Introduction to Machine Learning Methods in Condensed Matter Physics

LECTURE 8, FALL 2021

Yi Zhang (张亿)

International Center for Quantum Materials, School of Physics
Peking University, Beijing, 100871, China

Email: frankzhangyi@pku.edu.cn

Restricted Boltzmann Machine

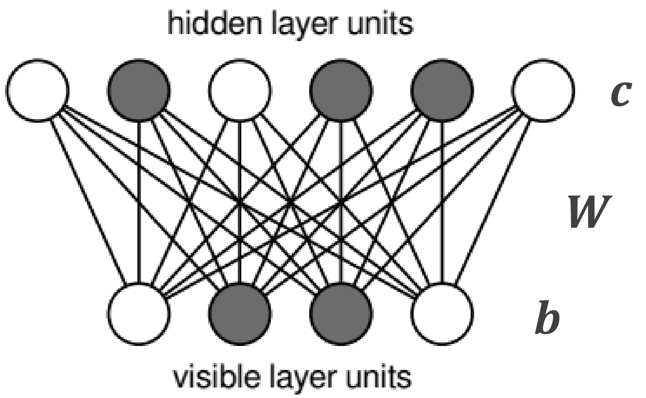
- A binary graphic model with a visible layer and hidden layer:

  The probability of a configuration: Boltzmann distribution

  $\boldsymbol{W}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ as model parameters

$$P(\boldsymbol{X},\boldsymbol{H}) = \frac{1}{Z}\exp\left(-E(\boldsymbol{X},\boldsymbol{H})\right)$$

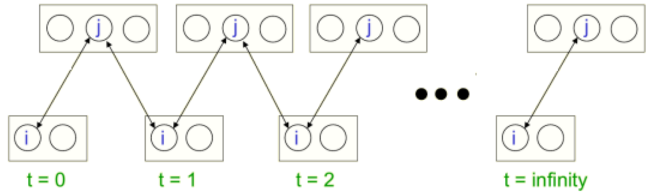
$$E(\boldsymbol{X},\boldsymbol{H}) = -\boldsymbol{X}^T\boldsymbol{b} - \boldsymbol{c}^T\boldsymbol{H} - \boldsymbol{X}^T\boldsymbol{W}\boldsymbol{H}$$



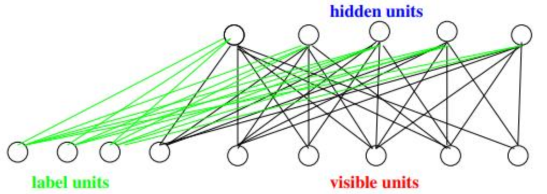
  Efficient sampling:

  Hidden layer:

  Visible layer:

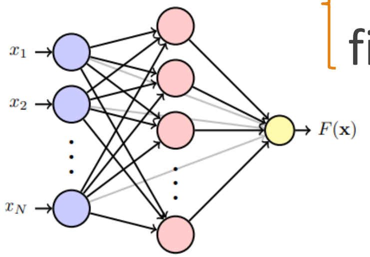


Train and interpolate with class labels:

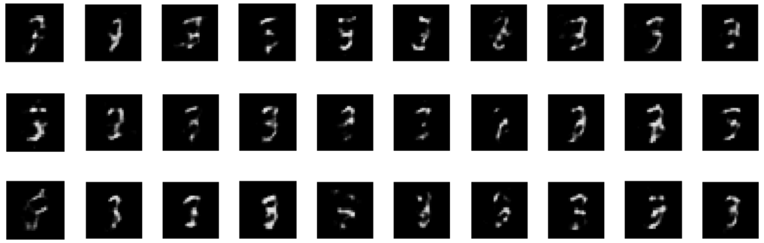


- Train RBM to: maximize the likelihood of the given data

  fit a given distribution



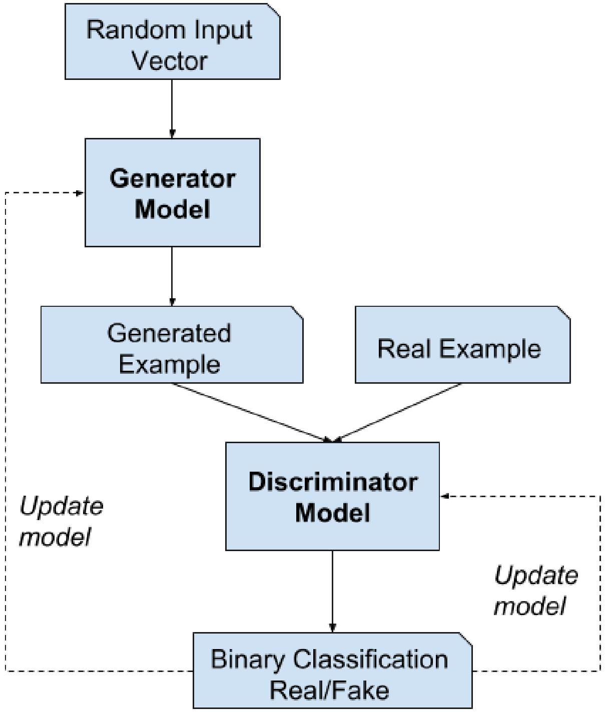
$$-F(\mathbf{x}) = \sum_{i=1}^{N} a_i x_i + \sum_{j=1}^{M} \ln(1 + e^{b_j + \sum_{i=1}^{N} x_i W_{ij}})$$



Generative Adversarial Network

- Two neural networks contest with each other in a zero-sum game:
  - Generator: generate new plausible examples from the training dataset
  - Discriminator. Model that is used to classify examples as real (from the training set) or fake (generated)
- Data augmentation: better performing models
  - Ideal convergence: the generator generates perfect replicas, and the discriminator cannot tell the difference

Example of the Progression in the Capabilities of GANs



- Also available: quantum GAN (QGAN)

Generative Adversarial Networks, Ian Goodfellow, et al. (2014);
NIPS 2016 Tutorial: Generative Adversarial Networks, Ian Goodfellow (2016).

Variational methods in quantum many-body physics

- An example of quantum many-body trial wave function:

$$|\Psi\rangle = \exp\left[-\frac{1}{2}\sum_{R,R'} v_{R,R'} S^z_R S^z_{R'}\right]|\Phi_{\rm cl}\rangle \qquad |\Phi_{\rm cl}\rangle = \prod_R \left(|\uparrow\rangle_R + e^{iQR}|\downarrow\rangle_R\right)$$

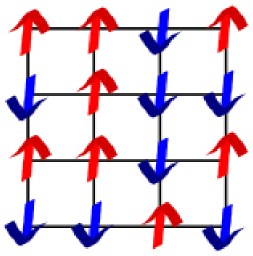
$v_{R,R'}$ (Jastrow factor), Q, etc. are our variational parameters.

or through Gutzwiller projection:

$$|\Psi\rangle = \mathcal{P}_G|\Phi_0\rangle$$

$$\mathcal{P}_G = \prod_R (n_{R,\uparrow} - n_{R,\downarrow})^2$$

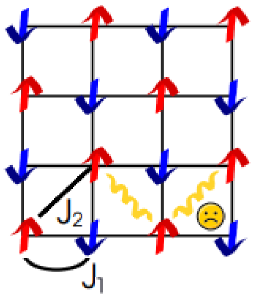
$$|\Phi_0\rangle = \exp\left\{\sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow}\right\}|0\rangle$$

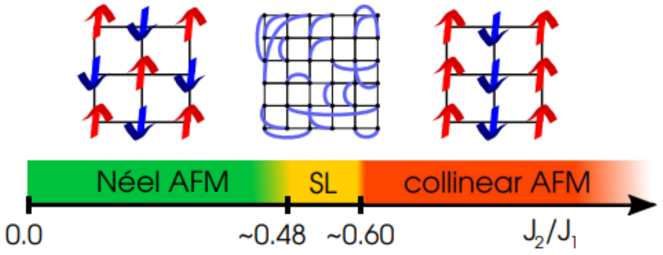


- An example of quantum many-body system (Hamiltonian):

$$\mathcal{H} = J_1 \sum_{\langle R,R'\rangle} \mathbf{S}_R \cdot \mathbf{S}_{R'} + J_2 \sum_{\langle\langle R,R'\rangle\rangle} \mathbf{S}_R \cdot \mathbf{S}_{R'}$$

One ‘simple’ job: find the quantum many-body state with the lowest $\langle H\rangle$.





Wen-Jun Hu, et al., Phys. Rev. B ***88****, 060402R (2013).*

Frustration → sign problem in quantum Monte Carlo methods

Sampling expectation values for quantum many-body state

- For an operator $\hat{O}$, its expectation value:

$$\langle \hat{O} \rangle = \langle \Psi | \hat{O} | \Psi \rangle = \sum_\alpha \langle \Psi | \alpha \rangle \langle \alpha | \hat{O} | \Psi \rangle = \sum_\alpha \langle \Psi | \alpha \rangle \langle \alpha | \Psi \rangle \cdot \frac{\langle \alpha | \hat{O} | \Psi \rangle}{\langle \alpha | \Psi \rangle}$$

$$= \sum_\alpha \underbrace{\langle \Psi | \alpha \rangle \langle \alpha | \Psi \rangle}_{\text{MC sampling: positive-definite normalized weight}} \cdot \sum_\beta \langle \alpha | \hat{O} | \beta \rangle \frac{\langle \beta | \Psi \rangle}{\langle \alpha | \Psi \rangle}$$

contribution of each sampled $|\alpha\rangle$

- The sampling only concerns with ratios of $\langle \alpha' | \Psi \rangle / \langle \alpha | \Psi \rangle$, $\langle \beta | \Psi \rangle / \langle \alpha | \Psi \rangle$ with local differences.

- Given one (formalism of) quantum many-body state, it is relatively easy and controlled to evaluate operator expectations.

$$|\Psi\rangle = \sum_{i_1 i_2 \ldots i_N} \Psi_{i_1 i_2 \ldots i_N} | i_1 \otimes i_2 \otimes \cdots \otimes i_N \rangle$$

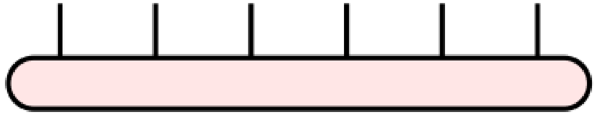
- Issue with the variational methods: even if we have hundreds of variational parameters, it is overwhelmingly dwarfed by the 2^N degrees of freedom of the quantum Many-body states

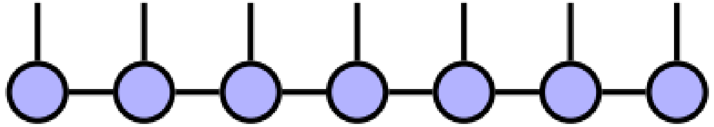
→ the "ground-state" energy may be just an upper bound, and the "ground-state" physics wrong.

Tensor network states as quantum many-body states

- A tensor network state is still a trial wavefunction representation with much fewer parameters:

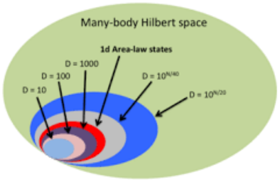
$$|\Psi\rangle = \sum_{i_1 i_2 \ldots i_N} \Psi_{i_1 i_2 \ldots i_N} |i_1 \otimes i_2 \otimes \cdots \otimes i_N\rangle \quad \Rightarrow \quad \sum_{s_1 \cdots s_N} \mathrm{Tr}(A^{s_1} \cdots A^{s_N}) |s_1 \cdots s_N\rangle$$

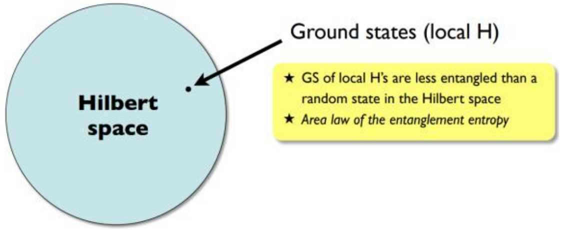


$\psi =$ 

- Why we can do this? Quantum entanglement suggests that ground states are special:

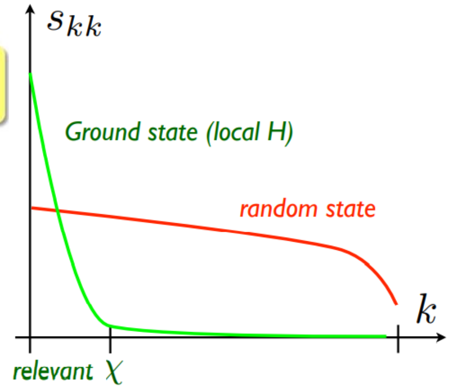
$$|\Psi\rangle = \sum_k^M s_{kk} |u_k\rangle |v_k\rangle$$





direct product state: $s_{11} = 1$ $\quad |\Psi\rangle = 1|u_1\rangle|v_1\rangle$

entangled state: $s_{11} = \frac{1}{\sqrt{2}}$ $\quad s_{22} = \frac{1}{\sqrt{2}}$ $\quad |\Psi\rangle = \frac{1}{\sqrt{2}}|u_1\rangle|v_1\rangle + \frac{1}{\sqrt{2}}|u_2\rangle|v_2\rangle$



- Ground-state protocol based upon tensor network states: density matrix renormalization group (DMRG)

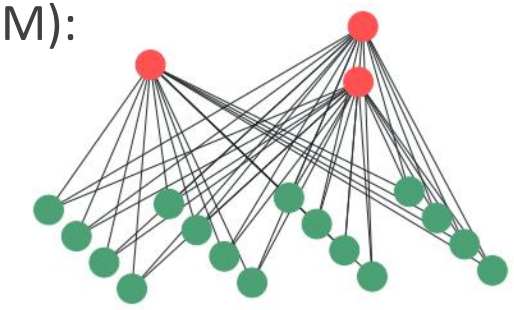
*Steven R. White, Phys. Rev. Lett. **69**, 2863 (1992).*

Neural network quantum states

- Representing the quantum many-body state with a neural network (RBM):

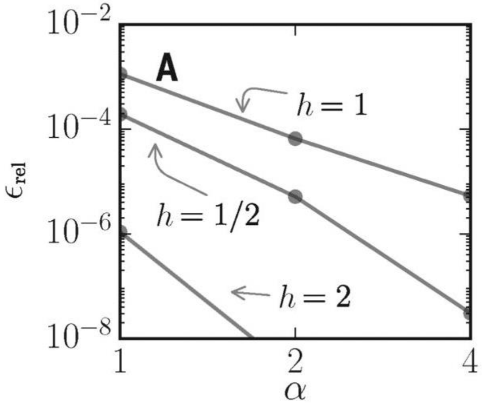
$$|\Psi_{\mathrm{RBM}}\rangle \propto \sum_{h_a=\pm1} \exp\left[\sum_{R,a} W_{R,a} S_R^z h_a + \sum_a b_a h_a\right] |\Phi_{\mathrm{cl}}\rangle$$

$$\propto \prod_a \exp\left\{\log\cosh\left[b_a + \sum_R W_{R,a} S_R^z\right]\right\} |\Phi_{\mathrm{cl}}\rangle$$



$\alpha = M/N$

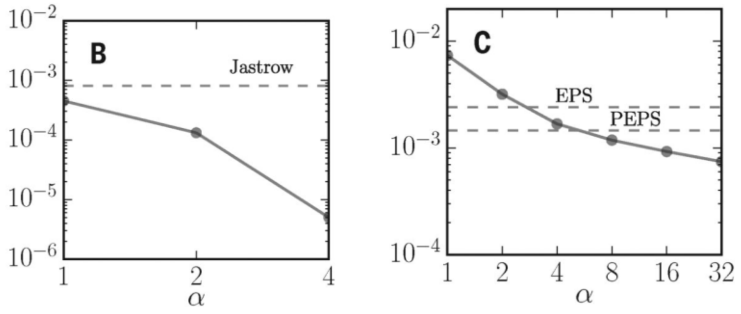
- Quality be systematically improved upon increasing the number of hidden neurons M.



1D transverse field Ising model:

$$\mathcal{H}_{\mathrm{TFI}} = -h\sum_i \sigma_i^x - \sum_{ij} \sigma_i^z \sigma_j^z$$

$$\epsilon_{\mathrm{rel}} = (E_{\mathrm{NQS}}(\alpha) - E_{\mathrm{exact}})/|E_{\mathrm{exact}}|$$



1D and 2D AFM Heisenberg model:

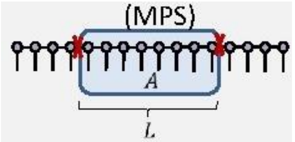
$$\mathcal{H}_{\mathrm{AFH}} = \sum_{ij} \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z$$

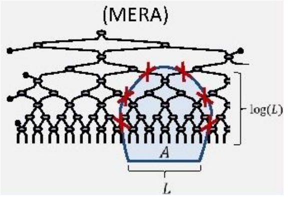
G. Carleo, et al., Science **355**,602–606 (2017).

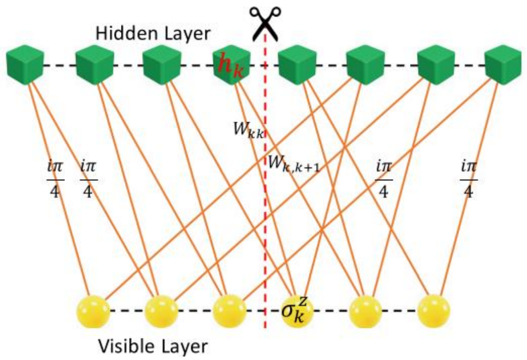
Neural network quantum states: pros and cons

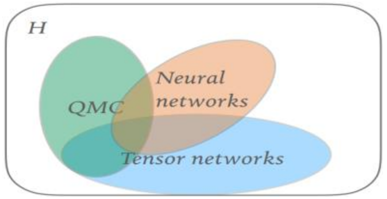
- Why it works: convertibility between neural network and tensor network representations
- Benefits: description beyond Area-law, etc. (chiral state)

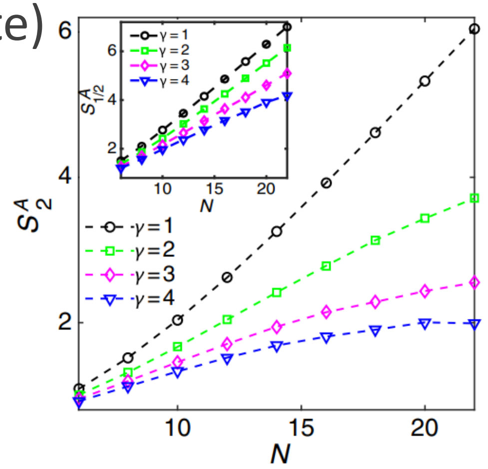
In comparison: MPS (area law) and MERA (area law with log correction):



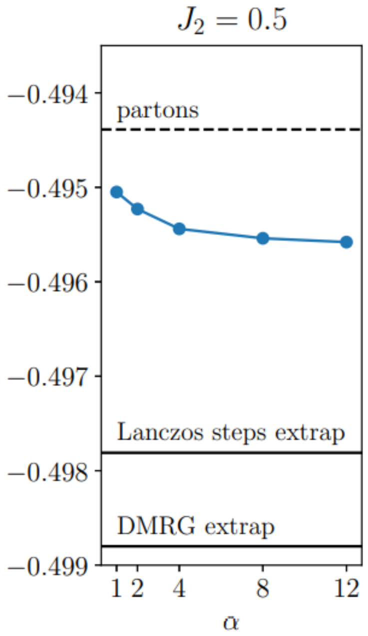








*Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma Phys. Rev. X **7**, 021021 (2017).*



- Potential problems:
  - difficult optimizations (local minima), e.g. perform relatively poorly on frustrated models: $J_1 - J_2$ model
  - preference for a specific spin-spin correlation
  - active degrees of freedom unclear → physics not transparent

Quantum state tomography

- Quantum state tomography reconstructs the quantum state by measurements on the system.

  necessary for quantum computation, quantum entanglement, error diagnosis, etc.

- Setup and requirements:
  - A quantum system ρ that can be prepared repeatedly
  - A set of quantum measurements
  - A model that analyze the outcome of the measurements to produce an estimate ρ^*

For N =8, a brute-force QST requires almost 10^6 measurements.

- Example: single spinor state $\langle\vec{\sigma}\rangle = tr(\rho\vec{\sigma})$

  (with infinite precision of $\langle\vec{\sigma}\rangle$) $\rho = aI + \mathbf{b}\cdot\boldsymbol{\sigma}$ $\Rightarrow$ $\begin{cases} a = \frac{1}{2}\operatorname{tr}\rho = \frac{1}{2} \\ b_i = \dfrac{1}{2}\operatorname{tr}(\sigma_i\rho) = \dfrac{1}{2}\langle\sigma_i\rangle \end{cases}$ $\Rightarrow$ $\rho = \dfrac{1}{2}\left(I + \langle\boldsymbol{\sigma}\rangle\cdot\boldsymbol{\sigma}\right)$

- Difficulty: exponential scaling both in the representation and the analysis, quantum and thermal fluctuations, etc.

  Limiting the scale of rapidly advancing quantum simulators:

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

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

Neural network quantum state tomography

- Define the quantum state:

$$\psi_{\lambda,\mu}(\mathbf{x}) = \sqrt{\frac{p_\lambda(\mathbf{x})}{Z_\lambda}}\, e^{i\phi_\mu(\mathbf{x})/2} \qquad \phi_\mu = \log p_\mu(\mathbf{x})$$

  represented by two RBMs: $p_\kappa(\boldsymbol{\sigma}, \mathbf{h}) = e^{\sum_{ij} W^\kappa_{ij} h_i \sigma_j + \sum_j b^\kappa_j \sigma_j + \sum_i c^\kappa_i h_i}$ one for amplitude, one for phase

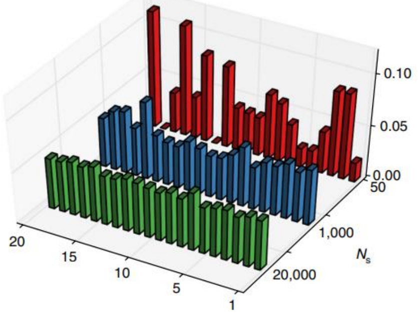
- Train to maximize the data-set likelihood of independent measurements: $|\psi_{\lambda,\mu}(\mathbf{x}^{[b]})|^2 \simeq |\Psi(\mathbf{x}^{[b]})|^2$

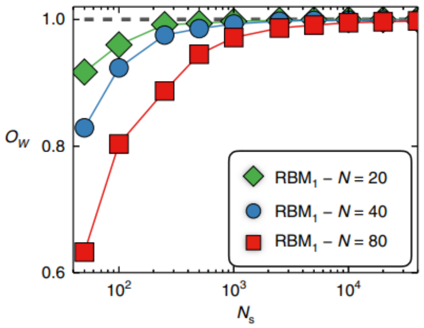
  (only raw data instead of expectation values). Also, KL divergence: $\Xi(\boldsymbol{\kappa}) \equiv \sum_{b=0}^{N_B} \mathrm{KL}^{[b]}_{\kappa} = \sum_{b=0}^{N_B} \sum_{\{\boldsymbol{\sigma}^{[b]}\}} P_b(\boldsymbol{\sigma}^{[b]}) \log \frac{P_b(\boldsymbol{\sigma}^{[b]})}{|\psi_\kappa(\boldsymbol{\sigma}^{[b]})|^2}$

- Example: the real W state: $|\Psi_W\rangle = \frac{1}{\sqrt{N}}(|100\ldots\rangle + \ldots + |\ldots001\rangle)$ (amplitude only)

Convergence as number of samples increases:

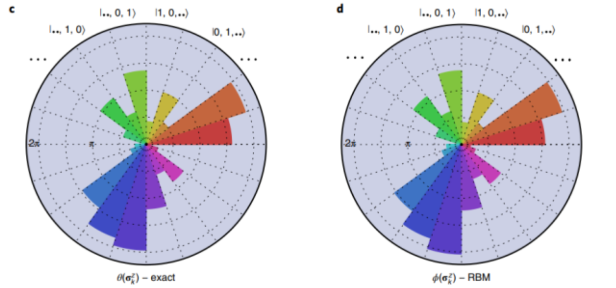
Dataset generated by sampling (synthetic measurements)





Now, with local phase variation:

$$\exp(i\theta(\boldsymbol{\sigma}^z_k)/2)$$



G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. ***14****, 447 (2018).*

Restricted Boltzmann Machine

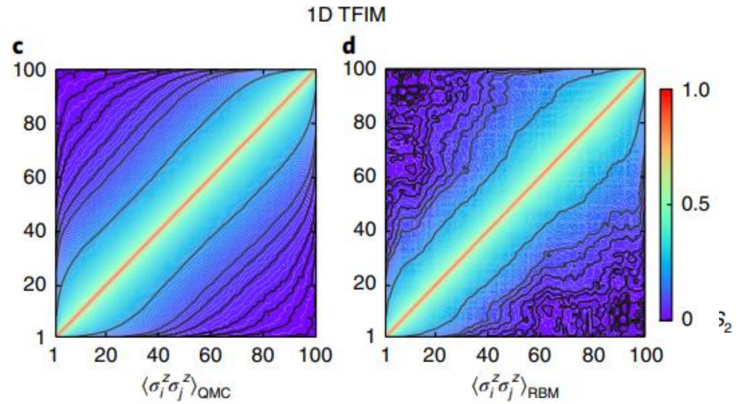
- Quantum state tomography on ground states of interacting many-body problems:

1D transverse-field Ising model:

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

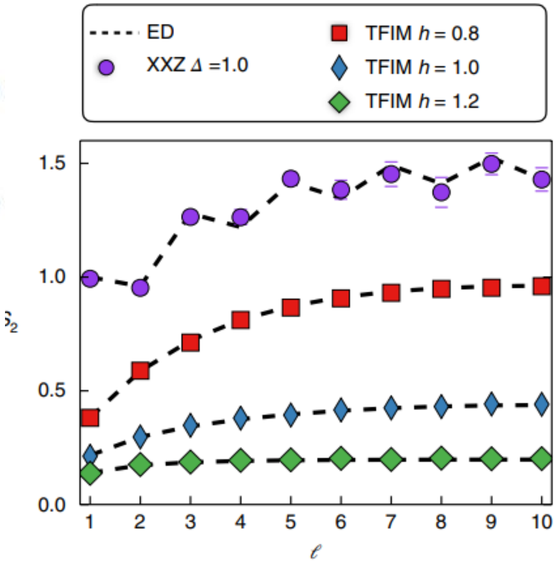
1D XXZ spin-1/2 model:

$$\mathcal{H} = \sum_{ij} \left[\Delta\left(\sigma_i^x\sigma_j^x + \sigma_i^y\sigma_j^y\right) + \sigma_i^z\sigma_j^z\right]$$



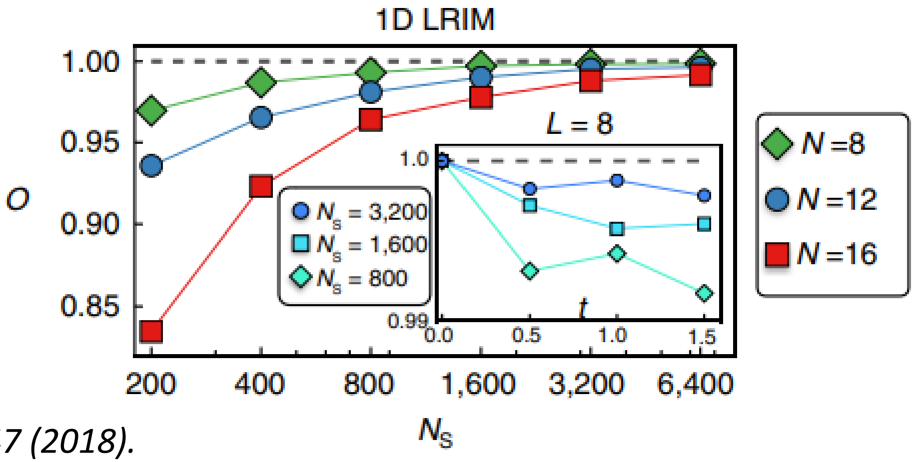
Consistent correlation and entanglement-entropy behavior

$$S_2(\rho_A) = -\log\left(\mathrm{Tr}\left(\rho_A^2\right)\right)$$



As well as time-evolved states:

$$|\Psi(t)\rangle = \exp(-i\mathcal{H}t)\,|\Psi_0\rangle \qquad \Psi_0 = |\rightarrow,\rightarrow,\ldots,\rightarrow\rangle$$



G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. ***14****, 447 (2018).*