Pilot Allocation Optimization by Hybrid Quantum-Classical Neural Network in CF-mMIMO

Doan Hieu Nguyen, Xuan Tung Nguyen, Won Joo Hwang

***Abstract*—In this paper, we employ Hybrid Quantum-Classical Neural Network to solve Pilot Allocation Optimization Problem in Cell-free massive Multiple Input Multiple Output system. The proposed model has shown the superior performance through practical stimulation in two scenarios.**

***Index Terms*—Cell-Free massive MIMO, Pilot Assignment, Quantum Machine Learning**

I. Introduction

CELL-Free (CF) massive multiple-input multiple-output (mMIMO) is an innovative wireless communication architecture that eliminates cell boundaries of conventional multicell mMIMO, offer significant enhancements [1]. Due to the limited pilot resources, pilots may need to be reused, causing pilot contamination (PC), raise interference in channel estimation. In order to enhance the performance of CF-mMIMO, the pilot assignment scheme (PAS) must be carefully designed. However, the PAS is in the class of NP-hard problem, and costs expensive to achieve the optimal with the exhaustive search. Several solution have been researched, such as *Greedy PAS* in [1], where the least interference pilot sequence is assigned to the UE which has the lowest data-rate value. The authors in [2] introduced *Master-AP PAS* to optimize selected signal of each UE for its strongest AP. In [3], authors designed a Convolutional Neural Network-based PAS (CNN-PAS) in a supervised manner with labels obtaining from exhaustive search, which provides theoretical upper bound capacity. Although CNN-PAS can reach 97% of optimal capacity gain, attaining learning labels from exhaustive search is impractical in real-time systems because it demands high computational resource and also limits the scalability of the systems.

Heuristic and Deep neural network-based (DNN) pilot assignment methods have limitations, such as focusing on fairness over data-rate or relying on computationally intense labels. Quantum machine learning (QML) has emerged as a promising solution, offering higher expressibility and lower computational complexity than classical models. By leveraging quantum properties like superposition, QML can exponentially expand representational capacity as qubits increase, making QML suited for pilot assignment task. On the other hand, QML requires parameter count scales logarithmically with the input size N, as $O(\log(N))$, which much more fewer than requirement in CNN. Hence, the computations needed reduce, result to lower computational complexity. Being inspired, we propose a Hybrid Quantum-Classical Convolutional Neural Network (HQCCNN) for PAS. We design an HQCCNN architecture derived from the parameter sharing feature of classical CNN kernel, reduces the training parameters significantly while still guaranteeing promising performance . To evaluate this architecture, we introduce two approaches: a supervised and an unsupervised models.

This work was supported by Quantum Computing based on Quantum Advantage challenge research(RS-2024-00408613) through the National Research Foundation of Korea(NRF) funded by the Korean government (Ministry of Science and ICT(MSIT)).

II. System Model and Problem Formulation

We consider a CF-mMIMO systems consists M L-antennas APs, K single-antenna UEs. k-th UE and m-th AP are denoted as U_k and A_m. UEs share τ_p pilots, each pilot is pairwisely orthogonal signal. PC issue happens when there are fewer pilot than UEs. The term $\mathcal{P}(k)$ stands for the set of UE share the same pilot with U_k. The channel vector between A_m and U_k is denoted as $\mathbf{g}_{mk} \in \mathbb{C}^L$, where $\mathbf{g}_{mk} \in \mathbb{C}\mathcal{N}(0, \beta_{mk}\mathbf{I}_L)$, where β_{mk} is the LSF coefficient comprising the path-loss and the shadowing.

In uplink training phases, UEs send assigned pilot sequence to APs in their connection set. The received pilot signal at A_m is formulated as $\mathbf{Y}_m = \sqrt{\tau_p \rho_p} \sum_{k=1}^{K} \mathbf{g}_{mk} \boldsymbol{\varphi}_k^H + \mathbf{W}_m$, where ρ_p is the normalized signal-to-noise ratio (SNR) of each pilot symbol; and $\mathbf{W}_{p,m} \in \mathbb{C}^{L \times \tau_p}$ is AWGN. Similar to [4], the channel coefficient between U_k and A_m is estimated through the minimum mean square error (MMSE) estimator, which is given as

$$\hat{\mathbf{g}}_{mk} = c_{mk} \mathbf{I}_L \left(\sqrt{\tau_p \rho_p} \left(\mathbf{g}_{mk} + \sum_{j \in \mathcal{P}(k)}^{K} \mathbf{g}_{mj} \right) + \mathbf{W}_m \boldsymbol{\varphi}_k \right), \tag{1}$$

where $c_{mk} \triangleq \frac{\sqrt{\tau_p \rho_p} \beta_{mk}}{\tau_p \rho_p \left(\beta_{mk} + \sum_{j \in \mathcal{P}(k)}^{K} \beta_{mj} \right) + 1}$. The mean-square of n-th component of channel estimation is denoted as $\gamma_{mk} \triangleq \mathbb{E}\left\{ |\hat{\mathbf{g}}_{mk}|^2 \right\} = \sqrt{\tau_p \rho_p} \beta_{mk} c_{mk}$.

In downlink of CF-mMIMO network, $\hat{\mathbf{g}}_{mk}$ is treated as true channel, the transmitted signal from A_m to U_k is given by,

$$\mathbf{x}_m = \sqrt{\rho_d} \sum_{k \in \mathcal{A}(m)}^{K} \sqrt{\eta_{mk}} \hat{\mathbf{g}}_{mk}^* q_k, \tag{2}$$

where ρ_d is the maximum normalized SNR transmit power at each AP, q_k is the intended downlink data symbols, and η_{mk} is the power control coefficients, satisfies $L \sum_{k=1}^{K} \eta_{mk} \gamma_{mk} \leq 1$. The received signal at U_k is represented as

$$r_k = \sqrt{\rho_d} \sum_{m=1}^{M} \sum_{j=1}^{K} \sqrt{\eta_{mk}} \mathbf{g}_{mk} \hat{\mathbf{g}}_{mj}^* q_j + w_k \tag{3}$$

The achievable downlink rate at U_k can be represented in closed-form as in (4) in the top of this page.

Pilot contamination factor directly affects the channel estimation error, which cause a observable degradation of net-

$$R_{d,k} = \log_2\left(1 + \frac{\rho_d L^2 \left(\sum_{m\in\mathcal{M}(k)}^{M} \sqrt{\eta_{mk}}\gamma_{mk}\right)^2}{\rho_d L^2 \sum_{j\in\mathcal{P}(k)}^{K}\left(\sum_{m\in\mathcal{M}(k)}^{M}\sqrt{\eta_{mj}}\gamma_{mj}\frac{\beta_{mk}}{\beta_{mj}}\right)^2 + \rho_d L\sum_{j=1}^{K}\sum_{m\in\mathcal{M}(k)}^{M}\eta_{mj}\gamma_{mj}\beta_{mk} + 1}\right) \tag{4}$$

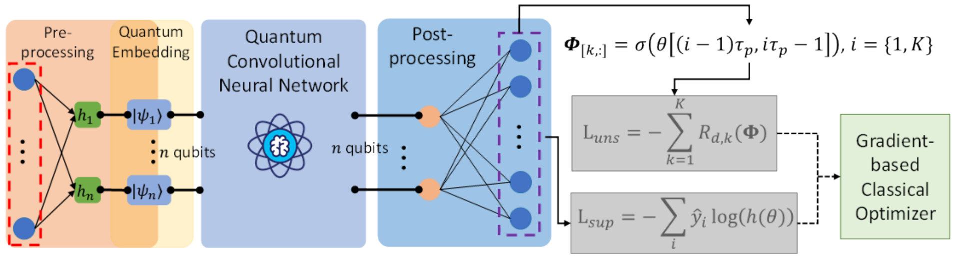

Fig. 1: Architecture HQCCNN: pre-processing layer for quantum state convertation; QCNN is computationally intensive component, post-processing layer to calculate result.

work throughput. Hence, we formulate the pilot assignment problem to maximize the sum-rate of system as the following optimization problem:

$$\underset{\Phi}{\text{maximize}} \sum_{k=1}^{K} R_{d,k}(\Phi) \tag{5}$$

$$\text{subject to} \quad \sum_{t=1}^{\tau_p} \Phi_{k,t} = 1 \quad \forall k, \tag{5.1}$$

The problem in (5) is in class of NP-hard, and difficult to obtain the optimal solution. The existing approaches primarily focus on some priorities such as lowering complexity or reducing interference in specific APs, while other DNN-based approaches are supervised models with a lot drawbacks. Accordingly, in the next section, we design a hybrid quantum neural network to solve the maximum downlink sum-rate optimization problem in CF-mMIMO.

III. Proposed Model: HQCCNN

We propose a Hybrid Quantum Convolutional Neural Network (HQCCNN) for PAS problem in CF-mMIMO systems. The design of HQCCNN, illustrated in Fig. 1, consists two classical linear layers as pre-processing and post-precessing layer with a QCNN layer as hidden layer.

A. Design of HQCCNN Model

The pre-processing layer consists of a linear layer and a quantum embedding layer. The linear layer adjusts the dimension of input data to n-dimension vector, where n is the pre-defined number of qubits used in quantum layer. This vector is convert to quantum states by angle embedding technique as

$$|\psi\rangle = \bigotimes_{i=1}^{n} (\cos(h_i)|0\rangle + \sin(h_i)|1\rangle) \tag{7}$$

where h_i is the i element of $\mathbf{h} = \mathbf{W}^T\mathbf{x} + \mathbf{b}$, is the output of first FC layer.

Quantum Convolution layer is inspired from the concept of classical Convolutional neural network (CNN) models. In CNN model, fixed-size kernels slide across all pixel of grid-structure data to extract features, construct lower-dimension feature maps. The key components are the convolutional layer and the pooling layer, illustrated in Fig. 2.a. This process helps in downsampling the data, reducing both the number of parameters and the computational cost, while also making the model more resilient to small variations in the input data. One of the key advantages of CNN lies in *parameter sharing*, significantly reduces the number of parameters.

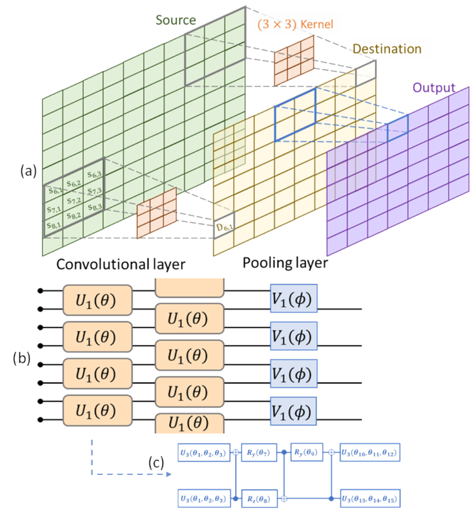

Fig. 2: (a) CNN structure. (b) QCNN structure. (c) VQC architecture.

Generally, the architecture of QCNN is similar to CNN, i.e. it has multiple quantum convolutional layers (QCL) and a final FC layer. In each layer of QCL, there is a convolutional layer and a pooling layer, illustrated in Fig. 2.b. A convolution layer applies single quasi-local unitary operators (U_i) for 2 adjacent qubits. In pooling layer, only a half of qubits are measured, the outcomes are determined by unitary rotations (V_j). At the end, a FC layer is applied. In our study, VQCs in a layer share a mutual 15-parameter set in each QCL. Total QCNN has $15\log_2(n_0)$ parameters, if one qubit remains. Reducing the number of learnable parameters makes our proposed model minimize computational cost, allowing effective operations. Various VQC designs are proposed in [5], with similar performance in expressibility. We recommend VQC9 for its overall performance, as shown in Fig. 2.c.

Post-processing layer is a FC layer used to process the outcomes of QCNN. A vector $(K \times \tau_p, 1)$ is calculated before being reshaped into a (K, τ_p) matrix for further pilot assignment decisions.

B. Training Procedure

In this work, HQCCNN model decides the pilot assignment based on LSF coefficients. The outcome state of l-th layer of QCNN can be expressed as

$$|\psi_i(\boldsymbol{\theta}_i)\rangle\langle\psi_i(\boldsymbol{\theta}_i)| = \mathrm{Tr}_{B_i}\left(U_i(\boldsymbol{\theta}_i)|\psi_{i-1}\rangle\langle\psi_{i-1}|U_i(\boldsymbol{\theta}_i)^{\dagger}\right), \quad (8)$$

where $\mathrm{Tr}_{B_i}(\cdot)$ is the partial trace operation over subsystem, U_i is the parameterized unitary gate operation that includes quantum convolution and the gate part of pooling, θ_i is the parameters of VQC of i-th quantum convolutional layer. After processing, the outcomes of QCNN are calculated as

$$\langle\mathcal{M}_i\rangle = \langle\boldsymbol{\psi}|U(\theta)^{\dagger}B_iU(\theta)|\boldsymbol{\psi}\rangle \quad (9)$$

where $|\boldsymbol{\psi}\rangle$ is the final quantum states, $U(\theta)$ is the product unitary gates, and the observable B_i are typically Pauli-operators and we choose Pauli-Z operators.

We introduce two training schemes as supervised and unsupervised. In terms of supervised process, we employ cross-entropy loss as loss function, given by

$$\mathrm{LF}_{\text{sup}} = -\parallel \sum_{k=1}^{K}\sum_{i=1}^{\tau_p} \mathbf{y}_{ki}\log(h(\boldsymbol{\beta};\theta)_{ki}) \parallel_1, \quad (10)$$

where $h(\boldsymbol{\beta};\theta)$ is prediction pilot assignment probability of model and θ is the trainable parameters.

In unsupervised scheme, The loss-function of unsupervised model is designed as the negative form of sum-rate of total system (4) under the effect of pilot selection probabilities

While classical layers are used conventional back-propagation to iteratively update weights, the parameter set of QCNN layer is updated following gradient-based method, called parameter shift rule [6], [7].

$$\frac{\partial\langle\mathcal{M}\rangle(\theta)}{\partial\theta} = 1/2\big(\langle\mathcal{M}\rangle(\theta + \frac{\pi}{2}) + \langle\mathcal{M}\rangle(\theta - \frac{\pi}{2})\big) \quad (11)$$

IV. Numerical Result

In this study, we consider the system configuration setup as in [1], with $M = \{30, 35, 40, 45\}$, $K = 20$, $\tau_p = 10$, $f = 1.9$ GHz, $B = 20$MHz, noise figure $= 9$dB, height of APs and UEs are 15m and 1.65m respectively; values of D, d_1, d_0 are 1000, 50, 10 (in meters). LSF coefficients are defined as $\beta_{mk} = 10^{\frac{\mathrm{PL}_{mk}}{10}}$, where PL_{mk} represents the path loss value defined by three-slope model in [1]. The corresponding normalized transmit SNRs ρ_d and ρ_p is calculated as $\rho_x = \frac{P_x}{P_N}$, where $x \in \{d, p\}$ correspond to downlink and pilot transmission, thermal noise $P_N = 2e10(W)$. This study is evaluated in two scenarios: first, by comparing HQCCNN with MLP and CNN in supervised learning; second, proposed model compares with R-PAS, Gr-PAS, Ma-PAS, MLP and CNN model. In terms of supervised scenario, we apply *MasterAP*-PAS to obtain labels of training data. Fig. 3 displays that HQCCNN is the fastest convergence and achieved the lowest final loss value, outperform in minimizing error compared to the CNN and MLP. In terms of unsupervised scenario, Fig. 4 depicts the comparison of average sum-rate values when changing the number of APs from 30 to 45. The proposed model outperforms all optimization algorithms and MLP models. The HQCCNN and CNN models perform similarly, achieving the highest average total data rate across different number of APs. However, HQCCNN model has significantly fewer parameters than CNN, as demonstrated in Table I. The experimental results suggest that HQCCNN is the best solution.

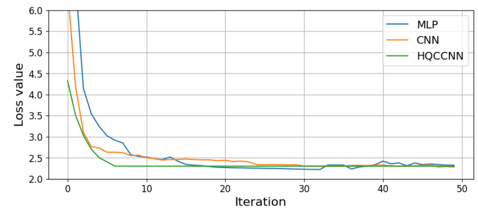

Fig. 3: Supervised training loss

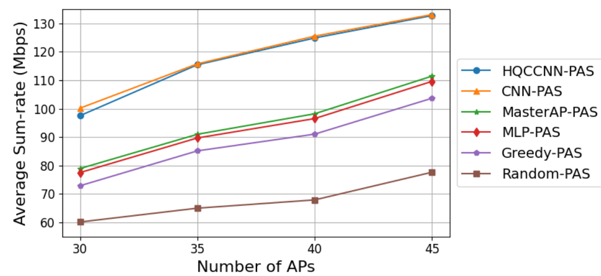

Fig. 4: Unsupervised performance

TABLE I: Parameter comparison

The proposed model	CNN: Kernel (3×3)
QCNN layer: 15	Conv. layer 1 $(1, 32)$: $1 \times (3 \times 3) \times 32$. Conv. layer 2 $(32; 64)$: $32 \times (3 \times 3) \times 64$
Pre-processing layer: $(M \times K + 1) \times n)$	FC layer 1: $((128 + 1) \times 128)$
Post-processing: $(n + 1) \times K \times \tau_p$	FC layer 2: $((128 + 1) \times (K \times \tau_p))$

V. Conclusion

In conclusion, we design a hybrid quantum-classical neural network model to solve pilot allocation in CF-mMIMO systems. Through experiments show the effectiveness of proposed model over several benchmarking technique.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, "Learning-based downlink power allocation in cell-free massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 174–188, 2022.

[3] K. Kim and J. Lee, "Deep cnn based pilot allocation scheme in massive mimo systems," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 14, no. 10, pp. 4214–4230, 2020.

[4] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive mimo," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, 2017.

[5] T. Hur, L. Kim, and D. K. Park, "Quantum convolutional neural network for classical data classification," *Quantum Machine Intelligence*, vol. 4, no. 1, p. 3, 2022.

[6] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, p. 032309, 2018.

[7] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Physical Review A*, vol. 99, no. 3, p. 032331, 2019.