Machine (Deep) Learning & explore nuclear structure with heavy ion collisions

Long-Gang Pang (庞龙刚)

Central China Normal University

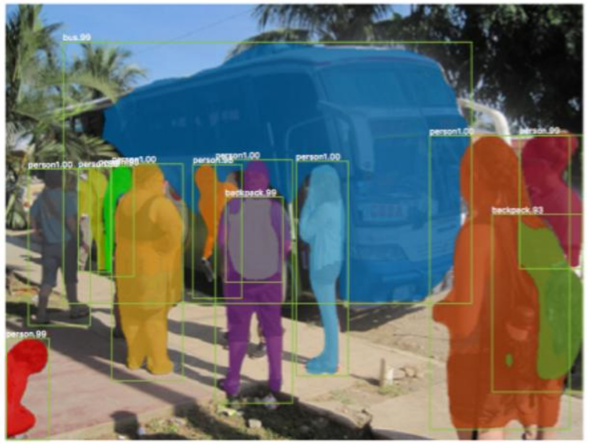
Exploring nuclear physics across energy scales

Peking University, 04/27/2024

Deep learning for AI

Face Identification

Segmentation what&where



NLP, NLG, Large Language Models

Chinese Poetry Generation with Planning based Neural Network

Zhe Wang†, Wei He‡, Hua Wu‡, Haiyang Wu‡, Wei Li‡, Haifeng Wang‡, Enhong Chen†
†University of Science and Technology of China, Hefei, China
‡Baidu Inc., Beijing, China

秋夕湖上 By a Lake at Autumn Sunset 一夜秋凉雨湿衣， A cold autumn rain wetted my clothes last night, 西窗独坐对夕晖。 And I sit alone by the window and enjoy the sunset. 湖波荡漾千山色， With mountain scenery mirrored on the rippling lake, 山鸟徘徊万籁微。 A silence prevails over all except the hovering birds.	秋夕湖上 By a Lake at Autumn Sunset 荻花风里桂花浮， The wind blows reeds with osmanthus flying, 恨竹生云翠欲流。 And the bamboos under clouds are so green as if to flow down. 谁拂半湖新镜面， The misty rain ripples the smooth surface of lake, 飞来烟雨暮天愁。 And I feel blue at sunset.

Table 6: A pair of poems selected from the blind test. The left one is a machine-generated poem, and the right one is written by Shaoti Ge, a poet lived in the Song Dynasty.

arXiv: 1610.09889v1

Machine Translation

Speech recognition

Natural Language Understanding, Processing and Generation

AlphaGo， Alpha Master， AlphaGo Zero， Alpha Zero …

Since Dec. 29, 2016, Master starts to beat top Go players secretly online. In one week, Master defeated all the top Go players from China, Korea and Japan.

60 wins, zero loss

AlphaGo@Twitter: “Now, I am the master”

Chat GPT， Clauder3， Kimi， 讯飞星火...

The common technique behind these AI applications: deep Learning!

DL: Neural Network with multi hidden layers

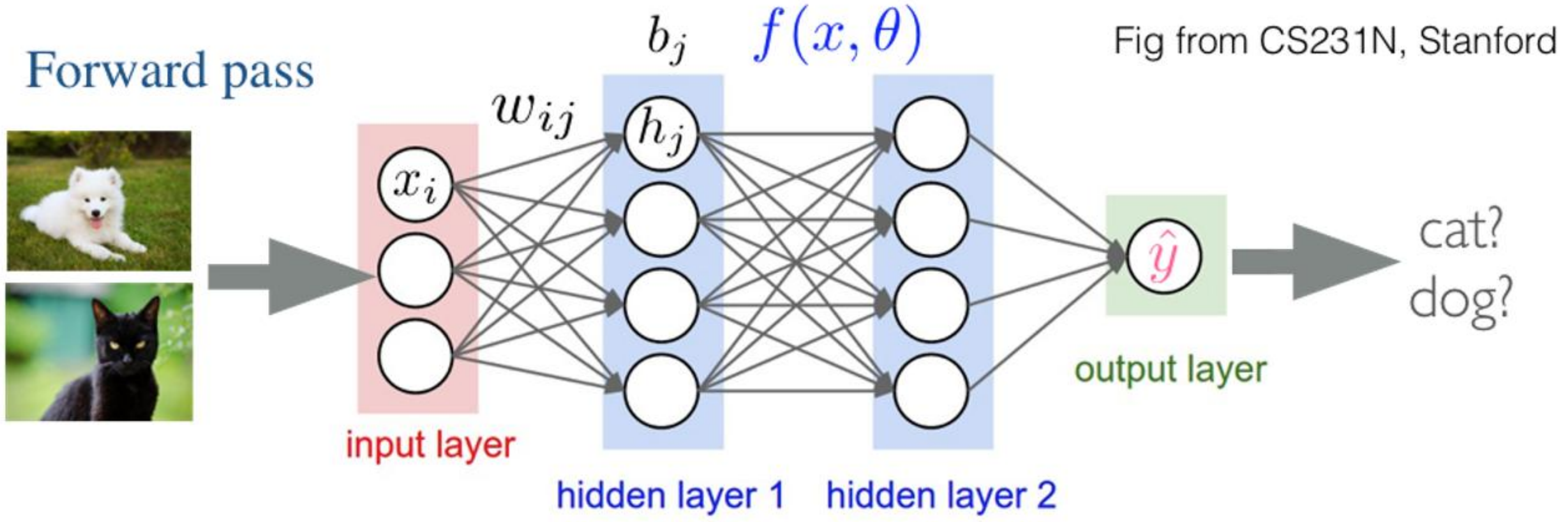


Fig from CS231N, Stanford

Linear operation

$$z_j = \sum_{i=1}^{N} x_i w_{ij} + b_j$$

scaling, rotating, boosting, changing dimensions

Non-linear activation function $h_j = \sigma(z_j)$

(a) Sigmoid

$$\sigma(z) = \frac{1}{1+\exp(-z)}$$

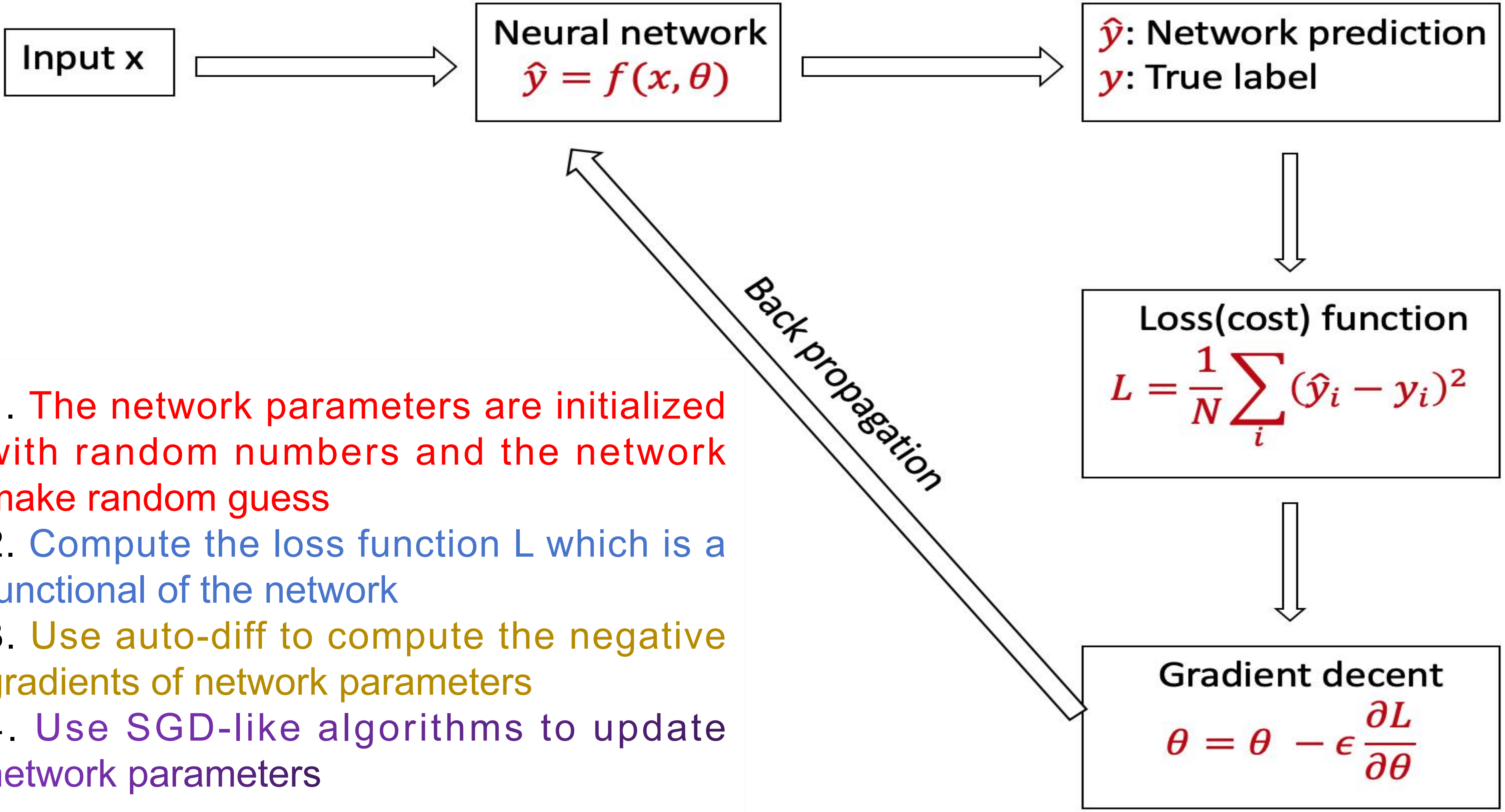
(b) ReLU

$$\sigma(z) = \begin{cases} z, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

(c) PReLU

$$\sigma(z) = \begin{cases} z, & z > 0 \\ az, & z \leq 0 \end{cases}$$

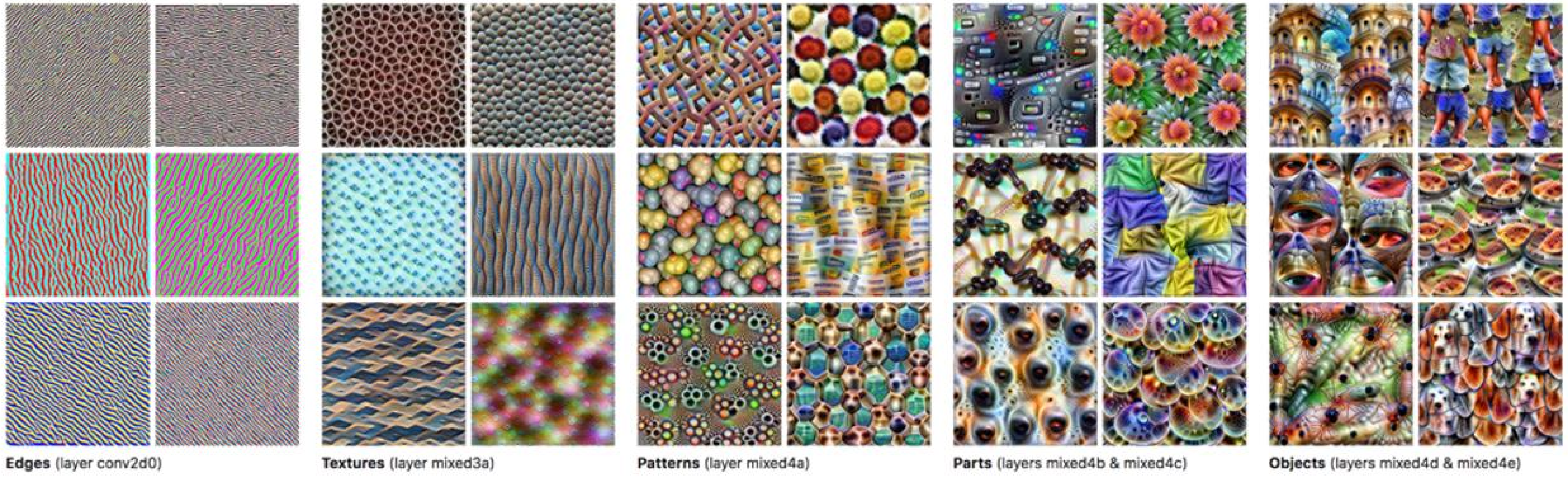
How does the network learn



1. The network parameters are initialized with random numbers and the network make random guess
2. Compute the loss function L which is a functional of the network
3. Use auto-diff to compute the negative gradients of network parameters
4. Use SGD-like algorithms to update network parameters

What has been learned by the deep neural network（Global interpretation）

Olah, et al., "Feature Visualization", Distill, 2017.



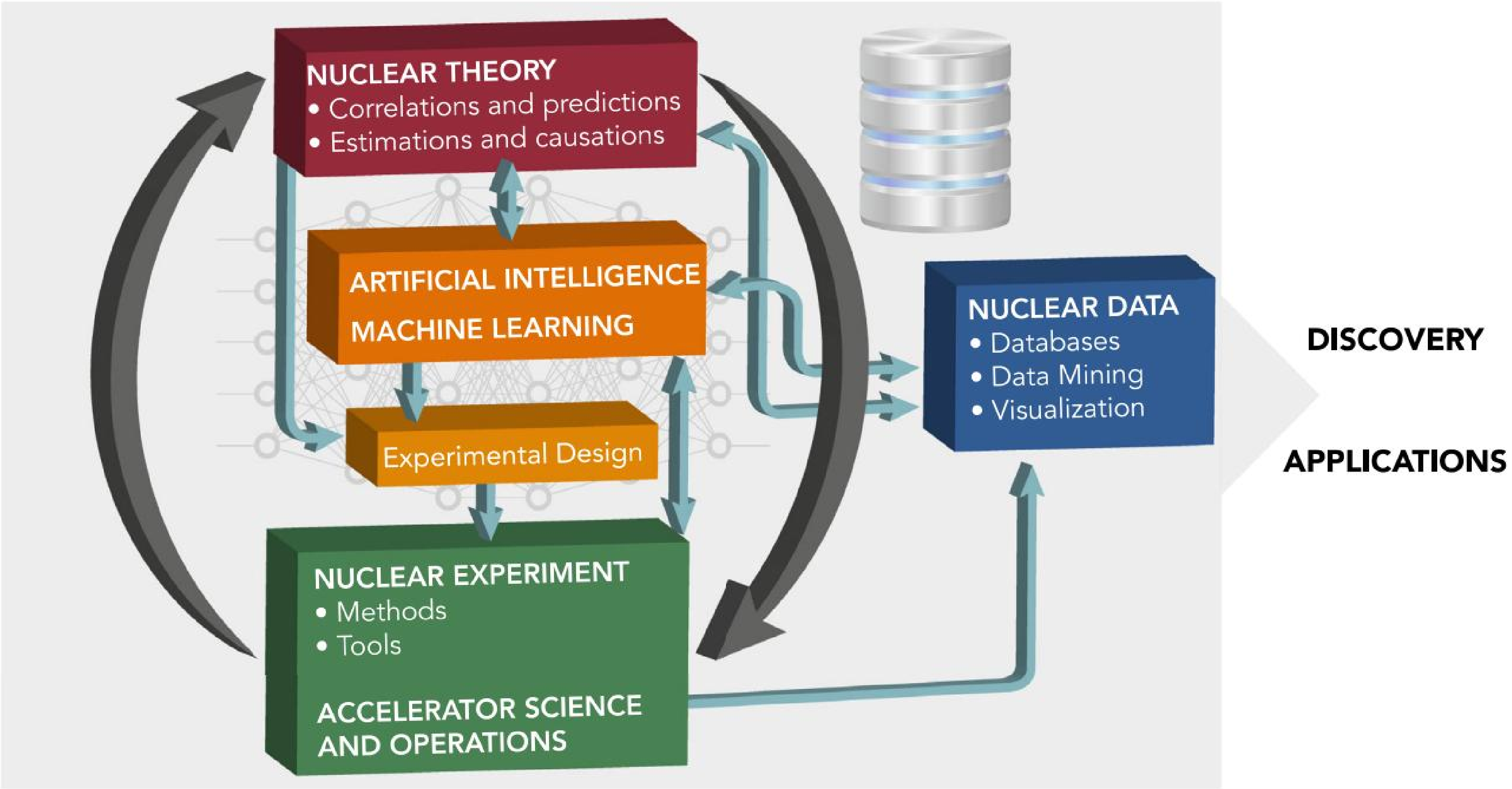
shallow layers

deep layers

What has been learned by the deep neural network (local interpretation)

- Ablation studies： LIME or Prediction Difference Analysis. M. Tulio Ribeiro, et. al. “Why should I trust you?”
- Class activation map： map the deep layers to the input image， look for the most important region for decision making. BoLei Zhou, et. al. “Learning Deep Features for discriminative localization”
- Layer-wise relevance propagation： set the relavance of the output layer to 1， propagate the relevance to the input data, to look for the most important region for decision making.

Machine Learning in Nuclear Physics



Machine Learning in Nuclear Physics，RMP 2022

Reviews

Colloquium: Machine learning in nuclear physics

Amber Boehnlein, Markus Diefenthaler, Nobuo Sato, Malachi Schram, Veronique Ziegler, Cristiano Fane
Morten Hjorth-Jensen, Tanja Horn, Michelle P. Kuchera, Dean Lee, Witold Nazarewicz, Peter Ostroumov
Orginos, Alan Poon, Xin-Nian Wang, Alexander Scheinker, Michael S. Smith, and Long-Gang Pang

Rev. Mod. Phys. **94**, 031003 – Published 8 September 2022

Article | References | No Citing Articles | PDF | HTML | Export Citation

ABSTRACT

Advances in machine learning methods provide tools that have broad applicability in scientific research. These techniques are being applied across the diversity of nuclear physics research topics, leading to advances that will facilitate scientific discoveries and societal applications. This Colloquium provides a snapshot of nuclear physics research, which has been transformed by machine learning techniques.

High energy nuclear physics meets Machine Learning

#1

Wan-Bing He (Fudan U., Shanghai and Fudan U.), Yu-Gang Ma (Fudan U., Shanghai and Fudan U.), Long-Gang Pang, Huichao Song (CCNU, Wuhan, Inst. Part. Phys. and Hua-Zhong Normal U., LQLP and Peking U.), Kai Zhou (Frankfurt U., FIAS) (Mar 12, 2023)

e-Print: 2303.06752 [hep-ph]

pdf

Exploring QCD matter in extreme conditions with Machine Learning

Kai Zhou (Frankfurt U., FIAS), Lingxiao Wang (Frankfurt U., FIAS), Long-Gang Pang (CCNU, Wuhan, Inst. Part. Phys.), Shuzhe Shi (Stony Brook U.)

Mar 27, 2023

146 pages

e-Print: 2303.15136 [hep-ph]

HEPML-LivingReview

A Living Review of Machine Learning for Particle Physics

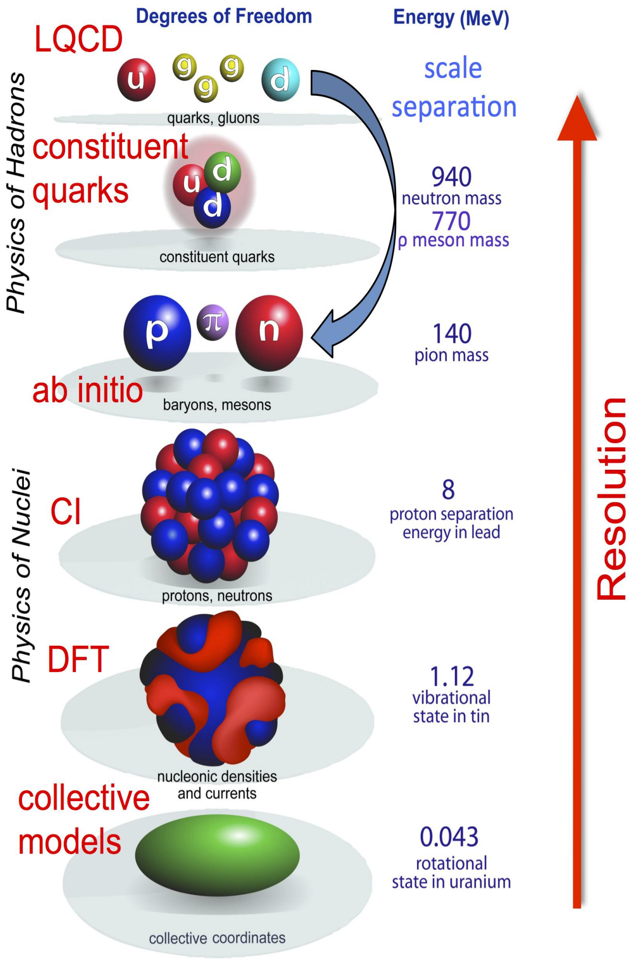
Modern machine learning techniques, including deep learning, is rapidly being applied, adapted, and developed for high energy physics. The goal of this document is to provide a nearly comprehensive list of citations for those developing and applying these approaches to experimental, phenomenological, or theoretical analyses. As a living document, it will be updated as often as possible to incorporate the latest developments. A list of proper (unchanging) reviews can be found within. Papers are grouped into a small set of topics to be as useful as possible. Suggestions are most welcome.

download review

The purpose of this note is to collect references for modern machine learning as applied to particle physics. A minimal number of categories is chosen in order to be as useful as possible. Note that papers may be referenced in more than one category. The fact that a paper is listed in this document does not endorse or validate its content - that is for the community (and for peer-review) to decide. Furthermore, the classification here is a best attempt and may have flaws - please let us know if (a) we have missed a paper you think should be included, (b) a paper has been misclassified, or (c) a citation for a paper is not correct or if the journal information is now available. In order to be as useful as possible, this document will continue to evolve so please check back before you write your next paper. If you find this review helpful, please consider citing it using \cite{hepmllivingreview} in HEPML.bib.

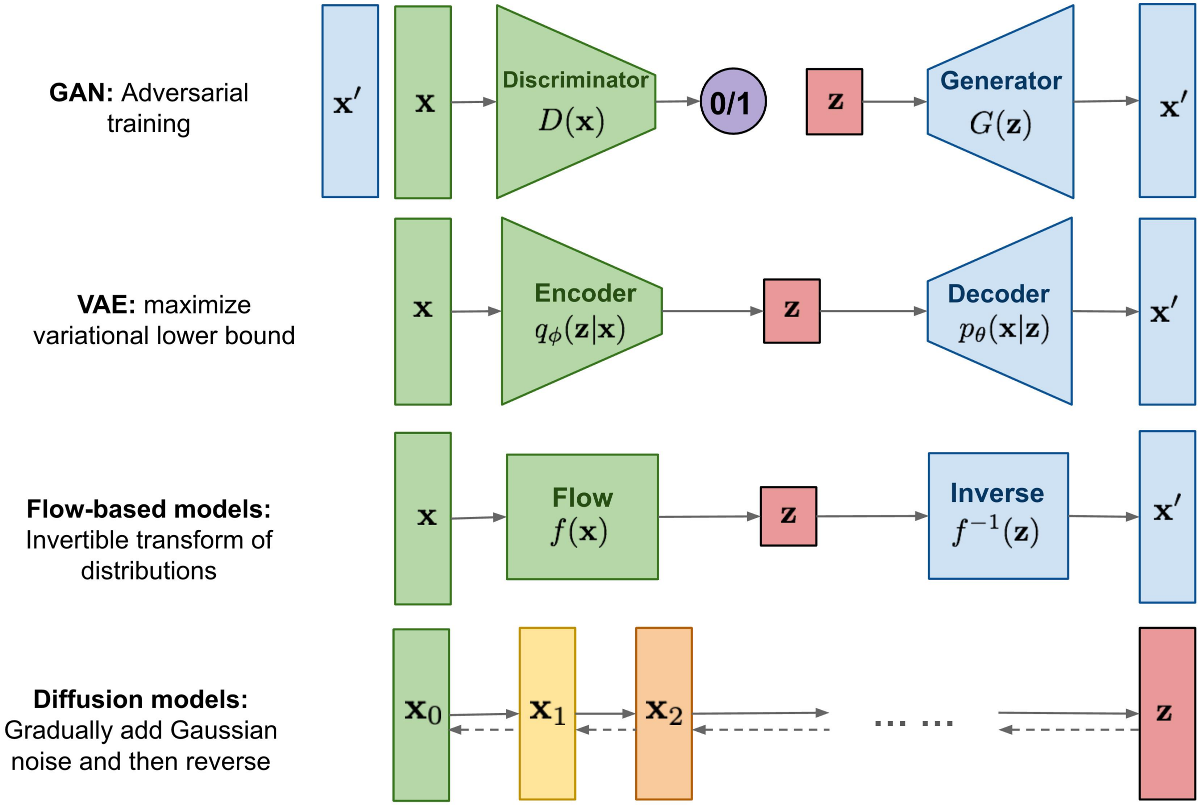
- Reviews
  - Modern reviews
    - Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning [DOI]
    - Deep Learning and its Application to LHC Physics [DOI]
    - Machine Learning in High Energy Physics Community White Paper [DOI]
    - Machine learning at the energy and intensity frontiers of particle physics
    - Machine learning and the physical sciences [DOI]
    - Machine and Deep Learning Applications in Particle Physics [DOI]
    - Modern Machine Learning and Particle Physics
    - Machine Learning in the Search for New Fundamental Physics
    - Artificial Intelligence and Machine Learning in Nuclear Physics

ML nuclear physics across energy scales

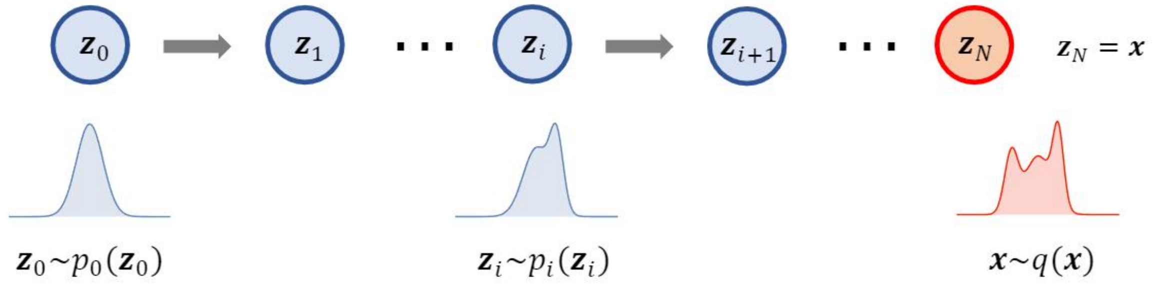


- Deep generative models (such as normalizing flow and the diffusion model) have been used to sample Field Configureations in Lattice QCD
- Deep learning is widely used to solve inverse problems of HIC to study the EoS of hot QCD matter, the phase transition, the transport coefficients eta/s, …
- Deep neural network is used to represent the many-body wave function of nucleus, to solve variational problems in ab initio calculations
- Deep learning is used to solve inverse problems of HIC to study the nuclear structure, for instance, the nuclear deformation, neutron skin, alpha cluster and short range correlation
- …

Generativive models: MC sampling

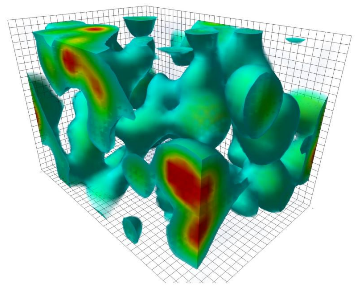


Similar to Box Muller algorithm



Samples Drawn from N-Dim Normal Distribution

Flow model or Diffusion models



Flow-based generative models for Markov chain Monte Carlo in lattice field theory
Albergo, Kanwar, Shanahan 1904.1207

Stacked U-net for relativistic fluid generation

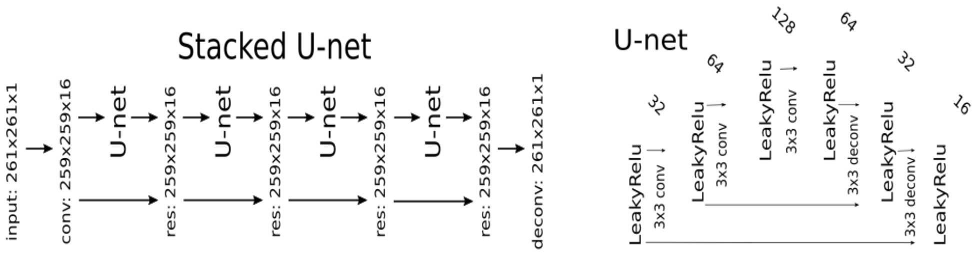
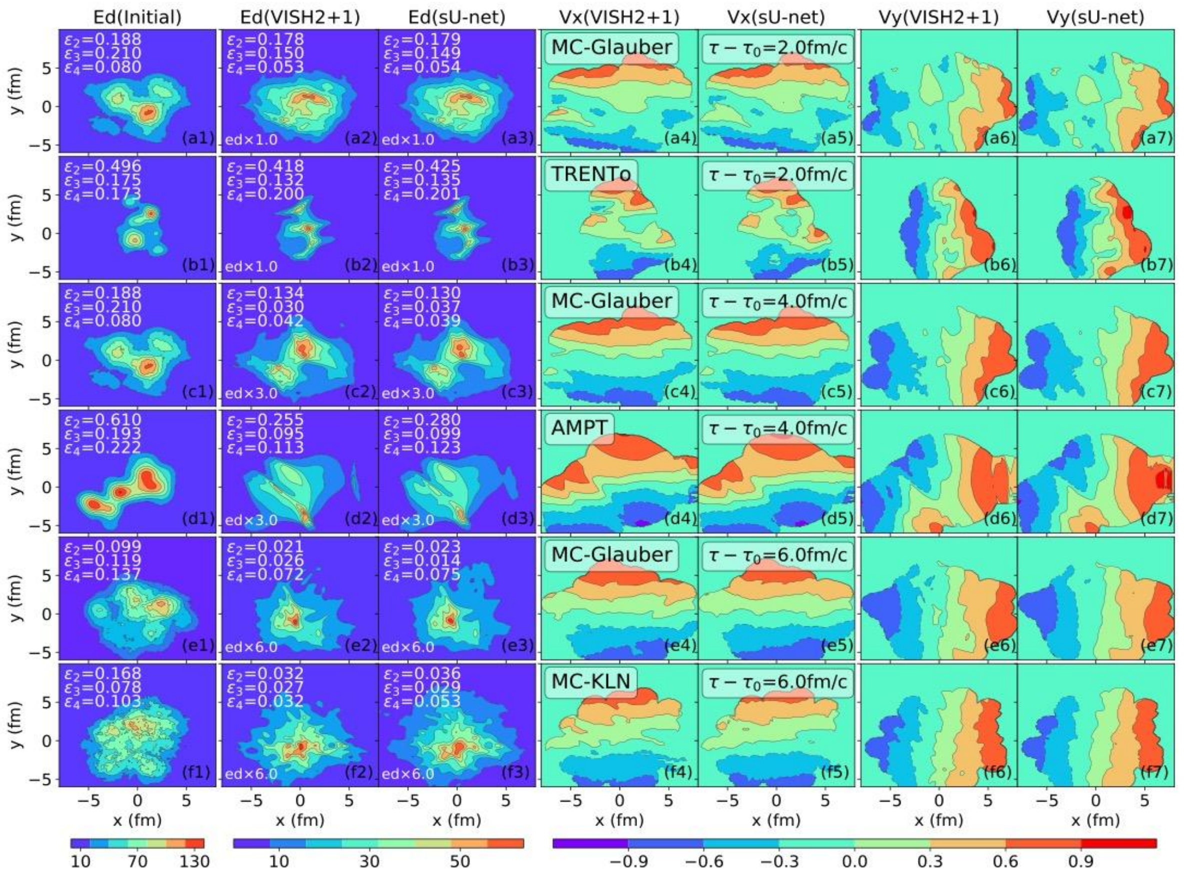


FIG. 1: An illustration of the encode-decode network, **stacked U-net**, which consists of the input and out layers and four residual U-net blocks. The right figure shows the U-net structure, and the depth of the hidden layer is written on the top of them.

The expansion of quark gluon plasma is learned in the image translation task using stacked UNET.

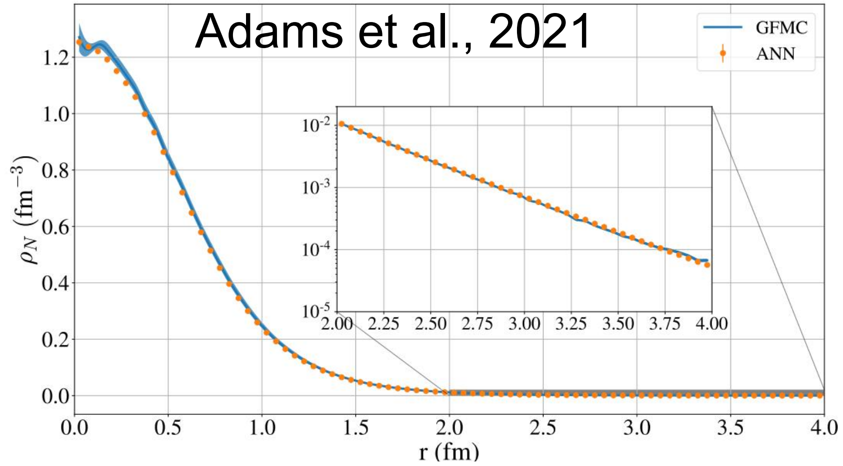
$$\nabla_\mu T^{\mu\nu} = 0$$



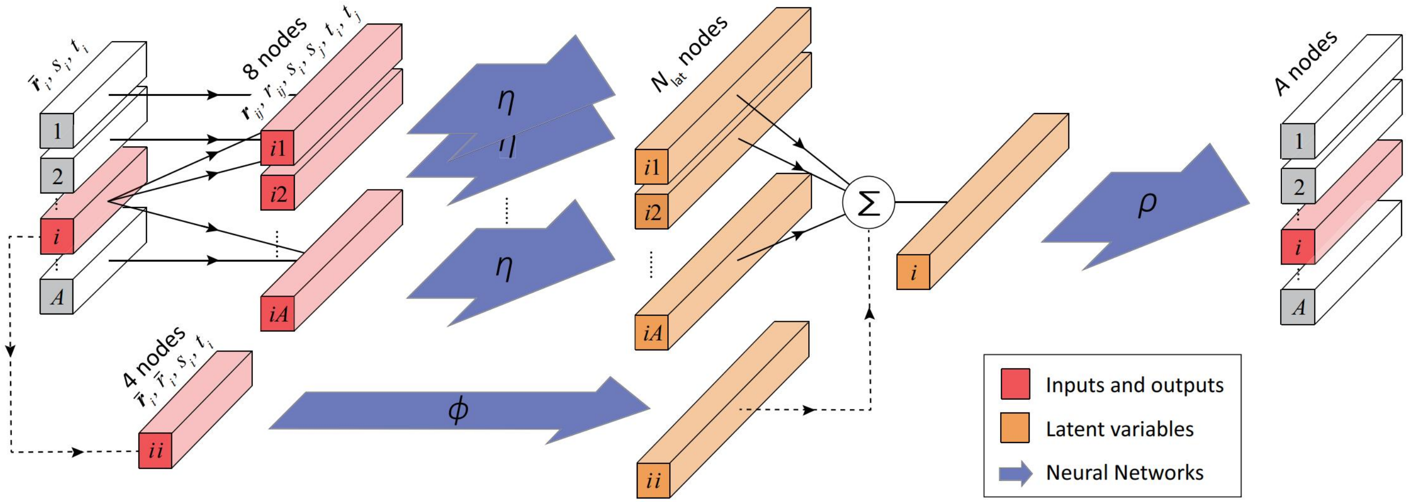
PRR3, 023256, H.Huang, B.Xiao, H.Xiong, Z.Liu, Z.Wu, Y. Mu and H.Song

Represent the many-body wave function

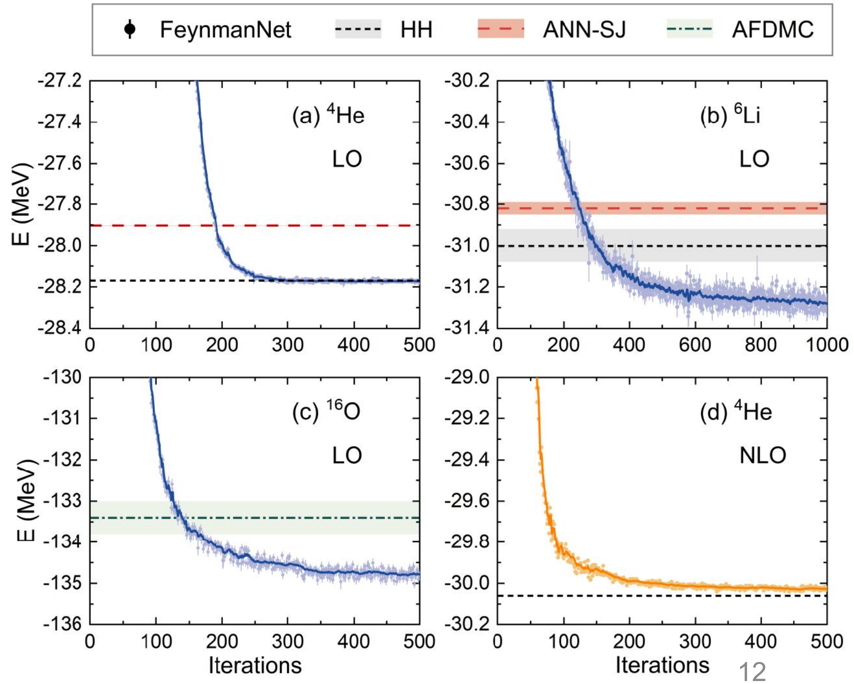
$$\hat{H} = \sum_{i=1}^{A} \frac{-\nabla_i^2}{2m_N} + \sum_{i<j} v_{ij} + \sum_{i<j<k} V_{ijk},$$



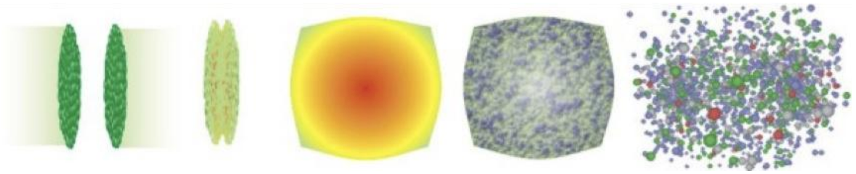
Adams et al., 2021

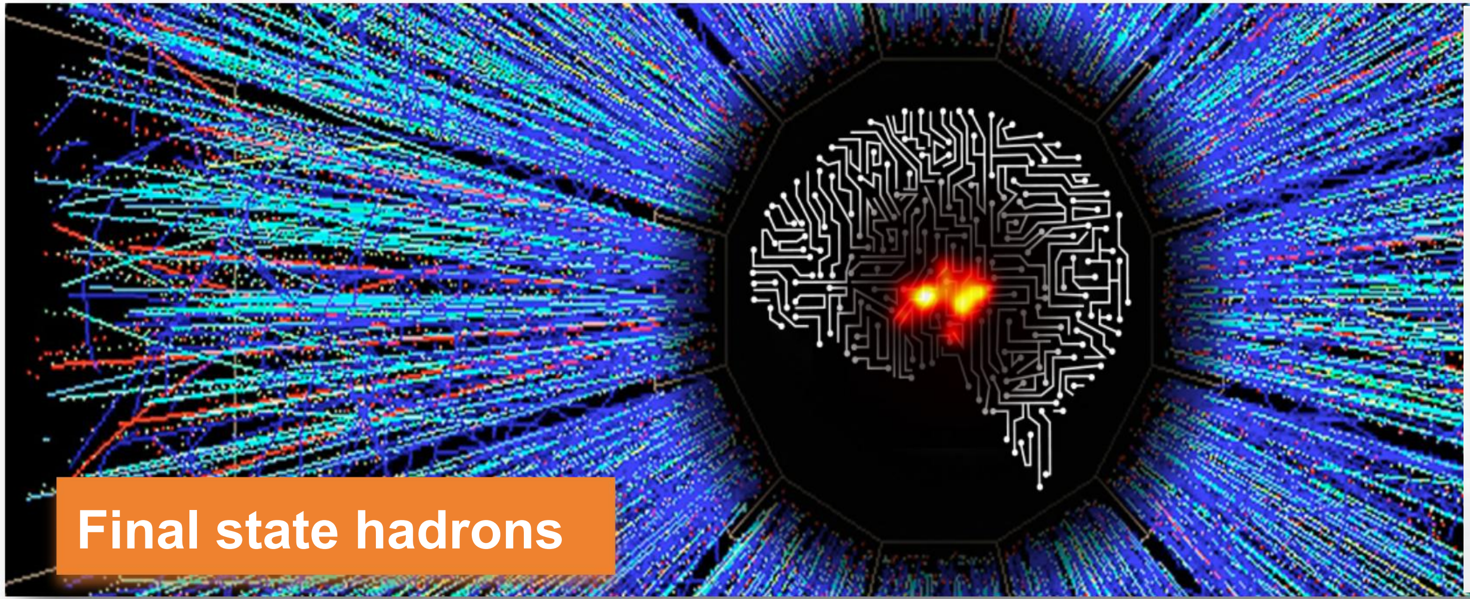


Y.L. Yang, P.W. Zhao, PRC 2023

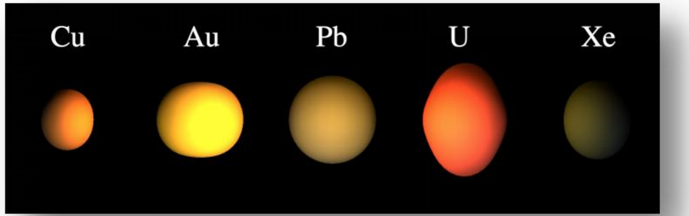


Inverse problems in HIC

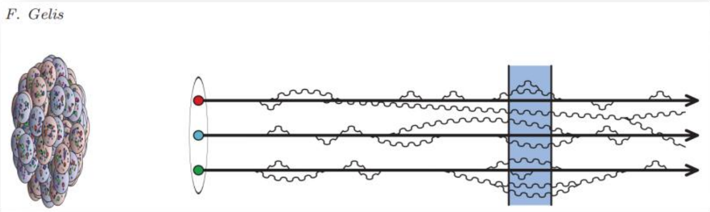




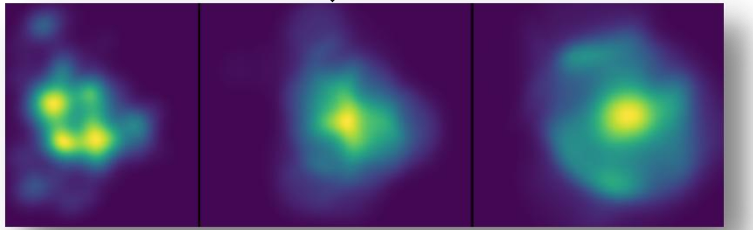
Non-linear mapping



(1) Nuclear Structure

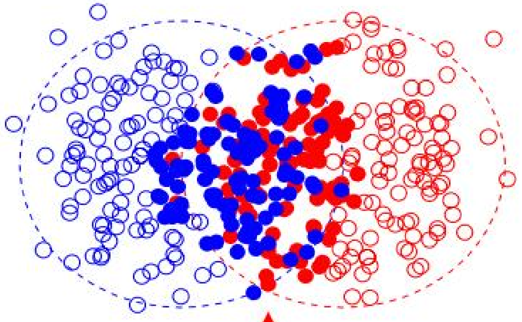


(2) Initial Parton Distribution



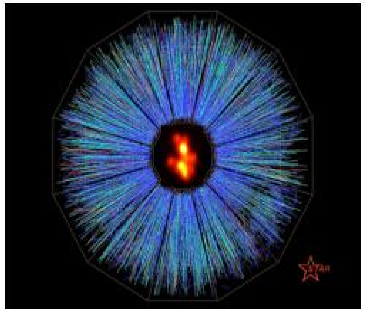
(3) QGP properties and EoS

Theoretical model: CLVisc



$$\nabla_\mu T^{\mu\nu} = 0$$

$$T^{\mu\nu} = (\varepsilon + P)u^\mu u^\nu - P g^{\mu\nu} + \pi^{\mu\nu}$$



Initial condition EoS Viscosity

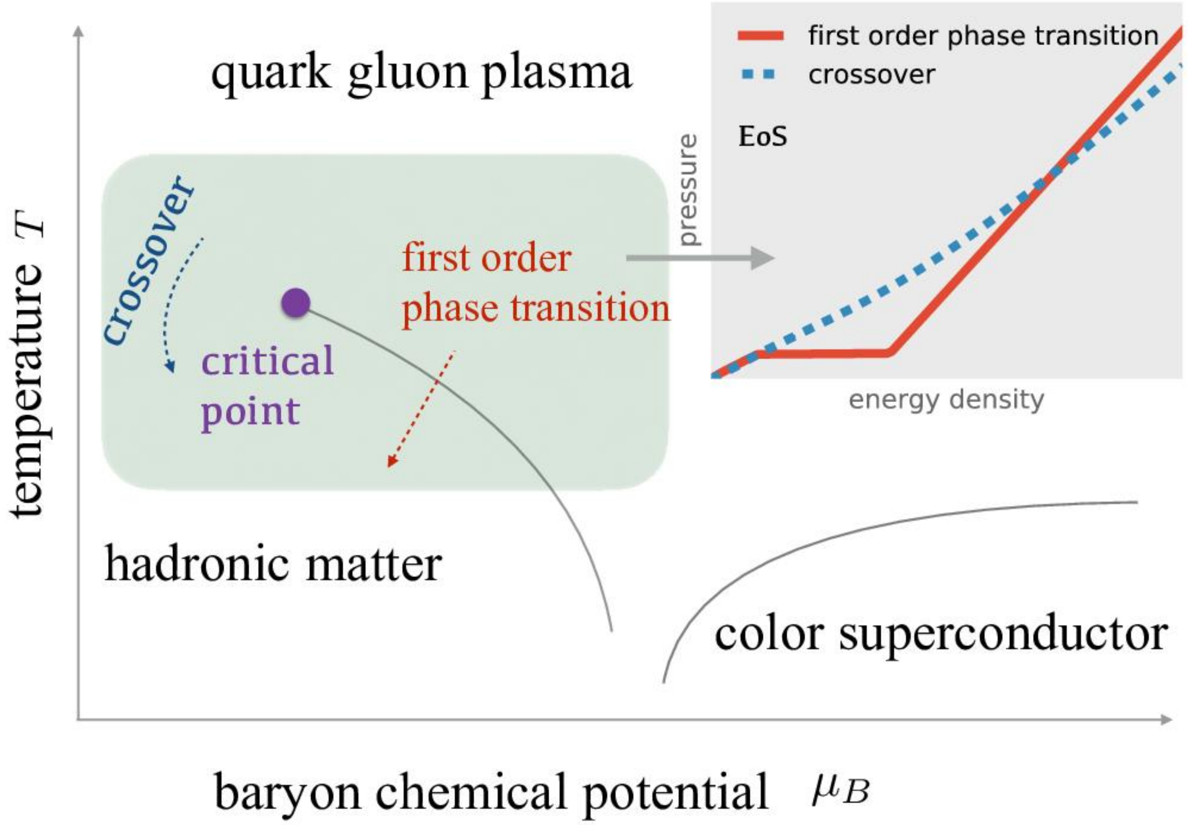
CLVisc： A 3+1D viscous hydro parallized on GPU using OpenCL

Purpose: Describe the non-equilibrium space-time evolution of hot QCD matter

Feature: 100 times faster than using a single core CPU.

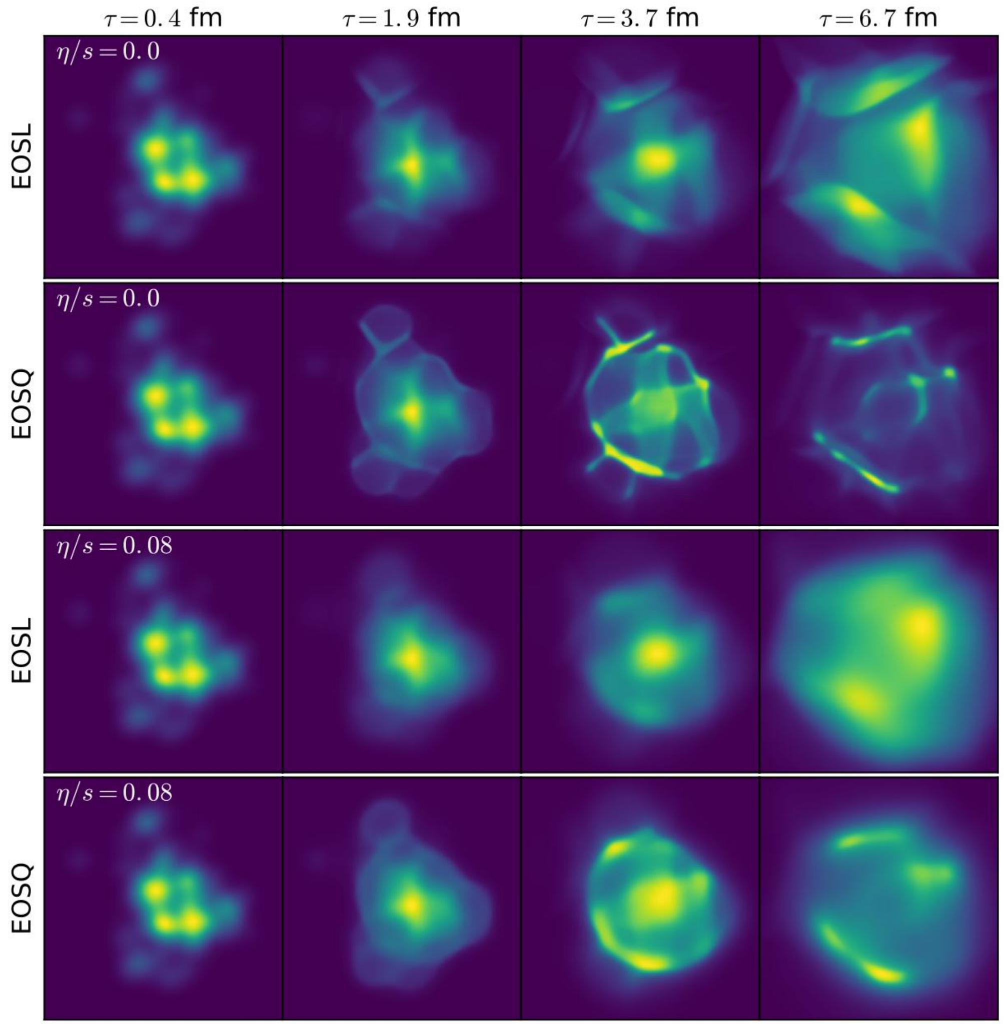
L.G. Pang, Q. Wang and X. N. Wang, PRC 86 (2012) 024911
L.G. Pang, B.W. Xiao, Y. Hatta, X.N.Wang, PRD 2015
L.G. Pang， H.Petersen, XN Wang, PRC97(2018)no.6,064918
XY Wu, GY Qin, LG Pang, XN Wang,PRC 105 (2022) 3, 034909

QCD phase diagram and nuclear EoS



- **Lattice QCD** predicts a **smooth cross over at 0 muB**
- **Taylor expansion** of Lattice QCD **gets EoS at small muB**
- **Sign problem** at large muB prevents the first principle calculation
- It is conjuctured there is a **1st order phase transition at large muB**
- Different phase transition types correspond to different EoS

CLVisc for different EoS



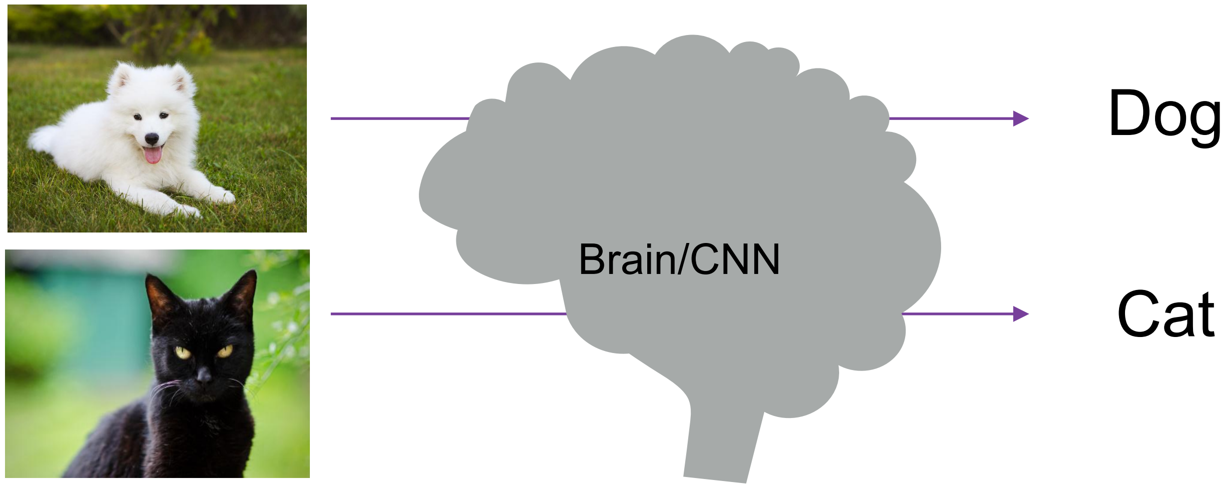
eta/s = 0
Lattice QCD EoS
(smooth cross over)

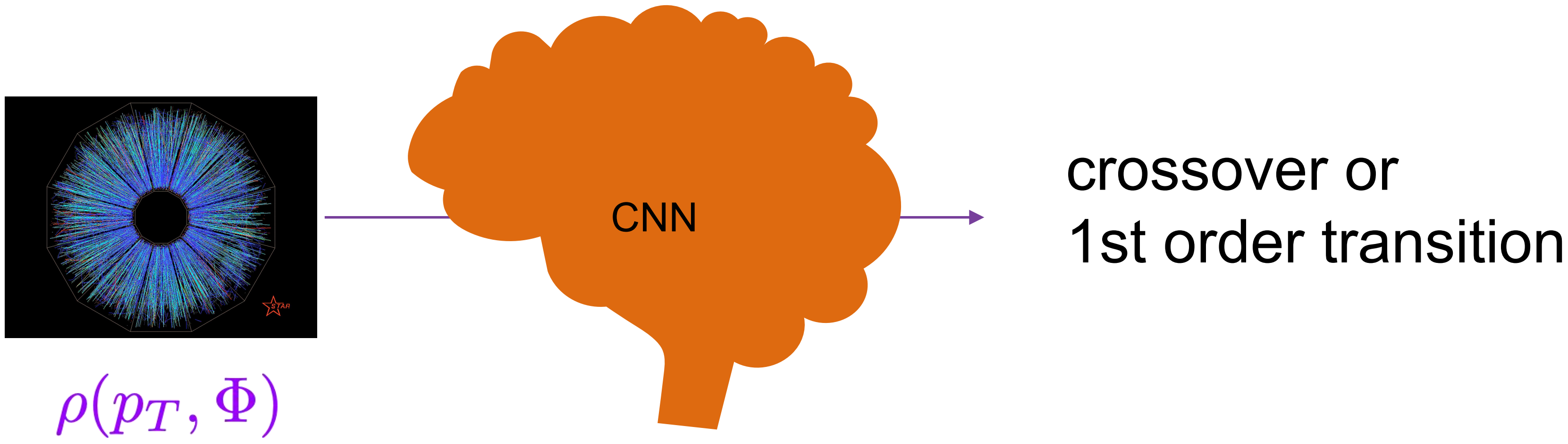
eta/s = 0
First order phase transition

eta/s = 0.08
Lattice QCD EoS

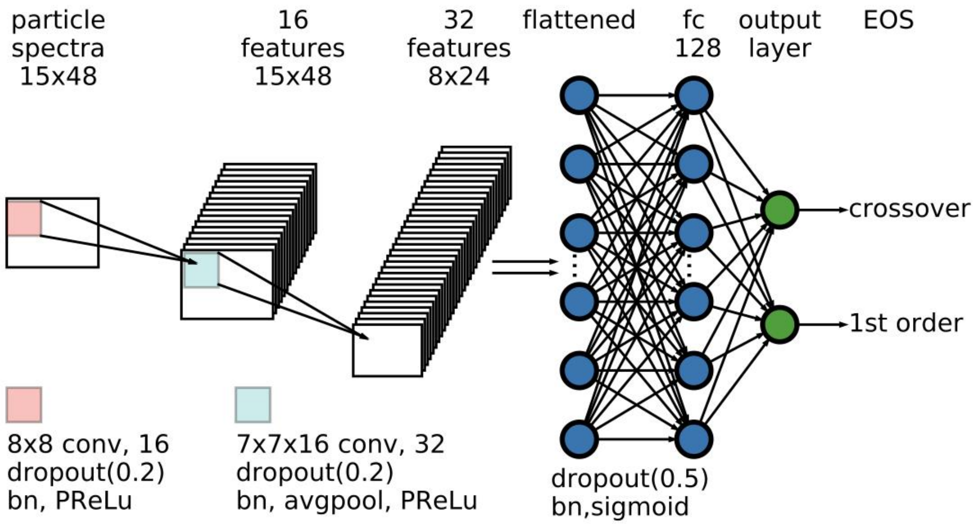
eta/s = 0.08
First order phase transition
eta/s: shear viscosity / entropy density

Powerful pattern recognition

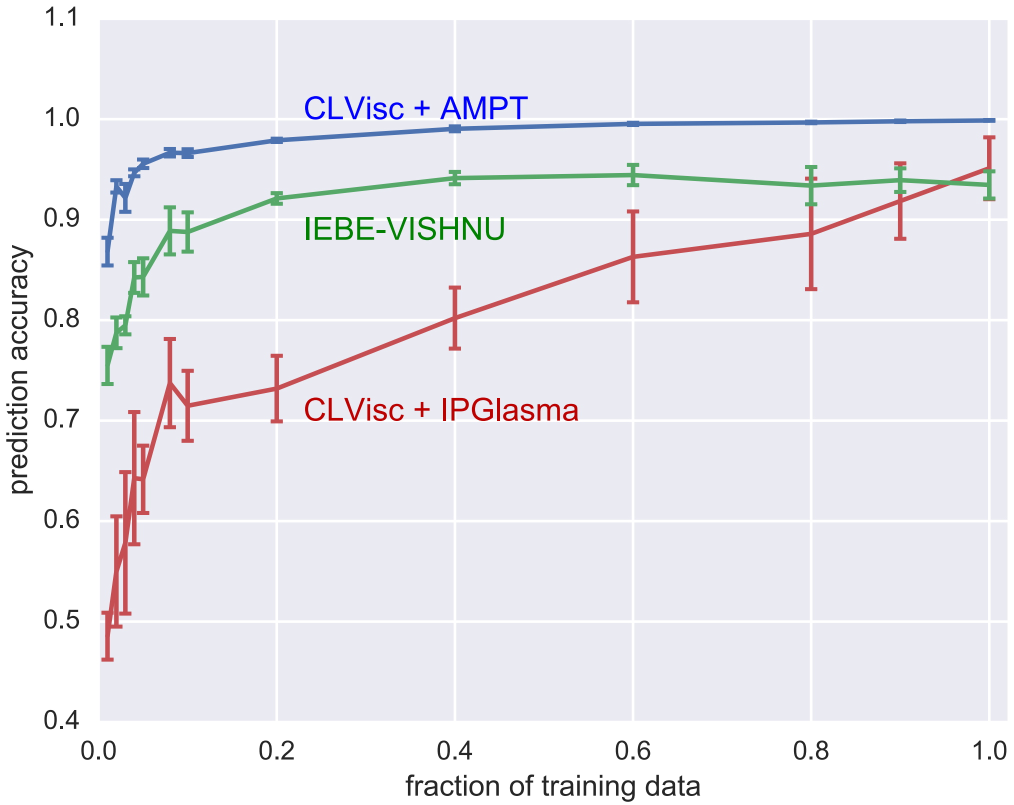




DL with CNN for EoS classification

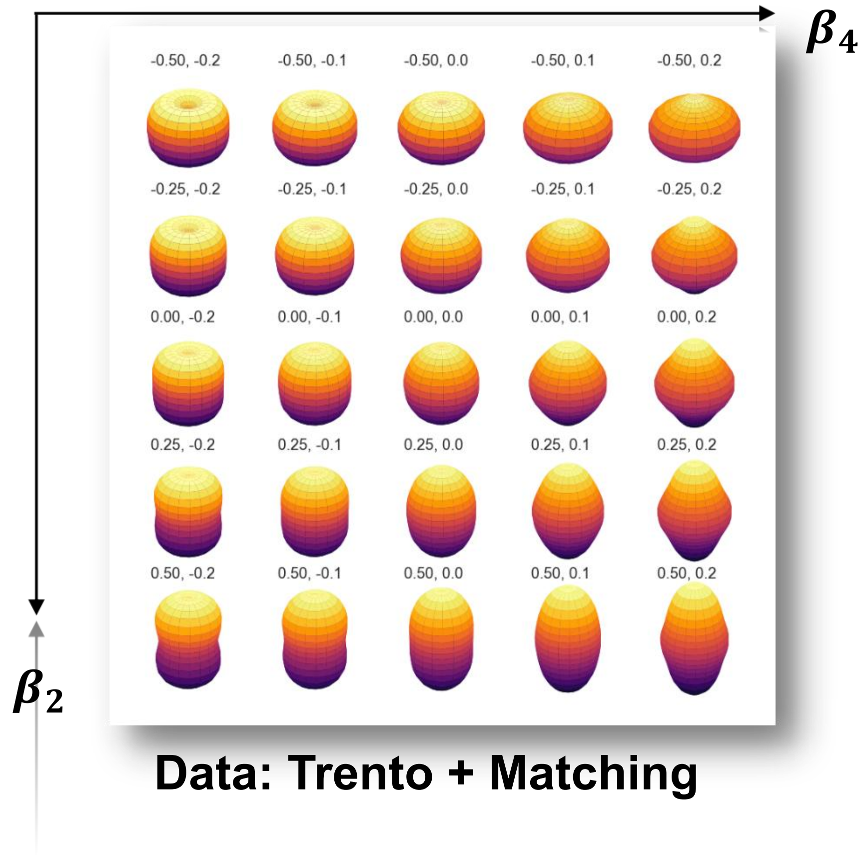


$$l(\theta) = \underbrace{-\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right]}_{\text{cross entropy loss}} + \underbrace{\lambda||\theta||_2^2}_{\text{L2 regularization}}$$

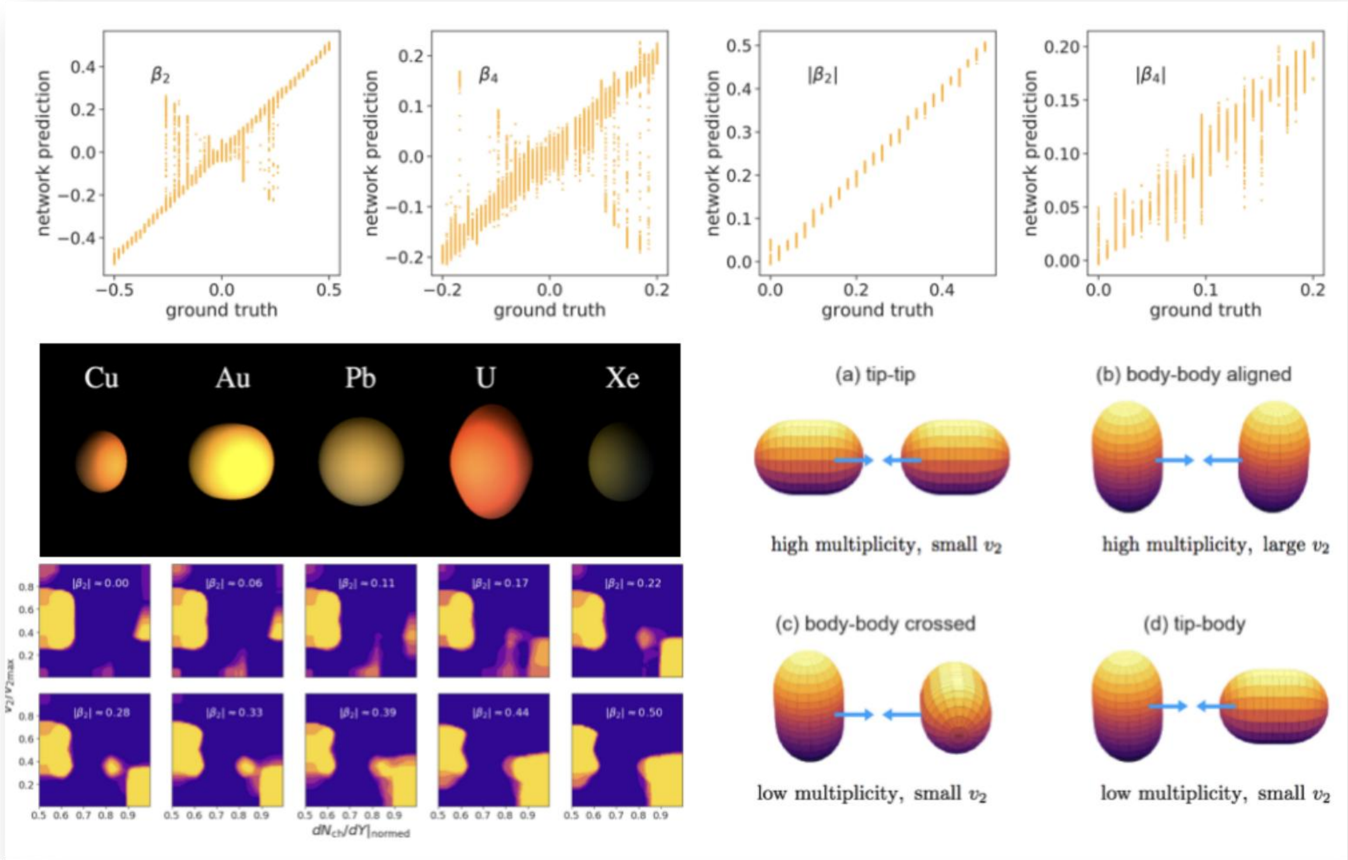


Nature Communications 2018, LG. Pang, K.Zhou, N.Su, H.Petersen, H. Stoecker, XN. Wang.

Determining nuclear deformation



Data: Trento + Matching



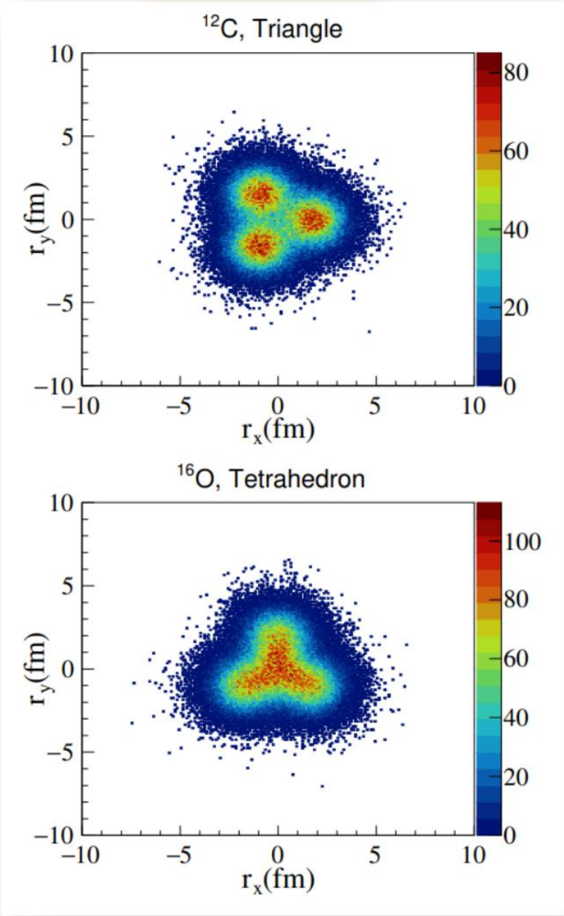
L.-G. Pang, K. Zhou and X.-N. Wang, arXiv:1906.06429

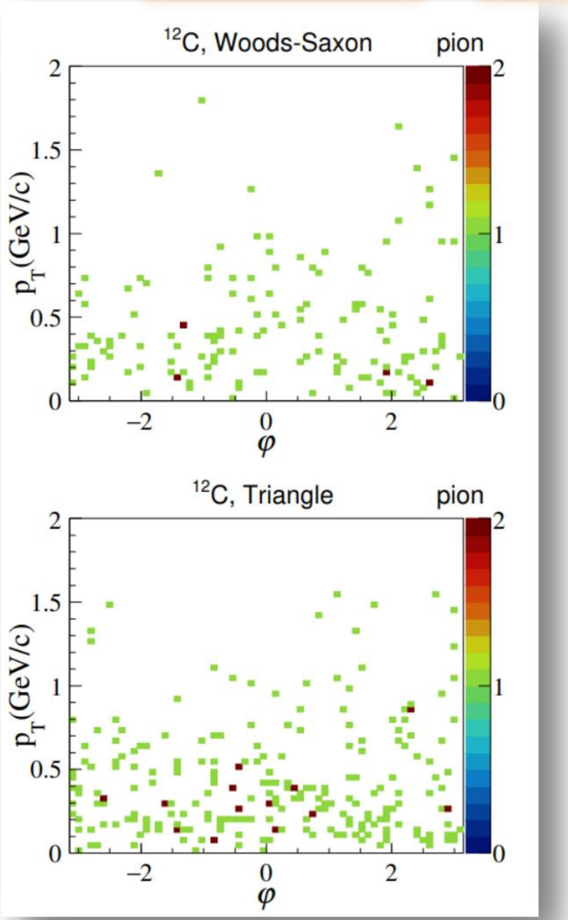
Identifying the α-clustering structure

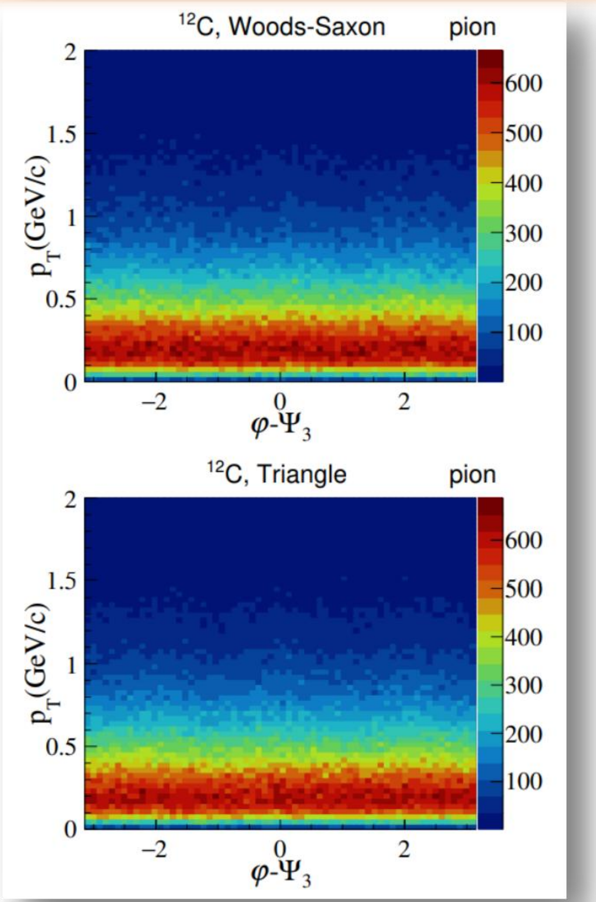
α clusters

Fail in EbE

Succeed with 4000-events average



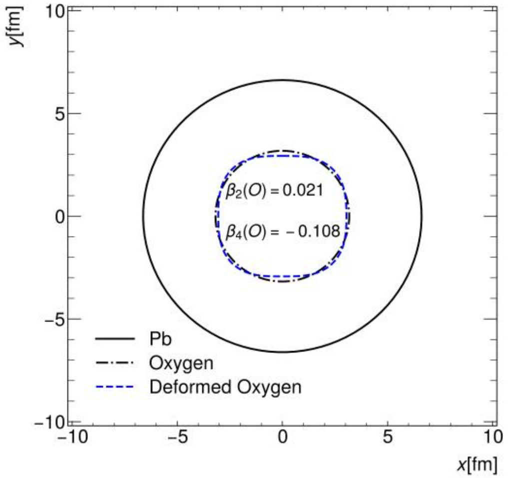




JJ He, WB He, YG Ma, S Zhang, PRC 104, 044902 (2021)

Alpha clusters in O+O collisions using CLVisc

WS, Deformed WS



4-α tetrahadron

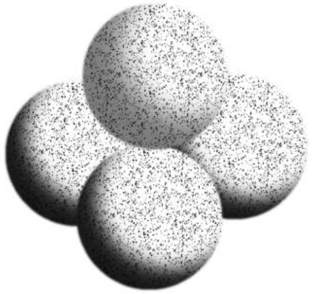
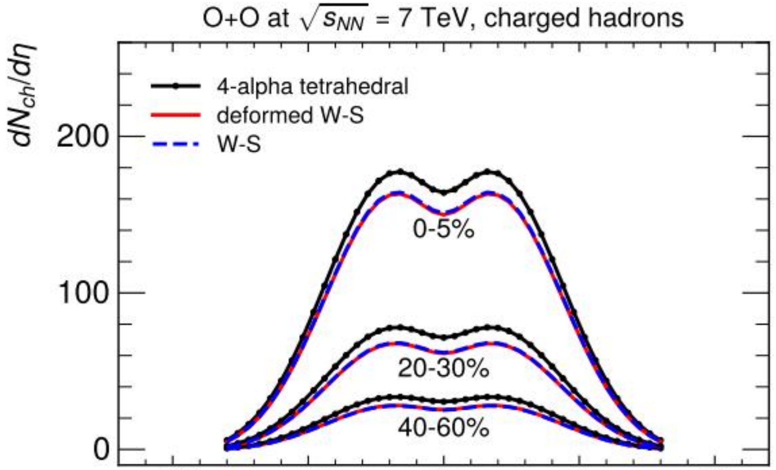
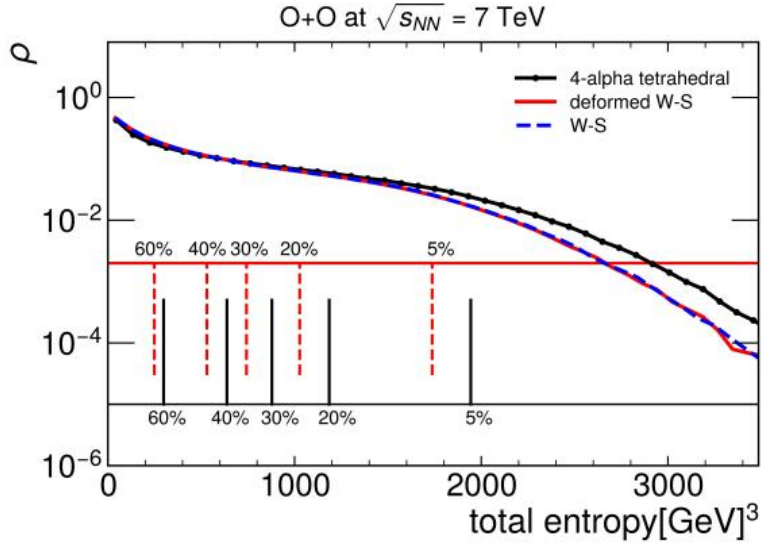


Table I. The ratio of charged multiplicity at middle pseudo-rapidity between different centralities.

Cent1/Cent2	WS	Deformed WS	Four-α
0-5%/20-30%	2.44	2.44	2.29
0-5%/40-60%	5.99	5.91	5.38

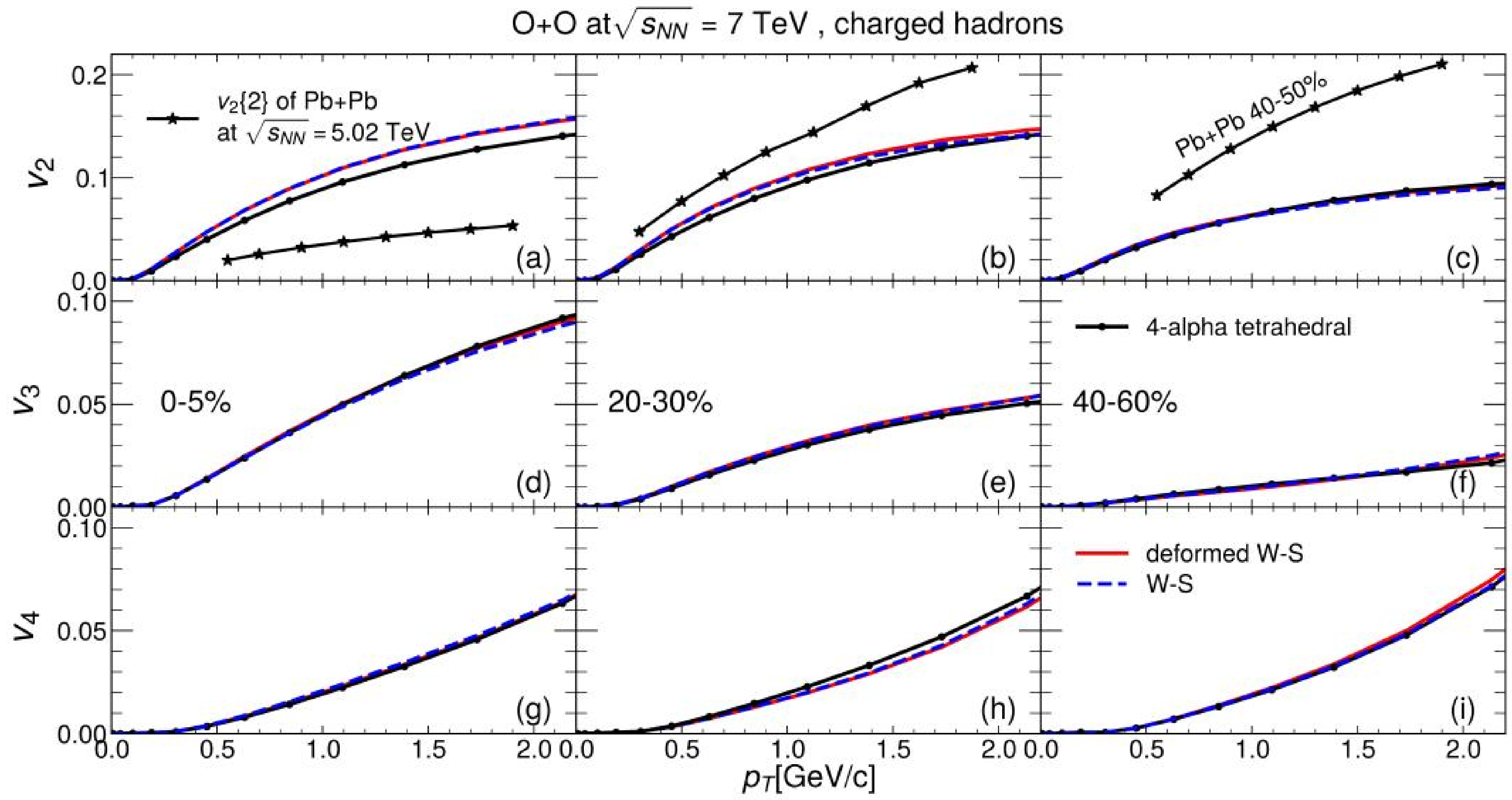
- In CLVisc simulations of 7 TeV O+O collisions, the centrality denpendence of charged multiplicity is quantitatively different for 4-α structure.



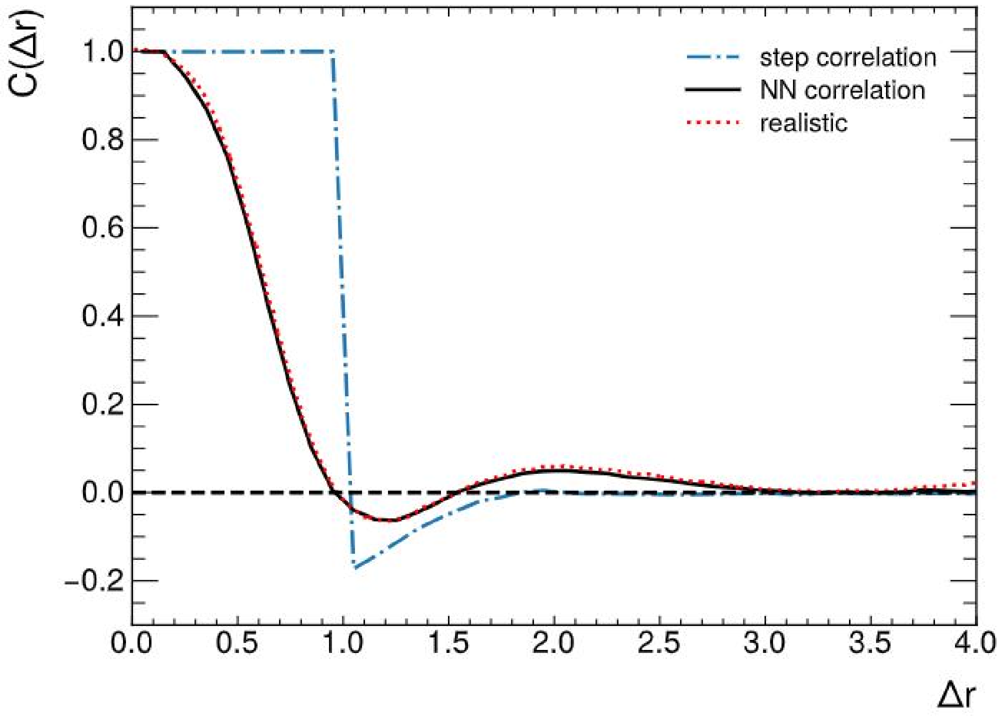


C.Ding, LG. Pang, S. Zhang and YG Ma, CPC 47 024105

The flow differences with alpha cluster in O16



Nucleon-Nuclon correlations



$$\rho(r) = \frac{\rho_0}{\exp(\frac{r-r_0}{d}) + 1},$$

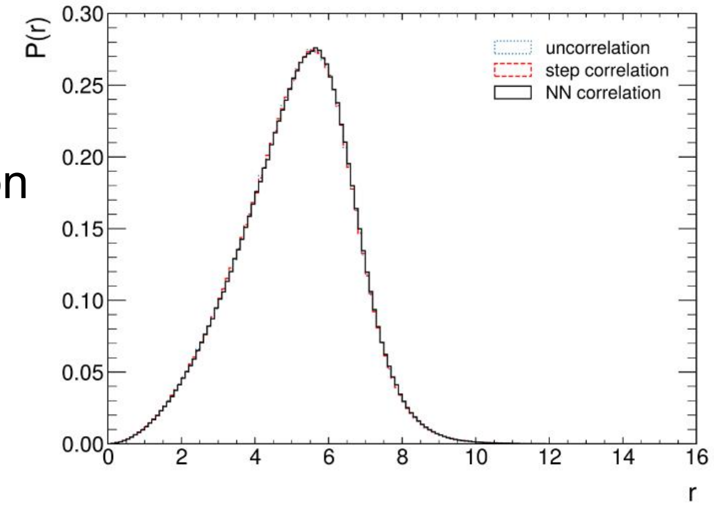
$$C(\Delta r) = 1 - \rho_c(\Delta r)/\rho_u(\Delta r).$$

Sample nucleons:
not only the single nucleon distribution,
but also the two-nucleon relative-distance distribution

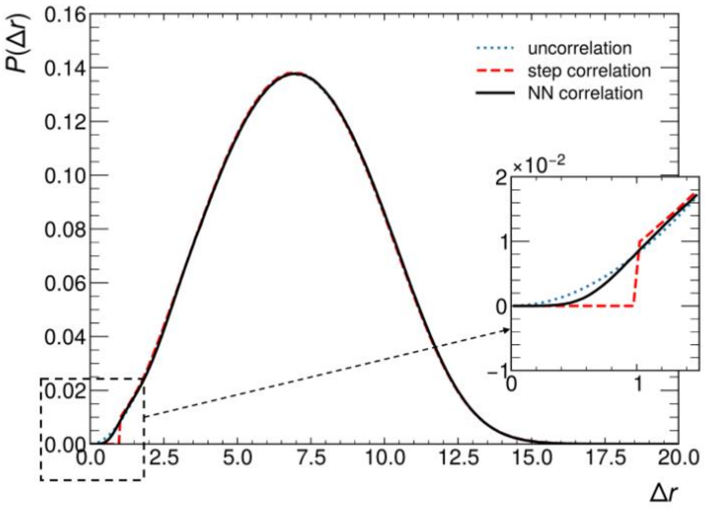
With YuJing Huang and Xin-Nian Wang, in preparation

The sampled results

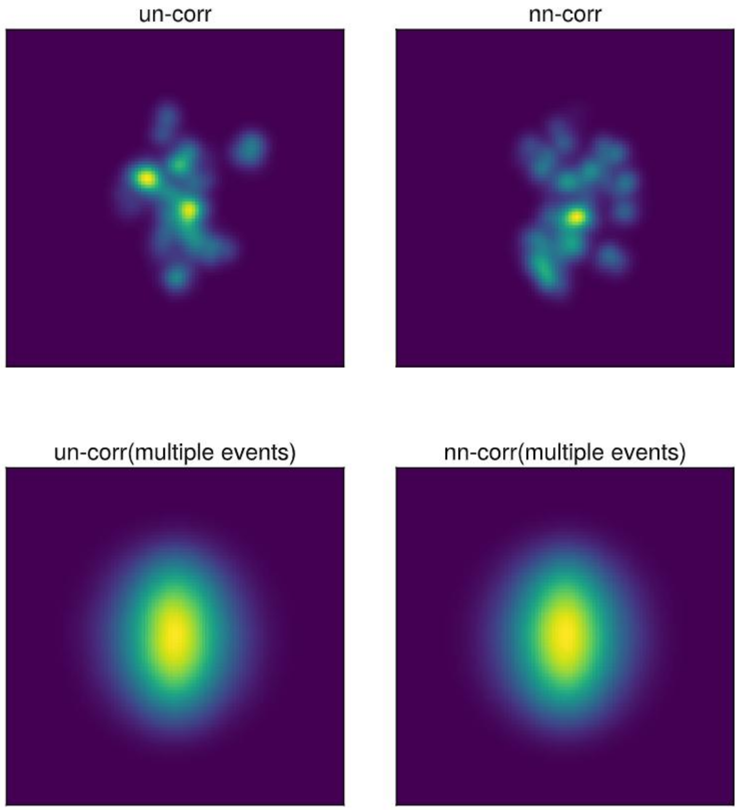
The single nucleon distributions keep the same.



The two nucleon distributions have small difference at short Δr

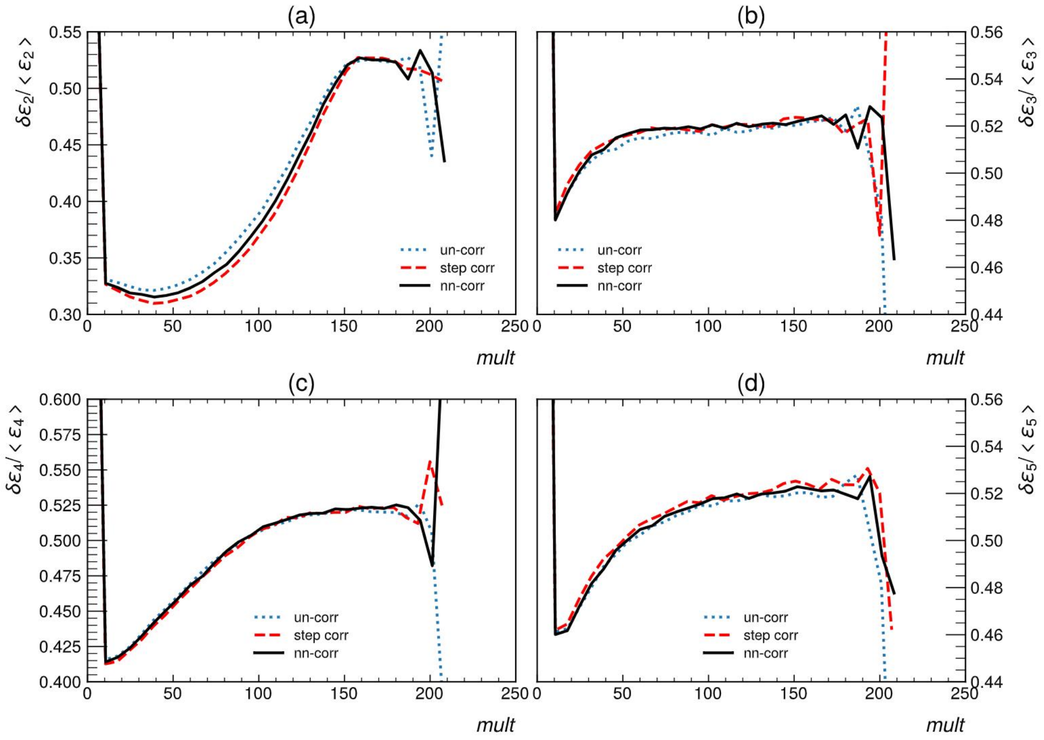


Visually no difference for initial energy density distribution

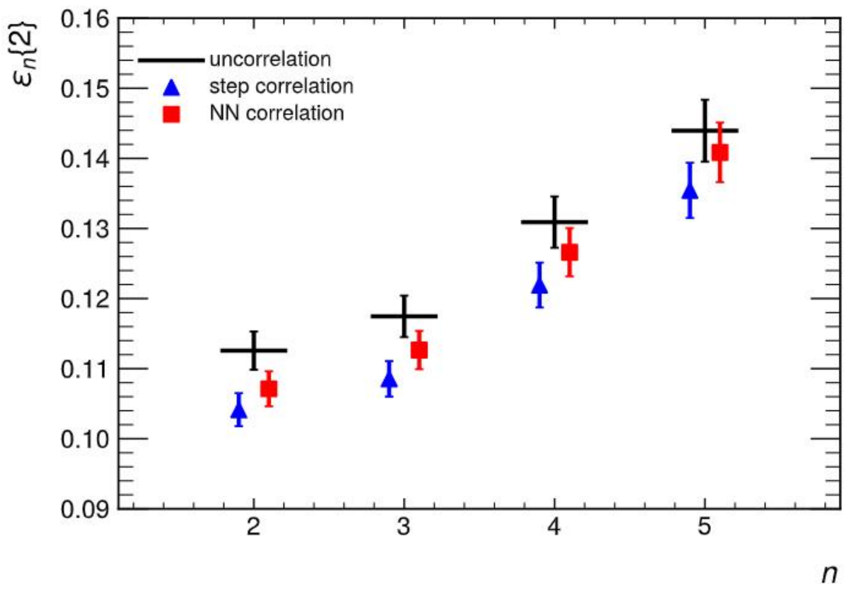


The effect on flow fluctuations (ini state)

Small but visible difference in the geometric eccentricity fluctuations at initial state.

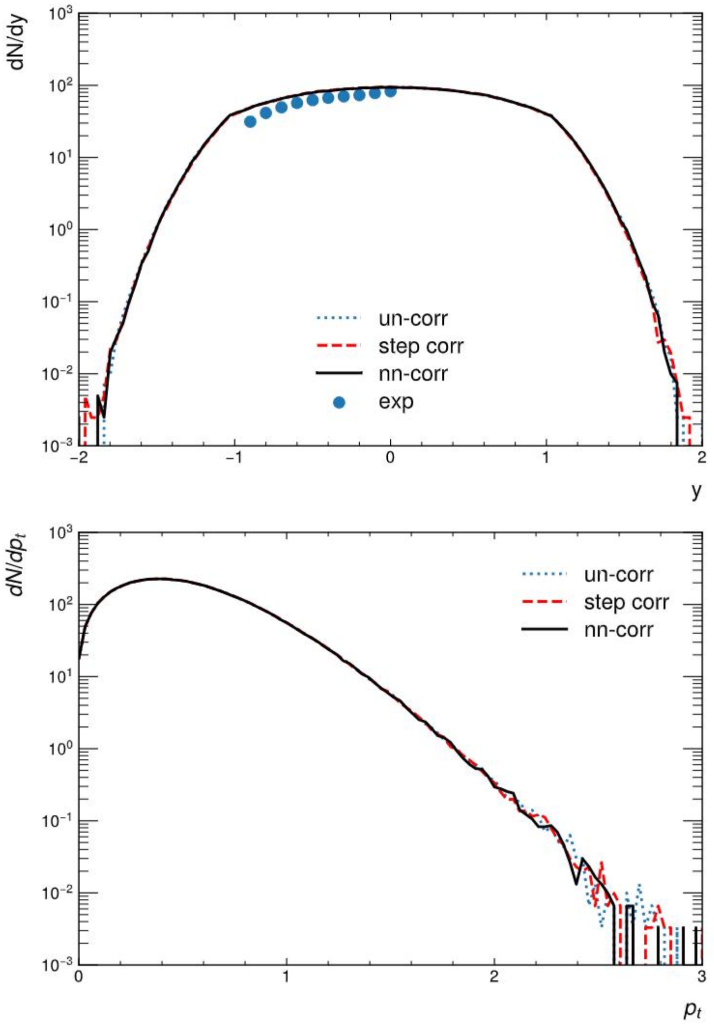


The v3 to v2 ratio puzzle for **ultra central collisions** is solved partially by two nucleon distribution.

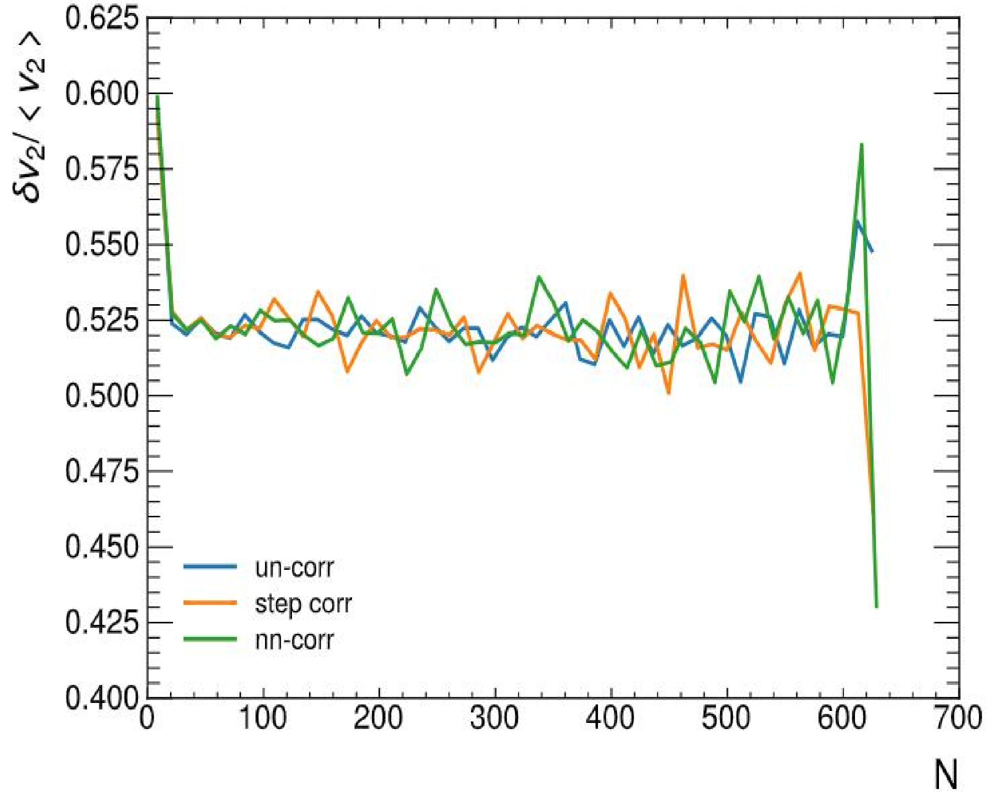


See also G. S. Denicol, C. Gale, S. Jeon, J. F. Paquet and B. Schenke, arXiv:1406.7792.

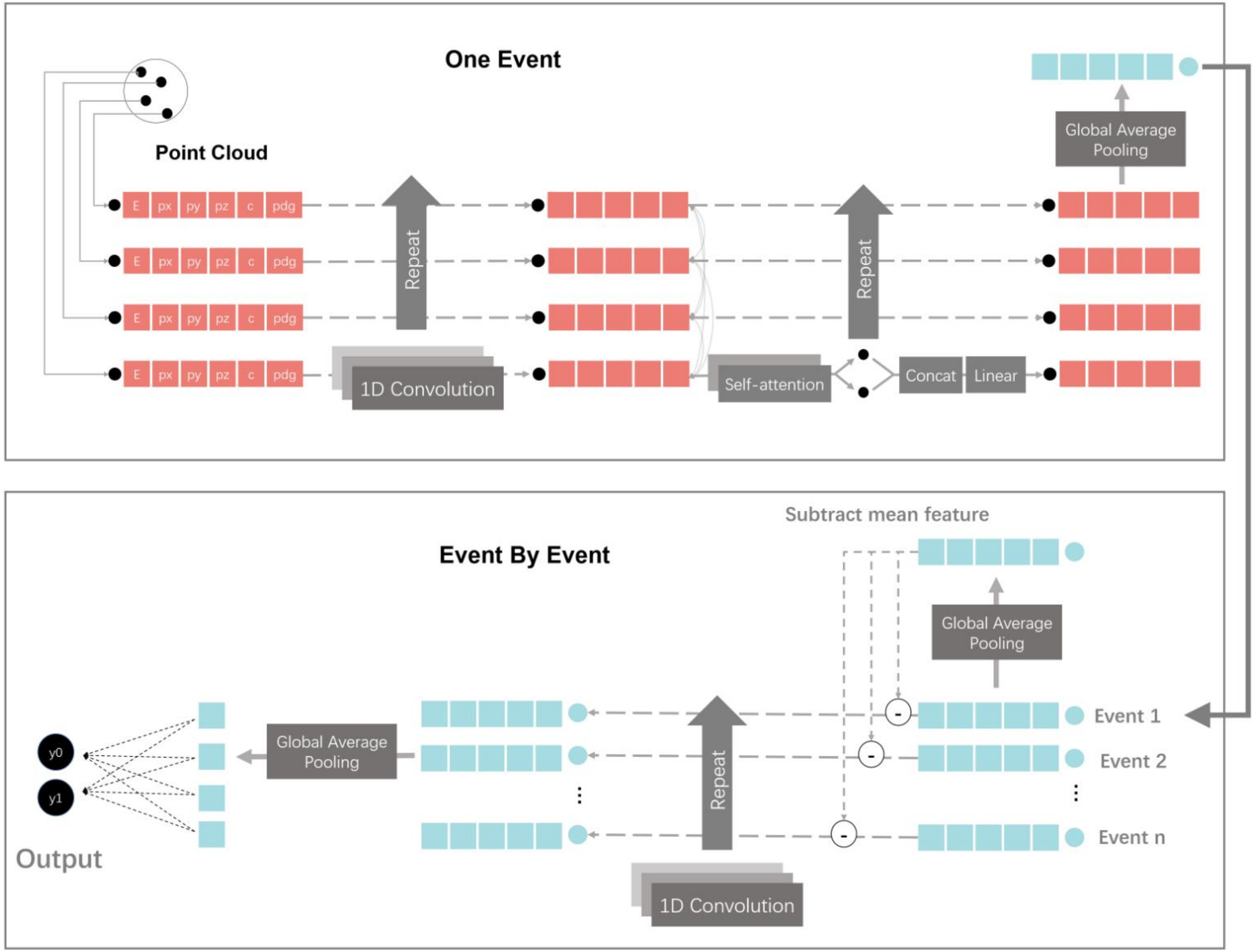
Effect of 2-nucleon dist. on final state obs.



No visible difference is observed using traditional observables.



Using a deep neural network



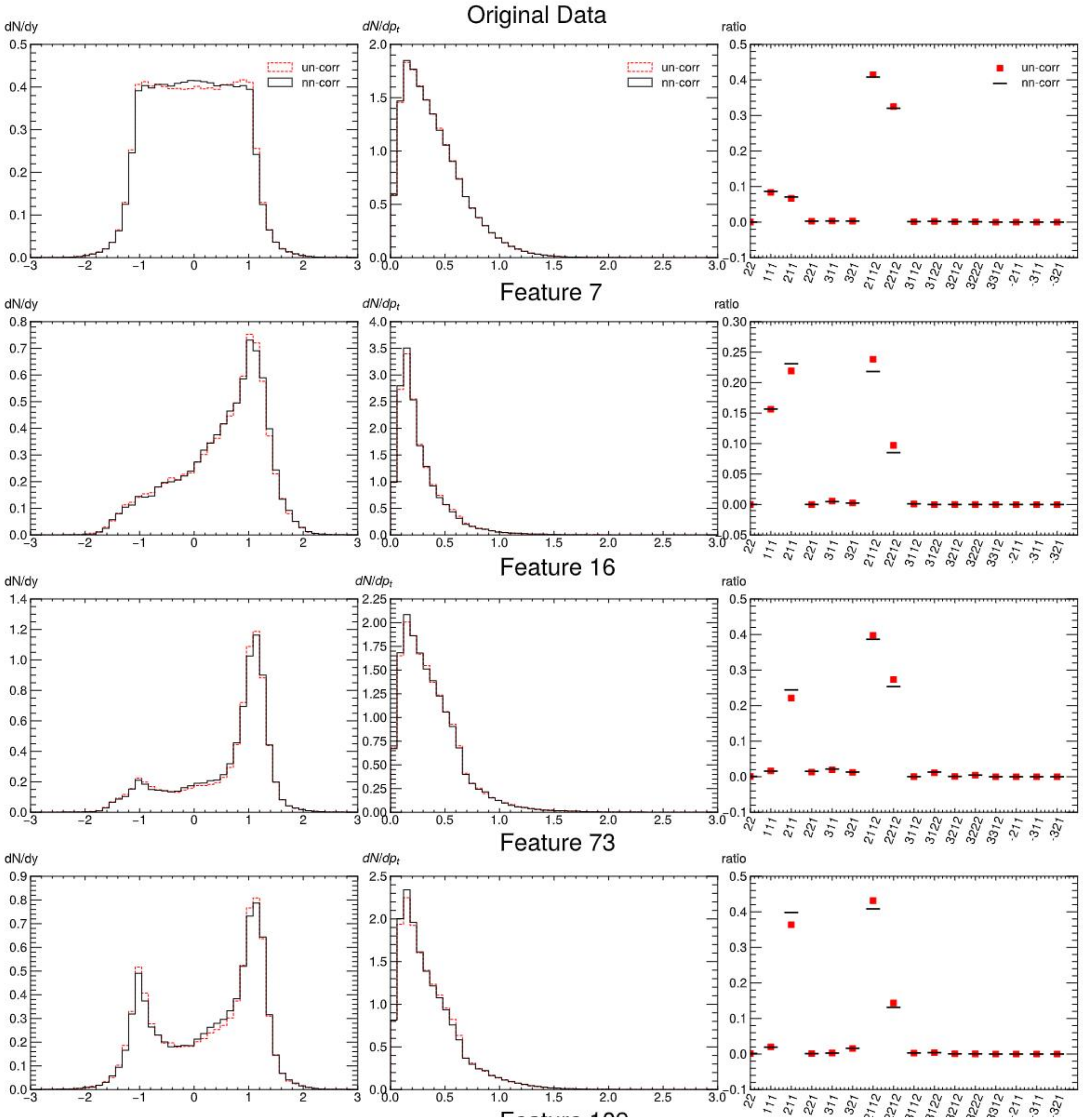
Au+Au $\sqrt{s_{NN}} = 3$ GeV, 0.5 Million events

TABLE III. Two-by-two classification accuracy for 50 combined events at different centralities in SMASH.

	0% - 20%	20% - 40%	40% - 95%
un-corr & step corr	85%	83%	69%
un-corr & nn-corr	70%	67%	56%
step corr & nn-corr	69%	68%	62%

- PointCloud for event-by-event classification and traditional multi-event mixing method fail!
- PointCloud Network + Self-attention + Statistical information of latent features in high dimensional space succeed.
- The classification accuracy is highest for central collisions using deep neural network!

Features learned by the network



- ➢ Interpretable ML can provide some inspiration what has been learned by the deep neural network
- ➢ By prediction difference analysis (through feature masking), we select most important features and visualize events and particles that maximizes these features.
- ➢ What deep learning tells us:
  - Low pt particles are important
  - Particles at large rapidity are important
  - Particle ratios are important

Summary

- DL is good at solving **inverse problems in HIC** to extract the nuclear structure
- Using the (v2, multiplicity) plot, **deep learning can predict the absolute values of nuclear deformation factors**
- Using **multi-event mixing**, the network can identify the alpha cluster in O using AMPT simulations of O+O collisions
- CLVisc simulations show that **4-alpha in O leads to different centrality dependencies** of charged multiplicity and anisotropic flows
- Two nucleon distribution is hard to identify using HIC.
- The DL method tells:
    1. Statistical information of high dimensional latent features are important for classification
    2. NN correlation signals are stronger in central collisions
    3. Look for particles(or their ratio) at small pt and large rapidity