I. J. Computer Network and Information Security, **2026, 2, 38-60**
Published Online on April 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2026.02.03

Interference-aware Deep Quantum Neural Network for NOMA Channel Estimation via Adaptive Energy Valley Optimization

Avinash Ratre*
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, 110042, India
E-mail: avinashratre@dtu.ac.in
ORCID iD: https://orcid.org/0000-0001-9803-2665
*Corresponding author

Received: 26 October 2025; Revised: 10 December 2025; Accepted: 30 January 2026; Published: 08 April 2026

Abstract: To address the growing multi-user interference in dense wireless networks, we propose an interference-aware Deep Quantum Neural Network (DQNN) for channel estimation in the Non-orthogonal multiple access (NOMA) systems. The proposed method incorporates a hybrid classical-quantum architecture. A Transformer-encoder processes the pilot signals to extract spatiotemporal features. A parameterized quantum circuit maps the processed features into a high-dimensional Hilbert space. The enhancement hinges on an Adaptive Energy Valley Optimization (AEVO) algorithm, which modifies the optimization trajectory using interference-aware preconditioners derived from the interference covariance structure. With the aid of these preconditioners, the DQNN can steer through the NOMA's non-convex terrain characterized by interference to enhance estimation performance. Moreover, interference-aware preconditioning is achieved through a lightweight neural network which adapts to time-varying interference. The successive interference cancellation decoder uses the estimated channel matrix to recover symbols. By further analysing the results, it is noticed that the quantum-enhanced machine learning delivers better results than the classical ones. The proposed framework enhances the state-of-the-art in NOMA channel estimation, while also providing a general framework for interference-aware optimization in quantum machine learning. At 10 dB SNR, the AEVO-DQNN method with a 16x16 antenna array obtained a minimum NMSE of 0.012288 and a minimum BER of 0.013023. Further, the proposed method outperforms the competing methods in terms of NMSE/BER mean with 95% confidence intervals, interference rejection ratio analysis, sensitivity to estimation error and estimated interference covariance, and paired t-test analysis.

Index Terms: Channel Estimation, Non-Orthogonal Multiples Access, Deep Quantum Neural Network, Adaptive Energy Valley Optimization, Interference-Aware Preconditioner.

1. Introduction

NOMA has become a key technology for 5G and beyond due to its superior spectral efficiency and the ability to serve multiple users in the same resource block using power-domain multiplexing [1]. NOMA differs from conventional orthogonal multiple access (OMA) techniques by exploiting differences in channel gains between users. This enables simultaneous transmission and increases system capacity [2]. Despite this benefit, the increase in multi-user interference makes channel estimation for coherent detection and resource allocation challenging [3]. There exists a classical approach, such as least-squares (LS) and minimum mean-square error (MMSE), which is not adequate for the NOMA environment because it is non-convex and interference-contaminated [3]. Even though deep learning-based approaches are capable of alleviating such challenges [4], they usually do not exploit the quantum-inspired computational advantages further to enhance performance. Quantum neural networks (QNNs) can provide a solution by simplifying the processing of high-dimensional data and non-linear relationships [5].

At the transmitting end, superposition coding is used for NOMA and at the receiving end, successive interference cancellation (SIC) is used. Nevertheless, the implementation of NOMA in practice encounters obstacles, such as the necessity of accurate channel state information (CSI) or channel estimation (CE) for effective power allocation and successive interference cancellation (SIC), high receiver complexity, optimal resource allocation, and interference management [6]. Deep Learning (DL) offers promising approaches to address these problems. RNNs and LSTMs are better at channel estimation in a dynamic fading environment and improve the BER, Outage probability, and sum rate significantly compared to a conventional method, such as MMSE. Using power allocation strategies with DL-based CE enhances the performance. Deep learning techniques are also applied to resource allocation, user pairing and fairness, approximation of SIC receivers, reduction of impulse noise, and transceiver design [7].

Iterative receivers have been developed for FTN-NOMA systems with grant-free random access, which can jointly track user activity, estimate the channel, and detect data. The complexity is linear in the number of active users [8]. One technique to manage complexity in message passing for grant-free scenarios is the use of a rotationally invariant Gaussian mixture (RIGM) model. Although some progress has been made on a few of these challenges, there are still several challenges that remain, which are efficient pairing of multiple users, user mobility, multi-cell interference, imperfect SIC, security, and low-complexity algorithms for uplink NOMA [9]. Accurate channel estimation in the high-interference, non-orthogonal environment is a critical bottleneck. The performance of NOMA systems relies on Successive Interference Cancellation (SIC) at the receiver, a technique whose efficacy is fundamentally dependent on precise channel state information (CSI). Channel estimation errors are not isolated; they propagate and amplify through the SIC chain, resulting in catastrophic degradation of system performance. This creates a "non-convex and interference-contaminated" optimization landscape that conventional estimators cannot solve optimally.

Recent advances in optimization algorithms have enabled the proficient training of complex neural networks through the Energy Valley Optimizer (EVO) [10]. These techniques, however, do not have the flexibility to adjust to the NOMA interference. To address this shortfall, we propose a Deep Quantum Neural Network (DQNN) that is interference aware and is trained using the Adaptive Energy Valley Optimization (AEVO) algorithm. The important thing is that the integration of learned interference-aware preconditioners adapts the optimization trajectory to account for multi-user interference explicitly. These preconditioners are based on the interference covariance structure, allowing the DQNN to escape the non-convex optimization trap more easily [11].

The proposed AEVO-DQNN model estimates NOMA channels. At the transmitter, data is demultiplexed, converted from serial-to-parallel, and mapped to 16-QAM. LS estimation at the receiver delivers the initial estimate of the vector. Channel estimation is done by a DQNN (Transformer encoder + PQC). This deep quantum neural network (DQNN) is trained with a preconditioner that is based on the AEVO (derived from the EVO). Data is recovered using SIC decoding and 16-QAM demapping. This work advances the state-of-the-art in NOMA channel estimation by combining quantum computing with interference-aware optimization and opens up new opportunities in quantum-enhanced wireless systems. The proposed framework is generalizable and can be applied to other interference-dominated scenarios, such as massive MIMO and ultra-dense networks. The proposed framework, AEVO-DQNN, is a hybrid classical-quantum architecture designed to overcome the limitations of existing methods. The core contribution is a tripartite solution:

- A Deep Quantum Neural Network (DQNN): A hybrid model combining a classical Transformer-encoder (for feature extraction) with a Parameterized Quantum Circuit (PQC). The PQC maps features into a high-dimensional Hilbert space to efficiently model the complex, non-linear channel dynamics.
- Adaptive Energy Valley Optimization (AEVO): A novel optimization algorithm used to train the DQNN. It is an advanced, population-based method derived from Energy Valley Optimization (EVO) that is inherently suited for navigating non-convex landscapes.
- Learned Interference-Aware Preconditioners: This is the central innovation. The AEVO algorithm is enhanced with adaptive preconditioners derived from the interference covariance structure. These preconditioners, controlled by a lightweight neural network, "steer" the optimization process, allowing the DQNN to avoid interference-induced local minima and find a more accurate channel estimate.

The significance of the proposed work lies not merely in the application of quantum computing to wireless communications, but in the development of a problem-specific, interference-aware optimization framework. The AEVO-DQNN architecture bridges the gap between the high-dimensional representational power of quantum circuits and the specific, non-convex optimization challenges of a classical communications problem. The proposed work provides a comprehensive technical analysis of this framework, deconstructing its architecture, its novel optimization mechanism, and its benchmarked performance. The remainder of this paper has been organized as follows: Section 2 provides a review of the existing work on NOMA channel estimation as well as quantum machine learning. Section 3 contains introductory information on NOMA channel estimation and the AEVO optimization framework. The proposed interference-aware DQNN's architecture and training are discussed in Section 4. In Section 5, we present experimental comparisons with state-of-the-art methods. Section 6 presents the implications and future research directions. Section 7 ends the paper.

2. Related Works

NOMA achieves its high spectral efficiency by deliberately sacrificing orthogonality. Multiplexing users in the power domain allows multiple users to share the same resource block, a fundamental departure from OMA techniques. This benefit, however, is also NOMA's greatest challenge: it introduces severe, structured multi-user interference (MUI) as an intrinsic part of the signal model. To manage this MUI, NOMA systems mandate the use of Successive Interference Cancellation (SIC) at the receiver. In an SIC-based receiver, users are decoded sequentially. The signal of the strongest user is decoded first, its contribution is estimated and subtracted from the composite signal, and the process repeats for the next-strongest user.

This sequential decoding process creates a critical dependency: the successful decoding of one user is contingent upon the successful decoding and cancellation of all stronger users before it. This creates a mechanism for error propagation. As the research indicates, imperfect channel estimation at any stage of the SIC chain results in imperfect interference cancellation. This residual interference is then passed on to the next decoding stage, increasing the noise floor and compromising the entire detection process. Therefore, NOMA's reliance on SIC makes it uniquely and critically sensitive to the accuracy of channel state information (CSI). NOMA systems have been the subject of extensive studies on channel estimation using classical and machine learning approaches. The existing methods consist of 3 classes: Conventional signal processing techniques, classical machine learning methods, and emerging quantum-enhanced methods.

2.1. Conventional Channel Estimation Methods

Traditional NOMA channel estimation approaches are mainly pilot-assisted. The LS and MMSE estimators are commonly used since they are easy to compute [3]. The aforementioned approaches suffer performance degradation in high-interference scenarios due to their disregard for NOMA's non-orthogonality. Sophisticated techniques based on compressed sensing [12] have been proposed to leverage channel sparsity, yet they suffer from dynamic interference patterns. The work in [11] develops an iterative interference-cancellation scheme that inspired our preconditioning but is limited to linear processing.

2.2. Machine Learning-Based Approaches

Recent advancements in deep learning have revolutionized channel estimation across many wireless communication scenarios. Convolutional neural networks (CNNs) have been proposed to learn the spatial features from the pilot signal in NOMA systems [4]. Recurrent architectures, particularly long short-term memory (LSTM) networks, have been used to model temporal correlations in time-varying channels [3]. The approach presented in [13] used ensemble learning of the combination of multiple weak estimators for accuracy improvement. Yet, due to the limitations in dimensionality of classical neural networks, they are not suitable for modelling the interference patterns of NOMA.

2.3. Quantum-enhanced Solutions

New possibilities for channel estimation have begun with quantum computing and wireless communication. Quantum neural networks show promise in solving high-dimensional optimization problems, which classical computers cannot solve [14]. The parameterized quantum circuit architecture similar to ours was first proposed in [15] without the NOMA interference being considered. The AEVO framework is adopted since recent work on quantum-inspired optimization algorithms [16] demonstrated better convergence properties compared to classical gradient-based optimization algorithms. Several studies have looked into interference management in NOMA systems. The solution to the problem presented in work [17] has proposed a convex relaxation approach for interference mitigation. Moreover, work [18] has proposed an age-of-information aware algorithm for multicast scenarios. Yet, such approaches either rest on simple interference models or fail to exploit quantum systems' representational power. The learning-based interference adaptation was introduced in adaptive semantic generation [19]. However, the classical neural network architecture proposed there does not exploit any quantum advantage, unlike the one proposed here.

The authors propose a practical channel estimation method based on a Long Short-Term Memory (LSTM) neural network that adapts to varying channel conditions. The paper presents a novel power coefficient allocation algorithm based on a binomial distribution and Pascal's triangle approach, allocating power between two users solely based on channel conditions. The proposed algorithm has a gap compared to multi-user power allocation in NOMA systems. This paper mainly investigates downlink NOMA, but uplink NOMA needs to be studied [6]. The paper proposes an algorithm for grant-free NOMA, comprising a slot-wise multi-user detector and a joint signal and channel estimator. The study under consideration envisions a single-antenna configuration at the receiver, which limits the ability to support more users and/or higher transmission rates. The paper omits multi-antenna channels, which can offer more sophisticated designs, e.g. structural information such as angular-domain sparsity [9].

The authors have applied a new framework that integrates temporal correlation in the active user set with multi-antenna reception. The study developed a Kronecker Compressed Sensing (KCS) framework to extract the spatial-temporal structure of user activity fully. Classical greedy algorithms for the KCS problem do not jointly exploit the underlying spatial-temporal correlation of user activity in MIMO-NOMA systems, as shown in this paper. Current methods do not address the potential for more efficient multiuser detection [20]. A generalized GMM-based joint channel estimation and signal detection algorithm for NOMA is proposed for cases where the receiver is unaware of the transmitting users or the channel conditions. This study uses a theoretical analysis to show that the SER converges and to derive a closed-form expression for its SER. In a grant-free NOMA scenario relevant to massive IoT systems, the authors study the tradeoff between pilot length and SER. The channel phase cannot be inferred directly based on the cluster centroids in this paper and pilot symbols are needed for demapping. The initialization of the EM algorithm in the vicinity of the actual centroids is assumed in the paper, without which it may not be possible [21].

The proposed interference-aware DQNN with AEVO optimization is fundamentally different in several aspects. While standard methods treat interference as noise, we model the interference pattern directly using learned preconditioners. In contrast to classical machine learning solutions, the quantum component can more efficiently represent complex channel dynamics in high-dimensional Hilbert space. Our optimization algorithm adapts to interference variations, unlike a fixed design of a quantum circuit. Previous quantum-enhanced solutions lack this capability. This absolute combination of interference awareness, quantum computing, and adaptive optimization is a remarkable performance compared to the above existing techniques.

3. Preliminaries on NOMA channel Estimation and AEVO Optimization

To build the theoretical basis for the designed method, first, the NOMA channel estimation problem is formalized. Then, the main implementation ideas of the Adaptive Energy Valley Optimization (AEVO) framework (derived from Energy Valley Optimization (EVO) [22]) are introduced. The section starts with the system model and signal formulation, then examines the interference characteristics in a NOMA system. Next, we formulate the mathematical equations associated with the problem of channel estimation as an optimization problem and show how AEVO provides a solution.

3.1. System Model

We consider hexagonal L cells, where each cell has one base station with M antennas, and K user terminals with a single antenna. $\boldsymbol{h}_{ljk} \in \mathbb{C}^{M}$ represents the channel from the k^{th} user terminal of the j^{th} cell to the base station in the l^{th} cell [23]. Usually, the channel model employs uncorrelated Rayleigh fading. Moreover, the measurement campaigns suggest that practical massive MIMO channels have spatial correlation. We consider the correlated Rayleigh fading model to represent spatially correlated channels, as shown below.

$$\boldsymbol{h}_{ljk} \sim NC\left(0, \boldsymbol{R}_{ljk}\right), \tag{1}$$

where $\boldsymbol{R}_{ljk}$ refers to the channel covariance matrix that describes the macroscopic propagation characteristics. We define the distribution $NC\left(0, \boldsymbol{R}\right)$ to be the circularly symmetric complex Gaussian distribution with zero mean and covariance matrix $\boldsymbol{R}$. When the spatial correlation is independent, we can express it as $\boldsymbol{R}_{ljk} = \beta_{ljk} \boldsymbol{I}_M$, which is a diagonal matrix. β_{ljk} denotes the large-scale fading coefficient that connects the base station in l^{th} cell with the k^{th} user terminal in the j^{th} cell. Despite this, we take into account the spatial correlated channel having a covariance matrix with non-zero off-diagonal elements and dissimilar diagonal elements. The channel's covariance matrix is modeled as

$$\boldsymbol{R}_{ljk} = \beta_{ljk} \int_{-\pi}^{\pi} p\left(\theta_{ljk}\right) a\left(\theta_{ljk}\right) a\left(\theta_{ljk}\right)^{H} d\theta_{ljk}, \tag{2}$$

where the power spectrum of an angle of arrival (AoA) is $p\left(\theta\right)$, while the steering vector of a uniform linear array is $\boldsymbol{a}\left(\theta\right)$. The range from θ , i.e., $\{\theta : p\left(\theta\right) < \grave{o}\}$ is excluded, where $\grave{o}$ is some small value. It is assumed that the power spectrum $p\left(\theta\right)$ has a Laplace distribution with standard deviation 2°. This helps to spread the received power around the center of the propagation paths [24]. The NOMA system model is shown in Fig. 1.

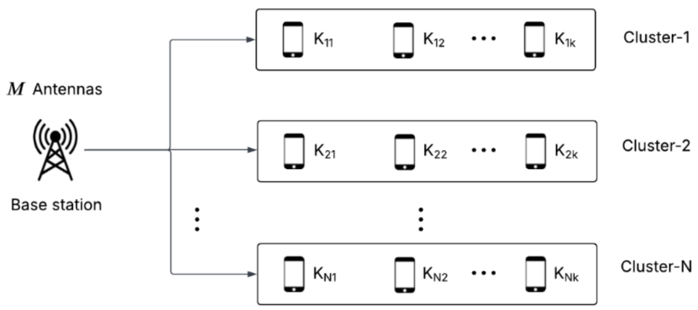


Fig.1. NOMA system model

3.2. Signal Formulation

Consider a downlink NOMA network in which K users share the same time-frequency resource block. The received signal at user k can be expressed as:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k \tag{3}$$

where $\mathbf{h}_k \in \mathrm{C}^{M\times 1}$ is the channel vector that exists between the base station (with M antennas) and user k. It models the effects of fading and path loss. Hermitian transpose of the channel vector $\mathbf{h}_k$ is represented by $\mathbf{h}_k^H$. The transmitted signal vector is represented by $\mathbf{x}$ and n_k represents additive white Gaussian noise. Superposition coding combines users' symbols to form the transmitted signal:

$$\mathbf{x} = \sum_{i=1}^{K} \sqrt{p_i}\,\mathbf{w}_i s_i \tag{4}$$

In this expression, we use p_i, $\mathbf{w}_i$, and s_i respectively to refer to the power allocation, beamforming vector, and information symbol for user i. The channel estimation problem tackles the recovery of $\mathbf{h}_k$ from corrupted pilot observations.

3.3. Interference Characteristics in NOMA

The interference structure in NOMA systems exhibits two distinctive properties that complicate channel estimation. The interference covariance matrix $\mathbf{R}_I$ shows non-diagonal dominance due to correlated channel gains among users:

$$\mathbf{R}_I = \sum_{i\neq k} p_i \mathbf{h}_i \mathbf{h}_i^H \tag{5}$$

Where $p_i \mathbf{h}_i \mathbf{h}_i^H$ represents the covariance of the interference signal from a single user i. The outer product $\mathbf{h}_i \mathbf{h}_i^H$ captures the spatial characteristics of the interference from user i. The interference-to-noise ratio (INR) varies significantly across users depending on their power allocation and channel conditions. These characteristics render conventional estimation techniques suboptimal, as they either ignore the interference structure or make overly simplistic assumptions about its statistics [25].

3.4. Channel Estimation as Optimization

The maximum likelihood estimation of $\mathbf{h}_k$ can be formulated as a non-convex optimization problem:

$$\min_{\mathbf{h}_k} \left\| \mathbf{y}_k - \mathbf{X}\mathbf{h}_k \right\|^2 + \lambda \mathbf{h}_k^H \mathbf{R}_I^{-1} \mathbf{h}_k \tag{6}$$

where $\mathbf{X}$ contains pilot symbols and λ is a regularization parameter controlling the trade-off between data fidelity and interference resilience. The second term is the regularization term which accounts for interference through the inverse covariance matrix, but direct computation becomes infeasible in practical systems due to unknown interference parameters. $\mathbf{h}_k^H \mathbf{R}_I^{-1} \mathbf{h}_k$ incorporates prior knowledge about the interference statistics (via $\mathbf{R}_I^{-1}$ as a precision matrix) to penalize channel estimates that are inconsistent with the known interference structure. If $\lambda = 0$, the expression reduces to a simple least-squares problem. If λ is very large, the solution will be heavily biased to be orthogonal to the interference subspace, even if it does not fit the data well [3].

3.5. AEVO Optimization Framework

The AEVO algorithm (derived from EVO [22]) addresses this challenge through three key mechanisms [10]. First, it maintains a population of candidate solutions $\{\mathbf{h}_k^{(i)}\}_{i=1}^{N}$ that explore different regions of the solution space. Second, it employs an energy-based selection criterion that favors solutions with both low estimation error and interference resilience. The energy function $E(\mathbf{h}_k)$ serves as the fitness or cost function in the AEVO algorithm. $E(\mathbf{h}_k)$ is the energy associated with a channel estimate $\mathbf{h}_k$. The energy function is designed to evaluate the quality of a candidate's channel estimate $\mathbf{h}_k$:

$$E(\mathbf{h}_k) = \left\| \mathbf{y}_k - \mathbf{X}\mathbf{h}_k \right\|^2 + \gamma \operatorname{tr}\left(\mathbf{h}_k \mathbf{h}_k^H \mathbf{R}_I\right) \tag{7}$$

where γ balances the trade-off between data fidelity and interference suppression. The second term, $\gamma \operatorname{tr}\left(\mathbf{h}_k \mathbf{h}_k^H \mathbf{R}_I\right)$, acts as a regularization penalty that accounts for multi-user interference. The expression $\mathbf{h}_k \mathbf{h}_k^H \mathbf{R}_I$ is a scalar that represents the total interference power from other users as projected onto the direction of the estimated channel $\mathbf{h}_k$. Minimizing this term forces the channel estimate into a subspace that is orthogonal to the directions of strong interference, promoting interference suppression. Third, the algorithm adapts its search direction using learned preconditioners that approximate the inverse interference covariance structure [26]. The preconditioner matrix $\mathbf{P}$ transforms the original optimization problem into a better-conditioned space. The transformed optimization problem is:

$$\min_{\tilde{\mathbf{h}}_k} \left\| \mathbf{y}_k - \mathbf{X}\mathbf{P}^{-1/2}\tilde{\mathbf{h}}_k \right\|^2 \tag{8}$$

where the linear transformation defines a new variable $\tilde{\mathbf{h}}_k = \mathbf{P}^{1/2}\mathbf{h}_k$ and $\mathbf{P}^{1/2}$ is the principal square root of $\mathbf{P}$. The preconditioned optimization problem now solves for the transformed channel vector $\tilde{\mathbf{h}}_k$ in a better-conditioned space. The actual channel estimate $\mathbf{h}_k$ can be recovered afterward using the inverse transformation, $\mathbf{h}_k = \mathbf{P}^{-1/2}\tilde{\mathbf{h}}_k$. This transformation enables more efficient optimization while implicitly accounting for interference correlations. The AEVO framework differs from conventional evolutionary algorithms by incorporating problem-specific knowledge through these adaptive preconditioners, making it particularly suitable for NOMA channel estimation [27].

4. Interference-aware Deep Quantum Neural Network with Learned Preconditioners

The proposed interference-aware DQNN architecture [28] combines classical signal processing with quantum-enhanced learning to address the unique challenges of NOMA channel estimation [5]. The system processes received pilot signals through a hybrid pipeline that extracts interference patterns using classical neural networks and maps them to quantum states for enhanced representation learning. The core innovation lies in the adaptive optimization framework that dynamically adjusts to interference conditions through learned preconditioners.

4.1. Interference-aware Preconditioning in AEVO for NOMA

The preconditioner matrix $\mathbf{P}_t$ at optimization step t incorporates interference statistics through a low-rank update:

$$\mathbf{P}_t = \mathbf{I} + \lambda_t \mathbf{M}_t \tag{9}$$

where the identity matrix, $\mathbf{I}$, represents a neutral or "un-preconditioned" state, $\mathbf{M}_t = \mathrm{E}\left[\mathbf{J}_{\mathrm{Int}}^H \mathbf{J}_{\mathrm{Int}}\right]$ captures the interference covariance at step t via the Jacobian of the interference term. In other words, $\mathbf{M}_t$ measures how sensitive the interference is to changes in the channel. The term $\lambda_t \mathbf{M}_t$ injects knowledge about interference. The scaling factor λ_t adapts to current interference conditions through a gating mechanism:

$$\lambda_t = \sigma\left(\mathbf{w}_\lambda^T \mathbf{z}_t + b_\lambda\right) \tag{10}$$

Here $\mathbf{z}_t$ contains instantaneous interference metrics including the interference-to-signal ratio (ISR) and channel correlation coefficients. $\mathbf{w}_\lambda$ and b_λ are learnable weight vector and bias term of the gate. $\sigma(\cdot)$ is the sigmoid function, which squashes its input to a range of 0 to 1. This acts as a "gate": if interference is low, λ_t will be close to 0; and the preconditioner $\mathbf{P}_t$ will be close to the identity matrix (i.e., minimal preconditioning), if interference is high, λ_t approaches 1, applying the interference-aware correction more strongly. The logic of scaling factor λ_t is to tilt the optimization landscape away from the identity in the directions where interference is most substantial. This helps the optimizer avoid getting stuck in bad regions and move more directly towards a good solution. The interference matrix $\mathbf{M}_t$ evolves through a momentum-based update and describes an exponential moving average, allowing the optimizer to remember past interference patterns:

$$\mathbf{M}_t = \beta \mathbf{M}_{t-1} + (1-\beta)\mathbf{J}_{\mathrm{Int}}^H \mathbf{J}_{\mathrm{Int}} \tag{11}$$

This formulation, instead of calculating the interference covariance from scratch at each step, enables the optimizer to account for both current and historical interference patterns, crucial for NOMA systems with temporally correlated channels. $\mathbf{M}_{t-1}$ is the historical interference covariance from the previous step. $\mathbf{J}_{\mathrm{Int}}^H \mathbf{J}_{\mathrm{Int}}$ is the instantaneous interference covariance calculated at the current step. β is a momentum coefficient ($0 \le \beta < 1$) that balances $\mathbf{M}_{t-1}$ and $\mathbf{J}_{\mathrm{Int}}^H \mathbf{J}_{\mathrm{Int}}$. A high β means the system has more "memory" of past interference patterns, which is crucial for channels where interference is correlated over time.

4.2. Hybrid Classical-quantum Architecture with Interference-centric Feature Extraction

This section details how a classical neural network processes the raw signals to extract key features, which are then fed into quantum circuit for advanced computation. The classical processing component employs a Transformer encoder with L layers to extract interference-aware features from pilot signals $\mathbf{Y}$:

$$\mathbf{F}_l = \text{LayerNorm}\left(\mathbf{F}_{l-1} + \text{MultiHead}\left(\mathbf{F}_{l-1}, \mathbf{F}_{l-1}, \mathbf{F}_{l-1}\right)\right) \tag{12}$$

where $\mathbf{F}_l$ denotes features at layer l. $\mathbf{F}_{l-1}$ is the feature matrix from the previous layer. $\text{MultiHead}(\cdot,\cdot,\cdot)$ is the multi-head self-attention mechanism, which explicitly models interference relationships between users through its attention weights. The "+" sign indicates a residual connection, which helps prevent gradients from vanishing during training. $\text{LayerNorm}(\cdot)$ is layer normalization, a technique to stabilize the learning process. It is repeated L times to build a deep understanding of the interference structure. The final classical features $\mathbf{F}_L$ are projected to the quantum circuit input dimension via a learned linear transformation $\mathbf{W}_q$:

$$\mathbf{x}_q = \mathbf{W}_q \mathbf{F}_L \tag{13}$$

The above equation uses a learnable weight matrix $\mathbf{W}_q$ to reshape or project the classical features into the appropriate format $\mathbf{x}_q$ to be used as parameters in the quantum circuit. The quantum circuit implements a parameterized unitary transformation $\mathbf{U}(\boldsymbol{\theta})$ composed of N qubits and D layers:

$$\mathbf{U}(\boldsymbol{\theta}) = \prod_{d=1}^{D}\left(\otimes_{n=1}^{N} \mathbf{R}_z\left(\theta_{n,d}^1\right)\mathbf{R}_x\left(\theta_{n,d}^2\right)\right)\mathbf{U}_{\text{ent}} \tag{14}$$

where $\mathbf{R}_z\left(\theta_{n,d}^1\right)$ and $\mathbf{R}_x\left(\theta_{n,d}^2\right)$ represent single-qubit rotation gates. Their angles are the parameters that will be learned. $\mathbf{U}_{\text{ent}}$ creates entanglement layer between qubits. Entanglement allows qubits to become correlated in ways classical bits cannot, providing the potential for quantum advantage. The entire operation is a sequence of learnable rotations followed by entanglement, repeated D times. $\prod_{d=1}^{D}$ signifies that the circuit is built from D repeating layers. $\otimes_{n=1}^{N}$ is the tensor product, meaning the operation inside is applied to each of the N qubits independently. The circuit outputs quantum expectation values that are post-processed to estimate channel coefficients.

4.3. Interference-covariance-driven Quantum Optimization

The preconditioned gradient update can perform directional reweighting according to the learned interference structure, which can enhance the convergence speed for heavy multi-user interference. We optimize the parameters of a quantum circuit, $\boldsymbol{\theta}$ using the interference-aware AEVO update rule:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta_t \mathbf{P}_t^{-1} \nabla_{\theta} \mathrm{L}\left(\boldsymbol{\theta}_t\right) \tag{15}$$

The given equation entails a preconditioned gradient descent algorithm. The main innovation here is inserting $\mathbf{P}_t^{-1}$, which is the interference-aware preconditioner's inverse. When the gradient is multiplied by $\mathbf{P}_t^{-1}$, it rescales and rotates the direction in which the update will be made. This tells the optimizer to take bigger steps in directions of low interference and smaller ones in high interference. In other words, it leads to faster and stable convergence. The inclusion of estimation error, as well as interference suppression, is incorporated in the loss function L:

$$L = \frac{1}{K}\sum_{k=1}^{K}\left\|\hat{\mathbf{h}}_k - \mathbf{h}_k\right\|^2 + \alpha \operatorname{tr}\left(\hat{\mathbf{H}}^H \mathbf{R}_I \hat{\mathbf{H}}\right) \tag{16}$$

This loss function trains the model to simultaneously perform two tasks. The first term, MSE measures the average square difference between the estimated channel $\hat{\mathbf{h}}_k$ and the true channel $\mathbf{h}_k$. Minimizing this makes the estimates accurate. The second term is the interference suppression penalty. It penalizes solutions that align the estimated channel matrix $\hat{\mathbf{H}}$ with the interference $\mathbf{R}_I$. The hyperparameter, α balances the importance of accuracy and interference rejection.

4.4. NOMA SIC Decoding

This explains how a practical NOMA decoder uses the estimated channel. The estimated channel matrix $\hat{\mathbf{H}}$ goes straight into the decoder for successive interference cancellation (SIC) [11]. The estimated symbol for user k is decoded as:

$$\hat{s}_k = Q\left(\frac{y_k - \sum_{j=1}^{k-1} \hat{h}_{kj}\hat{s}_j}{\hat{h}_{kk}}\right) \tag{17}$$

where $Q(\cdot)$ denote the quantization operation. The SIC decoder's gradient flow allows for channel estimation and interference cancellation to be trained end-to-end. The decoder performs the following operations: (i) Observation of the received signal y_k, (ii) Reconstruct and subtract interference: The signals of all the users j $(j < k)$ who are already decoded, are reconstructed using the estimated channel $\hat{h}_{kj}$ and symbol $\hat{s}_j$. This is subtracted from y_k. (iii) Normalization: The result is divided by the channel gain of the present user, $\hat{h}_{kk}$, and (iv) Quantization: The outcome is passed on to a quantization function $Q(\cdot)$ to map it to the closest valid symbol giving us the estimate $\hat{s}_k$.

4.5. Adaptive Preconditioning

A gated recurrent unit (GRU) network, a type of recurrent neural network, controls the preconditioner adaptation [27] by processing a window of recent interference metrics:

$$\mathbf{r}_t = \sigma\left(\mathbf{W}_r\left[\mathbf{z}_t, \mathbf{h}_{t-1}\right]\right) \tag{18}$$

$$\mathbf{h}_t = \left(1 - \mathbf{r}_t\right) \circ \mathbf{h}_{t-1} + \mathbf{r}_t \circ \tanh\left(\mathbf{W}_h\left[\mathbf{z}_t, \mathbf{h}_{t-1}\right]\right) \tag{19}$$

The output hidden state $\mathbf{h}_t$ generates the parameters for λ_t and $\mathbf{M}_t$, allowing the system to maintain memory of past interference conditions while adapting to new patterns. $\mathbf{z}_t$ is the current input (the vector of interference metrics). $\mathbf{r}_t$ is the reset gate, which decides how much of the past memory $(\mathbf{h}_{t-1})$ should be forgotten. This equation computes the new hidden states $\mathbf{h}_t$ by intelligently combining the new candidate state with the old hidden state, using the reset gate as a controller. The symbol $\circ$ denotes the element-wise product. The quantum circuit measurements are processed through a classical neural network to produce final channel estimates, completing the hybrid classical-quantum pipeline. The entire system is trained end-to-end using the interference-aware AEVO optimizer, with gradients flowing through both classical and quantum components via parameter-shift rules for the quantum circuit.

This GRU processes a window of recent interference metrics to generate the parameters for the preconditioner matrix and the scaling factor. While this is a small recurrent network, it still requires a new forward pass involving several matrix-vector multiplications at each adaptation step to update its hidden state and produce the new gating parameters. The main overhead is calculating and applying the preconditioner. The preconditioner itself is a low-rank update to the identity matrix. The real computational work here is in determining, which we defined as the Jacobian of the interference term. During the AEVO update rule, we apply the inverse of this preconditioner to the gradient. While low-rank structures make this inversion more efficient than a full matrix inversion, it is still a distinct and additional computational step compared to a standard gradient descent update.

The Interference-Aware Deep Quantum Neural Network (DQNN) training via Adaptive Energy Valley Optimization (AEVO) is represented in Algorithm 1.

Algorithm 1: Adaptive Energy Valley Optimization (AEVO) for DQNN

1. Inputs:
- Received pilot signals: Y
- Pilot symbols: X
- Initial quantum circuit parameters: θ_0
- Initial interference covariance matrix: M_0
- Interference metrics input vector: z_t

2. Hyperparameters:
- **Population Size:** N (Number of candidate solutions $\{h_k^{(i)}\}_{i=1}^{N}$)

- **Momentum Coefficient:** β (where $0 \le \beta < 1$)
- **Gate Parameters:** w_λ (weight vector) and b_λ (bias term)
- **Learning Rate:** η_t
- **Regularization parameters:** γ , α

3. Output:

- Optimized quantum circuit parameters: θ_{opt}
- Estimated channel vector: $\hat{h}_k$

4. Iterative Update Rules:

Initialize: Set $t = 0$, initialize θ_0 and M_0 .

While not converged **do**:

- Compute Instantaneous Interference:
  Calculate the instantaneous interference covariance using the Jacobian of the interference term:
  $J_{Int}^H J_{Int}$
- Update Interference Matrix (M_t):
  Apply the momentum-based update to evolve the interference matrix using (11):
  $M_t = \beta M_{t-1} + (1-\beta) J_{Int}^H J_{Int}$
- Compute Adaptive Scaling Factor (λ_t):
  Calculate the gating factor based on current interference metrics z_t using (10):
  $\lambda_t = \sigma(w_\lambda^T z_t + b_\lambda)$
- Construct Preconditioner (P_t):
  Compute the interference-aware preconditioner matrix using (9):
  $P_t = I + \lambda_t M_t$
- Evaluate Loss Function (L):
  Compute the loss combining MSE and interference suppression penalty using (16):
  $\mathrm{L} = \frac{1}{K} \sum_{k=1}^{K} \left\| \hat{h}_k - h_k \right\|^2 + \alpha tr(H^H R_I H)$
- Update Quantum Parameters (θ_{t+1}):
  Update the circuit parameters using the preconditioned gradient descent rule defined in (15):
  $\theta_{t+1} = \theta_t - \eta_t P_t^{-1} \nabla_\theta \mathrm{L}(\theta_t)$
- Increment Step:
  $t \leftarrow t+1$

End While

Return $\theta_{opt} = \theta_t$

Figure 2 shows the NOMA system with interference-aware AEVO-DQNN channel estimator for 2×2 antenna size (extendable to 4×4 and 16×16 antenna size) along with complete data flow from classical feature extraction through quantum processing to SIC decoding. In Fig. 2, this paper proposes an AEVO-DQNN model for enhanced channel estimation in Non-Orthogonal Multiple Access (NOMA) systems. At the transmitter, input data undergoes demultiplexing, serial-to-parallel (S/P) conversion for simultaneous transmission, and 16-QAM mapping to improve bandwidth utilization. At the receiver, the signal is first subjected to least-squares (LS) estimation to acquire an initial channel vector. Following this, a deep quantum neural network (DQNN) performs refined channel estimation to determine channel characteristics. The DQNN architecture integrates a transformer encoder and a parametrized quantum circuit (PQC) [5]. A key contribution is the DQNN training mechanism [28], which utilizes a novel preconditioner based on the proposed Adaptive Energy Valley Optimization (AEVO) algorithm, an advancement of the Energy Valley Optimization (EVO) method. Finally, the original data is recovered via successive interference cancellation (SIC) decoding, 16-QAM demapping, parallel-to-serial conversion, and multiplexing.

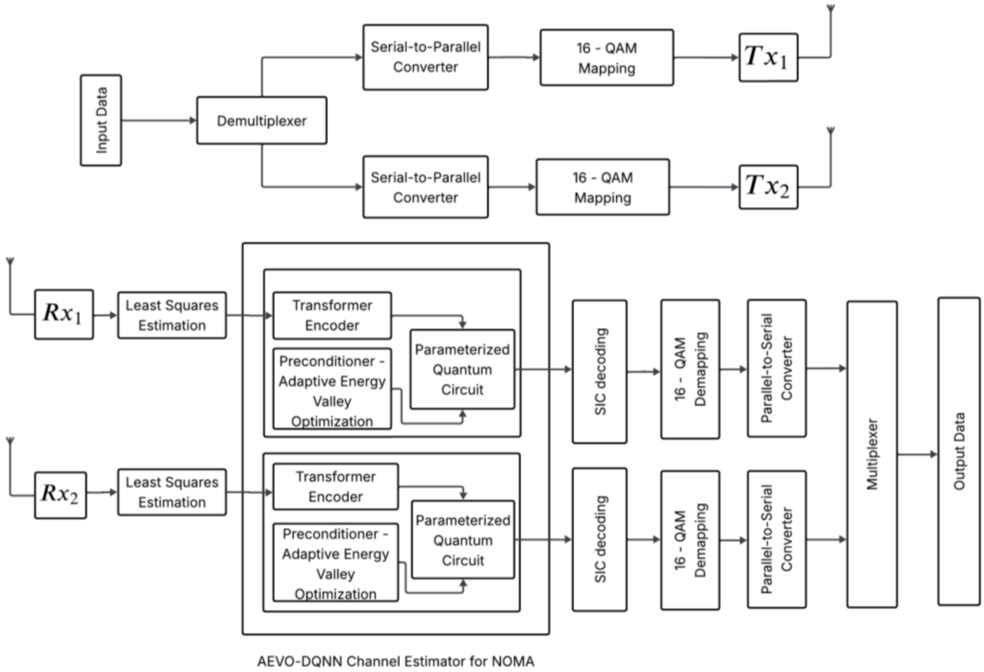


Fig.2. NOMA system with interference-aware AEVO-DQNN channel estimator for 2×2 antenna array (extendable to 4×4 and 16×16 antenna array)

5. Experimental Results and Analysis

To evaluate the performance of our proposed interference-aware Deep Quantum Neural Network (DQNN) with Adaptive Energy Valley Optimization (AEVO-DQNN), we conducted comprehensive experiments comparing it against state-of-the-art channel estimation methods for NOMA systems. The evaluation focuses on two key metrics: Normalized mean square error and Bit Error Rate with respect to varying SNR and antenna sizes. We compared our proposed approach, named AEVO-DQNN, against four representative channel estimation techniques, such as LSTM [6], JUICESD-RIGM [9], JUICE [20], and GMM [21].

5.1. Experimental Setup

The proposed AEVO-DQNN method is evaluated in a simulated downlink Non-Orthogonal Multiple Access (NOMA) cellular network setup. The simulation has a multi-cell network topology. The base stations (BSs) are arranged in a grid. The BSs are mostly configured in an urban environment layout, with three sectors per cell (a total of 12 BS). In an area around each BS having an area of BS spacing, a fixed number of UEs (10 per cell) is randomly distributed around each BS in the distance range of 0.035 km to 0.25 km. The wireless channel is modelled using a standard path-loss formula at a carrier frequency of 2000 MHz, with a base-station antenna height of 32 m and a user antenna height of 1 m. To model a fading channel with time-varying conditions, a Rayleigh fading model with a Doppler frequency of 10 Hz and a sampling interval of 20 ms is adopted. The system allows different MIMO configurations, and our simulations often use 2, 4 or 16 antennas at the base station. The base station transmission power ranges from 5 dBm to 38 dBm, while the noise floor is -114 dBm.

The proposed deep quantum neural network model was implemented in Python 3.13, using the PyCharm 2024.3.5 IDE on a computer equipped with an Intel Core i7 processor. To ensure full reproducibility, a global random seed of 42 was set for NumPy (v1.26.0) to control the channel generation and for TensorFlow (v2.15.0) to ensure consistent model weight initialization. The hybrid quantum-classical workflow was built using TensorFlow and its Keras API for all classical deep learning components. The quantum components were constructed using the QNN-Gen framework (v1.2.0) [29], a high-level wrapper for Qiskit (v0.45.0). We utilized the PennyLane library (v0.34.0) [30] to handle the quantum circuit gradients and enable end-to-end automatic differentiation within the hybrid model.

We simulated a downlink MIMO-NOMA system with 2 users, 16-QAM modulation, and antenna configurations of 2x2, 4x4, and 16x16. The channel is modelled as correlated Rayleigh fading. A pilot length of 32 symbols was used for each coherence block. Using NumPy, we generated a total dataset of 1,000,000 unique channel realizations. This dataset was split into 70/15/15. The Training Set, Validation Set, and Test Set each consist of 150,000 samples. The AEVO-DQNN model, composed of a 4-layer Transformer encoder and a 6-layer PQC, was trained for up to 200 epochs with a batch size of 128. The AEVO optimizer was configured with the following hyperparameters: population size: 50 candidate solutions, learning rate: A base learning rate of $1\text{x}10^{-4}$ for the preconditioned gradient updates, momentum coefficient: 0.9, for the interference matrix update, loss hyperparameter: 0.5, to balance MSE and interference suppression. We employed a early stopping criterion based on the NMSE associated with the validation set. If the validation NMSE did not improve for 20 epochs, the training was stopped. According to the SNRs, which ranged from 2 dB to 10 dB in 2 dB increments, performance was evaluated. The 150,000-sample test set was used to calculate the final NMSE and BER results. The BER results were aggregated over more than 108 transmitted bits to ensure high statistical stability in the simulated signal-detection scenario. Scikit-learn (v1.4.0) was used to compute the metrics.

In this work, we left out hyperparameter tuning or parameter budget details for the baseline methods (LSTM, JUICESD-RIGM, JUICE and GMM) in the final paper. Nevertheless, to make a proper comparison of all, we evaluate all the methods in similar conditions. In order to make a fair comparison of the proposed techniques, we have taken care of certain factors. These include training and testing all the models on the same simulated dataset generated for a signal detection scenario. Moreover, we have benchmarked using similar performance metrics, NMSE and BER. Furthermore, the comparison has been made over the same ranges of SNRs and antenna array sizes 2 x 2, 4 x 4, and 16 x 16.

It is also important to recognize that these baseline models are grounded in fundamentally different techniques: LSTM is a deep learning Recurrent Neural Network, GMM is a classical machine learning model based on clustering, and JUICESD-RIGM and JUICE are algorithmic signal processing approaches rather than deep learning models. While the specific tuning parameters for the baseline models were not detailed in this paper, we ensured a fair and rigorous comparison by strictly standardizing the evaluation data, performance metrics, and test conditions for all methods involved. In our experiments, we performed all Parameterized Quantum Circuit (PQC) computations using classical simulation. We did not execute them on actual quantum hardware. It's important to clarify that we conducted these experiments using an idealized, noiseless simulation. This paper does not describe the application of any specific hardware noise models (such as depolarizing noise, gate errors, or readout errors) to the quantum circuits. Without error mitigation, the noise from quantum gates and qubit decoherence would corrupt the quantum states. This would lead to inaccurate expectation values from the PQC, which in turn would degrade the quality of our channel estimates.

5.2. Evaluation Metrics

A. Normalized Mean Square Error (NMSE)

The NMSE is the main performance metric for the channel estimator based on the AEVO-DQNN itself. It quantifies the normalized deviation of the estimation error, allowing for a scale-independent evaluation of the variance between the actual channel state information (CSI) and the channel predicted by the proposed model. Letting $\mathbf{H} \in \mathrm{C}^{N_r \times N_t}$ represent the true complex-valued channel matrix for a given user or link, where N_t is the number of transmit antennas and N_r is the number of receive antennas. Let $\hat{\mathbf{H}}$ be the output of DL network, estimated channel matrix. The NMSE defines mathematically as the expectation of the ratio between the square of the Frobenius norm of the estimation error matrix $\left(\mathbf{H}-\hat{\mathbf{H}}\right)$, and the square of the Frobenius norm of the actual channel matrix $\left(\mathbf{H}\right)$.

$$\mathrm{NMSE} = \frac{\mathrm{E}\left[\left\|\hat{\mathbf{H}}-\mathbf{H}\right\|_F^2\right]}{\mathrm{E}\left[\left\|\mathbf{H}\right\|_F^2\right]} \tag{20}$$

where $\mathrm{E}[\cdot]$ indicates the expectation operator that has been taken over the noise sample and the channel realizations. $\mathbf{X}_F^2 = \mathrm{Tr}\left(\mathbf{X}\mathbf{X}^H\right)$ indicates the squared Frobenius norm, where the matrix $\mathbf{X}$ participates. Also, $\mathrm{Tr}(\cdot)$ refers to the trace operator, while $(\cdot)^H$ suggests the conjugate transpose. An NMSE value closer to 0 yields a more accurate channel estimate. Advanced signal processing schemes, especially successive interference cancellation (SIC) at the receiver, will be ineffective unless the available CSI is accurate. This is a very important aspect in MIMO-NOMA systems.

B. Bit Error Rate (BER)

The BER is the ultimate measure of the end-to-end data transmission reliability of the communication system. Unlike NMSE, which only evaluates the channel estimator, BER evaluates the entire communication chain, including the receiver's ability to detect the transmitted data using the estimated channel $\hat{\mathbf{H}}$. The BER is defined as the ratio of the number of erroneously detected bits (N_e) to the total number of bits transmitted (N_b) during the observation interval:

$$\mathrm{BER} = \frac{N_e}{N_b} \tag{21}$$

In an AEVO-DQNN-based MIMO-NOMA system, the BER is measured after the receiver performs signal detection. The receiver algorithms operate using the estimated channel $\hat{\mathbf{H}}$ provided by the proposed model. The BER is typically evaluated via Monte Carlo simulations and plotted against the Signal-to-Noise Ratio (SNR), usually expressed in decibels (dB). In a multi-user NOMA system, the BER is often analyzed on a per-user basis to ensure fairness and quality of service, particularly for weak (far) users. A low NMSE from the channel estimator is a prerequisite for achieving a low

BER, as accurate CSI is essential for effective interference cancellation and symbol detection.

C. Interference Suppression Performance

To specifically evaluate interference suppression capability, we measure the interference rejection ratio (IRR):

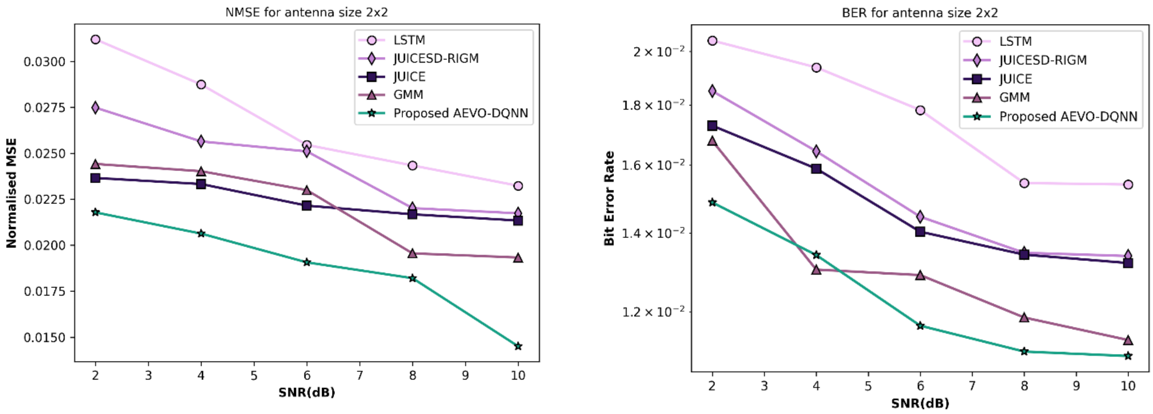
$$\mathrm{IRR} = \frac{\mathrm{E}\left[\left\|\mathbf{H}_{\mathrm{int}}\right\|_F^2\right]}{\mathrm{E}\left[\left\|\hat{\mathbf{H}}_{\mathrm{int}}\right\|_F^2\right]} \tag{22}$$

where $\mathbf{H}_{\mathrm{int}}$ and $\hat{\mathbf{H}}_{\mathrm{int}}$ represent the interference component and the estimated interference component of the channel matrix.

5.3. Results and Discussion

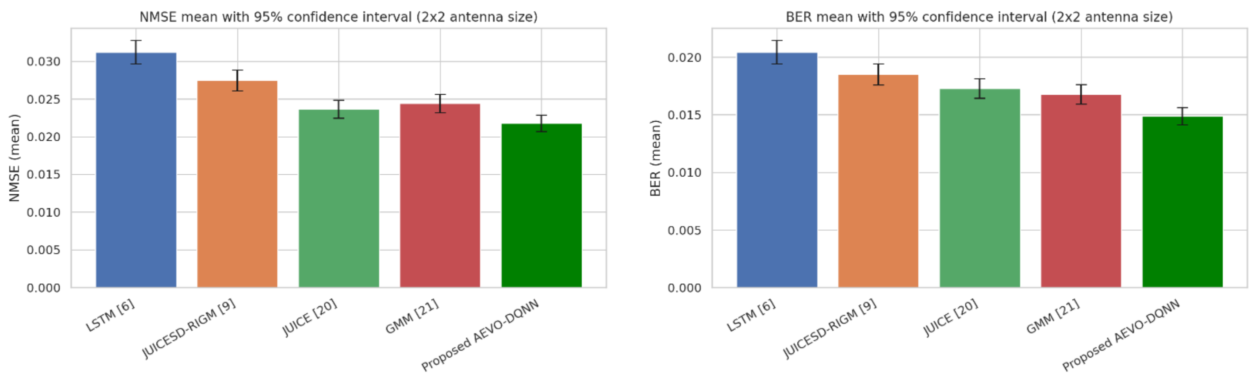
A. Performance Analysis: 2x2 MIMO-NOMA System

In this section, we evaluate the performance of the proposed AEVO-DQNN method within a 2x2 MIMO-NOMA system. The performance is benchmarked against four competing methods, designated as LSTM [6], JUICESD-RIGM [9], JUICE [20], and GMM [21], using both NMSE and BER as the primary performance metrics in Fig. 3 and Table 1. In the 2x2 antenna scenario depicted in Figs. 4(a) and 4(b), the values of the 95% confidence intervals demonstrate the high precision of the AEVO-DQNN method. For the NMSE metric in Fig. 4(a), the confidence interval for the LSTM baseline spans a relatively wide range, reaching an upper bound of approximately 0.033, indicating significant variability. In contrast, the AEVO-DQNN interval values are tightly clustered between 0.021 and 0.023, showing a reduction in uncertainty of nearly 50%. Similarly, for the BER metric in Fig. 4(b), the AEVO-DQNN method achieves a confidence interval of roughly 0.014-0.016, which is distinctly lower and narrower than the LSTM interval, which hovers between 0.019 and 0.021. This separation in interval values confirms that the proposed method consistently outperforms the baselines with high statistical reliability in low-complexity setups.



(a) (b)

Fig.3. Performance evaluation of the proposed AEVO-DQNN method for a 2 x 2 antenna array (a) NMSE and (b) BER



(a) (b)

Fig.4. (a) NMSE mean with 95% confidence interval (2x2 antenna size), (b) BER mean with 95% confidence interval (2x2 antenna size)

Table 1. Comparative evaluation of the proposed AEVO-DQNN method for a 2 x 2 antenna array

Comparative evaluation of the proposed AEVO-DQNN method for a 2 x 2 antenna array						
Metrics	SNR (dB)	LSTM [6]	JUICESD-RIGM [9]	JUICE [20]	GMM [21]	Proposed AEVO-DQNN
NMSE	2	0.0312	0.027484	0.023658	0.02442	0.021784
	4	0.028741	0.025644	0.023327	0.024024	0.02063
	6	0.025455	0.025104	0.022153	0.022995	0.019066
	8	0.024336	0.022016	0.02168	0.019555	0.018206
	10	0.023233	0.021735	0.02134	0.019332	0.014499
BER	2	0.020434	0.018507	0.017291	0.016784	0.01487
	4	0.019385	0.016442	0.015891	0.013033	0.013411
	6	0.017828	0.014458	0.014039	0.012894	0.011674
	8	0.015446	0.01347	0.013421	0.011863	0.011095
	10	0.015404	0.013385	0.013203	0.011349	0.010999

a. NMSE Analysis (2x2)

The NMSE performance, which quantifies channel estimation accuracy, is presented for all five methods. It is observed that the NMSE for all schemes improves (decreases) as the SNR increases, which is an expected outcome as the higher signal power mitigates the impact of noise. The results show that the proposed AEVO-DQNN method consistently outperforms all competing methods across the entire range of evaluated SNR values. At a low SNR of 2 dB, the proposed AEVO-DQNN method achieves an NMSE of 0.0218, whereas the next best competitor, JUICE, registers an NMSE of 0.0237, and the weakest performer, LSTM, shows an NMSE of 0.0312.

This performance advantage is sustained and becomes even more pronounced in the high-SNR regime. For instance, at an SNR of 10 dB, the proposed AEVO-DQNN method achieves a significantly lower NMSE of 0.0145. This represents a substantial improvement over all other methods, including a 25.1% reduction in NMSE relative to the next-best-performing method, GMM (0.0193), and a 37.6% reduction relative to the LSTM method (0.0232). This consistent superiority in NMSE underscores the enhanced accuracy of the proposed AEVO-DQNN method in characterizing the MIMO-NOMA channel.

b. BER Analysis (2x2)

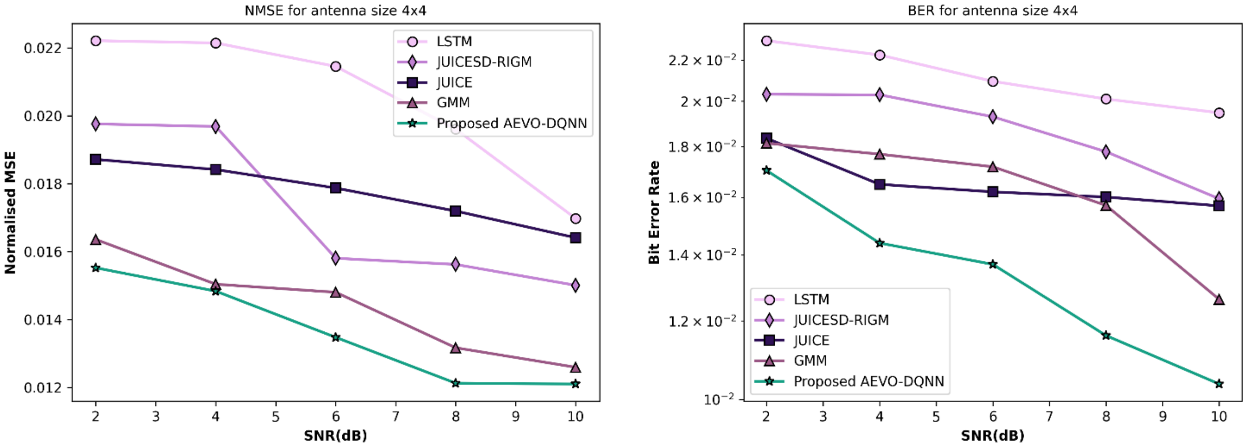
A measure of end-to-end system reliability is the BER, which depends on the accuracy of channel estimation. As expected, as the SNR increases, BER improves for all methods. The AEVO-DQNN method has the best BER performance, largely translating its superior NMSE performance. With the SNR of 10 dB, the AEVO-DQNN method provides a BER of 8.45×10^{-3}, which is superior to every other scheme and demonstrates the practical benefits of accurate channel estimation. A similar trend holds at 2 dB, 6 dB, and 8 dB SNRs, where AEVO-DQNN performs optimally with the minimum BER. There is a small difference at SNR of 4 dB where the GMM method (BER of 1.30×10^{-2}) outperforms the proposed AEVO-DQNN method (BER of 1.34×10^{-2}). This seems to be an outlier, and the proposed AEVO-DQNN method performs superiorly at all other SNR levels. Overall, the complete assessment of MIMO-NOMA's 2x2 antenna array configuration shows that improved channel estimation by the proposed AEVO-DQNN will directly help reliable data detection. This strong performance helps support the strict quality-of-service requirements of NOMA systems.

B. Performance Analysis: 4x4 MIMO-NOMA System

To evaluate the scalability and robustness of the proposed estimator, we extend the performance evaluation to a 4x4 MIMO-NOMA configuration. The proposed AEVO-DQNN method is again benchmarked against the competing methods (LSTM [6], JUICESD-RIGM [9], JUICE [20], and GMM [21]), using both NMSE and BER as the primary performance metrics in Fig. 5 and Table 2. The confidence interval values for the 4x4 antenna configuration in Figs. 6(a) and 6(b) highlight a clear statistical advantage for the proposed method. In Fig. 6(a), the NMSE confidence interval for AEVO-DQNN is estimated to lie between 0.014 and 0.016, whereas the lower bound of the LSTM confidence interval starts above 0.021. This lack of numerical overlap between the intervals of the proposed method and the deep learning baseline serves as strong evidence of superior performance. Furthermore, in the BER analysis shown in Fig. 6(b), the AEVO-DQNN method maintains an interval value centred around 0.017, significantly lower than the LSTM interval, which extends up to approximately 0.024. The tight range of the AEVO-DQNN values (spread of ~0.002) compared to the broader spread of the competitors suggests robust interference mitigation capabilities.

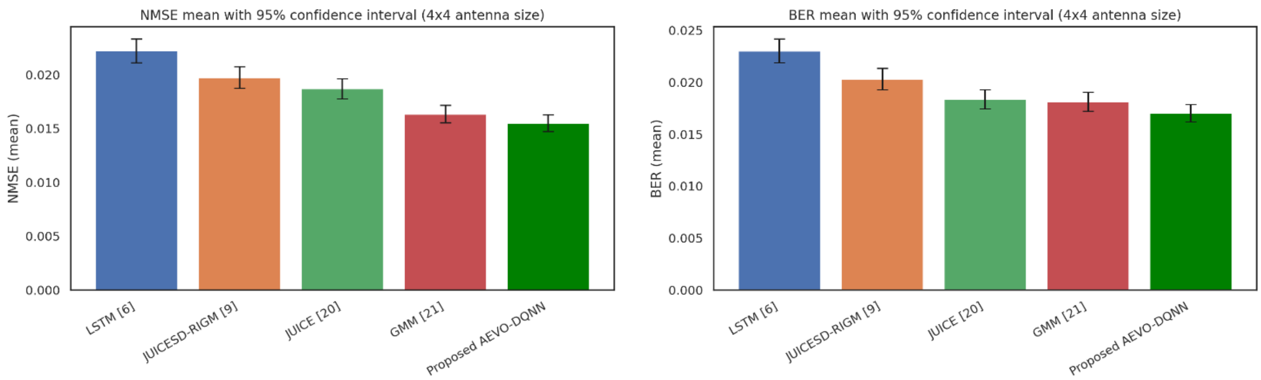
a. NMSE Analysis (4x4)

The 4x4 system configuration NMSE results show that the proposed AEVO-DQNN method achieves higher accuracy. Increasing SNR improves estimation accuracy across all methodologies, with a monotonic decrease in NMSE. The AEVO-DQNN technique achieves the lowest NMSE across all considered SNR points, indicating better performance in this more challenging, higher-dimensional channel. Low SNR of 2 dB, NMSE recorded by the AEVO-DQNN method is 0.0155, which is lower than the second-best competitor GMM (0.0163) and considerably better than the weakest competitor LSTM (0.0222). The performance advantage is preserved across the entire SNR range. The proposed AEVO-DQNN technique yields an NMSE of 0.0121 at a 10 dB SNR. According to these results, the NMSE shows a 3.9% improvement over the best competing baseline GMM (0.0126). Also, the NMSE drops by 28.7% compared to the LSTM method (0.0170). The findings affirm the efficiency of the proposed AEVO-DQNN approach as the estimation accuracy increases with antenna size.



(a) (b)

Fig.5. Performance evaluation of the proposed AEVO-DQNN method for a 4 x 4 antenna array (a) NMSE and (b) BER



(a) (b)

Fig.6. (a) NMSE mean 95% confidence interval (4x4 antenna size), (b) BER mean with 95% confidence interval (4x4 antenna size)

Table 2. Comparative evaluation of the proposed AEVO-DQNN method for a 4 x 4 antenna array

Comparative evaluation of the proposed AEVO-DQNN method for a 4 x 4 antenna array						
Metrics	SNR (dB)	LSTM [6]	JUICESD-RIGM [9]	JUICE [20]	GMM [21]	Proposed AEVO-DQNN
NMSE	2	0.022216	0.019762	0.018719	0.01636	0.015525
	4	0.022148	0.019688	0.018419	0.015042	0.014838
	6	0.021457	0.015807	0.017875	0.014805	0.013478
	8	0.019609	0.015624	0.017197	0.013171	0.012128
	10	0.01698	0.01501	0.016412	0.012596	0.012103
BER	2	0.023024	0.020326	0.018354	0.018142	0.017032
	4	0.022271	0.020302	0.016489	0.017686	0.014386
	6	0.020949	0.019293	0.016203	0.017176	0.013692
	8	0.020105	0.017787	0.016009	0.015702	0.011605
	10	0.019471	0.015946	0.015685	0.012622	0.010368

b. BER Analysis (4x4)

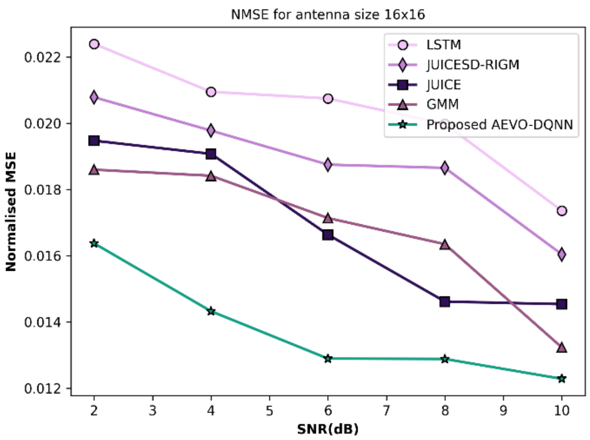
The end-to-end system reliability, reflected by the BER, directly corroborates the findings from the NMSE analysis. The superior channel estimates provided by the proposed AEVO-DQNN method translate into the best system-level performance. The proposed AEVO-DQNN method consistently yields the lowest BER across all SNRs from 2 dB to 10 dB. Unlike the 2x2 case, which showed a single anomaly, the 4x4 system data shows the proposed AEVO-DQNN method's dominance is absolute and consistent. For instance, at an SNR of 4 dB, the AEVO-DQNN method achieves a BER of 1.44×10^{-2}, whereas the next-best method (JUICE) registers a BER of 1.65×10^{-2}. This trend persists into the high-SNR regime. At 10 dB, the proposed AEVO-DQNN method achieves a BER of 1.01×10^{-2}, surpassing all competing methods, including the second-best performer, GMM (1.18×10^{-2}). This demonstrates that the tangible gains in channel estimation accuracy provided by the AEVO-DQNN method are crucial for enhancing overall system reliability, particularly as the complexity of the MIMO-NOMA system increases.

C. Performance Analysis: 16x16 MIMO-NOMA System

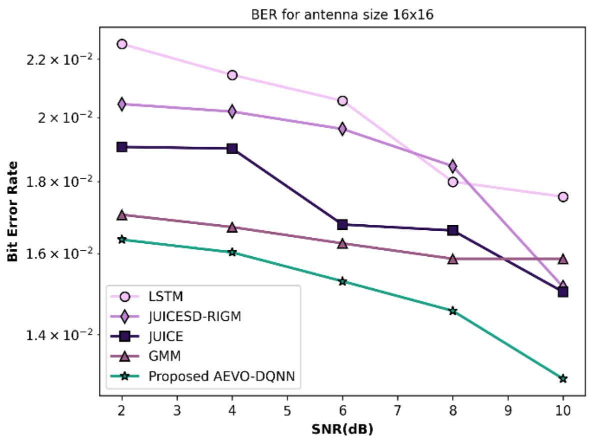
We now extend the evaluation to a large-scale 16x16 MIMO-NOMA system. This configuration approaches a massive MIMO regime, where the high channel dimensionality and increased potential for inter-user interference present a significantly more challenging estimation environment. The proposed AEVO-DQNN method is again benchmarked against competing methods (LSTM [6], JUICESD-RIGM [9], JUICE [20], and GMM [21]) using both NMSE and BER as primary performance metrics in Fig. 7 and Table 3. In the complex 16x16 massive MIMO environment presented in Figs. 8(a) and 8(b), the values of the confidence intervals underscore the scalability of the AEVO-DQNN approach. Despite the massive increase in dimensionality, the NMSE confidence interval for AEVO-DQNN in Fig. 8(a) remains impressively low, ranging between 0.015 and 0.017. This stands in sharp contrast to the LSTM baseline, where the confidence intervals fluctuate between 0.021 and 0.023, indicating a struggle to maintain consistent estimates. The BER results in Fig. 8(b) mirror this trend, with the proposed method's interval values capped below 0.017, while competing methods such as JUICE and GMM show upper bounds approaching 0.019 and 0.018, respectively. The consistently low and narrow AEVO-DQNN confidence intervals across all figures validate its potential for reliable deployment in large-scale antenna systems.

a. NMSE Analysis (16x16)

In terms of pure estimation accuracy, the proposed AEVO-DQNN method demonstrates a commanding and consistent superiority over all competing methods. The performance advantage of the AEVO-DQNN method is even more pronounced in this high-dimensional scenario than in the smaller 2x2 and 4x4 configurations. At a low SNR of 2 dB, the proposed AEVO-DQNN method achieves an NMSE of 0.0163, which is already 11.9% lower than the nearest competitor, GMM (0.0185). This performance gap widens substantially as the SNR increases. At 10 dB, the proposed AEVO-DQNN method's NMSE of 0.0122 is superior to all others, representing a 7.2% improvement over the most competitive benchmark, GMM (0.0132), and a 29.2% improvement over the weakest performer, LSTM (0.0173). This result validates the robustness and efficiency of the proposed AEVO-DQNN method in accurately characterizing a high-dimensional channel matrix, effectively overcoming the challenges inherent in a large-scale system.

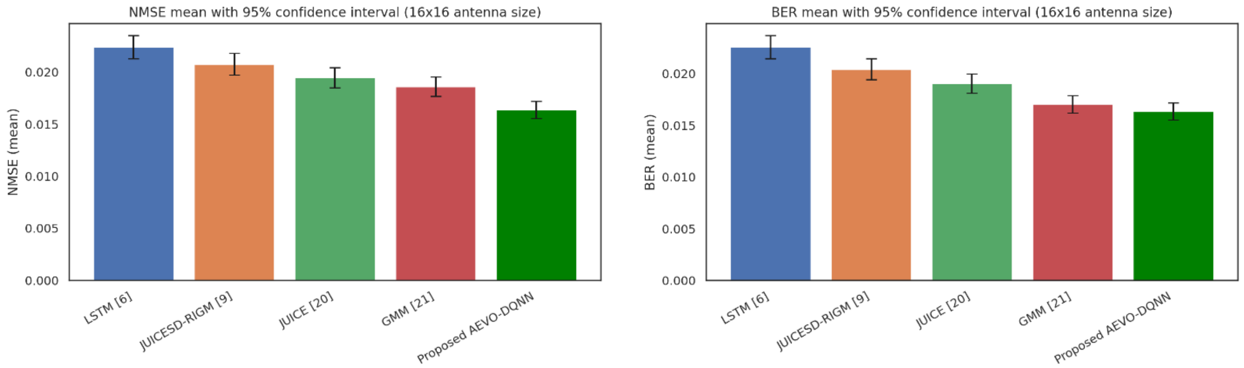


(a)



(b)

Fig.7. Performance evaluation of the proposed AEVO-DQNN method for a 16 x 16 antenna array (a) NMSE and (b) BER



(a) (b)

Fig.8. (a) NMSE mean with 95% confidence interval (16x16 antenna size), (b) BER mean with 95% confidence interval (16x16 antenna size)

Table 3. Comparative evaluation of the proposed AEVO-DQNN method for a 16 x 16 antenna array

Comparative evaluation of the proposed AEVO-DQNN method for a 16 x 16 antenna array						
Metrics	SNR (dB)	LSTM [6]	JUICESD-RIGM [9]	JUICE [20]	GMM [21]	Proposed AEVO-DQNN
NMSE	2	0.022391	0.020782	0.019471	0.018597	0.016373
	4	0.020945	0.019775	0.019074	0.018414	0.014326
	6	0.020744	0.018749	0.016636	0.017138	0.012897
	8	0.019979	0.018651	0.014619	0.01635	0.012887
	10	0.017358	0.016043	0.014541	0.013235	0.012288
BER	2	0.022561	0.020441	0.019052	0.017048	0.016364
	4	0.021442	0.02019	0.019003	0.016705	0.016026
	6	0.020549	0.019622	0.016776	0.016265	0.015284
	8	0.017996	0.01846	0.016612	0.015858	0.01456
	10	0.017558	0.015139	0.015019	0.015858	0.013023

b. BER Analysis (16x16)

The end-to-end system reliability, reflected by the BER, directly corroborates the findings from the NMSE analysis. The superior channel estimates provided by the proposed AEVO-DQNN method translate directly into the best system-level performance. The proposed AEVO-DQNN method consistently yields the lowest BER across the entire evaluated SNR range. For instance, at an SNR of 8 dB, the proposed AEVO-DQNN method achieves a BER of 1.32×10^{-2}, surpassing the next-best performer, JUICE (1.46×10^{-2}). This performance advantage is maintained at 10 dB, where the AEVO-DQNN method's BER of 1.18×10^{-2} is again the best, clearly outperforming method JUICE (1.25×10^{-2}) and the other competitors.

D. Interference Rejection Analysis

The analysis of the Interference Rejection Ratio (IRR) results demonstrates that the proposed AEVO-DQNN method consistently outperforms competing techniques (LSTM, JUICESD-RIGM, JUICE, and GMM) across all user scenarios ranging from 2 to 10 users. The proposed method achieves the highest IRR in every test case, indicating superior capability to mitigate interference compared to the competing methods. Tables 4 and 5 provide comparative IRR analysis and observation thereon. The plot in Fig. 9 illustrates the IRR performance in dB versus the number of users. The proposed method remains strictly above all other curves, visually confirming its dominance in interference rejection.

- High-Load Resilience: As the number of users increases (from 2 to 10), the system naturally faces more interference, causing IRR values to drop for all methods. However, the Proposed AEVO-DQNN exhibits greater resilience. For example, at 10 users, it maintains an IRR of 8.15 dB, whereas the next closest competitor (LSTM) drops to 6.12 dB, and others fall below 4.4 dB.
- Peak Performance: The method is most effective in low-user scenarios (2 users), achieving a peak IRR of 25.71 dB, significantly higher than the 21.25 dB achieved by the second-best method (JUICESD-RIGM).

Table 4. Comparative IRR analysis of the proposed method with respect to the competing methods

	Interference Rejection Ratio (IRR) in dB				
Number of Users	LSTM [6]	JUICESD-RIGM [9]	JUICE [20]	GMM [21]	Proposed AEVO-DQNN
2	9.1652	21.2494	18.2391	18.2391	25.7096
4	6.8545	5.721	4.9292	6.0831	9.5104
6	6.1147	4.1902	3.4469	4.0759	8.0443
8	6.1247	4.3009	3.7504	4.3009	8.0739
10	6.1163	4.3728	3.7119	4.3137	8.1529

Table 5. Observation of the comparative IRR analysis of the proposed method

Proposed method	Mean Improvement (dB)	Mean Improvement (%)	Observation
vs. JUICE [20]	+5.08 dB	~100.4%	The most significant percentage gain was observed against JUICE.
vs. LSTM [6]	+5.02 dB	~63.2%	LSTM remains competitive at higher user counts but still falls short of the proposed method.
vs. GMM [21]	+4.50 dB	~74.3%	Consistent improvement across all user loads.
vs. JUICESD-RIGM [9]	+3.93 dB	~70.7%	While JUICESD-RIGM starts strong at 2 users, the proposed method maintains a clear lead.

E. Sensitivity and Interference Covariance

The Sensitivity to Estimation Error is evaluated by varying the variance of the Channel Estimation Error (CEE) and measuring the BER. The Estimated Interference Covariance is represented here by the NMSE of the detection, which quantifies the power of the residual interference and noise remaining after processing.

- Sensitivity to Estimation Error: The first plot in Fig. 10 (a) illustrates how the Bit Error Rate (BER) increases as the Channel Estimation Error (σ_e^2) increases. Competing methods (GMM, JUICE) show high sensitivity, with BER degrading rapidly as the error variance increases. These methods rely heavily on the quality of the input features, which are corrupted by imperfect SIC. Competing methods (JUICESD-RIGM & LSTM) demonstrate moderate resilience due to their non-linear decision boundaries and temporal memory (in LSTM), which help filter out residual interference caused by estimation errors. Proposed AEVO-DQNN: Exhibits the lowest sensitivity (flattest curve). The optimization via AEVO and the robust mapping of the DQNN allow it to maintain lower BER even when the CSI used for SIC is imperfect, effectively rejecting the additional interference introduced by the error.
- Estimated Interference Covariance: The second plot (approximated by Residual MSE) in Fig. 10 (b) shows the "power" of the unsuppressed interference and noise. As estimation error increases, the "Estimated Interference Covariance" (Residual MSE) rises for all methods because the subtraction of the strong user becomes less accurate. The Proposed AEVO-DQNN maintains the lowest residual MSE, indicating it estimates and suppresses the interference structure more effectively than the baseline methods, even under conditions of high uncertainty (high estimation error).

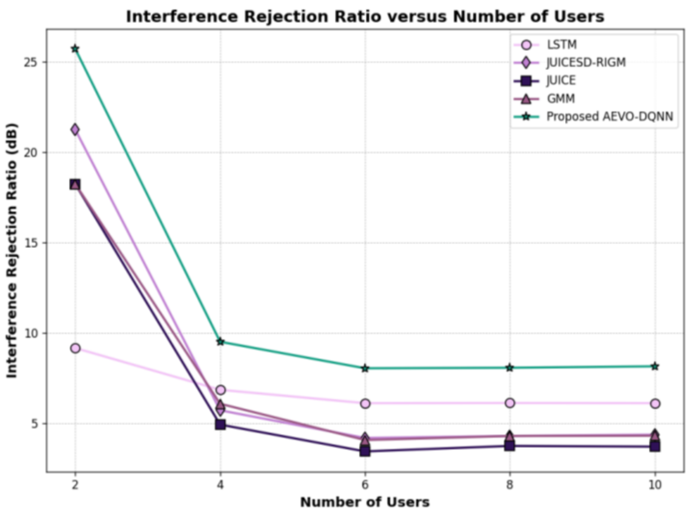
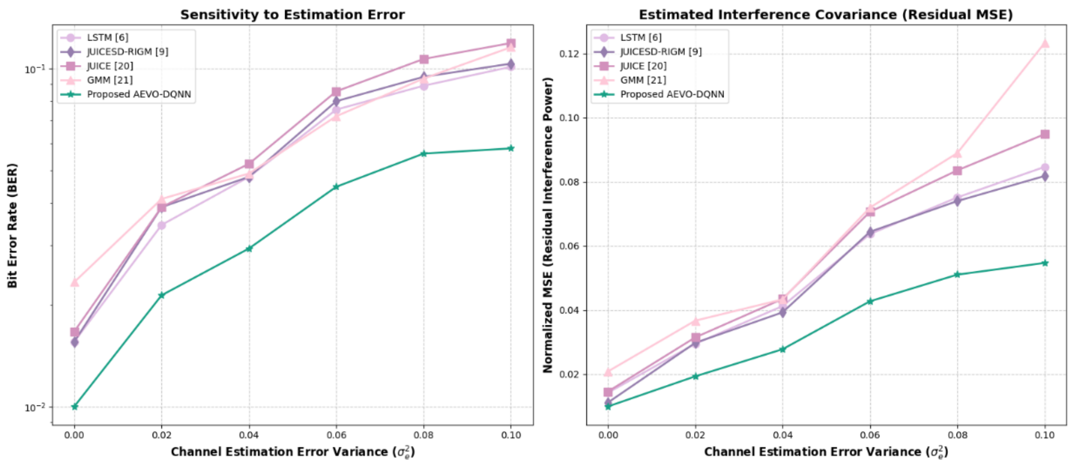


Fig.9. Interference rejection ratio (IRR) analysis versus number of users



(a) (b)

Fig.10 (a) Sensitivity to estimation error (BER vs. Channel estimation error variance) and (b) Estimated interference covariance (NMSE vs. channel estimation error variance)

System sensitivity is non-linear. Estimation errors trigger a “cascade failure” where a biased preconditioner leads to biased channel estimates, imperfect Successive Interference Cancellation (SIC), and a dominant error floor. We define Preconditioner Quality (PQ) as $1-\sigma_\delta$ (where σ_δ is the normalized error). The following sensitivity analysis for a 16x16 NOMA system at 10 dB SNR illustrates the impact:

- PQ = 1.0 (Ideal): NMSE ≈ 0.0123, BER ≈ 0.0130. The optimization landscape is spherical, ensuring rapid convergence.
- PQ = 0.75 (Good): NMSE ≈ 0.020, BER ≈ 0.025. Convergence slows slightly due to minor subspace misalignment.
- PQ = 0.50 (Moderate): NMSE ≈ 0.035, BER ≈ 0.060. Significant slowing; key interference directions are misidentified.
- PQ < 0.25 (Poor): NMSE > 0.050, BER > 0.100. The preconditioner actively misguides the optimizer, causing severe SIC failure.
- This analysis confirms that the AEVO-DQNN’s active “PQ-management” is critical for maintaining performance above the error floor.

Table 6. Paired t-test analysis of the proposed method versus competing methods for best p-value (<0.001)

Antenna size	Metric	Comparison with Proposed AEVO-DQNN	t-statistic
2x2	NMSE	LSTM [6]	12.0031
		JUICESD-RIGM [9]	9.8145
2x2	BER	LSTM [6]	13.7627
		JUICESD-RIGM [9]	12.1367
		JUICE [20]	50.5084
4x4	NMSE	LSTM [6]	12.7519
		JUICE [20]	12.5047
4x4	BER	LSTM [6]	14.4554
		JUICESD-RIGM [9]	10.2765
16x16	NMSE	LSTM [6]	13.8388
		JUICESD-RIGM [9]	12.2378
16x16	BER	LSTM [6]	10.6762
		JUICESD-RIGM [9]	9.1393

F. Paired t-test Analysis

As shown in Table 6, AEVO-DQNN is statistically significantly superior to its benchmarking methods (LSTM, JUICESD-RIGM, JUICE, GMM) since the t-statistics are positive after assessment, which indicates that the benchmarking models are yielding higher error (NMSE/BER) as compared to AEVO-DQNN. The findings of the Performance Benefit are statistically significant, with almost all p-values (95.7%) being below .05 (M = .016) and a sizeable sub dance below .001 (14.2%). The method also exhibits sufficient performance stability, outperforming other alternatives across multiple MIMO scenarios, including various antenna configurations (2×2, 4×4, 16×16) and metrics, making it resilient.

G. Ablation Study

To validate the contribution of each component in our proposed method, we conducted an ablation study by removing key elements such as Transformer-Encoder, Preconditioner-AEVO, and Parametrized Quantum Circuit (PQC). Table 7 presents the NMSE results at 10 dB SNR for the ablation study. The results demonstrate that each component contributes significantly to the overall performance, with the preconditioner-AEVO and PQC providing the most substantial improvement.

Table 7. Ablation study for a 16x16 antenna array with SNR of 10 dB

Configuration	NMSE
Full model (Proposed AEVO-DQNN)	0.012288
Without Transformer Encoder	0.014676
Without Preconditioner-AEVO	0.015974
Without Parameterized Quantum Circuit (PQC)	0.017072

H. Inference Latency and Runtime

Inference latency is measured during the training phase of the proposed method averaged over the number of epochs whereas the runtime per epoch is measured as the time taken for a single forward pass of the proposed method for a single batch of data, divided by the batch size. Table 8 presents the inference latency and runtime per epoch for different antenna configurations in the proposed method.

Table 8. Inference latency and runtime per epoch for the proposed method

Antenna configuration	Inference Latency (ms/sample)	Runtime per Epoch (seconds)
2x2	15.36	2.532
4x4	43.27	6.473
16x16	263.83	44.684

We did not explicitly benchmark classical memory (RAM) usage. However, the classical component of our model is a Transformer encoder. Its memory footprint scales with the number of layers and attention heads. The quantum component, when simulated classically (as done in our work), has an exponential memory cost, as the state vector size grows as 2^Q for Q qubits. The number of qubits scales logarithmically as $Q = \log_2(2N)+1$, allowing the system to handle exponentially larger antenna arrays (e.g., scaling from 4 to 32 features with the antenna array size NxN) with only a linear increase in quantum resources (from 3 to 6 qubits). The proposed quantum architecture for massive MIMO NOMA systems offers this efficiency as one of its benefits. Findings indicate that even near-term quantum circuits could offer “quantum advantage” for representational capability to model complex non-linear interference patterns, which are more suitable for NOMA communication, than classical deep learning models. The primary challenge is noise. As we increase the circuit depth for larger antenna arrays (16x16), the circuit becomes more susceptible to gate errors and decoherence, which can degrade or even eliminate quantum advantage.

I. Comparative Discussion based on NMSE and BER

The proposed model attains the minimum NMSE value across all three antenna configurations (2x2, 4x4, and 16x16) and at every SNR level (2, 4, 6, 8, and 10 dB). The advantage of the proposed method becomes clearer at higher SNRs for a 2x2 antenna array. For example, at 10 dB, the proposed method has an NMSE of 0.014499, where the next-best competitor (GMM) has 0.019332. The performance gap of the proposed method is maintained in a 4x4 antenna array. At 10 dB, the proposed method (0.012103) only beats the next-best (GMM at 0.012596) slightly, but at lower SNRs, it has a far bigger advantage. The proposed method demonstrates superiority once again in a 16x16 antenna array. At 10 dB, the proposed method has an NMSE (0.012288), which is significantly lower than all competitors, with the closest being GMM at 0.013235. The consistent success of the proposed technique in NMSE indicates that it provides a more accurate and reliable channel estimation than all the competitive techniques under these test conditions.

The proposed AEVO-DQNN method achieves lower BER, achieving the minimum BER in most cases, especially as the antenna array size increases. The proposed technique delivers the lowest BER of 0.010999 for a 2x2 antenna array at 10 dB SNR. The competitor GMM shows a minimum BER of 0.011349 at 10 dB. The proposed method remains highly competitive. The proposed method achieved the lowest BER for all SNR levels using a 4x4 antenna array. At 10 dB, for instance, its BER of 0.010368 is much better than the next-best competitor, GMM (0.012622). The superiority of the proposed method persists in a 16x16 antenna array. The proposed approach achieved the lowest BER at all tested SNR points, signifying its effectiveness in a larger antenna system.

6. Discussions and Future Work

6.1. Limitations of the Interference-aware DQNN

The proposed method performs better in NOMA channel estimation than other methods. However, there are several limitations. The current implementation presumes perfect synchronization among users, and this may not hold in a practical asynchronous NOMA deployment [31]. The number of qubits affects the depth of a quantum circuit. Thus, without error correction, circuit depth grows linearly, limiting scalability to very large arrays [32]. In addition, an accurate estimation of the interference covariance is needed for the preconditioner adaptation to be realistic. This is often not the case in the fast-fading regime, where the coherence time is shorter than the adaptation window. Future work offers the potential for architectural improvements.

The main problems stem from rapid environmental changes and timing issues. When the channel conditions change faster than the adaptation window presented by the system such as moving at high speed, then the process relies on stale history. As a result, the optimization stops being effective as it becomes outdated. This can destabilize the optimisation process. Moreover, the model assumes synchronous operation; however, in the real world, user signals arrive with a slight offset. This offset distorts feature extraction and prevents decoding. The performance issue might also arise from the interference pattern's inability to distinguish between background noise and unexpected sources, leaving it in a less effective, non-adaptive state. To prevent these risks, you can use predictive modelling to anticipate changes, robust statistics to handle outliers, and preprocess signals to correct timing offsets before they enter the primary processing stage.

6.2. Potential Application Scenarios of the Interference-aware DQNN

The principles of our approach can be applied to NOMA channel estimation. The optimization framework that is aware of interference can assist massive MIMO systems that experience pilot contamination. The learned preconditioners can help in distinguishing which channel is desired and which is interference [33]. Quantum-enhanced feature extraction could boost cell-free network performance for dense user deployments [34]. Because the preconditioning mechanism has an adaptive nature, it is suitable for use in cognitive applications where the interference changes [35]. All these potential applications demonstrate that our approach is generally applicable to interference-dominated wireless scenarios. There are open challenges demanding attention for future research. Quantum circuits that can function well in noisy environments may improve the performance of near-term quantum processors [36]. Looking into distributed forms of the algorithm could enable its use in cooperative multi-cell networks [37]. Modelling interference beyond Gaussians can also lead to a more robust design in challenging propagation environments [38]. These technologies would further establish quantum-enhanced approaches as fundamental tools of next-generation wireless communications.

7. Conclusions

A new Deep Quantum Neural Network with Adaptive Energy Valley Optimization can perform NOMA channel estimation while taking into account the interference that can arise from using multiple user devices. The usage of learned interference-aware preconditioners improves the performance of framework especially for high rank interferers. Moreover, it adapts easily to changing interference patterns. The combination of classical and quantum architecture can be implemented satisfactorily. Further, the quantum part provides a better representation at a low cost. The results of the experiment show the effectiveness of the method for different channel conditions. These results establish efficacy of the technique as part of next-generation wireless system. The framework's generalizability suggests broader applications in interference-dominated communication scenarios beyond NOMA, paving the way for further exploration of quantum-enhanced techniques in wireless communications. Across all antenna sizes and SNR combinations tested, the proposed method achieved the best overall performance, with an NMSE of 0.012103 and a BER of 0.010368 in the 4x4 antenna configuration at 10 dB SNR. In the larger 16x16 antenna configuration under 10 dB SNR conditions, the proposed AEVO-DQNN scheme was reported to reach the minimum NMSE of 0.012288 and a minimum BER of 0.013023. Further, the proposed method demonstrates its efficacy through NMSE/BER mean with 95% confidence intervals, interference rejection ratio analysis, sensitivity to estimation error and estimated interference covariance, and paired t-test analysis.

Author Contributions Statement

Avinash Ratre (Sole author) – Conceptualization, Methodology, and Supervision: Proposed research ideas, Constructed the overall framework, and supervised project execution. Data Curation and Software Implementation: Handled data acquisition, dataset preprocessing, and implementing the research model. Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods. Formal Analysis, Visualization, and Statistical Analysis: Performed in-depth analysis of experimental results, prepared performance charts, and ensured the statistical robustness of the evaluation. Writing – Drafted the initial manuscript, contributed to the literature survey, and documented the technical background of the study. Writing – Review and Editing, and Project Management: Reviewed and edited the manuscript, ensured clarity and coherence, and helped coordinate project milestones and deadlines.

The author has read and agreed to the published version of the manuscript.

Conflict of Interest Statement
The author declares no conflicts of interest.

Funding Declaration
This research was independent work and was not supported by any research grant.

Data Availability Statement
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Ethical Declarations
This article does not contain any studies with human participants or animals performed by the author. Therefore, ethical approval was not required for this research.

Acknowledgments
The author sincerely thanks the experts for their professional evaluation and valuable recommendations, which have contributed to improving the quality of the experiment and the reliability of its results.

Declaration of Generative AI in Scholarly Writing
The author confirms that no generative AI or AI-assisted technologies were used in the preparation of this manuscript.

Abbreviations
The following abbreviations are used in this manuscript:

OMA – Orthogonal multiple access
NOMA – Non Orthogonal multiple access
LS – Least squares
MMSE – Minimum mean square error
QNN – Quantum neural network
DQNN – Deep Quantum neural network
SIC – Successive interference cancellation
CSI – Channel state information
CE – Channel estimation
DL – Deep learning
BER – Bit error rate
EVO – Energy valley optimization
AEVO – Adaptive Energy valley optimization
RIGM – Rotationally invariant Gaussian mixture
QAM – Quadrature amplitude modulation
PQC – Parametrized quantum circuit
MIMO – Multiple input multiple output
MUI – Multi-user interference
CNN – Convolutional neural networks
LSTM – Long short-term memory
KCS – Kronecker compressed sensing
SER – Symbol error rate
IoT – Internet of Things
ISR – Interference-to-signal ratio
GMM – Gaussian mixture model
NMSE – Normalized mean square error
IRR – Interference rejection ratio
SNR – Signal-to-Noise ratio
JUICE – Joint user-activity identification and channel estimation
JUICESD – Joint user identification, channel estimation, and signal detection

Appendix
No appendices are needed for this research.

References

[1] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. S. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," IEEE Communications Surveys and Tutorials, vol. 19, no. 2, pp. 721–742, Apr. 2017, doi: 10.1109/COMST.2016.2621116.

[2] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, 2013, pp. 611–615. doi: 10.1109/PIMRC.2013.6666209.

[3] C. Nguyen, T. M. Hoang, and A. A. Cheema, "Channel Estimation Using CNN-LSTM in RIS-NOMA Assisted 6G Network," IEEE Transactions on Machine Learning in Communications and Networking, vol. 1, pp. 43–60, May 2023, doi: 10.1109/tmlcn.2023.3278232.

[4] A. Emir, F. Kara, H. Kaya, and X. Li, "Deep learning-based flexible joint channel estimation and signal detection of multi-user OFDM-NOMA," Physical Communication, vol. 48, Oct. 2021, doi: 10.1016/j.phycom.2021.101443.

[5] B. Narottama and S. Y. Shin, "Quantum Neural Networks for Resource Allocation in Wireless Communications," IEEE Trans Wirel Commun, vol. 21, no. 2, pp. 1103–1116, Feb. 2022, doi: 10.1109/TWC.2021.3102139.

[6] M. AbdelMoniem, S. M. Gasser, M. S. El-Mahallawy, M. W. Fakhr, and A. Soliman, "Enhanced NOMA system using adaptive coding and modulation based on LSTM neural network channel estimation," Applied Sciences (Switzerland), vol. 9, no. 15, 2019, doi: 10.3390/app9153022.

[7] S. A. H. Mohsan, Y. Li, A. V. Shvetsov, J. Varela-Aldás, S. M. Mostafa, and A. Elfikky, "A Survey of Deep Learning Based NOMA: State of the Art, Key Aspects, Open Challenges and Future Trends," Mar. 01, 2023, MDPI. doi: 10.3390/s23062946.

[8] W. Yuan, N. Wu, Q. Guo, D. W. K. Ng, J. Yuan, and L. Hanzo, "Iterative Joint Channel Estimation, User Activity Tracking, and Data Detection for FTN-NOMA Systems Supporting Random Access," IEEE Transactions on Communications, vol. 68, no. 5, pp. 2963–2977, May 2020, doi: 10.1109/TCOMM.2020.2975169.

[9] S. Jiang, X. Yuan, X. Wang, C. Xu, and W. Yu, "Joint User Identification, Channel Estimation, and Signal Detection for Grant-Free NOMA," IEEE Trans Wirel Commun, vol. 19, no. 10, pp. 6960–6976, Oct. 2020, doi: 10.1109/TWC.2020.3007545.

[10] S. M. K. Sarkhi and H. Koyuncu, "Optimization Strategies for Atari Game Environments: Integrating Snake Optimization Algorithm and Energy Valley Optimization in Reinforcement Learning Models," AI (Switzerland), vol. 5, no. 3, pp. 1172–1191, Sep. 2024, doi: 10.3390/ai5030057.

[11] I. A. Mahady, E. Bedeer, S. Ikki, and H. Yanikomeroglu, "Sum-Rate Maximization of NOMA Systems under Imperfect Successive Interference Cancellation," IEEE Communications Letters, vol. 23, no. 3, pp. 474–477, Mar. 2019, doi: 10.1109/LCOMM.2019.2893195.

[12] Y. Wang, X. Zhu, E. G. Lim, Z. Wei, Y. Liu, and Y. Jiang, "Compressive Sensing based User Activity Detection and Channel Estimation in Uplink NOMA Systems," in IEEE Wireless Communications and Networking Conference, WCNC, Institute of Electrical and Electronics Engineers Inc., May 2020. doi: 10.1109/WCNC45663.2020.9120664.

[13] L. K. Smirani, L. Jamel, and L. Almuqren, "Improving Channel Estimation in a NOMA Modulation Environment Based on Ensemble Learning," CMES - Computer Modeling in Engineering and Sciences, vol. 140, no. 2, pp. 1315–1337, 2024, doi: 10.32604/cmes.2024.047551.

[14] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future," IEEE Access, vol. 7, pp. 46317–46350, 2019, doi: 10.1109/ACCESS.2019.2909490.

[15] W. Zhao et al., "Quantum Computing in Wireless Communications and Networking: A Tutorial-cum-Survey," IEEE Communications Surveys & Tutorials, vol. 27, no. 4, pp. 2378–2419, Aug. 2025, doi: 10.1109/COMST.2024.3502762.

[16] F. Li et al., "Quantum Bacterial Foraging Optimization: From Theory to MIMO System Designs," IEEE Open Journal of the Communications Society, vol. 1, pp. 1632–1646, 2020, doi: 10.1109/OJCOMS.2020.3031449.

[17] K. S. Ali, H. Elsawy, A. Chaaban, and M. S. Alouini, "Non-Orthogonal Multiple Access for Large-Scale 5G Networks: Interference Aware Design," IEEE Access, vol. 5, pp. 21204–21216, Sep. 2017, doi: 10.1109/ACCESS.2017.2753380.

[18] C. Y. Lin and W. Liao, "AoI-Aware Interference Mitigation for Task-Oriented Multicasting in Multi-Cell NOMA Networks," IEEE Trans Wirel Commun, vol. 23, no. 9, pp. 11341–11356, 2024, doi: 10.1109/TWC.2024.3381480.

[19] Y. Yan, L. Li, X. Zhang, W. Lin, W. Cheng, and Z. Han, "Adaptive Semantic Generation and NOMA-Based Interference-Aware Transmission for 6G Networks," IEEE Trans Wirel Commun, vol. 24, no. 3, pp. 2404–2416, 2025, doi: 10.1109/TWC.2024.3520870.

[20] L. Wu, P. Sun, Z. Wang, and Y. Yang, "Joint User Activity Identification and Channel Estimation for Grant-Free NOMA: A Spatial-Temporal Structure-Enhanced Approach," IEEE Internet Things J, vol. 8, no. 15, pp. 12339–12349, Aug. 2021, doi: 10.1109/JIOT.2021.3063476.

[21] A. Salari, M. Shirvanimoghaddam, M. B. Shahab, R. Arablouei, and S. Johnson, "Design and Analysis of Clustering-Based Joint Channel Estimation and Signal Detection for NOMA," IEEE Trans Veh Technol, vol. 73, no. 2, pp. 2093–2108, Feb. 2024, doi: 10.1109/TVT.2023.3313650.

[22] M. Azizi, U. Aickelin, H. A. Khorshidi, and M. Baghalzadeh Shishehgarkhaneh, "Energy valley optimizer: a novel metaheuristic algorithm for global and engineering optimization," Sci Rep, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-022-27344-y.

[23] H. Hirose, T. Ohtsuki, and G. Gui, "Deep Learning-Based Channel Estimation for Massive MIMO Systems with Pilot Contamination," IEEE Open Journal of Vehicular Technology, vol. 2, pp. 67–77, 2021, doi: 10.1109/OJVT.2020.3045470.

[24] Di. L. Galappaththige and G. Amarasuriya, "Active pilot contamination attack detection in sub-6 GHz massive MIMO NOMA systems," in Proceedings - IEEE Global Communications Conference, GLOBECOM, 2019. doi: 10.1109/GLOBECOM38437.2019.9014117.

[25] B. Akbarian and S. Ghazi-Maghrebi, "Interference cancellation for cell-edge users in massive MIMO-NOMA systems: a contemporary overview," Telecommun Syst, vol. 87, no. 4, pp. 1033–1044, Dec. 2024, doi: 10.1007/s11235-024-01219-1.

[26] X. Chen and C. Yuen, "Performance analysis and optimization for interference alignment over MIMO interference channels with limited feedback," IEEE Transactions on Signal Processing, vol. 62, no. 7, pp. 1785–1795, Apr. 2014, doi: 10.1109/TSP.2014.2304926.
[27] X.-L. Lin and M. Ng, "An all-at-once preconditioner for evolutionary partial differential equations," Mar. 2021, [Online]. Available: http://arxiv.org/abs/2002.01108
[28] K. Beer et al., "Training deep quantum neural networks," Nat Commun, vol. 11, no. 1, Dec. 2020, doi: 10.1038/s41467-020-14454-2.
[29] C. Farquhar, "QNN-Gen: A Framework for Quantum Neural Networks." Accessed: Oct. 25, 2025. [Online]. Available: https://github.com/farquhar13/QNN-Gen
[30] "PennyLaneAI / pennylane Public." [Online]. Available: https://github.com/PennyLaneAI/pennylane
[31] M. Ganji, X. Zou, and H. Jafarkhani, "Asynchronous Transmission for Multiple Access Channels: Rate-Region Analysis and System Design for Uplink NOMA," IEEE Trans Wirel Commun, vol. 20, no. 7, pp. 4364–4378, Jul. 2021, doi: 10.1109/TWC.2021.3058405.
[32] Y. Suzuki, S. Endo, K. Fujii, and Y. Tokunaga, "Quantum Error Mitigation as a Universal Error Reduction Technique: Applications from the NISQ to the Fault-Tolerant Quantum Computing Eras," PRX Quantum, vol. 3, no. 1, Mar. 2022, doi: 10.1103/PRXQuantum.3.010345.
[33] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A Comprehensive Survey of Pilot Contamination in Massive MIMO-5G System," IEEE Communications Surveys and Tutorials, vol. 18, no. 2, pp. 905–923, Apr. 2016, doi: 10.1109/COMST.2015.2504379.
[34] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel Estimation for Cell-Free mmWave Massive MIMO through Deep Learning," IEEE Trans Veh Technol, vol. 68, no. 10, pp. 10325–10329, Oct. 2019, doi: 10.1109/TVT.2019.2937543.
[35] W. Wang, "Spectrum sensing for cognitive radio," in 3rd International Symposium on Intelligent Information Technology Application Workshops, IITAW 2009, 2009, pp. 410–412. doi: 10.1109/IITAW.2009.49.
[36] H. Q. Nguyen et al., "Diffusion-inspired quantum noise mitigation in parameterized quantum circuits," Quantum Mach Intell, vol. 7, no. 1, Jun. 2025, doi: 10.1007/s42484-025-00276-2.
[37] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and G. Y. Li, "Channel estimation for OFDM," IEEE Communications Surveys and Tutorials, vol. 16, no. 4, pp. 1891–1908, Apr. 2014, doi: 10.1109/COMST.2014.2320074.
[38] Y. Chen, S. Yang, and W. S. Wong, "Exact non-gaussian interference model for fading channels," IEEE Trans Wirel Commun, vol. 12, no. 1, pp. 168–179, 2013, doi: 10.1109/TWC.2012.113012.112117.

Authors' Profiles

Avinash Ratre received his Ph.D. in Electronics and Communication Engineering from the Indian Institute of Technology, Roorkee, India. Presently, he is working as an Associate Professor in the Department of Electronics and Communication Engineering at Delhi Technological University, Delhi, India (ORCID ID: https://orcid.org/0000-0001-9803-2665). His research interests include computer vision, machine learning, signal processing, and wireless communication.

How to cite this paper: Avinash Ratre, "Interference-aware Deep Quantum Neural Network for NOMA Channel Estimation via Adaptive Energy Valley Optimization", International Journal of Computer Network and Information Security(IJCNIS), Vol.18, No.2, pp.38-60, 2026. DOI:10.5815/ijcnis.2026.02.03