Application of Machine Learning tools in heavy-ion collisions at the Large Hadron Collider

GPU Day 2021
11 Nov 2021

Neelkamal Mallick
Indian Institute of Technology Indore, India
Neelkamal.Mallick@cern.ch

Outline

- Heavy-ion collisions
- Boosted Decision Trees

A. Estimation of impact parameter

B. Estimation of Transverse Spherocity

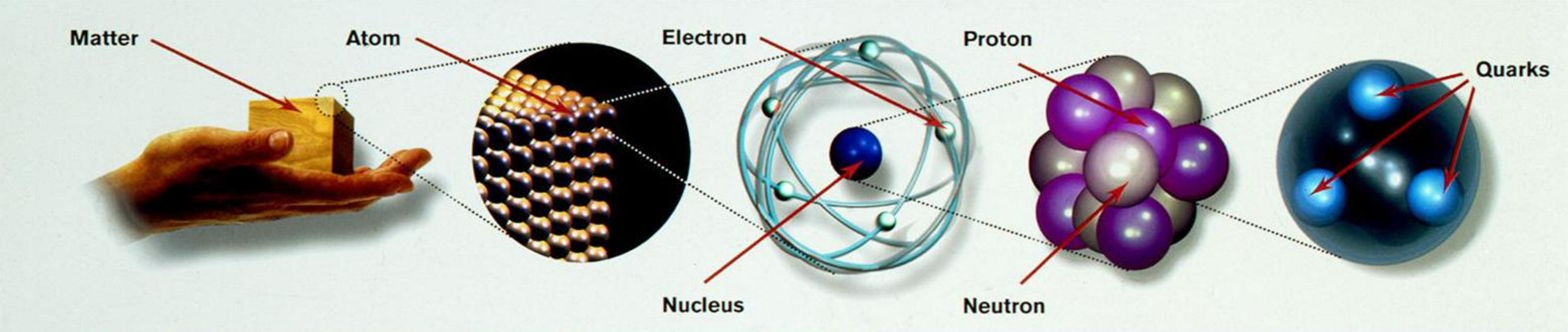
- Deep Neural Networks

C. Estimation of Elliptic flow (v_2)

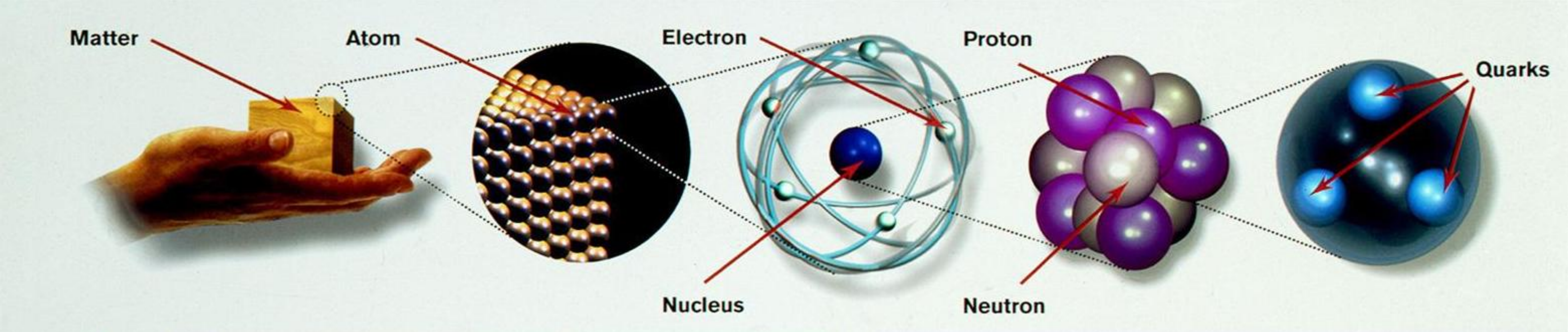
- Summary and outlook

Heavy-ion collisions

Heavy-ion collisions

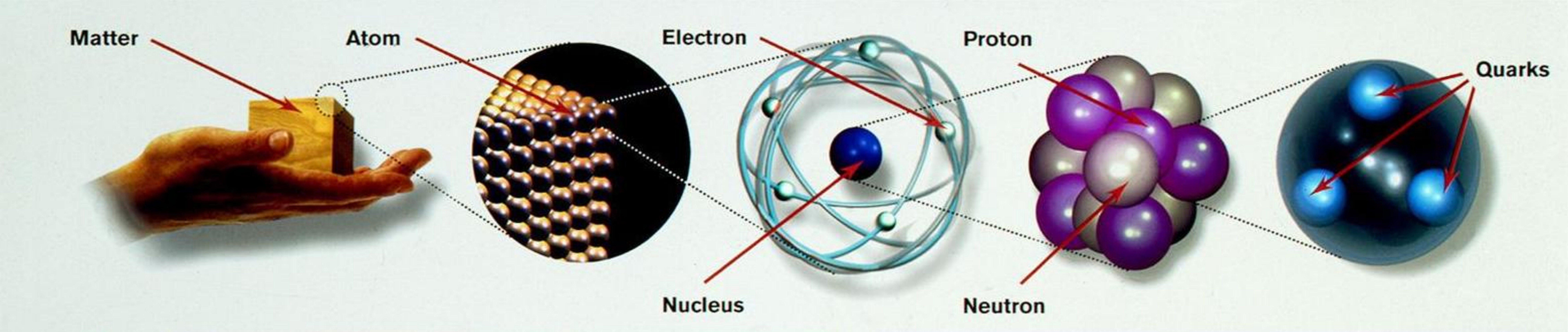


Heavy-ion collisions

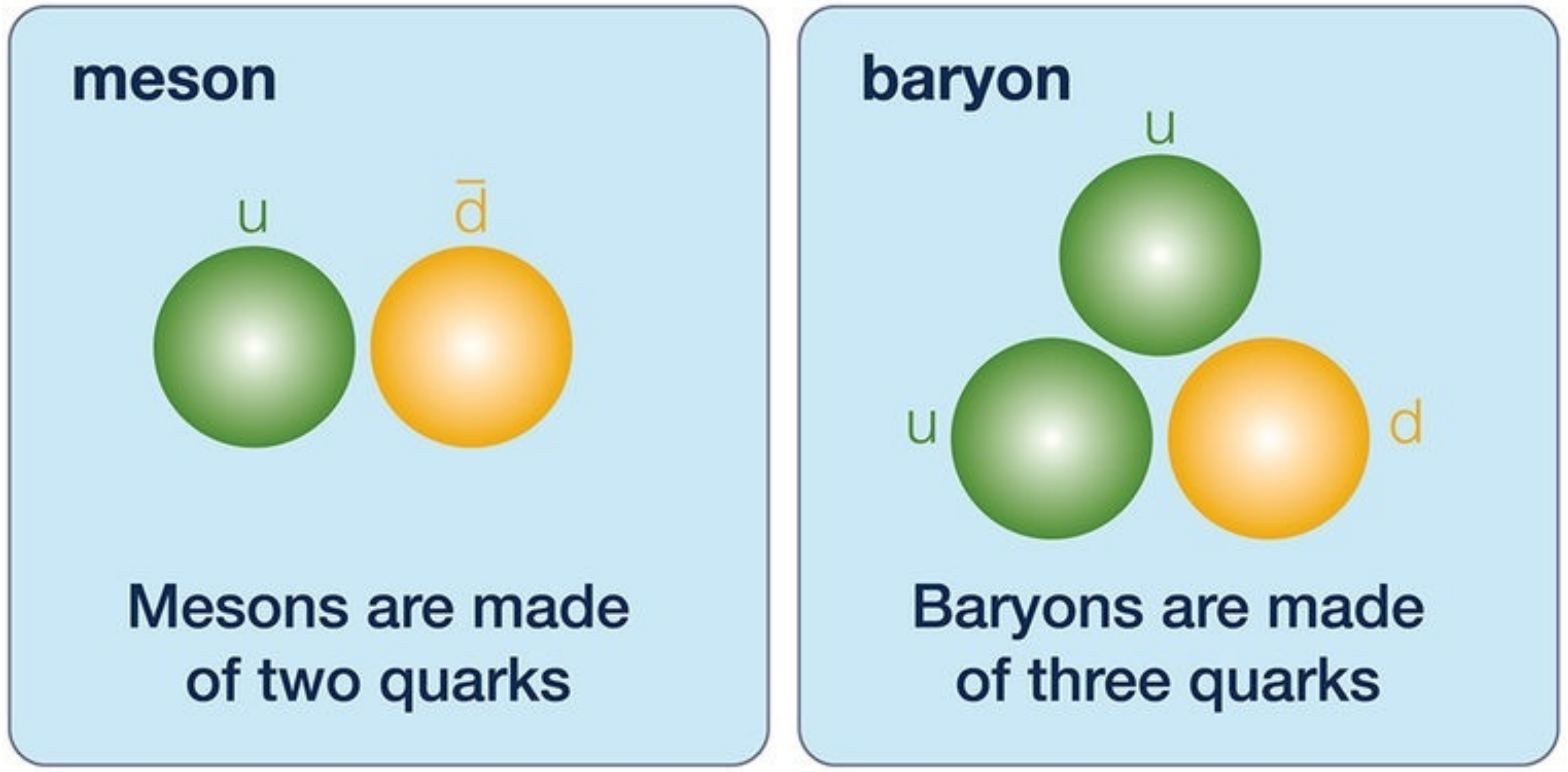


- Quarks: the fundamental bits of matter
- Gluons: carrier of strong force
- Theory of strong force: Quantum Chromodynamics (QCD)
- Quark confinement: free quarks and gluons can not exist
- Mesons ($q\bar{q}$) and Baryons (qqq) : quarks confined inside hadrons
- Heavy-ions: Pb, Au, Xe nucleus

Heavy-ion collisions



- Quarks: the fundamental bits of matter
- Gluons: carrier of strong force
- Theory of strong force: Quantum Chromodynamics (QCD)
- Quark confinement: free quarks and gluons can not exist
- Mesons ($q\bar{q}$) and Baryons (qqq) : quarks confined inside hadrons
- Heavy-ions: Pb, Au, Xe nucleus

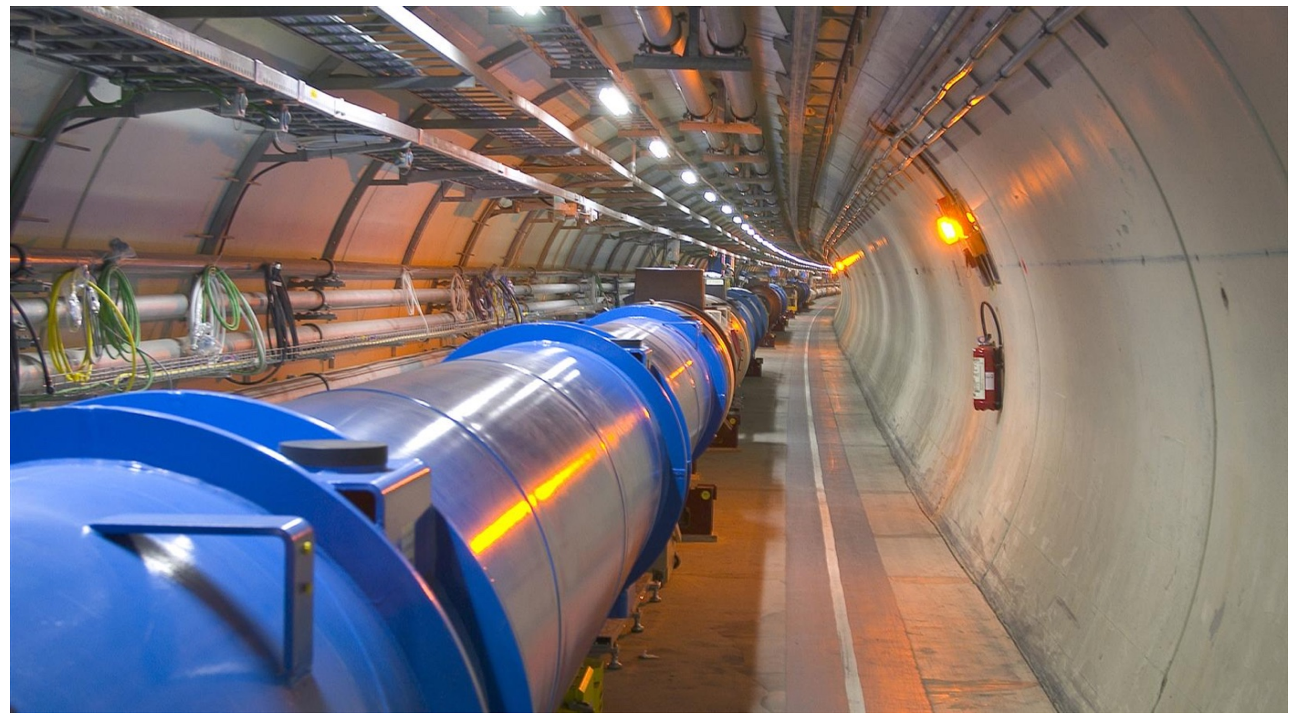


Heavy-ion collisions

Heavy-ion collisions

Large Hadron Collider, CERN, Switzerland

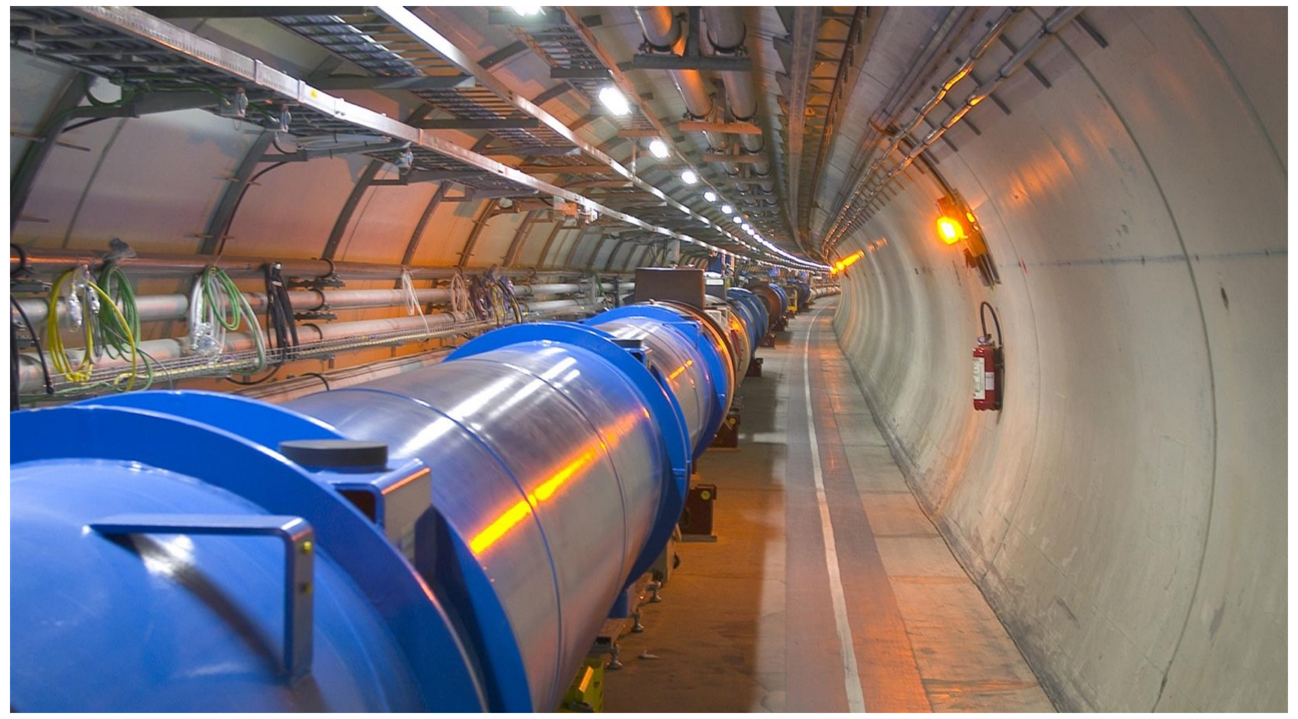
Heavy-ion collisions

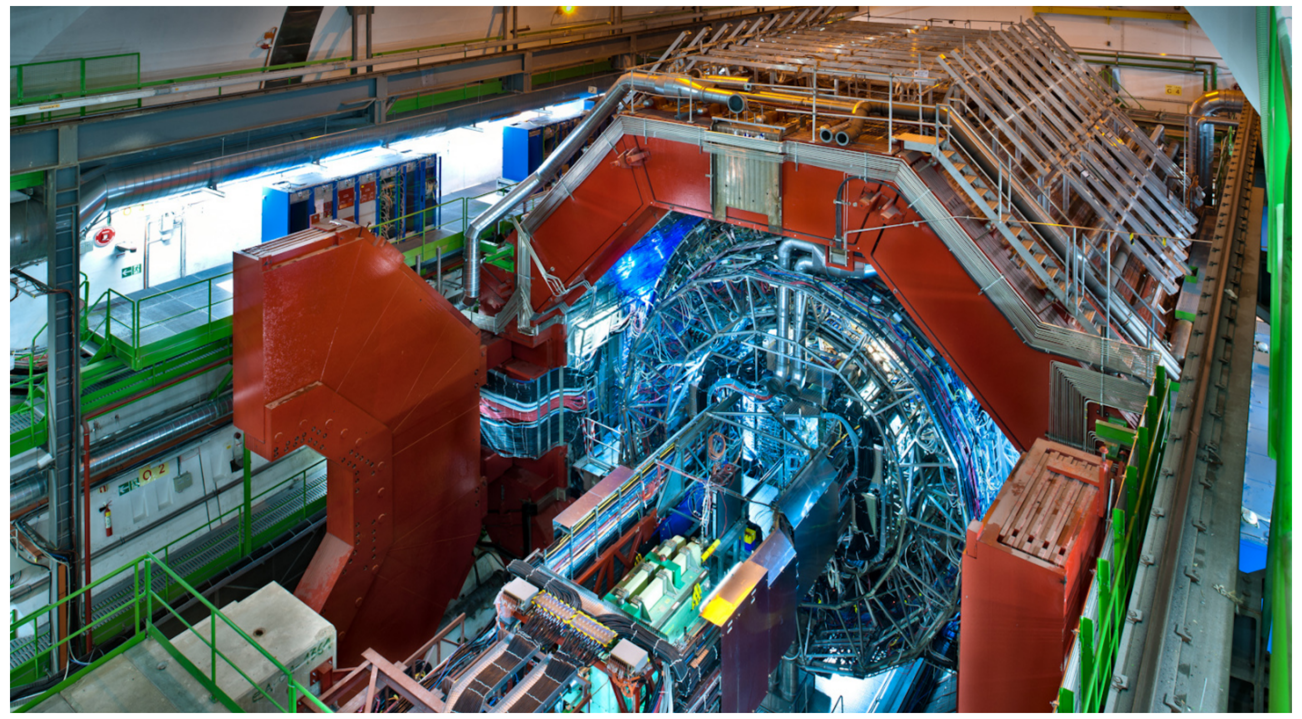


Large Hadron Collider, CERN, Switzerland

Heavy-ion collisions

Large Hadron Collider, CERN, Switzerland

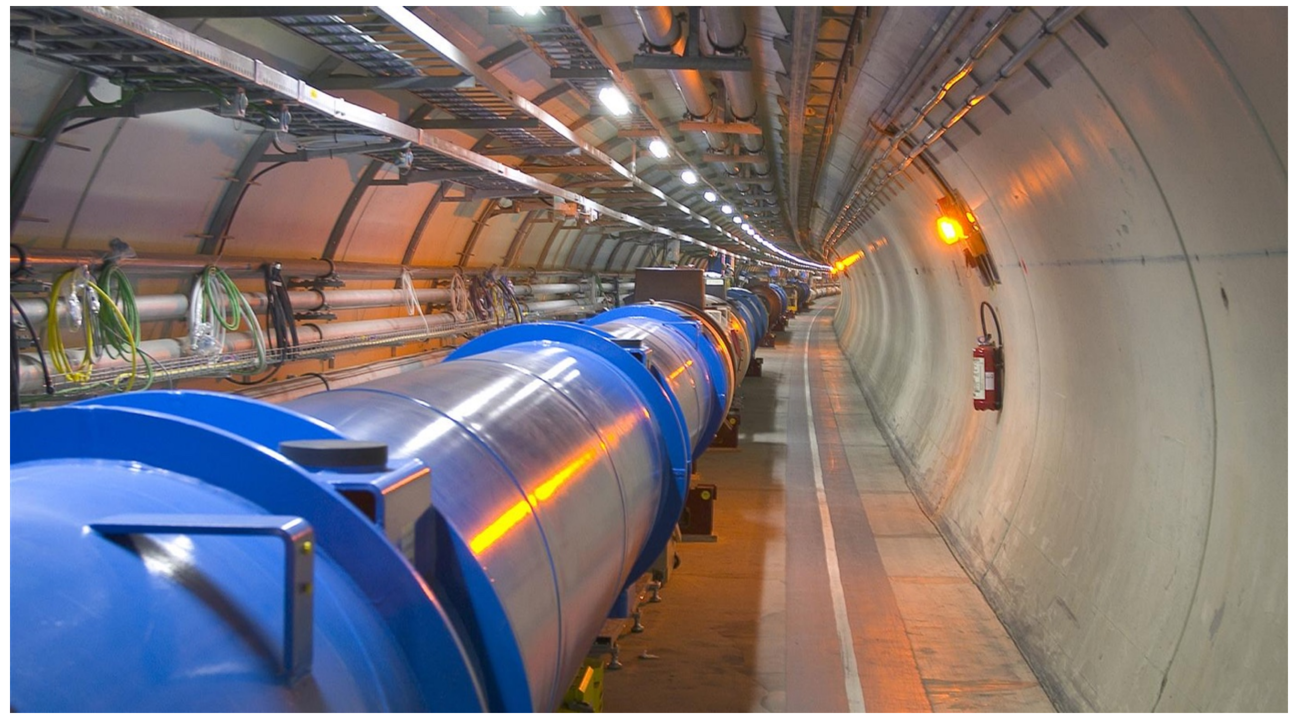


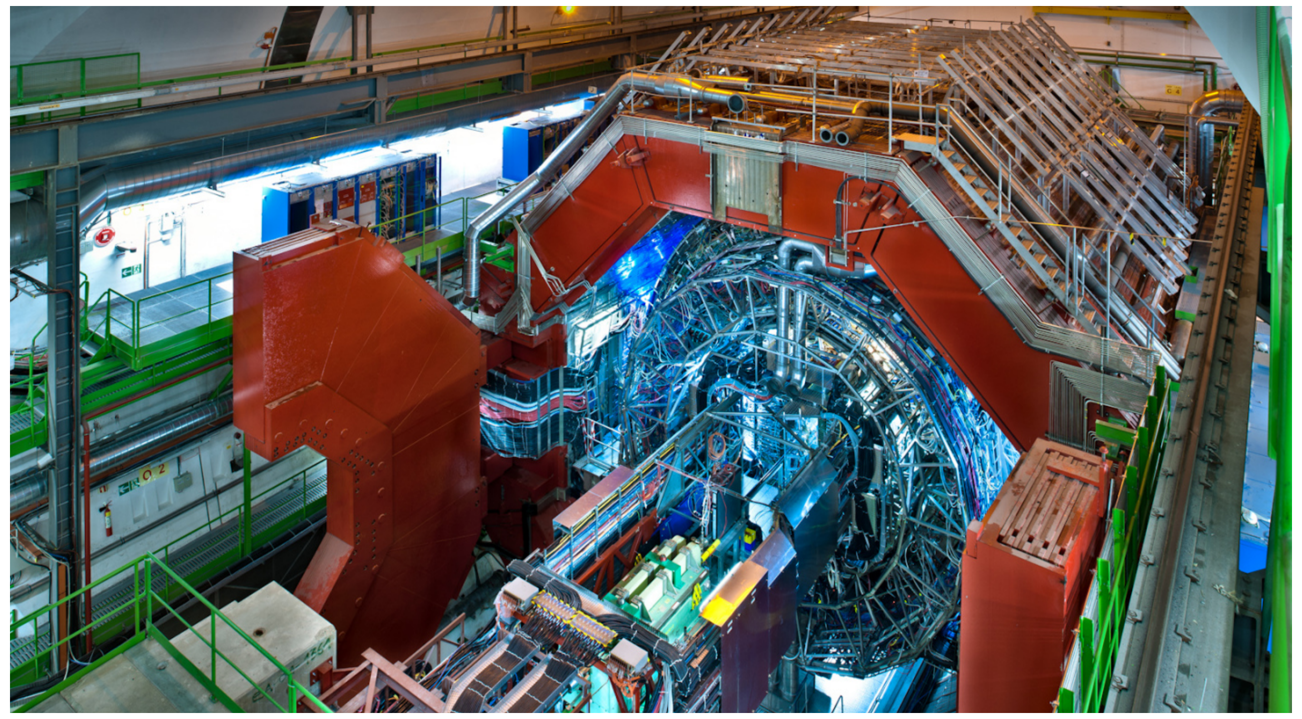


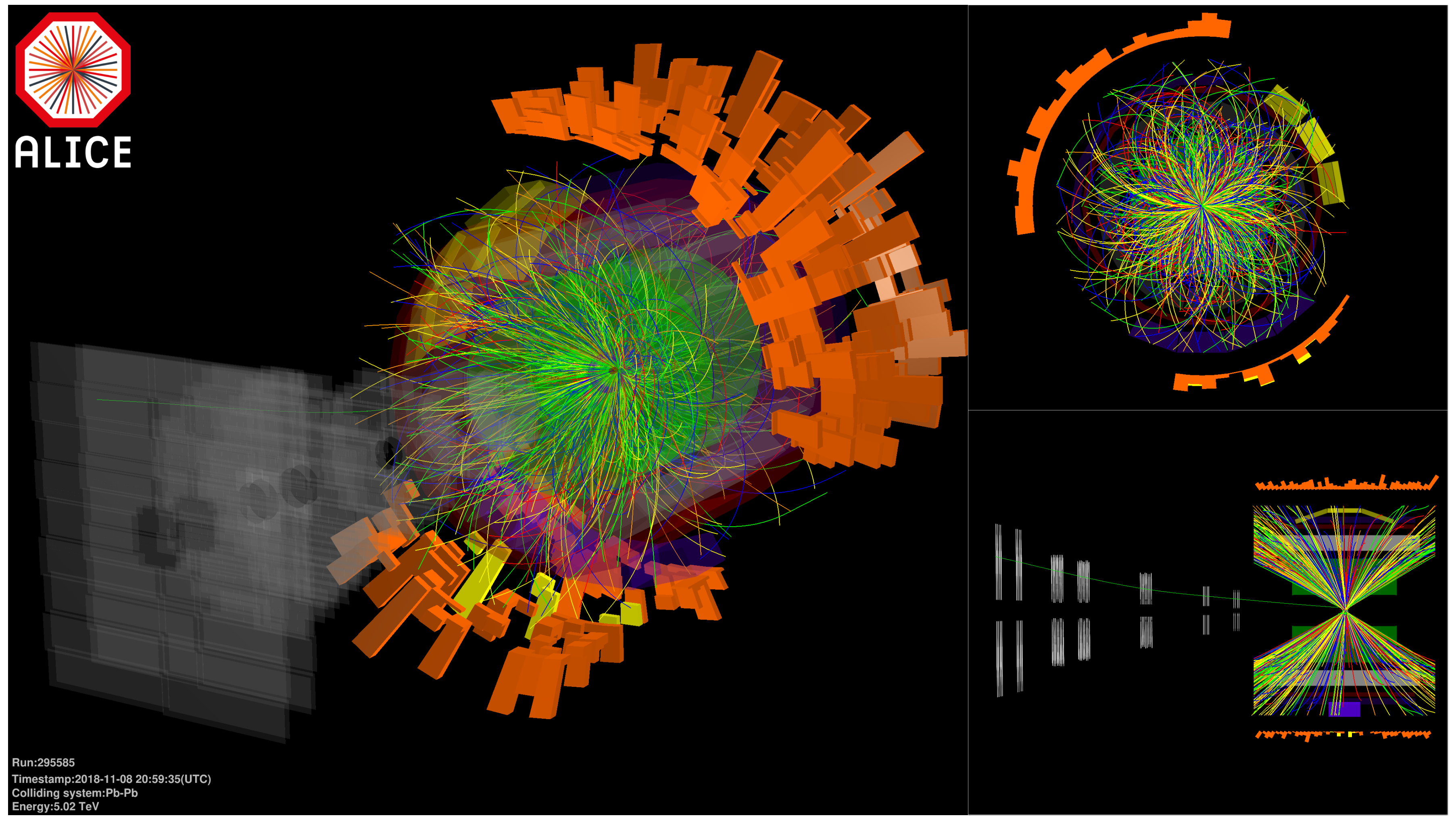
ALICE Detector

Heavy-ion collisions

Large Hadron Collider, CERN, Switzerland





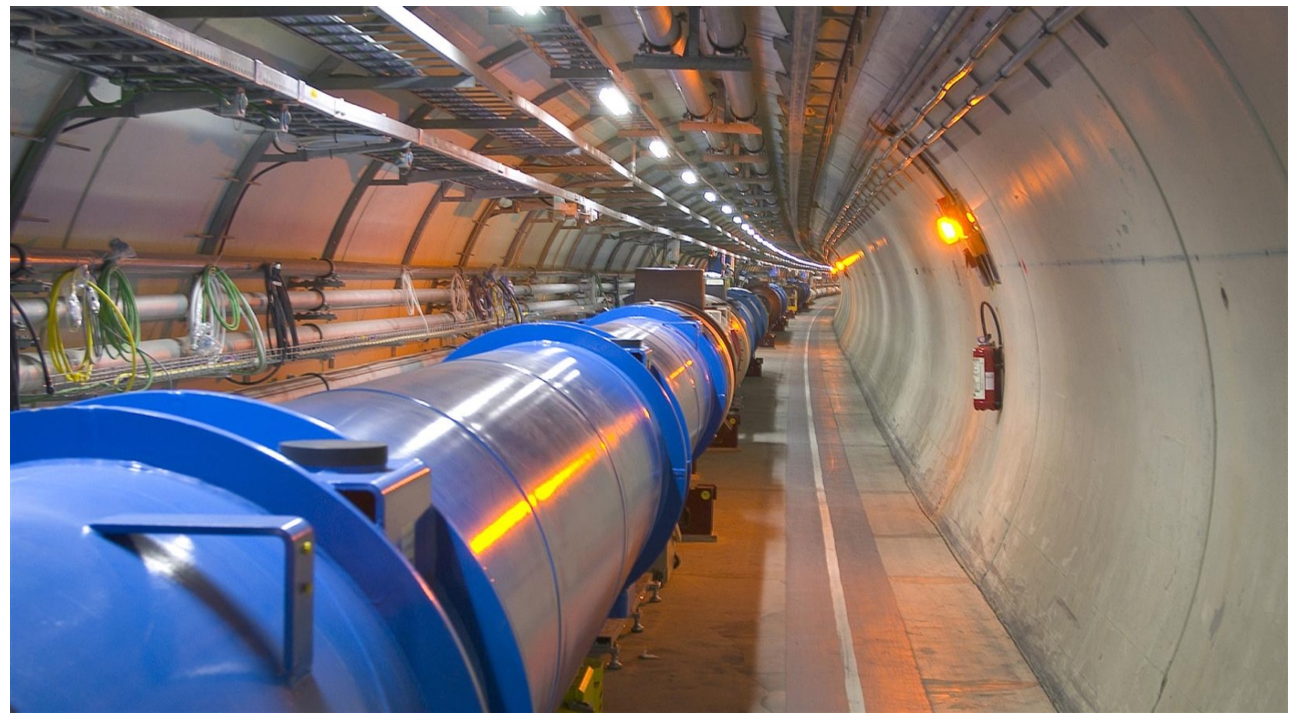


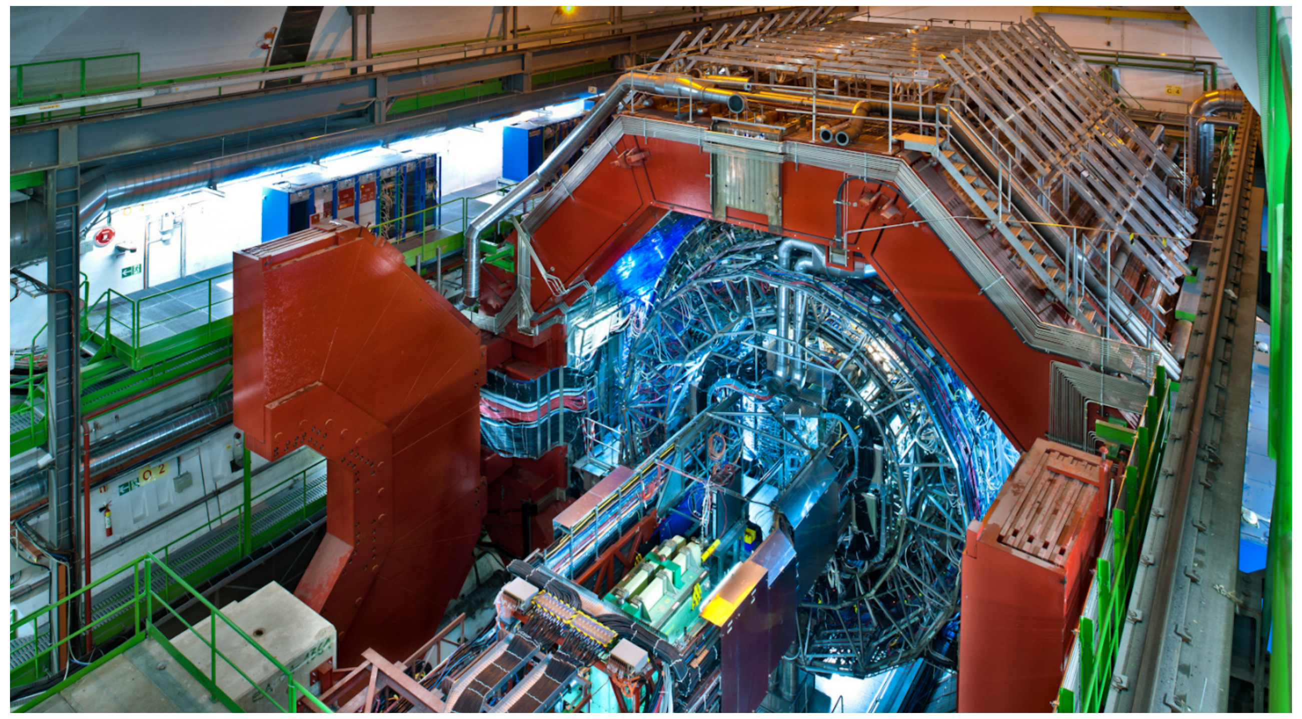
https://cds.cern.ch/record/2108293

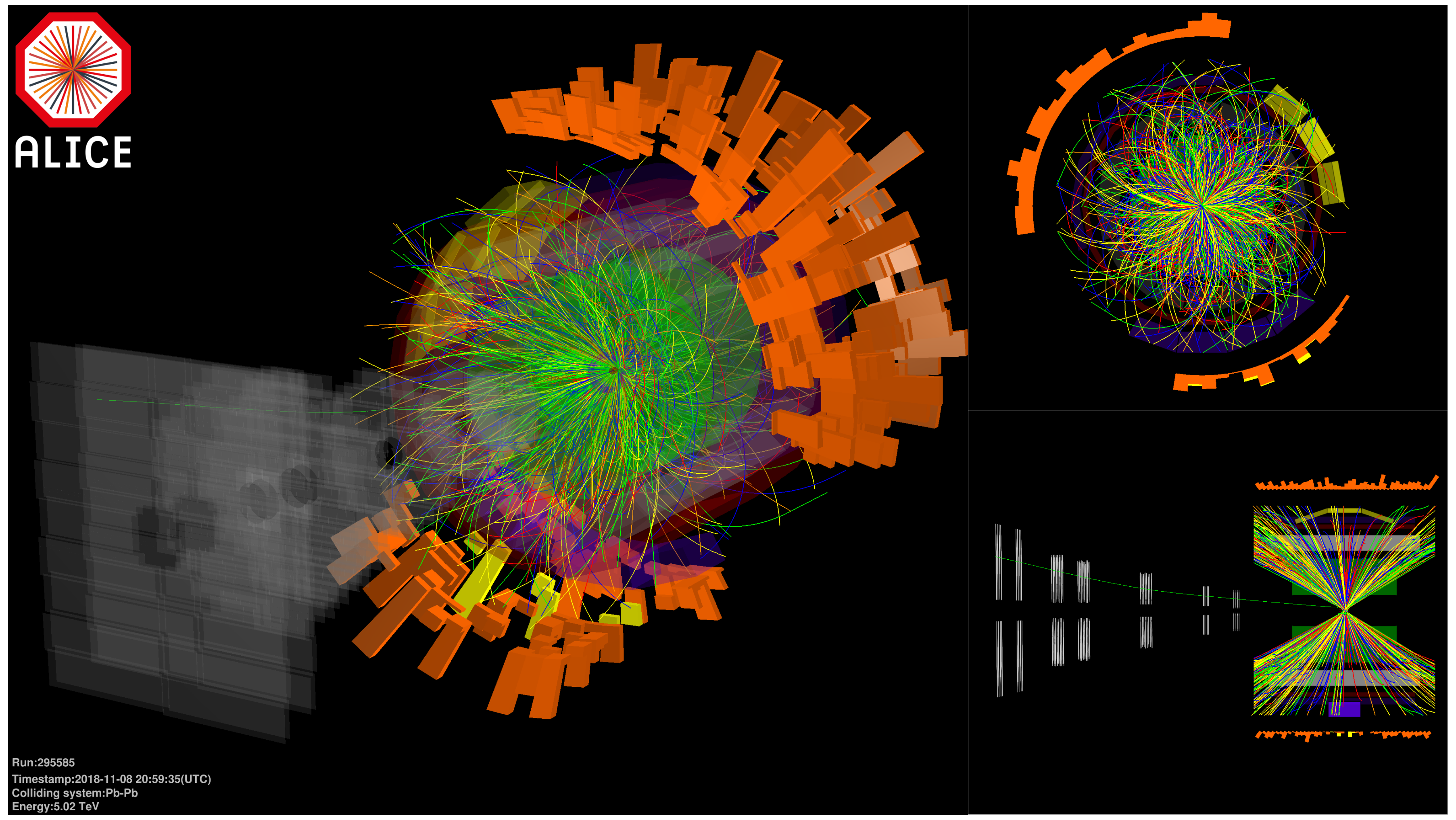
ALICE Detector

Heavy-ion collisions

Large Hadron Collider, CERN, Switzerland





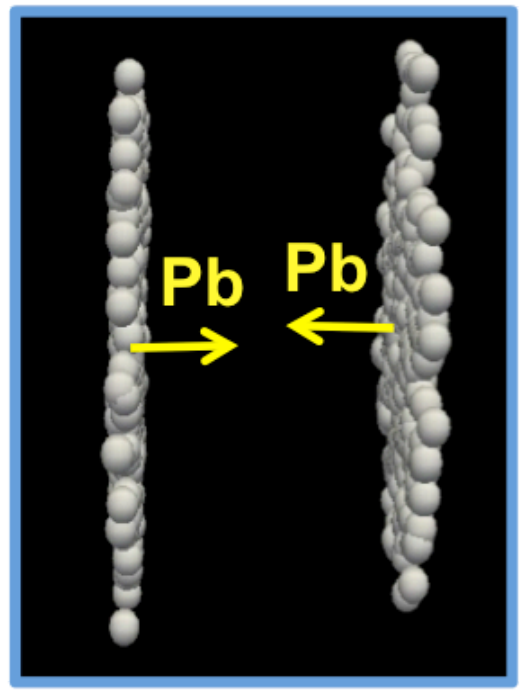


https://cds.cern.ch/record/2108293

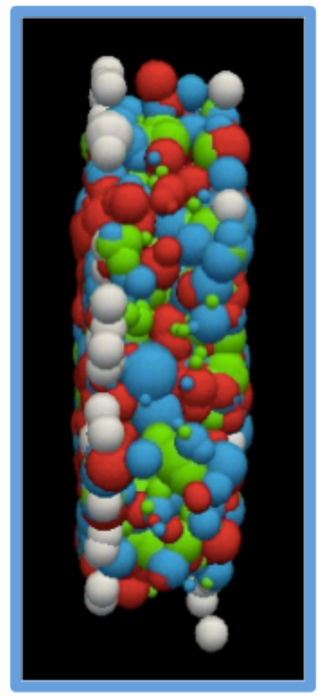
ALICE Detector

time

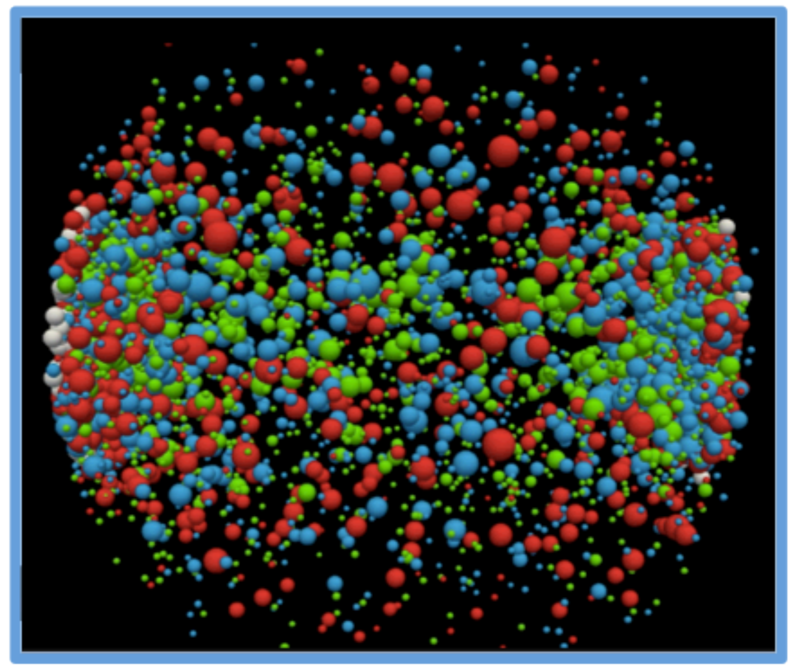
Pre-reaction



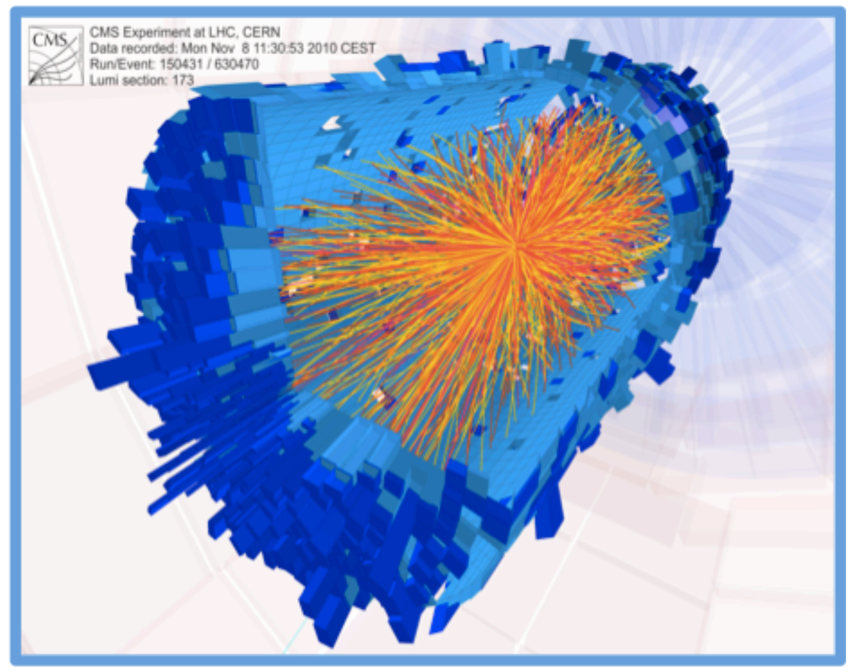
QGP



Hadronization

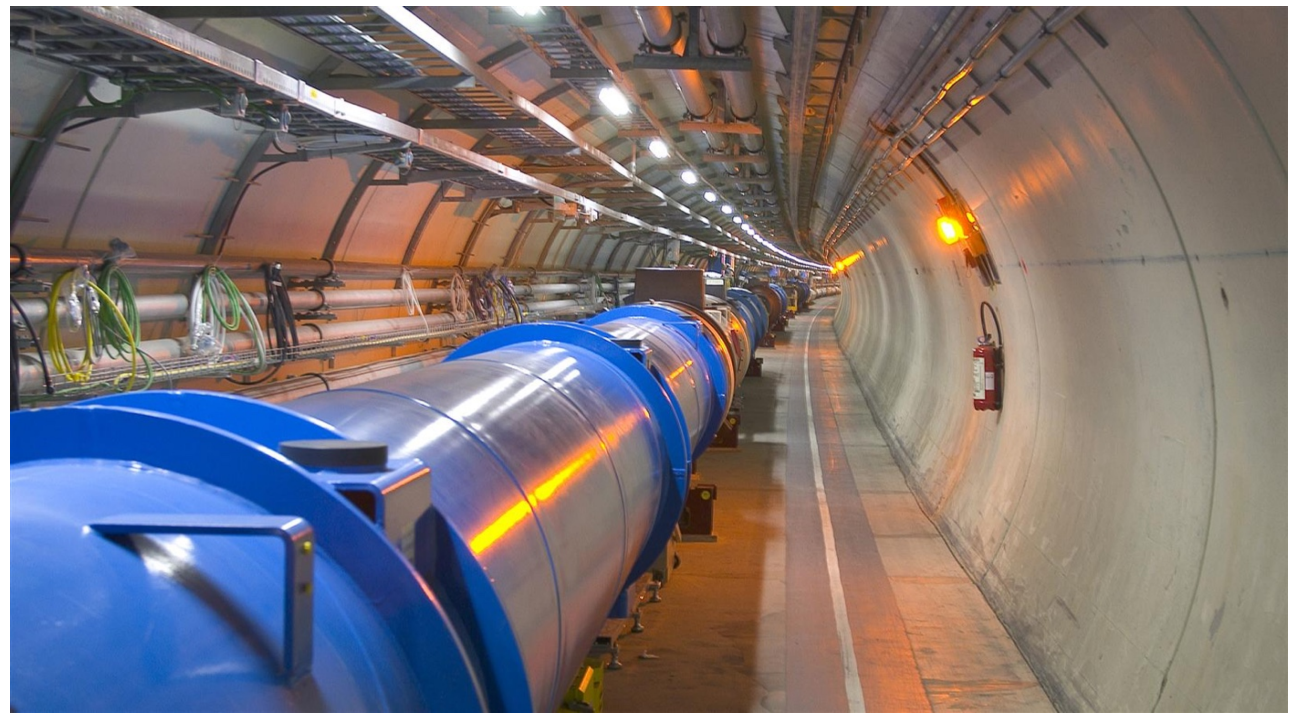


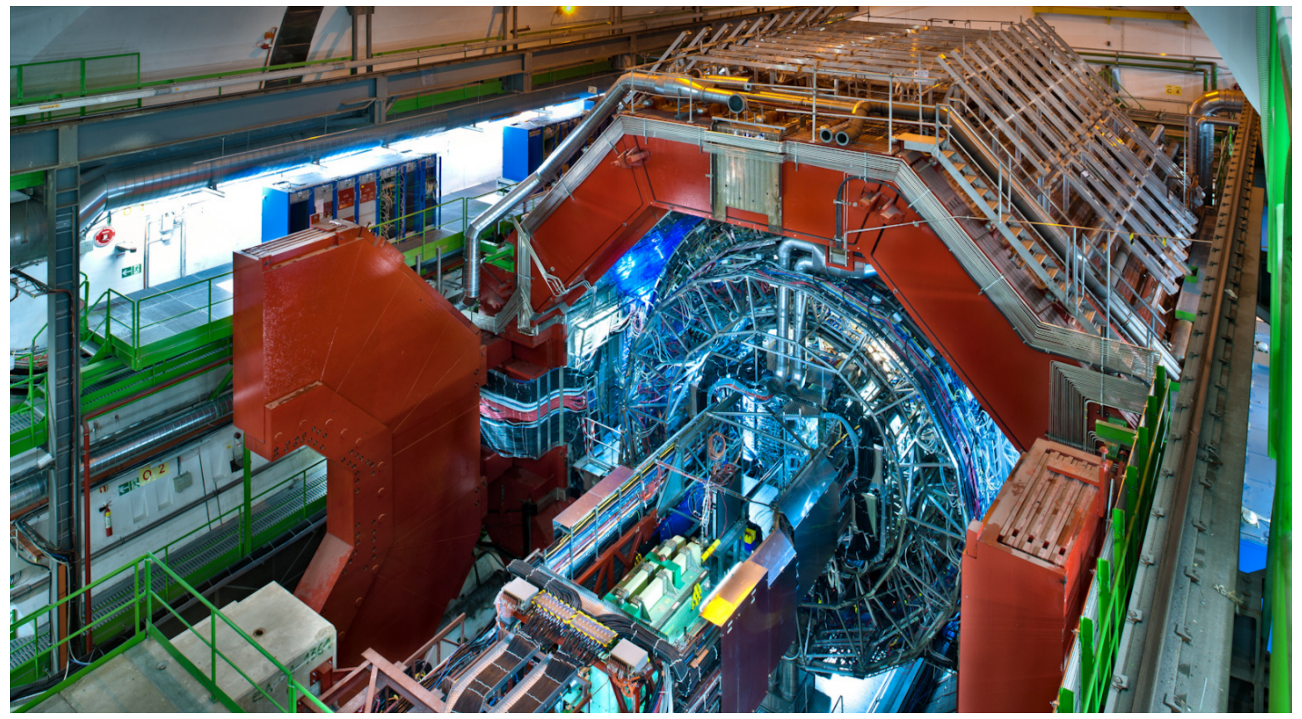
Detection

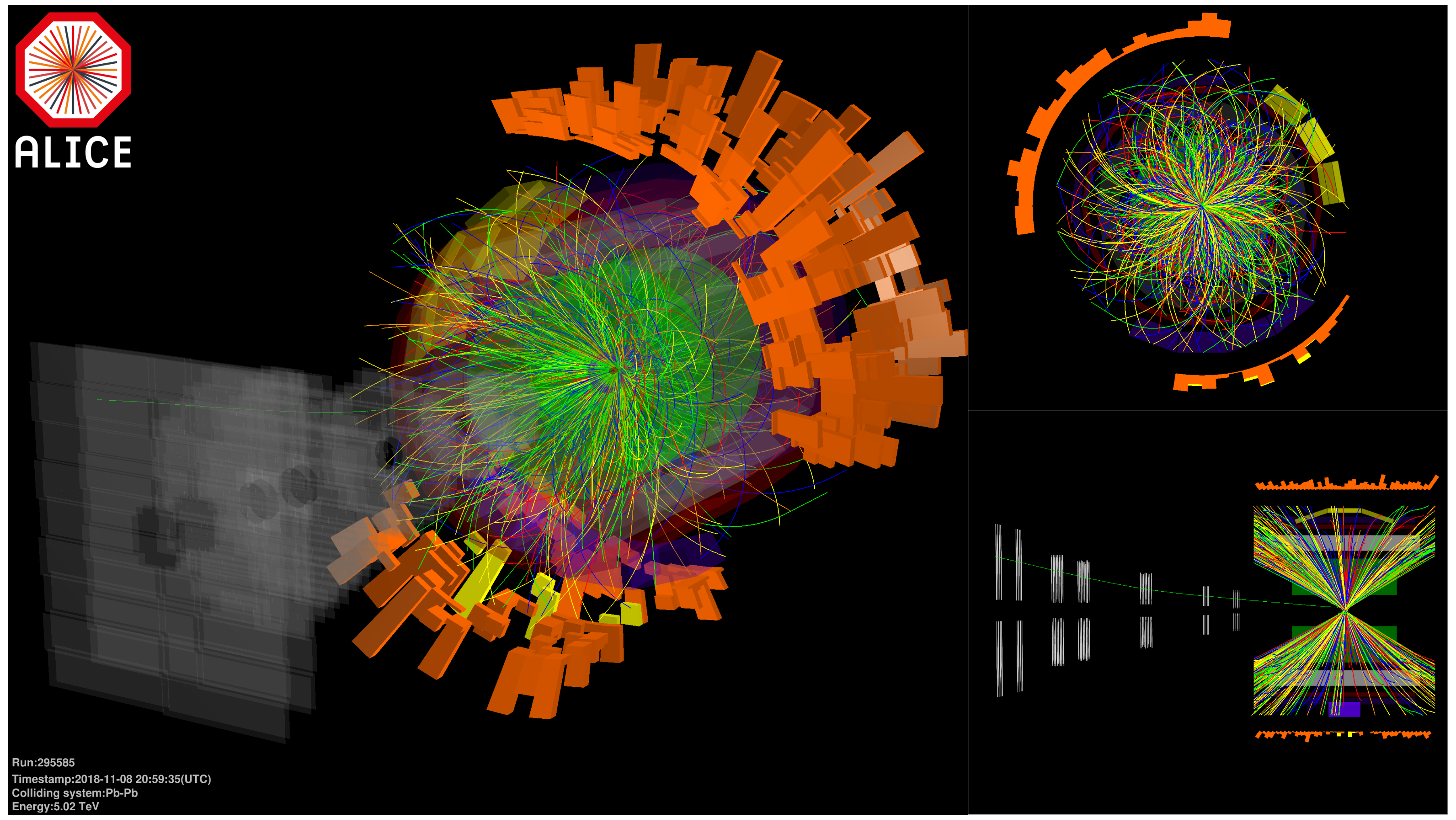


Heavy-ion collisions

Large Hadron Collider, CERN, Switzerland





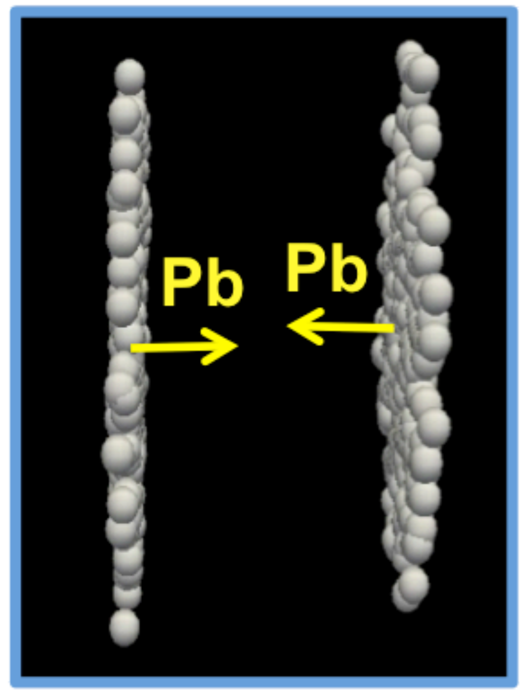


https://cds.cern.ch/record/2108293

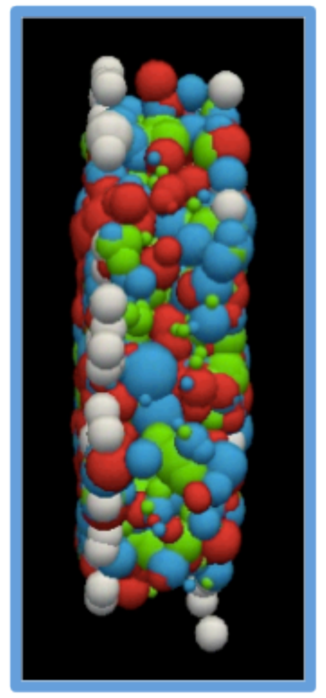
ALICE Detector

time

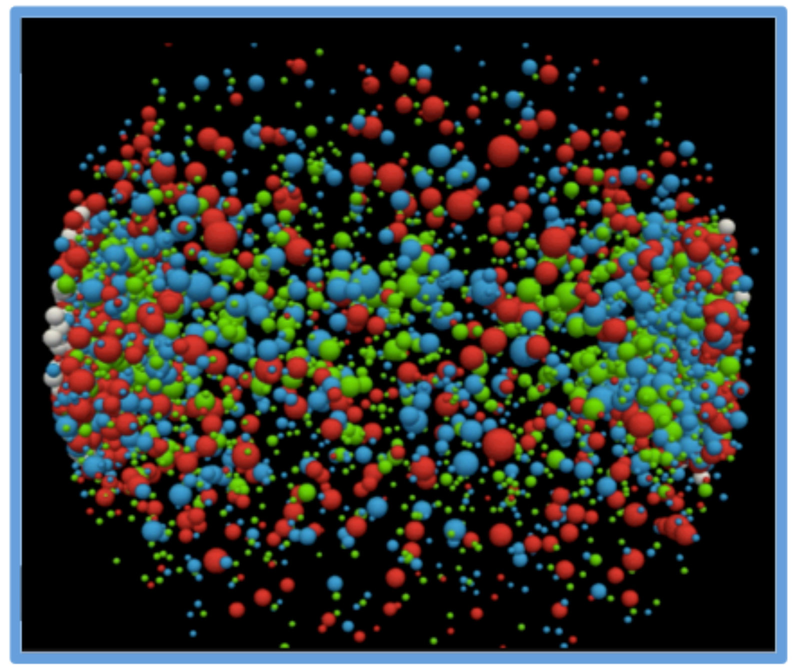
Pre-reaction



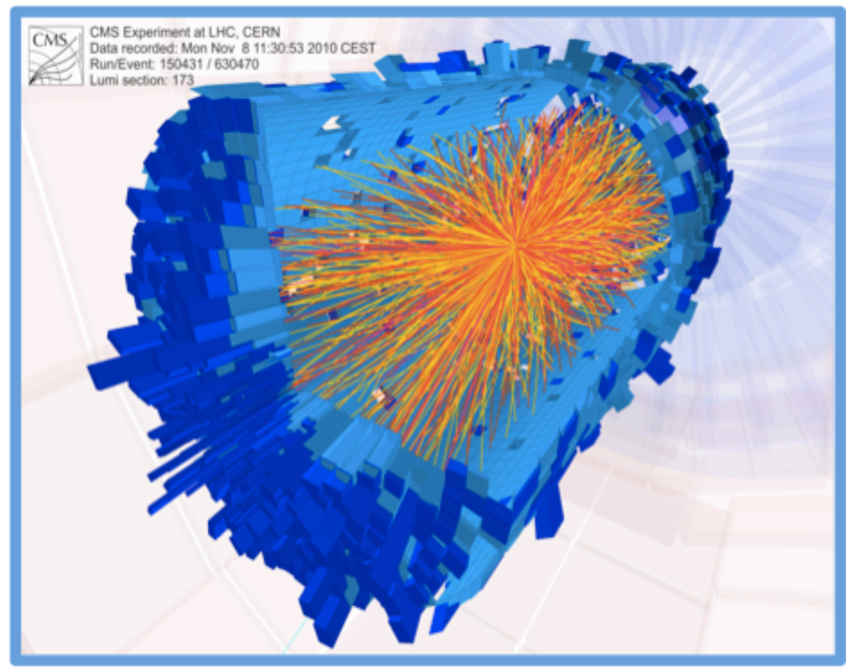
QGP



Hadronization



Detection



- Lorentz contracted discs
- Energy deposition: 0.5 GeV/fm^3
- Quark-gluon plasma

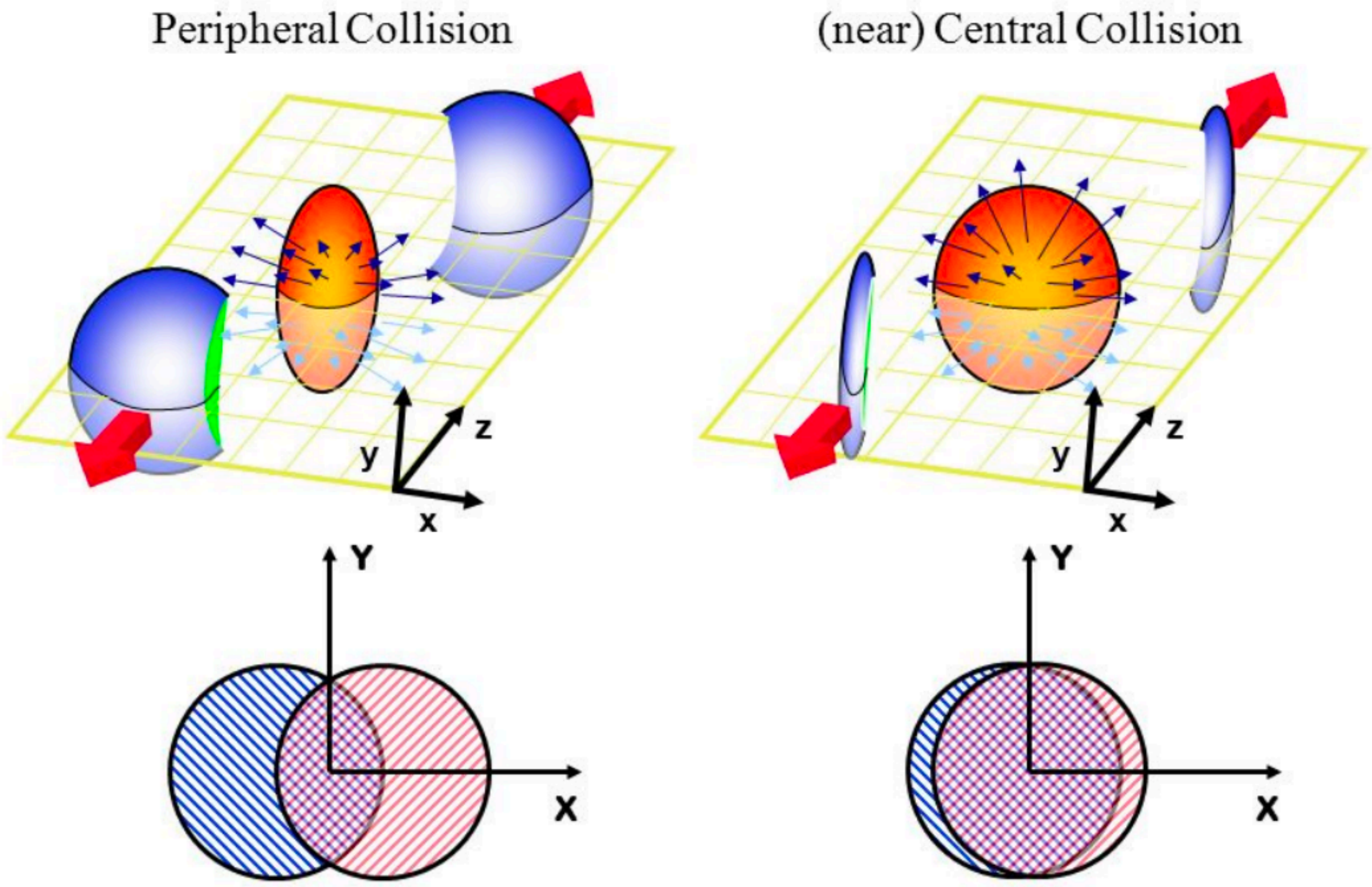
Impact parameter (b)

Impact parameter (b)

- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution

Impact parameter (b)

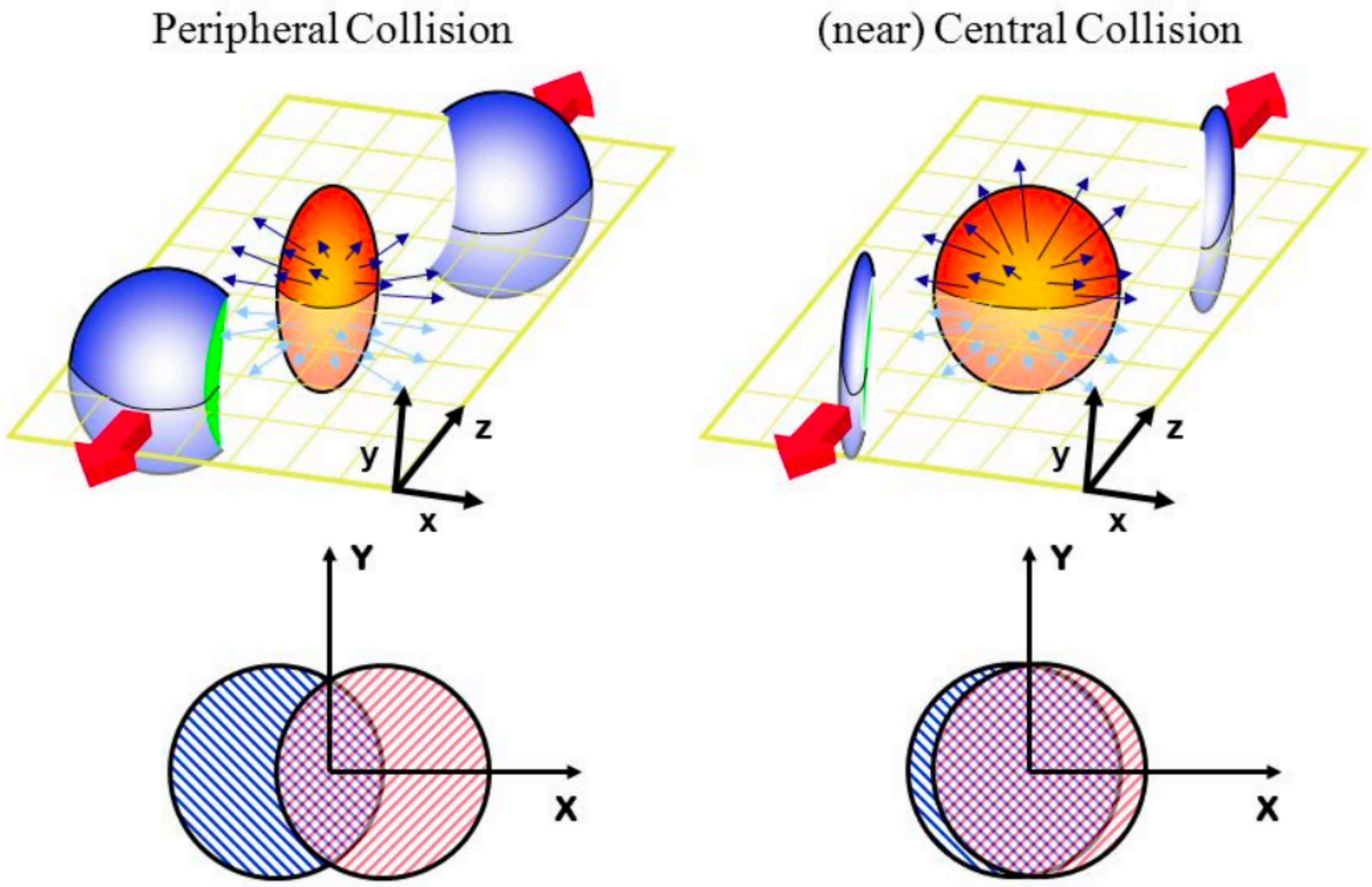
- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution



$0 \leq b \leq 2R$

Impact parameter (b)

- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution

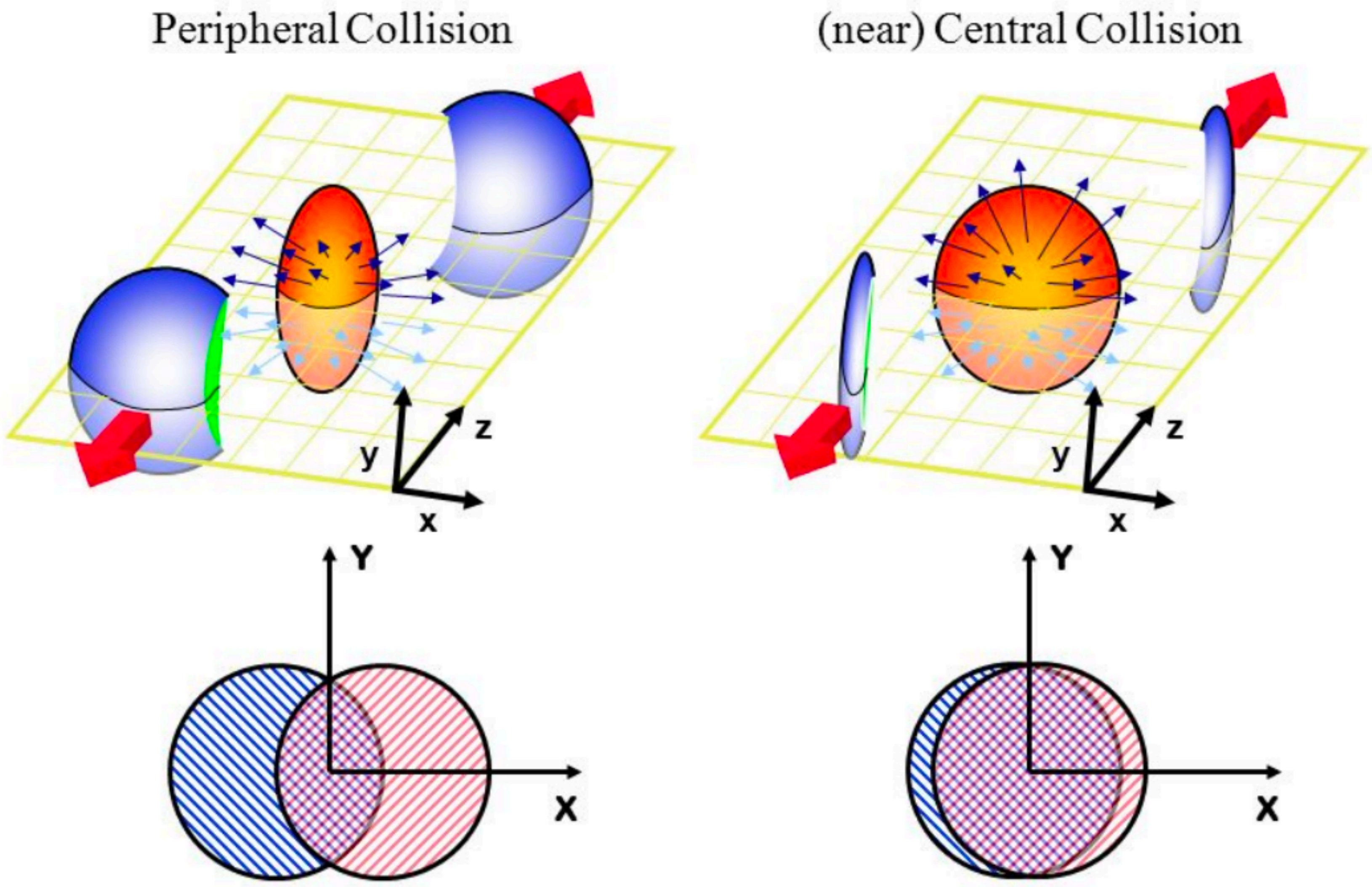


$0 \leq b \leq 2R$

Transverse Spherocity (S_0)

Impact parameter (b)

- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution



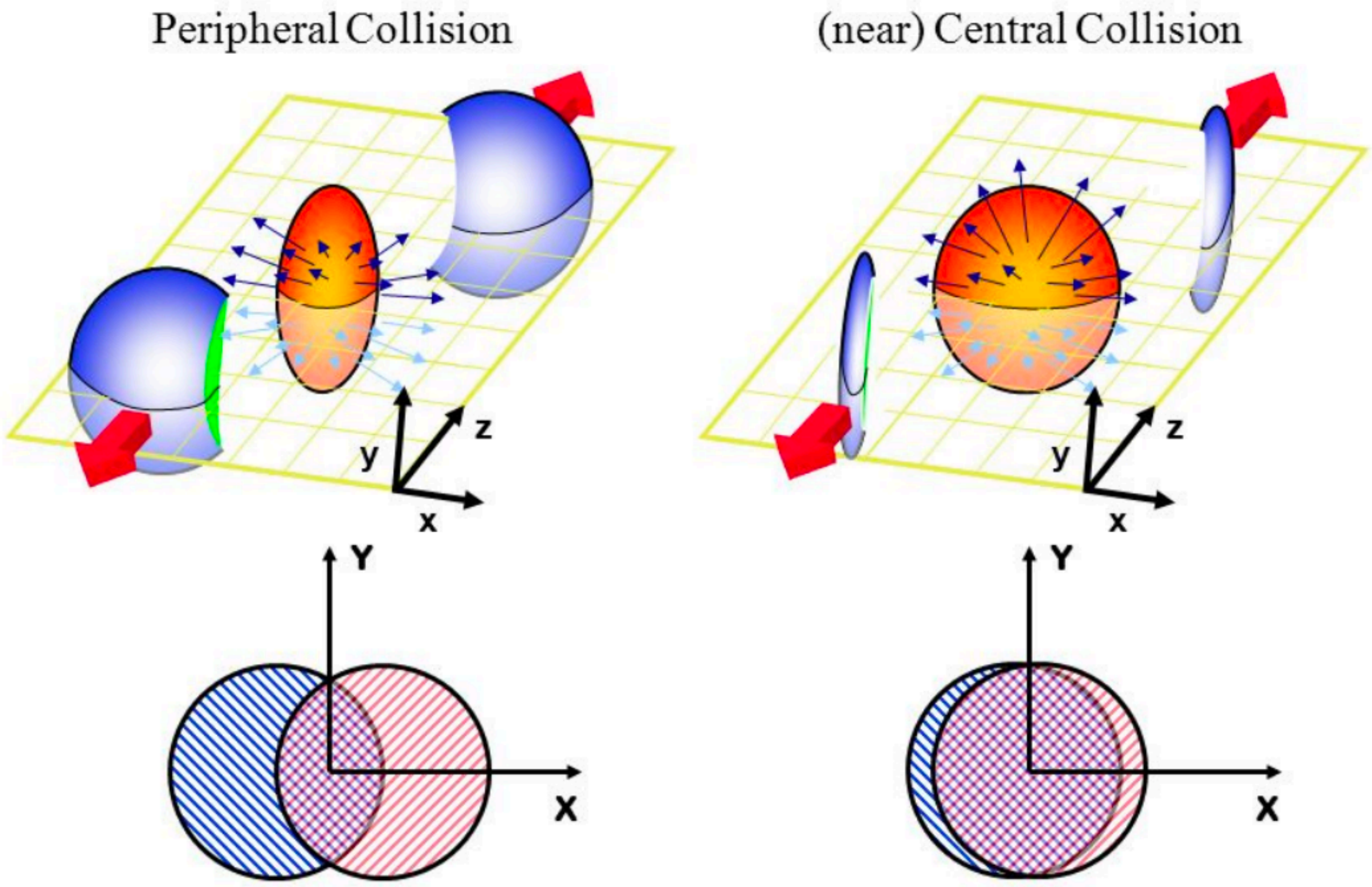
$0 \leq b \leq 2R$

Transverse Spherocity (S_0)

- Transverse Spherocity distinguishes hard and soft processes
- In pp collisions,
  1. Jetty: Back-to-back structure, indication of hard-QCD
  2. Isotropic: soft-QCD process
- Dominance of isotropic events in high multiplicity pp collisions
- $\langle p_T \rangle$ is higher for jetty events
- S_0 has multiplicity and centrality dependence

Impact parameter (b)

- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution



$0 \le b \le 2R$

Transverse Spherocity (S_0)

- Transverse Spherocity distinguishes hard and soft processes
- In pp collisions,
  1. Jetty: Back-to-back structure, indication of hard-QCD
  2. Isotropic: soft-QCD process
- Dominance of isotropic events in high multiplicity pp collisions
- $\langle p_T \rangle$ is higher for jetty events
- S_0 has multiplicity and centrality dependence

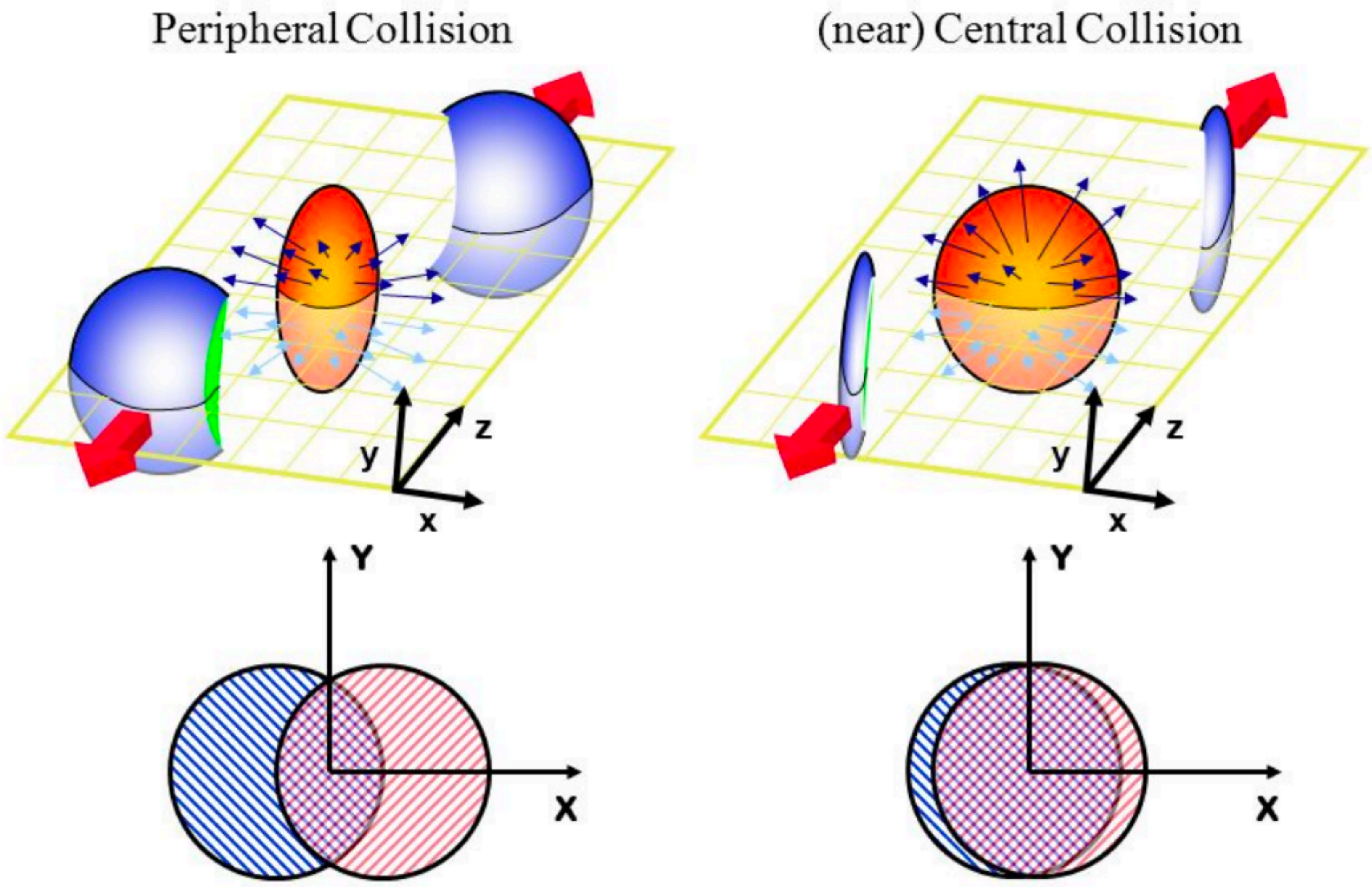
$$S_0 = \frac{\pi^2}{4} \times \min_{\hat{n} = (n_x, n_y, 0)} \left(\frac{\Sigma_i \, |\vec{p}_{T_i} \times \hat{n}|}{\Sigma_i \, \vec{p}_{T_i}} \right)^2$$

$$p_{\rm T} = \sqrt{p_x^2 + p_y^2}$$

A. Khuntia et al., J. Phys. G48, 035102

Impact parameter (b)

- Transverse distance between the centres of colliding nuclei
- Initial geometry affects the final state particle production
- Order of a few fermi (10^{-15}m)
- Impossible to estimate from experiments
- Could be inferred from charged particle multiplicity distribution



$0 \leq b \leq 2R$

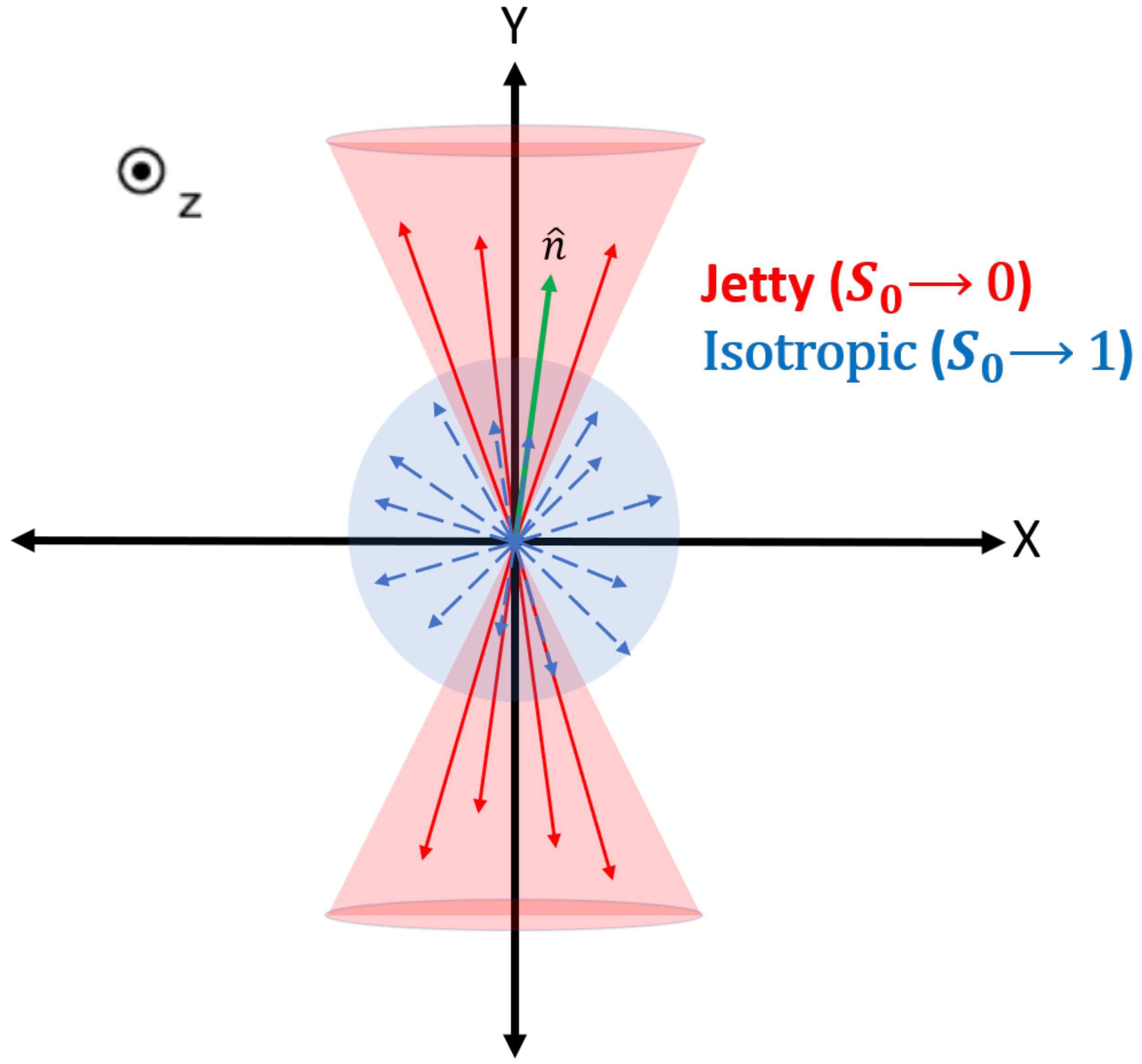
Transverse Spherocity (S_0)

- Transverse Spherocity distinguishes hard and soft processes
- In pp collisions,
  1. Jetty: Back-to-back structure, indication of hard-QCD
  2. Isotropic: soft-QCD process
- Dominance of isotropic events in high multiplicity pp collisions
- $\langle p_T \rangle$ is higher for jetty events
- S_0 has multiplicity and centrality dependence

$$S_0 = \frac{\pi^2}{4} \times \min_{\hat{n}=(n_x,n_y,0)} \left(\frac{\Sigma_i \, |\vec{p}_{T_i} \times \hat{n}|}{\Sigma_i \, \vec{p}_{T_i}} \right)^2$$

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$$

A. Khuntia et al., J. Phys. G48, 035102



Schematic picture showing possible jetty and isotropic event formations

Estimation of Impact parameter and spherocity

N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Estimation of Impact parameter and spherocity

- Pearsons correlation coefficient

$$\rho = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y}$$

- Defines the degree of correlation
- Input Variables: $\langle dN_{ch}/d\eta \rangle$, $\langle N_{ch}^{TS} \rangle$ and $\langle p_T \rangle$
  Output variable: b and S_0
- Good correlation is seen among chosen input and output variables
- The algorithm tries to understand the correlation and exploit the features to arrive on a conclusion (a number)

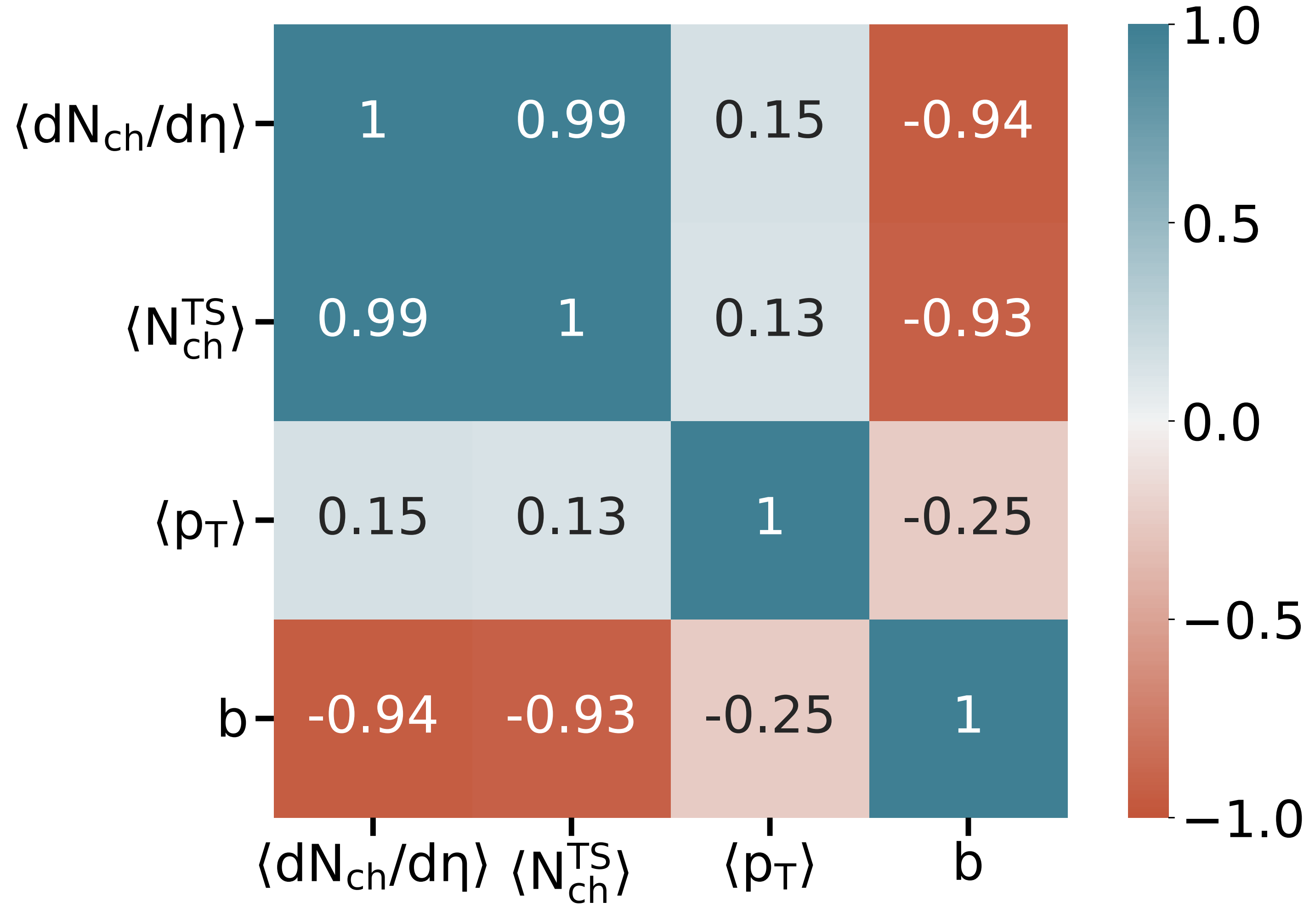
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Estimation of Impact parameter and spherocity

- Pearsons correlation coefficient

$$\rho = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y}$$

- Defines the degree of correlation
- Input Variables: $\langle dN_{ch}/d\eta \rangle$, $\langle N_{ch}^{TS} \rangle$ and $\langle p_T \rangle$
  Output variable: b and S_0
- Good correlation is seen among chosen input and output variables
- The algorithm tries to understand the correlation and exploit the features to arrive on a conclusion (a number)



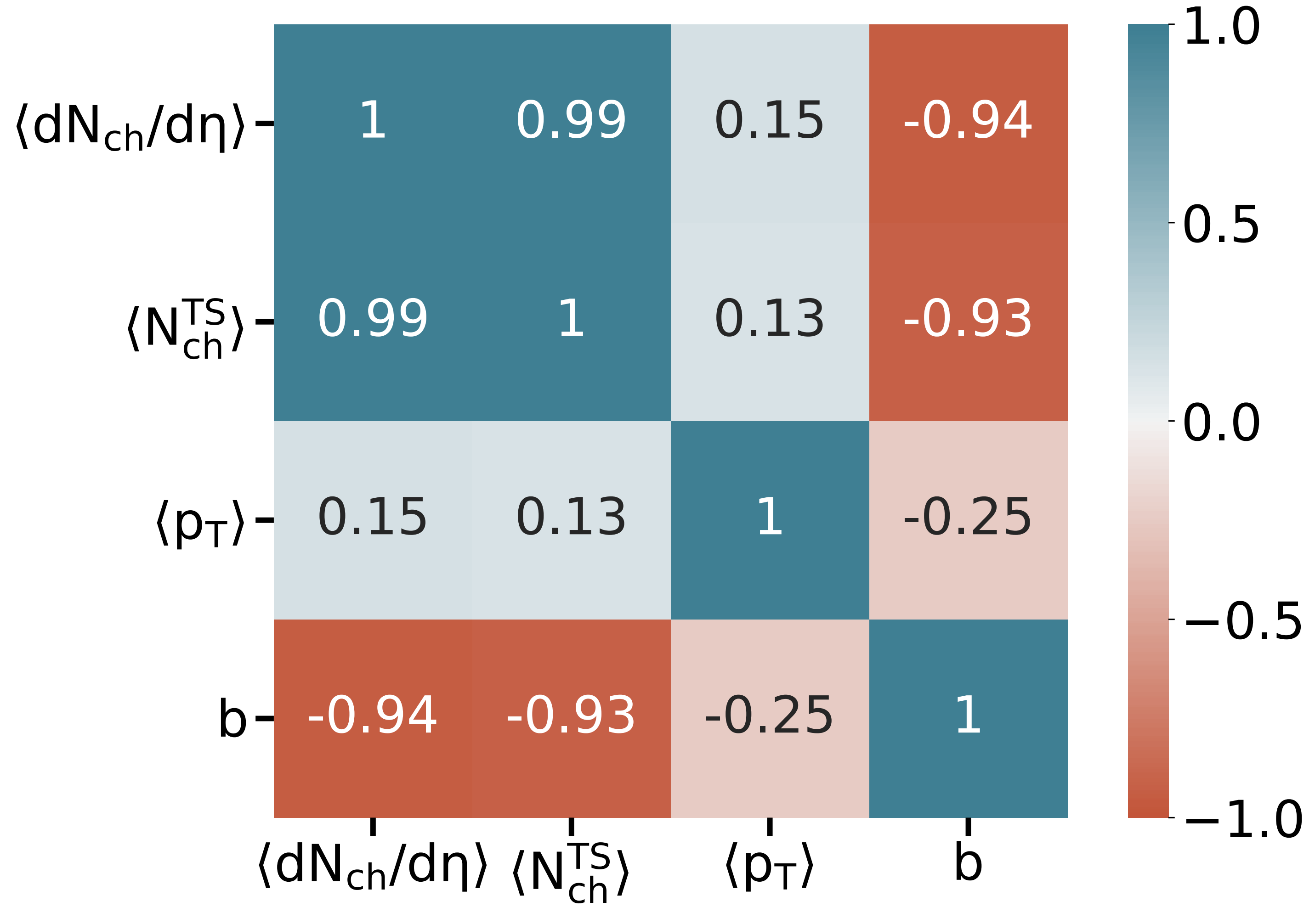
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

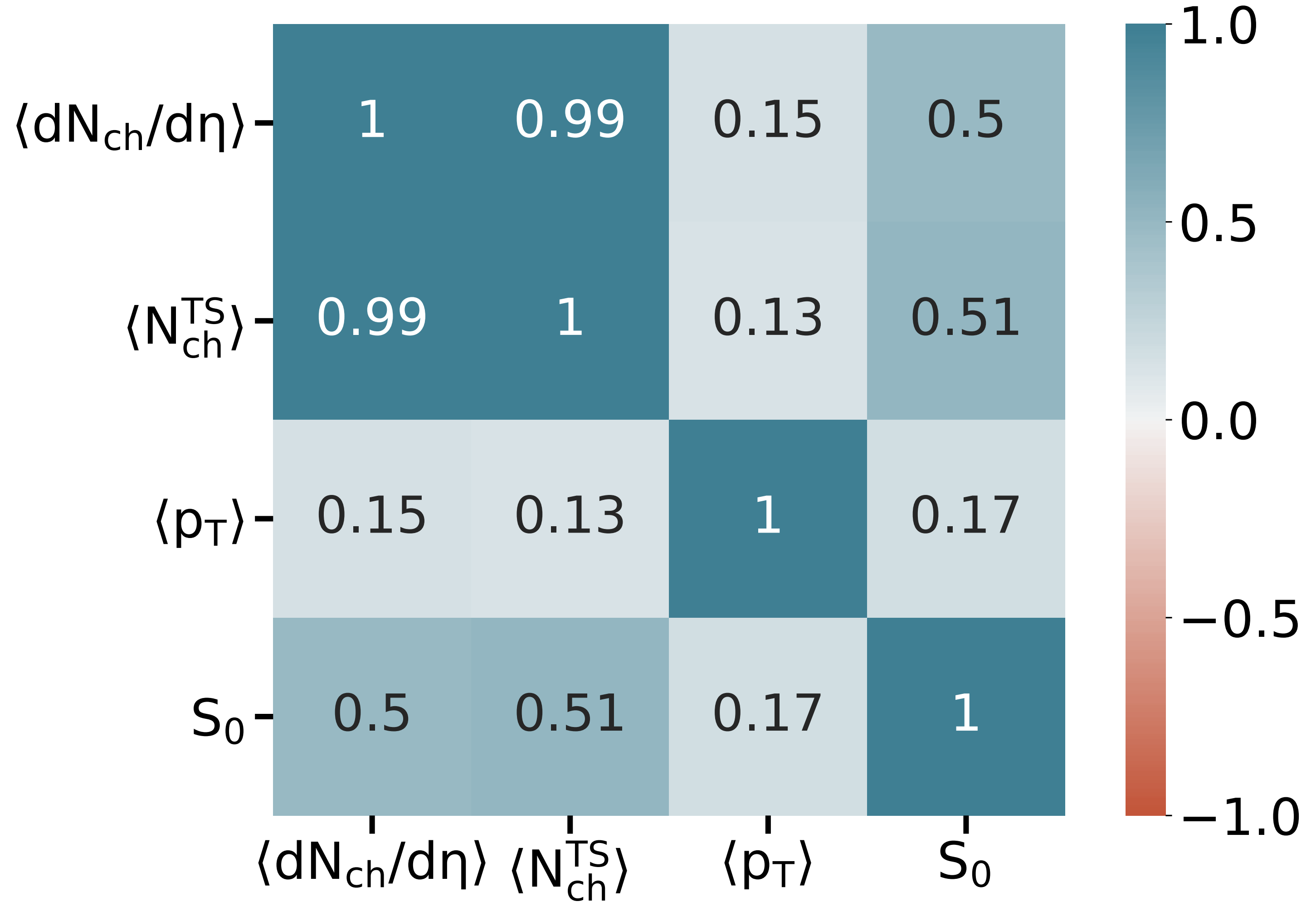
Estimation of Impact parameter and spherocity

- Pearsons correlation coefficient

$$\rho = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y}$$

- Defines the degree of correlation
- Input Variables: $\langle dN_{ch}/d\eta \rangle$, $\langle N_{ch}^{TS} \rangle$ and $\langle p_T \rangle$
  Output variable: b and S_0
- Good correlation is seen among chosen input and output variables
- The algorithm tries to understand the correlation and exploit the features to arrive on a conclusion (a number)





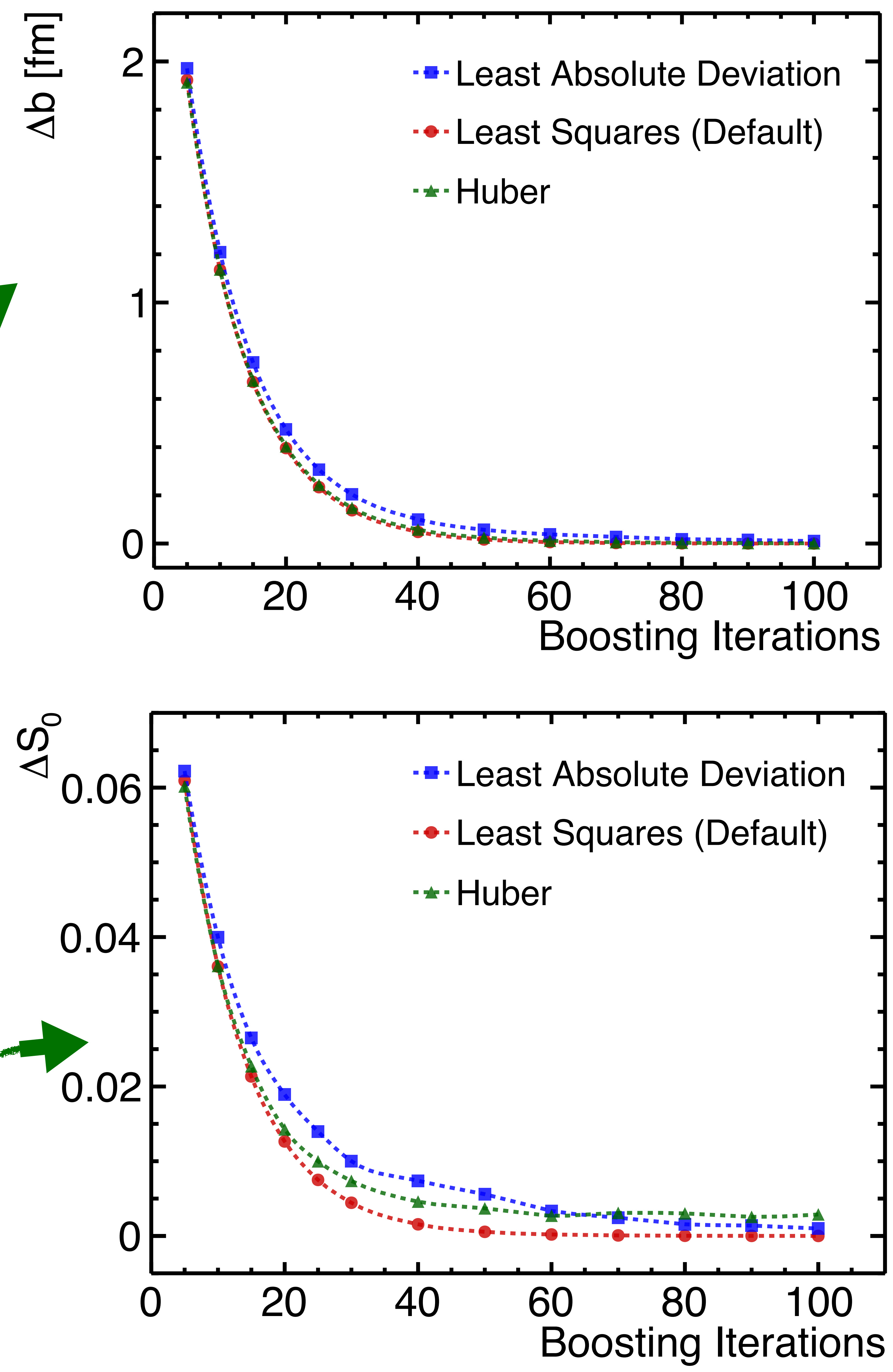
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Parameters in the GBDT-ML Model

- Loss Function: Least Square Loss
- Small learning rate = 0.1
- Number of trees = 100
- Training Size: 60,000 events (min. bias)

$$\Delta S_0 = \frac{1}{N_{\text{events}}} \sum_{n=1}^{N_{\text{events}}} | S_{0_n}^{true} - S_{0_n}^{pred.} |$$

Size of training data	2K	10K	20K	40K	50K	60K
Δb [fm] (Impact parameter)	0.71	0.62	0.58	0.53	0.52	0.52
ΔS_0 (Spherocity)	0.079	0.068	0.062	0.058	0.056	0.055



J. H. Friedman, Ann. Stat. 29, 1189 (2001).

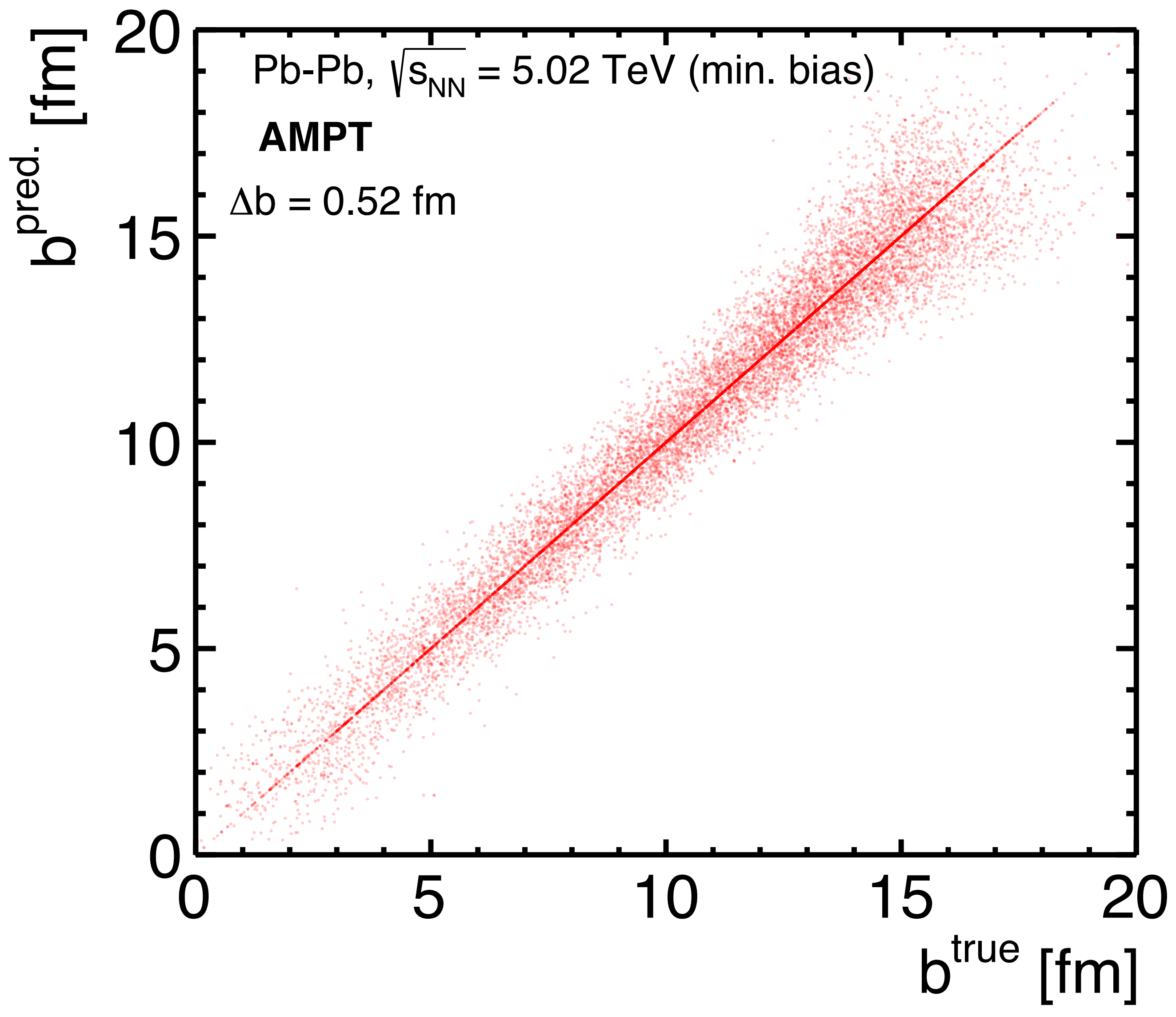
L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees (Wadsworth & Brooks/ Cole Advanced Books & Software, Monterey, CA, 1984), p. 358, https://doi.org/10.1002/cyto.990080516.

N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Results

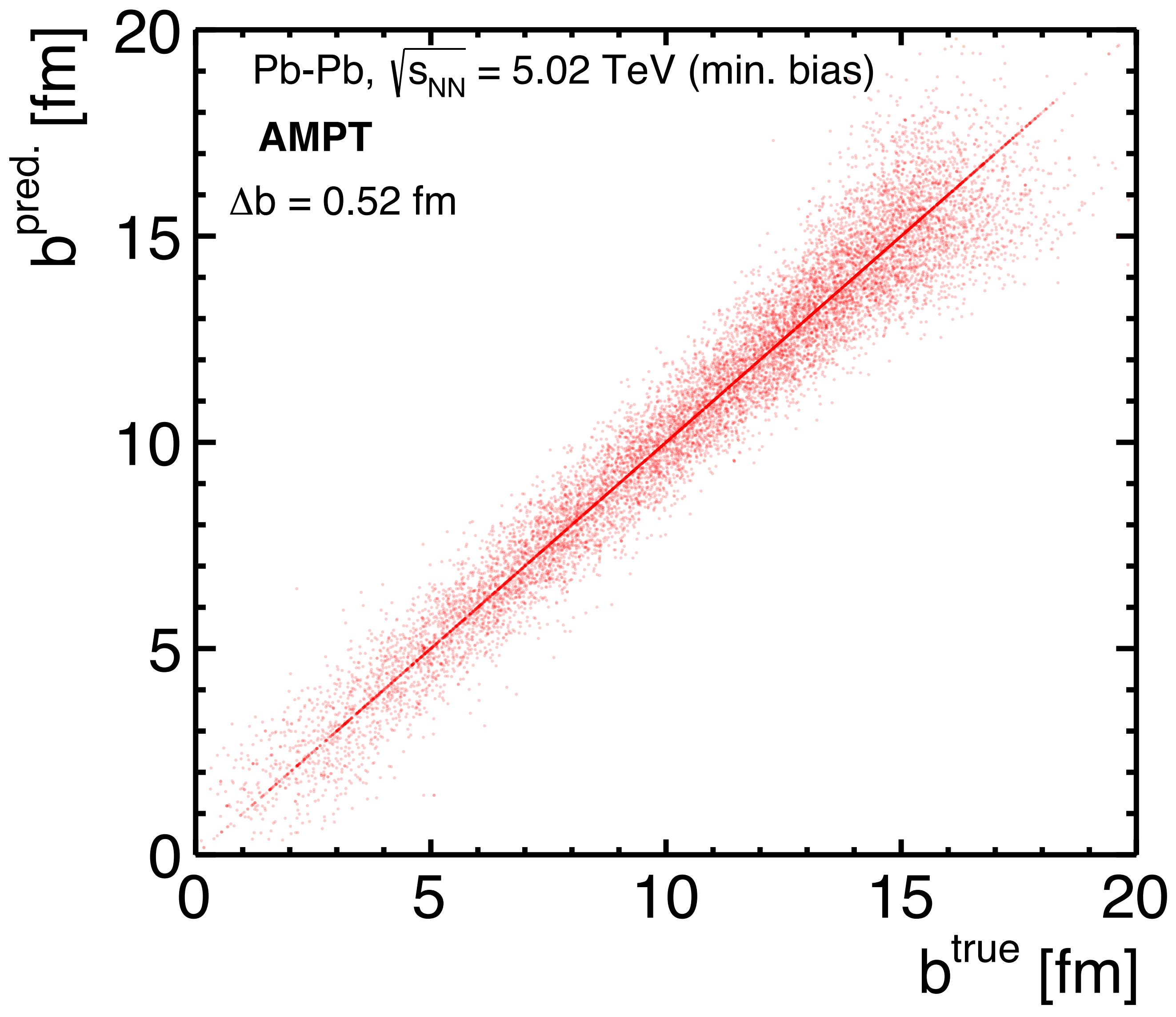
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

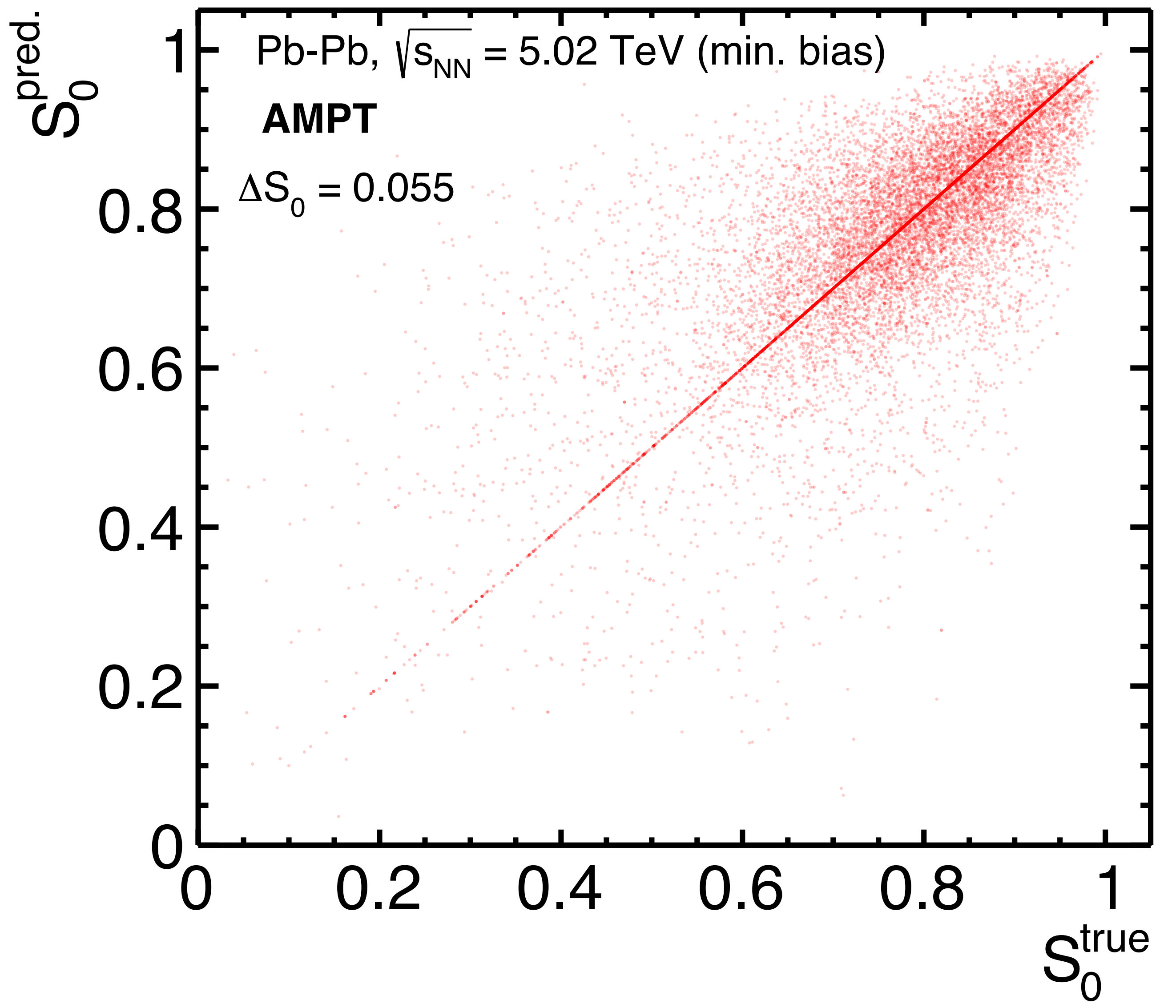
Results



N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

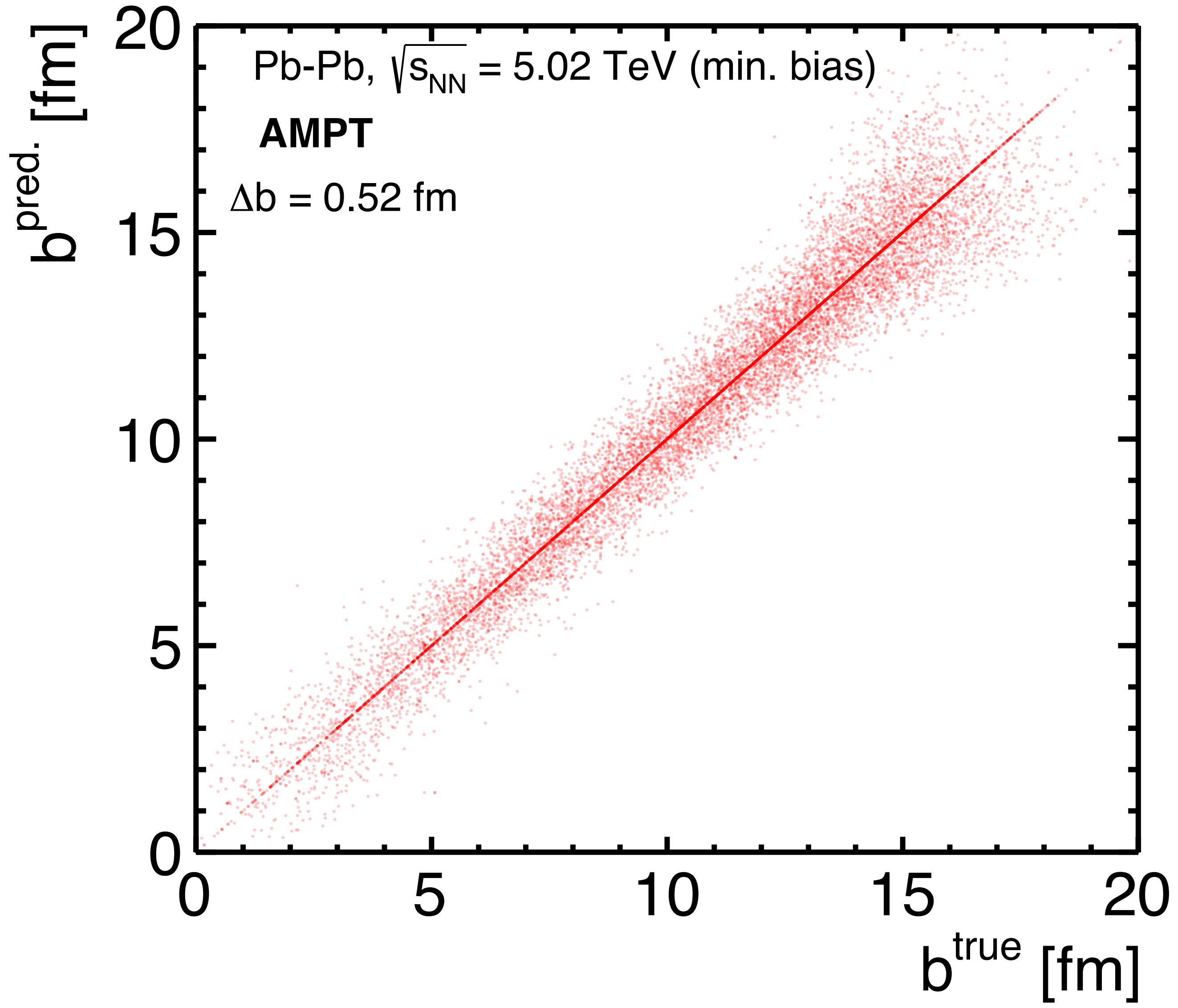
Results

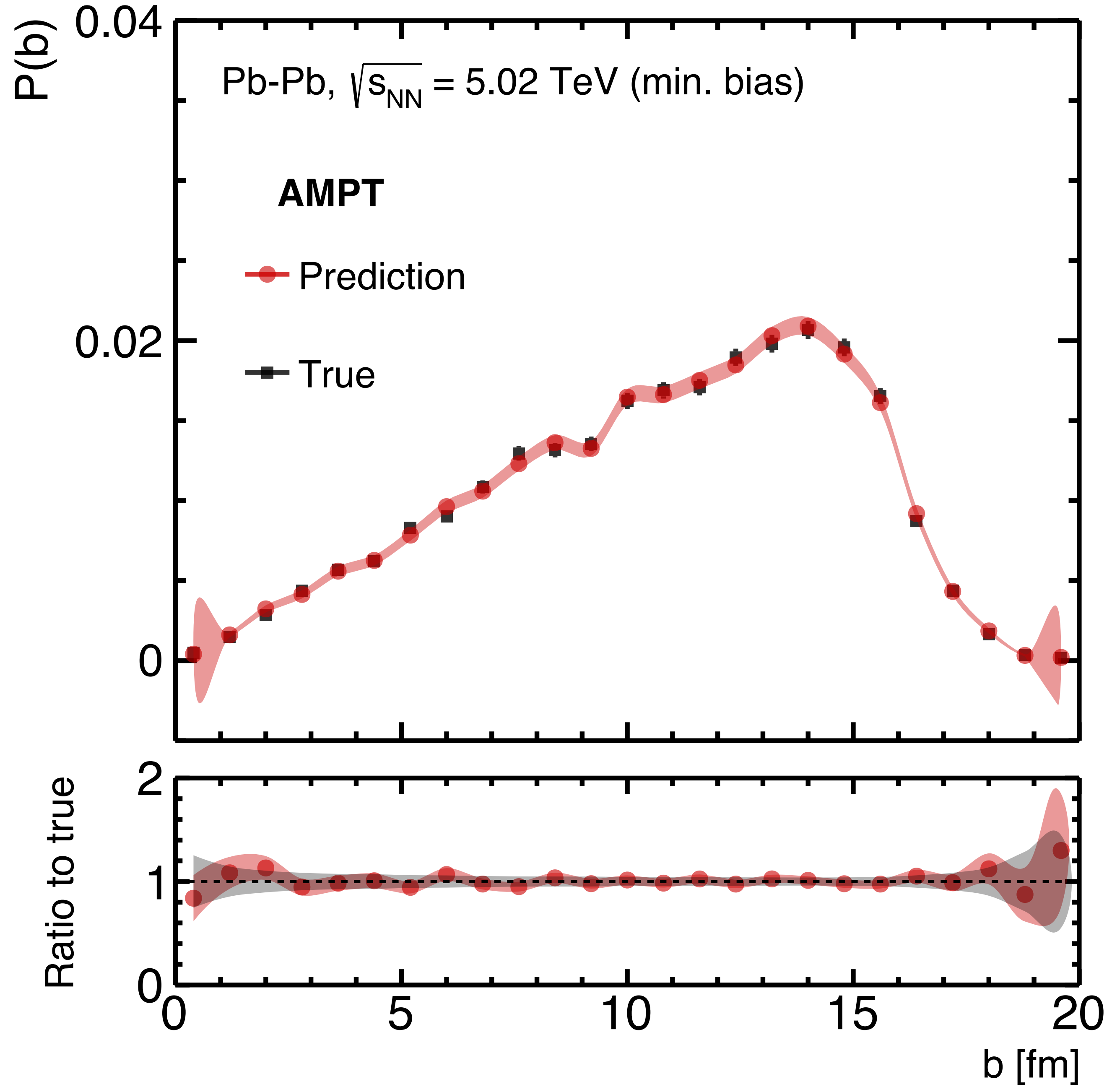


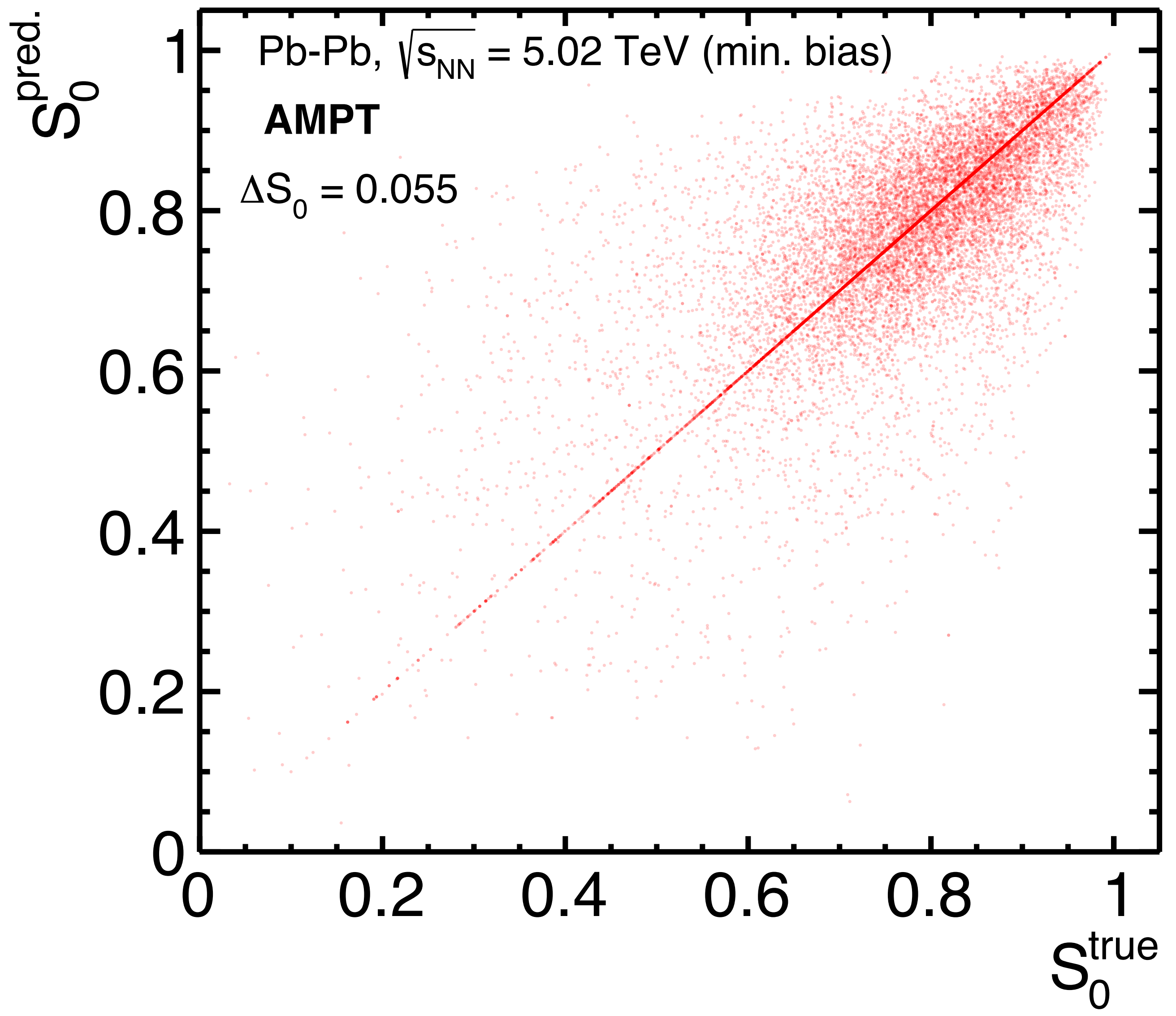


N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Results

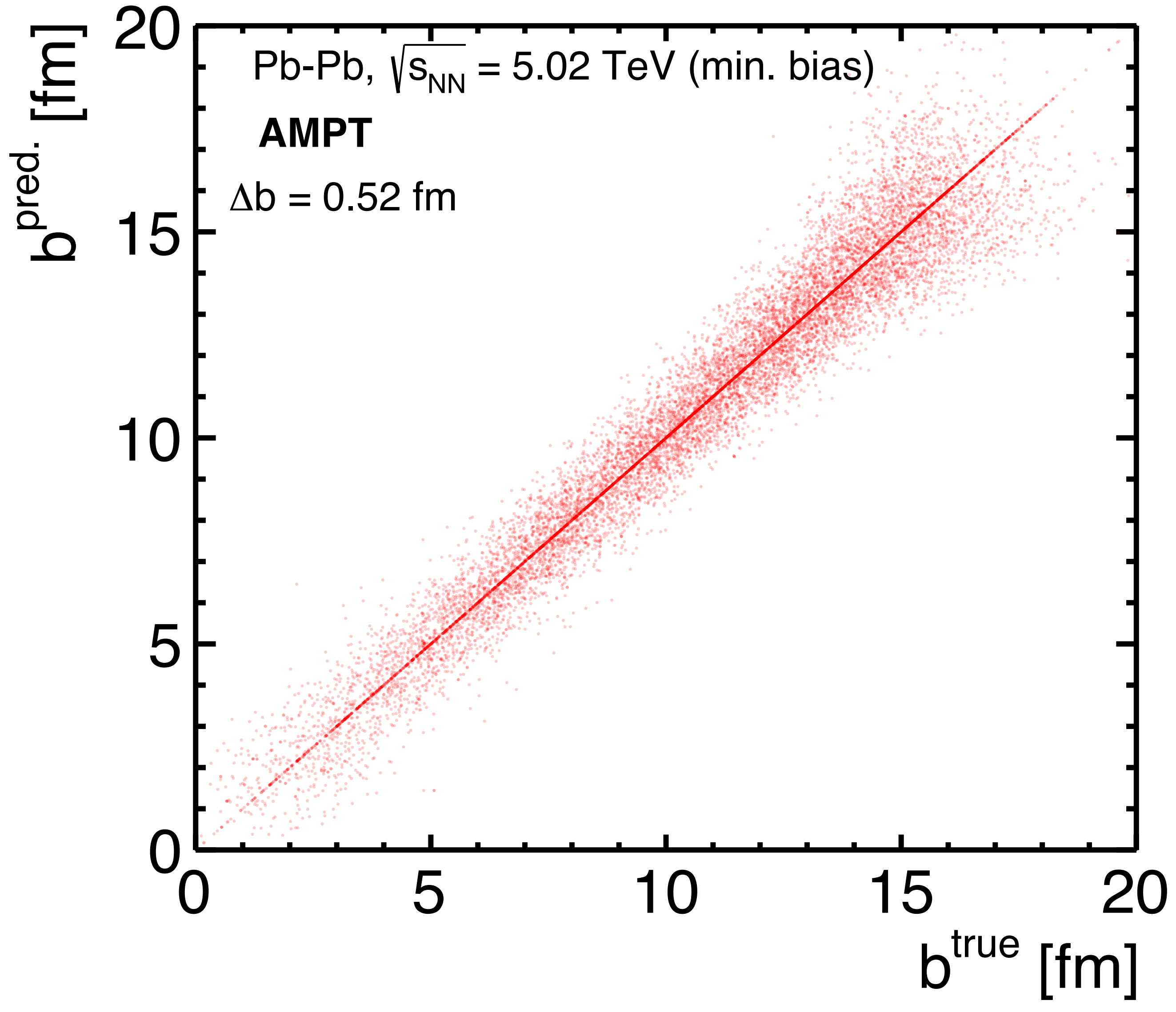


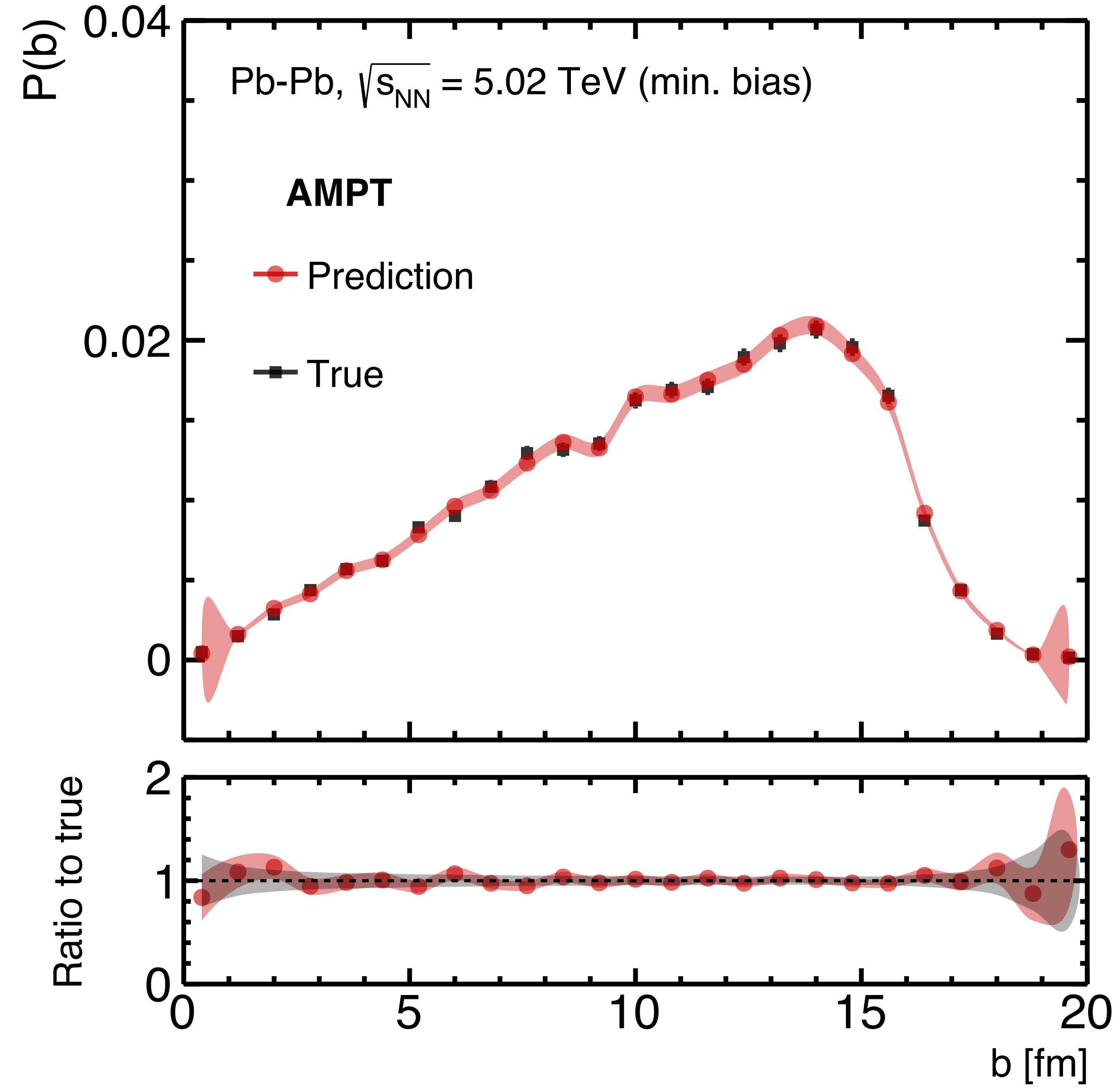


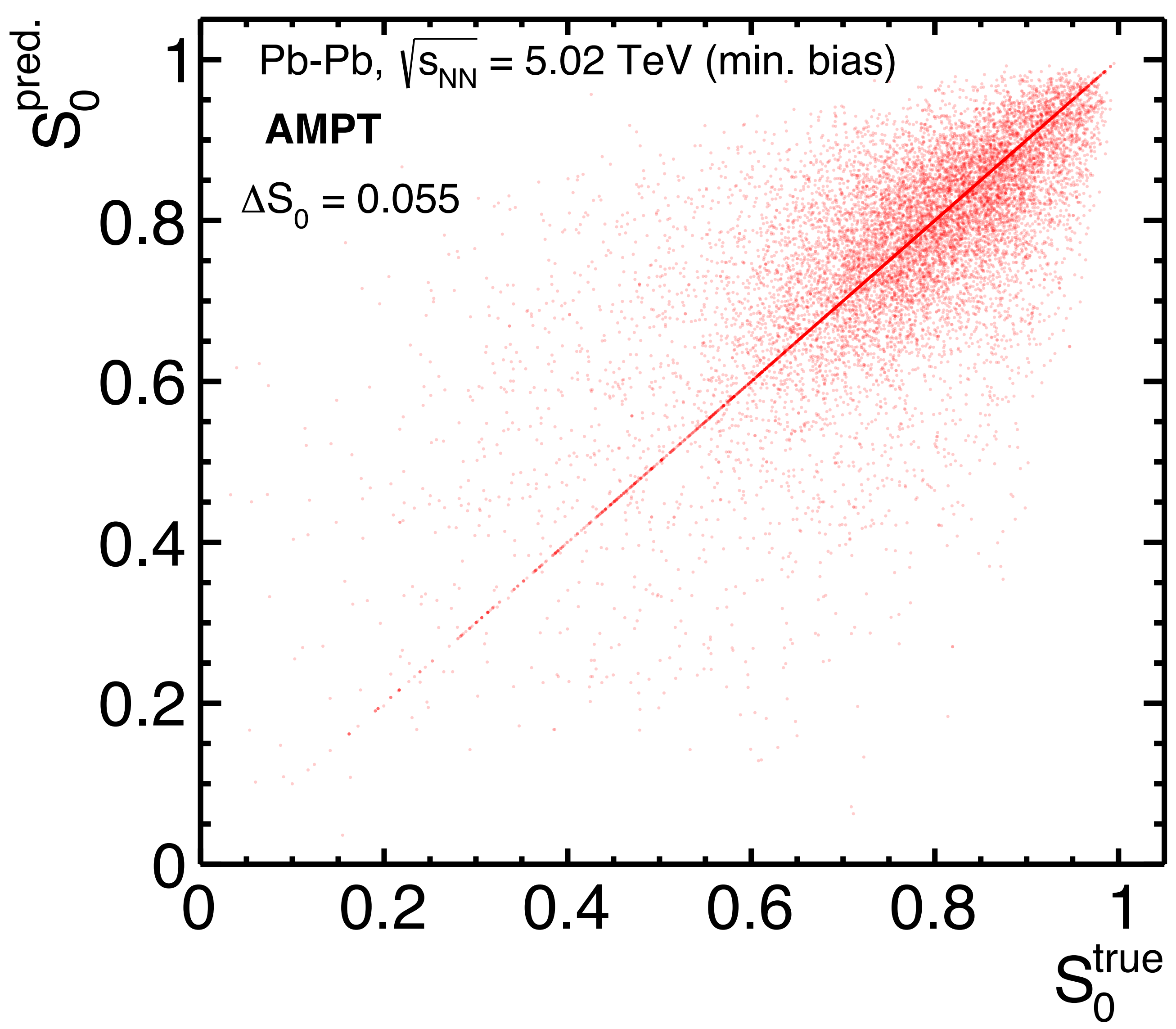


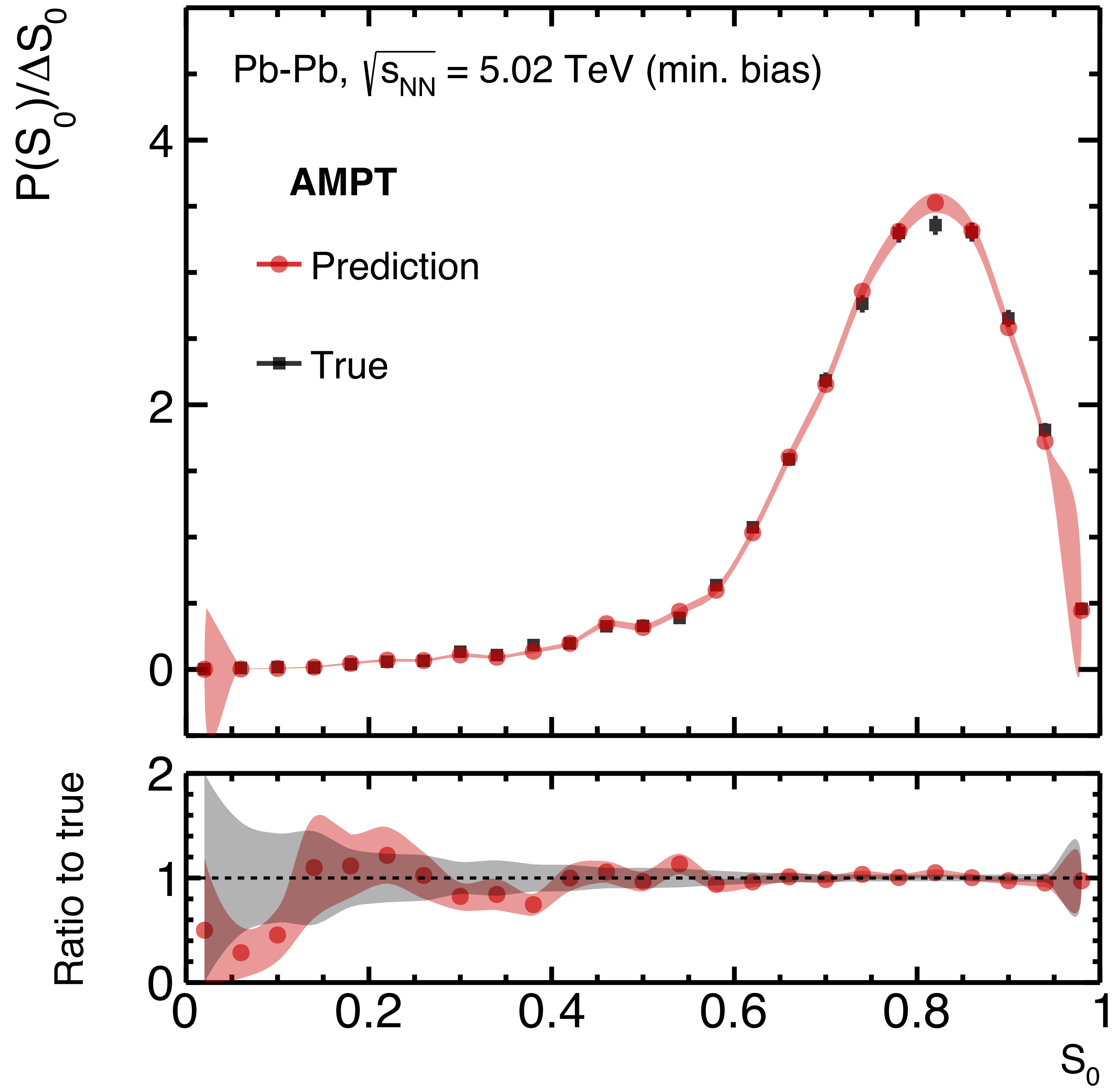
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Results



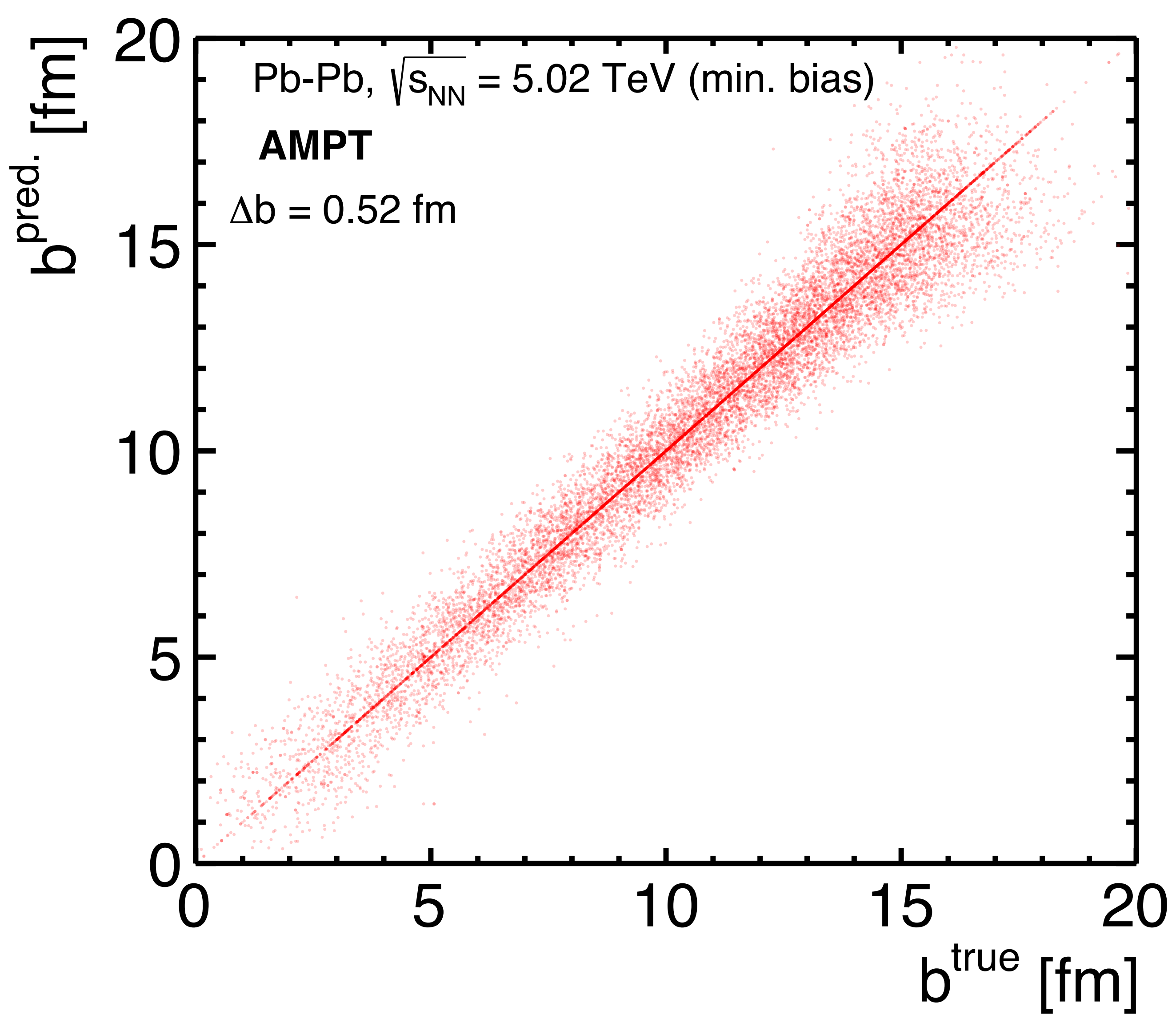


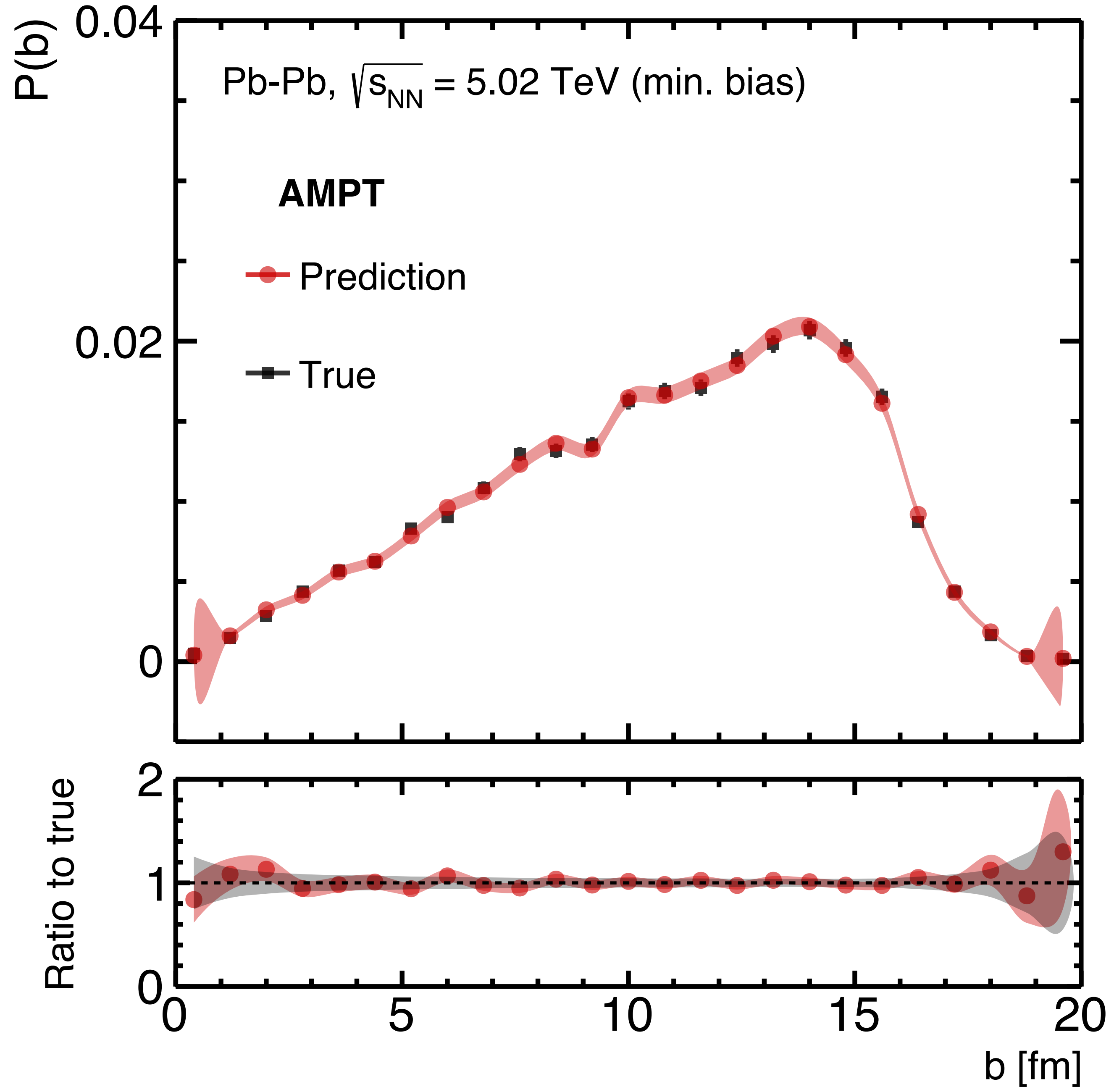


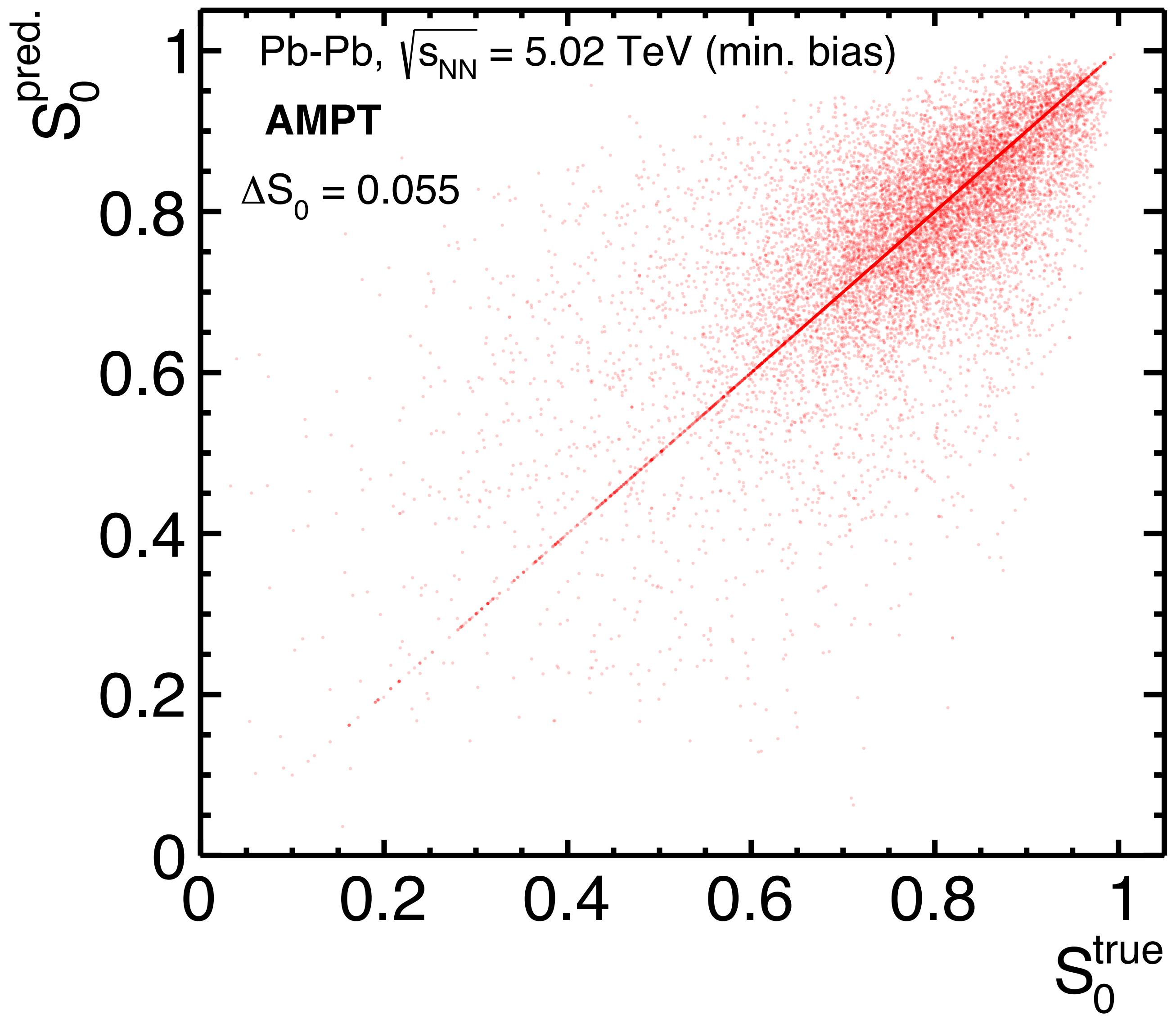


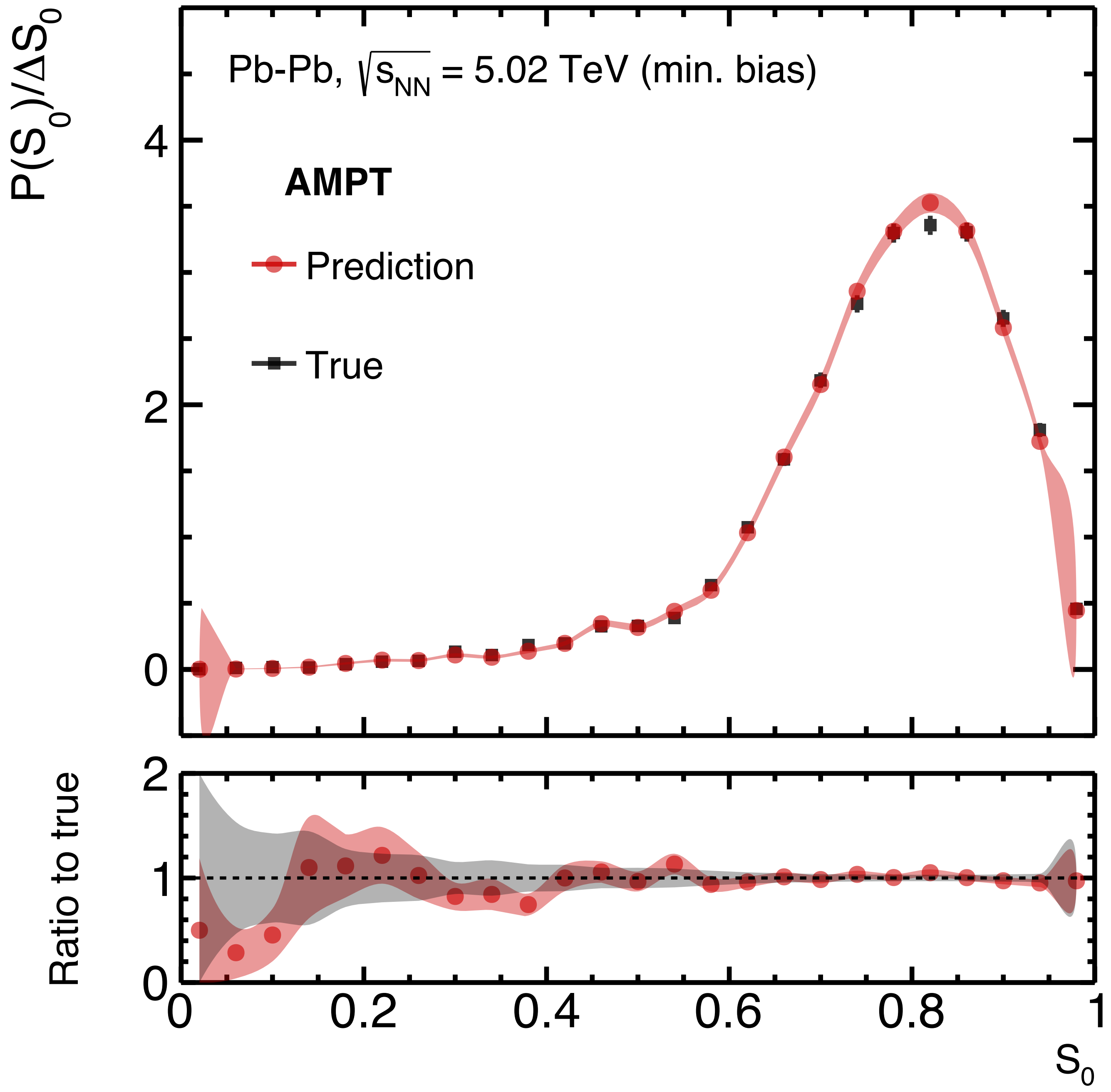
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Results





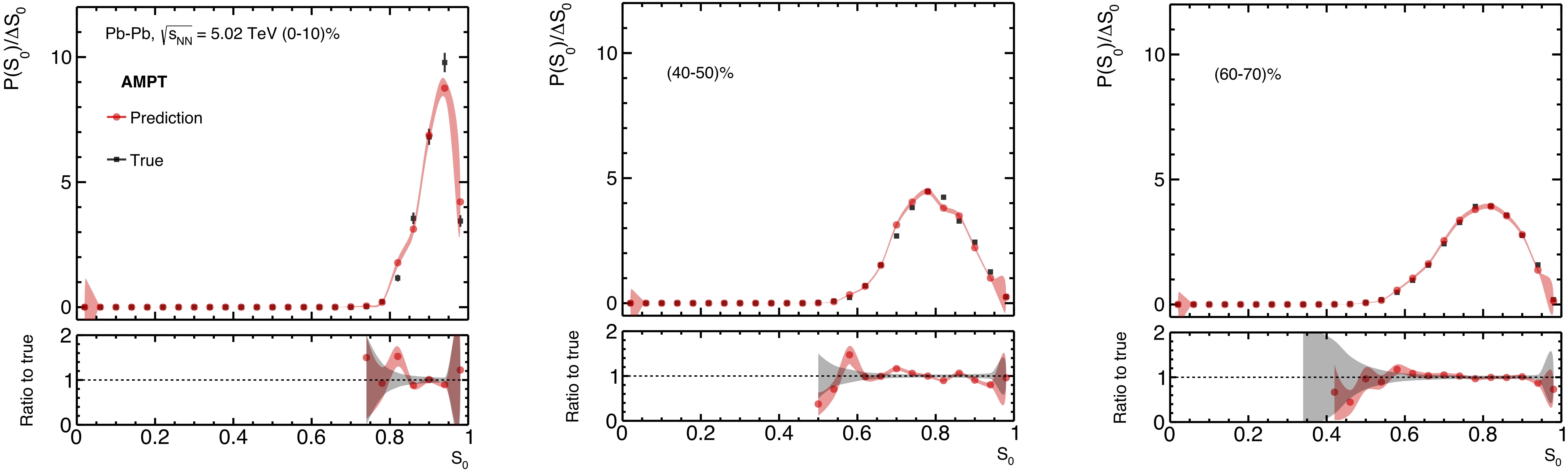




- The ML model trained with 5.02 TeV minimum bias simulated data
- Most of the points populate the straight line inclined at an angle $45°$ with the *x*-axis
- The predictions for both impact parameter and spherocity distributions are in good agreement with the simulated data

N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

Results

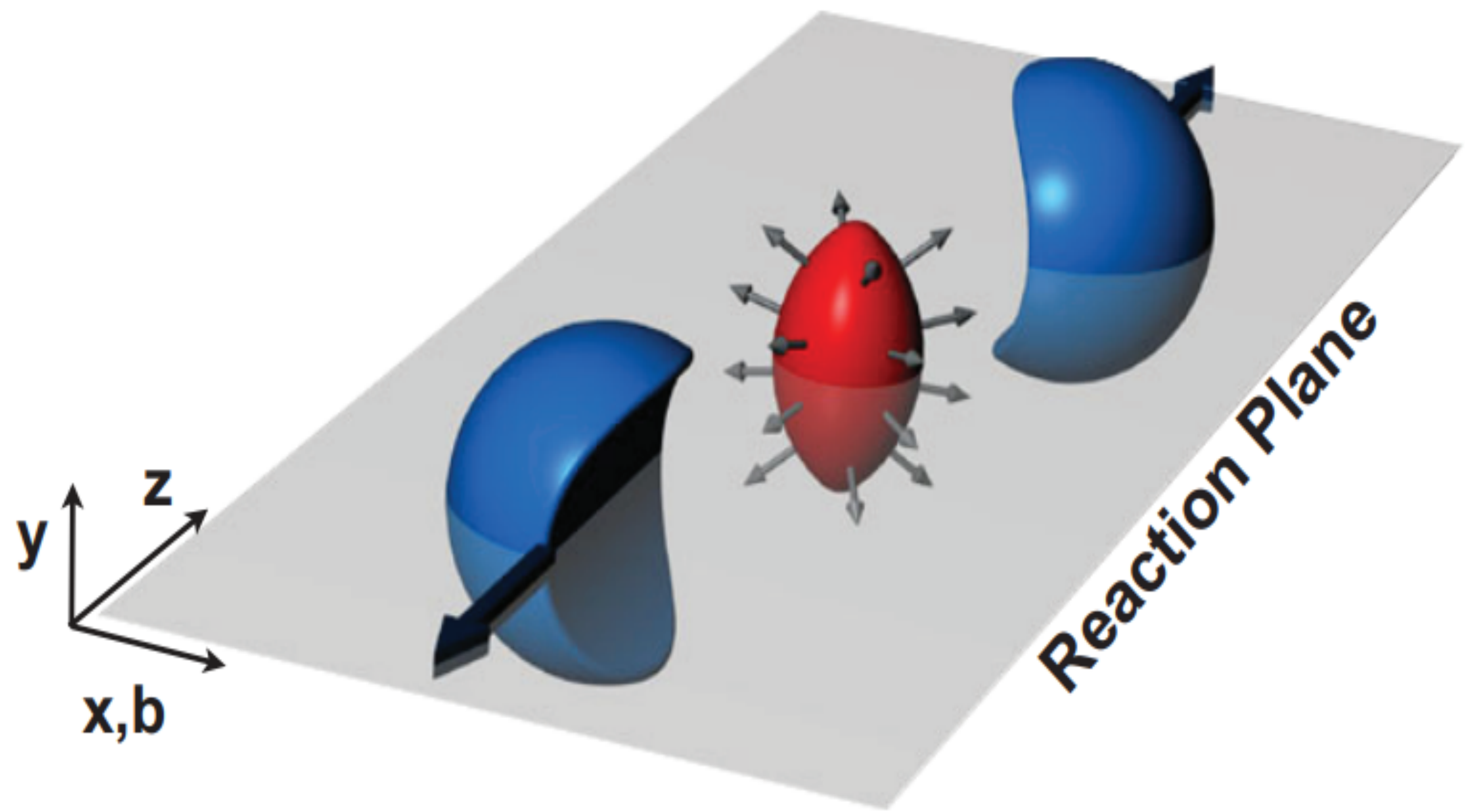


- Centrality wise spherocity distributions
- Training is done using minimum bias simulated data
- **BDT preserves the centrality (or multiplicity) dependence**

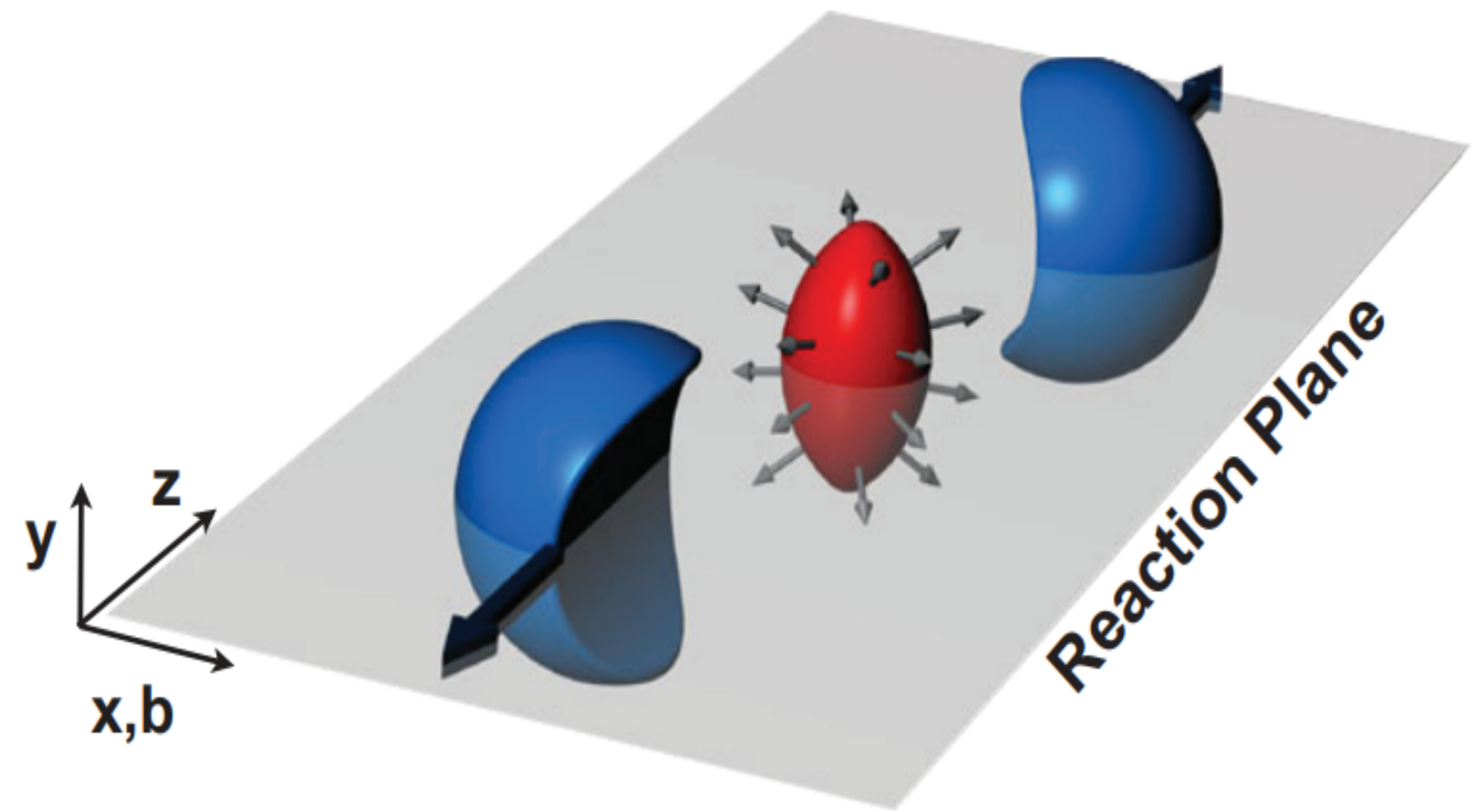
N. Mallick, S. Tripathy, A. N. Mishra, S. Deb, and R. Sahoo, Phys. Rev. D103, 094031 (2021)

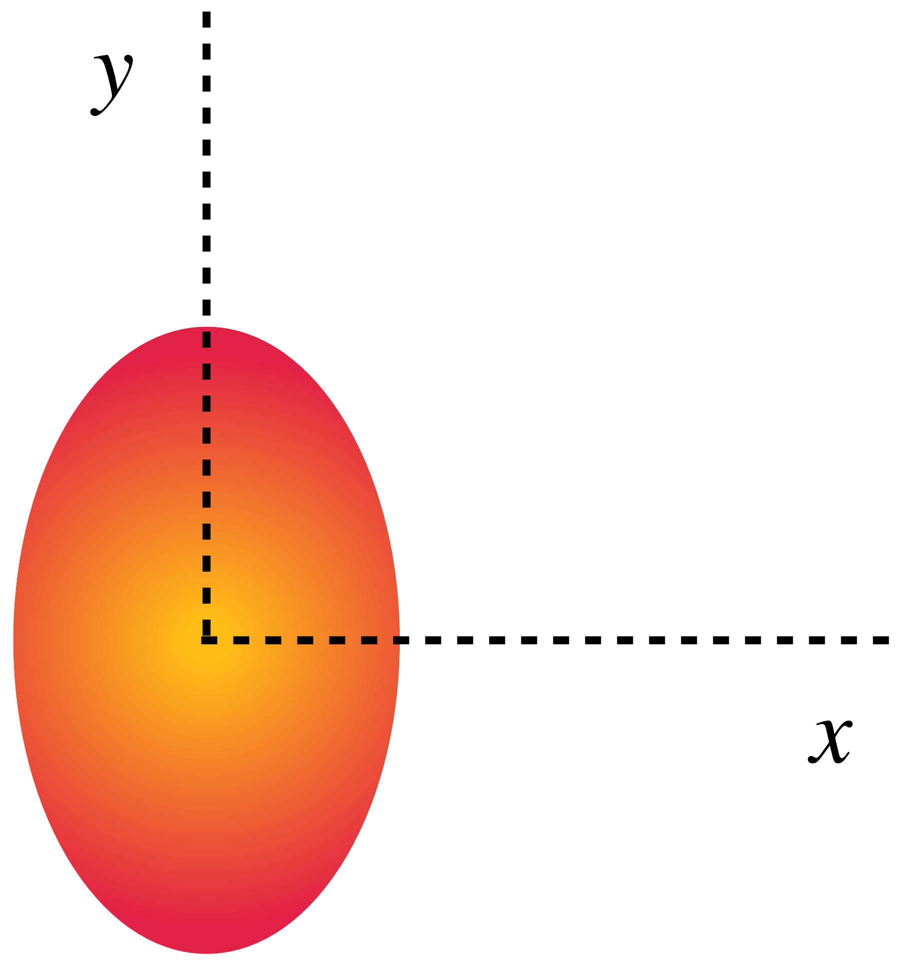
The Elliptic Flow (v_2)

The Elliptic Flow (v_2)

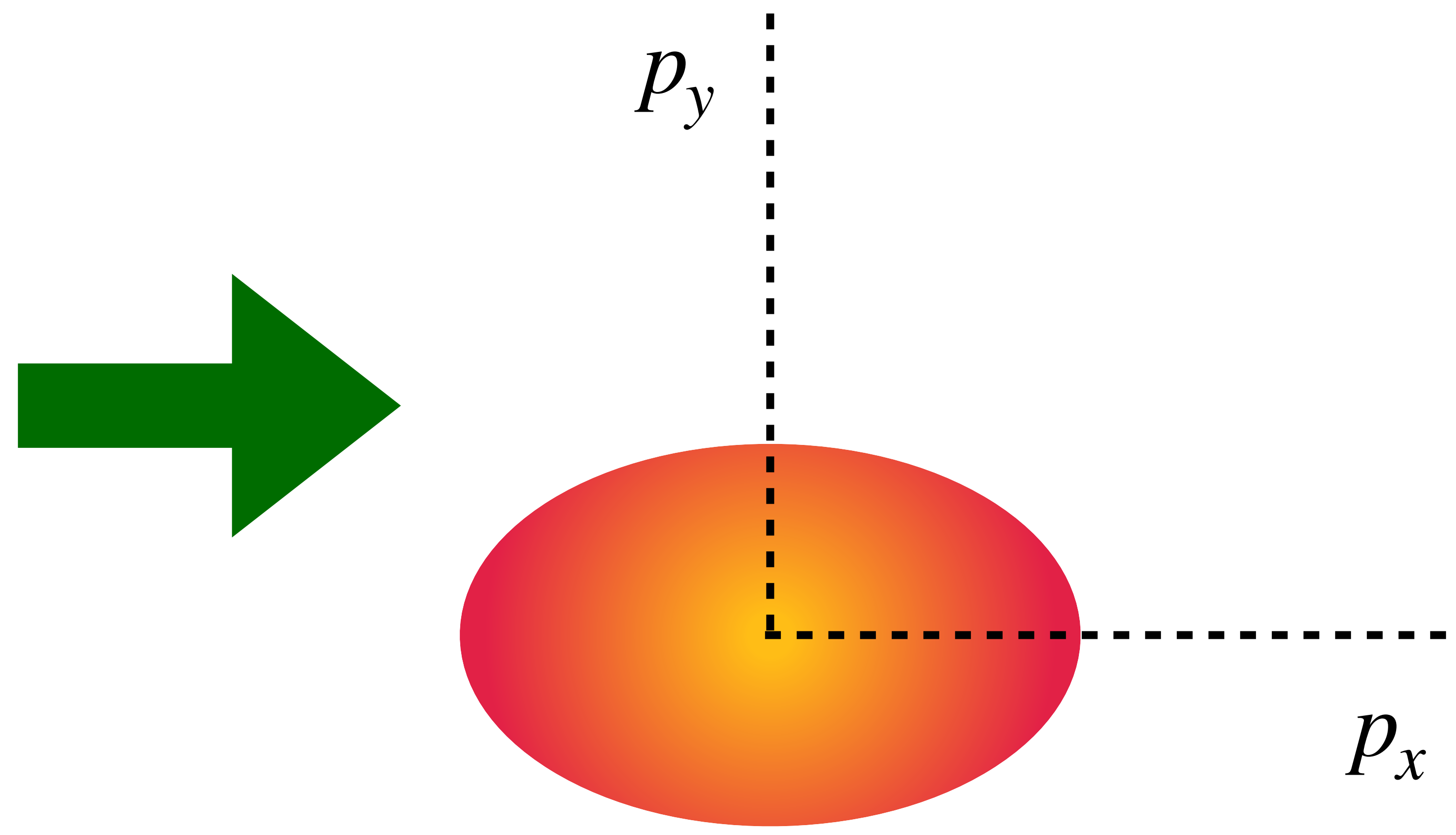


The Elliptic Flow (v_2)



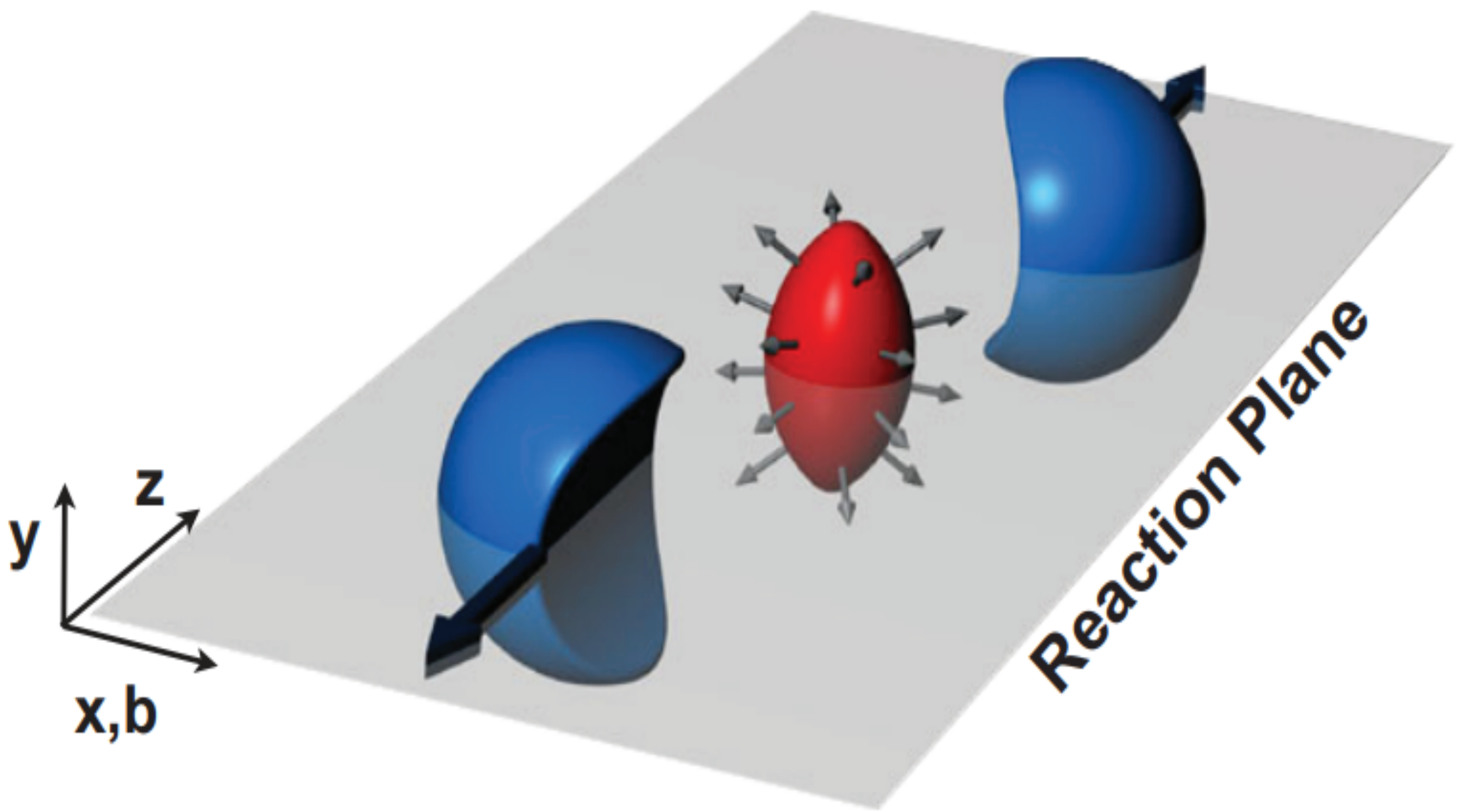


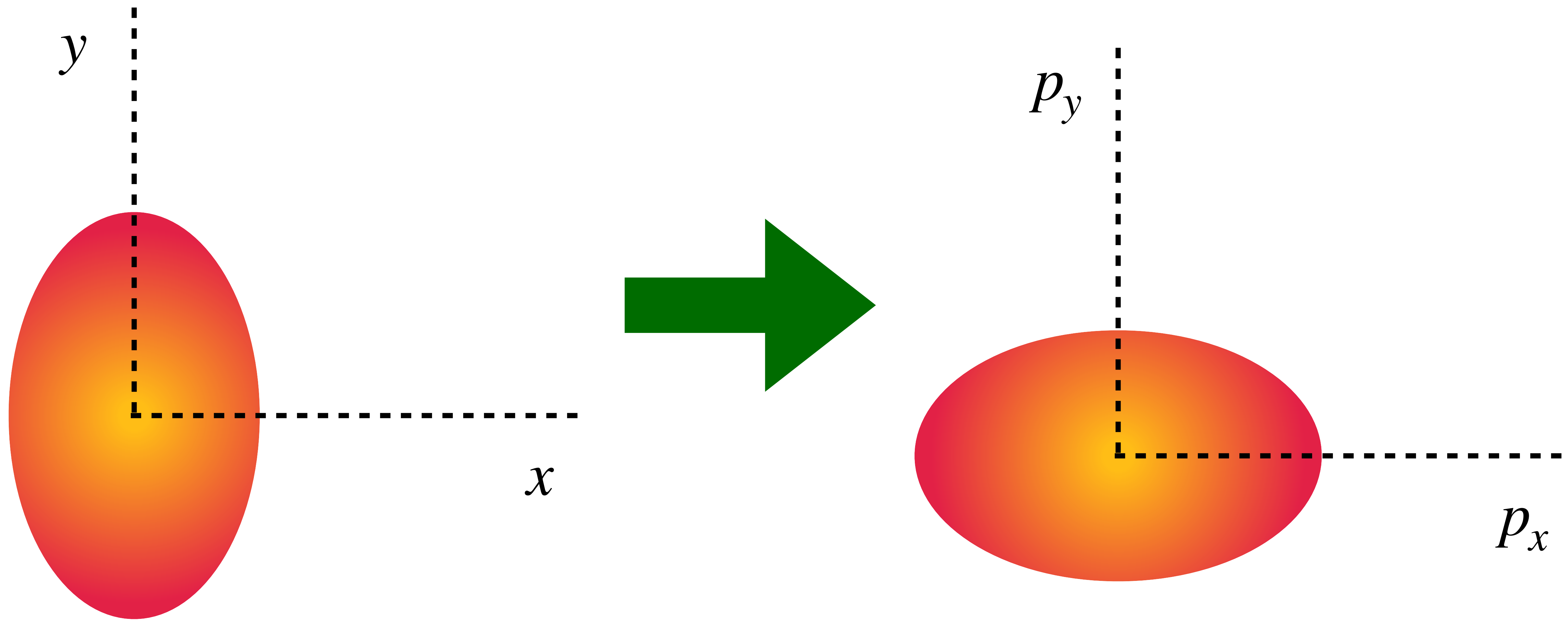
Spatial anisotropy



Azimuthal momentum space anisotropy

The Elliptic Flow (v_2)

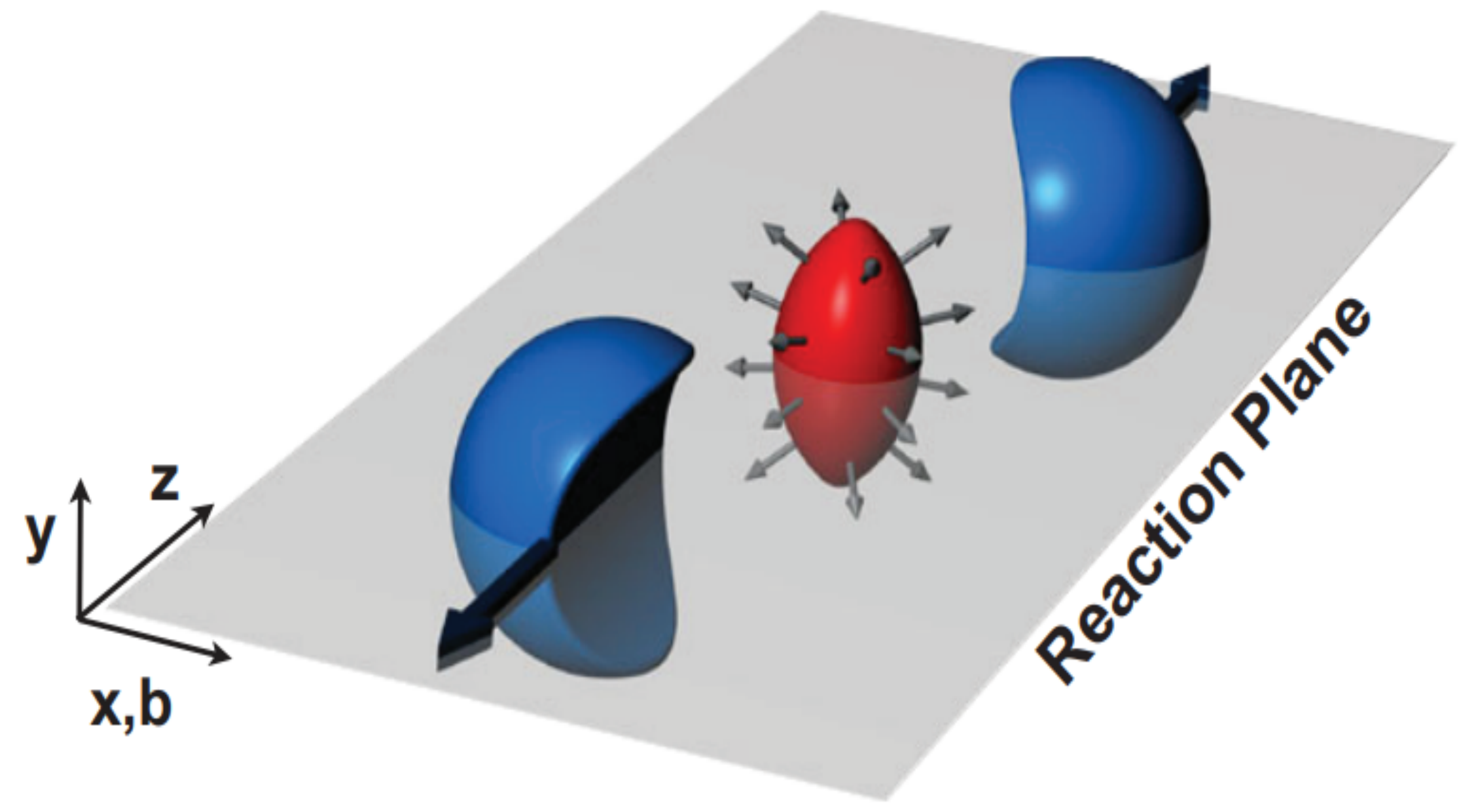


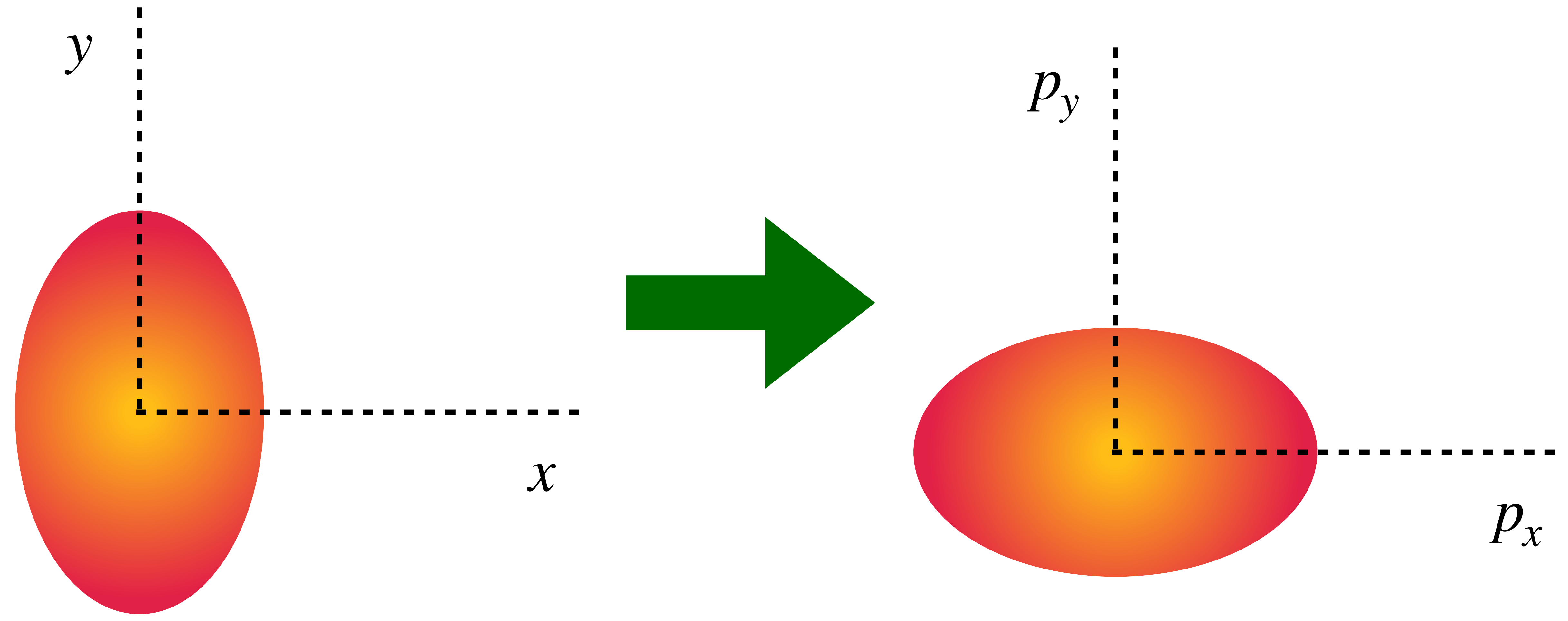


Spatial anisotropy | **Azimuthal momentum space anisotropy**

- Elliptic flow describes the azimuthal momentum space anisotropy of particle emission for a non-central heavy-ion collision
- The 2nd harmonic coefficient of the Fourier expansion of azimuthal momentum distribution $(dN/d\phi)$
- Directly reflects the initial spatial anisotropy of the nuclear overlap region in the transverse plane

The Elliptic Flow (v_2)





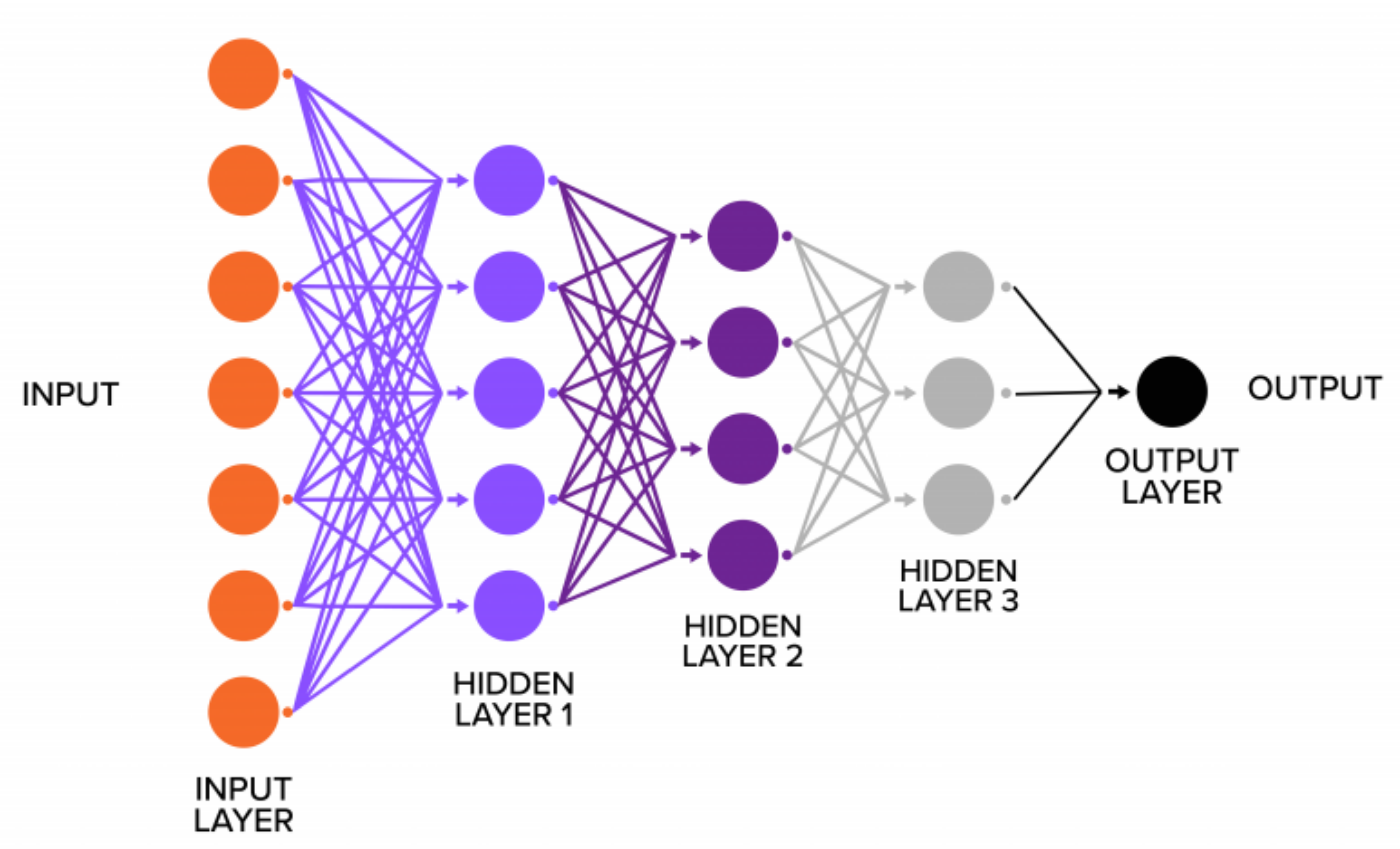
Spatial anisotropy **Azimuthal momentum space anisotropy**

- Elliptic flow describes the azimuthal momentum space anisotropy of particle emission for a non-central heavy-ion collision
- The 2nd harmonic coefficient of the Fourier expansion of azimuthal momentum distribution $(dN/d\phi)$
- Directly reflects the initial spatial anisotropy of the nuclear overlap region in the transverse plane

$$E\frac{d^3N}{dp^3} = \frac{d^3N}{p_\mathrm{T}dp_\mathrm{T}dyd\phi} = \frac{d^2N}{p_\mathrm{T}dp_\mathrm{T}dy}\frac{1}{2\pi}\left(1 + 2\sum_{n=1}^{\infty} v_n \cos[n(\phi - \psi_n)]\right)$$

$$v_2(p_\mathrm{T}, y) = \langle \cos(2(\phi - \psi_2)) \rangle$$
$$\phi = \tan^{-1}(p_y/p_x)$$

Deep Neural Network (DNN)



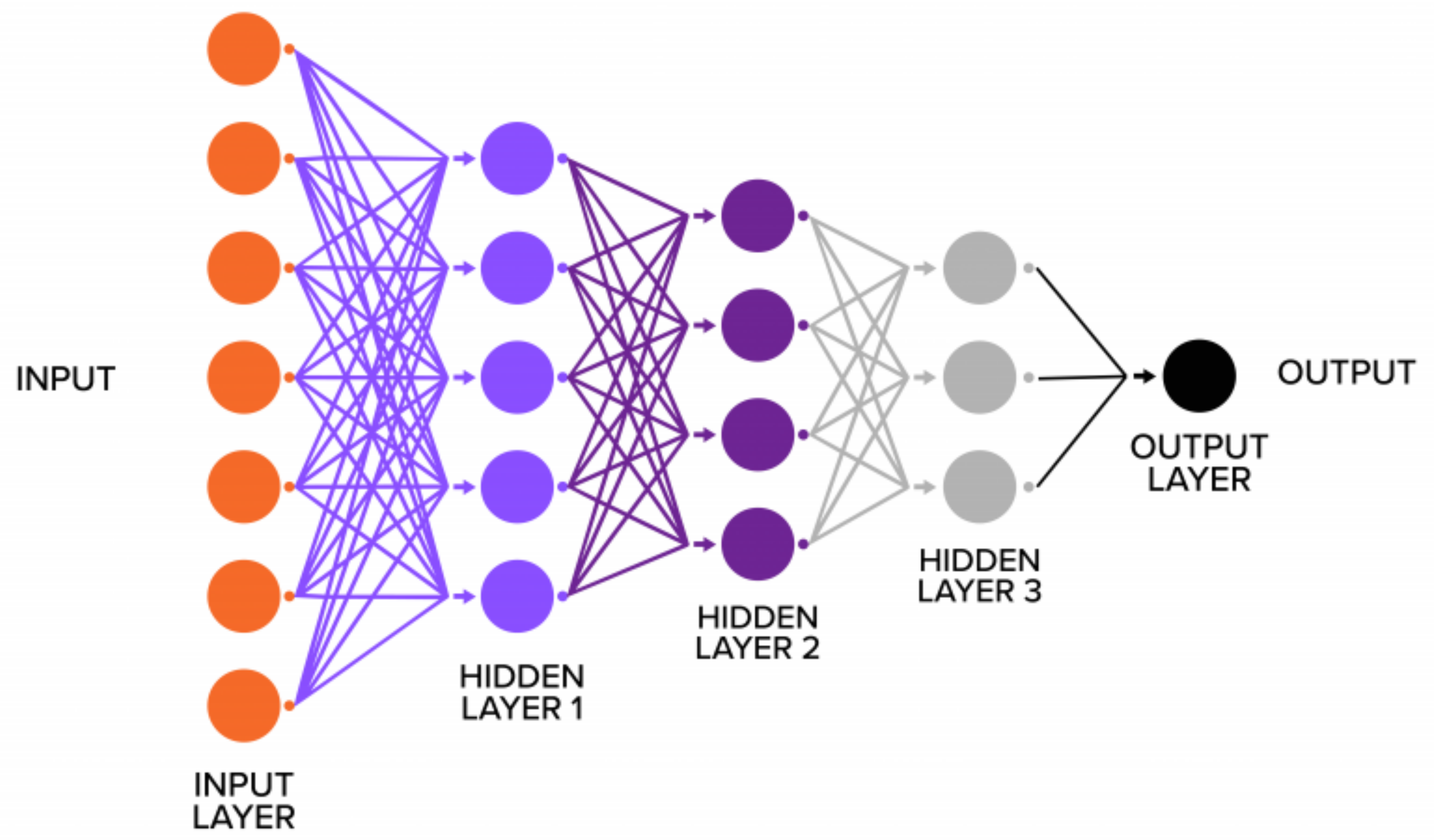
Deep Neural Network (DNN)

- ML Algorithm inspired from neurons in animal brains
- Three key layers

  Input: Takes the features as input

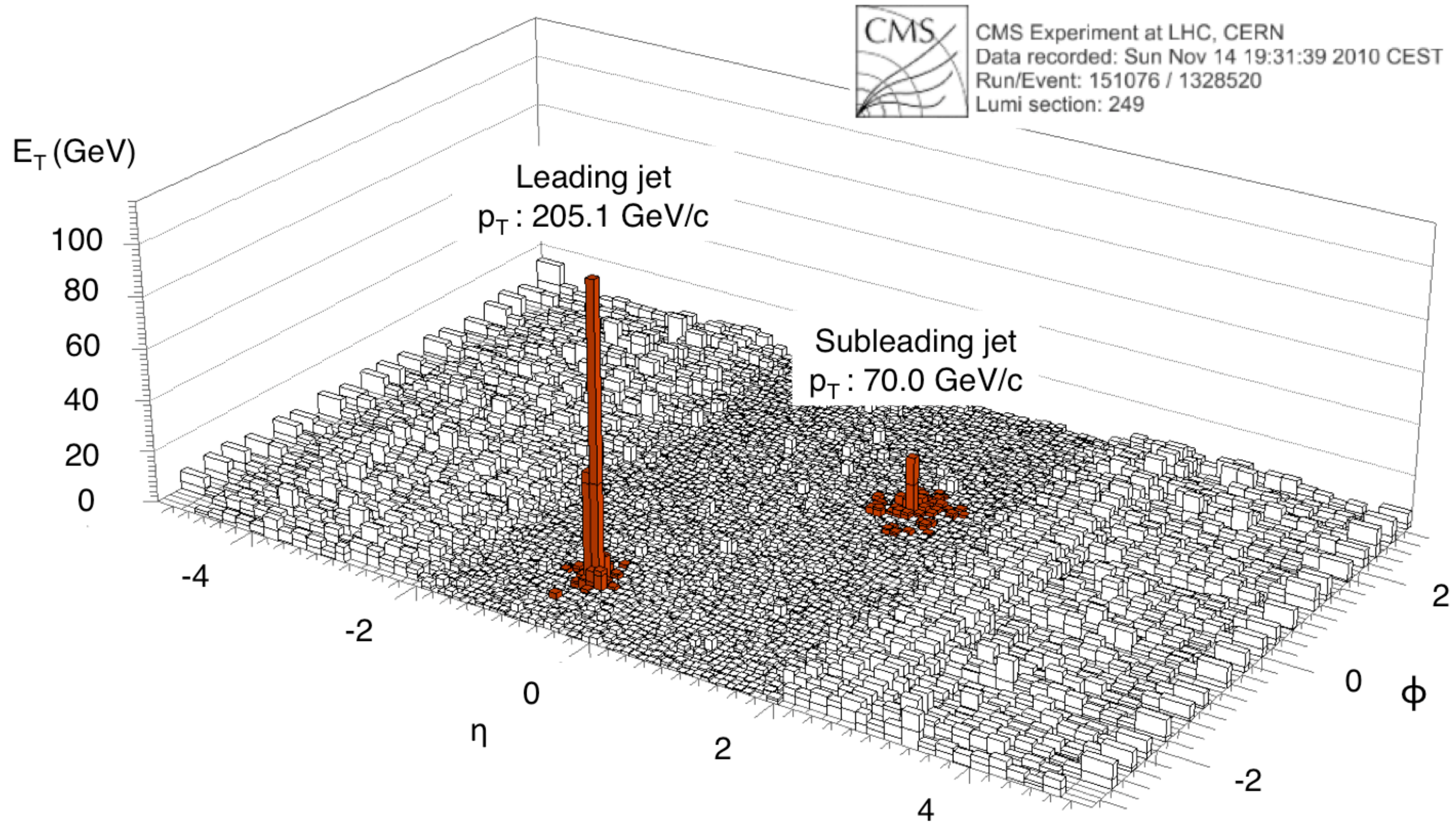
  Hidden layers: Connects to each neuron through different weights

  Output: Gives the result as a number or class
- **Weights** dictate the importance of an input → more important features get more weights
- **Activation function**: mathematical function that guides the outcome at each node → Standardize the values
- **Cost function**: Evaluates the accuracy between machine prediction and true value
- **Optimizer**: Method (or algorithm) that minimizes the cost function by automatically updating the weights



Estimation of elliptic flow (v_2)

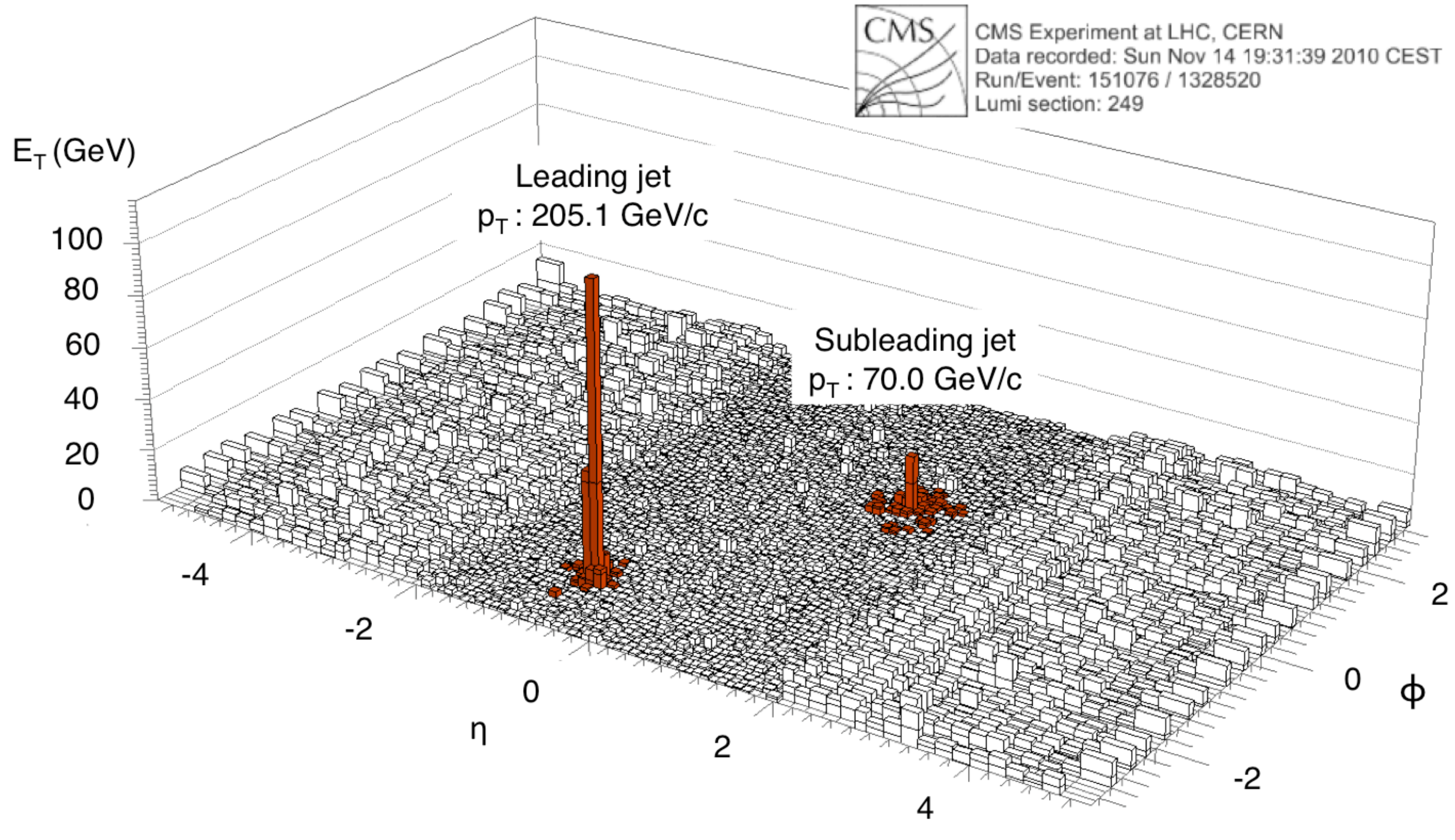
Estimation of elliptic flow (v_2)



Serguei Chatrchyan et al., Phys.Rev.C 84 (2011), 024906

Estimation of elliptic flow (v_2)

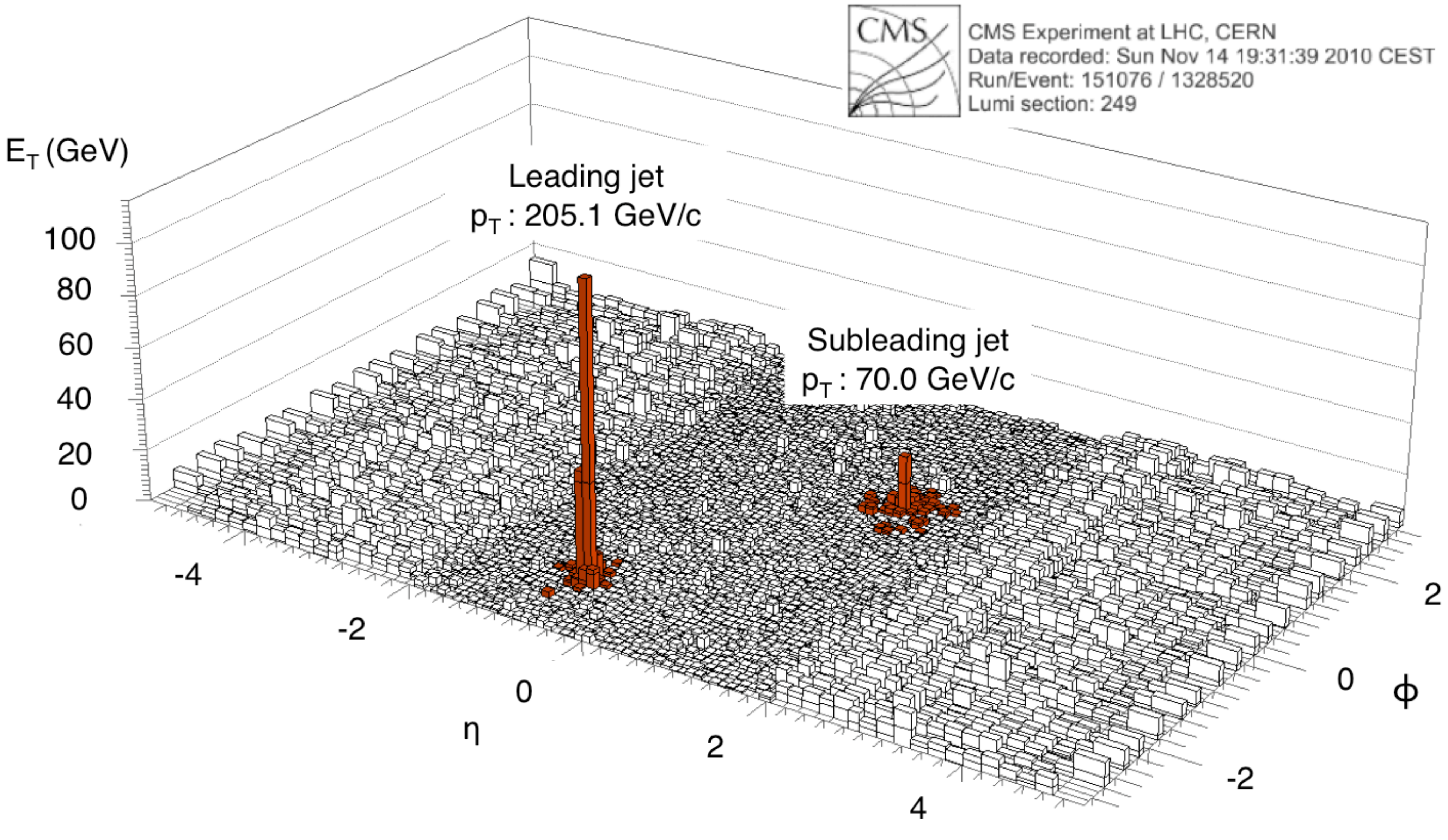
- Estimation of elliptic flow using Deep Neural Network
- Elliptic flow -> Event property
- Inputs ->Track property
- $(\eta - \phi)$ space could be taken as the primary input space
- Three layers having different weights
- p_{T}, mass and $\log(\sqrt{s_{\mathrm{NN}}}/s_0)$ weighted layers serve as the secondary input space



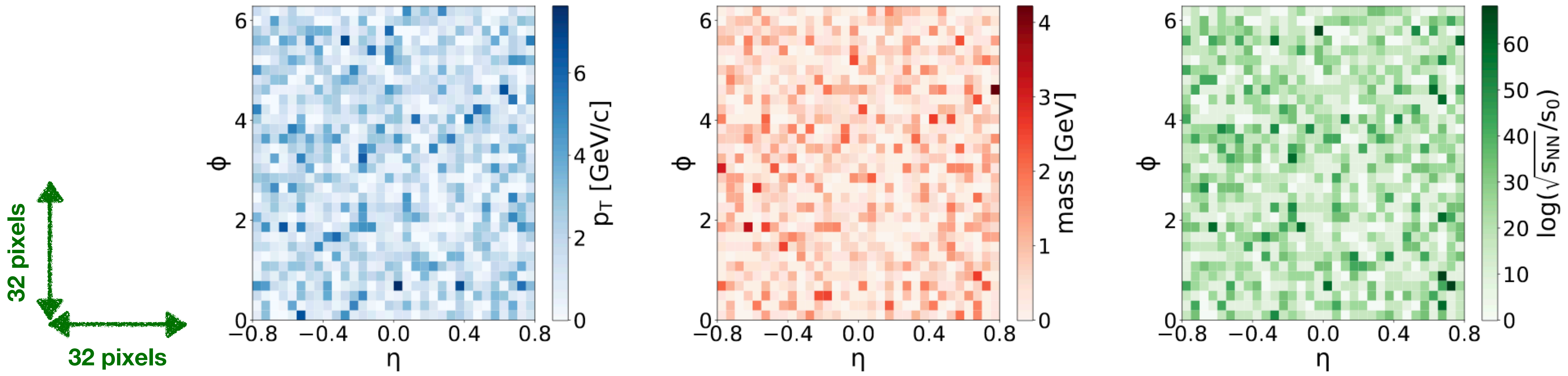
Serguei Chatrchyan et al., Phys.Rev.C 84 (2011), 024906

Estimation of elliptic flow (v_2)

- Estimation of elliptic flow using Deep Neural Network
- Elliptic flow -> Event property
- Inputs ->Track property
- $(\eta - \phi)$ space could be taken as the primary input space
- Three layers having different weights
- p_T, mass and $\log(\sqrt{s_{NN}}/s_0)$ weighted layers serve as the secondary input space



Serguei Chatrchyan et al., Phys.Rev.C 84 (2011), 024906



Pb-Pb, $\sqrt{s_{NN}} = 5.02$ TeV, **AMPT Simulation**

DNN Model:

DNN Model:

- Each space has 32×32 pixels (grids)
- Total number of pixel points = $32 \times 32 \times 3 = 3072$ for each event
- DNN with the following architecture

  Input Layer: 128 Nodes

  Three hidden layers: 256 Nodes each

  Final layer : 1 node (v_2)

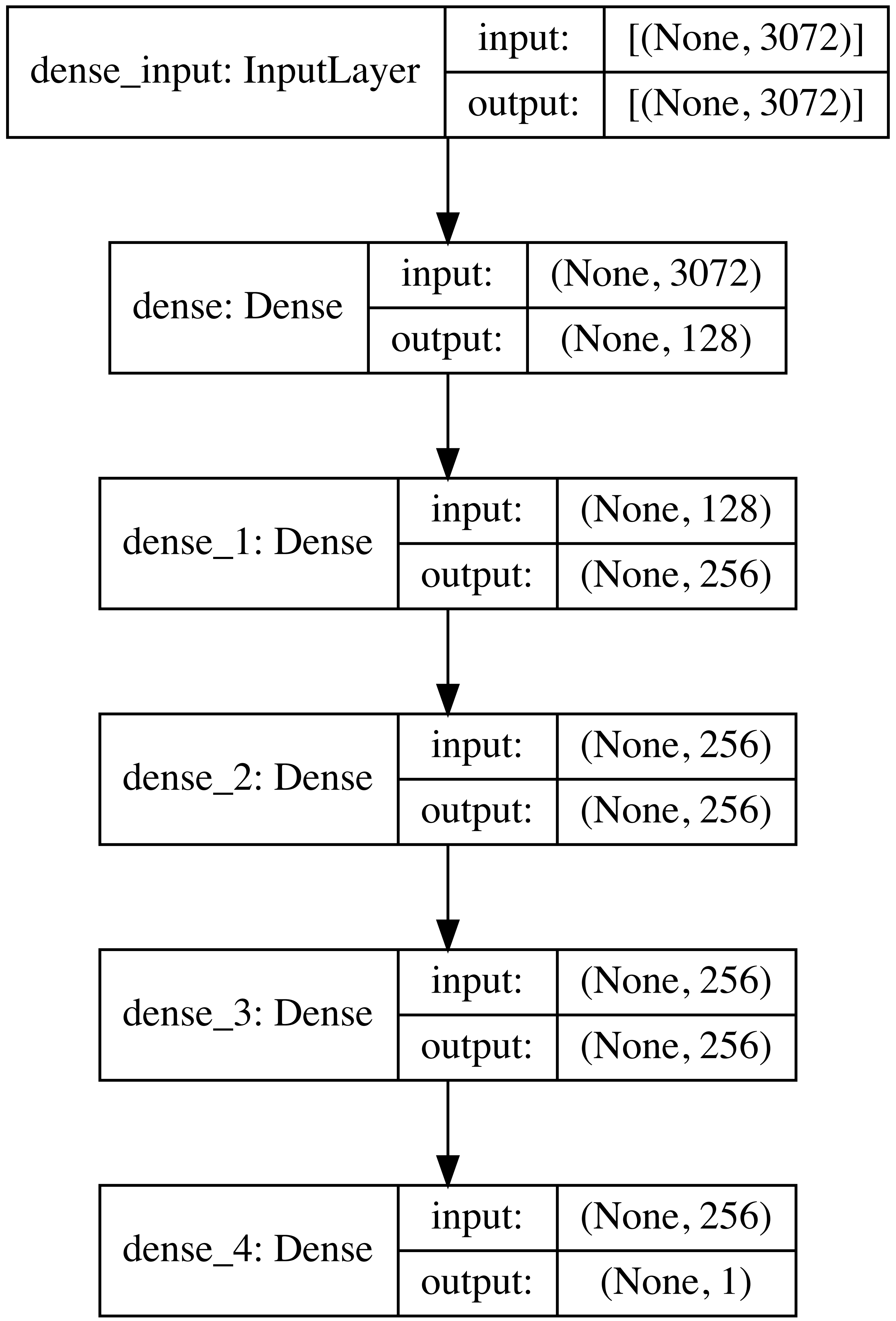
DNN Model:

- Each space has 32×32 pixels (grids)
- Total number of pixel points = $32 \times 32 \times 3 = 3072$ for each event
- DNN with the following architecture

  Input Layer: 128 Nodes

  Three hidden layers: 256 Nodes each

  Final layer : 1 node (v_2)



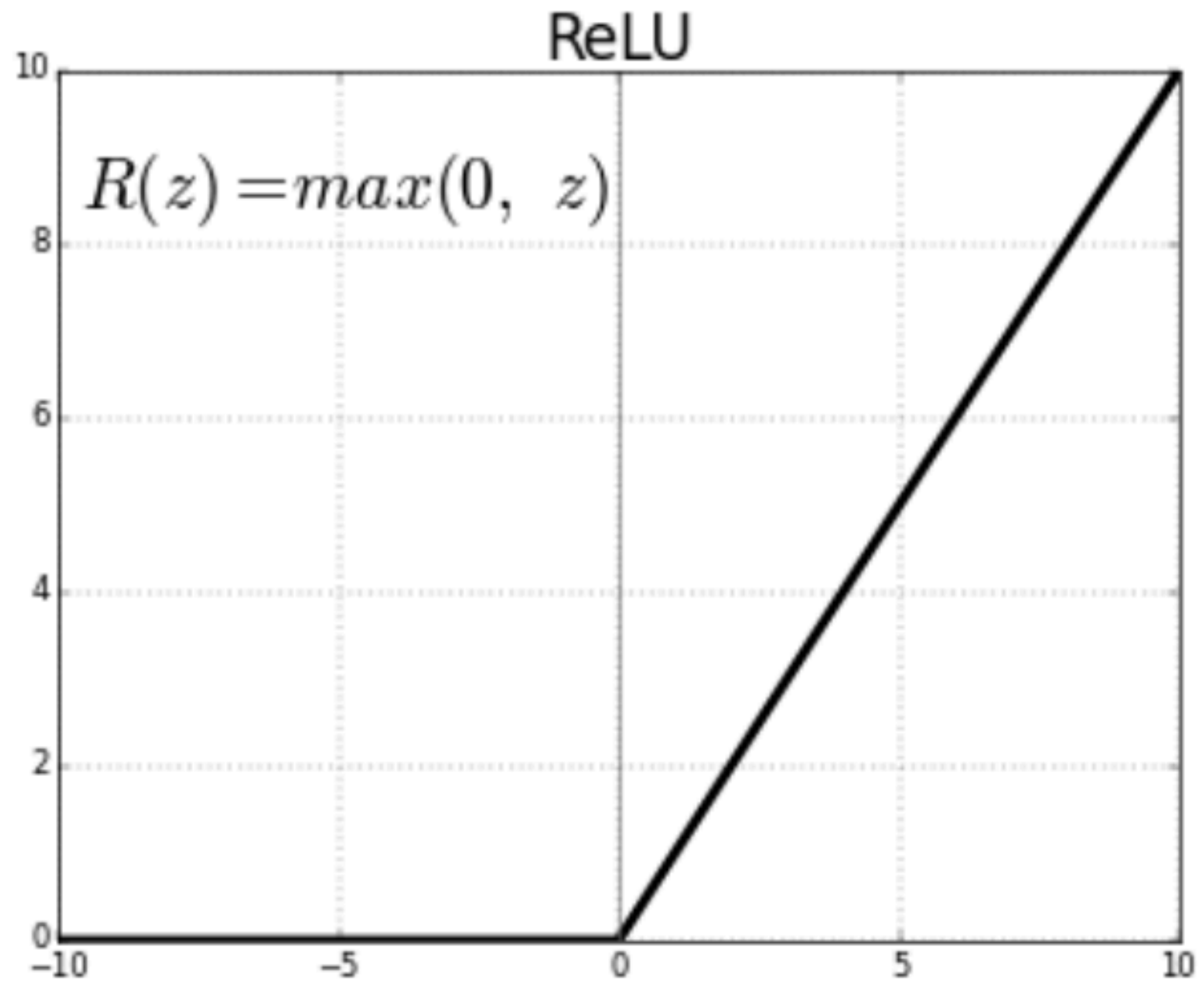
DNN Model:

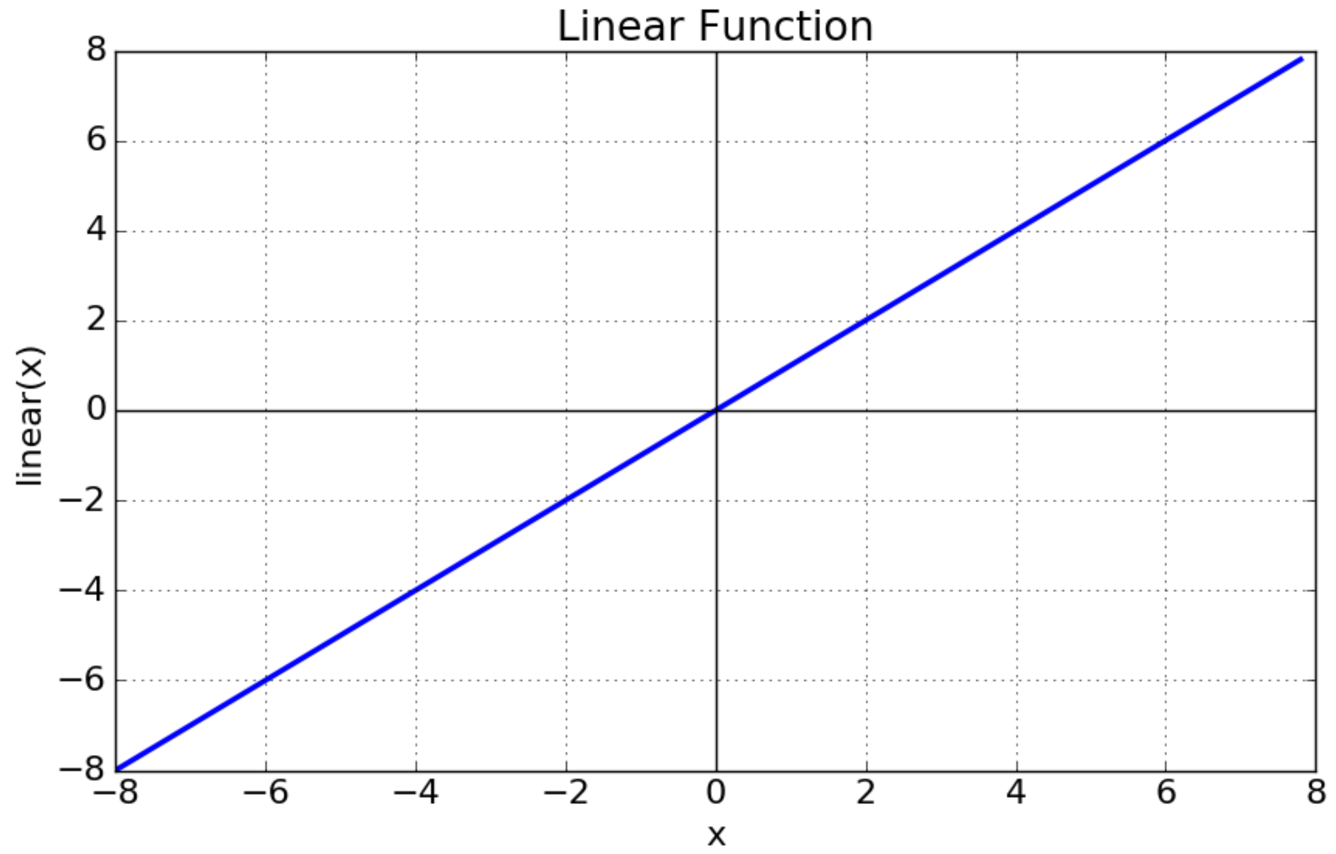
- Each space has 32×32 pixels (grids)
- Total number of pixel points = $32 \times 32 \times 3 = 3072$ for each event
- DNN with the following architecture

  Input Layer: 128 Nodes

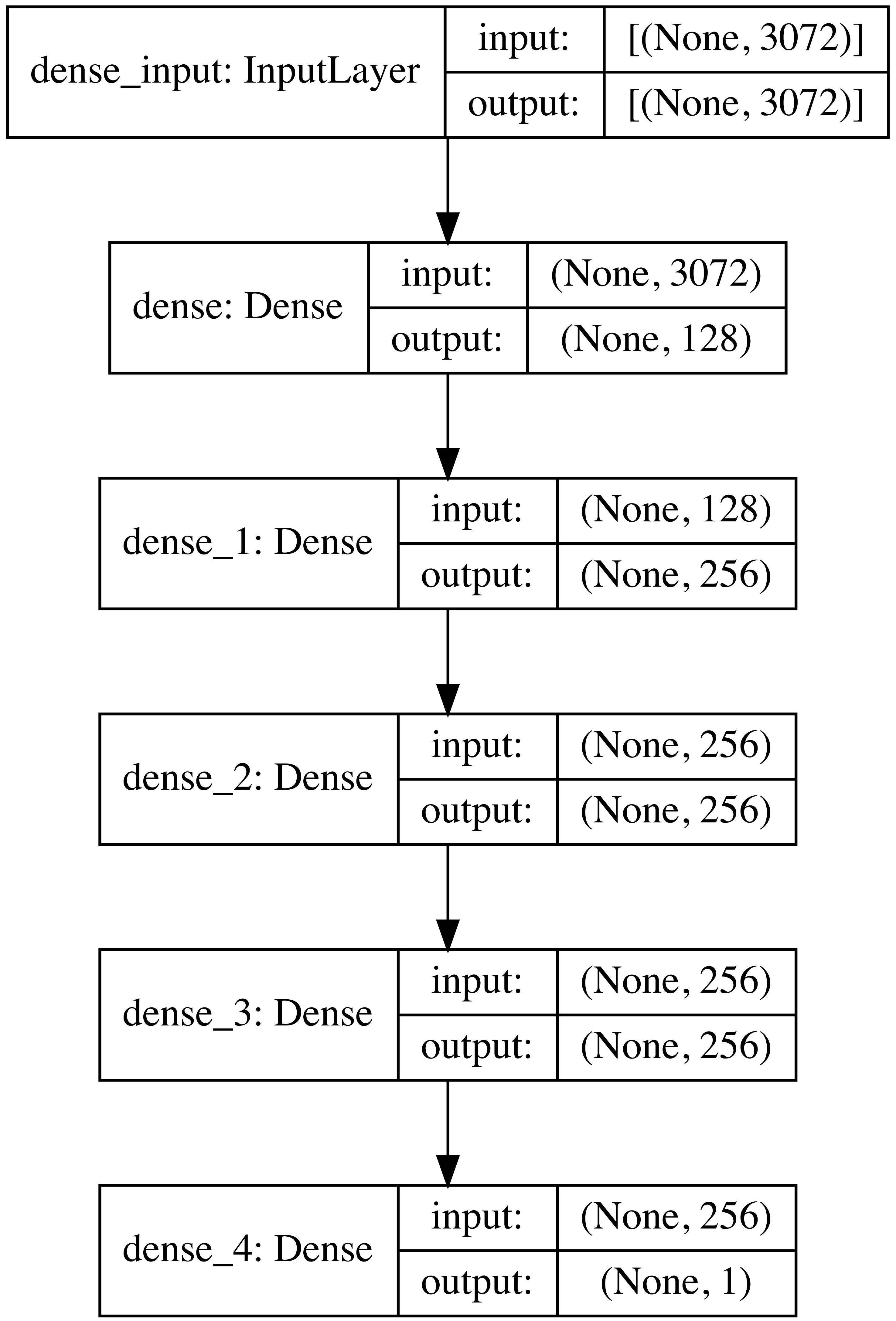
  Three hidden layers: 256 Nodes each

  Final layer : 1 node (v_2)

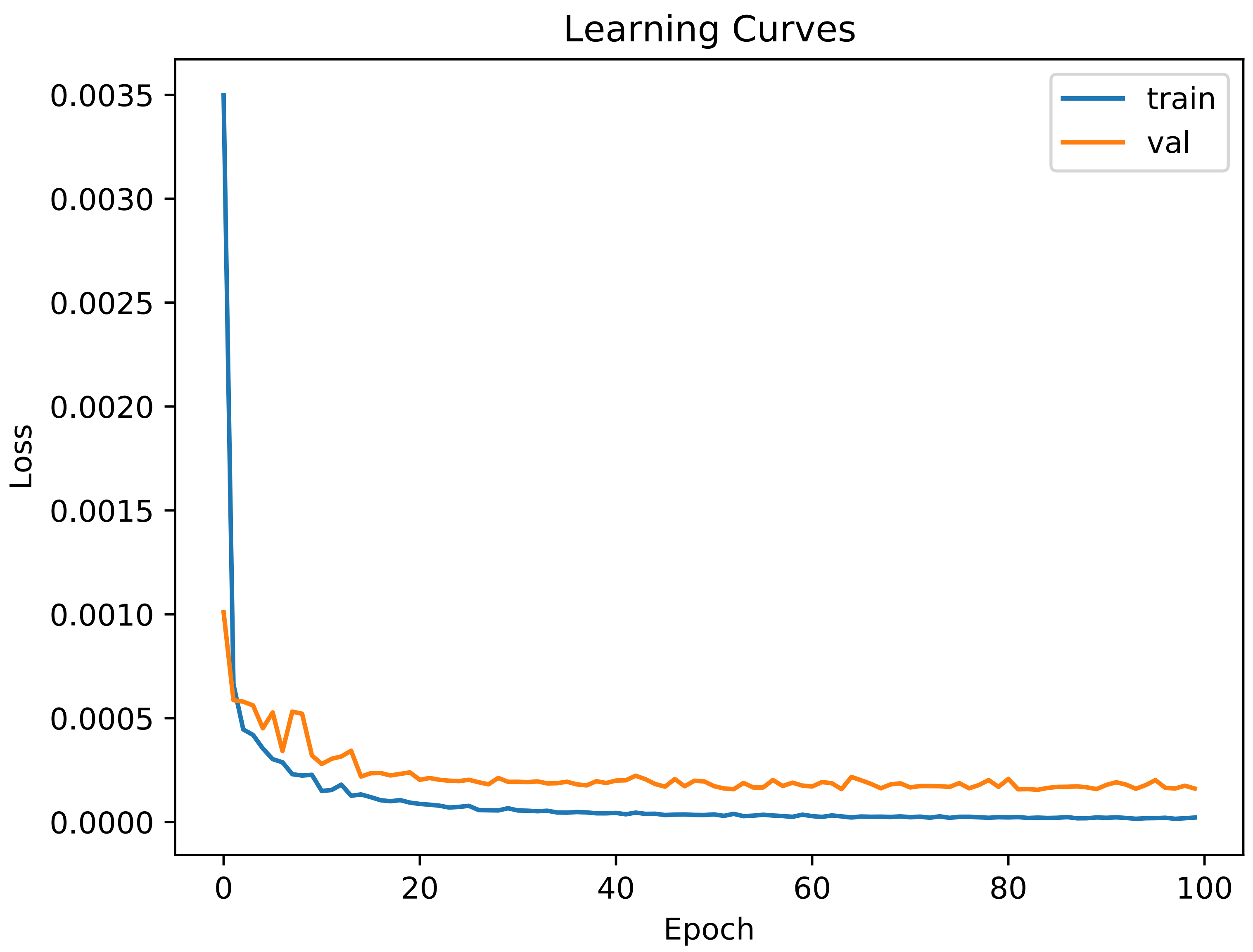


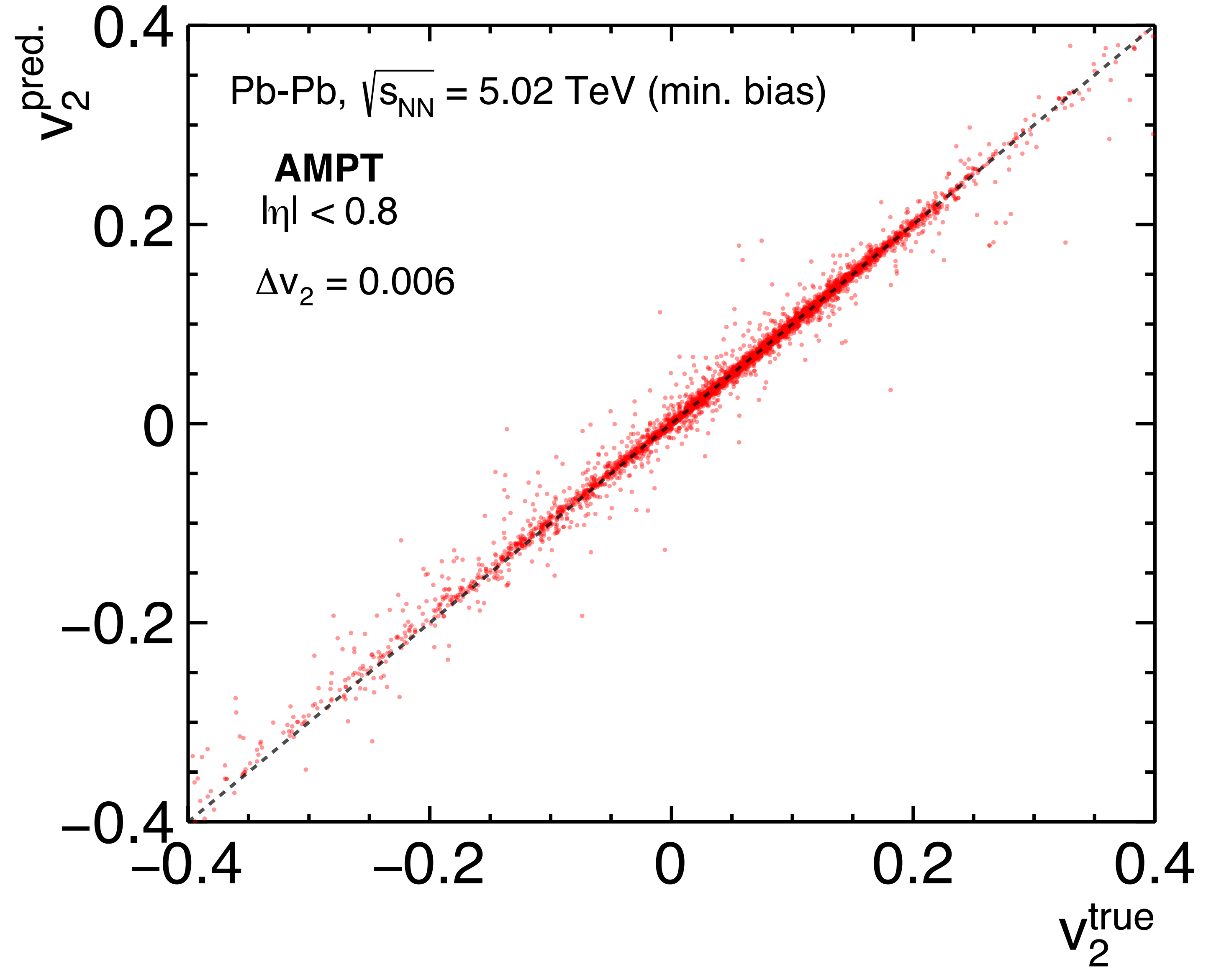


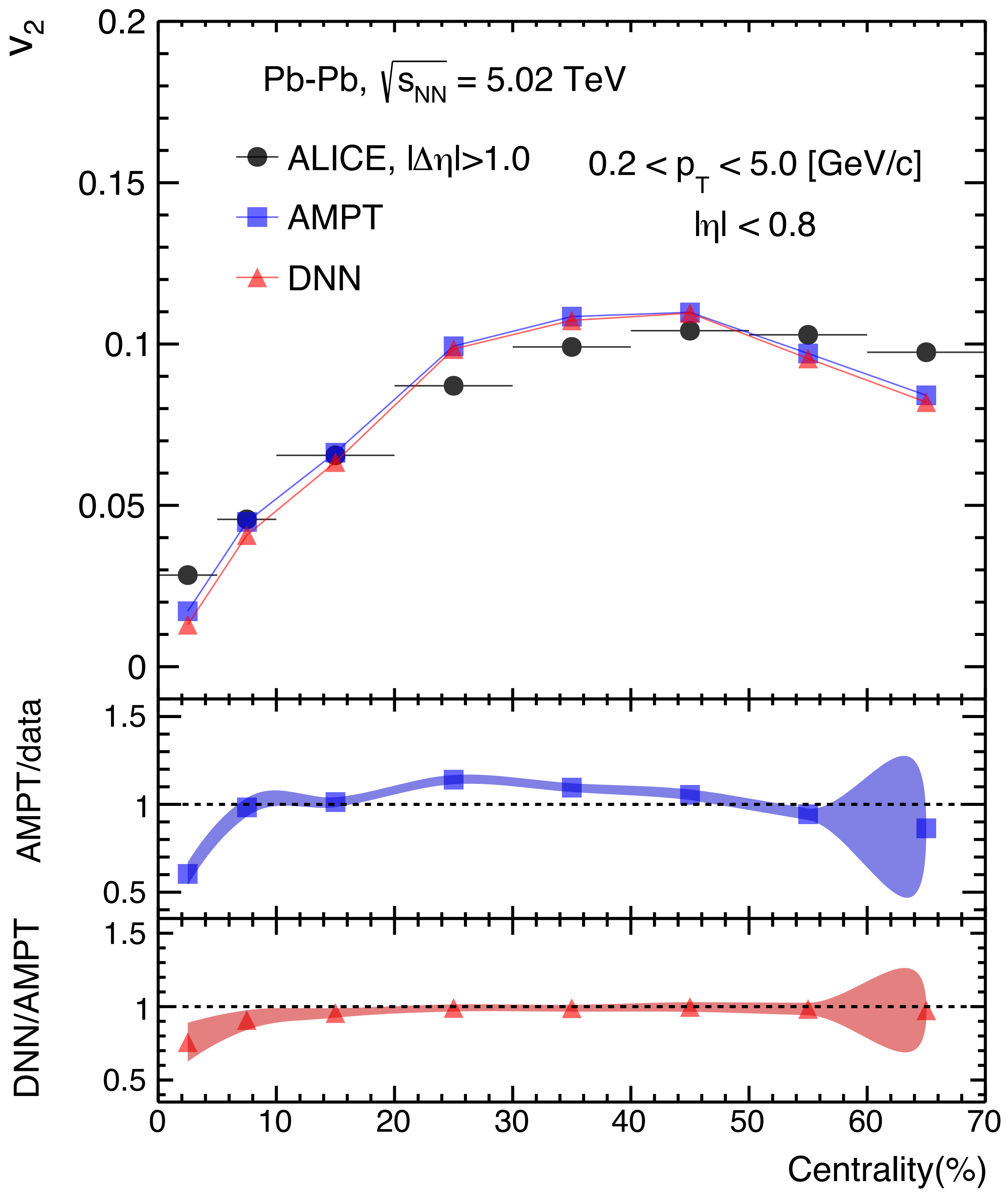
- Input and hidden layers have *ReLu* Activation
- Output layer has *Linear* activation
- Optimzer: $adam$, Loss function: mse
- Epoch: 100, Batch Size: 32
- Training: 10^5 Events (~25 GB)
- Validation: 10^4 Events

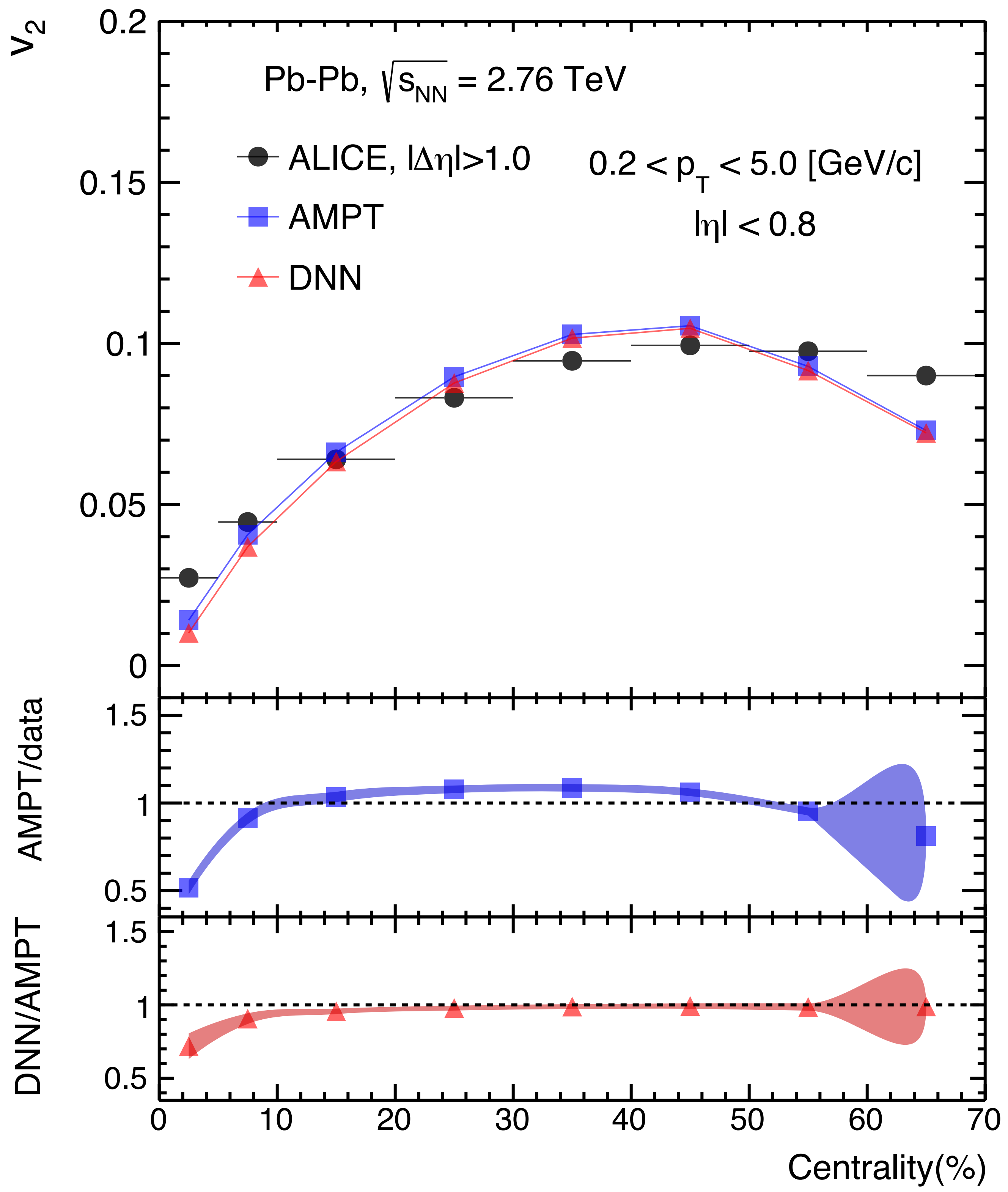


Results



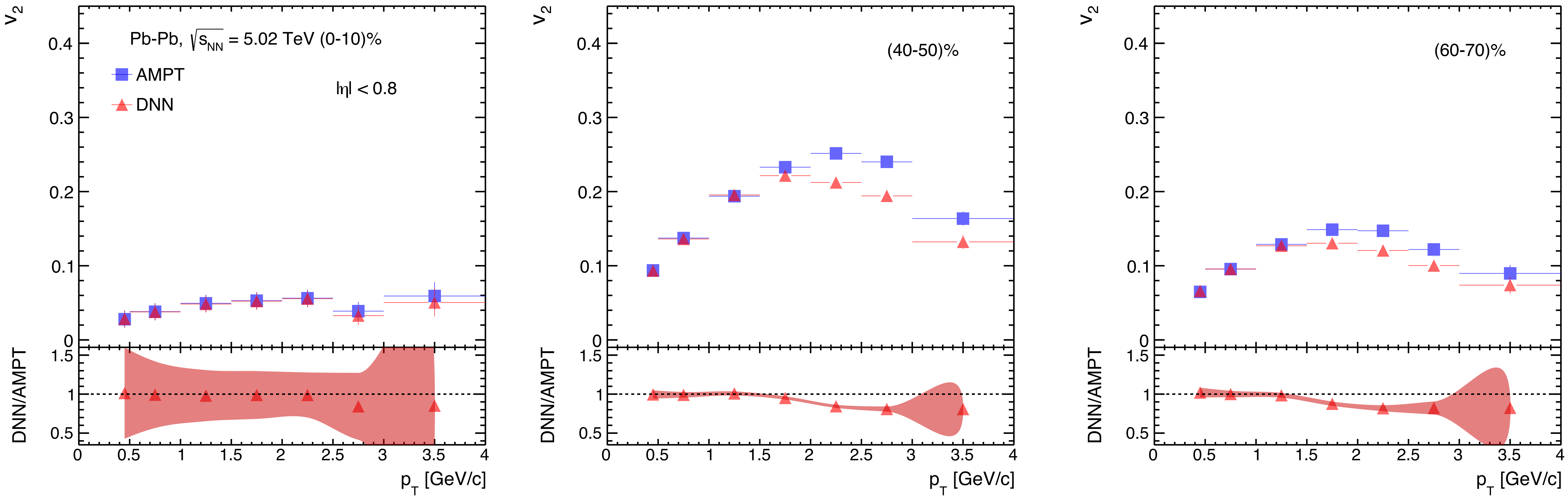






- DNN trained with 5.02 TeV minimum bias simulated data
- Good agreement between the simulated and predicted values of v_2
- ML model applied to lower energy
- **DNN preserves energy dependence of v_2**

Results



- Training done in the range: $0.2 < p_T < 5.0$ [GeV/c]
- Applied to different slices of p_T-bins: [0.4,0.5,1.0,1.5,2.0,2.5,3.0,4.0]
- Elliptic flow as a function of transverse momentum
- **DNN preserves the p_T dependence of v_2**
- For lower p_T, almost perfect prediction is achieved
- Slight mismatch is observed for higher p_T

Summary and outlook

- Implementation of ML tools for the estimation of impact parameter, transverse spherocity and elliptic flow
- ML preserves the energy, centrality and p_{T} dependence of particle production
- Good agreement observed between the simulation and prediction
- Efficient and faster application
- Final state particle information could be used
- To estimate the non-flow contributions and reduce it
- Estimation of identified particle elliptic flow
- Quark and gluon jets

Summary and outlook

- Implementation of ML tools for the estimation of impact parameter, transverse spherocity and elliptic flow
- ML preserves the energy, centrality and p_{T} dependence of particle production
- Good agreement observed between the simulation and prediction
- Efficient and faster application
- Final state particle information could be used
- To estimate the non-flow contributions and reduce it
- Estimation of identified particle elliptic flow
- Quark and gluon jets

Thank you!