Modeling N_{ch} distributions and p_T spectra in high-energy pp collisions with DNNs

Maria Alejandra Calmon Behling,
Jerome Jung, Mario Krüger, Henner Büsching

6th Inter-experiment Machine Learning Workshop, CERN

January 29, 2024

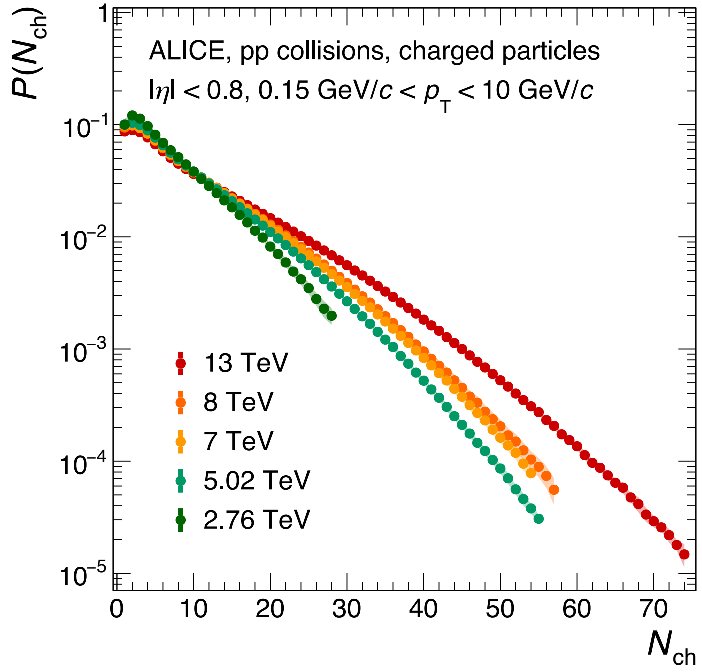
Institut für Kernphysik Frankfurt

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

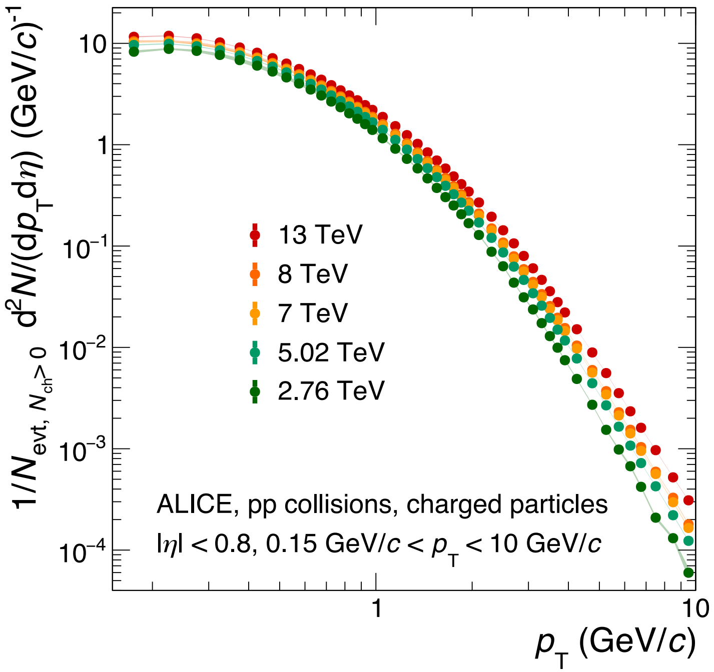
Motivation

- ALICE measurements [1] in pp collisions allow studying $\sqrt{s}$-dependent charged-particle production
- various center-of-mass energies available at LHC
- are observables at unmeasured energies predictable?

ALICE measurements



N_{ch} distributions



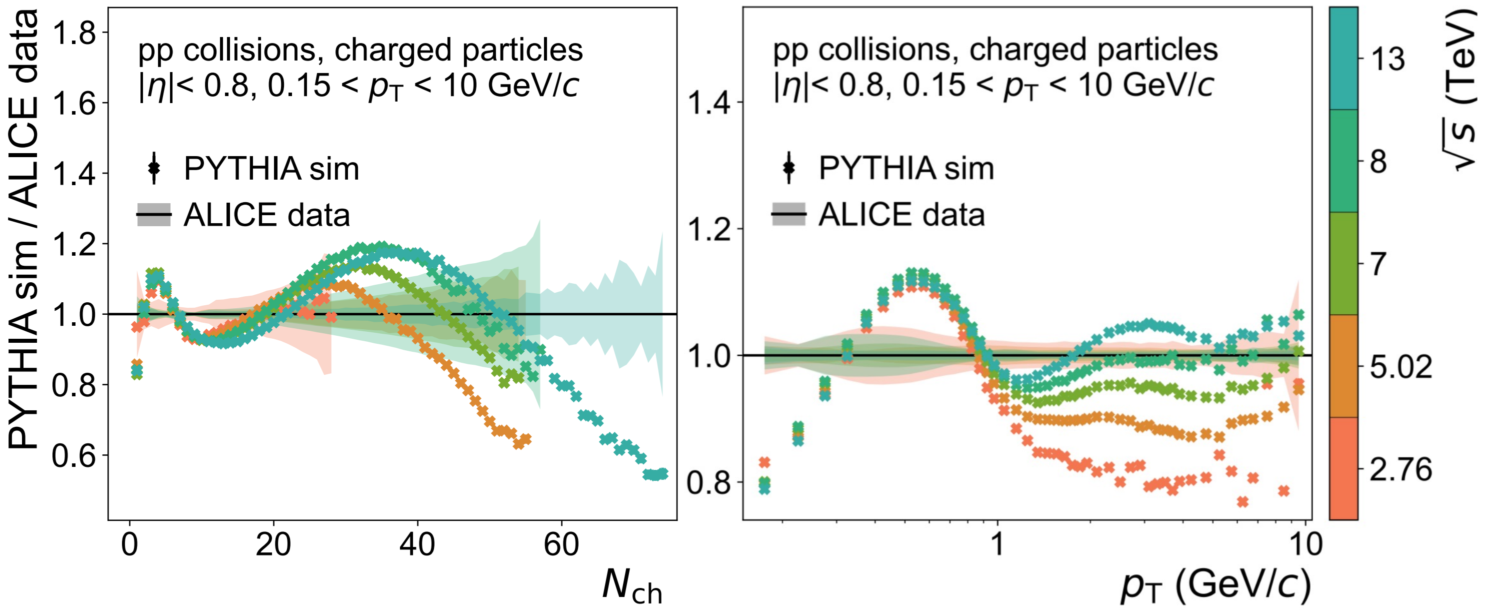
p_T spectra

[1] Phys. Lett. B Volume 845, 10 October 2023, 138110

Motivation

- ALICE measurements [1] in pp collisions allow studying $\sqrt{s}$-dependent charged-particle production
- various center-of-mass energies available at LHC
- are observables at unmeasured energies predictable?
- PYTHIA-simulated distributions not fully consistent with ALICE data

Ratio to PYTHIA Monash 2013



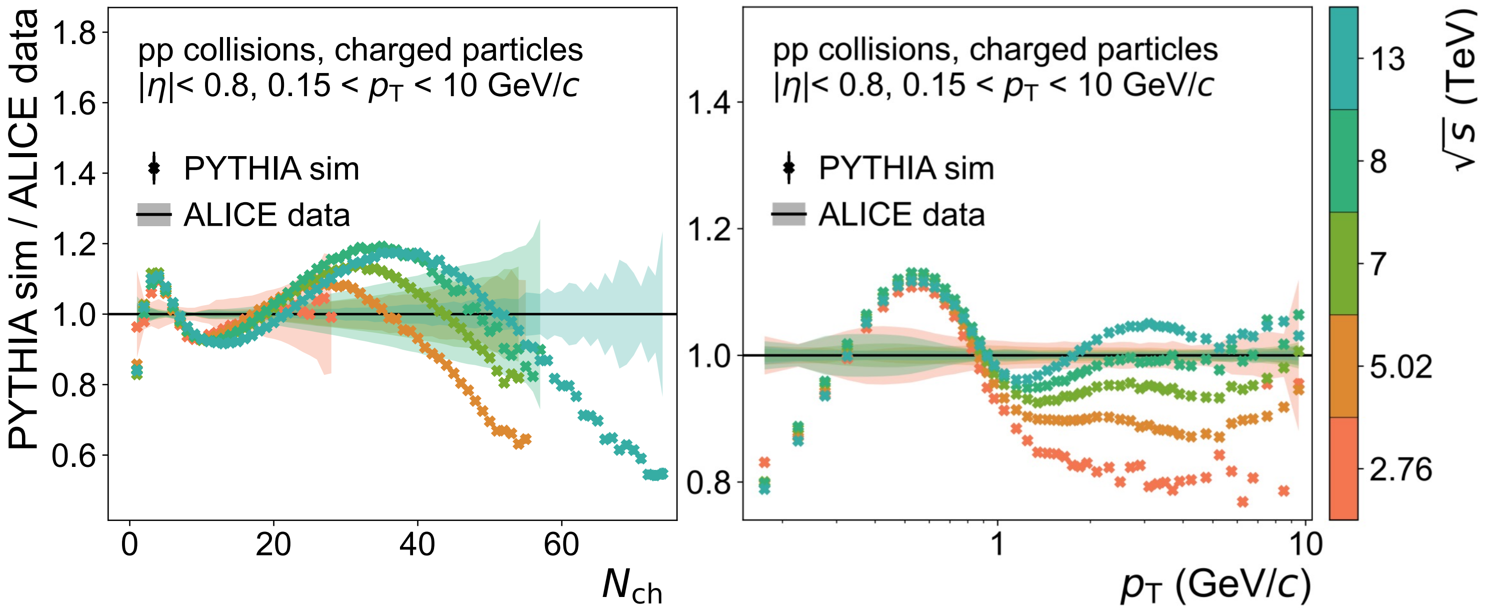
N_{ch} distributions

p_T spectra

Motivation

- ALICE measurements [1] in pp collisions allow studying $\sqrt{s}$-dependent charged-particle production
- various center-of-mass energies available at LHC
- are observables at unmeasured energies predictable?
- PYTHIA-simulated distributions not fully consistent with ALICE data

Ratio to PYTHIA Monash 2013



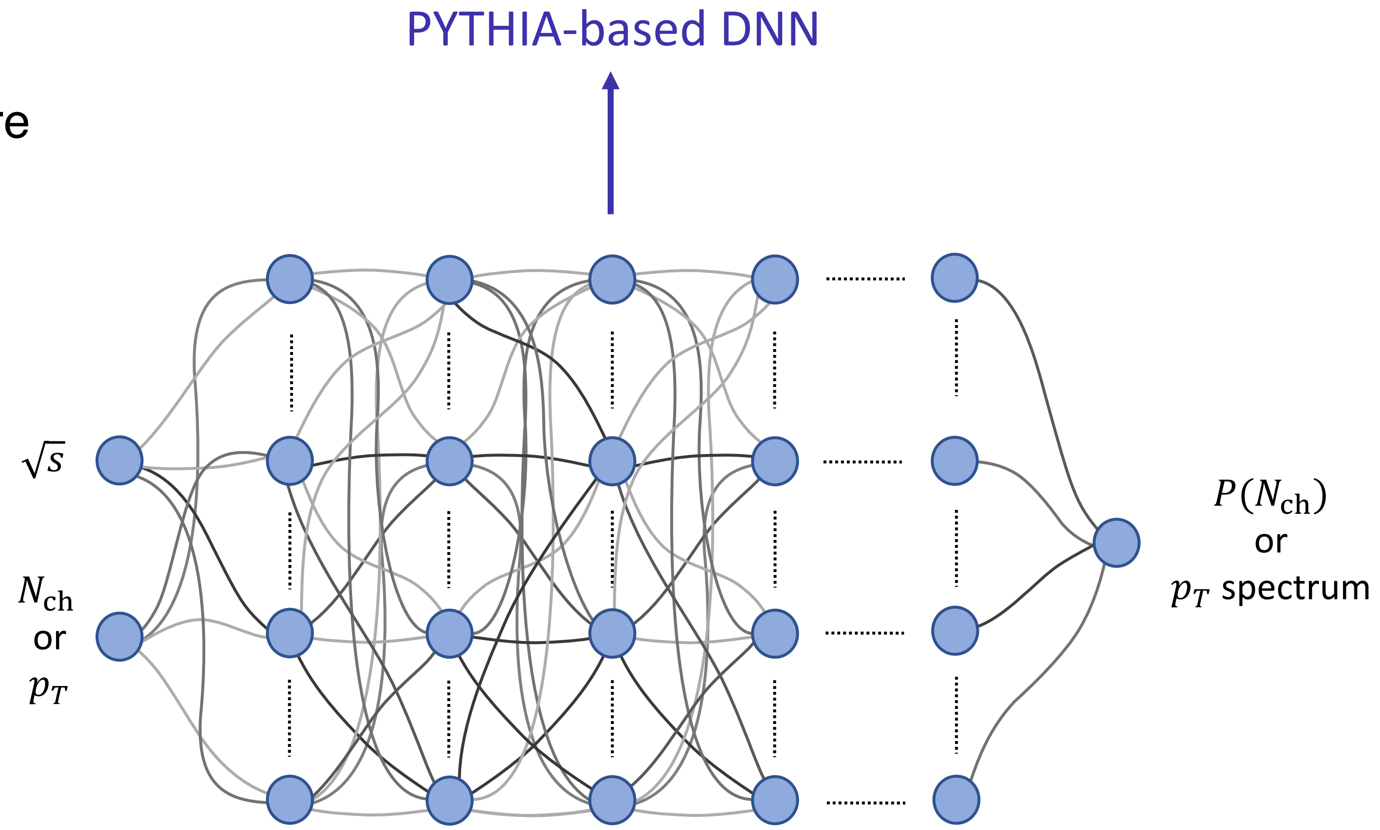
N_{ch} distributions **p_T spectra**

Alternative approach: DNNs
- parametrization of ALICE data
- predictions beyond discrete LHC energies

Model tuning

Hyperparameter Scan

- Bayesian-optimization search for best architecture
- models trained and evaluated with PYTHIA at:
  - ALICE-equivalent $\sqrt{s}$ (train./val.)
  - $\sqrt{s}$ = 0.5 TeV – 100 TeV (test)

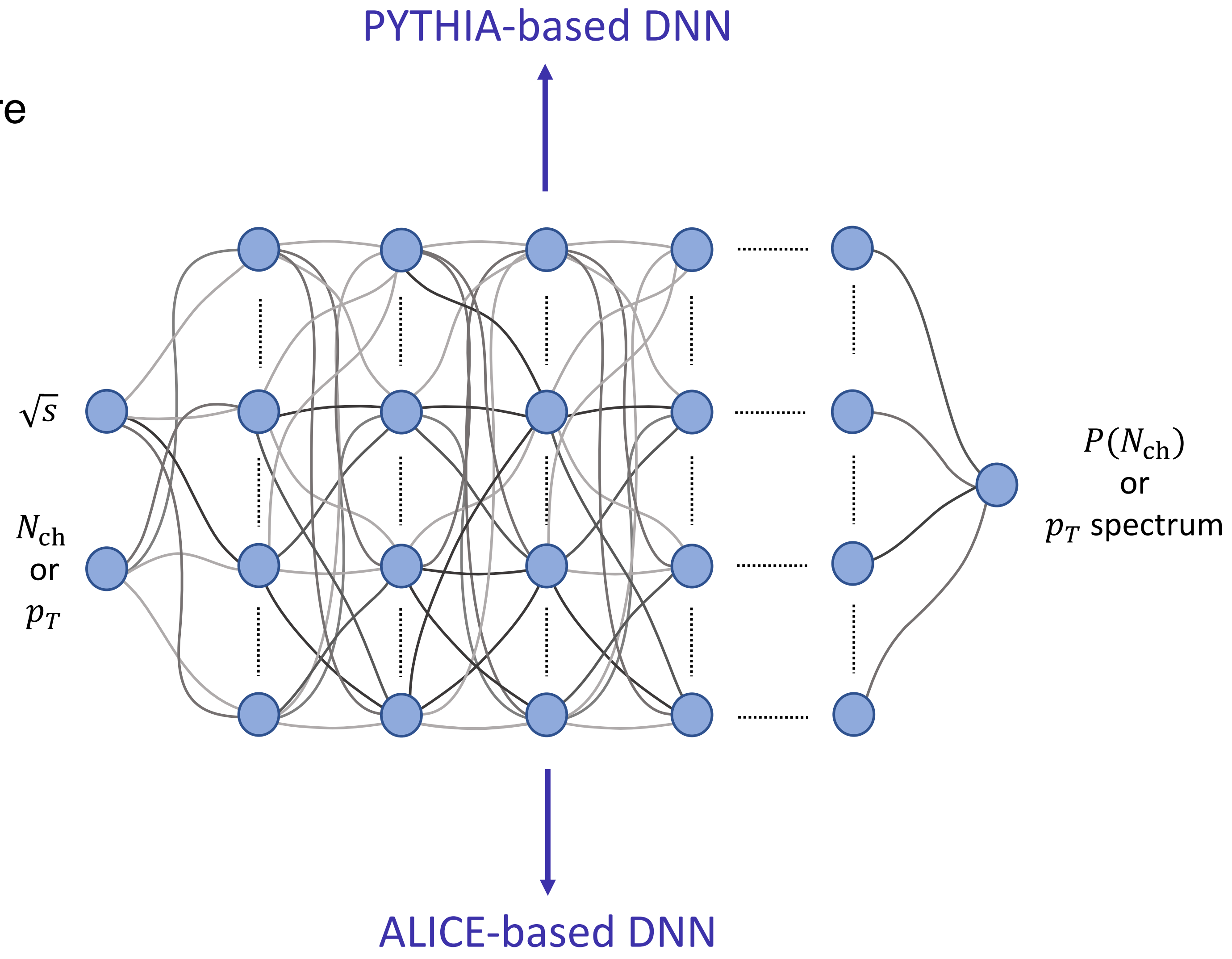


Model tuning

Hyperparameter Scan

- Bayesian-optimization search for best architecture
- models trained and evaluated with PYTHIA at:
  - ALICE-equivalent $\sqrt{s}$ (train./val.)
  - $\sqrt{s}$ = 0.5 TeV – 100 TeV (test)

➢ model architecture with best extrapolation performance chosen to train ALICE-based DNNs



Model tuning

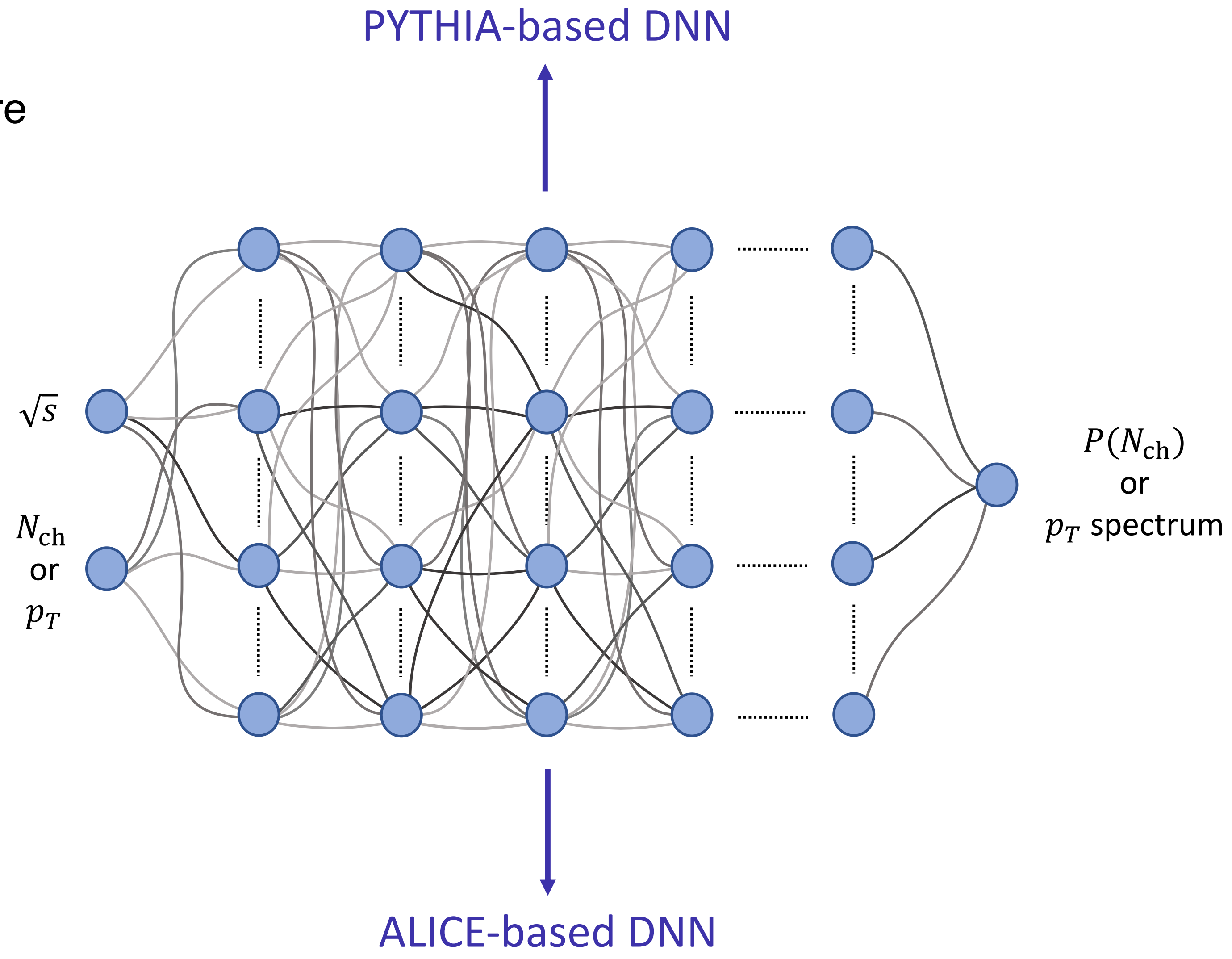
Hyperparameter Scan

- Bayesian-optimization search for best architecture
- models trained and evaluated with PYTHIA at:
  - ALICE-equivalent $\sqrt{s}$ (train./val.)
  - $\sqrt{s}$ = 0.5 TeV – 100 TeV (test)

➢ model architecture with best extrapolation performance chosen to train ALICE-based DNNs

Uncertainty estimation

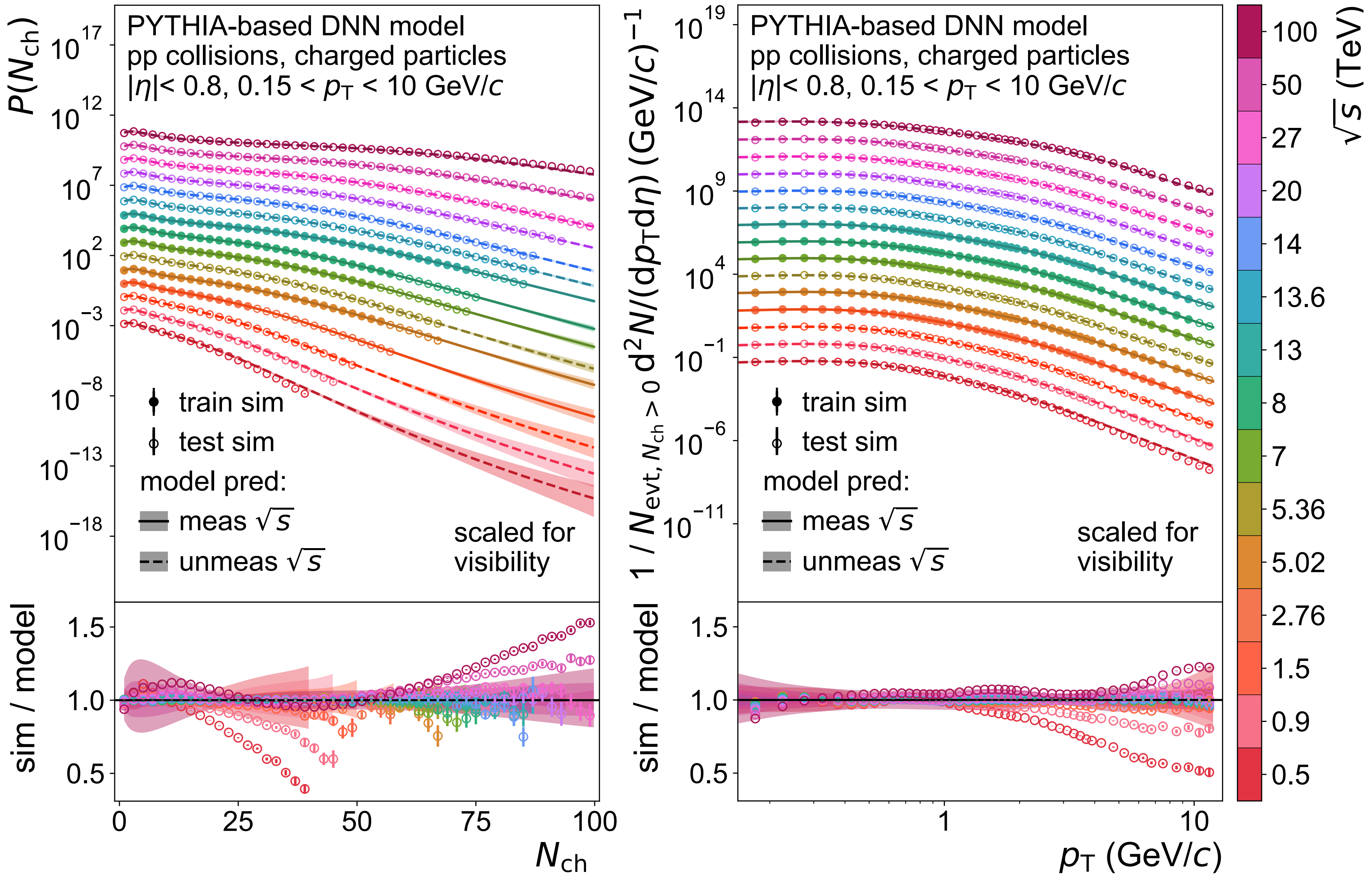
- different hyperparameter configurations
- random initialization and training-data selection



PYTHIA-based DNNs

- trained with simulations at $\sqrt{s}$ = 2.76, 5.02, 7, 8 and 13 TeV
- predictions within range $\sqrt{s}$ = 0.5 – 100 TeV

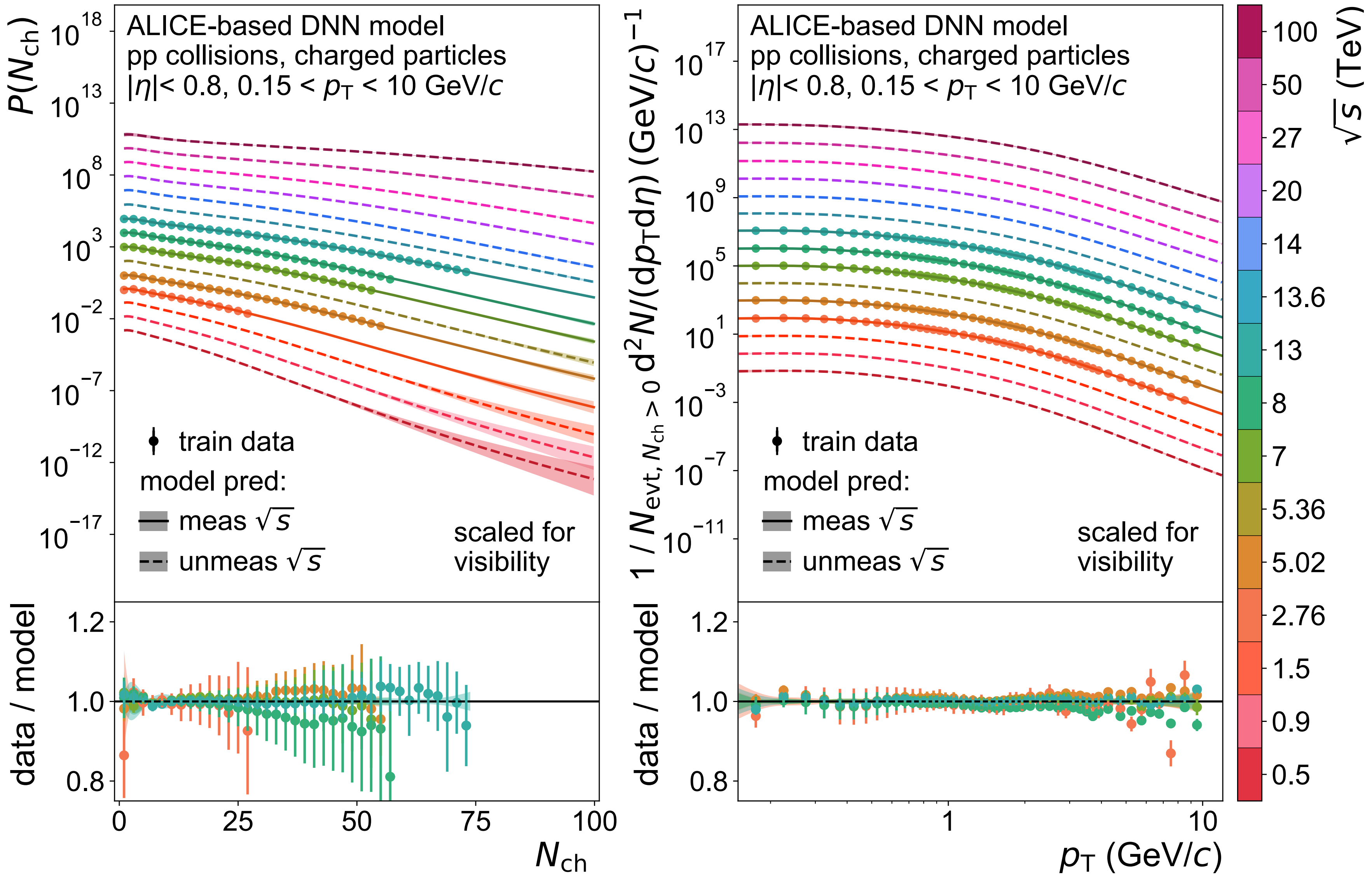
- $\sqrt{s}$-extrapolation highly consistent with test data within $\sqrt{s}$ = 1.5 – 27 TeV
- limitations at very high and very low $\sqrt{s}$



ALICE-based DNNs

- trained with ALICE data at $\sqrt{s}$ = 2.76, 5.02, 7, 8 and 13 TeV
- predictions within range $\sqrt{s}$ = 0.5 – 100 TeV

- excellent parametrization of ALICE measurements
- data-driven predictions for N_{ch} distributions and p_T spectra at arbitrary $\sqrt{s}$

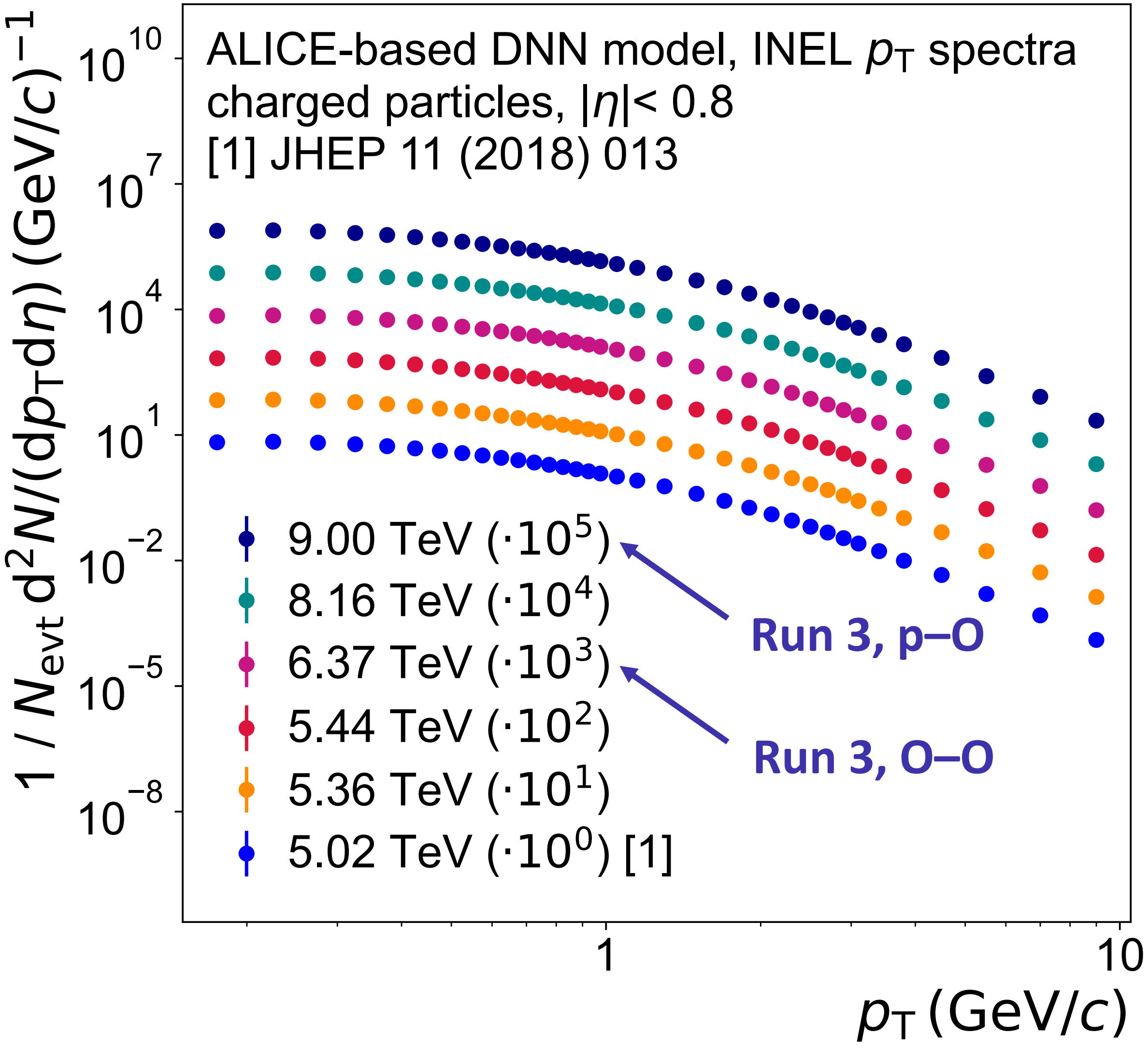


Interpolated pp reference p_T spectrum

- heavy-ion R_{AA} requires pp reference p_T spectrum at corresponding energy
- DNN provides p_T spectra, which can be converted to INEL through combination with published measurement at $\sqrt{s}$ = 5.02 TeV [1]

relevant for LHC Run 3:

- O–O at $\sqrt{s_{NN}}$ = 6.37 TeV, p–O at $\sqrt{s_{NN}}$ = 9 TeV
- no dedicated pp data-taking foreseen at these energies



[1] JHEP 11 (2018) 013

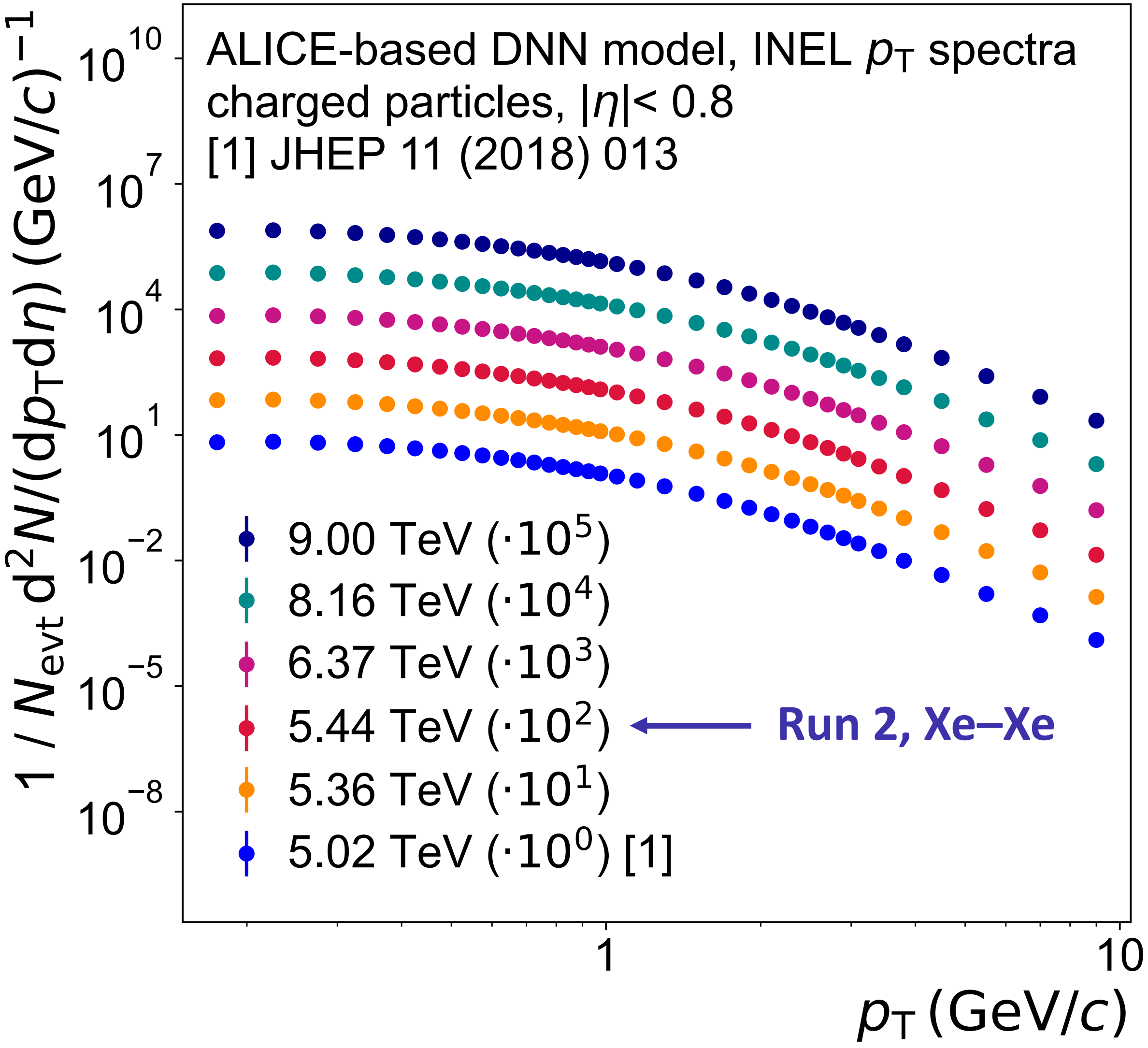
Interpolated pp reference p_T spectrum

- heavy-ion R_{AA} requires pp reference p_T spectrum at corresponding energy
- DNN provides p_T spectra, which can be converted to INEL through combination with published measurement at $\sqrt{s}$ = 5.02 TeV [1]

relevant for LHC Run 3:

- O–O at $\sqrt{s_{NN}}$ = 6.37 TeV, p–O at $\sqrt{s_{NN}}$ = 9 TeV
- no dedicated pp data-taking foreseen at these energies

➢ validation of DNN-predicted pp reference by comparing resulting R_{AA} for Xe–Xe at $\sqrt{s_{NN}}$ = 5.44 TeV to publication



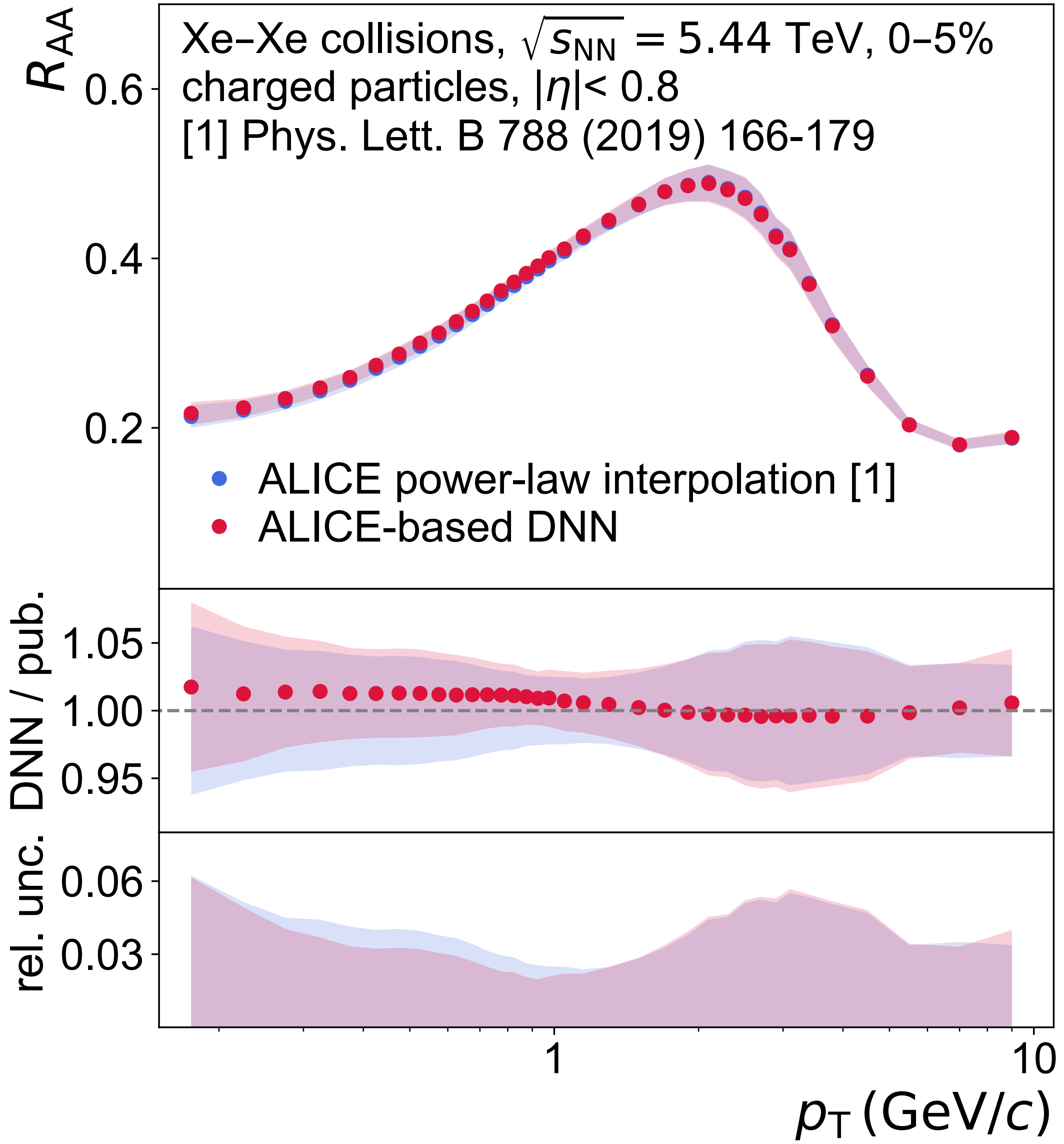
[1] JHEP 11 (2018) 013

R_{AA} for Xe–Xe

$$R_{AA} = \frac{dN^{AA}/dp_T}{\langle N_{coll} \rangle \, dN^{pp}/dp_T}$$

- publication: pp reference p_T spectrum obtained via power-law interpolation between two energies [1]
- ALICE-based DNN: pp reference p_T spectrum from parametrization of five energies + INEL measurement at $\sqrt{s}$ = 5.02 TeV

- R_{AA} determined with DNN highly consistent with publication
- slight deviations at low p_T, where DNN uncertainties are smaller



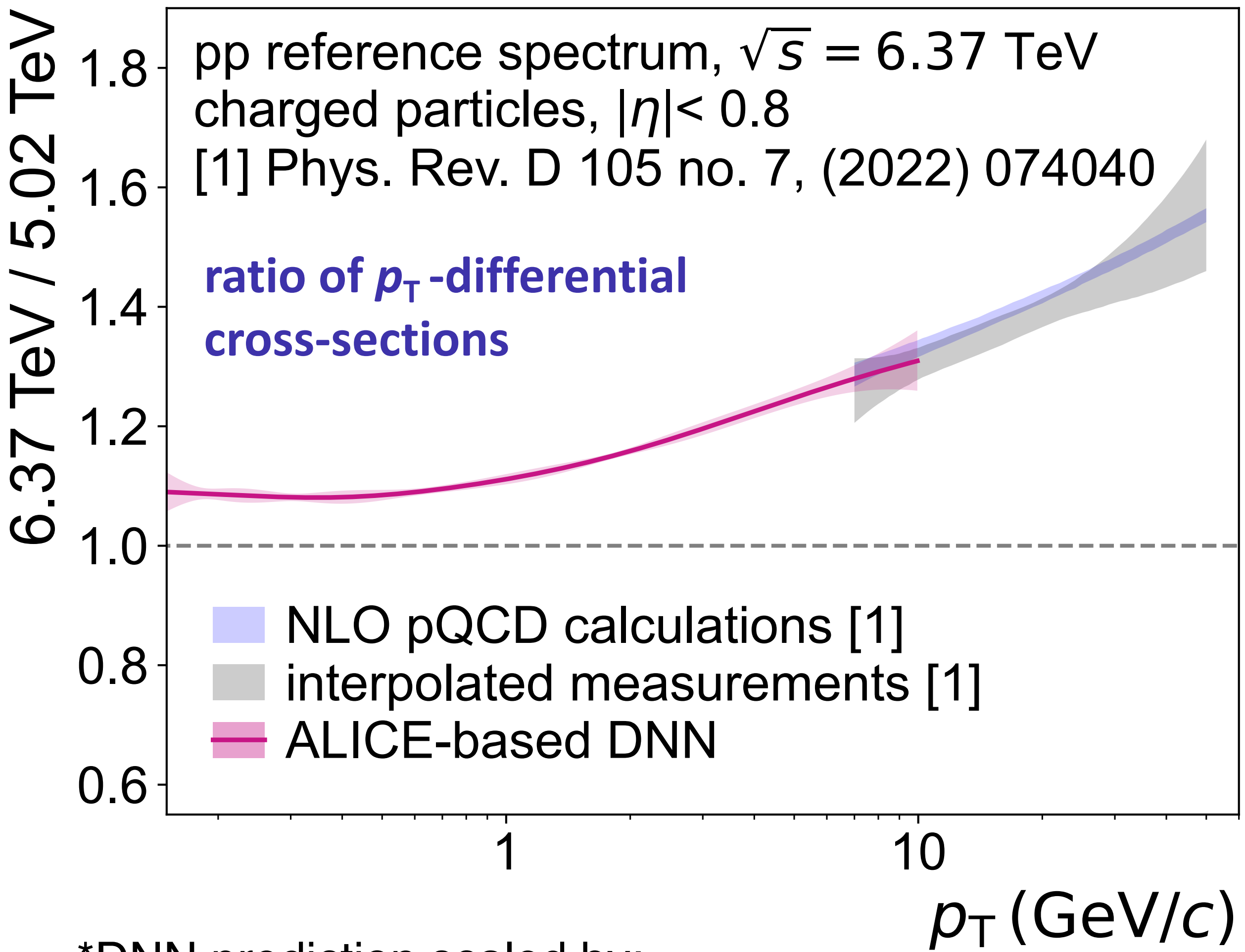
[1] Phys. Lett. B 788 (2019) 166-179

pp reference for O–O in LHC Run 3

- upcoming O–O data-taking in Run 3
- published ratios of p_{T}-differential cross-sections from NLO pQCD calculations and interpolated measurements [1]

➢ validation of DNN-predicted pp reference by comparing scaled* ratio to publication

- excellent agreement in overlap area $7\ \mathrm{GeV}/c < p_{\mathrm{T}} < 10\ \mathrm{GeV}/c$
- DNN successfully extends p_{T} range down to very low transverse momenta



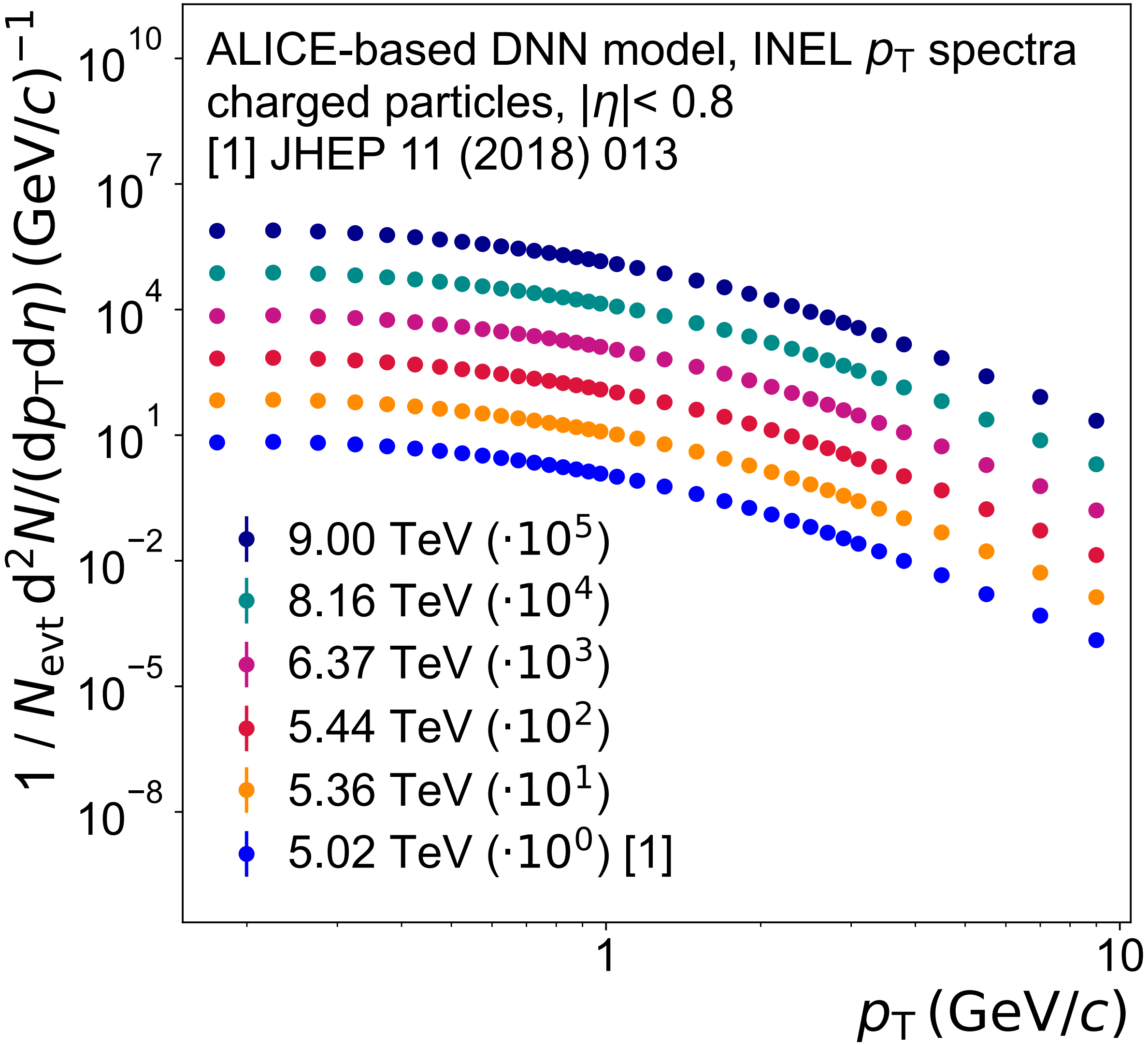
[1] Phys. Rev. D 105 no. 7, (2022) 074040

[2] Phys. Rev. C 97 no. 5, (2018) 054910

*DNN prediction scaled by:
$\sigma_{\mathrm{INEL}}(6.37\ \mathrm{TeV})\ /\ \sigma_{\mathrm{INEL}}(5.02\ \mathrm{TeV}) \approx 1.04$ [2]

Summary & Outlook

- ALICE measurements successfully parametrized by ALICE-based DNN
- $\sqrt{s}$-extrapolation evaluated via PYTHIA-based DNN
- predicted pp reference validated for Xe–Xe (Run 2)
- DNN extends projected pp reference for future O–O measurements (Run 3) to lower p_T



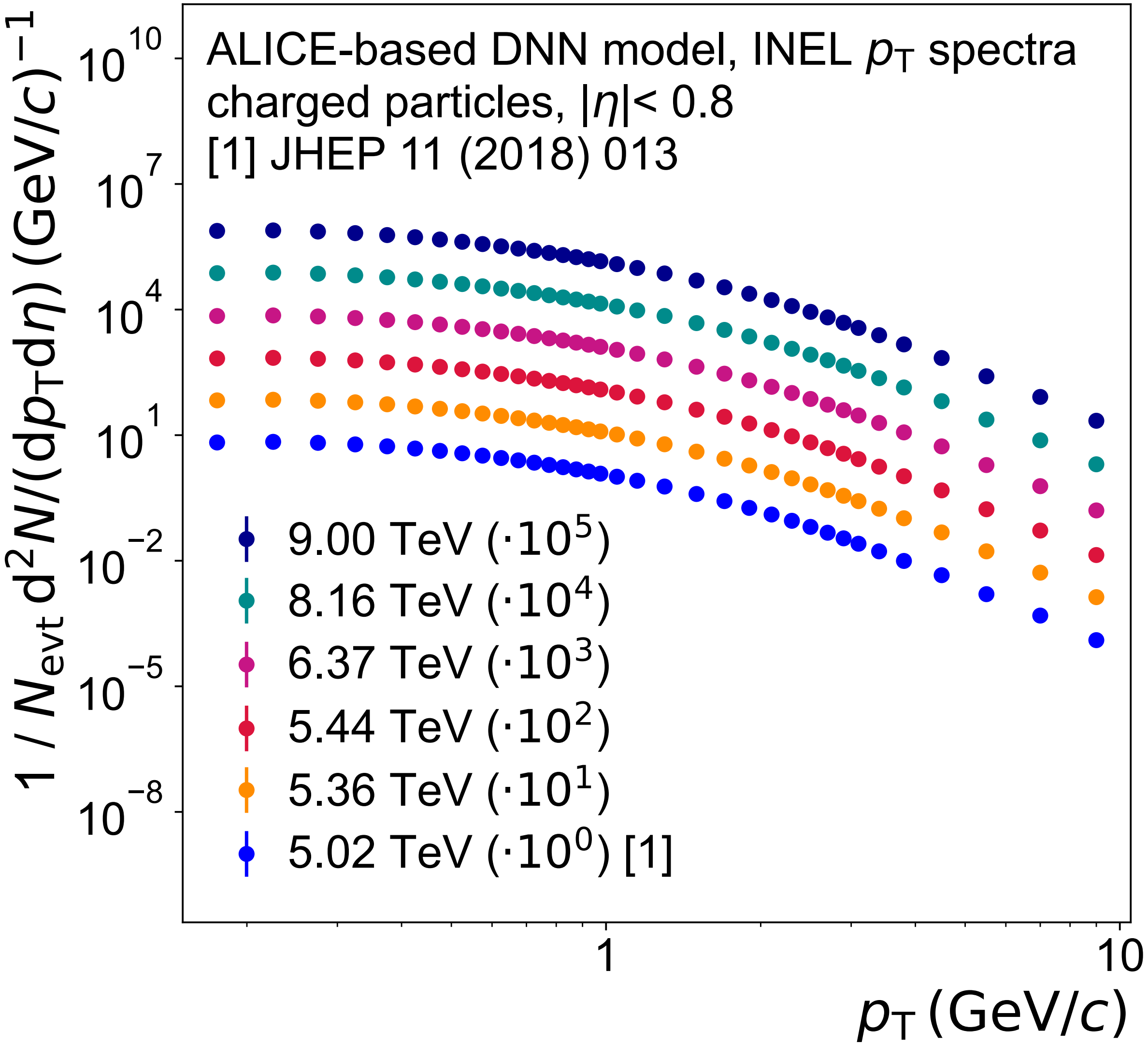
Outlook

- further details to be published soon
- explore more potential applications of DNNs for parametrizing particle spectra

Summary & Outlook

Poster session on Wednesday

- ALICE measurements successfully parametrized by ALICE-based DNN
- $\sqrt{s}$-extrapolation evaluated via PYTHIA-based DNN
- predicted pp reference validated for Xe–Xe (Run 2)
- DNN extends projected pp reference for future O–O measurements (Run 3) to lower p_T



Outlook

- further details to be published soon
- explore more potential applications of DNNs for parametrizing particle spectra

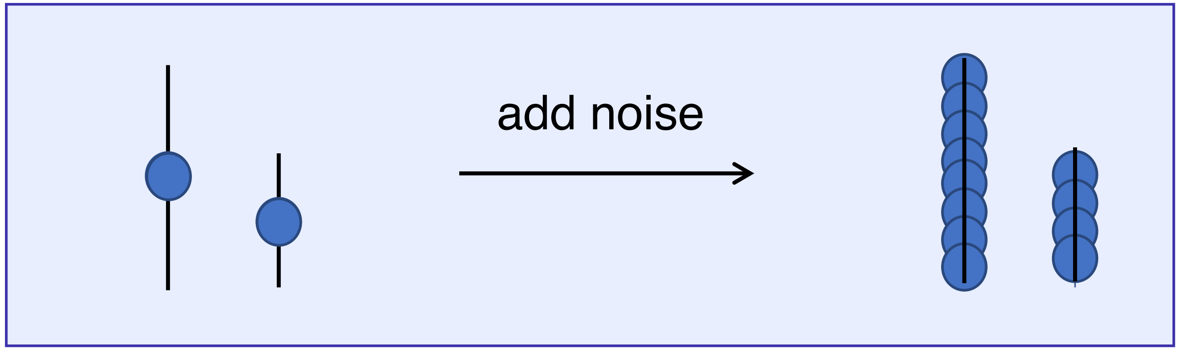
Backup

Data preparation

- spectral shapes span many orders of magnitude
  - model works best with smaller range of absolute values
  - **logarithmic transformation applied** to all data

$$\sqrt{s} \longrightarrow \log(\sqrt{s})$$
$$N_{ch} \longrightarrow \log(N_{ch})$$
$$P(N_{ch}) \longrightarrow \log(P(N_{ch}))$$
$$p_T \longrightarrow \log(p_T)$$
$$\text{yield}(p_T) \longrightarrow \log(\text{yield}(p_T))$$

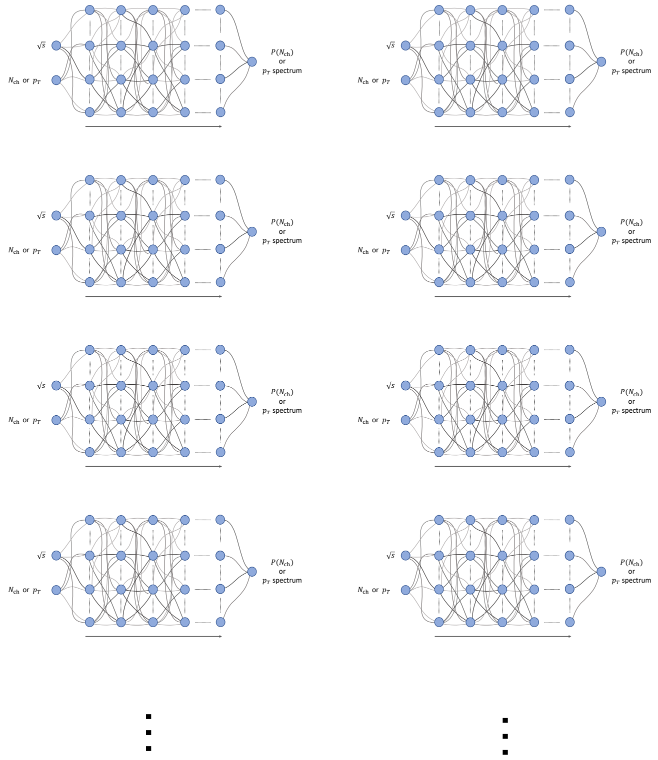
- DNNs tend to memorize training data
  - **split data** into 80% training, 20% validation
  - **create more datapoints** uniformly within systematic data uncertainty



DNN model uncertainties

Ensemble of 20 DNNs

same architecture, different initialization

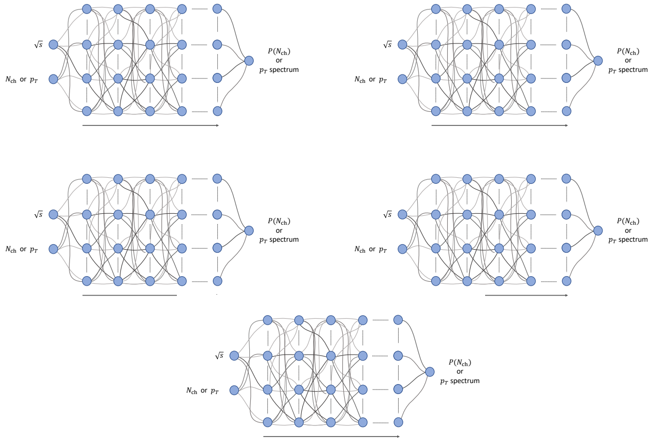


nominal predictions
+ ensemble uncertainties

hyperparameter
uncertainties

5 DNN architectures

best models from hyperparameter scan



DNN model architectures

Hyperparameter scan: Bayesian optimization

sampling: discrete values			
layers	neurons per layer	activation function	initializer
2	32	TanH (TH)	RandomUniform (RU)
3	64	ReLU (RE)	RandomNormal (RN)
4	128	SeLU (SE)	TruncatedNormal (TN)
5	256	Swish (SW)	GlorotUniform (GN)
	512	Mish (MI)	GlorotNormal (GU)
		Softplus (SP)	

sampling: intervals (logarithmic)			
	λ_1	λ_2	learning rate (lr)
min	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	$1 \cdot 10^{-5}$
max	$5 \cdot 10^{-1}$	$5 \cdot 10^{-1}$	$1 \cdot 10^{-3}$

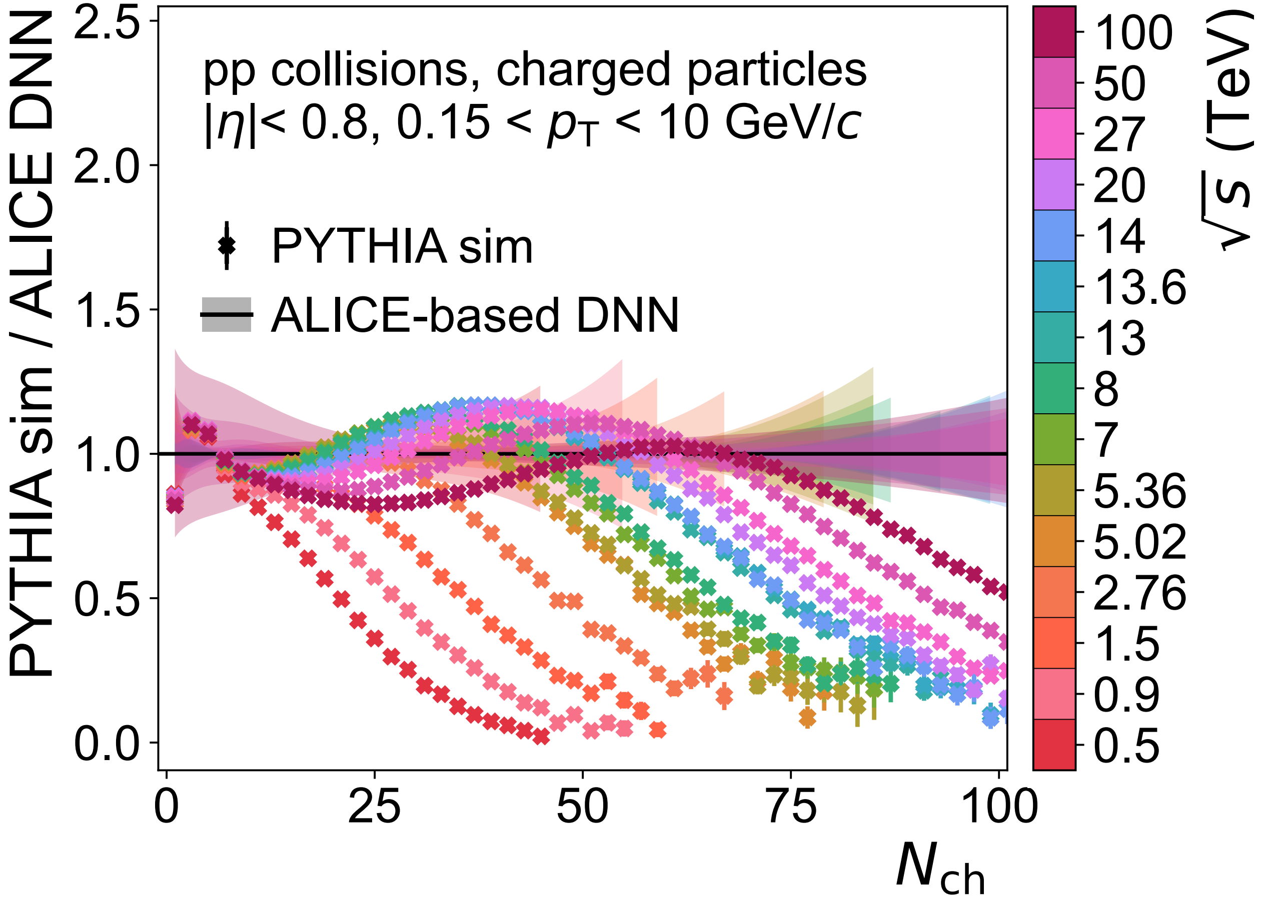
obs.	lay.	nod.	lr	act.	init.	λ_1	λ_2	objective
	5	**512**	$\mathbf{1 \cdot 10^{-5}}$	**SW**	**GU**	$\mathbf{5 \cdot 10^{-8}}$	$\mathbf{5 \cdot 10^{-8}}$	**0.121**
	5	128	$9 \cdot 10^{-5}$	SW	GU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.146
N_{ch}	2	512	$1 \cdot 10^{-4}$	SW	RU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.148
	4	512	$2 \cdot 10^{-5}$	MI	GU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.152
	5	512	$1 \cdot 10^{-5}$	MI	GU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.153
	5	**512**	$\mathbf{8 \cdot 10^{-5}}$	**SW**	**GU**	$\mathbf{5 \cdot 10^{-8}}$	$\mathbf{5 \cdot 10^{-8}}$	**0.080**
	5	512	$6 \cdot 10^{-5}$	SW	GU	$5 \cdot 10^{-8}$	$6 \cdot 10^{-6}$	0.085
p_{T}	5	512	$2 \cdot 10^{-4}$	SW	RU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.100
	5	512	$1 \cdot 10^{-4}$	MI	GU	$5 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	0.102
	5	512	$2 \cdot 10^{-x}$	SW	GU	$8 \cdot 10^{-7}$	$5 \cdot 10^{-8}$	0.110

PYTHIA vs. ALICE-based DNN

- ratio of PYTHIA-simulated to predicted N_{ch} spectra by ALICE-based DNN

- DNN reproduces trend observed for ALICE data:
  - $N_{ch} \lesssim 10$: almost $\sqrt{s}$ independent
  - $N_{ch} > 10$: energy ordering

trend of deviations observed between PYTHIA and ALICE data extended by ALICE-based DNN beyond LHC energies

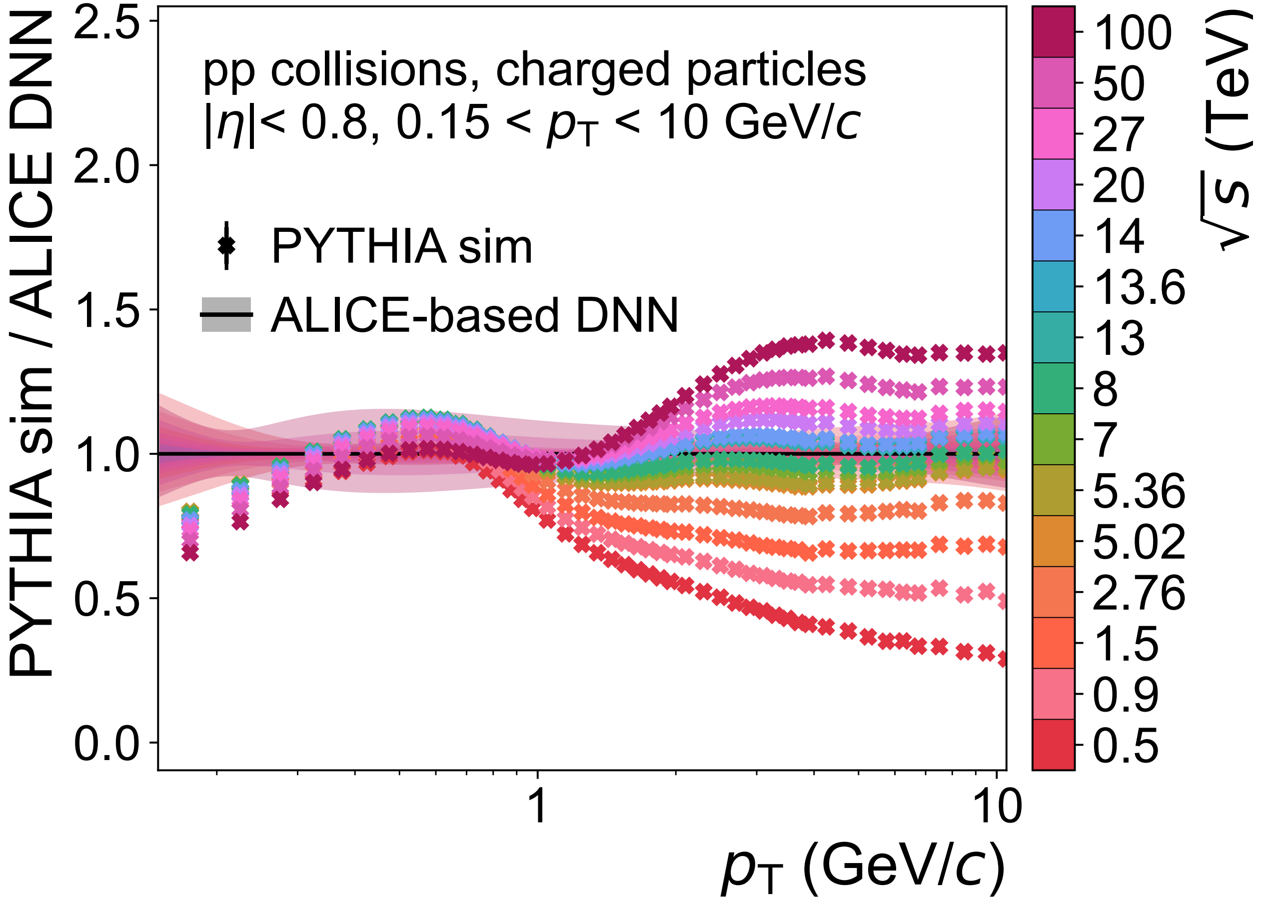


PYTHIA vs. ALICE-based DNN

- ratio of PYTHIA-simulated to predicted p_{T} spectra by ALICE-based DNN

- DNN reproduces trend observed for ALICE data:
  - $p_{\mathrm{T}} \lesssim 1$ GeV/c : almost $\sqrt{s}$ independent
  - $p_{\mathrm{T}} > 1$ GeV/c : energy ordering

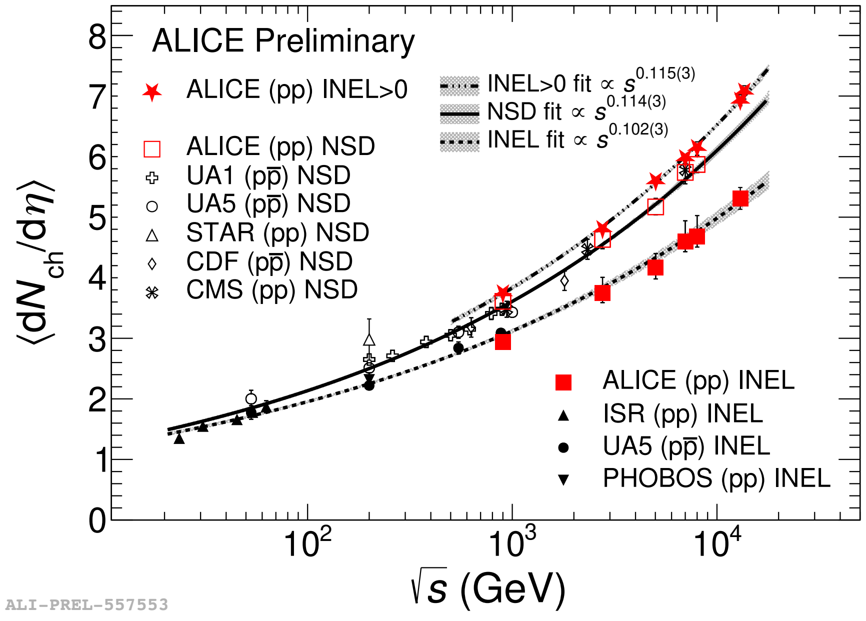
trend of deviations observed between PYTHIA and ALICE data extended by ALICE-based DNN beyond LHC energies



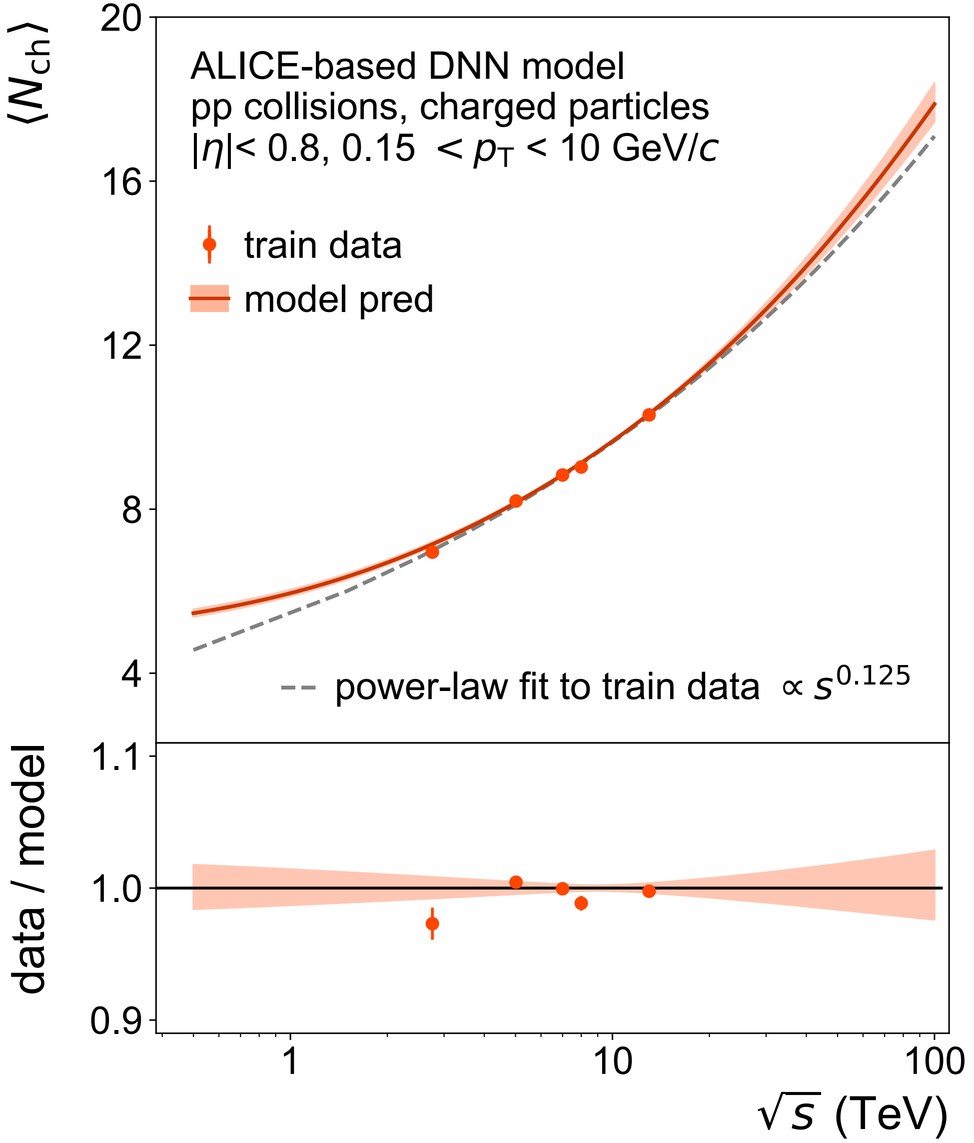
Energy-dependence of $\langle N_{ch} \rangle$

- measurements of $\langle dN_{ch}/d\eta \rangle$ follow power law
- $\langle N_{ch} \rangle$ predicted by ALICE DNN ($\sqrt{s}$ = 0.5 – 100 TeV)
- power-law fit applied to ALICE data compared to DNN

DNN compatible with power-law fit to ALICE data



power-law fit

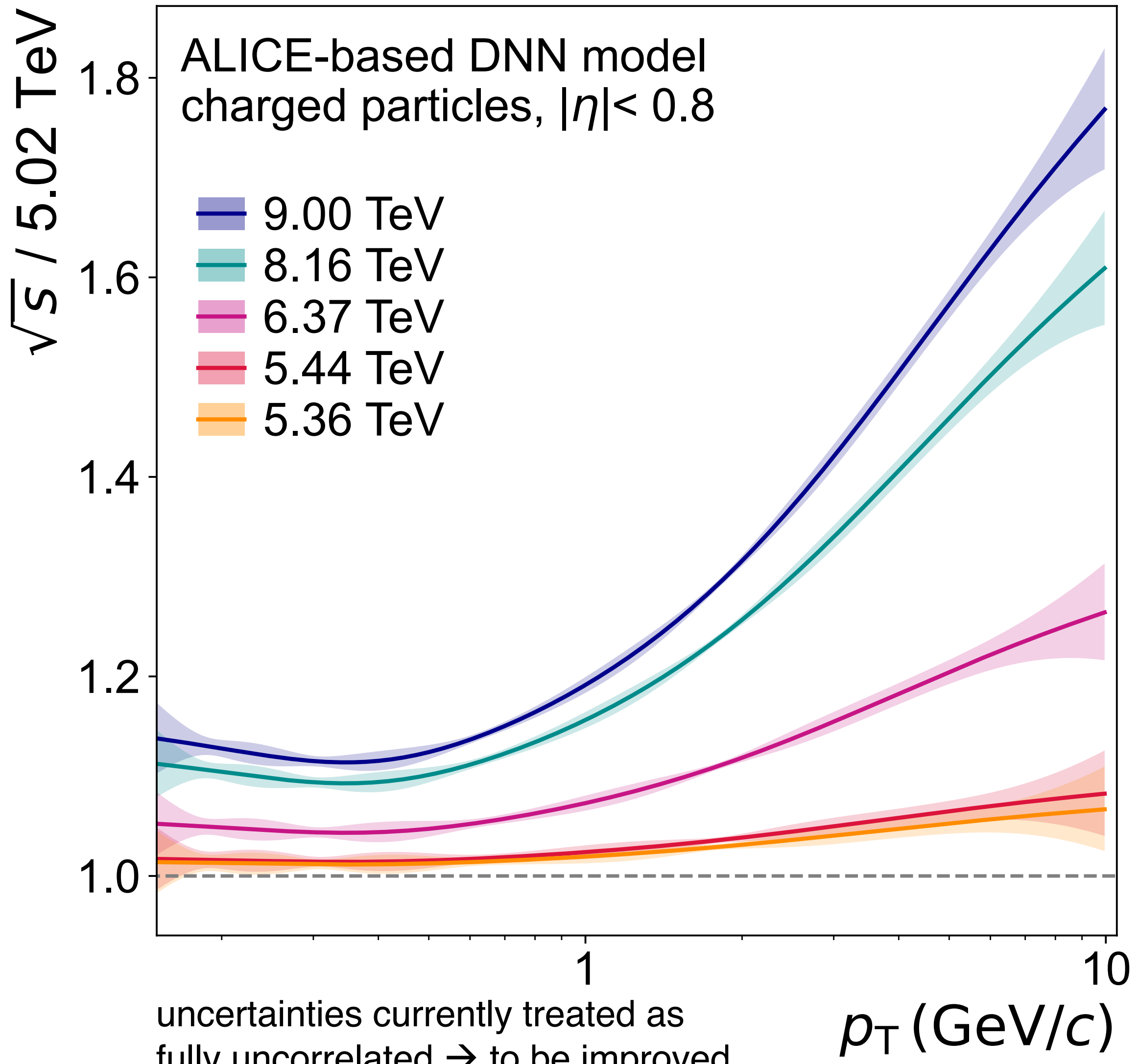


Interpolated pp reference p_T spectrum

- ratio of predicted p_T spectra by ALICE-based DNN at different $\sqrt{s}$ to 5.02 TeV
- shows energy dependence of spectral shape

R_{AA} for heavy-ion collisions

- pp reference p_T spectra needed at corresponding $\sqrt{s}$
- pp measurements not always available



uncertainties currently treated as fully uncorrelated → to be improved

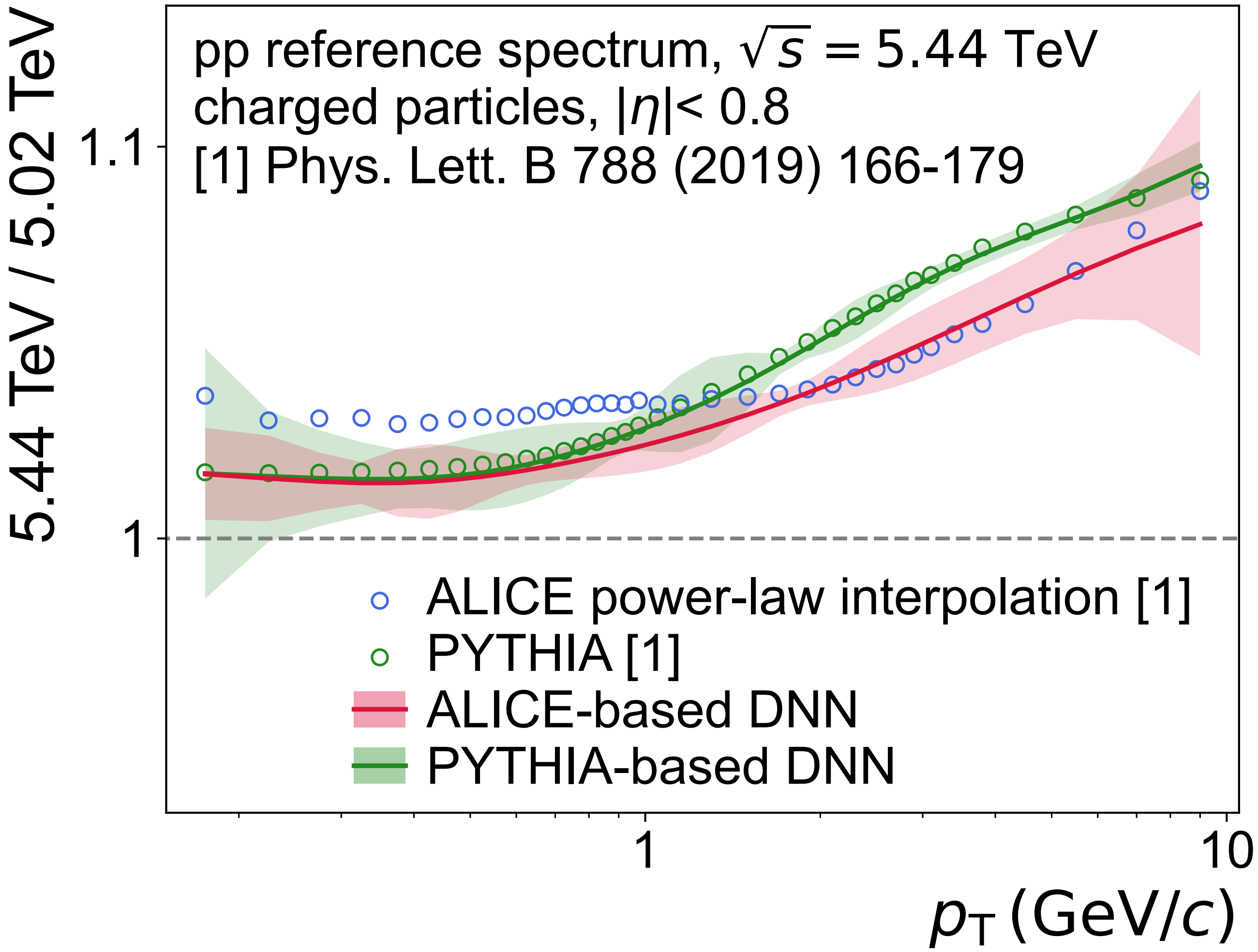
Interpolated pp reference p_T spectrum

- publication: Phys. Lett B 788 (2019) 166-179
  - pp reference: power-law interpolation
  - input: two energies ($\sqrt{s} = 5.02$, 7 TeV)

- comparison of DNN to publication:
  - ratio of $\sqrt{s_{NN}} = 5.44$ TeV to 5.02 TeV

- DNN:
  - consistent with PYTHIA and publication
  - input: five energies (LHC energies)

DNN successfully provides pp reference for heavy-ion collisions



High-p_T extrapolation: TCM fit to DNN

- Two-Component Model (TCM) well-established for parametrizing p_T spectra

$$\frac{d\sigma}{p_T dp_T} = A_e \exp(-E_{Tkin}/T_e) + \frac{A}{(1+\frac{p_T{}^2}{T^2 * n})^n}$$

- tail of spectrum at high p_T dominated by power law

➢ TCM fit to all DNN predictions up to $p_T = 10$ GeV/c, which is the p_T range of the ALICE measurements used for training

- excellent agreement up to $p_T = 10$ GeV/c, where predictions are well constrained by training data
- increasing deviations for higher p_T, where DNN has more freedom