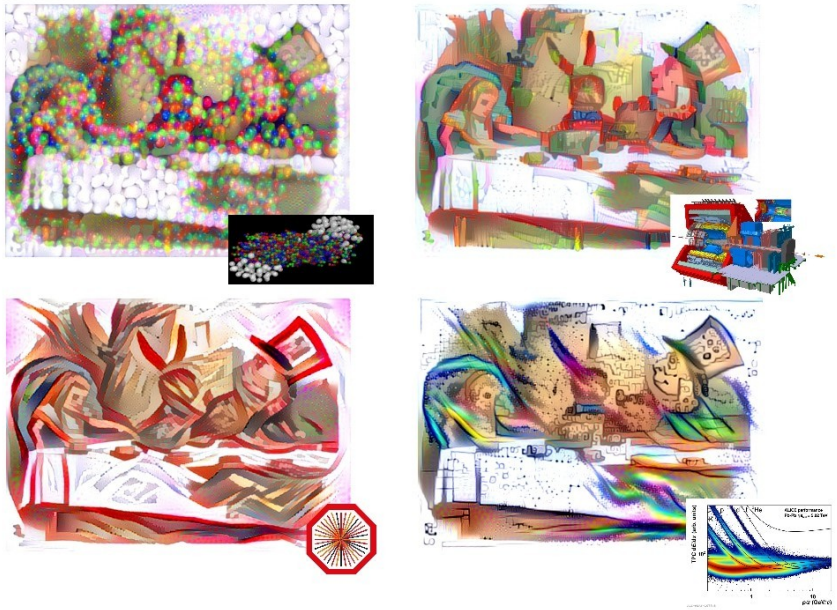
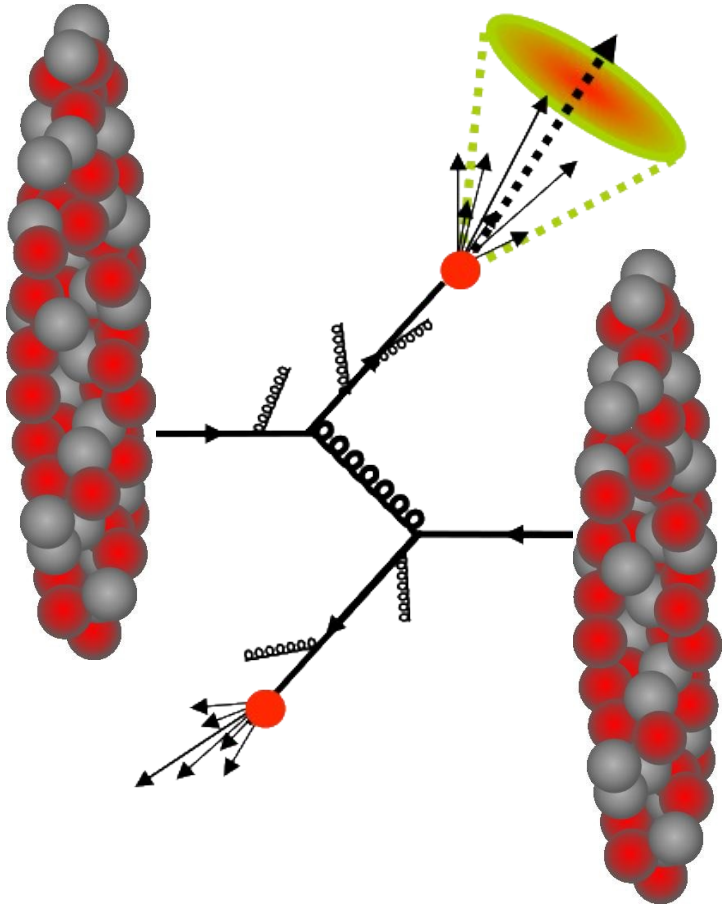
Machine Learning based jet momentum reconstruction in Pb-Pb collisions in ALICE

Rüdiger Haake (Yale University)
for the ALICE Collaboration



EPS-HEP 2019, Ghent, Belgium
(11.07.2019)

Motivation: Low-p_T jets in heavy-ion collisions



Broad bandwidth of jet measurements:

- Spectra, nuclear modification factors
- Correlation measurements
- Shapes, (sub)structure analyses

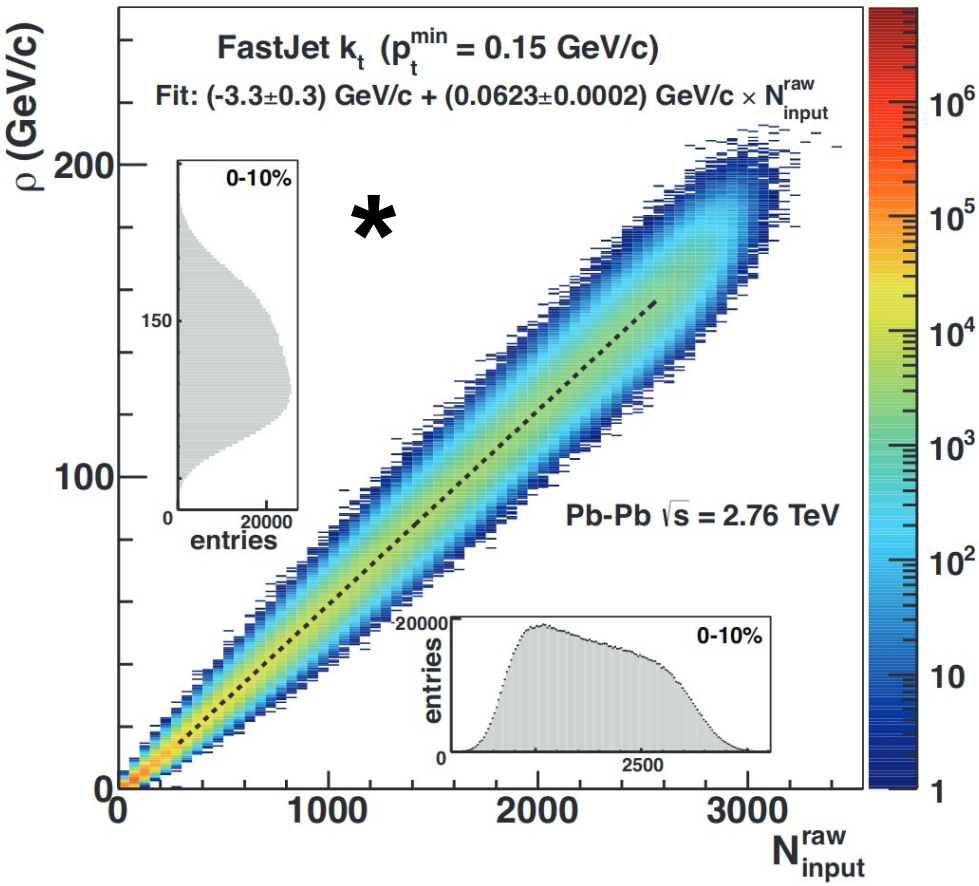
Particularly interesting:

Low transverse momenta

→ Medium effects strongest

Main obstacle: Overwhelmingly large background from underlying event of particles uncorrelated with the hard scattering

Jets in heavy-ion background I



Huge (charged) particle and energy density from soft processes:

→ Roughly 140 GeV/*c* per unit area

→ Tracks per event $N = O(2000)$

($p_{T,track}$ > 0.15 GeV/*c*, $\sqrt{s_{NN}}$ = 2.76 TeV)

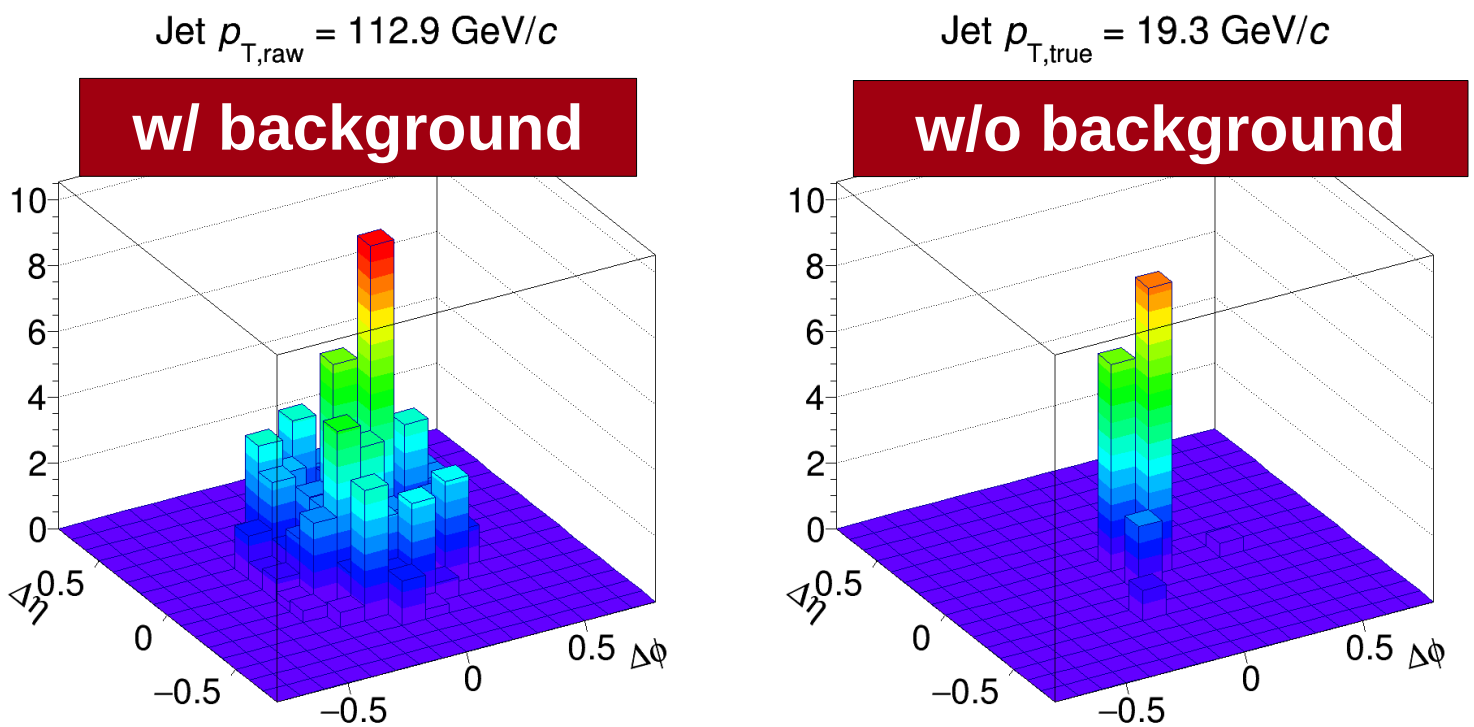
In addition: Large region-to-region fluctuations!

- Random poissonian
- Particle flow
- Detector inhomogenities
- ...

* **ALICE Collaboration JHEP 1203 (2012) 053**

Jets in heavy-ion background II

Background effect on jet in Pb-Pb (toy background):

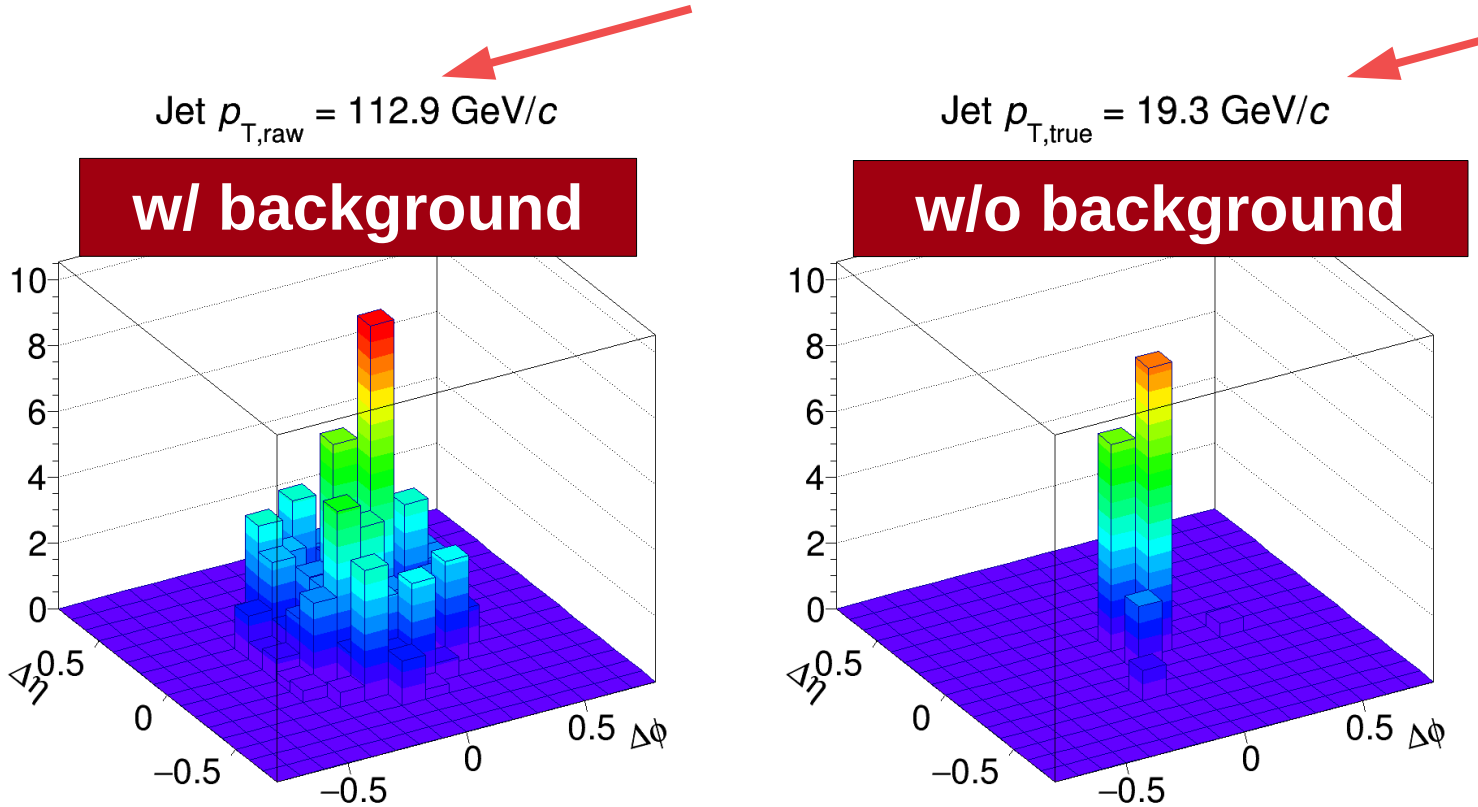


Accumulated particle momentum around jet axis

Background/fluctuations largely affect jet momenta and axes

Jets in heavy-ion background III

Background effect on jet in Pb-Pb (toy background):



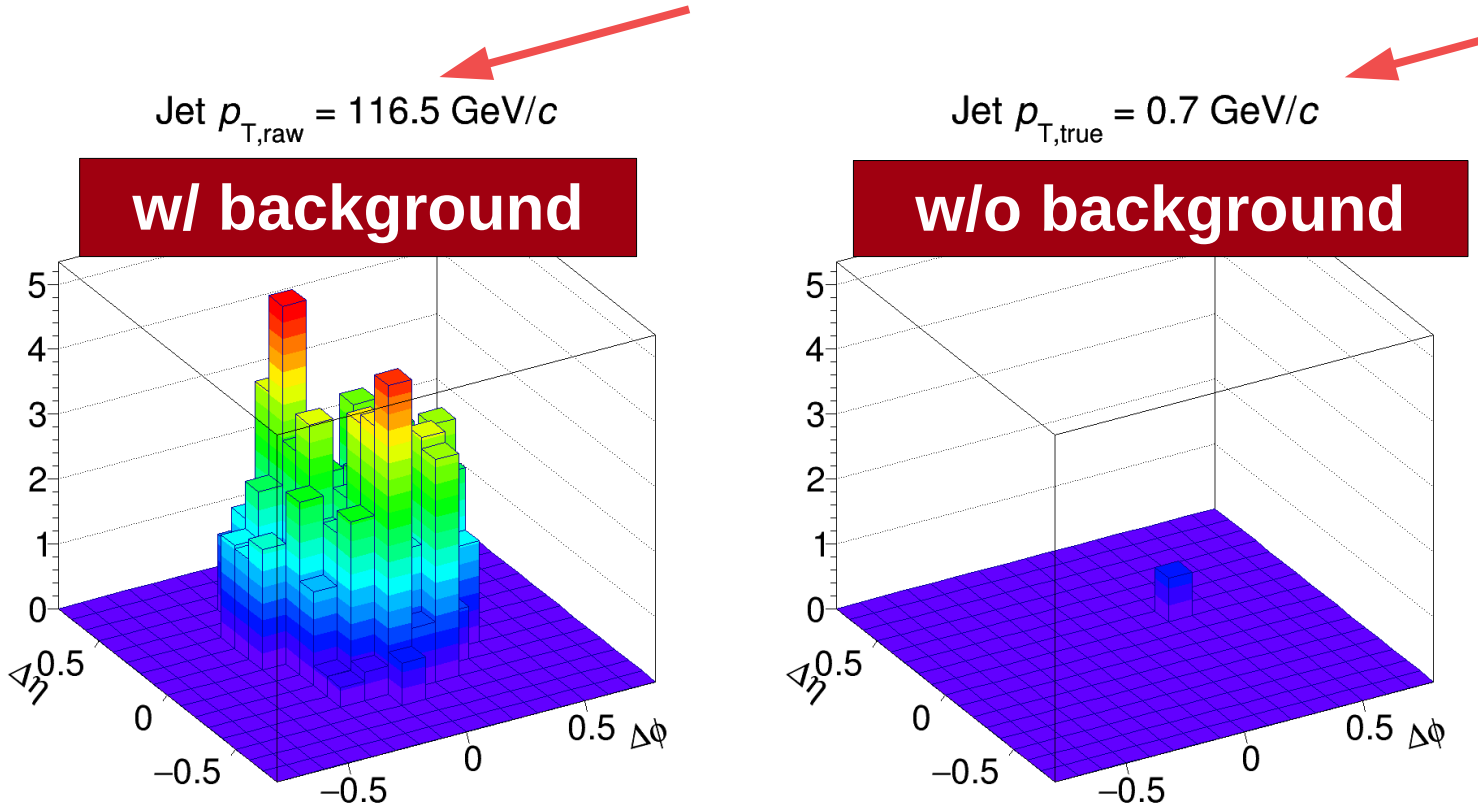
Accumulated particle momentum around jet axis

Example of fluctuated jet

Background/fluctuations largely affect jet momenta and axes

Jets in heavy-ion background IV

Background effect on jet in Pb-Pb (toy background):



Accumulated particle momentum around jet axis

Example of combinatorial jet

Background/fluctuations largely affect jet momenta and axes

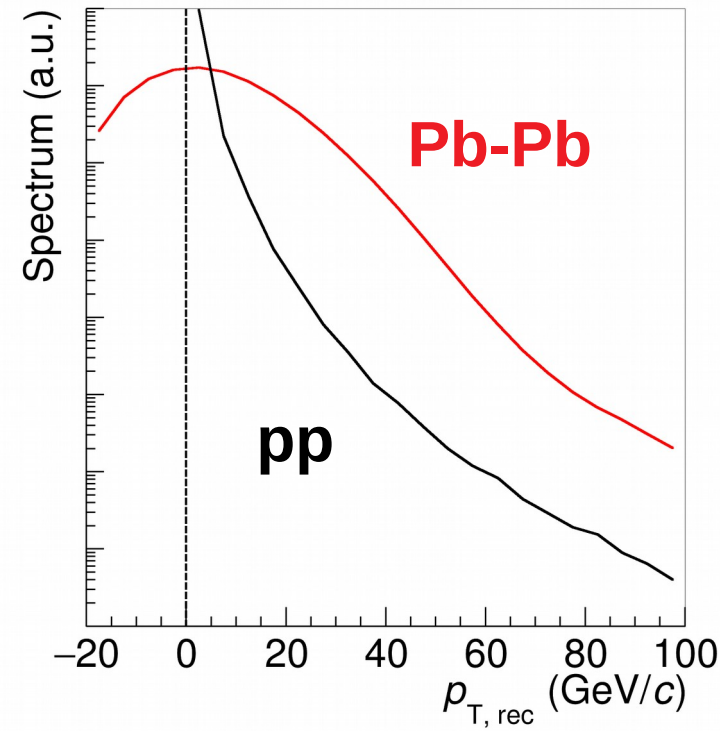
How to correct jets for background?

De facto standard method in ALICE:
Area-based correction method

- Event-by-event: Calculate mean background density ρ
- Jet-by-jet: Subtract background from jet depending on jet area A

$$p_{\mathrm{T,rec}} = p_{\mathrm{T,raw}} - \rho A$$

- Residual fluctuations usually treated statistically in unfolding



Main properties:

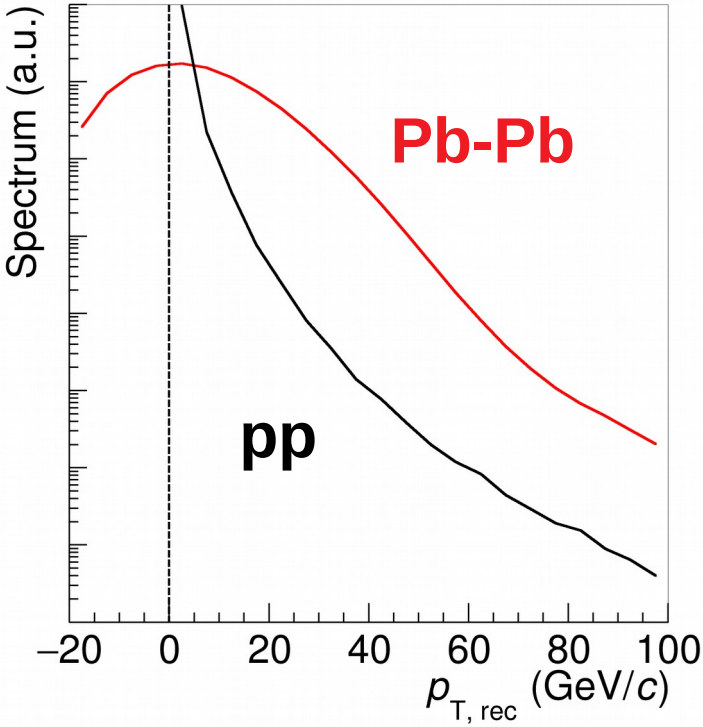
- Poor precision on jet energy scale at low-p_{T}
- Method is sensitive to energy-flow
  - → Independent on fragmentation changes within jet *(good)*
  - → Combinatorial jets not treated *(not so good)*

Can we do better?

ML background estimator: Introduction

Idea: Background different from signal
→ Different particle composition
→ Different spatial distribution

Ansatz: Estimate background jet-by-jet
→ Exploit more information from jet
→ Constituent properties, shapes, ...



Perfect candidate for high-dimensional parameter correlations:

Machine Learning (ML)

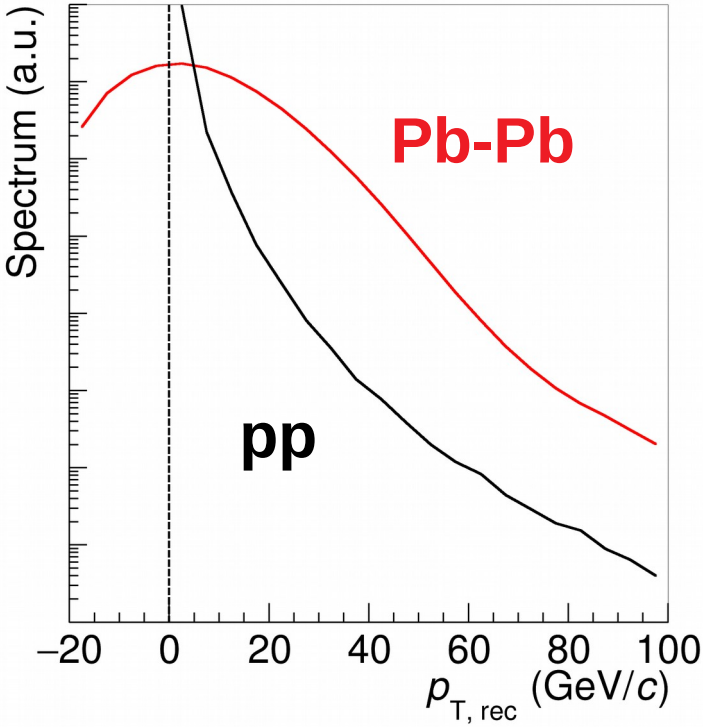
Instead of explicit modeling, we learn signal/background relation from jets embedded in background

ML background estimator: Introduction

Method comes with a very general trade-off:

→ Get better performance with more information

→ But also: Potential dependence on jet fragmentation!



- Trained on PYTHIA jets in background
  → Inherently assumes a certain fragmentation of jets
  → **Thus, we measure pp-like jets in Pb-Pb**

- Uncertainty covers realistic changes in fragmentation
  → We unfolded assuming pure quark-jet fragmentation

ML background estimator: Train the model

Supervised learning approach:

- Mapping raw ↔ corrected jet momentum learned from data
- Regression task: Numerical value is approximated

ML technique:

- We use a shallow neural network (100 → 100 → 50)
- Several algorithms (Random forest, linear regression, …) tested and found to be in agreement

Training data:

- PYTHIA8, $\sqrt{s}$ = 5.02 TeV, detector-level rec., embedded in thermal toy model inspired by real Pb-Pb distributions
- Input parameters (found by parameter ranking):
  Area-based jet momentum, N_{const}, angularity, eight leading jet constituent momenta

Results

Data & analysis setup

Dataset:

- Pb-Pb, $\sqrt{s_{NN}}$ = 5.02 TeV, taken in 2015
- Minimum bias, 0-10% + 30-50% most central collisions
- 0-10%: 7M events, 30-50%: 14M events
- pp reference taken from arXiv:1905.02536 [nucl-ex]

Jet definition:

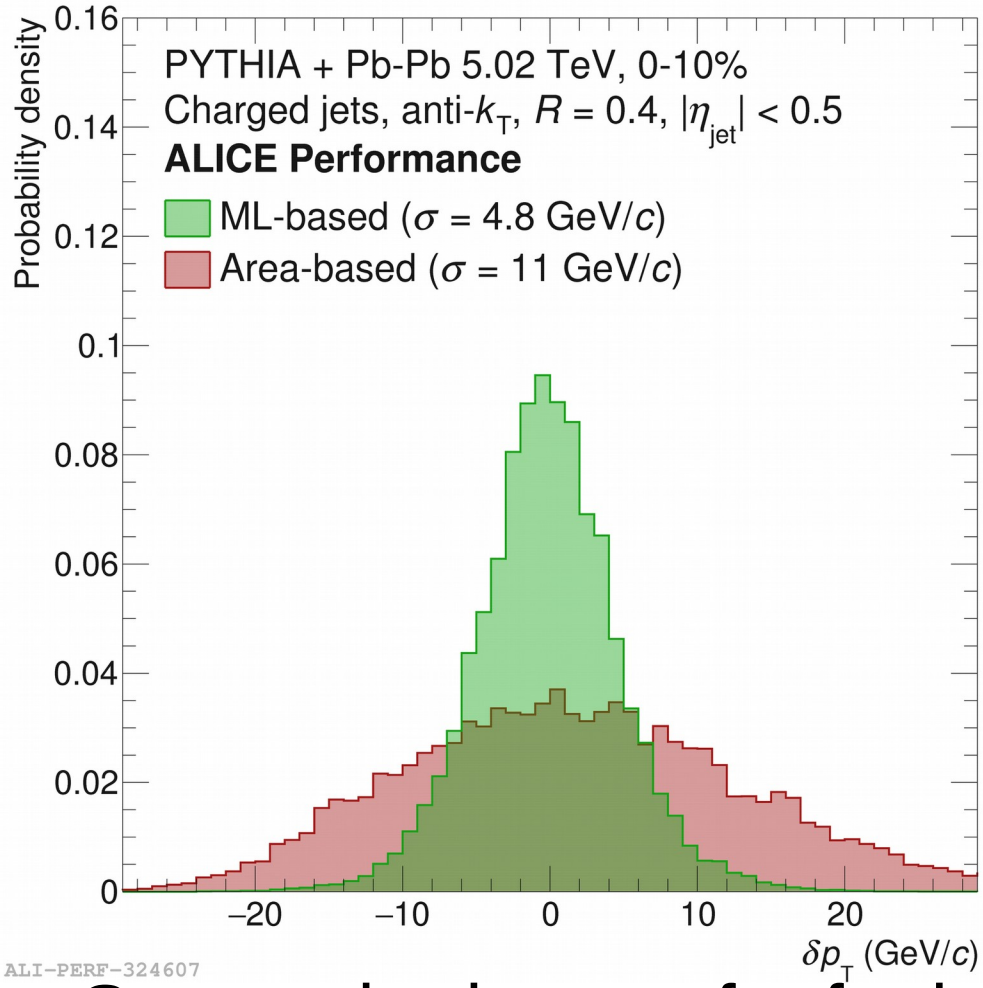
- Track-based jets (no neutral particles)
- Fastjet, anti-k_T jets, R = 0.2, 0.4, 0.6
- Jet cones fully-contained in acceptance, full azimuth

Unfolding for detector effects & residual fluctuations:

- SVD unfolding
- Prior: PYTHIA
- Detector response matrix from matched PYTHIA jets using a full GEANT3 detector model

Estimator performance

Direct measure for residual fluctuations: δp_T



- Calculated by embedding PYTHIA jets into Pb-Pb events
- δp_T = residual background fluctuation of probe

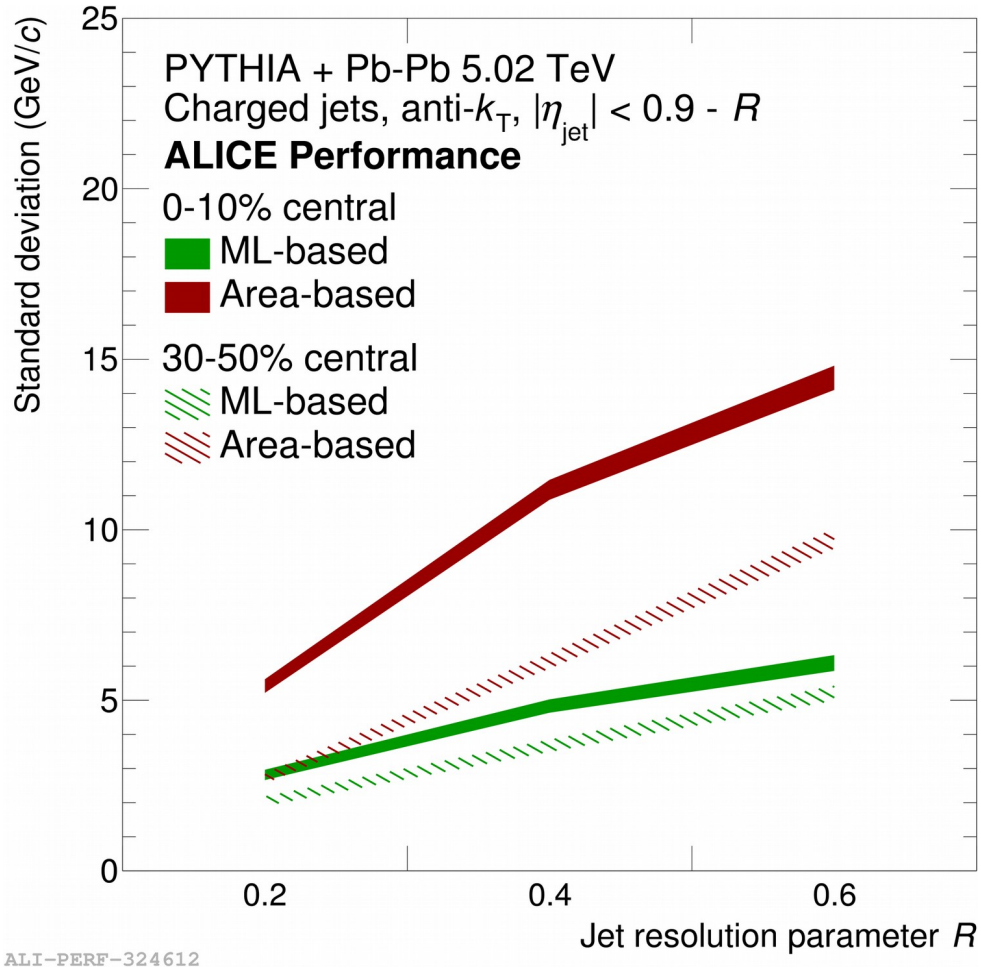
→ **Resolution much better for ML-based approach**

→ **Still centered around zero**

- See method paper for further toy studies on robustness and performance: R. Haake et al., PRC 99, 064904 (2019)

Estimator performance

Direct measure for residual fluctuations: δp_T



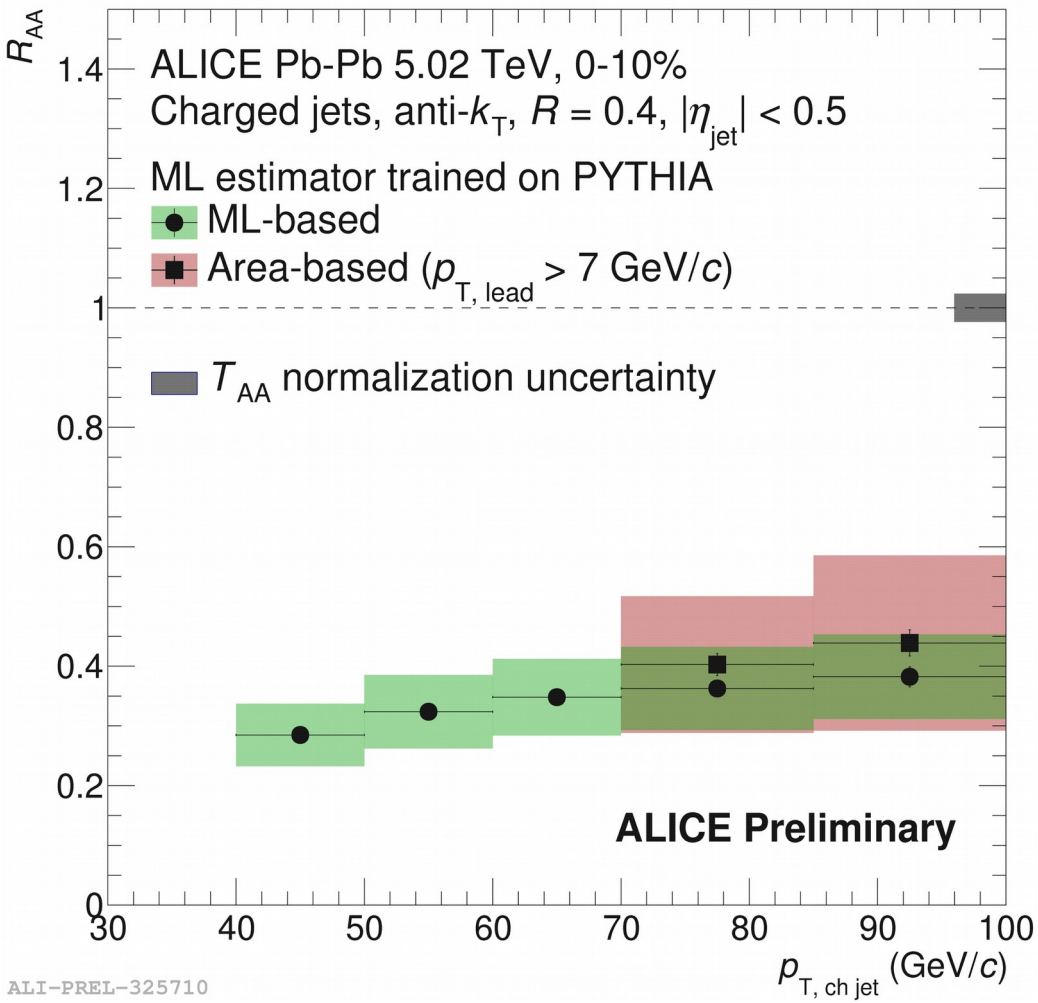
- Standard deviation = width of δp_T
- Smaller = better resolution
- Shown for several jet R

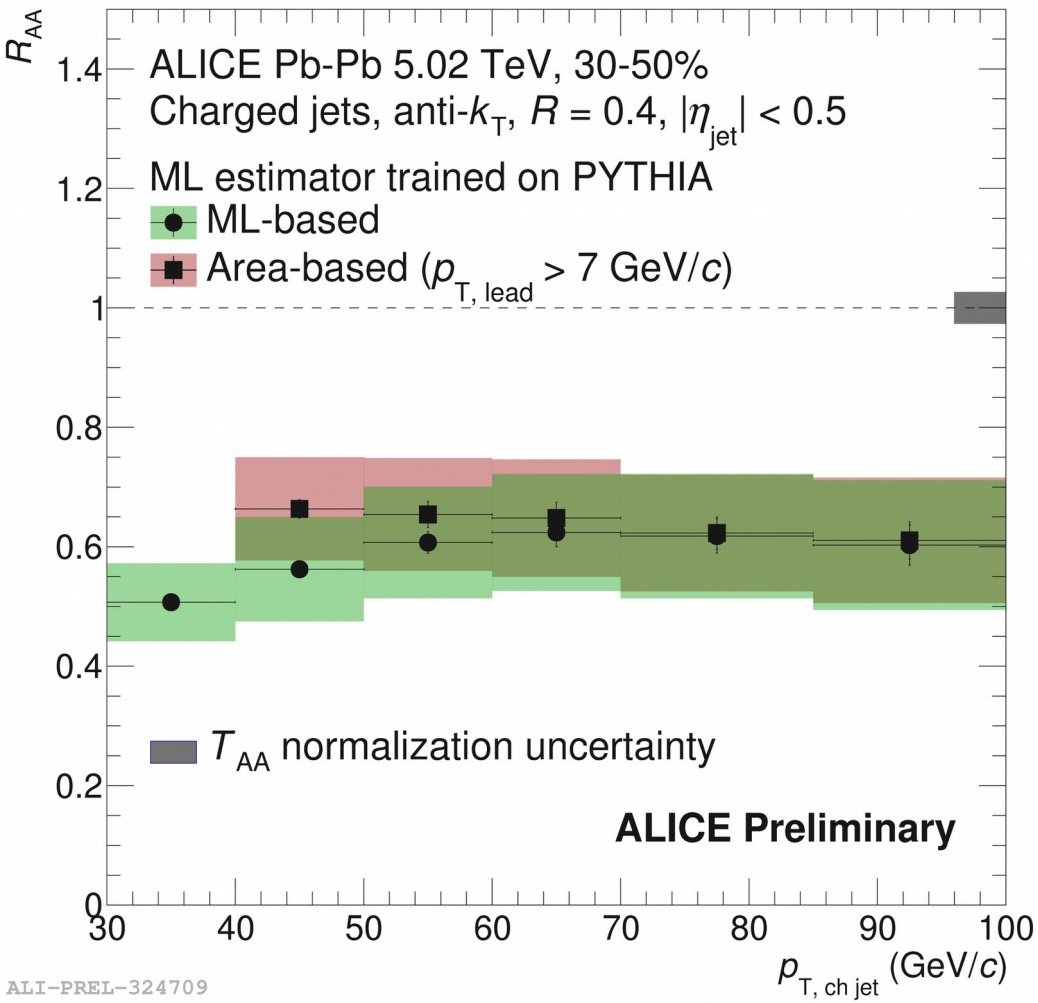
→ **Good ML estimator performance up to largest *R***

→ **Trend shows less progression towards higher *R***

- See method paper for further toy studies on robustness and performance: R. Haake et al., PRC 99, 064904 (2019)

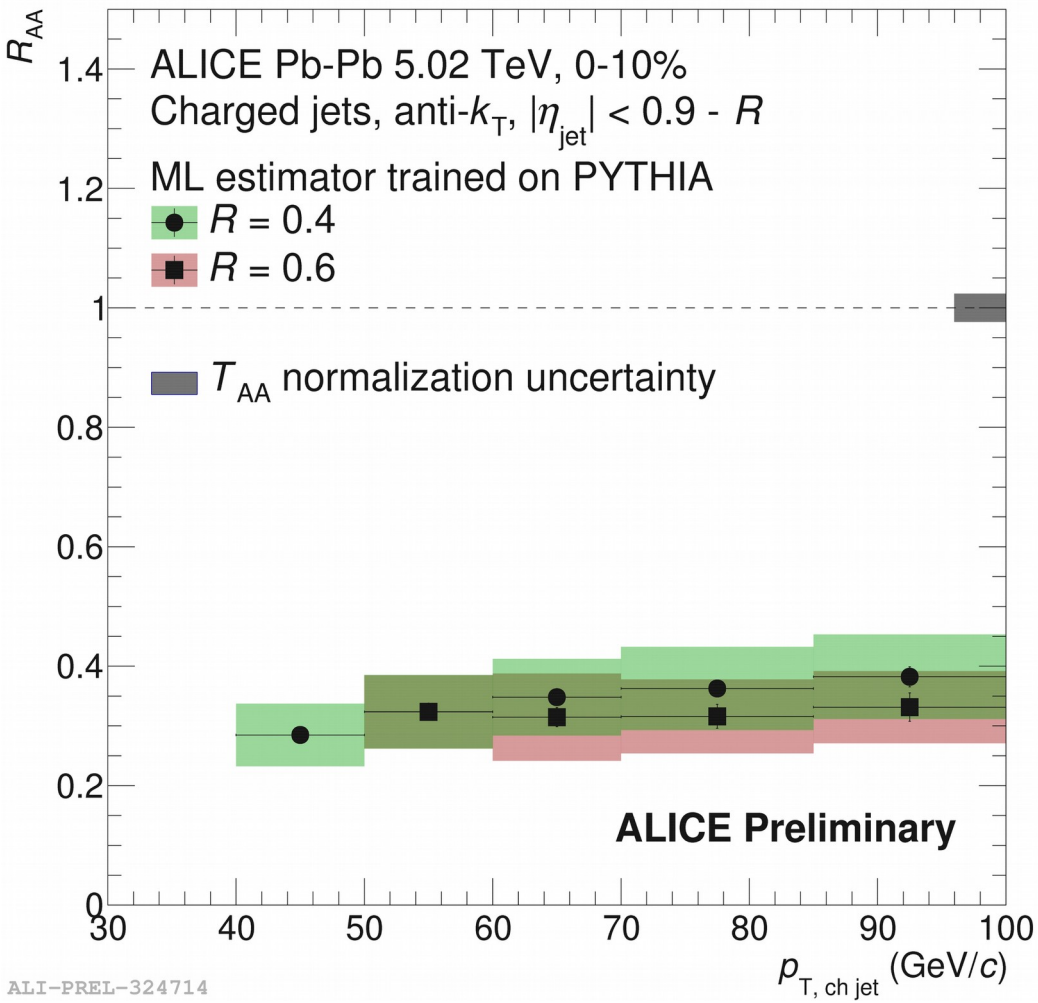
Nuclear modification factor R_{AA}

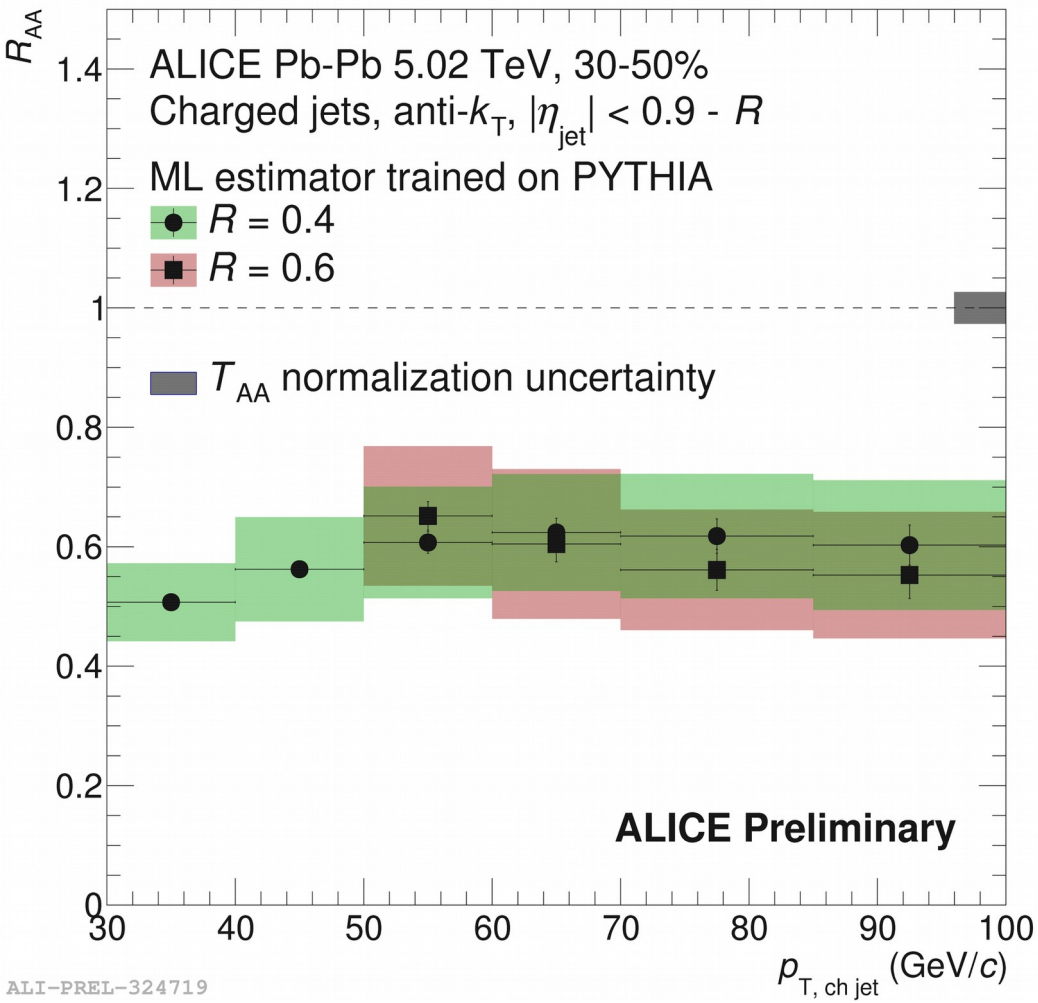




- Results fully compatible with established correction method
- **For 0-10%, we gain up to 30 GeV/*c* towards low p_T!**

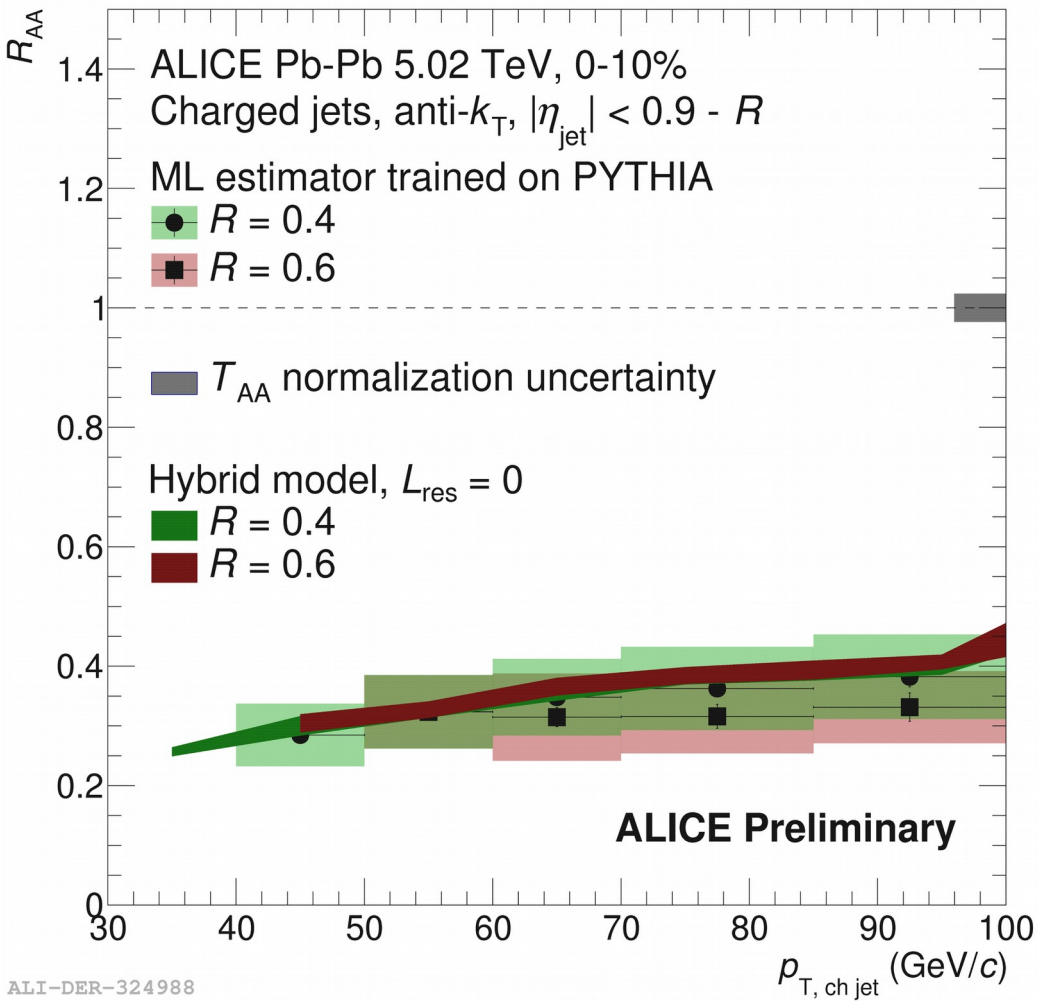
Nuclear modification factor R_{AA}

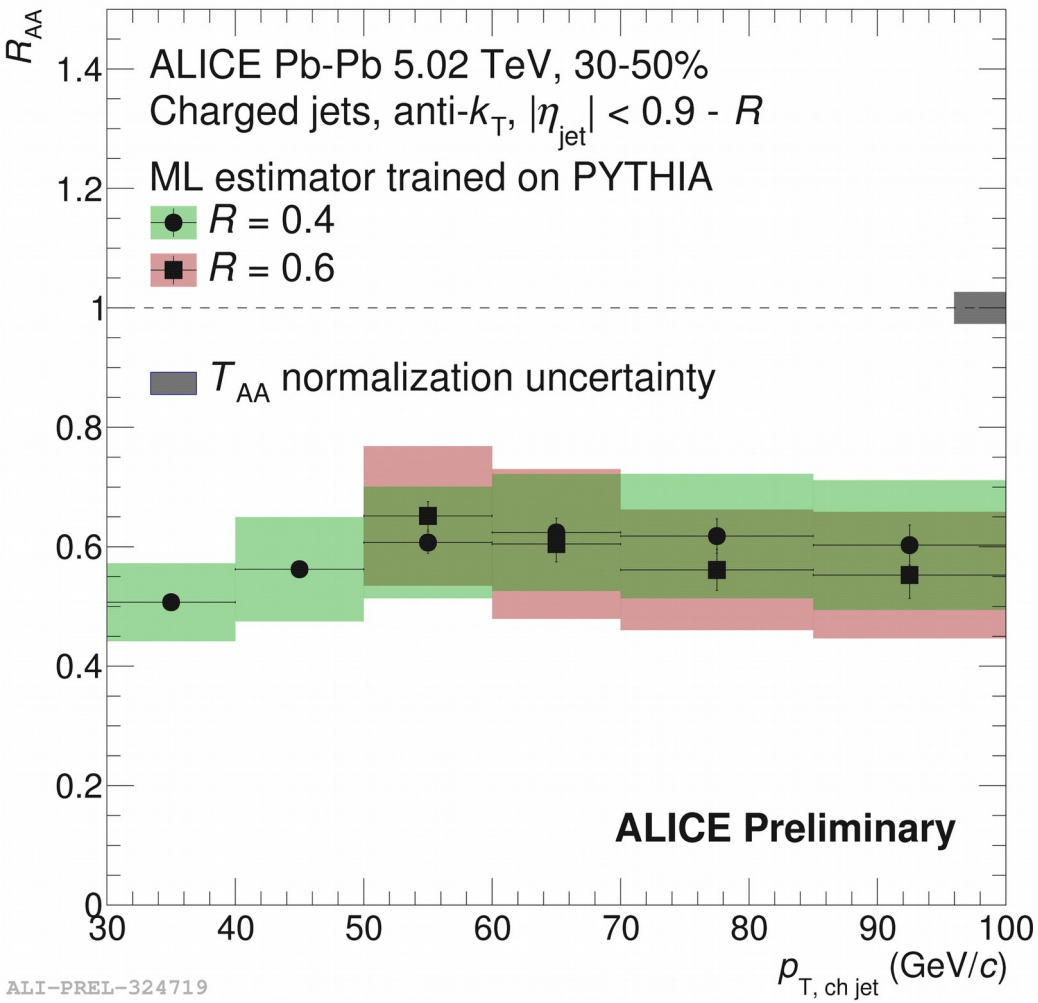




- **New at the LHC: R = 0.6 in Pb-Pb, down to 50 GeV/c!**
- Nuclear modification compatible for R = 0.4 & R = 0.6

Nuclear modification factor R_{AA}

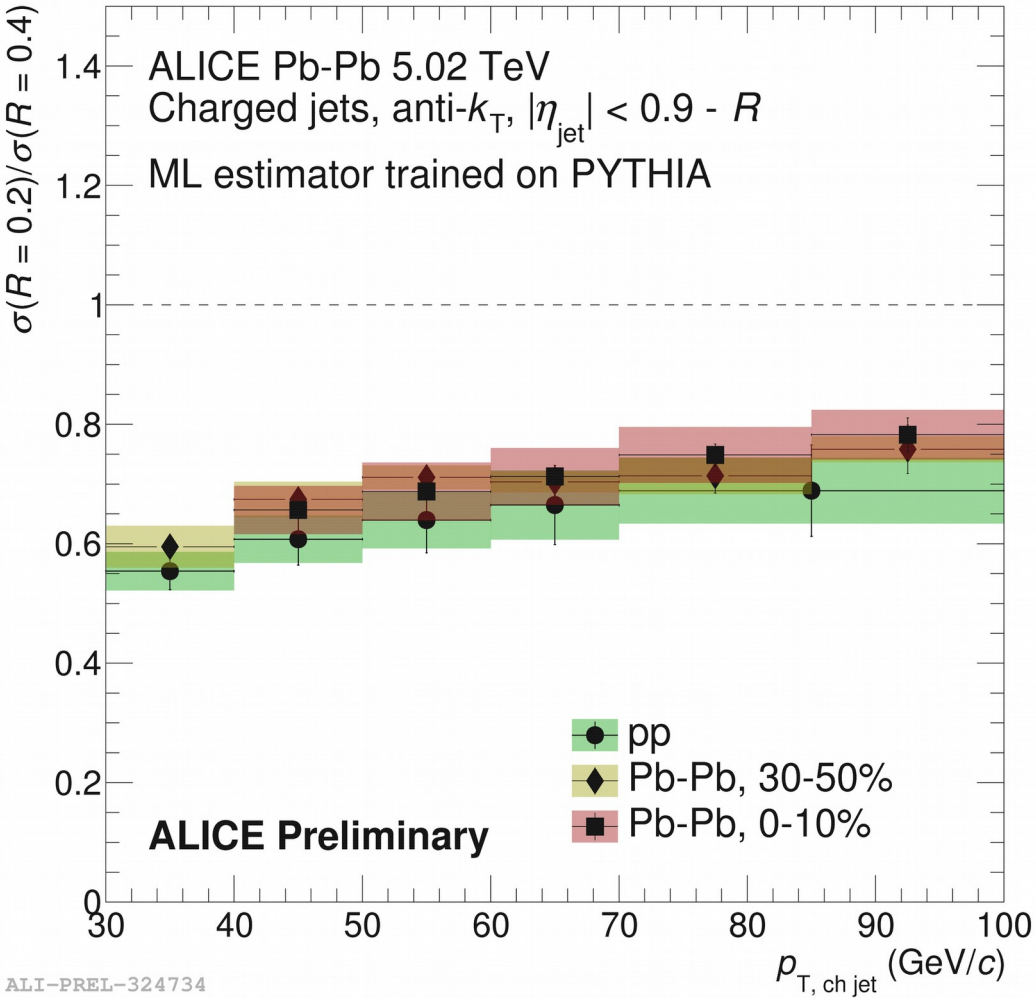


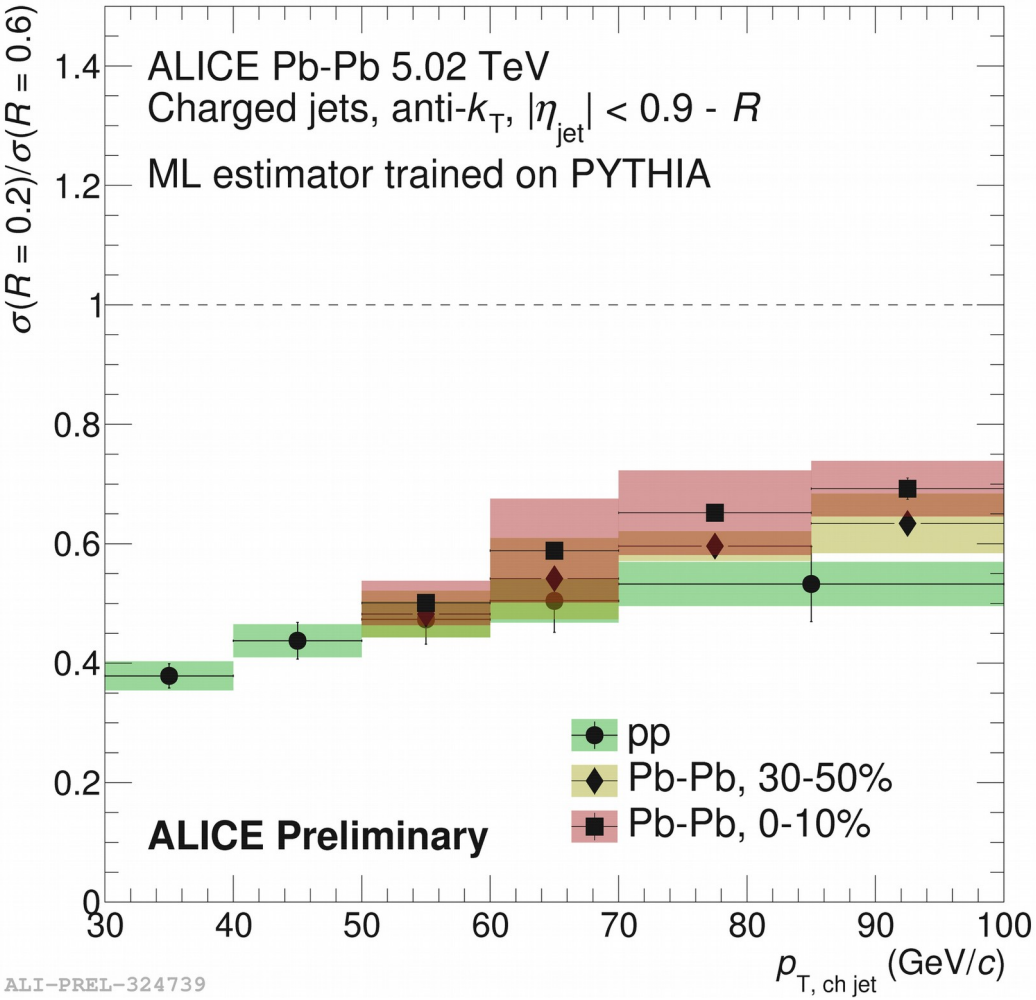


- **New at the LHC: R = 0.6 in Pb-Pb, down to 50 GeV/c!**
- R_{AA} for R = 0.6 still largely suppressed (assuming pp-like jets)

J. Casalderrey-Solana et al., JHEP (2014) 2014: 19

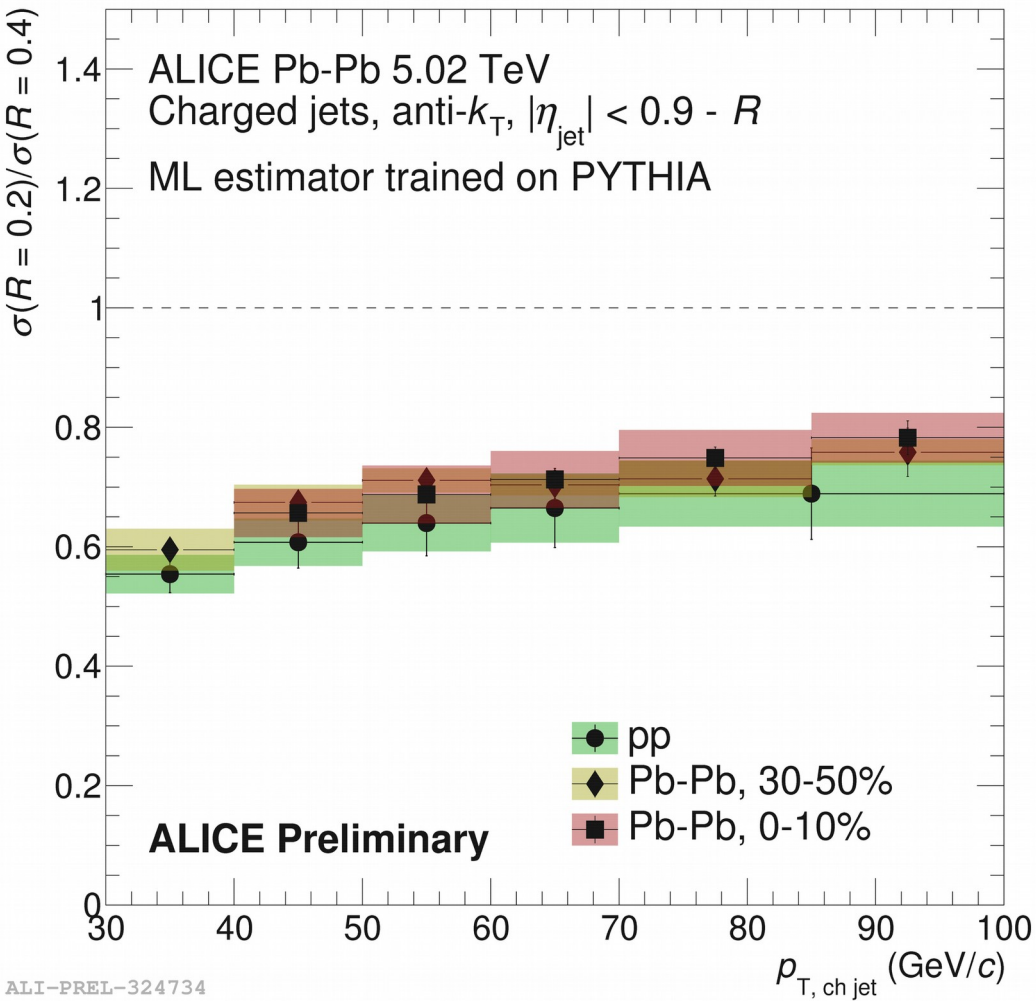
Jet cross-section ratios

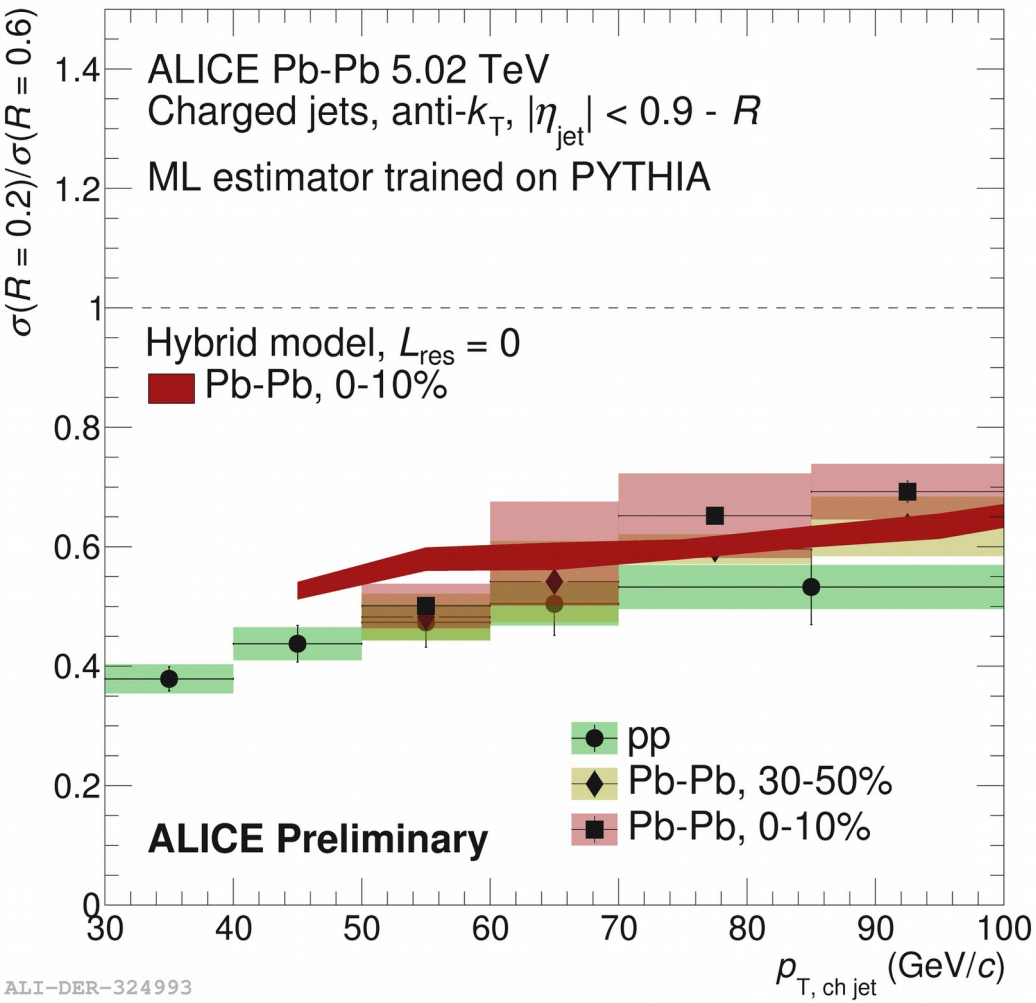




- **Cross-section ratio not strongly changed compared to pp**
- Hint to slightly higher ratio, but covered by uncertainties

Jet cross-section ratios





- **Cross-section ratio not strongly changed compared to pp**
- Hint to slightly higher ratio, but covered by uncertainties

J. Casalderrey-Solana et al., JHEP (2014) 2014: 19

Summary

Introduced method to correct jets for heavy-ion background:

- ML-based, exploits jet fragmentation properties
- Method shown to be compatible with established area-based correction technique where comparable

Measured charged jets in region of phase space, unconstrained so far by measurements:

- For R = 0.4: 40 GeV/c for 0-10%, 30 GeV/c for 30-50% most central collisions
- New in Pb-Pb at the LHC: jets with R = 0.6
- Nuclear modification factor R_{AA}: No strong radius dependence observed, continues to lower p_T
- Cross-section ratios not modified within uncertainties

Outlook

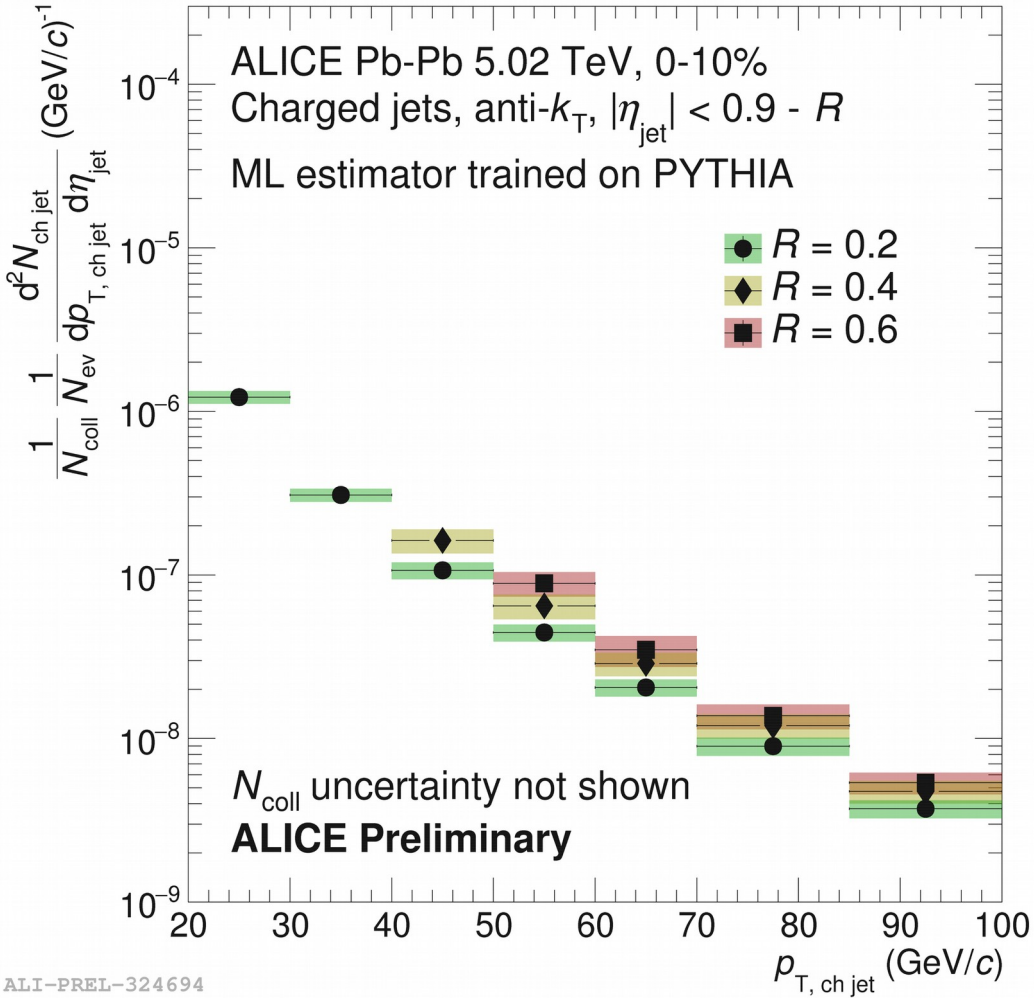
We still have a lot of ideas to improve the method:

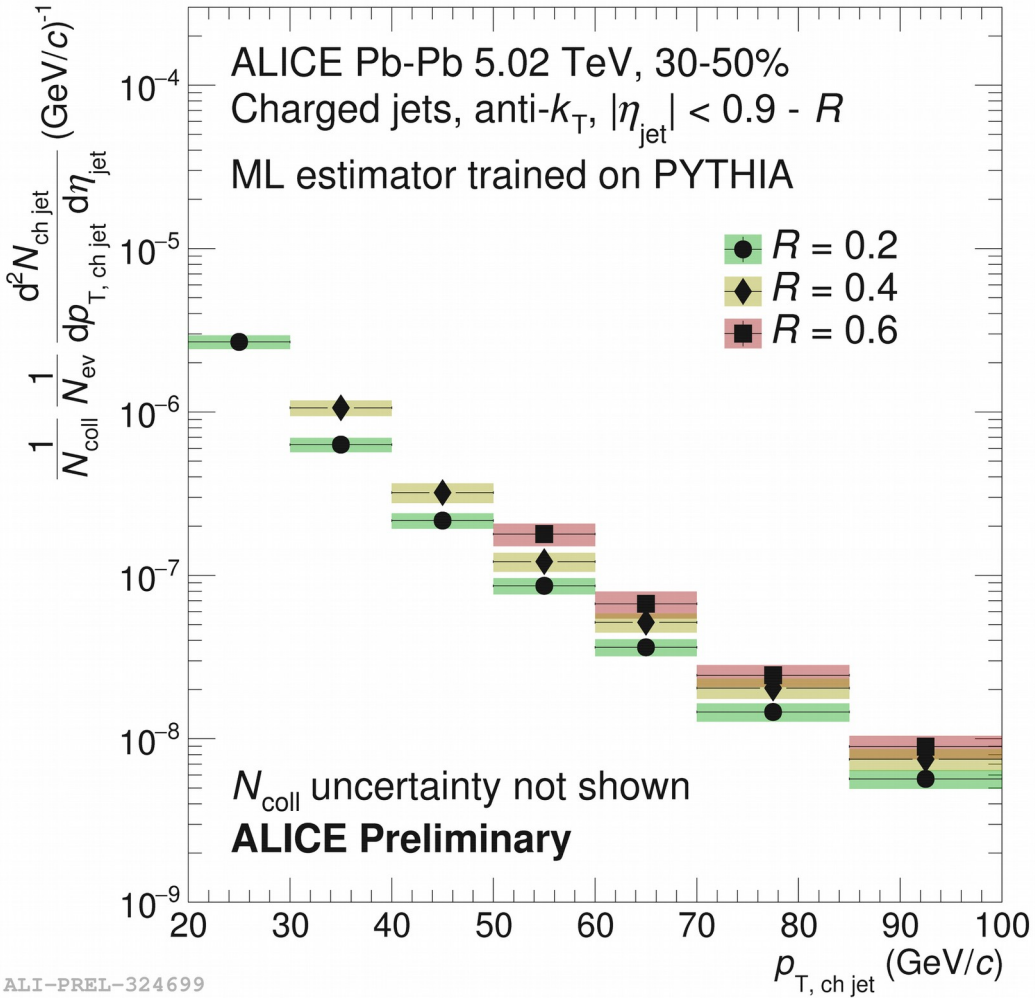
- Impact when unfolding or training assuming e.g. JEWEL fragmentation
- Build a classifier: Instead of approximating background, just discard combinatorial jets
  → Could solve main problem for unfolding
  → Might be less fragmentation dependent
- Try to train deep model that generalizes more on background than on the fragmentation
  → Train with wide variety of fragmentation models

Thank you for your attention!

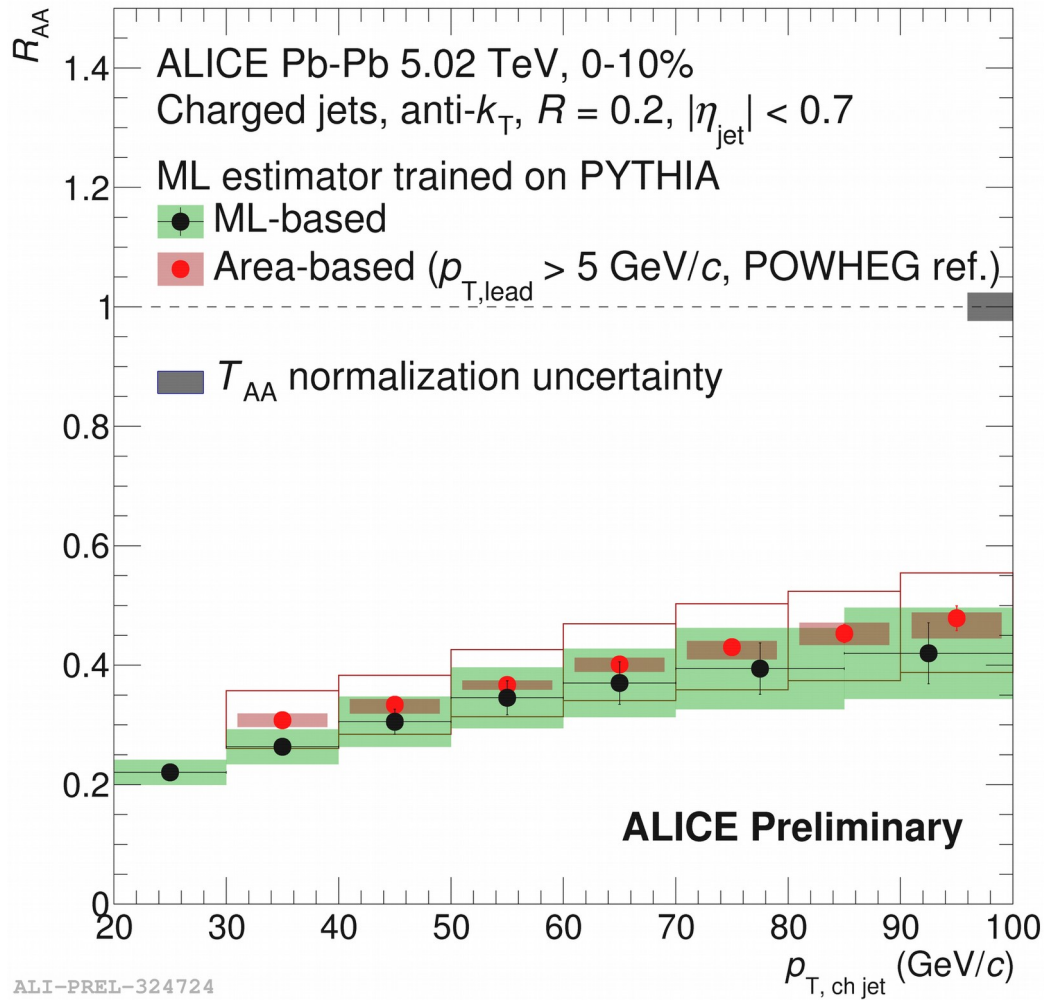
Backup

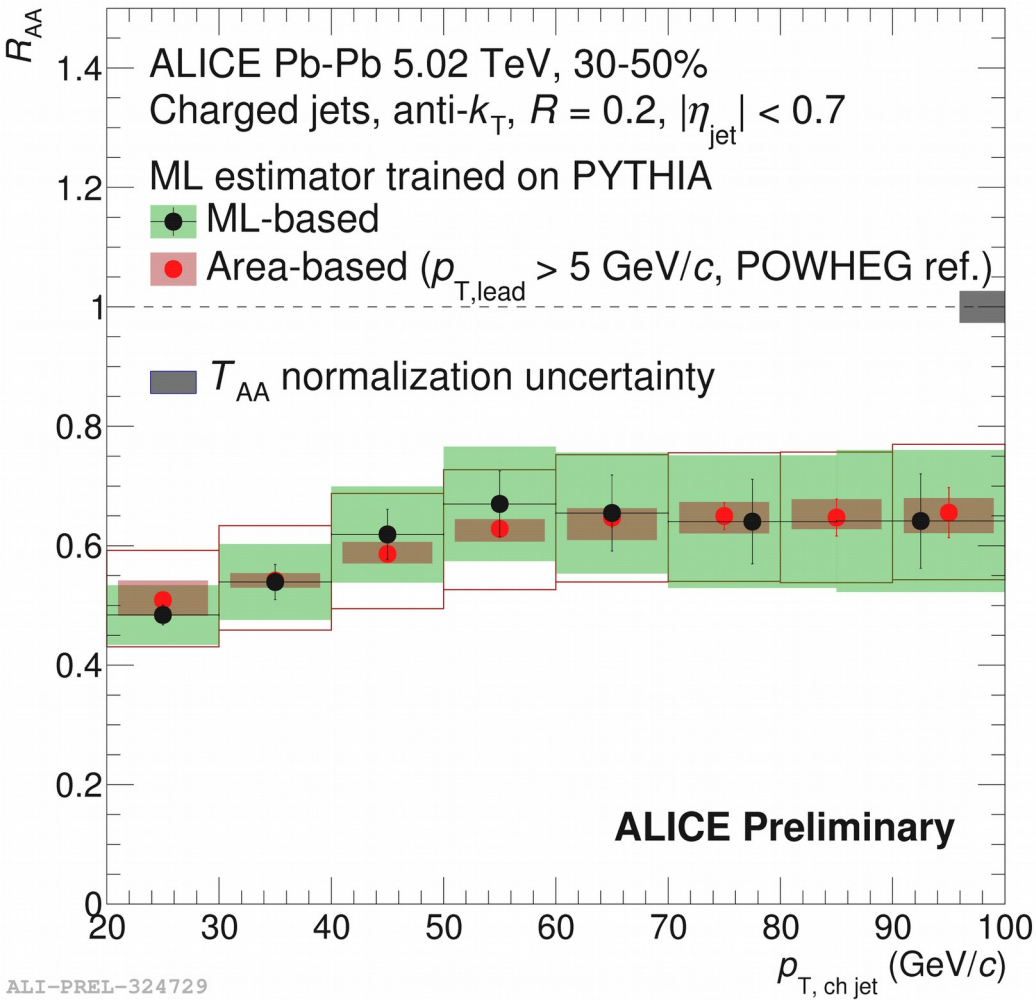
Fully corrected charged jet yields





Nuclear modification factor R_{AA} for R=0.2





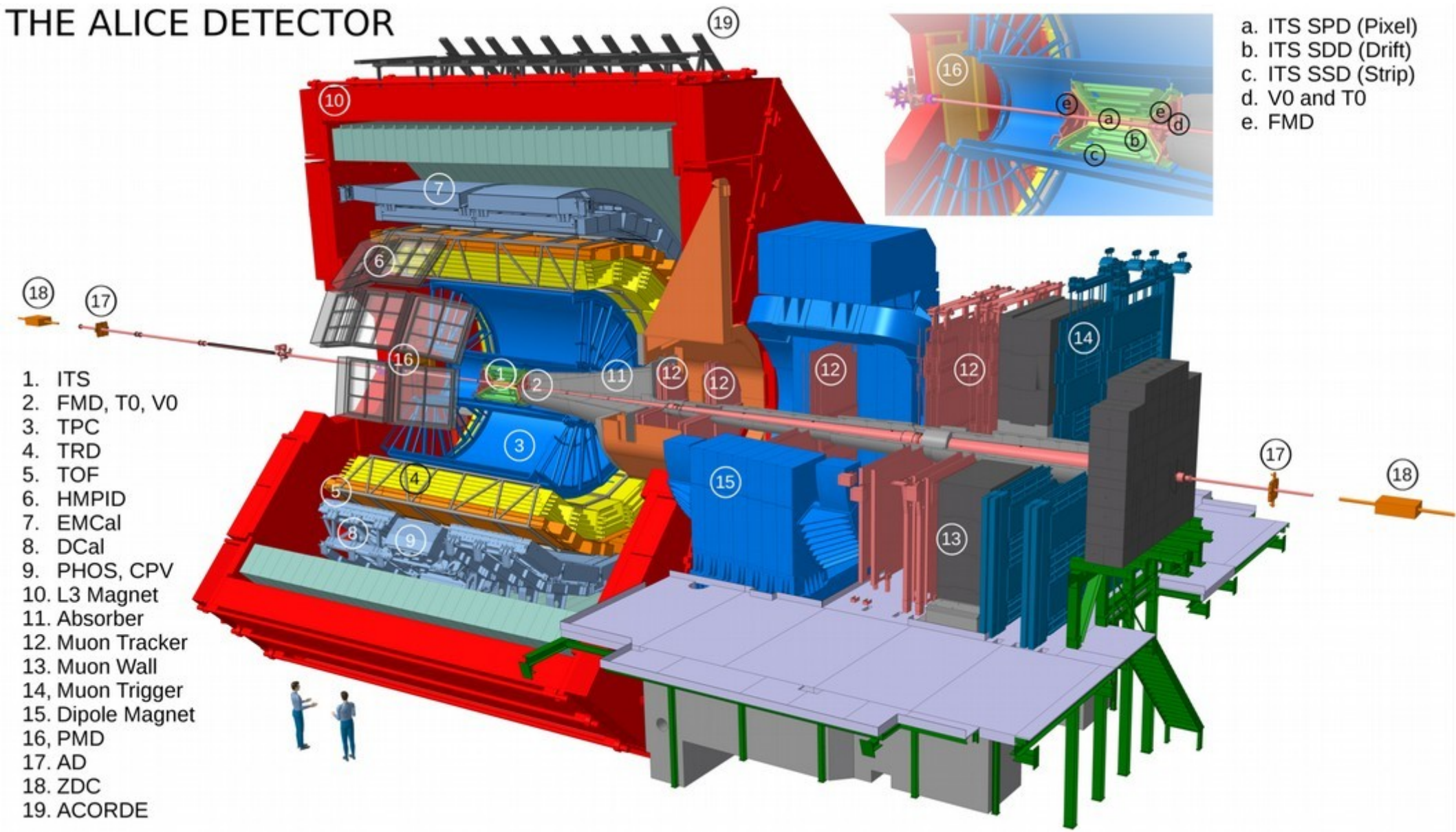
Systematic uncertainties

Uncertainties summed quadratically

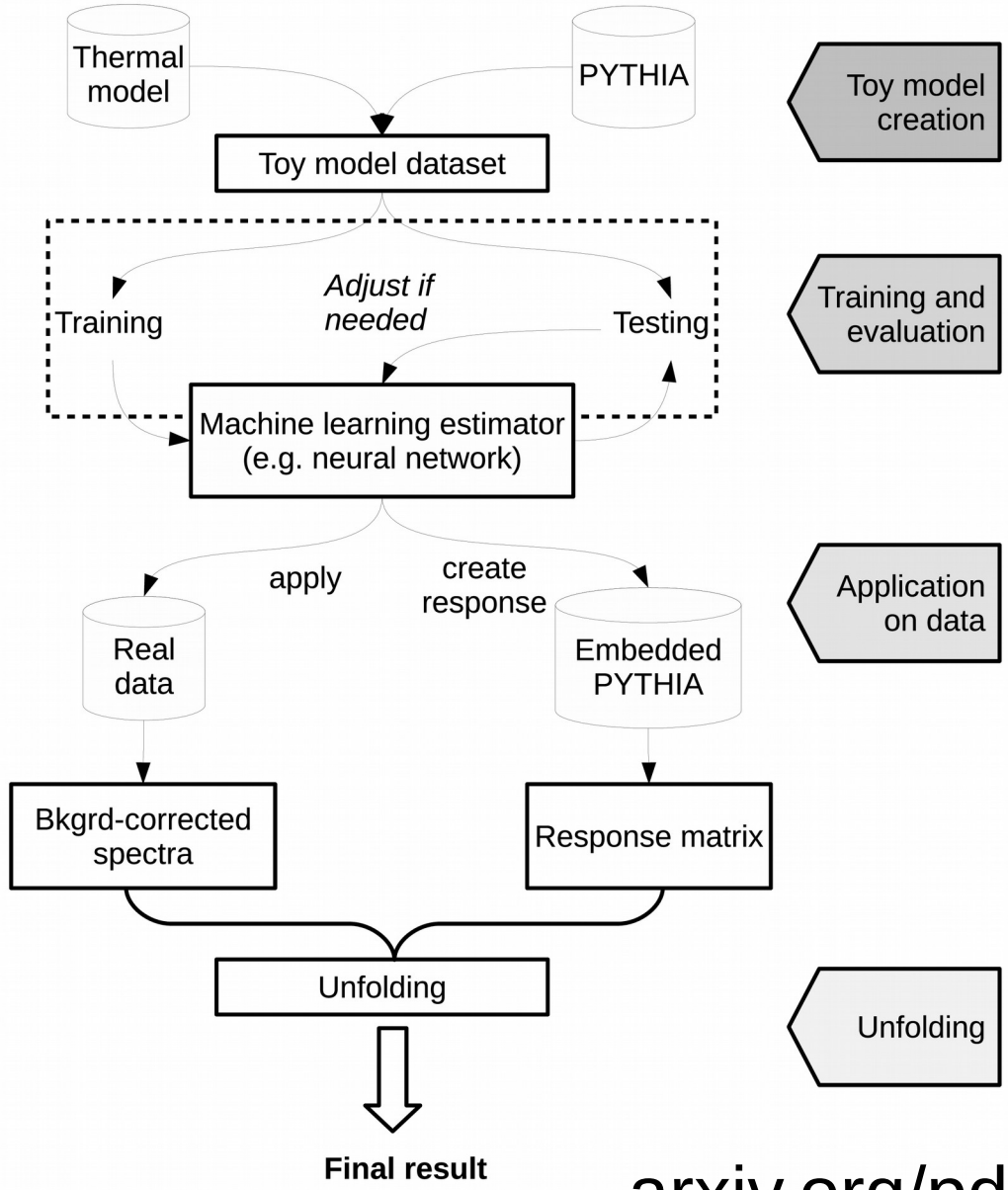
Variations taken into account

- Unfolding method (χ^2-unfolding)
- Unfolding regularization (higher/lower)
- Measured p_T range (higher/lower cut-off)
- Prior (PYTHIA for different radius, steeper/less steep)
- Tracking efficiency uncertainty (4% on tracks)
- Fragmentation uncertainty (response for quark-jets)

The ALICE detector



ML background estimator: Workflow



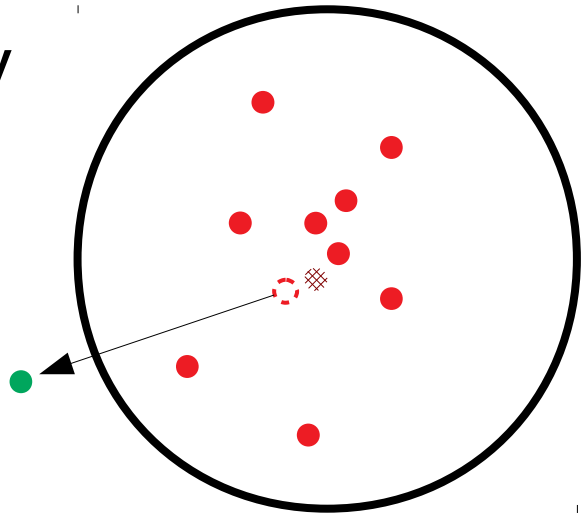
arxiv.org/pdf/1810.06324.pdf

Fragmentation dependence: Toy model

“Radiative broadening”:

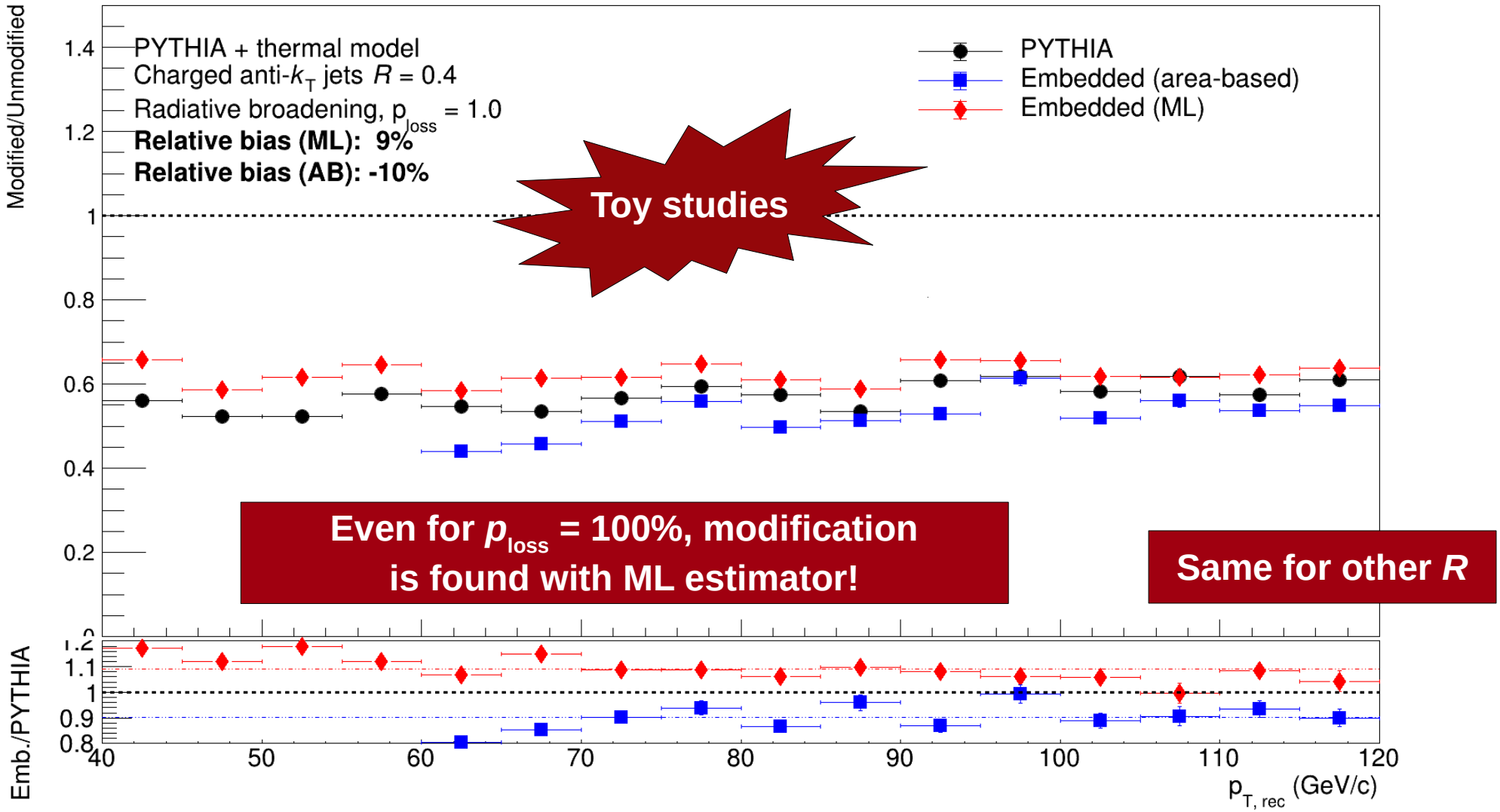
- For each PYTHIA constituent (above 1 GeV/c), radiate 1 GeV/*c* particle randomly $\Delta r = 0.6$ with given probability
- Vector calculation of kinematics → recoiling constituent
- Jet properties adjusted

- This simulates energy loss:
  1 GeV/*c* particles radiated dominantly outside jet cone



Would such an effect be correctly reconstructed by the ML estimator?
→ Yes, see next slide

Fragmentation dependence: Toy model



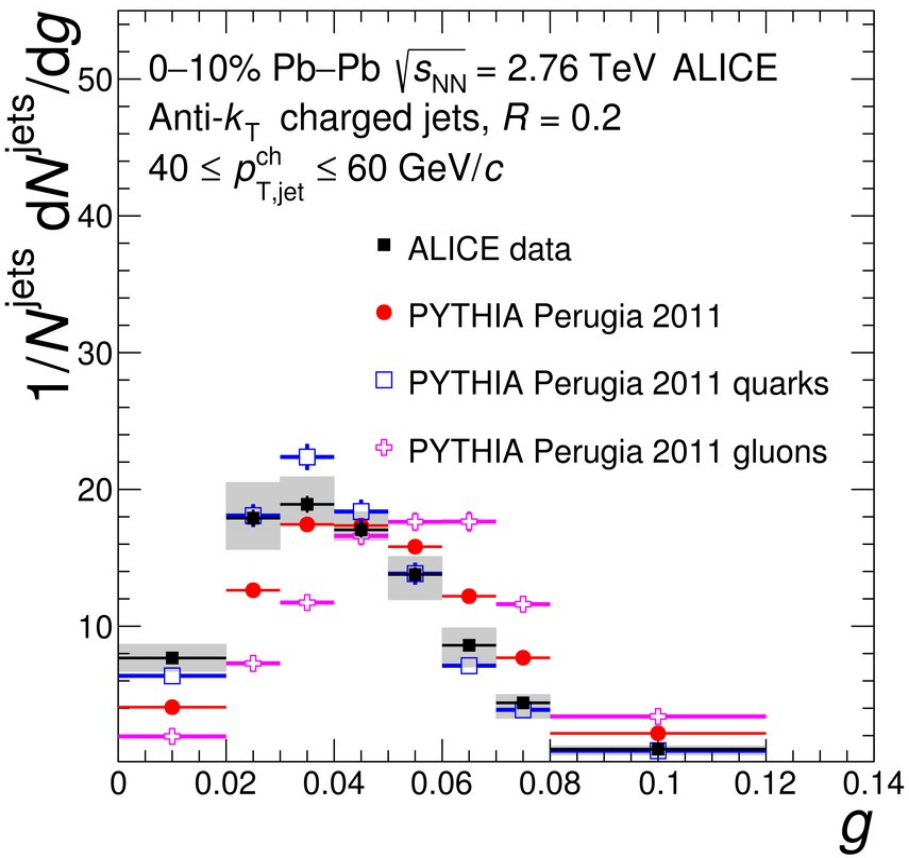
Even for p_{loss} = 100%, modification is found with ML estimator!

Same for other R

Fragmentation dependence

We measure pp-like jets → trained on PYTHIA8

- However, we checked the influence of a different jet assumption
  - → Unfold with response for quark-jets in PYTHIA
  - → Tested as well unfolding with gluon-jets only ✓
- Leads to a 10-20% uncertainty on the spectra



https://arxiv.org/pdf/1807.06854.pdf