Large-scale lattice QCD simulation and application of machine learning, Univ. of Tsukuba, 2023/11/25

Machine Learning Topological Sector of Yang-Mills Theory

Masakiyo Kitazawa (YITP, Kyoto)

Matsumoto, MK, Kohno, PTEP2021, 023D01

MK, Kohno, in preparation

Topological Charge in YM Theory

Topological Charge

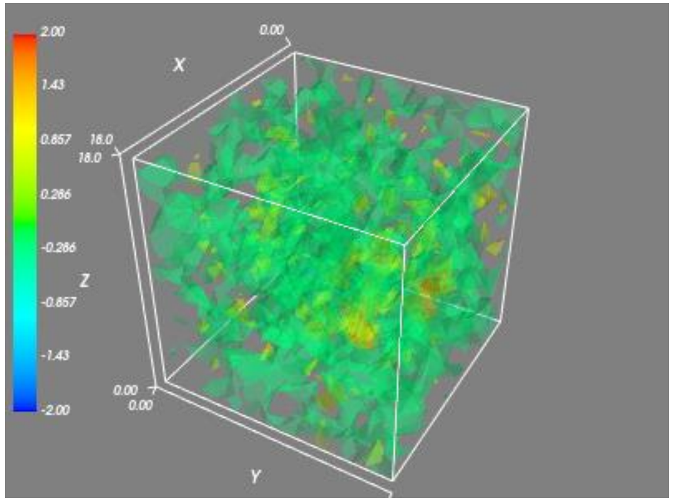
$$Q = \int d^4x \rho_q(x) \ : \text{integer}$$

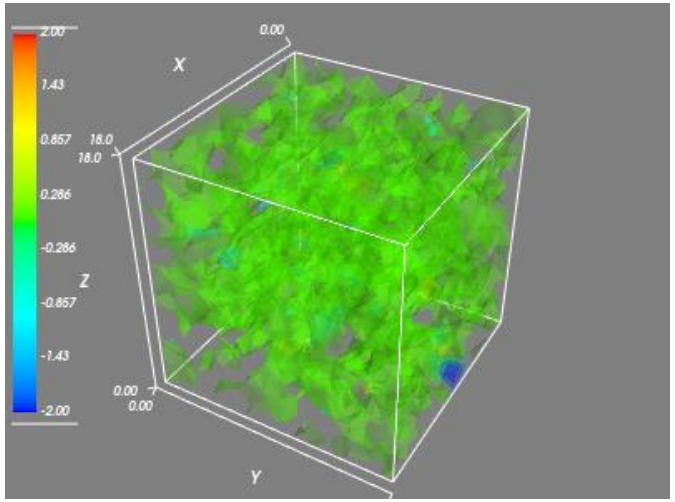
$$\rho_q(x) = -\frac{1}{32\pi^2}\mathrm{tr}[F_{\mu\nu}\tilde{F}_{\mu\nu}]$$

Motivations & Applications

- ☐ $U_A(1)$ anomaly
- ☐ Axion for cosmology
- ☐ Instantons
- ☐ Toplogical freezing in simulations

$\rho_q(x)$ in SU(3) YM

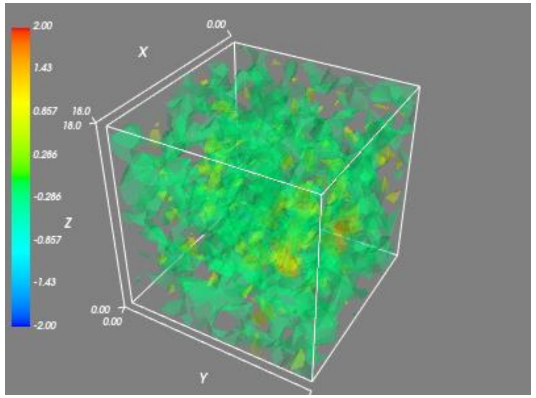




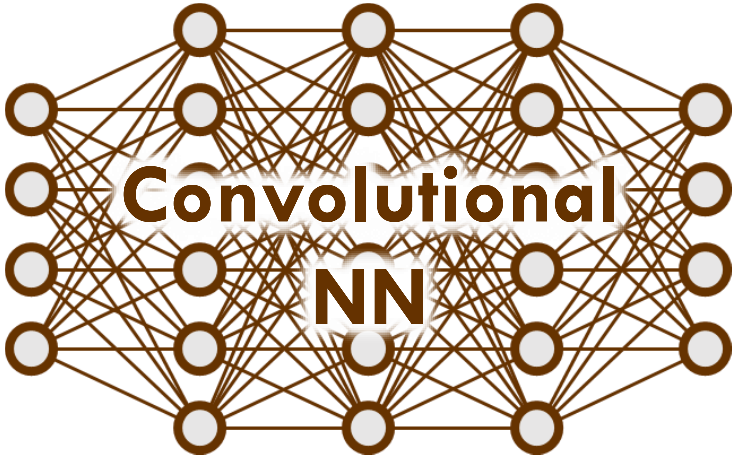
Machine Learning for measuring Q

Matsumoto, MK, Kohno, PTEP2021

Input



Observable on gauge conf.



Output

Q

Topological charge

Topological charge density on flowed field $\rho_q(x,t)$

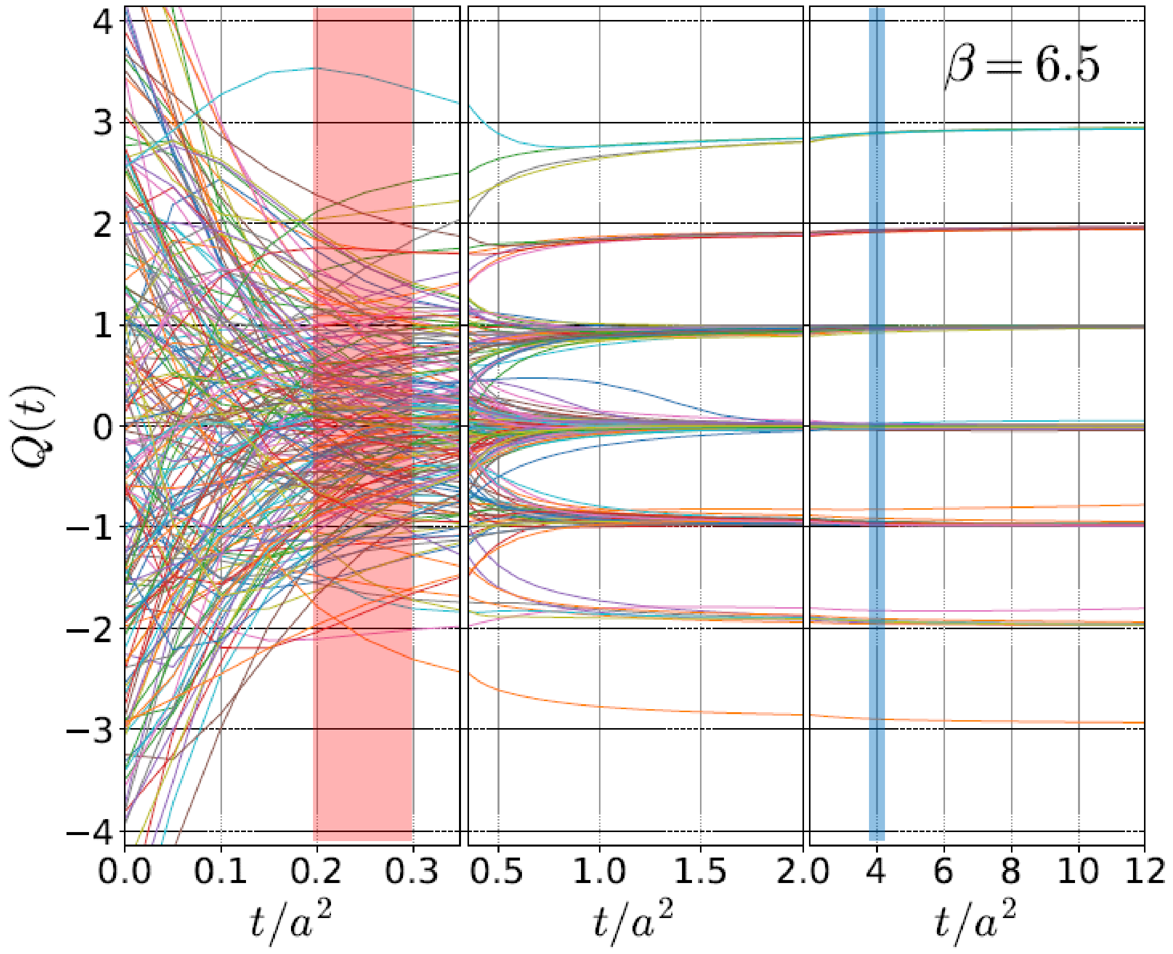
Q @ $t/a^2 = 4$

Aim:
- ☐ **Faster analysis of Q**
- ☐ **Recognizing objects in high-dim. space**

Input/Output Data

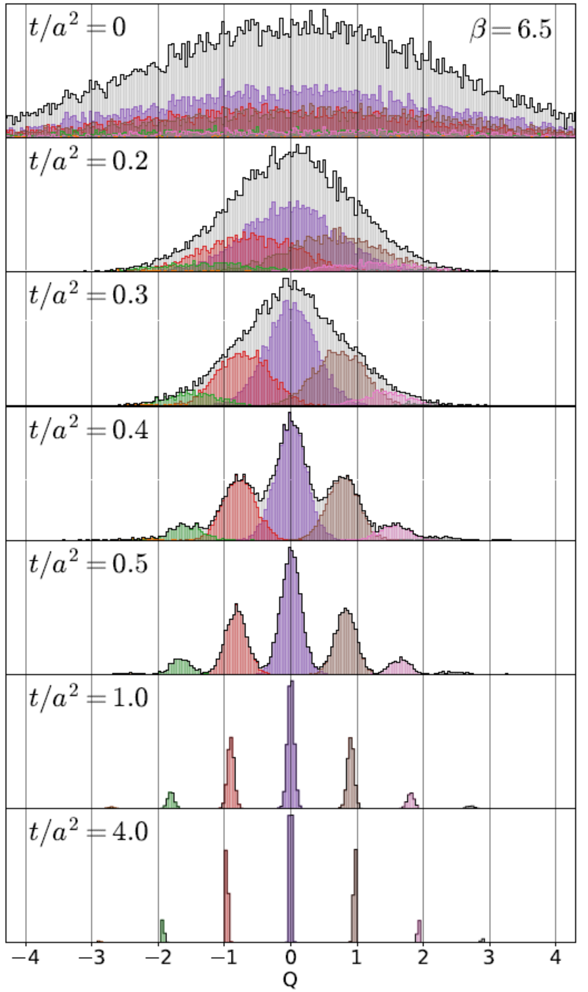
Matsumoto, MK, Kohno, PTEP2021

$Q(t)$ on the flowed field at flow time t



Input data:
topological charge density at small $t \simeq 0.2 - 0.3$

Answer:
spatial integral at $t/a^2 = 4$

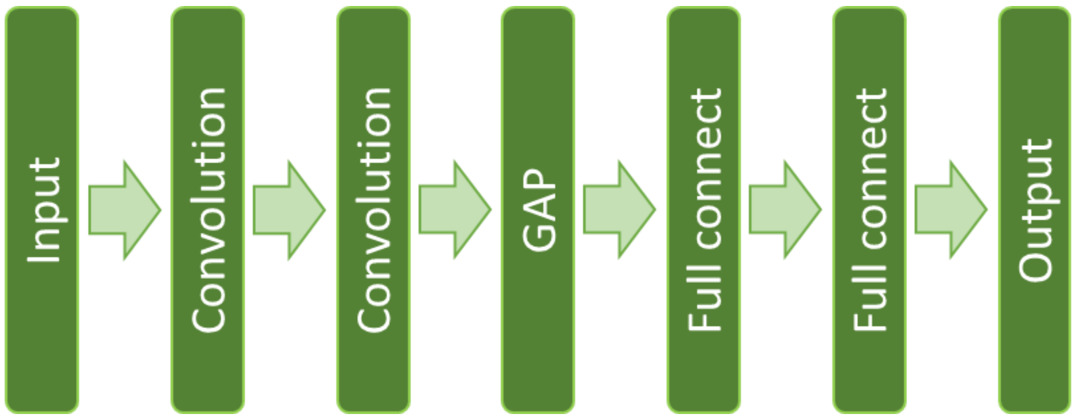


Setup/Result

Matsumoto, MK, Kohno, PTEP2021

Setup

- SU(3) YM Wilson action
- $\beta = 6.2, 6.5, 24^4$ $(T < T_c)$
- 20000 data
- average pooling to 12^4
- Multi-dim. CNN (Chainer)
- multiple t as multi-channel



Result

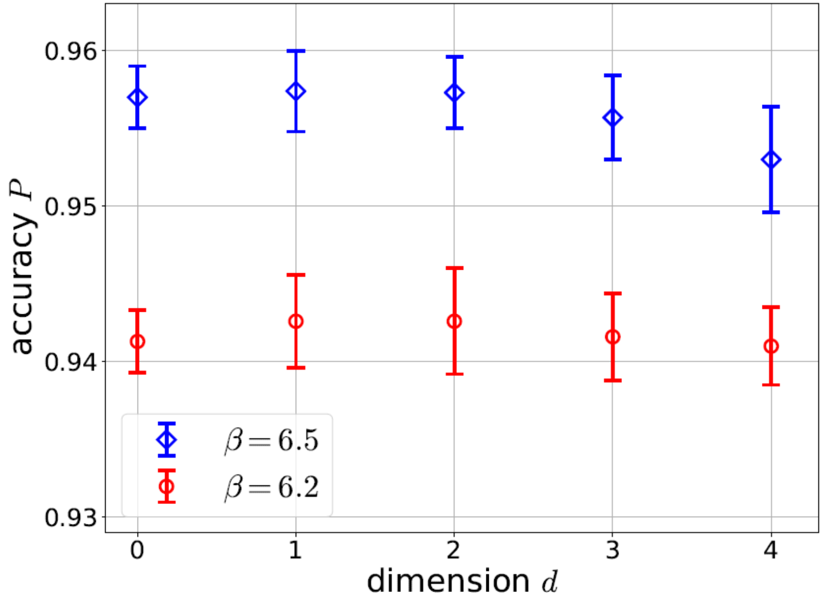
- **High accurate CNN model**
- Accuracy increases with increasing t.

Input t/a^2	$\beta = 6.2$	$\beta = 6.5$
0.45, 0.4, 0.35	0.974(2)	0.998(1)
0.4, 0.35, 0.3	0.975(2)	0.997(1)
0.35, 0.3, 0.25	0.967(2)	0.996(1)
0.3, 0.25, 0.2	**0.959(2)**	**0.990(2)**
0.25, 0.2, 0.15	0.939(3)	0.951(2)
0.2, 0.15, 0.1	0.864(3)	0.831(5)
0.15, 0.1, 0.05	0.692(4)	0.647(8)
0.1, 0.05, 0	0.538(5)	0.499(6)

Result

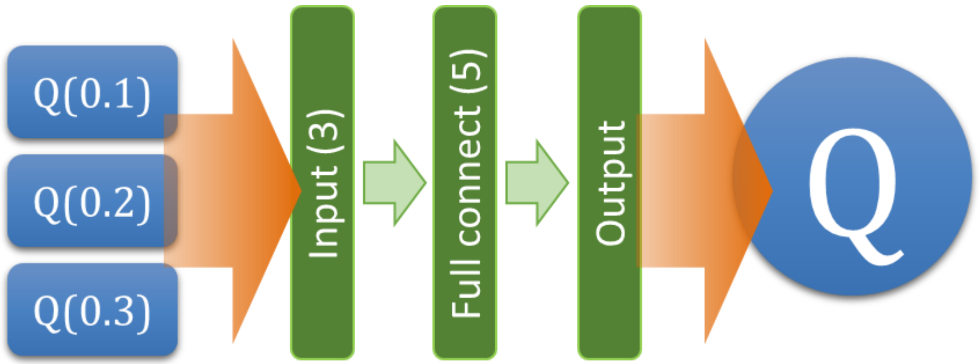
Matsumoto, MK, Kohno, PTEP2021

Dimensional Reduction



- ☐ Accuracy does not depend on d
- ☐ Spatial structure is not exploied

NN without spatial coordinates



High accuracy at small t:

Input t/a^2	$\beta=6.2$	$\beta=6.5$
0.45, 0.4, 0.35	0.974(2)	0.998(1)
0.4, 0.35, 0.3	0.975(2)	0.997(1)
0.35, 0.3, 0.25	0.967(2)	0.996(1)
0.3, 0.25, 0.2	**0.959(2)**	**0.990(2)**
0.25, 0.2, 0.15	0.939(3)	0.951(2)
0.2, 0.15, 0.1	0.864(3)	0.831(5)
0.15, 0.1, 0.05	0.692(4)	0.647(8)

NN just refers to spatial integral of $\rho_Q(t)$ → **Can NN recognize local structure?**

Strategy for 2nd Trial

MK, Kohno, in preparation

To realize the recognition of local structures

- **Normalization** (subtract the average) : $\tilde{\rho}_q(x) = \rho_q(x) - \bar{\rho}$
- **Data set @ $T > T_c$** → Range of T where DIGA is well justified
- Data at larger flow time t/a^2

Purpose

- Recognition of local structure in higher-dimensional space
- Direct confirmation of instantons
- Further applications in numerical studies

Setup

MK, Kohno, in preparation

- SU(3) YM
- Wilson gauge action, $\beta = 6.3$
- $36^3 \times 8$ $(T = 1.88T_c)$
- ~2500 data
- Approximately equal numbers for $|Q| = 0,1,2$

Q	-3	-2	-1	0	1	2	3	total
Data #	23	436	512	493	534	448	15	2461

- Average pooling: $36^3 \times 8 \rightarrow 18^3 \times 4$
- Padding by permutations of spatial axes by $\times 6$

Training Data

- $\rho_q(x)$ at $t/a^2 > 0$
- $t/a^2 = 0.3 - 0.5,\ 0.6 - 1.0$

→ **CNN** →

Answer

$Q(t = 4)$

CNN Models & Training

MK, Kohno, in preparation

Dimensional Reduction

- $3+1$: No reduction
- $3+0$: Integrate temporal direction
- $2+1$: Integrate one spatial direction

NN Structure

Layer	In ch	Out ch	Filter	Out size	Activation
conv1	1,2,3	4	3^4	4x4x18^3	sigmoid
conv2	4	8	3^4	8x4x18^3	sigmoid
conv3	8	16	3^4	16x4x18^3	sigmoid
GAP	-	-	4x18^3	16x1	-
full connect	-	-	16x1	1	-

4d CNN

Data Division

- Training: 1600×6
- Validation: 400×6
- Evaluation: 400×6

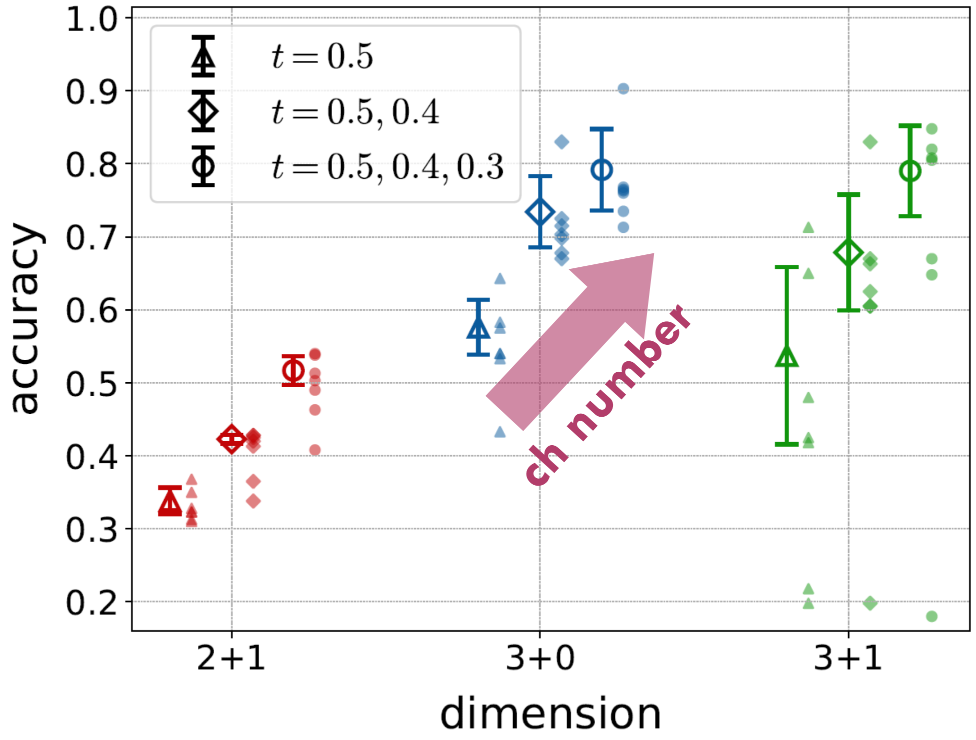
- **Data are divided randomly into training/validation/evalation**
- **Perform the training 7 times for each setting**

Setting/Parameters

Optimizer	Adam
Learning rate	0.001
Batchsize	16
Loss function	MSE
N_train	9600
N_val	1200
N_eval	1800
N_epoch	1500

Result: Accuracy

$t/a^2 \leq 0.5$



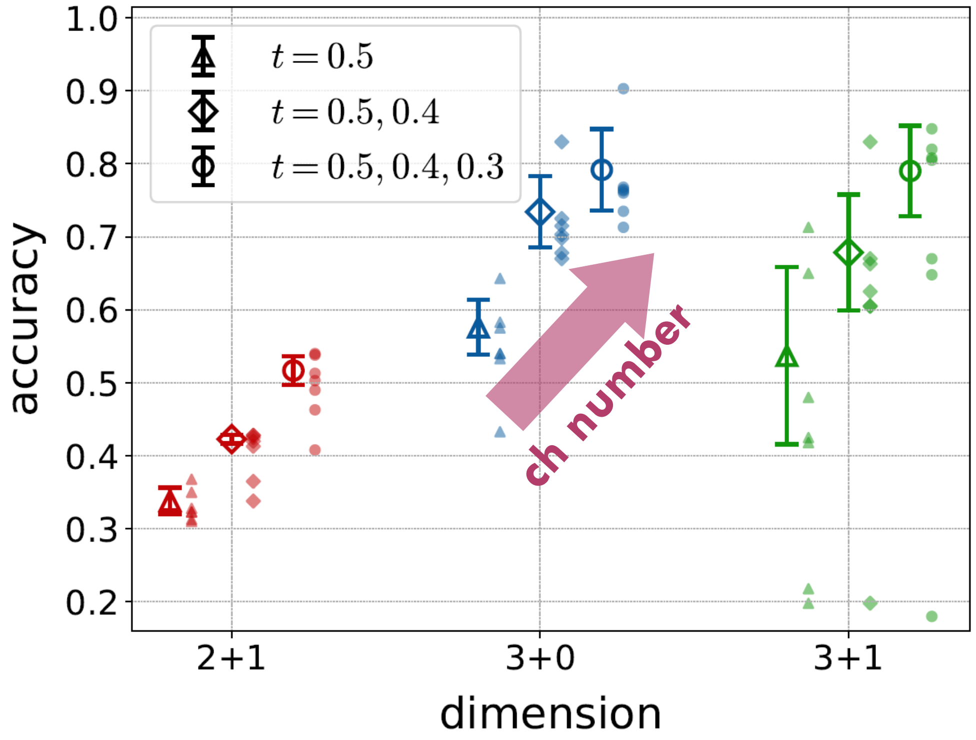
- Training for 7 times
- Big markers: Average and variance of top 5 results
- Small points: accuracy of each training

- Origin of errors
  - training data
  - procedure of training

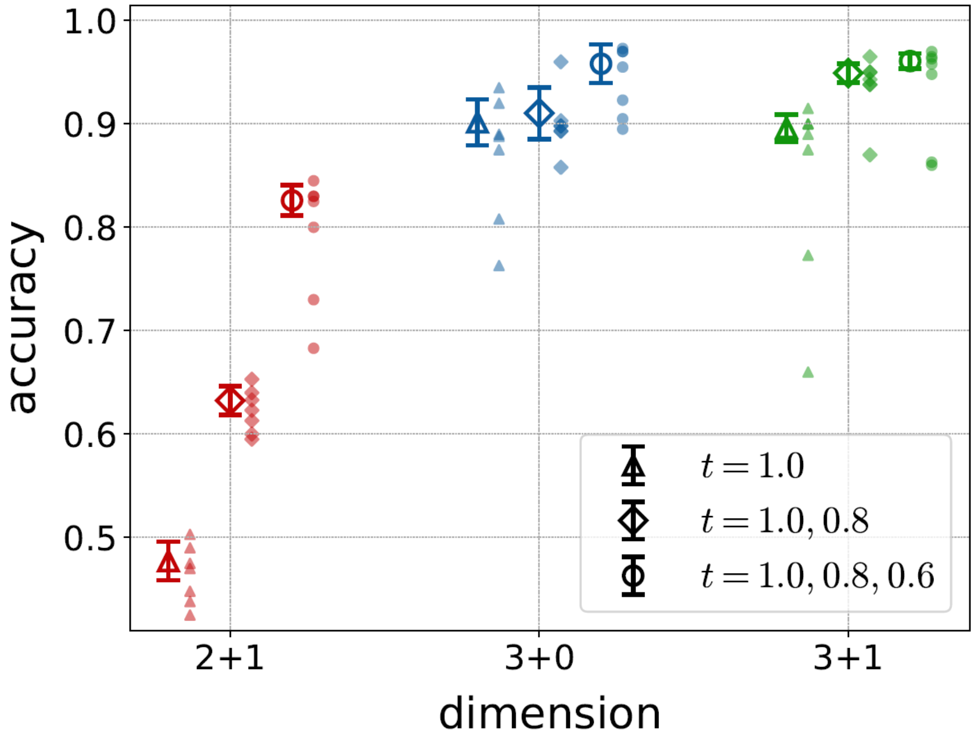
- worse accuracy at 2+1d ➔ **Recognition of 3d structure**
- increasing accuracy with multiple t ➔ **Recognition of 4d (space+flow) structure**
- unstable result for 3+1d? ➔Recognition of 5d (3+1+flow) space?

Result: Accuracy

$t/a^2 \leq 0.5$



$t/a^2 \leq 1.0$



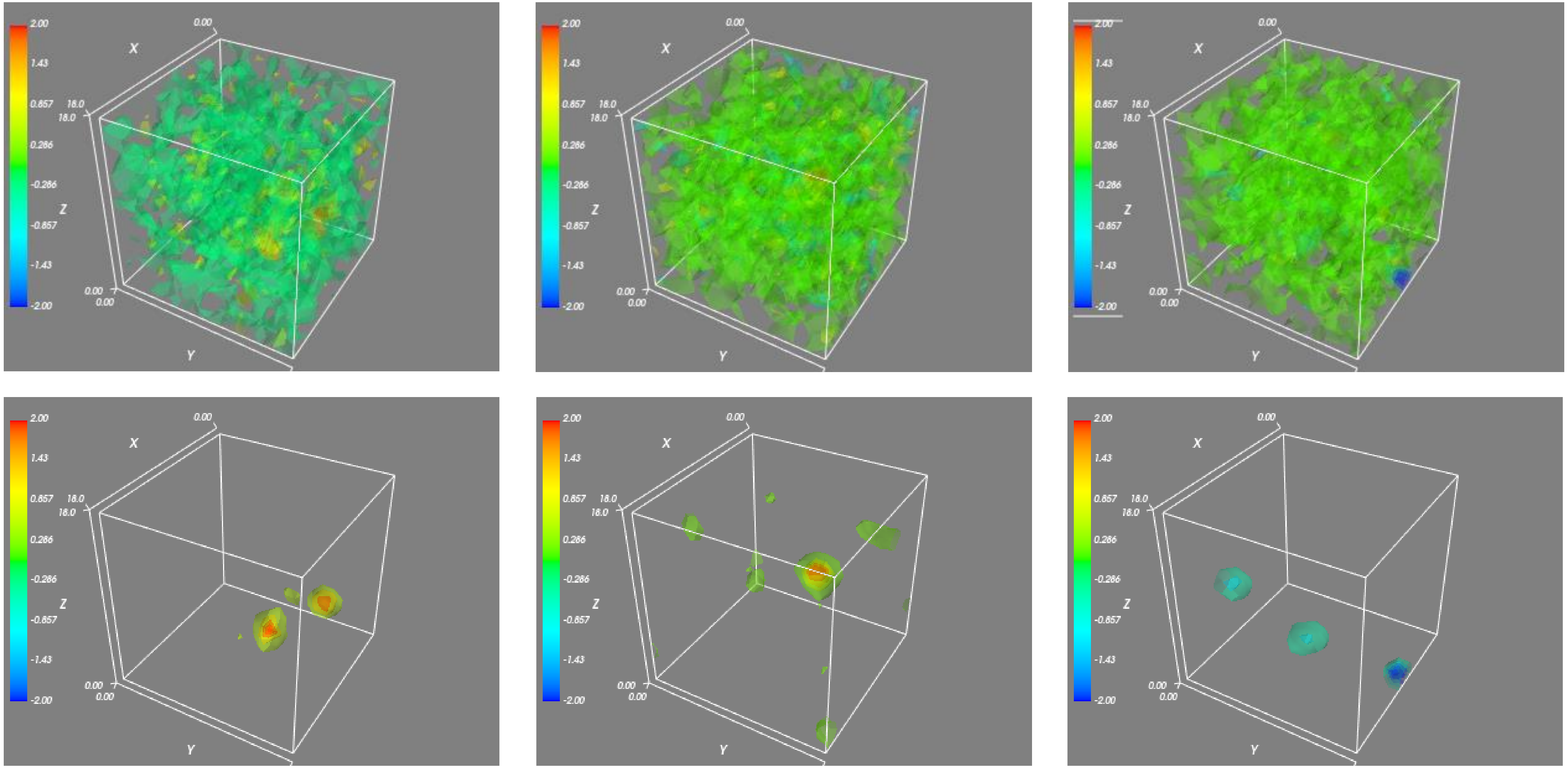
- worse accuracy at 2+1d ➔ **Recognition of 3d structure**
- increasing accuracy with multiple t ➔ **Recognition of 4d (space+flow) structure**
- unstable result for 3+1d? ➔Recognition of 5d (3+1+flow) space?
- $t/a^2 \leq 1$: saturation of accuracy

Topological Charge Density

$Q = 2$ $Q = 1$ $Q = -3$

$t/a^2 = 1.0$

$t/a^2 = 4.0$



Summary

- A CNN model to recognize local structure of the topological charge density
  - Supervised learning (Input: $\rho_Q(x,t)$, Output: $Q_{t=4}$)
  - SU(3) @ $T > T_c$
  - Removal the average from $\rho_Q(x,t)$ as preconditioning

- Accuracy increases as increasing the dimension of the data
- ➔Our CNN model recognizes structures in 4(+1)d space

Future

- Detailed structure of localized objects: location, size, form, etc.
- Training of vacuum configurations
  - understanding of vacuum structure of YM theory
- Other numerical applications