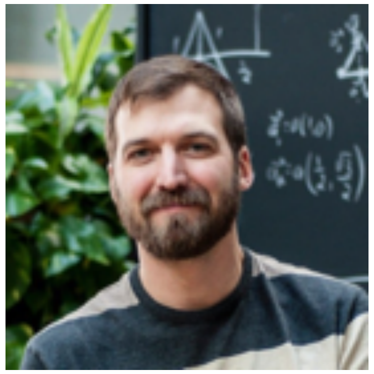
A machine learning perspective on the many-body problem in classical and quantum physics

Juan Carrasquilla

December 8th, 2017

Deep Learning for Physical Sciences
Workshop at the 31st NIPS Conference
Long Beach, California

COLLABORATORS



Roger Melko (U. Waterloo and Perimeter)

Giacomo Torlai (U. Waterloo and Perimeter)

Giuseppe Carleo (ETH)

Matthias Troyer (Microsoft, ETH)

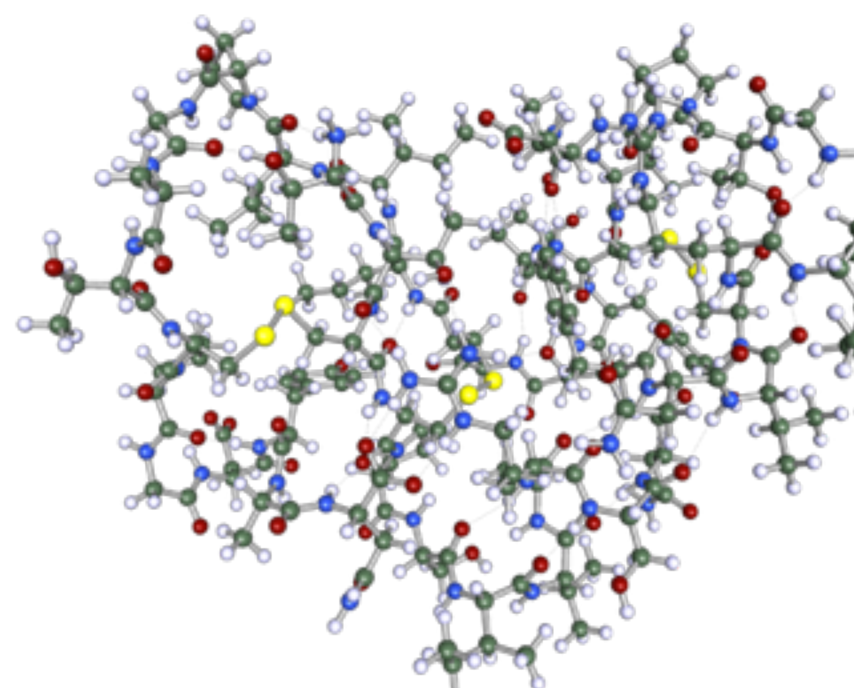
Guglielmo Mazzola (ETH)

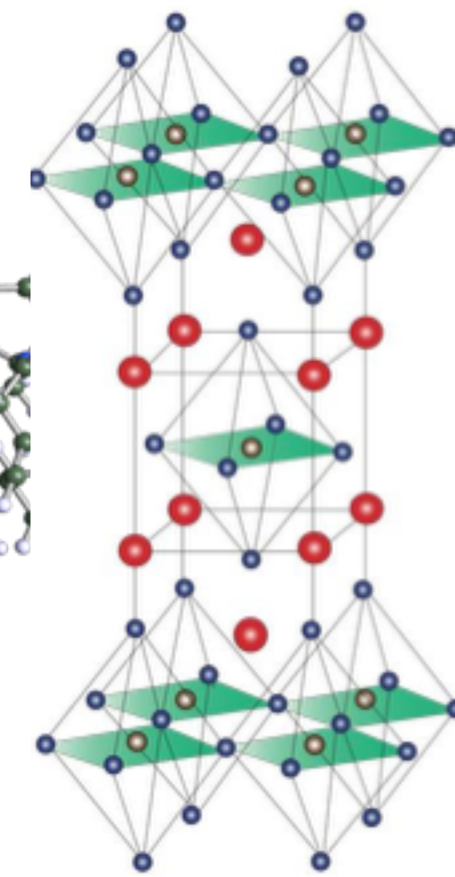
THE COMPLEXITY OF THE MANY-BODY PROBLEM IN CLASSICAL AND QUANTUM MECHANICS

THE MANY-BODY PROBLEM IN QUANTUM MECHANICS

- Generic specification of a quantum state requires resources exponentially large in the number of degrees of freedom N
- Today's best supercomputers can solve the wave equation **exactly** for systems with a maximum of ~45 particles.
- Storing the state of a 273 spin system requires a computer with more bits than there are atoms in the universe
- Yet, technologically relevant problems in chemistry, condensed matter physics, and quantum computing are much larger than 273.
- Quantum computing

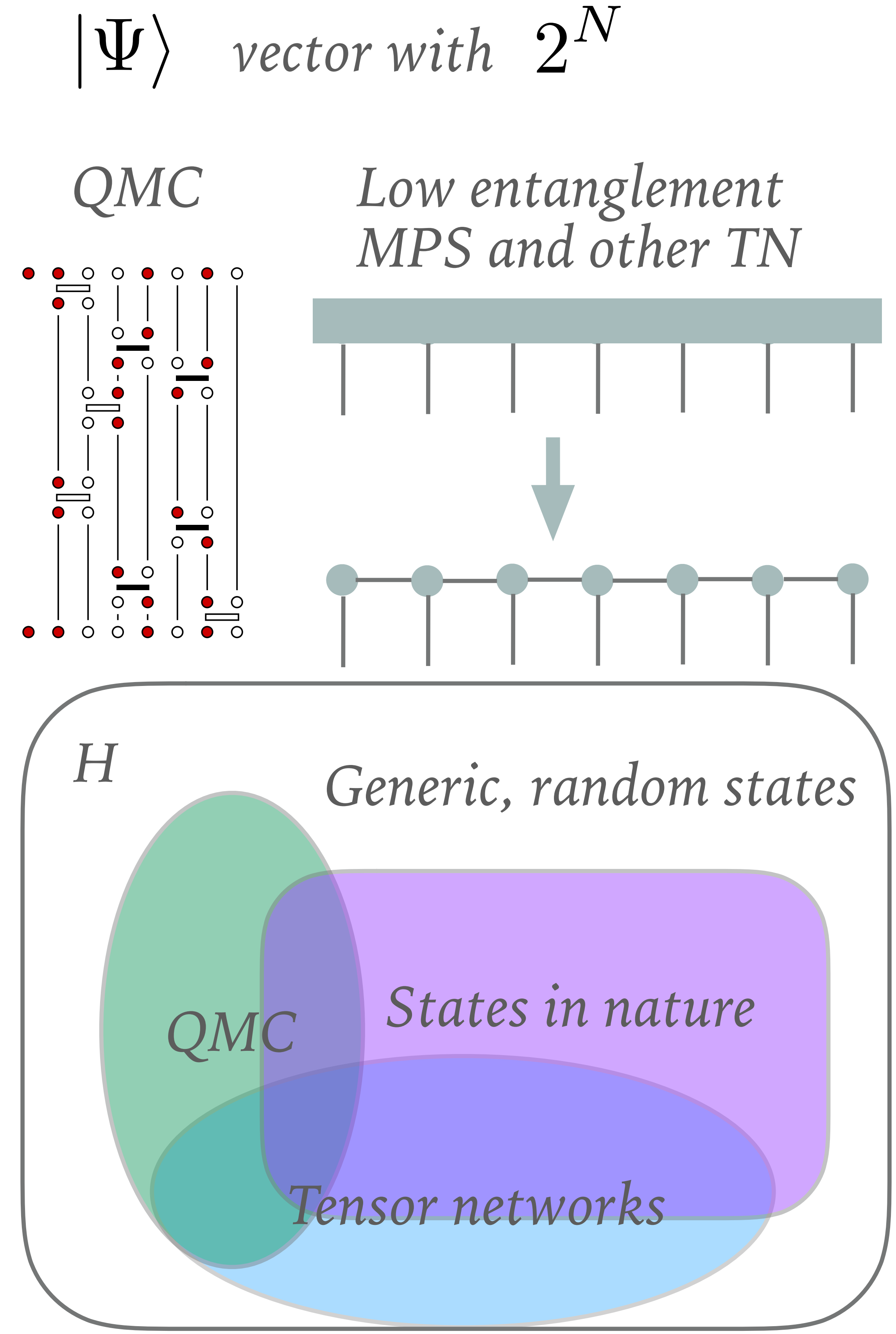
$|\Psi\rangle$ *vector with* 2^N





THERE IS STILL HOPE FOR CLASSICAL ALGORITHMS

- Nature is sometimes compassionate: many-body systems can be typically characterized by an amount of information smaller than the maximum capacity of the corresponding state space.
- Quantum Monte Carlo and other numerical methods based on Tensor Networks exploit this fact and are able to accurately study large quantum system in practice with limited amount of resources.
- Machine learning community deals with equally high dimensional problems and battle the curse of dimensionality successfully with impressive results in a wide spectrum of scientific and technologically relevant areas of research.

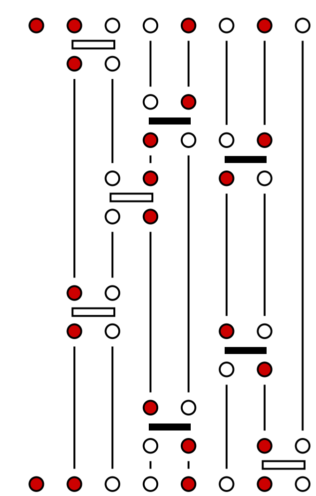


THERE IS STILL HOPE FOR CLASSICAL ALGORITHMS

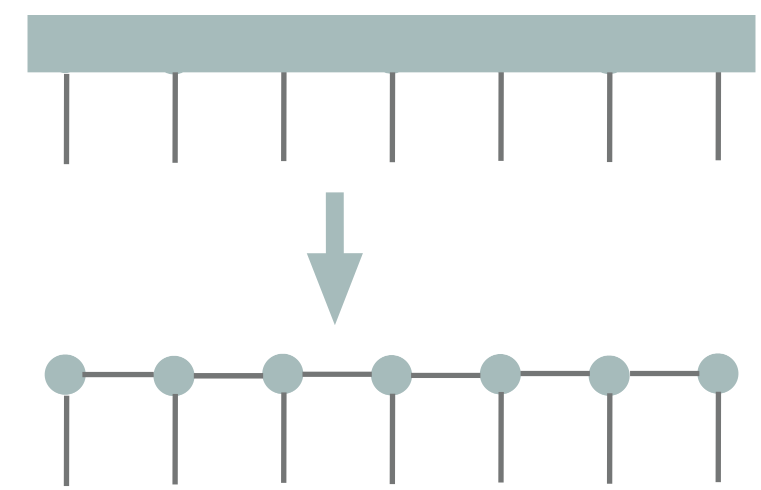
- Nature is sometimes benevolent: many-body systems can be typically characterized by an amount of information smaller than the maximum capacity of the corresponding state space.
- Quantum Monte Carlo and other numerical methods based on Tensor Networks exploit this fact and are able to accurately study large quantum system in practice with limited amount of resources.
- Machine learning community deals with equally high dimensional problems and battle the curse of dimensionality successfully with impressive results in a wide spectrum of scientific and technologically relevant areas of research.

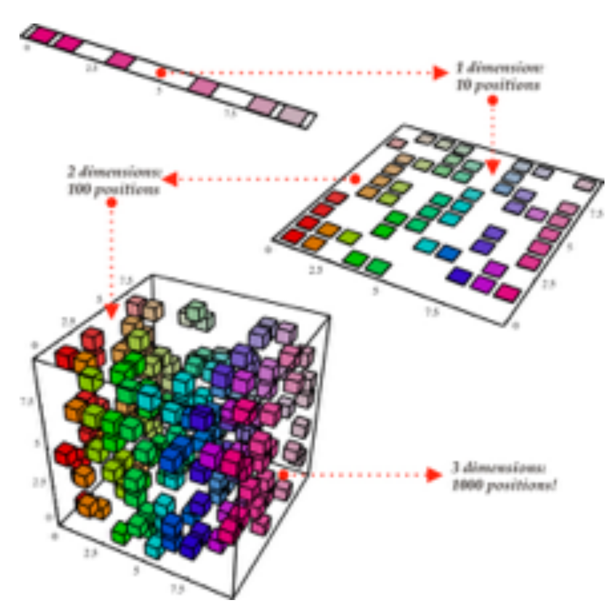
$|\Psi\rangle$ *vector with* 2^N

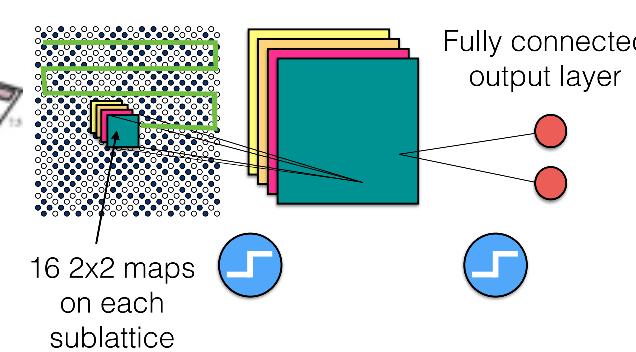
QMC



Low entanglement MPS and other TN







QUANTUM AND CLASSICAL MANY-BODY PHYSICS HAS NOT BEEN THE EXCEPTION

- ML phases of matter/phase transitions (Carrasquilla, Melko 1605.01735, Wang 1606.00318, Zhang, Kim, 1611.01518)
- New ML inspired ansatz for quantum many-body systems (Carleo, Troyer 1606.02318, Deng, Li, Das Sarma, 1701.04844, Deng, Li, Das Sarma 1609.09060, Carrasquilla, Melko 1605.01735)
- Accelerated Monte Carlo simulations (Huang, Wang 1610.02746)
- Quantum state preparation guided by ML (Bukov, Day, Sels, Weinberg, Polkovnikov and Mehta 1705.00565)
- Renormalization group analyses, RBMs, PCA (Bradde, Bialek 1610.09733, Koch-Janusz, Ringel 1704.06279,Mehta, Schwab,1410.3831)
- Quantum state tomography based on RBMs (Torlai, Mazzola, Carrasquilla, Troyer, Melko, Carleo, 1703.05334)
- ML based decoders for topological codes (Torlai, Melko 1610.04238, Varsamopoulos, Criger, Bertels, 1705.00857)
- Supervised Learning with Quantum-Inspired Tensor Networks (Stoudenmire, Schwab 1605.05775, Novikov, Trofimov, Oseledets, 1605.03795)
- Quantum Boltzmann machines (Amin, Andriyash, Rolfe, Kulchytskyy, Melko, 1601.02036, Kieferova, Wiebe, 1612.05204,)
- Quantum machine learning algorithms to accelerate learning (Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, Lloyd, 1611.09347)

And many more

IN THIS TALK

- I will discuss several applications of ML ideas to problems in many-body physics.
- Supervised learning approach to classical and quantum phases and phase transitions (Ising models)
- Interpreting wave function as a generative model: ground state of Kitaev's topological toric code using convolutional neural networks.
- Data intensive problem in quantum mechanics: quantum state tomography with neural networks (RBMs) $|\Psi\rangle$ 2^N

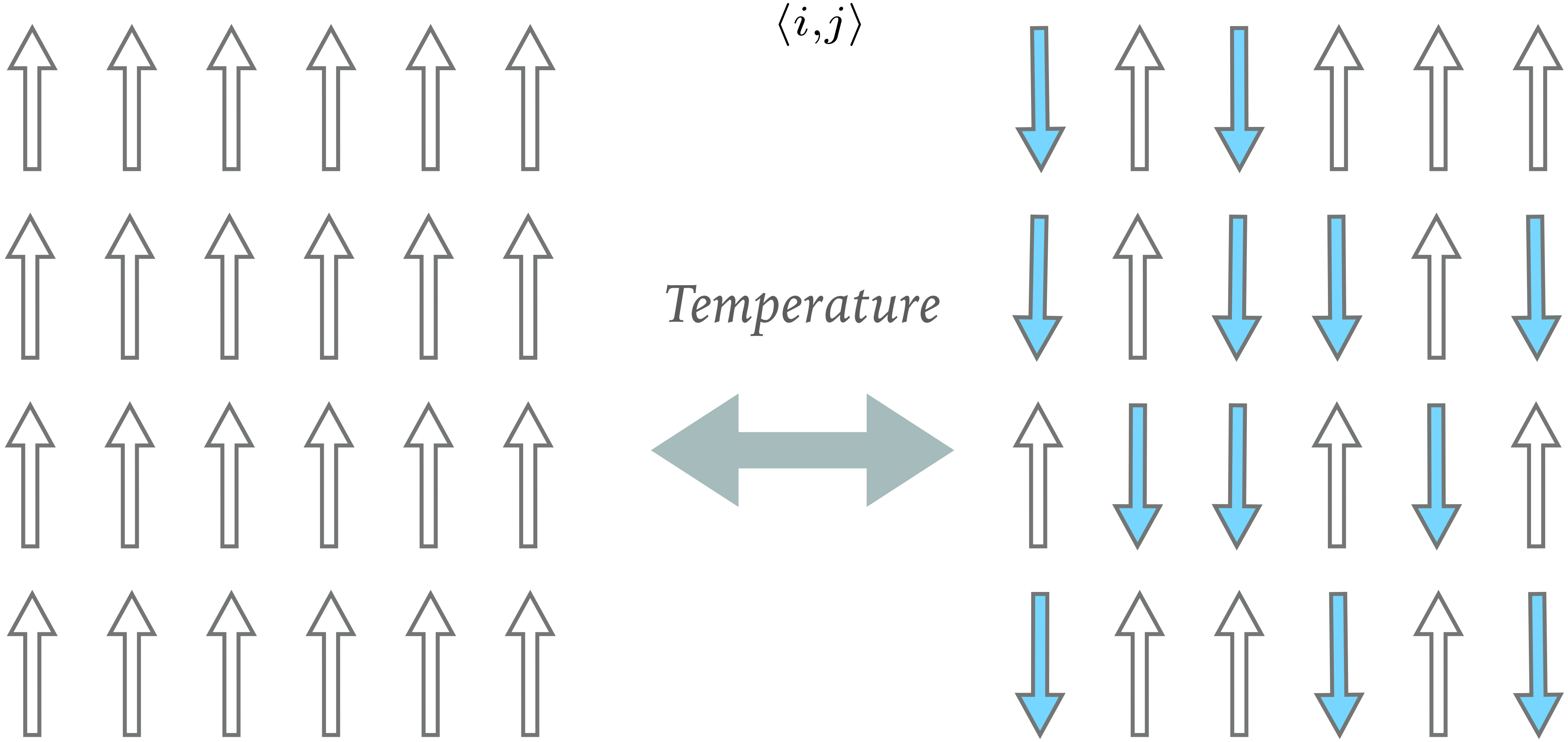
SUPERVISED LEARNING PERSPECTIVE OF PHASES OF MATTER

TOY PROBLEM: ISING MODEL

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ising ferromagnet in two dimensions

$$E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \qquad \sigma_i = \pm 1$$



Ferromagnet

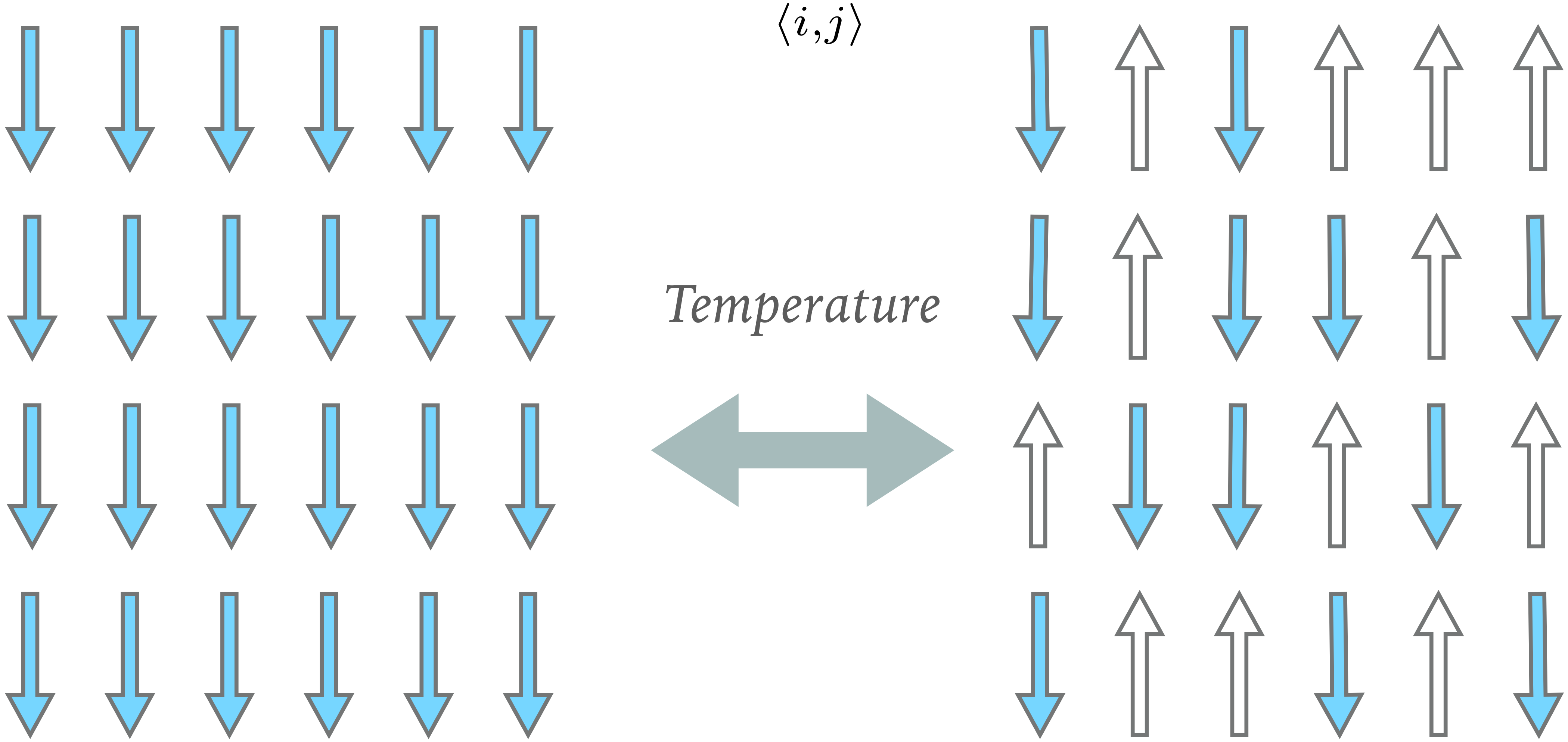
Lars Onsager Phys. Rev. 65, 117

Paramagnet

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ising ferromagnet in two dimensions

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

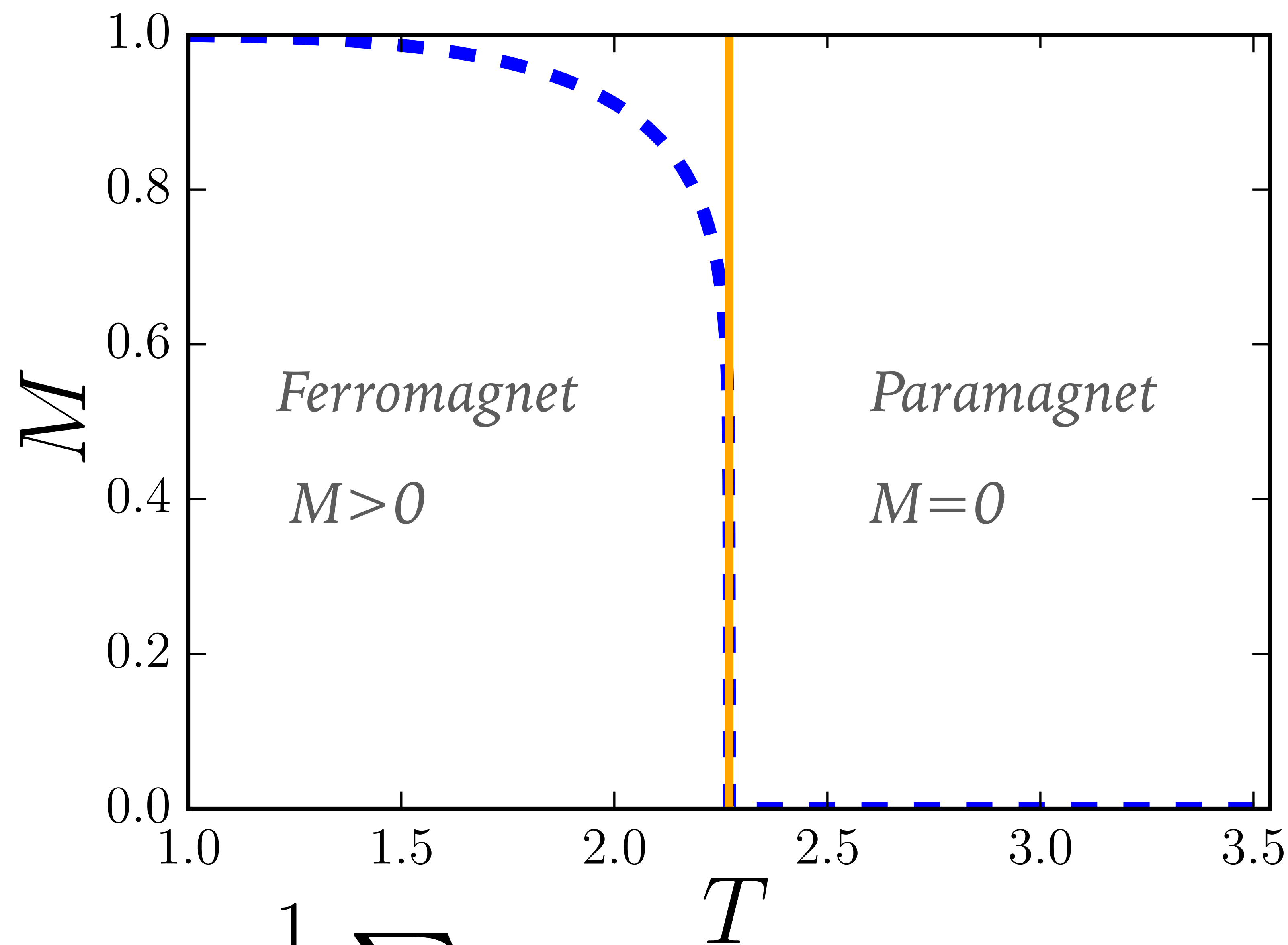


Ferromagnet

Paramagnet

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

Ferromagnetic transition: order parameter



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

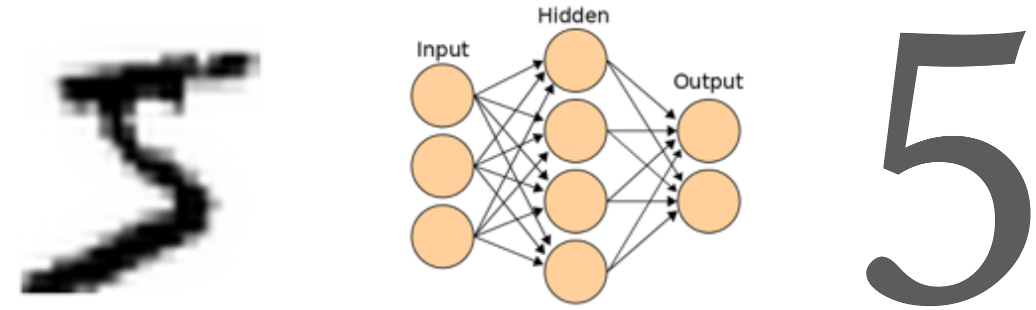
$$M = \frac{1}{N}\sum_i \langle \sigma_i \rangle, \qquad \sigma_i = \pm 1$$

Lars Onsager Phys. Rev. 65, 117 (1944)

WHAT DO I MEAN BY MACHINE LEARNING PHASES OF MATTER?

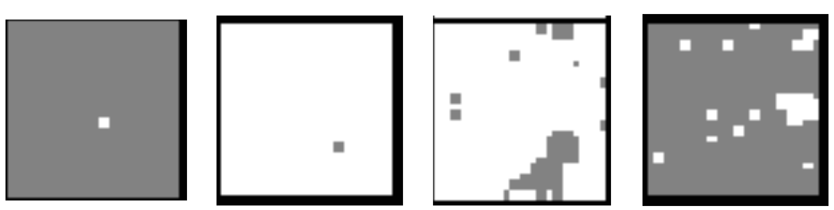
INSPIRATION: FLUCTUATIONS HANDWRITTEN DIGITS (MNIST)

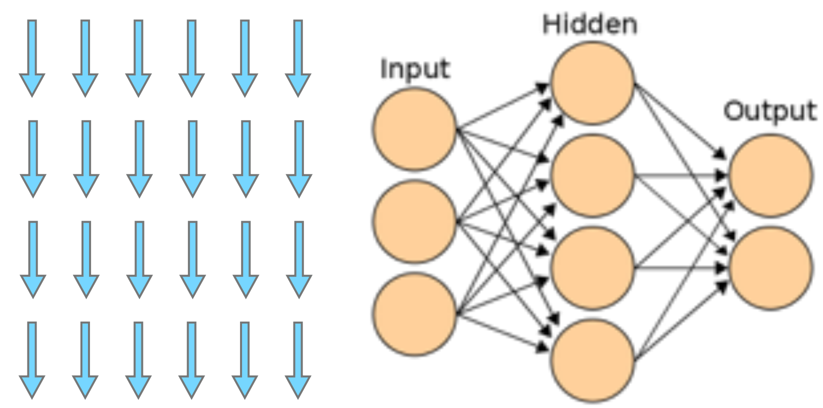
 $= 5 +$ *fluctuations*



ML community has developed powerful supervised learning algorithms

FM phase



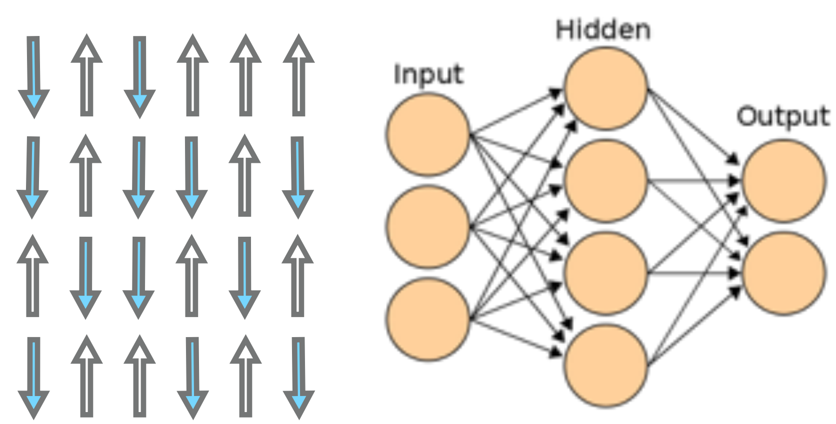


FM (0)

High T phase

gray=spin up

white=spin down

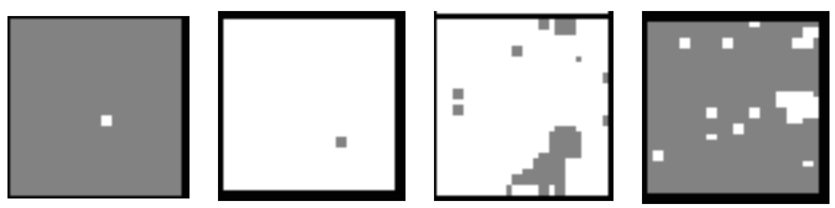


PM (1)

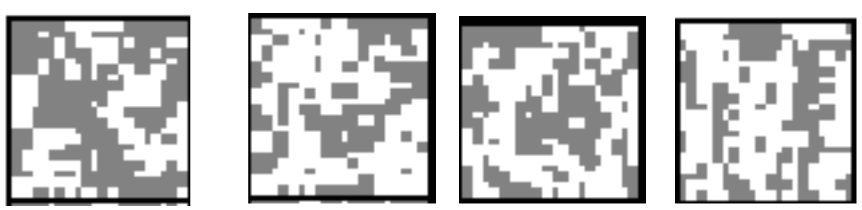
J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

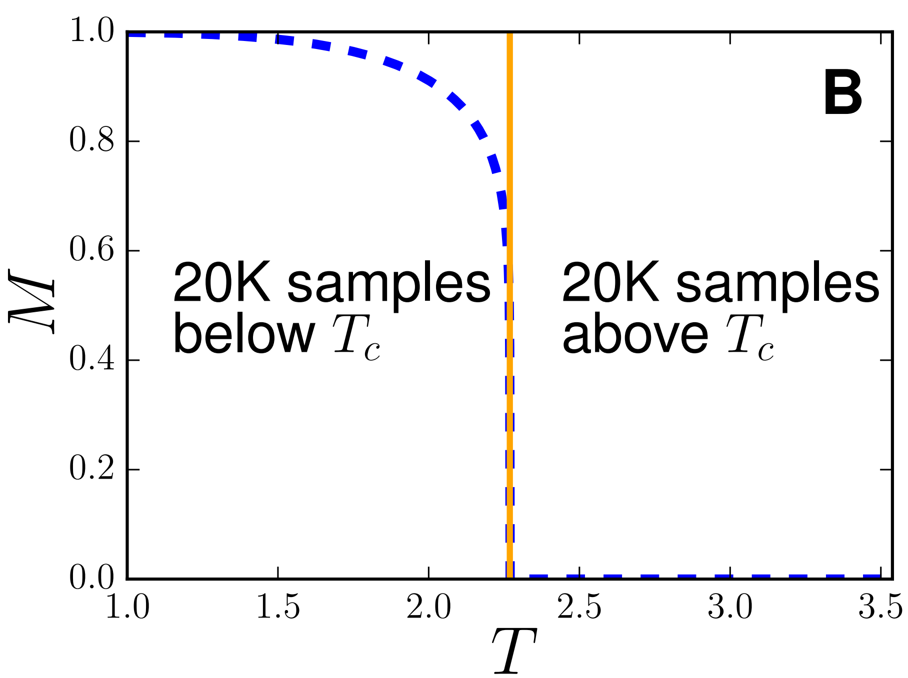
COLLECTING THE TRAINING/TESTING DATA: MC SAMPLING ISING MODEL AND LABELS

2D Ising model in the ordered phase



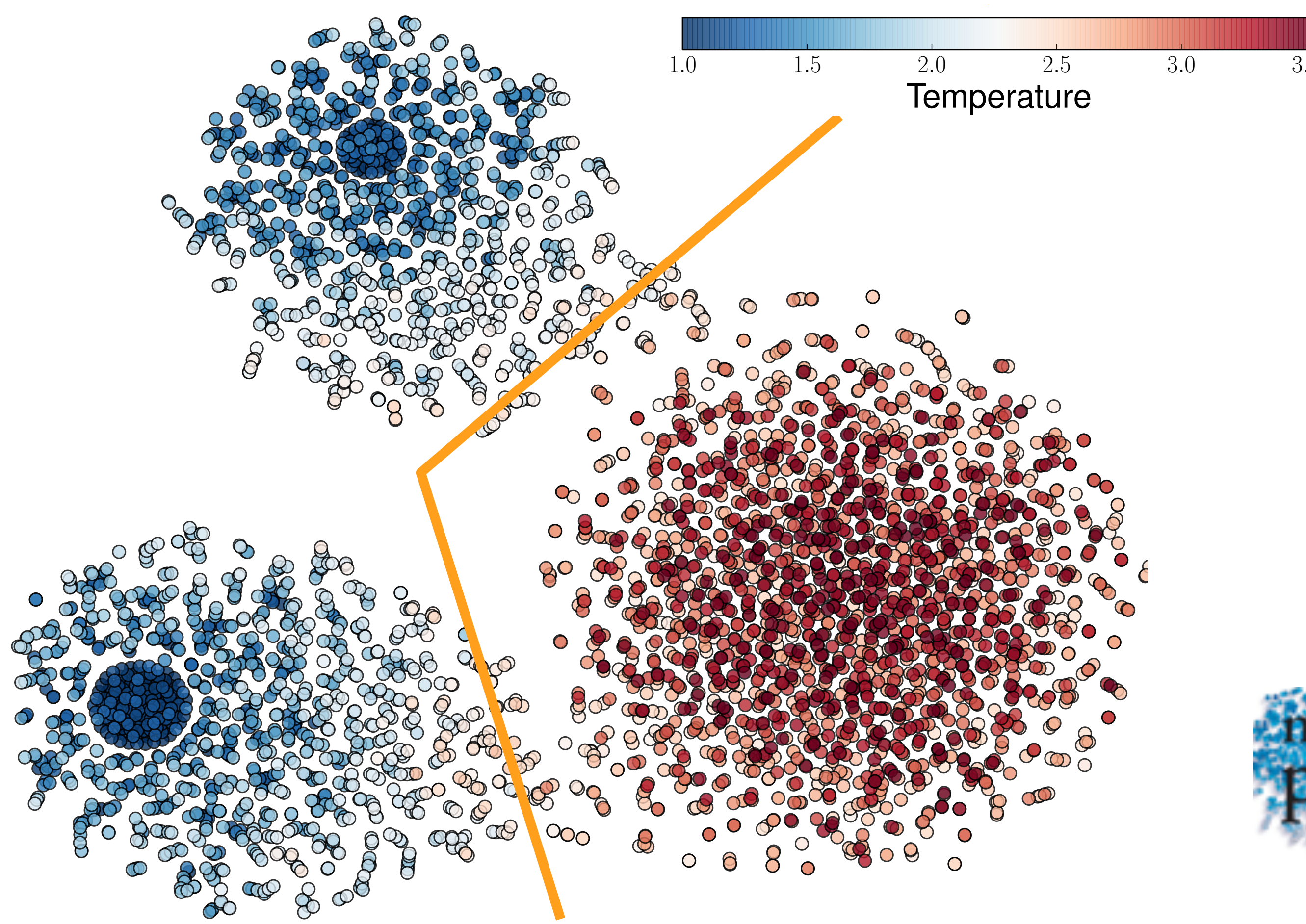
2D Ising model in the disordered phase





Training/testing data is drawn from the Boltzmann distribution

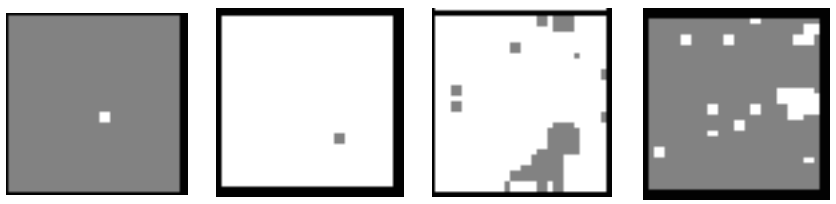
$$p(\sigma_1, \sigma_2, \ldots, \sigma_N) = \frac{e^{-\beta E(\sigma_1, \sigma_2, \ldots, \sigma_N)}}{Z(\beta)}$$



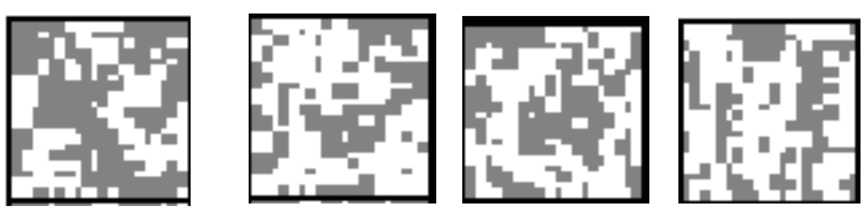
J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

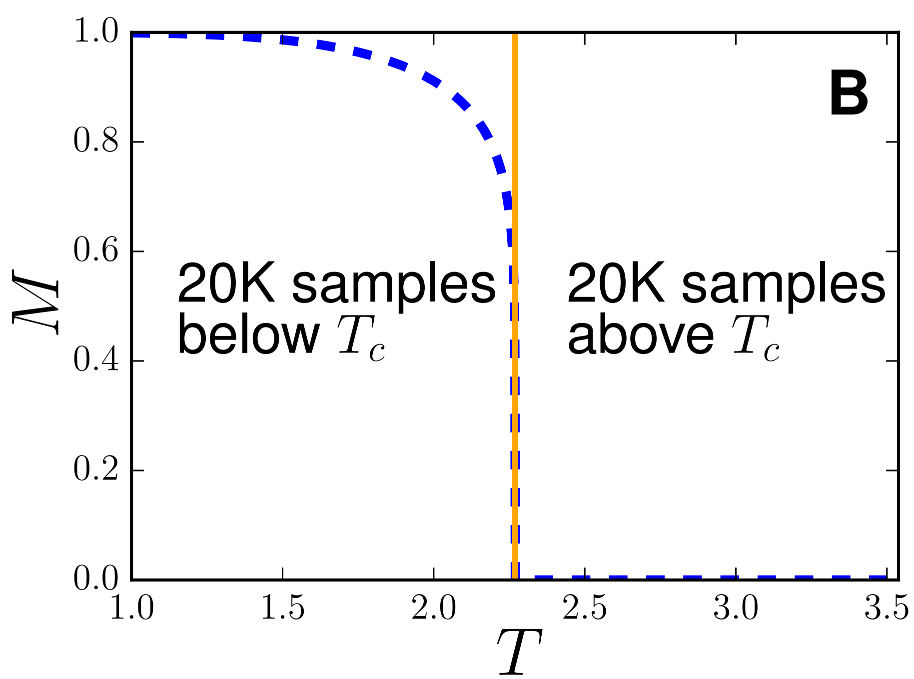
RESULTS: SQUARE LATTICE ISING MODEL (TEST SETS)

2D Ising model in the ordered phase



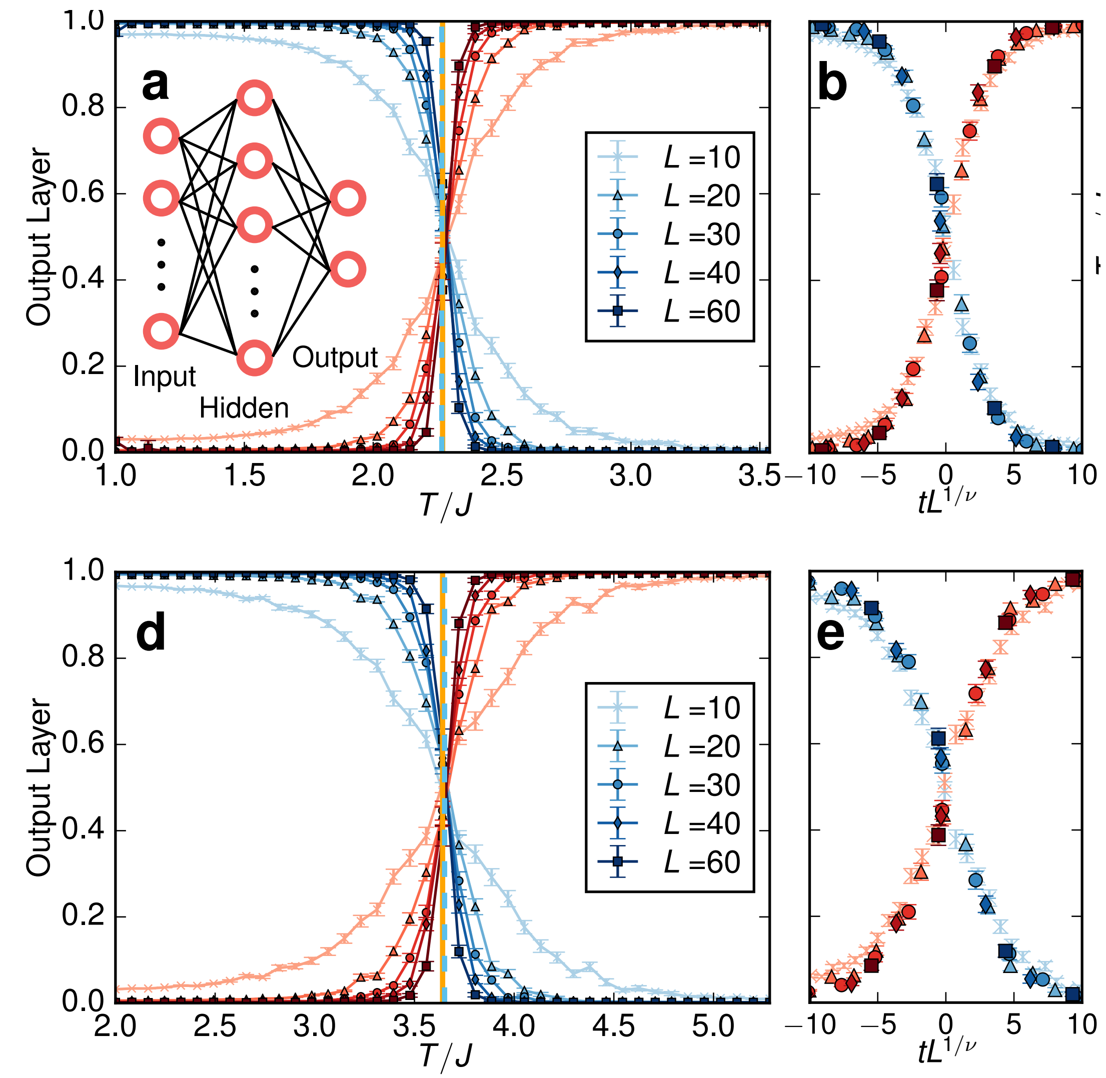
2D Ising model in the disordered phase





Analytical understanding: toy model with 3 analytically trained neutrons. NN relies on the **magnetization** of the system

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



Tc of Triangular within <1%

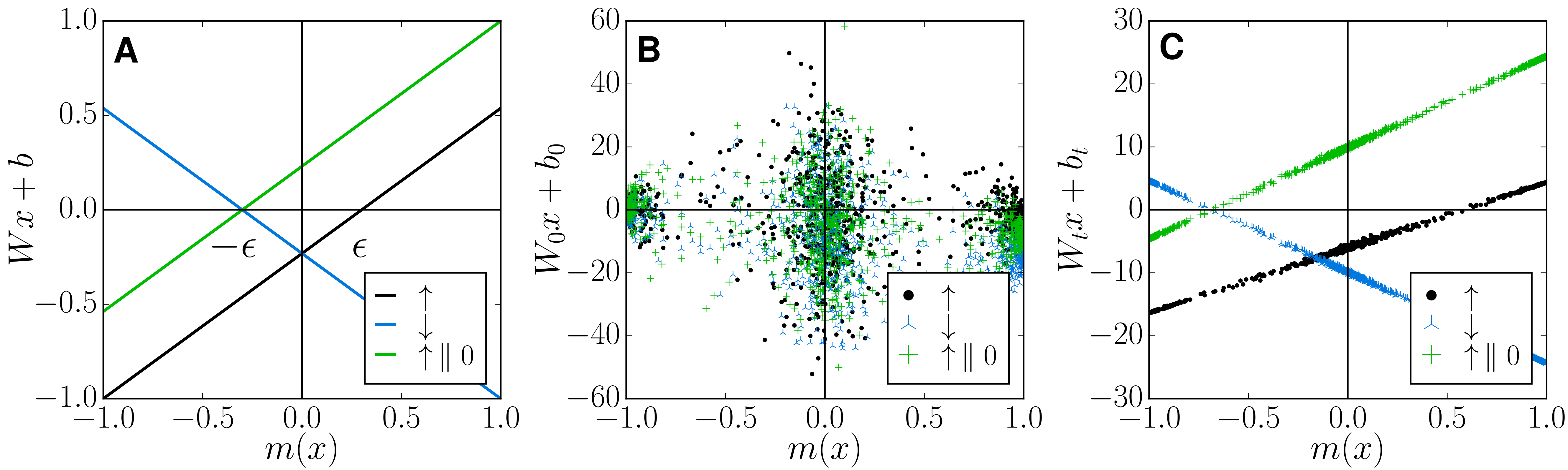
NN knows about criticality

$\nu \approx 1$

J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

ANALYTICAL UNDERSTANDING

Investigating the argument of the hidden layer during the training



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

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

J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

CAN WE DEAL WITH DISORDERED AND TOPOLOGICAL PHASES NOT DESCRIBED BY ORDER PARAMETERS?

PHASES OF MATTER WITHOUT AN ORDER PARAMETER AT T=0

- Topological phases of matter. Examples: Fractional quantum hall effect, quantum spin liquids, Ising gauge theory. Potential applications in topological quantum computing. Interestingly, these phases defy the Landau symmetry breaking classification.

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

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

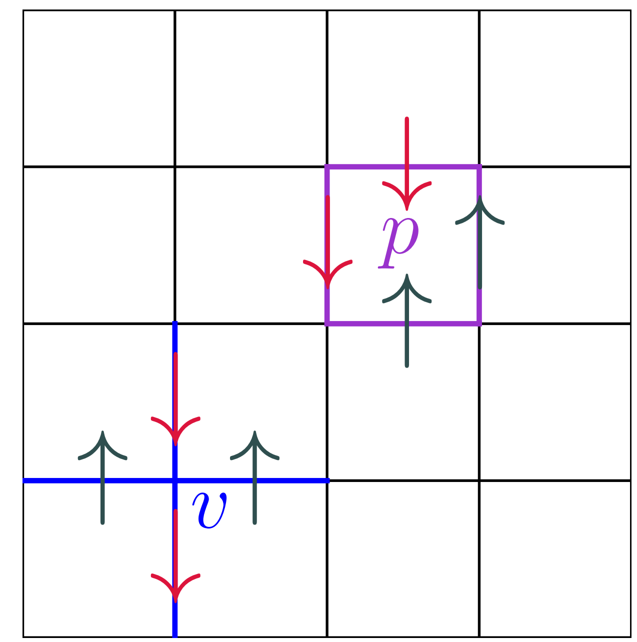
Wegner's Ising gauge theory:

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

F.J. Wegner, J. Math. Phys. 12 (1971) 2259

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

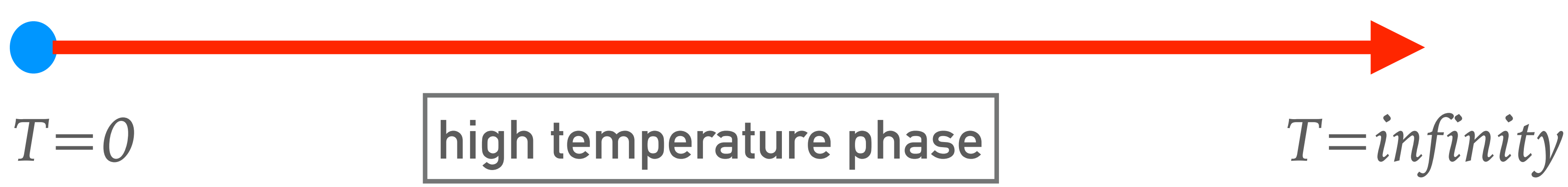
The ground state is a highly degenerate manifold with exponentially decaying spin–spin correlations.



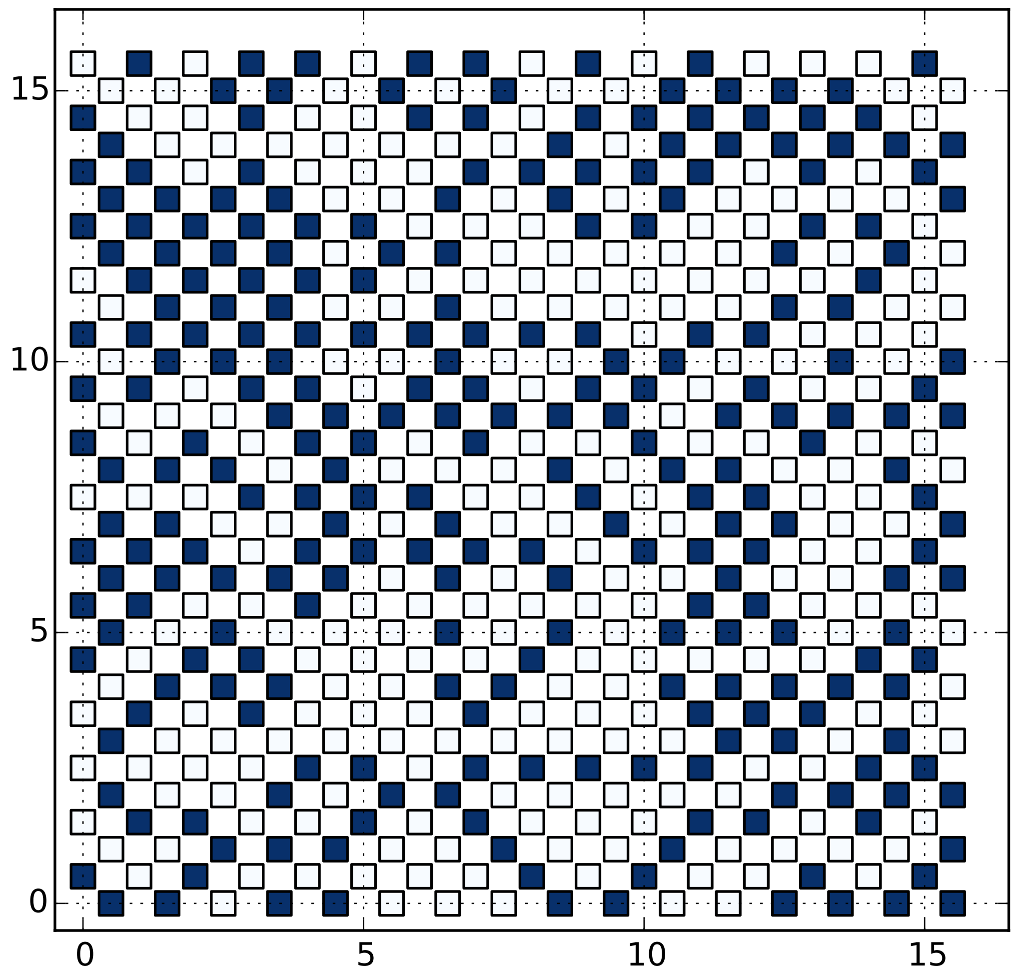
The grandmother of all lattice models for topological quantum computation

Ground state is a classical disordered topologically ordered phase

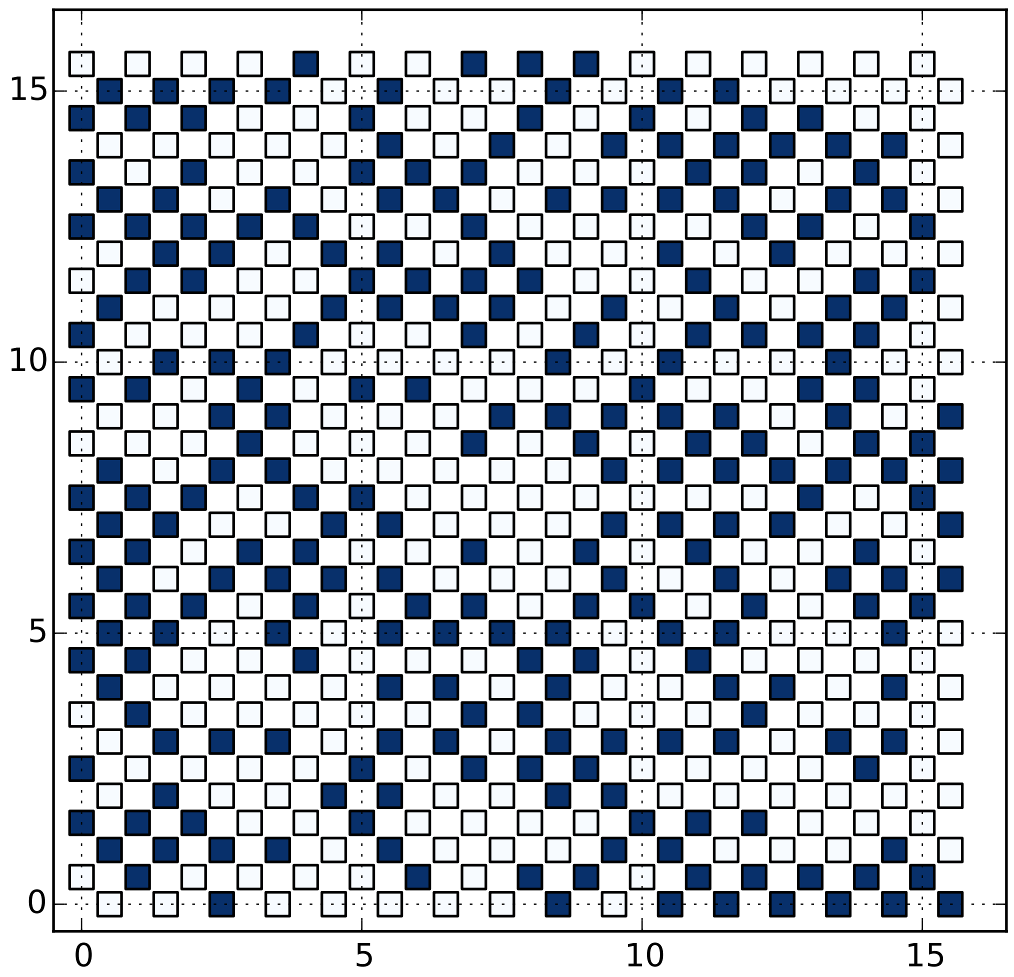
Castelnovo and Chamon Phys. Rev. B 76, 174416 (2007)



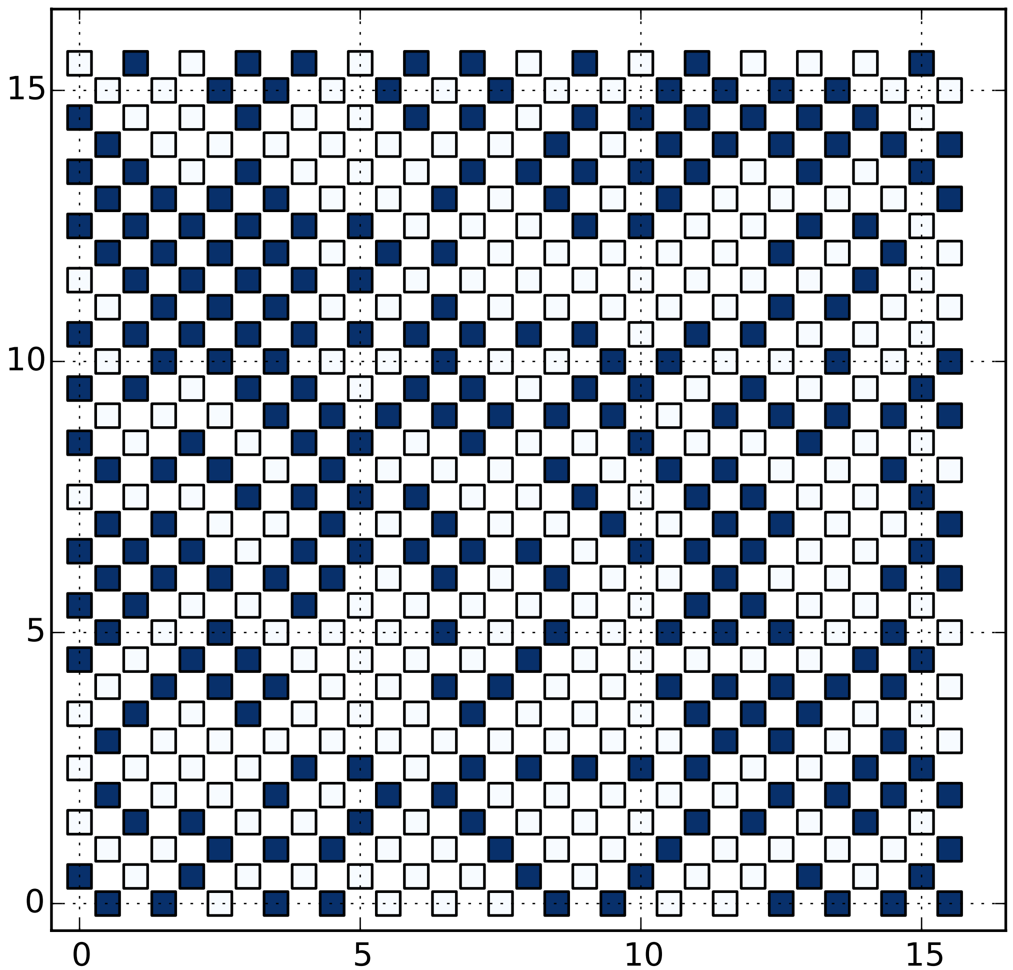
For two configurations



?

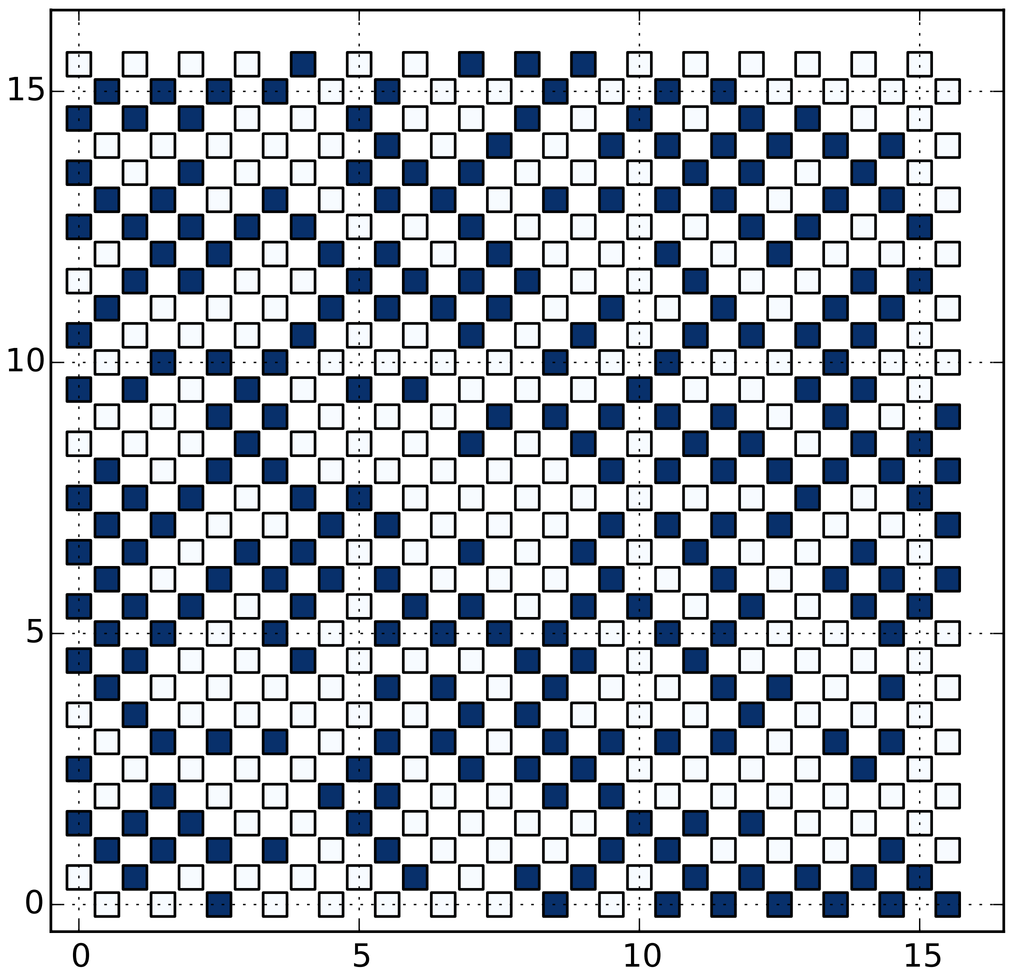


For two configurations



Ground state

?

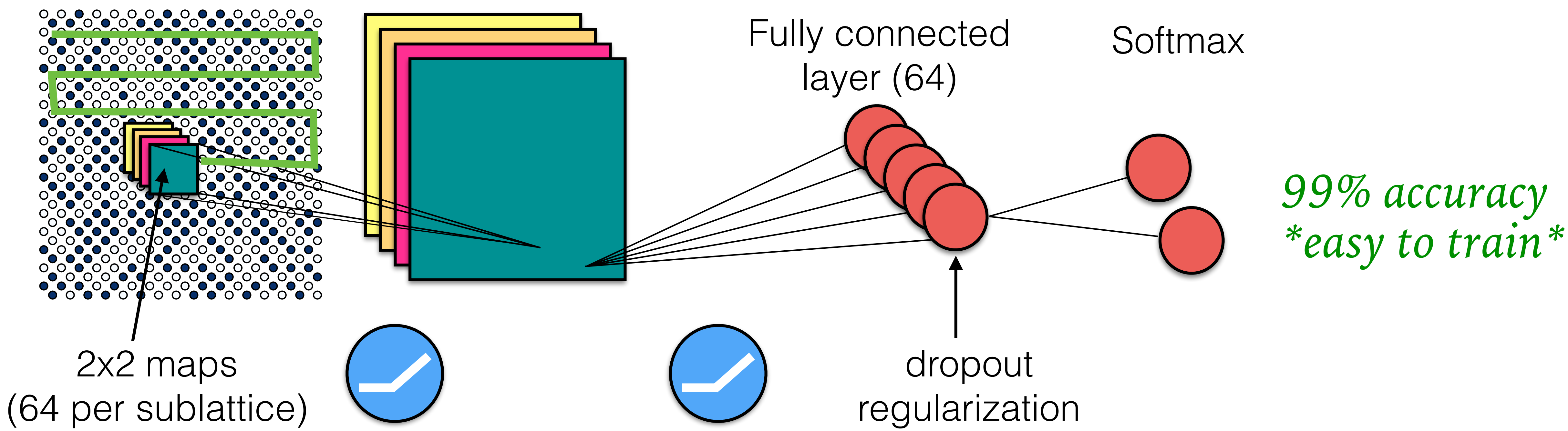


high-temperature state

Feedforward NN are difficult to apply to this problem and lead to 50% accuracy

ISING GAUGE THEORY *F.J. Wegner, J. Math. Phys. 12 (1971) 2259*

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

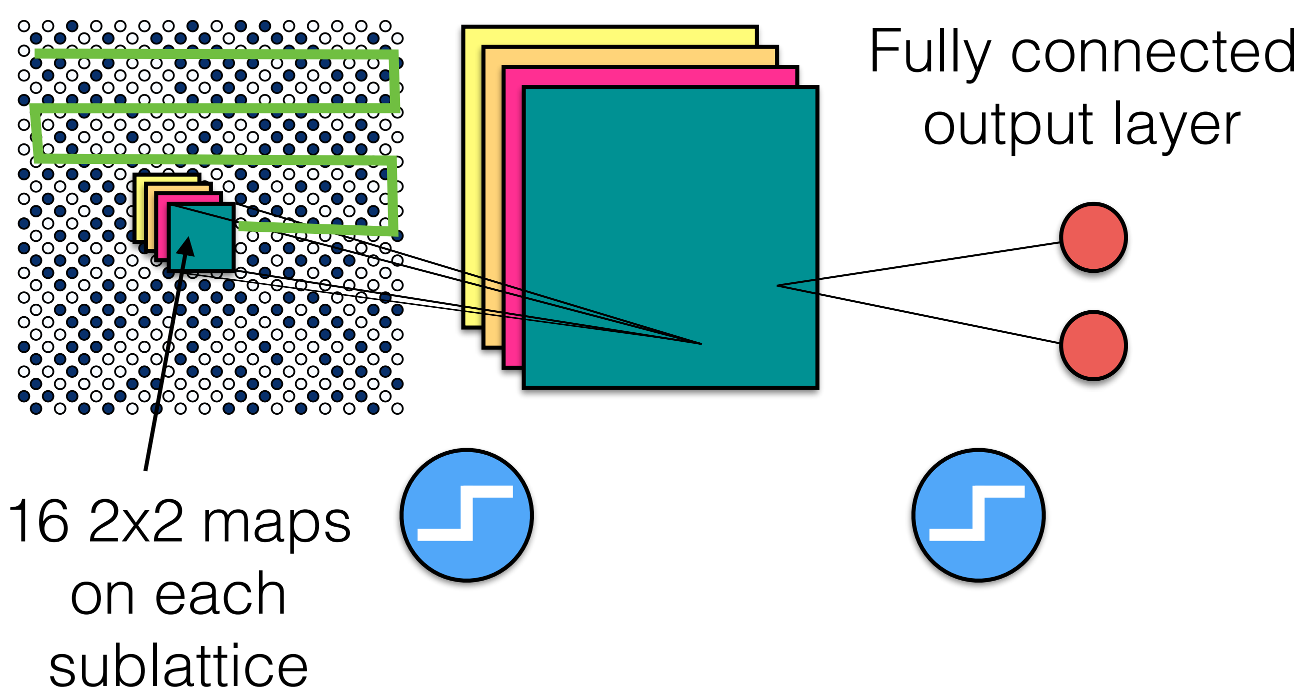


The picture we draw for what the CNN is using to distinguish the phases is that of the detection of satisfied local constraints. In few words, the neural network figures out the energy and uses it to classify states

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

- Based on this observation we derived the weights of a streamlined convolutional network *analytically* designed to work well for this problem:

$$O_{\text{cold}}(\sigma_1, \ldots, \sigma_N) \propto \lim_{\beta \to \infty} \exp \beta J \sum_p \prod_{i \in p} \sigma_i^z$$

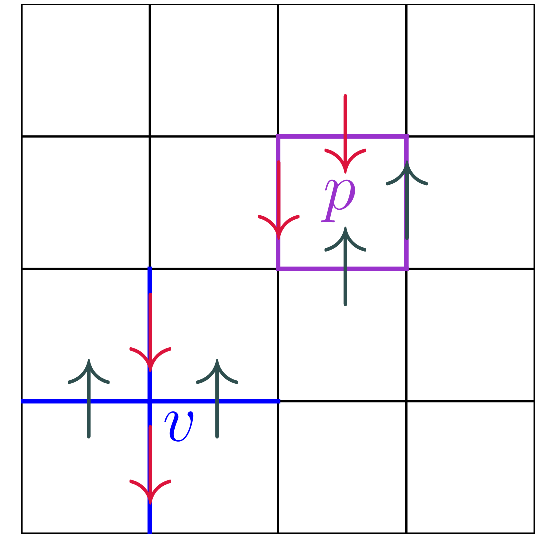


100% accuracy on test sets

J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

KITAEV'S QUANTUM ERROR CORRECTING CODE WITH CONVOLUTIONAL NEURAL NETWORKS

KITAEV'S TORIC CODE GROUND STATE

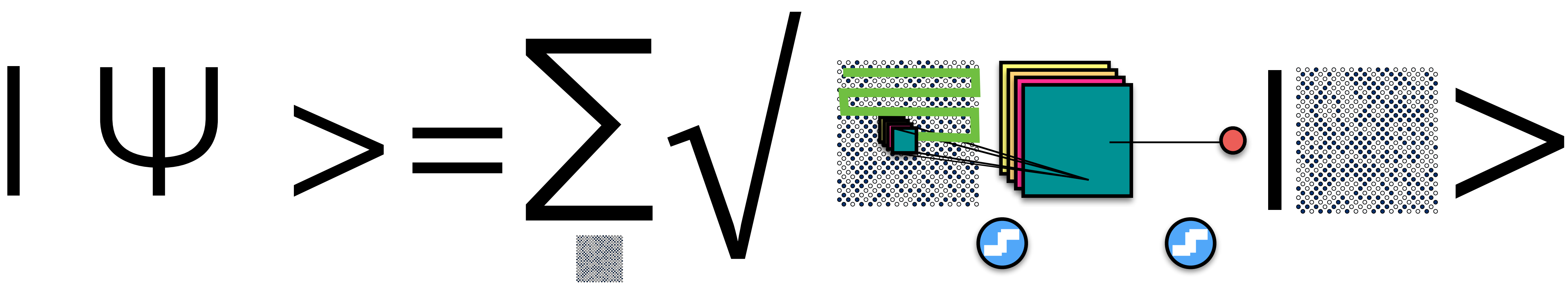


$$H = -J_p \sum_p \prod_{i \in p} \sigma_i^z - J_v \sum_v \prod_{i \in v} \sigma_i^x$$

$$|\Psi_{\mathrm{TC}}\rangle \propto \lim_{\beta \to \infty} \sum_{\sigma_1, \ldots, \sigma_N} e^{\frac{\beta}{2} J \sum_p \prod_{i \in p} \sigma_i^z} |\sigma_1, \ldots, \sigma_N\rangle$$

PEPS : F. Verstraete, M. M. Wolf, D. Perez-Garcia, J. I. Cirac Phys. Rev. Lett. 96, 220601 (2006).

$$O_{\mathrm{cold}}(\sigma_1, \ldots, \sigma_N) \propto \lim_{\beta \to \infty} \exp \beta J \sum_p \prod_{i \in p} \sigma_i^z$$



J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)
Dong-Ling Deng et al Phys. Rev. X 7, 021021 (2017)
Jing Chen, Song Cheng, Haidong Xie, Lei Wang, Tao Xiang arXiv:1701.04831 RBMs

MESSAGES

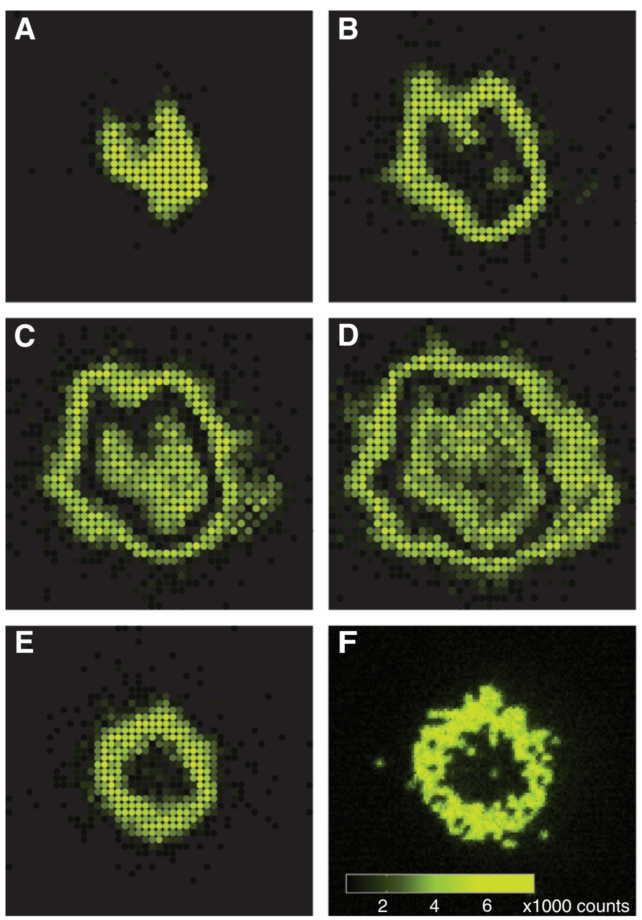
- With a neural network with a small number of parameters we are able to write down analytically the ground state of a system.
- Neural networks seem to enable very good compression quantum many-body states. (analogous to tensor networks)
- No limitations in the dimensionality of the systems
- More importantly, numerical procedures can be constructed to study other systems for where analytical results are elusive.
- Potential applications in materials physics, quantum chemistry, quantum state tomography, etc.

NEURAL-NETWORK QUANTUM STATE TOMOGRAPHY FOR LARGE MANY-BODY SYSTEMS

QUANTUM STATE TOMOGRAPHY

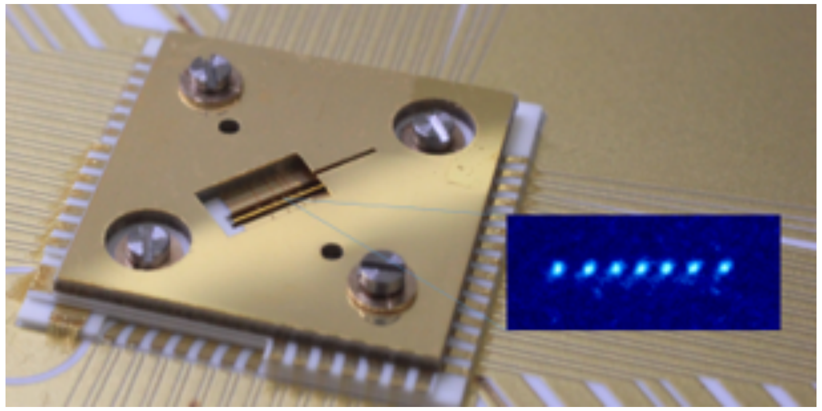
➤ Problem: Can we reconstruct the quantum state of a physical system from a limited set of experimentally accessible set of measurements?

Cold atomic gases



W. Bakr et al, Science (2010)

Trapped ions



ETH Trapped Ion Quantum Information Group

Quantum devices

QST is used as a diagnostic tool in experiments and implementation of technologically relevant quantum algorithms. Measurements required for QST are routinely available in these devices and other systems

THE PROBLEM AND THE REQUIREMENTS OF QST

- Problem: Can we reconstruct the quantum state of a physical system from a limited set of experimentally accessible set of measurements?

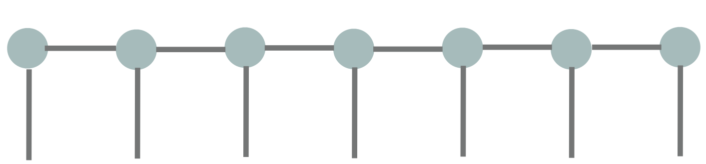
Requirements for QST of large systems (for small systems QST traditionally requires exponential resources)

-Efficient representation of the quantum state: Neural Networks, MPS.

-Set of projective measurements in different bases $|\psi_{\boldsymbol{\lambda},\boldsymbol{\mu}}(\boldsymbol{\sigma}^{[b]})|^2 \simeq P_b(\boldsymbol{\sigma}^{[b]})$

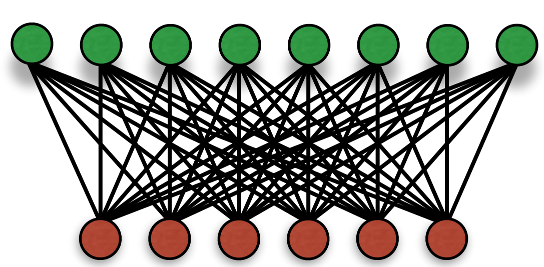
-A learning procedure that makes use of the data to learn the state. It is inherently a big-data problem: Unsupervised learning (maximum likelihood estimation MLE)

MPS QST



Cramer et al, Nat. Comm. (2010)

RBM QST

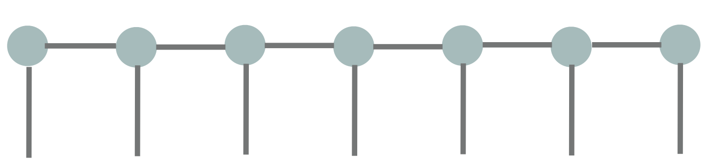


Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

BENEFITS OF RBM QST

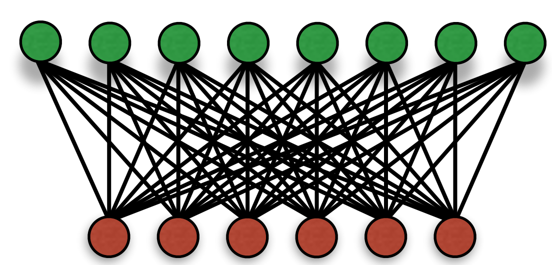
- Suitable for systems in any dimension
- Compact representation of the states
- Availability of a wide set of tools for the evaluation of the results (e.g. log-likelihood on test set, overfitting etc)
- RMBs can encode states with volume-law entanglement*
- Assumption: the state is pure

MPS QST



Cramer et al, Nat. Comm. (2010)

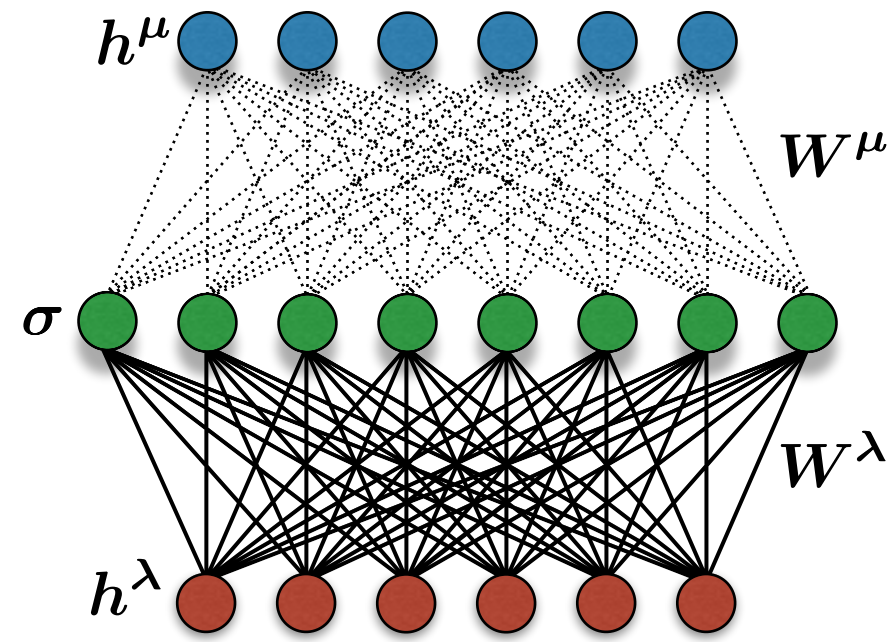
Let's introduce RBM QST



Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

*Dong-Ling Deng et al Phys. Rev. X 7, 021021 (2017)

RESTRICTED BOLTZMANN MACHINE WAVE FUNCTION



RBM probability distribution:

$$p_{\lambda}(\sigma) = e^{\sum_j b_j^{\lambda}\sigma_j + \sum_i \log\left(1+e^{c_i^{\lambda}+\sum_j W_{ij}^{\lambda}\sigma_j}\right)}$$

RBM wavefunction:

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

Widespread use of RBMs to solve many-body physics:

Variational ansatz for quantum wave-functions (Carleo & Troyer, Science 2017)

Exact representation of topological states (Deng, Li & Das Sarma, arXiv 2016)

Accelerate Monte Carlo simulations (Huang & Wang, PRB 2017)

Topological quantum error correction (GT & Melko, PRL)

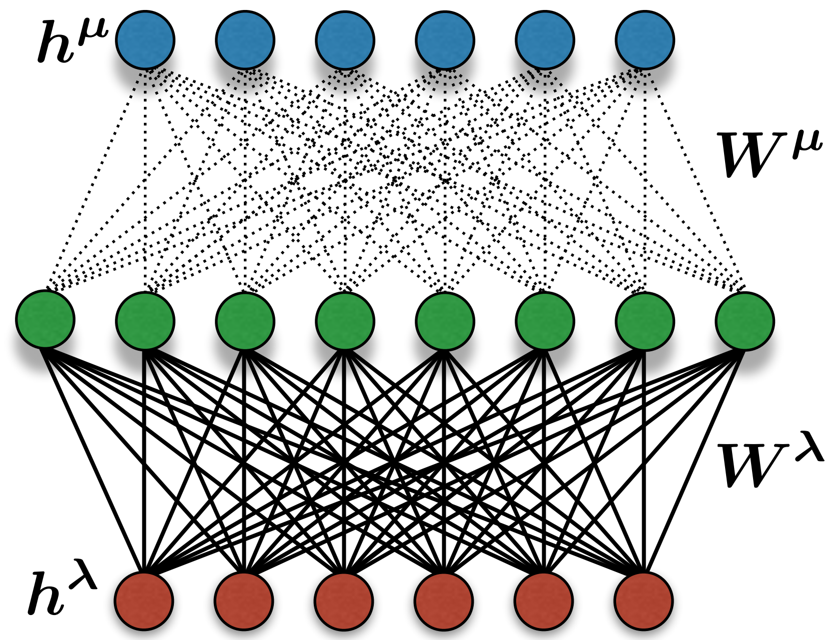
and more . . .

But other choices for the neural network are also possible (CNN, MLP etc)

Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

TRAINING THE RBM

Goal: find $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ such that $|\psi_{\boldsymbol{\lambda},\boldsymbol{\mu}}(\boldsymbol{\sigma}^{[b]})|^2 \simeq P_b(\boldsymbol{\sigma}^{[b]})$



Define the statistical divergence:

$$\Xi(\boldsymbol{\lambda}, \boldsymbol{\mu}) \equiv \sum_{b=0}^{N_B} \mathbb{KL}^{[b]} = \sum_{b=0}^{N_B} \sum_{\{\boldsymbol{\sigma}^{[b]}\}} P_b(\boldsymbol{\sigma}^{[b]}) \log \frac{P_b(\boldsymbol{\sigma}^{[b]})}{|\psi_{\boldsymbol{\lambda},\boldsymbol{\mu}}(\boldsymbol{\sigma}^{[b]})|^2}$$

Minimize the divergence averaged over the datasets in the various bases:

$$\Xi(\boldsymbol{\lambda}, \boldsymbol{\mu}) \simeq -\sum_{b=0}^{N_B} \frac{1}{|D_b|} \sum_{\boldsymbol{\sigma}^{[b]} \in D_b} \log |\psi_{\boldsymbol{\lambda},\boldsymbol{\mu}}(\boldsymbol{\sigma}^{[b]})|^2$$

Stochastic gradient descent:

$$\lambda_j \leftarrow \lambda_j - \eta \, \langle \nabla_{\lambda_j} \Xi(\boldsymbol{\lambda}, \boldsymbol{\mu}) \rangle$$

Natural gradient descent:

$$\mu_j \leftarrow \mu_j - \eta \sum_i \langle S_{ij}^{-1} \rangle \langle \nabla_{\mu_j} \Xi(\boldsymbol{\lambda}, \boldsymbol{\mu}) \rangle$$

W state

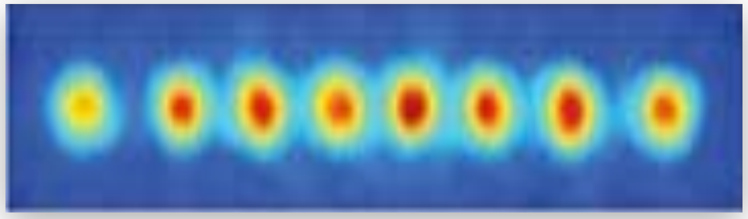
Target wavefunction:

$$|\Psi_W\rangle = \frac{1}{\sqrt{N}}\left(|100\dots\rangle + |010\dots\rangle + \dots |0\dots01\rangle\right)$$

Basis: $\{Z, Z, Z, Z, \dots\}$

$$O^2 = |\langle\psi_{\boldsymbol{\lambda}}|\Psi_W\rangle|^2 = \left\langle \frac{\Psi_W(\boldsymbol{\sigma})}{\sqrt{p_{\boldsymbol{\lambda}}(\boldsymbol{\sigma})}} \right\rangle_{p_{\boldsymbol{\lambda}}} \times \left\langle \frac{\sqrt{p_{\boldsymbol{\lambda}}(\boldsymbol{\sigma})}}{\Psi_W(\boldsymbol{\sigma})} \right\rangle_{|\Psi_W|^2}$$

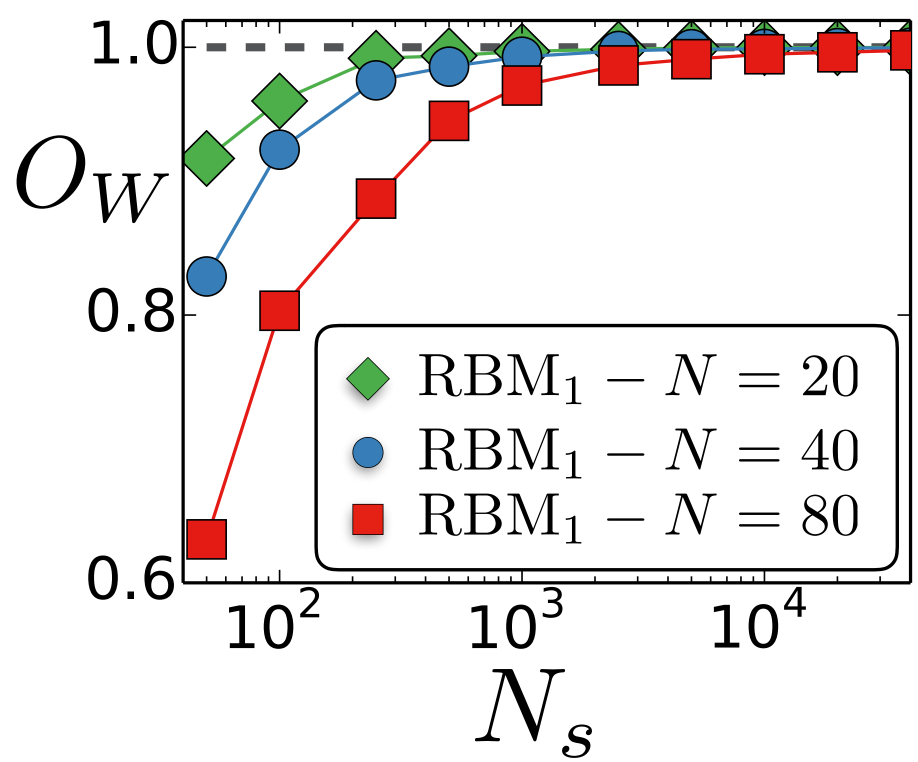
Full tomography

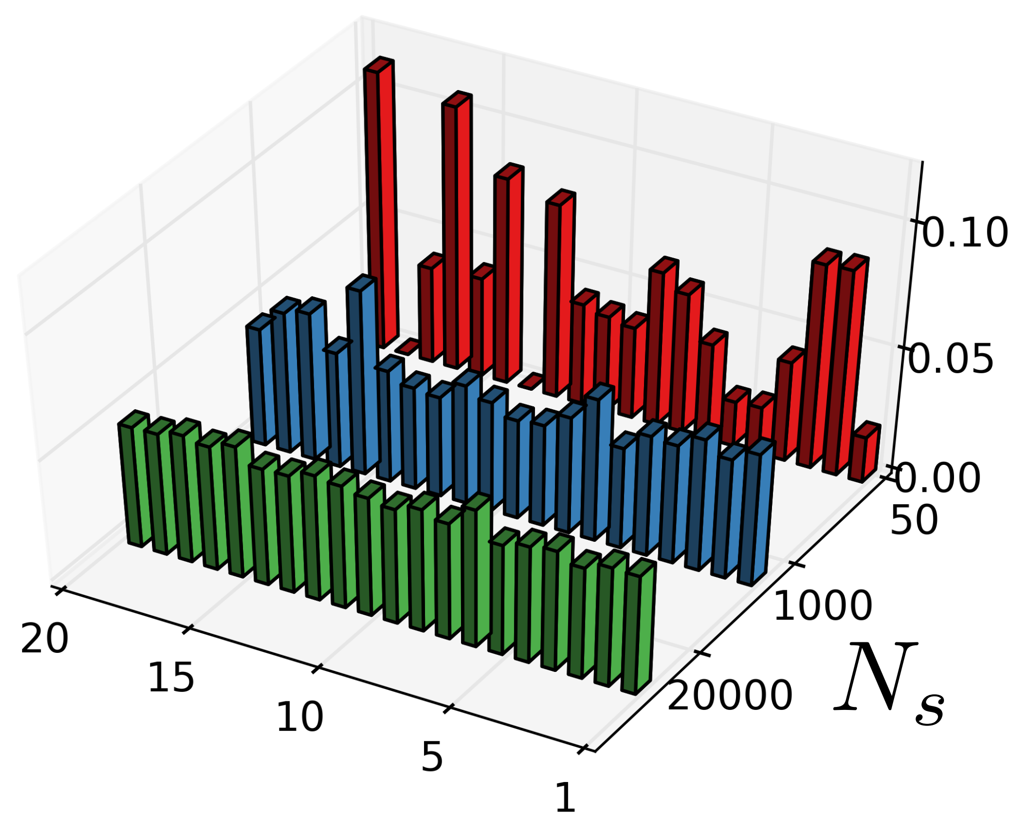


Häffner et al, Nature (2005)
W state trapped ions

MPS tomography

Cramer et al, Nat. Comm. (2010)





Amounts to learning a one-hot encoding vectors of N classes.

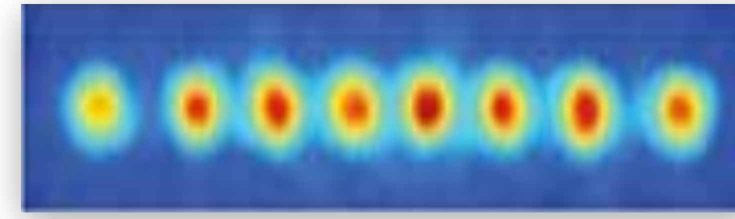
W state with random phases

$$|\tilde{\Psi}_W\rangle = \frac{1}{\sqrt{N}}\left(e^{i\theta_1}|100\ldots\rangle + e^{i\theta_2}|010\ldots\rangle + \cdots + e^{i\theta_N}|0\ldots01\rangle\right)$$

Bases: $\{X, X, Z, Z, \ldots\}, \{Z, X, X, Z, \ldots\}, \ldots$

$\{X, Y, Z, Z, \ldots\}, \{Z, X, Y, Z, \ldots\}, \ldots$

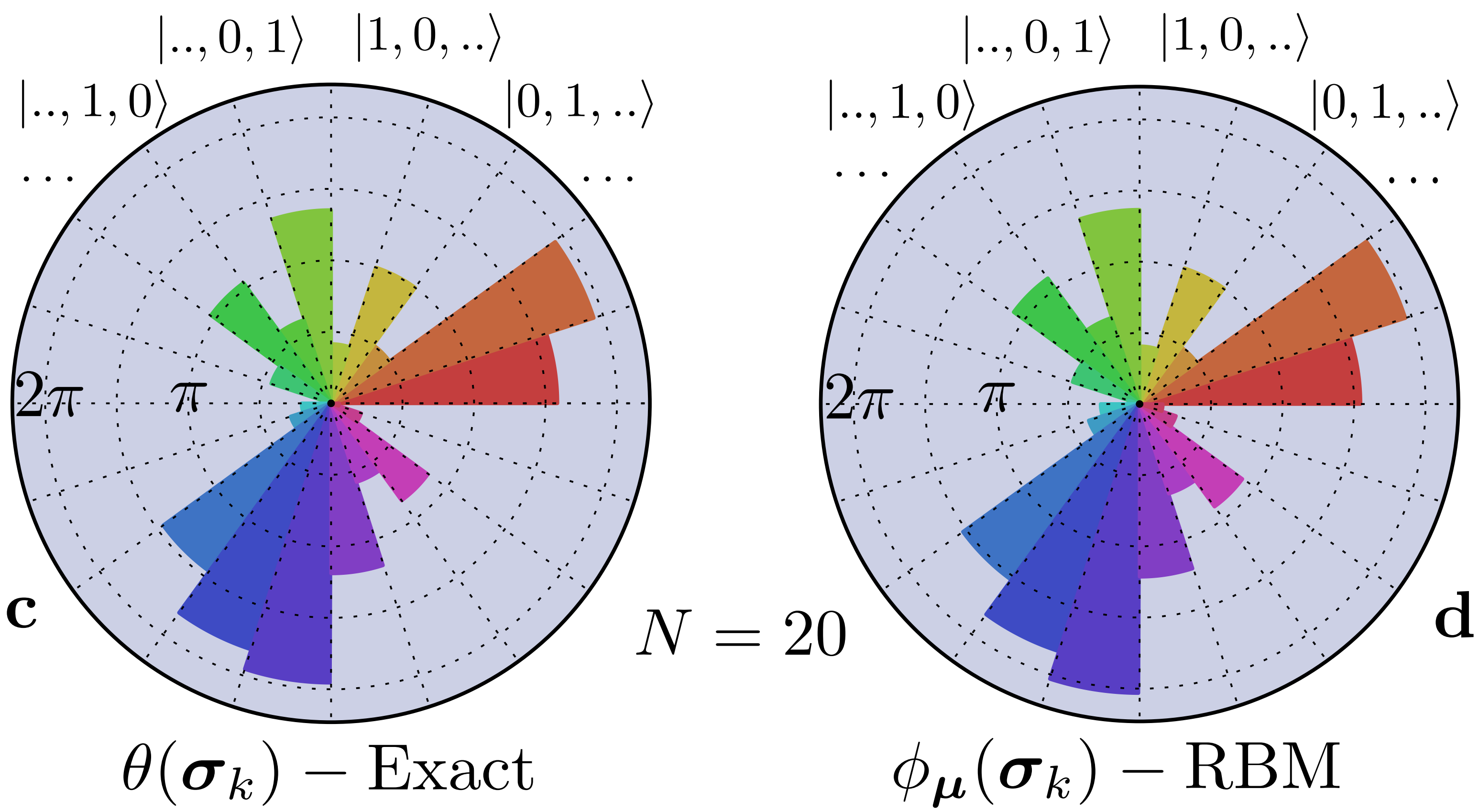
Full tomography



Häffner et al, Nature (2005)

MPS tomography

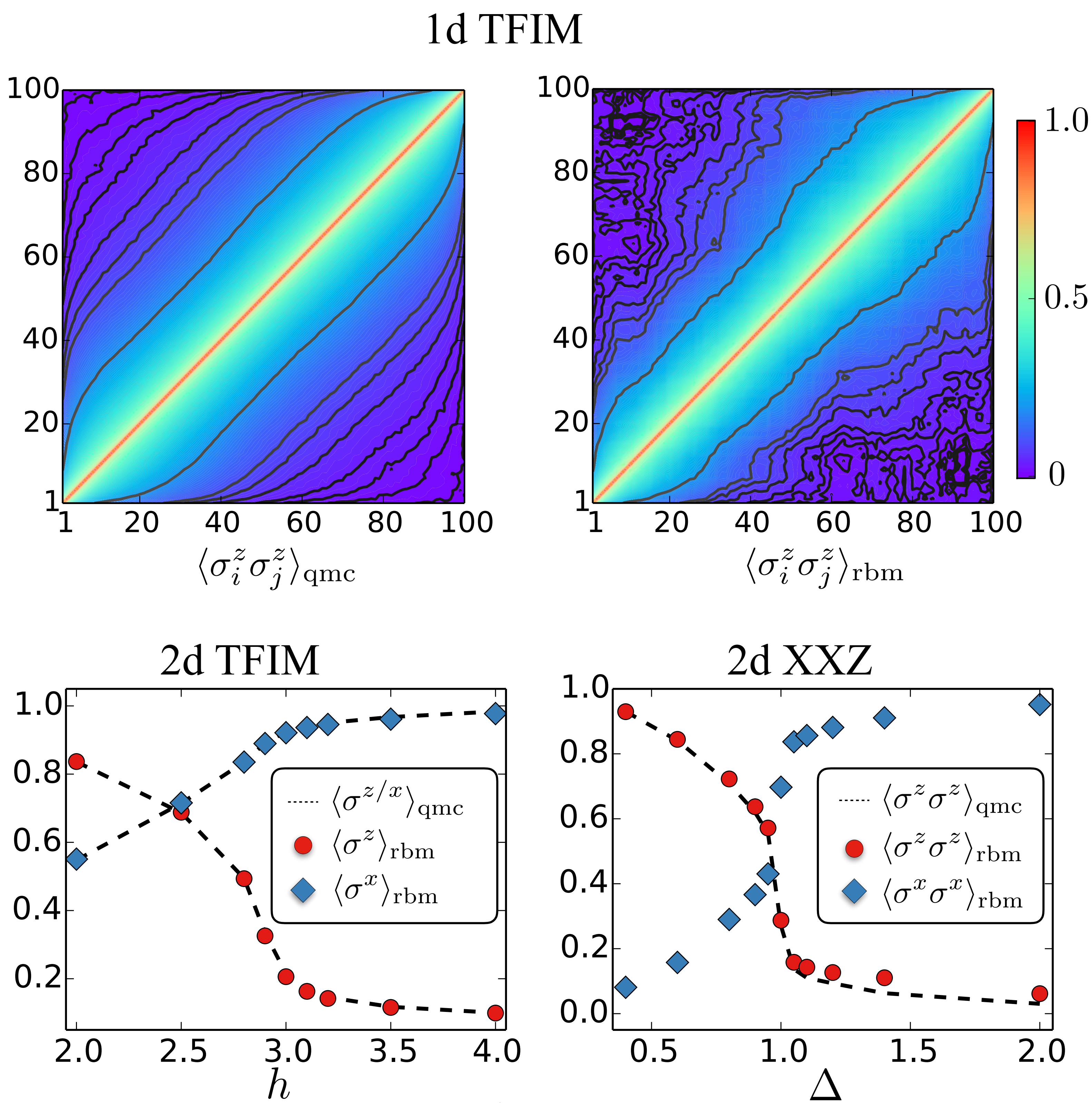
Cramer et al, Nat. Comm. (2010)



Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

Many-body Hamiltonians: ground state

Basis: $\{Z, Z, Z, Z, \ldots\}$



Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

Many-body Hamiltonians: entanglement

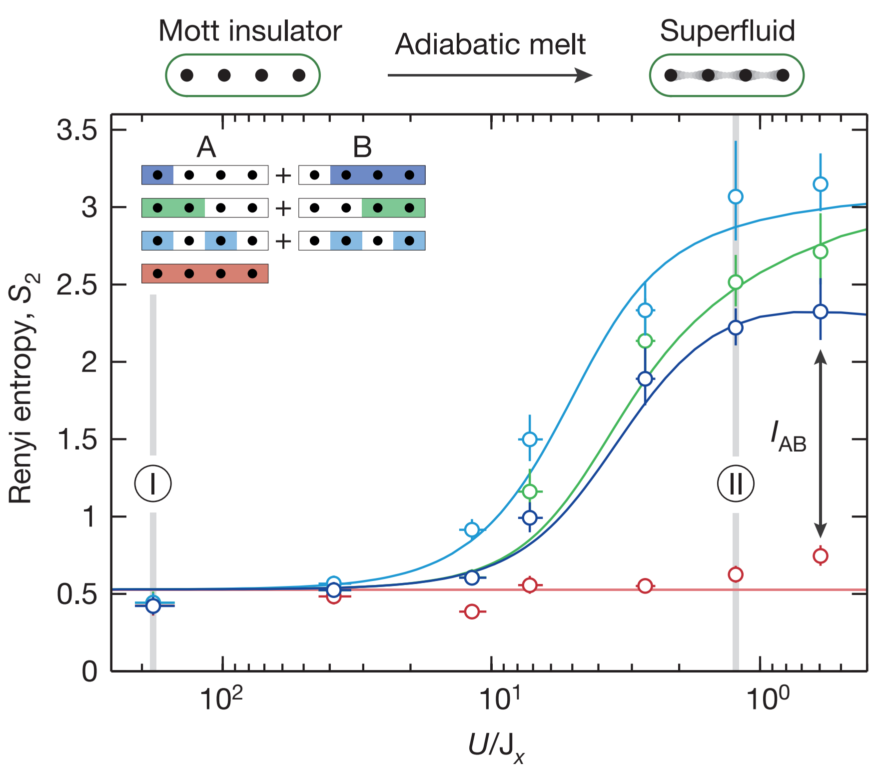
Basis: $\{Z, Z, Z, Z, \ldots\}$

Renyi entropies:

$$S_\alpha(\rho_A) = \frac{1}{1-\alpha}\log(\mathrm{Tr}(\rho_A^\alpha))$$

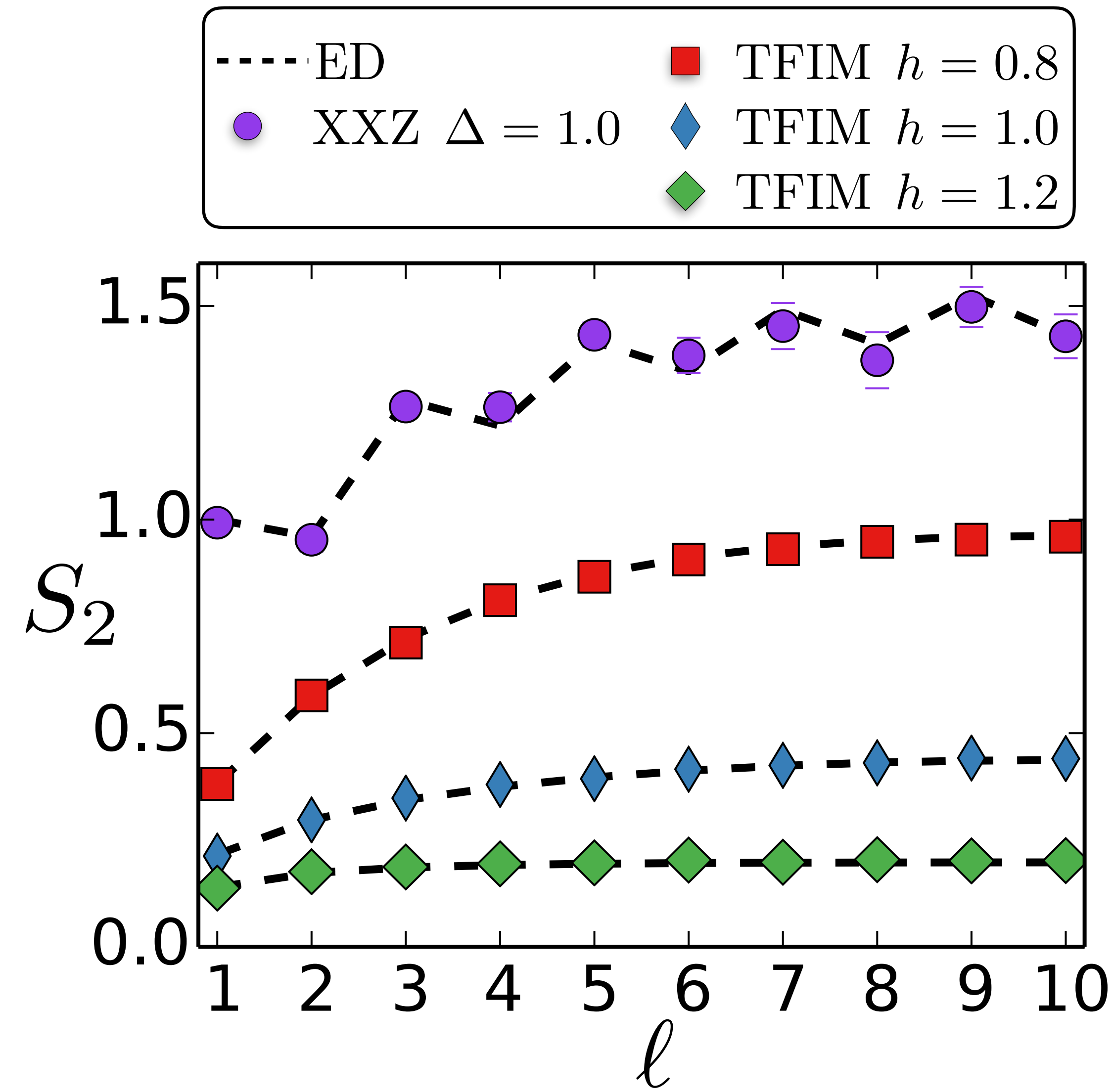
Hastings *et* al, PRL (2010)

$$S_2(\rho_A) = -\log(\langle \mathrm{Swap} \rangle)$$



R. Islam et al, Nature (2015) (4 particles)

$N = 20$



For experiments: unequivocal evidence of the quantum behaviour and some of the most profound and striking manifestations of quantum behaviour.

For RBMs: Shows that the learned RBMs generalize very well

Hard to get experimentally. Our approach suggests an experimentally viable way to do it.

Torlai, Mazzola, Carrasquilla, Troyer, Melko and Carleo 1703:05334

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.
- We have performed QST based on neural networks with results that are better than the state-of-the-art and enable the study of 2- and 3-dimensional quantum systems

Come Join Us at Vector Institute!

Research Scientists & Postdoctoral Fellows

The Vector Institute is building its team of Research Scientists and Postdoctoral Fellows who will have opportunities to contribute by:

- working and collaborating with other members of the Institute
- engaging in state-of-the-art research
- publishing at the highest international level
- contributing to the academic life and reputation of the Institute
- determining strategic areas of research

Vector also encourages its team to:

- develop entrepreneurial activities and interact with industry
- teach at summer schools offered by the Institute
- teach graduate and undergraduate courses at a university

Students

Students affiliated with the Vector Institute will be full-time graduate students at an affiliated Canadian university. The Vector Institute is not a degree-granting institution. Vector will partner with universities towards a goal to graduate 1,000 professional applied Master’s students in AI-related fields in Ontario per year within five years.

Faculty Members & Research Scientists

Geoffrey Hinton
Chief Scientific Advisor

Richard Zemel
Research Director

Learn more:

- http://vectorinstitute.ai/#people
- News Release: Vector Institute Doubles AI Faculty

Jimmy Ba

Juan Carrasquilla

David Duvenaud

Murat Erdogdu

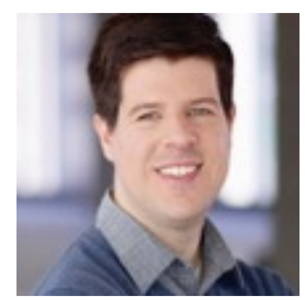
Sanja Fidler

David Fleet

Brendan Frey

Marzyeh Ghassemi

Anna Goldenberg



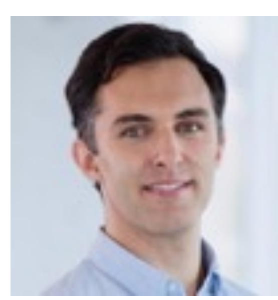
Roger Grosse

Alireza Makhzani

Quaid Morris

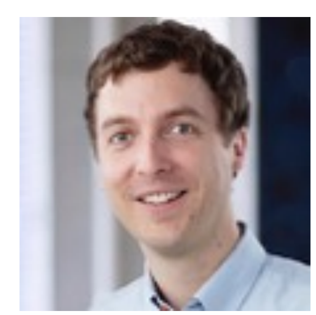
Sageev Oore

Pascal Poupart

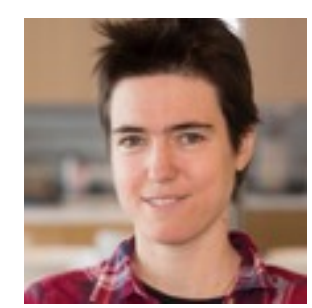


Daniel Roy

Frank Rudzicz



Graham Taylor



Raquel Urtasun

Neural-network quantum state tomography

The problem

How can we machine learn a quantum state $\Psi(\boldsymbol{\sigma})$ from a set of measurements?

We target a generic quantum state in a reference basis $\{\boldsymbol{\sigma}\}$:

$$|\Psi\rangle = \sum_{\boldsymbol{\sigma}} \Psi(\boldsymbol{\sigma})|\boldsymbol{\sigma}\rangle = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{i\theta(\boldsymbol{\sigma})}|\Psi(\boldsymbol{\sigma})|\,|\boldsymbol{\sigma}\rangle$$

Learning the phases

Perform the measurements in the rotated basis:

X rotation on site j :

$$\Psi(\sigma_0^z,\dots,\sigma_j^x,\dots) = \Psi(\sigma_0^z,\dots,\sigma_j^z=0,\dots)+(1-2\sigma_j^x)\Psi(\sigma_0^z,\dots,\sigma_j^z=1,\dots)$$

$$P(\sigma_0^z,\dots,\sigma_j^x,\dots) \sim \cos\big(\theta(\sigma_j^z=0)-\theta(\sigma_j^z=1)\big)$$

Y rotation on site j :

$$\Psi(\sigma_0^z,\dots,\sigma_j^y,\dots) = \Psi(\sigma_0^z,\dots,\sigma_j^z=0,\dots)+i(1-2\sigma_j^y)\Psi(\sigma_0^z,\dots,\sigma_j^z=1,\dots)$$

$$P(\sigma_0^z,\dots,\sigma_j^y,\dots) \sim \sin\big(\theta(\sigma_j^z=0)-\theta(\sigma_j^z=1)\big)$$