*Chaire de Physique
de la Matière Condensée
Antoine Georges*

Réseaux de Neurones, Apprentissage et Physique Quantique

Cours 3:

Introduction à la tomographie quantique utilisant les réseaux de neurones

Cycle 2022-2023
16 mai 2023

Image Credit: Javier Robledo-Moreno

Seminar (11h30): Giuseppe Carleo
Time-Dependent Neural Quantum States

Time-Dependent Neural Quantum States

Giuseppe Carleo
(École Polytechnique Fédérale de Lausanne)

In this presentation I will discuss how neural-network based representations of the many-body wave function can be used to study unitary dynamics of many-body quantum systems. I will first derive the general framework, introducing the time-dependent variational principle in the formulation of McLachlan and Dirac-Frenkel. I will then introduce and discuss stochastic approaches to solve the resulting equations of motion, in the framework of the time-dependent Variational Monte Carlo method. Applications to the exploration of out-of-equilibrium properties of two-dimensional spin systems will be discussed. The generalization of these methods to other types of dynamics, including general quantum circuits, and dissipative dynamics, will be also presented. I will show applications to the classical simulation of quantum algorithms and of measurement-induced dynamical phase transitions.

Mailing List

(Weekly announcement of lecture and seminar, etc.)

Send email to: listes-diffusion.cdf@college-de-france.fr
Subject line: subscribe chaire-pmc.ipcdf
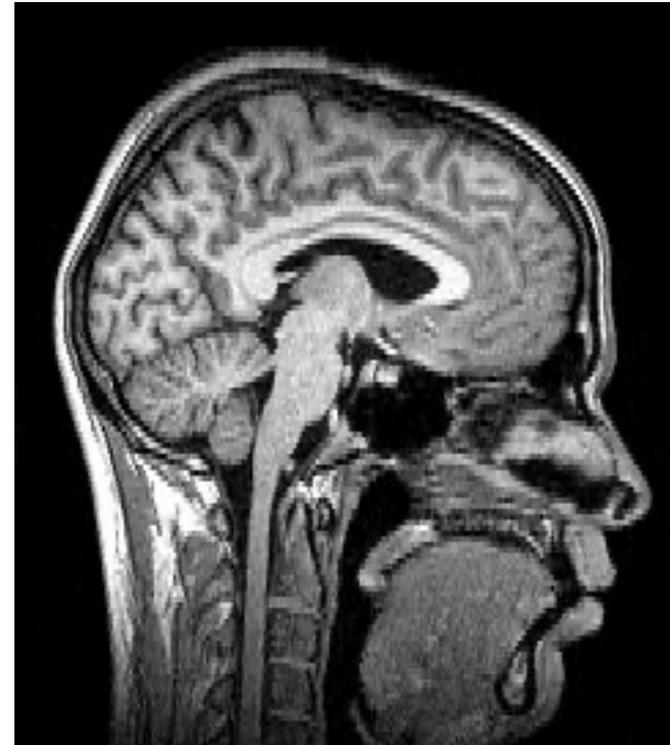…or: unsubscribe chaire-pmc.ipcdf

You can also just send me an email to be placed on the list

Website:

https://www.college-de-france.fr/site/antoine-georges/index.htm

Lectures are video recorded
and available on the website

Tomography = Reconstruct From Cuts (τόμος= to cut)

Projection

Cuts
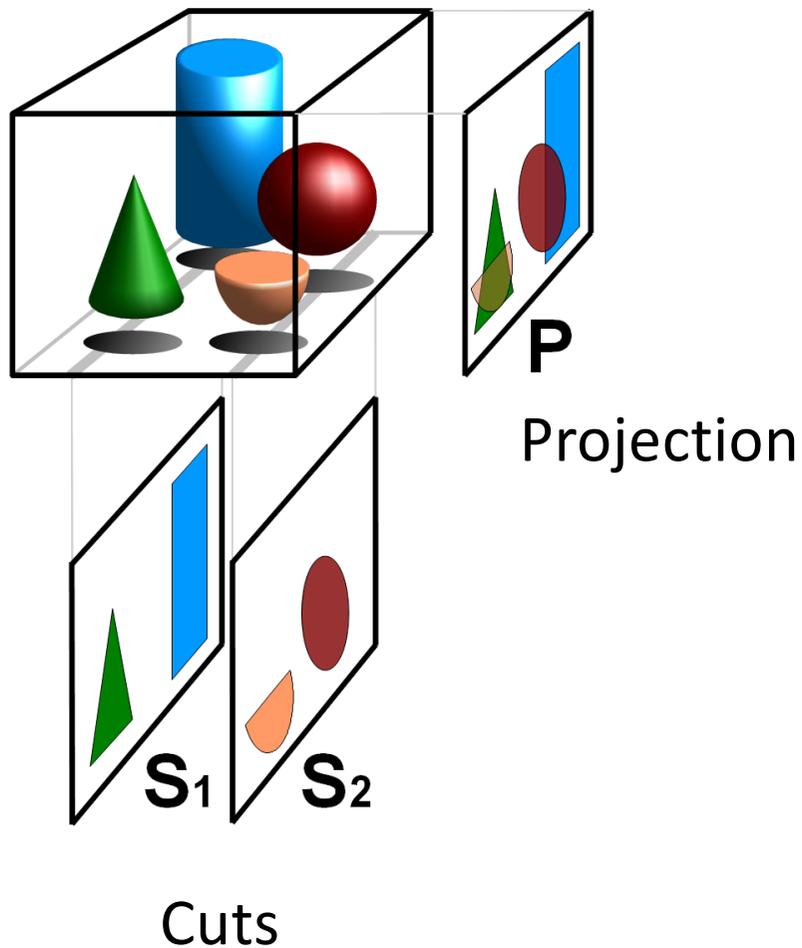
Images from Wikipedia `Tomography'

Reminder: Measurement in Quantum Mechanics

- Consider an observable to which we associate an Hermitian operator Ô (e.g spin along some axis: one of the Pauli operators)

- Spectrum of eigenstates of Ô: $\hat{O}|\omega_i\rangle = \omega_i|\omega_i\rangle$

- <u>Postulate</u>: If the system is in a (normalized) pure state, the result of a measure of Ô yields one of its eigenvalues ω_i with probability:

$$p(\omega_i) = |\langle\omega_i|\psi\rangle|^2$$

- <u>Postulate</u> (collapse of the state - ``réduction du paquet d'onde''): After the measurement the system is in the new projected state:

$$|\psi'\rangle = \frac{\hat{\Pi}_i|\psi\rangle}{\sqrt{p(\omega_i)}} \ , \ \hat{\Pi}_i \equiv |\omega_i\rangle\langle\omega_i|$$

If the system is in an ensemble of states described by a density matrix:

$$\hat{\rho} = \sum_{\alpha} p_{\alpha} \, |\psi_{\alpha}\rangle\langle\psi_{\alpha}| \quad \text{(in the basis where it is diagonal)}$$

The result of a measurement has probability (`Born rule'):

$$p(\omega_i) = \text{tr}\left(\hat{\Pi}_i \hat{\rho}\right)$$

$$= \sum p_{\alpha} \, |\langle\omega_i|\psi_{\alpha}\rangle|^2$$

Average value of O obtained in repeated measurements:

$$\langle\hat{O}\rangle = \sum_{\alpha} p_{\alpha} \sum_{i} \omega_i |\langle\omega_i|\psi_{\alpha}\rangle|^2 = \text{tr}\left(\hat{O}\,\hat{\rho}\right)$$

Measurement of the Wigner Distribution and the Density Matrix of a Light Mode Using Optical Homodyne Tomography: Application to Squeezed States and the Vacuum

D. T. Smithey, M. Beck, and M. G. Raymer

Department of Physics and Chemical Physics Institute, University of Oregon, Eugene, Oregon 97403

A. Faridani

Department of Mathematics, Oregon State University, Corvallis, Oregon 97331
(Received 16 November 1992)

We have measured probability distributions of quadrature-field amplitude for both vacuum and quadrature-squeezed states of a mode of the electromagnetic field. From these measurements we demonstrate the technique of optical homodyne tomography to determine the Wigner distribution and the density matrix of the mode. This provides a complete quantum mechanical characterization of the measured mode.

First appearance of `Tomography' in the context of quantum systems (according to Wikipedia…)

What is Quantum State Tomography?

- We want to reconstruct the quantum state of a system – the wavefunction if we know that the system is in a pure state or more generally the density matrix – using the result of measurements

- Repeated projective measurements will allow to estimate the probabilities (assuming that the density matrix stays constant during the measurement process…):

$$p_i = \text{tr}\left(\hat{\Pi}_i \hat{\rho}\right)$$

- For some numbers of operators

- Positive Operator-Valued Measure (POVM): it is convenient to chose the set of operators we measure to form a POVM – an ensemble of positive semi-definite operators M_i such that:

$$\sum_i \hat{M}_i = \mathbb{I}$$

QST by direct inversion

- Reconstruction by direct inversion:

$$p_i = \operatorname{tr}\left(\hat{M}_i\,\hat{\rho}\right)$$

- Can be viewed as a system of linear equations. The density matrix invoves $(D_H)^2$ complex numbers hence we need a priori to measure $(D_H)^2$ probabilities (-1 overall normalization)

- For N qubits: need to make 2^{2N} sets of repeated measurements

- Exponential cost!

- In practice, this method does not guarantee that the determined density matrix is positive (see below for an example)

Example: tomography of a single qubit

- Density matrix: positive semi-definite, Hermitian, unit trace:

$$\hat{\rho} = \hat{\rho}^{\dagger} \ , \ \ \mathrm{tr}\,\hat{\rho} = 1 \ , \ \ \forall u \ , \langle u|\hat{\rho}|u\rangle \geq 0 \ , \ \ \lambda_{\alpha}^{(\rho)} \geq 0$$

- General parametrization for a single `qubit':

$$\hat{\rho} = \frac{1}{2}\left[\mathbb{I} + \vec{\rho}\cdot\vec{\sigma}\right] = \frac{1}{2}\begin{pmatrix} 1+\rho_3 & \rho_1+i\rho_2 \\ \rho_1-i\rho_2 & 1-\rho_3 \end{pmatrix} \ ,$$

$$\rho_i \in \mathbb{R} \ , \ \ \vec{\rho}^2 \leq 1 \ , \ \ \lambda_{\pm} = \frac{1}{2}\left(1 \pm \sqrt{\vec{\rho}^2}\right)$$

- Pure state:

$$\hat{\rho} = |\psi\rangle\langle\psi| \ \Leftrightarrow \ \hat{\rho}^2 = \hat{\rho} \ \Leftrightarrow \ \vec{\rho}^2 = 1$$

$$|\psi\rangle = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\varphi}|-\rangle \ , \ \ \vec{\rho} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

The Bloch Sphere

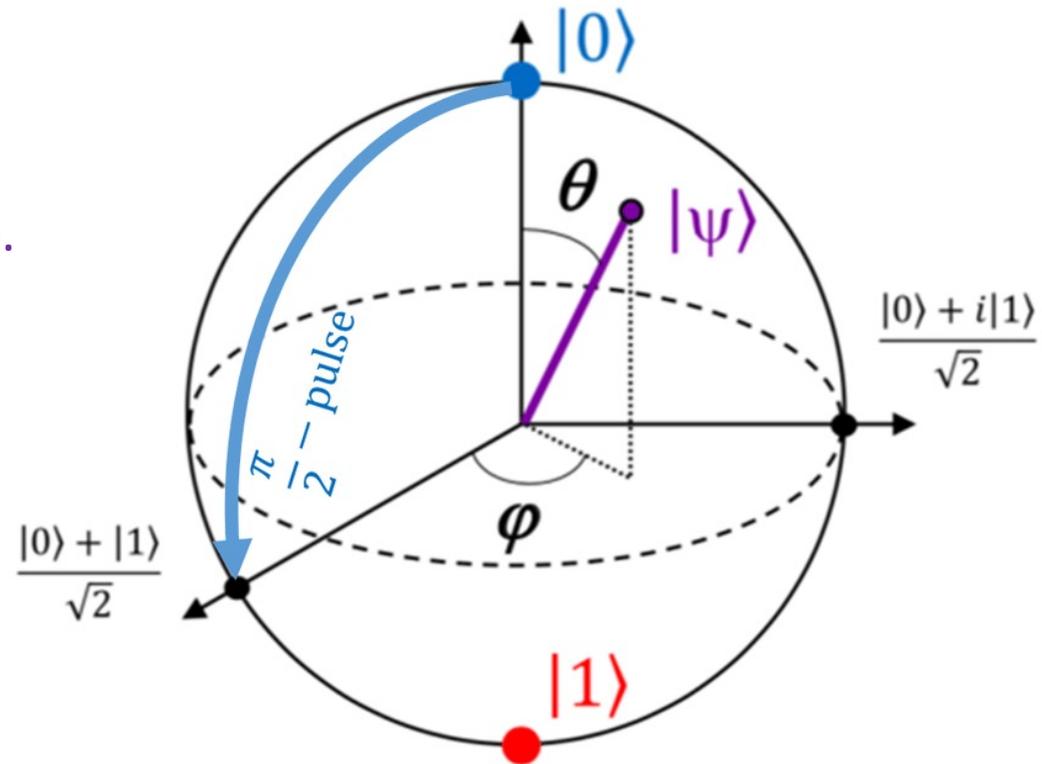Pure States have:

$||\rho||=1$
and hence live
on the surface of the sphere.

Otherwise interior
of the sphere:

$||\rho|| < 1$

$|0\rangle$

θ $|\psi\rangle$

$\frac{|0\rangle + i|1\rangle}{\sqrt{2}}$

$\frac{\pi}{2} - pulse$

φ

$\frac{|0\rangle + |1\rangle}{\sqrt{2}}$

$|1\rangle$

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

Single qubit: measurements

- Measure a specific polarization (think of photons) specified by a unit vector **n**

- Prepare N_n qubits identically and measure **n.σ**

$$p_{\vec{n}}(\uparrow) = \frac{N_n^{\uparrow}}{N_n^{\uparrow} + N_n^{\downarrow}} \quad , \quad p_{\vec{n}}(\downarrow) = \frac{N_n^{\downarrow}}{N_n^{\uparrow} + N_n^{\downarrow}} \quad , \quad \langle \vec{n} \cdot \vec{\sigma} \rangle = \frac{N_n^{\uparrow} - N_n^{\downarrow}}{N_n^{\uparrow} + N_n^{\downarrow}}$$

- Allows one to determine the component of the vector ρ along **n**: $\langle \vec{n} \cdot \vec{\sigma} \rangle = \mathrm{tr}\left(\vec{n} \cdot \vec{\sigma} \, \hat{\rho} \right) = \vec{n} \cdot \vec{\rho}$

- We need to measure along 3 non-collinear directions

- Note: this in fact corresponds to 6 projective measurements (hence overcomplete) – a POVM with 4 independent operators is actually enough

Issues with direct inversion

- For a finite set of measurements $\{N_x^\uparrow, N_x^\downarrow, N_y^\uparrow, N_y^\downarrow, N_z^\uparrow, N_z^\downarrow\}$
- The reconstructed density matrix:

  is not necessarily positive semi-definite:

$$\rho_i = \frac{N_i^\uparrow - N_i^\downarrow}{N_i^\uparrow + N_i^\downarrow} \;,\; (i = x, y, z) \qquad p_{i\uparrow} = \frac{N_i^\uparrow}{N_i^\uparrow + N_i^\downarrow} \;,\; p_{i\downarrow} = \frac{N_i^\downarrow}{N_i^\uparrow + N_i^\downarrow}$$

- Probability of a measurement:

$$P(\{N_x^\uparrow, N_x^\downarrow, N_y^\uparrow, N_y^\downarrow, N_y^\uparrow, N_y^\downarrow\}) = \binom{N_x}{N_x^\uparrow} p_{x\uparrow}^{N_x^\uparrow} p_{x\downarrow}^{N_x^\downarrow} \binom{N_y}{N_y^\uparrow} p_{y\uparrow}^{N_y^\uparrow} p_{y\downarrow}^{N_y^\downarrow} \binom{N_z}{N_z^\uparrow} p_{z\uparrow}^{N_z^\uparrow} p_{z\downarrow}^{N_z^\downarrow}$$

- For 30 measurements along each direction, the probability of an unphysical density matrix is 3.10^{-7} for a system with $\rho=\text{Id}/2$, but it is 98% for the pure state |z> !
- cf. R.Schmied *Quantum state tomography of a single qubit:comparison of methods*, J. Mod Optics, 2016 arXiv:1407.4759

QST with other methods than NNs (cont'd)

- Several methods have been proposed such as:
- `Maximum Likelihood' (Hradil, PRA55, R1561 (1997)) maximizes cost function subject to positivity constraint
- `Compressed sensing' Gross et al. PRL 105, 150401 (2010) Reconstruction of density matrix of dimension D and rank r using $O(rD \log^2 D)$ measurements
- Etc.

QST with NNs: Three articles discussed in some details in this lecture

[Torlai et al. 2018]

Neural-network quantum state tomography

Giacomo Torlai[1,2], Guglielmo Mazzola[3], Juan Carrasquilla[4,5], Matthias Troyer[3,6], Roger Melko[1,2] and Giuseppe Carleo[3,7]*

[Torlai et al. 2019]

Editors' Suggestion

Integrating Neural Networks with a Quantum Simulator for State Reconstruction

Giacomo Torlai[1,2,3], Brian Timar,[4] Evert P. L. van Nieuwenburg,[4] Harry Levine,[5] Ahmed Omran,[5] Alexander Keesling,[5] Hannes Bernien,[6] Markus Greiner,[5] Vladan Vuletić,[7] Mikhail D. Lukin,[5] Roger G. Melko,[2,3] and Manuel Endres[4]

Tutorial

[Carrasquilla-Torlai 2021]

How To Use Neural Networks To Investigate Quantum Many-Body Physics

Juan Carrasquilla[1] and Giacomo Torlai[2,3,]*

QST with a RBM for a `stoquastic' Hamiltonian

- *Stoquastic* Hamiltonian: sum of operators such that in some basis all off-diagonal matrix elements are negative or zero

$$\hat{H} = \sum_\alpha H_\alpha \ , \ \ (H_\alpha)_{i \neq j} \leq 0 \ , \ \ \forall i, j \qquad \text{S.Bravyi, 2015}$$

- Simulating H is sign-problem free

- The Boltzmann weight is a non-negative matrix:

$$e^{-\beta \hat{H}} = \mathbb{I} - \beta \hat{H} + \frac{1}{2}(\beta \hat{H})^2 + \cdots$$

$$\lim_{\beta \to \infty} e^{-\beta \hat{H}} = |\Psi_0\rangle\langle\Psi_0|$$

Example:
Transverse Field Ising Model

$$J \sum_{\langle ij \rangle} S_i^z S_j^z - h \sum_i S_i^x \ , \ \ (h \geq 0)$$

- Hence up to a global phase, the ground-state has non-negative amplitudes

QST/RBM – stoquastic system in a pure state

Assume that the system is in a pure state
(generalization to density matrix with RBMs: Torlai and Melko, PRL 2018)

RBM parametrization: $\quad |\Psi_\theta\rangle = \sum_s \sqrt{p_\theta(s)}\, |s\rangle \;,\;\; s \equiv (s_1, \cdots, s_N)$
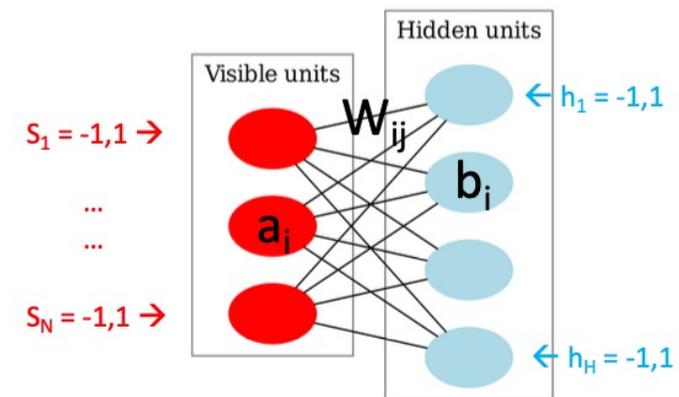
So here, we need to determine `only' 2^N positive real numbers and hence we can limit ourselves to measurements in the computational basis $|s\rangle$

$$p_\theta(s,h) = \frac{1}{Z_\theta} e^{\sum_i a_i s_i + \sum_j b_j h_j + \sum_{ij} W_{ij} s_i h_j}$$

$$\theta = \{a_i, b_j, W_{ij}\}$$

$$p_\theta(s) = \sum_h p_\theta(s,h)$$

$$= \frac{1}{Z_\theta} e^{\sum_{i=1}^N a_i s_i} \prod_{j=1}^H 2\cosh\left[b_j + \sum_i W_{ij} s_i \right]$$

Measurements dataset D is generated according to the (unknown) probability distribution q(σ) associated with the true state of the system. Kullback-Leibler divergence:

$$D_{\mathrm{KL}}(q|p_\theta) = \sum_s q(s) \ln \frac{q(s)}{p_\theta(s)}$$

$$D_{\mathrm{KL}}(q|p_\theta) \simeq \frac{1}{|D|} \sum_{\sigma \in D} \ln \frac{q(\sigma)}{p_\theta(\sigma)}$$

$$= \frac{1}{|D|} \sum_{\sigma \in D} \ln q(\sigma) - \frac{1}{|D|} \sum_{\sigma \in D} \ln p_\theta(\sigma)$$

Hence we have to minimize the cost function (second term only):

$$\boxed{C(\theta) = -\frac{1}{|D|} \sum_{\sigma \in D} \ln p_\theta(\sigma)}$$

Because the normalization matters, the computation of gradients require some care.

Notations: $$p_\theta(s) = \frac{1}{Z_\theta} e^{E_\theta(s)}$$

$$E_\theta = \sum_{i=1}^{N} a_i s_i + \sum_{j=1}^{H} \ln\left[2\cosh\Theta_j(s)\right]$$

$$\Theta_j(s) \equiv b_j + \sum_i W_{ij} s_i$$

Gradient:

$$\nabla_\theta \ln p_\theta(\sigma) = \nabla_\theta E_\theta(\sigma) - \frac{1}{Z_\theta} \sum_s \nabla_\theta E_\theta(s) \frac{1}{Z_\theta} e^{E_\theta(s)}$$

$$\boxed{\begin{aligned} \nabla_\theta \ln p_\theta(\sigma) &= \nabla_\theta E_\theta(\sigma) - \langle \nabla_\theta E_\theta \rangle_{p_\theta} \\ \nabla_\theta C(\theta) &\simeq \sum_{s/p_\theta(s)} \nabla_\theta E_\theta(s) - \frac{1}{|D|} \sum_{\sigma \in D} \nabla_\theta E_\theta(\sigma) \end{aligned}}$$

The second term is just a sum over the dataset, but the first one requires Monte Carlo sampling over configurations (s) generated by the RBM

Reminder – RBM: Explicit expression of the gradients

$$E_\theta = \sum_i a_i s_i + \sum_j \ln \left[2 \cosh \left(b_j + \sum_i s_i W_{ij} \right) \right]$$

$$\frac{\partial}{\partial a_i} E_\theta(\{s\}) = s_i$$

$$\frac{\partial}{\partial b_j} E_\theta = \tanh \Theta_j(\{s\})$$

$$\frac{\partial}{\partial W_{ij}} E_\theta = s_i \, \tanh \Theta_j(\{s\})$$

$$\Theta_j(\{s\}) \equiv b_j + \sum_i s_i W_{ij}$$

Natural Gradient Descent – Metric Tensor (Reminder from Lecture 1)

Using the KL divergence as a (pseudo-) distance in parameter space:

$$g_{lm} \simeq \frac{1}{|D|} \sum_{\sigma \in D} \frac{\partial \ln p_\theta(\sigma)}{\partial \theta_l} \frac{\partial \ln p_\theta(\sigma)}{\partial \theta_m}$$

NGD:
$$\delta\theta = -\epsilon \frac{g^{-1} \nabla_\theta C}{\sqrt{\langle \nabla_\theta C | g^{-1} | \nabla_\theta C \rangle}}$$

$$\nabla_\theta \ln p_\theta(\sigma) = \nabla_\theta E_\theta(\sigma) - \langle \nabla_\theta E_\theta \rangle_{p_\theta}$$

$$\nabla_\theta C(\theta) \simeq \sum_{s/p_\theta(s)} \nabla_\theta E_\theta(s) - \frac{1}{|D|} \sum_{\sigma \in D} \nabla_\theta E_\theta(\sigma)$$

Illustration on a model pure state
(from Torlai et al 2018)

$$|\Psi_W\rangle = \frac{1}{\sqrt{N}} \{|10\cdots0\rangle + |01\cdots0\rangle + \cdots + |0\cdots01\rangle\}$$

`Measurement' dataset generated by generating N_s random samples from this state
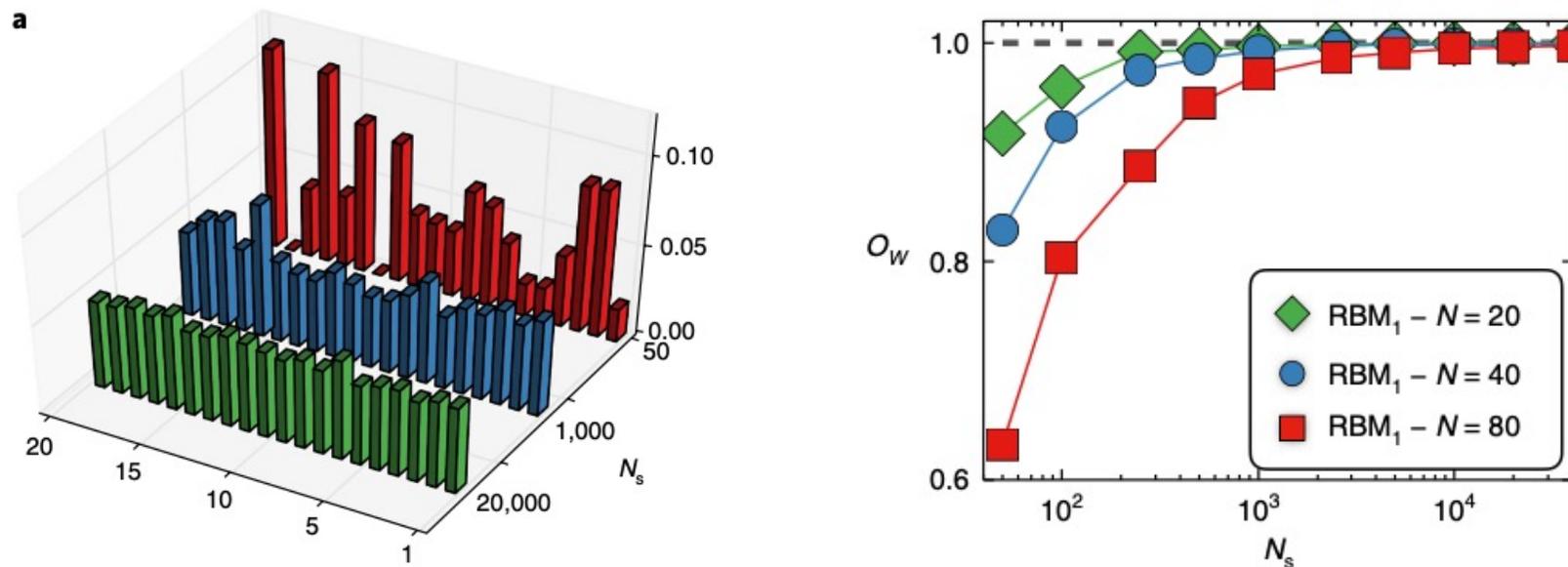
Reconstructed distribution over configurations
and overlap as a function of number of samples in dataset

Application to an array of Rydberg atoms (TFIM)

For a recent review of Rydberg atoms arrays, see:

Many-body physics with individually controlled Rydberg atoms

Antoine Browaeys [ID] * and Thierry Lahaye

Recent decades have witnessed great developments in the field of quantum simulation—where synthetic systems are built and studied to gain insight into complicated, many-body real-world problems. Systems of individually controlled neutral atoms, interacting with each other when excited to Rydberg states, have emerged as a promising platform for this task, particularly for the simulation of spin systems. Here, we review the techniques necessary for the manipulation of neutral atoms for the purpose of quantum simulation—such as quantum gas microscopes and arrays of optical tweezers—and explain how the different types of interactions between Rydberg atoms allow a natural mapping onto various quantum spin models. We discuss recent achievements in the study of quantum many-body physics in this platform, and some current research directions beyond that.

Also: Wu et al. arXiv:2012.10614 (Chinese Phys. B, 2021)

Rydberg atom: Highly excited state with very high principal quantum number

Large radius of electronic orbital (remember: radius scales as n^2)
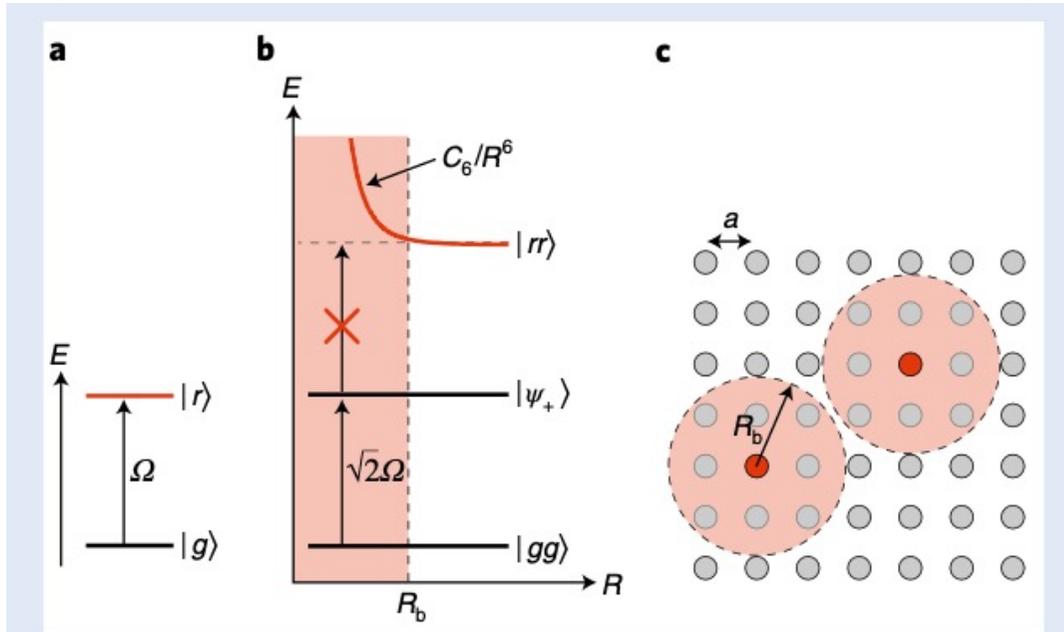
The Rydberg blockade. a, The ground and Rydberg states $|g\rangle$ and $|r\rangle$ are coupled by a resonant laser with Rabi frequency Ω. **b,** For two atoms separated by a distance $R < R_b$, the collective ground state $|gg\rangle$ is coupled only to $|\psi_+\rangle = (|gr\rangle + |rg\rangle)/\sqrt{2}$, but not to $|rr\rangle$, which is shifted out of resonance by the van der Waals interaction U_{vdW}. **c,** In a large ensemble of atoms, for example a regular array with spacing a, an atom excited in $|r\rangle$ (red dot) prevents the excitation of all the atoms contained in a sphere of radius R_b.

Two Rydberg atoms - Browaeys and Lahaye, Nat Phys(2020)
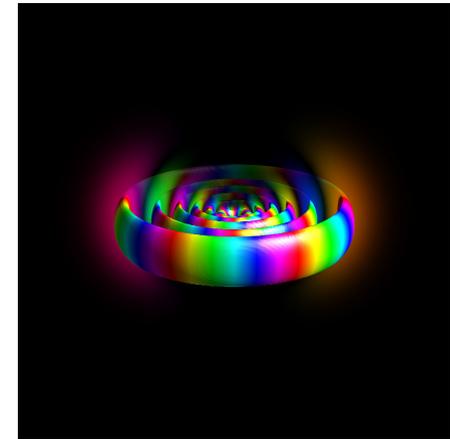
Electronic orbital n=12
(from Wikipedia)

Quantum information
with Rydberg atoms
Lukin and coworkers
PRL 85, 2208 (2000);
PRL 87, 037901 (2001)
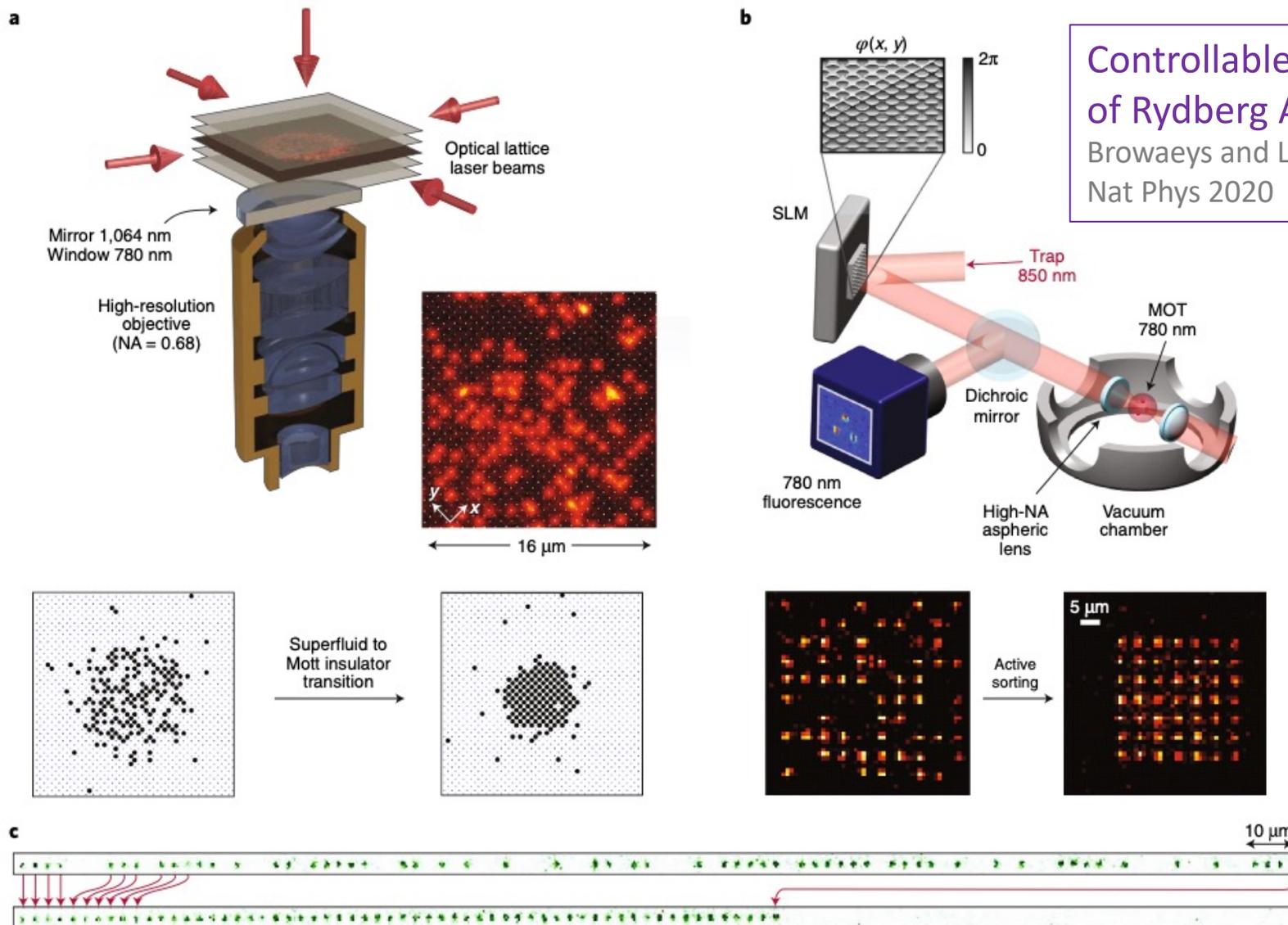
Controllable Arrays of Rydberg Atoms
Browaeys and Lahaye
Nat Phys 2020

Fig. 1 | Experimental platforms for realizing arrays of individually controlled neutral atoms. a, In a quantum gas microscope, a high-numerical-aperture (NA) objective is used to observe the fluorescence of ultracold atoms trapped in an optical lattice obtained by interfering several laser beams. To achieve a filling of exactly one atom per site, one drives the superfluid-to-Mott-insulator transition[34]. **b**, In the tweezer array platform, an SLM imprints an appropriate phase on a trapping beam before focusing with a high-NA lens, resulting in arrays of traps with almost arbitrary geometries[35]. Single, laser-cooled atoms are loaded in the optical tweezers from a magneto-optical trap, resulting in a random loading array at half filling, which can be actively reordered into a target array using a moving optical tweezers[37]. Bottom: single shot fluorescence image of an array of traps before and after assembly. **c**, Alternatively, in one dimension, the tweezers can be generated using an acousto-optic deflector fed with multiple radio-frequency tones, which allows rearrangment of the atoms in a single step[38]. Panels adapted from: **a**, ref. [34], Springer Nature Ltd; **b**, top, ref. [35] under a Creative Commons licence (https://creativecommons.org/licenses/by/3.0/); **b**, bottom, ref. [37], AAAS; **c**, ref. [38], AAAS.

Hamiltonian for atoms placed in the same Rydberg state:
Transverse field Ising model (TFIM)

Effective spin-1/2 degree of freedom on each site: $|g\rangle = |\downarrow\rangle \ , \ |r\rangle = |\uparrow\rangle$

$$\hat{n}_i = \frac{1}{2}(1 + \hat{S}_i^z) \ , \ \hat{n}|g\rangle = 0 \ , \hat{n}|r\rangle = |r\rangle \ V_{ij} = \frac{C_6}{R_{ij}^6}$$

$$\hat{H} = \frac{\hbar\Omega}{2}\sum_i \hat{S}_i^x - \hbar\delta \sum_i \hat{n}_i + \sum_{i<j} V_{ij}\hat{n}_i\hat{n}_j$$

Rabi drive

Detuning
(energy difference
Between g and r)

r-r
Interaction
(Rydberg
blockade)

If two different Rydberg states: XY model + fields

Pascal Scholl[1,6 ✉], Michael Schuler[2,6], Hannah J. Williams[1,6], Alexander A. Eberharter[3,6], Daniel Barredo[1,4], Kai-Niklas Schymik[1], Vincent Lienhard[1], Louis-Paul Henry[5], Thomas C. Lang[3], Thierry Lahaye[1], Andreas M. Läuchli[3] & Antoine Browaeys[1]
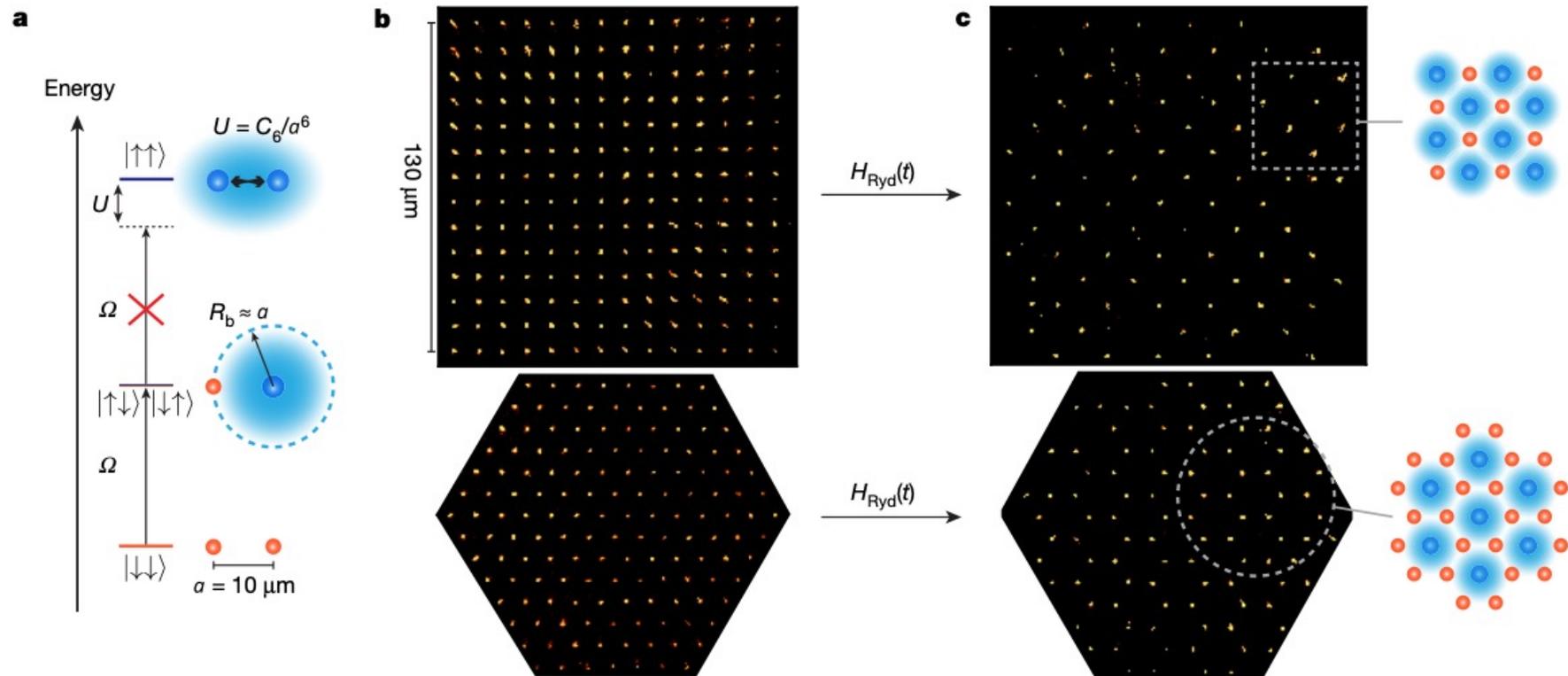
Fig. 1 | Emergence of AF ordering from the Rydberg blockade in square and triangular arrays. a, Illustration of the Rydberg blockade with two atoms, whereby the strong interactions prevent the simultaneous excitation of two atoms from the ground state (red circles) to the Rydberg state (blue circles) within the Rydberg blockade radius R_b at which $U = \hbar\Omega$. **b, c**, Single-shot fluorescence images of ground state ($|\downarrow\rangle$) atoms in a 14 × 14 square array (top) and a 147-atom triangular array (bottom) with an atomic separation of $a = 10\ \mu m$. **b**, Initial PM states. **c**, Nearly perfect AF ordering.
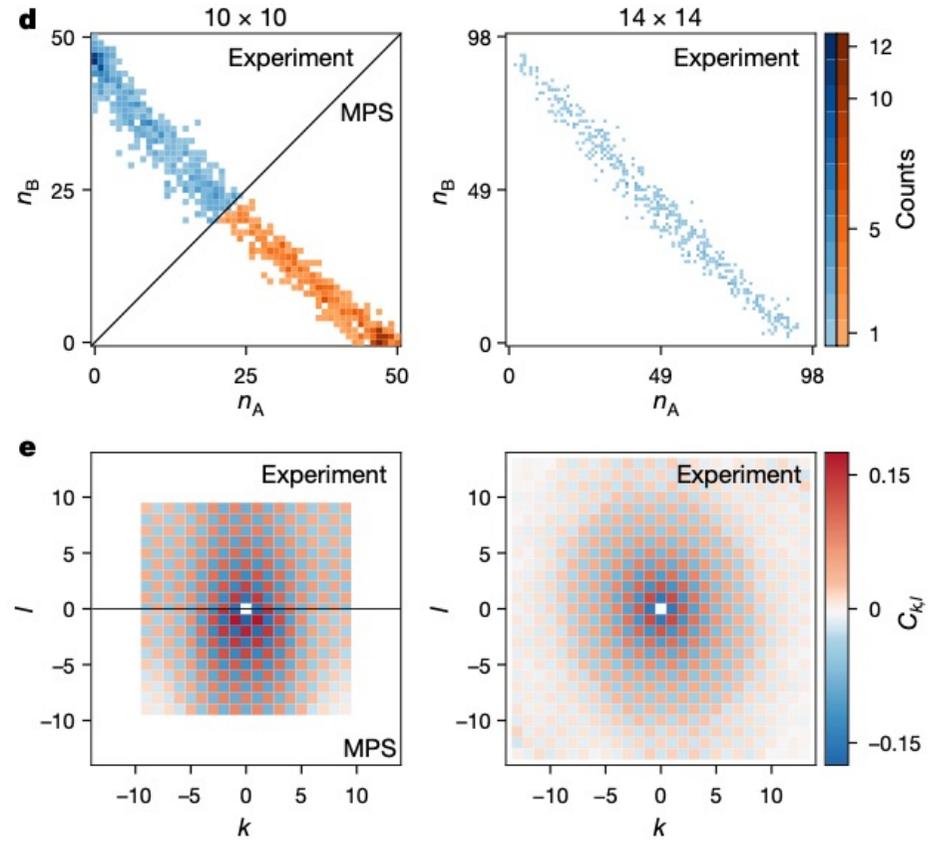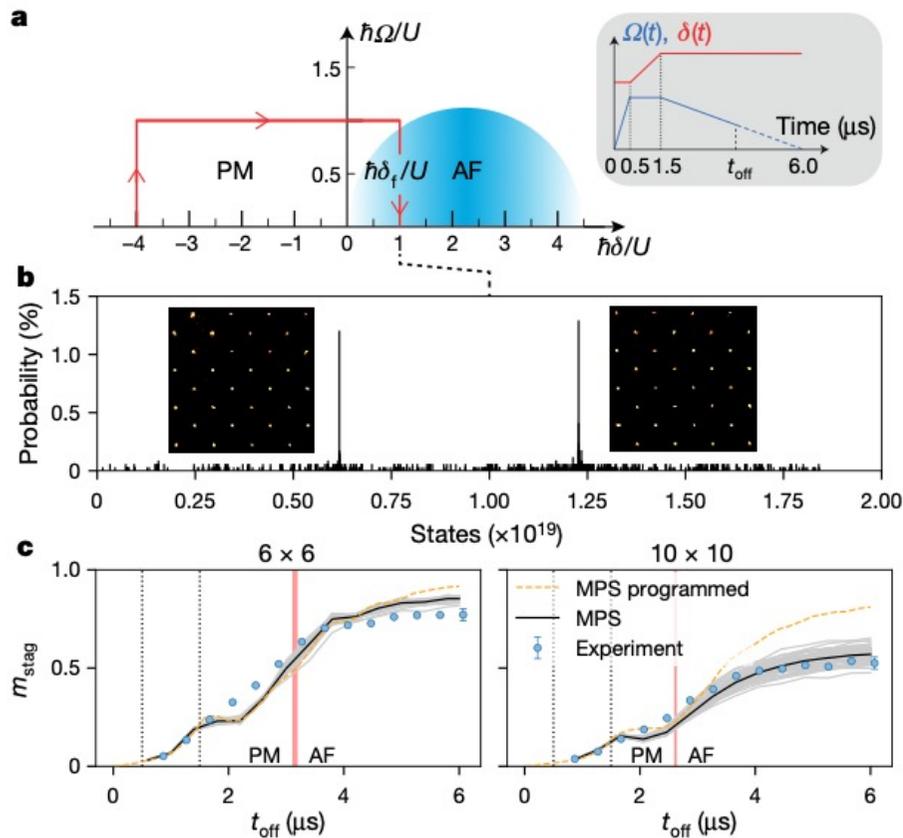
Fig. 2 | The Ising model on a square lattice. a, Sketched bulk phase diagram for the square lattice. The inset shows the sweep shape, with t_{off} the switch-off time of the excitation laser. The corresponding trajectory in the phase diagram is shown as a red arrow. **b**, State histogram for the 8×8 array at the end of the sweep. The insets show fluorescence images of the two perfect AF states, which are obtained with 2.5% probability. **c**, Growth of the staggered magnetization during the sweep for the 6×6 array (left) and the 10×10 array (right). The blue circles are experimental results with standard errors on the mean smaller than the markers size. The error bar on the final point is indicative of the long-term stability of the experimental setup (see 'Long-term stability' section). MPS simulations without (dashed line) and with (grey lines) experimental imperfections for which 50 (6×6 array) and 77 (10×10 array) disorder instances are shown, with their average shown in black. The vertical dotted lines correspond to the turning points in the sweep. The vertical red regions depict the finite-size phase boundaries between the PM and AF phases, obtained from the inflection point of m_{stag} using ground state density matrix renormalization group. **d**, Final staggered magnetization histograms for the 10×10 array (left) and the 14×14 array (right). **e**, The correlation maps for the 10×10 array (left) and the 14×14 array (right), with the MPS results shown in the lower half for the 10×10 array.

Scholl et al. Nature 2021

Tomography for an array of Rydberg atoms (TFIM) with simulated data

After Carrasquilla and Torlai PRXQ 2021 (code walk-through using NetKet); see Torlai et al 2018 for first study of TFIM

Define Hamiltonian: square L*L array L=8 (64 sites), $(V,\Omega,\delta)=(3,1,2)$MHz

```python
from mpi4py import MPI
import netket as nk

# Rydberg Hamiltonian parameters
L = 8          # linear size
V = 3.0        # Van der Waals interaction
Omega = 1.0    # Rabi frequency
delta = 2.0    # detuning
```

```
# Neural-network wavefunction
rbm = nk.machine.RbmSpinReal(hilbert=hilbert,
    alpha=1)
```

Density $\alpha = H/L^2$

```
# Monte Carlo sampler
sa = nk.sampler.MetropolisLocal(machine=rbm)


# Optimizer
op = nk.optimizer.AdaDelta(rho=0.95, epscut =
    1.0e-7)


# Initialize tomography object
qst = nk.Qsr(
        sampler = sa,
        optimizer = op,
        n_samples_data = 1000,
        n_samples = 2000,
        rotations = rotations,
        samples = samples,
        bases = bases,
        sr = None)
```

Number of training samples (data)
Number of Metropolis generated samples
to compute the gradients at each step

No basis rotations
(TFIM is *stoquastic*)

[Carrasquilla-Torlai PRXQ 2021]

```
# define observable for measurements
H = generatehamiltonian(hilbert, L, L, V, Omega,
      delta)
qst.add_observable(H, "H")

# run quantum state tomography
for ep in qst.iter(epochs):
    obs = qst.get_observable_stats()
```


FIG. 4. Quantum state tomography of an 8×8 array of Rydberg atoms with unsupervised learning of single-shot atomic occupation data. We compare various observables measured using the MPSs obtained from DMRG (black line) and the neural-network wave functions learned from data generated at different values of the detuning δ. We plot the average energy per spin (a) and the average magnetization along the z axis (b) and x axis (c). The inset in (a) shows the relative error in the energy $\epsilon_{rel} = |E_{RBM} - E_{DMRG}|/|E_{DMRG}|$. Error bars estimated via standard deviations are too small to be visible.

[Carrasquilla-Torlai 2021]

a 2D TFIM

b 2D XXZ

1D TFIM

c

d

Torlai et al Nat Phys 2018
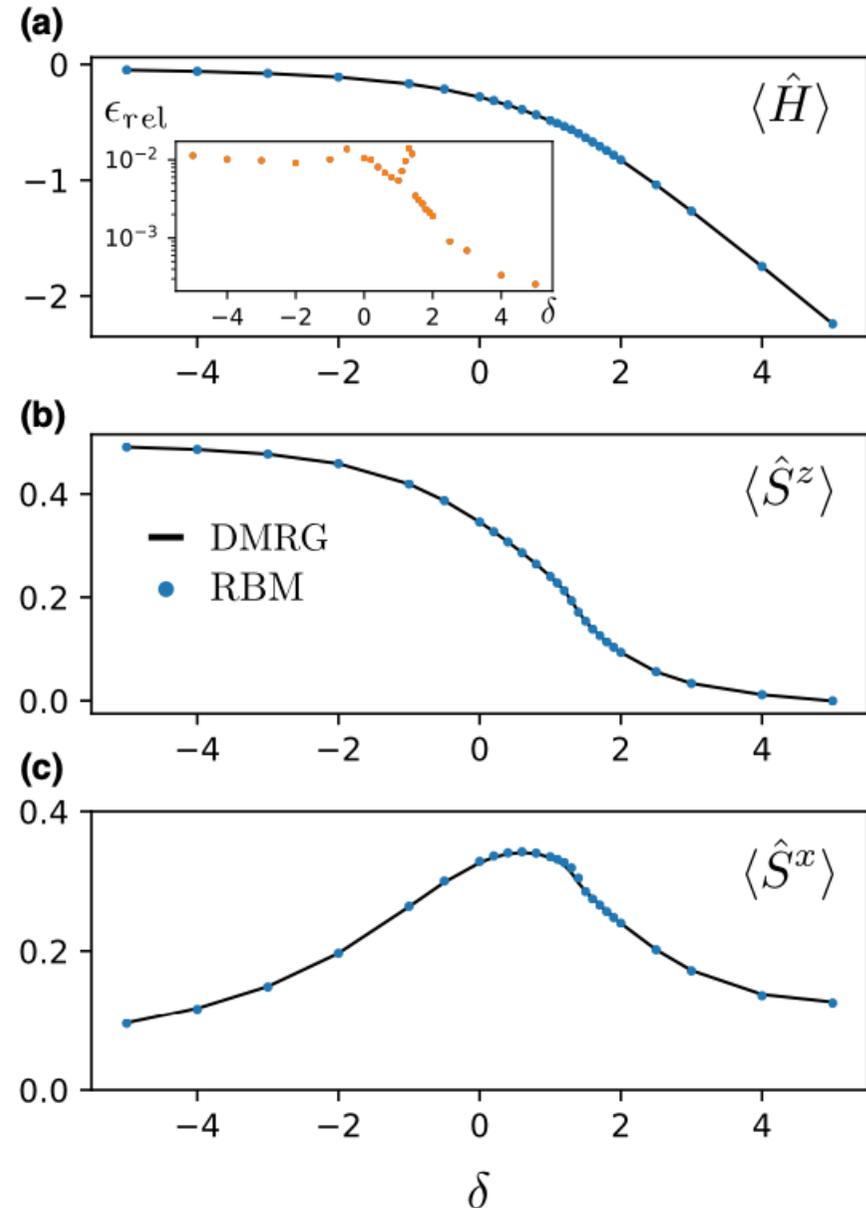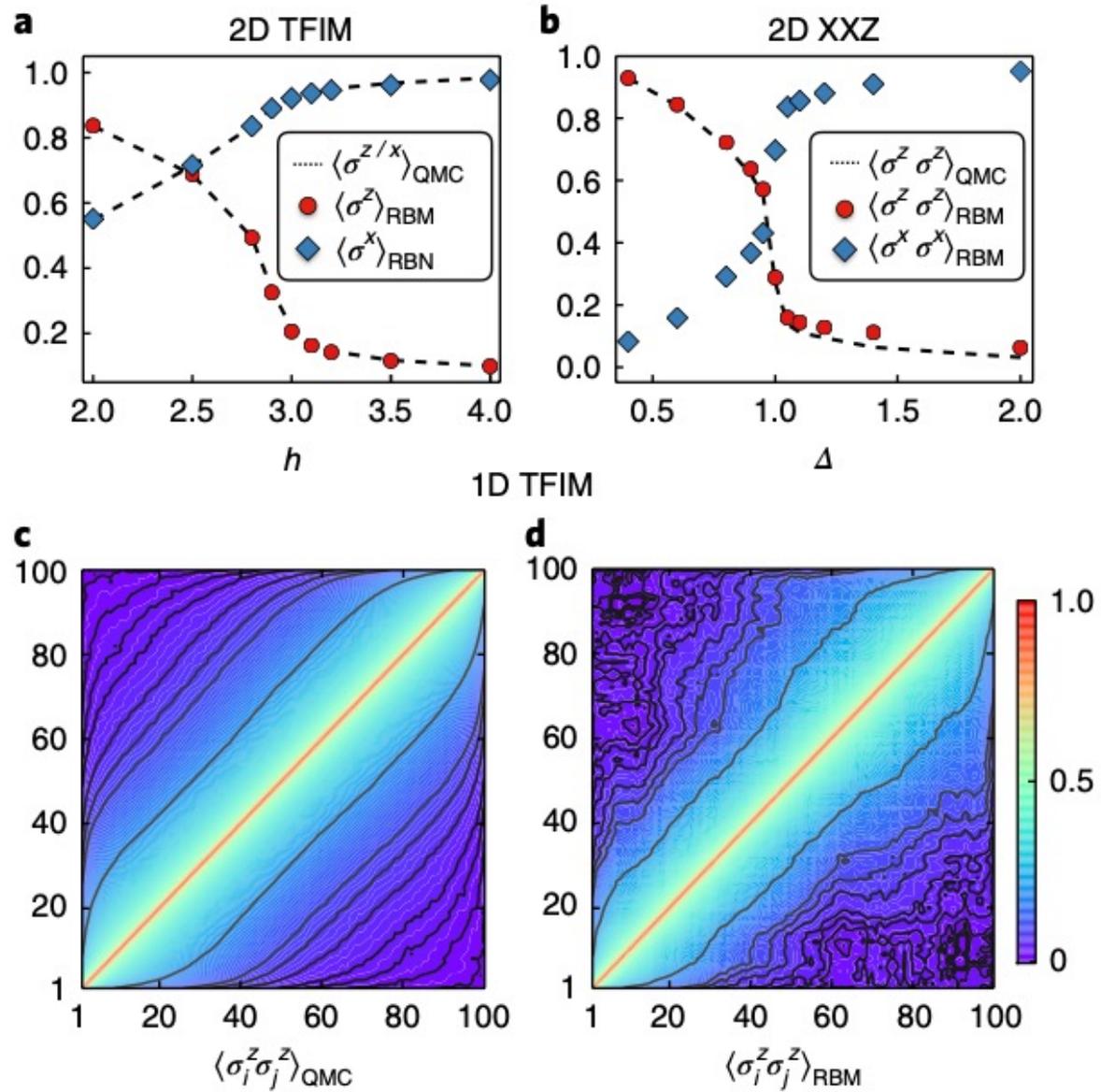
General case (not *stoquastic)*: Reconstructing the amplitude <u>and phase</u>

Parametrize both amplitudes and phase with two separate RBMs with real parameters (Torlai et al., 2018):

$$\psi_\theta(\mathbf{s}) = \sqrt{p_\theta(\mathbf{s})}\, e^{i\phi_{\theta'}(\mathbf{s})/2} \quad,\quad \phi_{\theta'}(\mathbf{s}) = \ln p_{\theta'}(\mathbf{s})$$
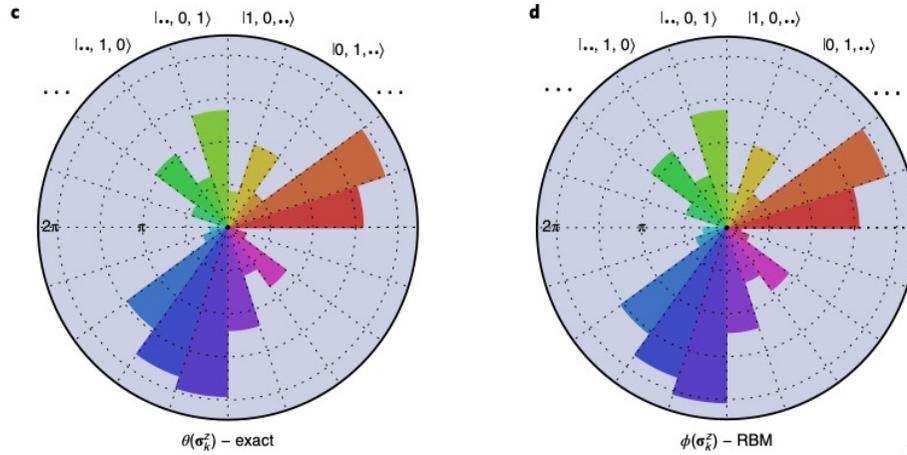
Or use complex biases and weights in the RBM (Carrasquilla and Torlai 2021)

We now need measurements in several bases (see above) and define the cost function (for each RBM) as:

$$C(\theta) = -\sum_{b=1}^{N_b} \frac{1}{|D_b|} \sum_{\sigma^{(b)} \in D_b} \ln p_\theta(\sigma^{(b)})$$

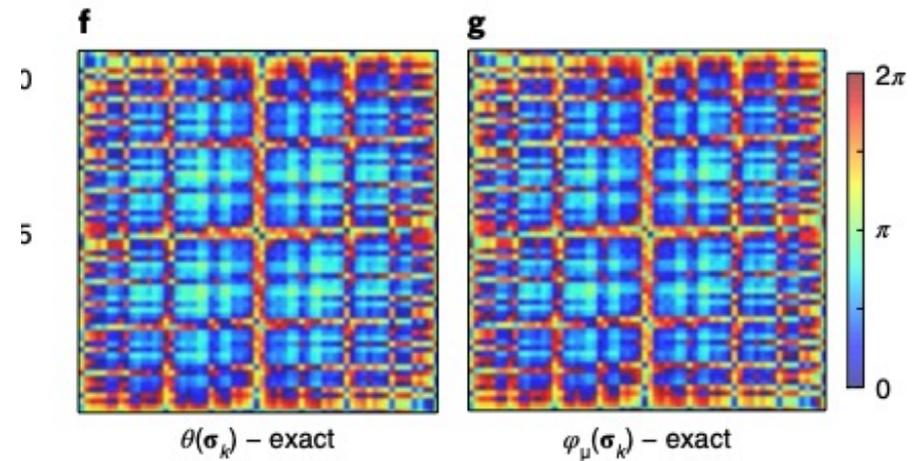To compute the gradients and metric, the NN wavefunction must be rotated to each basis according to:

$$\psi_\theta(\mathbf{s}^{(b)}) = \sum_{\mathbf{s}} U_b(\mathbf{s}, \mathbf{s}^{(b)})\, \psi_\theta(\mathbf{s})$$

Torlai et al., Nat Phys 2018

Toy problem: W-state
N=20 qubits with random phases
6400 samples/basis
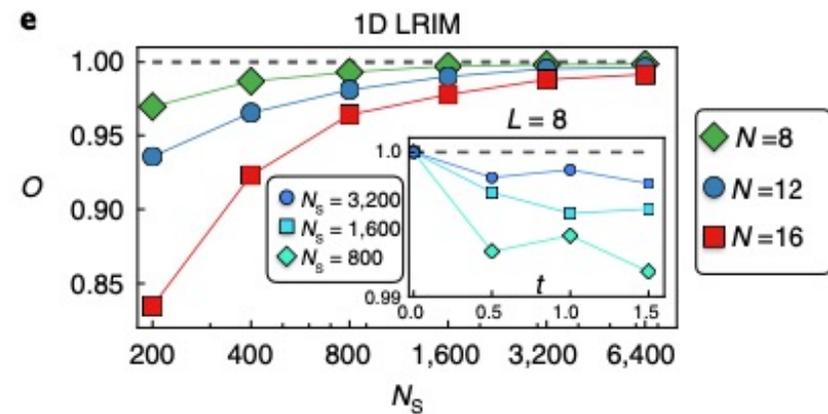
Training on projective
Measurements from
time evolution
TFIM with long-range
interactions

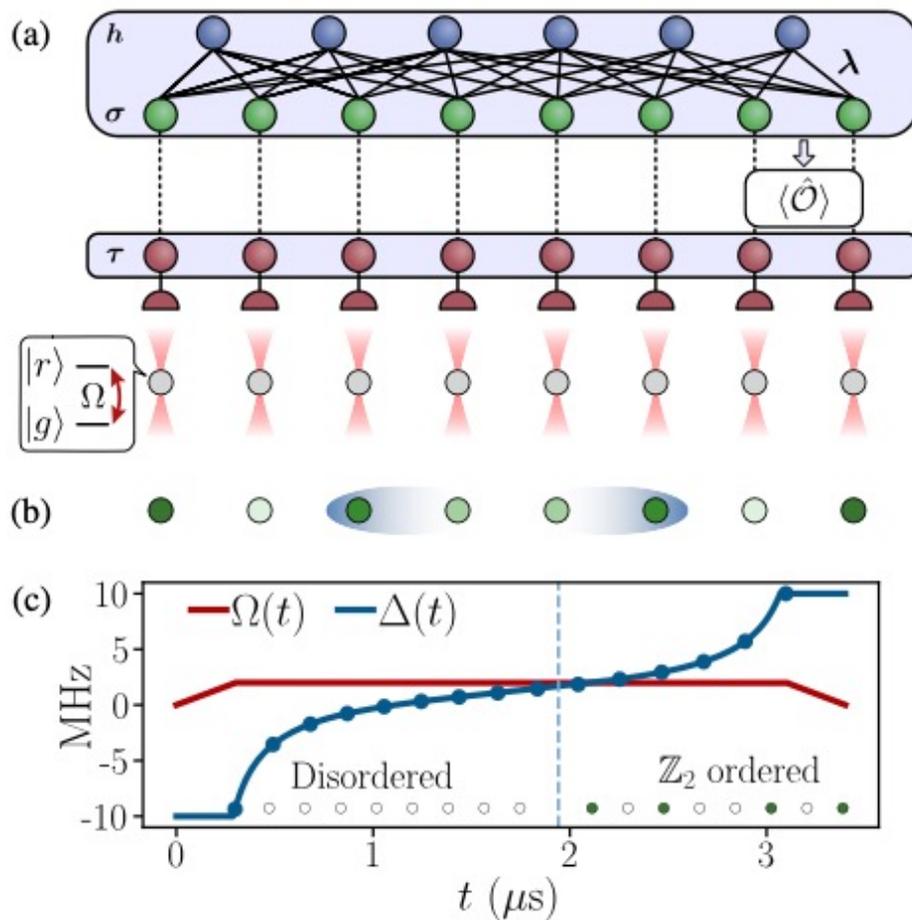# Rydberg arrays: tomography from experimental data Torlai et al. PRL 123, 230504 (2019) w/ M.Lukin's group

FIG. 1. Experiment and reconstruction. (a) Model of the reconstruction process. Individual ^{87}Rb atoms (gray circles) are trapped in an array of optical tweezers and coupled to a Rydberg state with Rabi frequency Ω. Site-resolved fluorescence imaging provides imperfect measurement in the $\hat{\sigma}^z$ basis. Our neural-network model describes the true quantum state as a RBM (blue and green neurons), while the binary data τ accessible to the experimentalist are included as an auxiliary "noise" layer (red neurons). By training on this data, the network learns parameters λ describing the experimental quantum state, which are subsequently used to compute observables $\langle \hat{\mathcal{O}} \rangle$. (b) Representation of the ordered state at the end of the adiabatic sweep; see Eq. (2).

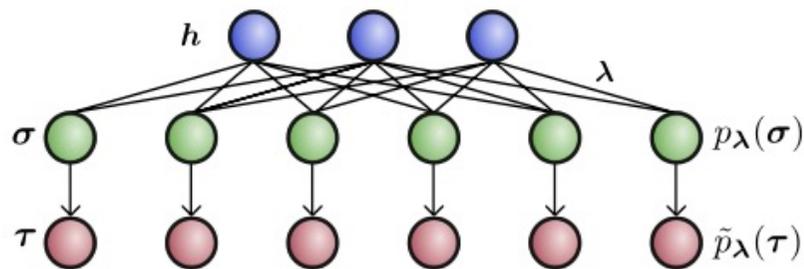# Measurement errors are taken into account by adding an intermediate layer

Figure 1. **Three layer model**. Schematic for how noise-corrupted data is modeled using a three-layer graph. The upper two layers h, σ constitute an RBM with trainable parameters λ, which defines a distribution $p_\lambda(\sigma)$ over the uncorrupted variables σ upon tracing out the hidden units h. The corrupted distribution is obtained through the noise process $p(\tau|\sigma)$ as $\tilde{p}_\lambda(\tau)$. The noise process is indicated here by arrows which link uncorrupted and corrupted variables at each site.

$$p(\tau|\sigma) = \frac{1}{\tilde{Z}}e^{\tilde{b}_\sigma \cdot \sigma + \tilde{b}_\tau \cdot \tau + \tilde{W}\sigma \cdot \tau}$$

$$\tilde{W} = \log \frac{p(1|1)p(0|0)}{p(1|0)p(0|1)}$$

$$\tilde{b}_{\sigma,i} = \log \frac{p(0|1)}{p(0|0)}$$

$$\tilde{b}_{\tau,i} = \log \frac{p(1|0)}{p(0|0)}$$

Conditional probabilities taken as given

The cost gradient takes a form nearly identical to that of the standard training method (8),

$$\nabla_\lambda \mathcal{L}_\lambda = \langle \nabla_\lambda \mathcal{E}_{\text{eff}}(\sigma) \rangle_{p_\lambda(\sigma)} - \frac{1}{|\mathcal{D}|} \sum_{\tau \in D} \langle \nabla_\lambda \mathcal{E}_{\text{eff}}(\sigma) \rangle_{\tilde{p}_\lambda(\sigma|\tau)} \quad (13)$$

The second term in the gradient update step is now computed not directly from the training set samples $\tau \in \mathcal{D}$, but rather from the Bayesian posterior distribution

$$\tilde{p}_\lambda(\sigma|\tau) = \frac{p(\tau|\sigma)p_\lambda(\sigma)}{\tilde{p}_\lambda(\tau)} \quad (14)$$

which the RBM assigns to visible states σ, given an observation τ in the noisy training set.
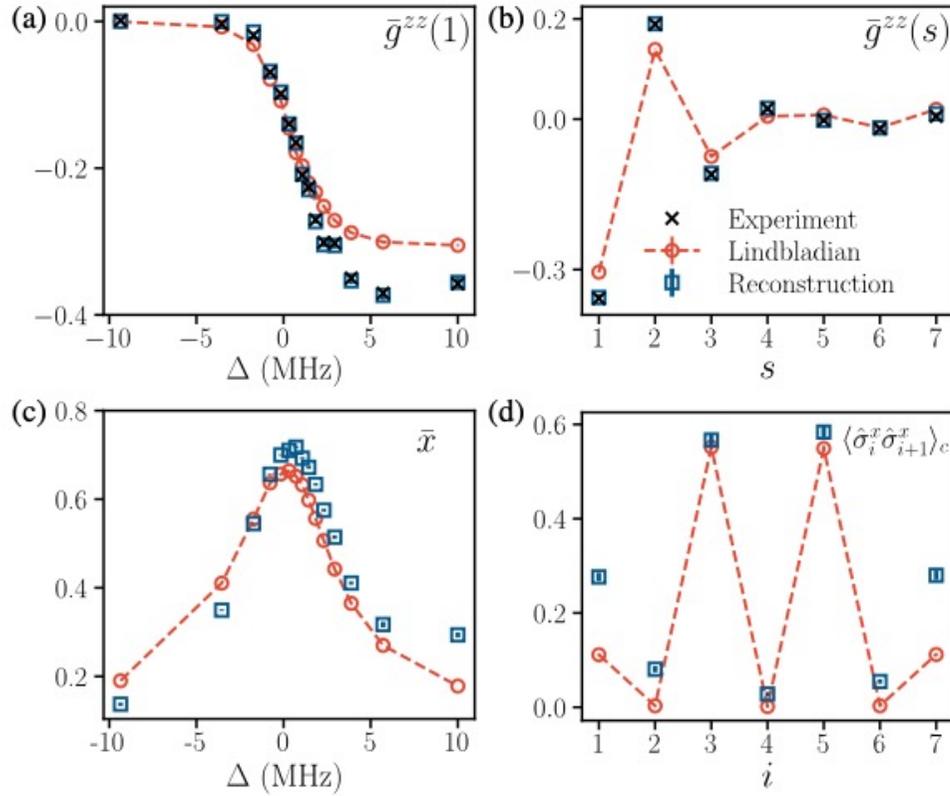
Torlai et al. PRL 2019

FIG. 3. Few-body observables. Comparison of the RBM reconstruction (squares) with the experiment results (crosses) and the predictions from the Lindbladian master equation (circles) [25]. In order to facilitate comparison with experiment, the values reported in (a) and (b) for the RBM and Lindbladian observables are computed including the known measurement error rates $p(0|1) = 0.04$, $p(1|0) = 0.01$. (a) Nearest-neighbor correlations $\bar{g}^{zz}(1)$ in the z basis, spatially averaged (see text for definition). (b) Average correlation $\bar{g}^{zz}(s)$ as a function of distance s for $\Delta = 10$ MHz. (c) Spatial average $\bar{x}$ of the transverse field $\langle \hat{\sigma}_i^x \rangle$. (d) Nearest-neighbor correlation $\langle \hat{\sigma}_i^x \hat{\sigma}_{i+1}^x \rangle_c$ as a function of position i for $\Delta = 10$ MHz. The two peaks correspond to the bonds highlighted in Fig. 1(b).

Torlai et al. PRL 2019

Some topics I have not discussed
in this introduction…

Further Reading…

Reconstruction of Entanglement/ Renyi Entropies [ask our speaker!]

$\psi_\lambda(\sigma_1) \otimes \psi_\lambda(\sigma_2)$ $\mathrm{Swap}_A \ \psi_\lambda(\sigma_1) \otimes \psi_\lambda(\sigma_2)$
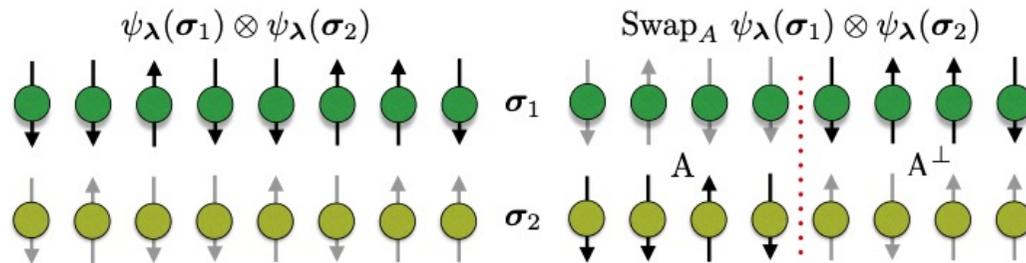
Figure 1. Replica trick for the evaluation of the second Renyi entropy.

$$S_\alpha(\rho_A) = \frac{1}{1-\alpha} \log[\mathrm{Tr}(\rho_A^\alpha)],$$

Fig. 3 | Reconstruction of the entanglement entropy for 1D lattice spin models. The second Renyi entropy as a function of the subsystem size ℓ for $N = 20$ spins. We compare results obtained using the the RBM wavefunctions (markers) with exact diagonalization (dashed lines) for the 1D TFIM at different values of the transverse magnetic field h and the 1D XXZ model with critical anisotropy $\Delta = 1$.

Torlai et al., Nat Phys 2018
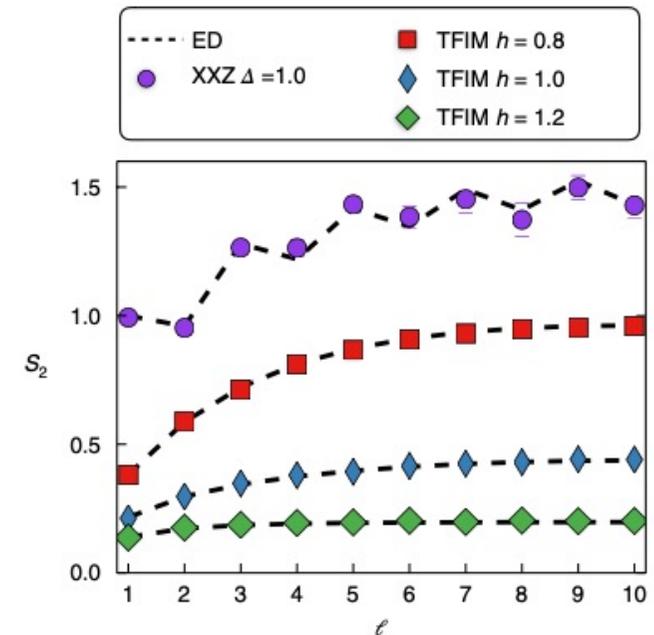
Density Matrix QST with NNs

Latent Space Purification via Neural Density Operators

Giacomo Torlai and Roger G. Melko

*Department of Physics and Astronomy, University of Waterloo, Ontario N2L 3G1, Canada,
and Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
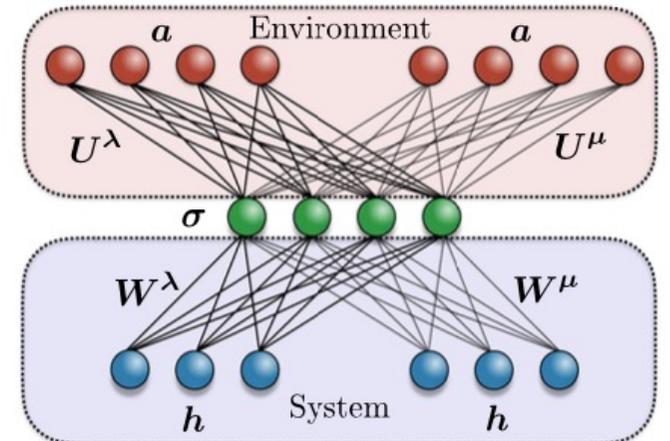
Machine learning is actively being explored for its potential to design, validate, and even hybridize with near-term quantum devices. A central question is whether neural networks can provide a tractable representation of a given quantum state of interest. When true, stochastic neural networks can be employed for many unsupervised tasks, including generative modeling and state tomography. However, to be applicable for real experiments, such methods must be able to encode quantum mixed states. Here, we parametrize a density matrix based on a restricted Boltzmann machine that is capable of purifying a mixed state through auxiliary degrees of freedom embedded in the latent space of its hidden units. We implement the algorithm numerically and use it to perform tomography on some typical states of entangled photons, achieving fidelities competitive with standard techniques.

$$\rho_\theta(\boldsymbol{\sigma}, \boldsymbol{\sigma}') = \sum_{\boldsymbol{a}} \psi_\theta(\boldsymbol{\sigma}, \boldsymbol{a}) \psi_\theta^*(\boldsymbol{\sigma}', \boldsymbol{a}).$$

$$\psi_{\lambda\mu}(\boldsymbol{\sigma}, \boldsymbol{a}) = Z_\lambda^{-\frac{1}{2}} \sqrt{p_\lambda(\boldsymbol{\sigma}, \boldsymbol{a})} e^{i\phi_\mu(\boldsymbol{\sigma}, \boldsymbol{a})},$$

$$p_\theta(\boldsymbol{\sigma}, \boldsymbol{a}) = e^{\sum_i \log(1 + e^{W_\theta^{[i]}\boldsymbol{\sigma} + c_\theta^{[i]}}) + \boldsymbol{a}^\top U_\theta \boldsymbol{\sigma} + b_\theta^\top \boldsymbol{\sigma} + d_\theta^\top \boldsymbol{a}},$$

Integrating NNs in quantum simulators

[ask our speaker!]

Precise measurement of quantum observables with neural-network estimators

Giacomo Torlai [1,*] Guglielmo Mazzola,[2] Giuseppe Carleo,[1] and Antonio Mezzacapo[3]
[1]*Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, USA*
[2]*IBM Research Zurich, Saumerstrasse 4, 8803 Ruschlikon, Switzerland*
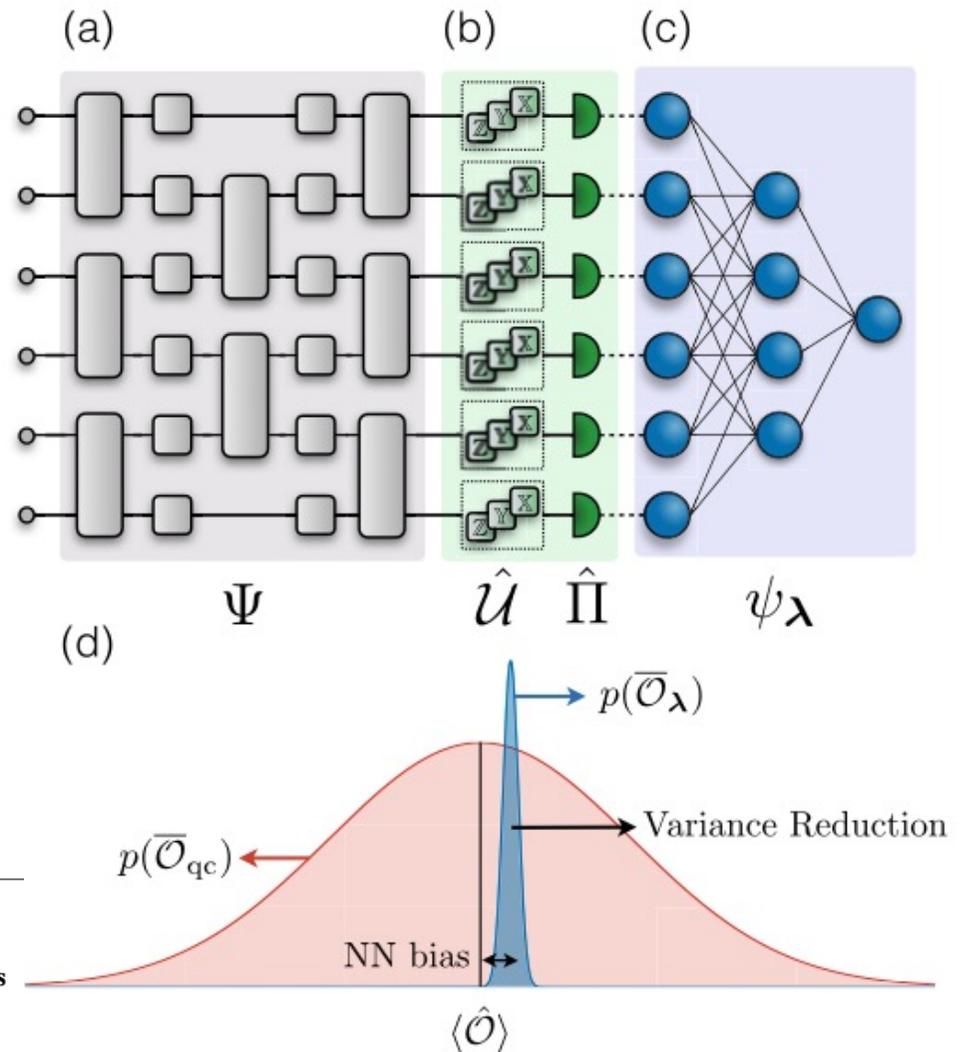[3]*IBM T.J. Watson Research Center, Yorktown Heights, New York 10598, USA*

FIG. 1. Measurements on quantum hardware with neural-network estimators. (a) A quantum circuit prepares a quantum state Ψ. (b) Single-qubit measurements, consisting of a local rotation $\hat{\mathcal{U}}$ and a projective measurement $\hat{\Pi}$. (c) A neural network is trained on the output of the measuring apparatus to discover a representation ψ_λ of the state Ψ that retrieves the expectation value of a target quantum observable $\hat{O}$. (d) The intrinsic measurement uncertainty is traded for a systematic reconstruction bias, leading to a measurement outcome distribution with lower variance.

Other NN architectures, e.g. CNN

Check for updates
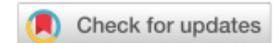
Efficient quantum state tomography with convolutional neural networks

Tobias Schmale[1]✉, Moritz Reh[1]✉ and Martin Gärttner[1,2,3]✉

Modern day quantum simulators can prepare a wide variety of quantum states but the accurate estimation of observables from tomographic measurement data often poses a challenge. We tackle this problem by developing a quantum state tomography scheme which relies on approximating the probability distribution over the outcomes of an informationally complete measurement in a variational manifold represented by a convolutional neural network. We show an excellent representability of prototypical ground- and steady states with this ansatz using a number of variational parameters that scales polynomially in system size. This compressed representation allows us to reconstruct states with high classical fidelities outperforming standard methods such as maximum likelihood estimation. Furthermore, it achieves a reduction of the estimation error of observables by up to an order of magnitude compared to their direct estimation from experimental data.