

Molecular docking via quantum approximate optimization algorithm

Qi-Ming Ding^{1,2,*}, Yi-Ming Huang^{1,2,†}, and Xiao Yuan^{1,2,‡}

¹Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China

²School of Computer Science, Peking University, Beijing 100871, China



(Received 22 August 2023; revised 27 January 2024; accepted 29 February 2024; published 19 March 2024; corrected 17 June 2024)

Molecular docking plays a pivotal role in drug discovery and precision medicine, furnishing insights into protein functionalities and fostering the development of therapeutics. Here, we introduce a potential alternative solution to this problem using a quantum approximate optimization algorithm (QAOA) based algorithm. Our method was applied to analyze diverse biological systems, including the SARS-CoV-2 M^{pro} complex with PM-2-020B, the DPP-4 complex with piperidine fused imidazopyridine 34, and the HIV-1 gp120 complex with JP-III-048. We found that the digitized-counterdiabatic QAOA (DCQAOA), which integrates the concept of counterdiabatic driving, surpasses the conventional QAOA in terms of quantum circuit depth and optimization efficiency. This is particularly evident in complex molecular-docking challenges, where DCQAOA delivers more accurate and biologically relevant results. Our research highlights the promising role of quantum computing in drug discovery and contributes significant insights towards the optimization of protein-ligand docking methodologies.

DOI: [10.1103/PhysRevApplied.21.034036](https://doi.org/10.1103/PhysRevApplied.21.034036)

I. INTRODUCTION

Quantum computing, a concept introduced in the 1980s by Feynman and other pioneers [1–3], has made significant strides, particularly with the developments of Shor's algorithm [4] and others [5]. Despite these advances, constructing a universal fault-tolerant quantum computer remains experimentally challenging. Currently, we are in the noisy intermediate-scale quantum (NISQ) era, where state-of-the-art quantum devices have a limited number of qubits and imperfect operations [6]. Since the NISQ hardware is weaker than a universal quantum computer [6–8], it is widely believed that NISQ hardware can only be used to solve specific problems. Therefore, the identification of tasks that could leverage quantum features of NISQ devices is of great significance [8–10].

One promising application of quantum computing lies in the realm of drug discovery [11,12], encompassing vital areas such as protein folding and molecular docking. Protein folding, commonly approached through two-dimensional (2D) or three-dimensional (3D) lattice models to minimize interaction energy, can be elegantly reframed as a Hamiltonian problem. By employing quantum annealing [13–15] or gate-based quantum

computation [16–19], we can reformulate protein folding as a ground-state problem. Significant progress has been made in this field, with explorations of system sizes up to nine amino acids utilizing various quantum hardware platforms [19]. Molecular docking, another potential field in quantum applications, can be broadly classified into score function-based approaches and score function-independent methods. Score function-based approaches involve scoring functions to evaluate docking poses, as demonstrated in recent studies [20,21]. On the other hand, score function-independent methods explore alternative strategies. For example, Gaussian-boson sampling enables the identification of docking postures based on the geometric relationship between proteins and ligands [22,23]. Another approach involves solving the molecular unfolding problem using quantum annealing, which serves as the first step in geometric molecular-docking techniques [24].

However, approaches based on scoring functions require precise computation of the binding energy between proteins and ligands. Even a slight error of 6 kJ/mol can cause significant deviations in the docking results [11], and achieving such accuracy with current quantum devices is difficult. Meanwhile, the other approach employing Gaussian-boson sampling does not provide a definitive guarantee of clique formation, and additional calculations are generally necessary to confirm its occurrence [22]. Additionally, in practical docking scenarios, the adjacency matrix often exhibits sparsity, which can lead to experimental parameters approaching zero. As a result, achieving

*dqiming94@pku.edu.cn

†ymhwang@pku.edu.cn

‡xiaoyuan@pku.edu.cn

§These authors contributed equally to this work.

these parameters poses a challenge in optical experiments [23].

A potential solution to these challenges is to exploit digital quantum computing to solve the maximum weighted clique of a given graph, thereby addressing the molecular-docking problem. Given the constraints of current quantum hardware, many NISQ algorithms adopt a hybrid quantum-classical strategy, known as variational quantum algorithms (VQAs) [25]. The quantum approximate optimization algorithm (QAOA) is a prominent type of VQA [26], which uses shallow quantum circuits and classical optimization to find the optimal state corresponding to a p -level parametrized quantum circuit. As $p \rightarrow \infty$, the final state approaches an adiabatically evolved state and becomes an exact maximal eigenstate. However, increasing p also leads to escalating classical optimization costs. Efforts to enhance QAOA primarily focus on reducing the quantum circuit depth [27], with improved versions such as ADAPT QAOA [28], digital-analog QAOA [29], and digitized-counterdiabatic QAOA(DCQAOA) [30]. In particular, DCQAOA, incorporates a counterdiabatic (CD) driving term to develop a superior ansatz, thereby enhancing global performance. The QAOA has been validated in various applications, including Ising spin models, entangled-state preparation, classical optimization problems, such as MaxCut and the SK model, and the P -spin model [26,31].

In this paper, we propose using digital quantum computing to tackle the molecular-docking problem, thereby circumventing the limitations inherent to the Gaussian-boson sampling method. Drawing inspiration from the previous literature [22,23,32], we improved the criterion of constructing a protein-ligand binding interaction graph (BIG) between specified proteins and ligands. Our objective is to identify the maximum vertex weight clique within this graph, which signifies the most probable docking posture. This approach uniquely translates a complex biological problem into a mathematical optimization problem that can be effectively addressed using digital quantum computers. We derive the cost Hamiltonian from mathematical optimization formulations of the maximum vertex weight clique and solve it using QAOA and DCQAOA. From the ground state, we can then extract the optimal binding conformation. Our method introduces an alternative paradigm for applying near-term digital quantum computers to the field of biopharmaceutics, underscoring the potential of quantum computing in solving intricate real-world problems.

II. METHODS

The exploration of drug discovery often involves the identification of small molecules, or ligands, that demonstrate high-affinity binding to specific target proteins. This binding can modulate the activity, functionality,

or interactions of the protein with other biomolecules, thus eliciting therapeutic effects. However, the conformational flexibility of proteins and ligands, coupled with the vast search space of potential binding poses, makes this a computationally challenging problem. To address this challenge, we first map the molecular-docking problem onto the maximum vertex weight clique problem, as previously outlined in Ref. [32]. This approach is particularly intriguing due to its potential applicability in quantum computing, notably in conjunction with Gaussian-boson sampling as explored in earlier studies. [22,23]. However, the exploration of the molecular-docking problem on general-purpose quantum computers, particularly those that are gate based or circuit based, remains limited. Here, we study alternative ways to solve molecular docking using digital circuit-based quantum computers.

A. Mapping molecular docking to maximum vertex weight clique

In the realm of protein-ligand docking, the significance of any atoms is not uniformly distributed. Thus, we focus on pharmacophores—specific sets of atoms that play a pivotal role in dictating a molecule’s pharmacological and biological interactions. These pharmacophores include elements with negative or positive charges, the hydrogen-bond donor or acceptor, the hydrophobe, and the aromatic ring. Given the constraints imposed by the current limitations of quantum devices, our study strategically narrows its focus to the most significant pharmacophores or adopts heuristic methods for selecting pharmacophore points.

To represent the geometric relationship of pharmacophores on proteins and ligands, we create a labeled distance graph (LDG). [Please refer to the Figs. 1(a) and 1(b)]. In the construction of the LDG for proteins and ligands, vertices represent selected pharmacophore points, with each vertex distinctly annotated to indicate its typology. The vertices are interconnected by edges, whose weights are determined by the Euclidean distances between corresponding pharmacophore points. This graphical representation enables the elucidation of spatial relationships amongst pharmacophores, particularly focusing on the docking points of interest on the protein or ligand under study.

After the selection and construction of the LDGs for the protein or ligand, the subsequent imperative involves the creation of a BIG. The primary objective of this BIG is to serve as a visual representation of the docking relationship, thereby elucidating potential interactions among pharmacophore points.

In the BIG, the vertex, denoted as (v_l, v_p) , encapsulates a pair of points, where v_l denotes a vertex originating from the ligand, and v_p signifies a vertex from the protein.

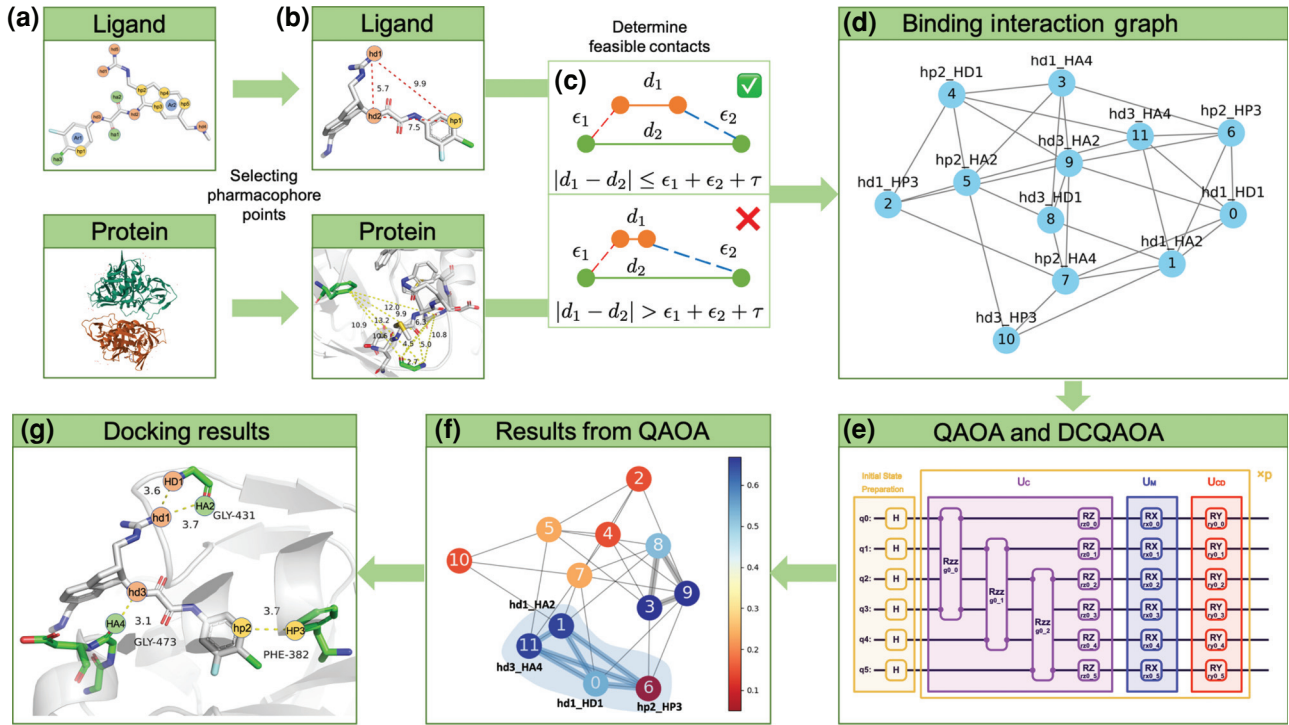


FIG. 1. Illustration of the protein-ligand docking processes with employing the QAOA and DCQAOA, exemplified by the complex formation involving HIV-1 gp120 and JP-III-048 (PDB ID: 5F4L). Panel (a) displays the structures of the protein and ligand. For analytical purposes, three pharmacophores on the ligand and four on the protein were identified, leading to the construction of LDGs for each, as depicted in (b). Subsequently, the BIG is defined, its vertices quantitatively correlating to the product of the number of selected pharmacophores on both ligand and protein. The establishment of edges between any two vertices follows the criteria illustrated in (c). Panel (d) illustrates the resultant graph. The application of both the QAOA and its derivative, DCQAOA algorithms, the latter incorporating CD driving terms, is then implemented. The results of these computational solutions are presented in (f), with the corresponding docking poses elucidated in (g).

Consequently, the total vertex count in the BIG is symbolized as $N = nm$, with n and m indicating the number of vertices in the LDG of the ligand and the protein, respectively. This integrative process results in a holistic representation, encompassing all possible theoretical interaction pairings, herein termed *connects*. Within this framework, each edge in the BIG, expressed as $(v_{l_i}, v_{p_i}) - (v_{l_j}, v_{p_j})$, delineates two vertices on the edge that are capable of co-existing in a potential docking posture. The presence of a clique in the BIG implies that every vertex within the clique can concurrently exist, thereby depicting a feasible docking posture. Furthermore, the resolution of the MVWCP within the BIG identifies the most probable docking posture from all potential configurations.

The inherent geometrical constraints of both the ligand and its binding site fundamentally limit the feasibility of certain combinations of contact points. This necessitates an initial assessment aimed at identifying contact points that are mutually compatible. The primary criterion for determining the compatibility of these contact points hinges on the geometric proximities between the ligand's and the binding site's respective contact points.

In this analysis, compatibility of a pair of contact points is determined based on the absolute difference in pharmacophore point distances, $|d_1 - d_2|$. This difference must meet a specific threshold defined as the sum of the flexibility constant τ and interaction distances ϵ_1 and ϵ_2 . As depicted in Fig. 1(c), a pair is deemed compatible if $|d_1 - d_2| \geq \tau + \epsilon_1 + \epsilon_2$; otherwise, it is incompatible. Here, d_1 is the Euclidean distance between ligands at vertices v_{l_1} and v_{l_2} , and d_2 is the distance between proteins at vertices v_{p_1} and v_{p_2} . The parameters τ , ϵ_1 , and ϵ_2 represent the flexibility allowance and allowed interaction distances.

Choosing a specific ϵ value involves combining empirical observations, experimental data, and theory. This choice depends on the pharmacophore's nature at both interaction ends, as different pharmacophores create unique interactions with their own typical distances. For example, hydrogen bonding occurs when a hydrogen atom in the ligand, attached to an electronegative atom like oxygen or nitrogen, forms a favorable interaction with a hydrogen-bond donor or acceptor in the protein. These bond lengths are usually between 2.5 to 3.3 Å. Hydrophobic interactions happen between nonpolar groups in the

ligand and hydrophobic protein residues, caused by water molecules being excluded from the hydrophobic area. The ideal distances for these interactions are generally between 3.3 to 4 Å. $\pi - \pi$ stacking happens when aromatic rings in the ligand and protein align, forming interactions through π -electron clouds, with typical distances around 3.5 to 4.5 Å.

Understanding these interactions leads to a flexible approach in selecting ϵ . If all pharmacophore points are hydrogen-bond receptors and donors, choosing a shorter distance can improve docking prediction accuracy. However, if the pharmacophore points involve different interaction types, selecting a fixed interaction distance can simplify the process. This adaptable method enhances docking accuracy by adjusting to various pharmacophore point interactions. The effectiveness of this approach will be shown in the upcoming examples in Sec. III.

B. Mathematical optimization formulations of maximum vertex weight clique

Considering a specific molecular-docking problem, the BIG is constructed, denoted as $BIG = (V, E)$, where V represents the set of vertices and E signifies the feasible set of edges. In this graph, each vertex is associated with a positive weight w_i , indicative of its pharmacophore potential within the context of molecular docking. The MVWCP in this context is formulated as a binary quadratic programming (BQP) problem, as delineated in Ref. [33]:

$$\begin{aligned} & \text{maximize} \quad \sum_{i=1}^N w_i x_i + \sum_{i=1}^N \sum_{j=1, j \neq i}^N w_{ij} x_i x_j, \\ & \text{subject to} \quad x_i \in \{0, 1\}, \quad \forall i \in \{1, \dots, N\}. \end{aligned} \quad (1)$$

Here, $N = |V|$ denotes the scale of the problem, equating to the number of qubits utilized for its resolution. The binary variable x_i corresponds to the vertex v_i . The weight $w_{ij} = P$ is assigned if $\{v_i, v_j\} \in \bar{E}$, and zero otherwise, where $\bar{E}$ represents the edge set of the complementary graph $\bar{G}$, and P is a predetermined negative penalty scalar. This BQP formulation, a nonlinear interpretation of MVWCP, parallels prior investigations, such as those outlined in Ref. [34]. The quadratic function yields an equivalent objective value to the linear form when all penalty terms are set to 0, signifying the fulfillment of all constraints. Theoretically, a higher absolute magnitude of $|P|$ simplifies attaining accurate results, as it deters the formation of nonpermissible edges. However, empirical observations from quantum computational simulations suggest a nonlinear relationship with P , indicating that excessive values can lead to suboptimal results. This can be attributed to the disproportionate reduction of w_i 's coefficient after normalization, which impedes the efficacy of QAOA in finding the optimal

solution. These phenomena are further elucidated through subsequent examples in Sec. III.

Notably, if $w_i = 1$ for all $i \in \{1, \dots, N\}$, Eq. (1) transforms into the linear model of the classical maximum clique problem (MCP). The MCP, as a decision problem, is established in the literature as NP complete, presenting significant computational challenges [35]. The MVWCP, as a more complex variant of the MCP, inherently poses computational obstacles that are at least as formidable. Consequently, efficiently and precisely solving an arbitrary MCP is improbable. Nonetheless, in the context of approximate solutions, quantum computing holds the potential for enhanced efficiency, as evidenced by prior research in graph theory concerning the maximum cut (MaxCut) and maximum independent set (MIS) problems [7, 26, 36–42], [42], in the realms of the Sherrington-Kirkpatrick (SK) model and the k -spin model within the theory of spin glasses and disordered systems [43–47], the problem of low autocorrelation binary sequences (LABS) [48], and, more broadly, in quadratic unconstrained binary optimization (QUBO) problems [49] (see also the review, Ref. [27]). Gratifyingly, the optimization problem represented by Eq. (1) can also be formulated as a QUBO problem, rendering it especially amenable to quantum computing approaches.

To construct the corresponding Hamiltonian, we map the binary variables x_i to σ_i^z matrices acting on individual qubits. Given the relationship as $x_i = (\sigma_i^z - 1)/2$, the objective function can be rewritten in terms of σ_i^z :

$$H = \frac{1}{2} \sum_{i \in V} w_i (\sigma_i^z - 1) + \frac{P}{4} \sum_{(i,j) \notin E, i \neq j} (\sigma_i^z - 1) (\sigma_j^z - 1). \quad (2)$$

Simplifying the above Hamiltonian yields

$$H_C = \sum_{i \in V} w_i \sigma_i^z + \frac{P}{2} \sum_{(i,j) \notin E, i \neq j} (\sigma_i^z \sigma_j^z - \sigma_i^z - \sigma_j^z) + \text{const.} \quad (3)$$

Here, the constant term does not influence the optimization process and can be omitted during the implementation of the quantum algorithm. By minimizing the expectation value of the cost Hamiltonian H_C with respect to a quantum state, we can find the ground state, which corresponds to the largest weighted clique in the graph.

C. QAOA and DCQAOA

The QAOA is a quantum algorithm tailored for solving combinatorial optimization problems. This method exploits the adiabatic principle, facilitated by a time-dependent Hamiltonian $H_{\text{adi}}(t)$, defined as

$$H_{\text{adi}}(t) = [1 - \lambda(t)]H_M + \lambda(t)H_C, \quad (4)$$

where $\lambda(t)$, a monotonic function within the range $[0, 1]$, orchestrates the annealing protocol over the interval

$t \in [0, T]$. The algorithm initializes the system in the ground state of a simpler Hamiltonian H_M and then adiabatically evolves towards the problem Hamiltonian H_C , whose ground state encodes the desired solution. According to the adiabatic theorem, if the evolution is sufficiently slow, particularly ensuring a nonvanishing energy gap between the ground state and the first excited state, the system predominantly remains in the ground state. The solution is subsequently extrapolated from the ground state of H_C .

The QAOA approximates this evolution by alternating unitary operations $U_M(\beta_j) = \exp(-i\beta_j H_M)$ and $U_C(\gamma_j) = \exp(-i\gamma_j H_C)$, with parameters β_j and γ_j . The process, limited to p iterations, is described as

$$U(\boldsymbol{\gamma}, \boldsymbol{\beta}) = U_M(\beta_p) U_C(\gamma_p) U_M(\beta_{p-1}) U_C(\gamma_{p-1}) \dots U_M(\beta_1) U_C(\gamma_1), \quad (5)$$

Here, p denotes the number of alternating applications of U_C and U_M , we also called the layers of QAOA. It is evident that the search space expanded by p layers of unitary operators U_C and U_M is more extensive than that of $p - 1$ layers.

The expected value of Hamiltonian H_C in the quantum state is given by $E_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$:

$$E_p(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \boldsymbol{\gamma}, \boldsymbol{\beta} | H_C | \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle. \quad (6)$$

The conventional QAOA commonly uses a single global parameter for $U_M(\beta)$ and $U_C(\gamma)$. To achieve better results, a significant number of layers p are often necessary. However, a recent enhancement known as the multiangle QAOA (MAQAOA), or full-parameter QAOA (FPQAOA), has been proposed [50,51], which can effectively reduce the number of layers at the cost of increasing the parameters optimized for each layer. Unless otherwise specified, all numerical simulations presented in this paper are conducted using this more intricate and physically representative full-parameter model.

Moreover, another variant, DCQAOA, incorporates CD terms [see Fig. 1(e)] [30]. Studies show that this addition significantly lowers the needed value of p for a near-optimal trial state [30]. Combining these benefits of two algorithms, we introduce the multiangle digitized-counterdiabatic QAOA (MADCQAOA). For simplicity, we refer to it as DCQAOA throughout our discussion.

The integration of the CD driving term, based on shortcuts to adiabaticity principles [52–54], mitigates nonadiabatic transitions and accelerates quantum processes, as noted in studies such as Refs. [55–59], with notable applications in quantum state preparation tasks [60]. In this paper, we follow the approach in Ref. [30] and adopt the nested commutator framework of the adiabatic gauge potential as introduced by Claeys et al. [61], to construct a pool of CD operators.

In the present study, we adopt the nested commutator framework for the adiabatic gauge potential, as introduced by Claeys *et al.* [61] and further explored in Chandarana *et al.* [30], which involves the construction of a comprehensive pool of CD operators. Specifically, we limit our focus to the second-order expansion within the nested commutator framework, denoted as $l = 2$. Consequently, the operator pool A is defined as $\{\sigma^y, \sigma^z \sigma^y, \sigma^y \sigma^z, \sigma^x \sigma^y, \sigma^y \sigma^x\}$, thereby encapsulating a range of both local and two-body interactions.

Choosing the right operator pool is useful and depends on the Hamiltonian being studied. Different problems may require different operators. For example, a key paper, Ref. [30], presents a method for selecting CD terms for the Ising model under various conditions. This method defines the Hamiltonian for the one-dimensional Ising spin model as $H_{\text{prob}}(\sigma) = -\sum_{(i,j)} J_{ij} \sigma_i^z \sigma_j^z - \sum_i h_i \sigma_i^z - \sum_i k_i \sigma_i^x$. In this formula, σ_i^d represents the Pauli matrices at each site, and (i,j) indicates nearest-neighbor interactions with the coupling constant J_{ij} .

This approach results in two scenarios. The first is the linear ferromagnetic Ising model (LFIM), where $k_i = 0$ and the CD term is $A_t = \sum_i \sigma_i^y$. The second scenario is the transverse-field Ising model (TFIM), which is defined when $h_i = 0$ and the CD term is $A_t = \sum_i \sigma_i^z \sigma_{i+1}^z$. In our example, we consider the Hamiltonian H_C for the MVWCP as shown in Eq. (3). In this case, the model matches LFIM, which leads to the CD term $A_t = \sum_i \sigma_i^y$.

The CD component in the quantum circuit is represented by the unitary operator

$$U_{\text{CD}}(\alpha) = \prod_{j=1}^L \exp[-i\alpha A_j^q], \quad (7)$$

where A_j^q is the q -local CD operator from the set A .

In circuit structure terms, DCQAOA improves on the original QAOA by adding an extra CD term. This change leads to a deeper circuit and a more complex optimization space. In a QAOA framework where each quantum gate has its own parameter, a DCQAOA circuit with the same number of layers will have more parameterized gates. This is especially true when the CD term is complex. However, this complexity does not mean it is less efficient. DCQAOA can achieve its goals with fewer layers, which is beneficial for reducing the number of gates. Sec. 3B of this study shows this advantage.

Moreover, the application of DCQAOA to intricate tasks such as molecular docking is further optimized by our specific CD term, denoted as $A_t = \sum_i \sigma_i^y$. Adding this term only inserts some single-qubit parameter gates after each QAOA layer, but can effectively reduce the circuit depth from the subsequent numerical simulations. This reduction is crucial as it avoids the need for more complex double-qubit gates, which is a key factor considering

the limitations of current quantum computing hardware. Additionally, this approach also simplifies optimization problems. Collectively, these advancements underscore the considerable potential of DCQAOA in effectively addressing large-scale molecular-docking challenges, thereby expanding its applicative scope and reinforcing its utility in the field of quantum computing.

III. NUMERICAL SIMULATION

In this section, we focus on the practical implementation of the QAOA-based algorithms in tackling three protein-ligand docking problems. The comprehensive approach used for protein-ligand docking through QAOA algorithms is illustrated in Fig. 1.

First, we demonstrate the complete process of our algorithm using a small-scale example, as illustrated in Fig. 1, and highlight the effect of flexible choice of ϵ . Next, we focus on a medium-scale docking problem within an eight-qubit quantum system, and conduct a thorough analysis to evaluate the effect of the different QAOA parameters for the docking process. This involved examining the number of QAOA layers (p), a penalty term (P), and evaluating the impact of quantum noise on the docking process. Finally, we applied our method to a challenging 12-qubit example, characterized by almost identical weights in the largest and second-largest cliques, which illustrates the advantages of DCQAOA in molecular-docking problems.

A. The impact of the flexible ϵ

We begin our numerical investigation with a six-qubit quantum system to explain the operational process of our proposed QAOA-based molecular-docking algorithm, and focus on empirically examining the influence of the ϵ parameter on docking results. The research delineates the complex interaction between the SARS-CoV-2 main protease (M^{pro}), a critical target in COVID-19 therapeutics, and a covalent pyrazoline-based inhibitor, PM-2-020B. This inhibitor has been identified through the Protein Data Bank (PDB) with the ID 8SKH, as reported in the seminal work [62].

The main protease (M^{pro}) of the coronavirus, a crucial enzyme in viral replication, presents a potential target for therapeutic intervention. Researchers have identified and optimized pyrazoline-based covalent inhibitors for SARS-CoV-2 M^{pro} using activity-based protein profiling. These inhibitors have shown promising efficacy against M^{pro} from SARS-CoV-2 and other coronaviruses, underscoring their potential as pan-coronavirus inhibitors. In this context, we delve into the docking process between the pyrazoline-based inhibitor PM-2-020B and SARS-CoV-2 M^{pro} , providing a detailed investigation of this interaction.

Our method begins with the extraction of all pharmacophore points on ligands and receptors using the Python

package *rdkit* [63]. We classify these points into four distinct categories: hydrogen-bond donor and acceptor (HD and HA), hydrophobe (HP), and aromatic ring (AR). Vertices on the ligand and receptor are denoted with lowercase and uppercase letters, respectively. This process yields seven pharmacophore points on the PM-2-020B ligand and 1422 points on the SARS-CoV-2 M^{pro} receptor.

The QAOA-based molecular-docking algorithm requires $N = m \times n$ qubits to find the maximum weighted clique, where n and m represent the vertex counts in the ligand and protein binding-site graphs, respectively. Given the constraints of current quantum devices, it is crucial to minimize the number of molecules to be paired. To this end, we select pharmacophore points on the protein within a 5-angstrom radius of the ligand, based on the actual binding pose. While the actual binding pose remains unknown in real scenarios, a similar set of points can be procured using drug-discovery knowledge. This includes heuristics selection of ligand pharmacophore points and the use of prior knowledge of binding-site locations to reduce receptor pharmacophore points. Sliding windows can further minimize receptor pharmacophore points by segregating different binding-site regions. As quantum devices evolve, such simplifications may become redundant.

In this example, we identified two pharmacophore points on PM-2-020B, specifically O2378 (ha1) and C2373 (hp2), as well as three pharmacophore points on SARS-CoV-2 M^{pro} , namely N1104 (HD1), N1108 (HD2), and N1114 (HD3). Here, the capital letter preceding the number indicates the atomic type of the pharmacophore, while the number corresponds to the atomic serial number in the PDB file. Once we obtained the two annotated distance graphs (one for the PM-2-020B ligand and another for the SARS-CoV-2 M^{pro} receptor), we proceeded to construct the binding interaction graph. As previously mentioned, the vertices in this graph denote potential connectivity modes. Given our selected pharmacophores, six possible vertices emerge in the binding interaction graph: ha1-HD1, ha1-HD2, ha1-HD3, hp2-HD1, hp2-HD2, and hp2-HD3.

The subsequent step entails assessing the coexistence of any two vertices. For this purpose, we propose two unique compatibility criteria. The first, a τ flexible approach drawn from existing literature, assumes a uniform distance $\epsilon = 4 \text{ \AA}$ across all pharmacophores [22]. The second approach adjusts $\tau + 2\epsilon$ in the τ flexible approach to $\tau + \epsilon_1 + \epsilon_2$, assigning specific distances according to different pharmacophore types. For vertices composed exclusively of hydrogen-bond acceptors and donors, we adopt a shorter interaction distance, $\epsilon_s \text{ \AA}$. In all other cases, the interaction distance remains at $\epsilon \text{ \AA}$. The adjacency matrix corresponding to these two scenarios is illustrated in Figs. 2(a) and 2(c), respectively.

Before solving the maximum weighted clique problem, we need to add weights to the vertices in the binding interaction graph based on the pharmacophore potential.

TABLE I. Knowledge-based pharmacophore potential. Data is derived from the PDBbind dataset from Refs. [65–67]. HD, hydrogen-bond donor; HA, hydrogen-bond acceptor; HP, hydrophobe; AR, aromatic.

Ligand/protein	HD	HA	HP	AR
hd	0.5244	0.6686	0.1453	0.1091
ha	0.6686	0.5478	0.2317	0.0770
hp	0.1453	0.2317	0.0504	0.0795
ar	0.1091	0.0770	0.0795	0.1943

Table I shows the pharmacophore potential we applied in the numerical experiments. We use the same data as in Ref. [22], and these experience parameters are derived from the PDBbind dataset [65–67].

In a scenario where the interaction distance between pharmacophore points remains invariant, the resulting

pharmacophore model is represented by a fully connected adjacency matrix. This is reflected in the maximum vertex weight clique, which encompasses all vertices in such a matrix. The corresponding docking results, under these conditions, are illustrated in Fig. 2(b). However, the situation becomes more complex when the interaction distances vary, as the adjacency matrix alone no longer provides immediate and clear insights.

To address this issue, we construct the Hamiltonian for the modified adjacency matrix, as outlined in Eq. (3), and apply both conventional QAOA and its enhanced variant, DCQAOA, to find solutions. For detailed data on the construction of LDGs and BIG, refer to Appendix B.

For this study, we adopt a three-layer QAOA circuit and configured with a penalty parameter set to $P = 6$, which has been empirically determined to yield optimal results. A detailed illustration of the quantum circuit is

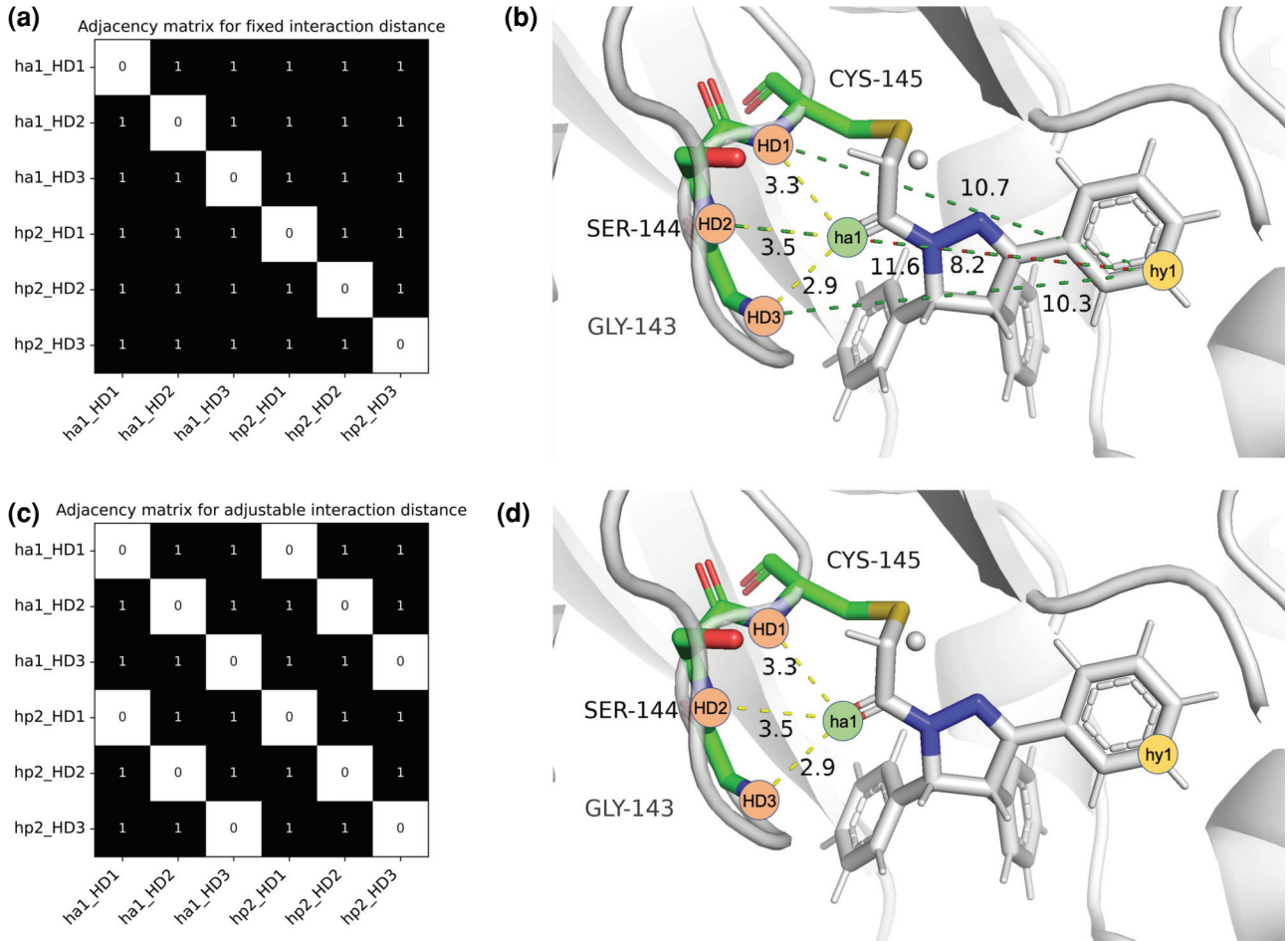


FIG. 2. The adjacency matrix corresponding to (a) a fixed distance ($\epsilon = 4 \text{ \AA}$) between all pharmacophores in the docking process. (b) illustrates the co-structure of SARS-CoV-2 M^{pro} (COVID-19) with covalent pyrazoline-based inhibitors PM-2-020B (PDB ID: 8SKH), assuming a fixed distance (ϵ) of $4 \text{ \AA}$ between all pharmacophores in the docking process. (c) A flexibly setting the distance parameters for different pharmacophore types (hydrogen-bond interactions are set at $3 \text{ \AA}$, while other interactions are set at $4 \text{ \AA}$). (d) demonstrates the docking results obtained by flexibly setting the distance parameters for different open pharmacophore types. Figures (b) and (d) show the schematic representation of the molecular-docking results, generated using the open-source software PyMOL [64].

presented in Fig. 9. Furthermore, the optimization process has been significantly enhanced through the integration of a quantum natural gradient-descent algorithm. The intricate details of the algorithm, including its specific parameter configurations, are comprehensively delineated in Appendix A.

The experimental methodology employed in this study involved a comprehensive evaluation of 500 distinct initial parameters. As depicted in Fig. 3(d), a meticulous analysis of the loss curve revealed that these randomly selected parameters exerted minimal impact on the experimental outcomes. Moreover, the study delved into examining the effects of varying parameters on the probability of achieving a successful outcome. To facilitate this analysis, for each initial parameters, the quantum circuit was reconfigured using a set of optimal parameters identified through optimization. Subsequently, for the final state of each quantum circuit, we sampled 5000 times under the $Z^{\otimes N}$ basis. The success probability for the state that corresponds to the maximum weighted clique is systematically illustrated in Fig. 4(a).

The docking results, both algorithms exhibit a remarkable consistency, particularly those involving the interactions ha1-HD1, ha1-HD2, and ha1-HD3, are depicted in Fig. 2(d). It is worth noting that we find that DCQAOA needs only 80 epochs to achieve a good loss, while the QAOA algorithm needs no less than 200 iterations under the same conditions, which proves that DCQAOA can effectively reduce the classical and quantum resources required for iterations.

Notably, when these results are juxtaposed against those obtained from the protein-ligand interaction profiler (PLIP) as referenced in Ref. [68], it is evident that our calculated docking results are more congruent with established biological principles. This contrast is especially marked when compared to the results derived under the condition of fixed interaction distances between pharmacophore points. This observation underscores the significance of appropriately modulating the interaction distances between pharmacophore points in order to obtain docking results that are more in harmony with biological laws.

B. Exploring the optimal parameter settings for the QAOA-based algorithm

Next, we begin a detailed analysis assessing the impact of various QAOA parameters on the docking process. We chose an eight-qubit system for its more complex solution space compared to a six-qubit system, while maintaining effective simulation across multiple trials.

We focus on the structure of dipeptidyl peptidase-4 (DPP-4) complexed with piperidine-fused imidazopyridine 34 (PDB ID: 3HAC). Piperidine-fused imidazopyridine 34 has been identified as a potent inhibitor of DPP-4 and shows promise in enhancing glycemic control [69].

Understanding its interactions with DPP-4 is vital for elucidating the drug's mechanism of action, optimizing its structure for increased efficacy, and potentially revealing alternative treatments for diabetes. This research is pivotal in advancing more effective therapies for diabetes management.

Unlike the previous example with six qubits, this example concentrates on a docking problem involving an aromatic ring. We selected two pharmacophore points on the piperidine fused imidazopyridine 34 ligand: N12298 (hd1), and the pseudoatoms at the center of the aromatic ring (ar2, the central coordinates of six atoms: C12301, C12302, C12306, N12307, C12308, and C12309). We also selected four pharmacophore points in DPP-4: O1453 (HA1), O1462 (HA2), C5171 (HA3), and the pseudoatoms in the center of the aromatic ring (AR4, the central coordinates of six atoms: C2683, C2684, C2685, C2686, C2687, and C2688). We then constructed the binding interaction graph, the adjacency matrix of which is depicted in Fig. 13(a) (Detailed calculations on LDGs and BIG can be found in Appendix C.)

In this example, we also implemented both the conventional QAOA and DCQAOA, and the quantum circuit diagram is shown in Fig. 12 in Appendix C. This investigation rigorously examines the disparities between these two algorithms and the impact of various parameters on the performance of algorithms. Key influencing factors considered include the layers p , the magnitude of the penalty term P , and the performance of the algorithm within the noise model.

To rigorously evaluate the efficacy of our algorithm, we adopted two distinct metrics. The first metric, designated as P_{true} , quantifies the success ratio of the algorithm. Specifically, it measures the likelihood that the postsampling string corresponding to the highest probability state aligns with the target state $\langle \psi_{\text{true}} \rangle$. The second metric assesses the overlap between the target state and the ground state ψ_g of problem Hamiltonian H_c , expressed as $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$.

Our analysis initially focused on evaluating the impact of the number of layers within QAOA and DCQAOA, on their respective performance results. We noted that both P_{true} and $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$ initially rose and then fell as more layers were added to the quantum circuit, as shown in Fig. 5(a). This trend is in line with theoretical predictions, suggesting that adding layers might mimic adiabatic quantum computing, improving results. However, too many layers add complexity and increase the risk of hitting local minima in the optimization process. Our findings imply that a balanced approach—increasing layers linearly with system size, following an $O(n)$ scale—may be most effective. This strategy balances efficiency and effectiveness, reducing the risks of too many layers. Thus, it appears as a valuable method for enhancing quantum algorithm performance in computational tasks.

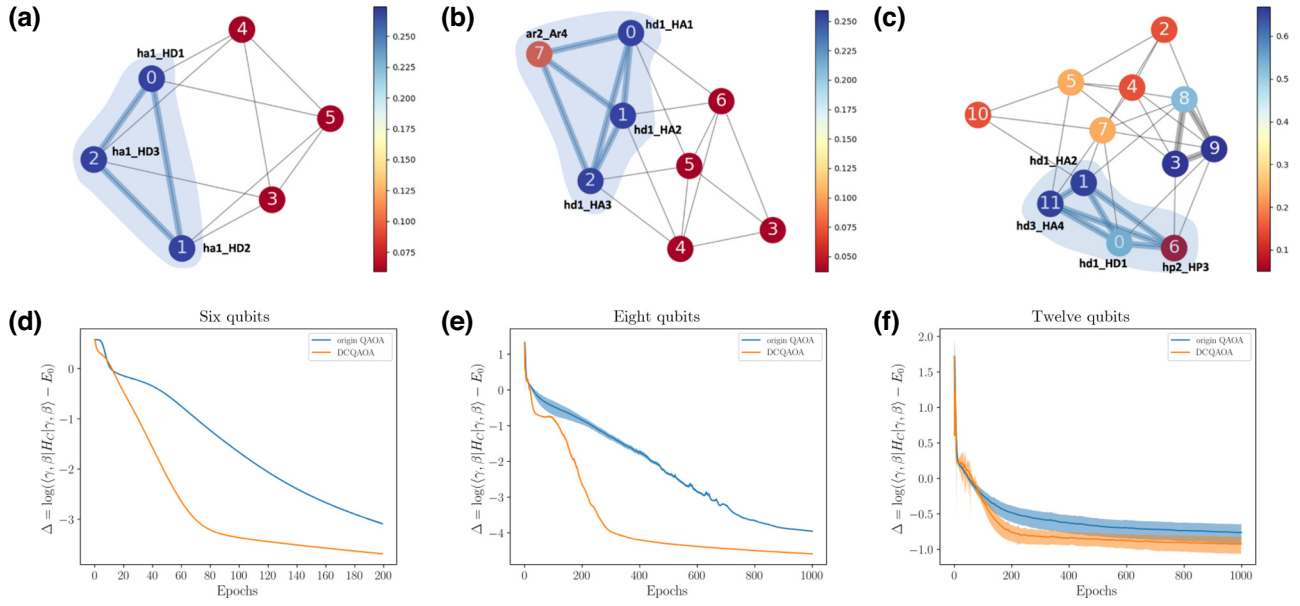


FIG. 3. In Panels (a)–(c), we present the results for three protein-ligand docking example form QAOA-based algorithms, specifically for PDB IDs: 8SKH (6 qubits), 3HAC (8 qubits), and 5F4L (12 qubits). Panels (d)–(f) illustrate the logarithmic disparity between the ground state energies determined by both QAOA approaches and those ascertained through precise diagonalization of the Hamiltonian, corresponding to the three protein-ligand docking cases. For each case, we generated and optimized 500 unique sets of initial parameters to demonstrate variance, denoted by a shaded region with 50% opacity.

This analysis categorically establishes the differences in resource allocation between DCQAOA and conventional QAOA, focusing particularly on gate requirements. Of note, integrating selected CD terms into DCQAOA does not require extra complex two-qubit gates, thus bolstering its experimental applicability. At $p = 8$ layers, DCQAOA reaches its optimal efficacy, employing only 200 single-qubit and 80 two-qubit parametric gates. Conversely, conventional QAOA attains its peak performance at a significantly higher layer of $p = 24$, necessitating 392 single-qubit and 240 two-qubit parametric gates. Moreover, the fidelity metric $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$ in QAOA achieves

a maximum near 0.8, reflecting its reduced effectiveness. Notably, the maximum iterations in our simulation are limited to 500, as DCQAOA achieves satisfactory results within this range. This comparative study illustrates that conventional QAOA requires twice as many single-qubit and 3 times as many two-qubit parametric gates as DCQAOA. This pronounced disparity in gate utilization highlights not only superior efficiency of DCQAOA but also suggests a heightened susceptibility of conventional QAOA to noise disturbances in quantum systems, attributed to its greater dependence on quantum gate complexity.

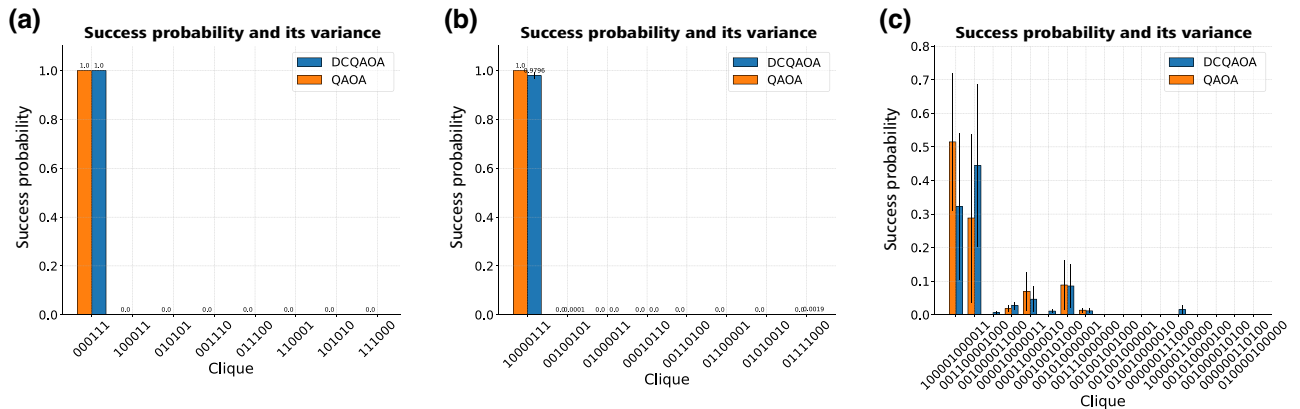


FIG. 4. The success probability of docking for complex with PDB IDs: 8SKH (6 qubits), 3HAC (8 qubits), and 5F4L (12 qubits). For each case, we generated and optimized 500 unique sets of initial parameters to demonstrate variance, denoted by vertical line.

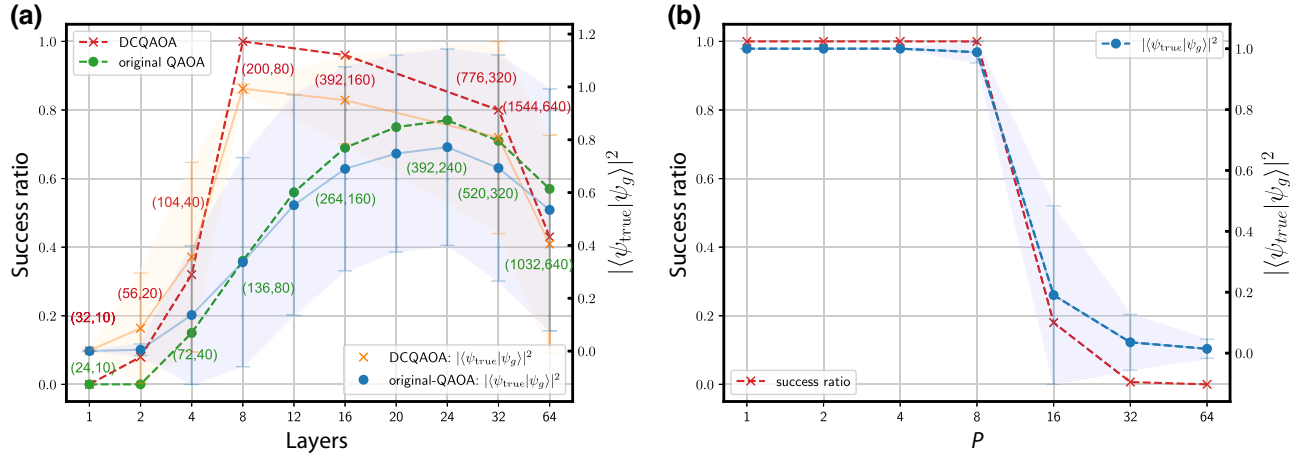


FIG. 5. Comparative analysis of the original QAOA and DCQAOA performance with varying numbers of layers and penalty term. The left panel delineates the success ratio P_{true} and the expected probability amplitude $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$ as a function of the quantum circuit layers, depicted on the x axis. The success ratio P_{true} for the original QAOA is denoted by a green dotted line with circular nodes, whereas the DCQAOA's is represented by a red dotted line with "x" nodes. The expected probability amplitude $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$ for the original QAOA is denoted by a solid yellow line, whereas the DCQAOA's is represented by a solid blue line, and the shaded regions reflect the statistical error margins computed from a set of 100 samples. The annotated numerical pairs (red for DCQAOA and green for QAOA) enumerate the quantum gate counts at varied layer depths, with the initial numeral indicating single-qubit gate counts and the trailing numeral specifying two-qubit gate counts. The right graph offers a focused view on the diminishing trend of success ratios P_{true} and probability amplitudes $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$ for DCQAOA as the penalty term P increases.

In the second analysis, we investigated the impact of varying the penalty term, P , on docking results, as illustrated in Fig. 5(b). Adjustments in P to values 1, 2, 4, or 8 yielded no significant alteration in success probability. However, when P values exceed 8, a significant decrease is evident in both P_{true} and $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$. This decline can be attributed to the increased penalty term, which diminishes the contribution of the dominant term in Eq. (3). As a result, it obscures the differentiation between optimal and suboptimal outcomes, consequently increasing the probability of errors.

Next, we investigated the impact of noise models on the effectiveness of docking. We introduced varying levels

of depolarizing noise into the quantum circuit, as illustrated in Fig. 6. Our experimental configuration utilized an eight-layer quantum circuit with an associated penalty term set to eight. The presented results are the product of an extensive series of 500 trials for each data point, with randomized parameter initialization in each trial. Our analysis reveals a clear decrease in both the success rate of docking and the probability amplitude, directly correlated with the increased intensity of noise.

Based on the analyses conducted above, we have employed a DCQAOA circuit featuring an eight-layer architecture and the inclusion of a penalty term set at $P = 8$. This configuration underwent a comprehensive

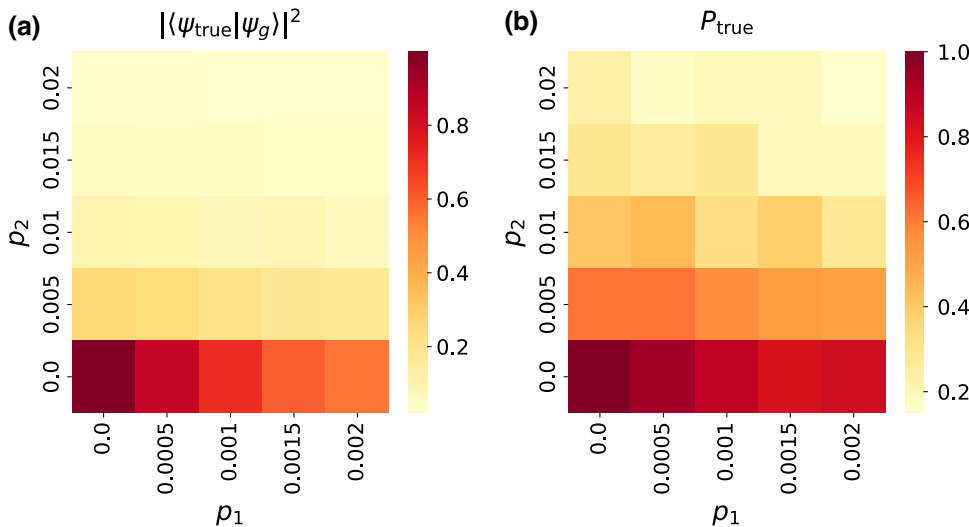


FIG. 6. Panel (a) and (b) represents the performance of DCQAOA across different noise level, where p_1 and p_2 refer to the depolarizing noise level of single and two qubit gates, respectively. Panel (a) shows the average overlap between the optimal states and DCQAOA output states, i.e., $|\langle \psi_{\text{true}} | \psi_g \rangle|^2$, in each block. Panel (b) the average success ratio of finding optimal solution over 500 runs.

evaluation encompassing the initialization of 500 distinct parameter sets selected randomly. The performance of each set was meticulously monitored, and the resulting loss curves are depicted in Fig. 3(e). Remarkably, these curves manifest a significant reduction in loss error when employing the DCQAOA. Furthermore, the DCQAOA enables us to attain the desired outcomes within a mere 200 iterations, marking a substantial departure from the requirement of 1000 iterations for the conventional QAOA. This decrease in the iteration count translates into a noteworthy reduction in classical and quantum resource consumption during the optimization process. Subsequently, we undertake a comprehensive analysis to assess the success probability, as illustrated in Fig. 4(b). In contrast to the approach adopted in Fig. 5, we established distinct iteration limits for DCQAOA and QAOA: the QAOA was granted a maximum of 1500 iterations, whereas the DCQAOA was constrained to a ceiling of 500 iterations. The DCQAOA algorithm exhibited an exceptionally high accuracy rate, approaching 100%, while the QAOA algorithm, although slightly less effective, still demonstrated considerable robustness in its results.

The docking outcomes, namely hd1-HA1, hd1-HA2, hd1-HA3, and ar2-AR4, have been illustrated in Fig. 13(b). It is worth highlighting that a comparative analysis performed in conjunction with data obtained from PLIP [68] demonstrates a notable congruence between our docking results and well-established biological principles. This concurrence not only fortifies the soundness of our methodology but also accentuates its potential to yield valuable insights within the domain of computational biology. All computational simulations were executed using the Mindquantum framework [70].

C. More qubits

In this section, we investigate the utilization of the QAOA-based molecular-docking algorithm in expanded quantum systems of 12 qubits.

We focus on the complex of HIV-1 gp120 and JP-III-048 (PDB ID: 5F4L), a key molecular interaction in HIV research with potential therapeutic significance. The glycoprotein HIV-1 gp120, a critical part of the human immunodeficiency virus type 1 envelope, is essential in attaching to the CD4 receptor of host cells, aiding in viral entry. JP-III-048, a compact synthetic molecule, intervenes in this process by adhering to gp120, consequently impeding the virus's capacity to infect cells and proliferate.

In this study, we have strategically identified three pharmacophore points on JP-III-048, denoted as N5302 (hd1), C5281 (hp2), and N5297 (hd3), and four distinct points on HIV-1 gp120, namely N2206 (HD1), 02209 (HA2), C5301 (HP3), and N2479 (HA4). Subsequently, we have meticulously constructed two ligand descriptors (LDGs),

one for the JP-III-048 ligand and another for the HIV-1 gp120 receptor. The relative distances between these pharmacophores have been vividly illustrated in Fig. 10. Furthermore, we have embarked upon the development of the BIG, the detailed construction of which can be found in Appendix D, with its corresponding adjacency matrix depicted in Fig. 17(a).

To commence our analysis, we have initially applied a classical algorithm to the BIG framework, aiming to identify and rank cliques consisting of three or more vertices based on their respective weights. The graphical representation of this data is presented in Fig. 15. It is worth noting that a remarkable observation emerged from this analysis; the weight differential between the largest and second-largest clique was found to be quite marginal, approximately at an order of magnitude of merely 10^{-2} . This subtle variation presents a noteworthy challenge to the effectiveness of our numerical algorithm.

Drawing upon insights derived from prior research, we have meticulously refined key parameters within our current model. These adjustments encompassed the calibration of the penalty term, denoted as P , and the configuration of the optimization layer. In particular, for this experimental setup, we selected a configuration with 13 layers, while rigorously fixing the penalty term at $P = 8$. The optimization process employed the quantum natural gradient-descent algorithm, and comprehensive details of the experimental circuit can be found in Appendix D.

We conducted a thorough evaluation of both the DCQAOA and the QAOA by subjecting them to simulations across 500 distinct sets of initial parameters. The results of these simulations, specifically the trends in loss functions and success probabilities, are presented graphically in Figs. 3(f) and 4(c). An in-depth analysis of the loss functions reveals a significant enhancement in the algorithm's efficiency in converging towards the ground state through the incorporation of the CD term. Quantitatively, we observed that the average loss for DCQAOA fell within the order of 10^{-1} , as opposed to an average loss of $10^{-1/2}$ for the standard QAOA. This disparity is even more conspicuous in terms of success probability: the DCQAOA algorithm achieved a noteworthy accuracy rate of 51%, whereas the conventional QAOA algorithm exhibited a considerably lower success rate of 32%, indicating a greater predisposition to converge towards suboptimal solutions. These findings provide compelling evidence that the efficacy of these algorithms is potentially influenced by the number of layers chosen and the precise implementation details of the algorithm. Nevertheless, one notable caveat of our approach is the algorithm's vulnerability to converging towards local optima within the vast parameter space. This aspect of the algorithm's performance necessitates further investigation to enhance its optimization capabilities.

IV. CONCLUSION

This work illuminates the potential of quantum computing, specifically the conventional QAOA and its variant, the digitized-counterdiabatic QAOA, in tackling the intricate problem of molecular docking. By transforming the molecular-docking problem into the maximum vertex weight clique problem, we demonstrate that quantum computing offers an innovative and efficient methodology for exploring the vast search space of protein-ligand interactions. Our investigation of various biological systems, including the SARS-CoV-2 M^{pro} complex with PM-2-020B, the DPP-4 complex with piperidine fused imidazopyridine 34, and the HIV-1 gp120 complex with JP-III-048, underscores the adaptability of this approach. The results obtained through QAOA and DCQAOA are in close alignment with biological principles, suggesting that quantum computing can provide invaluable insights into the design and optimization of therapeutic drugs.

Our research presents a transformative paradigm in constructing protein-ligand interaction graphs, emphasizing the pivotal role of pharmacophore interactions. This methodology has yielded consistently more accurate docking results, in alignment with established biological principles. In the realm of algorithmic development, our findings robustly demonstrate that the DCQAOA outperforms the conventional QAOA in identifying the ground states of problem Hamiltonians. This superiority is evident in two key aspects: DCQAOA exhibits higher accuracy at equivalent circuit depths and requires fewer iterations under the same parameter settings. The enhancement is attributed to the incorporation of the CD term in DCQAOA, which effectively minimizes transitions to excited states, thus enhancing the precision of ground-state identification. Moreover, in the current era of NISQ technologies, characterized by prevalent quantum noise, the lower quantum gate count of DCQAOA is particularly significant. This reduced gate count in DCQAOA facilitates the application of existing quantum devices to larger and more complex molecular systems, a crucial advancement given the challenges posed by quantum noise.

Notwithstanding the promising outcomes achieved, our investigation into a 12-qubit system with the QAOA-based algorithm has encountered substantial challenges. These impediments are likely attributed to the exponential increase in optimization parameters, a complexity that potentially impairs the efficacy of our optimization algorithm. Notably, even the integration of CD terms does not entirely mitigate this issue. The intrinsic optimization challenge within QAOA remains a pivotal area for further exploration, as alternative optimization strategies could be instrumental in transcending local optima.

In recent years, the field of quantum computing has experienced significant advancements, heralding a potential shift beyond the NISQ era. Remarkable progress

has been made in a variety of quantum systems, including superconducting [71], trapped-ion [72], neutral atomic [73], and optical quantum systems [74]. These advancements are highly significant as they facilitate the manipulation of an increased number of qubits with enhanced precision, laying the groundwork for revolutionary breakthroughs across multiple scientific fields. For example, the Quantinuum H2 quantum processor, a testament to Honeywell's technological ingenuity, has achieved single-qubit gate fidelity of 99.997% and two-qubit gate fidelity of 99.98% [72]. Our numerical simulations substantiate the potential of quantum systems of this caliber to execute molecular docking procedures with superior precision, even within an eight-qubit quantum framework.

In our proposed quantum algorithm for the docking problem, we have incorporated multiple approximation methodologies to confront this multifaceted challenge. As advancements in quantum technology continue, a diminished reliance on these approximation methods is anticipated. This progression is poised to yield improvements in both the precision and efficiency of quantum algorithms, particularly as we approach the post-NISQ era. This forthcoming era undeniably demands sustained and pioneering advancements in the field of quantum algorithms and their accompanying software frameworks. This scenario underscores the critical need for interdisciplinary collaboration and ongoing research endeavors. The fusion of diverse expertise, combined with cutting-edge innovations, represents a pivotal element in fully harnessing the potential of quantum computing. Such integration of resources and knowledge is positioned at the vanguard of reshaping the paradigms within various scientific and industrial domains, ultimately exerting a profound influence on the future trajectory of these disciplines.

In conclusion, our study emphasizes the transformative potential of quantum computing in the realm of molecular docking, thereby enhancing our comprehension of molecular interactions. This progress has significant ramifications for the development of effective therapeutic strategies. We argue that incorporating quantum computing throughout the drug-discovery process presents a significant opportunity to initiate a paradigm shift in biomedical research.

ACKNOWLEDGMENTS

We would like to express our gratitude to Dr. Shang Yu for providing Supplemental Material for their related paper Ref. [23]. This work is supported by the National Natural Science Foundation of China Grants (No. 12175003 and No. 12361161602), NSAF (Grant No. U2330201). The scientific calculations in this paper have been done on the HPC Cloud Platform of Shandong University, and supported by the High-performance Computing Platform of Peking University.

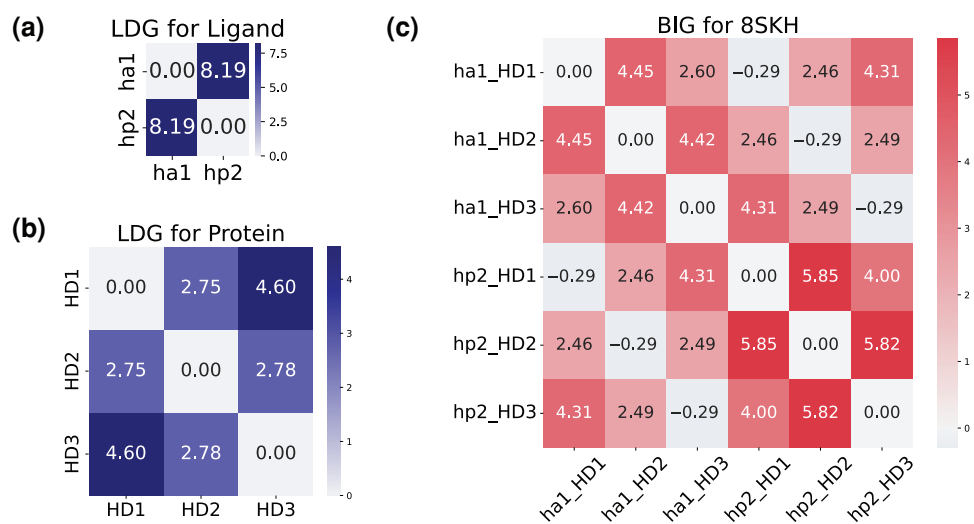


FIG. 7. Panels (a) and (b) present the distance data for the protein and ligand, respectively. Panel (c) illustrates the original data for the BIG of 8SKH, highlighting two pharmacophore points on PM-2-020B [O2378 (ha1) and C2373 (hp1)] and three pharmacophore points on SARS-CoV-2 M^{pro} [N1104 (HD1), N1108 (HD2), and N1114 (HD3)].

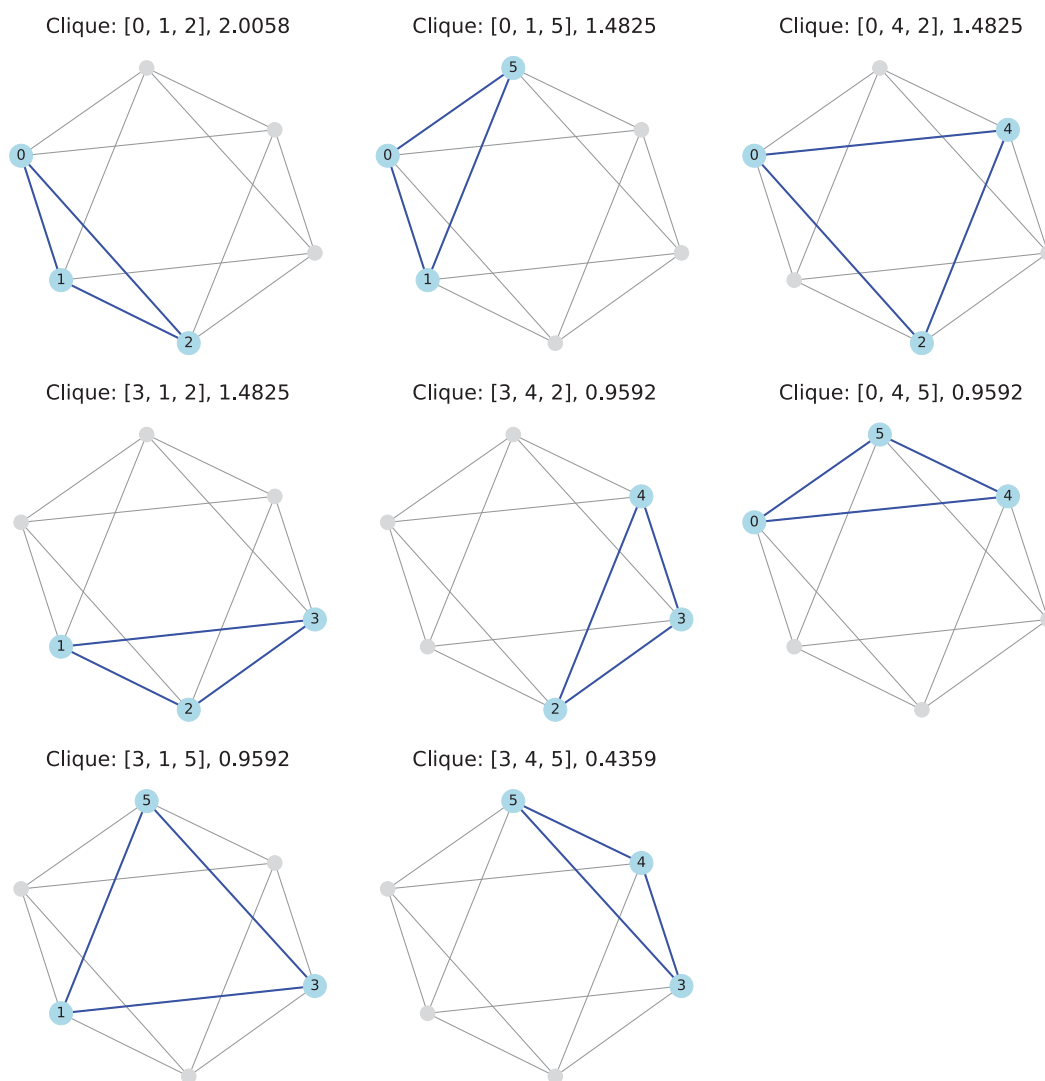


FIG. 8. All cliques with a greater number of vertices than 3 in BIG for 8SKH.

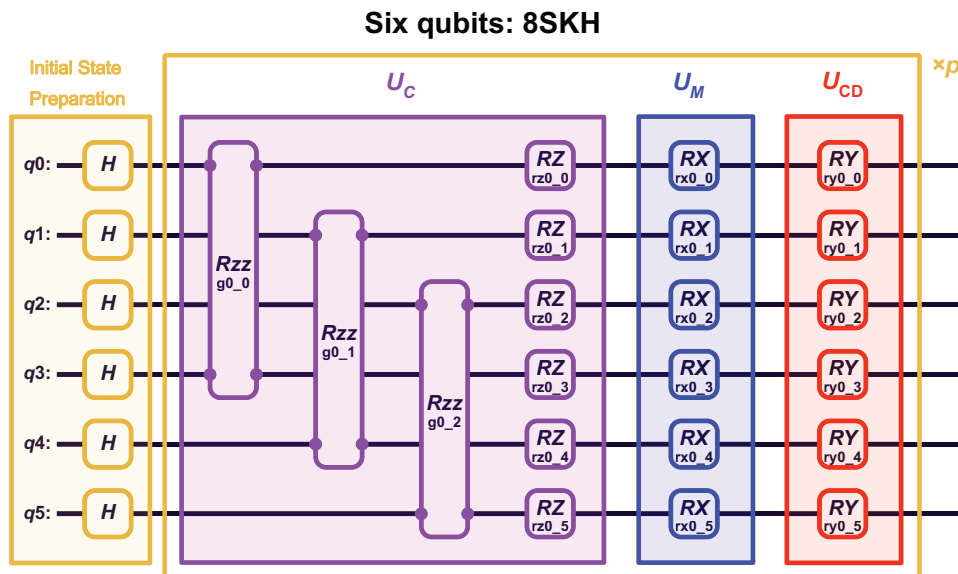


FIG. 9. The quantum circuit configuration for the QAOA-based algorithm for 8SKH. The quantum circuit is partitioned into distinct segments, each responsible for a separate phase of the algorithm’s execution. Initial quantum states are engineered into coherent superpositions through the application of the Hadamard gate (H) to each qubit. Following the state preparation, the circuit enters the control unitary (U_C) phase, characterized by the application of entangling two-qubit ZZ rotations (R_{zz}) and single-qubit Z rotations (R_z). Subsequent to U_C , the mixing unitary (U_M) is implemented using single-qubit X rotations (R_x) for each qubit. The circuit culminates with the controlled-drive unitary (U_{CD}), which consists of single-qubit Y rotations (R_y), whereas QAOA does not have U_{CD} part. The entire sequence encapsulates “ p ” iterations.

APPENDIX A: QUANTUM NATURAL GRADIENT-DESCENT ALGORITHM

As the natural gradient adjusts the learning rates for different parameters based on the curvature of the cost

function, it enables more efficient and faster convergence during training. The idea behind the natural gradient algorithm is to compute gradients by mapping the parameter to a Riemannian manifold on the parameter space

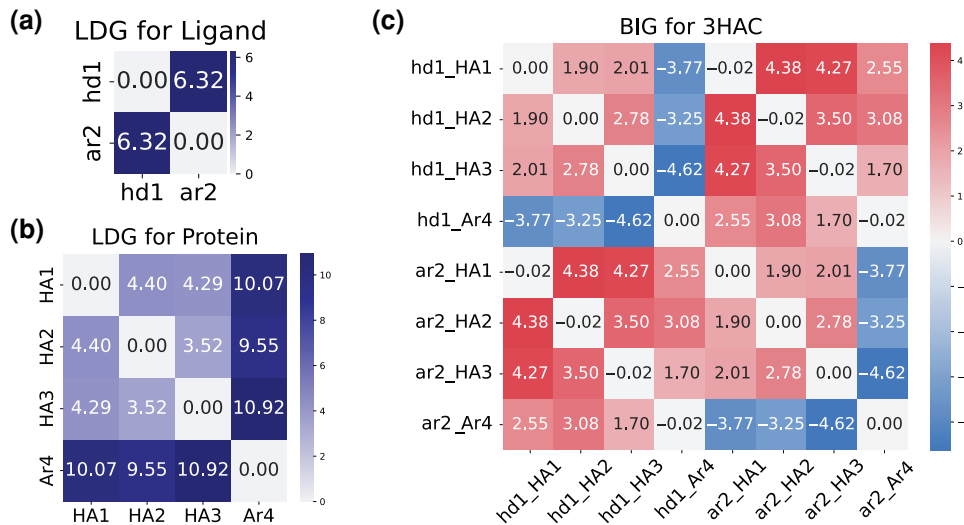


FIG. 10. Panels (a) and (b) present the distance data for the protein and ligand, respectively. Panel (c) illustrates the original data for the BIG of 8SKH, highlighting two pharmacophore points on the piperidine fused imidazopyridine 34 ligand: N12298 ($hd1$), and the pseudoatoms at the center of the aromatic ring ($ar2$, the central coordinates of six atoms: C12301, C12302, C12306, N12307, C12308, and C12309). We also selected four pharmacophore points in DPP-4: 01453 ($HA1$), 01462 ($HA2$), C5171 ($HA3$), and the pseudoatoms in the center of the aromatic ring ($Ar4$, the central coordinates of six atoms: C2683, C2684, C2685, C2686, C2687, and C2688).

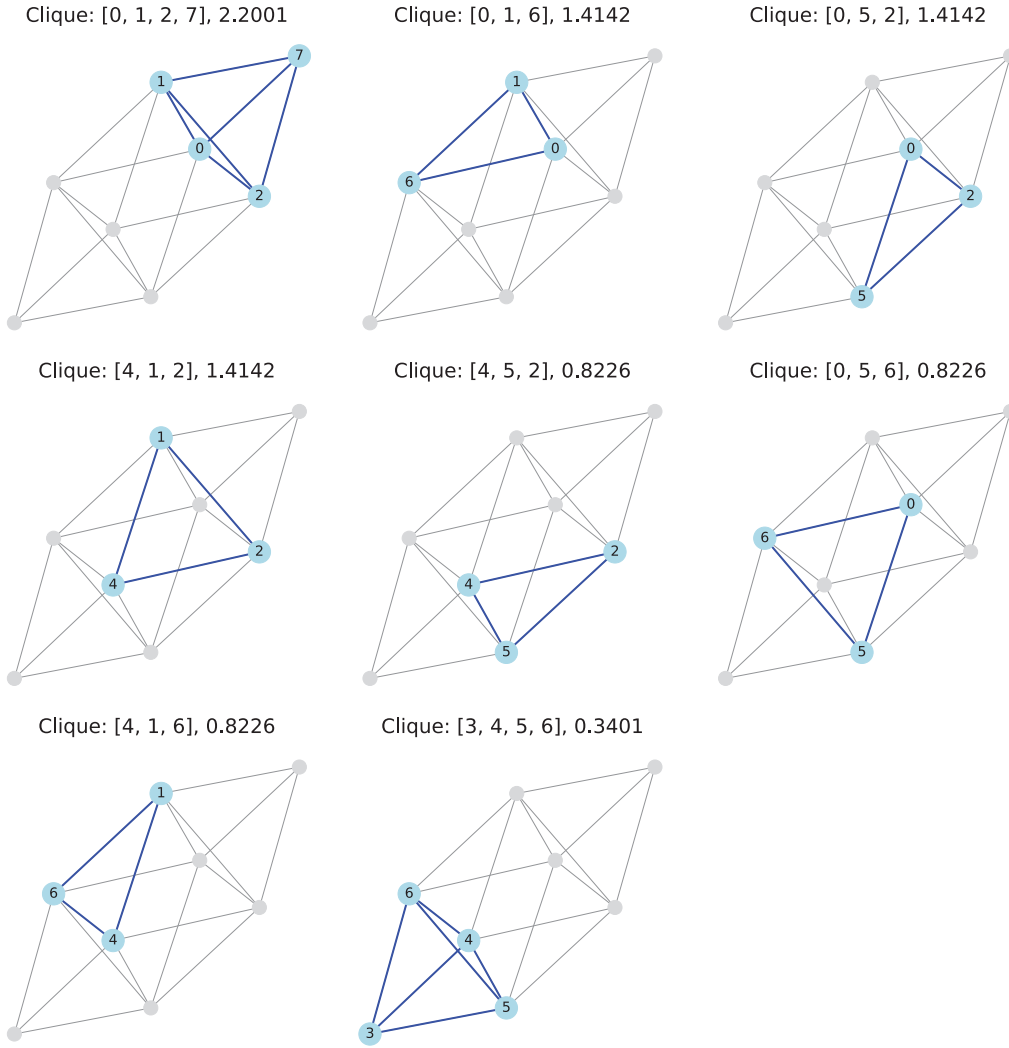


FIG. 11. Illustration of all cliques exceeding three vertices in the BIG for 3HAC.

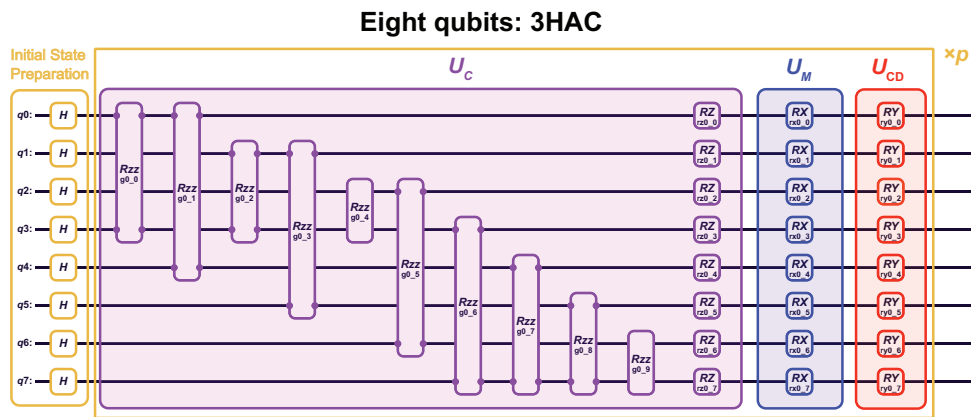


FIG. 12. Quantum circuit configuration for the QAOA-based algorithms applied to 3HAC. This circuit is segmented into distinct phases, each integral to the algorithm's operation. Initial quantum states are configured into coherent superpositions via the Hadamard gate (H) application on each qubit. Subsequently, the circuit progresses to the control unitary (U_C) phase, marked by entangling two-qubit ZZ rotations (R_{ZZ}) and single-qubit Z rotations (R_Z). Following U_C , the mixing unitary (U_M) is executed, employing single-qubit X rotations (R_X) for each qubit. The circuit concludes with the controlled-drive unitary (U_{CD}), comprising single-qubit Y rotations (R_Y), a component not present in the standard QAOA. This entire process encompasses “ p ” iterations.

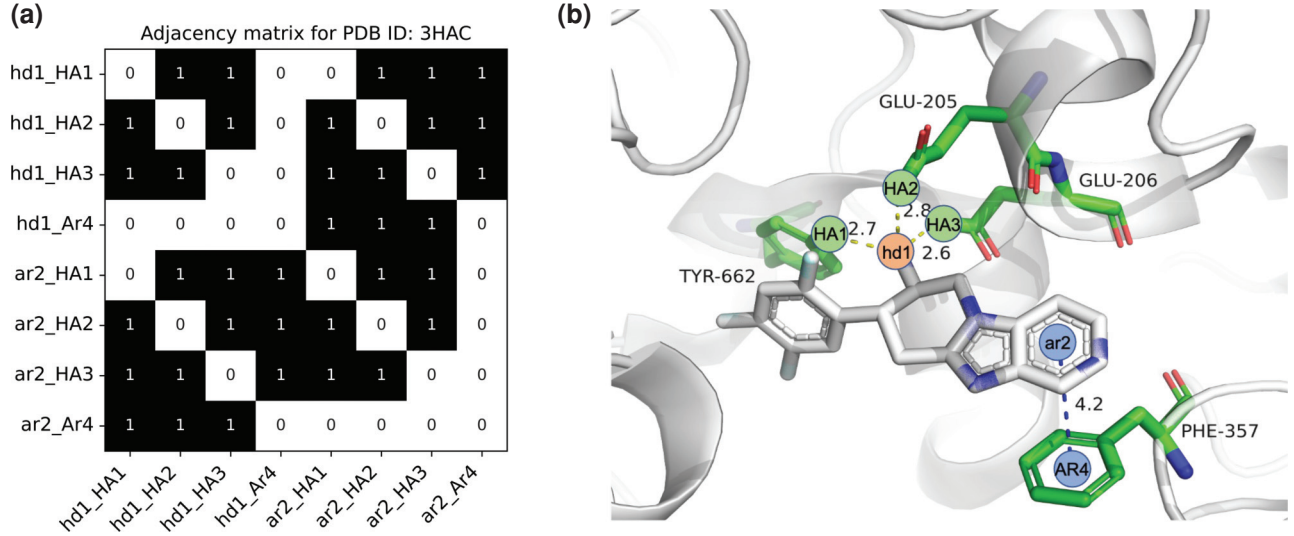


FIG. 13. Panel (a) depicts the adjacency matrix of the binding interaction graph for DPP-4 in complex with piperidine fused imidazopyridine 34 (PDB ID: 3HAC). Panel (b) presents the predicted protein-ligand docking visualization, showcasing interacting amino acid residues on the protein (green carbon atoms) and the ligand (white carbon atoms). Notably, the yellow bonds highlight crucial predicted interactions, namely hd1-HA1, hd1-HA1, hd1-HA1, and ar2-AR4.

instead of Euclidean space, where each point corresponds to a parameter vector. In that space, a metric matrix can be defined, reflecting the geometric structure of the parameter space. Unlike gradient-descent algorithms, the natural gradient algorithm differs by scaling the gradient with the Fisher information matrix, instead of $\mathbb{I}$, as the metric matrix $\mathbf{M}$. Thereby, it provides better control over the updating direction of model parameters. For efficiently training the model, we utilize the quantum natural gradient descent as the optimizer, which leverages the quantum Fisher information matrix to adjust learning rates for variational parameters in quantum circuits [75]. The parameter update rule of θ in i th iteration shows as follows:

$$\theta^{(i+1)} = \theta^{(i)} - \eta \cdot \mathbf{M}^{-1} \mathbf{C}, \quad (\text{A1})$$

where $\mathbf{C} \in \mathbb{C}^n$, $\mathbf{M} \in \mathbb{C}^{n \times n}$

$$\mathbf{C}_j = \partial_{\theta_j} \langle H \rangle_{\theta}, \quad (\text{A2})$$

$$\mathbf{M}_{jk} = \mathcal{R}[\langle \partial_{\theta_j} \psi | \partial_{\theta_k} \psi \rangle - \langle \partial_{\theta_j} \psi | \psi \rangle \langle \psi | \partial_{\theta_k} \psi \rangle]. \quad (\text{A3})$$

Here, unlike the matrix $\mathbf{M}$ is associated with the Fisher information matrix in natural gradient descent, it is related to the quantum geometric tensor (QGT), which is the Fubini-Study metric on quantum states in such a quantum counterpart. We perform all the numerical experiments of DCQAOA for molecular docking under $\eta = 0.1$ and randomly sample the gate parameters from a uniform distribution, $\theta \sim U(0, 1)$ to initialize the circuit.

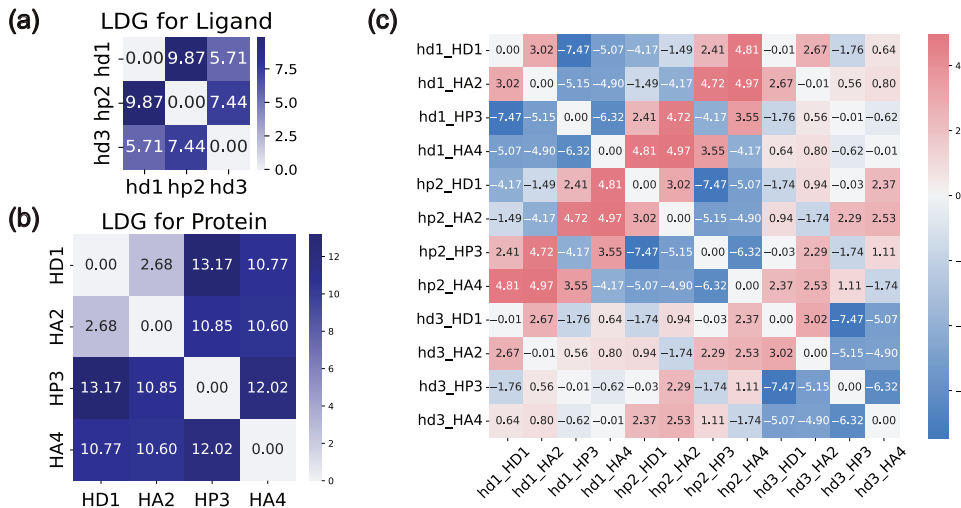


FIG. 14. Panels (a) and (b) present the distance data for the protein and ligand, respectively. Panel (c) illustrates the original data for the BIG of 8SKH, highlighting three pharmacophore points on JP-III-048: N5302 (hd1), C5281 (hp2), and N5297 (hd3), and four points on HIV-1 gp120: N2206 (HD1), 02209 (HA2), C5301 (HP3), and N2479 (HA4).

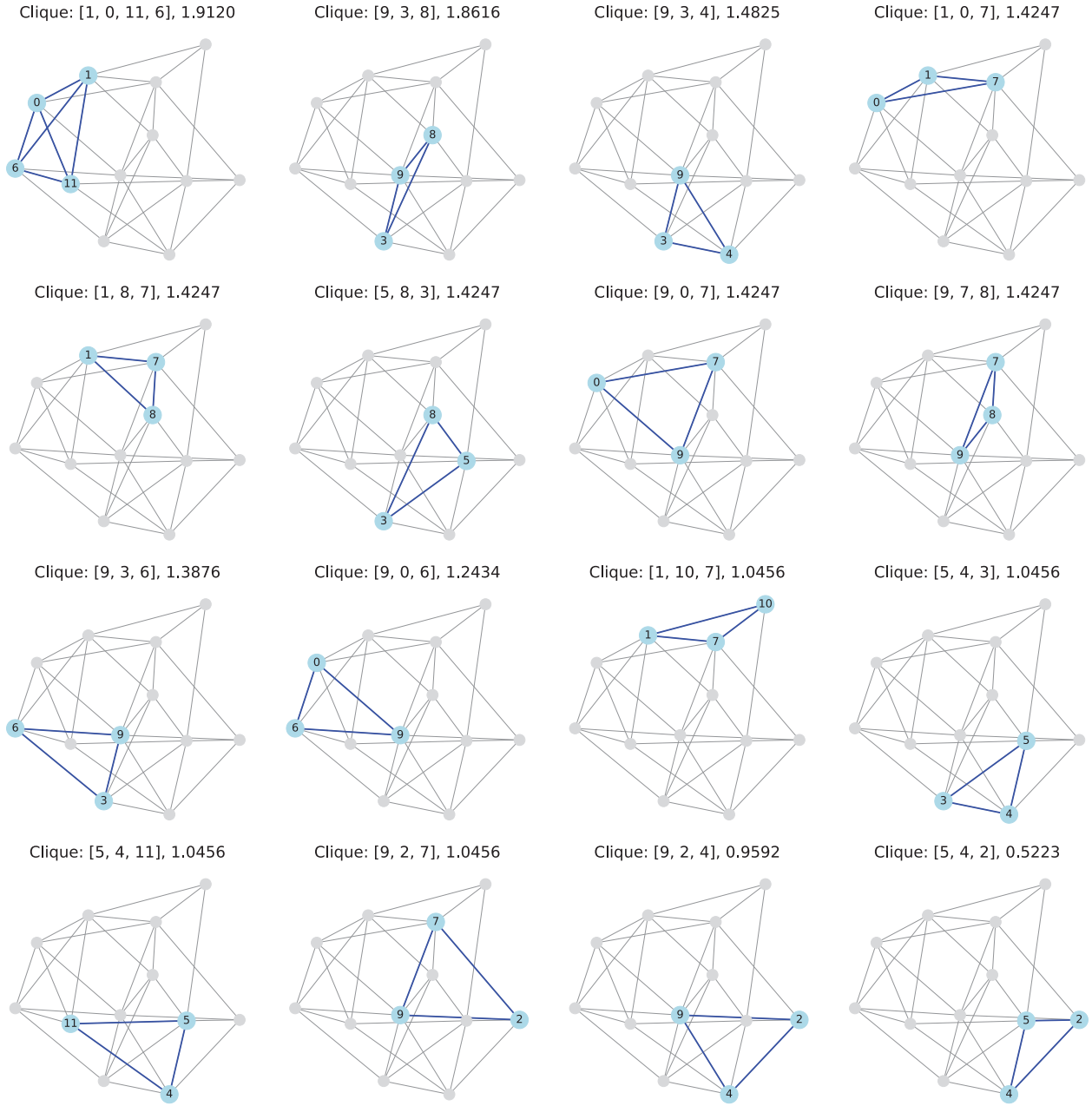


FIG. 15. Visualization of all cliques in BIG for 5F4L, encompassing vertices exceeding a count of three.

APPENDIX B: 8SKH

In this investigation, we enhance the τ flexible approach introduced in Ref. [22] by adjusting the parameter $\tau + 2\epsilon$ to $\tau + \epsilon_1 + \epsilon_2$. This adaptation permits the customization of specific interaction distances tailored to distinct pharmacophore categories. Precisely, for vertices composed exclusively of hydrogen-bond acceptors and donors, we employ a shorter interaction distance, denoted as ϵ_s Å. Conversely, for all other scenarios, we maintain the standard interaction distance of ϵ Å.

The determination of the connectivity within the BIG is calculated utilizing the subsequent equation:

$$In_{\text{str}A-\text{str}B} = (\tau + \epsilon_A + \epsilon_B) - |d_1 - d_2|. \quad (\text{B1})$$

In this study, the variable $In_{\text{str}A-\text{str}B}$ denotes the connectivity strength of a given edge within the BIG. This term is determined based on the ratio of $\text{str}A$ and $\text{str}B$, which represent the names of the two vertices forming an edge. Additionally, d_1 and d_2 are defined as the Euclidean distances of the ligand and the protein, respectively, located at the vertices of the edge in question.

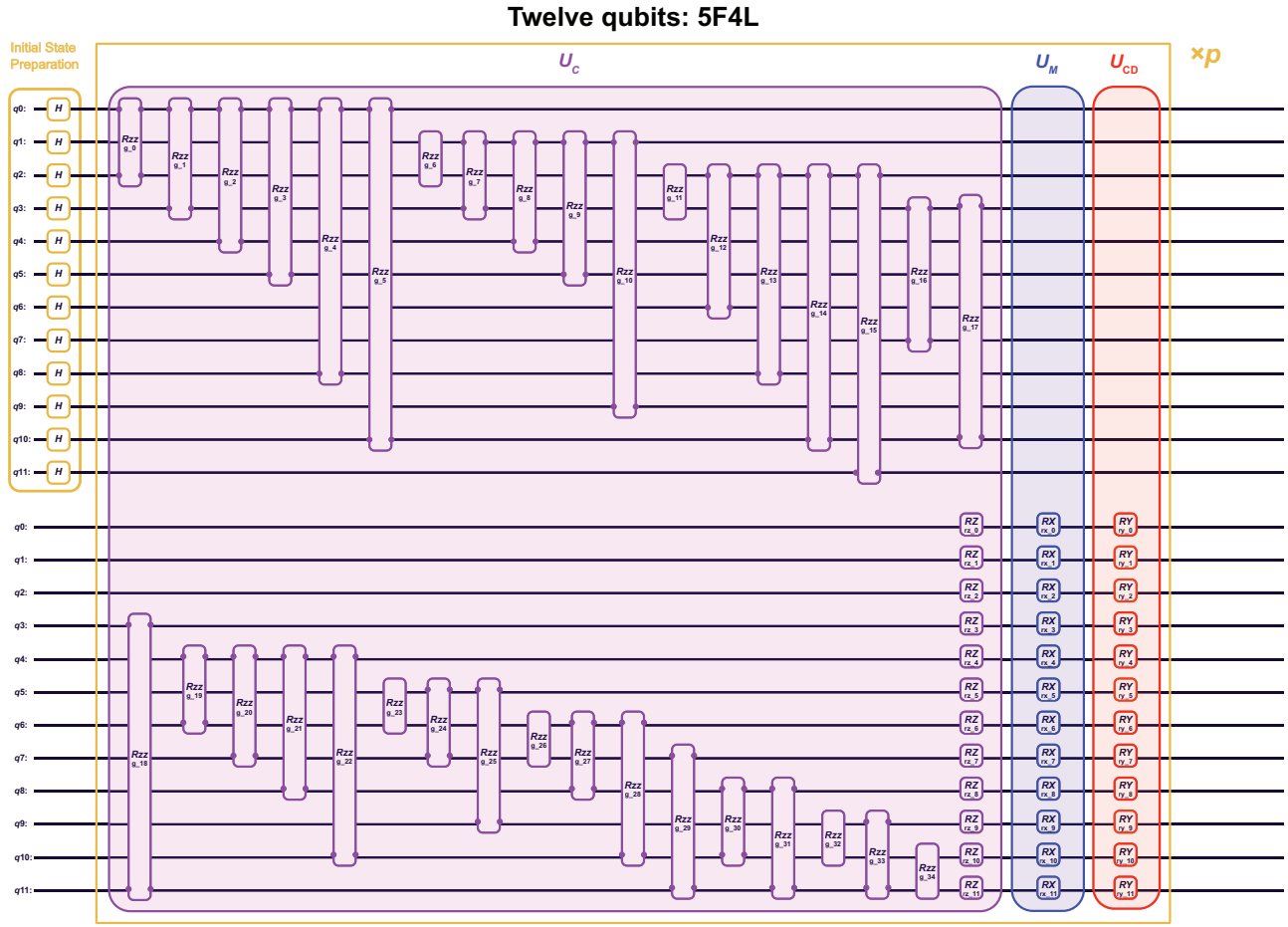


FIG. 16. DCQAOA quantum circuit configuration for 5F4L, segmented into distinct operational phases. The circuit initiates with the generation of initial quantum states, transforming each qubit into a coherent superposition via the Hadamard gate (H). The ensuing phase involves the control unitary (U_C), characterized by a series of entangling two-qubit ZZ rotations (R_{ZZ}) and single-qubit Z rotations (RZ). Following U_C , the mixing unitary (U_M) is executed, employing single-qubit X rotations (RX) for each qubit. The circuit culmination is marked by the implementation of the controlled-drive unitary (U_{CD}), constituted by single-qubit Y rotations (RY), a segment absent in the standard QAOA circuit. This sequence encompasses a total of “ p ” iterations.

Our proposed methodology involves a matrix representation, wherein each element corresponding to a pair of vertices is assigned a value based on the connectivity strength. Specifically, an element is set to 1 if $In_{strA-strB} > 0$, indicating the presence of an edge between the vertices. Conversely, an element is assigned a value of 0 if $In_{strA-strB} \leq 0$, signifying the absence of a connection. It is crucial to note that all diagonal elements of this matrix are set to 0. This specification is deliberately made to preclude the counting of self-interactions, thereby maintaining the integrity and accuracy of the interaction analysis in the BIG framework.

In the six-qubit example, considering the elements ha1-HD1, ha1-HD2, ha1-HD3, hp2-HD1, hp2-HD2, and hp2-HD3, and selecting values of $\epsilon = 4 \text{ \AA}$, $\epsilon_s = 3 \text{ \AA}$, and $\tau = 0.6 \text{ \AA}$, we can calculate the values for BIG. These values are visually represented in Fig. 7, and the corresponding adjacency matrix is illustrated in Fig. 2(c).

Simultaneously, we comprehensively enumerate all potential cliques employing the classical algorithm, and present those cliques consisting of more than three nodes in Fig. 8.

Furthermore, we present the quantum circuits of DCQAOA and QAOA, as depicted in Fig. 9. It is noteworthy that the QAOA circuit does not incorporate the U_{CD} component. The “number of layers” in the QAOA algorithm context refers to the frequency of repetition of the sequence highlighted in the yellow box within the circuit diagram.

APPENDIX C: 3HAC

In the investigation of the eight-qubit system, our analysis concentrated on the interactions involving hd1-HA1, hd1-HA2, hd1-HA3, hd1-Ar4, ar2-HA1, ar2-HA2, ar2-HA3, and ar2-Ar4. For this specific configuration, the parameters were meticulously defined as $\epsilon = 3.1 \text{ \AA}$,

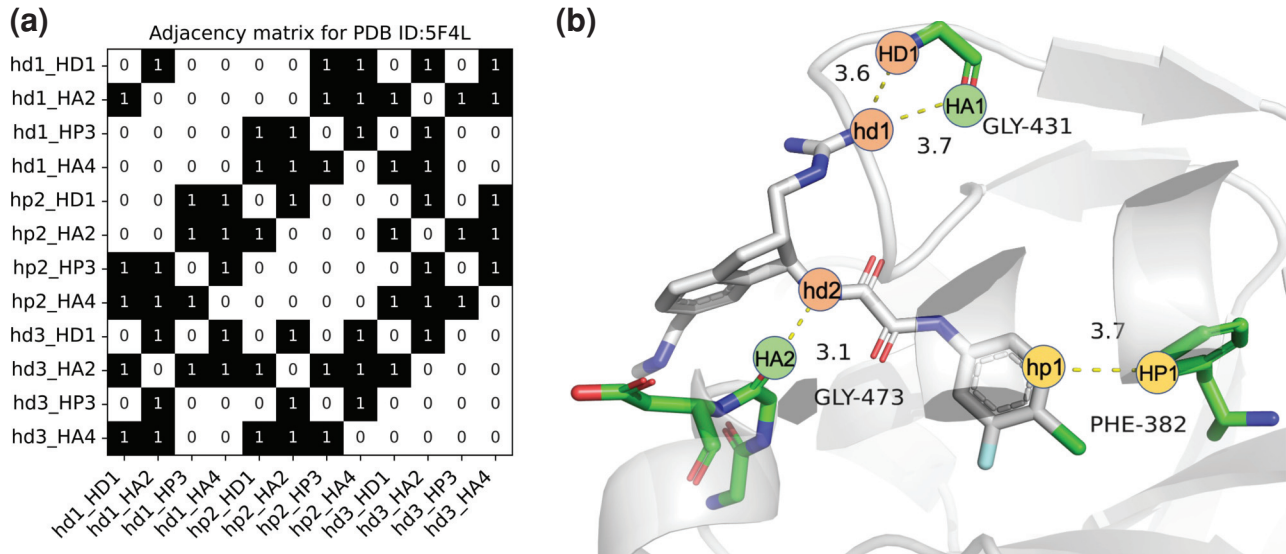


FIG. 17. (a) The adjacency matrix of the binding interaction graph of the HIV-1 gp120 complex with JP-III-048. (b) Predicted protein-ligand docking visualization: the interacting amino acid residues on the protein are shown with green carbon (C) atoms, while the ligand carbon atoms are depicted in white. The highlighted yellow bonds signify the predicted interactions, specifically, hd1-HD1, hd1-HA1, hd2-HA2, and HP-HP1. Notably, hd1 functions as both a hydrogen-bond donor and acceptor, connecting to HD1 and HA1 due to the positive charge on its nitrogen atom [Ng⁺]. Under these circumstances, the nitrogen atom primarily serves as a hydrogen-bond donor, while its capacity to act as a hydrogen-bond acceptor might be diminished.

$\epsilon_s = 2.5 \text{ \AA}$, and $\tau = 0.1 \text{ \AA}$. This approach represents a deviation from the conventional six-qubit model. The primary rationale for this deviation stems from the spatial compactness inherent in the selected pharmacophore points, which necessitated a precise adjustment of the parameters to accommodate the denser configuration. Following these parameter modifications, the BIG values were computed. The results of these computations are graphically represented in Fig. 10. In addition, the corresponding adjacency matrix, which provides an insightful visualization of the intermolecular interactions, is effectively depicted in Fig. 13(a). And, we also undertake a thorough enumeration of all possible cliques using the classical algorithm. Cliques comprising more than three nodes are depicted in Fig. 11.

Furthermore, this study delineates the quantum circuits employed for both the DCQAOA and traditional QAOA algorithms, as exemplified in Fig. 12. A pivotal distinction in the QAOA circuit configuration is the omission of the U_{CD} component. Within the QAOA algorithmic framework, the “number of layers” is defined as the count of sequential repetitions of the process enclosed in the yellow box in the circuit diagram.

APPENDIX D: 5F4L

In the context of a 12-qubit system, we designate the parameters $\epsilon = 2.8 \text{ \AA}$, $\epsilon_s = 2.5 \text{ \AA}$, and $\tau = 0.1 \text{ \AA}$. These values facilitate the computation of the BIG, as delineated

in Fig. 14. Correspondingly, the resulting adjacency matrix is elucidated in Fig. 17(a).

Moreover, we expound upon the quantum circuit architectures of both the DCQAOA and the standard QAOA, as depicted in Fig. 16. It is imperative to note that the QAOA circuit configuration excludes the inclusion of the U_{CD} component. Within the algorithmic framework of QAOA, the “number of layers” is indicative of the iteration frequency of the sequence encapsulated within the circuit diagram’s highlighted yellow box.

- [1] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [2] Y. Manin, Computable and uncomputable, Sovetskoye Radio, Moscow (1980), 128 p.
- [3] P. Benioff, The computer as a physical system: A microscopic quantum mechanical hamiltonian model of computers as represented by turing machines, *J. Stat. Phys.* **22**, 563 (1980).
- [4] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (IEEE, 1994), p. 124.
- [5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge university press, 2010).
- [6] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [7] B. Barak and K. Marwaha, in *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 215, edited by M. Braverman (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022), p. 14:1.

- [8] 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).
- [9] B. Cheng, X.-H. Deng, X. Gu, Y. He, G. Hu, P. Huang, J. Li, B.-C. Lin, D. Lu, Y. Lu, *et al.*, Noisy intermediate-scale quantum computers, *Front. Phys.* **18**, 21308 (2023).
- [10] Y. Kim, A. Eddins, S. Anand, K. X. Wei, v. d. B. Ewout, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala, Evidence for the utility of quantum computing before fault tolerance, *Nature* **618**, 500 (2023).
- [11] Y. Cao, J. Romero, and A. Aspuru-Guzik, Potential of quantum computing for drug discovery, *IBM J. Res. Dev.* **62**, 6 (2018).
- [12] M. Zinner, F. Dahlhausen, P. Boehme, J. Ehlers, L. Bieske, and L. Fehring, Quantum computing's potential for drug discovery: Early stage industry dynamics, *Drug Discov. Today* **26**, 1680 (2021).
- [13] A. Perdomo, C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik, Construction of model hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models, *Phys. Rev. A* **78**, 012320 (2008).
- [14] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, Finding low-energy conformations of lattice protein models by quantum annealing, *Sci. Rep.* **2**, 1 (2012).
- [15] A. Irbäck, L. Knuthson, S. Mohanty, and C. Peterson, Folding lattice proteins with quantum annealing, *Phys. Rev. Res.* **4**, 043013 (2022).
- [16] T. Babej, C. Ing, and M. Fingerhuth, Coarse-grained lattice protein folding on a quantum annealer, *ArXiv:1811.00713* (2018).
- [17] M. Fingerhuth, T. Babej, and C. Ing, A quantum alternating operator ansatz with hard and soft constraints for lattice protein folding, *ArXiv:1810.13411* (2018).
- [18] A. Robert, P. K. Barkoutsos, S. Woerner, and I. Tavernelli, Resource-efficient quantum algorithm for protein folding, *npj Quantum Inf.* **7**, 38 (2021).
- [19] 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).
- [20] F. D. Malone, R. M. Parrish, A. R. Welden, T. Fox, M. Degroote, E. Kyoseva, N. Moll, R. Santagati, and M. Streif, Towards the simulation of large scale protein–ligand interactions on nisq-era quantum computers, *Chem. Sci.* **13**, 3094 (2022).
- [21] J. J. Kirsopp, C. Di Paola, D. Z. Manrique, M. Krompiec, G. Greene-Diniz, W. Guba, A. Meyder, D. Wolf, M. Strahm, and D. Muñoz Ramo, Quantum computational quantification of protein–ligand interactions, *Int. J. Quantum Chem.* **122**, e26975 (2022).
- [22] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, Molecular docking with Gaussian boson sampling, *Sci. Adv.* **6**, eaax1950 (2020).
- [23] S. Yu, Z.-P. Zhong, Y. Fang, R. B. Patel, Q.-P. Li, W. Liu, Z. Li, L. Xu, S. Sagana-Stopfel, E. Mer, *et al.*, A universal programmable Gaussian boson sampler for drug discovery, *Nat. Comput. Sci.* **3**, 839 (2023).
- [24] K. Mato, R. Mengoni, D. Ottaviani, and G. Palermo, Quantum molecular unfolding, *Quantum Sci. Technol.* **7**, 035020 (2022).
- [25] 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).
- [26] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, *ArXiv:1411.4028* (2014).
- [27] 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, *ArXiv:2306.09198* (2023).
- [28] 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).
- [29] D. Headley, T. Müller, A. Martin, E. Solano, M. Sanz, and F. K. Wilhelm, Approximating the quantum approximate optimization algorithm with digital-analog interactions, *Phys. Rev. A* **106**, 042446 (2022).
- [30] 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).
- [31] V. Akshay, H. Philathong, E. Campos, D. Rabinovich, I. Zacharov, X.-M. Zhang, and J. D. Biamonte, Circuit depth scaling for quantum approximate optimization, *Phys. Rev. A* **106**, 042438 (2022).
- [32] F. S. Kuhl, G. M. Crippen, and D. K. Friesen, A combinatorial algorithm for calculating ligand binding, *J. Comput. Chem.* **5**, 24 (1984).
- [33] Y. Wang, J.-K. Hao, F. Glover, Z. Lü, and Q. Wu, Solving the maximum vertex weight clique problem via binary quadratic programming, *J. Comb. Optim.* **32**, 531 (2016).
- [34] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization* (Kluwer Academic Publishers, Boston, 1995).
- [35] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W.H. Freeman & Co., San Francisco, 1979).
- [36] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou, in *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 232, edited by F. Le Gall and T. Morimae (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2022), p. 7:1.
- [37] 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).
- [38] J. Choi and J. Kim, in *2019 International Conference on Information and Communication Technology Convergence (ICTC)* (IEEE, 2019), p. 138.
- [39] 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).

- [40] J. Wurtz and P. Love, Maxcut quantum approximate optimization algorithm performance guarantees for $p > 1$, *Phys. Rev. A* **103**, 042612 (2021).
- [41] K. Marwaha, Local classical max-cut algorithm outperforms $p = 2$ QAOA on high-girth regular graphs, *Quantum* **5**, 437 (2021).
- [42] E. Farhi, D. Gamarnik, and S. Gutmann, The quantum approximate optimization algorithm needs to see the whole graph: Worst case examples, *ArXiv:2005.08747* (2020).
- [43] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, *Quantum* **6**, 759 (2022).
- [44] A. S. Bandeira, D. Kunisky, and A. S. Wein, in *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 151, edited by T. Vidick (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2020), p. 78:1.
- [45] A. Montanari, Optimization of the Sherrington–Kirkpatrick Hamiltonian, *SIAM J. Comput.* **0**, FOCS19-1 (2021).
- [46] A. El Alaoui, A. Montanari, and M. Sellke, Optimization of mean-field spin glasses, *Ann. Probab.* **49**, 2922 (2021).
- [47] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, 2022), p. 335.
- [48] R. Shaydulin, *et al.*, Evidence of scaling advantage for the quantum approximate optimization algorithm on a classically intractable problem, *ArXiv:2308.02342* (2023).
- [49] A. Lucas, Ising formulations of many NP problems, *Front. Phys.* **2**, 5 (2014).
- [50] R. Herrman, P. C. Lotshaw, J. Ostrowski, T. S. Humble, and G. Siopsis, Multi-angle quantum approximate optimization algorithm, *Sci. Rep.* **12**, 6781 (2022).
- [51] K. Shi, R. Herrman, R. Shaydulin, S. Chakrabarti, M. Pistoia, and J. Larson, in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)* (IEEE, 2022), p. 414.
- [52] 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).
- [53] X. Chen, A. Ruschhaupt, S. Schmidt, A. del Campo, D. Guéry-Odelin, and J. G. Muga, Fast optimal frictionless atom cooling in harmonic traps: Shortcut to adiabaticity, *Phys. Rev. Lett.* **104**, 063002 (2010).
- [54] N. N. Hegade, K. Paul, Y. Ding, M. Sanz, F. Albarrán-Arriagada, E. Solano, and X. Chen, Shortcuts to adiabaticity in digitized adiabatic quantum computing, *Phys. Rev. Appl.* **15**, 024038 (2021).
- [55] A. del Campo, M. M. Rams, and W. H. Zurek, Assisted finite-rate adiabatic passage across a quantum critical point: Exact solution for the quantum Ising model, *Phys. Rev. Lett.* **109**, 115703 (2012).
- [56] M. Demirplak and S. A. Rice, Assisted adiabatic passage revisited, *J. Phys. Chem. B* **109**, 6838 (2005).
- [57] M. Demirplak and S. A. Rice, Adiabatic population transfer with control fields, *J. Phys. Chem. A* **107**, 9937 (2003).
- [58] M. V. Berry, Transitionless quantum driving, *J. Phys. A: Math. Theor.* **42**, 365303 (2009).
- [59] K. Z. Li and D. M. Tong, Geometric property of energetic cost in transitionless quantum driving, *Phys. Rev. A* **106**, 052414 (2022).
- [60] 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).
- [61] 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).
- [62] P. Moon, L. Boike, D. Dovala, N. J. Henning, M. Knapp, J. N. Spradlin, C. C. Ward, H. Wolleb, C. M. Zammit, D. Fuller, *et al.*, Discovery of potent pyrazoline-based covalent SARS-CoV-2 main protease inhibitors, *bioRxiv*, **2022** (2022).
- [63] G. Landrum, *et al.*, rdkit/rdkit: 2023_09_4 (q3 2023) release (2024).
- [64] L. Schrödinger and W. DeLano, Pymol, <http://www.pymol.org/pymol>.
- [65] R. Wang, X. Fang, Y. Lu, and S. Wang, The pdbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures, *J. Med. Chem.* **47**, 2977 (2004).
- [66] R. Wang, X. Fang, Y. Lu, C.-Y. Yang, and S. Wang, The pdbind database: Methodologies and updates, *J. Med. Chem.* **48**, 4111 (2005).
- [67] Z. Liu, Y. Li, L. Han, J. Li, J. Liu, Z. Zhao, W. Nie, Y. Liu, and R. Wang, Pdb-wide collection of binding data: Current status of the pdbind database, *Bioinformatics* **31**, 405 (2015).
- [68] M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. J. Haupt, and M. Schroeder, Plip 2021: Expanding the scope of the protein–ligand interaction profiler to DNA and RNA, *Nucleic Acids Res.* **49**, W530 (2021).
- [69] S. D. Edmondson, A. Mastracchio, J. M. Cox, G. J. Eiermann, H. He, K. A. Lyons, R. A. Patel, S. B. Patel, A. Petrov, G. Scapin, *et al.*, Aminopiperidine-fused imidazoles as dipeptidyl peptidase-iv inhibitors, *Bioorg. Med. Chem. Lett.* **19**, 4097 (2009).
- [70] M. Developer, Mindquantum, version 0.6.0, <https://gitee.com/mindspore/mindquantum> (2021).
- [71] S. Cao, B. Wu, F. Chen, M. Gong, Y. Wu, Y. Ye, C. Zha, H. Qian, C. Ying, S. Guo, *et al.*, Generation of genuine entanglement up to 51 superconducting qubits, *Nature* **619**, 738 (2023).
- [72] H2-quantum computing hardware, <https://www.quantinuum.com/hardware/h2> (2024).
- [73] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, *et al.*, Logical quantum processor based on reconfigurable atom arrays, *Nature* **626**, 58 (2023).
- [74] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, *et al.*, Quantum computational advantage with a programmable photonic processor, *Nature* **606**, 75 (2022).
- [75] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, Quantum natural gradient, *Quantum* **4**, 269 (2020).

Correction: The previously published Figures 3 and 5 and Figures 4 and 6 were interchanged during the proof production cycle and have been fixed.