SQM 2024

The 21st International Conference on Strangeness in Quark Matter
3-7 June 2024, Strasbourg, France

Deep learning predicted elliptic flow of identified particles in HIC at the RHIC and LHC

N. Mallick, A.N. Mishra, S. Pasad, R. Sahoo, **G.G. Barnaföldi**

Support: *Hungarian OTKA grants, K135515, NEMZ_KI-2022-00033,*
Wigner Scientific Computing Laboratory

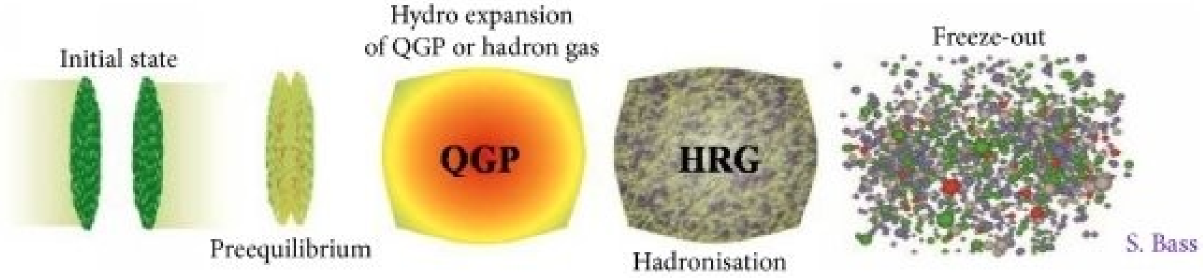
Refs.: PRD 105, 114022 (2022) & PRD 107, 094001(2023)

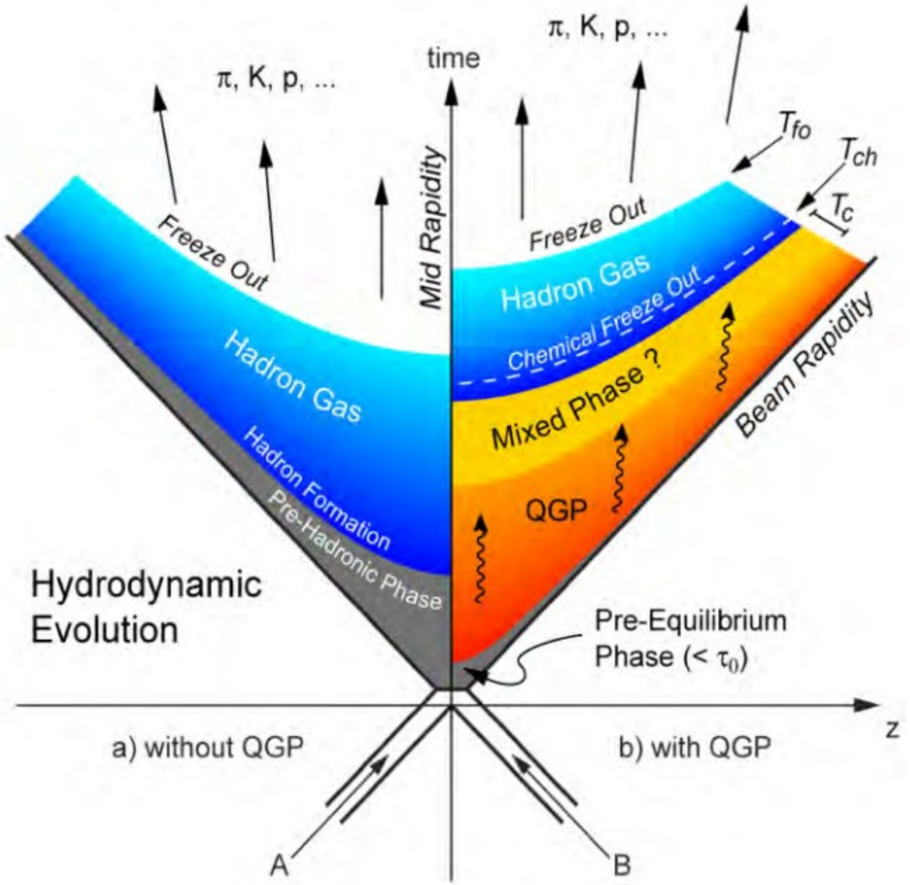
SQM 2024, Strassburg, France, 4th June 2024

Motivation & definitions

Primordial matter in heavy-ion collisions

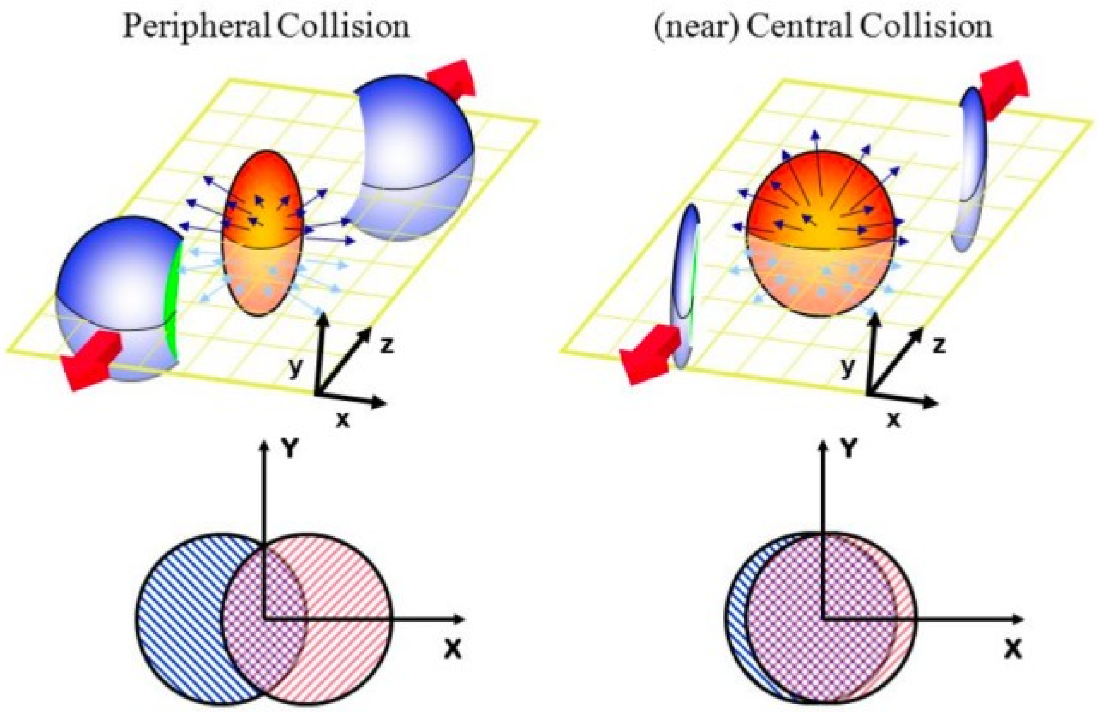
- **QGP in experimental vs theory points**
  - By colliding heavy-ions we can form small drop of the hot & dense primordial matter
  - No direct observations, just **signatures:** jet-quenching, correlations, collective effects, **anisotropic flow**…
  - Need a complex description, including QCD phenomenology, hydrodynamics, (non-equilibrium) thermodynamics





Elliptic flow (v_2) in heavy-ion collisions

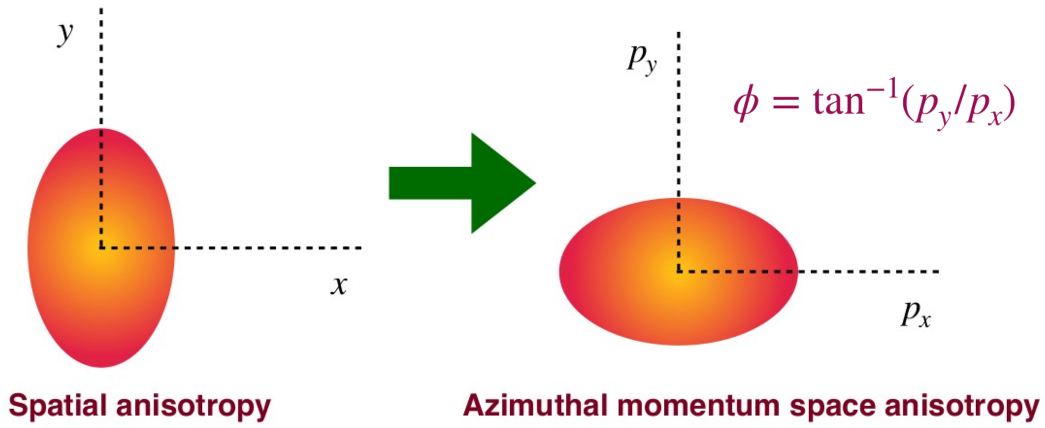
- **Experimental point:**
  - Elliptic flow describes the azimuthal momentum space anisotropy of particle emission for a non-central heavy-ion collision.



Elliptic flow (v_2) in heavy-ion collisions

- **Experimental point:**
  - 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:

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

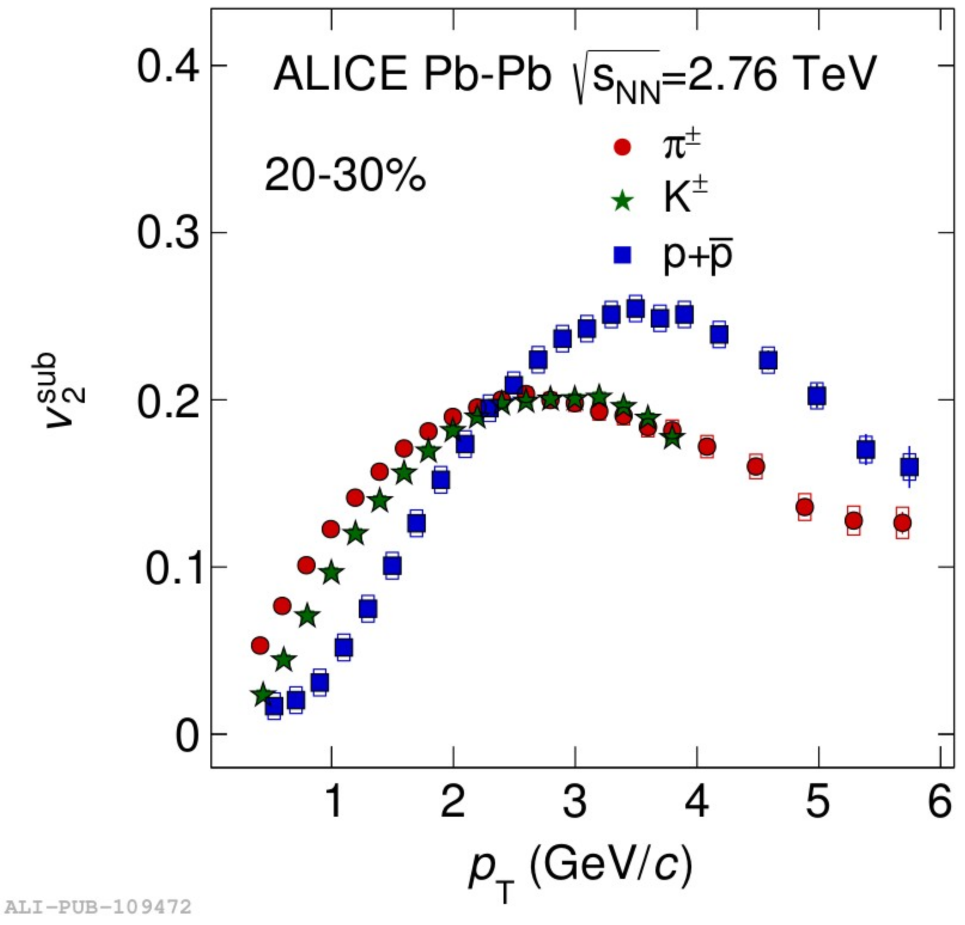


Elliptic flow (v_2) in heavy-ion collisions

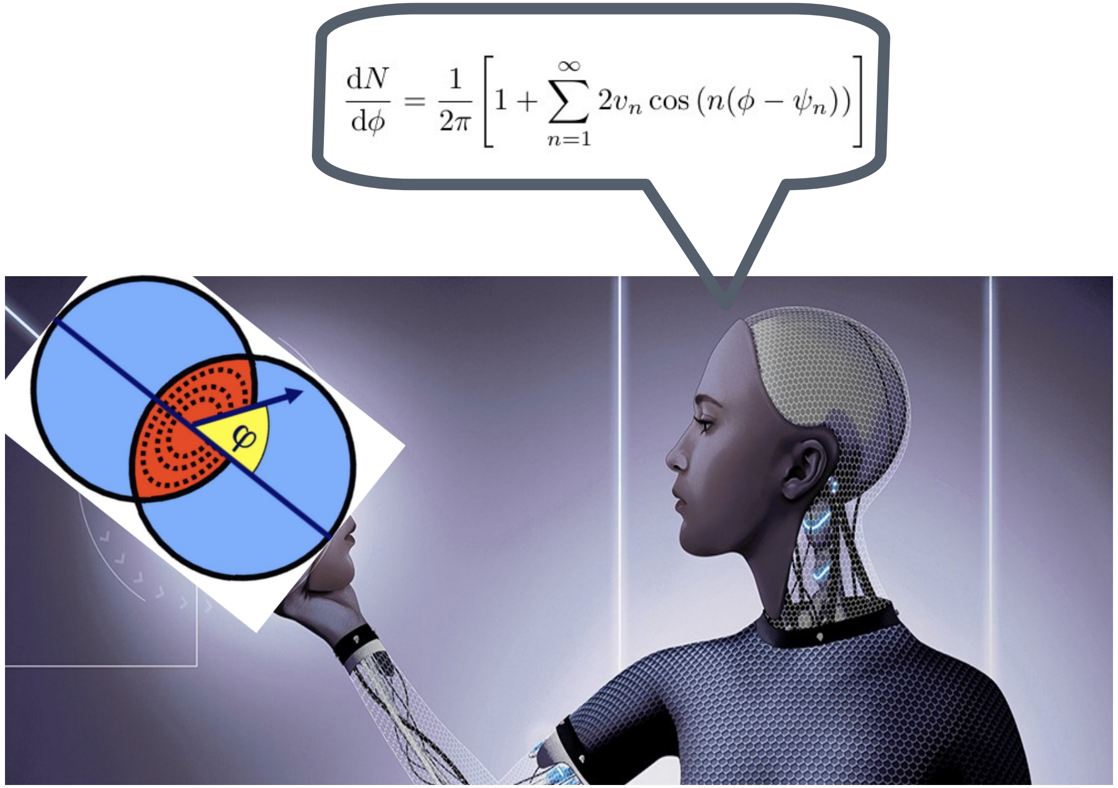
- **Experimental point:**
  - Elliptic flow describes the azimuthal momentum space anisotropy of particle emission for a non-central heavy-ion collision.
  - The 2nd harmonic coefficient of the Fourier expansion of azimuthal momentum distribution:

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

  - The $v_2(p_T, y) = \langle \cos(2(\phi - \psi_2)) \rangle$ directly reflects the initial spatial anisotropy of the nuclear overlap region in the transverse plane.



… and if the situation of calculating the v_2 is getting too problematic, then get…

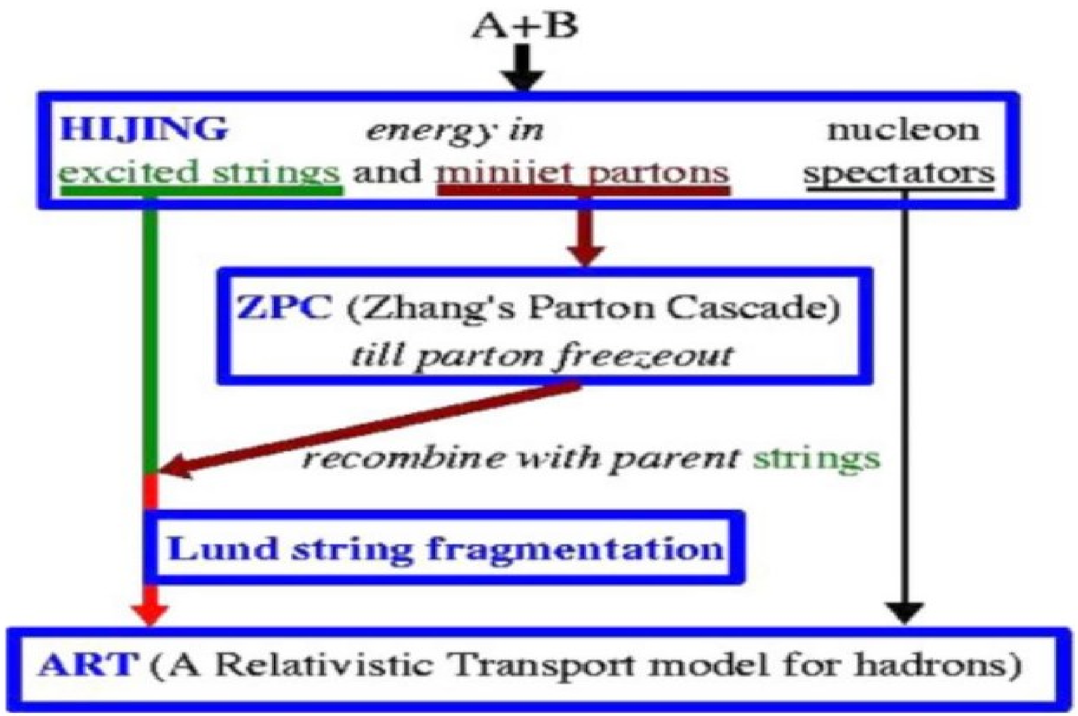


…*ex machina*

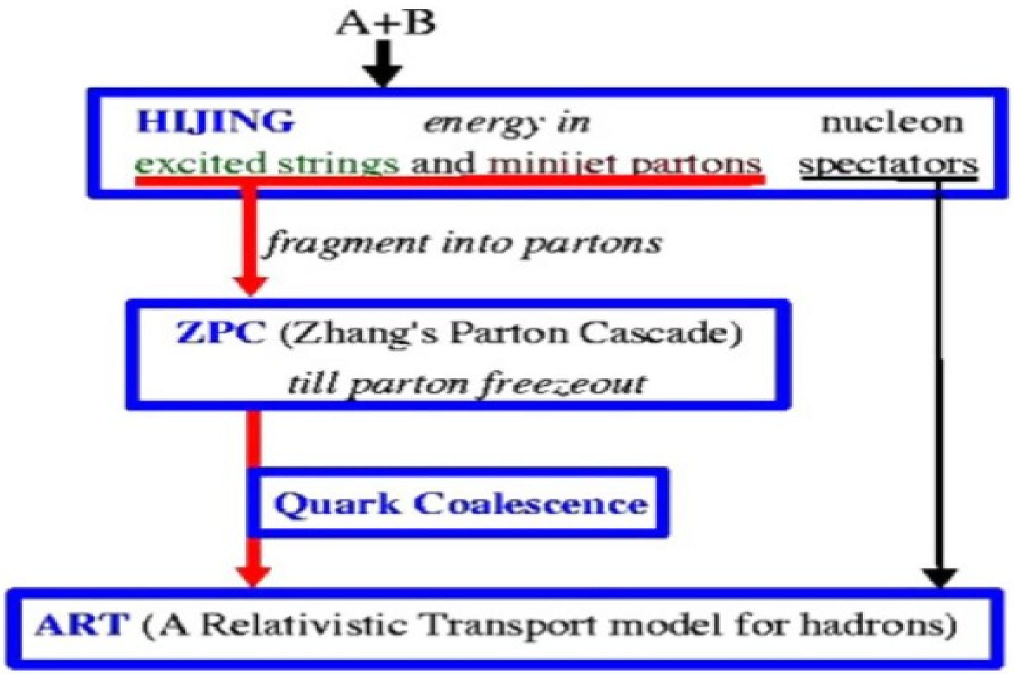
The input: MC-generated collisions

The AMPT model for Pb-Pb collisions

- **A Multi-phase transport model (AMPT):** MC event generator for simulating p-A and A-A collisions from RHIC to LHC energies.



Default Version



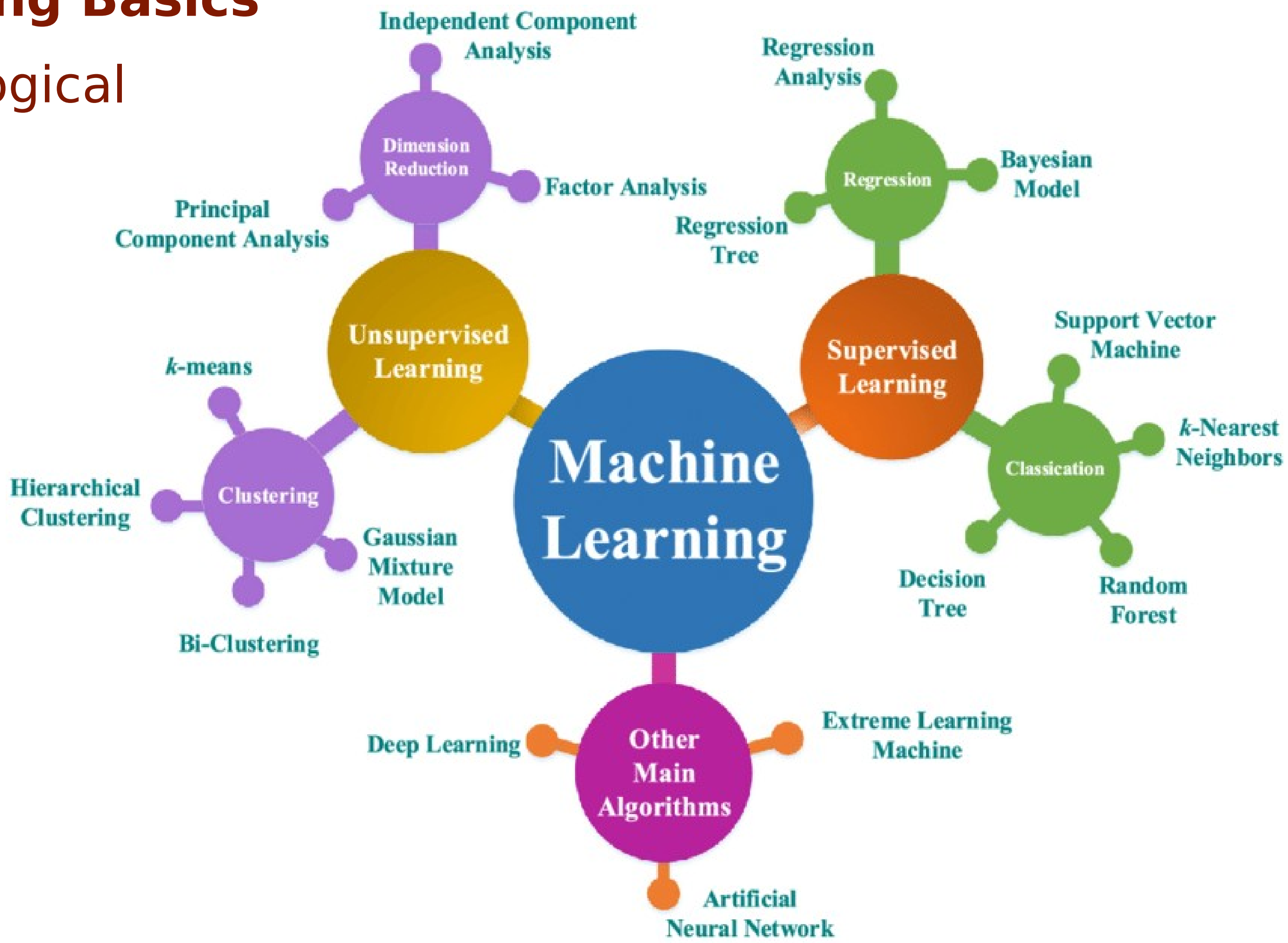
SM Version

Building up the Machine Learning: input, test, and model validation

Building up the ML structure

Machine Learning Basics

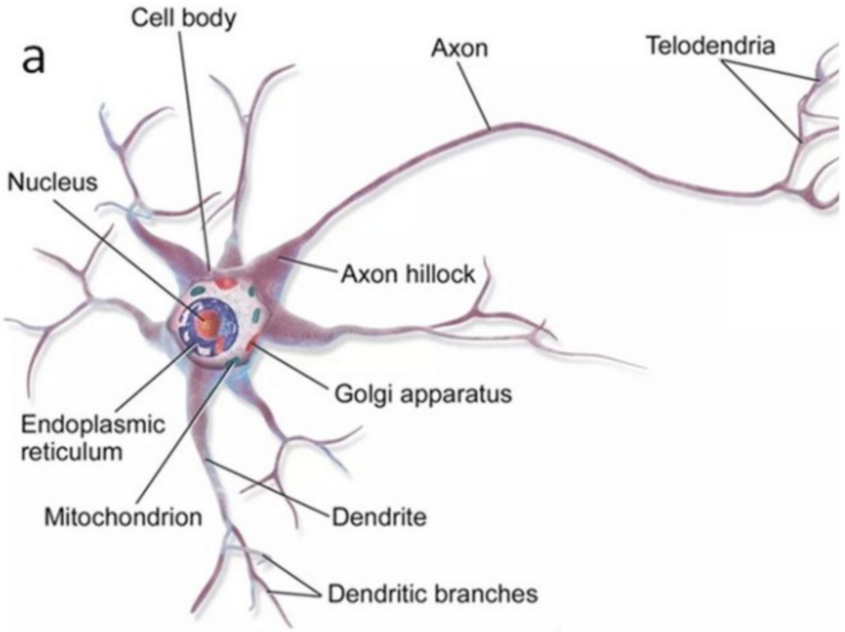
- **Neuron**: Biological



Building up the ML structure

Machine Learning Basics

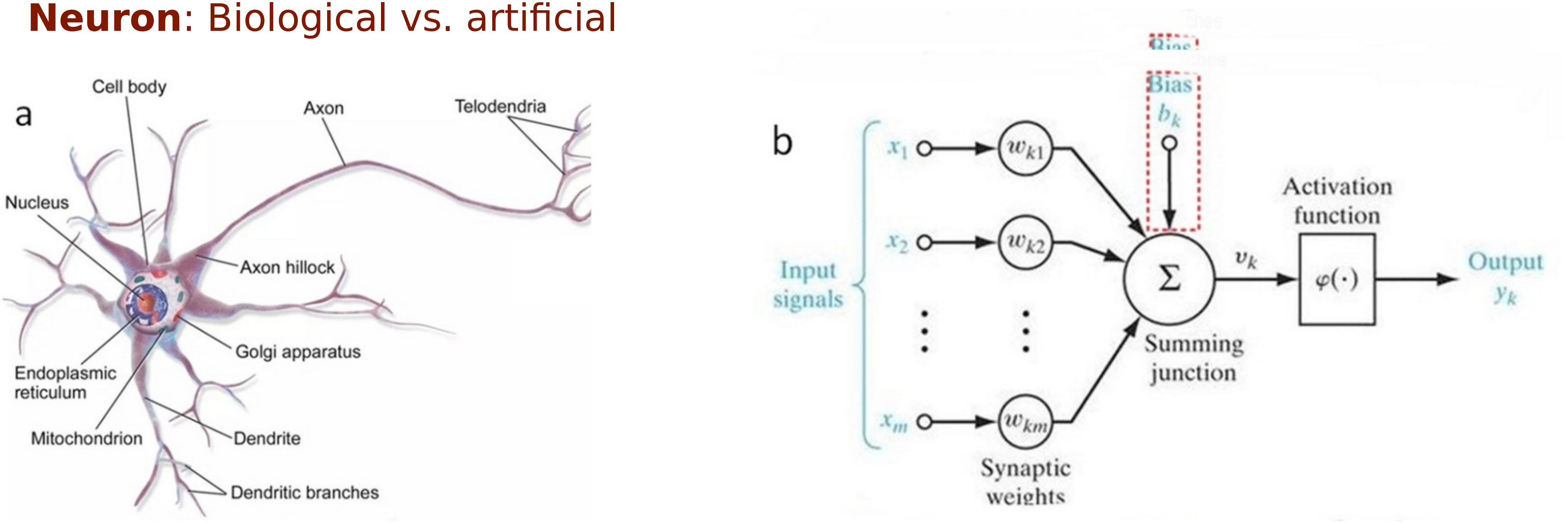
– **Neuron**: Biological



Building up the ML structure

Machine Learning Basics

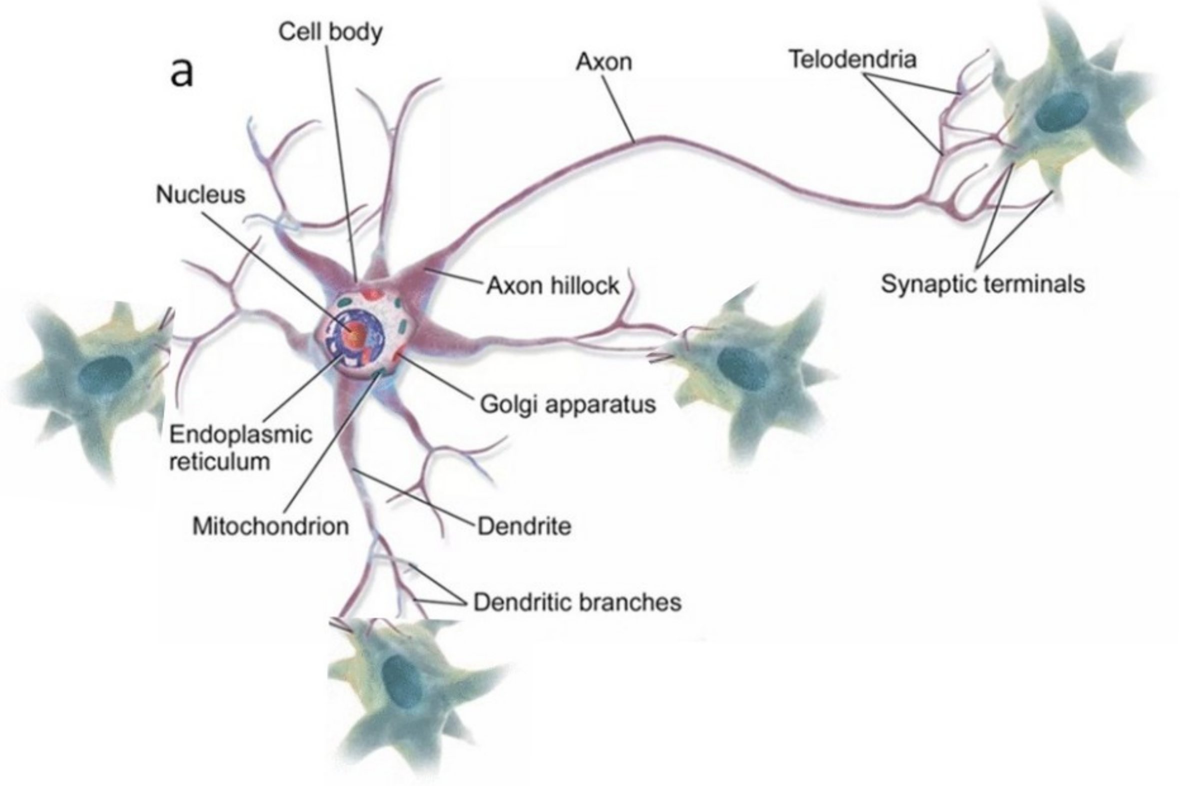
– **Neuron**: Biological vs. artificial

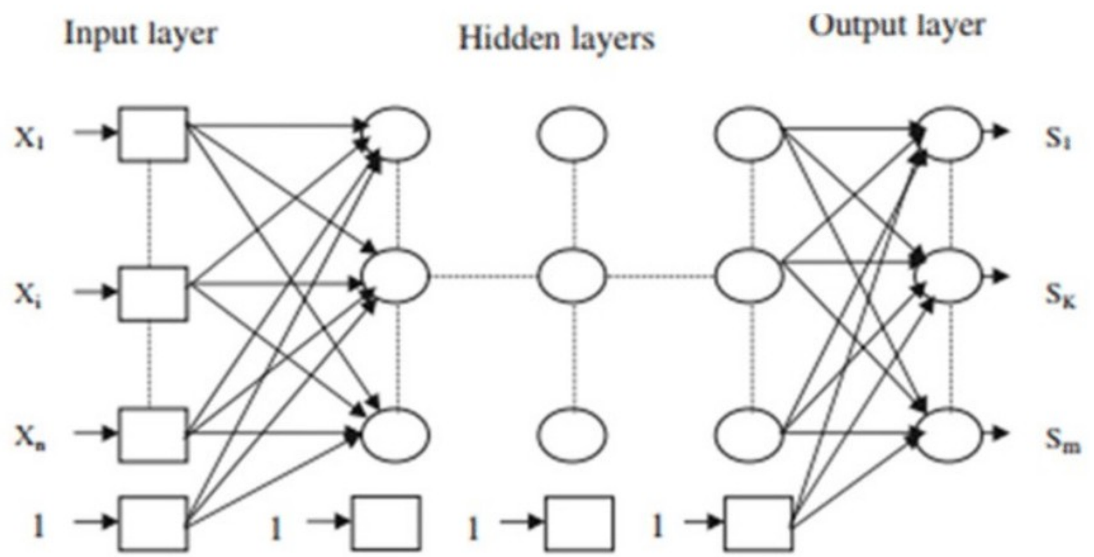


Building up the ML structure

Machine Learning Basics

- **ANN**: Artificial Neural Network





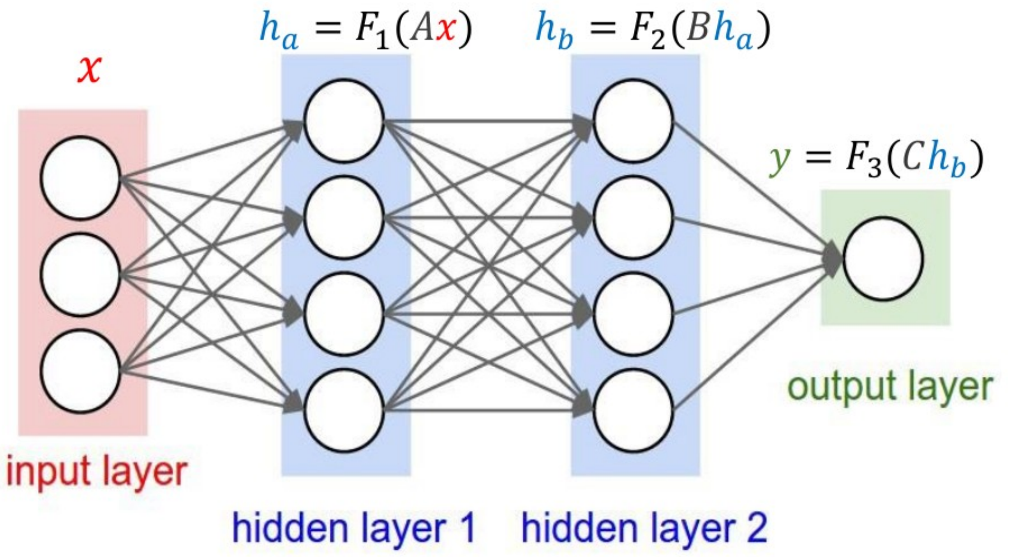
Building up the ML structure

Example: DNN with 2 layers

- **Input**: Takes the features as inputs
- **Hidden layers**: Connects to each neuron through different weights
- **Output**: Gives the result as a number or class

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

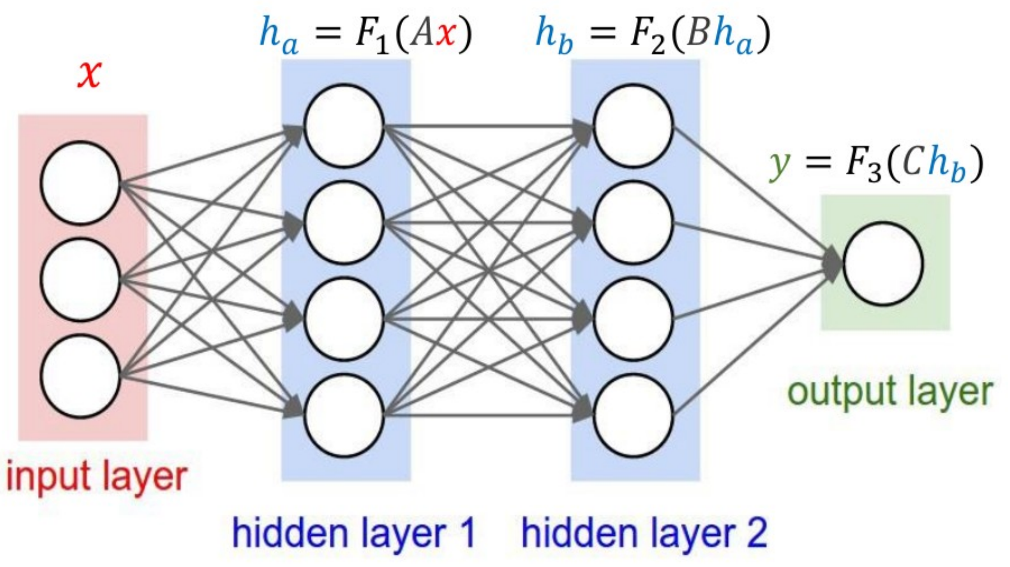
A, B, C represent the weight matrices
F_1, F_2, F_3 represent the activation functions



Building up the ML structure

Math & algorithms behind

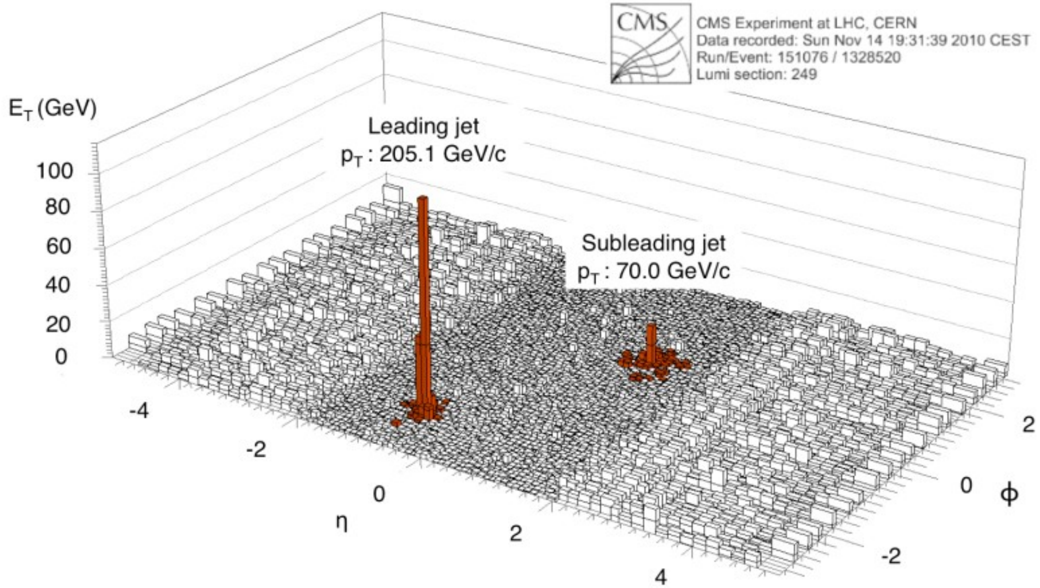
- **Weights** dictate the importance of an input
  → more important features get more weights
- **Activation function**: mathematical function that guides the outcome at each node
  → Standardize the values
- **Cost function**: Evaluates the accuracy between machine prediction and true value
- **Optimizer**: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



Building up the ML structure

Estimation of elliptic flow using DNN

- Elliptic flow → Event property
- Inputs → Track properties
- (η-ϕ) space is the primary input space

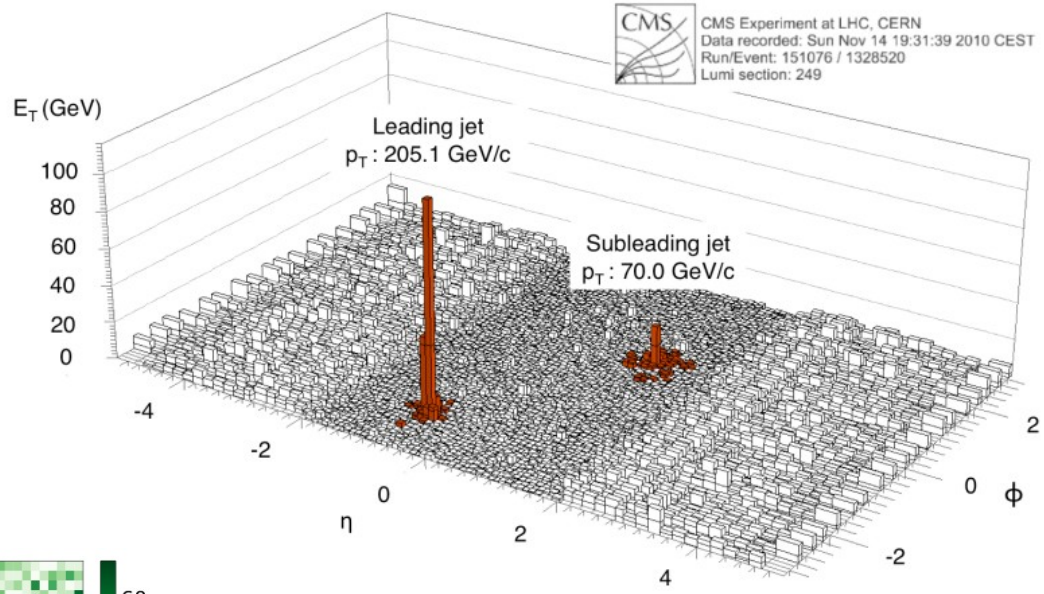


Serguei Chatrchyan et al., Phys.Rev.C 84 (2011), 024906

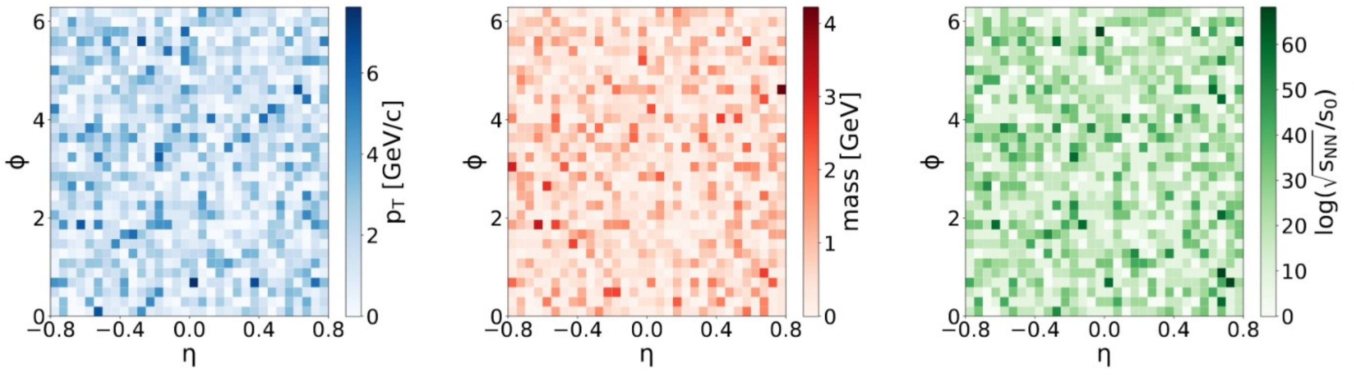
Building up the ML structure

Estimation of elliptic flow using DNN

- Elliptic flow → Event property
- Inputs → Track properties
- (η-ϕ) space is the primary input space
- Three layers having different weights: p_T, mass and $log(s_{NN}/s_0)$ 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**

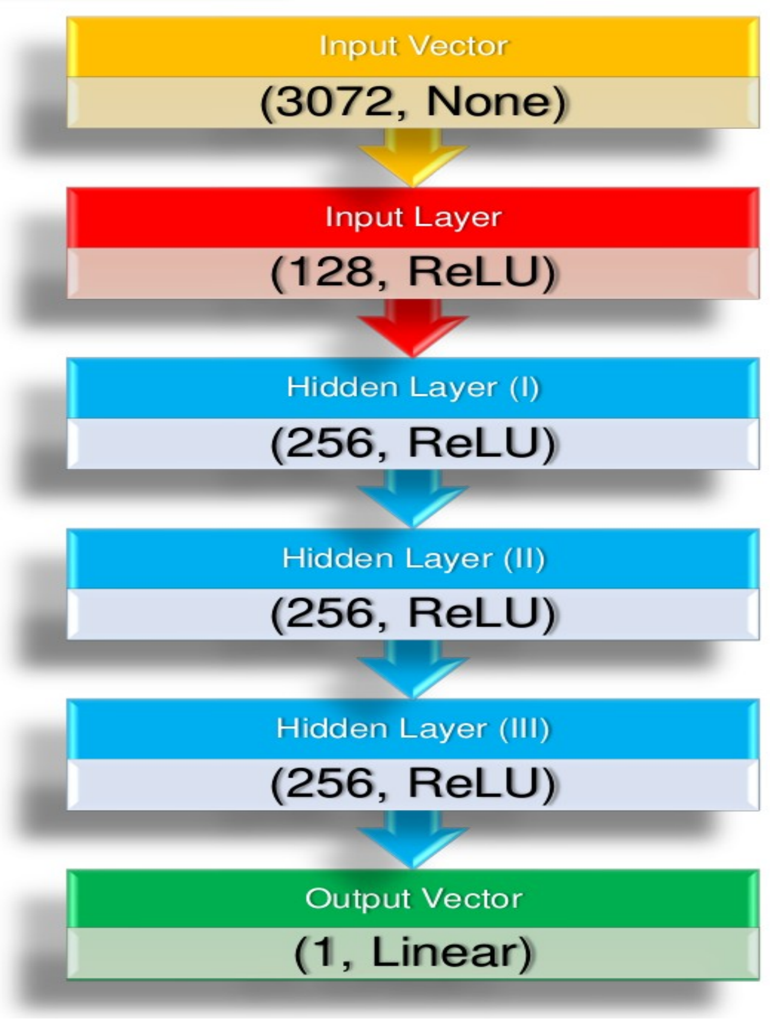
Building up the ML structure

Input "pictures" for DNN

- Each space has 32 × 32 pixels (grids)
- Total number of pixel points =
  32 × 32 × 3 = 3072 for each event

DNN with the following architecture

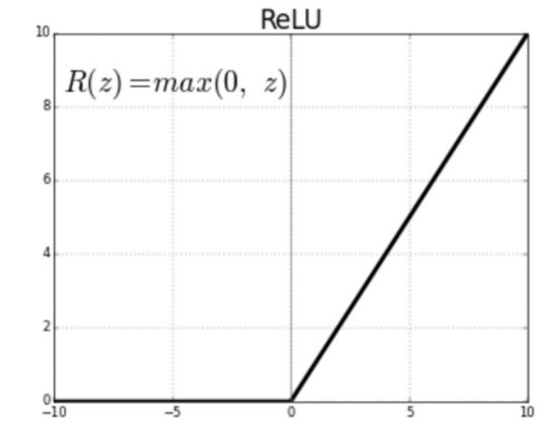
- Input Layer: 128 Nodes
- Three hidden layers: 256 Nodes each
- Final layer : 1 node (v_2)

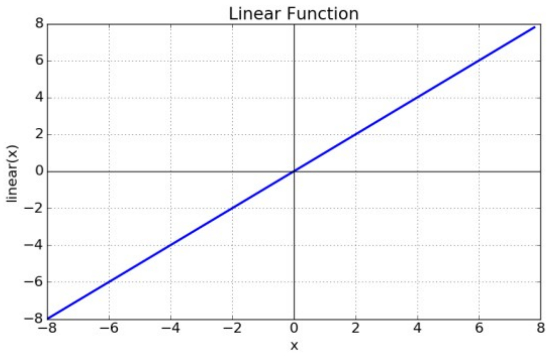


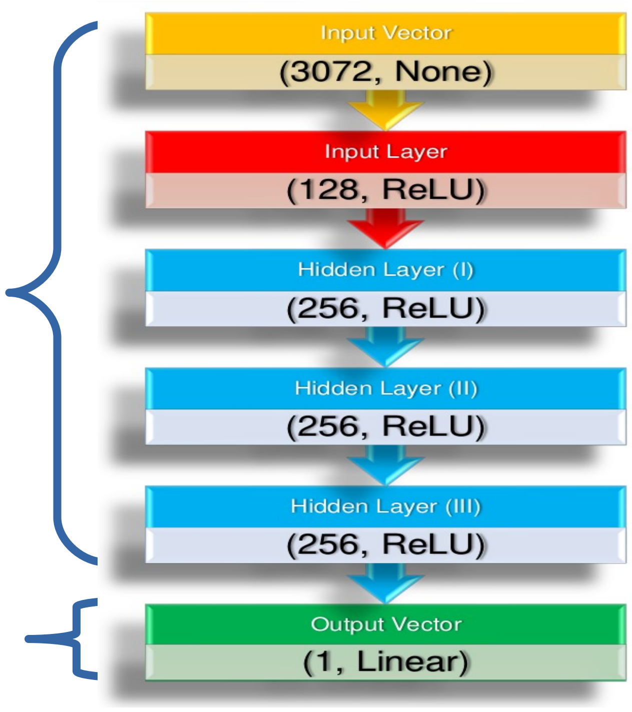
Building up the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse





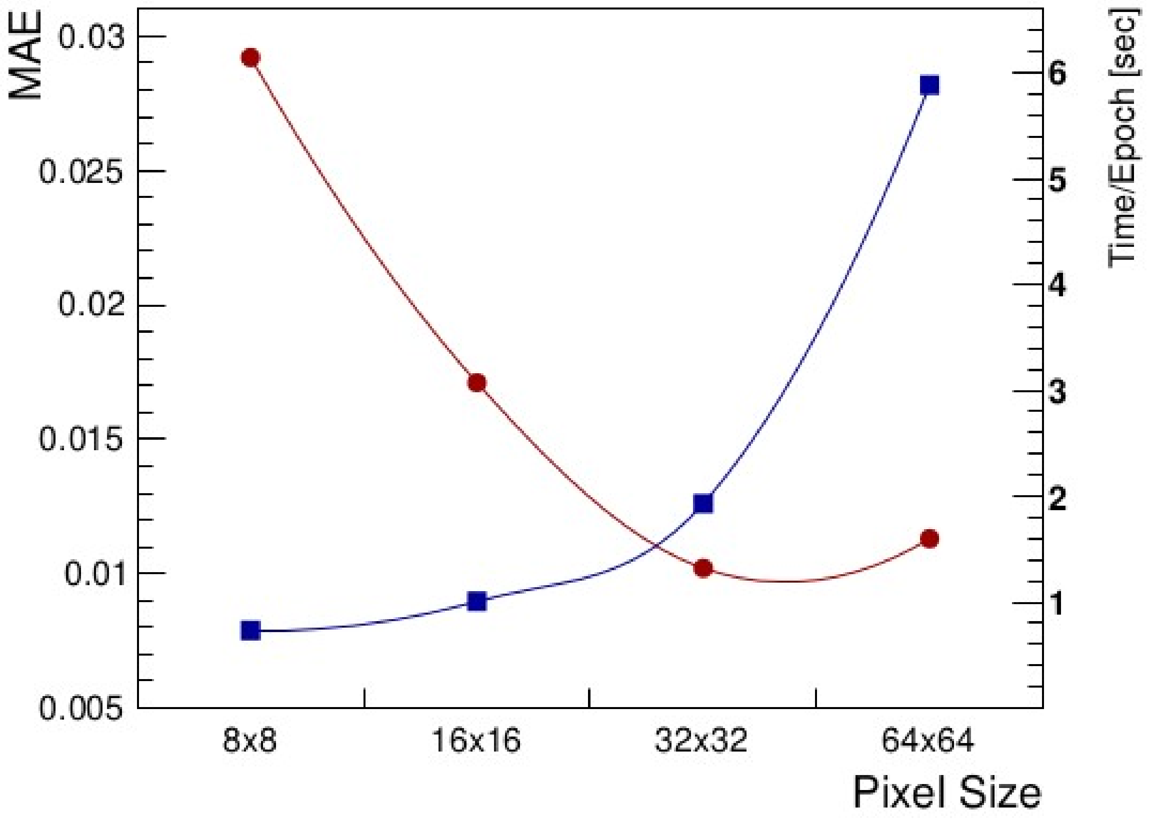


Optimalizing the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse
- Epoch: 30, Batch Size: 32x32
- Training: 10^8 Events (~25 GB)

Bin size	Input neurons	MAE	Epoch	$\frac{\text{Time (sec)}}{\text{Epoch}}$	Trainable parameters
8×8	192	0.0292	18	1.679	189,569
16×16	768	0.0171	28	1.909	263,297
32×32	3072	0.0102	30	2.684	558,209
64×64	12288	0.0113	60	6.001	1,737,857

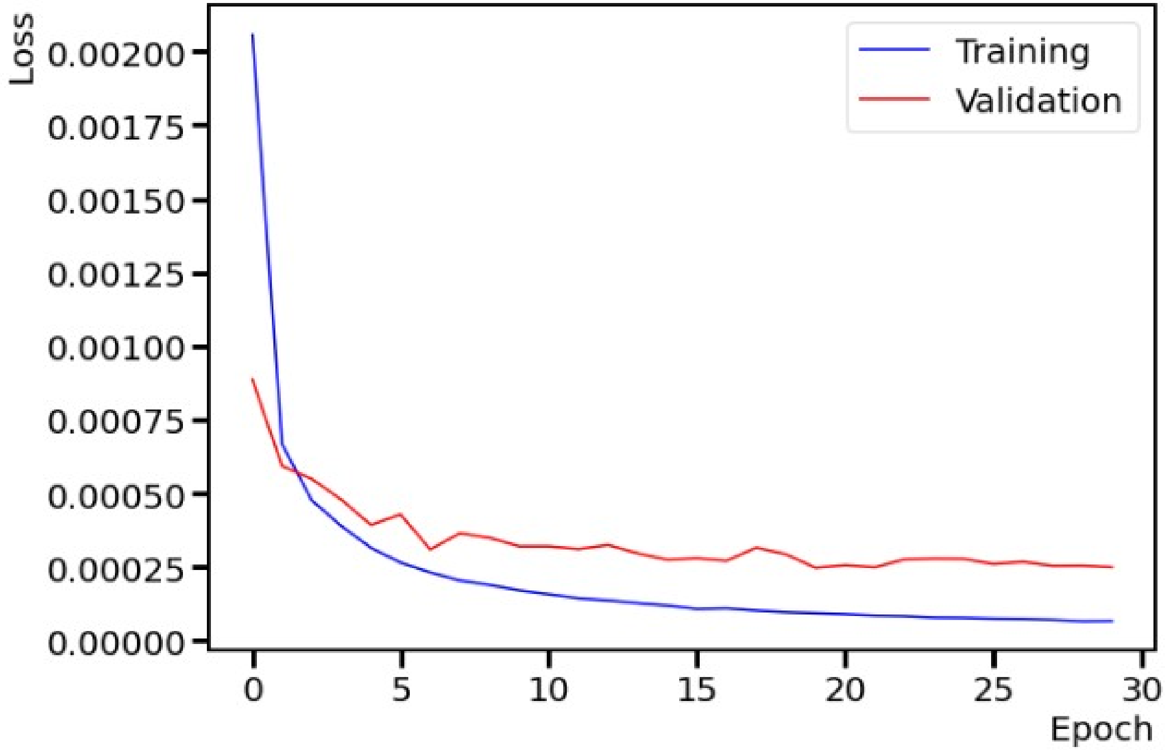


Optimalizing the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse
- Epoch: 30, Batch Size: 32x32
- Training: 10^8 Events (~25 GB)

Bin size	Input neurons	MAE	Epoch	$\frac{\text{Time (sec)}}{\text{Epoch}}$	Trainable parameters
8×8	192	0.0292	18	1.679	189,569
16×16	768	0.0171	28	1.909	263,297
32×32	3072	0.0102	30	2.684	558,209
64×64	12288	0.0113	60	6.001	1,737,857

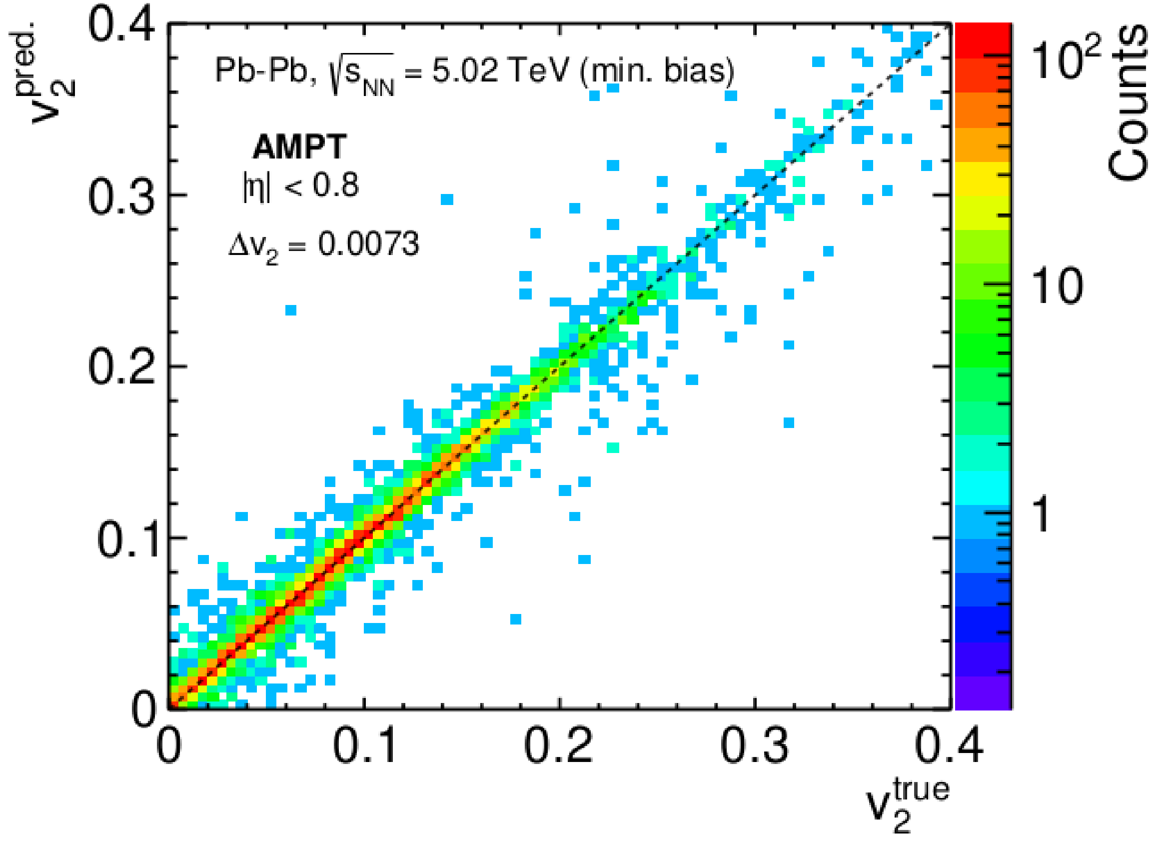


Testing the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse
- Epoch: 30, Batch Size: 32x32
- Training: 10^8 Events (~25 GB)
- Validation: 10^4 Events

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

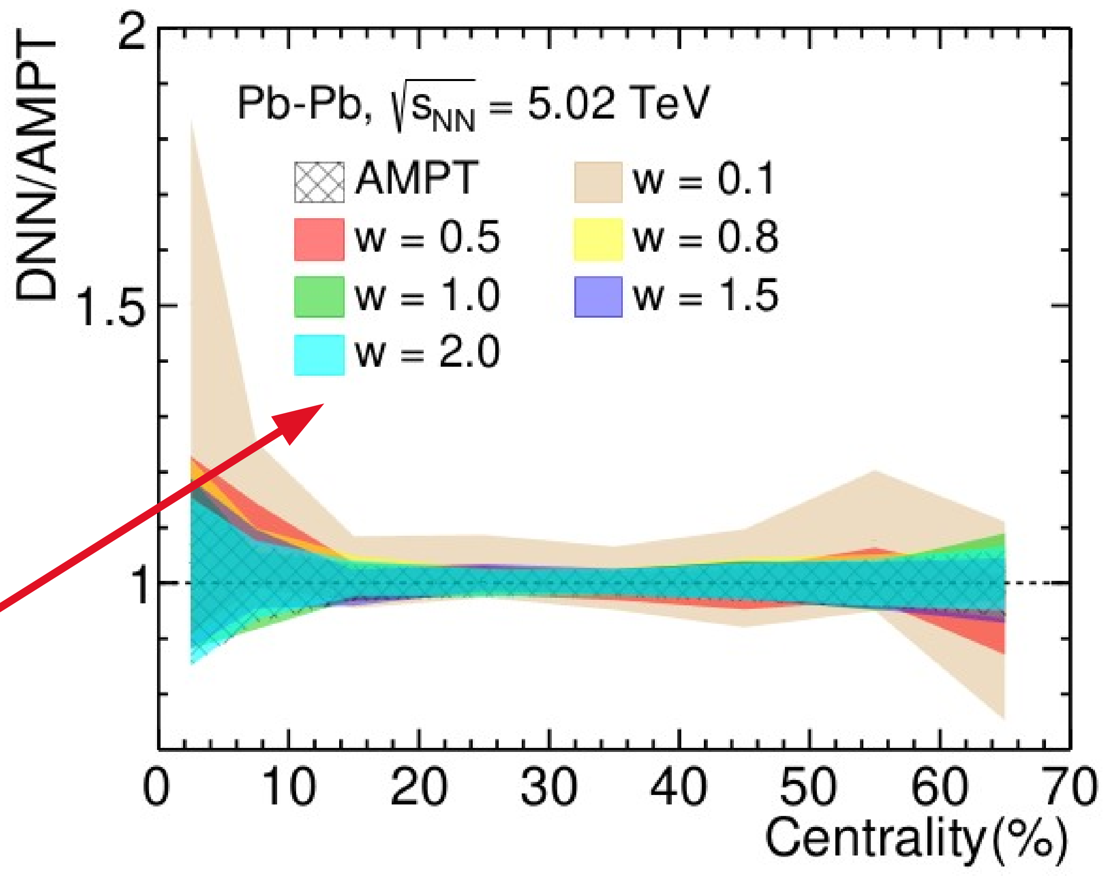


Testing the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse
- Epoch: 30, Batch Size: 32x32
- Training: 10^8 Events (~25 GB)
- Validation: 10^4 Events
- Error: effect of uncorrelated noise

$$F_{i,j} = F_{i,j} + X_{i,j}/\boxed{w}$$



v_2 ex machina

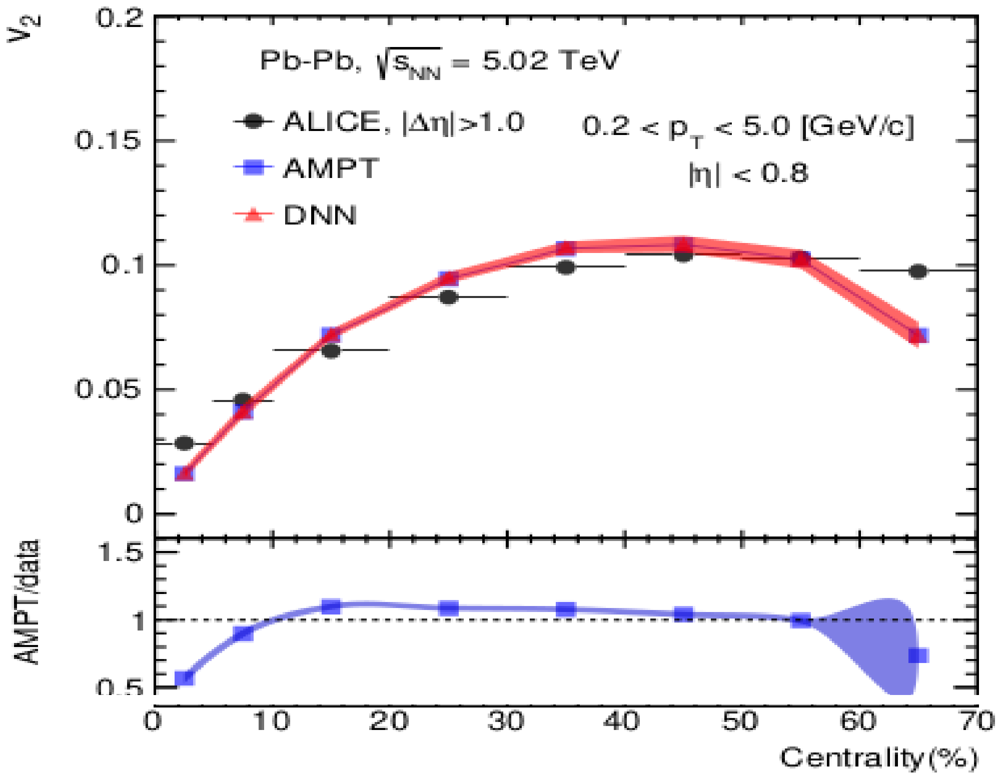
Results on v_2 *vs* centrality

Results on the training data & sets

– **AMPT simulation**: 5.02 TeV Pb-Pb

→ works well [10%:60%] centrality

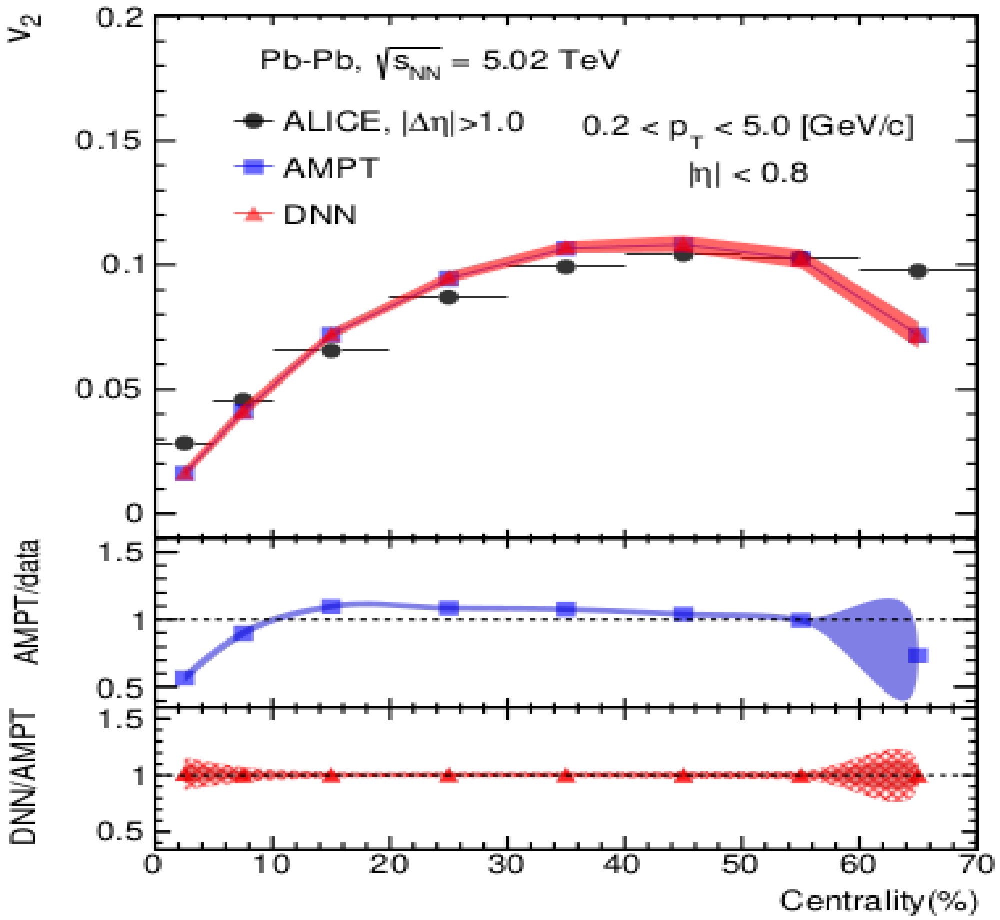
→ low statistics/v_2 values out of this



Results on v_2 *vs* centrality

Results on the training data & sets

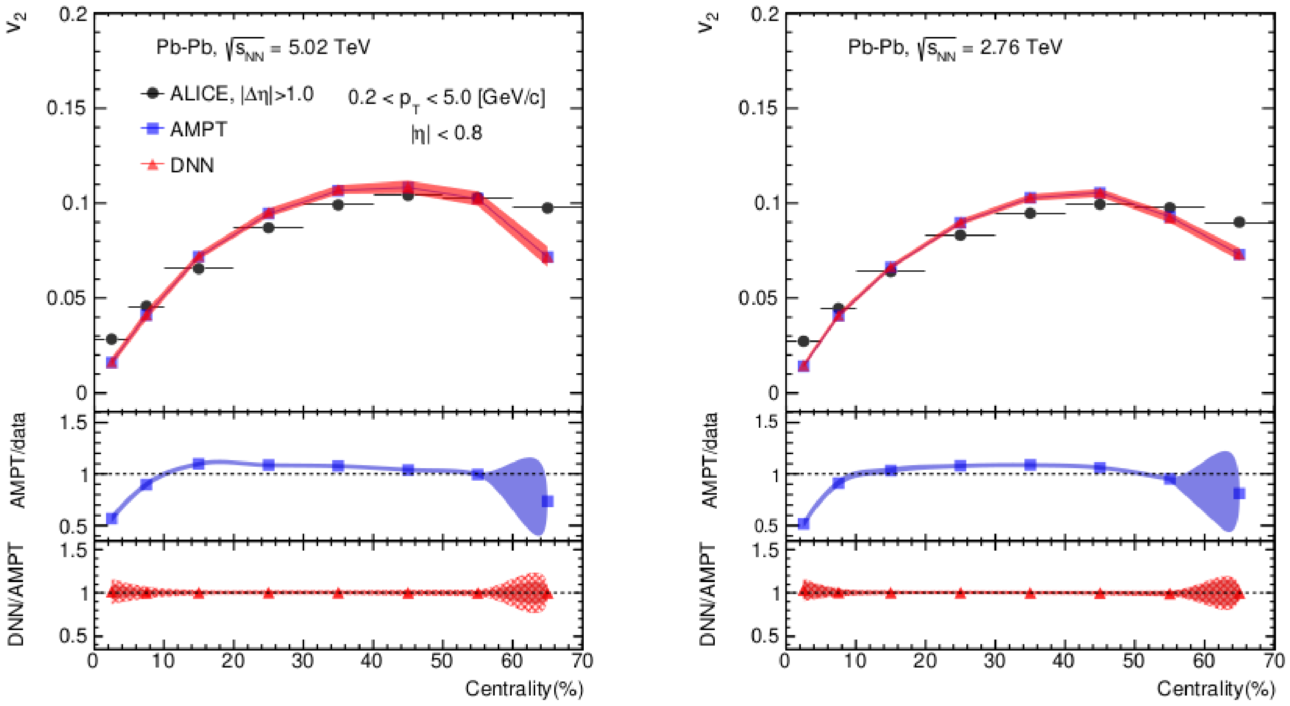
- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well [10%:60%] centrality
  → low statistics/v_2 values out of this
- **DNN simulation**: same parameters
  → Follows well the AMPT
  → Even including noise *w=0.5*



Results on v_2 *vs* c.m. energy

Results on the training data & sets

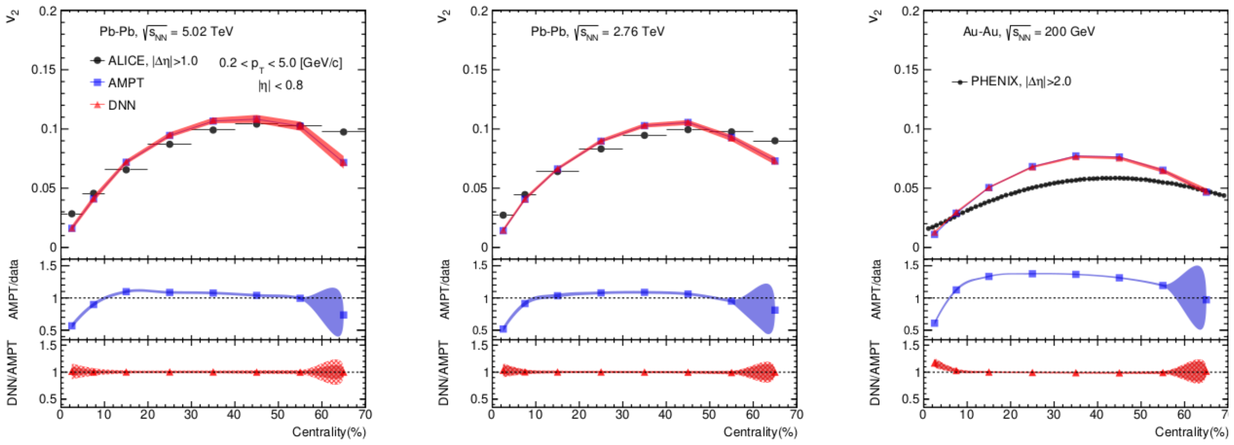
- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well [10%:60%] centrality
  → low statistics/v_2 values out of this
- **DNN simulation**: same parameters
  → Follows well the AMPT
  → Even including noise *w=0.5*
- **Predictions for other energies**
  → similar trends as on the training
  → AMPT tune for 200 GeV is different

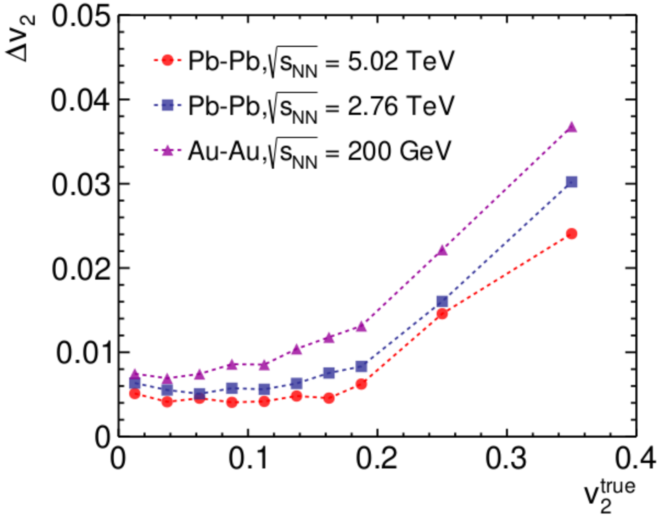


Results on v_2 *vs* c.m. energy

Results on the training data & sets

- **AMPT simulation**: 5.02 TeV Pb-Pb
  - → works well [10%:60%] centrality
  - → low statistics/v_2 values out of this
- **DNN simulation**: same parameters
  - → Follows well the AMPT
  - → Even including noise *w=0.5*
- **Predictions for other energies**
  - → similar trends as on the training
  - → AMPT tune for 200 GeV is different



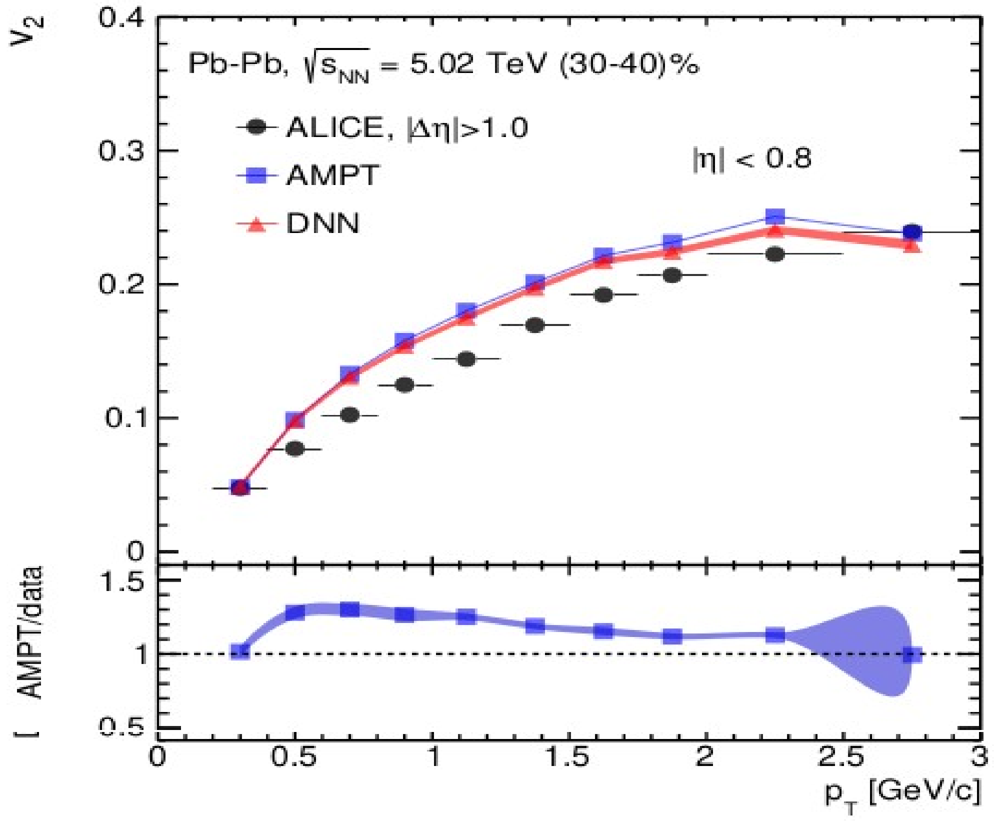


$$\Delta v_2 = \frac{1}{N_{\text{events}}} \sum_{n=1}^{N_{\text{events}}} |v_{2_n}^{\text{true}} - v_{2_n}^{\text{pred.}}|$$

Results on v_2 vs p_T

Results on the training data & sets

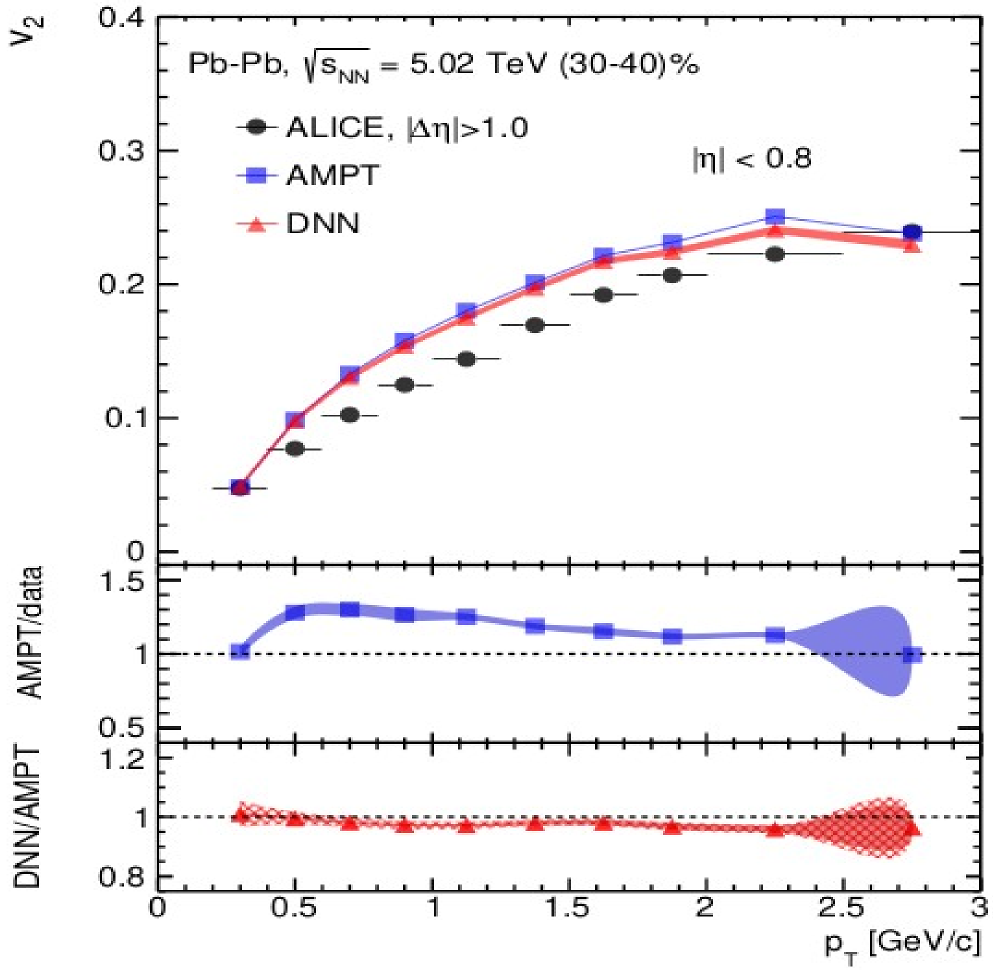
- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well at 30%-40% centrality
  → low statistics at high p_T



Results on v_2 vs p_T

Results on the training data & sets

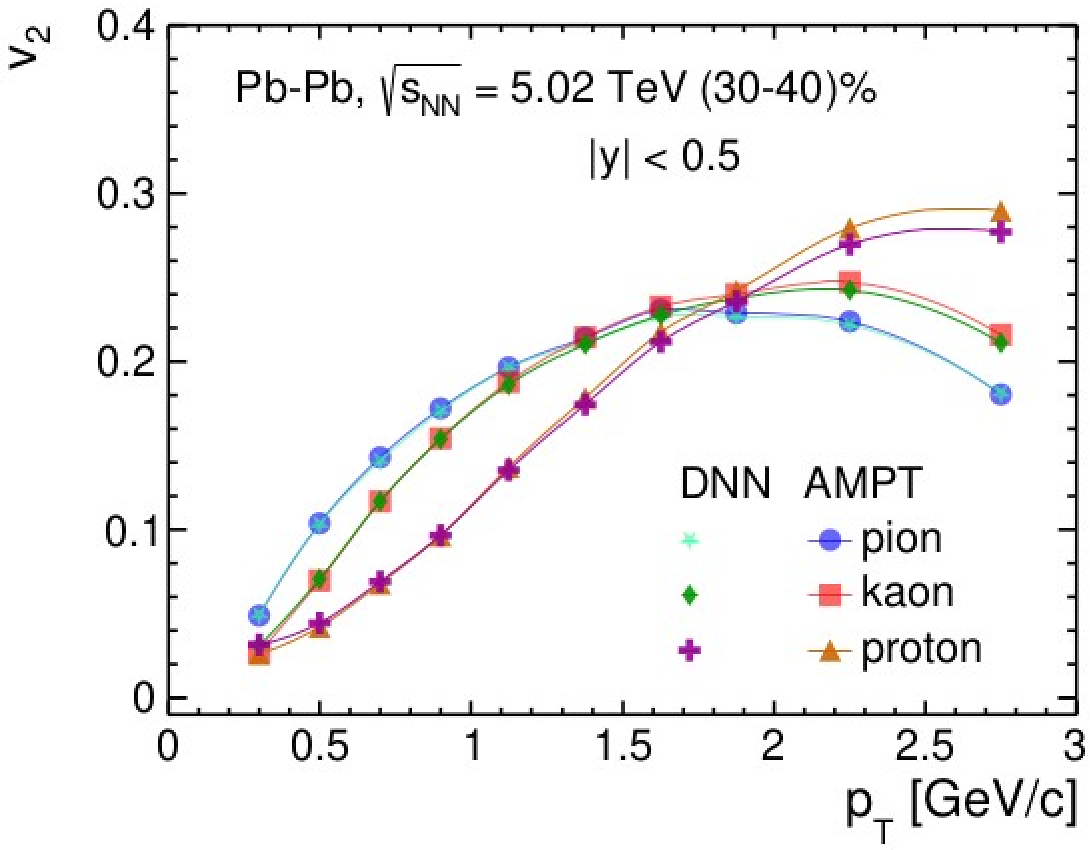
- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well at 30%-40% centrality
  → low statistics at high p_T
- **DNN simulation**: same parameters
  → Follows well the AMPT
  → Even including noise *w=0.5*



Results on v_2 vs p_T

Results on the training data & sets

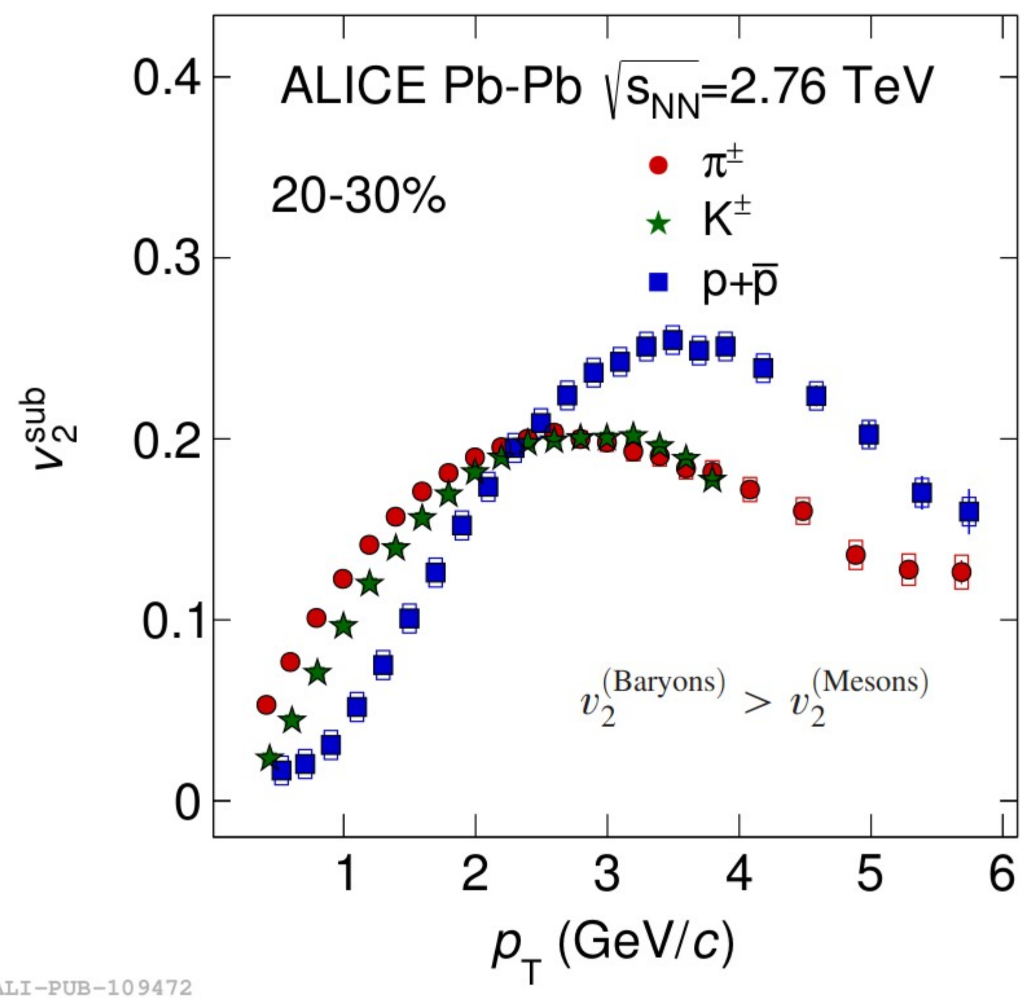
- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well at 30%-40% centrality
  → low statistics at high p_T
- **DNN simulation**: same parameters
  → Follows well the AMPT
  → Even including noise *w=0.5*
- **PID dependency** $v_2^{\pi^\pm} > v_2^{K^\pm} > v_2^{p+\bar{p}}$
  → DNN/AMPT=1 satisfied at low p_T
  → Interesting feature: turning point
  → Coalescence scaling



Results on v_2 vs p_T

Results on the training data & sets

- **AMPT simulation**: 5.02 TeV Pb-Pb
  → works well at 30%-40% centrality
  → low statistics at high p_T
- **DNN simulation**: same parameters
  → Follows well the AMPT
  → Even including noise *w=0.5*
- **PID dependency** $v_2^{\pi^\pm} > v_2^{K^\pm} > v_2^{p+\bar{p}}$
  → DNN/AMPT=1 satisfied at low p_T
  → Interesting feature: turning point
  → Coalescence scaling



Scaling properties with PID

- **Centrality**

  The largest in case 30%-40% mid-central

- **Collision energy**

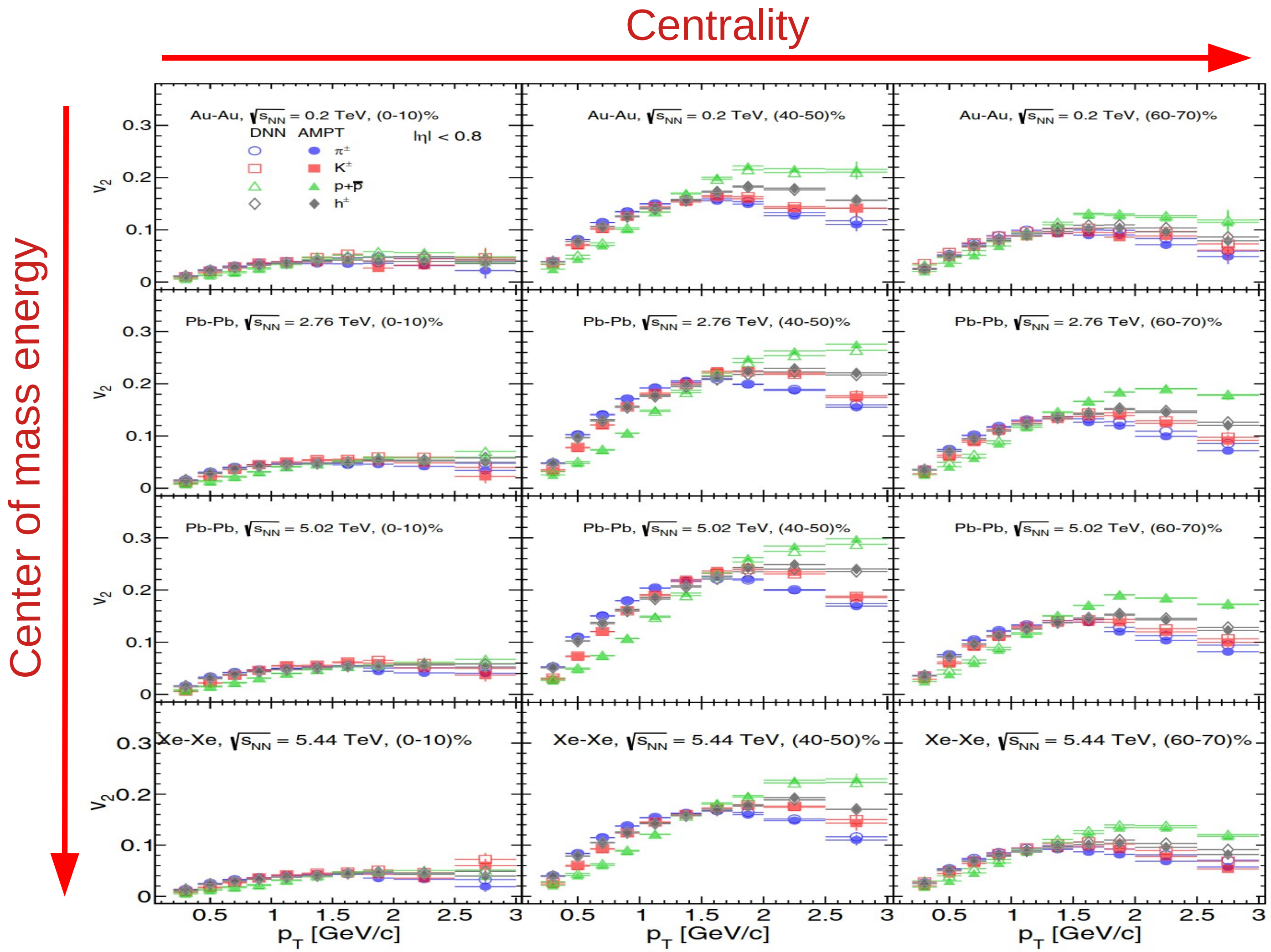
  The higher the energy higher effect.

- **System size**

  AuAu, PbPb, XeXe

- **DNN**

  Follows well the trends → scaling is encoded.



Scaling properties with PID

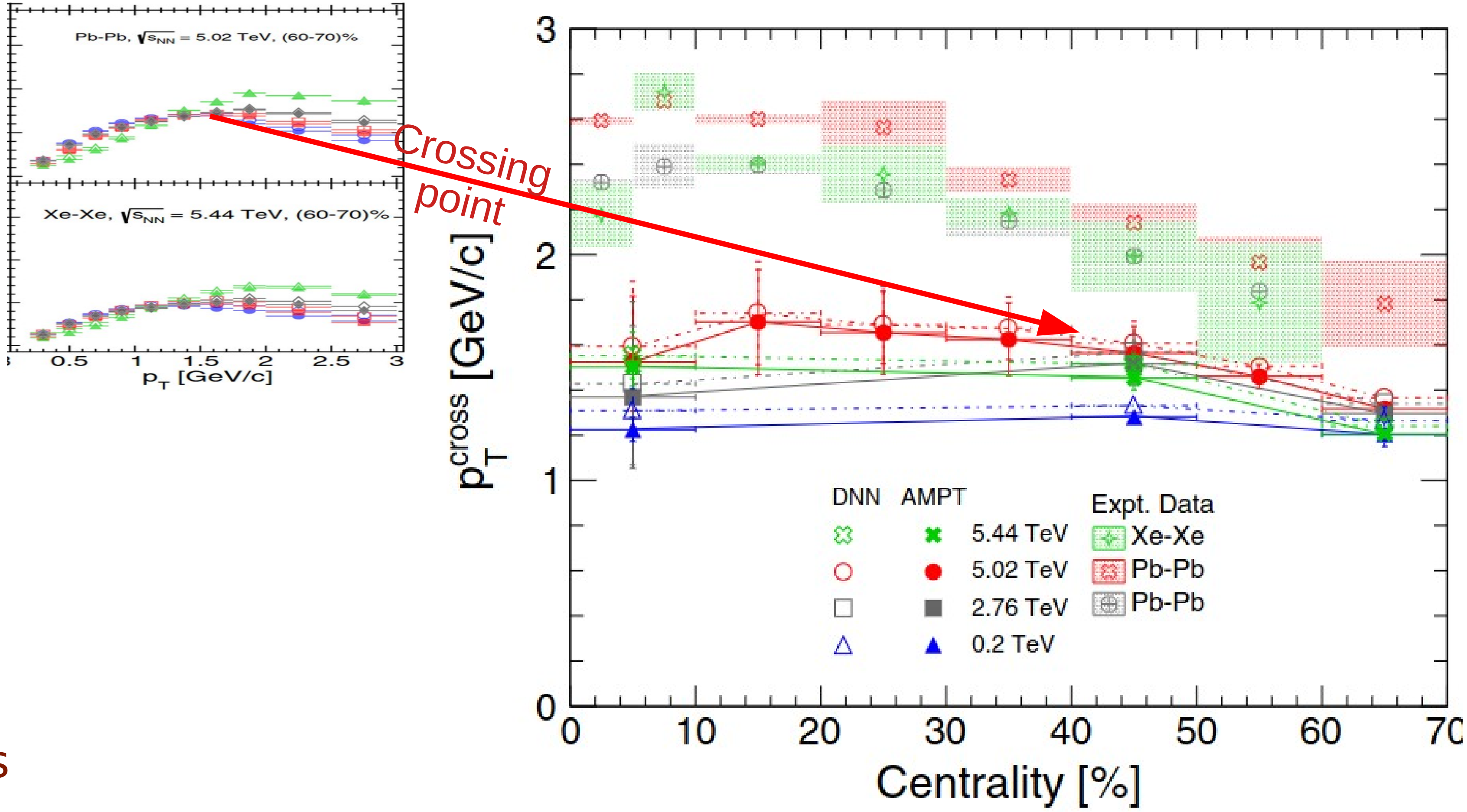
- **Particle ID (PID)**

  The highest crossing point in pT appears at the highest energies in case 10%-40%.

  The measured values are larger than the AMPT

- **DNN**

  Follows well the trends in the training AMPT simulations → scaling is encoded.



Scaling properties with PID

- **Centrality**

  The largest in case 30%-40% mid-central

- **Collision energy**

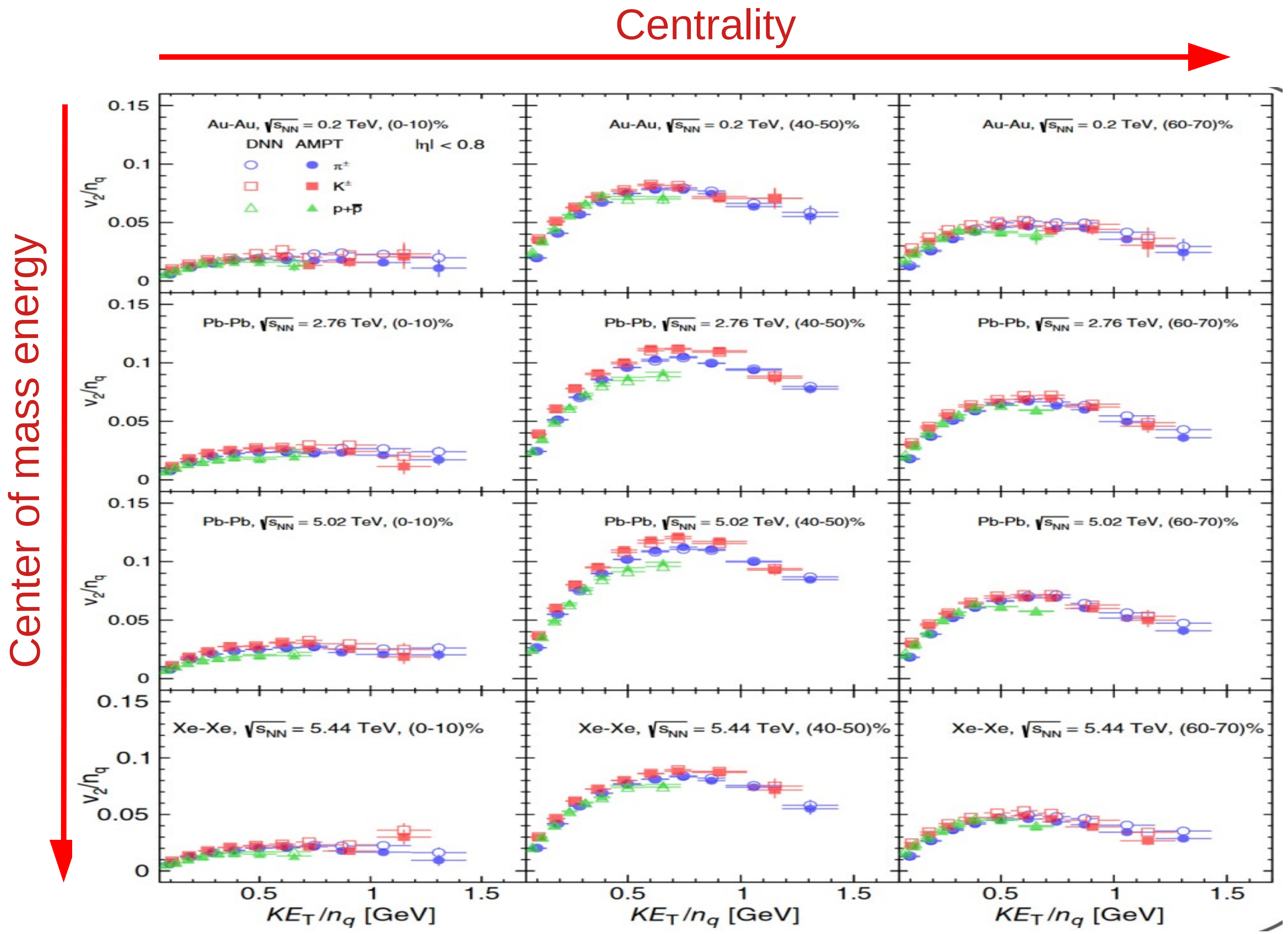
  The higher the energy higher effect.

- **System size**

  AuAu, PbPb, XeXe

- **DNN**

  Follows well the trends → scaling is encoded.



Where can we improve v_2 *ex machina* ?

Challenge #1: Results at higher p_T

- **AMPT vs Data**

  → Does not fit well above than few p_T Best at 30%-40% mid-central.

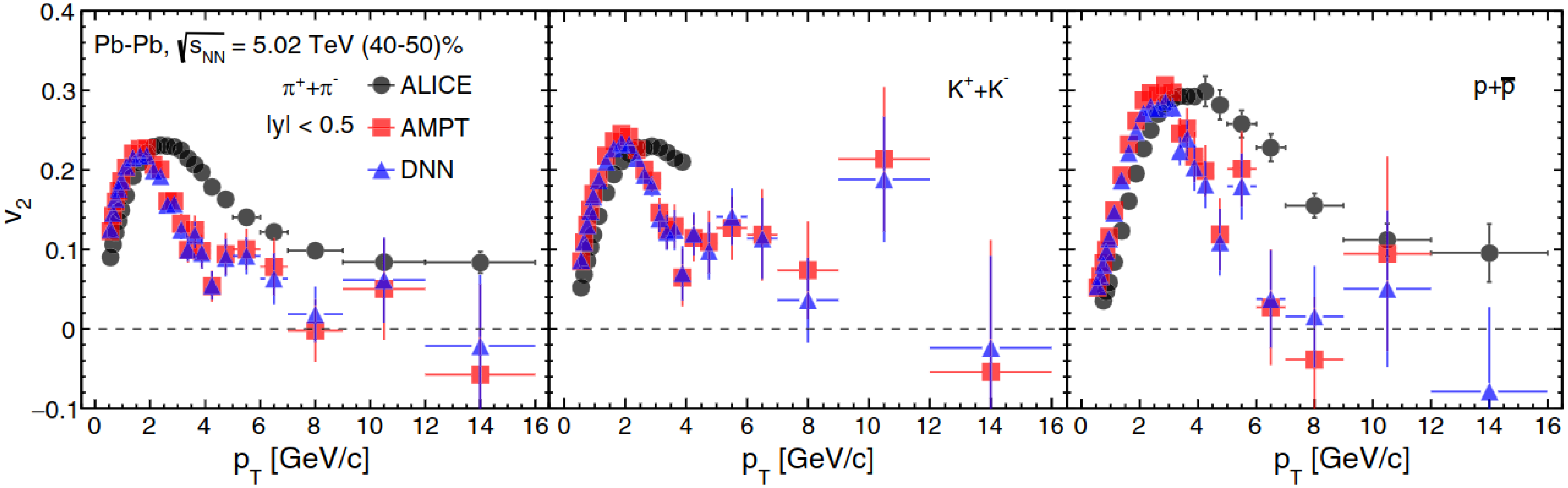
  → Need for more statistics

- **AMPT vs DNN**

  → DNN follows well the AMPT at any energy & centrality.

- **DNN**

  Follows well AMPT but NOT the high p_T data → need to improve!



$$v_2^{\pi^\pm} > v_2^{K^\pm} > v_2^{p+\bar{p}} \qquad v_2^{(\text{Baryons})} > v_2^{(\text{Mesons})}$$

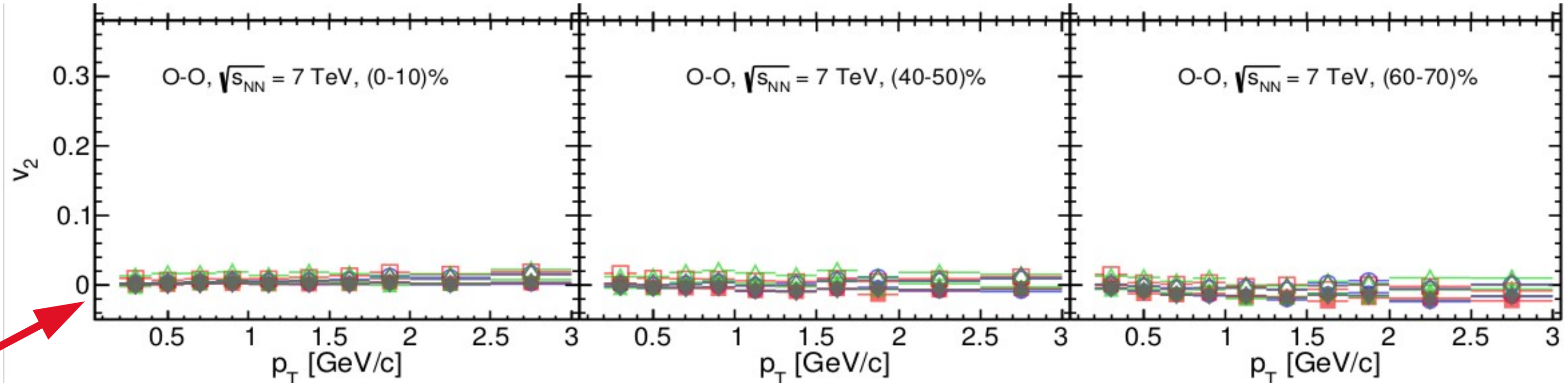
Challenge #2: Small Systems OO

- **Tiny System**

  In OO the event plane determination is a problematic → all flat

- **DNN**

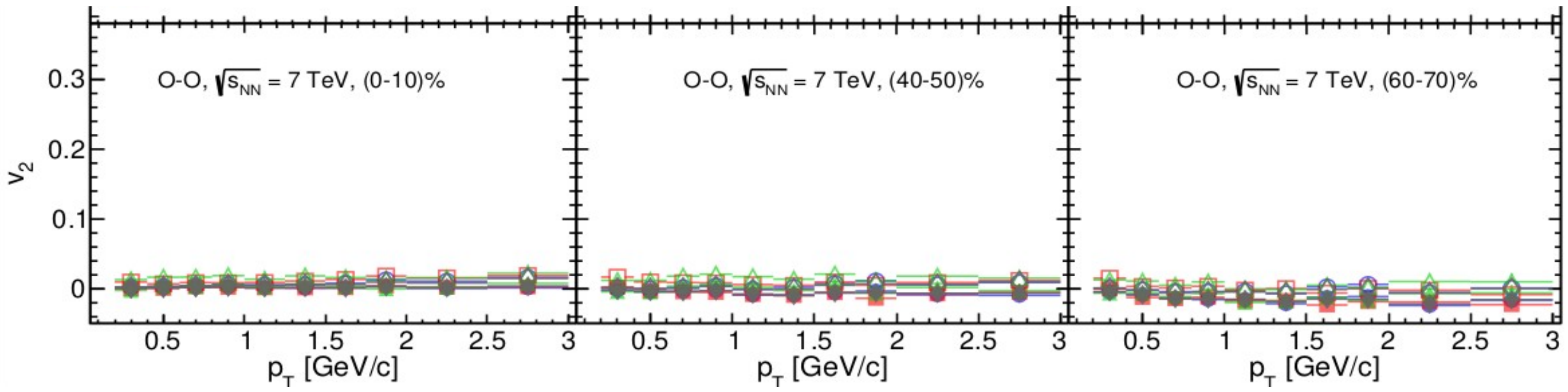
  Follows the fake trend → Fake scaling well :)



Challenge #2: Small Systems OO

- **Tiny System**

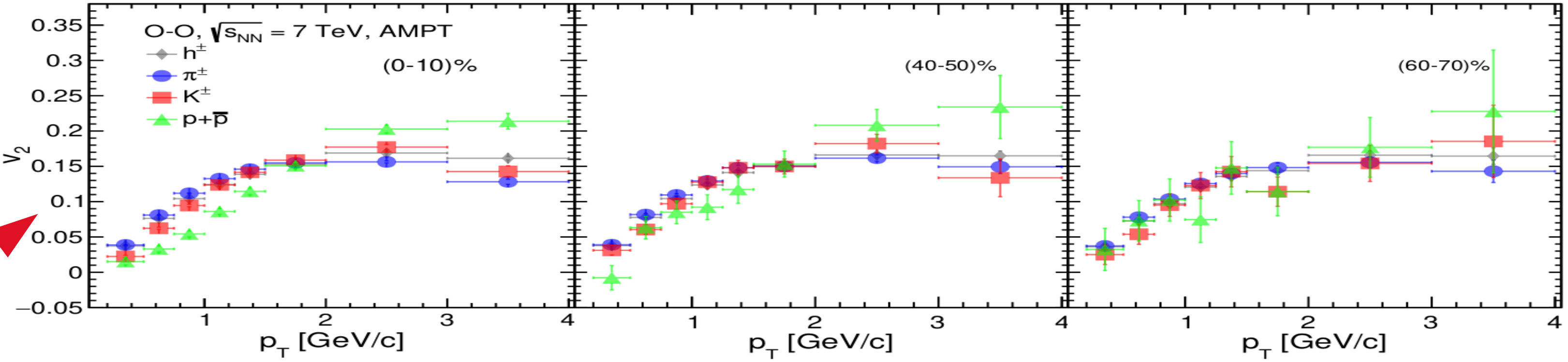
  In OO the event plane determination is a problematic → all flat



- **DNN**

  Follows the fake trend → Fake scaling well :)

- Other method gives correct v_2



Conclusions

- **Is it possible to estimate the elliptic flow by ML?**
  - Get best Min. Bias. Monte Carlo simulation data and train the well-designed DNN system…
    - → More sophisticated NN, the less epoch needs
    - → Un-correlated noise can be even $w=1$
    - → AMPT & DNN correlates well for all centrality
    - → Best correlation is for the highest statistic
    - → Energy scaling is well preserved (non-linear)
    - → The $v_2(p_T)$ is also preserved with PID & NCQ

Conclusions

- **Is it possible to estimate the elliptic flow by ML?**
  - Get best Min. Bias. Monte Carlo simulation data and train the well-designed DNN system…
    - → More sophisticated NN, the less epoch needs
    - → Un-correlated noise can be even *w=1*
    - → AMPT & DNN correlates well for all centrality
    - → Best correlation is for the highest statistic
    - → Energy scaling is well preserved (non-linear)
    - → The $v_2(p_T)$ is also preserved with PID & NCQ
- **What is missing…**
  - Test of correlated noise (detector setup, etc)
  - Train with real data (for high-p_T from ALICE)

Thank You!

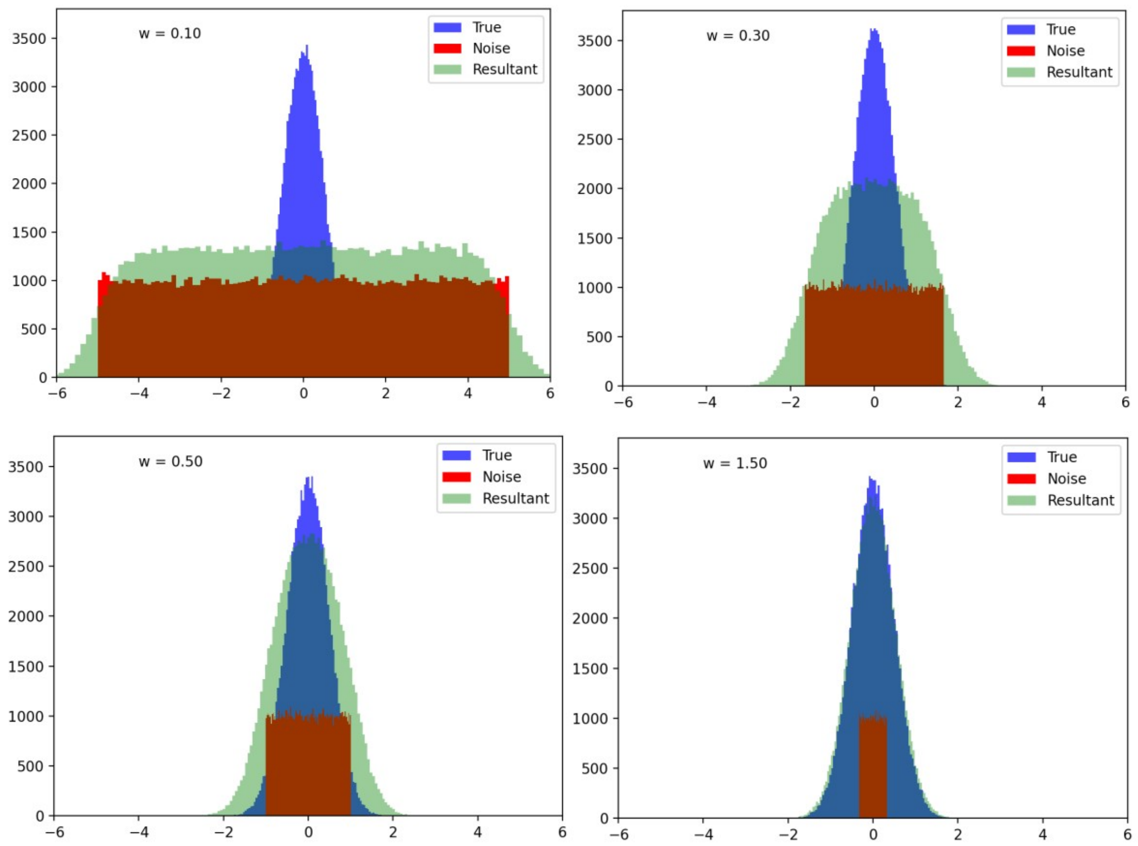
BACKUP

Testing the ML structure

Activation, optimalization, validation

- Input and hidden layers have ReLu Activation
- Output layer has Linear activation
- Optimizer: adam , Loss function: mse
- Epoch: 30, Batch Size: 32x32
- Training: 10^8 Events (~25 GB)
- Validation: 10^4 Events
- Error: effect of uncorrelated noise

$$F_{i,j} = F_{i,j} + X_{i,j}/w$$



Scaling properties are in the DNN

- **Centrality**

  The largest in case 30%-40% mid-central

- **Collision energy**

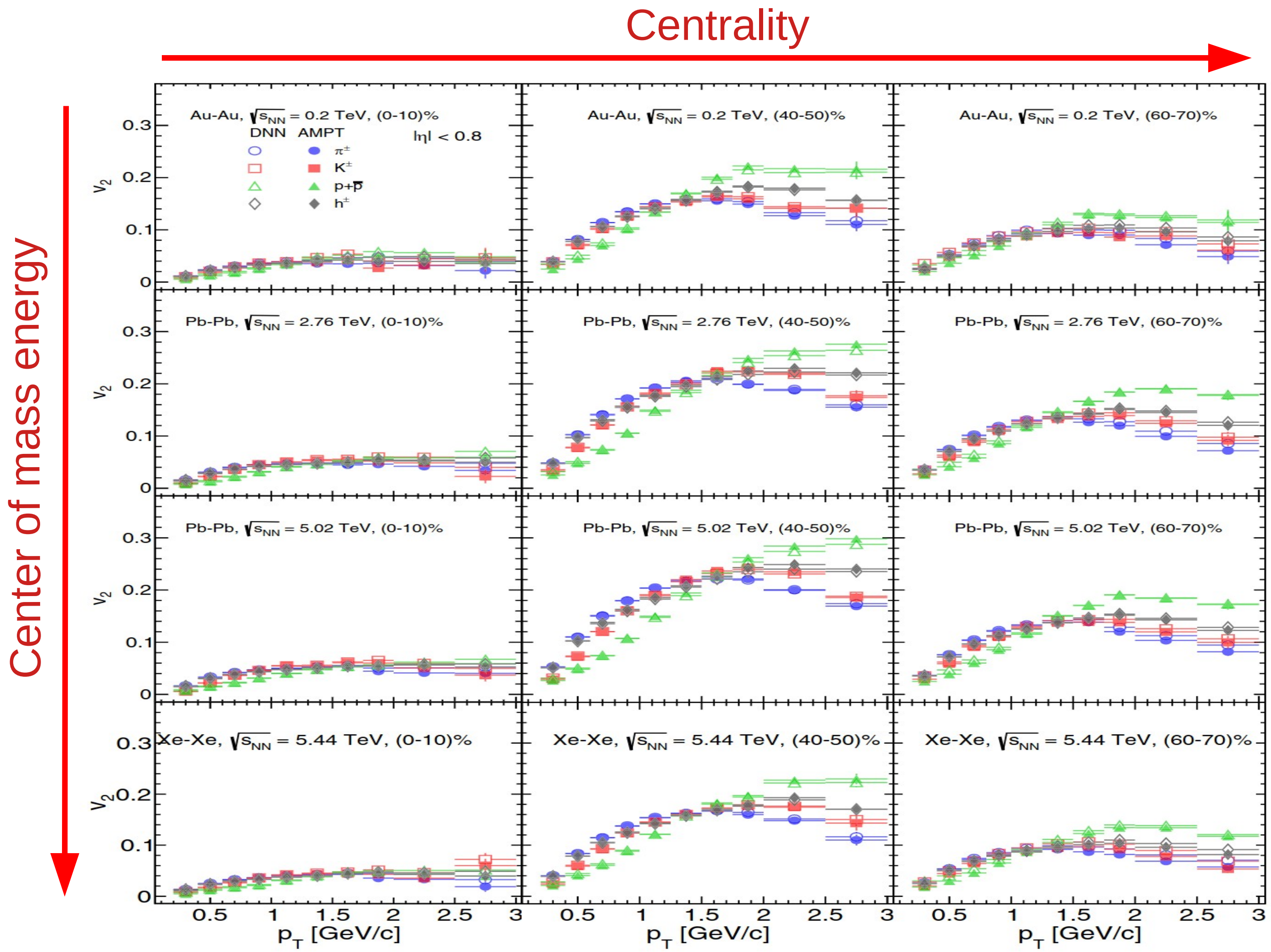
  The higher the energy higher effect.

- **System size**

  AuAu, PbPb, XeXe

- **DNN**

  Follows well the trends → scaling is encoded.



Scaling properties are in the DNN

- **Centrality**

  The largest in case 30%-40% mid-central

- **Collision energy**

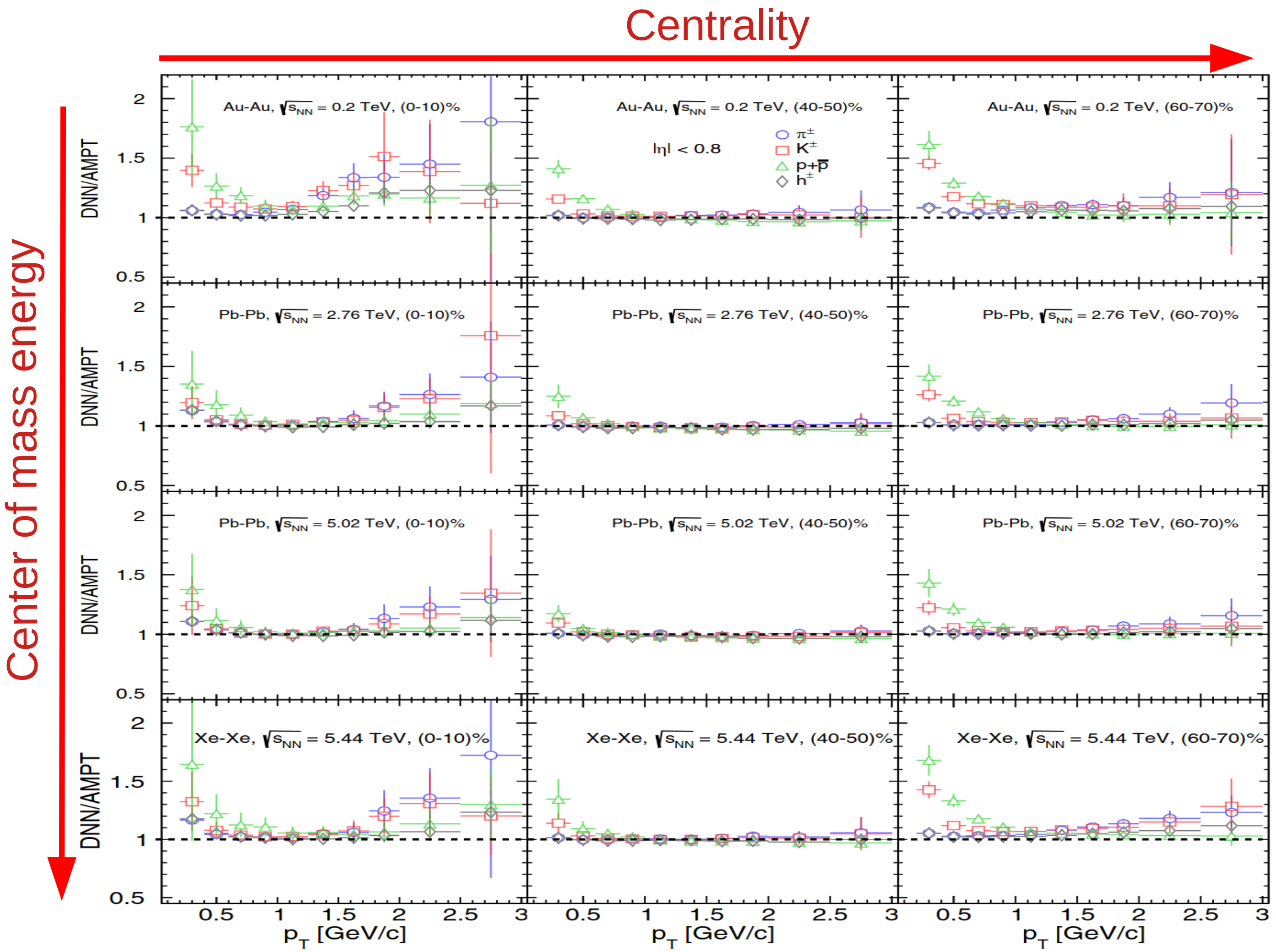
  The higher the energy higher effect.

- **System size**

  AuAu, PbPb, XeXe

- **DNN**

  Follows well the trends → scaling is encoded.



The AMPT model for Pb-Pb collisions

- **A Multi-phase transport model (AMPT):** MC event generator for simulating p-A and A-A collisions from RHIC to LHC energies.
  - **Fluctuating initial conditions:** Initialization of collision is done by obtaining the spatial and momentum distributions of the hard minijet partons and soft string excitations from the HIJING model. The inbuilt Glauber model is used to calculate and convert the cross-section of the produced mini-jets from pp to AA.
  - **Zhang's parton cascade (ZPC)** model is used to perform the partonic interactions and parton cascade which currently includes the two-body scatterings with cross-sections obtained from the pQCD with screening masses.
  - **Hadronization mechanism:** Lund string fragmentation model is used to recombine the partons with their parent strings and then the strings are converted to hadrons, whereas, in the string melting mode the transported partons are hadronized using a quark coalescence mechanism.
  - **Hadron cascade**: scattering among the produced hadrons are performed using a relativistic transport model (ART) by meson-meson, meson-baryon and baryon-baryon interactions.

Preliminary: results at higher p_T

- **AMPT vs Data**

  → Does not fit well above than few p_T Best at 30%-40% mid-central.

  → Need for more statistics

- **AMPT vs DNN**

  → DNN follows well the AMPT at any energy & centrality.

- **DNN**

  Follows well AMPT but NOT the high p_T data → need to improve!