


Extending relax-and-round combinatorial optimization solvers with quantum correlations

Maxime Dupont¹* and Bhuvanesh Sundar¹

Rigetti Computing, 775 Heinz Avenue, Berkeley, California 94710, USA

 (Received 25 July 2023; accepted 5 January 2024; published 24 January 2024)

We introduce a relax-and-round approach embedding the quantum approximate optimization algorithm (QAOA) with $p \geq 1$ layers. We show for many problems, including Sherrington-Kirkpatrick spin glasses, that at $p = 1$, it is as accurate as its classical counterpart, and maintains the infinite-depth optimal performance guarantee of the QAOA. Employing a different rounding scheme, we prove the method shares the performance of the Goemans-Williamson algorithm for the maximum cut problem on certain graphs. We pave the way for an overarching quantum relax-and-round framework with performance on par with some of the best classical algorithms.

DOI: [10.1103/PhysRevA.109.012429](https://doi.org/10.1103/PhysRevA.109.012429)

I. INTRODUCTION

Solving combinatorial optimization problems [1,2], or Ising models [3,4], is a formidable challenge connecting basic sciences, such as mathematical optimization, statistical physics, and condensed matter, with everyday life problems in logistics, scheduling, routing, finance, chemistry, biology, etc. [5]. The advent of controllable quantum simulators has inspired the development of quantum-based approaches for tackling combinatorial optimization. Notable examples include quantum annealing [6,7] and the quantum approximate optimization algorithm (QAOA) [8–11] for programmable quantum computers.

These quantum approaches have been successfully implemented at various scales on a wide range of platforms. For instance, superconducting quantum computers executed the QAOA for finding the ground state of spin glasses, solving the maximum cut problem, and performing a machine learning task for up to 23 qubits [12,13]. Trapped ions simulated the QAOA for solving a long-range Ising model on 40 spins [14]. Researchers observed a super-linear speedup in finding the maximum independent set on various graphs up to 289 vertices using ultracold Rydberg atoms [15–17]. A superconducting quantum annealer considered a spin glass with 5000 variables [18]. Despite ever-improving practical implementations as to the number of qubits, coherence, operating fidelity, and programmability, it remains an open question whether quantum machines can deliver—even in the fault-tolerant regime [19]—an advantage, in terms of speed, or the quality of the solution, versus the best classical methods.

The quality of a solution $\mathbf{z} \in \mathbb{Z}^N$ is typically characterized by the approximation ratio $\alpha = C(\mathbf{z})/C(\mathbf{z}_{\text{opt}})$, where $C(\mathbf{z})$ is a problem-dependent objective function that scores the solution and \mathbf{z}_{opt} is the optimal solution. Many combinatorial optimization problems are NP-hard with no efficient way of obtaining the optimal solution. Hence, an overarching goal is to develop approximate algorithms that return α as close to one as possible [20]—noting that it can also be NP-hard to

get α past a certain threshold. Celebrated classical examples include Goemans-Williamson algorithm for the maximum cut problem with $\alpha \simeq 0.878$ [21], Christofides-Serdyukov $\alpha = 3/2$ algorithm for the traveling salesman problem [22,23], and the $\alpha = 7/8$ approximation algorithm for the class MAX-E3-SAT of boolean satisfiability problems [20].

One of the most promising quantum algorithms for solving quadratic binary optimization problems on near-term quantum computers is the QAOA [8–11]. Given an objective function

$$C(\mathbf{z}) = \sum_{ij} W_{ij} z_i z_j, \quad (1)$$

where $W \in \mathbb{R}^{N \times N}$ is the adjacency matrix of an N -vertex undirected weighted graph encoding the problem, the optimization task is to minimize $C(\mathbf{z})$ over $\mathbf{z} \in \{\pm 1\}^N$. This is done by preparing a parameterized quantum state

$$|\Psi\rangle_p = \left[\prod_{\ell=1}^p e^{-i\beta_\ell \sum_{j=1}^N \hat{X}_j} e^{-i\gamma_\ell \hat{C}} \right] \hat{H}^{\otimes N} |0\rangle^{\otimes N}, \quad (2)$$

where \hat{H} is the one-qubit Hadamard gate, \hat{X}_i is the Pauli operator on qubit i , \hat{C} is the operator corresponding to the objective function of Eq. (1) obtained by replacing the binary variables z_i with Pauli operators \hat{Z}_i , and $\{\gamma_\ell, \beta_\ell\}$ are real-valued angles. These angles are optimal when they minimize the expectation value of the objective function $\langle \hat{C} \rangle_p$ over $|\Psi\rangle_p$ of Eq. (2).

The depth p acts as a control parameter of the quantum algorithm, such that the quality of the solutions improves as p is increased [8]. The solution from the QAOA is guaranteed to converge to the optimal solution ($\alpha = 1$) for $p \rightarrow +\infty$ due to the adiabatic theorem [8,24]. However, it is more difficult to ascertain its performance at low p . Yet, the low- p regime is particularly relevant for near-term quantum devices, which in the absence of quantum error correction [25], can only execute shallow algorithms. In some cases, the average approximation ratio is known for the QAOA at low p . For instance, the average approximation ratio at $p = 1$ for the maximum cut problem on random 3-regular graphs is $\alpha \simeq 0.692$ [8], and for ring graphs is $\alpha = (2p + 1)/(2p + 2)$ [8,26]. For paradigmatic Sherrington-Kirkpatrick (SK) spin glasses [27], the

*mdupont@rigetti.com

QAOA yields $\alpha \simeq 0.397$ at $p = 1$, and $\alpha \simeq 0.901$ at $p = 20$ [28,29]. Devising quantum algorithms with the lowest depth and highest performance compared to other known algorithms is a highly desired goal on the quest to achieving quantum advantage.

Here, we introduce an efficient quantum relax-and-round (QRR) algorithm that builds on top of QAOA to enhance its approximation ratio on a range of problems. QRR requires no more data from the quantum computer than what is already computed as part of the QAOA, making it an attractive plug-and-play addition to quantum optimization workflows. We show that at $p = 1$, QRR has the same performance as a classical relax-and-round algorithm on a large class of problem instances, much higher than that of the raw QAOA at $p = 1$. The solution from QRR converges asymptotically to the optimal solution for $p \rightarrow +\infty$ and displays robustness to certain types of quantum noise.

II. QUANTUM, RELAX, AND ROUND

A. The quantum relax-and-round (QRR) algorithm

Relax-and-round approaches are ubiquitous in approximate classical algorithms, such as those based on semidefinite programming [20]. Indeed, the difficulty in solving binary optimization problems comes from the solutions being restricted to the integer domain. If this constraint is relaxed (e.g., $\mathbf{z} \in \{\pm 1\}^N$ to $\mathbf{z} \in \mathbb{R}^N$ with $\|\mathbf{z}\| = \text{constant}$), the problem becomes an eigenvalue problem and is, therefore, solvable efficiently. A judicious rounding scheme is then employed to map the relaxed solution back to a valid one. For instance, for quadratic binary optimization problems [Eq. (1)] where \mathbf{W} is drawn from the Gaussian orthogonal ensemble, which corresponds to SK spin glasses [27], a relax-and-round scheme leads to an approximation ratio $\alpha = 2/\pi P^* \simeq 0.834$ [31–33], where P^* is the Parisi constant [30].

Here, we propose to perform a relax-and-round step on the correlation matrix resulting from the QAOA at depth p ,

$$\mathbf{Z}_{ij}^{(p)} = (\delta_{ij} - 1) \langle \hat{Z}_i \hat{Z}_j \rangle_p, \quad (3)$$

where δ_{ij} is the Kronecker delta and $\langle \hat{Z}_i \hat{Z}_j \rangle_p$ is the expectation value of the two-point correlation between qubits i and j over $|\Psi\rangle_p$ [Eq. (2)]. Specifically, we do an eigendecomposition of $\mathbf{Z}^{(p)}$ to obtain its eigenvectors $\{\mathbf{z} \in \mathbb{R}^N\}$. We round the eigenvectors entrywise to their sign $\{\mathbf{z} \leftarrow \text{sign}(\mathbf{z}) \in \{\pm 1\}^N\}$ to recover a valid solution to the original problem. Finally, the best rounded eigenvector with respect to the objective function is returned.

The intuition is that the correlation matrix elements encode the similarity between variables i and j : A positive matrix element means the variables z_i and z_j are negatively correlated in the ensemble of measurements, and a negative element means the variables are positively correlated in the ensemble. Thus implementing a relax-and-round algorithm on $\mathbf{Z}^{(p)}$ returns one solution \mathbf{z} such that pairs of variables (z_i, z_j) tend to minimize $\mathbf{Z}_{ij}^{(p)} z_i z_j$.

A similar argument explains the intuition for the classical relax-and-round algorithm [31–33]. However, the quantum relax-and-round algorithm is more powerful, since there can potentially be more nontrivial information in $\mathbf{Z}^{(p)}$ than in \mathbf{W} .

For example, consider a twofold degenerate optimal solution $\pm \mathbf{z}_{\text{opt}}$ because of the global \mathbb{Z}_2 sign flip symmetry $z_i \rightarrow -z_i \forall i$ for problems in the form of Eq. (1). In the $p \rightarrow +\infty$ limit, the correlation matrix becomes $\mathbf{Z}^{(\infty)} = \mathbb{I} - \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$, where \otimes denotes the outer product and \mathbb{I} is the identity matrix. The second term is a rank one matrix with eigenvector $\pm \mathbf{z}_{\text{opt}}/\sqrt{N}$, where the sign depends on the numerical solver. This eigenvector gets rounded to $\pm \mathbf{z}_{\text{opt}}$. For problems with degenerate optimal solutions beyond the global \mathbb{Z}_2 symmetry, vanishing one-body terms $\lim_{h_i \rightarrow 0} h_i z_i$ can be added to the original objective function to favor a single solution. For a nondegenerate solution, the rounding procedure should attempt both $\mathbf{z} \leftarrow \pm \text{sign}(\mathbf{z})$ to ensure the optimal solution is recovered in the infinite-depth limit.¹

B. Case study: Sherrington-Kirkpatrick spin glasses

We show that the QRR at $p = 1$ performs as well as its classical counterpart for a large class of problems. We exemplify this on SK spin glasses with random weights $\mathbf{W}_{i \neq j} = \pm 1$ [see Figs. 1(a) and 1(b)] and extend the analysis in Appendix C. The correlation matrix can be evaluated analytically through back-propagation at $p = 1$. At the optimal angles for the QAOA and employing a large- N expansion, the correlation matrix elements are given by (see Appendix C for details)

$$\lim_{N \rightarrow +\infty} \mathbf{Z}_{ij}^{(p=1)} \simeq \frac{\mathbf{W}_{ij}}{\sqrt{eN}} + \frac{\mathbf{N}_{ij}}{eN}, \quad (4)$$

where $\mathbf{N}_{ij} \equiv -[\mathbf{W}^2]_{ij}/2 = -\sum_k \mathbf{W}_{ik} \mathbf{W}_{kj}/2$ measures the sign imbalance in weights \mathbf{W}_{ik} and \mathbf{W}_{jk} for all nodes $k \neq i, j$. Because of the random nature of the weights, it is distributed similarly to a random walk with a total number $N/2 - 1$ of ± 1 steps. Consequently, \mathbf{W} and \mathbf{N} commute, and at the optimal angles and large N , the adjacency and correlation matrices also commute. Therefore they have the same eigenvectors and the QRR implemented on $\mathbf{Z}^{(1)}$ gives the same approximation ratio $\alpha = 2/\pi P^*$ as the classical relax-and-round algorithm based on \mathbf{W} . This is evidenced by the perfect agreement between marker symbols in Fig. 1(c), and between the bars for the SK model in Fig. 2. For comparison, the raw QAOA leads to $\alpha \simeq 0.397$ at $p = 1$ and crosses the threshold $\alpha \simeq 0.838 > 2/\pi P^*$ only for $p \geq 11$ [28].

To ascertain the performance of the QRR at finite $p > 1$, we perform numerical experiments (see Appendix D for details) for finite- N SK spin glasses and report the data in Fig. 1(c). We observe a systematic improvement as p increases, approaching the $p \rightarrow \infty$ value. Additional data analyses in Appendix E show that the QRR algorithm converges asymptotically faster to the optimal solution than the underlying QAOA at the same depth p .

¹This is because the correlation matrix $\mathbf{Z}^{(\infty)}$ captures \mathbf{z}_{opt} up to a global sign. In practice, the eigenvectors are real and defined up to a global ± 1 sign. Either can be returned depending on the numerical implementation of the eigendecomposition. Yet, only one corresponds to the nondegenerate optimal solution.

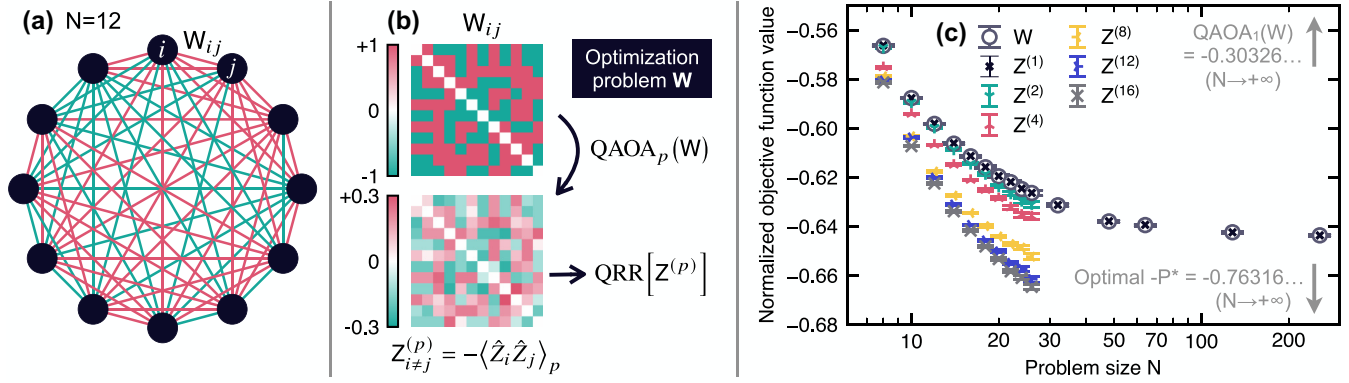


FIG. 1. (a) SK optimization problem instance [27] as a complete graph with $N = 12$ vertices. (b) The graph is encoded through adjacency matrix W . Running the QAOA at depth p gives access to a two-point correlation matrix $Z^{(p)}$ [Eq. (3)] between all variables i and j of the problem. A relax-and-round approach (QRR) is used on $Z^{(p)}$. (c) Numerical experiments at fixed optimal angles [28,29] for SK spin glasses. Objective function value normalized by $2N^{3/2}$ as a function of N . Each data point is averaged over 10^4 to 10^5 independent random problem instances. We report the value in the limit $N \rightarrow +\infty$ returned by the QAOA at $p = 1$ [28] and the optimal value P^* [30].

C. Algorithmic complexity

The QRR algorithm has a complexity of $O(N^2)$ for a fixed number n_{ex} of QAOA circuits executed. Executing one circuit takes a time $\sim pN$ [34] for SK instances. Building the correlation matrix takes a time $\sim N^2$, and assuming that the desired eigenvector is within the leading $k \ll N$ ones, the leading k eigenvectors of the correlation matrix $Z^{(p)}$ can be found in $\sim N^2$ operations. Finally, the rounding requires $\sim N$ steps and computing the objective function value $\sim N^2$, i.e., the number of edges in the graph. We study numerically the algorithmic time complexity of these steps in Appendix H. To date, the best performing classical algorithm for SK problems is an approximate message-passing algorithm that can return a solution with approximation ratio $\alpha = 1 - \varepsilon$ and complexity $\sim Q(\varepsilon)N^2$, where $Q(\varepsilon)$ is an inverse polynomial of ε controlling the desired accuracy [35]. We note that there

is a nontrivial control of the desired accuracy with the circuit depth p in the quantum case. In particular, there is an overhead in the QAOA algorithm for finding the optimal angles which are only known up to $p \leq 17$ [28,29]—an overhead which could scale exponentially with N and p [36,37]. Besides, we find in Appendix G that achieving optimal accuracy with the QRR algorithm requires a number of circuit executions for computing the correlation matrix elements going as $n_{\text{ex}} \sim N^\kappa$, where $\kappa \approx 1.5$ for $p = 1$ and where, by definition, $\kappa \rightarrow 0$ for $p \rightarrow +\infty$.

D. Robustness to noise and near-term friendliness

The QRR algorithm is robust to certain types of incoherent quantum noise such as those captured by a depolarizing noise channel. Under this channel, the system is in the mixed state [38], $\hat{\rho}_{p,F} = F|\Psi\rangle\langle\Psi|_p + (1-F)\hat{1}/2^N$, where $\hat{1}$ is the N -qubit identity matrix and $F \in [0, 1]$ the total circuit fidelity. The expectation value of the two-point correlation of Eq. (3) reads

$$Z_{ij}^{(p,F)} = (\delta_{ij} - 1) \text{tr}(\hat{\rho}_{p,F} \hat{Z}_i \hat{Z}_j) = F Z_{ij}^{(p,F=0)}, \quad (5)$$

which is just a rescaling of the noiseless correlation matrix by the total fidelity F . Therefore the eigenvectors of $Z_{ij}^{(p,F)}$ and the relax-and-round approach are unaffected by depolarizing noise. In practice, simulating an N -variable SK problem requires a minimum of pN^2 two-qubit gates [34]. Assuming the average gate fidelity is $f \in [0, 1]$, and that errors are uncorrelated, the circuit fidelity is $F \simeq f^{pN^2}$. Thus an exponentially large number of circuit executions n_{ex} is required to obtain a reliable signal-to-noise ratio $\sqrt{n_{\text{ex}}} f^{pN^2} \gg 1$. Depending on the problem (N), circuit (p), and hardware (f), the algorithm can remain practical in the absence of quantum error correction [25]. In particular, the algorithm is based on expectation values, enabling expectation-based error mitigation techniques [39] which are otherwise inapplicable with the QAOA for mitigating solutions. This reinforces its usability on the current generation of noisy quantum hardware.

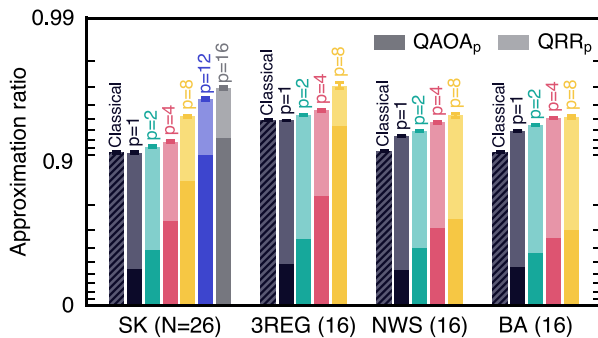


FIG. 2. Approximation ratio for different problem types at a fixed size N : SK with random ± 1 weights (SK), unit-weight random 3-regular graphs (3REG), random Newman-Watts-Strogatz small-world graphs with weights uniformly distributed on the unit line (NWS), and random Barabási-Albert graphs with weights drawn from a normal distribution (BA). Data are averaged between 10^3 to 10^4 independent random problem instances. For each case, the left-most dashed bar is the result of the classical relax-and-round on W . Subsequent bars show data from the QAOA and QRR algorithms at depth p .

E. Practical performance

We perform numerical experiments on various types of graph problems in the form of Eq. (1) (see Appendix A), and report the approximation ratio in Fig. 2. At $p = 1$, we empirically find that the QRR algorithm is always at least on par with its classical counterpart. For unit-weight random 3-regular graphs, this is because the adjacency and correlation matrices commute in the large N limit (see Appendix F), similar to SK problems. For the other two cases, the QRR algorithm outperforms the classical one. This is confirmed numerically on much larger problem instances up to $N = 256$ in Appendix I. As p increases and correlations get closer to that of the optimal solution, we find a systematic improvement in the QRR algorithm's results. Moreover, the approximation ratio is always larger than that of the underlying QAOA at the same depth p .

F. Quantum-flavored version of the Goemans-Williamson algorithm for the maximum cut problem

Noting that the classical relax-and-round algorithm that we used as a baseline does not have a performance guarantee for problems besides SK spin glasses, including the various types of graphs in Fig. 2, next we exemplify how the QRR algorithm can be modified such that it analogizes other classical algorithms with a performance guarantee. As a concrete example, we consider the Goemans-Williamson (GW) algorithm.

In the GW algorithm, the optimization task is to find the maximum cut of a graph, which is given by maximizing the objective function $C(\mathbf{z}) = \sum_{ij} L_{ij} z_i z_j / 4$ subject to $\mathbf{z} \in \{\pm 1\}^N$, where $L = D - W$ is the Laplacian matrix with $W_{ij} \geq 0$ and $D_{ij} = \delta_{ij} \sum_k W_{ik}$ is the degree matrix. The largest eigenvalue of $NL/4$ gives an upper bound to the actual maximum cut $C(\mathbf{z}_{\text{opt}})$ [40–43]. Noting that the objective function $C(\mathbf{z})$ is invariant under the transformation $D \rightarrow D + \text{diag}(\mathbf{u})$, where $\text{diag}(\mathbf{u})$ is a traceless diagonal matrix formed by the so-called correcting vector $\mathbf{u} \in \mathbb{R}^N$, it follows that the maximum cut is also bounded by $\max_{\|\mathbf{z}\|=1} \mathbf{z}^T \frac{N}{4} [D - W + \text{diag}(\mathbf{u})] \mathbf{z}$ for any vectors $\sum_i u_i = 0$ and $\mathbf{z} \in \mathbb{R}^N$. Thus one can make the eigenvalue bound tighter by implementing the relax-and-round algorithm on the following \mathbf{u} -augmented version. First, solve

$$\min_{\sum_i u_i = 0} \max_{\|\mathbf{z}\|=1} \mathbf{z}^T \frac{N}{4} [D - W + \text{diag}(\mathbf{u})] \mathbf{z}, \quad (6)$$

and then sign-round the leading eigenvector $\mathbf{z} \in \mathbb{R}^N$. This is equivalent to the GW algorithm and guarantees a maximum cut such that $\alpha \simeq 0.878$ [21,40–43]. A first task for this classical relaxed problem is to find the optimal correcting vector \mathbf{u}_{opt} minimizing the maximum eigenvalue, which is an upper bound to the optimal solution $C(\mathbf{z}_{\text{opt}})$. Finding \mathbf{u}_{opt} is a convex optimization problem, and thus numerically straightforward [41–43].

Equation (6) can naturally be adapted to our QRR algorithm by replacing W with $Z^{(p)}$. In some cases, the quantum relax-and-round algorithm with $D - Z^{(p)} + \text{diag}(\mathbf{u})$ can be analytically shown to perform at least as well as than the classical relax-and-round algorithm on $D - W + \text{diag}(\mathbf{u})$ for some correcting vectors \mathbf{u} .

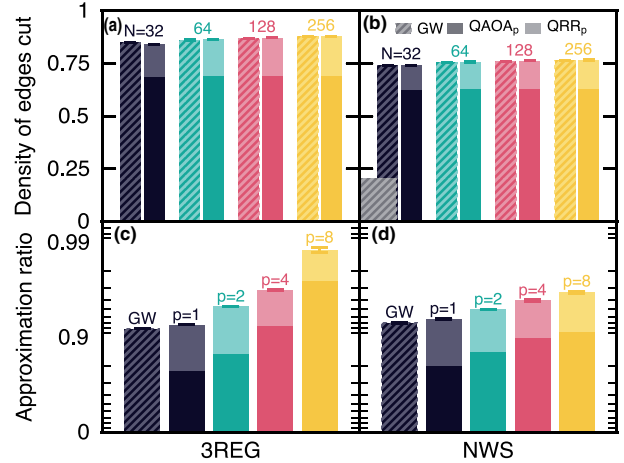


FIG. 3. Left: Maximum cut on unit-weight random 3-regular graphs (3REG). Right: Weighted maximum cut on random Newman-Watts-Strogatz small-world graphs with weights uniformly distributed on the unit line (NWS). Comparison of the GW algorithm with the QRR algorithm and the QAOA at depth p . (Top) Density of edges cut $C(\mathbf{z}) / \sum_{ij} W_{ij}$ for different problem sizes N at $p = 1$. (Bottom) Approximation ratio versus p at fixed $N = 16$. Data are averaged between 10^3 to 10^4 independent random problem instances.

As a first example, we show this in the $p \rightarrow +\infty$ limit, where $Z^{(\infty)} = I - \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$. Choosing $\text{diag}(\mathbf{u}) = \text{tr}(D)/N - D$ makes the eigenvalue problem trivial. The leading eigenvector is $\pm \mathbf{z}_{\text{opt}} / \sqrt{N}$, which gets rounded to $\pm \mathbf{z}_{\text{opt}}$.

Our second example is for certain vertex-transitive graphs. Due to the vertex-transitive nature, all the elements u_i are equal, and this is valid in both the classical and quantum algorithms. Since $\sum_i u_i = 0$, each element u_i is, therefore, zero [41–43]. For such graphs, the degree matrix D is proportional to the identity matrix. Thus the classical and QRR algorithms round the leading eigenvectors of $-W$ and $-Z^{(p)}$, respectively. The last step is to show that the leading eigenvector is the same in both cases, which is verified for ring and complete graphs (see also Appendix K).

We emphasize that the performance guarantee of the classical GW algorithm is agnostic to the graph, and showing whether a similarly strong statement can be made for the quantum version remains an open question for future work. Regardless, numerical experiments in Fig. 3 show that the QRR is a powerful heuristic even without theoretically established performance guarantees, on par with the classical version at $p = 1$ —and surpassing it for $p > 1$.

III. CONCLUSION

Our approach can embed other algorithms than the QAOA and directly applies to higher-order problems and those with one-body terms. For instance, we consider in Appendix J the maximum independent set problem on random unit disk graphs using a quantum annealing protocol. It is a straightforward testbed for Rydberg atoms thanks to their intrinsic blockade mechanism acting as a penalty for nonindependent set solutions [15–17,44,45]. A penalty disfavors invalid solutions but will not prevent them outside the $p \rightarrow +\infty$ limit. Here, the QRR algorithm provides a

heuristic to post-process for hard constraints: The squared values of the normalized eigenvectors of the correlation matrix $Z^{(p)}$ can be interpreted as probabilities to belong to their ± 1 sign-rounded groups. For instance, this probabilistic interpretation can guide a post-processing greedy approach for resolving out-of-constraint solutions, such as the one employed for maximum independent set problems with Rydberg atoms [15].

Furthermore, algorithms iteratively freezing variables to their classical values through consecutive executions of the QAOA on smaller and smaller problems [46–49] can also leverage this probabilistic interpretation as a natural freezing selection strategy.

A slightly different version of our algorithm would optimize the QAOA angles $\{\gamma, \beta\}_p$ with respect to the solution returned by relax-and-round step instead of $\langle \hat{C} \rangle_p$ directly. In addition to the proven guarantees at $p = 1$ and $p \rightarrow +\infty$, this variant would ensure that the QRR algorithm with p layers is necessarily at least as good as the QRR algorithm with $p - 1$ layers for all p , with the lower bound obtained by setting the variational angles $\gamma_p = \beta_p = 0$.

Relax-and-round strategies are at the roots of classical semidefinite programming methods—which are some of the best combinatorial problem solvers [20]. This makes our algorithm a fierce competitor and paves the way toward a quantum advantage for combinatorial optimization.

ACKNOWLEDGMENTS

We are grateful to A. Arrasmith, N. Didier, G. Enos, M. J. Hodson, M. Painsi, and M. J. Reagor for interesting discussions. This work is supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00112090058. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC Award No. DDR-ERCAP0024427.

APPENDIX A: DEFINITION OF PROBLEM AND GRAPH TYPES

We define the different problem and graph types considered throughout this work. The objective function of the following problems, that one seeks to minimize, is defined over a graph. The structure of the graph is encoded into its adjacency matrix W with $W_{ii}=0$ and $W_{i \neq j}$ the weight between vertices i and j (it is zero for nonadjacent vertices). The following problem types define effectively different adjacency matrices W . When needed for numerical experiments, we generate these graphs using the PYTHON package NETWORKX [50].

(1) *Sherrington-Kirkpatrick (SK) with random ± 1 weights.* SK problem instances [27] can be represented as complete graphs (i.e., all-to-all) with a random ± 1 weight of equal probability between all vertices.

(2) *Random 3-regular graphs with a unit weight.* Random 3-regular graphs correspond to graphs where each vertex is connected to three other vertices. The weight of the edges connecting those vertices is chosen to be one.

(3) *Random Newman-Watts-Strogatz small-world graphs with uniform random weights on the unit line.* The first step for building a Newman-Watts-Strogatz graph is to construct a ring graph. Then, an edge is added between all next-nearest neighbors, i.e., each node is connected to its $k = 4$ nearest neighbors. Then, with probability $p = 1/2$, we randomly pick an edge between vertices i and j , and rewire it to vertices i and k [51]. Finally, each edge is given a random weight drawn uniformly from the range $[0,1]$.

(4) *Random Barabási-Albert graphs with normally distributed random weights.* The first step for building a Barabási-Albert graph with N vertices is to generate a star graph with $N/4 + 1$ vertices. Then, the graph is grown by attaching new $3N/4 - 1$ nodes each with $m = N/4$ edges that are preferentially attached to existing nodes with high degree [52]. Finally, a random weight drawn from a normal distribution of mean zero and unit variance is given to each edge.

(5) *The ring graph.* The ring graph is a $d = 2$ regular graph with unit weight on the edges.

(6) *The Bethe lattice.* The Bethe lattice is an infinite connected cycle-free graph where all vertices have the same number of nearest neighbors. We consider the unit-weight Bethe lattice with d nearest neighbors.

(7) *The ring graph with additional next-nearest neighbor.* The starting point is the ring graph. Then, a connection between next-nearest neighbors is added. Thus each node has four connections. Each edge has unit weight.

(8) *The honeycomb lattice.* The honeycomb lattice is a hexagonal Bravais lattice with two nodes per unit cell. Each vertex has degree three. We either consider infinite honeycomb lattices, or lattices with periodic boundary conditions. Each edge has unit weight.

(9) *Random two-dimensional geometric graphs.* The graphs are generated by placing uniformly at random N vertices in the unit plane of dimensions 1×1 . Then, nodes are connected by an edge if they are within an euclidean distance of $r = \sqrt{\rho/N}$. Here ρ is a parameter of the graph leading to an average vertex degree for the nodes of $\pi\rho$. This graph is also known as random unit disk by rescaling $r \rightarrow 1$ and the dimension of the plane to $\sqrt{N/\rho} \times \sqrt{N/\rho}$.

APPENDIX B: ANALYTICAL EXPRESSION FOR THE EXPECTATION VALUES IN THE QAOA at $p = 1$

We calculate analytically the expectation value of the two-point correlation function $\langle \hat{Z}_i \hat{Z}_j \rangle_1$ from the quantum state $|\Psi\rangle_1$ resulting from a one-layer ($p = 1$) QAOA circuit. We denote $C(z) = \sum_{ij} W_{ij} z_i z_j$ the objective function for a graph problem with adjacency matrix W and Ising variables $z_i = \pm 1$. The quantum state reads

$$|\Psi\rangle_1 = e^{-i\beta_1 \sum_{j=1}^N \hat{X}_j} e^{-i\gamma_1 \hat{C}} \hat{H}^{\otimes N} |0\rangle^{\otimes N}, \quad (\text{B1})$$

where \hat{H} is the one-qubit Hadamard gate, \hat{X}_i is the Pauli operator on qubit i , \hat{C} is the operator corresponding to the objective function obtained by substituting the binary variables z_i for Pauli operators \hat{Z}_i , and $\{\gamma_1 \equiv \gamma, \beta_1 \equiv \beta\}$ are real-valued angles. Hence, the expectation value of the two-

point correlation function reads

$$\langle \hat{Z}_i \hat{Z}_j \rangle_1 = \langle + | e^{i\gamma \hat{C}} e^{i\beta \sum_{j=1}^N \hat{X}_j} \hat{Z}_i \hat{Z}_j e^{-i\beta \sum_{j=1}^N \hat{X}_j} e^{-i\gamma \hat{C}} | + \rangle, \quad (\text{B2})$$

where

$$|+\rangle \equiv \hat{H}^{\otimes N} |0\rangle^{\otimes N} = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes N}, \quad (\text{B3})$$

corresponds to an equal superposition of all the basis states. The inner term $e^{i\beta \sum_{j=1}^N \hat{X}_j} \hat{Z}_i \hat{Z}_j e^{-i\beta \sum_{j=1}^N \hat{X}_j}$ involves exponentials of sum of terms acting independently on the qubits and can be

easily expanded to

$$\begin{aligned} e^{i\beta \sum_{j=1}^N \hat{X}_j} \hat{Z}_i \hat{Z}_j e^{-i\beta \sum_{j=1}^N \hat{X}_j} \\ = [\cos(2\beta) \hat{Z}_i + \sin(2\beta) \hat{Y}_i][\cos(2\beta) \hat{Z}_j + \sin(2\beta) \hat{Y}_j], \end{aligned} \quad (\text{B4})$$

where \hat{Y}_i is the Pauli operator acting on qubit i . Expanding this product gives four terms. To compute $\langle \hat{Z}_i \hat{Z}_j \rangle$, one should first multiply these four terms on either side by $e^{i\gamma \hat{C}}$ and $e^{-i\gamma \hat{C}}$, and then compute the expectation value with respect to $|+\rangle$. We use the Baker-Campbell-Hausdorff formula to write

$$e^{i\gamma \hat{C}} \hat{Z}_i \hat{Z}_j e^{-i\gamma \hat{C}} = \hat{Z}_i \hat{Z}_j, \quad (\text{B5})$$

$$e^{i\gamma \hat{C}} \hat{Z}_i \hat{Y}_j e^{-i\gamma \hat{C}} = \hat{Z}_i \left[\cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j + \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{X}_j \right], \quad (\text{B6})$$

$$e^{i\gamma \hat{C}} \hat{Y}_i \hat{Z}_j e^{-i\gamma \hat{C}} = \left[\cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i + \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{X}_i \right] \hat{Z}_j, \quad (\text{B7})$$

$$e^{i\gamma \hat{C}} \hat{Y}_i \hat{Y}_j e^{-i\gamma \hat{C}} = \left[\cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i + \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{X}_i \right] \left[\cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j + \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{X}_j \right]. \quad (\text{B8})$$

In Eqs. (B5)–(B8), all the arguments inside the cosine and sine functions commute with each other. Therefore we can use standard trigonometric formulas to expand the cosines and sines. Specifically, $\cos(A+B) = \cos A \cos B - \sin A \sin B$ and $\sin(A+B) = \sin A \cos B + \cos A \sin B$.

Next, we compute expectation values of Eqs. (B5)–(B8) with respect to $|+\rangle$, one by one. The expectation value of Eq. (B5) is zero due to the global \mathbb{Z}_2 sign flip symmetry of the objective function, i.e. $\langle + | \hat{Z}_i \hat{Z}_j | + \rangle$. The expectation value of the first term in Eq. (B6) is also zero, because the expectation value must be real but \hat{Y}_j has a purely imaginary matrix. For the same reason, the first term in Eq. (B7), and two out of four terms in Eq. (B8) are zero

$$\langle + | \hat{Z}_i \cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j | + \rangle = 0, \quad (\text{B9})$$

$$\langle + | \cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i \hat{Z}_j | + \rangle = 0, \quad (\text{B10})$$

$$\langle + | \cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{X}_j | + \rangle = 0, \quad (\text{B11})$$

$$\langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{X}_i \cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j | + \rangle = 0. \quad (\text{B12})$$

The second term in Eq. (B6) is $\langle + | \hat{Z}_i \sin(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k) \hat{X}_j | + \rangle$. The state $|+\rangle$ is an eigenstate of \hat{X}_j , so we replace \hat{X}_j with its eigenvalue: one. We expand the sine function and obtain

$$\langle + | \hat{Z}_i \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) | + \rangle = \langle + | \hat{Z}_i \sin(2\gamma \mathbf{W}_{ij} \hat{Z}_i) \cos \left(\sum_{k \neq i, j} 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) + \hat{Z}_i \cos(2\gamma \mathbf{W}_{ij} \hat{Z}_i) \sin \left(\sum_{k \neq i, j} 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) | + \rangle. \quad (\text{B13})$$

The second term in Eq. (B13), $\langle + | \hat{Z}_i \cos(2\gamma \mathbf{W}_{ij} \hat{Z}_i) \cdots | + \rangle$ evaluates to zero due to the \mathbb{Z}_2 symmetry. The first term is nonzero. Using the identity that $\sin(2\gamma \mathbf{W}_{ij} \hat{Z}_i) = \hat{Z}_i \sin(2\gamma \mathbf{W}_{ij})$, it reduces to $\sin(2\gamma \mathbf{W}_{ij}) \langle + | \cos(\sum_{k \neq i, j} 2\gamma \mathbf{W}_{jk} \hat{Z}_k) | + \rangle$. Now, repeatedly expanding the cosine again, and using \mathbb{Z}_2 symmetry to send appropriate terms to zero at each step, we are left with

$$\langle + | \hat{Z}_i \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{X}_j | + \rangle = \sin(2\gamma \mathbf{W}_{ij}) \prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{jk}). \quad (\text{B14})$$

A similar argument to above yields that the second term in Eq. (B7) is

$$\langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{X}_i \hat{Z}_j | + \rangle \sin(2\gamma \mathbf{W}_{ij}) \prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{ik}). \quad (\text{B15})$$

The nonzero terms in Eq. (B8) are $\langle + | \cos(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k) \hat{Y}_i \cos(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k) \hat{Y}_j | + \rangle$ and $\langle + | \sin(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k) \hat{X}_i \sin(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k) \hat{X}_j | + \rangle$. Note that \hat{X}_i commutes with $\sin(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k)$ and recall that $| + \rangle$ is an eigenstate of \hat{X}_i . First, let us compute

$$\langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{X}_i \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{X}_j | + \rangle = \langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) | + \rangle. \quad (\text{B16})$$

We use the trigonometric formula $\sin A \sin B = \frac{1}{2} [\cos(A - B) - \cos(A + B)]$ and obtain

$$\begin{aligned} & \langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) | + \rangle \\ &= \frac{1}{2} \langle + | \cos \left[\sum_k 2\gamma (\mathbf{W}_{ik} + \mathbf{W}_{jk}) \hat{Z}_k \right] | + \rangle - \frac{1}{2} \langle + | \cos \left[\sum_k 2\gamma (\mathbf{W}_{ik} - \mathbf{W}_{jk}) \hat{Z}_k \right] | + \rangle. \end{aligned} \quad (\text{B17})$$

Repeatedly expanding the cosines and using \mathbb{Z}_2 symmetry at each step to keep only nonzero terms, it simplifies to

$$\begin{aligned} \langle + | \sin \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \sin \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) | + \rangle &= \frac{1}{2} \prod_k \cos 2\gamma (\mathbf{W}_{ik} - \mathbf{W}_{jk}) - \frac{1}{2} \prod_k \cos 2\gamma (\mathbf{W}_{ik} + \mathbf{W}_{jk}) \\ &= \frac{\cos^2(2\gamma \mathbf{W}_{ij})}{2} \prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{ik} - \mathbf{W}_{jk}) \\ &\quad - \frac{\cos^2(2\gamma \mathbf{W}_{ij})}{2} \prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{ik} + \mathbf{W}_{jk}). \end{aligned} \quad (\text{B18})$$

Next, let us compute

$$\langle + | \cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i \cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j | + \rangle = \langle + | \hat{Y}_i \cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j | + \rangle. \quad (\text{B19})$$

Using the trigonometric formula $\cos A \cos B = \frac{1}{2} [\cos(A + B) + \cos(A - B)]$, repeatedly expanding the cosines, and again using \mathbb{Z}_2 symmetry, we obtain

$$\begin{aligned} \langle + | \cos \left(\sum_k 2\gamma \mathbf{W}_{ik} \hat{Z}_k \right) \hat{Y}_i \cos \left(\sum_k 2\gamma \mathbf{W}_{jk} \hat{Z}_k \right) \hat{Y}_j | + \rangle &= \frac{\sin^2(2\gamma \mathbf{W}_{ij})}{2} \prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{ik} - \mathbf{W}_{jk}) - \frac{\sin^2(2\gamma \mathbf{W}_{ij})}{2} \prod_{k \neq i, j} \cos 2\gamma \\ &\quad \times (\mathbf{W}_{ik} + \mathbf{W}_{jk}). \end{aligned} \quad (\text{B20})$$

Finally, adding together the results from Eqs. (B9), (B14), (B15), (B18), and (B20), and multiplied by the right coefficients, the two-point correlation between variables i and j after a one-layer ($p = 1$) QAOA circuit reads

$$\begin{aligned} \langle \hat{Z}_i \hat{Z}_j \rangle_1 &= \sin(2\beta) \cos(2\beta) \sin(2\gamma \mathbf{W}_{ij}) \left[\prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{ik}) + \prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{jk}) \right] \\ &\quad - \frac{\sin^2(2\beta)}{2} \left[\prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{ik} + \mathbf{W}_{jk}) - \prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{jk} - \mathbf{W}_{ik}) \right]. \end{aligned} \quad (\text{B21})$$

It follows that the expectation value of the objective function is

$$\begin{aligned} \langle \hat{C} \rangle_1 &= \sum_{ij} \mathbf{W}_{ij} \left\{ \sin(2\beta) \cos(2\beta) \sin(2\gamma \mathbf{W}_{ij}) \left[\prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{ik}) + \prod_{k \neq i, j} \cos(2\gamma \mathbf{W}_{jk}) \right] \right. \\ &\quad \left. - \frac{\sin^2(2\beta)}{2} \left[\prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{ik} + \mathbf{W}_{jk}) - \prod_{k \neq i, j} \cos 2\gamma (\mathbf{W}_{jk} - \mathbf{W}_{ik}) \right] \right\}. \end{aligned} \quad (\text{B22})$$

APPENDIX C: ANALYTICAL RESULTS FOR THE QRR ALGORITHM

1. Sherrington-Kirkpatrick with random ± 1 weights

a. Rewriting the correlation matrix

We define the correlation matrix element $Z_{ij}^{(p=1)} = (\delta_{ij} - 1)\langle \hat{Z}_i \hat{Z}_j \rangle_1$, where δ_{ij} the Kronecker delta. For $N \rightarrow +\infty$, it is shown that the optimal QAOA angles are $\beta = -\pi/8$ and $\gamma \simeq 1/2\sqrt{N}$ [28]. Plugging this in Eq. (B21), it follows:

$$\begin{aligned} -Z_{i \neq j}^{(p=1)} &= \left(-\frac{\sqrt{2}}{2}\right) \left(\frac{\sqrt{2}}{2}\right) \sin\left(\frac{W_{ij}}{\sqrt{N}}\right) 2 \cos^{N-2}\left(\frac{1}{\sqrt{N}}\right) \\ &\quad - \frac{1}{4} \left[\prod_{k \neq i, j} \cos\left(\frac{W_{ik} + W_{jk}}{\sqrt{N}}\right) \right. \\ &\quad \left. - \prod_{k \neq i, j} \cos\left(\frac{W_{jk} - W_{ik}}{\sqrt{N}}\right) \right]. \end{aligned} \quad (C1)$$

We use $\sin x \sim x$ for $x \ll 1$ and $\cos^{x-2}(1/\sqrt{x}) \sim e^{-1/2}$ for $x \rightarrow +\infty$

$$\begin{aligned} Z_{i \neq j}^{(p=1)} &= W_{ij} \frac{1}{\sqrt{eN}} + \frac{1}{4} \left[\prod_{k \neq i, j} \cos\left(\frac{W_{ik} + W_{jk}}{\sqrt{N}}\right) \right. \\ &\quad \left. - \prod_{k \neq i, j} \cos\left(\frac{W_{jk} - W_{ik}}{\sqrt{N}}\right) \right]. \end{aligned} \quad (C2)$$

If we compare terms for a given k one by one in the two products above, one is equal to 1 while the other is equal to $\cos(2/\sqrt{N})$. We denote $K_{ij} \in [0, N-2]$ as the number of times $W_{jk} - W_{ik} \neq 0$ for the edge (i, j) . Then

$$\begin{aligned} Z_{i \neq j}^{(p=1)} &= W_{ij} \frac{1}{\sqrt{eN}} \\ &\quad + \frac{1}{4} \left[\cos^{N-2-K_{ij}}\left(\frac{2}{\sqrt{N}}\right) - \cos^{K_{ij}}\left(\frac{2}{\sqrt{N}}\right) \right]. \end{aligned} \quad (C3)$$

The distribution of K_{ij} follows from a discrete one-dimensional random walk with steps $+1$ or 0 of equal probability. On average, $\mathbb{E}[K_{ij}] = N/2 - 1$, and for a given walk, we denote $N_{ij} \in [-N/2 + 1, N/2 - 1]$ the distance from the expected average position. This is now a random walk on the variable N_{ij} with step ± 1 and a total number of step of $N/2 - 1$. We introduce

$$\begin{aligned} V_{i \neq j}(N) &= \frac{\sqrt{eN}}{4} \cos^{N/2-1}\left(\frac{2}{\sqrt{N}}\right) \\ &\quad \times \left[\cos^{-N_{ij}}\left(\frac{2}{\sqrt{N}}\right) - \cos^{N_{ij}}\left(\frac{2}{\sqrt{N}}\right) \right], \end{aligned} \quad (C4)$$

which implies that

$$\lim_{N \rightarrow +\infty} V_{i \neq j}(N) \sim \frac{N_{ij}}{\sqrt{eN}}. \quad (C5)$$

In the limit of large N , because N_{ij} is normally distributed with mean zero and variance $N/4$, V_{ij} also follows a normal

distribution with mean zero and variance $1/4e$. We write

$$Z_{i \neq j}^{(p=1)} = \frac{1}{\sqrt{eN}} (W_{ij} + V_{ij}). \quad (C6)$$

b. Equivalence with classical relax-and-round approach

In the large- N limit, we show that the matrices W and $Z^{(p=1)}$ share the same eigenvectors, and thus provide the same solution when used in the relax-and-round approach. To do so, we show that W and V commute (this also assumes that W has distinct eigenvalues, or that V does). For this, we first write $2N_{ij} = -\sum_k W_{ik} W_{kj}$. We have

$$\begin{aligned} [W, Z^{(p=1)}] &= [W, V] = [WV]_{ij} - [VW]_{ij} \\ &= \frac{1}{\sqrt{eN}} \left(\sum_k W_{ik} N_{kj} - \sum_k N_{ik} W_{kj} \right) \\ &= \frac{1}{2\sqrt{eN}} \left(\sum_{k,q} W_{iq} W_{qk} W_{kj} - \sum_{k,q} W_{ik} W_{kq} W_{qj} \right) \\ &= \frac{1}{2\sqrt{eN}} ([W^3]_{ij} - [W^3]_{ij}) = 0. \end{aligned} \quad (C7)$$

This proves that in the large- N limit, W and V share the same eigenvectors—although they may not be ordered in the same way with respect to their corresponding eigenvalues. Thus the relax-and-round approach based on $Z^{(p=1)}$ has the same solution as the classical RR algorithm based on W .

2. Unit-weight d -regular graphs

a. The correlation matrix

We consider unit-weight d -regular graphs, i.e., graphs where each node is connected to exactly d others. The weights W_{ij} of the adjacency matrix of the graph are 1 if two nodes are connected, and 0 otherwise. The correlation matrix element $Z_{ij}^{(p=1)} = (\delta_{ij} - 1)\langle \hat{Z}_i \hat{Z}_j \rangle_1$ where δ_{ij} is the Kronecker delta, reads

$$\begin{aligned} -Z_{i \neq j}^{(p=1)} &= W_{ij} \times \sin(2\beta) \cos(2\beta) \sin(2\gamma) \cos^{d-1}(2\gamma) \\ &\quad - \frac{\sin^2(2\beta)}{2} \left[\prod_{k \neq i, j} \cos 2\gamma (W_{ik} + W_{jk}) \right. \\ &\quad \left. - \prod_{k \neq i, j} \cos 2\gamma (W_{jk} - W_{ik}) \right]. \end{aligned} \quad (C8)$$

We introduce the scalar number

$$f = -2 \sin(2\beta) \cos(2\beta) \sin(2\gamma) \cos^{d-1}(2\gamma) \quad (C9)$$

and rewrite the correlation matrix as

$$\begin{aligned} Z_{i \neq j}^{(p=1)} &= f \left(W_{ij} + \frac{\tan(2\beta)}{4 \tan(2\gamma) \cos^d(2\gamma)} \right. \\ &\quad \times \left[\prod_{k \neq i, j} \cos 2\gamma (W_{ik} + W_{jk}) \right. \\ &\quad \left. \left. - \prod_{k \neq i, j} \cos 2\gamma (W_{jk} - W_{ik}) \right] \right). \end{aligned} \quad (C10)$$

For a pair of nodes (i, j) , we denote v_{ij} as the number of nodes $k \neq i, j$ which are nearest neighbors to both i and j but for which i and j are not nearest neighbors ($W_{ij} = 0$). Then,

(1) there will be a total of $2d - v_{ij} \geq 0$ nontrivial nonunit terms in the first product. When both i and j are nearest neighbors with k , we get a contribution $\cos 4\gamma$. When k is the nearest neighbor of either i or j , we get a contribution $\cos 2\gamma$. Else the contribution is trivially $\cos 0 = 1$. Thus $\prod_{k \neq i, j} \cos 2\gamma (W_{ik} + W_{jk}) = \cos^{v_{ij}}(4\gamma) \cos^{2d-2v_{ij}}(2\gamma)$.

(2) For the second product, we get nontrivial nonunit terms only if k is a nearest neighbor to either i or j , but not both. Thus $\prod_{k \neq i, j} \cos 2\gamma (W_{ik} - W_{jk}) = \cos^{2d-2v_{ij}}(2\gamma)$.

For a pair of nodes (i, j) , we denote λ_{ij} as the number of nodes $k \neq i, j$ which are nearest neighbors to both i and j and for which i and j are also nearest neighbors ($W_{ij} = 1$). Then,

(3) each node i and j has a total of $d - 1$ nearest neighbors others than j and i , respectively. When k is nearest neighbor with both i and j , the first product will give a contribution $\cos 4\gamma$. When k is the nearest neighbor of either i or j , but not both, we get a contribution $\cos 2\gamma$. Else, the contribution will be trivially $\cos 0 = 1$. Thus $\prod_{k \neq i, j} \cos 2\gamma (W_{ik} + W_{jk}) = \cos^{\lambda_{ij}}(4\gamma) \cos^{2d-2-2\lambda_{ij}}(2\gamma)$.

(4) For the second product, we get trivial unit contributions for triangles and when k is neither a nearest neighbor of i nor of j . Thus $\prod_{k \neq i, j} \cos 2\gamma (W_{jk} - W_{ik}) = \cos^{2d-2-2\lambda_{ij}}(2\gamma)$.

Plugging everything together, we get

$$\frac{Z_{i \neq j}^{(p=1)}}{f} = W_{ij} + \frac{\tan(2\beta)}{4 \tan(2\gamma) \cos^d(2\gamma)} \times \begin{cases} \cos^{2d-2v_{ij}}(2\gamma) [\cos^{v_{ij}}(4\gamma) - 1] \text{ if } W_{ij} = 0 \\ \cos^{2d-2-2\lambda_{ij}}(2\gamma) [\cos^{\lambda_{ij}}(4\gamma) - 1] \text{ if } W_{ij} = 1 \end{cases} \quad (\text{C11})$$

which we can rewrite in a single expression

$$\frac{Z_{i \neq j}^{(p=1)}}{f} = W_{ij} + \frac{\tan(2\beta) \cos^d(2\gamma)}{4 \tan(2\gamma)} \times \{(1 - W_{ij}) \cos^{-2v_{ij}}(2\gamma) [\cos^{v_{ij}}(4\gamma) - 1] + W_{ij} \cos^{-2-2\lambda_{ij}}(2\gamma) [\cos^{\lambda_{ij}}(4\gamma) - 1]\}. \quad (\text{C12})$$

We can express λ_{ij} and v_{ij} directly with the weights. We introduce $n_{ij} = \sum_k W_{ik} W_{kj}$, which counts the number of times nodes i and j share a common nearest neighbor k , and this is independent of whether i and j are nearest neighbors themselves. We note that n_{ij} is different than N_{ij} previously introduced for the SK problems. Therefore we can write

$$\lambda_{ij} = W_{ij} n_{ij}, \quad \text{and} \quad v_{ij} = (1 - W_{ij}) n_{ij} = n_{ij} - \lambda_{ij}, \quad (\text{C13})$$

and express $Z_{i \neq j}^{(p=1)}$ as a function of λ_{ij} and n_{ij} exclusively

$$\frac{Z_{i \neq j}^{(p=1)}}{f} = W_{ij} + \frac{\tan(2\beta) \cos^d(2\gamma)}{4 \tan(2\gamma)} \times \{(1 - W_{ij}) \cos^{-2n_{ij}+2\lambda_{ij}}(2\gamma) [\cos^{n_{ij}-\lambda_{ij}}(4\gamma) - 1] + W_{ij} \cos^{-2-2\lambda_{ij}}(2\gamma) [\cos^{\lambda_{ij}}(4\gamma) - 1]\}. \quad (\text{C14})$$

b. The ring graph

We consider the special case of a $d = 2$ regular graph with unit weight, which is a ring. In that case, $\lambda_{ij} = 0 \forall i, j$ and $n_{ij} = 1$ if and only if i and j are next-nearest neighbors, and zero otherwise. If i and j are next-nearest neighbors, then by definition, $W_{ij} = 0$. We get

$$\frac{Z_{ij}^{(p=1)}}{f} = W_{ij} + n_{ij} \frac{\tan(2\beta)}{4 \tan(2\gamma)} [\cos(4\gamma) - 1]. \quad (\text{C15})$$

Ordering the variables according to the ring structure makes the matrix $Z^{(p=1)}$ circulant with an entry when (i, j) are nearest- and next-nearest neighbors. Its eigenvectors are the Fourier modes. Because, the adjacency matrix W of the problem is also circulant, it shares the same eigenvectors as $Z^{(p=1)}$. Thus the QRR algorithm on $Z^{(p=1)}$ has the same solution as the classical RR algorithm on W . For an even number of variables, one of such Fourier modes is $|\phi\rangle = (+1, -1, +1, \dots, -1)^T / \sqrt{N}$, which solves the initial problem exactly when rounded. It follows that the QRR algorithm at $p = 1$ also solves the ring graph exactly. The result holds independently of the value of the QAOA angles β and γ . For completeness, the optimal angles are $\beta = \gamma = \pi/8$ and the correlation matrix takes the following form:

$$Z^{(p=1)} = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{8} & 0 & \cdots & 0 & -\frac{1}{8} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{8} & 0 & \cdots & 0 & -\frac{1}{8} \\ -\frac{1}{8} & \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{8} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\frac{1}{8} & 0 & \cdots & \cdots & -\frac{1}{8} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{8} & 0 & \cdots & 0 & -\frac{1}{8} & \frac{1}{2} & 0 \end{pmatrix}, \quad (\text{C16})$$

with minimum eigenvalue $4/5$ and corresponding eigenvector $|\phi\rangle = (+1, -1, +1, \dots, -1)^T / \sqrt{N}$. Thus the leading eigenvector of the correlation matrix is the optimal solution to the problem.

c. The Bethe lattice

The result for the Bethe lattice extends straightforwardly the result for the ring graph. We consider a unit-weight Bethe lattice where each node has d nearest neighbors. In that case, $\lambda_{ij} = 0 \forall i, j$ and $n_{ij} = 1$ if and only if i and j are next-nearest neighbors, and zero otherwise. If i and j are next-nearest neighbors, then by definition, $W_{ij} = 0$. We get

$$\frac{Z_{ij}^{(p=1)}}{f} = W_{ij} + \frac{\tan(2\beta) \cos^{d-2}(2\gamma)}{4 \tan(2\gamma)} \times [\cos(4\gamma) - 1] \left(\sum_k W_{ik} W_{kj} \right). \quad (\text{C17})$$

We now compute the commutator

$$\begin{aligned} [\mathbf{WZ}^{(p=1)}]_{ij} - [\mathbf{Z}^{(p=1)}\mathbf{W}]_{ij} &= f \frac{\tan(2\beta) \cos^{d-2}(2\gamma)}{4 \tan(2\gamma)} [\cos(4\gamma) - 1] \left(\sum_{k,q} \mathbf{W}_{ik} \mathbf{W}_{kq} \mathbf{W}_{qj} - \sum_{k,q} \mathbf{W}_{ik} \mathbf{W}_{kq} \mathbf{W}_{qj} \right) \\ &= f \frac{\tan(2\beta) \cos^{d-2}(2\gamma)}{4 \tan(2\gamma)} [\cos(4\gamma) - 1] ([\mathbf{W}^3]_{ij} - [\mathbf{W}^3]_{ij}) = 0. \end{aligned} \quad (\text{C18})$$

Thus the QRR algorithm on $\mathbf{Z}^{(p=1)}$ has the same solution as the classical RR algorithm on \mathbf{W} .

d. The ring graph with additional next-nearest neighbor connections

We consider the special case of a $d = 4$ regular graph with unit weight, which is a ring with additional next-nearest neighbor connections between the vertices. In that case, $\lambda_{ij} = 0, 1$, or 2 , and $v_{ij} = 0, 1$, or 2 . When limited to these values, we have the trigonometric identity

$$\begin{aligned} \cos^{-2v_{ij}}(2\gamma) [\cos^{v_{ij}}(4\gamma) - 1] &= -2v_{ij} \tan^2(2\gamma) \\ \text{for } v_{ij} \in \{0, 1, 2\} (v_{ij} \leftrightarrow \lambda_{ij}). \end{aligned} \quad (\text{C19})$$

Thanks to the linearization with respect to v_{ij} and λ_{ij} , we get

$$\begin{aligned} \frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} &= \mathbf{W}_{ij} - \frac{1}{2} \tan(2\beta) \cos^4(2\gamma) \tan(2\gamma) \\ &\times [(1 - \mathbf{W}_{ij})v_{ij} + \mathbf{W}_{ij} \cos^{-2}(2\gamma) \lambda_{ij}]. \end{aligned} \quad (\text{C20})$$

We recall that we consider a graph with unit weights where $\mathbf{W}_{ij} = 0, 1$. There is redundancy because $v_{ij} = 0$ when $\mathbf{W}_{ij} = 1$ and $\lambda_{ij} = 0$ when $\mathbf{W}_{ij} = 0$. Thus we drop \mathbf{W}_{ij} and $(1 - \mathbf{W}_{ij})$ in the second term

$$\begin{aligned} \frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} &= \mathbf{W}_{ij} - \frac{1}{2} \tan(2\beta) \cos^4(2\gamma) \tan(2\gamma) \\ &\times [v_{ij} + \cos^{-2}(2\gamma) \lambda_{ij}]. \end{aligned} \quad (\text{C21})$$

We now express v_{ij} and λ_{ij} in terms of n_{ij} and \mathbf{W}_{ij}

$$\begin{aligned} \frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} &= \mathbf{W}_{ij} - \frac{1}{2} \tan(2\beta) \cos^4(2\gamma) \tan(2\gamma) \\ &\times [(1 - \mathbf{W}_{ij})n_{ij} + \mathbf{W}_{ij} n_{ij} \cos^{-2}(2\gamma)] \\ &= \mathbf{W}_{ij} - \frac{1}{2} \tan(2\beta) \cos^4(2\gamma) \tan(2\gamma) n_{ij} \\ &\times \{1 + \mathbf{W}_{ij} [\cos^{-2}(2\gamma) - 1]\} \\ &= \mathbf{W}_{ij} - \frac{n_{ij}}{2} \tan(2\beta) \cos^4(2\gamma) \tan(2\gamma) \\ &\quad - \frac{n_{ij} \mathbf{W}_{ij}}{2} \tan(2\beta) \\ &\quad \times \cos^4(2\gamma) \tan(2\gamma) [\cos^{-2}(2\gamma) - 1]. \end{aligned} \quad (\text{C22})$$

Depending on (i, j) , there are four possible cases for large N :

(5) if i and j are nearest neighbors, then $\mathbf{W}_{ij} = 1$ and $n_{ij} = 2$,

(6) if i and j are next-nearest neighbors, then $\mathbf{W}_{ij} = 0$ and $n_{ij} = 1$,

(7) if i and j are next-next-nearest neighbors, then $\mathbf{W}_{ij} = 0$ and $n_{ij} = 2$,

(8) if i and j are next-next-next-nearest neighbors, then $\mathbf{W}_{ij} = 0$ and $n_{ij} = 1$.

Ordering the variables according to the ring structure makes the matrix $\mathbf{Z}^{(p=1)}$ circulant. Its eigenvectors are the Fourier modes. Because the adjacency matrix \mathbf{W} of the problem is also circulant, it shares the same eigenvectors as $\mathbf{Z}^{(p=1)}$. Thus the QRR algorithm based on $\mathbf{Z}^{(p=1)}$ has the same solution as the classical RR algorithm on \mathbf{W} . When N is a multiple of 4, it is known [53] that the ground state of the initial Ising model is the antiphase “ $\langle 2 \rangle$ ” state in the form $\pm \pm \mp \mp \pm \pm \dots \mp \mp$ with four-fold degeneracy by translating this state by one, two, or three entries. Because this sign structure is captured by a Fourier mode, the QRR algorithm based on $\mathbf{Z}^{(p=1)}$ solves this problem exactly, as does the RR algorithm on \mathbf{W} .

e. The complete graph

The adjacency matrix \mathbf{W} of a unit-weight complete graph is filled with ones except on its diagonal which is zero. The QAOA circuit conserves the symmetry of this graph, resulting in a correlation matrix $\mathbf{Z}^{(p)} \propto \mathbf{W}$. Hence, they trivially share the same eigenvectors and their eigenvalues are equivalent up to a global rescaling factor.

f. Circulant graphs

For circulant graphs, the QAOA circuit will lead to a correlation matrix $\mathbf{Z}^{(p)}$ which is also circulant. The eigenvectors of circulant matrices are Fourier modes. Thus both \mathbf{W} and $\mathbf{Z}^{(p)}$ will share the same eigenvectors. This is the property we used previously for the cycle graph and cycle graph with additional edges between next-nearest neighboring vertices.

g. The honeycomb lattice

We can extend the above calculations to a two-dimensional honeycomb lattice, either infinite or with periodic boundary conditions for simplicity. It has $d = 3$, $\lambda_{ij} = 0 \forall i, j$ and $n_{ij} = 1$ if and only if i and j are next-nearest neighbors, and zero otherwise. If i and j are next-nearest neighbors, then by definition, $\mathbf{W}_{ij} = 0$. We get

$$\begin{aligned} \frac{\mathbf{Z}_{ij}^{(p=1)}}{f} &= \mathbf{W}_{ij} + \frac{\tan(2\beta) \cos^2(2\gamma)}{4 \sin(2\gamma)} \\ &\times [\cos(4\gamma) - 1] \left(\sum_k \mathbf{W}_{ik} \mathbf{W}_{kj} \right). \end{aligned} \quad (\text{C23})$$

We now compute the commutator

$$\begin{aligned} [\mathbf{WZ}^{(p=1)}]_{ij} - [\mathbf{Z}^{(p=1)}\mathbf{W}]_{ij} &= f \frac{\tan(2\beta) \cos^2(2\gamma)}{4 \sin(2\gamma)} [\cos(4\gamma) - 1] \left(\sum_{k,q} \mathbf{W}_{ik} \mathbf{W}_{kq} \mathbf{W}_{qj} - \sum_{k,q} \mathbf{W}_{ik} \mathbf{W}_{kq} \mathbf{W}_{qj} \right) \\ &= f \frac{\tan(2\beta) \cos^2(2\gamma)}{4 \sin(2\gamma)} [\cos(4\gamma) - 1] ([\mathbf{W}^3]_{ij} - [\mathbf{W}^3]_{ij}) = 0. \end{aligned} \quad (\text{C24})$$

Thus the QRR based on $\mathbf{Z}^{(p=1)}$ has the same solution as the classical RR algorithm on \mathbf{W} for the unit-weight honeycomb lattice.

h. Random 3-regular graphs

By definition $d = 3$ for 3-regular graphs as all vertices have exactly three nearest neighbors. Because of that, $\lambda_{ij} = 0, 1, \text{ or } 2$, and $v_{ij} = 0, 1, 2, \text{ or } 3$. We can use the same trigonometric identity used for the ring graph with additional next-nearest neighbor, except for the $v_{ij} = 3$ case. We handle this case separately by introducing $\delta_{v_{ij}3}$ the Kronecker delta. We get

$$\frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} = \mathbf{W}_{ij} + \frac{\tan(2\beta) \cos^3(2\gamma)}{4 \tan(2\gamma)} \left\{ (1 - \mathbf{W}_{ij}) \left[\delta_{v_{ij}3} \frac{\cos^3(4\gamma) - 1}{\cos^6(2\gamma)} - (1 - \delta_{v_{ij}3}) 2v_{ij} \tan^2(2\gamma) \right] - \mathbf{W}_{ij} \frac{2\lambda_{ij} \tan^2(2\gamma)}{\cos^2(2\gamma)} \right\} \quad (\text{C25})$$

and expand the terms in the bracket

$$\begin{aligned} \frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} &= \mathbf{W}_{ij} + \frac{\tan(2\beta) \cos^3(2\gamma)}{4 \tan(2\gamma)} \left\{ \delta_{v_{ij}3} \frac{\cos^3(4\gamma) - 1}{\cos^6(2\gamma)} - (1 - \delta_{v_{ij}3}) 2v_{ij} \tan^2(2\gamma) - \mathbf{W}_{ij} \delta_{v_{ij}3} \frac{\cos^3(4\gamma) - 1}{\cos^6(2\gamma)} \right. \\ &\quad \left. + \mathbf{W}_{ij} (1 - \delta_{v_{ij}3}) 2v_{ij} \tan^2(2\gamma) - \mathbf{W}_{ij} \frac{2\lambda_{ij} \tan^2(2\gamma)}{\cos^2(2\gamma)} \right\}. \end{aligned} \quad (\text{C26})$$

Because we consider unit-weight graphs, we remove redundancies that may exist when multiplying \mathbf{W}_{ij} with λ_{ij} , v_{ij} , and $\delta_{v_{ij}3}$

$$\frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} = \mathbf{W}_{ij} + \frac{\tan(2\beta) \cos^3(2\gamma)}{4 \tan(2\gamma)} \left\{ \delta_{v_{ij}3} \frac{\cos^3(4\gamma) - 1}{\cos^6(2\gamma)} - 2v_{ij} \tan^2(2\gamma) + \delta_{v_{ij}3} 6 \tan^2(2\gamma) - \lambda_{ij} \frac{2 \tan^2(2\gamma)}{\cos^2(2\gamma)} \right\}. \quad (\text{C27})$$

Using the trigonometric identity $[\cos^3(4\gamma) - 1] \cos^{-6}(2\gamma) + 6 \tan^2(2\gamma) = -2 \tan^6(2\gamma)$, it can be further simplified to

$$\frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} = \mathbf{W}_{ij} - \frac{\tan(2\beta) \cos^2(2\gamma) \sin(2\gamma)}{2} [\delta_{v_{ij}3} \tan^4(2\gamma) + v_{ij} + \lambda_{ij} \cos^{-2}(2\gamma)] \quad (\text{C28})$$

and then to

$$\begin{aligned} \frac{\mathbf{Z}_{i \neq j}^{(p=1)}}{f} &= \mathbf{W}_{ij} - \frac{\tan(2\beta) \cos^2(2\gamma) \sin(2\gamma)}{2} \\ &\quad \times [n_{ij} + \delta_{n_{ij}3} \tan^4(2\gamma) + \mathbf{W}_{ij} n_{ij} \tan^2(2\gamma)]. \end{aligned} \quad (\text{C29})$$

Unlike other cases previously discussed, it cannot be easily shown that the last two terms commute with the adjacency matrix. However, relying on numerical experiments, we find in Sec. F that the operator norm of the commutator $[\mathbf{W}, \mathbf{Z}^{(p=1)}]$ slowly goes to zero as $N \rightarrow +\infty$. We find strong numerical evidence that the performance of the QRR algorithm is on par with that of the classical RR algorithm, similarly to SK models.

APPENDIX D: NUMERICAL EXPERIMENTS

The relax-and-round algorithm requires expectation values of Pauli \hat{Z} observables to construct the correlation matrix. When the optimal angles of the QAOA are unknown, these expectation values are also necessary to compute the cost in order to optimize the angles. Numerical experiments are performed using the PYTHON packages NUMPY [54], SCIPY [55], and NUMBA [56]. Graph problems are generated using the PYTHON package NETWORKX [50]. For the quantum relax-and-

round version of the Goemans-Williamson algorithm, which requires finding the optimal correction vector, we use the PYTHON convex optimization solver CVXPY [57,58].

1. Running the QAOA algorithm

For QAOA depth $p = 1$, we compute these expectation values using the analytical formulas of Sec. B. We use these formulas for up to several hundred variables N . For $p > 1$, we rely on state vector simulations, which explicitly compute the quantum state $|\Psi\rangle_p$ resulting from the QAOA circuit. We use state vector simulations for up to $N = 26$ variables and compute expectation values exactly without sampling the quantum state unless specified otherwise.

When needed, we optimize the angles of the QAOA circuit using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [59–62] with a maximum number of 100 iterations. For a problem of size N using the QAOA with depth p (there are $2p$ parametric angles), we run the BFGS algorithm independently $\min(2^{4+p}, 2^{10})$ times with random initial angles $[0, 2\pi]^{2p}$. The best result of these simulations is kept and the angles are considered as optimal for computing expectation values. However, it should be noted that there is no guarantee these angles are actually optimal.

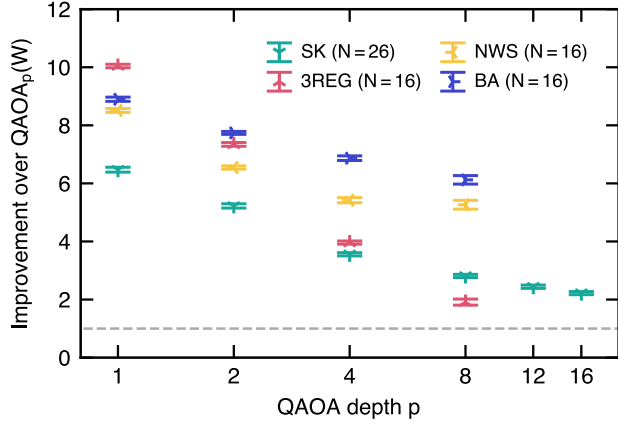


FIG. 4. Ratio of $(1 - \alpha)$ obtained from the relax-and-round approach to that obtained from the raw QAOA, where α is the approximation ratio. Problem instances considered: (a) Sherrington-Kirkpatrick with random ± 1 weights (SK), (b) random 3-regular graphs with unit weight (3REG), (c) random Newman-Watts-Strogatz small-world graphs with weights uniformly distributed on the unit line (NWS), and (d) random Barabási-Albert graphs with weights drawn from a normal distribution with mean zero and unit variance (BA). Each data point is averaged between 10^3 to 10^5 independent random problem instances. The dashed horizontal line is at 1 and corresponds to no improvements.

2. Collecting statistics

Unless specified otherwise, the data are averaged between 10^3 and 10^5 randomly generated problem instances.

3. Relax-and-round step

We use a standard numerical eigendecomposition method to perform the relax-and-round approach on the adjacency and correlation matrices W and $Z^{(p)}$. For an N -variable problem both matrices are of size $N \times N$, real, and symmetric. The eigendecomposition returns N eigenvectors $\{z\}$, which are real and normalized to unity. The eigenvectors are rounded entry-wise $\{z \leftarrow \text{sign}(z) \in \{\pm 1\}^N\}$. We also consider sign-flipped eigenvectors $\{-z\}$, also rounded entrywise. Therefore we obtain a total of $2N$ potential solutions to the initial problem. The objective value of each of these solutions is computed and the one extremizing the objective function is returned by the relax-and-round algorithm. A zero entry is rounded at random to ± 1 .

APPENDIX E: IMPROVEMENT OF THE APPROXIMATION RATIO FROM $QAOA_p(W)$ TO $QRR(Z^{(p)})$

In Fig. 4, we plot the ratio of $(1 - \alpha)$ obtained from the relax-and-round approach to that obtained from the raw QAOA. For all the cases studied here and up to $p \leq 8$, the numerical experiments show that the relax-and-round algorithm outperforms the QAOA at the same depth. Whether this remains true for larger N or p is an open question. In the limit $p \rightarrow +\infty$, both methods converge to the optimal answer where the ratio of $(1 - \alpha)$ obtained from the relax-and-round approach to that obtained from the raw QAOA is one.

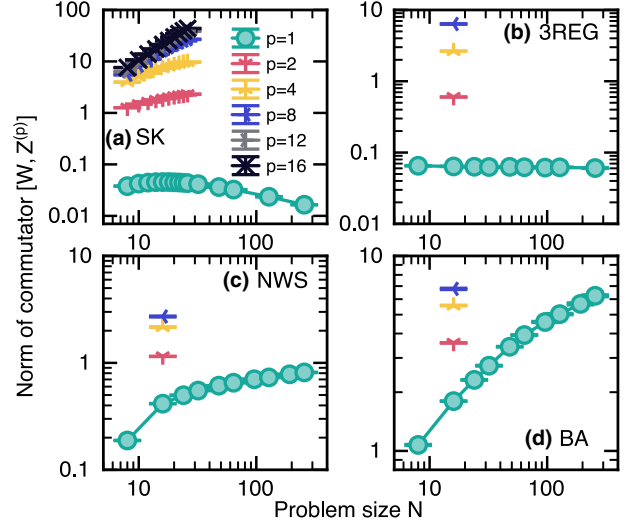


FIG. 5. Norm of the commutator [Eq. (F1)] between the adjacency matrix W and the correlation matrix $Z^{(p)}$ based on the QAOA at depth p as a function of the problem size N . Problem instances considered: (a) Sherrington-Kirkpatrick with random ± 1 weights (SK), (b) random 3-regular graphs with unit weight (3REG), (c) random Newman-Watts-Strogatz small-world graphs with weights uniformly distributed on the unit line (NWS), and (d) random Barabási-Albert graphs with weights drawn from a normal distribution with mean zero and unit variance (BA). Each data point is averaged between 10^3 to 10^5 independent random problem instances. Lines are a guide to the eye.

APPENDIX F: COMMUTATOR BETWEEN THE ADJACENCY AND THE CORRELATION MATRICES W AND $Z^{(p)}$

We consider the operator norm $\|\cdot\|_2$ of the commutator $[W, Z^{(p)}]$ between the adjacency and the correlation matrices W and $Z^{(p)}$

$$\|[W, Z^{(p)}]\|_2 = \|WZ^{(p)} - Z^{(p)}W\|_2 = \sigma_{\max}, \quad (\text{F1})$$

where σ_{\max} is the largest singular value of the commutator. We have used the commutator to prove analytically for several problem instances that the relax-and-round approach on W or $Z^{(p=1)}$ leads to the same result. In Fig. 5, we numerically compute and plot the norm as a function of the problem size N at various values of the QAOA depth p for different problem types. We find that the norm of the commutator increases with p .

APPENDIX G: EFFECT OF THE FINITE NUMBER OF CIRCUIT EXECUTIONS

We investigate the effect of a finite number of circuit executions n_{ex} to compute the entries of the correlation matrix $Z^{(p)}$ used in the relax-and-round approach. After collecting n_{ex} bit strings $\{b \in \{0, 1\}^N\}^{n_{\text{ex}}}$, expectation values are estimated as

$$\langle \hat{Z}_i \hat{Z}_j \rangle \approx \frac{1}{n_{\text{ex}}} \sum_{\{z\}} \langle b | \hat{Z}_i \hat{Z}_j | b \rangle, \quad (\text{G1})$$

with an error decreasing as $\sim 1/\sqrt{n_{\text{ex}}}$ as per the central limit theorem. The effect of a finite number of circuit executions on

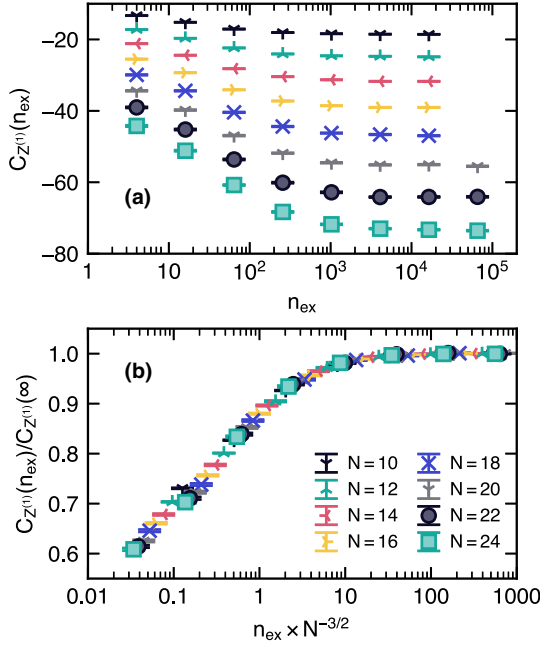


FIG. 6. Data for the relax-and round algorithm for SK problem instances based on the QAOA at $p = 1$. Each data point is averaged between 10^3 to 10^4 independent random problem instances. (a) Average objective function at finite n_{ex} for different problem sizes N (legend displayed in the other panel). (b) Average ratio of the objective function at finite n_{ex} over $n_{\text{ex}} = \infty$ versus $n_{\text{ex}} N^{-3/2}$.

the correlation matrix can be modeled by a random component

$$\mathbf{Z}_{i \neq j}^{(p)}(n_{\text{ex}}) = \mathbf{Z}_{i \neq j}^{(p)}(\infty) + \frac{\mathcal{N}(0, 1)}{\sqrt{n_{\text{ex}}}}, \quad (\text{G2})$$

with $\mathcal{N}(0, 1)$ a random normal variable of mean zero and variance unity such that $\mathbf{Z}^{(p)}$ is symmetric.

The robustness of performing an eigendecomposition of noisy matrices has a rich history. We suggest Ref. [63] for a recent review. The size N of the matrix, the strength of the noise $1/\sqrt{n_{\text{ex}}}$, the level spacing between eigenvalues, and the condition number of the matrix are parameters typically playing a role on the distance between the eigenvectors of $\mathbf{Z}^{(p)}(n_{\text{ex}})$ and those of $\mathbf{Z}^{(p)}(\infty)$. Here, the fact that we round the entries after computing the eigenvectors complicates the analysis. A noisy toy model for clustering, based on the stochastic block model [63], suggests that rounding enhances robustness to noise.

In the absence of a theoretical framework for a more general case, we rely on numerical experiments. We run the QAOA at $p = 1$ for SK problem instances with a finite number of circuit executions to evaluate the correlation matrix $\mathbf{Z}^{(p=1)}(n_{\text{ex}})$. We consider $n_{\text{ex}} = 4, 16, 64, 256, 1,024, 4,096, 16,384, \text{ and } 65,536$. We then perform the relax-and-round approach and report the data in Fig. 6. We find that the average objective function value follows as scaling relation relating n_{ex} to the number of variables N

$$C_{\mathbf{Z}^{(1)}}(n_{\text{ex}}) = C_{\mathbf{Z}^{(1)}}(\infty) \mathcal{G}(n_{\text{ex}} N^{-3/2}), \quad (\text{G3})$$

where $\mathcal{G}(x)$ is a scaling function such that $\mathcal{G}(x \rightarrow +\infty) = 1$. Hence, achieving a constant relative performance with respect

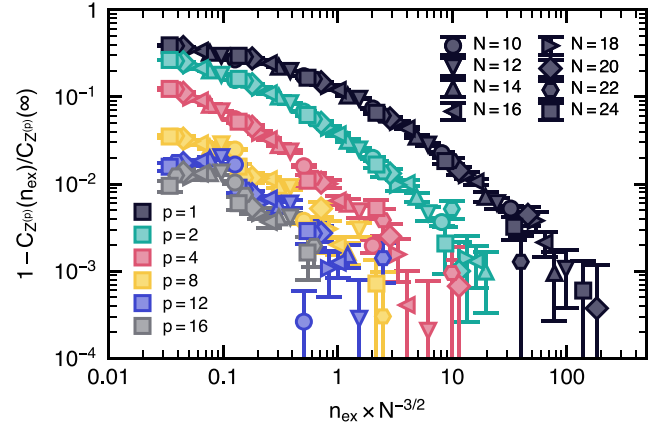


FIG. 7. Data for the relax-and round algorithm for SK problem instances based on the QAOA at depth p . Each data point is averaged between 10^3 to 10^4 independent random problem instances. Distance to optimality of the objective function value at finite n_{ex} over $n_{\text{ex}} = \infty$ versus $n_{\text{ex}} N^{-3/2}$. The same data for $p = 1$ are also displayed in Fig. 6.

to the ideal $n_{\text{ex}} = \infty$ case only requires a polynomial number of circuit executions with respect to the problem size N . We empirically find that this number is well-fitted by an exponent $\approx 3/2$, i.e., $n_{\text{ex}} \sim N^{3/2}$. It remains an open question to theoretically establish this scaling law, and in particular the exponent $\approx 3/2$.

We extend the analysis by running the same numerical experiments for various values of the QAOA depth p . In the $p \rightarrow +\infty$ limit, a single circuit execution is enough to capture the correlation matrix that will lead to the optimal solution. This corresponds to $\kappa = 0$. We show in Fig. 7 that for a fixed accuracy, the number of circuit executions required decreases as p increases. Moreover, the scaling relation with $\kappa \approx 1.5$ does not hold as p increases, in line with the expected behavior as $p \rightarrow +\infty$.

APPENDIX H: ALGORITHMIC TIME COMPLEXITY OF THE RELAX-AND-ROUND STEPS

We investigate the actual clock time it takes to run the classical relax-and-round steps of the QRR algorithm on SK problem instances of size N given n_{ex} input bit strings. The steps include: Building the $N \times N$ correlation matrix from the input of n_{ex} bit strings, finding its $k = 5$ leading eigenvectors, sign-rounding those eigenvectors, and computing the cost of each of these rounded eigenvectors in order to find the best one. Because we cannot easily sample bit strings from the QAOA at arbitrary N , for the practical purpose of this investigation, the bit strings used for building the correlation matrix are generated at random. However, the eigendecomposition is performed on the correlation matrix computed from the formulas of Appendix C, to ensure its genuine properties that will matter for the convergence of the eigensolver.

The dominant algorithmic complexity scaling of these steps is $O(N^2)$, as observed in Fig. 8. Besides, we note that the various steps can be massively parallelized to reduce the actual clock time, if needed. The total clock time for the QRR algorithm is the sum of the algorithmic time computed

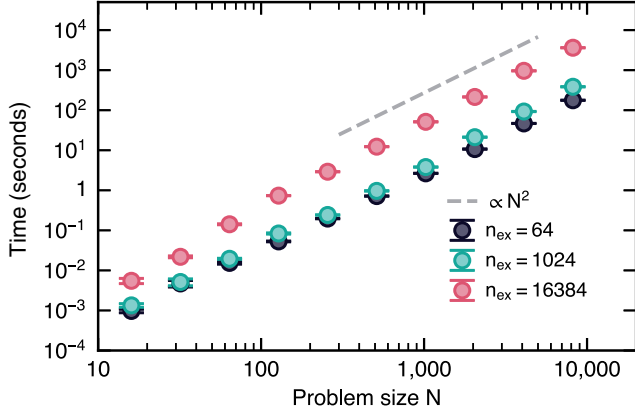


FIG. 8. Time in seconds to execute the relax-and-round steps as a function of the problem size N given an input of n_{ex} bit strings. SK problem instances are considered. The time is evaluated from a PYTHON code leveraging NUMPY and SCIPY running on an Apple M1 Max 64GB laptop (see Sec. D). The steps include: building the $N \times N$ correlation matrix from the input of n_{ex} bit strings, finding its $k = 5$ leading eigenvectors, sign-rounding those eigenvectors, and computing the cost of each of these rounded eigenvectors in order to find the best one. The dominant algorithmic complexity scaling of these steps is $O(N^2)$. Each data point is averaged over 10 to 10^3 independent runs.

here and the time taken to execute the n_{ex} circuits on actual quantum hardware. In addition, there might be an overhead in optimizing the variational parameters of the circuit when using the QRR in combination with the QAOA algorithm.

APPENDIX I: SIZE DEPENDENCE FOR THE DATA OF FIGURE 2 IN THE MAIN TEXT AT $p = 1$

Figure 4 of the main text suggests that, on average, the QRR algorithm on $Z^{(p=1)}$ performs better than the classical RR on W for the NWS and BA problem types at $N = 16$. Here, we consider different problem sizes $N \leq 256$ with the QAOA depth $p = 1$ using the analytical formula.

Data of Fig. 9 seems to confirm, on average, the superiority of the QRR algorithm for the NWS and BA problem types: the ratio of the objective values—or equivalently the ratio of the approximation ratios—is larger than 1. For the SK and 3REG problem instances, the two approaches perform, on average, equally well. We prove analytically in this work that this is expected for the SK models in the $N \rightarrow +\infty$ limit and numerical experiments show this result is robust to finite N .

APPENDIX J: RELAX-AND-ROUND FOR THE MAXIMUM INDEPENDENT SET PROBLEM WITH QUANTUM ANNEALING

The objective function to minimize for the weighted maximum independent set problem is given by [64]

$$\min_{z_i = \pm 1} C(z) = J \sum_{i < j} W_{ij} z_i z_j + \sum_i [J \deg(i) - 2u_i] z_i, \quad (\text{J1})$$

where $W_{ij} = 0$ or 1 depending whether two vertices of the graph problem are connected by an edge, $u_i \in \mathbb{R}^+$ a vertex-dependent weight, $\deg(i)$ is the degree of vertex i , and $J >$

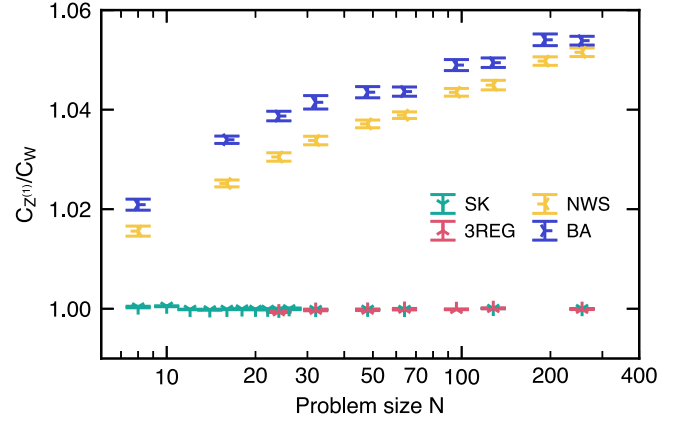


FIG. 9. Ratio of the objective values from the relax-and-round approach based on the correlation matrix $Z^{(p=1)}$ and the adjacency matrix W as a function of the problem size N . Problem instances considered: Sherrington-Kirkpatrick with random ± 1 weights (SK), random 3-regular graphs with unit weight (3REG), random Newman-Watts-Strogatz small-world graphs with weights uniformly distributed on the unit line (NWS), and random Barabási-Albert graphs with weights drawn from a normal distribution with mean zero and unit variance (BA). Each data point is averaged between 10^3 to 10^5 independent random problem instances.

$\min(u_i) \forall i$ a parameter. We draw $u_i \in [0, 1]$ at random from a uniform distribution and choose $J = 2$. The maximum independent set of the graph is given by the variables with the value $z_i = +1$ in the optimal solution z_{opt} . We consider random unit disks with N variables and density parameter $\rho = 7$.

Unlike the rest of this work focusing on the QAOA, we now use a quantum annealing protocol [6,7] to obtain the correlation matrix Z . We define the Hamiltonian

$$\hat{H}(T, t) = -\left(1 - \frac{t}{T}\right) \sum_{j=1}^N \hat{X}_j + \frac{t}{T} \hat{C}, \quad (\text{J2})$$

where the operator \hat{C} is obtained by substituting the binary variables z_j for Pauli operators \hat{Z}_j , \hat{X}_j is the Pauli operator, $t \in [0, T]$ the time, and T the total evolution time. The quantum state at time t reads

$$|\Psi\rangle_T = \mathcal{T} \exp\left[-i \int_0^T dt \hat{H}(T, t)\right] \hat{H}^{\otimes N} |0\rangle^{\otimes N}, \quad (\text{J3})$$

where \hat{H} is the Hadamard gate and \mathcal{T} indicates a time-ordered exponential. In the limit $T \rightarrow +\infty$, the quantum state will converge to the ground state of the objective function \hat{C} , i.e., the optimal solution z_{opt} to the combinatorial optimization problem. In practice, we discretize the above unitary by introducing a finite time step δt

$$|\Psi\rangle_T = \left[\prod_{\ell=1}^p e^{i\delta_t \beta_\ell \sum_{j=1}^N \hat{X}_j} e^{-i\delta_t \gamma_\ell \hat{C}} \right] \hat{H}^{\otimes N} |0\rangle^{\otimes N}, \quad (\text{J4})$$

where $p = T/\delta t$, $\beta_\ell = 1 - \ell\delta t/T$, and $\gamma_\ell = \ell\delta t/T$. From the quantum state time-evolved for a total time T , we compute the correlation matrix $Z^{(T)}$ on which to perform the QRR algorithm, just like we did with the QAOA.

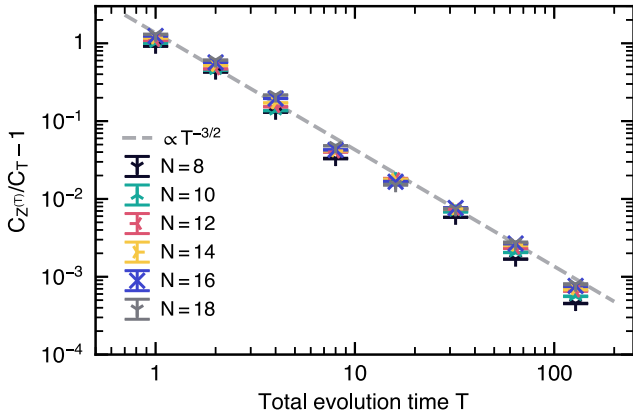


FIG. 10. Ratio of the objective values from the relax-and-round on $Z^{(T)}$ and the expectation value of the quantum state $\langle \hat{C} \rangle_T$ as a function of the total evolution time T for different problem sizes N . We consider random unit disk graphs with density $\rho = 7$ where the objective function is that of a weighted maximum independent set problem. Each data point is averaged over 10^3 to 10^4 independent random problem instances.

In the following, we use $\delta_i = 0.1$ and perform numerical experiments. We consider the objective function value for the relax-and-round divided by the expectation value $\langle \hat{C} \rangle_T$. Data plotted in Fig. 10 converge to zero as $T \rightarrow +\infty$ since both methods are expected to return the optimal solution. At finite T , the QRR algorithm systematically outperforms, on average, the quantum annealing protocol. This advantage increases with N . Note that we do not post-process for independent sets. We observe that the advantage of the QRR algorithm over the underlying quantum annealing protocol reduces algebraically with T , approximately as $\approx T^{-3/2}$.

APPENDIX K: A QUANTUM RELAX-AND-ROUND VERSION OF THE GOEMANS-WILLIAMSON ALGORITHM

1. The $p \rightarrow +\infty$ limit

In the $p \rightarrow +\infty$ limit, when the QAOA algorithm returns the optimal solution \mathbf{z}_{opt} , the correlation matrix takes the form $Z^{(\infty)} = \mathbb{I} - \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$. The matrix in the bracket becomes $\mathbb{D} - \mathbb{I} + \text{diag}(\mathbf{u}) + \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$ where $\mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$ is a rank one matrix. What is the optimal correcting vector \mathbf{u}_{opt} such that

the maximum eigenvalue is minimized? More generally, this eigenvalue problem is that of a diagonal matrix modified by a rank one matrix which has been studied mathematically. We denote $d_i = D_{ii} - 1 + u_i$ with $i = 1, \dots, N$ the diagonal entries of $\mathbb{D} = \mathbb{D} - \mathbb{I} + \text{diag}(\mathbf{u})$ and λ_i the eigenvalues of $\mathbb{D} + \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$. Eigenvalues are ordered such that $\lambda_i \leq \lambda_{i+1}$ and $d_i \leq d_{i+1}$. It can be shown that [65,66]

$$\begin{aligned} d_i &\leq \lambda_i \leq d_{i+1} \quad \text{for } i = 1, 2, \dots, N-1, \\ d_N &\leq \lambda_N \leq d_N + \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}. \end{aligned} \quad (\text{K1})$$

The goal is to minimize the eigenvalue λ_N bounded from below by $d_N = D_{NN} - 1 + u_N$, and where d_N is at least as large as d_{N-1} , etc. We recall the constraint $\sum_i u_i = 0$. Hence, the optimal solution is such that d_i are a constant for all i , which is given by $\mathbb{D} + \text{diag}(\mathbf{u}_{\text{opt}}) = \text{tr}(\mathbb{D})\mathbb{I}/N$. Thus the leading eigenvector for the optimal correcting vector is that of the matrix $-\mathbb{I} + \text{tr}(\mathbb{D})\mathbb{I}/N + \mathbf{z}_{\text{opt}} \otimes \mathbf{z}_{\text{opt}}$, which is simply $\pm \mathbf{z}_{\text{opt}}/\sqrt{N}$ as a constant diagonal matrix is irrelevant for computing eigenvectors.

Therefore the above algorithm substituting the adjacency matrix \mathbb{W} for the correlation matrix $Z^{(p)}$ solves the problem exactly in the limit $p \rightarrow +\infty$.

2. The case of finite p

It is more difficult to establish performance bounds for the finite p case for arbitrary graphs. We focus on vertex-transitive graphs. For such graphs, it is known that the optimal correcting vector for the Goemans-Williamson algorithm is null, i.e., $\mathbf{u}_{\text{opt}} = \mathbf{0}$ [41–43]. The correlation matrix $Z^{(p)}$ will have the same symmetries as the adjacency matrix: The correlation matrix can be pictured as the adjacency matrix of a graph which will also be vertex-transitive. It follows that the optimal vector is also $\mathbf{u}_{\text{opt}} = \mathbf{0}$. From there, demonstrating that \mathbb{W} and $Z^{(p)}$ share the same leading eigenvector is enough for the algorithm based on either \mathbb{W} or $Z^{(p)}$ to be equivalent. We have shown this is the case for different graphs such as the ring and complete graphs.

We emphasize that the performance guarantee of the classical Goemans-Williamson algorithm is agnostic to the graph, and showing whether a similarly strong statement can be made for a quantum relax-and-round version remains an open question. Here, we used the fact that we can show that the optimal correcting vector is the same in the classical and quantum versions, but this is not necessary for the two methods to have equivalent performance.

[1] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Courier, North Chelmsford, 1998).
 [2] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms* (Springer, Berlin Heidelberg New York, 2012), Vol. 2.
 [3] A. Lucas, *Front. Phys.* **2**, 5 (2014).
 [4] F. Glover, G. Kochenberger, and Y. Du, [arXiv:1811.11538](https://arxiv.org/abs/1811.11538).
 [5] G. Kochenberger, J.-K. Hao, F. Glover, M. Lewis, Z. Lü, H. Wang, and Y. Wang, *J. Comb. Optim.* **28**, 58 (2014).

[6] T. Kadowaki and H. Nishimori, *Phys. Rev. E* **58**, 5355 (1998).
 [7] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, *Science* **292**, 472 (2001).
 [8] E. Farhi, J. Goldstone, and S. Gutmann, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
 [9] E. Farhi, J. Goldstone, and S. Gutmann, [arXiv:1412.6062](https://arxiv.org/abs/1412.6062).
 [10] E. Farhi and A. W. Harrow, [arXiv:1602.07674](https://arxiv.org/abs/1602.07674).
 [11] K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, [arXiv:2306.09198](https://arxiv.org/abs/2306.09198).
 [12] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, M.

- Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney *et al.*, *Nat. Phys.* **17**, 332 (2021).
- [13] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, P. Karalekas, C. B. Osborn, A. Papageorge, E. C. Peterson, G. Prawiroatmodjo, N. Rubin, C. A. Ryan, D. Scarabelli, M. Scheer, E. A. Sete *et al.*, [arXiv:1712.05771](https://arxiv.org/abs/1712.05771).
- [14] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, L. T. Brady, A. Deshpande, F. Liu, S. Jordan, A. V. Gorshkov, and C. Monroe, *Proc. Natl. Acad. Sci. USA* **117**, 25396 (2020).
- [15] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler *et al.*, *Science* **376**, 1209 (2022).
- [16] A. Byun, M. Kim, and J. Ahn, *PRX Quantum* **3**, 030305 (2022).
- [17] M. Kim, K. Kim, J. Hwang, E.-G. Moon, and J. Ahn, *Nat. Phys.* **18**, 755 (2022).
- [18] A. D. King, J. Raymond, T. Lanting, R. Harris, A. Zucca, F. Altomare, A. J. Berkley, K. Boothby, S. Ejtemaee, C. Enderud, E. Hoskinson, S. Huang, E. Ladizinsky, A. J. R. MacDonald, G. Marsden, R. Molavi, T. Oh, G. Poulin-Lamarre, M. Reis, C. Rich *et al.*, *Nature (London)* **617**, 61 (2023).
- [19] Y. R. Sanders, D. W. Berry, P. C. Costa, L. W. Tessler, N. Wiebe, C. Gidney, H. Neven, and R. Babbush, *PRX Quantum* **1**, 020312 (2020).
- [20] D. P. Williamson and D. B. Shmoys, *The Design of Approximation Algorithms* (Cambridge University Press, New-York, 2011).
- [21] M. X. Goemans and D. P. Williamson, *J. ACM* **42**, 1115 (1995).
- [22] N. Christofides, *Worst case analysis of a new heuristic for the traveling salesman problem*, Tech. Rep. 388 (Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA, 1976).
- [23] R. van Bevern and V. A. Slugina, *Hist. Math.* **53**, 118 (2020).
- [24] J. Wurtz and P. J. Love, *Quantum* **6**, 635 (2022).
- [25] S. J. Devitt, W. J. Munro, and K. Nemoto, *Rep. Prog. Phys.* **76**, 076001 (2013).
- [26] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, *Phys. Rev. A* **97**, 022304 (2018).
- [27] D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [28] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, *Quantum* **6**, 759 (2022).
- [29] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou, in *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), edited by F. Le Gall and T. Morimae (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022), Vol. 232, pp. 7:1–7:21.
- [30] G. Parisi, *Phys. Rev. Lett.* **43**, 1754 (1979).
- [31] M. Aizenman, J. L. Lebowitz, and D. Ruelle, *Commun. Math. Phys.* **112**, 3 (1987).
- [32] A. Montanari and S. Sen, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, Cambridge MA USA, STOC '16* (ACM, Cambridge, USA, 2016), pp. 814–827.
- [33] A. S. Bandeira, D. Kunisky, and A. S. Wein, in *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 151 (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020), pp. 78:1–78:29.
- [34] B. O’Gorman, W. J. Huggins, E. G. Rieffel, and K. B. Whaley, [arXiv:1905.05118](https://arxiv.org/abs/1905.05118).
- [35] A. Montanari, *SIAM J. Comput.*, FOCS19-1 (2019).
- [36] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, *Phys. Rev. X* **10**, 021067 (2020).
- [37] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Nat. Commun.* **9**, 4812 (2018).
- [38] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, New-York, 2010).
- [39] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O’Brien, [arXiv:2210.00921](https://arxiv.org/abs/2210.00921).
- [40] B. Mohar and S. Poljak, *Czech. Math. J.* **40**, 343 (1990).
- [41] C. Delorme and S. Poljak, *Math. Program.* **62**, 557 (1993).
- [42] C. Delorme and S. Poljak, *Discrete Math.* **111**, 145 (1993).
- [43] S. Poljak and F. Rendl, *Discret. Appl. Math.* **62**, 249 (1995).
- [44] H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin, [arXiv:1808.10816](https://arxiv.org/abs/1808.10816).
- [45] M.-T. Nguyen, J.-G. Liu, J. Wurtz, M. D. Lukin, S.-T. Wang, and H. Pichler, *PRX Quantum* **4**, 010316 (2023).
- [46] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, *Phys. Rev. Lett.* **125**, 260505 (2020).
- [47] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, *Quantum* **6**, 678 (2022).
- [48] F. Wagner, J. Nüßlein, and F. Liers, [arXiv:2302.05493](https://arxiv.org/abs/2302.05493).
- [49] M. Dupont, B. Evert, M. J. Hodson, B. Sundar, S. Jeffrey, Y. Yamaguchi, D. Feng, F. B. Maciejewski, S. Hadfield, M. S. Alam, Z. Wang, S. Grabbe, P. A. Lott, E. G. Rieffel, D. Venturelli, and M. J. Reagor, *Sci. Adv.* **9**, eadi0487 (2023).
- [50] A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th PYTHONin Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, CA USA, 2008), pp. 11–15.
- [51] M. Newman and D. Watts, *Phys. Lett. A* **263**, 341 (1999).
- [52] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [53] W. Selke, *Phys. Rep.* **170**, 213 (1988).
- [54] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant *et al.*, *Nature (London)* **585**, 357 (2020).
- [55] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, *Nat. Meth.* **17**, 261 (2020).
- [56] S. K. Lam, A. Pitrou, and S. Seibert, in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC* (2015), pp. 1–6.
- [57] S. Diamond and S. Boyd, *J. Mach. Learn. Res.* **17**, 1 (2016).
- [58] A. Agrawal, R. Verschuere, S. Diamond, and S. Boyd, *J. Control. Decis.* **5**, 42 (2018).
- [59] C. G. Broyden, *IMA J. Appl. Math.* **6**, 76 (1970).
- [60] R. Fletcher, *Comput. J.* **13**, 317 (1970).

- [61] D. Goldfarb, *Math. Comp.* **24**, 23 (1970).
- [62] D. F. Shanno, *Math. Comp.* **24**, 647 (1970).
- [63] Y. Chen, Y. Chi, J. Fan, and C. Ma, *Found. Trends Mach. Learn.* **14**, 566 (2021).
- [64] V. Choi, *Quantum Inf. Process.* **7**, 193 (2008).
- [65] A. R. Mitchell, *Proc. Edinburgh Math. Soc.* **15**, 328 (1967).
- [66] G. H. Golub, *SIAM Rev.* **15**, 318 (1973).