

# Training-Free Quantum Architecture Search

Zhimin He<sup>1</sup>, Maijie Deng<sup>2</sup>, Shenggen Zheng<sup>3</sup>, Lvzhou Li<sup>4</sup>, Haozhen Situ<sup>5\*</sup>

<sup>1</sup>School of Electronic and Information Engineering, Foshan University, Foshan 528000, China

<sup>2</sup>School of Mechatronic Engineering and Automation, Foshan University, Foshan 528000, China

<sup>3</sup>Peng Cheng Laboratory, Shenzhen, 518055 China

<sup>4</sup>Institute of Quantum Computing and Computer Theory, School of Computer Science and Engineering, Sun Yat-Sen University, Guangzhou 510006, China

<sup>5</sup>College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China  
situhaozhen@gmail.com

## Abstract

Variational quantum algorithm (VQA) derives advantages from its error resilience and high flexibility in quantum resource requirements, rendering it broadly applicable in the noisy intermediate-scale quantum era. As the performance of VQA highly relies on the structure of the parameterized quantum circuit, it is worthwhile to propose quantum architecture search (QAS) algorithms to automatically search for high-performance circuits. Nevertheless, existing QAS methods are time-consuming, requiring circuit training to assess circuit performance. This study pioneers training-free QAS by utilizing two training-free proxies to rank quantum circuits, in place of the expensive circuit training employed in conventional QAS. Taking into account the precision and computational overhead of the path-based and expressibility-based proxies, we devise a two-stage progressive training-free QAS (TF-QAS). Initially, directed acyclic graphs (DAGs) are employed for circuit representation, and a zero-cost proxy based on the number of paths in the DAG is designed to filter out a substantial portion of unpromising circuits. Subsequently, an expressibility-based proxy, finely reflecting circuit performance, is employed to identify high-performance circuits from the remaining candidates. These proxies evaluate circuit performance without circuit training, resulting in a remarkable reduction in computational cost compared to current training-based QAS methods. Simulations on three VQE tasks demonstrate that TF-QAS achieves a substantial enhancement of sampling efficiency ranging from 5 to 57 times compared to state-of-the-art QAS, while also being 6 to 17 times faster.

## Introduction

The emergence of quantum computing introduces an entirely novel computing paradigm that leverages the principles of quantum mechanics, offering the potential for exponential acceleration (Nielsen and Chuang 2010). Variational quantum algorithms (VQAs) combining quantum hardware evaluation with classical optimization are anticipated to provide quantum advantage in the absence of fault-tolerant quantum error correction (Cerezo et al. 2021; Shi et al. 2022). VQAs have demonstrated their superiority over classical algorithms in certain fields, including integer factorization, combinatorial optimization, and quantum Hamiltonian ground state

\*Corresponding author.

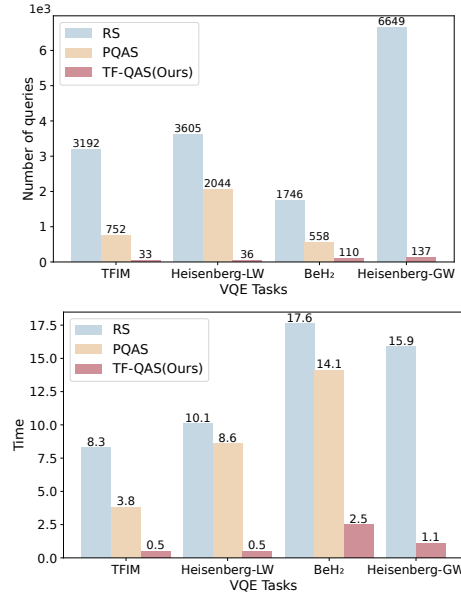


Figure 1: Average number of queries and time (in hours) required to achieve the ground states of 6-qubit TFIM, 5-qubit Heisenberg model, and 6-qubit BeH<sub>2</sub> molecule by random sampling (RS), PQAS (Zhang et al. 2021) and TF-QAS over 10 independent runs. Fewer queries indicate better capability to recognize high-performance circuits. LW and GW denote that the quantum circuit is generated through layerwise and gatewise pipelines, respectively.

estimation. The parametrized quantum circuit (PQC) acts as a bridge connecting quantum and classical computational resources. VQA involves executing a PQC on a quantum computer to calculate an objective function value, followed by refining the circuit parameters using optimization techniques on a classical computer.

The performance of VQA heavily relies on the structure of the used PQC. Automated design of PQC, also known as quantum architecture search (QAS) has gained increasing attention (Grimsley et al. 2019; Ostaszewski et al. 2021; Zhang et al. 2021, 2022; He et al. 2022; Du et al. 2022; Wu et al. 2023; Wang et al. 2023; Lu et al. 2023). Due to the constraints posed by the hardware-native topologies

and gate sets of quantum devices, manually designed templates or circuits for a specific problem require compilation before deployment onto the quantum device. This involves routing qubits with swap networks and gate synthesis, leading to a notable enlargement of the quantum circuit size, and consequently yields suboptimal circuits. Instead, QAS is an end-to-end approach dedicated to directly searching for an optimal quantum circuit under the constraints of the quantum device, enabling the discovery of more efficient circuit structures.

Neural architecture search (NAS), dedicated to the automated design of neural networks, exhibits parallels with the design of parametric quantum circuits. Inspired by NAS, a variety of quantum architecture search (QAS) algorithms have been proposed, including differentiable QAS (Zhang et al. 2022; Wu et al. 2023), QAS based on reinforcement learning (Ostaszewski et al. 2021), predictor-based QAS (Zhang et al. 2021) and weight-sharing QAS (Du et al. 2022). Training-free (or zero-shot) NAS has garnered considerable attention as it can incur significantly reduced computational costs while attaining competitive performances compared to training-based NAS algorithms. The key idea behind training-free NAS is to design proxies that predict the performance of neural networks without training the neural network. While training-free NAS prevails in the field of machine learning, the training-free paradigm has not been fully developed in QAS. The design of proxies for training-free NAS generally draws inspiration from theoretical analysis of deep neural networks. The inherent distinctions between quantum circuits and deep neural networks preclude the direct application of existing training-free proxies in NAS for the evaluation of quantum circuits. Training quantum circuits is less efficient than training neural networks, emphasizing the critical importance of the training-free approach for QAS. Consequently, it becomes necessary and valuable to explore effective training-free proxies for QAS. This motivates us to develop QAS algorithms in a training-free manner based on the unique characteristics of quantum circuits.

This paper proposes a straightforward yet potent QAS based on path-based and expressibility-based proxies that are training-free. The integration of these proxies not only improves the capability to identify high-performance circuits but also notably diminishes computational time, as depicted in Fig. 1. Prominent instances of VQAs include variational quantum eigensolvers (VQE) (McClean et al. 2016), quantum approximation optimization algorithms (QAOA) (Farhi, Goldstone, and Gutmann 2014), quantum neural networks (QNN) (Cong, Choi, and Lukin 2019) and quantum generative modeling (Lloyd and Weedbrook 2018; Situ et al. 2020). Despite variations in their objective functions (such as energy, mean squared error, and Kullback-Leibler divergence), these algorithms employ the same quantum subroutine to generate parameterized trial state or ansatz, allowing for tuning parameters to achieve the optimal or near-optimal values of the objective function. As a result, the performance of VQAs hinges upon the intrinsic expressibility of the chosen parameterized quantum circuit, defined as a circuit's ability to uniformly reach the entire Hilbert space (Sim, Johnson, and Aspuru-Guzik 2019). Since the assessment of ex-

pressibility solely entails collecting quantum states generated with random parameters, without any parameter optimization, the computational cost of circuit expressibility is significantly lower compared to achieving ground-truth performance through iterative optimization of gate parameters until convergence.

The search space of quantum circuits expands exponentially with the size of the circuit. For large-scale quantum circuits, it is necessary to evaluate a substantial number of circuits. Even though assessing the expressibility of a quantum circuit comes with a considerably lower computational cost compared to evaluating its ground-truth performance, calculating the expressibility for a multitude of quantum circuits remains unfeasible. To this end, we propose a zero-cost proxy to filter out the majority of the unpromising quantum circuits from the search space. The proxy is defined as the number of paths between the input and output nodes of the DAG representation of the quantum circuit. The number of paths can approximate the topological complexity of the quantum circuit, providing a rough indication of various properties like expressibility and entangling capacity. Although the number of paths only offers a rough insight into a quantum circuit's performance, it can serve as a preliminary filtering proxy to eliminate poor-performance circuits.

Overall, the main contributions are summarized below.

- We identify and investigate two training-free proxies to rank quantum circuits. The path-based proxy can be regarded as zero-cost, exhibiting a relatively weaker correlation with circuit performance. In contrast, the expressibility-based proxy is closely linked to circuit performance, albeit operating at a pace three orders of magnitude slower compared to the path-based proxy. As these proxies do not involve the objective functions of VQAs, they may easily transfer across different VQAs.
- This paper establishes a simple but effective progressive training-free QAS framework (TF-QAS) that effectively leverages the zero cost of the path-based proxy and the strong correlation of the expressibility-based proxy to circuit performance. The progressive (from rough to precise) strategy not only results in a substantial reduction in computational cost but also enables better identification of high-performance circuits by capturing various properties of the circuit.
- We evaluate the performance of TF-QAS on three VQE tasks for TFIM, Heisenberg, and  $\text{BeH}_2$  models. Simulation results reveal that TF-QAS achieves an average sampling efficiency 28 times greater than that of state-of-the-art (SOTA) QAS, reducing the average computational cost from 8.8 hours to 1.2 hours.

## Related Works

### Quantum Architecture Search Algorithms (QAS)

QAS is an automated process for designing quantum circuits. It comprises a search module and a performance evaluation module. Previous QAS algorithms primarily concentrated on the search module, employing various search strategies such as reinforcement learning algorithms (Ostaszewski et al. 2021; Kuo, Fang, and Chen 2021; Wang et al. 2023),

evolutionary algorithms (Las Heras et al. 2016; Romero, Olson, and Aspuru-Guzik 2017; Huang et al. 2022), and simulated annealing (Khatri et al. 2019; Cincio et al. 2021) to identify high-performance circuits. Although these QAS algorithms have demonstrated superior performance compared to manually designed circuits in various VQAs, they are computationally expensive due to their requirement for calculating the ground-truth performances of numerous circuits through circuit training during the search process. Circuit training entails iterative parameter optimizations for each gate until convergence, a highly time-consuming process. The computational bottleneck is a barrier to the applications of QAS algorithms, particularly for large-scale quantum circuits.

To enhance the search efficiency of QAS, researchers have directed their attention to the evaluation module, exploring two efficient evaluation methods: one-shot super-circuit (Du et al. 2022; Wang et al. 2022) and predictor-based QAS (Zhang et al. 2021; He et al. 2023). The super-circuit approach involves sharing gate parameters with sub-circuits after one-shot training to accelerate circuit evaluation. Nevertheless, optimizing the super-circuit presents a challenge, and furthermore, there is a weak correlation in performances between sub-circuits with inherited parameters and those with optimal parameters from individual training. Predictor-based QAS tries to directly predict circuit performance using a neural network. However, training the neural network requires a training set containing representative circuits and their corresponding ground-truth performances. To ensure a precise performance evaluation, a sufficient number of training samples is essential, particularly when dealing with the vast search space of large-scale circuits. Consequently, the process of data collection still incurs substantial computational costs. These strategies for performance evaluation involve circuit training. To further enhance efficiency, this paper proposes a QAS algorithm that eliminates the need for circuit training.

### Training-free Neural Architecture Search (NAS)

Training-free NAS has risen to prominence among NAS algorithms owing to its exceptional efficiency and competitive performance. It estimates the performances of neural architectures using intuitive or theoretical proxies without neural network (NN) training. NASWOT (Mellor et al. 2021) designed a proxy based on the number of linear regions that reflect the learning capability of an NN. Abdelfattah et al. assessed performance with a pruning-at-initialization criteria (Abdelfattah et al. 2021). Neural tangent kernel, which reflects the convergence speed of an NN, is utilized for performance evaluation (Chen, Gong, and Wang 2021; Shu et al. 2022). Gradients are also leveraged to develop proxies (Zhang and Jia 2021; Xu et al. 2021). Recently, the parameterized quantum circuit has been compared with classical neural networks and interpreted within their framework, where the gate parameters of circuits resemble the weights and biases of classical neural networks. However, these proxies are constructed based on the properties of neural networks. Due to the inherent differences between quantum circuits and neural networks, these proxies cannot directly gen-

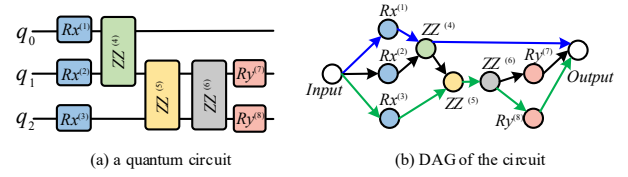


Figure 2: An illustration of a circuit’s DAG representation.

eralize to QAS. This challenge motivates us to delve deeper into the quantum circuit in order to develop QAS algorithms in a training-free manner.

## Methodology

We propose a training-free QAS algorithm that leverages two proxies (i.e., the number of paths and expressibility) to rank parameterized quantum circuits (PQCs) in place of the computationally intensive training step inherent in prevailing QAS algorithms. These proxies characterize the PQC in a more general way, independent of the specific VQAs. The proposed training-free QAS is highly appealing and practical thanks to its simplicity and low computational overhead.

### Number of Paths

A quantum circuit can be represented as a directed acyclic graph (DAG) (Nam et al. 2018) as depicted in Fig. 2. The DAG comprises an input node, a series of gate operation nodes, and an output node, interconnected in sequence. The directed edges encode the input/output relationships between quantum gates. DAG precisely captures the sequence of gate operations and the dependencies between quantum gates. The number of paths within a DAG is defined as the count of distinct paths from the input node to the output node. Fig. 2 illustrates two distinct input-to-output paths, depicted by blue and green lines. The number of paths is an indicator to assess the complexity of a quantum circuit. The computational complexity of calculating the number of paths in a quantum circuit is  $\mathcal{O}(|E|)$ , where  $|E|$  denotes the number of edges in the corresponding DAG. Consequently, the path-based proxy can be regarded as zero-cost.

Path counting provides a simple and efficient metric for quantifying the topological complexity of quantum circuits. Generally, a smaller number of paths might indicate a relatively simpler quantum circuit, whereas a larger number of paths implies a more complex circuit. Moreover, the number of paths correlates with the quantity of two-qubit gates and the qubits they operate on. The utilization of two-qubit gates notably amplifies the number of paths. However, it’s crucial to recognize that consecutively applying two-qubit gates to the same qubits does not contribute to an increased path count. For example, the two-qubit gate  $ZZ^{(4)}$  in Fig. 2 doubles the number of paths, whereas the  $ZZ^{(6)}$  has no impact on the path count. This observation implies that solely counting two-qubit gates cannot comprehensively encompass circuit complexity. Furthermore, the number of paths is also associated with entanglement capability. Quantum circuits with a larger number of paths generally imply a greater potential for entanglement, facilitating increased interaction

among qubits. Because the performance of a quantum circuit is affected by various factors such as the gate types and the arrangement of quantum gates, relying solely on the number of paths cannot precisely depict the circuit performance. Nevertheless, the information concerning circuit complexity and entanglement properties provided by the number of paths enables it to function as a preliminary filter for excluding poorly performing circuits.

### Expressibility

The expressibility of a parameterized quantum circuit (PQC) is defined as its capability to uniformly reach the entire Hilbert space (Sim, Johnson, and Aspuru-Guzik 2019). The state fidelity (i.e., the overlap of states) is defined as  $F = |\langle \psi_\theta | \psi_\phi \rangle|^2$ , where  $\theta$  and  $\phi$  are the gate parameters of the PQC. The expressibility of a PQC  $\mathcal{C}$  can be quantified by

$$\mathcal{E}(\mathcal{C}) = -D_{KL}(P(\mathcal{C}, F) \| P_{Haar}(F)), \quad (1)$$

where  $P(\mathcal{C}, F)$  is the resulting distribution of state fidelities generated by the sampled ensemble of parameterized states. In practice,  $P(\mathcal{C}, F)$  is approximated by the probability distribution of fidelities obtained through sampling parameterized states from the PQC with random parameters. The input state of the PQC is set as  $|0\dots 0\rangle$ .  $P_{Haar}(F)$  denotes the probability density function of fidelities for the ensemble of Haar random states. Its analytical form is known as  $P_{Haar}(F) = (N-1)(1-F)^{N-2}$ , where  $N$  is the dimension of the Hilbert space (Życzkowski and Sommers 2005). Once sufficient samples of state fidelities are gathered, the Kullback-Leibler (KL) divergence ( $D_{KL}$ ) can be utilized to quantify the difference between the estimated fidelity distribution of the PQC and that of the Haar distributed ensemble. A larger value of  $\mathcal{E}(\mathcal{C})$  indicates better expressibility. For an idle circuit with no quantum gate or a circuit without parameterized gate, the value of  $\mathcal{E}(\mathcal{C})$  reaches its minimum since the circuit consistently produces the same quantum state. Furthermore,  $\mathcal{E}(\mathcal{C})$  also quantifies the uniformity in the state distribution of a PQC.

Although various VQAs (such as VQE, QAOA, and QNN) employ different objective functions (such as energy and mean squared error), these objective functions are all derived from the output state of the PQC. VQAs use the same quantum subroutine to generate parameterized trial states or ansatz, aiming to achieve optimal or approximate optimal values of the objective functions by adjusting gate parameters. Circuits with high expressibility are capable of uniformly mapping the initial state to nearly all states in the Hilbert space possibly encompassing the optimal solutions (e.g., the ground states in VQE). As a result, the performance of VQA hinges on the inherent expressive power of the chosen PQC. Furthermore, PQCs with high expressibility generate quantum states that are distributed uniformly, an advantage particularly valuable in scenarios with limited prior knowledge.

### Two-stage Progressive TF-QAS

The path-based proxy, being zero-cost, provides a crude reflection of circuit performance, rendering it suitable for filtering out unpromising quantum circuits from the extensive

---

#### Algorithm 1: A two-stage progressive TF-QAS

---

**Input:** number of circuits sampled from the search space  $S$ ; number of remaining circuits after filtering based on the path-based proxy  $R$ , where  $R \ll S$

**Output:** the top- $K$  circuits

---

- 1: Randomly sample a set of circuits  $\mathcal{D} = \{\mathcal{C}_i\}_{i=1}^S$  from the search space.
  - 2: **for** each  $\mathcal{C}_i$  in  $\mathcal{D}$  **do**
  - 3:   Represent  $\mathcal{C}_i$  by a DAG and calculate the number of paths.
  - 4: **end for**
  - 5: Output  $R$  circuits with the largest numbers of paths  $\mathcal{D}_r = \{\mathcal{C}_{j_k}\}_{k=1}^R$ .
  - 6: **for** each  $\mathcal{C}_i$  in  $\mathcal{D}_r$  **do**
  - 7:   Calculate the expressibility  $\mathcal{E}_i$  of  $\mathcal{C}_i$  based on Eq. 1.
  - 8: **end for**
  - 9: **return** the top- $K$  circuits  $\mathcal{D}_K$  ranked by expressibility
- 

search space. In contrast, the expressibility-based proxy reliably indicates the performance of quantum circuits, facilitating the selection of high-performance circuits from the remaining circuits. Taking advantage of the zero cost of the path-based proxy and the strong rank correlation of the expressibility-based proxy with circuit performance, we devise a two-stage progressive training-free QAS (TF-QAS) as shown in Algorithm 1. Because the path-based proxy is zero-cost, we can sample a substantial number of circuits ( $\mathcal{D} = \{\mathcal{C}_i\}_{i=1}^S$ ) from the search space randomly, and calculate their respective numbers of paths. The generation of random circuits can be achieved through a gatewise or layerwise pipeline (Zhang et al. 2021). Subsequently,  $R$  circuits, possessing the highest numbers of paths, are selected to form a candidate set  $\mathcal{D}_r = \{\mathcal{C}_{j_k}\}_{k=1}^R$ , where  $R \ll S$ . In the second stage, we sort the circuits in  $\mathcal{D}_r$  according to their expressibility. As the expressibility is independent of specific VQAs, TF-QAS outputs the top- $K$  quantum circuits ( $\mathcal{D}_K$ ) characterized by the highest expressibility, rather than just the single highest one. For a particular VQA, the best circuit from  $\mathcal{D}_K$  can be chosen according to their performances on this VQA.

### Simulation Results

We evaluate the performance of TF-QAS across three variational quantum eigensolver (VQE) tasks for finding the ground states of the transverse field Ising model (TFIM), Heisenberg model, and BeH<sub>2</sub> molecule. Our method belongs to model performance inference (MPI)-based QAS, which rapidly assesses the predictive performance of various candidate circuits to identify high-performing ones. We compared our method to the state-of-the-art MPI-based QAS, namely PQAS (Zhang et al. 2021), and random sampling (RS). Additionally, we extended our comparisons to include the hardware-efficient ansatz (HEA) (Kandala et al. 2017). Simulations are executed on a computer equipped with an R9 7950X CPU @4.5GHz and a GeForce RTX 4090 GPU. All numerical simulations are implemented with the TensorCircuit package (Zhang et al. 2023) in Python. Each simula-

tion is independently run ten times, and the averaged results are reported.

## Settings

We consider VQE tasks for 6-qubit TFIM, 5-qubit Heisenberg model, and 6-qubit chemical Hamiltonian of  $\text{BeH}_2$  at bond distance. The Hamiltonian of TFIM and Heisenberg are defined as  $H_{\text{TFIM}} = \sum_{i=1}^n Z_i Z_{i+1} + X_i$  and  $H_{\text{Heis}} = \sum_{i=1}^n X_i X_{i+1} + Y_i Y_{i+1} + Z_i Z_{i+1} + \sum_{i=1}^n Z_i$ , where  $n$  denotes the number of qubits. The Hamiltonian of  $\text{BeH}_2$  is constructed using the coefficients and the Pauli terms provided by the supplement material in previous work (Kandala et al. 2017). We assume that the ground state energy is attained if the estimated value  $E$  is within the chemical accuracy, i.e.,  $E - E_0 \leq 1.6 \times 10^{-3}$  Hartree (Ha), where  $E_0$  is the exact ground state energy.

We follow the prior work (Zhang et al. 2021) to construct the search space of QAS. The native gate set is  $\mathcal{A} = \{R_x, R_y, R_z, XX, YY, ZZ\}$ .  $R_x$ ,  $R_y$  and  $R_z$  are single-qubit gates, where  $R_x(\theta) = e^{-i\theta\sigma_x/2}$  (similarly for  $R_y$  and  $R_z$ ).  $XX$ ,  $YY$  and  $ZZ$  are two-qubit gates, where  $XX(\theta) = e^{-i\theta\sigma_x\sigma_x/2}$  and similarly for  $YY$  and  $ZZ$ . The layouts of quantum circuits resulting from a uniformly random selection of gates are notably “chaotic”, often suffering from serious barren plateau problems. To mitigate this, we generate quantum circuits by the gatewise and layerwise pipelines with hierarchical sampling employed in prior research (Zhang et al. 2021), in order to create structured circuits. The gatewise pipeline constructs a quantum circuit by adding gates one by one, while the layerwise pipeline generates a circuit by iteratively adding half-layers (i.e.,  $n/2$ ) of quantum gates. As single-qubit gates are less susceptible to noise and easier to implement than two-qubit gates, we will exclude circuits with a higher number of two-qubit gates than single-qubit ones. In VQE for the TFIM, the quantum circuit begins with a layer of Hadamard gates applied to all qubits, consistent with prior study (Zhang et al. 2021).

The parameters of quantum circuits are initialized with random values within the range of  $[-2\pi, 2\pi]$ , and subsequently optimized using the Adam optimizer with a learning rate of 0.01 until convergence to calculate the ground-truth performance. For expressibility evaluation, 2000 state fidelities are sampled using random gate parameters to estimate the fidelity distribution of the circuit. The number of remaining circuits filtered by the path-based proxy  $R$  is set to 5000.

## Layerwise Search Space

For the layerwise pipeline, we select a gate type and then allocate it to either all the odd qubits or all the even qubits. For two-qubit gates, they are positioned on qubit pairs such as (2,3), (4,5)... or (1,2), (3,4)... In the layerwise search space, the total number of quantum gates in a circuit is consistently set as 36, 35, and 57 for the TFIM, Heisenberg, and  $\text{BeH}_2$  models. The max depths of circuits in these tasks are set as 10, 15, and 20. Any generated circuit surpassing the max depth is simply omitted.

The primary objective of VQE is to compute the ground states of a given system. In this context, Algorithm 1 is extended slightly: we keep querying the ground-truth performances of quantum circuits (from best to worst as ranked by TF-QAS) until the ground state energy within the range of chemical accuracy is obtained. Fig. 1 shows the average number of queries for the ground-truth performances required to achieve the ground state energy over 10 independent runs. In comparison to PQAS (Zhang et al. 2021), TF-QAS drastically reduces the number of queries for ground-truth performance, ranging from nearly 5 to 57 times, while also achieving a speed improvement of 6 to 17 times. Beyond the queries depicted in Fig. 1, PQAS requires another 300 queries for the ground-truth performances to construct the training set. Random sampling (RS) denotes the candidate circuits are randomly selected from the pre-defined search space. Compared to RS, TF-QAS improves the sampling efficiency by nearly 100 times in VQE for the TFIM and Heisenberg models.

We also show the probability of achieving the ground state energy over 10 independent runs, operating within a fixed budget of 100 queries for the ground-truth performances of quantum circuits in Table 1. TF-QAS outputs the top 100 circuits for the final validation. PQAS nearly fails to find out quantum circuits that achieve the ground state energy across all the VQAs. In contrast, under an equivalent number of quantum gates, TF-QAS finds out quantum circuits that attain the ground state energy with probabilities of 100%, 80%, and 50%, respectively. The average energy errors of TF-QAS fall within the scope of chemical accuracy ( $1.6 \times 10^{-3}$ ) in TFIM and Heisenberg models, and they marginally exceed the chemical accuracy in the  $\text{BeH}_2$  model.

Furthermore, we compare the performance with the hardware efficient ansatz (HEA) (Kandala et al. 2017), demonstrating the benefits of the automatic design of quantum circuits. HEA is constructed through the iterative replication of  $L$  identical unit layers, each comprising single-qubit operations followed by entangling two-qubit operations. Since HEA is a manually designed circuit with a fixed structure, we only need to optimize the gate parameters. 100% in the table denotes the successful attainment of the ground state energy by HEA, while 0% indicates the failure to achieve the ground state energy. Upon employing an identical or comparable number of quantum gates to that of TF-QAS, HEA fails to attain the ground state energy across all VQE tasks. In order to achieve the ground state energy, HEA requires twice the number of quantum gates used by TF-QAS in both the TFIM and Heisenberg models. Compared with manually designed circuits, TF-QAS not only employs significantly fewer gates to suppress noise and trainability issues but also upholds ample expressibility to contain the solution of VQA.

## Gatewise Search Space with IBM’s Topology

We use the topology of IBM’s 5-qubit quantum device called *ibmq\_quito*, which has limited connectivity between qubits as shown in Fig. 3. Our simulation is limited to VQE for the Heisenberg model as the TFIM and the  $\text{BeH}_2$  molecule are 6-qubit systems. In this context, PQAS is not applicable



VQE task	Method	#gates	Prob.	$E - E_0$	Design
TFIM	RS	36	0%	$1.2 \times 10^{-1}$	Auto
	PQAS	36	10%	$9.0 \times 10^{-2}$	Auto
	TF-QAS	36	100%	$1.4 \times 10^{-14}$	Auto
	HEA-2	36	0%	$1.2 \times 10^{-1}$	Manual
	HEA-3	54	0%	$1.1 \times 10^{-1}$	Manual
	HEA-4	72	100%	$4.5 \times 10^{-6}$	Manual
Heisenberg	RS	35	0%	$8.9 \times 10^{-2}$	Auto
	PQAS	35	0%	$8.9 \times 10^{-2}$	Auto
	TF-QAS	35	80%	$1.2 \times 10^{-3}$	Auto
	HEA-3	42	0%	$1.5 \times 10^{-1}$	Manual
	HEA-4	56	0%	$7.4 \times 10^{-3}$	Manual
	HEA-5	70	100%	$2.6 \times 10^{-5}$	Manual
BeH <sub>2</sub>	RS	57	0%	$5.0 \times 10^{-3}$	Auto
	PQAS	57	10%	$4.8 \times 10^{-3}$	Auto
	TF-QAS	57	50%	$1.8 \times 10^{-3}$	Auto
	HEA-3	54	0%	$1.1 \times 10^{-2}$	Manual
	HEA-4	72	100%	$6.2 \times 10^{-4}$	Manual

Table 1: Probability of attaining the ground state energy within 100 queries to the ground-truth performances of quantum circuits across 10 independent runs. HEA- $L$  is used to name a hardware efficient ansatz with  $L$  layers.  $E - E_0$  denotes the average energy error across 10 independent simulations, where  $E$  and  $E_0$  are the estimated energy and the exact ground state energy.

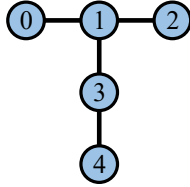


Figure 3: The topology of IBM’s 5-qubit quantum device (ibmq-quito).

as its image representation only allows two-qubit gates to act on adjacent qubits. The quantum circuit consists of 35 quantum gates, with a max circuit depth of 20.

TF-QAS also demonstrates satisfying performance in the gatewise search space with limited connectivity between qubits. As shown in Table 2, TF-QAS achieves the ground state energy with a probability of 50% using 35 quantum gates, whereas HEA with 3 layers fails to achieve the ground state energy even though it uses 42 gates. In order to achieve the ground state energy, TF-QAS requires 137 queries, significantly fewer than the 6649 queries required by RS as shown in Fig. 1. This reflects a 48-fold improvement in sampling efficiency.

### Rationality Analysis

Table 3 displays the average time of calculating the number of paths, expressibility, and ground-truth performance of a circuit generated through the layerwise pipeline. The com-

Method	#gates	Prob.	$E - E_0$	Design
RS	35	0%	$9.7 \times 10^{-2}$	Auto
TF-QAS	35	50%	$3.5 \times 10^{-3}$	Auto
HEA-3	42	0%	$1.2 \times 10^{-1}$	Manual
HEA-4	56	100%	$2.2 \times 10^{-4}$	Manual

Table 2: Probability of achieving the ground state energy of the Heisenberg model over 10 independent runs in the gate-wise search space using the topology of ibmq-quito.

VQE task	Path	Expressibility	Performance
TFIM	$1.8 \times 10^{-4}$	0.21	9.4
Heisenberg	$2.1 \times 10^{-4}$	0.21	10
BeH <sub>2</sub>	$7.8 \times 10^{-4}$	0.35	36

Table 3: Average time (in seconds) of calculating the number of paths, expressibility and ground-truth performance of a circuit. This average time is computed across 500 circuits.

putational cost of the path count is below  $7.8 \times 10^{-4}$  second in all VQE tasks, making it analogous to a zero-cost proxy. Although the computational cost of the expressibility is 45-103 times lower than that of the ground-truth performance, the computation of expressibility for a substantial number of quantum circuits remains time-consuming and inefficient.

Table 4 displays the rank correlation, assessed through both Spearman and Kendall’s Tau, between the values of the training-free proxies and the ground-truth performance. The expressibility-based proxy exhibits a strong correlation with the ground-truth performance, rendering it suitable for the precise selection of high-performance circuits from the candidate set. Despite the low correlation exhibited by the path-based proxy, it is capable of filtering out unpromising circuits from the extensive search space while retaining the optimal ones, as demonstrated in Fig. 4. We randomly select 50,000 circuits from the search space and sort them by the number of paths. Fig. 4 shows the proportion of circuits achieving the ground state energy within the top  $M$  quantum circuits out of all circuits attaining the ground state energy. Despite filtering out 90% of the circuits, more than 55% of the optimal circuits are still preserved. For the VQE for the Heisenberg model in gatewise search space, all the optimal circuits are preserved. Therefore, the two-stage progressive TF-QAS is rational. The fast but coarse path-based proxy is first used to filter out a substantial portion of unpromising circuits, followed by the use of the refined but relatively slow expressibility-based proxy to select high-performance circuits.

### Ablation Study

We assess the performance of QAS algorithms that exclusively depend on either expressibility-based or path-based proxies. This assessment aims to showcase the efficacy of merging these proxies in a two-stage progressive learning approach. In the case of the expressibility-based QAS and

VQE task	Proxy	Spearman	Kendall's Tau
TFIM	Path	0.21	0.16
	Expressibility	0.58	0.43
Heisenberg (Layerwise)	Path	0.54	0.39
	Expressibility	0.85	0.67
BeH <sub>2</sub>	Path	0.61	0.47
	Expressibility	0.82	0.64
Heisenberg (Gatewise)	Path	0.57	0.41
	Expressibility	0.81	0.62

Table 4: Correlation coefficients between the values of training-free proxies and the ground-truth performances, estimated from 6,000 circuits randomly selected from the search space.

VQE task	Method	#Queries	Time(h)
TFIM	Path	104.7	0.4
	Expressibility	73.4	2.6
	TF-QAS	33.5	0.5
Heisenberg (Layerwise)	Path	310.9	1.5
	Expressibility	85.8	2.5
	TF-QAS	36.5	0.5
BeH <sub>2</sub>	Path	189.0	3.5
	Expressibility	161.9	8.0
	TF-QAS	110.5	2.5
Heisenberg (Gatewise)	Path	702.0	2.3
	Expressibility	264.8	6.3
	TF-QAS	137.9	1.1

Table 5: Average number of queries to the ground-truth performances by different QAS algorithms required to achieve the ground state energy over 10 independent runs.

path-based QAS, we keep querying the ground-truth performances of quantum circuits from best to worst as ranked by the expressibility-based and path-based proxies until the ground state energy is achieved. Table 5 presents the average number of queries required by different QAS algorithms. In terms of computational time, the expressibility-based QAS exhibits the poorest performance, despite its inherent intuitive effectiveness. Due to the limited correlation between the path-based proxy and the ground-truth performance, the path-based QAS requires more queries for circuit performances than the other two methods. Although the path-based proxy is zero-cost, the path-based QAS generally demands more time to achieve ground state energy compared to TF-QAS, owing to its significantly higher query count. Interestingly, TF-QAS requires the fewest queries. This is because the performance of complex quantum circuits is determined by various properties, whereas a single proxy solely captures a solitary property of the quantum circuit. TF-QAS combines two training-free proxies to identify circuits that perform well across multiple training-free proxies, enabling better discovery of optimal circuits.

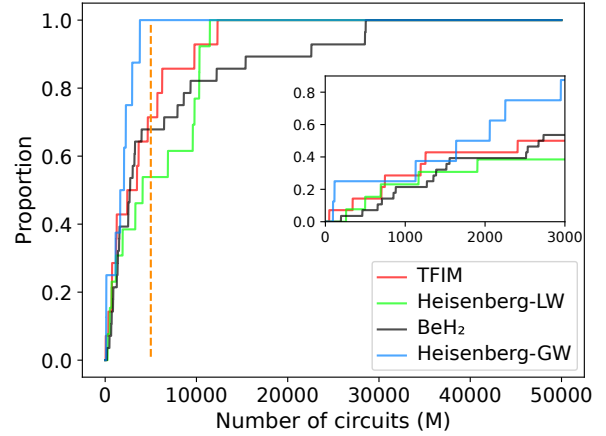


Figure 4: The variation of the proportion  $\tau = \eta_M / \eta$  w.r.t.  $M$ . The circuits are sorted by the number of paths, and the top  $M$  circuits are selected.  $\eta_M$  denotes the number of circuits achieving the ground state energy within the top  $M$  circuits.  $\eta$  represents the number of circuits achieving the ground state energy among all the 50,000 circuits. The values of  $\eta$  are 14, 13, 28, and 8 for TFIM, Heisenberg-LW, BeH<sub>2</sub>, and Heisenberg-GW, respectively. The inset enlarges the bottom-left portion of the figure.

## Conclusion

This paper employs two training-free proxies to rank quantum circuits, replacing the expensive circuit training process in current QAS algorithms. Through simulation results, we discern distinct properties of the path-based and expressibility-based proxies in estimating quantum circuit performance. Exploiting the zero cost associated with the path-based proxy and the strong correlation of the expressibility-based proxy, we devise a two-stage progressive training-free QAS (TF-QAS). VQE simulations involving the TFIM, Heisenberg model, and BeH<sub>2</sub> molecule demonstrate the superior capability of TF-QAS in identifying high-performance circuits, as well as the significantly lower computational cost compared to the SOTA PQAS. With these advantages, TF-QAS can be applied to VQAs involving larger-scale circuits, which remain challenging for training-based QAS algorithms. In future work, we will explore training-free QAS that takes into account trainability for large-scale QAS. Furthermore, the proposed proxies will be integrated into existing QAS algorithms including RL-based QAS and evolutionary QAS, providing feedback for these search strategies.

## Acknowledgments

This work is supported by Guangdong Basic and Applied Basic Research Foundation (2022A1515140116, 2022A1515010101), Innovation Program for Quantum Science and Technology (2021ZD0302901), National Natural Science Foundation of China (62272492, 61972091). SZ was supported by the Major Key Project of PCL.

## References

- Abdelfattah, M. S.; Mehrotra, A.; Dudziak, Ł.; and Lane, N. D. 2021. Zero-cost proxies for lightweight NAS. In *International Conference on Learning Representations*.
- Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S. C.; Endo, S.; Fujii, K.; McClean, J. R.; Mitarai, K.; Yuan, X.; Cincio, L.; et al. 2021. Variational quantum algorithms. *Nature Reviews Physics*, 3(9): 625–644.
- Chen, W.; Gong, X.; and Wang, Z. 2021. Neural architecture search on ImageNet in four GPU hours: a theoretically inspired perspective. In *International Conference on Learning Representations*.
- Cincio, L.; Rudinger, K.; Sarovar, M.; and Coles, P. J. 2021. Machine learning of noise-resilient quantum circuits. *PRX Quantum*, 2(1): 010324.
- Cong, I.; Choi, S.; and Lukin, M. D. 2019. Quantum convolutional neural networks. *Nature Physics*, 15(12): 1273–1278.
- Du, Y.; Huang, T.; You, S.; Hsieh, M.-H.; and Tao, D. 2022. Quantum circuit architecture search for variational quantum algorithms. *npj Quantum Information*, 8(1): 1–8.
- Farhi, E.; Goldstone, J.; and Gutmann, S. 2014. A quantum approximate optimization algorithm. *arXiv:1411.4028*.
- Grimsley, H. R.; Economou, S. E.; Barnes, E.; and Mayhall, N. J. 2019. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10(1): 1–9.
- He, Z.; Chen, C.; Li, L.; Zheng, S.; and Situ, H. 2022. Quantum architecture search with meta-learning. *Advanced Quantum Technologies*, 5(8): 2100134.
- He, Z.; Deng, M.; Zheng, S.; Li, L.; and Situ, H. 2023. GSQAS: graph self-supervised quantum architecture search. *Physica A: Statistical Mechanics and its Applications*, 630: 129286.
- Huang, Y.; Li, Q.; Hou, X.; Wu, R.; Yung, M.-H.; Bayat, A.; and Wang, X. 2022. Robust resource-efficient quantum variational ansatz through an evolutionary algorithm. *Physical Review A*, 105(5): 052414.
- Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J. M.; and Gambetta, J. M. 2017. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671): 242–246.
- Khatri, S.; LaRose, R.; Poremba, A.; Cincio, L.; Sornborger, A. T.; and Coles, P. J. 2019. Quantum-assisted quantum compiling. *Quantum*, 3: 140.
- Kuo, E.-J.; Fang, Y.-L. L.; and Chen, S. Y.-C. 2021. Quantum architecture search via deep reinforcement learning. *arXiv:2104.07715*.
- Las Heras, U.; Alvarez-Rodriguez, U.; Solano, E.; and Sanz, M. 2016. Genetic algorithms for digital quantum simulations. *Physical Review Letters*, 116(23): 230504.
- Lloyd, S.; and Weedbrook, C. 2018. Quantum generative adversarial learning. *Physical Review Letters*, 121(4): 040502.
- Lu, X.; Pan, K.; Yan, G.; Shan, J.; Wu, W.; and Yan, J. 2023. QAS-Bench: rethinking quantum architecture search and a benchmark. In *Proceedings of the International Conference on Machine Learning*, 22880–22898.
- McClean, J. R.; Romero, J.; Babbush, R.; and Aspuru-Guzik, A. 2016. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2): 023023.
- Mellor, J.; Turner, J.; Storkey, A.; and Crowley, E. J. 2021. Neural architecture search without training. In *International Conference on Machine Learning*, 7588–7598.
- Nam, Y.; Ross, N. J.; Su, Y.; Childs, A. M.; and Maslov, D. 2018. Automated optimization of large quantum circuits with continuous parameters. *npj Quantum Information*, 4(1): 1–12.
- Nielsen, M. A.; and Chuang, I. L. 2010. *Quantum computation and quantum information*. Cambridge University Press, 10th edition.
- Ostaszewski, M.; Trenkwalder, L. M.; Masarczyk, W.; Scerri, E.; and Dunjko, V. 2021. Reinforcement learning for optimization of variational quantum circuit architectures. In *Advances in Neural Information Processing Systems*, volume 34, 18182–18194.
- Romero, J.; Olson, J. P.; and Aspuru-Guzik, A. 2017. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2(4): 045001.
- Shi, J.; Wang, W.; Lou, X.; Zhang, S.; and Li, X. 2022. Parameterized Hamiltonian learning with quantum circuit. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5): 6086–6095.
- Shu, Y.; Cai, S.; Dai, Z.; Ooi, B. C.; and Low, B. K. H. 2022. NASI: label-and data-agnostic neural architecture search at initialization. In *International Conference on Learning Representations*.
- Sim, S.; Johnson, P. D.; and Aspuru-Guzik, A. 2019. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12): 1900070.
- Situ, H.; He, Z.; Wang, Y.; Li, L.; and Zheng, S. 2020. Quantum generative adversarial network for generating discrete distribution. *Information Sciences*, 538: 193–208.
- Wang, H.; Ding, Y.; Gu, J.; Lin, Y.; Pan, D. Z.; Chong, F. T.; and Han, S. 2022. QuantumNAS: noise-adaptive search for robust quantum circuits. In *International Symposium on High-Performance Computer Architecture*, 692–708.
- Wang, P.; Usman, M.; Paramalli, U.; Hollenberg, L. C.; and Myers, C. R. 2023. Automated quantum circuit design with nested monte carlo tree search. *IEEE Transactions on Quantum Engineering*.
- Wu, W.; Yan, G.; Lu, X.; Pan, K.; and Yan, J. 2023. QuantumDARTS: differentiable quantum architecture search for variational quantum algorithms. In *Proceedings of the International Conference on Machine Learning*, 37745–37764.
- Xu, J.; Zhao, L.; Lin, J.; Gao, R.; Sun, X.; and Yang, H. 2021. KNAS: green neural architecture search. In *Proceedings of the International Conference on Machine Learning*, 11613–11625.



- Zhang, S.-X.; Allcock, J.; Wan, Z.-Q.; Liu, S.; Sun, J.; Yu, H.; Yang, X.-H.; Qiu, J.; Ye, Z.; Chen, Y.-Q.; et al. 2023. Tensornetwork: a quantum software framework for the NISQ era. *Quantum*, 7: 912.
- Zhang, S.-X.; Hsieh, C.-Y.; Zhang, S.; and Yao, H. 2021. Neural predictor based quantum architecture search. *Machine Learning: Science and Technology*, 2(4): 045027.
- Zhang, S.-X.; Hsieh, C.-Y.; Zhang, S.; and Yao, H. 2022. Differentiable quantum architecture search. *Quantum Science and Technology*, 7(4): 045023.
- Zhang, Z.; and Jia, Z. 2021. GradSign: model performance inference with theoretical insights. In *International Conference on Learning Representations*.
- Życzkowski, K.; and Sommers, H.-J. 2005. Average fidelity between random quantum states. *Physical Review A*, 71(3): 032313.