東京大学
THE UNIVERSITY OF TOKYO

Topological Materials Science

YITP
YUKAWA INSTITUTE FOR THEORETICAL PHYSICS

University of Tokyo

Nobuyuki Yoshioka

Collaborators: Yutaka Akagi, Hosho Katsura

arXiv: 1709.05790

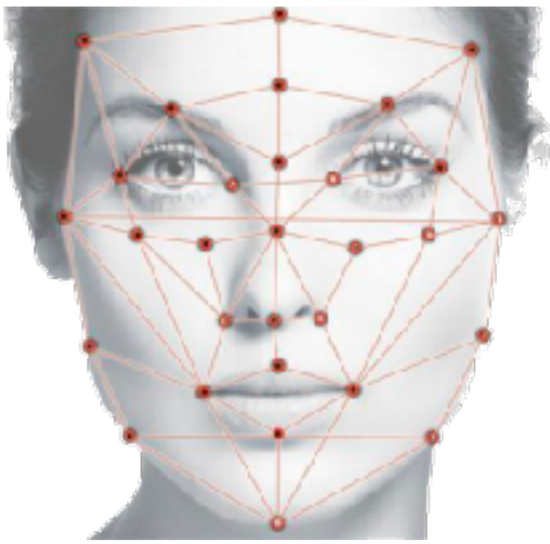
LEARNING DISORDERED TOPOLOGICAL PHASES BY STATISTICAL RECOVERY OF SYMMETRY

- **Introduction**
  - Objective of Machine Learning
  - Application to Physics
- **Method and Hamiltonian**
  - Problem set up
  - Classification by Artificial Neural Network
- **Result and Discussion**

- **Introduction**
  - Objective of Machine Learning
  - Application to Physics
- **Method and Hamiltonian**
  - Problem set up
  - Classification by Artificial Neural Network
- **Result and Discussion**

Machine Learning in Ordinary Life

Image recognition



UC Berkeley Computer Vision Group

Machine translation

Google translation

Well understanding on non-metamessage.

Triumph of Go/Shogi AI in 2017


ARTICLE

AlphaGo

doi:10.1038/nature24270

Mastering the game of Go without human knowledge

David Silver[1*], Julian Schrittwieser[1*], Karen Simonyan[1*], Ioannis Antonoglou[1], Aja Huang[1], Arthur Guez[1], Thomas Hubert[1], Lucas Baker[1], Matthew Lai[1], Adrian Bolton[1], Yutian Chen[1], Timothy Lillicrap[1], Fan Hui[1], Laurent Sifre[1], George van den Driessche[1], Thore Graepel[1] & Demis Hassabis[1]

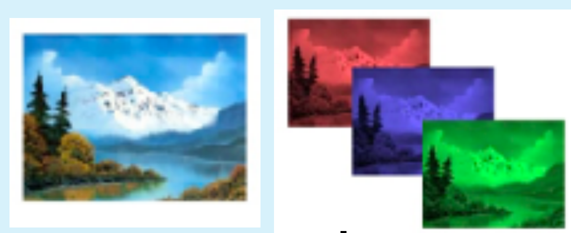
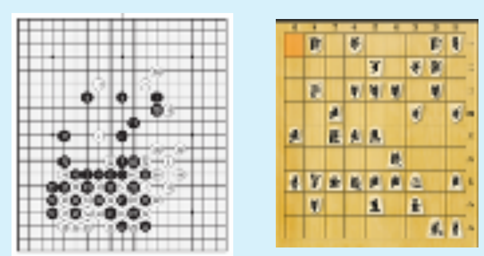
DeepMind group, Nature 550, 354 (2017).

Denou Sen Website
http://denou.jp/2017/

Element-wise understanding of ML

Machine Learning = Computer algorithm that gives prediction/knowledge from huge amount of data beyond human resources.

- **Machine task** = Construction of highly-nonlinear function.
  Recent progress: discovery of convolutional NN, ResNet etc.

	Input		Output
✓ Image recognition	 RGB values	⇨	Possibility for a label e.g.) mountain, tree, pipe
✓ Go/Shogi	 Piece configuration	⇨	1. Next move suggestion 2. Current win probability

- **Learning task** = Optimization of *"the Loss function"* i.e. the "performance" of the machine

e.g.) For data $\mathbf{x}$, label y, and some parametrized classifier $\mathcal{F}_p$, take

$$\mathcal{L}(\mathbf{x}) = |\mathcal{F}_p(\mathbf{x}) - y|$$

prediction error

and update $p \to p - \eta \partial_p \mathcal{L}$ in a stochastic manner.

Application to Physics

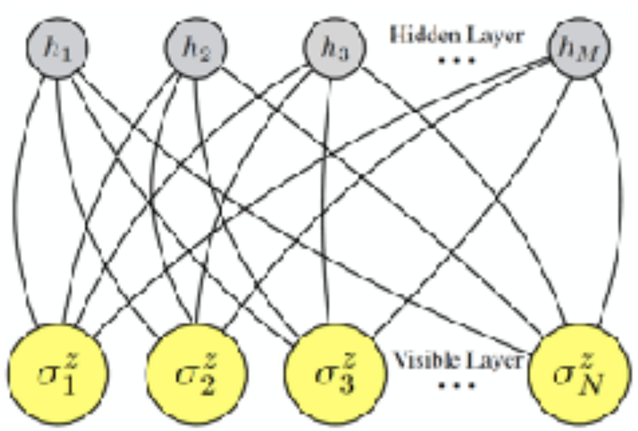
Three main approaches for <ML|cond-mat>

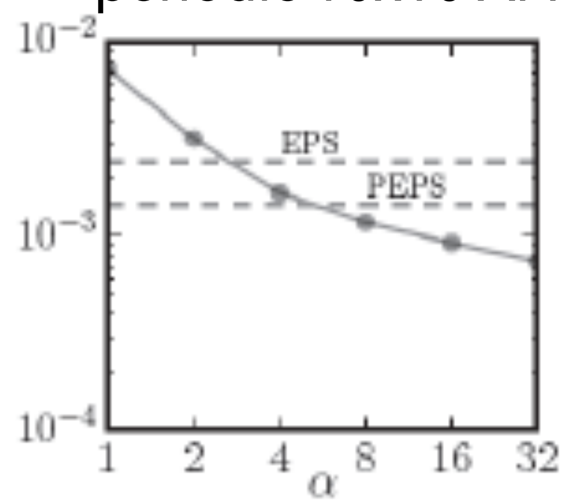
1. Neural network as wave func. ansatz

- Quantum many-body solver

Carleo&Troyer Science 355('17)
Nomura etal. arXiv:1709.06475

GS energy error of periodic 10x10 AFH





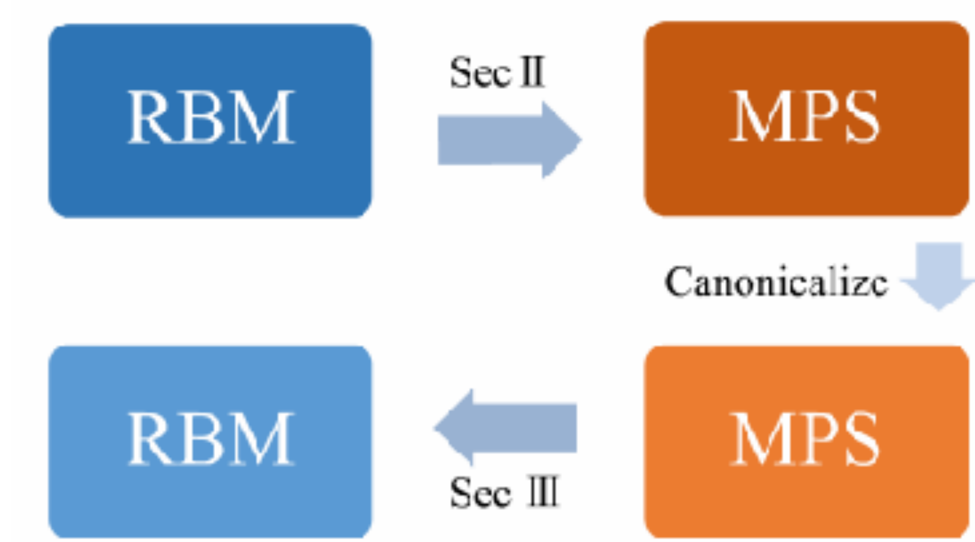
Restricted Boltzmann Machine (RBM)

$$\Psi_M(\mathcal{S};\mathcal{W}) = \sum_{\{h_i\}} e^{\sum_j a_j\sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$

Glasser etal. arXiv:1710.04045 ('17)
Clark arXiv:1710.03545 ('17)

- RBM <–> Tensor network

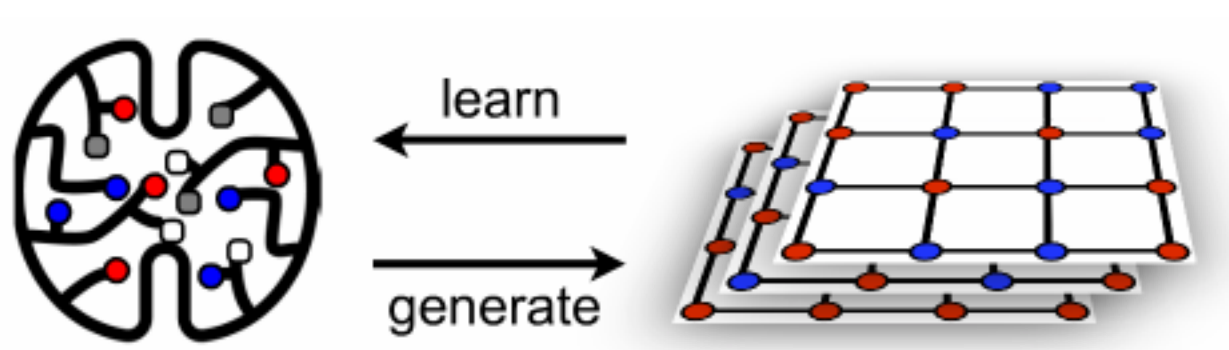
Chen etal arXiv:1701.04831 ('17)



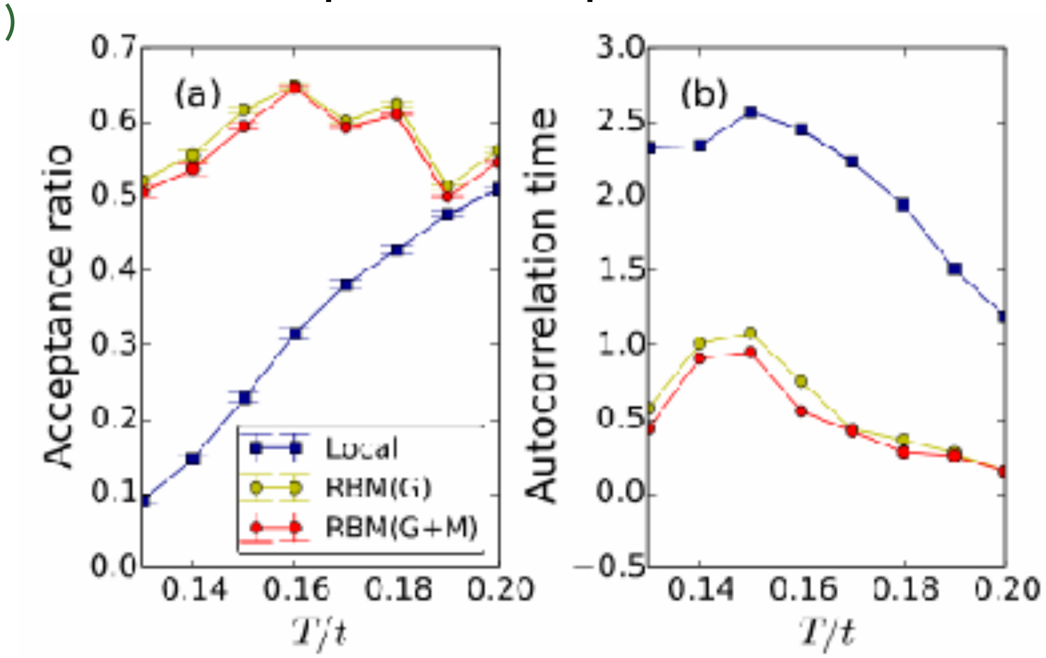
2. Speeding up Monte Carlo

Wang arXiv:1702.05856('17)
Huang&Wang PRB 95('17)

Efficient cluster-update by learning thermal distribution



Falicov-Kimball model, 8x8 periodic square lattice



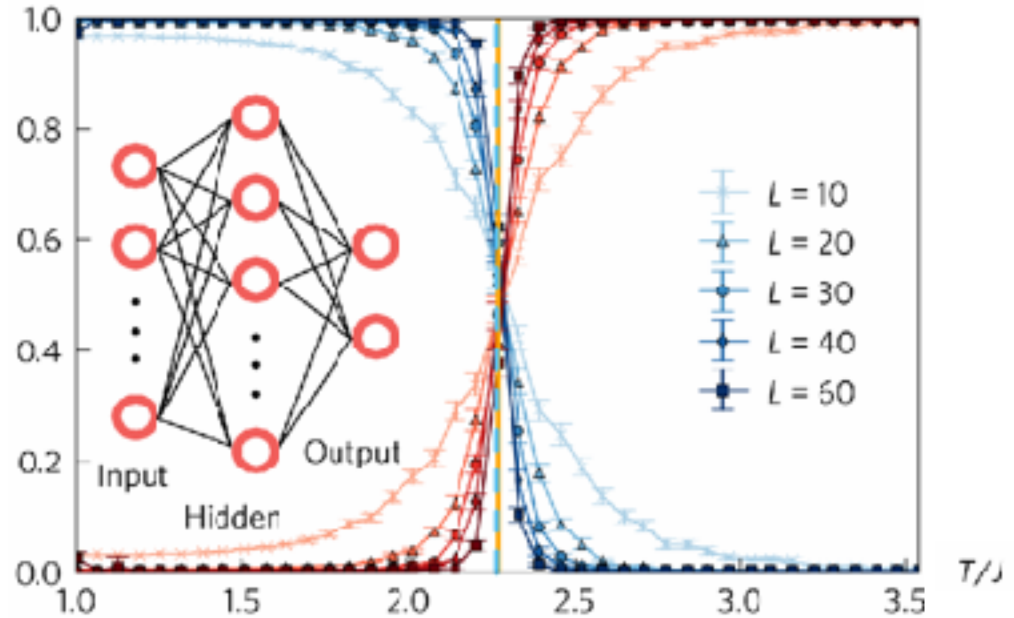
Three main approaches for <ML|cond-mat>

3. Classification of phases

Tanaka&Tomiya JPSJ 86 ('17) Broecker etal. SciRep 7 ('17)
Ch'ng etal. PRX 7 ('17) Zhang&Kim PRL 118 ('17)

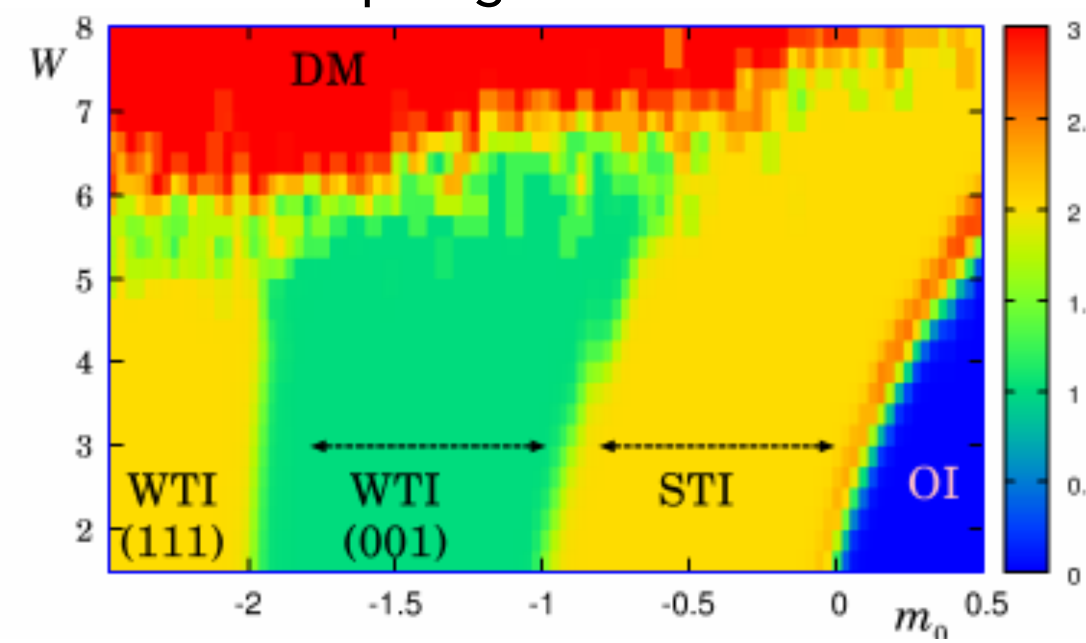
- Supervised learning : <u>Teach the pattern. Then let it predict.</u>

Possibility of low/high-temp. phase in 2D square Ising w/ Total Sz = 0



Carrasquilla&Melko NatPhys 13 ('17)

Possibility of disordered phases in topological insulator



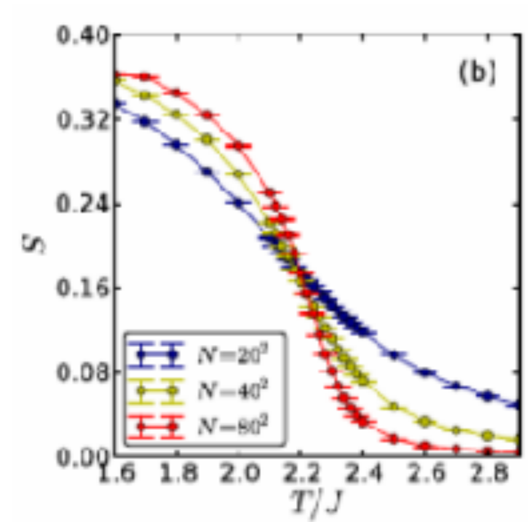
Ohtsuki&Ohtsuki JPSJ 85('16), 86('17)

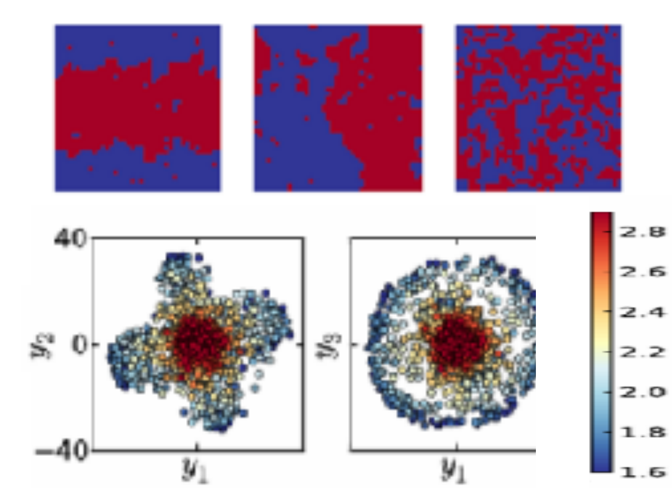
- Data science methods e.g. PCA, VAE, tSNE

L. Wang PRB 94 ('16), S. Wetzel PRE 96 ('17)
Ch'ng etal. ('17),

Extract patterns of spin configurations

Learning transition point *<u>without</u>* teaching the notion of "phase".





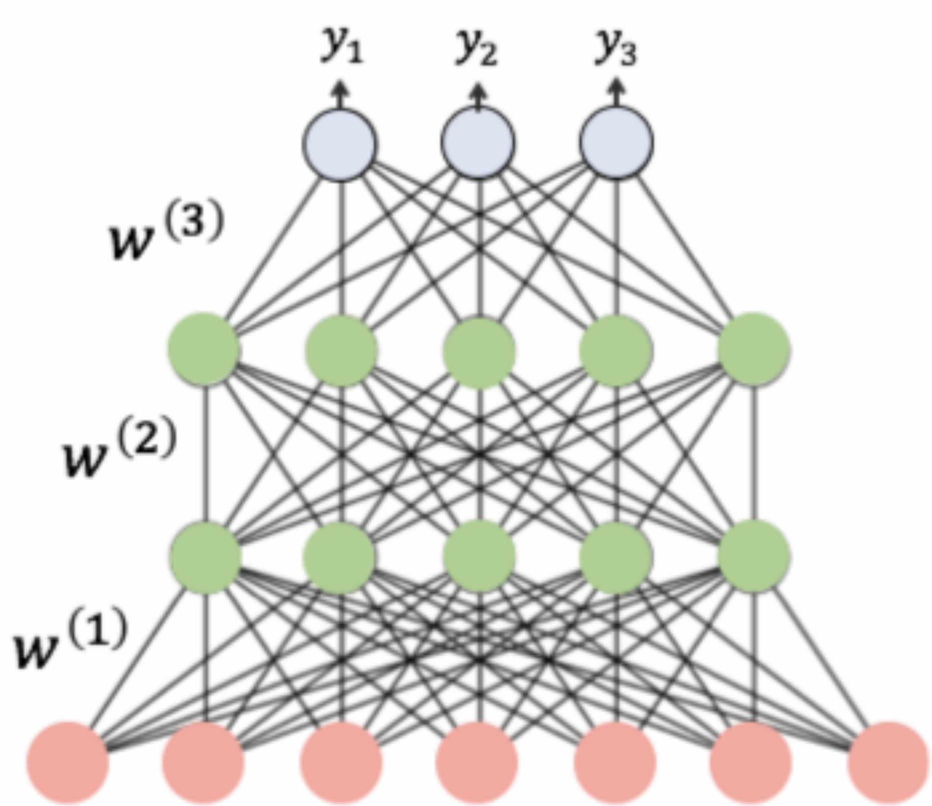
Today’s talk

- **Learn the clean, classify the disordered by Neural Network**

Goal Phase classification of disordered TSC (e.g. 2d, class DIII)

Obstacle Break down of well-known formulae in lattice system

Solution Use of NN.
Trained merely at clean limit, disordered phases classified correctly.
Phases not present at clean limit also correctly detected.



Output : Probability of the corresponding phases

Hidden Layers : Extract the feature of the data

Input : Statistical average of quasiparticle distribution

Introduction

Objective of Machine Learning

Application to Physics

Method and Hamiltonian

Problem set up

Classification by Artificial Neural Network

Result and Discussion

2d Class DIII Topological Superconductor

- **Class DIII in the "Periodic table"**

Schnyder etal. PRB 78 ('08)
Kitaev AIP Conf. Proc. 1134 ('09)

Symmetry				Dimension		
CAZ	TRS	PHS	CHS	1	2	3
A	0	0	0		$\mathbb{Z}$ (IQHE)	
AIII	0	0	1	$\mathbb{Z}$		$\mathbb{Z}$
D	0	+1	0			
DIII	−1	+1	1	$\mathbb{Z}_2$	$\mathbb{Z}_2$ (Today's topic)	$\mathbb{Z}$
AII	−1	0	0		$\mathbb{Z}_2$ (QSHE)	$\mathbb{Z}_2$ (3D TI)

- BdG system with TRS, w/o SU(2) symm.

TRS: $\Theta^{-1} H_{\mathbf{k}} \Theta = H_{-\mathbf{k}}, \Theta^2 = -1$

PHS: $\Xi^{-1} H_{\mathbf{k}} \Xi = -H_{-\mathbf{k}}, \Xi^2 = 1$

CHS: $\Gamma^{-1} H_{\mathbf{k}} \Gamma = -H_{\mathbf{k}}, \Gamma = i\Xi\Theta$

- Majorana edge mode as topo. feature

e.g.) $Cu_xBi_2Se_3$, Hasan group('10)

quasi-2D, surface modes

- **BdG Hamiltonian in k-space, square-lattice**

Sato&Fujimoto PRB 79('09)
Diez etal NewJPhys 16('14)

$$H_{\mathbf{k}} = \begin{pmatrix} \hat{\epsilon}_{\mathbf{k}} & \hat{\Delta}_{\mathbf{k}} \\ \hat{\Delta}^\dagger_{\mathbf{k}} & -\hat{\epsilon}_{-\mathbf{k}} \end{pmatrix}$$

$$\hat{\epsilon}_{\mathbf{k}} = 2t(\cos k_x + \cos k_y) - \mu$$

Kinetic — NN hopping — chemical pot.

$$\hat{\Delta}_{\mathbf{k}} = i\sigma_y \left(\Delta_p (\mathbf{d}_{\mathbf{k}} \cdot \boldsymbol{\sigma}) + \Delta_s \sigma_0\right)$$

$\mathbf{d}_{\mathbf{k}} = (\sin k_x, -\sin k_y, 0)$

Pairing — helical p-wave — s-wave

- Z2 invariant

Kane&Mele PRL 95('05)
Fu&Kane PRB 76('07)

Z_2 inv. — TRIM

$$(-1)^{\nu} = \prod_{i=1}^{4} \frac{\mathrm{Pf}[\omega(\Lambda_i)]}{\sqrt{\det \omega(\Lambda_i)}}$$

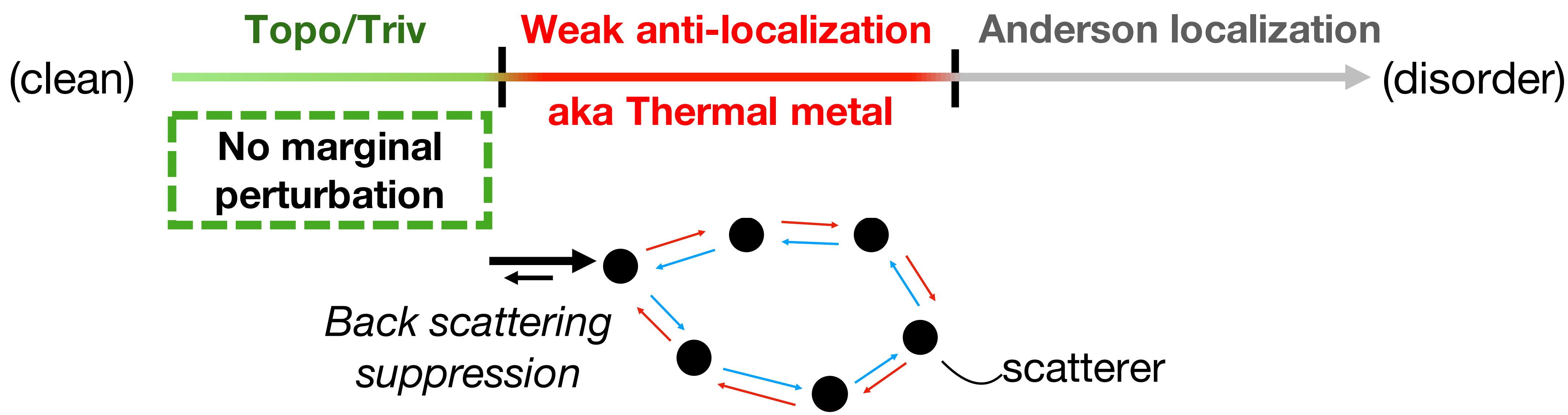
"Sewing matrix"

$$\omega_{\alpha\beta}(\mathbf{k}) = < u_\alpha(-\mathbf{k}) | \Theta | u_\beta(\mathbf{k}) >$$

2d Class DIII Topological Superconductor II: Dirty

- **Phases by intermediate disorder (NLSM analysis)**

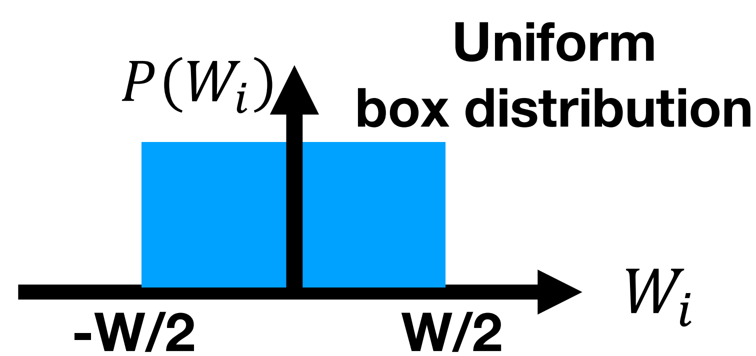
Hikami PRB 24 ('81)
Evers&Mirlin RevModPhys 80 ('08)



- **In lattice model…? For example,**

on-site randomness

$$H_W = \sum_{i,\sigma} W_i c^\dagger_{i,\sigma} c_{i,\sigma}$$



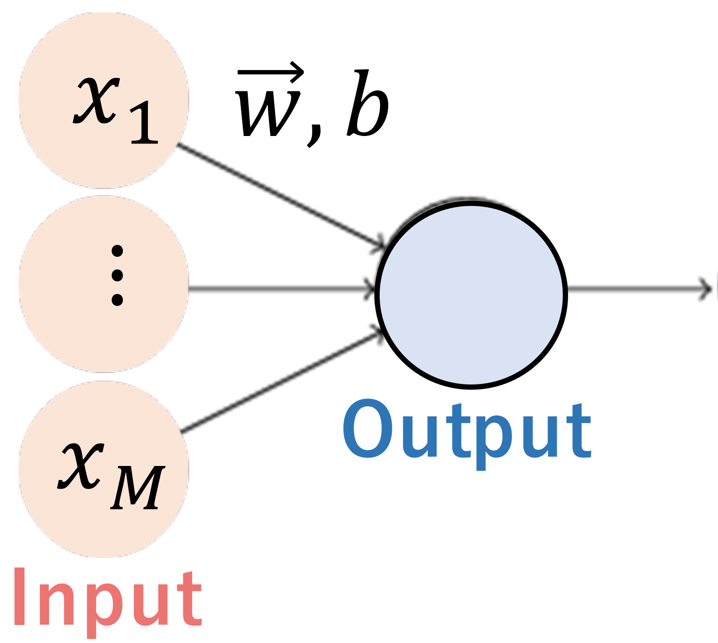
➡ Break down of top.inv. formulae (e.g. Kane-Mele, Niu-Thouless-Wu)…

New approach introduced. NN.

Classification by ANN

Single Neuron (perceptron)

Rosenblatt PsycoRev 65('58)



1. weighing -> linear combination of input

$$y = \sigma(\vec{w} \cdot \vec{x} + b)$$

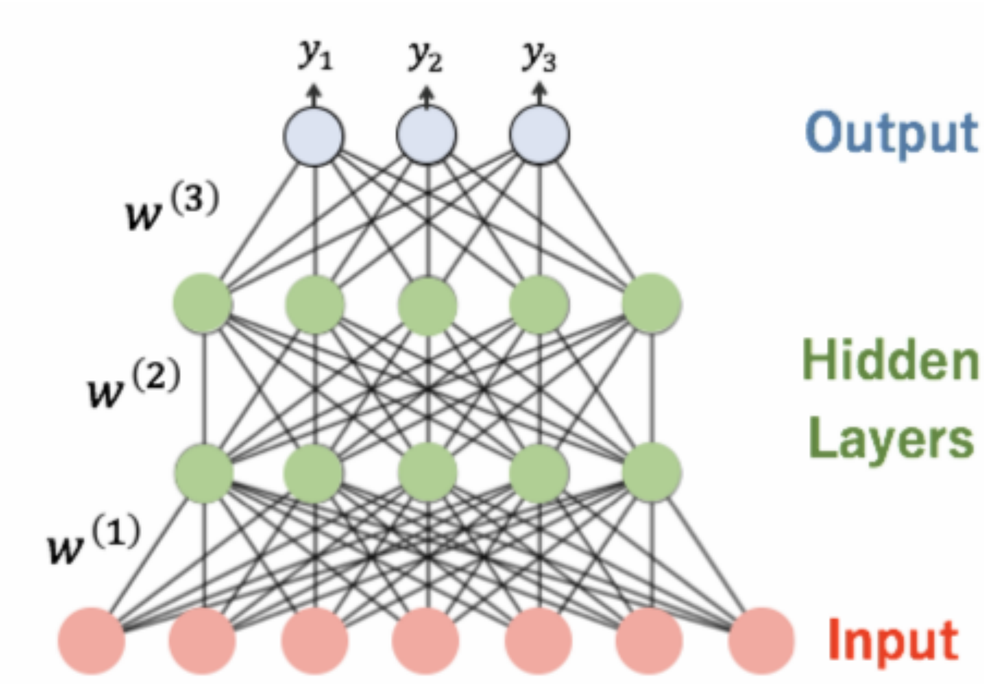
2. activation -> operate nonlinear function

e.g.)

$$\sigma(z) = \begin{cases} 1/(1+e^{-z}) & \text{(Sigmoid)} \\ max(0, z) & \text{(ReLU)} \end{cases}$$

Deep Neural Network

Rumelhart etal. Nature 323('86)
Hinton etal. Science 313 ('06)



$$\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^m$$

- **Weighing** and **activation** sequentially/simultaneously. ***Hidden Layers*** extracts abstract feature efficiently.
- Output:

$$y_i = e^{-z_i} / (\sum_j e^{-z_j}) \qquad \sum_i y_i = 1$$

probability

- Universal approximation theorem for multilayer NN

→ **Expression of any nonlinear function**

Cybbenko MathCon 2('89)
Hornik etal. Neural Network 2('89)

Supervised Learning

… Tune parameters by minimizing the "distance" btw output and correct label

$$\mathcal{W}^{(i)}_{j,k} \to \mathcal{W}^{(i)}_{j,k} - \eta \left(\partial \mathcal{L} / \partial \mathcal{W}^{(i)}_{j,k} \right)$$

Loss = Cross entropy + L2 regularization:

$$\mathcal{L}(w) = - \sum_{j=1}^{(\#\text{data})} \sum_{k=1}^{(\#\text{class})} \hat{y}^{(k)}_j \log y^{(k)}_j (\mathbf{x}_j; \mathbf{w}) / (\#\text{data})$$

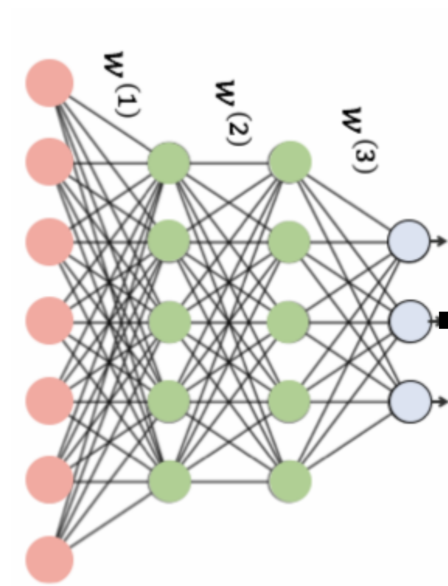
$$+ \lambda \sum_{i=1}^{(\#\text{layers})} |W^{(i)}|^2 .$$

(improves the generalization power)

Classification by Artificial Neural Network

- **Image -> Probability**

RGB per pixel



Probability/Confidence

Krizhevsky etal. ILSVRC 2012

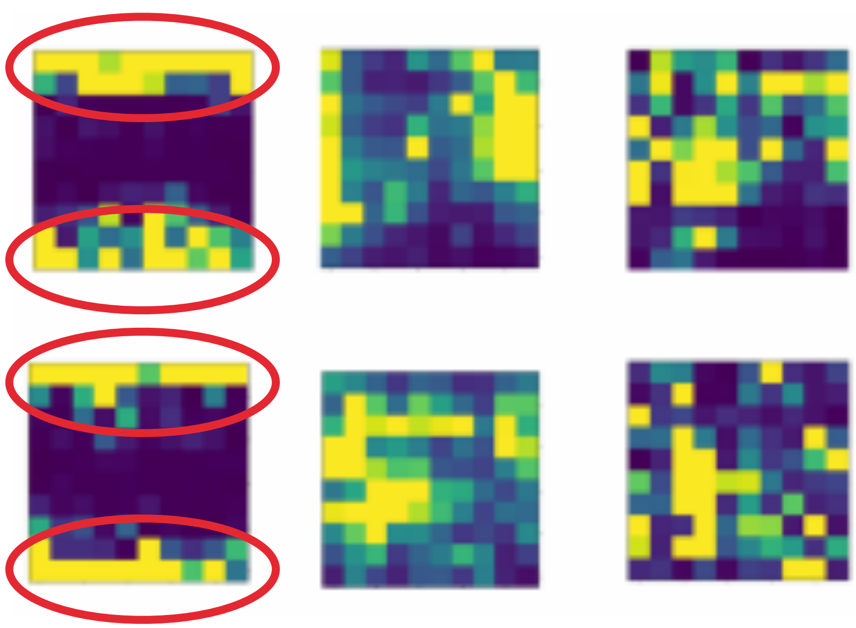
Classification by Artificial Neural Network

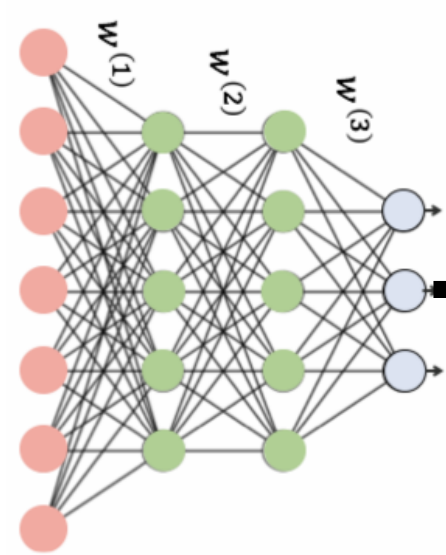
- **Quasiparticle dist. -> Probability**

$$H|\Psi_n> = E_n|\Psi_n>$$

Take first excitation state

$$P(\mathbf{r}) = |\psi^e_\uparrow(\mathbf{r})|^2 + |\psi^e_\downarrow(\mathbf{r})|^2 + |\psi^h_\uparrow(\mathbf{r})|^2 + |\psi^h_\downarrow(\mathbf{r})|^2$$





Probability/Confidence

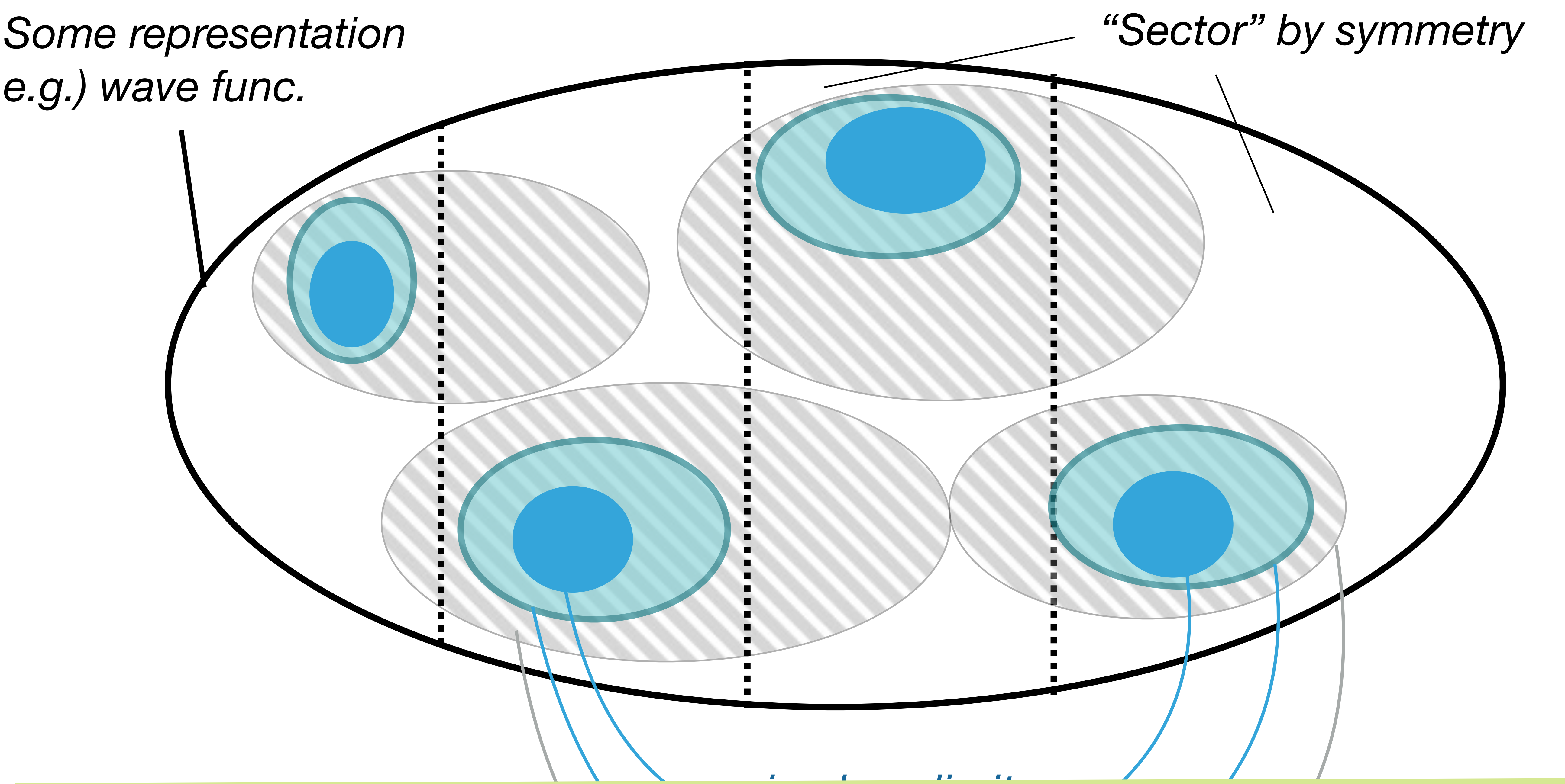
Phase		*Expected behavior*
Z_2	*->*	*Edge localized*
Trivial	*->*	*Bulk localized*
Thermal metal	*->*	*Extended*

“Training” requires knowledge on disorder phase boundary.

Possible to avoid it by statistical average!

"Statistical recovery" of translational symmetry

cf.) Recovery of TRS, Inv. Fulga *et al.*('14)



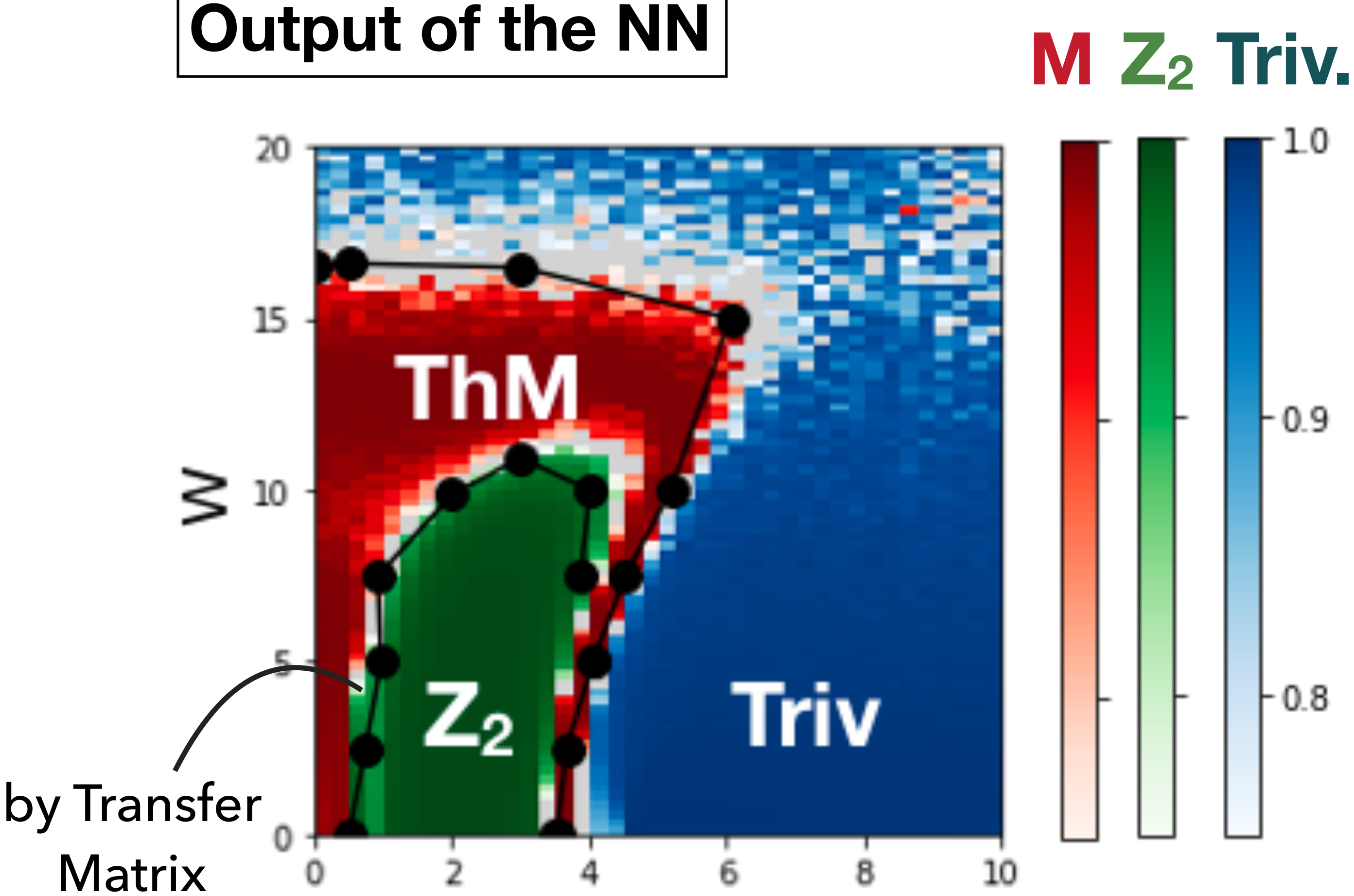
Learn in clean phase, classify dirty phase.

- **Introduction**
  - Objective of Machine Learning
  - Application to Physics
- **Method and Hamiltonian**
  - Problem set up
  - Classification by Artificial Neural Network
- **Result and Discussion**

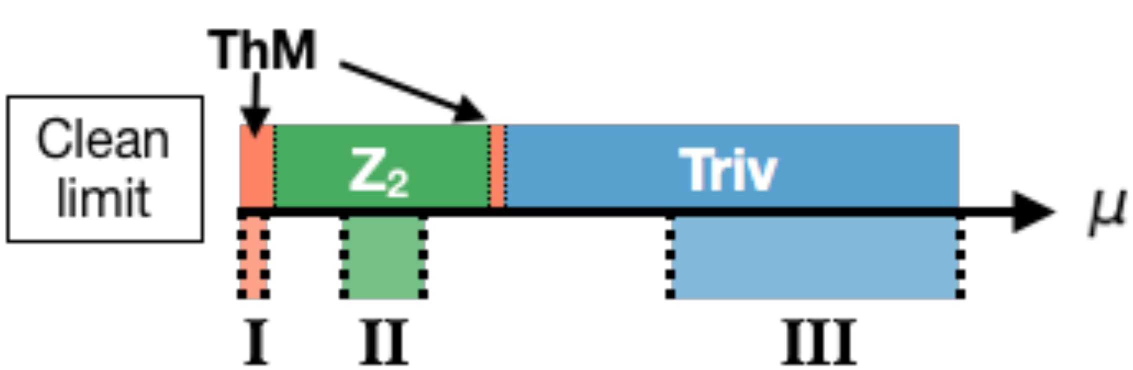
Result I: Ternary Classification

- **$t=1$, $\Delta_p=3$, $\Delta_s=2$, $L_x = L_y = 14$**

Output of the NN



Phase/training data at W=0

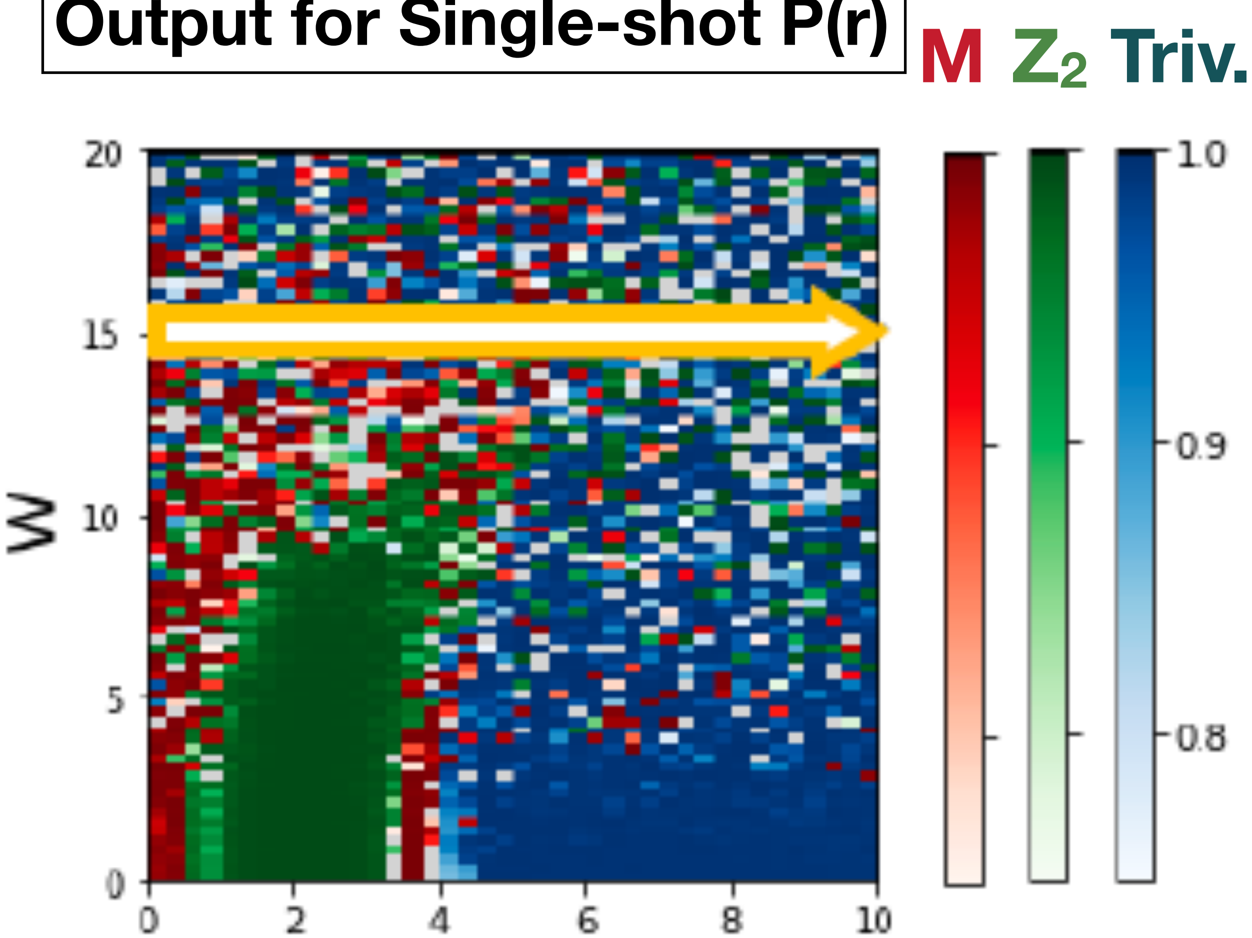


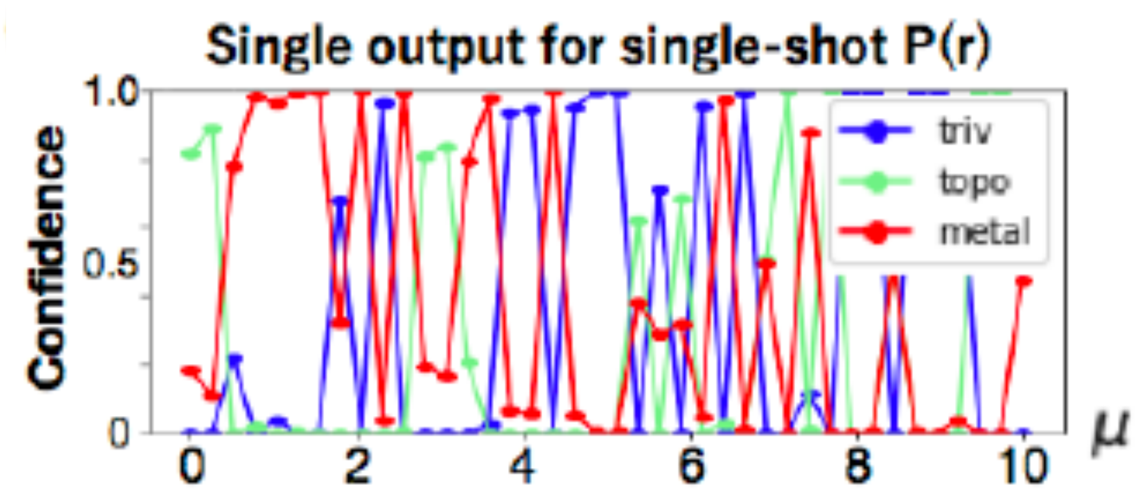
- **Phase boundary reproduced at W=0**
  - Accuracy>90% for test at $\mu \in [0,10]$.
  - Small window of ThM at $\mu \sim 3.5$ detected.
- **Consistency with Transfer Matrix at W>0**
  - ThM-Z2-ThM transition at $\mu \sim 3.5$.
  - Close boundaries of Z2-Thm, ThM-Triv.
  - Confusion(gray) at W~15 improved by increasing disorder average.

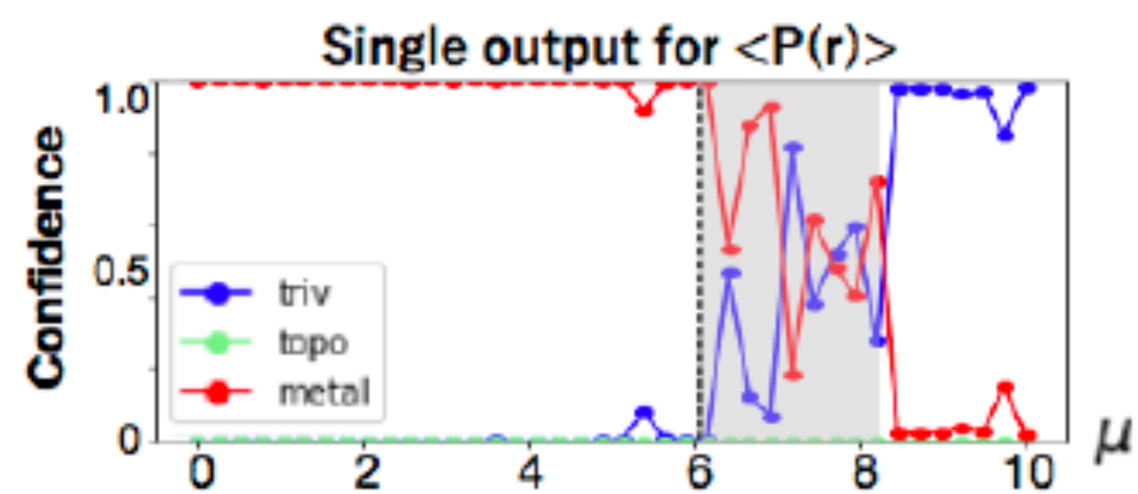
Fails without statistical symmetry recovery

- **t=1, Δ_p=3, Δ_s=2, Lx = Ly = 14**

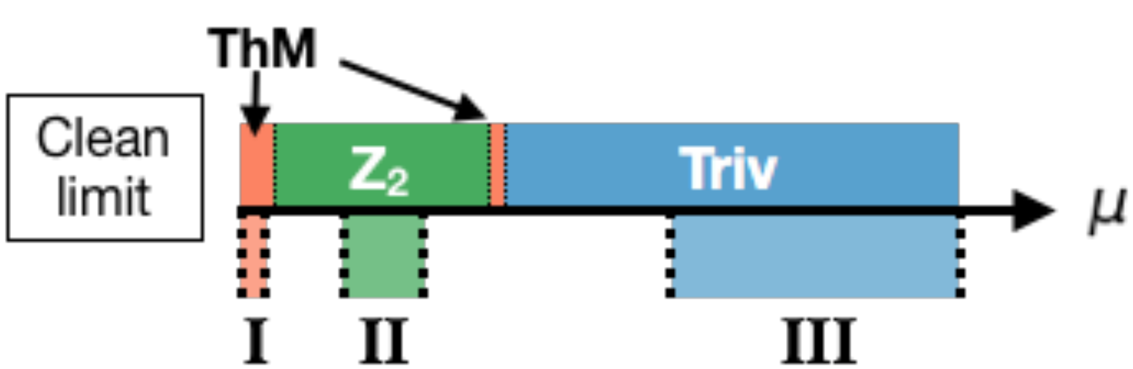
Output for Single-shot P(r)







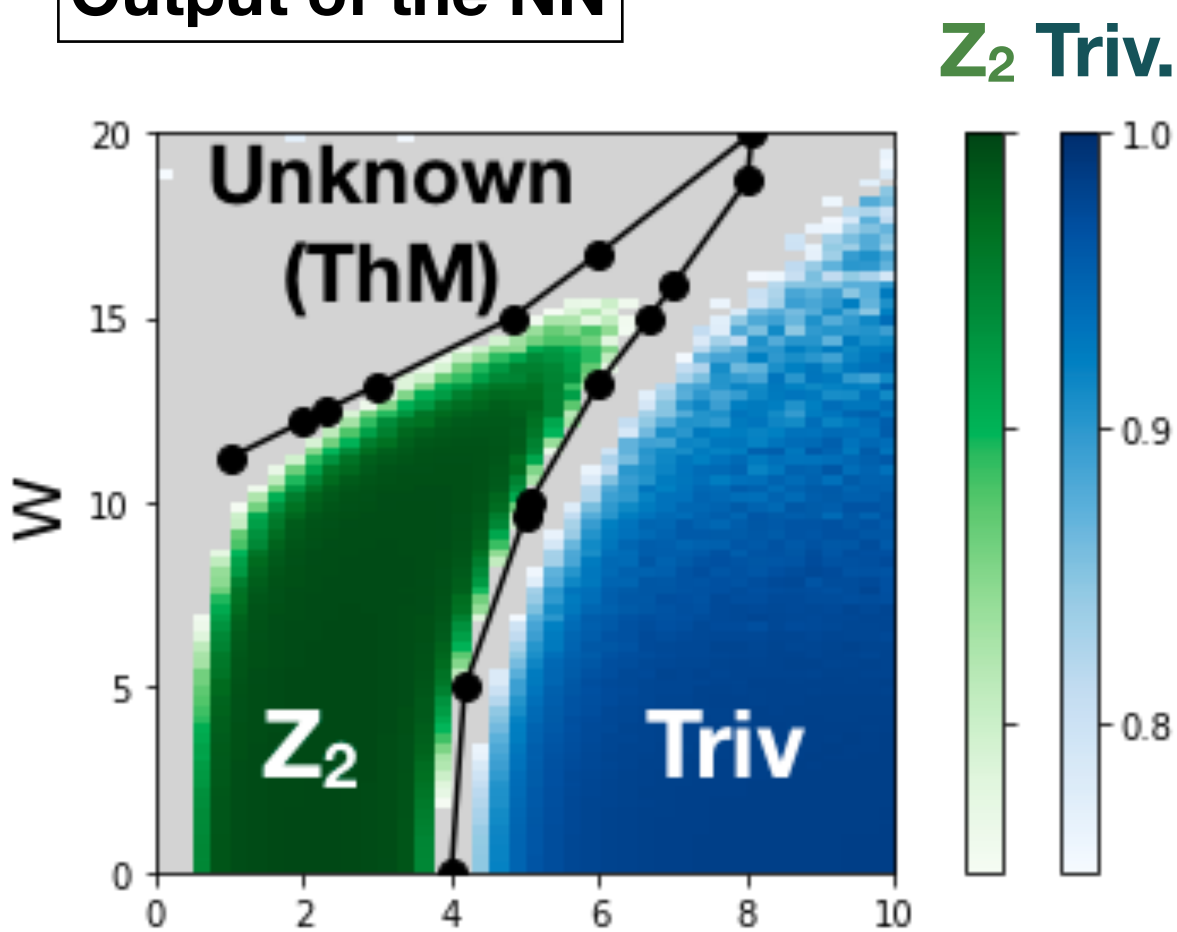
Phase/training data at W=0



Result II: Binary Classification

- **t=1, Δ_p=3, Δ_s=0, Lx = Ly = 14**

Output of the NN



Phase/training data at W=0

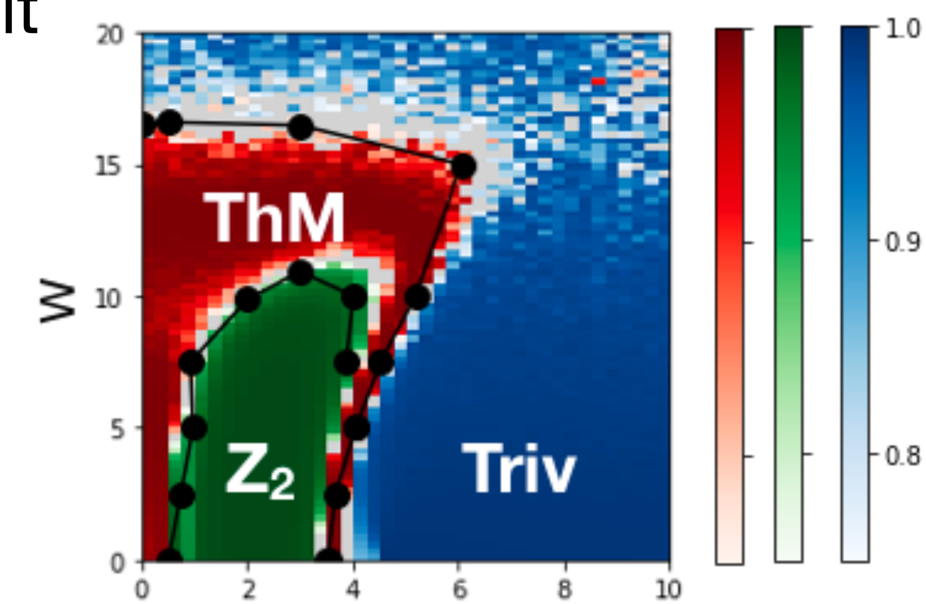
- **Consistency with TM**
  - Accuracy>95% for test at μ∈[0,10], W=0.
  - Z2-triv phase boundary reproduced.
  - Z2-Z2 boundary for confusion at μ~0

- **Confused region above Z2 phase**
  - Output convergence below 0.75.
  - ➡ **Detection of metallic phase.**
  - Shrink of Z2 phase due to finite-size effect.

Summary and Future works

☑ **Quantum phase diagram of class DIII by new method**

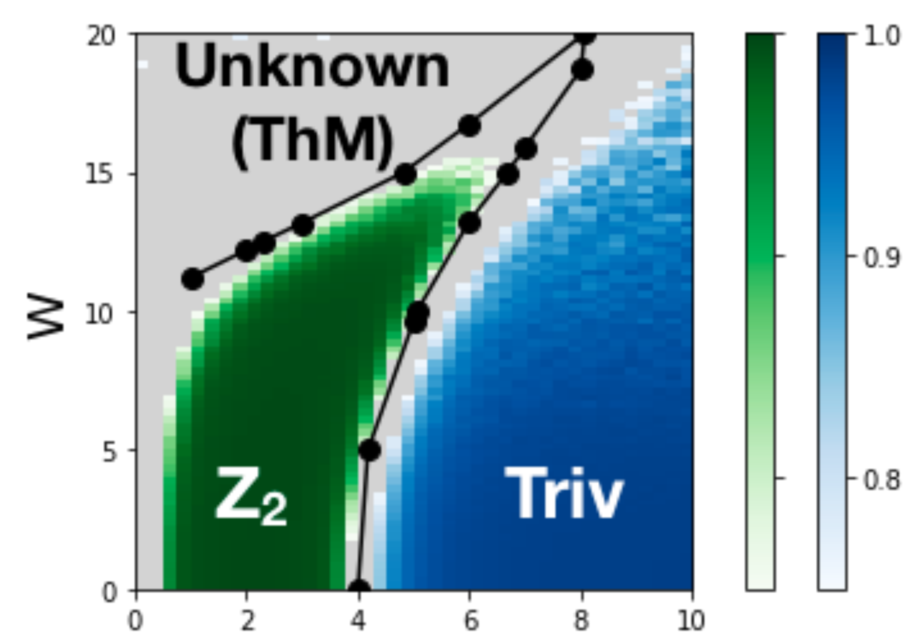
- Extension of phase boundary from clean limit
- Consistency with TM (and NCI)
- Higher precision by increasing samples



☐ Inclusion of higher moments

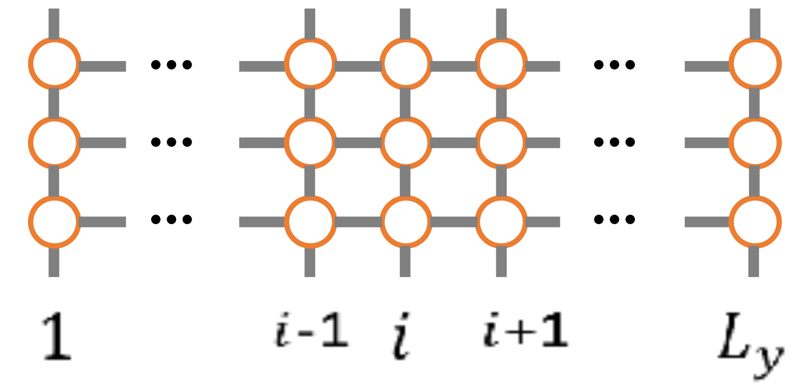
☐ Application to many-body system with disorder

☐ Further classification within the unknown phase



Supplement 1:Transfer Matrix

- **Localization length in quasi-1D system**



$$V_{i,i-1}\psi_{i-1} + H_i\psi_i + V_{i+1,i}\psi_{i+1} = E\psi_i$$

$$\begin{pmatrix} \psi_{i+1} \\ \psi_i \end{pmatrix} = \begin{pmatrix} V_{i+1,i}^{-1}(E - H_i) & V_{i+1,i}^{-1}V_{i-1,i} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_i \\ \psi_{i-1} \end{pmatrix}$$

$$= \cdots = M(E) \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}$$

Localization length

$$\lim_{L_x \to \infty} (MM^\dagger)^{1/2L_x} = U^\dagger \mathrm{diag}(e^{\pm L_y/\Lambda_1}, \ldots, e^{\pm L_y/\Lambda_{L_y}})U$$

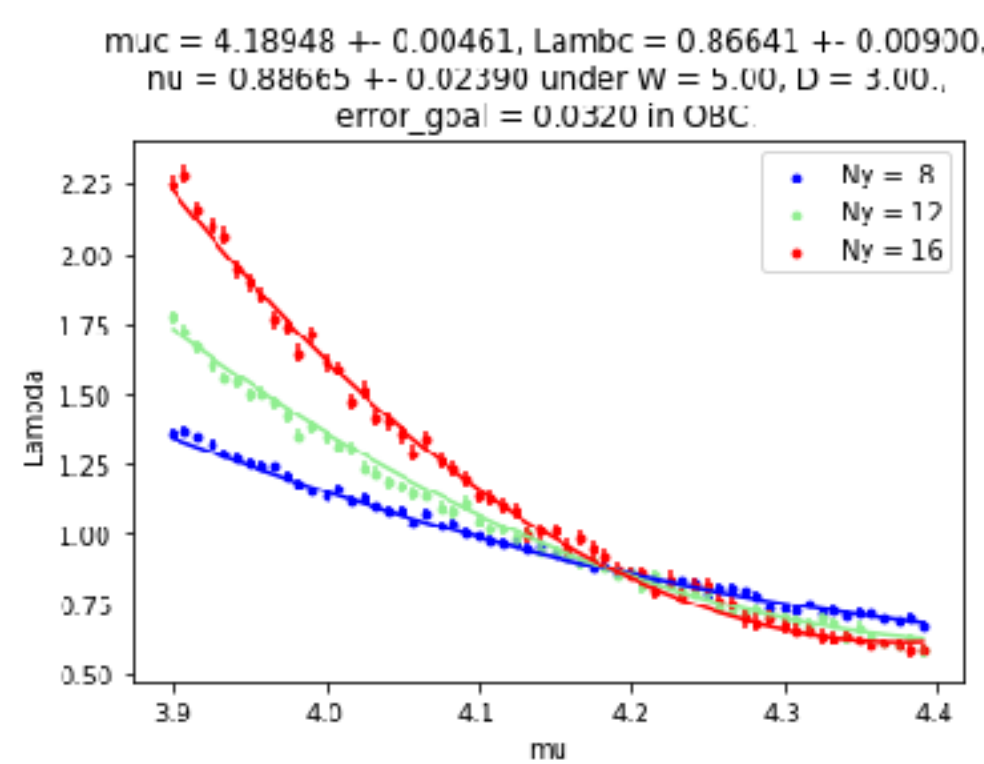
- **Finite-size scaling**

Scaling Λ equivalent to scaling conductance g.

MacKinnon & Kramer (1983), Yamakage *et al.* (2012)

$$\Lambda = \Lambda_0 + \sum_n a_n(q - q_c)^n L_y^{n/\nu}$$

μ or W critical exp.



Supp 2: Methods in real-space regularized system

○ Noncommutative Geometry *(new)*

$$\nu = \frac{1}{2}\dim\ker[\mathcal{A}-1] \bmod 2$$

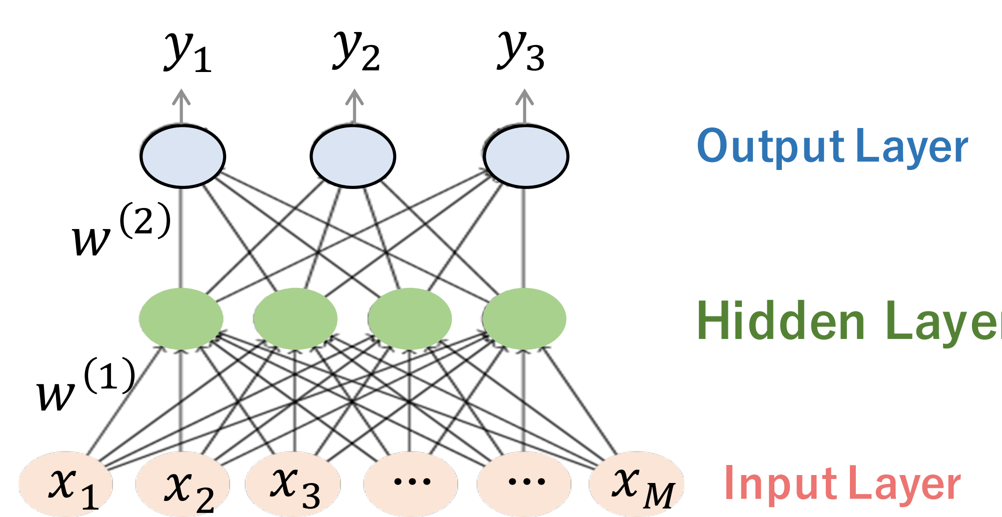
Proj. on Fermi sea.

$$\mathcal{A} = \sigma_3(P_F - D_a P_F D_a)$$

σ_3: Pauli mat. on aux. field; D_a: Dirac operator

- Proof in **infinite** system : Katsura&Koma('16)
- Demonstration in **finite** system : This work, Akagi *et al. (arXiv:1709.05853)*

○ Machine Learning *(new)*

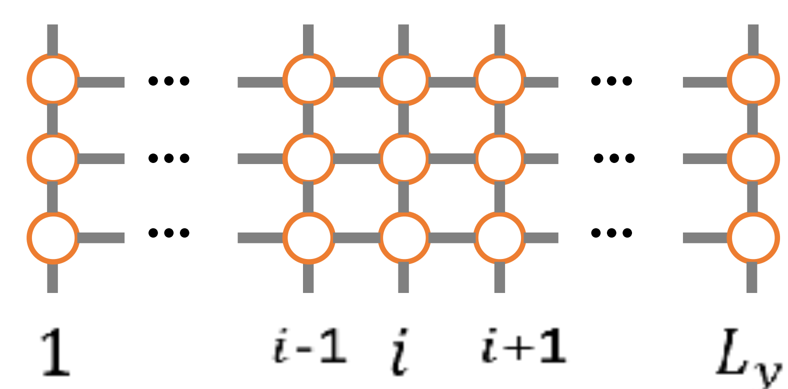


- Classification of phases by neural network.
- Learn clean phase, predict dirty phase.
- **Focused talk on 09/24 15:30**

☑ Transfer Matrix Method MacKinnon('83)

- Finite-size scaling of localization length in quasi-1D

$$\Lambda = \Lambda_0 + \sum_n a_n (q-q_c)^n L_y^{n/\nu}$$



Supplement 3: Defining Z2 topo. inv.

- **Noncommutative Geometry** Avron, Seiler & Simon ('94)

k-derivative → Commutator in real space

$$\frac{d}{dk} \to [iX, \ldots]$$

➡ Precise definition of topo. inv. **for any symmetry class!**

H. Katsura and T. Koma, arXiv:1611.01928.

In practice, results in counting #(eigenvalue = 1) of

$$\mathcal{A} = \sigma_3 [D_a(\vec{x}), D_a(\vec{x}) P_F]$$ (Commutator for space-dependent operator)

i.e.,

$$\nu = \frac{1}{2} \dim \ker[\mathcal{A} - 1] \bmod 2$$

where

$$D_a(\vec{x}) := \frac{1}{|\vec{x} - \vec{a}|} (\vec{x} - \vec{a}) \cdot \vec{\sigma}$$: Dirac operator

($D_a^2 = 1, D_a = D_a^\dagger$, σ for aux. field)

$$P_{\mathrm{F}} = \frac{1}{2\pi i} \oint_{\mathcal{C}} (z - H)^{-1} dz$$: Projection on Fermi sea