Isolated photons in ATLAS: performance of calorimetric isolation with 2015 data

Dafne Arias
M2 NPAC - Université Paris Diderot

IPN, June 2016

Advisor: **José Ocariz** (LPNHE)

Outline

- **Introduction**

- **PhotonID & Photon Isolation**

- **Efficiency & Purity**

- **Data vs MC comparisons of calorimetric isolation**

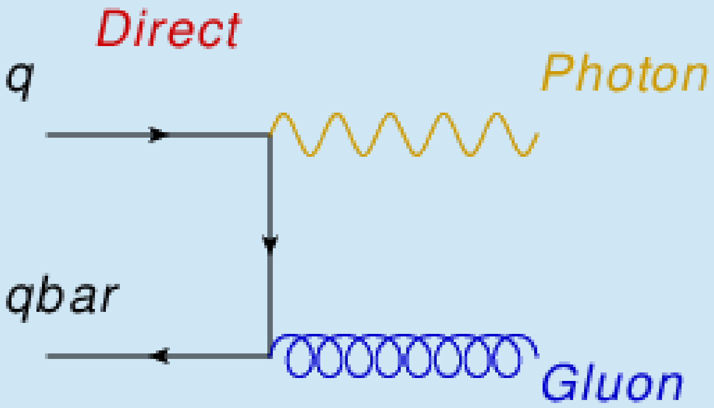
- **Conclusion**

Introduction

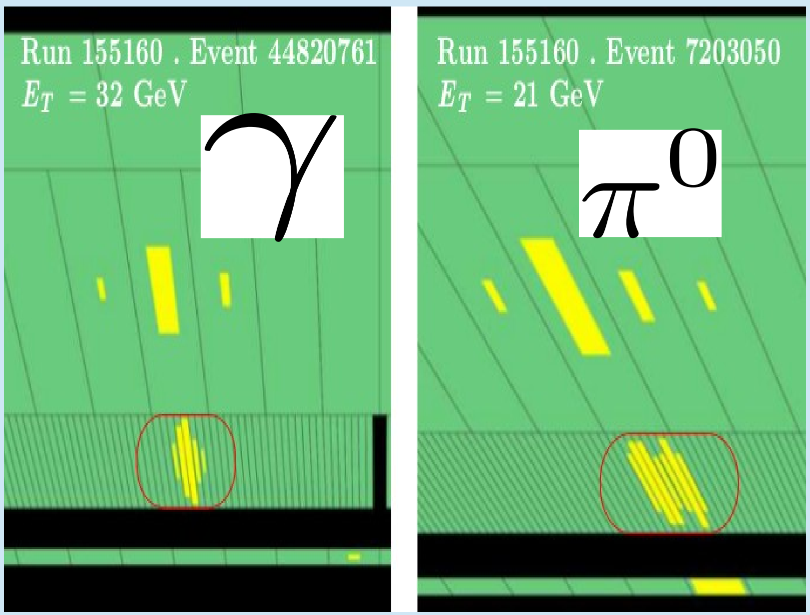
Studies of photon prodution at the LHC:

- Test of perturbative QCD.
- Higgs physics (H->gg).
- Backgrounds for (H->gg).
- Performance of photon reconstruction.

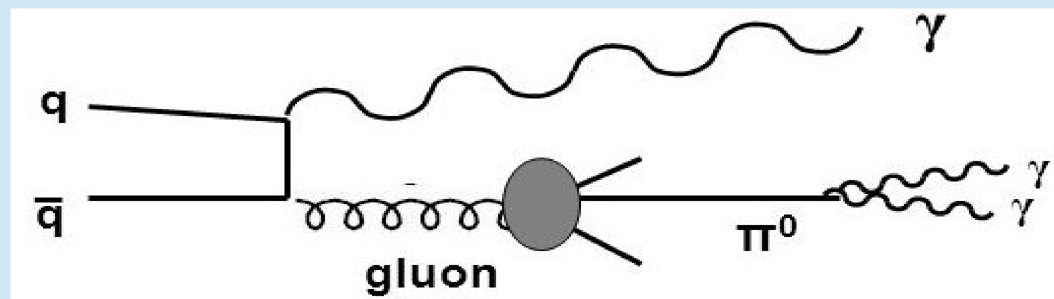
- Analysis of the data from the Run 2 recorded in 2015 at 13 TeV with ROOT.
  - Efficiency&Purity.
  - Evaluation of the agreement between the ATLAS data and the simulation (Sherpa and Pythia).



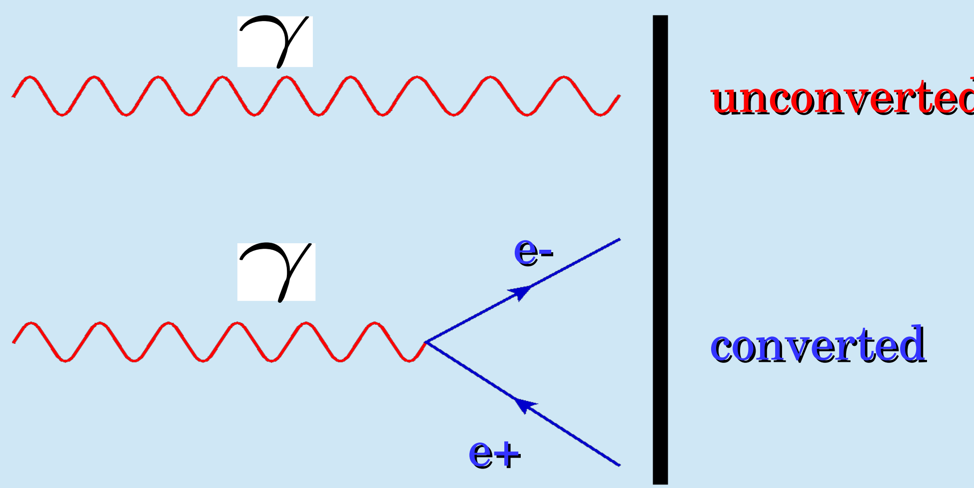
True and fake photons



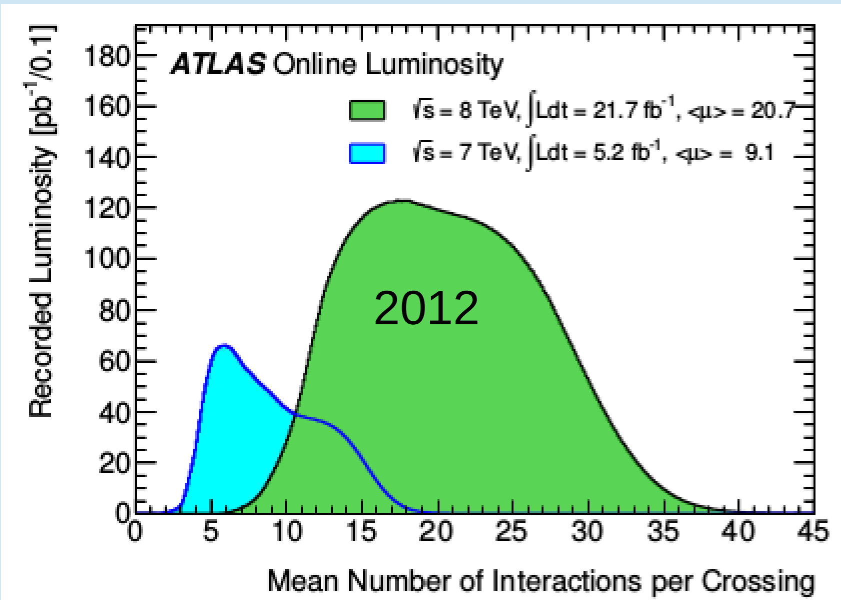
- **Direct photons** are produced in the hard processes.
- **Background to photons (fakes)** are jets with neutral hadrons.

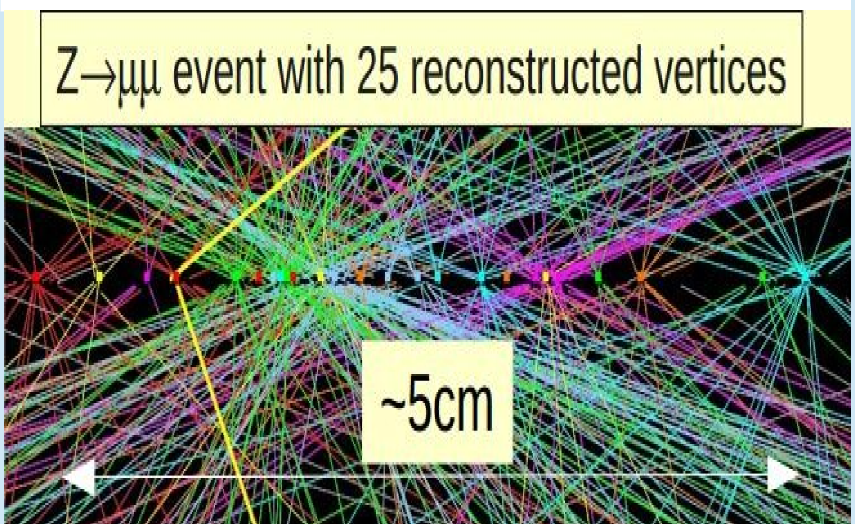


If there is no track associated the photon is classified as a **unconverted photon**. In the other case, is classified as a **converted photon**.



Pile-up





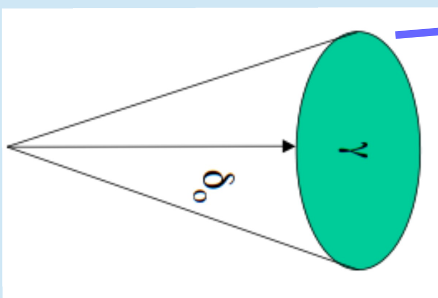
Number of proton collisions per bunch crossing.

- Data and Monte Carlo's pile-up are slightly different.

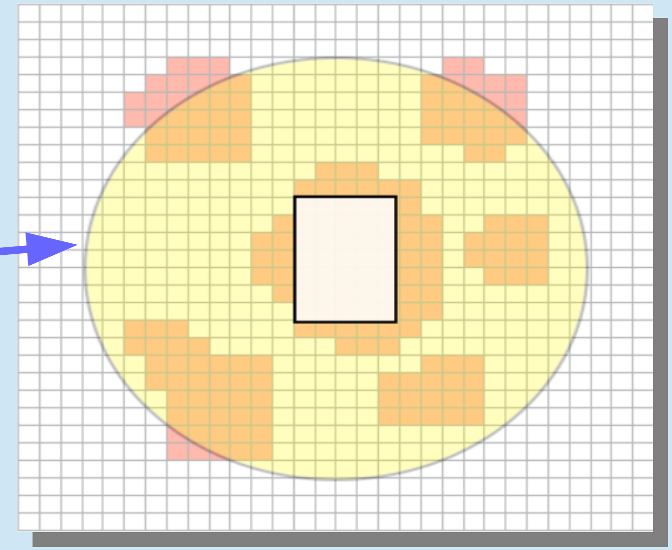
PU reweighting

PhotonID&Photon Isolation

- The photon ID algorithm in ATLAS relies on rectangular cuts using calorimetric variables.



$$E_T^{iso} \equiv \sum_{i}^{cone} E_T^i$$



To disentangle prompt photons from fakes, one applies the **isolation technique**:

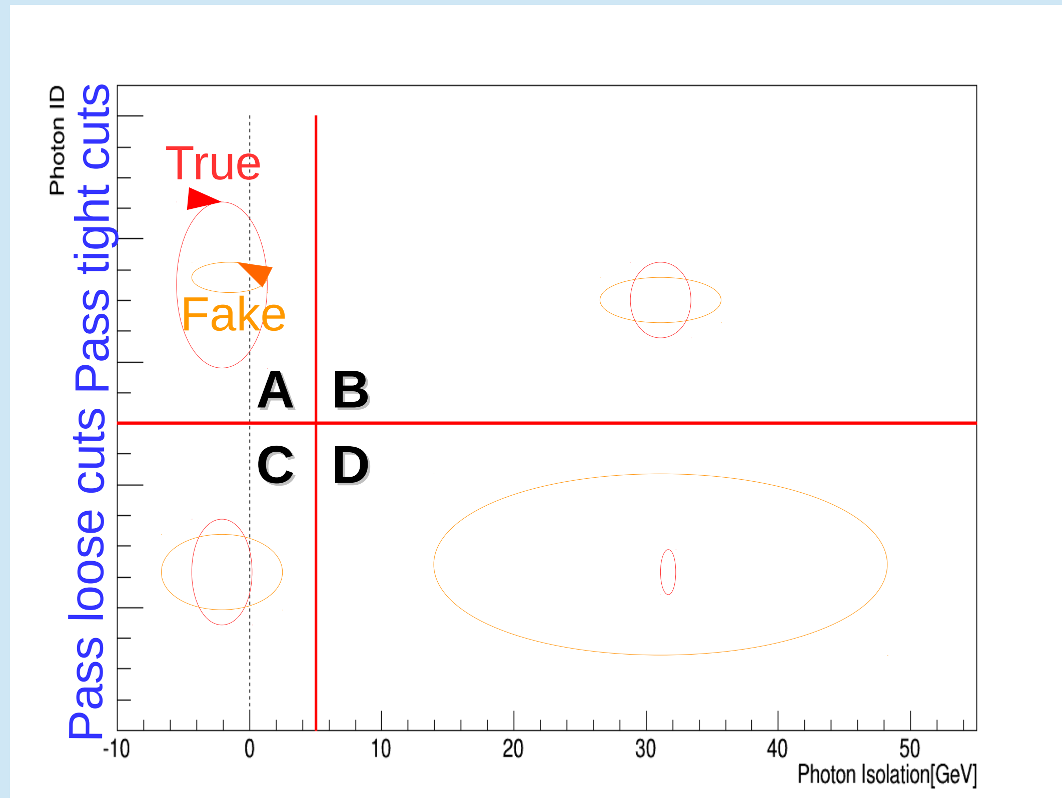
- The isolation is the sum of energies in a cone around the photon candidate.
- An isolated photon has smaller isolation that fakes.
- **Photon isolation is essential to separe real photons from fakes.**

In ATLAS, it is used the isolation variable topoEtcone40, where only cells belonging to clusters are considered.

$$\Delta R < 0.4$$

PhotonID&Photon Isolation

- Loose criterion favors large **efficiency**, while the tight criteria favours larger **purity**.
- **A:** tight photonID&tight isolation.
- B: tight photon&loose isolation.
- C: loose ID & tight isolation.
- D: loose ID&loose isolation.



Assumptions:

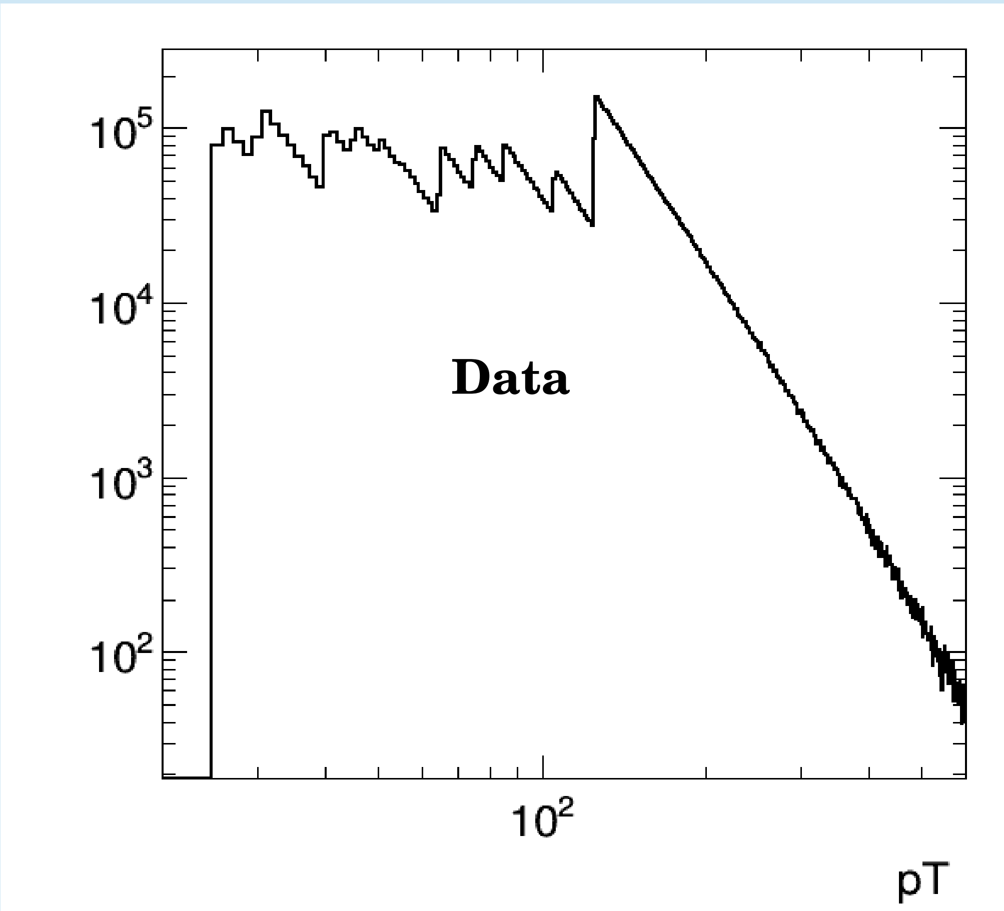
- No correlation for iso/ID for fakes.

$_{C} = {}_{B} = 0$

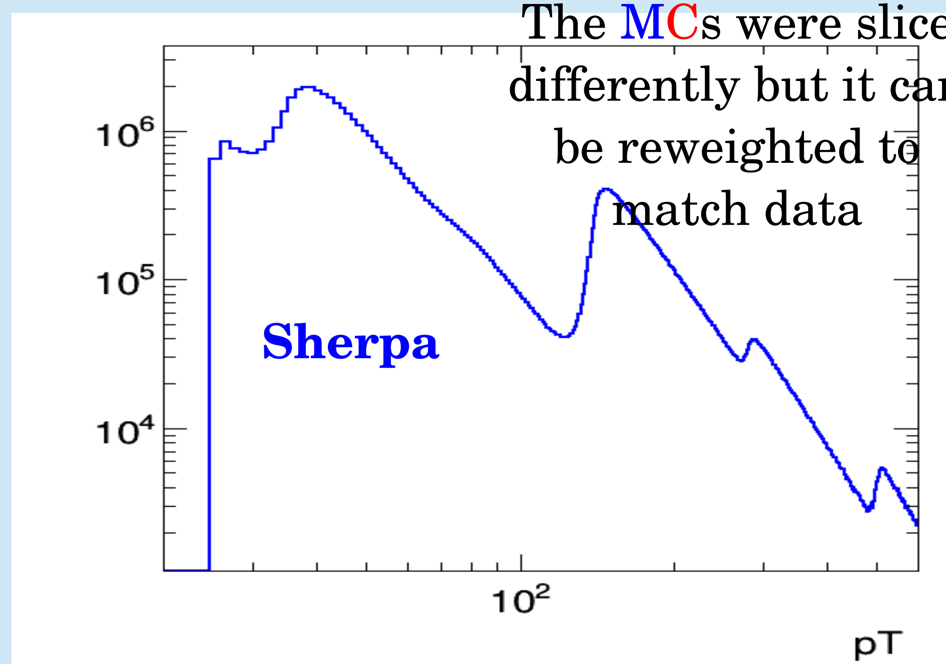
Hypothesis

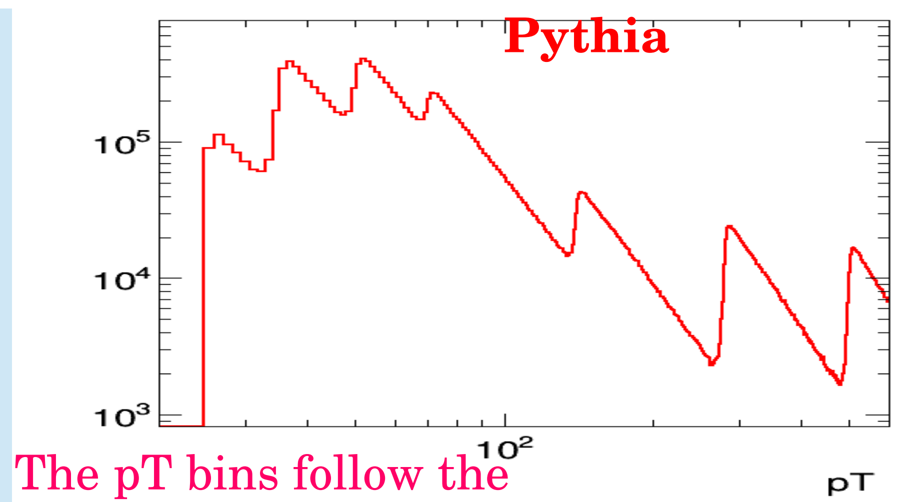
$$\frac{D}{C} \qquad \frac{B}{A}$$

Inclusive photon spectrum



The MCs were slice differently but it can be reweighted to match data





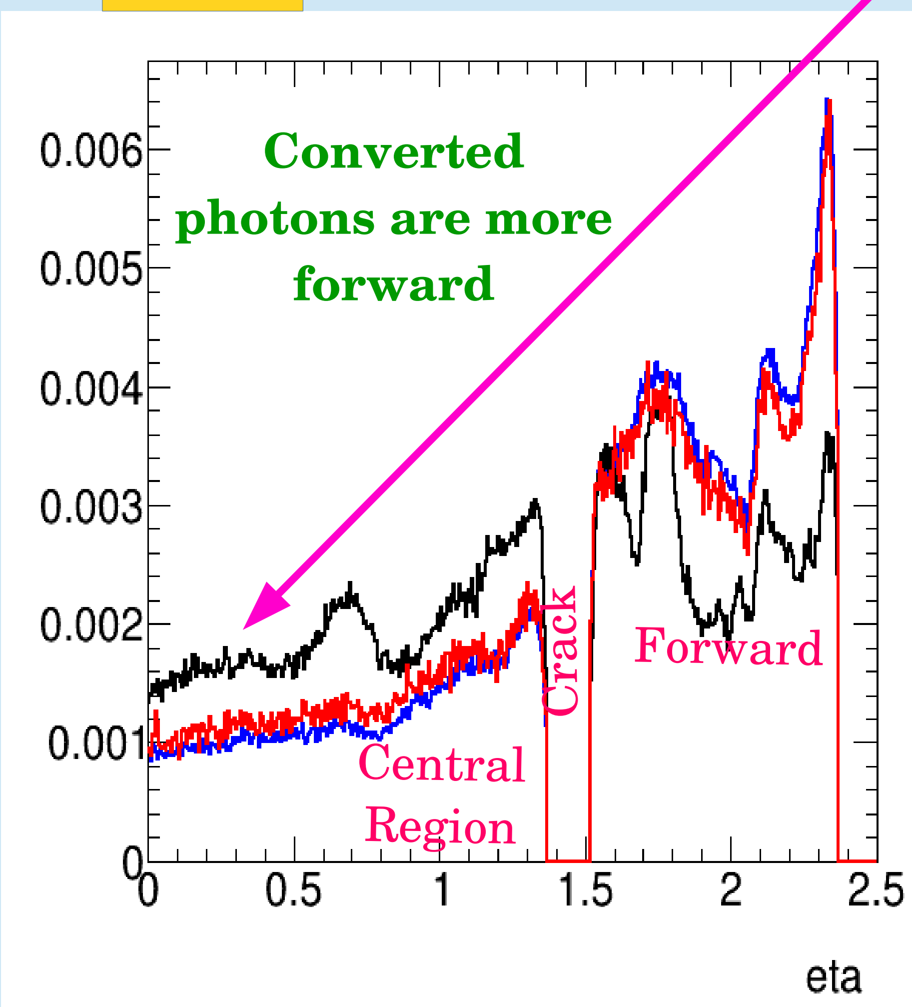
The pT bins follow the 2015 photon triggers

Ptbins: 25 , 30 , 40 , 45 , 50 , 55 , 65 , 75 , 85 , 105 , 125 , 145 , 205 , 300 , 500 , 999.

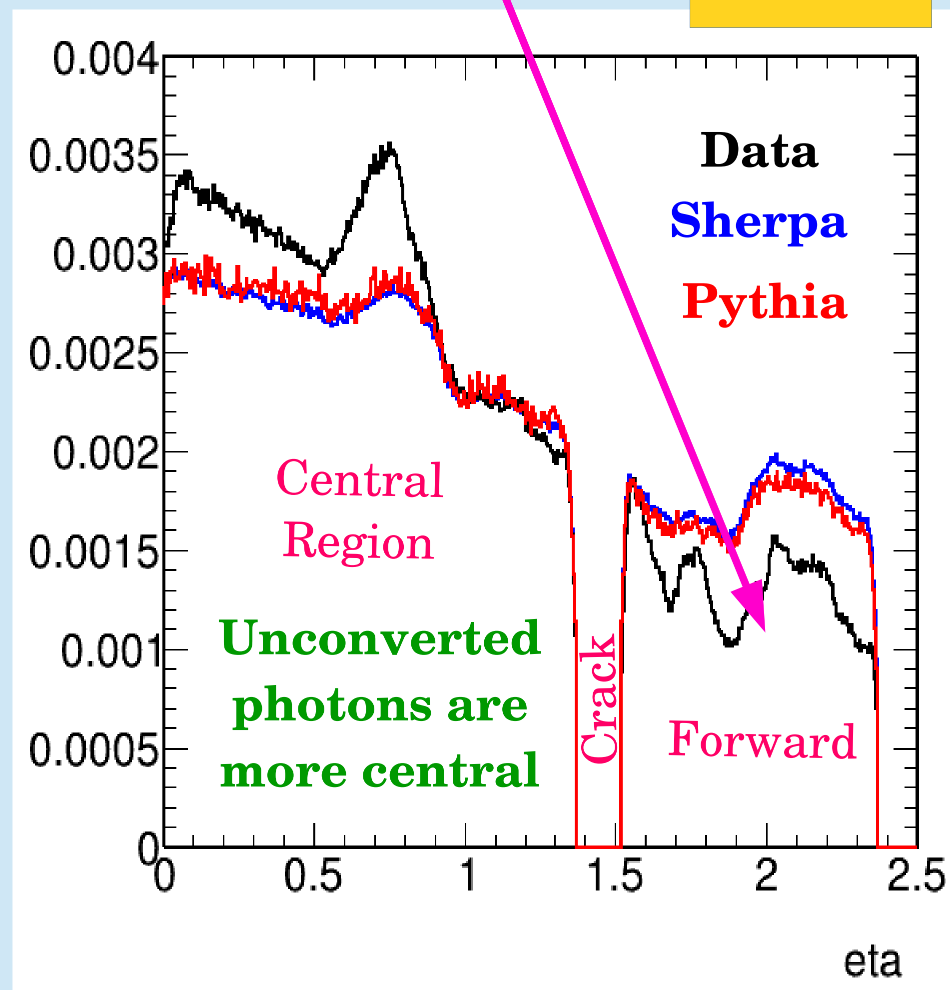
Pseudorapidity

|eTabins|: 0 , 0.6 , 1.37 , 1.52 , 1.81 , 2.37.

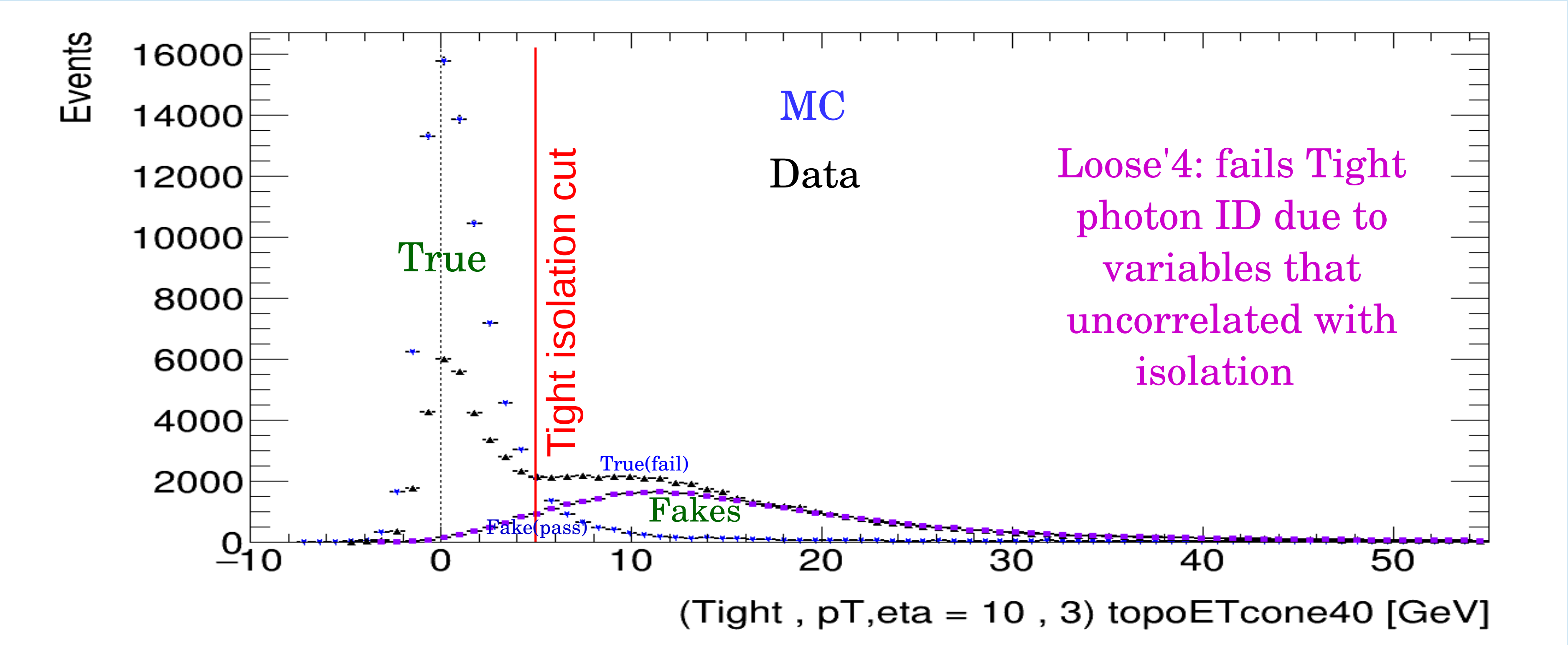
Conv



Unconv



Background Subtraction



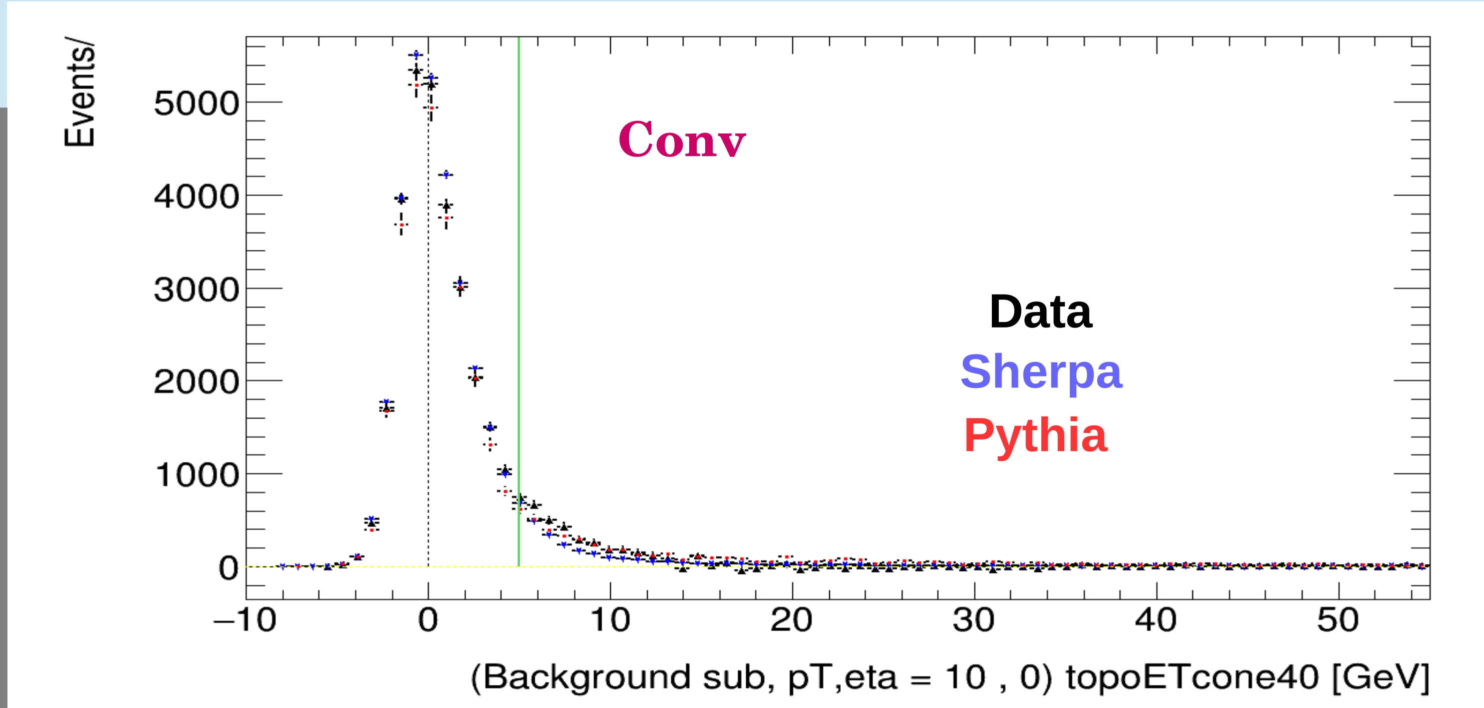
Hypothesis

$$tight = f_s signal + (1 - f_s) bg$$

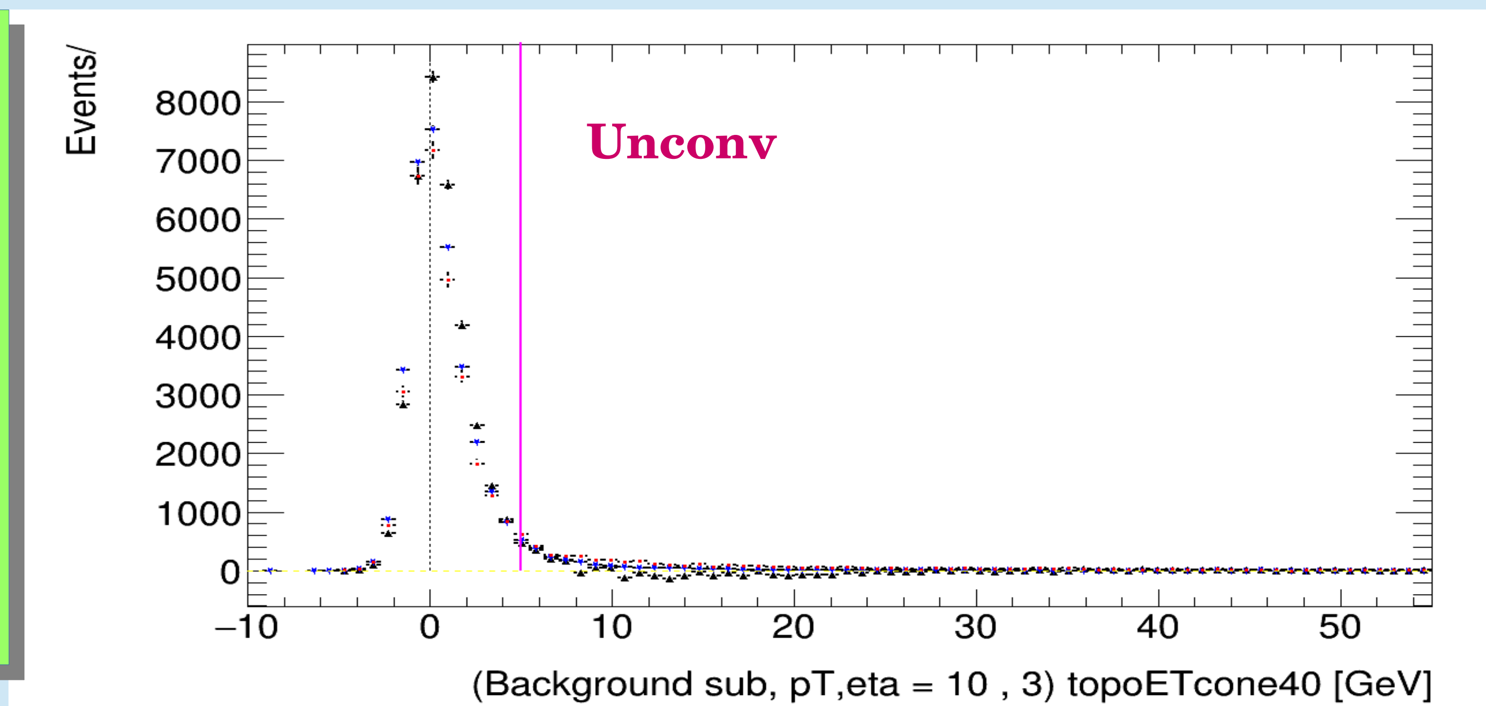
MC

Loose'4

Data tight&data loose'4 distributions are normalized using the hight isolation region(subtracting loose'4 from tight)



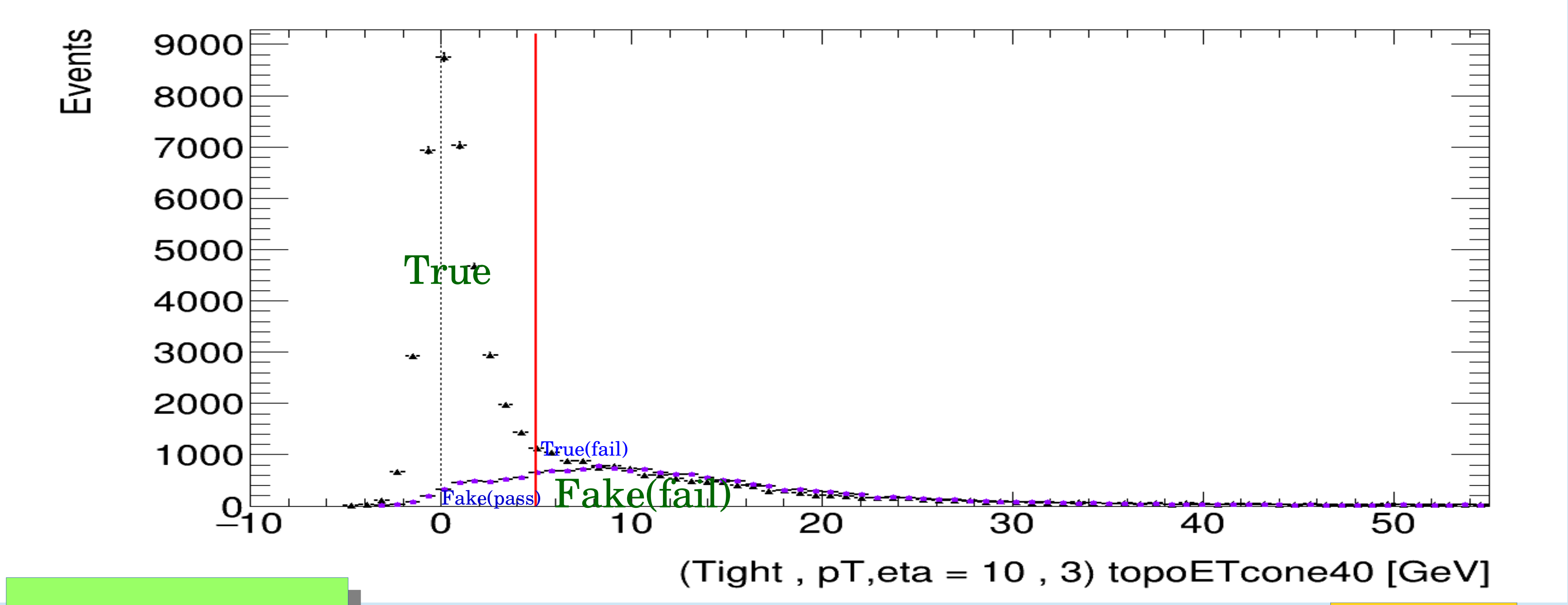
- For unconv the bg sub method works fine except for a shift.
- For conv the bg sub method indicates data-MC disagreement.



Outline

- **Introduction**
- **PhotonID & Photon Isolation**
- **Efficiency & Purity**
- **Data vs MC comparisons of calorimetric isolation**
- **Conclusion**

Efficiency & Purity



Efficiency

Purity

$$\epsilon^{MC}_{iso,tight} = \frac{N(true, iso < 5)}{N(true, all)}$$

$$P^{data}_{iso,tight} = \frac{N(sig, iso < 5)}{N(data, iso < 5)}$$

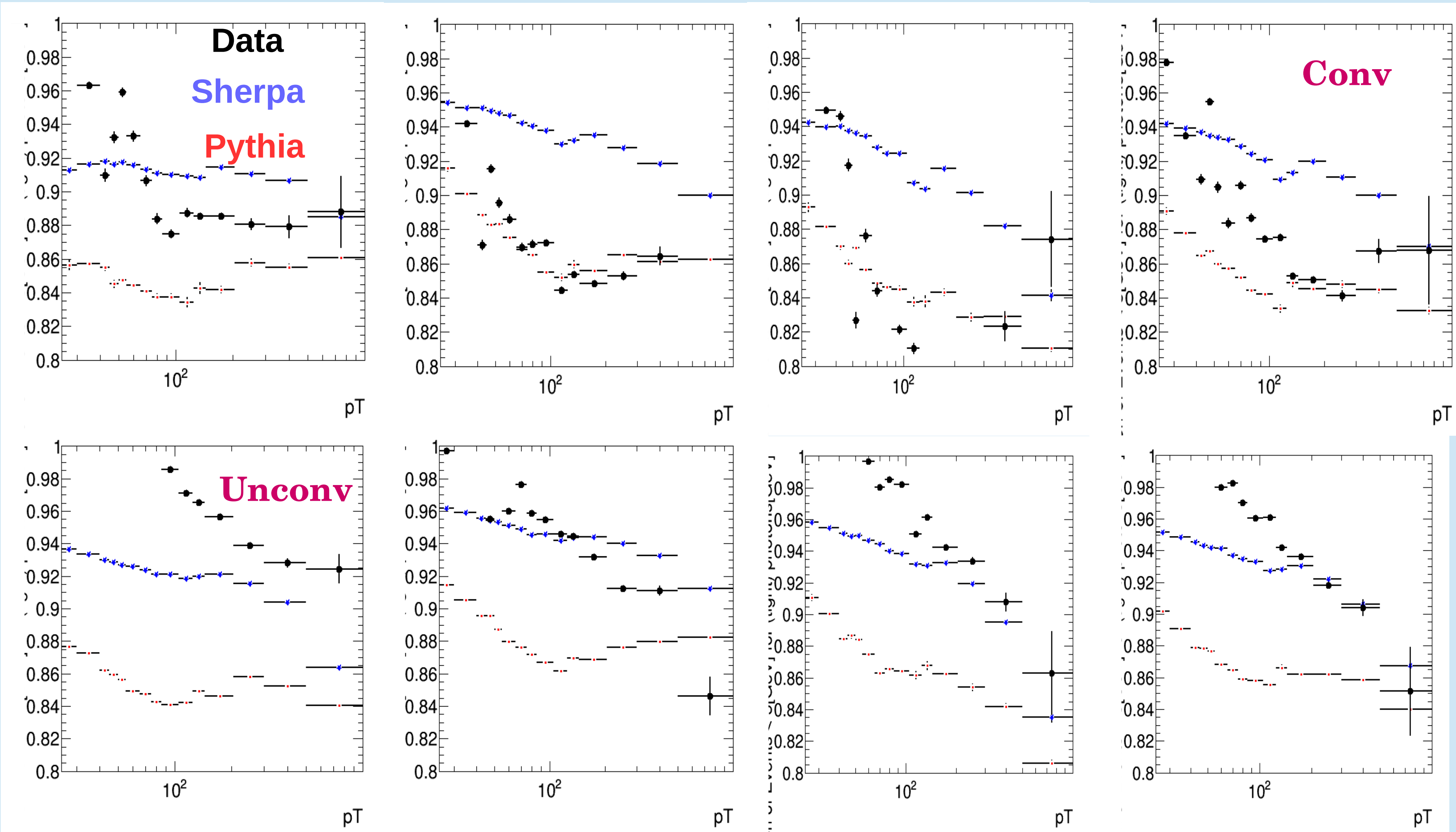
$$\epsilon^{data}_{iso,tight} = \frac{N(sig, iso < 5)}{N(sig, alliso)}$$

$$N(sig) = N(data) - N(bkg)$$

Efficiency

- Eff(SH) > Eff(py)
- Data and MC agree within +/- 5%
- Data efficiencies can be biased due to the bg sub (negative sub)

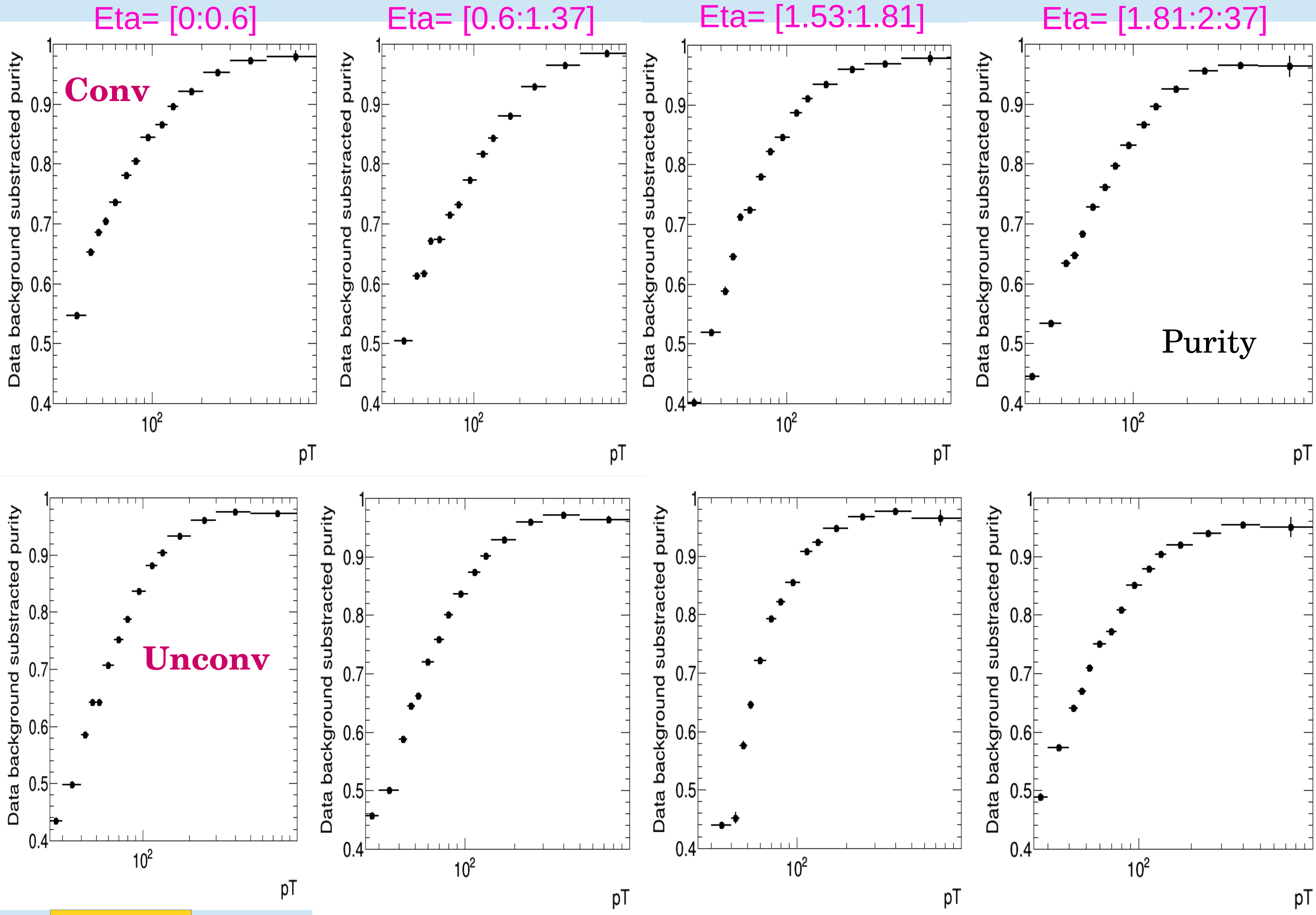
Eta= [1.81:2:37]



Eta= [0:0.6]

Eta= [0.6:1.37]

Eta= [1.53:1.81]



Purity

Purities increase significantly with pT

Outline

- **Introduction**
- **PhotonID & Photon Isolation**
- **Efficiency & Purity**
- **Data vs MC comparisons of calorimetric isolation**
- **Conclusion**

Pseudo-χ^2 Method

- Used for maximize the agreement between data and Monte Carlo.

- **Shift** isoMC by various steps of alpha.

$$\chi^2(\alpha) = \frac{(\alpha - \alpha_{min})^2}{\sigma_\alpha^2} + \chi^2_{min}$$

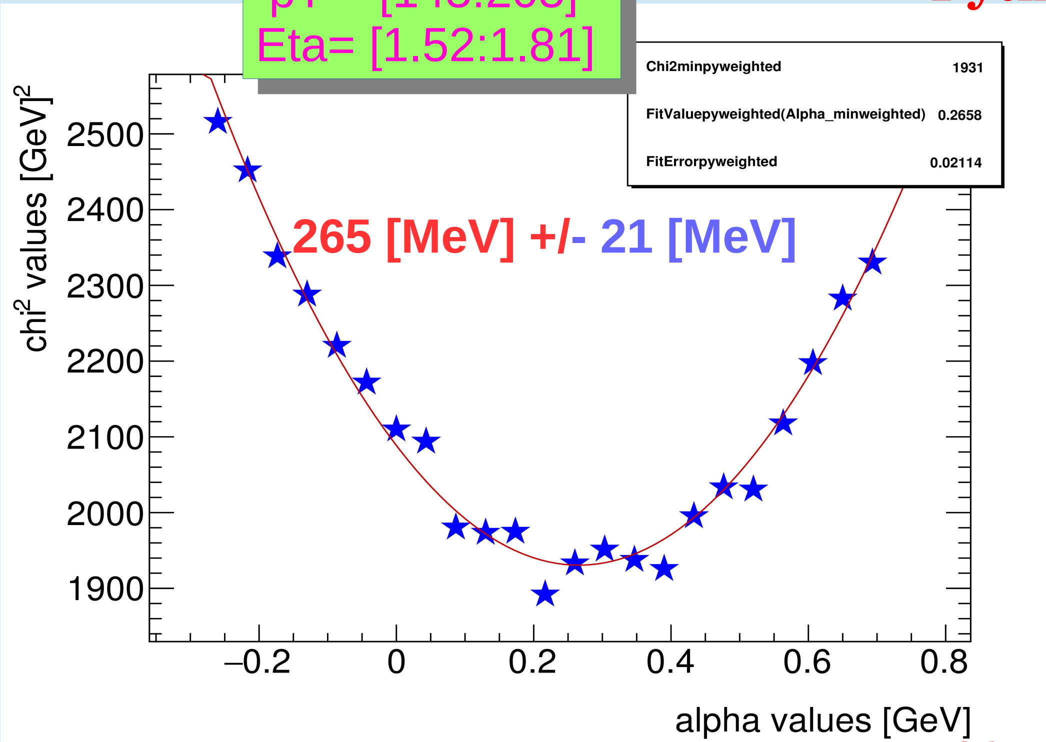
Minimum Value

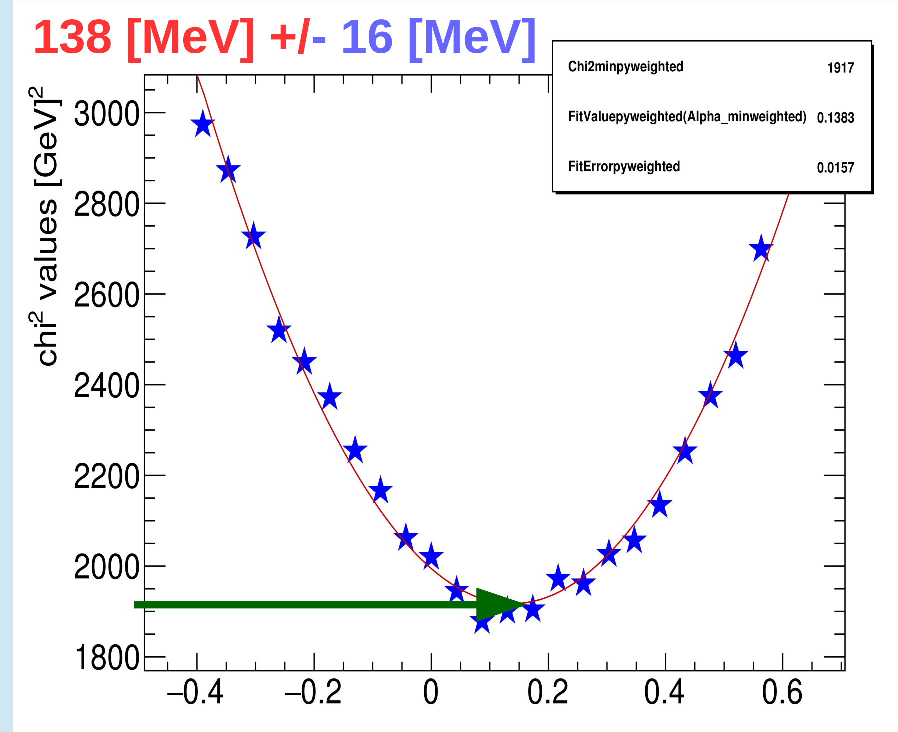
Error on shift

Best shift

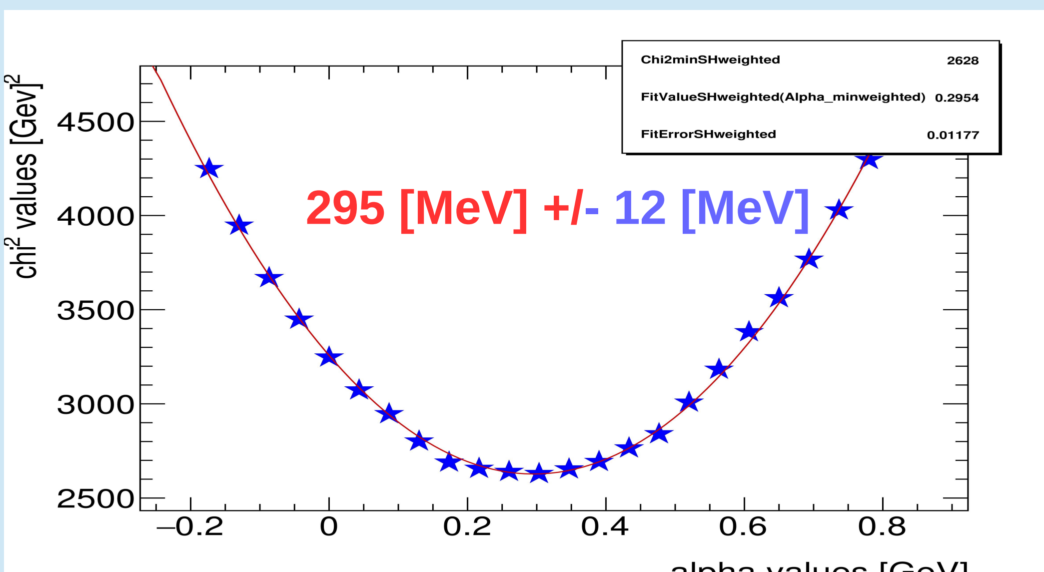
pT = [145:205]
Eta= [1.52:1.81]

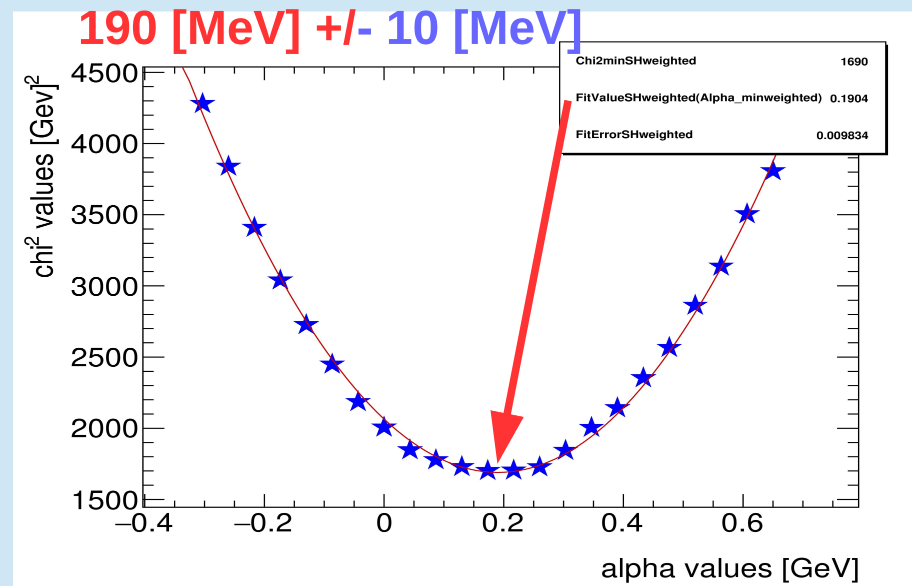
Pythia





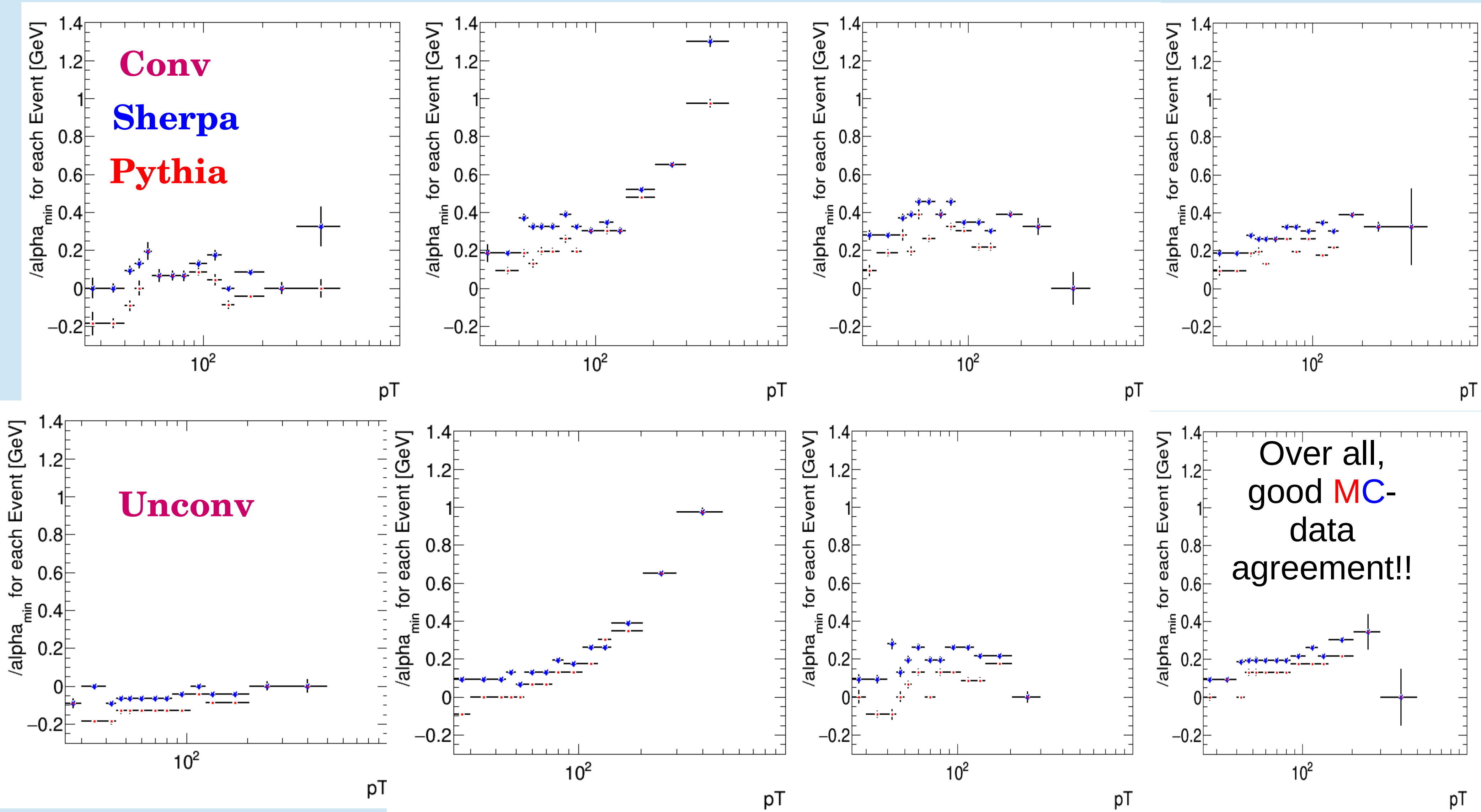
Sherpa





Conv **Unconv**

Small and independent of pT for central photons.	Larger and increasing with pT in the transition.	Slightly larger after the crack.	Slightly larger and increasing with pT in the forward.



Eta= [0:0.6] Eta= [0.6:1.37] Eta= [1.53:1.81] Eta= [1.81:2:37]

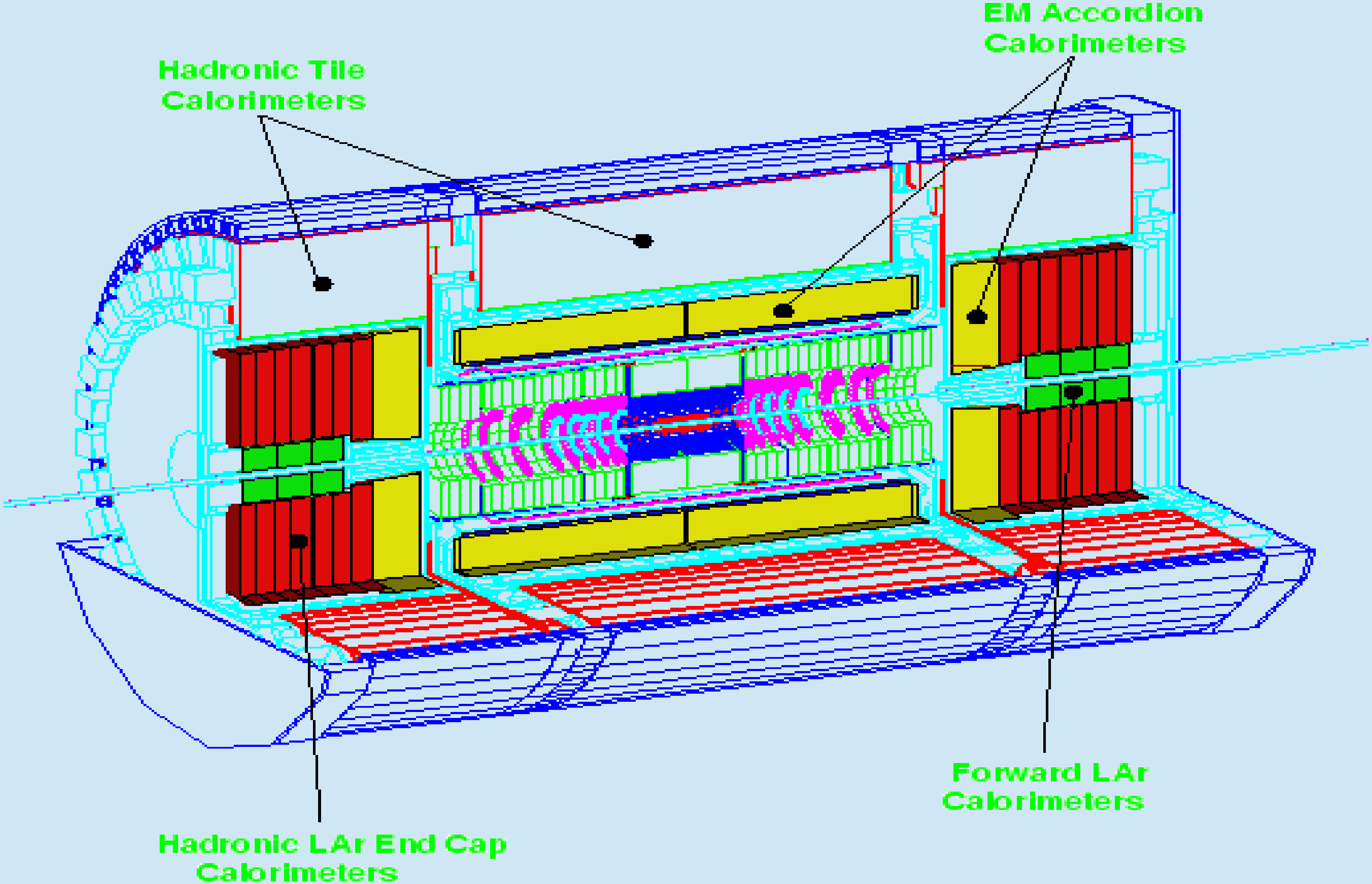
Conclusion

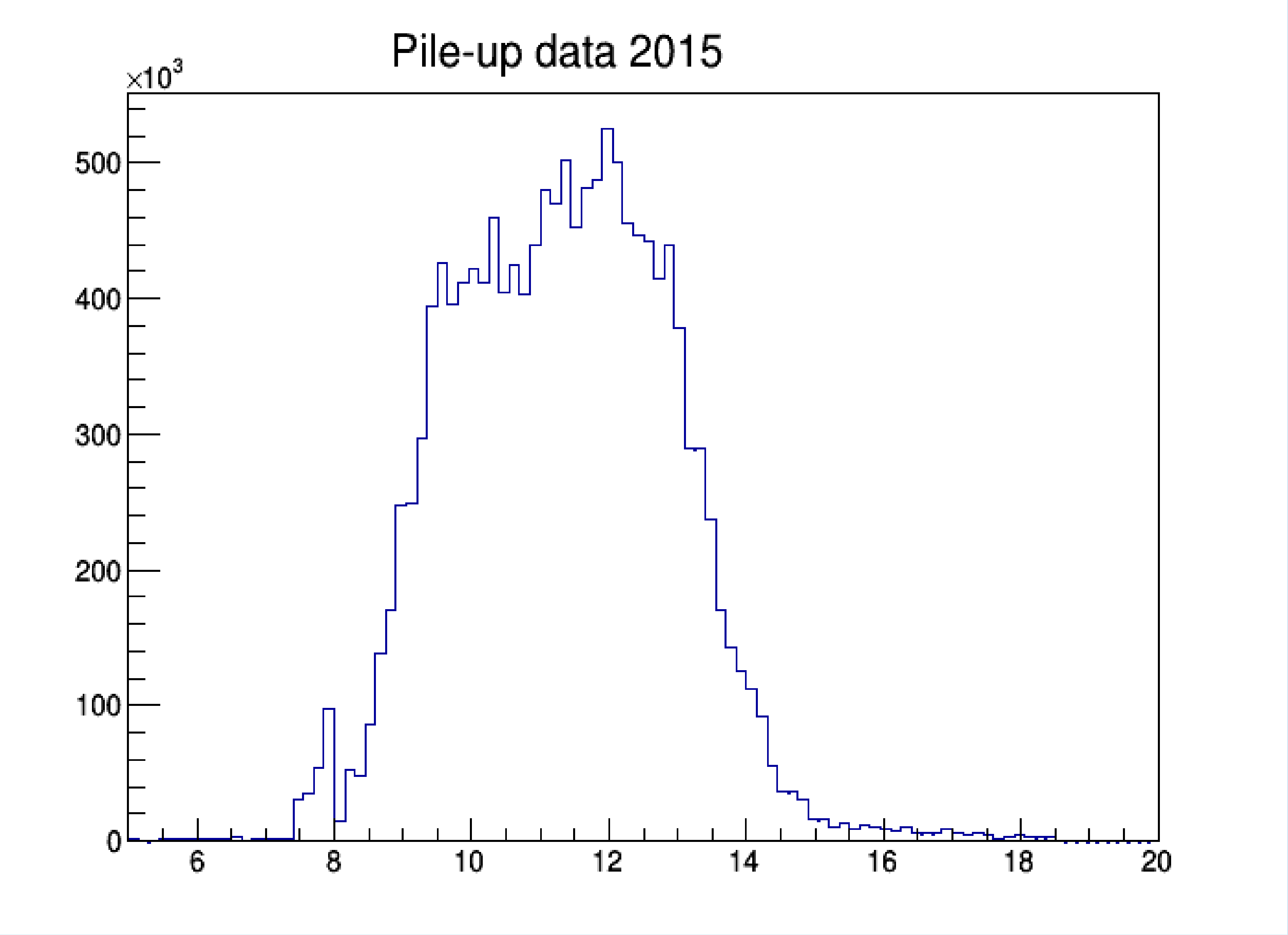
- Photon cadidates on 2015 Run II ATLAS were studied.
- Calorimetric isolation distributions were compared to MC simulations(Pythia and Sherpa).
- Signal distributions were obtaind with the data driven by bg sub method.
- Tight isolation efficiencies and purities were measured.
- Data MC agreements were estimated as shifts in the isolation distributions.
- Additional studies are needed to improve the data MC agreement.

Thank you!

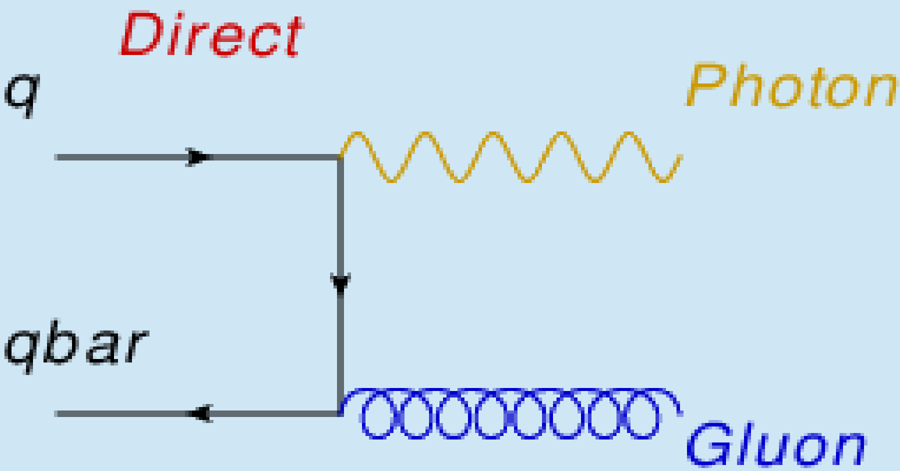
Backup slides

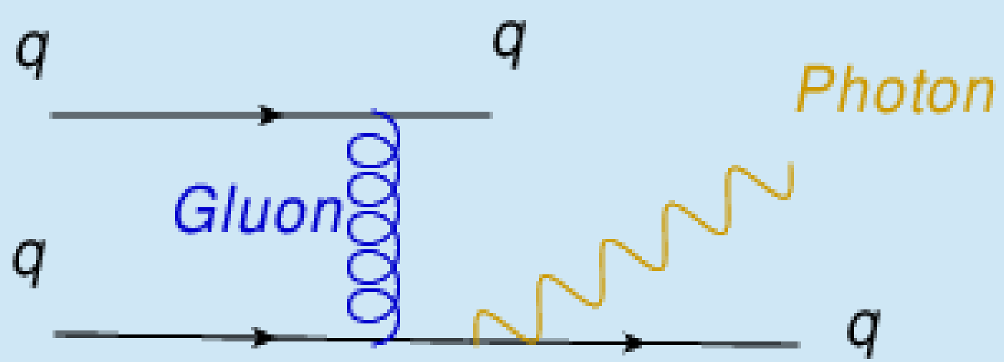
ATLAS Calorimetry (Geant)



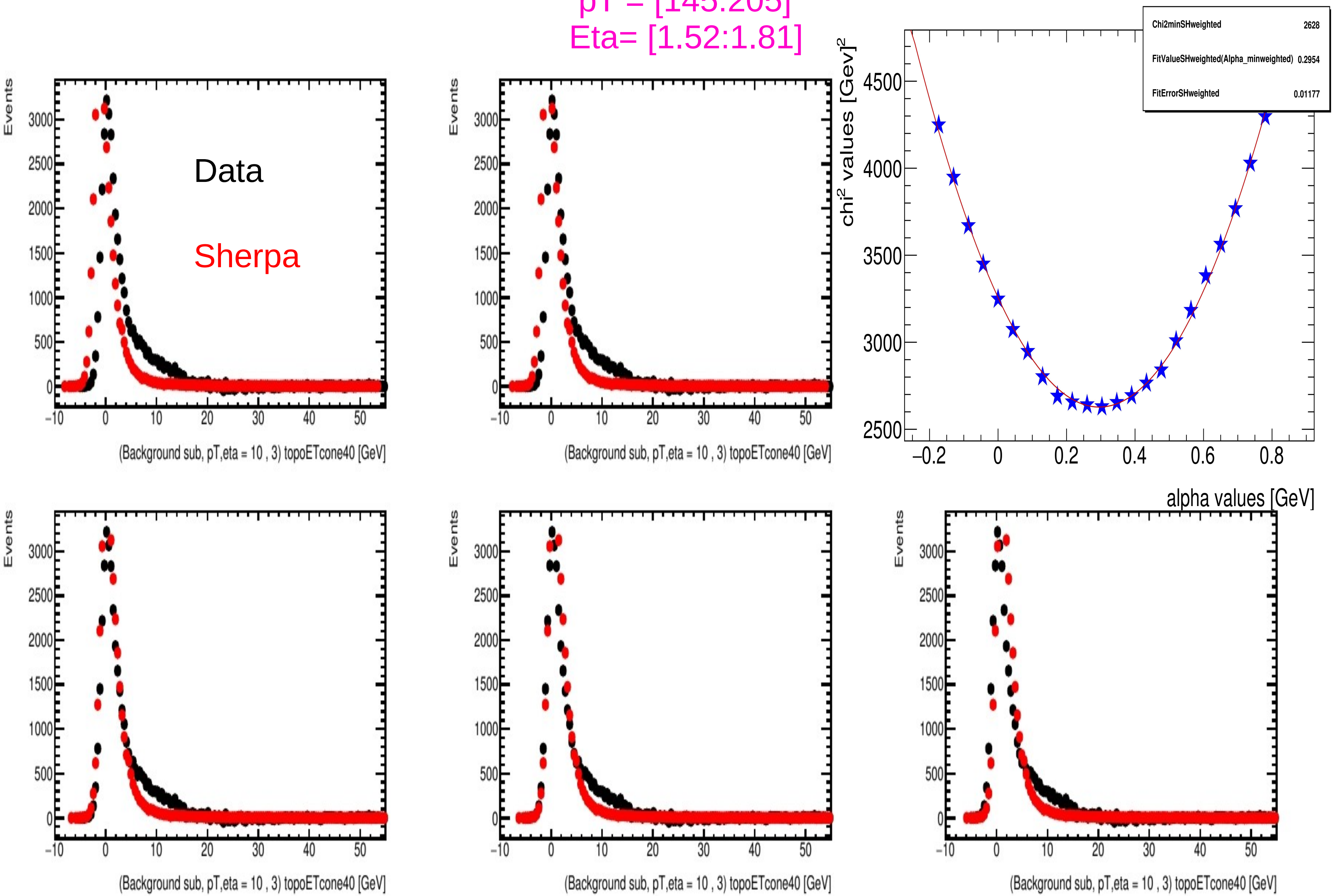

Pile-up data 2015
×10³
500
400
300
200
100
0
6
8
10
12
14
16
18
20

Direct and Brem





pT = [145:205]
Eta= [1.52:1.81]



Calorimetric variables

Category	Description	Name	Loose	Tight
Acceptance	$\|\eta\| < 2.37$, $1.37 < \|\eta\| < 1.52$ excluded	–	✓	✓
Hadronic leakage	Ratio of E_T in the first sampling of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$)	R_{had_1}	✓	✓
	Ratio of E_T in all the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{had}	✓	✓
EM Middle layer	Ratio in η of cell energies in 3×7 versus 7×7 cells	R_η	✓	✓
	Lateral width of the shower	w_{η_2}	✓	✓
	Ratio in ϕ of cell energies in 3×3 and 3×7 cells	R_ϕ		✓
EM Strip layer	Shower width for three strips around strip with maximum energy deposit	$w_{s\,3}$		✓
	Total lateral shower width	$w_{s\,\mathrm{tot}}$		✓
	Energy outside core of three central strips but within seven strips divided by energy within the three central strips	F_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer, and the energy reconstructed in the strip with the minimal value found between the first and second maxima	ΔE		✓
	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies	E_{ratio}		✓

Table 1: Variables used for loose and tight photon identification cuts.

- These variables describe the shape and structure of electromagnetic showers according to their propagation in the detector.

ATLAS-CONF-2012-123

Pseudo-χ² function & X² distribution

- **Pseudo Chi² method:** for a set n independent xi measurements is the desviation respecto a predicted value, in its invariance units.
- **Chi² distribution:** for n degrees of freedom this is the sum squares of a variable xi with a mean.

$$P\chi^2(\chi : n) = \sum_{i=1}^{n}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2$$

$$\chi^2(\alpha) = \frac{(\alpha - \alpha_{min})^2}{\sigma_\alpha^2} + \chi^2_{min}$$

Chi 2 min follows a Chi²
distribution