

QAS-BO : Quantum Architecture Search Based on Bayesian Optimization Applied to Variational Quantum Algorithms

Shuyan Chao¹ Yuxin Deng² Zhanou Liu Yuwei Zhang

Abstract—In the era of Noisy Intermediate-Scale Quantum (NISQ) computing, traditional quantum algorithms face the challenges of limited number of qubits, noise and decoherence. In order to address these issues, we propose a Quantum Architecture Search (QAS) method driven by Bayesian Optimization (BO), which is applied to variational quantum algorithms. In this work, QAS is regarded as a fixed-scale sampling problem. We innovatively propose a quantum gate pool and use a parameterized probabilistic model to dynamically determine the optimal quantum gate for each position in the quantum circuit, thus optimizing the circuit structure. Through using a gradient-free BO method based on radial basis function, we adaptively design end-to-end quantum circuits, significantly reducing circuit depths and improving computational accuracy. We conducted experiments on ground state energy estimation in quantum chemistry and combinatorial optimization problem. The experimental results show that our method is significantly superior to traditional methods and other heuristic search methods in accuracy and efficiency. Our method not only reduces the depth of quantum circuits by up to 85% under a certain accuracy, but also improves the accuracy rate to nearly 100% in combinatorial optimization problem. This provides a powerful and efficient tool for designing optimal quantum circuits and promotes the practical application of quantum algorithms in the NISQ era.

I. INTRODUCTION

The arrival of Noisy Intermediate-Scale Quantum (NISQ) era brings opportunities and challenges to quantum computing. Traditional quantum algorithms, such as the Quantum Phase Estimation (QPE) algorithm [1], require a large number of quantum gates and considerable circuit depth when solving complex quantum problems. Specifically, QPE is used to accurately measure the phase information of quantum states, and its operation involves a series of precise Quantum Fourier Transform (QFT) and control operations, which place high demands on the hardware. In the NISQ era, due to the limited number of qubits and the influence of noise and decoherence on the operational accuracy, it is difficult to achieve long-term quantum state maintenance and complex operation sequences. Therefore, when the QPE algorithm is executed on the existing NISQ hardware, the circuit depth often exceeds the capability limit of the hardware, which makes it difficult to be applied in practice.

In order to solve this problem, Peruzzo et al. proposed the Variational Quantum Eigensolver (VQE) to estimate the ground state energy of molecules, which greatly reduced the required circuit depth [2]. However, VQE still faces the

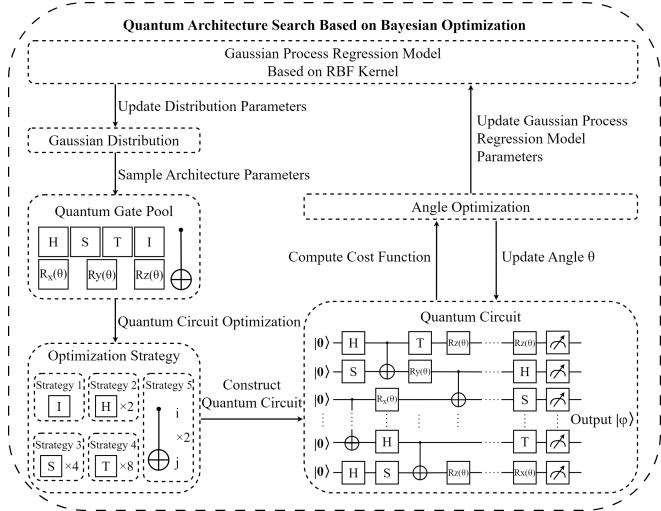


Fig. 1. QAS-BO method flow chart

Barren Plateau (BP) problem [3], where the cost function gradients become extremely small when optimizing large Parameterized Quantum Circuits (PQCs), making the optimization process difficult and inefficient.

In this paper, we propose a novel method to optimize PQCs by using Bayesian Optimization (BO) driven QAS (called QAS-BO) (Figure 1). This powerful gradient-free global optimization strategy is especially suitable for optimizing black-box functions that are expensive or difficult to evaluate directly. The whole algorithm regards QAS as a black box function of BO, and uses QAS-BO method to adjust the quantum circuit architecture. It automatically selects gates from quantum gate pool and optimizes the circuit by using optimization strategy. It makes full use of the previous observation information by establishing the probabilistic model of the objective function, and explores and utilizes the balance based on feedback adjustment, while processing the noisy data.

This method aims at automatically designing end-to-end quantum circuits and dynamically adjusting the quantum architecture. By introducing gradient-free optimization into quantum circuit design, our method alleviates the BP problem. Through heuristic search circuit and gaussian distribution to adjust the probability of quantum gate selection, our method does not need special domain knowledge, thus realizing the high automation of quantum circuit design process and the optimization of resource use and calculation efficiency.

¹Software Science and Technology, East China Normal University, Shanghai, China 51265902018@stu.ecnu.edu.cn

²East China Normal University, Shanghai, China yxdeng@sei.ecnu.edu.cn

In the experimental part, we mainly test our algorithm on two Variational Quantum Algorithms (VQAs):

Ground State Energy Estimation in Quantum Chemistry

We investigate the ground state energy estimation of different molecule ion systems, and our method keeps within the chemical accuracy, significantly reducing the number of quantum gates and circuit depth. For example, for LiH-4, the circuit depth is only 33, which reduces the resource consumption by nearly 85% compared with the Unitary Coupled Cluster Singles and Doubles (UCCSD) method. For small molecules, we can design quantum circuits with higher accuracy and lower depth than the UCCSD and the Hardware-Efficient (HE) method, such as HeH^+ and H_2^+ .

Combinatorial Optimization Problem

By applying our algorithm, a shallow and efficient quantum circuit for combinatorial optimization problem (unweighted Max-Cut problem, weighted Max-Cut problem and Not-All-Equal 3-Satisfiability) is successfully designed. The quantum circuit can efficiently find the solution of combinatorial optimization problem. The experimental results show that, compared with the commonly used QAOA circuits, the quantum circuit we searched out performs well under different dimensions and experimental settings. For example, in the unweighted Max-Cut experiment of 10 nodes, the circuit depth searched by our algorithm is 41, and the accuracy reaches 98.93%, far exceeding the accuracy of QAOA-4 and QAOA-5, and greatly reducing the circuit depth. Our algorithm achieves high accuracy, and only needs a few measurements, which reduces the calculation cost and improves the efficiency of resource use.

These experimental results show that our QAS framework can effectively find the optimal quantum circuit for a specific quantum computing task, which further promotes the development of quantum computing technology.

The rest of the paper is organized as follows: First of all, we review the work closely related to our research. Then, we introduce our method framework in detail and test its performance on ground state energy estimation in quantum chemistry and combinatorial optimization problem. Finally, the paper summarizes the research results and discusses the possible direction of future work.

II. RELATED WORKS

A. Quantum Architecture Search

The purpose of QAS is to automatically design PQCs to optimize and accelerate the performance of VQAs. Through QAS, the workload of manually designing quantum circuits can be effectively reduced, and at the same time, the best circuit architecture suitable for specific quantum tasks can be found.

Methods related to QAS have also developed rapidly in recent years. Among these advancements, the application of machine learning techniques in QAS has gained significant attention. Especially, the methods of Reinforcement Learning (RL) [4] and Generative Adversarial Networks (GANs) [5] have shown impressive results. The applications of heuristic and meta-heuristic algorithms include Genetic Algorithm

(GA) [6], Particle Swarm Optimization (PSO) [7] and Simulated Annealing (SA) [8]. These methods combine the advantages of classical computing and quantum computing, consider hardware characteristics and noise environment, and realize more efficient and low-error quantum circuit design through multi-level optimization strategies.

Zhang et al. proposed a Differentiable Quantum Architecture Search (DQAS) method based on the Gumbel-Softmax techniques [9]. By transforming the discrete design problem of quantum circuits into a differentiable continuous optimization problem, the automatic design of quantum circuit is realized. Additionally, DQAS sets a parameter sharing pool to reduce the search space.

The above method is improved and expanded to be the Q-DARTS [10]. The extended algorithm includes two versions: macro search and micro search, in which macro search directly searches the whole circuit and is suitable for small and medium-sized quantum circuit problems, while micro search shows its potential in large-scale QAS problems by inferring the sub-circuit structure from small-scale problems and then applying it to large-scale problems. And the setting of the parameter sharing pool is cancelled.

In our work, we regard the QAS as a fixed-scale sampling problem, in which each position only corresponds to a single parameter. Compared with the DQAS and Q-DARTS, this method significantly reduces the complexity of architecture search, which needs to optimize two parameters for each position.

B. Bayesian Optimization for Quantum Computing

In recent years, the application of BO in quantum computing has made remarkable progress. Many works have demonstrated the potential and effect of this optimization method in improving quantum algorithm and quantum circuit design.

Duong et al. improves the performance of BO in the search of Quantum Neural Network (QNN) architecture by introducing a quantum gate distance metric [11].

Nicoli et al. put forward a BO method driven by physical information in their research [12]. They combined the intrinsic physical characteristics of VQE, introduced a new VQE kernel function and a new acquisition function named EMICoRe, and proposed an angle optimization method, which effectively reduced the posterior uncertainty.

Dai et al. proposed a Quantum-Gaussian Process-Upper Confidence Bound (Q-GP-UCB) in their research [13]. By introducing Quantum Monte Carlo (QMC) subroutine and weighted Gaussian Process Regression (GPR), the upper bound of cumulative regret of BO is significantly improved. Q-GP-UCB can theoretically realize the upper bound of cumulative regret of $O(\text{polylog}T)$, which is significantly smaller than the lower bound of classical methods.

Our research is different from Duong et al., and it is aimed at optimizing the quantum circuit itself directly. Optimizing the circuit instead of its matrix representation can avoid the complexity of matrix calculation, especially in high-dimensional space, which can significantly reduce the com-

putational overhead. The optimization results can be directly mapped to the actual quantum hardware for verification and execution without complicated matrix-to-circuit conversion. By directly optimizing the quantum circuit architecture, our research puts forward a more efficient method, which fills a gap in the current research and provides new ideas and tools for the future quantum algorithm design.

III. METHOD FRAMEWORK

In this research, we regard QAS as a fixed-scale sampling problem. We use BO method to fine-tune the quantum architecture to optimize the circuit architecture. Then, combined with the prior knowledge in quantum field, we further optimize the circuit design and effectively reduce the circuit depth to resist the influence of quantum decoherence. Finally, we use COBYLA algorithm [14] to optimize the angle parameters in order to achieve higher efficiency and stability.

A. Quantum Gate Pool Setting and Mapping of Quantum Circuits

We propose an innovative framework for the selection and layout of quantum gates, aiming at automating the design process of quantum circuits. This framework uses parameterized probabilistic model to dynamically determine the optimal quantum gate for each position in the quantum circuit, thus optimizing the circuit structure.

Our method integrates a quantum gate pool (1). It is complete, which means that any unitary circuit can be represented only by the gates in the quantum gate pool. The quantum gate pool \mathcal{G} is composed of a universal and complete quantum gate set \mathcal{G}_1 [15], commonly used rotation gate set \mathcal{G}_2 and the I gate to ensure the flexibility and robustness of the architecture. The selection of each quantum gate is based on their key role in constructing complex quantum algorithms (Appendix VI).

$$\begin{aligned}\mathcal{G}_1 &= \{H, S, T, CNOT\}, \\ \mathcal{G}_2 &= \{Rx(\theta), Ry(\theta), Rz(\theta)\} \\ \mathcal{G} &= \mathcal{G}_1 \cup \mathcal{G}_2 \cup \{I\}\end{aligned}\quad (1)$$

For a quantum circuit with n qubits and m layers, any unitary operation U can be represented by Equation 2. The unitary matrix U_{ij} in the i -th qubit and j -th layer of the circuit is determined by the corresponding quantum gate C_{ij} .

$$\begin{aligned}U &= \prod_{j=1}^m \bigotimes_{i=1}^n U_{ij}, U_{ij} = \begin{cases} C_{ij}, C_{ij} \in \mathcal{G}_i \\ I, C_{ij} \notin \mathcal{G}_i \end{cases} \\ \mathcal{G}_i &= \{R_x(\theta), R_y(\theta), R_z(\theta), H, S, T, I\} \\ &\cup \{CNOT_{i,i+1}, \dots, CNOT_{i,n}\}\end{aligned}\quad (2)$$

The quantum gate C_{ij} for the i -th qubit and the j -th layer position is derived from the probability density function of the Gaussian distribution (3) (Figure 2).

$$\begin{aligned}p(C_{ij} = \mathcal{G}_i^{(k)}) &= \max_k f(k|\mu_{ij}, \sigma^2) \\ f(k|\mu_{ij}, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k-\mu_{ij})^2}{2\sigma^2}}\end{aligned}\quad (3)$$

where $\mathcal{G}_i^{(k)}$ is the k -th quantum gate in \mathcal{G}_i .

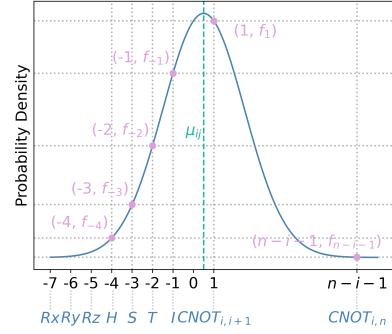


Fig. 2. Sampling of C_{ij} from the Gaussian distribution

Algorithm 1 Bayesian Optimization Algorithm with Noise

Input: Black box function $E(\mu)$, search space $\mu \in \mathbb{D}^{mn} \subset \mathbb{R}^{mn}$

Output: $E_m = \min_\mu E(\mu)$

- 1: Set the prior for $E(\mu)$ following the GPR model based on the RBF kernel.
 - 2: **for** $i = 1$ to N_C **do**
 - 3: Sample initial points $\mu_i \in \mathbb{D}^{mn}$ randomly.
 - 4: Initial training set $\Gamma \leftarrow \Gamma \cup \{\mu_i, E(\mu_i)\}$.
 - 5: **end for**
 - 6: Set $E_m = \min_j [\{E(\mu_i)\}_{i=1}^{N_C}]$.
 - 7: Optimize hyperparameters ℓ, σ^2 and σ_{NOISE}^2 according to Γ .
 - 8: **for** $n = 1$ to N_{BO} **do**
 - 9: Update the posterior distribution of $E(\mu)$ by using the hyperparameters ℓ, σ^2 and σ_{NOISE}^2 .
 - 10: Select the best sampling point $\tilde{\mu}$ by using the acquisition function based on posterior distribution.
 - 11: **if** $E(\tilde{\mu}) < E_m$ **then**
 - 12: $E_m \leftarrow E(\tilde{\mu})$.
 - 13: **end if**
 - 14: Update training set $\Gamma \leftarrow \Gamma \cup \{\tilde{\mu}, E(\tilde{\mu})\}$.
 - 15: Optimize hyperparameters ℓ, σ^2 and σ_{NOISE}^2 according to Γ .
 - 16: **end for**
-

B. Quantum Architecture Search Based on Bayesian Optimization

Bayesian optimization (BO) is an efficient global optimization method, especially suitable for optimizing complex black-box functions with high cost and noise. For a detailed algorithmic workflow, refer to Algorithm 1.

In our work, we comprehensively consider a variety of noise sources, including quantum decoherence, environmental noise, control error and the error of precise adjustment of pulse amplitude and duration due to the limitation of practical operation [16]. The representative model for these noise sources is $f_{NOISE} \sim \mathcal{GP}(0, \sigma_{NOISE}^2)$.

The optimization process revolves around adjusting the Gaussian distribution parameters representing the circuit architecture.

In this research, we establish the GPR model based on

the Radial Basis Function (RBF) kernel (4) as a proxy model to approximate the unknown objective function. The model is used to predict and update the optimization process of quantum circuit parameters in order to minimize the objective function. The GPR model based on the RBF kernel is a nonparametric prediction model widely used in machine learning and statistics. This model is especially suitable for dealing with complex and nonlinear data prediction problems, because it can flexibly adapt to potential relationships in data [17].

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

$$k(x, x') = \sigma^2 \exp\left(-\frac{\|x - x'\|^2}{2l^2}\right) + \sigma_{NOISE}^2 I \quad (4)$$

where x, x' are elements in the input space, and $m(x)$ is a mean function, which provides the expected output value for the input point x . Here σ^2 represents the fluctuation of the function between different node values, and l represents the degree of correlation between nodes. Note that σ_{NOISE}^2 represents the variance of the noise model, which describes the fluctuation of the observed values around the real values.

For the given training set $\Gamma = \{(\mu_i, E(\mu_i))\}_{i=1}^z$, the Log Marginal Likelihood (LML) of the GPR model is defined as in Equation 5:

$$\log p(\mathbf{E}(\mathbf{N}) | \mathbf{N}, \ell, \sigma^2, \sigma_{NOISE}^2)$$

$$= -\frac{1}{2}\mathbf{E}(\mathbf{N})^T \mathbf{K}^{-1} \mathbf{E}(\mathbf{N}) - \frac{1}{2} \log |\mathbf{K}| - \frac{z}{2} \log 2\pi \quad (5)$$

where $\mathbf{N} = (\mu_1, \dots, \mu_z)$ represents the matrix of input data, $\mathbf{E}(\mathbf{N}) = (E(\mu_1), \dots, E(\mu_z))^T$ represents the vector of the target output, $\mathbf{K} = (k(\mu_i, \mu_j))^{z \times z}$ represents the covariance matrix.

The optimization goal of the core hyperparameters $\ell, \sigma^2, \sigma_{NOISE}^2$ of a GPR model is to find the parameter set that can maximize the LML. This optimization process adopts the Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [18]. L-BFGS effectively promotes the optimization process by calculating the gradient of LML relative to hyperparameter and updating the hyperparameter values accordingly. Even though L-BFGS involves the gradient of LML relative to hyperparameter, this method has been proved to be effective in dealing with the BP problem of likelihood function [19]. By adopting this method, the kernel function of the GPR model not only captures the correlation between data points, but also naturally integrates the noise processing mechanism, thus simplifying the representation of the model and improving the calculation efficiency.

The new sampling point $\tilde{\mu}$ is selected by maximizing the acquisition function. There are three main acquisition functions: Expected Improvement (EI), Probability Improvement (PI) and Lower Confidence Bound (LCB) [20]. EI (6) tends to search carefully around the known optimal value.

$$EI(\mu) = \mathbb{E}[\max(E(\mu) - E(\tilde{\mu}), 0)]$$

$$= (m(\mu) - E(\tilde{\mu}) - \xi) \Phi(Z) + \sigma(\mu) \phi(Z) \quad (6)$$

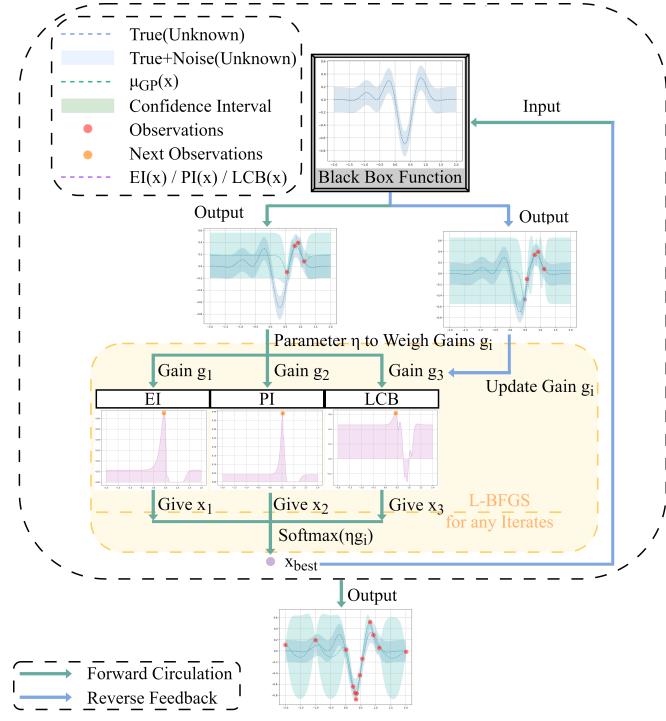


Fig. 3. Implementation of the adaptive acquisition function strategy based on MAB principle. The figure shows a one-dimensional special case, which is only used to clearly show the working principle of this strategy.

PI (7) tends to explore potential areas that have not been fully tested.

$$PI(\mu) = \Phi\left(\frac{m(\mu) - E(\tilde{\mu}) - \xi}{\sigma(\mu)}\right) \quad (7)$$

And LCB (8) provides a trade-off between exploration and utilization.

$$LCB(\mu) = m(\mu) - \kappa \sigma(\mu) \quad (8)$$

where $m(\mu)$ represents the predicted mean at point μ . $\sigma(\mu)$ represents the predicted standard deviation at point μ , and $Z = \frac{m(\mu) - E(\tilde{\mu})}{\sigma(\mu)}$. The term ξ is a small exploration parameter. $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and the probability density function of the standard normal distribution. κ is a coefficient that balances exploration and exploitation, dynamically adjusted based on the number of optimization iterations.

Our research adopts the adaptive acquisition function strategy based on Multi-Armed Bandit (MAB) principle (Figure 3) [21]. By maintaining the historical performance log of each strategy, and using the softmax function (9) to calculate the selection probability, the strategy with better performance has a higher selection probability, but all strategies have the possibility of being selected.

$$\text{softmax}(\eta g_i) = \frac{e^{\eta g_i}}{\sum_j e^{\eta g_j}} \quad (9)$$

where g_i is the gain weight of the i -th acquisition function, which represents the gain values of different acquisition functions (EI, PI, LCB). η is an adjustment parameter to

adjust the influence of different acquisition function weights. This method ensures the diversity and comprehensiveness of the selection of acquisition points by the acquisition function. The randomness is introduced to avoid over-reliance on deterministic selection, which is helpful to explore the optimal solution and reduce the risk caused by single strategy deviation.

The whole optimization framework not only pursues the efficiency of parameter optimization, but also aims to improve the robustness and reliability of the algorithm in the face of complex practical environment. In addition, we use multi-core parallel computing to accelerate the optimization process and improve the overall computing efficiency. All experiments were conducted on a server with an Intel Xeon Silver 4210R CPU at 2.40 GHz, 40 cores, and 512 GB RAM.

IV. APPLICATIONS

A. Ground State Energy Estimation in Quantum Chemistry

The main goal of VQE is to find the minimal energy state of a quantum system, named the energy of the ground state E_{min} . In this method, the system's total energy is described with the Hermitian operator Hamiltonian H , whose minimal eigenvalue corresponds to the ground state of the system. In VQE, the PQC performs the unitary transformation $U(\theta)$, known as *ansatz* [22]. Then, it uses the ansatz to prepare the parametrized quantum state $|\psi(\theta)\rangle = U(\theta)|\psi_0\rangle$, where $|\psi_0\rangle$ is the initial quantum state. The parameters θ of the quantum circuit are iteratively adapted by the classical optimization algorithm to a minimum of the energy expectation value E_{min} (10), which, in turn, corresponds to the energy of the system's ground state.

$$E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle \quad (10)$$

$$E_{min} \leq \min_{\theta} E(\theta)$$

VQE in quantum chemistry can be applied to calculate the ground state energy of molecules, which is one of the most important parameters for the prediction of the most stable structure of a molecule. The simulation of quantum mechanical behavior of molecules on a quantum computer allows for the prediction of the energy of molecules in different configurations and the finding of a stable structure with the lowest energy.

Selecting the appropriate ansatz is crucial for successfully simulating ground state energy. At present, the ansatz design is prominently divided into two categories: The Hardware-Efficient (HE) ansatz [23] and the Unitary Coupled Clustered (UCC) ansatz [24]. The goal of the HE ansatz is to reduce the number of quantum gates by simplifying the circuit structure, thus reducing the hardware resources and error rate required for quantum computing. This alignment reduces the overhead required for translation of more abstract, hardware-independent circuits into those efficient to execute on a particular device [25]. The most representative method in UCC is called the Unitary Coupled Cluster Singles and Doubles (UCCSD). In UCCSD, the theory of coupled clusters in quantum chemistry approaches the solution in ansatz

by including single-electron and double-electron excitation, which is usually recognized for its relatively high accuracy and adaptability to chemical problems [26]. However, it has high computational complexity and demands on resources, especially in cases with large molecular systems. On the contrary, the HE ansatz is characterized by low resource demand and rapid computation speed, making it particularly suitable for rapid prototyping and preliminary screening, though its accuracy is generally lower than that of the UCCSD ansatz.

In this research, our algorithm is compared with the traditional UCCSD method and the HE method in detail, and the energy accuracy and computational resource requirements when dealing with molecular ion systems with strong electron correlation or asymmetric electron distribution are emphatically evaluated. We carried out experiments across 10 different random seeds, with the results shown in Table I.(Note: Only partial data is displayed in the table, the remaining experimental results can be found in Appendix VII.)

As can be seen from Table I, our algorithm is significantly superior to HE in energy calculation accuracy. Compared with the UCCSD, our algorithm is only one order of magnitude worse in energy accuracy, and the results of each experiment are within the range of chemical accuracy (That is, the error is less than 0.0016 \AA). And when the simulation object becomes ions, especially systems with strong electron correlation or asymmetric electron distribution, the UCCSD ansatz shows obvious shortcomings. For example, the energy of HeH^+ under the UCCSD is 6.5×10^{-8} , while our algorithm shows lower energy deviation under different seeds, with an average energy deviation of 8.1×10^{-10} .

In terms of quantum circuit depth, our algorithm shows a significant resource efficiency advantage. Compared with the depth required by the UCCSD, our algorithm greatly reduces the number of quantum gates required. For example, the depth of LiH-4 under UCCSD is 222, while our algorithm only needs 33, which reduces the resource consumption by nearly 85%. This optimization makes the algorithm more feasible when dealing with large-scale molecular systems, especially under the existing quantum hardware conditions, the number and depth of gate operations directly affect the practical application of the algorithm.

In addition, we made a deep comparative research on the H_2 molecule, and compared our method with other heuristic search methods, including Q-DARTS, DQAS, QCAS [27] and RL [4] (Figure 4). The results show that our method can find quantum circuits with higher accuracy and shallower circuit depth than these known methods. In quantum computing, the deeper the circuit is, the more susceptible it is to noise, which will accelerate the process of quantum decoherence [28] and affect the accuracy and reliability of calculation. Our research results confirm that the optimized shallow circuit can effectively alleviate this problem and enhance the practicability of quantum computing.

TABLE I

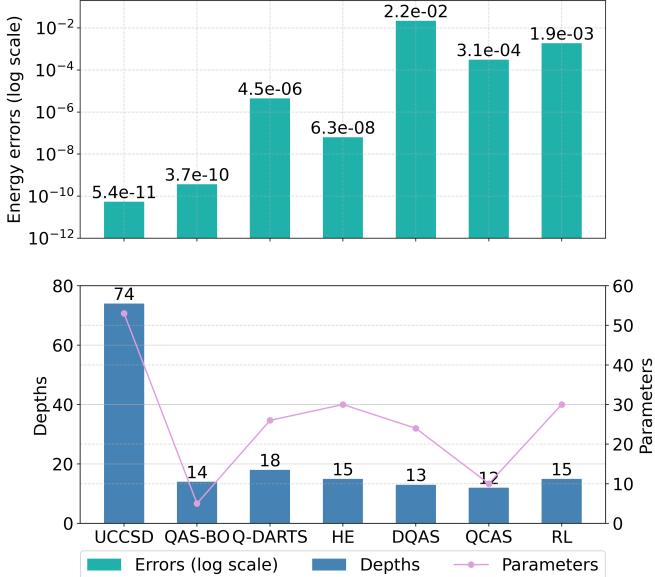
COMPARISON OF ESTIMATION RESULTS FOR VARIOUS MOLECULES ACROSS DIFFERENT RANDOM SEEDS AND METHODS (SEED 0, SEED 1, AND SEED 2 COLUMN REPRESENT THE ENERGY ERROR OF QAS-BO, THE UCCSD COLUMN REPRESENTS THE ENERGY ERROR OF UCCSD, AND THE HE COLUMN REPRESENTS THE ENERGY ERRORS OF HE.)

	Seed 0	Seed 1	Seed 2	Avg depth	UCCSD	Depth	HE	Depth
HeH ⁺	9.1×10^{-10}	7.6×10^{-10}	7.7×10^{-10}	9	6.5×10^{-8}	11	2.1×10^{-7}	17
H ⁺	1.2×10^{-9}	2.9×10^{-9}	4.5×10^{-9}	21	6.1×10^{-1}	9	2.1×10^{-6}	23
H ₂	3.6×10^{-10}	3.7×10^{-10}	7.7×10^{-10}	17	5.4×10^{-11}	74	8.4×10^{-6}	15
LiH-4	4.5×10^{-6}	5.3×10^{-5}	4.7×10^{-6}	33	2.3×10^{-7}	222	6.6×10^{-4}	19
LiH-6	3.1×10^{-4}	2.5×10^{-4}	2.1×10^{-4}	46	2.3×10^{-7}	317	1.7×10^{-3}	21

TABLE II

EXPERIMENTAL RESULTS OF COMBINATORIAL OPTIMIZATION PROBLEM: COMPARATIVE ANALYSIS OF ALGORITHM DEPTH AND ACCURACY RATE, WHERE QAOA- p REPRESENTS ANSATZ COMPOSED OF p LAYERS OF H_C AND H_M .

	Unweighted Max-Cut-4 nodes			Unweighted Max-Cut-6 nodes			Unweighted Max-Cut-8 nodes		
	QAS-BO	QAOA-3	QAOA-4	QAS-BO	QAOA-4	QAOA-5	QAS-BO	QAOA-4	QAOA-5
Depth Accuracy	24	44	57	27	87	106	40	63	76
	100.00%	87.50%	100.00%	100.00%	29.30%	46.38%	100.00%	90.43%	89.45%
	Weighted Max-Cut-4 nodes			Weighted Max-Cut-6 nodes			Weighted Max-Cut-8 nodes		
	QAS-BO	QAOA-6	QAOA-7	QAS-BO	QAOA-7	QAOA-8	QAS-BO	QAOA-8	QAOA-10
Depth Accuracy	27	47	54	30	144	163	46	94	114
	100.00%	94.82%	98.54%	100.00%	52.25%	27.25%	96.19%	20.41%	41.89%
	NAE 3-SAT-4 literals			NAE 3-SAT-6 literals			NAE 3-SAT-8 literals		
	QAS-BO	QAOA-4	QAOA-5	QAS-BO	QAOA-2	QAOA-3	QAS-BO	QAOA-4	QAOA-5
Depth Accuracy	10	30	37	19	46	57	50	68	81
	100.00%	98.73%	100.00%	100.00%	8.92%	12.32%	100.00%	4.59%	10.21%

Fig. 4. H₂ Estimation: Energy error, depth, and parameters comparison

B. Combinatorial Optimization Problem

Combinatorial optimization problem is an important field in mathematical optimization, which mainly involves finding the optimal solution in the set of discrete objects that meet specific constraints. This problem is widely used in many practical fields, such as network design, scheduling, routing planning and so on. As the problem size increases, the size of

the solution space typically grows exponentially. Therefore, such problems are classified as classic NP-hard problems.

The Quantum Approximate Optimization Algorithm (QAOA) is an approximate optimization algorithm with polynomial time, which is used to solve combinatorial optimization problems [29]. For graphs with n nodes, QAOA transforms combinatorial optimization problem into a large sparse diagonal matrix H_C (11), and its goal is to find the minimum eigenvalue of H_C . Classically, its time complexity is $O(2^n)$. From the quantum point of view, these steps are realized by quantum circuits.

$$H_C = - \sum_{(i,j) \in E} \frac{I - Z_i Z_j}{2} \quad (11)$$

QAOA's quantum circuit has a fixed architecture, and its initial state is generally a uniform superposition state (12).

$$|+\rangle = \frac{1}{\sqrt{2^n}} \sum_{z=0}^{2^n-1} |z\rangle \quad (12)$$

The quantum circuit is formed by the cost Hamiltonian H_C and the mixed Hamiltonian H_M (13) which interact repeatedly. The former contains graph information, while the latter explores 2^n different combinations.

$$H_M = \sum_i X_i \quad (13)$$

By minimizing the quantum circuit energy (14), the optimal

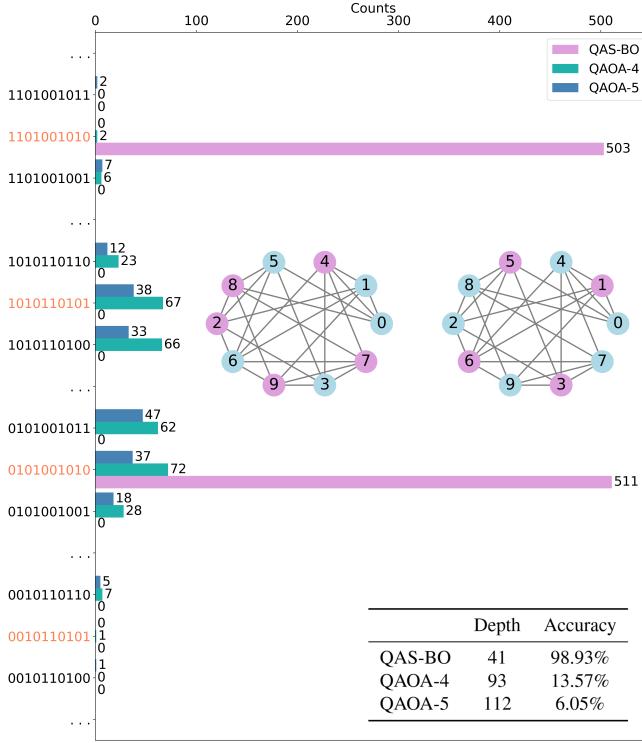


Fig. 5. The unweighted Max-Cut problem (10 nodes):Distribution of quantum states for different algorithms with 1024 measurements

solution of the target problem can be solved.

$$|\psi(\theta)\rangle = \prod_{l=1}^p e^{-i\beta_l H_M} e^{-i\gamma_l H_C} |+\rangle \quad (14)$$

$$E(\theta) = \langle \psi(\theta) | H_C | \psi(\theta) \rangle, \theta = (\beta, \gamma)$$

By utilizing our algorithm to adaptively search for quantum circuits for unweighted Max-Cut problem, weighted Max-Cut problem and Not-All-Equal 3-Satisfiability (NAE 3-SAT), we have successfully designed quantum circuits with lower depth and higher accuracy. The introduction of the problem and the experimental setup can be found in Appendix VIII. In 4-node, 6-node and 8-node problems (Table II), the accuracy of our method reaches 100%, and the required circuit depth is obviously lower than the QAOA algorithm. QAOA performs better on the unweighted Max-Cut problem under the same dimensions. However, its accuracy drops sharply for the weighted Max-Cut problem and the NAE-3SAT problem. This is because its ansatz fails to effectively capture critical information when the weights have large disparities or include negative values.

In the 10-node unweighted Max-Cut problem (Figure 5), the accuracy reaches 98.93%, far exceeding that of QAOA-4 (13.57%) and QAOA-5 (6.05%), while the depth is only 41, significantly lower than that of 93 and 112 of QAOA-4 and QAOA-5.

We only need a few measurements to get the correct results. Of course, our method can make the accuracy reach 100% under the condition of increasing the circuit depth. The lower circuit depth enhances the anti-noise ability of

the circuit and makes it more robust in the actual quantum computing environment. And, the dimension of the search space that our algorithm can accept can be higher.

V. DISCUSSION

The QAS method based on Bayesian Optimization proposed in this research shows excellent performance in dealing with complex quantum circuit design problems. The experimental results show that our method is significantly superior to the traditional method in computational accuracy and circuit depth, especially when dealing with molecular ion systems with strong electron correlation or asymmetric electron distribution. In addition, our method also shows high accuracy and resource efficiency on the combinatorial optimization problem. And compared with ground state energy estimation in quantum chemistry, the acceptable search space dimension of the combinatorial optimization problem can be higher.

The model performance tests of QAS-BO are provided in Appendix X. The appendix provides a detailed description of the comprehensive evaluation process of the QAS-BO model's performance, including the simulation of arbitrary quantum states, the discussion on the BP problem, and research based on different acquisition functions.

In the future, further research on this work can focus on the following aspects:

- The effectiveness of our algorithm will be tested on more VQA problems, and verified by experiments on actual quantum hardwares.
- Based on the characteristics of the quantum computing problem, the search strategy in Bayesian Optimization can be further optimized to improve the search efficiency and the accuracy of the results.
- According to the deployment requirements of actual quantum hardware, the searched quantum circuit can be optimized to ensure that it adapts to the characteristics and limitations of hardware.

ACKNOWLEDGMENT

I sincerely thank my advisor for their valuable guidance on this research and my lab colleagues for their support and assistance during the experiments and discussions.

APPENDIX

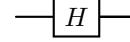
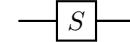
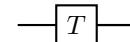
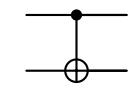
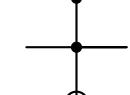
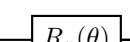
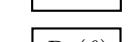
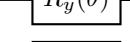
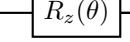
VI. SELECTION OF QUANTUM GATE POOL

The definition of universal quantum gate set requires its completeness, so it can be used to realize arbitrary unitary operation on n qubits. Common universal quantum gate sets mainly include $\mathcal{G} = \{H, S, T, CNOT\}$, $\mathcal{G} = \{H, S, CNOT, Toffoli\}$ and so on. *Toffoli* involves the entanglement of three qubits, increasing the difficulty in deploying architecture search. Therefore, $\mathcal{G} = \{H, S, T, CNOT\}$ is selected as a part of the quantum gate pool.

In order to enhance the flexibility of the architecture and make use of the latest progress of the VQE, we add the commonly used rotation gates $\mathcal{G}_2 = \{Rx(\theta), Ry(\theta), Rz(\theta)\}$ to the quantum gate pool. This extension not only enables our framework to adapt to a wider range of quantum computing tasks, but also facilitates more precise control of the operation of quantum states, which is very important for optimizing the design of quantum circuits for specific algorithm requirements.

Since we regard architecture search as a fixed-scale sampling problem, we need a placeholder unitary operation that does not significantly change the circuit performance. Obviously, the identity gate I is the most suitable choice. As a neutral element, it effectively maintains the structure of the quantum circuit without affecting its computational properties. The Dirac notation representations and gate circuit diagrams of the quantum gates involved in our paper are shown in Table III.

TABLE III: The Dirac notation representations and gate circuit diagrams discussed in our paper [1]

Gate	Dirac notation representation	Gate circuit diagram
H	$\frac{1}{\sqrt{2}}(0\rangle\langle 0 + 0\rangle\langle 1 + 1\rangle\langle 0 - 1\rangle\langle 1)$	
S	$ 0\rangle\langle 0 + i 1\rangle\langle 1 $	
T	$ 0\rangle\langle 0 + e^{i\pi/4} 1\rangle\langle 1 $	
$CNOT$	$ 0\rangle\langle 0 \otimes I + 1\rangle\langle 1 \otimes X$	
$Toffoli$	$ 00\rangle\langle 00 \otimes I + 01\rangle\langle 01 \otimes I + 10\rangle\langle 10 \otimes I + 11\rangle\langle 11 \otimes X$	
$R_x(\theta)$	$\cos \frac{\theta}{2}I - i \sin \frac{\theta}{2}X$	
$R_y(\theta)$	$\cos \frac{\theta}{2}I - i \sin \frac{\theta}{2}Y$	
$R_z(\theta)$	$\cos \frac{\theta}{2}I - i \sin \frac{\theta}{2}Z$	
I	$ 0\rangle\langle 0 + 1\rangle\langle 1 $	

VII. GROUND STATE ENERGY ESTIMATION IN QUANTUM CHEMISTRY EXPERIMENTAL SETUP

In the process of simulating the ground state energy of molecules, those orbits with low energy and inactive in chemical reactions are usually frozen to reduce the demand for computing resources. In this experiment, STO-3G basis set [30] is adopted, and the specific experimental parameters are shown in Table IV and Table V.

TABLE IV
MAPPER AND FROZEN ORBITAL FOR DIFFERENT MOLECULES

	Mapper	Frozen orbit
HeH ⁺	ParityMapper	None
H ₂ ⁺	JordanWignerMapper	None
H ₂	JordanWignerMapper	None
LiH-4	ParityMapper	[0,4,5]
LiH-6	JordanWignerMapper	[0,4,5]

TABLE V
ATOMIC COORDINATES OF DIFFERENT MOLECULES

Coordinate		
HeH ⁺	H (0.000,0.000,0.000)	He (1.000,0.000,0.000)
H ₂ ⁺	H (0.000,0.000,-0.350)	H (0.000,0.000,0.350)
H ₂	H (0.000,0.000,-0.350)	H (0.000,0.000,0.350)
LiH-4	Li (0.000,0.000,0.000)	H (0.000,0.000,2.200)
LiH-6	Li (0.000,0.000,0.000)	H (0.000,0.000,1.595)

Table VI presents the complete data of the ground state energy estimation experiment conducted under different random seeds. The table lists the energy error and corresponding circuit depth for several molecules (HeH⁺, H₂⁺, H₂, LiH-4, and LiH-6).

TABLE VI: Comparison of estimation results for various molecules across different random seeds

		HeH ⁺	H ₂ ⁺	H ₂	LiH-4	LiH-6
Seed 0	Energy error	9.1×10^{-10}	1.2×10^{-9}	3.6×10^{-10}	4.5×10^{-6}	3.1×10^{-4}
	Depth	9	20	18	33	47
Seed 1	Energy error	7.6×10^{-10}	2.9×10^{-9}	3.7×10^{-10}	5.3×10^{-5}	2.5×10^{-4}
	Depth	9	21	17	32	45
Seed 2	Energy error	7.7×10^{-10}	4.5×10^{-9}	7.7×10^{-10}	4.7×10^{-6}	2.1×10^{-4}
	Depth	9	21	16	34	46
Seed 3	Energy error	5.4×10^{-10}	1.1×10^{-9}	9.1×10^{-10}	9.3×10^{-7}	5.6×10^{-4}
	Depth	8	21	17	33	46
Seed 4	Energy error	1.2×10^{-9}	9.8×10^{-10}	2.5×10^{-10}	5.5×10^{-6}	2.4×10^{-4}
	Depth	8	19	16	33	45
Seed 5	Energy error	6.1×10^{-10}	3.5×10^{-9}	9.4×10^{-11}	6.9×10^{-7}	6.3×10^{-5}
	Depth	9	19	18	35	43
Seed 6	Energy error	3.6×10^{-10}	7.7×10^{-9}	2.3×10^{-10}	1.9×10^{-6}	3.4×10^{-4}
	Depth	9	21	16	32	45
Seed 7	Energy error	4.5×10^{-10}	4.4×10^{-9}	3.5×10^{-10}	1.7×10^{-6}	9.2×10^{-5}
	Depth	9	22	16	33	47
Seed 8	Energy error	8.2×10^{-10}	1.5×10^{-9}	3.4×10^{-10}	2.3×10^{-6}	3.4×10^{-4}
	Depth	11	20	17	33	44
Seed 9	Energy error	6.7×10^{-10}	1.2×10^{-10}	5.2×10^{-10}	7.1×10^{-6}	1.9×10^{-4}
	Depth	9	20	17	31	45

VIII. COMBINATORIAL OPTIMIZATION PROBLEM AND EXPERIMENTAL SETUP

Combinatorial optimization is an important field in mathematical optimization. The solution space of this kind of problem is limited (usually discrete), but with the increase of the scale of the problem, the size of the solution space usually increases exponentially. The goal of the problem is to find a solution that satisfies the constraints and is optimal among all possible solutions. Combinatorial optimization problems are widely used in various fields, including scheduling, routing planning, resource allocation, network design and so on.

A. The Max-Cut Problem

The Max-Cut problem is a classical combinatorial optimization problem with widespread applications in many fields, such as computer science, physics, and network science. The problem can be expressed as follows: Given an undirected graph $G = (V, E)$, where V is the set of vertices and E is the set of edges, and each edge $e \in E$ has a weight $w(e) \geq 0$. The goal is to partition the vertex set V into two subsets V_1 and V_2 , such that the sum of the weights of the edges crossing these two subsets is maximized.

The Max-Cut problem can be formulated as follows (15):

$$\text{Maximize} \quad \sum_{(i,j) \in E} w_{ij} \cdot \delta(x_i, x_j) = \sum_{(i,j) \in E} \frac{(1 - x_i x_j) w_{ij}}{2} \quad (15)$$

where w_{ij} is the weight of edge (i, j) . x_i, x_j are the subset labels of nodes i and j , $x_i, x_j \in \{-1, 1\}$. $\delta(x_i, x_j) = 1$ if $x_i \neq x_j$, otherwise $\delta(x_i, x_j) = 0$.

The objective Hamiltonian of the Max-Cut problem can be expressed as Equation 16. The derivation of the equivalence between Equation (15) and Equation (16) can be found in Appendix IX.

$$H_C = - \sum_{(i,j) \in E} \frac{(I - Z_i Z_j) w_{ij}}{2} \quad (16)$$

In this paper, the experiments on the unweighted Max-Cut problem ($w(e) = 1$) were conducted using the graphs shown in Figures 6 to 8.

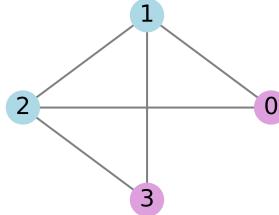


Fig. 6. The unweighted Max-Cut problem (4 nodes)

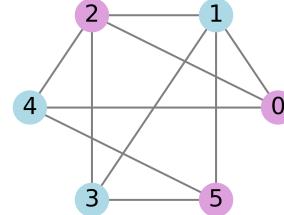


Fig. 7. The unweighted Max-Cut problem (6 nodes)

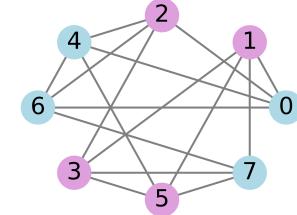


Fig. 8. The unweighted Max-Cut problem (8 nodes)

Next, we extend the unweighted Max-Cut problem to the weighted Max-Cut problem ($w(e) \geq 0$) and conduct experiments using the graphs shown in Figures 9 to 11.

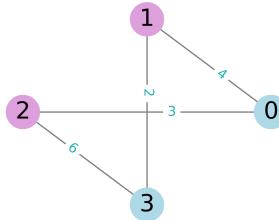


Fig. 9. The weighted Max-Cut problem (4 nodes)

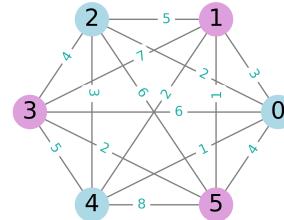


Fig. 10. The weighted Max-Cut problem (6 nodes)

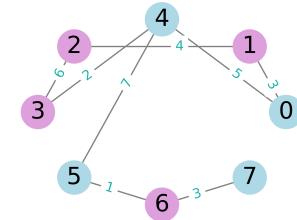


Fig. 11. The weighted Max-Cut problem (8 nodes)

B. The Not-All-Equal 3-Satisfiability Problem

The Not-All-Equal 3-Satisfiability (NAE 3-SAT) problem is a variant of the 3-Satisfiability (3-SAT) problem. 3-SAT is an NP-hard problem. Its definition is that : Given a Boolean expression, where the Boolean expression is a conjunction of multiple clauses, and each clause is the disjunction of three literals (or their negations), determine whether it is satisfiable. The following is an example of a 3-SAT Boolean expression (17):

$$(x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_1 \vee x_4 \vee x_5) \wedge (x_2 \vee \neg x_4 \vee \neg x_5) \quad (17)$$

where $(x_1 \vee x_2 \vee \neg x_3)$ is a clause. x_1 is a literal, and $\neg x_5$ is the negation of literal.

NAE 3-SAT requires that the three literals in each clause are not all identical. In other words, at least one literal must be true, and at least one literal must be false. NAE 3-SAT remains the NP-complete problem, which can be proven through a reduction from 3-SAT [?].

NAE 3-SAT problem can be represented using a spin model. The set of clauses in NAE 3-SAT be denoted as C . Each clause consists of three literals, and each literal is represented by a spin. The spin pointing upward ($s = 1$, qubit = 0) corresponds to the literal being false, while the spin pointing downward ($s = -1$, qubit = 1) corresponds to the literal being true. For any clause $(s_i, s_j, s_k) \in C$, the spins are constrained such that they cannot all be 1 or all be -1 simultaneously.

The Hamiltonian for the NAE 3-SAT problem is defined in Equation 18.

$$\begin{aligned} H_C &= \sum_{(i,j,k) \in C} \frac{[(s_i + s_j + s_k)^2 - 1]}{2} \\ &= \sum_{(i,j,k) \in C} (s_i s_j + s_j s_k + s_k s_i) + |C| \end{aligned} \quad (18)$$

Where $|C|$ denotes the number of clauses. When all clauses are satisfied, the Hamiltonian H_C reaches its minimum value of 0, and the corresponding qubit string represents the solution to the NAE 3-SAT problem.

The experiments in this section are shown in Table VII.

TABLE VII
NAE 3-SAT EXPERIMENTAL SETUP

Literal number	NAE 3-SAT
4	$(x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_3) \wedge$ $(x_1 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee \neg x_3 \vee \neg x_4) \wedge$ $(x_2 \vee x_3 \vee \neg x_4)$
6	$(x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_4) \wedge$ $(x_1 \vee \neg x_3 \vee x_5) \wedge (\neg x_1 \vee x_3 \vee \neg x_5) \wedge$ $(x_2 \vee \neg x_4 \vee x_6) \wedge (\neg x_2 \vee x_4 \vee \neg x_6) \wedge$ $(\neg x_3 \vee \neg x_5 \vee x_6) \wedge (x_3 \vee x_5 \vee \neg x_6) \wedge$ $(\neg x_1 \vee \neg x_4 \vee \neg x_6) \wedge (x_1 \vee x_4 \vee \neg x_5) \wedge$ $(x_2 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_5 \vee x_6) \wedge$ $(\neg x_1 \vee \neg x_3 \vee x_6)$
8	$(x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee x_4) \wedge$ $(x_1 \vee \neg x_3 \vee x_5) \wedge (\neg x_1 \vee x_3 \vee \neg x_5) \wedge$ $(x_2 \vee \neg x_4 \vee x_6) \wedge (\neg x_2 \vee x_4 \vee \neg x_6) \wedge$ $(\neg x_3 \vee \neg x_5 \vee x_7) \wedge (x_3 \vee x_5 \vee \neg x_7) \wedge$ $(x_1 \vee \neg x_2 \vee \neg x_8) \wedge (\neg x_1 \vee \neg x_6 \vee x_8) \wedge$ $(\neg x_4 \vee x_7 \vee \neg x_8) \wedge (x_2 \vee \neg x_3 \vee x_6) \wedge$ $(\neg x_7 \vee x_4 \vee \neg x_5) \wedge (x_1 \vee \neg x_8 \vee x_3) \wedge$ $(\neg x_1 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_6 \vee \neg x_7) \wedge$ $(x_5 \vee \neg x_7 \vee \neg x_8)$

IX. EQUIVALENCE DERIVATION OF THE MAX-CUT PROBLEM CLASSICAL TO QUANTUM TRANSFORMATION

Given an undirected graph $G = (V, E)$, divide its vertex set into two disjoint subsets $V = \{V_0, V_1\}$, so as to the sum of the weights of the edges spanning across the two subsets is maximized. If the node $v \in V_0$, it is denoted by the label -1, otherwise it is denoted by 1, thus defining the mapping from the node to the set $\{V_0, V_1\}$.

The transformation from classical to quantum is that the point labeled -1 corresponds to $|1\rangle$, and the point labeled 1 corresponds to $|0\rangle$ (19), then the objective function is mapped from $c(x)$ to $c(z)$ (20).

$$\begin{cases} x_i = -1 \rightarrow 1 \\ x_i = 1 \rightarrow 0 \end{cases} \iff x_i = (-1)^{z_i} \quad (19)$$

$$c(x) = - \sum_{(i,j) \in E} \frac{(1 - x_i x_j) w_{ij}}{2}, x_i \in \{-1, 1\} \quad (20)$$

$$\iff c(z) = - \sum_{(i,j) \in E} \frac{(1 - (-1)^{z_i + z_j}) w_{ij}}{2}, z_i \in \{0, 1\}$$

Then, replace each z_i in $c(z)$ with the Pauli Z gate acting on the i -th qubit, and construct the cost Hamiltonian H_C (21) of the quantum system.

$$H_C = - \sum_{(i,j) \in E} \frac{(I - Z_i Z_j) w_{ij}}{2} \quad (21)$$

Therefore, the Max-Cut problem is transformed into the form of finding the minimum eigenvalue $c(z_0)$ of the large sparse diagonal matrix H_C . The detailed derivation can be found in (22) and (23).

$$\begin{aligned} H_C |z\rangle &= - \sum_{i,j} \frac{(I - Z_i Z_j) w_{ij}}{2} |z_1, \dots, z_i, \dots, z_j\rangle \\ &= - \sum_{i,j} \frac{w_{ij}}{2} (|z_1, \dots, z_i, \dots, z_j\rangle \\ &\quad - (-1)^{z_i + z_j} |z_1, \dots, z_i, \dots, z_j\rangle) \\ &= - \sum_{i,j} \frac{w_{ij}}{2} (1 - (-1)^{z_i + z_j}) |z\rangle, z_i \in \{0, 1\} \\ &= c(z) |z\rangle \end{aligned} \quad (22)$$

Because of the particularity of the construction of H_C , only the diagonal element is not 0. However, in the classical computing framework, the time complexity of solving this problem is $O(2^n)$. In quantum computing, the solution to such problems is based on Theorem 1, namely the Variational Principle. It transforms the goal of calculating the minimum eigenvalue $c(z_0)$ of a large matrix into the ground state energy E_0 (23) of the Hamiltonian of the system by using quantum circuits. The advantage of this method lies in its potential for superpolynomial speedup.

Theorem 1. (Variational Principle) For a physical system with a time-independent Hamiltonian H , when any normalized trial wave function ψ is used to calculate the expectation value of the Hamiltonian $\langle\psi|H|\psi\rangle$, this expectation value is always greater than or equal to the ground state energy E_0 [31].

$$\langle H \rangle = \langle \psi | H | \psi \rangle \geq E_0$$

$$\begin{aligned} H_c |z_0\rangle &= c(z_0) |z_0\rangle \\ \langle z_0 | H_c | z_0 \rangle &= \langle z_0 | c(z_0) | z_0 \rangle \\ E_0 &= \langle z_0 | H_c | z_0 \rangle = c(z_0) \langle z_0 | z_0 \rangle = c(z_0) \\ E(\theta) &= \langle \psi(\theta) | H_C | \psi(\theta) \rangle \geq E_0 \end{aligned} \quad (23)$$

X. PERFORMANCE TESTING OF THE QAS-BO MODEL

A. Simulation of Arbitrary Quantum States

In this section, we validate the performance of the model by simulating arbitrary quantum states. We use the Python library `numpy` to randomly generate quantum states. The generation process involves creating a complex vector with random real and imaginary parts and normalizing it to ensure that the generated quantum states conform to the physical requirements of quantum mechanics. Additionally, we use the `Statevector` class from the `Qiskit` library to encapsulate the generated quantum states.

We use fidelity as a metric to evaluate the performance of the model in simulating arbitrary quantum states. Fidelity quantifies the similarity between two quantum states, with values ranging from 0 to 1. Fidelity close to 1 indicates that the model performs well in accurately simulating quantum states. The fidelity F between the target state $|\psi\rangle$ and the generated state $|\phi\rangle$ is defined as (24):

$$F(\psi, \phi) = |\langle \psi | \phi \rangle|^2 \quad (24)$$

where $\langle \psi | \phi \rangle$ is the inner product of the two quantum states.

In our experiments, we compared our model with RL, DQAS, and the fixed-architecture HE model. The depth and accuracy rate of the models are summarized in Table VIII, and the variation of fidelity with the number of iterations is shown in Figure 12.

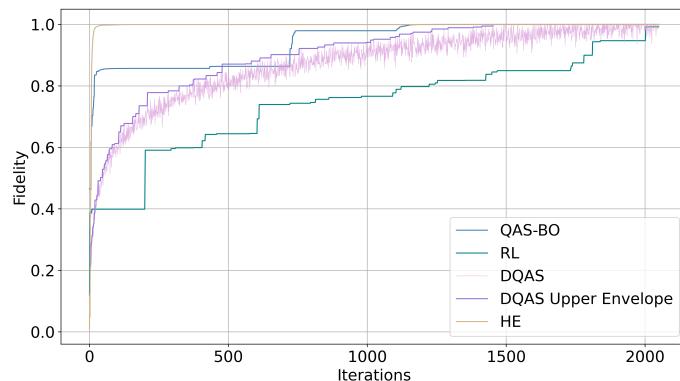


Fig. 12. 4-Qubit: Comparison of fidelity with iterations

TABLE VIII
4-QUBIT: COMPARISON OF DEPTH AND ACCURACY RATE

	Depth	Accuracy
QAS-BO	16	7.1×10^{-9}
RL	19	4.6×10^{-8}
DQAS	18	2.8×10^{-9}
HE	16	4.6×10^{-5}

According to the results in Figure 12 and Table VIII, QAS-BO model demonstrates superior fidelity during the iteration process, achieving high fidelity with a relatively shallow circuit depth. RL requires more computational resources, and its iteration process is relatively slow. Although DQAS achieves the highest fidelity, it requires more iterations to reach the optimal solution. The HE model, due to its fixed architecture, can quickly reach the optimal solution but exhibits the lowest fidelity performance.

B. Discussion on the Barren Plateau Problem

The Barren Plateau (BP) problem refers to the phenomenon where, during the training of PQCs, the gradient of the loss function approaches zero across the entire parameter space as the number of qubits and the circuit depth increase. This makes the optimization process nearly unable to progress and prevents the optimization algorithm from identifying the correct direction of adjustment, ultimately resulting in the model being unable to train effectively.

QAS-BO effectively mitigates the BP problem through improved circuit structure, adaptive optimization algorithms, and randomized strategies. QAS-BO is capable of generating quantum circuits with lower depth. By optimizing circuit structure, it reduces the number of parameters to be optimized, thereby mitigating the risk of gradient vanishing. The adaptive optimization algorithm enables QAS-BO to achieve significant progress even in the presence of small gradients, avoiding reliance on gradients. Additionally, the algorithm efficiently navigates complex search spaces, leverages historical information for decision-making, balances exploration and exploitation, and predicts unexplored regions, making it suitable for addressing high-dimensional and complex problems. Moreover, by introducing random perturbations during the optimization process, QAS-BO can explore new pathways within the barren plateau, effectively avoiding being trapped in local optima.

In contrast, benchmark methods such as Q-DARTS, DQAS, and QCAS rely on gradient-based optimization algorithms, which inevitably encounter the BP problem at higher precision levels. RL is gradient-free, but it requires significant resources and can only achieve the desired accuracy after extensive training, making it relatively inefficient. QAS-BO selects the sampling probabilities of quantum gates through heuristic circuit search and Gaussian distribution adjustment. This gradient-free sampling method does not rely on specific domain knowledge, enabling efficient automated quantum circuit design while optimizing resource utilization and computational efficiency.

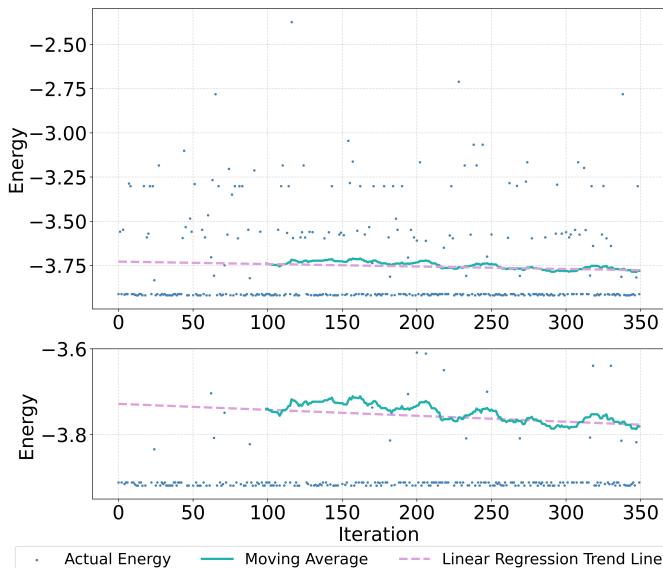


Fig. 13. HeH⁺ Estimation: With the increase of iterations, the value of the objective function tends to be stable and tends to decrease, which shows that our exploration is effective.

C. Research Based on Different Acquisition Functions

In the design of model architecture optimization for QAS-BO, comparative experiments were conducted on different sampling functions to select an appropriate optimization strategy.

Figure 14 illustrates the convergence of the objective function $\min(f(x))$ with respect to the number of calls n . The figure compares the performance of four different acquisition functions: LCB, EI, PI, and Adaptive acquisition function strategy. The x-axis represents the number of calls to the optimization function, and the y-axis represents the current best objective function value. The red dashed line represents the theoretical minimum of the objective function.

The adaptive acquisition strategy based on the multi-armed bandit principle demonstrates faster and more stable exploration capabilities in the optimization process of the objective function. In contrast, EI tends to perform fine-grained searches near the known optimum, focusing on local exploration. PI is inclined to explore potential but insufficiently tested areas, exhibiting stronger exploratory behavior but relatively lower efficiency. LCB achieves a balance between exploration and exploitation, combining the exploration of unknown areas with the utilization of known information during the search process. By combining the advantages of these three acquisition functions, adaptive acquisition function strategy can efficiently explore the potential optimal solution and quickly approach the global minimum.

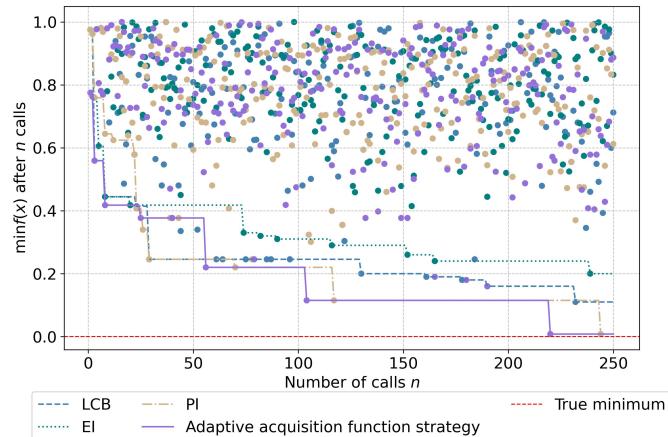


Fig. 14. Comparison of acquisition functions in objective function optimization

REFERENCES

- [1] M. A. Nielsen and I. L. Chuang, “Quantum Computation and Quantum Information,” Cambridge: Cambridge University Press, Jun. 2012, doi: 10.1017/cbo9780511976667.
- [2] Peruzzo, A., McClean, J., Shadbolt, P. et al.“ A variational eigenvalue solver on a photonic quantum processor,” Nat Commun 5, 4213 (2014). <https://doi.org/10.1038/ncomms5213>.
- [3] Arrasmith, Andrew, Marco Cerezo, Piotr Czarnik, Lukasz Cincio and Patrick J. Coles. “Effect of barren plateaus on gradient-free optimization,” Quantum 5 (2020): 558.
- [4] Mateusz Ostaszewski, Lea M. Trenkwalder, Wojciech Masarczyk, Eleanor Scerri, and Vedran Dunjko. 2021. “Reinforcement learning for optimization of variational quantum circuit architectures,” In Proceedings of the 35th International Conference on Neural Information Processing Systems (NIPS ’21). Curran Associates Inc., Red Hook, NY, USA, Article 1391, 18182–18194.
- [5] Anschuetz, Eric and Zanoci, Cristian. (2019). “Near-term quantum-classical associative adversarial networks,” Physical Review A. 100. 10.1103/PhysRevA.100.052327.
- [6] Lamata, Lucas, Unai Alvarez-Rodriguez, José David Martín-Guerrero, Mikel Sanz and Enrique Solano. “Quantum autoencoders via quantum adders with genetic algorithms.” Quantum Science and Technology 4 (2017): n. pag.
- [7] Dong, Yumin and Xie, Jianshe and Hu, Wanbin and Liu, Cheng and Luo, Yi. “Variational algorithm of quantum neural network based on quantum particle swarm,” Journal of Applied Physics. 132. 104401. 10.1063/5.0098702.
- [8] M. H. Cheng and K. E. Khosla and C. N. Self and M. Lin and B. X. Li and A. C. Medina and M. S. Kim. “Clifford Circuit Initialisation for Variational Quantum Algorithms,” 10.48550/arXiv.2207.01539.
- [9] Zhang, Shi-Xin and Hsieh, Chang-Yu and Zhang, Shengyu and Yao, Hong, “Differentiable quantum architecture search,” , vol. 7, no. 4, p. 045023, Aug. 2022, doi: 10.1088/2058-9565/ac87cd.
- [10] Wenjie Wu, Ge Yan, Xudong Lu, Kaisen Pan, and Junchi Yan. “QuantumDARTS: differentiable quantum architecture search for variational quantum algorithms,” ICML’23, Vol. 202. JMLR.org, Article 1573, 37745–37764.
- [11] Duong, Trong, Sang T. Truong, Minh Tam, Bao Gia Bach, Juno Ryu and June-Koo Kevin Rhee. “Quantum Neural Architecture Search with Quantum Circuits Metric and Bayesian Optimization,” ArXiv abs/2206.14115 (2022): n. pag.
- [12] K. A. Nicoli, C. J. Anders, L. Funcke, T. Hartung, K. Jansen, S. Kühn, K.-R. Müller, P. Stornati, P. Kessel, and S. Nakajima, “Physics-informed Bayesian optimization of variational quantum circuits,” in *Proc. NeurIPS*, 2023.
- [13] Z. Dai, G. K. R. Lau, A. Verma, Y. Shu, B. K. H. Low, and P. Jaillet, “Quantum Bayesian optimization,” 2023.
- [14] M. J. D. Powell, “A direct search optimization method that models the objective and constraint functions by linear interpolation,” in *Proc. Conf. Optimization Techniques and Applications*, 1994.
- [15] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, “Elementary gates for quantum computation,” *Phys. Rev. A*, vol. 52, no. 5, pp. 3457–3467, Nov. 1995, doi: 10.1103/PhysRevA.52.3457.

- [16] T. Alexander, N. Kanazawa, D. J. Egger, L. Capelluto, C. J. Wood, A. Javadi-Abhari, and D. C. McKay, "Qiskit pulse: Programming quantum computers through the cloud with pulses," vol. 5, no. 4, p. 044006, Aug. 2020, doi: 10.1088/2058-9565/aba404.
- [17] C. E. Rasmussen and C. K. I. Williams, "Gaussian Processes for Machine Learning," The MIT Press, Nov. 2005, doi: 10.7551/mitpress/3206.001.0001.
- [18] J. Nocedal, "Updating quasi-Newton matrices with limited storage," vol. 35, no. 151, p. 773, 1980.
- [19] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," vol. 45, pp. 503–528, 1989. Available: <https://api.semanticscholar.org/CorpusID:5681609>.
- [20] D. C. Liu and J. Nocedal, "Maximizing acquisition functions for Bayesian optimization," Neural Information Processing Systems (2018).
- [21] M. Hoffman, E. Brochu, and N. de Freitas, "Portfolio allocation for Bayesian optimization," In Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence (UAI'11). AUAI Press, Arlington, Virginia, USA, 327–336.
- [22] Various authors, *Qiskit Textbook*, GitHub, 2023. [Online]. Available: <https://github.com/Qiskit/textbook>.
- [23] R. J. Bartlett, S. A. Kucharski, and J. Noga, "Alternative coupled-cluster ansätze II. The unitary coupled-cluster method," Chemical Physics Letters 155 (1989): 133-140.
- [24] R. J. Bartlett, S. A. Kucharski, and J. Noga, "Alternative coupled-cluster ansätze II. The unitary coupled-cluster method," Chemical Physics Letters 190 (1992): 1-12.
- [25] M. Cerezo, A. Arrasmith and R. Babbush et al., "Variational quantum algorithms," Nat Rev Phys 3, 625–644 (2021). <https://doi.org/10.1038/s42254-021-00348-9>.
- [26] H. R. Grimsley, D. Claudino, S. E. Economou, E. Barnes, and N. J. Mayhall, "Is the Trotterized UCCSD ansatz chemically well-defined?," Journal of chemical theory and computation, 16(1), 1–6. <https://doi.org/10.1021/acs.jctc.9b01083>.
- [27] Y. Du, T. Huang, S. You, M.-H. Hsieh, and D. Tao, "Quantum circuit architecture search for variational quantum algorithms," npj Quantum Inf 8, 62 (2022). <https://doi.org/10.1038/s41534-022-00570-y>.
- [28] A. Ash-Saki, M. A. Alam, and S. Ghosh, "Study of decoherence in quantum computers: A circuit-design perspective," ArXiv abs/1904.04323 (2019).
- [29] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," arXiv: Quantum Physics (2014).
- [30] W. J. Hehre, R. F. Stewart, and J. A. Pople, "Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals," J. Chem. Phys. 51, 2657 (1969). 10.1063/1.1672392.
- [31] C. Cohen-Tannoudji, B. Diu, and F. Laloe, Quantum Mechanics, Volume 1 (Hermann, 2014).