Title: Machine Learning Phases of Matter

Date: Aug 11, 2016 11:45 AM

URL: http://pirsa.org/16080016

Abstract:

PERIMETER INSTITUTE FOR THEORETICAL PHYSICS

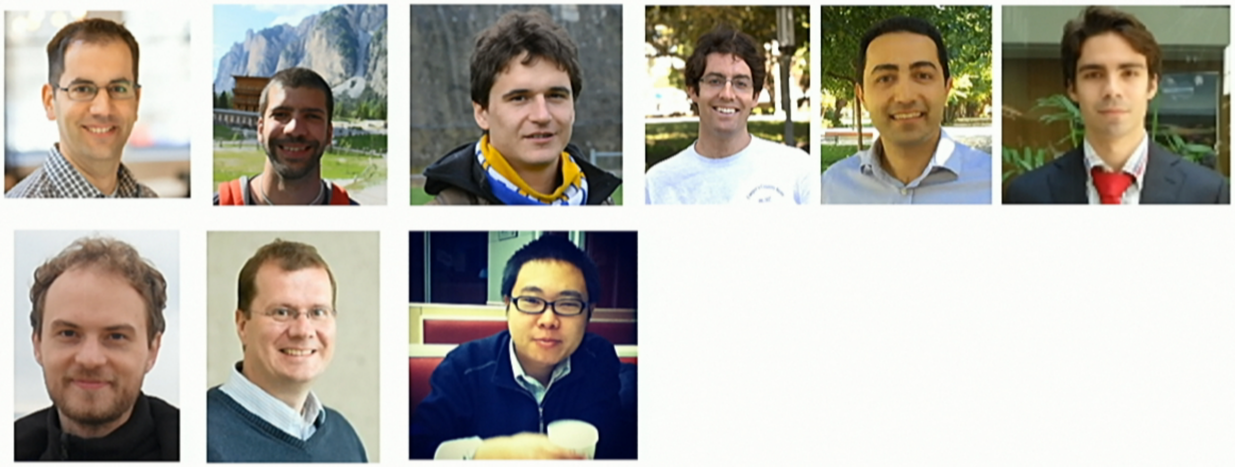
MACHINE LEARNING PHASES OF MATTER

Juan Carrasquilla
Perimeter Institute
Quantum Machine learning

COLLABORATOR

Roger Melko (Perimeter & U. Waterloo)

THANKS TO MY PHYSICS/ML FRIENDS FOR DISCUSSIONS AND COLLABORATIONS



CONTENTS

- Introduction: ML and condensed matter, phases, phase transitions.
- What we mean by learning phases of matter
- Ising model and conventional phases of matter
- Square ice, Ising gauge theory
- Sign problem
- Conclusions and outlook

CONTENTS

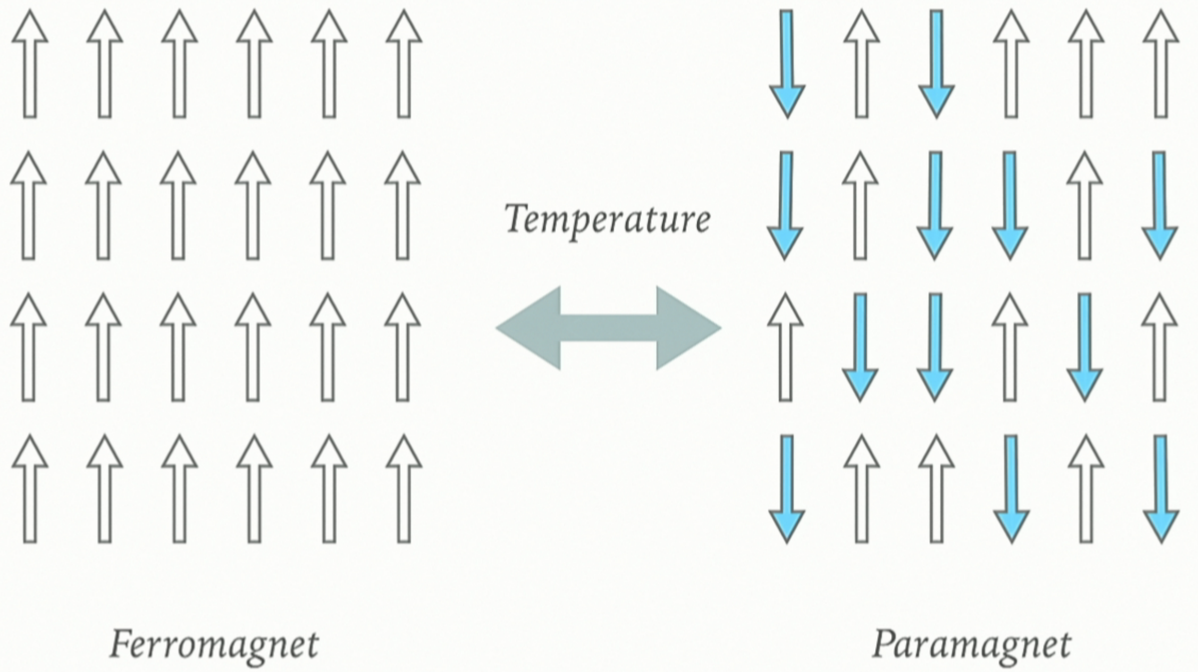
- Introduction: ML and condensed matter, phases, phase transitions.
- What we mean by learning phases of matter
- Ising model and conventional phases of matter
- Square ice, Ising gauge theory
- Sign problem
- Conclusions and outlook

CONTENTS

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ising ferromagnet in two dimensions

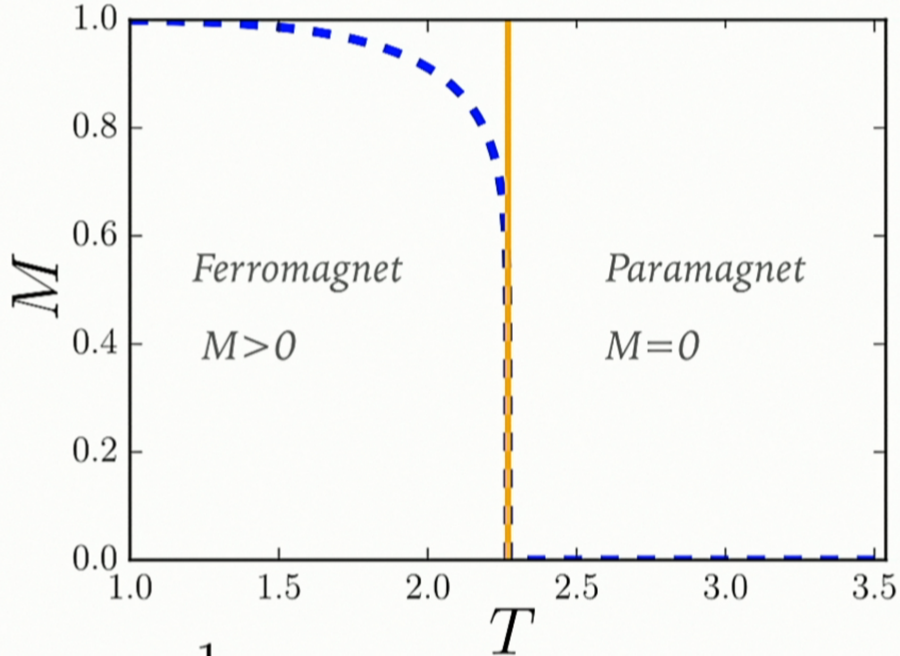
$$H = -J \sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z$$



Lars Onsager Phys. Rev. 65, 117

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ferromagnetic transition: order parameter



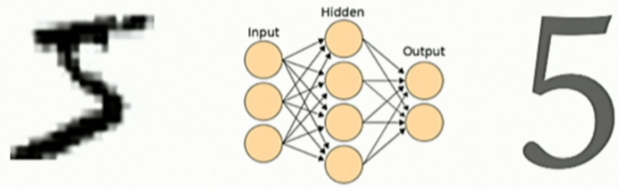
It is a measure of the degree of order in the system

$$M = \frac{1}{N}\sum_i \langle \sigma_i \rangle, \qquad \sigma_i = \pm 1$$

Lars Onsager Phys. Rev. 65, 117

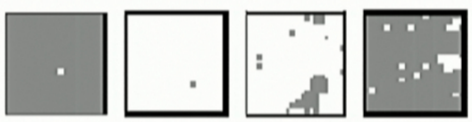
INSPIRATION: FLUCTUATIONS HANDWRITTEN DIGITS (MNIST)

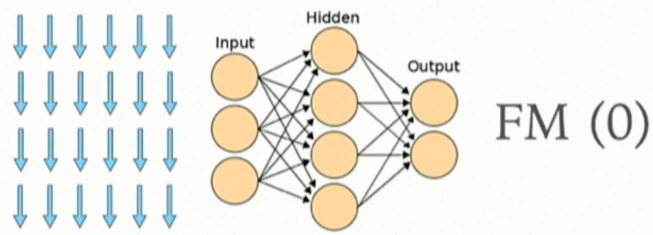
5 = 5 + *Fluctuations*



ML community has developed powerful supervised learning algorithms

FM phase

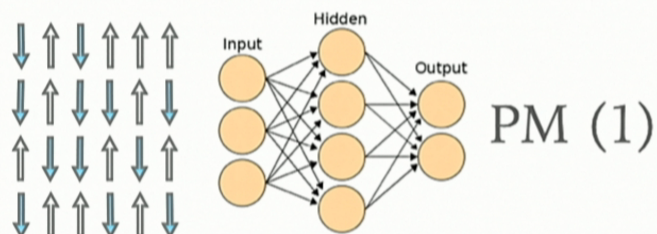




High T phase

gray=spin up

white=spin down

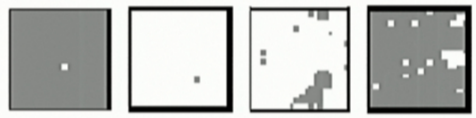


1605.01735

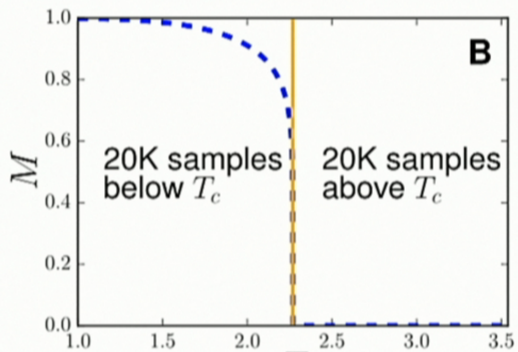
COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase



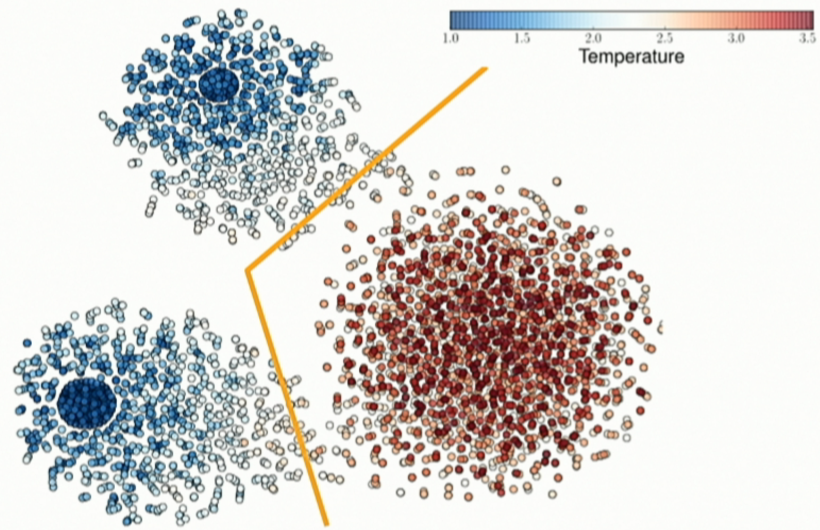
2D Ising model in the disordered phase



1605.01735

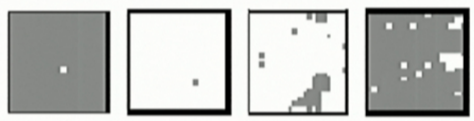
Training/testing data is drawn from the Boltzmann distribution

$$p(\sigma_1, \sigma_2, ..., \sigma_N) = \frac{e^{-\beta E(\sigma_1, \sigma_2, ..., \sigma_N)}}{Z(\beta)}$$

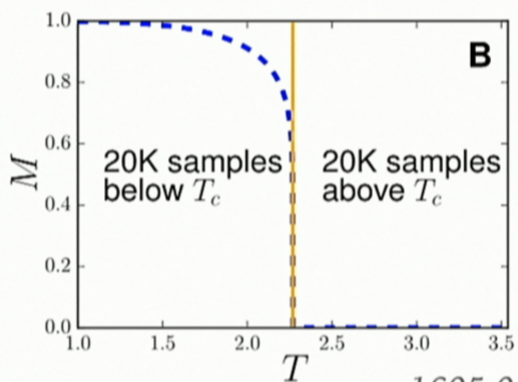


COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase



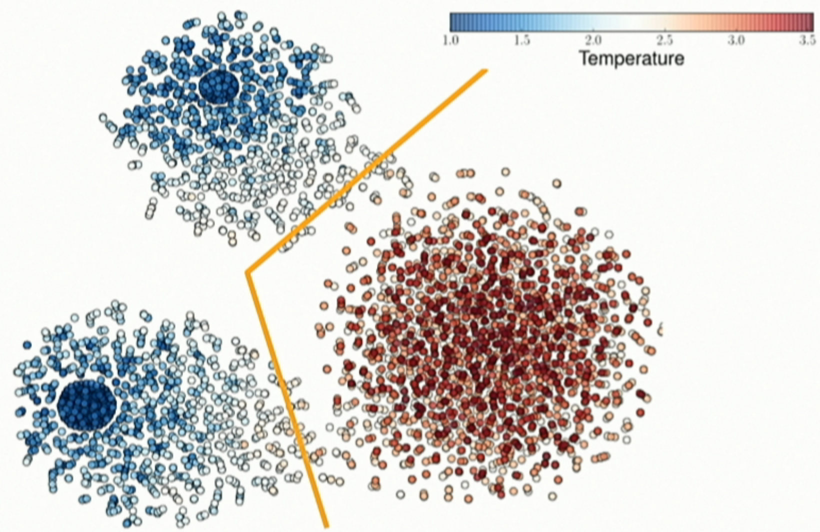
2D Ising model in the disordered phase



1605.01735

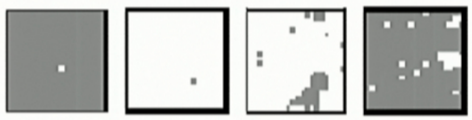
Training/testing data is drawn from the Boltzmann distribution

$$p(\sigma_1, \sigma_2, ..., \sigma_N) = \frac{e^{-\beta E(\sigma_1, \sigma_2, ..., \sigma_N)}}{Z(\beta)}$$

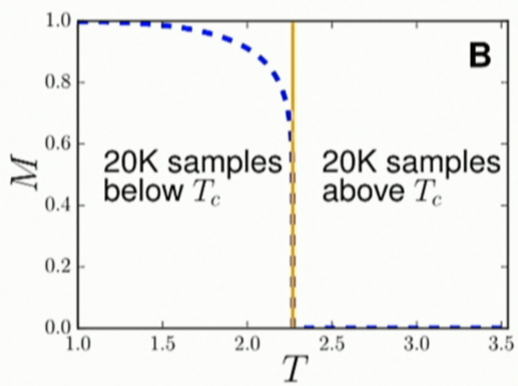


COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase

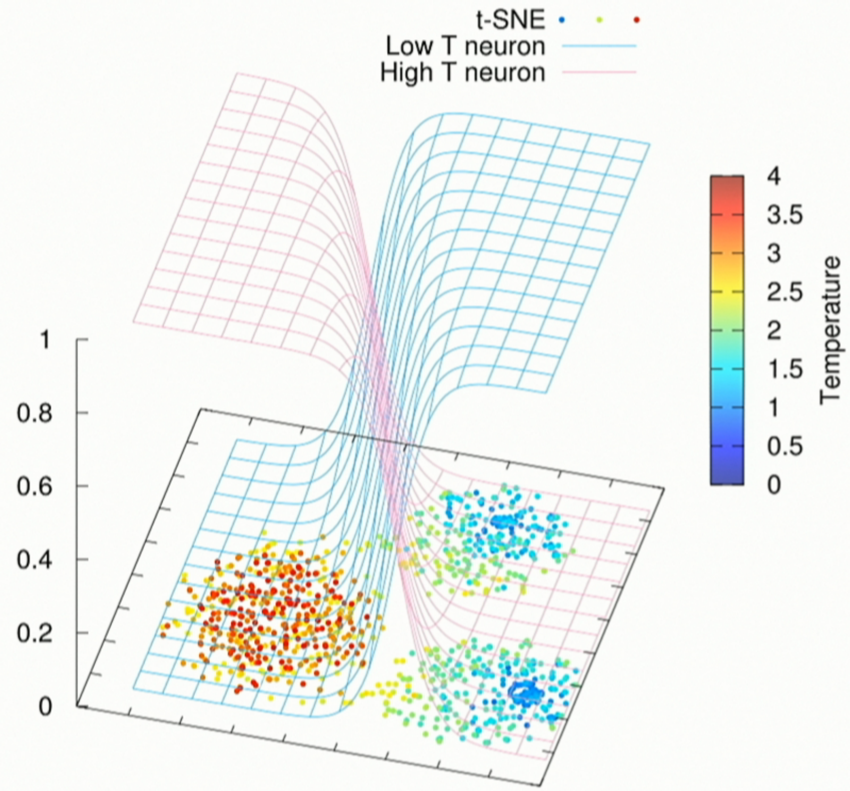


2D Ising model in the disordered phase



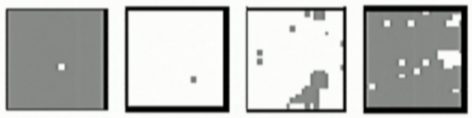
Successful training amounts to finding functions

$F_{\text{High }T}(\sigma_1, \ldots, \sigma_N)$ $F_{\text{Low }T}(\sigma_1, \ldots, \sigma_N)$

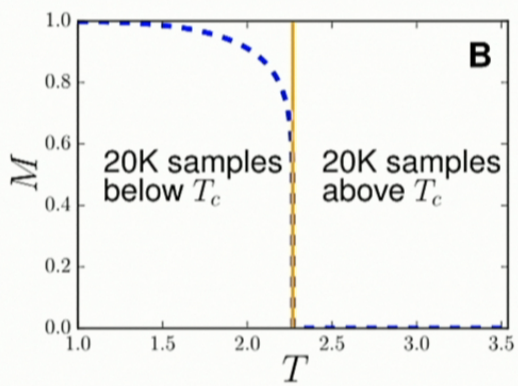


RESULTS: SQUARE LATTICE ISING MODEL (TEST SETS)

2D Ising model in the ordered phase

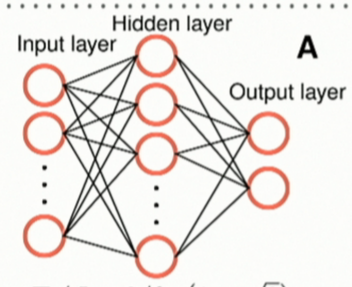


2D Ising model in the disordered phase

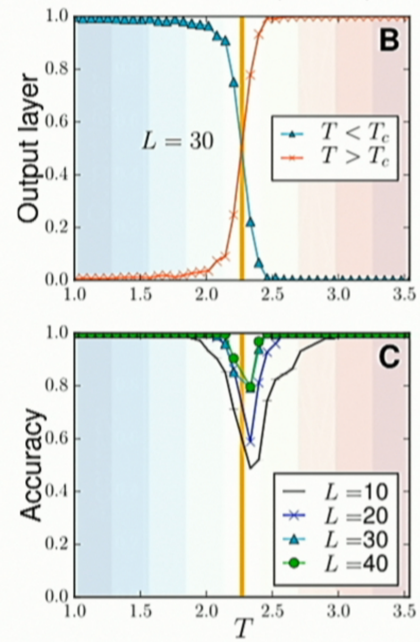


Critical

1605.01735

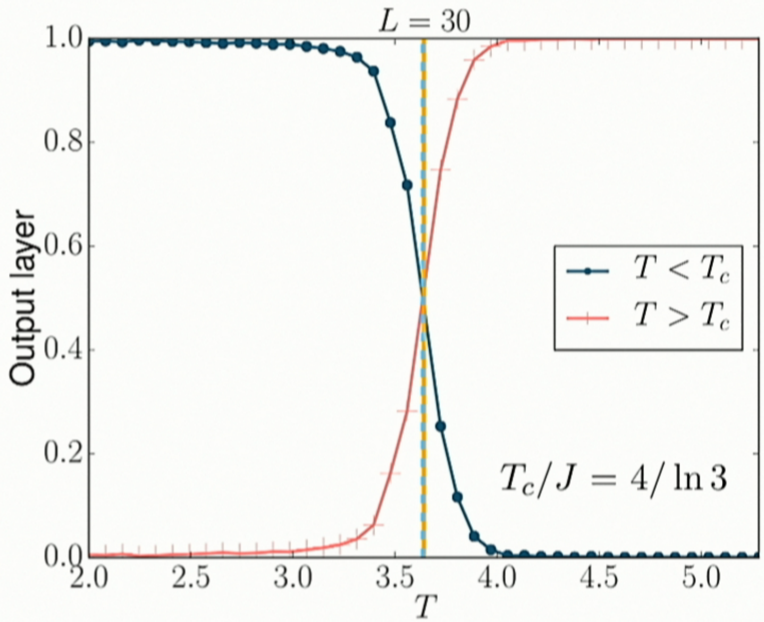


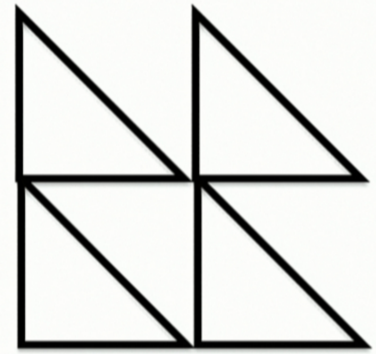
$T_c/J = 2/\ln(1+\sqrt{2})$



DO THE RESULTS EXTEND TO OTHER INTERESTING CASES?

Yes. We can obtain Tc in the triangular lattice from numerically trained model on the square lattice!

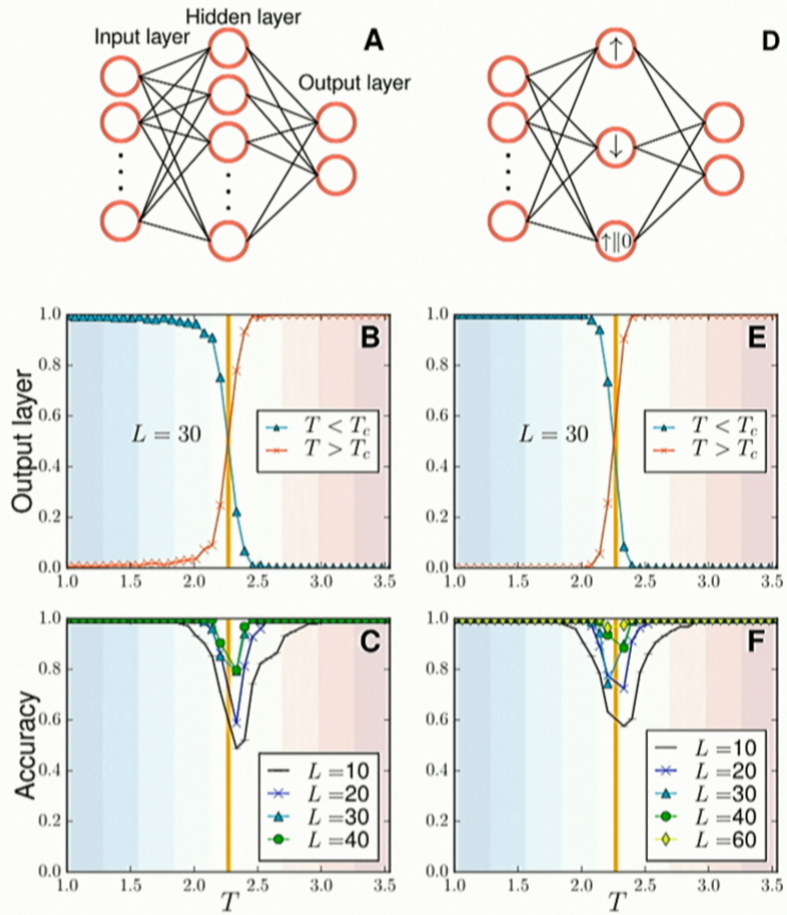




Tc within <1%!

1605.01735

ANALYTICAL UNDERSTANDING



Toy model: only three analytically "trained" perceptrons with precise functions: quantifying the magnetization of each configuration

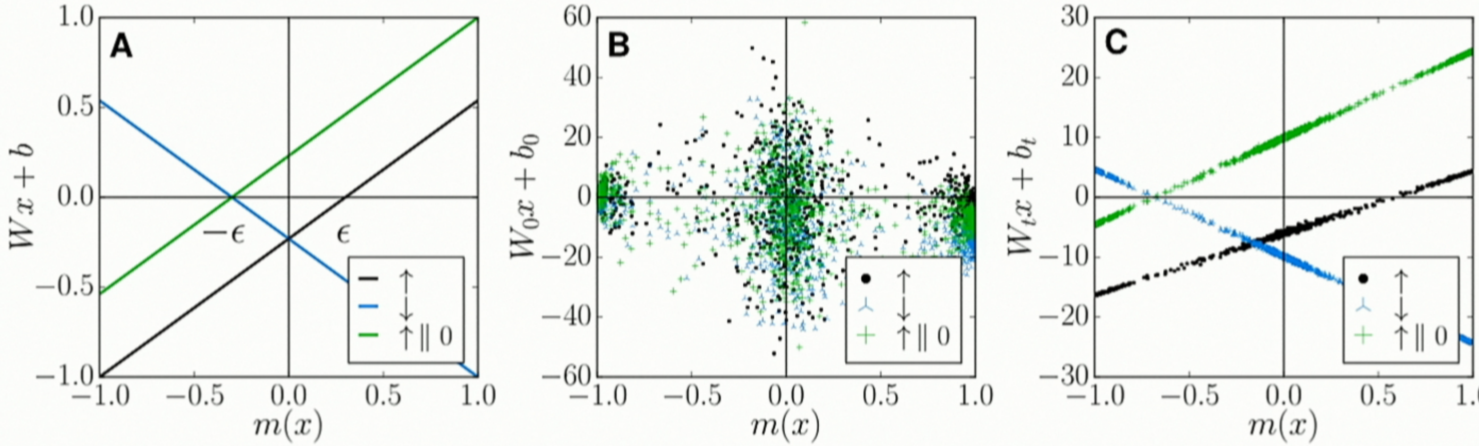
In general the hidden layer discovers the order parameter of the phase during the training

=>Works for AF Ising model

1605.01735

ANALYTICAL UNDERSTANDING

Investigating the argument of the hidden layer during the training



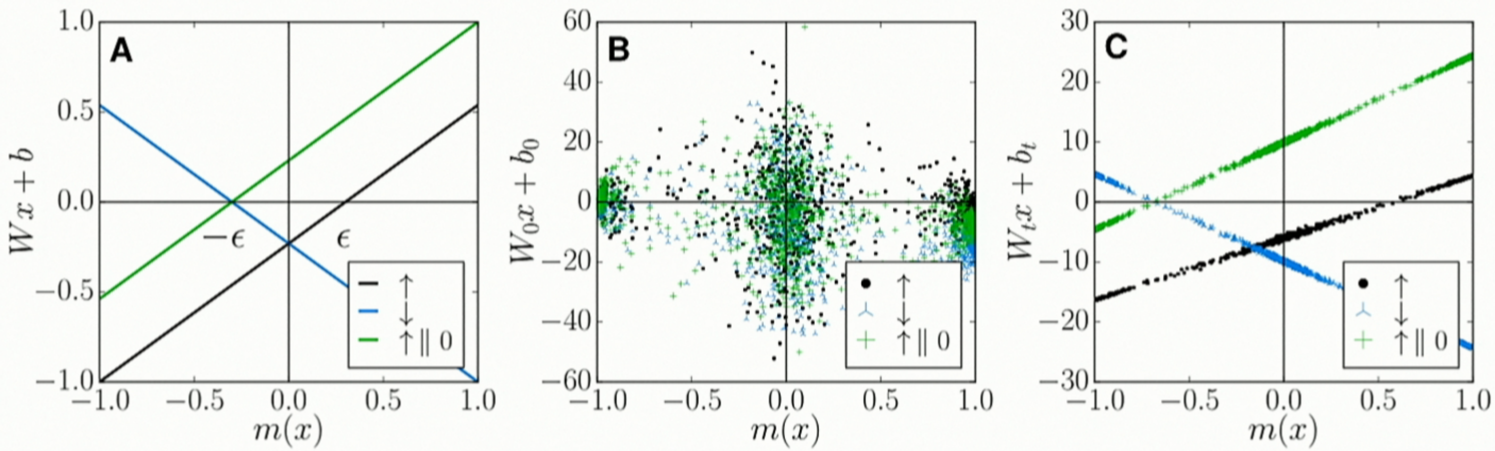
$$W = \frac{1}{N(1+\epsilon)} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \text{ and } b = \frac{\epsilon}{(1+\epsilon)} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \qquad Wx + b = \frac{1}{(1+\epsilon)} \begin{pmatrix} m(x) - \epsilon \\ -m(x) - \epsilon \\ m(x) + \epsilon \end{pmatrix},$$

$$x = [\sigma_1 \sigma_2, ..., \sigma_N]^{\mathrm{T}} \qquad m(x) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i$$

1605.01735

ANALYTICAL UNDERSTANDING

Investigating the argument of the hidden layer during the training



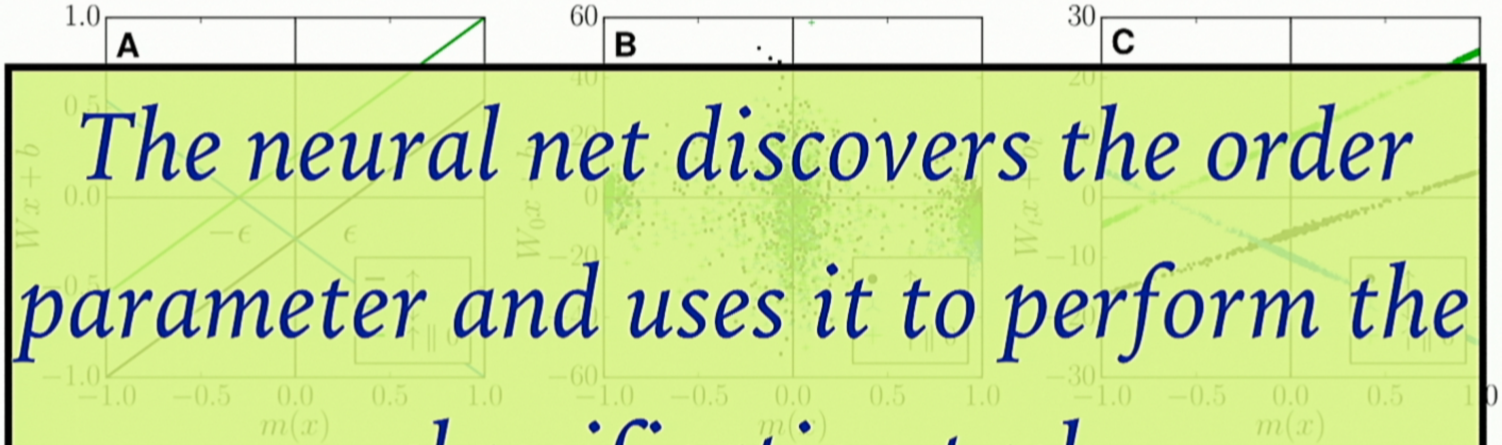
$$W = \frac{1}{N(1+\epsilon)} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \text{ and } b = \frac{\epsilon}{(1+\epsilon)} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \qquad Wx + b = \frac{1}{(1+\epsilon)} \begin{pmatrix} m(x) - \epsilon \\ -m(x) - \epsilon \\ m(x) + \epsilon \end{pmatrix},$$

$$x = [\sigma_1 \sigma_2, ..., \sigma_N]^{\mathrm{T}} \qquad m(x) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i$$

1605.01735

ANALYTICAL UNDERSTANDING

Investigating the argument of the hidden layer during the training



The neural net discovers the order parameter and uses it to perform the classification task

$$W = \frac{1}{N(1+\epsilon)} \begin{pmatrix} \cdots \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \text{ and } b = \frac{\epsilon}{(1+\epsilon)} \begin{pmatrix} \cdots \\ -1 \\ 1 \end{pmatrix}, \qquad Wx + b = \frac{1}{(1+\epsilon)} \begin{pmatrix} \cdots \\ -m(x) - \epsilon \\ m(x) + \epsilon \end{pmatrix},$$

$$x = [\sigma_1 \sigma_2, ..., \sigma_N]^{\mathrm{T}} \qquad m(x) = \frac{1}{N} \sum_{i=1}^{N} \sigma_i$$

SQUARE ICE, ISING GAUGE THEORY AND TOPOLOGICAL PHASES OF MATTER

PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0

- Topological phases of matter. Examples: Fractional quantum hall effect. Potential applications in topological quantum computing.

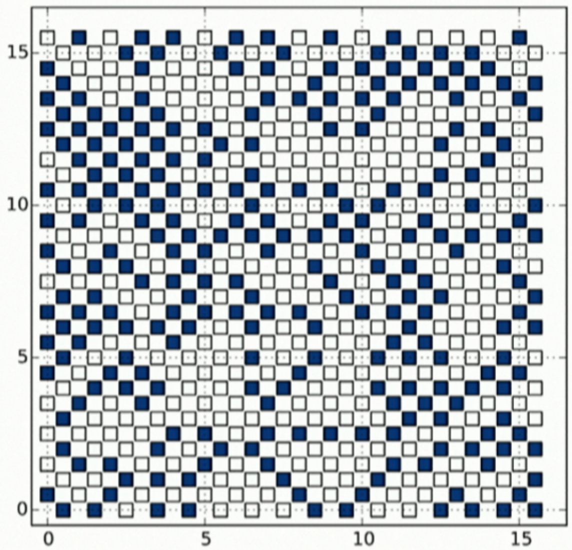
- Coulomb phases = Highly correlated "spin liquids" described by electrodynamics. Examples: Common water ice and spin ice materials ($Ho_2Ti_2O_7$ and $Dy_2Ti_2O_7$)

ST Bramwell, MJP Gingras Science 294 (5546), 1495-1501

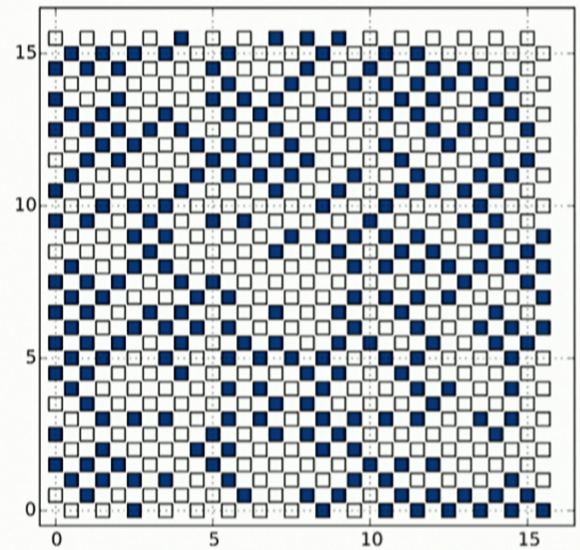
PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0

- Defy the Landau symmetry breaking classification. Neural nets capture the subtle differences between low- and high-temperature states successfully!
- Square ice: 99% accuracy
- Ising gauge theory: 50% (guessing) with a fully-connected neural net. Training fails. How to overcome this issue?

For two configurations

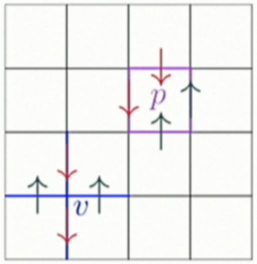


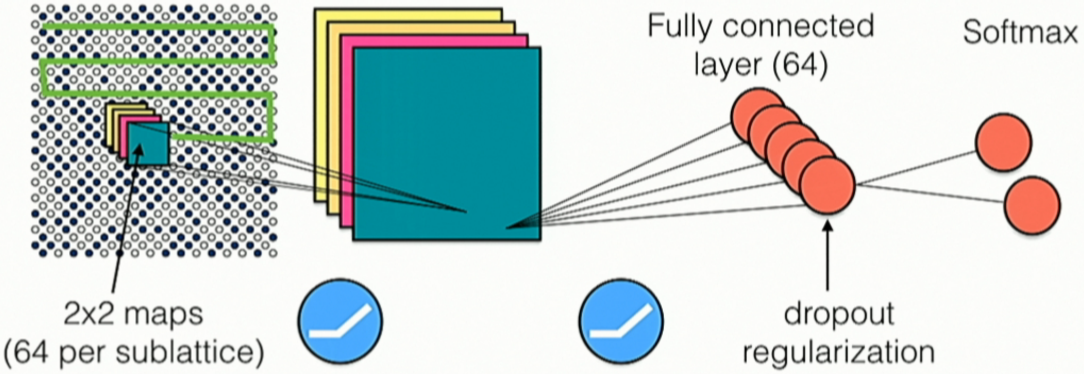
?



ISING GAUGE THEORY *(Kogut Rev. Mod. Phys. 51, 659 (1979))*

$$H = -J \sum_p \prod_{i \in p} \sigma_i^z$$

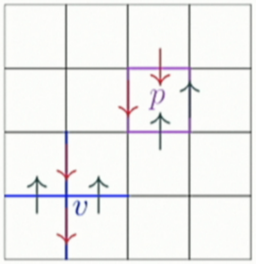


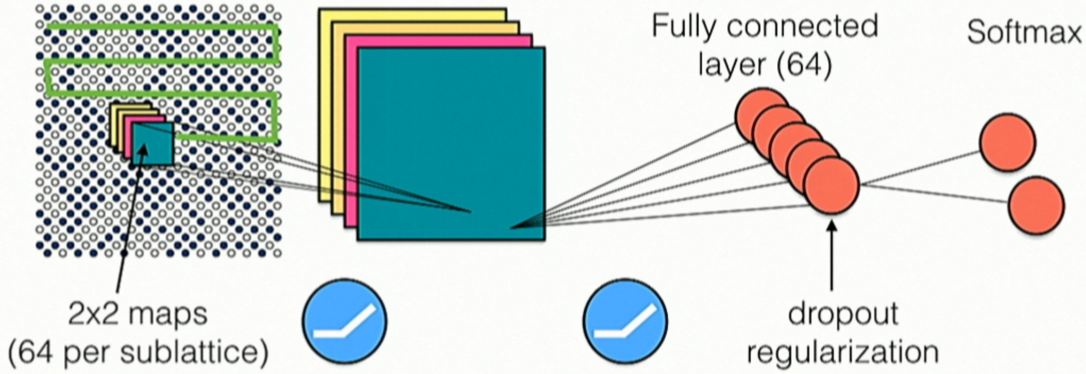


99% accuracy
easy to train

ISING GAUGE THEORY *(Kogut Rev. Mod. Phys. 51, 659 (1979))*

$$H = -J \sum_p \prod_{i \in p} \sigma_i^z$$

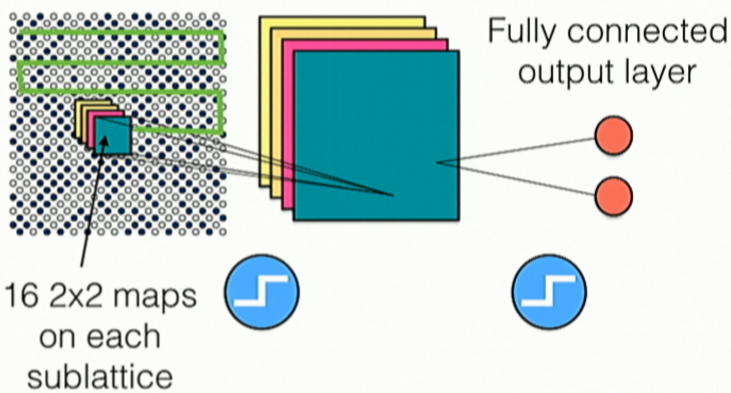




99% accuracy
easy to train

ANALYTICAL UNDERSTANDING: WHAT DOES THE CNN USE TO MAKE PREDICTIONS?

- The convolutional neural net relies on the detection of satisfied local constraints to make accurate predictions of whether a state is drawn at low or infinite temperature.
- Based on this observation we derived the weights of a streamlined convolutional network *analytically* that works perfectly on our test sets.



100% accuracy on test sets

1605.01735

ANALYTICAL MODEL FOR THE ISING GAUGE THEORY

Convolutional layer

f	s=A	s=B	f	s=A	s=B
1	$\begin{pmatrix}1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&1\\0&0\end{pmatrix}$	9	$\begin{pmatrix}1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&-1\\0&0\end{pmatrix}$
2	$\begin{pmatrix}-1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}-1&-1\\0&0\end{pmatrix}$	10	$\begin{pmatrix}1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&1\\0&0\end{pmatrix}$
3	$\begin{pmatrix}1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&-1\\0&0\end{pmatrix}$	11	$\begin{pmatrix}1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&1\\0&0\end{pmatrix}$
4	$\begin{pmatrix}1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}-1&1\\0&0\end{pmatrix}$	12	$\begin{pmatrix}-1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&1\\0&0\end{pmatrix}$
5	$\begin{pmatrix}-1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&1\\0&0\end{pmatrix}$	13	$\begin{pmatrix}-1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}-1&1\\0&0\end{pmatrix}$
6	$\begin{pmatrix}-1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&-1\\0&0\end{pmatrix}$	14	$\begin{pmatrix}-1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&-1\\0&0\end{pmatrix}$
7	$\begin{pmatrix}1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&-1\\0&0\end{pmatrix}$	15	$\begin{pmatrix}-1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&-1\\0&0\end{pmatrix}$
8	$\begin{pmatrix}-1&0\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&1\\0&0\end{pmatrix}$	16	$\begin{pmatrix}-1&0\\-1&0\end{pmatrix}$	$\begin{pmatrix}-1&-1\\0&0\end{pmatrix}$

$$b_c = -(2+\epsilon)\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}$$

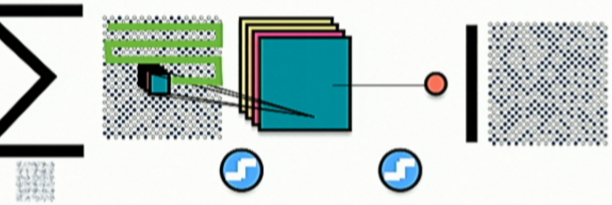
Output layer

$$W_o = \begin{pmatrix} \overbrace{1 \ \ldots \ 1}^{8L^2 \text{ terms}} & \overbrace{-L^2 \ldots - L^2}^{8L^2 \text{ terms}} \\ -1 \ \ldots \ -1 & L^2 \ldots L^2 \end{pmatrix}, \quad \text{and } b_o = \begin{pmatrix}0\\0\end{pmatrix}$$

ONLY QUANTUM SLIDE OF THE TALK

$$H = -J_p \sum_p \prod_{i \in p} \sigma_i^z - J_v \sum_v \prod_{i \in v} \sigma_i^x$$

$| \Psi > = \sum$



$| \quad >$

Neural net represents a ground state of the toric code: equal weight superposition of closed string states

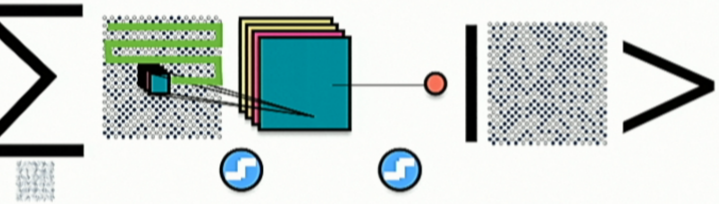
Non trivial ground states can be written as a NN

Optimize using VMC for other more challenging ground state problems (efficient sampling is possible)

ONLY QUANTUM SLIDE OF THE TALK

$$H = -J_p \sum_p \prod_{i \in p} \sigma_i^z - J_v \sum_v \prod_{i \in v} \sigma_i^x$$

$$| \Psi > = \sum$$



Neural net represents a ground state of the toric code: equal weight superposition of closed string states

Non trivial ground states can be written as a NN

Optimize using VMC for other more challenging ground state problems (efficient sampling is possible)

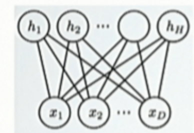
LET'S EXPLOIT THIS MORE

Condensed matter/stat mech

Machine learning

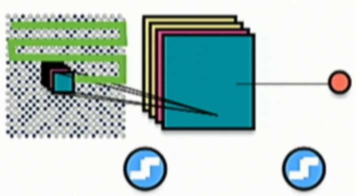
Stoudenmire, Schwab 1605.05775

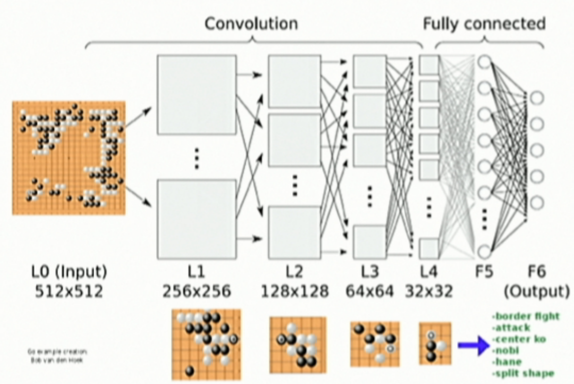
Energy-Based Models



Carleo, Troyer 1606.02318

Torlai, Melko 1606.02718





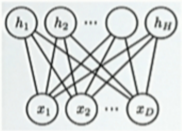
LET'S EXPLOIT THIS MORE

Condensed matter/stat mech

Machine learning

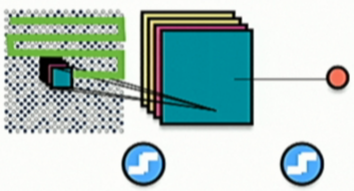
Stoudenmire, Schwab 1605.05775

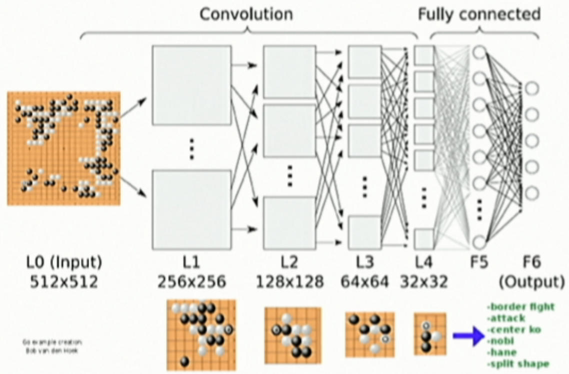
Energy-Based Models



Carleo, Troyer 1606.02318

Torlai, Melko 1606.02718





CONCLUSION

- We encode and discriminate phases and phase transitions, both conventional and topological, using neural network technology.
- We have a solid understanding of what the neural nets do in those cases through controlled analytical models.

OUTLOOK

- We expect a rapid adoption of ML techniques as a tool in condensed matter physics.
- Variational interpretation of CNNs and their optimization for ground state and real-time dynamics.
- Attempt at circumventing the sign problem (Check Peter Broecker's talk this afternoon)
- Establish a more practical connection between RG and deep learning and turn it into a useful tool.
- Train numerically using QMC and use machinery to analyze experimental data: 1. Phase transitions in quantum gas microscopes. 2. Measure temperature from experimental snapshots.