Quantum Neural Networks for Optimal Resource Allocation in Cell-Free MIMO Systems

Bhaskara Narottama*, *Member, IEEE*, and Trung Q. Duong†, *Fellow, IEEE*
*Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
†School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, UK
e-mail: *bhaskara@kumoh.ac.kr,†trung.q.duong@qub.ac.uk

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10001726

***Abstract*—In this paper, the potential benefit of employing quantum neural networks (QNNs) for cell-free MIMO is explored. In particular, QNN-based scheme are used to optimize transmitter-user assignment in cell-free MIMO. The objective of the optimization is to maximize the minimum achieved sum rate. Although QNN has received increasing research attention owing to the potential benefit of quantum computation, its utilization for multi-transmitter scenario is still limited. As such, in this paper, we consider the QNN-based algorithm for optimal resourcea allocation in cell-free MIMO systems. We also demonstrate the advantage of our proposed QNN-based algorithm through the numerical results.**

***Index Terms*—online learning, quantum neural networks, 6G, cell-free MIMO, wireless communications.**

I. Introduction

Cell-free multiple-input and multiple-output (MIMO), which can be operated without cell boundaries, can extends the coverage of the "conventional" MIMO network [1]. Optimization is critical for the performance of cell-free MIMO in 6G [2], [3]. However, as the scale of the cell-free MIMO network increases, it will be more challenging to provide mathematical-based optimal solutions [1].

To cope with this issue, neural network (NN)-based algorithm has been utilized to provide estimations for cell-free MIMO optimization [3]. In general, the complexity of the classical-bit-based NNs increases with the number of layers and neurons [4]. Motivated by its potentials such as higher performance [5] and low complexity [6], a quantum neural network (QNN)-based scheme can be employed to optimize the performance of cell-free MIMO systems [7]. Different from classical-based NN, QNN is a generalized model that consists of quantum gates and operations [5], [8]. In particular, parameterized gates can be used to tune the QNN. Very recently, quantum-inspired optimisation algorithm has been considered as promising candidate for optimal resource allocation of future networks [9].

The related works are summarized in Table I. In [3], NN-based scheme is used to optimize transmitter-user assignment. However, quantum-based neural networks was not investigated. In [10], ML-based optimization for cell-free MIMO is presented without the consideration of quantum-based ML. Very recently, in [11], quantum-based ML algorithm is employed for wireless optimization. In [6], QNN schemes are utilized for user grouping in non-orthogonal multiple access (NOMA). As presented in Table I, to the best of the authors' knowledge, the utilization of QNN for cell-free MIMO is still limited. Recent studies shows that quantum neural network (QNN), by taking the advantage of quantum mechanics such as quantum *superposition* and *entanglement*, has the potential to increase the training capability and reduce the training time [5], [8]. As such, in this paper, our attempt is to employ QNN-algorithm for optimal the resource allocation of cell-free MIMO system. We also provide the analysis of upper and lower bound for the considered QNN operation. Finally, the performance of the proposed scheme is investigated via simulation results.

Table I: Related Works

Ref.	QNN	Cell-Free MIMO	Takeaways
[3]	-	✓	NN-based scheme for transmitter-user assignment
[10]	-	✓	ML-based scheme for cell-free MIMO optimization
[11]	✓	-	employs quantum ML for wireless optimization
[6]	✓	-	QNN schemes for wireless optimization
this work	✓	✓	QNN for cell-free MIMO

The remainder of this work can be organised as follows. The assumed system model is presented in Section II. The proposed QNN-based scheme is described in Section III. The numerical result is analysed in Section IV. Finally, this study is concluded in Section V.

Notation: Frobenius norm and absolute value are presented as $\|\cdot\|_F$ and $|\cdot|$, respectively. $\mathbb{R}$ and $\mathbb{C}$ indicate real and complex numbers, respectively. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^\dagger$ indicate conjugate, transpose, and conjugate transpose operations, respectively. $\mathcal{CN}(\mu,\sigma^2)$ and $\mathcal{N}(\mu,\sigma^2)$ are complex and real -valued normal distribution, respectively, with mean μ and variance σ^2. $\mathbf{H}$, $\mathbf{C}_{\rm X}$, $\mathbf{C}_{\rm Z}$, and $\mathbf{R}_{\rm Z}$ are the Hadamard, controlled-X, controlled-Y, and rotate-Z quantum-based gates. Pauli rotation gates are indicated as σ_x, σ_y, and σ_z.

II. System Model

We consider a cell-free MIMO system consisting of $N_{\rm AP}$ transmitters serving as access points and $N_{\rm user}$ users. Let $\mathcal{A}=\{a_m\}_{m=1}^{N_{\rm AP}}$ and $\mathcal{U}=\{u_k\}_{k=1}^{N_{\rm user}}$ be the sets of available transmitters and users, respectively. The distance between mth transmitter and kth user is denoted as $d_{m,k}$. Each transmitter employs $N_{\rm Tx}$ antenna arrays.

Moreover, the imperfect channel information can be presented as $\hat{\boldsymbol{H}} = \boldsymbol{H} + \boldsymbol{N}_{\text{CSI}}$, where $\boldsymbol{N}_{\text{CSI}}$ indicates the imperfection in channel information [12]. The channel coefficients between a_m and u_k are presented as

$$\begin{aligned} \boldsymbol{h}_{m,k} &= \left\{ h_{m,k}^{[1]}, \ldots, h_{m,k}^{[N_{\text{Tx}}]} \right\}^T \in \mathbb{C}^{N_{\text{Tx}}}, \\ h_{m,k}^{[j]} &= \frac{1}{\sqrt{N_{\text{path}}}} \sum_{n=1}^{N_{\text{path}}} g_{m,k}^{[j]} a(\phi_n), \end{aligned} \tag{1}$$

where $g_{m,k}^{[j]}$ and $a(\phi_n)$ are the Rayleigh fading channel gain and steering variable (ϕ_n as antenna response angle), respectively. The Rayleigh channel can be presented as $g_{m,k}^{[j]} \sim \mathcal{CN}(0, d_{m,k}^{-\kappa})$, where κ is the pathloss exponent [13]. The steering variable can be presented as [14]

$$\begin{aligned} a(\phi_n) &= \{\exp(-\mathbf{i}2\pi\phi_n z)\}_{z\in\mathcal{Z}}, \\ \mathcal{Z} &= \{n - 0.5 \cdot (N_{\text{Tx}} - 1)\}_{n\in\{1,\ldots,N_{\text{Tx}}-1\}}. \end{aligned} \tag{2}$$

Let v_m be the precoding vector of the a_m. For initial assumption, each mth transmitter can only produce one directional beam, although multiple transmitters can cooperate to serve a user. To enable cooperation between transmitters, let us assume an ultra dense network where $N_{\text{AP}} \gg N_{\text{user}}$.

A. Achievable Rate

Let P_t be the total transmit power of each transmitter and σ^2 be the noise variance. Considering similar P_t for each transmitter, the achievable rate for a particular user u_k served by a single transmitter a_m can be expressed as

$$R_{m\to k} = \log_2\left(1 + \frac{P_t |\boldsymbol{h}_{m,k}^T v_m|^2}{\underbrace{P_t \sum_{n=1, n\neq m}^{N_{\text{AP}}} \mu_{n,k} |\boldsymbol{h}_{n,k}^T v_n|^2}_{\text{inter-transmitter interference}} + \sigma^2}\right), \tag{3}$$

where $\mu_{n,k} < 1$ is the interference factor from transmitter a_n to the user u_k. Now, let us consider that u_k can be served by multiple transmitters, which can be grouped as $\mathcal{A}_k$. Moreover, the transmit signal-to-noise ratio can be presented as $\rho = {}^{P_t}/_{\sigma^2}$, Given that the transmitters in $\mathcal{A}_k$ employs same resource block, the achievable rate of u_k can be expressed as

$$R_k = \log_2\left(1 + \frac{\sum_{m\in\mathcal{A}_k} \rho |\boldsymbol{h}_{m,k}^T v_m|^2}{\underbrace{\rho \sum_{n\in\mathcal{A}, n\notin\mathcal{A}_k} \mu_{n,k} |\boldsymbol{h}_{n,k}^T v_n|^2}_{\text{inter-transmitter interference}} + 1}\right). \tag{4}$$

Finally, the achieved sum rate of N_{user} users can be expressed as

$$R_{\text{sum}} = \sum_{k=1}^{N_{\text{user}}} R_k. \tag{5}$$

B. Objective

Let R_k be the achievable rate of the kth user. The objective function of the optimization problem, which is to maximize the lowest rate among the users,[1] can be expressed as:

$$\underset{\Phi}{\text{maximize}} \quad \min_{\{u_k\}_{k=1}^{N_{\text{user}}}} R_k(\Phi) \tag{6a}$$

$$\text{subject to} \quad \|v_m\|_F^2 \leq 1 \tag{6b}$$

where $\Phi = \left\{\gamma, \{v_m\}_{m=1}^{N_{\text{AP}}}\right\}$ is the tuple of solution which contains assigment policy γ and precoding $\{v_m\}_{m=1}^{N_{\text{AP}}}$. Here, the max-min optimization is considered to ensure fair rate among users (similar approach can be found in [14], [15]).

III. Proposed QNN-Based Scheme

A. Transmitter-User Assignment

The transmitter-user assigment problem can be described as follows. As presented in Fig. 1, the transmitter-user assignment can be presented as a bipartite graph. Similar to Eq. (6), the objective of this phase is to maximize the achived rate of the user with minimum rate, which can be expressed as:

$$\underset{\gamma}{\text{maximize}} \quad \min_{\{u_k\}_{k=1}^{N_{\text{user}}}} R_k(\gamma|v_m) \tag{7a}$$

$$\text{subject to} \quad \|v_m\|_F^2 \leq 1. \tag{7b}$$

Additionally, the maximum-ratio-based precoding, i.e., $v_m^{\text{MR}} = \frac{h_m^*}{\|h_m^*\|_F}$, $\forall m \in \{1, \ldots, N_{\text{AP}}\}$ can be considered for this phase [17]. Based on the objective in Eq. (7), the reward can be formulated as[2]

$$Q_{\text{assign}}\left(\gamma|\hat{\boldsymbol{H}}, v_m^{\text{MR}}\right) = -\min_{\{u_k\}_{k=1}^{N_{\text{user}}}} R_k\left(\gamma|v_m^{\text{MR}}\right). \tag{8}$$

The quantum-based scheme is presented in Fig. 2. For all kth user, given channel matrix $\hat{\boldsymbol{H}}$ as input feature, the assignment policy output can be expressed as[3]

$$\pi_{\text{assign}} : \hat{\boldsymbol{H}} \xrightarrow{U_{\text{QNN}}^{\text{cloud}}} \gamma, \tag{9}$$

[1]Here, max-min optimization is used to guarantee the fairness among the users by avoiding transmitters allocated only to the users with a higher channel gain.

[2]Considering r_{penalty} as penalty variable, let us assume $R_k\left(\gamma|v_m^{\text{MR}}\right) = -r_{\text{penalty}}$ is assumed instead of $R_k\left(\gamma|v_m^{\text{MR}}\right) = 0$ to introduce negative reward if no AP is assigned to kth user.

[3]For ease of implementation, the quantum-based computations can be performed via IBM Qiskit [18].

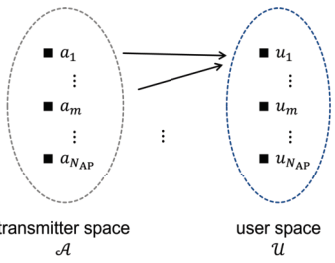

Figure 1: Transmitter-user assignment, presented as bipartite graph. Moreover, many-to-one relationship can be assumed in this study for the sake of simplicity [16].

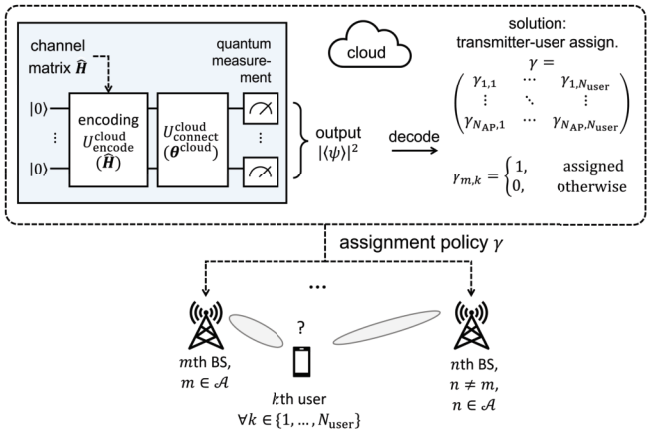

Figure 2: QNN-based optimization for transmitter-user assignment in cell-free MIMO.

where $U^{\text{cloud}}_{\text{QNN}}$ is the QNN operation at the cloud. The QNN outputs can be decoded as a feature matrix representing the transmitter-user assignment, which can be expressed as

$$\gamma \triangleq \begin{pmatrix} \gamma_{1,1} & \cdots & \gamma_{1,N_{\text{user}}} \\ \vdots & \ddots & \vdots \\ \gamma_{N_{\text{AP}},1} & \cdots & \gamma_{N_{\text{AP}},N_{\text{user}}} \end{pmatrix} \subset \{0,1\}^{N_{\text{AP}} \times N_{\text{user}}}, \quad (10)$$

where $\gamma_{m,k} = 1$ indicating that mth transmitter is assigned to kth user. The designated QNN operation for this task can be defined as follows.

The loss can be defined as the difference between the acquired reward and "ideal" performance, which can be expressed as

$$\text{L}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) = \left\| Q_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}, v_m^{\text{MR}}) - \Phi_{\text{assign}}(\hat{\boldsymbol{H}}) \right\|^2, \quad (11)$$

where the "ideal" performance can be expressed as[4]

$$\Phi_{\text{assign}}(\hat{\boldsymbol{H}}) = -\sum_{m=1}^{N_{\text{AP}}} \sum_{i=1}^{N_\lambda^{[m]}} \log_2 \left(1 + \frac{\lambda_i^{[m]}}{N_\lambda^{[m]}} \rho\right), \quad (12)$$

where $\{\lambda_i^{[m]}\}_{i=1}^{N_{\text{Tx}}}$ indicates the array of eigenvalues of $\boldsymbol{h}_{m,k}$ [19], [20].

Definition 1 (QNN operation − cloud)**.** The QNN operation in the cloud can expressed as

$$U^{\text{cloud}} = U^{\text{cloud}}_{\text{connect}}(\theta^{\text{cloud}}) U^{\text{cloud}}_{\text{encode}}(\hat{\boldsymbol{H}}) \, |0\rangle^{\otimes N_{\text{layer}} \times N_{\text{neuron}}}, \quad (13)$$

where $U^{\text{cloud}}_{\text{encode}}$ and $U^{\text{cloud}}_{\text{connect}}$ are the encoding and connection operations, respectively. Moreover, $|0\rangle^{\otimes N_{\text{layer}} \times N_{\text{neuron}}}$ indicates qubits preparation, where N_{layer} and N_{neuron} are the number of layers and neurons, respectively. Let θ^{cloud} be the parameter vector for the cloud. Based on [21], the encoding operation can be expressed as

$$U^{\text{cloud}}_{\text{encode}}(\hat{\boldsymbol{H}}) \triangleq \bigotimes_{m=1}^{N_{\text{AP}}} \Big(\underbrace{\prod_{k=1}^{N_{\text{user}}} \mathbf{R}_{\text{Z}}\big(\hat{\boldsymbol{h}}_{m,k}; \big|q_{1,n}^{\text{cloud}}\big\rangle\big) \mathbf{H}\big(\big|q_{1,n}^{\text{cloud}}\big\rangle\big)}_{\text{for } a_m} \Big), \quad (14)$$

where $\left|q_{l,n}^{\text{cloud}}\right\rangle$ indicates the assigned qubit, $l \in \{1, \ldots, N_{\text{layer}}\}$, $n \in \{1, \ldots, N_{\text{neuron},l}\}$. The main QNN operation can be expressed as [22]

$$\begin{aligned} U^{\text{cloud}}_{\text{connect}}(\theta^{\text{cloud}}) \triangleq & \bigotimes_{l=1}^{N^{\text{cloud}}_{\text{layer}}-1} \prod_{n=1}^{N^{\text{cloud}}_{\text{neuron},l}-1} \underbrace{\mathbf{C}_{\text{Z}}\big(\big|q_{l,n+1}^{\text{cloud}}\big\rangle; \big|q_{l,n}^{\text{cloud}}\big\rangle\big)}_{\text{connect neuron}} \\ & \bigotimes_{l=1}^{N^{\text{cloud}}_{\text{layer}}-1} \underbrace{\mathbf{C}_{\text{X}}\big(\big|q_{l+1,1}^{\text{cloud}}\big\rangle; \left|q_{l,N^{\text{cloud}}_{\text{neuron},l}}^{\text{cloud}}\right\rangle\big)}_{\text{connect layer}} \\ & \bigotimes_{l=1}^{N^{\text{cloud}}_{\text{layer}}} \bigotimes_{n=1}^{N^{\text{cloud}}_{\text{neuron},l}} \mathbf{R}_{\text{y}}\big(\theta_{l,n}^{\text{cloud}}\big), \end{aligned} \quad (15)$$

where $\theta_{l,n}^{\text{cloud}} \in \theta^{\text{cloud}}$ is the weight parameter for the nth neuron, lth layer.

After the measurement, we can obtain the QNN output as $o_m^{\text{cloud}} \in [0,1]$, $\forall m \in \{1, \ldots, N_{\text{AP}}\}$. Subsequently, the solution output for mth transmitter is given as $\hat{o}_m = \lfloor o_m^{\text{cloud}} N_{\text{user}} \rceil$.[5]

B. Algorithm

The proposed transmitter-user assignment scheme can be formulated as in Algorithm 1. The number of data training samples and episodes are denoted as N_{data} and N_{epoch}. As Algorithm 1 employs stochastic gradient descent, the number of training iterations is given as $N_{\text{train}} = N_{\text{data}} N_{\text{epoch}}$. The step size is denoted as μ. The channel information for training is presented as $\hat{\boldsymbol{H}}_{\text{train}}^{[i]}$.

C. Complexity analysis and required qubits

To facilitate the analysis, the complexities of Algorithm 1 is investigated based on the number of gradient calculation. Considering a single gradient calculation as $\mathcal{G}$, the complexities of the Algorithm 1 can be analyzed as $N_{\text{iter}}\mathcal{G}$.

Moreover, the number of required qubits for the considered QNNs can be analyzed as follows:

Lemma 1. *The operation of $U^{cloud}_{connect}$ and $U^{cloud}_{connect}$ require $\big((N^{cloud}_{layer} - 1) + (N^{cloud}_{layer} - 1) + N^{cloud}_{layer} N^{cloud}_{neuron}\big)$, respectively.*

Proof. As expressed in Eq. (15), $U^{\text{cloud}}_{\text{connect}}$ employs $N^{\text{cloud}}_{\text{layer}} - 1$ qubits to connect layers and $N^{\text{cloud}}_{\text{neuron}} - 1$ to connect neurons. Moreover, $N^{\text{cloud}}_{\text{layer}} N^{\text{cloud}}_{\text{neuron}}$ qubits are needed for weight parameters. In addition, the encoding operation $U^{\text{cloud}}_{\text{encode}}$ requires N_{AP}. Therefore, $U^{\text{cloud}}_{\text{connect}}$ employs $\big((N^{\text{cloud}}_{\text{layer}} - 1) + (N^{\text{cloud}}_{\text{layer}} - 1) + N^{\text{cloud}}_{\text{layer}} N^{\text{cloud}}_{\text{neuron}}\big)$ qubits. ■

[4]Hereafter, $\Phi_{\text{assign}}(\hat{\boldsymbol{H}})$ is used as ideal performance as each transmission channel $j, j \in \{1, \ldots, N_{\text{Tx}}\}$, is orthogonal to each other. Moreover, imperfect channel $\hat{\boldsymbol{H}}$ is assumed here. Hence, we can assume $\Phi_{\text{assign}}(\hat{\boldsymbol{H}})$ as the *ideal* capacity given imperfect channel condition.

[5]As example, $\hat{o}_1 = 2$ indicates that the second transmitter serves the first user.

Algorithm 1 Transmitter-User Assignment

Input: acquired channel information $\left\{\hat{\boldsymbol{H}}_{\text{train}}^{[i]}\right\}_{i=1}^{N_{\text{data}}}$
Output: assignment policy γ

Initialization:
1: Define available transmitters and users as $\mathcal{A}$ and $\mathcal{U}$, respectively.
2: Prepare the qubits (state $|0\rangle$ for all qubits).
Training:
3: **for** for $i_{\text{train}} \in \{1, \ldots, N_{\text{train}}\}$ **do**
4: Considering $\hat{\boldsymbol{H}}_{\text{train}}^{[i]}$ as input, employs U^{cloud} in Eq. (13).
5: Decode the output of U^{cloud}, obtain assignment policy γ in Eq. (9).
6: Obtain $Q_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}_{\text{train}}^{[i]}, v_m^{\text{MR}})$ and $\Phi_{\text{assign}}(\hat{\boldsymbol{H}}_{\text{train}}^{[i]})$.
7: Considering the loss $L_{\text{assign}}(\hat{\boldsymbol{H}}_{\text{train}}^{[i]})$ in Eq. (11), calculate $\nabla L_{\text{assign}}(\hat{\boldsymbol{H}}_{\text{train}}^{[i]})$ (Section A). Update parameter as $\theta^{\text{cloud}} \leftarrow \theta^{\text{cloud}} - \mu \nabla L_{\text{assign}}(\hat{\boldsymbol{H}}_{\text{train}}^{[i]})$. Rotosolve can also be employed (Section A).
8: **end for**
Deployment:
9: Considering instantaneous channel information $\hat{\boldsymbol{H}}$ as input, employs U^{cloud} in Eq. (13).
10: Decode the output of U^{cloud} to obtain assignment policy γ in Eq. (9).

Furthermore, the trainability of the cloud QNN can be presented through the lower and upper bounds of the norm of the gradient for the cloud QNN as

$$\frac{1}{8}\left\{ \text{Tr}[x_{\text{input}}]^2 + \text{Tr}[\sigma_z x_{\text{input}} \sigma_z]^2 \right\} \leq \mathbb{E}_{\theta^{\text{cloud}}}\left(\left\| \nabla_{\theta^{\text{cloud}}} \hat{\text{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) \right\|^2 \right) \leq N_{\text{layer}} N_{\text{neuron}}. \tag{16}$$

The proof can be found in Appendix B.

IV. Result

In this section, the performance of the proposed method is investigated. Unless otherwise stated, the simulation and training parameters are presented in Table II.[6] First, Algorithm 1 is employed to estimate the solution for transmitter-user assignment (Section III-A). Subsequently, the achieved rates of the users were measured. During the training, a descending learning rate was assumed. In particular, the learning step at the i_{train}th training iteration can be expressed as $\mu_{i_{\text{train}}} = \frac{\mu}{\sqrt{i_{\text{episode}}+1}}$.

[6]As current generation of quantum devices only employs a limited number of qubits [23], here this study only employs consider a handful of network elements. However, QNN employment to optimize a large-scale network can be considered for future work.

Table II: Parameters

Parameter	Value	Parameter	Value
$d_{m,k}$	0.5	μ	0.01
$\mu_{n,k}$	0.1	N_{epoch}	100
N_{AP}	4	r_{penalty}	-10
N_{Tx}	2	ϱ	$\pi/2$
N_{user}	3		

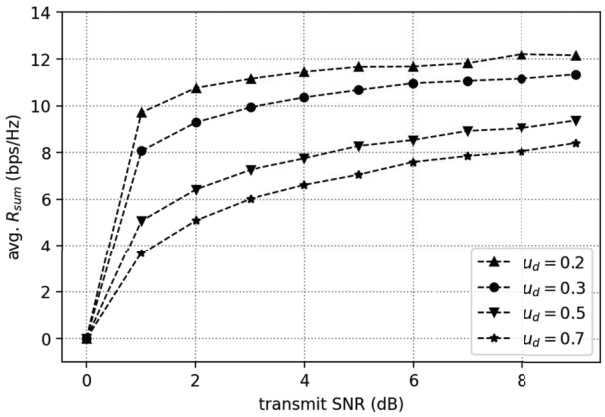

Figure 3: The achievable sum rate of transmitter-user assignment problem.

Considering Monte-Carlo simulation of 10^3 trials, the achievable sum rate of transmitter-user assignment problem, as showned in Eq. (5)), is illustrated in Fig. 3. As can be seen from this figure, the achieved sum rate increases with the received SINR ρ. Accordingly, the stepness of the figure is also decreased with respect to ρ due to the inter-transmitter interference, as described in Eq. (4). During the simulation, the user distance was obtained as $d \sim \mathcal{N}(\mu_d, \sigma_d^2)$, where $\sigma_d^2 = 0.1$ was assumed as an arbitrary value. Moreover, different values of μ_d, $\mu_d \in \{0.1, 0.3, 0.5, 0.7\}$, were also considered. In addition, antenna response ϕ_n was obtained uniformly between $-\pi/2$ and $\pi/2$. The loss during training in the cloud (Eq. (11)) is exhibited in Fig. 4. As can clearly be observed from the figure, the loss decreases with training episode. Moreover, the loss is converged at $\sim$ 30th training episode.

V. Conclusion

This study has explored the possibilities to leverage quantum-based machine learning for radio resource optimization of cell-

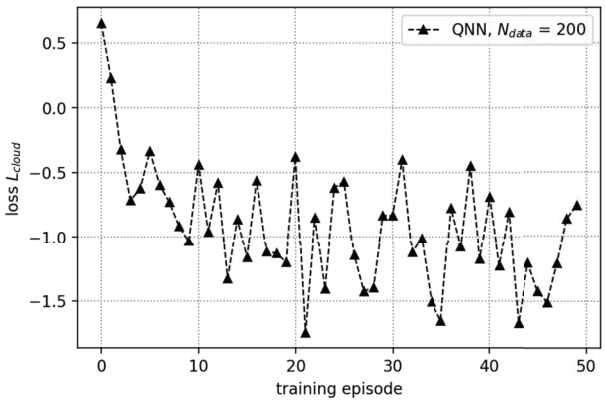

Figure 4: The training loss for the transmitter-user assignment.

free MIMO systems. Specifically, QNN-based optimization algorithm has been utilized for transmitter-user assignment problem. This study has investigated the trainability of the presented QNN model. Moreover, the required resources in terms of the number of required qubits has been also presented. Finally, numerical result of cell-free MIMO using the proposed QNN framework has been presented.

For further development, user deployment as relay to cover far users can be considered [24]. Moreover, quantum-based reinforcement learning can also be employed [11].

ACKNOWLEDGEMENT

The authors acknowledge the access of IBM Q quantum computer via IBM Qiskit in this work. IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

APPENDIX A
PARAMETER OPTIMIZATION

The gradient of parameter θ given $L \in \{L_{\text{assign}}, L_{\text{precode}}\}$, which is calculated by leveraging parameter-shift rule [25], is given as

$$\nabla_\theta L(\theta) = {}^1\!/_2 \sin(\varrho)\big(L(\theta-\varrho) - L(\theta+\varrho)\big), \tag{17}$$

where ϱ is the shift parameter.

APPENDIX B
BOUNDS OF THE DERIVATION

The bounds for transmitter-user assignment can be presented as follows. To simplify the analysis, let us consider the normalized loss as

$$\hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) = \left\| \frac{Q_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}, v_m^{\text{MR}}) - \Phi_{\text{assign}}(\hat{\boldsymbol{H}})}{\Phi_{\text{assign}}(\hat{\boldsymbol{H}})} \right\|^2, \tag{18}$$

so that $\hat{\mathrm{L}}_{\text{assign}} \in [0,1]$.

Lemma 2. *The bounds of the gradient norm for the cloud QNN can be presented as*

$$\begin{aligned} &\frac{1}{8}\Big\{ \mathrm{Tr}[x_{input}]^2 + \mathrm{Tr}[\sigma_z x_{input} \sigma_z]^2 \Big\} \\ &\quad \le \mathbb{E}_{\theta^{cloud}}\left(\left\| \nabla_{\theta^{cloud}} \hat{\mathrm{L}}_{assign}(\gamma|\hat{\boldsymbol{H}}) \right\|^2 \right) \le N_{layer} N_{neuron}. \end{aligned} \tag{19}$$

Proof. The upper bound can be analyzed as follows. Considering Eq. (18), we can infer that $\left(\nabla_{\theta^{\text{cloud}}_{l,n}} \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})\right)^2 \le 1$, where $\nabla_{\theta^{\text{cloud}}_{l,n}} \triangleq \frac{\delta \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})}{\delta \theta^{\text{cloud}}_{l,n}}$ [26]. Moreover, the norm of gradient can be expressed as

$$\left\| \nabla_{\theta^{\text{cloud}}} \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) \right\|^2 = \sum_{l=1}^{N_{\text{layer}}} \sum_{n=1}^{N_{\text{neuron}}} \left(\frac{\delta \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})}{\delta \theta^{\text{cloud}}_{l,n}} \right)^2, \tag{20}$$

so that

$$\left\| \nabla_{\theta^{\text{cloud}}} \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) \right\|^2 \le N_{\text{layer}} N_{\text{neuron}}, \tag{21}$$

where $N_{\text{neuron}} = N_{\text{AP}}$ can be considered.

The lower bound can be investigated as follows. the expectation of gradient norm with respect to parameter vector $\theta^{\text{cloud}} \in [0, 2\pi]$ is given as

$$\begin{aligned} &\mathbb{E}_{\theta^{\text{cloud}}}\left(\left\| \nabla_{\theta^{\text{cloud}}} \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}}) \right\|^2 \right) \\ &\quad = \sum_{l=1}^{N_{\text{layer}}} \sum_{n=1}^{N_{\text{neuron}}} \mathbb{E}_{\theta^{\text{cloud}}}\left(\left(\frac{\delta \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})}{\delta \theta^{\text{cloud}}_{l,n}} \right)^2 \right) \\ &\quad \ge \sum_{n=1}^{N_{\text{neuron}}} \mathbb{E}_{\theta^{\text{cloud}}}\left(\left(\frac{\delta \hat{\mathrm{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})}{\delta \theta^{\text{cloud}}_{l,1}} \right)^2 \right). \end{aligned} \tag{22}$$

where $m_{\text{input}} = |x_{\text{input}}\rangle \langle x_{\text{input}}|$ is the operator for input x_{input}. Here, the expectation of gradient norm of the first layer is considered as one-layer operation requires a lowest possible amount of qubits (only one qubit required) [26]. The function f^{cloud} is given as

$$\begin{aligned} f^{\text{cloud}} &= {}^1\!/_2 - {}^1\!/_2 \,\mathrm{Tr}\left[\sigma_{(3,3,\ldots,3)} \cdot U^{\text{cloud}} \cdot m_{\text{input}} \cdot U^{\text{cloud}\dagger} \right], \\ &= {}^1\!/_2 - {}^1\!/_2 \,\mathrm{Tr}\Big[\sigma_{(3,3,\ldots,3)} \cdot U^{\text{cloud}}_{\text{connect},N_{\text{layer}}+1} U^{\text{cloud}}_{\text{weight},N_{\text{layer}}} \cdots \\ &\quad U^{\text{cloud}}_{\text{weight},1} U^{\text{cloud}}_{\text{connect},1} \cdot m_{\text{input}} \cdot U^{\text{cloud}}_{\text{connect},1}{}^{\dagger} U^{\text{cloud}}_{\text{weight},1}{}^{\dagger} \cdots \\ &\quad U^{\text{cloud}}_{\text{weight},N_{\text{layer}}}{}^{\dagger} U^{\text{cloud}}_{\text{connect},N_{\text{layer}}+1}{}^{\dagger} \Big]. \end{aligned} \tag{23}$$

Subsequently, the partial derivation w.r.t. $\theta^{\text{cloud}}_{l,1}$ may be obtained as

$$\begin{aligned} D_l \triangleq \frac{\delta f^{\text{cloud}}}{\delta \theta^{\text{cloud}}_{l,1}} &= \mathrm{Tr}\Big[\sigma_{(3,3,\ldots,3)} \cdot U^{\text{cloud}}_{\text{connect},N_{\text{layer}}+1} \cdots \\ &\quad W_l U^{\text{cloud}}_{\text{connect},l} \cdots \cdot m_{\text{input}} \cdot \ldots U^{\text{cloud}}_{\text{connect},l}{}^{\dagger} U^{\text{cloud}}_{\text{weight},l}{}^{\dagger} \cdots \\ &\quad U^{\text{cloud}}_{\text{connect},N_{\text{layer}}+1}{}^{\dagger} \Big], \end{aligned} \tag{24}$$

where considering parameter shift rule, the following arbitrary variable W_l is given

$$W_l \triangleq \frac{U^{\text{cloud}}_{\text{weight},l}}{\delta \theta^{\text{cloud}}_{l,1}} = {}^1\!/_2\big(U^{\text{cloud}}_{\text{weight},l}(\theta^{\text{cloud}}_{l,1} - \varrho) - U^{\text{cloud}}_{\text{weight},l}(\theta^{\text{cloud}}_{l,1} + \varrho) \big), \tag{25}$$

where $\varrho = \pi/4$ is the shifting parameter.

At the minimum, the cloud QNN may served a single edge ($N_{\text{neuron}} = 1$) employing a single layer ($N_{\text{layer}} = 1$). For this scenario, the loss may be expressed as

$$\begin{aligned} f^{\text{cloud}}_{1,1} = \\ {}^1\!/_2 - {}^1\!/_2 \,\mathrm{Tr}\left[\sigma_{(z)} \cdot \mathrm{R}_y(\theta^{\text{cloud}}_{1,1}) \cdot m_{\text{input}} \cdot \mathrm{R}_y(\theta^{\text{cloud}}_{1,1})^{\dagger} \right]. \end{aligned} \tag{26}$$

The partial derivation of $f^{\text{cloud}}_{1,1}$ with respect to $\theta^{\text{cloud}}_{1,1}$ may be expressed as

$$\frac{\delta f^{\text{cloud}}_{1,1}}{\delta \theta^{\text{cloud}}_{1,1}} = \mathrm{Tr}\left[\sigma_{(z)} \cdot \frac{\delta \mathrm{R}_y(\theta^{\text{cloud}}_{1,1})}{\delta \theta^{\text{cloud}}_{1,1}} \cdot m_{\text{input}} \cdot \mathrm{R}_y(\theta^{\text{cloud}}_{1,1})^{\dagger} \right]. \tag{27}$$

Subsequently, the expectation of the norm of the previous loss can be expressed as

$$\begin{aligned}\mathbb{E}_{\theta}\left(\left(\frac{\delta f_{1,1}^{\text{cloud}}}{\delta\theta_{1,1}^{\text{cloud}}}\right)^2\right) &\triangleq \mathbb{E}_{\theta_{1,1}^{\text{cloud}}}\left(\left(\frac{\delta f_{1,1}^{\text{cloud}}}{\delta\theta_{1,1}^{\text{cloud}}}\right)^2\right)\\ &= \frac{1}{2\pi}\int_0^{2\pi} A\, d\theta_{1,1}^{\text{cloud}},\end{aligned} \tag{28}$$

where

$$A = \text{Tr}\left[\sigma_{(z)}\cdot\frac{\delta \text{R}_y(\theta_{1,1}^{\text{cloud}})}{\delta\theta_{1,1}^{\text{cloud}}}\cdot m_{\text{input}}\cdot \text{R}_y(\theta_{1,1}^{\text{cloud}})^{\dagger}\right]^2. \tag{29}$$

Considering $e^{-i\theta\sigma_i} = I\cos(\theta) - i\sigma_i\sin(\theta)$, the previous expresion can be presented as

$$\begin{aligned}\mathbb{E}_{\theta}\left(\frac{\delta f_{1,1}^{\text{cloud}}}{\delta\theta_{1,1}^{\text{cloud}}}\right) &= \frac{1}{2\pi}\int_0^{2\pi} d\theta_{1,1}^{\text{cloud}} A\\ &= \frac{1}{2\pi}\int_0^{2\pi} d\theta_{1,1}^{\text{cloud}}(\cos^2(\theta))^2\,\text{Tr}[m_{\text{input}}]^2\\ &- \frac{1}{2\pi}\int_0^{2\pi} d\theta_{1,1}^{\text{cloud}}(\sin(\theta)\cos(\theta))^2\,\text{Tr}[\sigma_z m_{\text{input}}]^2\\ &+ \frac{1}{2\pi}\int_0^{2\pi} d\theta_{1,1}^{\text{cloud}}(\sin(\theta)\cos(\theta))^2\,\text{Tr}[\sigma_z m_{\text{input}}]^2\\ &+ \frac{1}{2\pi}\int_0^{2\pi} d\theta_{1,1}^{\text{cloud}}(\sin^2(\theta))^2\,\text{Tr}[\sigma_z x_{\text{input}}\sigma_z]^2\\ &= \frac{1}{8}\,\text{Tr}[x_{\text{input}}]^2 + \frac{1}{8}\,\text{Tr}[\sigma_z m_{\text{input}}\sigma_z]^2.\end{aligned} \tag{30}$$

Therefore, the lower bound can be expressed as

$$\begin{aligned}&\mathbb{E}_{\theta^{\text{cloud}}}\left(\left\|\nabla_{\theta^{\text{cloud}}}\hat{\text{L}}_{\text{assign}}(\gamma|\hat{\boldsymbol{H}})\right\|^2\right)\\ &\qquad\geq \frac{1}{8}\Big\{\,\text{Tr}[m_{\text{input}}]^2 + \text{Tr}[\sigma_z m_{\text{input}}\sigma_z]^2\Big\}.\end{aligned} \tag{31}$$

■

REFERENCES

[1] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[2] H. Kim, T. Cho, J. Lee, W. Shin, and H. V. Poor, "Optimized shallow neural networks for sum-rate maximization in energy harvesting downlink multiuser NOMA systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 982–997, Apr. 2021.

[3] Y. Al-Eryani, M. Akrout, and E. Hossain, "Multiple access in cell-free networks: Outage performance, dynamic clustering, and deep reinforcement learning-based design," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1028–1042, Apr. 2021.

[4] M. Bianchini and F. Scarselli, "On the complexity of neural network classifiers: A comparison between shallow and deep architectures," *IEEE Trans. Neural Netw. Learning Sys.*, vol. 25, no. 8, pp. 1553–1565, Aug. 2014.

[5] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, "The power of quantum neural networks," *Nat. Comput. Sci.*, vol. 1, no. 6, pp. 403–409, Jun. 2021.

[6] B. Narottama and S. Y. Shin, "Quantum neural networks for resource allocation in wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1103–1116, Feb. 2022.

[7] 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. 46 317–46 350, Apr. 2019.

[8] N. H. Nguyen, E. C. Behrman, M. A. Moustafa, and J. E. Steck, "Benchmarking neural networks for quantum computations," *IEEE Trans. Neural Netw. Learning Sys.*, vol. 31, no. 7, pp. 2522–2531, Jul. 2020.

[9] T. Q. Duong, L. D. Nguyen, B. Narottama, J. A. Ansere, D. V. Huynh, and H. Shin, "Quantum-inspired real-time optimisation for 6G networks: Opportunities, challenges, and the road ahead," *IEEE Open J. Commun. Soc.*, pp. 1–1, Aug. 2022.

[10] T. T. Vu, D. T. Ngo, N. H. Tran, H. Q. Ngo, M. N. Dao, and R. H. Middleton, "Cell-free massive MIMO for wireless federated learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6377–6392, Oct. 2020.

[11] A. Jaiswal, S. Kumar, O. Kaiwartya, P. K. Kashyap, E. Kanjo, N. Kumar, and H. Song, "Quantum learning-enabled green communication for next-generation wireless systems," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 3, pp. 1015–1028, Sep. 2021.

[12] J. Zhao, X. Yue, S. Kang, and W. Tang, "Joint effects of imperfect CSI and SIC on NOMA based satellite-terrestrial systems," *IEEE Access*, vol. 9, pp. 12 545–12 554, Jan. 2021.

[13] M. B. Shahab, M. F. Kader, and S. Y. Shin, "A virtual user pairing scheme to optimally utilize the spectrum of unpaired users in non-orthogonal multiple access," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1766–1770, Dec. 2016.

[14] R. Jiao, L. Dai, W. Wang, F. Lyu, N. Cheng, and X. Shen, "Max-min fairness for beamspace MIMO-NOMA: From single-beam to multi-beam," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 739–752, Feb. 2022.

[15] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max–min SINR of cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2021–2036, Apr. 2019.

[16] F. Sun, V. O. Li, and Z. Diao, "Modified bipartite matching for multiobjective optimization: Application to antenna assignments in MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1349–1355, Mar. 2009.

[17] Z. Wang, Y. Cao, D. Zhang, X. Hua, P. Gao, and T. Jiang, "User selection for MIMO downlink with digital and hybrid maximum ratio transmission," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11 101–11 105, Oct. 2021.

[18] H. Abraham, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, G. Alexandrowics, E. Arbel, A. Asfaw, C. Azaustre, P. Barkoutsos, G. Barron *et al.*, "Qiskit: An open-source framework for quantum computing," *URL https://doi. org/10.5281/zenodo*, 2019.

[19] J.-Y. Ko and Y.-H. Lee, "Random beamforming in spatially correlated multiuser MISO channels," in *Proc. IEEE Veh. Tech. Conf.*, Calgary, Canada, Sep. 2008, pp. 1–5.

[20] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[21] R. LaRose and B. Coyle, "Robust data encodings for quantum classifiers," *Phys. Rev. A*, vol. 102, no. 3, p. 032420, Sep. 2020.

[22] C. Zoufal, A. Lucchi, and S. Woerner, "Quantum generative adversarial networks for learning and loading random distributions," *npj Quantum Inf.*, vol. 5, no. 103, pp. 1–9, Nov. 2019.

[23] J. Kusyk, S. M. Saeed, and M. U. Uyar, "Survey on quantum circuit compilation for noisy intermediate-scale quantum computers: Artificial intelligence to heuristics," *IEEE Trans. Quantum Eng.*, vol. 2, pp. 1–16, Mar. 2021.

[24] H. Masoumi, M. J. Emadi, and S. Buzzi, "Coexistence of D2D communications and cell-free massive MIMO systems with low resolution ADC for improved throughput in beyond-5G networks," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 999–1013, Feb. 2022.

[25] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Phys. Rev. A*, vol. 98, no. 3, p. 032309, Sep. 2018.

[26] K. Zhang, M.-H. Hsieh, L. Liu, and D. Tao, "Toward trainability of quantum neural networks," *arXiv preprint arXiv:2011.06258*, 2020.