

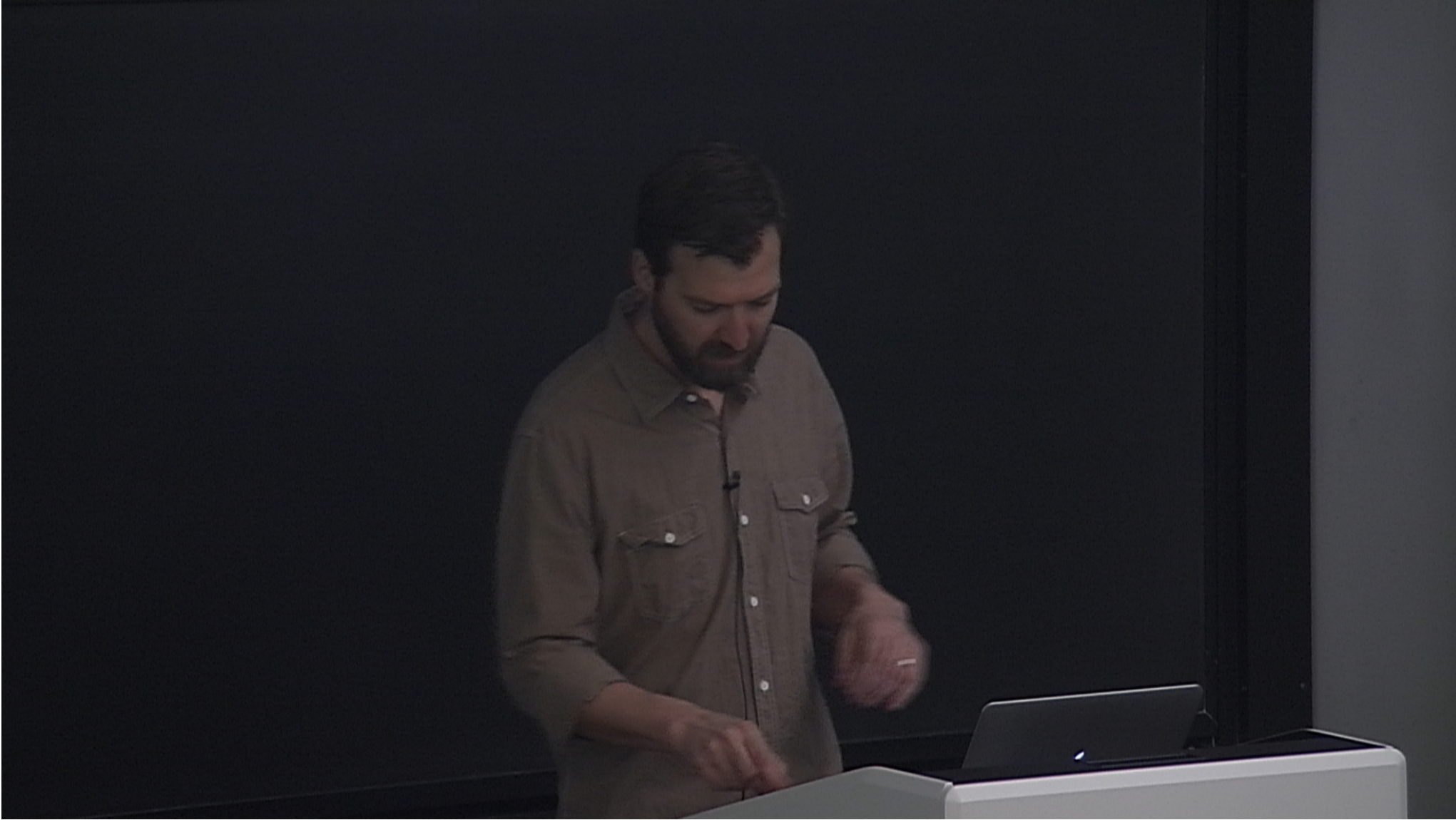
Title: PSI 2017/2018 - Machine Learning for Many Body Physics - Lecture 15

Date: Apr 25, 2018 09:00 AM

URL: <http://pirsa.org/18040066>

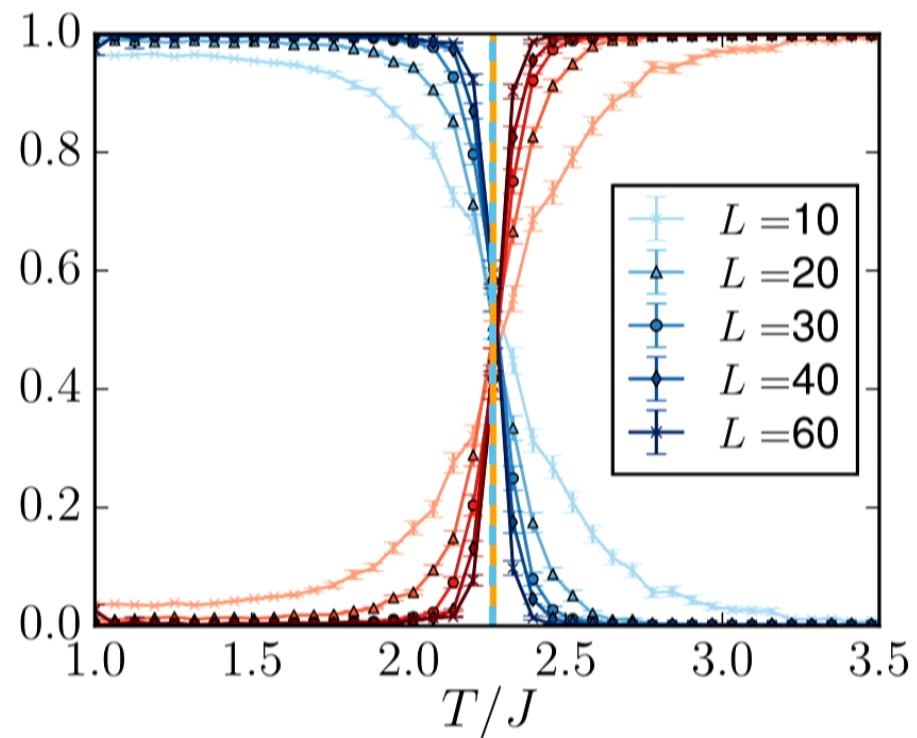
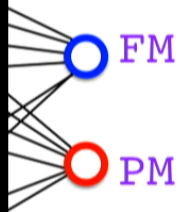
Abstract:

Research Frontiers **in Machine Learning for the** **Many Body Problem**

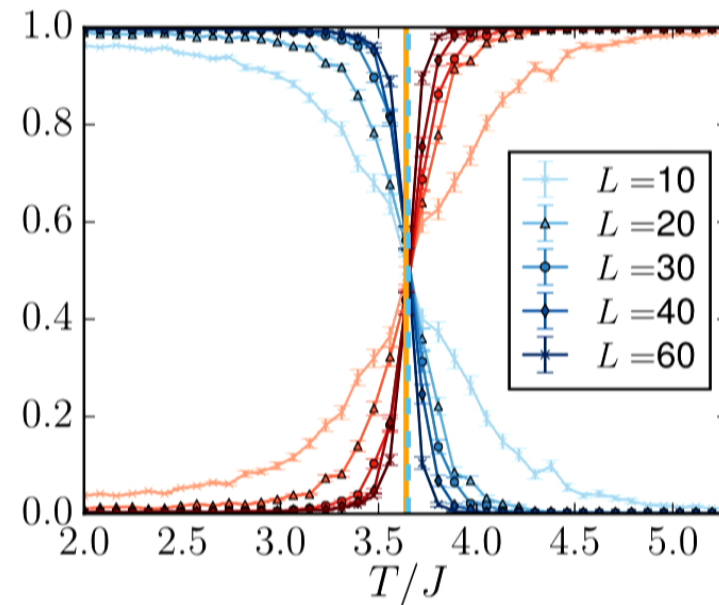
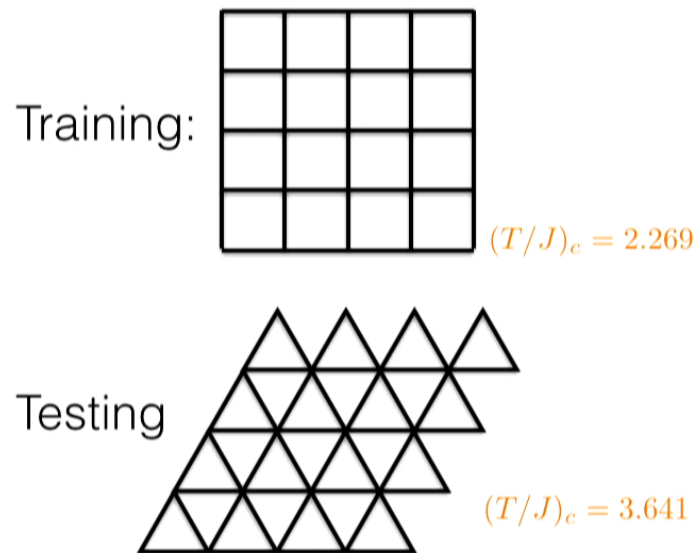


Supervised Learning

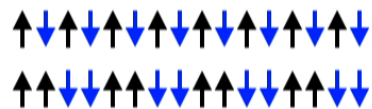
JC and RGM
Nature Physics **13**, 431 (2017)



Generalizing to other models

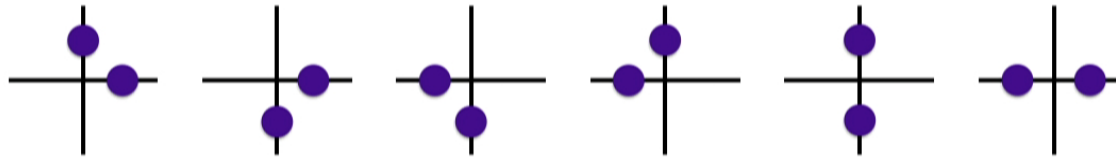


Other ordering patterns work similarly, with caveats...



Models without a groundstate order parameter:

Square ice: $H = \sum_v Q_v^2 \quad Q_v = \sum_{i \in v} S_i^z$



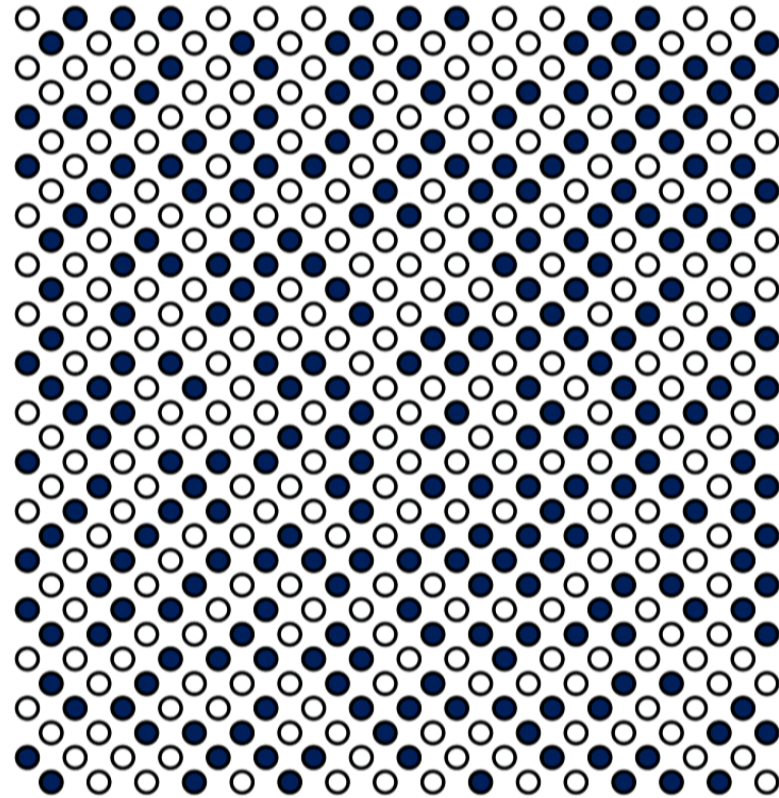
Coulomb phase at $T=0$

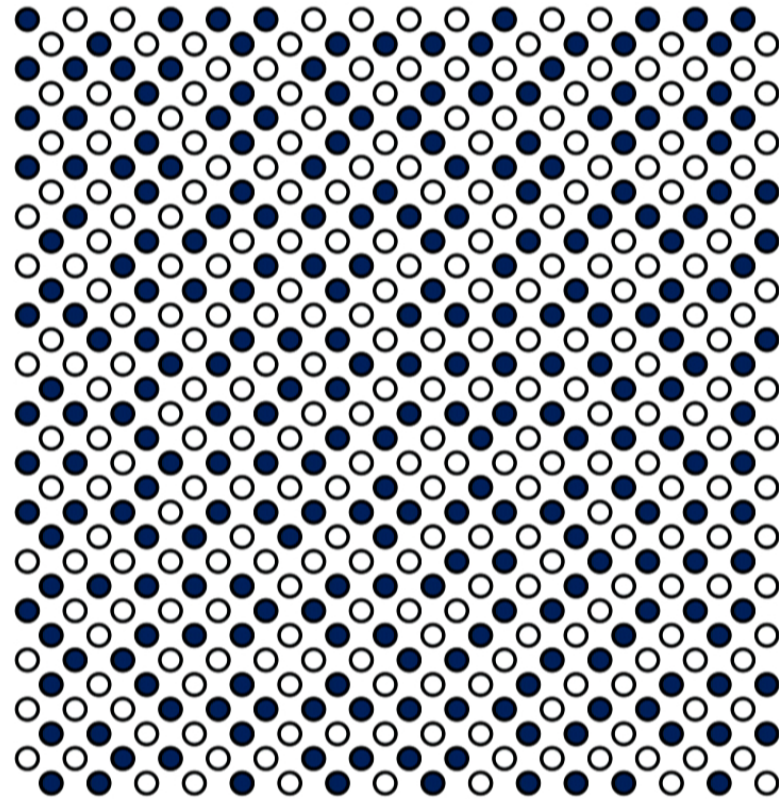
- Extensive groundstate degeneracy $S/N = \frac{3}{2} \log \frac{4}{3}$

Lieb Phys. Rev. **162**, 162 (1967)

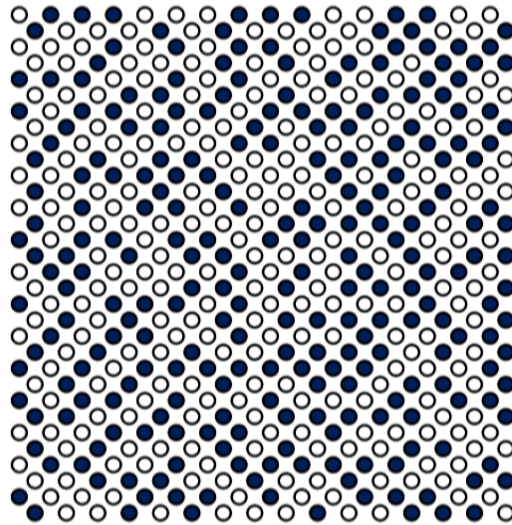
- Algebraic decay of correlations $\langle S_r^z S_0^z \rangle \sim \frac{1}{r^2}$

Train the neural network at two temperatures: $T = 0$ and $T = \infty$

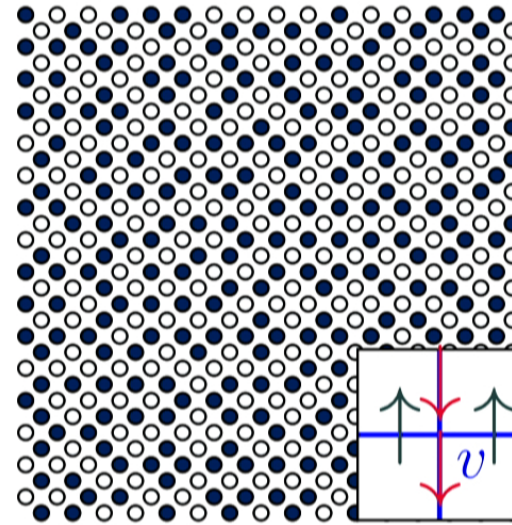




$$T = \infty$$



$$T = 0$$

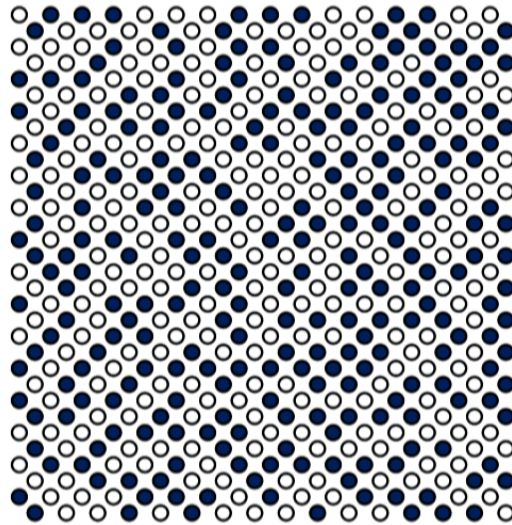


$$N = 2 \times 2 \times 16$$

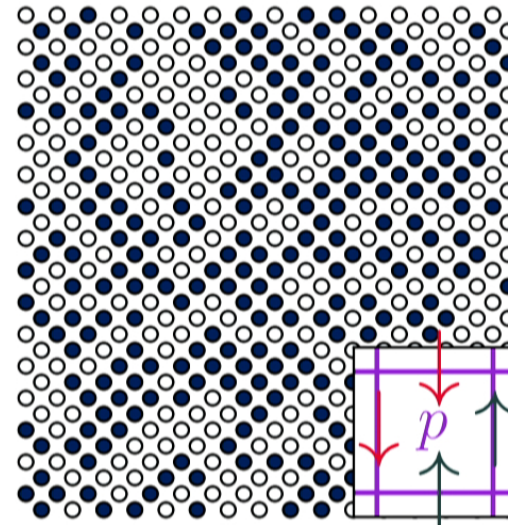
Using 100 hidden neurons, the feed forward neural network distinguishes the states with 99% accuracy.

Remember: the configurations are fed in as a 1D vector.

$$T = \infty$$



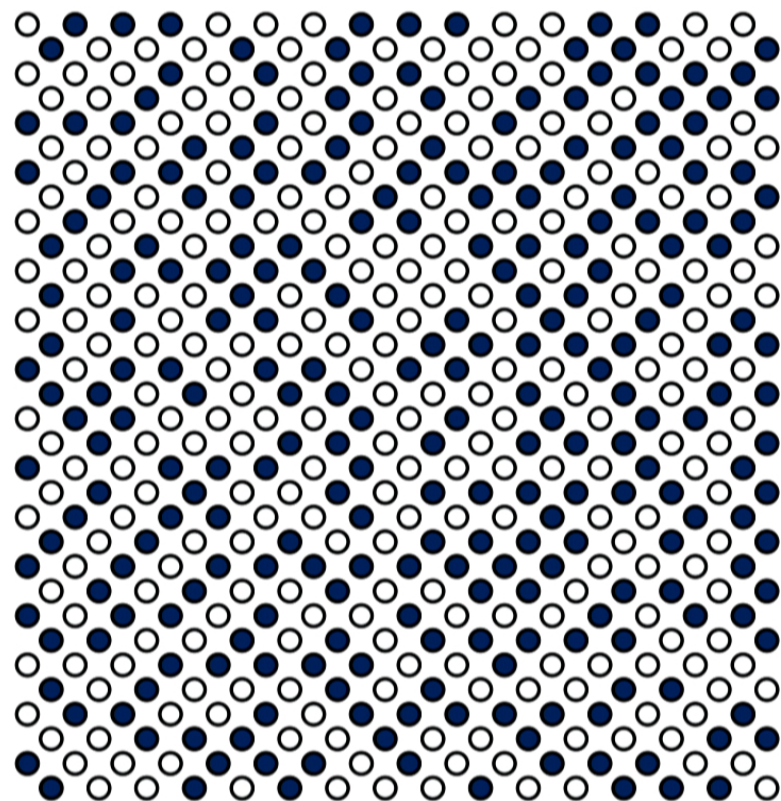
$$T = 0$$



$$N = 2 \times 2 \times 16$$

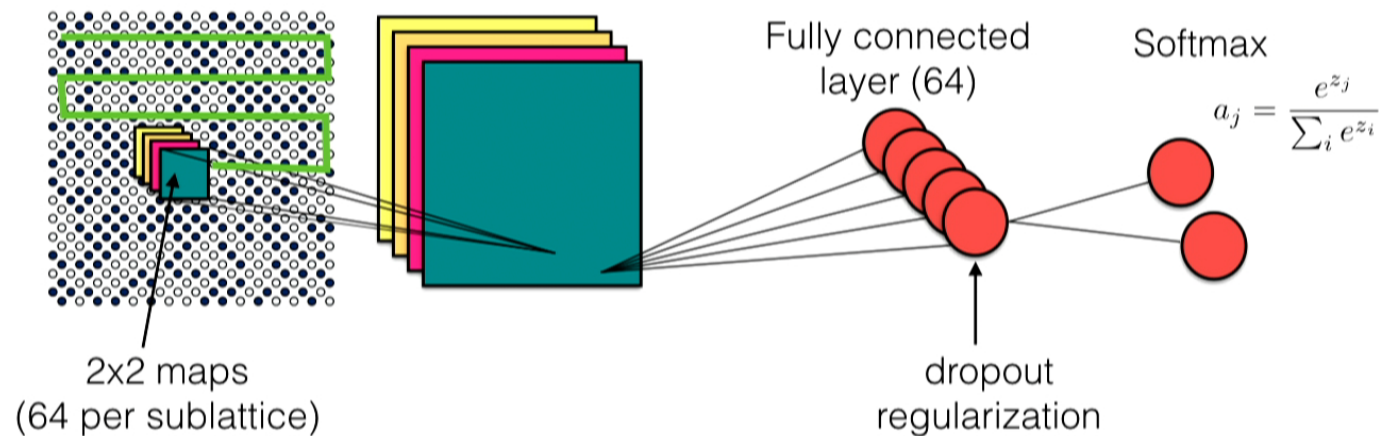
Using 100 hidden neurons, the feed forward neural network achieves only 50% accuracy - it is simply “guessing”

A failure of the simplest feed forward neural network...



Convolutional neural network

arXiv:1605.01735

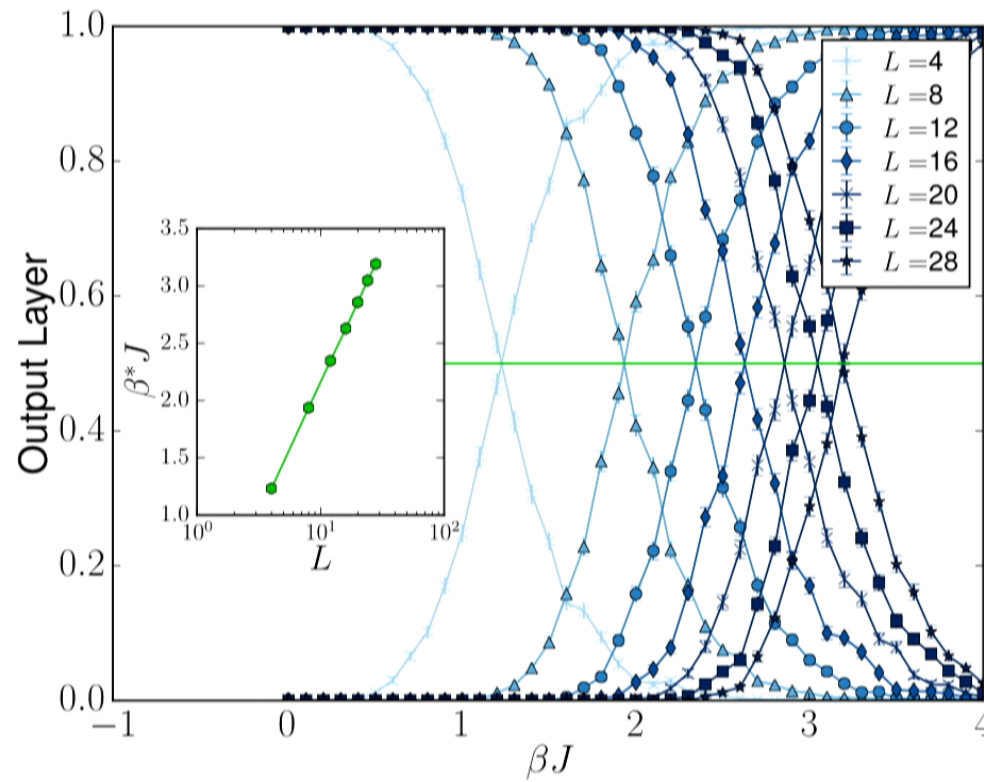
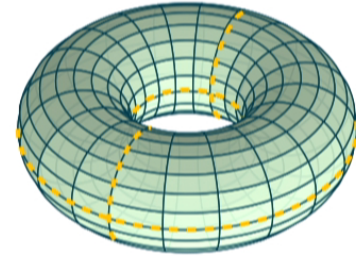


Restores information about dimension and locality -
discriminates high from low T states with near 100% accuracy

CNNs are the same technology behind many of the famous
machine/deep learning breakthroughs of the last few years

Crossover temperature for topological order in 2D

$$Le^{-J\beta} \sim 1 \quad T^*/J \sim \frac{1}{\log L}$$



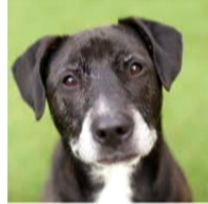
Other (very recent) applications of SUPERVISED machine learning in many- body physics



cat



cat



dog



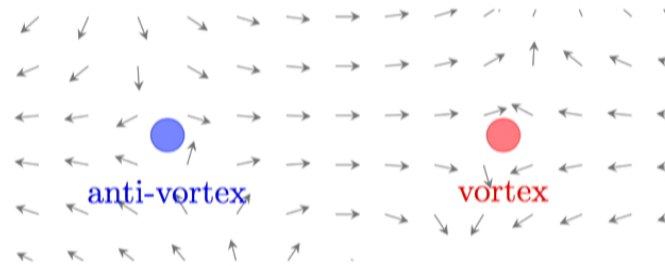
cat



muskrat

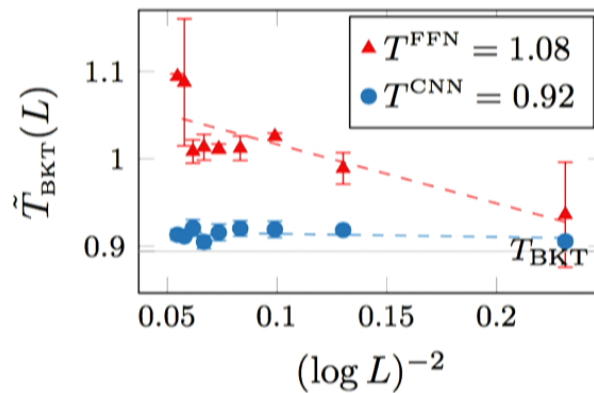
Can one “learn” the Kosterlitz Thouless transition?

Anna Golubeva Matt Beach

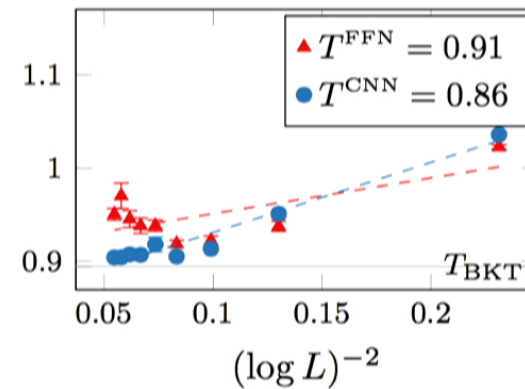


$$\mathcal{H}_{XY} = -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j)$$

$$T_{KT} = 0.893J$$



(a) Trained on raw spin configurations



(b) Trained on vortices

Identifying polymer states by machine learning

Qianshi Wei, Roger G. Melko, and Jeff Z. Y. Chen
Phys. Rev. E **95**, 032504 – Published 30 March 2017

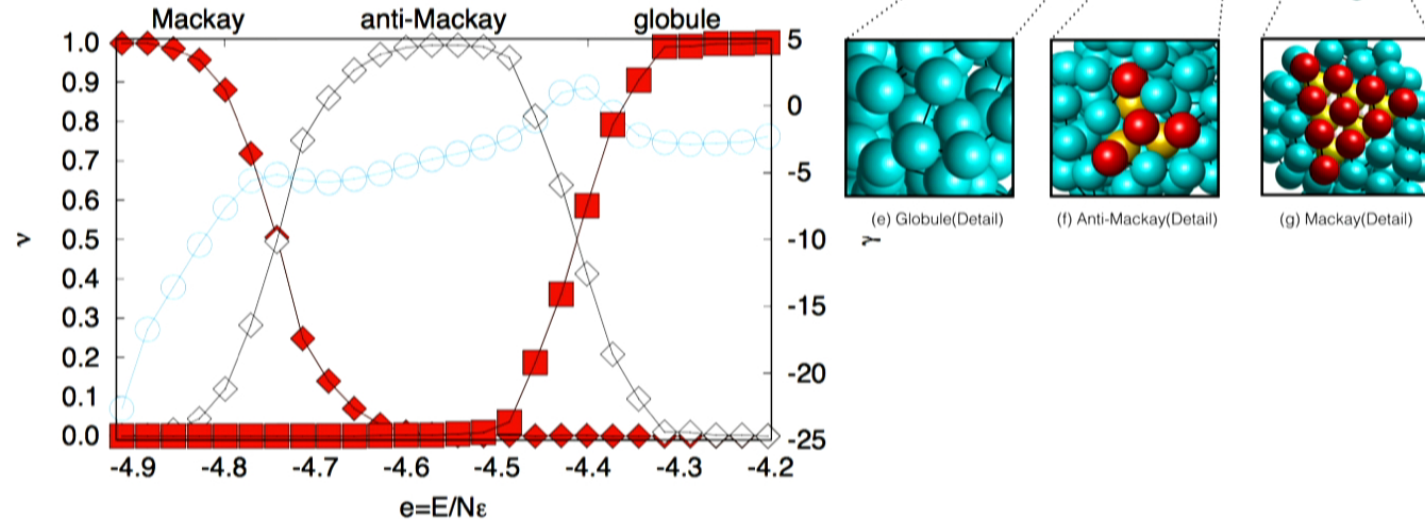
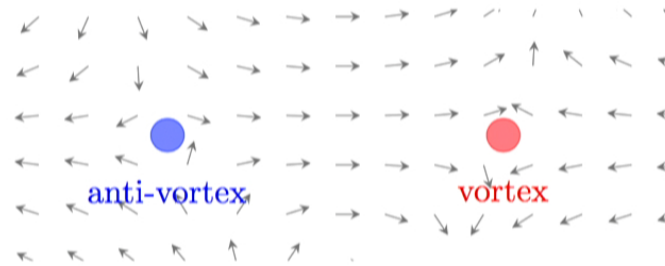


FIG. 4: Mean NN outputs ν (square for globule, open diamonds for anti-Mackay, and filled diamonds for Mackay) from the test samples, after the network is trained to recognizing these states in regimes where they are stable. In the background, the reduced specific heat-like γ (circles, to the right scale) was independently produced from the MC simulations.

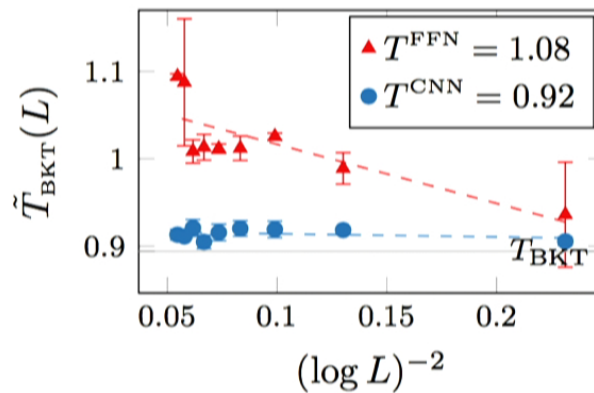
Can one “learn” the Kosterlitz Thouless transition?

Anna Golubeva Matt Beach

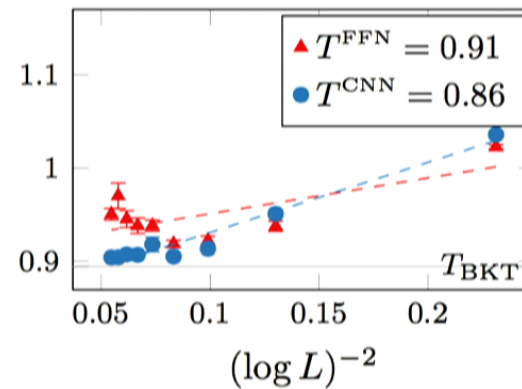


$$\mathcal{H}_{XY} = -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j)$$

$$T_{KT} = 0.893J$$



(a) Trained on raw spin configurations



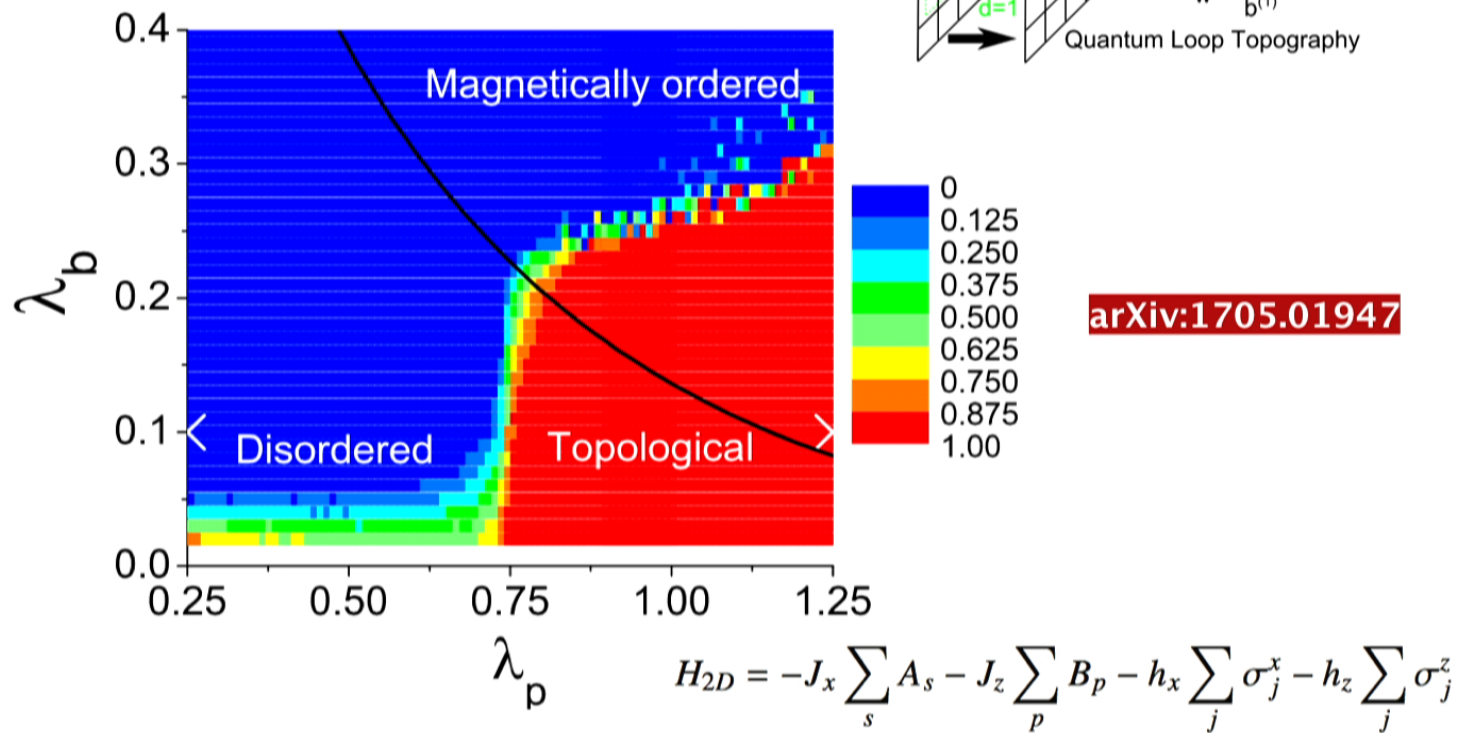
(b) Trained on vortices

Quantum Loop Topography for Machine Learning

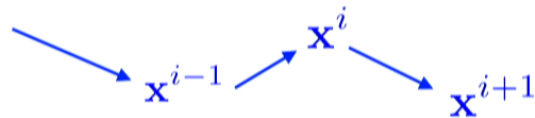
Yi Zhang and Eun-Ah Kim

Phys. Rev. Lett. **118**, 216401 – Published 22 May 2017

Physics See Viewpoint: [Neural Networks Identify Topological Phases](#)



Supervised learning is a powerful way to identify phases and phase transitions, in combination with conventional Monte Carlo sampling:



$$\mathbf{x}^i = [1, 0, 0, 1, 1, \dots]^T$$

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{MCS}}} \sum_i \mathcal{O}[\mathbf{x}^{(i)}]$$

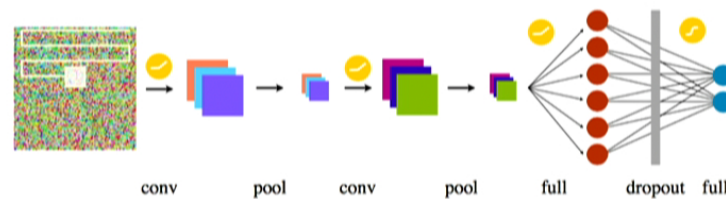
Quantum Monte Carlo: instances of the Markov Chain (images) are just $d+1$ dimensional configurations

$$\mathbf{x}^i = \begin{array}{|c|} \hline [1, 0, 0, 1, 1, \dots] \\ \hline \begin{array}{c} \text{Diagram of a 2D lattice configuration with paths} \end{array} \\ \hline [1, 0, 0, 1, 1, \dots] \\ \hline \end{array}$$

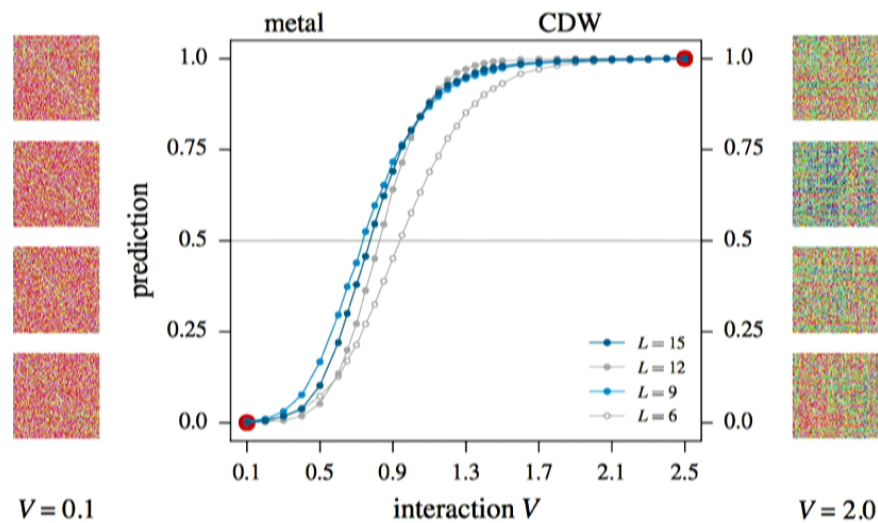
Machine learning quantum phases of matter beyond the fermion sign problem

Peter Broecker, Juan Carrasquilla, Roger G. Melko & Simon Trebst

Scientific Reports 7, Article number: 8823 (2017)



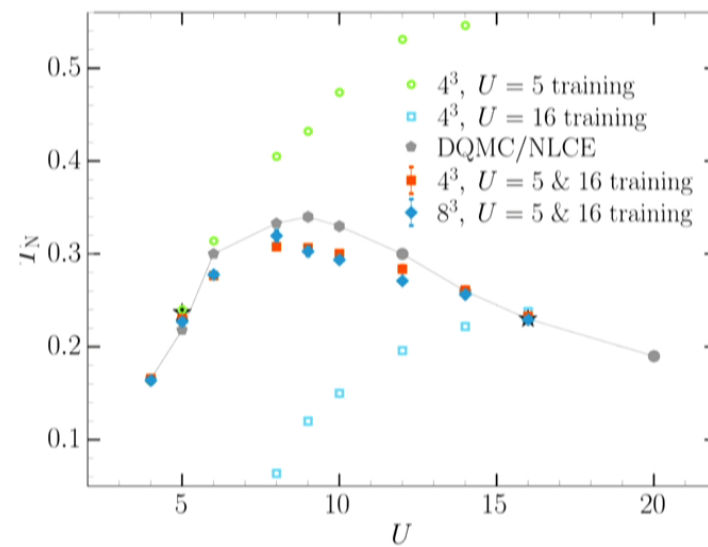
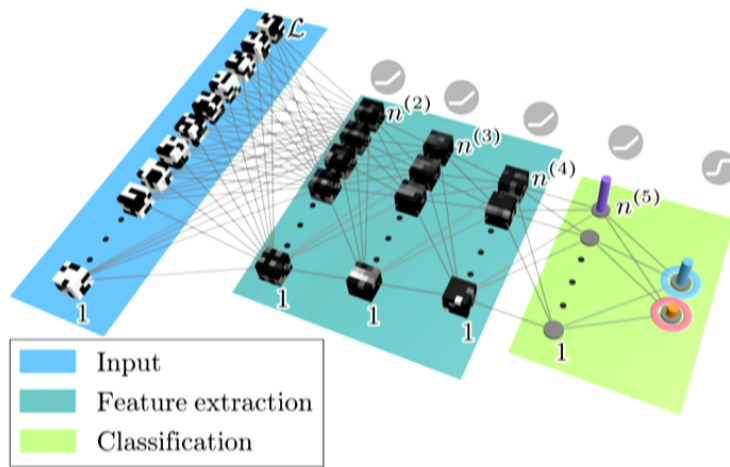
$$H = -t \sum_{\langle i,j \rangle} c_i^\dagger c_j + V \sum_{\langle i,j \rangle} n_i n_j$$



Machine Learning Phases of Strongly Correlated Fermions

Kelvin Ch'ng, Juan Carrasquilla, Roger G. Melko, and Ehsan Khatami
Phys. Rev. X **7**, 031038 – Published 30 August 2017

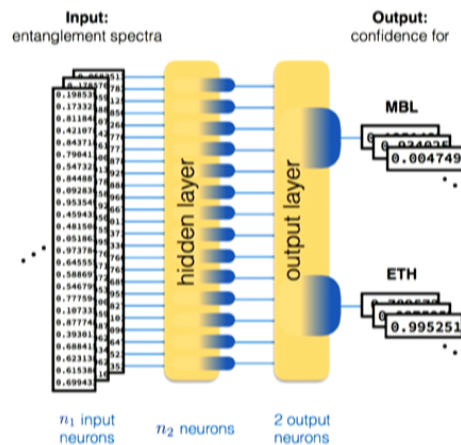
$$H = -t \sum_{\langle ij \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i \left(n_{i\uparrow} - \frac{1}{2} \right) \left(n_{i\downarrow} - \frac{1}{2} \right) - \mu \sum_{i\sigma} n_{i\sigma},$$



Probing many-body localization with neural networks

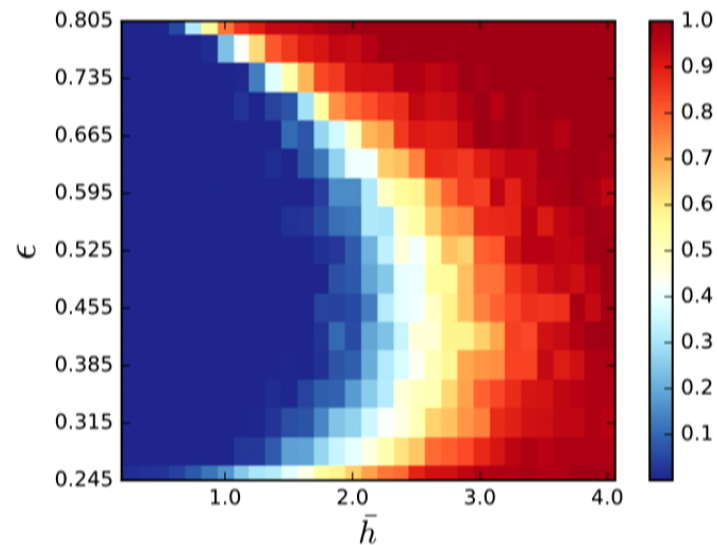
Frank Schindler, Nicolas Regnault, and Titus Neupert
Phys. Rev. B **95**, 245134 – Published 26 June 2017

We show that a simple artificial neural network trained on **entanglement spectra** of individual states of a many-body quantum system can be used to determine the transition between a many-body localized and a thermalizing regime.



$$\rho_A = \text{Tr}_B |\Psi\rangle \langle \Psi| \equiv e^{-H_e}$$

$$H = J \sum_{r=1}^{N-1} S_r \cdot S_{r+1} + \sum_{r=1}^N h_r S_r^z$$

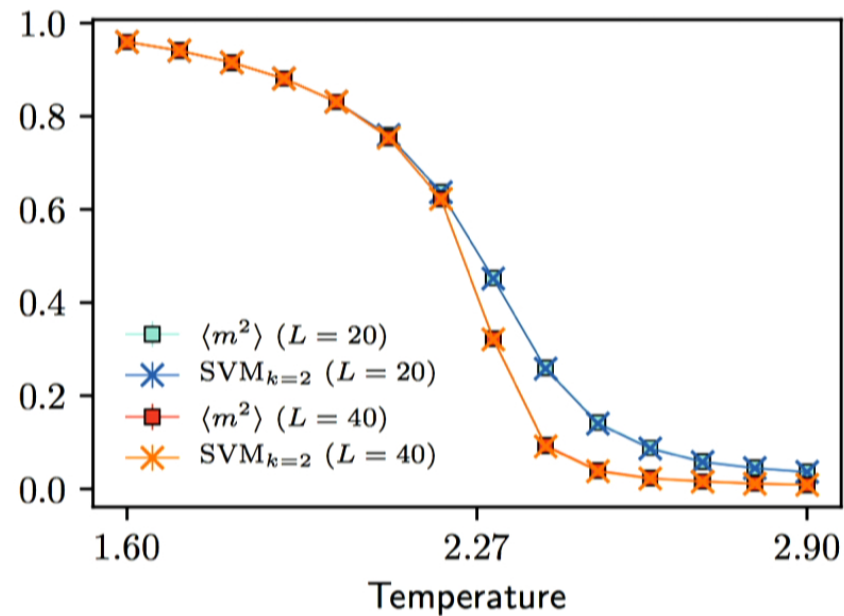


Kernel methods for interpretable machine learning of order parameters

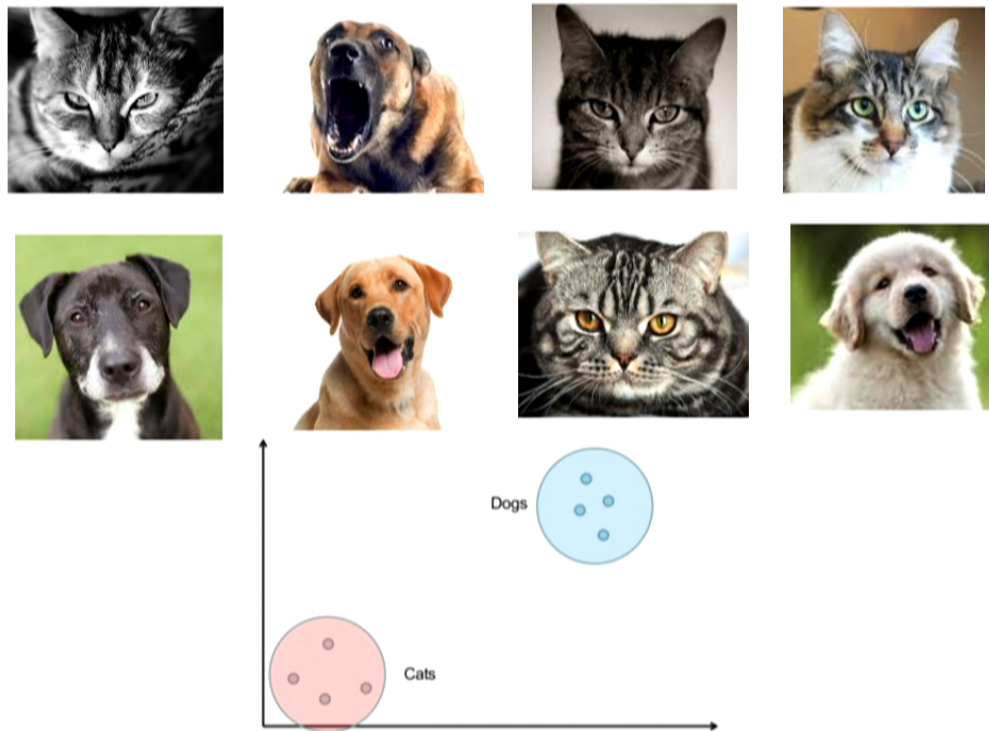
Pedro Ponte, Roger G. Melko

cond-mat > arXiv:1704.05848

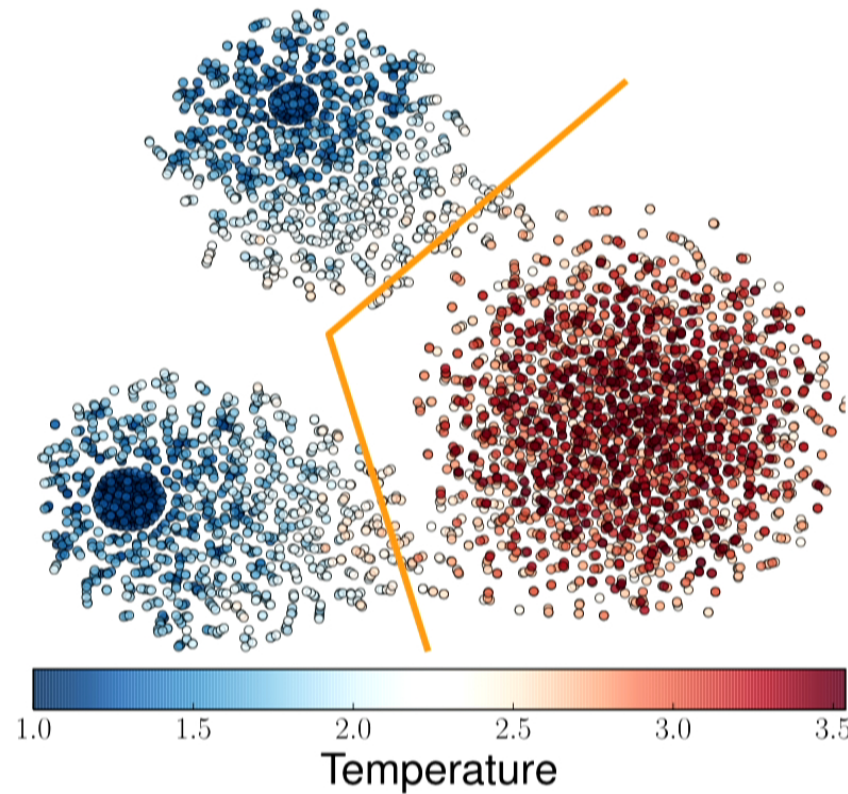
Support vector machines (SVMs) are a class of supervised kernel methods that can learn the mathematical form of physical discriminators, such as order parameters and Hamiltonian constraints.



Other (very recent) applications of UNSUPERVISED machine learning in many-body physics



The t-distributed stochastic neighbor embedding (t-SNE) technique, where high-dimensional data is embedded in two or three dimensions, so that data points close to each other in the original space are also positioned close to each other in the embedded low-dimensional space.



Discovering phase transitions with unsupervised learning

Lei Wang

Phys. Rev. B **94**, 195105 – Published 2 November 2016

The principal components are mutually orthogonal directions along which the variances of the data decrease monotonically.

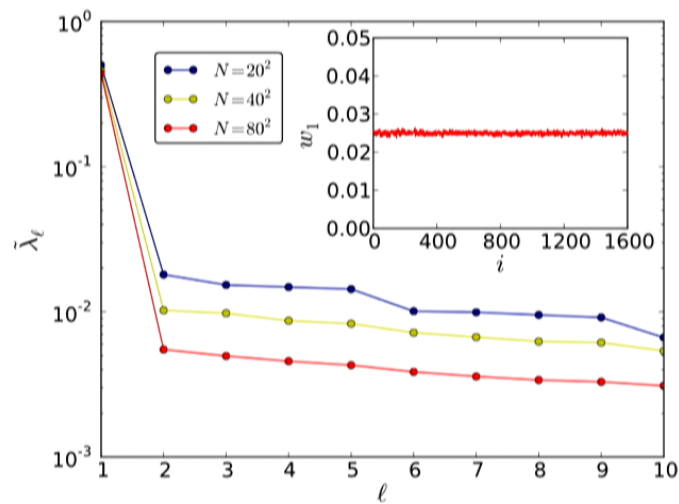


Figure 1: The first few explained variance ratios obtained from the raw Ising configurations. The inset shows the weights of the first principal component on an $N = 40^2$ square lattice.

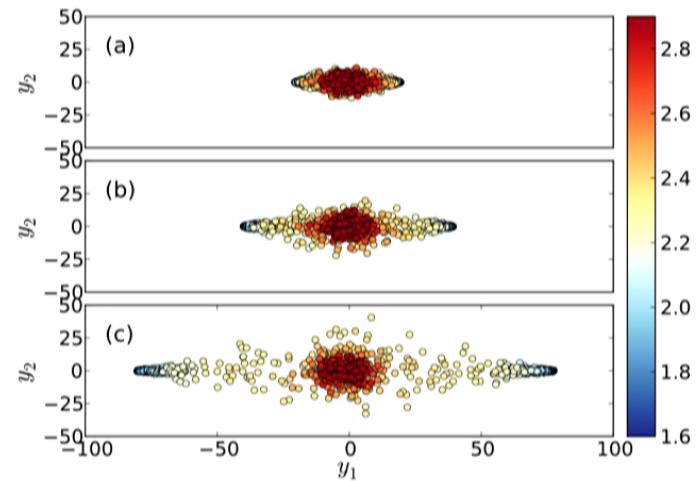


Figure 2: Projection of the samples onto the plane of the leading two principal components. The color bar on the right indicates the temperature T/J of the samples. The panels (a-c) are for $N = 20^2, 40^2$ and 80^2 sites respectively.

Unsupervised learning of phase transitions: from principal component analysis to variational autoencoders

Sebastian Johann Wetzel

arXiv:1703.02435

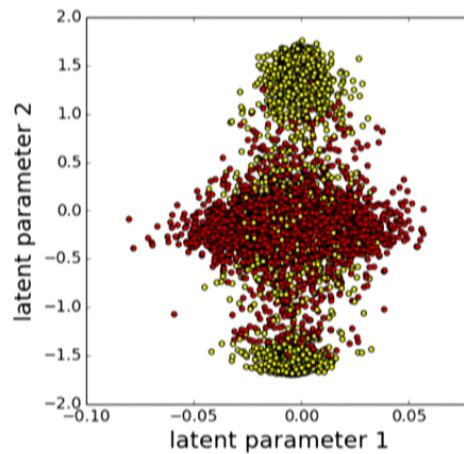
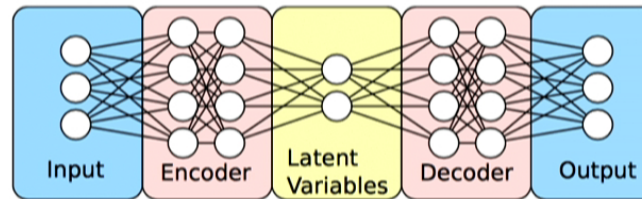
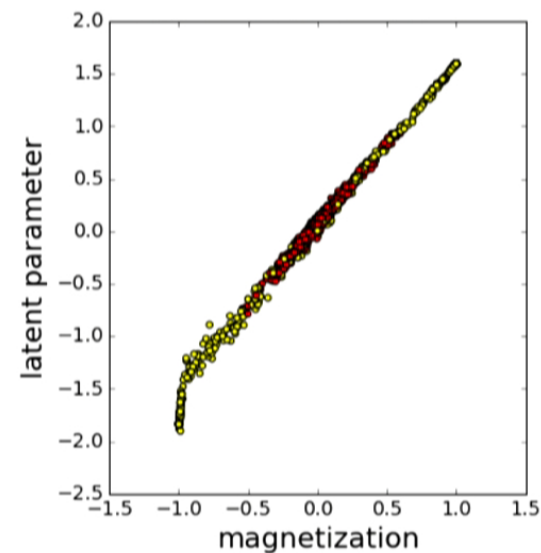


FIG. 5: Ferromagnetic Ising Model. Visualization of data in a two dimensional latent space. Red dots indicate points in the unordered phase, while yellow dots correspond to the ordered phase. The axis for parameter 1 has a smaller range than the axis for parameter 2.



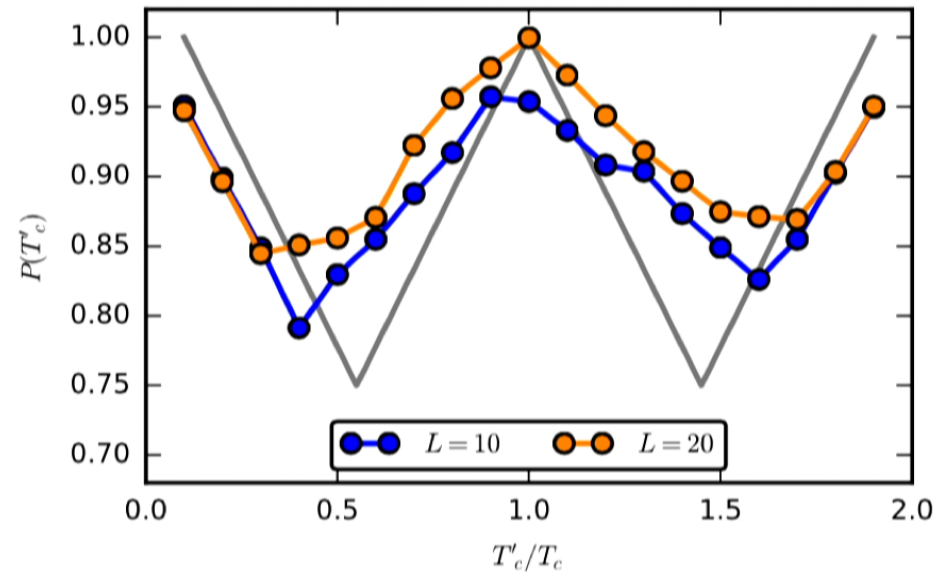
Autoencoder with one latent variable

Learning phase transitions by confusion

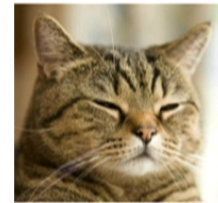
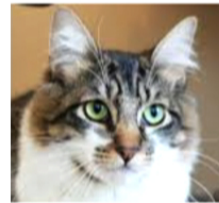
Evert P. L. van Nieuwenburg, Ye-Hua Liu & Sebastian D. Huber

Nature Physics **13**, 435–439 (2017)

Here, we propose to use a neural network based approach to find phase transitions depending on the performance of the neural network after training it with **deliberately incorrectly labelled data**. We demonstrate the success of this method on the topological phase transition in the Kitaev chain, the thermal phase transition in the classical Ising model, and the many-body-localization transition in a disordered quantum spin chain.

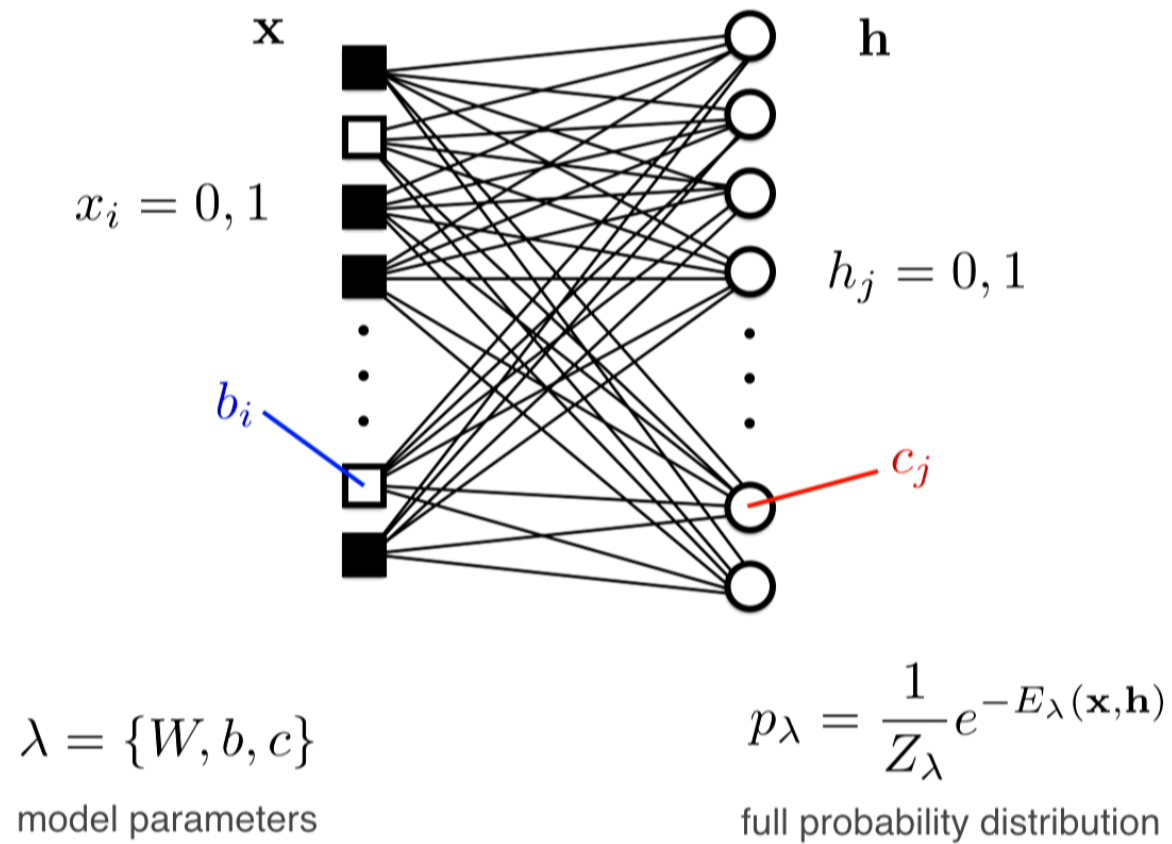


Other (very recent) applications of UNSUPERVISED machine learning in many-body physics



Restricted Boltzmann Machine

Smolensky, Hinton, Salakhutdinov, Bengio

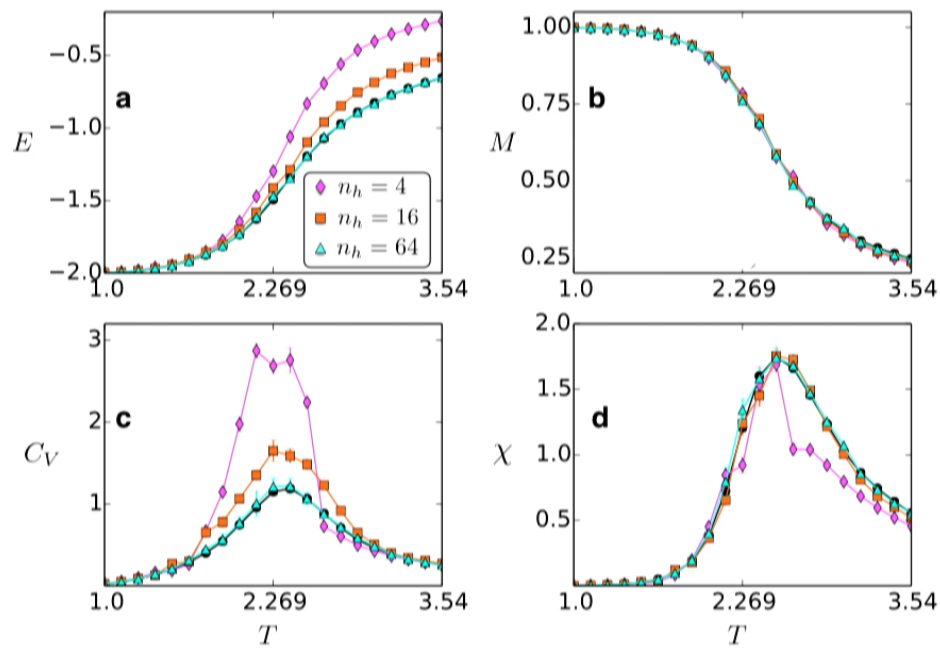


Learning thermodynamics with Boltzmann machines

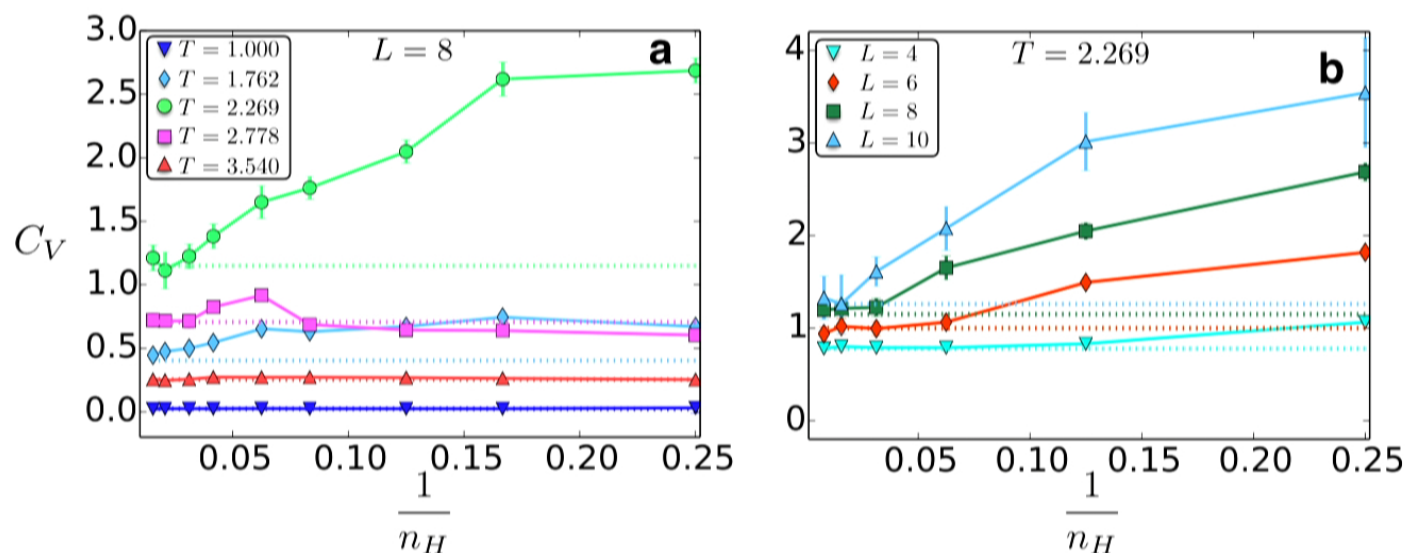
Giacomo Torlai and Roger G. Melko

Phys. Rev. B **94**, 165134 – Published 17 October 2016

Results from the generative model, after training: $\langle \mathcal{O} \rangle_\lambda = \frac{1}{N_{\text{BG}}} \sum_i \mathcal{O}[\mathbf{x}_\lambda^{(i)}]$



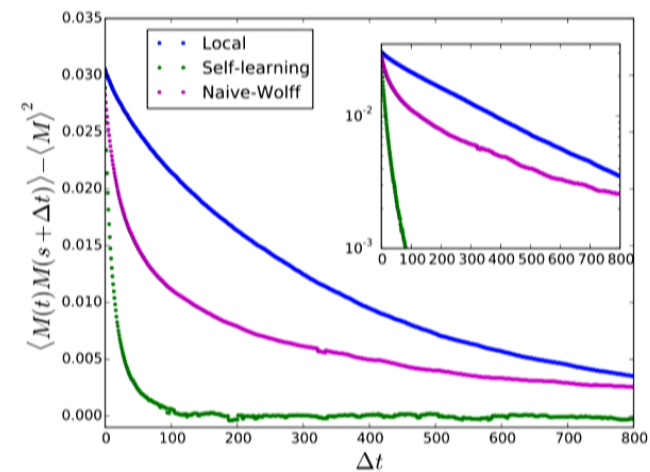
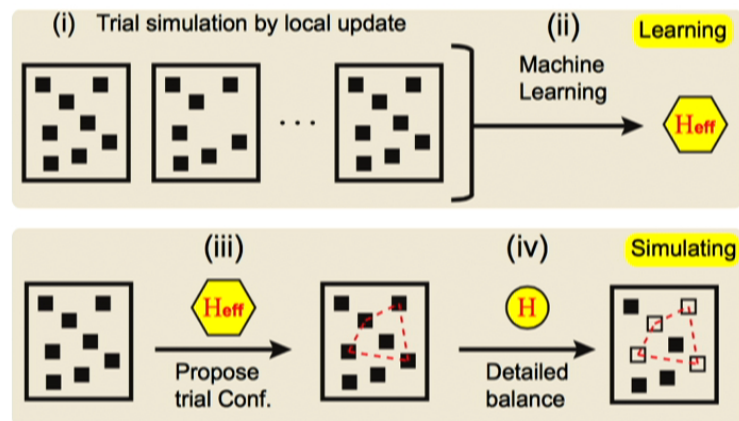
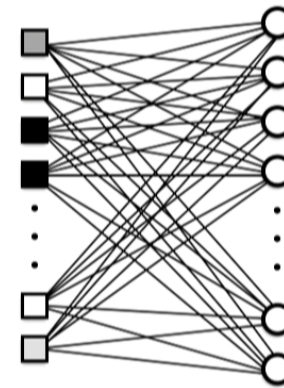
One can probe the dependence of physical observables on the hyper-parameters, e.g. n_H



The number of hidden units required for faithful reproduction of physical observables is *larger* at the critical point

Self-learning Monte Carlo method

Junwei Liu, Yang Qi, Zi Yang Meng, and Liang Fu
 Phys. Rev. B **95**, 041101(R) – Published 4 January 2017



Accelerated Monte Carlo simulations with restricted Boltzmann machines

Li Huang and Lei Wang

Phys. Rev. B **95**, 035105 – Published 4 January 2017

We fit the unnormalized probability of the physical model to a feedforward neural network and reinterpret the architecture as a restricted Boltzmann machine. Then, exploiting its feature detection ability, we utilize the restricted Boltzmann machine for efficient Monte Carlo updates and to speed up the simulation of the original physical system.

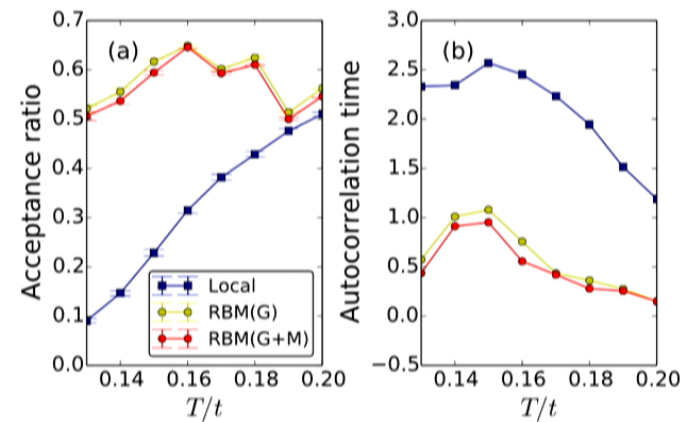
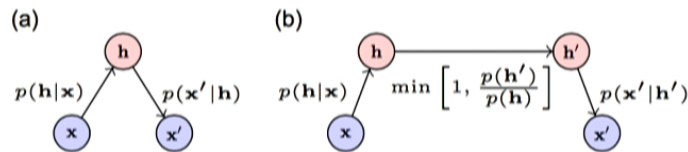
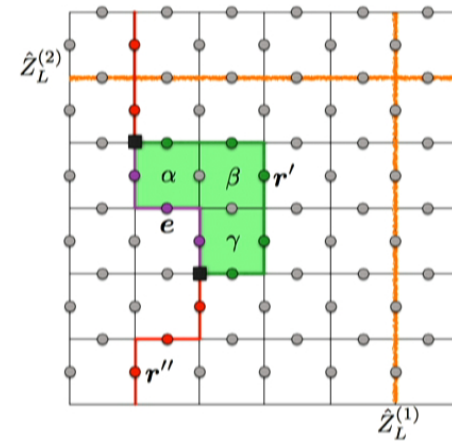
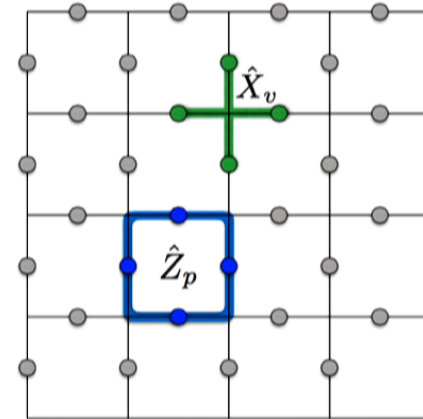
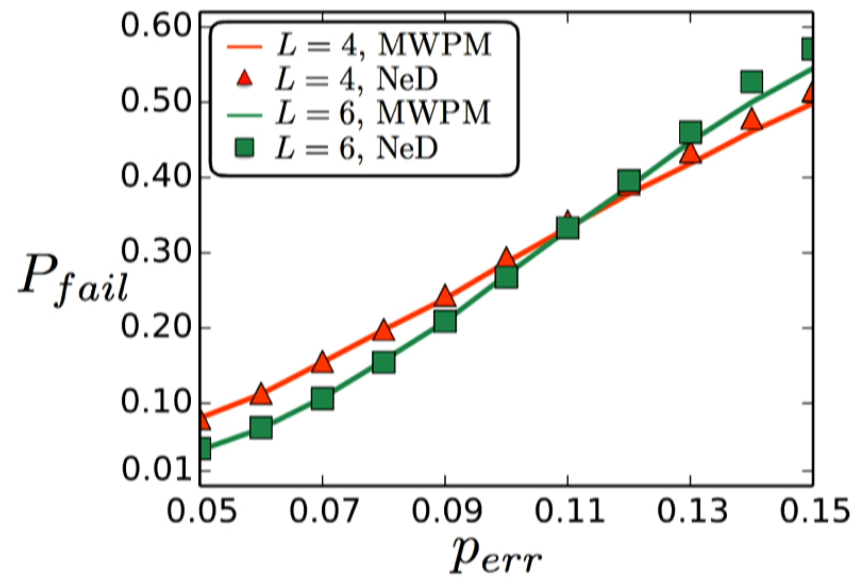


Figure 5. (a) The acceptance ratio and (b) the total energy autocorrelation time of the Falicov-Kimball model on a $N = 8^2$ square lattice. Blue squares denote results of local bit-flip updates. The yellow and red dots are using the RBM update schemes of Fig. 4(a,b) respectively. The critical temperature is at $T/t \approx 0.15$ [33–35]. The estimated physical observables agree within errorbars for all sampling approaches.

Neural Decoder for Topological Codes

Giacomo Torlai and Roger G. Melko

Phys. Rev. Lett. **119**, 030501 – Published 18 July 2017



Representing quantum wavefunctions with RBMs

A postulate of quantum mechanics:

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 dx = 1$$

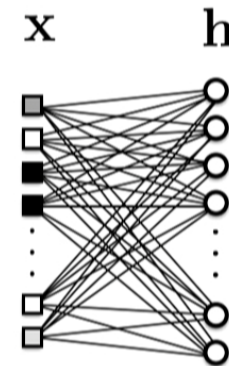
$$|\Psi(x)|^2 = \Psi^*(x)\Psi(x)$$

We can use an RBM to fully represent the wave function of a quantum system, and train an RBM (provided that data can be found; either synthetic or experimental).

$$p_{\lambda}(\mathbf{x}) = \sum_{\mathbf{h}} p_{\lambda}(\mathbf{x}, \mathbf{h}) = \frac{1}{Z_{\lambda}} e^{-F(\mathbf{x})}$$

$$\Psi(\mathbf{x}) \sim \sqrt{p_{\lambda}(\mathbf{x})}$$

(ok for positive definite wavefunctions)



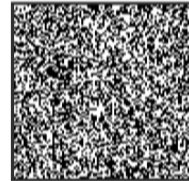
Many-body quantum state tomography with neural networks

Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, Giuseppe Carleo

arXiv:1703.05334

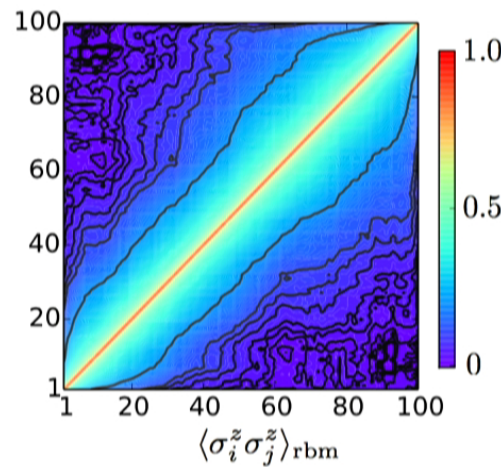
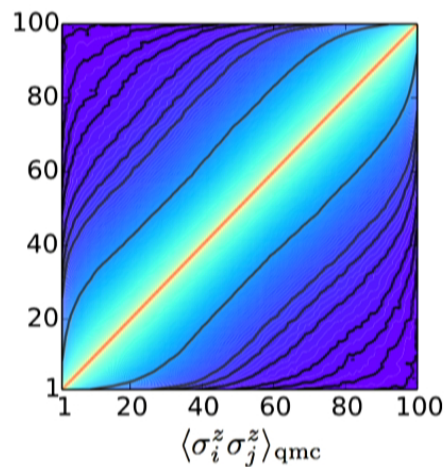
$$H = - \sum_{\langle ij \rangle} S_i^z S_j^z - \Gamma \sum_i S_i^x$$

e.g. sample quantum ising model
"wavefunction" in the S^z basis



$$\mathcal{O}_{\sigma\sigma'}^D = \mathcal{O}(\sigma) \delta_{\sigma\sigma'}$$

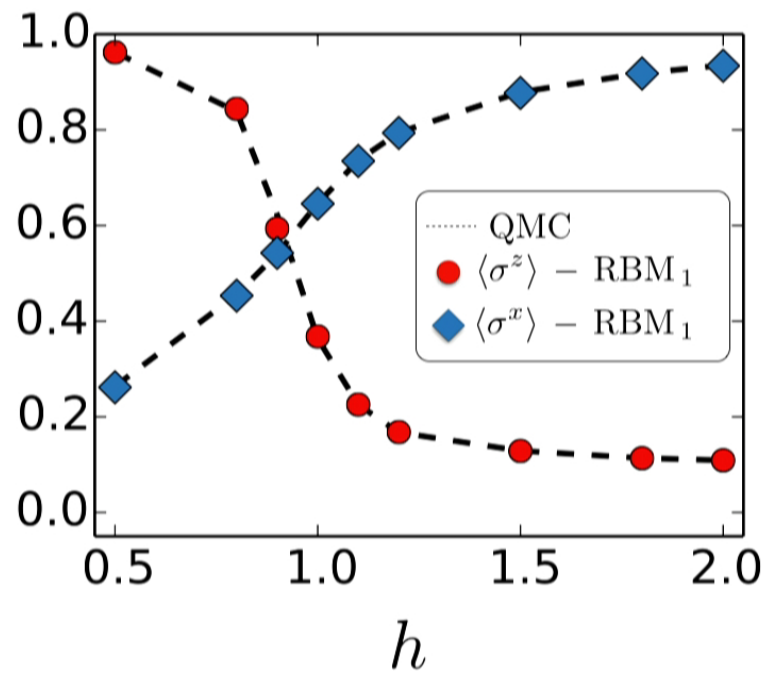
$$\mathcal{O} = \sum_{\sigma, \sigma'} \mathcal{O}_{\sigma\sigma'} |\sigma\rangle \langle \sigma'| \quad \langle \mathcal{O} \rangle = \sum_{\sigma, \sigma'} \psi_{\lambda}(\sigma) \psi_{\lambda}(\sigma') \mathcal{O}_{\sigma\sigma'}$$



$$\begin{aligned} \langle \mathcal{O}^D \rangle &= \sum_{\sigma} |\psi_{\lambda}(\sigma)|^2 \mathcal{O}(\sigma) \\ &= \frac{1}{Z_{\lambda}} \sum_{\sigma} p_{\lambda}(\sigma) \mathcal{O}(\sigma) \\ &\simeq \frac{1}{n} \sum_{k=1}^n \mathcal{O}(\sigma_k) \end{aligned}$$

Off-diagonal observables

$$\langle \mathcal{O}^{ND} \rangle = \sum_{\sigma} |\psi_{\lambda}(\sigma)|^2 \mathcal{O}_L(\sigma) \simeq \frac{1}{n} \sum_{k=1}^n \mathcal{O}_L(\sigma_k) \quad \mathcal{O}_L(\mathbf{x}) = \sum_{\mathbf{x}'} \sqrt{\frac{p_{\lambda}(\mathbf{x}')}{p_{\lambda}(\mathbf{x})}} \mathcal{O}_{\mathbf{x}\mathbf{x}'}$$



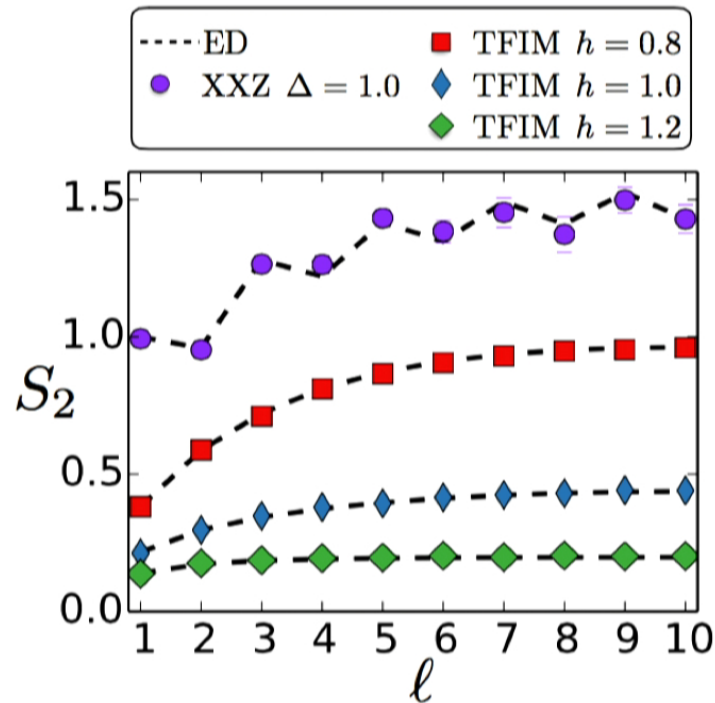
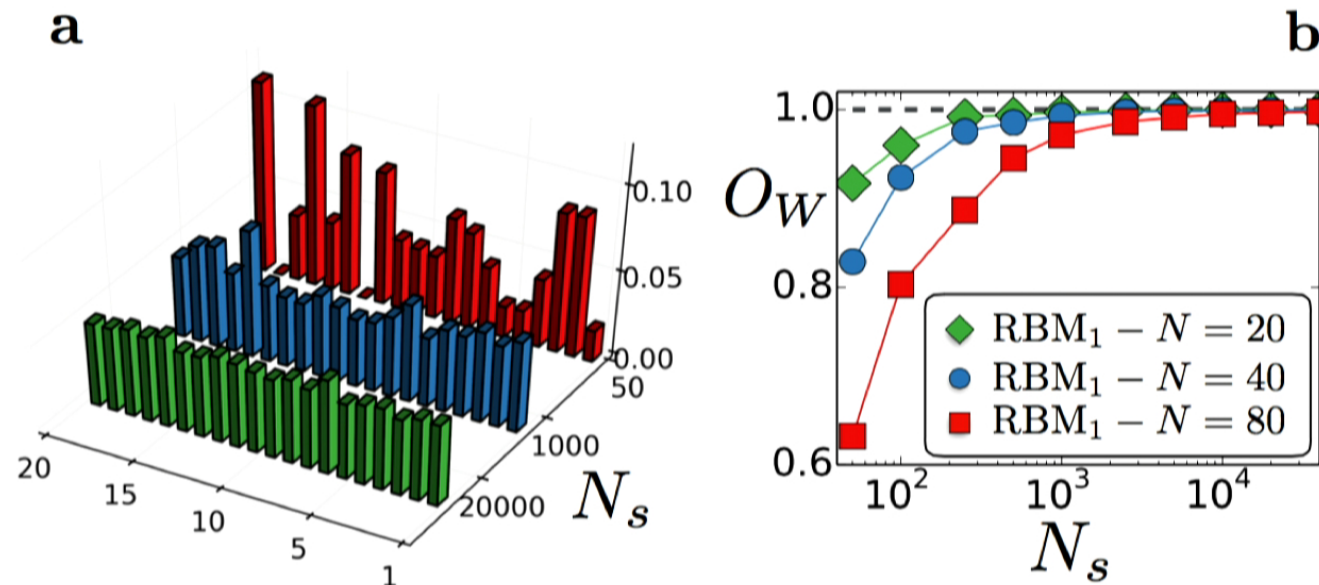


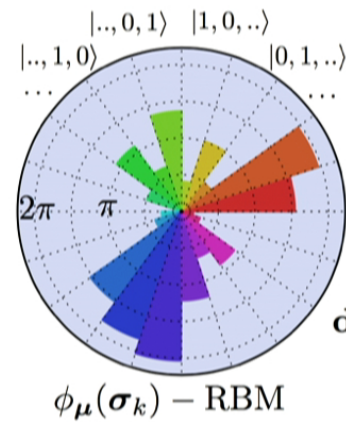
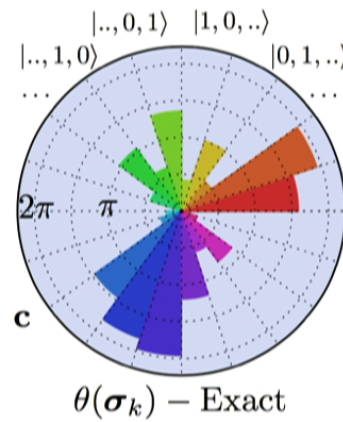
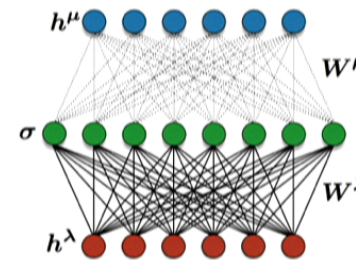
Figure 3. Reconstruction of the entanglement entropy. Second Renyi entropy as a function of the subsystem size ℓ for $N = 20$ spins. We compare results obtained using the RBM wave-functions (markers) with exact diagonalization (dashed lines) for the 1d-TFIM at different values of the transverse magnetic field h and the 1d-XXZ model with critical anisotropy $\Delta = 1$.

$$|\Psi_W\rangle = \frac{1}{\sqrt{N}} (|100\dots\rangle + \dots + |\dots 001\rangle)$$



$$|\tilde{\Psi}_W\rangle = \frac{1}{\sqrt{N}} \left(e^{i\theta_1} |100\dots\rangle + e^{i\theta_2} |010\dots\rangle + \dots + e^{i\theta_N} |0\dots 01\rangle \right)$$

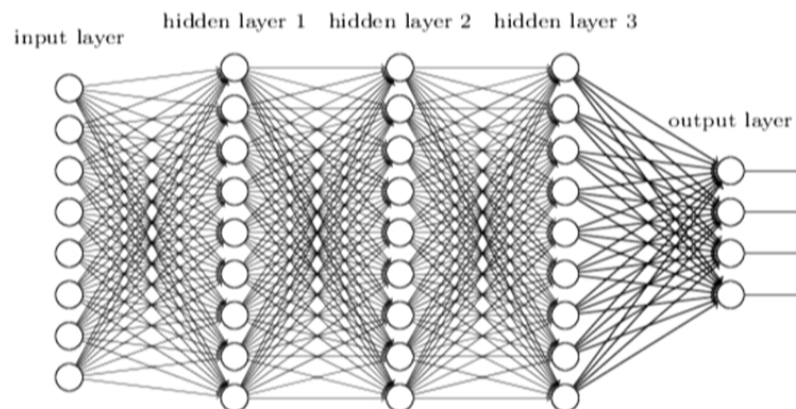
$$\Psi(\mathbf{x}) \sim \sqrt{p_\lambda(\mathbf{x})} e^{i\phi_\mu(\mathbf{x})}$$



$N = 20$ qubits
 99.7% fidelity
 2×10^5 measurements

Deep learning

It is empirically well supported that deep neural net architectures perform better for supervised learning of complex, hierarchal data

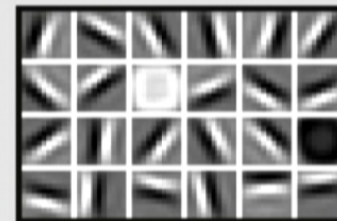


FACIAL RECOGNITION

Deep-learning neural networks use layers of increasingly complex rules to categorize complicated shapes such as faces.



Layer 1: The computer identifies pixels of light and dark.



Layer 2: The computer learns to identify edges and simple shapes.



Layer 3: The computer learns to identify more complex shapes and objects.

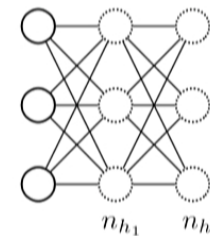
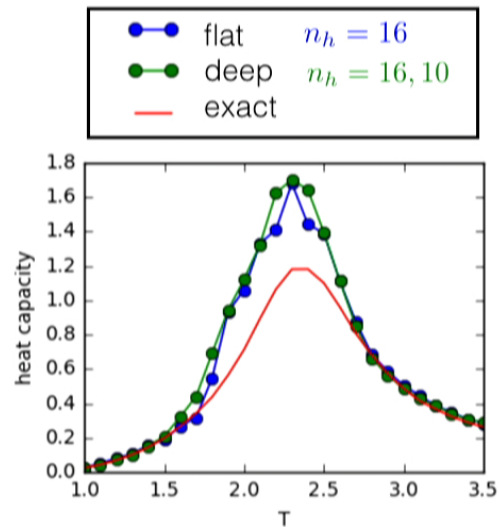
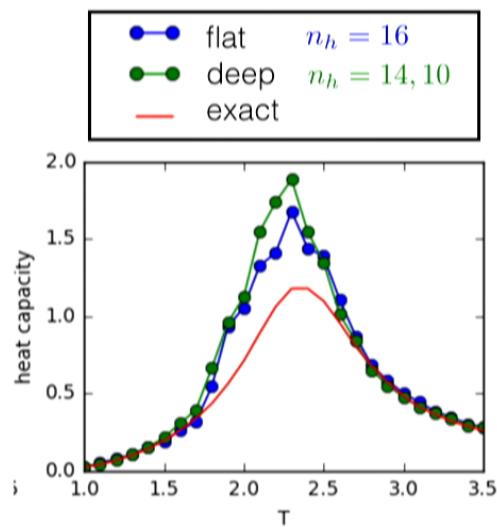


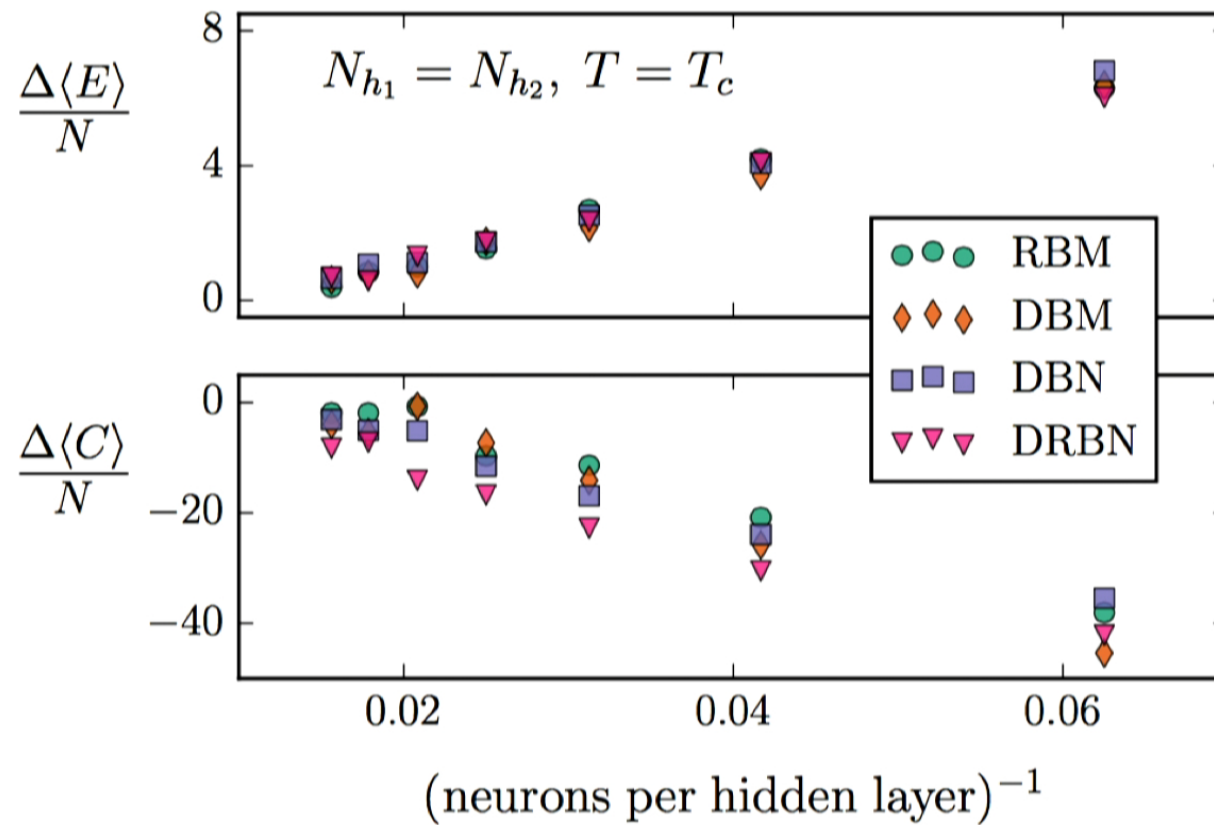
Layer 4: The computer learns which shapes and objects can be used to define a human face.

Does representational power depend on depth?

This can be concretely addressed in physical systems like the Ising model

Morningstar and RGM, arXiv:1708.04622



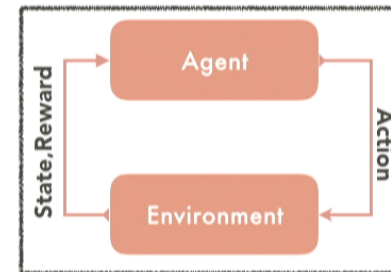


RBM based reinforcement learning

The goal is to find the optimal policy for an agent in a dynamic environment as evaluated by the total amount of received reward.

Q-learning: $Q(s_0, a_0) = \langle \sum_{n=1}^{\infty} \gamma^n R(s_n) \rangle$

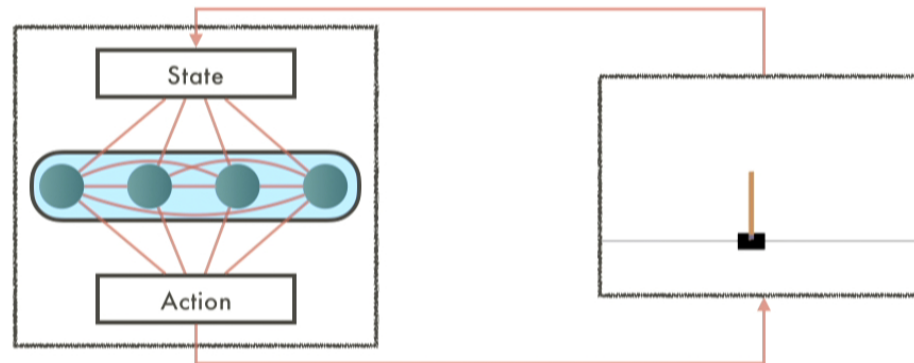
γ - discount factor
 $R(s_n)$ - reward



The state space can be exponentially large - parametrize it as a free energy of a Boltzmann machine

$$Q(s, a) = -\log \text{Tr}_h [e^{-\hat{H}(s, a)}]$$

$$\delta w \propto [R(s_0) + \gamma \max_a Q(s_1, a) - Q(s_0, a_0)] \frac{\partial}{\partial w} Q(s_0, a_0)$$

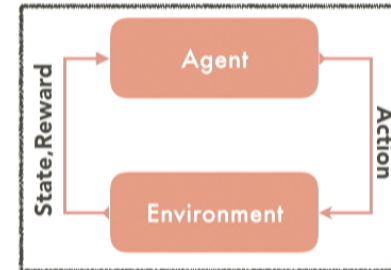


RBM based reinforcement learning

The goal is to find the optimal policy for an agent in a dynamic environment as evaluated by the total amount of received reward.

Q-learning: $Q(s_0, a_0) = \langle \sum_{n=1}^{\infty} \gamma^n R(s_n) \rangle$

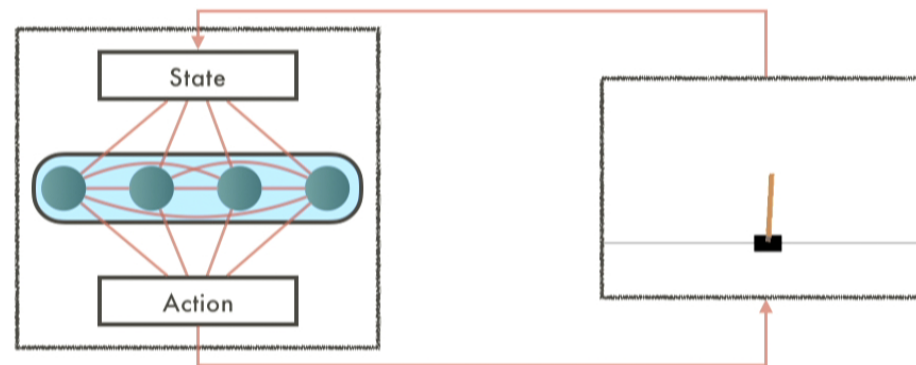
γ - discount factor
 $R(s_n)$ - reward



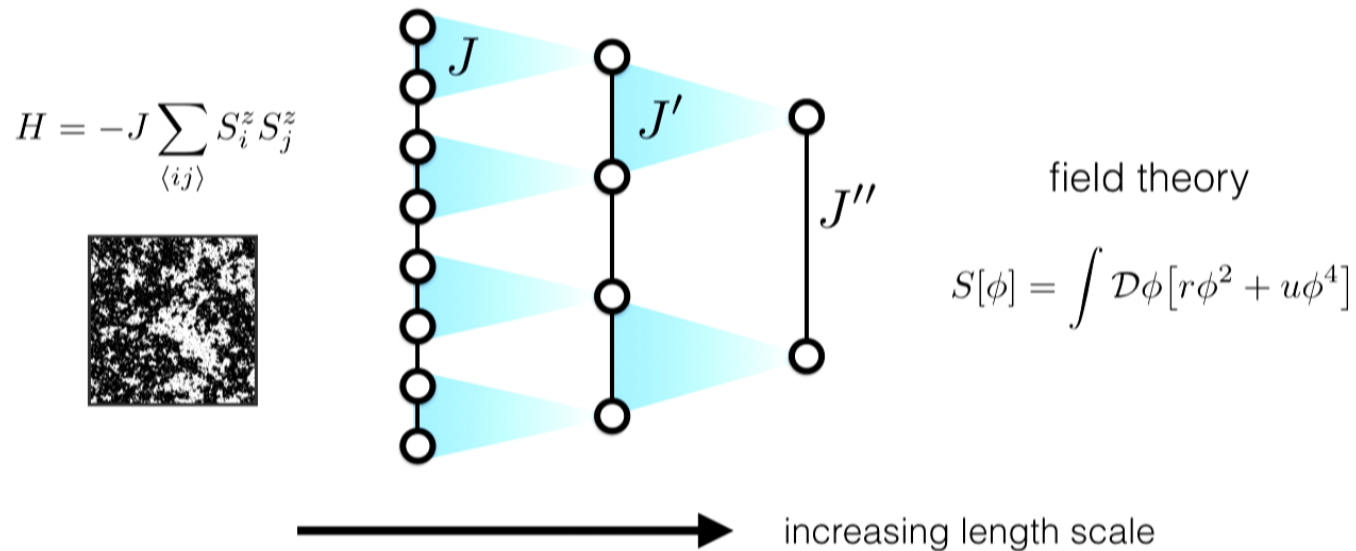
The state space can be exponentially large - parametrize it as a free energy of a Boltzmann machine

$$Q(s, a) = -\log \text{Tr}_h [e^{-\hat{H}(s, a)}]$$

$$\delta w \propto [R(s_0) + \gamma \max_a Q(s_1, a) - Q(s_0, a_0)] \frac{\partial}{\partial w} Q(s_0, a_0)$$



Q: are **deep networks** required for efficient representation & feature extraction at phase transitions?



Deep learning and the renormalization group
Cédric Bény, arXiv:1301.3124

An exact mapping between the Variational Renormalization Group and Deep Learning
P. Mehta, D. Schwab, arXiv:1410.3831

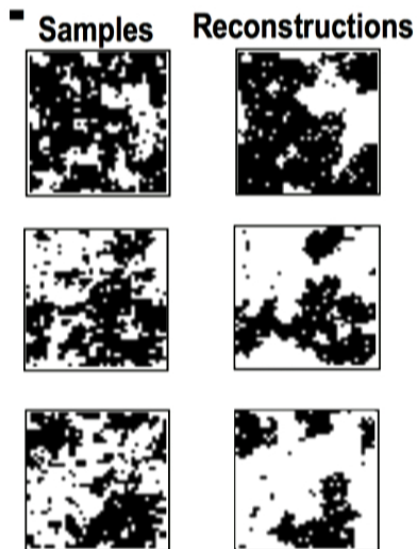
Why does deep and cheap learning work so well?
Lin and Tegmark, arXiv:1608.08225

An exact mapping between the Variational Renormalization Group and Deep Learning

Pankaj Mehta, David J. Schwab

arXiv:1410.3831

We construct an exact mapping from the variational renormalization group, first introduced by Kadanoff, and deep learning architectures based on Restricted Boltzmann Machines (RBMs)

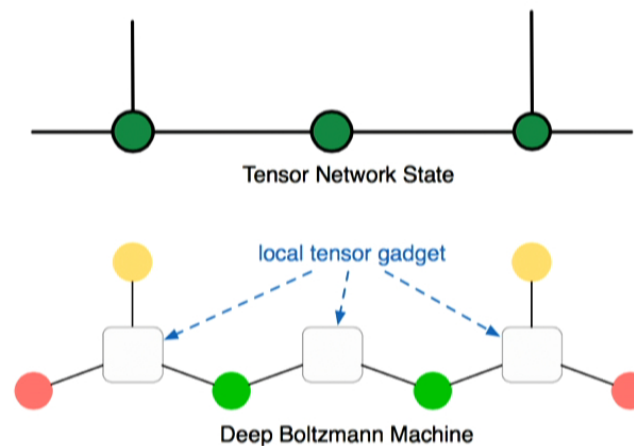


Efficient Representation of Quantum Many-body States with Deep Neural Networks

Xun Gao, Lu-Ming Duan

arXiv:1701.05039

Here, we give a rigorous proof that a deep neural network can efficiently represent most physical states, including those generated by any polynomial size quantum circuits or ground states of many-body Hamiltonians with polynomial-size gaps, while a shallow network through a restricted Boltzmann machine cannot efficiently represent those states unless the polynomial hierarchy in computational complexity theory collapses.

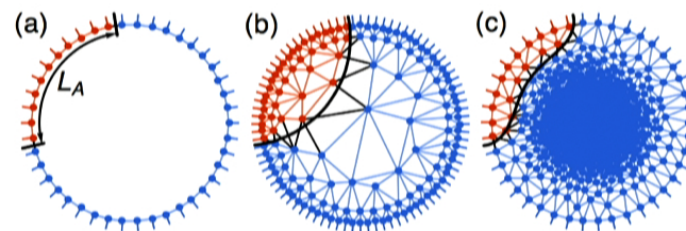
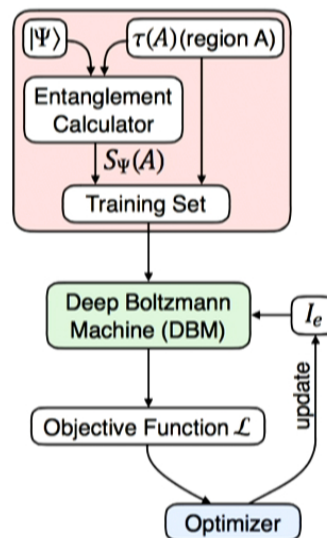


Machine Learning Spatial Geometry from Entanglement Features

Yi-Zhuang You, Zhao Yang, Xiao-Liang Qi

arXiv:1709.01223

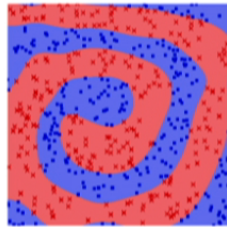
Motivated by the close relations of the renormalization group with both the holography duality and the deep learning, we propose that the holographic geometry can emerge from deep learning the entanglement feature of a quantum many-body state....We show that each RTN can be mapped to a Boltzmann machine, trained by the entanglement entropies over all subregions of a given quantum many-body state. The goal is to construct the optimal RTN that best reproduce the entanglement feature. T



Tensor Network Inspired Machine Learning

$$\Phi = \begin{matrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ \bigcirc & \bigcirc & \bigcirc & \bigcirc & \bigcirc & \bigcirc \\ \phi^{s_1} & \phi^{s_2} & \phi^{s_3} & \phi^{s_4} & \phi^{s_5} & \phi^{s_6} \end{matrix}$$

Feature map as a tensor product of local feature maps



Stoudenmire, Schwab,
arXiv:1605.05775



<https://physicsml.github.io/tensor-network.html>

Tensor Networks: Putting Quantum Wavefunctions into Machine Learning

by Miles Stoudenmire in Articles on 2016-11-07 | tags: tensor networks, quantum mechanics, machine learning, kernel learning, support vector machines

Near-term applications for ML many-body physics

- Classical simulators for quantum models
- The inverse problem - learning a Hamiltonian from the wf.
- Neural network decoders for error correcting protocols
- Tomography on near-term quantum devices
- Technology transfer between physics and machine learning