HQC-NBV: A Hybrid Quantum-Classical View Planning Approach

Xiaotong Yu[1,2*] Chang Wen Chen[1†]
[1] The Hong Kong Polytechnic University, Hong Kong SAR
[2] National University of Defense Technology, China

Abstract

Efficient view planning is a fundamental challenge in computer vision and robotic perception, critical for tasks ranging from search and rescue operations to autonomous navigation. While classical approaches, including sampling-based and deterministic methods, have shown promise in planning camera viewpoints for scene exploration, they often struggle with computational scalability and solution optimality in complex settings. This study introduces HQC-NBV, a hybrid quantum-classical framework for view planning that leverages quantum properties to efficiently explore the parameter space while maintaining robustness and scalability. We propose a specific Hamiltonian formulation with multi-component cost terms and a parameter-centric variational ansatz with bidirectional alternating entanglement patterns that capture the hierarchical dependencies between viewpoint parameters. Comprehensive experiments demonstrate that quantum-specific components provide measurable performance advantages. Compared to the classical methods, our approach achieves 7.9-49.2% higher exploration efficiency across diverse environments. Our analysis of entanglement architecture and coherence-preserving terms provides insights into the mechanisms of quantum advantage in robotic exploration tasks. This work represents a significant advancement in integrating quantum computing into robotic perception systems, offering a paradigm-shifting solution for various robot vision tasks.

1. Introduction

In unknown scene perception, determining where to move a camera next - known as the informative view planning problem - can mean the difference between success and failure in critical applications. For instance, in search and rescue operations, inefficient view planning can lead to crucial delays, where every minute matters for survival rates. Similar challenges exist in autonomous navigation and robotic manipulation, where systematic and efficient environment exploration directly impacts task completion time and resource utilization. The Next Best View (NBV) problem represents a fundamental challenge in computer vision and robotic exploration and perception, where the objective is to determine optimal sequential viewpoints to maximize information gained about the environment with each move. Solving the NBV problem effectively can significantly enhance the performance of robotic systems by ensuring that they gather the most relevant and useful visual data with minimal resources.

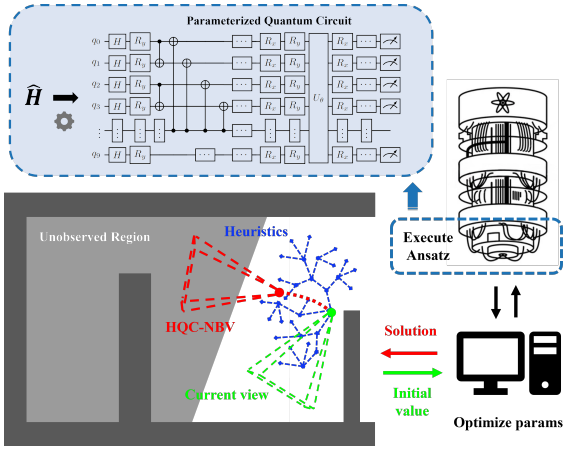

Figure 1. Execution logic of our HQC-NBV. Different from the classical approaches, we do not rely on heuristics but leverage quantum superposition to simultaneously evaluate multiple view parameters and quantum entanglement to capture complex dependencies between movement decisions.

Next-Best-View was initially introduced for exploring unknown areas using mobile robots [7, 21, 25, 28]. Early approaches can be primarily categorized into sampling-based and deterministic methods. Sampling-based approaches [4, 7, 22] typically employ Rapidly-exploring Random Trees (RRT) or RRT* within known free space [16, 19], generating candidate views and selecting the optimal one based on information gain versus cost metrics. While these methods have shown success in simple en-

*Part of this work was conducted during the visit to Nanyang Technological University. xiaotong.yu@connect.polyu.hk †Corresponding author. changwen.chen@polyu.edu.hk

vironments, they face significant scalability challenges in complex scenarios, often requiring exponentially increasing computational resources with environment size. Deterministic methods [26, 27, 31], on the other hand, rely on heuristics to guide viewpoint selection, focusing on specific metrics or uncertainty minimization. Despite their convenient and widespread deployment in real-world mobile platforms, these classical approaches suffer from fundamental limitations. Heuristic-based methods often struggle to find global optima, particularly in large-scale environments, while sampling techniques frequently result in suboptimal solutions due to their approximative nature of the solution space.

To address these challenges, we explore the potential of quantum computing in solving the NBV problem. Quantum computing has recently demonstrated promising results in various computer vision tasks, including multi-model fitting [12], multi-object tracking [30], motion segmentation [1], and graph matching [5, 6]. The quantum advantage derives from its ability to leverage quantum phenomena such as superposition and entanglement, enabling efficient exploration of vast solution spaces. The NBV problem is particularly well-suited for quantum approaches due to its combinatorial nature and the presence of complex parameter interdependencies that can be naturally encoded in quantum entanglement structures. Recent quantum implementations in the related problem [12] have shown the potential of adiabatic quantum computing (AQC) in disjoint set cover problems, suggesting a similar potential for view planning optimization.

Our work introduces a novel hybrid quantum-classical framework that combines the computational advantages of quantum systems with the robustness of classical optimization techniques, as shown in Figure 1. This hybrid approach aims to overcome the limitations of traditional methods while maintaining collocation with current quantum computing devices. Specifically, our contributions are as follows:

- A novel hybrid quantum-classical approach for informative view planning with a Hamiltonian formulation that maps robotic navigation into the quantum computing paradigm.
- A parameter-centric variational ansatz design with bidirectional alternating entanglement patterns that capture the hierarchical dependencies between view parameters for simultaneous exploration of movement directions, distances, and orientations.
- Comprehensive experimental validations demonstrate the contribution of quantum-specific components (i.e., entanglement architecture and coherence-preserving terms), as well as the robustness and effectiveness of our approach in different experimental settings, achieving 7.9-49.2% higher exploration efficiency compared to the classical methods.

To the best of our knowledge, this is the first study to propose a hybrid quantum-classical approach for informative view planning, opening new possibilities for efficient robot perception and navigation. The paper content is organised as follows: Section 2 provides an overview of the related work. Section 3 presents the problem formulation of NBV and our proposed hybrid quantum-classical approach. Section 4 presents the implementation of the proposed framework and an additional local strategy. The experimental setup and results are presented in Section 5. Finally, we analyse the results and draw the conclusion in Section 6.

2. Related Work

2.1. Informative View Planning

Informative view planning has evolved significantly since the 1980s [10, 20], with approaches falling into three main categories: surface-based, volumetric, and hybrid methods. Surface-based approaches [8, 16] represent environments as meshes and evaluate viewpoints through surface analysis, enabling direct quality assessment but requiring intensive visibility calculations [24]. Volumetric methods [13] employ voxel-based representations for computational efficiency, while hybrid approaches [17] combine both strategies for balanced accuracy and efficiency. The Next-Best-View (NBV) paradigm iteratively selects viewpoints to maximize information gain, with metrics evolving from simple voxel counting [2] to sophisticated entropy measures [11, 17]. Bircher et al. [7] introduced the receding horizon NBV (RH-NBV) approach, incorporating model predictive control to avoid local minima. Recent developments include ratio-based utility functions [23] and uncertainty-guided schemes [15]. Frontier-based methods [3, 9, 29] focus on exploration boundaries but lack NBV's flexible information gain formulation.

2.2. Quantum Computer Vision

Quantum computing offers unique advantages for computer vision through superposition, entanglement, and quantum parallelism. Recent works have applied quantum algorithms to various vision problems: Farina et al. [12] utilize quantum annealing for multi-model fitting, Zaech et al. [30] map multi-object tracking to Ising models, and Arrigoni et al. [1] reformulate motion segmentation for adiabatic quantum computing. Benkner et al. [5, 6] address quadratic assignment problems in shape matching through quantum annealing. Existing studies primarily focus on reformulating problems as Ising models or QUBO problems for adiabatic quantum computing. With the current Noisy Intermediate-Scale Quantum (NISQ) era [18], new opportunities emerge for solving complex problems using variational quantum algorithms and quantum approximate optimization.

3. Methodology

3.1. Problem Formulation

Consider a bounded 2D environment $S \subset \mathbb{R}^2$ containing obstacles $\mathcal{O} = \{O_1, O_2, ..., O_M\}$ to be explored. The camera's viewpoint is given by $v = (x, y, \theta) \in \mathcal{C}$. We aim to find the next best viewpoint that maximizes exploration while minimizing movement cost:

$$\min_{v \in \mathcal{C}} J(v) = -E(v) + \lambda_m M(v) \tag{1}$$

subject to:

$$P(v' \to v) \cap \mathcal{O} = \emptyset \tag{2}$$

where $E(v)$ denotes the exploration benefit function quantifying potential information gain, $M(v)$ denote the movement cost. λ_m is the weight parameter. And $P(v' \to v)$ represents the path between viewpoints. The exploration benefit function $E(v)$ measures the amount of new information gained by moving to viewpoint v, while $M(v)$ penalizes excessive movement. To evaluate $E(v)$ based on the historical and current views, we maintain an occupancy grid map $\mathcal{M}$ representing the accumulated knowledge of the environment, with each grid in a ternary state: unknown, free space, or occupied.

In practice, this formulation inherently becomes a combinatorial optimization problem. The continuous viewpoint space is discretized into a finite set $\mathcal{V} = \{v_1, ..., v_N\}$ of feasible positions. The collision-free constraint creates a finite feasible set $\mathcal{F} \subseteq \mathcal{V}$. The exploration benefit $E(v)$ depends on discrete visibility relationships derived from the $\mathcal{M}$, where each preference depends on the discrete distribution of unknown regions in the environment. Therefore, our problem becomes:

$$v^* = \arg\min_{v \in \mathcal{F}} J(v) \tag{3}$$

where $\mathcal{F}$ is a finite feasible set with solution space complexity $O(|\mathcal{F}|)$, making it naturally suited for quantum algorithms that can leverage superposition to explore multiple discrete solutions simultaneously.

3.2. Proposed Method

3.2.1. Problem Hamiltonian Formulation

We formulate the informative view planning problem as a combinatorial optimization task through a carefully designed Hamiltonian $\hat{H}$. This Hamiltonian is constructed as a weighted sum of Pauli operators, where each term encodes specific aspects of the exploration problem:

$$\hat{H} = \sum_i \alpha_i \hat{P}_i \tag{4}$$

Here, $\hat{P}_i$ represents a Pauli string (tensor product of Pauli matrices I, X, Y, Z), and α_i is the corresponding coefficient that determines the strength and direction of each term's contribution to the optimization objective. We decompose the Hamiltonian into five functional components:

$$\hat{H} = \hat{H}_{\text{dir}} + \hat{H}_{\text{dist}} + \hat{H}_{\text{adj}} + \hat{H}_{\text{orient}} + \hat{H}_{\text{coh}} \tag{5}$$

Our novel approach uses a 10-qubit system to encode viewpoint parameters for efficient exploration. The allocation of qubits is carefully designed to match the physical parameters' importance and range requirements. The main direction is encoded with the first 2 qubits. Distance and adjustment parameters are allocated 2 qubits each, providing sufficient precision for movement magnitudes while keeping the quantum circuit complexity manageable. The camera orientation angle receives 4 qubits to precisely direct the field of view toward information-rich regions. This allocation reflects the hierarchical nature of the exploration task while maintaining an efficient quantum representation with only 10 qubits total. This physical parameter encoding allows for effective quantum representation of navigation decisions.

The directional component $\hat{H}_{\text{dir}}$ encodes exploration preferences using Z operators on the first two qubits:

$$\hat{H}_{\text{dir}} = \sum_{i=0}^{1} \alpha_{dir,i} Z_i + \alpha_{ZZ} Z_0 Z_1 \tag{6}$$

Here Z_i denotes the Pauli-Z operator acting on qubit i. Identity operators on unspecified qubits are omitted for notational simplicity. The coefficients are directional exploration value derived from the occupancy map, e represents the unexplored density in each cardinal direction based on the area ratio, where E, N, W, and S correspond to East, North, West, and South directions, respectively:

$$\begin{aligned} \alpha_{dir,0} &= \lambda_{\text{dir},0} \cdot \tanh(e_W + e_S - e_E - e_N) \\ \alpha_{dir,1} &= \lambda_{\text{dir},1} \cdot \tanh(e_N + e_S - e_E - e_W) \\ \alpha_{ZZ} &= \lambda_{\text{diag}} \cdot \tanh(e_{SE} + e_{NE} - e_{SW} - e_{NW}) \end{aligned} \tag{7}$$

α_{ZZ} captures directional interdependencies. Here λ_{dir} and λ_{diag} are weighting parameters.

The distance component $\hat{H}_{\text{dist}}$ controls movement magnitude using Z operators on distance-encoding qubits:

$$\hat{H}_{\text{dist}} = \sum_{i=0}^{1} \alpha_{Z_{\text{dist},i}} Z_{i+2} \tag{8}$$

with coefficients proportional to observed obstacle proximity and bit significance:

$$\alpha_{Z_{\text{dist},i}} = \lambda_{\text{dist}} \cdot 2^{-(i+1)} \cdot \frac{d_{\text{obs}}}{d_{\max}} \tag{9}$$

The adjustment component $\hat{H}_{\text{adj}}$ provides fine-tuning of the movement direction through Z operators on adjustment-encoding qubits:

$$\hat{H}_{\text{adj}} = \sum_{i=0}^{1} \alpha_{Z_{\text{adj},i}} Z_{i+4} \tag{10}$$

where the coefficients are scaled by both bit significance and the remaining unexplored area:

$$\alpha_{Z_{\text{adj},i}} = \lambda_{\text{adj}} \cdot 2^{-(i+1)} \cdot (1-c) \tag{11}$$

Here, $(1-c)$ represents the proportion of unexplored environment, ensuring that directional adjustments become more precise as exploration progresses. The exponential term $2^{-(i+1)}$ maintains the binary significance hierarchy, with higher-order bits contributing more substantially to the directional refinement.

The orientation component $\hat{H}_{\text{orient}}$ combines target direction terms with exploration-promoting terms:

$$\hat{H}_{\text{orient}} = \sum_{i=0}^{3} \alpha_{Z_{\text{orient},i}} Z_{i+6} + \sum_{i=0}^{3} \alpha_{X_{\text{orient},i}} X_{i+6} + \sum_{m,n} \alpha_{ZZ_{\text{couple}}} Z_m Z_n \tag{12}$$

The target direction coefficients are:

$$\alpha_{Z_{\text{orient},i}} = \lambda_{\text{orient-Z}} \cdot 2^{-(i+1)} \cdot \rho \cdot (1-D) \cdot b_i \tag{13}$$

where ρ represents normalized point density, D is angular dispersion, and $b_i \in \{-1, 1\}$ encodes the target angle. Meanwhile, the ZZ operator handles the highest 2-bit coupling between direction and orientation with the coefficient:

$$\alpha_{ZZ_{\text{couple}}} = \lambda_{\text{orient-ZZ}} \cdot (1-D) \tag{14}$$

For high dispersion scenarios, the exploration coefficients are:

$$\alpha_{X_{\text{orient},i}} = \lambda_{\text{orient-X}} \cdot D \cdot (\gamma)^i \tag{15}$$

where $\lambda_{\text{orient-X}}$ is a weighting parameter and γ is a decay factor for higher-order bits. Finally, the coherence component $\hat{H}_{\text{coh}}$ maintains quantum advantage through X operators and entangling terms:

$$\hat{H}_{\text{coh}} = \sum_i \alpha_{X_i} X_i + \sum_{(i,j)\in\mathcal{P}} \alpha_{XX_{i,j}} X_i X_j \tag{16}$$

where $\mathcal{P}$ represents selected qubit pairs (based on the physical meaning and correlation between view parameters. i.e., direction-adjustment, distance-adjustment, direction-orientation). The entanglement coefficients scale with unexplored area:

$$\alpha_{XX_{i,j}} = \lambda_{\text{coh-XX}} \cdot (1-c) \tag{17}$$

And the X term promotes the exploration proportional to the coverage as well as the estimated information gain with Bresenham algorithm:

$$\alpha_{X_i} = \lambda_{\text{coh-X}} \cdot E_{exp} \cdot \gamma^i \cdot c \tag{18}$$

In our Hamiltonian formulation, the ground state of the Hamiltonian corresponds to the optimal next viewpoints for efficient environment exploration. This correspondence is established through the cost function encoding: since we minimize $J(v) = -E(v) + \lambda_m M(v)$ in the classical formulation, the Hamiltonian components are designed such that configurations yielding lower classical costs result in lower energy quantum states. Specifically, the directional terms $\hat{H}_{\text{dir}}$ favour movements toward unexplored regions, distance terms $\hat{H}_{\text{dist}}$ penalize excessive movement, and orientation terms $\hat{H}_{\text{orient}}$ promote information-rich viewing directions. The coherence terms $\hat{H}_{\text{coh}}$ maintain quantum superposition to explore multiple solutions simultaneously. Therefore, the ground state $|\psi_0\rangle$ satisfying $\hat{H}|\psi_0\rangle = E_0|\psi_0\rangle$ with minimum eigenvalue E_0 encodes the optimal viewpoint parameters that minimize the objective function $J(v)$. This ground state correspondence enables our variational quantum algorithm to approximate the optimal solution by minimizing the expectation value $\langle\psi(\vec{\theta})|\hat{H}|\psi(\vec{\theta})\rangle$, where the variational parameters $\vec{\theta}$ are optimized to approach the ground state configuration.

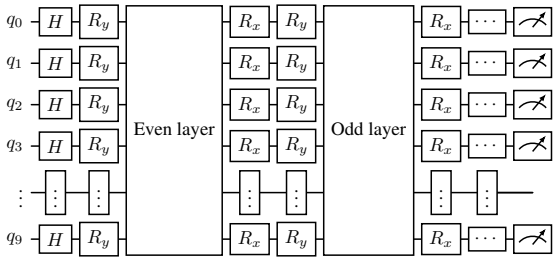

Figure 2. Block scheme of the proposed variational ansatz

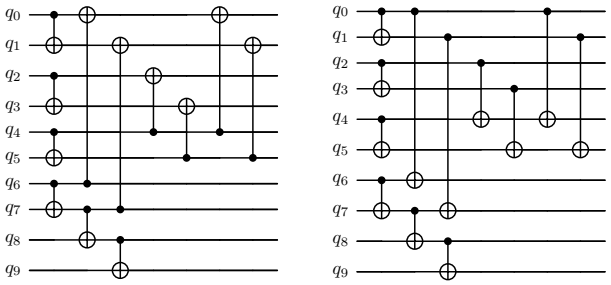

Figure 3. (Left) The block module with even index; (Right) The block module with odd index.

3.2.2. Variational Ansatz Design

We develop a multi-layered parameterized quantum circuit $U(\vec{\theta})$ that acts on n native qubits initialized in a uniform superposition state:

$$|\psi(\vec{\theta})\rangle = U(\vec{\theta})|+\rangle^{\otimes n} \tag{19}$$

where $|+\rangle^{\otimes n}$ represents the uniform superposition state obtained by applying Hadamard gates to all qubits in the $|0\rangle^{\otimes n}$ state.

The circuit architecture consists of $L = 5$ alternating layers of parameterized rotations and structured entanglement operations, the overview is shown in Figure 2:

$$U(\vec{\theta}) = U_L(\vec{\theta}_L) \cdots U_2(\vec{\theta}_2) U_1(\vec{\theta}_1) \tag{20}$$

Each layer $U_l(\vec{\theta}_l)$ comprises three key components:

$$U_l(\vec{\theta}_l) = U_l^{\text{rx}}(\vec{\theta}_l^{\text{rx}}) \cdot U_l^{\text{ent}} \cdot U_l^{\text{rot}}(\vec{\theta}_l^{\text{rot}}) \tag{21}$$

The rotational component $U_l^{\text{rot}}(\vec{\theta}_l^{\text{rot}})$ applies R_y rotations to encode the parameters into the quantum state. These rotations are partitioned according to the parameter groups:

$$U_l^{\text{rot}}(\vec{\theta}_l^{\text{rot}}) = \bigotimes_{i=0}^{1} R_y(\theta_{l,i}^{\text{dir}}) \otimes \bigotimes_{i=2}^{3} R_y(\theta_{l,i}^{\text{dist}}) \otimes \bigotimes_{i=4}^{5} R_y(\theta_{l,i}^{\text{adj}}) \otimes \bigotimes_{i=6}^{9} R_y(\theta_{l,i}^{\text{orient}}) \tag{22}$$

This structured encoding allows the circuit to independently modulate each parameter while maintaining correlations through subsequent entanglement operations.

The entanglement component U_l^{ent} establishes quantum correlations between qubits following a two-level hierarchical strategy, as is shown in Figure 3: intra-group entanglement followed by inter-group entanglement. The intra-group entanglement creates linear chains of CNOT gates within each parameter group:

$$U_l^{\text{intra}} = \prod_{g \in \{\text{dir,dist,adj,orient}\}} \prod_{i=l_g}^{l_g+n_g-2} \text{CNOT}_{i,i+1} \tag{23}$$

where l_g and n_g represent the starting position and size of group g, respectively.

The inter-group entanglement establishes connections between parameter groups, with the pattern alternating between even and odd layers:

$$U_l^{\text{inter}} = \begin{cases} \text{CNOT}_{l_{\text{dir}},l_{\text{adj}}} \cdot \text{CNOT}_{l_{\text{dist}},l_{\text{adj}}} \cdot \text{CNOT}_{l_{\text{dir}},l_{\text{orient}}}, & l\%2 = 0 \\ \text{CNOT}_{l_{\text{orient}},l_{\text{dir}}} \cdot \text{CNOT}_{l_{\text{adj}},l_{\text{dist}}} \cdot \text{CNOT}_{l_{\text{adj}},l_{\text{dir}}}, & l\%2 \neq 0 \end{cases} \tag{24}$$

This bidirectional entanglement pattern ensures information flow between parameter groups in both forward and reverse directions, facilitating complex correlations while maintaining circuit depth efficiency.

The final component in each layer applies R_x rotations to all qubits:

$$U_l^{\text{rx}}(\vec{\theta}_l^{\text{rx}}) = \bigotimes_{i=0}^{n-1} R_x(\theta_{l,i}^{\text{rx}}) \tag{25}$$

These rotations introduce non-commutativity with respect to the R_y rotations and the Z-based measurement observables, enhancing the circuit expressivity and enabling exploration of a larger subspace of the Hilbert space.

3.2.3. Optimization Process

The variational quantum circuit is optimized to minimize the expectation value of the cost Hamiltonian:

$$\vec{\theta}^* = \arg\min_{\vec{\theta}} \langle \psi(\vec{\theta}) | \hat{H} | \psi(\vec{\theta}) \rangle \tag{26}$$

where $\hat{H}$ is the problem-specific cost Hamiltonian incorporating exploration objectives, environmental constraints, and quantum coherence requirements. The optimization is performed using an adaptive Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm, which efficiently handles the high-dimensional parameter space of the variational circuit while being robust to the statistical noise inherent in quantum measurements. The adaptive learning rate mechanism adjusts the optimization step size based on progress metrics and stagnation detection:

$$\eta_{t+1} = \eta_t + \mu m_t + (1 - \mu)\Delta\eta_t \tag{27}$$

where μ is the momentum coefficient, m_t is the momentum term at iteration t, and $\Delta\eta_t$ is the learning rate adjustment based on recent optimization progress.

4. Implementation Details

The HQC-NBV system optimizes viewpoint selection for exploring unknown environments with obstacles. The process begins by initializing the scene with an initial viewpoint. In each iteration, before the coverage threshold is reached, the system updates the set of observed points by checking visibility from the current viewpoint. The variational ansatz is initialized for the current viewpoint and the problem Hamiltonian is constructed to encode the exploration objectives. The parameters are optimized using a Variational Quantum Eigensolver (VQE) with an adaptive SPSA optimizer, where the Hamiltonian-driven optimization is augmented with auxiliary constraints to ensure solution feasibility. The optimal parameters are decoded from Z expectation values to determine the next viewpoint, and trajectory validation ensures the new viewpoint lies within the observed area in $\mathcal{M}$ and avoids obstacle collisions. If the trajectory is invalid, we select the furthest valid position along the moving direction using a classical fallback strategy. The process iteratively executes to find out a sequence of optimal viewpoints, as detailed in Algorithm 1.

5. Experiments and Results

To demonstrate the effectiveness and robustness of the proposed HQC-NBV, we conduct a series of experiments on scenes with different areas and different obstacles. To examine the design of variational ansatz, we also conduct specialized experiments to isolate and quantify the contributions of key quantum components in our hybrid approach

Algorithm 1 Hybrid Quantum-Classical NBV System

Require: $S, v_0, \phi_{FOV}, d_{max}, \tau_{coverage}$
Ensure: $\mathcal{V} = \{v_0, v_1, \ldots, v_n\} : \mathcal{C}(\mathcal{V}) \geq \tau_{coverage}$

$\mathcal{V} \leftarrow \{v_0\}$, $\mathcal{M} \leftarrow$ InitializeMap(S)
$\mathcal{M} \leftarrow$ UpdateObservation$(\mathcal{M}, v_0)$
while $\mathcal{C}(\mathcal{M}) < \tau_{coverage}$ **do**
  $\hat{H} \leftarrow$ ConstructHamiltonian$(\mathcal{M}, v_t)$
  $U(\vec{\theta}) \leftarrow$ CreateParameterizedCircuit(n, L)
  $|\psi_0\rangle \leftarrow H^{\otimes n}|0\rangle^{\otimes n}$
  $\vec{\theta}_0 \leftarrow$ InitializeParameters()
  for $i = 0$ **to** $N_{iter} - 1$ **do**
    $c_i \leftarrow \langle\psi(\vec{\theta}_i)|\hat{H}|\psi(\vec{\theta}_i)\rangle + f(\vec{\theta}_i)$
    $g_i \leftarrow$ AdaptiveSPSA_GradientEsti.$(U, \hat{H}, \vec{\theta}_i)$
    $\eta_i \leftarrow$ AdaptiveLearningRate$(c_0, ..., c_i)$
    $\vec{\theta}_{i+1} \leftarrow \vec{\theta}_i - \eta_i \cdot g_i$
  end for
  $\vec{\theta}^* \leftarrow \vec{\theta}_{N_{iter}}$
  $\vec{z} \leftarrow$ Measure_Z_Expectations$(U(\vec{\theta}^*)|\psi_0\rangle)$
  $v_{next} \leftarrow$ DecodeParameters$(\vec{z}, v_t, \mathcal{M})$
  $valid \leftarrow$ ValidateTrajectory$(v_t, v_{next}, \mathcal{M})$
  if $valid$ **then**
    $v_{t+1} \leftarrow v_{next}$
  else
    $v_{t+1} \leftarrow$ ClassicalFallbackStrategy$(\mathcal{M}, v_t, v_{next})$
  end if
  $\mathcal{V} \leftarrow \mathcal{V} \cup \{v_{t+1}\}$
  $\mathcal{M} \leftarrow$ UpdateObservation$(\mathcal{M}, v_{t+1})$
end while
return $\mathcal{V}$

aiming to provide insights into how quantum characteristics—specifically entanglement patterns and coherence-preserving terms—impact exploration performance. In this study, the proposed method is implemented using the Qiskit framework, and all the experiments are performed on the Qiskit Aer backend simulator [14]. The camera parameters used in this study consistently respected a field of view (FOV) $2\pi/3$, and a maximum ray distance of 8 units. The starting view is initialized at a non-collision position.

5.1. Experimental Setup

We designed four distinct scenes with varying levels of complexity to comprehensively evaluate the robustness and scalability of the proposed methods: Scene 1 (S1) with surrounding obstacles, Scene 2 (S2) with central obstacle, Scene 3 (S3) with semi-enclosed complex walls and surrounding obstacles in 20×20 unit^2 areas, and Scene 4 (S4) with surrounding and central obstacles in a larger 40×40 unit^2 area. Visualized experimental scenes are provided in the supplementary material.

We comprehensively evaluate our approach against classical exploration methods (RH-NBV and frontier-based) and established optimization algorithms regarding exploration coverage ratio, path length, and exploration efficiency. Additionally, we conduct comparative experiments with established classical optimization algorithms to assess the advantage of our approach over general-purpose optimization approaches. To investigate the quantum-specific components, we conducted systematic ablation studies on both entanglement architecture variants (FA-Full Architecture, NE-Non-Entangled, IG-Intra-Group Only, EG-Inter-Group Only) and coherence-preserving term variants (CH-Complete Hamiltonian, NC-No Coherence Terms, SQX-Single-Qubit X Only). Detailed configurations are described in the supplementary material.

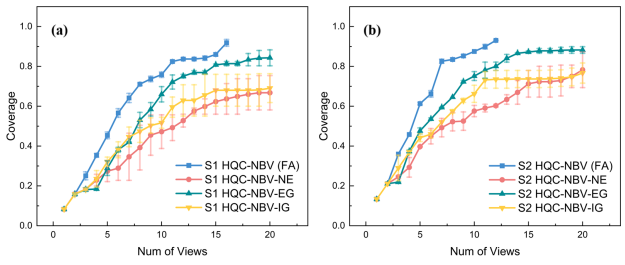

Figure 4. The effectiveness of entanglement architecture on the exploration performance: (a) coverage ratio in Scene 1; (b) coverage ratio in Scene 2.

In addition to these, we also comprehensively evaluate the performance of our approach against two classical exploration approaches, RH-NBV and the frontier-based method, regarding the exploration coverage ratio, path length and exploration efficiency. Additionally, we conduct comparative experiments with established classical optimization algorithms to assess the advantage of our approach over general-purpose optimization approaches.

5.2. Experimental Results

Figure 4 presents the comparative results, demonstrating that the full bidirectional entanglement architecture consistently outperformed reduced-entanglement variants. Notably, the non-entangled circuit required an average of 61.11% and 57.14% more views to achieve 65% coverage in S1 and S2 respectively, highlighting the significant role of quantum correlations in effective exploration planning. The inter-group-only variant performed better than the intra-group-only variant in both scenes, suggesting that maintaining cross-parameter entanglement between parameter groups is more critical than the entanglement within parameter groups. The intra-group entanglement architecture also contributes to the improvement compared to the non-entanglement variant because of the intrinsic connection between qubits within the logical groups of the infor-

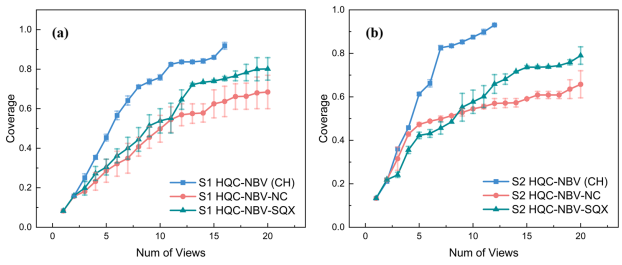

Figure 5. The evaluation of coherence-preserving term on the exploration performance: (a) coverage ratio in Scene 1; (b) coverage ratio in Scene 2.

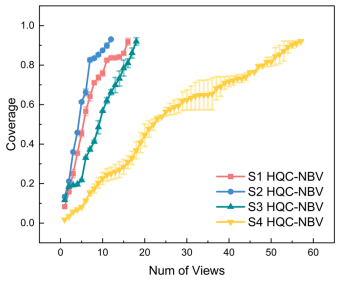

Figure 6. The coverage against the number of views of our approach in different scenes

mative view planning. Figure 5 illustrates that the absence of coherence-preserving terms led to frequent entrapment in local minima, with the no-coherence variant failing to achieve above 68.46% coverage and 65.77% on average in Scene 1 and Scene 2, respectively. The performance degradation was most pronounced in later exploration stages (coverage $>$ 50%), where remaining unexplored regions became sparse and disconnected. The single-qubit X-only variant demonstrated intermediate performance, maintaining reasonable exploration capabilities but showing reduced ability to escape local minima in complex scenarios. This suggests that while single-qubit superposition maintenance contributes to exploration effectiveness, the two-qubit coherence terms play a crucial role in coordinating parameter updates across different aspects of the navigation decision. Figure 6 demonstrates the robustness and scalability of our approach. Our approach performs an efficient exploration in S1, S2 and S3 within 15 viewpoints. The coverage growth in S2 is more dramatic than that in S1 and S3 due to the simplicity of the scene. The exploration in S4 requires 56 viewpoints to achieve comparable coverage, which is roughly four times the number required for S1, S2 and S3. This scaling factor meets the simple 4:1 ratio of environment sizes (S4 is four times larger in area than S1, S2 and S3), suggesting that the proposed approach does not degrade with the increase in environment size and complexity.

HQC-NBV consistently outperforms classical methods across all evaluation scenarios. In Scene 1, our method achieves 92.85% coverage within 16 viewpoints, while RH-NBV and frontier-based methods reach only 80.54% with the same views. Scene 2 demonstrates advantages with HQC-NBV achieving 93.02% coverage in 12 viewpoints compared to RH-NBV requiring twice as many viewpoints for 78.27% coverage. In the most challenging Scene 3, while frontier-based methods shows initial advantages in open spaces, they encounter early termination at 81.75% coverage due to disconnected unexplored regions. Our approach maintains consistent progress, achieving 91.97% coverage within 18 viewpoints as shown in Figure 7. Beyond coverage efficiency, HQC-NBV demonstrates 9.60-27.92% reduction in total path length and 16.19-30.75% higher exploration efficiency compared to classical approaches. The method also exhibits superior stability across multiple runs, particularly during the critical middle phase of exploration, as is shown in Figure 9. Comparison with traditional optimization methods (Powell and COBYLA) shows even more significant advantages, as shown in Figure 8. While classical optimizers frequently become trapped in local minima with coverage plateauing below 67.1%, HQC-NBV maintains superior performance across all complexity levels, highlighting the scalability advantages of our quantum-enhanced framework. The proposed approach outperforms traditional optimizers since we reformulate the conventional discrete information gain, which is ill-suited for optimization, into a continuous, differentiable Hamiltonian expectation. This expectation provides smooth gradients with respect to the variational parameters θ, while the underlying Hamiltonian encodes the exploration objectives through structured Pauli operators with smooth coefficient functions.

Figure 10 provides crucial insights into the optimization process underlying HQC-NBV performance. The left figure demonstrates our method's parallel search capability, taking the directional qubits as the example, where all four directional choices (East, North, West, South) are simultaneously evaluated through quantum superposition. Unlike classical methods that evaluate directions sequentially, the quantum approach maintains probability distributions for each direction throughout the optimization process. The evolution shows how initially uniform probabilities (0.25 each) gradually converge toward the optimal choice, with South direction emerging as the final selection with 59% probability after 500 iterations. The right figure denotes the quantum decision formation process, showing how quantum superposition gradually resolves into a definitive choice. The color-coded bars represent decisiveness levels ($P_{max} - P_{second}$), progressing from superposition to definitive decisions. This visualization captures the optimization transitions from exploring all possibilities simultaneously to converging on the

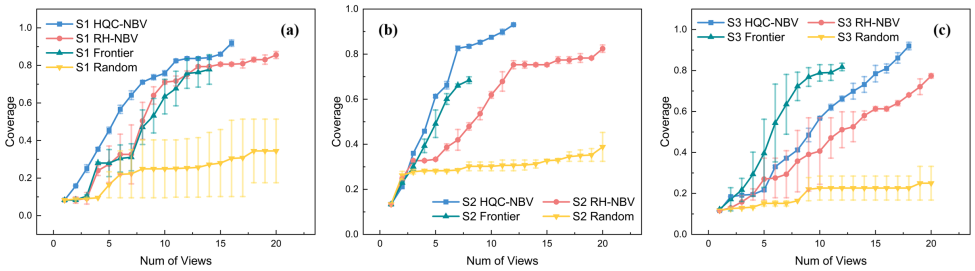

Figure 7. Comparison of coverage ratio progression between HQC-NBV, RH-NBV, and Frontier-based approaches in: (a) Scene 1; (b) Scene 2 and (c) Scene 3.

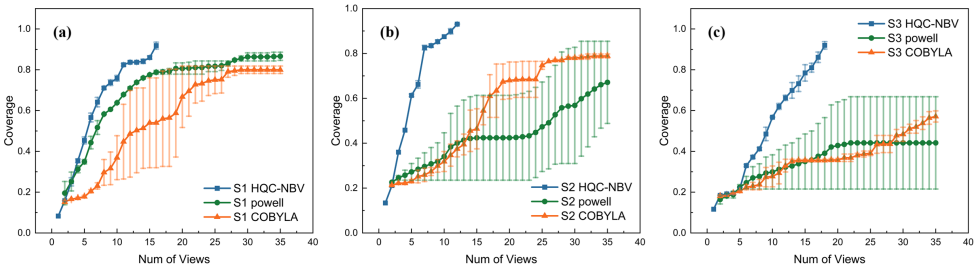

Figure 8. Comparison of coverage ratio progression between HQC-NBV and classical optimization methods Powell and COBYLA in: (a) Scene 1; (b) Scene 2 and (c) Scene 3.

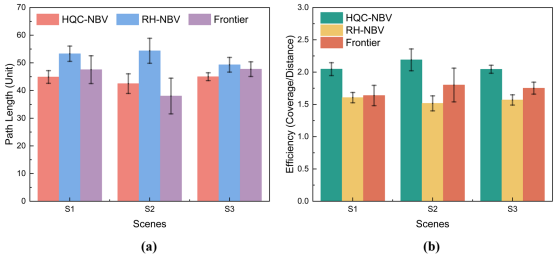

Figure 9. Performance metrics comparison: (a) Total path length across different scenes; (b) Exploration efficiency measured as coverage-to-distance ratio.

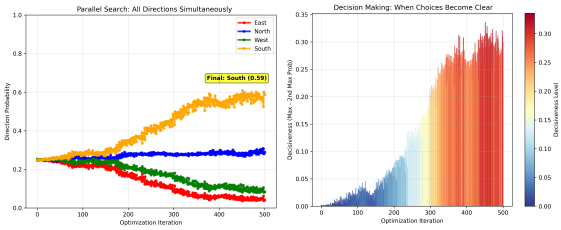

Figure 10. Optimization insight (direction qubits example): (Left) Evolution of directional probabilities; (Right) Decision formation process measured by decisiveness

optimal solution. This figure effectively illustrates the optimization simultaneous exploration of all possibilities followed by gradual convergence to optimal solutions. The smooth probability evolution and stable decision formation explain the observed performance improvements and reduced variance in exploration outcomes. More visualized experiments are presented in the supplementary materials.

6. Conclusion

In this paper, we introduced Hybrid Quantum-Classical Next-Best-View (HQC-NBV), a new paradigm for autonomous exploration. As a proof of concept, we restricted to a 2D setting with simplified sensing to validate the core idea in a controlled environment. Experiments show that quantum-specific components, especially entanglement and coherence-preserving terms, yield measurable gains: our method achieves up to 95.8% coverage and improves exploration efficiency relative to travel length by 7.9–49.2% over classical approaches, while maintaining scalability across environments of increasing complexity. This work represents a first step toward integrating quantum variational algorithms into robot vision problems. Future work will extend the framework to more realistic NBV settings, including 3D representations, richer view parameterizations, and uncertainty-aware formulations. Systematic evaluation in these settings is left as an important direction.

7. Acknowledgement

This research is supported by the Hong Kong Research Grants Council (GRF-15229423). The author Xiaotong Yu would like to thank Prof. Mile Gu and Ms. Wenyu Guo for their valuable suggestions.

References

[1] Federica Arrigoni, Willi Menapace, Marcel Seelbach Benkner, Elisa Ricci, and Vladislav Golyanik. Quantum motion segmentation. In *European Conference on Computer Vision*, pages 506–523. Springer, 2022. 2

[2] Joseph E Banta, LR Wong, Christophe Dumont, and Mongi A Abidi. A next-best-view system for autonomous 3-d object reconstruction. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 30(5): 589–598, 2000. 2

[3] Ana Batinovic, Tamara Petrovic, Antun Ivanovic, Frano Petric, and Stjepan Bogdan. A multi-resolution frontier-based planner for autonomous 3d exploration. *IEEE Robotics and Automation Letters*, 6(3):4528–4535, 2021. 2

[4] Ana Batinovic, Antun Ivanovic, Tamara Petrovic, and Stjepan Bogdan. A shadowcasting-based next-best-view planner for autonomous 3d exploration. *IEEE Robotics and Automation Letters*, 7(2):2969–2976, 2022. 1

[5] Marcel Seelbach Benkner, Vladislav Golyanik, Christian Theobalt, and Michael Moeller. Adiabatic quantum graph matching with permutation matrix constraints. In *2020 International Conference on 3D Vision (3DV)*, pages 583–592. IEEE, 2020. 2

[6] Marcel Seelbach Benkner, Zorah Lähner, Vladislav Golyanik, Christof Wunderlich, Christian Theobalt, and Michael Moeller. Q-match: Iterative shape matching via quantum annealing. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7586–7596, 2021. 2

[7] Andreas Bircher, Mina Kamel, Kostas Alexis, Helen Oleynikova, and Roland Siegwart. Receding horizon" next-best-view" planner for 3d exploration. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 1462–1468. IEEE, 2016. 1, 2

[8] SY Chen and YF Li. Vision sensor planning for 3-d model acquisition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 35(5):894–904, 2005. 2

[9] Titus Cieslewski, Elia Kaufmann, and Davide Scaramuzza. Rapid exploration with multi-rotors: A frontier selection method for high speed flight. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2135–2142. IEEE, 2017. 2

[10] Cl Connolly. The determination of next best views. In *Proceedings. 1985 IEEE international conference on robotics and automation*, pages 432–435. IEEE, 1985. 2

[11] Jeffrey Delmerico, Stefan Isler, Reza Sabzevari, and Davide Scaramuzza. A comparison of volumetric information gain metrics for active 3d object reconstruction. *Autonomous Robots*, 42(2):197–208, 2018. 2

[12] Matteo Farina, Luca Magri, Willi Menapace, Elisa Ricci, Vladislav Golyanik, and Federica Arrigoni. Quantum multimodel fitting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13640–13649, 2023. 2

[13] Armin Hornung, Kai M Wurm, Maren Bennewitz, Cyrill Stachniss, and Wolfram Burgard. Octomap: An efficient probabilistic 3d mapping framework based on octrees. *Autonomous robots*, 34:189–206, 2013. 2

[14] Ali Javadi-Abhari, Matthew Treinish, Kevin Krsulich, Christopher J. Wood, Jake Lishman, Julien Gacon, Simon Martiel, Paul D. Nation, Lev S. Bishop, Andrew W. Cross, Blake R. Johnson, and Jay M. Gambetta. Quantum computing with Qiskit, 2024. 6

[15] Liren Jin, Xieyuanli Chen, Julius Rückin, and Marija Popović. Neu-nbv: Next best view planning using uncertainty estimation in image-based neural rendering. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11305–11312. IEEE, 2023. 2

[16] Sertac Karaman and Emilio Frazzoli. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*, 30(7):846–894, 2011. 1, 2

[17] Simon Kriegel, Christian Rink, Tim Bodenmüller, and Michael Suppa. Efficient next-best-scan planning for autonomous 3d surface reconstruction of unknown objects. *Journal of Real-Time Image Processing*, 10(4):611–631, 2015. 2

[18] Jonathan Wei Zhong Lau, Kian Hwee Lim, Harshank Shrotriya, and Leong Chuan Kwek. Nisq computing: where are we and where do we go? *AAPPS bulletin*, 32(1):27, 2022. 2

[19] Steven LaValle. Rapidly-exploring random trees: A new tool for path planning. *Research Report 9811*, 1998. 1

[20] Jasna Maver and Ruzena Bajcsy. Occlusions as a guide for planning the next view. *IEEE transactions on pattern analysis and machine intelligence*, 15(5):417–433, 1993. 2

[21] Zehui Meng, Hailong Qin, Ziyue Chen, Xudong Chen, Hao Sun, Feng Lin, and Marcelo H Ang. A two-stage optimized next-view planning framework for 3-d unknown environment exploration, and structural reconstruction. *IEEE Robotics and Automation Letters*, 2(3):1680–1687, 2017. 1

[22] Victor Massagué Respall, Dmitry Devitt, Roman Fedorenko, and Alexandr Klimchik. Fast sampling-based next-best-view exploration algorithm for a mav. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 89–95. IEEE, 2021. 1

[23] Lukas Schmid, Michael Pantic, Raghav Khanna, Lionel Ott, Roland Siegwart, and Juan Nieto. An efficient sampling-based method for online informative path planning in unknown environments. *IEEE Robotics and Automation Letters*, 5(2):1500–1507, 2020. 2

[24] William R Scott, Gerhard Roth, and Jean-François Rivest. View planning for automated three-dimensional object reconstruction and inspection. *ACM Computing Surveys (CSUR)*, 35(1):64–96, 2003. 2

[25] Magnus Selin, Mattias Tiger, Daniel Duberg, Fredrik Heintz, and Patric Jensfelt. Efficient autonomous exploration plan-

ning of large-scale 3-d environments. *IEEE Robotics and Automation Letters*, 4(2):1699–1706, 2019. 1

[26] Marco Steinbrink, Philipp Koch, Bernhard Jung, and Stefan May. Rapidly-exploring random graph next-best view exploration for ground vehicles. In *2021 European Conference on Mobile Robots (ECMR)*, pages 1–7. IEEE, 2021. 2

[27] Dan Xiang, Hanxi Lin, Jian Ouyang, and Dan Huang. Combined improved a* and greedy algorithm for path planning of multi-objective mobile robot. *Scientific Reports*, 12(1): 13273, 2022. 2

[28] Zhefan Xu, Di Deng, and Kenji Shimada. Autonomous uav exploration of dynamic environments via incremental sampling and probabilistic roadmap. *IEEE Robotics and Automation Letters*, 6(2):2729–2736, 2021. 1

[29] Brian Yamauchi. A frontier-based approach for autonomous exploration. In *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97.'Towards New Computational Principles for Robotics and Automation'*, pages 146–151. IEEE, 1997. 2

[30] Jan-Nico Zaech, Alexander Liniger, Martin Danelljan, Dengxin Dai, and Luc Van Gool. Adiabatic quantum computing for multi object tracking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8811–8822, 2022. 2

[31] Tobias Zaenker, Julius Rückin, Rohit Menon, Marija Popović, and Maren Bennewitz. Graph-based view motion planning for fruit detection. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4219–4225. IEEE, 2023. 2