Introduction to Machine Learning Methods in Condensed Matter Physics

LECTURE 3, FALL 2021

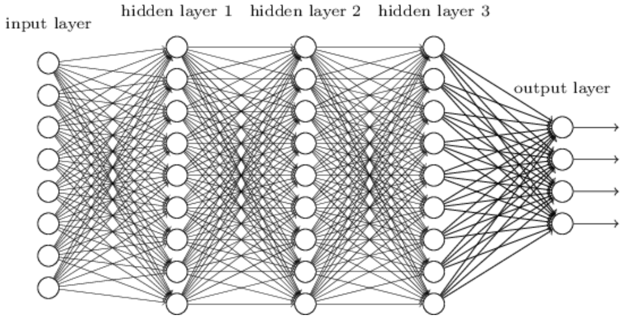
Yi Zhang (张亿)

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

Email: frankzhangyi@pku.edu.cn

Machine learning with artificial neural network

- Types
  - Classification: input → output: Class A $P(A)$, Class B $P(B)$
  - Regression: output: $f(x)$ with arbitrary range



- ***Training*** is the heavy lifting; the applications afterwards are relatively straightforward.
- No black magic: performance bounded from above by the ***quality of the samples***.

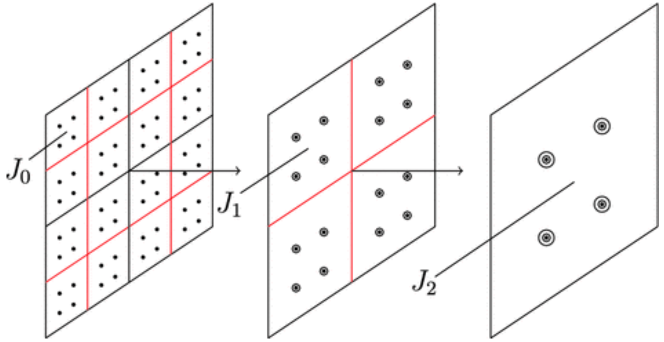
  'A good chef cannot make a decent meal with no ingredients.'

  Size
  Accuracy
  Diversity
- Even for the best case scenario, ANN is still an ***approximate*** function.

  Know your target and limitations!
- Sometimes, trying an idea out is the best way to verify its practicality.

What are phases?

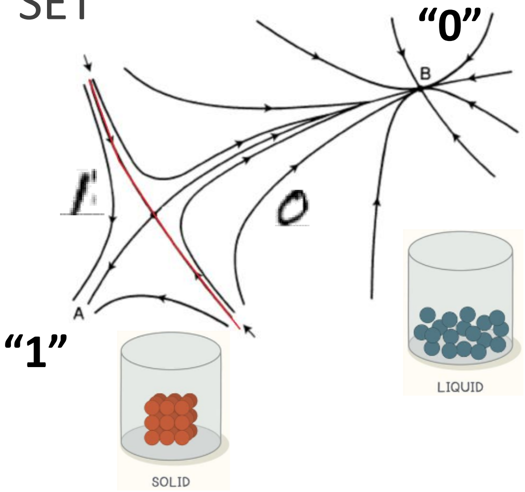
- Renormalization group stable fixed point
  - Conventional symmetry breaking phase ⇦ symmetries
  - Topological phases ⇦ topological invariants
  - Symmetry protected topological phases ⇦ both (also SET)



- Finer and more precise categories: solid ⇨ crystal symmetry group ⇨ SET

  even number ⇨ integer

- Uncertainties (noises) in measurements:
  - Model diversity (parameters, disorders)
  - Thermal fluctuations (configurational probability $e^{-\beta E}/Z$)
  - Quantum fluctuations (outcome a_n with probability $\langle\psi|P_n|\psi\rangle$)



Square-lattice ferromagnetic Ising model

$H = -J\sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z \qquad J = 1 \qquad \sigma_i^z = \pm 1$

At finite temperature: $P_\beta(\sigma) = \dfrac{e^{-\beta H(\sigma)}}{Z_\beta} \qquad Z_\beta = \sum_\sigma e^{-\beta H(\sigma)} \qquad \beta = (k_B T)^{-1}$

- Mean-field theory:

(single spin + environment)

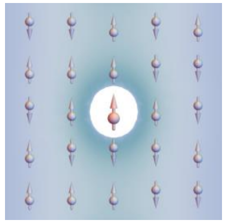
$\langle \sigma \rangle = m = \tanh \beta z J m$

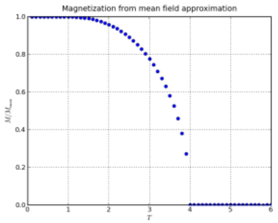
$T_c = zJ/k_B$

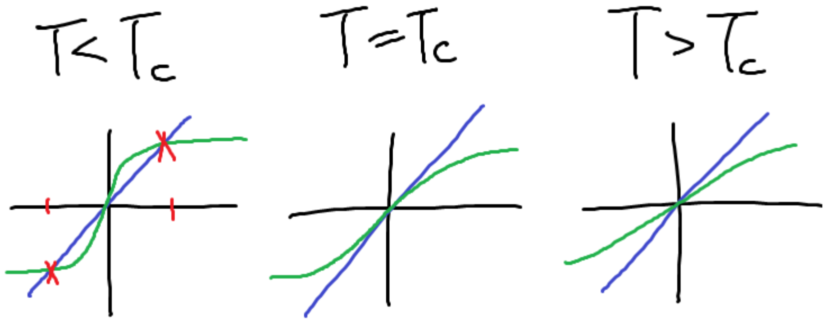
$z = 4$ for n.n. square lattice

$z = 6$ for n.n. triangle lattice

z: number of neighbors



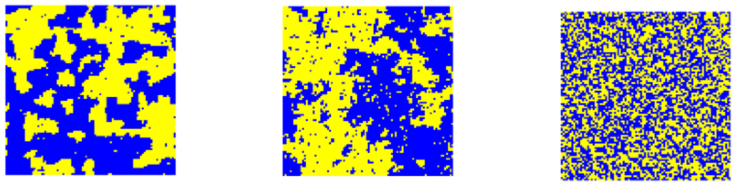




- Monte Carlo sampling:

Detailed balance:

$$A(\mu,\nu) = \begin{cases} e^{-\beta(H_\nu - H_\mu)}, & \text{if } H_\nu - H_\mu > 0 \\ 1 & \text{otherwise.} \end{cases}$$



For comparison, exact results:

square lattice: $T_c = 2J/\ln(1+\sqrt{2})$

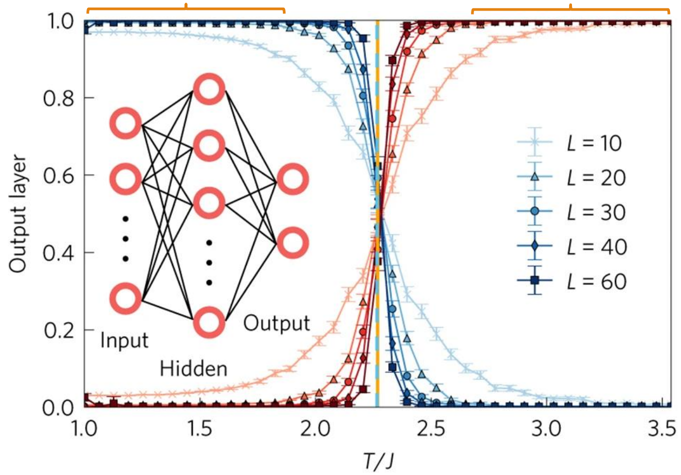
triangle lattice: $T_c = 4J/\ln 3$

Machine learning phases

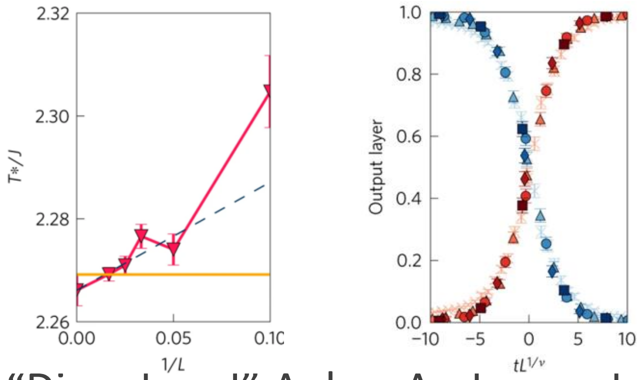
- Trained on square lattice

  Applied to square lattice

  Training sets:

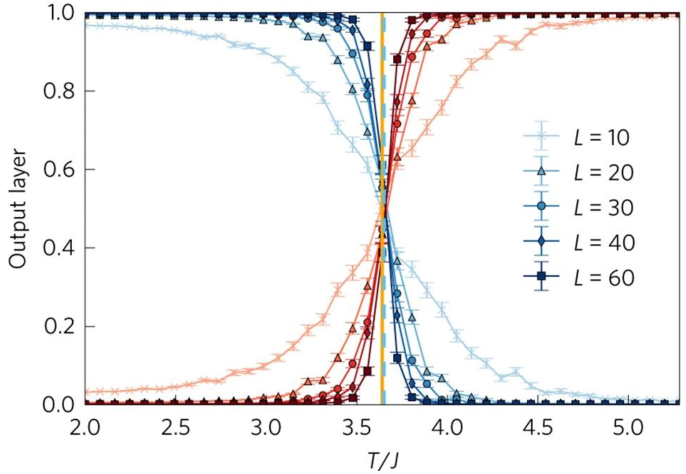


Finite size scaling? Data collapse? Really?



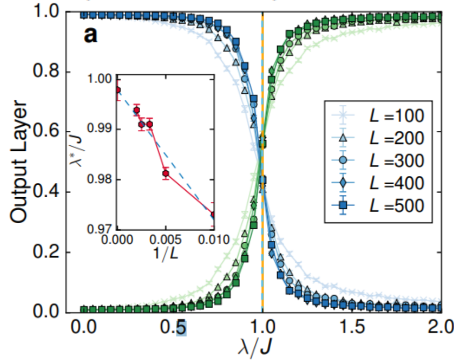
- Trained on square lattice

  Applied to triangle lattice



"Disordered" Aubry-Andre model:

$$H = -J\sum_i \left(c_i^\dagger c_{i+1} + \mathbf{h.c}\right) + 2\lambda \sum_i \cos(2\pi\phi i) c_i^\dagger c_i$$



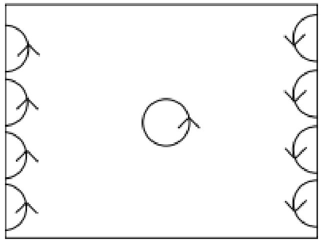
*Juan Carrasquilla, Roger G. Melko, Nature Physics **13**, 431–434 (2017).*

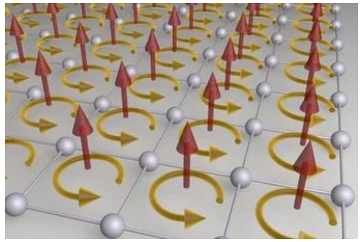
Quantum phases and topological states

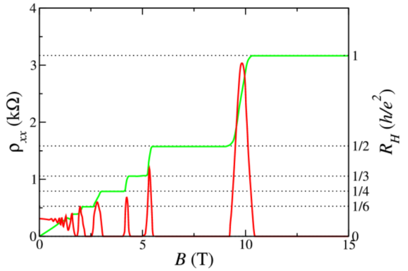
- Topological phases of matter

  Topological invariant   Chern number $\mathcal{A}_i(\mathbf{k}) = -i\langle u_{\mathbf{k}}|\frac{\partial}{\partial k^i}|u_{\mathbf{k}}\rangle$ $\mathcal{F}_{xy} = \frac{\partial \mathcal{A}_x}{\partial k^y} - \frac{\partial \mathcal{A}_y}{\partial k^x}$ $C = -\frac{1}{2\pi}\int_{\mathbf{T}^2} d^2k\, \mathcal{F}_{xy}$

  Edge states

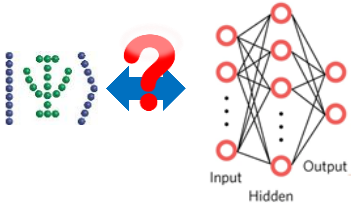
  

  Example: quantum Hall effect

  no symmetry-breaking order parameter

- The huge Hilbert space of a quantum many-body system $|\Psi\rangle = c_0|\uparrow\uparrow\uparrow\rangle + c_1|\uparrow\uparrow\downarrow\rangle + c_2|\uparrow\downarrow\uparrow\rangle + \cdots$

  Compatibility issue

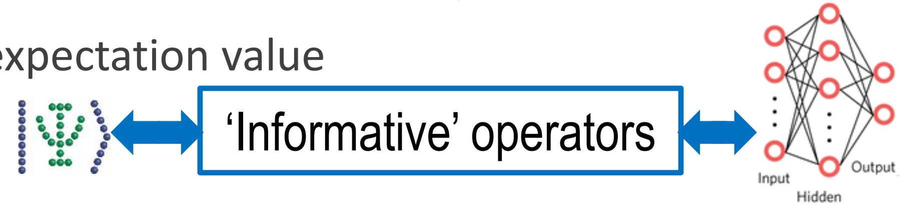
  

  exponentially growth is deleterious $2^{30} \sim 1000{,}000{,}000$

  $2^{1000} \sim \underbrace{100\cdots00}$ with 301 zeros

- Quantum fluctuations: measured value of observable $\neq$ expectation value

  Quantum operators map states to classical values

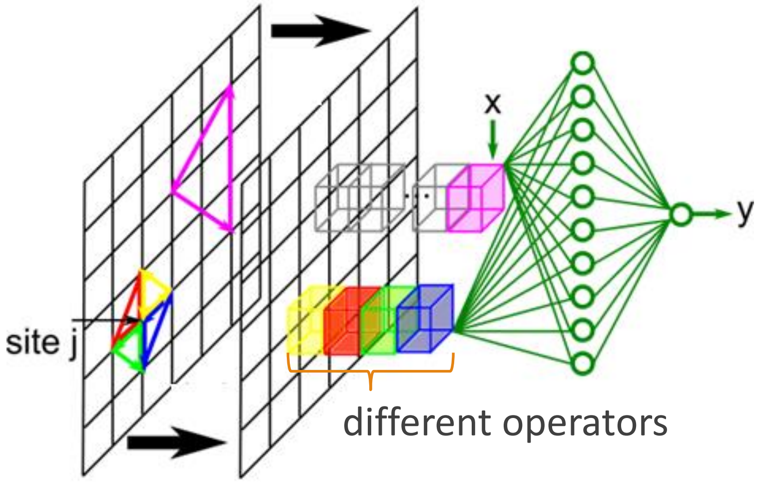
Machine learning quantum phases and topological states

- Intuition for operator choices from the Kubo formula:

$$\sigma_{xy} = \frac{ie^2\hbar}{N}\left[\sum_{n\neq 0}\frac{\langle\Phi_0\left|v_y\right|\Phi_n\rangle\langle\Phi_n\left|v_x\right|\Phi_0\rangle - x\leftrightarrow y}{\left(E_n-E_0\right)^2}\right]$$

$$= \frac{e^2}{h}\cdot\frac{1}{N}\sum 4\pi i\boxed{P_{jk}P_{kl}P_{lj}}S_{\triangle jkl} \qquad P_{ij}\equiv\langle c_i^\dagger c_j\rangle$$

Heavily reliant on the target phase ⇨ These operators must be important!



- In addition to a cut-off for more local operators, use single snapshots instead of expectation values:

$$\langle O\rangle = \langle\Phi|O|\Phi\rangle = \Sigma_\alpha\langle\Phi|\alpha\rangle\langle\alpha|O|\Phi\rangle = \Sigma_\alpha\langle\Phi|\alpha\rangle\langle\alpha|\Phi\rangle\times\boxed{\Sigma_\beta\langle\alpha|O|\beta\rangle\frac{\langle\beta|\Phi\rangle}{\langle\alpha|\Phi\rangle}}$$

$\langle\alpha|\Phi\rangle$: Slater determinant

expectation value

weight for $|\alpha\rangle$, probability density

train with the noise samples,

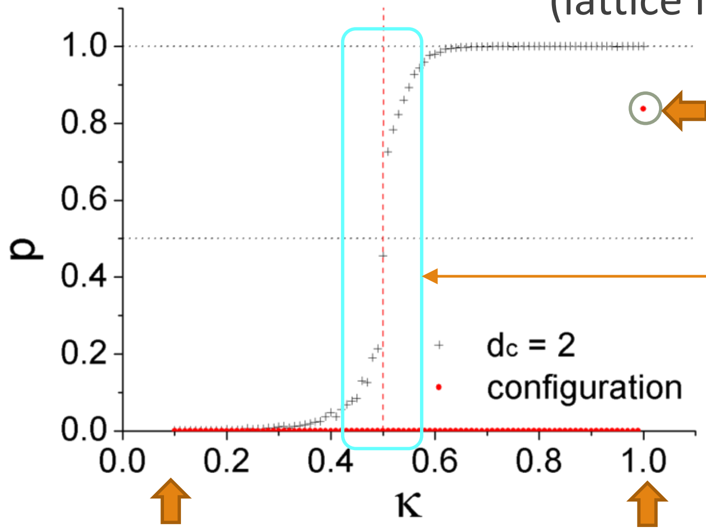
train for the noisy samples.

Yi Zhang, E.-A. Kim, Phys. Rev. Lett. ***118****, 216401 (2017).*

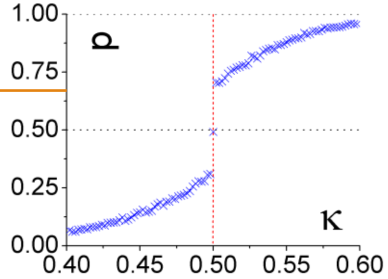
Machine learning quantum phases and topological states

- For a phase diagram between Chern insulators and normal insulators

(lattice integer quantum Hall insulator)

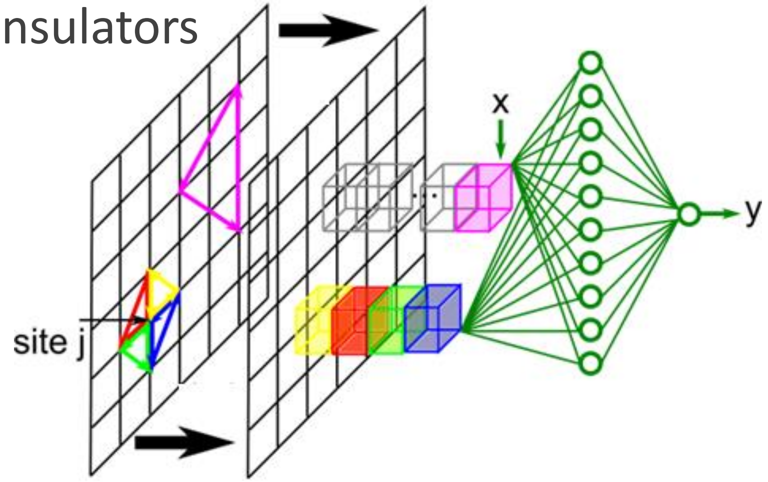


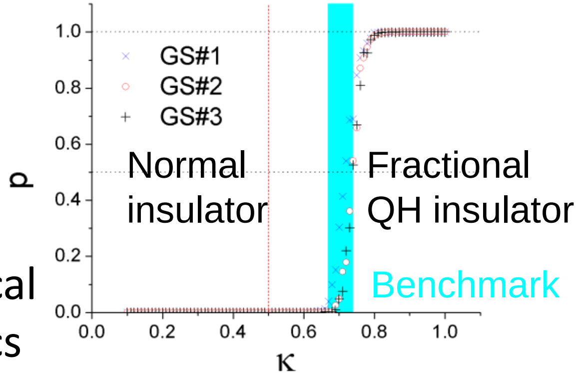
if electron number operators $n_r = c_r^\dagger c_r$ are used instead



Training sets

Training sets





- Strongly-correlated fractional quantum Hall states:

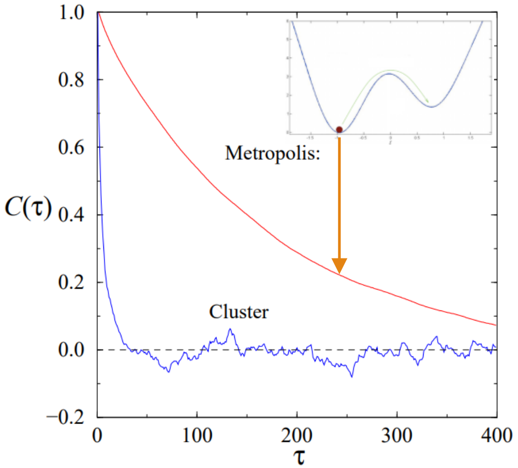
No available expression for Hall conductivity (topological invariant), but let's *try* the same operators for heuristics

Yi Zhang, E.-A. Kim, Phys. Rev. Lett. ***118****, 216401 (2017).*

Critical slowing down problem in Monte Carlo method

- Auto-correlation function as a measure of efficiency to de-correlate:

$$C(t) = \frac{\sum_i O_i O_{i+t}}{\langle O^2 \rangle} \propto e^{-t/\tau}$$



(a.) Metropolis: single spin flips ➡ encounter severe (critical) slowing down

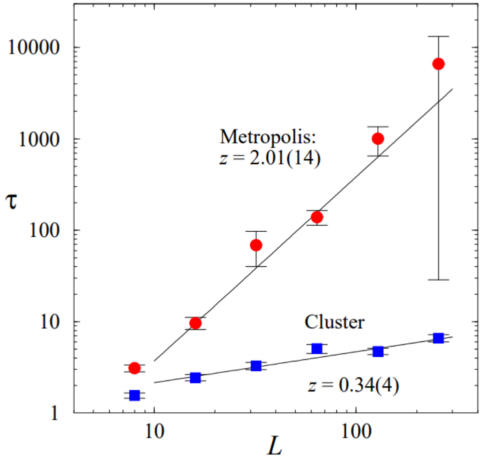
(b.) Wolff cluster algorithm: flip a single, randomly chosen cluster

1. Choose a random site. *U. Wolff, Phys. Rev. Lett.* ***62****, 361(1989).*

2. Add neighbor site into the cluster with probability $p = 1 - e^{-2\beta J}$ if the spin is identical.

3. Grow the cluster until all of cluster's neighbors and bonds are exhausted. Flip cluster.

heavily reliant on the model, yet 100% acceptance rate, very high efficiency



?(c.) Pick a random cluster. Accept with probability following detailed balance?

$$P(A \to B)/P(B \to A) = W(B)/W(A)$$

Self-learning Monte Caro methods

- For a given generic model, it is very difficult to design an efficient global update method.

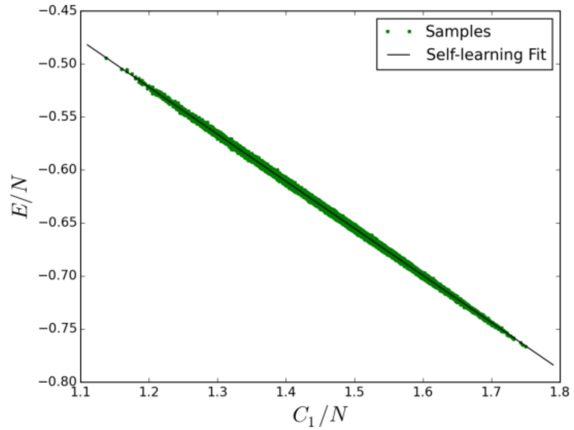
$$H = -J\sum_{\langle ij\rangle} S_i S_j - K \sum_{ijkl\in\square} S_i S_j S_k S_l$$

$$K/J = 0.2 \qquad J > 0$$

$$H_{\text{eff}} = E_0 - \tilde{J}_1 \sum_{\langle ij\rangle} S_i S_j$$

$$\tilde{J}_1 = 1.1064$$

Fit an effective model for a good description of the *low-energy physics*:



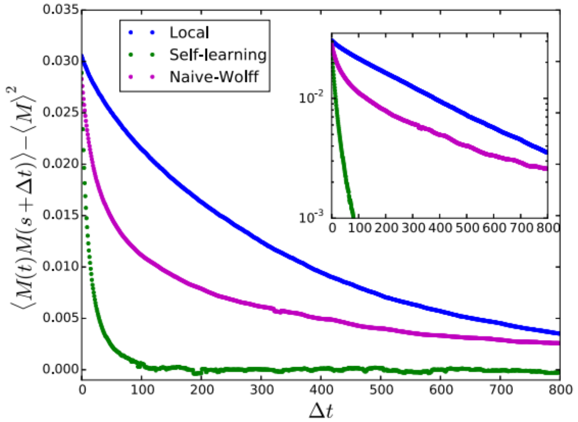
- Then, apply Wolff cluster algorithm to H_{eff}, with an acceptance rate:

$$\alpha(A \to B) = \min\{1, e^{-\beta[(E_B - E_B^{\text{eff}}) - (E_A - E_A^{\text{eff}})]}\}$$

Final result is exact after inclusion of α(A→B). T-dependent

Acceptance rate depends on approximation quality of effective model.

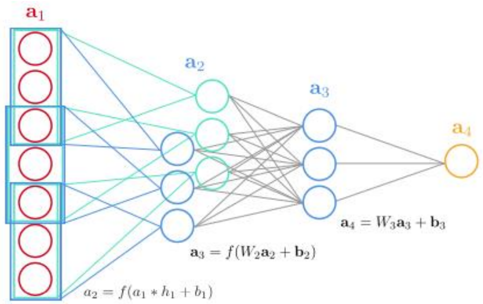
'Single-parameter' ($\tilde{J}_1$) machine learning, generalizable to other models

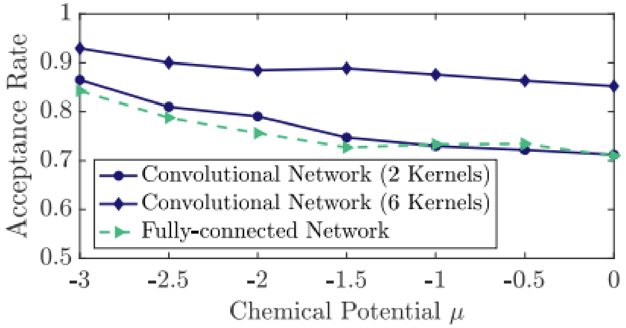


*Junwei Liu, Yang Qi, Zi Yang Meng, Liang Fu, Phys. Rev. B **95**, 041101(R) (2017).*

Accelerated Monte Carlo simulations with CNN and RBM

- Fitting $W(\sigma)$ with **convolutional neural network**:

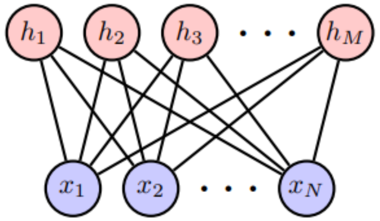




Issue: $\begin{cases}\text{not natural in proposing new clusters}\\ \text{learning slowdown in deep neural network}\end{cases}$

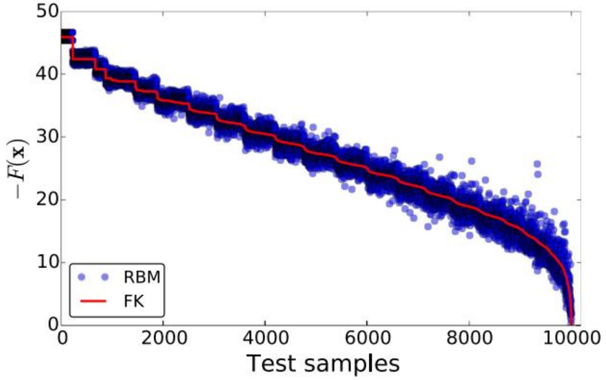
Huitao Shen, Junwei Liu, and Liang Fu,
*Phys. Rev. B **97**, 205140 (2018).*

- Fitting $W(\sigma)$ with **restricted Boltzmann machine**:



Generative model

propose new samples with (approximate) target probability



*Li Huang and Lei Wang, Phys. Rev. B **95**, 035105 (2017).*