Layerwise Federated Learning for Heterogeneous Quantum Clients Using Quorus

Jason Han, Nicholas S. DiBrita, Daniel Leeds, Jianqiang Li, Jason Ludmir, Tirthak Patel

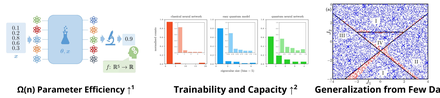
Rice University

Paper: **Code:**

Introduction

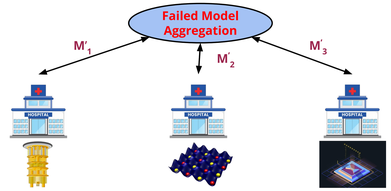
Why Quantum Computers for Machine Learning?

- For certain types of problems, quantum computers can be good alternatives to classical computers for ML



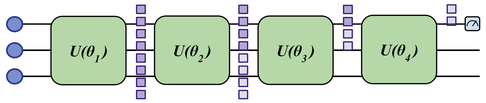
Why Federated Learning?

- Keeps user data private on-device while still allowing for distributed training, but model aggregation is challenging



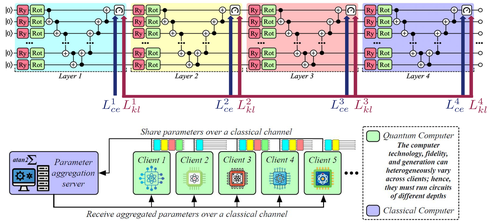
Why Can't Quantum Clients All Run The Same Model?

- Different device error rates imply that devices have different depths at which high-fidelity computation can run



How Can We Allow Different Depths In Federated Learning?

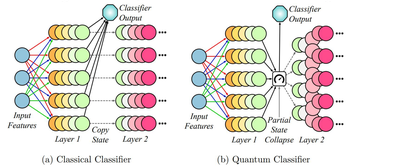
- Utilize layerwise loss + reverse distillation[4] for homogenous objectives, but unclear how to do early-exits in QNN's



Design

How To Perform Early-Exits in Quantum Classifiers?

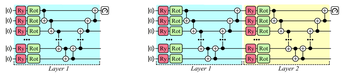
- In classical NN's, early-exits are obtained by copying the activations and feeding to a bottleneck layer
- In quantum NN's, arbitrary activations cannot be copied (no-cloning theorem), and measurement changes the state



(a) Classical Classifier (b) Quantum Classifier

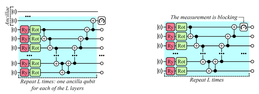
Design 1: Layerwise Measurement

- **Pros**: Early-exits are 'correctly' computed; rerun the circuit
- **Cons**: Shot count scales linearly with the number of layers



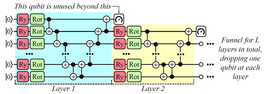
Designs 2/3: Ancilla/Blocking Measurement

- **Pros**: No additional shot-overhead to obtain early-exits
- **Cons**: Dephase first qubit, so unable to manipulate phase



Design 4: Funnel Measurement

- **Pros**: No shot-overhead, ancillae mid-circuit measurement
- **Cons**: Hilbert space is smaller for later layers



Each Design Is Uniquely Tailored for Hardware Constraints

- Each design has a clear use-case based on capabilities

Model	↑ Shot Budget	↑ Qubit Count	Midcirc. Meas.	↓ Hilbert Space
Layerwise	✓	✗	✗	✗
Ancilla	✗	✓	✗	✗
Blocking	✗	✗	✓	✗
Funnel	✗	✗	✗	✓

Results

Experimental Analysis:

- We evaluate Quorus to 3 baselines: (1) Q-HeteroFL, (2) Vanilla QFL, and (3) Standalone; we use 10 qubits

Quorus Beats Baseline in Most Cases (avg. 12.4% ↑)

- With modified loss + ensembled inference (see paper)

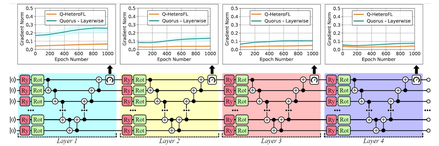
Capacity	Technique	MNIST 0/1	MNIST 3/4	MNIST 4/9	Fashion-MNIST Trouser/Boot	Fashion-MNIST Bag/Sandal	Fashion-MNIST Pullover/Coat
2L	Q-HeteroFL	90.3 ± 5.2 (↓ 7.9)	58.8 ± 12.6 (↓ 37.3)	59.3 ± 7.3 (↓ 20.7)	62.4 ± 33.5 (↓ 36.4)	67.1 ± 14.6 (↓ 24.8)	58.9 ± 6.1 (↓ 18.0)
	Vanilla QFL (2L)	99.2 ± 0.1 (↓ 0.0)	96.0 ± 1.2 (↓ 0.1)	**80.0 ± 0.9**	98.5 ± 1.1 (↓ 0.3)	**91.9 ± 1.2**	**76.9 ± 0.4**
	Standalone	**99.2 ± 0.3**	**96.1 ± 1.2**	78.5 ± 2.5 (↓ 1.5)	98.2 ± 0.9 (↓ 0.5)	91.2 ± 1.0 (↓ 0.7)	74.9 ± 1.8 (↓ 2.0)
	Quorus-Layerwise	97.0 ± 1.4 (↓ 1.2)	95.0 ± 1.2 (↓ 1.1)	78.2 ± 0.6 (↓ 1.8)	**98.8 ± 0.9**	86.1 ± 8.1 (↓ 5.8)	76.3 ± 1.4 (↓ 0.6)
3L	Q-HeteroFL	79.6 ± 14.6 (↓ 16.7)	65.0 ± 3.8 (↓ 31.9)	68.5 ± 4.7 (↓ 11.9)	76.9 ± 15.0 (↓ 22.3)	79.1 ± 12.0 (↓ 13.0)	59.5 ± 10.5 (↓ 19.1)
	Vanilla QFL (3L)	96.0 ± 1.8 (↓ 0.3)	96.0 ± 1.2 (↓ 0.9)	80.0 ± 0.9 (↓ 0.4)	98.3 ± 1.1 (↓ 0.7)	91.9 ± 1.2 (↓ 0.2)	76.9 ± 0.4 (↓ 1.7)
	Standalone	**98.3 ± 1.2**	95.5 ± 1.7 (↓ 1.4)	79.6 ± 3.4 (↓ 0.8)	99.1 ± 0.6 (↓ 0.1)	**92.1 ± 2.4**	76.5 ± 2.5 (↓ 2.1)
	Quorus-Layerwise	96.0 ± 1.0 (↓ 0.3)	**96.9 ± 0.7**	**80.4 ± 2.4**	**99.2 ± 0.4**	89.2 ± 5.6 (↓ 2.9)	**78.6 ± 1.0**
4L	Q-HeteroFL	80.4 ± 7.6 (↓ 17.9)	88.6 ± 7.3 (↓ 9.5)	68.8 ± 5.0 (↓ 13.1)	90.5 ± 6.6 (↓ 8.8)	88.6 ± 2.0 (↓ 5.1)	72.8 ± 3.2 (↓ 5.9)
	Vanilla QFL (2L)	99.2 ± 0.4 (↓ 0.1)	96.0 ± 1.2 (↓ 1.5)	80.0 ± 0.9 (↓ 1.9)	98.5 ± 1.1 (↓ 0.8)	91.9 ± 1.2 (↓ 1.6)	76.9 ± 0.4 (↓ 1.8)
	Standalone	98.0 ± 2.5 (↓ 1.3)	97.4 ± 0.5 (↓ 0.1)	81.2 ± 3.7 (↓ 0.7)	98.9 ± 1.1 (↓ 0.4)	**93.7 ± 1.1**	77.1 ± 1.0 (↓ 1.6)
	Quorus-Layerwise	**98.3 ± 0.9**	**97.5 ± 0.6**	**81.9 ± 2.2**	**99.3 ± 0.3**	91.5 ± 4.0 (↓ 2.2)	**78.7 ± 1.0**
5L	Q-HeteroFL	88.9 ± 3.7 (↓ 9.4)	88.5 ± 5.3 (↓ 9.0)	71.0 ± 4.4 (↓ 11.5)	94.6 ± 4.3 (↓ 4.7)	88.4 ± 3.9 (↓ 5.9)	72.8 ± 3.2 (↓ 6.0)
	Vanilla QFL (3L)	98.2 ± 0.4 (↓ 0.3)	96.0 ± 1.2 (↓ 1.5)	80.0 ± 0.9 (↓ 2.5)	98.5 ± 1.1 (↓ 0.8)	91.9 ± 1.2 (↓ 2.4)	76.9 ± 0.4 (↓ 1.9)
	Standalone	97.2 ± 1.7 (↓ 1.3)	96.1 ± 1.9 (↓ 1.1)	80.3 ± 3.2 (↓ 2.2)	98.5 ± 0.5 (↓ 0.9)	**94.3 ± 0.6**	77.6 ± 1.6 (↓ 1.2)
	Quorus-Layerwise	**98.5 ± 0.8**	**97.5 ± 0.4**	**82.5 ± 2.5**	**99.3 ± 0.2**	92.4 ± 2.6 (↓ 1.9)	**78.8 ± 1.1**
6L	Q-HeteroFL	88.6 ± 6.2 (↓ 10.0)	85.1 ± 5.9 (↓ 12.7)	73.9 ± 4.4 (↓ 9.2)	95.3 ± 0.8 (↓ 4.1)	92.1 ± 1.3 (↓ 3.2)	74.4 ± 0.9 (↓ 4.4)
	Vanilla QFL (3L)	98.2 ± 0.4 (↓ 0.1)	96.0 ± 1.2 (↓ 1.8)	80.0 ± 0.9 (↓ 1.3)	98.5 ± 1.1 (↓ 0.9)	91.9 ± 1.2 (↓ 3.4)	76.9 ± 0.4 (↓ 1.9)
	Standalone	98.3 ± 1.0 (↓ 0.3)	96.3 ± 0.8 (↓ 1.3)	80.8 ± 3.1 (↓ 2.7)	98.3 ± 0.8 (↓ 1.1)	**95.3 ± 1.0**	75.8 ± 1.5 (↓ 3.4)
	Quorus-Layerwise	**99.6 ± 0.8**	**97.8 ± 0.2**	**83.1 ± 2.4**	**99.4 ± 0.3**	92.7 ± 2.5 (↓ 2.6)	**78.8 ± 0.8**

Quorus Variants Have Similar Performance (avg. 0.9% diff.)

Capacity	Technique	MNIST 0/1	MNIST 3/4	MNIST 4/9	Fashion-MNIST Trouser/Boot	Fashion-MNIST Bag/Sandal	Fashion-MNIST Pullover/Coat
2L	Quorus-Layerwise	97.0 ± 1.4 (↓ 0.4)	95.0 ± 1.2 (↓ 0.3)	78.2 ± 0.6 (↓ 1.5)	**98.8 ± 0.9**	86.1 ± 8.1 (↓ 0.3)	76.3 ± 1.4 (↓ 0.2)
	Quorus-Ancilla/Blocking	97.0 ± 1.0 (↓ 0.4)	94.8 ± 1.5 (↓ 0.5)	78.3 ± 1.2 (↓ 1.4)	98.6 ± 0.9 (↓ 0.2)	86.2 ± 8.4 (↓ 0.2)	**76.5 ± 1.3**
	Quorus-Funnel	**97.4 ± 1.2**	**95.3 ± 1.5**	**79.7 ± 2.0**	98.1 ± 1.1 (↓ 0.7)	**86.4 ± 6.8**	76.4 ± 1.2 (↓ 0.1)
3L	Quorus-Layerwise	98.0 ± 1.0 (↓ 0.1)	96.9 ± 0.7 (↓ 0.0)	80.4 ± 2.4 (↓ 1.9)	**99.2 ± 0.4**	89.2 ± 5.9 (↓ 1.9)	78.6 ± 1.0 (↓ 0.1)
	Quorus-Ancilla/Blocking	97.9 ± 1.2 (↓ 0.2)	**96.9 ± 0.6**	81.4 ± 2.0 (↓ 0.9)	99.2 ± 0.5 (↓ 0.0)	88.7 ± 7.5 (↓ 1.5)	78.5 ± 1.2 (↓ 0.2)
	Quorus-Funnel	**98.1 ± 0.5**	96.3 ± 0.6 (↓ 0.6)	**82.3 ± 1.8**	99.1 ± 0.6 (↓ 0.1)	**90.2 ± 3.3**	**78.7 ± 1.1**
4L	Quorus-Layerwise	98.3 ± 0.9 (↓ 0.0)	**97.5 ± 0.6**	81.9 ± 2.2 (↓ 1.3)	**99.3 ± 0.3**	91.5 ± 4.0 (↓ 0.8)	78.7 ± 1.0 (↓ 0.7)
	Quorus-Ancilla/Blocking	98.1 ± 1.2 (↓ 0.2)	97.3 ± 0.5 (↓ 0.2)	81.5 ± 1.9 (↓ 1.7)	99.2 ± 0.5 (↓ 0.1)	90.3 ± 5.5 (↓ 2.0)	78.9 ± 1.0 (↓ 0.5)
	Quorus-Funnel	**98.3 ± 0.7**	97.1 ± 0.6 (↓ 0.4)	**83.2 ± 2.4**	99.0 ± 0.6 (↓ 0.3)	**92.3 ± 1.7**	**79.4 ± 1.0**
5L	Quorus-Layerwise	98.5 ± 0.8 (↓ 0.0)	**97.5 ± 0.4**	82.5 ± 2.5 (↓ 2.1)	**99.3 ± 0.2**	92.4 ± 2.6 (↓ 0.3)	78.8 ± 1.3 (↓ 1.5)
	Quorus-Ancilla/Blocking	98.3 ± 0.7 (↓ 0.2)	97.4 ± 0.5 (↓ 0.1)	81.9 ± 2.5 (↓ 2.7)	99.3 ± 0.3 (↓ 0.0)	91.1 ± 4.3 (↓ 1.6)	79.0 ± 1.4 (↓ 1.4)
	Quorus-Funnel	**99.5 ± 0.7**	97.1 ± 0.5 (↓ 0.4)	**84.6 ± 2.1**	99.1 ± 0.4 (↓ 0.2)	**92.7 ± 1.2**	**80.4 ± 1.0**
6L	Quorus-Layerwise	**98.6 ± 0.8**	**97.8 ± 0.2**	83.1 ± 2.4 (↓ 2.1)	**99.4 ± 0.3**	92.7 ± 2.5 (↓ 0.7)	78.8 ± 0.8 (↓ 1.5)
	Quorus-Ancilla/Blocking	98.4 ± 0.6 (↓ 0.2)	97.5 ± 0.5 (↓ 0.3)	82.2 ± 2.0 (↓ 3.0)	99.3 ± 0.3 (↓ 0.1)	91.4 ± 4.0 (↓ 2.0)	78.8 ± 1.1 (↓ 1.5)
	Quorus-Funnel	98.0 ± 0.7 (↓ 0.6)	97.1 ± 0.4 (↓ 0.7)	**85.2 ± 0.7**	99.1 ± 0.4 (↓ 0.3)	**93.4 ± 0.9**	**80.3 ± 0.9**

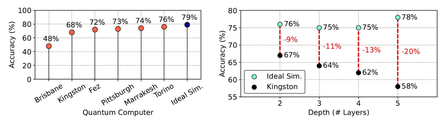
Quorus-Layerwise Achieves Higher Gradient Magnitude

- Due to early-exits, earlier parameters have ↑ gradients



Quorus-Funnel Is Validated on Real IBM QPU's

- Accuracy is heterogeneous across different QPU's
- Accuracy is heterogeneous across different model depths



(a) Same Model, Different Machines (b) Different Models, Same Machine

Additional Results

Quorus Beats Baseline in 4-Class Classification (avg. 24.0% ↑)

- See paper for details on resource costs for multiclass tasks

Capacity	Technique	MNIST 0/1/6/7	MNIST 0/2/4/7	MNIST 3/5/8/9	Fashion-MNIST Trouser/Sandal/Sneaker/Bag	Fashion-MNIST Pullover/Dress/Coat/Boot	Fashion-MNIST Top/Pullover/Coat/Shirt
2L	Q-HeteroFL	30.7 ± 15.0 (↓ 46.0)	33.9 ± 17.5 (↓ 31.3)	35.5 ± 9.6 (↓ 20.4)	32.2 ± 8.2 (↓ 46.4)	43.4 ± 12.8 (↓ 27.0)	36.4 ± 6.0 (↓ 12.5)
	Quorus-Layerwise	73.2 ± 5.2 (↓ 3.5)	64.6 ± 7.4 (↓ 0.6)	**55.9 ± 3.8**	65.5 ± 9.3 (↓ 13.1)	65.6 ± 4.2 (↓ 4.8)	44.7 ± 4.6 (↓ 4.2)
	Quorus-Funnel	**76.7 ± 3.7**	**65.2 ± 10.1**	55.0 ± 5.9 (↓ 0.9)	**78.4 ± 1.6**	**70.4 ± 1.6**	**49.9 ± 3.4**
3L	Q-HeteroFL	39.5 ± 12.9 (↓ 41.4)	46.3 ± 8.7 (↓ 32.9)	35.4 ± 5.0 (↓ 25.6)	50.3 ± 10.4 (↓ 30.5)	49.3 ± 6.9 (↓ 19.0)	38.8 ± 12.6 (↓ 35.3)
	Quorus-Layerwise	**80.9 ± 2.7**	71.1 ± 6.2 (↓ 8.1)	**61.0 ± 4.7**	71.4 ± 6.4 (↓ 9.4)	69.1 ± 3.8 (↓ 5.0)	52.2 ± 2.6 (↓ 3.9)
	Quorus-Funnel	80.1 ± 4.6 (↓ 0.8)	**79.2 ± 6.7**	60.6 ± 5.4 (↓ 0.4)	**80.8 ± 4.8**	**74.1 ± 1.2**	**56.1 ± 3.0**
4L	Q-HeteroFL	53.9 ± 13.9 (↓ 29.4)	50.6 ± 9.4 (↓ 27.7)	48.1 ± 4.7 (↓ 17.6)	56.4 ± 10.4 (↓ 25.5)	55.1 ± 6.9 (↓ 19.0)	40.1 ± 5.7 (↓ 17.0)
	Quorus-Layerwise	**83.3 ± 4.2**	76.5 ± 5.3 (↓ 1.8)	**65.7 ± 2.4**	75.5 ± 5.8 (↓ 6.4)	72.1 ± 3.8 (↓ 2.0)	52.1 ± 0.8 (↓ 3.0)
	Quorus-Funnel	83.1 ± 4.0 (↓ 2.2)	**78.3 ± 7.2**	63.6 ± 4.4 (↓ 2.1)	**81.8 ± 4.5**	**74.1 ± 2.5**	**57.1 ± 3.4**
5L	Q-HeteroFL	58.0 ± 20.7 (↓ 27.0)	54.4 ± 13.8 (↓ 24.2)	42.8 ± 3.5 (↓ 24.6)	54.6 ± 11.0 (↓ 27.7)	49.5 ± 10.5 (↓ 24.6)	47.3 ± 5.5 (↓ 27.6)
	Quorus-Layerwise	**85.0 ± 3.3**	78.0 ± 3.8 (↓ 0.6)	66.2 ± 2.7 (↓ 1.2)	78.5 ± 4.6 (↓ 3.8)	73.5 ± 2.9 (↓ 1.4)	55.4 ± 0.8 (↓ 2.7)
	Quorus-Funnel	80.4 ± 3.9 (↓ 4.6)	**78.6 ± 6.8**	**67.4 ± 3.1**	**82.3 ± 3.3**	**74.9 ± 3.2**	**58.1 ± 2.9**
6L	Q-HeteroFL	71.7 ± 6.3 (↓ 14.5)	71.4 ± 2.8 (↓ 8.1)	53.3 ± 2.5 (↓ 14.3)	71.3 ± 2.5 (↓ 11.2)	68.6 ± 7.5 (↓ 14.6)	46.9 ± 3.7 (↓ 10.5)
	Quorus-Layerwise	**86.2 ± 2.9**	**79.5 ± 3.3**	**67.6 ± 2.9**	80.1 ± 4.2 (↓ 2.4)	74.3 ± 2.2 (↓ 0.9)	56.0 ± 1.1 (↓ 1.4)
	Quorus-Funnel	79.3 ± 5.2 (↓ 7.9)	77.0 ± 7.7 (↓ 2.5)	66.0 ± 3.3 (↓ 1.6)	**82.5 ± 2.4**	**75.2 ± 3.8**	**57.4 ± 3.2**

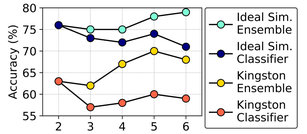
Quorus Beats Baseline with Non-IID Data (avg. 22.5% ↑)

- Non-IID data is distributed according to $p_c = \text{Dir}_k(\beta = 0.5)$

Capacity	Technique	MNIST (Non-IID) 0/1	MNIST (Non-IID) 3/4	MNIST (Non-IID) 4/9	Fashion-MNIST (Non-IID) Trouser/Boot	Fashion-MNIST (Non-IID) Bag/Sandal	Fashion-MNIST (Non-IID) Pullover/Coat
2L	Q-HeteroFL	89.1 ± 8.8 (↓ 8.6)	52.3 ± 24.0 (↓ 42.4)	60.0 ± 4.4 (↓ 16.0)	68.1 ± 38.1 (↓ 30.2)	70.2 ± 14.9 (↓ 17.7)	56.1 ± 6.0 (↓ 15.5)
	Quorus-Layerwise	**97.7 ± 0.8**	94.6 ± 3.3 (↓ 0.1)	74.2 ± 3.2 (↓ 1.8)	**98.3 ± 1.0**	**87.9 ± 3.1**	71.5 ± 7.2 (↓ 0.1)
	Quorus-Funnel	97.0 ± 2.0 (↓ 0.7)	**94.7 ± 1.5**	**76.0 ± 4.1**	98.0 ± 1.2 (↓ 0.3)	86.5 ± 8.8 (↓ 1.4)	**71.6 ± 4.8**
3L	Q-HeteroFL	77.0 ± 11.0 (↓ 21.4)	77.1 ± 16.7 (↓ 19.1)	57.3 ± 4.6 (↓ 22.1)	67.6 ± 12.1 (↓ 31.5)	62.1 ± 12.3 (↓ 28.4)	53.0 ± 3.6 (↓ 23.1)
	Quorus-Layerwise	**98.4 ± 0.5**	95.1 ± 3.5 (↓ 1.1)	79.1 ± 2.5 (↓ 0.3)	**99.1 ± 0.4**	90.5 ± 2.6 (↓ 0.0)	**76.1 ± 0.8**
	Quorus-Funnel	97.8 ± 1.3 (↓ 0.6)	**96.2 ± 1.1**	**79.4 ± 2.9**	98.6 ± 0.5 (↓ 0.5)	**90.5 ± 1.9**	75.0 ± 3.2 (↓ 1.1)
4L	Q-HeteroFL	67.8 ± 5.9 (↓ 30.6)	73.9 ± 17.9 (↓ 22.8)	56.9 ± 5.9 (↓ 25.3)	74.7 ± 19.7 (↓ 24.5)	70.7 ± 15.8 (↓ 20.8)	51.7 ± 15.8 (↓ 25.8)
	Quorus-Layerwise	**98.4 ± 0.4**	**96.7 ± 1.1**	80.6 ± 1.7 (↓ 1.6)	**99.2 ± 0.4**	**91.5 ± 3.0**	**73.5 ± 4.2**
	Quorus-Funnel	97.4 ± 2.5 (↓ 1.0)	96.5 ± 0.9 (↓ 0.2)	**82.2 ± 2.7**	98.8 ± 0.8 (↓ 0.4)	90.0 ± 3.4 (↓ 1.5)	72.4 ± 5.2 (↓ 1.1)
5L	Q-HeteroFL	80.6 ± 12.6 (↓ 17.9)	66.0 ± 16.2 (↓ 31.2)	52.7 ± 5.8 (↓ 25.7)	83.3 ± 17.0 (↓ 18.9)	72.8 ± 18.0 (↓ 19.3)	56.2 ± 5.3 (↓ 17.6)
	Quorus-Layerwise	**98.5 ± 0.4**	**97.2 ± 0.6**	**78.4 ± 1.8**	**99.2 ± 0.2**	**92.1 ± 2.7**	**73.8 ± 4.6**
	Quorus-Funnel	97.6 ± 2.1 (↓ 0.9)	96.5 ± 1.1 (↓ 0.7)	**83.4 ± 2.5**	98.7 ± 1.0 (↓ 0.5)	91.2 ± 3.4 (↓ 0.9)	72.6 ± 5.4 (↓ 1.2)
6L	Q-HeteroFL	82.1 ± 20.7 (↓ 16.6)	74.4 ± 12.7 (↓ 22.5)	59.6 ± 7.8 (↓ 22.7)	79.1 ± 17.8 (↓ 20.3)	68.6 ± 19.6 (↓ 23.6)	57.0 ± 7.9 (↓ 18.7)
	Quorus-Layerwise	**98.7 ± 0.4**	**96.9 ± 0.6**	81.9 ± 1.6 (↓ 0.4)	**99.4 ± 0.1**	**92.2 ± 1.7**	**75.7 ± 2.7**
	Quorus-Funnel	98.3 ± 1.8 (↓ 0.4)	96.1 ± 1.5 (↓ 0.8)	**82.3 ± 3.8**	98.7 ± 0.7 (↓ 0.7)	89.8 ± 8.3 (↓ 2.4)	72.7 ± 5.1 (↓ 3.0)

Quorus-Funnel Appears Robust to Noise due to Ensemble

- As each shot yields an ensemble of classifiers instead of a single classifier, Funnel seems to be more resilient to noise



Summary/Conclusions

- Introduced layerwise loss function for Heterogeneous QFL
- Introduced complementary circuit design strategies
- Improved accuracy by 12.4% on avg. over Q-HeteroFL (SOTA)
- Evaluated Quorus on real IBM superconducting QPU's

ICLR

References: [1] "Entanglement-induced provable and robust quantum learning advantages" (Zhao, Deng) (npj Quantum Info, 2025), [2] "The power of quantum neural networks" (Abbas et al.) (Nature Computational Science, 2021), [3] "Generalization in quantum machine learning from few training data" (Caro et al.) (Nature Communications, 2022), [4] "DepthFL : Depthwise Federated Learning for Heterogeneous Clients" (Kim et al.) (ICLR 2023)