

# Quantum for AI and AI for Quantum: Exploring the Benefits of Quantum Machine Learning in Representation Learning, Generalization, and Trust



**Pin-Yu Chen (IBM Research)**

**NSF Workshop on Post-Quantum AI**

**April 2024 @pinyuchenTW**

# Main Collaborators



Dr. Chao-Han Huck Yang  
Nvidia Research



Dr. Jun Qi  
Hong Kong Baptist University



Dr. Samuel Yen-Chi Chen  
Wells Fargo



# Quantum meets AI

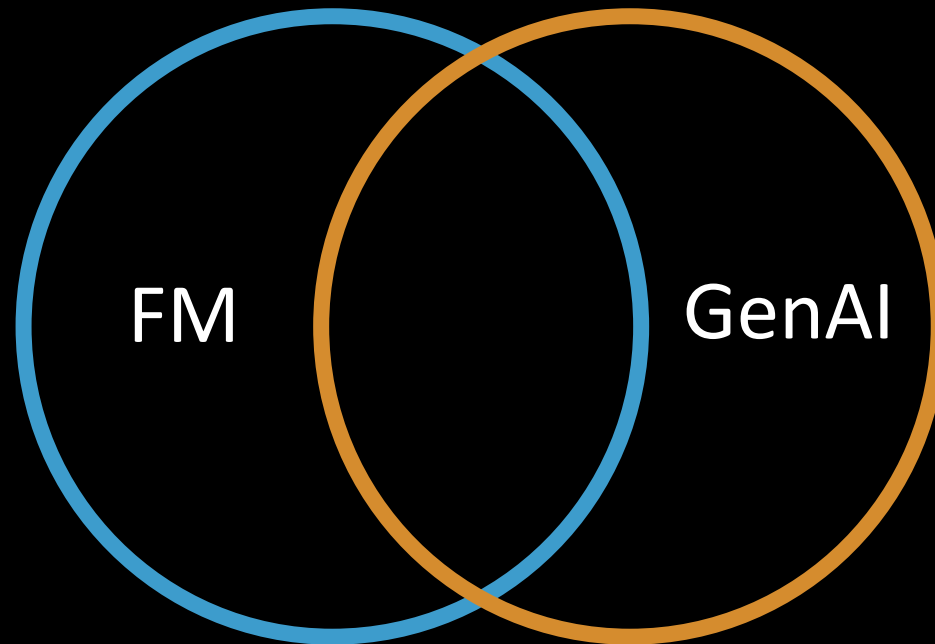
## AI for Quantum

- End-to-end Quantum Neural Networks for representation learning
- Generalization analysis of characterization of variational quantum circuit based functional regression

## Quantum for AI

- Noise can be a friend in trustworthy AI
- Use of quantum modules to improve privacy in deep learning
- (Free) certified robustness to input perturbations in quantum systems

What are Foundation Model and GenAI?



# On the Opportunities and Risks of Foundation Models

Rishi Bommasani\* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora  
 Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill  
 Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji  
 Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue  
 Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kawin Ethayarajh  
 Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman  
 Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt  
 Daniel E. Ho Jenny Hong Kyle Hsu Jing Huang Thomas Icard Saahil Jain  
 Dan Jurafsky Pratyusha Kalluri Siddharth Karamcheti Geoff Keeling Fereshte Khani  
 Omar Khattab Pang Wei Koh Mark Krass Ranjay Krishna Rohith Kuditipudi  
 Ananya Kumar Faisal Ladhak Mina Lee Tony Lee Jure Leskovec Isabelle Levent  
 Xiang Lisa Li Xuechen Li Tengyu Ma Ali Malik Christopher D. Manning  
 Suvir Mirchandani Eric Mitchell Zanele Munyikwa Suraj Nair Avaniika Narayan  
 Deepak Narayanan Ben Newman Allen Nie Juan Carlos Nieves Hamed Nilforoshan  
 Julian Nyarko Giray Ogut Laurel Orr Isabel Papadimitriou Joon Sung Park Chris Piech  
 Eva Portelance Christopher Potts Aditi Raghunathan Rob Reich Hongyu Ren  
 Frieda Rong Yusuf Roohani Camilo Ruiz Jack Ryan Christopher Ré Dorsa Sadigh  
 Shiori Sagawa Keshav Santhanam Andy Shih Krishnan Srinivasan Alex Tamkin  
 Rohan Taori Armin W. Thomas Florian Tramèr Rose E. Wang William Wang Bohan Wu  
 Jiajun Wu Yuhuai Wu Sang Michael Xie Michihiro Yasunaga Jiaxuan You Matei Zaharia  
 Michael Zhang Tianyi Zhang Xikun Zhang Yuhui Zhang Lucia Zheng Kaitlyn Zhou  
 Percy Liang\*<sup>1</sup>

Center for Research on Foundation Models (CRFM)  
 Stanford Institute for Human-Centered Artificial Intelligence (HAI)  
 Stanford University

:2108.07258v3 [cs.LG] 12 Jul 2022

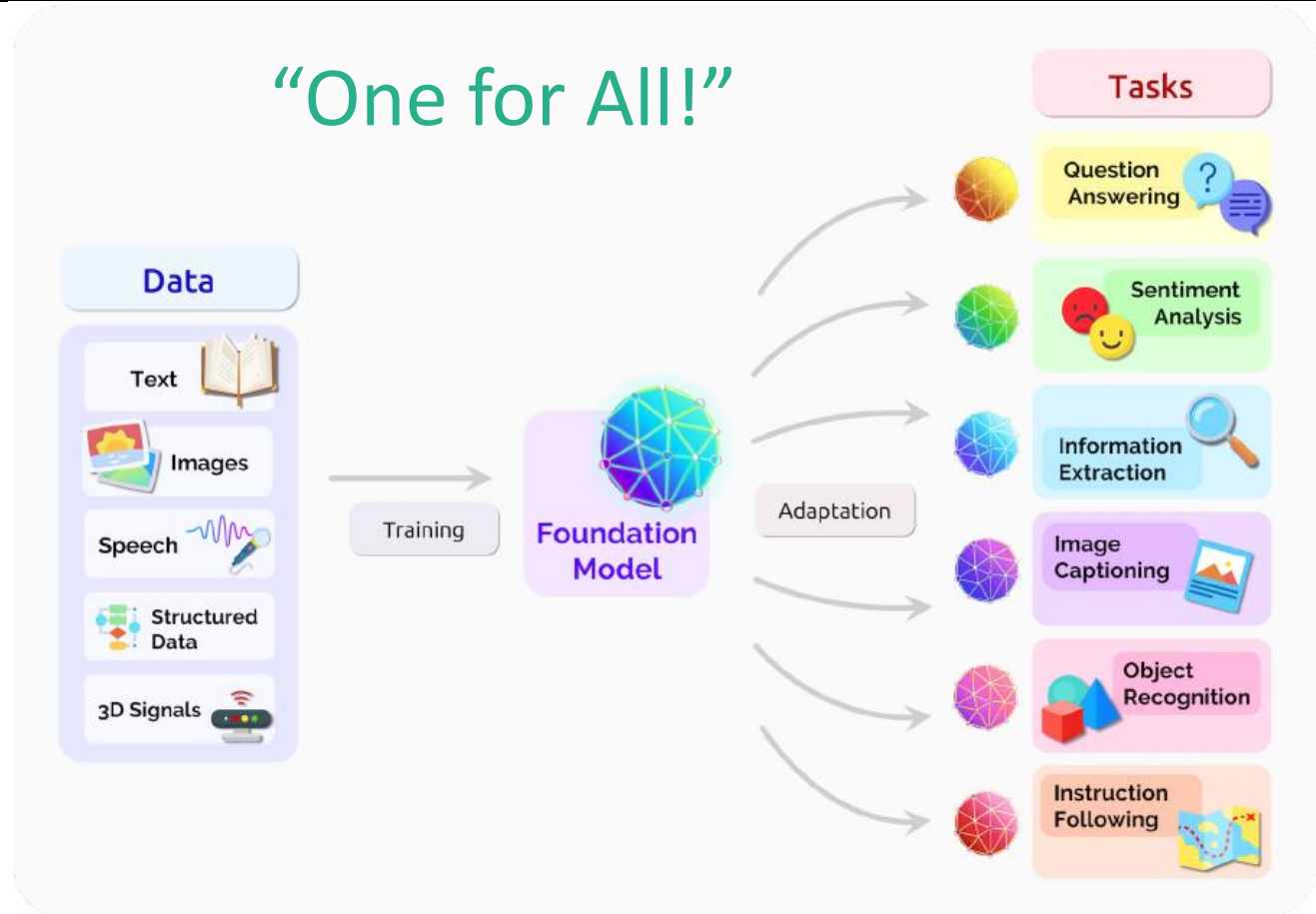
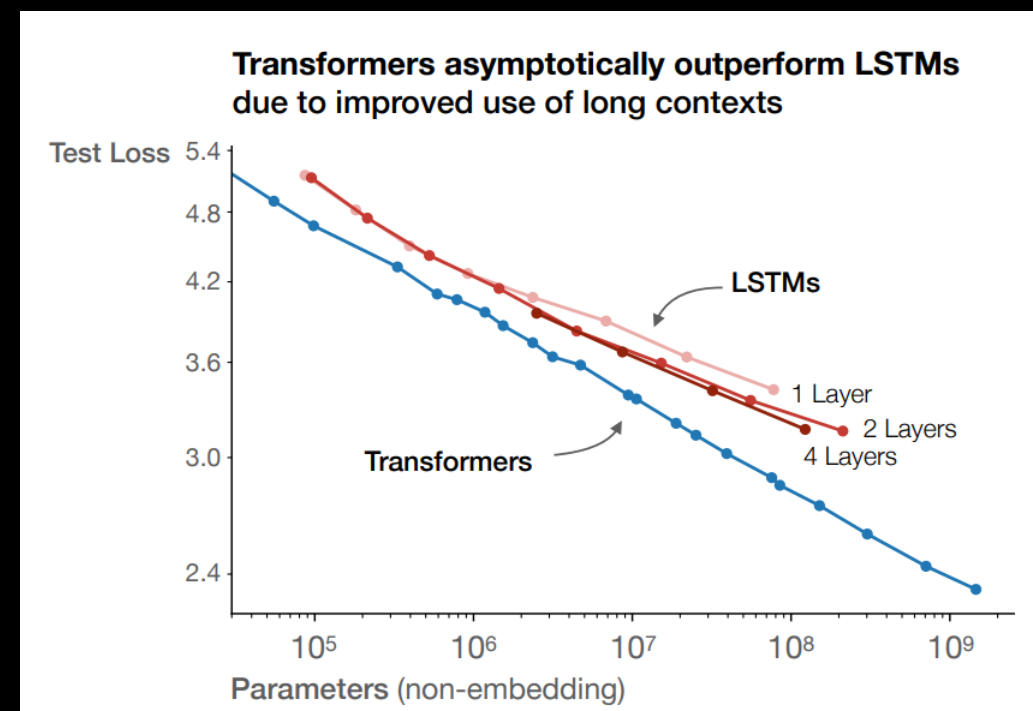
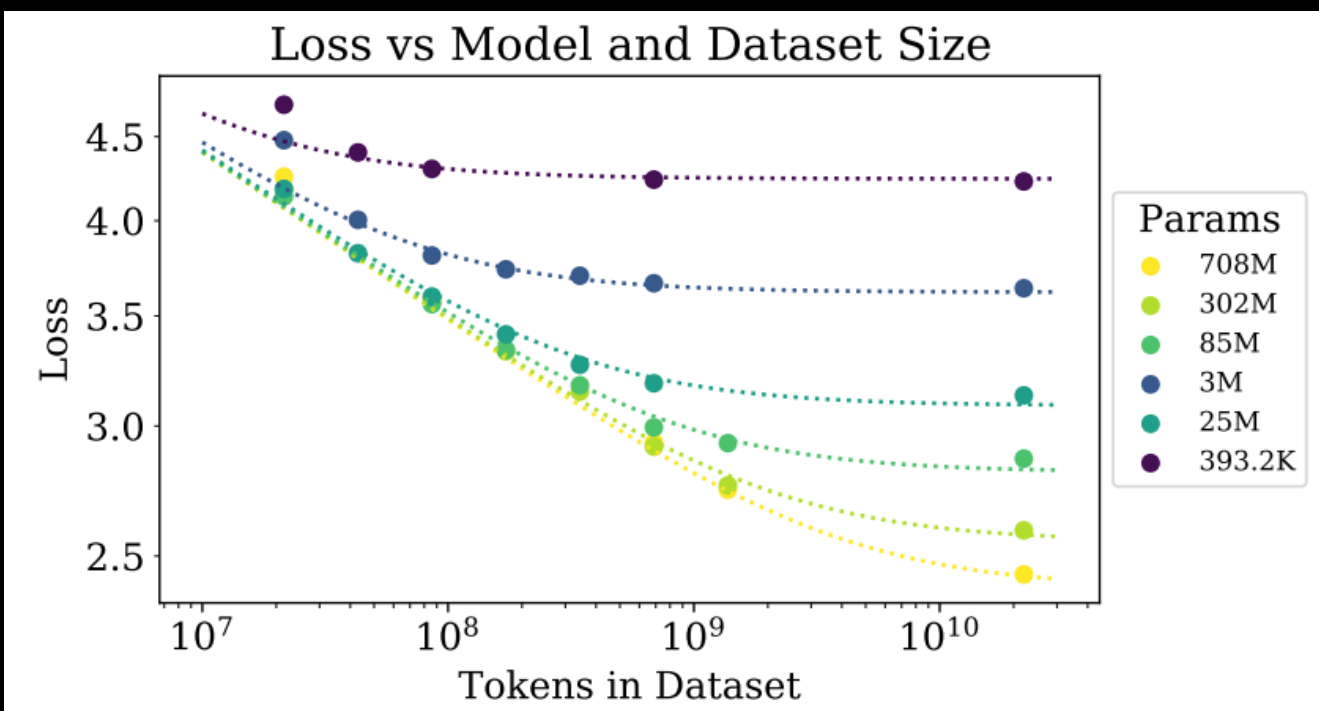


Fig. 2. A foundation model can centralize the information from all the data from various modalities. This one model can then be adapted to a wide range of downstream tasks.



# Neural Scaling Laws



Robustness at scale: data and compute hungry

# GenAI: the Beauty and the Ugly of AI-generated Content



News  
European Parliament

Headlines ▾ Press room ▾ Agenda ▾ FAQ Election Press Kit

Press room / MEPs ready to negotiate first-ever rules for safe and transparent AI

## MEPs ready to negotiate first-ever rules for safe and transparent AI

Press Releases [PLENARY SESSION](#) [IMCO](#) [LIBE](#) 14-06-2023 - 12:52

---

- Full ban on Artificial Intelligence (AI) for biometric surveillance, emotion recognition, predictive policing
- Generative AI systems like ChatGPT must disclose that content was AI-generated
- AI systems used to influence voters in elections considered to be high-risk



Administration Priorities The Record

JULY 21, 2023

## FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI

[BRIEFING ROOM](#) [STATEMENTS AND RELEASES](#)



**CDAO**  
Chief Digital and Artificial Intelligence Office

### Responsible Artificial Intelligence Test & Evaluation

Transforming the Department of Defense Through Assured AI

JATIC is developing an interoperable set of state-of-the-art software for rigorous AI model T&E



# QTN-VQC

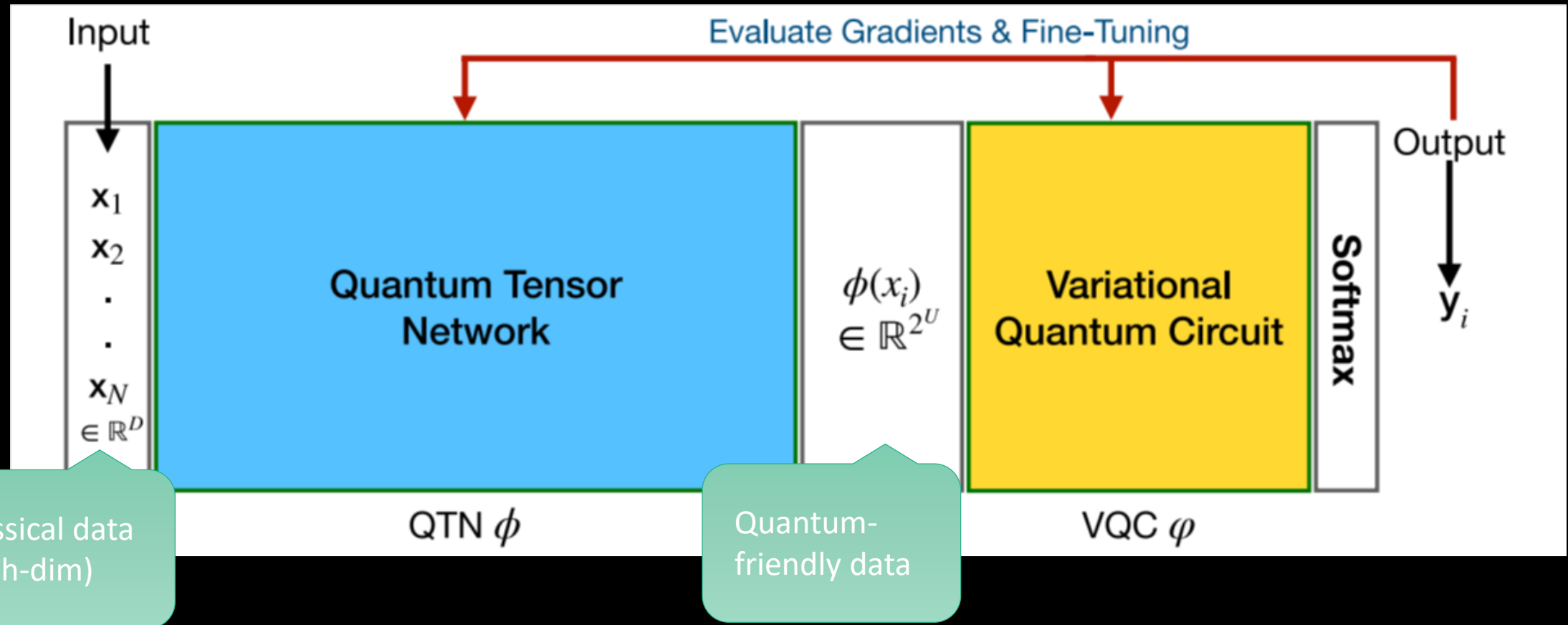
An end-to-end learning framework for quantum neural networks

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, 2023

Jun Qi, Chao-Han Huck Yang, and Pin-Yu Chen, "QTN-VQC: An End-to-End Learning framework for Quantum Neural Networks," Physica Scripta, 2023

# QTN-VQC Framework

- QTN: Quantum Tensor Network
- VQC: Variational Quantum Circuit



# TTN: Tensor-Train Network

- Encoding input N-dim data into tensor product of a unit-norm local feature map  $\phi \in R^d$

$$\Phi^{s_1 s_2 \dots s_N}(\mathbf{x}) = \phi^{s_1}(x_1) \otimes \phi^{s_2}(x_2) \otimes \dots \otimes \phi^{s_N}(x_N)$$

- Example:  $x_j$  is a pixel value ranging from [0,1]

$$\phi^{s_j}(x_j) = \left[ \cos\left(\frac{\pi}{2}x_j\right), \sin\left(\frac{\pi}{2}x_j\right) \right]$$

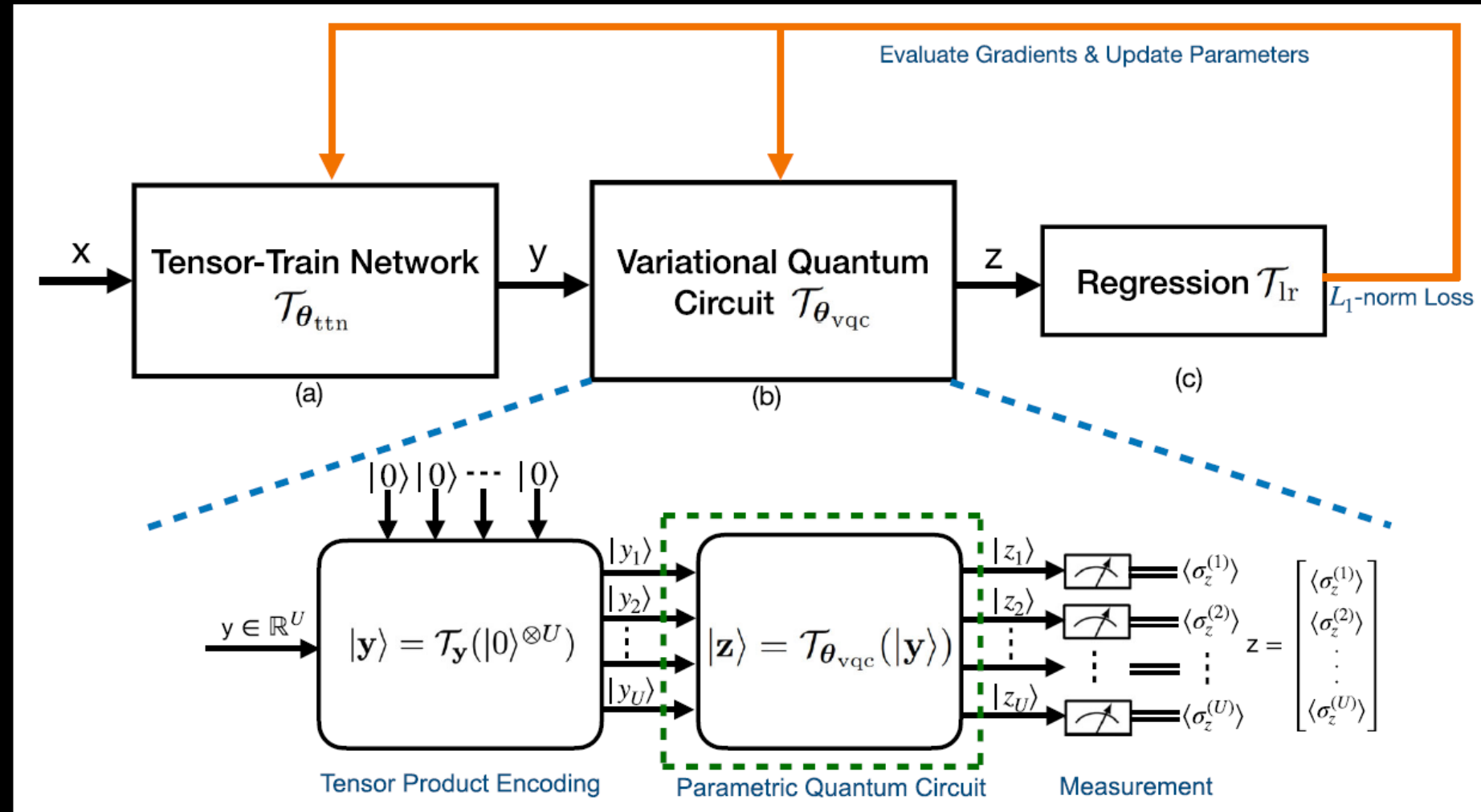
-> Normalized wavefunction of a single qubit

- Goal of TTN: Compute  $y = W \cdot \Phi(x)$  in a quantum-favored manner

# TTN-VQC: Hybrid Quantum Neural Nets

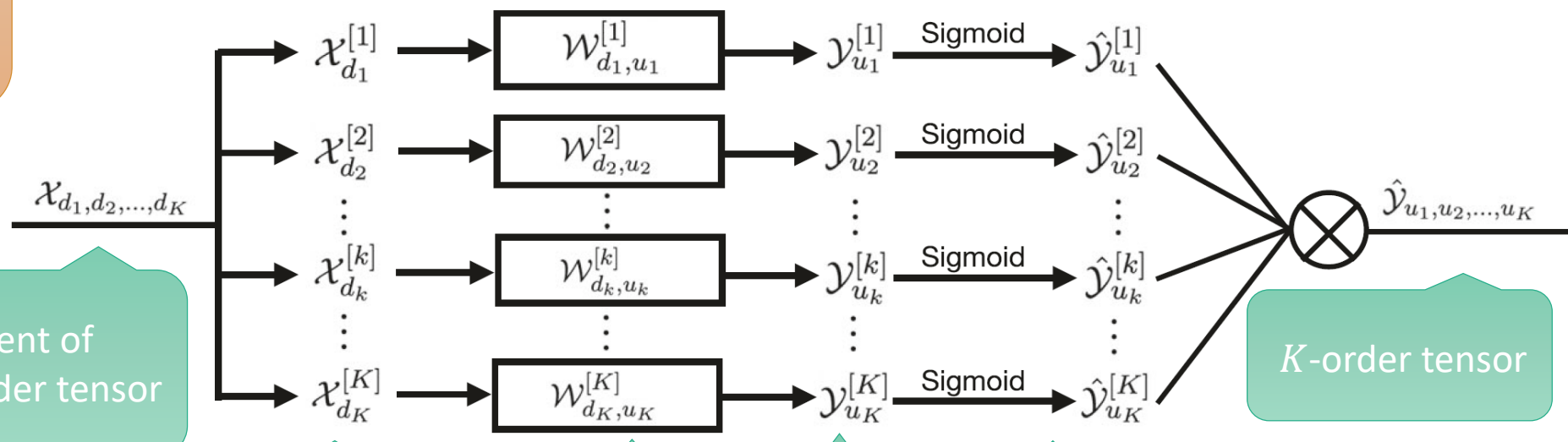
[npj Quantum Information]

- Tensor Product Encoding (TPE): converting a classical data sample  $x$  into a quantum state  $|x\rangle$
- The blocks can be repeated to build deeper networks



# TTN Model in Parallel

Convert a classical data sample of dim  $d$  into a  $K$ -order tensor  $X$



Element of  $K$ -order tensor

Factorization into  $K$  2-dim matrices  $\{X_{d_k}^{[k]}\}$  by Tensor-Train Decomposition

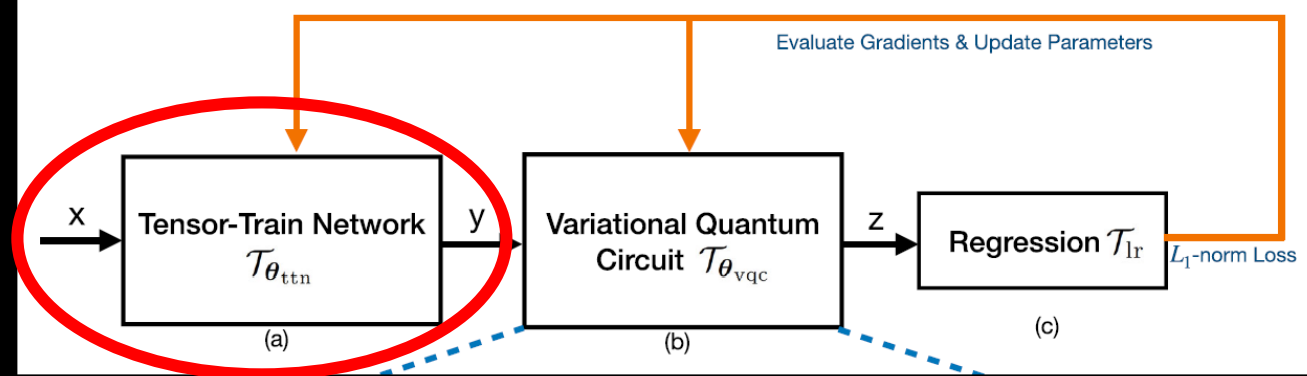
Trainable parameters of TTN

2-dim matrices

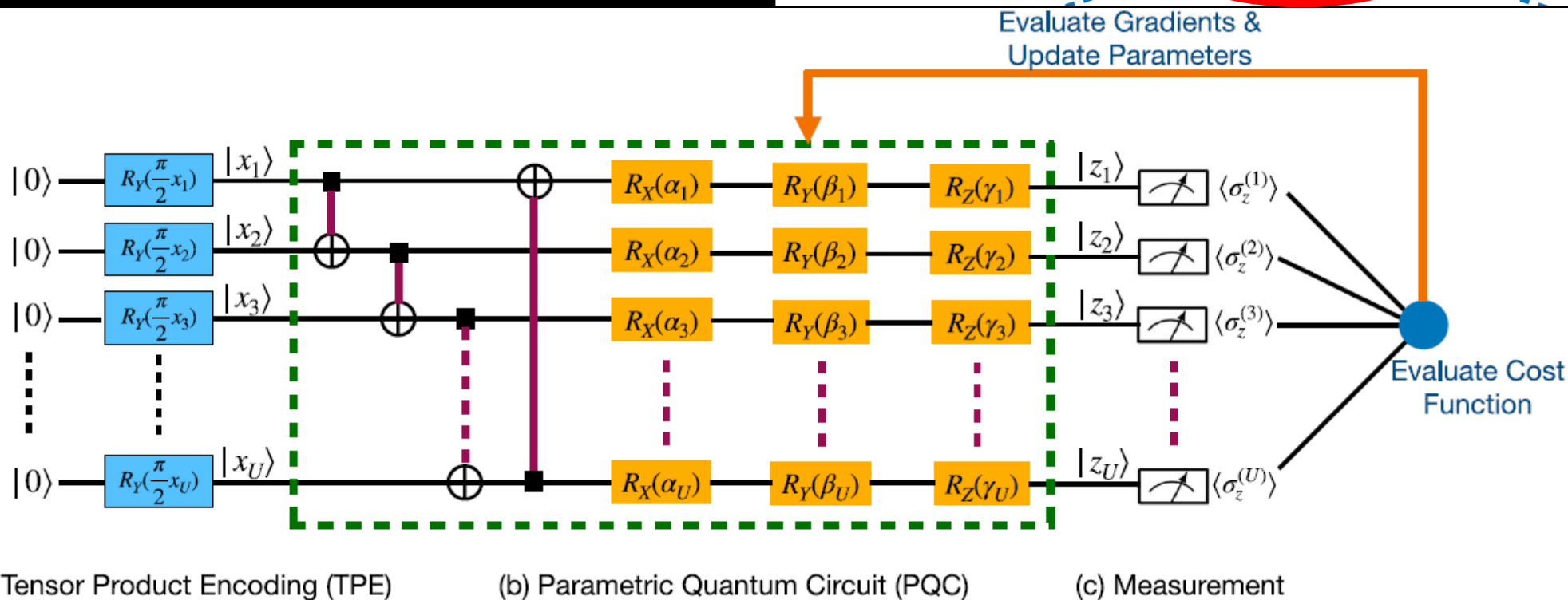
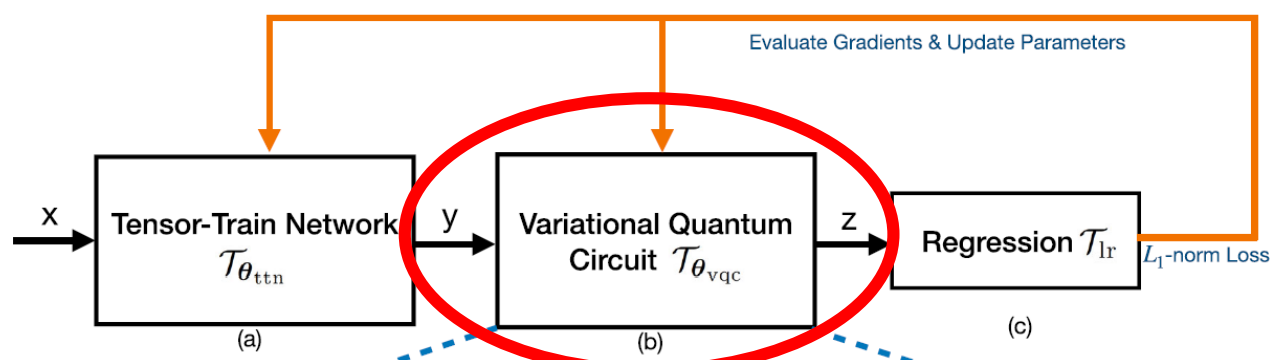
Sigmoid activation function

$$\hat{Y}(j_1, j_2, \dots, j_K) = \prod_{k=1}^K \text{Sigm}(\mathcal{Y}_k(j_k))$$

$$X_{d_1, d_2, \dots, d_K} = X_{d_1}^{[1]} X_{d_2}^{[2]} \dots X_{d_K}^{[K]}$$



# VQC Circuit

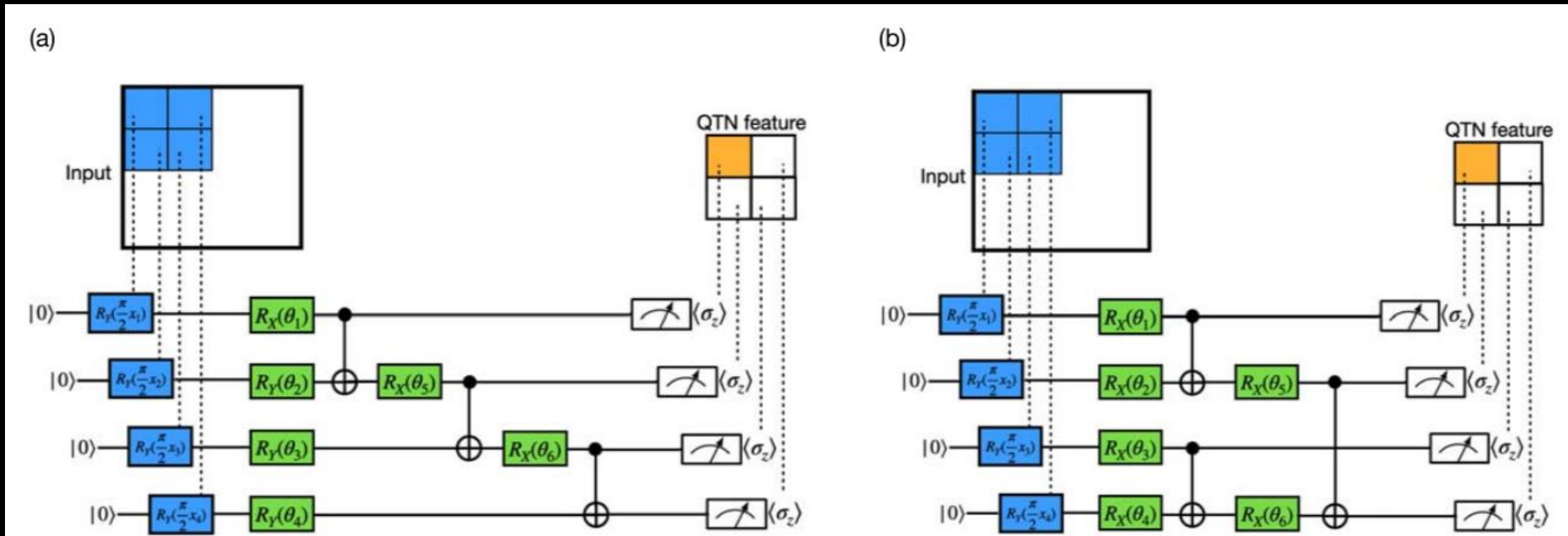


**Fig. 2 An illustration of three components in the VQC model.** The TPE employs a series of  $R_Y(\frac{\pi}{2}x_i)$  to transform classical data into quantum states. The PQC is composed of CNOT gates and single-qubit rotation gates  $R_X, R_Y, R_Z$  with free model parameters  $\alpha, \beta$ , and  $\gamma$ . The CNOT gates impose the operation of quantum entanglement among qubits, and the gates  $R_X, R_Y$ , and  $R_Z$  can be adjustable during the training stage. The PQC model in the green dash square is repeatably copied to build a deeper model. The measurement converts the quantum states  $|z_1\rangle, |z_2\rangle, \dots, |z_U\rangle$  into the corresponding expectation values  $\langle \sigma_z^{(1)} \rangle, \langle \sigma_z^{(2)} \rangle, \dots, \langle \sigma_z^{(U)} \rangle$ . The outputs  $\langle \sigma_z^{(1)} \rangle, \langle \sigma_z^{(2)} \rangle, \dots, \langle \sigma_z^{(U)} \rangle$  are connected to a loss function and the gradient descent algorithms can be used to update VQC parameters.

# QTN-VQC: Fully Quantum Neural Networks

[Physica Scripta]

- Two quantum circuit-implemented QTN architectures:
  1. ConvMPS : convolutional matrix product state
  2. ConvTTN: convolutional tree tensor network



**Figure 3.** An demonstration of two QTN architectures for the end-to-end quantum learning paradigm. (a) A convolutional matrix product state architecture. (b) A convolutional tree tensor network architecture. Similar to the VQC model, a TPE model is used to encode non-trainable  $R_Y(x_i)$  gates with the classical input  $x_i$ , and several trainable  $R_X(\theta_i)$  and  $R_Y(\theta_i)$  gates in the red dash square is associated with parameters  $\theta_i$ . The observables of the last quantum channel correspond to the expected values based on Pauli-Z operators and they are taken as QTN features for further processing. If a  $2 \times 2$  convolutional kernel is used, the inputs corresponding to  $4 \times 4$  matrices are associated with the outputs of the  $4 \times 2 \times 2$  matrices.

# Performance Evaluation

- MNIST: Handwritten digit classification dataset
  - Each sample is a 28 x 28 grayscale image
  - 10000 samples for training and 2000 samples for testing
- Methods
  - PCA-VQC (naïve dimension reduction)
  - TTN-VQC (hybrid; 8 qubits)
  - ConvMPS-VQC (quantum; 4+8 qubits)
  - ConvTTN-VQC (quantum; 4+8 qubits)
- Please refer to papers for detailed implementations

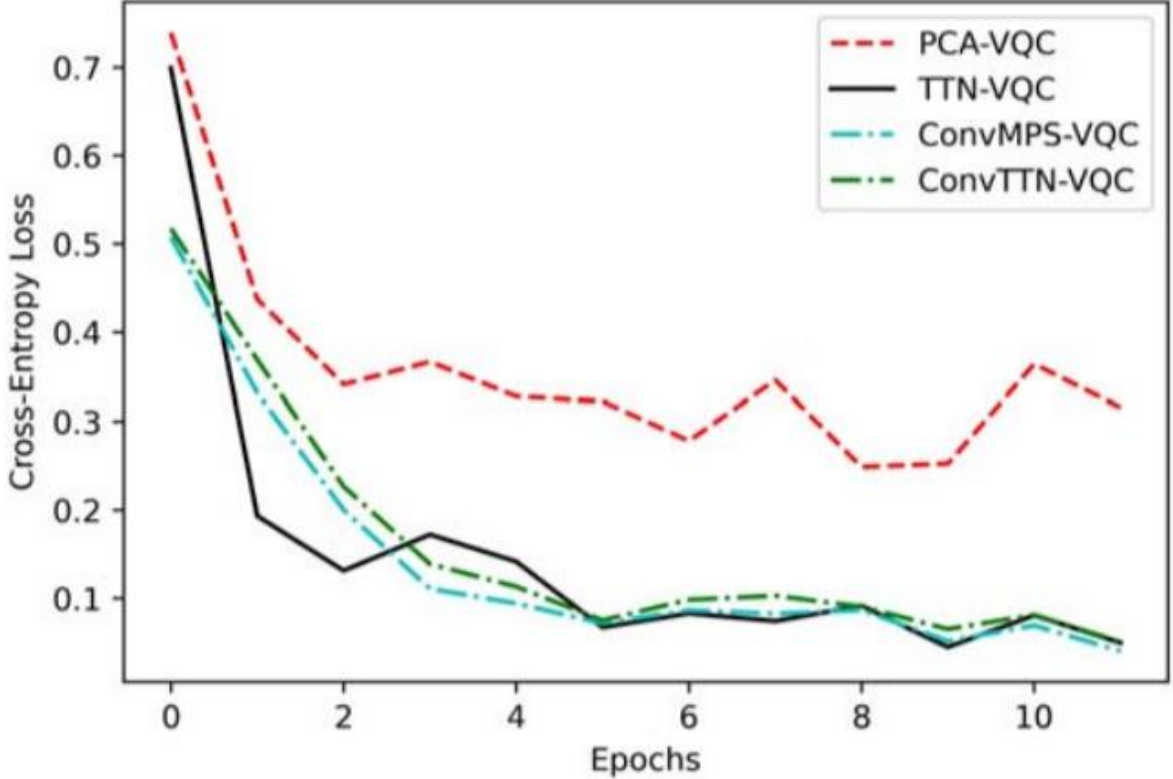




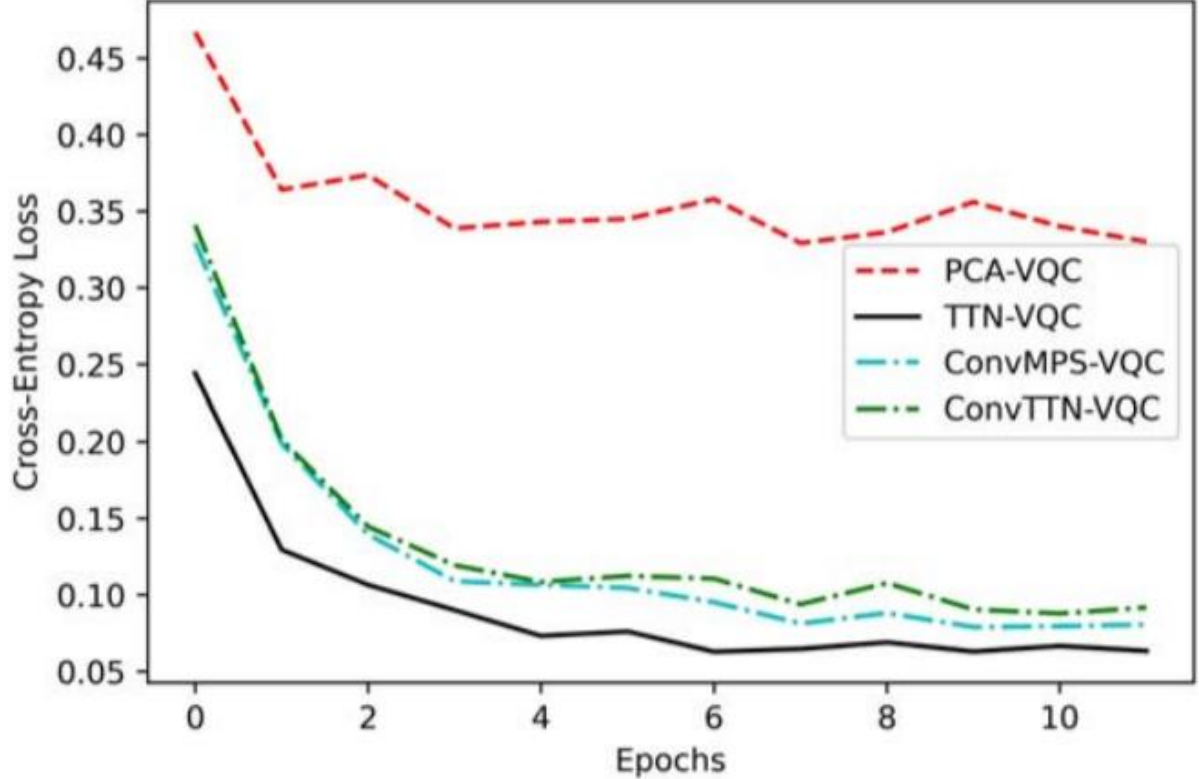
# Binary Classification on MNIST

Models	CE	Accuracy (%)
PCA-VQC	0.3301	87.7
TTN-VQC	0.0634	98.6
ConvTTN-VQC	0.0805	98.5
ConvMPS-VQC	0.0917	98.3

Training loss



Testing loss



# Representation Learning and Generalization of TTN-VQC and QTN-VQC (1)

- **Functional Regression:** Given a  $Q$ -dimensional input vector space  $R^Q$  and a measurable  $U$ -dimensional output vector space  $R^U$ , the TTN-VQC based vector-to-vector regression aims to find a TTN-VQC operator  $f: R^Q \rightarrow R^U$  such that the output vectors of  $f$  can approximate those of a desirable target function  $h^*$ .

- $D$ : data distribution; Draw  $N$  i.i.d. samples from  $D$

- $F_{TV}$ : TTN-VQC hypothesis space of TTN-VQC operators

- Population loss:

$$\mathcal{L}_D(f) := \mathbb{E}_{\mathbf{x} \sim D} [\ell(h_D^*(\mathbf{x}), f(\mathbf{x}))]$$

- Empirical loss:

$$\mathcal{L}_S(f) := \frac{1}{N} \sum_{n=1}^N \ell(h_D^*(\mathbf{x}_n), f(\mathbf{x}_n))$$

- $f_D^*$ : best hypothesis in  $F_{TV}$

# Representation Learning and Generalization of TTN-VQC and QTN-VQC (2)

- $f_S^*$ : best hypothesis based on empirical loss minimization
- $\bar{f}_S$ : returned hypothesis by gradient-based minimization
- Error decomposition:

$$\mathcal{L}_{\mathcal{D}}(\bar{f}_S) = \underbrace{\mathcal{L}_{\mathcal{D}}(f_{\mathcal{D}}^*)}_{\text{Approximation Error}} + \underbrace{\mathcal{L}_{\mathcal{D}}(f_S^*) - \mathcal{L}_{\mathcal{D}}(f_{\mathcal{D}}^*)}_{\text{Estimation Error}} + \underbrace{\mathcal{L}_{\mathcal{D}}(\bar{f}_S) - \mathcal{L}_{\mathcal{D}}(f_S^*)}_{\text{Training Error}}$$

- We derive theoretical upper bounds for each error term under proper conditions (e.g., smooth target function  $h_{\mathcal{D}}^*$ )

# Representation Learning and Generalization of TTN-VQC and QTN-VQC (3)

- **Representation Power**: upper bound on approximation error
- **Generalization power**: upper bound on the estimation error concerning the empirical Rademacher complexity
- **Optimization bias**: training error can be exponentially converged to a small loss value
- Similar analysis holds for QTN-VQC

$$\mathcal{L}_{\mathcal{D}}(\bar{f}_S) = \underbrace{\mathcal{L}_{\mathcal{D}}(f_{\mathcal{D}}^*)}_{\text{Approximation Error}} + \underbrace{\mathcal{L}_{\mathcal{D}}(f_S^*) - \mathcal{L}_{\mathcal{D}}(f_{\mathcal{D}}^*)}_{\text{Estimation Error}} + \underbrace{\mathcal{L}_{\mathcal{D}}(\bar{f}_S) - \mathcal{L}_{\mathcal{D}}(f_S^*)}_{\text{Training Error}}$$

Category	This work
Learning problem	Regression
Dimensionality reduction	TTN <span style="color: green;"># measurements</span>
Representation power	$\frac{\Theta(1)}{\sqrt{U}} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$ <span style="color: green;">↖ ↗</span>
Generalization power	<span style="color: green;"># qubits</span> $\frac{2P}{\sqrt{N}} \left( \sqrt{\sum_{k=1}^K \Lambda_k^2} + \Lambda' \right)$
Conditions for optimization bias	<span style="color: green;"># training samples</span> $\mu\text{-PL} + 1\text{-Lipschitz}$

# Quantum Foundation Models?

Foundation Models

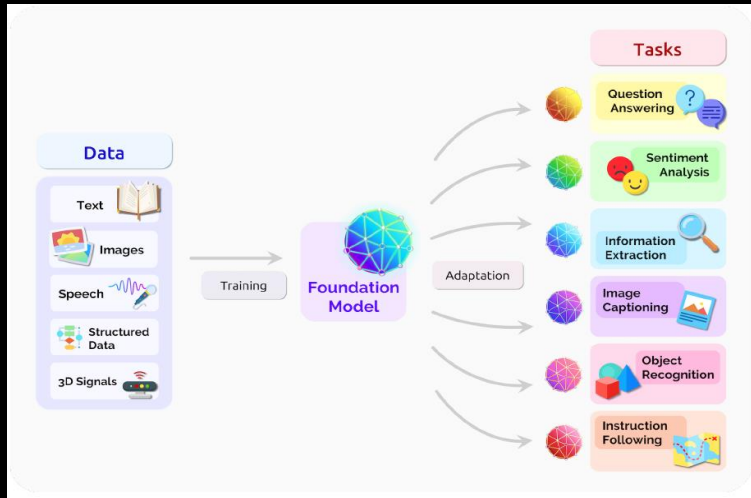
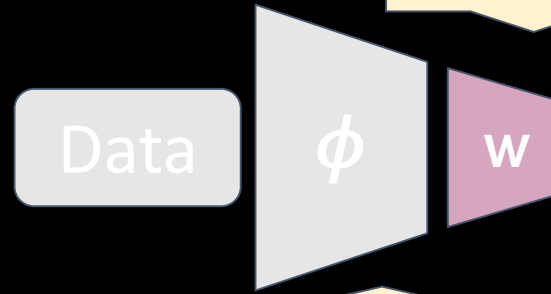


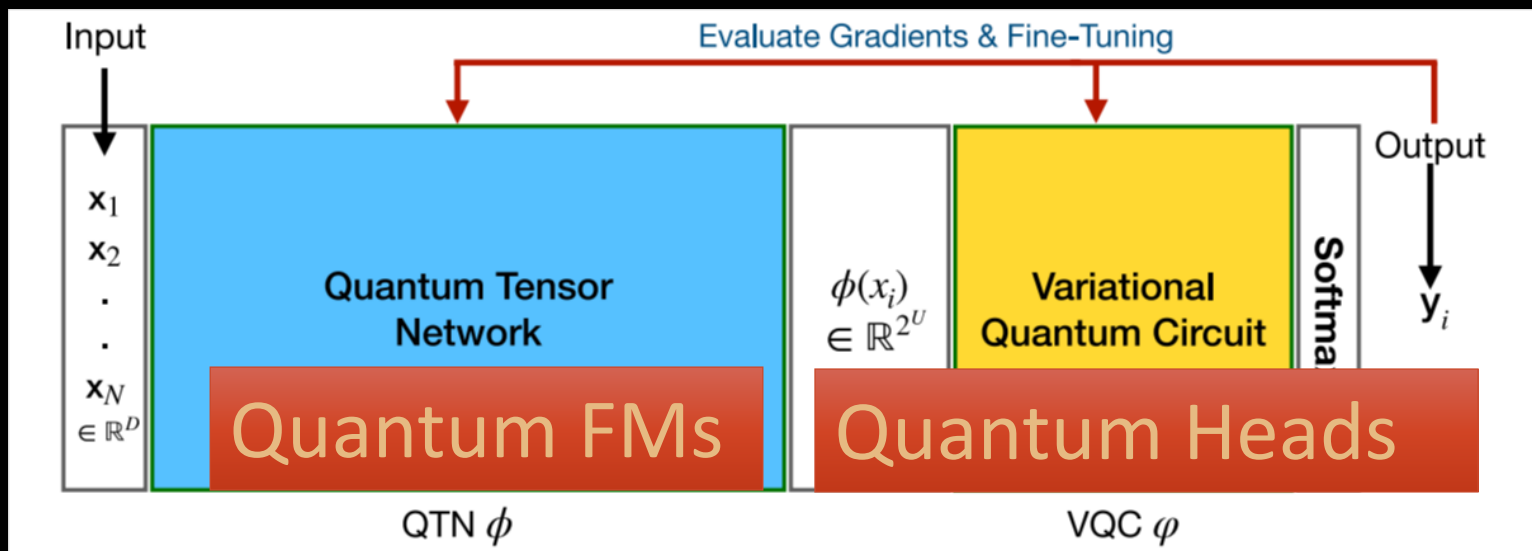
Fig. 2. A foundation model can centralize the information from all the data from various modalities. This one model can then be adapted to a wide range of downstream tasks.

Linear head / PEFT for downstream task



Encoder/Representation network from pretraining

QTN-VQC



# Exploring “Quantum Foundation Models” and “Quantum Transfer Learning” with TTN-VQC

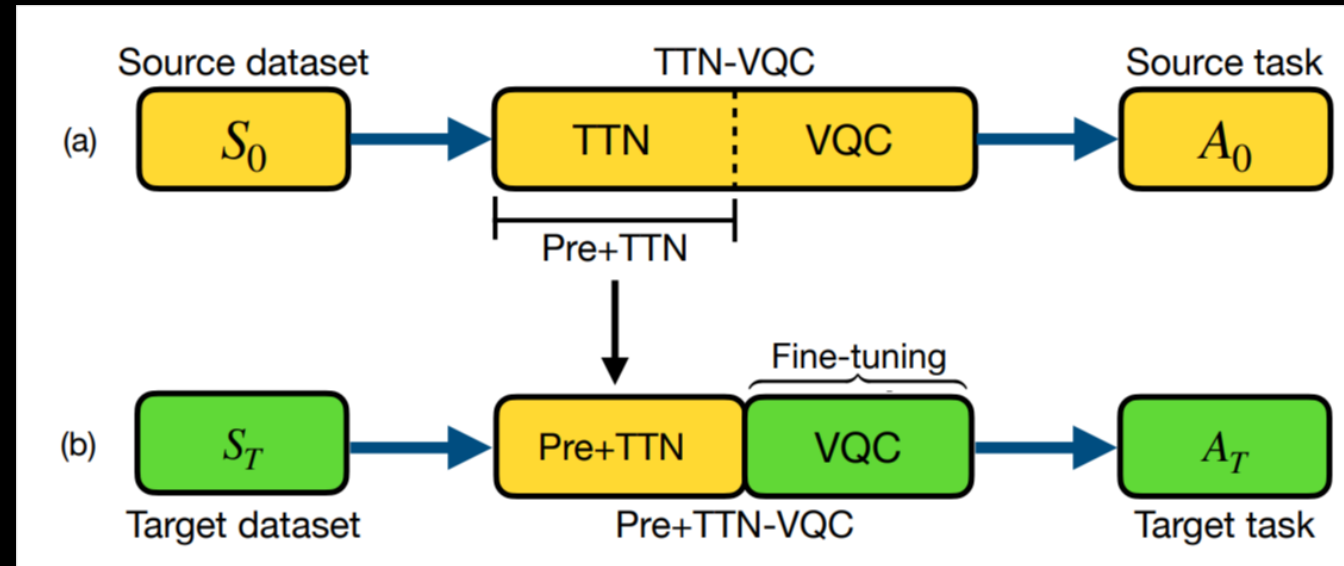
## Pre-training Tensor-Train Networks Facilitates Machine Learning with Variational Quantum Circuits

**Jun Qi**  
Georgia Institute of Technology  
jq41@gatech.edu

**Chao-Han Huck Yang**  
Georgia Institute of Technology  
huckiyang@gatech.edu

**Pin-Yu Chen**  
IBM Research  
pin-yu.chen@ibm.com

**Min-Hsiu Hsieh**  
Hon Hai Quantum Computing Research Center  
minhsiuh@gmail.coms



# Key Results

Category	Pre+TTN-VQC	TTN-VQC
Representation error <small># source samples</small>	$\frac{2P}{\sqrt{ S_0 }} \sqrt{\sum_{k=1}^K \Lambda_k^2} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$	$\frac{\Theta(1)}{\sqrt{U}} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$
Estimation error <small># target samples</small>	$\frac{2P\Lambda}{\sqrt{ S_T }}$	$\frac{2P}{\sqrt{N}} \left( \sqrt{\sum_{k=1}^K \Lambda_k^2} + \Lambda \right)$
Conditions for Optimization bias	without PL assumption	PL assumption
Optimization bias	$\beta R^2 + R \sqrt{\frac{L^2 + \beta^2 R^2}{T_{\text{sgd}}}}$	sufficient small

Binary MNIST  
classification

Models	Params (Mb)	CE	Accuracy (%)
PCA-VQC_8Qubit	0.080	0.5744	75.8
TTN-VQC_8Qubit	0.452	0.0625	98.7
Pre+TTN-VQC_8Qubit	0.452	0.0390	99.0

# co-TenQu: A quantum-classical collaborative training architecture [IEEE TQE 2024]

Quantum Software

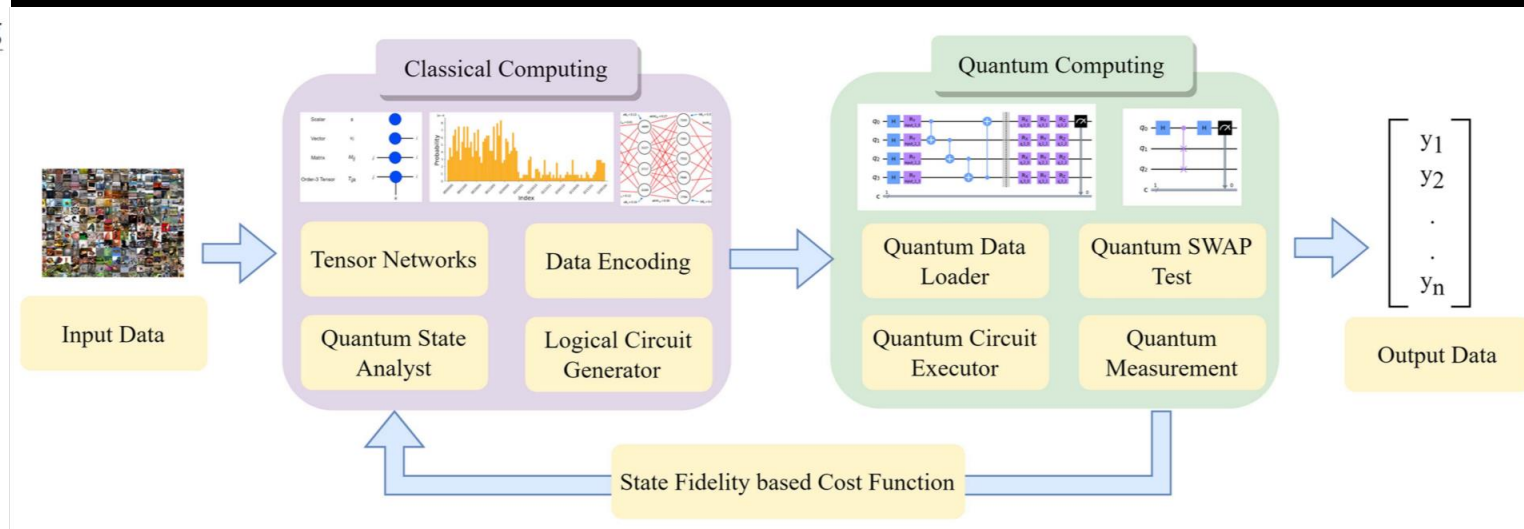
IEEE Transactions on Quantum Engineering

Received 9 November 2023; revised 18 January 2024; accepted 12 February 2024; date of publication 19 February 2024; date of current version 3 April 2024.  
Digital Object Identifier 10.1109/TQE.2024.3367234

## A Quantum-Classical Collaborative Training Architecture Based on Quantum State Fidelity

RYAN L'ABBATE<sup>1</sup>, ANTHONY D'ONOFRIO JR.<sup>1</sup>, SAMUEL STEIN<sup>2</sup>,  
SAMUEL YEN-CHI CHEN<sup>3</sup>, ANG LI<sup>2</sup>,  
PIN-YU CHEN<sup>4</sup> (Senior Member, IEEE), JUNTAO CHEN<sup>1</sup> (Member, IEEE),  
AND YING MAO<sup>1</sup>

<sup>1</sup>Computer and Information Science Department, Fordham University, Bronx, NY 10458 USA  
<sup>2</sup>Pacific Northwest National Laboratory, Richland, WA 99354 USA  
<sup>3</sup>Brookhaven National Laboratory, Upton, NY 11973 USA  
<sup>4</sup>IBM Research, Yorktown Heights, NY 10598 USA



- co-TenQu trains the model directly on quantum states aiming at accelerating the training process and improving performance
- co-TenQu with Qiskit and PennyLane:
  - Enhances a classical deep neural network by up to 41.72% in a fair setting.
  - Outperforms other quantum-based methods by up to 1.9X and achieves similar accuracy while utilizing 70.59% fewer qubits.



# Data Privacy and Robustness

# Data Privacy Breach in AI/ML

## Extracting Training Data from Large Language Models

USENIX 2021

Nicholas Carlini<sup>1</sup> Florian Tramèr<sup>2</sup> Eric Wallace<sup>3</sup> Matthew Jagielski<sup>4</sup>  
 Ariel Herbert-Voss<sup>5,6</sup> Katherine Lee<sup>1</sup> Adam Roberts<sup>1</sup> Tom Brown<sup>5</sup>  
 Dawn Song<sup>3</sup> Úlfar Erlingsson<sup>7</sup> Alina Oprea<sup>4</sup> Colin Raffel<sup>1</sup>

<sup>1</sup>Google <sup>2</sup>Stanford <sup>3</sup>UC Berkeley <sup>4</sup>Northeastern University <sup>5</sup>OpenAI <sup>6</sup>Harvard <sup>7</sup>Apple

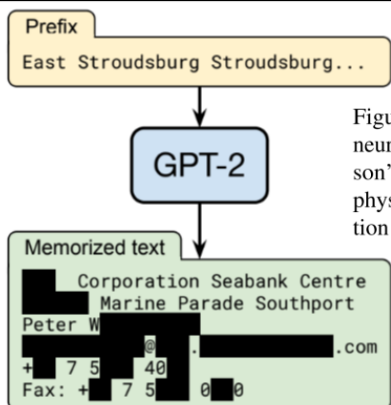


Figure 1: Our extraction attack. Given query access to a neural network language model, we extract an individual person's name, email address, phone number, fax number, and physical address. The example in this figure shows information that is all accurate so we redact it to protect privacy.

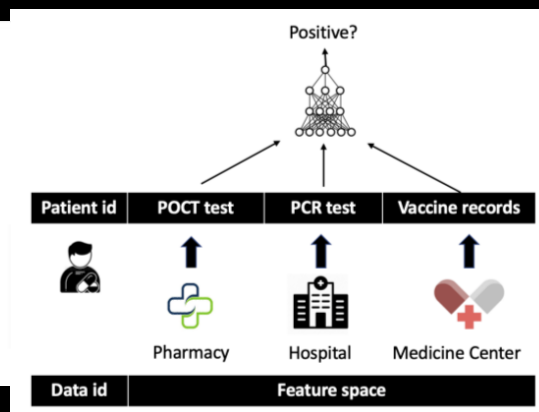
## CAFE: Catastrophic Data Leakage in Vertical Federated Learning

NeurIPS 2021

Xiao Jin Rensselaer Polytechnic Institute jinx2@rpi.edu  
 Pin-Yu Chen IBM Research pin-yu.chen@ibm.com  
 Chia-Yi Hsu National Yang Ming Chiao Tung University chiayihsu8315@gmail.com

Chia-Mu Yu National Yang Ming Chiao Tung University chiamuyu@gmail.com

Tianyi Chen Rensselaer Polytechnic Institute chent18@rpi.edu



- Clients submit gradients of private data for federated learning
- An honest but curious server (aggregator) in vertical federated learning can recover private client data without violating the protocol (without privacy protection mechanisms)



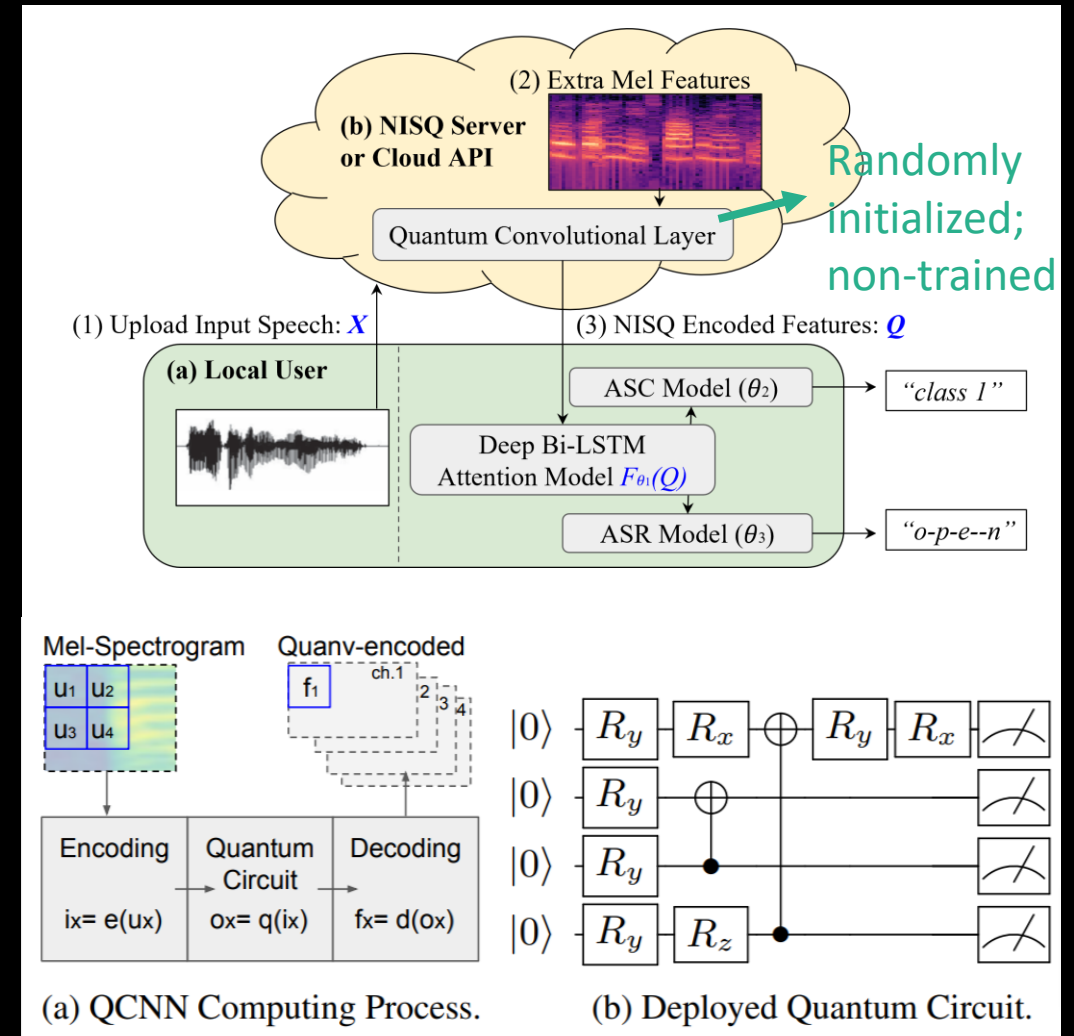
Real data image 6-25



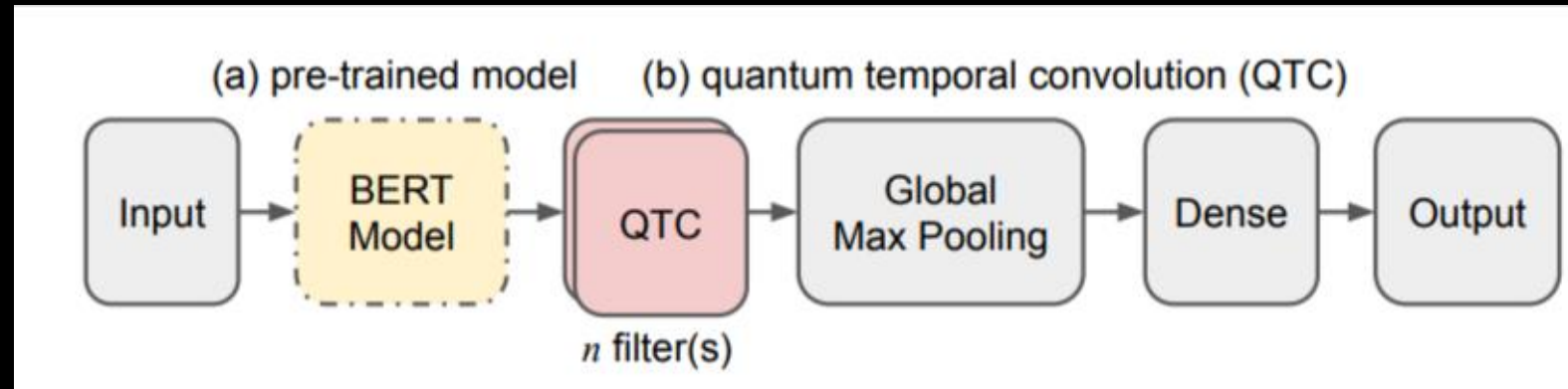
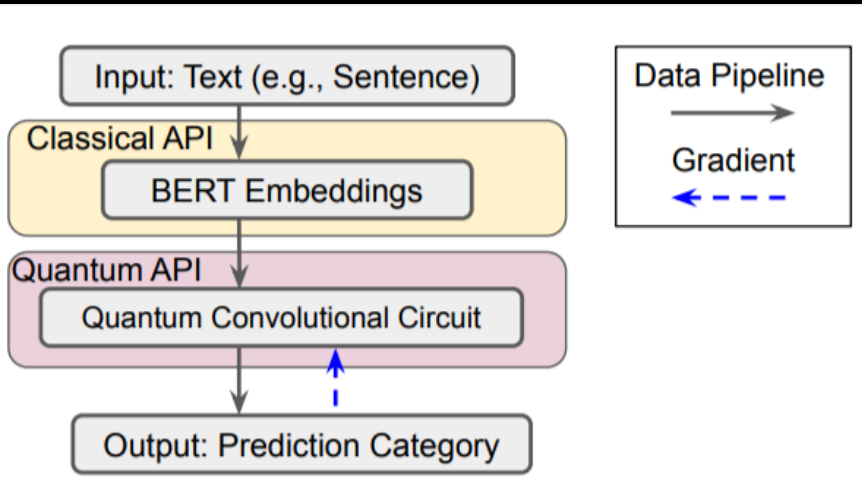
Recovered data image 6-25

# Quantum Layers for Data Protection

- Vertical federated learning
- Speech input is first processed into Mel spectrogram and then sent into a quantum layer for encoding (on the cloud)
- The encoded features are used to train the acoustic model (on user devices)
- Retain competitive accuracy and reduce model parameter leakage



# BERT (transformer models) with Quantum Temporal Convolution Layers



Embedding	word2vec				BERT				
	(n,k)	(1,4)	(2,2)	(2,3)	(2,4)	(1,4)	(2,2)	(2,3)	(2,4)
TCN		80.09	80.22	80.91	82.34	95.18	95.03	94.95	95.23
QTC		<b>81.42</b>	<b>82.49</b>	<b>83.82</b>	<b>83.95</b>	<b>96.69</b>	<b>96.92</b>	<b>96.32</b>	<b>96.98</b>

- $n$ : # of filters
- $k$ : kernel size
- TCN: random temporal convolutional network

# Robustness Challenges in AI/ML

$$(x, y) \sim D$$

$x$ : data sample

$y$ : groundtruth

$D$ : in-domain data distribution  
(in-distribution)

$\theta$ : model parameters

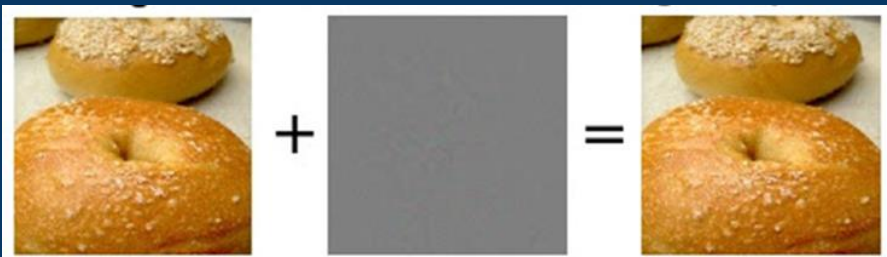
$f_{\theta}(x)$ : model prediction on  $x$

## Robustness Categories

- Adversarial robustness (worst-case performance)
- Out-of-distribution (OOD) generalization (domain shifts)
- Out-of-distribution detection (unknowns)

(I) Adversarial robustness:  
 $x'$  similar to  $x$ . Ideally,  $f_{\theta}(x') = f_{\theta}(x)$

$f_{\theta}(x) = \text{bagel}$      $\delta$ : adversarial perturbation     $f_{\theta}(x+\delta) = \text{piano}$



(II) OOD generalization:  
 $x' \sim D'$ , a shifted version of  $D$ . Ideally,  $f_{\theta}(x') = f_{\theta}(x)$



(III) OOD detection:  
 $x' \sim D'$ , a dissimilar or new domain w.r.t.  $D$ . Ideally,  $f_{\theta}(x') = \text{"Unknown"}$

$f_{\theta}(\cdot) \in \{\text{cat}, \text{dog}\}$

$f_{\theta}(\text{bagel}) = \text{"Unknown/OOD"}$

# Randomized Smoothing for Certified Robustness

- Rough idea: “Majority vote” with Gaussian perturbed input for classification gives guaranteed robustness
- $f$ : a base (deterministic) classifier

$$g(x) = \arg \max_{c \in \mathcal{Y}} \mathbb{P}(f(x + \varepsilon) = c) \quad (1)$$

where  $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$

**Theorem 1.** Let  $f : \mathbb{R}^d \rightarrow \mathcal{Y}$  be any deterministic or random function, and let  $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$ . Let  $g$  be defined as in (1). Suppose  $c_A \in \mathcal{Y}$  and  $\underline{p}_A, \overline{p}_B \in [0, 1]$  satisfy:

$$\mathbb{P}(f(x + \varepsilon) = c_A) \geq \underline{p}_A \geq \overline{p}_B \geq \max_{c \neq c_A} \mathbb{P}(f(x + \varepsilon) = c) \quad (2)$$

Then  $g(x + \delta) = c_A$  for all  $\|\delta\|_2 < R$ , where

$$R = \frac{\sigma}{2} (\Phi^{-1}(\underline{p}_A) - \Phi^{-1}(\overline{p}_B)) \quad (3)$$

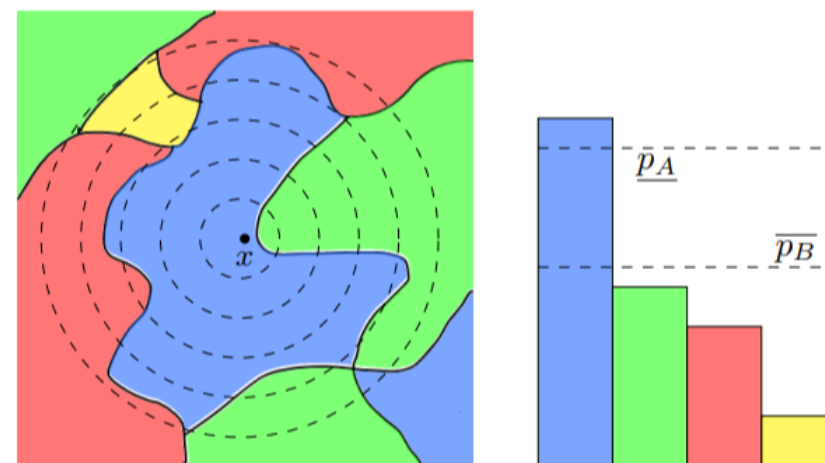


Figure 1. Evaluating the smoothed classifier at an input  $x$ . **Left:** the decision regions of the base classifier  $f$  are drawn in different colors. The dotted lines are the level sets of the distribution  $\mathcal{N}(x, \sigma^2 I)$ . **Right:** the distribution  $f(\mathcal{N}(x, \sigma^2 I))$ . As discussed below,  $\underline{p}_A$  is a lower bound on the probability of the top class and  $\overline{p}_B$  is an upper bound on the probability of each other class. Here,  $g(x)$  is “blue.”

# Randomized Smoothing for Quantum Circuits

- Theoretical proof that the addition of quantum random rotation noise can improve the robustness of quantum classifiers against adversarial perturbations
- Noise addition -> Quantum Differential Privacy -> Certified Robustness

---

**Algorithm 1** Quantum model under quantum noise rotation

---

**Input**  $\sigma$ : where  $\sigma$  is density matrix of n-dim data.

**Output**  $f(\theta^*, \sigma)$

1. For a chosen quantum classifier, add Pauli-X operators before each input qubit.
  2. Generate n random variables  $\theta_1, \theta_2, \dots, \theta_n$  subject to  $0 < h_1 < \tan \theta_i < h_2$  for all  $i \in \{1, 2, \dots, n\}$ .
  3. Set up rotation angles of additional Pauli-X operators with  $\theta_1, \theta_2, \dots, \theta_n$ .
  4. Execute the quantum classifier  $N$  times to get the score vector  $f(\theta^*, \sigma)$ .
- 

## Fruits for thoughts:

- Can Quantum-inspired noise processing improve certified robustness of classical ML classifiers?
- Large Language Models and GenAI technology are stochastic by design -> Quantum modeling?

# Concluding Remarks

- AI for Quantum: Tensor-Train Networks as powerful representation learning tools for quantum neural networks
- Quantum for AI: Quantum circuits and noises as privacy-enhanced and robustness-improved modules for classical and quantum ML
- Post Quantum AI: What can we do with Quantum Foundation Models and Quantum GenAI? (no, they are not buzz words)
  - Quantum-inspired trust and safety for AI
  - AI-inspired risks for Quantum AI



 **You**

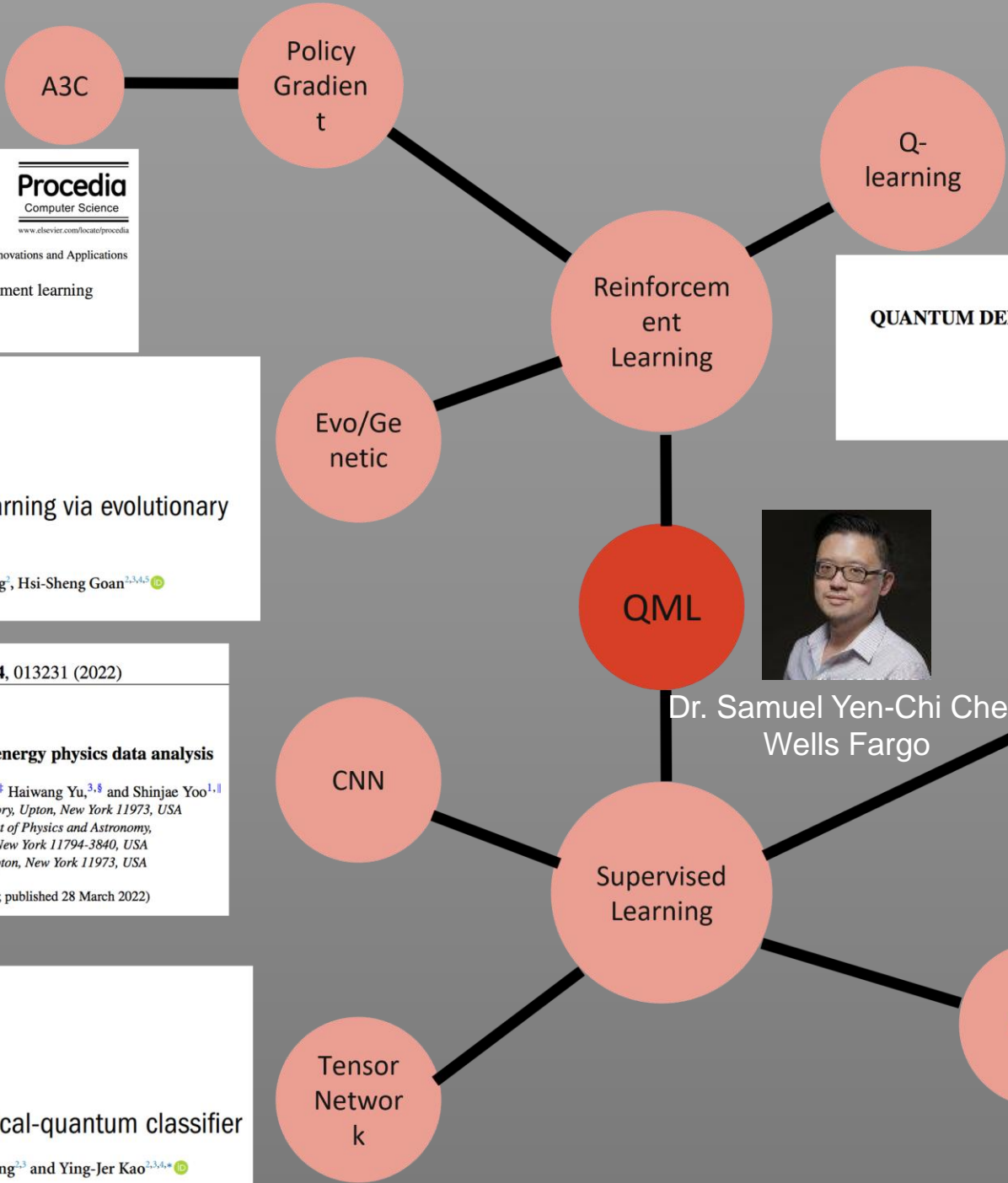
what can quantum computing help with generative AI?

 **Copilot**

Quantum computing has the potential to significantly impact generative AI in several ways:

1. Speeding up computations: Quantum computers can perform certain calculations much faster than classical computers, which could accelerate the training of generative models.
2. Handling complex problems: Quantum computing can process and analyze large datasets more efficiently, which is beneficial for training generative models on complex tasks.
3. Improving optimization: Quantum algorithms can potentially find better solutions to optimization problems, which are central to training generative AI models.
4. Enhancing creativity: With its ability to explore a vast search space, quantum computing could help generative AI produce more diverse and innovative outputs.
5. Energy efficiency: Quantum computers could use less energy for certain computations, making the process of training generative AI models more sustainable.

While quantum computing is still in its early stages, its integration with generative AI could lead to breakthroughs in the field, enabling the creation of more advanced and efficient AI systems.



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)  
**ScienceDirect**  
 Procedia Computer Science 222 (2023) 321–330  
**Procedia**  
 Computer Science  
[www.elsevier.com/locate/procedia](http://www.elsevier.com/locate/procedia)

International Neural Network Society Workshop on Deep Learning Innovations and Applications (INNS DLIA 2023)  
**Asynchronous training of quantum reinforcement learning**  
 Samuel Yen-Chi Chen<sup>\*,\*</sup>  
<sup>\*</sup>Wells Fargo, New York, NY 10017, USA

Received June 17, 2020, accepted July 6, 2020, date of publication July 20, 2020, date of current version August 12, 2020.  
 Digital Object Identifier 10.1109/ACCESS.2020.3100470

**Variational Quantum Circuits for Deep Reinforcement Learning**

SAMUEL YEN-CHI CHEN<sup>\*,1,2</sup>, CHAO-HAN HUCK YANG<sup>3</sup>, JUN QI<sup>\*,3</sup>, (Member, IEEE), PIN-YU CHEN<sup>4</sup>, (Member, IEEE), XIAOLI MA<sup>3</sup>, (Fellow, IEEE), AND HSI-SHENG GOAN<sup>\*,1,2,5</sup>

**QUANTUM DEEP RECURRENT REINFORCEMENT LEARNING**

Samuel Yen-Chi Chen  
 Wells Fargo

**MACHINE LEARNING**  
 Science and Technology

**PAPER**

**Variational quantum reinforcement learning via evolutionary optimization**

Samuel Yen-Chi Chen<sup>1,\*</sup>, Chih-Min Huang<sup>2</sup>, Chia-Wei Hsing<sup>2</sup>, Hsi-Sheng Goan<sup>2,3,4,5</sup> and Ying-Jer Kao<sup>2,3,4</sup>

**entropy**

Article

**Federated Quantum Machine Learning**

Samuel Yen-Chi Chen<sup>\*,\*</sup> and Shinjae Yoo

PHYSICAL REVIEW RESEARCH 4, 013231 (2022)

**Quantum convolutional neural networks for high energy physics data analysis**

Samuel Yen-Chi Chen<sup>\*,1,\*</sup>, Tzu-Chieh Wei<sup>\*,2,†</sup>, Chao Zhang<sup>\*,3,‡</sup>, Haiwang Yu<sup>\*,3,§</sup> and Shinjae Yoo<sup>\*,1,¶</sup>  
<sup>1</sup>Computational Science Initiative, Brookhaven National Laboratory, Upton, New York 11973, USA  
<sup>2</sup>C. N. Yang Institute for Theoretical Physics and Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York 11794-3840, USA  
<sup>3</sup>Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA

(Received 1 February 2021; accepted 29 January 2022; published 28 March 2022)



Dr. Samuel Yen-Chi Chen  
 Wells Fargo

BNL

**Quantum Long Short-Term Memory**

Samuel Yen-Chi Chen,<sup>\*</sup> Shinjae Yoo,<sup>†</sup> and Yao-Lung L. Fang<sup>‡</sup>  
 Computational Science Initiative, Brookhaven National Laboratory  
 (Dated: September 4, 2020)

**MACHINE LEARNING**  
 Science and Technology

**PAPER**

**An end-to-end trainable hybrid classical-quantum classifier**

Samuel Yen-Chi Chen<sup>1,\*</sup>, Chih-Min Huang<sup>2,3</sup>, Chia-Wei Hsing<sup>2,3</sup> and Ying-Jer Kao<sup>2,3,4,\*</sup>

**Reservoir Computing via Quantum Recurrent Neural Networks**

Samuel Yen-Chi Chen,<sup>1</sup> Daniel Fry,<sup>2</sup> Amol Deshmukh,<sup>2</sup> Vladimir Rastunkov,<sup>2</sup> and Charlee Stefanski<sup>1</sup>  
<sup>1</sup>Wells Fargo  
<sup>2</sup>IBM Quantum, IBM Research  
 (Dated: November 7, 2022)