

# Enhancing Quantum Approximate Optimization with CNN-CVaR Integration

**Pengnian Cai**

Anhui University

**Kang Shen**

Anhui University

**Tao Yang**

Anhui University

**Yuanming Hu**

Anhui University

**Bin Lv**

Anhui University

**Liuhan Fan**

Anhui University

**Zeyu Liu**

Anhui University

**Qi Hu**

Anhui University

**Yunlai Zhu**

Anhui University

**Zuheng Wu**

Anhui University

**Yuehua Dai**

Anhui University

**Fei Yang**

Anhui University

**Jun Wang**

Anhui University

**Zuyu Xu**

[zyxu@ahu.edu.cn](mailto:zyxu@ahu.edu.cn)

Anhui University

**Keywords:** Convolutional Neural Network, Conditional Value at Risk, QAOA optimization

**Posted Date:** June 5th, 2024

**DOI:** <https://doi.org/10.21203/rs.3.rs-4460928/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Quantum Information Processing on January 24th, 2025. See the published version at <https://doi.org/10.1007/s11128-025-04655-3>.

# Enhancing Quantum Approximate Optimization with CNN-CVaR Integration

Pengnian Cai<sup>1</sup>, Kang Shen<sup>1</sup>, Tao Yang<sup>1</sup>, Yuanming Hu<sup>1</sup>,  
Bin Lv<sup>1</sup>, Liuhuan Fan<sup>1</sup>, Zeyu Liu<sup>1</sup>, Qi Hu<sup>1</sup>, Yunlai Zhu<sup>1</sup>,  
Zuheng Wu<sup>1</sup>, Yuehua Dai<sup>1</sup>, Fei Yang<sup>1</sup>, Jun Wang<sup>1\*</sup>, Zuyu Xu<sup>1\*</sup>

<sup>1</sup>School of Integrated Circuits, Anhui University, Hefei, Anhui, 230601,  
China.

\*Corresponding author(s). E-mail(s): [iamwj7@163.com](mailto:iamwj7@163.com);  
[zyxu@ahu.edu.cn](mailto:zyxu@ahu.edu.cn);

## Abstract

The Quantum Approximate Optimization Algorithm (QAOA) represents a promising approach for tackling combinatorial optimization challenges on near-term quantum devices. Central to QAOA optimization is the minimization of the expectation of the problem Hamiltonian for parameterized trial quantum states, which motivates the exploration of advanced optimization techniques. In this study, we propose a novel combinatorial optimization strategy, CNN-CVaR-QAOA, which integrates a Convolutional Neural Network (CNN) with Conditional Value at Risk (CVaR) to optimize QAOA circuits. By replacing the traditional loss function with CVaR and leveraging CNN for variational quantum parameter optimization, we demonstrate the superior efficacy of CNN-CVaR-QAOA through experimental validation on Erdos-Renyi random graphs. Our results show better solutions across various graph configurations. Furthermore, we investigate the influence of the CVaR parameter ( $\alpha$ ) on algorithm performance, revealing that lower  $\alpha$  values lead to smoother objective functions and improved approximation ratios. This work indicates that CNN-CVaR-QAOA offers significant advantages in optimizing QAOA parameters, particularly in the context of Near-Term Intermediate-Scale Quantum era, highlighting its potential to enhance QAOA optimization efforts across diverse optimization domains.

**Keywords:** Convolutional Neural Network, Conditional Value at Risk, QAOA optimization

# 1 Introduction

With the advent of Near-Term Intermediate-Scale Quantum (NISQ) computing, the landscape of computational problem-solving is undergoing a profound transformation. Quantum computers are emerging as powerful tools poised to surpass classical counterparts in tackling complex real-world problems. Combinatorial optimization challenges, including Max Cut, Graph Coloring, Traveling Salesman, and Scheduling Management, find their solutions through Ising Hamiltonians, thus making them ripe for quantum computational approaches [1, 2, 3, 4].

Hybrid quantum-classical algorithms, particularly the Quantum Approximate Optimization Algorithm (QAOA), stand out as promising methodologies in this arena. By leveraging the principles of variation, QAOA offers an avenue to approximate solutions for problems encoded by Hamiltonians [5]. Recent advancements in both experimental implementations and theoretical underpinnings have propelled QAOA into the limelight [4, 6, 7, 8, 9, 10, 11, 12, 13, 14]. The tantalizing prospect of quantum advantage over classical algorithms beckons, with evidence mounting to support its universality and computational prowess [8, 15, 16, 17, 18].

Despite these advancements, the QAOA still faces several limitations and potential issues hindering its effectiveness. Notably, the trade-off between circuit depth and performance, constrained by coherence times in existing and near-term quantum processors, poses a significant challenge [19, 20, 21, 22]. Additionally, the complexity introduced by the QAOA ansatz, with its increasing number of variational parameters, presents challenges for classical optimizers [23, 24, 25]. Previous studies have shown that the optimal QAOA parameters exhibit specific patterns [26, 27], leading to depth-sequential strategies and machine learning-based methods for parameter initialization [10, 23, 28, 29, 30]. For instance, Alam et al. adopted a regression model to predict high-depth parameters from low-depth ones [26], while Amosy et al. applied iterative neural networks for parameter initialization, selecting the most promising parameter sets from cluster centers for the given problem [31]. However, existing machine learning approaches exhibit limitations in universality, necessitating numerous models for different problem sizes or QAOA depths, resulting in higher costs for preparing extensive training data for deeper QAOAs [26, 28]. Moreover, the reliance on expected value mean as the objective function may impede the optimization process [6, 32]. Therefore, it is crucial to further explore effective methods for dynamically controlling the expressiveness and trainability of QAOA for a given combinatorial problem to achieve optimal performance.

In this study, we propose a novel method, CNN-CVaR-QAOA, which synergistically integrates a classical convolutional neural network (CNN) with Conditional Value at Risk (CVaR) aggregation functions to effectively optimize variational parameters, thereby augmenting the performance of the QAOA. The adoption of the CNN architecture stems from the intrinsic similarities between pixel properties in images and QAOA parameters, such as their continuity and correlation among adjacent values. By employing CVaR instead of minimizing the expected value, our model iteratively optimizes by minimizing the CVaR, thereby aligning more closely with the objective of solving combinatorial optimization problems. Comparative evaluations

conducted on Max-Cut problem instances on Erdos-Renyi random graphs of different sizes demonstrate that our proposed CNN-CVaR-QAOA method outperforms alternative approaches, particularly at lower depths. This innovative framework not only advances the frontier of quantum optimization but also underscores the potential of synergistic integration between classical and quantum techniques in addressing complex computational challenges.

The paper is structured as follows: Section 2 provides an introduction to the Quantum Approximate Optimization Algorithm. In Section 3, we outline the implementation details of the CNN-CVaR method. Section 4 elaborates on the numerical and experimental study, followed by a conclusion where potential future steps are discussed.

## 2 Quantum Approximate Optimization Algorithm

The Quantum Approximate Optimization Algorithm (QAOA) stands as a pivotal solution tailored for tackling combinatorial optimization (CO) dilemmas, leveraging a blend of quantum and classical methodologies. Illustrated in Fig. 1, QAOA harnesses a parameterized quantum circuit denoted as  $U(\theta)$ , colloquially referred to as a variational form, to produce trial wavefunctions. The architecture of a p-layer circuit crafts states through the expression:

$$|\phi(\gamma, \beta)\rangle = \left( \prod_{p=1}^p e^{-i\beta_p H_M} e^{-i\gamma_p H_C} \right) |+\rangle^{\otimes N}, \quad (1)$$

where  $|+\rangle$  represents the eigenstate, the problem Hamiltonian  $H_C$  is used to encode the total energy of the system, and  $H_M$  is known as the mixer Hamiltonian, given by  $H_M = \sum_{n=1}^N \delta_n^X$ , where  $\delta_n^X$  is the Pauli-X operator acting on the qubit n [33]. Finally, the p-dimensional vectors  $\beta$  and  $\gamma$  serve as the variational parameters corresponding to  $H_M$  and  $H_C$ , respectively.

The QAOA operates by measuring the parameterized quantum state to derive a solution, with its parameter adjustment process guided by a classical optimization algorithm. The primary objective is to identify the optimal variational parameters ( $\beta^*$ ,  $\gamma^*$ ) that minimize the objective function  $F(z)$ :

$$F(z) = \langle \phi(\gamma, \beta) | H_C | \phi(\gamma, \beta) \rangle. \quad (2)$$

This study predominantly addresses the Max-Cut problem. In formal terms, let  $G(V, E)$  represent an undirected graph with a vertex set  $V = N$ , and  $E$  representing the edge set. The problem attempts to partition the nodes of a graph  $G$  into two sets  $[+1, -1]$ , maximizing the number of edges between nodes from these two sets. The optimal solution  $z^*$  for Max-Cut maximizes the objective function  $C(z)$ :

$$C(z) = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} (1 - (-1)^{z_i} (-1)^{z_j}), \quad (3)$$

where  $z$  represents a bit string. Theoretical studies have proven that determining  $z^*$  is an NP-hard problem, necessitating approximation techniques in most scenarios. A commonly used solver for Max-Cut is the Goemans-Williamson (GW) algorithm, which employs Semidefinite Programming (SDP) methodologies to achieve an approximation ratio of 0.879 [34].

To tackle combinatorial problems using physical systems, it's imperative to map the problem onto a Hamiltonian, where the ground state corresponds to the optimal solution [35]. In the case of the Max-Cut problem with  $n$  nodes, the cost Hamiltonian  $H_C$  is constructed using the Pauli-Z operators as follows:

$$H_C = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} (I - \delta_i^Z \delta_j^Z). \quad (4)$$

Here, we simplify by setting all edge weights to 1, allowing us to confine each  $\gamma$  and  $\beta$  within the intervals  $[0, \pi]$  and  $[0, \pi/2]$  respectively. For the Max-Cut problem, the commonly used performance metric for QAOA is the approximation ratio  $R$ , which is defined as:

$$R = \frac{F(z)}{C_{max}}, \quad (5)$$

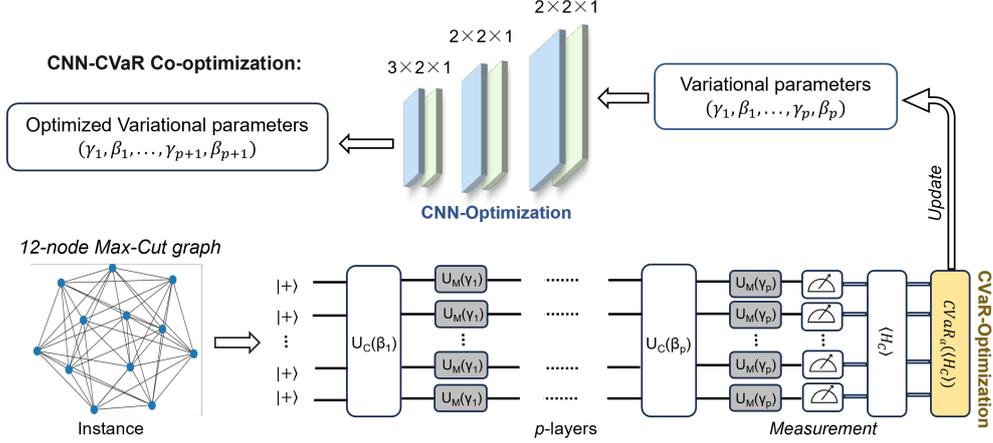
where  $C_{max}$  is the maximum cut of the graph.

### 3 CNN-CVaR Optimization

In this section, we propose to utilize CNN as a classical network parameter prediction architecture to achieve more efficient optimization of QAOA[36], with CVaR as the objective function[37]. The architecture of the proposed CNN model and the aggregation function for CVaR optimization are elucidated below.

Our CNN model aims to learn a mapping function from the parameters of  $p$ -depth QAOA to the parameters of  $p+1$ -depth. The  $p$ -depth and  $p+1$ -depth parameters are denoted as tensors of size  $1 \times 2 \times p$  and  $1 \times 2 \times p+1$ , respectively. We initialize the input by randomly setting the parameters of a QAOA circuit with a depth of 1 using the QN-SPSA optimizer, a stochastic optimizer belonging to the family of gradient descent methods [38]. As shown in Fig. 1, our CNN model architecture comprises two key components: up-sampling and down-sampling. The up-sampling segment of our CNN model is responsible for extracting features from the input QAOA parameters. This segment consists of two convolutional layers, both employing a stride of 1 and ReLU activation function [39] as the activation function. Each layer utilizes a  $2 \times 2$  kernel and provides one zero-padding for all four inputs. The first layer expands the input dimensions to 16, which are further augmented to 64 dimensions by the second layer. On the other hand, the down-sampling section involves simplifying the extracted features. Here, a single convolutional layer with a  $64 \times 3 \times 2$  filter is employed to reduce the dimensions from 64 to 1. This results in the final output of the CNN network, which produce the final optimized variational parameters for the QAOA circuit.

The objective of CVaR is to prioritize improving the best outcomes observed, rather than solely relying on the sample expected value as the traditional optimization objective function. The method of CVaR that we adopt involves selecting a set



**Fig. 1** A schematic diagram illustrating a p-level quantum approximate optimization algorithm along with the utilization of CNN-CVaR optimization. An example showcasing the Max-Cut problem on a 12-node graph, with a quantum circuit taking input  $|+\rangle$ . Alternate application of  $U_C(\beta)$  and  $U_M(\gamma)$ , followed by the measurement of the final state to obtain the expected value relative to the objective function  $H_C$ . Subsequently, CVaR optimization is employed to replace the objective function, and the resulting output is fed into a convolutional neural network to determine the optimal parameters  $\gamma$  and  $\beta$ .

of observation results within measurements, where the sample  $H_k$  is sorted in non-decreasing order, taking  $[H_1, \dots, H_k]$  as the objective function [40]. Simultaneously, the CVaR parameter  $\alpha \in (0, 1]$  is introduced, where  $\alpha$  represents further calculating the expected value within the  $\alpha$  proportion of  $[H_1, \dots, H_k]$ , thereby smoothing the objective function. The CVaR function, denoted as  $CVaR_\alpha$ , is defined as follows:

$$CVaR_\alpha = \frac{1}{\alpha K} \sum_{k=0}^{\alpha K} H_K, \quad (6)$$

where  $\alpha=1$  corresponds to the sample mean expected value. Decreasing  $\alpha$  biases the selection of observation samples increasingly towards the best outcomes. It's crucial to note that employing CVaR optimization modifies our objective function (2) to:

$$\min(CVaR_\alpha(F(z))), \quad (7)$$

$CVaR_\alpha$  considers only a subset of measurements, leading to an estimator variance of  $O(1/(K\alpha^2))$ , where  $K$  is the number of samples. As  $\alpha$  decreases, CVaR increasingly emphasizes the best observed samples while yielding a smoother and more manageable objective function. This enhanced smoothness facilitates more efficient convergence in optimization algorithms.

However, as indicated in Ref. [40], lowering the value of  $\alpha$  to increase the approximate ratio of QAOA results in performance degradation compared to the case of Variational Quantum Eigensolver (VQE). Presumably, this occurs because the state

vector obtained with QAOA becomes relatively flat, preventing a large overlap with the ground state. In this context, our CNN-CVaR co-optimization shows promise in addressing this issue.

## 4 Performance of CNN-CVaR Optimization on QAOA

### 4.1 Comparison between our heuristics and other optimizations

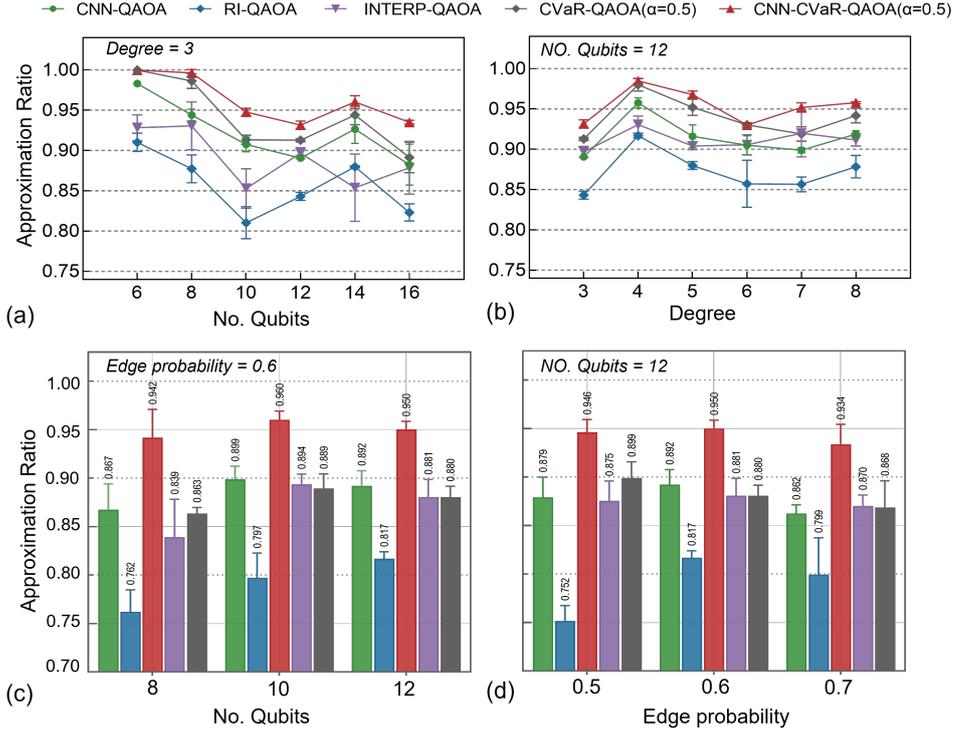
The analysis above indicates that the utilization of the CNN architecture for optimizing subsequent layer parameters within the QAOA. These parameters are incrementally predicted based on initial values. The CVaR aggregation function is employed to optimize the cost function for parameter updates, thereby enhancing the efficacy of optimal measurements through dynamic cost function adjustments. By amalgamating these methodologies with QAOA, we effectively address combinatorial optimization challenges, resulting in the development of our proposed approach, termed CNN-CVaR-QAOA. The validation of our approach is conducted through the resolution of the Max-Cut problem on regular graphs. We commence this validation process by delineating our experimental setup, encompassing the configuration of hyperparameters and the construction of the Max-Cut graph. Subsequently, to assess the proficiency of our proposed initialization methodology, we juxtapose the performance of our optimization strategy against standard strategies such as Random Initialization (RI) and interpolation (INTERP) [10].

To thoroughly evaluate the performance of our proposed CNN-CVaR-QAOA, we assembled a diverse collection of graphs with varying sizes, node degrees, and edge probabilities. For analysis, we employed Erdős-Rényi (ER) graphs, randomly allocating both degree and edge probability for each node. Multiple instances were generated for each configuration to derive statistical outcomes using the Max-Cut solver, with the average approximation ratio as the primary metric. In our experimental setup, for QAOA with a depth of 1, we consistently used the QN-SPSA optimizer to compute initial parameters conducive to convergence.

Firstly, we employed CNN-CVaR-QAOA to analyze various graph configurations, with the  $\alpha$  parameter set to 0.5. The numerical outcomes are visualized in Fig. 2, presenting the results derived from CNN-CVaR-QAOA( $\alpha=0.5$ ) alongside four heuristic Max-Cut solvers, namely Convolutional Neural Network prediction (CNN-QAOA), Random Initialization (RI-QAOA), heuristic Interpolation method (INTERP-QAOA) and Conditional Value at Risk optimization (CVaR-QAOA). Each scenario encompasses three distinct graph instances, where the number of nodes fluctuates within the range of [6,16], mirroring the number of quantum bits in the quantum circuit. The degree of nodes varies within the interval [3,8], and the probability of edge formation fluctuates within the interval [0.5,0.7]. To gauge the performance gap between CNN-CVaR-QAOA( $\alpha=0.5$ ) and the optimal classical outcomes, we employed the approximation ratio. For circuit depth 3, we validated our approach across diverse regular graphs, as shown in Fig. 2(a) and (b), affirming the superior approximation ratio achieved by our CNN-CVaR-QAOA( $\alpha=0.5$ ) strategy compared to the other four

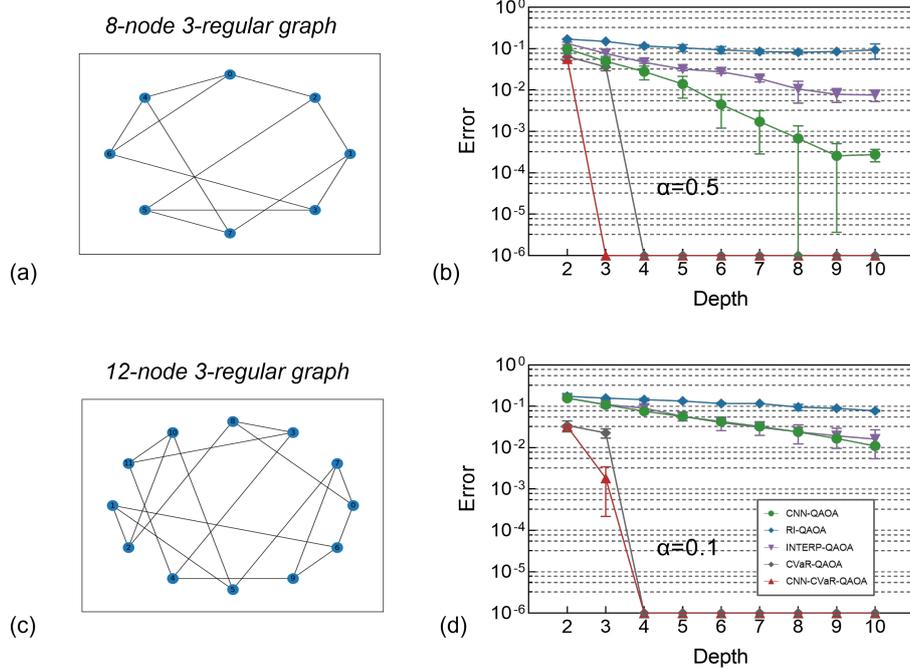
algorithms as qubit count and node degrees increased. Especially, the CNN-CVaR co-optimization significantly improves performance over the CVaR optimization of QAOA.

Additionally, with circuit depth set at 2 and  $\alpha$  at 0.5, we evaluated the algorithm across different configurations of ER random edge probability graphs, as indicated in Fig. 2(c) and (d). Our CNN-CVaR co-optimization method improved approximation ratio values in the range of 0.06-0.2 compared to other four methods. The findings consistently underscored the optimal performance of CNN-CVaR-QAOA( $\alpha=0.5$ ) across all cases. Thus, our heuristic algorithm represents a substantial enhancement in variational parameter optimization given the available resources.



**Fig. 2** The average performance of QAOA is depicted on the graph. (a) On 3-regular ER graphs, the approximation ratio obtained by each optimization method is presented as a function of the number of nodes (qubits) in the graph, with the number of layers  $p$  set to 3. (b) On 12-node ER graphs, the approximation ratio obtained by each optimization method is represented as a function of the number of degrees in the graph, with the layer count ( $p$ ) set to 3. (c) On ER sampled graphs with edge probability of 0.6, comparison of the relative performance of CNN-CVaR method with four other optimization methods on graphs containing 8, 10, and 12 qubits. (d) Comparison of the relative performance of the CNN-CVaR approach to four other optimization methods on ER sampled graphs with edge probabilities set to 0.5, 0.6, 0.7 for 12 qubits.

Then, we conducted a series of experiments utilizing benchmark instances from the Max-Cut dataset to investigate our optimization architecture by varying circuit depths. Through meticulous analysis, we aimed to elucidate performance nuances across different strategies and depths. Fig. 3(a) depicts randomly generated 8-node 3-regular graphs, serving as the focal point for comparative assessments among five distinct methods. Notably, as illustrated in Fig. 3(b), the CNN-CVaR-QAOA( $\alpha=0.5$ ) strategy exhibited rapid convergence to the optimal solution at a depth of 3. The CVaR( $\alpha=0.5$ ) method, without CNN, converges to the optimal result at the depth of 4.



**Fig. 3** (a) A set of diagrams extracted from the evaluation of an 8-node 3-regular graph instances. (b) On 8-node 3-regular graph instances, the errors during the optimization process of the CNN-CVaR( $\alpha=0.5$ ) method compared to four other QAOA optimization methods. The QAOA circuit varies from 2 layers to 10 layers. (c) A set of diagrams extracted from the evaluation of an 12-node 3-regular graph instances. (d) On 12-node 3-regular graph instances, the errors during the optimization process of the CNN-CVaR( $\alpha=0.1$ ) method compared to four other QAOA optimization methods. The QAOA circuit varies from 2 layers to 10 layers. The QAOA circuit varies from 2 layers to 10 layers.

Further insights were obtained from Fig. 3(c) and (d), revealing that CNN-CVaR-QAOA( $\alpha=0.1$ ) achieved optimality at a depth of 4 for the same 12-node 3-regular graphs. Due to the setting of the  $\alpha$  parameter, CVaR-QAOA( $\alpha=0.1$ ) also converges to the optimal result at this depth, but in comparison to CNN-CVaR-QAOA( $\alpha=0.1$ ), it

trails in the initial depths. Notably, superior performance was observed on unweighted graphs, potentially due to the inherent complexities involved in optimizing solutions for weighted counterparts. Furthermore, the circuit depth emerged as a critical factor influencing the effectiveness of our strategy. Particularly significant is our approach’s ability to produce comparable outcomes to alternative methods, even at shallower depths. For instance, in the 12-node 3-regular graphs scenario, the performance of CNN-CVaR-QAOA( $\alpha=0.1$ ) at a depth of 2 rivaled that of CNN-QAOA and INTERP-QAOA at depths as deep as 7. This finding demonstrates the potential of our approach to efficiently leverage quantum resources while approximating optimal solutions for the Max-Cut problem.

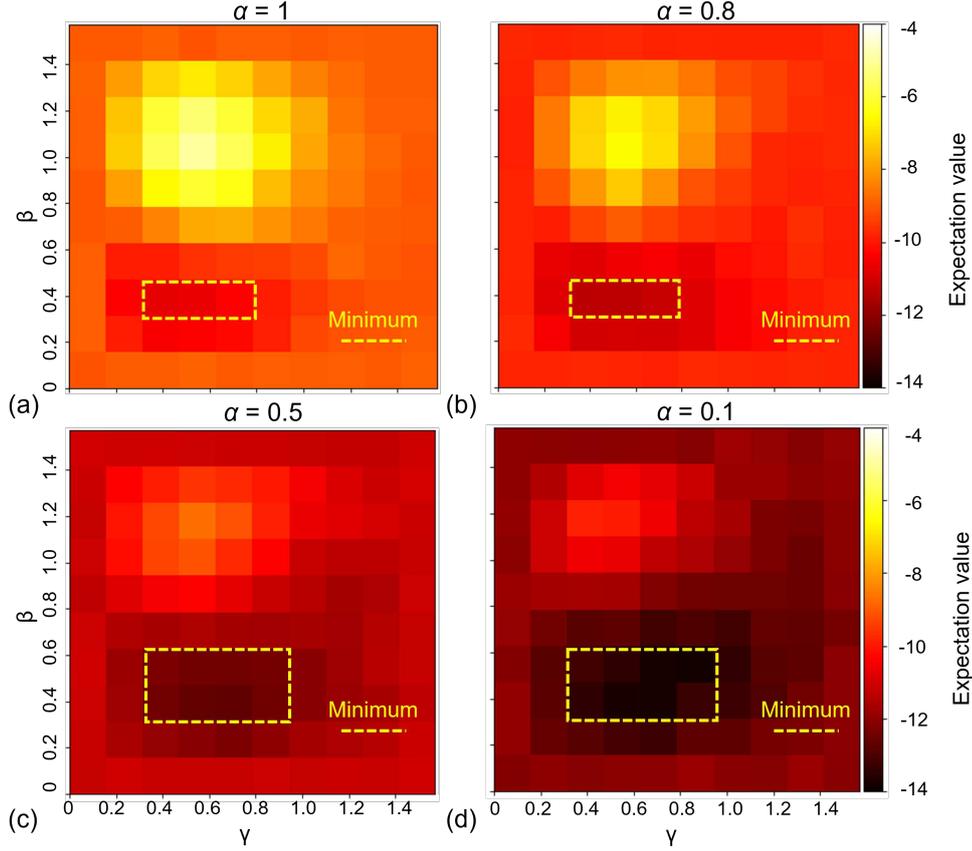
## 4.2 Analysis of hyperparameter in our co-optimization method

We examined the influence of hyperparameters, including the variational parameters  $\gamma$  and  $\beta$ , particularly the  $\alpha$  parameter of a CVaR aggregation function, on the performance of the QAOA. We employed various configurations of regular graphs and random edge probability graphs while maintaining a quantum circuit with a fixed layer count of  $p=2$ . Across these setups, we obtain the results from three distinct non-isomorphic instances.

In Fig. 4, we present a heat map depicting the relationship between expectation values and the parameters  $\gamma$ ,  $\beta$ , both ranging from 0 to  $\pi/2$ , for various CVaR parameter  $\alpha$ . As illustrated, the CVaR parameters are set to 1, 0.8, 0.5, and 0.1 on a 10-node 3-regular graphs, with  $\alpha = 1$  representing to the unoptimized cost function. By examining the convergence landscapes under various configuration scenarios, we can evaluate the quality of the results from the parameters. The yellow regions indicate the minimal values of the expectation, signifying the convergence of the random 12-node 3-regular graph. It is evident that the positions of the minimum remain almost consistent as the parameter  $\alpha$  changes, located in the ranges  $\beta \in [0.2, 0.6]$  and  $\gamma \in [0.3, 1.0]$ , as indicated by the yellow dashed lines in Figure 4. This demonstrates how the parameters  $\beta$  and  $\gamma$  distribute to achieve the optimal solutions for our CNN-CVaR co-optimization.

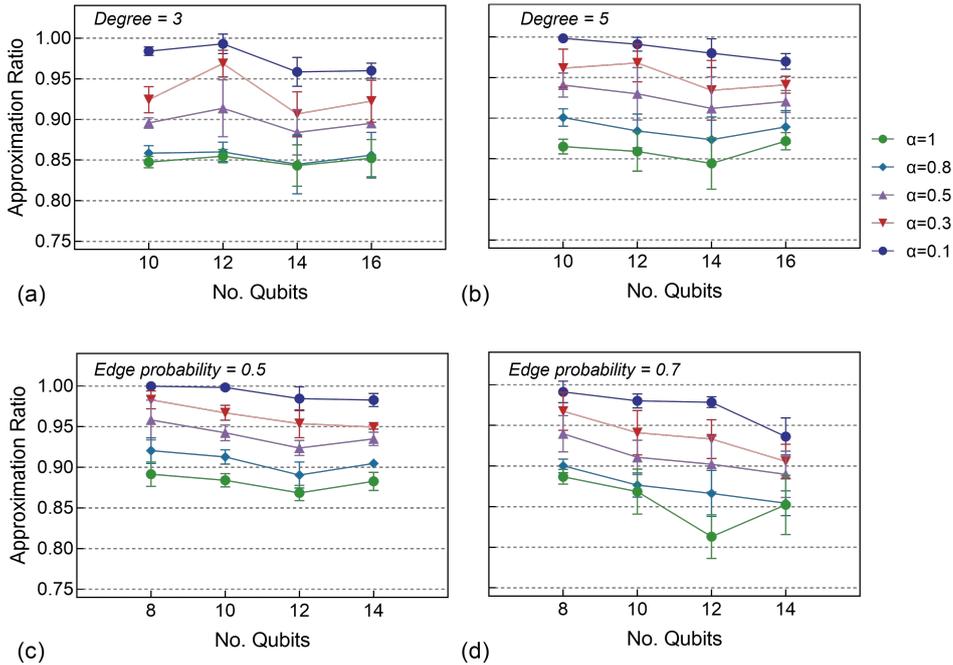
Regarding the CVaR parameter  $\alpha$ , our findings reveal that when a decrease in  $\alpha$  to 0.1 is associated with significantly lower energy values compared to the other three scenarios. The optimal variational parameters distribution expands as  $\alpha$  decreases, as shown in Figure 4,  $\beta \in [0.3, 0.5]$  and  $\gamma \in [0.3, 0.65]$  for  $\alpha = 1$  and 0.8, and  $\beta \in [0.3, 0.6]$  and  $\gamma \in [0.3, 1.0]$  for  $\alpha = 0.5$  and 0.1. This phenomenon can be attributed to the fact that, with  $\alpha = 0.1$ , the minimum energy value achieved by our CNN-CVaR co-optimization is closer to the ground state energy. Conversely, higher  $\alpha$  values correspond to higher energy values, indicating a lower probability of converging to the ground state. Additionally, as  $\alpha$  decreases, the upper bound of the expected value also significantly decreases. These insights underscore the critical role of the CVaR parameter in the optimization process, particularly in the search for superior energy states.

Additionally, the quantitative performance analysis of CNN-CVaR co-optimization under different  $\alpha$  parameters is shown in Fig. 5. The circuit depth  $p$  is set to 2, and on 10-node 3-regular graphs, we observed that under the parameter settings  $\alpha=0.1$ ,  $\alpha=0.3$ ,  $\alpha=0.5$ , and  $\alpha=0.8$ , the approximation ratio of CNN-CVaR-QAOA( $\alpha=0.1$ )



**Fig. 4** The landscape of the expectation for a random 12-node 3-regular graph with  $\gamma \in [0, \pi]$  and  $\beta \in [0, \pi]$ . (a) Convergence landscape of the parameters perturbations when  $\alpha=1$ . (b) Convergence landscape of the parameters perturbations when  $\alpha=0.8$ . (c) Convergence landscape of the parameters perturbations when  $\alpha=0.5$ . (d) Convergence landscape of the parameters perturbations when  $\alpha=0.1$ . As the CVaR parameter  $\alpha$  varies, with smaller values leading to convergence towards the optimal expectation value, following the optimal quantum optimization path. Conversely, the maximum value for  $\alpha=1$  deviates from the optimal path.

compared to the parameters of  $\alpha=0.3$ ,  $\alpha=0.5$ ,  $\alpha=0.8$ , and  $\alpha=1$  were on average improved by approximately 4.0%, 8.8%, 12.6%, and 13.7%, respectively. Similar improvements were observed on 12-node 3-regular graphs, 14-node 3-regular graphs, and 16-node 3-regular graphs. In Fig. 5(b), (c) and (d), more extensive testing was conducted on random regular graphs configurations 10-node 5-regular graphs, 12-node 5-regular graphs, 14-node 5-regular graphs, and 16-node 5-regular graphs, and on the random ER graphs, the edge probabilities are set to 0.5 and 0.7, while the number of nodes varies between 8 and 14, to assess the impact of  $\alpha$  on the overall performance of CNN-CVaR-QAOA model. These observations further confirm the aforementioned conclusions, which indicate that when executing CNN-CVaR-QAOA on a larger scale



**Fig. 5** The quantum approximation ratios obtained using classical simulation of CNN-CVaR-QAOA at different  $\alpha$  values, with a depth of 2. (a) Comparing the impact of different  $\alpha$  values on the overall performance for the Max-Cut problem on different random regular graphs with degree=3. (b) Comparing the impact of different  $\alpha$  values on the overall performance for the Max-Cut problem on different random regular graphs with degree=5. (c) Assessing the effect of different  $\alpha$  values on the overall performance on different ER random graphs with edge probabilities of 0.5. (d) Assessing the effect of different  $\alpha$  values on the overall performance on different ER random graphs with edge probabilities of 0.7. The error bars display the variance obtained through sampling over three non-isomorphic instances.

quantum system, a smaller value of  $\alpha$  would lead to the superior performance. These improvements can be explained by Eq.(6), which shows that a smaller  $\alpha$  value tends to favor optimal observation results, thereby smoothing the objective function.

## 5 Conclusions

In this study, we introduce CNN-CVaR-QAOA, a novel approach that integrates classical CNN architecture with CVaR aggregation to tackle the Max-Cut problem on small-scale quantum devices. Our method predicts QAOA parameters for each depth using a CNN network structure and incorporates the CVaR strategy as the cost function for the quantum circuit. Through comprehensive benchmark tests on diverse problem sets, we have demonstrated the superior performance of the CNN-CVaR-QAOA strategy compared to random initialization methods and interpolation

techniques. Our numerical results reveal compelling advantages, particularly evident in graph instances, where the depth-2 circuit of our approach exhibits performance levels akin to pure CNN optimization at a depth of 6. Notably, we observe that the  $\alpha$  parameter in the CNN-CVaR-QAOA strategy plays a significant role in performance enhancement, with smaller  $\alpha$  values emphasizing favorable observation samples, thereby leading to a smoother objective function and improved approximation ratios.

One of the distinguishing features of our proposed strategy is its ability to personalize predicted parameters for individual graph instances, enabling a broader exploration of optimal parameters. This flexibility becomes increasingly valuable as quantum computing hardware progresses towards handling larger-scale QAOA instances requiring  $p \gg 1$ . Our method's potential to match the approximation ratio of classical QAOA at lower depths positions it as a compelling alternative in the current NISQ era, characterized by the emergence of larger quantum devices grappling with high noise levels for executing deep circuits. While our focus lies on addressing the Max-Cut problem, we envision broader applicability of our method across various optimization problems. As quantum computers evolve and scale up in qubit count, the practicality of our approach becomes more pronounced, particularly given the finite constraints on the number of layers in QAOA for achieving feasible solutions. The CNN-CVaR-QAOA strategy represents a promising advancement in quantum optimization techniques, offering enhanced performance and adaptability that hold relevance not only for the Max-Cut problem but for a wide array of optimization challenges in the quantum computing landscape.

## Acknowledgements

- Funding : This work is supported by the the National Natural Science Foundation of China (Grant Nos. 61874001, 62004001, 62201005, 62004001, 62304001), the Anhui Provincial Natural Science Foundation under Grant No. 2308085QF213, 2308085QF195, and the Natural Science Research Project of Anhui Educational Committee under Grant No. 2022AH050106, 2023AH050072.
- Code availability : The data that support the findings of this study are available on request from the corresponding author upon reasonable request.
- Author contribution : All authors contributed to the study conception and design.

## References

- [1] Lucas A (2014) Ising formulations of many np problems. *Frontiers in Physics* 2. <https://doi.org/10.3389/fphy.2014.00005>, URL <https://www.frontiersin.org/articles/10.3389/fphy.2014.00005>
- [2] Ruan Y, Marsh S, Xue X, et al (2020) The quantum approximate algorithm for solving traveling salesman problem. *Computers, Materials & Continua* 63(3):1237–1247. <https://doi.org/10.32604/cmc.2020.010001>, URL <http://www.techscience.com/cmc/v63n3/38872>

- [3] Kochenberger, Hao G, Glover JK, et al (2014) The unconstrained binary quadratic programming problem: a survey. *Journal of Combinatorial Optimization* 28:58–81. <https://doi.org/10.1007/s10878-014-9734-0>, URL <https://doi.org/10.1007/s10878-014-9734-0>
- [4] Oh, Young-Hyun, Mohammadbagherpoor, et al (2019) Solving Multi-Coloring Combinatorial Optimization Problems Using Hybrid Quantum Algorithms [arXiv:1911.00595](https://arxiv.org/abs/1911.00595) [quant-ph]
- [5] Farhi E, Goldstone J, Gutmann S (2014) A quantum approximate optimization algorithm [arXiv:1411.4028](https://arxiv.org/abs/1411.4028) [quant-ph]
- [6] Barkoutsos, Kl. P, Nannicini G, et al (2020) Improving Variational Quantum Optimization using CVaR. *Quantum* 4:256. <https://doi.org/10.22331/q-2020-04-20-256>, URL <https://doi.org/10.22331/q-2020-04-20-256>
- [7] Jiang Z, Rieffel EG, Wang Z (2017) Near-optimal quantum circuit for grover’s unstructured search using a transverse field. *Phys Rev A* 95:062317. <https://doi.org/10.1103/PhysRevA.95.062317>, URL <https://link.aps.org/doi/10.1103/PhysRevA.95.062317>
- [8] Zhou Z, Du Y, Tian X, et al (2023) Qaoa-in-qaoa: Solving large-scale maxcut problems on small quantum machines. *Phys Rev Appl* 19:024027. <https://doi.org/10.1103/PhysRevApplied.19.024027>, URL <https://link.aps.org/doi/10.1103/PhysRevApplied.19.024027>
- [9] Zhu L, Tang HL, Barron GS, et al (2022) Adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer. *Phys Rev Res* 4:033029. <https://doi.org/10.1103/PhysRevResearch.4.033029>, URL <https://link.aps.org/doi/10.1103/PhysRevResearch.4.033029>
- [10] Zhou L, Wang ST, Choi S, et al (2020) Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Phys Rev X* 10:021067. <https://doi.org/10.1103/PhysRevX.10.021067>, URL <https://link.aps.org/doi/10.1103/PhysRevX.10.021067>
- [11] Wang Z, Hadfield S, Jiang Z, et al (2018) Quantum approximate optimization algorithm for maxcut: A fermionic view. *Phys Rev A* 97:022304. <https://doi.org/10.1103/PhysRevA.97.022304>, URL <https://link.aps.org/doi/10.1103/PhysRevA.97.022304>
- [12] Yang ZC, Rahmani A, Shabani A, et al (2017) Optimizing variational quantum algorithms using pontryagin’s minimum principle. *Phys Rev X* 7:021027. <https://doi.org/10.1103/PhysRevX.7.021027>, URL <https://link.aps.org/doi/10.1103/PhysRevX.7.021027>

- [13] Pagano G, Bapat A, Becker P, et al (2020) Quantum approximate optimization of the long-range ising model with a trapped-ion quantum simulator. Proceedings of the National Academy of Sciences 117(41):25396–25401. <https://doi.org/10.1073/pnas.2006373117>, URL <https://www.pnas.org/doi/abs/10.1073/pnas.2006373117>, <https://www.pnas.org/doi/pdf/10.1073/pnas.2006373117>
- [14] Crooks GE (2018) Performance of the quantum approximate optimization algorithm on the maximum cut problem [arXiv:1811.08419](https://arxiv.org/abs/1811.08419) [quant-ph]
- [15] Guerreschi AYG, G. Matsuura (2019) Qaoa for max-cut requires hundreds of qubits for quantum speed-up. Scientific Reports 9:6903. <https://doi.org/10.1038/s41598-019-43176-9>, URL <https://doi.org/10.1038/s41598-019-43176-9>
- [16] Morales, S.Biamonte ME, D.Zimbors J, et al (2020) On the universality of the quantum approximate optimization algorithm. Quantum Information Processing 19:291. <https://doi.org/10.1007/s11128-020-02748-9>, URL <https://doi.org/10.1007/s11128-020-02748-9>
- [17] Niu MY, Lu S, Chuang IL (2019) Optimizing qaoa: Success probability and runtime dependence on circuit depth [arXiv:1905.12134](https://arxiv.org/abs/1905.12134) [quant-ph]
- [18] Lloyd S (2018) Quantum approximate optimization is computationally universal [arXiv:1812.11075](https://arxiv.org/abs/1812.11075) [quant-ph]
- [19] Alexeev Y, Bacon D, Brown KR, et al (2021) Quantum computer systems for scientific discovery. PRX Quantum 2:017001. <https://doi.org/10.1103/PRXQuantum.2.017001>, URL <https://link.aps.org/doi/10.1103/PRXQuantum.2.017001>
- [20] Preskill J (2018) Quantum Computing in the NISQ era and beyond. Quantum 2:79. <https://doi.org/10.22331/q-2018-08-06-79>, URL <https://doi.org/10.22331/q-2018-08-06-79>
- [21] Herrman, Ostrowski R, Humble J, et al (2021) Lower bounds on circuit depth of the quantum approximate optimization algorithm. Quantum Information Processing 20:59. <https://doi.org/10.1007/s11128-021-03001-7>, URL <https://doi.org/10.1007/s11128-021-03001-7>
- [22] Holmes Z, Sharma K, Cerezo M, et al (2022) Connecting ansatz expressibility to gradient magnitudes and barren plateaus. PRX Quantum 3:010313. <https://doi.org/10.1103/PRXQuantum.3.010313>, URL <https://link.aps.org/doi/10.1103/PRXQuantum.3.010313>
- [23] Chandarana P, Vieites PS, Hegade NN, et al (2023) Meta-learning digitized-counterdiabatic quantum optimization. Quantum Science and Technology 8(4):045007. <https://doi.org/10.1088/2058-9565/ace54a>, URL <https://dx.doi.org/10.1088/2058-9565/ace54a>

- [24] Wierichs D, Gogolin C, Kastoryano M (2020) Avoiding local minima in variational quantum eigensolvers with the natural gradient optimizer. *Phys Rev Res* 2:043246. <https://doi.org/10.1103/PhysRevResearch.2.043246>, URL <https://link.aps.org/doi/10.1103/PhysRevResearch.2.043246>
- [25] Bravyi S, Kliesch A, Koenig R, et al (2020) Obstacles to variational quantum optimization from symmetry protection. *Phys Rev Lett* 125:260505. <https://doi.org/10.1103/PhysRevLett.125.260505>, URL <https://link.aps.org/doi/10.1103/PhysRevLett.125.260505>
- [26] Alam M, Ash-Saki A, Ghosh S (2020) Accelerating quantum approximate optimization algorithm using machine learning pp 686–689. <https://doi.org/10.23919/DATE48585.2020.9116348>
- [27] Lee X, Xie N, Cai D, et al (2023) A depth-progressive initialization strategy for quantum approximate optimization algorithm. *Mathematics* 11(9). <https://doi.org/10.3390/math11092176>, URL <https://www.mdpi.com/2227-7390/11/9/2176>
- [28] Xie N, Lee X, Cai D, et al (2023) Quantum approximate optimization algorithm parameter prediction using a convolutional neural network. *Journal of Physics: Conference Series* 2595(1):012001. <https://doi.org/10.1088/1742-6596/2595/1/012001>, URL <https://dx.doi.org/10.1088/1742-6596/2595/1/012001>
- [29] Sack SH, Serbyn M (2021) Quantum annealing initialization of the quantum approximate optimization algorithm. *Quantum* 5:491. <https://doi.org/10.22331/q-2021-07-01-491>, URL <https://doi.org/10.22331/q-2021-07-01-491>
- [30] Verdon G, Broughton M, McClean JR, et al (2019) Learning to learn with quantum neural networks via classical neural networks. *ArXiv abs/1907.05415*. URL <https://api.semanticscholar.org/CorpusID:195886305>
- [31] Amosy O, Danzig T, Porat E, et al (2022) Iterative-free quantum approximate optimization algorithm using neural networks URL <https://api.semanticscholar.org/CorpusID:251719266>
- [32] Li L, Fan M, Coram M, et al (2020) Quantum optimization with a novel gibbs objective function and ansatz architecture search. *Phys Rev Res* 2:023074. <https://doi.org/10.1103/PhysRevResearch.2.023074>, URL <https://link.aps.org/doi/10.1103/PhysRevResearch.2.023074>
- [33] Nielsen MA, Chuang IL (2011) *Quantum Computation and Quantum Information: 10th Anniversary Edition, 10th edn.* Cambridge University Press, USA
- [34] Goemans MX, Williamson DP (1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J ACM* 42(6):11151145. <https://doi.org/10.1145/227683.227684>, URL <https://doi.org/10.1145/227683.227684>

- [35] Glover F (2022) Quantum bridge analytics i: a tutorial on formulating and using qubo models. *Annals of Operations Research* 314:141–183. <https://doi.org/10.1007/s10479-022-04634-2>, URL <https://doi.org/10.1007/s10479-022-04634-2>
- [36] Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):8490. <https://doi.org/10.1145/3065386>, URL <https://doi.org/10.1145/3065386>
- [37] Lim AEB, Shanthikumar JG, Vahn GY (2011) Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Oper Res Lett* 39:163–171. URL <https://api.semanticscholar.org/CorpusID:8467251>
- [38] Gacon J, Zoufal C, Carleo G, et al (2021) Simultaneous Perturbation Stochastic Approximation of the Quantum Fisher Information. *Quantum* 5:567. <https://doi.org/10.22331/q-2021-10-20-567>, URL <https://doi.org/10.22331/q-2021-10-20-567>
- [39] Agarap AF (2018) Deep learning using rectified linear units (relu). *ArXiv abs/1803.08375*. URL <https://api.semanticscholar.org/CorpusID:4090379>
- [40] Kolotouros I, Wallden P (2022) Evolving objective function for improved variational quantum optimization. *Phys Rev Res* 4:023225. <https://doi.org/10.1103/PhysRevResearch.4.023225>, URL <https://link.aps.org/doi/10.1103/PhysRevResearch.4.023225>