

# FEDCOMPASS: FEDERATED CLUSTERED AND PERIODIC AGGREGATION FRAMEWORK FOR HYBRID CLASSICAL-QUANTUM MODELS

Yueheng Wang<sup>1</sup>, Xing He<sup>2</sup>, Zinuo Cai<sup>3</sup>, Rui Zhang<sup>1</sup>, Ruhui Ma<sup>3</sup>, Yuan Liu<sup>1</sup>, Rajkumar Buyya<sup>4</sup>

<sup>1</sup>School of Artificial Intelligence and Computer Science, Jiangnan University

<sup>2</sup>Antai College of Economics & Management, Shanghai Jiao Tong University

<sup>3</sup>School of Computer Science, Shanghai Jiao Tong University

<sup>4</sup>Cloud Computing and Distributed Systems Laboratory, University of Melbourne

## ABSTRACT

Federated learning enables collaborative model training across decentralized clients under privacy constraints. Quantum computing offers potential for alleviating computational and communication burdens in federated learning, yet hybrid classical-quantum federated learning remains susceptible to performance degradation under non-IID data. To address this, we propose FEDCOMPASS, a layered aggregation framework for hybrid classical-quantum federated learning. FEDCOMPASS employs spectral clustering to group clients by class distribution similarity and performs cluster-wise aggregation for classical feature extractors. For quantum parameters, it uses circular mean aggregation combined with adaptive optimization to ensure stable global updates. Experiments on three benchmark datasets show that FEDCOMPASS improves test accuracy by up to 10.22% and enhances convergence stability under non-IID settings, outperforming six strong federated learning baselines.

**Index Terms**— Federated Learning, Non-IID Data, Spectral Clustering, Circular Mean

## 1. INTRODUCTION

Nowadays, data privacy and potential leakage risks have become critical issues requiring urgent attention. Federated learning (FL) [1], as a privacy-preserving distributed learning paradigm, effectively reduces the privacy risk by training models locally on client devices and uploading only parameter updates instead of raw data. Thanks to this characteristic, FL has been widely adopted in scenarios such as mobile healthcare [2], IoT [3], and distributed sensing [4]. However, this privacy protection mechanism also introduces significant communication and computational overhead, posing serious challenges to training efficiency, especially in large-scale or resource-constrained edge environments [5].

At the same time, as an emerging computational paradigm, quantum computing [6] leverages quantum superposition and entanglement to enable parallel information processing. Since current quantum devices remain in the Noisy Intermediate-Scale Quantum (NISQ) era [7] with limited error correction capabilities and communication reliability issues [8], hybrid classical-quantum machine learning [9] has shown great potential. Such architecture employs classical neural networks for efficient feature extraction, and leverages quantum circuits to accelerate computation and enhance representational learning in specific tasks [10], thereby opening new avenues for improving federated learning efficiency.

However, federated learning often faces the challenge of non-IID data [11] in practice, which can easily lead to local model bias and difficulties in global convergence [12]. Although methods such as FedProx [13], FedBN [14], and FedPer [15] have achieved certain success in traditional scenarios, non-IID data still presents two major challenges in the classical-quantum federated learning setting. Firstly, significant differences in feature distributions among clients exacerbate the deviation in classical feature extraction layers, making it difficult to maintain global consistency after model aggregation. Secondly, quantum parameters are periodic [16] and sensitive to data distribution [17]. Direct arithmetic averaging can easily cause parameter period mismatch, aggravating training instability and leading to optimization conflicts between classical and quantum modules.

To address the aforementioned challenges, we propose FEDCOMPASS, a layered aggregation optimization framework in the hybrid classical-quantum federated learning setting. To mitigate model deviation caused by differences in client feature distributions, FEDCOMPASS employs spectral clustering based on client category statistics and performs weighted aggregation within clusters to generate cluster-level classical feature extractors. To tackle the periodicity issue specific to quantum parameters, we introduce a circular mean for periodic parameters based on the unit circle, combined with an adaptive optimizer to achieve robust global updates.

To validate the effectiveness and generalization capabil-

This work was supported by the Natural Science Foundation of Jiangsu Province under Grant BK20241606, and the National Natural Science Foundation of China under Grant 62472200.

\* Corresponding author: Rui Zhang(zhangruisg111@jiangnan.edu.cn).

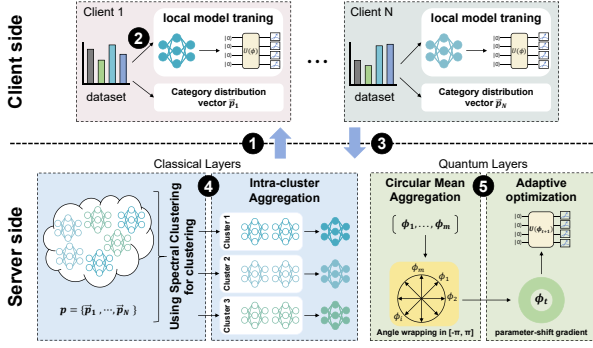


Fig. 1: Overview of our mechanism.

ity of the proposed framework, we conduct extensive experiments on three datasets: MNIST, Fashion-MNIST, and CIFAR-10. The results demonstrate that FEDCOMPASS significantly improves test accuracy and convergence stability across various non-IID settings, consistently outperforming six mainstream federated learning baseline methods. The framework fully exemplifies its comprehensive advantages in accuracy, stability, and privacy preservation within hybrid classical-quantum federated learning environments.

## 2. THE PROPOSED FEDCOMPASS ALGORITHM

### 2.1. Overview

Fig. 1 illustrates the overall workflow of FEDCOMPASS. The server first initializes a hybrid global model composed of a classical feature extraction layer and a quantum classifier, which is distributed to all clients (Step 1). Subsequently, each client conducts end-to-end training of the model based on their local data (Step 2). After training, clients upload the updated model parameters along with their data class distribution vectors to the server (Step 3). The server employs a hierarchical aggregation strategy, processing the classical and quantum layers separately. For the classical layer, clients are dynamically clustered based on data distribution similarity, and weighted aggregation is performed within each cluster to generate cluster-level classical feature extractors (Step 4). For the quantum layer, the circular mean method is applied to aggregate quantum parameters, combined with an adaptive optimization strategy to achieve robust global updates (Step 5). Finally, the server distributes the updated cluster-level classical model parameters and global quantum parameters to the clients for the next round of training.

### 2.2. Cluster-Based Aggregation for Classical Feature Extraction Layer

To address model bias caused by non-IID data distributions, we introduce a client clustering mechanism based on data distribution similarity in the classical network part,

enhancing the model’s adaptability to data heterogeneity. Throughout the federated learning process, no raw data is uploaded, thereby ensuring privacy protection. Let the client set be  $\{c_1, c_2, \dots, c_N\}$ , with corresponding datasets  $\{d_1, d_2, \dots, d_N\}$ , a total number of classes  $C$ , and a data concentration parameter  $\alpha$ . During data partitioning, the concentration parameter  $\alpha$  controls the degree of data heterogeneity: a smaller  $\alpha$  indicates a more heterogeneous data distribution. Simultaneously, each client uploads its class distribution vector  $\vec{p}$  to the server as a statistical representation of its data features. This vector is a  $C$ -dimensional vector representing the proportion of samples from each class in the client’s local data, i.e.,  $\vec{p}_i = (p_{i1}, p_{i2}, \dots, p_{iC})$ , where  $p_{ij}$  denotes the proportion of class  $j$  data in client  $c_i$ .

The server collects the local data distribution vectors  $\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N\}$  from the clients and computes a similarity matrix  $S$  based on these statistics:

$$S_{ij} = \exp \left( -\lambda_1 \text{JS}(\vec{p}_i, \vec{p}_j) - \lambda_2 \frac{|n_i - n_j|}{n_i + n_j} \right). \quad (1)$$

The similarity metric comprehensively evaluates the similarity between clients by considering both distribution divergence and sample size discrepancy. The Jensen–Shannon divergence  $\text{JS}(\cdot, \cdot)$  measures the difference in class distribution patterns between clients. The second term serves as the relative difference in sample size, where  $n_i$  denotes the sample size of client  $c_i$ , reflecting the impact of data volume disparity on model updates. The hyperparameters  $\lambda_1$  and  $\lambda_2$  balance the weights of the two terms: increasing  $\lambda_1$  places greater emphasis on distribution consistency, while increasing  $\lambda_2$  prioritizes the alignment of sample size scales.

Based on this, the server employs a spectral clustering algorithm grounded in the Normalized Cut criterion. It computes the normalized Laplacian matrix, performs eigenvalue decomposition, and applies K-means clustering to the top  $M$  eigenvectors to identify groups of clients with similar data distribution patterns. The clients are then grouped into  $M$  clusters  $\{C_1, C_2, \dots, C_M\}$ , with each cluster corresponding to a potential data distribution pattern, thereby achieving an effective partitioning of heterogeneous client groups.

During the aggregation phase, the server receives the classical feature extraction layer parameters  $\theta_c^{(i)}$  uploaded by the clients and performs a weighted average aggregation within each cluster:

$$\theta_c^{(m)} = \frac{\sum_{i \in C_m} n_i \cdot \theta_c^{(i)}}{\sum_{i \in C_m} n_i}, \quad (2)$$

where the weights are determined by the local sample size of each client, resulting in a cluster-shared classical model.

### 2.3. Global Aggregation Optimization for Periodic Quantum Classifier

As a globally shared module, the quantum classifier requires consistency and stability in its parameters across all clients.

---

**Algorithm 1** Quantum Parameter Update and Aggregation.

---

- 1: **Input:** Client parameters  $\{\phi_i\}_{i=1}^N$ , sample sizes  $\{n_i\}_{i=1}^N$ , previous global parameters  $\phi_t$ , FedAdam states  $m_{t-1}, v_{t-1}$ , hyperparameters  $\beta_1, \beta_2, \eta, \epsilon$
  - 2: **Output:** Updated global parameters  $\phi_{t+1}$ , updated states  $m_t, v_t$
  - 3: Compute client weights:  $\omega_i = n_i / \sum_j n_j$
  - 4: **for** each dimension  $j = 1, \dots, m$  **do**
  - 5:   Compute  $\bar{\phi}_j$  via Eq. 3 {Circular mean aggregation}
  - 6: **end for**
  - 7: Aggregated parameters:  $\bar{\phi} = (\bar{\phi}_1, \dots, \bar{\phi}_m)$
  - 8: Construct gradient:  $\mathbf{g}_t = \phi_t - \bar{\phi}$
  - 9: Update moment:  $m_t, v_t$  via Eq. 4
  - 10: Bias correction:  $\hat{m}_t, \hat{v}_t$  via Eq. 5
  - 11: Global update:  $\phi_{t+1}$  via Eq. 6
  - 12: **return**  $\phi_{t+1}, m_t, v_t$
- 

To achieve this, we design a quantum parameter aggregation method on the server side based on periodic averaging and adaptive updating. This approach first employs circular mean to resolve inconsistencies caused by the periodicity of rotation angles, and then introduces an adaptive update mechanism to enhance the stability of global convergence.

The overall quantum parameter aggregation and update process is described in Algorithm 1. Let the quantum parameters uploaded by client  $c_i$  be  $\phi_i = (\phi_i^{(1)}, \phi_i^{(2)}, \dots, \phi_i^{(m)})$ . For the  $j$ -th parameter dimension, the aggregation process is defined as:

$$\bar{\phi}_j = \text{atan2} \left( \sum_{i=1}^N \omega_i \sin(\phi_i^{(j)}), \sum_{i=1}^N \omega_i \cos(\phi_i^{(j)}) \right), \quad (3)$$

where  $\omega_i = n_i / \sum_j n_j$  is the client weight. This operation (line 5) maps angles to the unit circle for averaging before mapping them back to the angular space, thereby avoiding periodicity-induced inconsistencies.

To further enhance the convergence stability of the global quantum classifier, we employ an adaptive optimizer on the server side (lines 7–11) to update the aggregated parameters globally. First, the average gradient  $\mathbf{g}_t$  for the quantum parameters in round  $t$  is computed based on the global quantum parameters  $\mathbf{g}_t = \phi_t - \bar{\phi}$ . Subsequently, momentum update, bias correction, and parameter update are performed as follows:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \mathbf{g}_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) \mathbf{g}_t^2, \quad (4)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad (5)$$

$$\phi_{t+1} = \phi_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \quad (6)$$

where  $\beta_1$  and  $\beta_2$  are momentum hyperparameters,  $\eta$  is the learning rate, and  $\epsilon$  is a numerical stability constant. The

integration of the adaptive optimizer with the periodic constraints of quantum parameters ensures stable convergence of the quantum classifier under non-IID data distributions.

Finally, the server distributes the updated cluster-level classical model parameters  $\theta_c^{(m)}$  and the quantum classifier parameters  $\phi_{t+1}$  to the clients for the next round of training and aggregation. This design addresses the periodicity of quantum parameters while facilitating the transfer of discriminative capabilities across different clients through global sharing, thereby improving overall classification performance and convergence stability.

### 3. PERFORMANCE EVALUATION

#### 3.1. Experiments Setup

**Datasets.** We evaluate FEDCOMPASS and comparative methods on three datasets: MNIST, Fashion-MNIST, and CIFAR-10. From each dataset, we uniformly selected 4 classes to form a four-class classification task. The data was partitioned using a Dirichlet distribution [4] with two non-IID parameter settings,  $\alpha = 0.3$  and  $\alpha = 0.7$ .

**Models.** We adopt a hybrid classical-quantum architecture, where a classical network performs feature extraction and a quantum network carries out classification. For MNIST, we use LeNet followed by a parameterized quantum circuit. For more complex datasets such as CIFAR-10 and Fashion-MNIST, the first two layers of ResNet-18 are employed for feature extraction, and the features are then passed to a quantum convolutional network.

**Training Settings.** We simulate a federated learning environment with 10 clients. Each client performs 5 local epochs per communication round with a batch size of 32. Due to the high overhead of quantum training, the server conducts 5 global communication rounds in total. We use the Adam optimizer with a learning rate of 0.001 for updating both the local models and the server-side quantum parameters.

**Baselines.** We compare FEDCOMPASS with the following six classical federated learning methods: (1) FedAvg [18]; (2) FedProx [13]; (3) FedBN [14]; (4) FedPer [15]; (5) FedNova [19]; (6) Scaffold [20].

**Implementation.** We employ Ray as the distributed computing framework to coordinate multi-client parallel training tasks. The model construction and gradient calculation for the quantum part are implemented using the PennyLane library. The federated learning process is built and executed based on the Flower framework.

#### 3.2. Results and Discussion

We evaluate FEDCOMPASS on MNIST, Fashion-MNIST, and CIFAR-10 under two non-IID settings with  $\alpha = 0.3$  and  $\alpha = 0.7$ , comparing it against six baseline methods. As shown in the accuracy results (Table 1) and convergence curves

**Table 1:** Comparison of test accuracy of different federated learning algorithms across three datasets under non-IID settings. Best values are in **bold** and second best are underlined.

Dataset	MNIST		Fashion-MNIST		CIFAR-10	
Non-IID Degree	0.30	0.70	0.30	0.70	0.30	0.70
FedAvg [18]	<u>99.54</u>	99.49	<u>96.15</u>	<b>96.05</b>	66.78	<u>76.30</u>
FedProx [13]	74.96	50.69	93.18	93.03	55.20	69.68
FedBN [14]	50.83	50.83	23.98	25.55	<u>71.25</u>	57.63
FedPer [15]	33.68	26.65	48.50	69.83	66.55	57.68
FedNova [19]	40.32	<u>99.62</u>	4.40	25.25	23.13	19.10
Scaffold [20]	40.29	47.17	64.65	85.15	38.53	54.28
FEDCOMPASS (Ours)	<b>99.69</b>	<b>99.76</b>	<b>96.20</b>	<u>95.50</u>	<b>77.00</b>	<b>80.10</b>

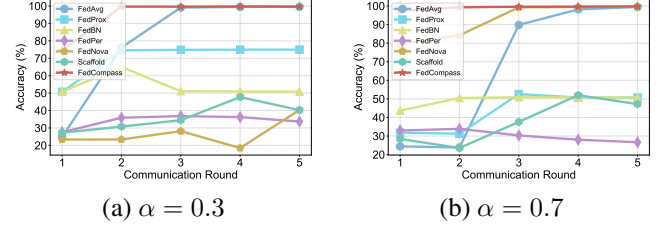
**Table 2:** Test accuracy of ablation study on CIFAR-10 across communication rounds.

Round	No Clustering	No Circular Mean	FedCompass
1	25.10	25.00	25.98
2	50.65	38.60	52.65
3	50.38	46.23	71.55
4	55.15	32.85	70.03
5	56.13	26.25	<u>77.00</u>

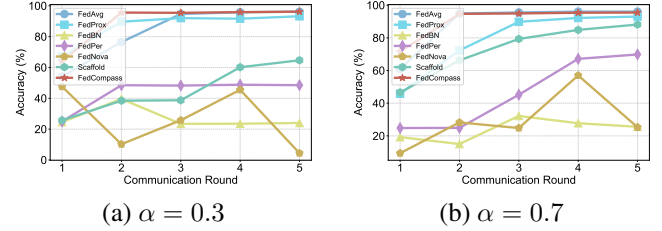
(Fig. 2 – 4), FEDCOMPASS consistently achieves the best performance in most scenarios.

FEDCOMPASS demonstrates the most significant improvement on the CIFAR-10 dataset. Under the condition of  $\alpha = 0.3$ , it achieves an accuracy of 77.00%, which is a 10.22% increase compared to FedAvg. This result depends on FEDCOMPASS’s clustering mechanism, which effectively groups clients with similar class distributions, thereby reducing discrepancies in classical feature learning. When  $\alpha = 0.7$ , the accuracy further improves to 80.10%, outperforming FedAvg by 3.80%, confirming the robustness of our method across varying degrees of non-IID data. On MNIST, its performance approaches the dataset’s upper limit of 99.7%, while on Fashion-MNIST, the improvement is relatively smaller, indicating the dataset’s lower sensitivity to distribution shifts. Nonetheless, FEDCOMPASS still maintains leading results. Convergence analysis shows that FEDCOMPASS exhibits faster and more stable convergence under different heterogeneity conditions, with the advantage being particularly pronounced at  $\alpha = 0.3$ .

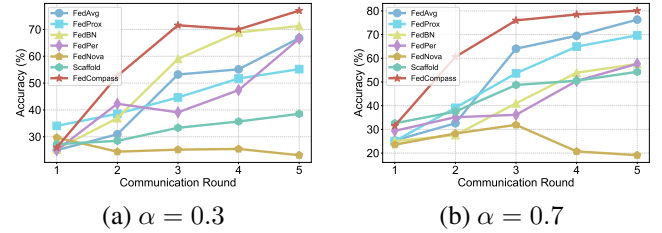
In the ablation study (Table 2), FEDCOMPASS achieved the highest test accuracy, with steady improvement as the number of communication rounds increased, demonstrating the effectiveness of the complete framework. Removing the clustering mechanism for the classical layers resulted in a significant performance drop, particularly in the later rounds. This indicates that the absence of clustering grouping leads to divergence in the feature extractors, adversely affecting global convergence. Omitting the circular mean aggregation for quantum parameters yielded the lowest and most unstable per-



**Fig. 2:** Convergence curves of test accuracy versus communication rounds on MNIST under non-IID settings.



**Fig. 3:** Convergence curves of test accuracy versus communication rounds on Fashion-MNIST under non-IID settings.



**Fig. 4:** Convergence curves of test accuracy versus communication rounds on CIFAR-10 under non-IID settings.

mance. The substantial fluctuations reflect the adverse impact of misaligned periodicity in quantum rotation angles, underscoring the necessity of circular aggregation for coordinating periodic updates and avoiding optimization conflicts.

## 4. CONCLUSION

This paper proposes FEDCOMPASS, a novel hybrid classical-quantum federated learning framework designed to address the challenges of non-IID data distribution. The method implements two key mechanisms, a spectral clustering-based client grouping strategy with within-cluster aggregation of classical feature extractors, and a circular mean aggregation method combined with adaptive optimization tailored for the periodic nature of quantum parameters. It provides an effective solution to data heterogeneity in hybrid federated learning while enhancing overall performance without compromising convergence.

## 5. REFERENCES

- [1] Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, and Yuan Gao, “A survey on federated learning,” *Knowledge-Based Systems*, vol. 216, pp. 106775, 2021.
- [2] Rodolfo Stoffel Antunes, Cristiano André da Costa, Arne Küderle, Imrana Abdullahi Yari, and Björn Eskofier, “Federated learning for healthcare: Systematic review and architecture proposal,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 13, no. 4, pp. 1–23, 2022.
- [3] Wei Yang Bryan Lim, Nguyen Cong Luong, Dinh Thai Hoang, Yutao Jiao, Ying-Chang Liang, Qiang Yang, Dusit Niyato, and Chunyan Miao, “Federated learning in mobile edge networks: A comprehensive survey,” *IEEE communications surveys & tutorials*, vol. 22, no. 3, pp. 2031–2063, 2020.
- [4] Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, and Yasaman Khazaeni, “Federated learning with matched averaging,” *arXiv preprint arXiv:2002.06440*, 2020.
- [5] Rui Zhang, Yueheng Wang, Xing He, Zinuo Cai, Yicheng Di, Jiayu Bao, Jiansong Fan, and Zhongle Qu, “Qfi-opt: Communication-efficient quantum federated learning via quantum fisher information,” *Software: Practice and Experience*, 2025.
- [6] Roman Rietsche, Christian Dremel, Samuel Bosch, Léa Steinacker, Miriam Meckel, and Jan-Marco Leimeister, “Quantum computing,” *Electronic Markets*, vol. 32, no. 4, pp. 2525–2536, 2022.
- [7] Jonathan Wei Zhong Lau, Kian Hwee Lim, Harshank Shrotriya, and Leong Chuan Kwek, “Nisq computing: where are we and where do we go?,” *AAPPS bulletin*, vol. 32, no. 1, pp. 27, 2022.
- [8] Liwei Lin, Rongbo Ma, Zejian Wang, Zinuo Cai, Haochen Xu, Baoheng Zhang, Ruhui Ma, and Rajkumar Buyya, “Hwdsqp: A historical weighted and dynamic scheduling quantum protocol to enhance communication reliability,” *IEEE Journal on Selected Areas in Communications*, 2025.
- [9] Gennaro De Luca, “A survey of nisq era hybrid quantum-classical machine learning research,” *Journal of Artificial Intelligence and Technology*, vol. 2, no. 1, pp. 9–15, 2022.
- [10] 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.
- [11] Qinbin Li, Yiqun Diao, Quan Chen, and Bingsheng He, “Federated learning on non-iid data silos: An experimental study,” in *2022 IEEE 38th international conference on data engineering (ICDE)*. IEEE, 2022, pp. 965–978.
- [12] Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, and Zhihua Zhang, “On the convergence of fedavg on non-iid data,” *arXiv preprint arXiv:1907.02189*, 2019.
- [13] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith, “Federated optimization in heterogeneous networks,” *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.
- [14] Xiaoxiao Li, Meirui Jiang, Xiaofei Zhang, Michael Kamp, and Qi Dou, “Fedbn: Federated learning on non-iid features via local batch normalization,” in *International Conference on Learning Representations*, 2021.
- [15] Manoj Ghuhan Arivazhagan, Vinay Aggarwal, Aaditya Kumar Singh, and Sunav Choudhary, “Federated learning with personalization layers,” *arXiv preprint arXiv:1912.00818*, 2019.
- [16] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer, “Effect of data encoding on the expressive power of variational quantum-machine-learning models,” *Physical Review A*, vol. 103, no. 3, pp. 032430, 2021.
- [17] Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R McClean, “Power of data in quantum machine learning,” *Nature communications*, vol. 12, no. 1, pp. 2631, 2021.
- [18] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Artificial Intelligence and Statistics*, 2017.
- [19] Jianyu Wang, Qinghua Liu, Hao Liang, Gauri Joshi, and H Vincent Poor, “Tackling the objective inconsistency problem in heterogeneous federated optimization,” *Advances in neural information processing systems*, vol. 33, pp. 7611–7623, 2020.
- [20] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh, “Scaffold: Stochastic controlled averaging for federated learning,” in *International conference on machine learning*. PMLR, 2020, pp. 5132–5143.