Symmetry-enhanced counterdiabatic quantum algorithm for qudits

Alberto Bottarelli,[1,2,*,†] Mikel Garcia de Andoin,[3,4,5,*] Pranav Chandarana,[4,6,5] Koushik Paul,[4,5,‡] Xi Chen,[7] Mikel Sanz,[4,5,8,9] and Philipp Hauke[1,2]

[1]*Pitaevskii BEC Center, CNR-INO and Dipartimento di Fisica, Universita di Trento, I-38123 Trento, Italy*
[2]*INFN-TIFPA, Trento Institute for Fundamental Physics and Applications, Trento, Italy*
[3]*TECNALIA, Basque Research and Technology Alliance (BRTA), 48160 Derio, Spain*
[4]*Department of Physical Chemistry, University of the Basque Country (UPV/EHU), 48940 Leioa, Spain*
[5]*EHU Quantum Center, University of the Basque Country (UPV/EHU), 48940 Leioa, Spain*
[6]*Kipu Quantum, Greifswalderstrasse 226, 10405 Berlin, Germany*
[7]*Instituto de Ciencia de Materiales de Madrid (CSIC), Cantoblanco, E-28049 Madrid, Spain*
[8]*IKERBASQUE, Basque Foundation for Science, 48009 Bilbao, Spain*
[9]*Basque Center for Applied Mathematics (BCAM), 48009 Bilbao, Spain*

(Received 26 May 2025; accepted 7 September 2025; published 8 October 2025)

Qubit-based variational quantum algorithms have undergone rapid development in recent years but still face several challenges. Here, we introduce a symmetry-based enhancement to digitized counterdiabatic quantum algorithms, applicable for qudits of any dimension. This approach offers three types of compression compared to conventional variational circuits. First, compression in the circuit depth is achieved by counterdiabatic protocols. Second, information about the problem is compressed by replacing qubits with qudits, allowing for a more efficient representation of the problem. Finally, the number of parameters is reduced by employing the symmetries of the system. We illustrate this approach by tackling a graph-based optimization problem MAX-3-CUT, a highly entangled state preparation, the qutrit W state, and a two-body only antiferromagnetic Ising problem. As our numerical results show, we achieve a better convergence with a lower circuit depth and less measurement overhead, albeit with some identified limitations for which we propose a work-around. This work leads to a better design of shallow variational quantum circuits, improving the feasibility of their implementation on near-term qudit devices.

DOI: 10.1103/6ldg-3w1f

I. INTRODUCTION

Over the past decade, quantum computing has made significant strides, showing great promise for advancing both scientific research and industrial applications [1–3]. This paves the way for solving complex computational problems, previously considered insurmountable using classical computers [4–7]. Quantum algorithms like Grover's search [8] and Shor's factorization [9] have been proposed to attain quantum advantage, particularly in a fault-tolerant quantum computing scenario. However, achieving practical quantum utility remains challenging in the current noisy intermediate scale quantum (NISQ) [1,10] era, despite several attempts [6,11]. This is due to limiting factors like decoherence, dephasing, and large error rates in the present quantum hardware. Most of the current quantum computation is dominated by a qubit-based model, developed to imitate the binary model of classical computation. This might not necessarily be optimal for maximizing computational efficiency and exploiting the full potential of quantum hardware capabilities.

In recent years, there has been a growing interest in utilizing quantum systems with multiple levels, known as qudits, as fundamental components in quantum computing. Qudits offer the potential to enhance computational power and efficiency by reducing the number of required computational units [12]. A qudit is formally defined as a quantum system with $d > 2$ levels, where the local dimension d varies depending on the physical nature of the system. Thanks to advances in quantum hardware and control techniques, universal qudit quantum computers have been realized with trapped $^{40}Ca^{+}$ ions [13], superconducting circuits [14], Rydberg atoms [15], and photonic devices [16].

Qudit-based quantum computing presents many advantages compared to the usual $d = 2$ counterpart [17]. Representing a given Hilbert space on a qudit quantum computer requires fewer particles by a factor of $\log_d 2$ than a qubit-based quantum computer. This is particularly favorable for NISQ devices since the increased number of qubits needed to represent a large Hilbert space becomes more vulnerable to noise and more challenging to control [18,19]. The high dimensionality of the qudits also makes them exploitable in different areas of quantum computing. For instance, in quantum simulation, they can be utilized to investigate high-order truncation of

*These authors contributed equally to this work.
†Contact author: alberto.bottarelli@unitn.it
‡Contact author: koushikpal09@gmail.com

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

quantum link models [20,21]. Moreover, many optimization problems like graph coloring and MAX-k-CUT, which are not based on binary variables, are naturally suited to qudit platforms. Both of these tasks can be studied either through qudit circuit-based variational quantum algorithms (VQAs) [21,22] or qudit-based adiabatic computing [23].

VQAs offer a promising approach to bridge the gap between current quantum capabilities and the goal of achieving quantum utility with the support of classical algorithms [24]. For instance, the quantum approximate optimization algorithm (QAOA) [25], derived from the principles of adiabatic quantum computing, has been proposed to address a wide range of problems, including combinatorial optimization [26]. However, these algorithms require substantial circuit depth to attain the desired solution, especially for large system sizes. The quality of solutions obtained from VQAs can be improved through various approaches like efficient parametrization and improved initialization strategies [27,28]. Apart from these, quantum control methods such as counterdiabatic (CD) protocols [29,30] can also be utilized to design shallow quantum circuits that reduce the resource requirement and obtain solutions [31].

In this article, we define a symmetry-enhanced CD-inspired variational ansatz for qudit systems. In Ref. [32], symmetries of graph-defined problems were used to improve the performance of VQAs and reduce the overhead on both the classical and quantum hardware. Utilizing symmetries is crucial because they reduce the algorithm's resources in the form of the number of parameters, which is a bottleneck for current VQAs. In this ansatz, the symmetries of the system are exploited to significantly reduce the number of free optimizable parameters. As the main contribution of this work, we define the theoretical framework that merges digitized counterdiabatic quantum optimization with graph permutation symmetries. This allows us to obtain a compression in three regimes. First, information compression is achieved by using qudits instead of qubits. Second, circuit compression is obtained using CD protocols. Third, compression in parameter space is achieved by leveraging symmetries. We demonstrate the efficiency of our method by implementing it on the MAX-k-CUT graph problem, the W state preparation, and an antiferromagnetic two-body only Ising Hamiltonian. We show that this combination significantly compresses the variational protocols. However, we find limitations on the performance on our ansätze, showing that using symmetries might reduce the expressivity of the circuit.

This article is organized as follows: In Sec. II, we review qudit-based quantum computing on trapped ions, variational algorithms, and CD protocols. In Sec. III, we derive the theory of symmetric CD terms. In Sec. IV, we numerically test the proposed symmetric CD ansätze, compare it to their fully parametrized counterparts and QAOA, and discuss their performance. Finally, we provide an overview of our results and conclude in Sec. V.

II. PRELIMINARIES

Propelled by recent advances in their experimental control [13–16], qudit-based algorithms and their applications in solving diverse problems are currently receiving significant interest in the quantum computing community [12]. This section provides an overview of the foundational elements of qudits, which outlines the definition of qudit operations crucial for the present study. Additionally, it explores their role in designing qudit-based VQAs, highlighting their potential for efficient problem-solving. Furthermore, the section briefly discusses the design of CD ansätze, an advanced technique aimed at enhancing VQAs.

A. Qudits gates and basis

A qudit is defined as a quantum system with d possible states and local Hilbert space with dimension C^d, with the local computational basis identified as $|0\rangle, ..., |d-1\rangle$. Unitaries acting on the qudit are part of the $U(d)$ unitary group and can be implemented with arbitrary error ϵ via a universal gate set. Just like the qubit case, this set will depend on the physical realization of the qudit. For example, qudits can be realized by splitting the $4^2S_{\frac{1}{2}}$ and $3^2D_{\frac{3}{2}}$ energy levels of trapped $^{40}Ca^+$ ions with a magnetic field, which results in a total of eight levels for the qudit [13]. The universal gate set for such a system is provided by equatorial rotations

$$R^{i,j}(\vartheta,\phi) = \exp\left(-i\theta\sigma_\phi^{i,j}/2\right) \tag{1}$$

combined with the entangling MS gate

$$\mathrm{MS}^{i,j}(\vartheta,\phi) = \exp\left(-\frac{i\theta}{4}\left(\sigma_\phi^{i,j}\otimes\mathbb{1}+\mathbb{1}\otimes\sigma_\phi^{i,j}\right)^2\right), \tag{2}$$

where $\sigma_\phi^{i,j} = \cos(\phi)\sigma_X^{i,j} \pm \sin(\phi)\sigma_Y^{i,j}$ and the indices i, j of the Pauli matrices are related to the levels of the qudits that are being rotated. Each two-level transition between the S and D levels can be addressed via a single narrowband laser at 729 nm. We encourage the reader to refer to Ref. [13] for all the details. In the computational basis, the angular momentum operators are defined as

$$L_z|m\rangle = m|m\rangle, \tag{3}$$

$$L_+|m\rangle = \sqrt{(l-m)(l+m+1)}\,|m+1\rangle, \tag{4}$$

$$L_-|m\rangle = \sqrt{(l+m)(l+m+1)}\,|m-1\rangle, \tag{5}$$

$$L_x = \frac{1}{2}(L_+ + L_-), \tag{6}$$

$$L_y = \frac{1}{2i}(L_+ - L_-), \tag{7}$$

where $l = (d-1)/2$ is the index of the angular momentum representation. We use these operators to describe our Hamiltonians in the rest of the work. Each unitary operator generated by this set can equally be generated via the universal gate set defined by Eqs. (1) and (2).

B. Qudit-based variational quantum algorithms

In VQAs [24], the ground state of a problem Hamiltonian H_P is obtained by utilizing both quantum and classical resources. The quantum part evolves an initial state $|\psi_0\rangle$ to a final state $|\psi_f\rangle$ by means of a parametrized ansatz $U(\boldsymbol{\theta})$. For the initial state, it is necessary to have an overlap with the target state, and for this, usually, the equal superposition state

is prepared. The classical part optimizes the set of parameters $\boldsymbol{\theta}$ such that $|\psi_f\rangle$ is close to the desired ground state. This task is accomplished by minimizing a cost function such as the expectation value of H_P, the fidelity with the ground state, etc. The performance of a VQA heavily depends upon the design of $U(\vec{\theta})$ and the cost function employed [33].

Among the several ways to define the ansatz, QAOA is a paradigmatic example [25]. QAOA takes inspiration from adiabatic quantum computing, where we have a Hamiltonian

$$H_a(t) = (1 - \lambda(t))H_0 + \lambda(t)H_P, \tag{8}$$

where $\lambda(t)$ is the scheduling function interpolating from H_0 to H_P. The system is initialized in the ground state of H_0 and, when the adiabatic condition is satisfied, we can find the ground state of H_P with a high success probability. In QAOA, one splits the continuous time evolution into discrete layers of alternating evolutions under H_0 and H_P and replaces the coefficients of both Hamiltonians with trainable parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. Therefore, the evolution $U(\boldsymbol{\theta})$ has the form

$$U(\boldsymbol{\theta}) = \prod_p e^{-i\theta_{1,p}H_0} e^{-i\theta_{2,p}H_P}, \tag{9}$$

where $\theta_{k,p}$ is the variational parameter for the kth term in the pth layer.

To minimize the cost function, the classical routines search for optimal parameters from the information obtained by computing its gradients [34] or by using other heuristics [35]. In most cases, one requires several circuit evaluations depending on the classical optimization method to calculate the parameters for the next step. For example, one requires at least one circuit evaluation per parameter to find gradients for the next step when using gradient-based optimization, yielding a parameter-based measurement overhead. Since the quantum resources in the NISQ era are limited, one has to design the ansatz with a minimum circuit depth and an optimal number of parameters $|\boldsymbol{\theta}|$ to balance the expressivity and measurement overhead.

Most of the previous works on VQAs have been conducted on qubit systems. Employing qudits is a natural extension, with the major advantage of opening up the possibility of finding more efficient encodings. In particular, qudit systems can enhance the performance of VQA when applied to problems that naturally present multidimensional elements, such as for lattice-gauge theories [22], integer optimization problems [21], or multiclass classification tasks [36]. However, similar to qubits, this approach faces the problem of designing ansätze with enough expressivity, while avoiding overparametrization [37] and the barren plateaus [38].

C. Digitized counterdiabatic quantum optimization

A key issue in adiabatic quantum algorithms for ground-state search is to avoid diabatic transitions to higher excited states. One way to mitigate this is CD driving, where the Hamiltonian $H_a(t)$ in Eq. (8) is augmented to $H(t)$ given by

$$H(t) = H_a(t) + \dot{\lambda}(t)A_\lambda^{(\ell)}, \tag{10}$$

where $A_\lambda^{(\ell)} = i\sum_{k=1}^{\ell} \alpha_k(t)Q_{2k-1}$ denotes the auxiliary CD term with $Q_k = [H_a(t), Q_{k-1}]$ and $Q_0 = \partial_\lambda H_a(t)$. Here, ℓ denotes the order of approximation. The term $\dot{\lambda}(t)A_\lambda^{(\ell)}$ introduces the CD correction to suppress nonadiabatic transitions.

The coefficients $\alpha_k(t)$ can be obtained in several ways including action minimization [39,40], variational optimization [41], machine learning [31,42], Krylov subspace methods [43], etc. In this work, we consider the action minimization method, where we have to find

$$\min_{\alpha_k} \mathrm{Tr}\big(G_\ell^2\big), \tag{11}$$

with $G_\ell = Q_0 - i[H_a(t), A_\lambda^{(\ell)}]$. The structure of the operator G_ℓ contains a term, $\partial_\lambda H_0$, that is independent of the weights of CD driving and a sum of parameters that depend on the weights. Thus, we can rewrite this operator as $G_\ell = B_0 + \sum_k \alpha_k B_k$. Note that for first-order CD terms ($\ell = 1$), the operators B_k will contain at most two-body terms when H_P is two-local and H_0 is local. This allows us to simulate the evolution under the CD Hamiltonian by using a decomposition of the usually available single-body and two-body native gate set, as it is the case in the experimental implementation described in Sec. II A. We will work with this setup hereafter.

Once we find the optimal CD terms, we discretize the evolution using the first-order Trotter decomposition given by

$$U_{\mathrm{Trott}}(0, T) = \prod_{j=1}^{n}\prod_{m=1}^{M} \exp[-i\Delta t C_m(j\Delta t)H^{(m)}] + O(\Delta t^2), \tag{12}$$

where n is the total number of Trotter steps, i.e., $T = n\Delta t$, $H^{(m)}$ is each of the terms of $H(t)$, with $C_m(t)$ its corresponding variational parameter, and M is the total number of terms present in $H(t)$. One can also utilize a hybrid quantum-classical algorithm known as digitized counterdiabatic QAOA (DCQAOA), where we set coefficients of each term in Eq. (10) as a trainable parameter optimized by classical optimization [44]. Therefore, one can implement additional unitary $U_{\mathrm{CD}} = e^{-i\sum_k \theta_k P_k}$ corresponding to the CD term in each layer. Here, P_k shows the kth term in the linear combination of Pauli operator tensor products corresponding to $A_\lambda^{(\ell)}$. The evolution $U(\boldsymbol{\theta})$ is of the form

$$U(\boldsymbol{\theta}) = \prod_p e^{-i\theta_{1,p}H_0} e^{-i\theta_{2,p}H_P} e^{-i\sum_{k=1}\theta_{2+k,p}P_k}. \tag{13}$$

In the impulse regime where $|\alpha_k(t)\dot{\lambda}(t)| \gg |\lambda(t)|$, we can neglect the adiabatic Hamiltonian $H_a(t)$ and implement evolutions corresponding to only CD terms [45,46]. The unitary evolution for this algorithm is then given by

$$U(\boldsymbol{\theta}) = \prod_p e^{-i\sum_{k=1}\theta_{k,p}P_k}. \tag{14}$$

Both DCQAOA and CD-inspired ansätze share a parametrization scaling of $O(N^2)$. For many optimizers, the circuit evaluations needed for each iteration depend upon the number of parameters. To address this challenge, in this work, we leverage symmetries to strategically reduce the number of trainable parameters in a problem-specific manner.

III. SYMMETRY-ENHANCED CD-INSPIRED ANSATZ

Symmetries in physics can be employed to condense the information of a system in the least amount of parameters. This is a property that is used to improve algorithms treating systems consisting of many particles, for instance, in the Hartree-Fock method [47], or density-matrix renormalization group (DMRG) [48]. In this section, we show how to exploit the symmetries of the system to reduce the number of parameters in the variational approach with CD terms.

Let us begin with the symmetries that act on the system and in particular, on the operator G_ℓ. Here and in the rest of this article, we focus only on the spatial symmetries of the system [49]. These symmetries π act on a Hermitian operator O permuting the positions of the qudits such that $\pi O \pi^{-1} = O$; i.e., the action of the symmetry changes the labels of the qudits but leaves the operator invariant.

We denote the set of symmetries of H_P with $S_P = \{\pi_q\}$, where q labels the elements of the set. It is straightforward to see that since H_0 is symmetric under the action of all the operations of the permutation group S_n, it will also have the ones of S_P, $S_P \subseteq S_n$. From this, it follows that the adiabatic Hamiltonian will have the same symmetries as H_P. Additionally, the CD Hamiltonian will inherit these symmetries, as the term $\partial_\lambda H(t)$ maintains them as well, $\partial_\lambda H(t) = \pi_q \partial_\lambda H(t) \pi_q^{-1} = \partial_\lambda(\pi_q H(t) \pi_q^{-1})$. Considering $\{A_k\}$ as the CD terms of the system, the action of each of the symmetries from S_P will only change the labels of these terms ($k \to \pi_q(k) = k_q$):

$$\begin{aligned} A_\lambda^{(\ell)} &= \sum_k \alpha_k A_k = \pi_q \left(\sum_k \alpha_k A_k \right) \pi_q^{-1} \\ &= \sum_{k_q} \alpha_{k_q} A_{k_q}. \end{aligned} \tag{15}$$

Now, we can identify each of the terms, such that the only difference after applying the symmetries is in the labels for the parameters. Thus, for any symmetry π_q we have

$$A_\lambda^{(\ell)} = \pi_q A_\lambda^{(\ell)} \pi_q^{-1} = \sum_k \alpha_{k_q} A_k. \tag{16}$$

Now, we formulate the main property of the symmetries. Informally, we say a set of parameters $\{\alpha_k\}_g$ can be grouped, i.e., have the same value, if their labels k form a closed set under the action of all symmetries in S_P. Rigorously, we define the set of grouped parameters $\{\alpha_k\}_g$ as the set that fulfills $\pi_q(\alpha_k \in \{\alpha_k\}_g) \in \{\alpha_k\}_g \; \forall \pi_q \in S_P$.

Interestingly for our problem, this allows us to reduce the total number of variational parameters. Recall the expression of operator G_ℓ, in which we can identify each of the operators B_k with a CD term A_k. The operators B_k inherit the symmetries of A_k as both the terms in $B_k \sim [H_0, A_k]$ share the same symmetries. Thus, the grouping of parameters effectively reduces the number of parameters in the action minimization problem by $|\{\alpha_k\}_g| - 1$ for each of the parameter groups.

To check this reduction, we have numerically computed a particular problem, the ZZ-Ising Hamiltonian $H_P = \sum_i \sigma_i^z + \sum_{i,j \in \mathcal{G}} \sigma_i^z \sigma_j^z$ with a local mixer Hamiltonian $H_0 = \sum_i \sigma_i^x$. We have calculated the number of parameters after applying the reduction technique and compare it with number of parameters in the original problem. As shown in Fig. 1, we have computed this for all nontrivial problems up to eight qubits and observed a reduction to $59 \pm 19\%$ of the original number of parameters. Additionally, the standard approach and the symmetry-enhanced method used to approximate the CD terms both converge to the same result. Therefore, this reduction of parameters can also be extended to the CD-inspired VQA. Detailed calculations are given in Appendix A.

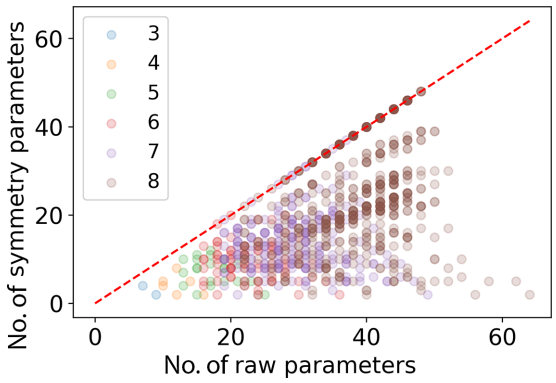


FIG. 1. Number of unique parameters in the CD ansatz after using the spatial permutation symmetries vs the total number of unique parameters without taking into account the symmetries. Here, we have checked all ZZ-Ising Hamiltonians as in Eq. (17) with a connected graph structure up to eight nodes that can be found in the database from Ref. [50]. The labels indicate the size of the system.

To give a clearer idea of the parameter reduction procedure, let us begin with an example of an $N = 4$ qudit system in which the problem Hamiltonian is given by

$$H_P = \sum_{i=1}^{4} L_{zi} + \sum_{i=1}^{3} L_{zi} L_{zi+1}, \tag{17}$$

i.e., a Hamiltonian with connectivity graph in the form of a linear path, which we denote as $\mathcal{G} = (V_\mathcal{G} = [1, 2, 3, 4], E_\mathcal{G} = [(1, 2), (2, 3), (3, 4)])$, with $V_\mathcal{G}$ and $E_\mathcal{G}$ the vertices and edges of $\mathcal{G}$, respectively. Given a mixer Hamiltonian of the form $H_0 = \sum_{i=1}^4 L_{xi}$, the CD Hamiltonian is given by

$$A_\lambda^{(1)} = \sum_{i \in V_\mathcal{G}} \alpha_i L_{y_i} + \sum_{\{i,j\} \in E_\mathcal{G}} (\alpha_{i,j} L_{y_i} L_{z_j} + \alpha_{j,i} L_{y_j} L_{z_i}). \tag{18}$$

Notably, even though the interactions of the problem Hamiltonian do not have any sense of directionality, the interactions in the CD Hamiltonian do. In particular, we cannot interchange $L_{y_j} L_{z_i} \leftrightarrow L_{y_i} L_{z_j}$ without swapping their corresponding parameters. Then, we can identify the terms in the CD Hamiltonian with a directed graph $\mathcal{H}$ in which each edge $E_\mathcal{G}$ is included as an arc of $A_\mathcal{H}$ in both directions as depicted in Fig. 2. As a result, we can rewrite the CD Hamiltonian as

$$A_\lambda^{(1)} = \sum_{i \in V_H} \alpha_i L_{y_i} + \sum_{\{i,j\} \in A_H} \alpha_{i,j} L_{y_i} L_{z_j}. \tag{19}$$

To reduce the number of parameters employing this technique, one has to identify the different groups. Above, we defined the grouping of parameters such that it coincides with the definition of the graph orbits [32,51]. For dealing with the cases in which the CD terms have directionality, we use the

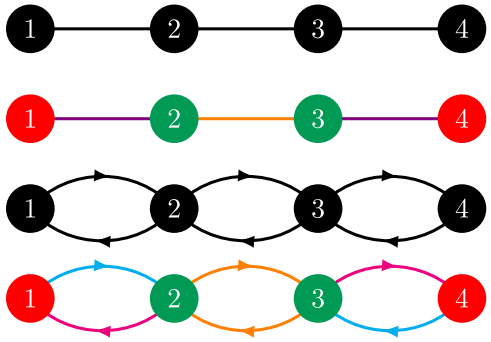


FIG. 2. Example of graphs and their symmetries. From top to bottom: $\mathcal{G}$, $\mathcal{G}$ with vertex and edge orbits, $\mathcal{H}$, and $\mathcal{H}$ with vertex and arc orbits. The colors indicate the different elements in each vertex and arc/orbits, which indeed indicate how the parameters are grouped.

extended notion of edge orbits to the arcs of a directed graph [52]. For calculating arc orbits, we obtain the edge orbits of the undirected graphs by employing the same algorithm as proposed in Ref. [32]. We detail the steps to group the parameters in Algorithm 1. For differentiating the edge orbits on directed graphs from the ones on undirected graphs, we will call the former arc orbits.

In our particular case, the vertex orbits of $\mathcal{G}$ are $[1, 4]$ and $[2, 3]$, and the edge orbits $[(1, 2), (3, 4)]$ and $[(2, 3)]$. Utilizing Algorithm 1, we then identify the arc orbits $[(1, 2), (4, 3)]$, $[(2, 1), (3, 4)]$, and $[(2, 3), (3, 2)]$. Subsequently, we can reduce the total number of variational parameters in the system from 10 to 5, as from the symmetries it follows that $\alpha_1 = \alpha_4$, $\alpha_2 = \alpha_3$, $\alpha_{12} = \alpha_{43}$, $\alpha_{21} = \alpha_{34}$, and $\alpha_{23} = \alpha_{32}$.

To sum up, the procedure to generate a symmetry-enhanced CD-inspired ansatz is as follows. First, obtain the CD terms $A_\lambda^{(\ell)}$ from H_P and H_0. Then, compute the arc and vertex orbits of the associated graph. Finally, when parametrizing the ansatz, include the information about the parameter groups obtained in the previous step. It is important to note that in tensor networks and DMRG, leveraging symmetries to reduce parameters may sometimes worsen the convergence [53,54]. However, in quantum systems, utilizing symmetries consistently decreases the number of measurements required, which is advantageous for near-term devices. In the next section, we illustrate with numerical examples that this reduction can lead to concrete improvements of key figures of merit of quantum optimization.

ALGORITHM 1. Algorithm for obtaining the arc orbits of the directed graph $\mathcal{H}$.

Get the undirected graph $\mathcal{G}$ from $\mathcal{H}$
Calculate vertex and edge orbits of $\mathcal{G}$
Define arc orbits as tuple (*arcs*, *tail vertex orbit*, *head vertex orbit*, *edge orbit*)
for each edge orbit **do**
  Identify the arcs (h, t) of $\mathcal{H}$ that fulfill $(h, t) = (i, j)$ or $(h, t) = (j, i)$, where (i, j) are the edges of $\mathcal{G}$
  for each arc **do**
    Identify the vertex orbit of the tail and head
    Add the arc to the arc orbit with the same tail vertex, head vertex, and edge orbits
  end for
end for

IV. NUMERICAL RESULTS

To test the validity of the proposed techniques for practical applications in quantum optimization, we need problems whose encoding in qudits is both efficient in terms of the number of particles and mapping of the cost function into a single- and two-body interaction Hamiltonian. Also, we require problems with naturally occurring symmetries. This is the case for some problems defined over lattices or graphs.

A. MAX-3-CUT

As a paradigmatic application of quantum optimization, we investigate a combinatorial optimization problem. In this regard, we are not particularly interested in optimization problems in which each instance is heavily determined by weights. Thus, we select the MAX-k-CUT problem [55], with instances for which no trivial or greedy solution exists. The objective of this problem is to find the best partition of the vertices into k colors, such that the number of edges connecting two vertices belonging to different colors is maximized. The objective is expressed as

$$\max_{\mathbf{c}} \sum_{\{i,j\}\in E} \begin{cases} 1, & \text{if } c_i \neq c_j, \\ 0, & \text{else,} \end{cases} \tag{20}$$

where $\mathbf{c} = \{c_i\}^N$ is the vector assigning each vertex i to the color $1 \leqslant c_i \leqslant k$ and the size of the problem N is the number of the vertices of the graph. Since the instances are uniquely defined by a graph, the symmetries from the graph will be used to reduce the number of parameters in our variational approach. The conversion of this classical problem into the problem of finding the ground state of a Hamiltonian can be done with two-local even-order terms. For the particular case of the MAX-3-CUT, we have

$$H_P = \sum_{\{i,j\}\in E} L_{z_i} L_{z_j} - 2\left(L_{z_i}^2 + L_{z_j}^2\right) + 3L_{z_i}^2 L_{z_j}^2. \tag{21}$$

Even though we limit our simulations to the MAX-3-CUT, the results here are easily extensible to any k, since to obtain the Hamiltonian for the general MAX-k-CUT problem one just needs to solve a linear system of $O(k^2)$ equations, as shown in Appendix B.

Using this problem, we have tested the performance of our proposal with two $p = 1$ layered ansätze against several similar approaches and the QAOA. We start with the DCQAOA ansätze at first order as proposed in Ref. [44]. Using our approach, we refine this ansätze by grouping the parameters of the CD terms. We also tested similar circuits in which we only use the CD terms, which are also typically known as CD-inspired ansätze [45]. As a baseline for comparison, we employed the usual QAOA for qudits (qudit-QAOA) as proposed originally in Refs. [22,25]. Using qudits also for qudit-QAOA allows us to make a comparison while having all algorithms working within the same Hilbert space of dimension d^N. For the benchmark, we have selected six instances

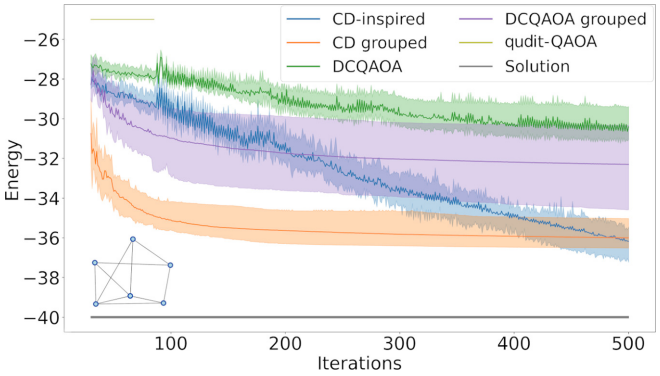


FIG. 3. Expected value of the energy at the end of each iteration of the variational procedure for a Max-3-Cut instance. We show the results for different ansätze for a single layer. We only show the results after 30 steps of the classical optimizer to discard the initial fluctuations. As shown, the CD ansatz with the grouped parameters shows a faster convergence compared to the fully parametrized CD ansatz. After a high number of iterations, the fully parametrized ansatz obtains better results due to its higher expressivity. These results are also true for the DCQAOA variants. QAOA fails to solve the problem due to its limited expressivity even though the depths of the ansätze are similar.

of the smallest nontrivial size with $N = 6$ qudits of $d = 3$ levels, commonly known as qutrits. The graphs for the instances have been selected from the database in Ref. [50], with the identifier (HoG IDs) 728, 220, 730, 758, 752, and 748. The selection criteria we employed are graphs with six nodes, fully connected, and nonplanar [56]. We have run up to 500 iterations of the COBYLA optimizer [57] for each of the classical update loops because this algorithm does not require computing costly derivatives. For each problem instance, we have run multiple different initializations with uniformly random parameters.

The results we have obtained for the expected value of the energy for the first graph are shown in Fig. 3. The full comparison is given in Appendix B. Additionally, approximation ratios are shown in Table I, which are defined as $\mathcal{R} = \langle H_P \rangle / E_0$ with E_0 the ground-state energy. We observe that all four algorithms implementing CD terms reach lower energies than QAOA with few iterations. Moreover, we see that the ansätze with grouped terms quickly reach energies close to the convergence energy. For both the grouped and nongrouped versions, the DCQAOA ansätze seem to need a higher number of iterations to converge. This is counterintuitive as the expressivity of the DCQAOA ansätze is the same as that of the CD ansätze if we set the QAOA parameters to 0. However, the reduction of the convergence speed might be due to the appearance of barren plateaus [38] and overparametrization effects [37]. We also stress that even though the performance of the grouped ansatz is slightly worse than the ungrouped one, we still achieve a good result with a reduced amount of variational parameters. Although the simulations have not been extensive, for all four problem instances checked, the grouping method consistently shows a performance improvement. Our approach not only improves the convergence speed or the approximation ratio as shown, but more importantly decreases the number of circuit evaluations.

B. *W* state preparation

As a second test, we study the ability of the CD symmetrized ansatz to prepare the W state of a system of N qutrits:

$$|W\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} P_i(|2\rangle^{\otimes N-1} |0\rangle), \tag{22}$$

where P_i indicates permutations of the labels of the local basis states. Note that the W state is also a qutrit Dicke state D_k^N with $k = 1$, as shown in Appendix C. This is an equal superposition of all the possible ground states of the Hamiltonian

$$H_P = \left(\frac{N}{2} - 1\right) \sum_{i=1}^{N} L_{zi} + \frac{1}{2} \sum_{j<i} L_{zi} L_{zj}. \tag{23}$$

This Hamiltonian can be represented by a fully connected graph with a local potential on each vertex. Grouping the coefficients, in this case, is trivial since the graph is fully connected and the resulting CD Hamiltonian is Eq. (18) with $\alpha_i = \alpha \ \ \forall \ \ i$ and $\alpha_{i,j} = \tilde{\alpha} \ \ \forall \ \ i, j$. This grouping of the coefficients enforces the symmetry of the unitary transformation on the system, which would be lost if all the parameters were left independent. Since the W state is the equal superposition of the N possible ground states, it is necessary to guide the evolution toward this specific state and not to a generic superposition of ground states. The metric with which we measure the efficiency of the state preparation is the fidelity of the final state with the state in Eq. (22). Given that the states are pure, the fidelity simply becomes the expectation value of the evolution operation over the chosen initial state $|\psi_0\rangle$ and the

TABLE I. Mean approximation ratios and standard deviation for the algorithms at the final iteration for the different problems. Values in bold showcase that the obtained result yields a better $\mathcal{R}$ value than the classical approximation ratio of 0.800 217 [58]. Underlined values mark the best performance for each problem in terms of the mean approximation ratio $\mathcal{R}$. QAOA converged to the same value of energy for each graph up to machine precision.

Problem graph ID	728	220	730	758	752	748
CD all	**0.90 ± 0.04**	**0.96 ± 0.01**	**0.85 ± 0.03**	**0.91 ± 0.03**	**0.92 ± 0.02**	**0.83 ± 0.01**
CD grouped	**0.90 ± 0.05**	**0.95 ± 0.03**	**0.92 ± 0.06**	**0.92 ± 0.03**	**0.93 ± 0.03**	**0.85 ± 0.03**
DCQAOA all	0.76 ± 0.03	**0.91 ± 0.04**	0.74 ± 0.04	**0.81 ± 0.01**	**0.87 ± 0.03**	**0.81 ± 0.01**
DCQAOA grouped	**0.81 ± 0.06**	**0.90 ± 0.04**	**0.81 ± 0.04**	**0.87 ± 0.06**	**0.85 ± 0.02**	**0.87 ± 0.06**
Qudit-QAOA	0.63	0.76	0.63	0.69	0.75	0.68

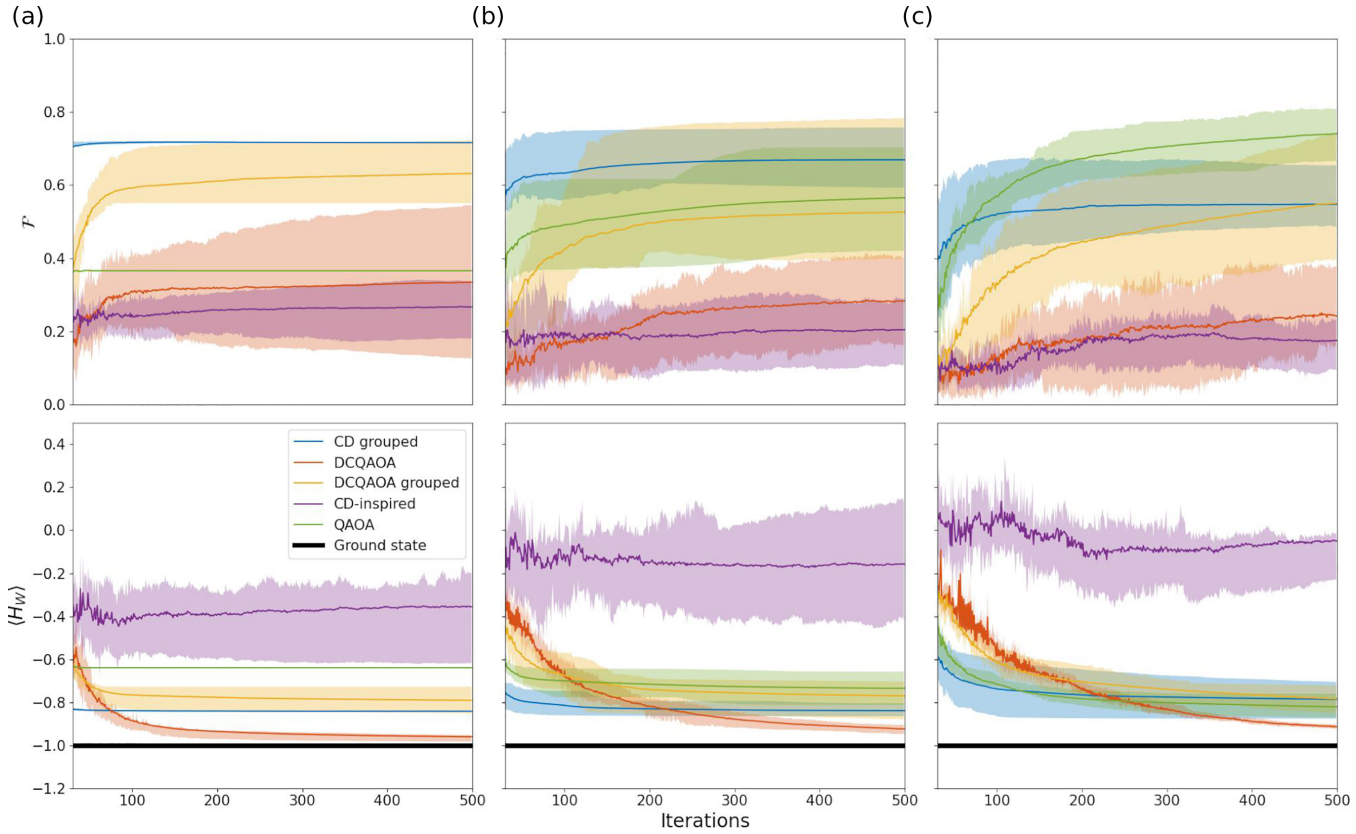


FIG. 4. Fidelity defined in Eq. (24) (top) and the expectation value of the Hamiltonian of Eq. (23) (bottom) for the ansätze considered, shown as a function of the optimizer iteration. The results refer to the case $N = 3$ and the three subfigures show the results for different numbers of layers: (a) $p = 2$, (b) $p = 3$, and (c) $p = 5$, as ordered from left to right. Also here, we omit the first 30 iterations to avoid showing initial fluctuations of the optimizer. We see that the grouped CD ansatz outperforms the other ones for a low number of layers, as expected.

target state:

$$\mathcal{F} = |\langle W|U(\boldsymbol{\theta})|\psi_0\rangle|^2. \tag{24}$$

For our study, we have used

$$|\psi_0\rangle = \frac{1}{\sqrt{d^N}} \bigotimes_{j=1}^{N} \sum_{i_j=0}^{d-1} |i_j\rangle\,. \tag{25}$$

The cost function that we minimize in the algorithm is the expectation value of H_P; thus, the optimization procedure will drive the state toward lower energy. In Appendix C, we also show the results that are obtained by maximizing the fidelity instead of minimizing the energy. The symmetry of the W state will be imposed on our proposed ansätze. Thus, we expect a better performance in terms of fidelity compared to the fully parametrized ones, even if the latter has final energy closer to the ground-state energy.

We executed the algorithm for the ansätze described in Sec. III from $p = 1$ to $p = 5$ layers, $N = 3$ qudits, and show the results in Fig. 4. The symmetry-enhanced CD ansatz outperforms QAOA for a small number of layers but gets worse as their number increases. This can be attributed to the fact that for low values of p we are in the impulse regime, while for larger p the ansatz is closer to the adiabatic regime. Furthermore, the fully parametrized ansätze approximate the ground-state energy faster but have worse results in terms of fidelity. Specifically, for just a few layers ($p = 2$) we can achieve fidelity with the W state of $\mathcal{F} \sim 71\%$.

C. When are symmetries useful for CD ansätze?

In two previous examples, we have observed that the use of the symmetries can be beneficial for the optimization problem, in particular, for improving the convergence speed of the optimization process while maintaining competitive results. This is explained through the reduction of the number of parameters of the circuit. However, there is still a question to answer about the expressivity of the circuits when we use the symmetries. For example, for the Max-3-Cut problem, the symmetries help to make the convergence much faster, implying an easier optimization landscape. At the same time, we also observe that the fully parametrized CD ansatz can reach better solutions, indicating a better expressivity, albeit with a harder optimization landscape.

Intuitively, when the problem Hamiltonian includes single- and two-body terms, the leading order of the nested commutator expansion (10) generates both single- and two-body terms. Using such terms in the approximated ansätze for Eq. (13) can effectively make up for the diabatic higher-order terms, even with a few layers. However, for problem

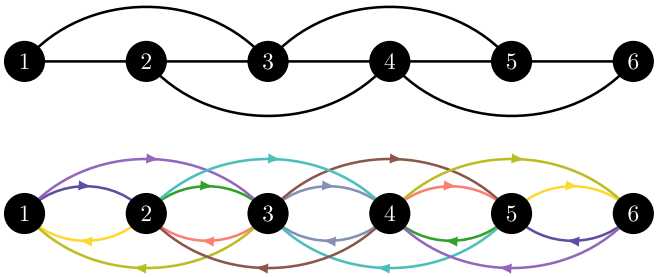


FIG. 5. Top: Interaction scheme for Hamiltonian (26). Bottom: Parameter reduction for the symmetrized DC ansatz. Each line represents a DC term in the Hamiltonian, with lines of the same color indicating terms optimized using the same parameter. This grouping reduces the number of optimized parameters from 18 to 9.

Hamiltonians that contain only two-body terms, the leading-order commutator with the driver Hamiltonian will generate only contributions from two-body terms. Using only these in the ansatz has less flexibility to compensate for the higher-order terms. As we will see in Eq. (13), adding single-body terms in the ansatz significantly improves the performance of the algorithm.

In terms of the hardness of the optimization, reducing the number of parameters is beneficial to ease the optimization landscape, specially true for single- and two-body Hamiltonians. However, the improvement is marginal where only two-body terms are involved. This can be attributed to the low number of parameters in both the symmetry-enhanced and fully parametrized ansätze. Consequently, use of the symmetries in these cases further restricts the ability of the ansatz to compensate higher-order contributions with the two-body terms.

As a rule of thumb, when working with single- and two-body Hamiltonians, there is a trade-off between the expressivity and the hardness of the optimization landscape. Thus, we can choose to utilize the symmetries based on which of the parameters we want to favor, expressivity or convergence speed. However, in cases where only two-body terms are present, the subtle advantage in convergence speed gained from using the symmetries does not justify the significant loss in expressivity.

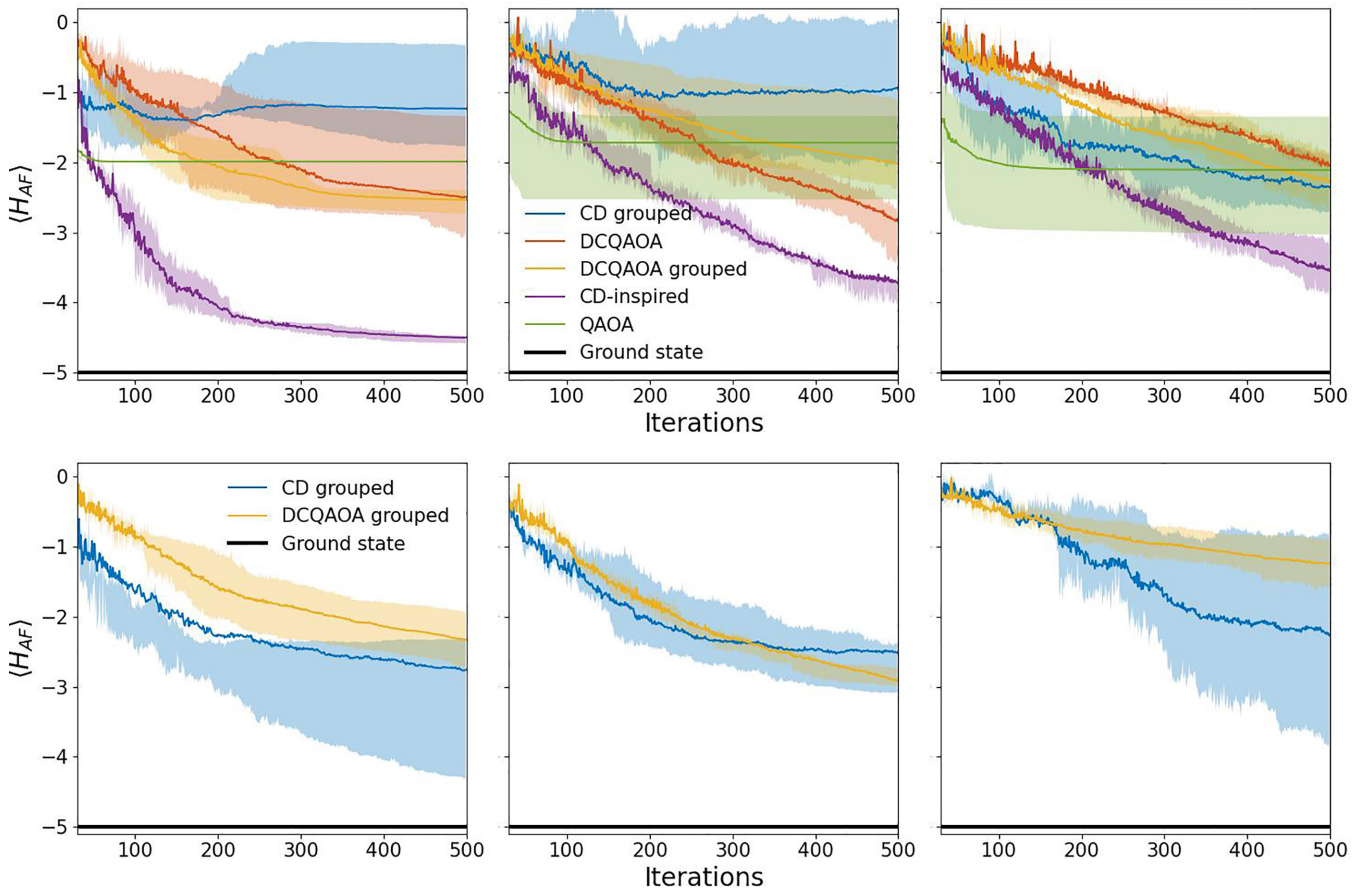


FIG. 6. Top: Expectation value of Hamiltonian (26) over optimizer iterations for all of the ansätze without artificial single-body terms. Results are shown for 2, 3, and 4 layers, from left to right. Bottom: Same figure of merit with grouped ansätze to which we add the layer of single-body terms $U = \prod_{i=0}^{N-1} e^{i\alpha_i L_{y_i}}$. It is possible to see that the fully parametrized CD-inspired ansatz converges to a vale closer to the round state in a smaller amount of iterations compared to the grouped CD and grouped CDQAOA ansätze. This difference is reduced when considering the symmetric ansätze enriched with single-body terms, even though the fully parametrized ansatz provides faster convergence to a value closer to the ground state.

In order to confirm the intuition that using the symmetries for only two-body Hamiltonians might not be adequate, in the following, we perform a test on an antiferromagnetic Ising Hamiltonian. Additionally, we introduce single-body terms that are not obtained from the nested commutator, without losing generality, in order to check the usefulness of single-body terms in enhancing expressivity.

D. Antiferromagnetic second-neighbor Ising Hamiltonian

To confirm the trade-off explained above, we study the ability of the CD symmetrized ansatz to optimize a next-to-nearest neighbor Ising model of N qutrits with open boundary conditions, whose Hamiltonian is defined as

$$H_{\mathrm{AF}} = \sum_{\langle\langle ij\rangle\rangle} L_{z_i} L_{z_j} + \sum_{\langle ij\rangle} L_{z_i} L_{z_j}. \tag{26}$$

The system is symmetric under the index permutation $i \to N - i$, and hence we can group the parameters according to Fig. 5.

In order to drive the system toward its ground state, we consider as objective function the expectation value of the Hamiltonian over the parametrized state:

$$C = \langle\psi_0|\, U^\dagger(\boldsymbol{\theta}) H_{\mathrm{AF}} U(\boldsymbol{\theta})\, |\psi_0\rangle\,. \tag{27}$$

As can be seen from the results in Fig. 6, there is a substantial difference in the performance of the grouped ansätze if single-body terms are artificially included in the ansatz. In particular, we see an improvement for the first layers (2, 3) while for the fourth layer the performance of the ansatzes with added single bodies is comparable to the original ansatz. Also, we see a great improvement of the CD ansatz with grouped parameters for all sizes. These results coincide with the expected performance discussed in the previous subsection.

V. CONCLUSION

In this article, we have introduced a symmetry-enhanced CD-inspired ansatz for variational qudit algorithms. The main advantage of incorporating CD terms lies in the digital circuit compression compared to adiabatic protocols. To further speed up the convergence rate of the algorithms, we have utilized the symmetries of the problem to group the variational parameters. The reduction of the parameters also decreases the measurement overhead needed to advance during the optimization process. This is achieved by analyzing how the symmetries are propagated from the problem Hamiltonian to the adiabatic gauge potential and finally to the CD ansätze.

To test the validity of our approach, we have addressed three problems: the classical optimization problem MAX-k-CUT, the preparation of a W state, and the ground-state preparation of an antiferromagnetic second-neighbor Ising Hamiltonian, using up to $p = 5$ layer ansätze. For the former, we have run a test on four nontrivial instances and compared our approach against various state-of-the-art algorithms. For single-layer ansatz, we have observed better performance of all approaches that employ CD terms. In the case of the preparation of the ground state of the Next-To-Nearest Ising Hamiltonian, i.e., the W state preparation, we have investigated the problem with the same suite of algorithms, demonstrating the utility of our strategy in preparing this state with fewer resources. Additionally, when studying the effect of the symmetries, we confirm a faster convergence and better performance on the first two problems, but not the latter, due to the lack of expressivity of only two-body term symmetric ansätze. This can be compensated by introducing additional single-body terms that also respect the symmetries.

The compression in circuit depth, information encoding, and parameters positions this approach as a feasible candidate for experimental implementations. Further works could also discuss if using approximate symmetries could be beneficial, as it would open more possibilities for reducing variational parameters with them in a wider set of problems. Another possibility would be to explore the possibilities of efficient qubit to qudit mappings and their impact on the performance of CD protocols.

ACKNOWLEDGMENTS

We thank Julian Ferreiro for insights about counterdiabatic driving and expressivity of the counterdiabatic terms. A.B. acknowledges funding from the Honda Research Institute Europe. P.H. acknowledges funding by the European Union under Horizon Europe Programme, Grant Agreement No. 101080086—NeQST. This project has received funding from the Italian Ministry of University and Research (MUR) through the FARE grant for the project DAVNE (Grant No. R20PEX7Y3A), was supported by the Provincia Autonoma di Trento, and Q@TN, the joint laboratory between University of Trento, FBK—Fondazione Bruno Kessler, INFN—National Institute for Nuclear Physics, and CNR—National Research Council. Project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—call for Tender No. 1031 of 17/06/2022 of Italian Ministry for University and Research funded by the European Union—NextGenerationEU (Project No. CN_00000013). Project DYNAMITE QUANTERA2_-00056 WAS funded by the Ministry of University and Research through the ERANET COFUND QuantERA II—2021 call and cofunded by the European Union (H2020, Grant Agreement No. 101017733). M.G.d.A., P.C., K.P., and M.S. acknowledge funding from OpenSuperQ+100 (Grant No. 101113946) of the EU Flagship on Quantum Technologies, EU FET-Open project EPIQUS (Grant No. 899368), Spanish Ramón y Cajal Grant No. RYC-2020-030503-I funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and "ERDF Invest in your Future," Spanish Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM ENIA project call—Quantum Spain, EU through the Recovery, Transformation and Resilience Plan—NextGenerationEU within the framework of the Digital Spain 2026 Agenda, Basque Government through Grant No. IT1470-22, and IKUR Strategy under the collaboration agreement between Ikerbasque Foundation and BCAM on behalf of the Department of Education of the Basque Government. M.G.d.A. acknowledges support from the UPV/EHU and TECNALIA 2021 PIF contract call, from the Basque Government through the "Plan complementario de comunicación cúantica"

(EXP.2022/01341) (A/20220551), from the Basque Government through the ELKARTEK program, project "KUBIT—Kuantikaren Berrikuntzarako Ikasketa Teknologikoa" (KK-2024/00105), and from the Spanish Ministry of Science and Innovation under the Recovery, Transformation and Resilience Plan (CUCO, MIG-20211005), Spanish CDTI through Plan complementario Comunicación cuántica (EXP. 2022/01341)(A/20220551).

Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Commission, the European Union, or of the Ministry of University and Research. Neither the European Union nor the granting authority can be held responsible for them.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: THEORY OF SYMMETRY GROUPED CD TERMS

Assume that the problem Hamiltonian H_P has some symmetries $S = \{\pi_k\}$ that can belong to the group of permutations. Now, the Hamiltonian is invariant under the action of S:

$$H_P = \pi_k H_P \pi_k^{-1} \ \forall k. \tag{A1}$$

From this, $H_a(t)$ will also BE symmetric:

$$H_a(t) = \pi_k H_a(t) \pi_k^{-1}. \tag{A2}$$

The objective is to speed up the usual annealing process by employing CD terms as detailed in the main text. Recall the expression of the CD term at first order:

$$Q_1(t) = [H(t), \partial_t H(t)]. \tag{A3}$$

The CD Hamiltonian will inherit the symmetries S exactly, since

$$\begin{aligned} Q_1(t) &= \pi_k Q_1(t) \pi_k^{-1} = \pi_k [H_a(t), \partial_t H_a(t)] \pi_k^{-1} \\ &= \left[\pi_k H_a(t) \pi_k^{-1}, \pi_k \partial_t H_a(t) \pi_k^{-1}\right] \\ &= \left[\pi_k H_a(t) \pi_k^{-1}, \partial_t \left(\pi_k H_a(t) \pi_k^{-1}\right)\right] \\ &= [H_a(t), \partial_t H_a(t)]. \end{aligned} \tag{A4}$$

The only effect of acting with the permutation symmetries is a change in the labels of the CD terms. To solve the problem of the original labeling of the qubits and the permuted one, one can identify the identical elements associated with the labels of the parameters and group them.

Example: Assume we have a CD Hamiltonian with seven different terms and a single symmetry π. When calculating Q_1, one gets the following structure:

$$\begin{aligned} Q_1 = \alpha_1 B_1 + \alpha_2 B_2 + \alpha_3 B_3 + \alpha_4 B_4 \\ + \alpha_5 B_5 + \alpha_6 B_6 + \alpha_7 B_7. \end{aligned} \tag{A5}$$

If we apply the symmetry and recalculate the Hamiltonian, we get

$$\begin{aligned} \pi Q_1 \pi^{-1} = \alpha_3 B_1 + \alpha_2 B_2 + \alpha_1 B_3 + \alpha_7 B_4 \\ + \alpha_6 B_5 + \alpha_5 B_6 + \alpha_4 B_7. \end{aligned} \tag{A6}$$

As both Hamiltonians are equal, we can identify term by term and reduce the number of parameters. In this case, we see that $\alpha_1 = \alpha_3$, $\alpha_4 = \alpha_7$, and $\alpha_5 = \alpha_6$.

APPENDIX B: FORMULATION OF MAX-k-CUT ON QUDITS AND RESULTS FOR $k = 3$ SIMULATIONS

We want to maximize the number of vertices connecting k different regions in a graph. For the usual MAX-2-CUT, the elements of the problem Hamiltonian corresponding to each of the edges is

$$H_{i,j} = 4 L_{zi} L_{zj} \tag{B1}$$

in the usual $d = 2$ qudits description. The pool of CD terms at first order is

$$Q_1 = \{L_{yi} L_{zj}\}. \tag{B2}$$

For MAX-3-CUT,

$$H_{i,j} = L_{zi} L_{zj} - 2\left(L_{zi}^2 + L_{zj}^2\right) + 3 L_{zi}^2 L_{zj}^2, \tag{B3}$$

$$Q_1 = \left\{L_{y_i} L_{z\,j}, L_z L_y + L_y L_z, L_{zi} L_{y\,j} + L_{y\,j} L_{zi}, L_z^2\right\}, \tag{B4}$$

with $d = 3$ qudits. For MAX-4-CUT,

$$\begin{aligned} H_{i,j} = \frac{365}{72} L_{zi} L_{zj} - \frac{5}{8}\left(L_{zi}^2 + L_{zj}^2\right) + \frac{1}{2} L_{zi}^2 L_{zj}^2 \\ - \frac{41}{18}\left(L_{zi}^3 L_{zj} + L_{zi} L_{zj}^3\right) + \frac{10}{9} L_{zi}^3 L_{zj}^3 \end{aligned} \tag{B5}$$

for $d = 4$ qudits.

We see that all terms involved in these problem Hamiltonians contain each an even number of L_z operators such that the maximum degree of this operator in a single qudit is $k - 1$. Calculating these coefficients is as difficult as solving a linear system of $\sum_{d=1}^{k} \lceil d/2 \rceil - 1$ equations, which scales as $O(k^2)$.

In Fig. 7, we show the results for the experimentation done solving the MAX-3-CUT problem for six instances generated from nonplanar graphs with six vertices. We show the results in terms of the energy of the final states at each iteration. This is indeed the cost function we employed for the classical optimization loop. Additionally, we show the probability of measuring the optimal solution for each problem. We observe that even in the worst cases the success probability is at least one order of magnitude higher than random guessing. This behavior is consistent with the fact that we are using the energy as the cost function and not the fidelity.

APPENDIX C: DICKE STATES AND OPTIMIZATION OF FIDELITY FOR W STATE PREPARATION

Dicke states are important metrological states due to their connection to the theory of quantum communication and multipartite entanglement [59]. For a set of N qudits, a possible

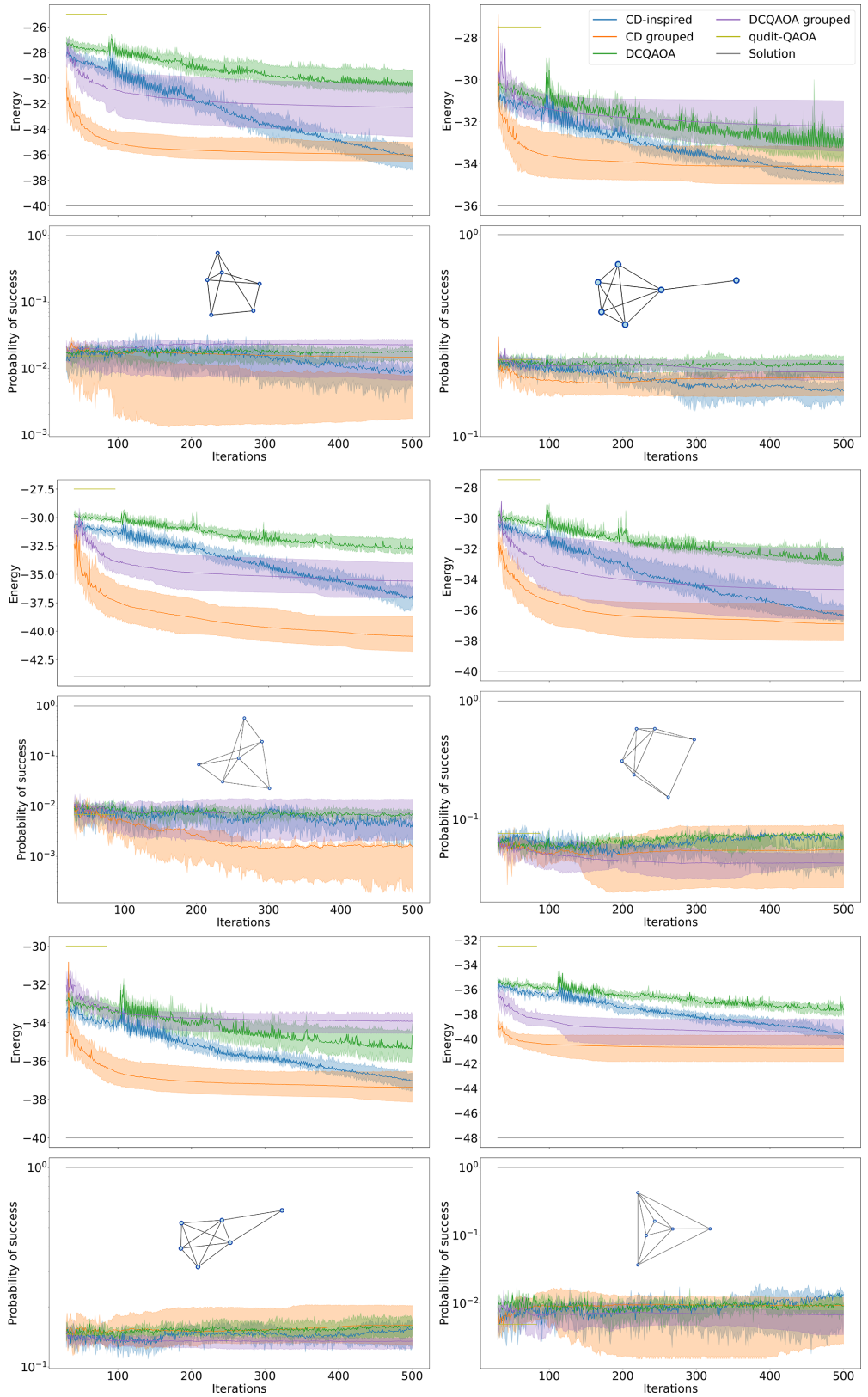


FIG. 7. Results for the different algorithms and instances for the Max-3-Cut. The solid lines correspond to mean values and the colored area corresponds to the interquartile range. The upper plots show the energy at each iteration step. The plots in the middle of each subfigure show the probability of measuring any of the states that correspond to the correct solution. The inset graphs correspond to each of the graphs that define the instances. We have run 20, 12, 12, 10, 11, and 9 runs, respectively, for each problem using uniformly random initial parameters.

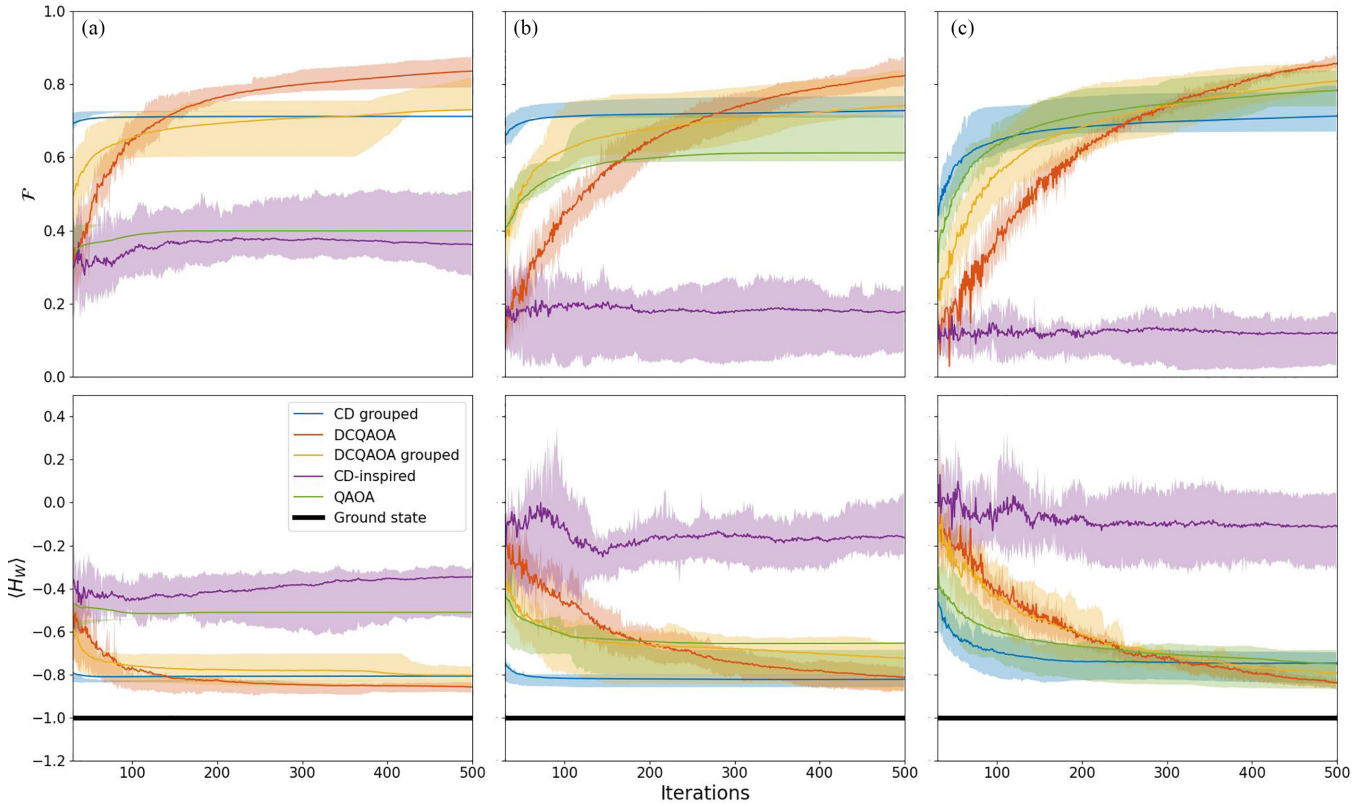


FIG. 8. Fidelity and the expectation value of the Hamiltonian of Eq. (23) (bottom) obtained by maximizing the fidelity, shown as a function of the optimizer iteration. Results for $N = 3$ and layers: $p = 2$ [panel (a)], $p = 3$ [panel (b)], and $p = 5$ [panel (c)]. The first 30 iterations are not shown.

choice of the Dicke state can be

$$D_\kappa^N = \binom{N}{\kappa}^{-\frac{1}{2}} \sum_i P_i(|d-1\rangle^{\otimes N-\kappa} |0\rangle^{\otimes \kappa}), \tag{C1}$$

where P_i indicates all the possible permutations of basis states. This is a ground state of the Hamiltonian

$$H_\kappa^N = \left(\frac{N}{2} - \kappa\right) \sum_{i=1}^{N} L_{zi} + \frac{1}{2} \sum_{j<i} L_{zi} L_{zj}. \tag{C2}$$

Considering a system of qutrits and $\kappa = 1$, from the previous Hamiltonian we obtain Hamiltonian (23). The corresponding Dicke state is

$$D_1^N = \frac{1}{\sqrt{N}} \sum_i P_i(|2\rangle^{\otimes N-1} |0\rangle), \tag{C3}$$

also known as the W states.

In the main text, we simulated variational algorithms by minimizing the expectation value of the energy in Eq. (23). Since our goal is to prepare a W state for three qudits, we try a direct approach by using the same variational ansätze but maximizing the fidelity in Eq. (24). A couple of considerations must be made in this procedure. This cost function can be represented as $\mathrm{Tr}[U(\boldsymbol{\theta}) |\psi_0\rangle \langle\psi_0| U(\boldsymbol{\theta})^\dagger |W\rangle \langle W|]$. Even though we cannot directly relate the Hamiltonian in Eq. (23), we can keep parallel with the original problem by considering that the W state of Eq. (22) is still a ground state of the original Hamiltonian and thus maximizing the fidelity will also minimize the energy. As shown in the results of Fig. 8, this algorithm obtains better results in approximating the W state ($\mathcal{F} \sim 85\%$ in the best case) and shows how CD terms have better performance than regular QAOA for few layers ($q \leqslant 4$).

[1] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[2] S. Bravyi, D. Gosset, R. König, and M. Tomamichel, Quantum advantage with noisy shallow circuits, Nat. Phys. **16**, 1040 (2020).

[3] M. E. Beverland, P. Murali, M. Troyer, K. M. Svore, T. Hoefler, V. Kliuchnikov, G. H. Low, M. Soeken, A. Sundaram, and A. Vaschillo, Assessing requirements to scale to practical quantum advantage, arXiv:2211.07629.

[4] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[5] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya,

M. Menotti, I. Dhand, Z. Vernon, N. Quesada, and J. Lavoie, Quantum computational advantage with a programmable photonic processor, Nature (London) **606**, 75 (2022).

[6] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, Nature (London) **618**, 500 (2023).

[7] H.-S. Zhong *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[8] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing STOC'96* (Association for Computing Machinery, New York, 1996), pp. 212–219.

[9] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comput. **26**, 1484 (1999).

[10] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[11] A. D. King, A. Nocera, M. M. Rams, J. Dziarmaga, R. Wiersema, W. Bernoudy, J. Raymond, N. Kaushal, N. Heinsdorf, R. Harris *et al.*, Beyond-classical computation in quantum simulation, Science **388**, 199 (2025).

[12] Y. Wang, Z. Hu, B. C. Sanders, and S. Kais, Qudits and high-dimensional quantum computing, Front. Phys. **8**, 589504 (2020).

[13] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, A universal qudit quantum processor with trapped ions, Nat. Phys. **18**, 1053 (2022).

[14] L. E. Fischer, A. Chiesa, F. Tacchino, D. J. Egger, S. Carretta, and I. Tavernelli, Universal qudit gate synthesis for transmons, PRX Quantum **4**, 030327 (2023).

[15] D. González-Cuadra, T. V. Zache, J. Carrasco, B. Kraus, and P. Zoller, Hardware efficient quantum simulation of non-Abelian gauge theories with qudits on Rydberg platforms, Phys. Rev. Lett. **129**, 160501 (2022).

[16] Y. Chi *et al.*, A programmable qudit-based quantum processor, Nat. Commun. **13**, 1166 (2022).

[17] X. Gao, P. Appel, N. Friis, M. Ringbauer, and M. Huber, On the role of entanglement in qudit-based circuit compression, Quantum **7**, 1141 (2023).

[18] S. Bravyi, O. Dial, J. M. Gambetta, D. Gil, and Z. Nazario, The future of quantum computing with superconducting qubits, J. Appl. Phys. **132**, 160902 (2022).

[19] A. Osman, J. Fernández-Pendás, C. Warren, S. Kosen, M. Scigliuzzo, A. Frisk Kockum, G. Tancredi, A. Fadavi Roudsari, and J. Bylander, Mitigation of frequency collisions in superconducting quantum processors, Phys. Rev. Res. **5**, 043001 (2023).

[20] S. Chandrasekharan and U. J. Wiese, Quantum link models: A discrete approach to gauge theories, Nucl. Phys. B **492**, 455 (1997).

[21] P. P. Popov, M. Meth, M. Lewestein, P. Hauke, M. Ringbauer, E. Zohar, and V. Kasper, Variational quantum simulation of U(1) lattice gauge theories with qudit systems, Phys. Rev. Res. **6**, 013202 (2024).

[22] Y. Deller, S. Schmitt, M. Lewenstein, S. Lenk, M. Federer, F. Jendrzejewski, P. Hauke, and V. Kasper, Quantum approximate optimization algorithm for qudit systems, Phys. Rev. A **107**, 062410 (2023).

[23] M. H. S. Amin, N. G. Dickson, and P. Smith, Adiabatic quantum optimization with qudits, Quantum Inf. Proccess. **12**, 1819 (2013).

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

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

[26] K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, A review on quantum approximate optimization algorithm and its variants, Phys. Rep. **1068**, 1 (2024).

[27] T. Volkoff and P. J. Coles, Large gradients via correlation in random parameterized quantum circuits, Quantum Sci. Technol. **6**, 025008 (2021).

[28] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, Quantum **3**, 214 (2019).

[29] A. Del Campo, Shortcuts to adiabaticity by counterdiabatic driving, Phys. Rev. Lett. **111**, 100502 (2013).

[30] X. Chen, I. Lizuain, A. Ruschhaupt, D. Guéry-Odelin, and J. G. Muga, Shortcut to adiabatic passage in two- and three-level atoms, Phys. Rev. Lett. **105**, 123003 (2010).

[31] J. Yao, L. Lin, and M. Bukov, Reinforcement learning for many-body ground-state preparation inspired by counterdiabatic driving, Phys. Rev. X **11**, 031070 (2021).

[32] F. Sauvage, M. Larocca, P. J. Coles, and C. Marco, Building spatial symmetries into parameterized quantum circuits for faster training, Quantum Sci. Technol. **9**, 015029 (2024).

[33] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, Phys. Rep. **986**, 1 (2022).

[34] Y. Suzuki, H. Yano, R. Raymond, and N. Yamamoto, Normalized gradient descent for variational quantum algorithms, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, USA, 2021), pp. 1–9.

[35] A. Kulshrestha, X. Liu, H. Ushijima-Mwesigwa, and I. Safro, Learning to optimize quantum neural networks without gradients, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, USA, 2023), Vol. 1, p. 263.

[36] A. Mandilara, B. Dellen, U. Jaekel, T. Valtinos, and D. Syvridis, Classification of data with a qudit, a geometric approach, Quantum Mach. Intell. **6**, 17 (2024).

[37] X. You, S. Chakrabarti, and X. Wu, A convergence theory for over-parameterized variational quantum eigensolvers, arXiv:2205.12481.

[38] M. Ragone, B. N. Bakalov, F. Sauvage, A. F. Kemper, C. Ortiz Marrero, M. Larocca, and M. Cerezo, A unified theory of barren plateaus for deep parametrized quantum circuits, Nat. Commun. **15**, 7172 (2024).

[39] D. Sels and A. Polkovnikov, Minimizing irreversible losses in quantum systems by local counterdiabatic driving, Proc. Natl. Acad. Sci. USA **114**, E3909 (2017).

[40] P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, Floquet-engineering counterdiabatic protocols in quantum many-body systems, Phys. Rev. Lett. **123**, 090602 (2019).

[41] D. Sun, P. Chandarana, Z.-H. Xin, and X. Chen, Optimizing counterdiabaticity by variational quantum circuits, Philos. Trans. R. Soc. A **380**, 20210282 (2022).

[42] A. Ferrer-Sánchez, C. Flores-Garrigos, C. Hernani-Morales, J. J Orquín-Marqués, N. N. Hegade, A. G. Cadavid, I. Montalban, E. Solano, Y. Vives-Gilabert, and J. D Martín-Guerrero, Physics-informed neural networks for an optimal counterdiabatic quantum computation, Mach. Learn.: Sci. Technol. **5**, 025035 (2024).

[43] K. Takahashi and A. del Campo, Shortcuts to adiabaticity in Krylov space, Phys. Rev. X **14**, 011032 (2024).

[44] P. Chandarana, N. N. Hegade, K. Paul, F. Albarrán-Arriagada, E. Solano, A. del Campo, and X. Chen, Digitized-counterdiabatic quantum approximate optimization algorithm, Phys. Rev. Res. **4**, 013141 (2022).

[45] P. Chandarana, N. N. Hegade, I. Montalban, E. Solano, and X. Chen, Digitized counterdiabatic quantum algorithm for protein folding, Phys. Rev. Appl. **20**, 014024 (2023).

[46] A. G. Cadavid, I. Montalban, A. Dalal, E. Solano, and N. N. Hegade, Efficient digitized counterdiabatic quantum optimization algorithm within the impulse regime for portfolio optimization, Phys. Rev. Appl. **22**, 054037 (2024).

[47] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, New York, 1996).

[48] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863 (1992).

[49] R. Zeier and T. Schulte-Herbrüggen, Symmetry principles in quantum systems theory, J. Math. Phys. **52**, 113510 (2011).

[50] K. Coolsaet, S. D'hondt, and J. Goedgebeur, House of graphs 2.0: A database of interesting graphs and more, Discrete Appl. Math. **325**, 97 (2023).

[51] J. Siemons, Automorphism groups of graphs, Arch. Math. **41**, 379 (1983).

[52] S. Foldes, Symmetries of directed graphs and the chinese remainder theorem, J. Comb. Theory B **28**, 18 (1980).

[53] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf, Strictly single-site DMRG algorithm with subspace expansion, Phys. Rev. B **91**, 155115 (2015).

[54] C. Hubig, Symmetry-protected tensor networks, Ph.D. thesis, Ludwig Maximilian University, 2017.

[55] C. H. Papadimitriou and M. Yannakakis, Optimization, approximation, and complexity classes, J. Comput. Syst. Sci. **43**, 425 (1991).

[56] F. Hadlock, Finding a maximum cut of a planar graph in polynomial time, SIAM J. Comput. **4**, 221 (1975).

[57] M. J. D. Powell, A direct search optimization method that models the objective and constraint functions by linear interpolation, in *Advances in Optimization and Numerical Analysis* (Springer, Dordrecht, 1994), p. 51.

[58] A. Frieze and M. Jerrum, Improved approximation algorithms for max*k*-cut and max bisection, Algorithmica **18**, 67 (1997).

[59] I. Bengtsson and K. Zyczkowski, A brief introduction to multipartite entanglement, arXiv:1612.07747.