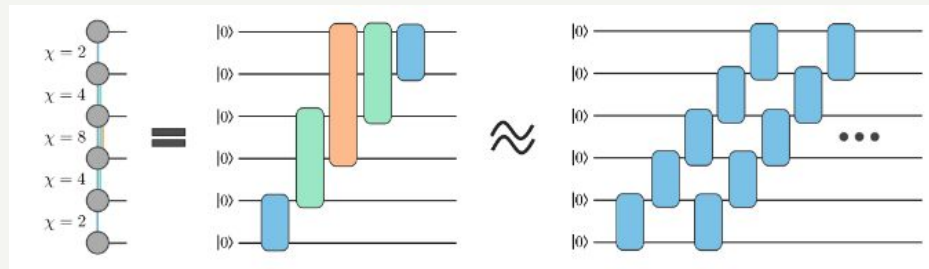
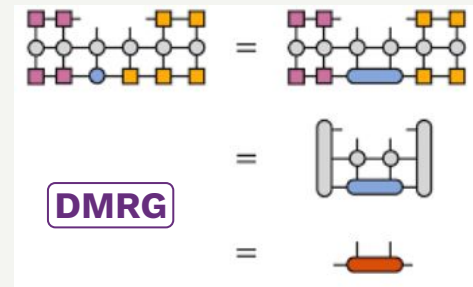
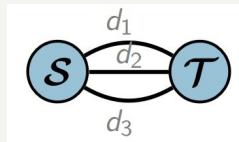
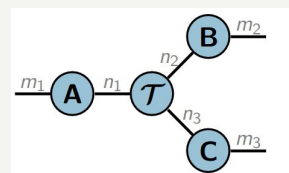
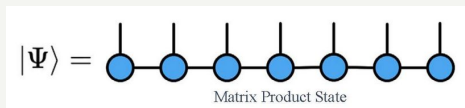
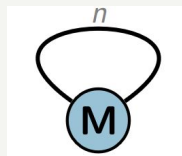
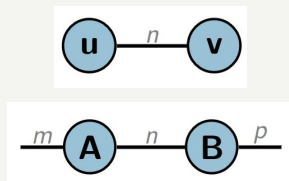


Tensor Networks, Machine Learning & Quantum

Guillaume Rabusseau, DIRO

What are Tensor Networks?

- Efficient parameterization of quantum states with low entanglement
- Set of tools to decompose tensors in small interacting blocks (and manipulate them)
- A useful framework to analyze and simulate quantum circuits



What are Tensor Networks?

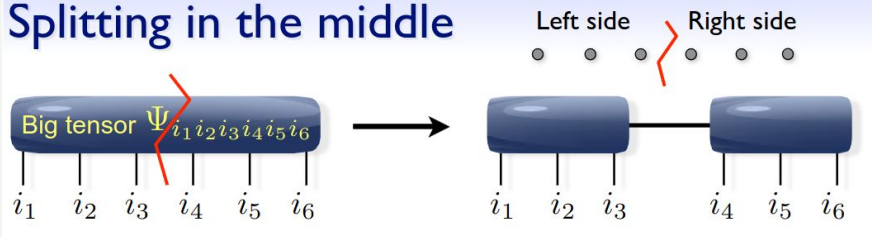
Matrix Product States / Tensor Train



d^n parameters vs. $O(dnR)$ parameters

OPERATION	OUTPUT RANKS	COMPLEXITY
$\mathbf{C} = \mathbf{A} \cdot \text{const}$	$r(\mathbf{C}) = r(\mathbf{A})$	$O(d r(\mathbf{A}))$
$\mathbf{C} = \mathbf{A} + \text{const}$	$r(\mathbf{C}) = r(\mathbf{A}) + 1$	$O(nd r^2(\mathbf{A}))$
$\mathbf{C} = \mathbf{A} + \mathbf{B}$	$r(\mathbf{C}) \leq r(\mathbf{A}) + r(\mathbf{B})$	$O(nd r^2(\mathbf{C}))$
$\mathbf{C} = \mathbf{A} \odot \mathbf{B}$	$r(\mathbf{C}) \leq r(\mathbf{A}) r(\mathbf{B})$	$O(nd r^2(\mathbf{A}) r^2(\mathbf{B}))$
$\mathbf{c} = M\mathbf{b}$	$r(\mathbf{c}) \leq r(M) r(\mathbf{b})$	$O(nd^2 r^2(M) r^2(\mathbf{b}))$
$\text{sum } \mathbf{A}$	–	$O(nd r^2(\mathbf{A}))$
$\ \mathbf{A}\ _F$	–	$O(nd r^3(\mathbf{A}))$
$\mathbf{C} = \text{round}(\mathbf{A}, \varepsilon)$	$r(\mathbf{C}) \leq r(\mathbf{A})$	$O(nd r^3(\mathbf{A}))$

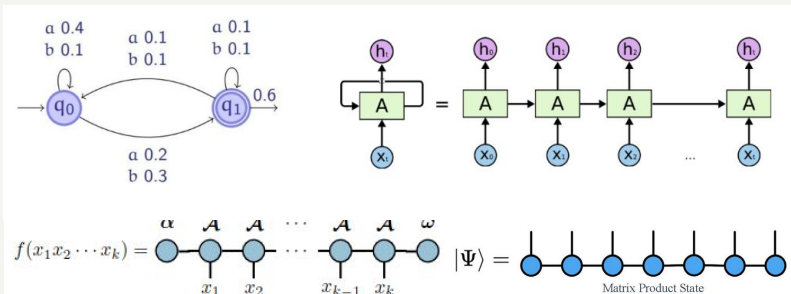
Splitting in the middle



Previous work Building Bridges

Connecting Weighted Automata, Tensor Networks and Recurrent Neural Networks through Spectral Learning

Tianyu Li · Doina Precup · Guillaume Rabusseau



Quantum Tensor Networks, Stochastic Processes, and Weighted Automata

Siddharth Srinivasan*

School of Computer Science and Engineering
University of Washington
sidsrini@cs.washington.edu

Sandesh Adhikary*

School of Computer Science and Engineering
University of Washington
adhikary@cs.washington.edu

Jacob Miller

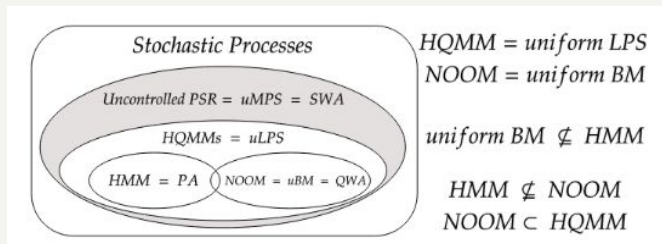
Mila & DIRO
Université de Montréal
jmjacobmiller@gmail.com

Guillaume Rabusseau

CCAI chair - Mila & DIRO
Université de Montréal
grabus@iro.umontreal.ca

Byron Boots

School of Computer Science and Engineering
University of Washington
bboots@cs.washington.edu



Classical vs. Quantum

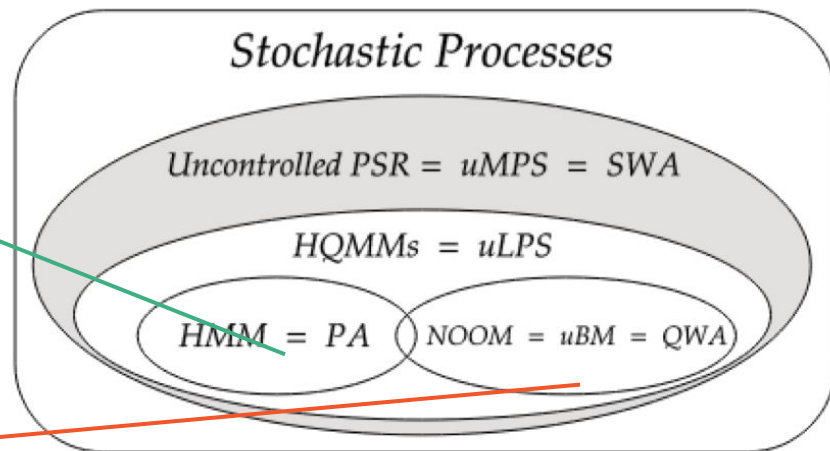
Only non-negative elements

$$\mathbb{P}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) =$$

Hidden Markov models

Born machines

$$\mathbb{P}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \left(\begin{array}{c} \alpha \text{---} R \text{---} A \text{---} R \text{---} A \text{---} R \text{---} A \text{---} R \text{---} A \text{---} R \text{---} \omega \\ | \quad | \quad | \quad | \\ x_1 \quad x_2 \quad x_3 \quad x_4 \end{array} \right)^2$$



- Coyle, Brian, et al. *npj Quantum Information* 6.1 (2020): 60.
- Adhikary, Sandesh, et al. *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021.

Previous work

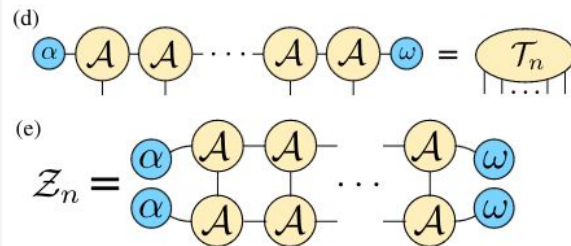
Quantum Inspired / Tensor Network ML models

Tensor Networks for Probabilistic Sequence Modeling

Jacob Miller
Mila and DIRO
Université de Montréal
jmjacobmiller@gmail.com

Guillaume Rabusseau
CCAI chair - Mila and DIRO
Université de Montréal
grabus@iro.umontreal.ca

John Terilla
CUNY and Tunnel
City University of New York
jterilla@gc.cuny.edu



Generative Learning of Continuous Data by Tensor Networks

Alex Meiburg,^{1,2,3} Jing Chen,¹ Jacob Miller,^{1,*} Raphaëlle Tihon,⁴ Guillaume Rabusseau,^{4,5} and Alejandro Perdomo-Ortiz⁶

¹Zapata AI, Boston, USA

²Perimeter Institute for Theoretical Physics, Waterloo, Canada

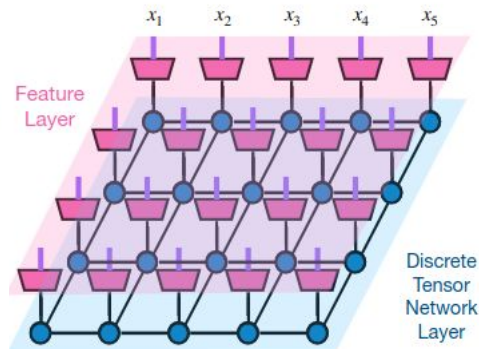
³Institute for Quantum Computing, University of Waterloo

⁴Mila and DIRO, Université de Montréal, Montréal, Canada

⁵CIFAR AI Chair

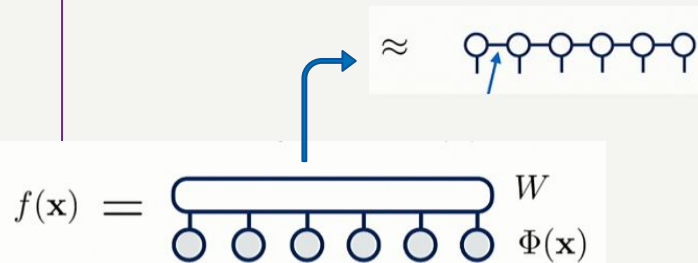
⁶Zapata AI, Toronto, Canada

(Dated: November 1, 2023)



Previous work

Learning Theory & Tensor Networks



Lower and Upper Bounds on the VC-Dimension of Tensor Network Models

Behnoush Khavari

DIRO & Mila

Université de Montréal

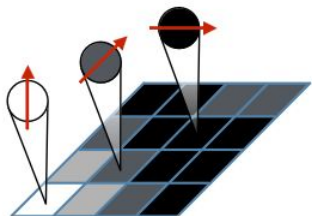
behnoush.khavari@umontreal.ca

Guillaume Rabusseau

DIRO & Mila, CIFAR AI Chair

Université de Montréal

grabus@iro.umontreal.ca



	rank one	CP	Tucker	TT / TR
Decomposition				
Lower Bound (condition)	$(d-1)p$	$rd \quad (r \leq d^{p-1})$	$r^p \quad (r \leq d)$	$r^2 d \quad (r \leq d^{\lfloor \frac{p-1}{2} \rfloor}, p \geq 3)$ $\frac{p(r^2 d - 1)}{3} \quad (r = d, p/3 \in \mathbb{N})$
Upper bound	$2dp \log(12p)$	$2prd \log(12p)$	$2(r^p + prd) \log(24p)$	$2pr^2 d \log(12p)$

Previous work

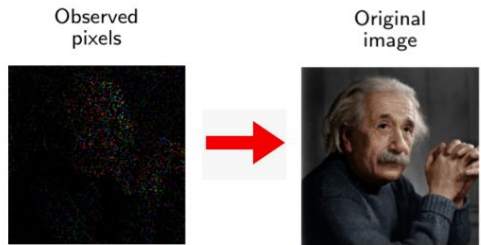
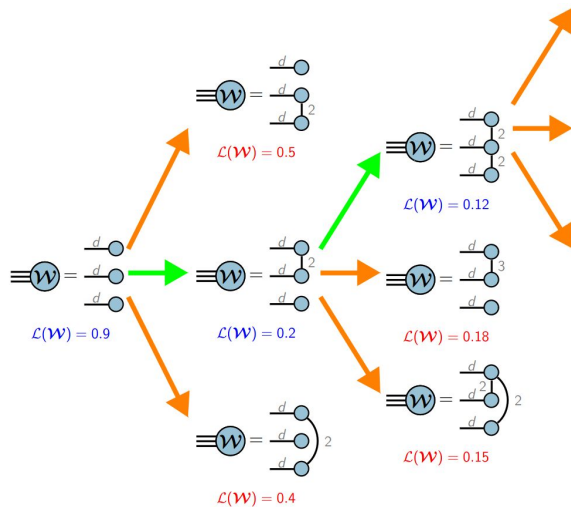
Towards Adaptive Tensor Network Structures

Adaptive Learning of Tensor Network Structures

Meraj Hashemizadeh
Mila & DIRO
Université de Montréal

Michelle Liu
Mila & DIRO
Université de Montréal

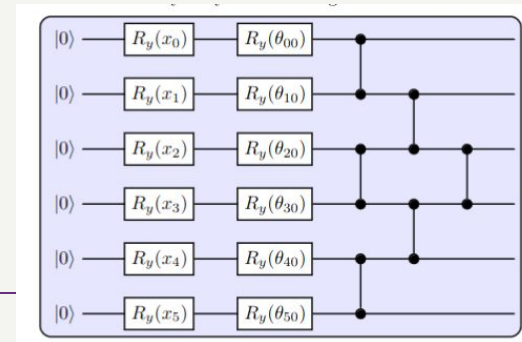
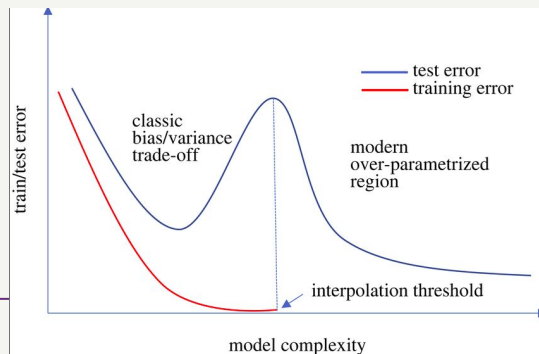
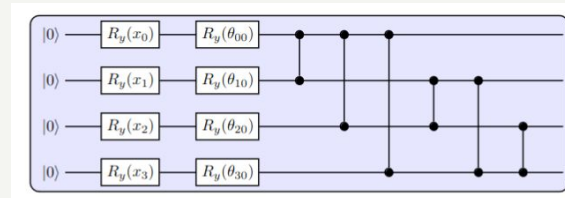
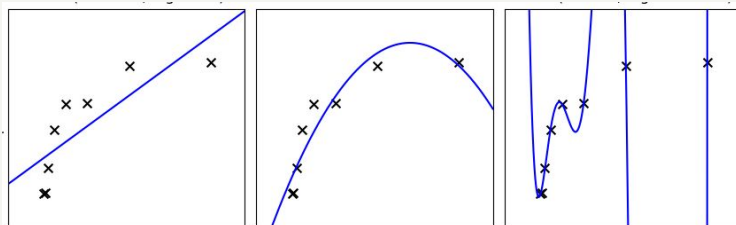
Jacob Miller
Mila & DIRO
Université de Montréal



Current interests

Limits and promises of Quantum ML (QML) through the lens of Tensor Networks

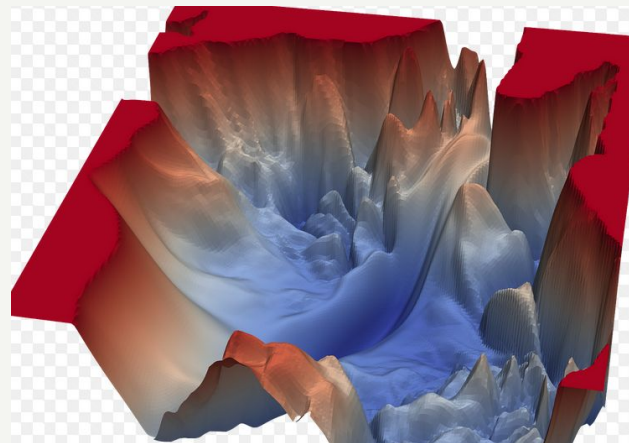
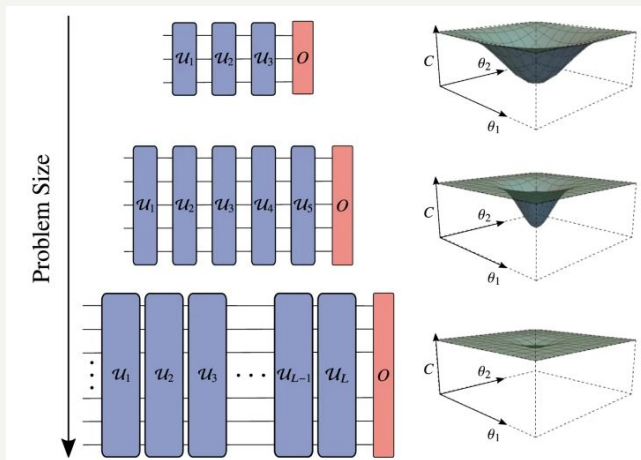
- Inductive biases and bias-variance trade-off in QML.
- Which quantum circuits can't be simulated by Tensor Networks? How are they relevant to ML?



Current interests

QML and Learning Theory

- Generalization bounds for QML.
- What is the role of the optimization algorithm in the generalization capabilities of QML models?
- Benefits of depth in Quantum circuits for ML.
- Difficulties of optimizing quantum circuits.



Current interests

Probabilistic Sequence Modeling on Quantum Hardware

- Going beyond Quantum-Inspired models
 - Tractable Operation with Tensor Networks \nRightarrow Tractable on Quantum Hardware
- Interconnected questions:
 - Which distributions can be **expressed** on quantum computers?
 - Which distributions can be **learned** on quantum computers?
 - Once a distribution is encoded as a quantum state, which **complex inference tasks** can we perform?
- Beyond sequences: When/How can quantum computing help model complex distributions over structured data?

Born machines

$$\mathbb{P}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \left(\alpha \overset{R}{\text{---}} \mathcal{A} \overset{R}{\text{---}} \mathcal{A} \overset{R}{\text{---}} \mathcal{A} \overset{R}{\text{---}} \mathcal{A} \overset{R}{\text{---}} \omega \right)^2$$

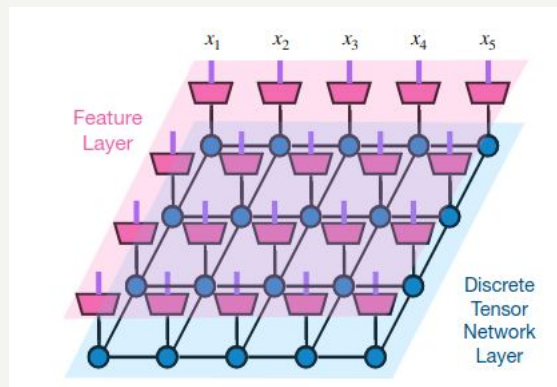
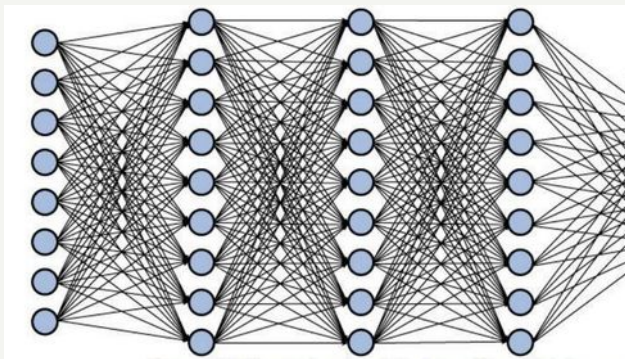
The diagram shows a sequence of operations: a parameter α connected by a horizontal line labeled R to a series of four \mathcal{A} blocks. Each \mathcal{A} block is connected vertically by a line labeled d to an input variable $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ respectively. The final \mathcal{A} block is connected by a horizontal line labeled R to a parameter ω . The entire expression is squared.



Current interests

Explainability of ML models

- Explainability and analysis of deep learning (DL) models through tensor networks and entanglement entropy
 - Particular interest in graph models
- Distilling DL models into tensor networks for efficient inference
- Analyzing DL models using TN (how much entanglement can DL models capture?)

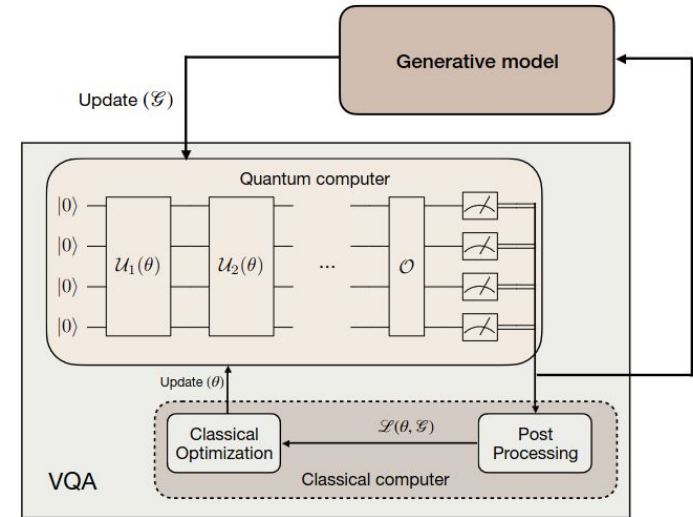
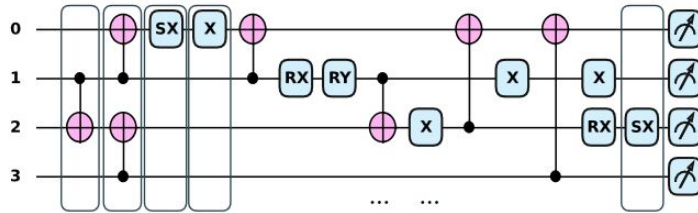


Current interests

Learning Quantum Circuit Designs

- Can we learn/select quantum circuits adaptively from data/experiments?
- The space of quantum circuits is combinatorial in nature... GFlowNets!

Enhancing Variational Quantum Algorithms: Effective Quantum Ansatz Design Using GFlowNets



Research Group

Post-doc

- Jun Dai

PhD students

- Maude Lizaire
- Beheshteh Tolouei Rakhshan
- [Andy Huang](#) (co-supervised with Reihaneh Rabbany)
- Farzaneh Heidari (co-supervised with [Jian Tang](#))
- Marawan Abdel Hameed
- Alireza Dizaji

MSc students

- Omar Chikhar (co-supervised with [Stefanos Kourtis](#))
- Michaël Rizvi-Martel
- Soroush Omranpour (co-supervised with Reihaneh Rabbany)

Undergrad students

Interns

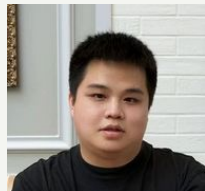
Visiting researcher

- Slimane Thabet (PhD, LIP6, France)



Farzaneh Heidari (PhD)

TN to explain graph learning models



Jun Dai (postdoc)

*Probabilistic Modeling on Quantum Hardware
GFlowNets for Quantum Design*

Past students

- Tianyu Li (PhD, now researcher at Samsung)
- [Jacob Miller](#) (postdoc, now at Zapata Computing)
- [Meraj Hashemizadeh](#) (MSc)
- Behnoush Khavari (MSc)