

## Optimization via quantum preconditioning

Maxime Dupont<sup>✉,\*</sup>, Tina Oberoi<sup>✉,†</sup> and Bhuvanesh Sundar<sup>✉</sup>  
*Rigetti Computing, 775 Heinz Avenue, Berkeley, California 94710, USA*



(Received 27 March 2025; revised 13 August 2025; accepted 12 September 2025; published 3 October 2025)

State-of-the-art classical optimization solvers set a high bar for quantum computers to deliver utility in this domain. Here, we introduce a quantum-preconditioning approach based on the quantum approximate optimization algorithm. This transforms the input problem into a more suitable form for a solver, with the level of preconditioning determined by the depth of the quantum circuit. We demonstrate that best-in-class classical heuristics such as simulated annealing and the Burer-Monteiro algorithm can converge more rapidly when given quantum-preconditioned input for various problems, including Sherrington-Kirkpatrick spin glasses, random 3-regular graph maximum-cut problems, and a real-world grid energy problem. Accounting for the additional time taken for preconditioning, the benefit offered by shallow circuits translates into a practical quantum-inspired advantage for random 3-regular graph maximum-cut problems through quantum circuit emulations. We investigate why quantum preconditioning makes the problem easier and test an experimental implementation on a superconducting device. We identify challenges and discuss the prospects for a hardware-based quantum advantage in optimization via quantum preconditioning.

DOI: [10.1103/9prw-684p](https://doi.org/10.1103/9prw-684p)

### INTRODUCTION

Optimization problems over discrete variables are ubiquitous [1–3]; however, they continue to pose a significant challenge for traditional computing paradigms. The absence of provably efficient algorithms in the general case unless  $P = NP$  [4,5] makes heuristic solvers the practical method of choice. Celebrated examples include physics-inspired approaches, such as simulated annealing (SA) [6] or parallel tempering [7], as well as heuristics that build on approximate methods with performance guarantees, like the Burer-Monteiro (BM) solver [8,9] based on the Goemans-Williamson (GW) algorithm for the maximum-cut problem [10,11]. These heuristics represent the state of the art in discrete optimization on a range of problems [12].

The advent of quantum technologies as a new computing paradigm [13,14] has sparked interest in exploring whether discrete optimization can leverage quantum phenomena for improved efficiency [15]. On the one hand, quantum adiabaticity [16,17] serves as the foundation for quantum annealers and, more broadly, analog quantum computation [18,19]. Analog platforms, through empirical observation, have demonstrated an asymptotic scaling advantage over classical methods such as SA and parallel

tempering on specific classes of problems [20–23]; however, an absolute advantage in computational time has not yet been realized at currently addressed problem sizes on current quantum hardware. On the other hand, digital quantum computers operating via programmable quantum logical operations typically leverage the variational principle to address discrete optimization [24]. In that context, a relevant algorithm is the quantum approximate optimization algorithm (QAOA) [25,26]. In certain limiting cases and for specific problem types, the QAOA can provide performance guarantees on the accuracy of the returned solution [25,27–30]. Additionally, the QAOA can be used as a heuristic, where classical emulations support evidence of a scaling advantage for certain problems that is yet to be realized in practice [31,32]. In addition to the QAOA, other quantum algorithms with a potential scaling advantage have been advanced [33–36]. While the evidence for a scaling advantage is promising, transforming this into a tangible advantage for quantum computers to deliver utility remains a substantial engineering challenge. Alongside hardware development, further progress in quantum algorithms is equally important in this endeavor.

Here, we introduce the idea of quantum preconditioning for discrete optimization problems (Sec. II). In the classical context, preconditioning seeks to transform a problem into a more suitable form for a solver. It aims to make the preconditioned problem easier to solve than its original version by, e.g., converging faster to a solution. A notable example arises in state-of-the-art classical linear system

\*Contact author: [mdupont@rigetti.com](mailto:mdupont@rigetti.com)

†Current affiliation: University of Chicago, Chicago, Illinois, USA.

solvers, where classical preconditioning is instrumental to their performance [37,38]. A related idea for solving mixed integer programming problems is “presolving” [39,40]: this involves preprocessing techniques being applied to the input problem to make it more tractable for a solver.

We propose a quantum preconditioner for quadratic unconstrained binary optimization problems [41] based on the QAOA with  $p$  layers, which works as follows. The goal is to extremize a quadratic objective function  $C(\mathbf{z})$  over  $N$  discrete variables  $\mathbf{z} = (z_1, z_2, \dots, z_N)^T$ , where the objective function is encoded by an adjacency matrix  $\mathbf{W} \in \mathbb{R}^{N \times N}$ . First, we execute the QAOA on the original  $N$ -variable problem encoded by  $\mathbf{W}$ . Then, we use the resulting quantum state to estimate two-point correlations between qubits as a pairwise correlation matrix  $\mathbf{Z}^{(p)} \in \mathbb{R}^{N \times N}$ . Finally, we replace the adjacency matrix  $\mathbf{W}$  with the correlation matrix  $\mathbf{Z}^{(p)}$  in the problem’s objective function. Hence, the quantum preconditioner preserves the structure of quadratic unconstrained binary optimization, and the correlation matrix is the quantum-preconditioned problem solved in place of the original problem using a suitable classical solver. This allows classical solvers to work with either the original input or a preconditioned version, potentially leading to a direct improvement over the current state-of-the-art classical approach.

Building on the adiabatic theorem, we demonstrate that in the adiabatic limit ( $p \rightarrow +\infty$ ) [16], the quantum transformation renders solving the problem optimally a trivial task. Any local updates that improve a candidate solution  $\mathbf{z}$  lead to the optimal solution  $\mathbf{z}_{\text{opt}}$ , thereby benefiting classical solvers like SA and parallel tempering. In the opposite limit, at a shallow circuit depth  $p \leq 2$ , we perform classical emulations of the QAOA to test the performance of the proposed quantum preconditioner. We find that best-in-class classical heuristics can converge more rapidly on an average case. In particular, SA and the BM algorithm can benefit from low-depth quantum preconditioning in solving a range of problems, such as Sherrington-Kirkpatrick spin glasses [42], random 3-regular graph maximum-cut problems, and a real-world grid energy problem [43].

Using a light-cone decomposition technique [44], we benchmark the proposed approach on problems with thousands of variables and investigate the effect of increased circuit depth on its convergence. For instance, on an ensemble of random 3-regular graph maximum-cut problems with  $N = 4096$  variables, quantum preconditioning can make BM and SA converge one order of magnitude faster to a typical solution  $\mathbf{z}$  with an average approximation ratio  $\alpha = C(\mathbf{z})/C(\mathbf{z}_{\text{opt}}) = 99.9\%$  (Sec. III). When accounting for the additional time taken for preconditioning, the benefit offered by shallow circuits remains and translates into a practical advantage for random 3-regular graph maximum-cut problems through classical QAOA emulations. To the best of our knowledge, this is the

first reported quantum-inspired advantage for optimization. The improved convergence observed upon increasing the circuit depth opens the door to a hardware-based quantum advantage using circuits that are beyond the reach of classical emulators for preconditioning.

In addition, we study why quantum preconditioning makes the problem easier by considering small- $N$  problems at large  $p$  (Appendix B). We test an experimental implementation of quantum preconditioning of the real-world grid energy problem [43] on a Rigetti Ankaa-3 superconducting device (Appendix D). Finally, we identify challenges and discuss prospects for quantum preconditioning to deliver utility (Sec. IV). We give our conclusions in Sec. V.

## QUANTUM PRECONDITIONING

We consider discrete optimization problems of the form [41]

$$C(\mathbf{z}) = \frac{1}{2} \mathbf{z}^T \mathbf{W} \mathbf{z} = \frac{1}{2} \sum_{i,j=1}^N W_{ij} z_i z_j, \quad (1)$$

with variables  $\mathbf{z} \in \{\pm 1\}^N$  and the symmetric matrix  $\mathbf{W} \in \mathbb{R}^{N \times N}$  encoding the problem. The quantum preconditioner computes the two-point correlation matrix  $\mathbf{Z}$  [see Fig. 1(a)]

$$W_{ij} \leftarrow Z_{ij}^{(p)} = (\delta_{ij} - 1) \langle \hat{Z}_i \hat{Z}_j \rangle_p, \quad (2)$$

where  $\delta_{ij}$  is the Kronecker delta and the expectation value involving Pauli- $Z$  operators is evaluated over the quantum state

$$|\Psi\rangle_p = \left[ \prod_{\ell=1}^p e^{-i\beta_\ell \sum_{j=1}^N \hat{X}_j} e^{-i\gamma_\ell \hat{C}} \right] \hat{H}^{\otimes N} |0\rangle^{\otimes N}, \quad (3)$$

where  $\hat{H}$ ,  $\hat{X}$ , and  $\langle \mathbf{z} | \hat{C} | \mathbf{z}' \rangle = C(\mathbf{z}) \delta_{\mathbf{z}\mathbf{z}'}$  are the Hadamard gate, Pauli- $X$  operator, and operator encoding the objective function, respectively. There is a one-to-one mapping between a variable and a qubit. Equation (3) corresponds to a standard  $p$ -layer QAOA circuit [25,26] parameterized via  $\gamma_\ell$  and  $\beta_\ell$ , which are optimal when extremizing  $\langle \hat{C} \rangle_p$ . In a similar form to the original problem of Eq. (1), the preconditioned problem’s objective function reads  $\tilde{C}^{(p)}(\mathbf{z}) = \frac{1}{2} \mathbf{z}^T \mathbf{Z}^{(p)} \mathbf{z}$ . In the following, quantum preconditioning is performed at optimal or near-optimal  $\gamma_\ell$  and  $\beta_\ell$  values with respect to extremizing  $\langle \hat{C} \rangle_p$  via the QAOA. Hence, the computational cost of preconditioning is the same as executing the QAOA. No additional information than what is available as part of the standard QAOA is needed.

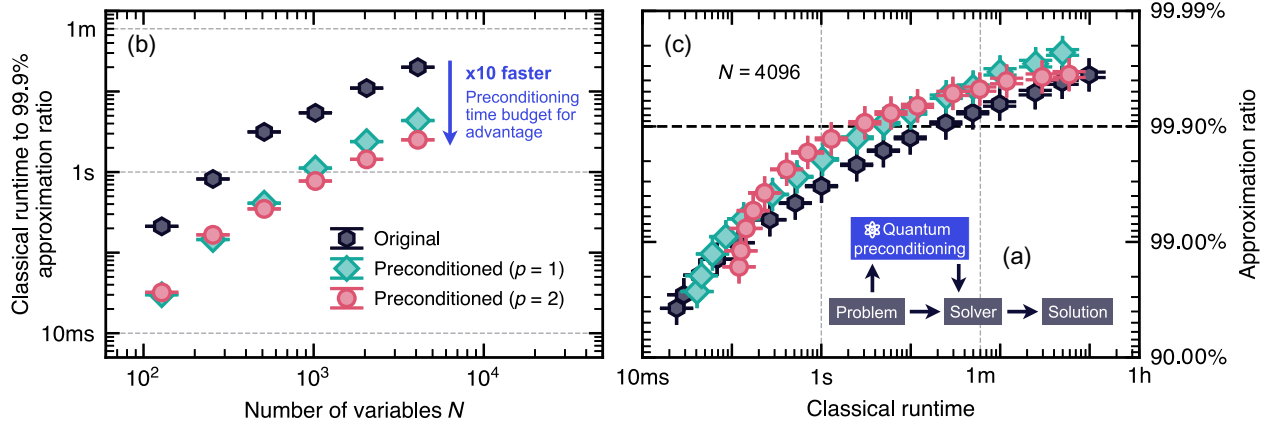


FIG. 1. Average performance for the maximum-cut problem on  $N$ -variable random 3-regular graphs via the classical simulated annealing (SA) solver based on the original and quantum-preconditioned problems. (a) Workflow diagram of quantum preconditioning. (b) Average run time for SA to get to an approximation ratio of 99.9% as a function of  $N$ . (c) Average approximation ratio versus SA run time for  $N = 4096$ . Each data point is averaged over 200 randomly generated problem instances. Run times correspond to a 64 GB MacBook Pro with an Apple M1 Max chip. Error bars indicate the standard error of the mean.

### The asymptotic infinite-circuit-depth limit

At infinite depth ( $p \rightarrow +\infty$ ), the QAOA converges to the optimal solution [25]. In the presence of a doubly degenerate optimal solution  $\pm \mathbf{z}_{\text{opt}}$  due to the  $\mathbb{Z}_2$  global sign-flip symmetry in the objective function of Eq. (1), the preconditioner takes the form  $\mathbf{Z}^{(\infty)} = \mathbf{I} - \mathbf{z}_{\text{opt}} \mathbf{z}_{\text{opt}}^T$ , where  $\mathbf{I}$  is an  $N \times N$  identity matrix. The objective function of the quantum-preconditioned problem now reads

$$\tilde{C}^{(\infty)}(\mathbf{z}) = \frac{1}{2} \mathbf{z}^T \mathbf{z} - |\mathbf{z}_{\text{opt}}^T \mathbf{z}|^2, \quad (4)$$

which is minimized for  $\mathbf{z} = \pm \mathbf{z}_{\text{opt}}$ . Therefore, both the original and preconditioned problems in the infinite-depth limit share the same optimal solution; however, while  $\mathbf{z}_{\text{opt}}$  is nontrivial to find based on the original problem of Eq. (1) in the general case, it can be straightforwardly obtained via local steps based on Eq. (4). Indeed,  $\tilde{C}^{(\infty)}(\mathbf{z}_{\text{opt}}) = -N(N-1)$  marks the absence of frustration in the quantum-preconditioned problem. Each variable  $z_i$  can be set to  $\pm 1$  to locally minimize each term  $\mathbf{Z}_{ij}^{(\infty)} z_i z_j$ , which minimizes the objective function globally. This signals the absence of glassy properties [45–47] in the quantum-preconditioned problem in the infinite-depth limit. Hence, we expect a strong benefit for solvers, such as SA [6], parallel tempering [7], and others that rely on local steps for convergence. This is evidenced by the decrease in frustration index of the preconditioned problem with QAOA circuit depth  $p$ , as shown in Appendix B for small- $N$  problem instances.

To extend the infinite-depth properties of quantum preconditioning to problems beyond the doubly degenerate optimal solution  $\pm \mathbf{z}_{\text{opt}}$ , another preconditioner is needed. For instance, for problems with a unique solution  $\mathbf{z}_{\text{opt}}$  due to an extra nonzero term of the form  $\sum_i h_i z_i$  with

$h_i \in \mathbb{R}$ , one may use  $\sum_i h_i z_i \leftarrow -\sum_i \langle \hat{Z}_i \rangle_p z_i$ . For problems with degenerate optimal solutions beyond  $\pm \mathbf{z}_{\text{opt}}$ , one can add vanishing one-body terms  $\lim_{h_i \rightarrow 0} h_i z_i$  to the original objective function to favor a single solution.

### At a shallow circuit depth

The performance at practical circuit depths may be independent of these infinite-depth considerations. In the following, we investigate the performance of the preconditioner of Eq. (2) on problems with at least a doubly degenerate optimal solution  $\pm \mathbf{z}_{\text{opt}}$  (potentially more) due to their global  $\mathbb{Z}_2$  sign-flip symmetry. Here, the optimal solution for  $\mathbf{Z}$  is generically different from the optimal solution for  $\mathbf{W}$ . Nevertheless, we find that this form of quantum preconditioning at shallow depth can still provide a benefit. Quantifying the quality of a solution  $\mathbf{z}$  through the approximation ratio  $\alpha = C(\mathbf{z})/C(\mathbf{z}_{\text{opt}})$ , we find that the classical solvers achieve  $\alpha \approx 1$  when working with the preconditioned problem, but they often do so with fewer iterations than with the original problem, which may translate to less computational time. This is the central result of this work. We emphasize that, although a solution  $\mathbf{z}$  is obtained by solving the original or preconditioned problem, the objective function  $C(\mathbf{z})$  and the optimal solution  $\mathbf{z}_{\text{opt}}$  are always with respect to the original problem for evaluating the approximation ratio  $\alpha$ .

The intuition behind the QAOA preconditioner of Eq. (2) at a shallow circuit depth  $p$  is that the correlation matrix contains similar information to the original problem. Indeed, locally, two variables  $z_i$  will tend to take values such that  $\mathbf{W}_{ij} z_i z_j$  is minimized, which translates into  $\text{sign}(\mathbf{W}_{ij}) \simeq -\text{sign}(\langle \hat{Z}_i \hat{Z}_j \rangle)$ . Thus, one may anticipate that solving a problem based on  $\mathbf{W}_{ij}$  or  $\mathbf{Z}_{ij}^{(p)}$  will lead to

similar solutions. A special case of quantum preconditioning is the quantum relax-and-round algorithm described in Refs. [44,48]. There, it was shown that a classical relax-and-round solver yields identical solutions for the original problem given by  $\mathbf{W}$  and the preconditioned problem at  $p = 1$  given by  $\mathbf{Z}^{(p=1)}$  on a range of problems, including Sherrington-Kirkpatrick spin glasses and random 3-regular graph maximum-cut problems.

The same intuition holds true for any classical solver. Here, we employ the state-of-the-art SA and BM solvers to solve Sherrington-Kirkpatrick spin glasses [42], random 3-regular graph maximum-cut problems, and a real-world grid energy problem [43]. We find that at small  $p$ , these solvers converge faster to a given value of the approximation ratio  $\alpha$  on the quantum-preconditioned versus the original problems, and conversely, they yield a higher approximation ratio  $\alpha$  in a given amount of time. In addition, in Appendix B, we investigate large- $p$  quantum preconditioning on small- $N$  problem instances.

## PRACTICAL PERFORMANCE

### Random 3-regular graph maximum-cut problems

#### *Classical simulated-annealing-solver baseline*

We investigate the average performance of SA [6] on hundreds of random  $N$ -variable 3-regular graph maximum-cut problems. SA is a physics-inspired solver treating an objective function such as Eq. (1) as a classical Ising model from which one wants to find the ground state. The search is based on a Metropolis-Hastings Markov chain Monte Carlo algorithm with a Boltzmann distribution [49,50]. It starts at a high temperature, which is reduced throughout the iterations according to a predefined schedule such that one samples low-energy states, and ideally the ground state (see Appendix E). The objective function of the maximum-cut problem is closely related to that of Eq. (1), as it seeks to maximize the cut number

$$C_{\text{MaxCut}}(\mathbf{z}) = \frac{1}{4} \sum_{i,j=1}^N W_{ij} (1 - z_i z_j), \quad (5)$$

where  $W_{ij} = 1$  if there is an edge between vertices  $i$  and  $j$  and zero otherwise. In Fig. 1(c), we evaluate the average approximation ratio  $\alpha$  as a function of the solver's run time for  $N = 4096$  variables. We estimate and report the run time to get to  $\alpha = 99.9\%$  as a function of  $N$  in Fig. 1(b).

#### *Performing quantum preconditioning*

Having defined a classical baseline, we now quantum precondition the problems using exact state-vector emulations. We employ a light-cone decomposition technique [44] to access two-point correlations  $\langle \hat{Z}_i \hat{Z}_j \rangle_p$  of Eq. (2) on problems much larger than qubits naively accessible.

A  $p$ -layer QAOA circuit induces quantum coherence around each qubit and its  $p$ -nearest neighbors. Therefore, one can estimate a two-point correlation between two qubits  $i$  and  $j$  via the QAOA by only considering the subset of qubits resulting from the intersection of their respective light cones. If the light cones do not intersect, then  $\langle \hat{Z}_i \hat{Z}_j \rangle_p = \langle \hat{Z}_i \rangle_p \langle \hat{Z}_j \rangle_p = 0$  due to the global sign-flip symmetry of Eq. (5). The topology of 3-regular graphs is such that intersecting light cones have at most  $1 + 6(2^p - 1)$  qubits, making  $p \leq 2$  readily accessible through standard circuit-emulation methods.

Moreover, because the average distance between two vertices grows as  $O(\ln N)$  in 3-regular graphs, only  $O(N)$  light cones will intersect in the limit  $p \ll \ln N$ , out of the  $O(N^2)$  naïve possibilities. In summary, we trade a single  $N$ -qubit QAOA circuit for  $O(N)$  QAOA circuits with  $O(\exp p)$  qubits each. Variational parameters are set to near-optimal tabulated values [29,51].

Quantum preconditioning increases the number of nonzero terms  $n$  in the problem because correlations extend beyond nearest-neighbor: There are  $n = 1.5N$  terms in the original 3-regular graph problem,  $n \simeq 4.5N$  at  $p = 1$ , and  $n \simeq 22.5N$  at  $p = 2$  in the large- $N$  limit (see Appendix G). Although classical solvers typically have a run time proportional to the number of terms  $n$  (see Appendix K), we show in the following that the increase in the number of terms is not a bottleneck for quantum preconditioning to deliver an advantage.

#### *Quantum-preconditioned simulated annealing solver*

In Fig. 1, we report the performance of quantum preconditioning on SA for hundreds of random 3-regular graph maximum-cut problems. We find that SA gets to an approximation ratio of 99.9% about one order of magnitude faster when working with the quantum-preconditioned input rather than the original problem. For the largest problem sizes considered, we also observe a separation between  $p = 1$  and  $p = 2$ , where increasing the circuit depth further accelerates the convergence of the classical solver. At  $p \rightarrow +\infty$ , we anticipate that SA will trivially converge to the optimal solution.

At a fixed number of variables ( $N = 4096$ ), Fig. 1(c) shows that quantum preconditioning can deliver a benefit across a range of approximation-ratio and run-time targets. As the run time increases, we observe that SA on the  $p = 2$  quantum-preconditioned problem saturates to an approximation ratio  $\alpha < 1$ . This is because, as mentioned earlier, at finite  $p$ , the optimal solutions of the original and quantum-preconditioned problems are generically different; however, it is interesting to see that in practice, their respective optimal solutions are spectrally close with the solver attaining  $\alpha \simeq 99.96\%$  at  $p = 2$ . We provide additional data for other values of  $N$  in Appendix J.

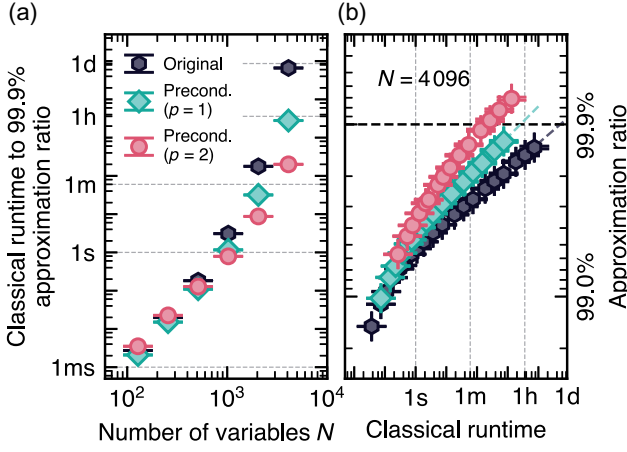


FIG. 2. Average performance for the maximum-cut problem on  $N$ -variable random 3-regular graphs via the classical BM solver based on the original and quantum-preconditioned problems. (a) Average run time for BM to get to an approximation ratio of 99.9% as a function of  $N$ . (b) Average approximation ratio versus BM run time for  $N = 4096$ . Each data point is averaged over 200 randomly generated problem instances. Run-times correspond to a 64 GB MacBook Pro with an Apple M1 Max chip. Error bars indicate the standard error of the mean.

### Quantum-preconditioned Burer-Monteiro solver

Next, we consider quantum preconditioning in the context of the BM solver [8,9]. The BM solver (see Appendix E) is inspired by semidefinite programming methods: it relaxes the discrete nature of the variables in the objective function, solves this modified problem, and rounds the solution back to the valid domain at the end. In practice, the BM algorithm is one of the best heuristics to date for the maximum-cut problem [12]. Depending on the problem size and the performance target, it may be a better choice than SA.

We observe in Fig. 2 that quantum preconditioning provides a run-time advantage to get to an average approximation ratio of  $\alpha = 99.9\%$  on the largest problem sizes. For example, for  $N = 4096$ , the BM solver takes only a few minutes to solve the preconditioned problem at  $p = 2$  and yield  $\alpha = 99.9\%$ , a notable improvement over 1 h required at  $p = 1$ , and about 1 day to solve the original problem. Although SA is more efficient than the BM solver in the  $\alpha = 99.9\%$  regime, including via quantum preconditioning, the hierarchy may change for a larger circuit depth  $p$ .

The performance boost offered by quantum preconditioning and observed on solvers as different as SA and BM suggests the wide applicability of the technique. In particular, the quantum relax-and-round solver [44,48] is another example of quantum preconditioning via a classical relax-and-round solver inspired by semidefinite programming methods. In that case, one can prove that the performance at  $p = 1$  matches that of the classical counterpart

and increases with  $p$  on classes of problems including Sherrington-Kirkpatrick spin glasses and random 3-regular graph maximum-cut problems [44,48].

### Sherrington-Kirkpatrick spin glasses

#### Problem and quantum preconditioning

We now turn our attention to another class of problems: Sherrington-Kirkpatrick (SK) spin glasses [42]. The goal is to minimize an objective function of the form of Eq. (1), where the adjacency matrix  $W$  encoding the problem is drawn from the Gaussian orthogonal ensemble. Namely, its entries are independent and identically distributed random variables from a normal distribution of zero mean and unit width  $W_{ij} \sim \mathcal{N}(0, 1)$ , with  $W_{ij} = W_{ji}$ . SK problem instances are quite different from the random 3-regular graph maximum-cut problems considered previously and constitute another standard optimization benchmark [27,52–55]. First, they are dense, with  $O(N^2)$  terms instead of  $O(N)$  for  $N$ -variable problems. Second, they are weighted problems with  $W_{ij} \neq 1$ .

We perform quantum preconditioning at  $p = 1$ , where one can compute the correlation matrix of Eq. (2) semi-analytically via a back-propagation technique [48] (reproduced in Appendix H for completeness). We use optimal QAOA angles  $\gamma = 1/2\sqrt{N}$  and  $\beta = \pi/8$  from the  $N \rightarrow +\infty$  limit [27], thus skipping variational optimization. Unlike the 3-regular graph maximum-cut problems, the resulting quantum-preconditioned SK spin glasses conserve the number of terms and the same statistics as the original  $W_{ij}$  [48]. Emulating the QAOA at  $p > 1$  for large  $N$  is currently not feasible.

#### Performance

In Fig. 3, we show the performance of SA and the BM solver on 200 randomly generated SK instances with  $N = 2048$  variables (additional data at other values of  $N$  are available in Appendix J). In particular, we plot the approximation ratio as a function of the number of iterations [Fig. 3(a)] and run time [Fig. 3(b)] with the relation between the two discussed in Appendix K.

We observe that for a large number of iterations (or run time), the approximation ratio obtained via the quantum-preconditioned problem saturates to  $\alpha \simeq 99.3\%$ . This, again, indicates that it does not share the same optimal solution as the original problem. For a small number of iterations, we find that the classical solver benefits from working with the preconditioned problem, achieving a higher approximation ratio at the same iteration count. This leads to a crossing of performance and an advantage window for quantum preconditioning on SK problems for both SA and BM solvers.

The overhead of about 2 s for the SA data in Fig. 3(b) originates from the  $O(N^2)$  heuristic used for setting the temperature schedule (see Appendix E). It is absent in the

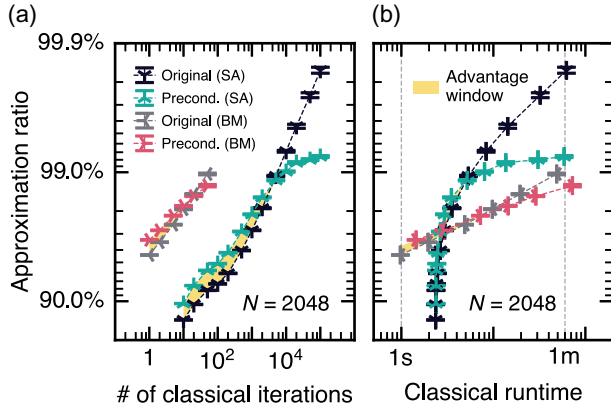


FIG. 3. Average performance for  $N = 2048$  Sherrington-Kirkpatrick spin-glass problems via the classical SA and Burer-Monteiro (BM) solvers based on the original and quantum-preconditioned ( $p = 1$ ) problems. (a) Average approximation ratio as a function of the number of iterations. (b) Average approximation ratio as a function of the run time. Each data point is averaged over 200 randomly generated problem instances. Run times correspond to a 64 GB MacBook Pro with an Apple M1 Max chip. Error bars indicate the standard error of the mean.

case of the BM solver, which does not have any tuned hyperparameters. The crossing of performance happens at slightly different approximation ratio values for SA and the BM algorithm. This can be understood as both solvers returning different distributions of solutions.

### Power-grid optimization problem: The maximum power exchange section

#### Problem definition

The last example considered is a real-world power-grid optimization problem seeking to compute the maximum power exchange section (MPES) of an energy network [56–59]. The MPES is a metric that provides information about the health and the power-delivery capability of the energy grid. We consider the realistic 500-bus power system dataset *ACTIVSg500*, which mimics the energy grid of the state of South Carolina in the United States of America [43]; see Fig. 4(a). The dataset was designed as part of the *ARPA-E's GRID DATA* program to be similar to the actual electric grid while containing no confidential information about critical energy infrastructure [60–64].

The MPES is a weighted maximum-cut problem. As such, it seeks to maximize the objective function

$$C_{\text{MPES}}(\mathbf{z}) = \frac{1}{4} \sum_{i,j=1}^N \mathbf{w}_{ij} (1 - z_i z_j), \quad (6)$$

where the weight  $\mathbf{w}_{ij}$  is given by the line impedance between two buses  $i$  and  $j$

$$\mathbf{w}_{ij} = \left( \mathbf{R}_{ij}^2 + \mathbf{X}_{ij}^2 \right)^{-1/2}, \quad (7)$$

with  $\mathbf{w}_{ji} = \mathbf{w}_{ij}$  and  $\mathbf{w}_{ij} = 0$  if buses  $i$  and  $j$  are not connected. Further,  $\mathbf{R}_{ij}$  and  $\mathbf{X}_{ij}$  correspond to the line resistance and reactance, respectively. Their values are provided as part of the dataset.

The original problem contains  $N = 410$  variables and  $n = 463$  terms. The corresponding graph problem contains dangling tree branches, which we remove. Because their topology is nonfrustrating, it is trivial to find the optimal value for the variables on those branches once a solution for the rest of the problem is found: one simply sets the value of a variable  $z_j = \pm 1$  such that  $\mathbf{w}_{ij} z_i z_j$  is locally minimized. The remaining problem has  $N = 201$  variables and  $n = 254$  terms; however, one finds that it actually consists of two disconnected and independent subgraphs. The first contains  $N = 180$  nodes and  $n = 226$  edges; the second contains  $N = 21$  nodes and  $n = 28$  edges and can, therefore, be solved via brute-force enumeration. In the following, we solve the larger of the two subproblems and trivially reconstruct a solution  $\mathbf{z}$  to the original problem of Eq. (6).

The optimal contribution of the dangling tree branches to the global objective value is approximately 50 778.093 880, the optimal contribution of the  $N = 21$  problem to the global objective value is approximately 3728.413 221, and we estimate that the global optimal solution has an objective value  $C_{\text{MPES}}(\mathbf{z}_{\text{opt}}) \simeq 61\,606.078\,827$ .

#### Quantum preconditioning

The  $N = 180$  graph problem is sufficiently sparse that we can employ the light-cone technique introduced in the context of the random 3-regular graph maximum-cut problems in Sec. III A 2. The graph has an average and maximum degree of approximately 2.51 and 7, respectively. At  $p = 1$  and  $p = 2$ , the light-cone-induced subgraphs for building the correlation matrix have at most 14 and 27 vertices, respectively. Both numbers are within the ability of classical state-vector emulators, which we use for preconditioning.

Unlike the previous graph problems investigated, near-optimal QAOA angles are not tabulated for this problem. As such, we perform a variational search for the QAOA angles extremizing the expectation value of the objective function  $\langle \hat{C} \rangle_p$ . We use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [65–68] algorithm, starting from random initial angles a few hundred times and consider the best output for computing the correlation matrix. We report the QAOA angles in Appendix A. The original and preconditioned problems have  $n = 226$ ,  $n = 600$  ( $p = 1$ ), and  $n = 1625$  ( $p = 2$ ) terms, respectively.

### Performance

We display the average approximation as a function of the number of iterations and run time for SA and the BM solver in Fig. 4. We compare the performance of these solvers on the original and preconditioned problems. We evaluate the average approximation ratio by sampling  $10^4$  solutions from the classical solvers for each number of iterations. We observe that for this problem in the range considered, the BM algorithm is more efficient than SA as it can reach a much higher approximation ratio for the same fixed run time.

For a given number of iterations, we find regions where the solvers on the preconditioned problems yield a higher approximation ratio than with the original one. This defines windows for an advantage. At a small number of iterations, we observe that the window is enlarged from  $p = 1$  to  $p = 2$ . Such a gain from increasing the amount of preconditioning was also observed for the random 3-regular graph maximum-cut problems. When substituting the number of iterations for the run time on a 64 GB MacBook Pro with an Apple M1 Max chip, there is still an advantage window, despite the quantum-preconditioned problems having more terms than the original one (see Appendix G). Finally, we note that  $p = 1, 2$  preconditioned problems can access approximation ratios  $\alpha \gtrsim 99.99\%$ , and that they do not share the same optimal solutions as the original problem, as exemplified by the saturation in approximation ratio observed in Fig. 4(d). In addition to classical state-vector emulations, we precondition the problem using an experimental implementation of the QAOA at  $p = 1$  and solve the

preconditioned problem. Results of this experiment are shown in Appendix D.

The existence of an advantage window for a real-world problem highlights the potential wide applicability of the proposed quantum-preconditioning method.

### QUANTUM-PRECONDITIONING BUDGET AND PROSPECTS FOR A QUANTUM ADVANTAGE

We have shown that quantum-preconditioning problems can benefit state-of-the-art classical solvers. While this is a nontrivial result on its own, a practical advantage can only be claimed if one also accounts for the preconditioning time itself—see, e.g., Fig. 1(b) defining a preconditioning time budget for random 3-regular graph maximum-cut problems via SA.

#### Quantum-inspired advantage for random 3-regular graph maximum-cut problems

We refer to a quantum-inspired advantage as obtained via a purely classical calculation or emulation of a quantum circuit, as opposed to execution on an actual quantum computer. This can be a first step toward genuine quantum utility if the proposed method has room to perform even better at a larger scale (more qubits, larger depth) beyond classical computational capabilities.

We focus on the random 3-regular graph maximum-cut problems, which present the largest advantage window compared to the SK spin glasses or the grid energy problem. We consider SA, which is more efficient than the BM

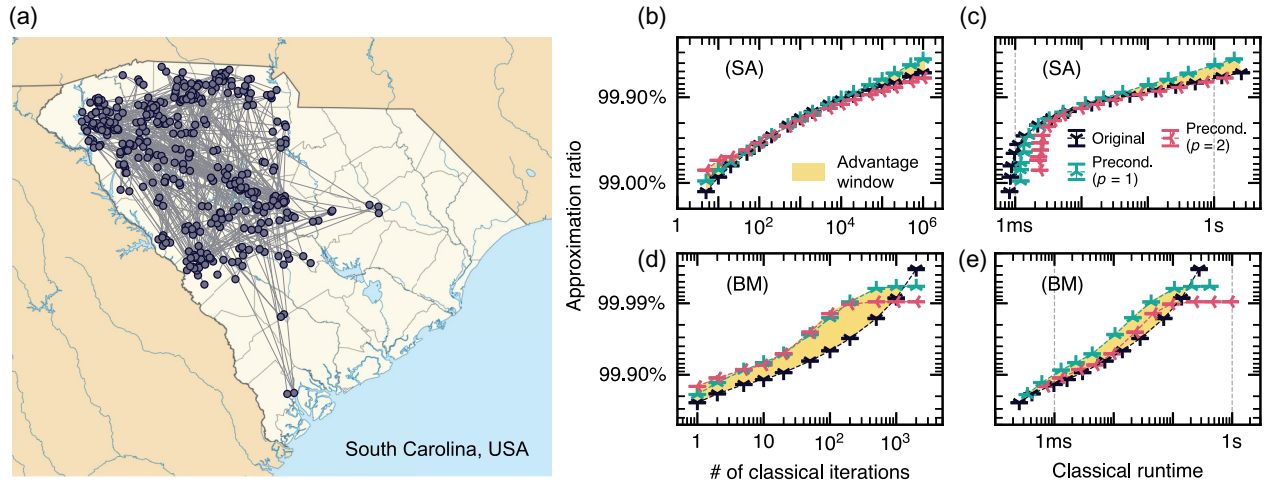


FIG. 4. Performance of quantum preconditioning for computing the maximum power exchange section of an energy-grid optimization problem. (a) Representation of the energy-grid optimization problem considered from the state of South Carolina in the United States of America. Vertices ( $N = 180$ ) and edges ( $n = 226$ ) correspond to buses and lines, respectively. (b) Average approximation ratio as a function of the number of iterations via SA. (c) Average approximation ratio as a function of run time via SA. (d) Average approximation ratio as a function of the number of iterations via the BM solver. (e) Average approximation ratio as a function of run time via the BM solver. Run times are based on a 64 GB MacBook Pro with an Apple M1 Max chip. Each data point is averaged over  $10^4$  samples. Error bars indicate the standard error of the mean.

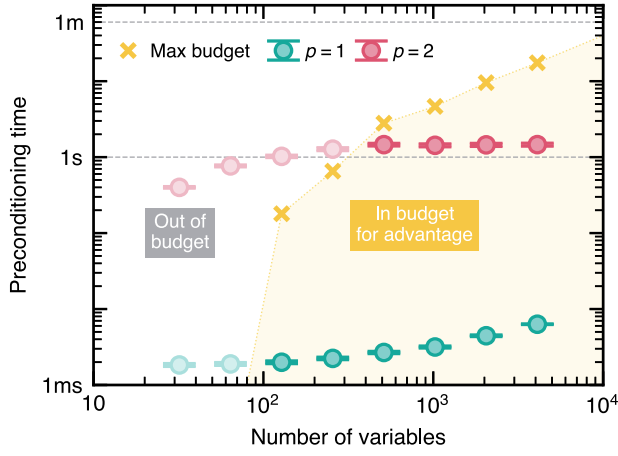


FIG. 5. Average time for preconditioning  $N$ -variable random 3-regular graph maximum-cut problems using classical state-vector emulations of the QAOA using the light-cone technique at  $p = 1$  and  $p = 2$ . The maximum preconditioning time budget for an advantage via SA for an approximation ratio of  $\alpha = 99.9\%$  is computed from Fig. 1. Run times correspond to a 64 GB MacBook Pro with an Apple M1 Max chip. Error bars indicate the standard error of the mean.

solver in the  $\alpha = 99.9\%$  regime (this value has been chosen arbitrarily, and others may provide a larger advantage window). The preconditioning time budget is defined in Fig. 1(b) as the difference in SA run time between the original and preconditioned problems to reach  $\alpha = 99.9\%$ . We report the budget between original and preconditioned at  $p = 2$  in Fig. 5. Because there is a difference of about one order of magnitude in runtime between the original and preconditioned problems, the budget is mostly dominated by the run time of the solver on the original problem.

We perform the preconditioning via state-vector emulations of the QAOA circuits on the relevant light-cone-induced subgraphs according to Sec. III A 2. Determining the light-cone-induced subgraphs of an  $N$ -variable problem scales as  $O(N)$ . The QAOA implementation benefits from the global  $\mathbb{Z}_2$  sign-flip symmetry of the problem, which reduces the overall emulation run time by a theoretical factor 2. Moreover, we use the fact that in the limit  $p \ll \ln N$ , most light-cone-induced subgraphs are trees (see Appendix G). Trees are  $N$ -variable graphs for which the number of edges is  $n = N - 1$ . As such, we cache the expectation value  $\langle \hat{Z}_i \hat{Z}_j \rangle$  of the trees such that the QAOA only needs to be emulated once on such subgraphs. This results in only emulating  $O(\exp p)$  QAOA circuits, independent of  $N$  (see Appendix G). The memory usage and algorithmic complexity of a QAOA emulation via the light-cone technique scale as  $O(\exp \exp p)$  and are classically manageable at  $p = 1$  and  $p = 2$ .

In Fig. 5, we show the time it takes to perform the quantum preconditioning at  $p = 1$  and  $p = 2$  via emulations. For large enough problems, the preconditioning run

time fits within the budget. Therefore, quantum preconditioning can deliver a quantum-inspired advantage over SA through classical emulations. This advantage extends for larger problem sizes to the BM solver, which has a much larger budget (see Fig. 2); however, the BM solver is not competitive against SA.

The change of slope observed at large  $N$  for  $p = 1$  is because the run time becomes dominated by the  $O(N)$  task of determining the light-cone-induced subgraphs rather than the QAOA emulations. We anticipate an analogous asymptotic behavior for  $p = 2$ .

### Prospects for hardware-based quantum utility for random 3-regular graph maximum-cut problems

We have shown in Fig. 5 that quantum preconditioning can provide a quantum-inspired advantage on random 3-regular graph maximum-cut problems via classical emulations at  $p = 1$  and  $p = 2$ . We have also shown in Fig. 1(b) that increasing the depth of the QAOA from  $p = 1$  to  $p = 2$  can accelerate the convergence. Hence, a question is whether  $p = 3$  could improve the convergence further; however, light-cone-induced subgraphs at  $p = 3$  have up to 43 qubits and, therefore, cannot be emulated with exact classical methods in a reasonable amount of time, calling for a new simulation platform, such as an actual quantum computer.

### Sampling from the wave function

When moving to quantum hardware for executing quantum circuits, one does not have access to the full quantum state  $|\Psi\rangle_p$  of Eq. (3) for computing expectation values but must sample bit strings  $\{\mathbf{z}^{(k)}\}_{k=1,\dots,K}$  to estimate them. After executing the circuit and measuring the qubits  $K$  times, one has collected  $K$  bit strings according to the probability  $|\langle \mathbf{z}^{(k)} | \Psi \rangle_p|^2$ . Two-point correlations are then estimated as

$$\langle \hat{Z}_i \hat{Z}_j \rangle_p \simeq \frac{1}{K} \sum_{k=1}^K z_i^{(k)} z_j^{(k)}. \quad (8)$$

The total run time for estimating the expectation value is proportional to  $K$ . For the classical state-vector emulations, we had access to the full quantum state  $|\Psi\rangle_p$  and did not need to sample bit strings, which corresponded to the  $K \rightarrow +\infty$  limit.

*a. Performance against sampling.* A first question is whether quantum preconditioning is robust to finite sampling. In Figs. 6(a)–6(d), we show the accuracy of SA and the BM solver on preconditioned  $N$ -variable random 3-regular graph maximum-cut problems where preconditioning was performed by sampling  $K$  bit strings to estimate correlations via Eq. (8). The correlation matrix

was computed via the classical emulation of light-cone-induced QAOA subgraphs, as was done previously in Sec. III A 2. We consider preconditioning at  $p = 1$  and  $p = 2$  and the performance of SA and BM with  $n_{\text{iter}} = 100$  and  $n_{\text{iter}} = 50$  iterations, respectively.

We do not observe a clear  $N$  or  $p$  dependence of the convergence of the average approximation ratio toward its infinite-sampling ( $K \rightarrow +\infty$ ) value; however, we can rule out a logarithmic convergence with  $K$  (i.e., an exponential number of samples needed to achieve a desired accuracy). Moreover, we note that a convergence to a relative error in average approximation ratio of 0.01% requires at most  $K \simeq 10^4$  bit strings, which is a typical number used for evaluating expectation values on quantum computers. The saturation observed at  $K \sim 10^4$  in Figs. 6(b) and 6(c) is due to finite statistics in the number of random problem instances considered and for estimating the approximation ratio of the individual problem instances. We expect all the curves to get asymptotically to zero as  $K \rightarrow +\infty$ . Therefore, finite sampling does not seem to be a barrier to the performance of quantum preconditioning.

*b. Sampling time.* The next question is the time needed for preconditioning via sampling. We consider the superconducting quantum computer Ankaa-3 developed by Rigetti Computing and accessed via cloud services. Superconducting qubits provide the fastest sampling rate to date thanks to their logical quantum operations in the tens-of-nanoseconds range, several orders of magnitude faster than atom- or ion-based quantum computers. Given that optimization solvers operate on an accuracy versus run time window, the speed of preconditioning is paramount to deliver any potential advantage.

In Fig. 6(e), we report the average sampling time for the  $p$ -layer QAOA on a typical  $N$ -vertex subgraph induced by the light cone on 3-regular graphs. An arbitrary QAOA circuit on  $N$  qubits with  $p$  layers executed on a linear chain topology of qubits through a naïve swap network [69] requires a total of  $3Np$  compact layers of two-qubit  $i\text{SWAP} = \exp[i\pi(\hat{X} \otimes \hat{X} + \hat{Y} \otimes \hat{Y})/4]$  gates and  $2(2N + 1)p + 1$  compact layers of one-qubit  $\text{RX}(\phi) = \exp(-i\hat{X}\phi/2)$  gates where  $\phi = \pm\pi/2$  [44].  $\hat{X}$ ,  $\hat{Y}$ , and  $\hat{Z}$  are Pauli operators. All the gates within a layer can be applied in parallel on independent qubits. We also make use of one-qubit  $\text{RZ}(\theta \in \mathbb{R}) = \exp(-i\hat{Z}\theta/2)$  gates, which are virtual and realized by a change of frame of the qubits tracked by the control system. The collection time of  $K$  samples is modeled by

$$t(K, N, p) = t_{\text{ovhead}}(K, N, p) + K t_{\text{circ}}(N, p), \quad (9)$$

where  $t_{\text{ovhead}}(K, N, p)$  indicates a classical overhead encompassing, e.g., network latency, job management, and

control systems. The other component,

$$t_{\text{circ}}(N, p) = 3Np t_{2Q} + [2(2N + 1)p + 1] t_{1Q} + t_{\text{mes}} + t_{\text{res}}, \quad (10)$$

relates to the circuit execution. It defines a minimum physical limit for sampling one quantum circuit. We have  $t_{2Q} \simeq 80$  ns as the typical duration of a two-qubit gate,  $t_{1Q} \simeq 40$  ns as the typical duration of a one-qubit gate,  $t_{\text{mes}} = 1$   $\mu$ s as the typical duration of measuring a qubit, and  $t_{\text{res}} = 200$   $\mu$ s as the typical duration for passively resetting a qubit prior to a subsequent circuit execution. These numbers provide a baseline, but they can typically be calibrated or fine-tuned.

For  $N = 4096$ , the average size of light-cone-induced subgraphs is  $N = 5.67(2)$  at  $p = 1$ ,  $N = 16.613(5)$  at  $p = 2$ , and  $N = 39.772(5)$  at  $p = 3$ . Hence, in this regime, the average circuit execution time  $t_{\text{circ}}$  is dominated by the passive reset time of the qubits. This is highlighted in Fig. 6(e), where for a large number of samples,  $t \simeq K t_{\text{res}}$ . For a small number of samples, the sampling time is instead dominated by the classical overhead. For a number of samples  $K \gtrsim 10^3$ , a first step for improving the sampling time is, therefore, to move to an active reset strategy for the qubits, which would yield  $t_{\text{mes}} + t_{\text{res}} \simeq 6$   $\mu$ s [70], a 30-fold reduction over passive reset.

At the moment, for  $K = 10^4$ , the total sampling time for a single light-cone-induced circuit takes about 2–3 s, as shown in Fig. 6(e). This is not competitive against classical state-vector emulations at  $p = 1$  and  $p = 2$ , which can precondition *all* light-cone-induced circuits in less time (see Fig. 5). Thus, improvements will be needed to reduce the sampling time and make the approach competitive and capable of delivering an advantage when accounting for the preconditioning time budget. Whether this will prove a bottleneck remains to be established: indeed, we observe in Fig. 5 that the budget increases with the number of variables  $N$  while the preconditioning is independent of  $N$  for  $p \ll \ln N$  thanks to caching of light-cone-induced subcircuits (see Appendix G)—strictly speaking, preconditioning has a classical  $O(N)$  dependence in finding the light-cone-induced subgraphs. If confirmed for much larger  $N$ , this scaling difference between the preconditioning budget and time could suggest a crossing value  $N^*$  for a quantum simulation method (hardware or classically emulated) to deliver an advantage, independently of its intrinsic run time.

### Performance of a noisy quantum computer

In the absence of quantum error correction, inherent hardware noise is another consideration for contemporary devices. To get a sense of the effect of noise through a back-of-the-envelope calculation, we assume that two-qubit gates are the dominant source of errors with an average fidelity  $f$ . An arbitrary QAOA circuit on  $N$  qubits with  $p$  layers executed on a linear chain topology of qubits

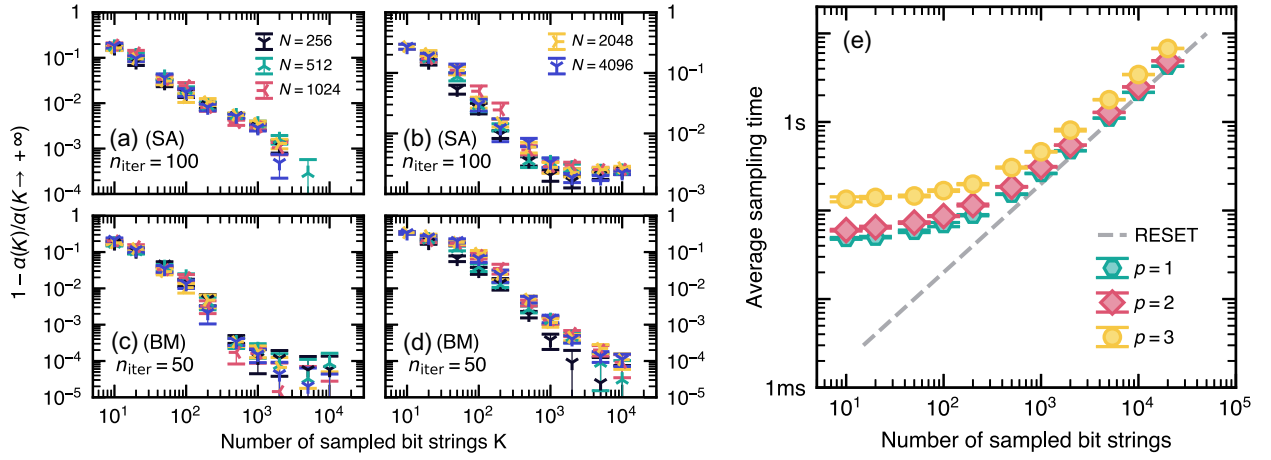


FIG. 6. (a)–(d) Random  $N$ -variable 3-regular graph maximum-cut problems are considered, showing the convergence of the average approximation ratio  $\alpha$  toward its infinite-sampling value, obtained via quantum preconditioning with a finite number of sampled bit strings  $K$ . Preconditioning is performed via the simulation of light-cone-induced  $p$ -layer QAOA circuits. (a) SA at  $p = 1$  with  $n_{\text{iter}} = 100$  iterations. (b) SA at  $p = 2$  with  $n_{\text{iter}} = 100$  iterations. (c) BM solver at  $p = 1$  with  $n_{\text{iter}} = 50$  iterations. (d) BM solver at  $p = 2$  with  $n_{\text{iter}} = 50$  iterations. Each data point is averaged over 200 randomly generated problem instances. (e) Average sampling time on Rigetti Ankaa-3 superconducting quantum chip for a light-cone-induced QAOA circuit for various numbers of QAOA layers  $p$  for random 3-regular graph maximum-cut problems. RESET corresponds to a passive reset time of 200  $\mu\text{s}$  of the qubits between two circuit executions. Each data point is averaged over 50 randomly selected light-cone-induced QAOA circuits. Error bars indicate the standard error of the mean.

through a naïve swap network requires  $n_{2Q} = 3pN^2/2$  two-qubit gates [69]. We have assumed three hardware-native two-qubit gates (such as  $i\text{SWAP}$  or  $\text{CZ}$ ) for compiling each two-qubit gate of the QAOA circuit. We suppose a fully depolarizing noise model parameterized by  $F \in [0, 1]$ , which results in the mixed state  $\hat{\rho}_{p,F} = F|\Psi\rangle\langle\Psi|_p + (1-F)\hat{I}/2^N$ , where  $\hat{I}$  is the identity matrix. Two-point correlations of interest read

$$\langle \hat{Z}_i \hat{Z}_j \rangle_{p,F} = \text{tr}(\hat{Z}_i \hat{Z}_j \hat{\rho}_{p,F}) = F \langle \hat{Z}_i \hat{Z}_j \rangle_{p,F=1}, \quad (11)$$

where  $F = f^{n_{2Q}}$  can be interpreted as the global quantum circuit fidelity, with individual errors happening independently and at random at the two-qubit gate level. To distinguish an  $O(1)$  expectation value with  $K$  sampled bit strings, one needs  $K \sim F^{-2}$ , or equivalently  $K \sim f^{-3pN^2}$ , according to the central limit theorem. By plugging in relevant numbers for  $f$ ,  $N$ , and  $p$ , one realizes that finite sampling by itself is likely not a bottleneck, but sampling for reliably estimating noisy and exponentially small expectation values might be. Therefore, hardware performance through  $f$  sets the expectation for what circuit can be successfully executed. We verify that such a depolarizing noise model is able to phenomenologically fit experimental data in Appendix D, as similarly observed in Ref. [54].

### Noisy quantum computer versus approximate classical tensor networks

Given the fidelity requirements to run the QAOA on large problem instances in the absence of quantum error correction, one will likely need to rely on a decomposition technique such as the light-cone one introduced in Sec. III A 2. Light-cone-induced subgraphs at  $p = 3$  require at most 43 qubits, far beyond what state-vector emulators can handle in a reasonable amount of time. Other candidate simulators for simulating such a quantum circuit include an actual quantum computer or classical approximate tensor-network methods, such as matrix product states [71,72]. In the absence of quantum error correction, quantum computers are imperfect and thus provide only an approximate simulation in some respects. A first question is whether quantum preconditioning via an approximately classically emulated preconditioner can deliver an advantage (see Appendix I).

Assuming an imperfect hardware output, a second question is what level of approximation in a tensor-network simulation matches the accuracy of the imperfect quantum computer? The level of approximation in tensor networks, encoded in a parameter  $\chi \in [1, 2^{N/2}]$ , controls their run time: the larger  $\chi$ , the less approximation is performed, and the longer the simulation takes. For instance, the algorithmic complexity of matrix product states goes as  $O(\chi^3)$ . Then, one can answer the final question: for a matched output accuracy, which of the approximate tensor-network simulator or quantum computer is faster? This will provide a basis for a hardware-based quantum utility [73–76].

### Beyond random 3-regular graph maximum-cut problems and challenges for a quantum optimization advantage

Many challenges faced by the proposed quantum-preconditioning method to deliver quantum utility are common to quantum optimization and are active areas of research. We discuss some of the main ones in the following.

#### *Compilation and efficient swap network for sparse graphs*

Superconducting quantum hardware has a given topology, with fixed connections between the qubits. This requires the use of two-qubit SWAP gates for achieving quantum logical operations between arbitrary non-neighboring qubits. Compiling QAOA circuits for sparse graph problems using a default swap network (i.e., on a linear arrangement of qubits [69,77]) is far from efficient in terms of circuit depth and number of two-qubit gates. This sets fidelity requirements much higher than what is actually needed under the right compilation strategy. An efficient swap network [78,79] for light-cone-induced subgraphs on random 3-regular graph maximum-cut problems can reduce the number of *i*SWAP gates by about 52%, 73%, and, 85% at  $p = 1$ ,  $p = 2$ , and  $p = 3$ , respectively [44]. Such smaller circuits would lead to much higher overall output fidelities. Therefore, the sampling requirements for reliably estimating expectation values would also be much lower—in addition to shorter circuits, reducing the execution time—thus greatly speeding up the preconditioning.

For instance, a light-cone-induced subgraph has at most  $N = 43$  vertices at  $p = 3$  for a 3-regular graph. Following Sec. IV B 2, this translates into about  $n_{2Q} \sim O(10^4)$  via a default swap network or  $n_{2Q} \sim O(10^3)$  with an efficient compilation [44,78,79]. Assuming an average two-qubit gate fidelity  $f = 99.5\%$ , the default strategy yields an overall fidelity of  $F \simeq 10^{-22}$ , and  $F \simeq 10^{-2}$  for the efficient one. This highlights the importance of the compilation step on noisy quantum devices.

However, finding an efficient swap network is a difficult task relating to graph isomorphism: solving graph isomorphism using a Boolean satisfiability (SAT) formulation is itself an NP-complete problem and scales exponentially with the size of the graph [44,78,79]. This calls for the development of computationally cheap and scalable approximate methods for encoding graphs onto the hardware-native topology of a quantum computer. In addition, while the focus is on the QAOA, it is possible that other variational quantum circuits could perform better.

#### *Decomposition technique*

We used the light-cone decomposition technique for classical emulations of QAOA circuits for sparse large- $N$  problems, such as random 3-regular graphs and the grid

energy problem (Sec. III A 2). At shallow circuit depth, this trades a single  $N$ -qubit QAOA circuit for  $O(N)$  QAOA circuits with  $O(\exp p)$  qubits each. While this can be advantageous for small values of  $p$ , it requires the execution of many subcircuits. For instance, the preconditioning time via state-vector emulations of the grid energy problem (Fig. 4) took 0.002 70(2) s at  $p = 1$  and 32.9(2) s at  $p = 2$  on a 64 GB MacBook Pro with an Apple M1 Max chip. The  $p = 1$  and  $p = 2$  preconditioning times are within and out of budget, respectively; however, they do not account for the variational search of optimal QAOA angles. A long preconditioning time could be mitigated by parallelizing the execution of subcircuits on quantum hardware if more qubits than necessary are available (e.g., about twenty 50-qubit circuits can be executed at a time on a 1000-qubit device).

Moreover, the light-cone decomposition technique is only applicable to sparse graphs. Dense graphs, such as the Sherrington-Kirkpatrick spin glasses considered, cannot benefit from it. Because running the QAOA on large, dense graph problems is out of scope for noisy devices due to the large number of qubits and gate requirements, this raises the question of whether other decomposition techniques could be envisioned in the context of quantum preconditioning [80–90].

#### *Variational optimization*

Another difficulty is the variational optimization of the QAOA or other quantum circuits used for preconditioning. Finding optimal angles is a known challenge for variational optimization due to Barren plateaus [91,92], whereas preconditioning has to happen fast to be within an advantage window.

For the random 3-regular graph maximum-cut problems and Sherrington-Kirkpatrick spin glasses, we employed tabulated near-optimal parameters [27,29,51], thus skipping altogether the variational optimization; however, the uniqueness of the grid energy optimization problem required a variational optimization of the circuit parameters. We note that classes of optimization problems have been found to share near-optimal angles and that such a strategy of transferring parameters from one problem to another could be leveraged [27–29,31,51,56,93–98]. Other strategies have also been advanced based on interpolation from lower  $p$  values and quantum annealing scheduling [94,99,100]; however, this raises the question of the performance of the quantum preconditioner based on nonoptimal angles.

### CONCLUSION

We have introduced a quantum-preconditioning method for quadratic unconstrained binary optimization problems. Here, preconditioning refers to the task of transforming the input problem into a more suitable form for a solver

by improving its convergence. The proposed quantum preconditioner is based the QAOA [25,26]. We demonstrated that state-of-the-art classical solvers such as simulated annealing and the Burer-Monteiro method benefited from working with the preconditioned input instead of the original one on a range of problems, including Sherrington-Kirkpatrick spin glasses, random 3-regular graph maximum cut, and a real-world grid energy problem. The amount of preconditioning is controlled by the quantum circuit depth and renders the problem trivial for a classical solver in the infinite-depth limit. We showed that an increased circuit depth for shallow circuits can improve the convergence. Because classical optimization heuristics operate on an accuracy versus run time basis, any additional time taken for preconditioning a problem should be given to the purely classical solution. Accounting for this, we showed that quantum preconditioning translates into a practical quantum-inspired advantage for random 3-regular graph maximum-cut problems through classical state-vector emulations of the quantum circuits. Finally, we discussed the challenges and prospects for delivering a hardware-based quantum utility in optimization via quantum preconditioning.

The results call for developing a better understanding of what makes a good quantum preconditioner. We found that quantum preconditioning via the QAOA reduces the problem's frustration, making local moves more efficient in approaching the global optimum. In this work, the use of the QAOA as a preconditioner was heuristic, and it is possible that other quantum circuits might be better suited.

It would be interesting to investigate quantum preconditioning in the context of constrained problems—here, we only considered unconstrained problems. Hard constraints are a part of many optimization problems; however, there is a lack of methods for efficiently handling them on contemporary, noisy, quantum computers. Constraints can be encoded in quantum optimization in several ways, and each has its unique challenges. For example, constraints are often encoded as penalty terms [41], but bit strings may violate the penalties. In some cases, special mixers are used to keep the state within the constraint-satisfying subspace of the Hilbert space [101,102], or the problem is encoded such that the Hilbert space contains only constraint-satisfying solutions [103–108]. In the former (latter) case, the mixer (phase separator) requires multiqubit gates. In our approach, one could quantum precondition a problem in any formulation of the constraints while letting a classical solver handle the hard constraints.

Improving the fidelity of noisy hardware remains the backbone for quantum computers to potentially deliver a quantum advantage in the absence of quantum error correction. While it is common to leverage error-mitigation techniques to enhance noisy expectation values [109–111], one should be mindful of the added time overhead of some of these techniques in a fair benchmark against

classical optimization solvers. Error-mitigation methods such as randomized compilation [112] or readout-error mitigation [70,113–123] may provide a benefit with little overhead. Furthermore, exploring quantum preconditioning or the QAOA [124] within the framework of quantum error correction would also be interesting, especially with the advent of fault-tolerant demonstrations [125–127]. In particular, quantum error correction induces a run-time overhead, which then raises the question of whether there might be an opportunity for a quantum advantage for quantum preconditioning in a fault-tolerant era.

## ACKNOWLEDGMENTS

We gratefully acknowledge B. Evert, M.J. Hodson, S. Jeffrey, F.B. Maciejewski, and D. Venturelli for discussions, input, and collaborations on related works. This work is supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Superconducting Quantum Materials and Systems Center (SQMS) under Contract No. DEAC02-07CH11359. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 using NERSC awards ASCR-ERCAP0028951 and ASCR-ERCAP0031818. The experiments were performed through Rigetti Computing Inc.'s Quantum Cloud Services QCS<sup>TM</sup> on the Ankaa<sup>TM</sup>-3 superconducting quantum processor developed, fabricated, and operated by Rigetti Computing Inc.

M.D. conceived and led the project with support from B.S. M.D. and T.O. performed simulations, data collection, and data analyses. M.D. wrote the manuscript with input from B.S. All coauthors contributed to the discussions leading to the completion of this project.

M.D. and B.S. are, have been, or may in the future be participants in incentive stock plans at Rigetti Computing Inc. M.D. and B.S. are inventors on two pending patent applications related to this work (No. 63/631,643 and No. PCT/US2024/033445). The other authors declare that they have no competing interests.

## DATA AVAILABILITY

The data that support the findings of this article are openly available [128].

## APPENDIX A: PROBLEM CLASSES AND INSTANCES

### Random 3-regular graph maximum-cut problems

Random 3-regular graph maximum-cut problems are a standard testbed for investigating the performance of

quantum optimization [25,29,44,48,51,54,129–137]. They have  $N$  vertices, each connected by an edge to three other vertices at random. The edges carry unit weight. These graphs are sparse, with the adjacency matrix  $W$  containing only three nonzero entries per row or column. Despite their apparently simple topology, it is NP-hard to devise an algorithm to find the maximum cut that can guarantee an approximation ratio of at least 99.7% [138], with the current best-known classical algorithms yielding  $\alpha \simeq 93.3\%$  [139]. We generated random 3-regular graphs using the PYTHON package NETWORKX [140–142].

The heuristics considered in this work have the potential to deliver a solution with a higher approximation ratio, but they cannot guarantee it *a priori*. Random 3-regular graph maximum-cut problems have been independently investigated on a range of solvers such as BM [143], extremal optimization [144], quantum annealing, and SA [134,145,146]—interestingly, it was shown that quantum annealing is less efficient than SA [145,146]. Classical quantum-inspired algorithms based on simulated coherent Ising machines [147] and simulated bifurcation [148,149] have also been considered [150]. Random 3-regular graph maximum-cut problems have also served as a testbed for combinatorial optimization solvers based on graph neural networks [151,152], although their performance is far from that of state-of-the-art methods [153–155]. We note that a one-to-one benchmark with the existing literature can be difficult as it has focused on the time-to-solution (TTS) metric [156], which relates to the probability of finding the optimal solution. In this work, we measure the performance through the average approximation ratio instead, as quantum preconditioning cannot guarantee the same optimal solution between the original and preconditioned problems, except in the asymptotic  $p \rightarrow +\infty$  limit. Furthermore, it might be interesting to investigate quantum preconditioning in the context of state-of-the-art classical solvers other than the BM solver and SA considered in this work, such as extremal optimization [144] and quantum-inspired algorithms [147–150,156].

### Sherrington-Kirkpatrick spin glasses

SK models [42] correspond to complete graphs, where the edge weights are independent and identically distributed random variables drawn from a normal distribution of zero mean and unit width  $W_{ij} \sim \mathcal{N}(0, 1)$  with  $W_{ij} = W_{ji}$ . As a result, the adjacency matrix  $W$  encoding the problem belongs to the Gaussian orthogonal ensemble. Such graphs present two stark differences with 3-regular graphs: first, they are as dense as possible because of the all-to-all connectivity between vertices; second, they are weighted graphs.

It is NP-hard to find the optimal solution to an SK spin glass; however, an approximate solution arbitrarily close to the optimal one can be obtained in polynomial

time through a recently introduced approximate message-passing algorithm [157]: a solution to an  $N$ -variable problem with an approximation ratio of at least  $\alpha = (1 - \varepsilon)$  for  $\varepsilon > 0$  can be found in  $O(P_\varepsilon N^2)$  with  $P_\varepsilon$  a polynomial in  $\varepsilon^{-1}$ . Nevertheless, SK problems remain highly relevant for developing an understanding thanks to their long-standing connection with statistical physics and benchmarking optimization solvers [27,48,52,53,88,89,158].

### Power-grid optimization problem

Quantum preconditioning on a real-world power-grid optimization problem seeking to compute the MPES of an energy network [56–59] was investigated in Sec. III C. We considered a realistic 500-bus power system dataset “ACTIVSg500” mimicking the energy grid of the state of South Carolina in the United States of America [43]. The dataset was designed as part of ARPA-E’s GRID DATA program to be similar to the actual electric grid while containing no confidential information about critical energy infrastructure [60–64]. In Table I, we present information regarding the QAOA settings that were used for preconditioning the  $N = 180$ -variable problem. The reported QAOA angles correspond to minimizing the objective function  $\tilde{C}_{\text{MPES}}(\mathbf{z}) = \frac{1}{2} \sum_{i,j=1}^N W_{ij} z_i z_j$ . The reported objective value is that of the underlying maximum-cut problem of Eq. (6).

## APPENDIX B: QUANTUM PRECONDITIONING EASES PROBLEM HARDNESS

### Large-depth quantum preconditioning at small scale

The focus of the main text was on large problems at shallow circuit depth  $p$ . We now turn our attention to quantum preconditioning of small-scale ( $N = 16$ ) random 3-regular graph maximum-cut problems and SK spin glasses for large values of  $p$ . We access large  $p$  values via classical emulations using tabulated near-optimal QAOA angles up to  $p \leq 11$  for random 3-regular graph maximum-cut problems [29,51] and  $p \leq 17$  for SK spin glasses [27,28]. We consider simulated annealing as a solver.

In Fig. 7, we plot the average approximation ratio as a function of the number of iterations in simulated annealing for a fixed value of  $p$  and as a function of  $p$  for a fixed number of iterations. We observe that increasing  $p$  in quantum preconditioning makes simulated annealing converge to an asymptotic approximation ratio value faster, i.e., in fewer iterations.

For a large number of iterations, we find that the approximation saturates to a value  $\alpha < 1$ , as anticipated, because the original and preconditioned problems do not share the optimal solution except in the asymptotic  $p \rightarrow +\infty$  limit. Nevertheless, we find that increasing  $p$  increases the asymptotic approximation ratio value for SK spin glasses.

TABLE I. Information regarding the MPES problem considered. The QAOA angles were found using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [65–68] algorithm. The number of nonzero terms correspond to the number of nonzero entries in the correlation matrix of Eq. (2).

| Number of QAOA layers | Optimized QAOA angle                                                                                                         | Objective value     | Number of nonzero terms |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------|---------------------|-------------------------|
| Original              | —                                                                                                                            | —                   | 226                     |
| $p = 1$               | $\gamma_1 \simeq -0.008\,418$<br>$\beta_1 \simeq 2.757\,259$                                                                 | $\simeq 59\,721.27$ | 600                     |
| $p = 2$               | $\gamma_1 \simeq -0.006\,541$<br>$\beta_1 \simeq 1.074\,866$<br>$\gamma_2 \simeq -0.012\,537$<br>$\beta_2 \simeq 1.307\,313$ | $\simeq 60\,202.88$ | 1625                    |

It is less clear for the random 3-regular graph maximum-cut problems [Fig. 7(a)]. We note that the random 3-regular graph maximum-cut problems are more likely to have degenerate solutions than the SK spin glasses beyond the global  $\mathbb{Z}_2$  spin flip symmetry, due to the nature of their weights  $W_{ij}$ : unity versus normally distributed. As such, the correlation matrix  $Z^{(p)}$  may not tend to the desired asymptotic  $p \rightarrow +\infty$  limit discussed in Sec. II. Another preconditioner may be more suited at large  $p$ .

### Solver-agnostic quantities

We consider solver-agnostic quantities related to the problem’s hardness for random 3-regular graph maximum-cut problems and SK spin glasses. First, we compute the frustration index [159–161]

$$f = \frac{1}{2} + \frac{\sum_{i,j=1}^N W_{ij} z_{\text{opt},i} z_{\text{opt},j}}{2 \sum_{i,j=1}^N |W_{ij}|}; \quad W_{ij} \leftrightarrow Z_{ij}^{(p)}, \quad (1)$$

where  $f \in [0, 1]$  measures how strongly the optimal solution  $z_{\text{opt}}$  is frustrated with respect to an input problem  $W$  or  $Z^{(p)}$ . Second, we compute the normalized gap

$$\Delta = (\sigma_2 - \sigma_1) / (\sigma_N - \sigma_1), \quad (2)$$

where  $\sigma_{i=1,2,\dots,N} \geq 0$  are the singular values sorted in ascending order of an input problem  $l + W$  or  $l + Z^{(p)}$ . The normalized gap  $\Delta \in [0, 1]$  indicates how accurately the matrix can be approximated by a rank-1 matrix. Indeed, in the asymptotic  $p \rightarrow +\infty$  limit in the presence of a doubly degenerate optimal solution  $\pm z_{\text{opt}}$ , one has  $l + Z^{(\infty)} = z_{\text{opt}} z_{\text{opt}}^T$ , which is a rank-1 matrix, where in that case, the normalized gap is  $\Delta = 1$ .

In Fig. 8, we plot these quantities on both the original problem  $W$  and the preconditioned ones  $Z^{(p)}$ . We find that for the largest  $p$  values, the frustration index and gap asymptotically reach their  $p \rightarrow +\infty$  value, i.e.,  $f \rightarrow 0$  and  $\Delta \rightarrow 1$ . This highlights the potential of quantum preconditioning to ease the hardness of challenging optimization problems. At a shallow circuit depth  $p$ , such as investigated in the rest of this work, the behavior of  $f$  and  $\Delta$  is much less clear. In particular, for random 3-regular graph maximum-cut problems, we observe an increase in frustration with  $p$  at small  $p$  in Fig. 8(a). This is counterintuitive, as the quantum-preconditioned problem is actually easier to solve for algorithms such as simulated annealing (Fig. 1) and Burer-Monteiro (Fig. 2); however, we observe a change of slope for  $p = 2$ , which indicates a drop in frustration for  $p \geq 3$ , which may further facilitate convergence at larger circuit depths for solvers relying on quantum preconditioning. Testing this hypothesis is out of scope for

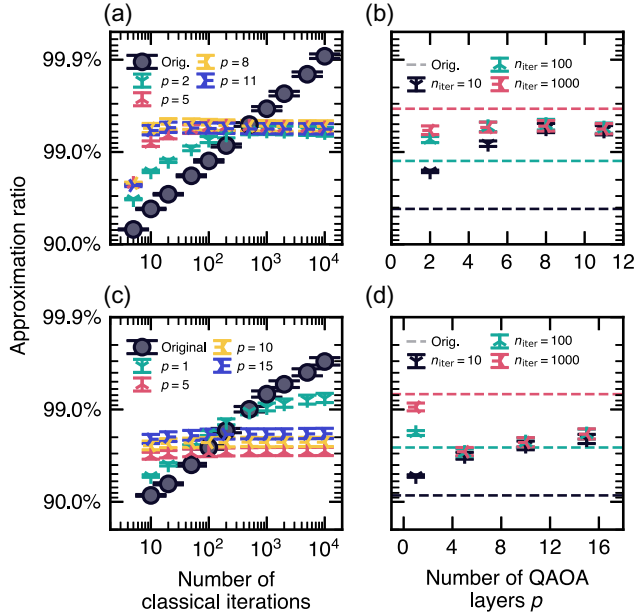


FIG. 7. (a),(b)  $N = 16$  random 3-regular graph maximum-cut problems. (c),(d)  $N = 16$  SK spin glasses. (a),(c) Average approximation ratio as a function of the number of classical iterations  $n_{\text{iter}}$  for simulated annealing, based on quantum preconditioning, for various values of  $p$ . (b),(d) Average approximation ratio as a function of the number of QAOA layers  $p$  used for quantum preconditioning for various number of iterations  $n_{\text{iter}}$  in simulated annealing. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

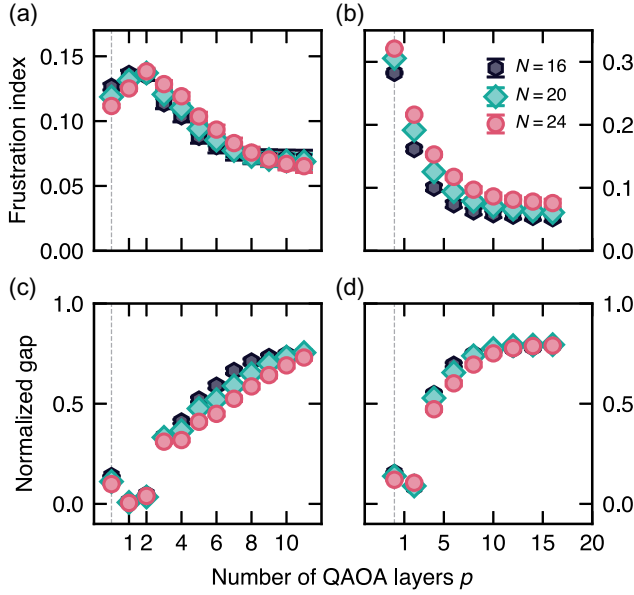


FIG. 8. (a),(c)  $N$ -variable random 3-regular graph maximum-cut problems. (b),(d)  $N$ -variable SK spin glasses. (a),(b) Average frustration index defined in Eq. (B1) as a function of the number of QAOA layers  $p$ . (c),(d) Average normalized gap defined in Eq. (B2) as a function of the number of QAOA layers  $p$ . Vertical dashed lines indicate the quantities evaluated on the original problem  $W$ . Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

this work given the challenges of scaling up quantum circuits to  $p = 3$  and beyond on relevant large problem sizes  $N$  (see discussions in Sec. IV B).

### Optimal solution overlap

Next, we consider the overlap of a sampled solution  $\mathbf{z}$  via simulated annealing with the optimal one  $\mathbf{z}_{\text{opt}}$ . This is defined as

$$q^2 = \left( \frac{1}{N} \sum_{i=1}^N z_{\text{opt},i} z_i \right)^2, \quad (3)$$

reminiscent of the Edwards-Anderson overlap [162–164]. In Fig. 9, we plot the overlap with the optimal solution as a function of the number of classical iterations in simulated annealing for  $N = 128$  random 3-regular graph maximum-cut problems and SK spin glasses.

In both cases, we find that for a fixed number of iterations of the classical solver, quantum preconditioning yields solutions  $\mathbf{z}$  with a greater overlap with the optimal solution than simulated annealing working on the original problem. The overlap follows the trend of the approximation ratio in Figs. 14(c) and 17(c), where we anticipate that the original problem will asymptotically reach  $q^2 \rightarrow 1$ , while the preconditioned problems, not sharing the same

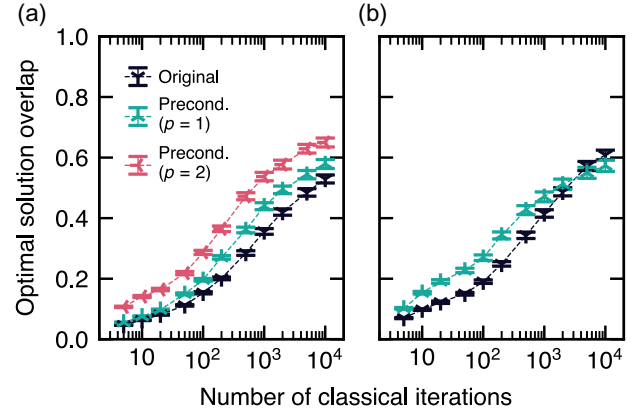


FIG. 9. Average optimal solution overlap as defined in Eq. (B3) as a function of the number of classical iterations in simulated annealing. (a)  $N = 128$ -variable random 3-regular graph maximum-cut problems. (b)  $N = 128$ -variable SK spin glasses. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

optimal solutions, will result in  $q^2 < 1$ . The larger overlap supports quantum preconditioning easing the convergence of classical solvers toward near-optimal solutions.

### APPENDIX C: ON THE LIGHT-CONE DECOMPOSITION TECHNIQUE

On sparse problem instances, such as random 3-regular graphs and the South Carolina grid network, one can leverage a light-cone decomposition technique for estimating expectation values such as  $\langle \hat{Z}_i \hat{Z}_j \rangle$  for preconditioning the problem.

Indeed, a  $p$ -layer QAOA circuit as defined in Eq. (3) induces quantum coherence around each qubit and its  $p$ -nearest neighbors, as per the underlying graph topology defined by  $\mathcal{G}$ . This defines a light-cone-induced subgraph  $\mathcal{G}_i$  for each of the qubits or problem variables  $i$ , which can straightforwardly be obtained. For computing an expectation value  $\langle \hat{Z}_i \hat{Z}_j \cdots \hat{Z}_m \rangle$ , one may only consider the union of the light-cone-induced subgraphs for each of the involved variables, i.e.,  $\mathcal{G}_{ij \dots m} = \mathcal{G}_i \cup \mathcal{G}_j \cdots \cup \mathcal{G}_m$ . Executing the QAOA on either  $\mathcal{G}$  or  $\mathcal{G}_{ij \dots m}$  will lead to the same expectation value  $\langle \hat{Z}_i \hat{Z}_j \cdots \hat{Z}_m \rangle$ .

The light-cone technique presents an advantage for computing  $\langle \hat{Z}_i \hat{Z}_j \cdots \hat{Z}_m \rangle$  when  $\mathcal{G}_{ij \dots m}$  is smaller than  $\mathcal{G}$ , as the resulting QAOA execution involves fewer qubits and gates. The price is that of a different QAOA execution for every single desired expectation value  $\langle \hat{Z}_i \hat{Z}_j \cdots \hat{Z}_m \rangle$  corresponding to different subgraphs  $\mathcal{G}_{ij \dots m}$ .

For a graph with  $N$  nodes with a maximum degree  $d$ , the number of nodes in a light-cone-induced subgraph by the QAOA with depth  $p$  for computing a two-point correlation

$\langle \hat{Z}_i \hat{Z}_j \rangle$  is at most

$$\min \left( 1 + \frac{2d[(d-1)^p - 1]}{d-2}, N \right). \quad (4)$$

For example, this upper bound is saturated for  $d$ -regular graphs such as the random 3-regular graph considered. For  $N$ -variable all-to-all graphs where  $d = N - 1$ , such as the SK problems considered, the light-cone subgraphs involve all the  $N$  nodes, meaning that the execution of the QAOA on the full graph is required. This worst-case bound is not saturated for the grid energy network. The light-cone technique is exact and leads to smaller simulations for sparse graphs at shallow QAOA depth  $p$  while mapping to the standard QAOA for dense graphs or large  $p$  values.

#### APPENDIX D: EXPERIMENTAL IMPLEMENTATION OF THE SOUTH CAROLINA MAXIMUM POWER EXCHANGE SECTION PROBLEM

We implement quantum preconditioning at  $p = 1$  on the Rigetti Ankaa-3 superconducting quantum computer for the South Carolina grid energy problem of Sec. III C. Using the light-cone decomposition technique for computing necessary expectation values, this requires executing 600 QAOA circuits ranging from  $N = 4$  to  $N = 14$  qubits (see Table I). We use the transpilation strategy of Sec. b in terms of one- and two-qubit gates. The average fidelity of the two-qubit  $i$ SWAP gates is approximately 98.6%, that of the one-qubit gate RX is approximately 99.8%, and that of the readout is approximately 96.7% [165]. Computations are performed on a linear chain of qubits, with the phase separator of the QAOA compiled via an efficient swap network [78,79]. Through a naïve swap network,  $N$ -qubit QAOA circuits at  $p = 1$  require  $3N^2/2$  two-qubit  $i$ SWAP gates [69]—i.e., 24 for  $N = 4$  and 294 for  $N = 14$ . Instead, the efficient swap network strategy makes this number range between 6 and 83 for the 600 circuits at hand—this is a three-to-four-fold improvement. We collect  $K = 10^4$  bit strings for each circuit for estimating expectation values.

In Fig. 10(a), we show the experimental two-point correlation versus the exact one obtained via classical state-vector emulations. As discussed in Sec. IV B 2, we anticipate a simple depolarizing noise model to rescale expectation values by a factor corresponding to the circuit fidelity. We find that the data roughly follows a linear relationship with a slope  $\mathcal{F} = 0.718(7)$  obtained by least-squares fitting. In Fig. 10(b), we rescale the exact value of each correlation by the corresponding circuit fidelity  $F_{ij}$ , which has been evaluated independently by multiplying the fidelity of all present individual operations, including one-qubit gates, two-qubit gates, and readout. The average fidelity over the 600 circuits is  $\mathbb{E}[F_{ij}] = 0.679(6)$ , which

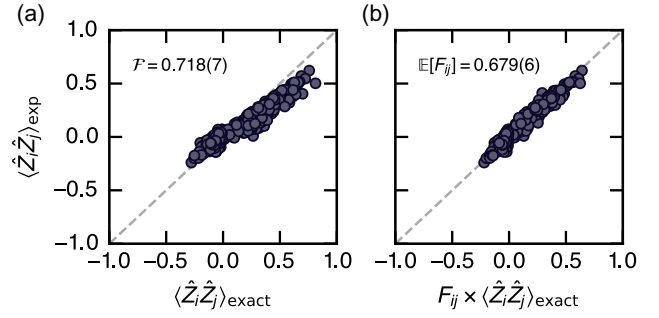


FIG. 10. There are 600 data points corresponding to  $p = 1$  QAOA circuits of the South Carolina grid energy problem. (a) Experimental two-point correlation versus the exact one. A least-squares fitting of the form  $y = \mathcal{F}x$  yields  $\mathcal{F} = 0.718(7)$ . (b) Experimental two-point correlation versus the exact one rescaled by the corresponding circuit fidelity  $F_{ij}$ . The average circuit fidelity across all the circuits is  $\mathbb{E}[F_{ij}] = 0.679(6)$ .

is relatively close to  $\mathcal{F}$  and explains why the rescaled data of Fig. 10(b) follows the  $y = x$  line more closely.

Having preconditioned the problem, we now solve it using the classical BM algorithm—a comparison against SA in Fig. 4 showed that BM outperformed it by one order of magnitude, thus favoring BM as the solver of choice. In Fig. 11, we plot the approximation ratio versus the number of iterations. We find that the experimentally preconditioned problem provides an advantage when compared to the original problem. We observe that the preconditioned problem rescaled by the circuit fidelity does not yield an

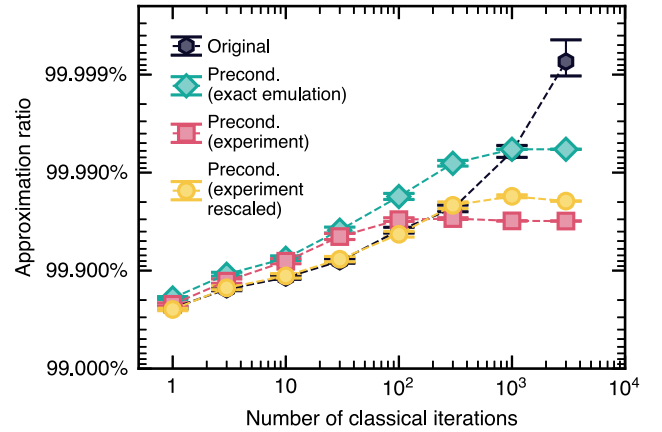


FIG. 11. Average approximation ratio against the number of iterations for the BM solver for the South Carolina grid energy problem. Different input problems are considered: the original problem,  $p = 1$  preconditioned problem via exact circuit emulations; the  $p = 1$  preconditioned problem via experimental computations; and the  $p = 1$  preconditioned problem via exact circuit emulations with values rescaled by the circuit fidelity. Each data point is averaged over 100 random independent runs of the BM solver. Error bars indicate the standard error of the mean.

advantage here. At current noise levels, the benefit of preconditioning by implementing the QAOA in experiment is lower than that of the exact classical state-vector emulation; however, we emphasize that we are achieving average approximation ratios above  $\alpha \simeq 99.97\%$  and that the observed differences are extremely small.

This highlights the experimental applicability of the proposed quantum-preconditioning method, although more work is needed for quantum preconditioning to potentially deliver utility. Understanding how noise affects quantum preconditioning would be valuable.

## APPENDIX E: IMPLEMENTATION OF CLASSICAL SOLVERS

In this work, we focus on two state-of-the-art classical optimization solvers, which are both heuristics: SA and the BM algorithm. We rely on existing efficient C++ implementations for both solvers.

### Simulated annealing

Simulated annealing is a physics-inspired approach to solving that treats an objective function such as Eq. (1) as a classical Ising model for which one wants to find the ground state. The search is based on a Metropolis-Hastings Markov-chain Monte Carlo algorithm [49,50] with a Boltzmann distribution. It starts at a high temperature, which is reduced throughout the iterations according to a predefined schedule such that one samples low-energy states, and ideally the ground state in the last step.

For an  $N$ -variable problem, an iteration, also known as a sweep, attempts to flip the sign of  $N$  variables uniformly at random. Each flip is accepted (or discarded) according to a Metropolis-Hastings probability. The initial configuration  $\mathbf{z}$  is generated at random among the  $2^N$  possibilities. Each sweep  $\ell$  works with a fixed temperature  $T_\ell$ . We employ a geometric temperature schedule such that for a total of  $M$  sweeps (defined by the user), the temperature at iteration  $\ell = 1, 2, \dots, M$  reads

$$T_\ell^{-1} = \exp \left( \ln T_{\text{hot}}^{-1} + \ell \frac{\ln T_{\text{cold}}^{-1} - \ln T_{\text{hot}}^{-1}}{M} \right), \quad (5)$$

where  $T_{\text{hot}} \geq T_{\text{cold}} \in \mathbb{R}$  are the initial and final temperatures, respectively. Finding optimal initial and final temperatures may require its own analysis and fine-tuning. The shape of the temperature schedule can also be optimized beyond a simple geometric descent. Here, we employ a heuristic to set  $T_{\text{hot}}$  and  $T_{\text{cold}}$  based on the  $n$  edges and weights of the input problem. This has an algorithmic complexity of  $O(n)$ , as visible in Appendix K. This step, which is only performed once at the beginning, is independent of the number of sweeps and typically not a bottleneck for SA.

First, we define the effective field experienced by a spin  $i$  due to two-body interactions with its nearest neighbors

$$h_i = \sum_{j=1}^N |W_{ij}|. \quad (6)$$

In the context of the preconditioned problems, one substitutes  $W_{ij}$  with  $Z_{ij}^{(p)}$  as defined in Eq. (2). Initially, when the temperature is hot, we want fast mixing such that all spins can be flipped with a probability of 50% (number chosen arbitrarily). Thus,  $0.5 = \exp(-2 \max_i h_i / T_{\text{hot}})$ , where  $2 \max_i h_i$  corresponds to the largest possible energy gap between two configurations after a spin flip. Inverting the relation, one gets

$$T_{\text{hot}} = 2 \max_i h_i / \ln 2. \quad (7)$$

At the coldest (and final) temperature, we bound the probability of exciting any of the  $N$  spins by 1%. Assuming that only spins with a minimal energy gap are excitable, one has  $0.01 = \nu \exp(-2 \min_{ij} |W_{ij}| / T_{\text{cold}})$ , where we have approximated the minimal energy gap by  $2 \min_{ij} |W_{ij}|$ . Inverting the relation, one gets

$$T_{\text{cold}} = 2 \min_{ij} |W_{ij}| / \ln(100\nu), \quad (8)$$

where  $\nu \in [1, N]$  counts, for  $i = 1, \dots, N$ , the number of values  $\min_i |W_{ij}|$  equal to  $\min_{ij} |W_{ij}|$ , i.e., spins with a minimal energy gap. For instance, for the original unit-weight random 3-regular graph maximum-cut problems,  $h_i = 3\forall i$ , such that  $T_{\text{hot}} = 6 \ln 2$  and  $T_{\text{cold}} = 6 / \ln(100N)$ , since  $\nu = N$  in this case.

The SA implementation used throughout this work relies on the PYTHON package DWAVE-SAMPLERS 1.4.0 [166], although the underlying code is in C++ for efficiency. The temperature-setting heuristic described above is the default one in this implementation.

### Burer-Monteiro solver

The BM solver eliminates the positive semidefinite constraint of the GW algorithm for the maximum-cut problem. The GW algorithm is based on semidefinite programming methods and the solver guarantees the highest possible approximation ratio  $\alpha \simeq 0.878$  [10,11]. Thus, the BM solver involves a nonconvex objective function, which may result in multiple local nonglobal minima. In practice, this is not an issue, and the BM solver is considered one of the best heuristics to date for the maximum-cut problem [12]. Its foundations are still intensively studied by the community; see, e.g., [167,168] and references therein.

We employ the MQLIB implementation of the BM solver [12], written in C++ for efficiency. This implementation does not readily expose any hyperparameters for the solver

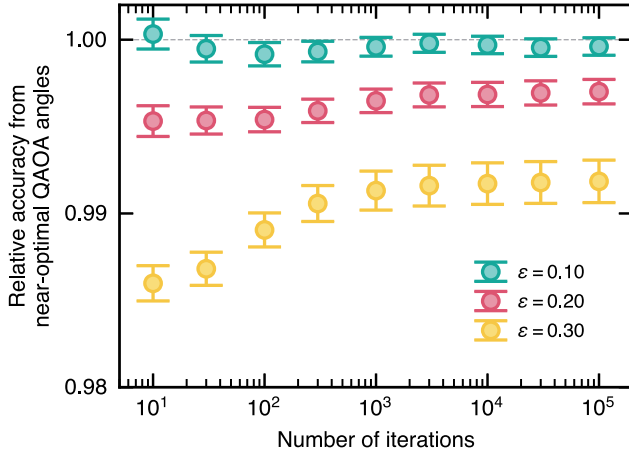


FIG. 12. Effect on the approximation ratio returned by simulated annealing versus the number of iterations due to variations in QAOA angles defined by  $\varepsilon > 0$ , relative to the near-optimal ( $\varepsilon = 0$ ) case. SK problem instances with  $N = 64$  variables at  $p = 1$  are considered, where each data point is averaged over 500 randomly generated instances. Error bars indicate the standard error of the mean.

and we keep the default values. This is why, unlike for SA, the run time is strictly proportional to the number of iterations (see Appendix K). Our only modification of the MQLIB code exposes the number of iterations performed by the solver in addition to its run time.

#### APPENDIX F: QUANTUM PRECONDITIONING PERFORMANCE AWAY FROM NEAR-OPTIMAL QAOA ANGLES

We investigate the performance of the proposed quantum-preconditioning approach away from near-optimal QAOA angles. We perform numerical simulations on  $N = 64$  SK problem instances at  $p = 1$  with QAOA angles defined as

$$\gamma = (\varepsilon \cos \theta + 0.5) / \sqrt{N}, \quad \beta = \varepsilon \sin \theta + \pi/8, \quad (9)$$

where  $\varepsilon$  measures the Euclidean distance to the near-optimal tabulated QAOA angles and  $\theta$  is a random number uniformly drawn from the range  $[0, 2\pi]$ . For  $\varepsilon = 0$ , one recovers the near-optimal tabulated angles [27].

In Fig. 12, we show the effect of  $\varepsilon > 0$  relative to the  $\varepsilon = 0$  case on the performance of simulated annealing versus the number of iterations, as measured by  $\alpha(\varepsilon)/\alpha(\varepsilon = 0)$ , where  $\alpha$  is the average approximation ratio over problem instances and samples. We observe that the quantum-preconditioning approach is quite robust to variations of QAOA angles. The sensitivity to  $\varepsilon$  for larger  $p$  remains open.

Finally, we note that for random 3-regular graphs, the tabulated angles that we used throughout this work are not optimal on an instance-by-instance basis but are

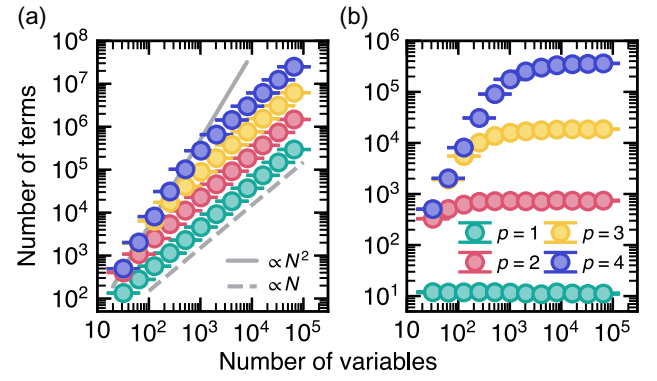


FIG. 13. (a) Average number of nonzero terms in quantum-preconditioned random 3-regular graph maximum-cut problems for various values of  $p$  as a function of the number of variables  $N$ . (b) Average number of terms to precondition in quantum-preconditioned random 3-regular graph maximum-cut problems for various values of  $p$  as a function of the number of variables  $N$ . It is smaller than or equal to the number of nonzero term via caching the preconditioned value of light-cone-induced tree subgraphs. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

only near-optimal on average [51], meaning that they correspond to an effective finite  $\varepsilon$  value; however, this near-optimality of the angles was enough to provide a quantum-inspired advantage via quantum preconditioning.

#### APPENDIX G: TERMS IN QUANTUM-PRECONDITIONED RANDOM 3-REGULAR GRAPH MAXIMUM-CUT PROBLEMS

Random 3-regular graph maximum-cut problems are sparse: the upper triangle of the adjacency matrix  $W$  contains  $3N/2$  entries. The QAOA with  $p$  layers induces nonzero correlations  $\langle \hat{Z}_i \hat{Z}_j \rangle$  at a distance  $2p$  between variables  $i$  and  $j$  via a light cone. This results in additional nonzero terms in the preconditioned problem when compared to the original one. In Fig. 13(a), we show the number of terms  $n$  as a function of the number of variables  $N$  for various values of  $p$ . We observe two asymptotic behaviors:  $n \sim O(N)$  for  $p \ll \ln N$  and  $n \sim O(N^2)$  for  $p \gg \ln N$ , understood through the topology of random 3-regular graphs having an average distance between two variables going as  $\ln N$ .

Although the number of nonzero terms grows with  $N$ , only an  $N$ -independent number  $O(\exp p)$  of terms needs to be individually preconditioned for  $p \ll \ln N$ . This is because asymptotically with  $N$ , most of the correlations  $\langle \hat{Z}_i \hat{Z}_j \rangle$  are identical: they correspond to light-cone-induced tree subgraphs. A tree is straightforwardly detected as a graph for which the number of edges  $n$  is  $N = n + 1$  and its corresponding correlation value  $\langle \hat{Z}_i \hat{Z}_j \rangle$  only needs to be computed once and cached.

TABLE II. Asymptotic behaviors from Fig. 13. Numbers of nonzero terms for  $p \ll \ln N$  extracted by a linear fit of the four largest problem sizes. The numbers of terms to precondition for  $p \ll \ln N$  correspond to the average value of the largest problem size with  $N = 2^{16}$  variables.

| Number of QAOA layers | Number of nonzero terms for $p \ll \ln N$ | Number of terms to precondition for $p \ll \ln N$ |
|-----------------------|-------------------------------------------|---------------------------------------------------|
| Original              | $1.5N$                                    | —                                                 |
| $p = 1$               | $4.500\,008(8)N$                          | $11.2(4)$                                         |
| $p = 2$               | $22.5000(2)N$                             | $740(17)$                                         |
| $p = 3$               | $94.500(2)N$                              | $18\,581(231)$                                    |
| $p = 4$               | $382.45(3)N$                              | $360\,937(2832)$                                  |

In Table II, we report asymptotic numerical values. We conjecture that the prefactor in the  $O(N)$  scaling for the average number of nonzero terms relates to a sequence of integers  $\{3, 9, 45, 189, 765, \dots\}$ , which is, however, not referenced in the online encyclopedia of integer sequences (OEIS).

#### APPENDIX H: SEMI-ANALYTICAL EXPRESSION FOR THE CORRELATION MATRIX AT $p = 1$

The proposed quantum preconditioning can be performed at  $p = 1$  for arbitrary  $N$ -variable problems classically and semianalytically using a back-propagation technique for evaluating expectation values  $\langle \hat{Z}_i \hat{Z}_j \rangle_{p=1}$  as a function of the QAOA parameters  $\gamma_1 \equiv \gamma$  and  $\beta_1 \equiv \beta$ , as defined in Eq. (3). For an arbitrary objective function (arbitrary adjacency matrix  $W$ ) of the form of Eq. (1), the complexity for computing the full correlation matrix is  $O(N^3)$ . We simply state

the result derived in Ref. [48] for completeness:

$$\begin{aligned}
 \langle \hat{Z}_i \hat{Z}_j \rangle_{p=1} = & -\sin(2\beta) \cos(2\beta) \sin(\gamma W_{ij}) \\
 & \times \left[ \prod_{k \neq i,j} \cos(\gamma W_{ik}) + \prod_{k \neq i,j} \cos(\gamma W_{jk}) \right] \\
 & - \frac{\sin^2(2\beta)}{2} \left[ \prod_{k \neq i,j} \cos \gamma (W_{ik} + W_{jk}) \right. \\
 & \left. - \prod_{k \neq i,j} \cos \gamma (W_{jk} - W_{ik}) \right]. \quad (10)
 \end{aligned}$$

We use Eq. (H1) for preconditioning the Sherrington-Kirkpatrick spin-glass problems studied in the main text.

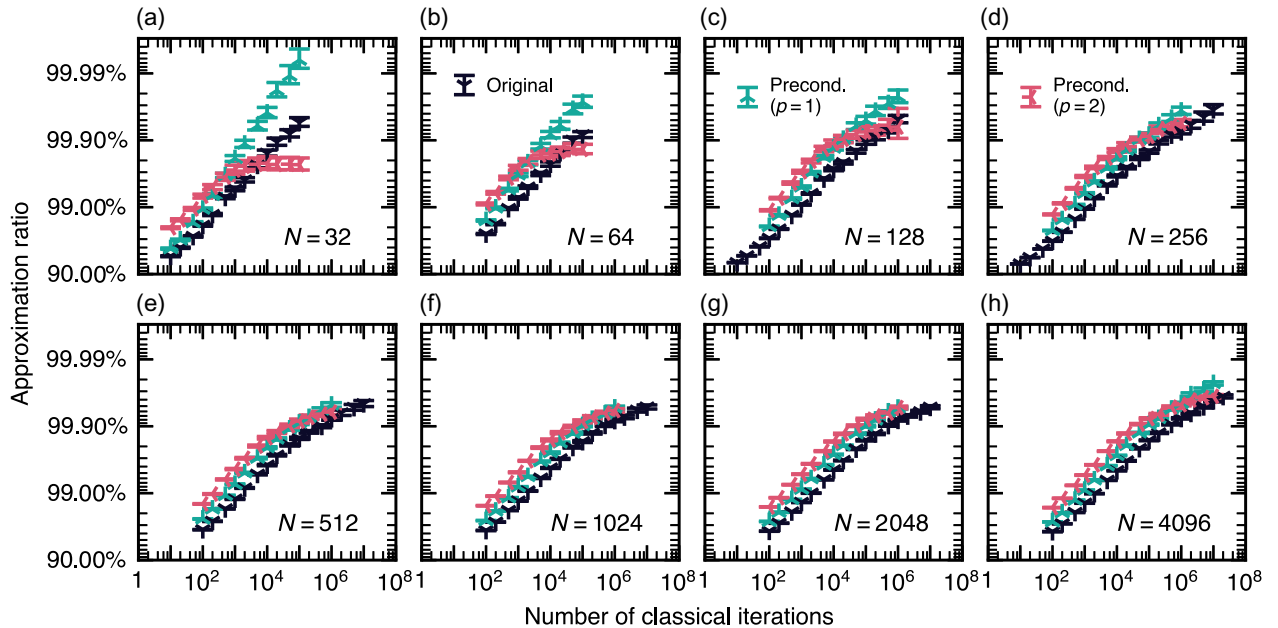


FIG. 14. Average performance (approximation ratio versus the number of iterations) for the maximum-cut problem on  $N$ -variable random 3-regular graphs via the classical SA solver based on the original and quantum-preconditioned problems. (a)  $N = 32$  variables. (b)  $N = 64$  variables. (c)  $N = 128$  variables. (d)  $N = 256$  variables. (e)  $N = 512$  variables. (f)  $N = 1024$  variables. (g)  $N = 2048$  variables. (h)  $N = 4096$  variables. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

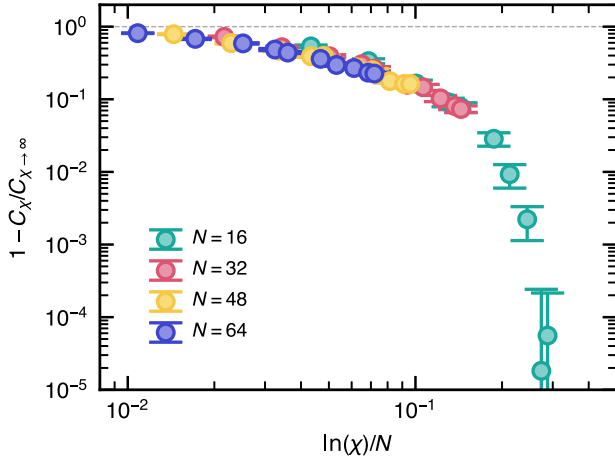


FIG. 15. Performance of simulated annealing for quantum-preconditioned SK problems with approximate MPS emulations. Ratio  $1 - C_\chi / C_{\chi \rightarrow \infty}$  against  $\ln \chi / N$ , where  $\chi$  is the MPS bond dimension and  $N$  is the number of variables. The value  $C_\chi$  is the average objective value of  $10^3$  solutions sampled with simulated annealing. Each data point is averaged over 50 randomly generated problem instances. Error bars indicate the standard error of the mean.

### APPENDIX I: PERFORMANCE AGAINST APPROXIMATE CLASSICAL TENSOR NETWORK EMULATIONS

To complement the discussion of Sec. IV B 3, we perform classical matrix-product-state (MPS) emulations of

the  $p = 1$  QAOA used to precondition an ensemble of  $N$ -variable SK problems. We employ the PYTHON library QUIMB [169] for the MPS emulations. The QAOA is emulated approximately, with the emulation quality controlled by the MPS bond dimension  $\chi$ . Then, we use simulated annealing with a fixed number of  $10^4$  iterations to evaluate the performance of the classical solver against approximately quantum-preconditioned problems.

Inspired by Ref. [73], and for each problem instance, we consider the average objective value  $C_\chi$  over  $10^3$  solutions sampled by simulated annealing. The comparison point is with respect to the solutions returned from perfectly quantum-preconditioned problems, i.e.,  $C_{\chi \rightarrow \infty}$ , which we can efficiently access at  $p = 1$  based on Sec. H.

We observe in Fig. 15 a data collapse for  $C_\chi / C_{\chi \rightarrow \infty}$  as a function of the rescaled variable  $\ln \chi / N$ . This behavior for the output of simulated annealing with quantum preconditioning was also observed for the standard QAOA [73], on which the quantum-preconditioning step is based.

For example, one needs  $\ln \chi / N \gtrsim 0.1$  for the average quality solutions of approximately quantum-preconditioned simulated annealing to be 90% as good as the exactly ( $\chi \rightarrow +\infty$ ) preconditioned one.

### APPENDIX J: PERFORMANCE VERSUS NUMBER OF ITERATIONS

#### Random 3-regular graph maximum-cut problems

We investigate the performance for the maximum-cut problem on  $N$ -variable random 3-regular graphs based on

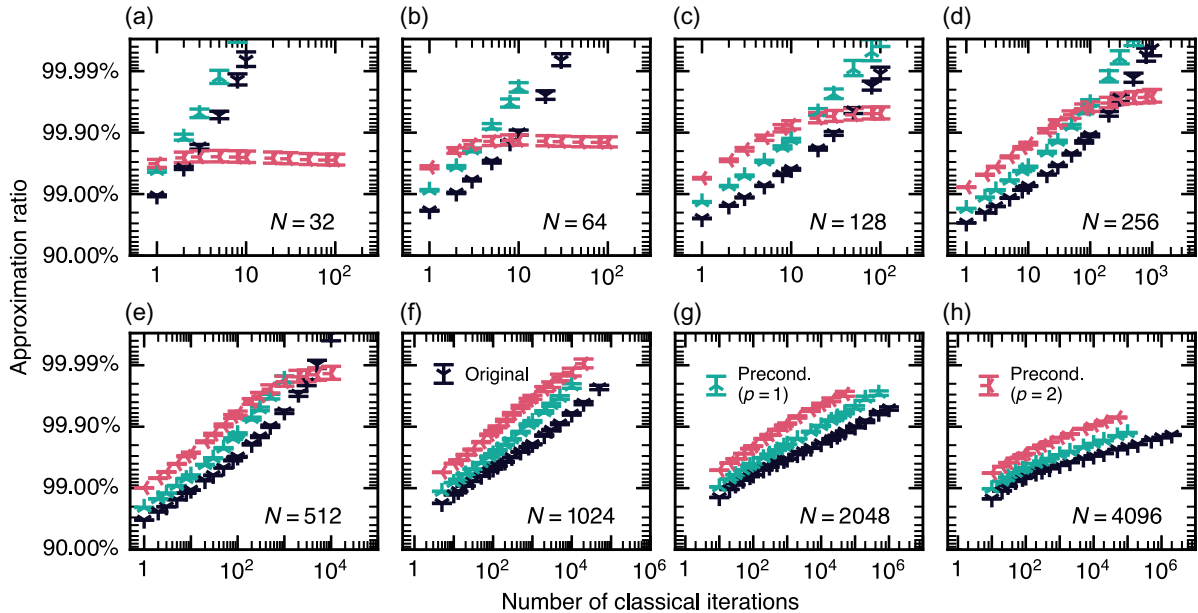


FIG. 16. Average performance (approximation ratio versus the number of iterations) for the maximum-cut problem on  $N$ -variable random 3-regular graphs via the classical BM solver based on the original and quantum-preconditioned problems. (a)  $N = 32$  variables. (b)  $N = 64$  variables. (c)  $N = 128$  variables. (d)  $N = 256$  variables. (e)  $N = 512$  variables. (f)  $N = 1024$  variables. (g)  $N = 2048$  variables. (h)  $N = 4096$  variables. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

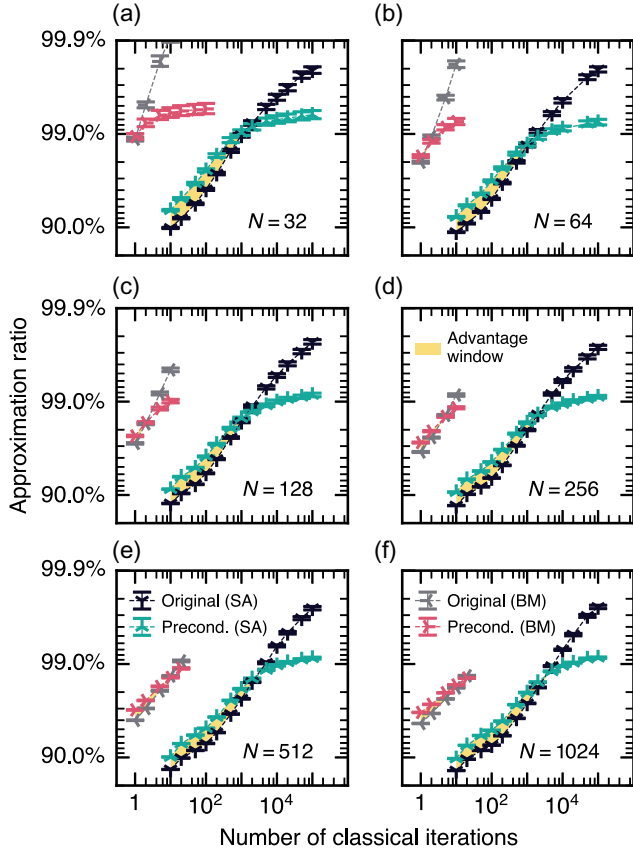


FIG. 17. Average performance (approximation ratio versus the number of iterations) for  $N$ -variable Sherrington-Kirkpatrick spin glass problems via the classical SA and BM solvers based on the original and quantum-preconditioned ( $p = 1$ ) problems. (a)  $N = 32$  variables. (b)  $N = 64$  variables. (c)  $N = 128$  variables. (d)  $N = 256$  variables. (e)  $N = 512$  variables. (f)  $N = 1024$  variables. Each data point is averaged over 200 randomly generated problem instances. Error bars indicate the standard error of the mean.

the original and quantum-preconditioned problems. We consider SA and the BM solver on problem sizes ranging from  $N = 32$  to  $N = 4096$ . For each problem size, we average the performance over 200 randomly generated instances. The preconditioning is performed using the light-cone technique described in Sec. III A 2.

We report data in Figs. 14 and 16. We observe an advantage for the preconditioned problems against the original one as a function of the number of iterations. This advantage is enhanced by going from  $p = 1$  to  $p = 2$ . We observe that for a large number of iterations, the approximation ratio for the quantum-preconditioned problem at  $p = 2$  saturates. This is because the original and preconditioned problems do not share the same optimal solutions. The approximation ratio saturates to an  $N$ -dependent value compatible with  $1 - \alpha_{\text{saturation}} \sim N^{-1}$  up to  $N = 1024$  (not enough iterations were performed to observe the saturation for larger values of  $N$ ).

The number of iterations can be converted to a run time following Appendix K. We present run-time data in the main text.

### Sherrington-Kirkpatrick spin glasses

Next, we provide additional data on Sherrington-Kirkpatrick spin glasses. We investigate the performance of SA and the BM solver on the original and quantum-preconditioned ( $p = 1$ ) problems. Problem sizes ranging from  $N = 32$  to  $N = 1024$  are considered ( $N = 2408$  was considered in the main text in Fig. 3). We report data in Fig. 17. We find that an advantage for the preconditioned problem for the BM solver only appears for large problem sizes. We highlight this advantage window in yellow. The average approximation ratio saturation value observed for a large number of iterations,  $\alpha \simeq 99.3\%$ , seems to be roughly independent of  $N$ .

### APPENDIX K: FROM NUMBER OF ITERATIONS TO RUN TIME

Instead of using the run time, the performance of the SA and BM solvers can be estimated through the number of iterations, which is a platform-independent number.

In both solvers, an iteration consists of  $O(n)$  steps, where  $n$  is the number of terms. For instance, in the case of SA, an iteration is also known as a sweep. For an  $N$ -variable problem, a sweep involves a local update on  $O(N)$  variables, which is either accepted or refused based on the updated objective value of the proposed solution. Computing the updated objective value involves an arithmetic operation per term relating to the updated variable. Hence, the complexity cost per iteration is  $O(n)$ . A similar argument can be developed for the BM solver.

The run time  $t$  of a solver scales asymptotically with  $N$  as

$$t/n = An_{\text{iter}} + B, \quad (11)$$

where  $n$  is the number of terms in the problem and  $n_{\text{iter}}$  is the number of iterations. Here,  $A$  and  $B$  are solver-, platform-, and problem-dependent constants. In the case of SA, the constant  $B$  is nonzero and corresponds to a one-time  $O(n)$  heuristic for determining the temperature schedule based on the problem's terms (see Appendix E). This is zero for the BM solver, which does not have any hyperparameters.

We show the relation between the run time and number of iterations in Fig. 18 for the random 3-regular graph maximum-cut problems and Sherrington-Kirkpatrick spin glasses for SA and the BM solver on the original and preconditioned problems. The run time is based on a 64 GB MacBook Pro with an Apple M1 Max chip. In the case of  $N$ -variable 3-regular graphs, we have  $n = 1.5N$  terms in the original problem,  $n \simeq 4.5N$  at  $p = 1$ , and

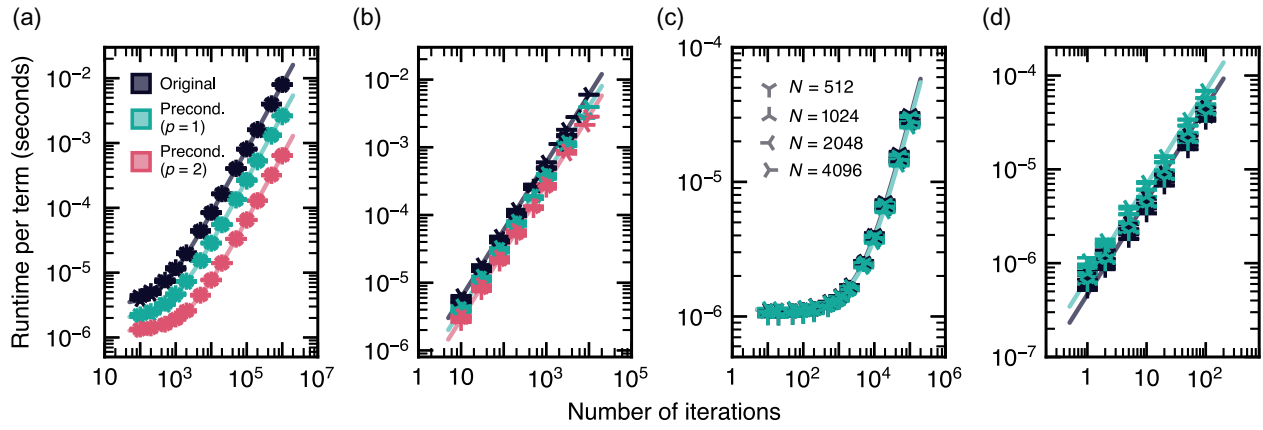


FIG. 18. Average run time per term in seconds (based on 64 GB MacBook Pro with an Apple M1 Max chip) versus the number of iterations for original and preconditioned problems. (a) SA solver for random 3-regular graph maximum-cut problems. (b) BM solver for random 3-regular graph maximum-cut problems. (c) SA solver for Sherrington-Kirkpatrick spin glasses. (d) BM solver for Sherrington-Kirkpatrick spin glasses. Error bars indicate the standard error of the mean. Bold lines are fits according to the main text.

$n \simeq 22.5N$  at  $p = 2$  in the large- $N$  limit (see Appendix G). For Sherrington-Kirkpatrick spin glasses, the number of terms is  $n = N(N - 1)/2$  in all cases. We reported the average approximation ratio as a function of both the number of iterations and run time for the grid energy problem in Fig. 4.

- [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] G. Kochenberger, J.-K. Hao, F. Glover, M. Lewis, Z. Lü, H. Wang, and Y. Wang, The unconstrained binary quadratic programming problem: A survey, *J. Comb. Optim.* **28**, 58 (2014).
- [4] C. H. Papadimitriou, in *Encyclopedia of Computer Science* (John Wiley and Sons Ltd., GBR, 2003), pp. 260–265.
- [5] F. Barahona, On the computational complexity of Ising spin glass models, *J. Phys. A Math. Theor.* **15**, 3241 (1982).
- [6] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 671 (1983).
- [7] K. Hukushima and K. Nemoto, Exchange Monte Carlo method and application to spin glass simulations, *J. Phys. Soc. Japan* **65**, 1604 (1996).
- [8] S. Burer, R. D. C. Monteiro, and Y. Zhang, Rank-two relaxation heuristics for max-cut and other binary quadratic programs, *SIAM J. Optim.* **12**, 503 (2002).
- [9] S. Burer and R. D. C. Monteiro, A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization, *Math. Program.* **95**, 329 (2003).
- [10] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability

problems using semidefinite programming, *J. ACM* **42**, 1115 (1995).

- [11] D. P. Williamson and D. B. Shmoys, *The design of approximation algorithms* (Cambridge University Press, 2011).
- [12] I. Dunning, S. Gupta, and J. Silberholz, What works best when? A systematic evaluation of heuristics for max-cut and QUBO, *INFORMS J. Comput.* **30**, 608 (2018).
- [13] A. Steane, Quantum computing, *Rep. Prog. Phys.* **61**, 117 (1998).
- [14] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, *Nature* **464**, 45 (2010).
- [15] A. Abbas *et al.*, Challenges and opportunities in quantum optimization, *Nat. Rev. Phys.* **6**, 718 (2024).
- [16] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, *arXiv:quant-ph/0001106*.
- [17] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, *Science* **292**, 472 (2001).
- [18] A. Das and B. K. Chakrabarti, Colloquium: Quantum annealing and analog quantum computation, *Rev. Mod. Phys.* **80**, 1061 (2008).
- [19] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, *Rep. Prog. Phys.* **83**, 054401 (2020).
- [20] T. Albash and D. A. Lidar, Demonstration of a scaling advantage for a quantum annealer over simulated annealing, *Phys. Rev. X* **8**, 031016 (2018).
- [21] S. Ebadi *et al.*, Quantum optimization of maximum independent set using Rydberg atom arrays, *Science* **376**, 1209 (2022).
- [22] A. D. King *et al.*, Quantum critical dynamics in a 5000-qubit programmable spin glass, *Nature* **617**, 61 (2023).
- [23] H. Muñoz-Bauza and D. Lidar, Scaling advantage in approximate optimization with quantum annealing, *Phys. Rev. Lett.* **134**, 160601 (2025).

- [24] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [25] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](#).
- [26] K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, A review on quantum approximate optimization algorithm and its variants, *Phys. Rep.* **1068**, 1 (2024).
- [27] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, *Quantum* **6**, 759 (2022).
- [28] 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.
- [29] J. Wurtz and P. Love, Maxcut quantum approximate optimization algorithm performance guarantees for  $p \geq 1$ , *Phys. Rev. A* **103**, 042612 (2021).
- [30] A. Montanaro and L. Zhou, Quantum speedups in solving near-symmetric optimization problems by low-depth QAOA, [arXiv:2411.04979](#).
- [31] R. Shaydulin et al., Evidence of scaling advantage for the quantum approximate optimization algorithm on a classically intractable problem, *Sci. Adv.* **10**, eadm6761 (2024).
- [32] J. A. Montañez-Barrera and K. Michielsen, Toward a linear-ramp QAOA protocol: Evidence of a scaling advantage in solving some combinatorial optimization problems, *npj Quantum Inf.* **11**, 131 (2025).
- [33] M. B. Hastings, A short path quantum algorithm for exact optimization, *Quantum* **2**, 78 (2018).
- [34] A. M. Dalzell, N. Pancotti, E. T. Campbell, and F. G. Brandão, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023), pp. 1131–1144.
- [35] E. Kapit, B. A. Barton, S. Feeney, G. Grattan, P. Patnaik, J. Sagal, L. D. Carr, and V. Oganessian, On the approximability of random-hypergraph MAX-3-XORSAT problems with quantum algorithms, [arXiv:2312.06104](#).
- [36] S. P. Jordan, N. Shutty, M. Wootters, A. Zalcman, A. Schmidhuber, R. King, S. V. Isakov, and R. Babbush, Optimization by decoded quantum interferometry, [arXiv:2408.08292](#).
- [37] E. Turkel, Preconditioning techniques in computational fluid dynamics, *Annu. Rev. Fluid Mech.* **31**, 385 (1999).
- [38] A. J. Wathen, Preconditioning, *Acta Numer.* **24**, 329 (2015).
- [39] E. D. Andersen and K. D. Andersen, Presolving in linear programming, *Math. Program.* **71**, 221 (1995).
- [40] G. Gamrath, T. Koch, A. Martin, M. Miltenberger, and D. Weninger, Progress in presolving for mixed integer programming, *Math. Program. Comput.* **7**, 367 (2015).
- [41] A. Lucas, Ising formulations of many NP problems, *Front. Phys.* **2**, 5 (2014).
- [42] D. Sherrington and S. Kirkpatrick, Solvable model of a spin-glass, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [43] ARPA-E GridData Program, ACTIVSg500: 500 bus synthetic grid on footprint of South Carolina (2017).
- [44] M. Dupont, B. Sundar, B. Evert, D. E. B. Neira, Z. Peng, S. Jeffrey, and M. J. Hodson, Benchmarking quantum optimization for the maximum-cut problem on a superconducting quantum computer, *Phys. Rev. Appl.* **23**, 014045 (2025).
- [45] K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, *Rev. Mod. Phys.* **58**, 801 (1986).
- [46] T. Castellani and A. Cavagna, Spin-glass theory for pedestrians, *J. Stat. Mech.: Theory Exp.* **2005**, P05012 (2005).
- [47] P. Charbonneau, E. Marinari, M. Mézard, G. Parisi, F. Ricci-Tersenghi, G. Sicuro, and F. Zamponi, *Spin Glass Theory and Far Beyond* (World Scientific, Singapore, 2023).
- [48] M. Dupont and B. Sundar, Extending relax-and-round combinatorial optimization solvers with quantum correlations, *Phys. Rev. A* **109**, 012429 (2024).
- [49] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**, 1087 (2004).
- [50] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**, 97 (1970).
- [51] J. Wurtz and D. Lykov, Fixed-angle conjectures for the quantum approximate optimization algorithm on regular MaxCut graphs, *Phys. Rev. A* **104**, 052419 (2021).
- [52] 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, Quantum-enhanced greedy combinatorial optimization solver, *Sci. Adv.* **9**, eadi0487 (2023).
- [53] F. B. Maciejewski, S. Hadfield, B. Hall, M. Hodson, M. Dupont, B. Evert, J. Sud, M. S. Alam, Z. Wang, S. Jeffrey, B. Sundar, P. A. Lott, S. Grabbe, E. G. Rieffel, M. J. Reagor, and D. Venturelli, Design and execution of quantum circuits using tens of superconducting qubits and thousands of gates for dense Ising optimization problems, *Phys. Rev. Appl.* **22**, 044074 (2024).
- [54] M. P. Harrigan et al., Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, *Nat. Phys.* **17**, 332 (2021).
- [55] D. Venturelli, S. Mandrà, S. Knysh, B. O’Gorman, R. Biswas, and V. Smelyanskiy, Quantum optimization of fully connected spin glasses, *Phys. Rev. X* **5**, 031040 (2015).
- [56] H. Jing, Y. Wang, and Y. Li, Data-driven quantum approximate optimization algorithm for power systems, *Comms. Eng.* **2**, 12 (2023).
- [57] H. Jing, Y. Wang, Y. Li, L. Du, and Z. Wu, in *2022 IEEE Transportation Electrification Conference & Expo (ITEC)* (IEEE, Detroit, MI, USA, 2022), pp. 574–578.
- [58] H. Jing, Y. Wang, Y. Li, L. Du, and Z. Wu, in *2022 IEEE Energy Conversion Congress and Exposition (ECCE)* (IEEE, Detroit, MI, USA, 2022), pp. 1–5.

- [59] N. Bauer, K. Yeter-Aydeniz, E. Kokkas, and G. Siopsis, Solving power grid optimization problems with Rydberg atoms, [arXiv:2404.11440](#).
- [60] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, in *2016 IEEE Power and Energy Conference at Illinois (PECI)* (IEEE, Urbana, IL, USA, 2016), pp. 1–6.
- [61] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye, and T. J. Overbye, Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies, *IEEE Trans. Power Syst.* **32**, 1502 (2017).
- [62] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, Grid structural characteristics as validation criteria for synthetic networks, *IEEE Trans. Power Syst.* **32**, 3258 (2017).
- [63] T. Xu, A. B. Birchfield, K. S. Shetye, and T. J. Overbye, in *The 10th Bulk Power Systems Dynamics and Control Symposium (IREP 2017)* (IEEE, 2017), pp. 1–6.
- [64] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, in *50th Hawaii International Conference on System Sciences (HICSS)* (IEEE, 2017).
- [65] C. G. Broyden, The convergence of a class of double-rank minimization algorithms 1. General considerations, *IMA J. Appl. Math.* **6**, 76 (1970).
- [66] R. Fletcher, A new approach to variable metric algorithms, *Comput. J.* **13**, 317 (1970).
- [67] D. Goldfarb, A family of variable-metric methods derived by variational means, *Math. Comput.* **24**, 23 (1970).
- [68] D. F. Shanno, Conditioning of quasi-Newton methods for function minimization, *Math. Comput.* **24**, 647 (1970).
- [69] B. O’Gorman, W. J. Huggins, E. G. Rieffel, and K. B. Whaley, Generalized swap networks for near-term quantum computing, [arXiv:1905.05118](#).
- [70] P. J. Karalekas, N. A. Tezak, E. C. Peterson, C. A. Ryan, M. P. da Silva, and R. S. Smith, A quantum-classical cloud platform optimized for variational hybrid algorithms, *Quantum Sci. Technol.* **5**, 024003 (2020).
- [71] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys. (N. Y.)* **349**, 117 (2014).
- [72] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, *Rev. Mod. Phys.* **93**, 045003 (2021).
- [73] M. Dupont, N. Didier, M. J. Hodson, J. E. Moore, and M. J. Reagor, Calibrating the classical hardness of the quantum approximate optimization algorithm, *PRX Quantum* **3**, 040339 (2022).
- [74] M. Dupont, N. Didier, M. J. Hodson, J. E. Moore, and M. J. Reagor, Entanglement perspective on the quantum approximate optimization algorithm, *Phys. Rev. A* **106**, 022423 (2022).
- [75] R. Sreedhar, P. Vikstål, M. Svensson, A. Ask, G. Johansson, and L. García-Álvarez, The quantum approximate optimization algorithm performance with low entanglement and high circuit depth, [arXiv:2207.03404](#).
- [76] T. Ayril, T. Louvet, Y. Zhou, C. Lambert, E. M. Stoudenmire, and X. Waintal, Density-matrix renormalization group algorithm for simulating quantum circuits with a finite fidelity, *PRX Quantum* **4**, 020304 (2023).
- [77] J. Weidenfeller, L. C. Valor, J. Gacon, C. Tornow, L. Bello, S. Woerner, and D. J. Egger, Scaling of the quantum approximate optimization algorithm on superconducting qubit based hardware, *Quantum* **6**, 870 (2022).
- [78] A. Matsuo, S. Yamashita, and D. J. Egger, A SAT approach to the initial mapping problem in swap gate insertion for commuting gates, *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **E106.A**, 1424 (2023).
- [79] S. H. Sack and D. J. Egger, Large-scale quantum approximate optimization on nonplanar graphs with machine learning noise mitigation, *Phys. Rev. Res.* **6**, 013223 (2024).
- [80] C.-Y. Liu and H.-S. Goan, Hybrid gate-based and annealing quantum computing for large-size Ising problems, [arXiv:2208.03283](#).
- [81] T. L. Patti, J. Kossai, A. Anandkumar, and S. F. Yelin, Variational quantum optimization with multibasis encodings, *Phys. Rev. Res.* **4**, 033142 (2022).
- [82] M. Bechtold, J. Barzen, F. Leymann, A. Mandl, J. Obst, F. Truger, and B. Weder, Investigating the effect of circuit cutting in QAOA for the MaxCut problem on NISQ devices, *Quantum Sci. Technol.* **8**, 045022 (2023).
- [83] M. Ponce, R. Herrman, P. C. Lotshaw, S. Powers, G. Siopsis, T. Humble, and J. Ostrowski, Graph decomposition techniques for solving combinatorial optimization problems with variational quantum algorithms, *Quantum Inf. Process.* **24**, 60 (2025).
- [84] A. Angone, X. Liu, R. Shaydulin, and I. Safro, in *2023 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, 2023), pp. 1–7.
- [85] M. Sciorilli, L. Borges, T. L. Patti, D. García-Martín, G. Camilo, A. Anandkumar, and L. Aolita, Towards large-scale quantum optimization solvers with few qubits, *Nat. Comm.* **16**, 476 (2025).
- [86] B. Bach, J. Falla, and I. Safro, in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Montreal, Quebec, Canada, 2024), Vol. 01, pp. 1–12.
- [87] J. Moondra, P. C. Lotshaw, G. Mohler, and S. Gupta, Promise of graph sparsification and decomposition for noise reduction in QAOA: Analysis for trapped-ion compilations, [arXiv:2406.14330](#).
- [88] B. Sundar and M. Dupont, Qubit-efficient quantum combinatorial optimization solver, [arXiv:2407.15539](#).
- [89] F. B. Maciejewski, B. G. Bach, M. Dupont, P. A. Lott, B. Sundar, D. E. B. Neira, I. Safro, and D. Venturelli, in *2024 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, 2024), pp. 1–10.
- [90] A. Acharya, R. Yalovetzky, P. Minssen, S. Chakrabarti, R. Shaydulin, R. Raymond, Y. Sun, D. Herman, R. S. Andrist, G. Salton, M. J. A. Schuetz, H. G. Katzgraber, and M. Pistoia, Decomposition pipeline for large-scale portfolio optimization with applications to near-term quantum computing, *Phys. Rev. Res.* **7**, 023142 (2025).
- [91] L. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard, *Phys. Rev. Lett.* **127**, 120502 (2021).
- [92] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Barren plateaus in variational quantum computing, *Nat. Rev. Phys.* **7**, 174 (2025).

- [93] R. Shaydulin, I. Safro, and J. Larson, in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, 2019), pp. 1–8.
- [94] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices, *Phys. Rev. X* **10**, 021067 (2020).
- [95] A. Galda, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2021), pp. 171–180.
- [96] A. Galda, E. Gupta, J. Falla, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, Similarity-based parameter transferability in the quantum approximate optimization algorithm, *Front. Quantum Sci. Technol.* **2**, 1200975 (2023).
- [97] T. Lubinski, C. Coffrin, C. McGeoch, P. Sathe, J. Apanavicius, D. Bernal Neira, and Q. E. D. C.-C. Collaboration, Optimization applications as quantum performance benchmarks, *ACM Trans. Quantum Comput.* **5**, 18 (2024).
- [98] R. Shaydulin, P. C. Lotshaw, J. Larson, J. Ostrowski, and T. S. Humble, Parameter transfer for quantum approximate optimization of weighted MaxCut, *ACM Trans. Quantum Comput.* **4**, 1 (2023).
- [99] S. H. Sureshbabu, D. Herman, R. Shaydulin, J. Basso, S. Chakrabarti, Y. Sun, and M. Pistoia, Parameter setting in quantum approximate optimization of weighted problems, *Quantum* **8**, 1231 (2024).
- [100] Z. He, R. Shaydulin, D. Herman, C. Li, R. Raymond, S. H. Sureshbabu, and M. Pistoia, in *Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design, ICCAD '24* (Association for Computing Machinery, New York, NY, USA, 2025).
- [101] S. Hadfield, Z. Wang, B. O'gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, *Algorithms* **12**, 34 (2019).
- [102] F. G. Fuchs, K. O. Lye, H. Møll Nilsen, A. J. Stasik, and G. Sartor, Constraint preserving mixers for the quantum approximate optimization algorithm, *Algorithms* **15**, 202 (2022).
- [103] A. Glos, A. Krawiec, and Z. Zimborás, Space-efficient binary optimization for variational computing, *npj Quantum Inf.* **8**, 39 (2022).
- [104] M. Schnaus, L. Palackal, B. Poggel, X. Runge, H. Ehm, J. M. Lorenz, and C. B. Mendl, in *2024 IEEE International Conference on Quantum Software (QSW)* (IEEE, Shenzhen, China, 2024), pp. 81–87.
- [105] Z. Tabi, K. H. El-Safty, Z. Kallus, P. Hágá, T. Kozsik, A. Glos, and Z. Zimborás, in *2020 IEEE international conference on quantum computing and engineering (QCE)*, Denver, CO, USA (IEEE, Denver, CO, USA, 2020), pp. 56–62.
- [106] F. G. Fuchs, H. Ø. Kolden, N. H. Aase, and G. Sartor, Efficient encoding of the weighted MAX  $k$ -CUT on a quantum computer using QAOA, *SN Comput. Sci.* **2**, 89 (2021).
- [107] S. Hadfield, On the representation of boolean and real functions as Hamiltonians for quantum computing, *ACM T. Quant. Comput.* **2**, 1 (2021).
- [108] R. Fakhimi, H. Validi, I. V. Hicks, T. Terlaky, and L. F. Zuluaga, Quantum-inspired formulations for the max  $k$ -cut problem, ISE Technical Report 21T-007 (2021).
- [109] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, *Rev. Mod. Phys.* **95**, 045005 (2023).
- [110] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
- [111] J. Robledo-Moreno, M. Motta, H. Haas, A. Javadi-Abhari, P. Jurcevic, W. Kirby, S. Martiel, K. Sharma, S. Sharma, T. Shirakawa, I. Sitdikov, R.-Y. Sun, K. J. Sung, M. Takita, M. C. Tran, S. Yunoki, and A. Mezzacapo, Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer, *Sci. Adv.* **11**, eadu9991 (2025).
- [112] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, *Phys. Rev. A* **94**, 052325 (2016).
- [113] A. Arrasmith, A. Patterson, A. Boughton, and M. Paini, Development and demonstration of an efficient readout error mitigation technique for use in NISQ algorithms, [arXiv:2303.17741](https://arxiv.org/abs/2303.17741).
- [114] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, Detector tomography on IBM quantum computers and mitigation of an imperfect measurement, *Phys. Rev. A* **100**, 052315 (2019).
- [115] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, *Quantum* **4**, 257 (2020).
- [116] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, Unfolding quantum computer readout noise, *npj Quantum Inf.* **6**, 84 (2020).
- [117] M. R. Geller, Conditionally rigorous mitigation of multi-qubit measurement errors, *Phys. Rev. Lett.* **127**, 090502 (2021).
- [118] K. E. Hamilton, T. Kharazi, T. Morris, A. J. McCaskey, R. S. Bennink, and R. C. Pooser, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Denver, CO, USA (IEEE, Denver, CO, USA, 2020), pp. 430–440.
- [119] M. R. Geller and M. Sun, Toward efficient correction of multiqubit measurement errors: Pair correlation method, *Quantum Sci. Technol.* **6**, 025009 (2021).
- [120] S. Bravyi, S. Sheldon, A. Kandala, D. C. McKay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, *Phys. Rev. A* **103**, 042605 (2021).
- [121] S. Seo, J. Seong, and J. Bae, Mitigation of crosstalk errors in a quantum measurement and its applications, [arXiv:2112.10651](https://arxiv.org/abs/2112.10651).
- [122] A. W. R. Smith, K. E. Khosla, C. N. Self, and M. S. Kim, Qubit readout error mitigation with bit-flip averaging, *Sci. Adv.* **7**, eabi8009 (2021).
- [123] E. Van Den Berg, Z. K. Mineev, and K. Temme, Model-free readout-error mitigation for quantum expectation values, *Phys. Rev. A* **105**, 032620 (2022).

- [124] Z. He, D. Amaro, R. Shaydulin, and M. Pistoia, Performance of quantum approximate optimization with quantum error detection, *Commun. Phys.* **8**, 217 (2025).
- [125] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. Bonilla Ataides, N. Maskara, I. Cong, X. Gao, P. Sales Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin, Logical quantum processor based on reconfigurable atom arrays, *Nature* **626**, 58 (2024).
- [126] R. Acharya et al., Quantum error correction below the surface code threshold, *Nature* **638**, 920 (2024).
- [127] B. W. Reichardt et al., Logical computation demonstrated with a neutral atom quantum processor, *arXiv:2411.11822*.
- [128] R. C. U. States, Data for “Optimization via quantum preconditioning” (2025), [doi:10.5281/zenodo.14921060](https://doi.org/10.5281/zenodo.14921060).
- [129] R. Shaydulin and M. Pistoia, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2023), Vol. 01, pp. 1074–1077.
- [130] S. A. Moses et al., A race-track trapped-ion quantum processor, *Phys. Rev. X* **13**, 041052 (2023).
- [131] M. DeCross, E. Chertkov, M. Kohagen, and M. Foss-Feig, Qubit-reuse compilation with mid-circuit measurement and reset, *Phys. Rev. X* **13**, 041057 (2023).
- [132] R. Tate and S. Eidenbenz, Theoretical approximation ratios for warm-started QAOA on 3-regular max-cut instances at depth  $p = 1$ , *arXiv:2402.12631*.
- [133] N. Sachdeva, G. S. Hartnett, S. Maity, S. Marsh, Y. Wang, A. Winick, R. Dougherty, D. Canuto, Y. Q. Chong, M. Hush, P. S. Mundada, C. D. B. Bentley, M. J. Biercuk, and Y. Baum, Quantum optimization using a 127-qubit gate-model IBM quantum computer can outperform quantum annealers for nontrivial binary optimization problems, *arXiv:2406.01743*.
- [134] C. C. McGeoch, K. Chern, P. Farré, and A. K. King, A comment on comparing optimization on D-Wave and IBM quantum processors, *arXiv:2406.19351*.
- [135] T. D. Morris, A. Kaushik, M. Roetteler, and P. C. Lotshaw, Performant near-term quantum combinatorial optimization, *arXiv:2404.16135*.
- [136] B. Augustino, M. Cain, E. Farhi, S. Gupta, S. Gutmann, D. Ranard, E. Tang, and K. V. Kirk, Strategies for running the QAOA at hundreds of qubits, *arXiv:2410.03015*.
- [137] D. Zhong, A. Francis, and E. Rrapaj, Classical optimization with imaginary time block encoding on quantum computers: The MaxCut problem, *arXiv:2411.10737*.
- [138] P. Berman and M. Karpinski, in *Automata, Languages and Programming*, edited by J. Wiedermann, P. van Emde Boas, and M. Nielsen (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999), pp. 200–209.
- [139] E. Halperin, D. Livnat, and U. Zwick, MAX CUT in cubic graphs, *J. Algorithms* **53**, 169 (2004).
- [140] A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th Python in Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, CA USA, 2008), pp. 11–15.
- [141] A. Steger and N. C. Wormald, Generating random regular graphs quickly, *Comb. Probab. Comput.* **8**, 377 (1999).
- [142] J. H. Kim and V. H. Vu, in *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, STOC ‘03* (Association for Computing Machinery, New York, NY, USA, 2003), pp. 213–222.
- [143] D. Lykov, J. Wurtz, C. Poole, M. Saffman, T. Noel, and Y. Alexeev, Sampling frequency thresholds for the quantum advantage of the quantum approximate optimization algorithm, *npj Quantum Inf.* **9**, 73 (2023).
- [144] S. Boettcher and A. G. Percus, Extremal optimization for graph partitioning, *Phys. Rev. E* **64**, 026114 (2001).
- [145] E. Farhi, D. Gosset, I. Hen, A. W. Sandvik, P. Shor, A. P. Young, and F. Zamponi, Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs, *Phys. Rev. A* **86**, 052334 (2012).
- [146] C.-W. Liu, A. Polkovnikov, and A. W. Sandvik, Quantum versus classical annealing: Insights from scaling theory and results for spin glasses on 3-regular graphs, *Phys. Rev. Lett.* **114**, 147203 (2015).
- [147] H. Goto, Bifurcation-based adiabatic quantum computation with a nonlinear oscillator network, *Sci. Rep.* **6**, 21686 (2016).
- [148] H. Goto, K. Endo, M. Suzuki, Y. Sakai, T. Kanao, Y. Hamakawa, R. Hidaka, M. Yamasaki, and K. Tatsumura, High-performance combinatorial optimization based on classical mechanics, *Sci. Adv.* **7**, eaab7953 (2021).
- [149] H. Goto, K. Tatsumura, and A. R. Dixon, Combinatorial optimization by simulating adiabatic bifurcations in nonlinear Hamiltonian systems, *Sci. Adv.* **5**, eaav2372 (2019).
- [150] Q.-G. Zeng, X.-P. Cui, B. Liu, Y. Wang, P. Mosharev, and M.-H. Yung, Performance of quantum annealing inspired algorithms for combinatorial optimization problems, *Commun. Phys.* **7**, 249 (2024).
- [151] M. J. A. Schuetz, J. K. Brubaker, and H. G. Katzgraber, Combinatorial optimization with physics-inspired graph neural networks, *Nat. Mach. Intell.* **4**, 367 (2022).
- [152] W. Yao, A. S. Bandeira, and S. Villar, in *Wavelets and Sparsity XVIII*, Vol. 11138, edited by D. V. D. Ville, M. Papadakis, and Y. M. Lu, International Society for Optics and Photonics (SPIE, San Diego, California, USA, 2019), p. 111380S.
- [153] M. C. Angelini and F. Ricci-Tersenghi, Modern graph neural networks do worse than classical greedy algorithms in solving combinatorial optimization problems like maximum independent set, *Nat. Mach. Intell.* **5**, 29 (2023).
- [154] S. Boettcher, Inability of a graph neural network heuristic to outperform greedy algorithms in solving combinatorial optimization problems, *Nat. Mach. Intell.* **5**, 24 (2023).
- [155] M. J. A. Schuetz, J. K. Brubaker, and H. G. Katzgraber, Reply to: Modern graph neural networks do worse than classical greedy algorithms in solving combinatorial optimization problems like maximum independent set, *Nat. Mach. Intell.* **5**, 32 (2023).
- [156] M. Aramon, G. Rosenberg, E. Valiante, T. Miyazawa, H. Tamura, and H. G. Katzgraber, Physics-inspired optimization for quadratic unconstrained problems using a digital annealer, *Front. Phys.* **7**, 48 (2019).
- [157] A. Montanari, in *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Baltimore, Maryland, 2019), pp. 1417–1433.

- [158] F. B. Maciejewski, J. Biamonte, S. Hadfield, and D. Venturelli, Improving quantum approximate optimization by noise-directed adaptive remapping, [arXiv:2404.01412](#).
- [159] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, England, 2009).
- [160] S. Aref, A. J. Mason, and M. C. Wilson, A modeling and computational study of the frustration index in signed networks, [Networks](#) **75**, 95 (2020).
- [161] Y. R. Pei, H. Manukian, and M. D. Ventra, Generating weighted MAX-2-SAT instances with frustrated loops: An RBM Case study, *J. Mach. Learn. Res.* **21**, 1 (2020).
- [162] S. F. Edwards and P. W. Anderson, Theory of spin glasses, [J. Phys. F: Met. Phys.](#) **5**, 965 (1975).
- [163] S. F. Edwards and P. W. Anderson, Theory of spin glasses. II, [J. Phys. F: Met. Phys.](#) **6**, 1927 (1976).
- [164] G. Parisi, Order parameter for spin-glasses, [Phys. Rev. Lett.](#) **50**, 1946 (1983).
- [165] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, [Phys. Rev. A](#) **77**, 012307 (2008).
- [166] D-Wave, dwave-samplers (2024).
- [167] I. Waldspurger and A. Waters, Rank optimality for the Burer–Monteiro factorization, [SIAM J. Optim.](#) **30**, 2577 (2020).
- [168] N. Boumal, V. Voroninski, and A. S. Bandeira, Deterministic guarantees for Burer–Monteiro factorizations of smooth semidefinite programs, [Commun. Pure Appl. Math.](#) **73**, 581 (2020).
- [169] J. Gray, quimb: A Python library for quantum information and many-body calculations, [J. Open Source Softw.](#) **3**, 819 (2018).