Machine learning based estimator for elliptic flow in heavy-ion collisions

Based on: N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi,
Phys. Rev. D. 105, 114022 (2022)

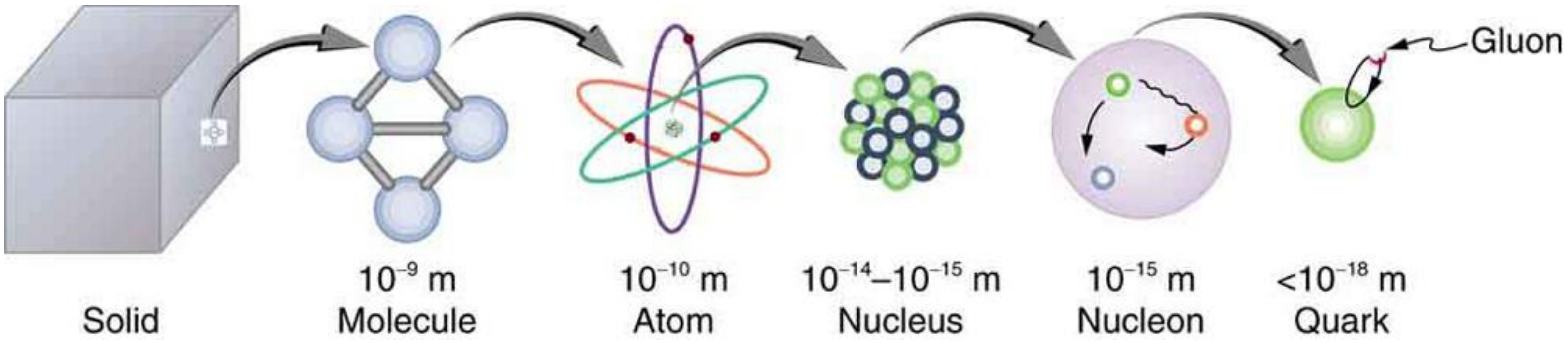
GPU Day 2022
Massive parallel computing for science and industrial application
20/06/2022

Suraj Prasad
Indian Institute of Technology Indore, India
Email: suraj.prasad@cern.ch

Outline

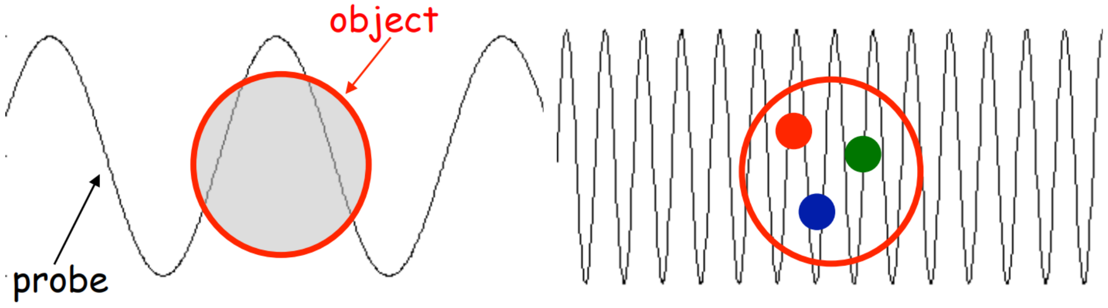
- Matter – from molecules to quarks
- Heavy-ion collisions and Quark gluon plasma
- Elliptic flow
- Deep Neural Network
- Input to the Model
- DNN Architecture
- Quality Assurance
- Results

Matter – from molecules to quarks



- Quarks – fundamental bits of matter; Gluons – mediators for strong interaction
- Quarks and gluons have color degrees of freedom and can't exist freely

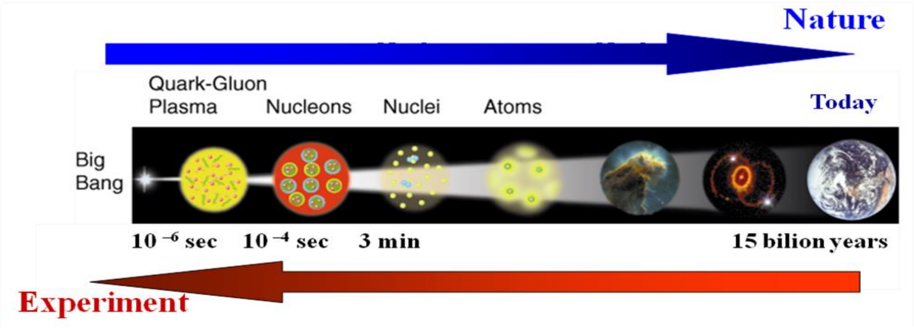
Hadrons
- Baryons (qqq)
- Mesons $(\bar{q}q)$

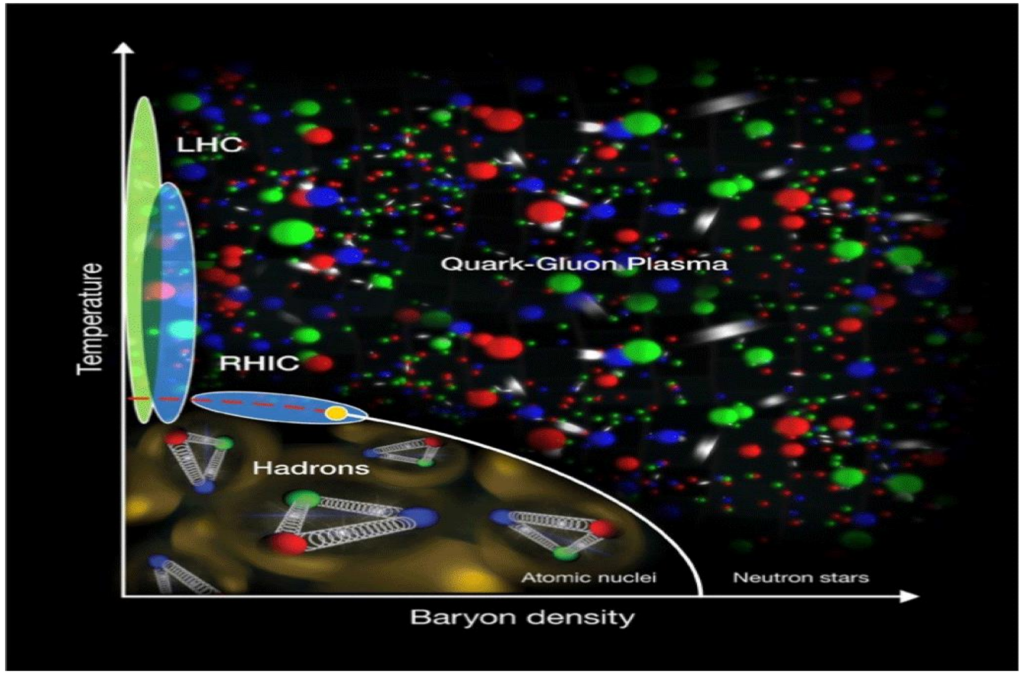


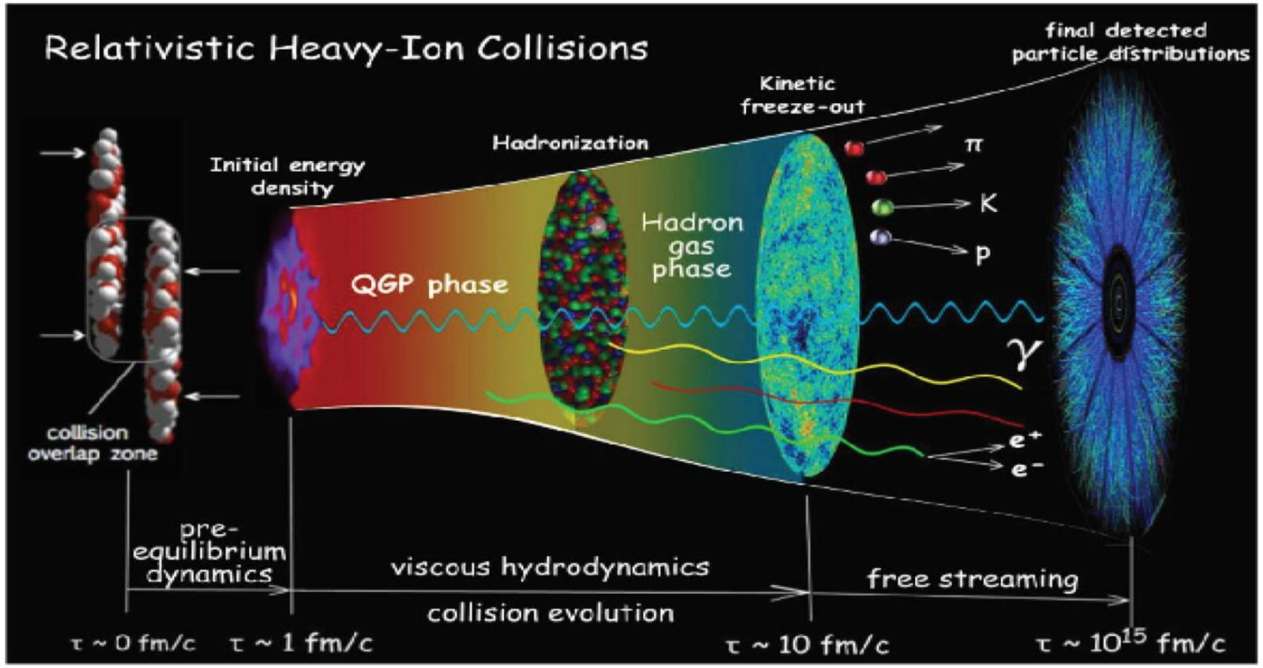
[1] R. Sahoo, "Relativistic Kinematics", [arXiv:1604.02651 [nucl-ex]]

Heavy-ion collisions (HIC) and Quark gluon plasma (QGP)

- Quark gluon plasma (QGP) is a hot and dense state of deconfined quarks and gluons in thermal equilibrium
- The value of the strong coupling constant decreases as the energy density increases: Asymptotic Freedom

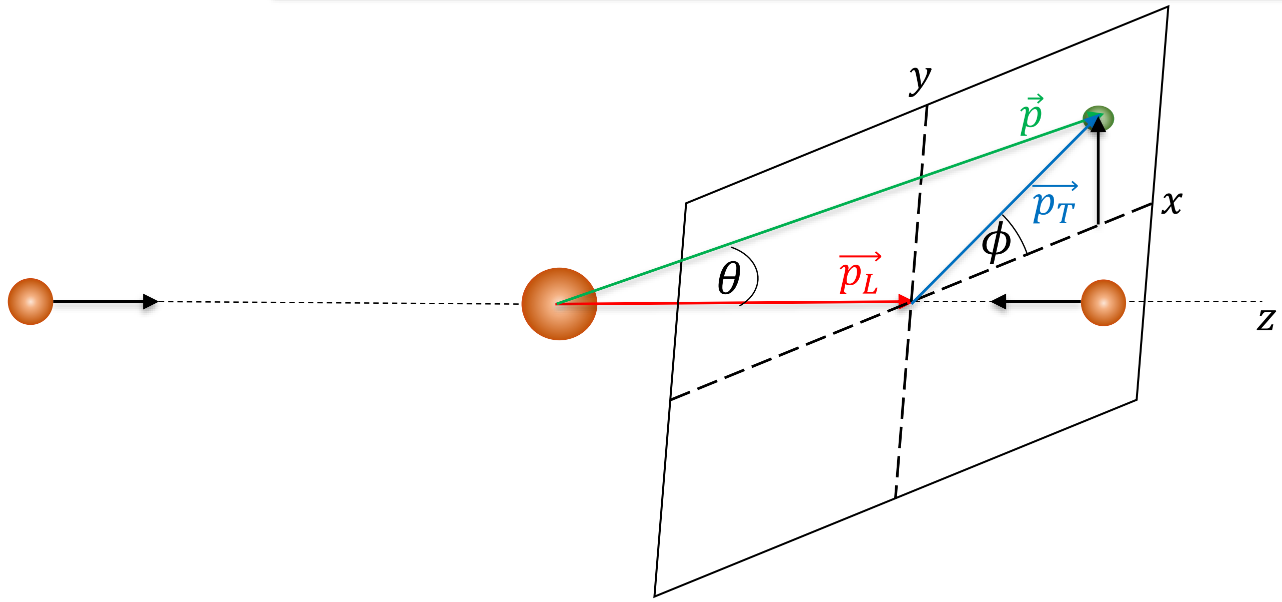


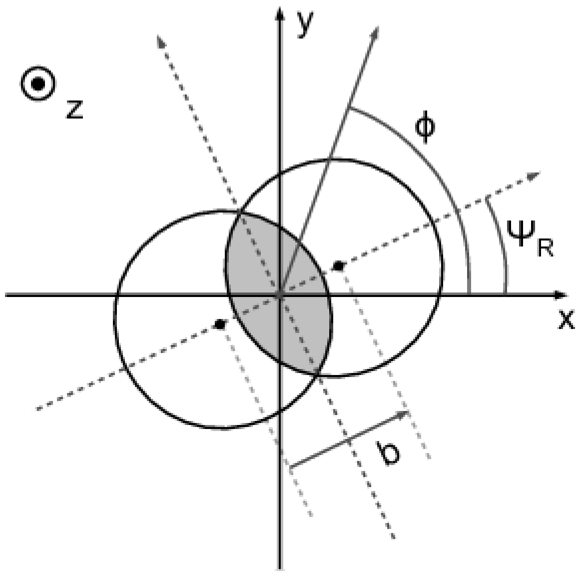




[1] R. Sahoo, AAPPS Bull. 29, 16 (2019).
[2] U. Heinz, Int. J. Mod. Phys. A 30, 1530011 (2015).
[3] R. Sahoo, and T. K. Nayak, Curr. Sci. 121, 1403 (2021).

Kinematic Observable in HIC

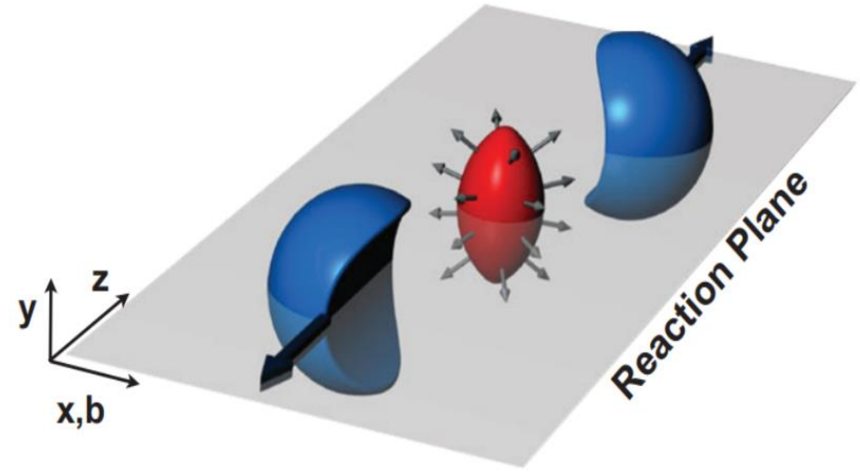


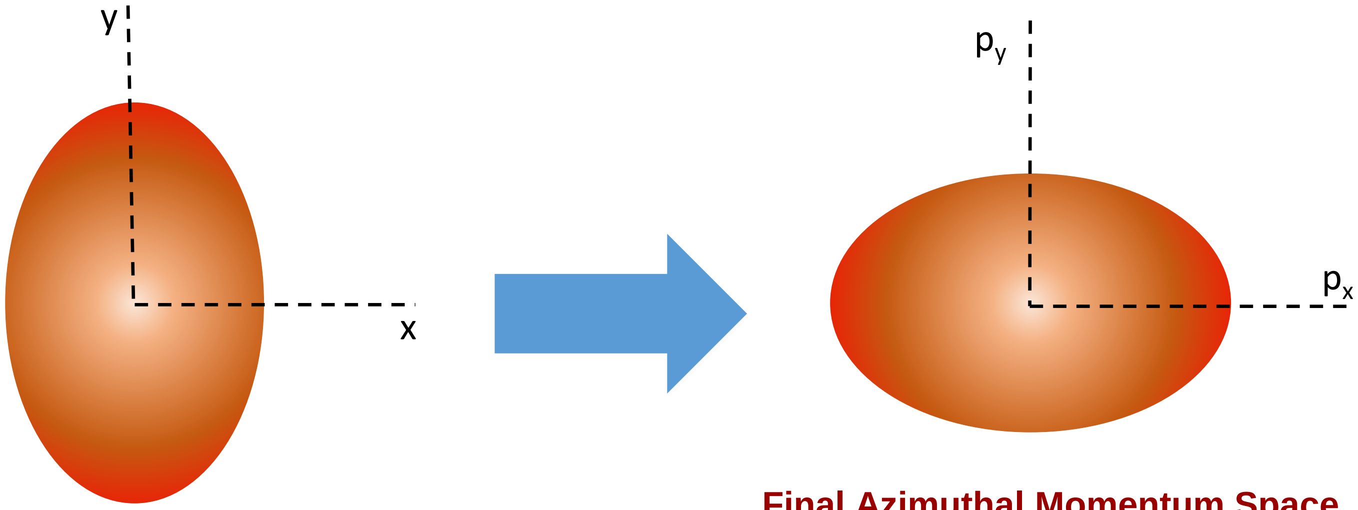


- Transverse Momentum, $p_T = \sqrt{p_x^2 + p_y^2}$
- Azimuthal Angle, $\phi = \tan^{-1}\left(\frac{p_y}{p_x}\right)$
- Polar angle, $\theta = \tan^{-1}\left(\frac{p_T}{p_z}\right)$
- Rapidity, $y = \frac{1}{2}\ln\left(\frac{E+p_Z}{E-p_Z}\right)$
- Pseudo-rapidity, $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$
- Reaction plane angle, ψ_R: Angle made by impact parameter (b) with x-axis

[1] R. Sahoo, "Relativistic Kinematics", [arXiv:1604.02651 [nucl-ex]]

Elliptic Flow (v_2)





Initial Spatial Anisotropy

Final Azimuthal Momentum Space Anisotropy

- Anisotropic flow: hydrodynamic response to spatial deformation of the initial density profile.

$$E\frac{d^3N}{dp^3} = \frac{d^2N}{2\pi p_T dp_T dy}\left(1 + 2\sum_{n=1}^{\infty} v_n \cos[n(\phi - \psi_n)]\right) = \frac{d^2N}{2\pi p_T dp_T dy}\left(1 + \underbrace{2v_1 \cos(\phi - \psi_1)}_{\text{Directed flow}} + \underbrace{2v_2 \cos[2(\phi - \psi_2)]}_{\text{Elliptic flow}} + \underbrace{2v_3 \cos[3(\phi - \psi_3)}_{\text{Triangular flow}} + \ldots]\right)$$

$$\boxed{v_2 = \langle \cos[2(\phi - \psi_2)] \rangle}$$

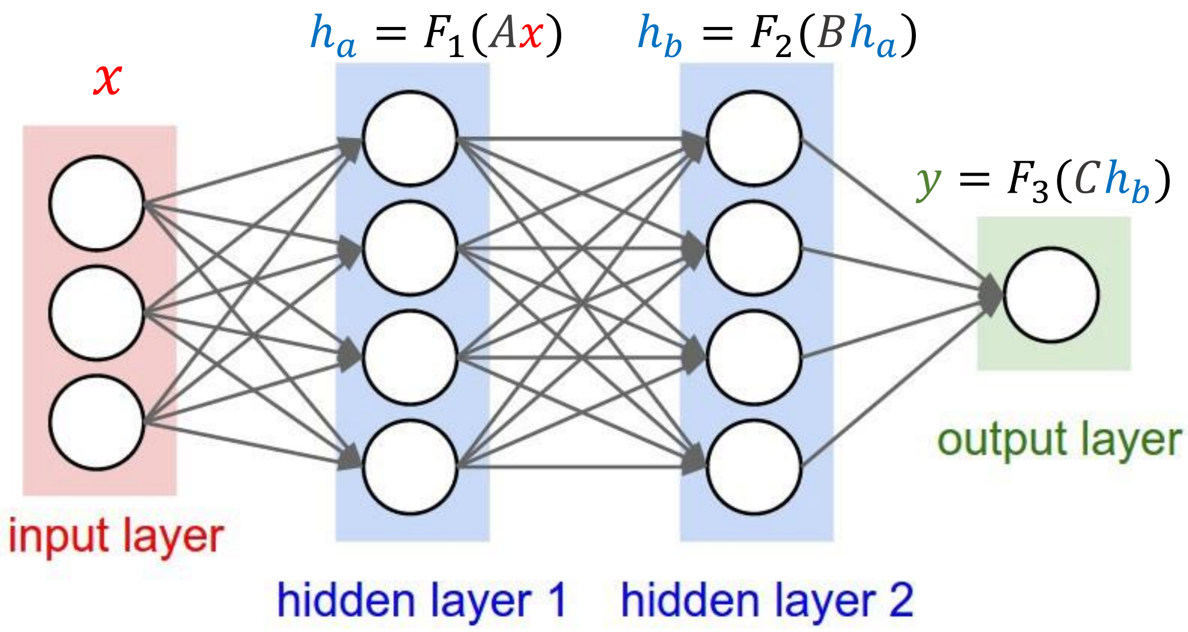
N. Mallick, R. Sahoo and S. Tripathy, and A. Ortiz, J. Phys. G 48, 045104 (2021)
B. B. Abelev et al. [ALICE Collaboration], JHEP 1506, 190 (2015)

Deep Neural Network (DNN)

- It is an ML algorithm inspired by the structure of the neurons in the animal brain
- Three key layers

  Input layer: Takes different features as input
  Hidden layer: Connects to different neurons to weights
  Output layer: gives the result as a number or class

- Weights: represents the importance of a neuron
- Activation function: Guides the outcome of each node
- Cost function: Evaluates the accuracy of the prediction
- Optimizer: Methods/Algorithms used to minimize the cost function by updating the weights



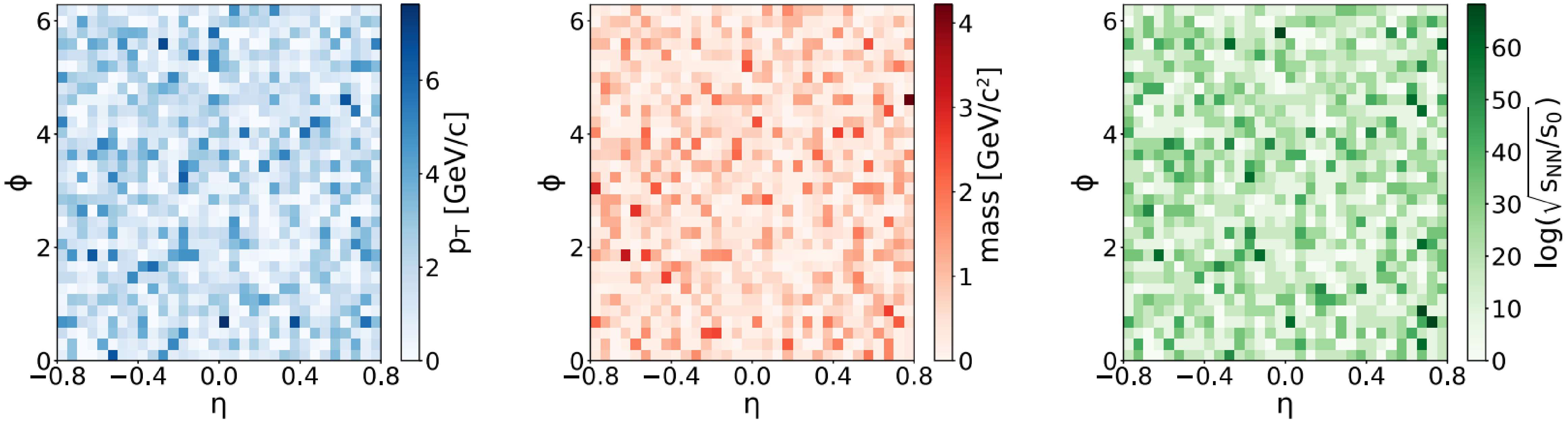
- For an input layer x, two hidden layers, the output y can be represented as:

$$y = F_3\left(CF_2\left(BF_1(Ax)\right)\right)$$

A, B, C represent the weight matrices
F_1, F_2, F_3 represent the activation functions

Input to the Model

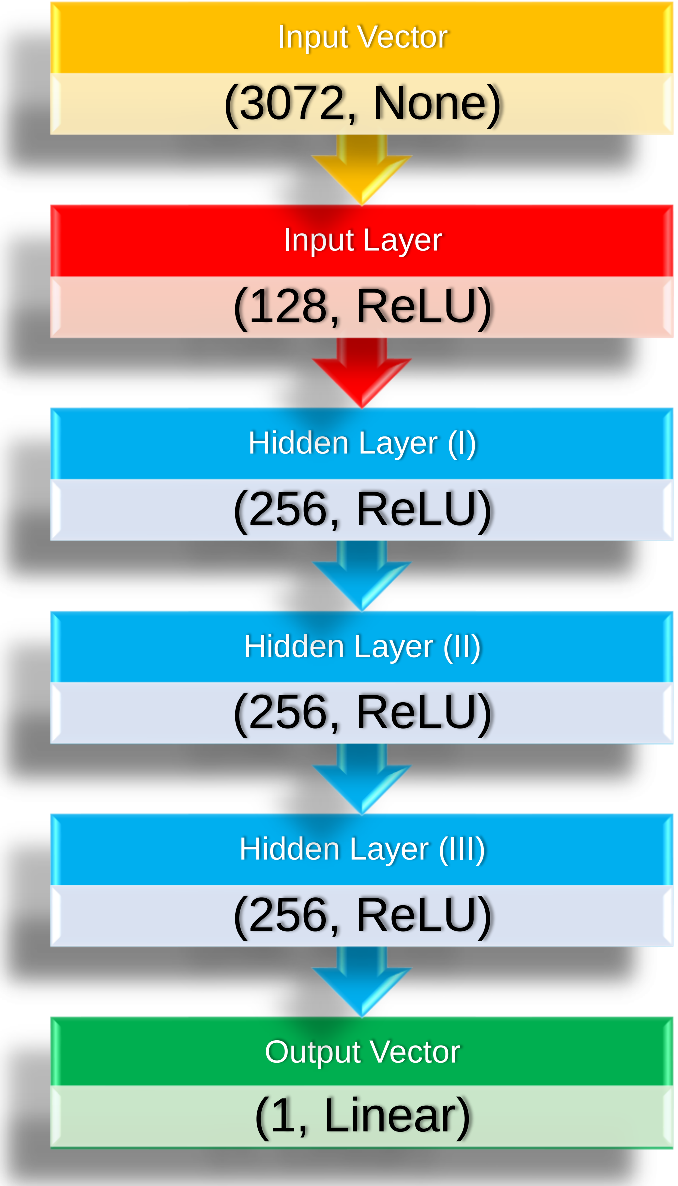
- Elliptic flow (v_2) is an event property, depends upon: η, mass, p_T, centrality, etc.
- v_2 is calculated using the event plane method and ψ_R is set to zero in AMPT. ($\Rightarrow\ v_2 = \langle\cos(2\phi)\rangle$)
- $(\eta - \phi)$ space is considered as primary input space, divided into (32x32) bin space
- Weighted p_T, mass and $\log(\sqrt{(s_{NN}/s_0)})$ are taken as input to the DNN architecture.
- $32 \times 32 \times 3\ (= 3072)$ features per event
- Trained with 150K minimum bias events of Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV simulated with AMPT.



N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi, Phys. Rev. D **105**, 114022 (2022)

DNN Architecture

- Each event has 3072 input features mapped into the first dense layer with 128 nodes and ReLU activation function
- 3 hidden layers with 256 nodes and ReLU activation function
- The final layer has single node as v_2 and linear activation function is used
- ReLU (x) = max$(0, x)$
- Cost Function: mean squared error
- Optimiser: adam
- Early stopping mechanism is used to stop the training
- Early stopping patience level of maximum 10 epochs
- Max Epoch=60, batch size = 32



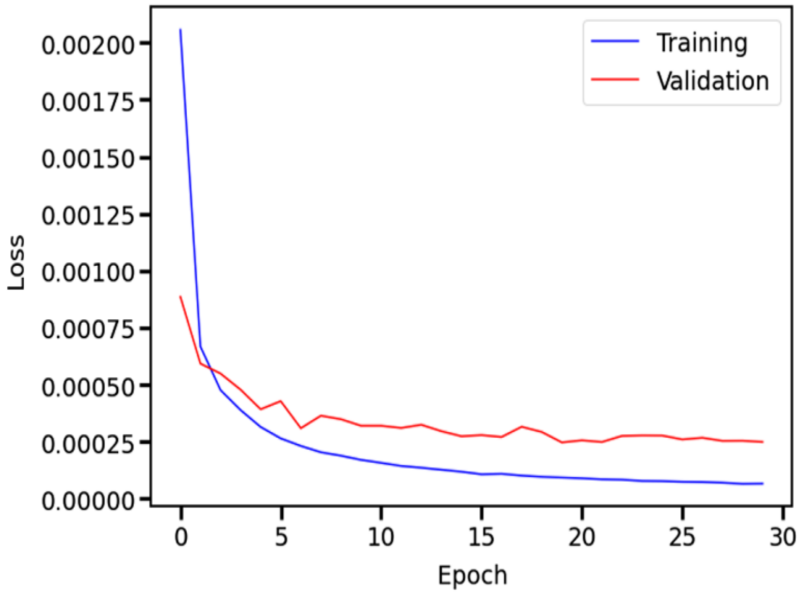
N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi, Phys. Rev. D **105**, 114022 (2022)

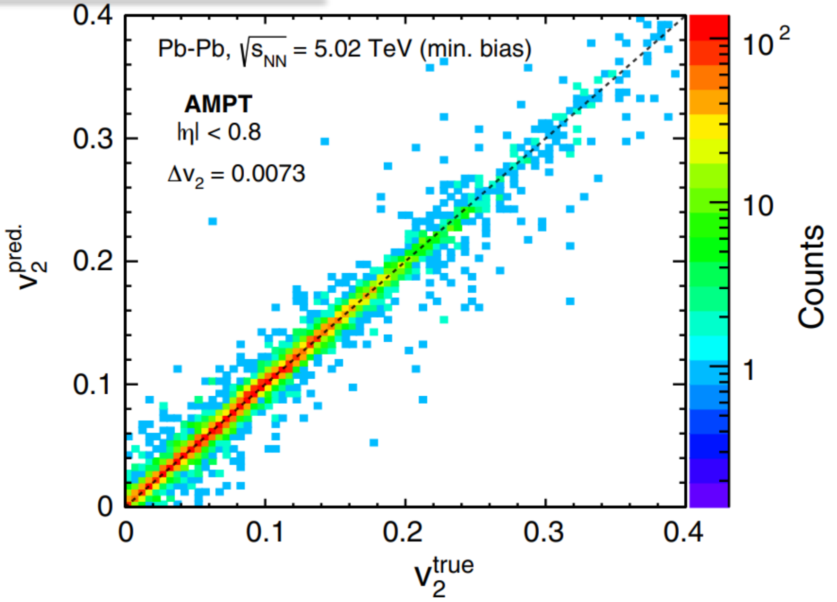
Quality Assurance

- Loss is the mean squared error of predicted v_2 w.r.t true v_2
- Loss is in the order of 10^{-4} : Good training
- Less to no overfitting/underfitting.
- Mean absolute error (MAE) of v_2 is defined as:

$$\Delta v_2 = \frac{1}{N_{\text{events}}} \sum_{n=1}^{N_{\text{events}}} |v_{2_n}^{\text{true}} - v_{2_n}^{\text{pred}}|$$

- $v_2^{true} = v_2^{pred}$ straight line is well populated
- 32x32 bin size is faster and has optimum performance compared to other bin size
- Model is quite sturdy and less sensitive to noise of different weights

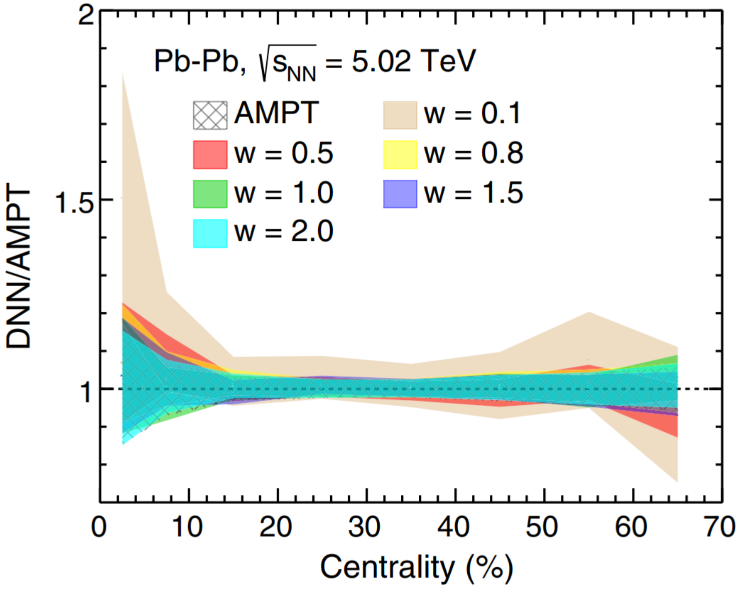




- Modified j^{th} feature of i^{th} event:

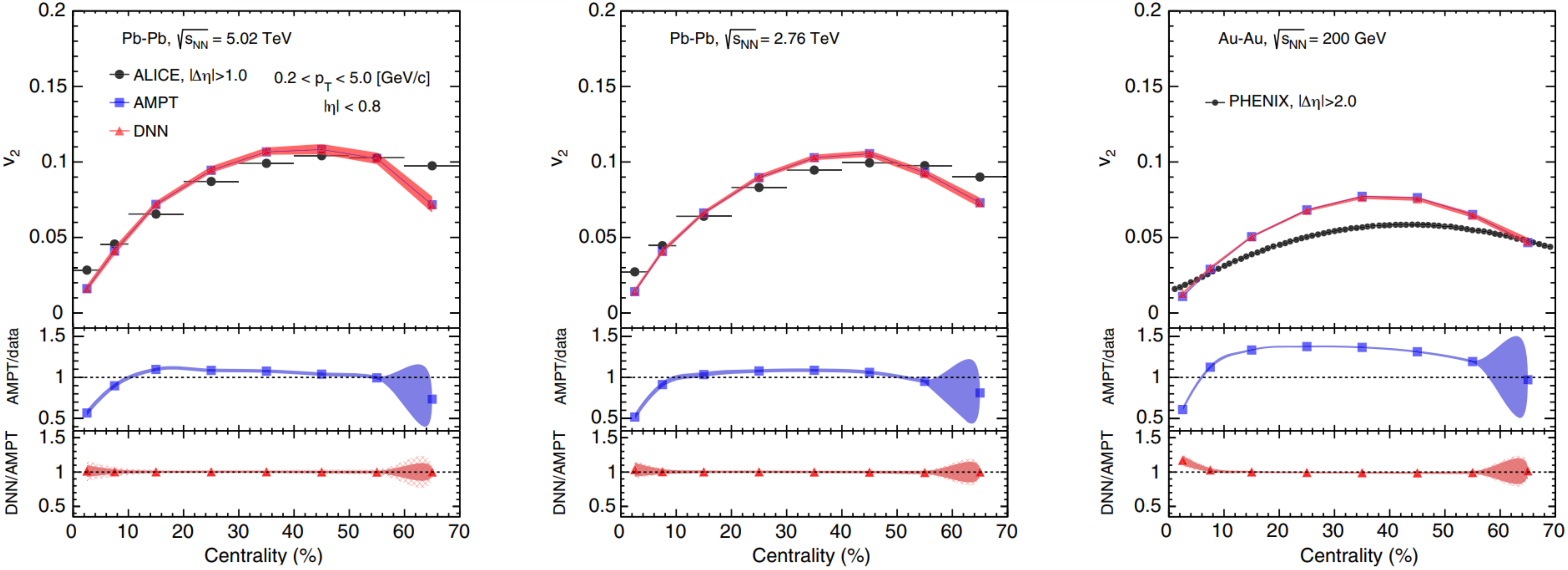
$$F'_{ij} = F_{ij} + X_{ij}/w$$

w: weight
$X_{ij} \in (-\sigma_j, \sigma_j)$ is a random number



N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi, Phys. Rev. D **105**, 114022 (2022)

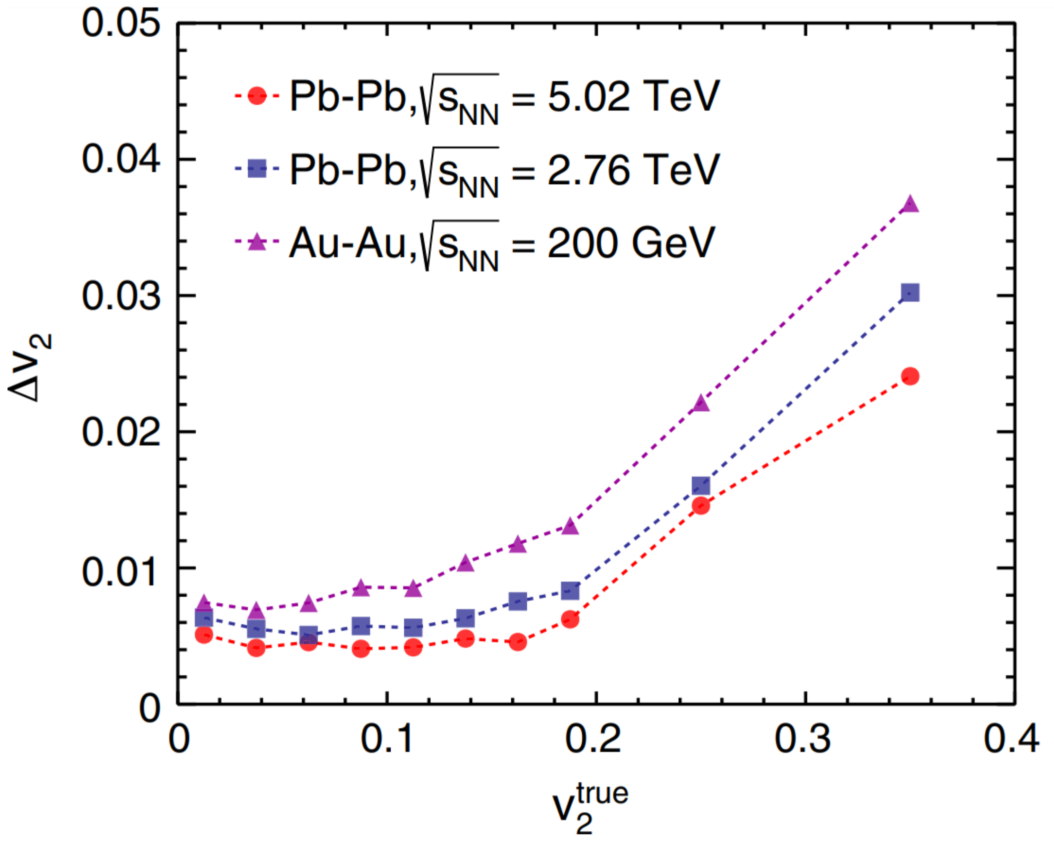
Results and Discussions

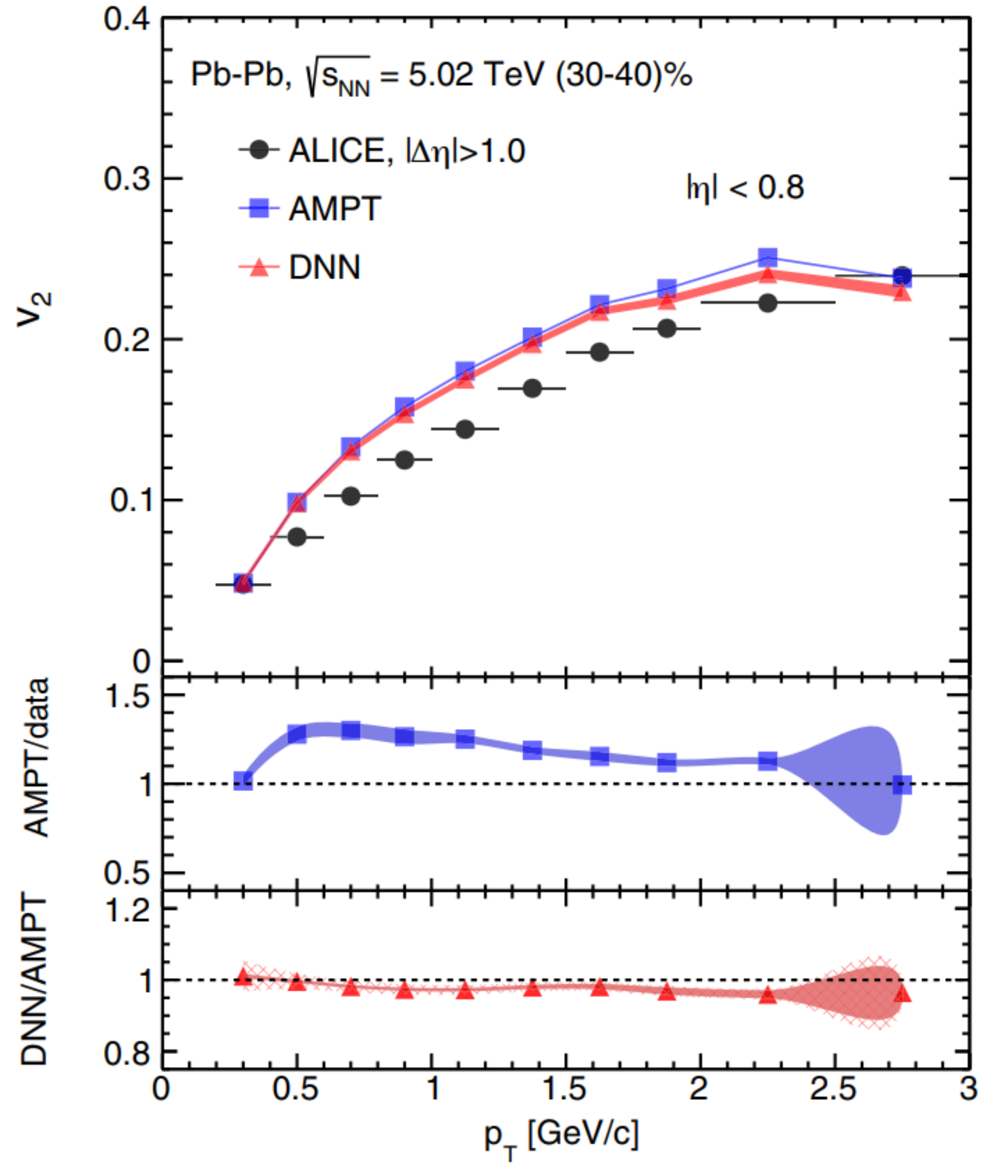


- AMPT results do not match with experimental data, which can be optimised with different settings available in AMPT
- DNN predictions almost match with the v_2 value from the AMPT, not only at LHC but also at RHIC energies, which is new to DNN

N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi, Phys. Rev. D **105**, 114022 (2022)

Results and Discussions





- Δv_2, is less than 10% of the true v_2
- Δv_2 is least for the case where it's trained
- DNN retains the p_T dependence, and matches with v_2 from AMPT in low p_T region.

N. Mallick, S. Prasad, A. N. Mishra, R. Sahoo, and G. G. Barnaföldi, Phys. Rev. D **105**, 114022 (2022)

Summary

- Elliptic flow is one of the important observables to understand the physics of QGP, and it can be successfully predicted with machine learning based DNN algorithm
- The DNN model retains the centrality, energy and p_T dependence and predicts with accuracy
- The DNN model is quite robust and it’s less sensitive to input data with uncorrelated noise

Outlook

- The DNN model is yet to be tested for particle dependence (mass dependence).
- In AMPT, we have set $\psi_R = 0$, The model is yet to be tested for random orientation of ψ and its dependence
- This DNN model takes nonflow effects into account while calculating v_2, which can be improved.

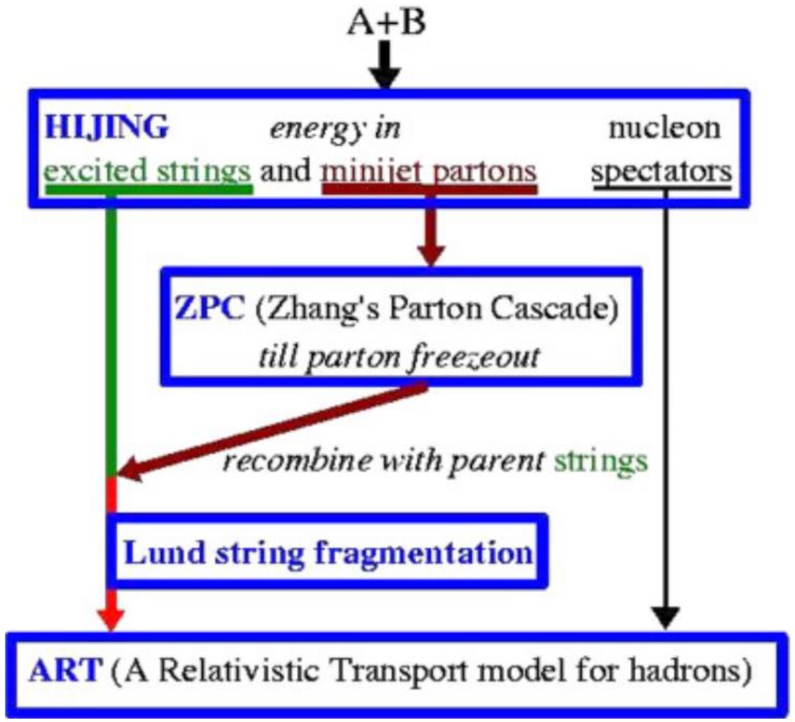
Thank You

For the attention

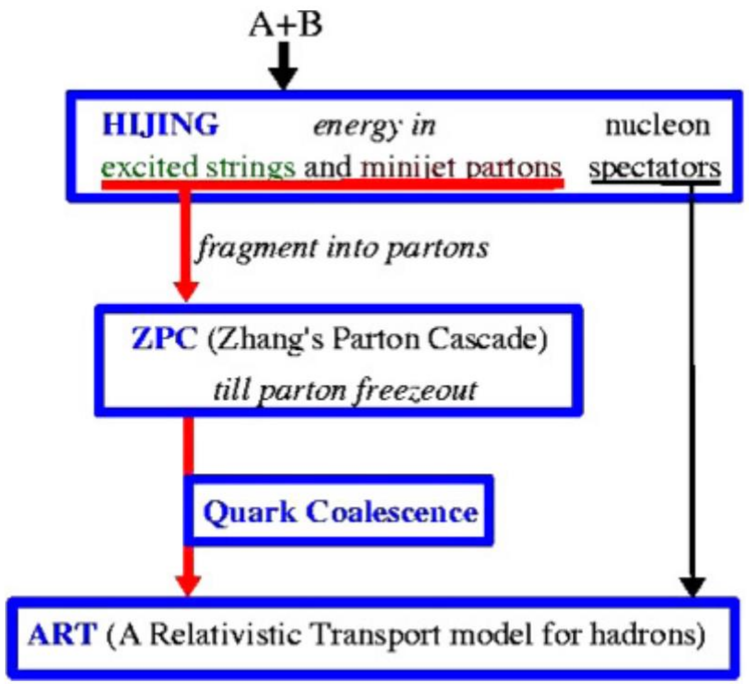
Backup Slides

A Multi-phase Transport Model (AMPT)

- Initialisation of collisions (HIJING)
- Parton Transport (ZPC)
- Hadronisation (Lund string fragmentation / Quark coalescence)
- Hadron Transport (ART)



Default Version



SM Version

Zi-Wei Lin, Che Ming Ko, Bao-An Li, Bin Zhang, and Subrata Pal, Phys. Rev. C 72, 064901 (2005)