Instituto de
Ciencias
Nucleares
UNAM

Multi-Parton Interactions from Machine Learnig-based regression

Antonio Ortiz (ICN, UNAM)
Erik Zepeda, Antonio Paz

Based on: arXiv:2004.03800

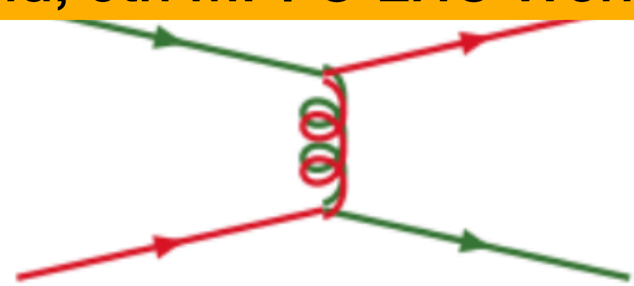
October 22, 2020

Motivation

MPI in HEP

- At high energies, the leading order cross-section for $2 \to 2$ parton scatterings with momentum transfer $Q > Q_{\mathrm{min}} \gg \Lambda_{\mathrm{QCD}}$ exceeds the total pp cross-section at a range of Q_{min}-values where perturbative QCD is applicable (at LHC, $Q_{\mathrm{min}} \approx 4$ GeV/c) [T. Sjöstrand and M. Zijil Phys. Rev. D36 (1987)]

T. Sjöstrand, 6th MPI @ LHC Workshop



Integrate QCD $2 \to 2$

$qq' \to qq'$

$q\bar{q} \to q'\bar{q}'$

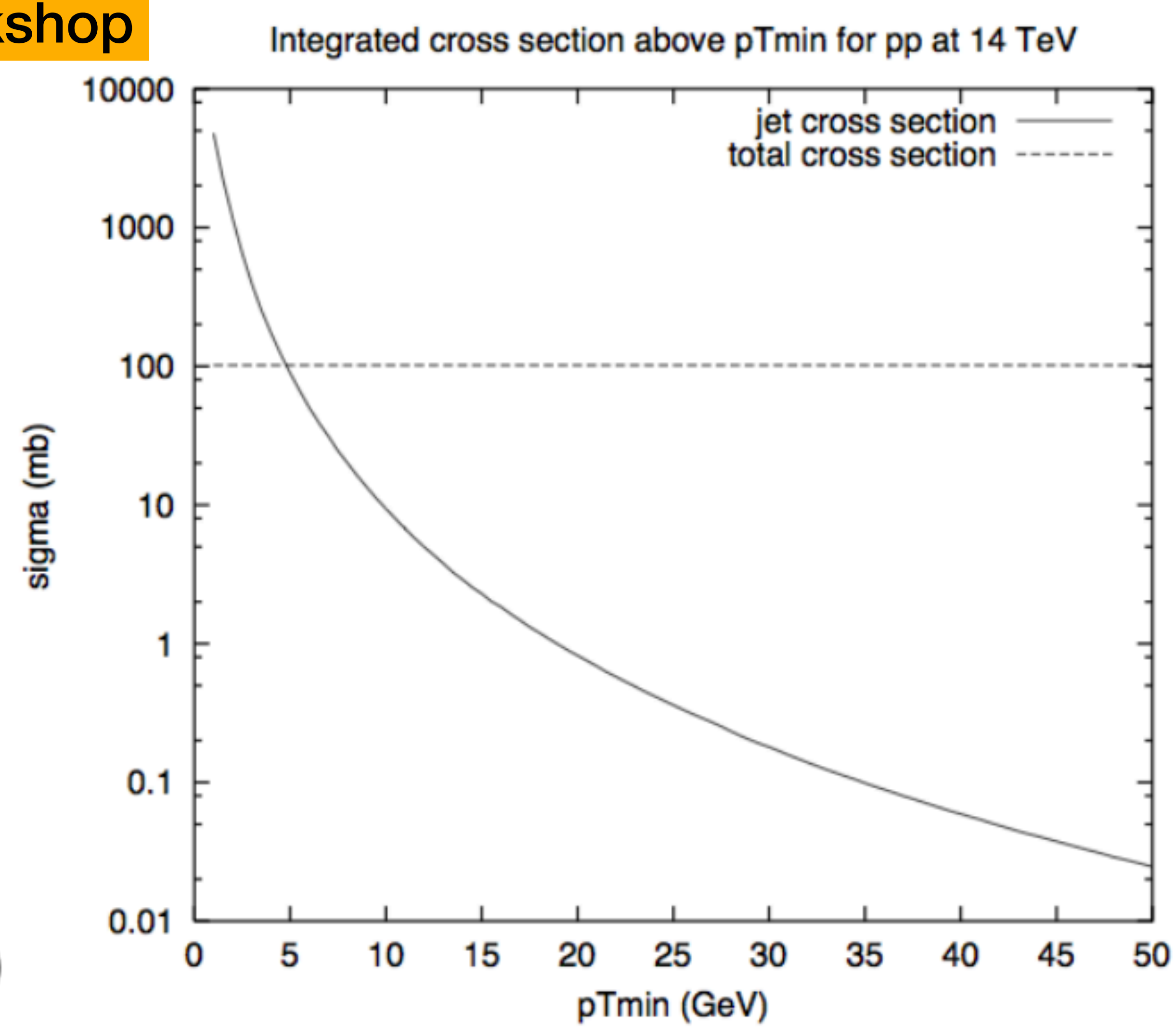
$q\bar{q} \to gg$

$qg \to qg$

$gg \to gg$

$gg \to q\bar{q}$

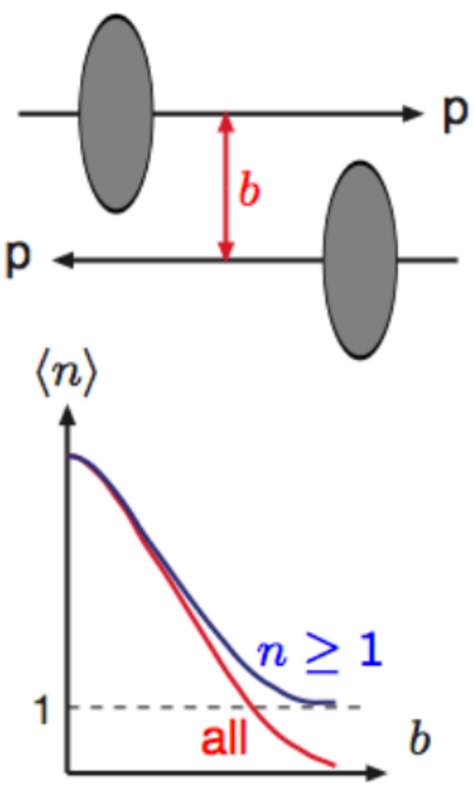
(with CTEQ 5L PDF's)



- At high energies, the leading order cross-section for $2 \to 2$

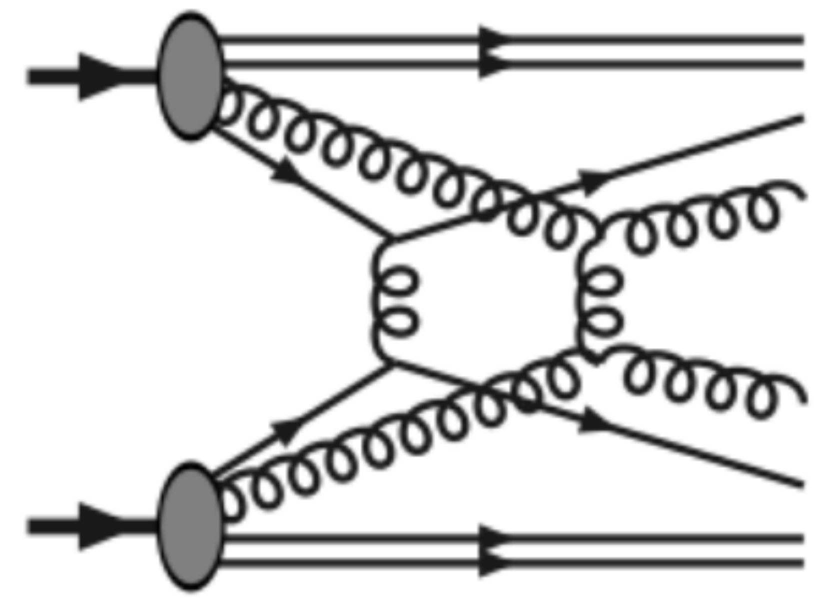
Interpretation: Many partonic scatterings per event: (MPI)

- MPI is a logical consequence of the composite nature of protons



T. Sjöstrand, ISAPP 2018

- In event generators like Pythia, an impact parameter dependence is considered

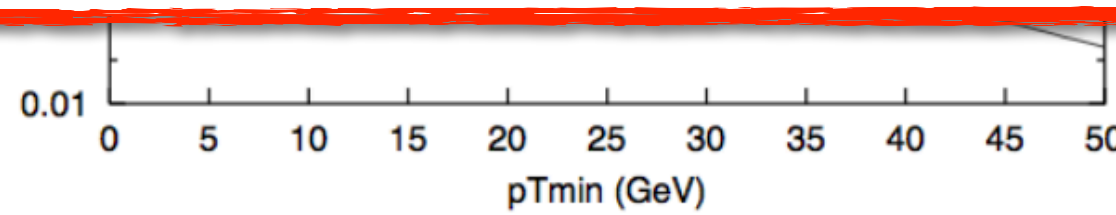


Overlap of protons during encounter is

$$\mathcal{O}(b) = \int \mathrm{d}^3\mathbf{x}\,\mathrm{d}t\ \rho_1(\mathbf{x},t)\,\rho_2(\mathbf{x},t)$$

where ρ is (boosted) matter distribution in p, e.g. Gaussian or more narrow peak.

$gg \to q\bar{q}$
(with CTEQ 5L PDF's)



MPI in HEP

- At high energies, the leading order cross-section for $2 \to 2$

Interpretation: Many partonic scatterings per event: (MPI)

MPI help to describe particle multiplicities in MB events

T. Sjöstrand and M. v. Zijl, PRD 36 (1987) 2019

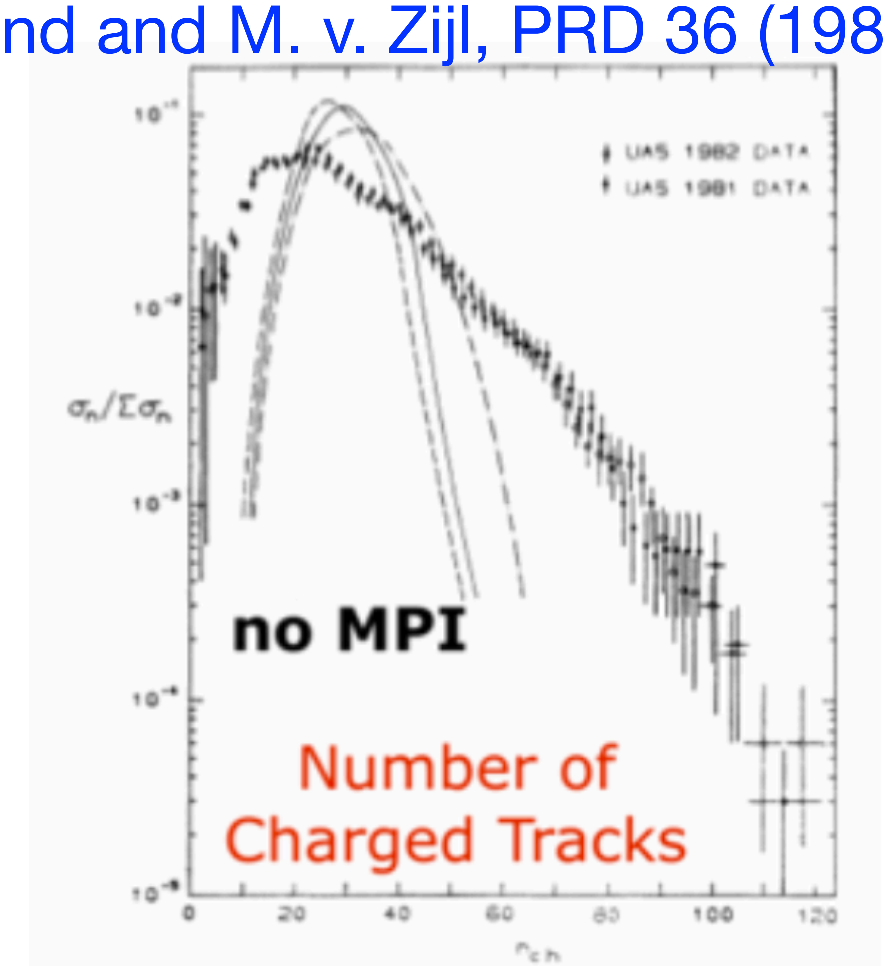


FIG. 3. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs simple models: dashed low p_T only, full including hard scatterings, dash-dotted also including initial- and final-state radiation.

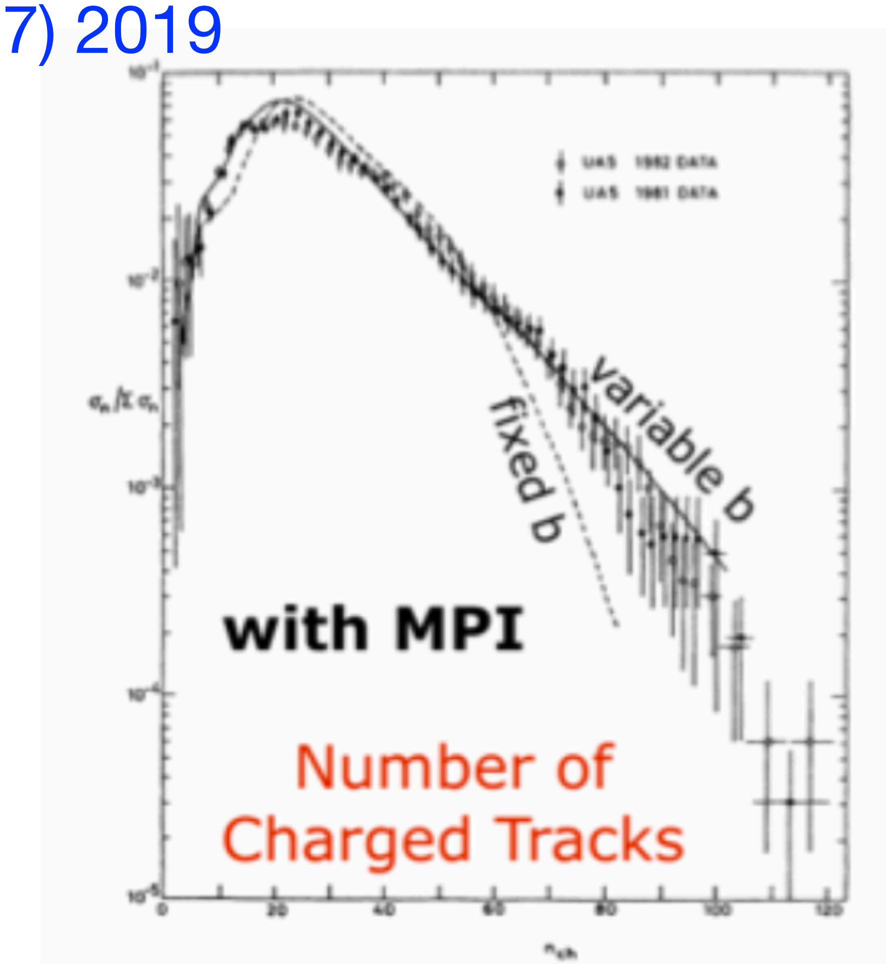
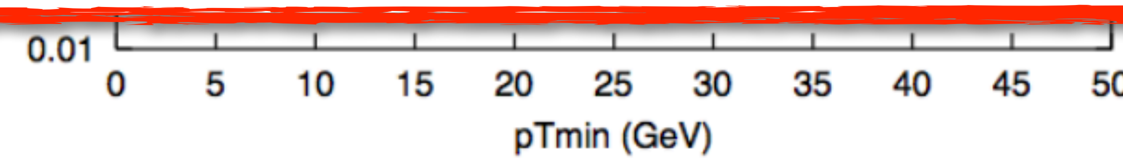


FIG. 12. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs multiple-interaction model with variable impact parameter: solid line, double-Gaussian matter distribution; dashed line, with fix impact parameter [i.e., $O_0(b)$].

(with CTEQ 5L PDF's)



MPI and HI-like effects

Instituto de Ciencias Nucleares UNAM

Striking similarities between numerous observables have been observed across different collision systems at both RHIC and LHC energies, when compared at similar multiplicity

Besides hydrodynamic description, calculations from transport models, hadronic re-scattering, as well as initial state effects have been investigated. Others like Multi-Parton Interactions (MPI), string rope and shoving can also explain some features of data

- Radial flow-like effects emerge in QCD-inspired event generators like Pythia due to color reconnection and MPI, PRL 111 (2013) 042001
- In a model based on the QCD theory of MPI, QCD interference is shown to give rise to values for $v_n\{2\}, v_n\{4\}$, n even, that persists in high N_{mpi} events: B. Block, C. D. Jäkel, M. Strikman, U. A. Wiedemann, JHEP 12 (2017) 074

MPI and HI-like effects

Instituto de Ciencias Nucleares UNAM

Can we infer N_{mpi} (target variable) from a given a set of input variables? → Regression problem

We use a multivariate regression technique based on Boosted Decision Trees (BDT) with gradient boosting training, which is implemented in TMVA (arXiv:physics/0703039)

We use the existing data on p_{T} spectra as a function of multiplicity [OK for MPI studies in minimum-bias pp collisions]

- **Input variables:** Event-by-event average p_{T} of charged particles / Multiplicity

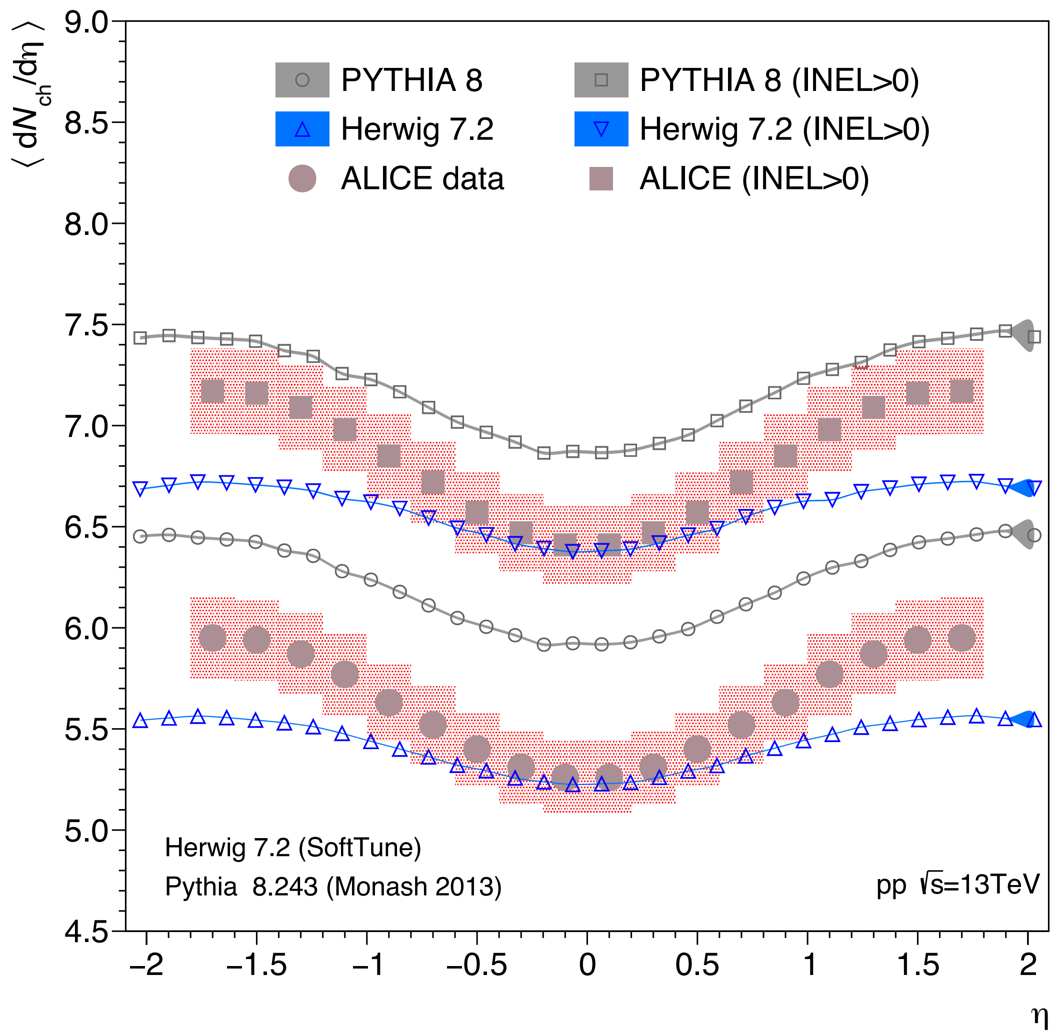
For systematic uncertainties other set of input variables was considered: Charged particle multiplicity in the pseudorapidity region covered by VZERO detector / Transverse spherocity

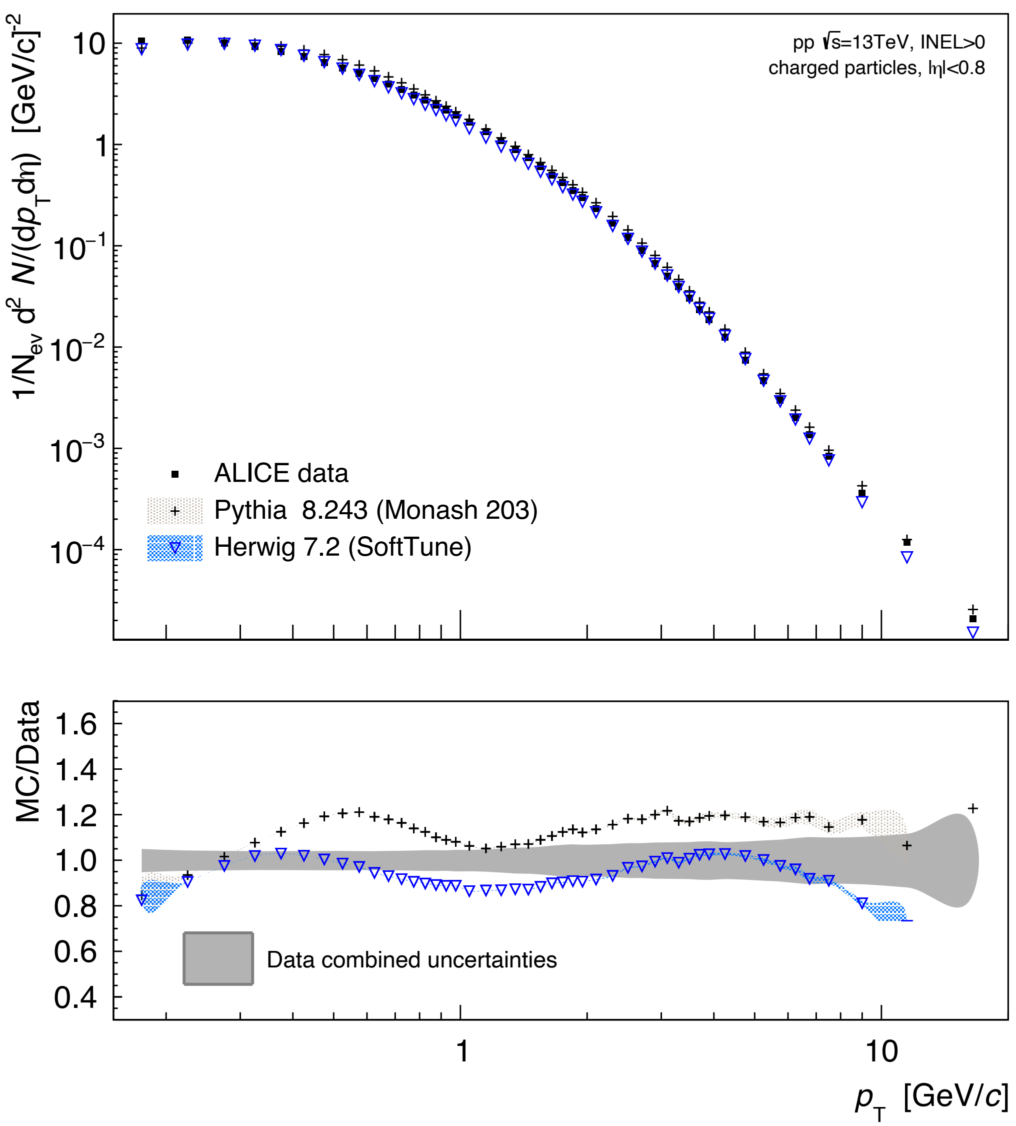
Validation: MB pp collisions

Event generators used for validation

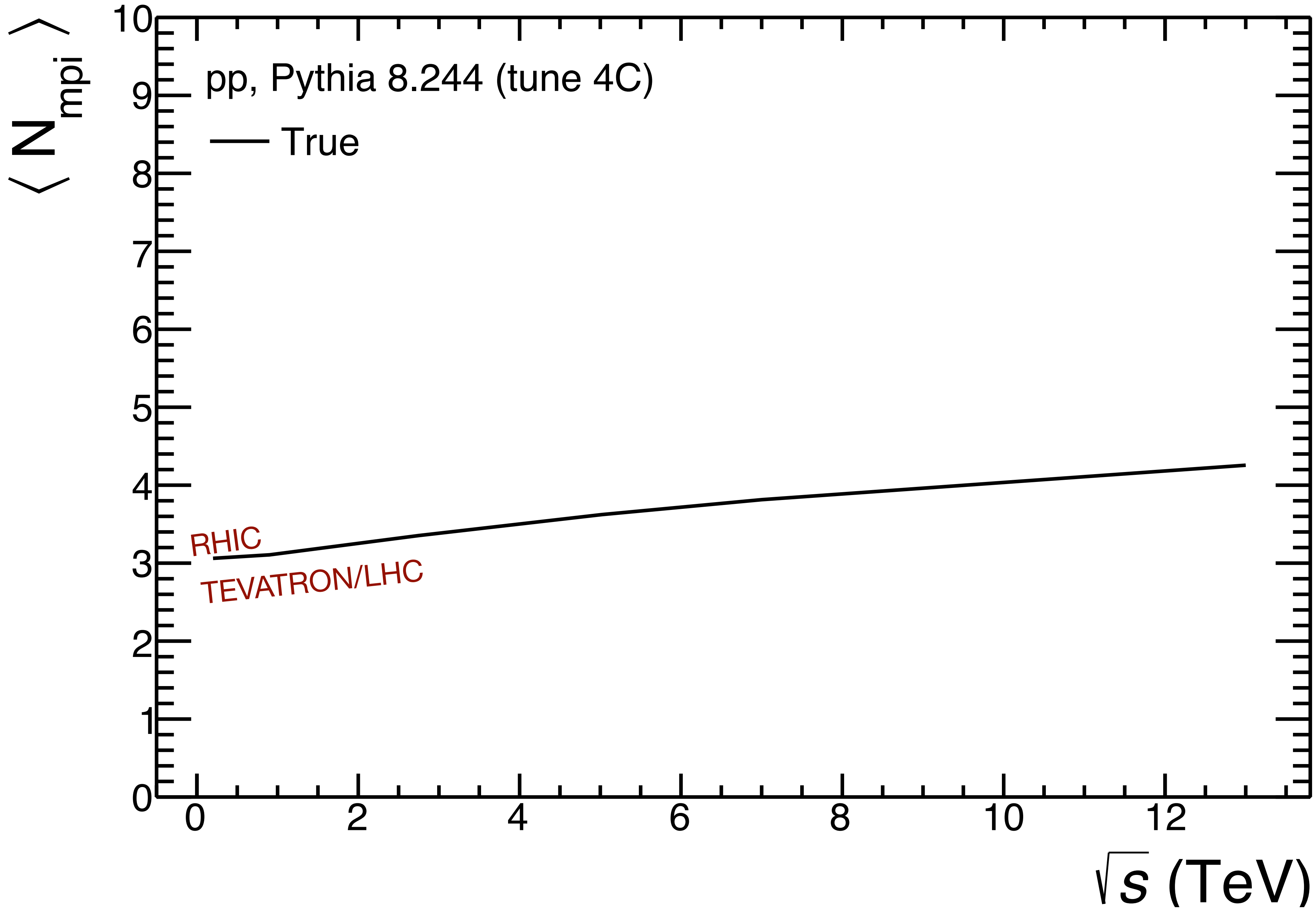
In order to investigate the model dependence (MPI and hadronization models), we use pp collisions simulated with **Herwig 7.2** (soft tune): S. Gieseke, C. Rohr and A. Siodmok, EPJC 72 (2012) 2225

Pythia 8.2: T. Sjöstrand et al., Comput. Phys. Commun. 191 (2015) 159-177

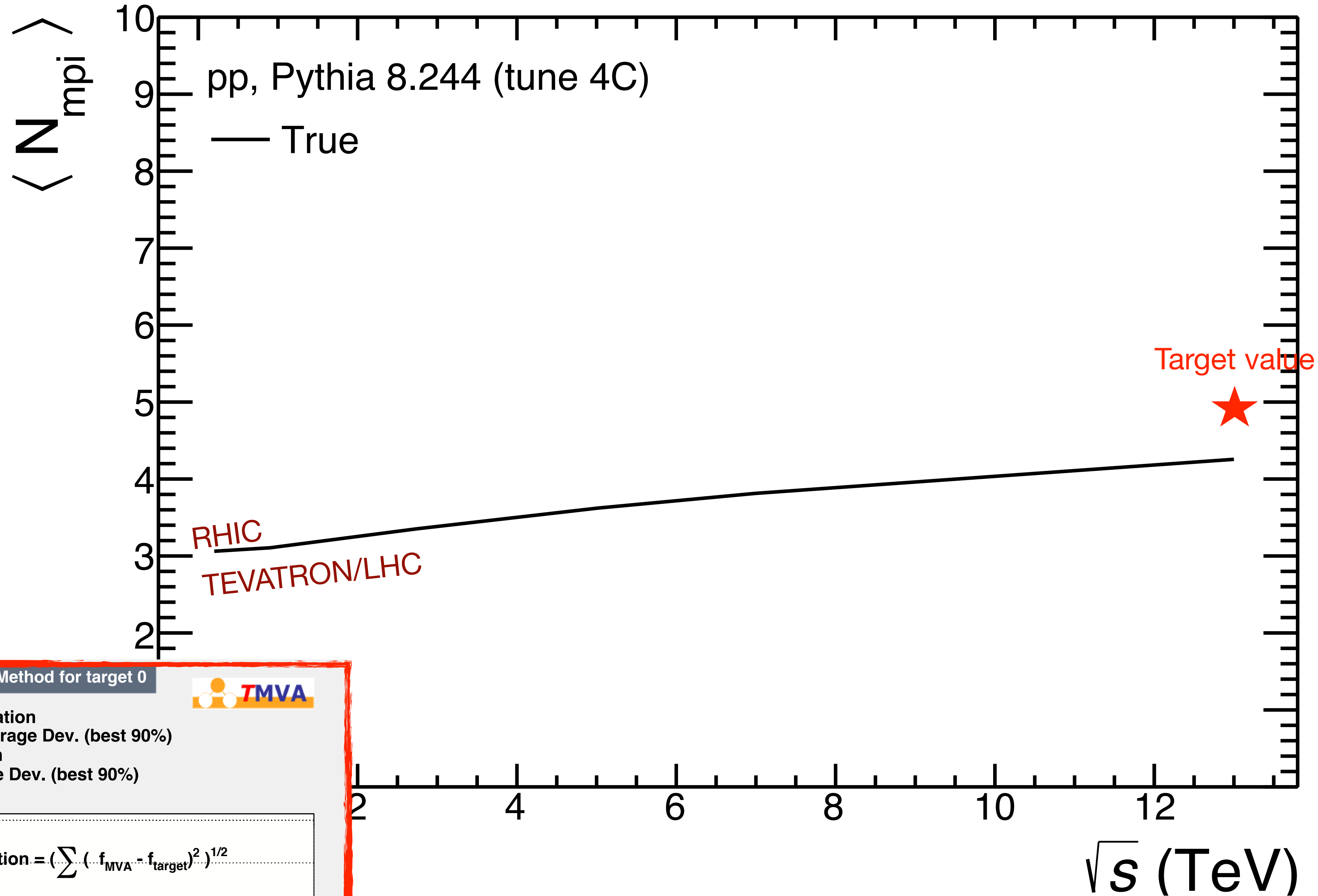


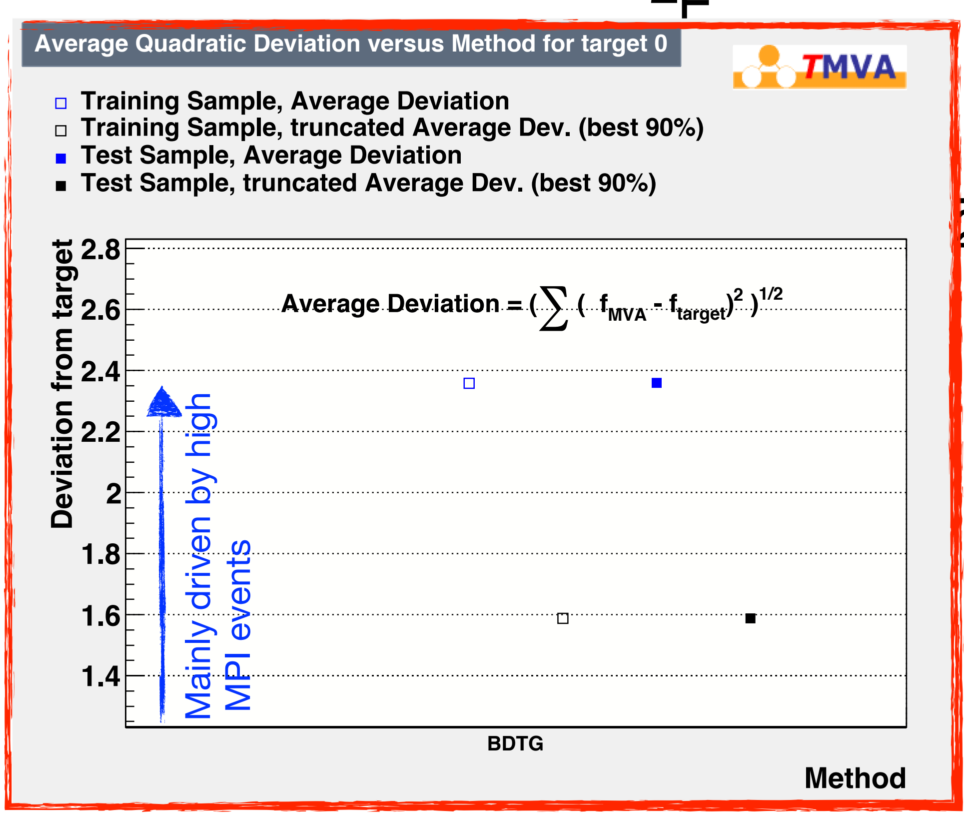


MPI vs $\sqrt{s}$ Pythia 8.244



MPI from BDT





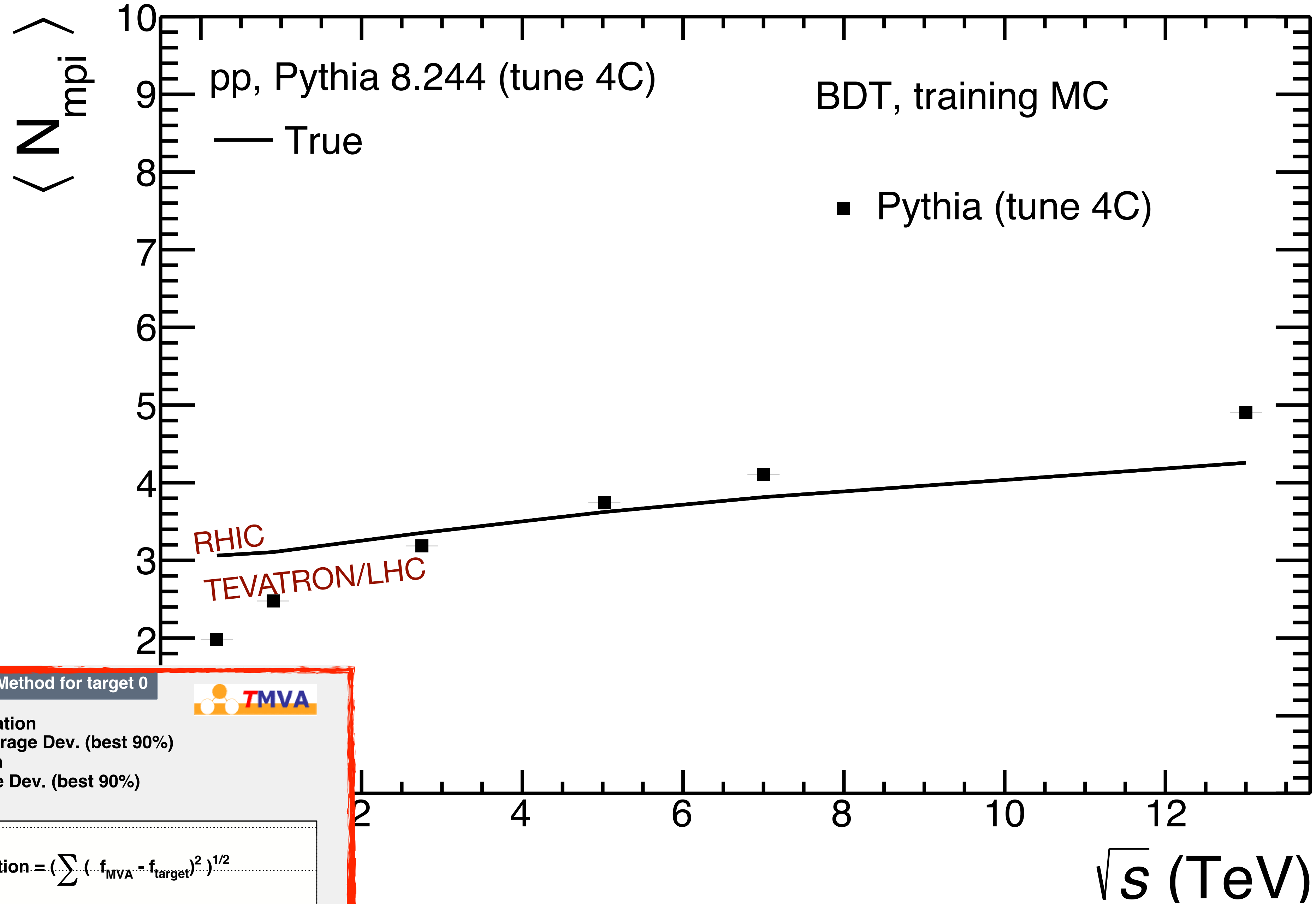
TRAINING (MB pp collisions at 13 TeV, Pythia tune 4C)

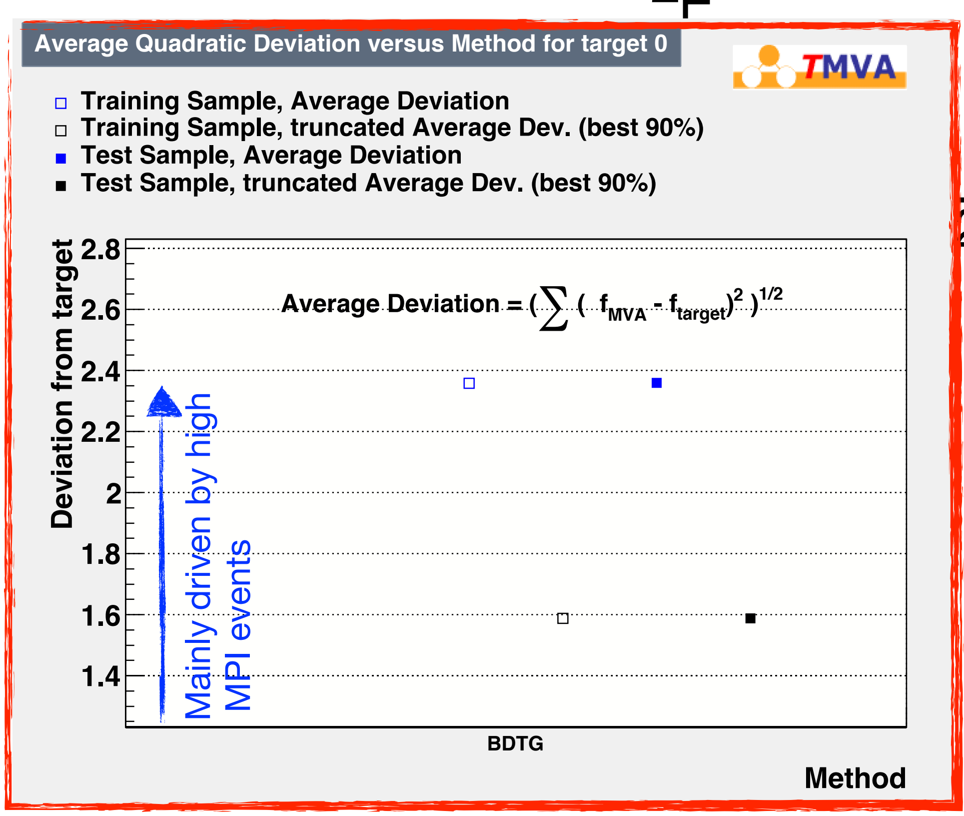
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_{\rm T} \geq 0.15$ GeV/c)
- multiplicity

TEST (MB pp collisions at $\sqrt{s} = 13$ TeV, Pythia tune 4C)

MPI from BDT





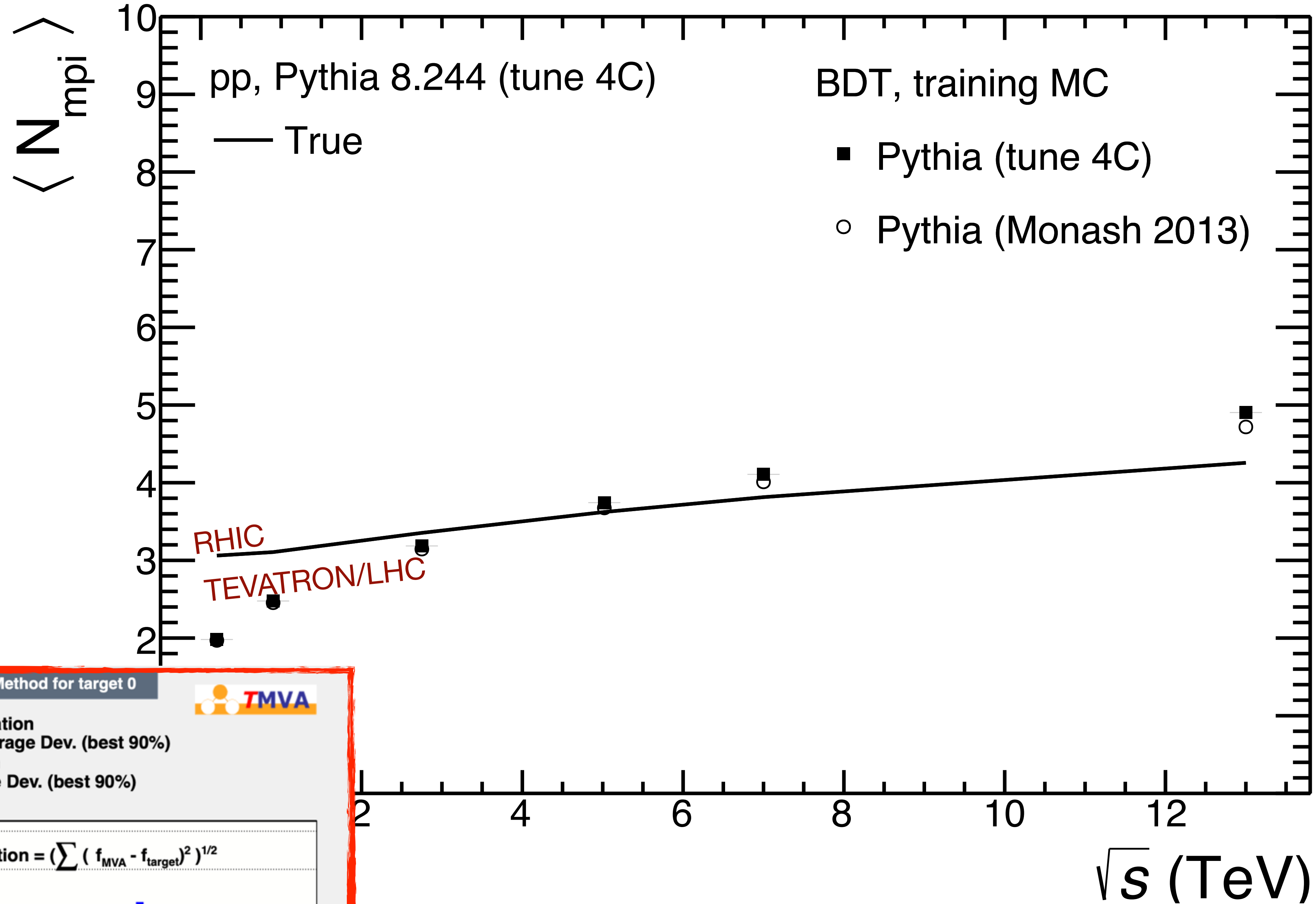
TRAINING (MB pp collisions at 13 TeV, Pythia tune 4C)

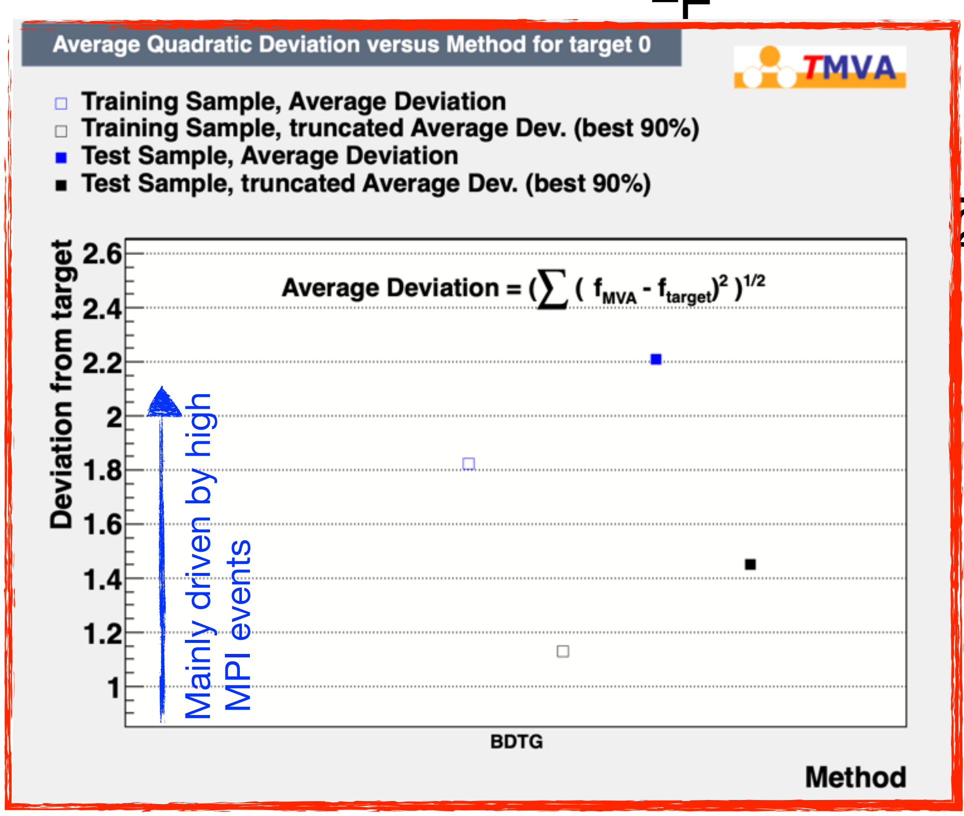
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_T \geq 0.15$ GeV/c)
- multiplicity

TEST (MB pp collisions at different $\sqrt{s}$, Pythia tune 4C)

MPI from BDT (model dependence)





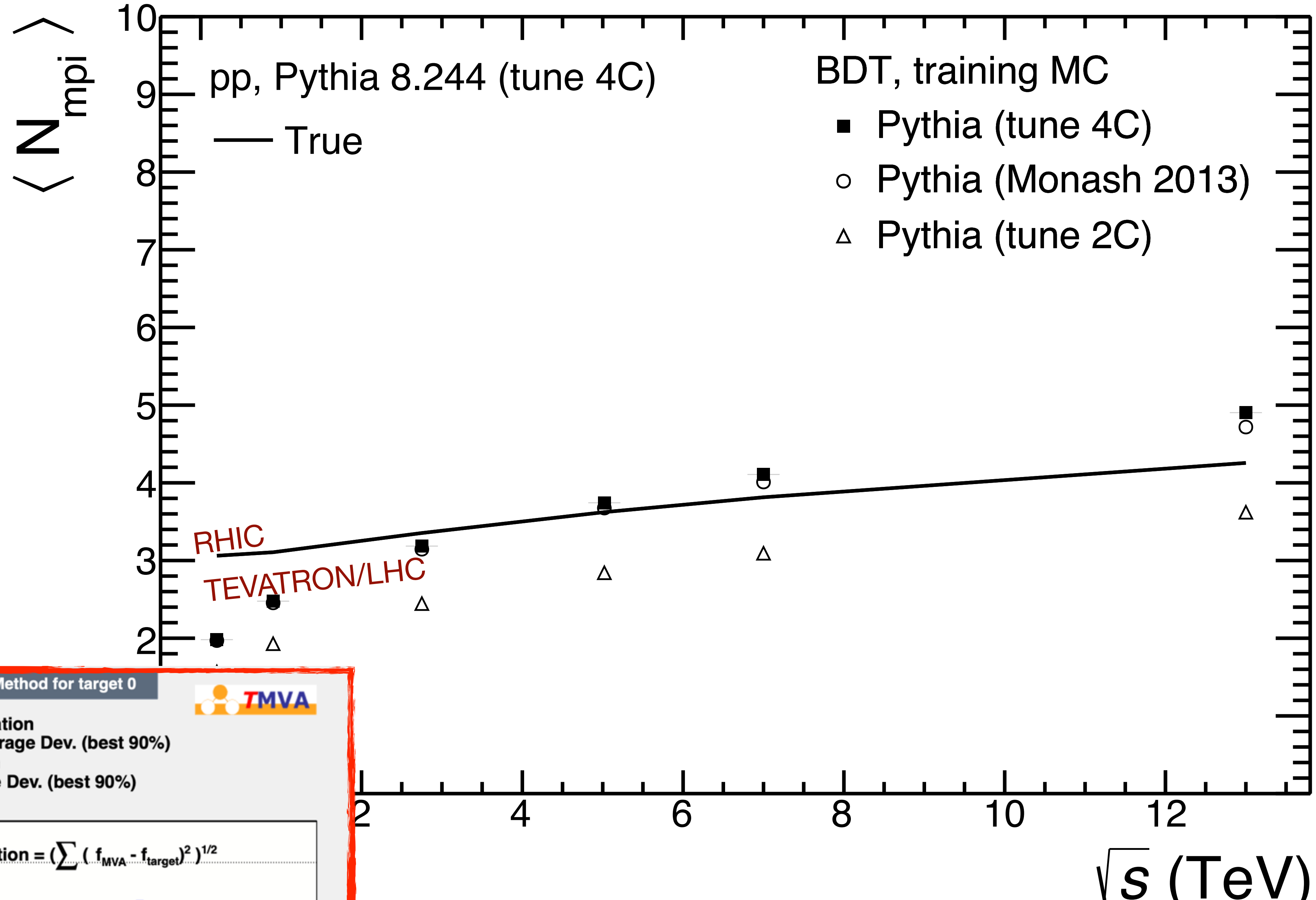
TRAINING (MB pp collisions at 13 TeV, Pythia Monash)

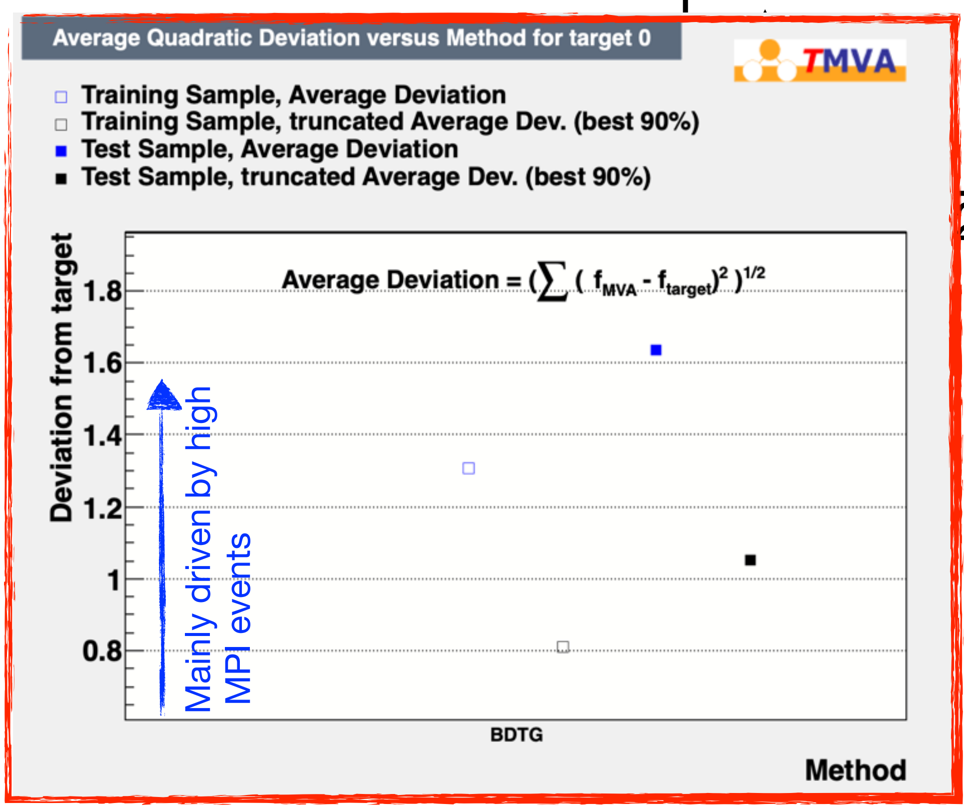
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_{\mathrm{T}} \geq 0.15$ GeV/c)
- multiplicity

TEST (MB pp collisions at different $\sqrt{s}$, Pythia tune 4C)

MPI from BDT (model dependence)





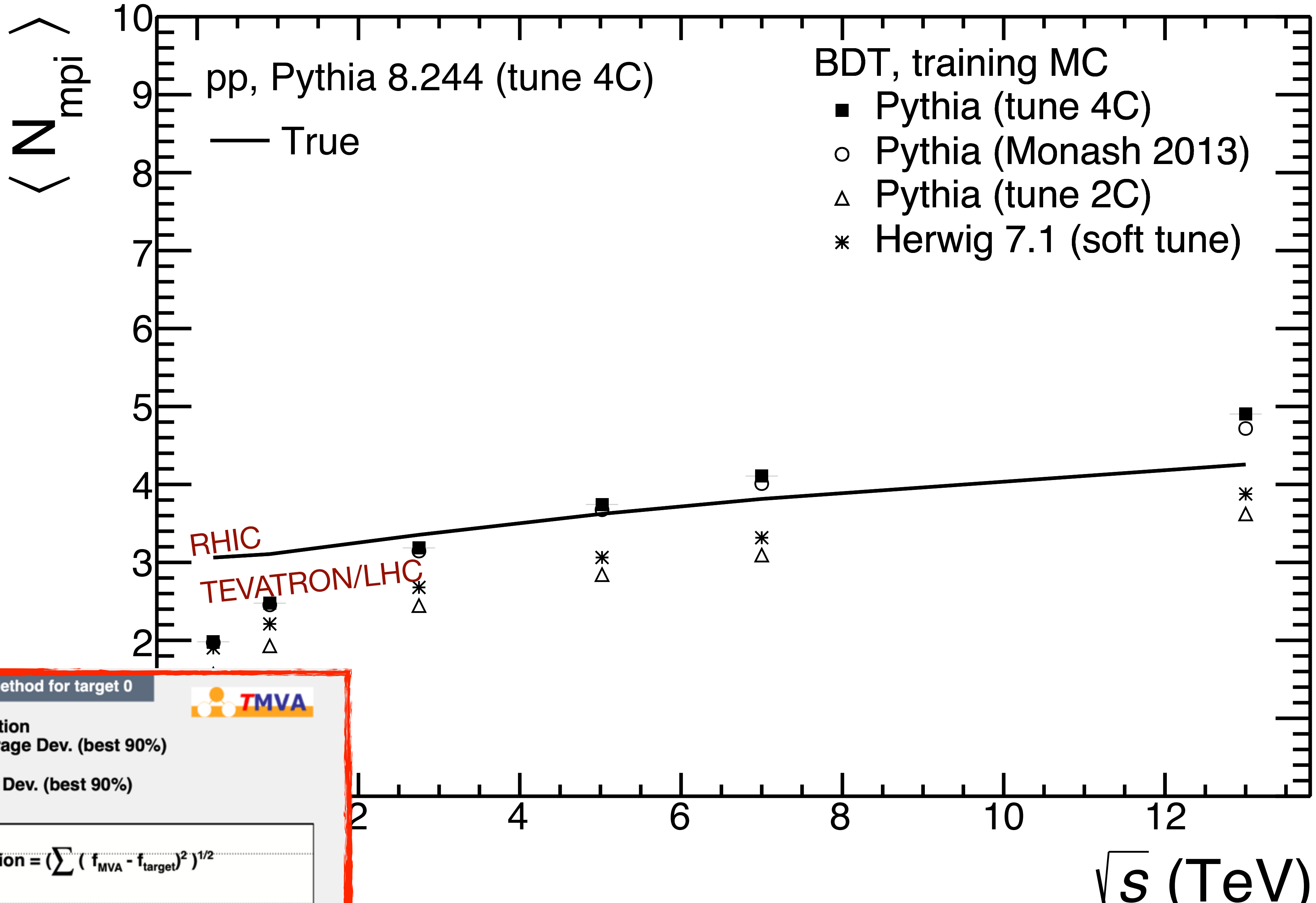
TRAINING (MB pp collisions at 13 TeV, Pythia tune 2C)

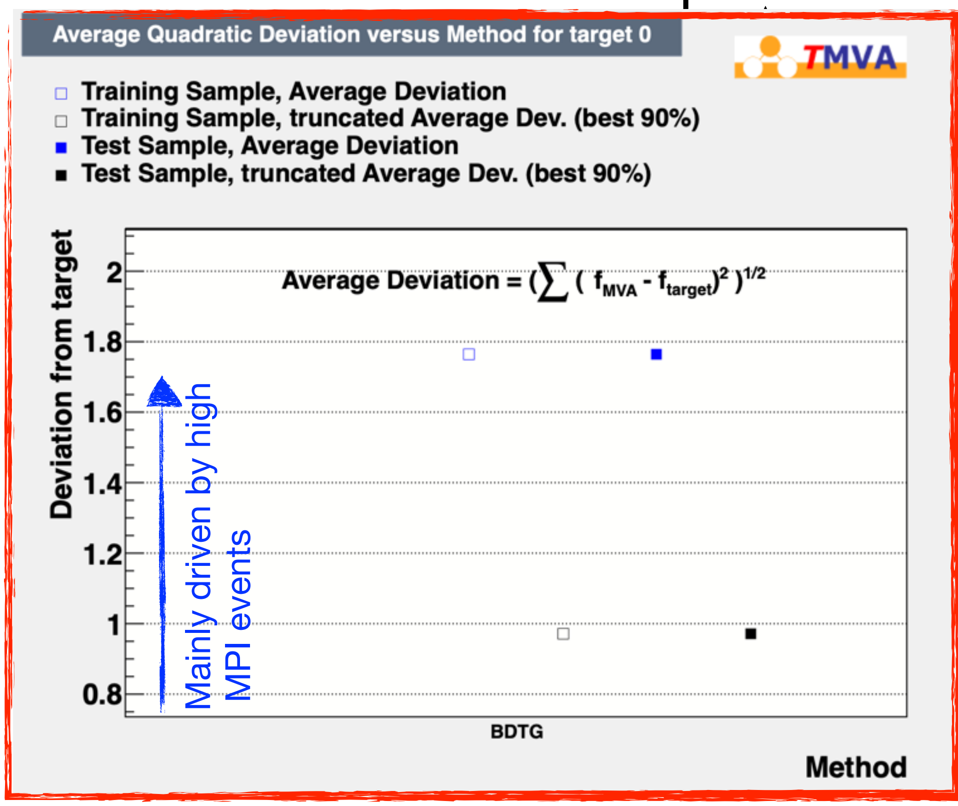
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_{\rm T} \geq 0.15$ GeV/c)
- multiplicity

TEST (MB pp collisions at different $\sqrt{s}$, Pythia tune 4C)

Effect of hadronization model I





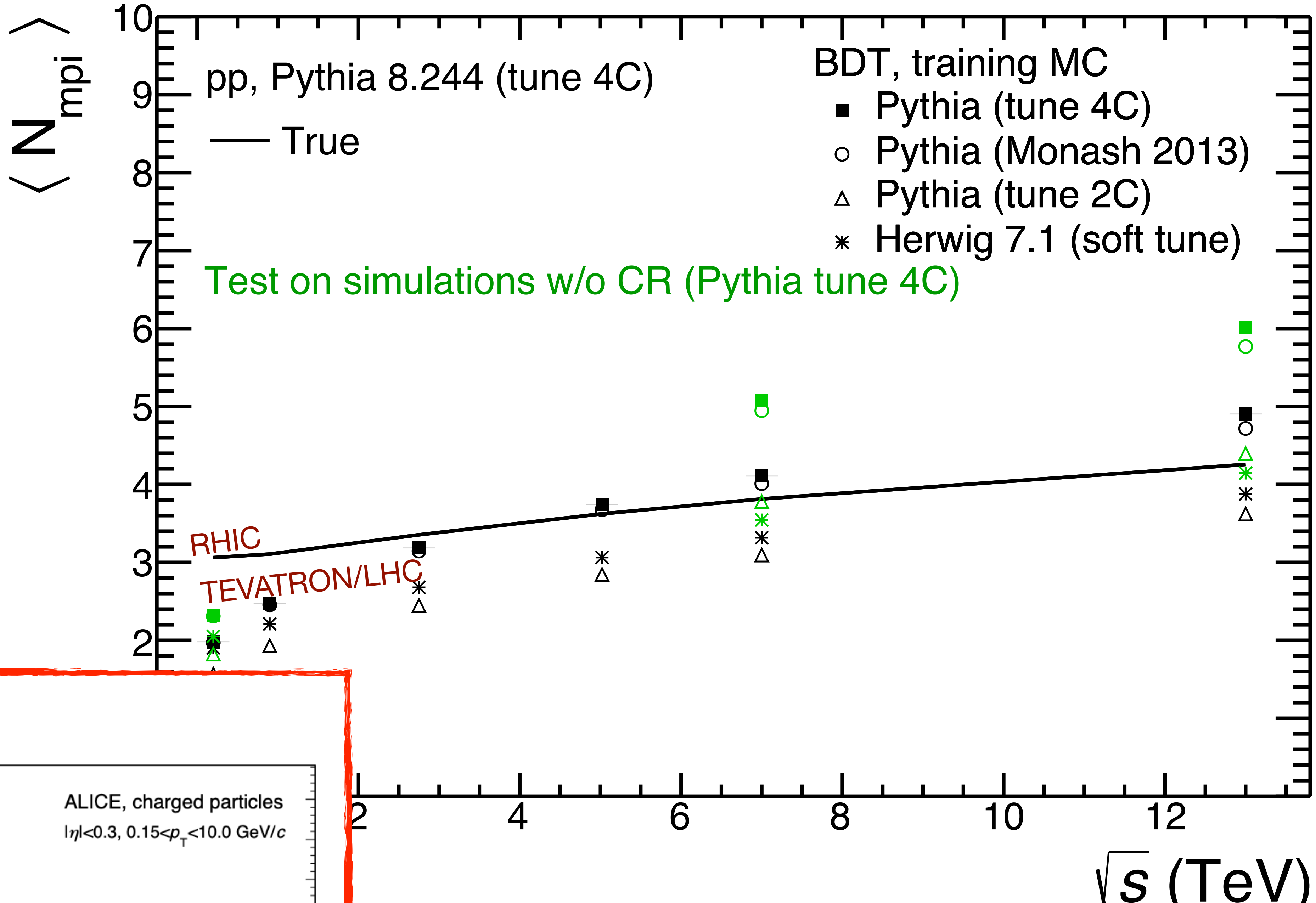
TRAINING (MB pp collisions at 13 TeV, Herwig soft tune)

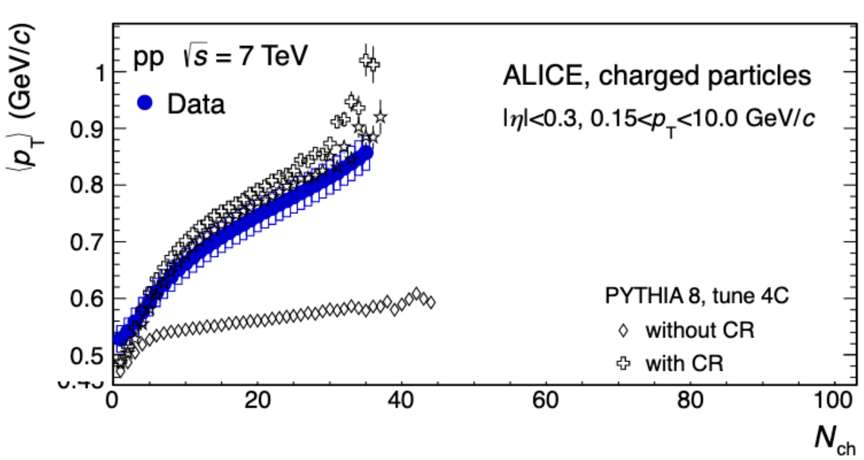
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_T \geq 0.15$ GeV/c)
- multiplicity

TEST (MB pp collisions at different $\sqrt{s}$, Pythia tune 4C)

Effect of hadronization model II





Simulations w/o CR can not describe the data

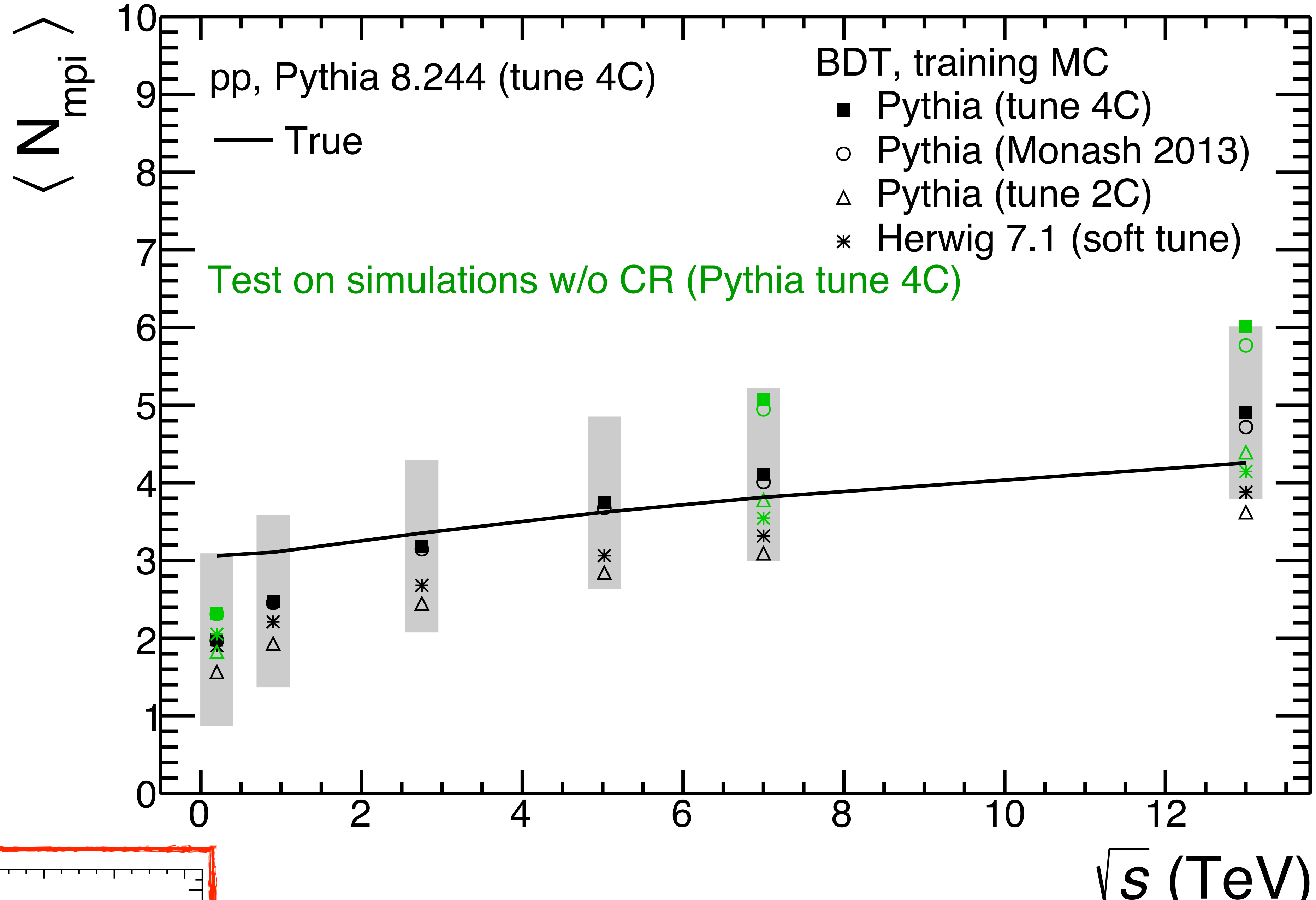
TRAINING (MB pp collisions at 13 TeV, Pythia tune 4C)

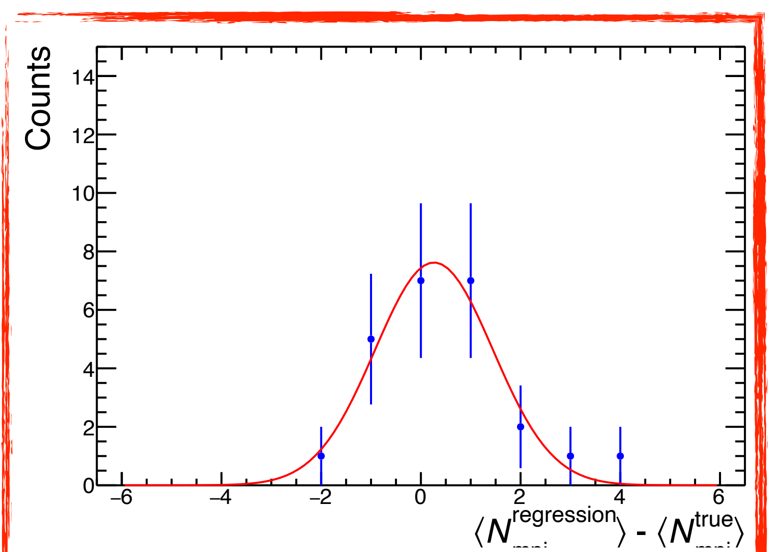
Input variables (primary charged particles $|\eta| < 0.8$):

- EbE average transverse momentum ($p_{\rm T} \geq 0.15$ GeV/c)
- multiplicity

TEST (Pythia, tune 4C w/o Color Reconnection)

Systematic uncertainties

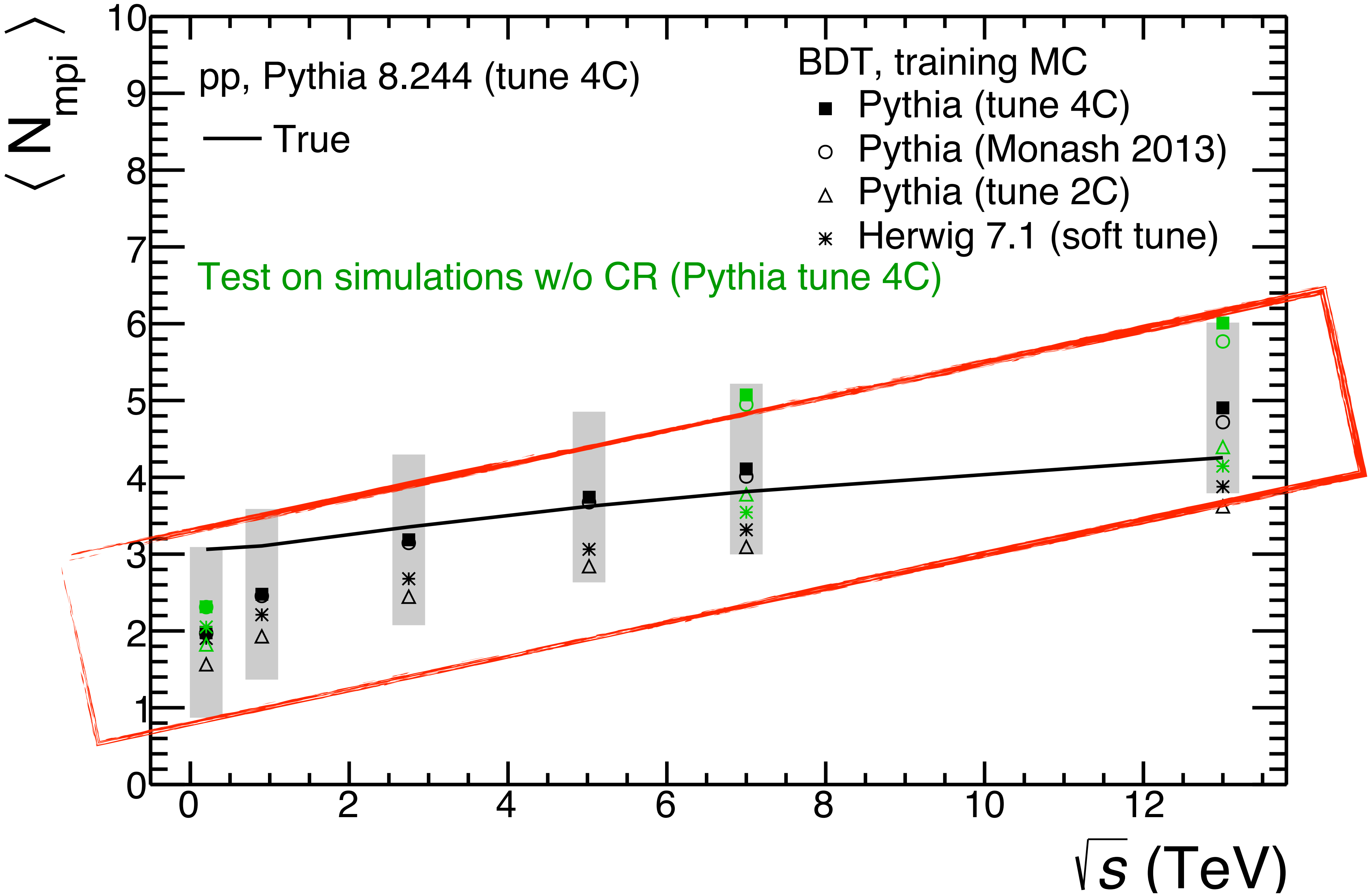




The analysis was repeated considering the following variations:

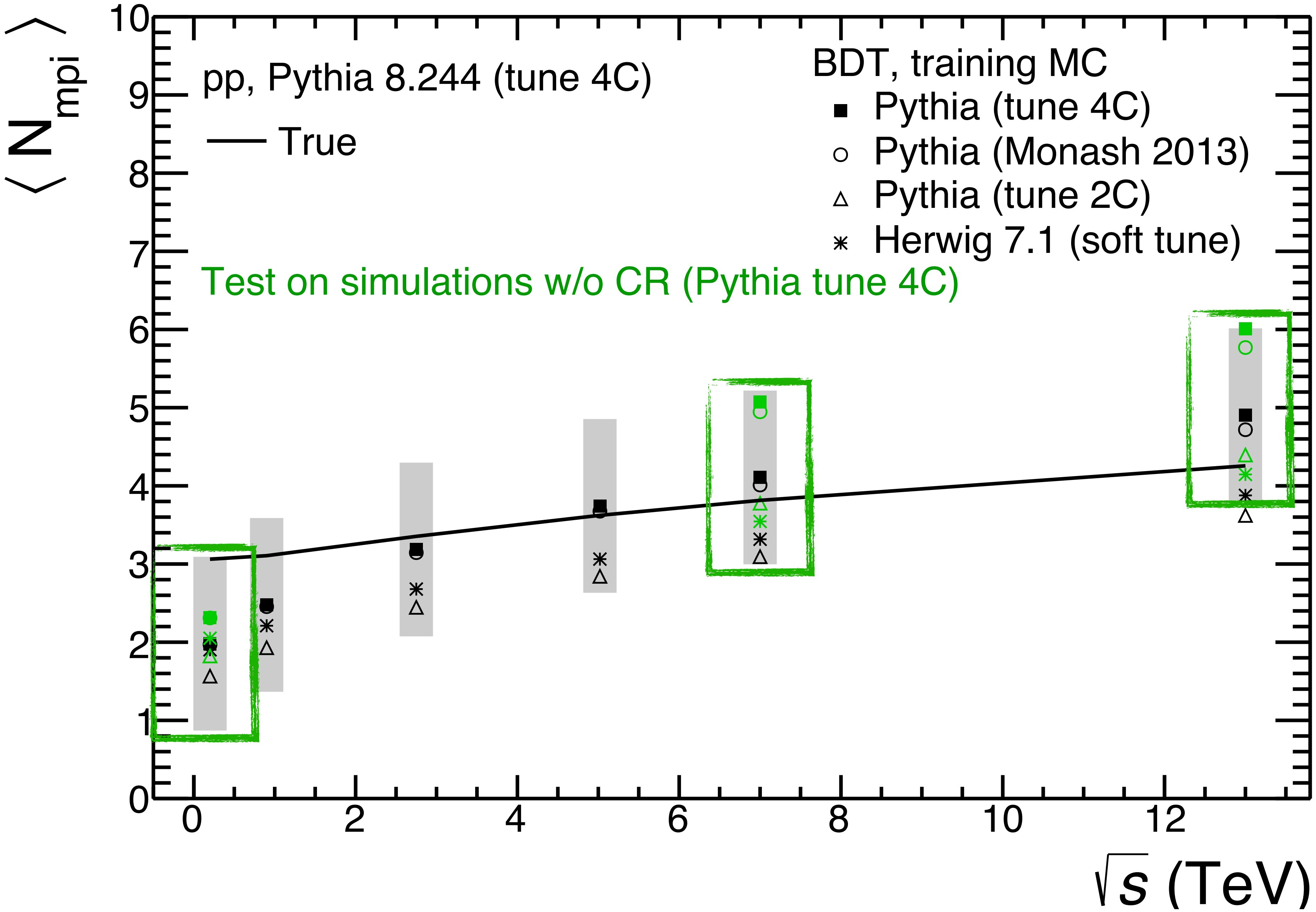
- A different set of input variables was used (spherocity, EbE average transverse momentum and multiplicity in the pseudorapidity region covered by the ALICE VZERO detector)
- Assuming a different MPI distribution (flat)
- All center-of-mass energies were considered

Systematic uncertainties



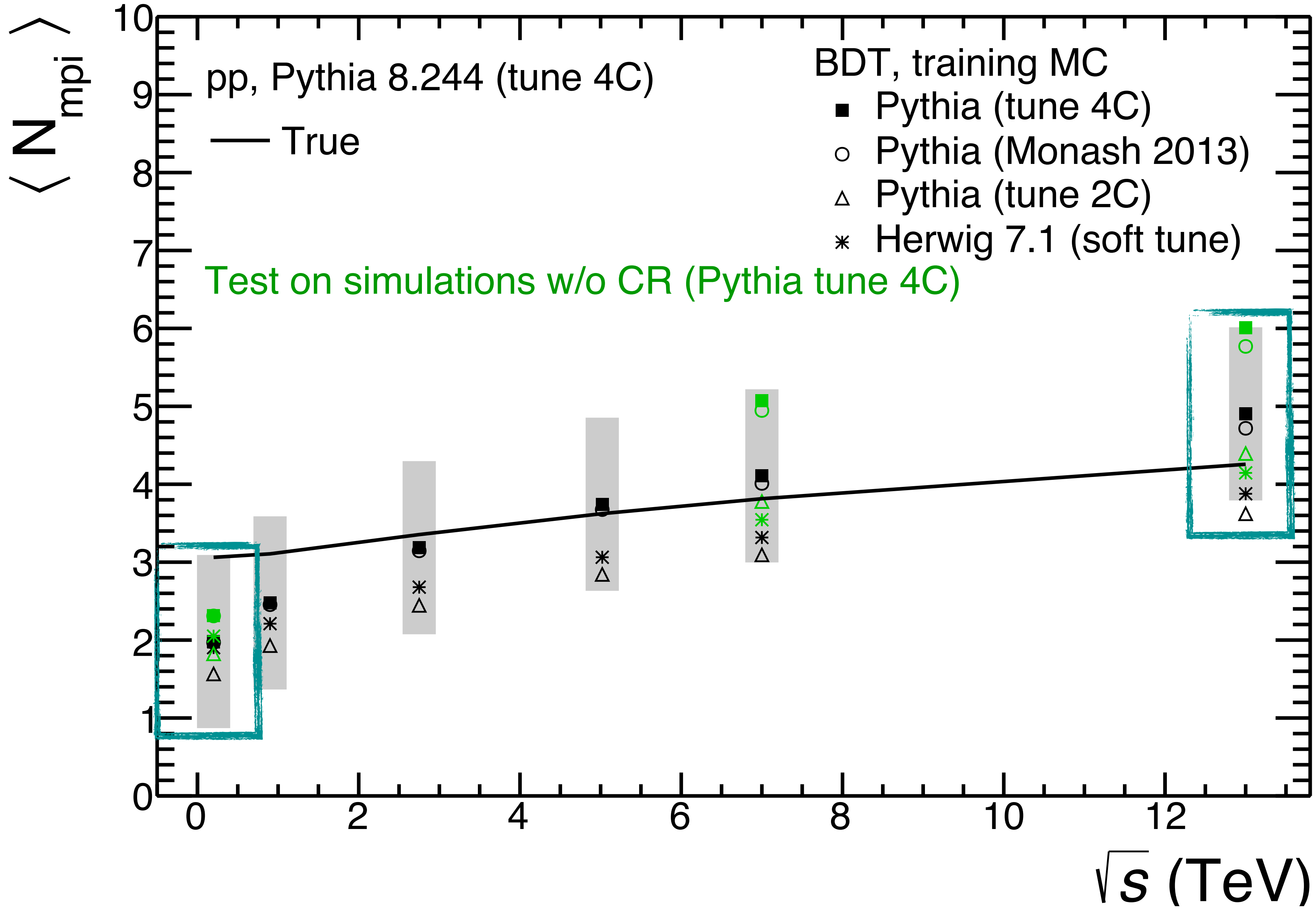
- Within uncertainties, we recover the modest center-of-mass energy dependence of $\langle N_{\mathrm{mpi}} \rangle$

Systematic uncertainties



- Within uncertainties, we recover the modest center-of-mass energy dependence of $\langle N_{\rm mpi} \rangle$
- Within uncertainties, $\langle N_{\rm mpi} \rangle$ is found to be independent of CR (expected behaviour)

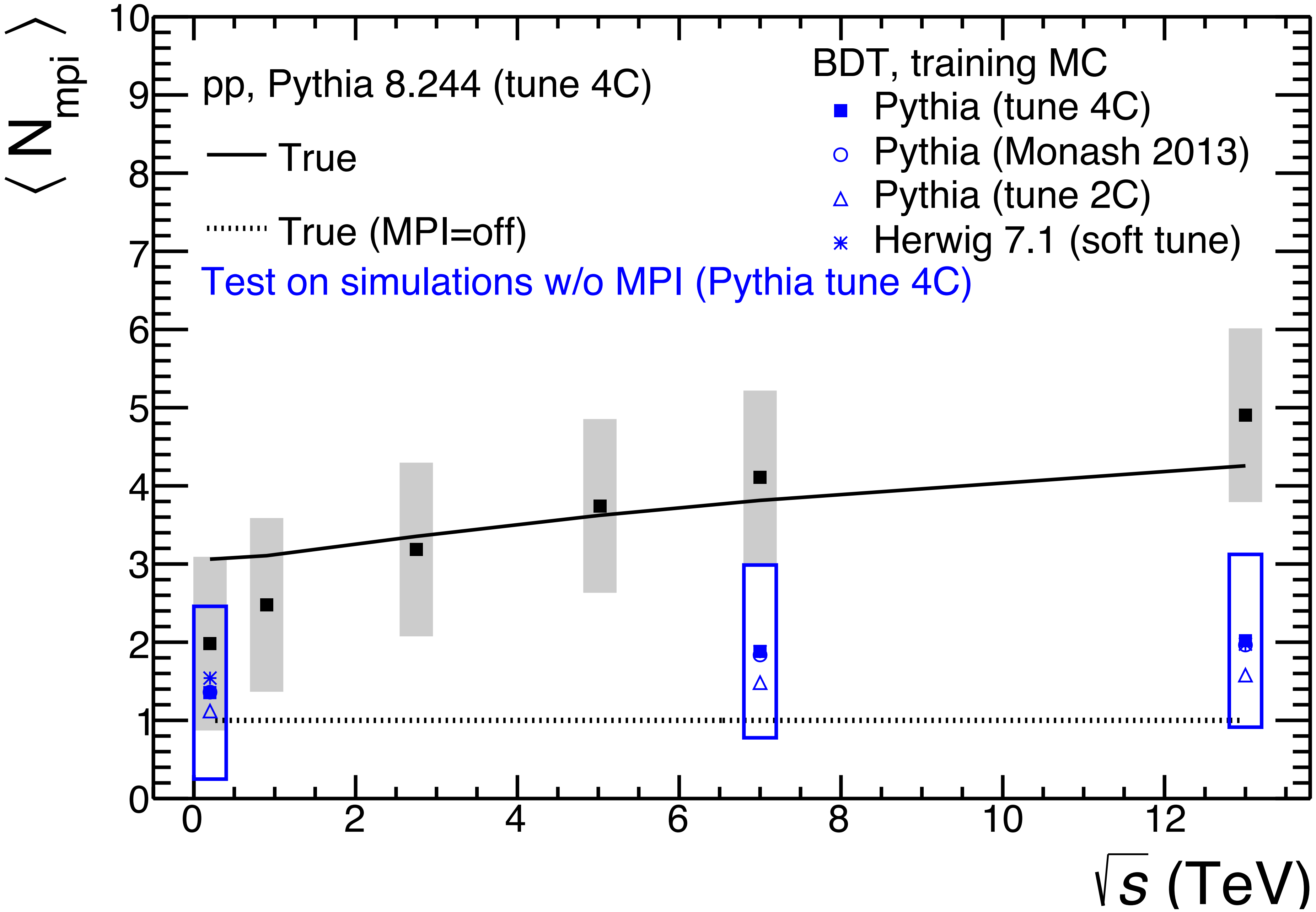
Systematic uncertainties



- Within uncertainties, we recover the modest center-of-mass energy dependence of $\langle N_{\mathrm{mpi}} \rangle$
- Within uncertainties, $\langle N_{\mathrm{mpi}} \rangle$ is found to be independent of CR (expected behaviour)
- The model dependence is well covered by the systematic uncertainties

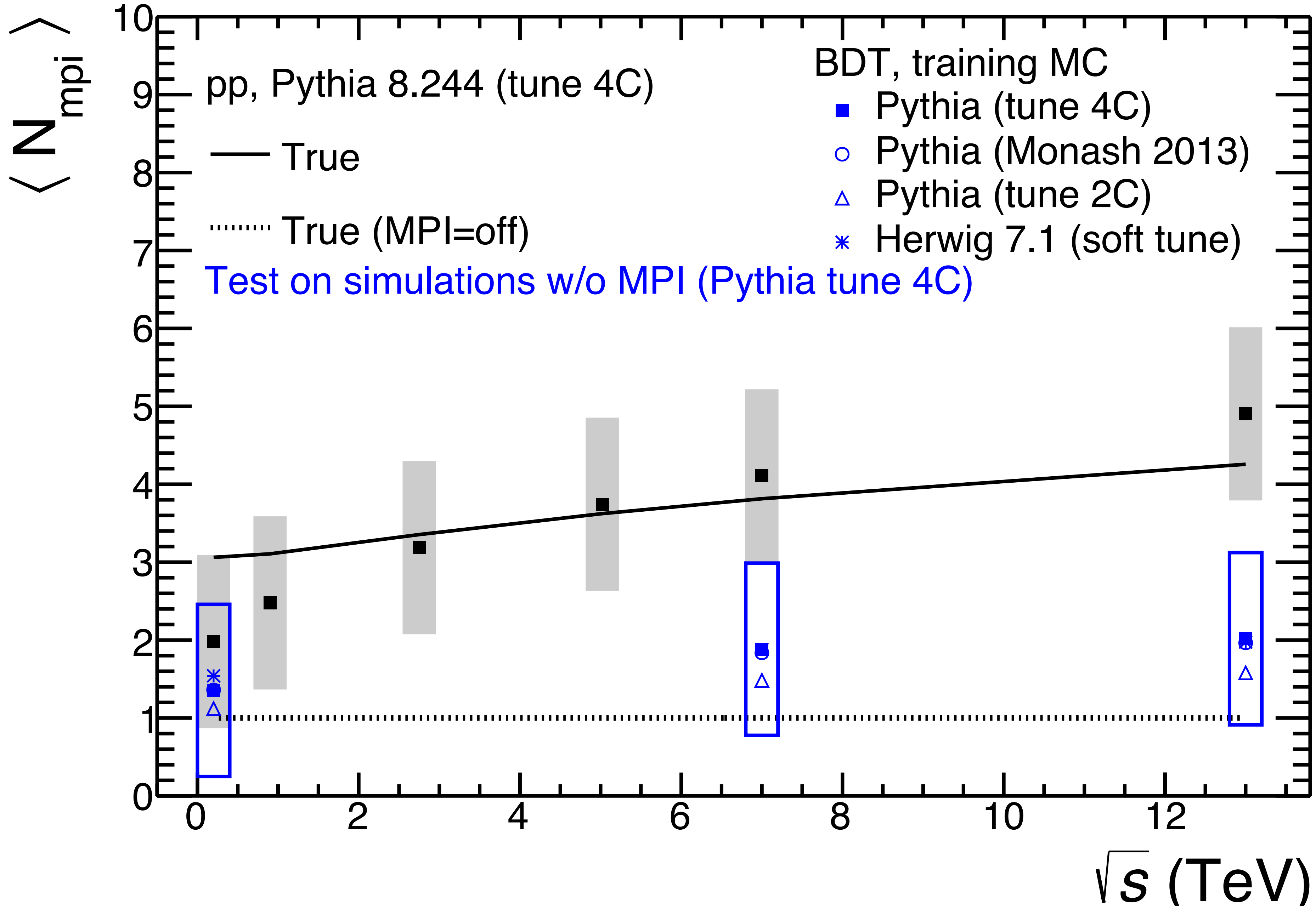
Systematic uncertainties

Instituto de Ciencias Nucleares UNAM



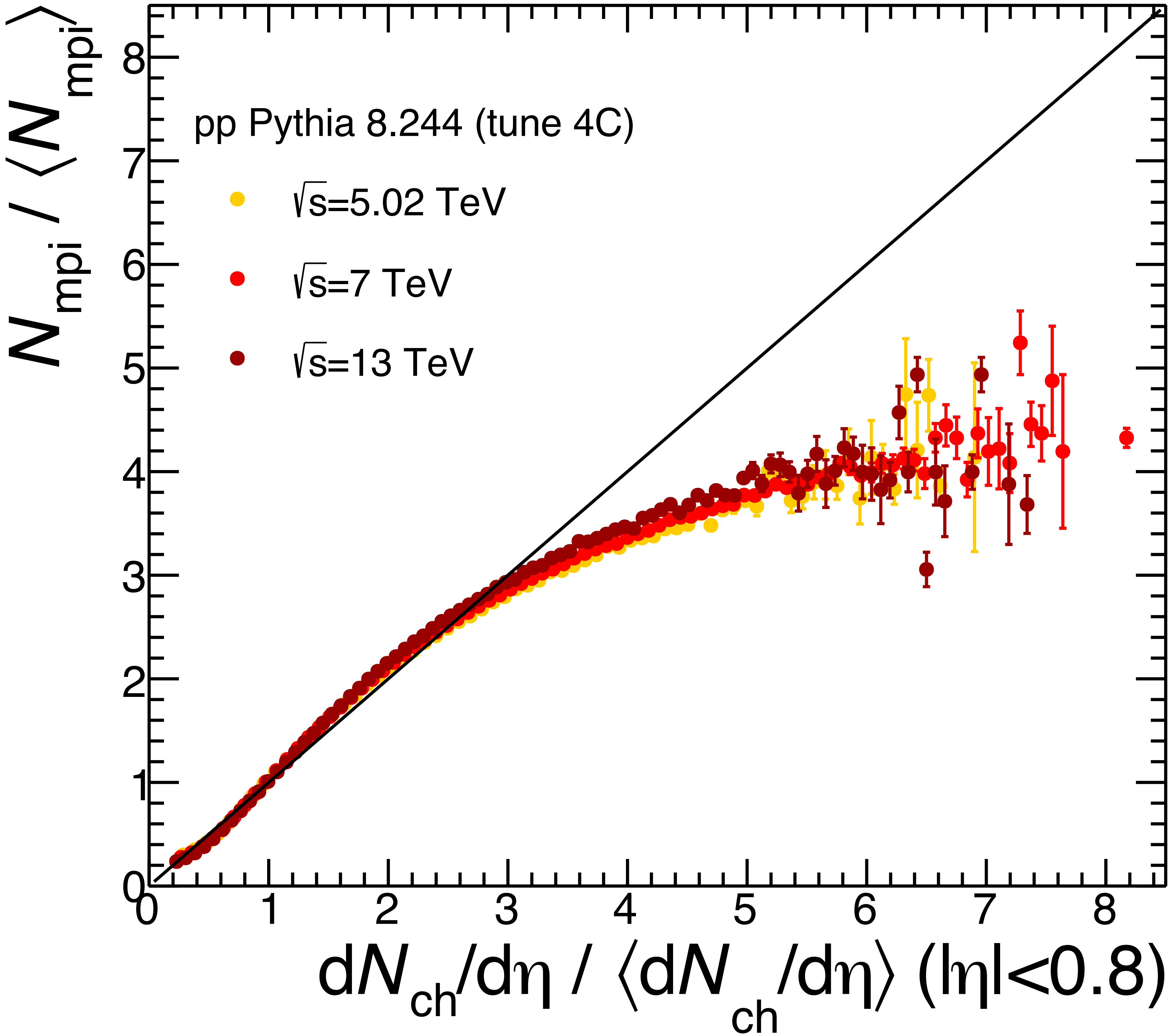
- Within uncertainties, we recover the modest center-of-mass energy dependence of $\langle N_{\mathrm{mpi}} \rangle$
- Within uncertainties, $\langle N_{\mathrm{mpi}} \rangle$ is found to be independent of CR (expected behaviour)
- The model dependence is well covered by the systematic uncertainties
- In simulations with MPI=off, within uncertainties, the regression value is consistent with one

Systematic uncertainties

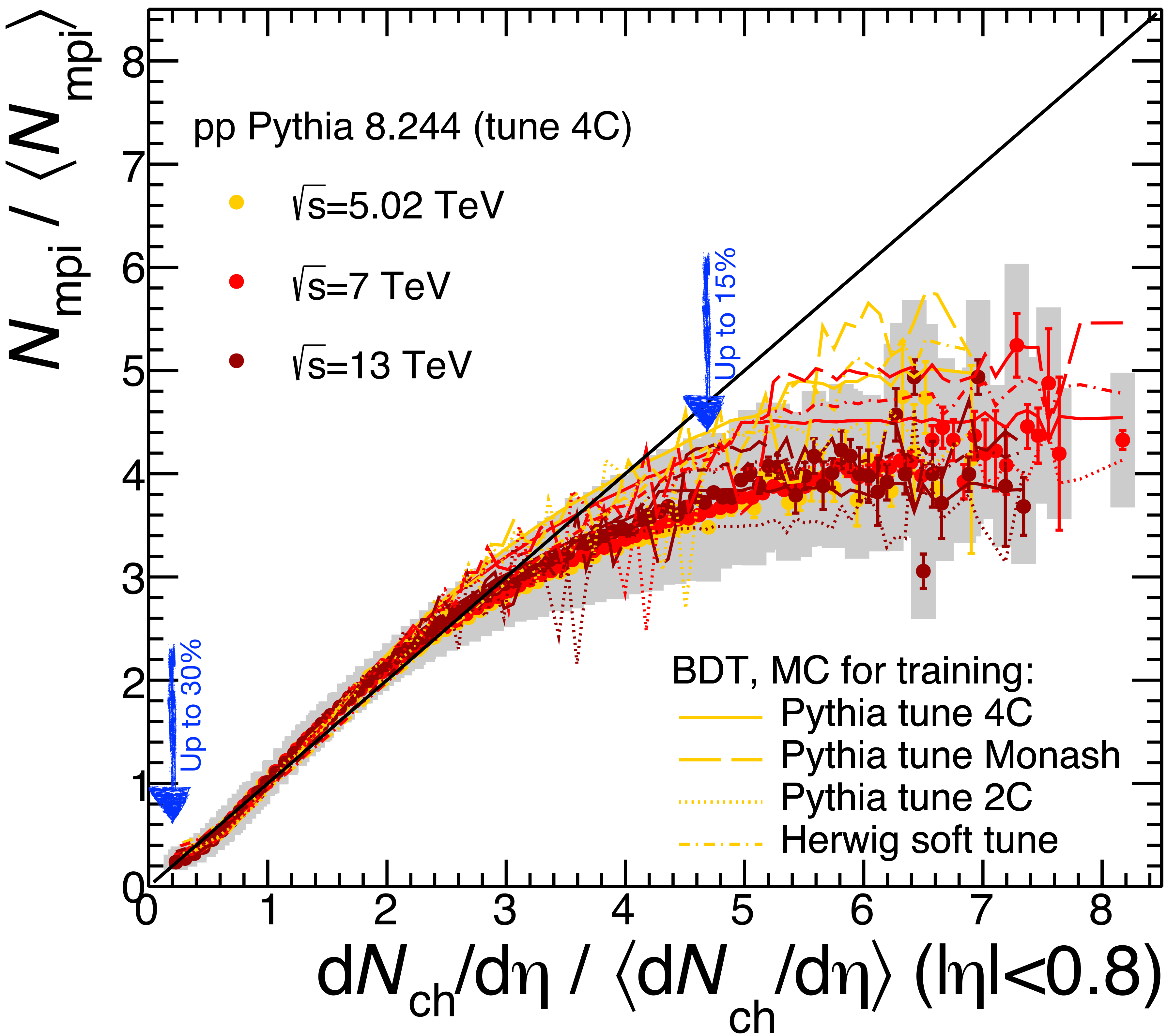


The method has been validated using minimum-bias pp collisions simulated with Pythia and Herwig, the systematic uncertainties cover the model dependence

N_{ch} dependence of N_{mpi}



MC closure test + Syst. Unc.

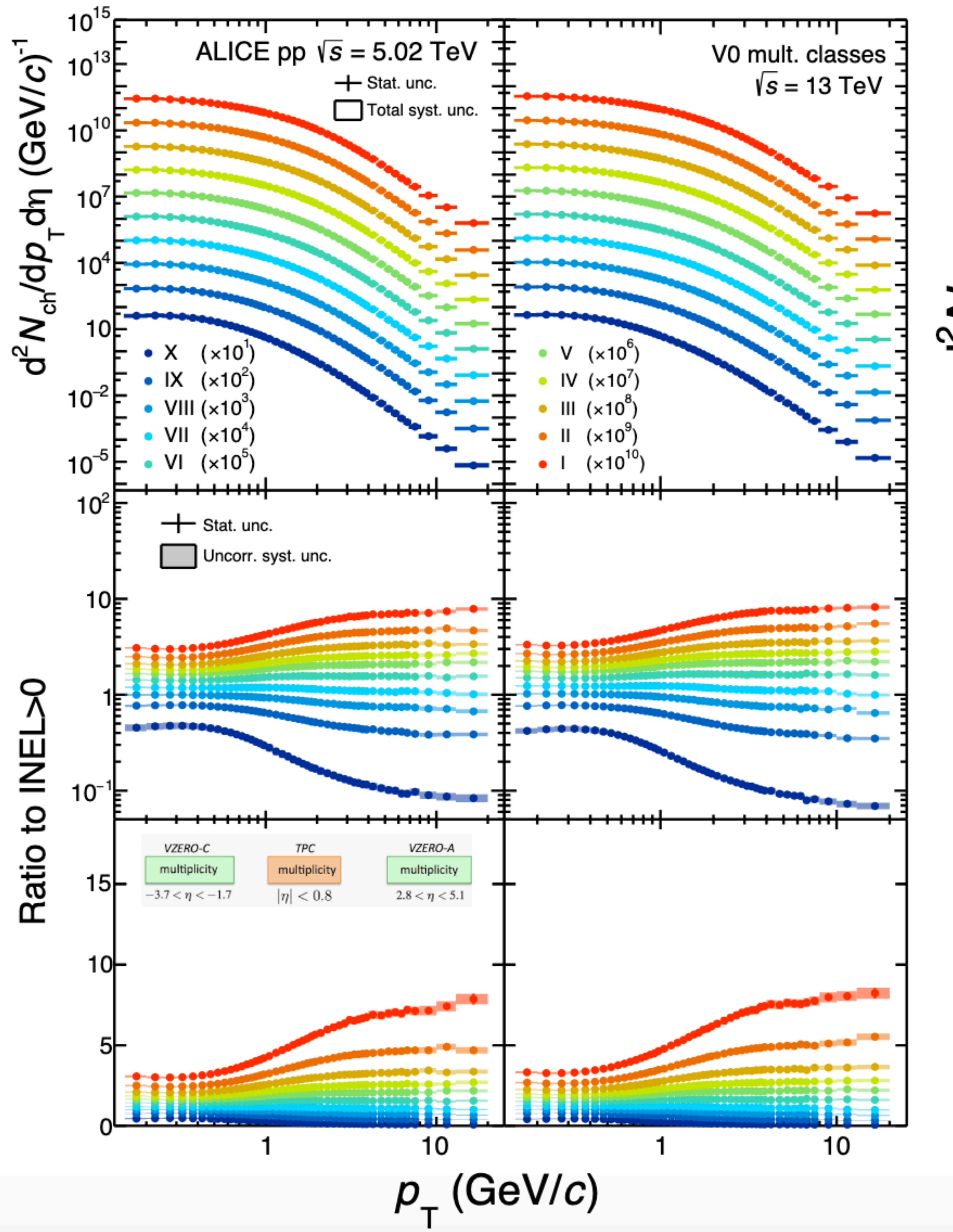


The systematic uncertainties cover the model dependence

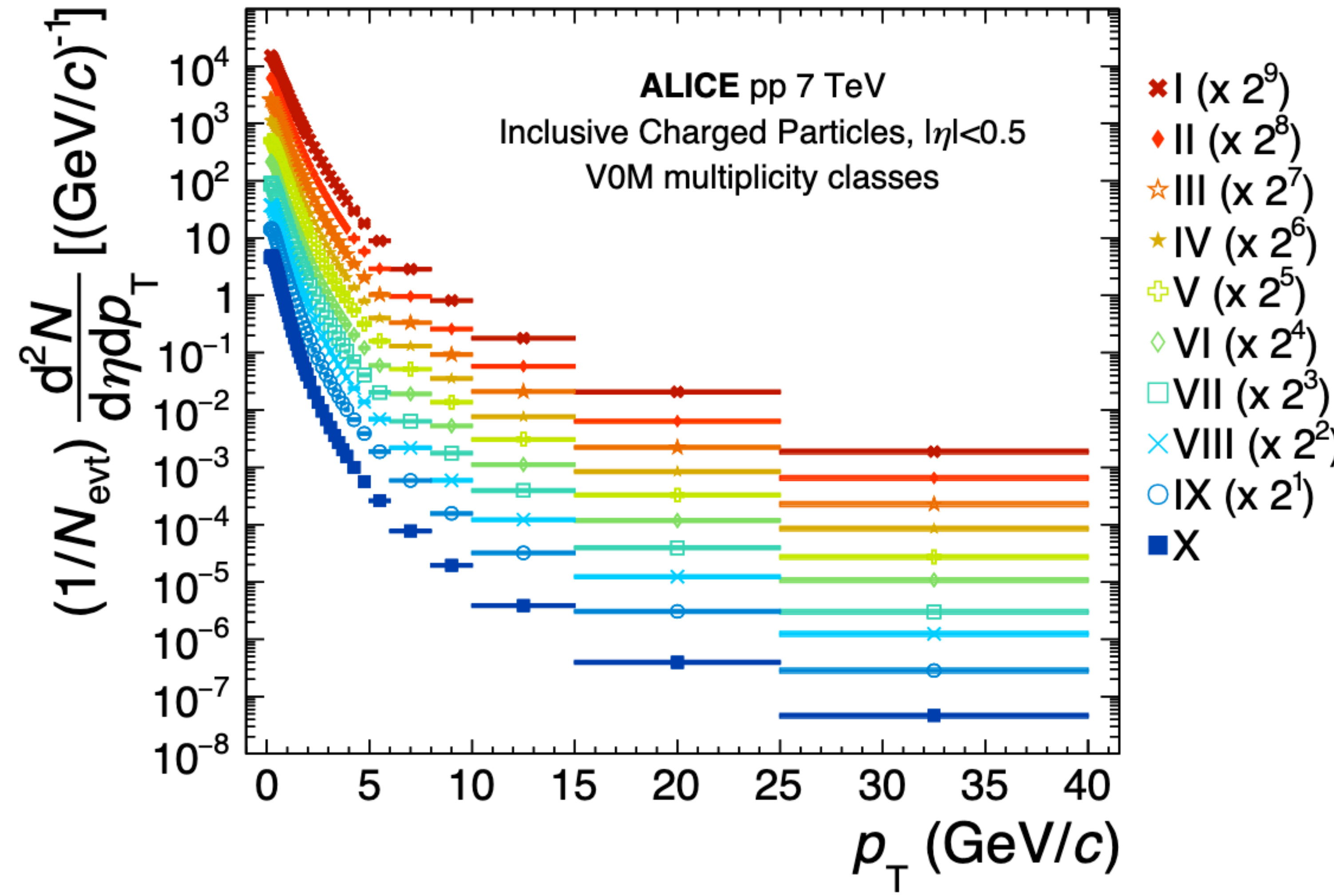
Extraction of N_{mpi} from existing LHC data

ALICE data

ALICE, EPJC 79 (2019) no. 10, 857



ALICE, PRC 99 (2019) 024906



Example, V0M multiplicity classes pp data at 7 TeV:

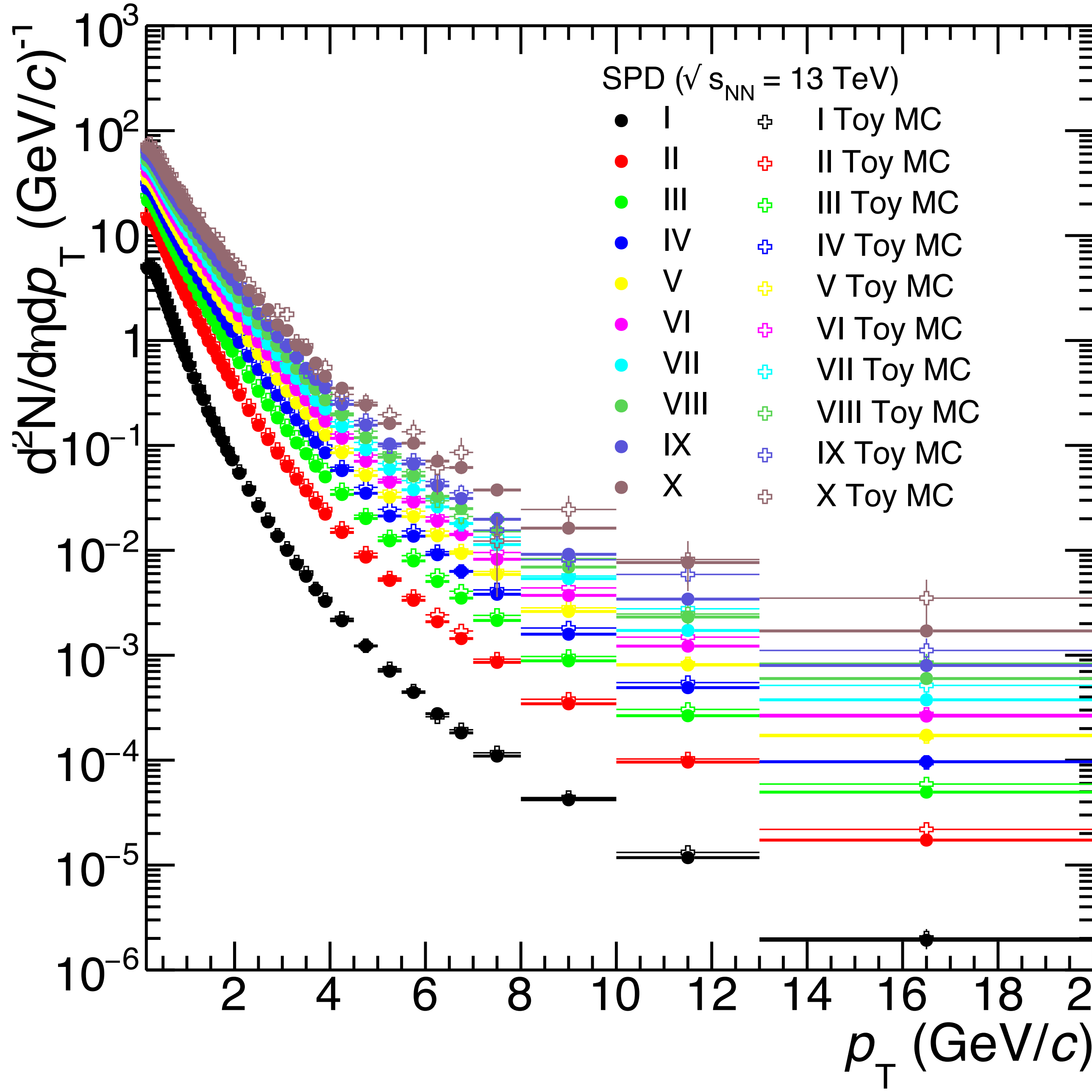
Multiplicity class	I	II	III	IV	V
$\sigma/\sigma_{INEL>0}$	0-0.95%	0.95-4.7%	4.7-9.5%	9.5-14%	14-19%
$\langle dN_{ch}/d\eta \rangle$	21.3 ± 0.6	16.5 ± 0.5	13.5 ± 0.4	11.5 ± 0.3	10.1 ± 0.3
Multiplicity class	VI	VII	VIII	IX	X
$\sigma/\sigma_{INEL>0}$	19-28%	28-38%	38-48%	48-68%	68-100%
$\langle dN_{ch}/d\eta \rangle$	8.45 ± 0.25	6.72 ± 0.21	5.40 ± 0.17	3.90 ± 0.14	2.26 ± 0.12

Instituto de Ciencias Nucleares UNAM

ALICE, EPJC 79 (2019) no. 10, 857

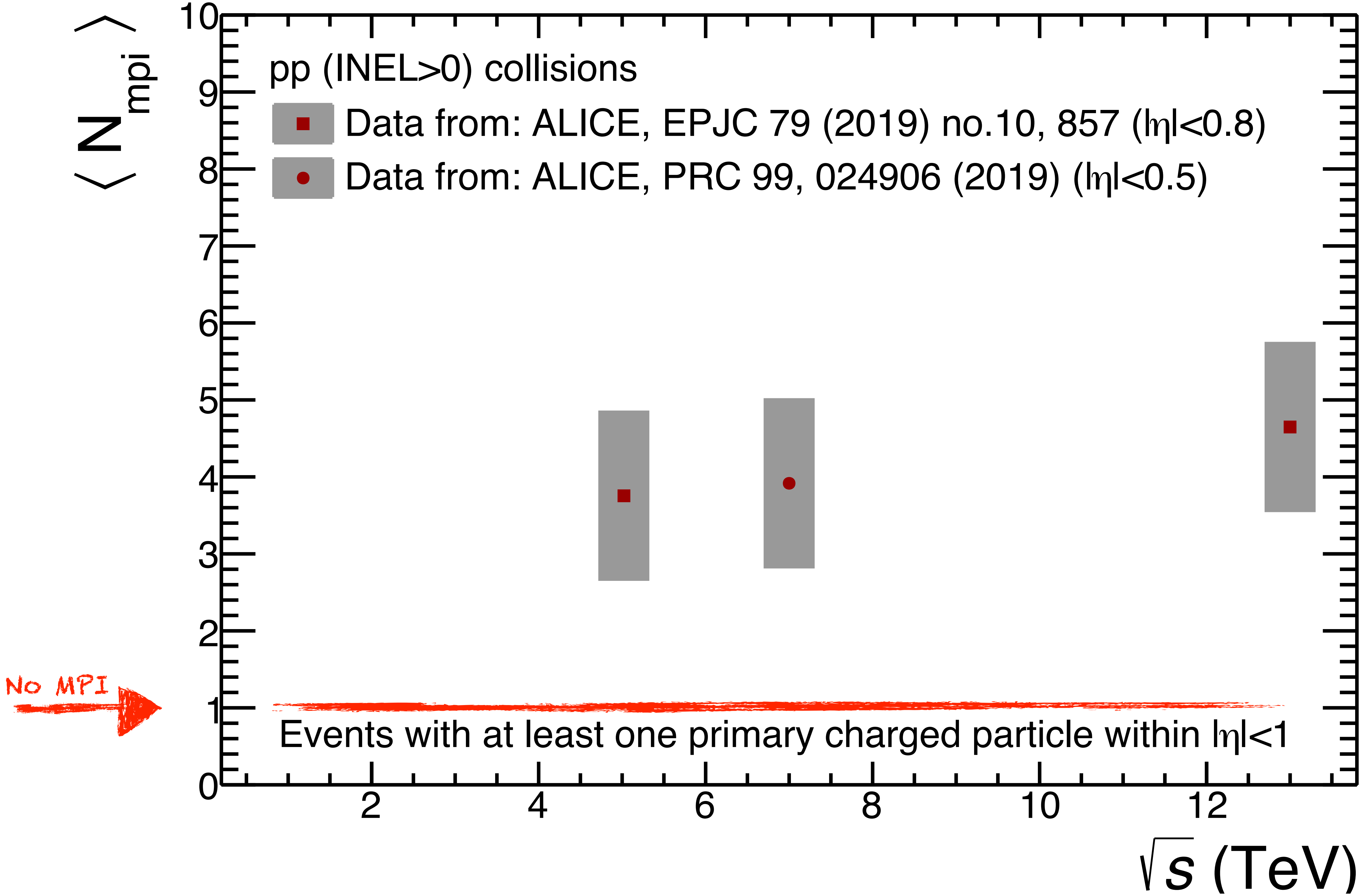
ALICE, PRC 99 (2019) 024906

We used a Toy MC generator aimed at reproducing the correlations between the event-by-event mean p_T and multiplicity. Pseudo-events were required to reproduce the p_T spectra as a function of multiplicity measured by ALICE. The events were further processed using the trained BDT. The performance (MC closure) was crosschecked with MC



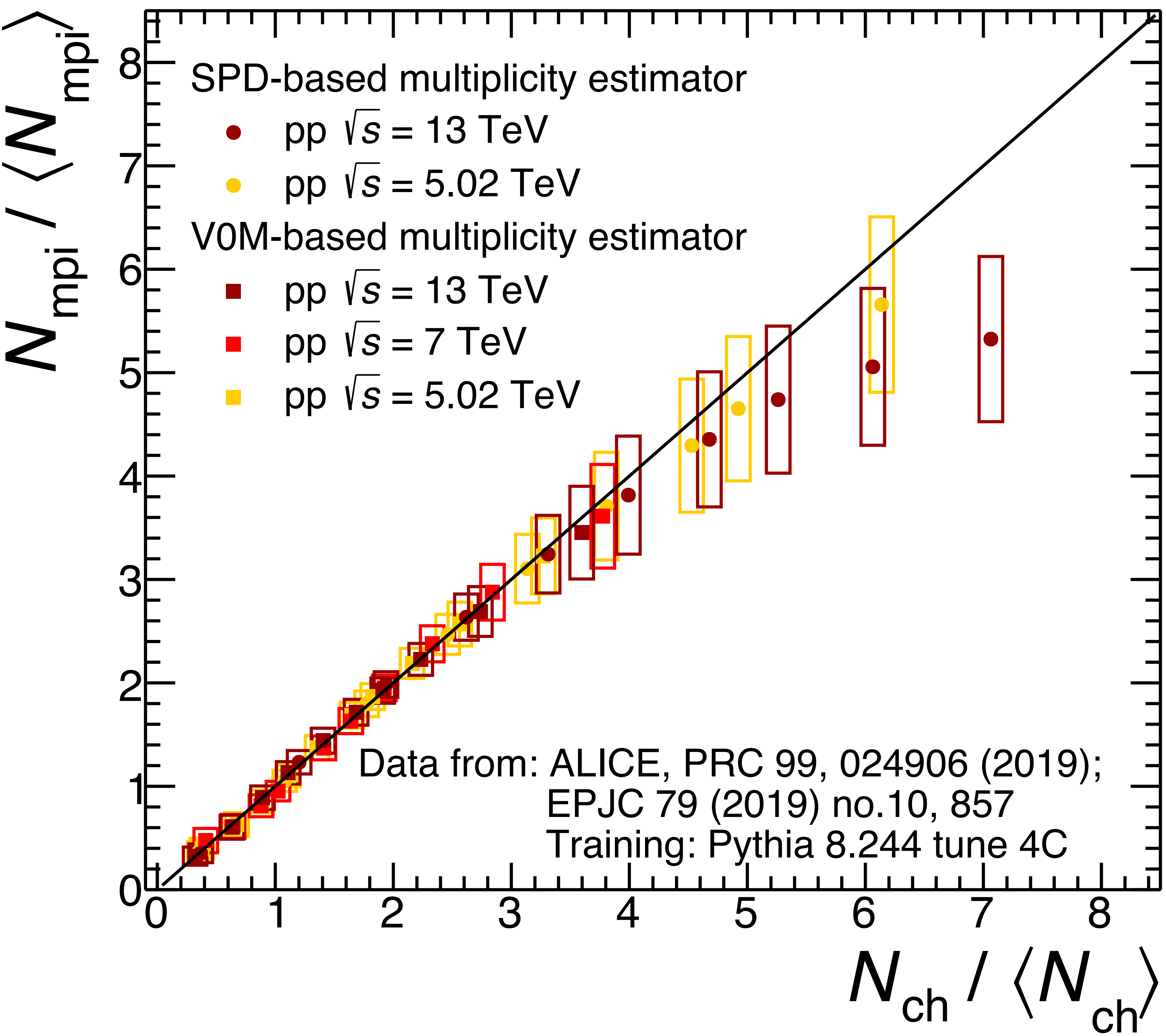
$\langle dN_{ch}/d\eta \rangle$	8.45±0.25	6.72±0.21	5.40±0.17	3.90±0.14	2.26±0.12

MPI from ALICE data



Data support the presence of MPI, regression value above 1!

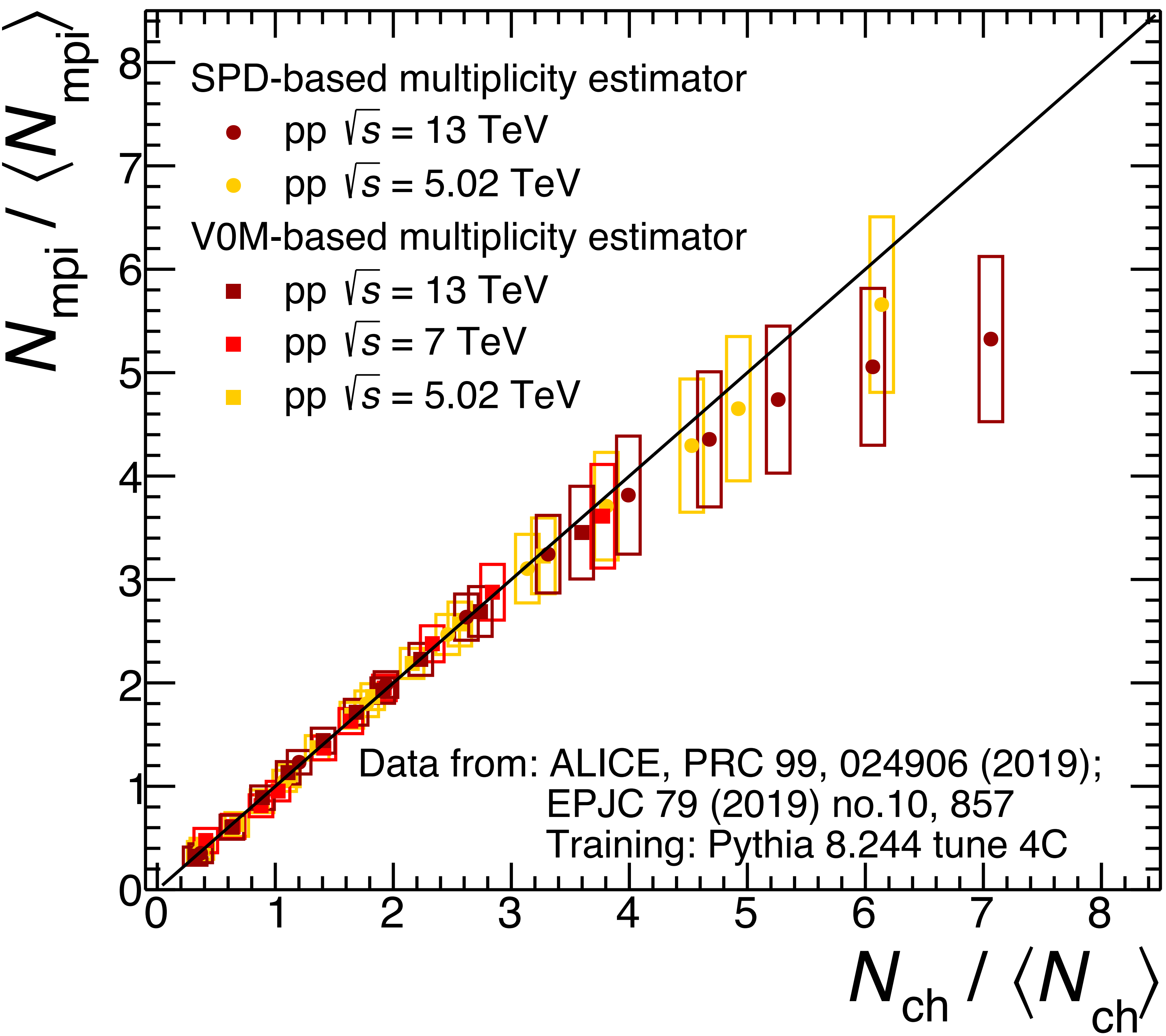
N_{ch} dependence of MPI

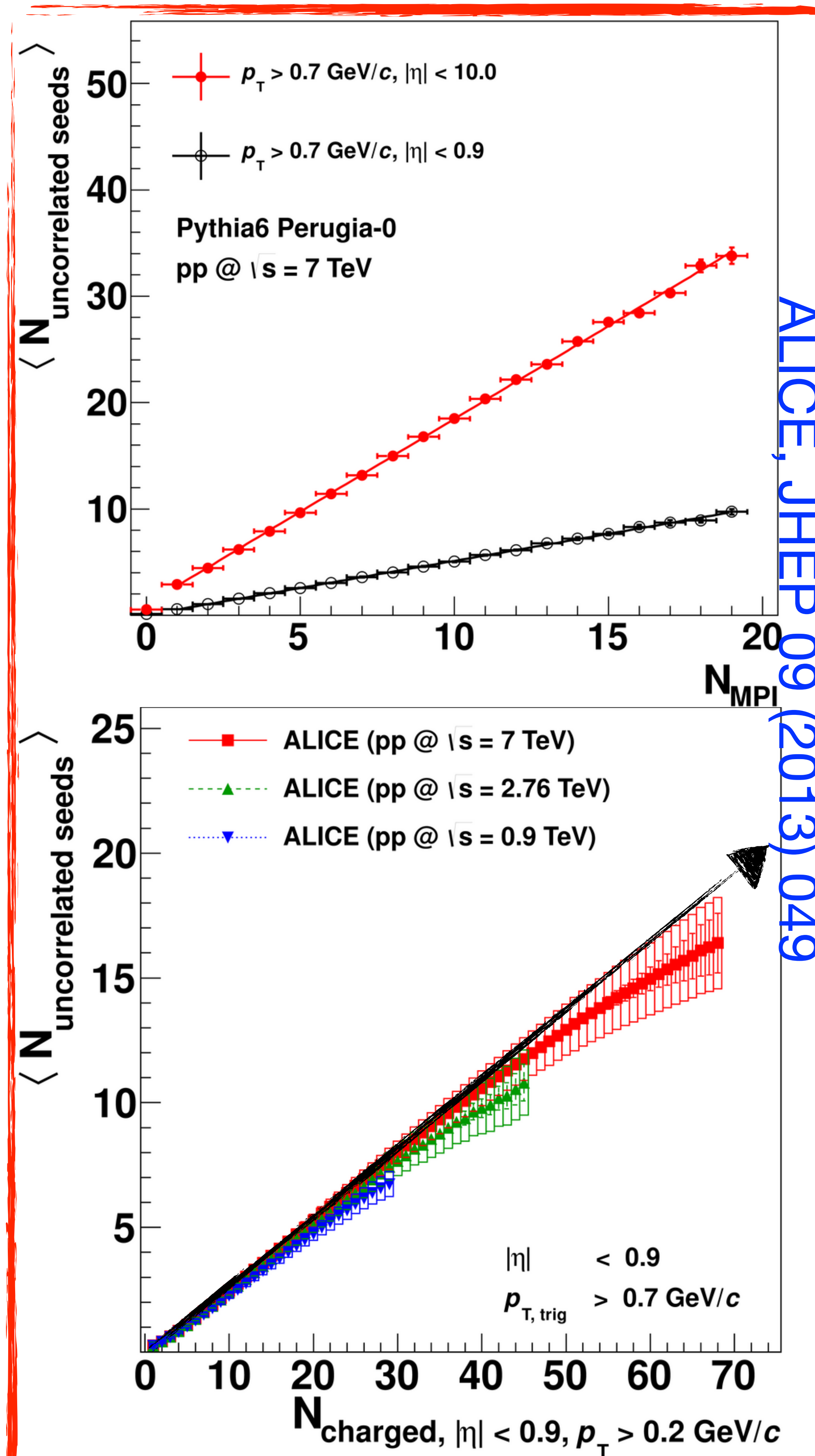


- N_{mpi} as a function of N_{ch} does not show a $\sqrt{s}$ dependence
- $N_{\mathrm{ch}} < 3 \times \langle N_{\mathrm{ch}} \rangle$: N_{ch} increases linearly with N_{mpi}
- $N_{\mathrm{ch}} > 3 \times \langle N_{\mathrm{ch}} \rangle$ can only be reached by selecting events with many high-multiplicity jets

N_{ch} dependence of MPI

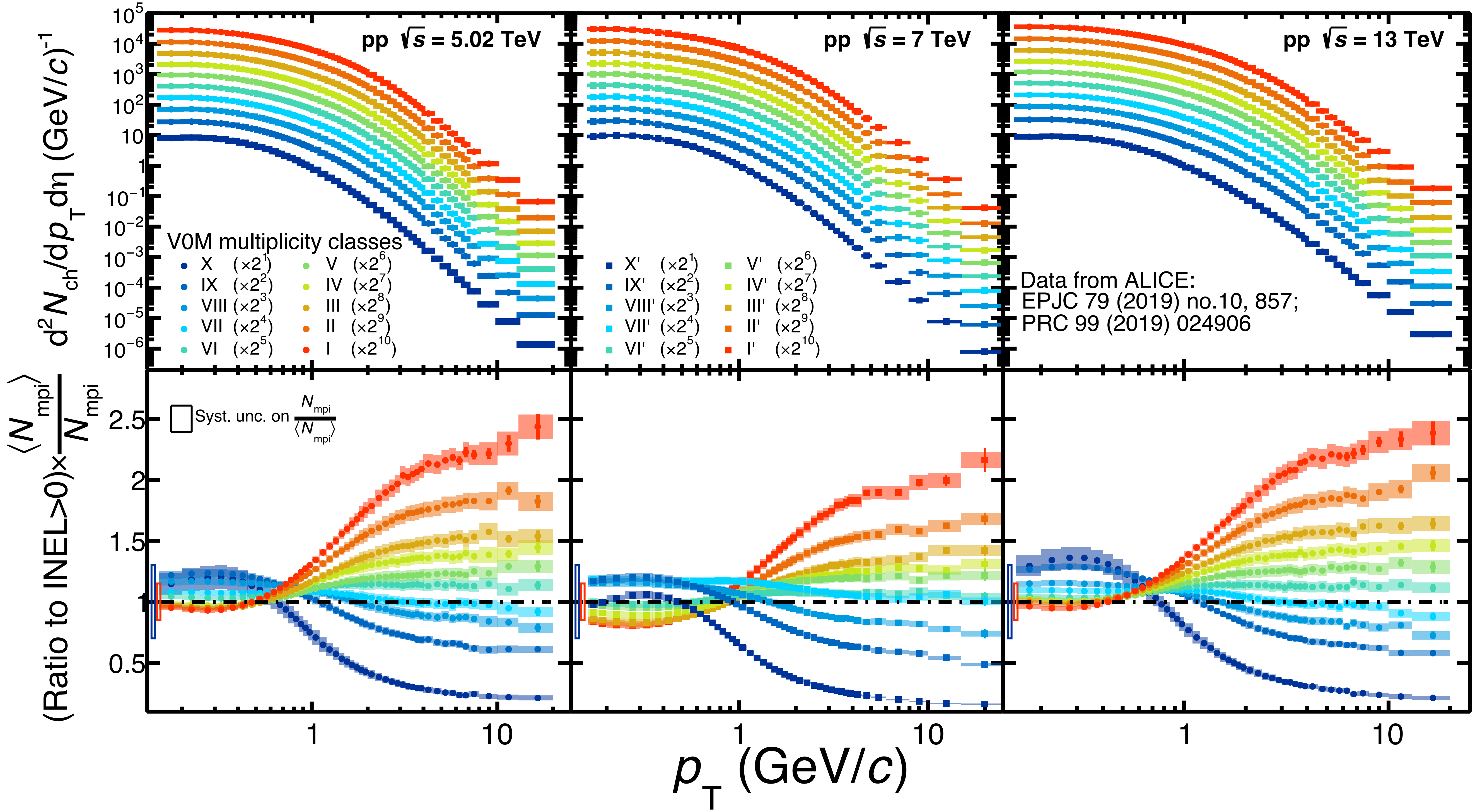
Instituto de Ciencias Nucleares UNAM





Our result is qualitatively consistent with the MPI-dedicated analysis

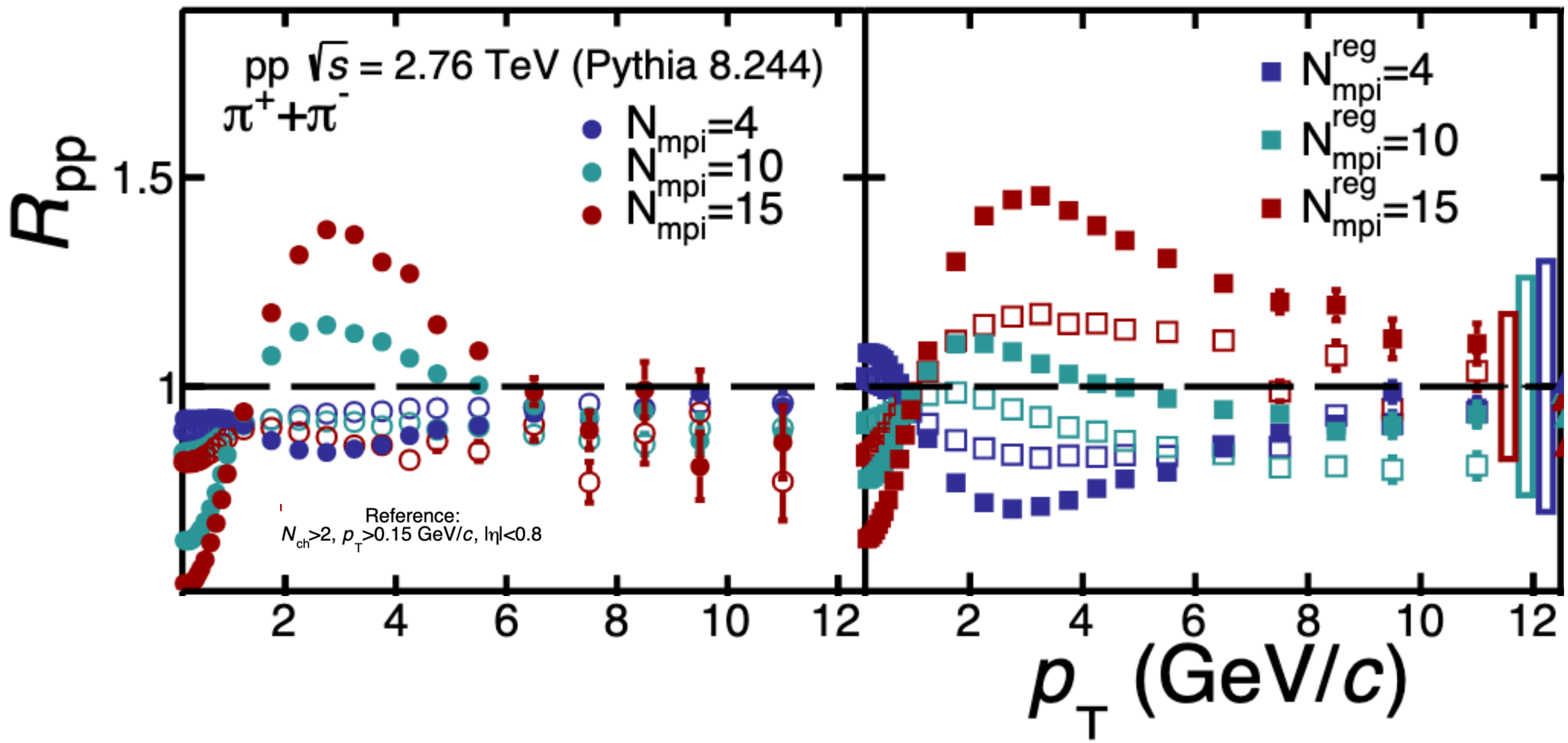
MPI scaling



Within uncertainties, the self-normalised p_T spectra are independent of center-of-mass energy

Other possibilities

Event-by-event determination of N_{mpi}



$$R_{\mathrm{pp}} = \frac{\mathrm{d}^2 N_{\mathrm{ch}}^{\mathrm{mpi}} / (\langle N_{\mathrm{mpi}} \rangle \mathrm{d}\eta \mathrm{d}p_{\mathrm{T}})}{\mathrm{d}^2 N_{\mathrm{ch}}^{\mathrm{MB}} / (\langle N_{\mathrm{mpi}}^{\mathrm{MB}} \rangle \mathrm{d}\eta \mathrm{d}p_{\mathrm{T}})}$$

Summary

Instituto de Ciencias Nucleares UNAM

- MPI are needed to describe the pp data, this mechanism can help to elucidate the origin of heavy-ion-like behaviour discovered in pp collisions

- In this work, we proposed to use ML-based regression in order to extract MPI

- The proposed strategy was validated using simulations of pp collisions at different center-of-mass energies. The model dependence was estimated processing Pythia 8 simulations with the BDT trained with Herwig 7 and *vice versa*

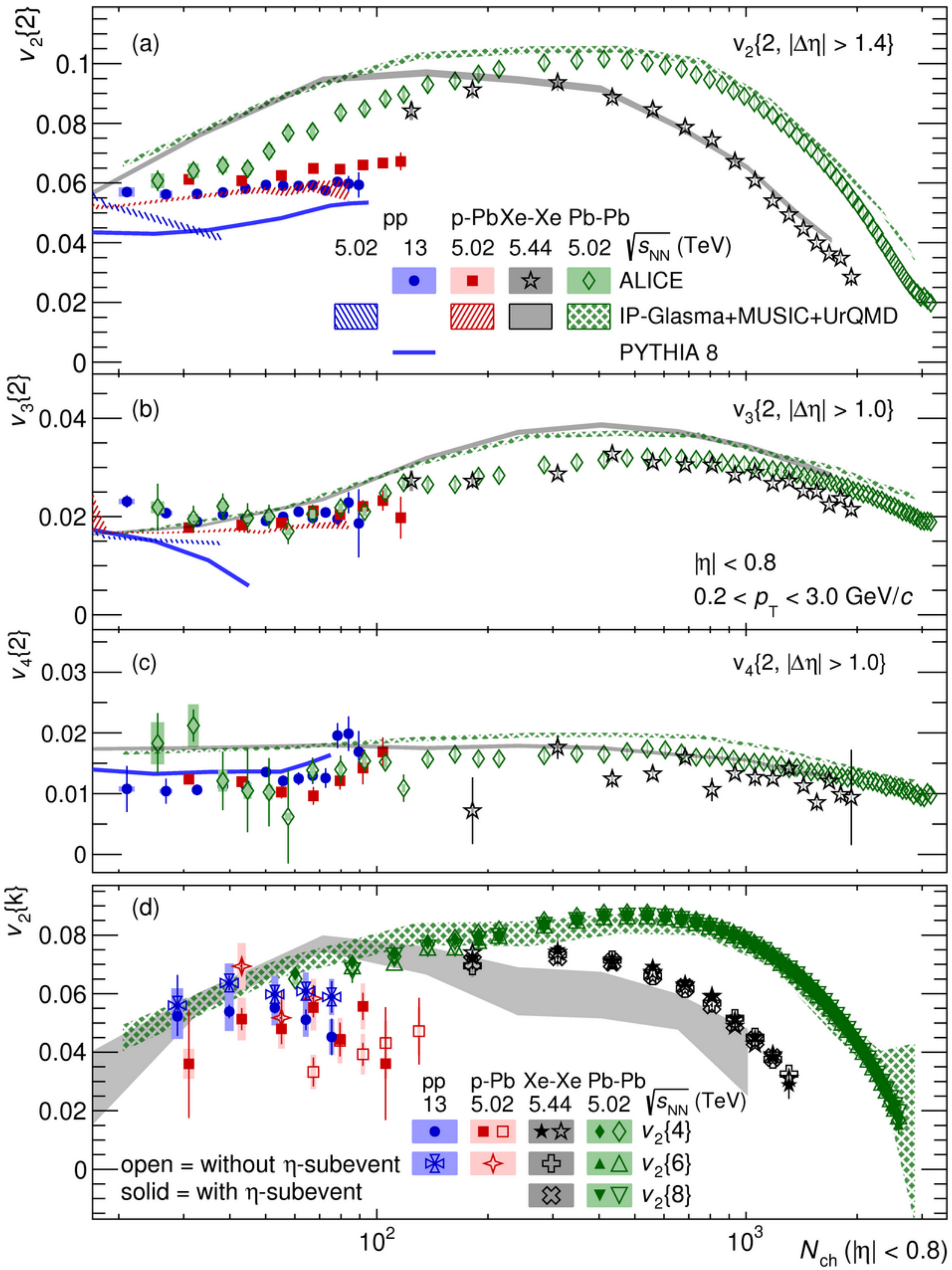
- We found that pp data at LHC is consistent with the presence of MPI, the N_{ch} dependence of N_{mpi} was also determined for pp collisions at 5.02, 7 and 13 TeV

- The proposed strategy can be used to determine N_{mpi} event-by-event

Backup

Collectivity in small systems

Instituto de Ciencias Nucleares UNAM



- Striking similarities between numerous observables have been observed across different collision systems at both RHIC and LHC energies, when compared at similar multiplicity
- Besides hydrodynamic description, calculations from transport models, hadronic re-scattering, Multi-Parton Interactions (MPI), string rope and shoving, as well as initial state effects have been investigated

e.g. ALICE, PRL 123, 142301 (2019)