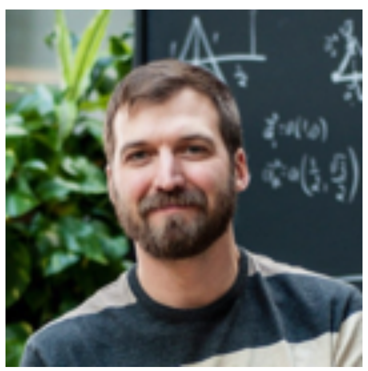
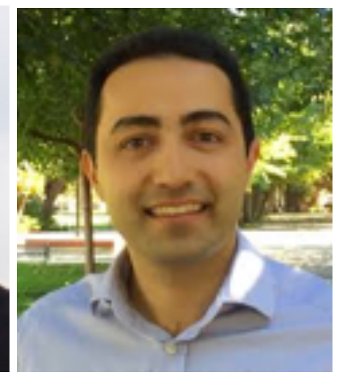
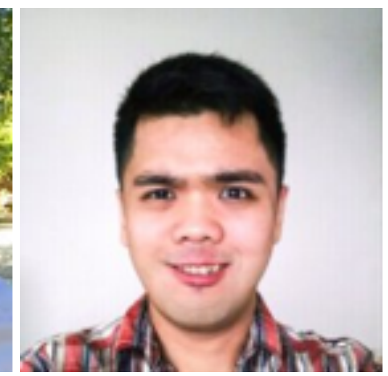
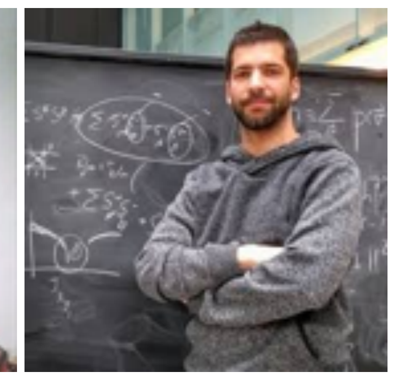
MACHINE LEARNING PHASES OF MATTER

Juan Carrasquilla
D-Wave systems

International workshop on numerical methods and simulations for materials design and strongly correlated quantum matters

COLLABORATORS



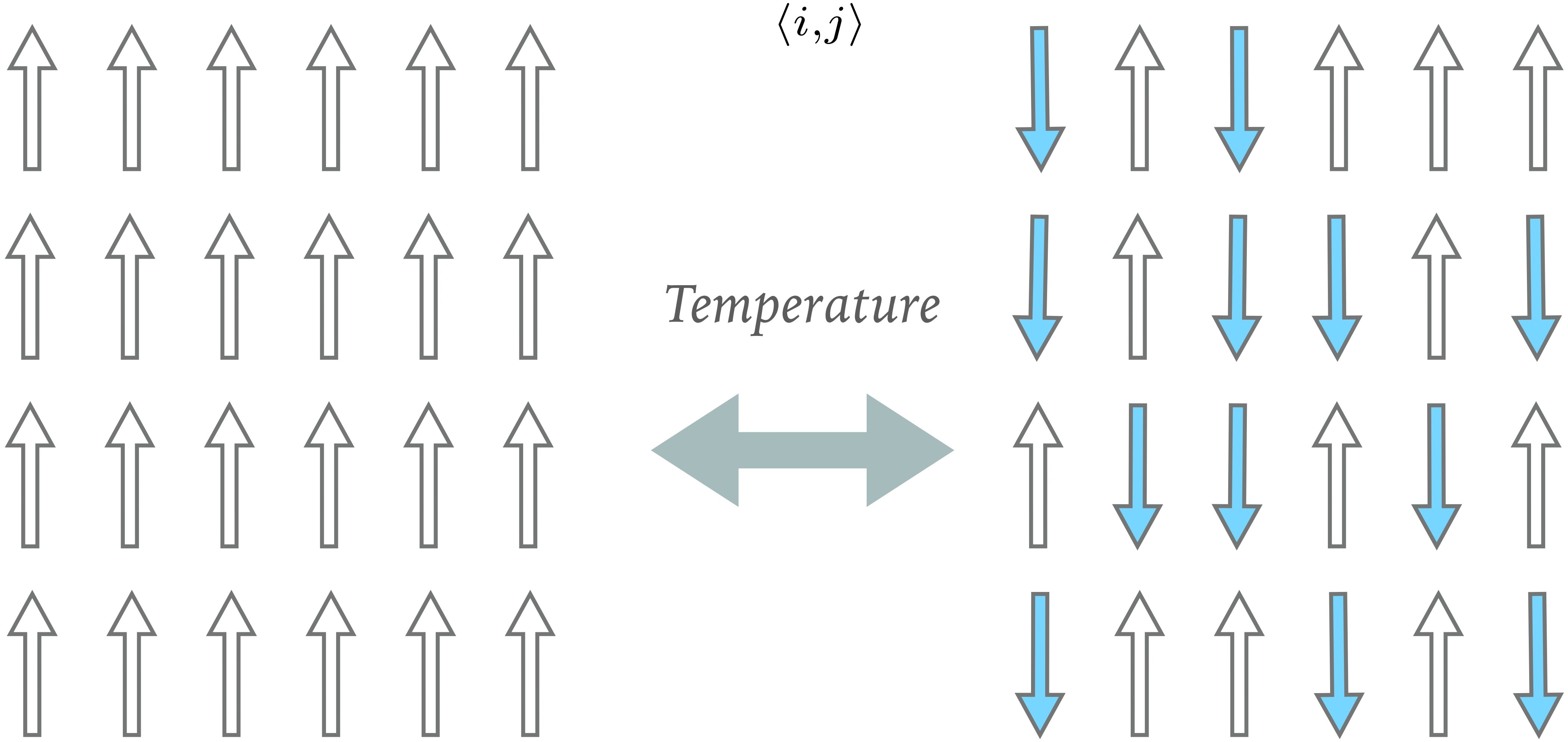
Roger Melko
Simon Trebst
Peter Broecker
Ehsan Khatami
Kelvin Chng
Giacomo Torlai
Giuseppe Carleo

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ising ferromagnet in two dimensions

$$E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$



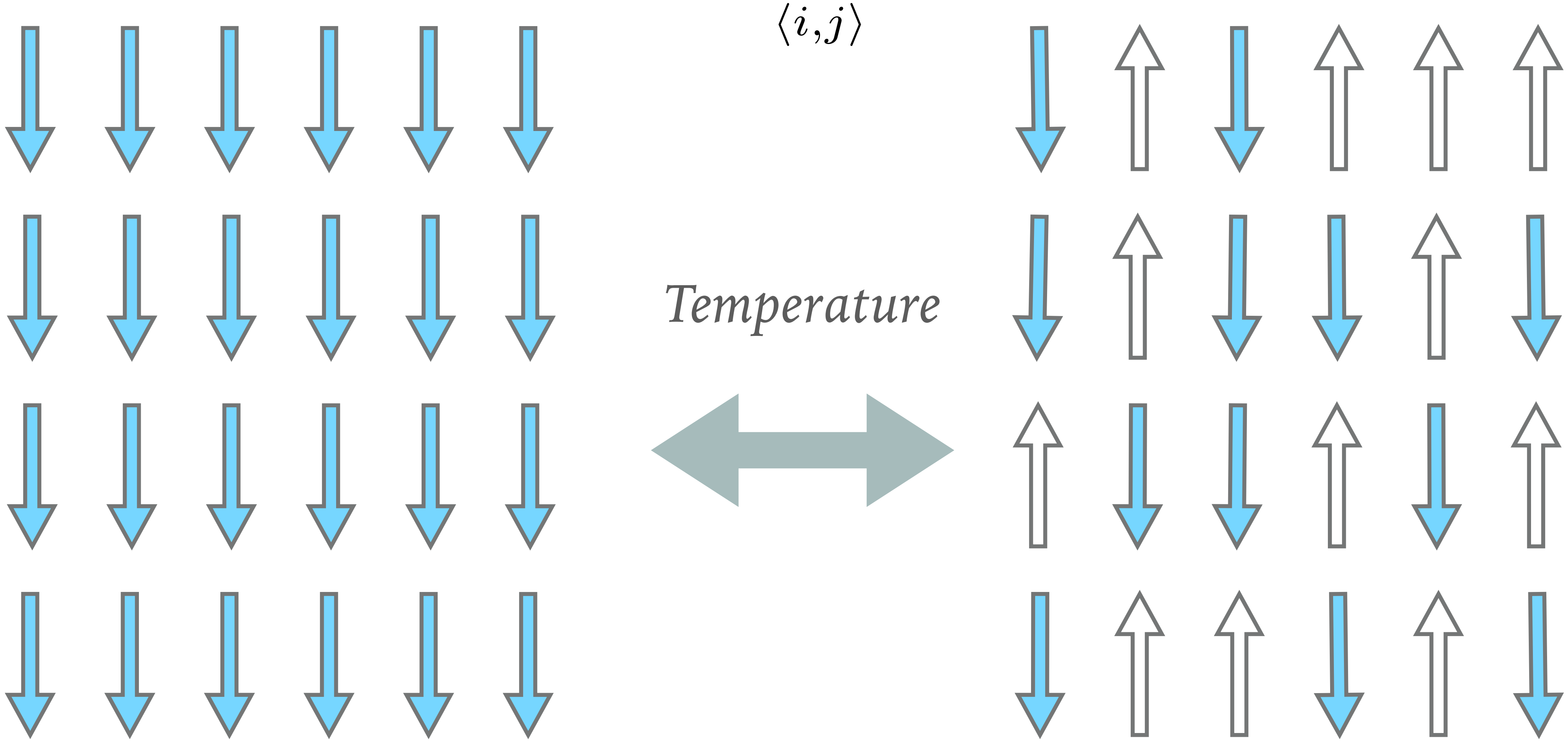
Ferromagnet *Paramagnet*

Lars Onsager Phys. Rev. 65, 117

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ising ferromagnet in two dimensions

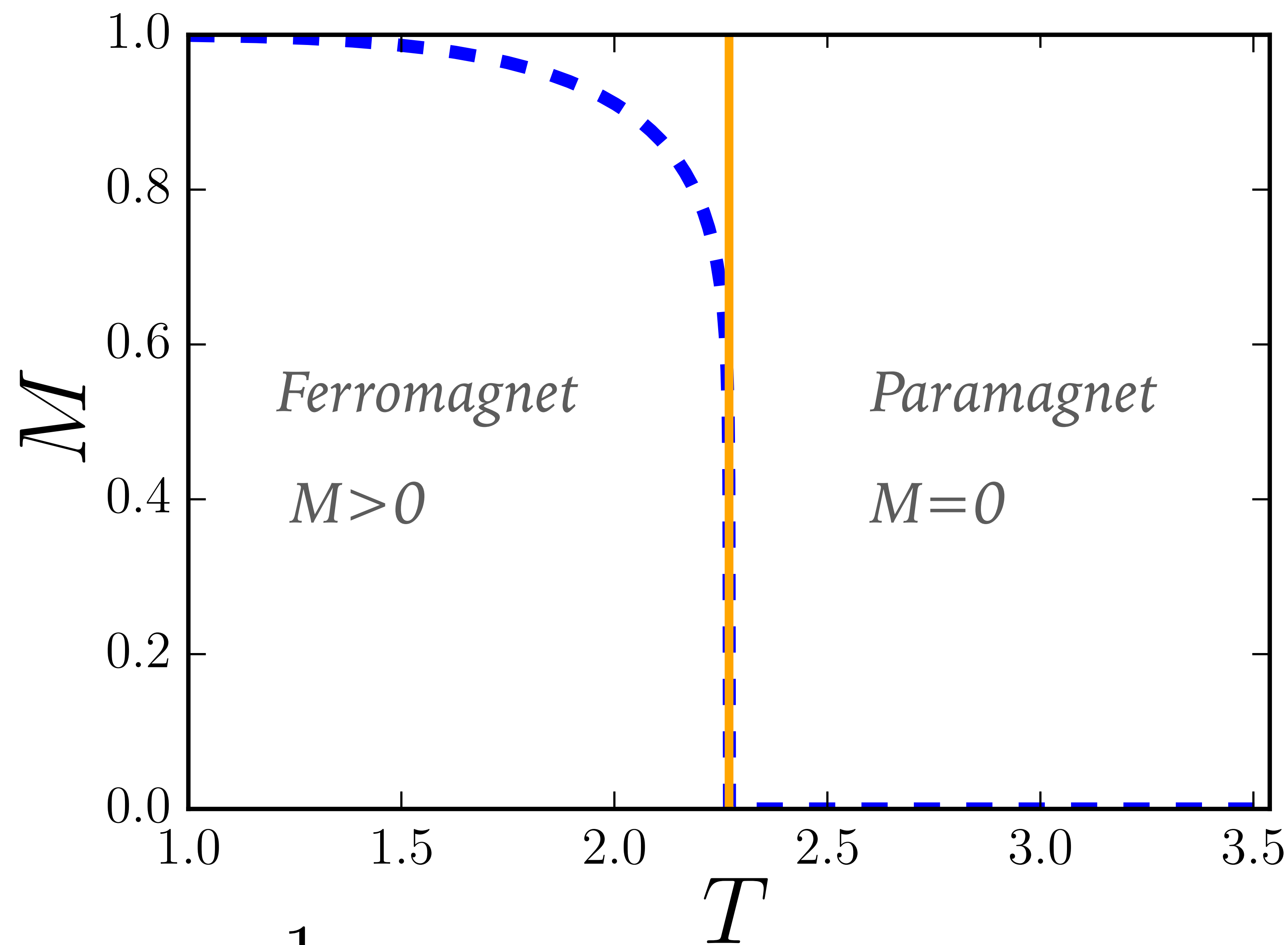
$$E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$



Ferromagnet *Paramagnet*

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ferromagnetic transition: order parameter



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

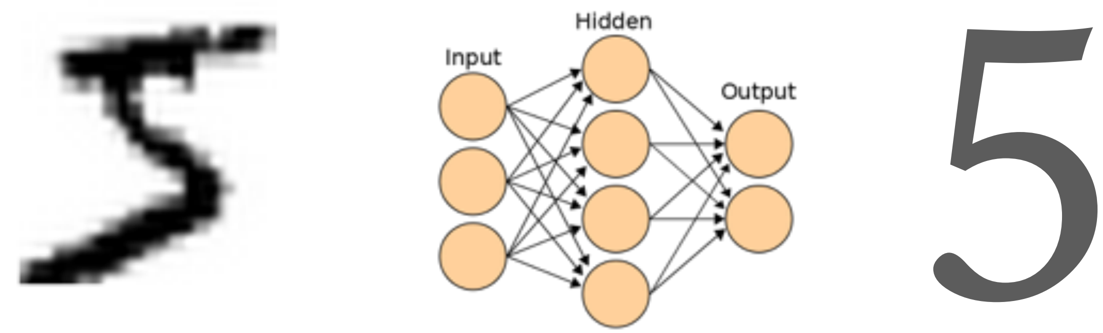
$$M = \frac{1}{N}\sum_i \langle \sigma_i \rangle, \qquad \sigma_i = \pm 1$$

Lars Onsager Phys. Rev. 65, 117

LEARNING PHASES OF MATTER: INSPIRATION FROM THE FLUCTUATIONS IN HANDWRITTEN DIGITS AND SUPERVISED LEARNING

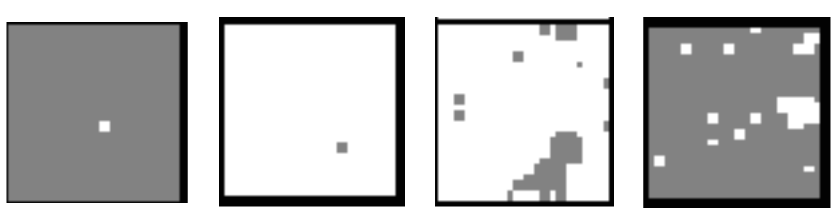
INSPIRATION: FLUCTUATIONS HANDWRITTEN DIGITS (MNIST)

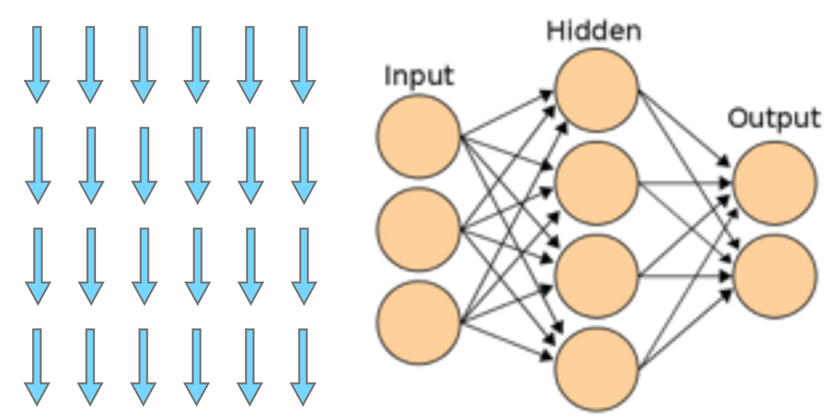
 $=5+$*Fluctuations*



ML community has developed powerful supervised learning algorithms

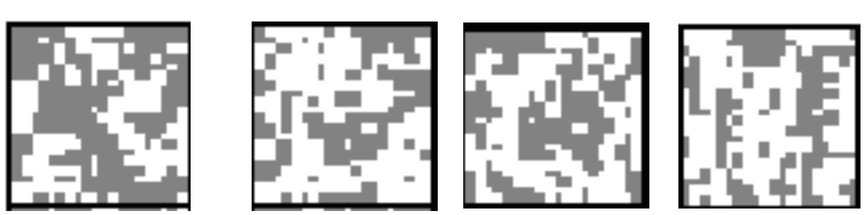
FM phase





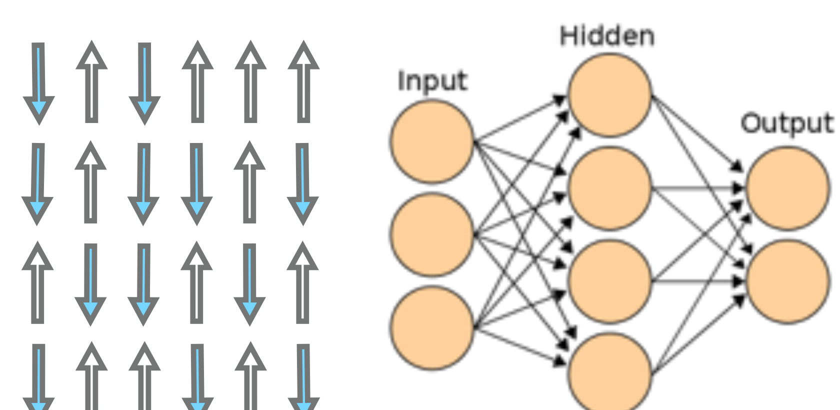
FM (0)

High T phase



gray=spin up

white=spin down



PM (1)

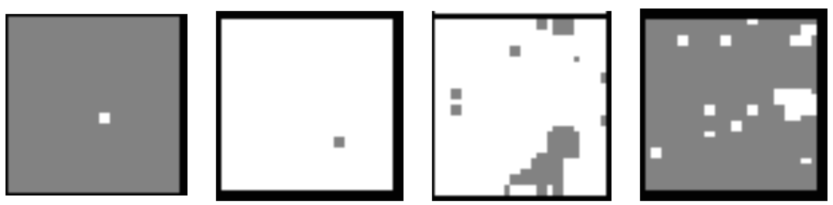
1605.01735

COLLECTING THE TRAINING/TESTING DATA:

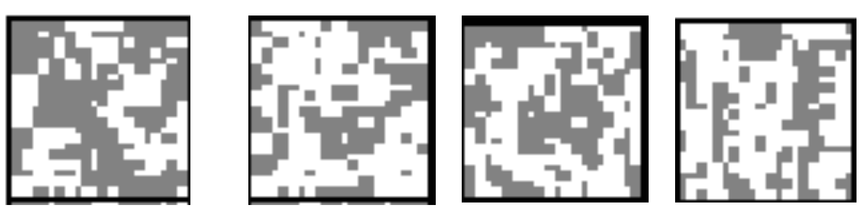
MC SAMPLING ISING MODEL AND LABELS

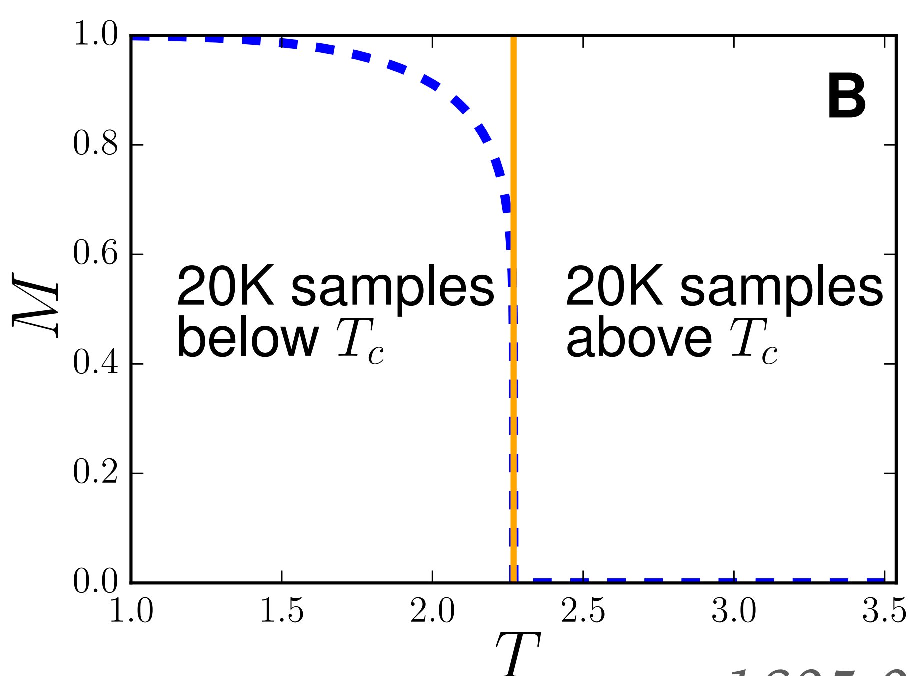
COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase



2D Ising model in the disordered phase

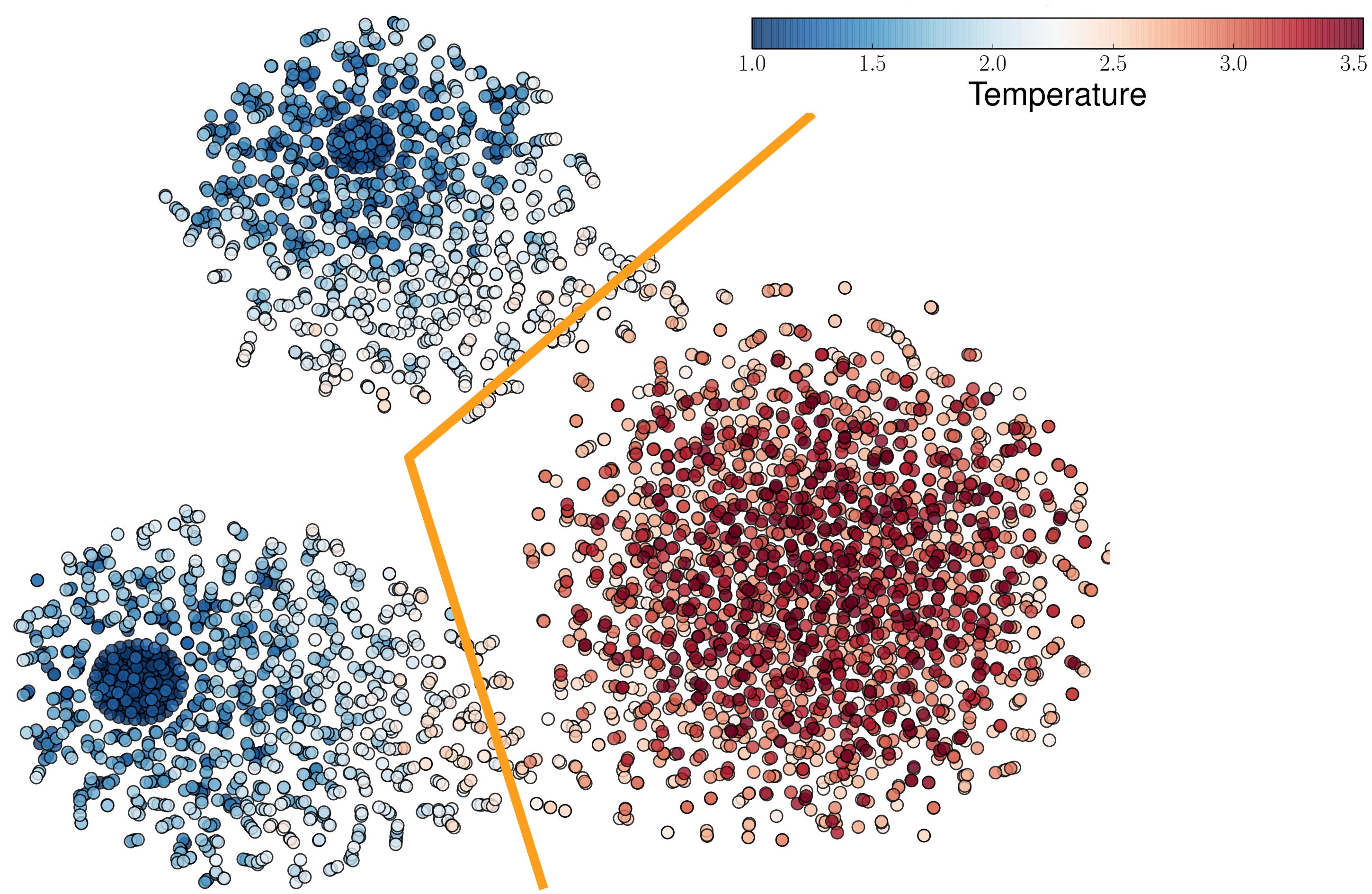




1605.01735

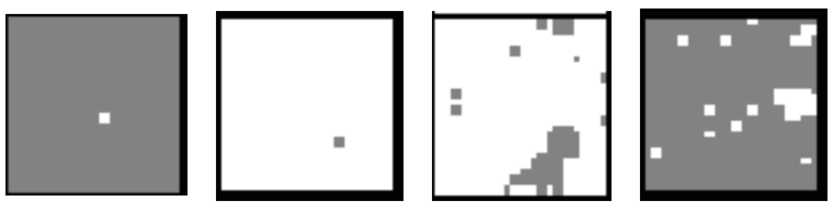
Training/testing data is drawn from the Boltzmann distribution

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

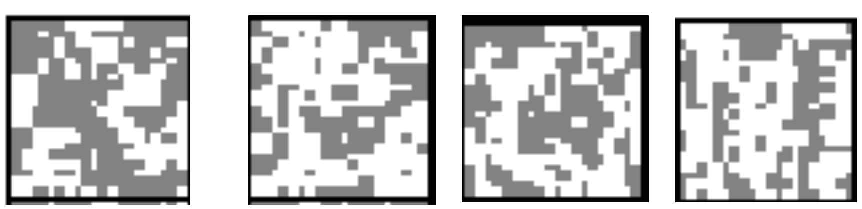


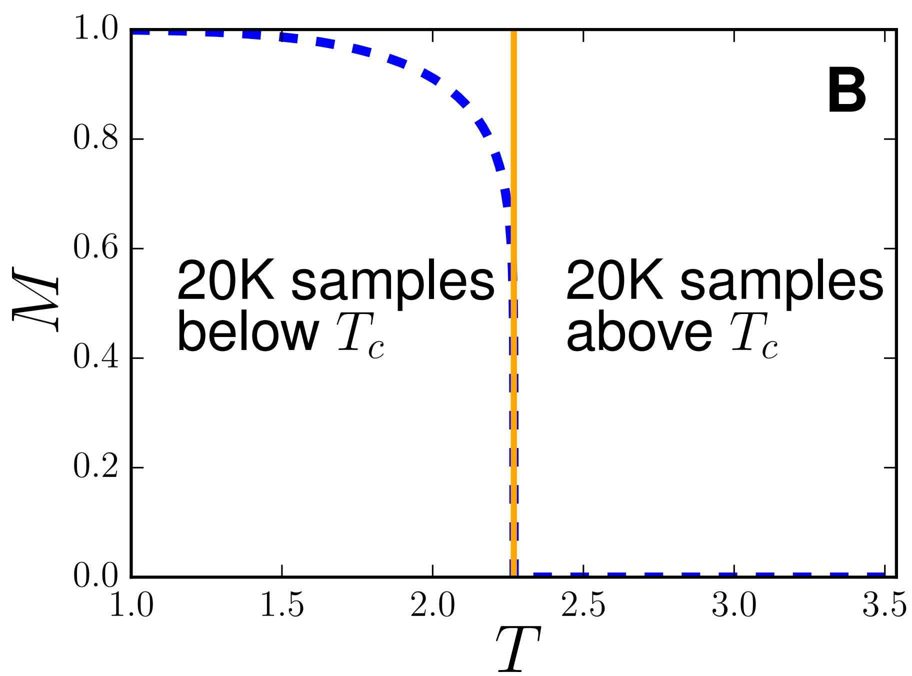
COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase



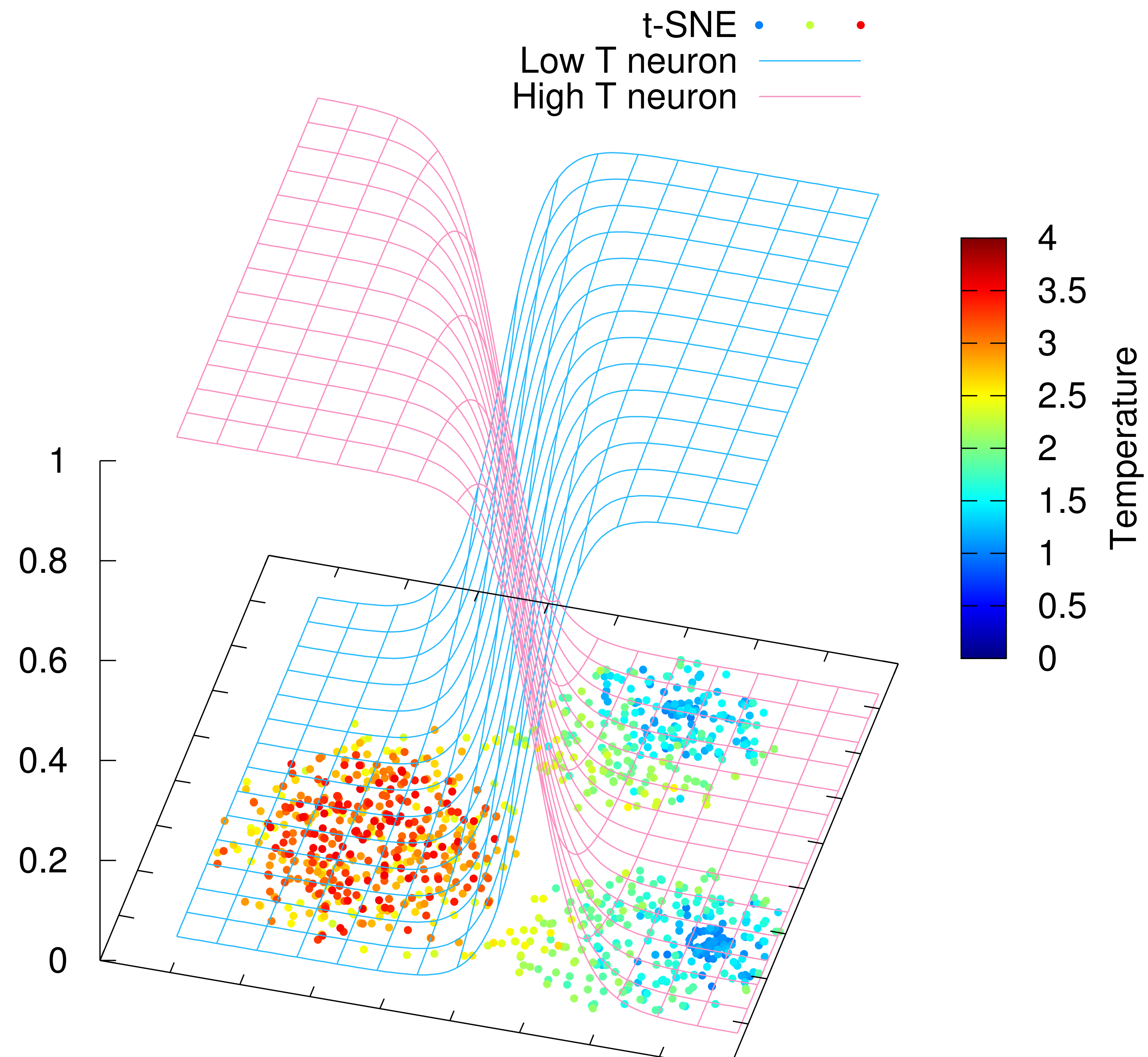
2D Ising model in the disordered phase





Successful training amounts to finding functions

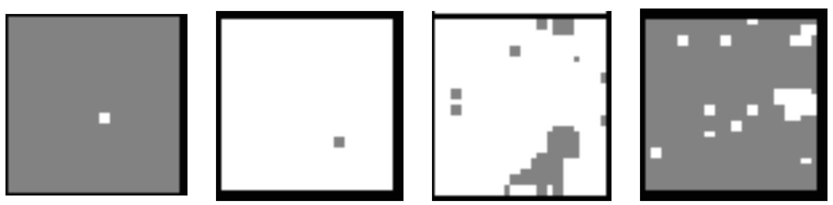
$$F_{\text{High }T}(\sigma_1, \ldots, \sigma_N) \qquad F_{\text{Low }T}(\sigma_1, \ldots, \sigma_N)$$



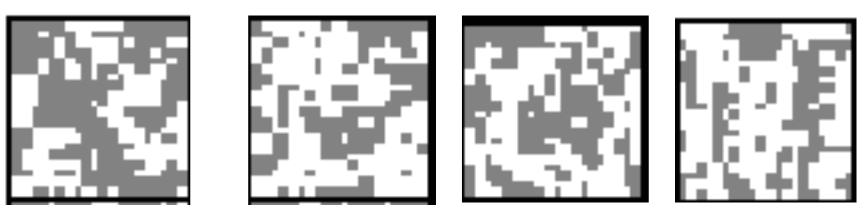
RESULTS: ISING MODEL ON SQUARE AND TRIANGULAR LATTICES

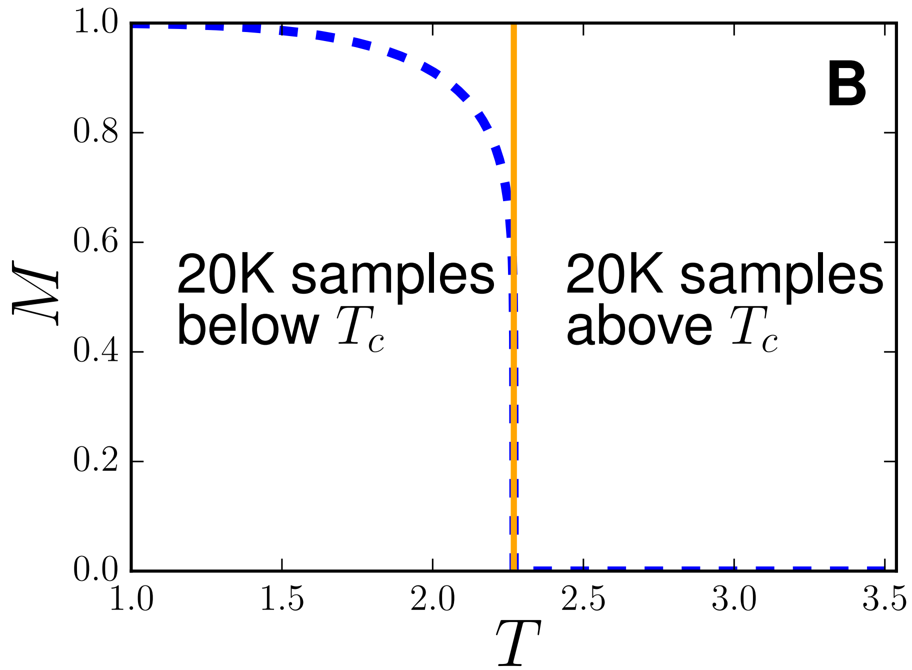
RESULTS: SQUARE LATTICE ISING MODEL (TEST SETS)

2D Ising model in the ordered phase

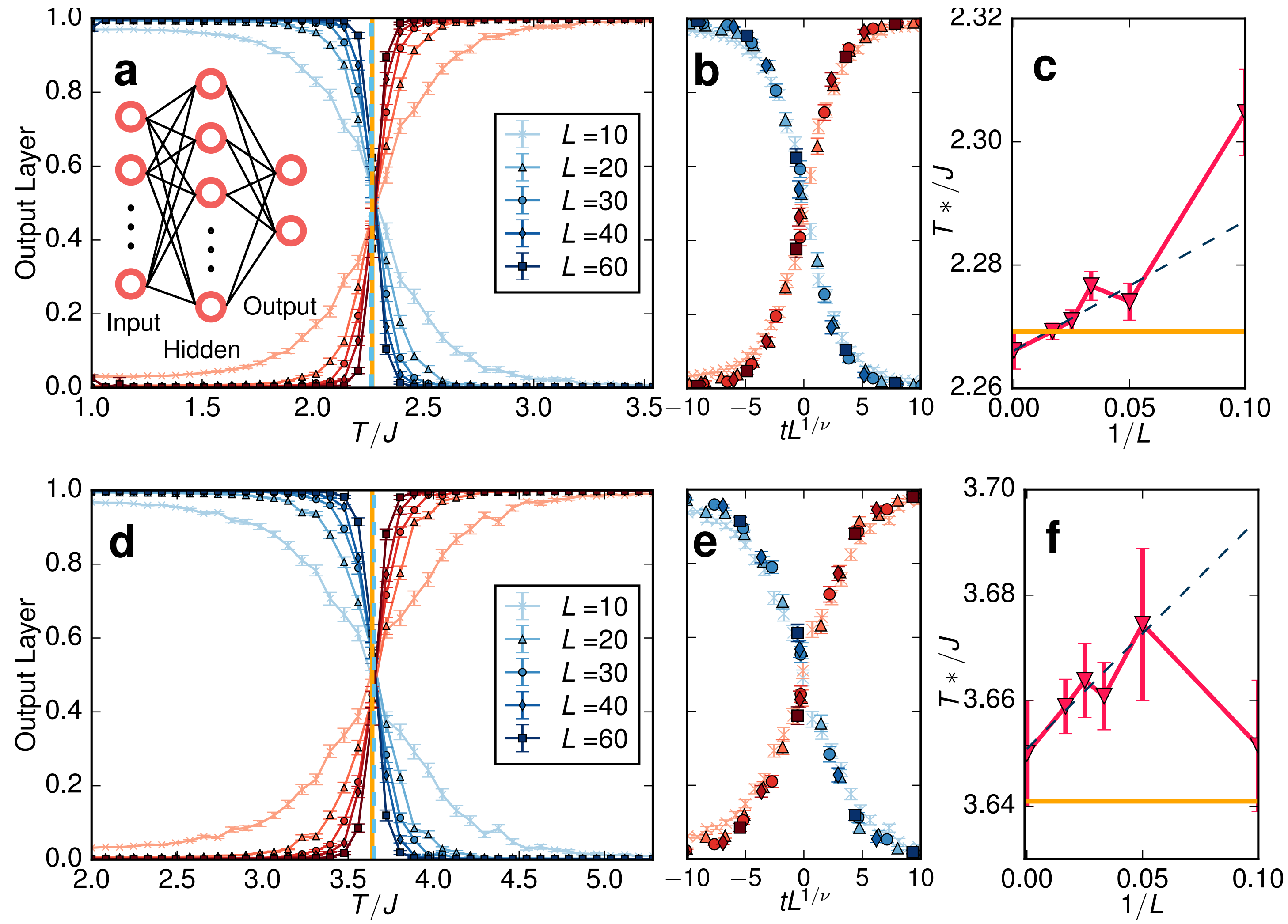


2D Ising model in the disordered phase





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



Critical

1605.01735, Nature Physics

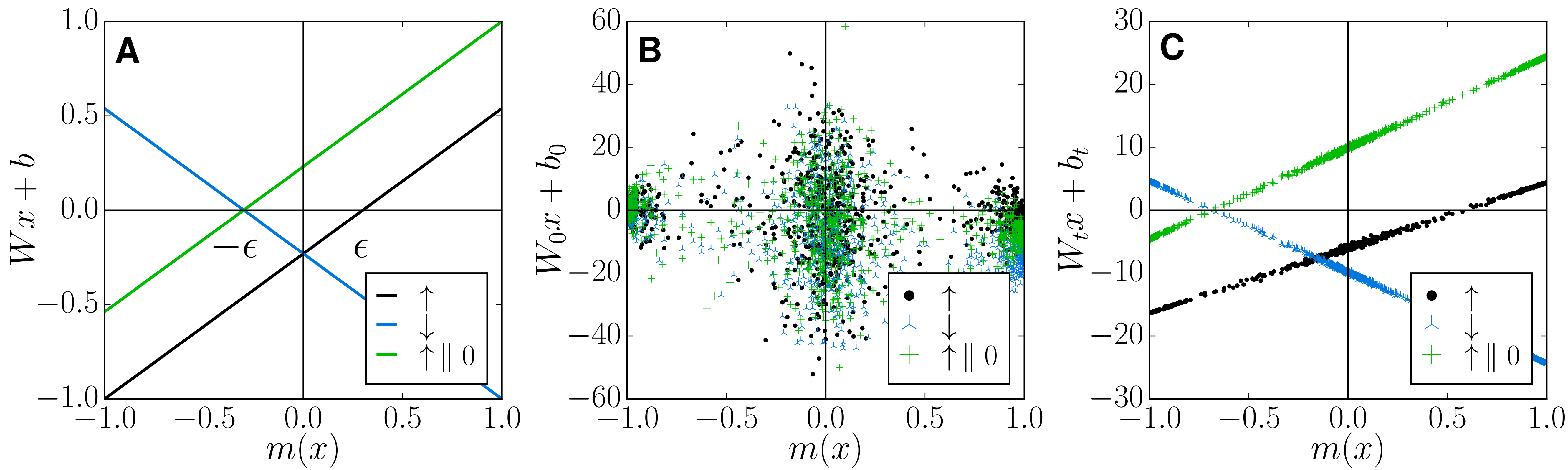
Tc of Triangular within <1%

NN knows about criticality

$\nu \approx 1$

ANALYTICAL UNDERSTANDING

Investigating the argument of the hidden layer during the training



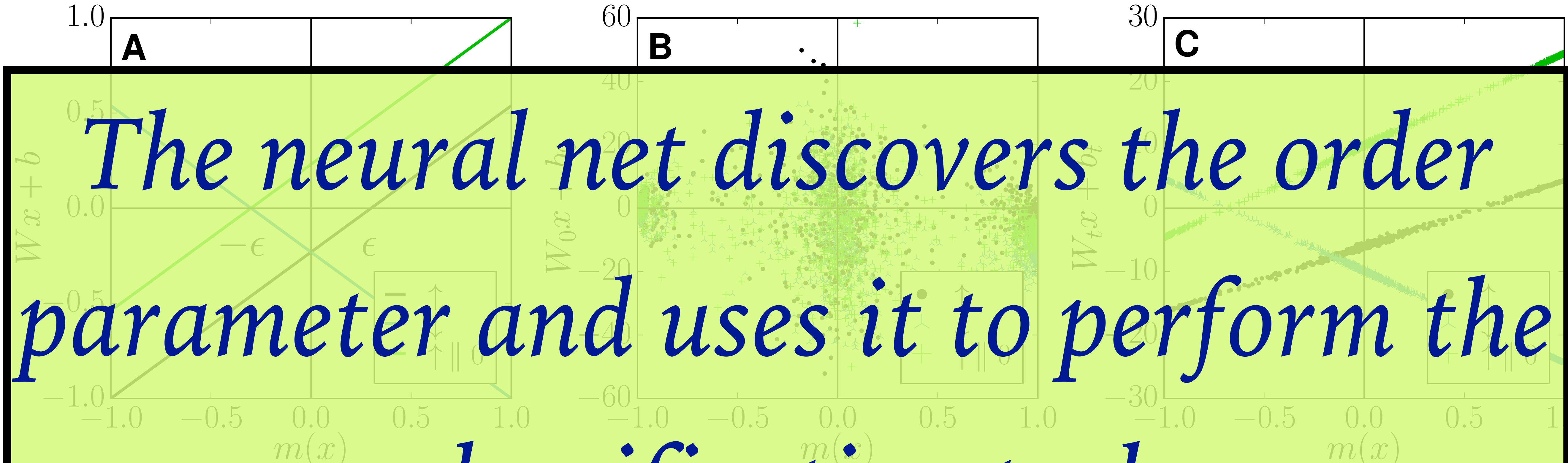
$$W = \frac{1}{N(1+\epsilon)} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \text{ and } b = \frac{\epsilon}{(1+\epsilon)} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \qquad Wx + b = \frac{1}{(1+\epsilon)} \begin{pmatrix} m(x) - \epsilon \\ -m(x) -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

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

1605.01735, Nature Physics

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

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

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

SQUARE ICE AND ISING GAUGE THEORY

PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0

- Topological phases of matter. Examples: Fractional quantum hall effect, Ising gauge theory. 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 $Dy_2Ti_2O_7$)

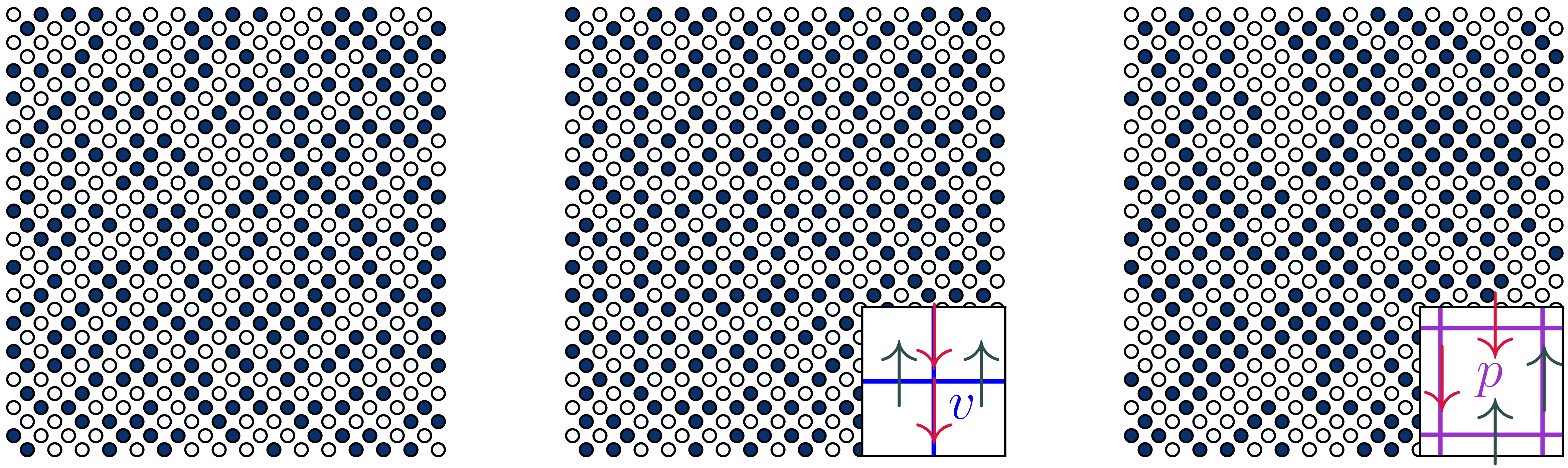
ST Bramwell, MJP Gingras Science 294 (5546), 1495-1501

PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0,∞

A High temperature state

B Ising square ice ground state

C Ising lattice gauge theory



Ising square ice

$$H = J \sum_v Q_v^2$$

$$Q_v = \sum_{i \in v} \sigma_i^z$$

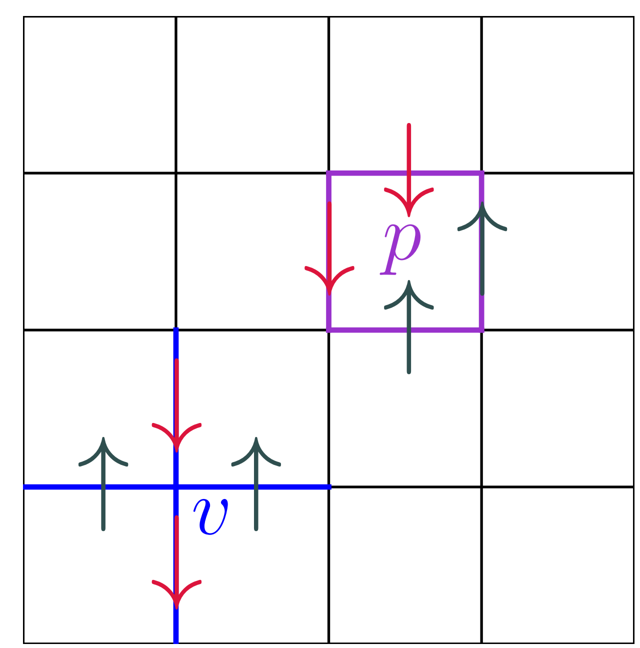
Loop update + spin flip MC

Ising gauge theory

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

(Kogut Rev. Mod. Phys. 51, 659 (1979))

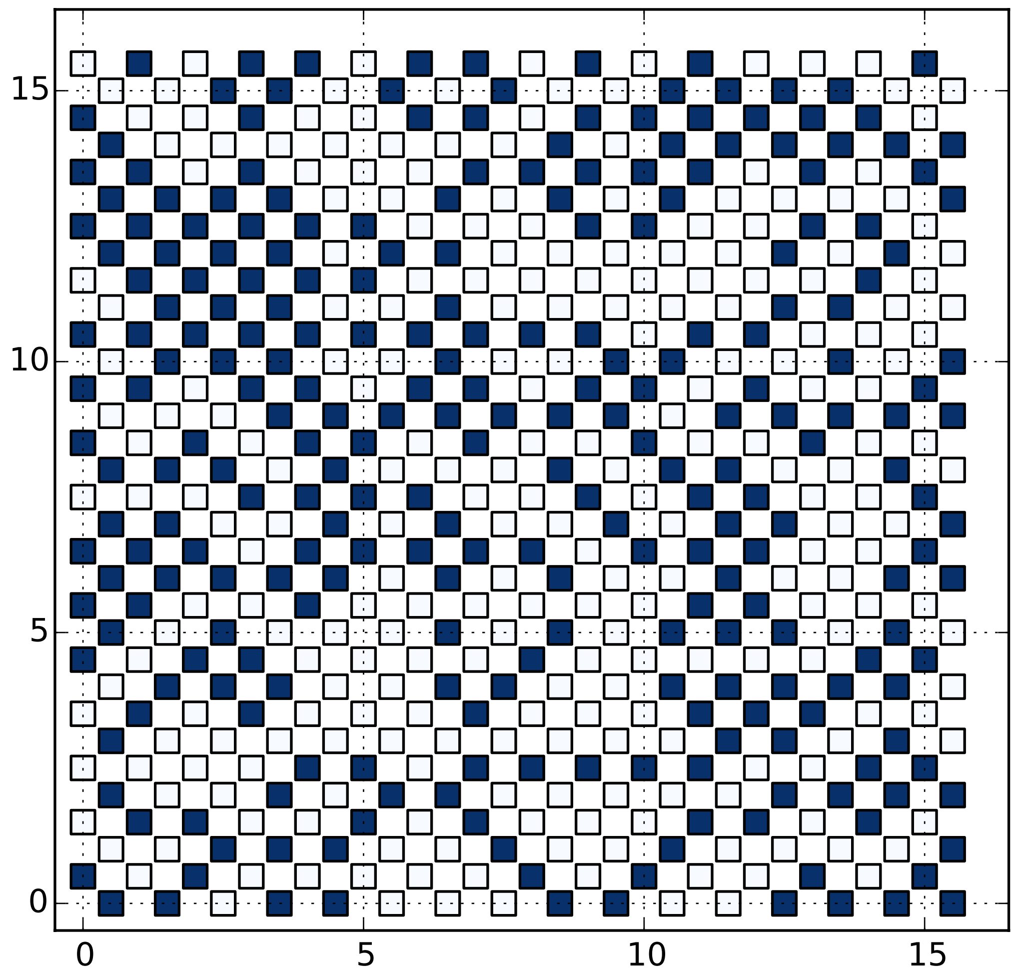
Gauge update + spin flip MC



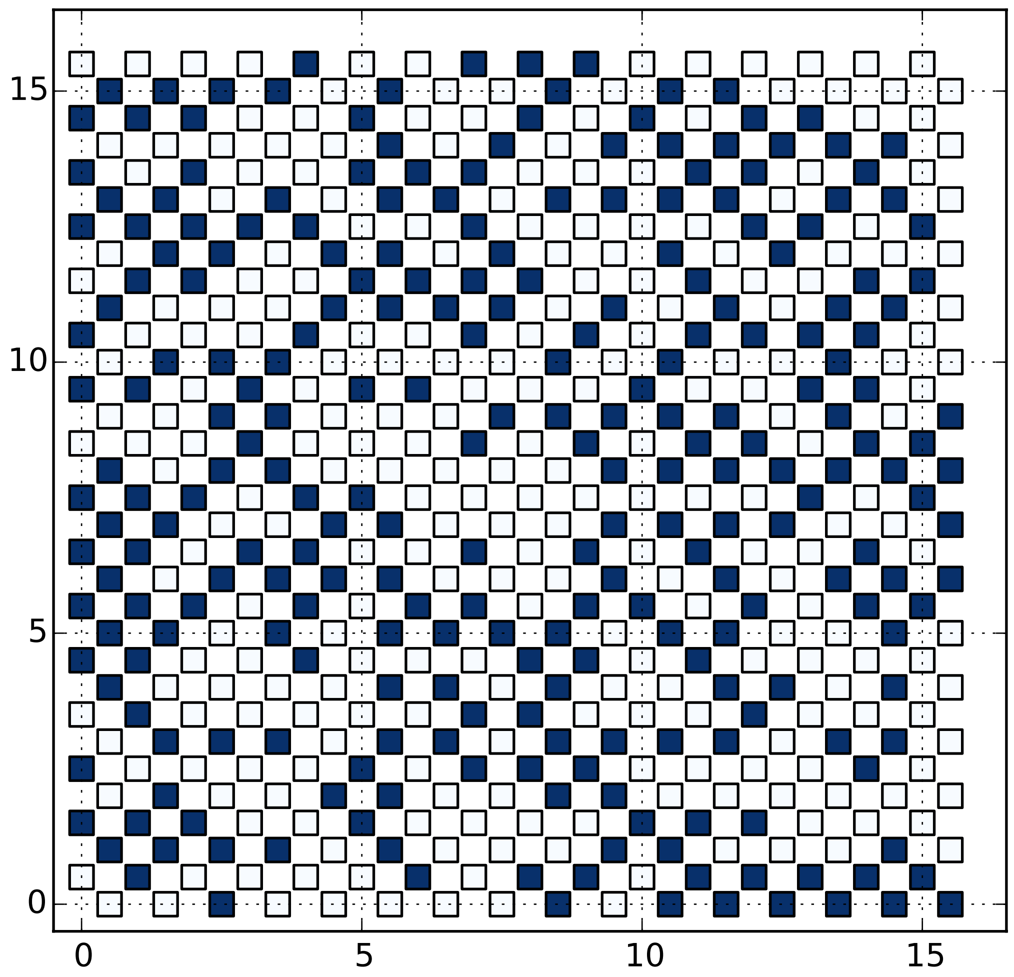
PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0

- 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?

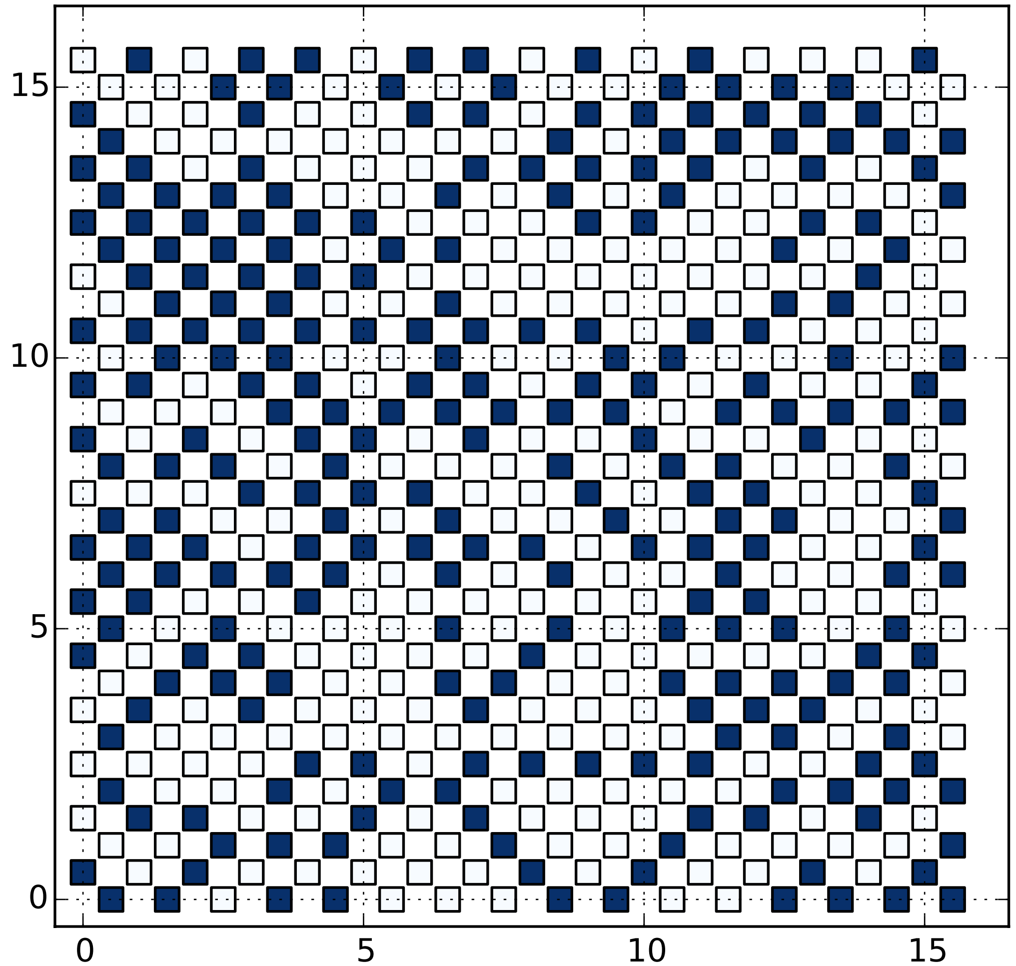
For two configurations



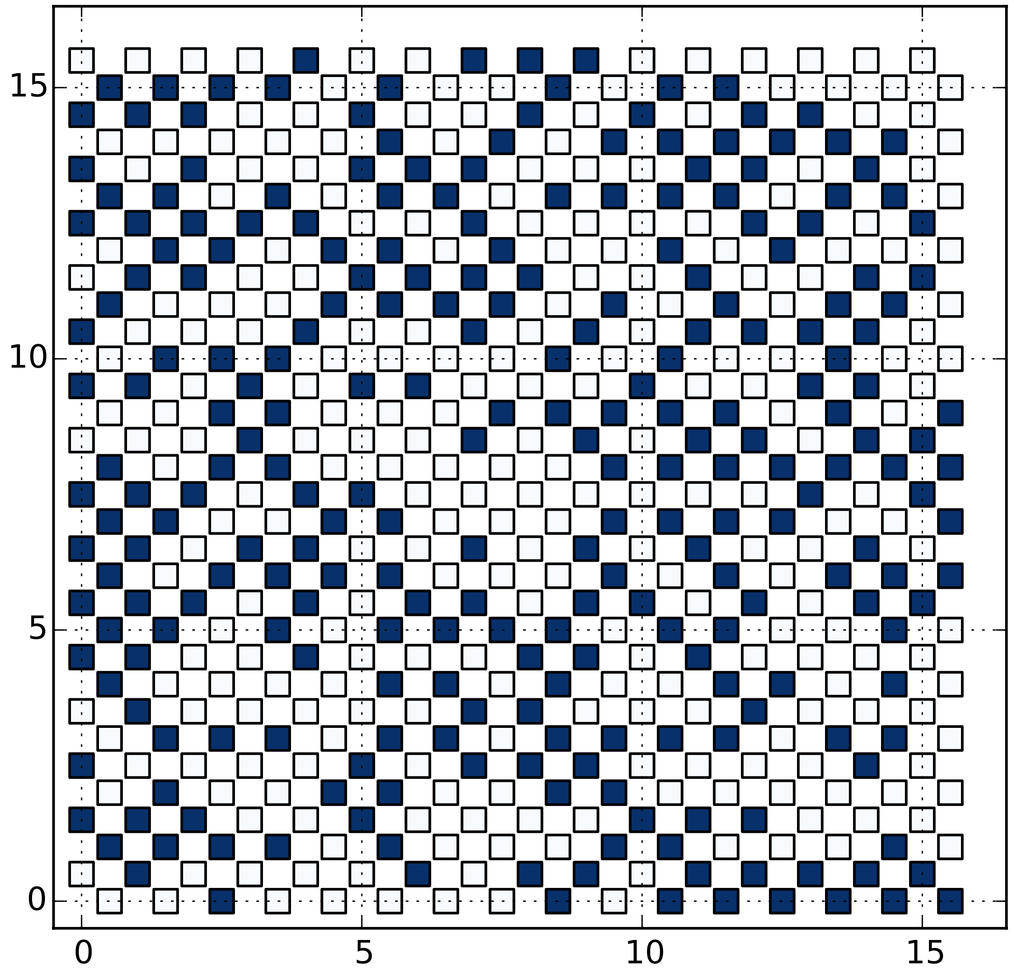
?



For two configurations



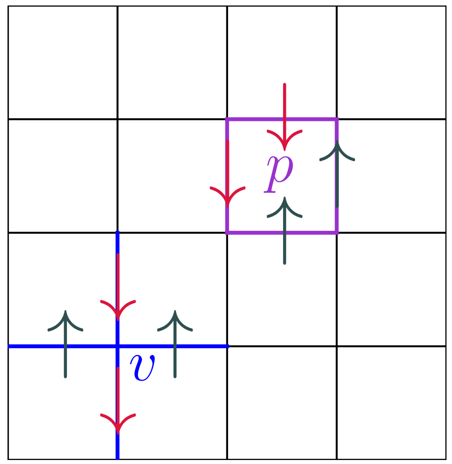
?

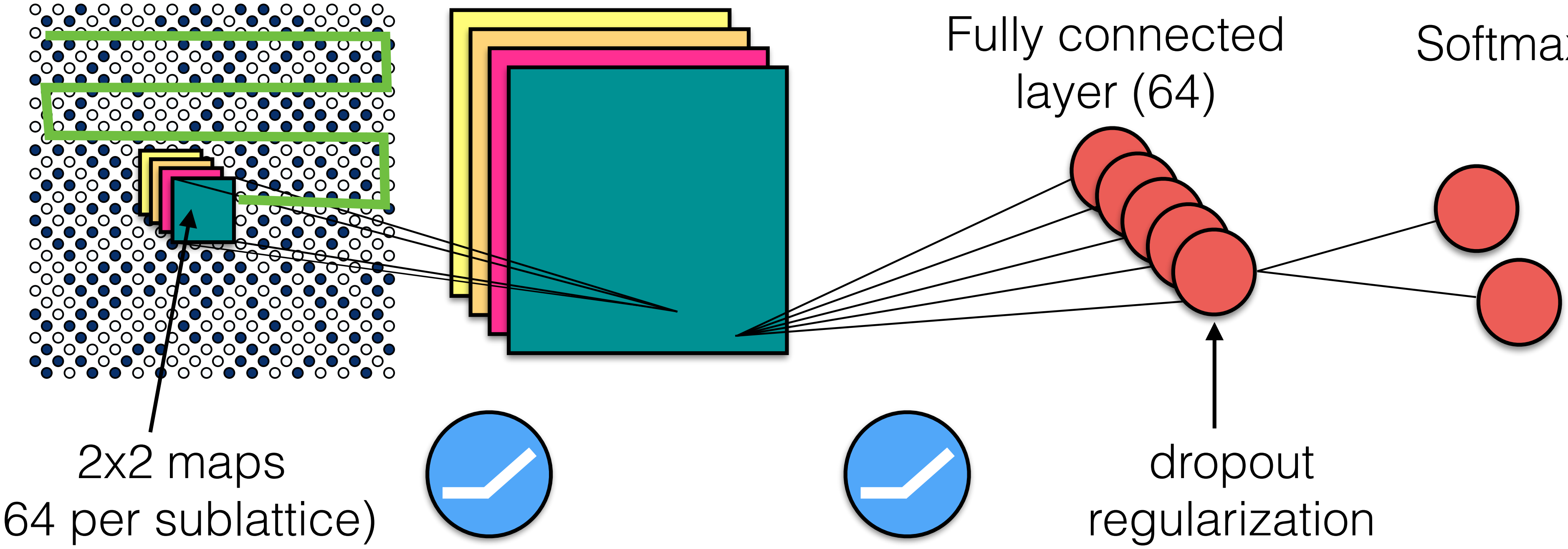


Ground state hahaha

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

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

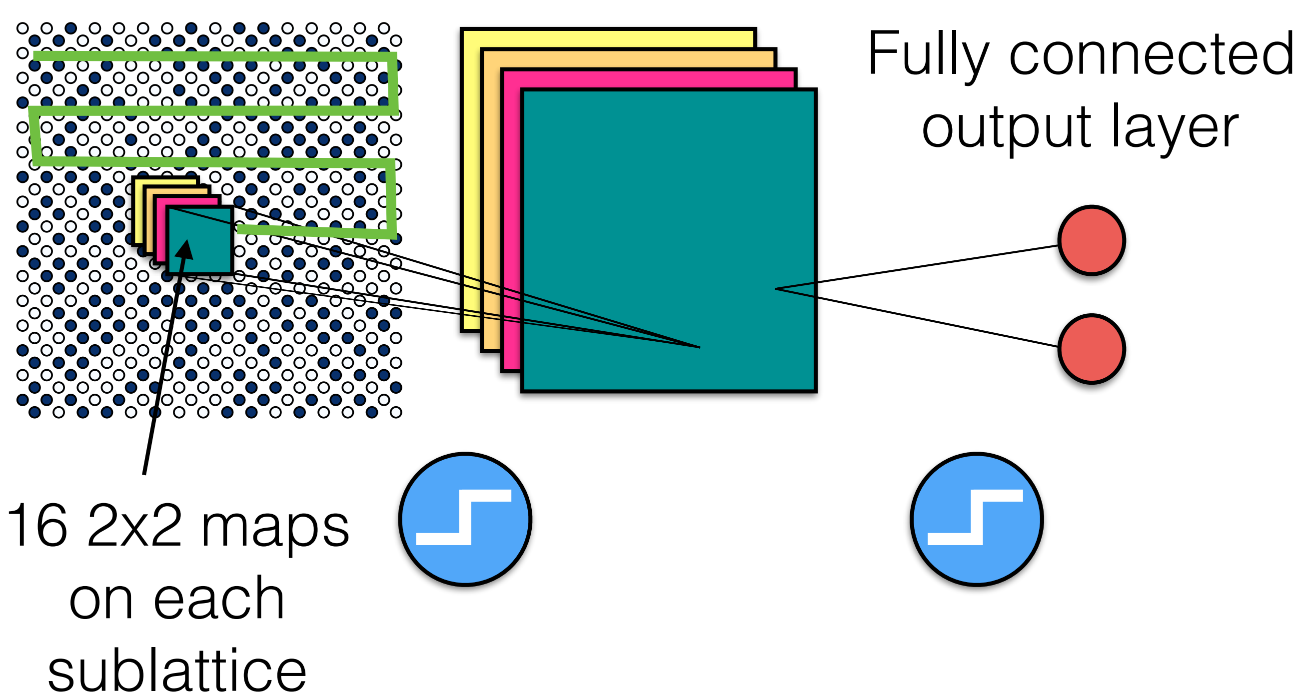




99% accuracy
easy to train

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, Nature Physics

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	$\begin{pmatrix}-1 & 0\\ -1 & 0\end{pmatrix}$	$\begin{pmatrix}-1 & -1\\ 0 & 0\end{pmatrix}$

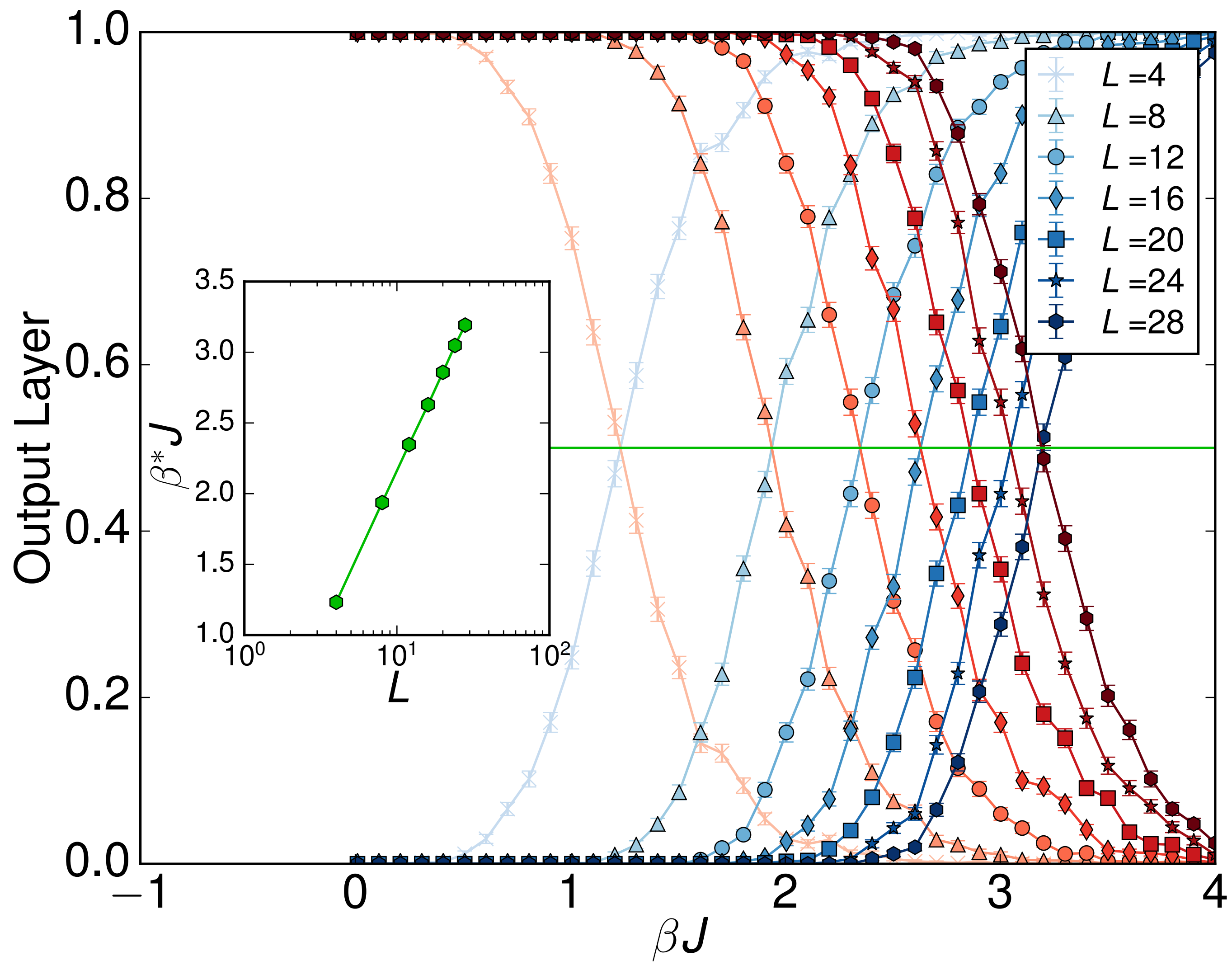
$$b_c = -(2+\epsilon)\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$$

fully-connected layer

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

Basically the Conv. layer encodes the Hamiltonian

LOGARITHMIC CROSSOVER OF THE ISING GAUGE THEORY



NN is confused=crossover temperature

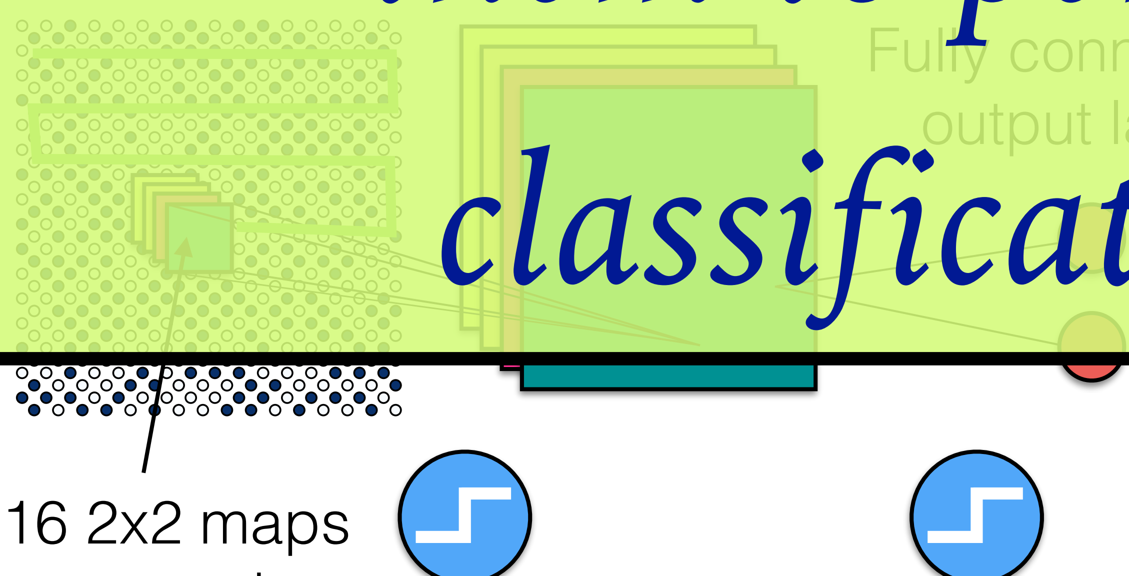
slowly cross over to the high-temperature phase. The cross-over temperature T^* happens as the number of thermally excited defects $\sim N\exp(-2J\beta)$ is of the order of one, implying $T^*/J \sim 1/\ln\sqrt{N}$.[23] As the presence of local defects is the mechanism through which the CNN decides

Only one defect on average distorts the Wilson loops. Same crossover observed in the topological entanglement entropy.

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.

The neural net discovers the local constraints induced by H and use them to perform the classification task



100% accuracy on test sets

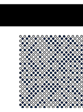
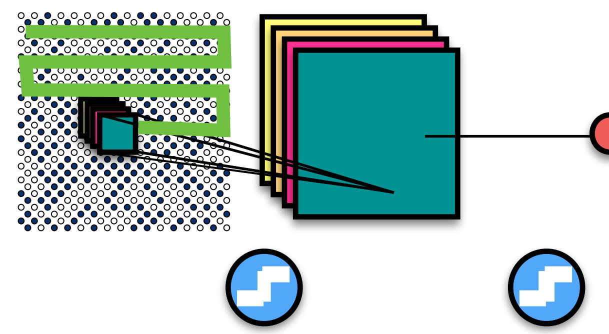
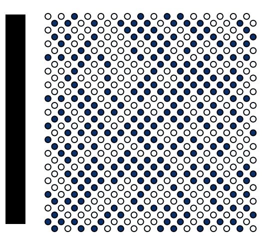
**Encodes a ground state of the toric code*

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

QUANTUM SYSTEMS

QUANTUM: GROUND STATE OF THE TORIC CODE

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

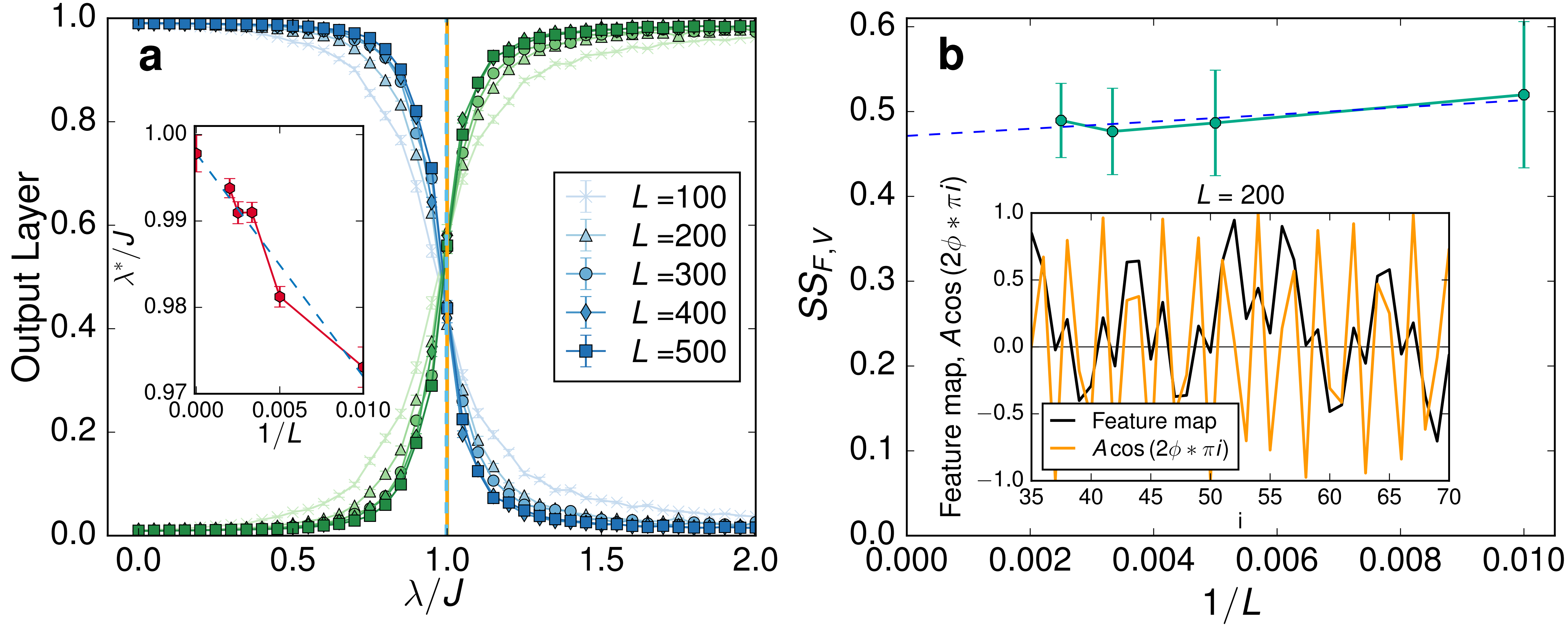
$| \Psi > = \sum$   $|$  $>$

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)

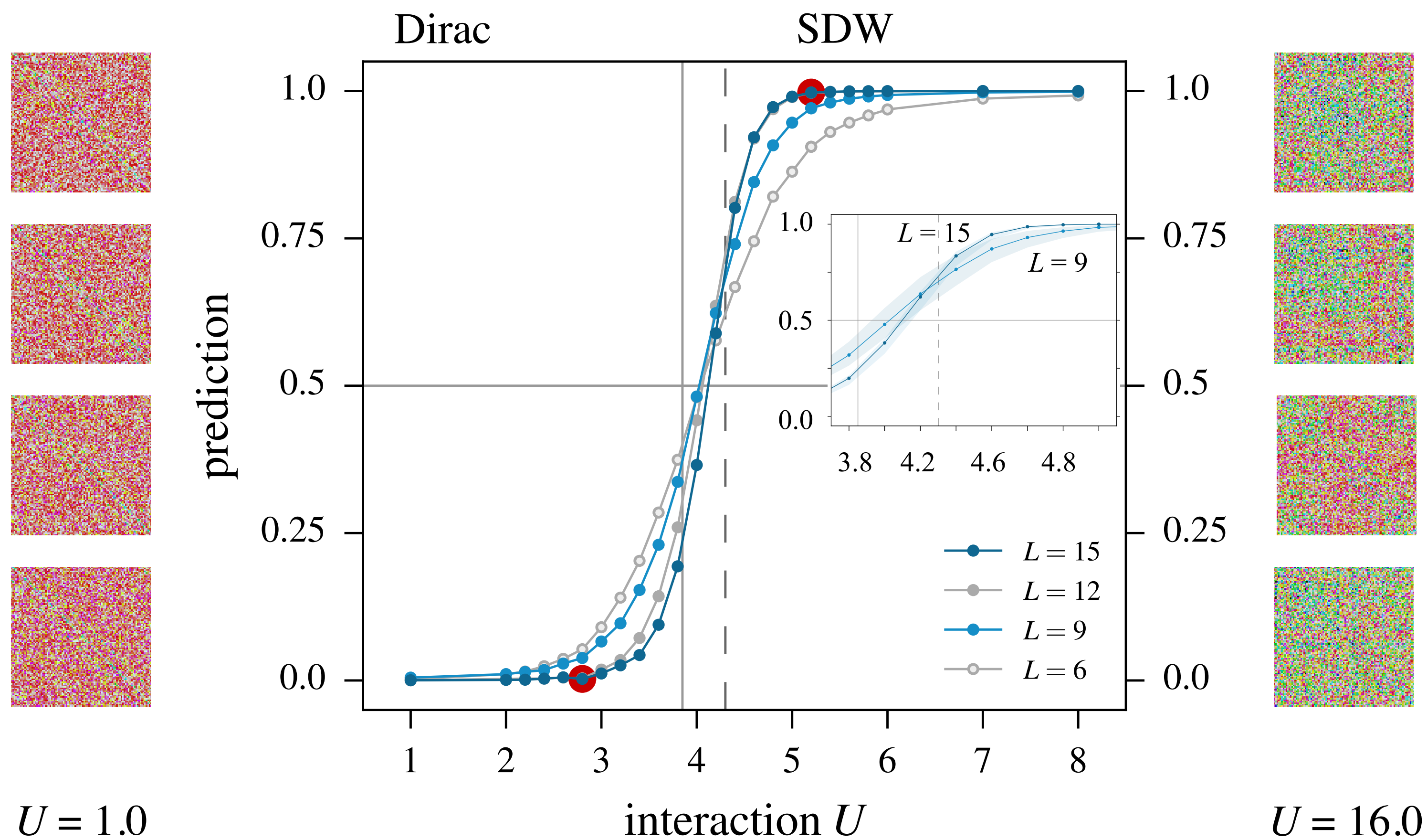
AUBRY ANDRE MODEL

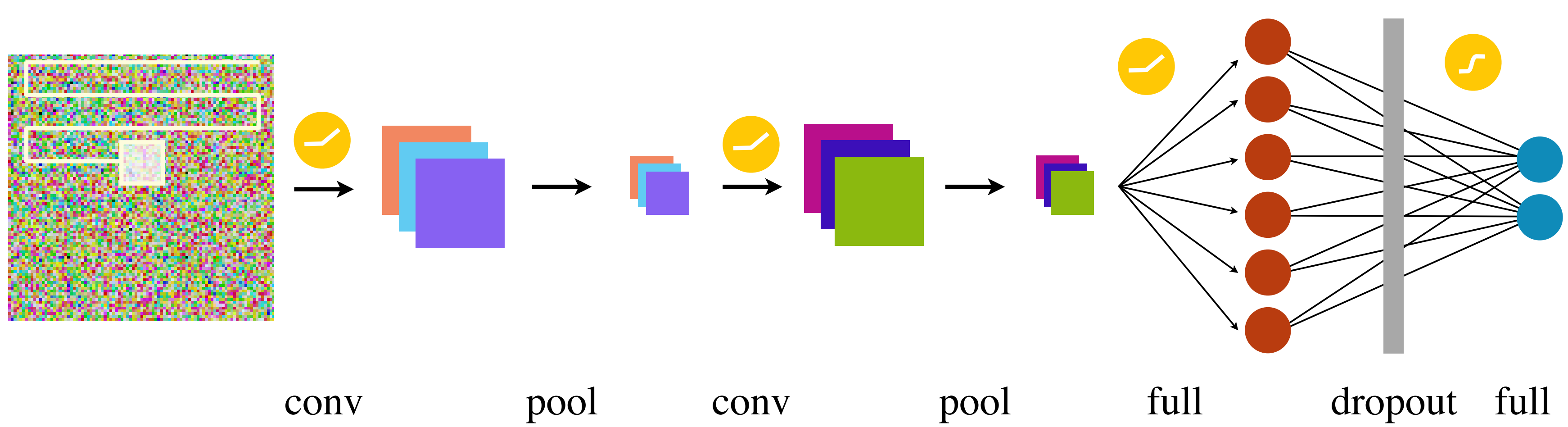


NN understands localization in a naive and fragile way: it basically learns the disordered potential

1605.01735, Nature Physics

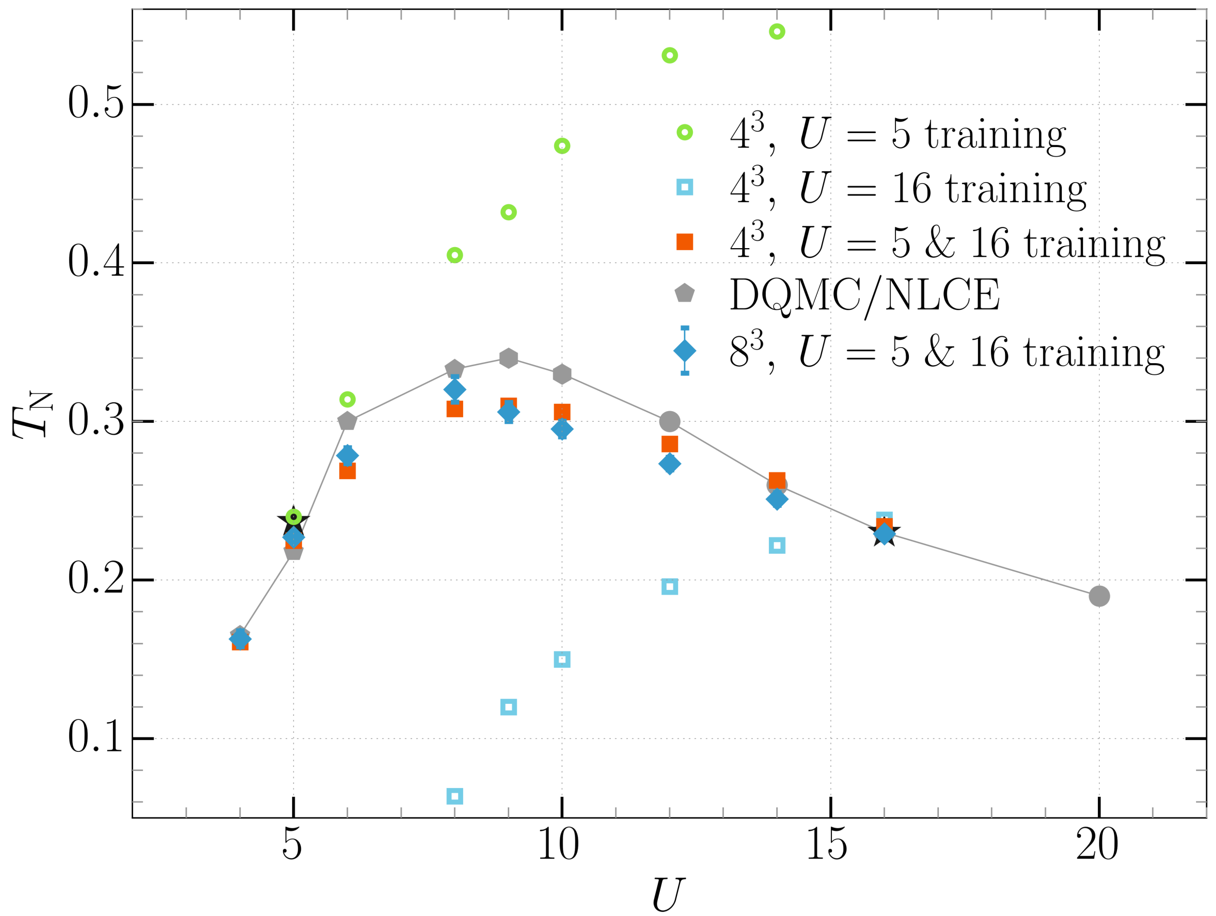
DIRAC FERMIONS ON THE HONEYCOMB LATTICE: MOTT TRANSITION

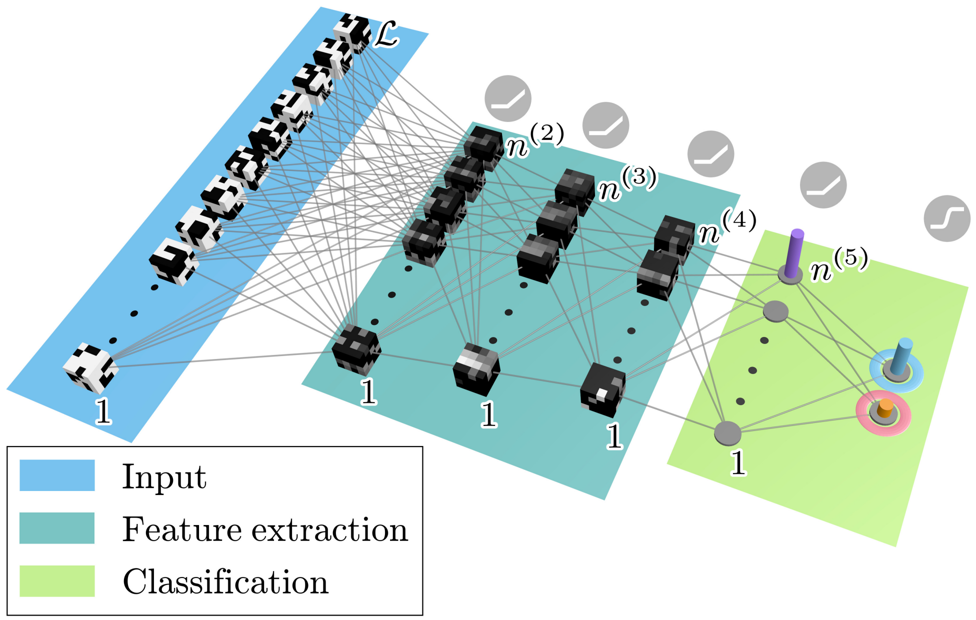




1608.07848

3D HUBBARD MODEL AT HALF FILLING: NEEL TRANSITION



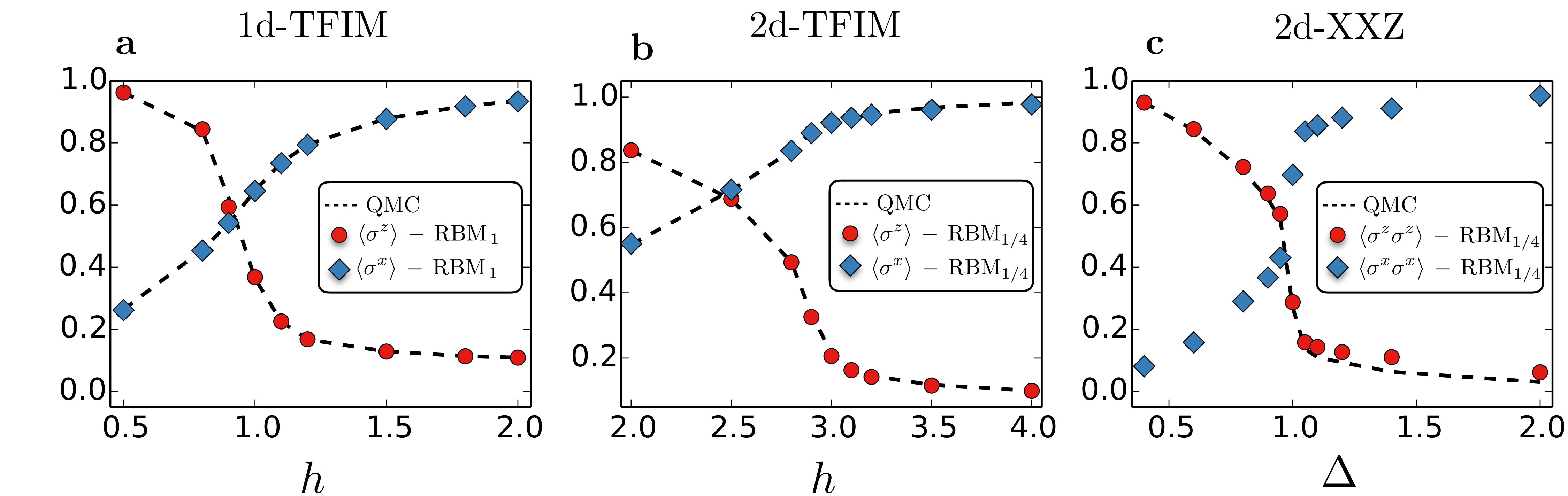


Patterns learned by the NN: AF correlations in the d+1 simulation cell. Much like the Ising examples

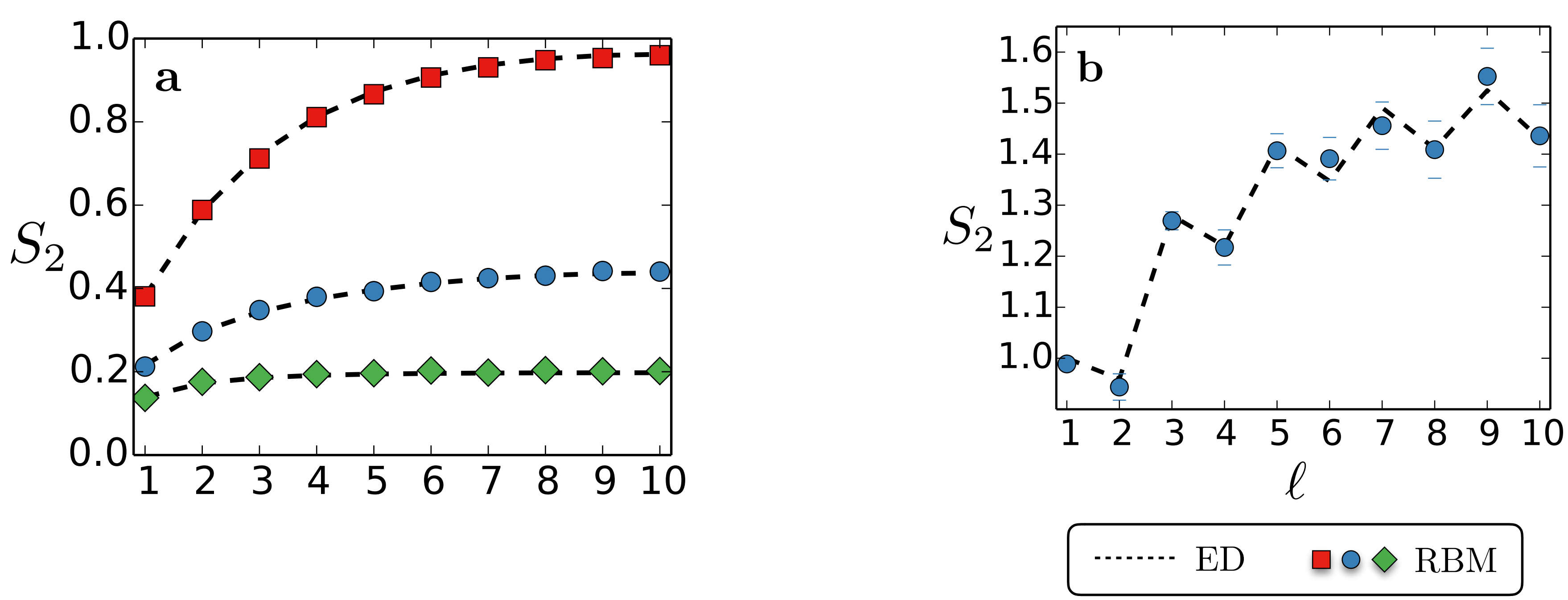
1609.02552

QUANTUM STATE TOMOGRAPHY WITH RBM

local observables



Second Renyi entropy

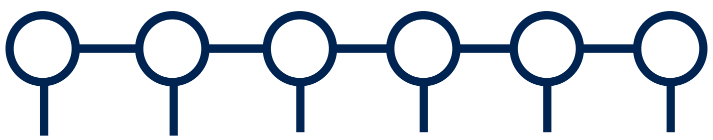


1703.05334

LET'S EXPLOIT THIS <ML|PHYSICS> MORE

Condensed matter/stat mech

Machine learning



Stoudenmire, Schwab 1605.05775

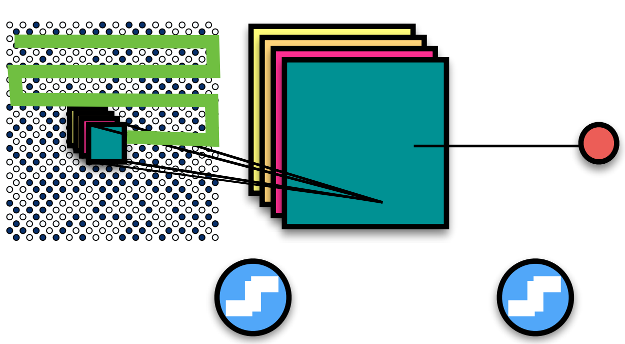
Energy-Based Models

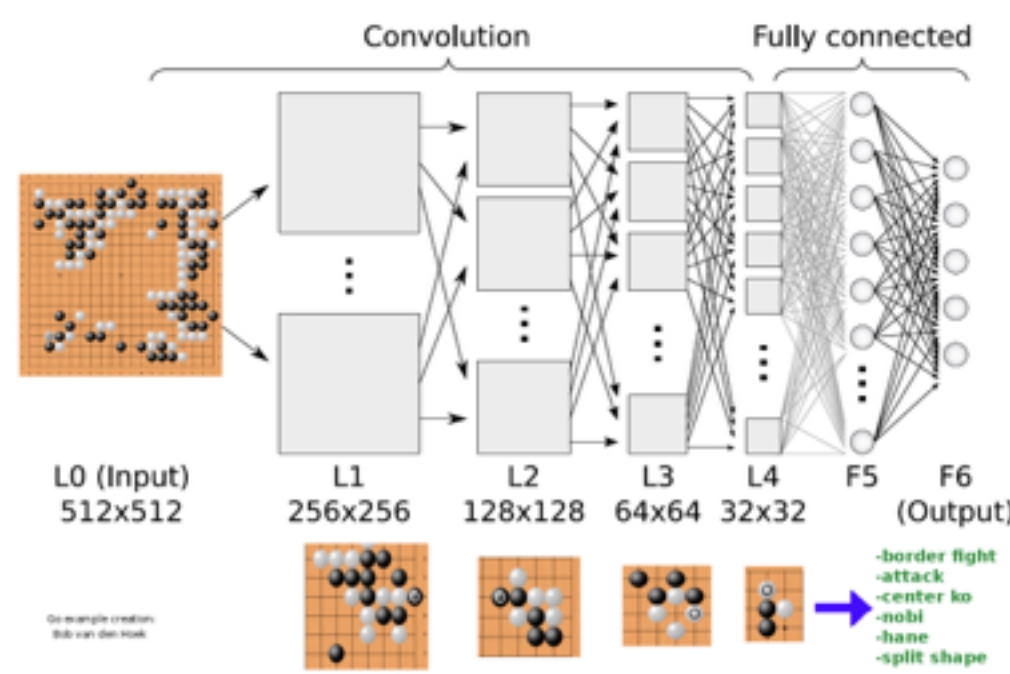
Carleo, Troyer 1606.02318

Torlai, Melko 1606.02718

Torlai et al 1703.05334

we are exploring it



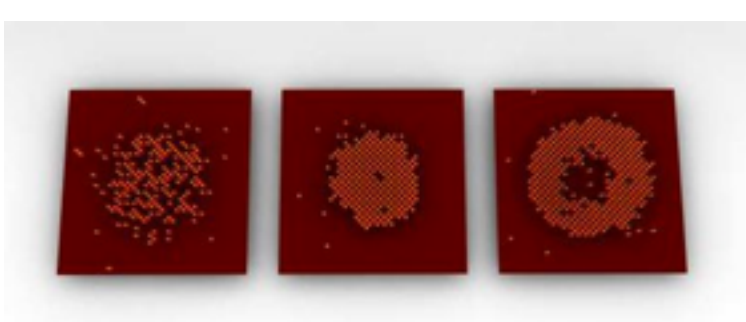


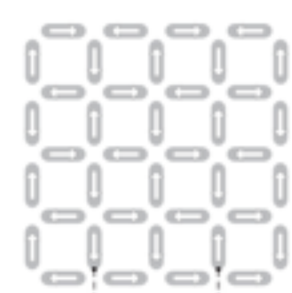
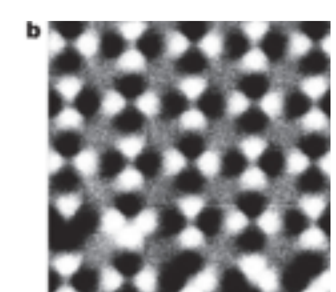
CONCLUSION

- We encode and discriminate phases and phase transitions, both conventional and topological, using neural network technology.
- We have an understanding of what the neural nets do in those cases through controlled analytical models.
- We have explored a somewhat new way to look at condensed matter systems. Hopefully this

FUTURE PROJECTS

- Variational interpretation of CNNs and their optimization for ground state
- Tomography with experimental data



QUANTUM MACHINE LEARNING POSITIONS AT D-WAVE

There are several positions at D-Wave at the intersection between

Machine learning, reinforcement learning, quantum physics, and quantum computing.

Friendly research atmosphere.

Great cities (Vancouver, Toronto)

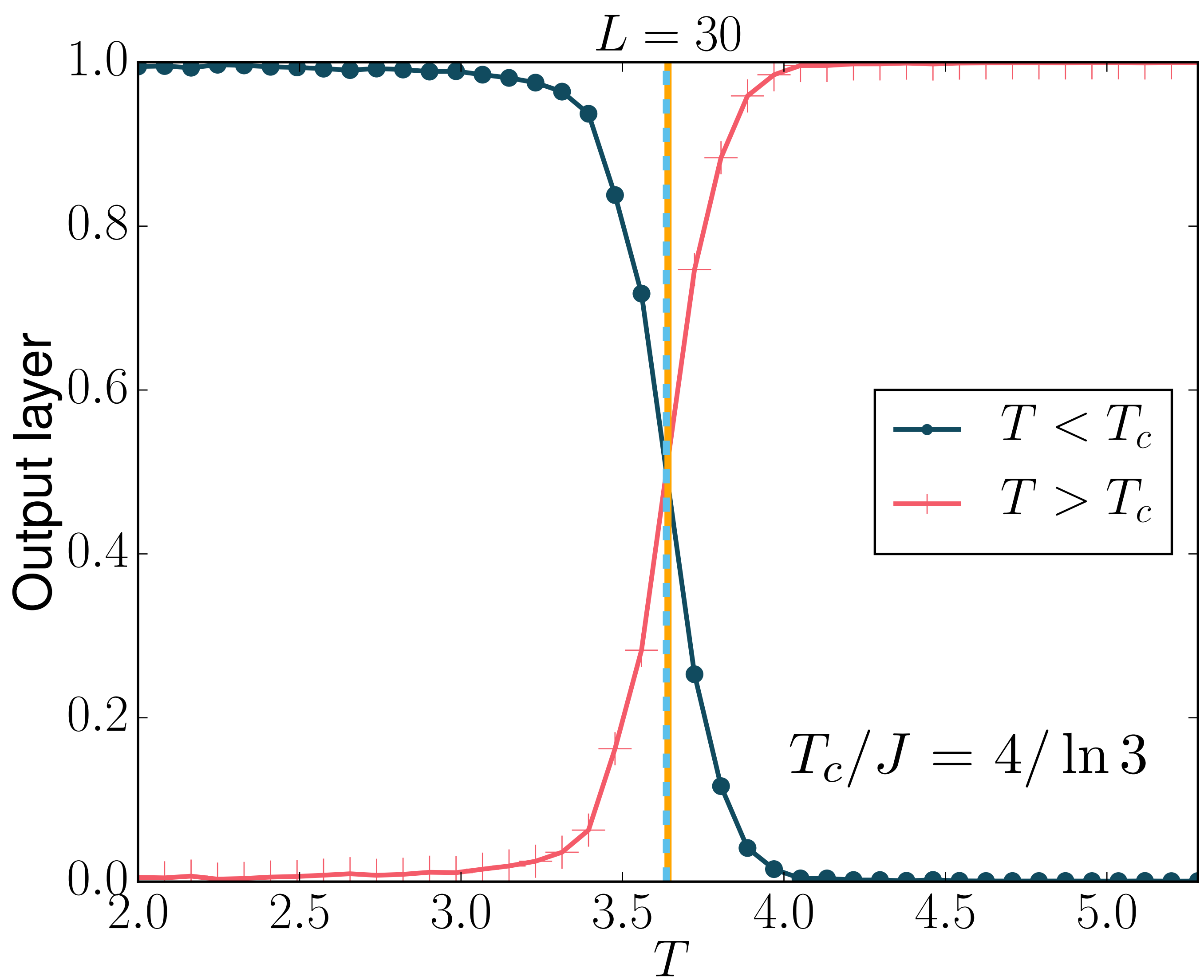
email me at: jcarrasquilla@dwavesys.com

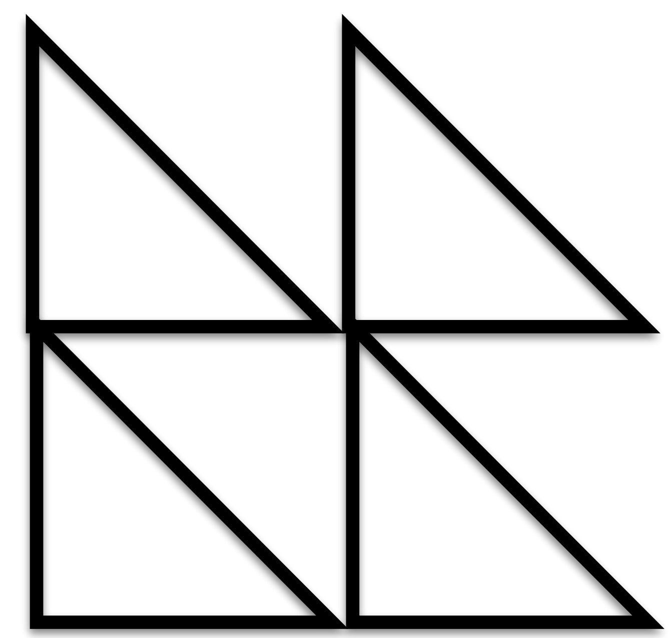
https://www.dwavesys.com/careers

D:WAVE
The Quantum Computing Company™

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!

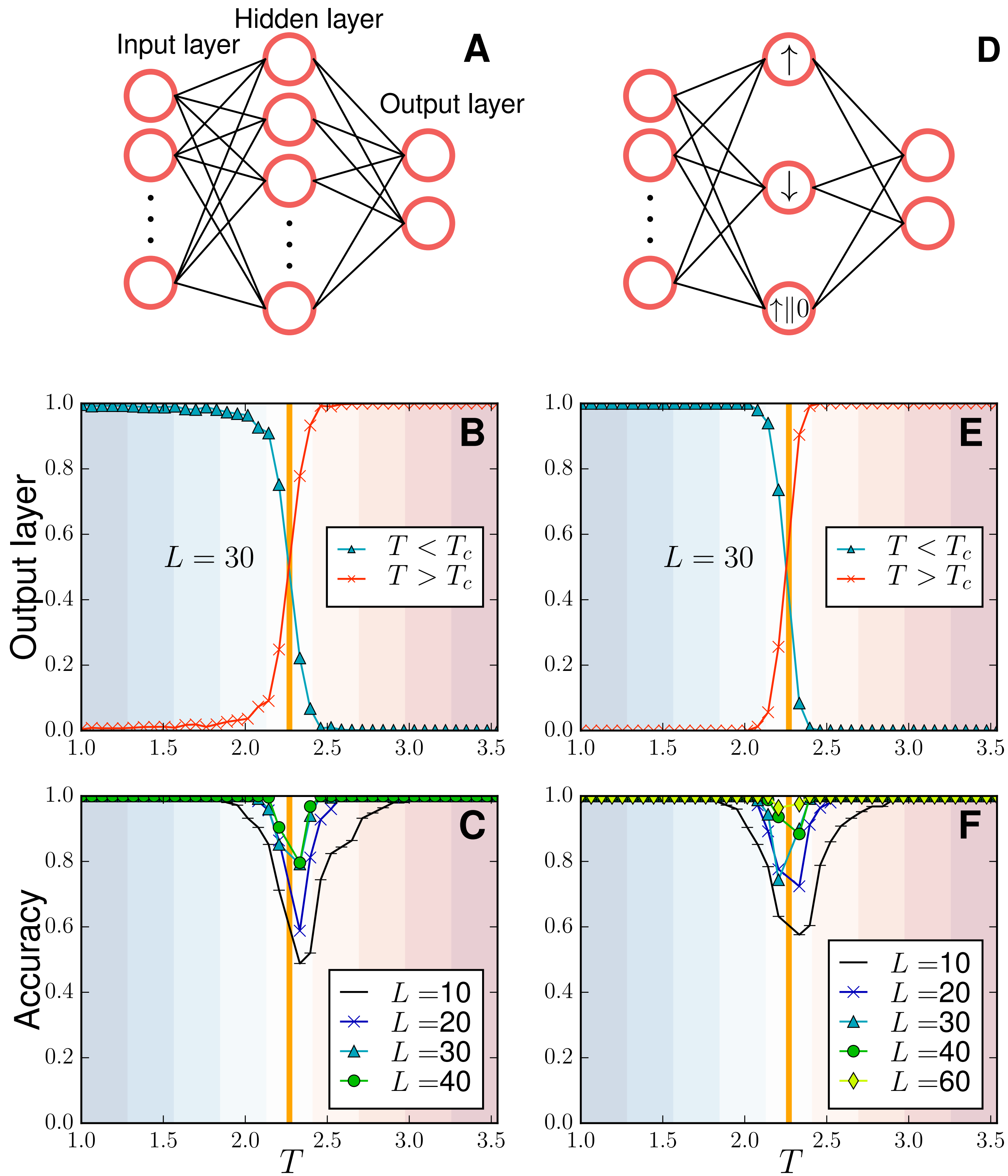




Tc within <1%!

1605.01735

ANALYTICAL UNDERSTANDING



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

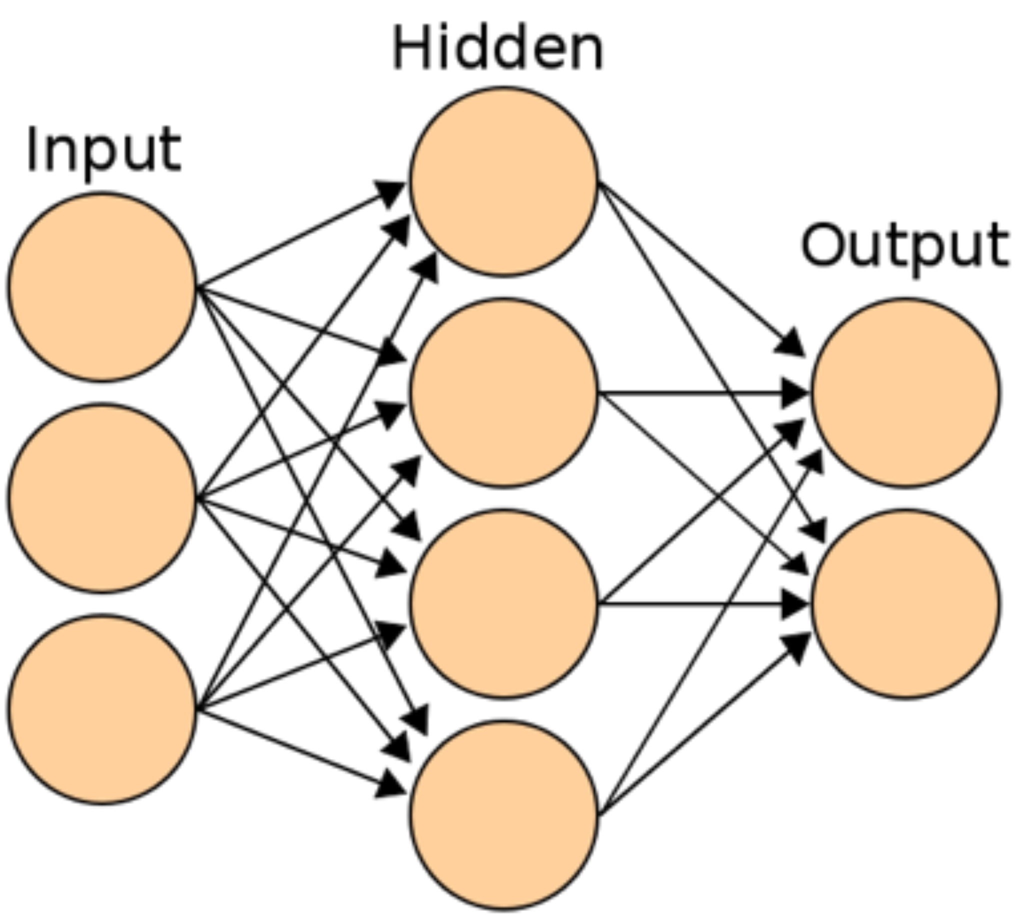
In general the hidden layer *discovers the order parameter* *of the phase during the training*

=>Works for AF Ising model

1605.01735

Artificial neural networks

Artificial neural networks are a family of models used to approximate functions that can take values on very high dimensional inputs. They are represented as interconnected "neurons" which are non-linear functions that exchange messages between each other



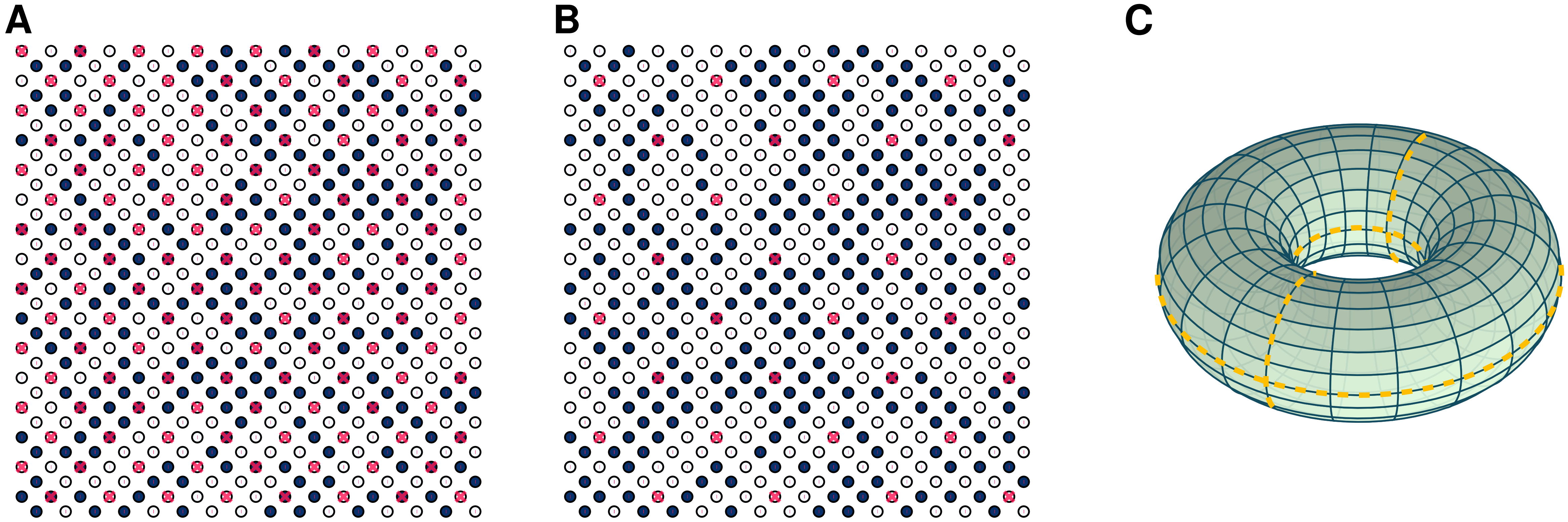
Connections= sets of adaptive weights, i.e. numerical parameters that are tuned by a learning algorithm

$$f : \mathbf{R}^n \to \mathbf{R}^m$$

Wikipedia

WHAT DOES THE CNN USE TO MAKE PREDICTIONS?

- It detects the local constraints induced by the Hamiltonian



ARTIFICIAL SQUARE ICE

- Wang and collaborators have used lithographic techniques to create a periodic two-dimensional array of single-domain sub micron ferromagnetic islands Nature 439, 303 (2006)

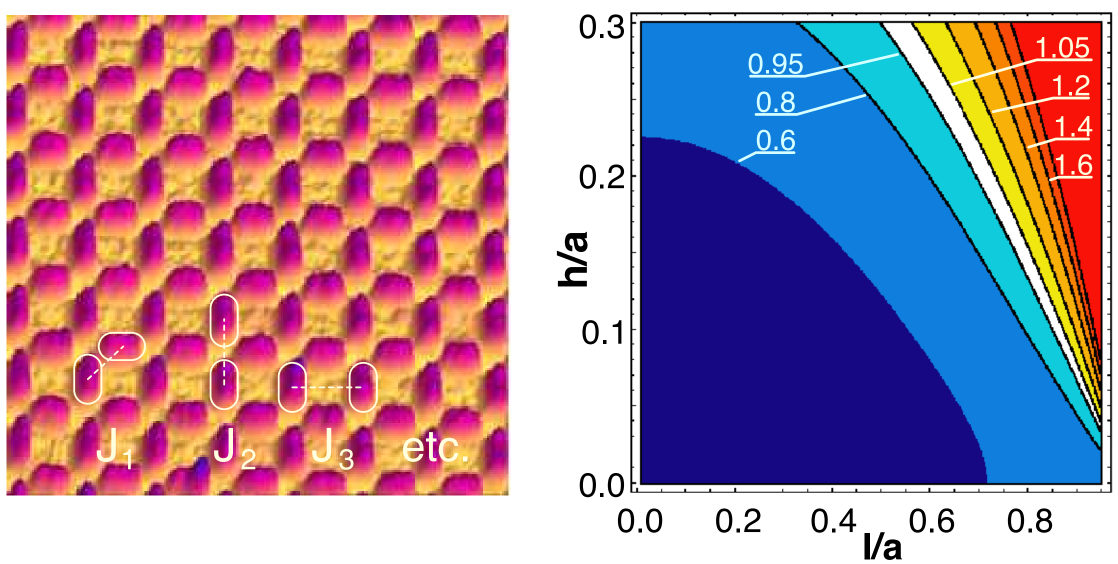


FIG. 1: (color online) Left: Atomic force microscope image of an array studied in Ref. 1. The islands have length $l = 220$nm, width 80nm and thickness 25nm. Right: Map of the ratio J_2/J_1 of the second to the first nearest neighbor interactions (highlighted in the left part) for different values of lattice constant, a, and sublattice height offset, h. In the white zone, $|J_2/J_1 - 1| < 5\%$. In the left (blue) region, the ordered state is antiferromagnetic, whereas it is ferromagnetic in the right (yellow-red) area.

INTERSECTION BETWEEN ML AND CONDENSED MATTER

Condensed matter
Statistical Mechanics ⟷ *Machine learning*

PHYSICS INSPIRED MACHINE LEARNING

MACHINE LEARNING INSPIRED PHYSICS

CONCEPTUAL CONNECTIONS

INTERSECTION BETWEEN ML AND CONDENSED MATTER

Condensed matter
Statistical Mechanics ⟷ *Machine learning*

PHYSICS INSPIRED MACHINE LEARNING
MACHINE LEARNING INSPIRED PHYSICS
CONCEPTUAL CONNECTIONS

-Boltzmann Machines (Hinton and Sejnowski) Dimensionality reduction, classification

-Spin glass interpretation of neural nets

-Quantum Boltzmann machine (1601.02036) quantum annealers, dwave

-Learning with Quantum-Inspired Tensor Networks (Stoudenmire, Schwab 1605.05775)

INTERSECTION BETWEEN ML AND CONDENSED MATTER

Condensed matter
Statistical Mechanics

Machine learning

PHYSICS INSPIRED MACHINE LEARNING
MACHINE LEARNING INSPIRED PHYSICS
CONCEPTUAL CONNECTIONS

-Phases and phase transitions (our work 1605.01735)

-Unsupervised learning phase transitions (Lei Wang 1606.00318)

-Ground state and real-time dynamics with Boltzmann Machines (Carleo, Troyer 1606.02318)

-Relaxation in glassy liquids (Nature Physics 12, 469–471 (2016))

-Materials design (npj Computational Materials 1, Article number: 15008 (2015))

- and many more …

INTERSECTION BETWEEN ML AND CONDENSED MATTER

Condensed matter
Statistical Mechanics

Machine learning

PHYSICS INSPIRED MACHINE LEARNING
MACHINE LEARNING INSPIRED PHYSICS
CONCEPTUAL CONNECTIONS

-Deep learning and the renormalization group (and MERA) (C. Beny 1301.3124)

-An exact mapping between the Variational Renormalization Group and Deep Learning (Schwab, Mehta 1410.3831)

-Could we transform this into a practical ML tool? (like Miles did with DMRG)

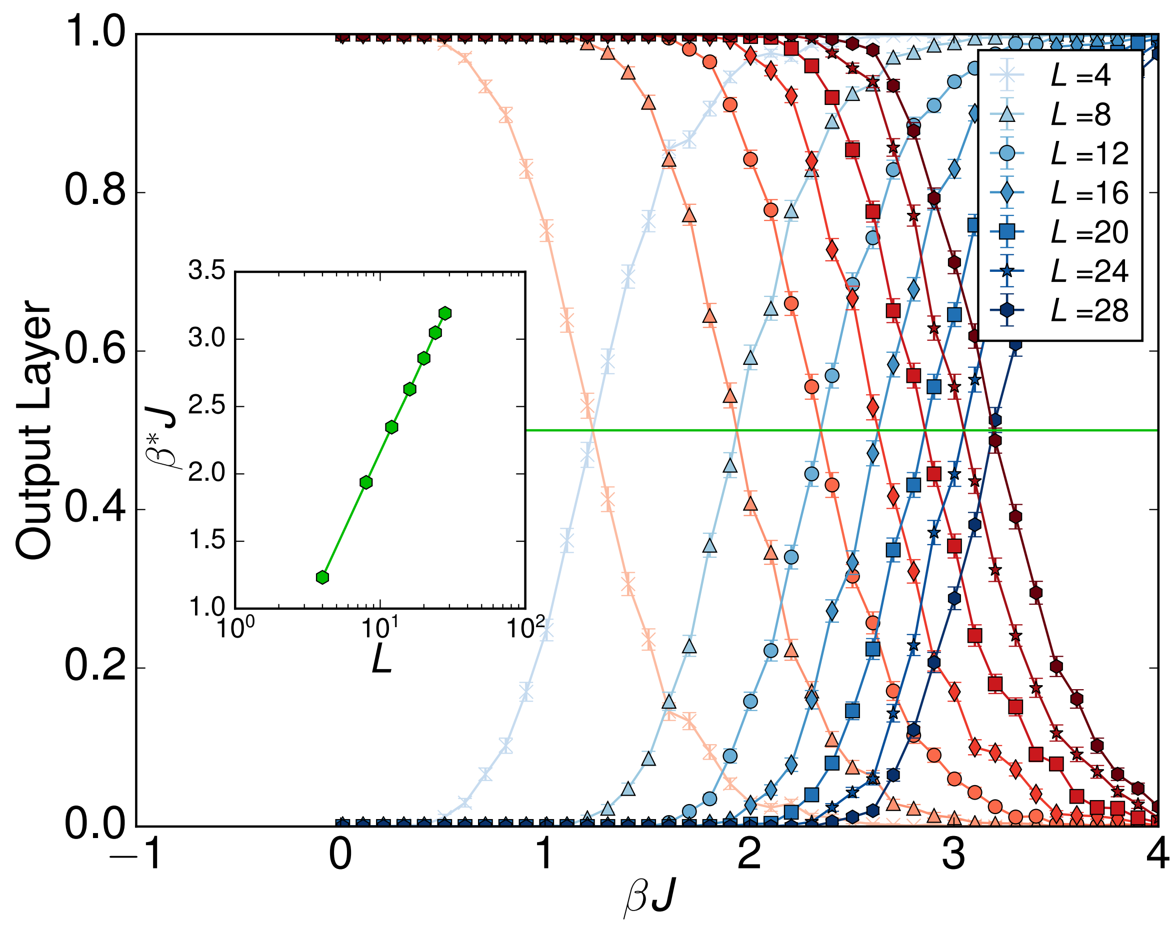


Figure 3 Detecting the logarithmic crossover temperatures in the Ising gauge theory. Output neurons for different system sizes averaged over test sets vs βJ. Linear system sizes $L = 4, 8, 12, 16, 20, 24,$ and 28 are represented by crosses, up triangles, circles, diamonds, squares, stars, and hexagons. The inset displays $\beta^* J$ (octagons) vs L in a semilog scale. The error bars represent the one standard deviation statistical uncertainty.

slowly cross over to the high-temperature phase. The cross-over temperature T^* happens as the number of thermally excited defects $\sim N \exp(-2J\beta)$ is of the order of one, implying $T^*/J \sim 1/\ln\sqrt{N}$.[23] As the presence of local defects is the mechanism through which the CNN decides

FOR A REAL NN THIS REMAINS TRUE