Quantum Machine Intelligence (2025) 7:12
https://doi.org/10.1007/s42484-025-00243-x

RESEARCH ARTICLE

Application of quantum-inspired tensor networks to optimize federated learning systems

Amandeep Singh Bhatia[1] · Mandeep Kaur Saggi[1] · Sabre Kais[1]

Received: 2 August 2024 / Accepted: 21 January 2025

Abstract
Federated learning (FL) has gained significant traction across diverse industries, which allows multiple clients or institutions to enhance model performance and outcomes while preserving data privacy collaboratively. In recent years, tensor networks (TNs) have become important in machine learning because they allow the compact representation of high-dimensional tensors by decomposing them into lower-dimensional components with polynomial complexity. The application of TNs in FL is a natural extension because of its flexible framework for representing and optimizing models. Inspired by quantum computing principles, we have integrated a quantum-inspired tensor network into the FL framework. This framework focuses on a one-dimensional matrix product state (MPS) tensor network (TN) in a federated setting (FedTN), with data distributed across homogeneous and heterogeneous partitions among clients. Our experiments demonstrate that tensor network-based federated learning can be made practical, as FedTN is robust to the unbalanced and non-IID data distributions typically encountered in such settings. Our research assessed the effectiveness and feasibility of comparing quantum-inspired TN and conventional methods, evaluating their performance, and exploring the benefits of incorporating quantum principles in FL settings. Furthermore, we have investigated its performance when training for many local epochs (large E) between the averaging steps.

Keywords Tensor networks · Federated learning · Quantum computing · Machine learning · Security · Quantum machine learning

1 Introduction

Protecting privacy has become essential in the current digital environment, especially within machine learning applications. In recent years, federated learning (FL) has emerged as a distributed learning paradigm where multiple clients (e.g., mobile devices or institutions) collaboratively train a model while keeping their data decentralized, as illustrated in Fig. 1. It allows clients to exchange model parameters with a global server without disclosing their private data (Li et al. 2021). This paradigm improves data privacy and security, making it ideal for use in healthcare, finance, and other sensitive areas. However, implementing FL comes with unique challenges. One significant challenge is the high communication overhead (Nguyen et al. 2021). Several efficient communication strategies and techniques, such as model compression and aggregation, have been proposed to improve communication overhead and resource allocation in computation efficiency. Another challenge is the system and data heterogeneity, leading to significant variance among clients and slower convergence rates (Mendieta et al. 2022). To address these non-IID (independent and identically distributed) data, advanced techniques are required to optimize federated learning effectively across distributed data sources.

Tensor networks are gaining prominence across various domains in data science. They represent a class of quantum-inspired algorithms and techniques that emulate the tensor operations of quantum computers while running on classical hardware. These networks are influenced by theories, models, and methods from quantum physics; they are commonly referred to as "quantum-inspired." By leveraging

✉ Amandeep Singh Bhatia
drasinghbhatia@gmail.com; asbhatia@ncsu.edu

✉ Sabre Kais
skais@ncsu.edu

Mandeep Kaur Saggi
mksaggi@ncsu.edu

[1] Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606, USA

Published online: 28 January 2025

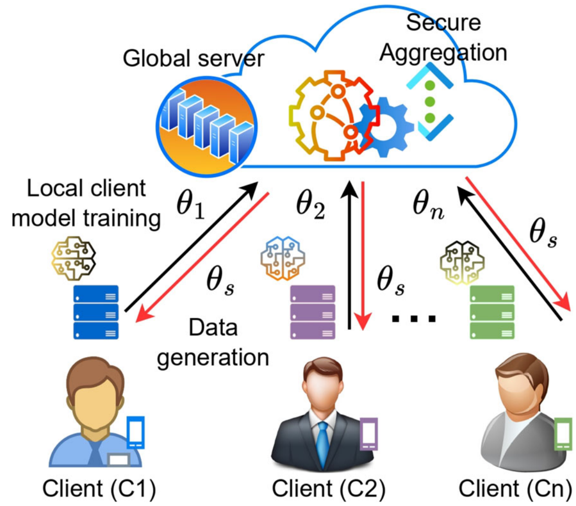

Fig. 1 An illustration of the federated learning process. Initially, the global model sends the current global weights (θ_s) to all clients ($C_1, C_2, ..., C_N$). Each client trains their local model on their respective data using these weights. After training, clients send their updated weights ($\theta_1, \theta_2, ..., \theta_n$) back to the global server. The global server then aggregates these weights to update the global model. This process repeats iteratively until the global model converges. The framework ensures data privacy as the clients do not share their raw data, only the model parameters

the properties of tensors, these techniques optimize computations, particularly in scenarios where accessing specific properties of the quantum state vector is sufficient, rather than dealing with the entire state vector (Ali et al. 2024). Quantum-inspired approaches have proven highly relevant to machine learning due to their ability to compress models effectively. By reducing memory requirements while maintaining a high level of precision, these methods offer a practical solution for handling complex computations efficiently. The MPS, a type of tensor network, is highly beneficial, as it decomposes a tensor into a one-dimensional structure through matrix product operations.

In recent years, quantum-inspired tensor network (QTN) approaches have been making significant strides in the field of data science (Cichocki 2014), offering novel solutions to traditional computational challenges in condensed matter physics (Fernández-González et al. 2015), and quantum computer science (Jahn and Eisert 2021; Grant et al. 2018; Bhatia et al. 2019). In the context of tensor networks (TNs), a *tensor* refers to a multi-dimensional array that generalizes scalars, vectors, and matrices to higher dimensions. A tensor with rank r and dimensions $d_1 \times \cdots \times d_r$ are elements of the tensor product vector space $\bigotimes_{i=1}^{r} \mathbb{C}^{d_i}$. These tensors are represented by components $V_{i_1 \cdots i_r}$, where i_j takes values from $[d_j]$ (Monturiol et al. 2024). Tensor networks use these tensors to capture and model complex relationships within the data.

Integrating tensor networks into deep learning pipelines offers multiple advantages, potentially boosting the performance and capabilities of existing models with tensorization and tensor decomposition. The process of tensorizing lowers the storage memory needed by decreasing the number of elements, and it also enhances computation efficiency. There are different layouts of TNs: grid (matrix product state (MPS) for 1D or projected entangled-pair states (PEPS) for 2D systems) or hierarchical (tree tensor network (TTN) and multi-scale entanglement renormalization ansatz (MERA)). One-dimensional MPS is the most widely studied TN and a versatile method used extensively in the study of 1D quantum many-body systems. It provides an efficient way to represent the state of a quantum system by decomposing it into a the product of matrices, which simplifies the handling and analysis of complex quantum states.

An MPS representation expresses a quantum state as a sequence of tensors (or matrices) connected in a linear fashion, as shown in Fig. 2. Consider a quantum state $|\psi\rangle$ of a system with N sites, each having a local Hilbert space dimension d (Sengupta et al. 2022). The MPS representation of $|\psi\rangle$ is given by

$$|\psi\rangle = \sum_{s_1, s_2, \ldots, s_N} A^{s_1} A^{s_2} \cdots A^{s_N} |s_1, s_2, \ldots, s_N\rangle, \tag{1}$$

where A^{s_k} are $D \times D$ matrices (with boundary conditions requiring the first and last matrices to be vectors of dimension D). Here, D is the bond dimension, a parameter that controls the accuracy of the MPS representation, as shown in Fig. 2. A larger bond dimension allows for a more accurate representation but at the cost of increased computational resources.

A tensor network's layout specifies the maximum entanglement or internal correlation it can accommodate. In the case of MPS, the entanglement adheres to an area law, indicating that the entanglement between a sub-network and its environment is proportional to its boundary size. Therefore, for a 1D MPS, the entanglement remains constant (Rieser et al. 2023). Due to the quantum-inspired formulation of TNs, the information cannot be shared across branches in

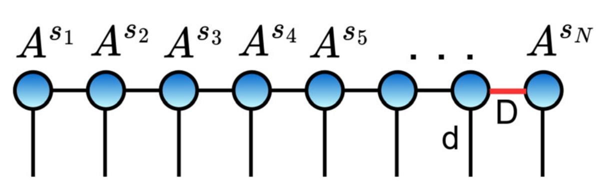

Fig. 2 Matrix product state (MPS) representation for a vector. The MPS representation of a quantum state $|\psi\rangle$ for a system with N sites, each with a local Hilbert space dimension d and bond dimension D. It is expressed as a sequence of tensors (or matrices) A^{s_k}, which are connected linearly. Each tensor is primarily entangled with its neighbors

the same way that neural networks use the information to trigger neuron activation.

Based on the current literature, TNs-based machine learning models are applied to various tasks including classification (Grant et al. 2018; Bhatia et al. 2019; Huggins et al. 2019; Wang et al. 2021), compression (Selvan et al. 2020), entanglement-based feature extraction (Liu et al. 2021), and generative TNs (Wall et al. 2021). Due to the structural similarity between TN architectures and neural networks, several studies have explored the mapping of one onto the other to compare different network configurations. For instance, recurrent neural networks can efficiently simulate MPS in certain scenarios (Wu et al. 2023), while TNs can be used to implement convolutional neural networks efficiently (Levine et al. 2019). Recently, Bhatia and Neira (2024) introduced the first quantum federated learning framework based on tensor networks for the healthcare sector and investigated its performance in various types of medical images. Additionally, the authors provided a comparison of the methods based on communication overhead and robustness against local differential privacy.

Bhatia et al. (2024a) improved QFL efficiency by reducing communication rounds and boosting performance through quantum natural gradient descent optimization and the use of variational quantum circuits on diverse healthcare datasets. Saggi et al. (2024a) introduced FedQNN, a hybrid quantum-classical neural network framework, for the classification of human DNA sequence datasets. Li et al. (2024) proposed advanced quantum protocols incorporating quantum communication to tackle federated learning challenges, enhance privacy, and improve communication efficiency using the gradient hiding technique. Javeed et al. (2024) developed a framework combining quantum computing, federated learning, and 6G wireless networks to bolster the security of IoT ecosystems.

Federated learning based on the MPS quantum-inspired network has the potential to be extended to various industries beyond the scope of this study. Quantum-inspired tensor networks are well-suited for industrial applications, offering the capability to tackle highly complex and large-scale problems with remarkable efficiency. Healthcare and drug discovery: The lightweight MPS-based federated learning framework can be applied to privacy-sensitive tasks in healthcare, such as collaborative training of models for diagnosing diseases (Bhatia et al. 2023a, 2024a, 2023b), predicting patient outcomes, or analyzing genomic data (Saggi et al. 2024a; Bhatia et al. 2023c; Saggi et al. 2024b). Its ability to process high-dimensional data efficiently aligns well with applications like medical imaging, electronic health records, and drug discovery. Finance: This framework can enable secure and efficient collaborative learning for fraud detection, credit scoring, portfolio optimization, risk assessment, and financial predictions, where client data must remain private due to stringent regulatory requirements (Orús et al. 2019; Paquet et al. 2022; Brandhofer et al. 2022). Manufacturing and supply chain: It can be extended to predict equipment failure, optimize supply chains, and detect anomalies in manufacturing processes while keeping proprietary operational data confidential (Bhatia et al. 2024b). Quantum-inspired models like MPS often have lower computational and memory overheads compared to traditional deep learning models, making them ideal for industries prioritizing sustainability, energy efficiency, and security.

1.1 Challenges

The following are the key challenges encountered during the implementation of the quantum-inspired MPS tensor network in federated learning:

- The decomposition of tensors in the MPS framework requires sophisticated algorithms and careful optimization to avoid excessive computational burden, particularly when applied to real-world, large-scale datasets in a federated setting.
- FL typically involves frequent communication between clients and the central server. The quantum-inspired MPS approach, while reducing model size, still requires careful handling of tensor operations to minimize communication costs, which can be substantial in large-scale federated settings.
- FL environments often deal with data that is non-IID (non-independent and identically distributed). Ensuring that the quantum-inspired MPS model generalizes well across diverse client data without overfitting is a key concern.

1.2 Motivation

Given the remarkable capability of tensor networks to handle and represent high-dimensional data through advanced tensor decompositions, integrating quantum-inspired tensor networks (QTNs) into federated learning environments emerges as a compelling objective. Based on the principles of quantum computing such as entanglement and superposition, tensor networks offer an advanced framework for representing and managing intricate relationships and dependencies within data. This makes them well-suited for the heterogeneous data encountered in FL settings. To the best of our knowledge, no published studies have implemented or evaluated the utility of quantum-inspired tensor train networks in federated settings.

1.3 Main contributions

To summarize, this paper makes the following contributions:

- We propose a quantum-inspired federated learning framework, FedTN, which leverages tensor train networks for collaborative training across multiple clients, allowing TNs to effectively handle non-IID data.
- Our experiments demonstrate that FedTN is capable of efficiently integrating new clients into the existing federated learning system by aligning with the local data distributions of these clients. Additionally, FedTN can achieve convergence of the global model without the need for extra adjustments or fine-tuning.
- Our experiments show that FedTNs provide better classification accuracy on test data than federated multilayer perceptron (FedMLP). Moreover, the global TN federated model exhibited enhanced classification accuracy and generalizability, consistently surpassing the performance of locally trained models, even when data is unevenly distributed among clients.

The remainder of the article is organized as follows: Section 2 discusses quantum-inspired tensor network-based federated learning. Section 3 presents the data settings and discusses the findings and outcomes. Finally, the concluding remarks are reported in Section 4. The list of acronyms used throughout the paper is given in Table 1.

2 Federated learning optimization for MPS tensor networks

In this paper, we focused on optimizing one-dimensional MPS tensor networks for image classification tasks (MNIST and FMNIST) in the federated learning environment.

The MPS tensor network, representing vector in the image space, consists of 2^N components, with each one described by the product of N matrices. Before describing the federated learning approach based on MPS TN employed in this study, it is important to first explain the data encoding process into the MPS.

Table 1 List of acronyms used throughout the paper

Abbr	Explanation
FL	Federated learning
non-IID	Non-independent and identically distributed
TN	Tensor networks
QML	Quantum machine learning
QTN	Quantum-inspired tensor networks
MPS	Matrix product state
FedMLP	Federated multi-layer perceptron
FedCNN	Federated convolutional neural networks
FedTN	Federated tensor networks

First, each pixel of the image is mapped into its own two-dimensional pixel space according to a local feature transformation as

$$\Phi(p) = (\cos(\pi p/2), \sin(\pi p/2)) \tag{2}$$

where $p \in [0, 1]$ represents the normalized pixel value. In the experiments, we used $\Phi(p)_i$, where the index i can take two values. For the linear feature transformation, we have $\Phi(p)_0 = 1 - p$ and $\Phi(p)_1 = p$. The overall image space is created by taking the tensor product of all individual pixel spaces. On changing a single pixel value from black to white results in a vector that is orthogonal to the original vector (Rieser et al. 2023). An image pixel $(p_1, p_2, ..., p_n)$ is represented in the image space as

$$(p_1, p_2, \ldots, p_n) \mapsto \Phi(p_1) \otimes \Phi(p_2) \otimes \cdots \otimes \Phi(p_n) \tag{3}$$

It is also referred to as a data tensor and has 2^n components, represented as $\Phi(p_1)_{i_1}\Phi(p_2)_{i_2}\cdots\Phi(p_n)_{i_n}$. Suppose an input image $x = (p_1, p_2, \ldots, p_n) \in [0, 1]^n$ represents a flattened image and corresponding label $y \in \{0, 1, ..., m-1\}$ (with n=784 and m=10 for MNIST and FMNIST). First, we determine the inner product between the encoded image vector and a variational MPS (Eq. 1) as

$$f^{(m)}(x) = \sum_{i_1, i_2, \ldots, i_n = 0}^{1} |\psi\rangle \Phi(p_1)_{i_1}\Phi(p_2)_{i_2}\cdots\Phi(p_n)_{i_n} \tag{4}$$

It should be noted that every pixel indices of MPS is contracted with the data, except the index (0, 1,...., 9) which is left free to distinguish between 10 labels. For example, the middle one index of MPS is set free to distinguish whether an image belongs to class 0 or 1, as shown in Fig. 3c.

Following conventional machine learning methods, the variational parameters A^{s_i} defining MPS $|\psi\rangle$ (see Eq. 1) will be optimized to minimize an objective function in the training set. The optimization of parameters in MPS tensor networks can be achieved using any gradient-based or gradient-free method. In all experiments across models, we opted for multi-class cross-entropy (MCE) and ADAM optimizer to optimize our training set is as

$$\text{MCE} = -\sum_{(x_i, y_i) \in N} \log \text{softmax}\, f^{(y_i)}(x_i), \quad \text{where} \tag{5}$$

$$\text{softmax}\, f^{(y_i)}(x_i) = \frac{e^{f^{(y_i)}}(x_i)}{\sum_{l=0}^{m-1} e^{f^{(l)}(x_i)}} \tag{6}$$

2.1 Federated learning setup

We incorporated the federated averaging (FedAvg) algorithm with MPS tensor networks for image classification tasks

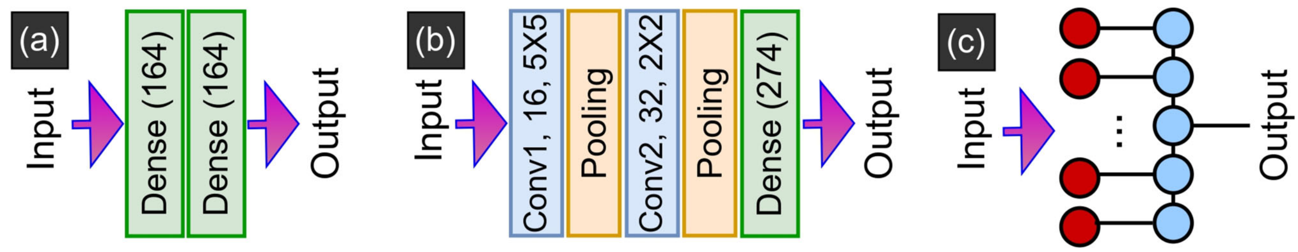

Fig. 3 Schematic depictions of the considered architectures in federated environment. **a** a Multi-layer perceptron (MLP) with two hidden layers, **b** a convolutional neural network (CNN) with two 5x5 convolutional layers, followed by a fully connected layer with 274 units, **c** a matrix product state (MPS) tensor network with 10 bond dimensions and output dimensions equal to the number of classes to predict

among the clients. It is widely used to train machine learning models collaboratively across multiple clients while keeping data decentralized. The steps of federated averaging are summarized as

1. Client-side computation

Each client c has a local dataset $\mathcal{D}_c$ with total n_c samples and updates its local model parameters as

- **Initial setup**: The global server model at round t is denoted as w_t.
- **Local gradient computation**: Each client c computes the gradient of its local loss function with respect to the current model parameters w_t. For a model with multi-class cross-entropy loss, the gradient g_c for client c is

$$g_c = \frac{1}{n_c} \sum_{(x_i, y_i) \in \mathcal{D}_c} \nabla_w \mathrm{MCE}(x_i, y_i, w_t),$$

  where $\mathrm{MCE}(x_i, y_i, w_t)$ is the cross-entropy loss for sample (x_i, y_i), and ∇_w represents the gradient with respect to the TN model parameters w.
- **Local update**: Each client performs a local update using a fixed learning rate η_c:

$$w_t^{(c)} = w_t - \eta_c g_c.$$

2. Aggregation on server-side

The server aggregates the locally updated models from all clients to compute the global model update as

- **Aggregation**: The server computes the weighted average of the updated models $w_t^{(c)}$ based on the number of samples n_c at each client as

$$w_{t+1} = \frac{1}{n} \sum_{c=1}^{C} n_c w_t^{(c)},$$

  where n is the total number of samples across all clients, given by $n = \sum_{c=1}^{C} n_c$.

3. Increase computation on the client side

To improve performance, each client can perform multiple local gradient descent steps or local epochs (E) before sending updates to the server for aggregation. For each epoch (e), the client updates its model parameters as

$$w_t^{(e)} = w_t^{(e-1)} - \eta \frac{1}{n_c} \sum_{(x_i, y_i) \in \mathcal{D}_c} \nabla_w \mathrm{MCE}(x_i, y_i, w_t^{(e-1)}),$$

where $w_t^{(e-1)}$ is the model weights from the previous epoch. After E local epochs, each client sends the final model $w_t^{(E)}$ to the server for aggregation. FedAvg achieves collaborative MPS TN model training across multiple clients by iterating this approach, all the while upholding data privacy and taking advantage of decentralized data.

3 Experimental results and setup

3.1 Datasets and models

In this paper, we focus on image classification tasks and utilize two datasets: MNIST and FMNIST. We train three different models in a federated environment: (1) a simple multilayer perceptron with 2-hidden layers with 164 units each using ReLu activations (157,450 total parameters), which we refer to as the FedMLP in Fig. 3a. (2) A CNN with two 5x5 convolution layers (the first with 32 channels, the second with 64, each followed with 2x2 max pooling), a fully connected layer with 274 units and ReLu activation, and a final softmax output layer (156,560 total parameters) (McMahan et al. 2017), which we refer to as the FedCNN in Fig. 3b. (3) A quantum-inspired tensor network with 10 bond dimensions and output dimensions embeds each pixel value into a vector space of dimension 2, using a random initialization method, and obtains an output vector with the probabilities after tensor contraction (157,820 total parameters), which we refer to as the FedTN in Fig. 3c. To investigate federated optimization, it is essential to define how the data is distributed across the clients. Each dataset is converted into a federated version by randomly distributing the training data among 100 clients, where each client is assigned at least 500 samples. To adjust the level of non-IIDness, we use latent Dirichlet allocation (LDA) on the labels, where the Dirichlet concentration parameter α controls the degree of class heterogeneity. The distribution of classes for different α values is depicted in Fig. 4a and d. Finally, for the number of rounds R, we use 500 rounds for the heterogeneous Dirichlet distribution with $\alpha = 0.1$ and 300 rounds for the homogeneous Dirichlet distribution with $\alpha = 1.0$, applied across all models for the MNIST and FMNIST datasets.

3.2 Performance evaluation

To make the comparison fair, we use multilayer perceptron (MLP) and convolutional neural network (CNN) with the same number of model parameters and entire data training as performance upper bound, which serves as a target that we strive to achieve in our experiments. We set the number of clients to 100. From these, we randomly select 10% to participate. Each participating client undergoes E local epochs of training on their dataset, where E=1 for results in Fig. 4, mini-batch gradients of size $b = 64$, and η_c is set as 0.001. We investigated the performance of different federated models with different degrees of heterogeneity by varying the concentration parameter $\alpha \in \{1, 0.1\}$. We ensured that all models have an equal number of trainable parameters, so the total amount of data exchanged between the data center and the clients is nearly identical across all models (Table 2).

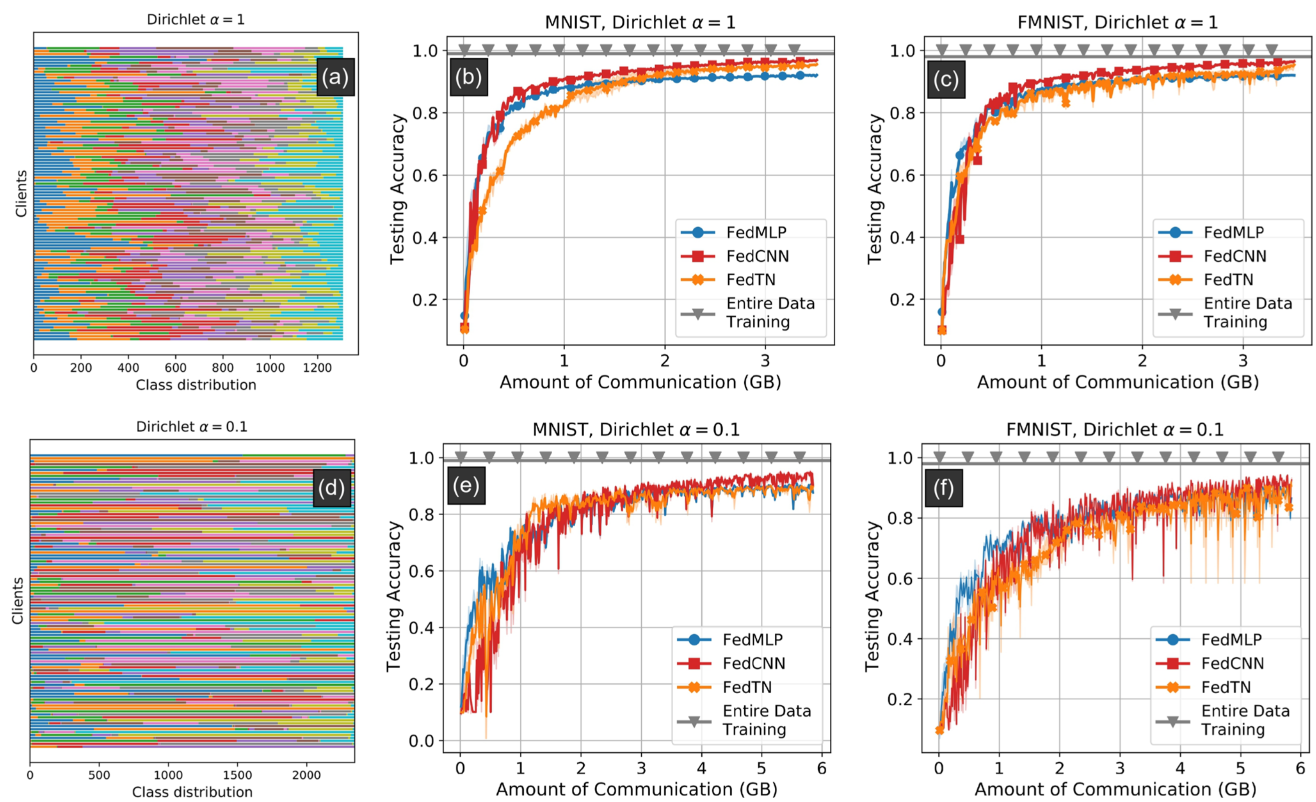

Fig. 4 Convergence rates and the effect of heterogeneity on different methods, induced by the Dirichlet distribution in the federated learning scenario. **a**, **d** The level of heterogeneity resulting from different concentration parameters $\alpha \in \{0.1, 1.0\}$ for the Dirichlet distribution, applied to the MNIST and FMNIST datasets (10 classes, represented by 10 colors) across 100 clients (100 rows on the y-axis). **b**, **c** For this task, with $\alpha = 1.0$ (i.e., closer to IID), all federated models perform better when trained for fixed 300 rounds, as compared to $\alpha = 0.1$, as expected, and data consumption is also less. **e**, **f** With $\alpha = 0.1$, which represents a highly non-IID, a significant drop in performance is observed across all methods, when trained for fixed 500 communication rounds. Experiments are repeated for two independent trials. The average is plotted, with a shaded area indicating the standard deviation

Table 2 Performance (testing accuracy) of different federated models on MNIST and FMNIST datasets

Non-IIdness	Models	Parameters	MNIST	FMNIST
$\alpha = 1$	FedMLP	157450	91.8	92.0
	FedCNN	156560	96.9	96.5
	FedTN	157820	95.5	95.3
$\alpha = 0.1$	FedMLP	157450	87.3	89.1
	FedCNN	156560	92.1	91.3
	FedTN	157820	89.9	90.7

Figure 4b and c presents the results of training MNIST and FMNIST datasets with federated MLP, CNN, and TN models, distributing the data among 100 clients with $\alpha = 1.0$. The gray line represents the target accuracy achieved by centralized training of the TN model. In contrast to MNIST and FMNIST, the FedTN model demonstrates superior performance compared to the FedMLP model, as shown in Table 1. FedCNN performed marginally better than other methods. Lastly, in Fig. 4e and f, results $\alpha = 0.1$ are illustrated. We noticed a significant drop in performance in all federated models due to a high level of heterogeneity in data among the clients. When the MPS tensor network (MPS TN) is trained for MNIST classification, it achieves an accuracy of 95.3%, outperforming the MLP model, which achieves 92%. However, FedCNN achieves a higher accuracy of 96.5%. For FMNIST, on the other hand, FedTN achieves 90.7%, outperforming FedMLP at 89.6% and closely approaching the performance of FedCNN, which stands at 91.3% on training with fixed 500 communication rounds, as shown in Fig. 5a and b. Moreover, FedTN demonstrates effective generalization across clients, with the global model significantly outperforming the local average accuracy of the client models, regardless of how the data is distributed among them, as shown in Fig. 5a and b. We observed that all models exhibit variability in their results due to the non-IID nature of the data distributed among the clients, as shown in Fig. 4e and f.

3.3 Existing federation to accept new clients

In our FedTN environment, the strategy of incrementally adding new clients proved effective in evaluating the system's scalability and performance. Initially, we trained the TN model with the first 10 clients for 20 rounds, which allowed us to establish a baseline for performance and data transmission. By introducing an additional 10 clients every 20 rounds, we tested the model's ability to generalize and aggregate updates from an increasing number of clients. This incremental approach not only provided insights into how well the system handles increasing client diversity but also revealed the impact of additional clients on data transmission and computational demands. For 100 rounds, we evaluated the performance of FedTN for MNIST and FMNIST with Dir $\alpha \in \{0.1, 1.0\}$, as shown in Fig. 6. This methodology demonstrated that while adding more clients improved model

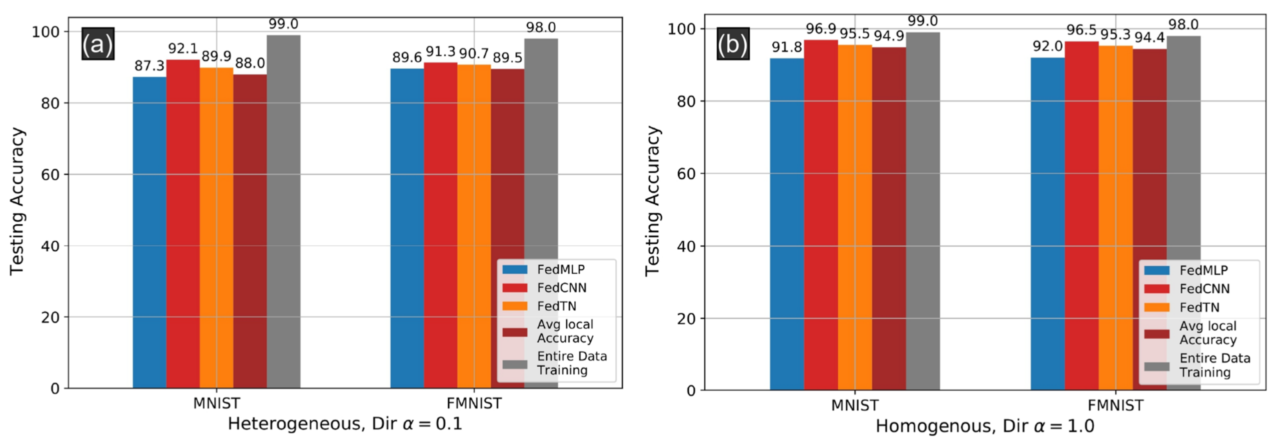

Fig. 5 Comparison among various federated learning methods when trained with a limited number of communication rounds. **a** Heterogeneous data partition (Dir $\alpha = 0.1$): The performance of the FedTN global model outperformed that of the FedMLP and the average local accuracy of the local clients, regardless of how the data was distributed among the clients. The gray color indicates the centralized performance of the TN model, and the number of trainable parameters is similar across all methods. **b** Homogeneous data partition (Dir $\alpha = 1.0$): All federated models perform better, as compared to non-IID distribution $\alpha = 0.1$. The performance of FedTN is on par with well-known classical CNN and MLP models

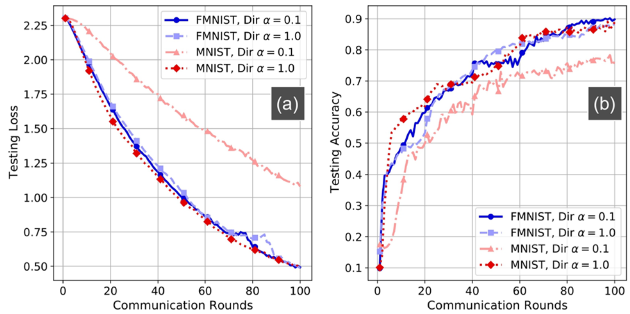

Fig. 6 Performance of the existing federation to accommodate new clients. Incrementally adding new clients to the existing federated system. An additional 10 clients are added after every 20 rounds to demonstrate the TN model's ability to generalize and aggregate updates from a growing number of clients

robustness and generalization, it also increased the total data transmitted, highlighting the trade-offs between client participation and communication efficiency in federated learning.

3.4 Effect of local training epochs

Next, we evaluated the impact of local training epochs on the FedTN model for the FMNIST dataset with a Dirichlet distribution of $\alpha = 0.1$. Nevertheless, the improvement in convergence speed appears minimal relative to the additional computation time needed; specifically, using E=20 approximately doubles the total wall-clock time compared to using E=10, as shown in Fig. 7. We find that extending the training duration enhances FedTN performance, supporting our assumption that FedTN is most effective when applied to local client models of superior quality. However, FedTN achieves strong performance even with just one local epoch per client ($E = 1$), demonstrating its effectiveness, in Fig. 4.

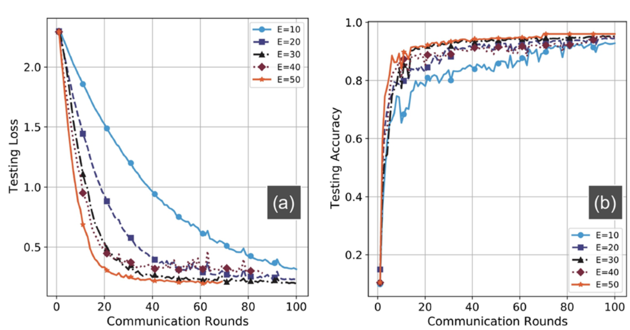

Fig. 7 Effect of local training epochs. The effect of the different number of local epochs (E), while fixing the experimental setup to be the classification of FMNIST Dir $\alpha = 0.1$ distribution, trained with a FedTN. Using a higher E leads to slightly faster convergence

4 Conclusion

We presented federated learning based on tensor networks (FedTNs), which has shown the capability to enhance the effectiveness and convergence of federated learning TN models, ultimately leading to improved performance across diverse client data. Our proposed FedTN model surpasses multiplayer perceptron and closely approaches convolutional neural networks' performance in fixed communication rounds. We ensured that all models have an equal number of trainable parameters, so the total amount of data exchanged between the data center and the clients is nearly identical across all models. The global FedTN model significantly outperforms the average accuracy of locally trained models, achieving a testing accuracy of 95% and demonstrating smoother convergence. As research advances, the potential benefits of applying quantum-inspired methods to FL are likely to become increasingly significant. In future work, we want to extend FedTN to color images, thereby addressing higher-dimensional image spaces. These enhancements will broaden the applicability and robustness of FedTN across various complex datasets and sensitive applications.

Acknowledgements We would like to acknowledge the partial internal support from the ECE Department at NC State University and the U.S. Department of Energy (DOE) (Office of Basic Energy Sciences), under Award No. DE-SC0019215.

Author Contributions A.S.B. supervised and designed the algorithm and wrote the manuscript, while M.K.S. contributed to the manuscript and handled the implementation under the supervision of S.K.

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Ali A, Delgado I, Leceta A (2024) Quantum-inspired techniques in tensor networks for industrial contexts. ArXiv Preprint ArXiv:2404.11277

Bhatia A, Saggi M, Kumar A, Jain S (Jun.2019) Matrix product state-based quantum classifier. Neural Comput 31(6):1499–1517

Bhatia A, Kais S, Alam M (2023) Federated quanvolutional neural network: a new paradigm for collaborative quantum learning. Quant Sci Technol 8

Bhatia A, Kais S, Alam M (2023) Handling privacy-sensitive clinical data with federated quantum machine learning. APS March Meet Abstr 2023:T70-007

Bhatia A, Kais S, Alam M (2024) Robustness of quantum federated learning (QFL) against "Label Flipping Attacks" for lithography hotspot detection in semiconductor manufacturing. 2024 IEEE international reliability physics symposium (IRPS), pp 1-4

Bhatia A, Neira D (2024) Federated learning with tensor networks: a quantum AI framework for healthcare. Mach Learn Sci Technol 5

Bhatia A, Saggi M, Kais S (2023) Quantum machine learning predicting ADME-Tox properties in drug discovery. J Chem Inf Model 63:6476–6486

Bhatia A, Saggi M, Kais S (2024) Communication-efficient quantum federated learning optimization for multi-center healthcare data. IEEE-EMBS International Conference on Biomedical and Health Informatics

Brandhofer S, Braun D, Dehn V, Hellstern G, Hüls M, Ji Y, Polian I, Bhatia A, Wellens T (2022) Benchmarking the performance of portfolio optimization with QAOA. Quant Inf Process 22:25

Cichocki A (2014) Tensor networks for big data analytics and large-scale optimization problems. arXiv preprint arXiv:1407.3124

Fernández-González C, Schuch N, Wolf M, Cirac JI, Perez-Garcia D (2015) Frustration free gapless Hamiltonians for matrix product states. Commun Math Phys 333(1):299–333

Grant E, Benedetti M, Cao S, Hallam A, Lockhart J, Stojevic V, Green A, Severini S (2018) Hierarchical quantum classifiers. npj Quant Inf 4:65

Huggins W, Patil P, Mitchell B, Whaley K, Stoudenmire E (2019) Towards quantum machine learning with tensor networks. Quant Sci Technol 4(2)

Jahn A, Eisert J (2021) Holographic tensor network models and quantum error correction: a topical review. Quant Sci Technol 6(3)

Javeed D, Saeed M, Ahmad I, Adil M, Kumar P, Islam A (2024) Quantum-empowered federated learning and 6G wireless networks for IoT security: concept, challenges and future directions. Futur Gener Comput Syst

Levine Y, Sharir O, Cohen N, Shashua A (2019) Quantum entanglement in deep learning architectures. Phys Rev Lett 122:065301

Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, Liu X, He B (2021) A survey on federated learning systems: vision, hype and reality for data privacy and protection. IEEE Trans Knowl Data Eng 35(7):3347–3366

Li C, Kumar N, Song Z, Chakrabarti S, Pistoia M (2024) Privacy-preserving quantum federated learning via gradient hiding. Quant Sci Technol 9

Liu Y, Li W, Zhang X, Lewenstein M, Su G, Ran S (2021) Entanglement-based feature extraction by tensor network machine learning. Front Appl Math Stat 7

McMahan B, Moore E, Ramage D, Hampson S, Arcas B (2017) Communication-efficient learning of deep networks from decentralized data. Artif Int Stat, pp 1273-1282

Mendieta M, Yang T, Wang P, Lee M, Ding Z, Chen C (2022) Local learning matters: rethinking data heterogeneity in federated learning. in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 8397-8406

Monturiol J, Pérez-García D, Pozas-Kerstjens A (2024) TensorKrowch: smooth integration of tensor networks in machine learning. Quantum 8:1364

Nguyen D, Ding M, Pathirana P, Seneviratne A, Li J, Poor H (2021) Federated learning for internet of things: a comprehensive survey. IEEE Commun Surv Tutorials 23(3):1622–1658

Orús R, Mugel S, Lizaso E (2019) Quantum computing for finance: overview and prospects. Rev Phys 4

Paquet E, Soleymani F (2022) QuantumLeap: hybrid quantum neural network for financial predictions. Expert Syst Appl 195:116583

Rieser H, Köster F, Raulf A (2023) Tensor networks for quantum machine learning. Proc R Soc A 479(20230218)

Saggi M, Bhatia A, Isaiah M, Gowher H, Kais S (2024) Multi-Omic and quantum machine learning integration for lung subtypes classification. ArXiv Preprint ArXiv:2410.02085

Saggi M, Bhatia A, Kais S (2024) Federated quantum machine learning for drug discovery and healthcare. Ann Rep Comput Chem 20:269–322

Selvan R, Ørting S, Dam E (2020) Multi-layered tensor networks for image classification. arXiv preprint arXiv:2011.06982

Sengupta R, Adhikary S, Oseledets I, Biamonte J (2022) Tensor networks in machine learning. Eur Math Soc Mag, pp 4-12

Shin S, Teo Y, Jeong H (2024) Dequantizing quantum machine learning models using tensor networks. Phys Rev Res 6(023218)

Wall M, Abernathy M, Quiroz G (2021) Generative machine learning with tensor networks: benchmarks on near-term quantum computers. Phys Rev Res 3

Wang K, Xiao L, Yi W, Ran S, Xue P (2021) Experimental realization of a quantum image classifier via tensor-network-based machine learning. Photonics Res 9:2332–2340

Wu D, Rossi R, Vicentini F, Carleo G (2023) From tensor network quantum states to tensorial recurrent neural networks. Online Available: https://arxiv.org/abs/2206.12363