Reconstructing quantum states with generative models[1]

Author: Juan Carrasquilla, Giacomo Torlai, Roger G. Melko and Leandro Aolita

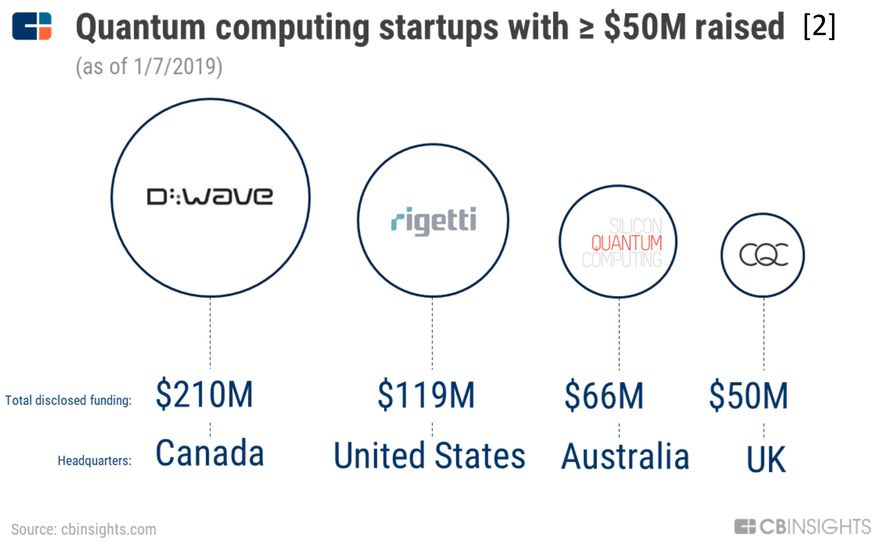
Published in: *nature machine intelligence (March 2019)*

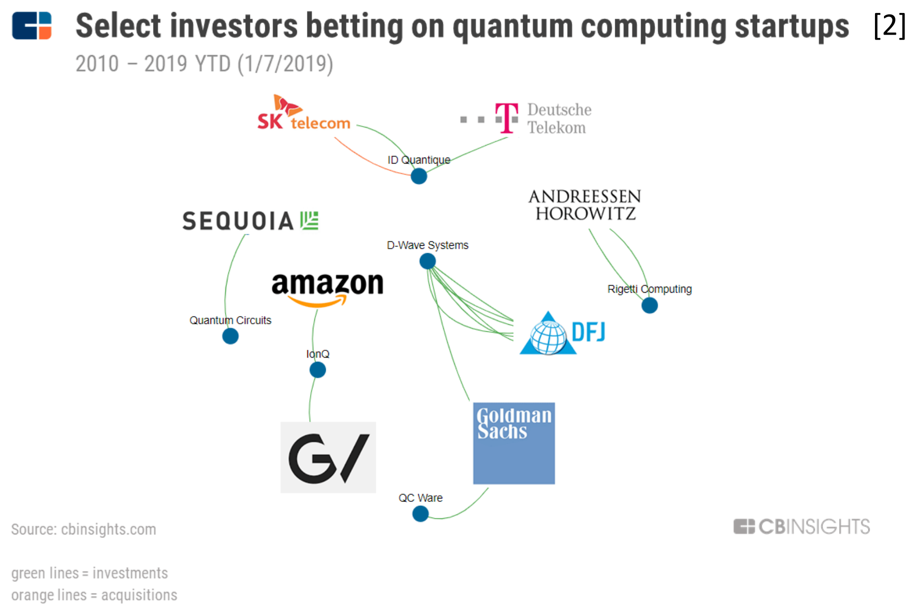
Structure

- Quantum computing / Quantum state tomography
- IC-POVMs
- General idea / Methods
- Results
- Summary

A general introduction to quantum computing

- Ongoing field of research since the 80th
- Qubit = two state system
  - Use $|0\rangle$ and $|1\rangle$ instead of classical bits 0 and 1
- Research by: Google, Microsoft, Intel, IBM and Alibaba
- Application:
  - Cryptography
  - Material science
  - Stock market
  - …





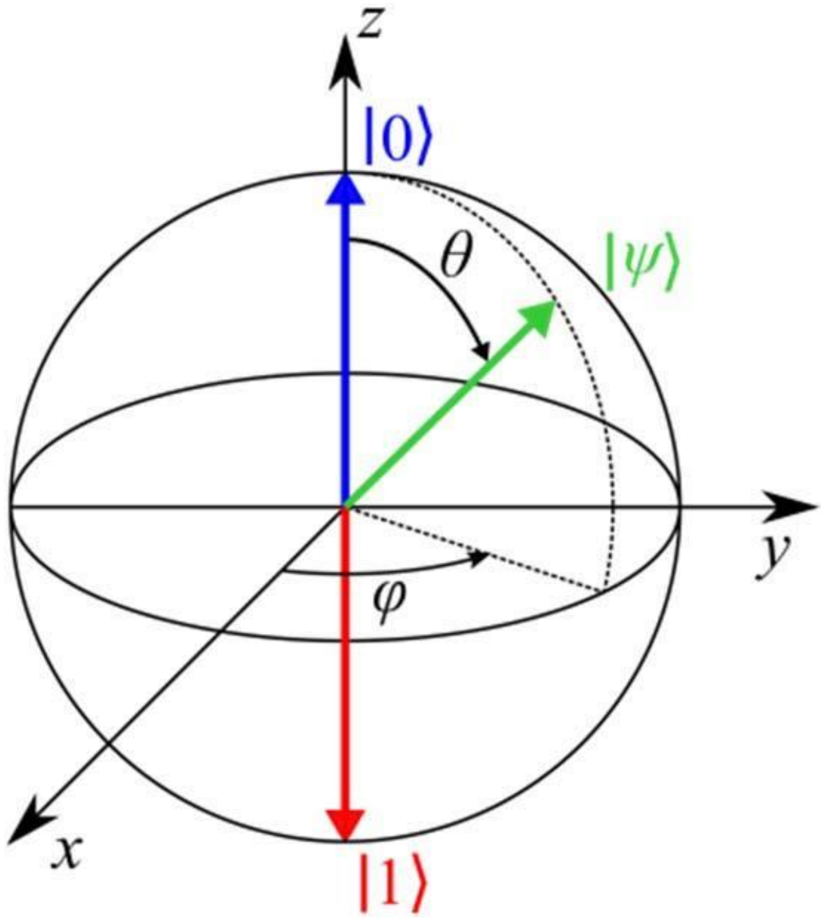
The Bloch sphere

- A general single qubit state can be visualized on the Bloch sphere by the Bloch vector

$$\vec{v} = \begin{pmatrix} \langle \sigma_x \rangle \\ \langle \sigma_y \rangle \\ \langle \sigma_z \rangle \end{pmatrix}$$

- For a pure state the angles θ and φ can be obtained by the parametrisation (up to a gobal phase):

$$|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + e^{i\phi}\sin(\frac{\theta}{2})|1\rangle$$

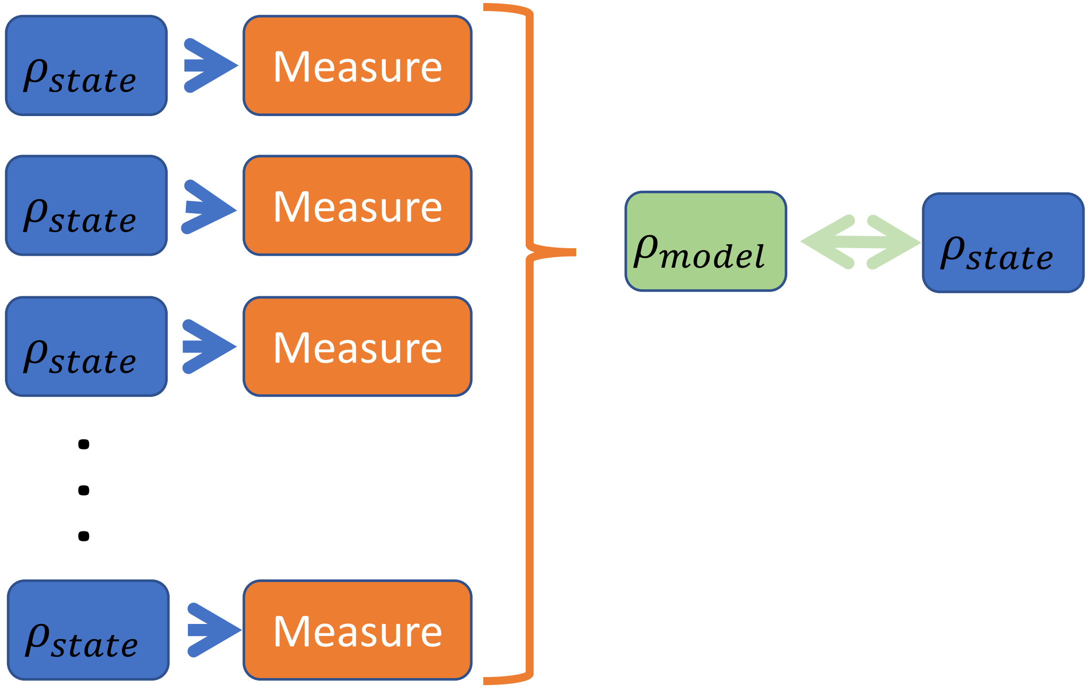


The representation of a state ψ on the Bloch sphere [3]

Quantum state tomography

how good does a quantum computer work?

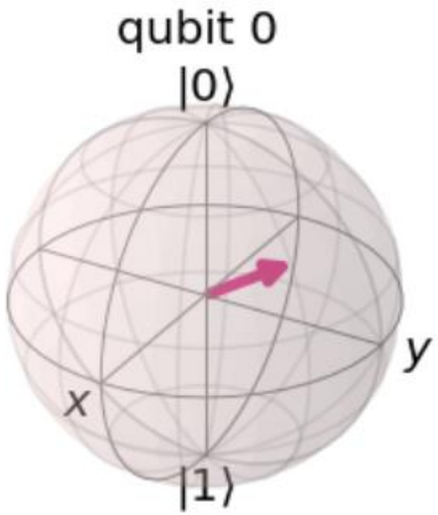
- Quantum state tomography is the method to benchmark the system
- Many measurements on a well-known state are used to reconstruct the state
- The reconstructed state is compared to the original one to quantify the quality of the quantum computer



Scheme of quantum state tomography

EXAMPLE: Single Qubit state tomography

Initial state ρ_{state}

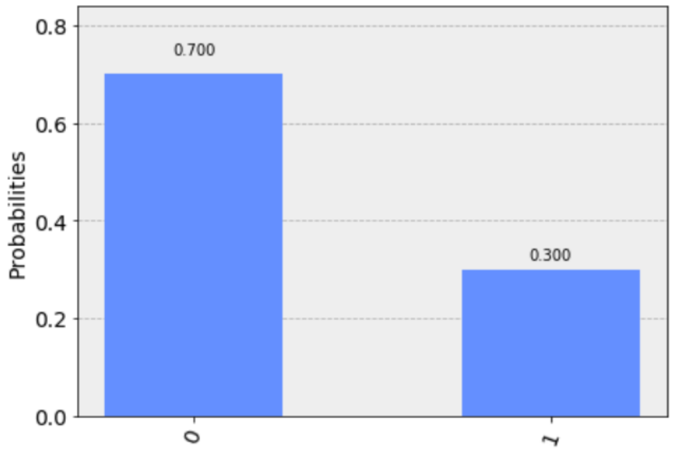


- An arbitrary single qubit density matrix can be written as:

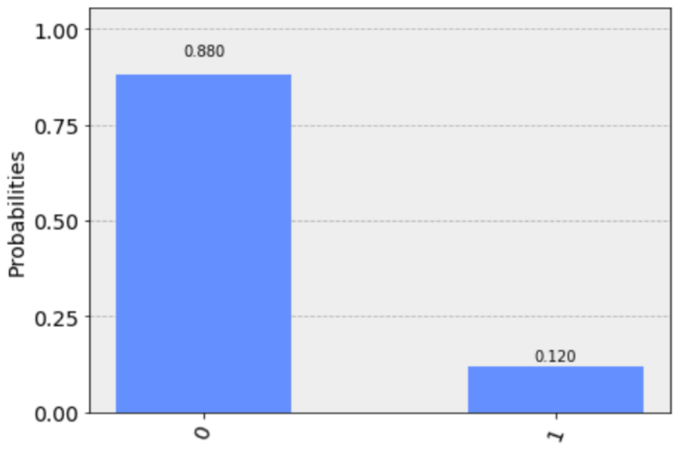
$$\hat{\rho} = \frac{1}{2}\sum_{i=0}^{3} S_i \hat{\sigma}_i \qquad \hat{\sigma}_0 = \hat{\mathbb{1}} \qquad S_i = \mathrm{Tr}(\hat{\sigma}_i \hat{\rho})$$

- Only the expectation values are required
- IBM provides an online platform that provides access to a few qubit quantum computer, so we can try it out!
  (use 100 measurements for every S_i)

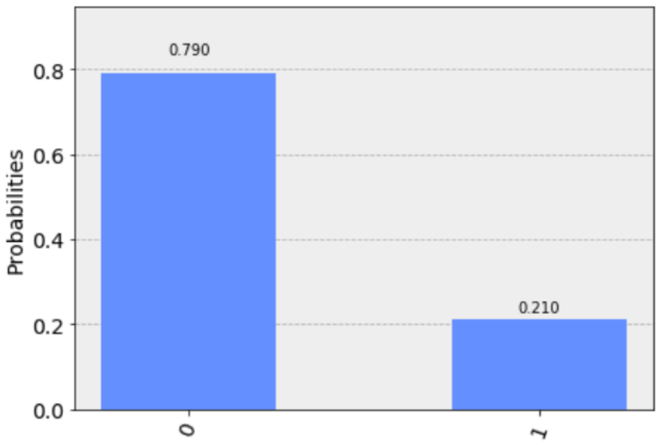
σ_1



σ_2

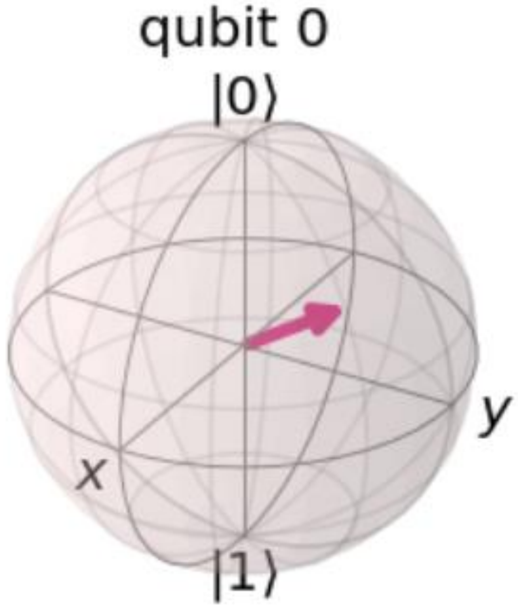


σ_3

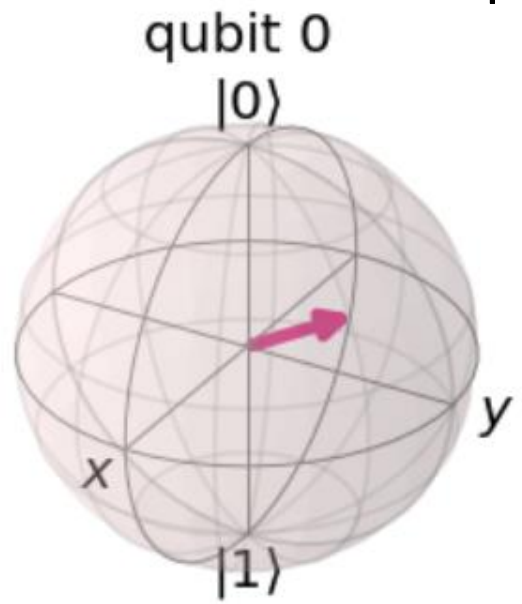


EXAMPLE: Single Qubit state tomography

Initial pure state ρ_{state}



reconstructed state ρ_{model}



Quality criterium: $F(\rho_{state}, \rho_{model}) = \mathrm{Tr}[\sqrt{\sqrt{\rho_{state}}\rho_{model}\sqrt{\rho_{state}}}]$ = 0.982

Problem: In general there are 4^N -1 free parameter which have to be obtained from measurements:

- Makes exact state tomography impractical for large N: *curse of dimensionality*

Information Complete Positive Operator Valued Measure (IC-POVM)

Definition POVM:

- Set of Measurements **M** = $\{M^{(\alpha)}\}$
- Each $M^{(\alpha)}$ is a positive semidefinite hermitian operator
- $\sum_\alpha M^{(\alpha)} = \mathbb{1}$

$$1 = Tr(\rho) = \sum_i \langle i|\rho|i\rangle$$
$$= \sum_i \langle i| \sum_\alpha \rho M^{(\alpha)} |i\rangle$$
$$= \sum_\alpha Tr(\rho M^{(\alpha)})$$
$$= \sum_\alpha P_\alpha \text{ with } P_\alpha := Tr(\rho M^{(\alpha)}) \geq 0$$

Definition IC-POVM:

- the POVM spans the whole space of bounded norm operators on the observed hilbert space

A probability can be assigned to each measurement $M^{(\alpha)}$!

Example: Thetrahedral single qubit IC-POVM

- 4 rank-1 projectors are used (see picture)

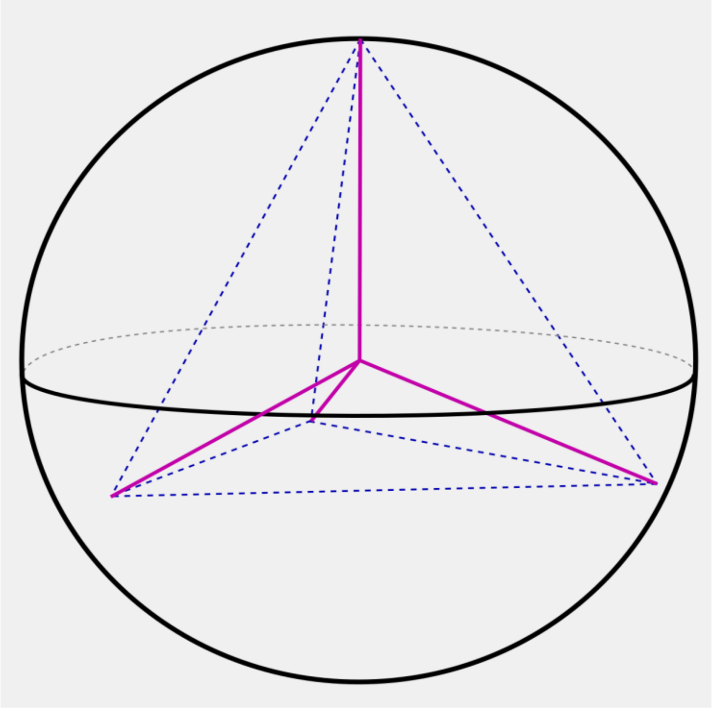
$$M_{\text{tetra}} = \left\{ M^{(a)} = \frac{1}{4}(\mathbb{1} + \boldsymbol{s}^{(a)} \cdot \boldsymbol{\sigma}) \right\}_{a \in \{0,1,2,3\}}$$

$$s^0 = (0, 0, 1)$$

$$s^1 = (\tfrac{2\sqrt{2}}{3}, 0, -\tfrac{1}{3})$$

$$s^2 = (-\tfrac{\sqrt{2}}{3}, \sqrt{\tfrac{2}{3}}, -\tfrac{1}{3})$$

$$s^3 = (-\tfrac{\sqrt{2}}{3}, -\sqrt{\tfrac{2}{3}}, -\tfrac{1}{3})$$



Visualization of M_{thetra} [4]

$$\boxed{P_\alpha = \mathrm{Tr}(\rho M^{(\alpha)})}$$

How can P_α be measured?

- Couple to an ancillary qubit
- Apply a unitary operator U in a way, that the new state shows P_α

$$|1\rangle, |0\rangle \xrightarrow{U} |00\rangle, |01\rangle, |10\rangle, |11\rangle$$

$$\psi = \sqrt{P_0}|00\rangle + \sqrt{P_1}|01\rangle + \sqrt{P_2}|10\rangle + \sqrt{P_3}|11\rangle$$

Single-qubit M_{thetra} $\Longrightarrow$ Multi-qubit M_{Thetra}

Natural extension:

- Apply M_{thetra} on each single qubit

- $M \to M_1 \otimes M_2 \otimes M_3 \ldots$

Requires only single-qubit measurements:

- Easy implementation in real experiment

How a measurement looks like for M_{thetra}:

$$\begin{array}{c} \text{qubit 1} \\ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{array} \otimes \begin{array}{c} \text{qubit 2} \\ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{array} \otimes \begin{array}{c} \text{qubit 3} \\ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \end{array} \otimes \ldots$$

IC-POVMs: some mathtools

How can we model ρ with an IC-POVM?

$$\rho \overset{IC}{=} \sum_{\alpha'} c_{\alpha'} M^{(\alpha')} = \vec{c} \cdot \vec{M}$$

$$\vec{P} = \{P_\alpha\} = \mathrm{Tr}\big(\rho \vec{M}\big) = \vec{c} \cdot \mathbf{T}$$

$$\text{with } \mathbf{T}_{\alpha,\beta} := \mathrm{Tr}(M^{(\alpha)} M^{(\beta)})$$

$$\vec{c} = \vec{P} \cdot \mathbf{T^{-1}}$$

$$\boxed{\rho = \vec{P} \cdot \mathbf{T^{-1}} \vec{M} = \mathbb{E}_P\big[\mathbf{T^{-1}} \vec{M}\big]}$$

Multi-qubit $\mathbf{M_{thetra}}$:

$$\boxed{\rho = \mathbb{E}_P\big[\mathbf{T}^{-1} \vec{M}_1 \otimes \mathbf{T}^{-1} \vec{M}_2 \otimes \dots\big]}$$

How can we estimate arbitrary operators *O*?

$$O \overset{IC}{=} \sum_{\alpha'} d_{\alpha'} M^{(\alpha')} = \vec{d} \cdot \vec{M}$$

$$\vec{d} = \mathrm{Tr}(O\vec{M}) \cdot \mathbf{T^{-1}}$$

- determines each d univocally

$$\boxed{\langle O \rangle = \mathrm{Tr}(\rho O) = \vec{d} \cdot \vec{P} = \mathbb{E}_P\big[\vec{d}\big]}$$

$c, d \in \mathbb{R}; \quad M^{(\alpha')}, O =$ operators

$\vec{M} =$ vector of operators

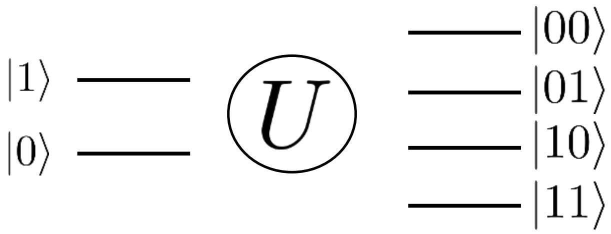
$\vec{c}, \vec{d} =$ vector of coefficients

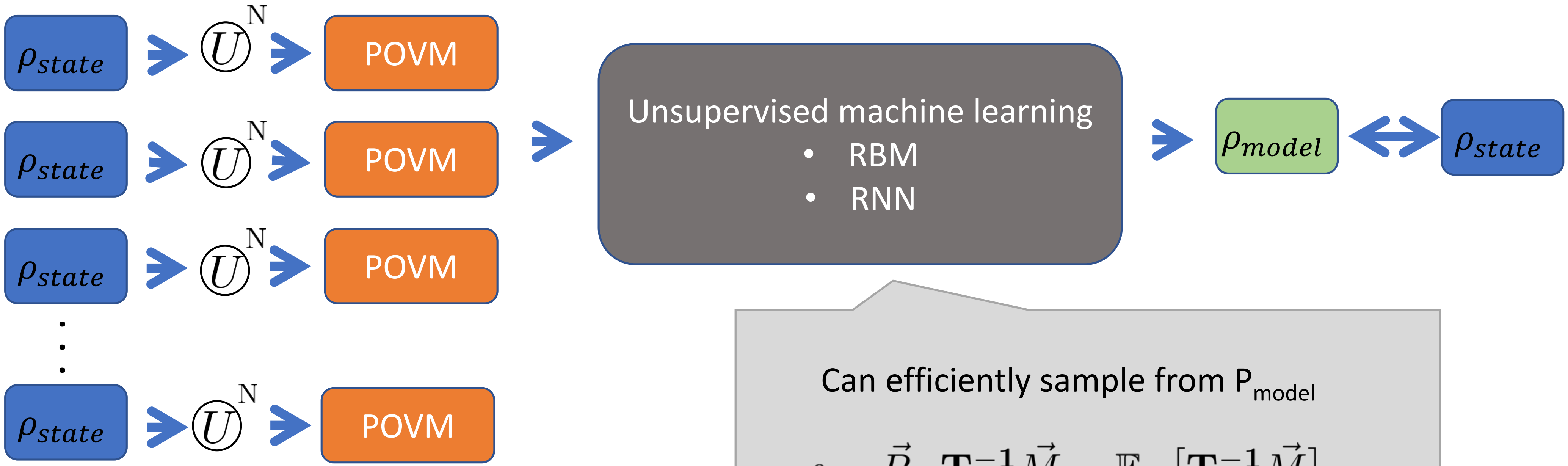
General idea

One measurement of M_{thetra}:

$$\begin{array}{ccccccc} \text{qubit 1} & & \text{qubit 2} & & \text{qubit 3} & & \\ \begin{pmatrix}1\\0\\0\\0\end{pmatrix} & \otimes & \begin{pmatrix}0\\0\\0\\1\end{pmatrix} & \otimes & \begin{pmatrix}1\\0\\0\\0\end{pmatrix} & \otimes & \dots \end{array}$$

Reminder:





N_s measurements on N qubits

Can efficiently sample from P_{model}

$$\rho = \vec{P} \cdot \mathbf{T^{-1}} \vec{M} = \mathbb{E}_P\left[\mathbf{T^{-1}} \vec{M}\right]$$

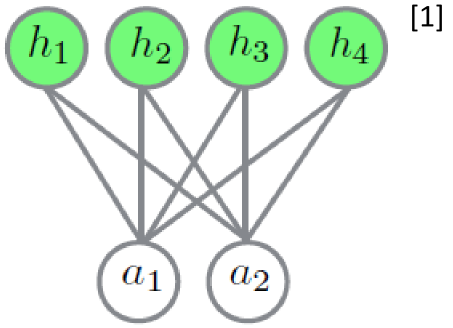
$$\langle O \rangle = \mathrm{Tr}(\rho O) = \vec{d} \cdot \vec{P} = \mathbb{E}_P\left[\vec{d}\right]$$

Restricted Boltzmann machine (RBM)

Energy: $E(v,h) = -\sum_{i,j,k,l} W_{ij}^{kl} v_i^k h_j^l - \sum_{i,k} b_i^k v_i^k - \sum_{j,l} a_j^l h_j^l$

- *W = Weights* — *upper indices = dimension of v, h*
- *v, h = visible and hidden units*
- *b, a = biases for v and h respectively*

Hidden units h_j of dimension L



Visible units v_i of dimension M, where M is the number of measurements in the POVM

Training:

Gibbs sampling (1 step):

v is the measurement
P(h|v) is calculated, h is drawn from this distribution
P(v|h) is calculated, v_{new} is drawn from this distribution

Update general parameter $\theta \in (W, b, a)$ by contrastive divergence

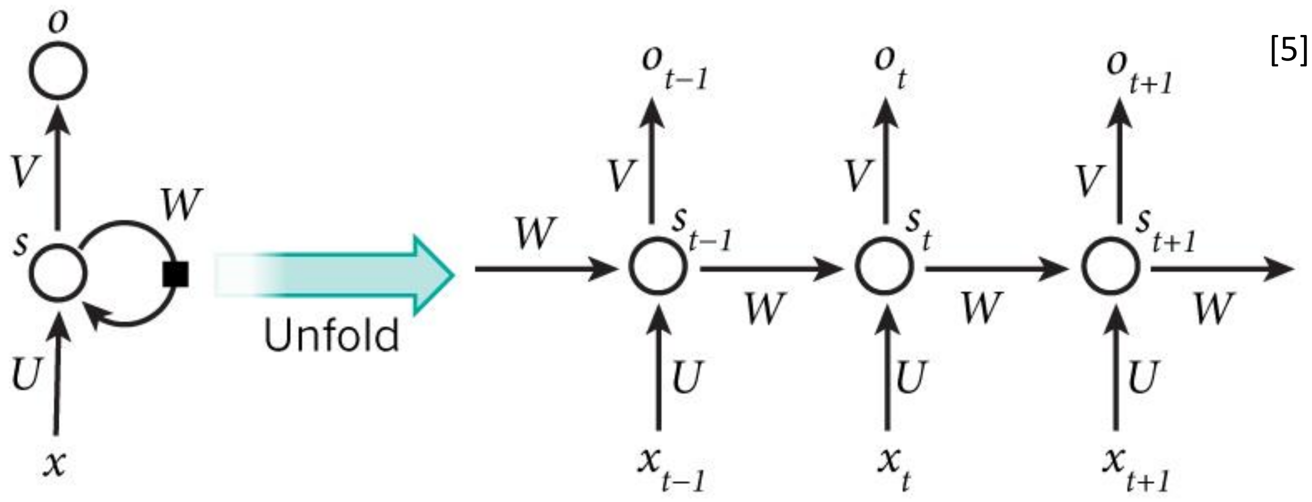
$$\mathcal{L} = \langle \log(p_\theta(x)) \rangle_{data}$$

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = \mathbb{E}\left[\frac{\partial E(\mathrm{v},\theta)}{\partial \theta_i}\right] - \mathbb{E}\left[\frac{\partial E(\mathrm{v_{new}},\theta)}{\partial \theta_i}\right]$$

Sampling:

Gibbs sampling after training (with optimized parameter)

Recurrent neural network (RNN)



Variables:

- Inputs *x*
- Input weights *U*
- Outputs *o*
- Hidden state *s (memory)*
- Hidden weights *W*

RNN:

- Optimized to process sequential data with *memory*
- Each decision depends on the decision before:
  $s_t = f(Ws_{t-1} + Ux_t)$
  e.g. *f* is a ReLU or tanh

For this paper:

- Sequence step = one qubit
- Output = single qubit prob. dist.
- Hidden state = contains all relevant information for the output

1. P_1 is calculated (first qubit)
2. P_2 is calculated with the memory of P_1 : $P_2|P_1$ is obtained
3. …

This is done for every qubit, then we get

$$\boxed{P_{model} = P_1 \times (P_2|P_1) \times \ldots \times (P_N|P_1, P_2 \ldots P_{N-1})}$$

How do we measure the quality of the reconstruction?

- **Kullback-Leibler divergence**

$$D_{KL}(P_{state}||P_{model}) = \mathbb{E}_{P_{state}}\left[\log \frac{P_{state}}{P_{model}}\right]$$

  - Measures how much P_{model} diverges from P_{state}

- **Classical Fidelity**

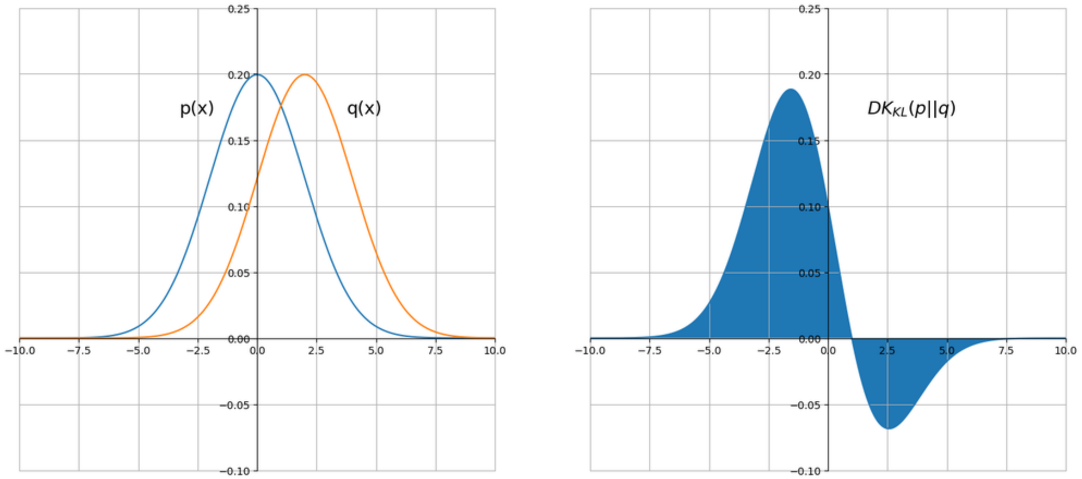
$$F_C(P_{state}, P_{model}) = \mathbb{E}_{\alpha \sim P_{state}}\left[\sqrt{\frac{P_{model}(\alpha)}{P_{state}(\alpha)}}\right]$$

- **Quantum Fidelity**

$$F(\rho_{state}, \rho_{model}) = \mathrm{Tr}\left[\sqrt{\sqrt{\rho_{state}}\rho_{model}\sqrt{\rho_{state}}}\right]$$

In the case of pure states this simplifies to

$$F(\psi_{state}, \psi_{model}) = \langle \psi_{state} | \psi_{model} \rangle$$



KL divergence for different gaussians. [6]

Which reconstruction criterium should be used?

- Measure the different reconstruction quality criteria using an RBM with the N=2 GHZ state
- As POVM, the multi-qubit thetrahedral POVM is used
- Add single-qubit depolarisation probability *p*

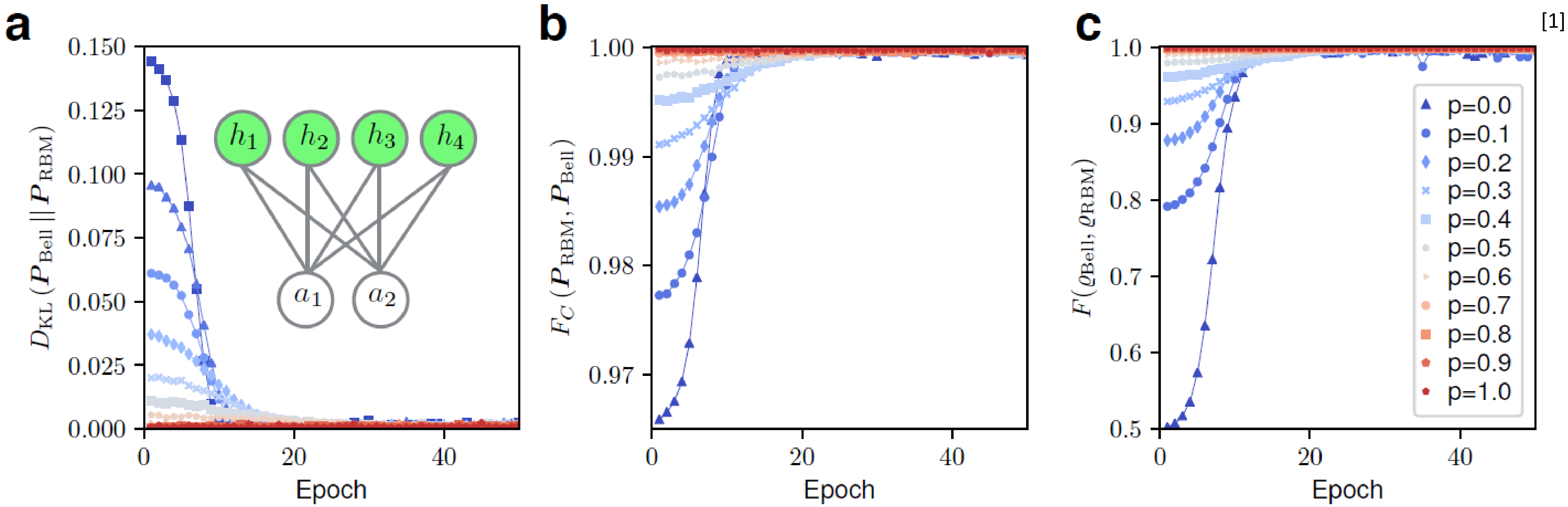
- Greenberger-Horne-Zeilinger (GHZ) state:

$$|\Psi\rangle_{GHZ} = \frac{1}{\sqrt{2}}(|0\rangle^N + |1\rangle^N)$$

- Depolarisation of a single qubit:

$$\rho \rightarrow \mathrm{p} \cdot \left(\frac{1}{2} \cdot \mathbb{1}\right) + (1-p) \cdot \rho$$

Which reconstruction criterium should be used?



Epoch: 1 step of improving the weights and biases after Gibbs sampling

F_C and F approach unity in a similar behaviour

F_C for different number of training samples N_s

- F_C is measured for different number of qubits N and for the depolarization probability *p* = 0.0 or 0.4

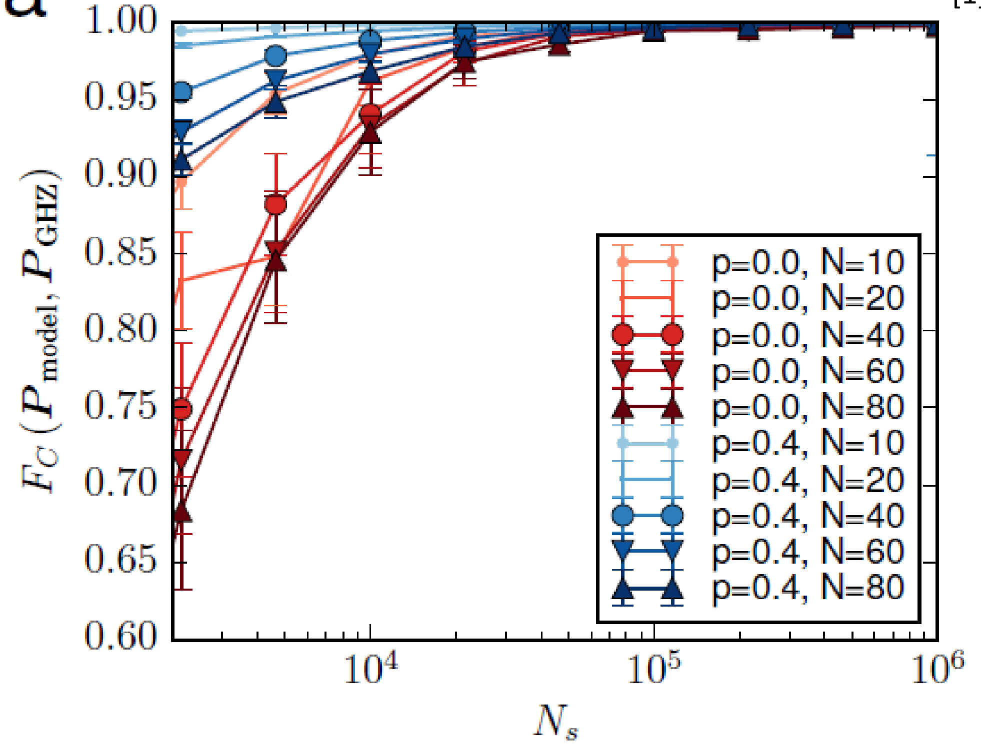
- Use the General GHZ state
  $|\Psi\rangle_{GHZ} = \frac{1}{\sqrt{2}}(|0\rangle^N + |1\rangle^N)$

- Advantage of F_C: It only requires P_{model} and can be efficiently estimated for large N by simply sampling and averaging from P_{model}

- RBM is difficult to train; RNN is used

$$F_C(P_{state}, P_{model}) = \mathbb{E}_{\alpha \sim P_{state}}\left[\sqrt{\frac{P_{model}(\alpha)}{P_{state}(\alpha)}}\right]$$

a



[1]

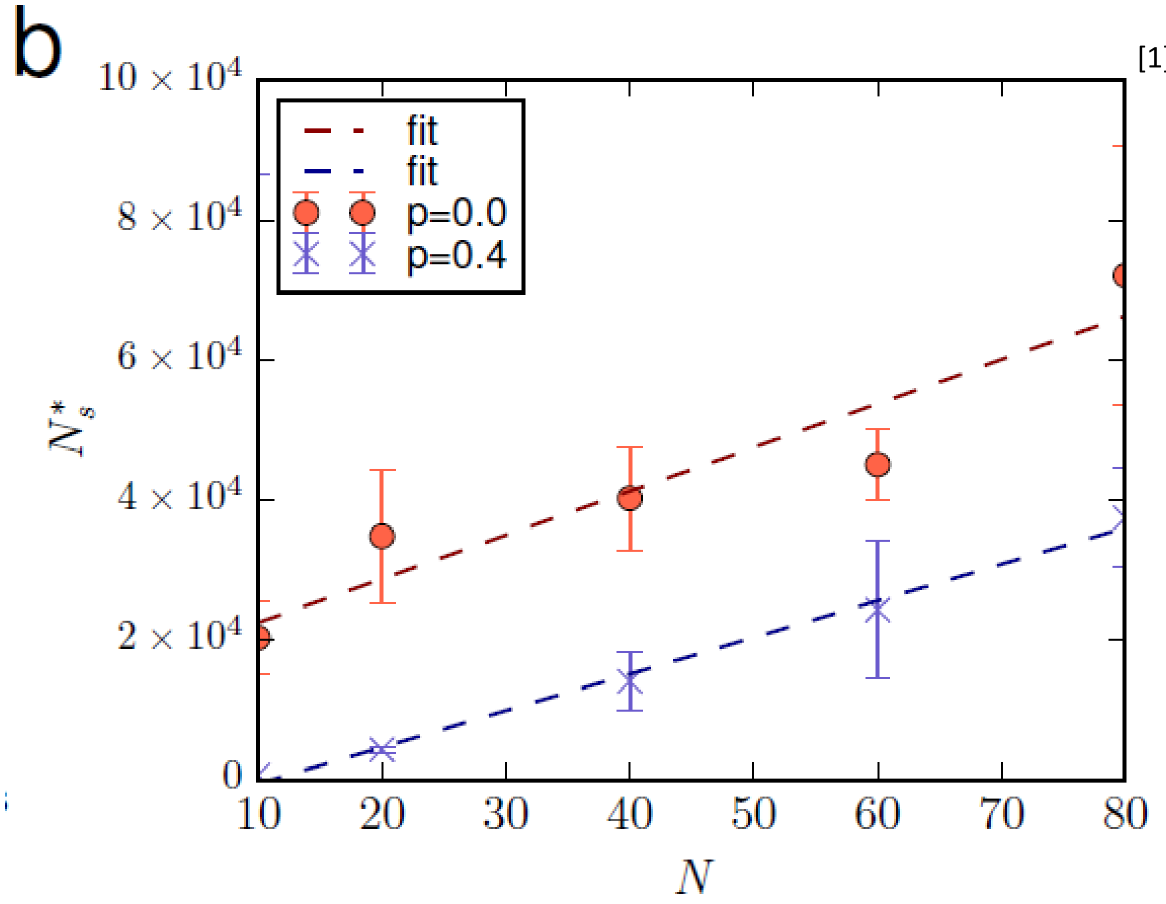
How does the number of samples N_s scale with the number of qubits for a given F_C?

- A Fidelity value is fixed

$$N_s^* := F_C(N_s^*) = 0.99$$

- The scaling is approximately linear!

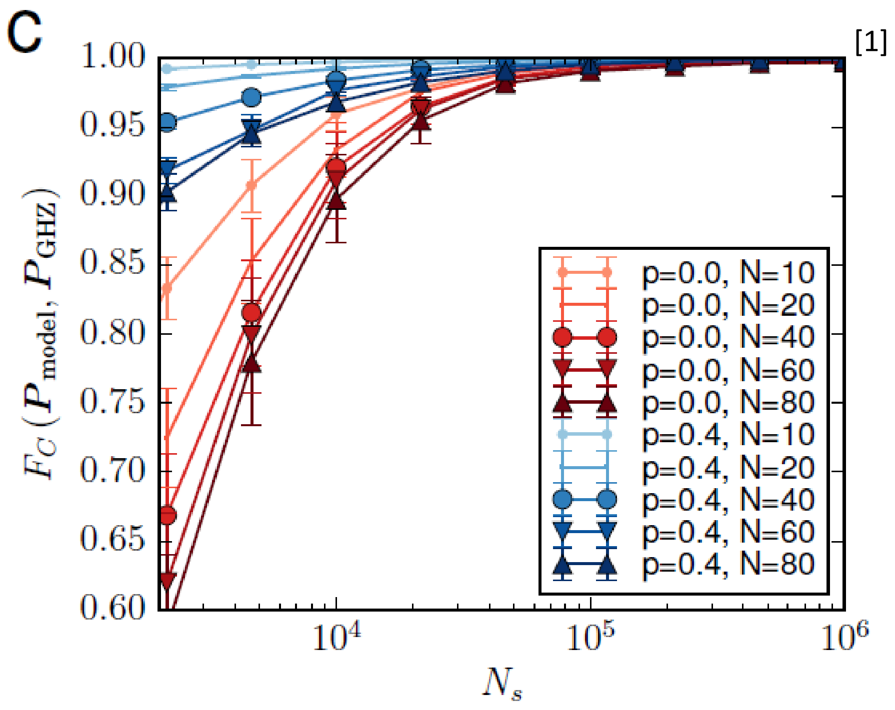
- Pure luck? Is it just a pecularity of M_{thetra} ?

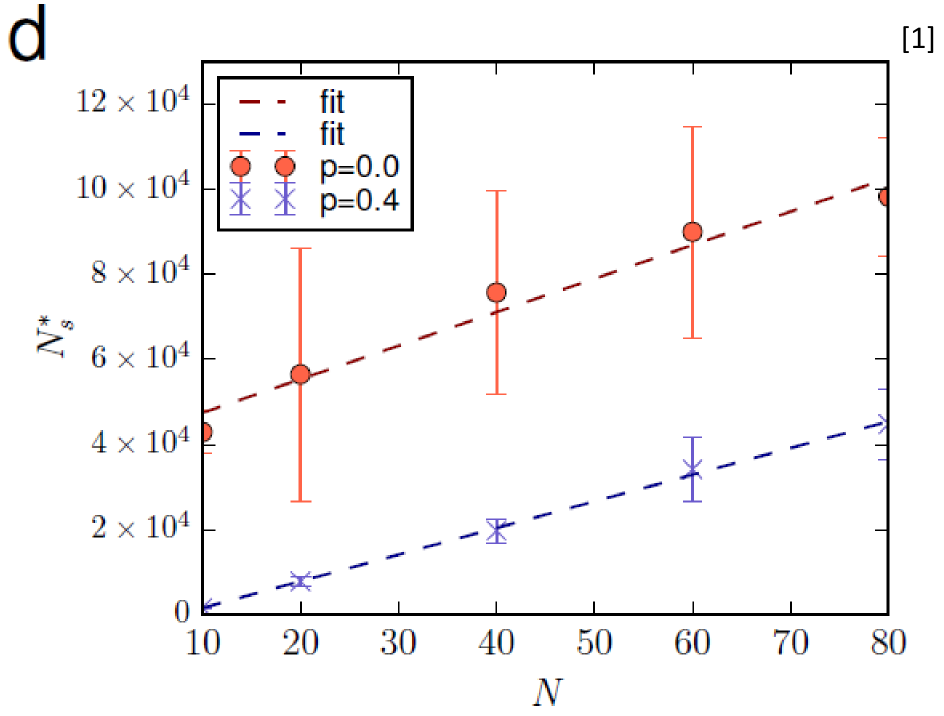


[1]

Pauli M_6 POVM

- Set of 6 measurements
- Projections on the 6 eigenstates of the three pauli matrices





- Even though 6 different measurements per single-qubit are used, a linear behaviour can still be observed!

Antiferromagnetic 1D transverse field ising model

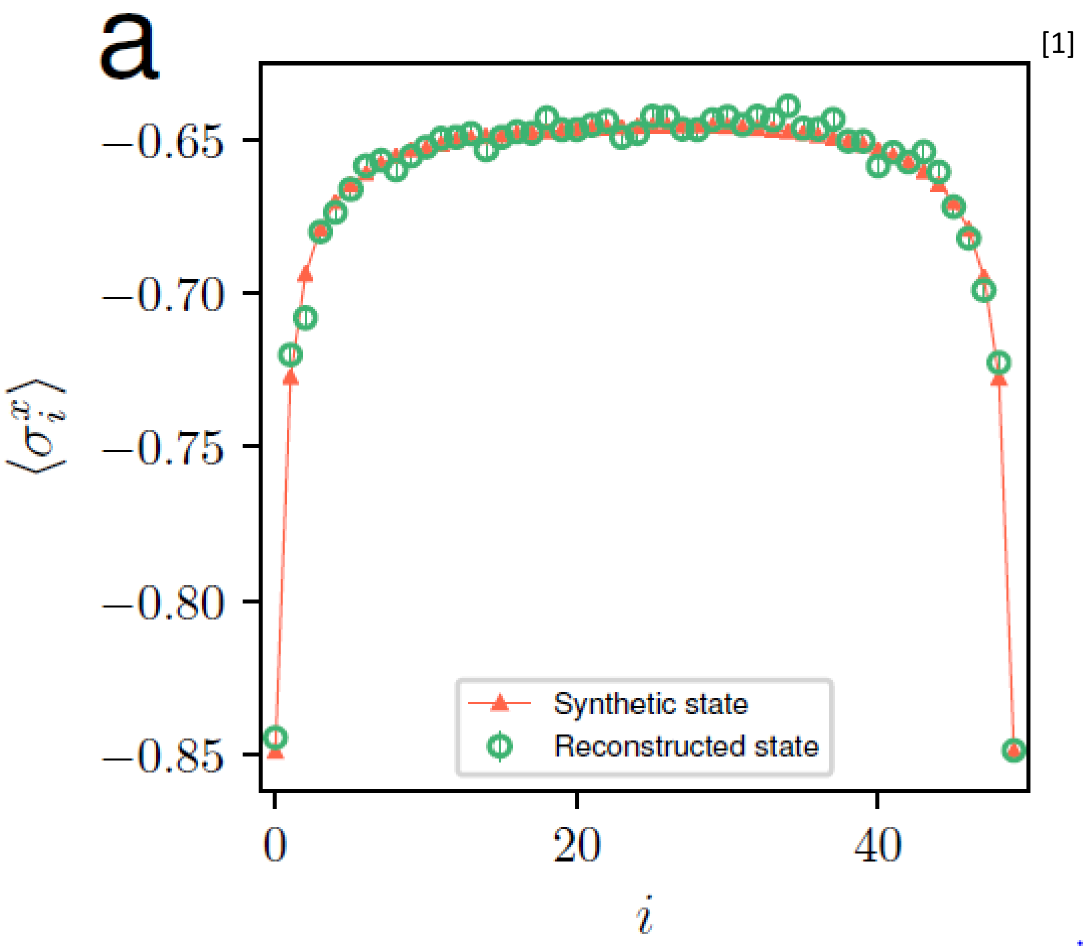
$$\mathcal{H} = J \sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z + h \sum_i \sigma_i^x$$

$J \gg h$: The ground state is $|\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \ldots\rangle$

$J \ll h$: The ground state is $|\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \ldots\rangle$

- The phase transition at J = h is observed
- The problem can be solved analytically and is therefore the testbed par excellence for quantum simulators.

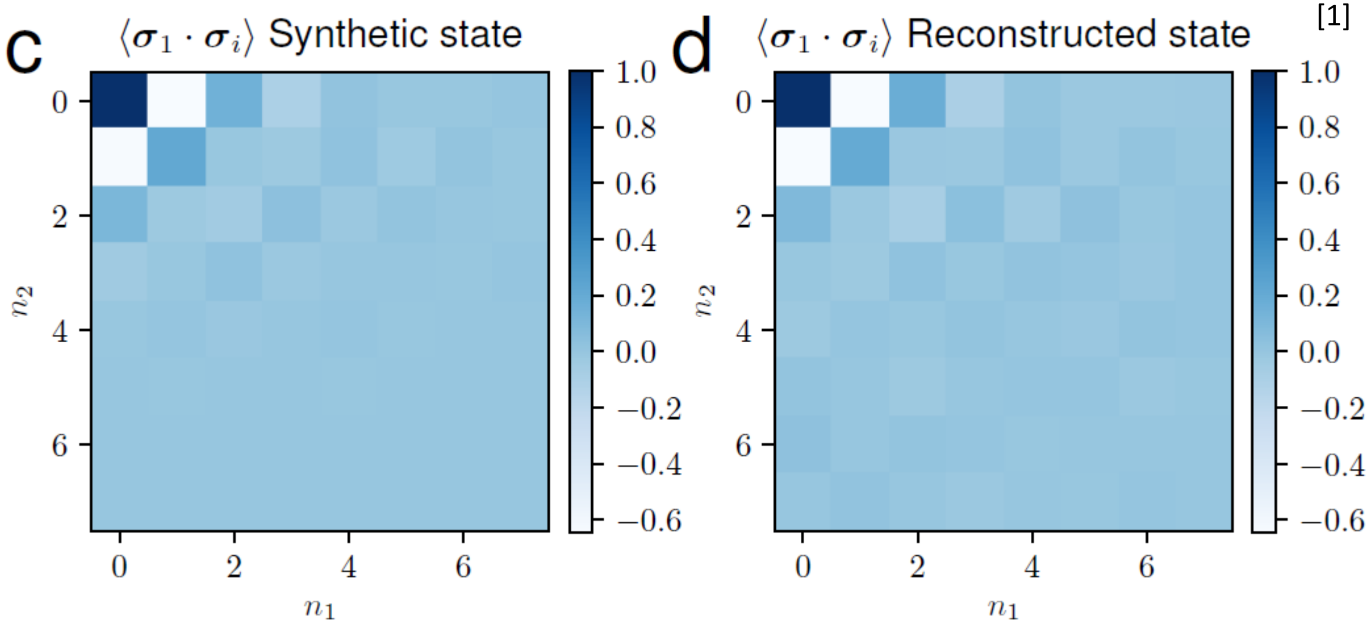
- $N_{qubit} = 50$; $N_s = 10^6$
- $F_C = 0.998$
- Also investigate 1-body correlation function



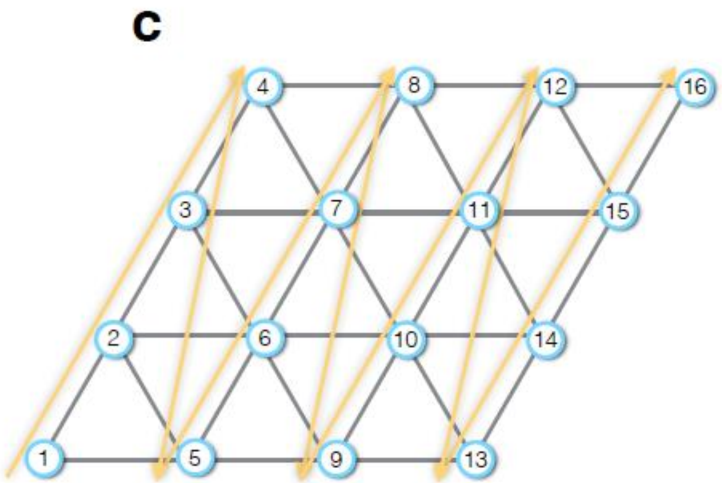
Antiferromagnetic Heisenberg model triangular lattice

$$\mathcal{H} = \sum_{\langle ij \rangle} \vec{\sigma}_i \vec{\sigma}_j$$

- No analytical solution
- A 8 x 8 lattice is used (64 Qubits)
- $F_C = 0.98$
- Remarkable agreement in the correlation function



[1]



[1]

Model for the path used by the RNN (4 x 4 lattice)

$$\boxed{\langle O \rangle = \mathrm{Tr}(\rho O) = \vec{d} \cdot \vec{P} = \mathbb{E}_P \left[\vec{d} \right]}$$

- Can be easily and efficiently computed!

Conclusion

- Method: Quantum state tomography based on probability distribution estimation using unsupervised machine learning
- Could be certified on pure and mixed states, as well as on ground states of local hamiltonians
- A linear scaling between N and N_s^* is observed!
- Can efficiently estimate F_c and expectation values of local operators, e. g. $\langle \sigma_x \rangle$

References

- [1]: arXiv:1810.10584: Reconstructing quantum states with generative models
- [2]: https://www.cbinsights.com/research/report/quantum-computing/
- [3] https://www.researchgate.net/figure/Bloch-sphere-representation-of-a-qubit_fig1_317573486
- [4] https://en.wikipedia.org/wiki/SIC-POVM#/media/File:Regular_tetrahedron_inscribed_in_a_sphere.svg
- [5] http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/
- [6] https://www.science-emergence.com/Articles/How-to-calculate-and-visualize-Kullback-Leibler-divergence-using-python-/

Limitation:

- Expectation values of non-local operators can not always be well estimated by sampling, e.g. F