Quantum-Connected Collaborative Learning Workflow for Optimizing Wireless Internet-of-Everything under Imperfections

Goki F. R. Purba, Bhaskara Narottama, *Member, IEEE,* Saeed R. Khosravirad, *Senior Member, IEEE,* Dac-Binh Ha, Haejoon Jung, *Senior Member, IEEE,* Simon L. Cotton, *Fellow, IEEE,* Trung Q. Duong, *Fellow, IEEE*

***Abstract*—The integration of quantum-enabled learning and quantum communication offers new opportunities for secure collaborative optimization in future wireless networks. As these networks become denser and more functionally integrated, distributed quantum learning units may be used to solve complex optimization problems, leveraging unique quantum properties, while exchanging learning information through quantum communication. However, we typically encounter practical challenges, quantum channel noise and device imperfections, which distort the exchanged information. Consequently, it also degrades both the learning process and the resulting optimization performance. This paper, therefore, proposes a quantum-connected collaborative learning (QCCL) workflow which incorporates parameterized quantum circuits (PQCs) and superdense coding protocol to optimize wireless Internet-of-Everything (WIoE) networks. Specifically, we introduce an additional unit called a quantum-classical router (QCR) into the proposed workflow to mitigate bit-flip channel and coherence errors caused by device imperfections. We then apply the QCCL to centralized power allocation in a distributed wireless network to maximize the minimum user rate under transmit-power and minimum-rate constraints. Using the extra QCR unit, it reduces the information-retrieval error probability by approximately 50.41% at a bit-flip probability of 0.25. Moreover, we also demonstrate the reliability of the proposed QCCL for the centralized power allocation under mild quantum-channel and device imperfections.**

***Index Terms*—Quantum machine learning, quantum communication, quantum superdense coding, wireless Internet-of-Everything.**

G. F. R. Purba and B. Narottama are with the Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada (e-mail: {gfrpurba,bnarottama}@mun.ca).

S. R. Khosravirad is with Nokia Bell Laboratories, Murray Hill, NJ 07974, USA (email: saeed.khosravirad@nokia-bell-labs.com).

D.-B. Ha is with the Faculty of Electrical-Electronic Engineering, Duy Tan University, Da Nang 550000, Vietnam (e-mail: hadacbinh@duytan.edu.vn).

H. Jung is with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin-si 17104, South Korea (e-mail: haejoonjung@khu.ac.kr).

S. L. Cotton is with the Centre for Wireless Innovation (CWI), School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT3 9DT, UK (e-mail: simon.cotton@qub.ac.uk).

T. Q. Duong is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1C 5S7, Canada, and with the Centre of Wireless Innovation (CWI), School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K, and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, South Korea (e-mail: tduong@mun.ca).

This paper was accepted in part for presentation at IEEE Asilomar Conference on Signals, Systems, and Computers, Oct. 2025, Pacific Grove, CA.

The work of Dac-Binh Ha was supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant 102.04-2025.72. The work of S. L. Cotton and T. Q. Duong were funded in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) through the EPSRC Hub on All Spectrum Connectivity under Grant EP/X040569/1 and Grant EP/Y037197/1. The work of Haejoon Jung was supported in part by MSIT, South Korea; in part by the National Research Foundation of Korea under Grant RS-2025-23323081; in part by the ITRC Support Programs under Grant IITP-2026-RS-2021-II212046; and in part by the Convergence Security Core Talent Training Business Support Program Supervised by IITP under Grant IITP-2023-RS-2023-00266615. The work of G. F. R. Purba, B. Narottama, and T. Q. Duong were supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109, in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant Program RGPIN-2025-04941, and in part by the NSERC CREATE program (Grant number 596205-2025).

Corresponding authors are Trung Q. Duong and Haejoon Jung.

I. Introduction

THE DEVELOPMENT of sixth-generation (6G) communication technology is facilitating convergence between communication systems and emerging quantum technologies [1], [2]. This reflects that leveraging more quantum properties to help actualize 6G applications is an emerging technology. For example, integrating existing networks with quantum to achieve ultra-secure communication [3], and extending computational capability through distributed quantum computing [2], [4]. Furthermore, the inherent quantum properties such as quantum entanglement and superposition also enable emerging communication approaches in which information can be effectively processed via quantum bits (qubits) [2]. These approaches are further motivated by the slowdown of conventional transistor scaling associated with Moore's law, as transistor dimensions approach the atomic scale, thereby implying the need for an alternative computing and communication paradigm [5]. A major challenge facing current quantum technologies, however, is the absence of fault-tolerant quantum computers, that is, processors capable of detecting and fixing errors using quantum error correction (QEC) [6]. Besides its current limitation, many studies have shown potential benefits in using quantum technologies, particularly in certain tasks that may be infeasible for classical algorithms. However, such analysis is often conducted under the assumption that quantum properties can be utilized perfectly and that algorithms are computed on an ideal, isolated quantum computer [7], which signifies prospects for forthcoming technology.

The realization of 6G is expected to deliver unprecedented capabilities in terms of data rate, latency, reliability, and other major indicators [8]. To meet these expectations, 6G strives to establish a unified intelligent network capable of real-time adaptation and autonomous optimization across diverse communication, sensing, and computing domains. Attaining this vision is typically referred to as realizing a *network of everything*, in which heterogeneous applications and functionalities are deeply interconnected within a cross-domain infrastructure. In particular, wireless Internet-of-Everything (WIoE) emerges as a key application thereof, because it allows not only the collection and transmission of large data pools, but also their processing and integration based on contextual intelligence within wireless networks. To support the operation of such networks, artificial intelligence (AI) and machine learning (ML) serve as key enablers by providing the intelligence required. Typically, pertinent network management and optimization tasks, including resource allocation, channel estimation, and beamforming [9]–[13], can be addressed using AI and ML. However, the associated computational and energy burdens of managing and optimizing these networks are expected to grow with their scale. These challenges are further intensified by the physical limits of classical computing hardware, particularly the slowdown in transistor scaling [14].

As network complexity and data volume are expected to grow, quantum computing and communication offer compelling approaches to information processing, managing vast amounts of information through quantum operations. Quantum-enabled learning has also attracted significant interest owing to its potential to capture nonlinearity and intricate data correlations [15], which are highly relevant to solving the aforementioned optimization problems in Internet-of-Everything (IoE) [16]. This quantum-enabled learning approach also provides a viable practical means on current noisy intermediate-scale quantum (NISQ) computers, where noise and limited qubit resources are present [17]. For example, in the wireless communication field, many studies have investigated integrating deep learning and parameterized quantum circuits (PQCs) to optimize resource allocation, power control, or beamforming configurations by learning channel information patterns [18], [19]. In addition, [20] and [21] demonstrate the utilization of quantum machine learning to estimate channel information and manage fluid antenna systems, respectively. Collectively, these studies highlight the practicality of variational quantum circuits for quantum-enabled learning on near-term quantum hardware, including NISQ computers.

Recent academic and industrial efforts have accelerated quantum computing beyond conceptual constructs, toward early practical realization, notably through cloud-based services, making quantum processors more accessible. Currently, cloud-accessible quantum computers can handle computational tasks involving up to hundreds of qubits, although preserving quantum information at such large scales remains a challenge due to decoherence and the presence of noise [22], [23]. While current NISQ devices remain noise-prone, notable initiatives drive the notion of distributed learning workflows, in which multiple quantum nodes (agents with embedded quantum processor units) jointly execute computational tasks. Such interconnected nodes, constituting distributed quantum systems [22], inherently enable collaborative learning across quantum nodes, where information or learned representations are shared rather than computed centrally. Inspired by hybrid classical-quantum federated learning [24], [25], the same concept can support collaborative learning for early-stage 6G and IoE applications, where, in particular, a central quantum processing node can act as a *teacher* by generating soft targets or distilled knowledge for distributed quantum learners, e.g., mobile edge computing (MEC) units, enabling efficient knowledge transfer without full quantum state sharing.

As connected nodes, they may utilize quantum communication to share information, exploiting quantum entanglement or other protocols, and facilitating visions such as quantum internet [26]. However, such inter-node transmission may encounter loss and attenuation, which may compromise the integrity of the quantum states during communication [27]. Transmitting quantum information often involves sending photons through optical fiber or free-space channels, as typically assumed in protocols such as quantum key distribution (QKD) and quantum superdense coding. Unlike classical channels, quantum channels are governed by quantum mechanics, which introduces distinct concepts of information and channel capacity [27]. The associated state evolution can be particularly described using Kraus operators, with the evolved state remaining completely positive trace-preserving (CPTP). In practice, local computations can be performed using superconducting technologies, while quantum information transmission can be realized through photonic optical networks due to their low latency and communication benefits.

There are possible applications of connecting quantum-enabled learning nodes. One example is quantum-enabled federated learning, wherein local data is processed on quantum edge devices, and only the updates thereof (e.g., gradients or parameters) are aggregated by central servers, to construct global models. More broadly, here we refer to such non-centralized quantum-enabled learning as quantum-connected collaborative learning (QCCL), which includes quantum federated learning, ensemble learning, and related workflows. Despite its benefits, QCCL may face security concerns: i) It is especially vulnerable to Byzantine attacks, in which adversarial nodes send malicious information to disturb the learning process. In this context, the adversarial nodes may deceptively pose as honest nodes. ii) In addition, model evasion attacks allow the adversaries to tamper with the testing process of the trained learning model, giving inaccurate measures of how well the model performs. One may notice that these risks share a common pipeline: attackers manipulate or extract learning information, be it regarding the learning models (e.g., the depth of the circuits), or relevant to them (e.g., the datasets used for training). Accordingly, in optimizing future 6G scenarios such as WIoE, security is no longer only about protecting transmitted data, but must also safeguard the learning process itself. In particular, exchanging model updates and control information over classical channels can compromise the trustworthiness of the learning process. For example, studies such as [28] have classified ML-related attack categories in 6G/ Internet of Things (IoT), including communication-centric threats that arise during the learning stage.

To secure this critical exchange of information, the use of quantum communication protocols has emerged as a solution, utilizing quantum properties to provide secure communication between parties. Developing such protocols typically requires generating quantum entanglement, particularly in the form of Bell states, where any malicious interventions would immediately alter the correlations, exposing the adversarial attempts. Not to mention, the no-cloning theorem prevents any direct copying of qubits, thereby preventing attackers from duplicating transmitted information, particularly models' gradients and prediction outputs. That said, quantum communication realized through quantum optics and photonic technology may encounter device imperfections, leading to non-ideal operation. In real-world experiments, such imperfections may result in detector inefficiency, optical interferometer drift, and beam-splitter imbalance. While non-coherent errors affect the quality of quantum information during quantum transmission via the quantum channel, hardware impairments associated with using photonic technology should also be considered, as many quantum applications, such as teleportation, superdense coding, and entanglement swapping, depend on high-fidelity Bell state measurements.

A. Motivations and Objectives

Collaborative learning strategies aimed at developing efficient, robust, and precise computing paradigms have been discussed in survey papers such as [24], [29]. Yet, while the exchange of learning information (e.g., model updates or parameters) continues to rely on classical channels, the use of quantum channels for this purpose remains relatively underexplored. For example, quantum servers that collaborate with classical clients, while exchanging classical snapshots and local gradients, yet still use classical transmission, are presented in [30]. Likewise, [31] adds classical Laplace-distributed differential privacy noise to protect transmitted gradients from clients to the server, yet the transmission occurs over classical channels. In addition to quantum communication, recent studies further indicate that quantum-enabled learning and quantum communication are beginning to intersect. For example, [32] uses quantum circuit learning for quantum entanglement distillation and quantum state discrimination, aiming to maximize the conditional fidelity and success probability for noisy shared entangled states, respectively. [33] employs a quantum autoencoder approach where encoder and decoder quantum-enabled learning circuits are jointly trained to adapt to a noisy channel. However, quantum communication, as an interconnection between quantum-enabled learning units that directly participate in optimization, remains less explored. Capitalizing on the security benefits of quantum communication and quantum-enabled learning, therefore, this study introduces QCCL, utilizing protocols, particularly superdense coding, to connect different learning models, as in Fig. 1.

As we briefly mentioned, noise remains a major problem in quantum systems, including quantum communication, and it is typically handled by QEC or error mitigation, which is mostly assumed to apply in fault-tolerant systems. The use of quantum learning models, such as quantum neural networks (QNNs), in addition, has also been explored in the noisy settings, primarily for optimization and noise characterization. For example, [34] uses QNN for high-fidelity reconstruction of distorted orbital angular momentum (OAM) beam profile images, affected by perturbative channels. QNNs have the potential to be used in a quantum environment for learning in the presence of quantum noise-induced errors: QNNs can be used to characterize noise and recover distorted images [34]. Another work, [35], uses a QNN to profile quantum noise and leverage the information to optimize noise effects by mapping highly influenced output logical qubits onto reliable physical qubits. Accordingly, this study introduces quantum-classical router units (QCRs) that use a quantum model to learn the noise characteristics and thus facilitate both communication and noise mitigation. As a proof-of-concept, it further leverages QCCL to maximize clients' achievable sum rate in WIoE, under different quantum channel conditions. Our initial study indicates that, even without such noise mitigation models, QCCL can attain reasonable performance under low error rates, albeit with the expected performance degradation [36].

In optimizing WIoE networks amid imperfections in quantum channels, this paper aims to deliver the following contributions. i) This paper proposes the QCCL workflow in which distributed quantum-enabled learning units are connected through a quantum communication protocol, particularly superdense coding. ii) We develop an imperfection-aware quantum communication model by considering both quantum-channel-induced bit-flip errors and coherent Bell-operation errors caused by device imperfections. iii) We design a learnable QCR unit based on a PQC model to mitigate these imperfections and recover the transmitted information more reliably before it is used at the MEC unit. iv) We apply the proposed QCCL workflow to a distributed WIoE network and formulate a $\max\text{-}\min$ power-control problem to improve the minimum user rate under AP power and service constraints. v) This paper numerically shows that the proposed QCR-assisted QCCL workflow reduces quantum-link decoding errors and improves the minimum-rate and sum-rate performance compared with equal-power allocation under imperfect quantum communication.

B. The Paper's Organization and Notation

What follows outlines the rest of this paper. In Section II, we discuss the quantum channel and errors associated with imperfect deployment. Then, the propositions for connecting quantum learning models by means of quantum communi-

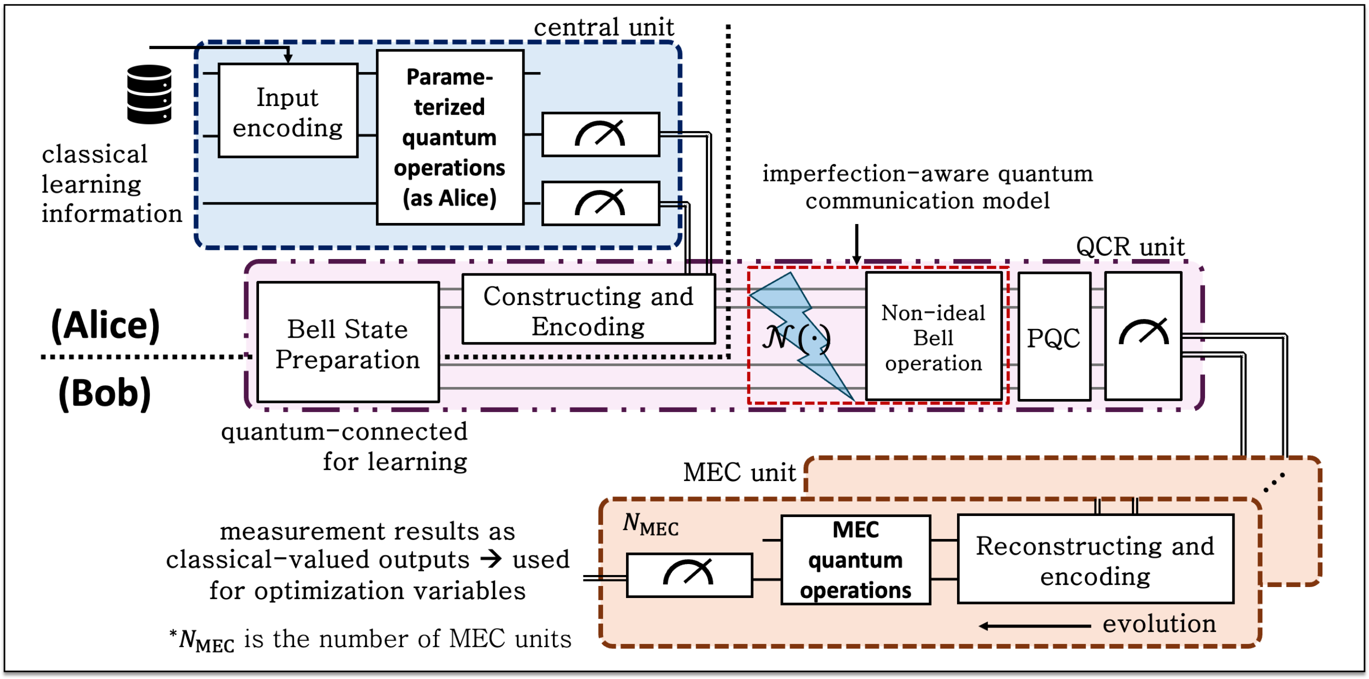

Fig. 1: The proposed QCCL workflow between Alice and Bob. The central unit processes classical learning information and sends the quantized outputs to Bob via superdense coding. During transmission, the encoded state passes through an imperfection-aware quantum communication model (red dashed box), which is then mitigated by the QCR unit before the reconstructed values are used as optimization variables.

cation, particularly as QCCL, are detailed in Section III. We discuss the system model of the wireless network in Section IV, with particular attention given to the optimization objective and the performance indicators. In Section V, we demonstrate the utility of QCCL with a performance analysis. In Section VI, we conclude this paper and offer future research directions.

Throughout the paper, we use a variety of quantum gate notations, particularly $\mathsf{R}_{\mathsf{x}}(\theta), \mathsf{R}_{\mathsf{y}}(\theta), \mathsf{R}_{\mathsf{z}}(\theta)$ denote rotations around the x-, y-, and z-axes, with θ representing the rotation angle. Controlled-X gates are indicated by $\mathsf{C}_{\mathsf{x}}(|\psi_{\mathsf{c}}\rangle, |\psi_{\mathsf{t}}\rangle)$, with $|\psi_{\mathsf{c}}\rangle$ and $|\psi_{\mathsf{t}}\rangle$ indicating the states of the control and the target qubits, respectively. $\mathsf{U}^{\otimes q}$ stands for parallel operations of unitary U using q qubits. H denotes the Hadamard gate notation. $\mathbb{I}$ denotes the identity gate, and a vector and a matrix are respectively indicated through bold lowercase and uppercase lettering. The operators of transpose and Hermitian are given by $(\cdot)^T$ and $(\cdot)^H$, respectively. The absolute value is denoted by $|\cdot|$. $(\cdot)^\dagger$ denotes the Hermitian conjugate of a matrix or operator. $\mathbb{C}$ and $\mathbb{R}$ denote the fields of complex and real numbers, respectively, while $\mathfrak{Re}(\cdot)$ and $\mathfrak{Im}(\cdot)$ convey the real and imaginary values of a complex number, respectively.

II. Preliminary on Quantum Computing and Channel

The following section details the fundamental concepts of quantum computing and communication. In particular, it also discusses the potential errors pertaining to quantum communication, highlighting the relevance of addressing such errors using quantum learning models.

The basic unit of a quantum system is typically referred to as a qubit, represented as a Dirac vector that describes its state. For example, the computational basis states of a single qubit can be expressed as $|0\rangle = [1, 0]^T$ and $|1\rangle = [0, 1]^T$. Quantum processors store and process information distinctly from their classical counterparts, owing to inherent quantum phenomena such as superposition and entanglement. In particular, quantum superposition permits a quantum system to exist in multiple states simultaneously. For a single qubit, this can be expressed as $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle \in \mathbb{C}^2$, where α_0 and α_1 are the complex amplitudes, of which the squared magnitudes, $|\alpha_0|^2$ and $|\alpha_1|^2$, must add up to 1, to preserve the probabilistic interpretation of quantum mechanics. Generalizing this formulation, for a register of N_q qubits, the state of the system is given by $|\psi\rangle = \sum_{i=0}^{2^{N_q-1}} \alpha_i|i\rangle \in \mathbb{C}^{2^{N_q}}$, where the squared magnitudes of the amplitudes satisfy the normalization condition $\sum_{i=0}^{2^{N_q}-1} |\alpha_i|^2 = 1$, and $|i\rangle$ denotes the computational basis state corresponding to the binary representation of the integer i. In influencing the states of the qubits, one may apply different quantum gates. For example, the controlled-Z gate, which can be defined as $\mathsf{U}_{\mathsf{C_Z}}$ represents a two-qubit operation that applies a phase flip only when the states of both qubits are $|1\rangle$, explicitly, $|11\rangle \rightarrow -|11\rangle$. Another quantum phenomenon of interest is quantum entanglement, in which two or more qubits are linked such that the state of a qubit depends on the state of another, regardless of the spatial distance between them. Accordingly, measuring one qubit allows us to infer information about the other instantly. Here, "measuring" refers to the act of performing a quantum measurement, whereby classical outcomes are obtained by observing a quantum system. Further discussion on this is provided in Section III-B. As this paper focuses on applying quantum principles to the benefits of quantum collaborative learning and communications, those interested in the foundational theories are encouraged to consult treatises such as [37].

A. The Evolution of Quantum State Under Quantum Channel

A quantum channel can be described as a physical transformation that maps a quantum state operator, a quantum density operator, to another operator, over a Hilbert space. This transformation can be mathematically described as a linear mapping: $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, in which $\mathcal{H}$ denotes a Hilbert space, and $\mathcal{L}(\mathcal{H})$ denotes a space of linear operators on $\mathcal{H}$. This maps operators linearly, and represents a quantum system evolution, which may be unitary or non-unitary, depending on the system interaction. The unitary dynamics are ideal evolutions because they are reversible and do not destroy information. Typically, to obtain this type of evolution, the system interaction ought to be isolated from the external environment, a condition usually referred to as a closed quantum system. However, inevitable interaction with the environment causes the evolution of the system of interest to become non-unitary due to induced errors. However, it can still be represented as a unitary operation, by recasting it within an extended system, which includes the environment and the system of interest, known as the Stinespring dilation theorem. It follows that, in evaluating a valid physical transformation, the said evolution must be a trace-preserving and completely positive map [27].

A quantum channel, resembling a state evolution, can be described using the following operator-sum representation

$$\mathcal{E}(\rho_{\text{in}}) = \sum_i \mathcal{M}_i \rho_{\text{in}} \mathcal{M}_i^{\dagger}, \tag{1}$$

where $\{\mathcal{M}_i\}_i$ represents the Kraus operators, satisfying $\sum_i \mathcal{M}_i^{\dagger}\mathcal{M}_i = \mathbb{I}$. This evolution can be derived by tracing out the environment from the extended system, containing the system of interest and an auxiliary. As noted earlier, while the reduced dynamics of the system makes the evolution of the interested system is non-unitary, the evolution of the extended system remains unitary $\rho_{\text{exp}} = \mathsf{U}(\rho_{\text{in}} \otimes |0\rangle\langle 0|_{\text{env}})\mathsf{U}^{\dagger}$. The use of an auxiliary state $|0\rangle\langle 0|_{\text{env}}$, also known as an ancilla, helps describe a system's interaction with its environment. Accordingly, any such interaction, including those wherein systems are entangled with others, is represented by a completely positive (CP) map. Therefore, we can describe the evolution of the system of interest as

$$\mathcal{E}(\rho_{\text{in}}) = \text{Tr}_{\text{env}}\left[\mathsf{U}(\rho_{\text{in}} \otimes |0\rangle\langle 0|_{\text{env}})\mathsf{U}^{\dagger}\right], \tag{2}$$

$$= \sum_i (\mathbb{I}_{\text{in}} \otimes \langle \xi_i|_{\text{env}}) V \rho_{\text{in}} V^{\dagger} (\mathbb{I}_{\text{in}} \otimes |\xi_i\rangle_{\text{env}}), \tag{3}$$

where $\mathsf{U}\mathsf{U}^{\dagger} = \mathbb{I}$, $\text{Tr}_{\text{env}}[\cdot]$ denotes a partial trace operator over an environment subsystem, which operates on the joint state ρ_{exp}. The term $V = \mathsf{U}\mathsf{W}$ captures an isometry operator, where $\mathsf{W} : |\xi\rangle_{\text{in}} \mapsto |\xi\rangle_{\text{in}} \otimes |\xi\rangle_{\text{env}}$ represents an embedding isometry, $V^{\dagger}V = \mathbb{I}$, and $\mathcal{E}(\rho_{\text{in}}) \succeq 0$ [27]. Following this equation, it implies that Eqs. (1) and (3) are equivalent, for the i-th Kraus operator is given by $\mathcal{M}_i = (\mathbb{I}_{\text{in}} \otimes \langle \xi_i|_{\text{env}})V$, $\forall i$, where $|\xi_i\rangle_{\text{env}}$ forms a basis for the said environment.

Based on the above characterization, we now discuss the evolution of quantum channels, which generally comprises i) decoherence and ii) coherence errors.

As the interaction of the system with the environment induces errors, these errors can be modeled using Kraus operators, where each operator represents a physical error. Therefore, we can use operator sum representation to model the errors arising from such interaction. For example, a bit-flip channel with its parameter ζ results an output density state ψ_{out}, as expressed in the following:

$$\mathcal{N}(\psi_{\text{in}}) = \psi_{\text{out}} = \zeta \sigma_{\text{x}} \psi_{\text{in}} \sigma_{\text{x}} + (1-\zeta)\psi_{\text{in}}, \tag{4}$$

where σ_{x} indicates the Pauli-X operator corresponding to the bit-flip error, with ζ indicating its probability. Here, the gates with probabilities still represent Kraus forms, with $\mathcal{M}_{\text{x}} = \sqrt{\zeta}\sigma_{\text{x}}$ and $\mathcal{M}_{\text{I}} = \sqrt{1-\zeta}\mathbb{I}$. We can also say that, due to the interaction, the channel output becomes a probabilistic mixture, which is also known as quantum decoherence. Although a practical quantum communication channel may involve various error conditions, including phase-flip and depolarizing channels, throughout this paper, we consider the bit-flip channel as a representative Pauli error model because it can be interpreted as a logical-bit mismatch at the receiver, and it provides a simple and interpretable first-order abstraction for analyzing how quantum-channel errors propagate through the proposed QCCL workflow. The evolution of the Bell density matrix, after transmission through a bit-flip channel, has been discussed in studies such as [38], and extensions to more general quantum channels are left for future work.

A coherence error is a form of error arising during a quantum operation, in which the implemented state rotation deviates from the ideal, either through under-rotation, over-rotation, or rotation about an incorrect axis. Typically, hardware impairments or miscalibrations can lead to this category of error. Such interactions can be described as an incorrect Hamiltonian, which still generates a unitary evolution, yet deviates from the intended unitary. For example, the ideal/logical C_{z} Hamiltonian comes from $H_{\mathsf{C}_{\mathsf{z}}} = \pi|11\rangle\langle 11|$, and in the Pauli form it can be written as

$$H_{\mathsf{C}_{\mathsf{z}}} = \frac{\pi}{4}\left[\mathbb{I} \otimes \mathbb{I} - \mathbb{I} \otimes \mathsf{Z} - \mathsf{Z} \otimes \mathbb{I} + \mathsf{Z} \otimes \mathsf{Z}\right]. \tag{5}$$

Therefore, the ideal unitary can be written as $\mathsf{U}_{\mathsf{C}_{\mathsf{z}}} = e^{-iH_{\mathsf{C}_{\mathsf{z}}}\tau}$. An imperfect entangling angle error can be illustrated by an extra ϵ_{ϕ} in the Hamiltonian, and it can be written as $H_{\mathsf{C}_{\mathsf{z}}} = \pi(1+\epsilon_{\phi})|11\rangle\langle 11|$. Following this notion, we abstract the imperfections of defective devices by modeling the hardware-dependent coherence at the qubit level. Since we focus on Bell-state operations, we form non-ideal unitaries for Hadamard

("H") and CNOT gates by adding extra terms to their ideal Hamiltonians. In particular, the ideal Hamiltonian for the Hadamard gate can be articulated as $H_{\mathsf{H}} = \pi/2\sqrt{2}(\mathsf{X}+\mathsf{Z})$, and we use

$$H_{\mathsf{H}}^{(\mathsf{err})} = \frac{\pi}{2}(1+\epsilon)\left(\frac{\mathsf{X}+\mathsf{Z}}{\sqrt{2}}\right) + \frac{\delta}{2}\mathsf{Z} \tag{6}$$

to represent an imperfect device when applying the unitary operation corresponding to this Hamiltonian. In this Hamiltonian, ϵ and δ represent over/under rotation and equivalent mismatch of the Hadamard operation, respectively. With respect to the CNOT operation, we may build it leveraging H and $\mathsf{C_z}$, in situations where $\mathsf{C_z}$ is more accessible at the hardware level. To be precise, we can obtain the equivalent of CNOT by placing H after and before $\mathsf{C_z}$ on the target qubit: $\mathsf{CNOT} = (\mathbb{I}\otimes\mathsf{H})\mathsf{U}_{\mathsf{C_z}}(\mathbb{I}\otimes\mathsf{H})$, where U_{CZ} is the ideal operation. Since we consider imperfect devices, the $\mathsf{C_z}$ Hamiltonian can be represented as $H_{\mathsf{C_z}}^{(\mathsf{err})} = \pi(1+\epsilon_\phi)|11\rangle\langle 11| + \delta_1\mathsf{Z}\otimes\mathbb{I} + \delta_2\mathbb{I}\otimes\mathsf{Z}$. We add extra Z terms on both the control and the target qubits herein to illustrate an equivalent mismatch of phase drift error within Bell-state operations. The imperfect operation of the Bell unitary can therefore be described by

$$\mathsf{U}_{\mathrm{Bell}}^{(\mathsf{err})} = \mathsf{U}_{H_{\mathsf{H}}}^{(\mathsf{err})}\mathsf{U}_{H_{\mathsf{CNOT}}}^{(\mathsf{err})}, \tag{7}$$

where $\mathsf{U}_{H_{\mathsf{H}}}^{(\mathsf{err})}$ and $\mathsf{U}_{H_{\mathsf{CNOT}}}^{(\mathsf{err})}$ represent the imperfect Hadamard and CNOT unitary operators, respectively. The time evolution of a closed quantum system can be generated by its Hamiltonian, to describe its time evolution. Generally, a unitary operator can be constructed as $\mathsf{U}_H(\tau) = e^{-iH\tau}$, where H denotes the corresponding Hamiltonian matrix generating the evolution, and τ conveys the evolution duration [37]. Likewise, we adopt this construction to represent the imperfect Bell gate in Eq. (7).

The following discussions detail the impacts of such imperfections. Initially, suppose we have a pure quantum state, for which we have complete knowledge of the quantum system, and the measurement outcomes are fully determined by the quantum randomness. After it goes through the quantum channel in Eq. (4), the state transforms into a statistical mixture state. In this manner, the randomness of the system is no longer fully quantum in origin. Rather, it contains classical probabilistic uncertainties, as we can observe from the Kraus operators. In other words, the purity of the quantum state (the interference terms) shall be reduced or may even vanish when noise affects the system. By contrast, the state remains pure after undergoing an imperfect unitary operation.

In classical information theory, the maximum extractable information from a classical channel is quantified by the mutual information between inputs A and B, which is generally described as $I(A:B)$. The amount of information that can be conveyed through the channel is quantified through the entropy of the random source variable, signified as $H(A)$. In a quantum system, quantifying the uncertainty of a quantum state is typically used to capture both entanglement, the fundamental property of quantum mechanics, and extracting classical information from a quantum system. Such a quantification can be done by calculating the quantum entropy, known as the von Neumann entropy, as follows.

$$S(\rho) = -\operatorname{Tr}(\rho\log_2\rho) = -\sum_i \lambda_i\log_2\lambda_i, \tag{8}$$

where $\{\lambda_i\}$ are the corresponding eigenvalues of the given state, represented by the density matrix ρ [39]. For example, for a pure quantum state, which contains no uncertainty of itself, the entropy is $S(\rho)=0$. For maximally entangled states, where d represents the dimension of the Hilbert space, the entropy reaches $S(\rho) = \log_2(d)$, representing a completely random state. For general mixed states, the entropy holds $S(\rho) > 0$.

We now discuss the transmission of classical information through a quantum channel, where classical information is encoded into quantum states using an ensemble of quantum states $\{p_i, \rho_i\}$. The amount of classical information that can be extracted is restricted by the Holevo quantity, as follows.

$$I(A:B) \le \mathcal{X} \le H(A), \tag{9}$$

where

$$\mathcal{X} := S(\rho_{\mathrm{esm}}) - \sum_i p_i S(\rho_i) \tag{10}$$

represents a general definition of the Holevo quantity. Here, $\rho_{\mathrm{esm}} = \sum_i p_i\rho_i$ captures the ensemble-averaged states, where ρ_i denotes the i-th encoded state with probability p_i [37]. Based on Eq. (10), the use of mixed states in the ensemble generally reduces the amount of accessible information. Besides the choice of encoding states, the non-orthogonality between the encoded states further decreases the accessible information because of the state overlap. Such an overlap is a property that limits the distinguishability of quantum states. For the case of two encoded states, this minimum unavoidable error is bounded by the Helstrom limits: $P_{\mathrm{e}}^{\min} = \frac{1}{2}\left(1 - \|p_0\rho_0 - p_1\rho_1\|_1\right)$, where $\|\cdot\|_1$ represents a trace norm. Therefore, the maximum accessible information cannot exceed classical entropy (see Eq. (9)). This bound can be achieved by an ensemble comprising pure and orthogonally encoded states [37], [39]. For this reason, we select pure, orthonormal basis states as our encoding states. Further details on quantum communication security, pertinent to this functionality, appear in Section V.

III. Quantum-Connected Collaborative Learning

In this work, we use the term quantum-connected collaborative learning to describe two or more quantum systems linked through a quantum communication protocol. Such a connection is designed to securely share meaningful values during and after the learning process, leveraging it within a secure distributed quantum optimization system. Figure 1 illustrates a quantum-connected learning between the central unit, as "Alice", and the MEC unit, as "Bob", where the MEC unit may have multiple units under this study. This workflow can be directly adapted to other distributed quantum learning architectures discussed in Section I. For example, when transmitting local quantized gradients or model parameters in quantum federated learning, the QCR unit can be placed within a quantum communication link to mitigate channel- and

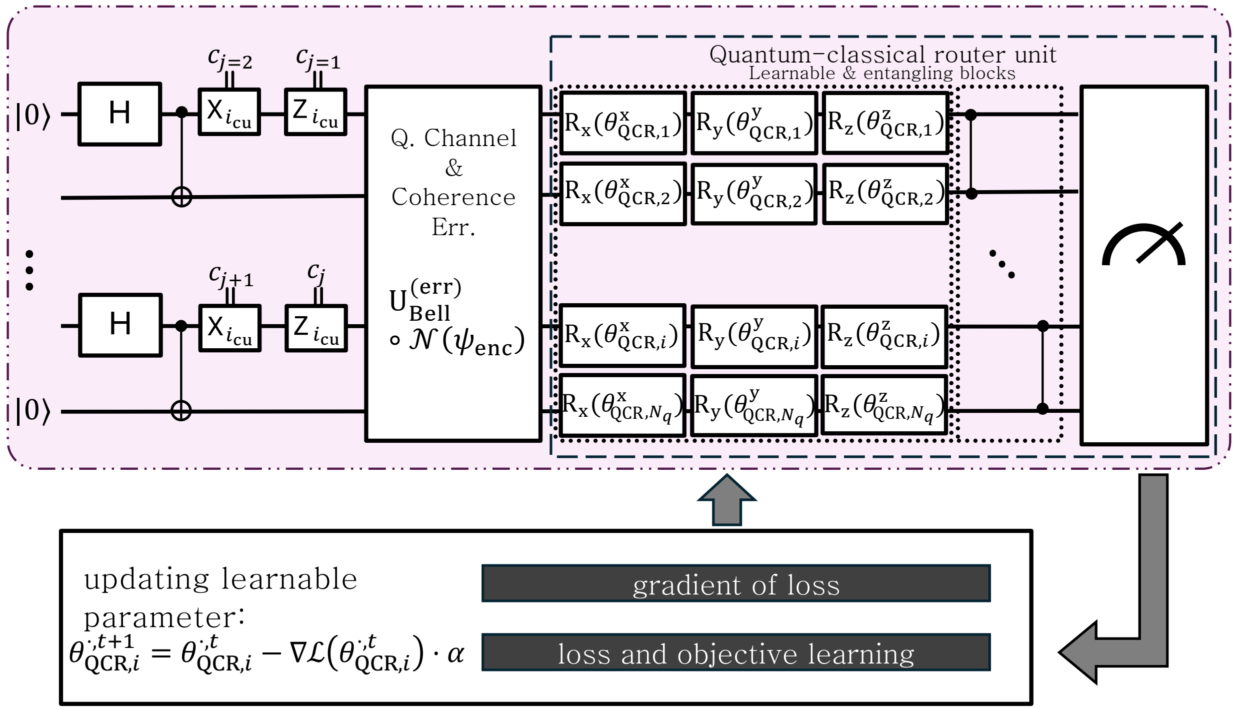

Fig. 2: Training workflow of the proposed QCR unit. The noisy superdense-coded state is processed by trainable rotation and entangling blocks, and the measured output is used to compute the loss gradient and update the circuit parameters.

device-induced imperfection before recovering the information for global aggregation [24], [25]. Thus, the workflow can be considered as a link-level quantum communication and mitigation interface for distributed quantum learning architectures. In the following paragraphs, we first discuss the employed quantum communication protocol and then explain how the central and MEC units learn collaboratively.

A. Quantum Communication and Quantum-Classical Router Unit

One prominent quantum communication protocol is the superdense coding, which facilitates the transmission of two classical bits using a single qubit. Now we explain the use of this protocol, assuming a simple scenario using only one pair of entanglement states. Nevertheless, it is also possible to extend this scenario to a higher number of entanglement pairs, using a similar process. i) The process of quantum superdense coding requires the central unit and the router unit to share a Bell state to establish a link. In this context, we assume that both the central unit and the router unit always share ideal Bell state qubits during the learning and optimization process, as they are used as entangled resources for secure communication. ii) Having prepared the state $|\Phi^+_{\mathrm{c;r}}\rangle$, the central unit's qubit is encoded using the superdense coding protocol to send its classical bits by following protocol $\mathrm{U}_{\mathrm{SDC}} = \left(X_{i_{\mathrm{cu}}}^{c_{j+1}} \cdot Z_{i_{\mathrm{cu}}}^{c_j}\right)$, where X and Z are the Pauli gates on the i_{cu}-th central unit's qubits. iii) After the preparation and encoding process, the whole state at this stage can be written as

$$|\psi_{\mathrm{enc}}\rangle = \mathsf{U}_{\mathrm{SDC}}|\Phi^+_{\mathrm{c;r}}\rangle. \tag{11}$$

In the ideal case, the prepared state can be characterized either by measuring its Pauli correlators $\{\mathsf{XX}, \mathsf{YY}, \mathsf{ZZ}\}$ or by applying a Bell-basis transformation (Bell operation) and then performing computational-basis measurements. However, since the ideal case is hard to achieve in a quantum communication scenario, considering noise and non-ideal hardware operation is hardly avoidable. iv) We therefore introduce an extra operation to the Bell operation circuit to characterize the noisy prepared state, which we define within the QCR unit. The discrete measurement outcomes (collapsed state) from this unit will be demapped classically to a predefined continuous value, subsequently used as input data during execution at the MEC unit. We discuss this approach in more detail in the following.

Previously, we discussed in general the interaction between the environment and the quantum system of interest, which can be described by a unitary transformation in Eq. (2). The interaction also describes the evolution of the quantum state, especially during the transmission of the encoded qubits after the central unit applies the superdense-coding operation, which can be written as $\mathcal{E}(\psi_{\mathrm{enc}})$, where $\psi_{\mathrm{enc}} = |\psi_{\mathrm{enc}}\rangle\langle\psi_{\mathrm{enc}}|$. Following the bit-flip channel model in Eq. (4), errors occurring during transmission not only perturb the isolated qubits, but also alter the Bell-state correlation that is later decoded at the QCR. On top of that, we also use an abstraction of the imperfect Bell operation as the effect of coherence noise (see

Eq. (7)). Hence, as we use quantum communication to facilitate secure transmission under such impairments, the evolution of the original information state is affected by both errors, $\mathsf{U}^{(\mathsf{err})}_{\text{Bell}} \circ \mathcal{N}(\psi_{\text{enc}})$, which substantially affects the subsequent optimization stages. Therefore, this study specifically aims to address these imperfections during quantum communication by executing an additional parameterized circuit inherently within the QCR.

Fig. 2 depicts the integration of a parameterized circuit within a QCR to support quantum communication, and further illustrates the QCR learning process, in which each circuit parameter is updated according to the gradient of the loss function defined by the learning objective. During the learning process, this circuit is pre-trained, as illustrated in Fig. 1, prior to its deployment as a quantum-enabled learning component within the overall QCCL workflow. Typically, trainable quantum circuits, such as PQCs and QNNs, consist of two main blocks: parameterized rotation blocks and qubit entangling blocks. These blocks form the core functionalities of quantum-enabled learning, and thus can be used for addressing impairments within learning pipelines. Accordingly, we utilize these blocks jointly to learn noise characteristics and subsequently mitigate errors. Specifically, the goal of the QCR circuit is to construct a trainable unitary approximator capable of representing an ideal Bell transformation, thereby facilitating information retrieval in the presence of noise. For an N_q-qubit QCR system, we design the PQC as

$$\mathsf{U}_{\text{pqc,QCR}}(\boldsymbol{\theta}_{\text{QCR}}) = \prod_{i}^{N_q-1} \mathsf{C}_{\mathsf{z}}(|\psi_i\rangle, |\psi_{i+1}\rangle)\Bigg(\bigotimes_{i=1}^{N_q} \mathsf{R}_{\mathsf{z}}(\theta^{\mathsf{z}}_{\text{QCR},i})\mathsf{R}_{\mathsf{y}}(\theta^{\mathsf{y}}_{\text{QCR},i})\mathsf{R}_{\mathsf{x}}(\theta^{\mathsf{x}}_{\text{QCR},i})\Bigg), \tag{12}$$

where $\boldsymbol{\theta}_{\text{QCR}} = [\boldsymbol{\theta}^{\mathsf{z}}_{\text{QCR}}, \boldsymbol{\theta}^{\mathsf{y}}_{\text{QCR}}, \boldsymbol{\theta}^{\mathsf{x}}_{\text{QCR}}]^{\mathsf{T}} \in \mathbb{R}^{3N_q}$ and $\boldsymbol{\theta}^{\mathsf{z}}_{\text{QCR}}, \boldsymbol{\theta}^{\mathsf{y}}_{\text{QCR}}, \boldsymbol{\theta}^{\mathsf{x}}_{\text{QCR}} \in \mathbb{R}^{1\times N_q}$ are the row vectors for particular Pauli-gates rotations. This combination of rotations provides a general single-qubit unitary structure that enables the circuit to mitigate amplitude, phase, and axis-misalignment effects induced by imperfect quantum operations. Meanwhile, the entangling block is included to capture the two-qubit correlation required for the Bell state. This circuit can be stacked or applied multiple times to form layers when quantum communication and hardware-related errors become sufficiently complex. As ansatz design in quantum-enabled learning requires balancing model expressiveness and trainability [40], a shallow ansatz is considered in this work. Here, the protocol requires the number of qubits to be a multiple of two, with a larger number of qubits affecting the quantum communication effectiveness pertinent to the quantization level. For instance, when both the quantization level γ and the number of transmitted qubits N_q are equal to 4, then the quantum-connected system transmits only a single piece of information from the central unit (namely, the measurement outcome used subsequently as an optimization variable later, see III-B) to the MEC. Accordingly, reducing γ decreases quantization resolution, while reducing N_q, albeit potentially improving qubit efficiency, increases the computational complexity of the learnable QCR unit.

During the QCR learning process, we train the model using a predefined subset of classical bit inputs to estimate noise characteristics by updating the parameters via minimizing the loss function, as illustrated in Fig. 2. When $N_q = 2$, for example, we train the model on the classical input set $\mathcal{C}_{\text{QCR}} = \{00, 11\}$, which generates the Bell states belonging to different parity sectors (correlated and anti-correlated) in the Z basis. Using this scenario, we can train the QCR unit while simultaneously demonstrating quantum communication, after which it is used as an additional operation to mitigate errors. The learning objective of the QCR circuit is thus to counteract the effect of imperfections in the Bell unitary (see Eq. (7)), such that the trained PQC in Eq. (12) represents the ideal Bell transformation. This objective can be described by the following expression.

$$\mathsf{U}_{\text{pqc,QCR}}(\boldsymbol{\theta}^*_{\text{QCR}}) \circ \mathsf{U}^{(\mathsf{err})}_{\text{Bell}} \circ \mathcal{N}(\psi_{\text{enc}}) \approx \mathsf{U}_{\text{Bell}}(\psi_{\text{enc}}). \tag{13}$$

Using the representative Bell states in the predefined set, the QCR can learn to identify the coherent misalignment. Once the compensating transformation is approximated, the learned correction naturally extends to the remaining Bell states, owing to their Pauli symmetry. Therefore, we define a loss function and aim to minimize it by adjusting the entries of $\boldsymbol{\theta}_{\text{QCR}}$ over the predefined input sets during the learning stage. The optimization reads:

$$\begin{aligned} \underset{\boldsymbol{\theta}_{\text{QCR}}}{\text{minimize}} \quad & \mathcal{L}_{\text{QCR}}(\boldsymbol{\theta}_{\text{QCR}}) = C_{00}(\boldsymbol{\theta}_{\text{QCR}}) + C_{11}(\boldsymbol{\theta}_{\text{QCR}}), \\ \text{subject to} \quad & 0 \preceq \boldsymbol{\theta}_{\text{QCR}} \preceq 2\pi, \end{aligned} \tag{14}$$

where each entry of $\boldsymbol{\theta}_{\text{QCR}}$ holds non-negative value, and is constrained to the interval $[0, 2\pi]$, corresponding to the rotation phases of quantum rotation gates. The term C_j denotes the cost associated with the error probability of the quantum-connected circuit, for transmitting the set of classical information j, which can be written as $C_j(\boldsymbol{\theta}_{\text{QCR}}) = (1 - P(\hat{j}|j))^2$, where $\hat{j}$ is the predicted output. Specifically, the terms $P(00|00)$ and $P(11|11)$ indicate the probabilities that the classical information 00 and 11 are correctly retrieved under the measurement $\langle\cdot\rangle$. $C_{00}(\boldsymbol{\theta}_{\text{QCR}})$ and $C_{11}(\boldsymbol{\theta}_{\text{QCR}})$ respectively indicate the error rates pertinent to transmitting classical information 00 and 11, with respect to the parameter set $\boldsymbol{\theta}_{\text{QCR}}$.

B. Central and MEC units within a collaborative learning workflow

As discussed, the collaborative approach involves both the central and MEC units sharing the same learning objective, connected through quantum channels, exploiting quantum security, i.e., security inherited from quantum properties, such as the no-cloning theorem. The collaborative approach has two options for applying quantum-connected collaborative learning [36]: i) store and process all information at the central unit, or ii) partially store and process the information at both the central and the MEC units. In this work, we focus on the first option, as it is more practical at the time this work is conducted. At the central unit, the circuit comprises three main parts: an encoding circuit, a parameterized quantum circuit, and the measurement part. In the encoding process, the central unit encodes classical information (in this work, wireless channel state information of the users) into Hilbert spaces. This process transforms classical data entries into quantum states, facilitating the utilization of quantum algorithms to solve optimization problems. We perform the angle encoding techniques by adopting the angles of the rotation gates based on the classical data, which is formulated as:

$$\mathsf{U}_{\text{enc,A}} = \bigotimes_{i=1}^{N_q} \Big[\mathsf{R}_\mathsf{z}\big(\tanh(\mathfrak{Im}(x_i))\pi\big) \cdot \mathsf{R}_\mathsf{y}\big(\tanh(\mathfrak{Re}(x_i))\pi\big) \Big].$$

Since we admit wireless channel information as classical data input (defined in Section IV), we accordingly separate each complex-valued data point into two entries, the real and the imaginary components, encoded into a single qubit by applying two subsequent gates $\mathsf{R}_\mathsf{z}(\cdot)$ and $\mathsf{R}_\mathsf{y}(\cdot)$, respectively, instead of encoding one feature per qubit. Each data entry is first normalized into a typical quantum angle range using $\tanh(\cdot)\pi$, to avoid unnecessary rotations. The state after the encoding process, therefore, can thus be written as $|\psi_{\text{enc;A}}\rangle = \mathsf{U}_{\text{enc;A}}|0\rangle^{\otimes N_q}$, where N_q represents the number of qubits. For simplicity, this work assumes that the qubits are initialized in the state $|0\rangle$, typical in quantum circuit models.

For the parameterized quantum circuit, we train the encoded state using

$$\mathsf{U}_{\text{pqc,A}}(\boldsymbol{\theta}_\text{A}) = \left[\bigotimes_{i=1}^{N_q} \mathsf{R}_\mathsf{y}\left(\tanh(\theta_{\text{A},i})\right) \right] \Big[\mathsf{C}_\mathsf{x}(|\psi_{N_q}\rangle, |\psi_1\rangle) \Big(\prod_{i=1}^{N_q-1} \mathsf{C}_\mathsf{x}(|\psi_i\rangle, |\psi_{i+1}\rangle) \Big) \Big] \tag{15}$$

as a learning circuit, where $\boldsymbol{\theta}_\text{A}$ is a set of parameters vector. During the training phase, we initialized the parameters randomly with $\boldsymbol{\theta}_\text{A} \sim \mathcal{U}[-\pi, \pi]$. In the PQC circuit, we apply the same learning operation circuit blocks, rotation and entangling blocks, which are required to influence the model output. To alleviate the issue of barren plateaus, indicated by the vanishing variance of the training loss gradients, the learning circuit dimension should lie in a sweet spot between being excessively simple and excessively expressive, in terms of the Lie algebra dimension.

The last part of the connection is the MEC unit circuit. We express this circuit as an execution circuit since its outputs are directly used as optimized outputs for the wireless system model. The MEC unit circuit comprises the encoding and the entangling parts, which can be written as

$$\mathsf{U}_\text{B} = \Big[\otimes_{i=1}^{N_q} \mathsf{R}_\mathsf{z}\big(\tanh(e^*_{\text{QCR},i})\pi\big) \Big] \left[\prod_{i=1}^{N_q-1} \mathsf{C}_\mathsf{x}(|\psi_i\rangle, |\psi_{i+1}\rangle) \right], \tag{16}$$

where $e^*_{\text{qcr};i}$ denotes a reconstruction value from the QCR's output, used as input at the i-th qubit at the MEC unit, representing the quantization and de-quantization process (detailed in the following paragraph). While the MEC unit circuit could be extended into a learnable design, similar to a PQC circuit, representing a multi-stage optimization approach [41], it does not contain a learning component for generalization purposes in this work.

To obtain task-related classical outputs from the quantum state, we perform quantum measurements and evaluate the expectation value of an observable $\hat{M}$. Since the required outputs as solutions are bounded within the range of the continuous value $[0, 1]$ (detailed in Section IV-B), we define the projection operator as $\hat{M} = |1\rangle\langle 1|$, applied to both the central and the MEC unit circuits. The resulting expectation value $\langle\psi|\hat{M}|\psi\rangle$ consequently corresponds to the probability of the target state. In particular, we quantify $e_{\text{A},i} \leftarrow \langle \hat{M}_{\text{A},i}\rangle$, which represent the probability of measuring state $|1\rangle$ at the i-th central unit. This captures the desired information transmitted through quantum communication. The presented QCCL, however, involves a quantization process to map this value into a sequence of classical bits, for transmission via the superdense protocol, as discussed in Section III-A.

In this work, the quantization process essentially transforms the value according to the quantization level γ into binary data. The value of γ depends on the chosen quantization strategy, and leads to an overhead in the number of qubits when the level is high. For example, using $\gamma = 2$, we can transform $e_i < 0.5$ into $c = 0$, while $e_i \geq 0.5$ is quantized into $c = 1$. This requires the same number of qubits as the number of e_i values due to the low quantization level. However, if the quantization resolution is too coarse, it leads to less precise reconstruction of the QCR outputs back to continuous values (see Eq. (16)). The reconstruction process represents a linear mapping from a single bit or a sequence of bits into specified continuous values, akin to a code-book mapping, e.g., $00 \rightarrow 0.25$ and $10 \rightarrow 0.75$.

IV. Wireless System Model and Problem Formulation

To analyze QCCL utility in improving wireless system performance, this section describes a WIoE network model, with particular emphasis on a distributed transmitter system with centralized power control. Subsequently, we consider the practical limitations and formulate the optimization problem for analyzing the performance of the QCCL workflow.

A. WIoE Network with Centralized Power Coordination

This work considers a network with L distributed cooperative access points (APs) jointly serving K user equipments (UEs) distributed across the network. Each AP and UE is assumed to be equipped with M antennas and a single antenna, respectively, and thus $M \geq K$. Regarding QCCL integration into WIoE networks: the APs are connected to the MEC unit via high-capacity links, which enable user-data sharing and coherent joint transmission. Therefore, we can model the channel between the l-th AP and the k-th UE as

$$\mathbf{h}_{k,l} = \sqrt{\beta_{k,l}}\mathbf{R}_{k,l}^{1/2}\mathbf{g}_{k,l} \in \mathbb{C}^{M\times 1}, \tag{17}$$

Algorithm 1 Training Procedure of the Proposed QCCL Workflow

Require: Acquired wireless channel information $\hat{\mathbf{H}}$; predefined QCR bit-label set $\mathcal{C}_{\text{QCR}}$; quantization level γ
Ensure: Trained parameters $\boldsymbol{\theta}_{\text{A}}^{*}$, $\boldsymbol{\theta}_{\text{QCR}}^{*}$, and optimized t^{*}
Stage 1: QCR pre-training
Initialize QCR parameters $\boldsymbol{\theta}_{\text{QCR}}$
for each QCR training iteration **do**
    Transmit predefined bit labels $c \in \mathcal{C}_{\text{QCR}}$ through superdense coding under quantum-channel and Bell-operation imperfections
    Update $\boldsymbol{\theta}_{\text{QCR}}$ by minimizing the decoding loss $\mathcal{L}_{\text{QCR}}$
end for
Fix the trained QCR parameters $\boldsymbol{\theta}_{\text{QCR}}^{*}$
Stage 2: WIoE optimization
Initialize the central-unit parameters $\boldsymbol{\theta}_{\text{A}}$ and t
for each WIoE training iteration **do**
    Encode $\hat{\mathbf{H}}$ at the central unit, obtain measurement outputs, and quantize the outputs using γ.
    Transmit the quantized value through the noisy quantum link
    Apply the QCR unit with the trained parameters and reconstruct the corrected values at the MEC unit
    Obtain the power-control variables and evaluate the WIoE loss $\mathcal{L}_{\text{WIoE}}$
    Update $\boldsymbol{\theta}_{\text{A}}$ using the parameter-shift rule
    Apply the updated $\boldsymbol{\theta}_{\text{A}}$. Obtain the MEC unit's output to update t using its gradient
end for
return $\boldsymbol{\theta}_{\text{A}}^{*}$, $\boldsymbol{\theta}_{\text{QCR}}^{*}$, and t^{*}

where $\beta_{k,l}$ denotes the large-scale channel gain component that captures distance-dependent path loss and shadowing effects, $\mathbf{g}_{k,l} \sim \mathcal{CN}(0, \mathbf{I}_M)$ marks the small-scale fading random component that captures channel characteristics such as reflections or scatterers, primarily due to multipath links and other factors, and $\mathbf{R}_{k,l}^{1/2}$ represents a correlated component across element antennas. For the large-scale component, we model it as $\beta_{k,l} = 10^{-\frac{\text{PL}_{k,l}}{10}}$, where the path loss model is given by $\text{PL}_{k,l} = \text{PL}_0 + 10 e_{\text{PL}} \log_{10}(\frac{d_{k,l}}{d_0})$, with $d_{k,l}$ denoting the distance between the k-th UE and the l-th AP, PL_0 denoting reference path loss, d_0 denoting reference distance, and e_{PL} denoting the path loss exponent [42]. We model the spatial correlation across the antenna elements of a uniform linear array (ULA), at the l-th AP, using the following structured correlation matrix: $[\bar{\mathbf{R}}_{k,l}]_{m,n} = r_{k,l}^{|m-n|}$, where $r_{k,l} \in [0, 1)$ represents the correlation coefficient associated with the l-th AP and the k-th UE, in which a larger value indicates a higher correlation level. The channel correlation can be captured as $\mathbf{R}_{k,l} \triangleq \mathbb{E}[\mathbf{h}_{k,l}\mathbf{h}_{k,l}^H] \in \mathbb{C}^{M \times M}$. We can also define the covariance matrix of the channel model expressed in Eq. (17) as the product of the large-scale and structure components, as follows.

$$\mathbf{R}_{k,l} = \beta_{k,l} \bar{\mathbf{R}}_{\text{norm}}, \tag{18}$$

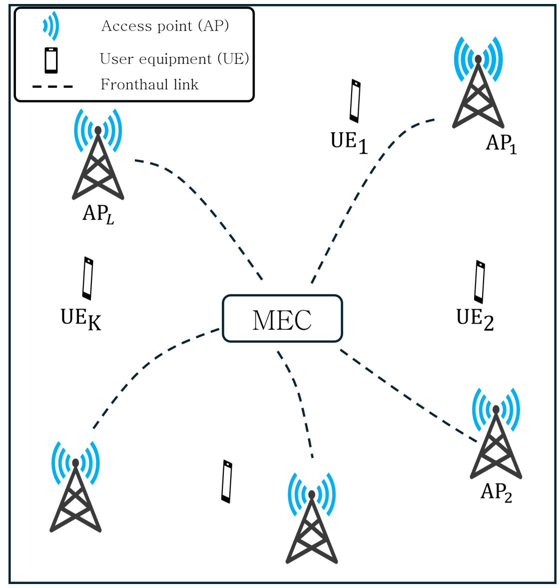

Fig. 3: A depiction of the assumed WIoE network, featuring an MEC for centralized power coordination. L APs are deployed to serve K UEs.

where $\bar{\mathbf{R}}_{\text{norm}} = \bar{\mathbf{R}}_{k,l}(\frac{M}{\mathbf{Tr}(\bar{\mathbf{R}}_{k,l})})$ conveys the normalized structural matrix. The normalization maintains $\text{Tr}(\mathbf{R}_{k,l}) = M\beta_{k,l}$, capturing the total average power, which is well explained in [43]. Since $\mathbf{R}_{k,l} \succeq 0$ is a positive semi-definite matrix, we can perform eigen-decomposition $\mathbf{R}_{k,l} = \mathbf{U}_{k,l}\boldsymbol{\Lambda}_{k,l}\mathbf{U}_{k,l}^H$, to construct a valid square-root matrix $\mathbf{R}_{k,l}^{1/2}$ in Eq. (17), by computing $\mathbf{R}_{k,l}^{1/2} = \mathbf{U}_{k,l}\boldsymbol{\Lambda}_{k,l}^{1/2}\mathbf{U}_{k,l}^H$. Here, $\mathbf{U}_{k,l}$ represents a unitary matrix that contains the corresponding eigenvectors.

Fig. 3 illustrates the assumed WIoE network model, where L distributed APs are employed to serve K users. The figure also depicts a single MEC unit connected to the central unit for evaluating the QCCL model, which is defined in Sec. III. As detailed therein, our QCCL model can be extended to incorporate multiple MEC units, where each unit is linked to the central unit through quantum communication protocols (see Fig. 1). The MEC unit in the WIoE network is responsible for controlling the transmit power of all APs. We assume that each AP stores data pertinent to the UEs. In this context, fronthaul links connect the MEC with all APs, to facilitate data sharing. Centralized multi-AP coordination is also considered in emerging WLANs, where a local controller exchanges channel and control information with connected APs over high-capacity links to coordinate their transmissions [44]. Additionally, the discussion on how the MEC unit coordinates transmit power levels for all APs in the system can be found in Section III-B. Such a network model is commonly adopted in scenarios such as cell-free multiple-input multiple-output (MIMO) systems, where the access points are distributed to provide better fairness to all users [45]. Data sharing among users can also enable coherent joint transmission, providing beamforming gains across the network [46].

In the downlink scenario, each AP transmits a superposition signal to K UEs simultaneously. With respect to the l-th AP, such a transmitted signal can be represented as

$$\mathbf{x}_l = \sum_{k=1}^{K} \mathbf{w}_{k,l} s_k \ \in \mathbb{C}^{M\times 1}, \tag{19}$$

where s_k denotes the symbol designated to the k-th UE, with $\mathbb{E}[|s_k|^2] = 1$. The notation $\mathbf{w}_{k,l} \in \mathbb{C}^{M\times 1}$ denotes a beamforming vector intended to the k-th UE, by the l-th AP, and can be decomposed as

$$\mathbf{w}_{k,l} = \sqrt{p_{k,l}}\mathbf{v}_{k,l}. \tag{20}$$

Here, $\mathbf{v}_{k,l}$ denotes the corresponding beamforming vector at the k-th UE and the l-th AP; hence, the term $\|\mathbf{v}_{k,l}\|_2 = 1$ conveys the corresponding beam direction. The term $p_{k,l}$ indicates the transmit power associated with the l-th AP and the k-th UE, where it must be ≥ 0; the total power at the l-th AP must satisfy $\sum_k^K p_{k,l} \leq \hat{P}_l^{\max}$, where $\hat{P}_l^{\max}$ denotes the maximum transmit power. This paper assumes that the beam directions are set using the zero-forcing technique, for the sake of generality. Accordingly, the channel matrix between the l-th AP and K users can be defined as $\mathbf{H}_l = [\mathbf{h}_{1,l}, \mathbf{h}_{2,l}, \ldots, \mathbf{h}_{K,l}] \in \mathbb{C}^{M\times K}$. We can then compute the zero-forcing beam direction matrix of the l-th AP to K UEs as $\mathbf{V}_l = \mathbf{H}_l(\mathbf{H}_l^H\mathbf{H}_l)^{-1} \in \mathbb{C}^{M\times K}$. In addition, to preserve unit-norm beamforming vectors, each (k,l)-th beamformer can be normalized as follows: $\mathbf{v}_{k,l} = \frac{\tilde{\mathbf{v}}_{k,l}}{\|\tilde{\mathbf{v}}_{k,l}\|}$. Therefore, the received signal at the k-th user can be described as

$$\begin{aligned} y_k &= \sum_{l=1}^{L} \mathbf{h}_{k,l}^H \mathbf{x}_l + n_k \\ &= \underbrace{\sum_{l=1}^{L} \mathbf{h}_{k,l}^H \mathbf{w}_{k,l} s_k}_{\text{desired signals}} + \underbrace{\sum_{i\neq k}\sum_{l=1}^{L} \mathbf{h}_{k,l}^H \mathbf{w}_{i,l} s_i}_{\text{interference signals}} + n_k, \end{aligned} \tag{21}$$

where $n_k \sim \mathcal{CN}(0,\sigma^2)$, denotes the additive white Gaussian noise, σ^2 is the noise variance. Since the received signal contains the desired signal and interference from other users, the data rate of the k-th user shall depend on the received signal-to-interference-plus-noise ratio (SINR), which can be written as

$$\text{SINR}_k = \frac{|\sum_{l=1}^{L} \mathbf{h}_{k,l}^H \mathbf{w}_{k,l}|^2}{\sum_{i\neq k} |\sum_{l=1}^{L} \mathbf{h}_{k,l}^H \mathbf{w}_{i,l}|^2 + \sigma^2}. \tag{22}$$

The rate of the k-th user, therefore, is given by

$$R_k = \log_2\left(1 + \text{SINR}_k\right). \tag{23}$$

This work also considers imperfect channel estimation processes to account for limitations in practical deployments, such as limited pilot resources. Therefore, as we assume the presence of residual error, the actual channel can be modeled as

$$h_{k,l} = \hat{h}_{k,l} + e_{k,l}, \tag{24}$$

where $e_{k,l} \sim \mathcal{CN}(0,\sigma_e^2)$ indicates the stochastic error, with variance σ_e^2, and $\hat{h}_{k,l}$ denotes the estimated channel state information [43]. Accordingly, as each AP applies the zero-forcing technique based on the estimated channel state information, perfect interference nulling cannot be guaranteed under imperfect channel information. Each l-th AP then computes $\hat{\mathbf{H}}_l = [\hat{\mathbf{h}}_{1,l}, \hat{\mathbf{h}}_{2,l}, \ldots, \hat{\mathbf{h}}_{K,l}] \in \mathbb{C}^{M\times K}$. The beam-direction matrix can then be obtained as $\hat{\mathbf{V}}_l = \hat{\mathbf{H}}_l(\hat{\mathbf{H}}_l^H\hat{\mathbf{H}}_l)^{-1} \in \mathbb{C}^{M\times K}$, with each column normalized to construct $v_{k,l}$. Thus, for $h_{k,l}^H v_{i,l} \neq 0, \forall i \neq k$, the residual inter-user interference remains in the received SINR.

B. Max-Min Problem Formulation

The objective of the WIoE optimization is to maximize the minimum user rate by jointly optimizing AP transmit power levels. Such an objective emphasizes user fairness while promoting overall rate improvement without neglecting the performance of users experiencing poor channel conditions, which is crucial for communication reliability. We formulate this objective subject to AP power restriction and service constraint, as follows:

$$\begin{aligned} \underset{p_{k,l},\cdots,p_{K,L}}{\text{maximize}} \quad & \min_{k=1,\ldots,K} R_k \qquad , \\ \text{subject to} \quad & p_l \leq \hat{P}_l^{\max}, \forall l \in \{1,\cdots,L\}, \\ & R_k \geq R_k^{\min}, \forall k \in \{1,\cdots,K\}, \end{aligned} \tag{25}$$

where $p_l = \sum_k p_{k,l}$. The first constraint dictates that the selected power levels at each AP do not go beyond the maximum allowable power level. The second constraint ensures that the service received by each UE, particularly attained user rate, shall meet a certain threshold, where $R_k^{\min}$ defines the minimum allowable rate.

C. Reformulated Problem

To avoid the non-smooth function $\max\min$, we reformulate our problem in Eq. (25) into an epigraph formula by introducing an auxiliary variable t. The optimization problem becomes:

$$\begin{aligned} \underset{t,\boldsymbol{\theta}_{\text{A}}}{\text{maximize}} \quad & t \\ \text{subject to} \quad & t \leq R_k(\boldsymbol{\theta}_{\text{A}}), \forall k \in K, \\ & R_k^{\min} - R_k(\boldsymbol{\theta}_A) \leq 0, \forall k \in K, \\ & 0 \leq p_l(\boldsymbol{\theta}_{\text{A}}) \leq \hat{P}_l^{\max}, \forall l \in L. \end{aligned} \tag{26}$$

Applying the first constraint, we maximize $\min[R_k(\boldsymbol{\theta}_{\text{A}})]$ via the variable t. Using this form, our problem becomes easier to solve while remaining aligned with the original problem. We also redefine the second constraint as a standard inequality so that we can construct the training loss using regularization terms. As our QCCL model employs a gradient-based method to optimize the variables, we transform the constrained problem into an unconstrained one by adding penalty terms based on the constraints. The penalized objective function can be written as

$$\begin{aligned} f(\boldsymbol{\theta}_{\text{A}}, t) = t - \lambda_1 \sum_{k=1}^{K} \left[\max\left(0,\, t - R_k(\boldsymbol{\theta}_{\text{A}})\right)\right]^2 - \\ \lambda_2 \sum_{k=1}^{K} \left[\max\left(0,\, R_k^{\min} - R_k(\boldsymbol{\theta}_{\text{A}})\right)\right]^2, \end{aligned} \tag{27}$$

where λ_1 and λ_2 are the penalty coefficients (larger coefficient values impose greater priority to satisfy the constraint). The penalty term is activated only when the particular constraint conditions are violated, to ensure that the model accounts for the constraints in the original problem. This approach is commonly adopted in learning-based constrained optimization, including constrained reinforcement learning and deep reinforcement learning workflows [47], [48].

V. RESULTS

This section discusses the feasibility of the presented QCCL workflow for optimizing wireless communications, particularly corresponding to WIoE networks. In this wise, it first defines the training loss function and the associated gradient computation. Subsequently, the feasibility of integrating QCR units is demonstrated through numerical results. For this experiment, we use $\hat{\mathbf{H}}$, generated in Section IV, as our unlabeled data during the WIoE optimization. The workflow follows an objective-driven unsupervised learning, to directly optimize the data rate. Predefined labels are used only during QCR pre-training, to evaluate the decoding probability of the noisy superdense-coding links.

A. Training Loss and Gradient Computations

We define a training loss for optimizing the WIoE network based on the problem formulation in Eq. (25) and the reformulated Eq. (27) as

$$\begin{aligned}\mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}_{\mathrm{A}}, t) &= -f(\boldsymbol{\theta}_{\mathrm{A}}, t)\\ &= -t + \lambda_1 \sum_{k=1}^{K} \left[\max\left(0,\, t - R_k(\boldsymbol{\theta}_{\mathrm{A}})\right)\right]^2\\ &+ \lambda_2 \sum_{k=1}^{K} \left[\max\left(0,\, R_k^{\min} - R_k(\boldsymbol{\theta}_{\mathrm{A}})\right)\right]^2. \end{aligned} \tag{28}$$

Since this work is also inspired by gradient descent to optimize the penalized objective, we redefine the objective as the negative of the penalized objective function. The final objective for optimizing our WIoE network is defined as:

$$\min_{\boldsymbol{\theta}_{\mathrm{A}}, t} \mathcal{L}_{\mathrm{WIoE}}. \tag{29}$$

In this stage, we have two different variables to be optimized using the final training loss: t and $\boldsymbol{\theta}_{\mathrm{A}}$. To optimize t, we simply use the first-order iterative optimization technique to update its value. Hence, we need the loss gradient calculation with respect to t. In addition, since t is an explicit variable, and $R_k(\boldsymbol{\theta}_{\mathrm{A}})$ does not depend on it directly, we can compute its gradient as:

$$\frac{\partial \mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}_{\mathrm{A}}, t)}{\partial t} = -1 + 2\lambda_1 \sum_{k:\, t > R_k(\boldsymbol{\theta}_{\mathrm{A}})} \big(t - R_k(\boldsymbol{\theta}_{\mathrm{A}})\big), \tag{30}$$

where the second term is activated if the condition $t > R_k(\boldsymbol{\theta}_{\mathrm{A}})$ is satisfied. For the second variable, as it represents the sequence of weights, as in quantum machine learning, in the quantum circuit we leverage the parameter-shift rule method for parameter updates, given its wide adoption in quantum-enabled learning workflows. As discussed in Section III-B, the outputs from the learnable circuit are used for power allocation. Therefore, we treat the quantum circuit as a model for optimizing the final training loss while optimizing for the power allocation. By doing that, we can obtain, for example, the gradient of a single parameter of the quantum model with respect to the training loss at step ℓ by computing

$$\begin{aligned}\nabla_{\theta^{\ell}_{\mathrm{A},i}} \mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}}, t) = \frac{\partial \mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}})}{\partial (\theta^{\ell}_{\mathrm{A},i})} = \frac{1}{2\sin(s)}\\ \left[\Big(-f_{\theta^{\ell}_{\mathrm{A},i}+s}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}}, t^{\ell})\Big) - \Big(-f_{\theta^{\ell}_{\mathrm{A},i}-s}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}}, t^{\ell})\Big)\right],\end{aligned} \tag{31}$$

where s is a predefined shift coefficient [49]. During training, we also update the parameters of each rotational gate using a first-order gradient-based method. The iterative update for the i-th parameter, $\theta^{\ell+1}_{\mathrm{A},i}$, can be defined as

$$\theta^{\ell+1}_{\mathrm{A},i} = \theta^{\ell}_{\mathrm{A},i} - \alpha \nabla_{\theta^{\ell}_{\mathrm{A},i}} \mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}}, t), \tag{32}$$

where α specifies the learning rate for optimizing the circuit parameters, and $\nabla_{\theta^{\ell}_{\mathrm{A},i}} \mathcal{L}_{\mathrm{WIoE}}(\boldsymbol{\theta}^{\ell}_{\mathrm{A}}, t)$ represents the partial derivative of the training loss function computed at ℓ, with respect to $\theta^{\ell}_{\mathrm{A},i} \in \boldsymbol{\theta}^{\ell}_{\mathrm{A}}$.

The last constraint in Eq. (26) ensures that the total power of a single AP, p_l, cannot exceed the maximum power, $\hat{P}^{\max}_l$. As the output of the quantum circuit is $e_i \leftarrow \langle \hat{M}_i \rangle$, we allocate $p_{k,l} = e_i \cdot \hat{P}^{\max}_l$. To ensure the allocated power always satisfies the constraint conditions, we perform power scaling when the total transmit power of an AP is larger than the maximum power, as in $p_l \leq \hat{P}^{\max}_l$, where $p_l = \sum_k p_{k,l}$. The scaled power allocation can be computed as: $\hat{p}_{k,l} = \hat{P}^{\max}_l \cdot \frac{p_{k,l}}{p_l}$. Otherwise, the original power allocation is used. By doing this, we can guarantee that the last constraint from Eq. (26) is always satisfied during the optimization process. During the learning process, we use an alternating optimization approach, where the first variable is optimized before moving to the subsequent variables. For example, we first set t to an initial value, then perform the learning process to update $\boldsymbol{\theta}_{\mathrm{A}}$ for the quantum model. Subsequently, after updating all quantum model parameters, we update t using classical optimization while simultaneously using the updated quantum model parameters in the training loss. This process is also detailed in Algorithm 1.

As we use encoding states from the Bell states $|\psi_{\mathrm{enc}}\rangle \in \{|\Phi^{\pm}\rangle, |\Psi^{\pm}\rangle\}$ which are orthogonal (see Section III-A), the entropy of each of these states is $S(\cdot) = 0$, as they represent pure states. However, when the subsystems are observed individually, each appears completely random, hence the entropy of the reduced state is $S(\cdot) = 1$. In this wise, neither qubit carries information, as shown in Lemma 1. We can therefore utilize this property, whereby the global system remains perfectly determined while each individual subsystem appears random, improving the security of information exchange. The same argument applies to the other Bell states. For example, for $|\Psi^{+}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$, the reduced state, $\rho_A = \mathrm{Tr}_B(|\Psi^{+}\rangle\langle\Psi^{+}|) = \frac{1}{2}I$, also gives $S(\rho_A) = 1$. Hence, Bell states yield maximally mixed reduced state and the individual qubit carries no accessible information by itself.

Lemma 1. *The entropy of the reduced state corresponding* *to a maximally entangled state is inherently equal to* 1, *representing completely random information.*

Proof. Consider $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ as the encoded state. The density matrix of the composite system is thereby $\rho_{AB} = |\Phi^+\rangle\langle\Phi^+|$. We understand, from Eq. (8), that $S(\rho_{AB}) = 0$ because the state is pure. Assume that an attacker can only observe one individual subsystem, for the qubits are transmitted individually. The reduced state of ρ_{AB} is therefore $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \frac{1}{2}\cdot\mathbb{I}$, with eigenvalues $\{\lambda_1 = \lambda_2 = \frac{1}{2}\}$. The entropy of this state can thus be calculated as $S(\rho_A) = 1$. ■

B. Performance results

The following details the parameters associated with coherence errors in the Bell operation: Regarding the detuning error in the Pauli-Z operation, denoted by Z, we assume the unwanted frequency drift is set as $\delta = 2\pi\cdot 1$ MHz. In this abstraction, we set the error associated with amplitude miscalibration as $e_{\text{amp}} = 10\%$. Regarding the construction of the controlled-Z gate, we consider that the detuning parameters, affecting both the target and control qubits, are equal to the previously set δ. Regarding the miscalibration error in this abstraction, we assume a 5% over-rotation error, and the quantization level used in the QCCL transmission process is set to $\gamma = 4$. As we focus on the optimization performance, we assume certain physical or hardware conditions. For example, rather than expressing these values in Watts, we adopt normalized quantities relative to noise power. More precisely, we assume a ratio between the transmit and noise power values, which, under ideal conditions, with a noise power of $\sigma^2 = 1$ and a maximum power of $\hat{P}_l^{\max} = 200$, corresponds to a signal-to-noise ratio of around 23 dB. Accordingly, we normalize the channel coefficient in Eq. (17) to obtain consistent rate calculation under such normalization. To do so, we normalize the channel by considering effective channel gains to preserve the physical characteristics, such as frequency and distance, during simulation. The effective channel gain is computed from the aggregated channel gains of each user: $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K]$, where $\eta_k = \sum_l |h_{k,l}|^2$. Therefore, the effective channel gain and the normalized channel coefficient can be written as

$$\eta_{\text{eff}} = \text{median}(\boldsymbol{\eta}) \tag{33}$$

and $h'_{k,l,m} = \frac{h_{k,l,m}}{\sqrt{\eta_{\text{eff}}}}$, respectively, where $h_{k,l,m}$ denotes the m-th element of $\mathbf{h}_{k,l}$, $\forall k$, $\forall l$. During simulation, we consider a distributed cooperative network with $L = 4$ APs and $K = 3$ UEs, where each AP is equipped with $M = 8$ antenna elements. The APs and UEs are independently placed at random within a 200 m $\times$ 200 m area, following a uniform distribution. In generating the wireless channel, we use a path-loss exponent of $e_{\text{PL}} = 2.3$ and include log-normal shadowing with a standard deviation of 7 dB. For the spatial correlation model, we set the correlation coefficient to $r_{k,l} = 0.75$, representing a moderately high correlation among the antenna elements at each AP. For the constraint-quality service, the minimum rate is defined by a target of 1 bps/Hz, which is converted to the corresponding minimum rate based on the adopted bandwidth.

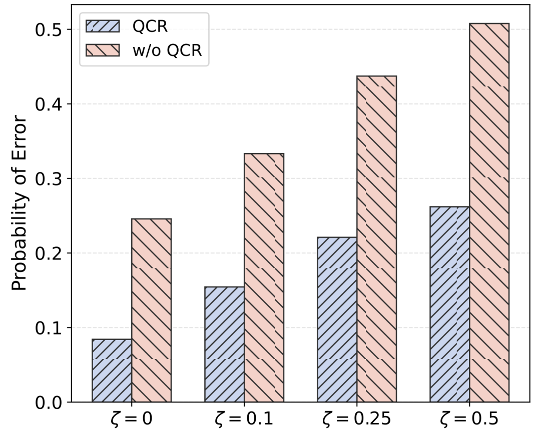

Fig. 4: Decoding error probability of the quantum-connected link with and without the proposed QCR unit under different bit-flip error levels.

Figure 4 demonstrates the link-level benefits of integrating a QCR unit into the QCCL workflow. Here, the estimated decoding error probabilities under different bit-flip error levels ζ are depicted; the error probability is obtained by averaging the decoding errors observed across simulation trials. The result compares performance under the imperfection-aware quantum communication for two cases: a quantum-connected link with and without the proposed QCR unit. Here, increasing ζ leads to a higher decoding error probability since the transmitted Bell-state correlations are more frequently disturbed. Nevertheless, the QCR-assisted case consistently achieves lower error probability across all considered noise levels. For example, employing the QCR unit reduces the decoding error probability at $\zeta = 0.25$ by approximately 50.41% compared to the case without the QCR-assisted. This quantitative reduction confirms that the QCR improves the reliability of the decoded information, or in other words, the learned QCR circuit can partially compensate for quantum-channel and device-induced imperfections, before it is used by the MEC unit for power-control optimization.

Figure 5 illustrates the training loss in Eq. (28) for optimizing the WIoE network under different levels of quantum-channel imperfection. While the training loss decreases noticeably over the training iterations, the loss exhibits fluctuations as the number of iterations increases. This indicates that learning is occurring despite imperfect conditions, and the fluctuation behavior arises from penalty terms and quantum randomness. For illustration, when the optimization approaches the feasible boundary, small variations in the decoded-and-reconstructed values from the QCR can slightly change the user rates $R_k(\theta_A)$, which may activate or deactivate the penalty terms associated with $(t - R_k(\theta_A))$ and $(R_k^{\min} - R_k(\theta_A))$. These variations can occur due to noisy quantum communication, finite-shot measurements, and quantization followed by reconstruction into continuous values.

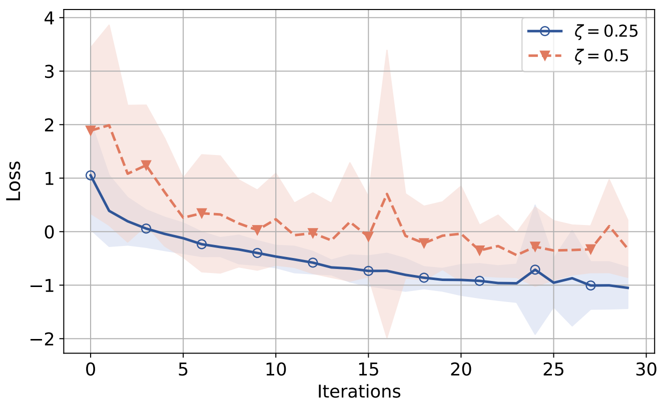

Fig. 5: Training-loss evolution of the proposed QCR-assisted QCCL workflow under mild and severe quantum-channel imperfections.

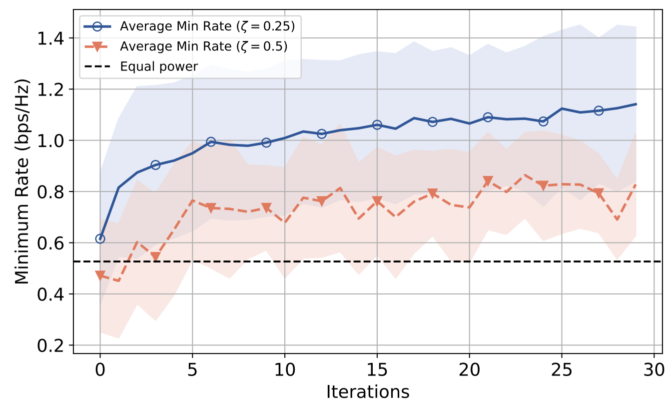

Fig. 6: Lowest user rate attained by the proposed QCCL workflow compared with equal-power allocation.

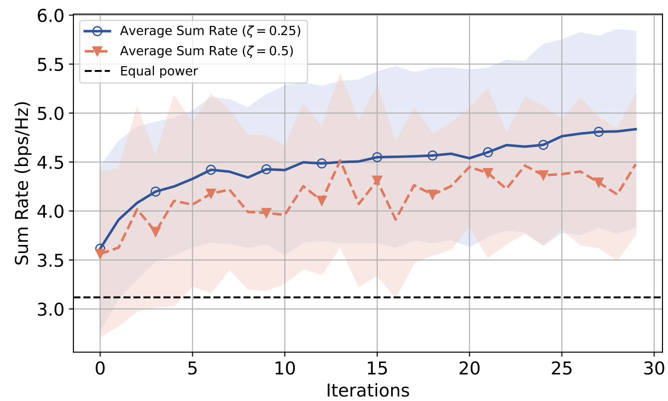

Fig. 7: Sum rate performance of the proposed QCCL workflow compared with equal-power allocation.

As a result, the training loss is not monotonic, but the learning remains stable under mild-noise ($\zeta = 0.25$) setting. Compared with the severe noise when ($\zeta = 0.5$), the loss has a larger variance and higher value, but still maintains a learning process, demonstrating the resilience of the proposed workflow.

While the results in Fig. 5 indicate that the proposed QCCL workflow can continue to be useful under imperfect quantum communication, this benefit is accompanied by practical trade-offs. In particular, the use of shared entangled Bell pairs is required before transmission via the superdense protocol. The shared pairs currently introduce practical challenges such as resource overhead, synchronization requirements, and possible latency, especially in larger quantum-connected networks, which is a common consideration in entanglement routing studies [50]. In addition, the purpose of these results is to evaluate the effectiveness of the proposed QCR-assisted QCCL workflow under imperfect quantum communication. Equal-power allocation is therefore used as a non-learning baseline to verify whether the proposed workflow can learn useful power-control behavior while mitigating quantum-channel and device imperfections.

Figures 6 and 7 further evaluate the system-level performance of the proposed QCCL workflow in terms of the minimum user rate and the sum rate, respectively. In Fig. 6, the proposed workflow improves the minimum user rate over iterations under both severe and mild noise settings. Comparison with the equal-power baseline is also included. While some fluctuations are observed during training, the model can still allocate power control more effectively than the baseline. As a result, the weaker user, with severe wireless channel conditions, is still supported. This promotes rate fairness. In addition, Figure 7 complements this observation by showing the aggregate network performance. A common concern in $\max\text{-}\min$ optimization is that improving the worst-user rate may reduce the overall sum rate. The overall sum-rate trend for both settings, however, maintains a higher sum rate than the equal-power baseline.

The impact of quantum-channel imperfections can also be observed by comparing the mild and severe noise settings. When the bit-flip probability increases from $\zeta = 0.25$ to $\zeta = 0.5$, the decoded information becomes less reliable, which degrades the quality of the power-control decisions at the MEC unit. This explains the decline observed in both the minimum-rate and sum-rate curves. Nevertheless, the proposed workflow remains competitive relative to equal-power allocation.

VI. Conclusion and Future Works

This paper employs quantum learning models to address imperfections in quantum communication channels and deployment while maximizing WIoE performance, in terms of user rate. It presents QCCL workflows in which distributed nodes, each capable of executing quantum computing and communications, host separate quantum learning models, connected via quantum communication protocols, leveraging quantum superdense coding. Such separated yet connected architectures facilitate improved security and computational load balancing among nodes. However, quantum communication channels are expected to be affected by imperfections, such as bit-flip errors and coherence errors, which compromise state fidelity and disrupt training progression. In response, this paper introduces the use of quantum-classical router units capable of learning and adapting to quantum communication channel conditions. It computes parameterized circuits tasked with minimizing the information misalignment caused by such errors. Numerical results confirm the effectiveness of this approach. Further, as this study explores the integration of quantum communication protocols with variational quantum circuits, prospective research paths can thus be investigated as follows.

Future studies may also discuss the utility of quantum entanglement routing within networks of quantum communication nodes. Akin to networks of classical communications, various key performance indicators can be improved through quantum-enabled learning approaches. In particular, we may aim to reduce latency or balance computational and communication burdens across nodes. In addition, entanglement routing potentially involves entanglement swapping, with approaches such as heralded swapping, which requires nodes to perform swapping operations according to the swapping outcomes of other nodes, or unheralded swapping, which does not require waiting for such outcomes [50]. It is worth mentioning that swapping operations inherently have certain success probabilities, which might need to be accounted for in optimization or analysis.

At the link level, it is possible to design quantum protocols that can accommodate multi-user communication. With respect to QCCL, such protocols facilitate aggregation-based distributed learning approaches, whereby training outcomes of several edge units are aggregated at central nodes, leveraging multiple entanglement links. These protocols may additionally involve multi-user query management and key exchanges as well, as highlighted in studies such as [51]. Furthermore, given the increasing relevance of joint quantum-classical communications, multi-user protocols exploiting quantum-classical communication channels can be developed as well, accommodating functionalities such as interference mitigation.

Finally, we concur that the presented QCCL workflow is not restricted to WIoE performance maximization, and can be useful for the other implementations as well: Notably, when considering large-scale quantum-wireless networks, graph representation and learning can be used to jointly optimize network topology and associated performance metrics. For instance, quantum-enabled graph representation can be used to identify node activation sequences and routing patterns that yield higher success transmission probability, subject to qubit processing capability thresholds and state fidelity. Furthermore, abstraction learned from large-scale network topologies can be adopted to smaller networks, using approaches such as relational knowledge distillation [52]. For example, through such methods, once a particular quantum learning model identifies the favorable configurations of a wide-area network, this representation can be distilled to construct small-scale models, which in turn are used to maximize the performance of local entanglement networks.

References

[1] X. Zhou, A. Shen, S. Hu, W. Ni, X. Wang, and E. Hossain, "Toward quantum-native communication systems: State-of-the-art, trends, and challenges," *IEEE Commun. Surv. Tutor.*, vol. 28, pp. 1436–1482, Jun. 2026.

[2] P. Popovski *et al.*, "1Q: First-generation wireless systems integrating classical and quantum communication," *IEEE Veh. Technol. Mag.*, vol. 20, no. 4, pp. 18–33, Dec. 2025.

[3] P. Sharma, A. Agrawal, V. Bhatia, S. Prakash, and A. K. Mishra, "Quantum key distribution secured optical networks: A survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2049–2083, Aug. 2021.

[4] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, "Towards a distributed quantum computing ecosystem," *IET Quantum Commun.*, vol. 1, no. 1, pp. 3–8, Jul. 2020.

[5] M. M. Waldrop, "More than Moore," *Nature*, vol. 530, no. 7589, pp. 144–148, Feb. 2016.

[6] D. Gottesman, "An introduction to quantum error correction and fault-tolerant quantum computation," Apr. 2009. [Online]. Available: https://arxiv.org/abs/0904.2557

[7] Y. Li, H. Zhu, W. Luo, H. Cai, M. F. Karim, X. Luo, F. Gao, X. Wu, X. Zhou, Q. Song *et al.*, "Realizing ultrahigh capacity quantum superdense coding on quantum photonic chip," *npj Quantum Inf.*, vol. 11, no. 1, p. 49, Mar. 2025.

[8] ITU-R, "Framework and overall objectives of the future development of IMT for 2030 and beyond," International Telecommunication Union, Recommendation ITU-R M.2160-0, Nov. 2023. [Online]. Available: https://www.itu.int/rec/R-REC-M.2160-0-202311-I

[9] M. Eisen, C. Zhang, L. F. O. Chamon, D. D. Lee, and A. Ribeiro, "Learning optimal resource allocations in wireless systems," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2775–2790, May 2019.

[10] S. Verma, T. K. Rodrigues, N. Kato, M. Ariyoshi, and Y. Hasegawa, "Multi layer jamming detection and classification in 6G-enabled satellite-air-ground integrated networks using machine learning," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 8224–8240, Mar. 2026.

[11] J. Gao, C. Zhong, G. Y. Li, and Z. Zhang, "Deep learning-based channel estimation for massive MIMO with hybrid transceivers," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5162–5174, Jul. 2022.

[12] J. Guo, Y. Liu, and A. Nallanathan, "A graph neural network for learning beamforming in pinching antenna systems (PASS)," *IEEE Wireless Commun. Lett.*, vol. 14, no. 12, pp. 3962–3966, Dec. 2025.

[13] I. Batool, M. Fouda, M. Ismail, S. Verma, Z. M. Fadlullah, and N. Kato, "Privacy-embedded spectrum coordination in 6G V2X networks: An attention-driven multi-agent learning approach," *IEEE Netw.*, pp. 1–9, May 2026.

[14] Z. Niu, S. Zhou, and N. Crespi, "Greening 6G," in *Shaping Future 6G Networks: Needs, Impacts, and Technologies*, E. Bertin, N. Crespi, and T. Magedanz, Eds. Wiley-IEEE Press, Nov. 2021, ch. 4, pp. 39–53.

[15] M. Schuld and N. Killoran, "Quantum machine learning in feature Hilbert spaces," *Phys. Rev. Lett.*, vol. 122, no. 4, p. 040504, Feb. 2019.

[16] F. Chaouech *et al.*, "Quantum-based QoE optimization in advanced cellular networks: Integration and cloud gaming use case," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 9604–9618, Nov. 2025.

[17] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018.

[18] B. Narottama, A. U. Haq, J. A. Ansere, N. Simmons, B. Canberk, S. L. Cotton, H. Shin, and T. Q. Duong, "Quantum deep reinforcement learning for digital twin-enabled 6G networks and semantic communications: Considerations for adoption and security," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 2053–2076, Sep. 2025.

[19] W. Zhao, T. Weng, Y. Ruan, Z. Liu, X. Wu, X. Zheng, and N. Kato, "Quantum computing in wireless communications and networking: A tutorial-cum-survey," *IEEE Commun. Surv. Tutor.*, vol. 27, no. 4, pp. 2378–2419, Nov. 2025.

[20] N. Q. T. Thoong, A. A. Cheema, B. Canberk, D. T. Tran, O. A. Dobre, and T. Q. Duong, "Channel estimation for reconfigurable intelligent surface-aided 6G NOMA systems: A quantum machine learning approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 2197–2218, Sep. 2025.

[21] O. Recy, B. Narottama, and T. Q. Duong, "Topology-aware quantum graph neural networks for sum-rate maximization in fluid antenna systems," *IEEE Commun. Lett.*, vol. 30, pp. 912–916, Jan. 2026.

[22] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, and G. Bianchi, "Quantum internet: Networking challenges in distributed quantum computing," *IEEE Netw.*, vol. 34, no. 1, pp. 137–143, Jan. 2020.

[23] M. Fellous-Asiani, J. H. Chai, R. S. Whitney, A. Auffèves, and H. K. Ng, "Limitations in quantum computing from resource constraints," *PRX Quantum*, vol. 2, no. 4, p. 040335, Nov. 2021.

[24] M. R. Uddin, S. Shaon, R. Rahman, D. C. Nguyen, O. A. Dobre, and D. Niyato, "Quantum federated learning: A comprehensive survey," *IEEE Commun. Surv. Tutor.*, vol. 28, pp. 3942–3975, Dec. 2026.

[25] M. Chehimi and W. Saad, "Quantum federated learning with quantum data," in *Proc. 2022 IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2022, pp. 8617–8621.

[26] A. Singh, K. Dev, H. Siljak, H. D. Joshi, and M. Magarini, "Quantum internet—applications, functionalities, enabling technologies, challenges, and research directions," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2218–2247, Sep. 2021.

[27] L. Gyongyosi, S. Imre, and H. V. Nguyen, "A survey on quantum channel capacities," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 2, pp. 1149–1205, Jan. 2018.

[28] M. A. Ferrag *et al.*, "Edge learning for 6G-enabled internet of things: A comprehensive survey of vulnerabilities, datasets, and defenses," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 4, pp. 2654–2713, Sep. 2023.

[29] C. Qiao, M. Li, Y. Liu, and Z. Tian, "Transitioning from federated learning to quantum federated learning in internet of things: A comprehensive survey," *IEEE Commun. Surv. Tutor.*, vol. 27, no. 1, pp. 509–545, Feb. 2025.

[30] Y. Song *et al.*, "A quantum federated learning framework for classical clients," *Sci. China Phys. Mech. Astron.*, vol. 67, no. 5, p. 250311, Mar. 2024.

[31] L. Chen, L. Yan, and S. Zhang, "Robust quantum federated learning with noise," *Phys. Scr.*, vol. 99, no. 7, p. 076003, Jun. 2024.

[32] H. H. S. Chittoor and O. Simeone, "Quantum machine learning for distributed quantum protocols with local operations and noisy classical communications," *Entropy*, vol. 25, no. 2, p. 352, Feb. 2023.

[33] L. Rathi, S. DiAdamo, and A. Shabani, "Quantum autoencoders for learning quantum channel codes," in *Proc. 16th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2024, pp. 988–993.

[34] G. Manavalan and S. Arnon, "Quantum neural network-based compensation of distorted orbital angular momentum beams in complex media," *Sci. Rep.*, vol. 16, p. 1361, Jan. 2026.

[35] S. Sajadimanesh, H. Aghaee Rad, J. Paul Latyr Faye, and E. Atoofian, "NR-QNN: Noise-resilient quantum neural network," *IEEE Access*, vol. 13, pp. 40 185–40 197, Mar. 2025.

[36] G. F. R. Purba, B. Narottama, S. R. Khosravirad, S. L. Cotton, and T. Q. Duong, "Quantum connected collaborative learning with superdense coding for wireless internet-of-everything networks," in *Proc. 59th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2025, pp. 388–393.

[37] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[38] A. Raina and S. G. Srinivasa, "Quantum communication over bit flip channels using entangled bipartite and tripartite states," in *Proc. 52nd Annu. Allerton Conf. Commun., Control, and Comput. (Allerton)*, Oct. 2014, pp. 1368–1375.

[39] S. Imre and L. Gyongyosi, *Advanced quantum communications: an engineering approach*. John Wiley & Sons, 2012.

[40] L. Banchi, J. L. Pereira, S. T. Jose, and O. Simeone, "Statistical complexity of quantum learning," *Adv. Quantum Technol.*, vol. 8, no. 12, p. 2300311, Apr. 2025.

[41] B. Narottama, B. Canberk, S. L. Cotton, H. Shin, G. K. Karagiannidis, and T. Q. Duong, "Non-centralized quantum neural networks for cell-free MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 9971–9985, Dec. 2025.

[42] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.

[43] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, no. 3-4, pp. 154–655, 2017.

[44] S. Verma, T. K. Rodrigues, Y. Kawamoto, M. M. Fouda, and N. Kato, "A survey on multi-AP coordination approaches over emerging WLANs: Future directions and open challenges," *IEEE Commun. Surv. Tutor.*, vol. 26, no. 2, pp. 858–889, 2nd Quarter 2024.

[45] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[46] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

[47] C. Tessler, D. J. Mankowitz, and S. Mannor, "Reward constrained policy optimization," in *Int. Conf. Learn. Represent.*, Dec. 2019. [Online]. Available: https://openreview.net/forum?id=SkfrvsA9FX

[48] S. C. Prabhashana, M.-H. T. Nguyen, V. Sharma, T. X. Vu, B. Canberk, H. Shin, and T. Q. Duong, "Quantum machine learning for wireless-powered UAV positioning in 6G digital twin SAGIN with cooperative nano-satellite constellations," *IEEE J. Sel. Areas Commun.*, vol. 44, pp. 5028–5042, Jun. 2026.

[49] D. Wierichs, J. Izaac, C. Wang, and C. Y.-Y. Lin, "General parameter-shift rules for quantum gradients," *Quantum*, vol. 6, p. 677, Mar. 2022.

[50] A. Abane, M. Cubeddu, V. S. Mai, and A. Battou, "Entanglement routing in quantum networks: A comprehensive survey," *IEEE Trans. Quantum Eng.*, vol. 6, pp. 1–39, Feb. 2025.

[51] A. Mahmoud As'Ad Alnaser, H. M. S. Hatamleh, N. A. Almolhis, S. Duraibi, and Y. Alqahtani, "Secure quantum communication with multi-users in quantum networks," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 2397–2419, Apr. 2025.

[52] W. Park, D. Kim, Y. Lu, and M. Cho, "Relational knowledge distillation," in *Proc. IEEE/CVF Conf. Comput. Vision, Pattern Recog.*, Jun. 2019, pp. 3962–3971.