

Problem-size-independent angles for a Grover-driven quantum approximate optimization algorithmDavid Headley^{1,2,*} and Frank K. Wilhelm^{2,3}¹*Mercedes-Benz AG, Stuttgart, Germany*²*Theoretical Physics, Saarland University, 66123 Saarbrücken, Germany*³*Institute for Quantum Computing Analytics (PGI 12), Forschungszentrum Jülich, 52425 Jülich, Germany*

(Received 17 October 2022; accepted 15 December 2022; published 9 January 2023)

The quantum approximate optimization algorithm (QAOA) requires that circuit parameters are determined that allow one to sample from high-quality solutions to combinatorial optimization problems. Such parameters can be obtained using either costly outer-loop optimization procedures and repeated calls to a quantum computer or, alternatively, via analytical means. In this work, we consider a context in which one knows a probability density function describing how the objective function of a combinatorial optimization problem is distributed. We show that, if one knows this distribution, then the expected value of strings, sampled by measuring a Grover-driven, QAOA-prepared state, can be calculated independently of the size of the problem in question. By optimizing this quantity, optimal circuit parameters for average-case problems can be obtained on a classical computer. Such calculations can help deliver insights into the performance of and predictability of angles in QAOA in the limit of large problem sizes, in particular, for the number partitioning problem.

DOI: [10.1103/PhysRevA.107.012412](https://doi.org/10.1103/PhysRevA.107.012412)**I. INTRODUCTION**

Variational quantum algorithms aim to exploit the power of noisy, intermediate-scale quantum (NISQ) computers [1] through the use of parametrized quantum circuits. The quantum approximate optimization algorithm (QAOA) [2] provides a universal [3,4] *Ansatz* state, efficiently preparable on a quantum computer, measurements from which might provide a useful heuristic method for the approximate solution of problems in combinatorial optimization.

QAOA circuits utilizing standard—single-qubit- X —drivers applied to Grover-like marked-state problems, recover the asymptotic scaling of Grover’s algorithm [5]. Furthermore, it has been shown in the same unstructured search problem context that parameters in QAOA concentrate [6]. That is, for increasing problem size, the variance in the optimal angles for QAOA circuits tends to zero and optimal angles concentrate at terminal values [7]. When considering QAOA in its original incarnation of single-qubit- X drivers and problem Hamiltonians encoding combinatorial optimization problems such as Max-Cut, Max-3-Lin-2, Max- k -SAT, numerical studies have provided evidence that parameters also tend to concentrate, and, as such, can be determined for classes of similar problems without access to a quantum processing unit using classically tractable tensor network approaches [8].

That optimal angles in QAOA concentrate, for certain classes of problem, is useful due to the exponential cost of using classical optimizers in the number of parameters over which they operate. Should QAOA provide a good heuristic method for a number of layers greater than ≈ 10 , it is likely

that naive global classical optimization with repeated calls to the quantum computer to estimate and extremize the expectation value of the problem Hamiltonian is not the method used, due to the excessively large number of calls to the quantum computer required. Alternative methods to calculate optimal angles might use numerical methods on smaller or simplified problem sizes, as aforementioned, or use the averaged behavior of problems at limiting size. Using this second approach, Farhi *et al.* demonstrated that for Sherrington Kirkpatrick type spin glasses and similar models, using single-qubit- X drivers, one can average over model instances to find optimal angles in a calculation that is independent of the number of spins in a model [9]. This provides a formula that has been numerically optimized to a depth of $p = 10$, a result later generalized [10] to show that the same technique extends to problems involving clauses incorporating more than two variables, providing an expression numerically tractable up to a depth $p = 20$ with the specific hardware used. This work was later widened to the context of Max-Cut problems on varying random graph families [11], with Claes *et al.* considering QAOA on infinite size mixed-spin SK models [12] and showing that the expected performance of QAOA also concentrates in this context at depth 1.

This paper takes an alternative approach to finding angles in the large problem-size limit. By using a Grover-style driver that bestows maximal permutation symmetry to the expectation value of a problem Hamiltonian under the QAOA state produced, we simplify the calculation of optimal angles at large n . This procedure requires that the probability density function (pdf) or density of states of the problem to be solved is known. In particular, we provide a method to find optimal angles and their corresponding expectation values on a classical computer, such that high-quality solutions can be sampled from a quantum computer without the need for outer-loop

*DavidKeithHeadley@gmail.com

optimization of the on-device QAOA state. We find that this method can be applied to the number partitioning problem (NPP) and other problems where the probability density function is known or can be approximated.

The *Ansätze* used in variational quantum algorithms, in general, must have problem specificity to avoid the fate of barren plateaus [13,14]. In QAOA, this problem specificity enters into the *Ansatz* twice. First, the problem Hamiltonian takes its eigenvalues directly from the objective function of the problem to be solved. Second, and more subtly, the driver Hamiltonian biases certain transitions between computational basis states over others. While this does not mean that a driver Hamiltonian is, in general, taken to be dependent on the problem solved, the driver induces a distance metric on the solution space of the problem, with respect to which the problem might exhibit structure that the algorithm can utilize. In the case of the single-qubit- X driver, this induced metric is that of the Hamming distance between two strings. The use of Grover drivers removes the second of these problem dependencies—inducing a metric under which any pair of nonidentical states can be considered as equidistant—and allows one to quantify where the performance in QAOA lies between that of algorithms that use amplitude amplification-like structure agnostic speedups to solve problems and those which derive performance from classically exploitable problem structure with respect to the single-qubit basis.

Other works have used Grover drivers in the context of QAOA. It has been shown that combinatorial optimization problems can be approximately solved using Grover's algorithm via the compilation of their objective function to a threshold function. Grover's algorithm is asymptotically optimal for the Grover problem [15], with the performance of the threshold-based algorithm numerically observed to outperform the nonthreshold version, while, however, incurring additional compilation overhead [16–18]. Though one might not *a priori* know which threshold to use for such a compiled version, knowledge of the pdf of the problem, as used in this work, would certainly be sufficient knowledge for the selection thereof.

While it is not expected that Grover-driven QAOA should provide superior performance to that of well-suited problem-specific drivers, it is expected that Grover-driven QAOA will not suffer from the issue of exponentially diminishing gradients—barren plateaus—in the case that the drivers are not well suited to the problem. As such, Grover-driven QAOA provides a point of reference to the performance of a QAOA-type algorithm in the very average case of driver-problem synergy.

II. THEORY

QAOA is performed via p applications of driver and problem Hamiltonians applied for parametrized times $\vec{\beta}, \vec{\gamma}$. Applying such Hamiltonians to a state prepared in an equal superposition of computational basis states, one produces

$$|\vec{\beta}, \vec{\gamma}\rangle = \prod_{p'=1}^p e^{i\beta_{p'}\hat{H}_D} e^{i\gamma_{p'}\hat{H}_P}|+\rangle^{\otimes n}, \quad (1)$$

which is the problem-dependent QAOA *Ansatz* state, depending on $2p$ free parameters. One is then required to find angles such that when sampling from the computational basis measurement outcome distribution of the *Ansatz* state, one finds configurations corresponding to small or large eigenvalues of \hat{H}_P depending on whether a problem is one of maximization or minimization. To this end, the expectation,

$$E_p(\vec{\beta}, \vec{\gamma}) = \langle \hat{H}_P \rangle_{\vec{\beta}, \vec{\gamma}} = \langle \vec{\beta}, \vec{\gamma} | \hat{H}_P | \vec{\beta}, \vec{\gamma} \rangle, \quad (2)$$

should be extremized with respect to the variational parameters. Measurements of the optimized *Ansatz* state then correspond to high-quality solutions to a binary-variable combinatorial optimization problem specifying \hat{H}_P . A combinatorial optimization problem on n binary variables has objective function $C(z)$ with $z \in \{0, 2^n - 1\}$ from which the problem Hamiltonian is defined as

$$\hat{H}_P = \sum_z C(z) |z\rangle \langle z|, \quad (3)$$

but more typically written in a Hadamard-transformed clause representation as

$$\hat{H}_P = \sum_{k=0}^{2^n-1} g(k) Z^{k_1} Z^{k_2} \dots Z^{k_n}, \quad (4)$$

where k_i is the i th bit of the binary representation of k . Typical problems for QAOA utilize low-weight interactions such that compilation to near-term quantum computers allows for state preparation within the coherence time of a device. Such problems include Erdős Rényi random Max-Cut problems [19] in which one takes

$$g(k) \sim \begin{cases} \text{Bernoulli}(f), & \text{if } \text{weight}(k) = 2, \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

for filling factor f and $\text{weight}(k)$ denoting the bit-weight of the label k . SK models are defined similarly with weights being sampled from a Gaussian distribution rather than the Bernoulli distribution of Max-Cut [20]. Also of interest are max- k -SAT type problems [21] in which terms up to weight k are present, taking integer values. Number partitioning problems, considered in Sec. VI B, consist of all-to-all coupled Mattis type Ising spin glasses in which the weights of ZZ interactions take the product of random variables associated with the two constituent bits. The descriptions $g(k)$ and $C(z)$ are related via an n -dimensional binary-variable Fourier transform, otherwise known as the Walsh-Hadamard transform [22].

A. Properties of Grover drivers

The driver Hamiltonian \hat{H}_D typically consists of single-qubit X operations,

$$\hat{H}_D = \sum_{i=1}^n X_i. \quad (6)$$

Various works have investigated the use of drivers that differ from this standard single-qubit driver. For example, a version of QAOA that samples from a differing selection of higher-order drivers for each layer has been shown to improve QAOA convergence [23,24]. Drivers can be used to shift complexity

from problem to driver Hamiltonian in the case of problems of hard and soft constraints [25].

In this work, an alternate driver—that which generates the Grover mixing operator—is used. The Grover driver can be written as a projector onto a Hadamard basis product state $|+\rangle^{\otimes n}$:

$$\hat{H}_G = |+\rangle^{\otimes n} \langle +|^{\otimes n} = \left[\frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n-1} |z\rangle \right] \left[\frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n-1} \langle z| \right]. \quad (7)$$

Importantly, this Hamiltonian is invariant under the re-labeling of any pair of states, i.e., under the permutation operator $\hat{U}_{j \leftrightarrow k}$ with

$$\hat{U}_{j \leftrightarrow k} = \mathcal{I} - |j\rangle \langle j| - |k\rangle \langle k| + |k\rangle \langle j| + |j\rangle \langle k|, \quad (8)$$

and so it holds that

$$\hat{H}_G = \hat{U}_{j \leftrightarrow k} \hat{H}_G \hat{U}_{j \leftrightarrow k} \quad (9)$$

for any permutation. Moreover, the initial state of QAOA bears the same property; as an eigenstate of any permutation operator one can write

$$|+\rangle^{\otimes n} = \hat{U}_{j \leftrightarrow k} |+\rangle^{\otimes n}. \quad (10)$$

Of course, a problem Hamiltonian is, in general, not invariant under any permutation. However, alongside a Grover-QAOA state [defined as in Eq. (1) using with $H_D = H_G$] its expectation is, since permutation operators may commute through the driver to annihilate or act as the identity on the initial state. So, for any permutation of a problem Hamiltonian \hat{H}_P , given by $\hat{P}^\dagger \hat{H}_P \hat{P}$ where

$$\hat{P} = \hat{U}_{j_m \leftrightarrow k_m} \cdots \hat{U}_{j_2 \leftrightarrow k_2} \hat{U}_{j_1 \leftrightarrow k_1} \quad (11)$$

is a sequence of permutations, the QAOA problem expectation produced by such a permuted problem Hamiltonian can be shown to be equivalent to that of the unpermuted Hamiltonian,

$$\langle \hat{P}^\dagger \hat{H}_P \hat{P} \rangle_{\bar{\beta}, \bar{\gamma}} = \langle \hat{H}_P \rangle_{\bar{\beta}, \bar{\gamma}}, \quad (12)$$

as shown in Appendix A. The consequence of this symmetry is that any two problems which have the same distribution of energy levels will have identical expectation values under Grover-driven QAOA states.

B. Problem objective functions as random variables

The only property of the problem Hamiltonian that therefore affects the expectation is its spectrum. The spectrum of a problem Hamiltonian can be seen as sampling from some random distribution with probability density function $f(c)$ —describing the relative likelihood that the random variable C , modeling a problem objective function, takes real value c —analogous to the density of states' function of solid state physics. Our key result is that as this pdf can be, in general, independent of the size of a problem, angles can be calculated directly from the pdf, rather than for any model instance with a given size. By computing angles for the large- n limit of Grover driven QAOA, one can determine the asymptotic performance of the algorithm for large system sizes without suffering from unfavorable scaling in the number of qubits. The computed angles are for “typical” instances. However, as

the problem size increases, the law of large numbers rapidly reduces the sample-to-sample fluctuations.

Consider a problem's objective function that has been sorted from low to high, denoted by $C_{\text{sort}}(z)$, such that

$$C_{\text{sort}}(x) \geq C_{\text{sort}}(y) \quad \text{if } x > y. \quad (13)$$

To this quantity, we can associate a quantile distribution function (qdf)

$$F^{-1}(p) = C_{\text{sort}}(z) \quad \text{for } p = \frac{z}{N} \quad (14)$$

denoting the value of C at its p th quantile. The inverse of the quantile distribution function is the cumulative distribution function (cdf) $F(c)$ with an associated pdf $f(c)$ in the continuum limit of large n :

$$F(c) = P(C < c) = \sum_{c'=-\infty}^c P(C = c') \approx \int_{-\infty}^c f(c') dc'. \quad (15)$$

The quantity on which the Grover-driven QAOA expectation value depends is the Fourier transform of this pdf, known as the characteristic function Γ which for a random variable X is defined as the expected value of $e^{i\gamma X}$ and so for the random variable C modeling an ensemble of objective functions for a given problem:

$$\Gamma(\gamma) = E[e^{i\gamma C}]. \quad (16)$$

This quantity, associated with the random variable C , can be expressed in terms of the pdf $f(c)$, the cdf $F(c)$, and the qdf $F^{-1}(p)$ as

$$\begin{aligned} \Gamma(\gamma) &= \int_{-\infty}^{\infty} f(c) e^{i\gamma c} dc \\ &= \int_{-\infty}^{\infty} e^{i\gamma c} dF(c) = \int_0^1 e^{i\gamma F^{-1}(p)} dp \end{aligned} \quad (17)$$

and is related via a Wick rotation and scaling factor of the system size to the partition function of the model.

A finite-size, n -qubit problem considered for QAOA is modeled here as taking $2^n = N$ samples from the ensemble random variable C , and as such the approximation made in this work is that the mean exponentiated objective function value can be replaced by the ensemble characteristic function Γ . That is, that

$$\Gamma(\gamma) \approx \sum_z \frac{e^{i\gamma C(z)}}{N} \quad (18)$$

with equality in the infinite N limit. Due to the exponential growth of N with the number of qubits, such an approximation rapidly improves with increasing system sizes. As such, it suffices to replace the finite sum in Eq. (18) by the characteristic function of the continuous pdf—from which the problem is modeled to sample from—in calculations.

III. DEPTH 1

In this section, we wish to obtain an expression for the expectation value of the depth $p = 1$ Grover-QAOA problem Hamiltonian $E_1(\gamma, \beta)$ in which the problem is expressed only

via a characteristic function Γ . We denote the mean value of a random variable C as

$$\bar{C} = \int_{-\infty}^{\infty} c f(c) dc \stackrel{n \rightarrow \infty}{\cong} \frac{1}{N} \sum_z C(z) = -i\Gamma'(0), \quad (19)$$

which can in general be set to zero by adding a constant to the objective function. Such a shift corresponds only to a global phase of the effected problem Hamiltonian unitary $e^{-i\bar{C}\gamma\mathcal{I}}$. The expression derived in full detail in Appendix B consists of factors for which Eq. (18) can immediately be substituted, with, however, one instance of $C(z)$ as a multiplicative factor rather than in an exponent. To substitute this factor, one may differentiate under the sum or integral in the characteristic function as

$$\begin{aligned} \frac{1}{N} \sum_z C(z) e^{iyC(z)} &= -\frac{i}{N} \frac{d}{d\gamma} \sum_z C(z) e^{-iyC(z)} \\ &= -i \frac{d}{d\gamma'} \Gamma(\gamma') \Big|_{\gamma} = -i\Gamma'(\gamma). \end{aligned} \quad (20)$$

Substituting the characteristic function Γ and its derivative above, one obtains the expression

$$E_1(\gamma, \beta) = \bar{C}(1 + BB^*\Gamma\Gamma^*) + 2 \operatorname{Im}(B^*\Gamma^*\Gamma'), \quad (21)$$

in which the first term can be set to zero by zeroing the mean value of the problem.

IV. DEPTH 2

The expectation value of the problem Hamiltonian for a $p = 2$ Grover-driven QAOA state depends on four parameters and can be expressed as

$$E_2(\gamma_1, \gamma_2, \beta_1, \beta_2) = \langle \gamma_1, \gamma_2, \beta_1, \beta_2 | \hat{H}_P | \gamma_1, \gamma_2, \beta_1, \beta_2 \rangle, \quad (22)$$

which, if one assumes that the mean of the distribution is zero, can be expanded in 10 terms seen in Sec. V with the help of computer algebra [26]. Further simplification and substitution of the depth 1 expression yields

$$\begin{aligned} E_2(\gamma_1, \gamma_2, \beta_1, \beta_2) &= E_1(\gamma_1, \beta_1) + E_1(\gamma_1 + \gamma_2, \beta_2) \\ &\quad + |B(\beta_1)|^2 |\Gamma(\gamma_1)|^2 E_1(\gamma_2, \beta_2) \\ &\quad + 2 \operatorname{Im} B(\beta_1)^* B(\beta_2)^* \Gamma(\gamma_1)^* \Gamma(\gamma_2)^* \Gamma'(\gamma_1 + \gamma_2) \\ &\quad + 2 \operatorname{Im} B(\beta_1) B(\beta_2)^* \Gamma(\gamma_1) \Gamma'(\gamma_2) \Gamma(\gamma_1 + \gamma_2)^*, \end{aligned} \quad (23)$$

which can be numerically extremized on a classical computer to find optimal parameters given a characteristic function Γ .

V. DEPTH p

For arbitrary depth p , we have the problem Hamiltonian expectation value of

$$E_p(\vec{\gamma}, \vec{\beta}) = \langle + | \prod_{j=-1}^{-p} \hat{U}_P(\gamma_j) \hat{U}_D(\beta_j) \hat{H}_P \prod_{i=1}^p \hat{U}_D(\beta_i) \hat{U}_P(\gamma_i) | + \rangle, \quad (24)$$

where we have introduced the convention that negative indices simply add a negative sign to the value, as

$$\gamma_{-i} = -\gamma_i, \quad \beta_{-i} = -\beta_i. \quad (25)$$

This can be expanded and expressed as a formula summing over 2^{2^p} terms expressed in factors of the characteristic function Γ and the driver-dependent function B . The full derivation and explanation of the notation used can be found in Appendix D with the expression for the arbitrary depth expected value of the problem Hamiltonian taking the form

$$\begin{aligned} E_p(\vec{\gamma}, \vec{\beta}) &= \sum_{k_{\text{bra}} < k_{\text{ket}} = 0}^{2^p - 1} 2 \operatorname{Im} \prod_{P \in \mathcal{P}_{\text{bra}}} \Gamma \left(\sum_{i \in P} \gamma_i \right) \Gamma' \left(\sum_{i \in \mathcal{P}_{\text{central}}} \gamma_i \right) \\ &\quad \times \prod_{P \in \mathcal{P}_{\text{ket}}} \Gamma \left(\sum_{i \in P} \gamma_i \right) \prod_{-j | k_{\text{bra}}^j = 1 \text{ or } j | k_{\text{ket}}^j = 1} B_j. \end{aligned} \quad (26)$$

VI. PROBLEM DISTRIBUTIONS

To find suitable problems for this algorithm, one requires problems for which the probability density function is known, but for which no efficient classical algorithms for finding states of extreme energy exist. The standard QAOA problems such Max-Cut, k -SAT, and SK models approximately follow binomial and Gaussian distributions, but the presence of frustration in these problems results in distributions that, to the knowledge of the authors, do not have simple analytical forms. The random cost model and number partitioning problem do not feature the same frustration, with the random variables by which they are defined entering into the problem via single-qubit terms, with the number partitioning problem made nontrivial via a global constraint enforcing positivity of the cost function. As such, the distributions for the latter problems are known.

A. Random cost model

The random cost model (RCM) is a toy model used as a testing ground for techniques in the study of disordered systems as is the simplest model exhibiting a phase transition [20]. The model is not a description of any physical system, nor does it describe a hard problem. However, it does provide an example for which our procedure generates a concise analytical expression at low depth. The RCM is defined as one in which each energy level or objective function value samples from a normal distribution, so

$$C \sim \text{normal}(\mu = 0, \sigma^2 = 1), \quad f(c) = \frac{1}{\sqrt{2\pi}} e^{-\frac{c^2}{2}}. \quad (27)$$

When using the Grover driver, a problem with this pdf can also be obtained via n independent spins with Hamiltonian

$$\sum_i g_i Z_i = \sum_z \sum_i (-1)^{z_i} g_i |z\rangle \langle z|, \quad (28)$$

in which each Pauli spin weight is set as $g_i \sim \text{normal}(\mu = 0, \sigma^2 = 1/n)$ such that the resulting angles and energy expectations are independent of problem size. This Hamiltonian is, of course, structured and for another choice of driver could not be usefully modeled as a Gaussian pdf alone; the Grover driver, however, does not see this additional structure. This problem yields the desired pdf as a normalized Gaussian with which is associated a characteristic function via Fourier

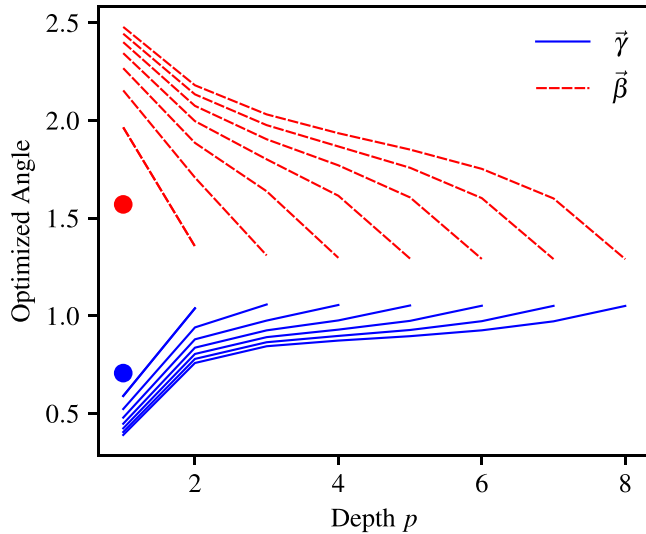


FIG. 1. Angles for the optimized infinite-size RCM are seen for QAOA depths 1 to 8. As the depth p increases, a similar pattern of optimal parameters for the limiting distribution of the random cost model emerges with the problem Hamiltonian's angles monotonically increasing and the driver Hamiltonian angles monotonically decreasing. Such behavior invites comparison to adiabatic quantum computation, overlooking the nonzero start point of γ and end point of β

transform with

$$\Gamma(\gamma) = e^{-\frac{\gamma^2}{2}}. \quad (29)$$

When substituted into Eq. (21), this characteristic function produces a $p = 1$ Grover-QAOA problem expectation value of

$$E_1(\gamma, \beta) = 2\gamma e^{-\gamma^2} \sin(\beta) \quad (30)$$

for which minimization yields optimal angles of

$$\gamma = \frac{\sqrt{2}}{2}, \quad \beta = \frac{\pi}{2}. \quad (31)$$

Angles for higher depth QAOA states can be found in Fig. 1 and the associated expectation values in Fig. 2. The best-known angles are determined via numerical optimization with BFGS [27] on a wide range of start points, in which the consistent behavior of the angles with increasing depth implies that with increasing depth one tends towards a continuous terminal schedule.

B. Number partitioning problem

The number partitioning problem (NPP) is that of finding a partition $z \in \{0 \dots N - 1\}$ of a set of n positive real numbers $\{x_0, \dots, x_{n-1}\}$ indexed by the set $\{0, \dots, n - 1\}$ into two subsets with indices $\{i|z_i = 0\}$ and $\{i|z_i = 1\}$ such that the sum of the numbers indexed by each subset is equal. NPP has applications ranging from task scheduling in multicore processing to choosing equally matched teams from a selection of players. It is an example of a problem for which classical heuristics such as simulated annealing do not perform well, notably, due to the similarity of the problem to totally unstructured cost

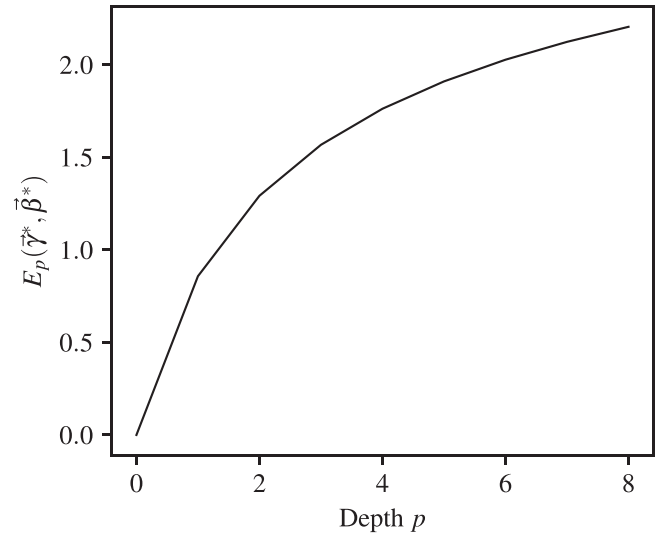


FIG. 2. Expectation value of the QAOA state for the random cost model at optimal angles exhibits an increasing solution quality at a decreasing rate.

problems [28]. The problem has seen study in the context of adiabatic quantum computing [29] due to the relative ease of computing-related properties in the large problem size limit, thus allowing the calculation of the minimum energy gap and asymptotic time complexity in this context. In the large problem-size limit, adiabatic quantum computing returns a Grover-like quadratic speedup in run time over random guessing for this problem.

One may define the decision version of the NPP as to satisfy the equality:

$$\sum_{i|z_i=0} x_i = \sum_{i|z_i=1} x_i, \quad (32)$$

which, for an approximate minimization version of the problem, is characterized by an objective function:

$$C(z) = \left| \sum_{i|z_i=0} x_i - \sum_{i|z_i=1} x_i \right| = \left| \sum_i (-1)^{z_i} x_i \right|. \quad (33)$$

When considering this problem in the context of QAOA, or other quantum optimization contexts in which many-body interactions are nontrivial to implement, one can consider an equivalent minimization problem defined by the squared problem $C^2(z)$. To express this as an Ising spin glass Hamiltonian, one can first consider the signed single qubit problem Hamiltonian,

$$\sum_i x_i Z_i = \sum_z \left(\sum_i (-1)^{z_i} x_i \right) |z\rangle \langle z|, \quad (34)$$

of which the square results in the problem Hamiltonian

$$\begin{aligned} & \sum_i x_i Z_i \sum_j x_j Z_j \\ &= \sum_i x_i^2 I + 2 \sum_{i<j} x_i x_j Z_i Z_j \end{aligned}$$

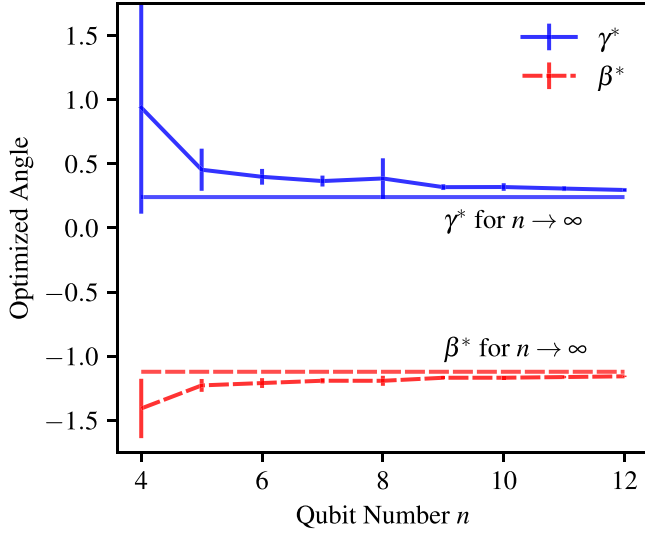


FIG. 3. For $p = 1$, optimal parameters of the NPP converge to the infinite size angle for increasing n . QAOA states for 30 problem instances are optimized and angles averaged for each point.

$$\begin{aligned}
 &= \sum_{z, z'} \left(\sum_i (-1)^{z_i} x_i \right) |z\rangle \langle z| \left(\sum_i (-1)^{z'_i} x_i \right) |z'\rangle \langle z'| \\
 &= \sum_z \sum_{i,j} (-1)^{z_i + z_j} x_i x_j |z\rangle \langle z| \\
 &= \sum_z C^2(z) |z\rangle \langle z|, \tag{35}
 \end{aligned}$$

which is a problem Hamiltonian encoding for the squared residue of the number partitioning problem, as

$$\begin{aligned}
 C^2(z) &= \left| \sum_i (-1)^{z_i} x_i \right| \left| \sum_j (-1)^{z_j} x_j \right| \\
 &= \sum_{i,j} (-1)^{z_i + z_j} x_i x_j. \tag{36}
 \end{aligned}$$

Consequently, the NPP Hamiltonian is an all-to-all coupled Mattis-type Ising spin glass in which couplings consist of the products of random variables. Now we must determine the pdf $f(c)$ and the associated partition function Γ for this problem. For convenience, we consider the numbers to be partitioned to be sampled from distributions such that the resulting cost function distribution has a unit mean. As such the numbers x_i are allowed to take values between 0 and $x_{\max} = 1/\sqrt{3n}$ uniformly, i.e., $x_i \sim U(0, x_{\max}) \forall i \in [1 \dots n]$. The probability density function for the individual variables is then defined as

$$f(c) = \frac{1}{x_{\max}} [\Theta(c) - \Theta(c - x_{\max})], \tag{37}$$

where Θ is the Heaviside step function. As such, the pdf for the distribution of residues defined as C_{SQ} , prior to squaring the cost function, is distributed as

$$\begin{aligned}
 C_{SQ} &\sim \sum_i (-1)^{z_i} x_i = \sum_i (-1)^{z_i} U(0, x_{\max}) \\
 &= \sum_i U(-x_{\max}, x_{\max}), \tag{38}
 \end{aligned}$$

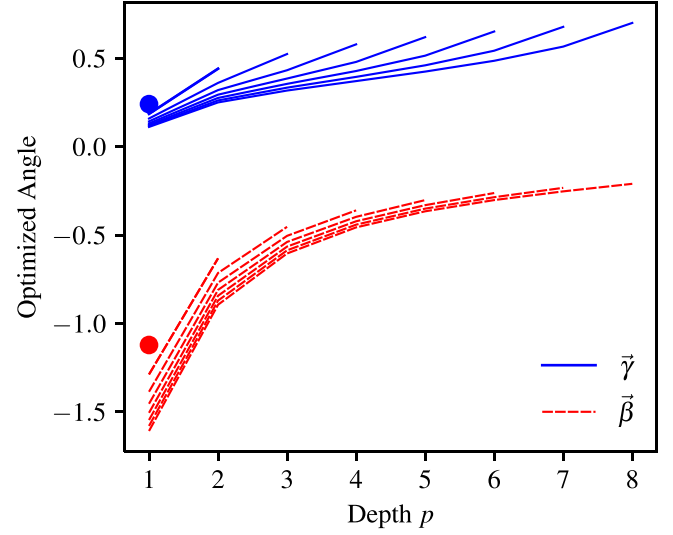


FIG. 4. Angles for the optimized infinite-size number partitioning problem are seen for QAOA depths 1 to 8. Optimal parameters follow annealing-reminiscent schedules and vary smoothly. Values are obtained here through the optimization of the formula derived in Appendix D via the BFGS optimizer and a random start point strategy.

where the factor $(-1)^{z_i}$ imparts a minus sign with probability $\frac{1}{2}$ and results in a sum over uniform distributions over the symmetric interval $[-x_{\max}, x_{\max}]$. This sum of independent, identically distributed random variables, by the central limit theorem, approaches a Gaussian distribution in the large n limit. So, in the large n limit

$$\begin{aligned}
 C_{SQ} &\sim \text{normal} \left(\sum_i E[U(-x_{\max}, x_{\max})], \right. \\
 &\quad \left. \sum_i \text{Var}[U(-x_{\max}, x_{\max})] \right) \\
 &= \text{normal}(\mu = 0, \sigma^2 = 1). \tag{39}
 \end{aligned}$$

For the squared single-qubit cost function $C_{sq}^2(z) = C^2(z)$, one can use the standard result that the square of a Gaussian distributed random variable is the χ^2 distribution with a single degree of freedom with pdf:

$$f(c) = \begin{cases} (2\pi c e^c)^{-\frac{1}{2}}, & c > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{40}$$

The characteristic function of this pdf can readily be calculated as

$$\Gamma(\gamma) = \sqrt{\frac{1}{1 - 2i\gamma}}. \tag{41}$$

For $p = 1$, one can numerically optimize the expectation value $E_1(\gamma, \beta)$ to find minimizing angles of

$$\gamma^* = 0.241, \quad \beta^* = 5.162 \tag{42}$$

to three decimal places. These angles can be seen in Fig. 3 in which the value of β has been shifted by 2π for convenience. Such angles yield an expected value of $E_1(\tilde{\gamma}^*, \tilde{\beta}^*) = 0.557$.

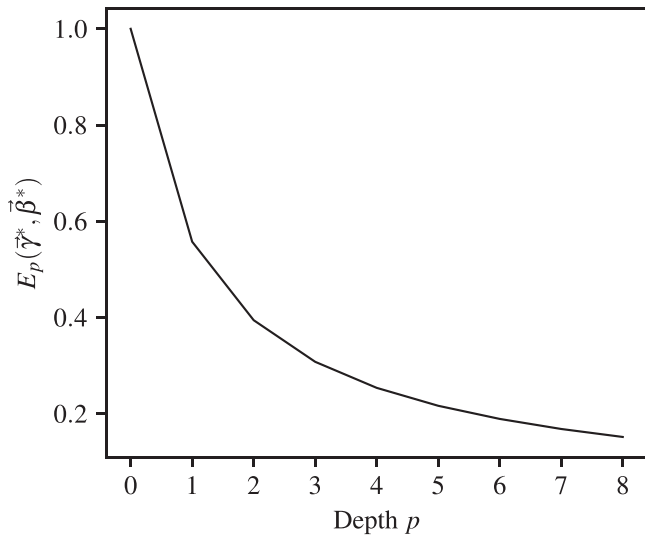


FIG. 5. Optimizing the Grover-QAOA state for the number partitioning problem at increasing depth p results in monotonically decreasing residue. The depth 0 value of 1 is added as a reference point attained by random sampling from the computational basis.

The expectation value of the problem Hamiltonian for this characteristic function can be minimized to obtain angles that are thought to be optimal; the best known angles can be seen in Fig. 4. The associated qualities of the solutions attained by these angles can be seen in Fig. 5. To demonstrate that these angles are appropriate for finite-size instances of the number partitioning problem, we plot in Figs. 3 and 6 the average result when optimizing finite-size instances at increasing size for depths 1 and 2, respectively. This demonstrates that with increasing problem size, the optimal angles for finite-size instances rapidly approach those of the analytically derived, ensemble average optimal angles. Finally, Fig. 7 demonstrates the convergence of the landscape defined by the problem

Hamiltonian expectation value of the QAOA state at a given point in the QAOA parameter space. As the number of qubits in the problem becomes larger, the QAOA solution quality landscape becomes indistinguishable from the analytically derived version.

VII. COMPILING THE GROVER DRIVER

Compilation of the Grover mixing oracle can be attained with or without ancillary qubits. Without ancillary qubits, the operation can be performed using an $(n - 1)$ -controlled-Z rotation. Such a gate can be compiled into two n -qubit Toffoli gates alongside three two-qubit gates [30]. These Toffoli gates can in turn be compiled using a number of two-qubit gates and depth quadratic and linear, respectively, in the number of qubits [31,32].

VIII. CONCLUSION

The power of QAOA as a heuristic for solving problems in combinatorial optimization is not yet certain and derives from some combination of Grover-like speedup alongside performance attributable to the synergy of the problem and driver used. In this work, we demonstrate a method to determine likely optimal QAOA angles in the case of a driver with average-case synergy, in which the problem structure cannot be exploited. For such problem-driver combinations, concentrations of optimal angles occur for increasing problem sizes and tend to a limit calculable via the pdf of the problem. It is expected that barren plateaus, limiting the ability of an optimizer to find good angles for QAOA, will affect problem driver combinations in which synergy is not present (for example, problems with high-order terms and single-qubit drivers). The problem of barren plateaus—in essence, the No Free Lunch Theorem of quantum optimization [33]—necessitates that one finds heuristics for which drivers, and angles to use for a given problem instance. For this, the performance of a driver in comparison to the Grover version is instructive as to whether a strategy exploits problem structure and to what degree. The strategy of optimizing QAOA angles for statistically defined ensembles of problems, rather than instances thereof, comes with the added advantage of smoother QAOA objective function landscapes with fewer local extrema being the average over instance landscapes, as evidenced by Fig. 7.

This work originated from asking the question as to whether a single-qubit- X driver should outperform the Grover driver as an *Ansatz* for the infinite range models in the work of Farhi *et al.* [9]. While this work has not directly compared the performance of Grover QAOA on the SK model, a future direction could compare this, via the application of methods applied to infinite-range SK models, to the number partitioning problem, or by finding or approximating an expression for the limiting-size pdf of an SK-type model. Future work could also investigate the performance of subspace-type drivers, e.g., XY type in problems of hard and soft constraints of limiting size [25], in which the form of the drivers on the valid-solution subspace differs from that of Grover drivers or standard single-qubit mixers. Work could also determine the dependence of QAOA performance on the shape of the

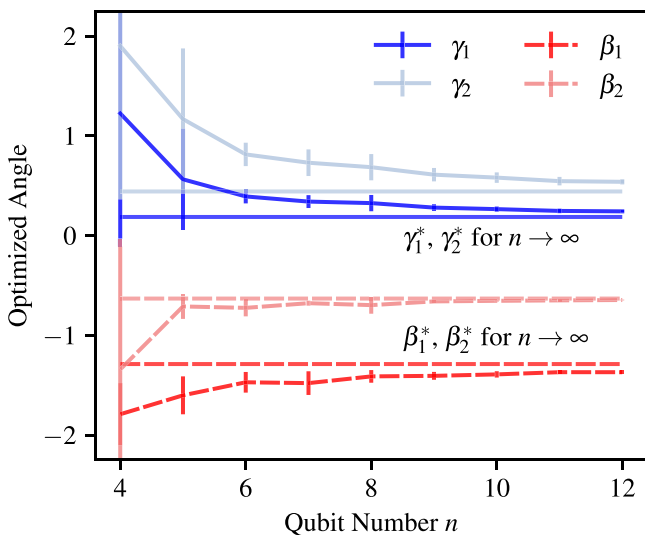


FIG. 6. For $p = 2$, optimal parameters of the NPP converge to the infinite size angle for increasing n . QAOA states for 30 problem instances are optimized and angles averaged for each point.

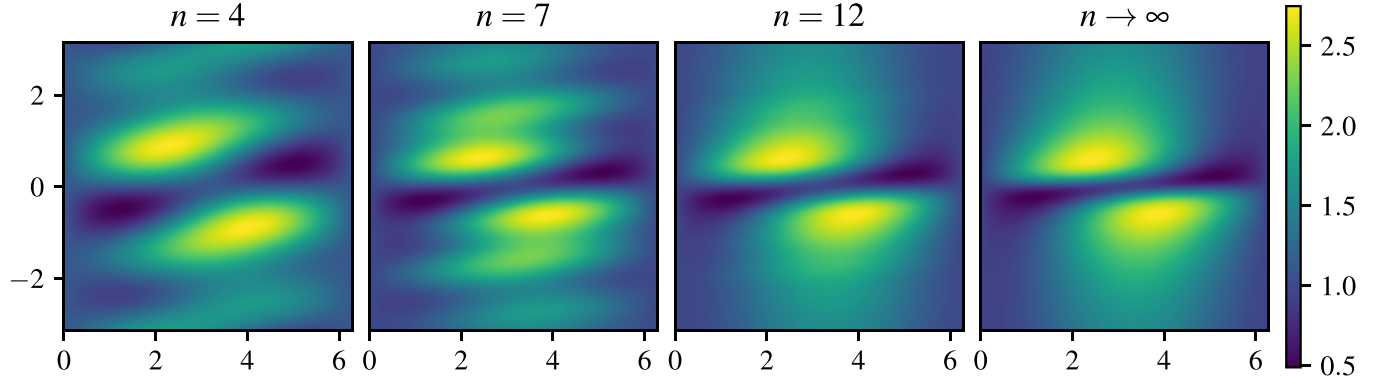


FIG. 7. For $p = 1$ the number partitioning problem landscape can be seen to converge to that of the large- n limit for $n = 5, 7, 12$. Periodicity is always observed in β , but as n increases, for low n , local maxima and minima resurge in γ due to common multiples in problem energy eigenvalues. For large n , increasing γ past a region near zero results in the expectation monotonically returning to the mean value of 1.

distribution of the unstructured random variable on which it acts, as has been considered in the context of nonadiabatic annealing [34]. Various works have provided analytical results for the use of Grover-type problems or mixing operators to solve search problems in the context of QAOA, quantum random walks, and adiabatic quantum computing [5,6,35,36], with this work demonstrating that similar analytical procedures can be applied to find optimal angles for Grover-driven QAOA with nonoracle problems. The work in this paper could also be applied to other contexts, such as that of quantum random walks, in which fully connected graphs have been previously considered for numerical studies on finding ground states of spin glasses [37].

ACKNOWLEDGMENTS

We acknowledge the support of Mercedes-Benz AG and the German Federal Ministry of Education and Research in the funding program “quantum technologies—from basic research to market” (Contract No. 13N15582). We thank D. Bagrets for bringing the number partitioning problem to our attention as a hard problem for which the probability density function is known.

APPENDIX A: PERMUTATION SYMMETRY OF PROBLEM HAMILTONIAN UNDER GROVER-QAOA STATE

The invariance of a QAOA problem Hamiltonian under permutation operator

$$\hat{P} = \hat{U}_{j_m \leftrightarrow k_m} \dots \hat{U}_{j_2 \leftrightarrow k_2} \hat{U}_{j_1 \leftrightarrow k_1} \quad (\text{A1})$$

can be demonstrated using the identities

$$\hat{P}\hat{P}^\dagger = \hat{I}, \quad e^{\hat{P}^\dagger \hat{O} \hat{P}} = \hat{P}^\dagger e^{\hat{O}} \hat{P}, \quad (\text{A2})$$

and therefore

$$\begin{aligned} \langle \hat{P}^\dagger \hat{H}_P \hat{P} \rangle_{\tilde{\beta}, \tilde{\gamma}} &= \langle + | \otimes^n \prod_{p'=1}^p [e^{i\gamma_{p'} \hat{P}^\dagger \hat{H}_P \hat{P}} \hat{e}^{i\beta_{p'} \hat{H}_D}] P^\dagger \hat{H}_P \hat{P} \prod_{p'=1}^p [e^{i\beta_{p'} \hat{H}_D} e^{i\gamma_{p'} \hat{P}^\dagger \hat{H}_P \hat{P}}] | + \rangle^{\otimes n} \\ &= \langle + | \otimes^n \prod_{p'=1}^p [\hat{P}^\dagger e^{i\gamma_{p'} \hat{H}_P} \hat{P} \hat{e}^{i\beta_{p'} \hat{H}_D}] P^\dagger \hat{H}_P \hat{P} \prod_{p'=1}^p [e^{i\beta_{p'} \hat{H}_D} \hat{P}^\dagger e^{i\gamma_{p'} \hat{H}_P} \hat{P}] | + \rangle^{\otimes n} \\ &= \langle + | \otimes^n P^\dagger \prod_{p'=1}^p [e^{i\gamma_{p'} \hat{H}_P} \hat{P} e^{i\beta_{p'} \hat{H}_D} \hat{P}^\dagger] \hat{H}_P \prod_{p'=1}^p [\hat{P} e^{i\beta_{p'} \hat{H}_D} \hat{P}^\dagger e^{i\gamma_{p'} \hat{H}_P}] \hat{P} | + \rangle^{\otimes n} \\ &= \langle + | \otimes^n P^\dagger \prod_{p'=1}^p [e^{i\gamma_{p'} \hat{H}_P} e^{i\beta_{p'} \hat{P} \hat{H}_D \hat{P}^\dagger}] \hat{H}_P \prod_{p'=1}^p [e^{i\beta_{p'} \hat{P} \hat{H}_D \hat{P}^\dagger} e^{i\gamma_{p'} \hat{H}_P}] \hat{P} | + \rangle^{\otimes n} \\ &= \langle \hat{H}_P \rangle_{\tilde{\beta}, \tilde{\gamma}}, \end{aligned} \quad (\text{A3})$$

demonstrating that the expectation value is invariant under any permutation.

APPENDIX B: CALCULATION OF DEPTH 1 EXPECTATION VALUE

We can write down the unitaries for the problem Hamiltonian,

$$\hat{U}_P(\gamma) = \sum_z e^{i\gamma C(z)} |z\rangle\langle z|, \quad (\text{B1})$$

and driver Hamiltonian,

$$\hat{U}_D(\beta) = e^{i\beta|+\rangle\langle +|} = \mathcal{I} + i\beta|+\rangle\langle +| + \frac{(i\beta)^2}{2}|+\rangle\langle +|\dots = \mathcal{I} + (e^{i\beta} - 1)|+\rangle\langle +| = \mathcal{I} + B(\beta)|+\rangle\langle +|, \quad (\text{B2})$$

with $B(\beta) = e^{i\beta} - 1$. For $p = 1$ the problem expectation can be written as

$$\begin{aligned} \langle \hat{H}_P \rangle_{\gamma, \beta} &= \langle + | \hat{U}_P^\dagger(\gamma) \hat{U}_D^\dagger(\beta) \hat{H}_P \hat{U}_D(\beta) \hat{U}_P(\gamma) | + \rangle \\ &= \sum_{z_{-1}, z_0, z_1} \langle + | e^{-i\gamma C(z_{-1})} | z_{-1} \rangle \langle z_{-1} | [\mathcal{I} + B^* | + \rangle \langle + |] C(z_0) | z_0 \rangle \langle z_0 | [\mathcal{I} + B | + \rangle \langle + |] e^{i\gamma C(z_1)} | z_1 \rangle \langle z_1 | | + \rangle \\ &= \frac{1}{N} \sum_{z_{-1}, z_0, z_1} e^{-i\gamma C(z_{-1})} \langle z_{-1} | [\mathcal{I} + B^* | + \rangle \langle + |] C(z_0) | z_0 \rangle \langle z_0 | [\mathcal{I} + B | + \rangle \langle + |] e^{i\gamma C(z_1)} | z_1 \rangle, \end{aligned} \quad (\text{B3})$$

which yields four terms as follows:

$$\bar{C} + \frac{1}{N^2} B B^* \bar{C} \sum_z e^{i\gamma C(z)} \sum_z e^{-i\gamma C(z)} + \frac{1}{N^2} B \sum_z e^{i\gamma C(z)} \sum_z C(z) e^{-i\gamma C(z)} + \frac{1}{N^2} B^* \sum_z e^{-i\gamma C(z)} \sum_z C(z) e^{i\gamma C(z)}. \quad (\text{B4})$$

One then can substitute the characteristic function of the ensemble to obtain an expression for the expectation value of the QAOA state, with

$$\frac{1}{N} \sum_z C(z) \rightarrow \bar{C}, \quad \sum_z \frac{e^{i\gamma C(z)}}{N} \rightarrow \Gamma(\gamma). \quad (\text{B5})$$

To express the term containing $C(z)$ not in the exponent, one may differentiate under the sum or integral as

$$\sum_z C(z) e^{i\gamma C(z)} = -i \frac{d}{d\gamma} \sum_z C(z) e^{-i\gamma C(z)} \rightarrow -i \frac{d}{d\gamma'} \Gamma(\gamma') \Big|_{\gamma} = -i \Gamma'(\gamma). \quad (\text{B6})$$

Resulting in the expression

$$\langle \hat{H}_P \rangle_{\gamma, \beta} = \bar{C}(1 + B B^* \Gamma \Gamma^*) + 2 \text{Im}(B^* \Gamma^* \Gamma'). \quad (\text{B7})$$

APPENDIX C: CALCULATION OF DEPTH 2 EXPECTATION VALUE

The $p = 2$ expression can be calculated similarly to the $p = 1$ expression, however, with more terms and thus requiring the use of a computer algebra package [26]. The result of expanding the expression for the expectation value, setting the mean \bar{C} to zero, and substituting the characteristic function and its derivative is the following expression of 10 terms:

$$\begin{aligned} E_2(\vec{\gamma}, \vec{\beta}) &= iB(\beta_1)B(\beta_2)\Gamma(\gamma_1)\Gamma(\gamma_2)B(\beta_1)^*\Gamma(\gamma_1)^*\Gamma(\gamma_2)^* - iB(\beta_1)\Gamma(\gamma_1)B(\beta_1)^*B(\beta_2)^*\Gamma(\gamma_1)^*\Gamma(\gamma_2)^*\Gamma'(\gamma_2) \\ &\quad - iB(\beta_1)^*\Gamma(\gamma_1)^*\Gamma'(\gamma_1) + iB(\beta_1)\Gamma(\gamma_1)\Gamma'(\gamma_1)^* + iB(\beta_1)B(\beta_2)\Gamma(\gamma_1)\Gamma(\gamma_2)\Gamma'(\gamma_1 + \gamma_2)^* \\ &\quad - iB(\beta_1)^*B(\beta_2)^*\Gamma(\gamma_1)^*\Gamma(\gamma_2)^*\Gamma'(\gamma_1 + \gamma_2) \\ &\quad - iB(\beta_1)\Gamma(\gamma_1)B(\beta_2)^*\Gamma'(\gamma_2)\Gamma(\gamma_1 + \gamma_2)^* + iB(\beta_2)B(\beta_1)^*\Gamma(\gamma_1)^*\Gamma'(\gamma_2)^*\Gamma(\gamma_1 + \gamma_2) \\ &\quad - iB(\beta_2)^*\Gamma(\gamma_1 + \gamma_2)^*\Gamma'(\gamma_1 + \gamma_2) + iB(\beta_2)\Gamma(\gamma_1 + \gamma_2)\Gamma'(\gamma_1 + \gamma_2)^*. \end{aligned} \quad (\text{C1})$$

This expression can be further simplified via the substitution of the $p = 1$ expectation resulting in the final expression in Eq. (23).

APPENDIX D: CALCULATION OF ARBITRARY DEPTH EXPECTATION VALUE

Using the intuition from the $p = 1$ and $p = 2$ cases, we can write down the expression for any term in the arbitrary depth expectation value of the problem Hamiltonian. We can write the full expectation $E_p(\vec{\gamma}, \vec{\beta})$ as

$$\langle \vec{\gamma}, \vec{\beta} | \hat{H}_P | \vec{\gamma}, \vec{\beta} \rangle = \langle + | \prod_{j=-p}^{-1} \hat{U}_P(\gamma_j) \hat{U}_D(\beta_j) \hat{H}_P \prod_{i=1}^p \hat{U}_D(\beta_i) \hat{U}_P(\gamma_i) | + \rangle, \quad (\text{D1})$$

where we have introduced the convention that negative indices simply add a negative sign to the value, as

$$\gamma_{-i} = -\gamma_i, \quad \beta_{-i} = -\beta_i. \quad (\text{D2})$$

The expectation value can be expanded, substituting the expressions for the driver and problem Hamiltonians as

$$\sum_{\vec{z}} \langle + | \prod_{j=-1}^{-p} [e^{i\gamma_j C(z_j)} |z_j\rangle \langle z_j| (I + B_j |+\rangle \langle +|)] C(z_0) |z_0\rangle \langle z_0| \prod_{i=1}^p [(I + B_i |+\rangle \langle +|) e^{i\gamma_i C(z_i)} |z_i\rangle \langle z_i|] |+\rangle, \quad (\text{D3})$$

where the sum over $2p + 1$ indices $\vec{z} = (z_{-p}, \dots, z_0, \dots, z_p)$ runs over the set $\{0, N - 1\}$ for each index. Expanding the products in this expression will yield 2^{2p} terms that can be sorted in the number of occurrences of a B variable in the bra or ket.

For the one term with zero B incidences, one finds, for example,

$$\frac{1}{N} \sum_{\vec{z}} e^{\sum_{j=-1}^{-p} i\gamma_j C(z_j)} C(z_0) e^{\sum_{i=1}^p i\gamma_i C(z_i)} \langle z_{-1} | \langle z_{-2} | \langle z_{-2} | \dots | z_{-p} \rangle \langle z_{-p} | z_0 \rangle \langle z_0 | z_p \rangle \langle z_p | \dots | z_2 \rangle \langle z_2 | z_1 \rangle, \quad (\text{D4})$$

in which the orthogonal bra-kets evaluate to zero leaving only one index yielding a nonzero term evaluating to the mean of the objective function \bar{C} . For weight-1 terms one has B in a position l

$$\sum_{\vec{z}} \langle + | \prod_{j=-1}^{-p} [e^{-i\gamma_j C(z_j)} |z_j\rangle \langle z_j|] C(z) |z_0\rangle \langle z_0| \prod_{i=l}^p [|+\rangle \langle +| e^{i\gamma_i C(z_i)} |z_i\rangle \langle z_i|] B_l \prod_{i=1}^{l-1} [|+\rangle \langle +| e^{i\gamma_i C(z_i)} |z_i\rangle \langle z_i|] |+\rangle, \quad (\text{D5})$$

for which the orthogonal bra-kets can be evaluated as

$$\frac{1}{N^2} \sum_{\vec{z}} e^{\sum_{j=-1}^{-p} i\gamma_j C(z_j)} C(z_0) e^{\sum_{i=l}^p -i\gamma_i C(z_i)} B_k \delta_{z_{-1}, z_{-2}} \dots \delta_{z_{-p}, z_0} \delta_{z_0, z_p} \dots \delta_{z_{k+2}, z_{k+1}} \delta_{z_k, z_{k-1}} \dots \delta_{z_2, z_1}, \quad (\text{D6})$$

in which the sum is eliminated for all but two partitions giving

$$\frac{1}{N^2} \sum_z e^{\sum_{j=-1}^{-p} i\gamma_j C(z)} C(z) e^{\sum_{i=k+1}^p i\gamma_i C(z)} \sum_z e^{\sum_{i=1}^k i\gamma_i C(z)} B_k \quad (\text{D7})$$

and leaving two variables over which we sum and leaving the nonexponentiated $C(z)$ and therefore the differentiated characteristic function Γ' in the factor corresponding to the partition containing the central variable z_0 :

$$-i B_k \Gamma' \left(\sum_{i=-1}^{-p} \gamma_i + \sum_{i=k+1}^p \gamma_i \right) \Gamma \left(\sum_{i=1}^k \gamma_i \right). \quad (\text{D8})$$

To expand the expression more generally, we can introduce indices $k_{\text{bra}} \in \{0 \dots 2^p - 1\}$ and $k_{\text{ket}} \in \{0 \dots 2^p - 1\}$ that are each p -bit numbers such that the presence of a 1 in the binary representation of the index implies a factor of $B_i |+\rangle \langle +|$ at the bit position in the bra or ket for the relevant term. The formula can be rewritten as

$$\sum_{k_{\text{bra}}, k_{\text{ket}}=0}^{2^p-1} \sum_{\vec{z}} \langle + | \prod_{j=-p}^{-1} [e^{i\gamma_j C(z_j)} |z_j\rangle \langle z_j| (B_j |+\rangle \langle +|)^{k_{\text{bra}}^j}] C(z_0) |z_0\rangle \langle z_0| \prod_{i=1}^p [(B_i |+\rangle \langle +|)^{k_{\text{ket}}^i} e^{i\gamma_i C(z_i)} |z_i\rangle \langle z_i|] |+\rangle. \quad (\text{D9})$$

When the index $k_{\text{ket}, \text{bra}}$ is zero, this expression $B_i |+\rangle \langle +|$ is the identity and adjacent nonidentical z indices result in a Kronecker delta, with $\langle z_j | z_k \rangle = \delta_{z_j, z_k}$ collapsing a sum over one of the z indices. As such, these terms are collected into a single sum. When a bit in the index $k_{\text{ket}, \text{bra}}$ is one, the projector $|+\rangle \langle +|$ annihilates with a basis state vector $\langle z_i |$ and simply produces a factor of $1/N$, with $\langle + | z \rangle = 1/\sqrt{N} \forall z$, leaving the sums over z intact.

To substitute the characteristic function, we note that unbroken chains of 0's of length n^* in the indices $k_{\text{ket}, \text{bra}}$ will result in factors of the form

$$\begin{aligned} \sum_{z_1, z_2, \dots, z_{n^*}} e^{i\gamma_1 C(z_1)} |z_1\rangle \langle z_1| e^{i\gamma_2 C(z_2)} |z_2\rangle \langle z_2| \dots e^{i\gamma_{n^*} C(z_{n^*})} |z_{n^*}\rangle \langle z_{n^*}| &= \sum_{z_1, z_2, \dots, z_{n^*}} \delta_{z_1, z_2} \delta_{z_2, z_3} \dots \delta_{z_{n^*-1}, z_{n^*}} e^{i\gamma_1 C(z_1)} e^{i\gamma_2 C(z_2)} \dots e^{i\gamma_{n^*} C(z_{n^*})} \\ &= \sum_z e^{iC(z) \sum_{i \in [1 \dots n^*]} \gamma_i} = \Gamma \left(\sum_{i \in [1 \dots n^*]} \gamma_i \right). \end{aligned} \quad (\text{D10})$$

To substitute the characteristic function for all incidences of the objective function $C(z)$, it is convenient to define a set of boundary indices between those that are grouped as above. We can define these as

$$S_{\text{bra}} = \{0\} + \{i | k_{\text{bra}}^i = 1\}, \quad S_{\text{ket}} = \{0\} + \{i | k_{\text{ket}}^i = 1\}, \quad (\text{D11})$$

defined by the locations of 1's in the binary representation of $k_{\text{bra}}, k_{\text{ket}}$. The additional boundary at 0 represents the boundary imposed by the initial state of QAOA. In the partitions formed by these boundaries, we collect the indices contained in the i th partition of the bra or ket as $P_{\text{bra}}^i, P_{\text{ket}}^i$ with

$$P_{\text{bra}}^i = \{-j | j > k_{\text{bra}}^i, j \leq k_{\text{bra}}^{i+1}\}, \quad P_{\text{ket}}^i = \{j | j > k_{\text{ket}}^i, j \leq k_{\text{ket}}^{i+1}\} \quad (\text{D12})$$

and collect these into a set of sets:

$$P_{\text{bra}} = \{P_{\text{bra}}^i | i \in [0, \text{weight}(k_{\text{bra}})]\}, \quad P_{\text{ket}} = \{P_{\text{ket}}^i | i \in [0, \text{weight}(k_{\text{ket}})]\}. \quad (\text{D13})$$

Finally, one defines a central partition spanning the indices between the last occurrence of 1 in the binary representation of $k_{\text{bra}}, k_{\text{ket}}$ and containing the nonexponentiated factor of the objective function

$$P_{\text{central}} = \{-j | j > \max S_{\text{bra}}\} + \{j | j > \max S_{\text{ket}}\}. \quad (\text{D14})$$

The expectation value after the substitution of the characteristic function and its derivative can then be expressed as

$$E_p(\vec{\gamma}, \vec{\beta}) = \sum_{k_{\text{bra}}, k_{\text{ket}}=0}^{2^p-1} -i \prod_{P \in P_{\text{bra}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \Gamma'\left(\sum_{i \in P_{\text{central}}} \gamma_i\right) \prod_{P \in P_{\text{ket}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \prod_{-j | k_{\text{bra}}^j=1} B_j \prod_{j | k_{\text{ket}}^j=1} B_j, \quad (\text{D15})$$

for which exchanging $k_{\text{bra}}, k_{\text{ket}}$ conjugates the term, allowing the expression to be split into two parts—one in which $k_{\text{bra}} = k_{\text{ket}}$ and one otherwise, so that

$$E_p(\vec{\gamma}, \vec{\beta}) = \sum_{k=0}^{2^p-1} -i \prod_{P \in P_{\text{bra}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \Gamma\left(\sum_{i \in P} \gamma_i\right)^* \Gamma'(0) \prod_{j | k^j=1} B_j B_j^* + \sum_{k_{\text{bra}} < k_{\text{ket}}=0}^{2^p-1} -i \prod_{P \in P_{\text{bra}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \Gamma'\left(\sum_{i \in P_{\text{central}}} \gamma_i\right) \prod_{P \in P_{\text{ket}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \prod_{-j | k_{\text{bra}}^j=1} B_j \prod_{j | k_{\text{ket}}^j=1} B_j + \text{conjugate}. \quad (\text{D16})$$

If the problem mean is zeroed by global phase, $\Gamma'(0) = 0$ and the first sum disappears, leaving

$$E_p(\vec{\gamma}, \vec{\beta}) = \sum_{k_{\text{bra}} < k_{\text{ket}}=0}^{2^p-1} 2 \text{Im} \prod_{P \in P_{\text{bra}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \Gamma'\left(\sum_{i \in P_{\text{central}}} \gamma_i\right) \prod_{P \in P_{\text{ket}}} \Gamma\left(\sum_{i \in P} \gamma_i\right) \prod_{-j | k_{\text{bra}}^j=1 \text{ or } j | k_{\text{ket}}^j=1} B_j. \quad (\text{D17})$$

A Python script to evaluate the Grover-QAOA objective function $E_p(\vec{\gamma}, \vec{\beta})$ for a given characteristic function is available upon request to the author.

[1] J. Preskill, *Quantum* **2**, 79 (2018).
 [2] E. Farhi, J. Goldstone, and S. Gutmann, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
 [3] S. Lloyd, [arXiv:1812.11075](https://arxiv.org/abs/1812.11075).
 [4] M. E. Morales, J. D. Biamonte, and Z. Zimborás, *Quantum Inf. Process.* **19**, 291 (2020).
 [5] Z. Jiang, E. G. Rieffel, and Z. Wang, *Phys. Rev. A* **95**, 062317 (2017).
 [6] V. Akshay, D. Rabinovich, E. Campos, and J. Biamonte, *Phys. Rev. A* **104**, L010401 (2021).
 [7] D. Rabinovich, R. Sengupta, E. Campos, V. Akshay, and J. Biamonte, *Mathematics* **10**, 2601 (2022).
 [8] M. Streif and M. Leib, *Quantum Sci. Technol.* **5**, 034008 (2020).
 [9] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, *Quantum* **6**, 759 (2022).
 [10] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou, [arXiv:2110.14206](https://arxiv.org/abs/2110.14206).
 [11] S. Boulebnane and A. Montanaro, [arXiv:2110.10685](https://arxiv.org/abs/2110.10685).
 [12] J. Claes and W. van Dam, *Quantum* **5**, 542 (2021).
 [13] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Nat. Commun.* **9**, 4812 (2018).
 [14] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, *Nat. Commun.* **12**, 1791 (2021).
 [15] C. Zalka, *Phys. Rev. A* **60**, 2746 (1999).
 [16] A. Bärttschi and S. Eidenbenz, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2020), pp. 72–82.
 [17] J. Golden, A. Bärttschi, D. O'Malley, and S. Eidenbenz, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2021), pp. 137–147.
 [18] J. Golden, A. Bärttschi, S. Eidenbenz, and D. O'Malley, [arXiv:2202.00648](https://arxiv.org/abs/2202.00648).
 [19] G. E. Crooks, [arXiv:1811.08419](https://arxiv.org/abs/1811.08419).
 [20] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).
 [21] V. Akshay, H. Philathong, M. E. S. Morales, and J. D. Biamonte, *Phys. Rev. Lett.* **124**, 090504 (2020).
 [22] S. A. Hadfield, *Quantum Algorithms for Scientific Computing and Approximate Optimization* (Columbia University, New York, 2018).
 [23] L. Zhu, H. L. Tang, G. S. Barron, F. A. Calderon-Vargas, N. J. Mayhall, E. Barnes, and S. E. Economou, *Phys. Rev. Res.* **4**, 033029 (2022).
 [24] Y. Chen, L. Zhu, C. Liu, N. J. Mayhall, E. Barnes, and S. E. Economou, *Quantum 2.0 Conference and Exhibition* (Optica Publishing Group, 2022), paper QM4A.2.

- [25] S. Hadfield, Z. Wang, E. G. Rieffel, B. O’Gorman, D. Venturelli, and R. Biswas, in *Proceedings of the Second International Workshop on Post Moores Era Supercomputing* (Association for Computing Machinery, New York, 2017), pp. 15–21.
- [26] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, v. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz, *PeerJ Comput. Sci.* **3**, e103 (2017).
- [27] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, 2nd ed. (Springer New York, 1999).
- [28] S. Mertens, *Phys. Rev. Lett.* **84**, 1347 (2000).
- [29] V. N. Smelyanskiy, U. v. Toussaint, and D. A. Timucin, [arXiv:quant-ph/0202155](https://arxiv.org/abs/quant-ph/0202155).
- [30] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, *Phys. Rev. A* **52**, 3457 (1995).
- [31] M. Saeedi and M. Pedram, *Phys. Rev. A* **87**, 062318 (2013).
- [32] Y. He, M.-X. Luo, E. Zhang, H.-K. Wang, and X.-F. Wang, *Int. J. Theor. Phys.* **56**, 2350 (2017).
- [33] D. H. Wolpert and W. G. Macready, *IEEE Trans. Evol. Comput.* **1**, 67 (1997).
- [34] B. Yan and N. A. Sinitsyn, *Nat. Commun.* **13**, 2212 (2022).
- [35] J. G. Morley, N. Chancellor, S. Bose, and V. Kendon, *Phys. Rev. A* **99**, 022339 (2019).
- [36] J. Roland and N. J. Cerf, *Phys. Rev. A* **65**, 042308 (2002).
- [37] A. Callison, N. Chancellor, F. Mintert, and V. Kendon, *New J. Phys.* **21**, 123022 (2019).