

LEARNING QUANTUM STATES WITH GENERATIVE MODELS

Juan Carrasquilla

Vector Institute
University of Waterloo
Canada CIFAR AI chair

At the Crossroad of Physics and Machine Learning
KITP, Santa Barbara, CA Feb. 14th 2019



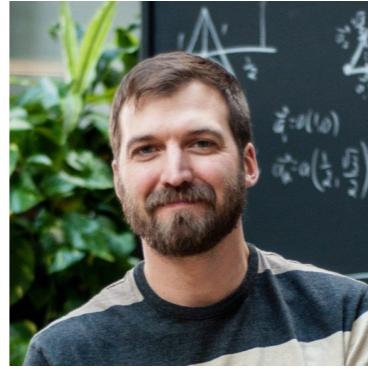
COLLABORATORS



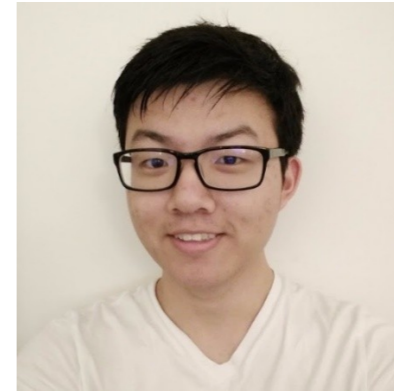
Leandro Aolita
(Universidade Federal do
Rio de Janeiro, ICTP-SAIFR)



Giacomo Torlai
(Flatiron Institute)



Roger Melko (U.
Waterloo and Perimeter)

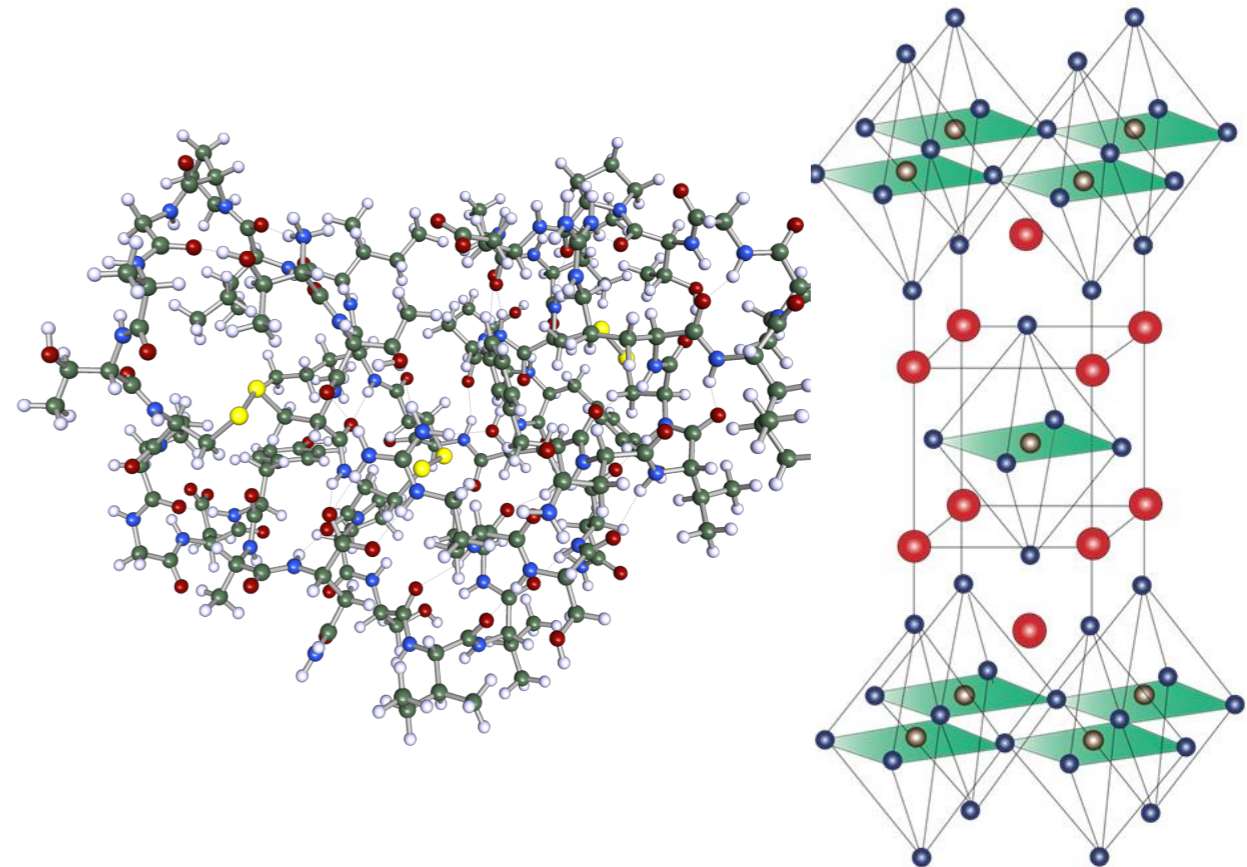


Andrew Tan (U of
Toronto)

THE MANY-BODY PROBLEM IN QUANTUM MECHANICS

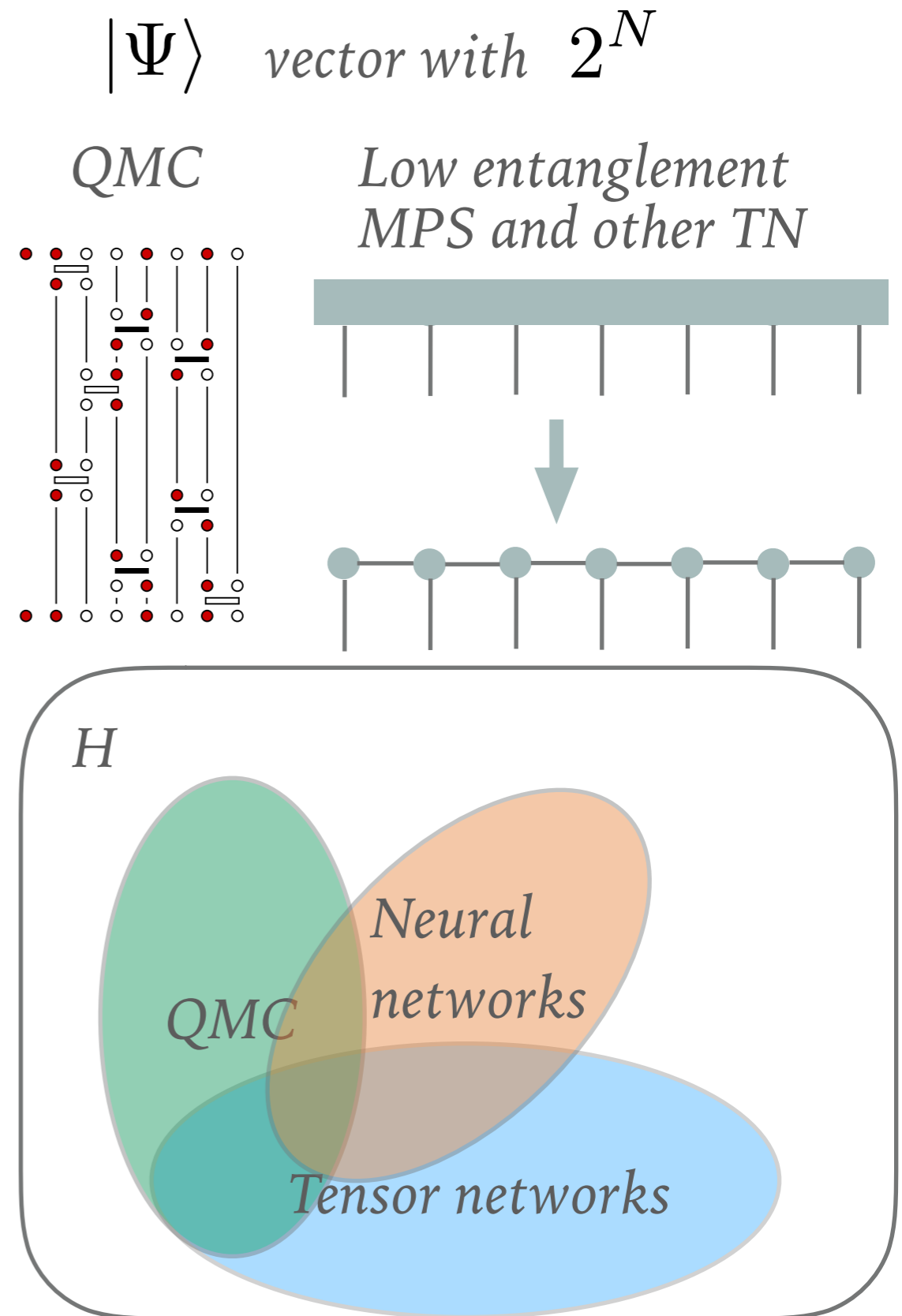
- Generic specification of a quantum state requires resources exponentially large in the number of degrees of freedom N
- Today's best supercomputers can solve the wave equation **exactly** for systems with a maximum of ~ 45 spins.
- Yet, technologically relevant problems in chemistry, condensed matter physics, and quantum computing are much larger than 45.
- Quantum computing

$|\Psi\rangle$ vector with 2^N



THERE IS STILL HOPE FOR CLASSICAL ALGORITHMS

- Nature is sometimes compassionate: amount of information smaller than the maximum capacity.
- Quantum Monte Carlo and other numerical methods based on Tensor Networks exploit this fact.
- Machine learning community deals with equally high dimensional problems ($\sim 2^{28 \times 28}$)
- **Curse of dimensionality** but still successful in **wide spectrum of scientific and technologically relevant areas of research.**

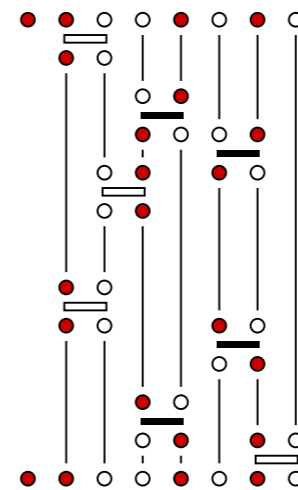


THERE IS STILL HOPE FOR CLASSICAL ALGORITHMS

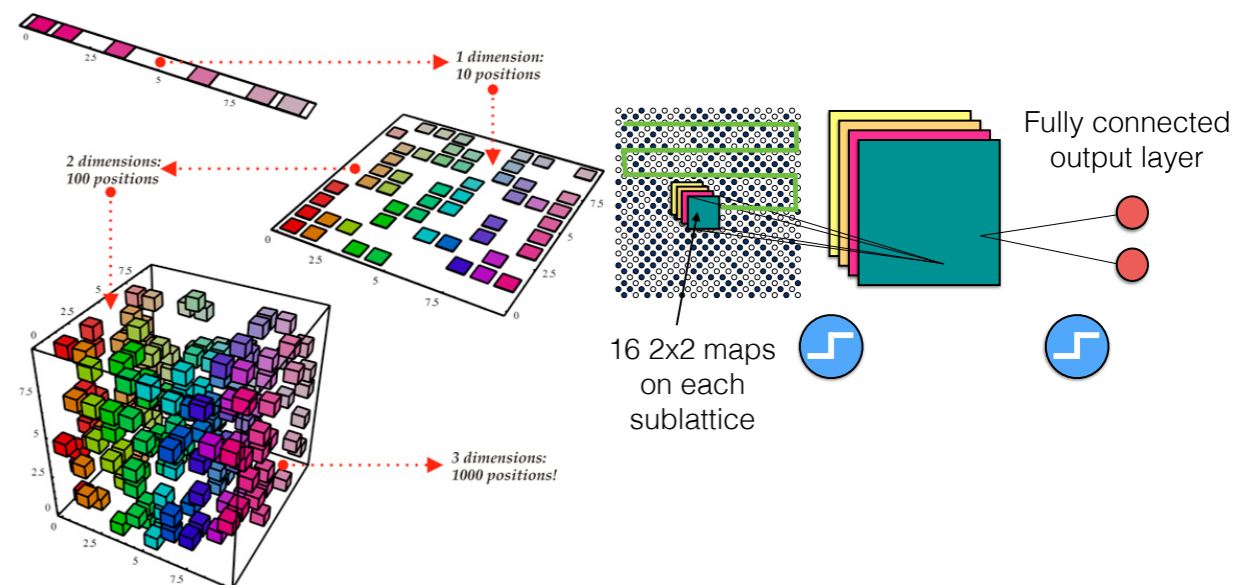
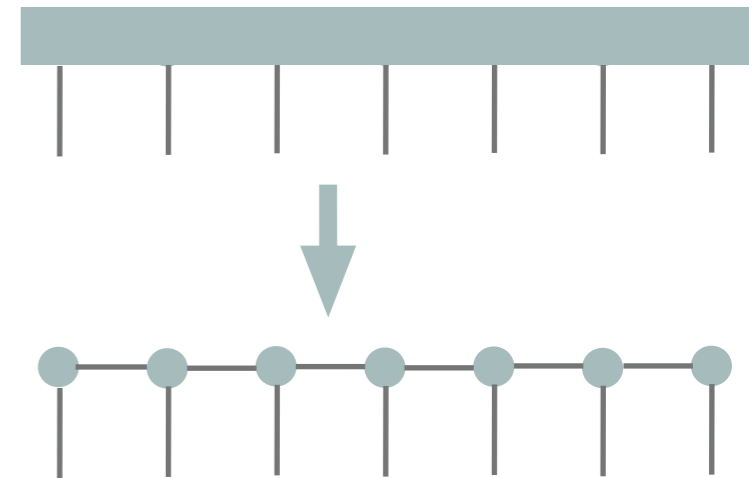
- Nature is sometimes compassionate: amount of information smaller than the maximum capacity.
- Quantum Monte Carlo and other numerical methods based on Tensor Networks exploit this fact.
- Machine learning community deals with equally high dimensional problems ($\sim 2^{28 \times 28}$)
- **Curse of dimensionality** but still successful in **wide spectrum of scientific and technologically relevant areas of research.**

$|\Psi\rangle$ vector with 2^N

QMC



Low entanglement
MPS and other TN



QUANTUM AND CLASSICAL MANY-BODY PHYSICS HAS NOT BEEN THE EXCEPTION

- [ML phases of matter/phase transitions](#) (Carrasquilla, Melko 1605.01735, Wang 1606.00318, Zhang, Kim, 1611.01518)
- [New ML inspired ansatz for quantum many-body systems](#) (Carleo, Troyer 1606.02318, Deng, Li, Das Sarma, 1701.04844, Deng, Li, Das Sarma 1609.09060, Carrasquilla, Melko 1605.01735)
- [Accelerated Monte Carlo simulations](#) (Huang, Wang 1610.02746)
- [Quantum state preparation guided by ML](#) (Bukov, Day, Sels, Weinberg, Polkovnikov and Mehta 1705.00565)
- [Renormalization group analyses, RBMs, PCA](#) (Bradde, Bialek 1610.09733, Koch-Janusz, Ringel 1704.06279, Mehta, Schwab, 1410.3831)
- [Quantum state tomography based on RBMs](#) (Torlai, Mazzola, Carrasquilla, Troyer, Melko, Carleo, 1703.05334)
- [ML based decoders for topological codes](#) (Torlai, Melko 1610.04238, Varsamopoulos, Criger, Bertels, 1705.00857)
- [Supervised Learning with Quantum-Inspired Tensor Networks](#) (Stoudenmire, Schwab 1605.05775, Novikov, Trofimov, Oseledets, 1605.03795)
- [Quantum Boltzmann machines](#) (Amin, Andriyash, Rolfe, Kulchytskyy, Melko, 1601.02036, Kieferova, Wiebe, 1612.05204,)
- [Quantum machine learning algorithms to accelerate learning](#) (Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, Lloyd, 1611.09347)

And many more <https://physicsml.github.io/pages/papers.html>

**LEARNING QUANTUM
STATES AND QUANTUM
STATE TOMOGRAPHY**

QUANTUM STATE TOMOGRAPHY

Quantum state tomography is the process of reconstructing the quantum state by measurements on the system. It “is the gold standard for verification and benchmarking of quantum devices”*

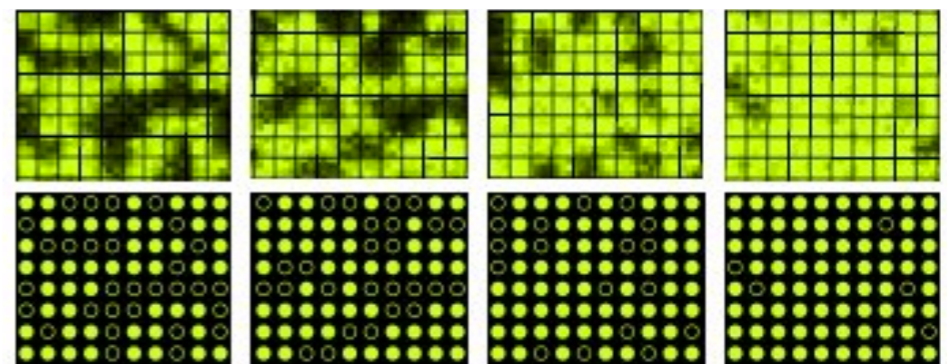
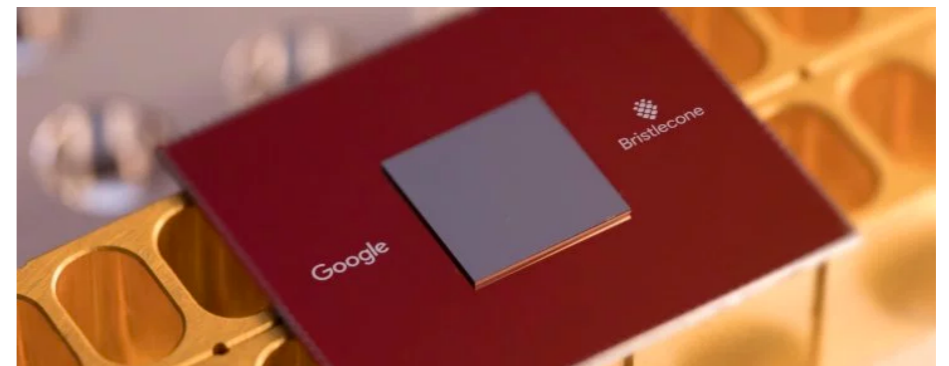
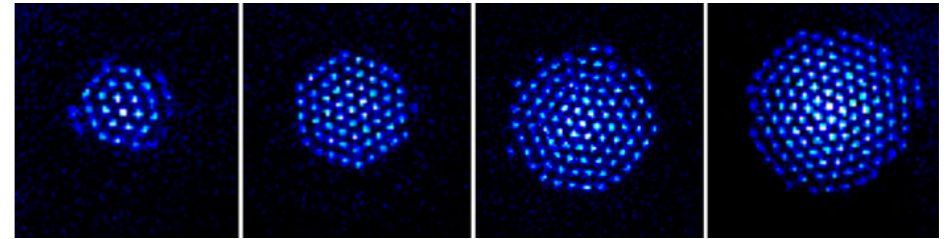
Useful for:

- Characterizing optical signals
- Diagnosing and detecting errors in state preparation, e.g. states produced by quantum computers reliably.
- Entanglement verification

* Efficient quantum state tomography. Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin & Yi-Kai Liu. Nature Communications volume 1, Article number: 149

NEED TO GO BEYOND STANDARD QUANTUM STATE TOMOGRAPHY

- Progress in controlling large quantum systems.
- Availability of arbitrary measurements performed with great accuracy.
- The bottleneck limiting progress in the estimation of states: **curse of dimensionality**.



SYNTHETIC QUANTUM DEVICES ARE GROWING FAST

nature.com > nature > articles > article



Article | Published: 29 November 2017

Probing many-body dynamics on a 51-atom quantum simulator

Hannes Bernien, Sylvain Schwartz, Alexander Keesling, Harry Levine, Ahmed Omran, Hannes Pichler, Soonwon Choi, Alexander S. Zibrov, Manuel Endres, Markus Greiner✉, Vladan Vuletić✉ & Mikhail D. Lukin✉

Nature **551**, 579–584 (30 November 2017) | [Download Citation](#) ↓

nature.com > nature > letters > article



Letter | Published: 29 November 2017

Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator

J. Zhang✉, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong & C. Monroe

Nature **551**, 601–604 (30 November 2017) | [Download Citation](#) ↓

nature.com > nature > letters > article



Letter | Published: 22 August 2018

Observation of topological phenomena in a programmable lattice of 1,800 qubits

Andrew D. King✉, Juan Carrasquilla, [...] Mohammad H. Amin

Nature **560**, 456–460 (2018) | [Download Citation](#) ↓

PHYSICAL REVIEW X

[Highlights](#) [Recent](#) [Subjects](#) [Accepted](#) [Collections](#) [Authors](#) [Referees](#) [Search](#) [Press](#) [About](#) [Staff](#)

Open Access

Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator

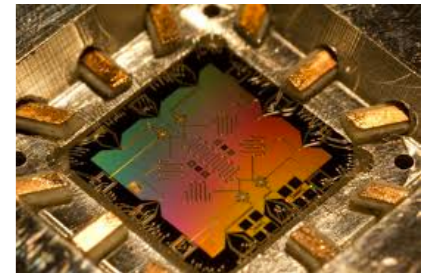
Cornelius Hempel, Christine Maier, Jonathan Romero, Jarrod McClean, Thomas Monz, Heng Shen, Petar Jurcevic, Ben P. Lanyon, Peter Love, Ryan Babbush, Alán Aspuru-Guzik, Rainer Blatt, and Christian F. Roos
Phys. Rev. X **8**, 031022 – Published 24 July 2018



QUANTUM STATE TOMOGRAPHY

Required ingredients

- A quantum system that can be prepared repeatedly
- Set of measurements.
- A training procedure and a model (full density matrix, MPS, MPO, neural networks)
- Certification



A typical state tomography protocol prepares many copies of ρ which are measured in diverse ways, and finally the outcomes of those measurements (data) are analyzed to produce an estimate ρ^* .

TRADITIONALLY QST REQUIRES EXPONENTIAL RESOURCES

Examples

- **Linear inversion method** requires inverting a large matrix and results in an explicit representation of the density matrix

$$P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}} \quad \text{experimental histogram} \quad \longrightarrow \quad p_{\text{hist}}(a)$$

$$\hat{\rho} = \sum_{a, a'} T_{a, a'}^{-1} p_{\text{hist}}(a') M^{(a)}$$

- **Issues:** Exponential scaling both in the representation and inversion problem
- Potentially unphysical density matrices $\hat{\rho}$

TRADITIONALLY QST REQUIRES EXPONENTIAL RESOURCES

Examples

- Maximum likelihood estimation. Requires an explicit “physical” density matrix  representation scales poorly

$$L(\hat{\rho}) = \prod_a P(a)^{f_a}$$

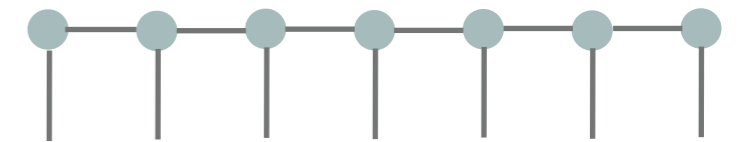
Maximize probability of observed data with respect to a parametrization of $\hat{\rho}$

- **Issues:** Exponential scaling in the parametrization
- Estimation of errors due to finite statistics in the measurements is difficult

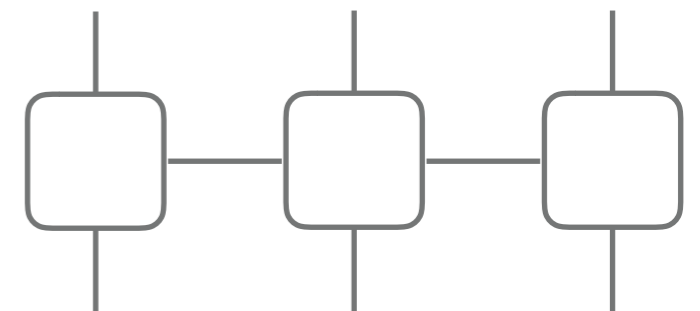
HOW TO MAKE QST EFFICIENT?

- Introduce a parametrization of the quantum state with good scaling if non-trivial structural information on the quantum systems under consideration is utilized: [MPS\[1\]](#) and [MPO\[2\]](#) tomography

[1] **Efficient quantum state tomography.** Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin & Yi-Kai Liu. Nature Communications volume 1, Article number: 149



[2] **A scalable maximum likelihood method for quantum state tomography** T Baumgratz¹, A Nüßeler, M Cramer and M B Plenio. New Journal of Physics, Volume 15, December 2013

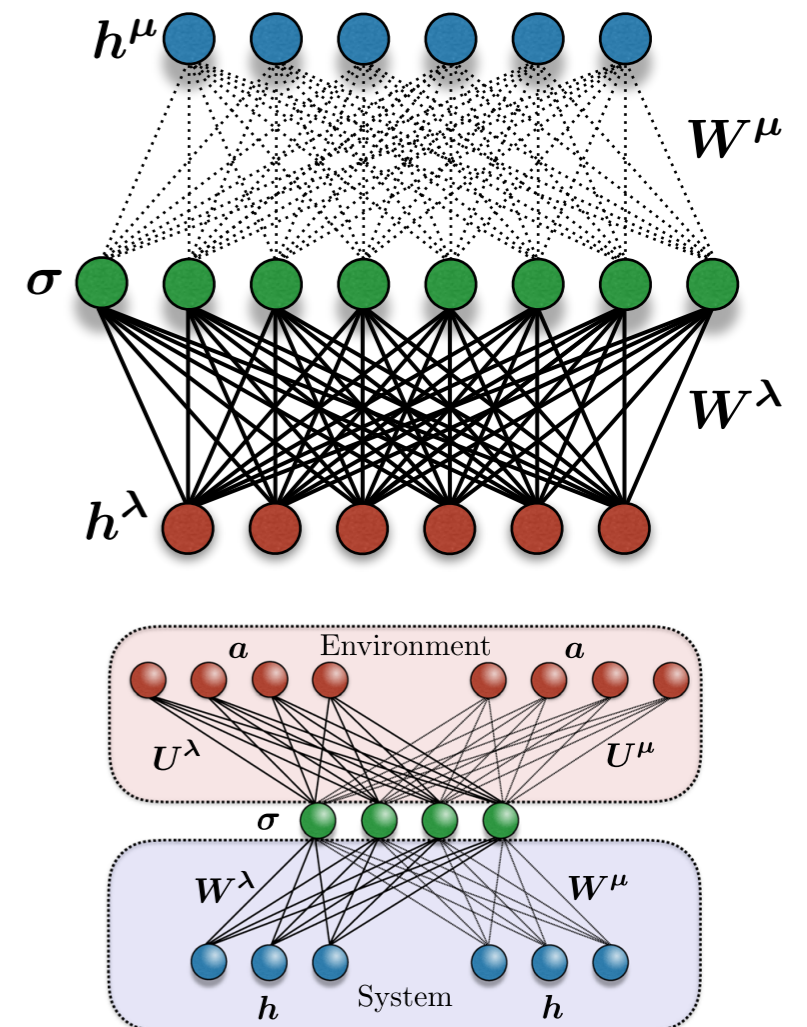


HOW TO MAKE QST EFFICIENT?

- Introduce a parametrization of the quantum state with good scaling if non-trivial structural information on the quantum systems under consideration is utilized: **Restricted Boltzmann machines both for pure[3] states and mixed states[4]**

[3] Neural-network quantum state tomography. G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. 14, 447 (2018).

[4] Latent Space Purification via Neural Density Operators. Giacomo Torlai and Roger G. Melko. Phys. Rev. Lett. 120, 240503 (2018)



HOW TO MAKE QST EFFICIENT?

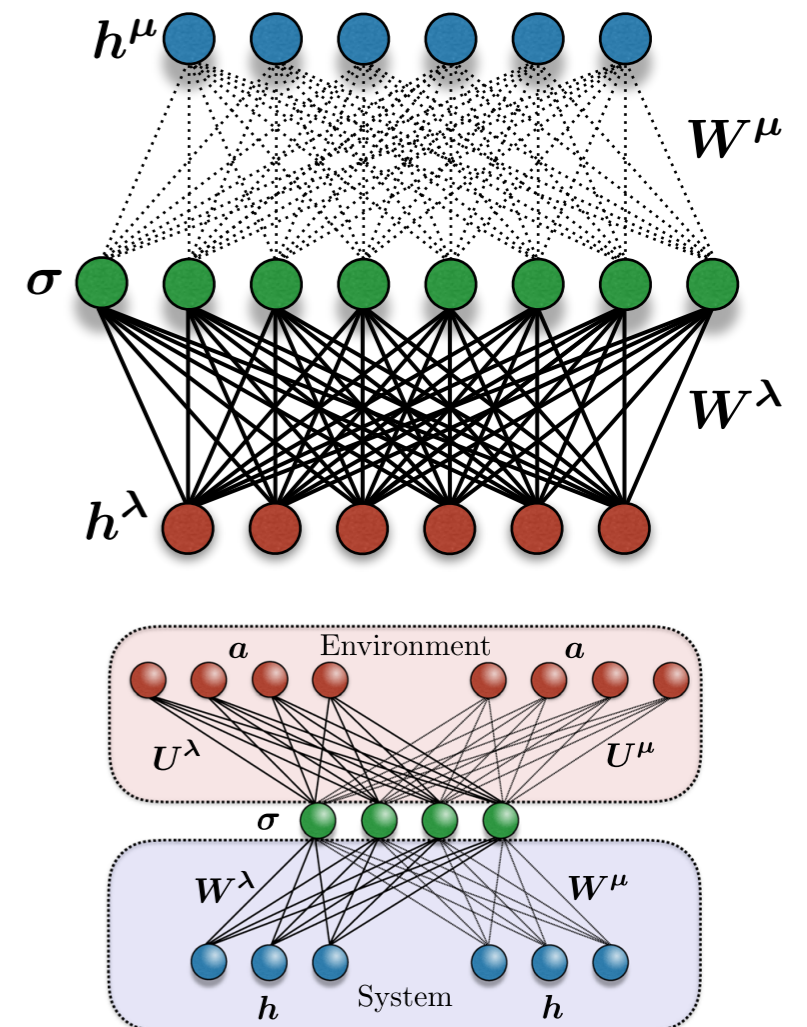
- Introduce a parametrization of the quantum state with good scaling if non-trivial structural information on the quantum systems under consideration is utilized: **Restricted Boltzmann machines both for pure[3] states and mixed states[4]**

[3] Neural-network quantum state tomography. G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. 14, 447 (2018).

Pure states with structure, good scaling

[4] Latent Space Purification via Neural Density Operators. Giacomo Torlai and Roger G. Melko. Phys. Rev. Lett. 120, 240503 (2018)*

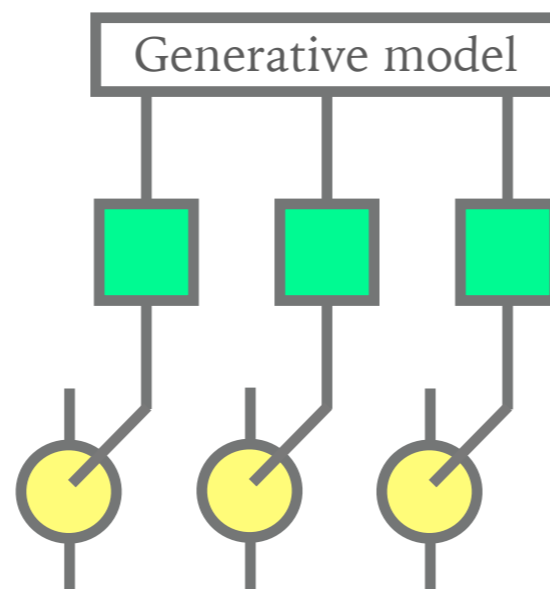
Mixed states with structure, bad scaling



IN THIS TALK

- ~~Parametrize the quantum state~~ We parametrize the measurement statistics of an informationally complete positive operator valued measure (POVM) which is given by Born's rule. We use **generative models**.
- Use this idea to learn states from synthetic measurements mimicking experimental data

$$\rho_{\text{model}} = (\mathbf{T}^{-1} P_{\text{model}})^T \mathbf{M}$$



IN THIS TALK

- ~~Parametrize the quantum state~~ We parametrize the measurement statistics of an informationally complete positive operator valued measure (POVM) which is given by Born's rule. We use **generative models**.
- Use this idea to learn states from synthetic measurements mimicking experimental data

Mixed states with structure, good scaling

SETTING

- We want to know if a large quantum device works as intended
- We think this device can produce nontrivial quantum states
- We want to certify that the system works in some “simple” cases (like MPS)
- This is only useful in the near term devices

MEASUREMENTS: POSITIVE OPERATOR-VALUED MEASURE (POVM)

- POVM elements $\mathbf{M} = \{M^{(a)} \mid a \in \{1, \dots, m\}\}$
- Positive semidefinite operators
- $\sum_i M^{(a)} = \mathbb{1}$

EXAMPLE: MEASUREMENT IN THE COMPUTATIONAL BASIS

$$M^{(0)} = |0\rangle\langle 0| \quad M^{(1)} = |1\rangle\langle 1| \quad \mathbb{I} = |0\rangle\langle 0| + |1\rangle\langle 1|$$

STATE $|\Psi\rangle = a|0\rangle + b|1\rangle$

$$\left. \begin{aligned} p(0) &= \langle \Psi | M^{(0)} | \Psi \rangle = |a|^2 \\ p(1) &= \langle \Psi | M^{(1)} | \Psi \rangle = |b|^2 \end{aligned} \right\}$$

BORN RULE: PROVIDES A LINK BETWEEN QUANTUM THEORY AND EXPERIMENT

MEASUREMENTS: POSITIVE OPERATOR-VALUED MEASURE (POVM)

POVM ELEMENTS $\mathbf{M} = \{M^{(a)} \mid a \in \{1, \dots, m\}\} \quad \sum_i M^{(a)} = \mathbb{1}$

$P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}}$ Born rule. Defines a distribution over the generalized measurements => link between quantum theory and experimental outcome

INFORMATIONALLY COMPLETE POVM

- The measurement statistics $P(\mathbf{a})$ contains all of the information about the state.
- If m is at least $D^2 = 2^{2N}$ and M **span** the entire Hilbert space

$$O = \sum_{\mathbf{a}} O(\mathbf{a}) M^{(\mathbf{a})}$$

- Relation between ρ and distribution $P(\mathbf{a})$ can be inverted.

MEASUREMENTS: POSITIVE OPERATOR VALUED MEASURES (POVM)

CONSTRUCTING POVMS: TAKE A SINGLE QUBIT POVM AND MAKE A TENSOR PRODUCT OF MANY

Pauli measurement for one qubit

$$\begin{aligned} \mathbf{M}_{\text{Pauli}} := & \{ M^{(0)} := p(3) \times |0\rangle\langle 0|, M^{(1)} := p(3) \times |1\rangle\langle 1|, \\ & M^{(+)} := p(1) \times |+\rangle\langle +|, M^{(-)} := p(1) \times |-\rangle\langle -|, \\ & M^{(r)} := p(2) \times |r\rangle\langle r|, M^{(l)} := p(2) \times |l\rangle\langle l| \}, \end{aligned}$$

For multiqubit systems $\mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}}$

we will consider

$$\mathbf{M} = \{ M^{(a_1)} \otimes M^{(a_2)} \otimes \dots \otimes M^{(a_N)} \}_{a_1, \dots, a_N}$$

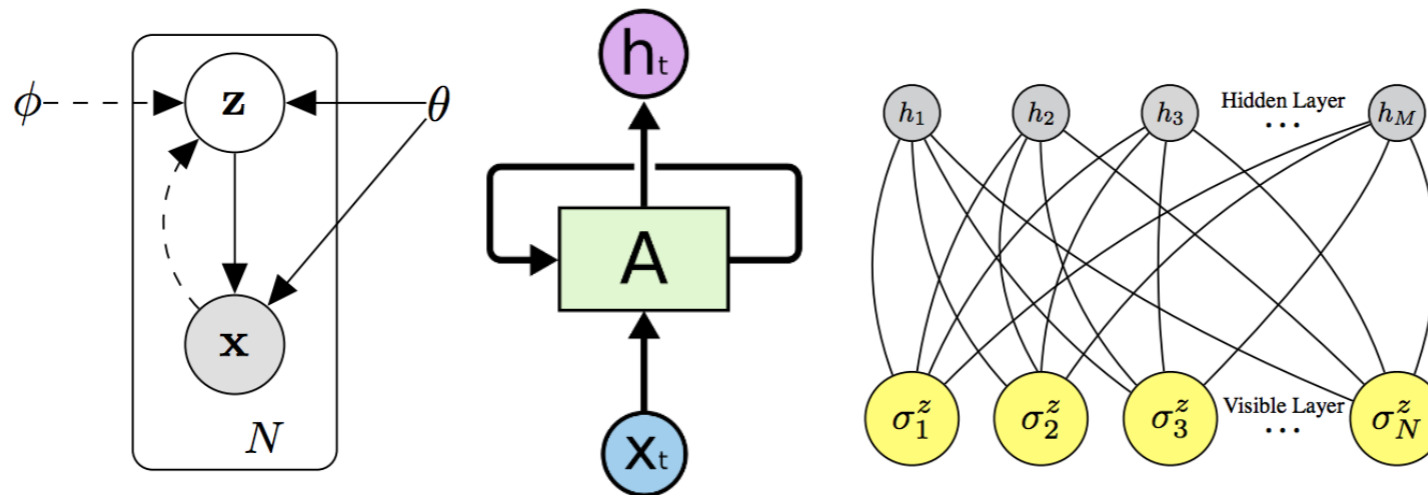
Experimental realization: pick a random direction with probability 1/3, then measure in that direction

Easy to implement in gate-based QC (Qiskit, Cirq, Rigetti, etc.)

KEY INSIGHT: PARAMETRIZE STATISTICS AND INVERT

$$P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}} \quad \Rightarrow \text{Unsupervised learning of } P(\mathbf{a})$$

$P_{\text{model}}(\mathbf{a}) \longrightarrow$ VAE, RBM, autoregressive models, etc



Goodfellow et al. Deep Learning (2016)

WHAT ARE THESE MODELS TYPICALLY USED FOR?

$P_{\text{model}}(\mathbf{a}) \longrightarrow$ VAE, RBM, GAN, autoregressive models, etc

- Understand probability distributions defined over high-dimensional data
- Language translation $P(\text{English}|\text{French})$
- Image generation
- Density estimation
- Anomaly/novelty detection



Source	When asked about this, an official of the American administration replied: "The United States is not conducting electronic surveillance aimed at offices of the World Bank and IMF in Washington."
PBMT	Interrogé à ce sujet, un responsable de l'administration américaine a répondu : "Les Etats-Unis n'est pas effectuer une surveillance électronique destiné aux bureaux de la Banque mondiale et du FMI à Washington".
GNMT	Interrogé à ce sujet, un fonctionnaire de l'administration américaine a répondu: "Les États-Unis n'effectuent pas de surveillance électronique à l'intention des bureaux de la Banque mondiale et du FMI à Washington".
Human	Interrogé sur le sujet, un responsable de l'administration américaine a répondu: "les Etats-Unis ne mènent pas de surveillance électronique visant les sièges de la Banque mondiale et du FMI à Washington".

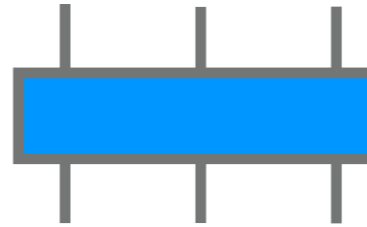
<https://arxiv.org/pdf/1812.04948.pdf>

<https://thispersondoesnotexist.com/>

PENROSE GRAPHICAL NOTATION

Density matrix

ρ

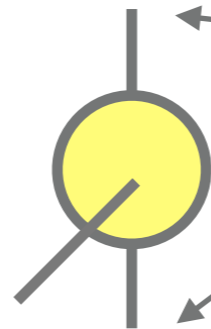
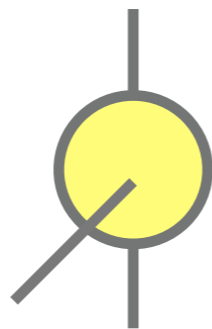
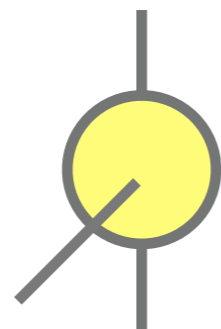


Statistics of measurement

$P(\mathbf{a})$



M



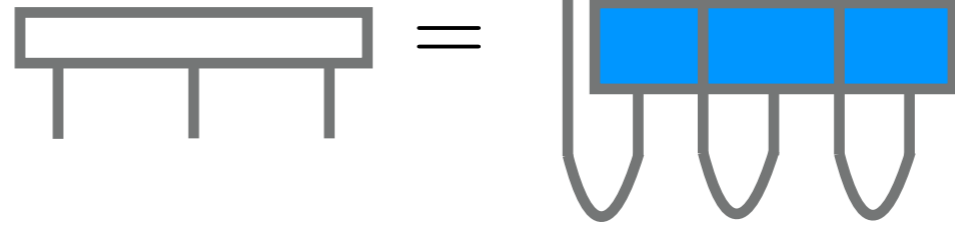
Physical indices

POVM outcome indices

$$M = \{ M^{(a_1)} \otimes M^{(a_2)} \otimes \dots \otimes M^{(a_N)} \}_{a_1, \dots, a_N}$$

GRAPHICAL NOTATION AND INVERSE

Born rule $P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}}$



If the POVM is informationally complete then

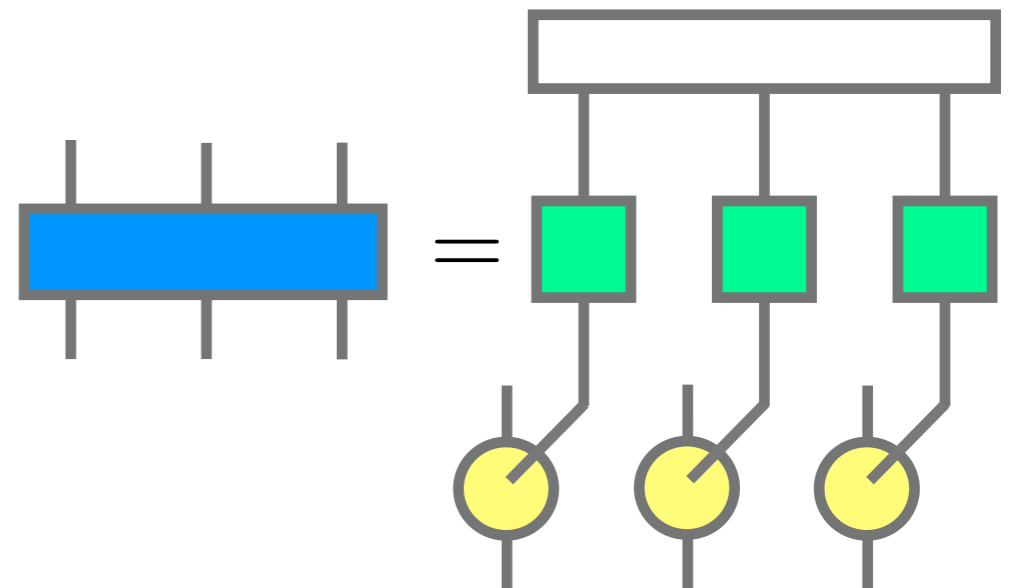
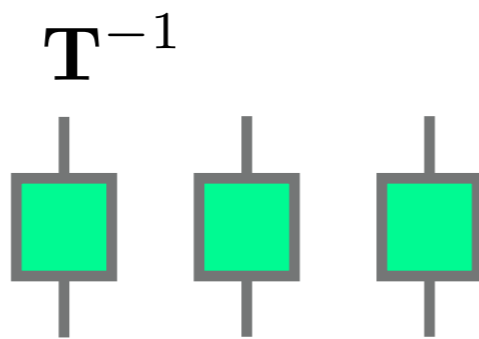
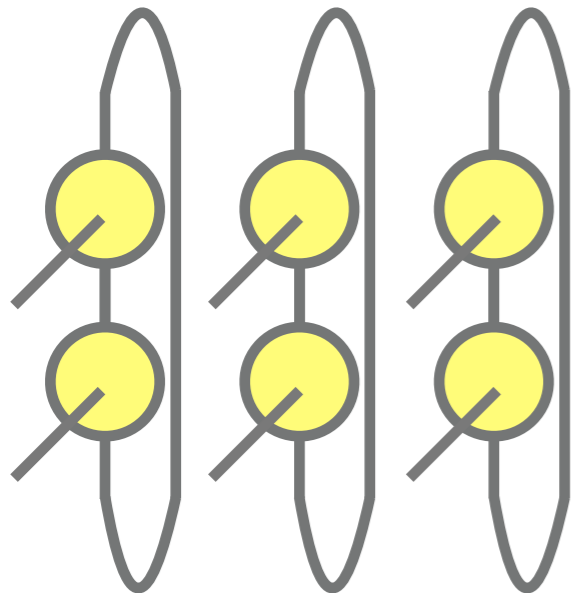
$$\rho = \sum_a O_\rho(a) M^{(a)}$$

Insert this relation into Born's rule $P(a) = \sum_{a'} O_\rho(a') \text{Tr}[M^{(a)} M^{(a')}] = \sum_{a'} O_\rho(a') T_{a'a}$

$$\rho = \sum_{a,a'} T_{a,a'}^{-1} P(a') M^{(a)}$$

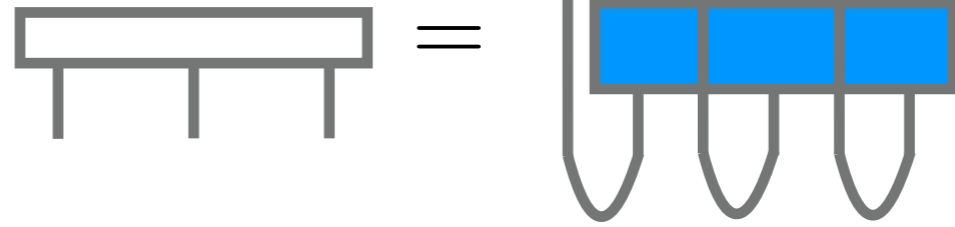
$$\rho = \sum_{a,a'} T_{a,a'}^{-1} P(a') M^{(a)}$$

$$T_{\alpha,\beta} = \text{Tr } M^\alpha M^\beta$$



GRAPHICAL NOTATION AND INVERSE

Born rule $P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}}$



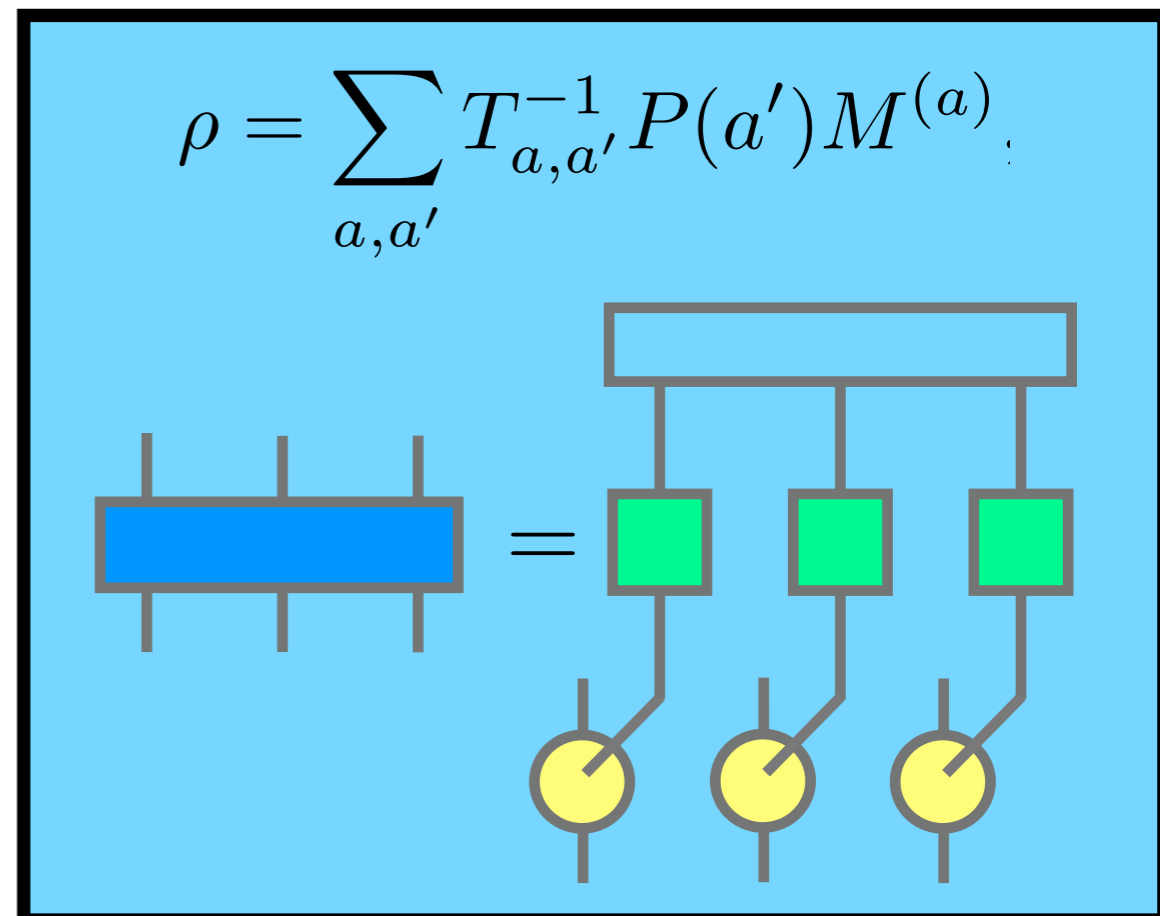
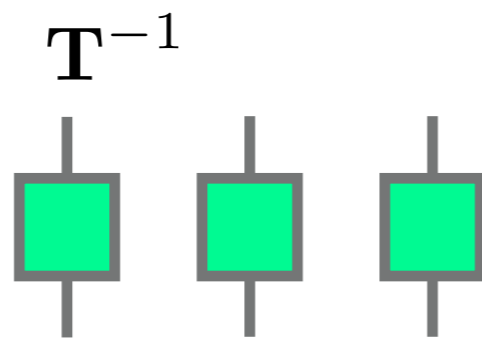
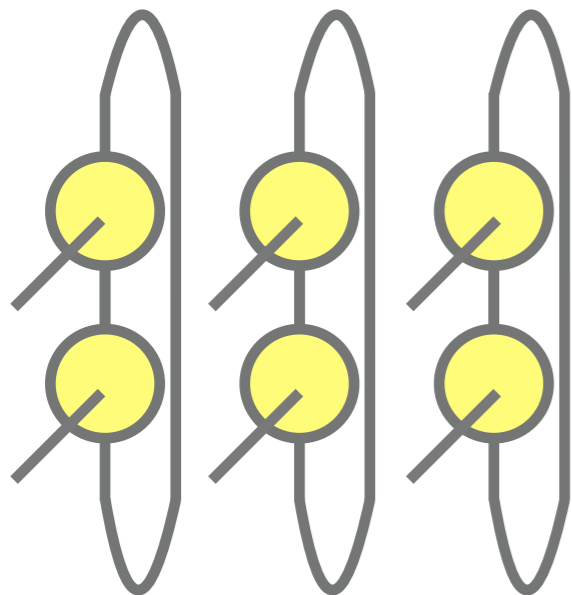
If the POVM is informationally complete then

$$\rho = \sum_a O_\rho(a) M^{(a)}$$

Insert this relation into Born's rule $P(a) = \sum_{a'} O_\rho(a') \text{Tr}[M^{(a)} M^{(a')}] = \sum_{a'} O_\rho(a') T_{a'a}$

$$\rho = \sum_{a,a'} T_{a,a'}^{-1} P(a') M^{(a)}$$

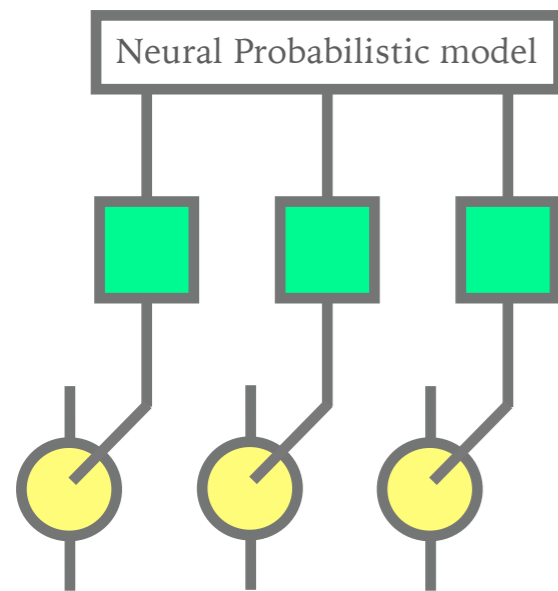
$$T_{\alpha,\beta} = \text{Tr } M^\alpha M^\beta$$



LEARNING MIXED STATES USING POSITIVE OPERATOR VALUED MEASURE

$$\rho_{\text{model}} = (\mathbf{T}^{-1} P_{\text{model}})^T \mathbf{M}$$

The statistics of the measurements is given by **Born rule**

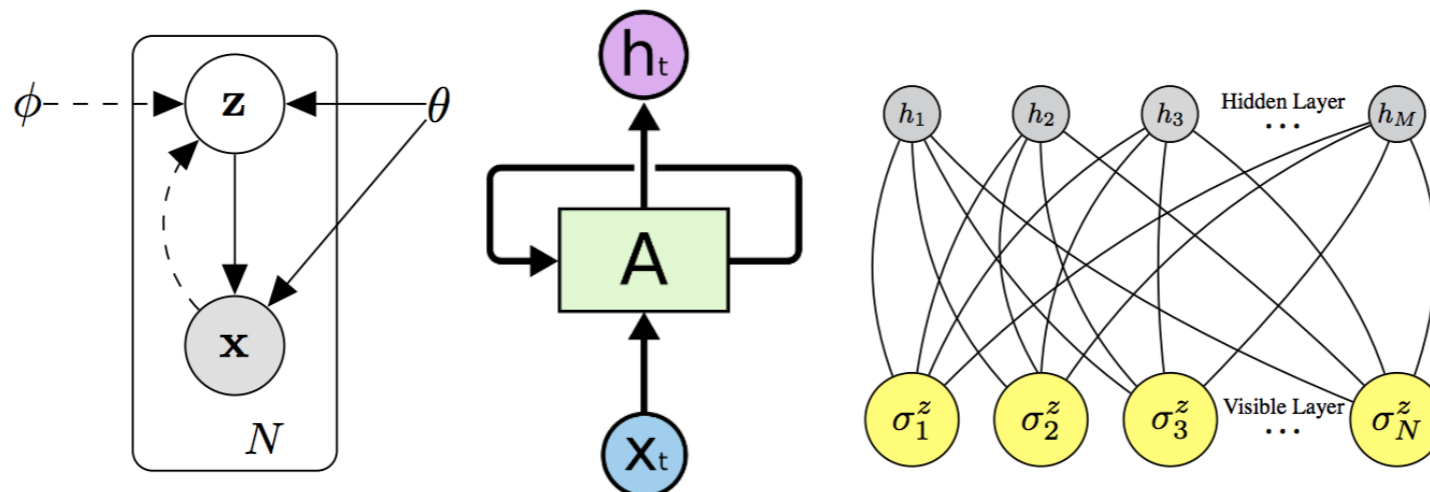


$$P(\mathbf{a}) = \text{Tr} \rho M^{\mathbf{a}}$$

↖ *Statistics of the measurements*
↗ *Measurements*
↘ *quantum state*

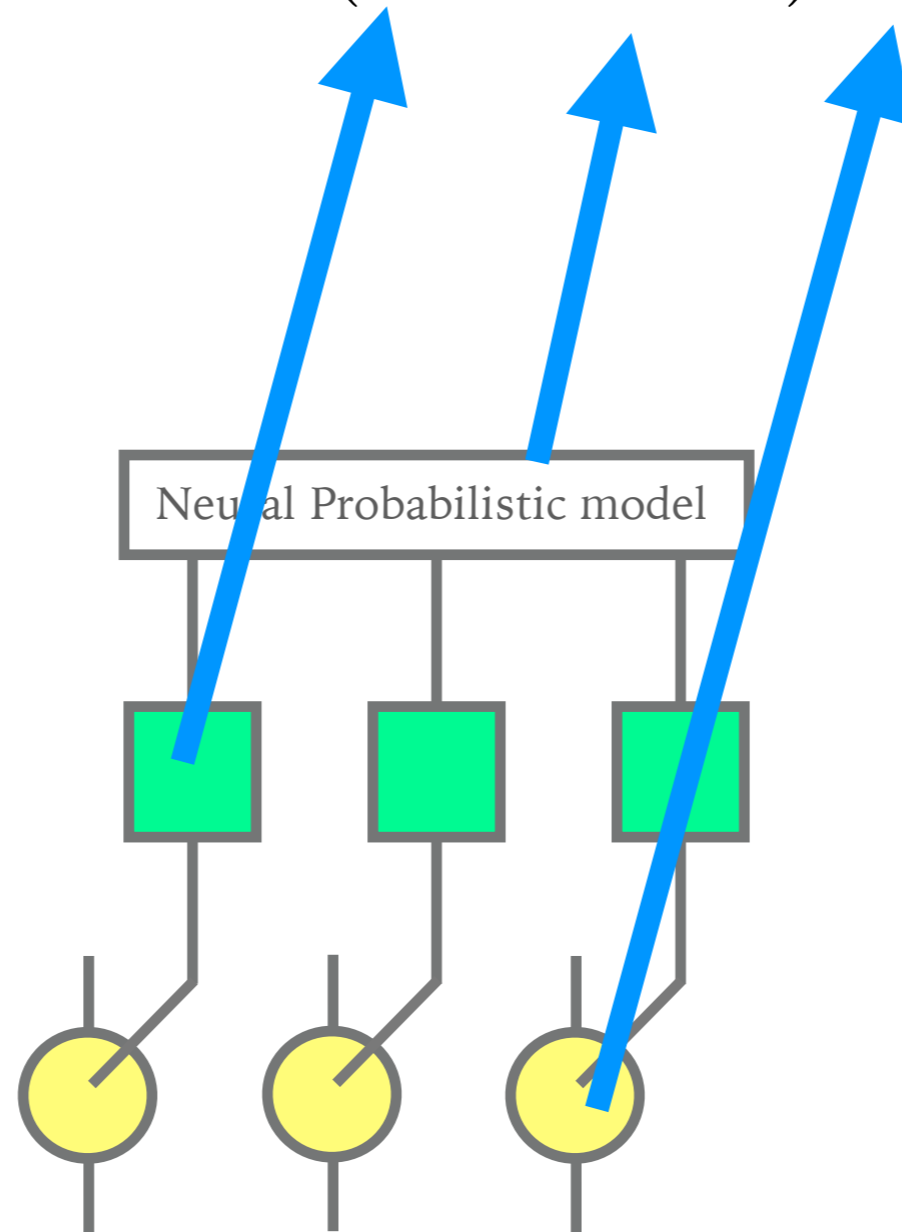
=> Unsupervised learning of P(a)

$P_{\text{model}}(\mathbf{a}) \longrightarrow$ VAE, RBM, GAN, autoregressive models, any model



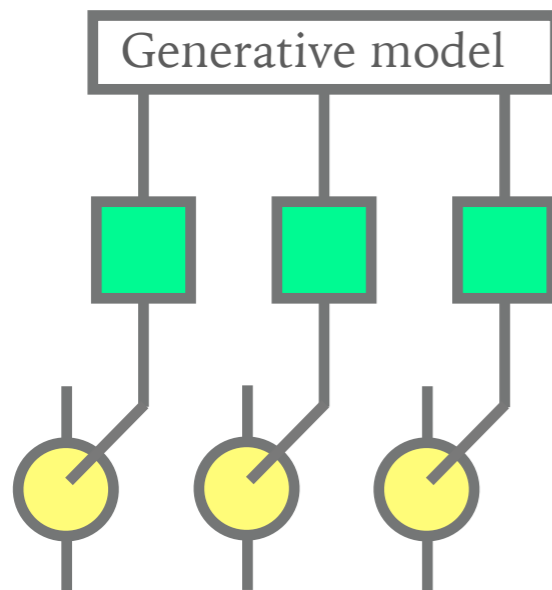
BORN RULE IS A LINEAR RELATION AND CAN BE INVERTED

$$\rho_{\text{model}} = (\mathbf{T}^{-1} P_{\text{model}})^T \mathbf{M}$$



MODEL FOR THE DENSITY MATRIX

$$\rho_{\text{model}} = (\mathbf{T}^{-1} P_{\text{model}})^T \mathbf{M}$$



- Factorization of the state in terms of a probability distribution and a set of tensors
- All the entanglement and potential complexity of the state comes from the structure of the $P(a)$
- $P(a)$ is approximated by powerful neural network probabilistic models

CERTIFICATION: FIDELITY, CLASSICAL FIDELITY, CORRELATION FUNCTIONS

- Fidelity
- Classical fidelity in POVM space
- Kullback–Leibler divergence in POVM space
- Correlation functions

$$F(\rho, \sigma) = \text{Tr} \left[\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right]^2$$

$$KL(P_{\text{model}}|P) = - \sum_a P(a) \log \frac{P_{\text{model}}(a)}{P(a)}$$

$$F_{\text{Classical}} = \sum_a \sqrt{P(a) P_{\text{model}}(a)}$$

LEARNING MIXED STATES USING POVM: PROCEDURE

- Prepare a desired quantum state repeatedly
- Perform generalized measurements
- Reconstruct $P(a)$ using unsupervised learning
- “Invert” $\rho = \sum_{a,a'} T_{a,a'}^{-1} P(a') M^{(a)}$,
- Perform some sort of certification via fidelity or classical fidelity or measure correlation functions etc.

Combines the idea of MLE with linear inversion in one method (observables/certification)

RESULTS

**GHZ + NOISE, GROUND STATES OF QUANTUM SYSTEMS
APPROXIMATED BY MPS STATES**

RESULTS ON SYNTHETIC DATASETS FOR GHZ STATES

Pure GHZ

$$|\Psi_0\rangle \equiv \alpha |0\rangle^{\otimes N} + \beta |1\rangle^{\otimes N}$$

$$\rho_0 := |\Psi_0\rangle \langle \Psi_0|$$

$$= |\alpha|^2 |0\rangle \langle 0|^{\otimes N} + |\beta|^2 |1\rangle \langle 1|^{\otimes N} + \left(\alpha\beta^* |0\rangle \langle 1|^{\otimes N} + \text{h.c.} \right)$$

GHZ with

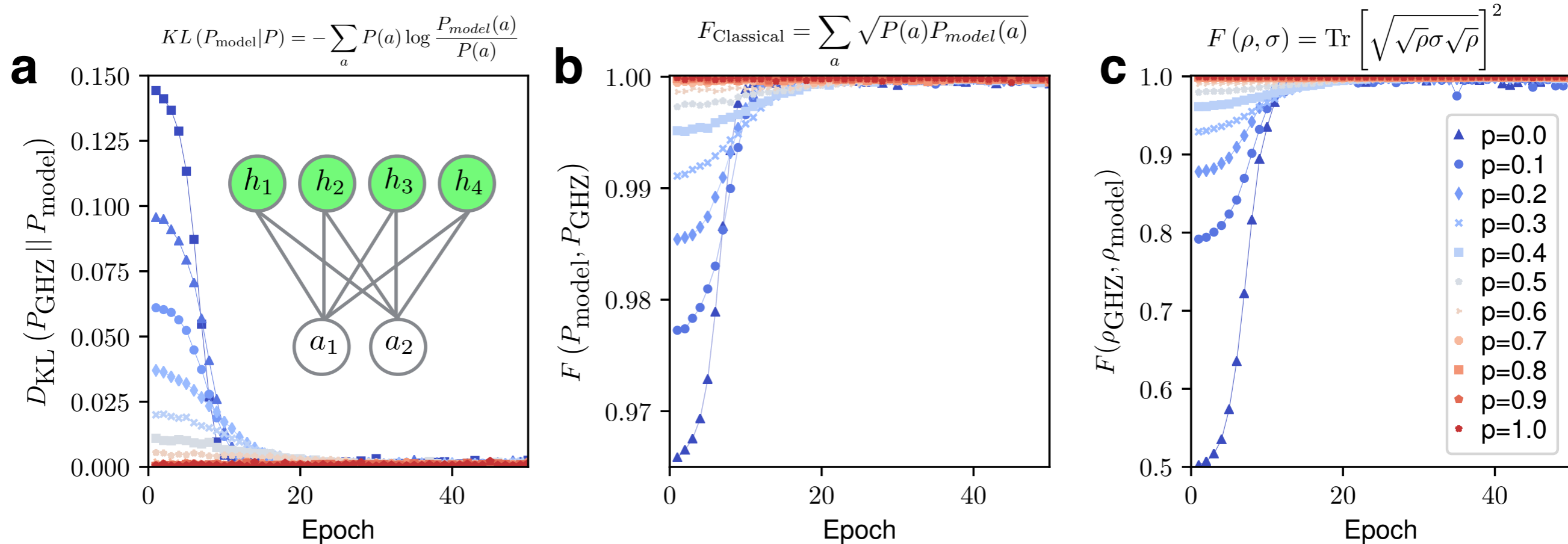
$$\mathcal{E}_i \rho_0 = (1 - p) \rho_0 + \frac{p}{3} \left(\sigma_i^{(1)} \rho_0 \sigma_i^{(1)} + \sigma_i^{(2)} \rho_0 \sigma_i^{(2)} + \sigma_i^{(3)} \rho_0 \sigma_i^{(3)} \right)$$

Locally depolarized

generalized GHZ

states : a model of a decohering qubit where with probability $1 - p$ the qubit remains intact, while with probability p an “error” occurs.

LEARNING 2 QUBIT LOCALLY DEPOLARIZED GHZ



This result is obtained by parametrizing the $P(a)$ with an RBM with multinomial visible units (4 states per qubit for the corresponding to the outcomes of the POVMs)

$$v_i = 0, 1, 2, 3, 4$$

$$h_j = 0, 1$$

RECURRENT NEURAL NETWORK MODEL AND RESULTS

GHZ results

40 qubit $p=0$

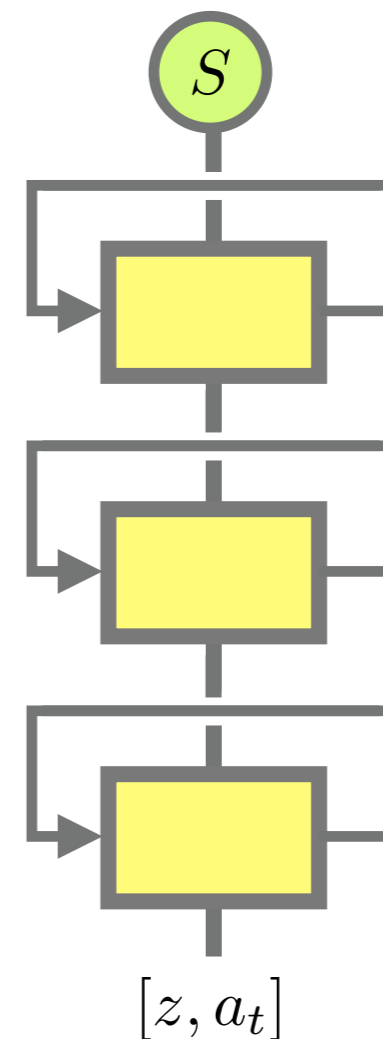
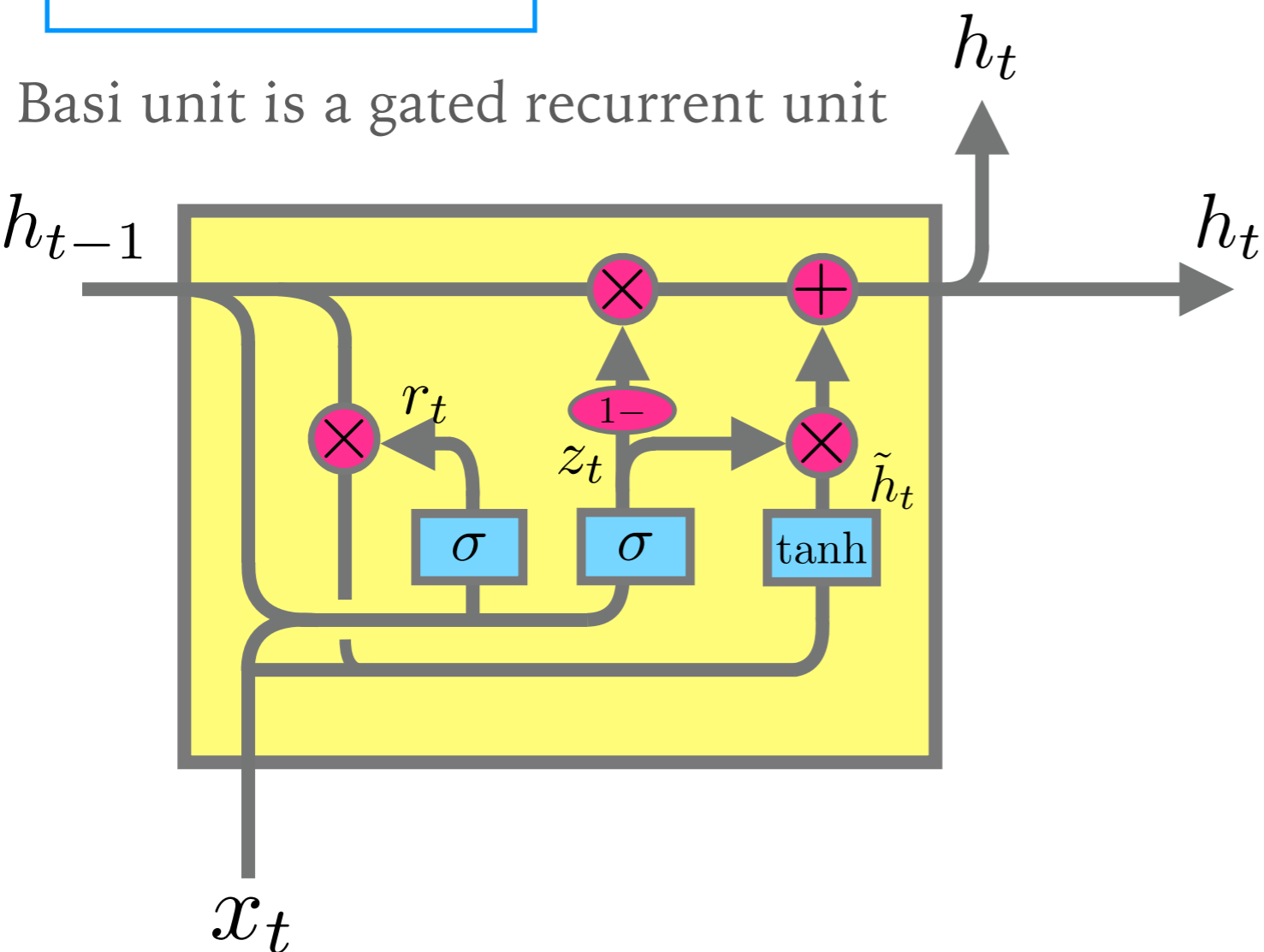
$F_c = 0.9992(4)$

80 qubits $p = 0.01$

$F_c = 0.9988(1)$

$KL = 0.0050(2)$

Full model stacks three of these units and adds a softmax dense layer at each time step



RECURRENT NEURAL NETWORK MODEL AND RESULTS

GHZ results

40 qubit $p=0$

$F_c = 0.9992(4)$

80 qubits $p = 0.01$

$F_c = 0.9988(1)$

Full model stacks three of these units and adds a softmax dense layer at each time step

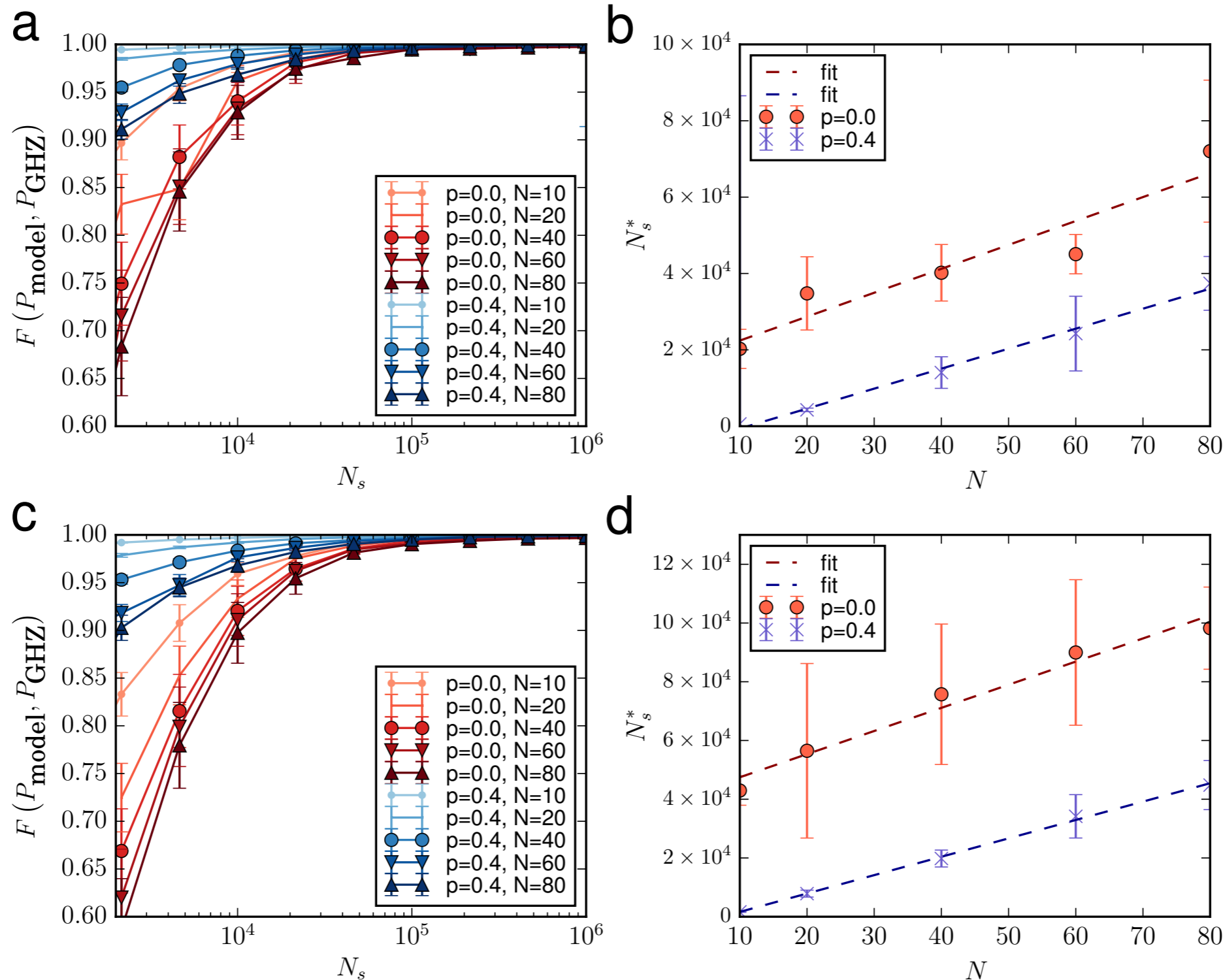
Using this model was important for the success of the method

- Tractable likelihood
- Exact sampling

x_t

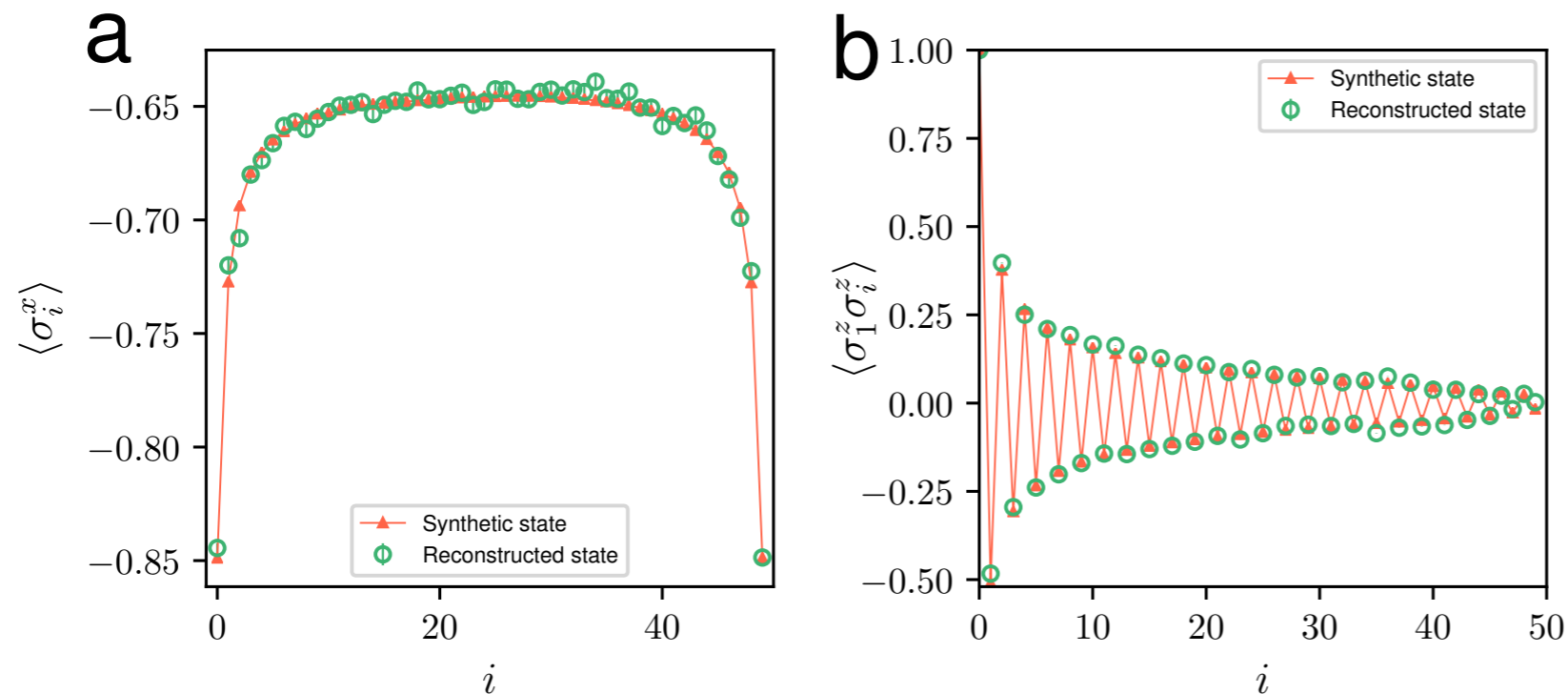
$[z, a_t]$

NUMERICAL INVESTIGATION OF THE SAMPLE COMPLEXITY OF LEARNING



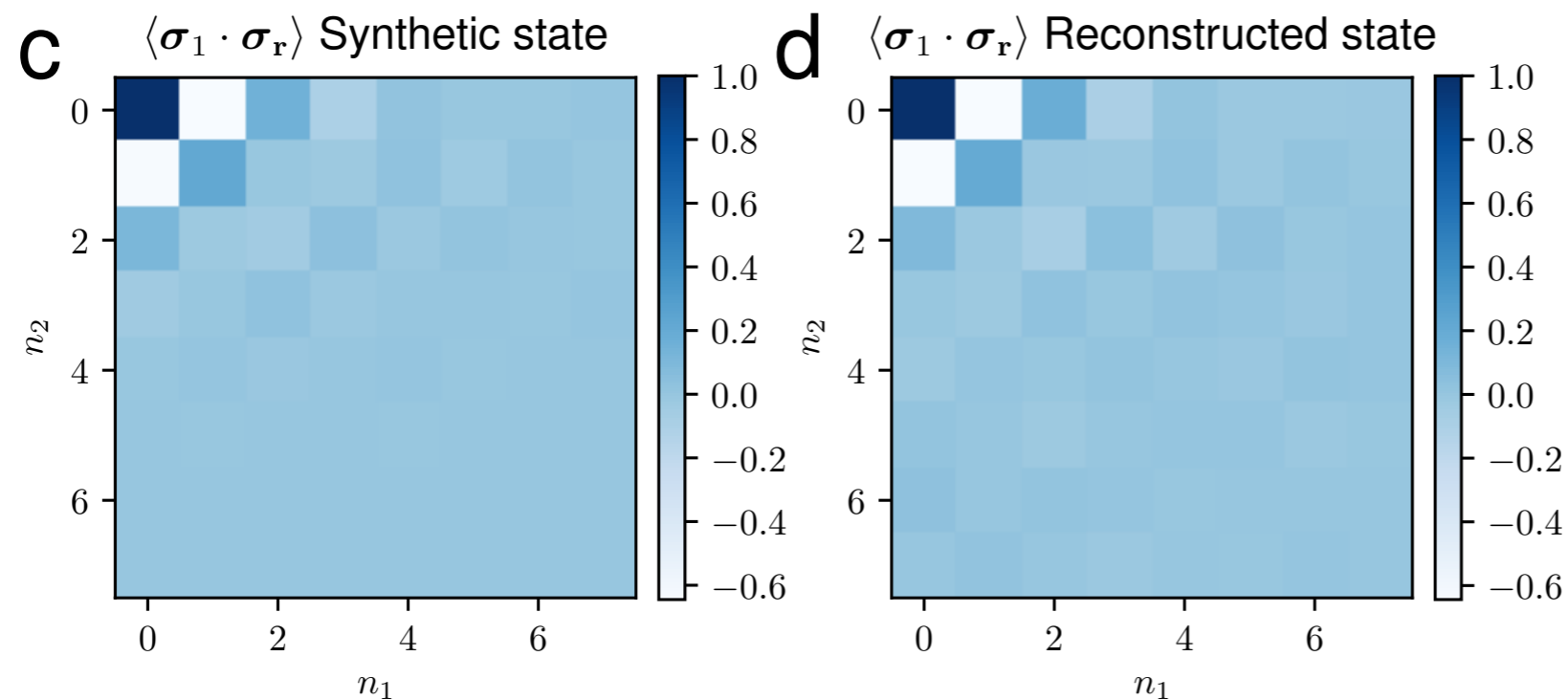
Favorable scaling. arXiv:1810.10584 Carrasquilla, Torlai, Melko, Aolita

GROUND STATES OF LOCAL HAMILTONIANS

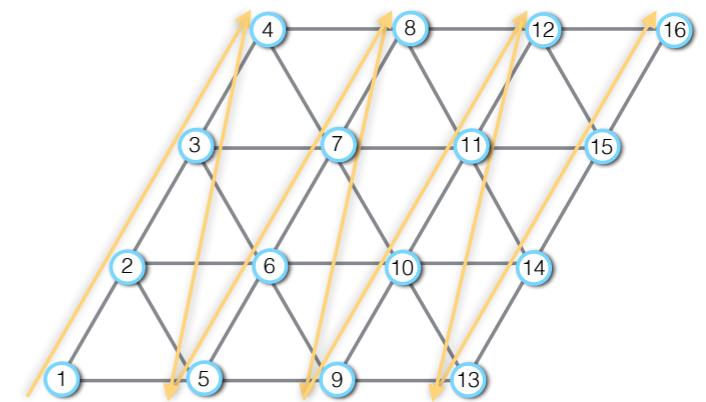


$$\mathcal{H} = J \sum_{ij} \sigma_i^z \sigma_j^z + h \sum_i \sigma_i^x$$

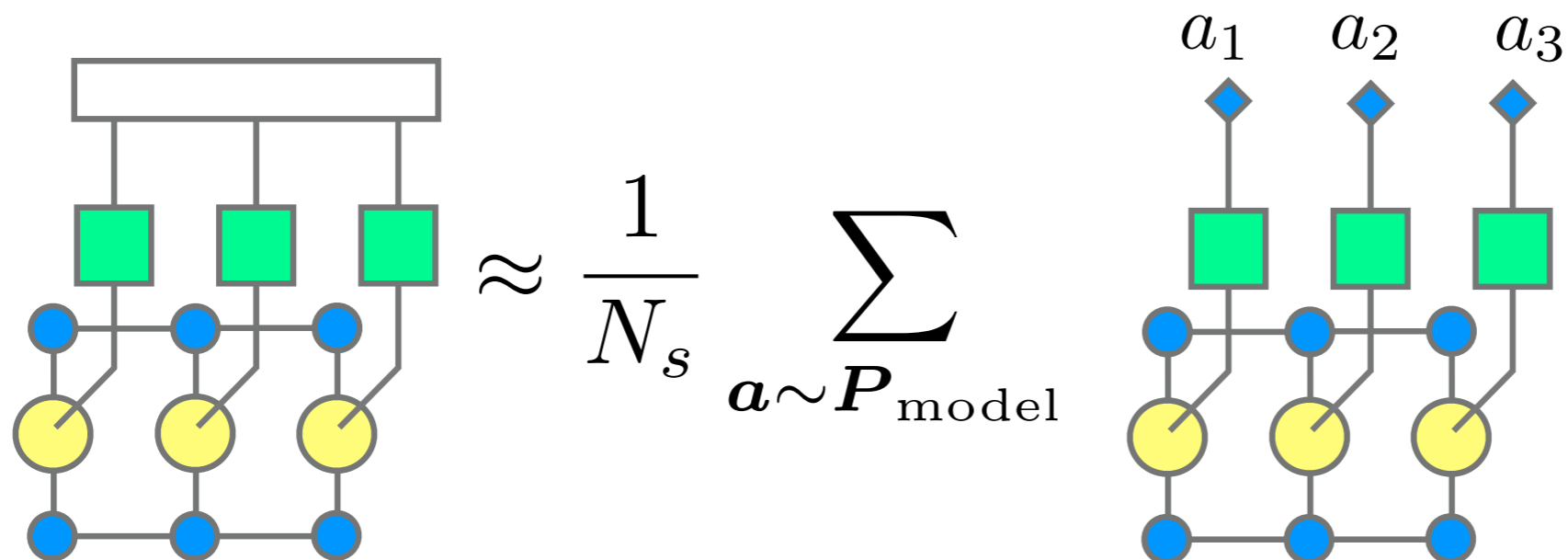
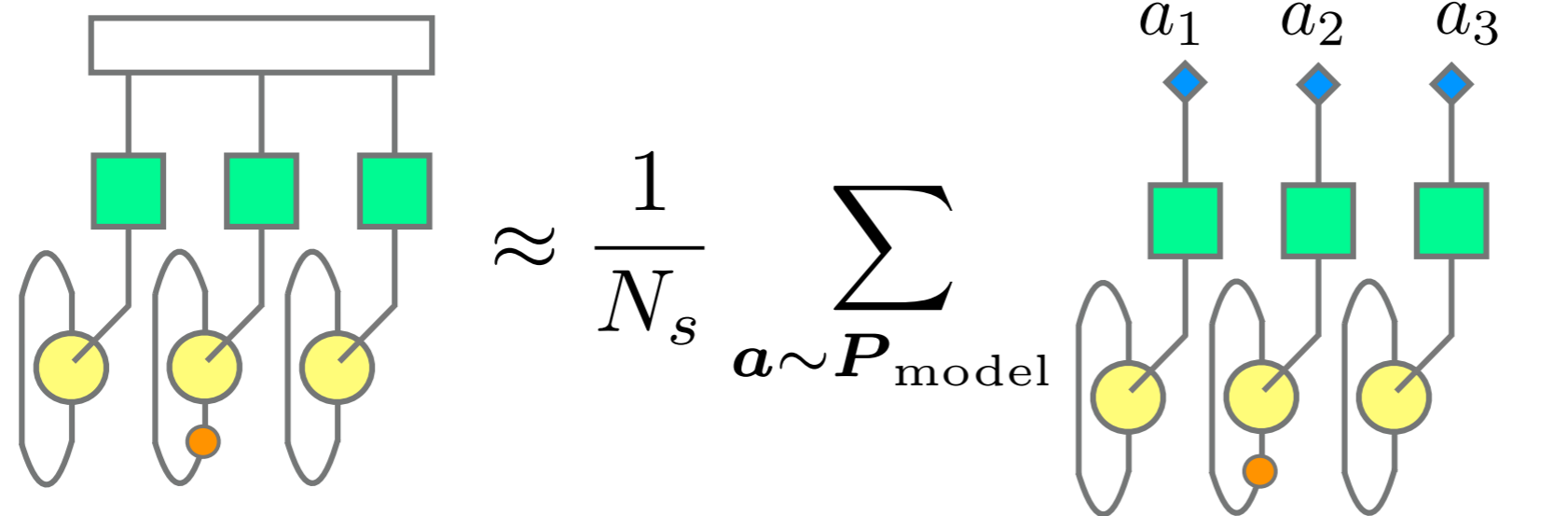
N=50 spins. P(a) is a deep (3 layer GRU) recurrent neural network language model.



$$H = J \sum_{i,j} \sigma_i \cdot \sigma_j$$

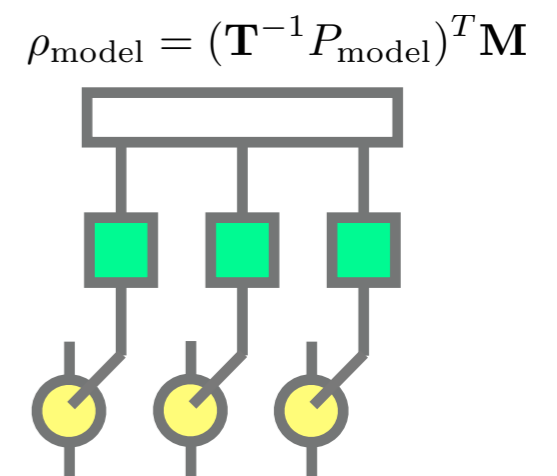


OBSERVABLES AND FIDELITY W.R.T. AN MPS



ONGOING

- Apply to an experimental reconstruction
- Simulate real-time dynamics
- Simulate quantum circuits
- Energy minimization
- But challenging regularization issues, positivity is challenging to impose in a scalable way (—>Andrew Tan's work)

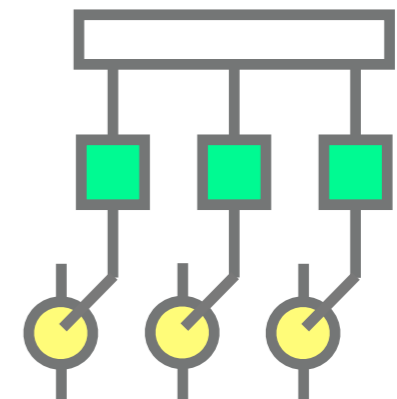


CONCLUSION

- We are starting to explore a machine learning perspective on many-body problems in classical and quantum physics.
- We have performed QST based on neural networks which enable us to study of 2-dimensional quantum systems
- I believe there are a lot of opportunities for us physicists.

⟨ **PHYSICS | MACHINE LEARNING** ⟩

$$\rho_{\text{model}} = (\mathbf{T}^{-1} P_{\text{model}})^T \mathbf{M}$$

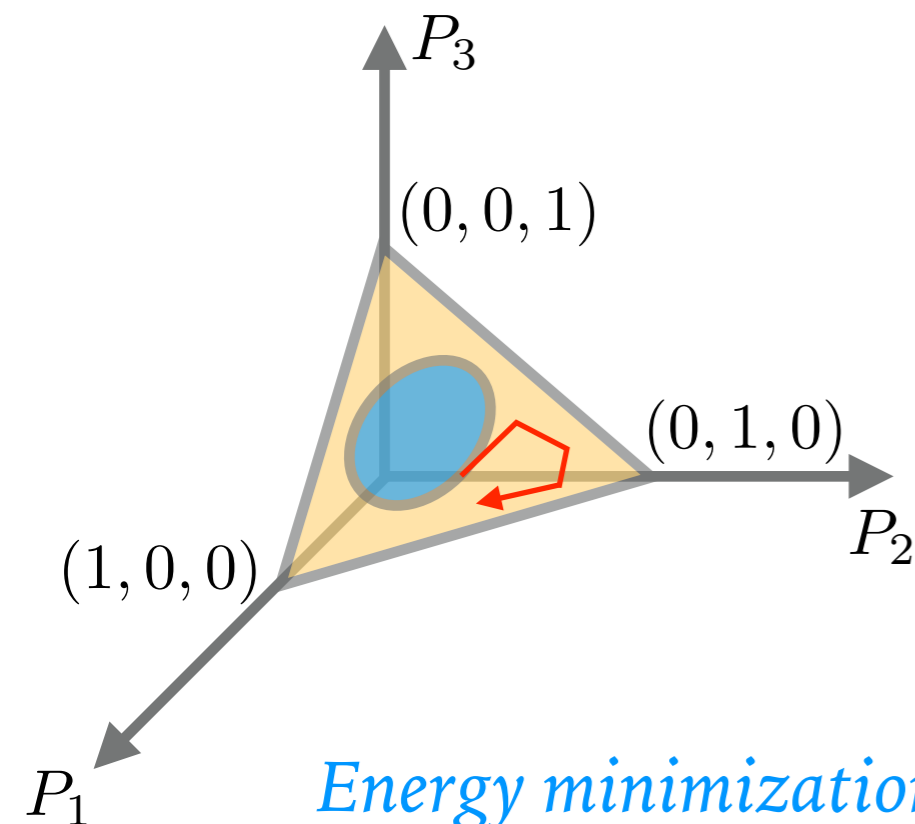


REGULARIZATION ISSUES

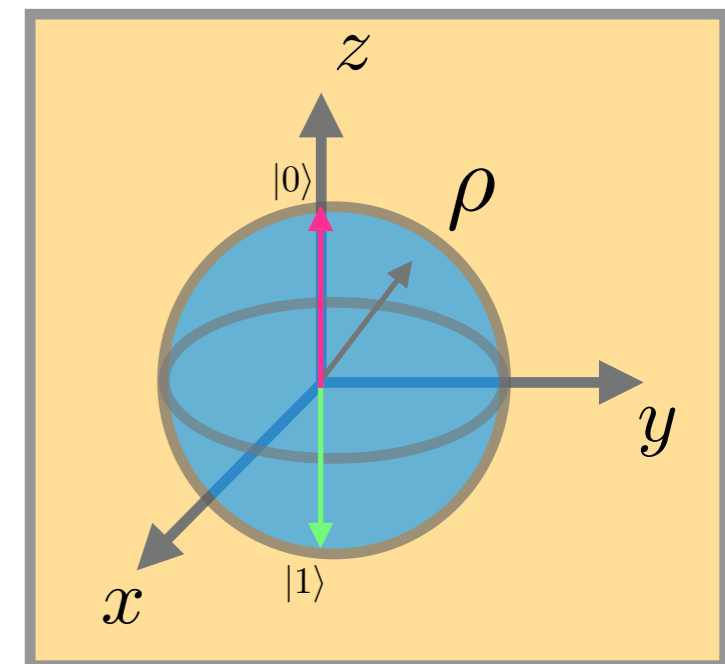
Requisite to specifying a quantum state through $P(a)$, one must have an understanding of the allowed probabilities since not every choice of $P(a)$ will give rise to positive semi-definite density matrix.

Requires regularization that is tricky to perform for large systems

For this tomography problem in particular we just use a lot of data



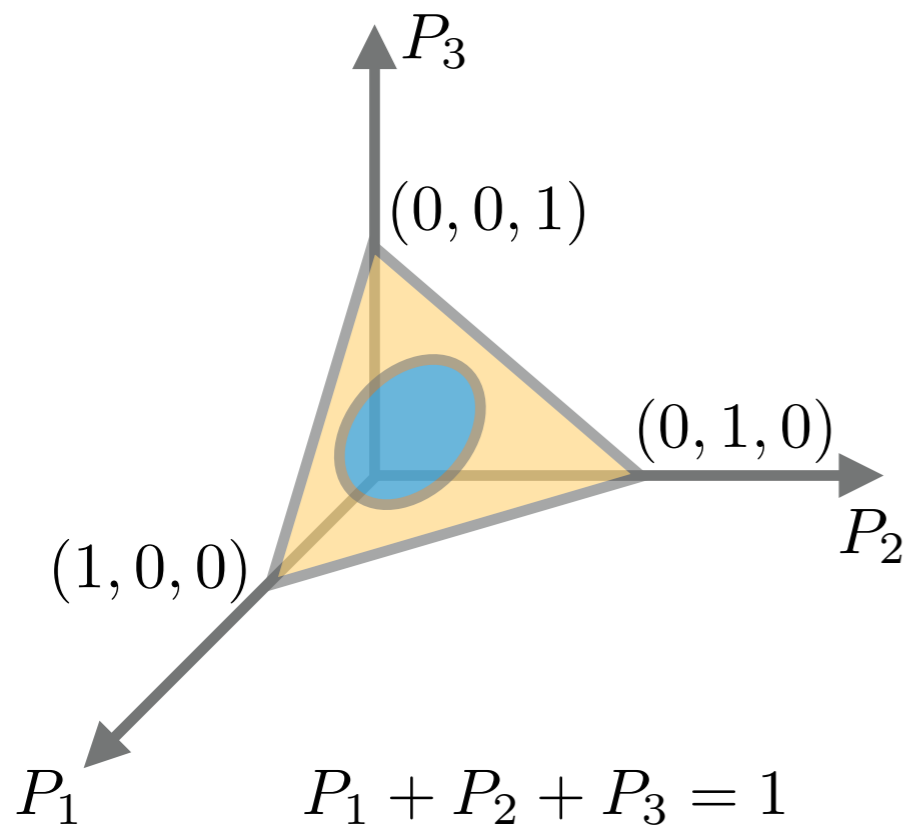
$$\rho = \sum_{a, a'} T_{a, a'}^{-1} P(a') M^{(a)}$$



Energy minimization is not variational. How to fix this?

THE STANDARD SIMPLEX AND QUANTUM STATES

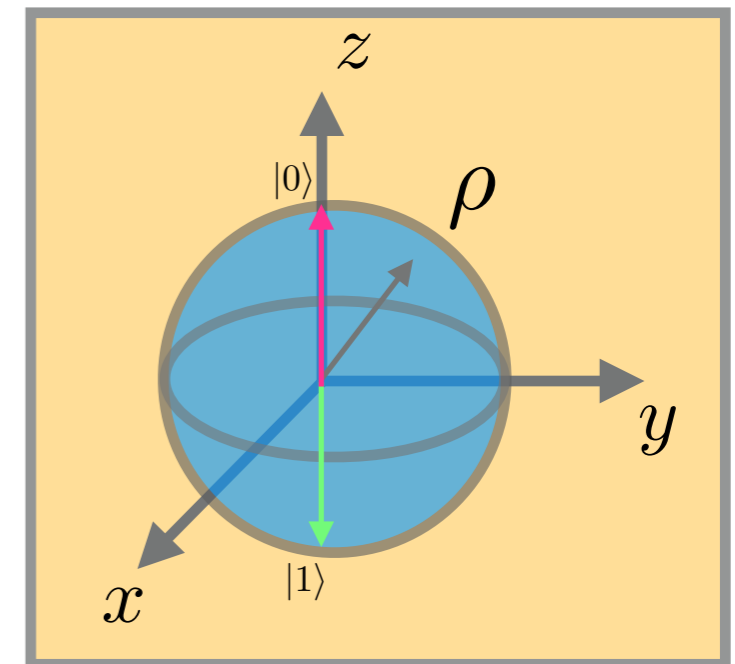
In probability, the points of the standard n -simplex in $(n + 1)$ -space are the space of possible parameters (probabilities) of the **categorical distribution** on $n + 1$ possible outcomes.



$$P(\mathbf{a}) = \text{Tr } \rho M^{\mathbf{a}}$$



$$\rho = \sum_{a, a'} T_{a, a'}^{-1} P(a') M^{(a)}$$



THE STANDARD SIMPLEX AND QUANTUM STATES

$$i \frac{\partial P(\mathbf{a}'')}{\partial t} = \sum_{\mathbf{a}, \mathbf{a}'} T_{\mathbf{a}, \mathbf{a}'}^{-1} P(\mathbf{a}') \text{Tr} \left([H, M^{(\mathbf{a}'')}] \right)$$

$$P_U(\mathbf{a}'') = \sum_{\mathbf{a}'} P(\mathbf{a}') O_{\mathbf{a}', \mathbf{a}''}$$

$$O_{\mathbf{a}', \mathbf{a}''} = \sum_{\mathbf{a}} \text{Tr}(U M^{(\mathbf{a})} U^\dagger M^{(\mathbf{a}'')}) T_{\mathbf{a}, \mathbf{a}'}^{-1}$$

Generalization of stochastic matrices:

Somewhat stochastic matrices:

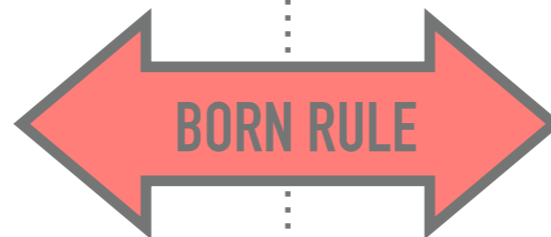
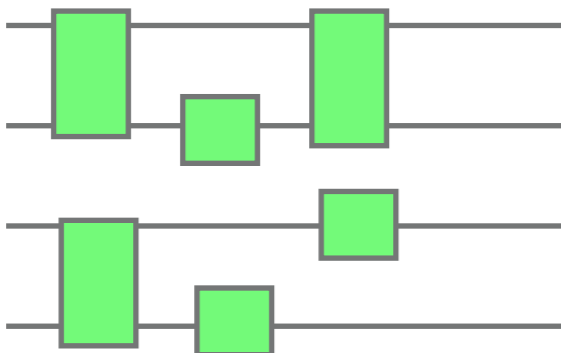
<https://arxiv.org/abs/0709.0309>

Somewhat stochastic matrices

Branko Ćurgus, Robert I. Jewett

(Submitted on 3 Sep 2007)

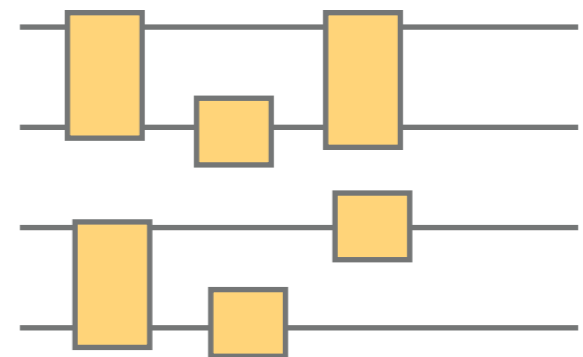
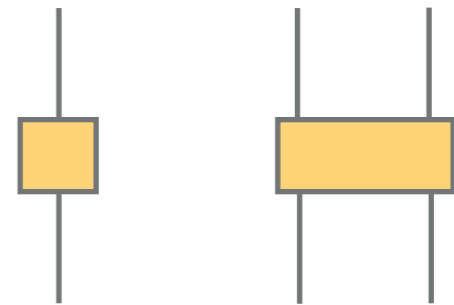
The standard theorem for regular stochastic matrices is generalized to matrices with no sign restriction on the entries. The condition that column sums be equal to 1 is kept, but the regularity condition is replaced by a condition on the ℓ_1 -distances between columns.



$$i \frac{\partial \rho}{\partial t} = [H, \rho]$$

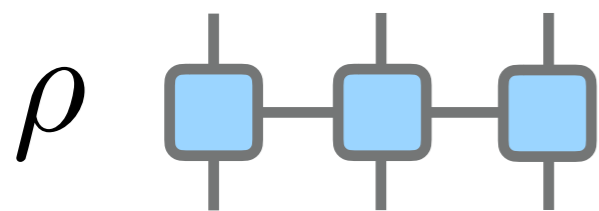
$$\rho_U = U \rho U^\dagger$$

Unitary matrices U

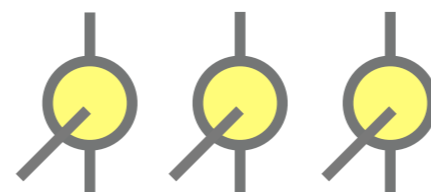


FOR LARGER GHZ STATES USE MPS/MPO

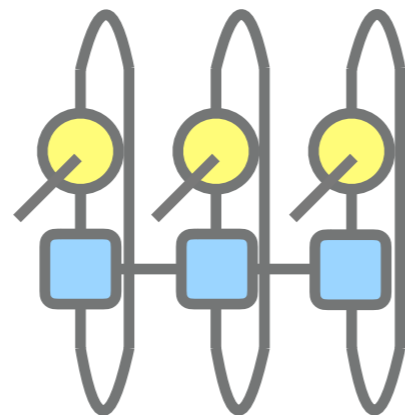
GENERATING **SYNTHETIC** DATA FOR BIG SYSTEMS EFFICIENTLY USING TENSOR NETWORKS



M

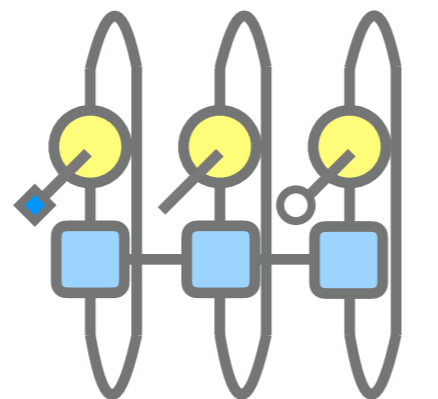


$$P(\mathbf{a}) = \text{Tr} \rho M^{\mathbf{a}}$$

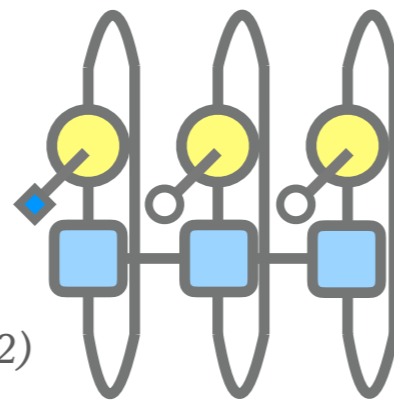


Can be sampled in polynomial time
exactly (zipper algorithm)

$$P(a_i | a_{<i}) = \frac{\sum_{a_{j>i}} P(\mathbf{a})}{\sum_{a_{j \geq i}} P(\mathbf{a})}$$



$$P(a_2 | a_1)$$



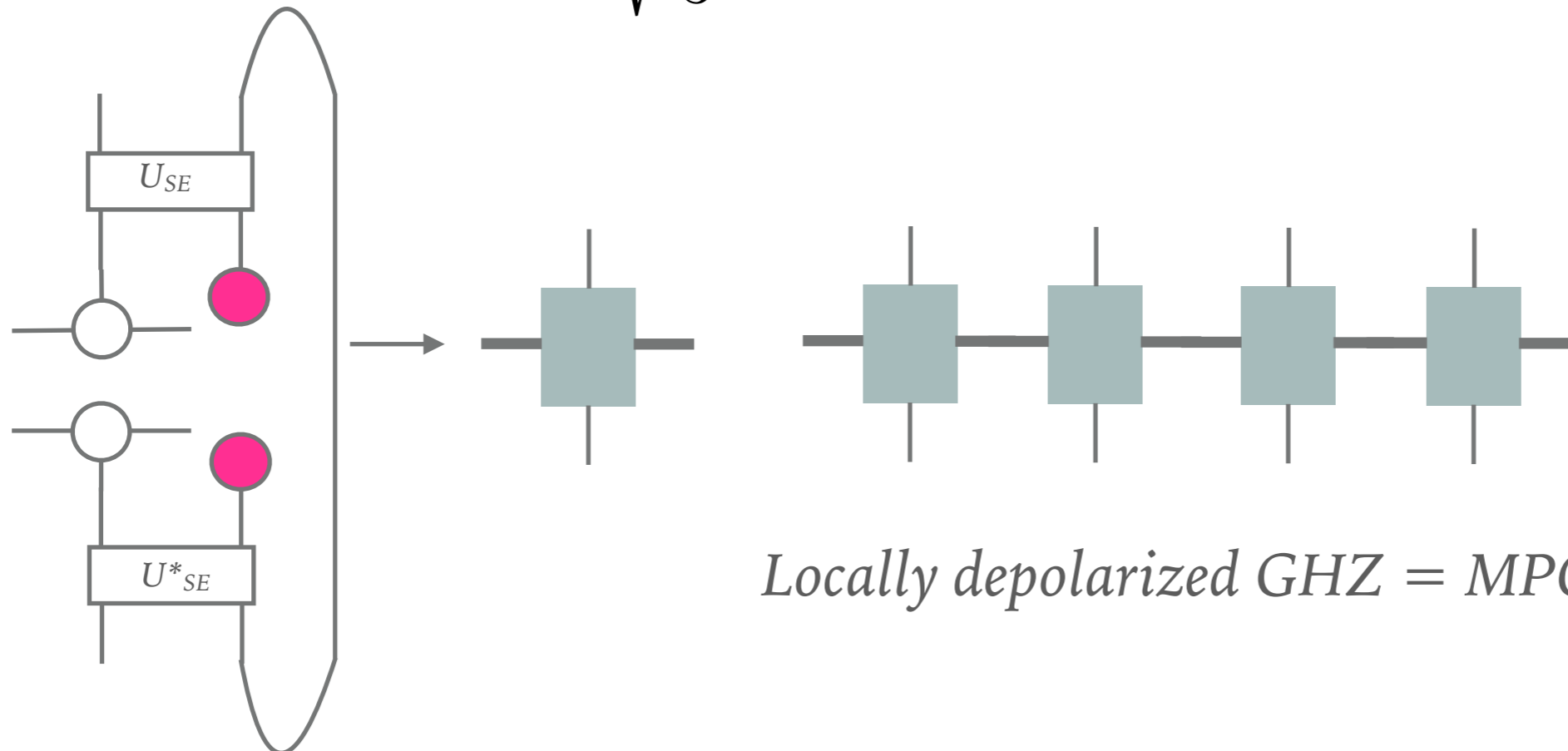
FOR LARGER GHZ STATES USE MPS/MPO

Pure GHZ = MPS


$$\begin{array}{c} 1 \\ \hline \bigcirc \\ \hline 0 \end{array} \begin{array}{c} 1 \\ \hline \end{array} = \begin{array}{c} 2 \\ \hline \bigcirc \\ \hline 1 \end{array} \begin{array}{c} 2 \\ \hline \end{array} = 2^{-1/(2N)}$$

Locally depolarized GHZ

$$U_{SE} = \mathbb{1}_S \otimes |0\rangle\langle 0|_E + \sqrt{\frac{p}{3}} (\sigma_S^x \otimes |1\rangle\langle 0|_E + \sigma_S^y \otimes |2\rangle\langle 0|_E + \sigma_S^z \otimes |3\rangle\langle 0|_E)$$



Locally depolarized GHZ = MPO

 $= (1 \ 0 \ 0 \ 0)$

DEPOLARIZING CHANNEL

3.4.1 Depolarizing channel

The *depolarizing channel* is a model of a decohering qubit that has particularly nice symmetry properties. We can describe it by saying that, with probability $1 - p$ the qubit remains intact, while with probability p an “error” occurs. The error can be of any one of three types, where each type of error is equally likely. If $\{|0\rangle, |1\rangle\}$ is an orthonormal basis for the qubit, the three types of errors can be characterized as:

1. Bit flip error: $\begin{matrix} |0\rangle \mapsto |1\rangle \\ |1\rangle \mapsto |0\rangle \end{matrix}$ or $|\psi\rangle \mapsto \sigma_1|\psi\rangle$, $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,
2. Phase flip error: $\begin{matrix} |0\rangle \mapsto |0\rangle \\ |1\rangle \mapsto -|1\rangle \end{matrix}$ or $|\psi\rangle \mapsto \sigma_3|\psi\rangle$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,
3. Both: $\begin{matrix} |0\rangle \mapsto +i|1\rangle \\ |1\rangle \mapsto -i|0\rangle \end{matrix}$ or $|\psi\rangle \mapsto \sigma_2|\psi\rangle$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$.

If an error occurs, then $|\psi\rangle$ evolves to an ensemble of the three states $\sigma_1|\psi\rangle$, $\sigma_2|\psi\rangle$, $\sigma_3|\psi\rangle$, all occurring with equal likelihood.

Unitary representation. The depolarizing channel mapping qubit A to A can be realized by an isometry mapping A to AE , where E is a four-dimensional environment, acting as

$$\begin{aligned} U_{A \rightarrow AE} : |\psi\rangle_A \mapsto & \sqrt{1-p} |\psi\rangle_A \otimes |0\rangle_E \\ & + \sqrt{\frac{p}{3}} (\sigma_1|\psi\rangle_A \otimes |1\rangle_E + \sigma_2|\psi\rangle_A \otimes |2\rangle_E + \sigma_3|\psi\rangle_A \otimes |3\rangle_E). \end{aligned} \tag{3.83}$$

NEUMARK'S DILATION THEOREM

How one can implement a general quantum measurement by performing a unitary on the system of interest and an ancilla, followed by a von Neumann measurement of the ancilla.

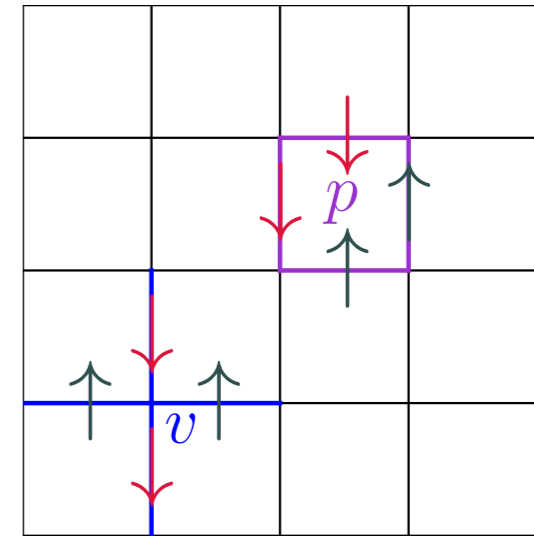
$$P(a) = \text{Tr} \rho M^{(a)} = \text{Tr} \left[\mathbf{I}_s \otimes |a\rangle\langle a|_p U_{sp} \rho \otimes |0\rangle\langle 0|_p U_{sp}^\dagger \right]$$

- Requires projective measurements in only one basis set

**KITAEV'S QUANTUM ERROR
CORRECTING CODE WITH
CONVOLUTIONAL NEURAL
NETWORKS**

KITAEV'S TORIC CODE GROUND STATE

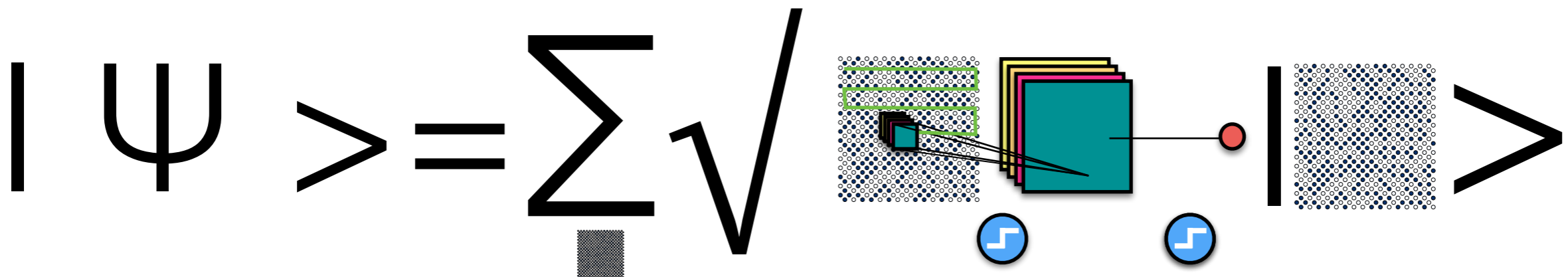
$$H = -J_p \sum_p \prod_{i \in p} \sigma_i^z - J_v \sum_v \prod_{i \in v} \sigma_i^x$$



$$|\Psi_{\text{TC}}\rangle \propto \lim_{\beta \rightarrow \infty} \sum_{\sigma_1, \dots, \sigma_N} e^{\frac{\beta}{2} J \sum_p \prod_{i \in p} \sigma_i^z} |\sigma_1, \dots, \sigma_N\rangle$$

PEPS : F. Verstraete, M. M. Wolf, D. Perez-Garcia, J. I. Cirac *Phys. Rev. Lett.* 96, 220601 (2006).

$$O_{\text{cold}}(\sigma_1, \dots, \sigma_N) \propto \lim_{\beta \rightarrow \infty} \exp \beta J \sum_p \prod_{i \in p} \sigma_i^z$$



J. Carrasquilla and R. G. Melko. *Nature Physics* 13, 431–434 (2017)

Dong-Ling Deng et al *Phys. Rev. X* 7, 021021 (2017)

Jing Chen, Song Cheng, Haidong Xie, Lei Wang, Tao Xiang *arXiv:1701.04831 RBMs*


SYNTHETIC QUANTUM DEVICES ARE GROWING FAST




nature
International journal of science

Letter | Published: 22 August 2018

Observation of topological phenomena in a programmable lattice of 1,800 qubits

Andrew D. King , Juan Carrasquilla, [...] Mohammad H. Amin

Nature **560**, 456–460 (2018) | [Download Citation](#) 

- “These are really rather beautiful pieces of science,” says physicist Seth Lloyd of MIT
- “This work presents the cleanest data for a KT transition that I have seen in any quantum simulator to date. It is the most advanced use I have to date seen of a superconducting quantum simulator.” Matthias Troyer of ETH and Microsoft
- “This paper represents a breakthrough in the simulation of physical systems which are otherwise essentially impossible,” said 2016 Nobel laureate Dr. J. Michael Kosterlitz.
- “**They haven’t addressed at all whether it’s a quantum system they’ve got,**” says physicist Graeme Smith of the JILA research center in Boulder

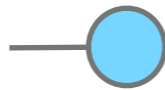
NOTATION SLIDE

S



Scalar

V_i



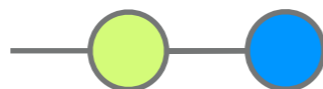
Vector

$W_{i,j}$

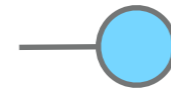


Matrix

$$C_k = \sum_j W_{k,j} V_j$$



=



Matrix vector

Multiplication

$$|C|^2 = \sum_k C_k C_k^*$$

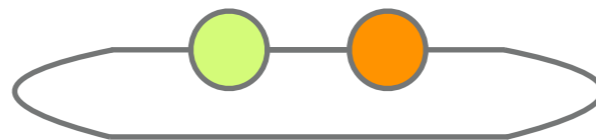


=



norm

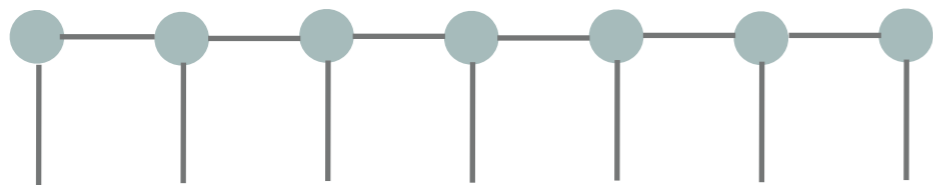
$$\text{Tr } W M = \sum_{k,j} W_{k,j} M_{j,k}$$



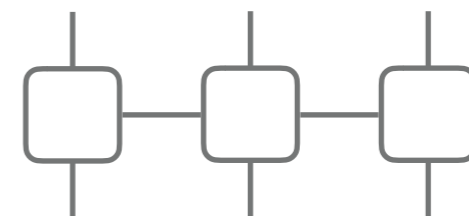
=



Matrix Product States (MPS)



Matrix Product Operators (MPO)



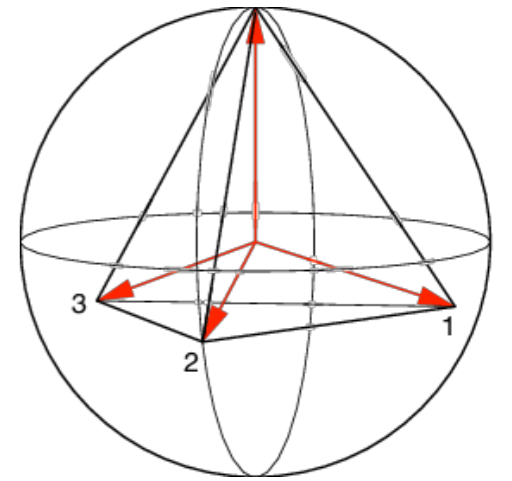
MEASUREMENTS: POSITIVE OPERATOR VALUED MEASURES (POVM)

.....
CONSTRUCTING POVMS: TAKE A SINGLE QUBIT POVM AND MAKE A TENSOR PRODUCT OF MANY

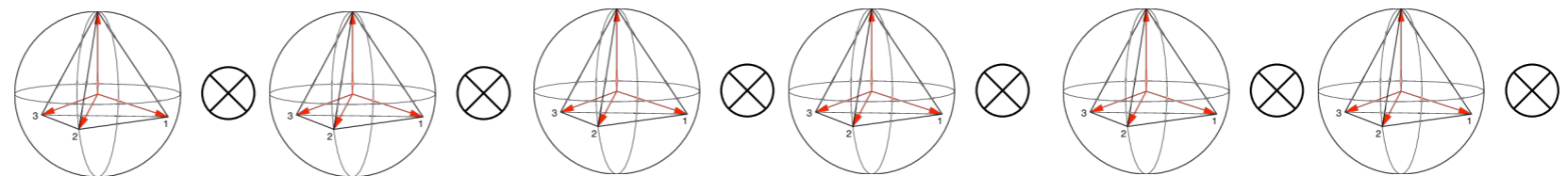
Tetrahedral measurement for one qubit

$$M^{(i)} = \frac{1}{4} \left(\mathbb{1} + \mathbf{s}^{(i)} \cdot \boldsymbol{\sigma} \right), \quad i = 1, \dots, 4$$

$$\mathbf{s}^{(1)} = (0, 0, 1), \quad \mathbf{s}^{(2)} = \left(\frac{2\sqrt{2}}{3}, 0, -\frac{1}{3} \right), \quad \mathbf{s}^{(3)} = \left(-\frac{\sqrt{2}}{3}, \sqrt{\frac{2}{3}}, -\frac{1}{3} \right), \quad \mathbf{s}^{(4)} = \left(-\frac{\sqrt{2}}{3}, -\sqrt{\frac{2}{3}}, -\frac{1}{3} \right)$$



For multiqubit systems



$$\mathbf{M} = \left\{ M^{(a_1)} \otimes M^{(a_2)} \otimes \dots \otimes M^{(a_N)} \right\}_{a_1, \dots, a_N}$$

MEASUREMENTS: POSITIVE OPERATOR VALUED MEASURES (POVM)

CONSTRUCTING POVMS: TAKE A SINGLE QUBIT POVM AND MAKE A TENSOR PRODUCT OF MANY

Pauli-4 measurement for one qubit

$$\mathbf{M}_{\text{Pauli-4}} = \left\{ M^{(0)} = \frac{1}{3} \times |0\rangle\langle 0|, M^{(1)} = \frac{1}{3} \times |1\rangle\langle 1|, M^{(2)} = \frac{1}{3} \times |+\rangle\langle +|, M^{(3)} = \frac{1}{3} \times (|-\rangle\langle -| + |r\rangle\langle r| + |l\rangle\langle l|) \right\}$$

For multiqubit systems $\mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}} \otimes \mathbf{M}_{\text{Pauli}}$

$$\mathbf{M} = \left\{ M^{(a_1)} \otimes M^{(a_2)} \otimes \dots \otimes M^{(a_N)} \right\}_{a_1, \dots, a_N}$$

Experimental realization: pick a random direction with probability 1/3, then measure in that direction + **postprocess**

Easy to implement in gate-based QC (Qiskit, Cirq, Rigetti, etc.)

OBSERVABLES AND FIDELITY W.R.T. AN MPS

