Deep-Learning-assisted chiral magnetic effect search in heavy-ion collisions

Yuan-Sheng Zhao(赵渊晟)
Fudan University

with Xu-Guang Huang, Kai Zhou

HENPIC online seminar, Feb. 24, 2021

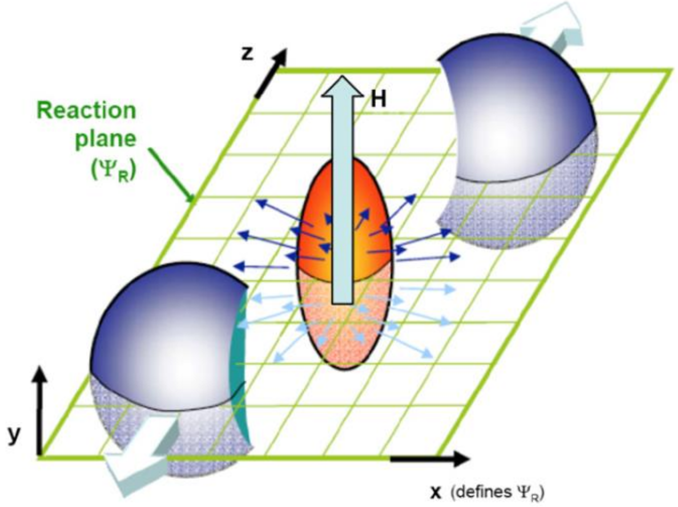
Outline

- Motivation
  - CME and its current observables
- Introduction
  - Deep learning & Convolutional neural network (CNN)
  - Our target
  - AMPT
- Modifications to the training procedure
- Results
- Summary and outlook

Motivation

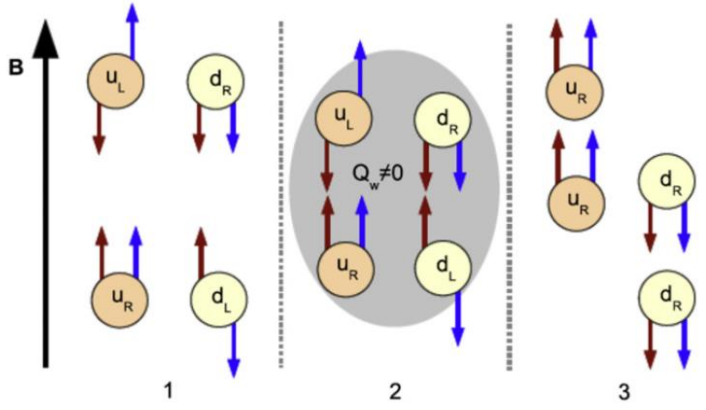
- CME and its current observables

Chiral magnetic effect(CME)



Strong Magnetic Field

$\mu_5 \neq 0$ topological non-trivial bubbles



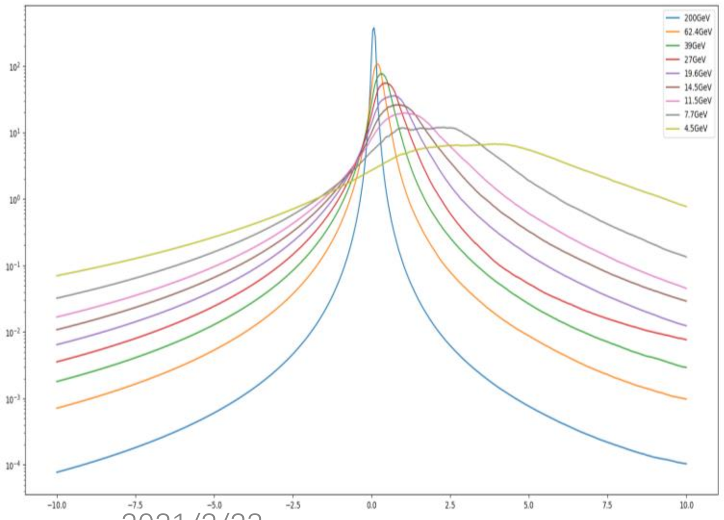
PhysRevD.78.074033
K Fukushima, DE Kharzeev, HJ Warringa

$$\overrightarrow{J_5} = \frac{e^2 \mu_5}{2\pi^2} \vec{B}$$

Chiral Magnetic Effect (CME)

Evolution

Charge separation (CS)



B v.s. time

Observables

- γ, $\Delta\gamma$
- Event-shape-engineering
- ΔS
- Invariant mass
- Spectator event plane
- …

Background effect

- Transverse momentum conservation
- Local charge conservation
- Elliptic flow
- …

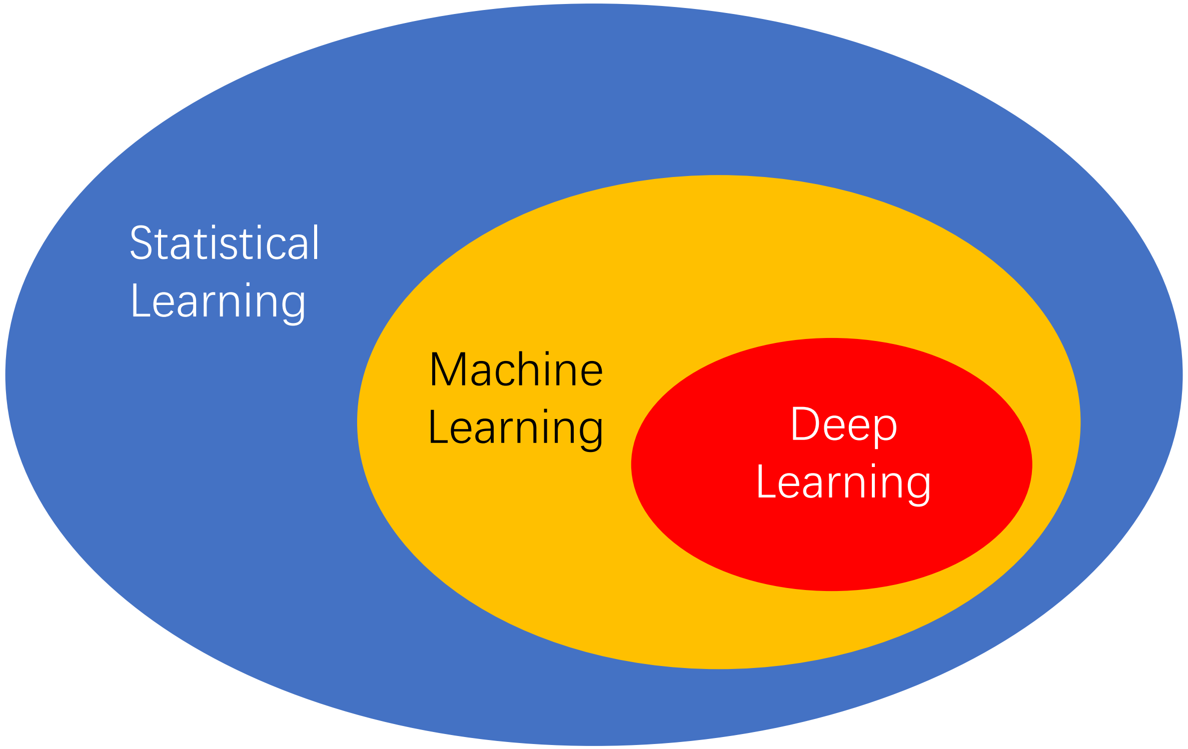
What dose CME remain after freeze out?

Introduction

- Deep learning & Convolutional neural network (CNN)
- Our target
- AMPT

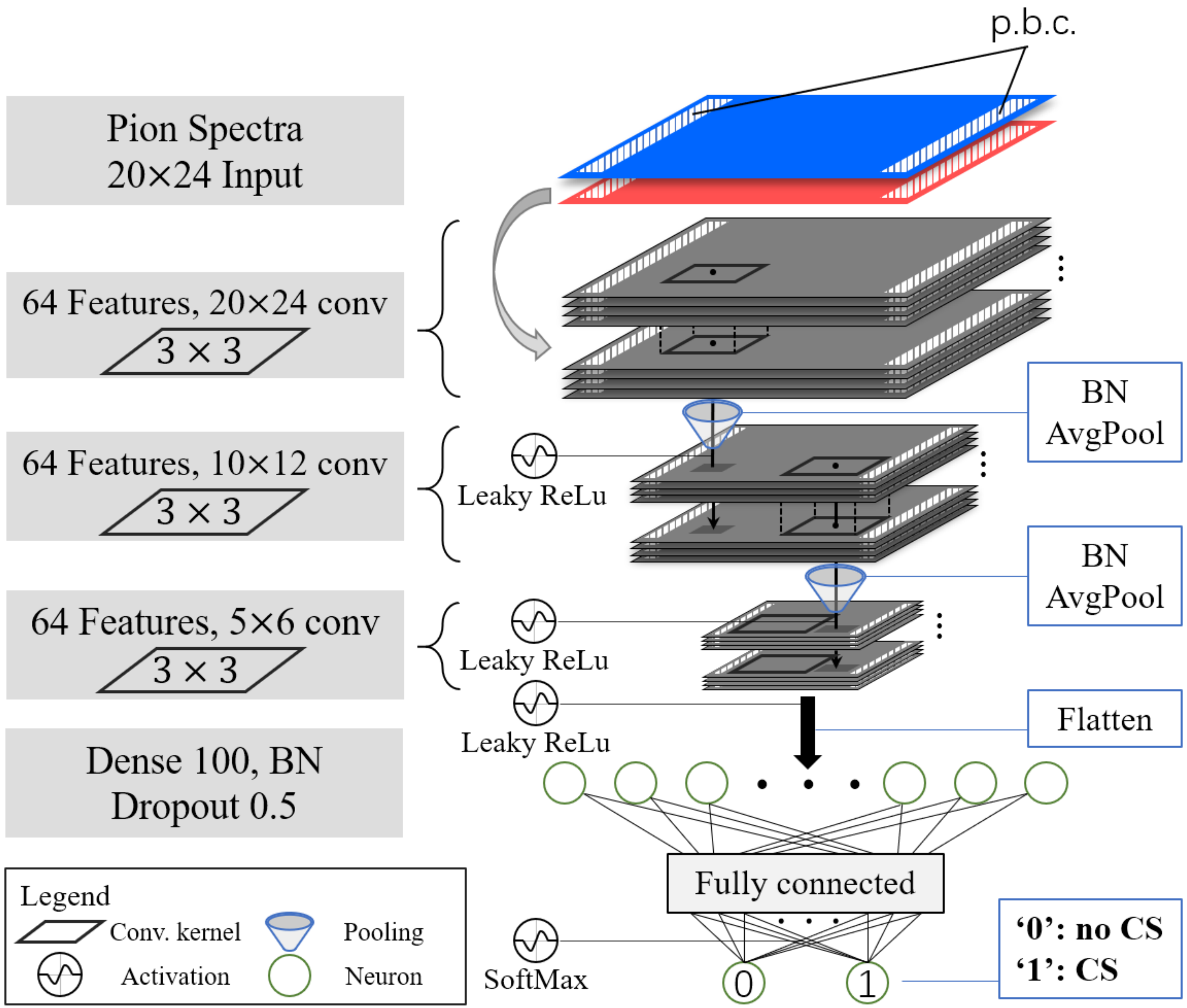
Deep learning

- Statistical learning: Model fixed, fit parameters, like Bayesian analysis.
- Machine learning: Neurons(linear)+Activations(non-linear). No fixed model.
- Deep learning: multiple neural layers



CNN

- Universal Approximation theorem
- Any f with proper NN



Deep learning × HIC: Previous research

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-02726-3 OPEN

An equation-of-state-meter of quantum chromodynamics transition from deep learning

Applications of deep learning to relativistic hydrodynamics

Hengfeng Huang,[1,2] Bowen Xiao,[3] Huixin Xiong,[1] Zeming Wu,[1,2] Yadong Mu,[3,4,*] and Huichao Song[1,2,5,†]

epartment of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing 100871, China
[2]Collaborative Innovation Center of Quantum Matter, Beijing 100871, China
[3]Institute of Computer Science and Technology, Peking University, Beijing 100080, China
[4]Center for Data Science, Peking University, Beijing 100871, China
[5]Center for High Energy Physics, Peking University, Beijing 100871, China

(Dated: April 24, 2018)

Relativistic hydrodynamics is a powerful tool to simulate the evolution of the quark gluon plasma (QGP) in relativistic heavy ion collisions. Using 10000 initial and final profiles generated from 2+1-d relativistic hydrodynamics VISH2+1 with MC-Glauber initial conditions, we train a deep neural network based on stacked U-net, and use it to predict the final profiles associated with various initial

A fast centrality-meter for heavy-ion collisions at the CBM experiment

Manjunath Omana Kuttan,[1,2,*] Jan Steinheimer,[1] Kai Zhou,[1,†] Andreas Redelbach,[1,3] and Horst Stoecker[1

[1]Frankfurt Institute for Advanced Studies, D-60438 Frankfurt am Main, Germany
[2]Institut für Theoretische Physik, Johann Wolfgang Goethe Universität, D-60438 Frankfurt am Main, Germany
[3]Institut für Informatik, Johann Wolfgang Goethe Universität, D-60438 Frankfurt am Main, Germany
[4]GSI Helmholtzzentrum für Schwerionenforschung GmbH, D-64291 Darmstadt, Germany

(Dated: October 29, 2020)

A new method of event characterization based on Deep Learning is presented. The PointNet models can be used for fast, online event-by-event impact parameter determination at the CBM experiment. For this study, UrQMD and the CBM detector simulation are used to generate Au+Au collision events at 10 *A*GeV which are n used to train and evaluate PointNet based architectures. The models can be trained on features like the position of particles in the CBM detector planes, tracks reconstructed from the hits or combinations thereof. : Deep Learning models reconstruct impact parameters from 2-14 fm with a mean error varying from -0.33

Deep learning jet modifications in heavy-ion collisions

Yi-Lun Du, Daniel Pablos and Konrad Tywoniuk

Deep learning × HIC: Previous research

- Phase transition
  - L. -G Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, Classify QCD phase transition with deep learning.
- Determine impact parameter / centrality
  - M. O. Kuttan, J. Steinheimer, K. Zhou, A. Redelbach, H. Stoecker, Deep Learning Based Impact Parameter Determination for the CBM Experiment.
- Equation of state
  - L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stocker, X.-N. Wang, An equation-of-state-meter of quantum chromodynamics transition from deep learning, Nature Commun. 9 (1) (2018) 210.
- Hydrodynamics
  - H. Huang, B. Xiao, H. Xiong, Z. Wu, Y. Mu, H. Song, Applications of deep learning to relativistic hydrodynamics.
- Jet
  - Y.-T. Chien, R. Kunnawalkam Elayavalli, Probing heavy ion collisions using quark and gluon jet substructure.
- …

CNN: how to work

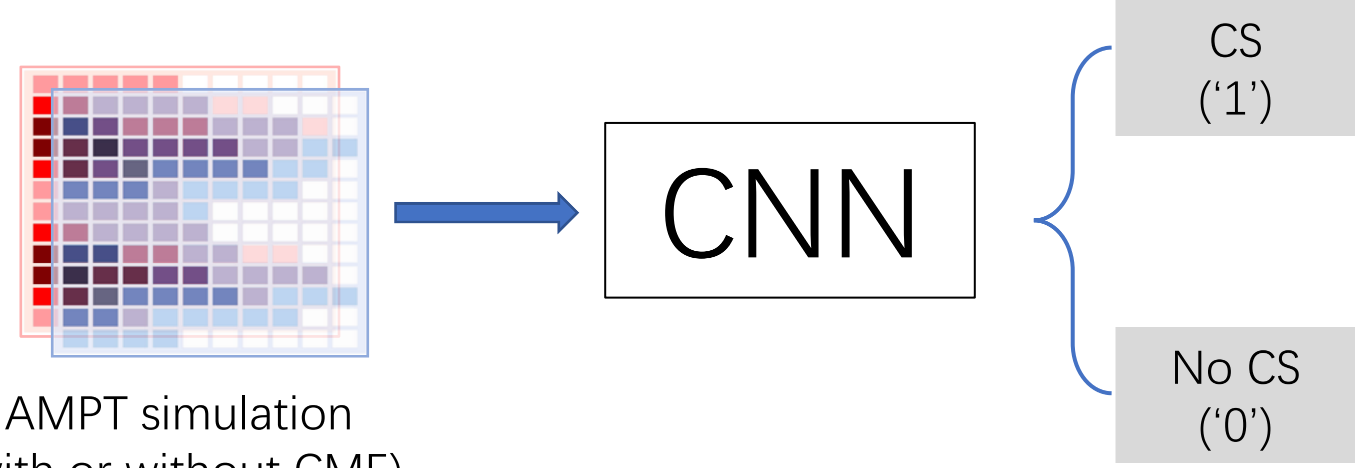
- Neural network
- Loss function: Cross entropy

$$H = \frac{1}{N}\sum -y_i \log(p_i)$$

- Optimizer: Adam, with tuned learning rate schedule
- Output: SoftMax
- Data: Training set / Validation set / Test set
- Periodic boundary condition
  - Cylindrical & torus Conv2D layers

Our target

- A supervised learning that can distinguish whether an event(Au+Au) has CS
- Insights into the trained network for physical understandings



AMPT simulation
(with or without CME)

CNN: how to work

- Neural network
- Loss function: Cross entropy

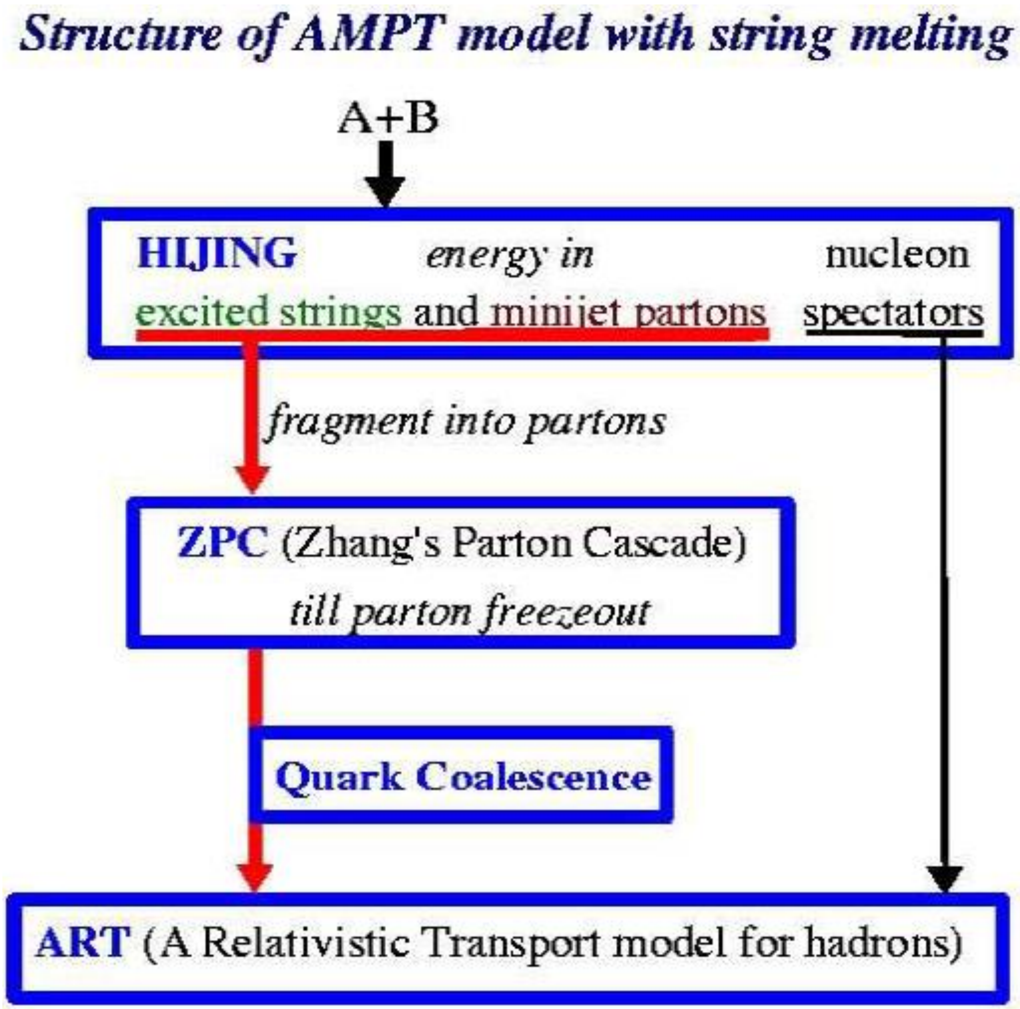
$$H = \frac{1}{N} \sum -y_i \log(p_i)$$

- Optimizer: Adam, with tuned learning rate schedule
- Output: SoftMax
- Data: Training set / Validation set / Test set
- Periodic boundary condition
    - Cylindrical & torus Conv2D layers

A multiphase transition(AMPT) model

- Simulation of nuclear-nuclear collision event
- CME not included

- The method by Guo-Liang Ma and Bin Zhang: switching p_y of a certain fraction of partons before ZPC



Zi-Wei Lin *etc.*, arXiv:nucl-th/0411110v3

Modifications with respect to CME

- Training set
- Boundary condition

Training set

- Events are pre-processed into the spectra of π^+ and π^-(20*24):

$$\rho^{\pm}(p_T, \phi)$$

- Generating at training:
  - For every batch, randomly pick a set of simulation condition(a Blue Box)
  - From the 50,000 events in the chosen Blue Box, randomly pick 100 events' pion spectra, and average them into a mixed event.

f		$\sqrt{s_{NN}}$ (GeV)						
Centrality		11.5	14.5	19.6	27	39	62.4	200
	0-10							
	10-20							
	20-30							
	30-40							
	40-50							
	50-60							

Every Orange or Blue Box corresponding to 50,000 single events

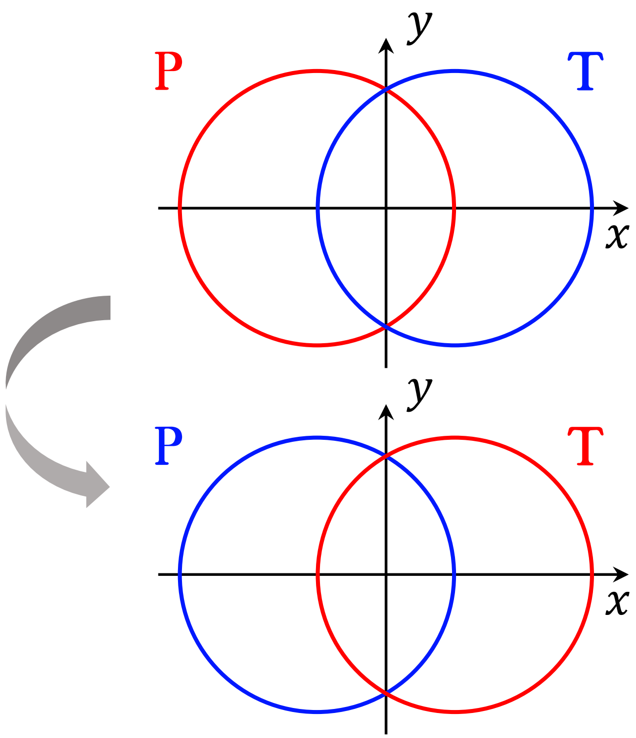
$f = 0$: No CME, Label '0'

$f > 0$: With CME, Label '1'

- Large fluctuation → Statistically better
- Batch average v.s. pre-averaged data
- 'Average knowledge' or 'typical behavior' of charge separation.

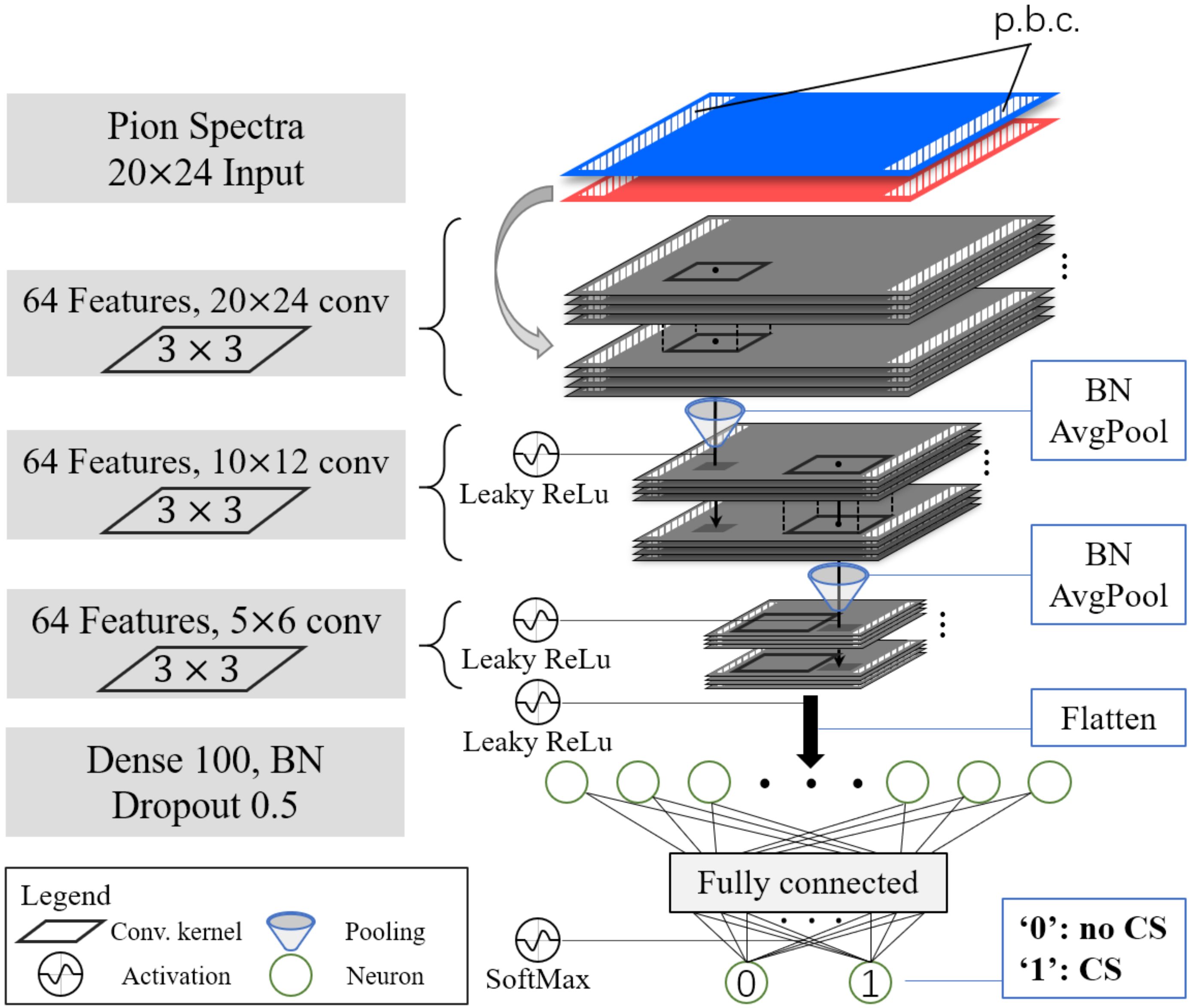
Training set

- Mirror symmetry along y-axis
  - Corresponding exchanging target and projectile
- Normalization
- Validation set
  - Drag events from every Orange and Blue Box
  - 100 average events for every box



Boundary condition

- $\rho^{\pm}(p_T, \phi)$, angular distribution
- Periodic boundary condition for ϕ
  - Cylindrical Conv2D layers



Results

- Accuracy of two trained NN: 0+5% and 0+10%
- Robustness against centrality and $\sqrt{\boldsymbol{s_{NN}}}$
- More CS tests
- Comparison to $\Delta\gamma$
- Prediction against elliptic flow
- Isobar results
- Visualization-Deep Dream

Accuracy and Prediction($\mathbf{P_1}$)

- The output of NN: (P_0, P_1) for a single event
  - P_0 is the probability of 'no initial CS'
  - P_1 is the probability of 'undergone CS'
  - $P_0 + P_1 = 1$
- '0' if $P_0 > P_1$, '1' if $P_1 > P_0$
  - P_1 can be a measure of CS strength
- Accuracy is defined as:

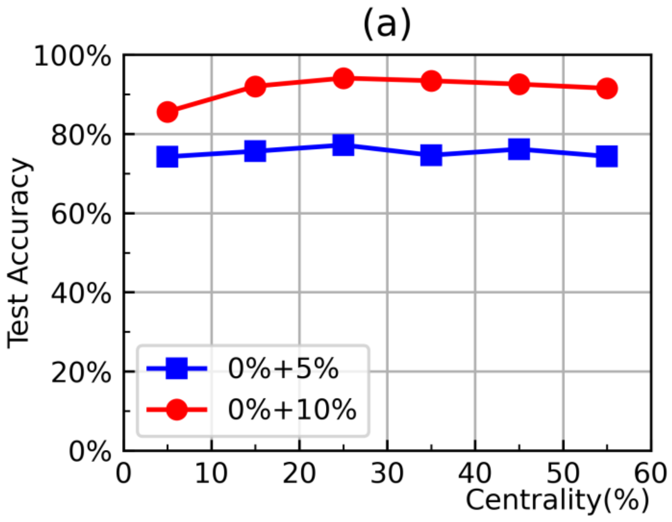
$$\frac{No.\,correct\ tests}{No.\,tests}$$

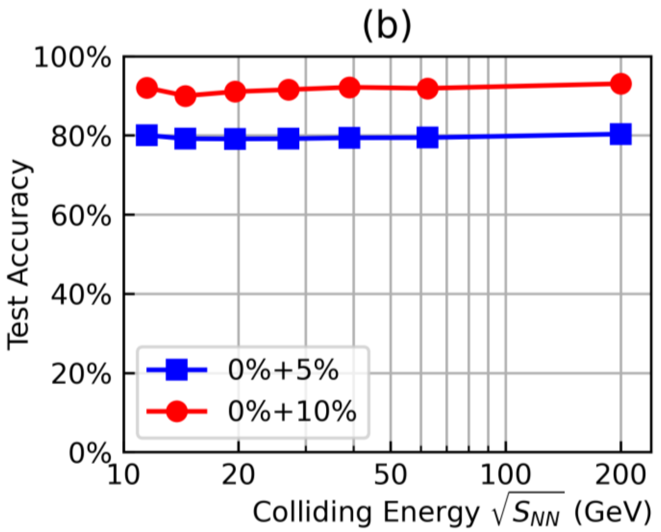
Accuracy & Robustness

NN	0+5%	0+10%
Accuracy (Under training cond.)	~80%	~92%

- $f = 5\%$ samples have larger similarity with $f = 0$ samples

Left: $\sqrt{s_{NN}}$ are mixed
Right: centrality are mixed

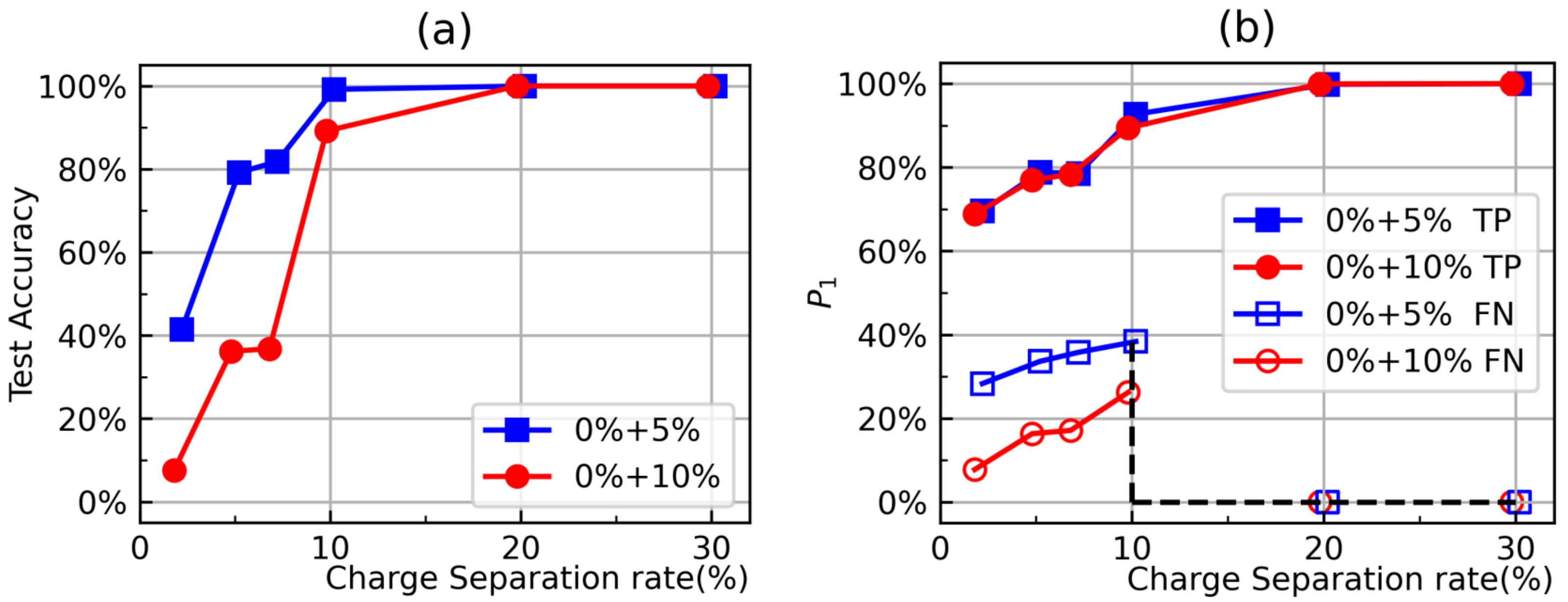




More CS tests

- f = 2%, 5%, 7%, 10%, 20%, 30%, all in '1' class

Left: Accuracy v.s. f
Right: P_1 v.s. f



$\sqrt{s_{NN}} = 39GeV$

Comparison to $\Delta\gamma$

$$\gamma_{same} = \left\langle \cos\left(\phi_\alpha^{(\pm)} + \phi_\beta^{(\pm)} - 2\Phi_R\right)\right\rangle$$

$$\gamma_{opp} = \left\langle \cos\left(\phi_\alpha^{(\pm)} + \phi_\beta^{(\mp)} - 2\Phi_R\right)\right\rangle$$

$$\Delta\gamma = \gamma_{opp} - \gamma_{same}$$

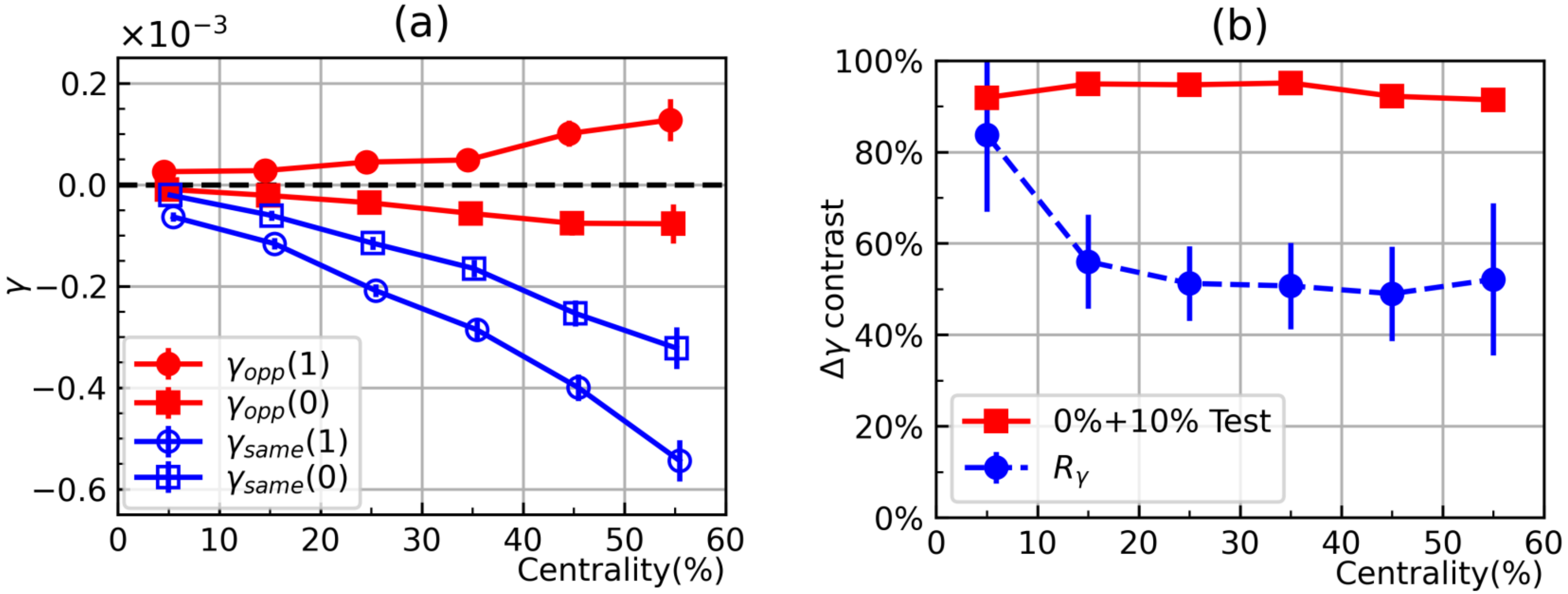
- Contrast of $\Delta\gamma$

$$R_\gamma = \frac{|\Delta\gamma(1) - \Delta\gamma(0)|}{|\Delta\gamma(1)| + |\Delta\gamma(0)|}$$

Comparison to $\Delta\gamma$

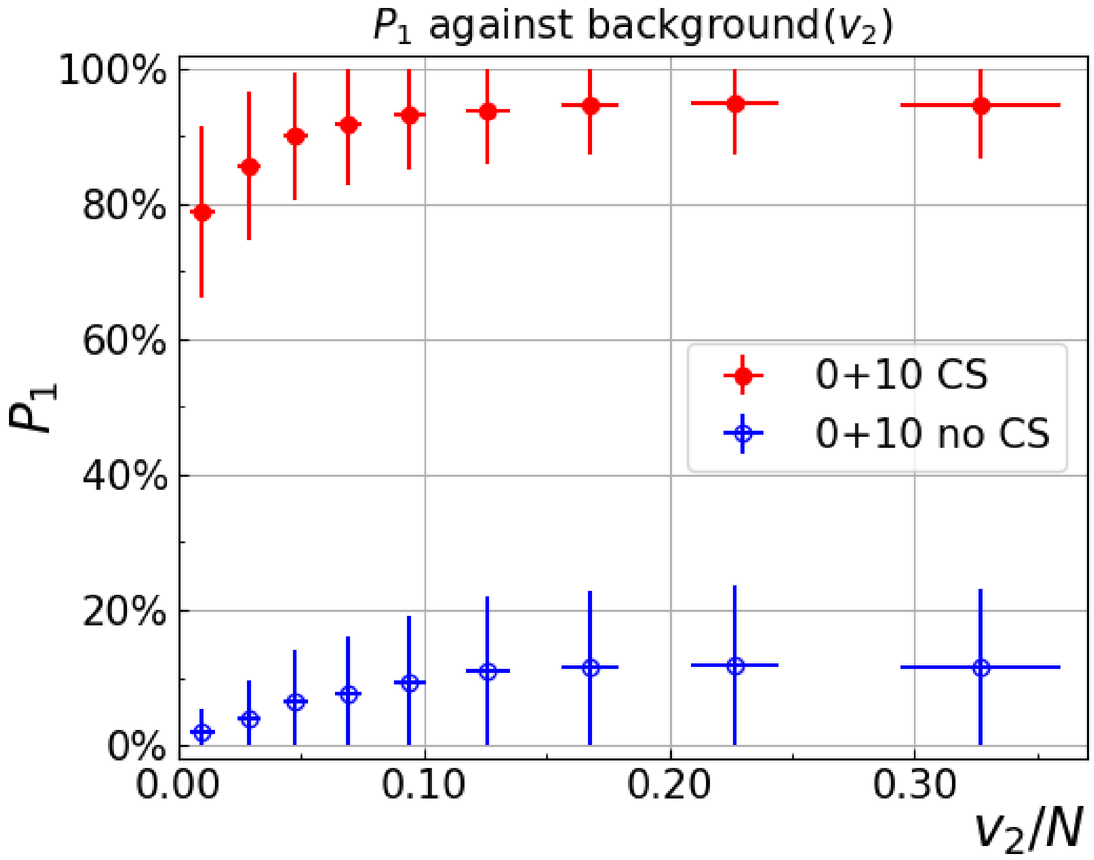
- $R_\gamma = \frac{|\Delta\gamma(1) - \Delta\gamma(0)|}{|\Delta\gamma(1)| + |\Delta\gamma(0)|}$, $R_{CNN} = \left|\frac{\langle P_1(1)\rangle - \langle P_1(0)\rangle}{\langle P_1(1)\rangle + \langle P_1(0)\rangle}\right|$
- 0%+10%

$\sqrt{s_{NN}} = 39GeV$



Compare to v_2

- CME pattern is correlated with v_2 (both indicating anisotropy)
- Lower v_2, smaller P_1, more uncertainty for CS class
- Significance: $\sim 5\sigma$ at large v_2



Isobar results

- ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$
- Same nuclei number, different proton number $\rightarrow$
- Same background, different magnetic field $\rightarrow$
- Different CS!

Overfitting

- More parameters, more likely to be overfitting
- Generalization to not learned data is subtle.(Which shall base on large number of training samples & ML techniques)

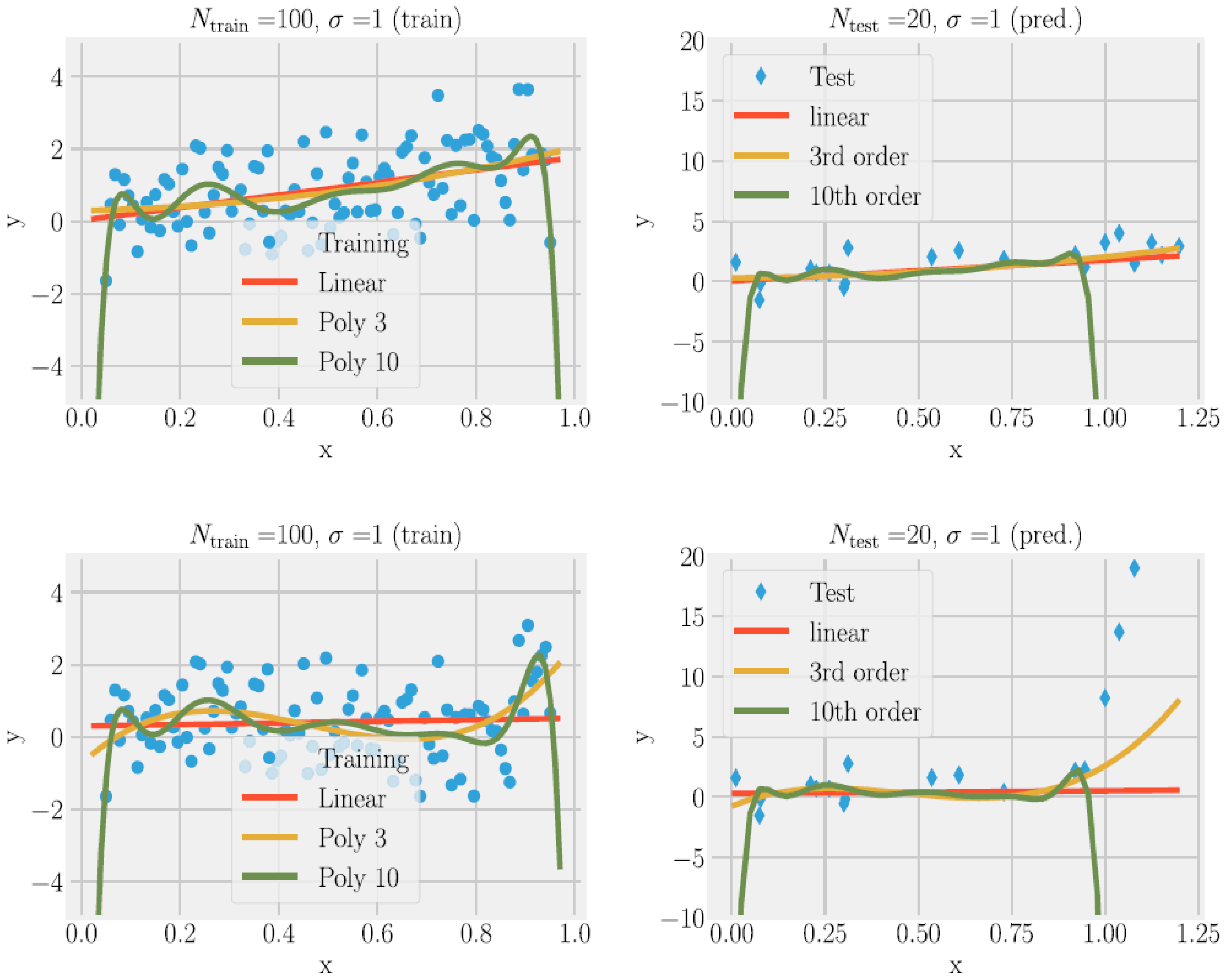
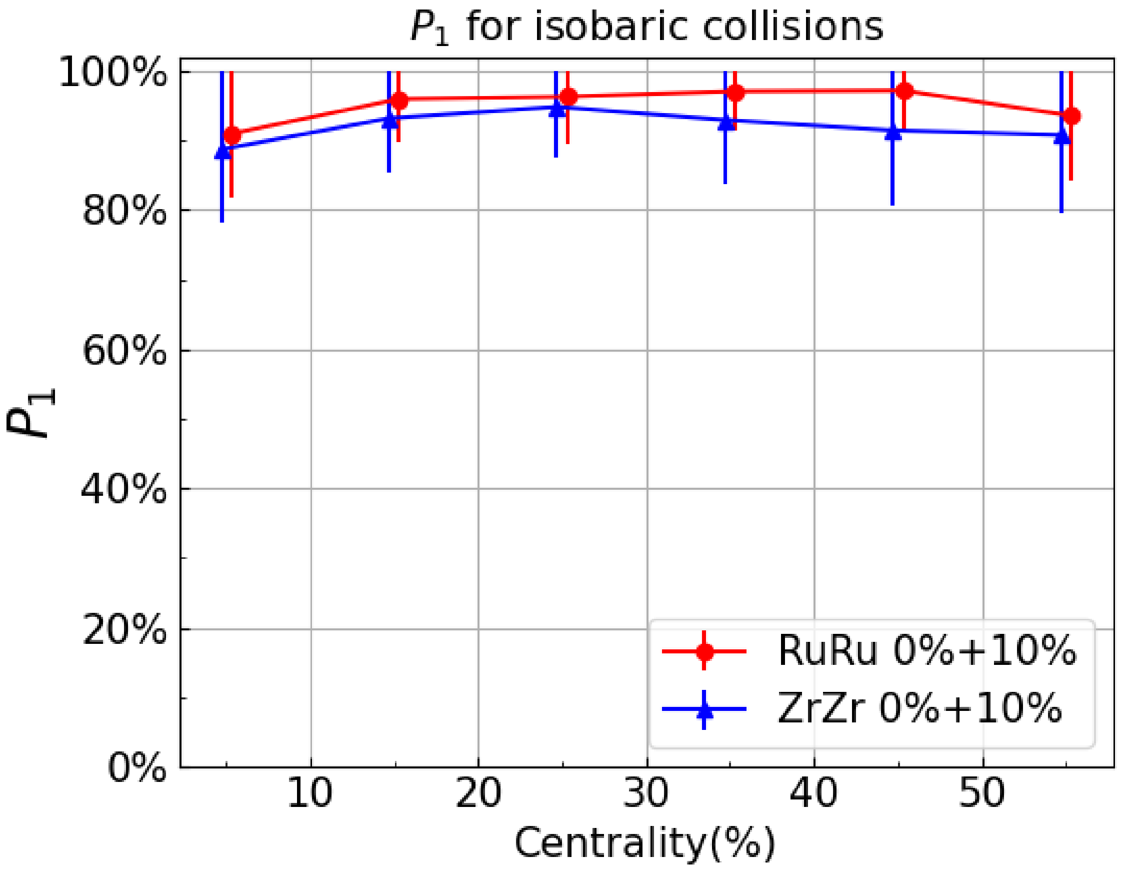


Fig. 2. Fitting versus predicting for noisy data. $N_{\mathrm{train}} = 100$ noisy data points ($\sigma = 1$) in the range $x \in [0, 1]$ were generated from a linear model (top) or tenth-order polynomial (bottom). This data was fit using three model classes: linear models (red), all polynomials of order 3 (yellow), all polynomials of order 10 (green) and used to make prediction on $N_{\mathrm{test}} = 20$ new data points with $x_{\mathrm{test}} \in [0, 1.2]$(shown on right). Notice that even when the data was generated using a tenth order polynomial, the linear and third order polynomials give better out-of-sample predictions, especially beyond the x range over which the model was trained.

Pankaj Mehta *etc.*, Phys. Reports, 810 (2019) 1–124

Isobar results

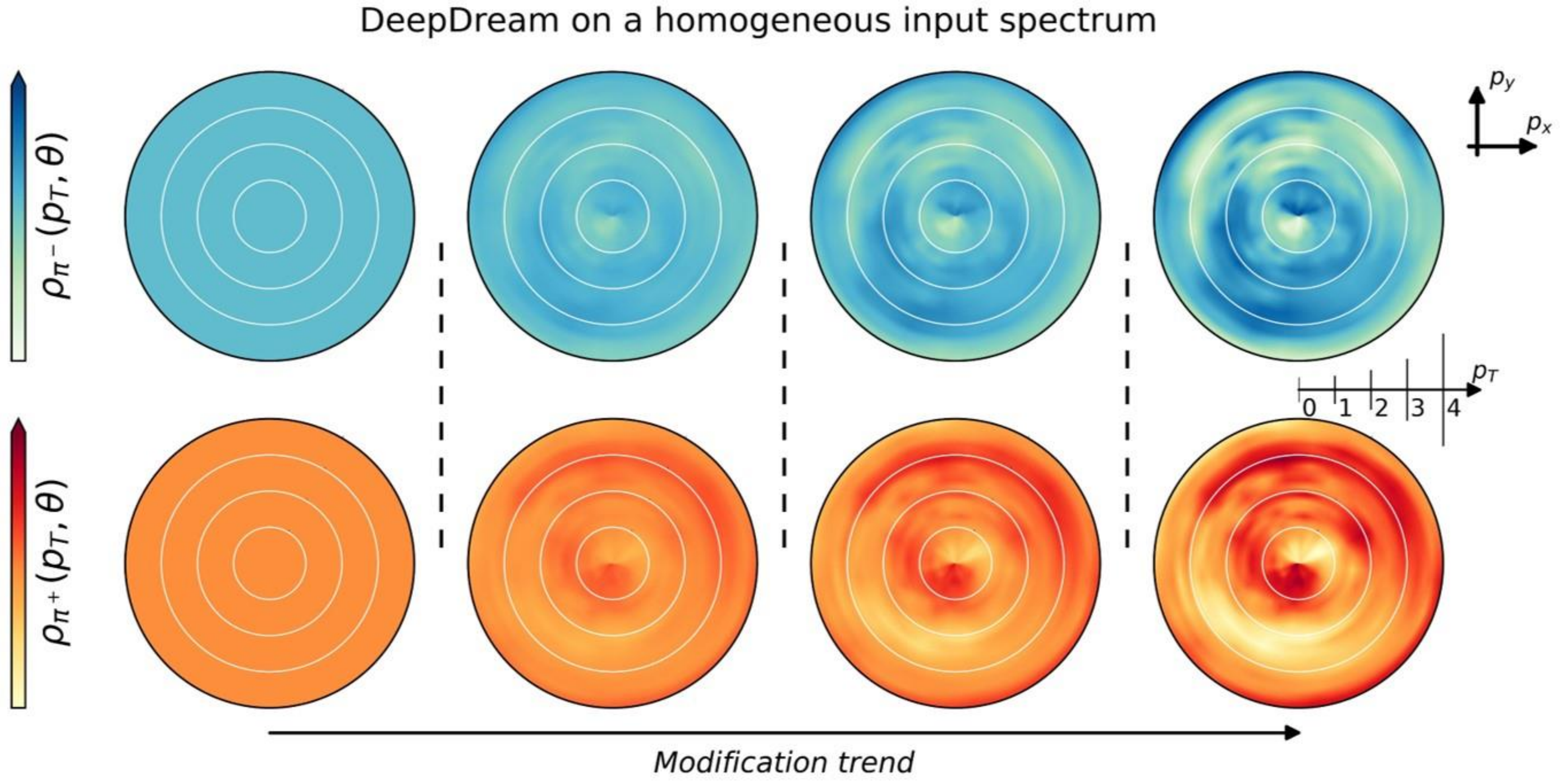
- $P_{1,RuRu} > P_{1,ZrZr}$, both $\sim 90\%$
- Problem: the lines are close
  - Possible solution: change the activation function to make the NN work more reliably at large P_1.



${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$
Simulated by AMPT, both @ 200GeV

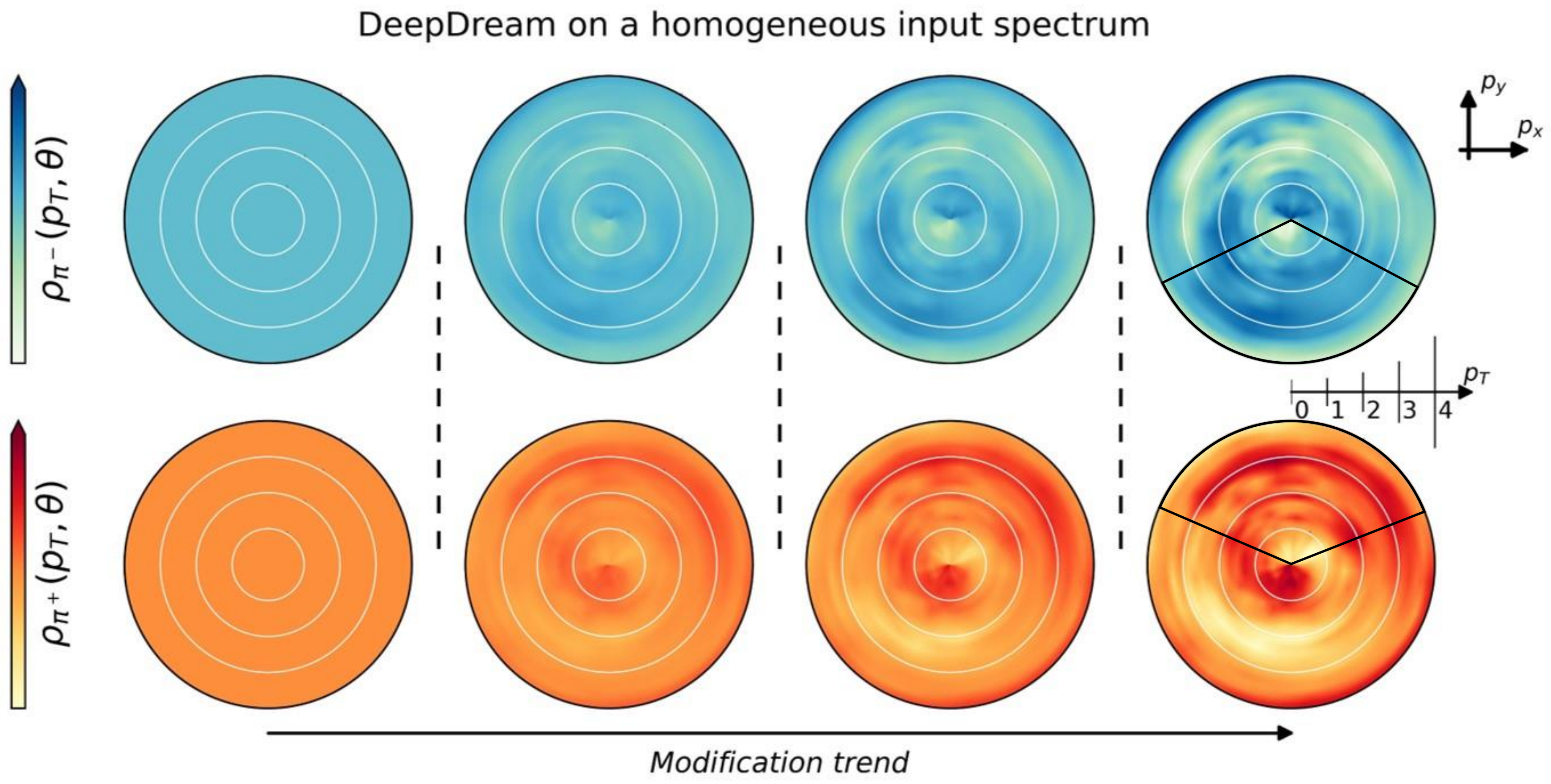
Visualization-DeepDream

- Fix NN, modify the input



Visualization-DeepDream

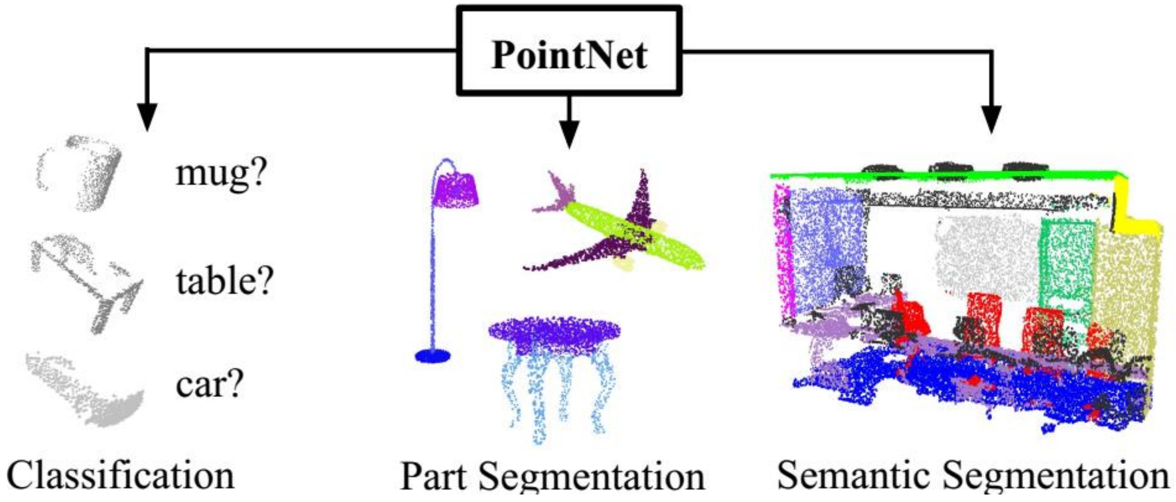
- Fix NN, modify the input



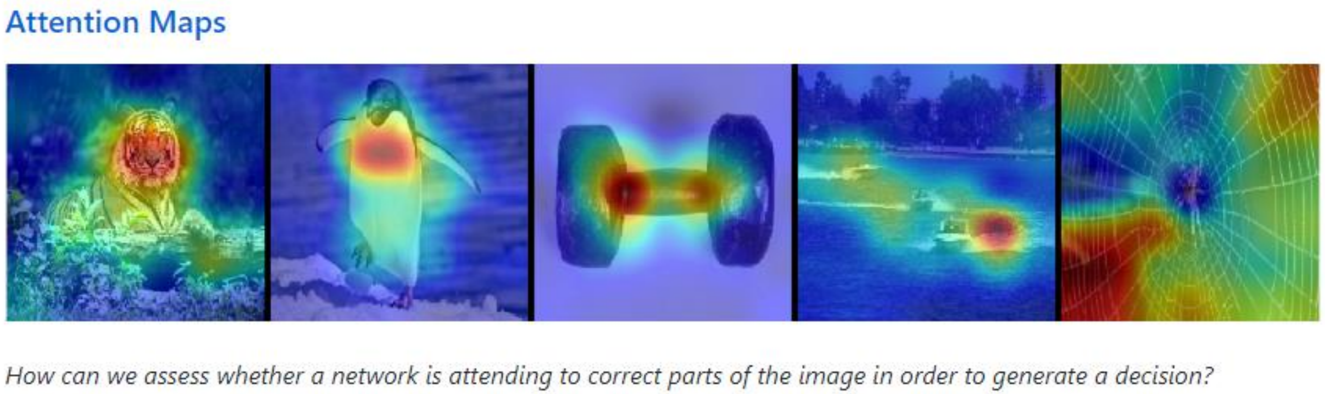
Summary

- DL is capable of distinguishing the pattern of charge separation
- And this pattern is robust against the background in the final states
- Transportable to a series of collision systems

Outlook



- Better y-axis mirror symmetry
  - Tuning initial weights by hand
- From NN to analytic observable
- Attention mechanism / importance mechanism
- PointNet and single event spectra



From github/keras-vis

Thank you