ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 979-8-3503-4485-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICASSP48485.2024.10447516

QUANTUM FEDERATED LEARNING WITH QUANTUM NETWORKS

Tyler Wang[1], *Huan-Hsin Tseng*[2], *Shinjae Yoo*[2]

[1] Department of Computer Science, Stony Brook University, Stony Brook, NY 11794
[2] Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11793

ABSTRACT

A major concern of deep learning models is the large amount of data that is required to build and train them, much of which is reliant on sensitive and personally identifiable information that is vulnerable to access by third parties. Ideas of using the quantum internet to address this issue have been previously proposed, which would enable fast and completely secure online communications. Previous work has yielded a hybrid quantum-classical transfer learning scheme for classical data and communication with a hub-spoke topology. While quantum communication is secure from eavesdrop attacks and no measurements from quantum to classical translation, due to no cloning theorem, hub-spoke topology is not ideal for quantum communication without quantum memory. Here we seek to improve this model by implementing a decentralized ring topology for the federated learning scheme, where each client is given a portion of the entire dataset and only performs training on that set. We also demonstrate the first successful use of quantum weights for quantum federated learning, which allows us to perform our training entirely in quantum.

1. INTRODUCTION

Federated learning has recently emerged as a major tenet of machine learning and deep learning due to growing privacy concerns associated with the large amounts of user data that these models depend on. Such information should not be accessible to third parties, who, upon penetrating the network, can compromise and leak confidential data (e.g., financial data, medical records, etc.). Federated learning mitigates this concern by focusing on a decentralized computer architecture, which keeps each client's data independent of and private to outsiders. Individuals can train their own models and contribute to the global model without revealing the contents of the model.

Quantum federated learning (QFL) integrates federated learning with quantum machine learning (QML) [1]. The major benefit of QFL is the quantum no-cloning theorem, which states that two identical and independent quantum states may not exist, adding to the security of client data. If a client network were penetrated by an unauthorized third party, they would not be able to copy down the information, as its state would change immediately after, rendering the attacker's information useless.

Prior research in the area of QFL has shown its feasibility and advantages in training large machine learning models. The work in [2] demonstrated a hybrid quantum-classical learning scheme, extracting data from a pre-trained classical neural network and encoding it into a quantum state. These data were then used in the QFL model to make predictions for image classification. Additionally, the authors of [3] develop a QFL framework that utilizes quantum convolutional neural networks for classification tasks. They were also able to generate the first purely quantum federated dataset.

We build on this work by providing the first implementation of quantum weights in QFL, which are communicated using quantum teleportation methods between clients. We evaluate the performance of this model against two equivalent QFL and classical federated learning (CFL) models which use classical weights and classical communication. A general outline of the QFL network is provided in Figure 1.

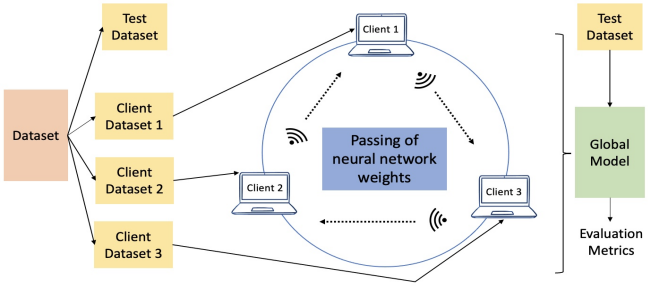


Fig. 1. Decentralized ring QFL framework. The training data is divided among each client, who each trains their portion of the data separately, updating and passing weights to the next client until all training rounds have finished. At the end of training, we use the last client's model (in this example, client 3) and its parameters to evaluate the model's accuracy.

The rest of this paper is organized as follows: Section 2 describes the design of our QFL models. Section 3 shows the

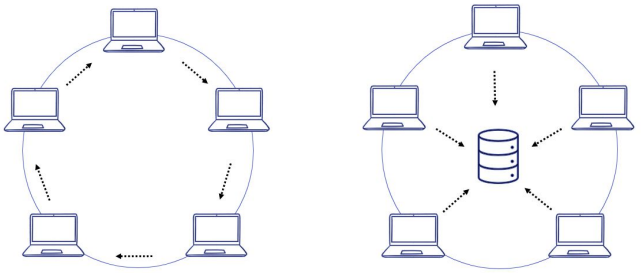
This work was supported by the U.S. DOE, under award DE-SC-0012704 and BNL's LDRD #24-061. This project also was supported in part by the DOE-SC's Office of WDTS under the SULI. This research used resources of the NERSC, under Contract No. DE-AC02-05CH11231 using NERSC award HEP-ERCAP0023403.

performance of all three federated learning models. Section 4 contains further discussions about these results. We draw conclusions in Section 5.

2. PROBLEM SETUP

2.1. Federated Learning Topology

For both the classical and quantum neural networks, we use a *ring network topology*, which has been shown to yield better accuracy of federated learning in neural networks [4, 5, 6, 7, 8]. Each client in the ring only obtains a subset of the entire dataset and performs their training only on that set; additionally, only one set of parameters is needed for training. Once one client has finished its round, the updated parameters are passed to the next client for use, and so on; once completed, these parameters are used along with the test data to compute the accuracy of the model. This differs from a hub-spoke topology, which includes a central server node that performs an averaging process to update the parameters after each training round, which are then distributed to all clients in the network. These two network architectures are contrasted in Figure 2.



(a) Ring Topology (b) Hub-Spoke Topology

Fig. 2. Topology of a ring network (a) compared to a hub-spoke network (b). Note that the hub-spoke topology contains an additional central server node to which clients pass their parameters after training, while in the ring topology, clients simply exchange parameters with other clients.

2.2. Quantum Neural Networks

Quantum neural networks (QNN), or variational quantum circuits (VQC), contains circuit parameters that can be updated using gradient-based methods. Recent research underscores the superior expressiveness of VQCs compared to classical neural networks [9, 10, 11, 12]. Moreover, VQCs exhibit efficacy in training with smaller datasets, as demonstrated by Caro et al. [13]. VQC applications span various domains in QML, including classification, reinforcement learning, natural language processing, and sequence modeling [14, 15, 16, 3, 17, 18, 19, 20, 21, 22, 23, 24].

The VQC contains three essential parts: a quantum encoder that provides a routine for encoding classical data into a quantum state as well as a variational quantum circuit block with the learnable parameters. The final part of the VQC consists of measuring the qubits, which gives classical values that can further be processed in other quantum or classical components during the training and inference process.

2.2.1. *Quantum Encoder*

To train our QFL model on a classical dataset, we use the *variational encoding* method to transform the classical data into quantum states. As seen in Figure 3, the encoding part consists of R_x gates, which rotate the qubit along the x-axis by an angle specified by the input data x_i.

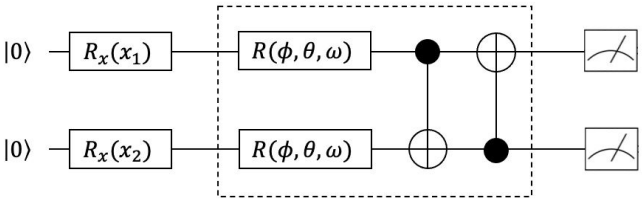


Fig. 3. Architecture for the variational quantum circuit. The encoding layer of the VQC consists of single-qubit rotational R_x gates, whose rotation angles are given by the input data. The grouped box represents the variational quantum layer, which consists of general parametrized unitary gates $R(\phi, \theta, \omega)$. These parameters ϕ, θ, and ω are subject to the optimization procedures. Finally, the quantum measurement component outputs the Pauli-Z expectation values of the qubits.

2.2.2. *Quantum Layer*

The variational layer is the block of the VQC that contains the learnable parameters in the network. We use CNOT gates between each pair of neighboring qubits to entangle the qubits followed by a series of single qubit unitary gates parametrized by ϕ, θ, and ω, see Figure 3. After these operations have been applied, measurements are taken on the two qubits.

2.2.3. *Gradient Calculation Using Parameter Shift*

Given a Hermitian operator B (observable), a fixed encoder $E(x)$ and a unitary gate $R_i \in U(2)$ of free parameters $\theta_i \in \mathbb{R}^3$, we can construct a scalar-valued quantum circuit function defined by

$$\begin{aligned} f(x, \theta_i) &= \langle 0 |\, E^\dagger(x)\, R_i^\dagger(\theta_i)\, B\, R_i(\theta_i)\, E(x)\, |0\rangle \\ &:= \langle x |\, R_i^\dagger(\theta_i)\, B\, R_i(\theta_i)\, |x\rangle \end{aligned} \tag{1}$$

where x is originally a classical data point of vector forms converted into a quantum state denoted as $|x\rangle := E(x)\,|0\rangle$

by a fixed embedding operator $E(x)$. Denoting $M_{\theta_i}(B) := R_i^\dagger(\theta_i)\, B\, R_i(\theta_i)$, the gradient of the scalar function Eq. (1) with respect to θ_i can be written as:

$$\nabla_{\theta_i} f(x, \theta_i) = \langle x | \nabla_{\theta_i} M_{\theta_i}(B) | x \rangle \tag{2}$$

For multiple parameterized gates, consider a product of unitary operators $R_i(\theta_i)$ with learnable parameters θ_i:

$$R(x, \theta) := R_n(\theta_n) R_{n-1}(\theta_{n-1}) ... R_i(\theta_i) ... R_1(\theta_1) R_0(x) \tag{3}$$

Then we may define a scalar function of multiple parameterized gates by $f(x, \theta) = \langle 0 | R^\dagger(x, \theta)\, B\, R(x, \theta) | 0 \rangle$ whose gradient is given by

$$\nabla_{\theta_i} f(x, \theta) = \langle \phi_{i-1} | \nabla_{\theta_i} M_{\theta_i}(B_{i+1}) | \phi_{i-1} \rangle \tag{4}$$

where

$$M_{\theta_i}(B_{i+1}) := R_i^\dagger(\theta_i)\, B_{i+1} R_i(\theta_i)$$
$$B_{i+1} := R_{i+1}^\dagger(\theta_{i+1}) \cdots R_n^\dagger(\theta_n) B\, R_n(\theta_n) \cdots R_{i+1}(\theta_{i+1})$$

and a relationship is to be satisfied:

$$\nabla_{\theta_i} M_{\theta_i}(B) = c\,(M_{\theta_i + s}(B) - M_{\theta_i - s}(B)) \tag{5}$$

This is called the *parameter-shift rule* for updating the parameters θ_i.

2.3. Quantum Weights

Prior QFL frameworks have used classical communication of classical weights between clients and server [2, 3, 25, 26, 27, 28, 29, 30, 31]. In this work, we demonstrate the first QFL model that uses quantum weights and quantum communication. To accomplish this, the weights are transformed into quantum states with a similar encoding process as for the input data of the network, obtaining the expectation values of the qubits. These values are used as the weights in our training and updated with the parameter-shift rule that we previously derived. Then, quantum teleportation sends the parameters to the next client. This process is depicted in Figure 4.

3. RESULTS

We perform a binary classification task on the simple make-circles dataset[1] of 1200 data points, separated into 960 training and 240 testing samples. The training data are equally distributed among three clients for the ring-fashioned training. Each individual runs a model for five local epochs before passing the updated parameters onto the next. The final client is tasked to calculate the model's accuracy with testing data. All of our three models were computed for 100 training rounds with results summarized in Figure 5 and Table 1.

[1]A generated dataset commonly used in the Scikit-Learn package.

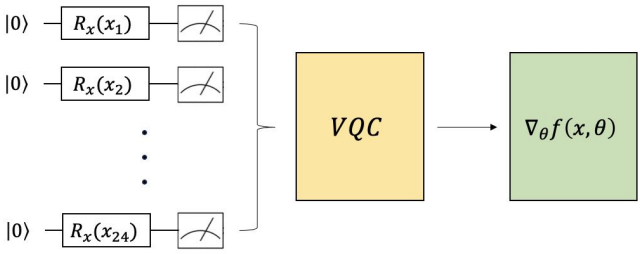


Fig. 4. Implementation of quantum weights. The classical weights are encoded using quantum R_x gates and used in the VQC to train the model. These weights are updated using the gradient $\nabla_\theta f(x, \theta)$, obtained using the parameter-shift rule.

	CFL	QFL (C)	QFL (Q)
Train Loss	0.4445	0.6449	0.6459
Valid. Loss	0.3145	0.6382	0.6423
Test Loss	0.3185	0.6389	0.6435
Test Accuracy	99.58%	100.0%	91.67%

Table 1. Comparison of performance in different federated learning schemes on the make-circles dataset. It is noted that the quantum models obtain higher losses than the classical one.

4. DISCUSSION

4.1. Decentralization

The ring topology intends to offer more security in data communication than the hub-spoke topology. As it was mentioned in [2], without the presence of a central node the risk of external access to the entire dataset can be lowered. For an unauthorized third party to penetrate a client in the ring, they can only gain access to partial data, not the whole one.

This network structure also eliminates the need for quantum memory. Indeed, according to the quantum no-cloning theorem, it would not be possible for the central node of a hub-spoke topology to have several quantum copies for distribution. As a result, the ring topology sharing only one model between clients serves as an alternative to avoid cloning model parameters. Consequently, the lack of quantum memory should not be a concern.

4.2. Quantum Training

One notable contribution of this study is the implementation of quantum weights in QFL, which removes the overlap between the quantum and classical realms. This allows entire model training in the quantum approach to avoid eavesdropping attacks. Moreover, the communication error between the quantum and classical systems can also be reduced.

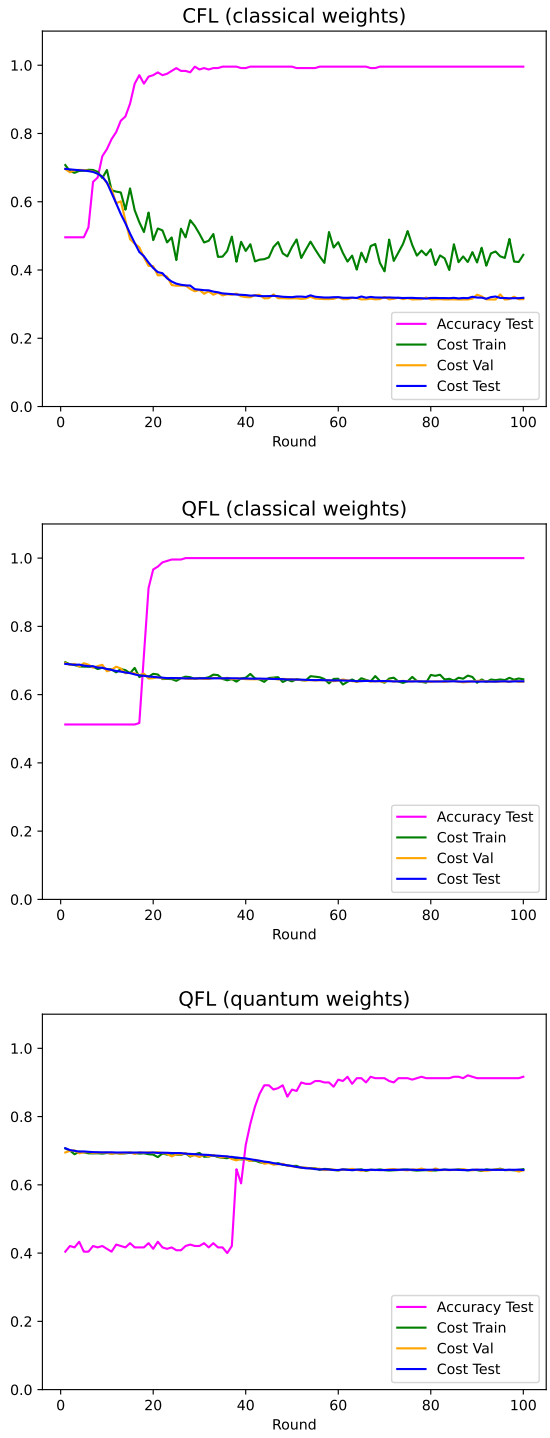


Fig. 5. Results. A slight quantum advantage can be seen between the classical communication models in (a) and (b), while the quantum communication takes much longer to converge as seen by graph (c). This is expected, since we only run in a simulated environment, with the assumption that this model will perform similarly on a real quantum computer once we have access to one.

However, the noise level of current quantum computers in the noisy intermediate-scale quantum (NISQ) era remains acknowledged. Since the noise level tends to increase as the number of qubits increases, designing a QFL model using large qubits will be more difficult.

4.3. Future Research

4.3.1. Quantum Teleportation

We only mimic the use of quantum teleportation for the weights in our model; these techniques, while feasible, do not truly demonstrate teleportation of quantum weights. The authors of the work in [32] have developed a novel quantum network simulation framework that offers more powerful quantum capabilities, allowing for a more accurate illustration of quantum teleportation between two client nodes. However, our current model is sufficient to show that it is possible to use quantum weights in the training of neural networks.

Regarding teleportation scheme, the work in [33] proves a deterministic and secure way to send a six-qubit state by using a six-qubit cluster state as quantum channel. We plan to implement this in our network by teleporting the weights of each layer separately to the next client in the training.

4.3.2. Differential Privacy

A central drawback of QML/QFL models is the lack of privacy-preserving features, which can be exploited by membership inference and similar attacks. While malicious adversaries are unable to directly access the model and its data, they could still deduce whether or not a given data entry is used in the training process [34]. Such attacks are very simple and commonly used, which raises concerns with current quantum models.

One potential solution is to incorporate differential privacy in the training routine. This has been shown by the work in [29] to be successful when incorporated in QML models, as it is able to protect the client data without hindering the model's performance. We plan to adapt a similar differentially private scheme with our QFL model, utilizing classical methods to generate noise within the models. This integration would ensure even greater security of client information from outside attacks.

5. CONCLUSION

Our study establishes the quantum advantage of our QFL model over classical FL models using the make-circles dataset. We also demonstrate the success of a decentralized ring topology for training and introduce the first QFL model utilizing pure quantum weight communication through quantum teleportation methods. This work contributes to large-scale quantum internet development and enhances data privacy in ML models.

6. REFERENCES

[1] Mahdi Chehimi, Samuel Yen-Chi Chen, Walid Saad, Don Towsley, and Mérouane Debbah, "Foundations of quantum federated learning over classical and quantum networks," *IEEE Network*, 2023.

[2] Samuel Yen-Chi Chen and Shinjae Yoo, "Federated quantum machine learning," *Entropy*, vol. 23, no. 4, pp. 460, 2021.

[3] Mahdi Chehimi and Walid Saad, "Quantum federated learning with quantum data," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8617–8621.

[4] Zhao Wang, Yifan Hu, Shiyang Yan, Zhihao Wang, Ruijie Hou, and Chao Wu, "Efficient ring-topology decentralized federated learning with deep generative models for medical data in ehealthcare systems," *Electronics*, vol. 11, no. 10, pp. 1548, 2022.

[5] Lu Han, Xiaohong Huang, Dandan Li, and Yong Zhang, "Ringffl: A ring-architecture-based fair federated learning framework," *Future Internet*, vol. 15, no. 2, 2023.

[6] Guanghao Li, Yue Hu, Miao Zhang, Ji Liu, Quanjun Yin, Yong Peng, and Dejing Dou, "Fedhisyn: A hierarchical synchronous federated learning framework for resource and data heterogeneity," 2022.

[7] Zimu Xu, Wei Tian, Yingxin Liu, Wanjun Ning, and Jingjin Wu, "A ring topology-based communication-efficient scheme for d2d wireless federated learning," 2023.

[8] Jin woo Lee, Jaehoon Oh, Sungsu Lim, Se-Young Yun, and Jae-Gil Lee, "Tornadoaggregate: Accurate and scalable federated learning via the ring-based architecture," 2021.

[9] Sukin Sim, Peter D Johnson, and Alán Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Advanced Quantum Technologies*, vol. 2, no. 12, pp. 1900070, 2019.

[10] Trevor Lanting, Anthony J Przybysz, A Yu Smirnov, Federico M Spedalieri, Mohammad H Amin, Andrew J Berkley, Richard Harris, Fabio Altomare, Sergio Boixo, Paul Bunyk, et al., "Entanglement in a quantum annealing processor," *Physical Review X*, vol. 4, no. 2, pp. 021041, 2014.

[11] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, and Dacheng Tao, "Expressive power of parametrized quantum circuits," *Physical Review Research*, vol. 2, no. 3, pp. 033125, 2020.

[12] Amira Abbas, David Sutter, Christa Zoufal, Aurélien Lucchi, Alessio Figalli, and Stefan Woerner, "The power of quantum neural networks," *Nature Computational Science*, vol. 1, no. 6, pp. 403–409, 2021.

[13] Matthias C Caro, Hsin-Yuan Huang, Marco Cerezo, Kunal Sharma, Andrew Sornborger, Lukasz Cincio, and Patrick J Coles, "Generalization in quantum machine learning from few training data," *Nature communications*, vol. 13, no. 1, pp. 1–11, 2022.

[14] Kosuke Mitarai, Makoto Negoro, Masahiro Kitagawa, and Keisuke Fujii, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, pp. 032309, 2018.

[15] Jun Qi, Chao-Han Huck Yang, Pin-Yu Chen, and Min-Hsiu Hsieh, "Theoretical error performance analysis for variational quantum circuit based functional regression," *npj Quantum Information*, vol. 9, no. 1, pp. 4, 2023.

[16] Samuel Yen-Chi Chen, Chih-Min Huang, Chia-Wei Hsing, and Ying-Jer Kao, "An end-to-end trainable hybrid classical-quantum classifier," *Machine Learning: Science and Technology*, vol. 2, no. 4, pp. 045021, 2021.

[17] Samuel Yen-Chi Chen, Chao-Han Huck Yang, Jun Qi, Pin-Yu Chen, Xiaoli Ma, and Hsi-Sheng Goan, "Variational quantum circuits for deep reinforcement learning," *IEEE Access*, vol. 8, pp. 141007–141024, 2020.

[18] Samuel Yen-Chi Chen, "Quantum deep recurrent reinforcement learning," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[19] Chao-Han Huck Yang, Jun Qi, Samuel Yen-Chi Chen, Pin-Yu Chen, Sabato Marco Siniscalchi, Xiaoli Ma, and Chin-Hui Lee, "Decentralizing feature extraction with quantum convolutional neural network for automatic speech recognition," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 6523–6527.

[20] Chao-Han Huck Yang, Jun Qi, Samuel Yen-Chi Chen, Yu Tsao, and Pin-Yu Chen, "When bert meets quantum temporal convolution learning for text classification in heterogeneous computing," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8602–8606.

[21] Riccardo Di Sipio, Jia-Hong Huang, Samuel Yen-Chi Chen, Stefano Mangini, and Marcel Worring, "The dawn of quantum natural language processing," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8612–8616.

[22] Shuyue Stella Li, Xiangyu Zhang, Shu Zhou, Hongchao Shu, Ruixing Liang, Hexin Liu, and Leibny Paola Garcia, "Pqlm-multilingual decentralized portable quantum language model," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[23] Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L Fang, "Quantum long short-term memory," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 8622–8626.

[24] Jun Qi and Javier Tejedor, "Classical-to-quantum transfer learning for spoken command recognition based on quantum neural networks," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 8627–8631.

[25] Weikang Li, Sirui Lu, and Dong-Ling Deng, "Quantum federated learning through blind quantum computing," *Science China Physics, Mechanics and Astronomy*, vol. 64, no. 10, 2021.

[26] Qi Xia and Qun Li, "Quantumfed: A federated learning framework for collaborative quantum training," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.

[27] Won Joon Yun, Jae Pyoung Kim, Soyi Jung, Jihong Park, Mehdi Bennis, and Joongheon Kim, "Slimmable quantum federated learning," 2022.

[28] Rui Huang, Xiaoqing Tan, and Qingshan Xu, "Quantum federated learning with decentralized data," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 4, pp. 1–10, 2022.

[29] William M Watkins, Samuel Yen-Chi Chen, and Shinjae Yoo, "Quantum machine learning with differential privacy," *Scientific Reports*, vol. 13, no. 1, pp. 2453, 2023.

[30] Mahdi Chehimi, Samuel Yen-Chi Chen, Walid Saad, and Shinjae Yoo, "Federated quantum long short-term memory (fedqlstm)," *arXiv preprint arXiv:2312.14309*, 2023.

[31] Jun Qi, Xiao-Lei Zhang, and Javier Tejedor, "Optimizing quantum federated learning based on federated quantum natural gradient descent," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[32] Ryosuke Satoh, Michal Hajdušek, Naphan Benchasattabuse, Shota Nagayama, Kentaro Teramoto, Takaaki Matsuo, Sara Ayman Metwalli, Poramet Pathumsoot, Takahiko Satoh, Shigeya Suzuki, and Rodney Van Meter, "Quisp: a quantum internet simulation package," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, 2022, pp. 353–364.

[33] Xiaoqing Tan, Xiaoqian Zhang, and Tingting Song, "Deterministic quantum teleportation of a particular six-qubit state using six-qubit cluster state," *International Journal of Theoretical Physics*, vol. 55, no. 1, pp. 155–160, 2016.

[34] Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov, "Membership inference attacks against machine learning models," in *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2017, pp. 3–18.