Symmetry-informed transferability of optimal parameters in the quantum approximate optimization algorithm

(article starts on next page)

Symmetry-informed transferability of optimal parameters in the quantum approximate optimization algorithm

Isak Lyngfelt ⬡* and Laura García-Álvarez ⬡

Department of Microtechnology and Nanoscience (MC2), Chalmers University of Technology, SE-412 96 Göteborg, Sweden

One of the main limitations of variational quantum algorithms is the classical optimization of the highly dimensional nonconvex variational parameter landscape. To simplify this optimization, we can reduce the search space using problem symmetries and typical optimal parameters as initial points if they concentrate. In this article, we consider typical values of optimal parameters of the quantum approximate optimization algorithm for the MAXCUT problem with d-regular tree subgraphs and reuse them in different graph instances. We prove symmetries in the optimization landscape of several kinds of weighted and unweighted graphs, which explains the existence of multiple sets of optimal parameters. However, we observe that not all optimal sets can be successfully transferred between problem instances. We find specific transferable domains in the search space and show how to translate an arbitrary set of optimal parameters into the adequate domain using the studied symmetries. Finally, we extend these results to general classical optimization problems described by Ising Hamiltonians, the Hamiltonian variational ansatz for relevant physical models, and the recursive and multiangle quantum approximate optimization algorithms.

I. INTRODUCTION

Variational quantum algorithms (VQAs) are among the most popular and accessible quantum algorithms for solving classically hard problems, with a notable example being the quantum approximate optimization algorithm (QAOA) [1], used to tackle combinatorial optimization problems (COPs). The QAOA is a hybrid algorithm consisting of a quantum circuit with several layers of parameter-dependent operations applied iteratively, followed by a classical optimization of the parameters. The quality of the solutions output by this algorithm depends on these classically optimized variational parameters and the depth of the quantum circuit. In theory, the solution improves as the circuit depth increases but, in practice, deep circuits introduce noise that affects the quantum state. However, shallow depth QAOA is often inferior to state-of-the art classical algorithms [2–6]. Therefore, we have to choose an algorithm depth that achieves an acceptable solution while maintaining quantum coherence. Alternatively, one can modify the standard QAOA to mitigate the effects of decoherence. Multiple alternative QAOA inspired algorithms adapted to specific COPs show better theoretical performance at lower depths, like the recursive QAOA (RQAOA) [7] or the multiangle QAOA (ma-QAOA) [8]. Yet their success also depends on finding good variational parameters.

The number of variational parameters scales linearly with the depth of the circuit, and the number of random initialization points needed to find the global optima is exponential in the number of parameters [9]. All QAOA variants suffer from this curse of dimensionality and, in general, the classical optimization routine of QAOA is known to be an NP-hard problem [10]. Barren plateaus in the optimization landscape further increase the difficulty in finding the optimal parameters [10–13], and it has been suggested that problems that are classically hard to simulate will always have these plateaus [14]. As a result, efficient strategies for finding optimal parameters are vital in the viability of QAOA [15].

Optimizing the variational parameters is an active field of research, and a multitude of techniques have been studied, including standard classical optimizers [16], gradient-free optimizers [16,17], classical genetic algorithms [18], machine-learning assisted optimization [19–22], and extrapolation techniques [9,23]. Furthermore, we can benefit from prior knowledge of the problem, either by initializing a quantum state close to the solution [24–26], finding a better initial ansatz by an efficient classical optimization over Clifford gates [27,28], or bypassing the optimization by transferring pretrained optimal parameters [29–38]. The optimization can also be simplified by employing fixed schemes such as Fourier-series extrapolation [9], discrete adiabatic paths [9,23], or linear ramp schedules [39]. While the main purpose of QAOA is finding approximate solutions to COPs, the sampling from the output quantum state is related to pseudo-Boltzmann distributions that are classically intractable [40,41]. In these cases, the quality of the algorithm parameters affects the effective temperature of the probability distribution, with the optima leading to the lowest temperature.

In some cases, the optimal parameters concentrate on typical values in parameter space. This behavior was initially

---

*Contact author: isak.lyngfelt@chalmers.se

observed for MAXCUT on 3-regular graphs [42], i.e., graphs where every vertex is connected to 3 vertices, and later for MAXCUT on other kinds of graphs [9,33,40,43]. This phenomenon has been linked with the locality of QAOA at shallow depths, where the cost Hamiltonian of a graph can be reformulated as a sum of local Hamiltonians of subgraphs [1,9]. The typical optimal parameters are those optimal to the local Hamiltonian that dominate the expectation value of the total cost Hamiltonian [29–31,44]. At algorithm depth $p = 11$, numerical studies show that MAXCUT on 3-regular graphs with typical parameters has performance guarantees superior to the best classical algorithms [45], and a recent work by Ref. [46] suggests that the QAOA with typical parameters already has an advantage at $p = 8$. Furthermore, it is possible to achieve near-optimal QAOA performances by transferring the typical values from a donor problem instance to a receiver problem instance drawn from a different ensemble. These successful transfers have been numerically confirmed between regular graphs with either even or odd degrees [30,47], between the Sherrington-Kirkpatrick model and MAXCUT on high-girth graphs [42,43], between unweighted and weighted graphs [32,33], and between the MAXCUT problem and other COPs [34]. Additionally, the transferred parameters can be used as initial guesses for a warm-start local optimization of QAOA parameters, significantly accelerating the classical loop [9,19,35,48]. Warm-start optimization is a promising heuristic method to mitigate the challenge of barren plateaus. Unlike other methods to avoid barren plateaus which often inadvertently imply that the problem at hand is efficiently solvable by classical computers, warm-start techniques are not believed to have this limitation [14].

The classical optimization can also be simplified by identifying symmetries in the parameter space and limiting the search space accordingly [9,37,49]. In principle, any optimal parameters related by symmetry operations output the same quantum state and, thus, perform equally when solving the classical problem. However, when transferring these parameters to other instances, not all sets of optimal parameters are suitable [30,47]. In particular, one should consider domains in the search space containing optimal parameters for both the donor and the receiver instances. To this end, it is essential to study the symmetries in both the donor and receiver problem instances.

In this article, we identify symmetries in the optimization landscape of QAOA for Ising Hamiltonian with certain integer coefficients previously not addressed in the literature. In addition, we study the performance of QAOA for MAXCUT using transferred parameters between problem instances for circuit depths $p = 1$ and 2. We define a function characterizing the transferability of parameters and provide an analytical closed-form expression for triangle-free regular graphs. Such graphs are said to have a girth higher than 3, i.e., the shortest cycle length is longer than 3. We observe that, in general, these transferred parameters reach a high performance if they belong to certain domains in the search space. Our analysis extends to d-regular graphs with integer weights and unweighted random graphs from different models: graphs with a fixed number of vertices, and edges generated by different probability distributions. We show that when specific parameters are transferable, all parameters can be effectively transferred

by applying translations based on the symmetries of the variational state's expectation value. This result appears to extend to higher algorithm depths, suggesting a scalable approach to parameter optimization. Furthermore, while our numerical study focuses on particular problem instances, our work provides tools to analyze the symmetries and transferability of parameters, which can help to identify successful initialization or optimization schemes in general VQAs. Building on this idea, we extend our symmetry results to the Hamiltonian variational *Ansatz* (HVA), a relevant quantum heuristic for physical models in the context of variational quantum eigensolvers (VQEs), as well as the RQAOA and the ma-QAOA. For the latter two, we also discuss the potential of parameter transferability.

Different initialization or optimization schemes have unique strengths and weaknesses that make them more suitable for specific VQAs. For hardware-efficient *Ansätze*, a classically efficient search over Clifford-gate constructed *Ansätze* has proven effective for finding good initializations [28]. However, this approach is less suitable for structured *Ansätze* that utilize problem-specific features, such as QAOA or HVA circuits. Indeed, in QAOA circuits, Clifford gates constrain the parameters to integer multiples of $\pi/4$, which are typically far from the optimal values [27], unless one considers hardware-efficient QAOA variants, like ADAPT-QAOA [50]. Conversely, our results are not directly applicable to hardware-efficient *Ansatz* circuits, as we rely on integer-valued coupling constants in the spin model to derive parameter space symmetries. These assumptions suggest that our findings will improve parameter training in structured quantum circuits for specific problems with integer-valued coupling constants in the Hamiltonians.

The article is structured as follows. In Sec. II, we review the formulation of QAOA and the COP MAXCUT, as well as the known symmetries of the variational parameters. We prove a generalization of these symmetries to include a wider family of graphs. Then, in Sec. III, we analyze the transferability of optimal variational parameters and identify which optimal sets are suitable for reuse. In Sec. IV, we generalize our symmetries to encompass more VQAs and discuss the parameters' transferability in those cases. Lastly, we summarize our results in Sec. V.

II. SYMMETRIES OF QAOA FOR MAXCUT

We begin by reviewing the formulation of QAOA, an algorithm that can be easily adapted to tackle the optimization of classical functions, and introducing the notation we use throughout the paper. In this work, we consider a fundamental NP-complete problem in graph theory, namely MAXCUT, whose solution is also hard to approximate. This problem seeks to divide the vertices of an undirected graph into two sets to maximize the total weight of edges, or the number of edges in the case of unweighted graphs, crossing between these sets. That is, given an undirected graph $G = (V, E)$ with weights w_{ij} on the edges $(i, j) \in E$ joining the vertices V, we seek to maximize

$$C(x) = \frac{1}{2} \sum_{(i,j) \in E} w_{ij}(1 - x_i x_j), \qquad (1)$$

with $x \in \{-1, 1\}^n$ the length-n string to be optimized, and n the order of the graph $|V|$. For a given problem instance, the quality of a feasible solution x can be quantified by the approximation ratio

$$r(x) = \frac{C(x) - C_{\min}}{C_{\max} - C_{\min}}, \qquad (2)$$

which is equal to one for optimal solutions and approaches zero as the quality decreases. For MAXCUT, the best classical algorithms guarantee an approximation ratio of at least 0.8785 [51].

The MAXCUT problem can be encoded on a quantum computer by rewriting the cost function of Eq. (1) in terms of quantum operators. We replace each variable x_i by the Pauli operator $\hat{\sigma}_z^i$ to obtain the cost Hamiltonian

$$\hat{H}_C = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} \big(\mathbb{1} - \hat{\sigma}_z^i \hat{\sigma}_z^j\big), \qquad (3)$$

where the problem's solutions are related to the measurement outcomes ± 1 of each qubit in the computational basis. The QAOA leverages the problem structure by encoding it in the cost unitary $\hat{U}_C(\gamma) = e^{-i\gamma \hat{H}_C}$. The algorithm iteratively applies this cost unitary and a mixer unitary $\hat{U}_M(\beta) = e^{-i\beta \hat{H}_M}$, with the mixer Hamiltonian $\hat{H}_M = \sum_i \hat{\sigma}_x^i$, to the symmetric superposition of computational basis states $|+\rangle^{\otimes n}$ to prepare the n-qubit variational state $|\gamma, \beta\rangle$. The state depends on the real-valued parameters $\gamma = (\gamma_1, \ldots, \gamma_p)$ and $\beta = (\beta_1, \ldots, \beta_p)$ assigned to the cost and mixing operations for the total p iterations, depth p, and aims to optimize the expectation value of the cost Hamiltonian $\hat{H}_C$,

$$C(\gamma, \beta) = \langle \gamma, \beta | \hat{H}_C | \gamma, \beta \rangle. \qquad (4)$$

We can thus consider the approximation ratio over the distribution of QAOA solution outputs as a quality measure

$$r(\gamma, \beta) = \frac{C(\gamma, \beta) - C_{\min}}{C_{\max} - C_{\min}}. \qquad (5)$$

The cost landscape represents the relationship between the objective function of Eq. (4) and the parameters γ and β. The classical optimization of γ and β is generally challenging due to highly nonconvex landscapes containing several local minima and maxima. The landscape symmetries help simplify the optimization process by restricting the search space. In other words, these symmetries help divide the search space into domains containing optimal parameters that perform equally when finding good approximate solutions to a COP or creating final states from which one can perform "pseudothermal" sampling [40,41].

First, we notice that the cost and mixer Hamiltonians are real symmetric matrices in the computational basis and, therefore, QAOA exhibits a "time-reversal" symmetry $(\gamma, \beta) \rightarrow (-\gamma, -\beta)$ [9,37]. This symmetry inverts the sign of all elements of the vectors simultaneously and is valid for any graph with real weights. Second, we consider the standard mixer Hamiltonian $\hat{H}_M = \sum_i \hat{\sigma}_x^i$ and MAXCUT problem instances with integer weights so that the cost Hamiltonian has integer eigenvalues and only two-qubit interactions. Therefore, we begin by restricting the parameter optimization to the domain $\mathcal{A}^p$, such that $\gamma \in [-\pi, \pi)^p$ and $\beta \in [-\pi/4, \pi/4)^p$ (as shown
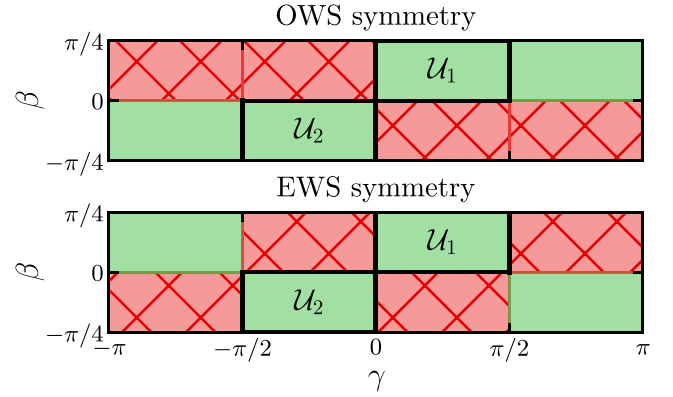

FIG. 1. Symmetries for optimal QAOA parameters (γ, β) at depth $p = 1$ for unweighted regular graphs with odd (top) and even (bottom) degrees, which follow OWS and EWS symmetries, respectively. Each solid green area contains one set of optimal parameters, while they do not appear in red-crossed areas. The domains $\mathcal{U}_1$ and $\mathcal{U}_2$ contain optimal parameters in both cases.

in Fig. 1 for $p = 1$) since the QAOA cost function for MAX-CUT remains unchanged when any γ_i and β_i parameters shift by integer multiples of 2π and $\pi/2$, respectively.

In the following, we review additional symmetries of the optimal parameters within $\mathcal{A} = [-\pi, \pi) \times [-\pi/4, \pi/4)$, summarized in Table I, and extend the known results for MAXCUT unweighted graph instances [9,37] to specific families of weighted graphs.

Graphs where the sum of weights connected to each vertex is even. The cost in Eq. (4) is invariant under parameter changes $\gamma_i \rightarrow \gamma_i \pm \pi$ for any layer i in QAOA, for graphs where the sum of the weights connected to each vertex is even, i.e., $\Sigma_j w_{jk} \in 2\mathbb{Z}, \ \forall \ k \in V$ (see proof in Appendix B). This even weight sum (EWS) symmetry captures the previously known result for unweighted graphs where all vertices have even degree [37]. More examples that exhibit this parameter pattern include graphs with only even weights and graphs where all vertices have even degrees with only odd weights, for instance, $w_{ij} \in \{-1, 1\}$.

Graphs where the sum of weights connected to each vertex is odd. The cost in QAOA is invariant under the simultaneous parameter change $\gamma_i \rightarrow \gamma_i \pm \pi$ and $\beta_j \rightarrow -\beta_j$ in all layers

TABLE I. Transformations of $(\gamma, \beta) \rightarrow (\gamma', \beta')$ such that $C(\gamma, \beta) = C(\gamma', \beta')$. The EWS symmetry refers to integer-weighted graphs where the sum of weights connected to each vertex is even, i.e., $\Sigma_j w_{jk} \in 2\mathbb{Z}, \ \forall \ k \in V$. Analogously, OWS symmetry refers to integer-weighted graphs where the sum of weights connected to each vertex is odd, i.e., $\Sigma_j w_{jk} \in \mathbb{Z} \setminus 2\mathbb{Z}, \ \forall \ k \in V$.

Graph	Symmetry
EWS	$(\gamma_1, ..., \gamma_i, ..., \gamma_p) \rightarrow (\gamma_1, ..., \gamma_i \pm \pi, ..., \gamma_p)$
OWS	$\begin{cases} (\gamma_1, ..., \gamma_i, ..., \gamma_p) \rightarrow (\gamma_1, ..., \gamma_i \pm \pi, ..., \gamma_p) \\ (\beta_1, ..., \beta_i, ..., \beta_p) \rightarrow (\beta_1, ..., -\beta_i, ..., -\beta_p) \end{cases}$
Integer weights	$(\gamma_1, ..., \gamma_i, ..., \gamma_p) \rightarrow (\gamma_1, ..., \gamma_i \pm 2\pi, ..., \gamma_p)$
	$(\beta_1, ..., \beta_i, ..., \beta_p) \rightarrow (\beta_1, ..., \beta_i \pm \pi/2, ..., \beta_p)$
Real weights	$\begin{cases} (\gamma_1, ..., \gamma_i, ..., \gamma_p) \rightarrow -(\gamma_1, ..., \gamma_i, ..., \gamma_p) \\ (\beta_1, ..., \beta_i, ..., \beta_p) \rightarrow -(\beta_1, ..., \beta_i, ..., \beta_p) \end{cases}$

$j \geqslant i$, for graphs where the sum of the weights connected to each vertex is odd, i.e., $\Sigma_j w_{jk} \in \mathbb{Z} \setminus 2\mathbb{Z}, \forall k \in V$ (see proof in Appendix B). This odd weight sum (OWS) symmetry contains the case of unweighted graphs when all vertices have odd degree [37]. Moreover, we also include graphs where each vertex has an odd degree and all weights are odd, for example, $w_{ij} \in \{-1, 1\}$.

As previously mentioned, these symmetries relate to patterns in the optimization landscape and help identify sets of optimal parameters that lead to the same QAOA performance for a given MAXCUT problem instance. The optimal parameters are known analytically for unweighted d-regular triangle-free graphs at depth $p = 1$ [52], and empirically for the unweighted 3-regular tree subgraph up to $p = 11$ [45]. Similar empirical results are known for unweighted regular graphs of degree $d \leqslant 11$ at depth $p = 2$ [45], and exhaustive sets of small graphs with $n \leqslant 9$ vertices at depths $p \leqslant 3$ [37]. In Fig. 1, we show the parameter space domains that contain the optimal sets for unweighted regular graphs with even and odd degrees. We observe that, in both cases, the domains $\mathcal{U}_1 = [0, \pi/2) \times [0, \pi/4)$ and $\mathcal{U}_2 = [-\pi/2, 0) \times [-\pi/4, 0)$ contain a set of optimal parameters.

Symmetry patterns at higher depths. At higher algorithm depths p, graphs that obey either the even weight sum (EWS) symmetry or odd weight sum (OWS) symmetry will have 2^{p+1} areas containing an optimal set of parameters within $\mathcal{A}^p$. As shown in Table I, the cost function will be invariant under the mth algorithm layer parameter changes $\gamma_m \rightarrow \gamma_m \pm \pi$ for instances with EWS symmetry, and $\gamma_m \rightarrow \gamma_m \pm \pi$ and $\beta_i \rightarrow -\beta_i$ for $i \geqslant m$ for instances with OWS symmetry. These symmetries provide two optimal parameter choices at each layer, either the initial set or the symmetric one, which results in 2^p possible combinations of optimal parameters. Lastly, the cost also remains invariant under the "time-reversal symmetry," which changes the sign of all parameters in the algorithm, thus doubling the number of optimal choices. In total, there are 2^{p+1} sets of optimal parameters within $\mathcal{A}^p$ providing the same QAOA performance. We could visualize them as different domains in the energy landscape, as shown in Fig. 1 for $p = 1$.

In summary, we have identified parameter symmetries in QAOA for weighted MAXCUT instances that explain optimal parameter patterns from numerical simulations. Given the proven symmetries, for a given instance, all sets of optimal parameters can be retrieved knowing a single set, provided there is a single optimum in the domain $[0, \pi/2) \times [-\pi/4, \pi/4)$. The choice of parameters in the algorithm determines the quantum circuit. Therefore, a higher number of optimal parameter sets results in more choices of quantum circuits and gates, potentially improving the hardware implementation of the algorithm.

III. TRANSFERABILITY OF OPTIMAL PARAMETERS

In this section, we analyze the performance of QAOA when transferring the different sets of optimal parameters identified in Sec. II from one problem instance to another. To that end, we define a figure of merit to quantify how transferable the parameters are. We study this transferability property both analytically and numerically for several families of graphs, with different degree and weight distributions.

Previous research showed the viability of transferring optimal parameters of QAOA for MAXCUT between unweighted 3-regular graphs of different sizes [29,45], from unweighted to weighted graphs [32,33], between graphs with different degrees [30,47], and between different COPs [34,39]. The transferability of optimal parameters can be explained by the decomposition of a given graph instance into local subgraphs at shallow QAOA depths. Following a light cone argument, one can show that the algorithm sees a sum of local subgraphs instead of the whole graph. Thus, the QAOA cost function only depends on these subgraphs, and the algorithm will perform similarly for two different problem instances with similar local subgraphs. For example, the cost function of Eq. (4) for all unweighted 3-regular graphs for QAOA with depth $p = 1$ is a sum of the cost of three subgraphs, times their respective multiplicity [1]. Such locality arguments for shallow algorithm depths also explain concentration properties of QAOA, and have been linked to performance limitations for certain problems [2,5,7,14].

The number of possible subgraphs for a d-regular graph grows exponentially with the algorithm depth p, if QAOA does not see the whole graph, i.e., if $p \sim O(\log n)$ with n the number of vertices. However, despite this exponential growth, in a sufficiently large random d-regular graph, almost all subgraphs will be tree subgraphs [2,25,44], i.e., connected subgraphs graphs with no cycles. Consequently, the optimal QAOA parameters for a d-regular graph will be close to the optimal parameters of the d-regular tree subgraph, which explains the successful parameter transferability between them at shallow depths. However, this argument fails to explain whether we can reuse optimal parameters between graphs with different degrees (or average degrees). Particularly, prior research suggests that, on average, we can successfully transfer any optimal parameters between regular graphs with the same parity (both have either odd or even degrees). However, for graphs with opposite parity, we can only reuse a subset of the optimal parameters [30,47].

Transferability error. To measure the success of reusing parameters of QAOA, we define the transferability error $\eta_{R,D}$ as

$$\eta_{R,D} = r_R(\boldsymbol{\gamma}_R^*, \boldsymbol{\beta}_R^*) - r_R(\boldsymbol{\gamma}_D^*, \boldsymbol{\beta}_D^*). \tag{6}$$

The error $\eta_{R,D}$ is the difference between the approximation ratios r_R for a receiver problem instance R, defined in Eq. (5), when using its optimal parameters $(\boldsymbol{\gamma}_R^*, \boldsymbol{\beta}_R^*)$, and when using parameters that are optimal to a donor instance D, $(\boldsymbol{\gamma}_D^*, \boldsymbol{\beta}_D^*)$. A low error indicates a successful transfer.

Analytic expression of the transferability error for simple cases at $p = 1$. We begin our study of the parameters' transferability with the simplest example of unweighted d-regular graphs, for which we can rely on previously known analytical results. Given the locality of QAOA at shallow depths, we choose the d'-regular tree subgraph, the most common subgraph, as the donor instance. Indeed, using the optimal parameters for the 3-regular tree subgraph is known to give performance guarantees for arbitrary 3-regular graphs at $p = 1$ [1] and $p = 2$ [45]. Furthermore, the cost value for MAXCUT with triangle-free graphs at depth $p = 1$ has a known closed form [52]. Using the closed form, we can derive an analytical expression for the transferability error
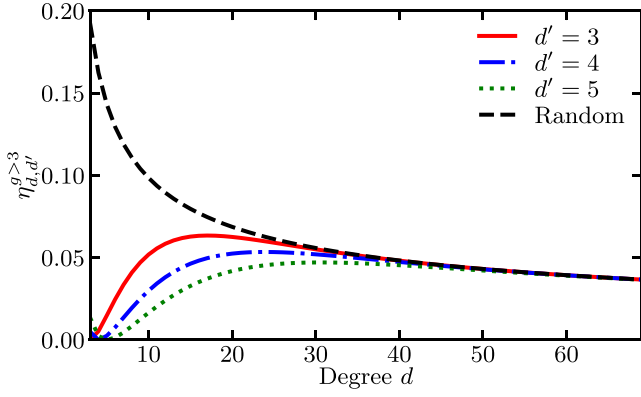
FIG. 2. Transferability error $\eta_{d,d'}^{g>3}$ from Eq. (7) for d-regular bipartite graphs ($k = 1$) with a donor tree subgraph with degrees $d' = 3$ (red solid line), $d' = 4$ (blue dashed-dotted line), and $d' = 5$ (green dotted line). The black dashed line represents the transferability error when QAOA with donated parameters is equivalent to random guessing.

FIG. 3. Average transferability error $\eta_{d,d'}$ of Eq. (6) at depth $p = 1$ for d-regular receiver graphs with $n = 120$ vertices and donor tree subgraphs with degrees $d' = 3$ (red circle) and $d' = 4$ (blue cross). The error bars overlap with the symbols and denote the 25th and 75th percentiles. Comparison with the analytic results $\eta_{d,3}^{g>3}$ (red solid line) and $\eta_{d,4}^{g>3}$ (blue dashed-dotted line) of Eq. (7). The black dashed line indicates the average transferability error when QAOA with donated parameters performs as random guessing.

between a d'-regular tree subgraph and a d-regular graph with girth $g > 3$, i.e., triangle-free regular graphs. Thus, following our definition of Eq. (6), and identifying the receiver (R) and donor (D) graphs by their degree d and d', we have for depth $p = 1$

$$\eta_{d,d'}^{g>3} = \frac{1}{2k\sqrt{d}}\left[\left(\frac{d-1}{d}\right)^{\frac{d-1}{2}} - \sqrt{\frac{d}{d'}}\left(\frac{d'-1}{d'}\right)^{\frac{d-1}{2}}\right], \quad (7)$$

with $k = C_{\max}/|E|$, the fraction of edges cut in the maximum cut. In general, $k \in (\frac{1}{2}, 1]$ and $k = 1$ for bipartite graphs, and Eq. (7) is valid for sufficiently large graphs [1], that is, graphs of order larger than $n \sim (d-1)^{2p}$ (see Appendix E for a detailed derivation).

The function in Eq. (7) allows us to calculate the transferability error for regular graphs up to an arbitrary degree. We visualize this expression in Fig. 2 for bipartite receiver graphs with degrees up to $d = 70$ and tree donor subgraphs of degree $d' \in \{3, 4, 5\}$. We also include a reference line where QAOA with donated parameters is equivalent to random guessing, that is, when QAOA gives an approximation ratio of $\frac{1}{2}$, or equivalently $\eta_{R,D} = r_R(\gamma_R^*, \beta_R^*) - \frac{1}{2}$. Note that if we shift this line by $\frac{1}{2}$, it corresponds to the approximation ratio of QAOA with optimal parameters, $r_R(\gamma_R^*, \beta_R^*)$, for increasing graph degrees. The transferability error $\eta_{d,d'}^{g>3}$ is minimized for donor and receiver graphs with the same degree $d = d'$, in agreement with previous results [29,30]. We observe that as the difference between the receiver and donor degrees d and d' grows, QAOA with the parameters transferred from the tree subgraphs performs as random guessing with an approximation ratio of $r = \frac{1}{2}$ [53]. Therefore, while the error tends to 0 as $d \to \infty$, transferring parameters becomes pointless in the large-d limit. Moreover, when the error eventually reaches 0 for large d, it also indicates that the approximation ratio of QAOA with depth $p = 1$ for any optimal parameters is $r \approx \frac{1}{2}$ [2].
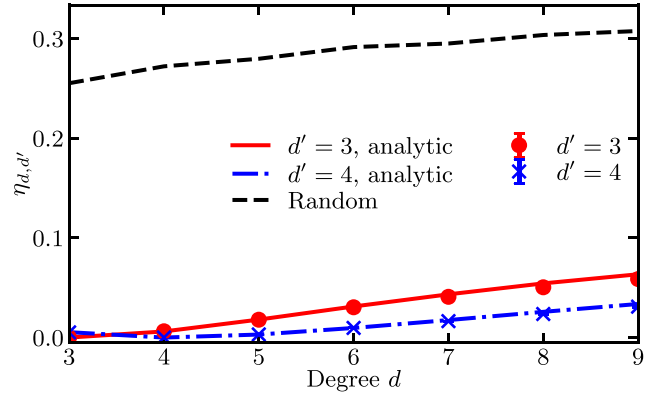
A. Numerical analysis

Here, we simulate and study the transferability of optimal QAOA parameters, characterized by Eq. (6), for different families of graphs.

Transferability from the d'-regular tree subgraph to d-regular unweighted graphs at $p = 1$. This case is approximated by the analytic expression in Eq. (7), visualized in Fig. 2. Here, we consider 10 instances of order $n = 120$ of receiver regular graphs for each degree ranging from $d = 3$ to 9. We use the optimal parameters for the triangle-free graphs, equivalent to the optimal parameters of the tree subgraphs, $\gamma_{d'}^* = \arctan(1/\sqrt{d'-1})$ and $\beta_{d'}^* = \pi/8$ [52]. In Fig. 3, we observe that Eq. (7) reproduces the numerical simulations of Eq. (6), and that the transferability error increases when the difference between the degrees of the donor and receiver graphs grows.

We note that, in this case, the success of the transfer does not depend on the relative parity between the degrees d and d'. Indeed, the set of optimal parameters from the donor graph lies in the domain $\mathcal{U}_1$ (see Fig. 1) and, as discussed in Sec. II, $\mathcal{U}_1$ contains optimal parameters for regular graphs with both odd and even degree. Previous works also suggest that only parameters from such shared domains, $\mathcal{U}_1$ and $\mathcal{U}_2$, are the ones transferable between unweighted regular graphs with degrees of different parity [30]. However, considering the problem symmetries, if we instead have access to an optimal set of parameters from a different domain, it can be easily translated into either $\mathcal{U}_1$ or $\mathcal{U}_2$ following the rules of Table I. In view of this, knowing the problem symmetries and the shared domains between instances, we can transfer any optimal parameter between any d-regular graph independently of the graphs' degrees.

Transferability from the d'-regular tree subgraph to random graphs at $p = 1$. In this case, we use again the unweighted d'-regular tree subgraph as a donor graph, for general random receiver graphs. Unlike regular graphs, random graphs

do not have a fixed degree and, at low depths, QAOA sees a local subgraph with a varying degree around each vertex. The random graph models will have different distributions of degrees. Here, we study three models of random graphs: the Erdős-Rőnyi (ER) model [54], the Barabási-Albert (BA) model [55], and the Watts-Strogatz (WS) model [56]. While these models can be tuned to have the same average degree, they exhibit distinct properties that make them suitable for describing different systems. The ER model generates random graphs where edges appear independently with equal probability, resulting in a binomial degree distribution. In contrast, the BA model generates scale-free networks characterized by large hubs, where a few vertices have many connections and most have few, following a power-law degree distribution. The WS model produces graphs with local clusters of connected vertices by rewiring edges in a ring lattice with fixed probability. Both BA and WS graphs are representative of real-world network structures.

We use ten instances of each type of random receiver graph with order $n = 200$, and the same average degree $\bar{d} \approx 6$. The degree of the donor tree subgraph takes the values $d' = \{3, ..., 9\}$. In contrast to the previous case, these random unweighted graphs have neither the OWS nor the EWS symmetry described in Sec. II, and we find that a transfer is only successful if the donor parameters are from the domains $\mathcal{U}_1$ or $\mathcal{U}_2$. As we observe in Fig. 4, the transferability error is low for parameters from $\mathcal{U}_1$ and increases for parameters from neither $\mathcal{U}_1$ nor $\mathcal{U}_2$. Given the degree distribution of the models, the random graphs have both odd and even degree subgraphs. As a result, even the parameters optimal to the 6-regular tree subgraph perform poorly if they are not included in $\mathcal{U}_1$ or $\mathcal{U}_2$. We can compare this result to Ref. [30], where they fix the percentage of even-regular subgraphs in 20-node graphs and find different QAOA performances for parameters transferred from each domain. As in the previous case, the transferability error is minimized when the donor and receiver graphs have similar degrees $d' = \bar{d}$, that is, $\eta_{\bar{d},6} \approx 0$ when the donor is a 6-regular tree in our analysis. Furthermore, for the random graph models we studied, the impact of the variance of the degree distribution compared to the average degree is negligible. Additionally, the relative parity between the degree of the donor subgraph and the average degree only affects the transferability for the BA graphs, when using optimal parameters outside $\mathcal{U}_1$ or $\mathcal{U}_2$.

Transferability from the d'-regular tree subgraph to d-regular graphs with integer weights at $p=1$. We study the transferability of parameters from unweighted tree donor subgraphs with degrees $d' = 3$ and 4 to d-regular graphs with uniformly distributed integer weights $w_{ij} \in \{-1, 1\}$, and degrees d varying from 3 to 9. To this end, we use 10 instances of 20-node receiver graphs for each degree d. We choose the optimal parameter set of the donor graphs from the domain $\mathcal{U}_1$. Following the discussion in Sec. II, we know that the receiver regular graphs with this weight distribution follow the EWS or the OWS symmetry for even or odd degrees, respectively.

In Fig. 5, we observe a low transferability error, especially when $d \approx d'$, confirming that the donor parameters are transferable to the integer weighted graphs. The small size of the simulated graphs ($n = 20$) increases the numerical results' variability and moves us out of the large graph regime with
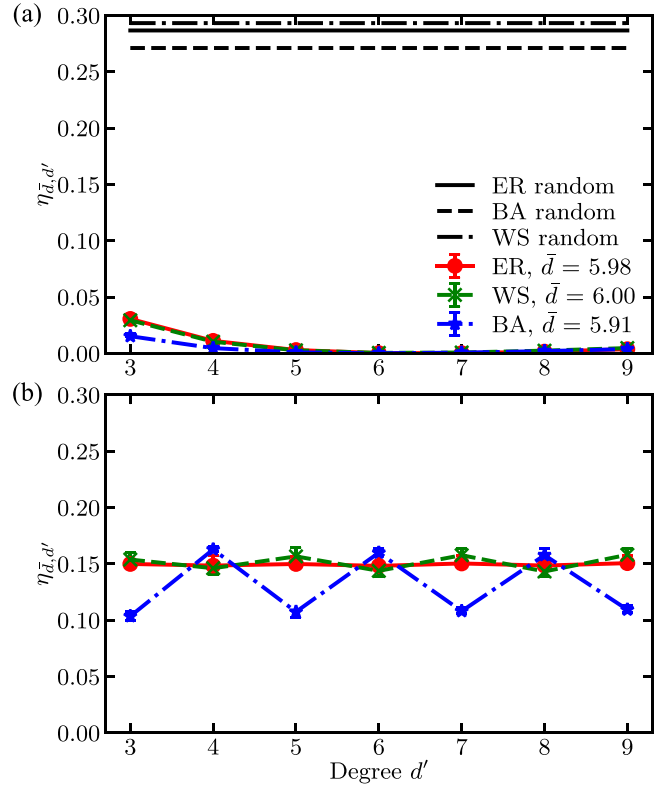

FIG. 4. Average transferability error $\eta_{\bar{d},d'}$ of Eq. (6) at depth $p = 1$ for random receiver graphs with $n = 200$ vertices and average degree $\bar{d} \approx 6$, sampled from the ER model (red solid line), the WS model (green dashed line), and the BA model (blue dashed-dotted line). We consider donor regular tree subgraphs with degrees $3 \leqslant d' \leqslant 9$ and their optimal parameters from (a) $\mathcal{U}_1$, and (b) neither $\mathcal{U}_1$ nor $\mathcal{U}_2$. The error bars overlap with the symbols and denote the 25th and 75th percentiles. The black lines in (a) represent the average transferability error when QAOA with donated parameters is equivalent to random guessing for ER (solid line), WS (dashed line), and BA (dashed-dotted line).
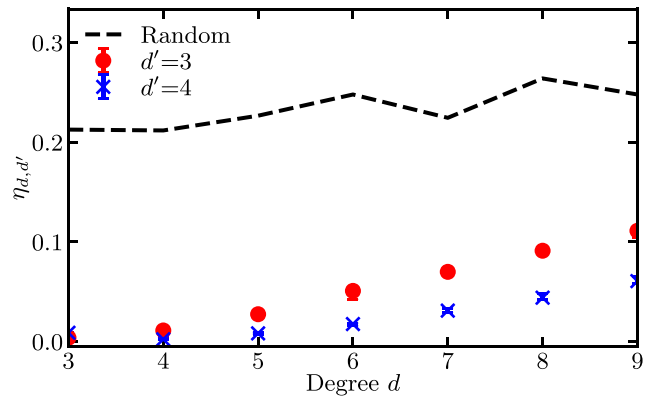

FIG. 5. Average transferability error $\eta_{d,d'}$ of Eq. (6) at $p = 1$ for d-regular receiver graphs with $n = 20$ vertices, and uniformly distributed weights $w_{ij} \in \{-1, 1\}$. The donor graphs are unweighted tree subgraphs with degrees $d' = 3$ (red circle) and $d' = 4$ (blue cross). The error bars overlap with the symbols and denote the 25th and 75th percentiles. The black dashed line shows the average transferability error when QAOA with transferred parameters behaves as random guessing.
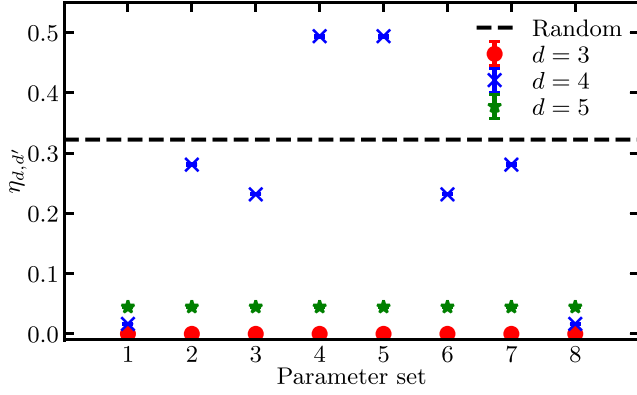
FIG. 6. Average transferability error $\eta_{d,d'}$ of Eq. (6) at depth $p = 2$ for unweighted regular receiver graphs of degrees $d = 3$ (red circles), $d = 4$ (blue crosses), and $d = 5$ (green stars), with an unweighted tree donor subgraph of degree $d' = 3$. The error bars overlap with the symbols and denote the 25th and 75th percentiles. We use the eight optimal sets of parameters of the donor graph (see Table II), numbered from one to eight, as shown on the x axis. The sets 1 and 8 lie in $\mathcal{U}_1$ and $\mathcal{U}_2$, respectively. The black dashed line shows the average transferability error when QAOA with transferred parameters performs as random guessing.

triangle-free graphs [2], especially for the higher degrees. Without this large graph approximation, the tree subgraph is a less suitable donor. As in all previous examples, the error increases as the difference between d and d' increases and, in this case, it grows faster than for the unweighted d-regular graphs. The smaller sizes of the receiver graphs can contribute to this faster deterioration of the transferability. Additionally, we confirm with an initial numerical exploration that the transferability error increases when using parameters outside $\mathcal{U}_1$ and $\mathcal{U}_2$. Since the symmetries that apply are the same as for the unweighted d-regular graphs, we can follow a similar argumentation as in the unweighted graphs case. Namely, the universally transferable domains for the weighted graphs are those universal to the unweighted d-regular graphs, i.e., $\mathcal{U}_1$ and $\mathcal{U}_2$. However, we know d-regular graphs with these integer weights satisfy either the OWS or the EWS symmetry, and using the rules in Table I we can shift a given set of parameters so that the transferability becomes independent of the relative parity between d and d'.

Transferability at higher algorithm depths. As we show in Sec. II, at depth $p = 2$, QAOA for MAXCUT on d-regular graphs has eight sets of optimal parameters in $\mathcal{A}^2$, instead of the four sets at $p = 1$. We present these parameters in Appendix F (see Table II) for the 3-regular tree subgraph. Here, we study the performance of QAOA for unweighted regular graphs of order $n = 120$ and degrees $d \in \{3, 4, 5\}$, considering 10 graphs per degree, and using, without shifting to a common domain, the eight sets of optimal parameters of the 3-regular tree donor subgraph.

In Fig. 6, we observe an analogous behavior to the case of $p = 1$. That is, only parameters from the higher-dimensional domains $\mathcal{U}_1^2 = [0, \pi/2)^2 \times [0, \pi/4)^2$ and $\mathcal{U}_2^2 = [-\pi/2, 0)^2 \times [-\pi/4, 0)^2$ can be transferred (without using symmetry rules) independently of the relative parity between the donor and receiver. When transferring these parameters

to a receiver graph of degree $d = 3$ or 5, i.e., with the same parity, the error remains low for all eight sets of parameters. Conversely, when transferring to 4-regular graphs, the error is significantly higher for parameters from neither $\mathcal{U}_1^2$ nor $\mathcal{U}_2^2$. We notice that these six sets of parameters could be translated into $\mathcal{U}_1^2$ or $\mathcal{U}_2^2$ using the symmetry relations of the cost function. Unlike the case of tree subgraphs at $p = 1$, there is no analytically known set of optimal parameters in $\mathcal{U}_1$ for a general d-regular tree at $p = 2$. However, as mentioned in Sec. II, previous studies show numerical evidence for optimal parameters in $\mathcal{U}_1^2$ for $d \leqslant 11$ and $p = 2$ [45]. When restricting to the 3-regular tree subgraph, the same authors provide numerically optimized values in $\mathcal{U}_1^p$ at even higher depths, $p \leqslant 11$ [45]. Moreover, other works that consider interpolation and pseudoadiabatic paths for finding parameters at higher depths use the initial optimal parameters in $\mathcal{U}_1$ and find that $(\boldsymbol{\gamma}, \boldsymbol{\beta}) \in \mathcal{U}_1^p$ [9,23]. These adiabatic interpolation techniques fail when considering initial optimal parameters to 3-regular graphs outside $\mathcal{U}_1$ and $\mathcal{U}_2$, $(\gamma^*, \beta^*) \in [\pi/2, \pi) \times [0, \pi/4)$ [23].

Summarizing, prior work confirms the existence of optimal parameters in $\mathcal{U}_1^p$ for the tree subgraph up to $p = 11$. However, more work is necessary to conclude whether these parameters are universally transferable at depths $p > 2$ for regular graphs of any degree, either odd or even. For depths $p > 11$, it remains an open question whether optimal parameters (transferrable or not) exist in $\mathcal{U}_1^p$. For instance, if we consider a linear ramp interpolating the optimal values of the parameters, we can expect that their values will eventually lie outside $\mathcal{U}_1^p$. Nevertheless, any optimal parameter at those depths will follow the symmetry rules discussed in Sec. II.

IV. PARAMETER SYMMETRIES IN OTHER VARIATIONAL QUANTUM ALGORITHMS

The quantum circuit of many hybrid classical-quantum algorithms can be built through several iterations of similar quantum operations, e.g., any layer of QAOA contains a parametrized mixing unitary operation and a parametrized cost unitary operation. Such structured quantum circuit *Ansätze* also appear in realizations of VQEs for finding ground states of quantum many-body problems, as the HVA. Compared to the hardware-efficient *Ansatz*, the problem-specific HVA circuit provides an easier optimization, reducing the challenges of barren plateaus while maintaining a high expressivity [57]. While in this paper our focus mainly lies on the QAOA for different MAXCUT instances and the transferability of its optimal parameters between various ensembles of graphs, we discuss here how these results can be extended to other COPs and VQAs. The EWS and OWS symmetries shown in Sec. II for the MAXCUT problem also apply to all Ising Hamiltonians of the form $\hat{H}_C = \sum_{i<j} w_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z / 2 + \sum_i h_i \hat{\sigma}_z^i / 2$ (see detailed derivation in Appendix C). The condition on the coefficients w_{ij} and h_i change to $\sum_j w_{jk} + h_k \in 2\mathbb{Z}$, $\forall\, k \in V$ and $\sum_j w_{jk} + h_k \in \mathbb{Z} \setminus 2\mathbb{Z}$, $\forall\, k \in V$ for EWS and OWS, respectively. This extension covers a wide range of problems, including Karp's 21 NP-complete problems [58] and finding ground states of spin-glass models.

The quantum unitary operations of the HVA use terms of many-body Hamiltonians, analogously to the cost function in QAOA and adiabatic quantum computation. Therefore, we can identify similar symmetries in these VQAs. A notable example is the HVA for the transverse-field Ising model (TFIM), which coincides with the QAOA for solving MAXCUT in the "ring of disagrees," a 2-regular unweighted ring graph [1]. Therefore, the HVA for the TFIM exhibits EWS, and the parameters can be optimized over a smaller domain. In Appendix D, we show that EWS is valid for *Ansätze* constructed with operators of the form

$$\hat{U}_{\alpha\alpha}(\gamma) = \prod_{(i,j)\in E} \exp\left(-i\frac{w_{ij}}{2}\gamma\hat{\sigma}_\alpha^i\hat{\sigma}_\alpha^j\right), \qquad (8)$$

where $\alpha \in \{x, y, z\}$, given that the weights w_{ij} fulfill the criteria of EWS described in Sec. II. In other words, any VQE *Ansatz* with these terms and conditions follows the EWS.

We extend our results to other HVA based on different Hamiltonian terms. In particular, we consider the general case of a Heisenberg model on high-dimensional spaces, where the connectivity between the spins is defined by the set of edges E of a given graph. Moreover, we allow the coupling constants between the spins to vary from site to site, thus introducing disorder. The Hamiltonian reads as

$$\hat{H} = \frac{1}{2} \sum_{(i,j)\in E} \left(w_{ij}^x\hat{\sigma}_x^i\hat{\sigma}_x^j + w_{ij}^y\hat{\sigma}_y^i\hat{\sigma}_y^j + w_{ij}^z\hat{\sigma}_z^i\hat{\sigma}_z^j\right), \qquad (9)$$

with w_{ij}^x, w_{ij}^y, and w_{ij}^z the coupling strengths associated with the x, y, and z components of the spin interaction between the sites i and j, given by the edge (i, j). This general case reduces to well-known models for studying magnetic systems. For instance, we can consider a quantum spin chain with real-valued coupling constants, such that the Hamiltonian becomes

$$\hat{H}_{XYZ} = \sum_{i=1}^n \left(J_x\hat{\sigma}_x^i\hat{\sigma}_x^{i+1} + J_y\hat{\sigma}_y^i\hat{\sigma}_y^{i+1} + J_z\hat{\sigma}_z^i\hat{\sigma}_z^{i+1}\right), \qquad (10)$$

in which we assume periodic boundary conditions ($1 \equiv n + 1$). This one-dimensional Heisenberg XYZ model corresponds to a particular case of Eq. (9) for a 2-regular ring graph with n vertices defining the connectivity of the interactions, and uniform coupling constants $w_{ij}^\alpha/2 = J_\alpha$, with $\alpha \in \{x, y, z\}$. When $J_x = J_y$, the model is called the one-dimensional Heisenberg XXZ model, studied in the context of HVA [57]. Without loss of generality, we can take $J_x = J_y = 1$ and $J_z = \Delta$, with Δ controlling the spin anisotropy. The critical point $\Delta = 1$ becomes harder to approximate classically [57]. Interestingly, optimal HVA parameters have been successfully transferred from small to large instances of the Heisenberg XXZ model for parameters in the domain $\mathcal{U}_1^p$ [38]. We derive symmetry results for all these models in the case of integer-valued coupling constants w_{ij}^α. In Appendix D, we base our derivations on the fact that the HVA for the general Hamiltonian of Eq. (9) comprises products of operators given by Eq. (8).

Finally, we also consider different versions of the QAOA, namely, the RQAOA [7] and the ma-QAOA [8]. The RQAOA reduces the original optimization problem size by imposing constraints obtained from recursive applications of QAOA.

With each algorithm iteration, the weights and regularity of the problem graph change, altering the cost Hamiltonian for the QAOA subroutine. As we show in Appendix D, RQAOA preserves the EWS of the original cost Hamiltonian throughout every iteration, each reduced cost Hamiltonian adheres to the EWS but breaks the OWS after the first iteration. This behavior implies that all optimal parameters are directly transferable if the original graph obeys EWS. Otherwise, for initial problem graphs with OWS, we can see from our discussion on transferability to random model graphs in Sec. III, that only optimal parameters from the domains $\mathcal{U}_1$ or $\mathcal{U}_2$ can be reused between iterations. Another version of QAOA is the ma-QAOA, which uses a higher number of parameters at shallow depths to create a more expressive *Ansatz*. Unlike standard QAOA, which requires EWS or OWS for the entire graph, with ma-QAOA we can focus on the weight sums of individual vertices to analyze the transferability of the gate parameters related to Hamiltonian terms of specific vertices and edges. Both RQAOA and ma-QAOA improve the performance of the standard QAOA at shallow depths, albeit with increased demands on parameter optimization [7,8,59,60]. Therefore, insights from transferability analyses to improve these optimization loops are especially valuable.

V. CONCLUSIONS AND OUTLOOK

Optimizing the variational circuits remains a big obstacle for realizable and successful VQAs. One useful simplification in the optimization task is limiting the search space using symmetries in the energy landscape. Additionally, the optimal parameters concentrate on typical values for different instances of certain problems. Thus, we can use these typical values as pretrained parameters and transfer them between instances, either to be used directly, or as initialization points for a warm-start optimization. In this article, we extend the known energy landscape symmetries of MAXCUT and show how to leverage them to transfer optimal QAOA parameters. While previous studies focus on unweighted graphs, we find that for all graphs where the sum of the weights connected to each vertex is even, the landscape follows the EWS symmetry (see Table I). Analogously, the landscape obeys the OWS symmetry for all graphs where the sum of the weights connected to each vertex is odd. Among the examples following these symmetries, we find regular graphs with weights $w_{ij} \in \{-1, 1\}$.

Regarding the optimal parameters for $p = 1$, it is known that unweighted triangle-free regular graphs have an optimal set in the domain $\mathcal{U}_1 = [0, \pi/2] \times [0, \pi/4]$ [52], and from the symmetries, other optimal sets in the green solid areas of Fig. 1, depending on the parity of the degree. Here, we analyze the transferability of these optimal parameters, i.e., those of a d'-regular tree subgraph, to triangle-free d-regular graphs, d-regular graphs, different kinds of random graphs, and d-regular graphs with integer weights. Our work shows that, in all cases, the graphs are viable recipients for the transferred parameters.

To measure the success of transferring parameters, we define a transferability error in Eq. (6), $\eta_{R,D}$, as a function of the approximation ratios of QAOA with the optimal parameters of the receiver (R) and donor (D) graphs. Using an

analytically known set of optimal parameters to the d'-regular tree subgraphs, we derive a closed-form expression for the error when transferring to triangle-free d-regular graphs at $p = 1$. The expression allows us to analyze the performance of transferability for large graphs with high d without needing to simulate large QAOA circuits. Furthermore, this approximated expression reproduces the numerical simulations of low-girth graphs. The transferability error is minimized and close to zero if $d \approx d'$ for d-regular graphs with a d'-regular donor tree subgraph. We also find that the error goes to zero in the $d \to \infty$ limit. However, we can explain the latter result by QAOA performing equally poorly for all parameters.

We observe that the success of a transfer is independent of the relative parity between the degree of the donor and receiver when the transferred set of parameters comes from $\mathcal{U}_1$ or $\mathcal{U}_2$ (see Fig. 1), as already reported in previous works [30]. However, we provide the rules to translate any optimal set of parameters into these domains using the donor graph symmetry, thus making all sets of optimal parameters transferable in practice.

Similarly, our results with random receiver graphs show that a transfer is only successful if the donor parameters come from $\mathcal{U}_1$ or $\mathcal{U}_2$. In all the cases with random graphs, generated from the ER model, the BA model, and the WS model, the success depends on the closeness between the degree of the donor and the average degree of the receiver graph, rather than on the degree distribution of the model. In this case, we find that a transfer is only successful if the donor parameters come from $\mathcal{U}_1$ or $\mathcal{U}_2$, independently of the relative parity between donor and receiver graphs. When considering weighted graphs obeying either the EWS or the OWS symmetry, the transferability error behaves similarly to the case of unweighted graphs. That is, the minimum error occurs when $d = d'$, and the domains $\mathcal{U}_1$ and $\mathcal{U}_2$ contain optimal parameters that work regardless of the relative parity between the degrees of the donor and receiver graphs.

At the higher algorithm depth $p = 2$, we also find successful transfers between any d-regular and d'-regular graph, if the parameters come from the higher-dimensional domains $\mathcal{U}_1^2 = [0, \pi/4]^2 \times [0, \pi/2]^2$ or $\mathcal{U}_2^2 = [-\pi/4, 0)^2 \times [-\pi/2, 0)^2$. Similarly, in this case, we can translate any optimal set of parameters to these domains using symmetries. We notice that the number of domains that contain an optimal set of parameters within $\mathcal{A}^p = [-\pi, \pi)^p \times [-\pi/4, \pi/4)^p$ scale as 2^{p+1}, and our numerical simulations suggest that the number of shared domains suitable for transferring parameters is two. However, more studies are needed to conclude this.

Finally, we discuss the applicability of our results to other COPs and VQAs. We find conditions such that adding a local field term to the cost Hamiltonian in QAOA preserves the two symmetries EWS and OWS, thus including general COPs described by Ising Hamiltonians [58]. Additionally, we show that the EWS is valid for all unitaries associated with Hamiltonians with $\hat{\sigma}_x \hat{\sigma}_x$, $\hat{\sigma}_y \hat{\sigma}_y$, or $\hat{\sigma}_z \hat{\sigma}_z$ interactions. Among these Hamiltonians, we consider the general case of the Heisenberg model on high-dimensional spaces with integer-valued couplings. For specific coupling values, this model reduces to the XYZ and XXZ models, previously studied with HVA [57]. We also examine two QAOA variants, RQAOA and ma-QAOA,

which improve the performance of the standard QAOA but require more parameter optimizations. We find that the iterative process of RQAOA only maintains the EWS, but breaks the OWS after the first iteration. Furthermore, we observe that the ma-QAOA could benefit from specific transferability schemes based on the symmetries of individual subgraphs, instead of the entire graph.

Our insights on the conditions of transferred parameters and their relation to the problem symmetries set the stage for a more comprehensive study of symmetry-based optimization in quantum algorithms. Future work may seek to numerically investigate the benefits of reducing the landscape size with these symmetries and performing warm-start local optimization with transferred parameters. For instance, the next steps would involve characterizing the number of iterations in the classical optimization compared to the current standard approaches.

APPENDIX A: SIMULATION METHODS

In this Appendix, we detail the methods used to simulate the QAOA and perform the classical optimization. All graphs are generated using the Python package NETWORKX [61], and the figures are created with the Python package MATPLOTLIB [62]. We simulate MAXCUT with d-regular graphs and random graphs on QAOA using the tensor-network-based package QTENSOR [63,64]. The optimization of the parameters was done with the QTENSOR implementation of the Python package PYTORCHS [65] optimizer RMSPROP, with 50 optimization steps.

We simulated MAXCUT on weighted d-regular graphs with QAOA using QISKIT [66], with 3000 shots per circuit evaluation. Here, the parameters' optimization was carried out with the Python package SCIPYS [67] and the optimizer differential_evolution with a population size of 15, and a maximum of 1000 function calls. For all optimizations, we used the optimal parameters to the relevant d-regular tree subgraphs as initial guesses.

APPENDIX B: SYMMETRIES OF OPTIMAL QAOA PARAMETERS ON MAXCUT

Here, we provide detailed proofs for the QAOA parameter symmetries presented in Sec. II, extending the results presented in Ref. [37] for unweighted graphs where all vertices have either odd or even degree. We identify two symmetries for regular graphs with integer weights that can be classified

according to the parity of the weights' sum on each vertex. In the following subsections, we present the results for graphs in which the weights of the edges connected to each vertex sum to an even and odd quantity, respectively.

1. Symmetries for graphs where the sum of weights connected to each vertex is even

In this section, we prove the symmetry called EWS in Sec. II. We consider a variational state constructed by a QAOA circuit of depth p. We focus on the mth iteration of the algorithm, in which the operation $\hat{U}_M(\beta_m)\hat{U}_C(\gamma_m)$ acts on the intermediate variational state

$$|\Psi_{m-1}\rangle = \prod_{j=m-1}^{1} e^{-i\beta_j\hat{H}_M} e^{-i\gamma_j\hat{H}_C}|+\rangle. \tag{B1}$$

Let us now assume that all eigenvalues of the diagonal cost Hamiltonian $\hat{H}_C$ are even, that is, for a computational basis n-qubit state $|y\rangle$, with $y = y_1y_2\ldots y_n$ using the binary representation $\hat{H}_C|y\rangle = 2\ell_y|y\rangle$, for $\ell_y \in \mathbb{Z}$. In this case, the operations $\hat{U}_C(\gamma_m)$ and $\hat{U}_C(\gamma_m \pm \pi)$ are equivalent, as we can observe by their action on $|\Psi_{m-1}\rangle = \sum_{y=0}^{2^n-1} c_y|y\rangle$. Explicitly,

$$
\begin{aligned}
e^{-i\gamma_m\hat{H}_C}\sum_y c_y|y\rangle &= \sum_y c_y e^{-i\gamma_m 2\ell_y}|y\rangle \\
&= \sum_y c_y e^{-i\gamma_m 2\ell_y} e^{\pm i2\ell_y\pi}|y\rangle \\
&= e^{-i(\gamma_m\pm\pi)\hat{H}_C}\sum_y c_y|y\rangle, \tag{B2}
\end{aligned}
$$

where we have introduced the global phases $e^{\pm i2\ell_y\pi} = 1$. That is, the output of the QAOA is invariant under changes $\gamma_m \to \gamma_m \pm \pi$ if the costs associated with each computational state $|y\rangle$ are even. For a graph instance of order n, the cost associated with the MAXCUT problem introduced in Eq. (1) evaluates to zero for the string of n ones $x^1 = 1^n$, $C(x^1) = 0$. Consequently, in the quantum formulation, $\hat{H}_C$ will have the even eigenvalue zero for the product state $|1\rangle^{\otimes n}$, $\hat{H}_C|1\rangle^{\otimes n} = 0$. In the following, we consider an arbitrary string $x \in \{-1, 1\}^n$ and show that changing the value of one bit changes its cost by an even amount. Therefore, since the cost of x^1 is even, all the costs are even in both the classical and quantum formulations, which in turn proves the symmetry.

The costs of Eq. (1) for two strings x and x' that differ only by the kth bit, such that $x'_k = -x_k$, are

$$C(x) = \frac{1}{2}\sum_{i\neq k}\sum_{j>i} w_{ij}(1 - x_i x_j) + \frac{1}{2}\sum_j w_{kj}(1 - x_k x_j)$$

and

$$C(x') = \frac{1}{2}\sum_{i\neq k}\sum_{j>i} w_{ij}(1 - x_i x_j) + \frac{1}{2}\sum_j w_{kj}(1 - x'_k x_j),$$

respectively. Thus, the cost difference is

$$C(x) - C(x') = \sum_j w_{kj}x'_k x_j. \tag{B3}$$

Since $x'_k x_j = \pm 1$, this difference will be an even quantity if the sum of all weights connected to a vertex k is even. This

requirement is equivalent to saying that the number of edges with odd integer weights connected to a vertex k is even. For the symmetry to hold for all bit changes, we require this criterion to hold for each vertex.

Particularly, any graph with only even integer weights will show the parameter symmetry $\gamma_m \to \gamma_m \pm \pi$, for any layer m, regardless of each vertex degree. Moreover, graphs with only odd weights where all vertices have even degree (thus including regular graphs of even degree) will share the same symmetry, with the specific examples of all odd weights $w_{ij} \in \{-1, 1\}$, and unweighted graphs when all vertices have even degree. This last example is already proven in Ref. [37]. In Sec. III A, we study numerically the transferability of the parameters for regular graphs of even degree with weights $w_{ij} \in \{-1, 1\}$.

2. Symmetries for graphs where the sum of weights connected to each vertex is odd

In this section, we prove the symmetry called OWS in Sec. II. As in the previous case, we consider the general case of a QAOA circuit of depth p and how the change $\gamma_m \to \gamma_m \pm \pi$ in the mth layer changes the final output state. Given the cost Hamiltonian $\hat{H}_C$ of Eq. (3) for MAXCUT on a weighted graph, the algorithm operation of layer m is

$$
\begin{aligned}
&\hat{U}_C(\gamma_m \pm \pi) \\
&= \prod_{(i,j)\in E}\exp\left[-i\frac{w_{ij}}{2}(\gamma_m \pm \pi)\left(\mathbb{1} - \hat{\sigma}_z^i\hat{\sigma}_z^j\right)\right] \\
&= \hat{U}_C(\gamma_m)\prod_{(i,j)\in E}\exp\left(\mp i\frac{w_{ij}\pi}{2}\right)\exp\left(\mp i\frac{w_{ij}\pi}{2}\hat{\sigma}_z^i\hat{\sigma}_z^j\right). \tag{B4}
\end{aligned}
$$

Here, we can ignore the global phase factors $\exp(\mp i w_{ij}\pi/2)$ and focus on the operational terms differing from $\hat{U}_C(\gamma_m)$, which can be rewritten as

$$\exp\left(\mp i\frac{w_{ij}\pi}{2}\hat{\sigma}_z^i\hat{\sigma}_z^j\right) = \cos\left(\frac{w_{ij}\pi}{2}\right) \mp i\sin\left(\frac{w_{ij}\pi}{2}\right)\hat{\sigma}_z^i\hat{\sigma}_z^j.$$

We only consider graphs with integer weights w_{ij}, which can be either even or odd. If a given weight w_{ij} is even, only the cosine differs from zero, which just adds a global phase. Otherwise, if a weight w_{ij} is odd, only the sine function is different from zero, and the resulting operation of the mth layer in QAOA changes. For convenience, we rewrite Eq. (B4) as

$$\hat{U}_C(\gamma_m \pm \pi) = \hat{U}_C(\gamma_m)e^{i\kappa}e^{i\delta}\prod_{\substack{(i,j)\in E \\ w_{ij}\text{ odd}}}\hat{\sigma}_z^i\hat{\sigma}_z^j, \tag{B5}$$

where the global phases encompass all the phase factors that do not change the final variational state. These phases include the contribution of all edges with even weights and the phase factors from the sine functions,

$$e^{i\kappa} = \prod_{(i,j)\in E}\exp\left(\mp i\frac{w_{ij}\pi}{2}\right)\prod_{\substack{(k,\ell)\in E \\ w_{\ell k}\text{ even}}}\cos\left(\frac{w_{\ell k}\pi}{2}\right), \tag{B6}$$

$$e^{i\delta} = \prod_{\substack{(i,j)\in E \\ w_{ij}\text{ odd}}} i\sin\left(\mp\frac{w_{ij}\pi}{2}\right). \tag{B7}$$

As we can observe from Eq. (B5), we recover the symmetry $\gamma_m \to \gamma_m \pm \pi$ for any graph with even integer weights, already proven in the previous section.

In the following, we address the case in which a graph can have both even and odd integer weights. We denote the set of edges connected to a vertex k as E_k and focus on that vertex and the weights of E_k. The contribution of these edges to the product in Eq. (B5) is

$$\prod_{\substack{(i,k)\in E_k \\ w_{ik}\ \text{odd}}} \hat{\sigma}_z^i \hat{\sigma}_z^k = \left(\hat{\sigma}_z^k\right)^{\mu_k} \prod_{\substack{i|(i,k)\in E_k \\ w_{ik}\ \text{odd}}} \hat{\sigma}_z^i, \qquad \text{(B8)}$$

where μ_k is the number of weights of the set E_k with an odd integer value. Now, we require that μ_k is either even or odd for all vertices in the graph, $k \in V$, that is, that every vertex has the same parity for the number of odd weighted edges connected to it. We can simplify Eq. (B5) as

$$\hat{U}_C(\gamma_m \pm \pi) = \hat{U}_C(\gamma_m)e^{i\kappa}e^{i\delta}\prod_{k\in V}\left(\hat{\sigma}_z^k\right)^{\mu_k}, \qquad \text{(B9)}$$

and we have two cases: if μ_k is even, then $(\hat{\sigma}_z^k)^{\mu_k} = \mathbb{1}$; otherwise, if μ_k is odd, then $(\hat{\sigma}_z^k)^{\mu_k} = \hat{\sigma}_z^k$. In the first case, with μ_k even for all $k \in V$, we recover again the symmetry rule $\gamma_m \to \gamma_m \pm \pi$ described in the previous section since the sum of all weights connected to each vertex is even. We note that if the cost Hamiltonian was expressed with $\hat{\sigma}_x\hat{\sigma}_x$ or $\hat{\sigma}_y\hat{\sigma}_y$ products instead of $\hat{\sigma}_z\hat{\sigma}_z$, similar derivations would hold. We discuss this fact further in our analyses of other VQAs in Appendix D.

We explore now the case in which μ_k is odd for all vertices $k \in V$ and include the action of the entire mth QAOA layer, also with the mixer unitary $\hat{U}_M(\beta_m)$. In particular, we consider changing its parameter $\beta_m \to \beta_m'$ such that we find algorithm symmetries. The complete algorithm step is then given by

$$\hat{U}_M(\beta_m')\hat{U}_C(\gamma_m \pm \pi) = \prod_{k\in V}\left[\cos(\beta_m') - i\,\sin(\beta_m')\hat{\sigma}_x^k\right]$$
$$\times\, e^{i\kappa}e^{i\delta}\hat{\sigma}_z^k\hat{U}_C(\gamma_m). \qquad \text{(B10)}$$

Using the relation $\hat{\sigma}_z^k\hat{\sigma}_x^k = -\hat{\sigma}_x^k\hat{\sigma}_z^k$, and $e^{i\theta} = e^{i\kappa}e^{i\delta}$, we can rewrite Eq. (B10) as

$$\hat{U}_M(\beta_m')\hat{U}_C(\gamma_m \pm \pi) = e^{i\theta}\prod_{k\in V}\hat{\sigma}_z^k\hat{U}_M(-\beta_m')\hat{U}_C(\gamma_m), \quad \text{(B11)}$$

which allows us now to study the action of these parameter changes in the final QAOA expectation value of the objective Hamiltonian $C(\boldsymbol{\gamma}, \boldsymbol{\beta})$ in Eq. (4). In particular, we recall the intermediate variational state $|\Psi_{m-1}\rangle$ of Eq. (B1), and construct the final state for the modified parameters as

$$|\boldsymbol{\gamma}', \boldsymbol{\beta}'\rangle = e^{i\theta}\prod_{j=p}^{m+1}\hat{U}_M(\beta_j')\hat{U}_C(\gamma_j)\prod_{k\in V}\hat{\sigma}_z^k$$
$$\times\, \hat{U}_M(-\beta_m')\hat{U}_C(\gamma_m)|\Psi_{m-1}\rangle, \qquad \text{(B12)}$$

with $\boldsymbol{\gamma}' = (\gamma_1, \ldots, \gamma_{m-1}, \gamma_m \pm \pi, \gamma_{m+1}, \ldots, \gamma_p)$ and $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_{m-1}, \beta_m', \beta_{m+1}', \ldots, \beta_p')$. Using again the commutation relations of the Pauli operators, we can rewrite

Eq. (B12) as

$$|\boldsymbol{\gamma}', \boldsymbol{\beta}'\rangle = e^{i\theta}\prod_{k\in V}\hat{\sigma}_z^k\prod_{j=p}^{m}\hat{U}_M(-\beta_j')\hat{U}_C(\gamma_j)|\Psi_{m-1}\rangle. \qquad \text{(B13)}$$

The expectation value of the objective Hamiltonian for the new parameters $C(\boldsymbol{\gamma}', \boldsymbol{\beta}')$ can be constructed from the probabilities of measuring each computational state $|y\rangle$, $P'(y) = \langle\boldsymbol{\gamma}', \boldsymbol{\beta}'||y\rangle\langle y||\boldsymbol{\gamma}', \boldsymbol{\beta}'\rangle$, with

$$P'(y) = \langle\Psi_{m-1}|\prod_{j=m}^{p}\hat{U}_C^\dagger(\gamma_j)\hat{U}_M^\dagger(-\beta_j')\prod_{k\in V}\hat{\sigma}_z^k|y\rangle$$
$$\times\, \langle y|\prod_{k\in V}\hat{\sigma}_z^k\prod_{j=p}^{m}\hat{U}_M(-\beta_j')\hat{U}_C(\gamma_j)|\Psi_{m-1}\rangle. \quad \text{(B14)}$$

Given that the $\hat{\sigma}_z^k$ operators commute with $|y\rangle\langle y|$, and $\prod_{k\in V}(\hat{\sigma}_z^k)^2 = \mathbb{1}$, the probabilities become

$$P'(y) = \langle\boldsymbol{\gamma}, \boldsymbol{\beta}'||y\rangle\langle y||\boldsymbol{\gamma}, \boldsymbol{\beta}'\rangle, \qquad \text{(B15)}$$

with

$$\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p),$$
$$\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_{m-1}, \beta_m', \beta_{m+1}', \ldots, \beta_p').$$

We observe that the QAOA symmetry, for which $P'(y) = P(y)$, with $P(y) = |\langle y|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle|^2$, occurs then for the simultaneous parameter changes

$$\boldsymbol{\gamma} \to (\gamma_1, \ldots, \gamma_m \pm \pi, \ldots, \gamma_p),$$
$$\boldsymbol{\beta} \to (\beta_1, \ldots, \beta_{m-1}, -\beta_m, -\beta_{m+1}, \ldots, -\beta_p). \quad \text{(B16)}$$

Thus, graphs where every vertex has an odd number of odd weighted edges connected to it are symmetric under the parameter changes of Eq. (B16). In other words, graphs where the sum of weights connected to each vertex is odd will exhibit these parameter symmetries.

In particular, graphs with only odd weights where all the vertices have odd degree (thus also regular graphs of odd degree) will exhibit this symmetry, which includes the unweighted graphs studied in Ref. [37]. We study numerically the transferability of the parameters for regular graphs of odd degree with weights $w_{ij} \in \{-1, 1\}$ in Sec. III A.

APPENDIX C: SYMMETRIES OF OPTIMAL QAOA PARAMETERS ON GENERAL COPs

In this Appendix, we extend the results of Appendix B and find the conditions for which the symmetries EWS and OWS, introduced in Sec. II, still apply if the cost Hamiltonian includes local field terms. In other words, a general cost Hamiltonian in the form of an Ising Hamiltonian can result in the parameter landscape having EWS or OWS. Many COPs can be mapped to an Ising Hamiltonian, including all of Karp's 21 NP-complete problems [58].

A general Ising Hamiltonian with local field terms reads as

$$\hat{H}_C = \frac{1}{2}\sum_{(i,j)\in E} w_{ij}\hat{\sigma}_z^i\hat{\sigma}_z^j + \frac{1}{2}\sum_{i\in V} h_i\hat{\sigma}_z^i, \qquad \text{(C1)}$$

where E is the set of edges in an interaction graph, and V is the set of spins. The QAOA unitary operation associated to

this cost Hamiltonian is

$$\hat{U}_C(\gamma_m) = \prod_{(i,j)\in E} \exp\left(-i\frac{w_{ij}}{2}\gamma_m\hat{\sigma}_z^i\hat{\sigma}_z^j\right)$$
$$\times \prod_{i\in V} \exp\left(-i\frac{h_i}{2}\gamma_m\hat{\sigma}_z^i\right). \qquad (C2)$$

We now consider the parameter shift $\gamma_m \pm \pi$, so that

$$\hat{U}_C(\gamma \pm \pi) = \hat{U}_C(\gamma_m)\prod_{(i,j)\in E} \exp\left(\mp i\frac{w_{ij}\pi}{2}\hat{\sigma}_z^i\hat{\sigma}_z^j\right)$$
$$\times \prod_{i\in V} \exp\left(\mp\frac{ih_i\pi}{2}\hat{\sigma}_z^i\right). \qquad (C3)$$

Analogously to Appendix B, we assume that the coefficients w_{ij} and h_i are integers. Then, we can rewrite the terms of Eq. (C3) as

$$\prod_{(i,j)\in E} \exp\left(\mp i\frac{w_{ij}\pi}{2}\hat{\sigma}_z^i\hat{\sigma}_z^j\right) = e^{i\varepsilon}e^{i\delta}\prod_{k\in V}(\hat{\sigma}_z^k)^{\mu_k}, \qquad (C4)$$

$$\exp\left(\mp\frac{ih_i\pi}{2}\hat{\sigma}_z^i\right) = \cos\left(\frac{h_i\pi}{2}\right) \mp i\,\sin\left(\frac{h_i\pi}{2}\right)\hat{\sigma}_z^i, \quad (C5)$$

where

$$e^{i\varepsilon} = \prod_{\substack{(k,\ell)\in E \\ w_{\ell k}\text{ even}}} \cos\left(\frac{w_{\ell k}\pi}{2}\right), \qquad (C6)$$

and $e^{i\delta}$ is given by Eq. (B7). For integer coefficients h_i, the right-hand side of Eq. (C5) is $\mathbb{1}$ if h_i is even, and $\mp i\hat{\sigma}_z^i$ if h_i is odd. Thus, we get

$$\prod_{i\in V} \exp\left(\mp\frac{ih_i\pi}{2}\hat{\sigma}_z^i\right) = e^{i\lambda}e^{i\nu}\prod_{k\in V}(\hat{\sigma}_z^k)^{h_k}, \qquad (C7)$$

with the global phases

$$e^{i\lambda} = \prod_{\substack{i\in V \\ h_i\text{ odd}}} i\,\sin\left(\mp\frac{h_i\pi}{2}\right) \text{ and } e^{i\nu} = \prod_{\substack{i\in V \\ h_i\text{ odd}}} \cos\left(\frac{h_i\pi}{2}\right).$$

We can finally rewrite Eq. (C3) as

$$\hat{U}_C(\gamma_m \pm \pi) = e^{i\alpha}\hat{U}_C(\gamma_m)\prod_{k\in V}(\hat{\sigma}_z^k)^{\mu_k+h_k}, \qquad (C8)$$

with the global phase $e^{i\alpha} = e^{i\varepsilon}e^{i\delta}e^{i\lambda}e^{i\nu}$. Now, we can follow a similar argument as in Appendix B considering the quantity $\mu_k + h_k$, instead of just μ_k, which was the number of odd-valued weights connected to vertex k. We conclude that the EWS applies if $\mu_k + h_k$ is even valued, and OWS applies if $\mu_k + h_k$ is odd valued.

APPENDIX D: SYMMETRIES OF OPTIMAL PARAMETERS IN OTHER VQAs

Here, we connect and extend the symmetries derived for the QAOA for the MAXCUT problem to the HVA for condensed matter problems, such as those described by Eq. (9) in Sec. IV. Moreover, we include a detailed derivation for similar

symmetry results in iterative quantum algorithms such as the RQAOA.

1. Symmetries for interaction graphs where the sum of weights connected to each vertex is even

We consider the unitary operator from Eq. (8),

$$\hat{U}_{\alpha\alpha}(\gamma) = \prod_{(i,j)\in E} \exp\left(-i\frac{w_{ij}}{2}\hat{\sigma}_\alpha^i\hat{\sigma}_\alpha^j\right), \qquad (D1)$$

with $\alpha = \{x, y, z\}$, and E the set of edges in an interaction graph. We assume that the weights satisfy $\sum_j w_{jk} = 2\mathbb{Z}$, $\forall\,k \in V$, where V is the set of all spins with an interaction. Now, following the same line of arguments made in Appendix B, we find that

$$\hat{U}_{\alpha\alpha}(\gamma \pm \pi) = \prod_{(i,j)\in E} \exp\left(-i\frac{w_{ij}}{2}\hat{\sigma}_\alpha^i\hat{\sigma}_\alpha^j\right)$$
$$= \hat{U}_{\alpha\alpha}(\gamma)\prod_{(i,j)\in E} \exp\left(\mp i\pi\frac{w_{ij}}{2}\hat{\sigma}_\alpha^i\hat{\sigma}_\alpha^j\right)$$
$$= \hat{U}_{\alpha\alpha}(\gamma)e^{i\kappa}e^{i\delta}\prod_{\substack{(i,j)\in E \\ w_{ij}\text{ odd}}} \hat{\sigma}_\alpha^i\hat{\sigma}_\alpha^j$$
$$= \hat{U}_{\alpha\alpha}(\gamma)e^{i\kappa}e^{i\delta}\prod_{k\in V}(\hat{\sigma}_\alpha^k)^{\mu_k}. \qquad (D2)$$

The global phases $e^{i\kappa}$ and $e^{i\delta}$ are given by Eqs. (B6) and (B7), respectively, μ_k is the number of odd-valued weights of the set E_k. The initial assumption made on the weights imply that μ_k is even, so $(\hat{\sigma}_\alpha^k)^{\mu_k} = \mathbb{1}$, and we retrieve that, up to a global phase,

$$\hat{U}_{\alpha\alpha}(\gamma \pm \pi) = \hat{U}_{\alpha\alpha}(\gamma). \qquad (D3)$$

It follows that *Ansätze* made up of products of operators on the form of Eq. (D1) will have EWS, like the one-dimensional Heisenberg XYZ model.

2. Symmetries for RQAOA

In this section, we show that RQAOA preserves EWS, but breaks OWS. The RQAOA is a recursive algorithm in which, after each iteration, one of the vertices with the largest correlation magnitude is effectively removed from the graph. Formally, if the maximal correlation is $|\mathcal{M}_{k\ell}|$, with $\mathcal{M}_{k\ell} = \langle\boldsymbol{\gamma}, \boldsymbol{\beta}|\hat{\sigma}_z^k\hat{\sigma}_z^\ell|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle$, one of the vertices k or ℓ is removed by imposing the constraint $\hat{\sigma}_z^\ell = \text{sgn}(M_{k\ell})\hat{\sigma}_z^k$. For our purpose, it is sufficient to consider $\hat{\sigma}_z^\ell = \pm\hat{\sigma}_z^k$.

With RQAOA, the dimension of the Hilbert space decreases every iteration. Thus, we can identify each distinct cost Hamiltonian of different iterations by the systems' size or number of qubits n. Here, we label the initial Hamiltonian as $\hat{H}_n$, the Hamiltonian after the recursive step $\hat{H}_{n-1}$, and so on. We are interested in how the sum of integer coefficients (weights and local field) connected to each vertex changes with each iteration. We recall from Sec. IV that, for graphs with EWS, the sum of coefficients connected to any vertex k is even, $\Sigma_j w_{jk} + h_k \in 2\mathbb{Z}$, $\forall\,k \in V$. In contrast, for graphs

with OWS, the sum of coefficients connected to any vertex k is odd, $\Sigma_j w_{jk} + h_k \in \mathbb{Z} \setminus 2\mathbb{Z}, \; \forall \; k \in V$.

Now, we consider an initial cost Hamiltonian $\hat{H}_n$ related to a graph that obeys either EWS or OWS. After the first iteration, we impose the constraint $\hat{\sigma}_z^\ell = \pm\hat{\sigma}_z^k$ such that the vertex ℓ is removed from the graph, or qubit ℓ is removed from the system. To analyze the symmetry of the new Hamiltonian $\hat{H}_{n-1}$, we first examine the terms of the initial cost Hamiltonian related to each vertex i, that is, the terms in which qubit i participates. We denote such terms as $\hat{H}_n^i = \sum_j w_{ij} \hat{\sigma}_z^i \hat{\sigma}_z^j + h_i$, and the total cost Hamiltonian can be written then as $\hat{H}_n = \sum_i \hat{H}_n^i / 2$.

For a given vertex i, we find that there are three possibilities: either i share no edge with the removed vertex ℓ, or i share one edge with the removed vertex ℓ, or $i = k$, meaning that it shares one edge with the removed vertex ℓ, and that edge is removed by the constraint. In the first case, if the vertex i does not share an edge with vertex ℓ, we observe that $\hat{H}_{n-1}^i = \hat{H}_n^i$. Second, if the vertex i shares an edge with vertex ℓ with the weight $w_{i\ell}$, after the RQAOA iteration, this vertex i will share a new edge with vertex k instead with weight $\pm w_{i\ell}$. Then, the initial Hamiltonian terms

$$\hat{H}_n^i = \sum_{j \neq \ell} w_{ij} \hat{\sigma}_z^i \hat{\sigma}_z^j + w_{i\ell} \hat{\sigma}_z^i \hat{\sigma}_z^\ell + h_i \hat{\sigma}_z^i \quad \text{(D4)}$$

will change to

$$\hat{H}_{n-1}^i = \sum_{j \neq \ell} w_{ij} \hat{\sigma}_z^i \hat{\sigma}_z^j \pm w_{i\ell} \hat{\sigma}_z^i \hat{\sigma}_z^k + h_i \hat{\sigma}_z^i. \quad \text{(D5)}$$

Let us consider the difference in the sum of coefficients between Eqs. (D4) and (D5). The weights of the edges not connected to vertex ℓ and the local field h_i remain unchanged, so the difference is given by $w_{i\ell} \pm w_{i\ell}$. This quantity is always even, as it is either equal to zero, or to $2w_{i\ell}$. Thus, the parity of the sum of coefficients around a vertex i in this second case does not change by the constraint imposed in the recursive algorithm. Lastly, we consider the vertices directly subject to the constraint, ℓ and k. The contribution to the initial cost Hamiltonian related to vertex k is

$$\hat{H}_n^k = \sum_{j \neq \ell} w_{jk} \hat{\sigma}_z^j \hat{\sigma}_z^k + w_{\ell k} \hat{\sigma}_z^\ell \hat{\sigma}_z^k + h_k \hat{\sigma}_z^k, \quad \text{(D6)}$$

while for vertex ℓ we have

$$\hat{H}_n^\ell = \sum_{j \neq k} w_{j\ell} \hat{\sigma}_z^j \hat{\sigma}_z^\ell + w_{\ell k} \hat{\sigma}_z^k \hat{\sigma}_z^\ell + h_\ell \hat{\sigma}_z^\ell. \quad \text{(D7)}$$

As the graph follows either the EWS or the OWS, the parity of the sums of coefficients in Eqs. (D6) and (D7) is the same. After the first RQAOA iteration, we get

$$\hat{H}_{n-1}^k = \sum_{j \neq \ell} w_{jk} \hat{\sigma}_z^j \hat{\sigma}_z^k \pm \sum_{j \neq k} w_{j\ell} \hat{\sigma}_z^j \hat{\sigma}_z^k + (w_{\ell k} \pm w_{\ell k})$$
$$+ (h_k \pm h_\ell) \hat{\sigma}_z^k, \quad \text{(D8)}$$

and $\hat{H}_{n-1}^\ell = 0$. The term $(w_{\ell k} \pm w_{\ell k})$ in Eq. (D8) is a constant energy shift, and not a weight on an edge.

We focus now on the sum of coefficients connected to vertex k in the Hamiltonian $\hat{H}_{n-1}^k$ of Eq. (D8), denoted as $s_{n-1}(k)$, which is

$$s_{n-1}(k) = \sum_{j \neq \ell} w_{jk} \pm \sum_{j \neq k} w_{j\ell} + (h_k \pm h_\ell)$$
$$= \sum_{j \neq \ell} w_{jk} + w_{\ell k} + h_k \pm \sum_{j \neq k} w_{j\ell} \pm w_{\ell k} \pm h_\ell$$
$$- w_{\ell k} \mp w_{\ell k}, \quad \text{(D9)}$$

where we have added and subtracted $w_{\ell k}$ twice. The first three terms of the second line in Eq. (D9) are equal to $s_n(k)$, the sum of coefficients that connected to vertex k in $\hat{H}_n^k$. Similarly, the next three terms correspond to $\pm s_n(\ell)$. We get

$$s_{n-1}(k) = s_n(k) \pm s_n(\ell) - w_{\ell k} \mp w_{\ell k}. \quad \text{(D10)}$$

As the graph in $\hat{H}_n$ is assumed to be either EWS or OWS, $s_n(k)$ and $s_n(\ell)$ have the same parity, so the sum, or difference, of these quantities must be even. The quantity $-w_{\ell k} \mp w_{\ell k}$ is also even, making Eq. (D10) even.

To summarize, after the constraint is imposed, all vertices except one are related to coefficients that sum to a value with unchanged parity. The exception is vertex k, the one directly affected by the constraint in the previous example. Regardless of the symmetry of the initial Hamiltonian $\hat{H}_n$, the sum of coefficients associated with it after the iteration, $s_{n-1}(k)$, is even. As such, RQAOA preserves EWS, but breaks OWS.

APPENDIX E: DERIVATION OF THE ANALYTIC EXPRESSION OF THE TRANSFERABILITY ERROR

In this Appendix, we show a detailed derivation of the closed-form expression for $\eta_{d,d'}^{g>3}$ from Eq. (7). The expression is exact for unweighted triangle-free d-regular graphs, shown in Fig. 2, and approximate for unweighted low-girth d-regular graphs, shown in Fig. 3.

From Ref. [52], the objective function of an unweighted triangle-free graph is given by

$$C_d = \frac{|E|}{2}[1 + \sin 4\beta \sin \gamma \cos^{d-1} \gamma], \quad \text{(E1)}$$

evaluated at $(\gamma_d^*, \beta_d^*) = (\arctan \frac{1}{\sqrt{d-1}}, \frac{\pi}{8})$ yields

$$C_d^* = \frac{|E|}{2}\left[1 + \frac{1}{\sqrt{d}}\left(\frac{d-1}{d}\right)^{\frac{d-1}{2}}\right]. \quad \text{(E2)}$$

The maximum number of edges that can be cut is $|E|$, and the minimum number of edges that can be cut is $\frac{|E|}{2}(1 + \frac{1}{d}) - \frac{1}{4}$. For our purpose, it is sufficient to lower bound the cut to half the edges, so $\frac{C_{\max}}{|E|} = k \in (0.5, 1]$. The minimum cut value is $C_{\min} = 0$, so $r_d = \frac{C_d}{C_{\max}}$. From the definition of $\eta_{d,d'}$ from Eq. (6) we get

$$\eta_{d,d'} = \frac{1}{k|E|}[C_d(\gamma_d^*, \beta_d^*) - C_d(\gamma_{d'}^*, \beta_{d'}^*)]. \quad \text{(E3)}$$

Since we consider triangle-free receiver graphs, we replace $C_d(\gamma_d^*, \beta_d^*)$ with Eq. (E2). The donor graph is the d'-regular tree subgraph so $C_d(\gamma_{d'}^*, \beta_{d'}^*)$ is Eq. (E1) evaluated

TABLE II. Eight sets of optimal parameters to the 3-regular tree subgraph at depth $p = 2$ within $\mathcal{A}^2 = [-\pi, \pi)^2 \times [-\pi/4, \pi/4)^2$. Sets 1 and 8 are in the domains $\mathcal{U}_1^2$ and $\mathcal{U}_2^2$, respectively.

	γ_1	β_1	γ_2	β_2
Set 1	0.156π	0.177π	0.286π	0.0933π
Set 2	0.156π	0.177π	-0.714π	-0.0933π
Set 3	0.844π	0.177π	0.714π	-0.0933π
Set 4	0.844π	0.177π	-0.286π	0.0933π
Set 5	-0.844π	-0.177π	0.286π	-0.0933π
Set 6	-0.844π	-0.177π	-0.714π	0.0933π
Set 7	-0.156π	-0.177π	0.714π	0.0933π
Set 8	-0.156π	-0.177π	-0.286π	-0.0933π

at $(\gamma_{d'}^*, \beta_{d'}^*) = (\arctan \frac{1}{\sqrt{d'-1}}, \frac{\pi}{8})$. Then, we get

$$
\eta_{d,d'}^{g>3} = \frac{1}{2k} \left[\frac{1}{\sqrt{d}} \left(\frac{d-1}{d} \right)^{(d-1)/2} \right.
$$
$$
\left. - \cos^{d-d'} \gamma_{d'}^* \frac{1}{\sqrt{d'}} \left(\frac{d'-1}{d'} \right)^{(d'-1)/2} \right]. \quad \text{(E4)}
$$

Using the trigonometry relation

$$
\cos^n(\arctan x) = \left(\frac{1}{1+x^2} \right)^{\frac{n}{2}},
$$

we get the final expression

$$
\eta_{d,d'}^{g>3} = \frac{1}{2k\sqrt{d}} \left[\left(\frac{d-1}{d} \right)^{(d-1)/2} - \sqrt{\frac{d}{d'}} \left(\frac{d'-1}{d'} \right)^{(d'-1)/2} \right]. \quad \text{(E5)}
$$

Note that the exponent is $\frac{d-1}{2}$ for both terms.

APPENDIX F: OPTIMAL PARAMETERS FOR THE 3-REGULAR TREE SUBGRAPH AT $p = 2$

Here, we present the eight sets of optimal parameters to the 3-regular tree subgraph at $p = 2$ that we used as donors in Fig. 6 of Sec. III. The sets are taken from Ref. [45], where the authors define the parameters $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) \in [0, 2\pi)^2 \times [0, \pi/2)^2$. However, to follow the convention used in this article, we have translated the parameters into $\mathcal{A}^p$ using the OWS symmetry. We provide the optimal parameter values in Table II.

[1] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028

[2] E. Farhi, D. Gamarnik, and S. Gutmann, The quantum approximate optimization algorithm needs to see the whole graph: Worst case examples, arXiv:2005.08747.

[3] E. Farhi, D. Gamarnik, and S. Gutmann, The quantum approximate optimization algorithm needs to see the whole graph: A typical case, arXiv:2004.09002.

[4] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2022), pp. 335–343.

[5] A. Anshu and T. Metger, Quantum **7**, 999 (2023).

[6] G. De Palma, M. Marvian, C. Rouzé, and D. S. França, PRX Quantum **4**, 010309 (2023).

[7] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, Phys. Rev. Lett. **125**, 260505 (2020).

[8] R. Herrman, P. C. Lotshaw, J. Ostrowski, T. S. Humble, and G. Siopsis, Sci. Rep. **12**, 6781 (2022).

[9] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Phys. Rev. X **10**, 021067 (2020).

[10] L. Bittel and M. Kliesch, Phys. Rev. Lett. **127**, 120502 (2021).

[11] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Nat. Rev. Phys. **3**, 625 (2021).

[12] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, PRX Quantum **3**, 010313 (2022).

[13] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. O. Marrero, M. Larocca, and M. Cerezo, Nat. Commun. **15**, 7172 (2024).

[14] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability?

Or, why we need to rethink variational quantum computing, arXiv:2312.09121.

[15] J. Rajakumar, J. Golden, A. Bärtschi, and S. Eidenbenz, in *Proceedings of the 21st ACM International Conference on Computing Frontiers, CF '24* (ACM, New York, 2024), pp. 199–206.

[16] M. Fernández-Pendás, E. F. Combarro, S. Vallecorsa, J. Ranilla, and I. F. Rúa, J. Comput. Appl. Math. **404**, 113388 (2022).

[17] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).

[18] G. Acampora, A. Chiatto, and A. Vitiello, Appl. Soft Comput. **142**, 110296 (2023).

[19] M. Alam, A. Ash-Saki, and S. Ghosh, in *Proceedings of the 23rd Conference on Design, Automation and Test in Europe, DATE '20* (EDA Consortium, San Jose, CA, 2020), pp. 686–689.

[20] P.-Y. Wang, M. Usman, U. Parampalli, L. C. L. Hollenberg, and C. R. Myers, IEEE Trans. Quantum Eng. **4**, 3100620 (2023).

[21] A. Deshpande and A. Melnikov, Symmetry **14**, 2593 (2022).

[22] L. Cheng, Y.-Q. Chen, S.-X. Zhang, and S. Zhang, Commun. Phys. **7**, 83 (2024).

[23] X. Lee, N. Xie, D. Cai, Y. Saito, and N. Asai, Mathematics **11**, 2176 (2023).

[24] D. J. Egger, J. Mareček, and S. Woerner, Quantum **5**, 479 (2021).

[25] M. Cain, E. Farhi, S. Gutmann, D. Ranard, and E. Tang, The QAOA gets stuck starting from a good classical string, arXiv:2207.05089.

[26] R. Tate and S. Eidenbenz, Guarantees on warm-started QAOA: Single-round approximation ratios for 3-Regular MAXCUT and higher-round scaling limits, arXiv:2402.12631.

[27] M. H. Muñoz-Arias, S. Kourtis, and A. Blais, Phys. Rev. Res. **6**, 023294 (2024).

[28] G. S. Ravi, P. Gokhale, Y. Ding, W. M. Kirby, K. N. Smith, J. M. Baker, P. J. Love, H. Hoffmann, K. R. Brown, and F. T.

Chong, Cafqa: A classical simulation bootstrap for variational quantum algorithms, in *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS 2023* (Association for Computing Machinery, New York, NY, 2022), Vol. 1, pp. 15–29.

[29] F. G. S. L. Brandão, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, For fixed control parameters the quantum approximate optimization algorithm's objective function value concentrates for typical instances, arXiv:1812.04170.

[30] A. Galda, E. Gupta, J. Falla, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, Similarity-based parameter transferability in the quantum approximate optimization algorithm, Front. Quantum Sci. Technol. **2**, 1200975 (2023).

[31] M. Streif and M. Leib, Quantum Sci. Technol. **5**, 034008 (2020).

[32] R. Shaydulin, P. C. Lotshaw, J. Larson, J. Ostrowski, and T. S. Humble, ACM Trans. Quantum Comput. **4**, 1 (2023).

[33] S. H. Sureshbabu, D. Herman, R. Shaydulin, J. Basso, S. Chakrabarti, Y. Sun, and M. Pistoia, Quantum **8**, 1231 (2024).

[34] J. A. Montañez-Barrera, D. Willsch, and K. Michielsen, Transfer learning of optimal QAOA parameters in combinatorial optimization, arXiv:2402.05549.

[35] J. Falla, Q. Langfitt, Y. Alexeev, and I. Safro, Quantum Machine Intelligence **6**, 46 (2024).

[36] R. Shaydulin, C. Li, S. Chakrabarti, M. DeCross, D. Herman, N. Kumar, J. Larson, D. Lykov, P. Minssen, Y. Sun, Y. Alexeev, J. M. Dreiling, J. P. Gaebler, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, N. Hewitt, C. V. Horst, S. Hu *et al.*, Sci. Adv. **10**, eadm6761 (2024).

[37] P. C. Lotshaw, T. S. Humble, R. Herrman, J. Ostrowski, and G. Siopsis, Quantum Inf. Proc. **20**, 403 (2021).

[38] A. A. Mele, G. B. Mbeng, G. E. Santoro, M. Collura, and P. Torta, Phys. Rev. A **106**, L060401 (2022).

[39] J. A. Montañez-Barrera and K. Michielsen, Towards a universal QAOA protocol: Evidence of quantum advantage in solving combinatorial optimization problems, arXiv:2405.09169.

[40] P. Díez-Valle, D. Porras, and J. J. García-Ripoll, Phys. Rev. Lett. **130**, 050601 (2023).

[41] P. C. Lotshaw, G. Siopsis, J. Ostrowski, R. Herrman, R. Alam, S. Powers, and T. S. Humble, Phys. Rev. A **108**, 042411 (2023).

[42] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou (unpublished).

[43] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, Quantum **6**, 759 (2022).

[44] D. Lykov, J. Wurtz, C. Poole, M. Saffman, T. Noel, and Y. Alexeev, npj Quantum Inf. **9**, 73 (2023).

[45] J. Wurtz and D. Lykov, Phys. Rev. A **104**, 052419 (2021).

[46] E. Wybo and M. Leib, Missing puzzle pieces in the performance landscape of the quantum approximate optimization algorithm, arXiv:2406.14618.

[47] A. Galda, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Piscataway, NJ, 2021).

[48] R. Shaydulin, I. Safro, and J. Larson, in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, Piscataway, NJ, 2019).

[49] R. Shaydulin and S. M. Wild, IEEE Trans. Quantum Eng. **2**, 3101409 (2021).

[50] 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).

[51] M. X. Goemans and D. P. Williamson, J. ACM **42**, 1115 (1995).

[52] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Phys. Rev. A **97**, 022304 (2018).

[53] R. Diestel, *Graph Theory*, 3rd ed., Graduate Texts in Mathematics (Springer, Berlin, 2006).

[54] P. Erdős and A. Rényi, Publicat. Math. Debrecen **6**, 290 (2022).

[55] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[56] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[57] R. Wiersema, C. Zhou, Y. de Sereville, J. F. Carrasquilla, Y. B. Kim, and H. Yuen, PRX Quantum **1**, 020319 (2020).

[58] A. Lucas, Front. Phys. **2**, 5 (2014).

[59] E. Bae, H. Kwon, V. Vijendran, and S. Lee, Improved recursive qaoa for solving max-cut on bipartite graphs, arXiv:2408.13207.

[60] I. Gaidai and R. Herrman, Sci. Rep. **14**, 18911 (2024).

[61] A. Hagberg, P. Swart, and D. S Chult, Exploring network structure, dynamics, and function using NetworkX, Technical Report, Los Alamos National Lab. (LANL), Los Alamos, NM, 2008 (unpublished).

[62] J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).

[63] D. Lykov, A. Galda, and Y. Alexeev, QTENSOR, https://github.com/danlkv/QTensor.

[64] D. Lykov, A. Chen, H. Chen, K. Keipert, Z. Zhang, T. Gibbs, and Y. Alexeev, in *2021 IEEE/ACM Second International Workshop on Quantum Computing Software (QCS)* (IEEE, Piscataway, NJ, 2021), pp. 27–34.

[65] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai *et al.*, PyTorch: An imperative style, high-performance deep learning library, arXiv:1912.01703.

[66] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta, Quantum computing with QISKIT, https://github.com/Qiskit/qiskit.

[67] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, Nat. Methods **17**, 261 (2020).