Azimuthal correlations between Λ_c baryons and charged particles in pp collisions at $\sqrt{s}$ = 13 TeV

Samrangy Sadhu

Jadavpur University, Kolkata

ALICE-STAR India Collaboration Meeting, November 2023

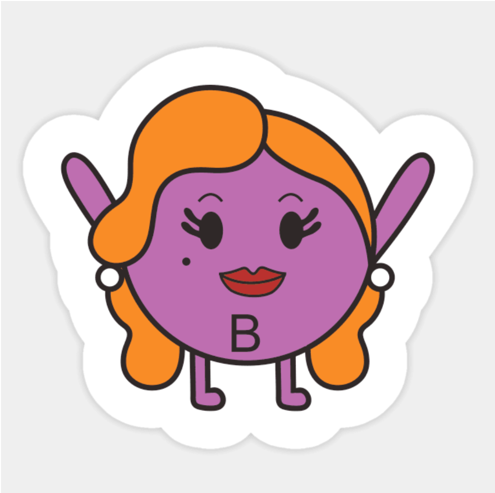
Heavy quarks : a unique probe

- Heavy quarks: charm and beauty, predominantly produced by the parton-parton hard scattering -> perturbative QCD can be applied.
- In heavy-ion collisions : Quark-gluon plasma (QGP) produced
  - -> heavy quarks are produced before QGP (t_{QGP} ~ 1 fm/c and $t_Q = 1/2m_Q \leq 0.1$ fm/c)
  - -> Experience the complete evolution of QGP medium

Charm

m_c ~ 1.3 GeV/c^2

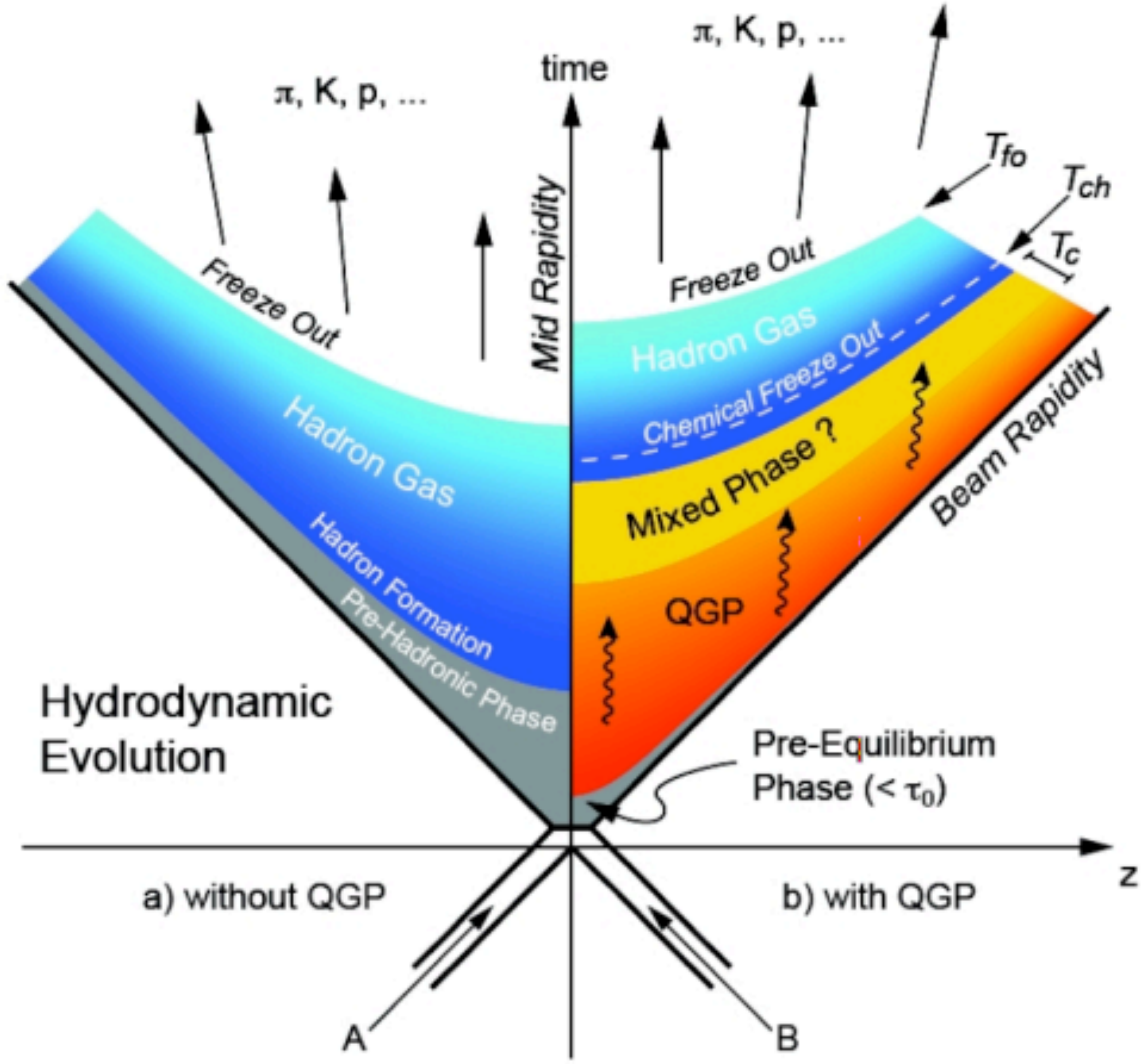
t_c ~ 0.08 fm/c



Beauty

m_b ~ 4.2 GeV/c^2

t_b ~ 0.03 fm/c



- QCD energy loss is expected to occur via both inelastic (radiative energy loss via medium-induced gluon radiation) and elastic (collisions with the QGP constituents) processes.

★ **Therefore, heavy quarks act as important tools for characterizing the medium formed in heavy-ion collisions.**

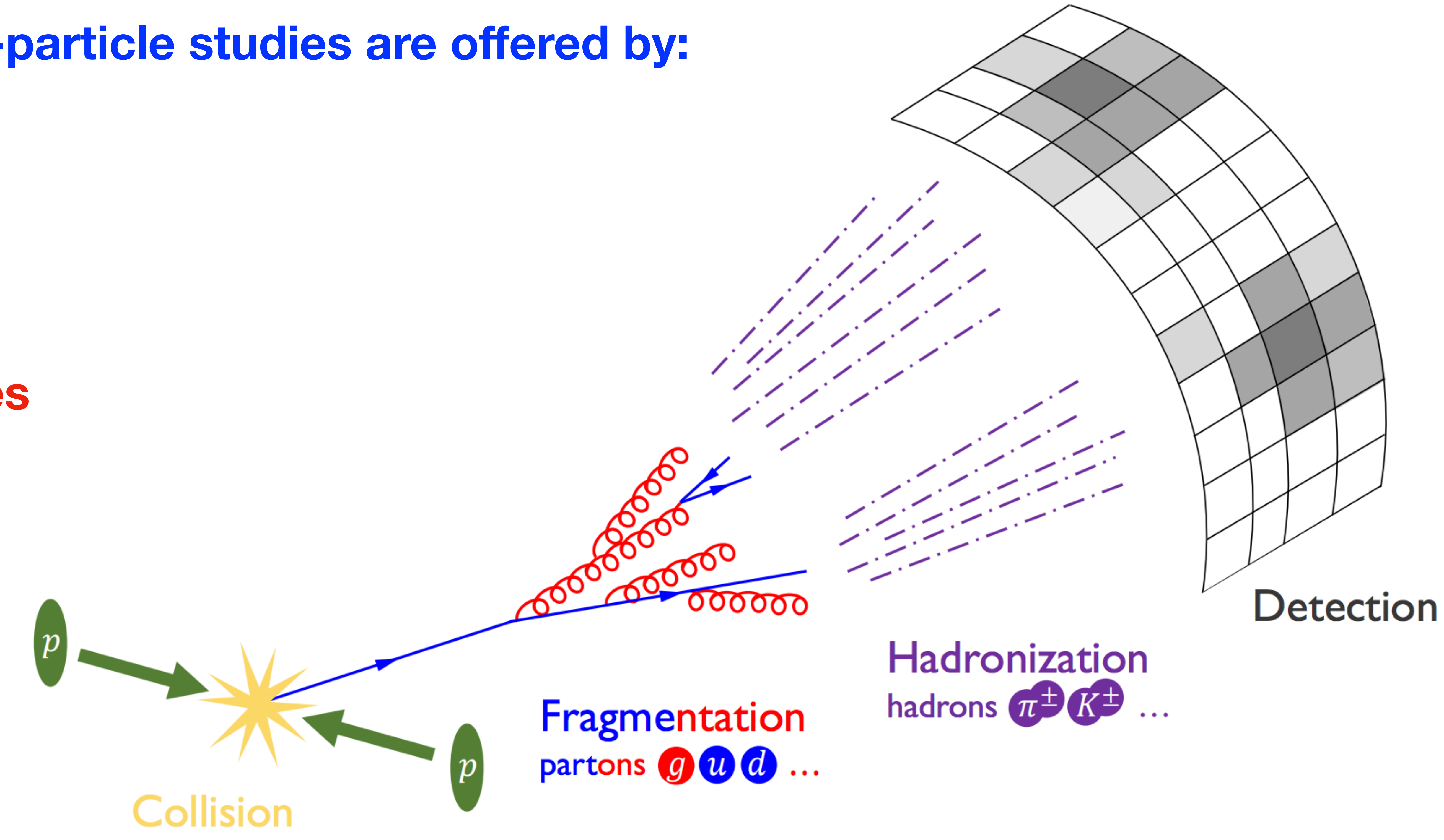
Heavy-flavour fragmentation through correlation and jet measurements

Fragmentation -> shower of quarks and gluons due to the strong force, mediated by gluons -> a jet of collimated particles

Regarding fragmentation, additional insights compared to single-particle studies are offered by:

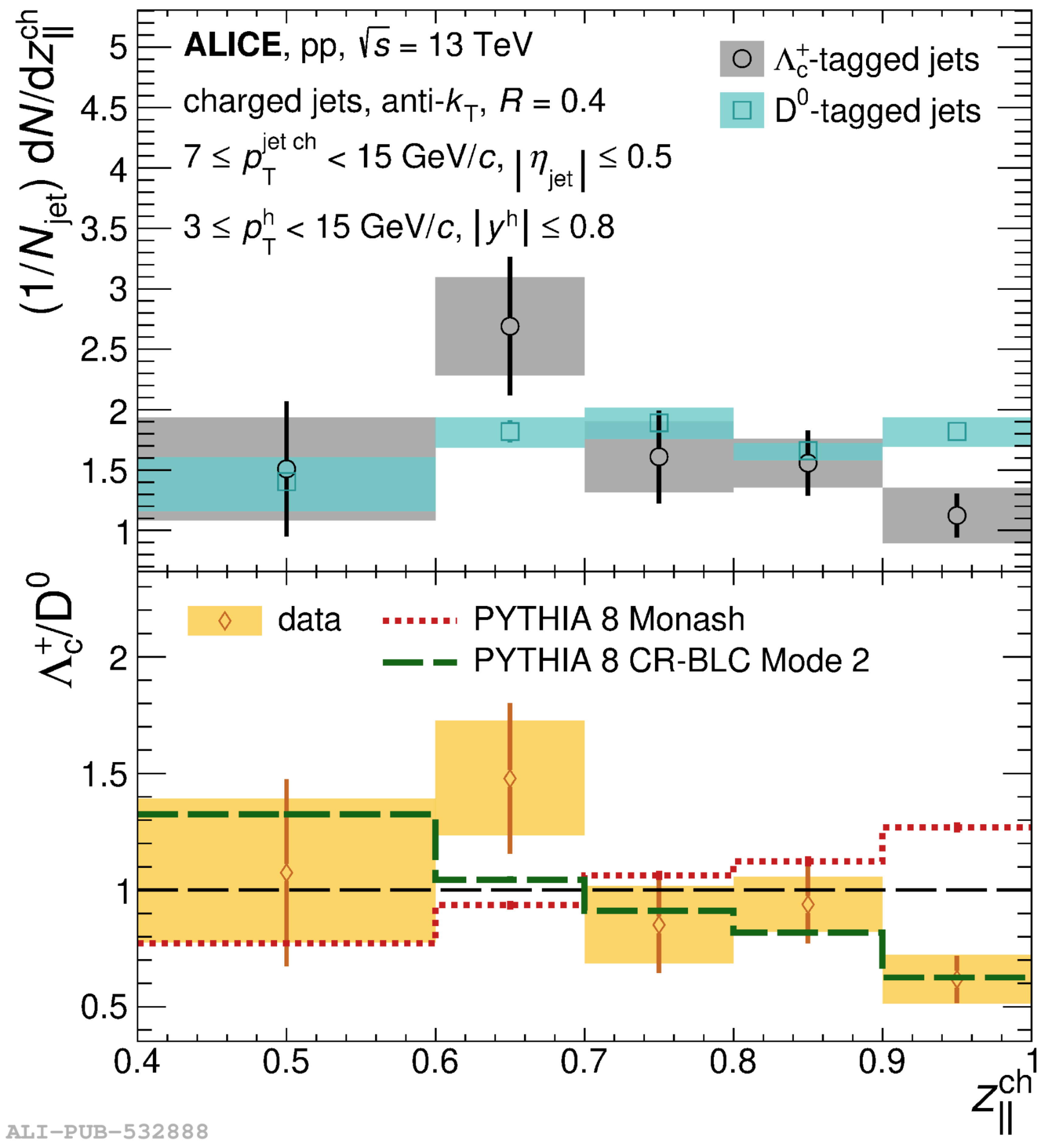
→ Charm-hadron tagged jets:
- **→ access to the original parton kinematics**
- **→ constrain the fragmentation functions**

→ Azimuthal correlations of charm hadrons with charged particles
- **→ description of the jet shape and its particle composition**
- **→ sensitivity to production mechanisms**



Exploring Λ_c-h azimuthal correlations

arXiv:2301.13798



Why do we study HF azimuthal correlations in pp collisions?

- → Characterise the charm production
- → Detail the internal structure of charm-induced jets
- → Validate MC models reproducing the charm showering
- → Reference for larger collision systems

First measurement of Λ_c-h azimuthal correlations:

→ detail charm fragmentation when the final state is a baryon
  → hint of softer fragmentation in Λ_c-jets than D^0-jet

→ influence of different production mechanisms
(coalescence, color reconnection, decay from higher-mass charm states)

→ benchmark for MC models

Analysis Strategy

- **Selection of Λ_c^+ baryon candidates through Machine Learning**
- **Angular correlation distributions between Λ_c candidates and "associated" particles**
- **Corrections:**
  - Λ_c ("trigger") and associated track efficiency → online weighting for Λ_c^+ and associated part reconstruction
  - Event mixing and sideband subtraction → detector inhomogeneities and background contribution
  - secondary track contamination → secondary particles (interaction with detector material …)
  - $b \rightarrow \Lambda_c$ bias from decay topology → bias in correlation induced by the candidate selection
  - beauty feed-down → subtraction via Pythia8 templates
  - Soft-π → Remove contributions from $\Sigma_c^{0,++}(2455) \rightarrow \Lambda_c^+ + \pi^{\mp}$
- **Systematics**
- **Study of correlation properties**

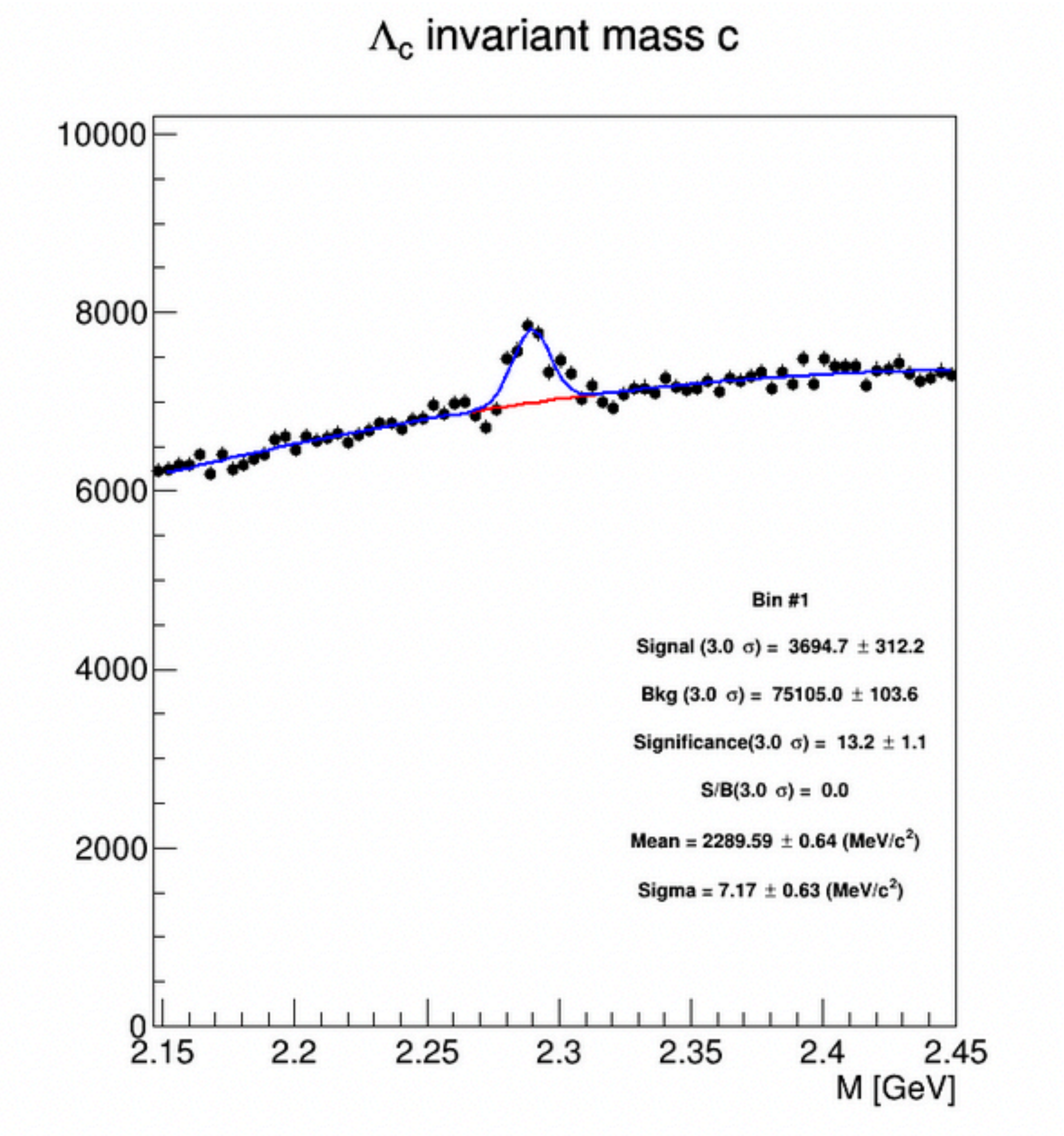
Data periods :
2016, 2017, 2018

MonteCarlo productions:
HF enriched:
For ML and trigger efficiency estimation
General-purpose:
For associated track efficiency estimation

Reconstruction of Λ_c baryons

- Decay channel : $\Lambda_c \rightarrow pK^-\pi^+$

- Selection of Λc candidates based on **Machine Learning** multi-classification algorithm - XGBoost (integrated in hipe4ml package)
- Preselection: single-track quality criteria, loose topological and PID selections
- The second stage of the selection foresees the use of a binary classification algorithm to separate the two contributions, prompt Λc and combinatorial background



Fit:
Background: 2^{nd} order pol.
Signal: Gaussian.

S and B, extracted from the efficiency weighted distributions, are used in the correlation extraction

Two-particle azimuthal correlations

Two- particle angular correlation function is characterized by per-trigger yield of associated charged particle :

Associated hadron yield per trigger:

$$\frac{1}{N_{trig}}\frac{d^2N^{pair}}{d\Delta\eta d\Delta\phi} = B(0,0) \times \frac{S(\Delta\eta,\Delta\phi)}{B(\Delta\eta,\Delta\phi)}$$

$\Delta\eta = \eta_1 - \eta_2$
$\Delta\varphi = \varphi_1 - \varphi_2$

where,

$$S(\Delta\eta,\Delta\phi) = \frac{1}{N_{trig}}\frac{d^2N^{same}}{d\Delta\eta d\Delta\phi}$$

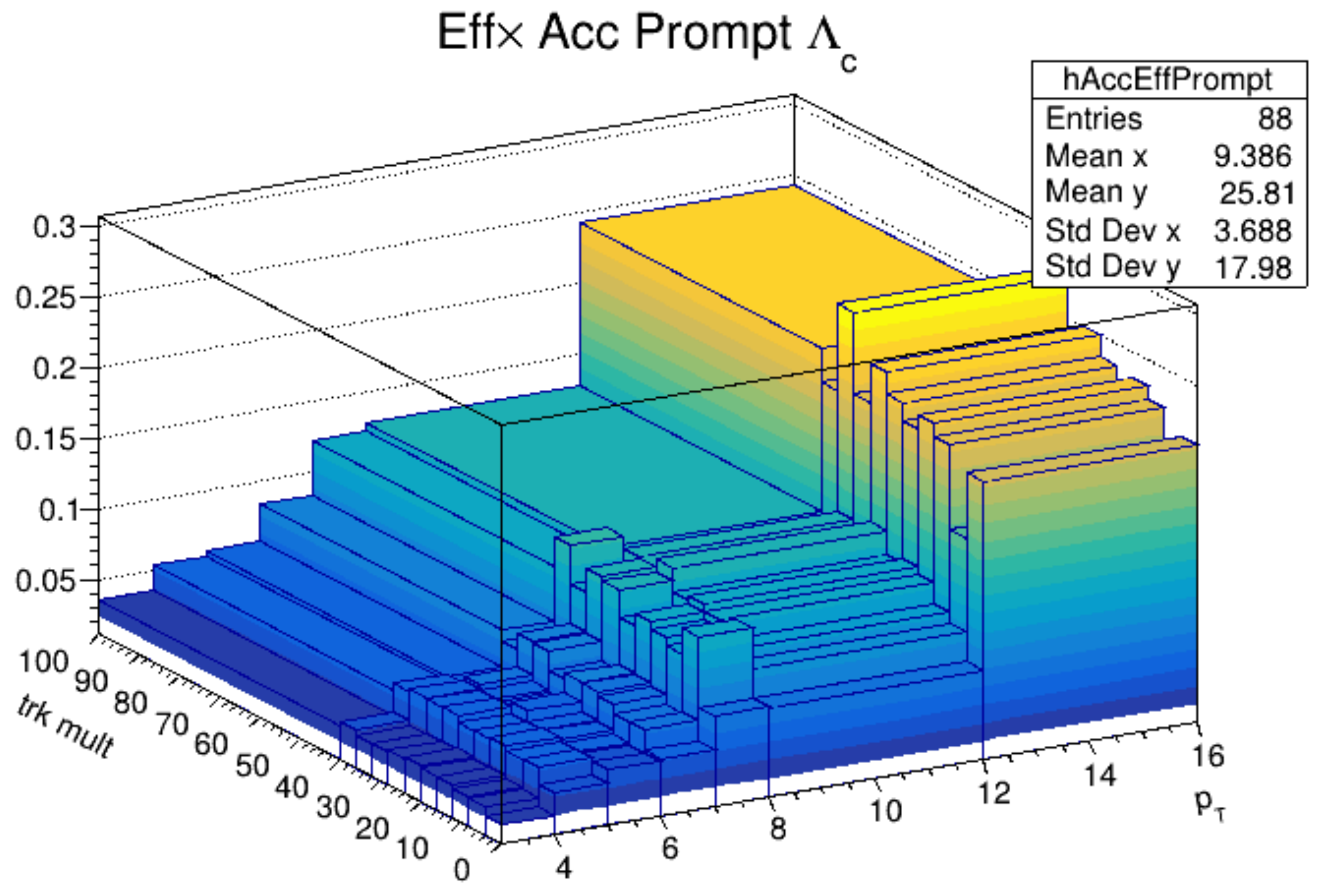
→ **The differential measure of per-trigger distribution of the associated charged particles in the same event**

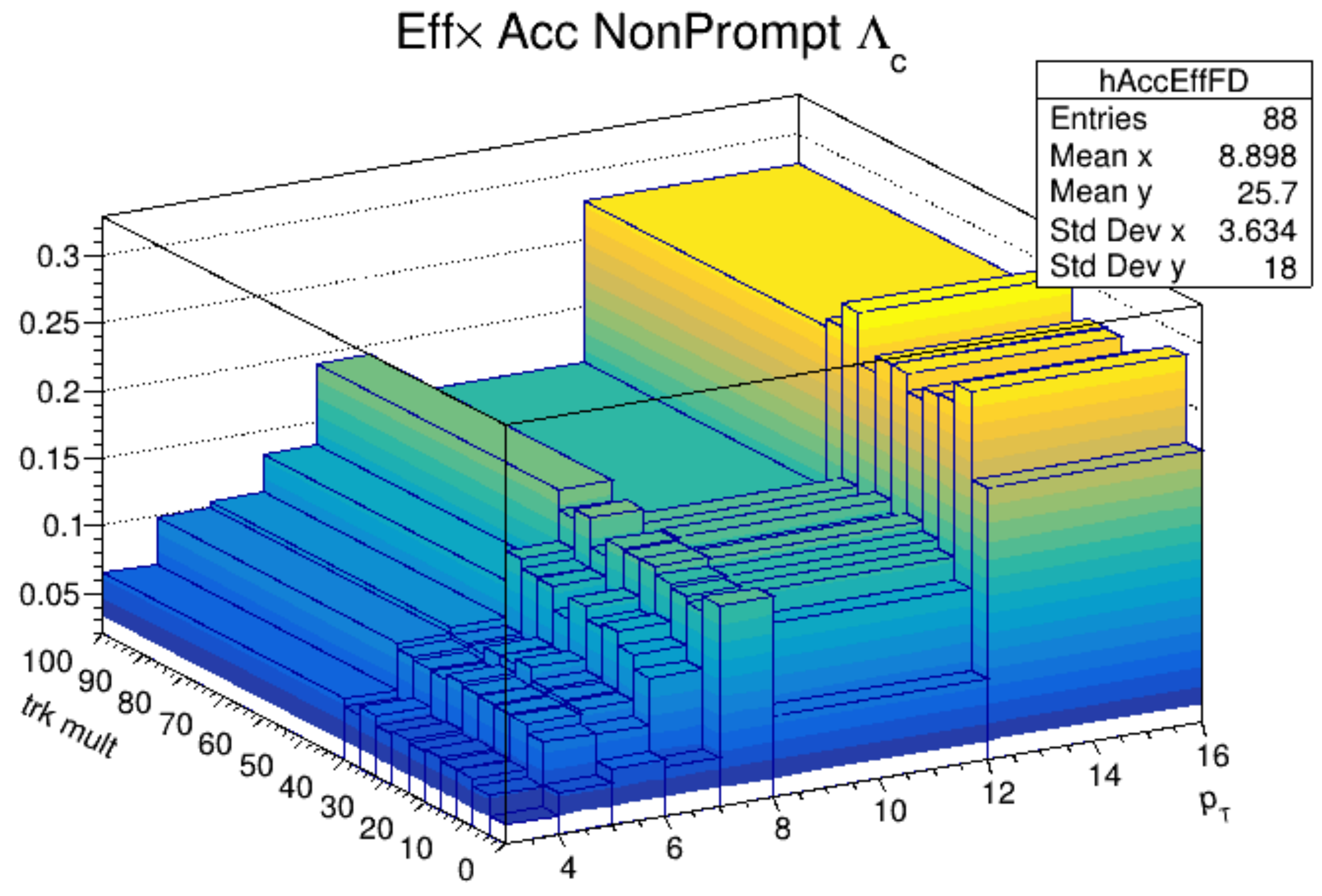
$$B(\Delta\eta,\Delta\varphi) = \frac{d^2N^{mixed}}{d\Delta\eta d\Delta\varphi}$$

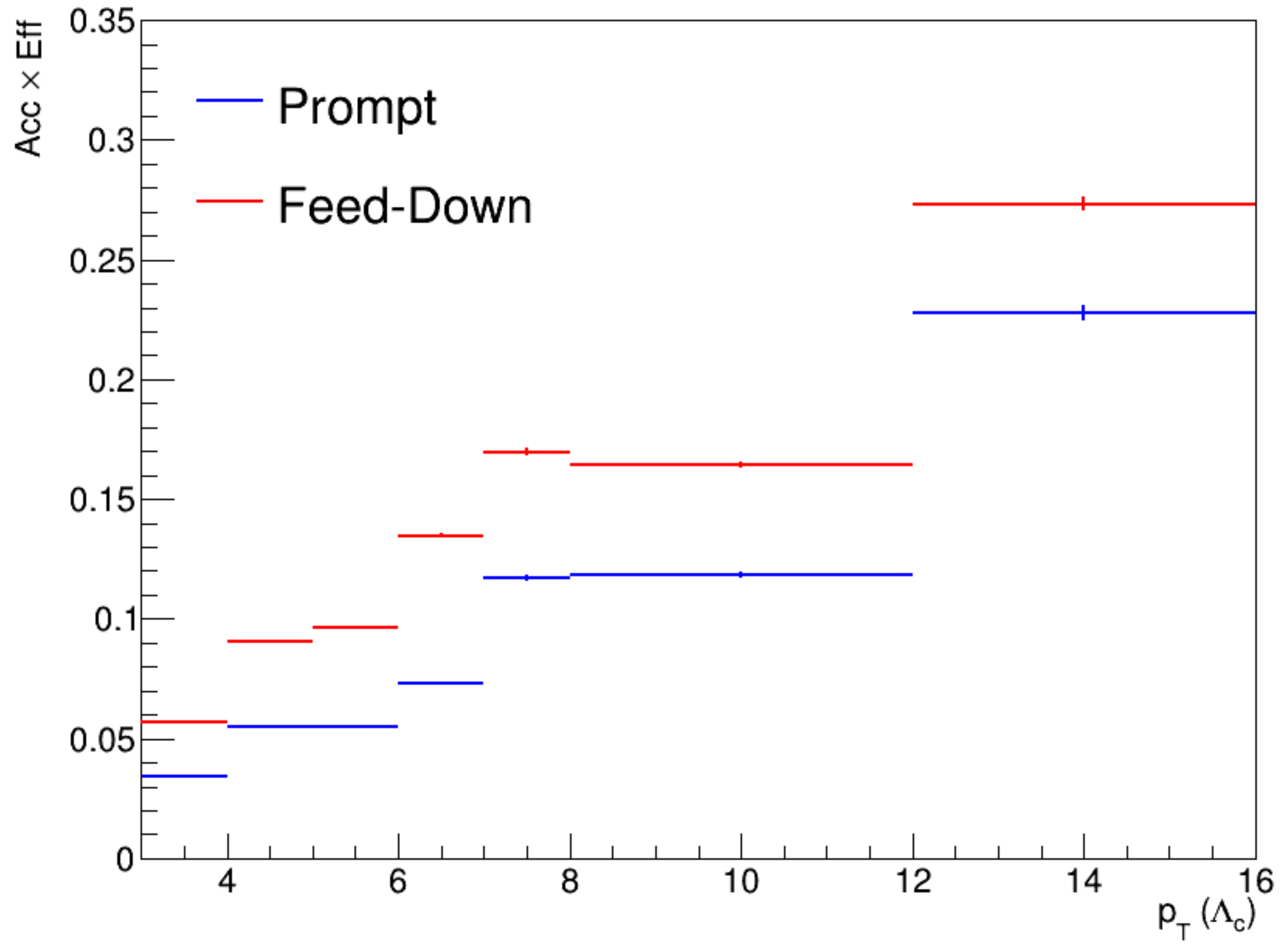
→ **The background distribution function, where N^{mixed} is the number of mixed event pairs**

Efficiency correction

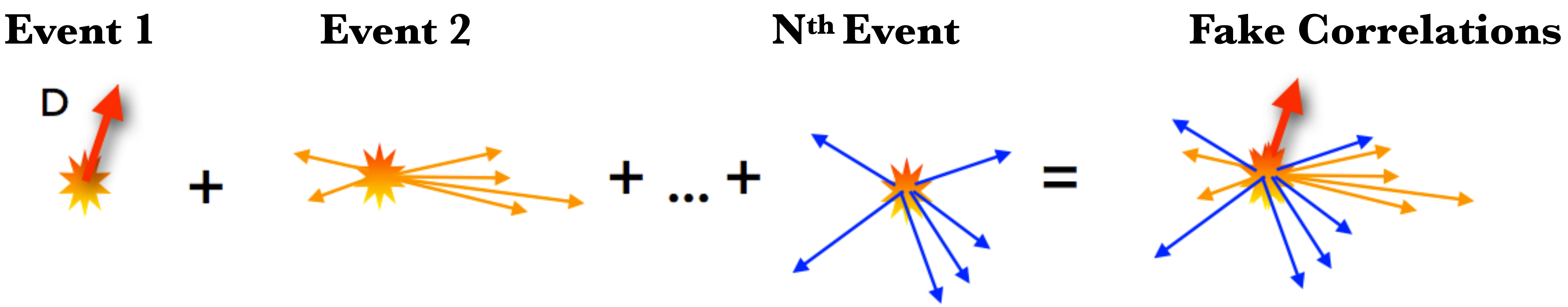
- **The limited detector acceptance and efficiency for the reconstruction and selection of trigger candidates and associated particles has been accounted.**
- **2-D trigger efficiency maps are prepared as a function of pT and multiplicity (for trigger efficiency correction) and 3-D tracking efficiency maps are prepared as a function of pT, η and z-vtx position (for tracking efficiency correction).**
- **Enriched and general purpose MC samples are used for trigger and tracking efficiency corrections respectively.**
- **Correlation entries are weighted by 1/[ε(trig)* ε(track)]**







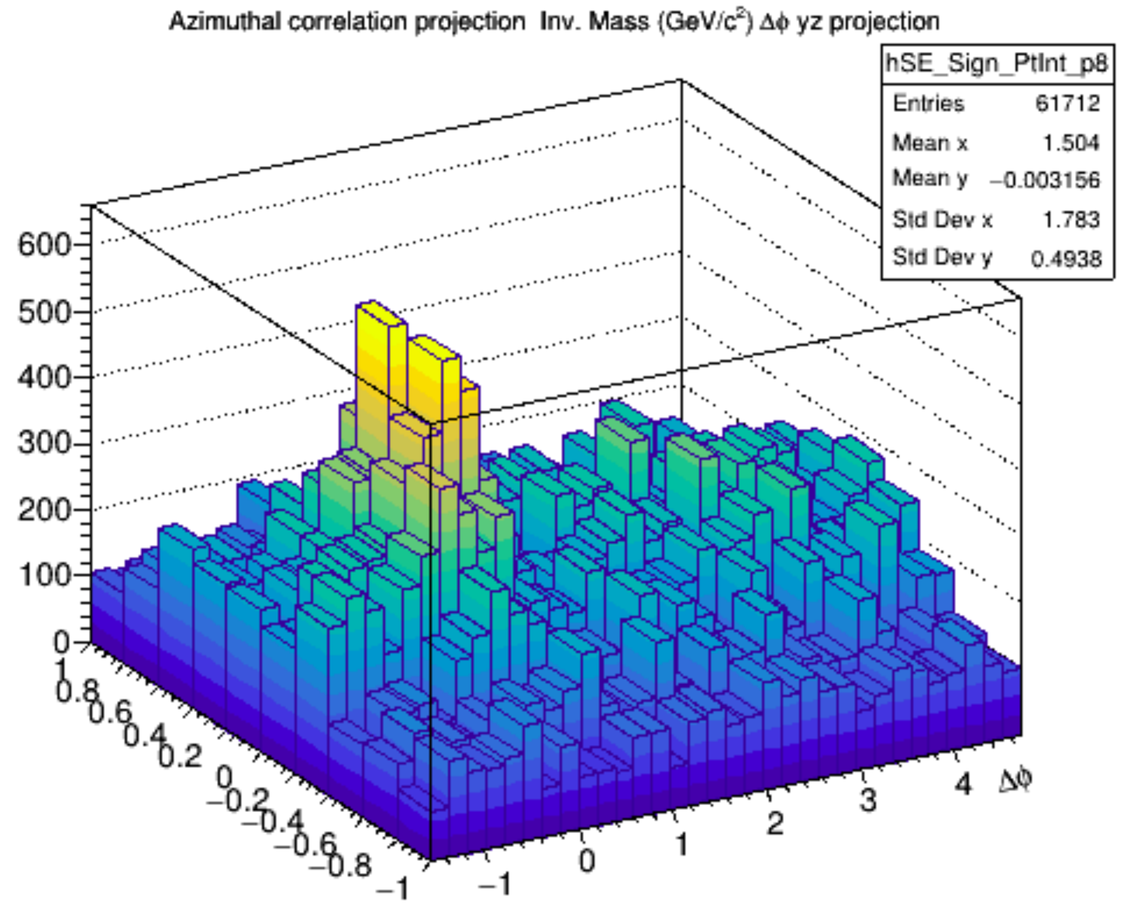
Event mixing technique

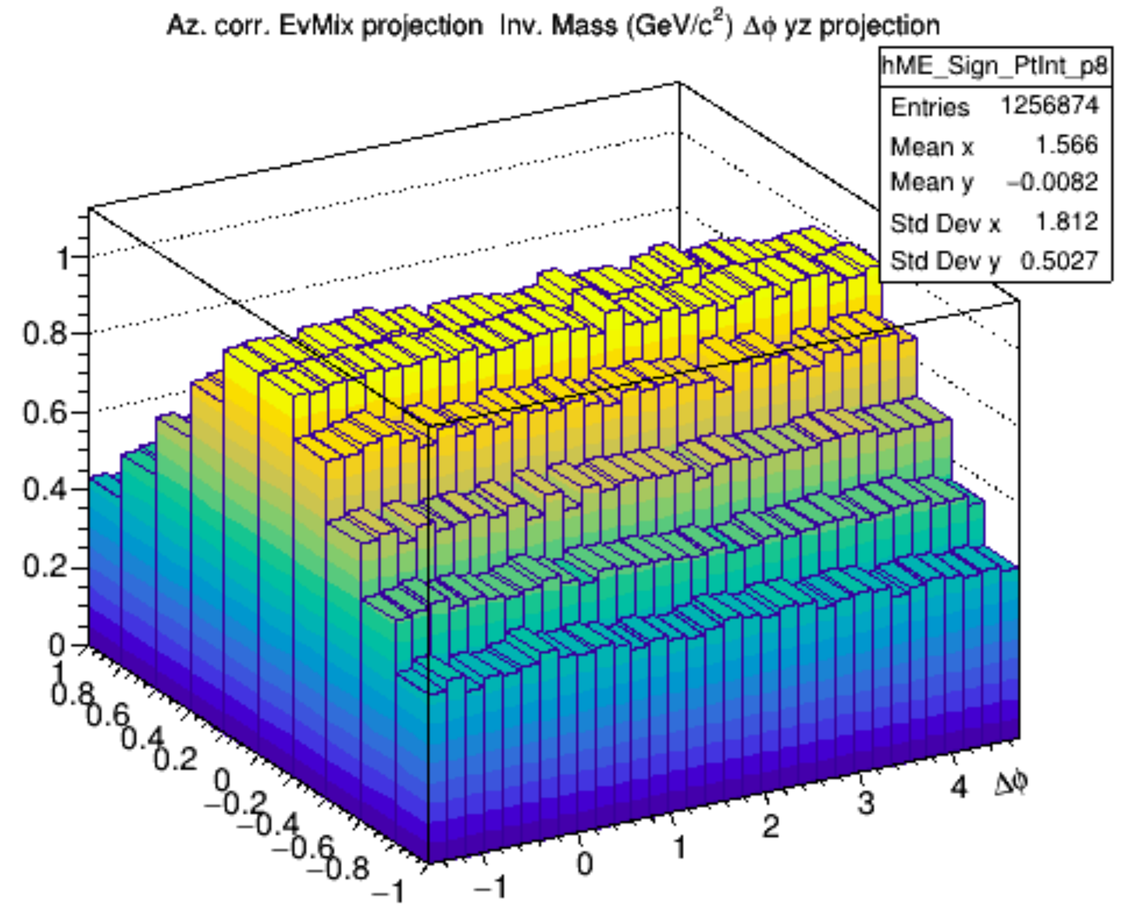


- The correlation distributions are corrected for the limited detector acceptance and detector spatial inhomogeneities using the event-mixing technique.
- Mixed events are obtained by taking the trigger candidate from the event N and the associated tracks from other preceding selected events.
- Every processed event is stored in an event pool based on its topology. The pools store events based on selection of multiplicity and primary vertex position.

Event mixing correction and Background subtraction

The correlation distributions are corrected for the limited detector acceptance and detector spatial inhomogeneities using the **event mixing technique**.





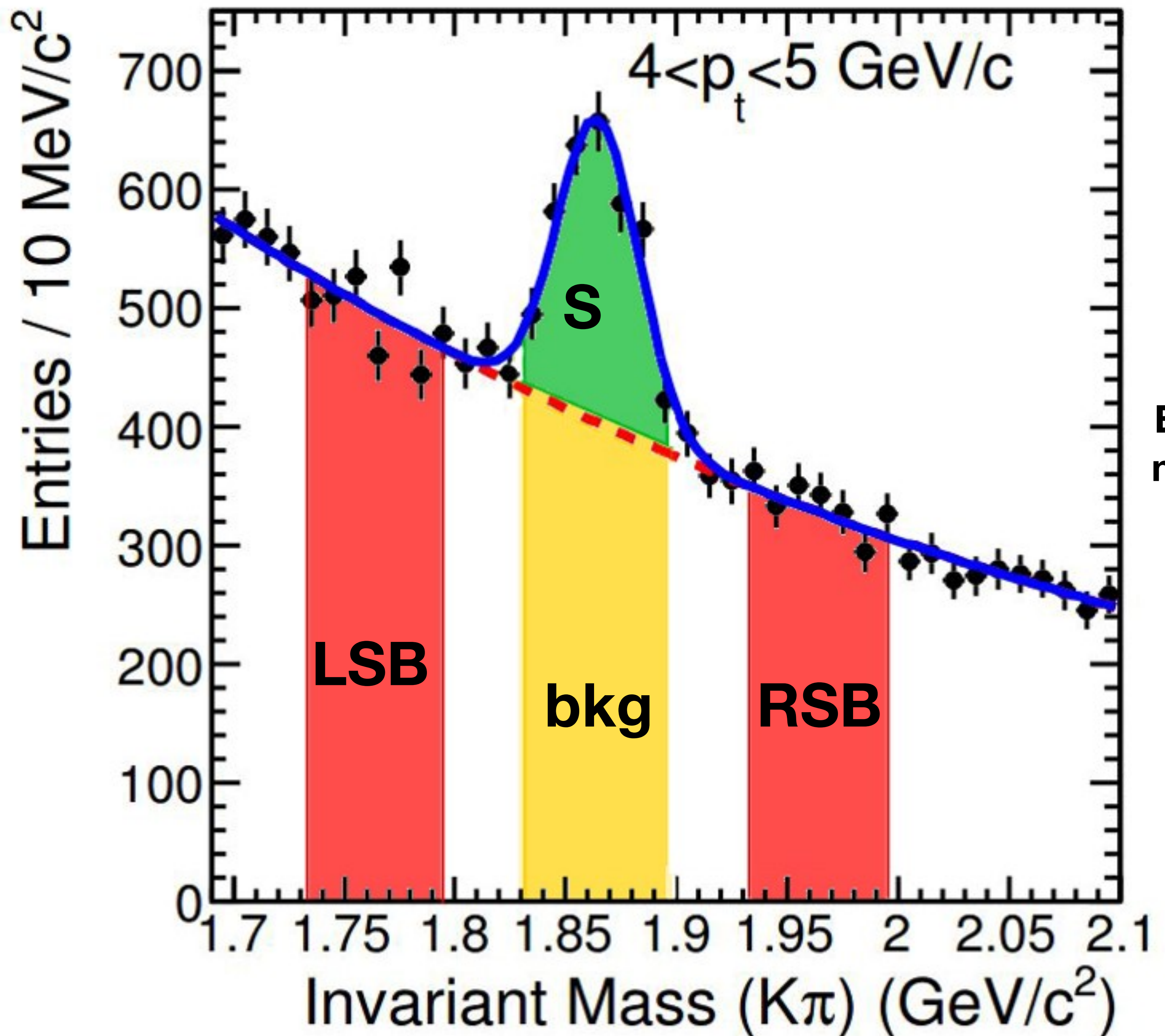
$$\frac{d^2N^{corr}(\Delta\varphi\Delta\eta)}{d\Delta\varphi d\Delta\eta} = \frac{\frac{d^2N^{SE}(\Delta\varphi\Delta\eta)}{d\Delta\varphi d\Delta\eta}}{\frac{d^2N^{ME}(\Delta\varphi\Delta\eta)}{d\Delta\varphi d\Delta\eta}} \frac{d^2N^{ME}(0,0)}{d\Delta\varphi d\Delta\eta}$$

Removal of background under the signal peak : **Side-Band technique**

$$\left(\frac{d^2N^{corr}}{d\Delta\varphi d\Delta\eta}\right) = \left(\frac{d^2N^{corr}}{d\Delta\varphi d\Delta\eta}\right)_{s+bkg} - \frac{B}{B_{sb}}\left(\frac{d^2N^{corr}}{d\Delta\varphi d\Delta\eta}\right)_{LSB+RSB}$$

B=>The background under the signal mass peak that we estimate from the invariant mass fit

B_{sb} =>The integral of the invariant mass distribution in the side-band region



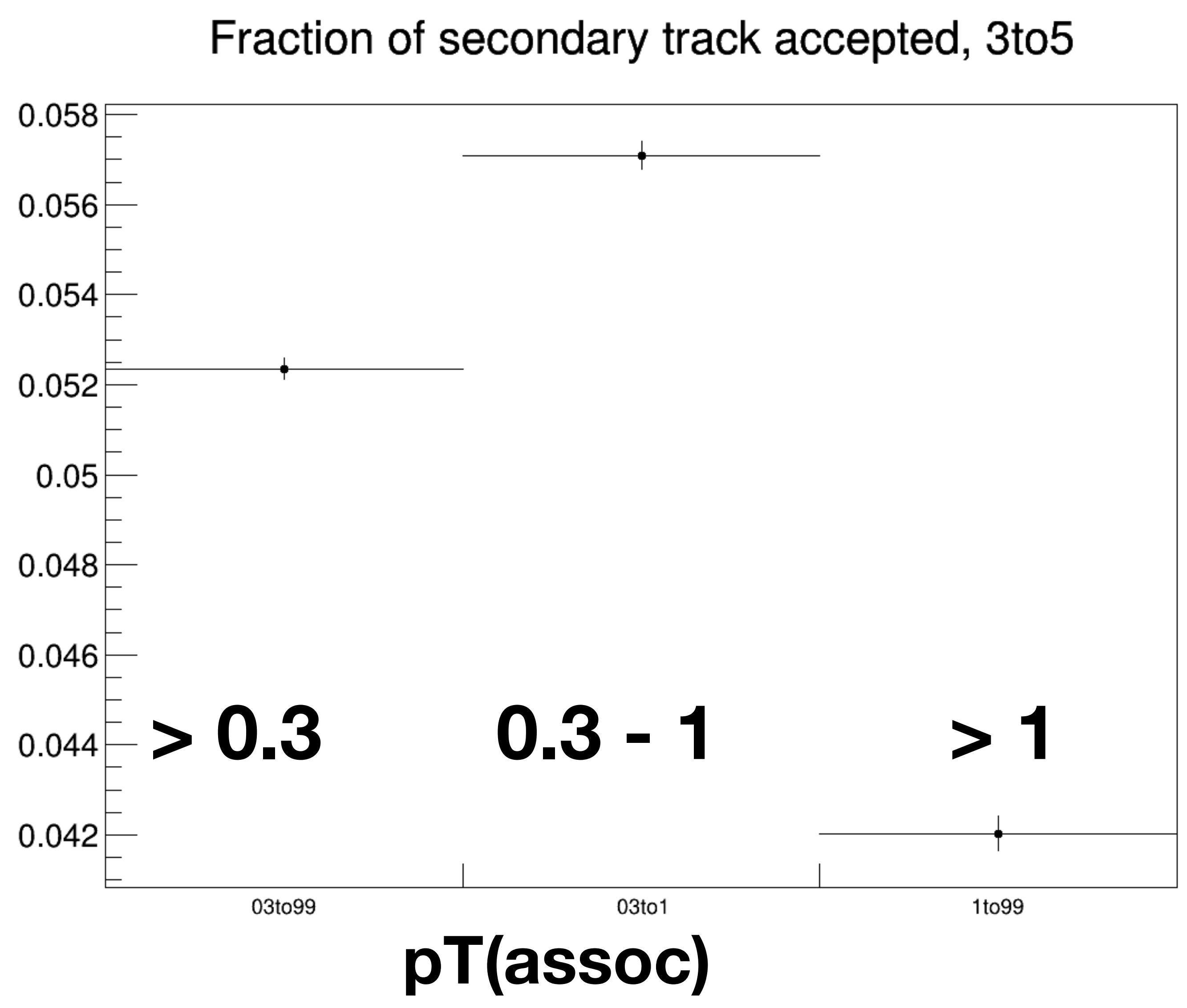
Example plot from D0 meson invariant mass

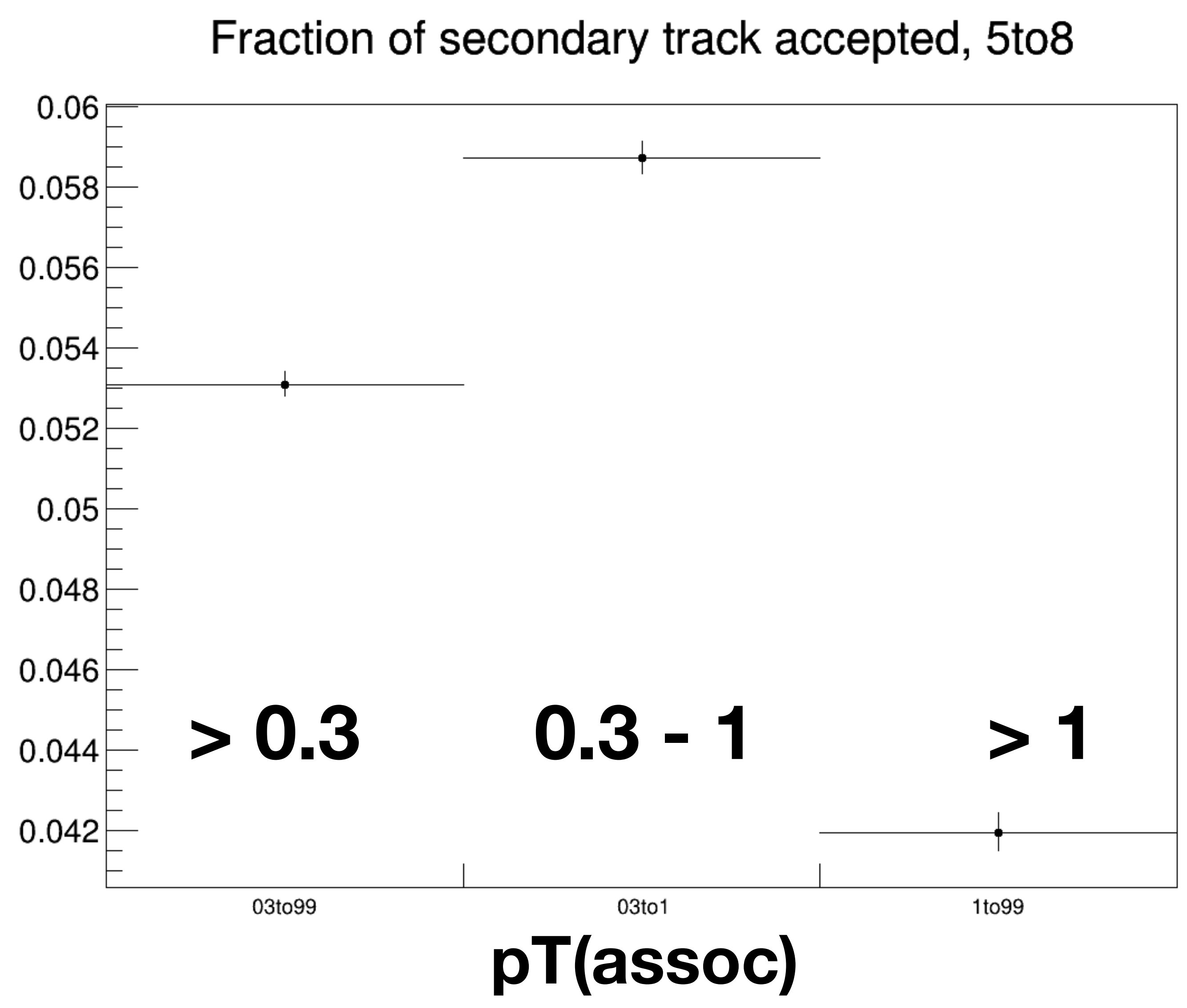
Removal of secondary particle contamination

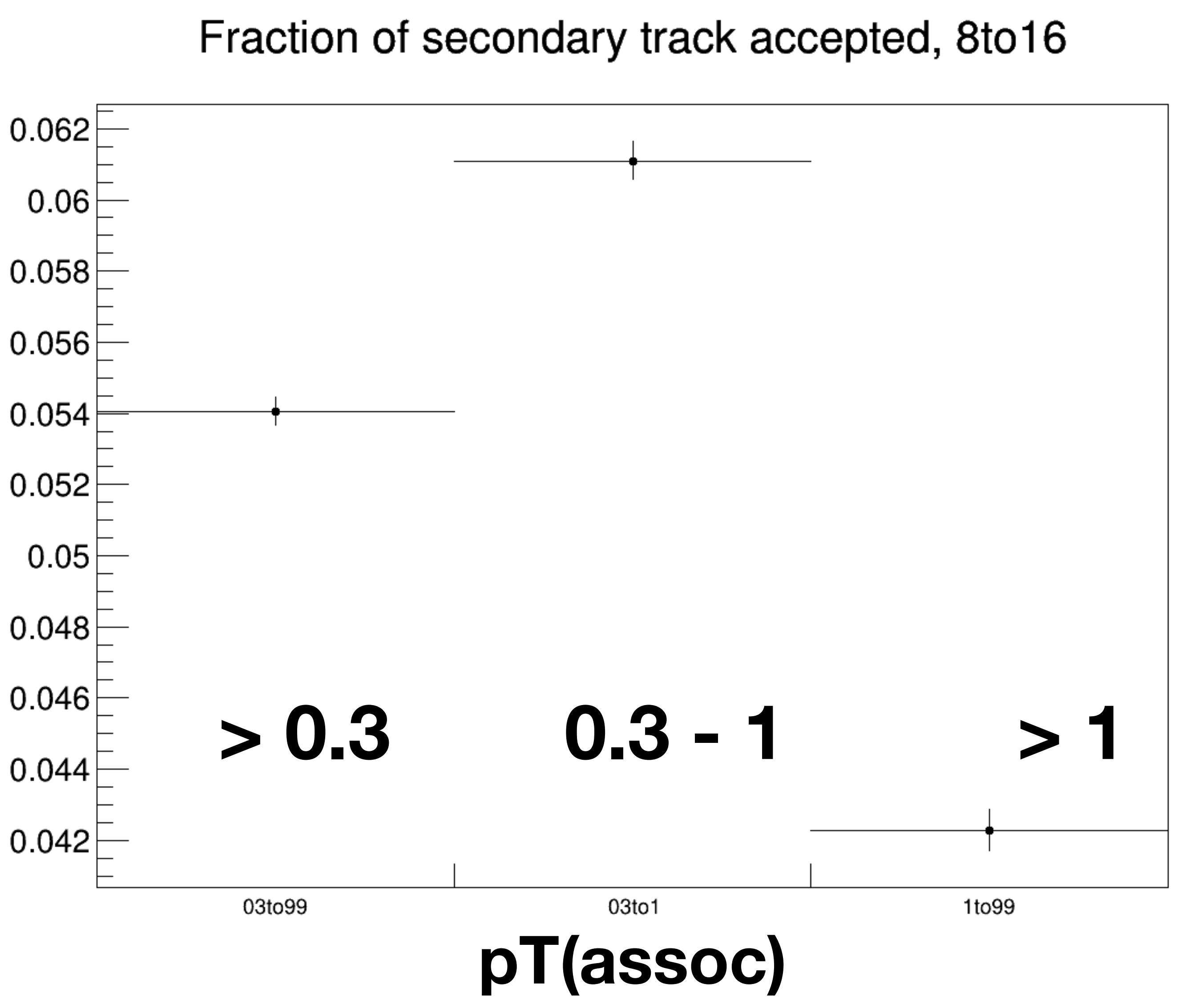
- Secondary particles surviving DCA cuts, generally coming from:
  – interaction of primary tracks with the detector material
  – decays of strange hadrons

- Ratio of $\Delta\varphi$ distribution of primary track accepted over all tracks distribution.
- From the ratio between the distribution
  → fraction of residual secondary track is flattish along $\Delta\varphi$
  → inhomogeneities not larger than ~1%

- A bin-by-bin correction was applied by multiplying the azimuthal correlation distribution to the moving average of secondary track $\Delta\varphi$ template.







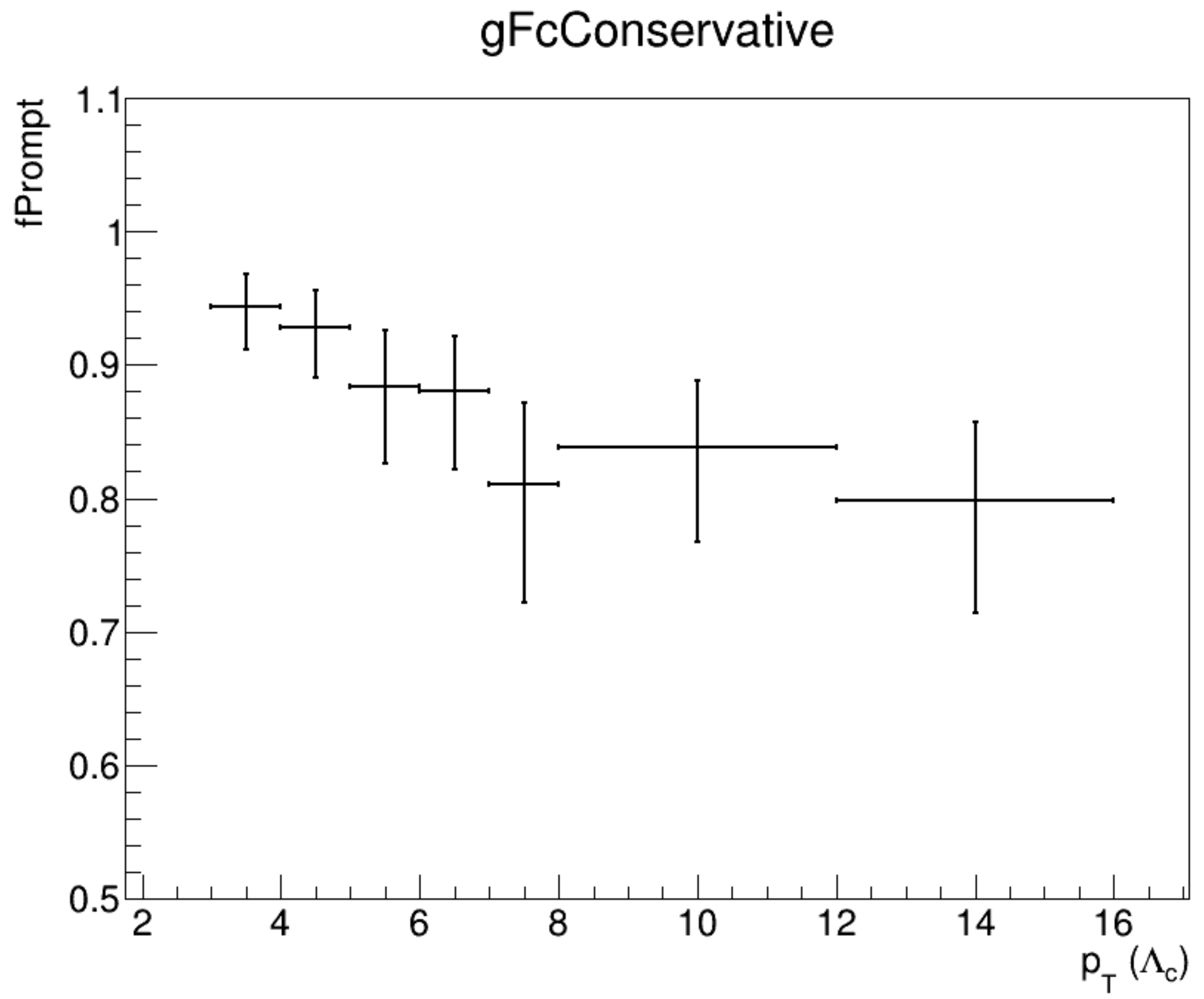
Feed-Down subtraction

f_{prompt} => fraction of prompt Λ_c in the reconstructed sample

$$f_{prompt} = 1 - \frac{N^{\Lambda_c\, from\, b}}{N^{\Lambda_c\, inclusive}}$$

$$where, \qquad N^{\Lambda_c\, from\, b} = 2\frac{d\sigma^{\Lambda_c\, from\, b}_{FONLL}}{dp_T}\Delta y \Delta p_t\,(Acc\times\varepsilon)_{feed-down}\,BR.L_{int}$$

$$f_{prompt} = \frac{(Acc\times\varepsilon)_{prompt}\frac{d\sigma^{\Lambda_c\, prompt}_{FONLL}}{dp_T}}{(Acc\times\varepsilon)_{prompt}\frac{d\sigma^{\Lambda_c\, prompt}_{FONLL}}{dp_T} + (Acc\times\varepsilon)_{feed-down}\frac{d\sigma^{\Lambda_c\, from\, b}_{FONLL}}{dp_T}}$$

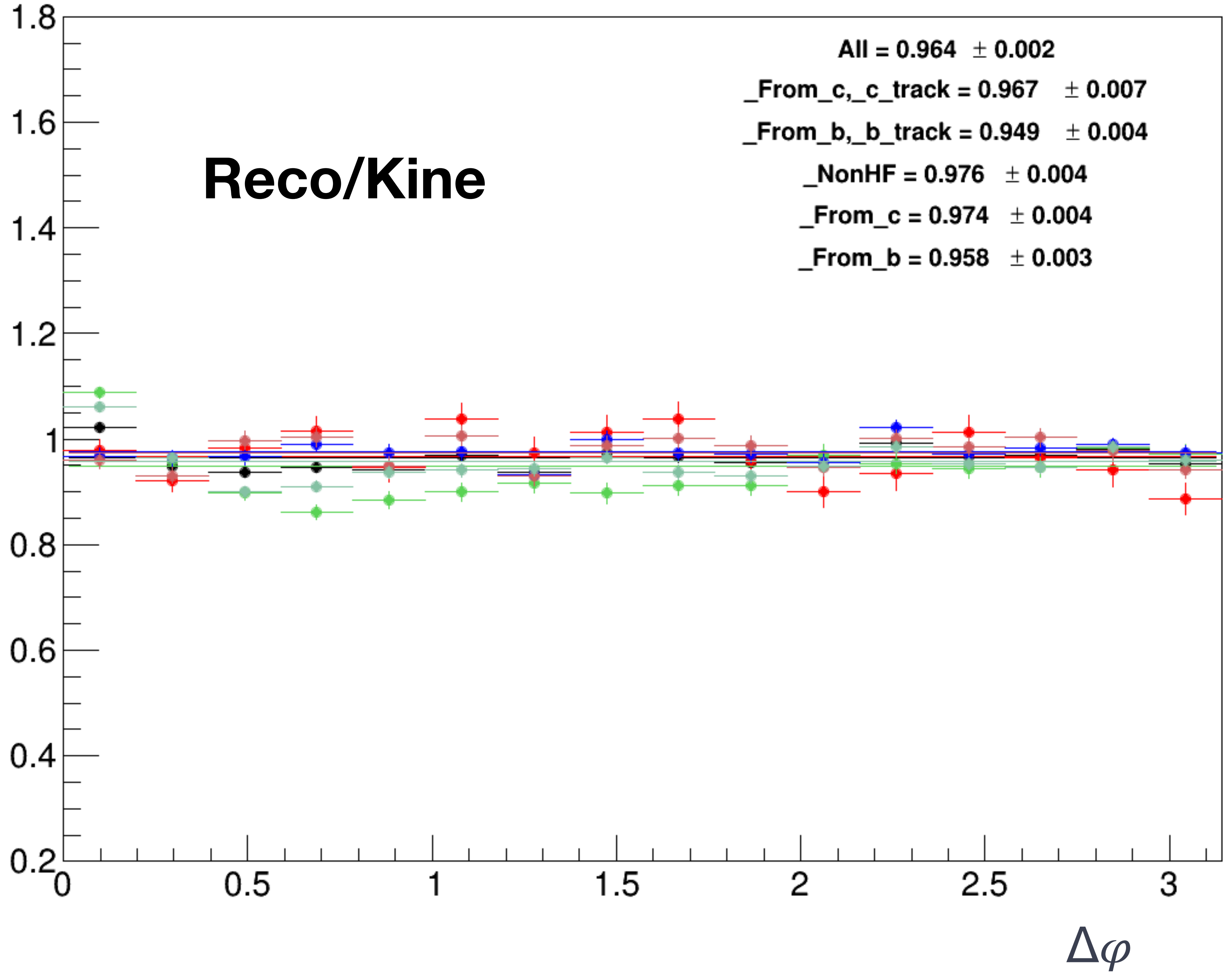


Corrected distribution :

$$\tilde{C}_{\mathrm{prompt}\ \Lambda_c}(\Delta\varphi) = \frac{1}{f_{\mathrm{prompt}}}\left(\tilde{C}_{\mathrm{inclusive}}(\Delta\varphi) - (1 - f_{\mathrm{prompt}})\tilde{C}^{\mathrm{MC\ templ}}_{\mathrm{feed-down}}(\Delta\varphi)\right)$$

Correction for b->D topological bias : MC Closure test

- **Kinematic level** : only acceptance cuts were applied on the D mesons and the associated particles
- **Reconstructed level** : the analysis was performed as if it were executed on data, applying the event selection, followed by all other corrections.



$b \rightarrow \Lambda_c$ bias induced by topological variables

– b-origin correlation pairs at the centre of the NS region

→ More prominent at low-$p_T(\Lambda_c)$;

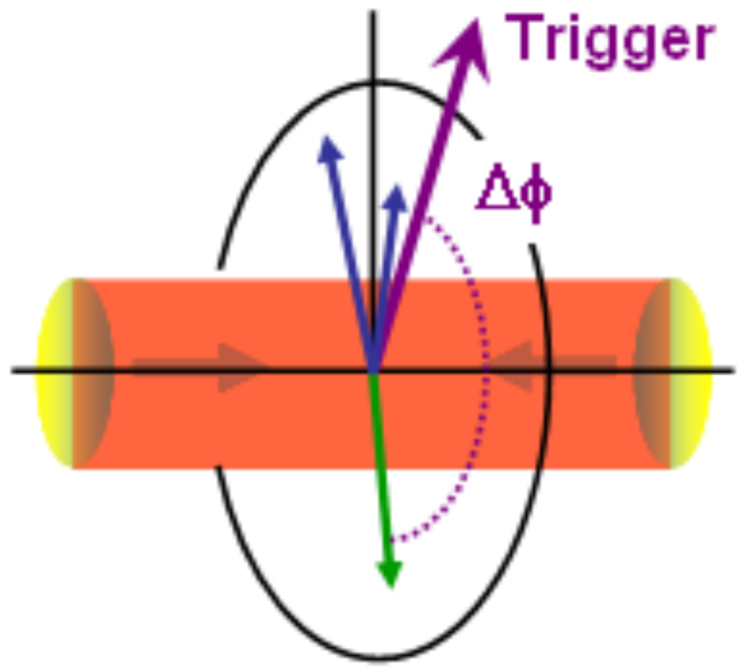
→ Larger effect with increasing p_T(assoc);

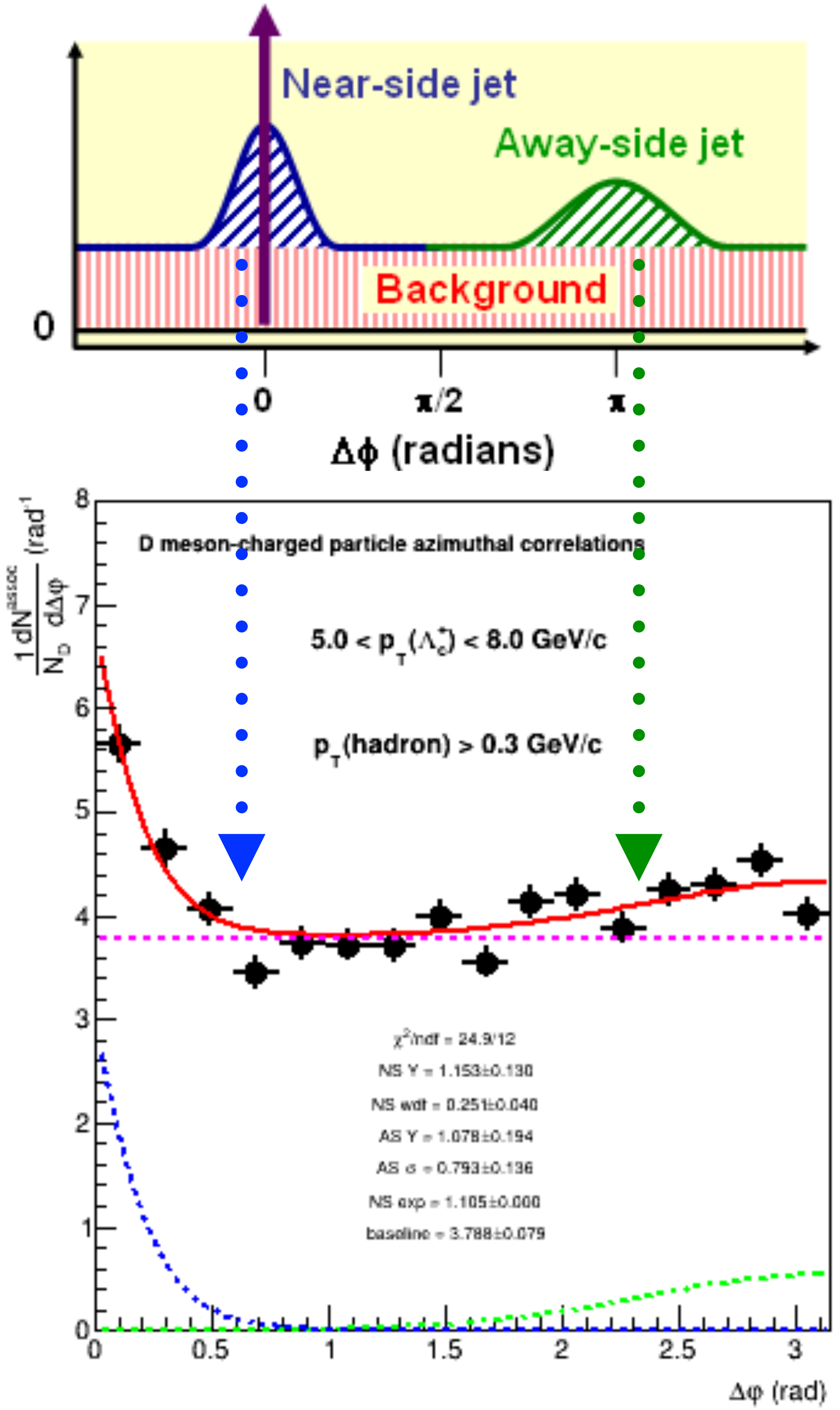
Correction formula:

$$C(\Delta\varphi)_{\text{corr}} = C(\Delta\varphi)_{\text{raw}} \cdot \left[\frac{\mathrm{c} \rightarrow \mathrm{D}|_{\text{amplit}}}{(\mathrm{B}+\mathrm{c}) \rightarrow \mathrm{D}|_{\text{amplit}}} \cdot f_{\text{prompt}} + \frac{\mathrm{B} \rightarrow \mathrm{D}|_{\text{amplit}}}{(\mathrm{B}+\mathrm{c}) \rightarrow \mathrm{D}|_{\text{amplit}}} \cdot (1 - f_{\text{prompt}}) \cdot \frac{1}{\mathbf{modul}} \right]$$

modul => The excess at the reconstructed level for the b-origin correlated pairs (dark green case), evaluated bin-by-bin for the NS peak region

Fitting : Calculation of Physical observables





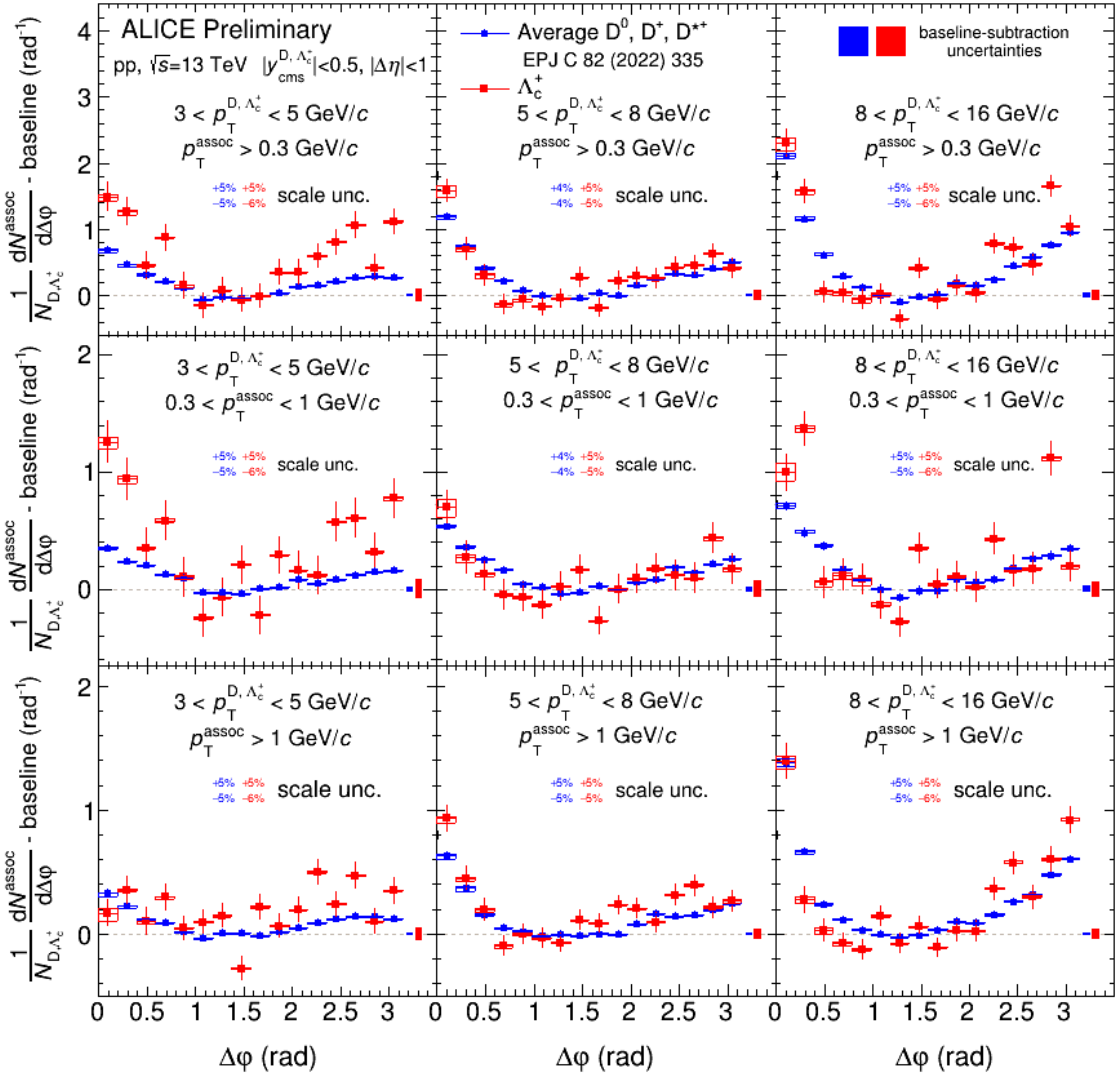
Fit function :

$$f(\Delta\varphi) = a + \frac{Y_{NS} \times \beta}{2\alpha\,\Gamma(1/\beta)} \times e^{-\left(\frac{\Delta\varphi}{\alpha}\right)^{\beta}} + \frac{Y_{AS}}{\sqrt{2\pi}\sigma_{AS}} \times e^{-\frac{(\Delta\varphi-\pi)^2}{2\sigma_{AS}^2}}$$

Characterization of the jet shape and its composition:

- Near Side (NS): fragmentation of the tagged charm quark;
- Away Side (AS): fragmentation of the other charm quark;
- Transverse Region: sensitivity to underlying event

Results : $\Delta\varphi$ distributions



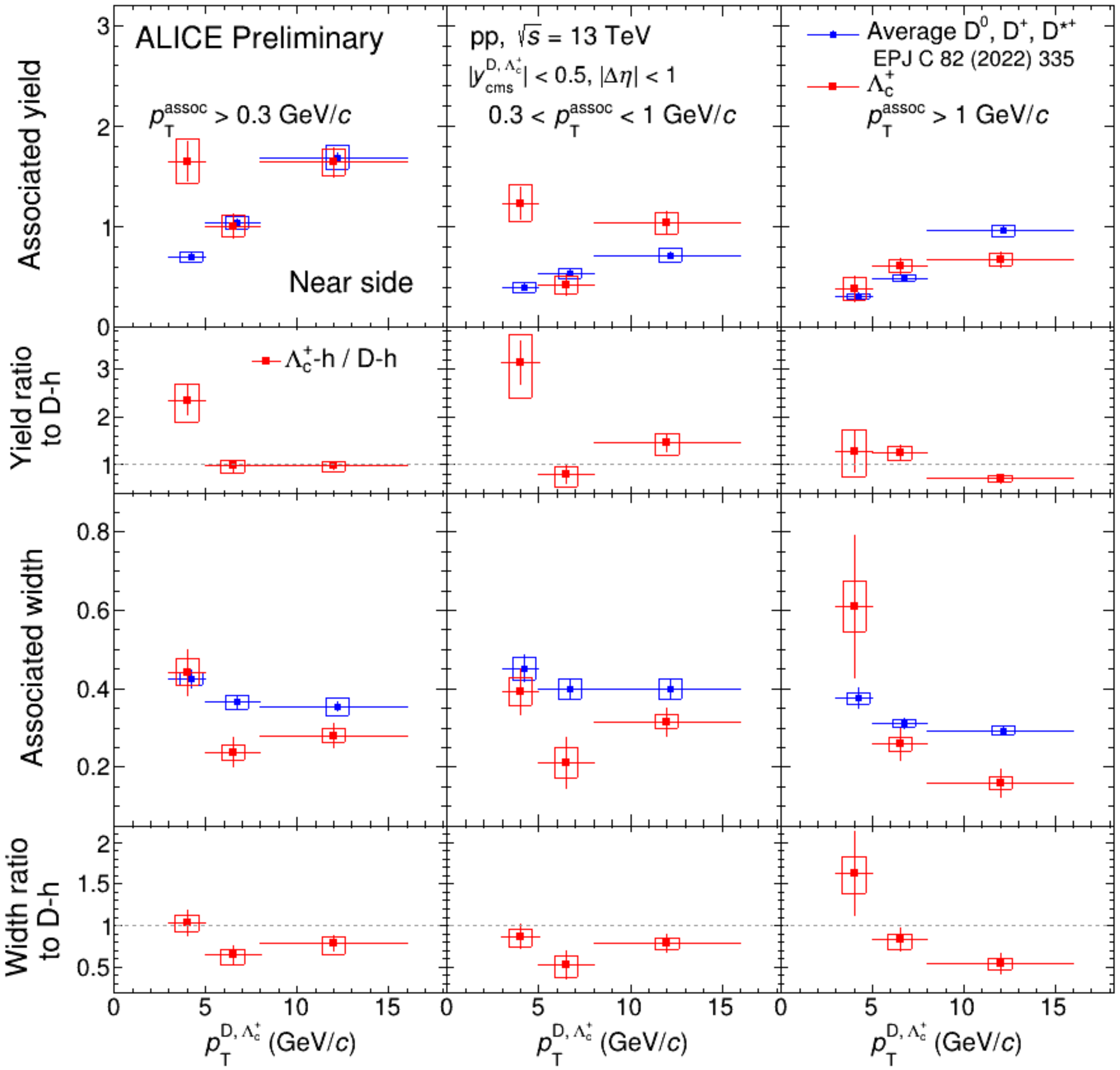
Compared with the results from D-h correlation analysis EPJC 82 (2022) 335

From the comparison:

→ discrepancy in the low-$p_T(\Lambda_c^+)$ region for low-momentum associated particles

→ good agreement between the $\Delta\varphi$ distribution in other kinematic ranges

Results : Near-side peak observables



Compared with the results from D-h correlation analysis EPJC 82 (2022) 335

From D-h correlations, increasing p_T^{HF}:

→ More energetic parton
→ increasing yields

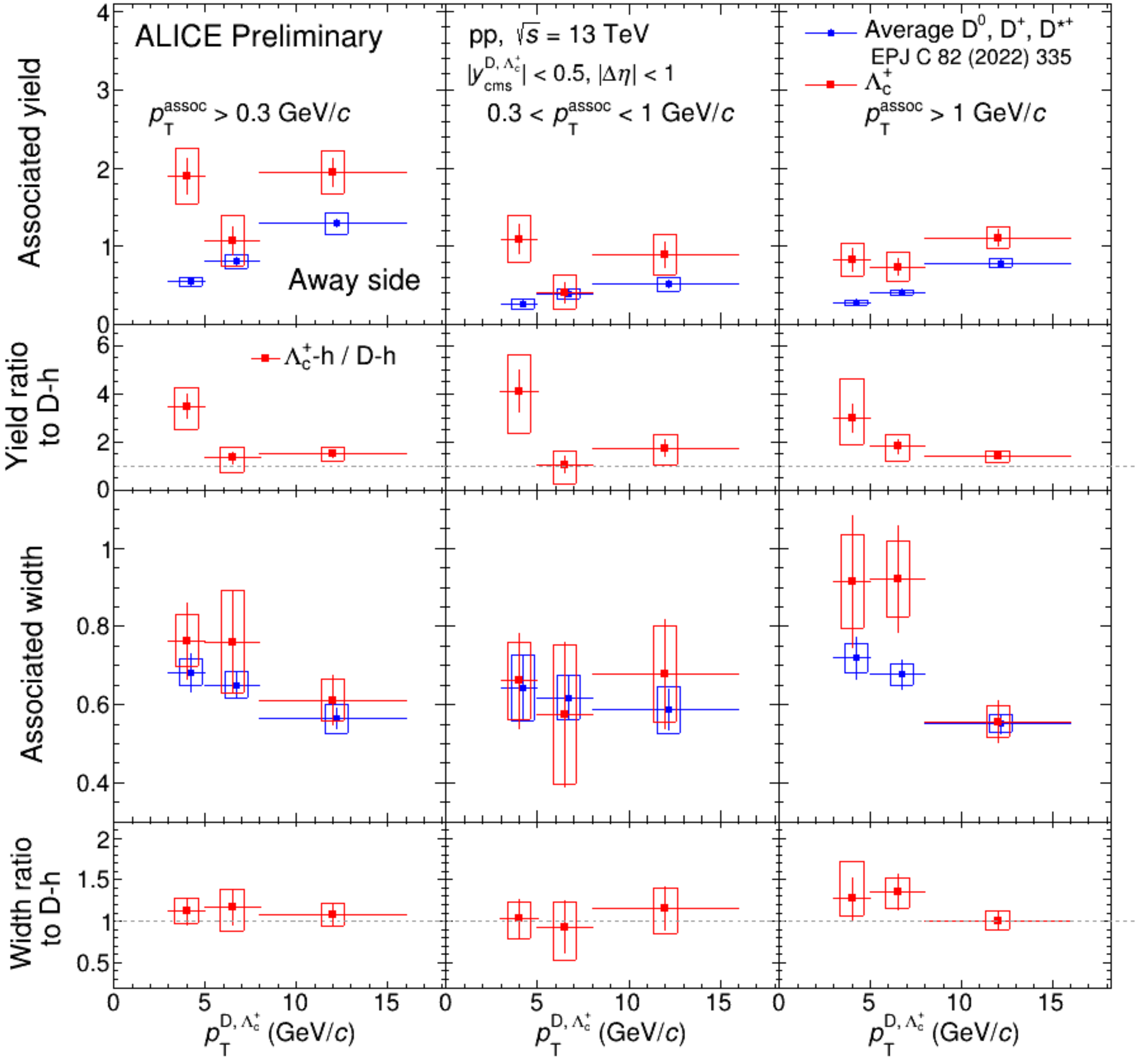
→ Larger heavy-quark boost
→ more collimated shower
→ sharpening of the peak

Higher NS yields in Λ_c^+-h than D-h at low-p_T:

→ different energy of the charm quark as a consequence of a softer Λ_c^+ fragmentation

→ decay of higher mass charm states

Results : Away-side peak observables

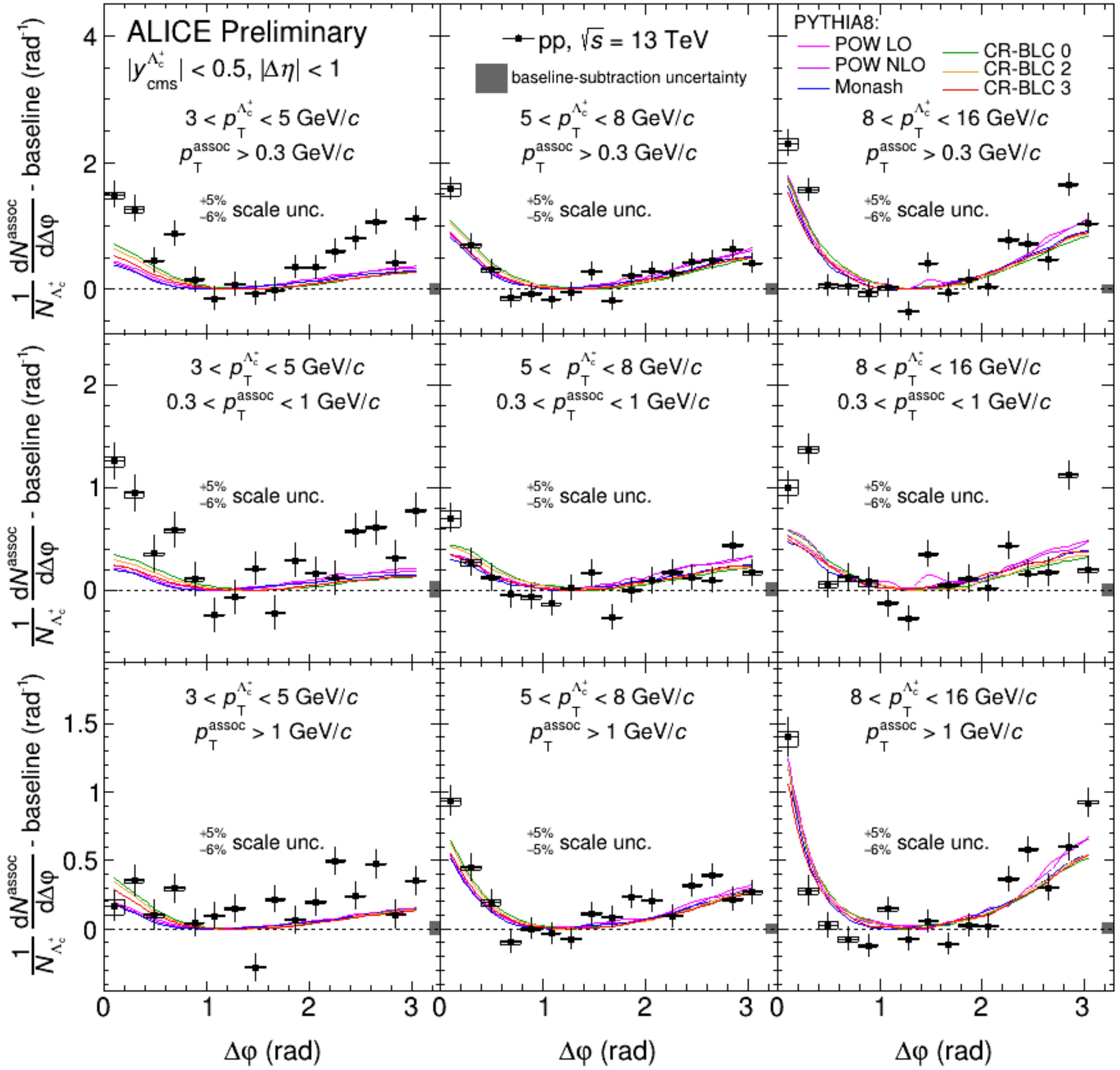


Compared with the results from D-h correlation analysis EPJC 82 (2022) 335

Enhancement also in the AS region:

→ possibly justified by a softer Λ_c^+ fragmentation

Results : comparison with MC generators

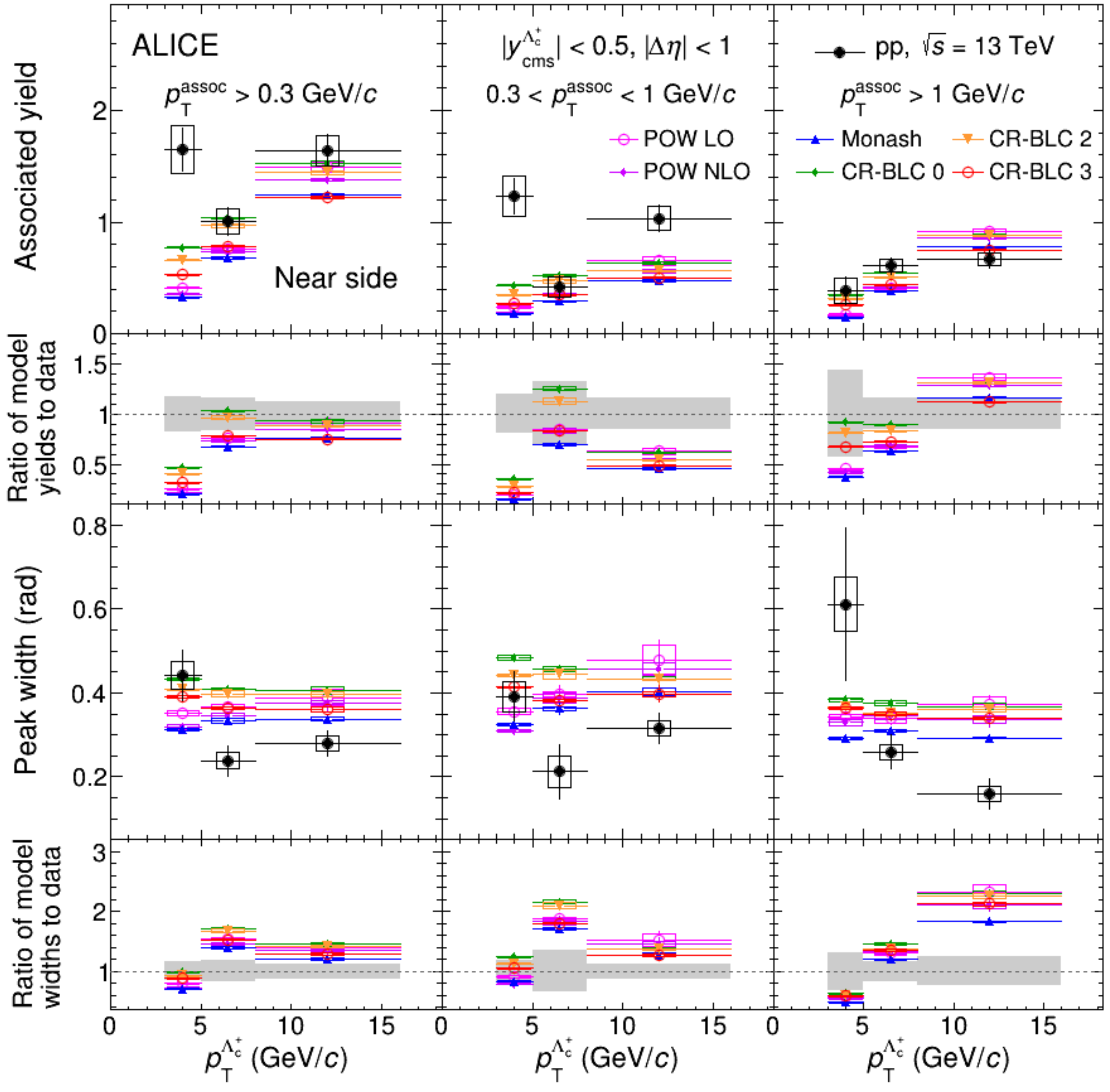


→ discrepancy in p_T (Λ_c) 3-5 GeV/c

→ better description of the data in p_T (Λ_c) 5-8 and 8-16 GeV/c

Results : comparison with MC generators

Near Side

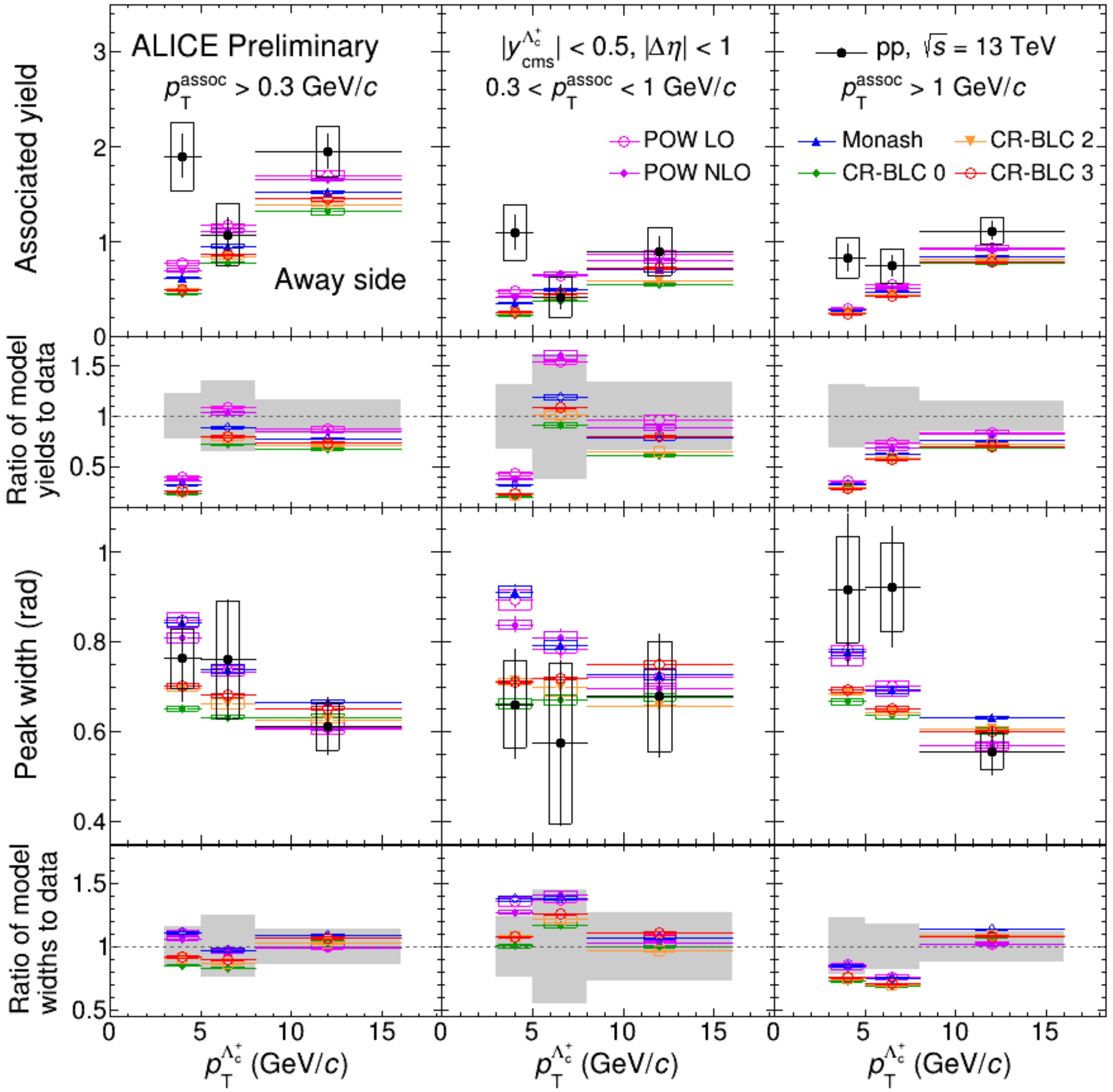


- Increasing Yields and flat widths vs p_T (Λ_c)

    → higher particle multiplicity in the jet cone including color reconnection

- Large discrepancies in low-p_T(Λ_c) yields

Away Side



- Increasing Yields and decreasing widths vs p_T (Λ_c)
- Large discrepancies in low-p_T(Λ_c) yields

    → both p_T^{assoc} sub-ranges are underestimated by models

Summary

Paper proposal placed : https://indico.cern.ch/event/1326692/

Analysis note : https://alice-notes.web.cern.ch/node/1349

What is missing:

Monte Carlo models:

- JetScape + modified fragmentation: some iterations needed with the authors
- Catania: will be included if their timeline is compatible, no news for the moment

Ongoing:

PYTHIA8: adding decay of higher charm states to quantify possible impact on correlations

Thank you

Backup

Data and MC samples

Data periods :

2016: LHC16d,e,g,h,j,k,l,o,p
2017: LHC17e,f,g,h,i,j,k,l,m,o,r
2018: LHC18b,d,e,f,g,h,i,k,l,m,n,o,p

Total number of selected events (from AliNormalizationCounter): ~ 1.70B

MonteCarlo productions:

HF enriched:
ML: LHC20l1a, LHC20l1b, LHC20l1c; Eff.Corrections: LHC20f4a, LHC20f4b, LHC20f4c;

General-purpose:
2016: LHC17f6, LHC17f9, LHC17d17, LHC17f5, LHC17e5, LHC18f1, LHC18d8, LHC18d16, LHC18d18;
2017: LHC17h1, LHC17h3, LHC18c12, LHC17k4, LHC17h11, LHC18c13, LHC18a8, LHC17l5, LHC18a9, LHC18a1;
2018: LHC18g4, LHC18g5, LHC18g6, LHC18h2, LHC18h4, LHC18j1, LHC18j4, LHC18k1, LHC18k2, LHC18k3.

Reconstruction of Λ_c baryons

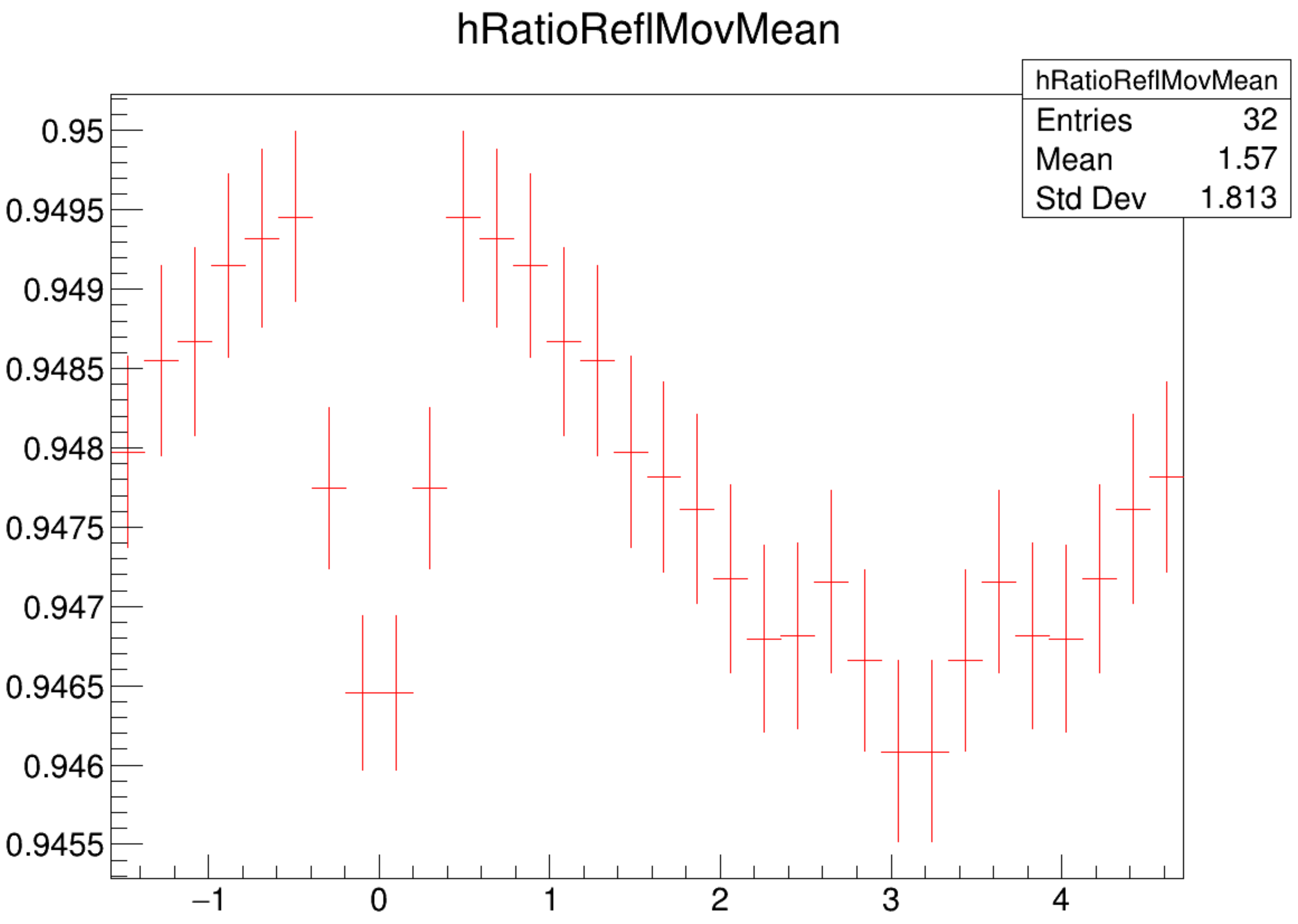
Model training : a supervised machine learning, as XGBOOST is, uses algorithms to train a model to find patterns in a dataset with labels and features and then uses the trained model to predict the labels by evaluating the new dataset feature two labelled dataset are built, respectively for the training of the models and for its performance evaluation by comparison with the truth labels. Two class of candidates are considered: prompt Λc, extracted from an HF-enriched MC production and combinatorial background candidates, selected from the invariant mass distribution sidebands of real data after having applied of the aforementioned preselections

Choice of training variable : The topological variables used to train the models are those usually used for the reconstruction of charmed hadrons without ML approach

After the training, the model is applied to both training and test datasets in order to extract the ML score distributions and verify the level of agreement between the two samples. In binary classification cases, one score is provided by the model, representing the probability for a candidate to be signal or background. The closer the score is to 1, the higher the probability for the candidate to be a true Λc baryon rather than a combinatorial triplet of uncorrelated charged tracks.

ML Working point : The choice of the ML output scores to be applied for candidate selection on data was performed by estimating the expected significance (S/√S + B, signal-over-background ratio (S/B)

Removal of secondary particle contamination



ratio of Δϕ distribution of primary track accepted over all tracks distribution is moving by three points and this distribution is multiplied to the data Δϕ correlation distributions.

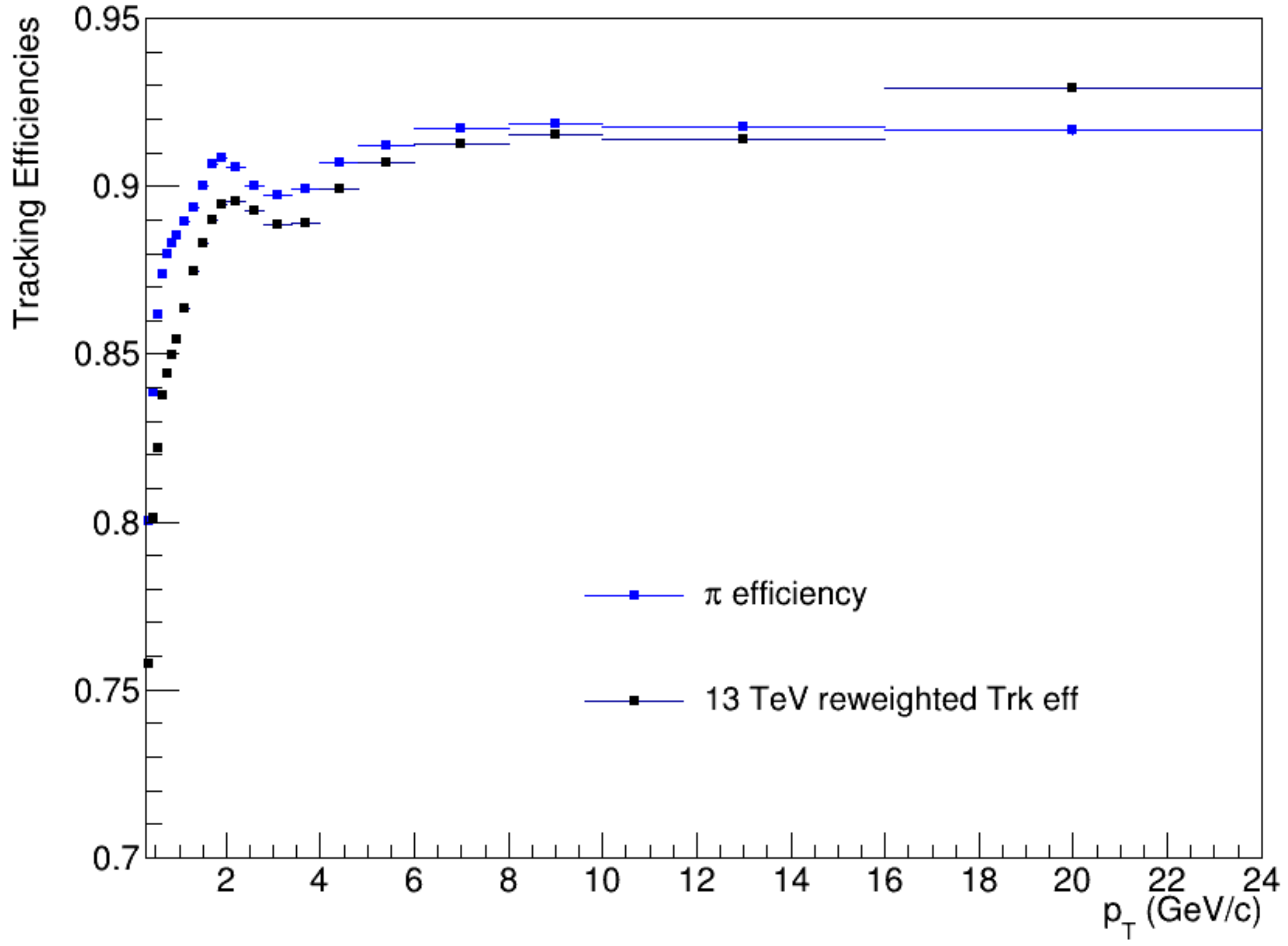
Soft Pion removal

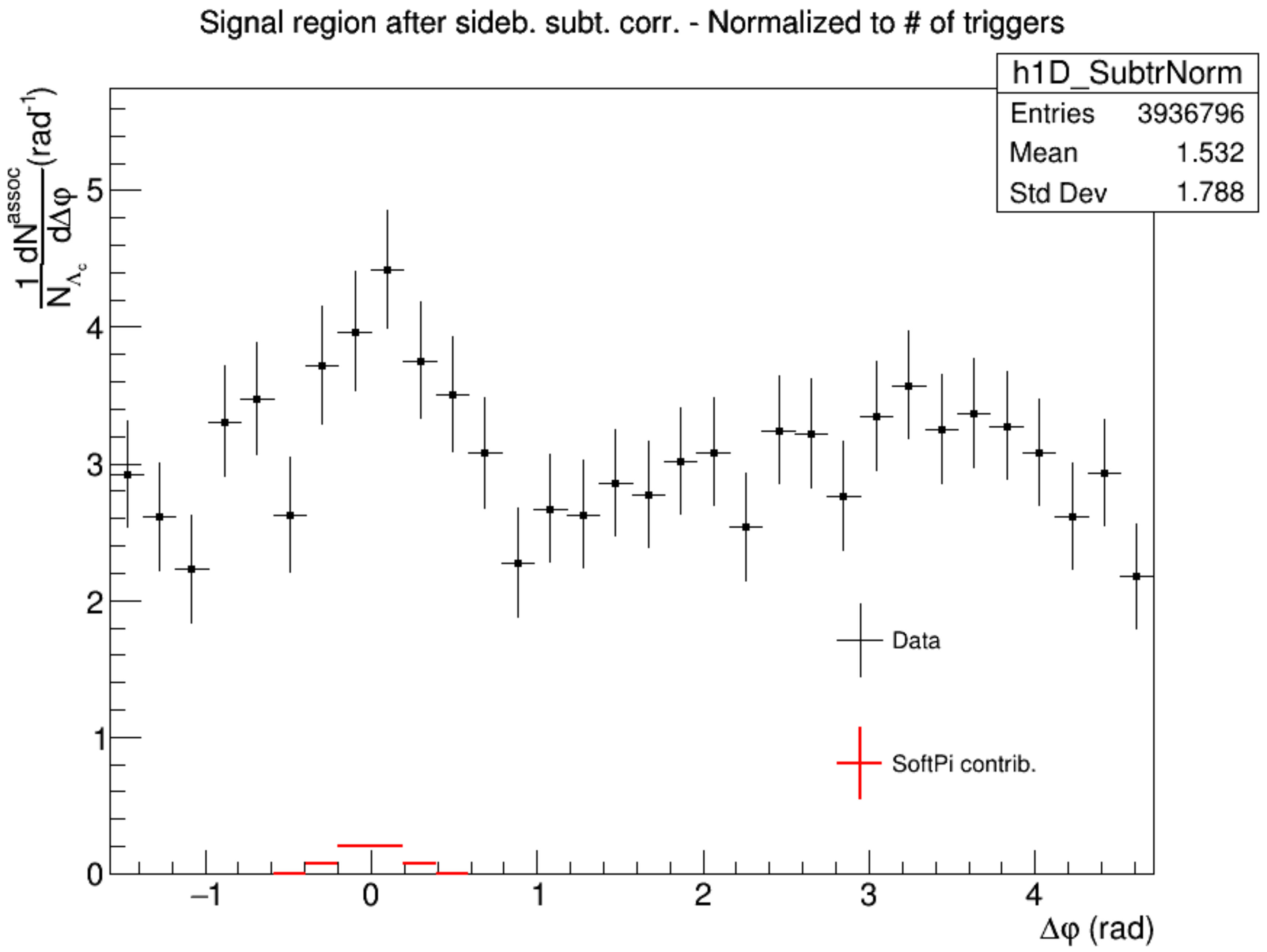
→ Correction for π coming from $\Sigma_c^{0,++}(2455) \rightarrow \Lambda_c^+ \pi^{\mp}$

Monte Carlo based approach:

- → $\Sigma_c^{0,++} \rightarrow \Lambda_c^+ \pi^{\mp}$ decay kinematics simulation
- → Extraction and correction of $\Delta\varphi_{MC}(\Lambda_c^+ - \pi^{\mp})$
    - → p_T ($\Lambda_c^+ \leftarrow \Sigma_c^{0,++}$) rescaled to measurements;
    - → Eff_{π}/Eff_{track} reweighting
- → Product of the per-trigger-normalized $\Delta\varphi_{MC}^{\Lambda c - \pi}$ with

The corrected $\Delta\varphi$ ($\Lambda_c^+ - \pi^{\mp}$) is then subtracted from the fully corrected azimuthal correlation distribution





Systematics on $\Delta\varphi$ shape

Those above are Δϕ-correlated (act as scale factor on the corr. distribution).

In addition, Δϕ-uncorrelated uncertainties from feed-down subtraction, $b \rightarrow \Lambda_c^+$ bias and soft-π (Δϕ~0) (max effect, under the peaks, 3-4%)

Sources	pp (Λ_c)								
	p_T(assoc)>0,3			0,3< p_T(assoc)<1			p_T(assoc)>1		
p_T ranges	3-5	5-8	8-16	3-5	5-8	8-16	3-5	5-8	8-16
Λ_c-cut stability	2	2	2	2	2	2	2	2	2
Bkg corr. shape	1	0.5	1	1	0.5	1	2	1	1
Yield extraction	2	1	1	2	1	1	2	2	1
Track Efficiency	4	4	4	4	4	4	4	4	4
Purity	2	1	2	2	2	2	1	2	2