

Expectation Values from the Single-Layer Quantum Approximate Optimization Algorithm on Ising Problems

Asier Ozaeta,^{1,*} Wim van Dam,^{1,2,3,†} and Peter L. McMahon^{4,1,‡}

¹*QC Ware Corp., 550 Hamilton Ave, Palo Alto, CA 94301, USA*

²*Department of Computer Science, University of California, Santa Barbara, CA 93106, USA*

³*Department of Physics, University of California, Santa Barbara, CA 93106, USA*

⁴*School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA*

(Dated: December 8, 2020)

We report on the energy-expectation-value landscapes produced by the single-layer ($p = 1$) Quantum Approximate Optimization Algorithm (QAOA) when being used to solve Ising problems. The landscapes are obtained using an analytical formula that we derive. The formula allows us to predict the landscape for any given Ising problem instance and consequently predict the optimal QAOA parameters for heuristically solving that instance using the single-layer QAOA. We have validated our analytical formula by showing that it accurately reproduces the landscapes published in recent experimental reports. We then applied our methods to address the question: *how well is the single-layer QAOA able to solve large benchmark problem instances?* We used our analytical formula to calculate the optimal energy-expectation values for benchmark MAX-CUT problems containing up to 7 000 vertices and 41 459 edges. We also calculated the optimal energy expectations for general Ising problems with up to 100 000 vertices and 150 000 edges. Our results provide an estimate for how well the single-layer QAOA may work when run on a quantum computer with thousands of qubits. In addition to providing performance estimates when optimal angles are used, we are able to use our analytical results to investigate the difficulties one may encounter when running the QAOA in practice for different classes of Ising instances. We find that depending on the parameters of the Ising Hamiltonian, the expectation-value landscapes can be rather complex, with sharp features that necessitate highly accurate rotation gates in order for the QAOA to be run optimally on quantum hardware. We also present analytical results that explain some of the qualitative landscape features that are observed numerically.

CONTENTS

I. Introduction	2
II. Analytical Formulae for Expected Energies	2
A. Formula for General Case with Variable Couplings	4
B. Formulae for Ising Problems with Constant Couplings	4
III. Numerical Results and discussion	5
A. Numerical Results on Google MAX-CUT and Sherrington-Kirkpatrick (SK) Instances	5
B. Numerical Results on G-Set MAX-CUT Instances and Ising Instances with External Fields	9
C. Numerical and Analytical Results on Ising Problem Expectation Value Landscapes	12
IV. Conclusions	14
V. Acknowledgements	14
References	14
A. Derivation of Analytical Expression for General Expectation	15
B. Estimation of Optimal Energies	18
C. Normalized Expectation-Value Landscapes	18

* asier.ozaeta@qware.com

† wim.van.dam@qware.com

‡ pmcmahon@cornell.edu

I. INTRODUCTION

The Quantum Approximate Optimization Algorithm (QAOA) [1] is a variational quantum algorithm [2] that is able to find approximate solutions to combinatorial-optimization problems [3], and can also be used to heuristically find optimal solutions [4].

The basic structure of the QAOA, and some precursor quantum-optimization algorithms [5, 6], is the application of an alternating series of unitary operations. These unitary operations correspond to applications of the cost function (cost Hamiltonian) that one is attempting to optimize for, and of a *mixer* or *driver* Hamiltonian. The application of these unitaries is intended to evolve the state of the quantum computer such that it has substantial overlap with computational-basis states representing low-energy configurations of the variables being optimized over. In the QAOA, each unitary is parameterized by a single real number (often referred to in the QAOA literature as an *angle*), and these parameters are chosen to optimize the quality of the solutions output by the algorithm. A crucial question is how the choice of the QAOA parameters affects the solutions the QAOA is able to find, and hence how to best pick the parameters to be used when running the QAOA.

In this paper we will discuss the QAOA in the setting of solving (classical) Ising optimization problems, i.e., where the optimization problem's cost function takes the form $\sum_{(i,j) \in E} J_{ij} s_i s_j + \sum_{i \in V} h_i s_i$, and $s_i \in \{-1, +1\}$. We note that MAX-CUT problems are a subclass of Ising-optimization problems, for which the external-field terms h_i are zero and the couplings J_{ij} are either 0 or -1 . Our results are however applicable to the solution of any NP-hard problem by the QAOA, via an appropriate mapping [7].

The QAOA has recently been experimentally demonstrated with up to 23 qubits on Ising-optimization problems with no external-field terms [8]. The experimental results in Ref. [8] include optimization landscapes for the case where the QAOA has only a single *layer* (a single application of the cost-function Hamiltonian, followed by a single application of the mixer Hamiltonian). These landscapes are plots of the expected value of the cost-function Hamiltonian as a function of the two QAOA parameters. In our work we present an analytical formula that is able to compute the energy landscape of the single-layer QAOA for any Ising-problem instance. We compare our analytical results with the numerical results from Ref. [8] (which include both numerical results from classical simulations of the QAOA, and measured data from running the QAOA on a quantum processor), and find excellent agreement—showing that our analytical formula is able to predict the energy landscape. Our analytical formula gives us the ability to analytically optimize the QAOA parameters—this enables one to compute the optimal QAOA parameters classically, and then only run the QAOA on a quantum processor with the optimal set of angles [9]. This yields a large speedup versus the current standard mode of operating the single-layer QAOA, where the parameters are optimized numerically via an iterative process or via performing a brute-force search of the landscape.

In addition to replicating the landscapes for the three problem instances given in Ref. [8], we also explore the landscapes our analytical formula obtains for a range of other Ising problems, some of which we discover have qualitatively different features than the landscapes published in the QAOA literature so far.

One of the key benefits of having an analytical formula is that we can study the performance of the QAOA on problems using far more qubits than we can classically simulate or that we have experimental hardware containing. We use our analytical formula to predict the performance of the QAOA on the G-set MAX-CUT benchmark instances [10], which includes problems having up to 7 000 vertices (qubits), as well as on Ising instances with up to 100 000 vertices. These results give us an indication of how well the QAOA may perform on problem instances with sizes far beyond the ~ 50 spins that is the current limit for direct testing of the QAOA.

There has been previous work [11] on studying the QAOA parameters and energy-expectation landscapes analytically in the single-layer setting. Ref. [11] derives a formula for the energy expectation as a function of the QAOA parameters for the specific case of MAX-CUT on unweighted graphs, and Ref. [12] generalized this formula for the energy expectation for the QAOA to the case of weighted graphs. Our work builds on this: we include in our analysis the case of non-zero external-field terms in the Ising Hamiltonian. This allows us to assess the energy expectation for the single-layer QAOA solving any Ising problem.

II. ANALYTICAL FORMULAE FOR EXPECTED ENERGIES

Given an undirected graph $G = (V, E)$ defined by vertices $V = \{1, \dots, n\}$ and edges $E \subset V \times V$, as well an external field h_i associated with each vertex, and a coupling strength (edge weight) J_{ij} associated with each edge, we aim to find the configuration $s \in \{-1, +1\}^n$ that minimizes the (Ising) cost function

$$C = \sum_{i \in V} h_i s_i + \sum_{(i,j) \in E} J_{ij} s_i s_j. \quad (1)$$

This cost function is equivalent to the cost Hamiltonian

$$H_C = \sum_{i \in V} C_i + \sum_{(i,j) \in E} C_{ij} \quad (2)$$

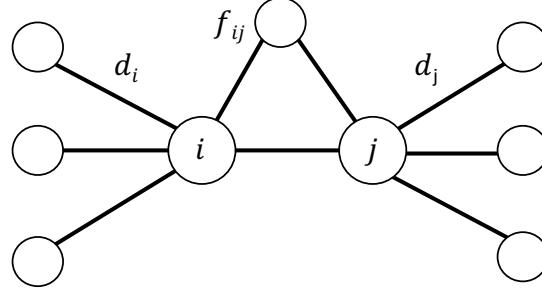


FIG. 1. The adjacent vertices of vertex i are denoted by $A_i = \{k \in V \mid (i, k) \in E\}$ and similarly for A_j . The number d_i of vertices connected to i is thus $|A_i \setminus A_j| = 4$, and for j we have $d_j = |A_j \setminus A_i| = 4$. The number f_{ij} of shared vertices between i and j equals $f_{ij} = |A_i \cap A_j| = 1$.

where

$$\begin{cases} C_i = h_i \sigma_i^z \\ C_{ij} = J_{ij} \sigma_i^z \sigma_j^z \end{cases}. \quad (3)$$

The unitary transformation U for a QAOA circuit with depth $p = 1$ is defined as

$$U = U_B^\beta U_C^\gamma \quad (4)$$

where

$$U_C^\gamma = e^{-i\gamma H_C} = e^{-i\gamma(\sum_{ij} J_{ij} \sigma_i^z \sigma_j^z + \sum_i h_i \sigma_i^z)} \quad (5)$$

$$= \prod_{(i,j) \in E} (\cos(J_{ij}\gamma) \sigma_i^0 \sigma_j^0 - i \sin(J_{ij}\gamma) \sigma_i^z \sigma_j^z) \prod_{i \in V} (\cos(h_i\gamma) \sigma_i^0 - i \sin(h_i\gamma) \sigma_i^z) \quad (6)$$

$$U_B^\beta = e^{-i\beta H_B} = e^{-i\beta \sum_i \sigma_i^x} \quad (7)$$

$$= \prod_{i \in V} (\cos(\beta) \sigma_i^0 - i \sin(\beta) \sigma_i^x). \quad (8)$$

Applying $U = U_B^\beta U_C^\gamma$ to an initial, uniform superposition of all bit strings, $|\psi_0\rangle$, we obtain the final QAOA state

$$|\beta, \gamma\rangle = U |\psi_0\rangle. \quad (9)$$

The expectation value of H_C for this final state is

$$F(\beta, \gamma) := \sum_{i \in V} \langle C_i \rangle + \sum_{(i,j) \in E} \langle C_{ij} \rangle \quad (10)$$

where

$$\begin{cases} \langle C_i \rangle = \langle \beta, \gamma | C_i | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_i U | \psi_0 \rangle = h_i \text{Tr}[\rho_0 U^\dagger \sigma_i^z U] \\ \langle C_{ij} \rangle = \langle \beta, \gamma | C_{ij} | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_{ij} U | \psi_0 \rangle = J_{ij} \text{Tr}[\rho_0 U^\dagger \sigma_i^z \sigma_j^z U] \end{cases} \quad (11)$$

and $\rho_0 = |\psi_0\rangle \langle \psi_0|$ is the initial density matrix.

A. Formula for General Case with Variable Couplings

In Appendix A we derive explicit expressions for the two traces of Equation 11, giving us the results

$$\left\{ \begin{array}{l} \langle C_i \rangle = h_i \sin(2\beta) \sin(2\gamma h_i) \prod_{(i,k) \in E} \cos(2\gamma J_{ik}) \\ \langle C_{ij} \rangle = J_{ij} \sin(2\beta) \cos(2\beta) \sin(2\gamma J_{ij}) \left[\cos(2\gamma h_i) \prod_{\substack{(ik) \in E \\ k \neq j}} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{\substack{(jk) \in E \\ k \neq i}} \cos(2\gamma J_{jk}) \right] \\ - \frac{J_{ij}}{2} (\sin(2\beta))^2 \prod_{\substack{(ik) \in E \\ (jk) \notin E}} \cos(2\gamma J_{ik}) \prod_{\substack{(jk) \in E \\ (ik) \notin E}} \cos(2\gamma J_{jk}) \\ \times \left[\cos(2\gamma(h_i + h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} + J_{jk})) - \cos(2\gamma(h_i - h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} - J_{jk})) \right]. \end{array} \right. \quad (12)$$

Equation 12 is the contribution of a single vertex and connected subgraphs in the graph. Connected subgraphs are formed by two vertices (i, j) connected by an edge and the nearest neighbours of the i and j vertices, as shown in Fig. 1. From this general formula we will derive the expressions corresponding to simpler problem instances and graphs, including previously derived results.

Without loss of generality we can assume the graph to be complete as we can use $J_{ij} = 0$ for absent edges. Doing so simplifies the above expression to

$$\begin{aligned} \langle C_{ij} \rangle_K &= J_{ij} \sin(2\beta) \cos(2\beta) \sin(2\gamma J_{ij}) \left[\cos(2\gamma h_i) \prod_{k \neq i,j} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{k \neq i,j} \cos(2\gamma J_{jk}) \right] \\ &- \frac{J_{ij}}{2} (\sin(2\beta))^2 \left[\cos(2\gamma(h_i + h_j)) \prod_{k \neq i,j} \cos(2\gamma(J_{ik} + J_{jk})) - \cos(2\gamma(h_i - h_j)) \prod_{k \neq i,j} \cos(2\gamma(J_{ik} - J_{jk})) \right]. \end{aligned} \quad (13)$$

If, additionally, one assumes the absence of local fields, i.e., $h_i = 0$ and $h_j = 0$, this further simplifies to

$$\begin{aligned} \langle C_{ij} \rangle_{K,h=0} &= J_{ij} \sin(2\beta) \cos(2\beta) \sin(2\gamma J_{ij}) \left[\prod_{k \neq i,j} \cos(2\gamma J_{ik}) + \prod_{k \neq i,j} \cos(2\gamma J_{jk}) \right] \\ &- \frac{J_{ij}}{2} (\sin(2\beta))^2 \left[\prod_{k \neq i,j} \cos(2\gamma(J_{ik} + J_{jk})) - \prod_{k \neq i,j} \cos(2\gamma(J_{ik} - J_{jk})) \right]. \end{aligned} \quad (14)$$

This last expression has been derived previously and appears in Lemma C.1 of Ref. [12].

Returning to the general Equation 12, we can also consider the case of triangle-free graphs, i.e., graphs where i and j share no common vertices k with both $(i, k) \in E$ and $(j, k) \in E$, for which the expectation reduces to

$$\langle C_{ij} \rangle_{\text{tf}} = J_{ij} \sin(2\beta) \cos(2\beta) \sin(2\gamma J_{ij}) \left[\cos(2\gamma h_i) \prod_{(ik) \in E} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{(jl) \in E} \cos(2\gamma J_{jl}) \right] \quad (15)$$

$$+ J_{ij} (\sin(2\beta))^2 \sin(2\gamma h_i) \sin(2\gamma h_j) \prod_{\substack{(ik) \in E \\ (jk) \notin E}} \cos(2\gamma J_{ik}) \prod_{\substack{(jk) \in E \\ (ik) \notin E}} \cos(2\gamma J_{jk}). \quad (16)$$

B. Formulae for Ising Problems with Constant Couplings

Next, we will derive the expressions of the expectation value for some special-case Hamiltonians where the couplings are constant ($J_{ij} = 1$ or $J_{ij} = -1$). For the NP-hard ‘‘P5’’ Ising problem of Ref. [13] for a graph (V, E) the goal is to minimize the Hamiltonian

$$H_{\text{P5}} = \sum_{(i,j) \in E} \sigma_i^z \sigma_j^z + \sum_{i \in V} \sigma_i^z \quad (17)$$

With d_i and d_j the degrees of i and j , and f_{ij} the number of shared vertices between i and j , it is easy to see that for the expectation we now have

$$\begin{cases} \langle C_i \rangle_{\text{P5}} = \sin(2\beta) \sin(2\gamma) (\cos(2\gamma))^{d_i} \\ \langle C_{ij} \rangle_{\text{P5}} = \sin(2\beta) \cos(2\beta) \sin(2\gamma) \left[(\cos(2\gamma))^{d_i} + (\cos(2\gamma))^{d_j} \right] \\ \quad + \frac{1}{2} (\sin(2\beta))^2 (\cos(2\gamma))^{d_i+d_j-2f_{ij}-2} (1 - (\cos(4\gamma))^{f_{ij}+1}). \end{cases} \quad (18)$$

For the MAX-CUT problem, with no local fields and couplings $J_{ij} = -1$, the expectation reads

$$\begin{cases} H_{\text{MC}} = - \sum_{(i,j) \in E} \sigma_i^z \sigma_j^z \\ \langle C_i \rangle_{\text{MC}} = 0 \\ \langle C_{ij} \rangle_{\text{MC}} = \sin(2\beta) \cos(2\beta) \sin(2\gamma) [(\cos(2\gamma))^{d_i-1} + (\cos(2\gamma))^{d_j-1}] \\ \quad - \frac{1}{2} (\sin(2\beta))^2 (\cos(2\gamma))^{d_i+d_j-2f_{ij}-2} (1 - (\cos(4\gamma))^{f_{ij}}). \end{cases} \quad (19)$$

This is the result obtained by Wang et al. [14].

The expression in the absence of spin-spin couplings reads,

$$H_{\text{NC}} = \sum_{i \in V} h_i \sigma_i^z \quad (20)$$

$$\langle C_i \rangle_{\text{NC}} = h_i \sin(2\beta) \sin(2\gamma h_i) \quad (21)$$

$$\langle C_{ij} \rangle_{\text{NC}} = 0. \quad (22)$$

The solution to this optimization problem is trivial, however, the determination of the angles for the optimal expectation value is not.

We will use the analytical formulae derived in this section to gain insight into the behavior of expectation value landscapes for QAOA $p = 1$. First, we will study MAX-CUT instances, as the paper progresses we will gradually increase the complexity of the instances we study until we reach general Ising problems.

III. NUMERICAL RESULTS AND DISCUSSION

A. Numerical Results on Google MAX-CUT and Sherrington-Kirkpatrick (SK) Instances

Eq. 12 allows us to replicate the expectation-value landscapes of the recent experimental work of Ref. [8], which provides numerical confirmation of the validity of our analytical results. In this section, we present our investigation of the problems studied in Ref. [8], as well as some similar but larger problem instances.

The problems studied in Ref. [8] were weighted MAX-CUT instances with weights -1 and 1 assigned at random; we obtained the three specific instances used in Ref. [8] from the *ReCirq* repository [16]. We numerically evaluated Eq. 12 for these three problem instances. This gave us the energy-expectation landscapes for the single-layer QAOA for these instances, i.e., the numerical evaluation of the function $F(\beta, \gamma)$ for varying values of both β and γ . We have plotted these energy-expectation landscapes in Fig. 2. We note that these landscapes, which we obtained using the analytical formula Eq. 12, are in excellent agreement with the brute-force-classical and quantum-hardware evaluations of the landscapes shown in Ref. [8]. Observe, also, that we plot $-F(\beta, \gamma)$ due to the normalization of the previous evaluations.

The analytical formula Eq. 12 allows us to explore the landscapes of problem instances far larger than those possible to study on current experimental hardware. We have computed the landscapes for two 100-spin (i.e., 100-qubit) Ising problem instances: one instance has a 3-regular graph structure, and the other is a Sherrington-Kirkpatrick instance. The graphs and their corresponding landscapes are shown in Fig. 3.

We will now comment briefly on the features that one can observe in the landscapes shown in Fig. 2 and Fig. 3. In Fig. 2, the peaks are clearly narrower and closer to the edge for the 11-spin SK instance than for the two (23-spin and 14-spin) sparsely-connected instances. Furthermore, for the SK instance, additional peaks appear at $\gamma = \pi/4$.

In Fig. 3, we see that the landscape for the sparsely-connected (3-regular) instance with 100 spins is the same as the landscape for the 3-regular instance with 14 spins—i.e., as the number of spins has increased, the landscape has not changed. In contrast, the landscape for the 100-spin SK instance is qualitatively different than the landscape for the 11-spin SK instance: the peaks

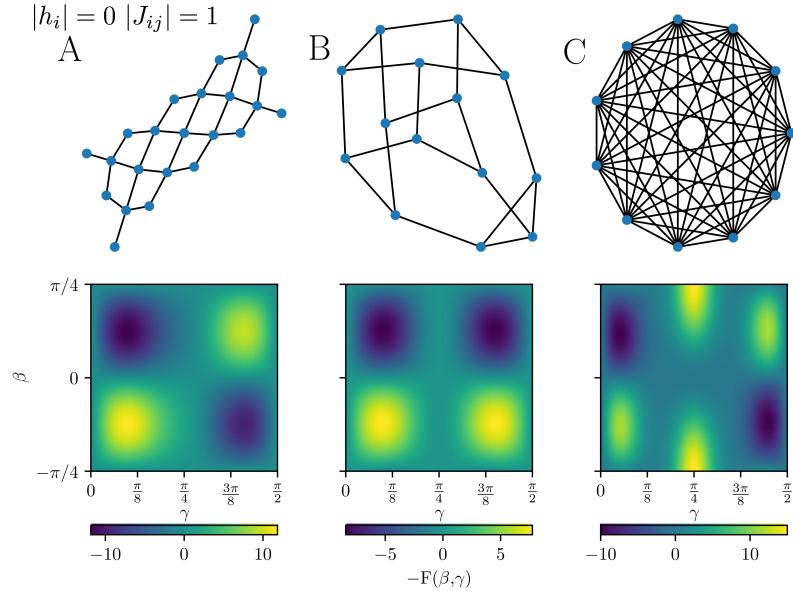


FIG. 2. Negative expectation-value landscapes, obtained using our analytical formulae, Eq. 12, for the same problem instances studied experimentally in Ref. [8]. The graphs for each instance are shown above their respective landscape. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 14 qubits. **C:** Fully connected SK-model with 11 qubits.

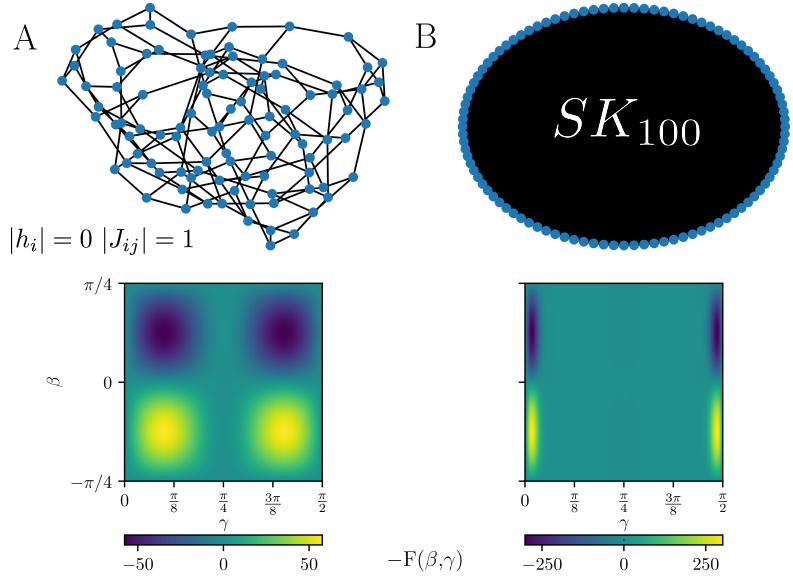


FIG. 3. Negative expectation-value landscapes for two 100-qubit instances. **A:** 3-regular graph with 100 qubits. **B:** Fully connected SK-model with 100 qubits.

in the landscape for the 100-spin instance are narrower and closer to the edges of the γ range, and the peaks at $\gamma = \pi/4$ for the 11-spin instance are not present for the 100-spin instance.

An SK instance with n spins is an n -regular graph. We can understand the narrowing and shifting of the peaks between the $n = 11$ and $n = 100$ SK instances we studied as an example of a more general phenomenon for regular graphs: the larger the degree of a regular graph, the narrower the peaks in the landscape will be. This is explored in Sec. III C. This phenomenon has an important experimental consequence: the peak sharpness places a constraint on how accurately the angle γ needs to be implemented in hardware, and so the necessary precision for specifying γ will increase with the *degree* of a regular graph.

While the expectation-value landscapes presented in Fig. 2 already show some qualitative differences from instance to instance, we will show later in this paper that landscapes can vary far more dramatically depending on the parameters of the cost

Name	$ V $	$ E $	Sum of weights	Best-known Ising	QAOA expectation	Ratio Ising	$1.18\sqrt{ E \times V }$	Ratio exp.	Weights
G1	800	19 176	19 176	4072	1898.955	0.466	4621.745	0.411	{0, 1}
G2	800	19 176	19 176	4064	1898.98	0.467	4621.745	0.411	{0, 1}
G3	800	19 176	19 176	4068	1901.214	0.467	4621.745	0.411	{0, 1}
G4	800	19 176	19 176	4116	1897.055	0.461	4621.745	0.41	{0, 1}
G5	800	19 176	19 176	4086	1899.615	0.465	4621.745	0.411	{0, 1}
G6	800	19 176	154	4202	1675.603	0.399	4621.745	0.363	{-1, 0, +1}
G7	800	19 176	-150	4162	1679.381	0.404	4621.745	0.363	{-1, 0, +1}
G8	800	19 176	-170	4180	1676.031	0.401	4621.745	0.363	{-1, 0, +1}
G9	800	19 176	-64	4172	1676.523	0.402	4621.745	0.363	{-1, 0, +1}
G10	800	19 176	-160	4160	1677.502	0.403	4621.745	0.363	{-1, 0, +1}
G11	800	1600	34	1094	519.61	0.475	1335.018	0.389	{-1, 0, +1}
G12	800	1600	-4	1116	519.61	0.466	1335.018	0.389	{-1, 0, +1}
G13	800	1600	34	1130	519.61	0.46	1335.018	0.389	{-1, 0, +1}
G14	800	4694	4694	1434	888.003	0.619	2286.644	0.388	{0, 1}
G15	800	4661	4661	1439	877.073	0.61	2278.592	0.385	{0, 1}
G16	800	4672	4672	1432	878.629	0.614	2281.279	0.385	{0, 1}
G17	800	4667	4667	1427	878.527	0.616	2280.058	0.385	{0, 1}
G18	800	4694	64	1920	718.027	0.374	2286.644	0.314	{-1, 0, +1}
G19	800	4661	-113	1925	709.963	0.369	2278.592	0.312	{-1, 0, +1}
G20	800	4672	-46	1928	711.113	0.369	2281.279	0.312	{-1, 0, +1}
G21	800	4667	-67	1929	708.482	0.367	2280.058	0.311	{-1, 0, +1}
G27	2000	19 990	-42	6724	2691.716	0.4	7461.109	0.361	{-1, 0, +1}
G28	2000	19 990	-104	6700	2696.452	0.402	7461.109	0.361	{-1, 0, +1}
G29	2000	19 990	80	6730	2692.164	0.4	7461.109	0.361	{-1, 0, +1}
G30	2000	19 990	100	6726	2694.92	0.401	7461.109	0.361	{-1, 0, +1}
G31	2000	19 990	-80	6700	2693.874	0.402	7461.109	0.361	{-1, 0, +1}
G59	5000	29 570	96	12 076	4450.69	0.369	14 348.043	0.31	{-1, 0, +1}
G61	7000	17 148	362	11 230	4590.256	0.409	12 928.191	0.355	{-1, 0, +1}
G64	7000	41 459	527	16 975	6232.192	0.367	20 102.054	0.31	{-1, 0, +1}

TABLE I. Computational results of QAOA $p = 1$ expectation values in Ising energies for a subset of the G-set instances. Column *Sum of weights* shows the sum of all J_{ij} values. Column *Best-known Ising* shows the best classical results for the Ising energy, column *QAOA expectation* gives the best QAOA expectation value. The best-known classical results were obtained from Ref. [15]. The columns *Ratio Ising* and *Ratio exp.* show the corresponding ratio of the QAOA value versus the best classical result and versus the analytical formula value, respectively.

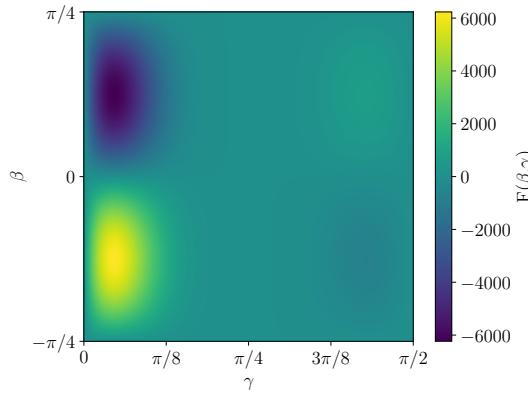


FIG. 4. Expectation-value landscape for the $G64$ instance with 7000 vertices and 41459 edges, obtained using the analytical formula Eq. 12.

Hamiltonian. Furthermore, we will study how the optimal γ value, γ_{\min} , that minimizes the expectation value, changes with those parameters. So far we have seen evidence that increased degree of regular graphs can cause γ_{\min} to be closer to 0 or $\pi/2$, but we will find that the value of γ_{\min} can have a rather complex dependence on the cost-Hamiltonian parameters.

One of the potential benefits of using an analytical formula to determine the optimal angles instead of performing a brute-force search is speed, since the former requires on classical-computing calculations whereas the latter uses quantum simulations or runs on quantum hardware. This potential benefit is seen already with the instances studied in this section. To illustrate this,

Name	$ V $	$ E $	Sum of weights	Best-known cut	QAOA expectation cut	Ratio cut	Time (s)	Weights
G1	800	19 176	19 176	11 624	10 537.478	0.907	28 322	{0, 1}
G2	800	19 176	19 176	11 620	10 537.49	0.907	28 562	{0, 1}
G3	800	19 176	19 176	11 622	10 538.607	0.907	28 332	{0, 1}
G4	800	19 176	19 176	11 646	10 536.528	0.905	28 430	{0, 1}
G5	800	19 176	19 176	11 631	10 537.807	0.906	29 039	{0, 1}
G6	800	19 176	154	2178	914.801	0.42	31 953	{-1, 0, +1}
G7	800	19 176	-150	2006	764.69	0.381	32 223	{-1, 0, +1}
G8	800	19 176	-170	2005	753.016	0.376	32 685	{-1, 0, +1}
G9	800	19 176	-64	2054	806.262	0.393	33 056	{-1, 0, +1}
G10	800	19 176	-160	2000	758.751	0.379	33 681	{-1, 0, +1}
G11	800	1600	34	564	276.805	0.491	1193	{-1, 0, +1}
G12	800	1600	-4	556	257.805	0.464	1188	{-1, 0, +1}
G13	800	1600	34	582	276.805	0.476	1194	{-1, 0, +1}
G14	800	4694	4694	3064	2791.002	0.911	4808	{0, 1}
G15	800	4661	4661	3050	2769.037	0.908	4982	{0, 1}
G16	800	4672	4672	3052	2775.315	0.909	5076	{0, 1}
G17	800	4667	4667	3047	2772.764	0.91	5056	{0, 1}
G18	800	4694	64	992	391.014	0.394	5125	{-1, 0, +1}
G19	800	4661	-113	906	298.481	0.329	5187	{-1, 0, +1}
G20	800	4672	-46	941	332.556	0.353	5186	{-1, 0, +1}
G21	800	4667	-67	931	320.741	0.345	5237	{-1, 0, +1}
G27	2000	19 990	-42	3341	1324.858	0.397	22 103	{-1, 0, +1}
G28	2000	19 990	-104	3298	1296.226	0.393	21 916	{-1, 0, +1}
G29	2000	19 990	80	3405	1386.082	0.407	21 741	{-1, 0, +1}
G30	2000	19 990	100	3413	1397.46	0.409	22 292	{-1, 0, +1}
G31	2000	19 990	-80	3310	1306.937	0.395	22 460	{-1, 0, +1}
G59	5000	29 570	96	6086	2273.345	0.374	39 462	{-1, 0, +1}
G61	7000	17 148	362	5796	2476.128	0.427	14 794	{-1, 0, +1}
G64	7000	41 459	527	8751	3379.596	0.386	57 647	{-1, 0, +1}

TABLE II. Computational results of QAOA $p = 1$ expectation values in cut values for a subset of the G-set instances. Column *Sum of weights* shows the sum of all J_{ij} values. Column *Best-known cut* shows the best classical results for the cut value, column *QAOA expectation cut* gives the best QAOA expectation value. The best-known classical results were obtained from Ref. [15]. The column *Ratio cut* shows the corresponding ratio of the QAOA value versus the best classical result for the cut value. *Time (s)* column shows the time required to obtain the full landscape and optimal angles, in seconds, on a *ac5n.large* AWS instance.

Name	$ V $	$ E $	Sum of weights	QAOA expectation	$1.18\sqrt{ V \times (E + V)}$	Ratio exp.	Time (s)	Weights
I10	800	19 176	-144	1710.04	4717.167	0.363	31 376	{-1, 0, +1}
I20	800	4672	-128	789.409	2468.88	0.32	5269	{-1, 0, +1}
I30	2000	19 990	86	2821.631	7825.455	0.361	22 648	{-1, 0, +1}
I59	5000	29 570	156	4935.399	15 513.747	0.318	40 740	{-1, 0, +1}
I64	7000	41 459	661	6914.468	21 732.929	0.318	59 894	{-1, 0, +1}
I99	70 000	140 000	-62	54 813.021	143 067.397	0.383	125 363	{-1, 0, +1}
I100	100 000	150 000	956	72 928.233	186 574.382	0.391	137 774	{-1, 0, +1}

TABLE III. Computational results of QAOA $p = 1$ expectation values in Ising energies for a subset of the G-set instances with external fields, the I-set. Plus two more custom generated instances, *I99*, and *I100*. Column *Sum of weights* shows the sum of all h_i and J_{ij} values, column *QAOA expectation* gives the best QAOA expectation value. The column *Ratio exp.* corresponds to the ratio of the QAOA energy expectation value versus the optimal energy predicted by the formula $1.18\sqrt{|V| \times (|E| + |V|)}$. *Time (s)* column shows the time required to obtain the full landscape and optimal angles, in seconds, on an *ac5n.large* AWS instance.

we simulated the expectation-value landscape of fully connected graphs with 23 qubits using *Cirq* [17]; these calculations took approximately 1 min to complete. In comparison, using the analytical formula we could compute the landscape for 100-qubit problems in approximately 4 seconds. Both calculations were performed on a *c5n.large* AWS instance.

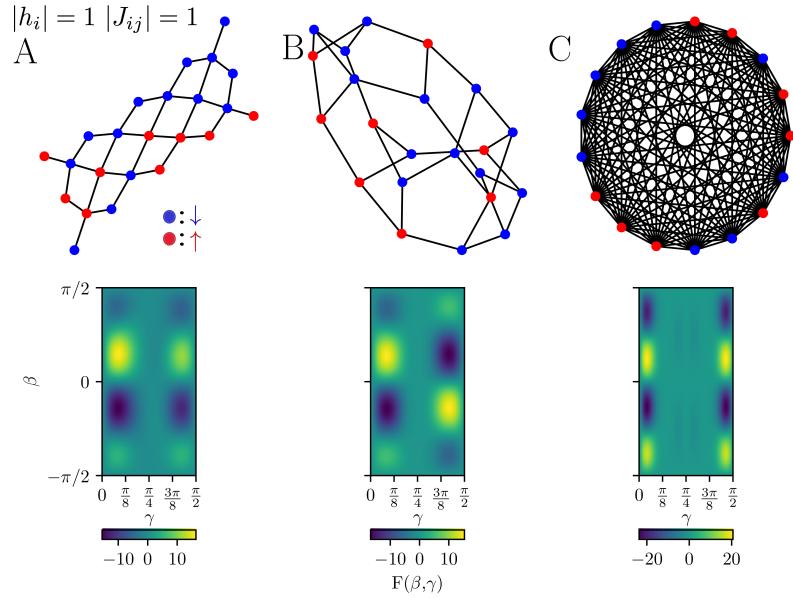


FIG. 5. Expectation-value landscapes obtained for problem instances with $h_i, J_{ij} \in \{-1, 1\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -1$, and red for those with $h_i = 1$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.

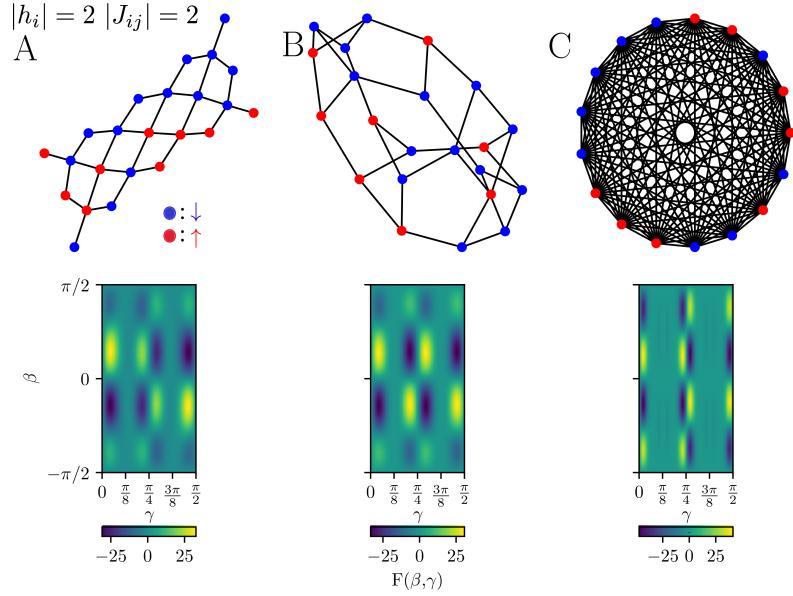


FIG. 6. Expectation-value landscapes obtained for problem instances with $h_i, J_{ij} \in \{-2, 2\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -2$, and red for those with $h_i = 2$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.

B. Numerical Results on G-Set MAX-CUT Instances and Ising Instances with External Fields

The G-set [10] collection of benchmark instances for MAX-CUT is widely used in the comparison of classical MAX-CUT heuristics [15, 18, 19]. We endeavored to predict how well the single-layer QAOA would perform on these instances, which affords a comparison between the single-layer QAOA running on a (future) quantum computer and the best-known classical methods for heuristically solving MAX-CUT problems.

We calculated the optimal QAOA expectation values for a subset of the G-set instances having 800, 2000, 5000, or 7000

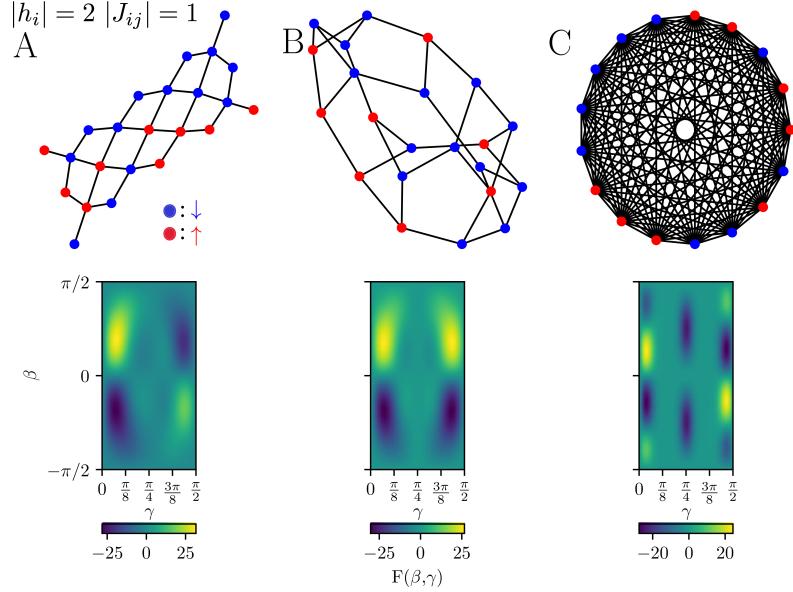


FIG. 7. Expectation-value landscapes obtained for problem instances with $h_i \in \{-2, 2\}$ and $J_{ij} \in \{-1, 1\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -2$, and red for those with $h_i = 2$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.

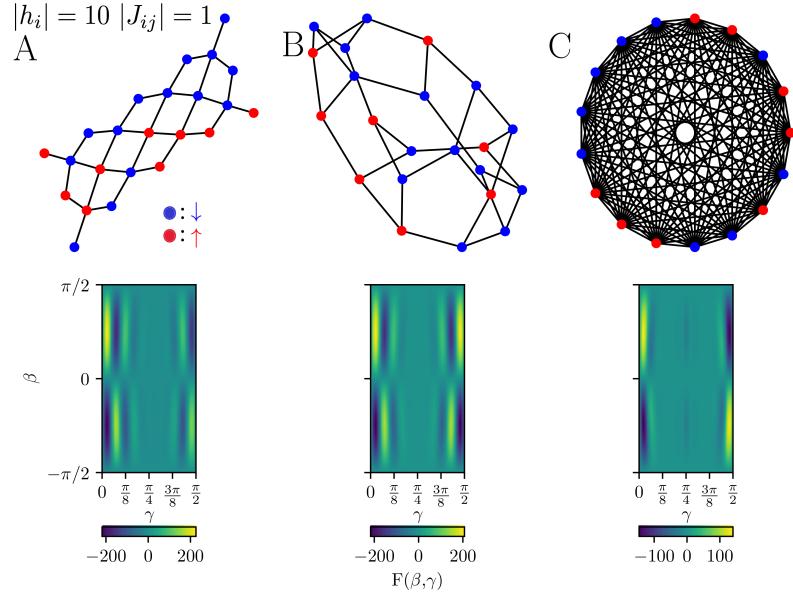


FIG. 8. Expectation-value landscapes obtained for problem instances with $h_i \in \{-10, 10\}$ and $J_{ij} \in \{-1, 1\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -10$, and red for those with $h_i = 10$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.

vertices and up to 41459 edges. The results are divided into two tables; Table I, where results are expressed in terms of Ising energies, and Table II, where we show them as cut values. The instances have weights that are drawn from two possible sets, depending on the instance: $\{0, 1\}$ and $\{-1, 0, +1\}$. For Table I, we derive in the appendix an analytical formula that allows us to estimate the optimal energy of a G-set instance, $E_{\text{opt}} \approx 1.18\sqrt{|E| \times |V|}$. The ratio between the theoretical maximum and the QAOA expectation, $r_{\text{exp.}} = F_{\text{max}}/E_{\text{opt}}$, ranges between 0.31–0.41. This ratio is similar to the formula of the approximation ratio, $r = F_{\text{max}}/C_{\text{max}}$, that is frequently used in the literature to determine the efficiency of the QAOA [1, 8, 11]. We just substitute the real maximum of the cost function by our theoretical approximation. The ratio for the best-known classical solution to the

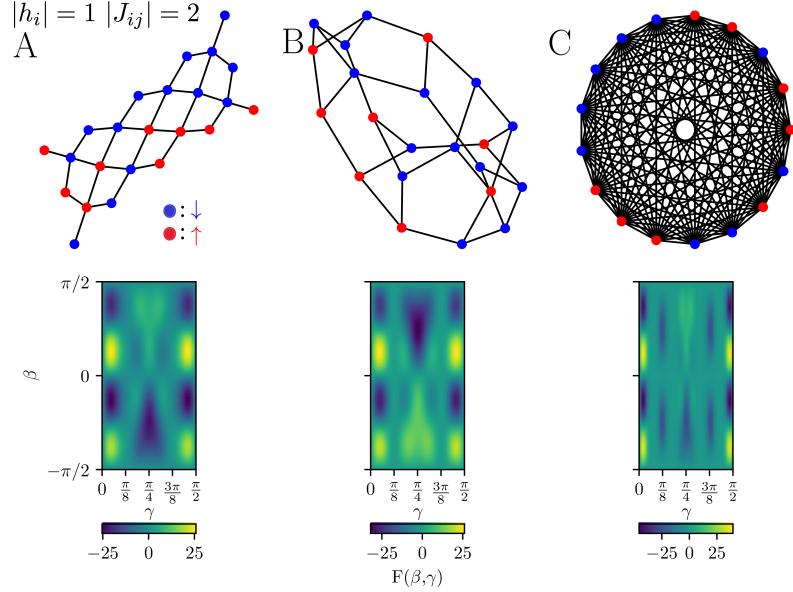


FIG. 9. Expectation-value landscapes obtained for problem instances with $h_i \in \{-1, 1\}$ and $J_{ij} \in \{-2, 2\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -1$, and red for those with $h_i = 1$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.

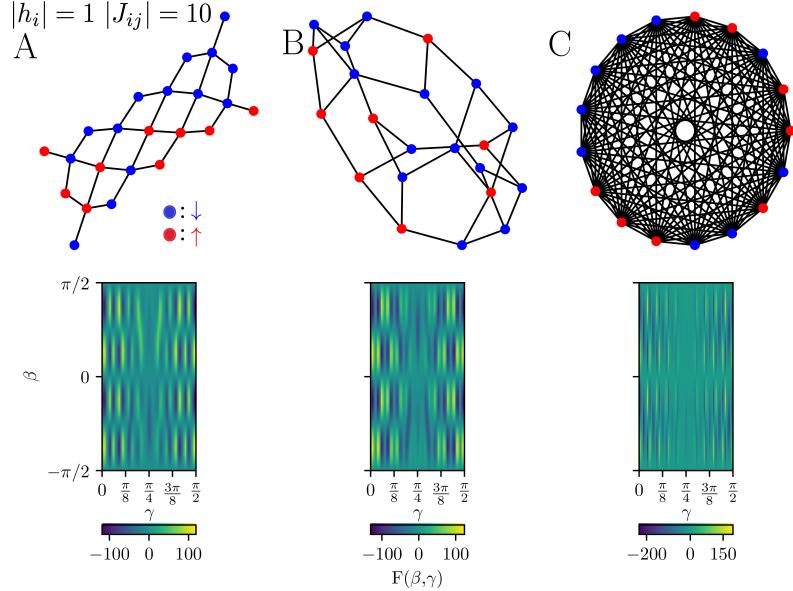


FIG. 10. Expectation-value landscapes obtained for problem instances with $h_i \in \{-1, 1\}$ and $J_{ij} \in \{-10, 10\}$, using the analytical formula Eq. 12. Problem graphs are shown above, with blue nodes representing those with $h_i = -1$, and red for those with $h_i = 1$. **A:** Fully connected SK-model with 17 qubits.

QAOA expectation, $r_{\text{Ising}} = F_{\max}/E_{\max}^{\text{cl}}$, ranges between $0.36 - 0.62$. Note how the theoretical maximum is close to the classical best-known solution for most instances except $G14 - G17$. For this set the *Ratio Ising* and *Ratio cut* are also particularly high. This leads us to believe that the best-known classical solution is not as good for these particular instances. This same logic applies to other instances with high values of *Ratio Ising* e.g. $G11$. Additionally, the QAOA expectation value is expected to be higher for instances with small amount of edges and common neighbours between vertices. These ratios can be used as a method to quickly determine the QAOA expectation of a given instance; one can use the analytical formula to determine the theoretical maximum, and use the ratio range to determine the QAOA expectation. The expectation value landscapes are very similar for the G-set instances, with similar optimal angles as predicted in Ref. [20]. Fig. 4 shows the landscape of instance $G64$.

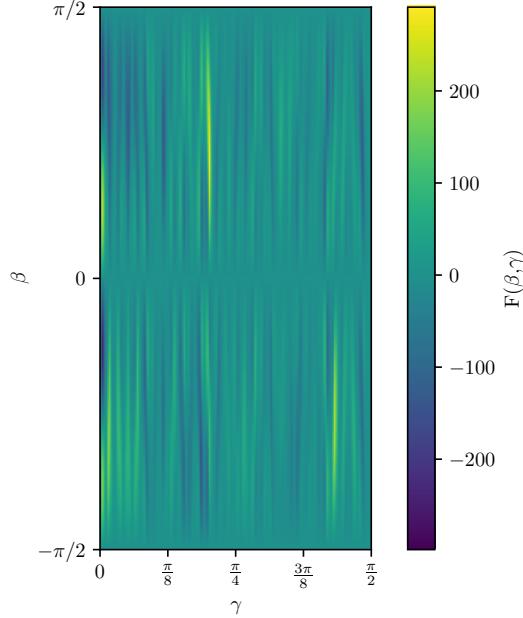


FIG. 11. Expectation-value landscape, obtained using the analytical formula Eq. 12, for a 3-regular graph problem instance that has 16 qubits, and $h_i \in \{-30, -29, \dots, 29, 30\}$ and $J_{ij} \in \{-30, 30\}$.

We also calculated the best QAOA expectation values for 7 Ising instances in Table III. These instances were generated by adding weight $\{-1, +1\}$ external fields at random to instances of the G-set. We denote these instances as Ix where x is the G instance label in which it is based. Additionally, two more instances, $I99$ and $I100$, were custom generated by assigning $\{-1, +1\}$ weights at random. The number of vertex and edges goes up to 100000 and 150000, respectively, in $I100$. An approximation for the optimal Ising energy of these instances was obtained using the analytical formula $E_{\text{opt}} \approx 1.18\sqrt{|V| \times (|E| + |V|)}$, that is derived in the appendix. We obtained a ratio $r_{\text{exp.}}$ that ranges between 0.32 – 0.39 when comparing the theoretical optima and the QAOA expectation. The values of the $r_{\text{exp.}}$ ratios are very similar to those obtained for MAX-CUT instances. The *Time* column displays in seconds the time required to obtain the full expectation value landscape and optimal angles.

C. Numerical and Analytical Results on Ising Problem Expectation Value Landscapes

In this section we will show the expectation-value landscapes of simple Ising problem instances with different weights, and derive analytic expressions that describe some of the observed phenomena.

First, we add external fields and study the case in which h_i and J_{ij} take only the values $\{-h, h\}$ assigned at random, where h can take any (real) value, and the spin connectivity is that of a d -regular graph. For this particular case we can approximate γ_{\min} as

$$\gamma_{\min} \approx \frac{1}{2h} \arctan \frac{1}{\sqrt{d}}, \frac{1}{2h} (\pi - \arctan \frac{1}{\sqrt{d}}). \quad (23)$$

We previously observed how the regularity of the graph affects the γ_{\min} value, and now we see that increasing the h value will also bring γ_{\min} closer to the edges ($\gamma = 0$ and $\gamma = \pi/2$). The h value also affects the γ -periodicity of the landscape: the period of the oscillations in the landscape as a function of γ is proportional to $1/h$. For example, peaks that for $h = 1$ were located at $\gamma = \pi/2$, for $h = 2$ will be located at $\gamma = \pi/4$. The expectation value of γ_{\min} for the single-node contribution in Eq. 12 reads,

$$\langle C_i \rangle(\beta_{\min}, \gamma_{\min}) = \frac{\sqrt{d}}{d+1} (1 + \frac{1}{d})^{-(d-1)/2}. \quad (24)$$

This value decreases monotonically with d . Here we operate under the assumption that the single-node term will be the main contribution to the expectation value. We observe the predicted phenomena in Fig. 5 and 6 where we plot the expectation landscapes for the cases $h = 1$ and $h = 2$, respectively. The first plot is, as expected, similar to what we encountered in the absence of external fields. Due to the presence of local minima and maxima with different expectation values, and additional

features, the β values of the plots now range between $-\pi/2$ and $\pi/2$. For $h = 2$ we observe that the periodicity in γ is divided by half, and there is a narrowing of the minimum and maximum peaks in γ . We had already seen (e.g., in Figures 2 and 3) that increasing number of edges per node causes the peaks in the landscape to become narrower in γ , and we now see that the peaks also become narrower when h is increased.

The difference in expectation values for different minima is going to arise from the summation terms in Eq. 12. For the general case these terms have the form

$$[\cos(2\gamma h)^o \cos(2\gamma J)^p + \cos(2\gamma h)^q \cos(2\gamma J)^r]. \quad (25)$$

If by our choice of o, p, q, r, h, J parameters we can make one term positive and another negative for the same value of γ , this will cause different expectation values for different minima.

Consider the case of problem instances with $h_i \in \{-rJ, rJ\}$ and $J_{ij} \in \{-J, J\}$ assigned at random, where the graph is again d -regular. In this case $J_{ij} \leq h_i$ for all i and j (since r is an integer). The following are two different approximations for γ_{\min} one can make depending on the value of r . For $r \leq 5$,

$$\gamma_{\min} \approx \frac{1}{2J} \arctan\left(\frac{1}{\sqrt{d}}(1 - 0.15(r - 1))\right). \quad (26)$$

For $r > 5$,

$$\gamma_{\min} \approx \frac{1}{2r^{0.6}J} \arctan\left(\frac{1}{\sqrt{d}}\right). \quad (27)$$

These equations show that the relative increase of h_i with regards to J_{ij} is going to make γ_{\min} closer to the edges. This phenomenon is observed in the $r = 2, 10$ cases shown in Fig. 7 and 8, respectively. The $r = 2$ case shows the peaks taking different shapes to what we previously observed. And for $r = 10$, the generation of additional peaks at different locations in the landscape, and the narrowing of those. By evaluating the expectation value in γ_{\min} we observe that it decreases monotonically with r . Thus, an increase in r will not only reduce γ_{\min} but it will generate additional peaks, the narrowing of those, a reduced optimal expectation value, and generate new shapes.

Alternatively, we can study the case $J_{ij} \geq h_i$, where we set $h_i \in \{-h, h\}$ and $J_{ij} \in \{-rh, rh\}$ assigned at random in a d regular graph. For this case we can approximate γ_{\min} ,

$$\gamma_{\min} \approx \frac{1}{2rh} \arctan\left(\frac{1}{\sqrt{d}}\right) \quad (28)$$

This equation shows how an increase in r will again lead to γ_{\min} being closer to the edge. The expectation value of γ_{\min} for the single node contribution Eq. 12 reads,

$$\langle C_i \rangle(\beta_{\min}, \gamma_{\min}) = \left(1 + \frac{1}{d}\right)^{-\frac{d}{2}} \sin\left(\frac{\arctan(1/\sqrt{d})}{r}\right). \quad (29)$$

This expression decreases monotonically with r , as well as d , leading to a worse expectation value with increasing r .

Fig. 9 and 10 show the $r = 2, 10$ cases respectively. The $r = 2$ case shows additional peaks, narrowing of those, and new shapes. While for $r = 10$ we have a complex landscape with a big number of narrow peaks with different shapes. Thus, an increase of r for these kind of problem instances will lead to a γ_{\min} closer to the edge, a reduced optimal expectation value, additional peaks, narrowing of these, and new shapes.

All these phenomena are very similar to those observed in the $J_{ij} \leq h_i$ case. However, unlike previously, the number of additional peaks is larger, the narrowing is much more noticeable, and the new shapes are easily observable. Thus, the increase of the relative J_{ij} values has a bigger impact in generating wild expectation value landscapes. Notice also that the expectation value of some of these additional peaks is very similar to that of the optima. For the instance displayed in Fig. 10 the peaks are already very narrow and difficult to distinguish. This can cause difficulties when trying to set the optimal angles in an experimental setup due to precision issues. Our method can be useful in this case to identify the most suitable set of optimal angles that are experimentally more feasible.

We have shown that varying the parameters of the cost Hamiltonian the expectation value landscapes can look very different to the ones presented at the beginning of this paper; one example of this is shown in Fig. 10. This means that for certain Ising instances the computation of optimal angles is not straightforward and further study is required. As another example, consider the class of Ising instances where h_i can take any integer value at random between h and $-h$, i.e., $h_i \in \{-h, \dots, h\}$, and $J_{ij} \in \{-h, h\}$. The landscape of one such instance is shown in Fig. 11 for $h = 30$. Here we no longer have any of the recognisable patterns that we had in the first landscapes, and the optimal angles are not at the edges.

More general and practically relevant problems may have complicated landscapes and require their own study. As we have shown, QAOA landscapes can vary substantially depending on the class of problem instance, and finding the optimal angles using classical methods requires some care to be taken.

IV. CONCLUSIONS

We introduced an analytical formula for the expectation value of single-layer ($p = 1$) QAOA circuits solving general Ising Hamiltonians, which allows for an efficient classical calculation of the optimal angles β_{\min} , γ_{\min} . We validated our analytical formula by numerically evaluating the QAOA expectation values of several Ising problems studied in a recent experimental paper [8]. Furthermore, we predicted the energy landscapes for single-layer QAOA with 100 qubits, and benchmarked the expectation values for the G-set MAX-CUT problems and custom Ising problems with hundreds to thousands of vertices/spins/qubits. For the G-set benchmark instances the ratio between the theoretical estimated maximum Ising energy and the QAOA expectation for Ising energy was in the range $0.31 - 0.4$; for the Ising problems with external fields that we tested, the ratio was in the range $0.32 - 0.39$. In order to compute these ratios—since we don't know the true optimal energy for each of such large Ising instances—we derived an analytical formula that allows one to estimate the optimal energy of an Ising or MAX-CUT instance with weights $\{-1, 0, +1\}$. The largest instance that we estimated the accuracy of the single-layer QAOA for had 100 000 vertices and 150 000 edges, and found that it achieved a ratio of ≈ 0.39 .

For the G-set MAX-CUT instances, we also computed the ratio between the expected cut value from the single-layer QAOA and the best-known cut value (obtained using classical heuristic methods) for each instance. For the G-set instances that are *unweighted* (i.e., $J_{ij} \in \{0, 1\}$), the ratios are all ≥ 0.9 —meaning that we can expect the QAOA to find MAX-CUT solutions for these instances that have at least 90% of the maximum (known) cut value. The accuracy of the solutions one would obtain by running the single-layer QAOA on such large instances in practice (under the assumption of noiseless quantum hardware) should be even higher, because each run of the QAOA will yield a single sample from a distribution that has a *mean* solution value with $\sim 90\%$ accuracy, and so multiple runs of the QAOA are likely to yield some samples that have solution values better than that of the mean value. Two related open questions of great interest are *what the shape of the output-sample distribution from the QAOA is and what the probabilities of measuring solutions with energies much better than the mean energy are* [21].

We also note that the multi-layer QAOA cannot perform worse than the single-layer QAOA [22], so our results provide a lower bound for the performance of the multi-layer QAOA.

While the numerical evaluation of our analytical formula has polynomial scaling, the wall-clock times for our current classical code implementation are typically $\gg 1000$ seconds for Ising instances comprising $|V| \gg 100$ spins (exact times are given in Tables II and III). An area for future work is to increase the speed of the classical code for computing optimal angles.

Finally, we studied the effect that the values of the external-field terms h_i and spin-spin-coupling terms J_{ij} have on the expectation-value landscape and the optimal γ parameter, γ_{\min} . We showed that an increase of the external fields or spin-spin couplings appear, in general, to lead to the generation of additional peaks in the landscape, the narrowing of these peaks, new shapes, a reduction in optimal expectation value, and γ_{\min} being closer to the edges. A relative increase in J_{ij} has a bigger impact in the generation of complex expectation landscapes. We found that a broad class of Ising Hamiltonians have highly complex landscapes that are very dissimilar to the ones that we are aware of in the literature. Our analytical methods for finding optimal angles may be particularly useful in these cases, especially since angle searches performed *in situ* using quantum hardware may easily “miss” very narrow features in the landscapes.

Note added in preparation: Related work studying the single-layer QAOA numerically appeared on the arXiv on 29 November 2020: Ref. [23], which investigates the performance of the QAOA on a particular subclass of MAX-CUT instances related to graph coloring.

V. ACKNOWLEDGEMENTS

The authors thank V. Putz, J. I. Adame, and R. Parrish for useful discussions, and J. I. Adame for a thorough reading of a draft of this paper. PLM acknowledges membership of the CIFAR Quantum Information Science Program as an Azrieli Global Scholar.

-
- [1] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithm,” *arXiv:1411.4028*, 2014.
 - [2] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, *et al.*, “Quantum optimization using variational algorithms on near-term quantum devices,” *Quantum Science and Technology*, vol. 3, no. 3, p. 030503, 2018.
 - [3] G. E. Crooks, “Performance of the quantum approximate optimization algorithm on the maximum cut problem,” *arXiv:1811.08419*, 2018.
 - [4] 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,” *arXiv:1812.01041*, 2018.
 - [5] T. Hogg and D. Portnov, “Quantum optimization,” *Information Sciences*, vol. 128, no. 3-4, pp. 181–197, 2000.
 - [6] C. A. Trugenberger, “Quantum optimization for combinatorial searches,” *New Journal of Physics*, vol. 4, no. 1, p. 26, 2002.

- [7] A. Lucas, “Ising formulations of many NP problems,” *Frontiers in Physics*, vol. 2, p. 5, 2014.
- [8] F. C. Arute, K. Arya, R. Babbush, D. Bacon, J. Bardin, R. Barends, S. Boixo, M. B. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, J. Chen, Y. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. R. Foxen, C. M. Gidney, M. Giustina, R. Graff, S. Habegger, S. Hong, L. Ioffe, S. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Leib, M. Lindmark, E. Lucero, O. Martin, J. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, F. Neukart, H. Neven, M. Y. Niu, T. E. O’Brien, B. O’Gorman, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. Rubin, D. Sank, K. Satzinger, A. Skolik, V. Smelyanskiy, D. Strain, M. Streif, K. J. Sung, M. Szalay, A. Vainsencher, T. White, J. Yao, A. Zalcman, and L. Zhou, “Quantum approximate optimization of non-planar graph problems on a planar superconducting processor,” *arXiv:2004.04197*, 2020.
- [9] M. Streif and M. Leib, “Training the quantum approximate optimization algorithm without access to a quantum processing unit,” *Quantum Science and Technology*, vol. 5, no. 3, p. 034008, 2020.
- [10] Y. Ye, “Gset test problems,” 2003. <https://web.stanford.edu/~yyye/yyye/Gset/>.
- [11] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, “Quantum approximate optimization algorithm for maxcut: A fermionic view,” *Physical Review A*, vol. 97, no. 2, p. 022304, 2018. *arXiv:1706.02998*.
- [12] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, “Obstacles to state preparation and variational optimization from symmetry protection,” *arXiv:1910.08980*, 2019.
- [13] F. Barahona, “On the computational complexity of ising spin glass models,” *Journal of Physics A: Mathematical and General*, vol. 15, pp. 3241–3253, oct 1982.
- [14] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, “Quantum approximate optimization algorithm for maxcut: A fermionic view,” *Phys. Rev. A*, vol. 97, p. 022304, Feb 2018.
- [15] Y. Matsuda, “Benchmarking the MAX-CUT problem on the simulated bifurcation machine,” *Medium*, 2019. <https://medium.com/toshiba-sbm/benchmarking-the-max-cut-problem-on-the-simulated-bifurcation-machine-e26e1127c0b0>.
- [16] Google Quantum AI Team and collaborators, “Recirq,” Oct. 2020. <https://doi.org/10.5281/zenodo.4091470>.
- [17] Google Quantum AI Team and collaborators, “Cirq,” Oct. 2020. <https://doi.org/10.5281/zenodo.4062499>.
- [18] U. Benlic and J.-K. Hao, “Breakout local search for the max-cut problem,” *Engineering Applications of Artificial Intelligence*, vol. 26, p. 1162–1173, 03 2013.
- [19] T. Leleu, Y. Yamamoto, P. L. McMahon, and K. Aihara, “Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity,” *Physical Review Letters*, vol. 122, no. 4, p. 040607, 2019.
- [20] F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, “For fixed control parameters the quantum approximate optimization algorithm’s objective function value concentrates for typical instances,” *arXiv:1812.04170*, 2018.
- [21] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyriyanidis, W. L. Tan, C. Baldwin, et al., “Quantum approximate optimization of the long-range Ising model with a trapped-ion quantum simulator,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 41, pp. 25396–25401, 2020.
- [22] Since one can always choose to set the angles for all layers after the first one to be equal to 0, and hence have the layers after the first layer perform the identity operation.
- [23] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, “Hybrid quantum-classical algorithms for approximate graph coloring,” *arXiv:2011.13420*, 2020.
- [24] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, “The quantum approximate optimization algorithm and the sherrington-kirkpatrick model at infinite size,” *arXiv preprint arXiv:1910.08187*, 2019.
- [25] A. Montanari, “Optimization of the sherrington-kirkpatrick hamiltonian,” in *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 1417–1433, IEEE, 2019.
- [26] G. Parisi, “Infinite number of order parameters for spin-glasses,” *Physical Review Letters*, vol. 43, no. 23, p. 1754, 1979.

Appendix A: Derivation of Analytical Expression for General Expectation

Given an undirected graph $G = (V, E)$ with vertices $V = \{1, \dots, n\}$ and edges $E \subset V \times V$, combined with external fields h_i for the vertices, and coupling strengths J_{ij} on the edges, we aim to find the configuration $s \in \{-1, +1\}^n$ that minimizes the cost function

$$C = \sum_{i \in V} h_i s_i + \sum_{(i,j) \in E} J_{ij} s_i s_j. \quad (\text{A1})$$

This cost function is equivalent to the cost Hamiltonian

$$H_C = \sum_{i \in V} C_i + \sum_{(i,j) \in E} C_{ij}, \quad (\text{A2})$$

where

$$\begin{cases} C_i = h_i \sigma_i^z \\ C_{ij} = J_{ij} \sigma_i^z \sigma_j^z \end{cases} \quad (\text{A3})$$

The unitary transformation U for a QAOA circuit with depth $p = 1$ is defined as

$$U = U_B^\beta U_C^\gamma \quad (\text{A4})$$

where

$$U_C^\gamma = e^{-i\gamma H_C} = e^{-i\gamma(\sum_{ij} J_{ij} \sigma_i^z \sigma_j^z + \sum h_i \sigma_i^z)} \quad (\text{A5})$$

$$= \prod_{(i,j) \in E} (\cos(J_{ij}\gamma) \sigma_i^0 \sigma_j^0 - i \sin(J_{ij}\gamma) \sigma_i^z \sigma_j^z) \prod_{i \in V} (\cos(h_i\gamma) \sigma_i^0 - i \sin(h_i\gamma) \sigma_i^z) \quad (\text{A6})$$

$$U_B^\beta = e^{-i\beta H_B} = e^{-i\beta \sum_i \sigma_i^x} \quad (\text{A7})$$

$$= \prod_{i \in V} (\cos(\beta) \sigma_i^0 - i \sin(\beta) \sigma_i^x). \quad (\text{A8})$$

Applying $U = U_B^\beta U_C^\gamma$ to an initial, uniform superposition of all bit strings, $|\psi_0\rangle$, we obtain the final QAOA state

$$|\beta, \gamma\rangle = U |\psi_0\rangle. \quad (\text{A9})$$

The expectation value of H_C for this final state reads

$$F(\beta, \gamma) = \sum_{i \in V} \langle C_i \rangle + \sum_{(i,j) \in E} \langle C_{ij} \rangle \quad (\text{A10})$$

where

$$\begin{cases} \langle C_i \rangle = \langle \beta, \gamma | C_i | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_i U | \psi_0 \rangle = h_i \text{Tr}[\rho_0 U^\dagger \sigma_i^z U] \\ \langle C_{ij} \rangle = \langle \beta, \gamma | C_{ij} | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_{ij} U | \psi_0 \rangle = J_{ij} \text{Tr}[\rho_0 U^\dagger \sigma_i^z \sigma_j^z U] \end{cases} \quad (\text{A11})$$

and $\rho_0 = |\psi_0\rangle \langle \psi_0|$ is the initial density matrix.

Resolving the U_B unitaries, the terms inside the traces read

$$\begin{cases} U^\dagger \sigma_i^z U = U_C^{-\gamma} (\cos(2\beta) \sigma_i^z + \sin(2\beta) \sigma_i^y) U_C^\gamma \\ U^\dagger \sigma_i^z \sigma_j^z U = U_C^{-\gamma} (\cos(2\beta) \sigma_i^z + \sin(2\beta) \sigma_i^y) (\cos(2\beta) \sigma_j^z + \sin(2\beta) \sigma_j^y) U_C^\gamma \end{cases} \quad (\text{A12})$$

As a result, we can break down the computation into the evaluation of the following 6 traces, which we group according to the number of σ^y terms:

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^z U_C^\gamma], \quad \text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^z \sigma_j^z U_C^\gamma], \quad (\text{A13})$$

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y U_C^\gamma], \quad (\text{A14})$$

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y \sigma_j^z U_C^\gamma], \quad \text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^z \sigma_j^y U_C^\gamma], \quad (\text{A15})$$

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y \sigma_j^y U_C^\gamma]. \quad (\text{A16})$$

Note that the terms in $U_C^{\pm\gamma}$ that do not involve the vertices i and j commute with $\sigma_i^* \sigma_j^*$, hence those terms will cancel each other. When evaluating $U_C^{-\gamma} \sigma_i^* \sigma_j^* U_C^\gamma$ we therefore only have to keep track of the terms of C involving h_i , h_j , J_{ij} , J_{ik} , and J_{jk} (with $k \neq i, j$). Also, because $\rho_0 = 2^{-n} \prod_{j=1}^n (\sigma_j^0 + \sigma_j^x)$, the trace $\text{Tr}[\rho_0 M]$ only depends on the $\{\sigma^0, \sigma^x\}^n$ components of M , that is:

$$\text{Tr} \left[\rho_0 \sum_{s \in \{0,x,y,z\}^n} c_s \sigma_1^{s_1} \cdots \sigma_n^{s_n} \right] = \sum_{s \in \{0,x\}^n} c_s. \quad (\text{A17})$$

Going back to the six traces, because σ_i^z and σ_j^z commute with U_C we see that the traces of Equation A13 equal 0.

To tackle the traces of Equation A14 we perform the calculation

$$U_C^{-\gamma} \sigma_i^y U_C^\gamma = \prod_{(i,k) \in E} (\cos(J_{ik}\gamma) \sigma_i^0 \sigma_k^0 + i \sin(J_{ik}\gamma) \sigma_i^z \sigma_k^z) \cdot (\cos(h_i\gamma) \sigma_i^0 + i \sin(h_i\gamma) \sigma_i^z) \cdot \sigma_i^y \quad (\text{A18})$$

$$\times (\cos(h_i\gamma) \sigma_i^0 - i \sin(h_i\gamma) \sigma_i^z) \prod_{(i,k) \in E} (\cos(J_{ik}\gamma) \sigma_i^0 \sigma_k^0 - i \sin(J_{ik}\gamma) \sigma_i^z \sigma_k^z) \quad (\text{A19})$$

$$= \prod_{(i,k) \in E} (\cos(J_{ik}\gamma) \sigma_i^0 \sigma_k^0 + i \sin(J_{ik}\gamma) \sigma_i^z \sigma_k^z) \cdot (\sin(2h_i\gamma) \sigma_i^x + \cos(2h_i\gamma) \sigma_i^y) \quad (\text{A20})$$

$$\times \prod_{(i,k) \in E} (\cos(J_{ik}\gamma) \sigma_i^0 \sigma_k^0 - i \sin(J_{ik}\gamma) \sigma_i^z \sigma_k^z). \quad (\text{A21})$$

Hence, relying on Equation A17 and some minimum algebra we have for Equation A14:

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y U_C^\gamma] = \text{Tr} \left[\rho_0 \sin(2h_i \gamma) \prod_{(i,k) \in E} ((\cos(J_{ik}\gamma))^2 - (\sin(J_{ik}\gamma))^2) \sigma_i^x \sigma_k^0 \right] \quad (\text{A22})$$

$$= \sin(2h_i \gamma) \prod_{(i,k) \in E} \cos(2J_{ik}\gamma). \quad (\text{A23})$$

For trace Expressions A15 we use the fact that $\rho_0 U_C^{-\gamma} \sigma_i^y \sigma_j^z U_C^\gamma = \rho_0 U_C^{-\gamma} \sigma_i^y U_C^\gamma \sigma_j^z$, and by some additional algebra similar to the calculation above we obtain

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y U_C^\gamma \sigma_j^z] = \text{Tr} \left[\rho_0 \cos(2h_i \gamma) 2 \sin(J_{ij}\gamma) \cos(J_{ij}\gamma) \prod_{\substack{(i,k) \in E \\ k \neq j}} \cos(2J_{ik}\gamma) \sigma_i^x \sigma_j^z \sigma_k^0 \cdot \sigma_j^z \right] \quad (\text{A24})$$

$$= \cos(2h_i \gamma) \sin(2J_{ij}\gamma) \prod_{\substack{(i,k) \in E \\ k \neq j}} \cos(2J_{ik}\gamma). \quad (\text{A25})$$

By symmetry, we have, $\rho_0 U_C^{-\gamma} \sigma_i^z \sigma_j^y U_C^\gamma = \rho_0 U_C^{-\gamma} \sigma_j^y U_C^\gamma \sigma_i^z$ and,

$$\text{Tr}[\rho_0 \sigma_i^z U_C^{-\gamma} \sigma_j^y U_C^\gamma] = \cos(2h_j \gamma) \sin(2J_{ij}\gamma) \prod_{\substack{(j,k) \in E \\ k \neq i}} \cos(2J_{jk}\gamma). \quad (\text{A26})$$

We thus have to calculate the remaining case, Expression A16. We used Mathematica to perform and verify the algebraic manipulations necessary to calculate Expression A16 and obtain the result of Eq. A28, since a large number of terms is involved.

$$\text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^y \sigma_j^y U_C^\gamma] = -\frac{1}{2} \prod_{\substack{(ik) \in E \\ (jk) \notin E}} \cos(2\gamma J_{ik}) \prod_{\substack{(jk) \in E \\ (ik) \notin E}} \cos(2\gamma J_{jk}) \quad (\text{A27})$$

$$\times \left[\cos(2\gamma(h_i + h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} + J_{jk})) - \cos(2\gamma(h_i - h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} - J_{jk})) \right] \quad (\text{A28})$$

Resulting in the expression,

$$\left\{ \begin{array}{l} \langle C_i \rangle = h_i \sin(2\beta) \sin(2\gamma h_i) \prod_{(i,k) \in E} \cos(2\gamma J_{ik}) \\ \langle C_{ij} \rangle = J_{ij} \sin(2\beta) \cos(2\beta) \sin(2\gamma J_{ij}) \left[\cos(2\gamma h_i) \prod_{\substack{(ik) \in E \\ k \neq j}} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{\substack{(jk) \in E \\ k \neq i}} \cos(2\gamma J_{jk}) \right] \\ - \frac{J_{ij}}{2} (\sin(2\beta))^2 \prod_{\substack{(ik) \in E \\ (jk) \notin E}} \cos(2\gamma J_{ik}) \prod_{\substack{(jk) \in E \\ (ik) \notin E}} \cos(2\gamma J_{jk}) \\ \times \left[\cos(2\gamma(h_i + h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} + J_{jk})) - \cos(2\gamma(h_i - h_j)) \prod_{\substack{(ik) \in E \\ (jk) \in E}} \cos(2\gamma(J_{ik} - J_{jk})) \right]. \end{array} \right. \quad (\text{A29})$$

Appendix B: Estimation of Optimal Energies

Here we will describe an estimate of the maximal value of a polynomial energy function in n spins. In general, for vertices $V = \{1, \dots, n\}$, consider the cost function $C : \{-1, +1\}^V \rightarrow \mathbb{R}$ defined by the polynomial

$$C(z_1, \dots, z_n) = \sum_{S \subseteq V} c_S z_S, \quad (\text{B1})$$

where z_S is shorthand for the product $\sum_{i \in S} z_i$ and the polynomial has 2^n coefficients $c_S \in \mathbb{R}$. Without loss of generality we assume that $c_\emptyset = 0$, i.e. the function has zero offset. For a uniform 2^{-n} distribution over the possible strings z , the expected function value will be 0, and the expected variance will be the sum of squares of its coefficients, i.e.

$$\langle C \rangle = \mathbb{E}[C(Z)] = 2^{-n} \sum_{z \in \{-1, +1\}^n} C(z) = 0 \quad (\text{B2})$$

$$\langle C^2 \rangle = \mathbb{E}[C^2(Z)] = 2^{-n} \sum_{z \in \{-1, +1\}^n} C^2(z) = \sum_{S \subseteq V} c_S^2. \quad (\text{B3})$$

This is an exact result that holds for all polynomials C , and hence also for quadratic Ising problems. The maximum of interest is defined by

$$M = \max_{z \in \{-1, +1\}^n} C(z). \quad (\text{B4})$$

The general idea that we use here is that M can be approximated by the expected maximum among 2^n independent samples of a distribution that mimics picking a random string $z \in \{-1, +1\}^n$ and then returning the value $C(z)$. More precisely let Z be the random variable that produces strings from $\{0, 1\}^n$ with uniform probability 2^{-n} , then we want to estimate M using the expectation

$$\hat{M} = \mathbb{E}[\max(C(Z_1), \dots, C(Z_{2^n}))]. \quad (\text{B5})$$

Next we will argue that the distribution of the random variable C will approximate a normal distribution and that we can estimate M by the expected maximum of 2^n samples of this normal distribution.

Assume that we can model the random variable $C(Z)$ by the summation $\sum_{S \subseteq V} c_S R_S$ with R_S a uniformly random spin value $\{-1, +1\}$. This model is not exact as it ignores the correlations between terms like $z_{\{i,j\}}$ and $z_{\{j,k\}}$ and $z_{\{i,k\}}$, but it can nevertheless give us intuition about what to expect about the values of C . The expectation of $c_S R_S$ is 0 and its variance is c_S^2 . Hence for large n the central limit theorem tells us that this sum will approximate the normal distribution $N(0, \sigma^2)$ with $\sigma^2 = \sum_S c_S^2$.

What remains is to calculate what to expect for the maximum among 2^n samples from $N(0, \sigma^2)$. When sampling k times the normal distribution $N(0, 1)$, the expected maximum of these samples can be approximated by $\sqrt{2 \ln k}$. In our case here we have $k = 2^n$ and the scale has to be multiplied by the standard deviation σ . Hence it we reached the estimate

$$\hat{M} \approx \sigma \sqrt{2n \ln 2} = \sqrt{2n \ln 2} \sqrt{\sum_{S \subseteq V} c_S^2}. \quad (\text{B6})$$

For G-set instances with $n = |V|$ and $\sum_S c_S^2 = |E|$ we thus have

$$\hat{M} \approx \sqrt{2|V| \ln 2} \sqrt{|E|} \approx 1.18 \sqrt{|V||E|}. \quad (\text{B7})$$

For I-set instances with $n = |V|$ and $\sum_S c_S^2 = |E| + |V|$ we have

$$\hat{M} \approx \sqrt{2|V| \ln 2} \sqrt{|E| + |V|} \approx 1.18 \sqrt{|V|(|E| + |V|)}. \quad (\text{B8})$$

We reiterate that the above is an informal argument and can almost certainly be improved upon. Indeed the literature on random Sherrington-Kirkpatrick models strongly suggests that the 1.18 constant should in fact be smaller. See for example [24, 25]. A future version of this article will include the estimates that follow from the body of work on the Parisi formula [26].

Appendix C: Normalized Expectation-Value Landscapes

In this section we present several expectation-value landscapes normalized by our analytical approximation of the optimal energy, namely $E_{\text{opt}} \approx 1.18 \sqrt{|E| \times |V|}$ and $E_{\text{opt}} \approx 1.18 \sqrt{|V| \times (|E| + |V|)}$. The new normalized figures are Fig. 12, Fig. 13, Fig. 14, and Fig. 15 that correspond to Fig. 2, Fig. 3, Fig. 4, and Fig. 5 in the paper, respectively.

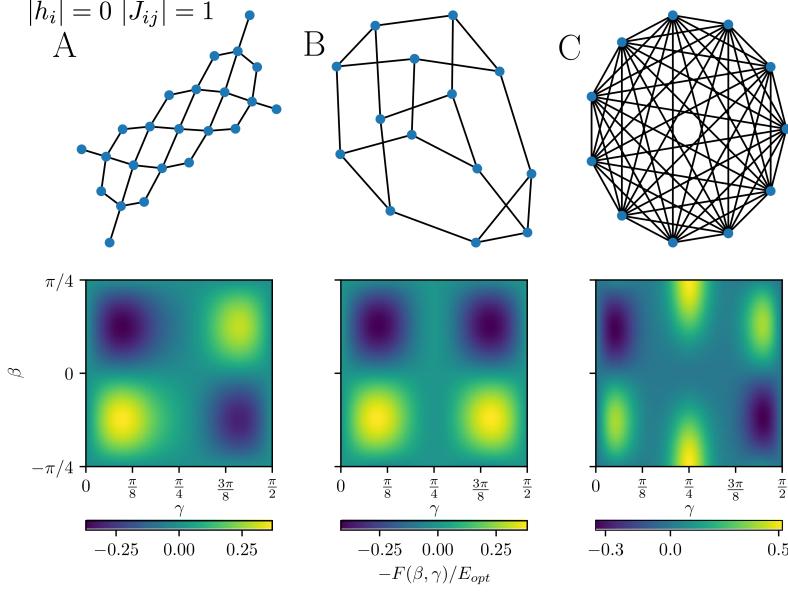


FIG. 12. Negative expectation-value landscapes, obtained using our analytical formulae, Eq. 12, for the same problem instances studied experimentally in Ref. [8]. Normalized by the optimal energy E_{opt} . The graphs for each instance are shown above their respective landscape. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 14 qubits. **C:** Fully connected SK-model with 11 qubits.

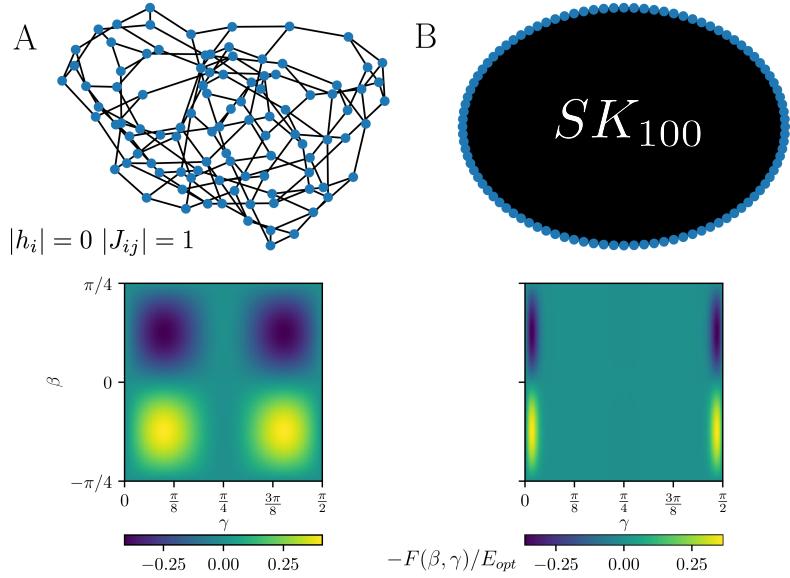


FIG. 13. Negative expectation-value landscapes for two 100-qubit instances. Normalized by the optimal energy E_{opt} . **A:** 3-regular graph with 100 qubits. **B:** Fully connected SK-model with 100 qubits.

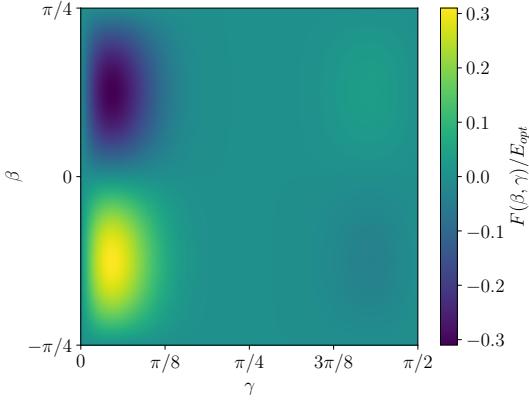


FIG. 14. Expectation-value landscape for the $G64$ instance with 7000 vertex and 41459 edges, obtained using the analytical formula Eq. 12. Normalized by the optimal energy E_{opt} .

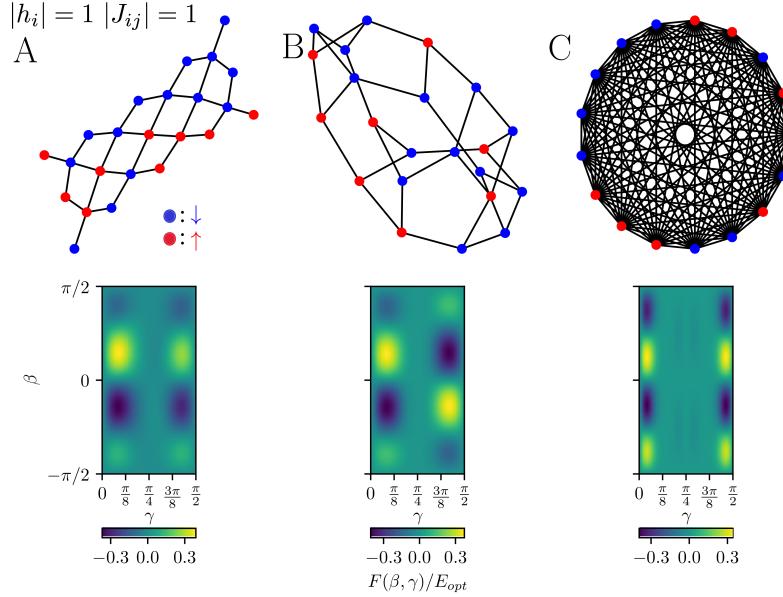


FIG. 15. Expectation-value landscapes obtained for problem instances with $h_i, J_{ij} \in \{-1, 1\}$, using the analytical formula Eq. 12. Normalized by the optimal energy E_{opt} . Problem graphs are shown above, with blue nodes representing those with $h_i = -1$, and red for those with $h_i = 1$. **A:** Sycamore grid problem with 23 qubits. **B:** 3-regular graph with 22 qubits. **C:** Fully connected SK-model with 17 qubits.