Title: Neural network enhanced cross entropy benchmark for monitored circuits

Speakers: Yangrui Hu

Collection/Series: Special Seminars

Subject: Other

Date: May 15, 2025 - 9:30 AM

URL: https://pirsa.org/25050019

Abstract:

We explore the interplay of quantum computing and machine learning to advance experimental protocols for observing measurement-induced phase transitions (MIPT) in quantum devices. In particular, we focus on trapped ion monitored circuits and apply the cross entropy benchmark recently introduced by [Li et al., Phys. Rev. Lett. 130, 220404 (2023)], which can mitigate the postselection problem. By doing so, we reduce the number of projective measurements -- the sample complexity required per random circuit realization, which is a critical limiting resource in real devices. Since these projective measurement outcomes form a classical probability distribution, they are suitable for learning with a standard machine learning generative model. In this work, we use a recurrent neural network (RNN) to learn a representation of the measurement record for a native trapped-ion MIPT, and show that using this generative model can substantially reduce the number of measurements required to accurately estimate the cross entropy. This illustrates the potential of combining quantum computing and machine learning to overcome practical challenges in realizing quantum experiments.

Neural Network Enhanced Cross Entropy Benchmark for Monitored Circuits

Yangrui Hu (BIMSA)
2025/05/15
IVADO-Perimeter-IC Online Seminar

Based on arXiv: 2501.13005 [quant-ph] with Yi Hong Teoh, William Witczak-Krempa, and Roger G. Melko

1

The theme: AI + Quantum

Quantum Computing ⟷ Machine Learning

↘ ↙

to advance experimental protocols

- Generative models are capable of learning complex probability distributions
- Their ability to scale and generalize makes them well-suited for applications in quantum sciences, in particular, quantum state representations, quantum error-correction, and experimental protocols [a large body of literature]

[1] J. Kaplan, et al., Scaling Laws for Neural Language Models, arXiv:2001.08361 [cs.LG].
[2] J. Wei, et al., Emergent abilities of large language models, arXiv:2206.07682.
[3] J. Carrasquilla, Machine learning for quantum matter, Advances in Physics: X 5, 1797528 (2020).
[4] A. Dawid et al., Modern applications of machine learning in quantum sciences (2022) arXiv:2204.04198 [quant-ph].
[5] R. G. Melko and J. Carrasquilla, Language models for quantum simulation, Nature Computational Science 4, 11 (2024)

The question: measurement-induced phase transitions

- We explore the interplay of quantum computing and machine learning to advance experimental protocols for observing measurement-induced phase transitions (MIPT) in quantum devices
- MIPT is an entanglement phase transition in monitored quantum circuits

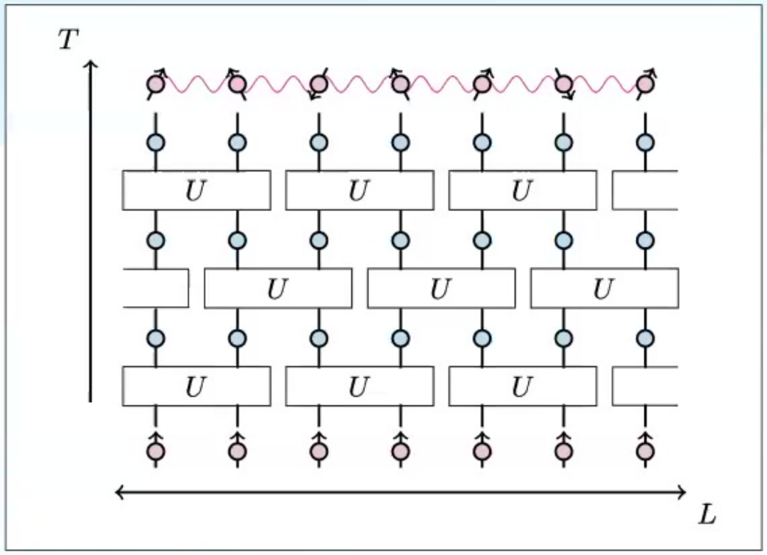
quantum computers + (classical) observers

  - offer an opportunity to explore new physics, including phases and phase transitions
  - many contexts have been studied extensively! [a large body of literature]

[6] B. Skinner, J. Ruhman, and A. Nahum, Measurement-Induced Phase Transitions in the Dynamics of Entanglement, Phys. Rev. X 9, 031009 (2019)
[7] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, Phys. Rev. B 100, 134306 (2019).

The question: MIPT

- MIPT is an entanglement phase transition in monitored quantum circuits
- Consider monitored circuits in the following brick-layer fashion

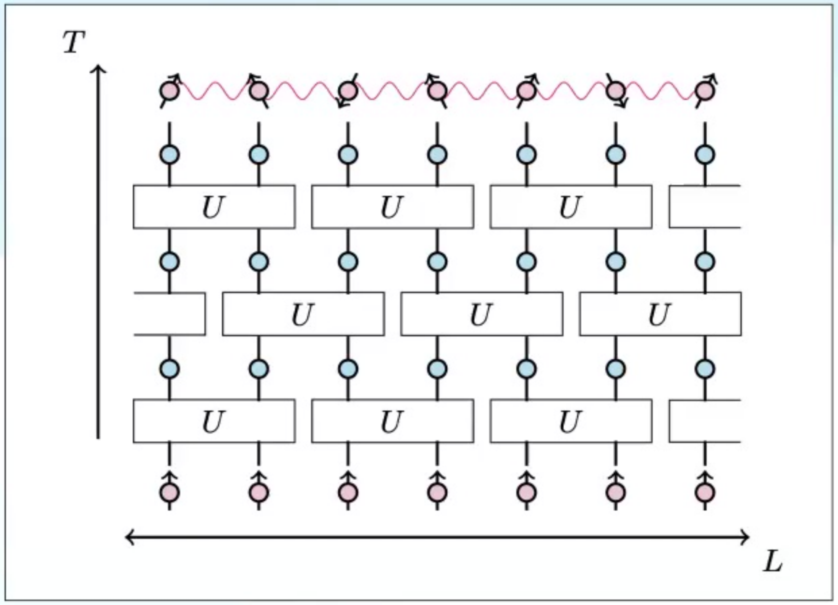


- Schematically, $|\psi(t)\rangle \propto \left(\prod_{l=1}^{t} P_l U_l\right) |\psi(0)\rangle$
- repeated layers of local random unitary dynamics
$$U_l = \bigotimes_{i \in \text{even/odd}} U_l^{[i,i+1]}$$
- projective measurements with a measurement rate p
$$P_l = \bigotimes_i P_{m_i}^{[i]}, \; P_{\pm}^{[i]} = \frac{I \pm \sigma_z^{[i]}}{2}$$

[6] B. Skinner, J. Ruhman, and A. Nahum, Measurement-Induced Phase Transitions in the Dynamics of Entanglement, Phys. Rev. X 9, 031009 (2019)
[7] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, Phys. Rev. B 100, 134306 (2019).

The question: MIPT

Later we will check in classical simulations



- It takes depth $T = O(L)$ for the dynamics to reach equilibrium and the steady-state ensemble exhibits an entanglement phase transition
- The steady states $\{ |\psi(t)\rangle \}$ form an ensemble
- Consider the trajectory-averaged entanglement entropy $\bar{S}(A) = \mathbb{E}_{\{U_l\},\{P_l\}} S_A(|\psi(T)\rangle)$
  - $p = 0 : \bar{S}(A) \propto L_A$
  - $p < p_c : \bar{S}(A) \propto L_A$ volume law
  - $p = 1 : \bar{S}(A) = 0$
  - $p > p_c : \bar{S}(A) = O(1)$ area law
  - $p = p_c : \bar{S}(A) = \alpha \log L_A + O(1)$

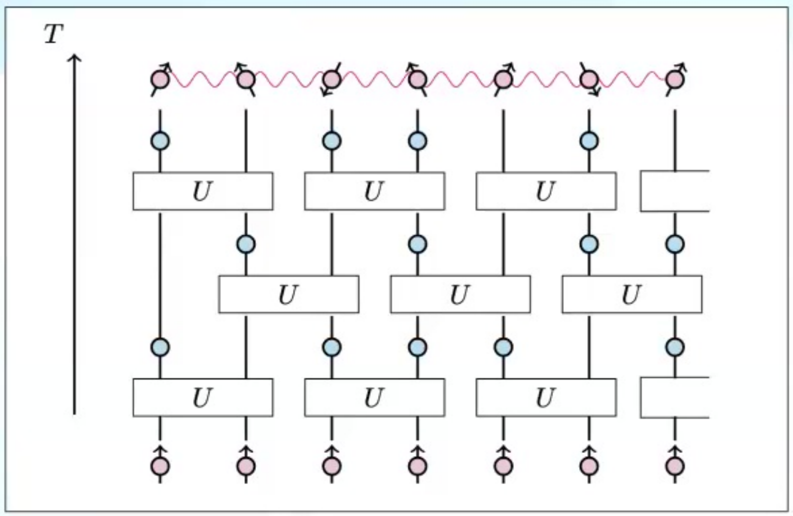
depend on the ensemble

The question: MIPT & post-selection problem

- To compute the entropy, the state reconstruction is needed

$$S_A = -\mathrm{Tr}(\rho_A \log \rho_A)$$

- To prepare the same final state multiple times (i.e. reproduce the same trajectory), the number of circuit runs grows exponentially with the system size



- the number of measurement sites $N \sim pLT$
- For a specific circuit (with fixed unitaries and measurement sites), the number of possible configurations $\sim 2^N$

The question: post-selection problem

- To mitigate the post-selection problem:
  - Look at quantities other than entanglement entropy
  - Apply machine learning!

For example, cross-entropy benchmark

quantum components + classical components

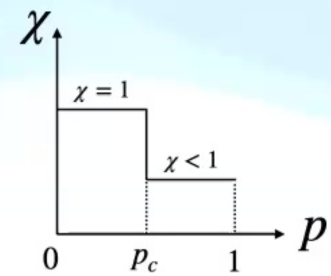
↓

a playground for exploration with machine learning, e.g. [8]: map measurement outcomes in monitored circuits to the reference qubit configuration

[8] H. Dehghani, A. Lavasani, M. Hafezi, and M. J. Gullans, Neural-network decoders for measurement induced phase transitions, Nature Commun. 14, 2918 (2023)

Cross-entropy Benchmark

- Consider the cross entropy $\chi = \mathbb{E}_C \chi_C$ between the probability distributions formed by measurement outcomes with two different input states
- $\chi = \mathbb{E}_C \chi_C$ is an order paramter for MIPT [9]
  - The volume law phase (low-p): the "coding" property $\rightarrow \chi = 1$
  - The area law phase (high-p): $\chi < 1$



ρ or σ → Monitored circuit C $\{m_1, m_2, \ldots, m_N\}$ → $p^{\rho}_{\vec{m}}$ or $p^{\sigma}_{\vec{m}}$

$$\chi_C = \frac{\sum_{\vec{\mathbf{m}}} p^{\rho}_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}}}{\sum_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}}} = \frac{\left\langle p^{\sigma}_{\vec{\mathbf{m}}} \right\rangle_{\rho}}{\left\langle p^{\sigma}_{\vec{\mathbf{m}}} \right\rangle_{\sigma}}$$

[9] Y. Li, Y. Zou, P. Glorioso, E. Altman, and M. P. A. Fisher, Cross Entropy Benchmark for Measurement-Induced Phase Transitions, Phys. Rev. Lett. 130, 220404 (2023)

Cross-entropy Benchmark

- When we compute χ_C, $p_{\vec{m}}$'s are infeasible
- To approximate (the law of large numbers):

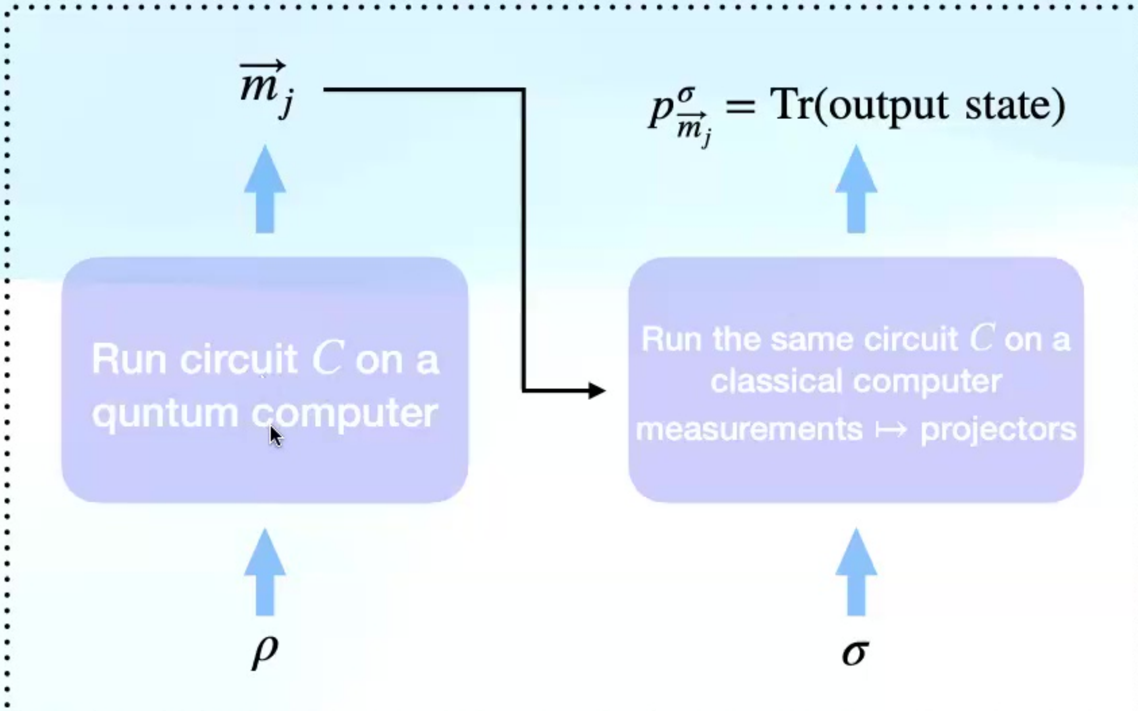
$$\left\langle p^{\sigma}_{\vec{m}} \right\rangle_{\rho} = \sum_{\vec{\mathbf{m}}} p^{\rho}_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}} = \lim_{M\to\infty} \frac{1}{M} \sum_{j=1, \vec{m}_j \sim p^{\rho}_{\vec{m}}}^{M} p^{\sigma}_{\vec{m}_j} ,$$

$$\left\langle p^{\sigma}_{\vec{m}} \right\rangle_{\sigma} = \sum_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}} p^{\sigma}_{\vec{\mathbf{m}}} = \lim_{M\to\infty} \frac{1}{M} \sum_{k=1, \vec{m}_k \sim p^{\sigma}_{\vec{m}}}^{M} p^{\sigma}_{\vec{m}_k}$$

- [Li et al., Phys. Rev. Lett. **130**, 220404 (2023)] proposed the cross entropy benchmark for observing MIPTs on quantum computers, which mitigates the post-selection problem

Cross-entropy Benchmark

- [Li et al., Phys. Rev. Lett. **130**, 220404 (2023)] protocol:



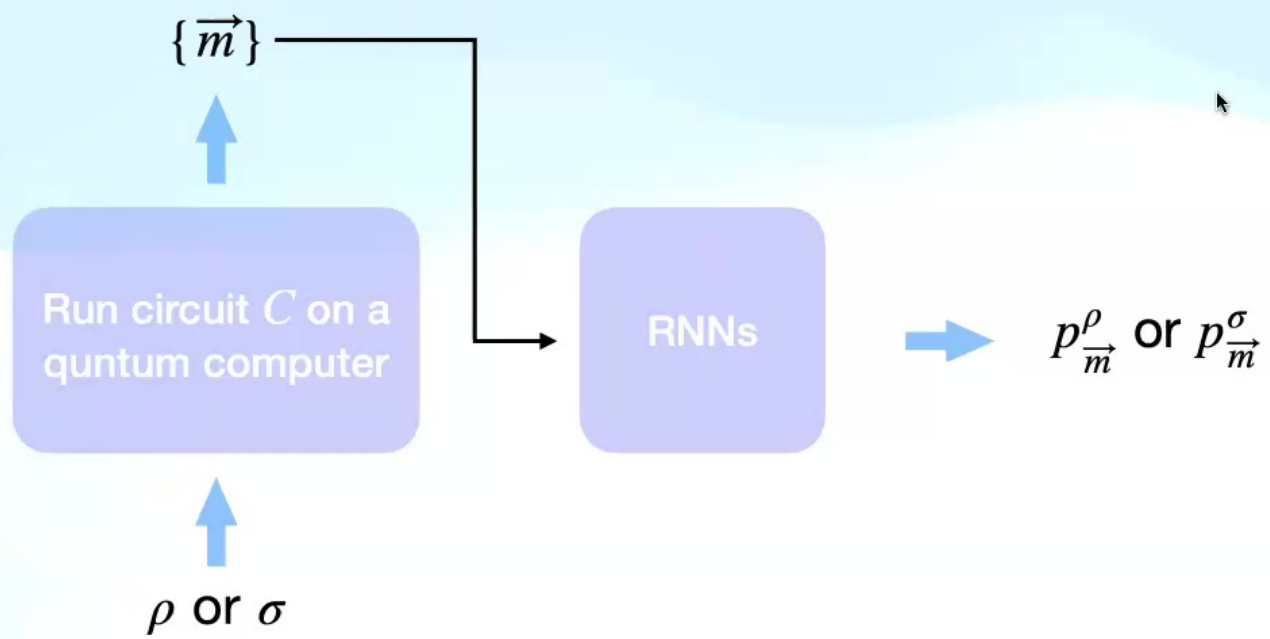
- repeat M times to approximate
$$\left\langle p^{\sigma}_{\vec{m}} \right\rangle_{\rho} = \lim_{M\to\infty} \frac{1}{M} \sum_{j=1,\vec{m}_j\sim p^{\rho}_{\vec{m}}}^{M} p^{\sigma}_{\vec{m}_j}$$
- Similarly, we get $\left\langle p^{\sigma}_{\vec{m}} \right\rangle_{\sigma}$ and compute χ_C
- Average over circuits: Repeat the procedure for M_C sampled ρ-circuits and compute the average of χ_C across these circuits

Cross-entropy Benchmark

- [Li et al., Phys. Rev. Lett. **130**, 220404 (2023)]'s protocol
  - For generic non-stabilizer circuits, the scalability is restricted by classical simulations
  - The sample complexity — the number of measurement runs M required to obtain an accurate estimation of the cross entropy — is relatively unexplored
  - The XEB relies solely on measurement outcomes and their associated probabilities, which naturally aligns with the framework of language models

→ Neural Network enhanced protocol

NN-enhanced protocol



Outline

- Setup: trapped ion monitored circuit
- Classical simulation: validate the benchmark
- Vanilla RNNs
- RNN-enhanced protocol
- Training results & data complexity
- Remarks

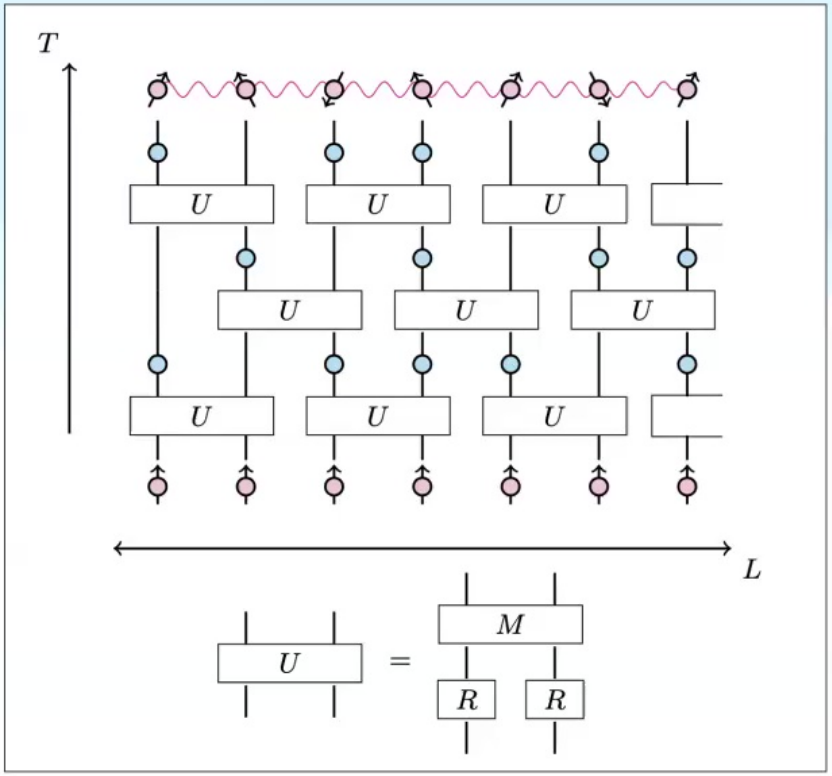
Trapped-ion monitored circuit

- Trapped ions have become a key experimental testing ground for the feasibility of monitored circuit experiments [10]
  - the precise and programmable control achievable over individual qubits [11,12]
  - pure quantum states can be prepared consistently [13,14]
  - single-qubit measurements with a constant rate are feasible through optical addressing [15,16]

[10] S. Czischek, G. Torlai, S. Ray, R. Islam, and R. G. Melko, Simulating a measurement-induced phase transition for trapped ion circuits, Phys. Rev. A 104, 062405 (2021).
[11] A. H. Myerson, D. J. Szwer, S. C. Webster, D. T. C. Allcock, M. J. Curtis, G. Imreh, J. A. Sherman, D. N. Stacey, A. M. Steane, and D. M. Lucas, High-fidelity readout of trapped-ion qubits, Phys. Rev. Lett. 100, 200502 (2008).
[12] J. E. Christensen, D. Hucul, W. C. Campbell, and E. R. Hudson, High-fidelity manipulation of a qubit enabled by a manufactured nucleus, npj Quantum Information 6, 35 (2020).
[13] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, Demonstration of a small programmable quantum computer with atomic qubits, Nature 536, 63–66 (2016).
[14] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao, Programmable quantum simulations of spin systems with trapped ions, Rev. Mod. Phys. 93, 025001 (2021).
[15] S. Crain, E. Mount, S. Baek, and J. Kim, Individual addressing of trapped 171yb+ ion qubits using a microelectromechanical systems-based beam steering system, Applied Physics Letters 105, 181115 (2014)
[16] C.-Y. Shih, S. Motlakunta, N. Kotibhaskar, M. Sajjan, R. Habľutzel, and R. Islam, Reprogrammable and highprecision holographic optical addressing of trapped ions for scalable quantum control, npj Quantum Inf. 7, 57 (2021).

Setup: trapped-ion monitored circuit

- Consider the same hybrid circuit setup in [10]
  - L trapped atomic ions
  - open boundary condition
- Entangling unitaries:
  - Mølmer-Sørensen (MS) gate
    $M_{j,k}(\theta) = \cos\theta \, \mathbb{I}_j \otimes \mathbb{I}_k - i \sin\theta \, X_j \otimes X_k$
  - Single-qubit rotations
    $R_j(\varphi_j) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -ie^{-i\varphi_j} \\ -ie^{i\varphi_j} & 1 \end{pmatrix}$
- Projective measurements along Z-axis



15

[10] S. Czischek, G. Torlai, S. Ray, R. Islam, and R. G. Melko, Simulating a measurement-induced phase transition for trapped ion circuits, Phys. Rev. A 104, 062405 (2021).

Trapped-ion monitored circuit

- Randomness: fixing the MS gate while incorporating random R_j's is sufficient [10]
  - each φ_j is randomly chosen from $\varphi_j \in \left\{0, \frac{\pi}{2}, \frac{\pi}{4}\right\}$

    $0, \frac{\pi}{2}$: 90-degree rotations about the x- and y-axes

    $\frac{\pi}{4}$: non-Clifford

  - [17] provides a mathematical framework for similar random circuit models (entangling gates are fixed and the randomness only comes from single-qubit gates)

[10] S. Czischek, G. Torlai, S. Ray, R. Islam, and R. G. Melko, Simulating a measurement-induced phase transition for trapped ion circuits, Phys. Rev. A 104, 062405 (2021).
[17] L. Kong, Z. Li, and Z.-W. Liu, Convergence efficiency of quantum gates and circuits, (2024), arXiv:2411.04898 [quant-ph].

Classical Simulations

- [10] analyzed the entanglement entropy and identified a MIPT in this system.
  - The critical measurement rate $p_c \approx 0.17 \pm 0.02$
- Validate the XEB
  - Introduce additional encoding layers following the procedure outlined in [9]
  - Steady states
  - Simulate the XEB

[9] Y. Li, Y. Zou, P. Glorioso, E. Altman, and M. P. A. Fisher, Cross Entropy Benchmark for Measurement-Induced Phase Transitions, Phys. Rev. Lett. 130, 220404 (2023)
[10] S. Czischek, G. Torlai, S. Ray, R. Islam, and R. G. Melko, Simulating a measurement-induced phase transition for trapped ion circuits, Phys. Rev. A 104, 062405 (2021).

Classical Simulations: encoding layers

- To obtain stable numerical results
- Encoding layers consisting solely of random unitaries

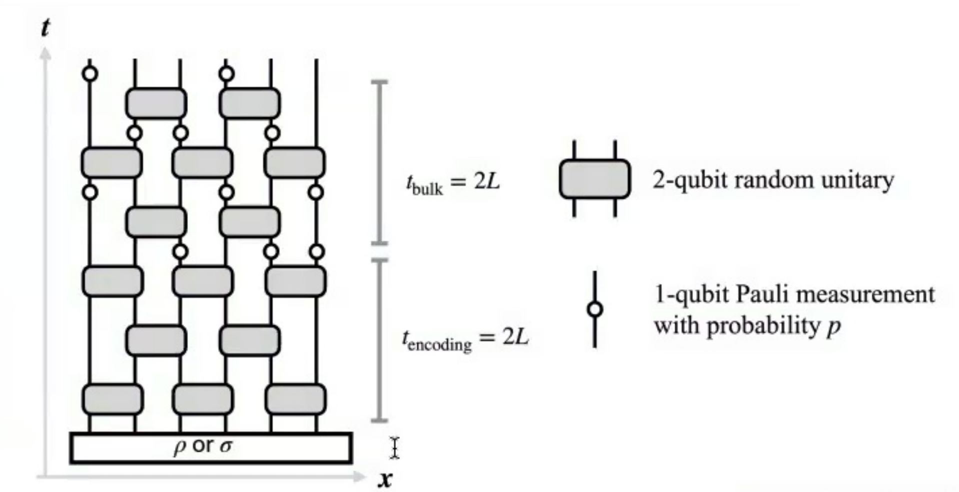


Figure obtained from [10]

18

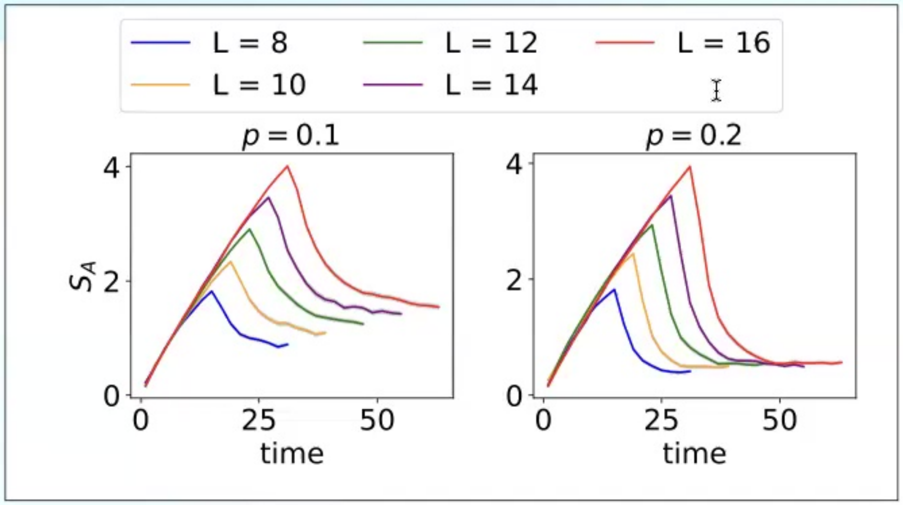
[10] Y. Li, Y. Zou, P. Glorioso, E. Altman, and M. P. A. Fisher, Cross Entropy Benchmark for Measurement-Induced Phase Transitions, Phys. Rev. Lett. 130, 220404 (2023)

Classical Simulations: steady states

- we study the evolution of the half-chain von Neumann entropy

$$S_A = -\mathrm{Tr}(\rho_A \log \rho_A)$$ A is chosen as the first $L/2$ qubits

- S_A is simply computed using the exact diagonalization of the density matrix



- For each case, we simulate $M_C = 100$ circuit realizations
- $T_{\text{encoding}} = T_{\text{bulk}} = 2L$
- Consistent with the encoding-bulk circuit structure

Classical Simulations: cross-entropy

- Recall: the experimental protocol for observing MIPT on a quantum device proposed in [10]

1. Run the ρ-circuit
2. Simulate the σ-circuit with same parameters
3. Compute χ_C: perform M measurement runs
4. Average over circuits: Repeat the procedure for M_C sampled ρ-circuits

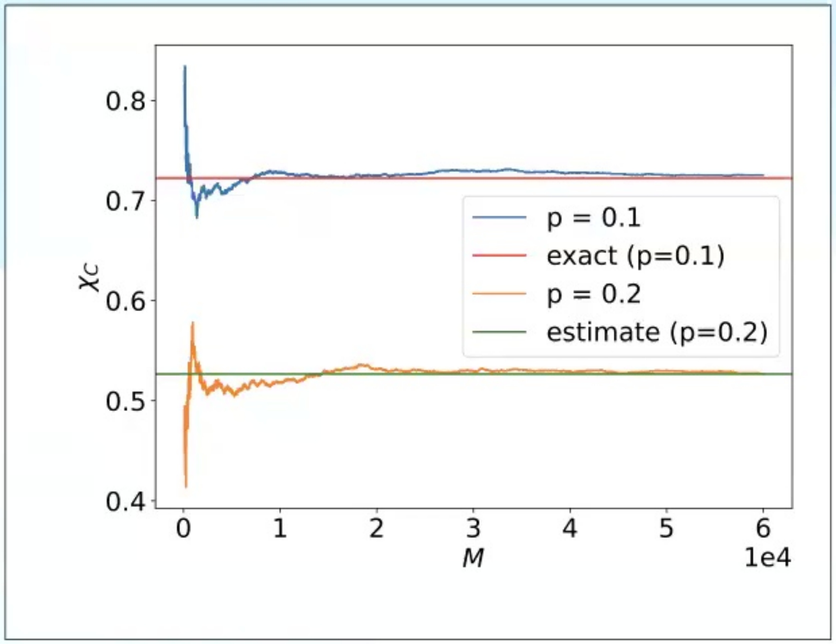
What M we should choose?

$$\rho = |+\rangle^{\otimes L},\ \sigma = |0\rangle^{\otimes L}$$

[10] Y. Li, Y. Zou, P. Glorioso, E. Altman, and M. P. A. Fisher, Cross Entropy Benchmark for Measurement-Induced Phase Transitions, Phys. Rev. Lett. 130, 220404 (2023)

IVADO
Page 21 of 33

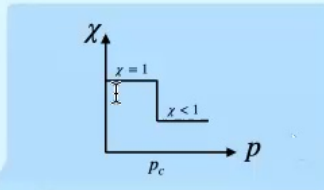
Classical Simulations: cross-entropy



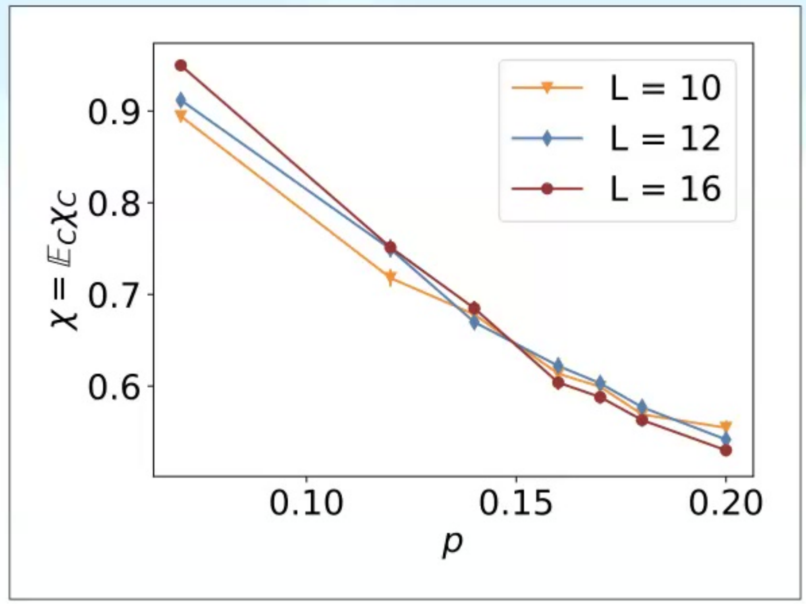
- Consider two example circuits with $L = 8$
  - $p = 0.1$, $N = 12$
  - $p = 0.2$, $N = 24$
- χ_C initially fluctuates before gradually converging to a constant value as M increases.

21

Classical Simulations: cross-entropy



- Focus on small systems (finite-size effect)
- Consider $M = 5 \times 10^3$ and $M_C = 100$
- The crossing behavior confirms that the XEB is effective for trapped ion hybrid circuits
- Consistent with $p_c \approx 0.17 \pm 0.02$ obtained in [10]



[10] S. Czischek, G. Torlai, S. Ray, R. Islam, and R. G. Melko, Simulating a measurement-induced phase transition for trapped ion circuits, Phys. Rev. A 104, 062405 (2021).

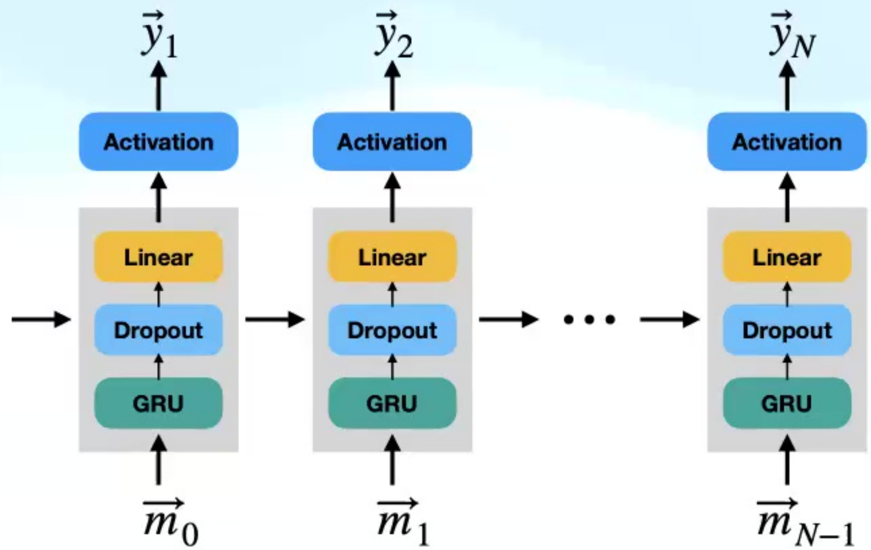
Vanilla RNNs for probability distributions

- We consider a similar architecture for simple (vanilla) RNNs as in [18]

- Input: one-hot encoding of the measurement outcomes $\{\vec{m}_1, \vec{m}_2, \ldots, \vec{m}_N\}$

$$m_i = \begin{cases} 0 \\ 1 \end{cases} \mapsto \ \vec{m}_i = \begin{cases} (1,0) \\ (0,1) \end{cases} , \quad \vec{m}_0 = (0,0)$$

- RNN cell:
  - A gated recurrent unit (GRU) layer $\mapsto$ hidden state (N_h)
  - A dropout layer for regularization to prevent overfitting
  - A linear layer $\mapsto$output logits ($N_o = 2$)



[18] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, Recurrent neural network wave functions, Physical Review Research 2, 023358 (2020)

Vanilla RNNs for probability distributions

- A Softmax activation layer $\mapsto \vec{y}_i$ (conditional probability distribution of the outcomes at site i)

e.g. $\vec{y} = \begin{pmatrix} 0.3 \\ 0.7 \end{pmatrix}$

$$\text{Softmax}(x_n) = \frac{\exp(x_n)}{\sum_i \exp(x_i)}$$

- Given $\{m_1, \ldots, m_{i-1}\}$, the conditional probability of obtaining m_i reads

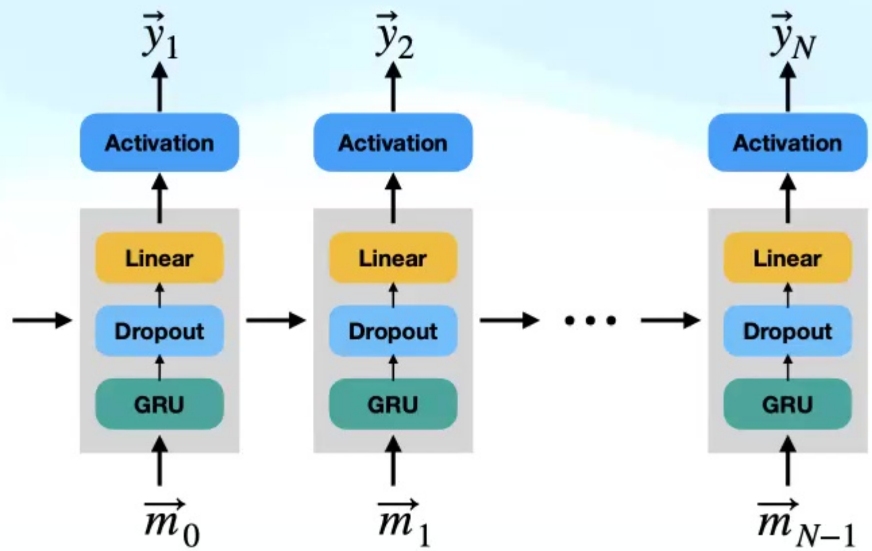
$$P(m_i \,|\, m_1, m_2, \ldots, m_{i-1}) = \vec{y}_i \cdot \vec{m}_i$$

- The probability of observing the full configuration

$$P(m_1, m_2, \ldots, m_N) = \prod_{i=1}^{N} \vec{y}_i \cdot \vec{m}_i$$

- We employ the Negative Log Likelihood as the loss function

$$\mathscr{L} = -\frac{1}{N} \text{avg}_{\{\vec{m}\}} \log P(\vec{m})$$



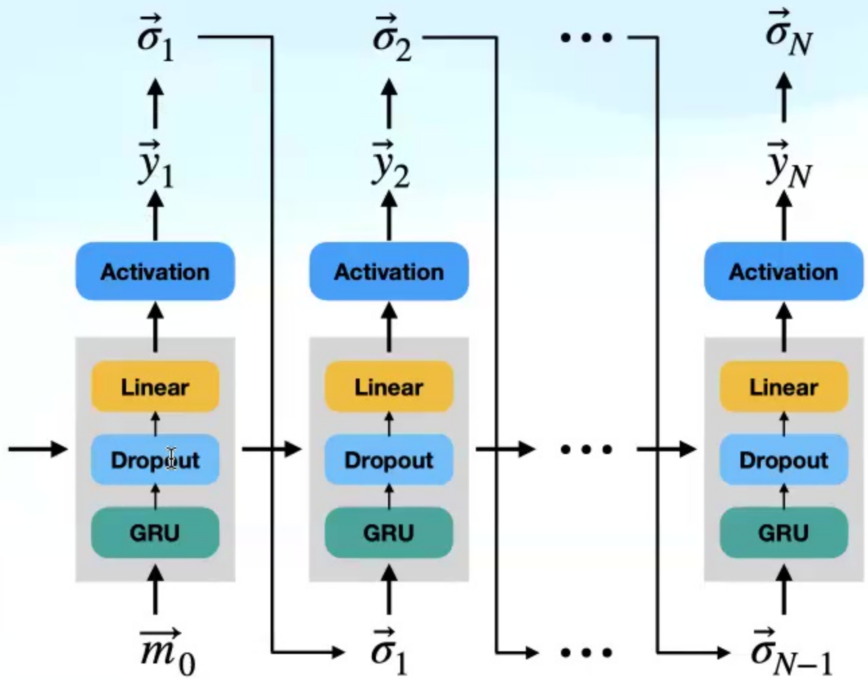
24

[18] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, Recurrent neural network wave functions, Physical Review Research 2, 023358 (2020)

IVADO
Page 25 of 33

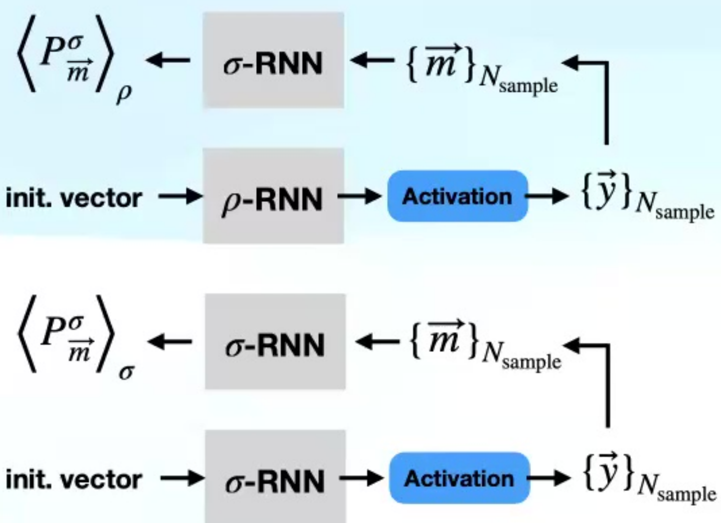
Vanilla RNNs for probability distributions

- Sampling: the autoregressive property of the model yields a recursive sampling process



25

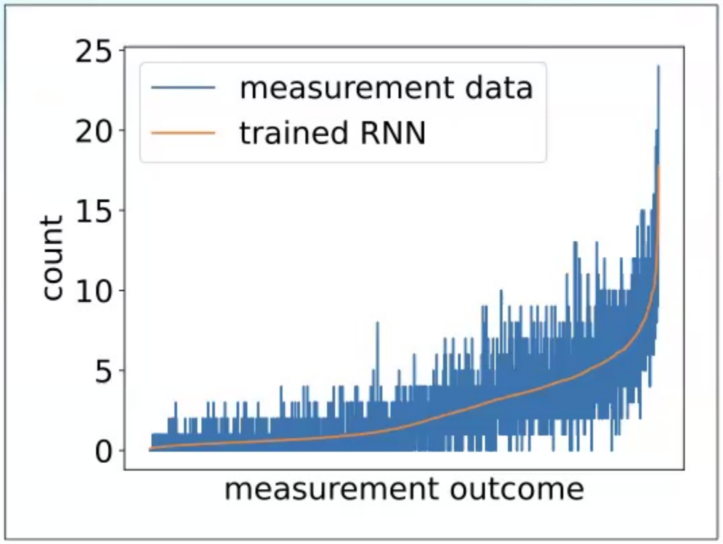
RNN-enhanced protocol



1. Sample a random circuit C with a given system size L and measurement rate p
2. Run ρ- and σ-circuits on quantum computers and record the measurement outcomes
3. Train ρ- and σ-RNNs $\mapsto p^{\rho}_{\vec{m}}$ and $p^{\sigma}_{\vec{m}}$
4. Compute $\chi_C = \left\langle p^{\sigma}_{\vec{\mathbf{m}}} \right\rangle_{\rho} \Big/ \left\langle p^{\sigma}_{\vec{\mathbf{m}}} \right\rangle_{\sigma}$
5. Average the cross entropy over circuits.

Training Results

- The distribution formed by the raw data vs associated RNN model
  - Frequency plot formed by the measurement data: blue lines
  - RNN: orange curve



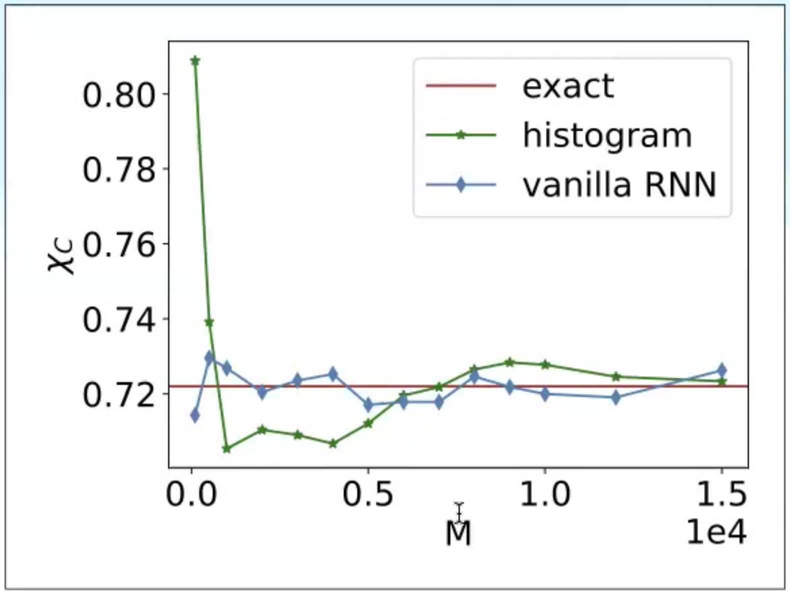
Parametrized models "defuzzify" the histogram

27

ρ-RNN: $M = 10^4$ and $N = 12$

Training Results

- The evolution of estimated χ_C as a function of the number of measurement runs M using the raw data versus RNNs.
- χ_C obtained using the RNN and histogram approaches fluctuate around a constant value at which they converge.
- RNN-enhanced protocol: smaller fluctuations



an guilding example : $N = 12$

IVADO
Page 29 of 33

Data complexity

- Define the accuracy ε of the estimation:

$$\varepsilon \ := \ |\chi_C^{\text{exact}} - \chi_C|$$

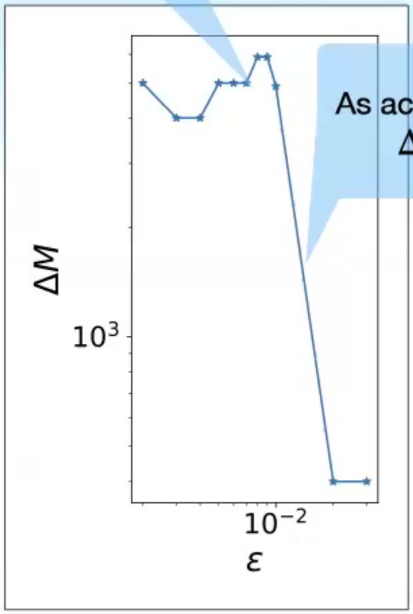
- Sample complexity [19] : the minimum number of measurement runs M_{min} required to achieve a target accuracy ε

- The improvement offered by the RNN-enhanced protocol is quantified by the relative reduction in sample complexity:

$$\Delta M \ := \ M_{\text{min}}^{\text{histogram}} \ - \ M_{\text{min}}^{\text{RNN}}$$

  - ΔM is always positive, indicating that the RNN-enhanced protocol outperforms the histogram approach

High accuracy (small ε): the improvement is substantial



As accuracy ↓(ε ↑), ΔM drops

an guilding example : $N = 12$

[19] D. Iouchtchenko, J. F. Gonthier, A. Perdomo-Ortiz, and R. G. Melko, Neural network enhanced measurement efficiency for molecular groundstates, Mach. Learn. Sci. Tech. 4, 015016 (2023)

Closing Remarks

- In this work, we examine the cross entropy benchmark (XEB) for a measurement-induced phase transition in trapped ion monitored circuits
- We propose a recurrent neural network (RNN) strategy to enhance the XEB signal obtained in a data-limited setting
  - the distribution learned by the RNN is an approximation of the true one
  - the scalability of the RNN-enhanced protocol depends on extending the generative models to handle a larger number of qubits (larger N).

Closing Remarks

Future routes:

- A systematic assessment of the RNN-enhanced protocol in terms of sample complexity
- Evaluate different autoregressive models (in particular, the state-of-the-art transformers) and assess their efficiency in the context of our problem.
- Generalizations in the model architecture (one natural direction would be to incorporate the circuit parameters as an encoder)
- As a next step toward benchmarking quantum devices, it would be interesting to introduce noise models in the simulations
  - conducting simulations with larger system sizes and training the models on real experimental data

IVADO
Page 32 of 33

Thank You