

**Hamiltonian-oriented homotopy quantum approximate optimization algorithm**Akash Kundu <sup>1,2,\*</sup>, Ludmila Botelho <sup>1,2,\*†</sup> and Adam Glos <sup>1,3</sup><sup>1</sup>*Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, Gliwice, Poland*<sup>2</sup>*Joint Doctoral School, Silesian University of Technology, Akademicka 2A, Gliwice, Poland*<sup>3</sup>*Algorithmiq Ltd, Kanavakatu 3C 00160 Helsinki, Finland*

(Received 14 March 2023; revised 30 November 2023; accepted 10 January 2024; published 20 February 2024)

The classical homotopy optimization approach has the potential to deal with highly nonlinear landscapes, such as the energy landscape of quantum approximate optimization algorithm (QAOA) problems. Following this motivation, we introduce Hamiltonian-oriented homotopy QAOA (HOHo-QAOA), a heuristic method for combinatorial optimization using QAOA, based on classical homotopy optimization. The method consists of a homotopy map that produces an optimization problem for each value of the interpolating parameter. Therefore, HOHo-QAOA decomposes the optimization of QAOA into several loops, each using a mixture of the mixer and the objective Hamiltonian for cost function evaluation. Furthermore, we conclude that the HOHo-QAOA improves the search for low-energy states in the nonlinear energy landscape and outperforms other variants of QAOA.

DOI: [10.1103/PhysRevA.109.022611](https://doi.org/10.1103/PhysRevA.109.022611)**I. INTRODUCTION**

The speedup of practical applications is yet to be realized for quantum devices as they operate on a small number of qubits and the devices are noise prone. The limitations of the available hardware initiated the noisy intermediate scale quantum (NISQ) era [1]. The NISQ algorithms [2] can operate on a limited amount of resources, in particular, by distributing tasks between quantum and classical devices. Many of those algorithms are represented by a broad class of variational quantum algorithms (VQAs) [3]. Their generic structure consists of two subroutines: a parametric quantum circuit (PQC) is implemented on quantum hardware that generates a quantum state and classical hardware calculates the cost function and optimizes the parameters of PQC. One of the advantages of VQAs is that they can be easily adapted to various computational problems as long as the Hamiltonian can be designed whose ground state corresponds to the solution of the problem. To mention a few, VQAs have potential applications in finding the ground state of a molecule [4], solving linear [5] and nonlinear [6] systems of equations, quantum state-diagonalization [7], and quantum device certification [8]. A detailed review can be found in [3].

The quantum approximate optimization algorithm (QAOA) [9] is a variational quantum algorithm dedicated to combinatorial optimization problems. The PQC in QAOA is a trotterized adiabatic evolution, i.e., the circuit consists of interchangeably applied so-called mixer and problem Hamiltonians. It has a potential application in solving problems like graph coloring [10–12], MaxE3Lin2 [13], Max- $K$ -Vertex Cover [14], or the traveling

salesman problem [12,15]. To improve the performance of QAOA, multiple optimization strategies have been introduced [16–24]. This is because, given the limited resources of quantum computers, it is essential to effectively explore the landscape of cost function for PQC. However, the landscape of the energy function in QAOA is highly nonlinear and to deal with such complicated landscapes, sophisticated methods are necessary.

This motivates us to formulate a heuristic optimization strategy that uses classical homotopy optimization for QAOA. The homotopy optimization has potential applications in dealing with highly nonlinear functions [25]. The homotopy method comprises a homotopy map, which for each value of interpolating parameter  $\alpha \in [0, 1]$  outputs an optimization problem. In particular, for  $\alpha = 0$ , the problem is easy to solve, and for  $\alpha = 1$  the homotopy map returns the problem of interest. During the interpolation process, which changes the value of  $\alpha$  from 0 to 1, the solution continuously changes and is expected to be optimal, or close to optimal for the intermediate problems. If the intermediate optimization succeeds, in the end, we obtain the optimum of the target problem. One can see quantum annealing as a particular type of homotopy optimization. A homotopy optimization for VQE was already proposed in [26] and improved in [27,28]. However, its applicability for QAOA was only briefly mentioned in [29].

The introduced Hamiltonian-oriented homotopy QAOA (HOHo-QAOA), illustrated in Fig. 1, decomposes the optimization into several loops. The homotopy map smoothens between the mixer Hamiltonian and the problem Hamiltonian during the optimization and each loop uses the mixture of these two Hamiltonians for cost function evaluation. In each loop, the quantum state is optimized with respect to such intermediate cost functions. This strategy simplifies the search for good QAOA parameters while keeping the PQC unchanged. To show this, we investigate the weighted Max-Cut problem

\*These authors contributed equally to this work.

†Corresponding author: [lbotelho@iitis.pl](mailto:lbotelho@iitis.pl)

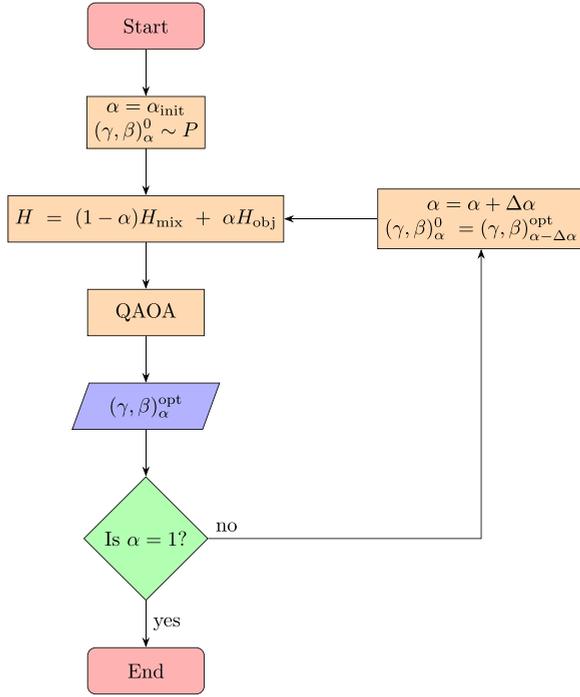


FIG. 1. Schematic representation of HOHo-QAOA. The algorithm starts with choosing an initial value of parameters  $(\gamma, \beta)$  according to some probability distribution  $P$ , and optimizing them for the initial Hamiltonian  $H$  for  $\alpha = \alpha_{\text{init}}$  with the chosen classical optimization procedure. Then, the optimal parameters for the ansatz are iteratively used as the initial parameters for the consecutive optimization routines for  $H$  with an increased value of  $\alpha$ , i.e.,  $\alpha = \alpha + \Delta\alpha$ . The procedure stops at  $H = H_{\text{obj}}$  for  $\alpha_{\text{init}} = 1$ , which is the objective Hamiltonian. Throughout the article,  $\alpha$  is referred to as a homotopy parameter.

on Barabási-Albert graphs. First, we empirically analyze the impact of the choice of the homotopy parameters: the initial  $\alpha_{\text{init}}$  value and the step parameter  $\alpha_{\text{step}}$  which defines the difference between two consecutive  $\alpha$  values. Although theoretically, a choice of  $\alpha_{\text{init}}$  and  $\alpha_{\text{step}}$  very close to zero provides a better approximation to the optimal solution, empirically we show that one can still get a good approximation to the optimal solution even if  $\alpha_{\text{init}}$  and  $\alpha_{\text{step}}$  are detached from zero. This hugely reduces the computational cost of HOHo-QAOA. Finally, we compare HOHo-QAOA with other commonly used QAOA optimization strategies [9,22].

The rest of the paper is organized in the following way. In Sec. II, we provide a brief overview of the adiabatic quantum computing, variants of QAOA, and the homotopy method. Throughout Sec. III, we numerically investigate the efficient settings of the homotopy parameters. Furthermore, we compare HOHo-QAOA with the other variants of QAOA considered in the literature. Finally, we conclude the article in Sec. V.

## II. PRELIMINARIES

### A. QAOA

The core concept of adiabatic quantum computing (AQC) lies in the adiabatic theorem. Let us consider  $H(s) = H(t/T)$ ,

a time-dependent smoothly varying Hamiltonian for all  $t \in [0, T]$ , i.e.,  $s \in [0, 1]$ , where  $T$  is the total time of evolution. Let us denote by  $|E_i(s)\rangle$  an eigenvector of  $H(s)$  with corresponding eigenvalue  $E_i(s)$ , where we assume  $E_0(s) \leq E_1(s) \leq \dots$ . The adiabatic theorem roughly states that a system that is initially prepared in  $|E_0(0)\rangle$  of  $H(t=0)$ , after time evolution that is piloted by the Schrödinger equation with the given Hamiltonian  $H(s)$ , will approximately keep the state of the system in the  $|E_0(1)\rangle$  at  $t = T$ , provided that the change in  $H(s)$  is “sufficiently slow.” Traditionally, the sufficiently slow change is given by the condition [30,31]

$$T \gg \Delta^{-2} \max_{s \in [0,1]} \left\| \left[ \frac{\partial H(s)}{\partial t} \right]^2 \right\|, \quad (1)$$

where  $\Delta = \min_s [E_1(s) - E_0(s)]$  is the spectral gap. A class of independent conditions on  $T$  was discussed in [32–35]. AQC has the potential to take an initial Hamiltonian, say  $H_{\text{mix}}$ , whose ground state is easy to prepare to the ground state of a computationally hard problem Hamiltonian  $H_{\text{obj}}$ . A particular time-dependent Hamiltonian interpolates between the  $H_{\text{mix}}$  and  $H_{\text{obj}}$  as

$$H(s) = (1 - s)H_{\text{mix}} + sH_{\text{obj}}, \quad (2)$$

AQC in the form of quantum annealing has been used for a variety of applications, including real-world problems [36–41] and in quantum chemistry [42]. For a rigorous review of AQC, check [31,43].

The QAOA uses the first-order Suzuki-Trotter transformation of  $\exp[-iH(s)]$  as the variational ansatz to solve combinatorial optimization problems. The trotterization gives rise to the operators  $\exp(-i\gamma_j H_{\text{obj}})$  and  $\exp(-i\beta_j H_{\text{mix}})$ , where  $\gamma_j$  is the parameter corresponding to the objective Hamiltonian and  $\beta_j$  corresponds to the mixer Hamiltonian for the  $j$ th step. The mixer Hamiltonian is traditionally expressed as  $H_{\text{mix}} = -\sum_i X_i$ , where  $X_i$  is the Pauli  $X$  operator acting on the  $i$ th qubit and  $H_{\text{obj}}$  is the objective Hamiltonian whose ground state encodes the optimal solution of the problem. This results in state

$$|\vec{\gamma}, \vec{\beta}\rangle = \prod_{j=1}^L \exp(-i\beta_j H_{\text{mix}}) \exp(-i\gamma_j H_{\text{obj}}) |+\rangle^{\otimes N}, \quad (3)$$

where  $N$  is the number of qubits,  $L$  is the number of layers that defines the number of repeated applications of the mixer and objective Hamiltonian, and  $|+\rangle^{\otimes N}$  is the ground state of  $-\sum_i X_i$ . The algorithm utilizes quantum hardware to evaluate the energy expectation value  $E(\vec{\gamma}, \vec{\beta}) = \langle \vec{\gamma}, \vec{\beta} | H_{\text{obj}} | \vec{\gamma}, \vec{\beta} \rangle$ . Then the parameters  $\vec{\gamma}$  and  $\vec{\beta}$  are optimized using classical optimization methods so that the energy is minimized. This energy evaluation along with classical optimization QAOA is well defined for any combinatorial optimization problems as long as  $H_{\text{obj}}$  can be implemented efficiently. While the proposed  $X$  mixer combined with the two-local Ising model is frequently used in the literature, different choices were also considered [12,15,44–46].

Heuristic learning of QAOA was explored in trajectories QAOA (T-QAOA) [22]. T-QAOA is a heuristic strategy that utilizes the interpolation-based prediction of good QAOA parameters. With the random initialization, the cost for op-

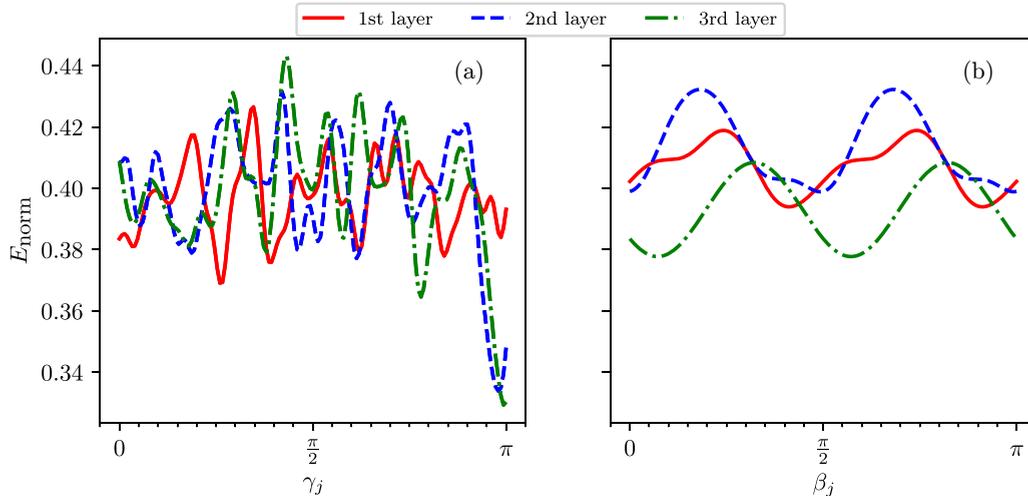


FIG. 2. Illustration of the highly nonlinear energy landscape of QAOA for Max-Cut for ten nodes with weighted Barabási-Albert graph for (a) objective Hamiltonian and (b) mixer Hamiltonian.  $E_{\text{norm}}$  is a normalized energy of the objective Hamiltonian, described in Eq. (11), so that the eigenvalues lie in  $[0,1]$ .

timization of QAOA is exponential in the number of layers of QAOA [22]. However, with the increased number of layers,  $H_{\text{mix}}$  may gradually turn off while the  $H_{\text{obj}}$  turns on, which is reminiscent of AQC. However, QAOA could learn via following a diabatic path to achieve higher success probability [31,47,48], which is beyond the adiabatic process natural for AQC. This fact was used in T-QAOA by reusing the optimal angles found for  $L$  layers in the  $(L + 1)$ -layers PQC.

The T-QAOA variant considered in this paper runs as follows. It starts with a number of layer  $L_0$  and finds the locally optimal parameters  $(\vec{\gamma}^{L_0}, \vec{\beta}^{L_0})$ . Then it uses the optimal parameters of layer  $L_0$  to construct the initial parameters for the layer  $L_0 + 1$  by sampling the last entries of  $\vec{\gamma}^{L_0+1}$  from a uniform random distribution and setting  $\vec{\beta}^{L_0+1} = 0$ . With such initialization, the  $(L_0 + 1)$ -th layer PQC is optimized and the procedure is repeated until a final number of layer  $L$  is reached. Note that different interpolation methods can be used [22].

Note that, for QAOA, the energy landscape with respect to a single parameter  $\theta$  is related to the following process. First, an initial quantum state is prepared. Then, if applicable, all the unitary operations that precede the  $\theta$ -dependent operation are applied, which transforms the initial state into a different state (possibly a mixed state for noisy evolution). Afterwards, under an assumption of pure evolution, a unitary  $\exp(-i\theta H)$  for the mixer or objective Hamiltonian  $H$  is applied. Finally, the remaining operations are applied and the energy estimation with respect to the observable is conducted. As shown in Appendix B, the energy function with respect to  $\theta$  takes the form

$$C + \sum_{i>j} A_{i,j} \cos[\theta(E_i - E_j) + B_{i,j}], \quad (4)$$

in which  $\{E_i\}$  is the set of all eigenvalues of the operator  $H$  and real parameters  $C, A_{i,j}, B_{i,j}$  depend on the initial state, observable, and  $\theta$ -independent quantum operations. Note that Eq. (4) is highly nonlinear, therefore, its optimization may be difficult in practice. This is in contrast to the typically used

VQE approaches in which the parameter-dependent unitary can be reduced to a single-qubit gate, which, in turn, may result in a simple, yet powerful gradient-free optimization technique [49,50].

Unfortunately, the number of cosines in Eq. (4) may grow quadratically with the number of distinct eigenvalues of the considered Hamiltonian. In the case of the objective Hamiltonian, the number may be particularly high. While for many simple problems like unweighted Max-Cut or Max-SAT the number of different eigenvalues usually grows polynomially with the size of the data, for weighted Max-Cut each partition may result in a different objective value, which may give  $O(2^n)$  different energies in general. A complicated energy landscape can be seen already even for a small and simple instance, see Fig. 2. For problems generating such complicated landscapes, more sophisticated methods may be at hand.

### B. Homotopy optimization method

One of the well-known methods to solve a system of highly nonlinear problems is *homotopy optimization*, where a homotopy map is constructed between two systems. The solution corresponding to one of the systems is transformed into the solution of the other system. For example, consider the function  $f_{\text{targ}}(x)$  which encodes a computationally hard problem, and  $f_{\text{init}}(x)$  which is a problem with an easy-to-find solution. Then the particular homotopy map between the systems can be given as

$$\mathcal{F}(\alpha, x) = g_1(\alpha)f_{\text{targ}}(x) + g_2(\alpha)f_{\text{init}}(x), \quad 0 \leq \alpha \leq 1, \quad (5)$$

where

$$\begin{aligned} g_1(0) &= 0, & g_2(0) &= 1, \\ g_1(1) &= 1, & g_2(1) &= 0. \end{aligned} \quad (6)$$

Here, we get a family of problems corresponding to  $\min_x \mathcal{F}(\alpha, x) = 0$  for each  $\alpha$  value from 0 to 1. We track the optimized solutions starting from  $(\alpha, x) = (0, x_0)$ , as  $\alpha$  moves from 0 to 1, which, for a successful homotopy map, leads to

$(\alpha, x) = (1, x_1)$ , where  $x_1$  is ideally the optimal solution of  $f_{\text{targ}}$ .

The state-of-the-art approach is to start from  $(\alpha_{\text{init}}, x_{\text{init}})$  with  $x_{\text{init}}$  minimizing  $\mathcal{F}(0, x) = f_{\text{init}}(x)$ . Then the problem  $\min_x \mathcal{F}(\alpha + \alpha_{\text{step}}, x) = 0$  is iteratively solved using the solution of  $\min_x \mathcal{F}(\alpha, x)$  as a starting point for sufficiently small  $\alpha_{\text{step}} > 0$  [25].

### III. METHODS

#### A. Hamiltonian-oriented Homotopy QAOA

The Hamiltonian-oriented homotopy QAOA decomposes the optimization process of the objective Hamiltonian into several optimization loops. Each loop optimizes the energy

$$E_\alpha(\vec{\gamma}, \vec{\beta}) = \langle \vec{\gamma}, \vec{\beta} | H(\alpha) | \vec{\gamma}, \vec{\beta} \rangle, \quad (7)$$

where  $H(\alpha)$  encodes the homotopy map

$$H(\alpha) = g_1(\alpha)H_{\text{mix}} + g_2(\alpha)H_{\text{obj}}, \quad 0 \leq \alpha \leq 1. \quad (8)$$

For  $\alpha = 1$  the expectation value in Eq. (7) is the energy corresponding to the  $H_{\text{obj}}$ . There is a freedom in the choice of  $g_1$  and  $g_2$  with the necessary conditions  $H(0) = H_{\text{mix}}$  and  $H(1) = H_{\text{obj}}$ . The simplest and most frequent formulation of  $g_1$  and  $g_2$  is [25,51]

$$g_1(\alpha) = 1 - \alpha, \quad g_2(\alpha) = \alpha. \quad (9)$$

During the optimization process, we choose an initialization of mixer and objective parameters (at  $\alpha = 0$ ) in such a way that the parameters corresponding to the mixer are sampled from the uniform random distribution  $U(a, b)$  in an interval  $[a = 0, b = 2\pi]$  and the objective parameters are all set to 0. With this initialization, we make sure that the homotopy starts from the exact ground state of the mixer on a noise-free setting, as the application of the mixer on its eigenstate does not change the state. For  $\alpha' > \alpha \geq 0$  the initial parameters are chosen as

$$(\vec{\gamma}, \vec{\beta})_{\alpha'}^{\text{init}} = (\vec{\gamma}, \vec{\beta})_\alpha^*, \quad (10)$$

here  $*$  denotes the optimal parameters for  $\alpha$ .

Note that, by the very definition of the homotopy process, the initial quantum state should be a ground energy of the Hamiltonian  $H(0) = H_{\text{mix}}$ . This suggests that not all generalizations of QAOA are a suitable choice for this kind of problem, as, for example, in [44] the authors started with an arbitrary feasible state. However, one can still find an example of quantum states and mixers that could be used here, like the  $XY$  mixer [46] defined for one-hot vectors or even the Grover mixer for a general initial state [45]. In addition, it is required that estimating the energy of the  $H_{\text{mix}}$  can be done efficiently, which is true for both referenced models [45,46].

It should be noted that each run of HOHo-QAOA follows the generic structure of the homotopy process as in Eq. (8) where the ‘‘run-time’’ of HOHo-QAOA is characterized by the  $\alpha_{\text{step}}$ , for a fixed  $\alpha_{\text{init}}$ . The parameter  $\alpha_{\text{init}}$  fixes the initial  $\alpha$  value. Generally, it can be inferred that a better approximation to the optimal solution can be achieved if we choose a sufficiently small value of  $\alpha_{\text{step}}$  and  $\alpha_{\text{init}}$ . They can be described in a more elaborate way as follows. The small value of  $\alpha_{\text{step}}$  helps

us realize the homotopy of Eq. (8) and at the same time if we initiate with  $\alpha_{\text{init}} \rightarrow 0$ , it becomes easier to find the ground state for the first step. To show this, throughout the paper, we investigate the normalized energy

$$E_{\text{norm}}[E_\alpha(\vec{\gamma}, \vec{\beta}), \alpha] = \frac{E_\alpha(\vec{\gamma}, \vec{\beta}) - \min H(\alpha)}{\max H(\alpha) - \min H(\alpha)}, \quad (11)$$

with respect to the parameters of HOHo-QAOA, where  $E_{\text{norm}}(\alpha) = 0$  is the normalized ground energy for any  $\alpha \in [0, 1]$ , and  $\min H(\alpha)$  ( $\max H(\alpha)$ ) denotes the minimum (maximum) of  $H(\alpha)$ . It should be noted that, for each value of  $\alpha$ , the homotopy method generates a new Hamiltonian and we get a new optimal energy. Hence, without normalization, we would not be able to see how close to the ground energy of the intermediate Hamiltonian we are. In addition, normalizing the energy allows a fair comparison of the values obtained for different problem instances.

#### B. HOHo-QAOA initialization strategy

In the following, first, we numerically discuss proposed settings for initial QAOA parameters  $(\vec{\gamma}, \vec{\beta})^{\text{init}}$ . With this setting, we show that the homotopy parameters, i.e.,  $\alpha_{\text{init}}$ ,  $\alpha_{\text{step}}$  can be chosen detached from zero without compromising the efficiency of the method. We consider an optimized energy  $E_{\text{norm}}^*$ , or in the case of HOHo-QAOA also an intermediate optimized step energy  $E_{\text{norm}}^*(\alpha)$ . In the numerical results the  $E_{\text{norm}}^*$  is averaged over 100 noiseless simulations. Each noiseless simulation of HOHo-QAOA is uniquely characterized by the random graph  $G = (V, E)$ , which is chosen from the Barabási-Albert distribution with 6, 8,  $\dots$ , 18 nodes and with  $m = 2$ , where  $m$  defines the number of edges to be attached from a new node to existing ones. The weights corresponding to the edges are picked up from a uniform set of integer weights  $w_{jj'} \in \{1, \dots, 10\}$  for each edge  $\{j, j'\}$ .

For generating the objective Hamiltonian, we started by generating graph objects which we later converted to Pauli operators objects and Hamiltonian matrices with QISKIT [52]. We generated 100 graphs for each number of nodes. Then we sampled the initial optimization parameters  $\vec{\gamma}, \vec{\beta}$  in one of the ways introduced in the next paragraph. We took an exact expectation energy and gradient of the state during the optimization. We choose the L-BFGS [53] algorithm implemented in Julia’s OPTIM package as a subroutine. The optimization has no periodic or bound conditions. We set OPTIM with absolute tolerance, relative tolerance, and absolute tolerance in a gradient equal to  $10^{-9}$ . We allow steps that increase the objective value and the maximum number of iterations is 10 000.

For the numerical investigation of optimal QAOA parameters, which is illustrated in Fig. 3, we consider three possible initialization choices of the mixer and objective parameters at  $\alpha = \alpha_{\text{init}}$  as follows.

(1) Random random (RR): When the parameters corresponding to the mixer and objective Hamiltonians are chosen from a uniform random distribution  $U(0, 2\pi)$ , i.e.,  $\gamma_j^{\text{init}} \sim U(0, 2\pi)$ ,  $\beta_j^{\text{init}} \sim U(0, 2\pi)$ .

(2) Near-zero random (NZR): The parameters corresponding to the mixer Hamiltonian are chosen from  $U(0, 2\pi)$ , but

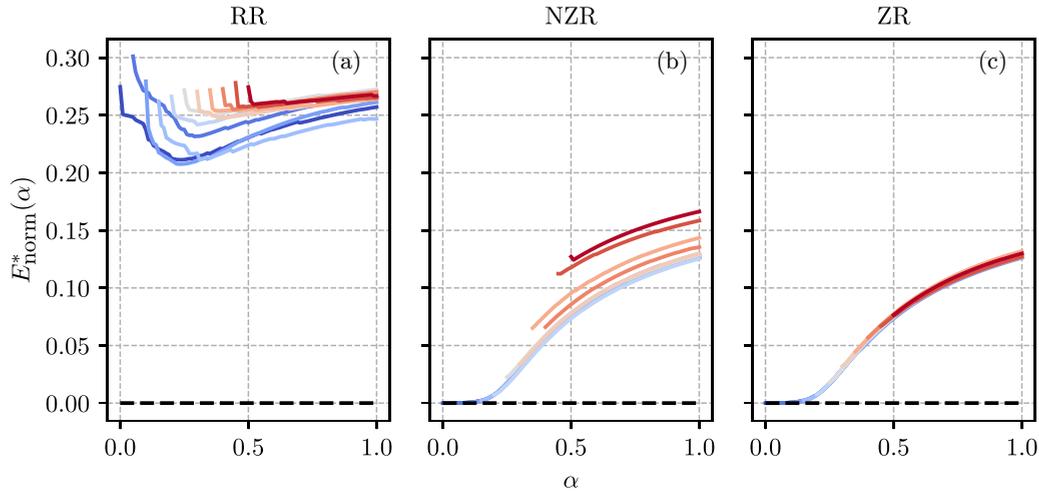


FIG. 3. The impact of different methods of initialization of  $\gamma_j, \beta_j$  on HOHo-QAOA for (a) RR (Random Random), (b) NZR (Near-Zero Random) with parameter  $v = 0.05$ , and (c) ZR (Zero Random) initialization, respectively, see Sec. III B for details. It is visible that the ZR outperforms the other two initializations. The  $\alpha_{\text{init}}$  is chosen between 0.0 (blue) to 0.5 (red) in an interval of 0.05 and the blue to red color scheme, and the moment lines start w.r.t. the X axis, represents the different values of  $\alpha_{\text{init}}$ . The performance of NZR and ZR are comparable for  $\alpha_{\text{init}} \leq 0.3$ , but for  $\alpha_{\text{init}} > 0.3$ , the minima for NZR scatters in the region  $0.05 < E_{\text{norm}} < 0.15$  whereas the minima for ZR clusters in a very narrow  $E_{\text{norm}}$  width. For the purpose of this work, we chose  $v = 0.05$ , which turned out to be a good transition case from the RR to the ZR initialization.

the objective parameters are sampled from the values very close to zero, i.e.,  $\gamma_j^{\text{init}} \sim U(0, v)$ ,  $\beta_j^{\text{init}} \sim U(0, 2\pi)$ ,

(3) Zero random (ZR): Mixer parameters are sampled from  $U(0, 2\pi)$  chosen and the objective is all zeros, i.e.,  $\gamma_j^{\text{init}} = 0$ ,  $\beta_j^{\text{init}} \sim U(0, 2\pi)$  as proposed before.

From Fig. 3 we conclude that ZR gives the best approximation to the ground state. This is because, under the ZR setting, the initial parameters of QAOA always start corresponding to the exact ground state of  $H_{\text{mix}}$  while the  $H_{\text{obj}}$  is turned off. This is within the spirit of the homotopy optimization, in which starting with the optimal solution of the initial system is critical. Hence this good approximation to the initial parameters leads us to a better solution to the ground state of  $H_{\text{obj}}$ . Keeping in mind that if we sample  $\alpha_{\text{init}}$  in the range  $0 \leq \alpha_{\text{init}} \leq 0.2$ , we see that NZR shows comparable performance to ZR and the choice of initialization of  $\gamma_j, \beta_j$  can be either one of them, relaxing the conditions on the choice of  $\vec{\gamma}$  and  $\vec{\beta}$ . In the remainder of this paper, all the numerical results are initialized with the ZR setting.

Now we move to the analysis of the choice of suitable  $\alpha_{\text{init}}$ . In Fig. 4 we investigate the  $\alpha_{\text{init}}$  dependency of the  $E_{\text{norm}}^*$ , where the energy is averaged over 100 noiseless simulations. From Fig. 4(a) we take three layers of HOHo-QAOA and observe that the mean optimal energy and the corresponding standard deviation remain unchanged (which we term as the *region of stability*) with respect to  $\alpha_{\text{init}}$  in the range  $0 \leq \alpha_{\text{init}} \leq 0.5$ . With an increase in the number of nodes from 6 to 16, the *region of stability* shifts upwards but remains in the range  $0 \leq \alpha_{\text{init}} \leq 0.5$ . This observation leads us to conclude that  $\alpha_{\text{init}}$  can be chosen detached from zero without degrading the performance of HOHo-QAOA, or that at least the region of stability does not shrink rapidly with the increased size of the problem. So setting  $\alpha_{\text{init}}$  in the region of stability along with  $\gamma_j^{\text{init}} = 0, \beta_j^{\text{init}} \sim U(0, 2\pi)$  yields a solution with particularly small energy value.

In Fig. 4(b), we investigate how the efficiency of the optimization depends on the  $\alpha_{\text{step}}$ . During this investigation, we take ten layers of HOHo-QAOA. We observe that, in the range  $10^{-4} \leq \alpha_{\text{step}} < 0.5$ , the approximation to the ground energy and the corresponding standard deviation with increasing  $\alpha_{\text{step}} \rightarrow 0$  remains almost unchanged, giving rise to a region of stability concerning  $\alpha_{\text{step}}$ . Several factors could contribute to the observed stability, such as the specific problem instances or the optimization algorithm utilized (in our paper, L-BFGS). Moreover, it is possible that, within the specified number of layers, the HOHo-QAOA method attains the optimal achievable energy for the ansatz. Consequently, altering the  $\alpha_{\text{step}}$  may only marginally enhance the optimized states. Still, the behavior of  $E_{\text{norm}}^*$  with  $\alpha_{\text{step}}$  is similar to what we observe for  $\alpha_{\text{init}}$ . This leads us to a conclusion that one can choose  $\alpha_{\text{step}}$  detached from zero for HOHo-QAOA.

Meanwhile, we can also observe a sharp increase in the normalized energy in Fig. 4(b) when the  $\alpha_{\text{step}}$  changes from 0.5 to 1. This is because at  $\alpha_{\text{init}} = 0$  no optimization is done as the ansatz starting in ZR initialization is already in the ground state of the mixer, and then for  $\alpha = \alpha_{\text{init}} + \alpha_{\text{step}} = 1$  the Hamiltonian  $H(\alpha)$  is just an objective function. However, the optimized expectation value at  $\alpha_{\text{step}} = 0.5$  suddenly decreases. This confirms that, even for such large  $\alpha_{\text{step}}$ , when we conduct just a single optimization of homotopy parameters, our algorithm gives a clear benefit.

It should be noted that, due to the high simulation cost for 16, qubits are halted at the  $\alpha_{\text{step}} = 10^{-2}$ , whereas the investigation for six and ten qubits is extended to  $10^{-4}$ .

The discussion and numerical results from the previous paragraphs give us the following initialization rules of HOHo-QAOA, which leads to a high efficiency of the method.

(1) The parameters of the mixer and objective should be initialized with the ZR setting, i.e.,  $\gamma_j^{\text{init}} = 0, \beta_j^{\text{init}} \sim U(0, 2\pi)$ .

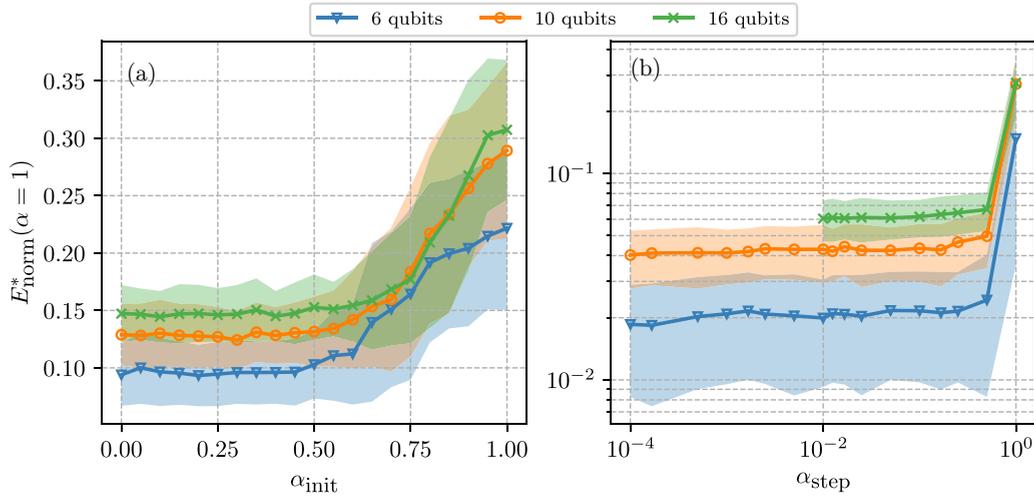


FIG. 4. We illustrate the dependency of  $E_{\text{norm}}^*$  with  $\alpha_{\text{init}}$  and  $\alpha_{\text{step}}$ . In (a) the variation of  $E_{\text{norm}}^*$  with  $\alpha_{\text{init}}$  for three layers of HOHo-QAOA is presented, with  $\gamma_j^{\text{init}} = 0$ ,  $\beta_j^{\text{init}} \sim U(0, 2\pi)$ . In the figure, we see a *region of stability* of HOHo-QAOA with respect to  $\alpha_{\text{init}}$  in the range 0.0 to 0.50. In (b) we present  $E_{\text{norm}}^*$  versus  $\alpha_{\text{step}}$  using ten layers of HOHo-QAOA. Just like in the case of  $\alpha_{\text{init}}$ , for  $\alpha_{\text{step}}$  a similar *region of stability* can be observed. This gives us the preference on the choice of *step parameter* while utilizing HOHo-QAOA. It should be noted that the y axis in (a) is in linear scale whereas in (b) it is in log scale. The lines in both the plots are taken  $\alpha_{\text{init}}$  and  $\alpha_{\text{step}}$  wise and are the mean of 100 noiseless simulations. The areas under the plots are the standard deviation of energies.

(2) Although one can infer that  $\alpha_{\text{init}} \rightarrow 0$  along with  $\alpha_{\text{step}} \rightarrow 0$  gives the best result, our investigations show that one can choose the homotopy parameters detached from zero. This greatly reduces the cost of simulating HOHo-QAOA.

#### IV. RESULTS

In this section, we analyze the performance of the introduced algorithm with respect to other optimization strategies introduced above. While it is natural for HOHo-QAOA to initialize using ZR strategy, it is unclear whether this choice will improve or worsen the results for QAOA or T-QAOA. Therefore, before comparing state-of-the-art methods to the introduced one, we verify whether there is any difference in

the performance of QAOA and T-QAOA with respect to the initialization of the optimized angles. In Fig. 5 we investigate state-of-the-art methods for parameters  $(\gamma_j, \beta_j)^{\text{init}}$  initialized with RR and ZR strategy. We observe that the performance of QAOA and T-QAOA is not influenced by the chosen strategies. This justifies using the ZR strategy when comparing QAOA, T-QAOA, and HOHo-QAOA.

Note that for QAOA we are observing undesired nonmonotonic behavior with respect to the number of layers. We claim that this is caused because of a complicated landscape of the energy function, which makes it difficult to optimize it if no information about the problem instance is used during the initialization from a large number of nodes. This argument complies with the good performance of T-QAOA where the

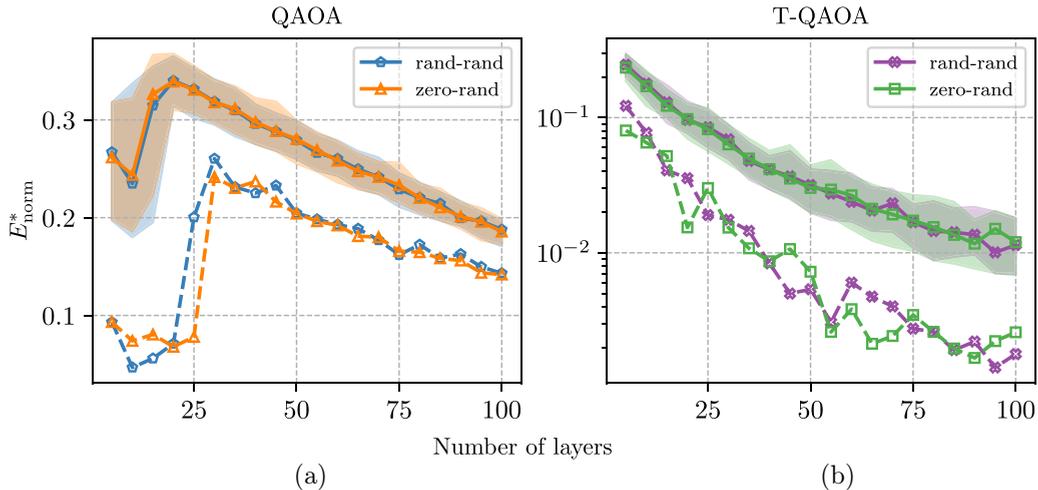


FIG. 5. Comparison of different initialization for QAOA and T-QAOA. In (a) we illustrate how the  $E_{\text{norm}}^*$  changes with increasing number of layers in QAOA under the RR and ZR settings whereas in (b) we conduct the experiment with similar settings for T-QAOA. The solid line is the median energy over 100 noiseless simulations, meanwhile, the dashed line represents the best case, taken layer-wise and node-wise by choosing the minimum energy among all the results. The areas are delimited by the first and third quartiles.

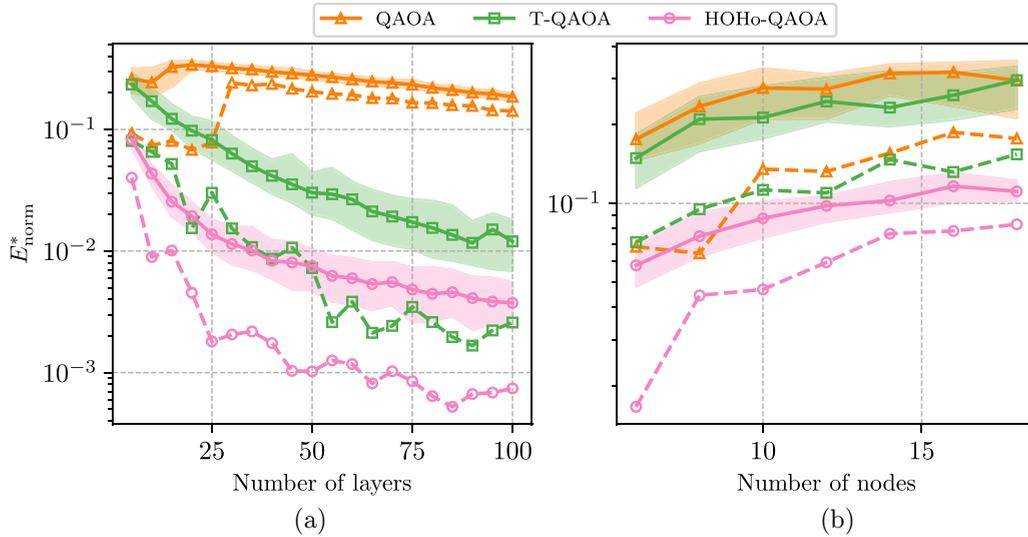


FIG. 6. Performance of HOHo-QAOA compared to QAOA and T-QAOA. In both figures, for all the QAOA methods, we applied the ZR settings. The areas are delimited by the first and third quartiles. The solid line presents  $E_{\text{norm}}^*$  median over 100 noiseless simulations for the left figure and 50 for the right figure, and the dashed line represents the best case, taken layer-wise and node-wise by choosing the minimum energy among all the simulations. In (a), the number of nodes is fixed to 10. Meanwhile in (b), the number of layers is fixed to 5 and the Max-Cut instances are chosen within 6 to 18 nodes. The homotopy parameters are set as  $\alpha_{\text{init}} = 0$  and  $\alpha_{\text{step}} = 0.01$ . One can see that in both cases the averaged energy as well as the best case of HOHo-QAOA outperforms the other variants of QAOA.

initial parameters of the  $(L + 1)$ -layer step are evaluated based on local optimal solutions of the  $L$ -layers step.

In Fig. 6 we compare the performance of HOHo-QAOA with the other variants when  $(\gamma_j, \beta_j)^{\text{init}}$  are initialized using the ZR setting. In the first simulation, we run the algorithms with a fixed number of nodes while increasing the number of layers. In the second, the number of layers is fixed while we vary the number of nodes. The plots present optimized energy values, averaged respectively over 100 and 50 instances. The data show that the introduced HOHo-QAOA gives us significantly smaller energy in both setups. Good improvements remain as more layers of the HOHo-QAOA are used and also outperform the other variants of QAOA for a higher number of nodes. These conclusions remain valid also for the best case chosen (dashed line). It should be noted that the HOHo-QAOA outperforms QAOA and the T-QAOA in each and every layer starting from the initial layer 5 to the final layer 100.

## V. CONCLUSION

In the article, we present a combination of homotopy optimization with an application in QAOA suitable for combinatorial optimization. In our method, the observable used for computing the energy is changed during the optimization process. The process starts with the observable being a mixer, for which the initial state of QAOA is a ground state, and is slowly moved into the objective Hamiltonian. In addition, we verify that, although traditionally in the homotopy method, the initial value of the transition parameter  $\alpha$  should be 0 and the step should be as small as possible, for QAOA the value of the considered parameters can be detached from 0. Our investigation of different initializations of HOHo-QAOA helped us to conclude that the zero random (ZR) initialization is the

optimal choice for the weighed Max-Cut problem. However, the optimal choice of the hyperparameters, or whether we can indeed detach  $\alpha_{\text{init}}$  from 0, may depend on the particular problem at hand and the size of its instances.

A homotopy optimization is an algorithm dedicated to nonlinear optimized functions, and since even a simple QAOA landscape is a linear combination of many (for some problems exponentially many) sinusoidal functions, our approach is well motivated for such energy function. This is in contrast to the typical VQE optimization process in which the function landscape with respect to a single parameter is just a sine. By comparing our approach and QAOA algorithm with the typical choice of optimization strategies we numerically confirm that our method outperforms state-of-the-art approaches.

While our algorithm was only presented for QUBO and the  $X$  mixer, it is not restricted to it. In particular, if the transition function is of the form  $H(\alpha) = g_1(\alpha)H_{\text{mix}} + g_2(\alpha)H_{\text{obj}}$ , we only require the energy of the  $H_{\text{mixer}}$  to be efficiently computable and the ground state to be easily prepared. Both the  $XY$  mixer [46] and the Grover mixer [45] satisfy these conditions. Moreover, our approach remains also valid for higher-order binary problems [11,15,54] and more advanced pseudo-code-based QAOA Hamiltonian implementation [12].

The introduced HOHo-QAOA uses the same quantum circuit as the standard QAOA, yet it allows reaching quantum states with much lower expectation values. Hence, using this method does not increase the impact of the noise. Compared to T-QAOA, where new layers are added one by one, our algorithm is from the very beginning working on the full circuit. One could expect therefore that HOHo-QAOA overshoots with the number of QAOA layers, as it is not chosen adaptively as for T-QAOA. However, as we observed for the given number of layers, HOHo-QAOA explores the ansatz much better than T-QAOA. Hence, repeating HOHo-QAOA from scratch with a gradually increasing number of layers may

lead in fact to shorter (and thus more noise-robust) quantum circuits compared to the T-QAOA method.

One could expect that the HOHo-QAOA algorithm will be the slowest method of all the considered ones because of the classical optimization being executed for all intermediate  $\alpha$ . However, if the functions  $g_1, g_2$  which defines the combination of mixer and objective Hamiltonians are not varying extensively, one should expect that the ground states for most of the intermediate Hamiltonians should be close to neighboring ones. Therefore, the number of steps to be taken by the classical optimizer with each  $\alpha$  change is expected to be much smaller compared to adding a new layer as in T-QAOA with randomly chosen QAOA parameters. Whether this phenomenon will make up for the fact that the total optimization time will be comparable to the time required for T-QAOA would require investigating larger instances.

Data and code available at [55].

#### ACKNOWLEDGMENTS

A.K., A.G., and L.B. were partially supported by the Polish National Science Center under Grant Agreement No. 2019/33/B/ST6/02011. A.G. acknowledges support from

the Polish National Science Center under Grant Agreement No. 2020/37/N/ST6/02220. The authors would like to thank Zoltán Zimborás, Özlem Salehi, and Jarosław A. Miszcza for valuable discussions and comments on the manuscript.

#### APPENDIX A: CODE FOR NOISELESS SIMULATION

To enable the simple reproduction of our results, we publish our code in [55]. The algorithms for generating data and plotting were implemented in JULIA and PYTHON programming languages. Versions of the software and additional packages are listed in [56].

#### T-QAOA

For T-QAOA implementation, we initialize with a minimum number of levels  $L_0 = 4$  and run the optimization similarly to the state of art QAOA with the given parameters' initiation strategy. The method proceeds to check the convergence of the solution and moves to the next layer  $L_0 + 1$ , using the previously optimized parameters with the addition of a zero for the mixer Hamiltonian and a value sampled from a uniform random distribution  $U(0, 2\pi)$ .

#### APPENDIX B: PROOF OF NONLINEAR LANDSCAPE FOR QAOA

*Theorem 1.* Let  $\rho$  be an arbitrary quantum state,  $H$  be an arbitrary Hamiltonian with spectrum set  $\{E_1, \dots, E_k\}$ , and  $O$  be an arbitrary observable. Then

$$\text{tr}[\exp(-i\theta H)\rho \exp(i\theta H)O] = C + \sum_{i>j} A_{i,j} \cos[\theta(E_i - E_j) + B_{i,j}], \quad (\text{B1})$$

for some real values  $C, A_{i,j}, B_{i,j}$ .

*Proof.* Let  $U$  be a unitary that diagonalizes the Hamiltonian  $H$ . Then we have

$$\begin{aligned} \text{tr}[\exp(-i\theta H)\rho \exp(i\theta H)O] &= \text{tr} \left[ \sum_{i=1}^k (U e^{-i\theta E_i} |i\rangle\langle U|^\dagger) \rho \sum_{j=1}^k (U e^{i\theta E_j} |j\rangle\langle U|^\dagger) O \right] \\ &= \sum_{i=1}^k \sum_{j=1}^k e^{i\theta(E_j - E_i)} \text{tr}[U|i\rangle\langle U|^\dagger \rho U|j\rangle\langle U|^\dagger O] \\ &= \sum_{i=1}^k \sum_{j=1}^k e^{i\theta(E_j - E_i)} \text{tr}[|i\rangle\langle \rho'|j\rangle\langle O'] \\ &= \sum_{i=1}^k \sum_{j=1}^k e^{i\theta(E_j - E_i)} \langle i|\rho'|j\rangle \langle j|O'|i\rangle, \end{aligned} \quad (\text{B2})$$

where  $\rho' = U^\dagger \rho U$  and  $O' = U^\dagger O U$ . Since  $\rho'$  is a hermitian operator and therefore  $\langle i|\rho'|j\rangle = \overline{\langle j|\rho'|i\rangle}$ , and similarly for  $O'$ , therefore for any  $i, j$  the term for  $i > j$  is a conjugate of the term  $i < j$ . Hence

$$\sum_{i=1}^k \sum_{j=1}^k e^{i\theta(E_j - E_i)} \langle i|\rho'|j\rangle \langle j|O'|i\rangle = \sum_{i=1}^k \langle i|\rho'|i\rangle \langle i|O'|i\rangle + 2 \sum_{i>j} \text{Re}[e^{i\theta(E_j - E_i)} \langle i|\rho'|j\rangle \langle j|O'|i\rangle]. \quad (\text{B3})$$

Note that the left-hand side sum in the above above is a free term and is a real number. Starting from now we will assume that the Hamiltonian  $H$  is nondegenerate, otherwise, the corresponding element of the right sum will contribute to the free term. Taking

$x_{i,j} + iy_{i,j} := \langle i|e^{i\theta}|j\rangle\langle j|O|i\rangle$  for some real  $x_{i,j}, y_{i,j}$  we have

$$\begin{aligned} \operatorname{Re}[e^{i\theta(E_j-E_i)}\langle i|e^{i\theta}|j\rangle\langle j|O|i\rangle] &= \operatorname{Re}\{\cos[\theta(E_j-E_i)] + i\sin[\theta(E_j-E_i)]\}(x_{i,j} + iy_{i,j}) \\ &= x_{i,j}\cos[\theta(E_j-E_i)] - y_{i,j}\sin[\theta(E_j-E_i)] \\ &= \sqrt{x_{i,j}^2 + y_{i,j}^2} \left( \frac{x_{i,j}}{\sqrt{x_{i,j}^2 + y_{i,j}^2}} \cos[\theta(E_j-E_i)] - \frac{y_{i,j}}{\sqrt{x_{i,j}^2 + y_{i,j}^2}} \sin[\theta(E_j-E_i)] \right) \\ &= \sqrt{x_{i,j}^2 + y_{i,j}^2} \{\cos(\alpha_{i,j})\cos[\theta(E_j-E_i)] - \sin(\alpha_{i,j})\sin[\theta(E_j-E_i)]\}, \end{aligned} \quad (\text{B4})$$

where  $\alpha_{i,j}$  is such a real number for which the above transformation holds. Note that such a number  $\alpha$  can always be found as the replaced fraction squared sum to 1 and one can use Pythagorean trigonometric identity. Finally, we have

$$\sqrt{x_{i,j}^2 + y_{i,j}^2} \{\cos(\alpha_{i,j})\cos[\theta(E_j-E_i)] - \sin(\alpha_{i,j})\sin[\theta(E_j-E_i)]\} = \sqrt{x_{i,j}^2 + y_{i,j}^2} \cos[\theta(E_j-E_i) + \alpha_{i,j}], \quad (\text{B5})$$

which proves the statement of the theorem. ■

Note that the case of Hamiltonian with two different eigenvalues was already presented in [49,50].

- 
- [1] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [2] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke *et al.*, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [3] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [4] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [5] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, Variational quantum linear solver, *Quantum* **7**, 1188 (2023).
- [6] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, *Phys. Rev. A* **101**, 010301(R) (2020).
- [7] R. LaRose, A. Tikku, É. O’Neel-Judy, L. Cincio, and P. J. Coles, Variational quantum state diagonalization, *npj Quant. Info.* **5**, 1 (2019).
- [8] A. Kundu and J. A. Miszczak, Variational certification of quantum devices, *Quant. Sci. Technol.* **7**, 045017 (2022).
- [9] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [10] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem, *Science* **292**, 472 (2001).
- [11] Z. Tabi, K. H. El-Safty, Z. Kallus, P. Hága, T. Kozsik, A. Glos, and Z. Zimborás, Quantum optimization for the graph coloring problem with space-efficient embedding, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2020), pp. 56–62.
- [12] B. Bakó, A. Glos, Ö. Salehi, and Z. Zimborás, Near-optimal circuit design for variational quantum optimization, [arXiv:2209.03386](https://arxiv.org/abs/2209.03386).
- [13] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem, [arXiv:1412.6062](https://arxiv.org/abs/1412.6062).
- [14] J. Cook, S. Eidenbenz, and A. Bärttschi, The quantum alternating operator ansatz on maximum k-vertex cover, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2020), pp. 83–92.
- [15] A. Glos, A. Krawiec, and Z. Zimborás, Space-efficient binary optimization for variational quantum computing, *npj Quantum Inf.* **8**, 1 (2022).
- [16] M. Medvidović and G. Carleo, Classical variational simulation of the quantum approximate optimization algorithm, *npj Quantum Inf.* **7**, 1 (2021).
- [17] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Quantum approximate optimization algorithm for MaxCut: A fermionic view, *Phys. Rev. A* **97**, 022304 (2018).
- [18] M. Alam, A. Ash-Saki, and S. Ghosh, Accelerating quantum approximate optimization algorithm using machine learning, in *2020 Design, Automation & Test in Europe Conference & Exhibition (DATE)* (IEEE, New York, 2020), pp. 686–689.
- [19] R. Shaydulin, I. Safro, and J. Larson, Multistart methods for quantum approximate optimization, in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (IEEE, New York, 2019), pp. 1–8.
- [20] N. Hegade, P. Chandarana, K. Paul, X. Chen, F. Albarrán-Arriagada, and E. Solano, Portfolio optimization with digitized-counterdiabatic quantum algorithms, *Phys. Rev. Res.* **4**, 043204 (2022).
- [21] L. Zhu, H. L. Tang, G. S. Barron, F. A. Calderon-Vargas, N. J. Mayhall, E. Barnes, and S. E. Economou, Adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer, *Phys. Rev. Res.* **4**, 033029 (2022).
- [22] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices, *Phys. Rev. X* **10**, 021067 (2020).

- [23] Z. Zhou, Y. Du, X. Tian, and D. Tao, QAOA-in-QAOA: Solving large-scale MaxCut problems on small quantum machines, *Phys. Rev. Appl.* **19**, 024027 (2023).
- [24] Y. J. Patel, S. Jerbi, T. Bäck, and V. Dunjko, Reinforcement learning assisted recursive QAOA, *EPJ Quantum Tech.* **11**, 6 (2024).
- [25] L. T. Watson and R. T. Haftka, Modern homotopy methods in optimization, *Comput. Methods Appl. Mech. Eng.* **74**, 289 (1989).
- [26] A. Garcia-Saez and J. Latorre, Addressing hard classical problems with adiabatically assisted variational quantum eigen-solvers, [arXiv:1806.02287](https://arxiv.org/abs/1806.02287).
- [27] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
- [28] S. M. Harwood, D. Trenev, S. T. Stober, P. Barkoutsos, T. P. Gujarati, S. Mostame, and D. Greenberg, Improving the variational quantum eigensolver using variational adiabatic quantum computing, *ACM Trans. Quant. Comp.* **3**, 1 (2022).
- [29] J. R. McClean, M. P. Harrigan, M. Mohseni, N. C. Rubin, Z. Jiang, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Low-depth mechanisms for quantum optimization, *PRX Quantum* **2**, 030312 (2021).
- [30] A. Messiah and G. Temmer, *Quantum Mechanics*, Vol. 1 in Quantum Mechanics (North-Holland, Amsterdam, 1961).
- [31] T. Albash and D. A. Lidar, Adiabatic quantum computation, *Rev. Mod. Phys.* **90**, 015002 (2018).
- [32] K.-P. Marzlin and B. C. Sanders, Inconsistency in the application of the adiabatic theorem, *Phys. Rev. Lett.* **93**, 160408 (2004).
- [33] D. M. Tong, K. Singh, L. C. Kwek, and C. H. Oh, Quantitative conditions do not guarantee the validity of the adiabatic approximation, *Phys. Rev. Lett.* **95**, 110407 (2005).
- [34] J. Du, L. Hu, Y. Wang, J. Wu, M. Zhao, and D. Suter, Experimental study of the validity of quantitative conditions in the quantum adiabatic theorem, *Phys. Rev. Lett.* **101**, 060403 (2008).
- [35] J. Wu, M. Zhao, J. Chen, and Y. Zhang, Adiabatic condition and quantum geometric potential, *Phys. Rev. A* **77**, 062114 (2008).
- [36] Ö. Salehi, A. Glos, and J. A. Mischak, Unconstrained binary models of the travelling salesman problem variants for quantum optimization, *Quantum Info. Proc.* **21**, 67 (2022).
- [37] K. Domino, M. Koniorczyk, K. Krawiec, K. Jałowiecki, S. Deffner, and B. Gardas, Quantum annealing in the NISQ era: railway conflict management, *Entropy* **25**, 191 (2023).
- [38] K. Domino, A. Kundu, Ö. Salehi, and K. Krawiec, Quadratic and higher-order unconstrained binary optimization of railway rescheduling for quantum computing, *Quantum Info. Proc.* **21**, 337 (2022).
- [39] M. Borowski, P. Gora, K. Karnas, M. Błajda, K. Król, A. Matyjasek, D. Burczyk, M. Szewczyk, and M. Kutwin, New hybrid quantum annealing algorithms for solving vehicle routing problem, in *International Conference on Computational Science* (Springer, New York, 2020), pp. 546–561.
- [40] A. Arya, L. Botelho, F. Cañete, D. Kapadia, and Ö. Salehi, *Applications of Quantum Annealing to Music Theory* (Springer, Cham, Germany, 2022), pp. 373–406.
- [41] A. Glos, A. Kundu, and Ö. Salehi, Optimizing the production of test vehicles using hybrid constrained quantum annealing, *SN Comput. Sci.* **4**, 609 (2023).
- [42] R. Babbush, P. J. Love, and A. Aspuru-Guzik, Adiabatic quantum simulation of quantum chemistry, *Sci. Rep.* **4**, 1 (2014).
- [43] A. Das and B. K. Chakrabarti, Colloquium: Quantum annealing and analog quantum computation, *Rev. Mod. Phys.* **80**, 1061 (2008).
- [44] S. Hadfield, Z. Wang, B. O’gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, From the quantum approximate optimization algorithm to a quantum alternating operator ansatz, *Algorithms* **12**, 34 (2019).
- [45] A. Bärtshi and S. Eidenbenz, Grover mixers for QAOA: Shifting complexity from mixer design to state preparation, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2020), pp. 72–82.
- [46] Z. Wang, N. C. Rubin, J. M. Dominy, and E. G. Rieffel, XY-mixers: Analytical and numerical results for the quantum alternating operator ansatz, *Phys. Rev. A* **101**, 012320 (2020).
- [47] E. Crosson, E. Farhi, C. Y.-Y. Lin, H.-H. Lin, and P. Shor, Different strategies for optimization using the quantum adiabatic algorithm, [arXiv:1401.7320](https://arxiv.org/abs/1401.7320).
- [48] S. Muthukrishnan, T. Albash, and D. A. Lidar, Tunneling and speedup in quantum optimization for permutation-symmetric problems, *Phys. Rev. X* **6**, 031010 (2016).
- [49] M. Ostaszewski, E. Grant, and M. Benedetti, Structure optimization for parameterized quantum circuits, *Quantum* **5**, 391 (2021).
- [50] K. M. Nakanishi, K. Fujii, and S. Todo, Sequential minimal optimization for quantum-classical hybrid algorithms, *Phys. Rev. Res.* **2**, 043158 (2020).
- [51] D. M. Dunlavy and D. P. O’Leary, Homotopy optimization methods for global optimization, Technical Rep. SAND2005-7495 (Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA, 2005), <https://www.osti.gov/biblio/876373%7D>.
- [52] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen *et al.*, QISKIT: An open-source framework for quantum computing, Accessed on: Nov 22, 2023.
- [53] D. C. Liu and J. Nocedal, On the limited memory bfgs method for large scale optimization, *Math. Program.* **45**, 503 (1989).
- [54] C. Campbell and E. Dahl, QAOA of the highest order, in *2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C)* (IEEE, New York, 2022), pp. 141–146.
- [55] <https://doi.org/10.5281/zenodo.7585691>.
- [56] <https://github.com/iitis/hoho-qaoa-code>.