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 7000 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.
B. Estimation of Optimal Energies

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. In order for QAOA to be effective it suffices to find good enough parameters, not necessarily optimal. However, the effectiveness of the QAOA will be directly related to the quality of these parameters. 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 \(-1\). 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 \[8\] 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-layer 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 \[8\] (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 \[8\], 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 7000 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 on studying the QAOA parameters and energy-expectation landscapes analytically in the single-layer setting: \[11\] 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 \[12\] generalized this formula for the energy expectation for the QAOA to the case of weighted graphs. Our work builds on this as we include in our analysis the case of non-zero external-field terms in the Ising Hamiltonian. This generalization 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 \( V = \{1, \ldots, n\} \) and edges \( E \subset V \times V \), as well as 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

\[
C = \sum_{i \in V} h_i s_i + \sum_{(i,j) \in E} J_{ij} s_i s_j ,
\]
FIG. 1. A graph with 5 vertices \( V = \{1, \ldots, 5\} \) and 5 edges \( E = \{(12), (13), \ldots, (24)\} \). The product \( \prod_{(1k) \in E, k \neq 2} \) involves the vertices connected to 1 that are not 2, i.e. the vertices 3 and 5. The product \( \prod_{(1k) \in E, (2k) \in E} \) involves the vertices that are connected to 1 but that are not connected to 2, which is the single vertex 5. (Note that it is implicit that \( k = 2 \) should be excluded.) The product \( \prod_{(1k) \in E, (2k) \in E} \) involves the vertices that are connected to both 1 and 2, which is the vertex 3.

This cost function is equivalent to the cost Hamiltonian

\[
H_C = \sum_{i \in V} C_i + \sum_{(i,j) \in E} C_{ij}
\]

where

\[
\begin{align*}
C_i &= h_i \sigma^z_i \\
C_{ij} &= J_{ij} \sigma^z_i \sigma^z_j
\end{align*}
\]

Here \( \sigma^x_i, \sigma^y_i, \sigma^z_i \) correspond to the Pauli matrices. The unitary transformation \( U \) for a QAOA circuit with depth \( p = 1 \) is defined as

\[
U = U^\beta_B U^\gamma_C
\]

where

\[
U^\gamma_C = e^{-i \gamma H_C} = e^{-i \gamma \sum_{(i,j) \in E} J_{ij} \sigma^z_i \sigma^z_j + \sum h_i \sigma^z_i}
\]

\[
= \prod_{(i,j) \in E} (\cos(J_{ij} \gamma) \sigma^0_i \sigma^0_j - i \sin(J_{ij} \gamma) \sigma^z_i \sigma^z_j) \prod_{i \in V} (\cos(h_i \gamma) \sigma^0_i - i \sin(h_i \gamma) \sigma^z_i)
\]

\[
U^\beta_B = e^{-i \beta H_B} = e^{-i \beta \sum_i \sigma^x_i}
\]

\[
= \prod_{i \in V} (\cos(\beta) \sigma^0_i - i \sin(\beta) \sigma^x_i)
\]

Applying \( U = U^\beta_B U^\gamma_C \) to an initial, uniform superposition of all bit strings, \( |\psi_0\rangle \), we obtain the final QAOA state

\[
|\beta, \gamma\rangle = U |\psi_0\rangle.
\]

The expectation value of \( H_C \) for this final state is

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

where

\[
\begin{align*}
\langle C_i \rangle &= \langle \beta, \gamma | C_i | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_i U | \psi_0 \rangle = h_i \text{Tr}[\rho_0 U^\dagger \sigma^z_i U] \\
\langle C_{ij} \rangle &= \langle \beta, \gamma | C_{ij} | \beta, \gamma \rangle = \langle \psi_0 | U^\dagger C_{ij} U | \psi_0 \rangle = J_{ij} \text{Tr}[\rho_0 U^\dagger \sigma^z_i \sigma^z_j U]
\end{align*}
\]
and \( \rho_0 = |\psi_0 \rangle \langle \psi_0 | \) is the initial density matrix. The expectation value can be computed classically, however, a quantum computer is still required to obtain the bitstring with the relevant energy.

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

\[
\langle C_i \rangle = h_i \sin(2\beta) \sin(2\gamma h_i) \prod_{(ik) \in E} \cos(2\gamma J_{ik})
\]

\[
\langle C_{ij} \rangle = \frac{J_{ij} \sin(4\beta)}{2} \sin(2\gamma J_{ij}) \left[ \cos(2\gamma h_i) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}) \right] - \frac{J_{ij}}{2} (\sin(2\beta))^2 \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik} - J_{jk}) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik} + J_{jk}) \right].
\]

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 the complete graph \( K_n \) as we can use \( J_{ij} = 0 \) for absent edges. Doing so simplifies the above expressions to

\[
\langle C_i \rangle_{K} = h_i \sin(2\beta) \sin(2\gamma h_i) \prod_{k \neq i} \cos(2\gamma J_{ik})
\]

\[
\langle C_{ij} \rangle_{K} = \frac{J_{ij} \sin(4\beta)}{2} \sin(2\gamma J_{ij}) \left[ \prod_{k \neq i,j} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{k \neq i,j} \cos(2\gamma J_{ik}) \right] - \frac{J_{ij}}{2} (\sin(2\beta))^2 \prod_{k \neq i,j} \cos(2\gamma (J_{ik} + J_{jk}) - \prod_{k \neq i,j} \cos(2\gamma (J_{ik} - J_{jk})).
\]

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

\[
\langle C_{ij} \rangle_{K, h=0} = \frac{J_{ij} \sin(4\beta)}{2} \sin(2\gamma J_{ij}) \left[ \prod_{k \neq i,j} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{k \neq i,j} \cos(2\gamma J_{ik}) \right] - \frac{J_{ij}}{2} (\sin(2\beta))^2 \prod_{k \neq i,j} \cos(2\gamma (J_{ik} + J_{jk})) - \prod_{k \neq i,j} \cos(2\gamma (J_{ik} - J_{jk})).
\]

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

\[
\langle C_{ij} \rangle_{\text{tf}} = \frac{J_{ij} \sin(4\beta)}{2} \sin(2\gamma J_{ij}) \left[ \cos(2\gamma h_i) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}) + \cos(2\gamma h_j) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}) \right] + \frac{J_{ij} (\sin(2\beta))^2 \sin(2\gamma h_i) \sin(2\gamma h_j) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}) \prod_{(ik) \in E \setminus (jk)} \cos(2\gamma J_{ik}).
\]

(15)
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 ($J_{ij} = \pm 1$). For the NP-hard “P5” Ising problem of [13] for a graph $(V, E)$ the goal is to minimize the Hamiltonian

$$H_{P5} = \sum_{(i,j) \in E} \sigma_i^z \sigma_j^z + \sum_{i \in V} \sigma_i^z$$ \hfill (16)

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 from Equation 12

$$\begin{align*}
\langle C_i \rangle_{P5} &= \sin(2\beta) \sin(2\gamma) (\cos(2\gamma))^{d_i} \\
\langle C_{ij} \rangle_{P5} &= \sin(4\beta) \cdot 2 \sin(2\gamma) \left[(\cos(2\gamma))^{d_i} + (\cos(2\gamma))^{d_j}\right] \\
&\quad + \frac{1}{2} (\sin(2\beta))^2 (\cos(2\gamma))^{d_i} + (\cos(2\gamma))^{d_j} - 2f_{ij} - 2(1 - (\cos(4\gamma))^{f_{ij}+1}).
\end{align*}$$ \hfill (17)

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

$$\begin{align*}
H_{MC} &= - \sum_{(i,j) \in E} \sigma_i^z \sigma_j^z \\
\langle C_i \rangle_{MC} &= 0 \\
\langle C_{ij} \rangle_{MC} &= \sin(4\beta) \cdot 2 \sin(2\gamma) \left[(\cos(2\gamma))^{d_i} + (\cos(2\gamma))^{d_j}\right] \\
&\quad - \frac{1}{2} (\sin(2\beta))^2 (\cos(2\gamma))^{d_i} + (\cos(2\gamma))^{d_j} - 2f_{ij} - 2(1 - (\cos(4\gamma))^{f_{ij}}).
\end{align*}$$ \hfill (18)

This is the result obtained by Wang et al. [11].

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

$$\begin{align*}
H_{NC} &= \sum_{i \in V} h_i \sigma_i^z \\
\langle C_i \rangle_{NC} &= h_i \sin(2\beta) \sin(2\gamma h_i) \\
\langle C_{ij} \rangle_{NC} &= 0.
\end{align*}$$ \hfill (19-21)

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.

C. Formulae for Standard Sherrington-Kirkpatrick Model without Local Field

To analyze the Sherrington-Kirkpatrick model, Equation 14 we assume no local field, i.e. $h_i = 0$, and that the all-to-all coupling values $J_{ij}$ are distributed according to an independent and identical distribution with mean 0. Taking the average over this distribution and using the facts that all $J_{ij}$ distributions are independent and that $E[J_{ij}] = 0$ we are interested in the expectation of expectations

$$\mathbb{E}_J[\langle C_{ij} \rangle] = (\sin(4\beta)) \cdot \mathbb{E}_J[J \sin(2\gamma J)] \cdot (\mathbb{E}_J[\cos(2\gamma J)])^{n-2}. \hfill (22)$$

Often one is interested in the $n \to \infty$ behaviour of this model, which is easiest to discuss when we scale the $J_{ij}$ random variables by $1/\sqrt{n-1}$ such that we have

$$J_{ij} = \frac{J_{ij}'}{\sqrt{n-1}}. \hfill (23)$$
where $J'$ is a random variable with mean 0 and variance $\sigma^2 \in O(1)$ independent of $n$. Doing so will result in an $O(1)$ energy per spin, i.e. $(C/n) = (n-1)/2 \cdot (C_{ij})$ will be of order 1. We use $\sqrt{n-1}$ instead of the more common $\sqrt{n}$ to make the finite $n$ expressions work out nicer.

For the normal distribution $J' \sim N(0, \sigma^2)$ we have $J \sim N(0, \sigma^2/(n-1))$, and hence

$$E_{J \sim N(0, \sigma^2/(n-1))}[J \sin(2\gamma J)] = 2\gamma \sigma^2 \cdot (n-1)^{-1} \cdot e^{-2\gamma^2 \sigma^2/(n-1)} \quad (24)$$

$$E_{J \sim N(0, \sigma^2/(n-1))}[\cos(2\gamma J)] = e^{-2\gamma^2 \sigma^2/(n-1)}. \quad (25)$$

The expectation of the expectations of the energy per spin therefore equals

$$E_{J' \sim N(0, \sigma^2)}[(C/n)] = \sin(4\beta) \cdot \gamma \sigma^2 e^{-2\gamma^2 \sigma^2} \quad (26).$$

Note that this is an exact expression for finite $n$ and it allows us to compute the optimal angles $\beta, \gamma$ that minimize the expectation as a function of $n$ and $\sigma$:

$$\beta_{\min} = -\pi/8, \quad \gamma_{\min} = \frac{1}{2\sigma}, \quad \text{with} \quad E_{J' \sim N(0, \sigma^2)}[(C/n)] = -\frac{\sigma}{2\sqrt{\pi}} \approx -0.303\sigma. \quad (27)$$

All of the above results agree with those in [14] Section 4. Note, however, that the authors in [14] take the equivalent approach of normalizing the entire cost function by the $\sqrt{n}$ factor, not just $J$. Moreover, [14] shows that in the $n \to \infty$ limit this result is independent of the details of the distribution $J'$, given that it has mean 0 and variance 1. As an example, consider the bimodal distribution $B$ where $Pr[J=\pm \sigma] = 1/2$ for some fixed $\sigma$, such that $E[J] = 0$ and $E[J^2] = \sigma^2$. As a result, we have $Pr[J=\pm \sigma/\sqrt{n-1}] = 1/2$ with $E_j[J \sin(2\gamma J)] = (\sigma/\sqrt{n-1}) \sin(2\gamma \sigma/\sqrt{n-1})$ and $E_j[\cos(2\gamma J)] = \cos(2\gamma \sigma/\sqrt{n-1})$, and hence

$$E_{J' \sim B}[(C/n)] = \frac{1}{2} \sin(4\beta) \cdot \sigma \sqrt{n-1} \cdot \sin(2\gamma \sigma/\sqrt{n-1}) \cdot (\cos(2\gamma \sigma/\sqrt{n-1}))^{n-2}. \quad (28)$$

In the $n \to \infty$ limit for large systems we thus get again

$$\lim_{n \to \infty} E_{J' \sim B}[(C/n)] = \sin(4\beta) \cdot \gamma \sigma^2 e^{-2\gamma^2 \sigma^2}, \quad (29)$$

which is identical to Equation 26.

It is important to realize, though, that there are distributions over $J'$ with mean 0 and variance $\sigma^2$ that give different results for the optimal angles and expectations. Consider the setting where we want to describe a spin glass where each vertex has a $b$-modal distribution over coupling values $J \in \{-1, 0, +1\}$ with $Pr[J=\pm 1] = d/(2n-2)$ and $Pr[J=0] = 1 - d/(n-1)$ such that $E_j J = 0$ and $E_j J^2 = \sigma^2$. As a result, we have $E_{J' \sim T}[J \sin(2\gamma J)] = d/(n-1) \sin(2\gamma)$, and $E_{J' \sim T}[\cos(2\gamma J)] = 1 + d(\cos(2\gamma) - 1)/(n-1)$ and hence

$$E_{J' \sim T}[(C/n)] = \frac{1}{2} \sin(4\beta) \cdot d \cdot \sin(2\gamma) \cdot (1 - d/(n-1) + d/(n-1) \cos(2\gamma))^{n-2}. \quad (30)$$

In the $n \to \infty$ limit this gives

$$E_{J' \sim T}[(C/n)] = \frac{1}{2} \sin(4\beta) \cdot d \cdot \sin(2\gamma) \cdot e^{-2d(\sin(\gamma))}. \quad (31)$$

If we set $d = 1$ the Equation 31 is maximized by $\beta = -\pi/8$, $\gamma = 0.452278$, such that $E_{J' \sim T}[(C/n)] = -0.268281$. Note now that this derived optimum is different from the results of Equation 27. This difference is explained by the fact that the related distribution of $J' = J/\sqrt{n-1}$ (with mean 0 and variance 1) is no longer independent of $n$, as it is defined by $Pr[J'=\pm \sigma/\sqrt{n-1}] = 1/(2n-2)$ and $Pr[J'=0] = (n-2)/(n-1)$. Because of this $n$-dependence it no longer holds that the $n \to \infty$ properties are described by Equations 26 and 27.

More specifically, Equation 26 will hold as long as we have the expectations in the $n \to \infty$ limit:

$$E_j [n J^2] \to \sigma^2, \quad \text{and} \quad E_j [n J'] \to 0, \quad \text{for } r \in \{3, 4, \ldots\}. \quad (32)$$

That way we have

$$\lim_{n \to \infty} E_j [(C/n)] = \lim_{n \to \infty} \frac{n-1}{2} \cdot \sin(4\beta) \cdot E_j [J \sin(2\gamma J)] \cdot (E_j [\cos(2\gamma J)])^{n-2} \quad (33)$$

$$= \sin(4\beta) \cdot \gamma \sigma^2 e^{-2\gamma^2 \sigma^2}. \quad (34)$$

The above trimodal distribution does not fit this mold as it has $E_j [n J^3] \to d$. 


D. Formulae for the Sherrington-Kirkpatrick Model with Constant Local Field

Consider now a variation of the SK model with a constant local field $h$ and with $J_{ij}$ couplings that again are i.i.d. with mean zero, the case shown in Equation 13. The local energies can now be expressed as

\[
\begin{align*}
\mathbb{E}_{j}[(C_{i})] &= h \sin(2\beta) \cdot \mathbb{E}_{j}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-1} \\
\mathbb{E}_{j}[(C_{ij})] &= \sin(4\beta) \cdot \mathbb{E}_{j}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-2}.
\end{align*}
\]  

(35)

The per spin energy thus equals

\[
\mathbb{E}_{j}[(C/n)] = h \sin(2\beta) \cdot \mathbb{E}_{j}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-1} + \frac{n-1}{2} \sin(4\beta) \cdot \mathbb{E}_{j}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-2}.
\]  

(36)

Assuming again that the $J$ values are normally distributed according to $J = J'/\sqrt{n-1}$ with $J' \sim N(0,\sigma^{2})$ we get for the energy per spin

\[
\mathbb{E}_{j' \sim N(0,\sigma^{2})}[(C/n)] = (2\gamma \sigma^{2} \cos(2h\gamma) \sin(4\beta) + h \sin(2\beta) \sin(2h\gamma))e^{-2\gamma^{2}\sigma^{2}}.
\]  

(37)

As an example, if we take $\sigma = 1$ and $h = 1$ we have the optimal angles $\beta \approx -0.54854$ and $\gamma \approx 0.3962$, such that $\mathbb{E}[(C/n)] \approx -0.62791$. Compare this with previous result for $(\sigma, h) = (1, 0)$ for which the optimal angles are $\beta = -\frac{\pi}{4} \approx -0.3927$ and $\gamma = 0.5$, with $\mathbb{E}[(C/n)] = -\frac{8}{\sqrt{2}} \approx -0.3033$. This shows that the presence of a local field does change not only the optimal angles but the expectation values in a significant way.

E. Formulae for the Sherrington-Kirkpatrick Model with Normal Local Field

Next, we will study the SK model with local fields distributed according to an independent and identical distribution with mean 0, again the case of Equation 13. The local energies can be expressed as

\[
\begin{align*}
\mathbb{E}_{j,h}[(C_{i})] &= \sin(2\beta) \cdot \mathbb{E}_{h}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-1} \\
\mathbb{E}_{j,h}[(C_{ij})] &= \sin(4\beta) \cdot \mathbb{E}_{h}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-2}.
\end{align*}
\]  

(38)

The per spin energy thus equals

\[
\mathbb{E}_{j,h}[(C/n)] = \sin(2\beta) \cdot \mathbb{E}_{h}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-1} + \frac{n-1}{2} \sin(4\beta) \cdot \mathbb{E}_{h}[(2\gamma h) \cdot \mathbb{E}_{j}[(2\gamma J)]]^{n-2}.
\]  

(39)

(40)

Assume now that the $h_{i}$ random variables will have the normal distribution $h \sim N(0,\sigma_{h}^{2})$:

\[
\mathbb{E}_{h \sim N(0,\sigma_{h}^{2})}[(2\gamma h)] = 2\gamma \sigma_{h}^{2} \cdot e^{-2\gamma^{2}\sigma_{h}^{2}}
\]  

(41)

\[
\mathbb{E}_{h \sim N(0,\sigma_{h}^{2})}[(cos(2\gamma h))] = e^{-2\gamma^{2}\sigma_{h}^{2}}.
\]  

(42)

The expectation of the expectations of the energy per spin equals

\[
\mathbb{E}_{h' \sim N(0,\sigma^{2})}[(C/n)] = (2\sigma_{h}^{2} \sin(2\beta) + \sigma^{2} \sin(4\beta)) \cdot e^{-2\gamma^{2}(\sigma^{2} + \sigma_{h}^{2})}.
\]  

(43)

We then compute the optimal angles $\beta$, $\gamma$,

\[\beta_{\text{min}} = -\arccos \left(\frac{\sqrt{4 - \sigma^{2} + 2\sigma^{2} + 2\sigma_{h}^{2}}}{2\sqrt{2}}\right), \quad \gamma_{\text{min}} = \frac{1}{2\sqrt{\sigma^{2} + \sigma_{h}^{2}}}\]  

(44)
We can reduce the expression of the expectation of the expectations of the energy per spin by making both parameters have the same distribution $\sigma = \sigma_h$

$$\mathbb{E}_{J' \sim \mathcal{N}(0, \sigma^2)} \mathbb{E}_{h \sim \mathcal{N}(0, \sigma^2)} [(C/n)] = (2\sin(2\beta) + \sin(4\beta)) \cdot \gamma \sigma^2 e^{-4\gamma^2 \sigma^2}. \quad (45)$$

And obtain the optimal angles $\beta, \gamma$,

$$\beta_{\min} = -\frac{\pi}{6}, \quad \gamma_{\min} = \frac{1}{2\sqrt{2}\sigma}, \quad \text{with} \quad \mathbb{E}_{J' \sim \mathcal{N}(0, \sigma^2)} \mathbb{E}_{h \sim \mathcal{N}(0, \sigma^2)} [(C/n)] = -\frac{3}{4} \sqrt{\frac{3}{2\sigma}} \approx -0.557\sigma. \quad (46)$$

### F. Formulae for Random Spin Glasses on Regular Graphs

When considering random $J$ couplings between vertices of a constant degree $d$ graph $G = (V, E)$ and no local field, the expectation of expectations for edges $(ij) \in E$ derived from Equation (12) will be

$$\mathbb{E}_{j} [(C_{ij})] = \sin(4\beta) \cdot \mathbb{E}_{j} [J \sin(2\gamma J)] \cdot (\mathbb{E}_{j} [\cos(2\gamma J)])^d. \quad (47)$$

where $d$ is the degree of the graph. For non-edges $(ij) \notin E$ we obviously have $\mathbb{E}_{j} [(C_{ij})] = 0$. If we assume a normal distribution of $J$ couplings over the edges and for each vertex $i$ the expectation $\mathbb{E}_{j} [\sum_{j} J_{ij}^2] = \sigma^2$, we should take $J \sim \mathcal{N}(0, \sigma^2/d)$. With this distribution we get

$$\mathbb{E}_{J \sim \mathcal{N}(0, \sigma^2/d)} [J \sin(2\gamma J)] = 2\gamma \sigma^2 \cdot d^{-1} \cdot e^{-2\gamma^2 \sigma^2/d} \quad (48)$$

and hence an expectation of the expectations of the energy per spin of

$$\mathbb{E}_{J \sim \mathcal{N}(0, \sigma^2/d)} [(C/n)] = \sin(4\beta) \cdot \frac{n-1}{d} \cdot \gamma \sigma^2 \cdot e^{-2\gamma^2 \sigma^2}. \quad (50)$$

The corresponding optimal values are thus

$$\beta_{\min} = -\frac{\pi}{8}, \quad \gamma_{\min} = \frac{1}{2\sigma}, \quad \text{and} \quad \mathbb{E}_{J' \sim \mathcal{N}(0, \sigma^2)} [(C/n)] = -\frac{n-1}{d} \cdot \frac{\sigma}{2\sqrt{e}}. \quad (51)$$

which for $d = n - 1$ returns the standard SK result.

Next, we add local fields. We obtain for degree $d$ graphs with normal distribution $h \sim \mathcal{N}(0, \sigma^2)$,

$$\left\{\begin{array}{l}
\mathbb{E}_{J, h} [(C_i)] = \sin(2\beta) \cdot \mathbb{E}_{h} [h \sin(2\gamma h)] \cdot \mathbb{E}_{j} [\cos(2\gamma J)]^d \\
\mathbb{E}_{J, h} [(C_{ij})] = \sin(4\beta) \cdot \mathbb{E}_{h} [\cos(2\gamma h)] \cdot \mathbb{E}_{j} [J \sin(2\gamma J)] \cdot \mathbb{E}_{j} [\cos(2\gamma J)]^d \end{array}\right. \quad (52)$$

The expectation of the expectations of the energy per spin equals

$$\mathbb{E}_{J' \sim \mathcal{N}(0, \sigma^2)} \mathbb{E}_{h \sim \mathcal{N}(0, \sigma^2)} [(C/n)] = (2\sin(2\beta) + \frac{n-1}{d} \sin(4\beta)) \cdot \gamma \sigma^2 e^{-4\gamma^2 \sigma^2}. \quad (53)$$

Here the optimal angle $\gamma_{\text{min}}$ remains constant as we modify $d$, while $\beta_{\text{min}}$ changes.

### III. NUMERICAL RESULTS AND DISCUSSION

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

The equations above allow us to replicate the expectation-value landscapes of the recent experimental work by Harrigan et al. [8], which provides numerical support for the validity of our analytical results. In this section, we present our investigation
TABLE I. Summary of the parameters of the figures present in the text.

| Figure | $h_i$ | $J_{ij}$ | n   | d   |
|--------|-------|----------|-----|-----|
| Fig. 2 | 0     | $\{-1,1\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 3 | 0     | $\{-1,1\}$ | 100 | 3 / 99 |
| Fig. 4 | 0     | $\{-1,1\}$ | 7000 | 11.85 |
| Fig. 5 | $\{-1,1\}$ | $\{-1,1\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 6 | $\{-2,2\}$ | $\{-2,2\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 7 | $\{-2,2\}$ | $\{-1,1\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 8 | $\{-10,10\}$ | $\{-1,1\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 9 | $\{-1,1\}$ | $\{-2,2\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 10 | $\{-1,1\}$ | $\{-10,10\}$ | 23 / 14 / 11 Sycamore / 3 / 10 |
| Fig. 11 | $\{-30, -29, \ldots, 29, 30\}$ | $\{-30, 30\}$ | 16 | 3 |

of the problems studied in Ref. [8], as well as some similar but larger problem instances. The expectation value landscapes are obtained evaluating our analytical formula in a grid of $500 \times 500$ points, and from the results we obtain the optimal angles, and optimal expectation value. For clarity, Table I shows a summary of the parameters of the figures found in the text.

The three problems studied in [8] are different weighted MAX-CUT instances with weights $-1$ and $+1$. The three specific instances used in [8] can be obtained from the ReCirq repository [17]. We numerically evaluate Eq. (12) for these instances, which gave us the energy-expectation landscapes for the single-layer QAOA. The numerical evaluation of the expectations $F(\beta, \gamma)$ for varying values of both $\beta$ and $\gamma$ gives an energy-expectation landscapes, shown in Fig. 2. (Note that we plot $-F(\beta, \gamma)$ due to the normalization by a negative quantity of the expectation values in [8].) We observe 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 [8].

In addition, the analytical formulas allow 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.

In relation with the previous section, note how the models shown in Fig. 2 C and Fig. 3 B correspond to the cases studied in Section II C. We predicted that $\sqrt{n - 1} \gamma_{\min} = 0.5$ for this bimodal distribution, and how that quantity will remain constant for any SK instances. For $n = 11$ we obtain $\sqrt{10} \gamma_{(11)}^{(11)} \approx \sqrt{10} \cdot 0.158 \approx 0.5$, and for $n = 100$, $\sqrt{99} \gamma_{(100)}^{(100)} \approx 99 \cdot 0.0527 \approx 0.5$, confirming our theoretical prediction.

On the other hand, Fig. 2 B and Fig. 3 A correspond to the cases studied in Section II F. Here we showed that $\sqrt{d} \gamma_{\min} = 0.5$
FIG. 3. Negative expectation-value landscapes for two 100-qubit instances. Normalized by the optimal energy $E_{\text{opt}}$. (A) 3-regular graph with 100 qubits. (B) Fully connected SK-model with 100 qubits.

FIG. 4. Negative expectation-value landscape for the G64 instance with 7000 vertices and 41 459 edges, obtained using the analytical formula Eq. 12. Normalized by the optimal energy $E_{\text{opt}}$.

for a bimodal distribution and $d$ regular graphs. In the first case, with $n = 14$, we obtain $\sqrt{3} \gamma_{\text{min}}^{(14)} \approx \sqrt{3} \cdot 0.316 \approx 0.5$. While for $n = 100$, the equality reads $\sqrt{3} \gamma_{\text{min}}^{(100)} \approx \sqrt{3} \cdot 0.306 \approx 0.5$.

We will now comment briefly on the features that one can observe in the landscapes shown in Figs. 2 and 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 $\gamma = \pi/4$. These peaks will vary in height depending on the distribution of the $\pm 1$ values in the graph.

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 in the landscape for the 100-spin instance are narrower and closer to the edges of the $\gamma$ range, and the peaks at $\gamma = \pi/4$ for the 11-spin instance are not present for the 100-spin instance. The peaks at $\gamma = \pi/4$ also become less pronounced as we increase the size of the SK graph. From the results derived in theory Sections II C and II F, we showed above that $\gamma_{\text{min}}$ will remain constant for 3-regular graphs, while decreasing with the number of spins for SK instances. The landscapes that we observe are a direct effect of this phenomenon.

The couplings of 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 phenomenon is explored in Sec. III C and it has an important experimental consequence: the peak sharpness places a constraint on how accurately the angle $\gamma$ needs to be implemented in hardware, and so the necessary precision for specifying $\gamma$ 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 Hamiltonian. Furthermore, we will study how the optimal $\gamma$ value, $\gamma_{\text{opt}}$, that minimizes the expectation value, changes with those parameters. So far we have seen evidence that increased degree of regular graphs can cause $\gamma_{\text{opt}}$ to be closer to 0 or $\pi/2$, but we will find that the value of $\gamma_{\text{opt}}$ 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, we simulated the expectation-value landscape of fully connected graphs with 23 qubits using Cirq [18]; these calculations took approximately 1 minute 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. [16]

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

The G-set [18] collection of benchmark instances for MAX-CUT is widely used in the comparison of classical MAX-CUT heuristics [15,19,20]. We chose to study single-layer QAOA’s performance on G-set MAX-CUT instances primarily because this enables comparison of QAOA’s predicted performance against a wide range of heuristics, including the best-known classical methods. We also studied QAOA on Ising instances having non-zero external-field terms. We generated these instances for this

| 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            | 4072            | 1482.034        | 0.364      | 4621.745        | 0.321      | {0, 1}  |
| G2   | 800  | 19   | 176            | 4064            | 1480.775        | 0.364      | 4621.745        | 0.321      | {0, 1}  |
| G3   | 800  | 19   | 176            | 4068            | 1478.236        | 0.363      | 4621.745        | 0.321      | {0, 1}  |
| G4   | 800  | 19   | 176            | 4116            | 1481.303        | 0.36        | 4621.745        | 0.321      | {0, 1}  |
| G5   | 800  | 19   | 176            | 4086            | 1479.496        | 0.362      | 4621.745        | 0.321      | {0, 1}  |
| G6   | 800  | 19   | 176            | 4202            | 1679.595        | 0.4        | 4621.745        | 0.363      | {0, ±1} |
| G7   | 800  | 19   | 176            | 4162            | 1674.41         | 0.402      | 4621.745        | 0.362      | {0, ±1} |
| G8   | 800  | 19   | 176            | 4180            | 1676.846        | 0.401      | 4621.745        | 0.363      | {0, ±1} |
| G9   | 800  | 19   | 176            | 4172            | 1676.17         | 0.402      | 4621.745        | 0.363      | {0, ±1} |
| G10  | 800  | 19   | 176            | 4160            | 1675.396        | 0.403      | 4621.745        | 0.363      | {0, ±1} |
| G11  | 800  | 19   | 176            | 1094            | 519.61          | 0.475      | 1335.018        | 0.389      | {0, ±1} |
| G12  | 800  | 19   | 176            | 1116            | 519.61          | 0.466      | 1335.018        | 0.389      | {0, ±1} |
| G13  | 800  | 19   | 176            | 1130            | 519.61          | 0.46        | 1335.018        | 0.389      | {0, ±1} |
| G14  | 800  | 19   | 176            | 1334            | 577.546         | 0.403      | 2286.644        | 0.253      | {0, 1}  |
| G15  | 800  | 19   | 176            | 1339            | 575.252         | 0.4        | 2278.592        | 0.252      | {0, 1}  |
| G16  | 800  | 19   | 176            | 1342            | 573.294         | 0.401      | 2281.279        | 0.252      | {0, 1}  |
| G17  | 800  | 19   | 176            | 1327            | 573.562         | 0.402      | 2280.058        | 0.252      | {0, 1}  |
| G18  | 800  | 19   | 176            | 1344            | 714.097         | 0.372      | 2286.644        | 0.312      | {0, ±1} |
| G19  | 800  | 19   | 176            | 1920            | 710.511         | 0.369      | 2278.592        | 0.312      | {0, ±1} |
| G20  | 800  | 19   | 176            | 1928            | 709.332         | 0.368      | 2281.279        | 0.311      | {0, ±1} |
| G21  | 800  | 19   | 176            | 1929            | 710.908         | 0.369      | 2280.058        | 0.312      | {0, ±1} |
| G22  | 2000 | 19   | 176            | 6724            | 2693.005        | 0.401      | 7461.109        | 0.361      | {0, ±1} |
| G23  | 2000 | 19   | 176            | 6700            | 2695.314        | 0.402      | 7461.109        | 0.361      | {0, ±1} |
| G24  | 2000 | 19   | 176            | 6730            | 2693.685        | 0.4        | 7461.109        | 0.361      | {0, ±1} |
| G25  | 2000 | 19   | 176            | 6726            | 2693.264        | 0.4        | 7461.109        | 0.361      | {0, ±1} |
| G26  | 2000 | 19   | 176            | 6700            | 2691.868        | 0.402      | 7461.109        | 0.361      | {0, ±1} |
| G27  | 5000 | 29   | 570            | 12076           | 4452.877        | 0.369      | 14345.043       | 0.31       | {0, ±1} |
| G28  | 700  | 17   | 148            | 11230           | 4590.429        | 0.409      | 12928.191       | 0.355      | {0, 1}  |
| G29  | 700  | 14   | 359            | 16975           | 6235.328        | 0.367      | 20102.054       | 0.31       | {0, ±1} |

Table II. 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 [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.
TABLE III. 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 [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 $c5n.large$ AWS instance, which uses a 3.0 GHz Intel Xeon Platinum Processor. [16]

| Name | $|V|$ | $|E|$ | Sum of weights | Best-known cut | QAOA expectation cut | Ratio cut | Time (s) |
|------|------|------|----------------|----------------|----------------------|-----------|----------|
| G1   | 800  | 19176| 19176          | 11624          | 10329.017            | 0.889     | 28.322   |
| G2   | 800  | 19176| 19176          | 11620          | 10328.387            | 0.889     | 28.562   |
| G3   | 800  | 19176| 19176          | 11622          | 10327.118            | 0.889     | 28.332   |
| G4   | 800  | 19176| 19176          | 11646          | 10328.651            | 0.887     | 28.430   |
| G5   | 800  | 19176| 19176          | 11631          | 10327.748            | 0.888     | 29.039   |
| G6   | 800  | 19176| 154            | 2178           | 916.797              | 0.42      | 31.953   |
| G7   | 800  | 19176| -150           | 2006           | 762.205              | 0.38      | 32.223   |
| G8   | 800  | 19176| -170           | 2005           | 753.423              | 0.376     | 32.685   |
| G9   | 800  | 19176| -64            | 2054           | 806.085              | 0.392     | 33.056   |
| G10  | 800  | 19176| -160           | 2000           | 757.698              | 0.379     | 33.681   |
| G11  | 800  | 1600 | 34             | 564            | 276.805              | 0.491     | 1193.0   |
| G12  | 800  | 1600 | -4             | 556            | 257.805              | 0.464     | 1188.0   |
| G13  | 800  | 1600 | 34             | 582            | 276.805              | 0.476     | 1194.0   |
| G14  | 800  | 4694 | 4694           | 3064           | 2635.773             | 0.86      | 4808.0   |
| G15  | 800  | 4661 | 4661           | 3050           | 2618.126             | 0.858     | 4982.0   |
| G16  | 800  | 4672 | 4672           | 3052           | 2623.147             | 0.859     | 5076.0   |
| G17  | 800  | 4667 | 4667           | 3047           | 2620.281             | 0.86      | 5066.0   |
| G18  | 800  | 4694 | 64             | 992            | 389.049              | 0.392     | 5125.0   |
| G19  | 800  | 4661 | -113           | 906            | 298.755              | 0.33      | 5187.0   |
| G20  | 800  | 4672 | -46            | 941            | 331.666              | 0.352     | 5186.0   |
| G21  | 800  | 4667 | -67            | 931            | 321.954              | 0.346     | 5237.0   |
| G22  | 2000 | 19990| -42            | 3431           | 1325.952             | 0.397     | 22103.0  |
| G23  | 2000 | 19990| -104           | 3298           | 1295.657             | 0.393     | 21916.0  |
| G24  | 2000 | 19990| 80             | 3405           | 1386.082             | 0.407     | 21741.0  |
| G25  | 2000 | 19990| 100            | 3413           | 1396.632             | 0.409     | 22292.0  |
| G26  | 2000 | 19990| -80            | 3310           | 1305.934             | 0.395     | 22460.0  |
| G27  | 5000 | 29570| 96             | 6086           | 2274.439             | 0.374     | 39492.0  |
| G28  | 7000 | 17148| 362            | 5796           | 2476.214             | 0.427     | 14794.0  |
| G29  | 7000 | 41459| 527            | 8751           | 3381.164             | 0.586     | 57847.0  |

TABLE IV. 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 a $c5n.large$ AWS instance, which uses a 3.0 GHz Intel Xeon Platinum Processor. [16]

| Name | $|V|$ | $|E|$ | Sum of weights | QAOA expectation | 1.18$\sqrt{|V| \times (|E| + |V|)}$ | Ratio exp. | Time (s) |
|------|------|------|----------------|-----------------|--------------------------|-----------|----------|
| I10  | 800  | 19176| -144           | 1706.929        | 4717.167                | 0.362     | 31.376   |
| I20  | 800  | 4672 | -128           | 787.059         | 2468.88                 | 0.319     | 5269.0   |
| I30  | 2000 | 19990| 86             | 2829.98         | 7825.455                | 0.36      | 22648.0  |
| I59  | 5000 | 29570| 156            | 4940.085        | 15513.747               | 0.318     | 40740.0  |
| I64  | 7000 | 41459| 661            | 6914.15         | 21732.929               | 0.318     | 59894.0  |
| I99  | 70000| 14000| -62            | 54809.093       | 143067.397              | 0.383     | 125363.0 |
| I100 | 100000| 150000| 956           | 72847.211       | 186574.382              | 0.39      | 137774.0 |

study (details provided later); this was motivated by the fact that we are unaware of a standard benchmark set for such Ising instances.

We calculated the optimal QAOA expectation values for a subset of the G-set instances having 800, 2000, 5000, or 7000 vertices and up to 41 459 edges. The results are divided into two tables; Table [II] where results are expressed in terms of Ising energies, and Table [III] 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 [II] we derive in the appendix an analytical formula that allows us to estimate the optimal energy of a G-set instance, $E_{opt} \approx 1.18 \sqrt{|E| \times |V|}$. The ratio between the theoretical maximum and the QAOA expectation, $f_{exp.} = f_{max.} / E_{opt}$, ranges between $0.252 - 0.389$. This ratio is similar to the formula of the approximation
We denote these instances as $I_G$-set instances, with similar optimal angles as predicted in [21]. Fig. 4 shows the landscape of instance $G64$.

Higher for instances with small amount of edges and common neighbours between vertices. These ratios can be used as a method in the analytical formula expectation difference.

The values of the $\beta$-intercept and $\gamma$-intercept are expected to be $0.3\%$ for most instances except $G14-G17$. For this set the $R\text{atio } IS\text{ing}$ and $R\text{atio cut}$ 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 $R\text{atio IS\text{ing}}$ 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 to determine the QAOA expectation. The expectation value landscapes are very similar for the $G$-set instances, with similar optimal angles as predicted in [21]. Fig. 4 shows the landscape of instance $G64$.

We also calculated the best QAOA expectation values for the 7 Ising instances listed in Table [V]. The first five of these instances were generated by adding weight $\{-1,+1\}$ external fields at random to instances of the $G$-set benchmark instances. We denote these instances as $xG$ where $x$ is the $G$ instance label on which it is based. Additionally, two more instances, $I99$ and $I100$, were generated by assigning $\{-1,+1\}$ weights and fields at random. The number of vertices 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 $0.318 - 0.39$ when comparing the theoretical optimal and the QAOA expectation. The values of the $r_{\text{exp}}$ ratios are very similar to those obtained for MAX-CUT instances. The $\text{Time}$ column displays in seconds the time required to obtain the full expectation value landscape and optimal angles.

In Table [V] we compare the optimal angles for a subset of the $G$-set instances with the universal angles suggested in [14], which are only strictly valid for regular graphs. To approximate this, we calculate the average degree $d_{\text{ave}}$ of the graph, and obtain the universal angles using this metric. We obtain the universal angles $\beta_{\text{uni}}, \gamma_{\text{uni}}$ using the formulas $\beta_{\text{uni}} = -\pi/8$ and $\gamma_{\text{uni}} = \pi/8$. The results are shown in Table [V].
FIG. 5. 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.

$\gamma_{uni} = 1/(2\sqrt{d_{ave}})$ derived in [14]. The true optimal angles for each instance of the G-set are obtained using a brute force approach such that the optimality of the angles is guaranteed up to the resolution of the search. The values displayed in columns $\beta_{dif}$ and $\gamma_{dif}$ correspond to the absolute difference between these two of angles. The standard deviation of the degrees of the vertices is captured by the standard expression $\sigma_d = (\sum_{i \in V} (d_i^2 - d_{ave}^2)/|V|)^{1/2}$. The $\sigma_d$ of an instance graph allows us to quantify how the graph is to a regular graph, which is defined by having zero deviation. Furthermore, we compare the difference in expectation values in the last column expectation $dif$ by calculating the percentage difference using the formula, $\% = 200|F_{opti} - F_{uni}|/(F_{opti} + F_{uni})$, where $F_{opti}$ and $F_{uni}$ are the expectation values of the optimal and universal angles respectively.
$|h_i| = 2 \mid J_{ij} \mid = 1$

A

B

C

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.

$|h_i| = 10 \mid J_{ij} \mid = 1$

A

B

C

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.

Overall we observe that the approach of using the average degree as a proxy for a regular degree is very successful, and that the obtained universal angles are similar to the optimal angles of the individual instances. With an observed maximum value of the expectation dif. of 1.5% for all the instances with most showing values smaller or equal to 1%. The cases with graphs closer to regular graphs do not show smaller differences in angles and expectations. How close the instance is to being regular is quantified by $\sigma_d$ as regular graphs are defined by $\sigma_d = 0$. Somewhat surprisingly, even for highly non-regular graphs (like G14–G21, G59, G64) the universal angles for regular graphs still work really well.

In general, higher $d_{ave}$ values will cause the differences in angles to affect more strongly the difference in expectation value because the area with relevant expectation values in the landscape will be narrower as we increase $d_{ave}$. However, we do not
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.

observe this phenomena in our results. Perturbations in $\gamma$ angles will also influence the difference in expectation value in a more relevant way than $\beta$ perturbations, due to the shape of the relevant expectation value area being longer in the $\beta$ direction. We also observe that having instances with weights $\{0, \pm 1\}$ or $\{0, 1\}$ does not affect the expectation value difference in a relevant way. The described approach can be applied to any other problem instances either as a good starting point for the expectation value optimization or as a set of good angles to use in the QAOA without the need for optimization.

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) Sycamore grid problem with 23 qubits. (B) 3-regular graph with 22 qubits. (C) Fully connected SK-model with 17 qubits.
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 $\gamma_{\text{min}}$ as

$$\gamma_{\text{min}} \approx \frac{1}{2h} \arctan \frac{1}{\sqrt{d}} \left( \frac{1}{2h} \arctan \frac{1}{\sqrt{d}} \right). \quad (54)$$

We previously observed how the regularity of the graph affects the $\gamma_{\text{min}}$ value, and now we see that increasing the $h$ value will also bring $\gamma_{\text{min}}$ closer to the edges ($\gamma = 0$ and $\gamma = \pi/2$). The $h$ value also affects the $\gamma$-periodicity of the landscape: the period of the oscillations in the landscape as a function of $\gamma$ 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 $\gamma_{\text{min}}$ for the single-node contribution in Eq. 12 reads,

$$\langle C_i \rangle (\beta_{\text{min}}, \gamma_{\text{min}}) = \frac{\sqrt{d}}{d+1} \left( 1 + \frac{1}{d} \right)^{-\left( d-1 \right)/2}. \quad (55)$$

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 $\beta$ values of the plots now range between $-\pi/2$ and $\pi/2$. For $h = 2$ we observe that the periodicity in $\gamma$ is divided by half, and there is a narrowing of the minimum and maximum peaks in $\gamma$. 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 $\gamma$, 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

$$[\cos(2\gamma h)^o \cos(2\gamma J)^p + \cos(2\gamma h)^q \cos(2\gamma J)^r]. \quad (56)$$

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 $\gamma$, 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 \( \gamma_{\min} \):

\[
\gamma_{\min} \approx \frac{1}{2J} \arctan \left( \frac{1}{\sqrt{d}} (1 - 0.15(r - 1)) \right).
\] (57)

For \( r > 5 \),

\[
\gamma_{\min} \approx \frac{1}{2r^{0.6} J} \arctan \left( \frac{1}{\sqrt{d}} \right).
\] (58)

These equations show that the relative increase of \( h_i \) with regards to \( J_{ij} \) is going to make \( \gamma_{\min} \) 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 \( \gamma_{\min} \) we observe that it decreases monotonically with \( r \). Thus, an increase in \( r \) will not only reduce \( \gamma_{\min} \) 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 \( \gamma_{\min} \),

\[
\gamma_{\min} \approx \frac{1}{2rh} \arctan \left( \frac{1}{\sqrt{d}} \right).
\] (59)

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

\[
\langle C_i \rangle \beta_{\min, \gamma_{\min}} = \left( 1 + \frac{1}{d} \right)^{-\frac{1}{2}} \sin \left( \frac{\arctan(1/\sqrt{d})}{r} \right).
\] (60)

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 \( \gamma_{\min} \) 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., \( h_i \in \{-h, \ldots, h\} \), 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, \( \text{QAOA} \) 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.

The objective of this paper is not to claim quantum advantage at \( p = 1 \) but to present tools to investigate the expectation value landscapes present in \( \text{QAOA} \). Although our paper is limited to the study of a single layer, we see value in applying the findings presented here to higher \( p \) versions of \( \text{QAOA} \).

### IV. CONCLUSIONS

We introduced an analytical formula for the expectation value of single-layer \((p = 1)\) \( \text{QAOA} \) circuits solving general Ising Hamiltonians, which allows for an efficient classical calculation of the optimal angles \( \beta_{\min}, \gamma_{\min} \). 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 0.252 – 0.389; for the Ising problems with external fields that we tested, the ratio was in the range 0.318 – 0.39. In order to compute these ratios—since we do not know the true optimal energy for each of such large Ising instances—we derive 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 \( \approx 0.39 \).

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 \( \geq 0.85 \)—meaning that we can expect the QAOA to find MAX-CUT solutions for these instances that have at least 85\% 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 \( \sim 85\% \) 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 [22].

We also note that the multi-layer QAOA cannot perform worse than the single-layer QAOA, 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, 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 III and IV). 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 \( \gamma \) parameter, \( \gamma_{\text{min}} \). 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 \( \gamma_{\text{min}} \) 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: [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, R. Parrish and E. Farhi 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.

VI. DATA AVAILABILITY

The I-set dataset, which comprises Ising instances generated from a subset of the G-set [10] instances by adding external fields, and two additional large Ising instances unrelated to G-set, is available at [https://doi.org/10.5281/zenodo.5258011](https://doi.org/10.5281/zenodo.5258011).

[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.
Appendix A: Derivation of Analytical Expression for General Expectation

Given an undirected graph \( G = (V, E) \) with vertices \( V = \{1, \ldots, n\} \) 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

\[
C = \sum_{i \in V} h_i s_i + \sum_{(i,j) \in E} J_{ij} s_i s_j, \tag{A1}
\]

This cost function is equivalent to the cost Hamiltonian

\[
H_C = \sum_{i \in V} C_i + \sum_{(i,j) \in E} C_{ij}, \tag{A2}
\]

where

\[
\begin{cases}
C_i = h_i \sigma_i^z \\
C_{ij} = J_{ij} \sigma_i^z \sigma_j^z
\end{cases} \tag{A3}
\]
The unitary transformation $U$ for a QAOA circuit with depth $p = 1$ is defined as

$$ U = U_B^g U_C^g $$

where

$$ U_C^g = e^{-i H_C} = e^{-i (\sum_i J_{ij} \sigma_i^z \sigma_j^z + \sum h_i \sigma_i^z)} 
= \prod_{(i,j) \in E} \left( \cos(J_{ij} \gamma) \sigma_i^0 \sigma_j^0 - i \sin(J_{ij} \gamma) \sigma_i^z \sigma_j^z \right) 
\prod_{i \in V} \left( \cos(h_i \gamma) \sigma_i^0 - i \sin(h_i \gamma) \sigma_i^z \right) $$

$$ U_B^g = e^{-i \beta H_B} = e^{-i \beta \sum_i \sigma_i^z} 
= \prod_{i \in V} \left( \cos(\beta) \sigma_i^0 - i \sin(\beta) \sigma_i^z \right). $$

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

$$ |\beta, \gamma\rangle = U |\psi_0\rangle. $$

The expectation value of $H_C$ for this final state reads

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

where

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

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

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

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 $\sigma^y$ terms:

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

Note that the terms in $U_C^{-\gamma}$ that do not involve the vertices $i$ and $j$ commute with $\sigma_i^z \sigma_j^z$, hence those terms will cancel each other. When evaluating $U_C^{-\gamma} \sigma_i^z \sigma_j^z U_C^\gamma$, 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 $\rho_0 = 2^{-n} \prod_{j=1}^n (\sigma_j^0 + \sigma_j^y)$, the trace $\text{Tr}[\rho_0 M]$ only depends on the $\{\sigma^0, \sigma^y\}^n$ components of $M$, that is:

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

Going back to the six traces, because $\sigma_i^z$ and $\sigma_j^z$ 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

\[ U_C^{-\gamma} \sigma_i^0 U_C^\gamma = \prod_{(i,k) \in E} (\cos(J_{ik}\gamma)\sigma_i^0 \sigma_k^0 - \cos(h_{ik}\gamma)\sigma_i^0 \sigma_k^0) \cdot (\cos(J_{ik}\gamma)\sigma_i^0 + \cos(h_{ik}\gamma)\sigma_i^0) \cdot (\cos(h_{ik}\gamma)\sigma_i^0 - \cos(J_{ik}\gamma)\sigma_i^0) \]

\[ \times (\cos(h_{ik}\gamma)\sigma_i^0 - i \sin(h_{ik}\gamma)\sigma_i^z) \prod_{(i,k) \in E} (\cos(J_{ik}\gamma)\sigma_i^0 \sigma_k^0 - i \sin(J_{ik}\gamma)\sigma_i^z \sigma_k^z) \]

\[ = \prod_{(i,k) \in E} (\cos(J_{ik}\gamma)\sigma_i^0 \sigma_k^0 + i \sin(J_{ik}\gamma)\sigma_i^z \sigma_k^z) \cdot (\cos(2h_{ik}\gamma)\sigma_i^z \sigma_k^z \prod_{(i,k) \in E} ((\cos(J_{ik}\gamma))^2 - (\sin(J_{ik}\gamma))^2)) \]

\[ = 2(\cos(2h_{ik}\gamma)\sigma_i^z \sigma_k^z - \sin(2h_{ik}\gamma)\sigma_i^0 \sigma_k^0) \prod_{(i,k) \in E} \cos(J_{ik}\gamma) \sin(J_{ik}\gamma) \]

Hence, there is only term that has a non-zero trace and we obtain the following result for Equation A14:

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

\[ = \sin(2h_{ik}\gamma) \prod_{(i,k) \in E} \cos(2J_{ik}\gamma). \]

For trace Expressions A15 we use the fact that \( \rho_0 U_C^{-\gamma} \sigma_i^0 \sigma_j^0 U_C^\gamma = \rho_0 U_C^{-\gamma} \sigma_i^0 U_C^\gamma \sigma_j^0 \), and by some additional algebra similar to the calculation above we obtain

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

\[ = \cos(2h_{ij}\gamma) \sin(2J_{ij}\gamma) \prod_{(i,k) \in E, \ k \neq j} \cos(2J_{ik}\gamma). \]

By symmetry, we have, \( \rho_0 U_C^{-\gamma} \sigma_i^0 \sigma_j^0 U_C^\gamma = \rho_0 U_C^{-\gamma} \sigma_j^0 U_C^\gamma \sigma_i^0 \) and,

\[ \text{Tr}[\rho_0 \sigma_i^0 U_C^{-\gamma} \sigma_j^0 U_C^\gamma] = \cos(2h_{ij}\gamma) \sin(2J_{ij}\gamma) \prod_{(j,k) \in E, \ k \neq i} \cos(2J_{jk}\gamma). \]

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. A30 since a large number of terms is involved.

\[ \text{Tr}[\rho_0 U_C^{-\gamma} \sigma_i^0 \sigma_j^0 U_C^\gamma] = - \frac{1}{2} \prod_{(i,k) \in E} \cos(2\gamma J_{ik}) \prod_{(j,k) \in E} \cos(2\gamma J_{jk}) \]

\[ \times \left[ \cos(2\gamma (h_i+h_j)) \prod_{(i,k) \in E} \cos(2\gamma (J_{ik}+J_{jk})) - \cos(2\gamma (h_i-h_j)) \prod_{(i,k) \in E} \cos(2\gamma (J_{ik}-J_{jk})) \right] \]

For this derivation the following trigonometric identities were used,
The distribution that mimics picking a random string $z$ is defined by

$$\text{This is an exact result that holds for all polynomials } C.$$ 

Assume that $c$ be the random variable that produces strings from $\{-1, +1\}$ with uniform probability $2^{-n}$, then we want to estimate $M$ using the expectation

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

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 $\{-1, +1\}^n$ with uniform probability $2^{-n}$, then we want to estimate $M$ using the expectation

$$\hat{M} = \mathbb{E} \left[ \max(C(Z_1), \ldots, C(Z_{2^n})) \right].$$

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 an intuition 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 $\sigma$. Hence we reached the estimate

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

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

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

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

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

We reiterate that the above is an informal argument and can be made more rigorous in different ways, depending on the characteristics of the problem instance. For example, in the case of quadratic cost functions with coefficients $c_S \in \{0, 1\}$, in the limit of large $|V|$ and with $P^* \approx 0.76321$ the Parisi constant, we know through the work of [24] that the maximum $M$ can be estimated by

$$\hat{M} = 2P^* \sqrt{|V||E||(1 - 2|E|/|V|^2) \approx 1.52642 \sqrt{|V||E|}}, \text{ for sparse graphs with } 2|E|/|V|^2 \approx 0. \quad (B9)$$

For quadratic cost functions with coefficients $c_S \in \{-1, 0, +1\}$ and with the $-1$ and $+1$ couplings evenly distributed, another estimator directly follows from [25]

$$\hat{M} = P^* \sqrt{2|V||E|} \approx 1.07934 \sqrt{|V||E|}. \quad (B10)$$

Note how the informal estimator of Equation (B7) sits between the rigorous ones of Equations (B9) and (B11). A third estimator would also be needed for the I-set of Table IV. In this article we sidestep this diversity of rigorous estimators by using the informal estimators of Equations (B7) and (B8).