The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

Towards Real-Time Neutral Atom Array Assembly via Unsupervised Hologram Generation and Path Optimization

Ge Yan[1*], **Yuchen Wang**[2*], **Junchi Yan**[2†]
[1]College of Computing and Data Science, Nanyang Technological University, Singapore
[2]School of Artificial Intelligence and School of Computer Science, Shanghai Jiao Tong University, Shanghai, China
ge.yan@ntu.edu.sg, {wizihzkd, yanjunchi}@sjtu.edu.cn

Abstract

The rapid and reliable assembly of defect-free atom arrays poses a fundamental challenge for neutral atom quantum computing. While parallel rearrangement methods using spatial light modulators show promise, they suffer from significant overhead in two sub-tasks: atom-site matching and hologram generation. We propose a framework to address these bottlenecks and enhance the efficiency and fidelity of the assembly process. It features a new optimization objective for atom-site matching that minimizes the longest movement path, and a Fourier U-Net model that integrates Fourier operators with image-to-image translation to enable real-time hologram generation. The model is trained in a fully self-supervised paradigm, leveraging the physical properties of holography to remove the need for costly ground-truth labels. Experimental results show our framework not only significantly outperforms the state-of-the-art supervised CNN-based model but also achieves an inference speed orders of magnitude faster than traditional iterative algorithms, enabling real-time, dynamic atom rearrangement.

Introduction

The pursuit of practical quantum computation has ignited exploration across a diverse landscape of physical systems. Among these, neutral atom platforms have rapidly emerged as a leading contender, offering inherent scalability, flexible array reconfigurability, and exceptionally long coherence times (Bluvstein et al. 2022; Evered et al. 2023; Manetsch et al. 2024). Despite this potential, the field faces a fundamental challenge: the rapid and reliable assembly of a large-scale, defect-free atom array, which is a dynamic and repetitive process that consumes a significant fraction of the available coherence budget. Traditional methods using acoustically-driven deflectors (AODs) to move atoms serially suffer from severe efficiency and precision limitations as array sizes scale (Endres et al. 2016; Barredo et al. 2016, 2018). While global parallel movement with spatial light modulators (SLMs) (Lee, Kim, and Ahn 2017; Kim et al. 2019b) bypasses this bottleneck, their reliance on computationally intensive, on-the-fly hologram calculation introduces a critical latency, rendering them antithetical to this real-time requirement. Thus, developing a highly parallel assembly method that is also computationally efficient remains a pivotal, unresolved problem.

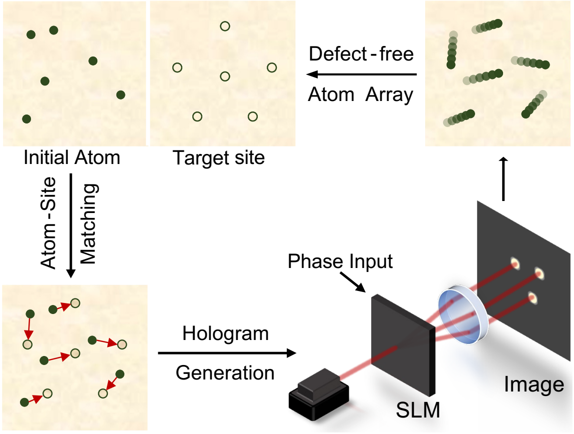

Figure 1: The defect-free atom array assembly pipeline.

Deep generative models offer a compelling paradigm to overcome this computational barrier. Architectures such as U-Nets (Ronneberger, Fischer, and Brox 2015) and Fourier Neural Operators (FNOs) (Li et al. 2020) excel at learning complex, non-linear mappings from high-dimensional inputs. Crucially, after an offline training phase, these models can generate high-fidelity solutions at an inference speed far surpasses traditional physics-based solvers. Inspired by these breakthroughs, we propose an AI-powered framework designed to efficiently address the two primary sub-tasks in neutral atom array assembly: atom-site matching and hologram generation (Figure 1).

For atom-site matching, we introduce a novel optimization objective that minimizes the longest single-atom path (a minimax approach) alongside the total path length (Lee, Kim, and Ahn 2017). This strategy reduces the number of parallel rearrangement steps, enhancing atom survival by minimizing assembly time. For rapid hologram generation, we propose the Fourier U-Net. This architecture embeds an FNO into the U-Net bottleneck, enabling it to capture the global, non-local correlations of light propagation while handling the sparse, discrete nature of target intensity

---

[*] These authors contributed equally.

[†] Corresponding author.

graphs. Finally, we leverage holographic physics for a fully self-supervised training paradigm. This framework trains the model using only randomly sampled target intensity graphs, eliminating the need for ground-truth phase maps and thus significantly reducing data acquisition and computational overhead. **Our contributions are:** 1) We propose a new atom-site matching optimization objective that minimizes the longest movement path to improve assembly efficiency and atom survival rate. 2) We introduce the Fourier U-Net, a model combining U-Net and FNO, to address the inherent global and sparse data challenges of hologram generation. 3) We establish an unsupervised training paradigm, leveraging physics for self-supervision, which significantly reduces data labeling and training costs.

Relation to prior work: This paper differs from previous work (Lin et al. 2025; Kim et al. 2019b) in two key aspects. First, we employ a minimax objective to minimize the longest individual atom path—the metric dictating total process time in parallel SLM rearrangement—rather than the aggregate travel distance. Second, our model enables direct, WGS-free inference, permitting real-time operation for large systems of up to 1000 atoms. In contrast, while the AI method of (Lin et al. 2025) also achieves real-time performance, it requires iterative WGS for input generation. The purely WGS-based approach (Kim et al. 2019b) is consequently non-real-time and limited to $\sim$100 atoms.

Preliminaries

In this section, we will provide former definitions of the two sub-tasks in atom array generation and essential algorithms for a better understanding of our proposed model.

Atom-Site Matching Given randomly loaded N atoms forming a set $\mathcal{A}$ and the target site set as $\mathcal{B}$, we can thus model the problem as a bipartite matching problem. Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be an undirected bipartite graph, where $\mathcal{V} = \mathcal{A} \cup \mathcal{B}$. Each edge $e_{ij} \in \mathcal{E}$ between $v_i \in \mathcal{A}$ and $v_j \in \mathcal{B}$ represents a feasible mapping between a randomly loaded atom and a target site. The edge weight d_{ij} for edge e_{ij} denotes the distance between node v_i and node v_j. The problem of finding optimal mappings is defined as follows in (Lin et al. 2025):

Definition 1. *Given a bipartite graph $\mathcal{G}(\mathcal{A} \cup \mathcal{B}, \mathcal{E})$, find a perfect matching $\mathcal{M} \subseteq \mathcal{E}$, such that*

$$|\mathcal{M}| = N, \quad \min \sum_{e_{ij} \in \mathcal{M}} d_{ij}. \tag{1}$$

Hungarian algorithm (Kuhn 1955) has been widely used in finding the minimum/maximum matching for a weighted bipartite graph. The key step in the Hungarian algorithm is utilizing the augmenting paths to extend the current matching. The time complexity to find the optimal solution is $O(N^3)$. (Lee, Kim, and Ahn 2017) first utilized this algorithm to solve the atom-slot matching problem in defect-free array generation. The authors mitigate the issue of crossing paths generated by the Hungarian algorithm by raising the distance in the loss function to a power. However, a formal proof for this approach is not provided in the paper.

Hologram Generation To precisely control the atoms during the moving process, we linearly interpolate the coordinates from $\mathbf{r}_1$ (initial position) to $\mathbf{r}_2$ (target position) with T steps. SLM receives an input laser beam and modulates the light by applying a spatially varying phase shift across the beam to produce a desired output pattern. The core challenge is to rapidly determine the phase distribution that, when applied to the input laser, will produce the desired intensity pattern at the image plane. Those positions with high intensity will form optical tweezers that trap atoms. We denote the intensity and phase distribution in the SLM 2D space as $I(x, y)$ and $\phi(x, y)$, the resulting light field is $U_{in} = \sqrt{I(x, y)} \cdot e^{i\phi(x,y)}$. The intensity at the image plane can be obtained via $\mathcal{F}$, the Fourier transformation, $\hat{I}(x, y) = |\mathcal{F}[U_{in}(x, y)]|^2$. Without loss of generosity, we set the initial intensity of the laser beam as a uniform distribution. Thus, the essence of hologram generation is to calculate the corresponding phase distribution $\phi(x, y)$ for each target intensity graph, which adds up to T phase distributions for one complete movement.

The Gerchberg-Saxton (GS) algorithm (Gerchberg 1972) is a foundational iterative method for phase retrieval, widely adopted for generating optical trap arrays (Kim et al. 2019a). The algorithm operates by alternating between the spatial domain (trap plane) and the Fourier domain (SLM plane). In each iteration, the known intensity constraints (e.g., the target trap locations and the SLM amplitude) are applied in their respective domains, while the phase is allowed to evolve. This process progressively refines the phase distribution to find a solution that satisfies the intensity constraints in both planes. However, the standard GS algorithm does not guarantee intensity uniformity across the generated traps. *The Weighted GS (WGS)* algorithm was introduced to address this (Kim et al. 2019a). WGS modifies the core GS iteration by introducing an adaptive weighting factor applied to the target intensity constraint. Conceptually, this factor serves as a feedback mechanism: trap locations with intensities higher than the array average are penalized (i.e., their target intensity is lowered), while dimmer spots are boosted. This iterative re-weighting effectively enforces a much more uniform intensity distribution across the final array.

Methodology

In this section, we will propose our approaches to solve the two main problems mentioned above.

Atom-Site Matching

The conventional objective for atom-site matching is to minimize the sum of Euclidean distances (as in Def. 1). However, this metric is ill-suited for neutral atom array assembly. In this process, all atoms are moved synchronously, and all individual paths are discretized into an identical number of steps. Consequently, the total assembly time is dictated by the single longest path (a minimax problem), not the aggregate distance traveled. The standard sum-of-squares objective does not guarantee this property. For example, consider $S_1 = (0, 0)$, $S_2 = (1, 0)$, $T_1 = (1, 10)$, and $T_2 = (10, 0)$, where S_i and T_i denotes the starting and terminal positions for atom i. The sum-of-squares optimal pairing P_1 has cost $C_1 = \|S_1 - T_1\|^2 + \|S_2 - T_2\|^2 = 182$, while P_2 has cost $C_2 = \|S_1 - T_2\|^2 + \|S_2 - T_1\|^2 = 200$. The algorithm selects P_1, but P_1 introduces a maximum path length of $\sqrt{101}$, which is longer than the $\sqrt{100}$ in P_2. In contrast to the standard formulation (Definition 1), we first introduce a new problem definition explicitly designed to address the min-max criterion:

Definition 2. *Given a bipartite graph $\mathcal{G}(\mathcal{A} \cup \mathcal{B}, \mathcal{E})$, find a perfect matching $\mathcal{M} \subseteq \mathcal{E}$, such that*

$$|\mathcal{M}| = N, \quad \min \max_{e_{ij} \in \mathcal{M}} d_{ij}. \tag{2}$$

To tackle this alternation, we propose a two-stage algorithm. For the fully connected bipartite graph $\mathcal{G}$, we first find the minimum and maximum distance (d_{min} and d_{max}, respectively). We conduct a binary search for the minimum max distance d_{opt} such that if we delete all the edges with distance larger than d_{opt}, we can still have a perfect matching. To efficiently check for perfect matches under the d_{opt} constraint, we resort to the Hopcroft-Karp (HK) algorithm, an $O(E\sqrt{N})$ algorithm to find the maximum matching in unweighted bipartite graphs. The time complexity for the combination of binary search and the HK algorithm is $O(\log d_{max} \times E \times \sqrt{N})$, which is bounded by the $O(N^3)$ complexity of the Hungarian Algorithm. Once we determine the minimized maximum distance d_{opt}, we move on to the second stage. We eliminate all the edges with a distance larger than d_{opt}, and conduct a Hungarian search for the optimal matching with minimum total cost.For this second stage, the Hungarian algorithm optimizes a sum-of-squares cost ($\min \sum_{e_{ij} \in \mathcal{M}} d_{ij}^2$), a secondary objective consistent with (Lee, Kim, and Ahn 2017; Lin et al. 2025).

While the authors of (Lee, Kim, and Ahn 2017) proposed this metric to help avoid atom collisions, we find that this squaring does not strictly guarantee collision-free paths. For example, consider $S_1 = (1, 0)$, $S_2 = (4, 1)$, $T_1 = (0, -1)$, and $T_2 = (0, -2)$. The pairing P_2, which involves a collision, has a smaller cost $C_2 = \|(1, 2)\|^2 + \|(4, 2)\|^2 = 25$, while the collision-free pairing P_1 has a larger cost $C_1 = \|(1, 1)\|^2 + \|(4, 3)\|^2 = 27$. Instead, its practical benefit is the maximization of the minimum instantaneous distance between atoms during assembly. This consideration is critical, as excessively close tweezers can induce premature and undesired interactions between the trapped atoms. We first analyze the geometric results of the widely used sum-of-squares cost. Without loss of generality, we consider the two-point case, as scenarios with more atoms can be decomposed into this case. The two pairs are denoted as $P_1 = \{(S_1, T_1), (S_2, T_2)\}$ with cost $C_1 = \|S_1 - T_1\|^2 + \|S_2 - T_2\|^2$, and $P_2 = \{(S_1, T_2), (S_2, T_1)\}$ with cost $C_2 = \|S_1 - T_2\|^2 + \|S_2 - T_1\|^2$.

Lemma 3. *The difference between the costs of the two pairings is equal to twice the negative dot product of their corresponding initial and final separation vectors: $C_1 - C_2 = -2(\overrightarrow{S_1S_2} \cdot \overrightarrow{T_1T_2})$.*

This result is a direct consequence of Euler's Quadrilateral Theorem (Debnath 2009), and it relates the algorithmic choice to the geometry of the separation vectors.

Corollary 4. *Let $\mathbf{u} = \overrightarrow{S_1S_2}$ be the initial separation vector, and the corresponding final separation vectors be $\mathbf{v}_1 = \overrightarrow{T_1T_2}$ and $\mathbf{v}_2 = \overrightarrow{T_2T_1}$. The pairing that minimizes the sum-of-squares cost is the one for which the dot product of the initial and final separation vectors is non-negative (i.e., $\mathbf{u} \cdot \mathbf{v}_{opt} \geq 0$).*

Proof. The algorithm selects the optimal pairing that minimizes the cost. We examine both possible outcomes:
Case 1: P_1 selected. This implies $C_1 \leq C_2$.

$$C_1 - C_2 \leq 0 \quad -2(\mathbf{u} \cdot \mathbf{v}_1) \leq 0 \quad \mathbf{u} \cdot \mathbf{v}_1 \geq 0$$

This pairing satisfies the non-obtuse condition.
Case 2: P_2 selected. This implies $C_2 \leq C_1$.

$$C_1 - C_2 \geq 0 \quad -2(\mathbf{u} \cdot \mathbf{v}_1) \geq 0 \quad \mathbf{u} \cdot \mathbf{v}_1 \leq 0$$

We must now check the dot product for the chosen pairing, P_2. The relevant dot product is $\mathbf{u} \cdot \mathbf{v}_2$:

$$\mathbf{u} \cdot \mathbf{v}_2 = \mathbf{u} \cdot (-\mathbf{v}_1) = -(\mathbf{u} \cdot \mathbf{v}_1)$$

Since this case requires $\mathbf{u} \cdot \mathbf{v}_1 \leq 0$, it follows that $-(\mathbf{u} \cdot \mathbf{v}_1) \geq 0$. Thus, $\mathbf{u} \cdot \mathbf{v}_2 \geq 0$.

In both cases, the pairing selected by the algorithm ensures that the dot product of its respective separation vectors is non-negative, implying an acute or right angle. □

Theorem 5. *The pairing that minimizes the sum of squared path lengths (and thus enforces an acute-angle relationship between separation vectors) also maximizes the minimum instantaneous distance between the moving points.*

Proof. Let $\mathbf{u} = \overrightarrow{S_1S_2}$, the final separation vector $\mathbf{v} = \overrightarrow{T_1T_2}$ for P_1. The instantaneous separation vector at a time fraction $\tau \in [0, 1]$ is given by the linear interpolation

$$\mathbf{d}_1(\tau) = (1 - \tau)\mathbf{u} + \tau\mathbf{v}.$$

Similarly, we have $\mathbf{d}_2(\tau) = (1 - \tau)\mathbf{u} - \tau\mathbf{v}$ for P_2. The difference between these squared distances can be derived as

$$\|\mathbf{d}_1(\tau)\|^2 - \|\mathbf{d}_2(\tau)\|^2 = 4\tau(1 - \tau)(\mathbf{u} \cdot \mathbf{v}).$$

From Corollary 4, the sum-of-squares optimal pairing is the one for which $\mathbf{u} \cdot \mathbf{v} > 0$. Since $\tau(1 - \tau) > 0$ for any $\tau \in (0, 1)$, the entire difference term is positive. This implies that $\|\mathbf{d}_1(\tau)\|^2 > \|\mathbf{d}_2(\tau)\|^2$ for the entire movement duration.

Consequently, the minimum squared distance for the optimal pairing must be greater than or equal to the minimum for the suboptimal pairing:

$$\min_{\tau \in [0,1]} \|\mathbf{d}_1(\tau)\|^2 \geq \min_{\tau \in [0,1]} \|\mathbf{d}_2(\tau)\|^2$$

Thus, the pairing that minimizes the sum of squared path lengths simultaneously maximizes the minimum separation distance between the moving atoms. □

Theorem 5 thus confirms that our two-stage approach—which integrates the primary minimax constraint (Stage 1) with the sum-of-squares optimization (Stage 2)—is guaranteed to simultaneously minimize the maximum atom movement and maximize the minimum instantaneous atom separation.

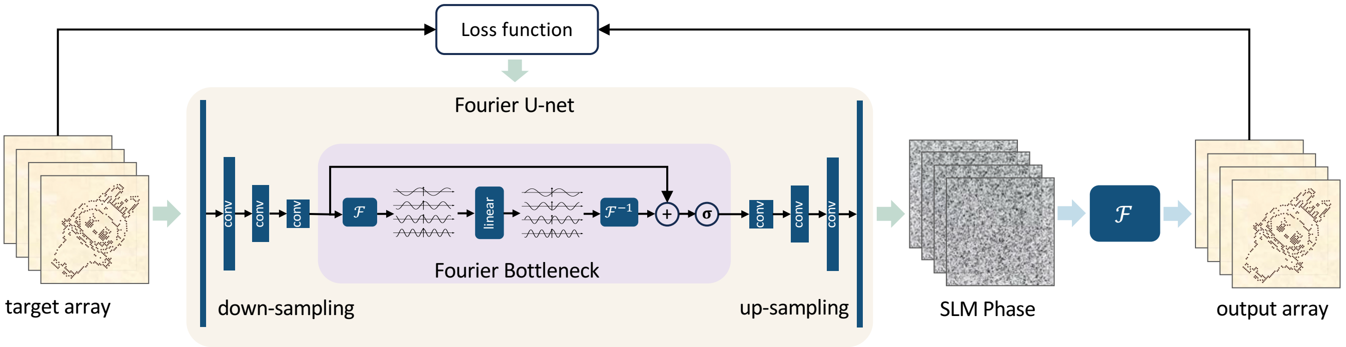

Figure 2: Fourier U-Net for self-supervised phase retrieval. The network employs a convolution encoder-decoder structure with a Fourier bottleneck. We first utilize a convolution layer with a stride = 2 to downsample the input intensity graph. A Fourier bottleneck applies a linear transformation in the frequency domain to extract the global features, and the outcome is upsampled to generate the phase image for SLM. Combined with the unified input intensity, we can reconstruct the target array with our phase output. An NPCC loss is applied to update the network.

Hologram Generation

The optimal atom rearrangement requires a series of steps, each demanding a new hologram to precisely control the optical tweezers. To efficiently generate the required phase-only hologram, $\phi(x, y)$, from a target intensity graph, $\hat{I}_{target}(x, y)$, we propose a self-supervised Fourier U-Net architecture. This self-supervised approach bypasses the need for ground-truth phase data. Instead, the model is trained by comparing the target intensity, $\hat{I}_{target}$, to a reconstructed intensity, $\hat{I}_{recon}$. This reconstruction is simulated by propagating the complex field from the SLM, $\hat{I}_{recon} = |\mathcal{F}[\sqrt{I(x, y)} \cdot e^{i\phi(x,y)}]|^2$, where $\phi(x, y)$ is the predicted phase and $I(x, y)$ is the incident laser intensity. Standard metrics like Mean Squared Error (MSE) fail to converge on the sparse, discrete targets of tweezer arrays. We therefore employ the Normalized Pearson Correlation Coefficient (NPCC) as the loss function,

$$\mathcal{L}_{\text{NPCC}} = 1 - \frac{\sum(\hat{I}_{\text{recon}} - \mu_{\text{recon}})(\hat{I}_{\text{target}} - \mu_{\text{target}})}{\sqrt{\sum(\hat{I}_{\text{recon}} - \mu_{\text{recon}})^2 \sum(\hat{I}_{\text{target}} - \mu_{\text{target}})^2}}, \tag{3}$$

where μ_{recon} and μ_{target} are the mean intensities calculated over the entire image. Equation 3 prioritizes the structural fidelity and uniformity of the reconstructed pattern over absolute intensity values, which is critical for generating high-uniformity traps.

As illustrated in Figure 2, our network uses a classic U-Net architecture (Ronneberger, Fischer, and Brox 2015) to map the input intensity graph to the output phase-only hologram. The final layer employs a hyperbolic tangent activation function to scale the output phase to the required $[-\pi, \pi]$ range. However, standard CNNs in U-net face two unique challenges in this task. First, the far-field intensity is a non-local manifestation of the SLM phase, a global relationship that local convolutional kernels cannot capture. Second, the target intensity graphs are sparse and discrete, making it difficult for standard CNNs to learn the long-range phase correlations required to form sharp, isolated peaks.

To address these limitations, we embed a Fourier Neural Operator (FNO) (Li et al. 2020) into the network's bottleneck, where feature density is highest. The FNO operates in the frequency domain: it applies a 2D Fourier transform to the feature map, applies a learnable linear transform to the low-frequency modes, and then applies an inverse Fourier transform.

$$v_{\text{out}} = \sigma\Big(\mathcal{F}^{-1}\big(R_\phi \cdot (\mathcal{F} v_{\text{in}})\big) + W(v_{\text{in}})\Big), \tag{4}$$

where R_ϕ is the learnable Fourier-space transform, W is a spatial-space linear transform, and σ is the activation function. This allows the model to efficiently capture the global dependencies that local kernels miss. Therefore, the proposed hybrid architecture uniquely combines the local feature extraction of the U-Net with the global, non-local modeling of the FNO.

Numerical Experiments

Atom-Site Matching

Datasets We designed two classes of rearrangement tasks representative of the primary applications for neutral atom quantum computing: large-scale array formation and programmable graph-problem encoding.

The first class, Square Array Assembly, simulates the compaction of atoms from a stochastically-loaded source grid (e.g., 60×60) into a smaller, dense, defect-free target (e.g., 45×45). This scenario benchmarks the algorithm's efficiency in scaling to large, uniform qubit registers. This capability is a fundamental prerequisite for fault-tolerant quantum computing, where logical qubits are encoded using surface or LDPC codes with dense, reconfigurable 'patches' of physical atoms (Bluvstein et al. 2024, 2025b).

The second class, UDG Preparation, addresses a key benchmark for neutral atom processors: solving hard combinatorial optimization problems (Pichler et al. 2018; Ebadi et al. 2022). This task involves arranging atoms into specific, arbitrary patterns representing Unit Disk Graphs (UDGs),

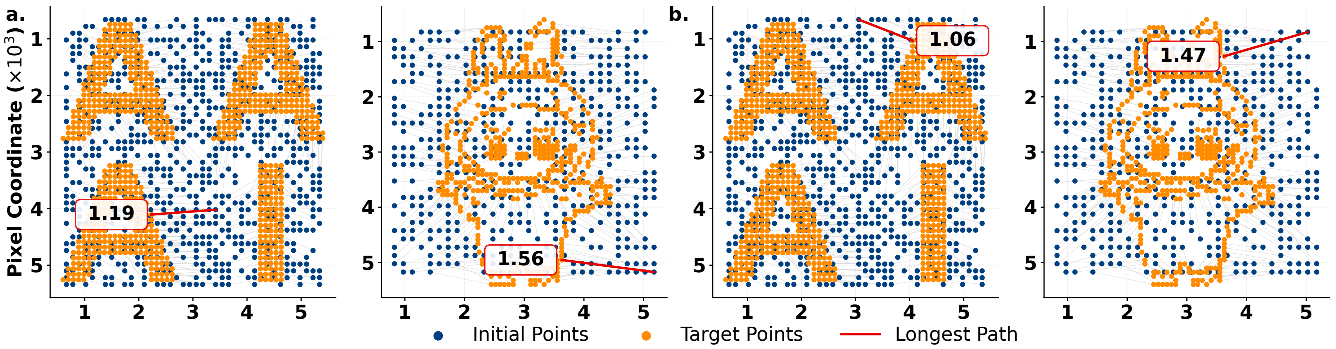

Figure 3: Visual comparison of solutions generated by the baseline TDH method (a) and our proposed MMH method (b) for two complex, non-grid target geometries: "AAAI" (left panel) and "labubu" (right panel). Blue dots are the initial atom positions, while orange dots form the target pattern. The red line in each subplot highlights the single longest path (d_{max}).

which are the native geometry for the Rydberg blockade. Such configurations are the standard input for studying the Maximum Independent Set problem, where neutral atom platforms have demonstrated a significant performance advantage. Our UDG datasets include configurations with 100 and 200 atoms. For each scenario, algorithms were tested against a fixed target configuration, with results averaged over 20 randomly generated initial atom distributions to ensure fair comparison.

Evaluation Metrics Performance is evaluated primarily by the Total Survival Probability (P_{total}), our core objective metric. This single metric is designed to model the final, physically-relevant outcome of a rearrangement, as it accounts for both movement-induced and collision-induced atom loss. The survival probability for a single atom i, denoted $P_{\text{single},i}$, is modeled as the product of two factors accounting for the primary sources of atom loss: movement-induced loss ($P_{\text{moving},i}$) and collision-induced loss ($P_{\text{cross},i}$). The first factor, $P_{\text{moving},i} = e^{-\beta \mathcal{T} d_i^2}$, accounts for loss due to the movement itself, where $\mathcal{T}$ denotes the number of steps. This loss depends on the squared Euclidean distance d_i^2 traveled by atom i and the total number of control steps $\mathcal{T}$. The second factor, $P_{\text{cross},i} = 1 - e^{-\gamma \delta^2_{\text{min},i}}$, represents the survival probability from potential collisions and is a function of the minimum squared distance $\delta^2_{\text{min},i}$ between the path of atom i and any other atom's path. The total survival probability can thus be given by: $P_{\text{total}} = \sum_{i=1}^{N} \left(P_{\text{moving},i} \cdot P_{\text{cross},i} \right)$.

To provide a comprehensive diagnosis of how an algorithm achieves its P_{total} score, we supplement this primary metric with three complementary diagnostics:

1) Maximum Moving Length (d_{max}), the longest single-atom path, which directly determines the total transport time.

2) Number of Steps ($\mathcal{T}$), the number of discrete control frames required.

3) Average Minimum Distance ($\bar{\delta}_{\text{min}}$), the mean clearance between all atom paths, serving as a proxy for the overall collision risk.

Implementation Details To ensure a direct comparison with prior work, the physical parameters for the survival probability calculation are set consistently with (Lee, Kim, and Ahn 2017): movement-induced loss coefficient $\beta = 0.0076/a^2$, collision-induced loss coefficient $\gamma = 80/a^2$, and lattice constant $a = 4.0$ µm. In our simulation coordinate system, 1 unit corresponds to 0.1 µm, and the maximum travel speed per control step was set to $v_{\text{max}} = 5.0$ units/step (0.5 µm/step). All reported results are averaged over the 20 random samples for each scenario.

Results and Discussion We evaluated our proposed Min-Max Hungarian (MMH) method against the Total-Distance Hungarian (TDH) baseline. As summarized in Table 1, MMH consistently and significantly outperforms TDH across all scenarios and key metrics.

The primary advantage of MMH is its ability to deliver a higher Total Survival Probability (P_{total}). This improvement is substantial, reaching a 3% relative increase in final atom yield in the challenging 60×60 compression task. This performance gain stems directly from MMH's core objective. By minimizing the maximum moving length (d_{max}) by 18-23% compared to the baseline, MMH significantly reduces the required number of control steps (N_{steps}). This reduction in rearrangement time is the dominant factor in mitigating time-dependent atom loss.

Crucially, MMH achieves this via an advantageous trade-off. It accepts a marginal (typically $< 1\%$) decrease in the average minimum distance between paths in exchange for the large reduction in global movement steps. This demonstrates that prioritizing the minimization of the temporal bottleneck (d_{max}) is a more effective strategy for maximizing final array fidelity than a strategy focused on total path length.

The effectiveness of this approach is qualitatively demonstrated in Figure 3 for complex "AAAI" and "labubu" target geometries. The visuals clearly show that while the TDH baseline creates excessively long bottleneck moves, our MMH method finds solutions with shorter maximum paths. This confirms that MMH is not only quantitatively superior but also better suited for assembling arbitrary, complex atom arrangements, making it a practical and compelling choice for quantum applications.

Hologram Generation

Datasets We created six datasets to test our method's performance under different conditions. The datasets are split into

Method & Metric		20x20 to 15x15	40x40 to 32x32	60x60 to 45x45	UDG 100	UDG 200
MMH	P_{total}	**196.90**	**933.06**	**1684.81**	**81.96**	**168.25**
	Survival ratio %	**87.51**	**91.12**	**83.20**	**81.96**	**84.12**
	$d_{\max}$ (μm)	**8.10**	**7.60**	**9.40**	**8.92**	**8.73**
	$\bar{\delta}_{\min}$ (μm)	3.86	3.81	3.85	3.77	3.77
	$\mathcal{T}$	**16.8**	**15.7**	**19.2**	**18.0**	**17.6**
TDH	P_{total}	193.48	915.76	1618.76	78.75	162.22
	Survival ratio %	85.99	89.43	79.94	78.75	81.11
	$d_{\max}$ (μm)	9.95	9.54	12.23	11.38	11.00
	$\bar{\delta}_{\min}$ (μm)	**3.87**	**3.82**	**3.86**	**3.79**	**3.78**
	$\mathcal{T}$	20.5	19.5	24.9	23.1	22.2

Table 1: A comprehensive comparison of our proposed MMH method and the baseline TDH method across all datasets and metrics. For each comparison, the superior result is highlighted in bold.

two scenarios: Static Distribution and Dynamic Rearrangement, each with 100, 400, and 900 trap arrays. For the static datasets, trap positions were chosen randomly on a grid. For the dynamic datasets, which simulate moving atoms, we defined the paths using linear interpolation between random starting points and a fixed target layout.

The approaches differ significantly in their data requirements. Our self-supervised model learns directly from the target coordinates and does not require pre-computed holograms. In contrast, the supervised CNN requires a complex data preparation pipeline. We follow the methodology in (Lin et al. 2025) to generate the labels for the CNN model.

Baselines We evaluate against two primary baseline models to provide a comprehensive performance comparison.

WGS: A widely-used iterative algorithm that serves as a strong classical benchmark. We run the WGS algorithm for 30 iterations on each test case.

Supervised CNN (Lin et al. 2025): The state-of-the-art deep learning model. This model is trained on pre-computed ground-truth holograms and requires a specialized input format consisting of separate amplitude and phase images.

Evaluation Metrics To quantify the performance of the generated holograms, we employ the following three standard metrics (Di Leonardo, Ianni, and Ruocco 2007).

1) Efficiency (e): the total optical power delivered to the traps $e = \sum_m I_m$, where I_m denotes the intensity of trap m. Higher efficiency generally creates deeper traps.

2) Uniformity (u): the evenness of the optical field amplitudes ($\sqrt{I_m}$) at the trap plane $u = 1 - \frac{\max[\sqrt{I_m}]-\min[\sqrt{I_m}]}{\max[\sqrt{I_m}]+\min[\sqrt{I_m}]}$, where a value of 1 is ideal. High uniformity is essential for quantum computing, as intensity variations cause unequal AC Stark shifts, a primary source of quantum gate error.

3) Percent Standard Deviation (s): measurement of the intensity variation across the traps $s = 100\frac{\sqrt{\langle(\sum_m(I_m-\langle I\rangle)^2)\rangle}}{\langle I\rangle}$, where $\langle\cdot\rangle$ denotes the mean. Lower s values indicate that all traps provide nearly identical trap depths.

Implementation Details All models were implemented in PyTorch and trained on an NVIDIA RTX 4090 GPU. The static (3K samples) and dynamic (2K samples) datasets were partitioned with a 90/10 training-to-testing ratio. Our Fourier U-Net was trained from scratch for 200 epochs using the Adam optimizer (betas=(0.9, 0.999)) with a batch size of 2. A cosine annealing scheduler decayed the learning rate from an initial 10^{-4} to 10^{-7}.

Model	Metric	Static-100	Static-400	Static-900
U-Net	e (%)	7.46	7.16	9.76
	u (%)	96.19	85.59	88.11
	s (%)	3.03	13.19	7.60
Fourier U-net	e (%)	**9.40**	**8.21**	**12.44**
	u (%)	**96.78**	**87.48**	**89.66**
	s (%)	**2.76**	**8.21**	**6.08**

Table 2: Ablation study comparing Fourier U-net against a standard U-Net across the static distribution datasets.

The supervised CNN baseline (Lin et al. 2025) is designed to predict the phase graph at the target plane. It then requires an additional inverse Fourier transform to compute the final hologram for the SLM. This inverse Fourier transform step is an intrinsic part of that model's architecture. In contrast, our Fourier U-Net directly predicts the phase graph at the SLM plane, requiring no such extra transformation. However, to ensure a fair comparison under an equivalent computational pipeline, we were compelled to add a round of Fourier transform and inverse Fourier transform to our process only when benchmarking against (Lin et al. 2025).

Ablation Study To rigorously validate that our architectural innovation is the primary source of performance gain, we conducted a targeted ablation study comparing our Fourier U-Net against a **Standard U-Net**. Both models were trained under identical conditions, using the NPCC loss. The results are summarized in Table 2. The relatively low overall efficiency of both models is an expected consequence of the NPCC loss, which directs the network to prioritize pattern uniformity over total power throughput. However, even under this constraint, the Fourier U-net consistently achieves significantly higher efficiency, superior uniformity, and lower standard deviation across all datasets. This demonstrates that the FNO block provides a fundamentally more effective architecture for learning the complex, global phase correlations required for this task.

Results and Discussion Having established the FNO block's foundational advantage, we now compare our pro-

	Unif. (%) (u) ↑				**Eff. (%) (e) ↑**				**PSD (%) (s) ↓**			
Dataset	**Fourier U-Net**	**CNN**	**WGS (2 iter)**	**WGS (30 iter)**	**Fourier U-Net**	**CNN**	**WGS (2 iter)**	**WGS (30 iter)**	**Fourier U-Net**	**CNN**	**WGS (2 iter)**	**WGS (30 iter)**
S-100	96.69 ± 5.73	76.43 ± 3.47	75.85 ± 2.67	94.53 ± 0.84	82.89 ± 0.79	79.89 ± 0.71	81.05 ± 0.55	85.55 ± 0.32	2.54 ± 4.21	18.68 ± 1.65	20.13 ± 1.56	0.43 ± 0.12
S-400	95.87 ± 8.37	71.96 ± 3.16	73.17 ± 2.07	93.82 ± 0.68	82.35 ± 1.25	80.44 ± 0.39	80.65 ± 0.26	85.12 ± 0.16	2.70 ± 5.33	16.24 ± 0.82	19.60 ± 0.72	4.10 ± 0.24
S-900	94.51 ± 10.22	83.18 ± 1.25	69.46 ± 2.14	92.77 ± 0.63	82.68 ± 1.09	84.85 ± 0.08	80.98 ± 0.18	85.64 ± 0.10	3.41 ± 6.31	10.61 ± 0.28	20.39 ± 0.53	4.35 ± 0.17
D-100	86.21 ± 5.18	84.87 ± 6.41	81.72 ± 3.59	96.50 ± 1.36	79.48 ± 0.87	78.83 ± 0.74	79.69 ± 0.86	82.60 ± 1.77	10.67 ± 3.88	11.71 ± 4.95	16.43 ± 2.54	2.75 ± 1.02
D-400	85.17 ± 6.67	82.40 ± 6.10	80.04 ± 4.56	96.17 ± 1.56	79.02 ± 0.48	78.18 ± 1.09	79.37 ± 0.78	82.05 ± 1.80	9.69 ± 4.32	11.27 ± 3.62	15.98 ± 2.19	2.49 ± 0.97
D-900	81.14 ± 7.64	80.56 ± 8.47	77.16 ± 5.17	95.29 ± 1.70	79.13 ± 0.53	79.13 ± 1.68	79.57 ± 0.84	82.60 ± 1.84	11.09 ± 4.56	10.93 ± 3.98	16.61 ± 2.29	2.83 ± 0.96

Table 3: Comprehensive performance comparison across all static and dynamic datasets, using abbreviations for metrics: Uniformity (Unif.), Efficiency (Eff.), and Power Spectral Density (PSD). Our Fourier U-Net is compared against the supervised CNN and WGS algorithm with varying iterations. All values are reported as mean (top) and standard deviation (bottom).

posed Fourier U-Net against the WGS and supervised CNN baselines. The comprehensive results across all static and dynamic datasets are presented in Table 3.

The analysis reveals two conclusions. First, when compared to the SOTA supervised CNN, our Fourier U-Net demonstrates a clear and consistent performance advantage. Across all six test scenarios, our model achieves significantly higher uniformity (u) and a drastically lower percent standard deviation (s), while also delivering superior diffraction efficiency (e). This indicates that our architecture, combined with an unsupervised training strategy, is fundamentally better suited for generating high-quality holograms.

Second, the comparison with the WGS algorithm highlights the crucial advantage of deep learning methods: inference speed. A highly optimized 30-iteration WGS algorithm requires approximately 150 ms per hologram. In contrast, both our Fourier U-Net (7.358 ms) and the CNN (4.781ms) generate holograms in single-digit milliseconds. This speed is not just an improvement but a transformative capability, enabling real-time applications like dynamic atom rearrangement that are impractical with slow iterative methods.

To create a comparison based on a fair computational budget, we should consider our method's total runtime, which is the sum of the network inference time and the back-and-forth Fourier transformation that follows. This is roughly equivalent to the time required for two iterations of the classical WGS algorithm. The data in Table 3 reveals that for a nearly identical time cost, our method significantly outperforms the 2-iteration WGS, particularly in uniformity, where it is often more than 20% points higher in static cases.

The results also suggest that the proposed model can be set as an initializer for classical phase retrieval methods, which have better convergence guarantee. Instead of being initialized with a random phase, the Fourier U-net provides a highly structured starting point that is already close to an optimal solution. This allows a single, fast refinement step to achieve a quality that would otherwise require many more costly iterations. While a heavily iterated WGS can eventually achieve very high efficiency, the computational cost at ~ 150 ms is prohibitive for real-time applications.

Conclusion

Neutral atoms have emerged as a leading quantum platform (Bluvstein et al. 2022; Evered et al. 2023; Bluvstein et al. 2024), leveraging record coherence times (Manetsch et al. 2024) and Rydberg-mediated interactions for high-fidelity computation and large-scale analog simulations like MIS (Ebadi et al. 2022; Pichler et al. 2018). A prerequisite for these applications is the efficient, defect-free assembly of large arrays. Historically, Acousto-Optic Deflectors moved atoms serially (Endres et al. 2016; Barredo et al. 2016, 2018). As recent experiments scale to thousands of atoms (Chiu et al. 2025; Bluvstein et al. 2025a), the linear time complexity of AODs presents a fundamental bottleneck. This has driven the field towards fully parallel rearrangement using SLMs (Lin et al. 2025).

We have addressed the critical operational bottleneck in neutral atom quantum computing: the rapid and high-fidelity assembly of large-scale arrays. This work introduces a comprehensive framework that systematically solves this challenge. Our Min-Max Hungarian (MMH) algorithm provides a rearrangement strategy that is physically optimal, maximizing survival fidelity by minimizing the longest path. Concurrently, our self-supervised Fourier U-Net generates the requisite high-uniformity holograms at real-time speeds. The synergy of these two components provides a complete, fast, and high-fidelity solution to the problem.

One key limitation is that the results of our Fourier U-Net are based on simulation and have not yet been validated in a physical experiment. However, the upper side of this limitation is the successful abstraction of the complex physical task into a computational problem. By framing the challenge this way, our model facilitates more comprehensive and deliberate exploration for the AI community.

Acknowledgments

This work was supported by NSFC 92370201.

References

Barredo, D.; De Léséleuc, S.; Lienhard, V.; Lahaye, T.; and Browaeys, A. 2016. An atom-by-atom assembler of defect-free arbitrary two-dimensional atomic arrays. *Science*, 354(6315): 1021–1023.

Barredo, D.; Lienhard, V.; De Leseleuc, S.; Lahaye, T.; and Browaeys, A. 2018. Synthetic three-dimensional atomic structures assembled atom by atom. *Nature*, 561(7721): 79–82.

Bluvstein, D.; Evered, S. J.; Geim, A. A.; Li, S. H.; Zhou, H.; Manovitz, T.; Ebadi, S.; Cain, M.; Kalinowski, M.; Hangleiter, D.; et al. 2024. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997): 58–65.

Bluvstein, D.; Geim, A. A.; Li, S. H.; Evered, S. J.; Ataides, J.; Baranes, G.; Gu, A.; Manovitz, T.; Xu, M.; Kalinowski, M.; et al. 2025a. Architectural mechanisms of a universal fault-tolerant quantum computer. *arXiv preprint arXiv:2506.20661*.

Bluvstein, D.; Geim, A. A.; Li, S. H.; Evered, S. J.; Bonilla Ataides, J. P.; Baranes, G.; Gu, A.; Manovitz, T.; Xu, M.; Kalinowski, M.; et al. 2025b. A fault-tolerant neutral-atom architecture for universal quantum computation. *Nature*, 1–3.

Bluvstein, D.; Levine, H.; Semeghini, G.; Wang, T. T.; Ebadi, S.; Kalinowski, M.; Keesling, A.; Maskara, N.; Pichler, H.; Greiner, M.; et al. 2022. A quantum processor based on coherent transport of entangled atom arrays. *Nature*, 604(7906): 451–456.

Chiu, N.-C.; Trapp, E. C.; Guo, J.; Abobeih, M. H.; Stewart, L. M.; Hollerith, S.; Stroganov, P.; Kalinowski, M.; Geim, A. A.; Evered, S. J.; et al. 2025. Continuous operation of a coherent 3,000-qubit system. *arXiv preprint arXiv:2506.20660*.

Debnath, L. 2009. The legacy of Leonhard Euler–a tricentennial tribute. *International Journal of Mathematical Education in Science and Technology*, 40(3): 353–388.

Di Leonardo, R.; Ianni, F.; and Ruocco, G. 2007. Computer generation of optimal holograms for optical trap arrays. *Optics express*, 15(4): 1913–1922.

Ebadi, S.; Keesling, A.; Cain, M.; Wang, T. T.; Levine, H.; Bluvstein, D.; Semeghini, G.; Omran, A.; Liu, J.-G.; Samajdar, R.; et al. 2022. Quantum optimization of maximum independent set using Rydberg atom arrays. *Science*, 376(6598): 1209–1215.

Endres, M.; Bernien, H.; Keesling, A.; Levine, H.; Anschuetz, E. R.; Krajenbrink, A.; Senko, C.; Vuletic, V.; Greiner, M.; and Lukin, M. D. 2016. Atom-by-atom assembly of defect-free one-dimensional cold atom arrays. *Science*, 354(6315): 1024–1027.

Evered, S. J.; Bluvstein, D.; Kalinowski, M.; Ebadi, S.; Manovitz, T.; Zhou, H.; Li, S. H.; Geim, A. A.; Wang, T. T.; Maskara, N.; et al. 2023. High-fidelity parallel entangling gates on a neutral-atom quantum computer. *Nature*, 622(7982): 268–272.

Gerchberg, R. W. 1972. A practical algorithm for the determination of plane from image and diffraction pictures. *Optik*, 35(2): 237–246.

Kim, D.; Keesling, A.; Omran, A.; Levine, H.; Bernien, H.; Greiner, M.; Lukin, M. D.; and Englund, D. R. 2019a. Large-scale uniform optical focus array generation with a phase spatial light modulator. *Optics letters*, 44(12): 3178–3181.

Kim, H.; Kim, M.; Lee, W.; and Ahn, J. 2019b. Gerchberg-Saxton algorithm for fast and efficient atom rearrangement in optical tweezer traps. *Optics express*, 27(3): 2184–2196.

Kuhn, H. W. 1955. The Hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2): 83–97.

Lee, W.; Kim, H.; and Ahn, J. 2017. Defect-free atomic array formation using the Hungarian matching algorithm. *Physical Review A*, 95(5): 053424.

Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2020. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*.

Lin, R.; Zhong, H.-S.; Li, Y.; Zhao, Z.-R.; Zheng, L.-T.; Hu, T.-R.; Wu, H.-M.; Wu, Z.; Ma, W.-J.; Gao, Y.; et al. 2025. AI-enabled parallel assembly of thousands of defect-free neutral atom arrays. *Physical Review Letters*, 135(6): 060602.

Manetsch, H. J.; Nomura, G.; Bataille, E.; Leung, K. H.; Lv, X.; and Endres, M. 2024. A tweezer array with 6100 highly coherent atomic qubits. *arXiv preprint arXiv:2403.12021*.

Pichler, H.; Wang, S.-T.; Zhou, L.; Choi, S.; and Lukin, M. D. 2018. Quantum optimization for maximum independent set using Rydberg atom arrays. *arXiv preprint arXiv:1808.10816*.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.