Detection of Phase Transition via Convolutional Neural Networks

Akio TOMIYA

CCNU→ RIKEN(Tokyo)→ RIKEN(BNL, NY)

A. Tanaka (RIKEN AIP center)

Based on JPSJ86, 063001 (2017)
(arXiv:1609.09087)

 akio.tomiya@mail.ccnu.edu.cn

Self-introduction

I am interested in machine learning with/for physics

I am working in CCNU(Central China Normal Univ., Wuhan).

Self-introduction

I am interested in machine learning with/for physics

I am working in CCNU(Central China Normal Univ., Wuhan).

1. One of the hottest place in China (三大火炉)
2. Ancient battlefield, the red cliff in 3 kingdoms

We are here

Self-introduction

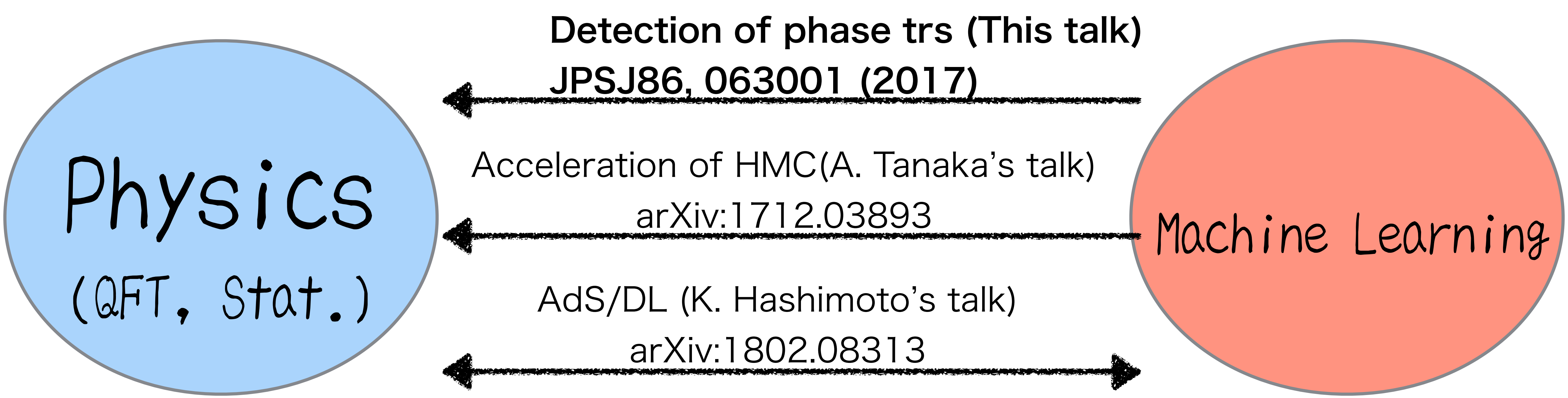
I am interested in machine learning with/for physics

I am working in CCNU(Central China Normal Univ., Wuhan).

Ph.D: Theory of high energy particle physics

Interest: Lattice gauge theory(QCD) at finite temperature, phase transitions, quantum entanglement

My viewpoint to ML&physics:



Summary : Motivations and Results

Neural networks (NN) can detect the critical point

Motivation: NN can recognize images and configurations of spin system are similar to images.

Can NN recognize phase or phase transition?

Summary : Motivations and Results

Neural networks (NN) can detect the critical point

Motivation: NN can recognize images and configurations of spin system are similar to images.
Can NN recognize phase or phase transition?

Method: By employing Convolutional NN(CNN), we try to detect phase transition in the 2D Ising model.

Summary : Motivations and Results

Neural networks (NN) can detect the critical point

Motivation: NN can recognize images and configurations of spin system are similar to images.
Can NN recognize phase or phase transition?

Method: By employing Convolutional NN(CNN), we try to detect phase transition in the 2D Ising model.

Results: CNN can detect phase transition in few % accuracy without explicit information of Tc

System size	β_c (CNN)
8×8	0.478915
16×16	0.448562
32×32	0.451887

$$\beta_c^{\text{Exact}} = \frac{1}{2}\log(\sqrt{2}+1)$$
$$\sim 0.440686$$

Introduction

2 dimensional Ising model

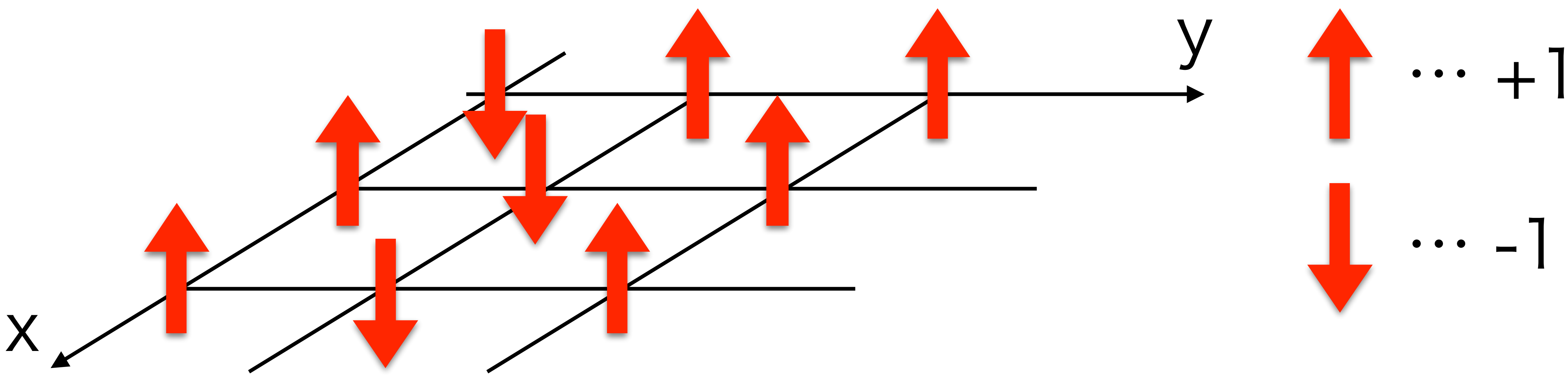
Ising model and phase transition(1/4)

arXiv:1609.09087, A. Tomiya

2 dimensional ferromag. Ising model in the statistical mechanics

Wilhelm Lenz (1920)

+1 or -1 on each point in 2 dim. lattice (grid)



A toy model of ferro-magnetism, similar to quantum field theory (QFT) and exactly solvable

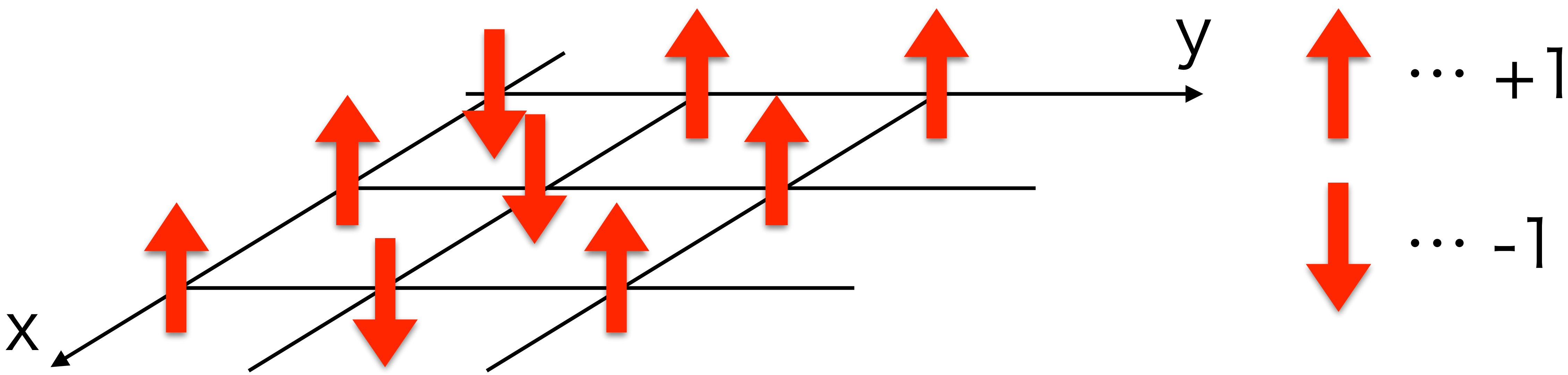
i.e. good testing ground for new methods towards applications to QFT

Ising model and phase transition(2/4)

2 dimensional Ising model in the statistical mechanics

Wilhelm Lenz (1920)

+1 or -1 on each point in 2 dim. lattice (grid)



$$\sigma_{x,y} = -1 \text{ or } +1,$$

$\{\sigma\}$: a spin configuration, which is given by in a probability,

$$P\big[\{\sigma\}\big] \propto \exp\big[-\beta H[\{\sigma\}]\big] \qquad \beta \propto 1/T$$

$$H = -\sum_{x,y} \sigma_{xy}(\sigma_{(x+1),y} + \sigma_{x,(y+1)}),$$

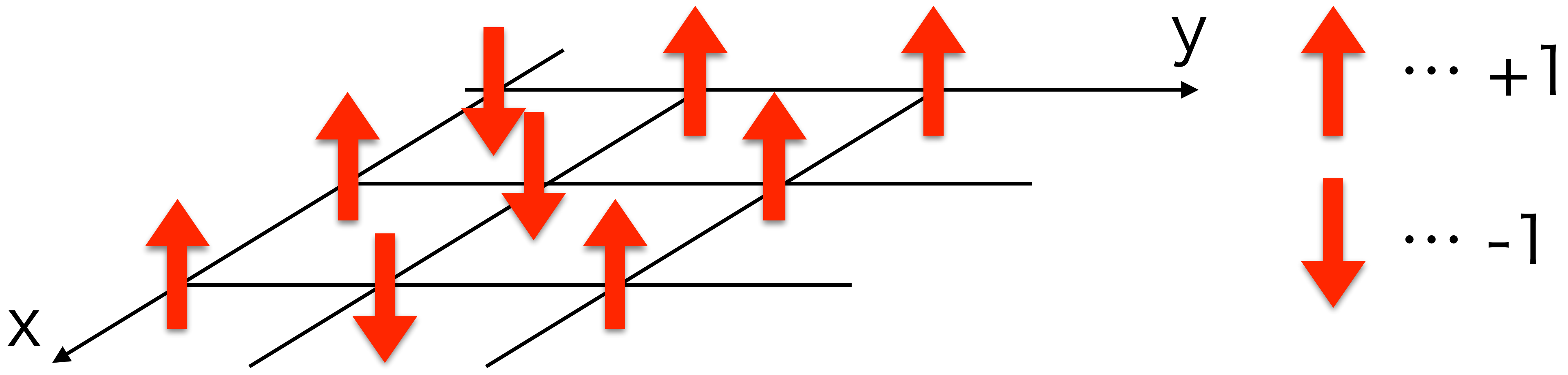
Energy function

Ising model and phase transition(3/4)

arXiv:1609.09087, A. Tomiya

2 dim. Ising model is the simplest model which has a phase transition

Wilhelm Lenz (1920)



The expectation value of magnetization $<\sigma>$ at inv. temp. $\beta \propto 1/T$:

$$\langle\sigma\rangle \propto \boxed{\sum_{\{\sigma\}}} \sum_{x,y} \sigma_{x,y} \exp\left[-\beta H[\{\sigma\}]\right]$$

Sum over all possible spin combinations: Difficult

Done by Onsager(1944), Nambu(1950), ⋯

$\langle\sigma\rangle$ is zero above Tc and non-zero below Tc

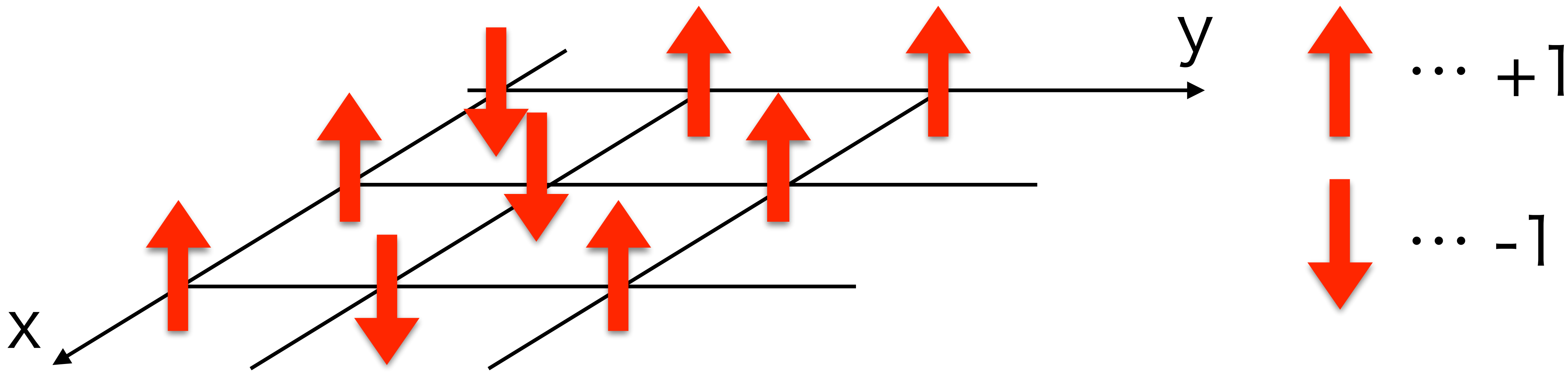
$$\beta_c^{\text{Exact}} = \tfrac{1}{2}\log(\sqrt{2}+1)$$
$$\sim 0.440686$$

Ising model and phase transition(4/4)

arXiv:1609.09087, A. Tomiya

We have to choose order parameters to define phase transitions

Wilhelm Lenz (1920)



$$\langle \sigma \rangle \propto \sum_{\{\sigma\}} \sum_{x,y} \boxed{\sigma_{x,y}} \exp\left[-\beta H[\{\sigma\}] \right]$$

$$H = -\sum_{x,y} \sigma_{xy}(\sigma_{(x+1),y} + \sigma_{x,(y+1)}),$$

H is invariant under $\sigma \to -\sigma$ but $<\sigma>$ can be non-zero.

σ is called "order parameter" (spontaneous symmetry breaking)

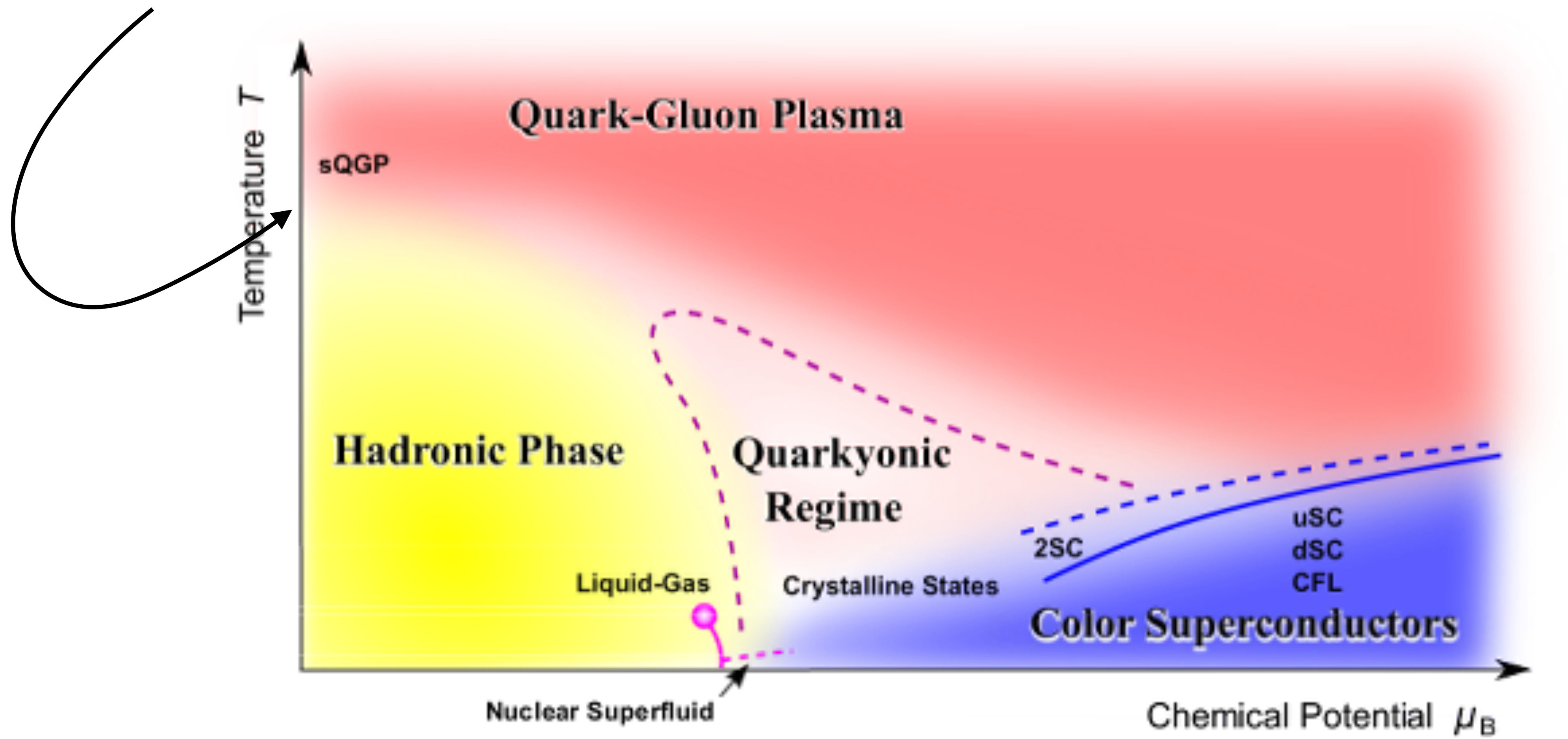
Q. To define phase transition, we need to specify order parameters.

Without teaching order parameter, can we find phase transitions?

Context of this work(1/2)

Deconfinement phase transition in QCD (Quantum Chromo-Dynamics)

Nucleons (hadrons) dissociate to quarks/gluons above
T ~ 150 MeV ~ 10^11 Kelvin (ALICE expriment in LHC)
= Confinement/deconfinement phase transition in QCD

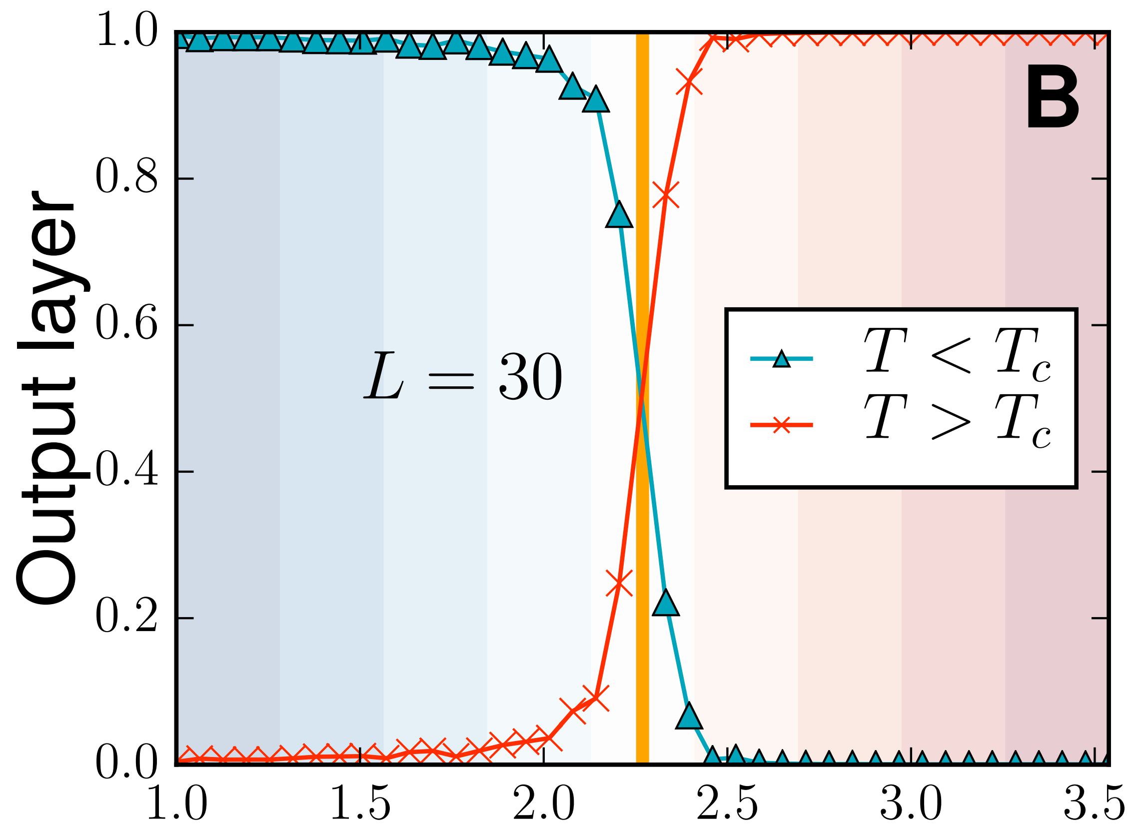


If we ignore quarks, Polyakov loops give an order parameter but

QCD phase trs. with quarks does not have clear order parameters…

Context of this work(2/2)

A previous work by a Perimeter group



Juan Carrasquilla and Roger G. Melko
arXiv: 1605.01735

2D Ising model.

Input configurations, phases as labels (2 labels).

Output layer exhibits the critical temperature

Teaching the phase -> implicitly teaching Tc

Outline

1. Introduction
2. Convolutional NN
3. Our works & results
4. Summary

Convolutional NN

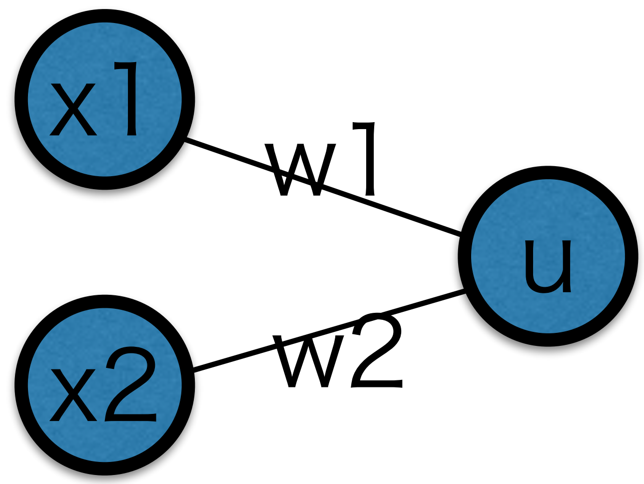
Neural networks

Elements in our neural networks

arXiv:1609.09087, A. Tomiya

x1, x2:

Inputs

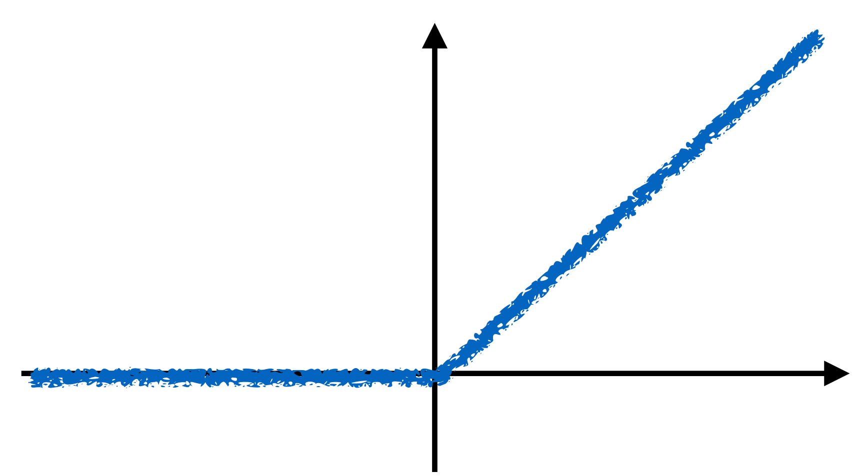


w: fit parameters

$$u = \mathrm{ReLu}(w_1 x_1 + w_2 x_2)$$

Bias terms can be included

$$\mathrm{ReLu}(x) = \max(x, 0)$$

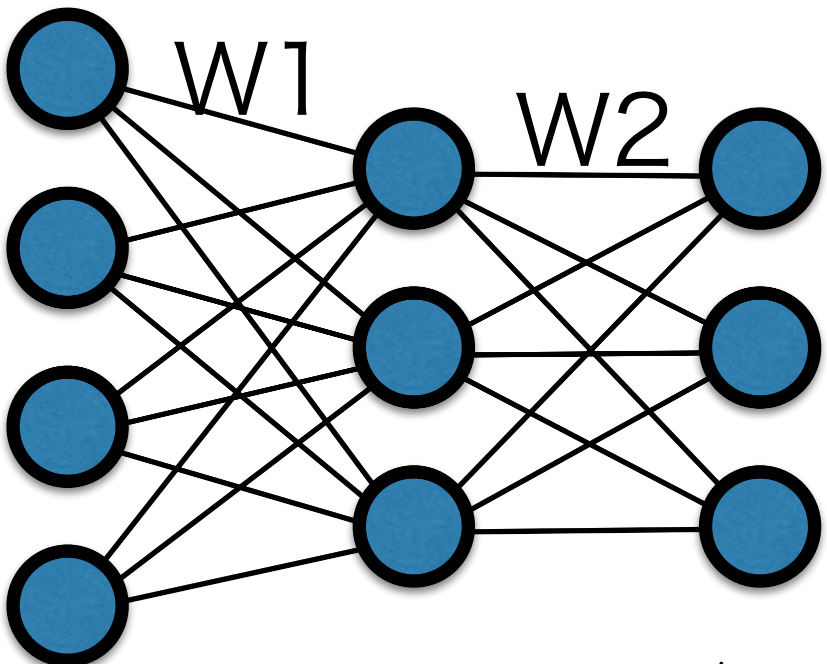


Neural networks

Neural networks

W: Weight matrices

x: Input



$\vec{x}$ $\vec{u}$ $\vec{f}(\vec{x}, W)$ $\vec{y}$

y: labels of x

try to minimize

$E(\vec{f}(\vec{x}, W), \vec{y})$: Error function, “distance” between f and y

Minimizing E by tuning weight matrices W = Learning of NN

Convolutional Neural networks

It improves image recognition

Convolution = Filtering with fitted filters

= special sparse weights

p11	p12	p13	p14
p21	p22	p23	p24
p31	p32	p33	p34
p41	p42	p43	p44

Image: p_ij

*

Filter

f11	f12
f21	f22

f_st ∈ R

f also will be trained

Output: $$u_{ij} = \sum_s \sum_t f_{st} p_{i+s,j+j}$$

1. Essentially same as the convolution in the Fourier transformation.
2. It helps to solve image recognition problem

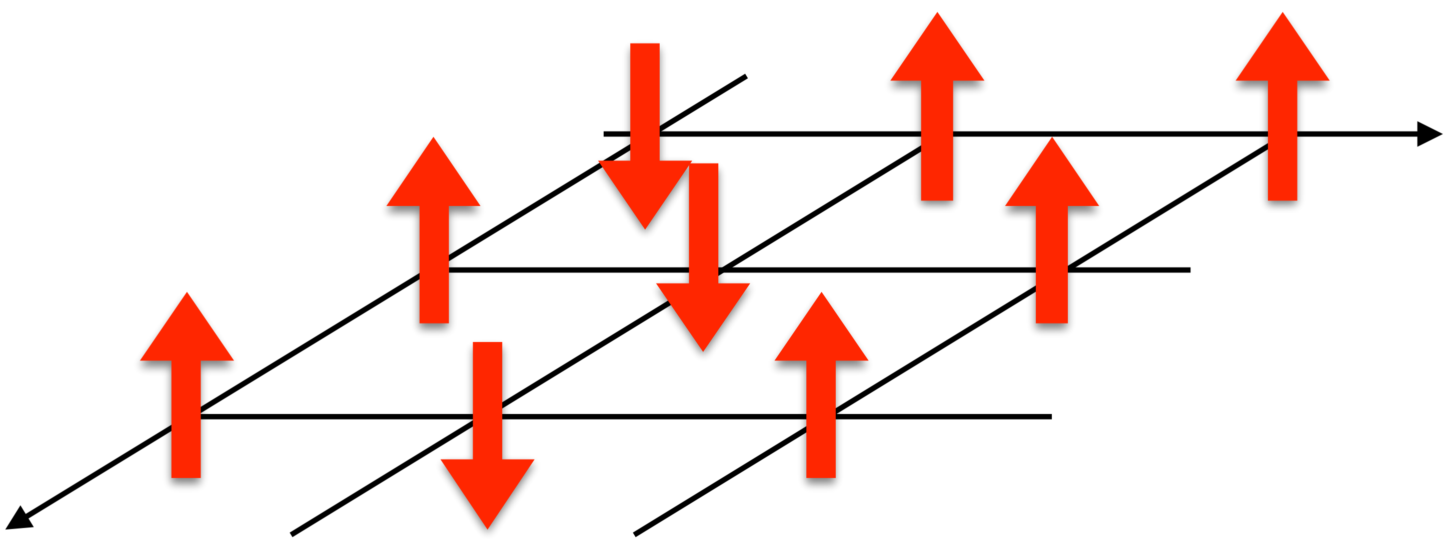
Our works & results

Idea of this work

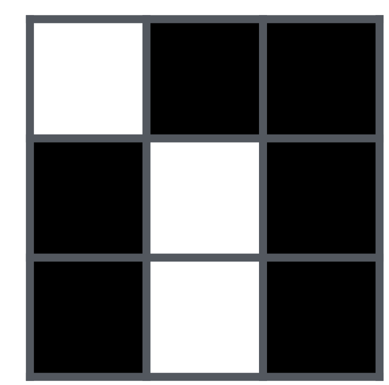
Neural networks

arXiv:1609.09087, A. Tomiya

By employing Convolutional NN(CNN), we try to detect phase transition in the 2D Ising model.



<~>



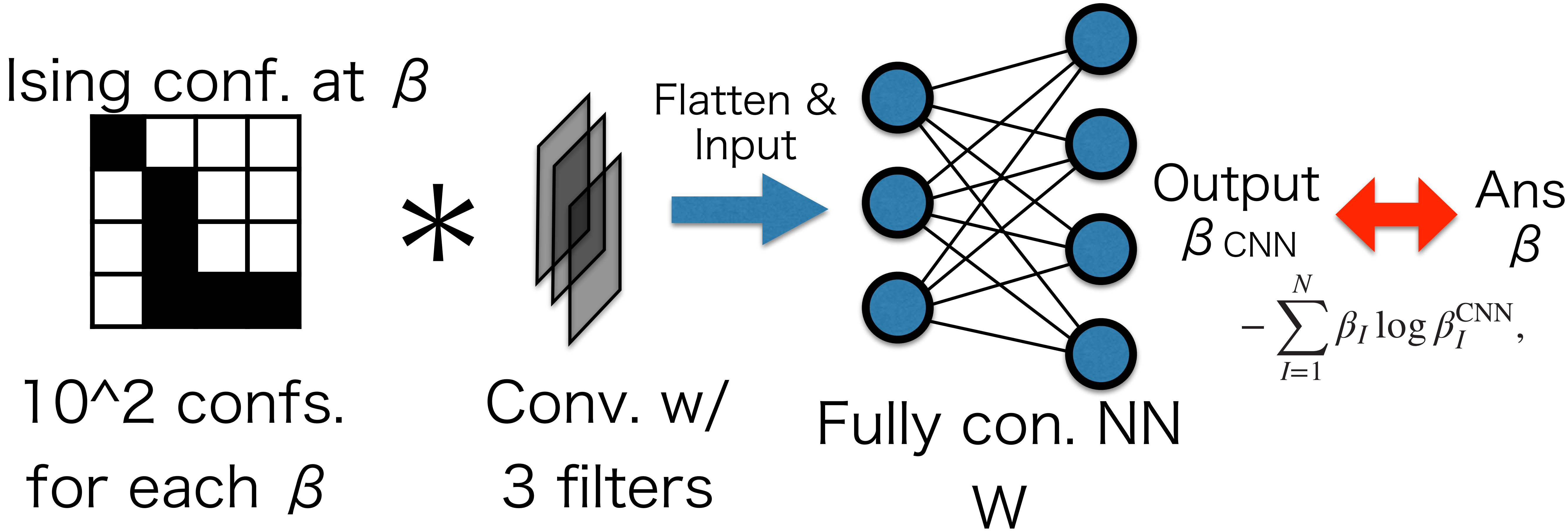
Q. To define phase transition, we need to specify order parameters.

Without teaching order parameter, can we find phase transitions?

A. Yes (in some sense)

Setup

Convolutional neural net as a “thermometer”



Inv. β is
discretized for
labeling
(one-hot)

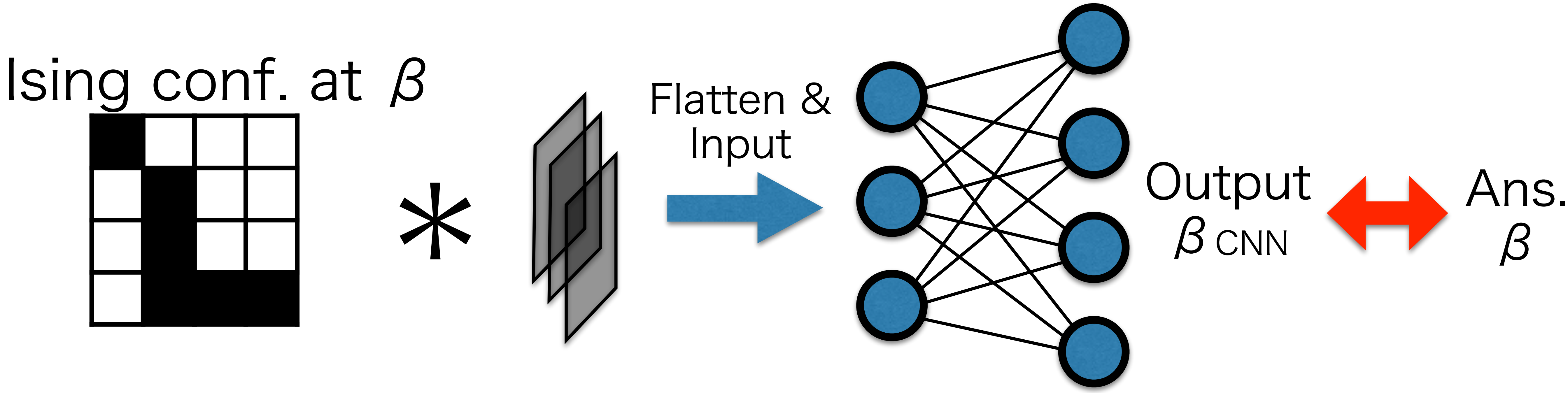
$$\vec{\beta} = \begin{cases} (1,0,\ldots,0,0) & \text{for } \beta < 0 \\ (0,1,\ldots,0,0) & \text{for } \beta \in \left[0, \frac{1}{N-2}\right) \\ \cdots \\ (0,0,\ldots,1,0) & \text{for } \beta \in \left[\frac{N-3}{N-2}, 1\right) \\ (0,0,\ldots,0,1) & \text{for } 1 \le \beta \end{cases}.$$

$0.2 < \beta < 10,$

$N = 100$

Setup

Convolutional neural net as a “thermometer”



Training process:

$\beta_c^{\mathrm{Exact}} \sim 0.440686$

0. Configurations for $\beta \in [0.2, 10]$ are prepared by MCMC

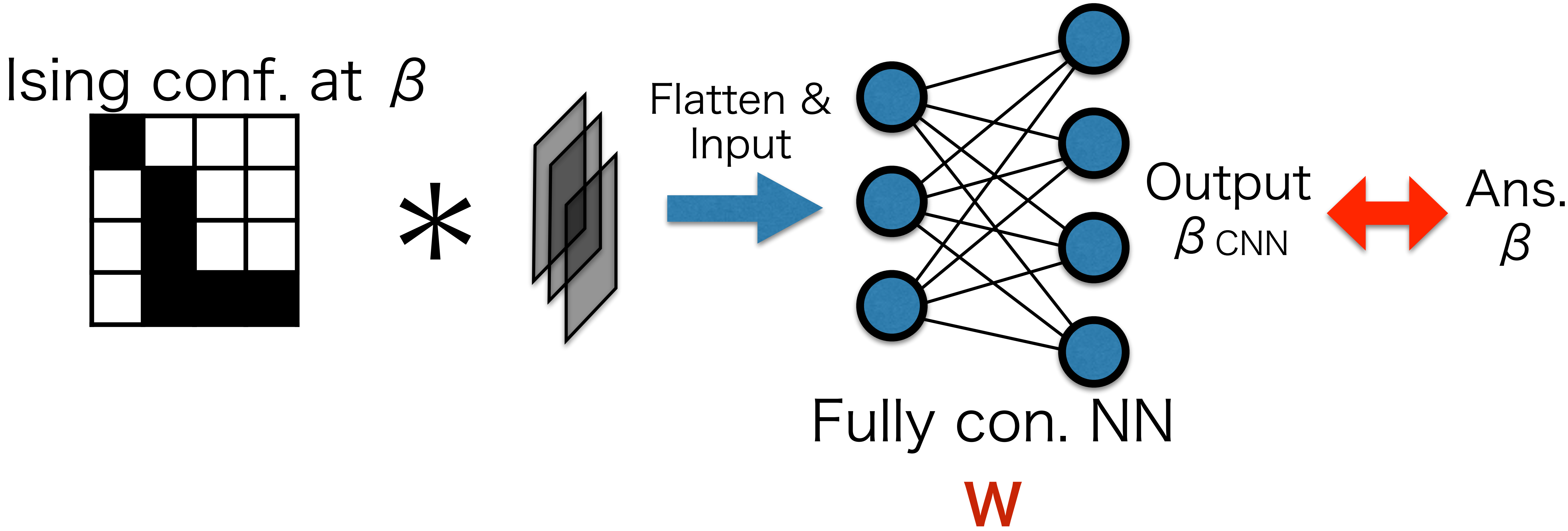
1. Chose one dataset for β, i.e. configurations(β) and label(β)

2. Training with configurations with the β

3. Back 1 and repeat for 10^5.

Setup

Convolutional neural net as a “thermometer”

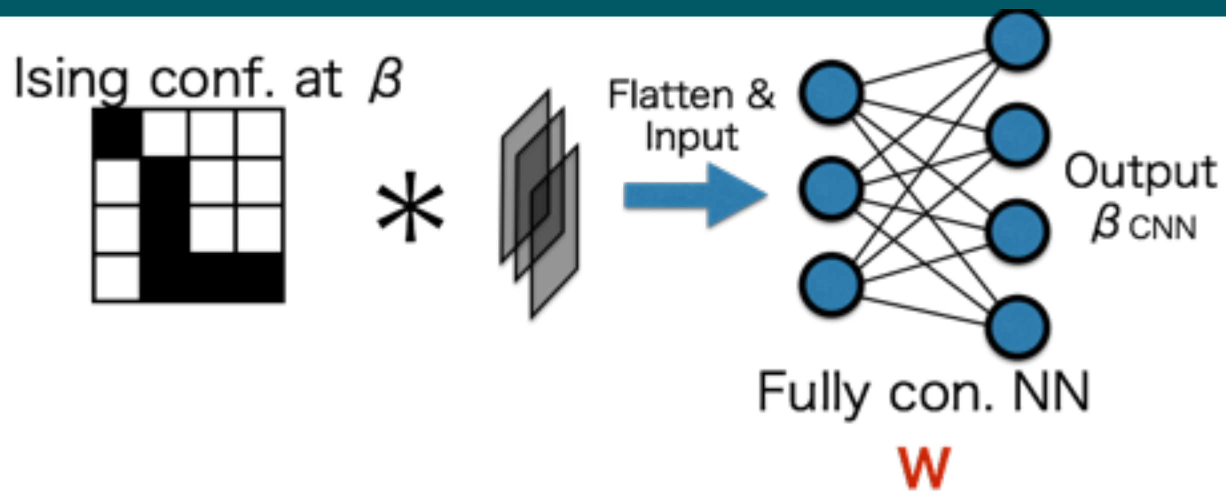


W connects “Spin configurations” to “temperature”

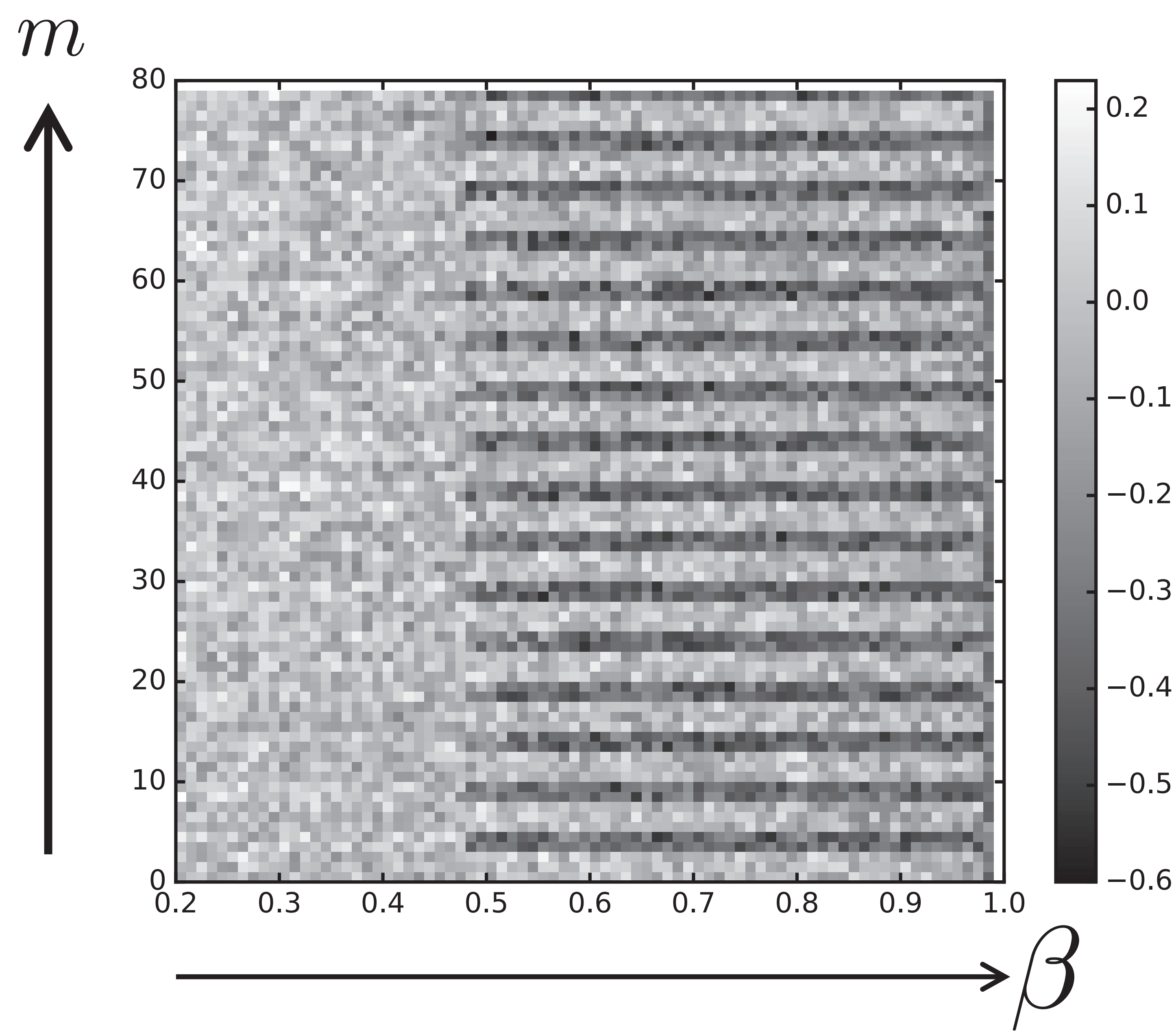
We plot **W** as a heat map after training.

Results (CNN)

Weight W has a pattern

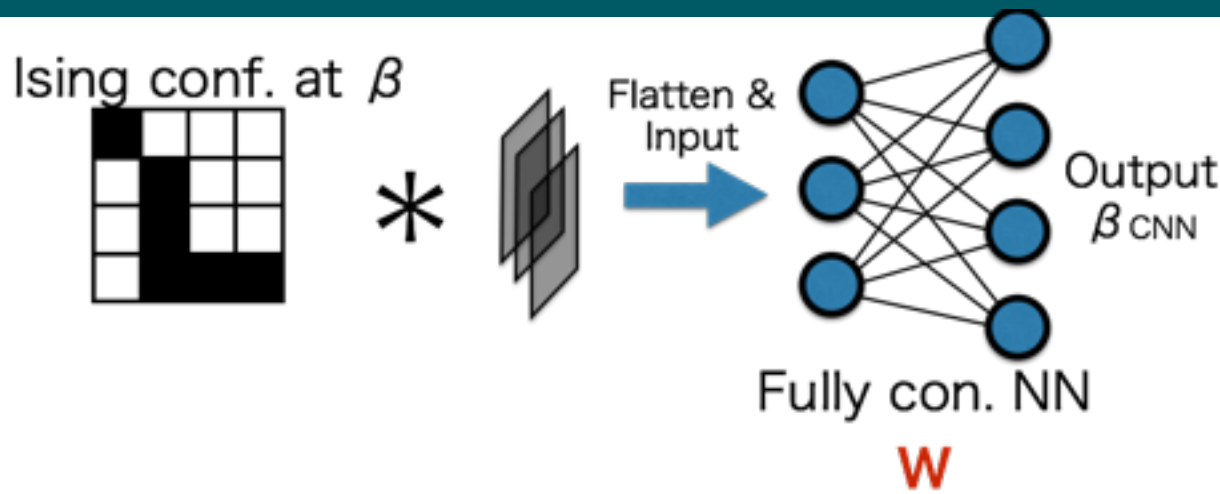


Heat map of W
(After whole training)



Results (CNN)

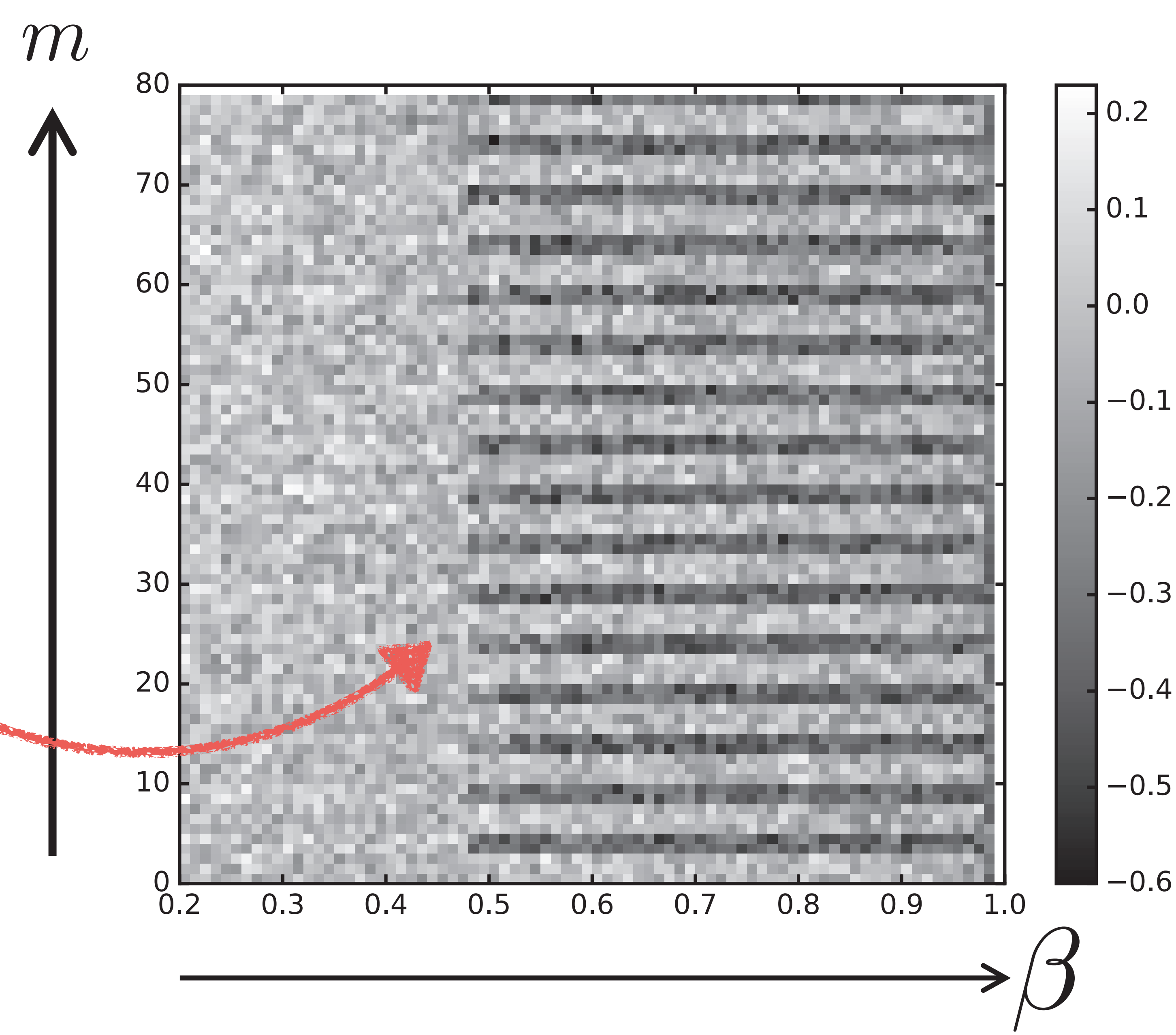
Weight W has a pattern



Heat map of W
(After whole training)

The weight matrix
has a pattern

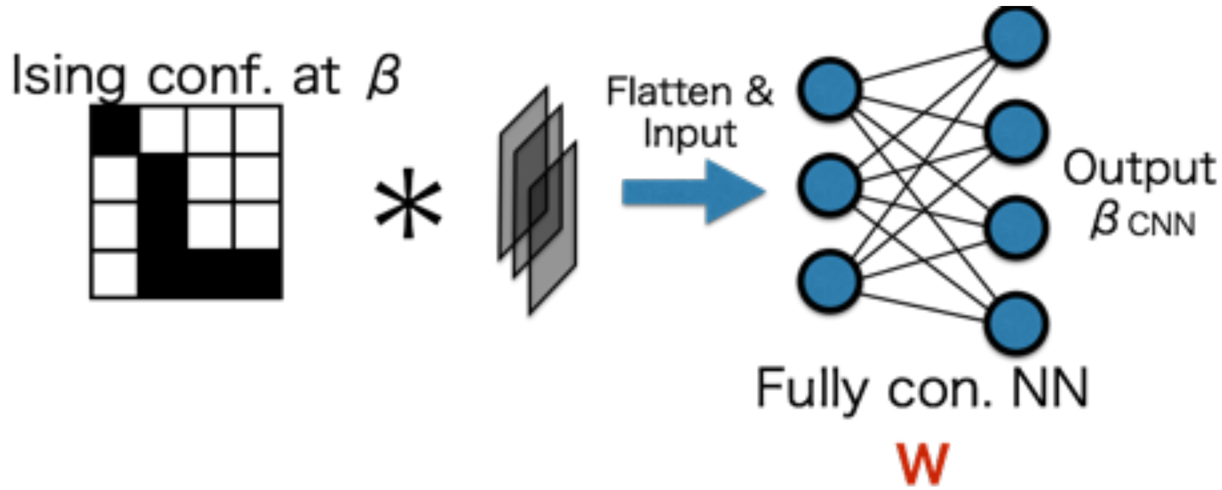
$$\beta_c^{\text{Exact}} = \frac{1}{2}\log(\sqrt{2}+1)$$
$$\sim 0.440686$$



Results (CNN)

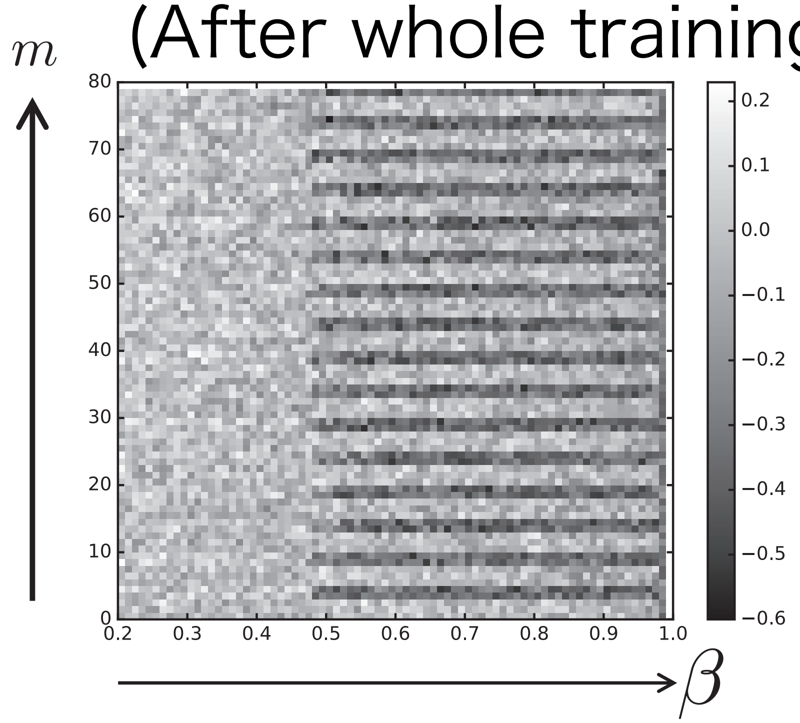
arXiv:1609.09087, A. Tomiya

Weight W has a pattern

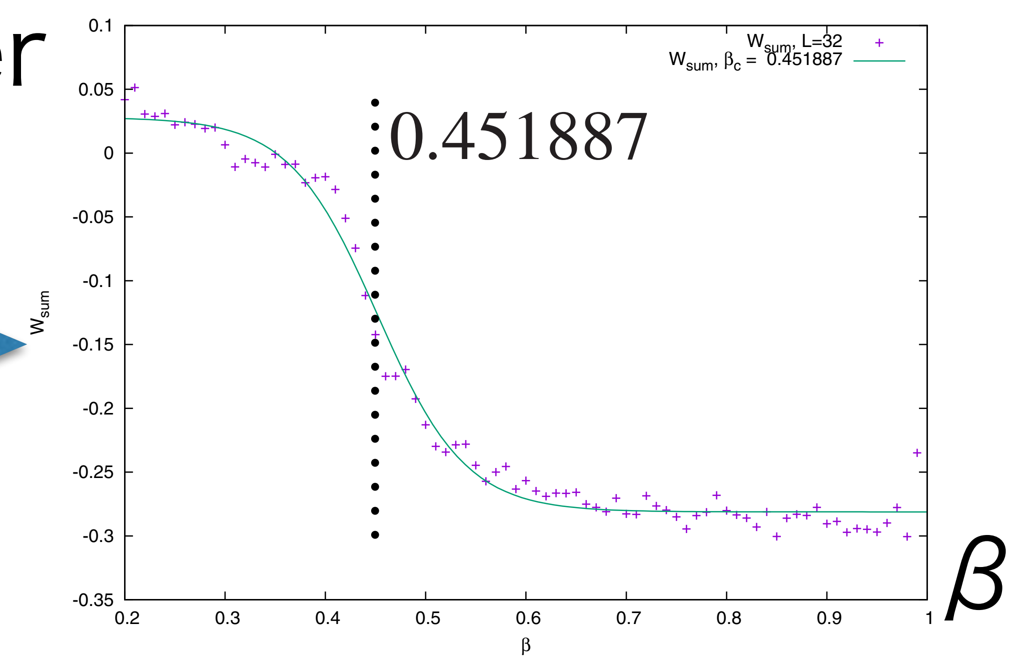


Heat map of W

(After whole training)



Summing W over m direction



$$\beta_c^{\text{Exact}} = \frac{1}{2}\log(\sqrt{2}+1) \sim 0.440686$$

Discussions

W carries information of criticality?

System size	β_c (CNN)
8×8	0.478915
16×16	0.448562
32×32	0.451887

- We do not know why W has information of criticality
- This behavior is stable for L=8, 16, 32.
- It looks like behavior of magnetization along with the temperature but there are no (explicit) reasons for such coincidence.
- Filter F seems to capture information of magnetization but FNN can capture Tc.
- On the other hand, this network is not a good thermometer
- There are still some discussions:
  - Weights cannot carry information! (K-I. Aoki et al.) Using RG
  - Weights can carry information! (Iso et al.) Using RG

Related works

Detecting phase transitions

1. Ising models: Juan Carrasquilla and Roger G. Melko arXiv: 1605.01735
2. Transverse Ising model (**Quantum phase trs**), S. Arai et. al. JPSJ 87, 033001 (2018)
3. XY model
    1. W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E95, 062122 (2017).
    2. S. J. Wetzel, Phys. Rev. E 96, 022140 (2017).
    3. C. Wang and H. Zhai, Phys. Rev. B 96, 144432 (2017).
    4. M. J. S. Beach, A. Golubeva, and R. G. Melko, Phys. Rev.B 97, 045207 (2018).
4. Heisenberg model
    1. I. A. Iakovlev, O. M. Sotnikov, and V. V. Mazurenko, arXiv:1803.06682
5. RG:
    1. S. Iso, S. Shiba, and S. Yokoo, Phys. Rev. E 97, 053304 (2018).
    2. K-I. Aoki et al, Learning from estimation of temperature (Only in Japanese)
6. Skyrmion, Vinit Kumar Singh, Jung Hoon Han,arXiv:1806.03749

Summary : Motivations and Results

Neural networks (NN) can detect the critical point

Motivation: NN can recognize images and configurations of spin system are similar to images.
Can NN recognize phase or phase transition?

Method: By employing Convolutional NN(CNN), we try to detect phase transition in the 2D Ising model.

Results: CNN can detect phase transition in few % accuracy without explicit information of Tc

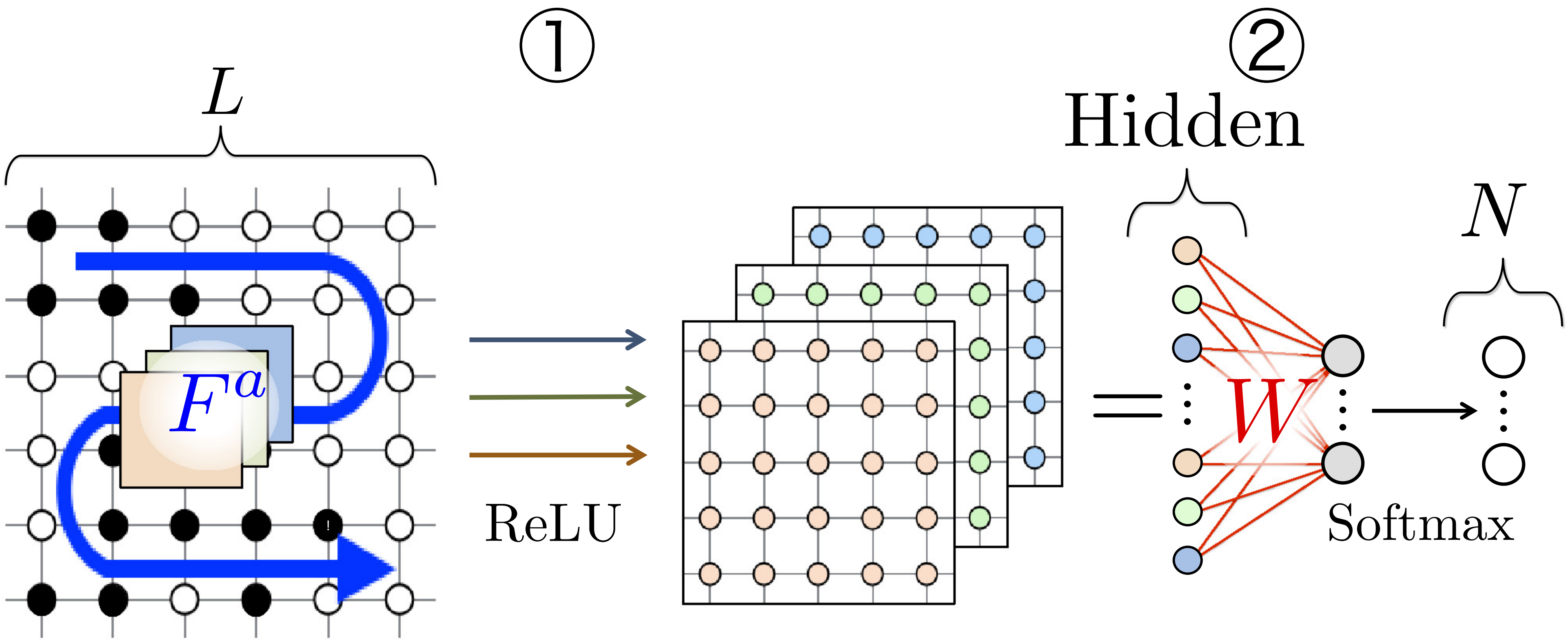
System size	β_c (CNN)
8×8	0.478915
16×16	0.448562
32×32	0.451887

$$\beta_c^{\mathrm{Exact}} = \frac{1}{2}\log(\sqrt{2}+1)$$
$$\sim 0.440686$$

Backups

Setup (detailed)

Our Neural networks



① $\{\sigma_{xy}\}$

$$\xrightarrow{\text{conv}} \sum_{i,j=1}^{N_f} \sigma_{(sX+i)(sY+j)} F^a_{ij} = \Sigma^a_{XY}$$

$$\xrightarrow{\text{ReLU}} \max(0, \Sigma^a_{XY}) = u^a_{XY}$$

$$\xrightarrow{\text{flatten}} \vec{u} = [u^1_{11}, u^2_{11}, \dots u^C_{11}, u^1_{21}, u^2_{21}, \dots, u^C_{21}, \dots] = [u_m].$$

② $[u_m]$

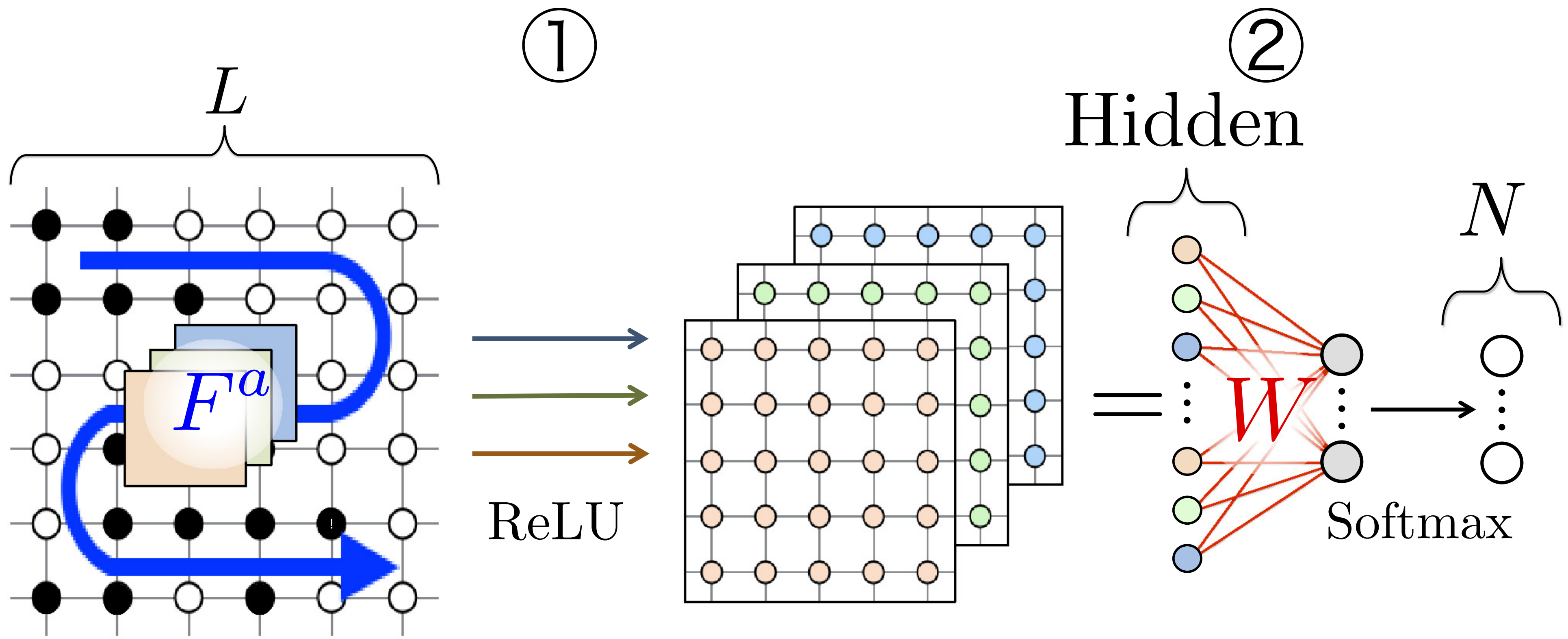
$$\xrightarrow{\text{fully-connected}} \sum_{m=1}^{L^2/s^2 \times C} W^m_I u_m = z_I$$

$$\xrightarrow{\text{Softmax}} \frac{e^{z_I}}{\sum_{J=1}^{N} e^{z_J}} = \beta^{\text{CNN}}_I.$$

$C = 5$, and $s = L/4$ $\quad N_f = 3$,

Setup (detailed)

Our Neural networks



F_{ij}^{a} in convolution, W_{I}^{m} in fully connected layer.

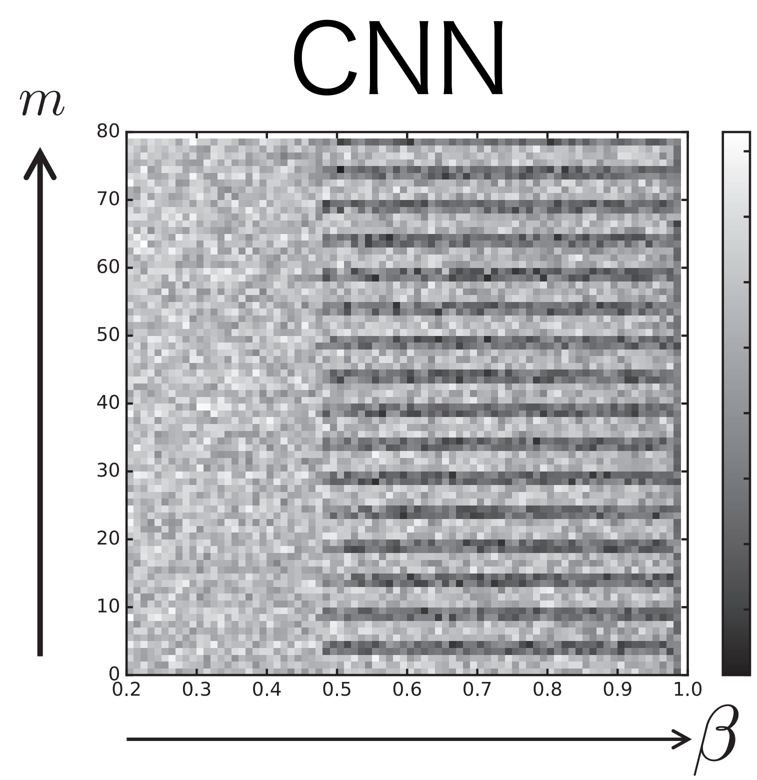
$$\begin{bmatrix} a = 1, \ldots, C \\ i, j = 1, \ldots, N_f \end{bmatrix} \quad \begin{bmatrix} m = 1, \ldots, L^2/s^2 \times C \\ I = 1, \ldots N \end{bmatrix}$$

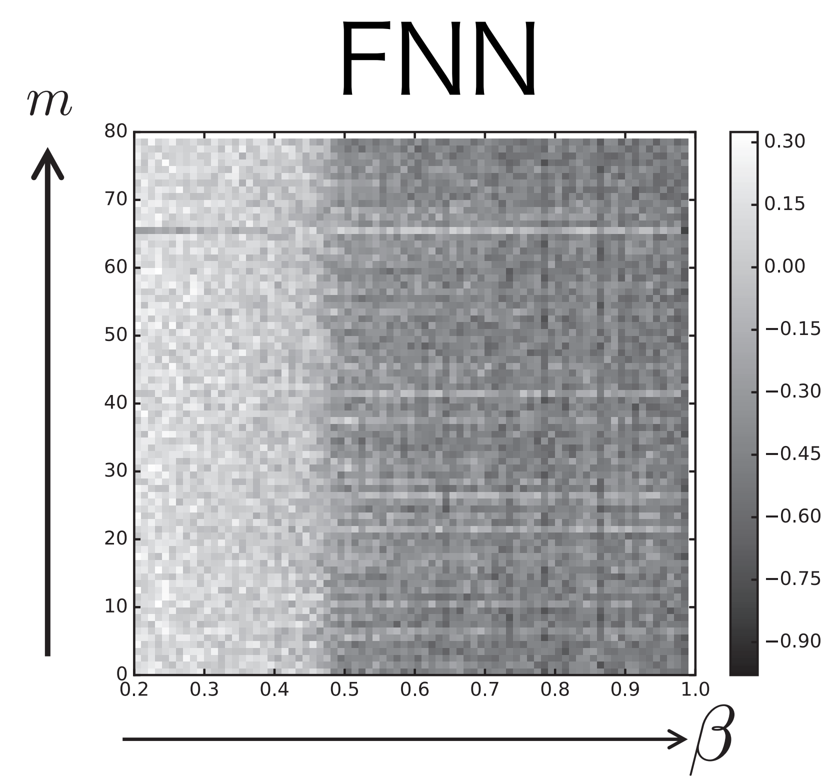
Fully connected results

arXiv:1609.09087, A. Tomiya

FNN gives Tc but CNN gives precise Tc

$$\left[\begin{array}{l} \mathcal{I} = \{\sigma_{xy} | \text{Ising config on } L \times L \text{ lattice.}\} \\ \downarrow \left\{\begin{array}{l} \text{Flatten} \\ \text{Fully connected } F \\ \text{Softmax} \end{array}\right. \\ [0,1]^{\text{Hidden}=80} \\ \downarrow \left\{\begin{array}{l} \text{Fully connected } W \\ \text{Softmax} \end{array}\right. \\ [0,1]^{N=100} = \mathcal{O} \end{array}\right]$$





System size	β_c (CNN)	β_c (FC)
8×8	0.478915	0.462494
16×16	0.448562	0.433915
32×32	0.451887	0.415596
$L \to \infty$	$\beta_c^{\text{Exact}} \sim 0.440686$	

CNN gives better Tc

What's the role of filters?

If filters have positive values, W in CNN has domain structure

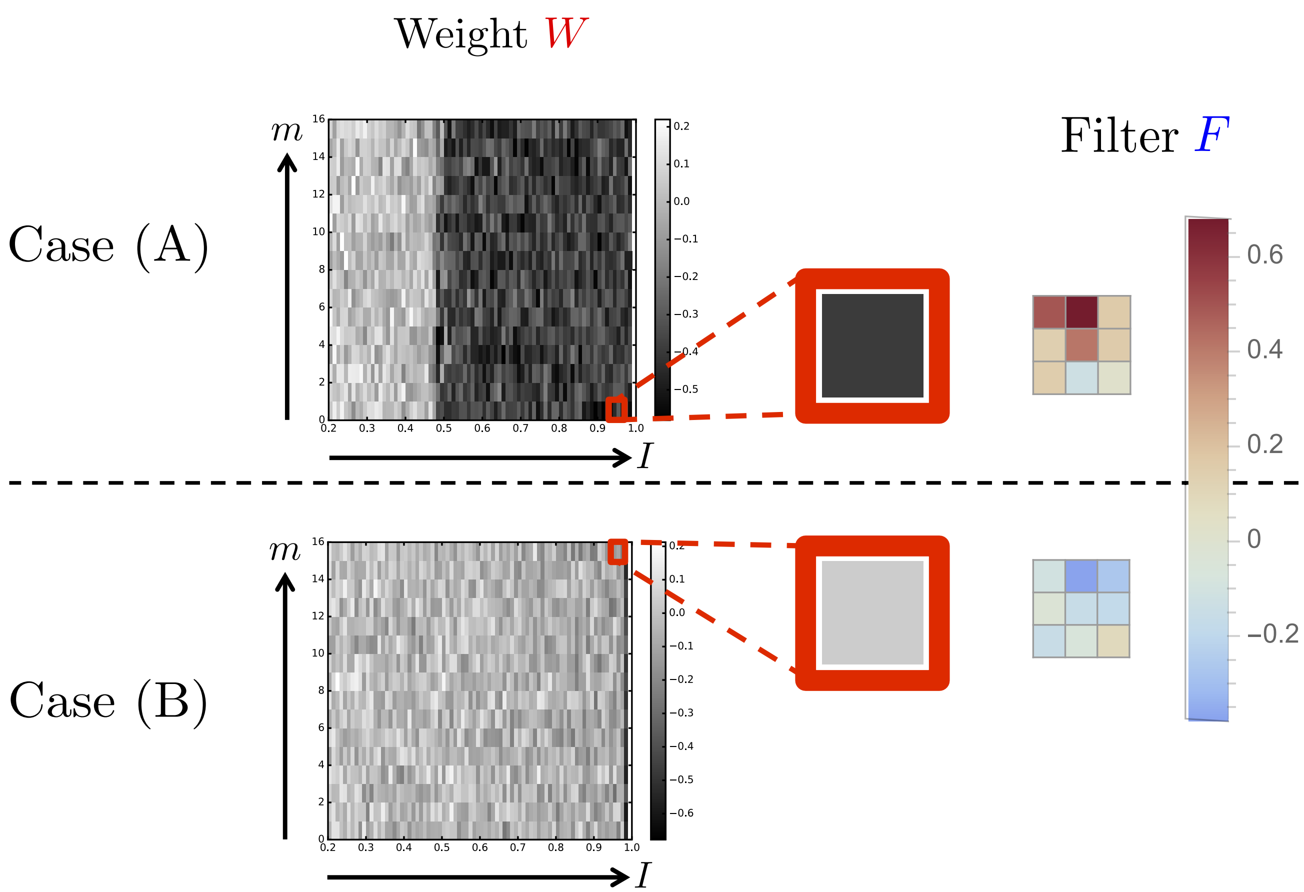


Fig. 3. (Color online) Heat maps of W_I^m and F_{ij} for the CNN with one filter. In case (A), there always exist two distinct regions (black and gray). In case (B), there is no such clear decomposition.

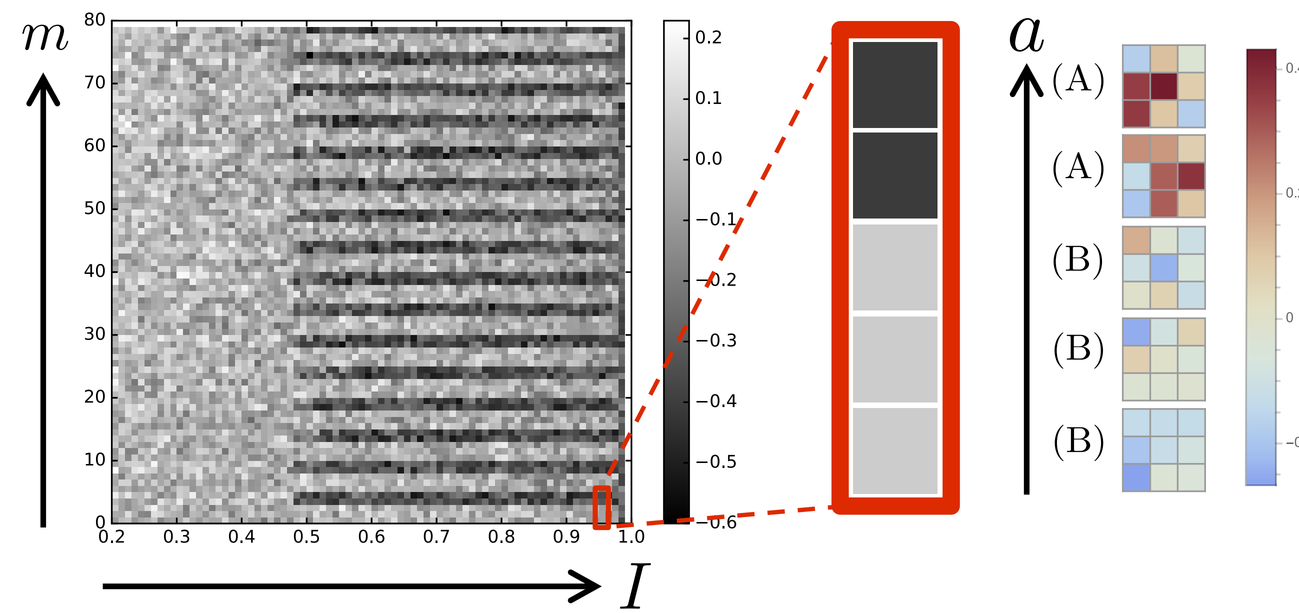


Fig. 4. (Color online) Heat maps of W_I^m and F_{ij}^a with five filters.