Studying photons in Atlas

Omar Medina Bautista[1,2]

Supervisors: Luis Pascual Dominguez[3], José Ocariz[2]
Romain Van Den Broucke[2]

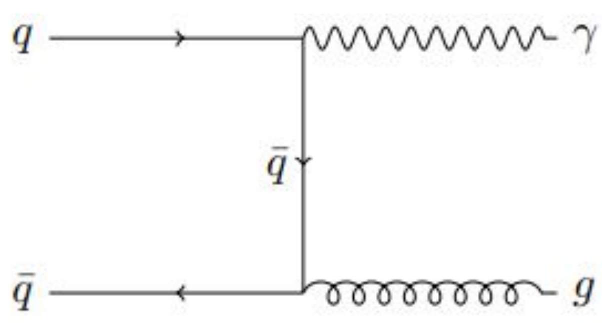
[1]UNMSM, [2]LPNHE Paris, [3]Tel Aviv University,

Latin American alliance for
Capacity buildiNG in Advanced physics
LA-CoNGA physics

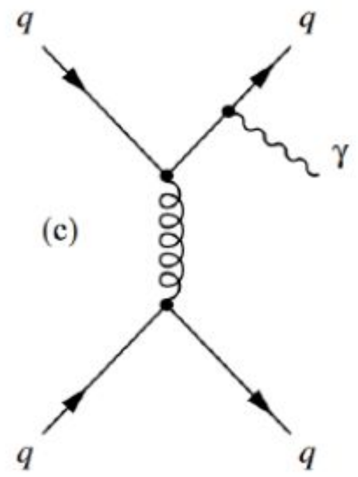
Cofinanciado por el
programa Erasmus+
de la Unión Europea

Preliminary Concepts: Photons

- Single photons are produced through several processes in **hard proton-proton collisions.** To the study of photon isolation, two type of processes are of interest, both occurring at the QCD level.



Prompt Photon
High energy
dominant



Brem Photon
Low energy
dominant

ATLAS Calorimeter

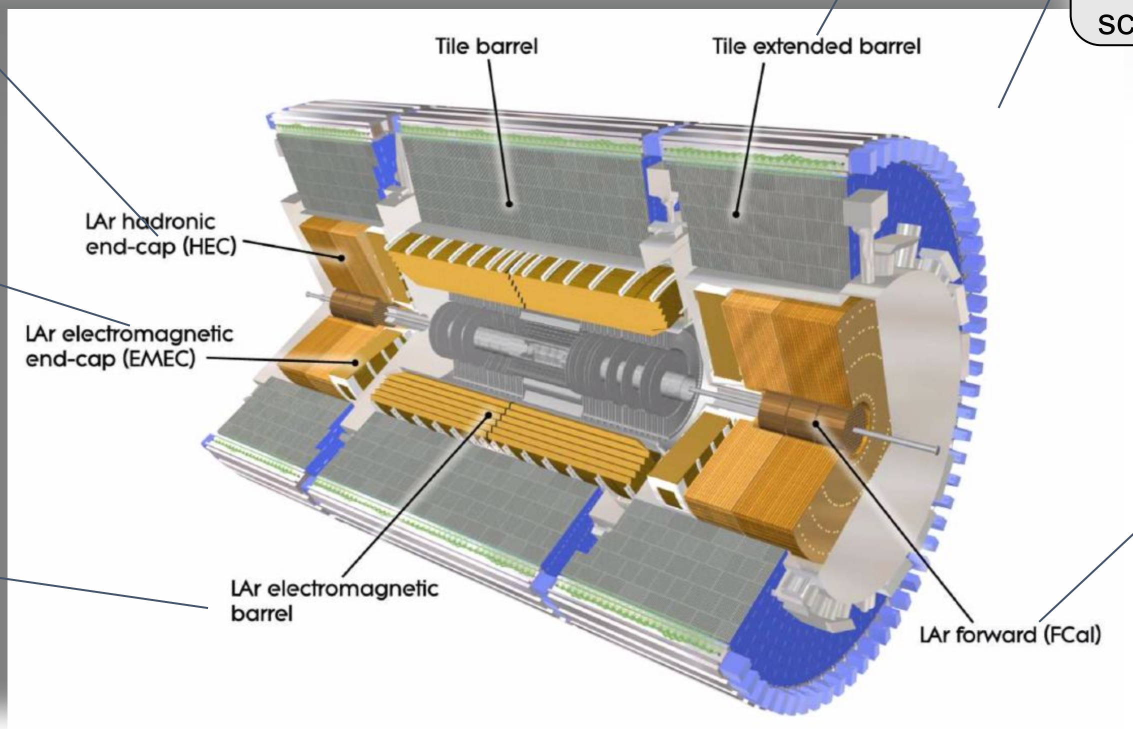
After going through the inner detector, incoming particles will reach the calorimeter which is composed of an absorber material and an active media. The first one stops particles and induced the creation of a shower of secondary-low energy particles which ionize the second media or make it radiate. Both phenomena can be translated into a signal from which the energy of the main particle can be deduced.

Hadronic calorimeter, with steel as the absorber and plastic scintillator as the active medium. Radiation produce in the scintillator is collected to produce a signal.

Hadronic calorimeter, but with LAr technology. Instead of lead used in the EM calorimeter, the absorber material is cooper.

Similar to LAr electromagnetic barrel, but disk-shaped.

Accordion-shaped calorimeter having lead as the absorber material and LAr as the active medium. Contains three layers

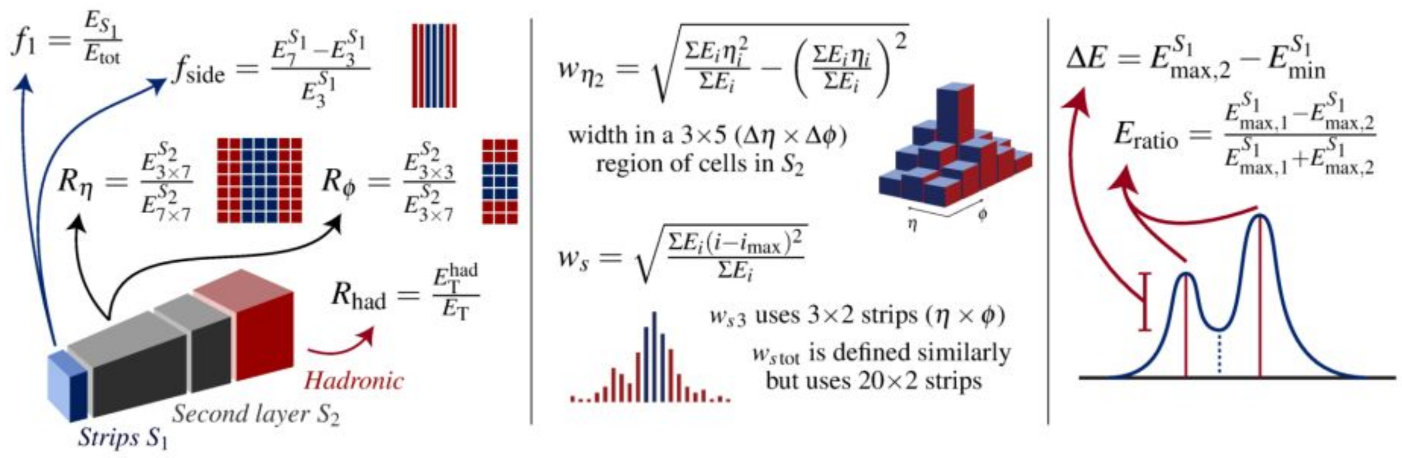


LAr calorimeter to detect both hadrons and EM particles. The first layer has an absorber cooper media (EM) and the other two layers have tungsten absorber (hadrons)

Figure 3: G Aad and Bentvelsen. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: JINST 3 (2008). Also published by CERN Geneva in 2010, S08003. 437 p. doi: 10.1088/1748-0221/3/08/S08003. url: https://cds.cern.ch/record/1129811 (cit. on pp. 39, 40, 42, 47–53).

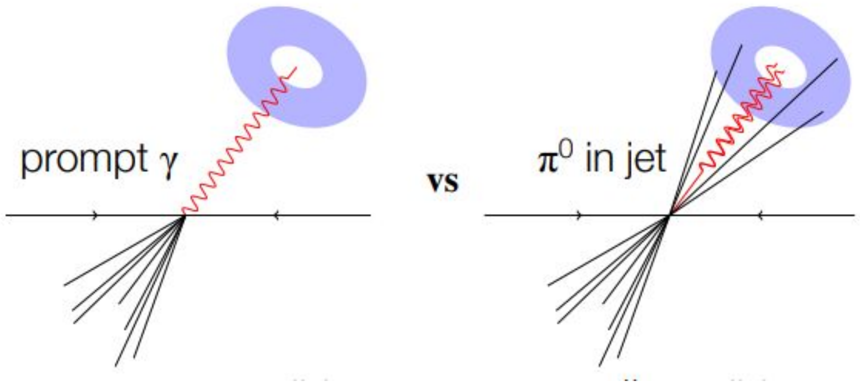
Preliminary Concepts: Photon ID

- A photon candidate is defined by cuts on the shower shape variables.
- Photons candidates are classified as Loose or Tight.
- There is an intermediate classification, Loose'4 for photons failing Tight due to cuts on 4 strip shower shape variable assumed to be uncorrelated with isolation.



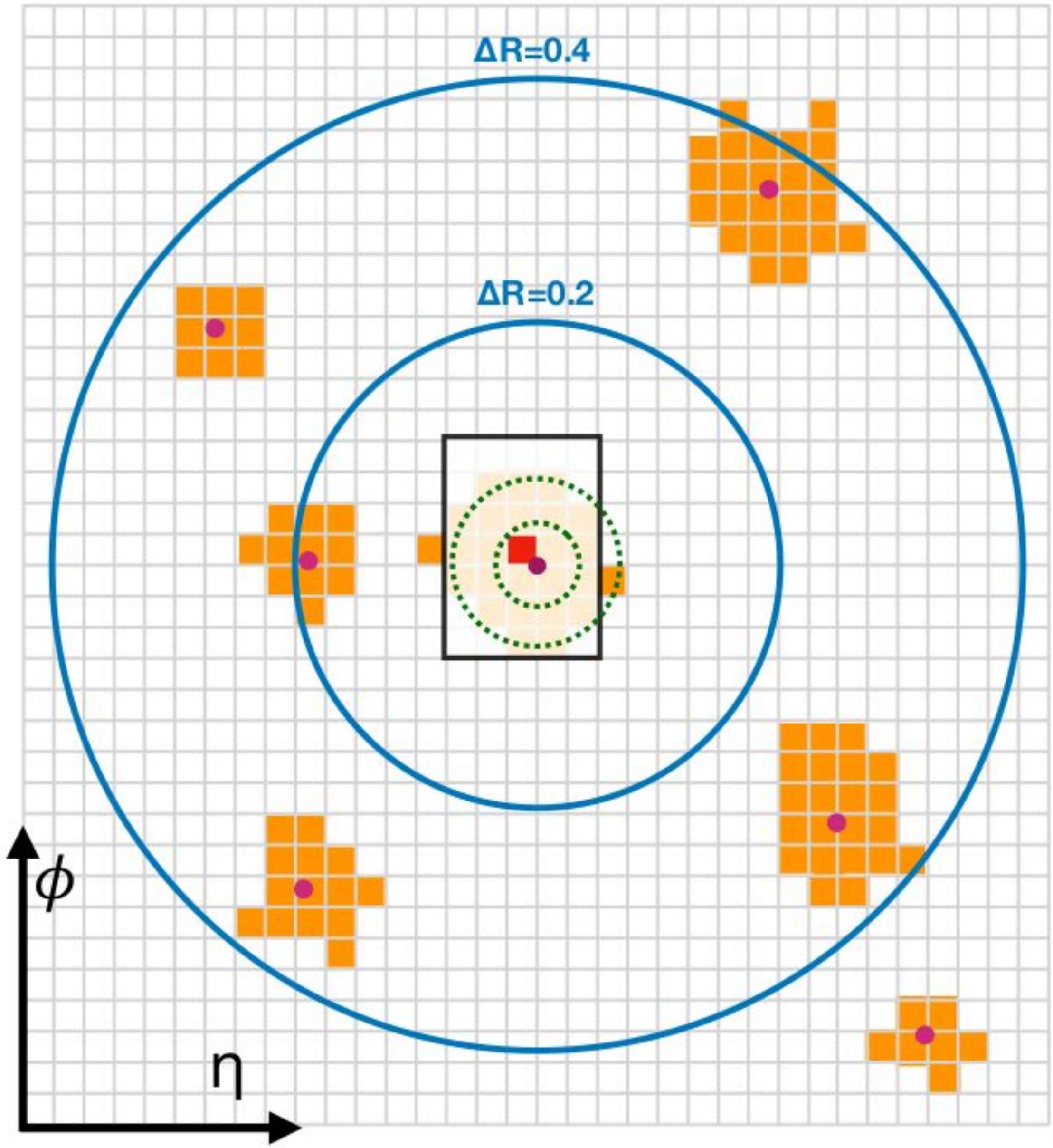
Preliminary Concepts: Fake photons

- Prompt photons such as the prompts are characterized for their almost null hadronic activity in their surroundings. On the contrary, fake photons(background) usually have an important flow of energy around them.



- This hadrons also deposit their energy on the calorimeters(LAr and hadronic). This extra energy is called isolation energy, and the method of photon isolation consists of using this quantity to distinguish photons from fake ones.

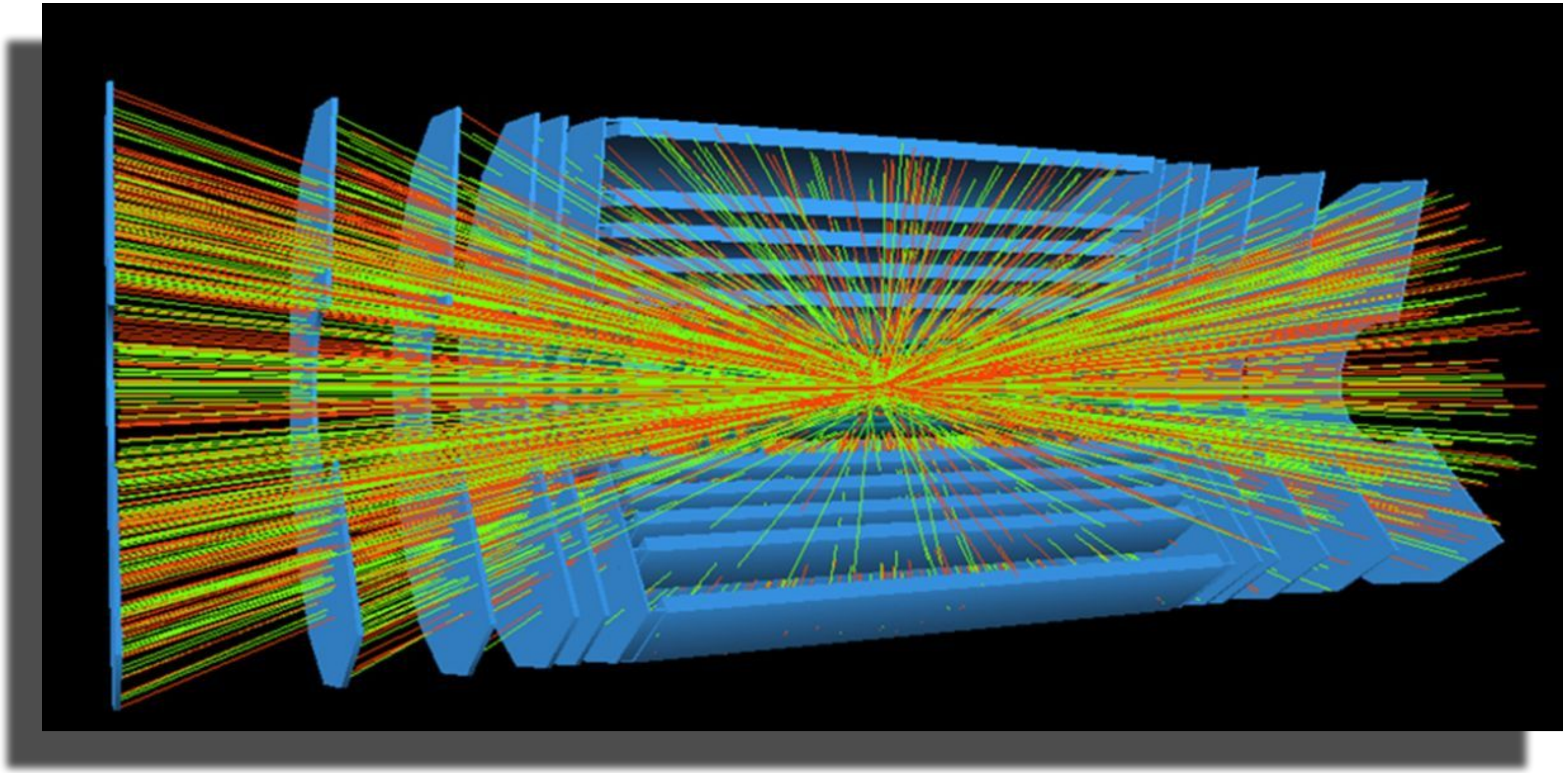
Photon candidate on the calorimeter



- An incident photon candidate activates several clusters of cells called **topoclusters.**
- Each topocluster has a "center of energy" called barycenter **(purple dot)**, which defines the center of cone of radius 0.2 or 0.4.
- The **total energy** is the sum of all topoclusters whose barycenter are inside the cone.
- The most energetic cell of the photon candidate topocluster defines the center of a **mask** of 7x5 cells which contains the core energy of the photon candidate.
- A small fraction of the photon energy **leaks** outside the mask and needs to be subtracted from the isolation energy.

Pileup and Pileup Profile

- **Pileup** is the number of proton-proton collisions besides the collision of interest (hard collisions) which can take place within the same bunch crossing or before and after the collisions due to the recovery time of the detectors (25 ns).
- **Pileup Profile** is the average number of interactions per bunch crossing (**<μ>**).



Isolation energy definition

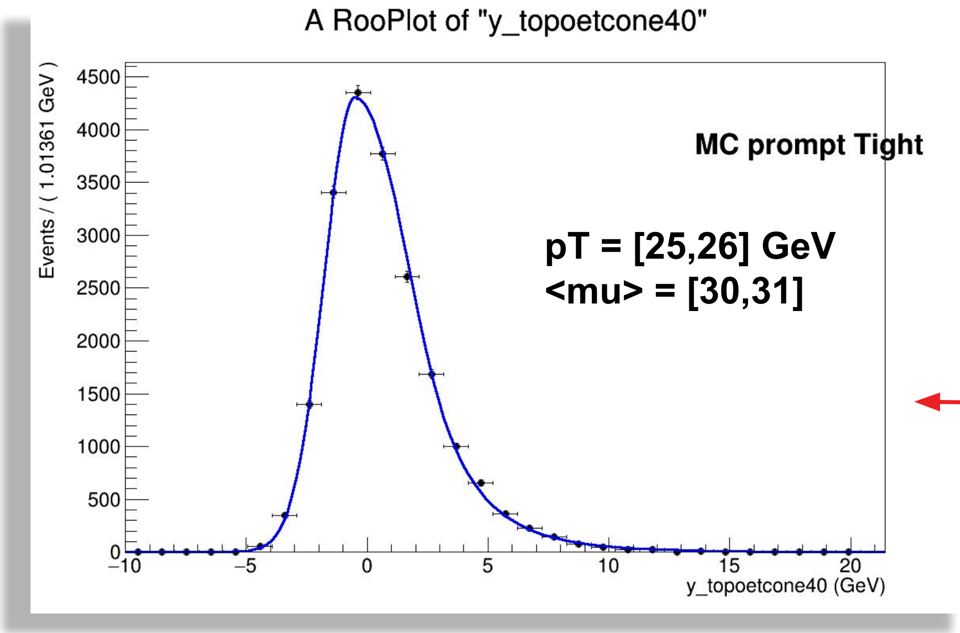
- To find the extra energy **(isolation energy)**, the total energy must be subtracted by the core energy, its leakage, and pileup energy distribution.

$$E_T^{iso}(R) = \left(\sum_{r_i<R}^{clusters} E_T^i\right) - E_T^{core} - E_T^{leakage}(E_T^\gamma, |\eta^\gamma|) - E_T^{pileup}(|\eta^\gamma|, \langle\mu\rangle)$$

Average corrections

leads to

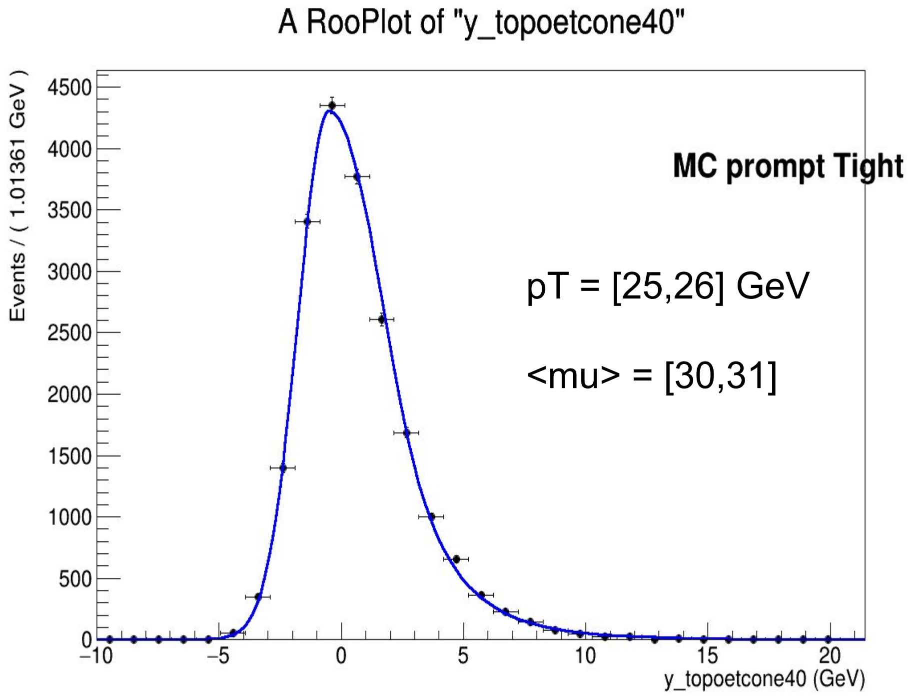
Fluctuations



PDFs

An ACB function is used to fit the isolation energy distribution:Signal MC

$$f(x;\mu_P,\sigma_L,\sigma_R,\alpha,n) = N\cdot\begin{cases}\exp\left(-\frac{(x-\mu_P)^2}{2\sigma_L^2}\right) & for\, x<\mu_P\\ \exp\left(-\frac{(x-\mu_P)^2}{2\sigma_R^2}\right) & for\, x\geq\mu_P\, and\, \frac{x-\mu_P}{\sigma_R}<\alpha\\ \frac{A}{\left(B+\frac{x-\mu_P}{\sigma_R}\right)^n} & for\, \frac{x-\mu_P}{\sigma_R}>\alpha\end{cases}$$



σR and σL -> fluctuations on leakage and pileup corrections (GeV)

μ -> the peak position (GeV)

α -> “Gaussian shape” (No units)

Parameters depend on pT and <mu>!!

Montecarlo Weighted PDFs: (Ft = Tight Prompt Fraction, FL = Lose Brem Fraction)

$$PDF_{MC,tight} = (F_T)ACB_{MC,prompt,tight} + (1-F_T)ACB_{MC,brem,tight}$$

$$PDF_{MC,loose} = (F_L)ACB_{MC,prompt,loose} + (1-F_L)ACB_{MC,brem,loose}$$

Fitting process

- Fit the MC distributions in 4 regions (Prompt-Tight, Prompt-Loose, Brem-Tight, Brem-Loose) in bins of p_T and pileup with ACB PDF

- Fit the fitted PDF parameters with analytical 2D functions of p_T and pileup
    - parameter(pT,mu) = Function(pT,mu)

- Fit MC + data simultaneously, in bins of p_T and pileup
    - Cross Section weighted of MC Prompt and Brem
    - MC Tight Signal: Constrained to the analytical functions, unless normalization and tight fraction
    - Data: Data Tight Signal constrained to the MC Tight Signal, except peak position for MC (= data-driven shift)

- Fit the data-driven shift to get an analytical function of p_T and pileup

More PDFs and Data Driven Shift

Extended PDFs: (g = Lose Data Fraction, p = purity, f = Tight Fraction)

Assumption: PDF background is ACB

$$PDF_{MC,prompt,tight} = fN_{MC,prompt}ACB_{MC,prompt,tight}$$

$$PDF_{MC,prompt,loose} = (1-f)N_{MC,prompt}ACB_{MC,prompt,loose}$$

$$PDF_{data,loose} = (1-\frac{(1-f)(1-g)p}{f})N_{data}PDF_{bkg} + \frac{(1-f)(1-g)}{f}N_{data}PDF_{loose,signal}$$

$$PDF_{data,tight} = (1-g)(1-p)N_{data}PDF_{bkg} + (1-g)pN_{data}PDF_{signal}$$

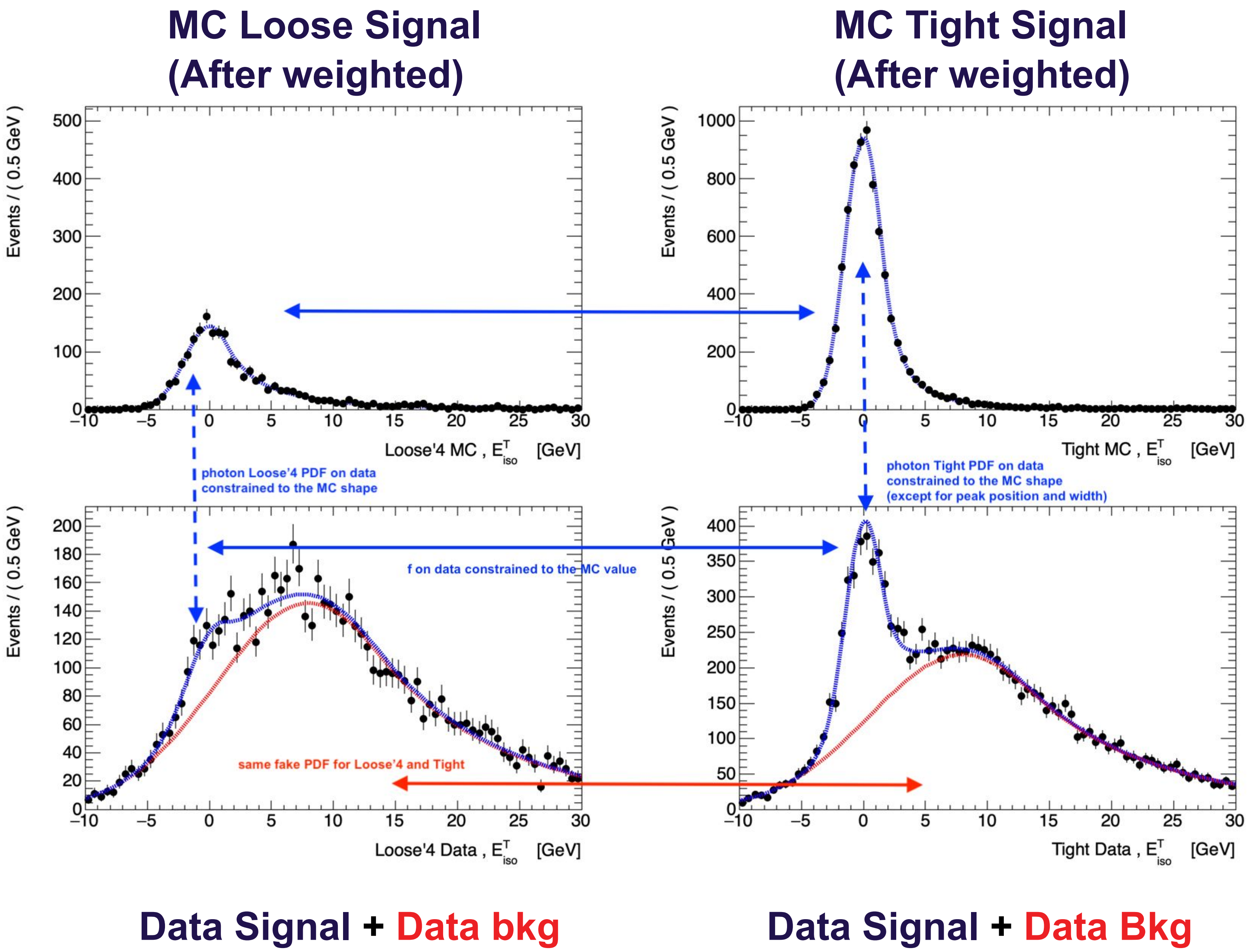
Definition Data Driven Shift

$$PDF_{MC,tight} = PDF_{MC,tight}(\mu) \qquad PDF_{signal} = PDF_{MC,tight}(\mu + d)$$

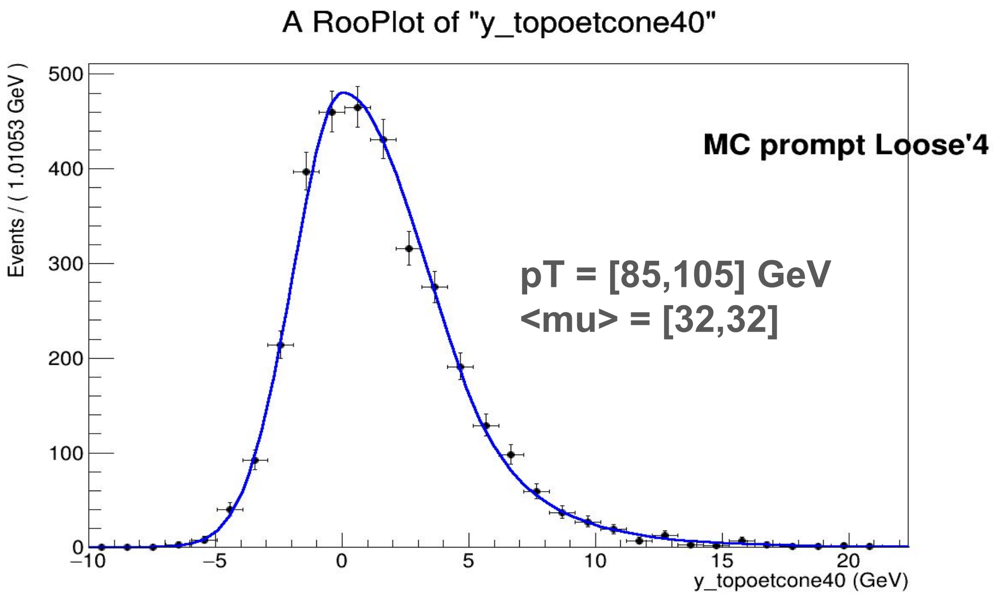
d = Data Driven Shift

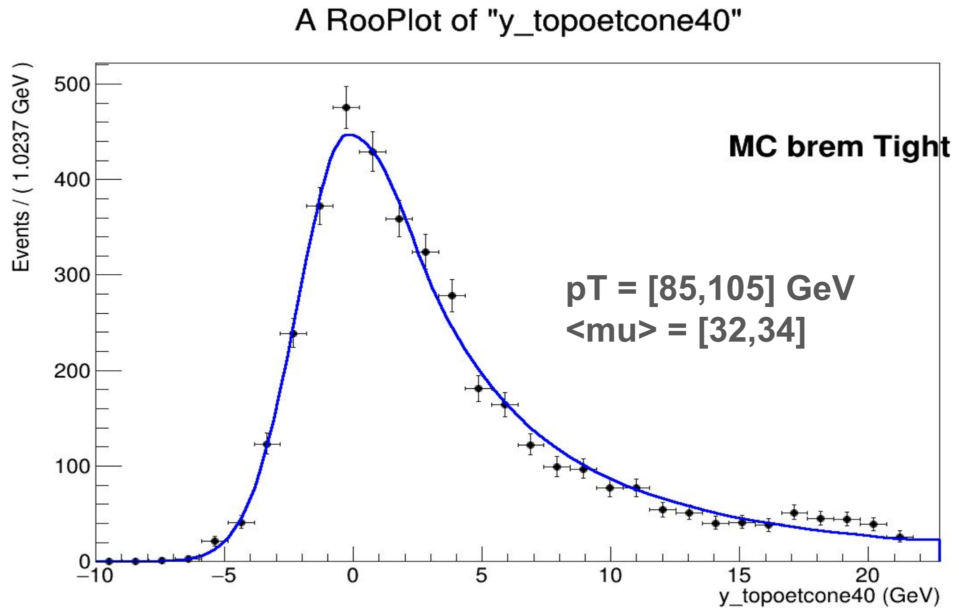
Unbinned combined fit

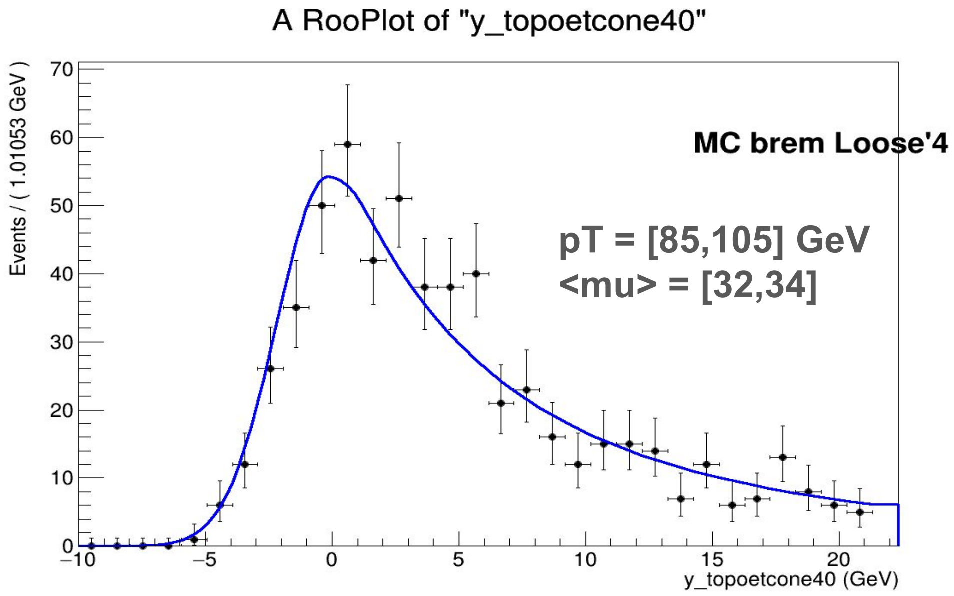
- Fit in a bin of isolation energy (a year, a region of the calorimeter, a photon energy, a pileup range, etc.)
- 6 regions: MC Prompt / MC Brem / Data, Tight / Loose’4
- MC constrained to the previous fit, except for normalization and fraction of Tight/Loose’4
- Data
  - Signal constrained to MC (except for peak position in Tight region), with Prompt and Brem in good proportions
  - Same background in Loose’4 and Tight
  - Fraction of Tight/Loose’4 constrained to MC value

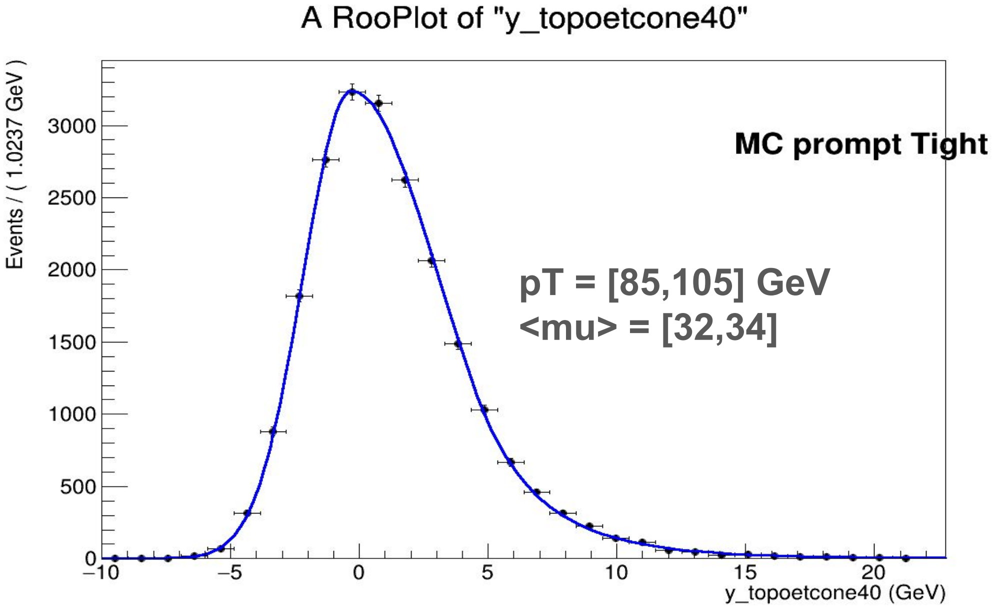


Fitting MC Distributions

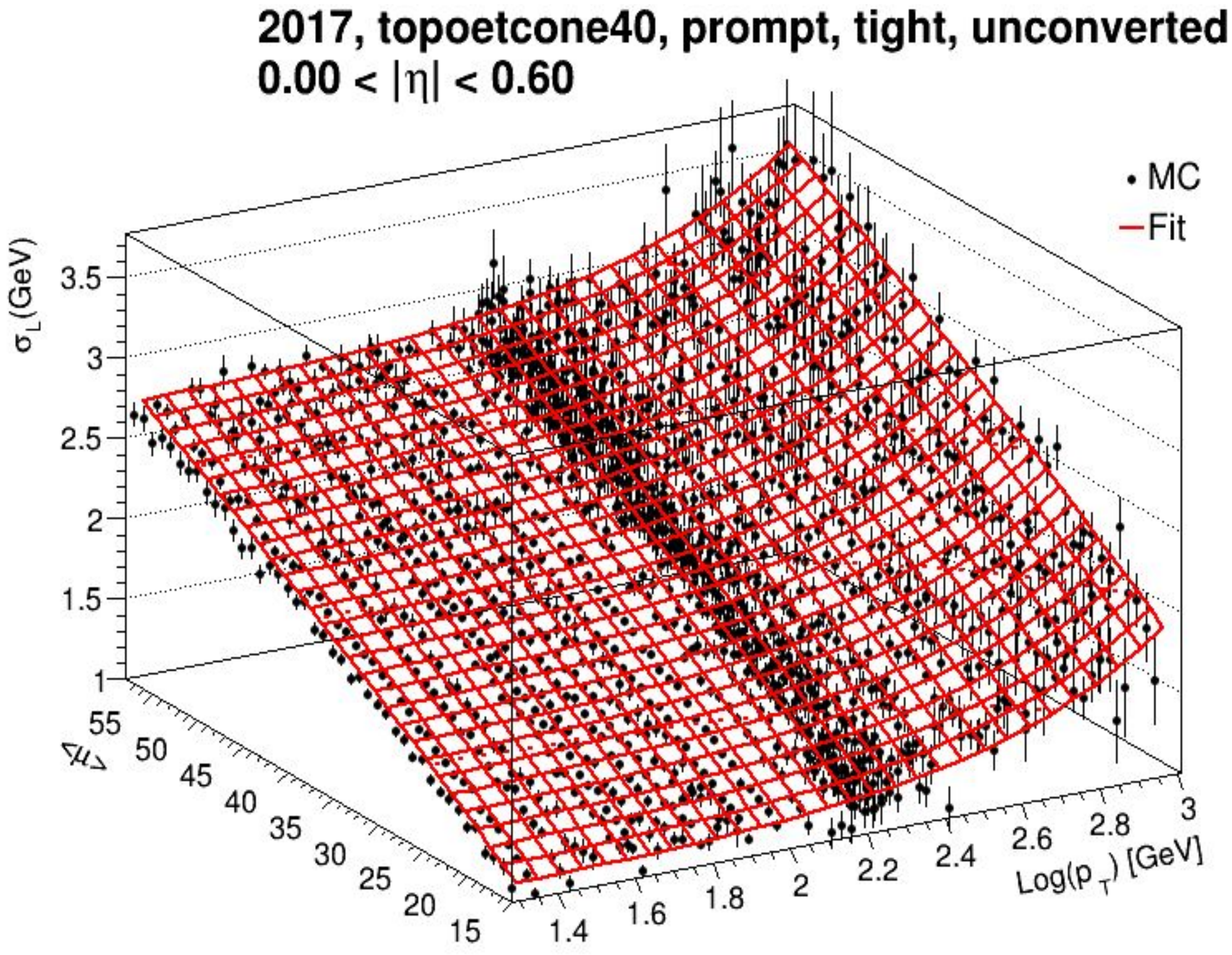








σL-PromptTight Histogram-Fit



Observation: σL increases for both pT and <μ>.

Explanation: σL quantifies fluctuations, and those increases in principle as the square root of the variables.

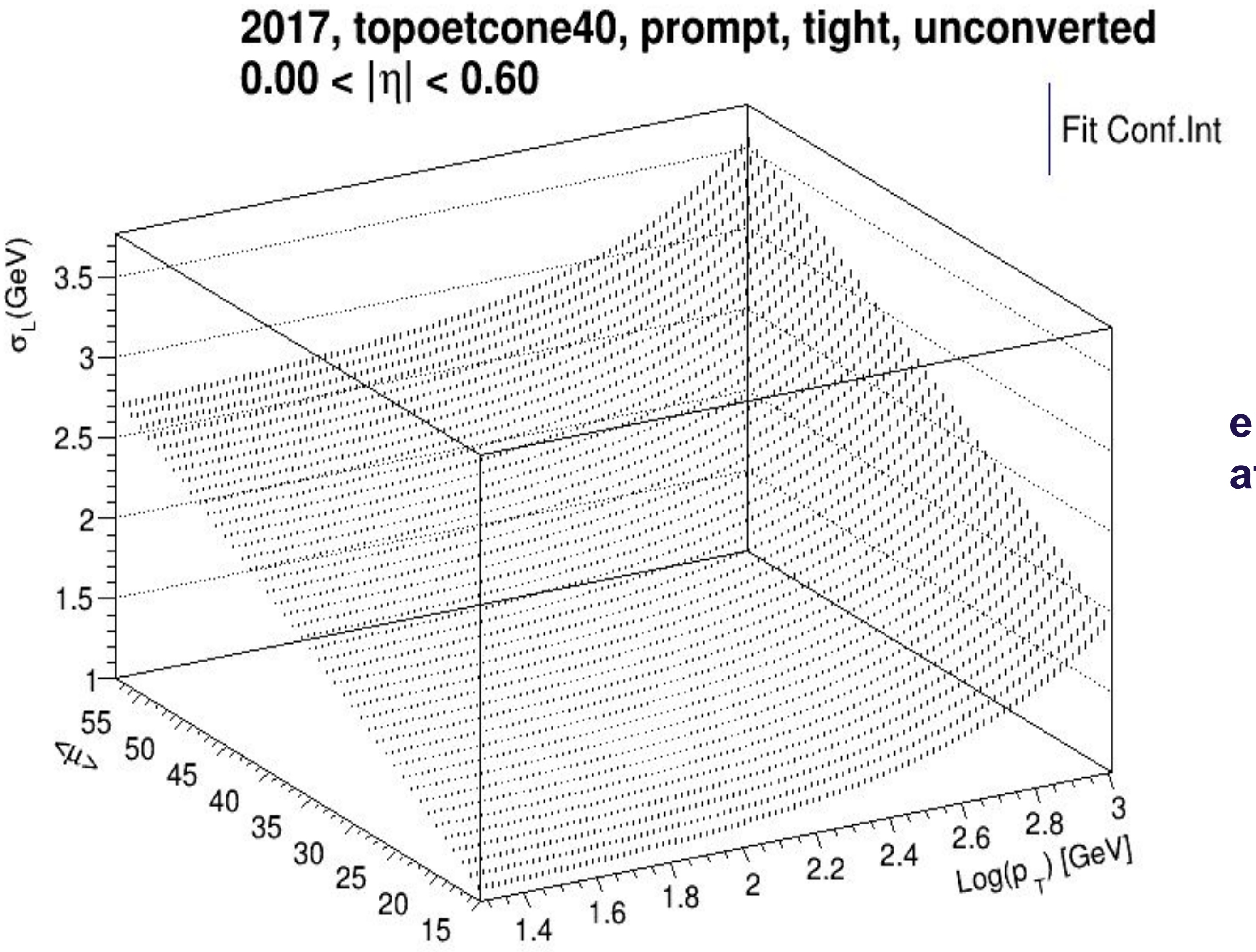
```
Chi2                       =     1226.92
NDf                        =        1186
```

```
param_0          =      0.77806   +/-    0.0330487
param_1          =     0.115704   +/-    0.0723065
param_2          =     0.110863   +/-    0.0176515
param_3          =  0.000828416   +/-    1.43453e-05
```

$$\sigma_L = p_2 < \mu >^{p_0} + p_1 + p_3 10^{p_T}$$

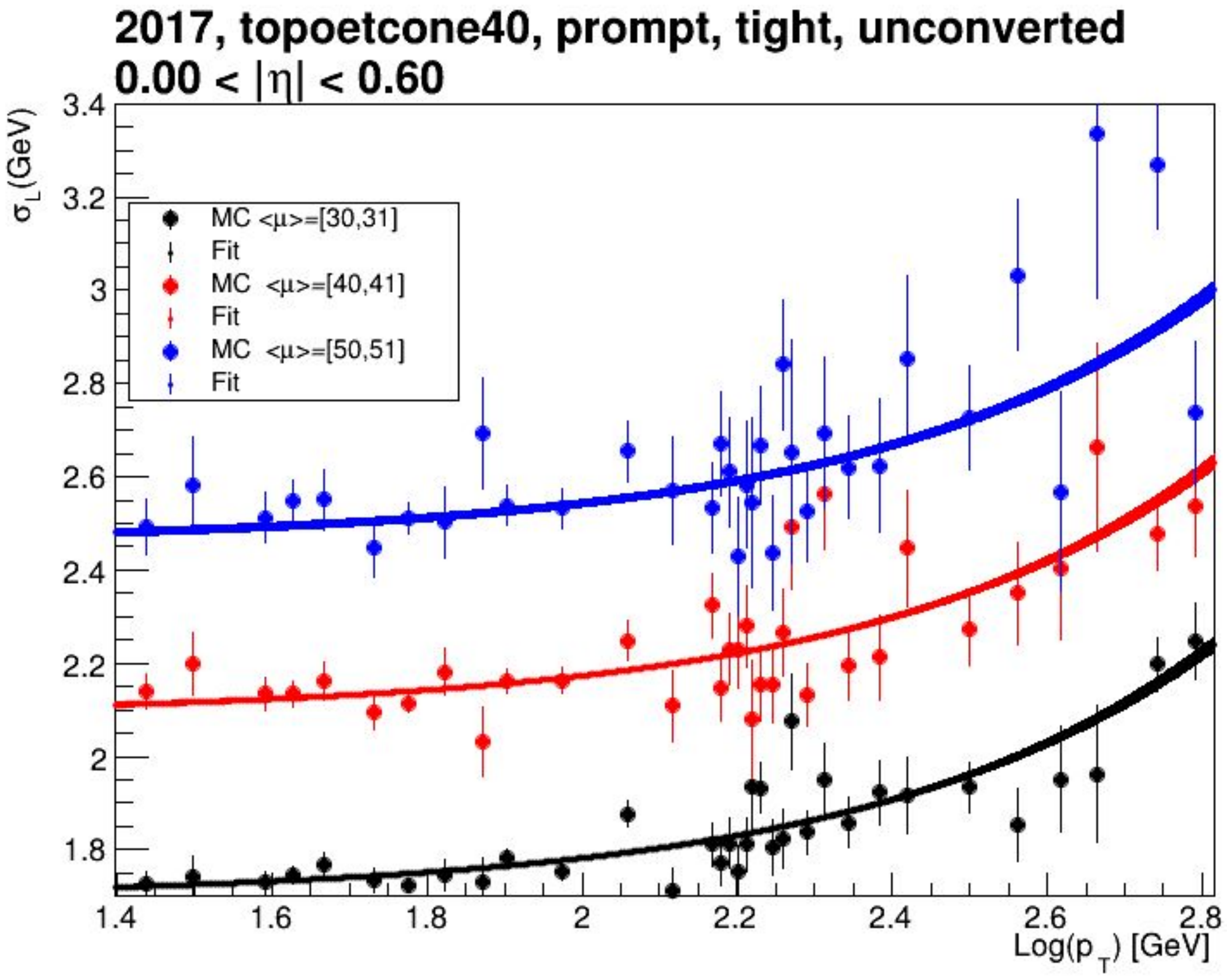
p-Value 0.2

σL-PromptTight Fitting Function Confidence Interval

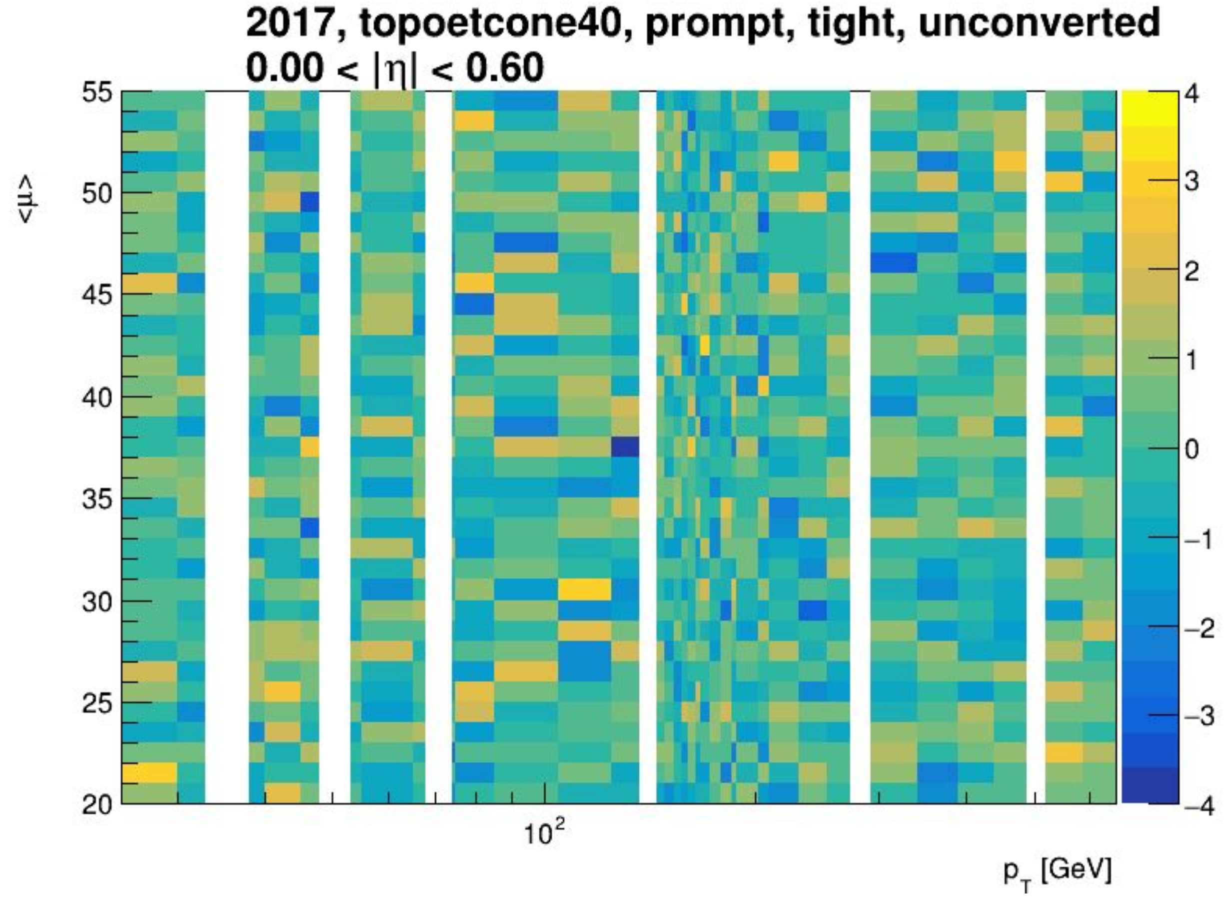


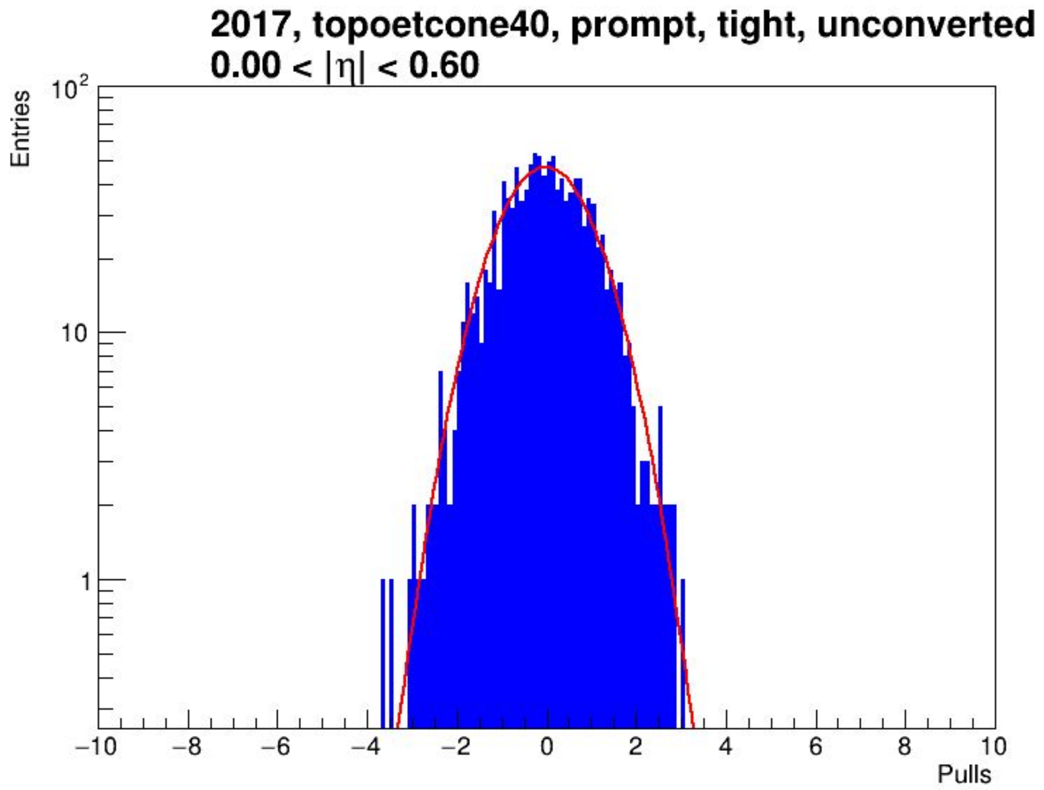
error =0.0158 GeV
at pT 39.8 GeV and <mu> = 54

σL-PromptTight 1D Trends (MC + Confidence)



σL-PromptTight Pulls



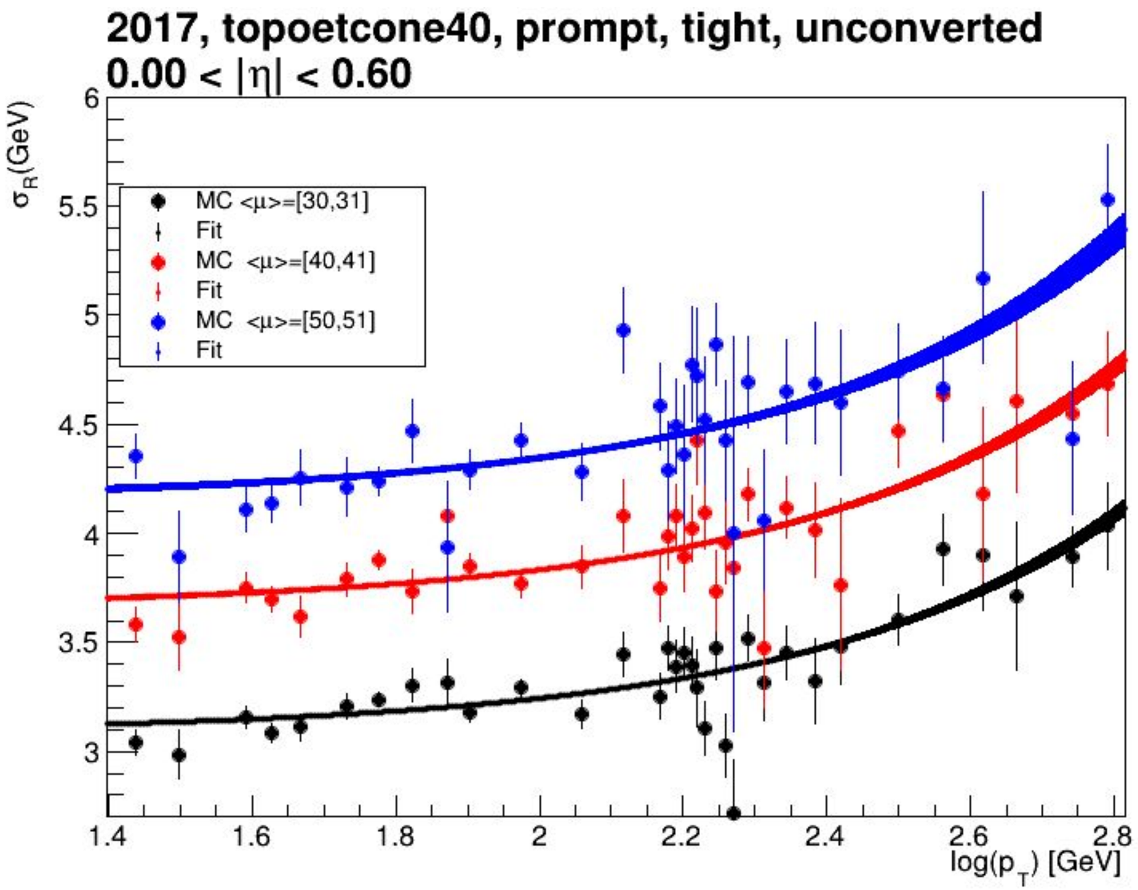


$$pull = \frac{z_{ij} - f(x_i, y_j)}{\sigma_{ij}}$$

Pulls appear randomly distributed
Mean compatible with zero
RMS compatible with one

Mean = -0.0120328 +/- 0.0294464
RMS = 1.01512 +/- 0.0209234

σR-PromptTight Fit Results and 1DTrends



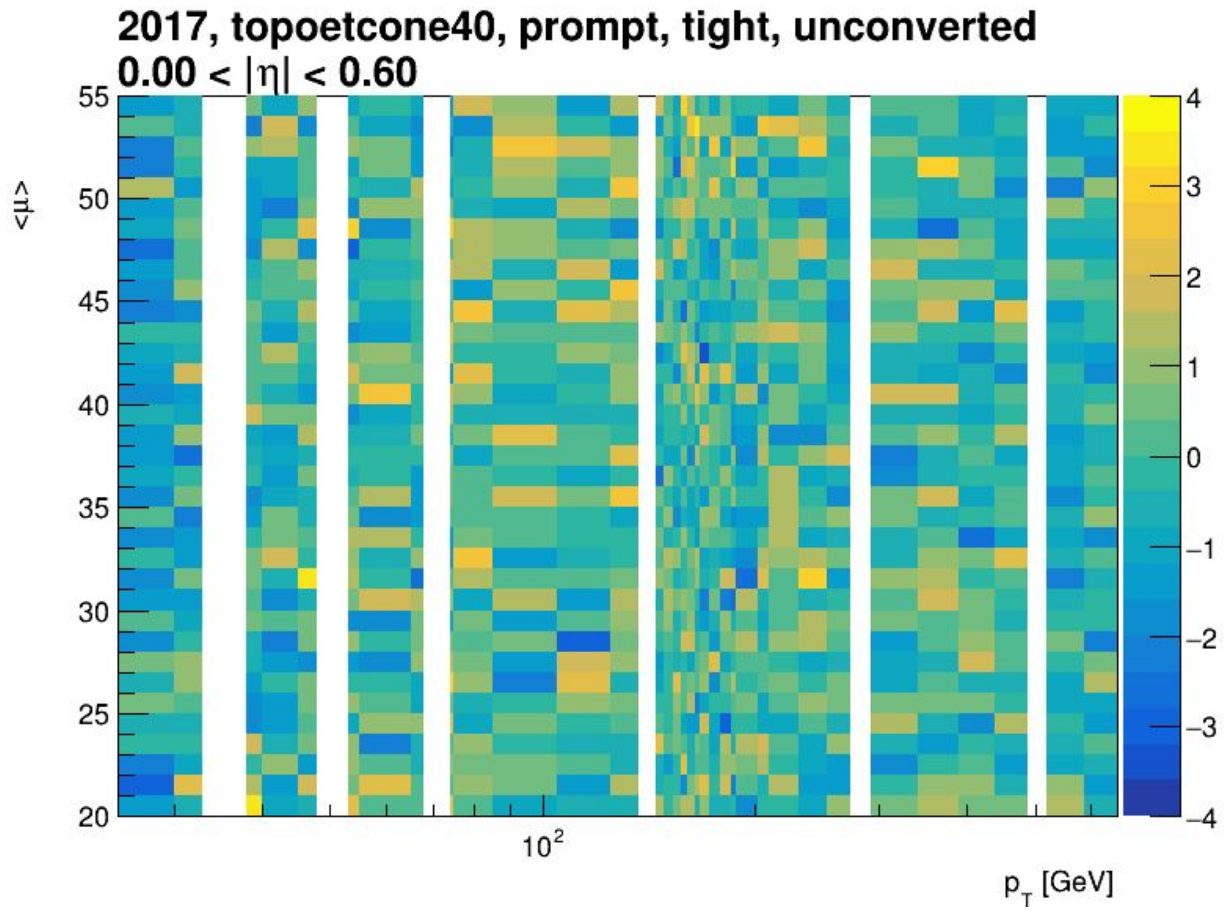

```
Chi2                          =      1482.84
NDf                           =         1186
```

```
param_0        =   0.432838    +/-   0.00141554
param_1        =   -1.30397    +/-   0.0228829
param_2        = 1.5741e-05    +/-   3.64176e-06
param_3        = 0.00109904    +/-   0.00012762
```

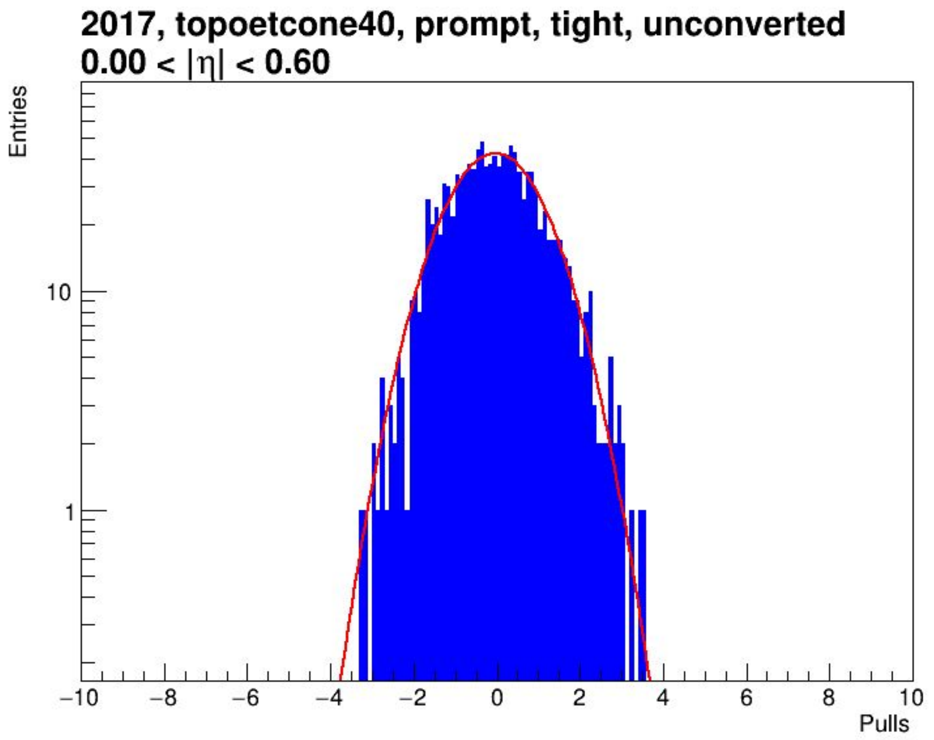
Observation: σL increases for both pT and <μ>.

Explanation: same reason as for σL

σR-PromptTight Pulls



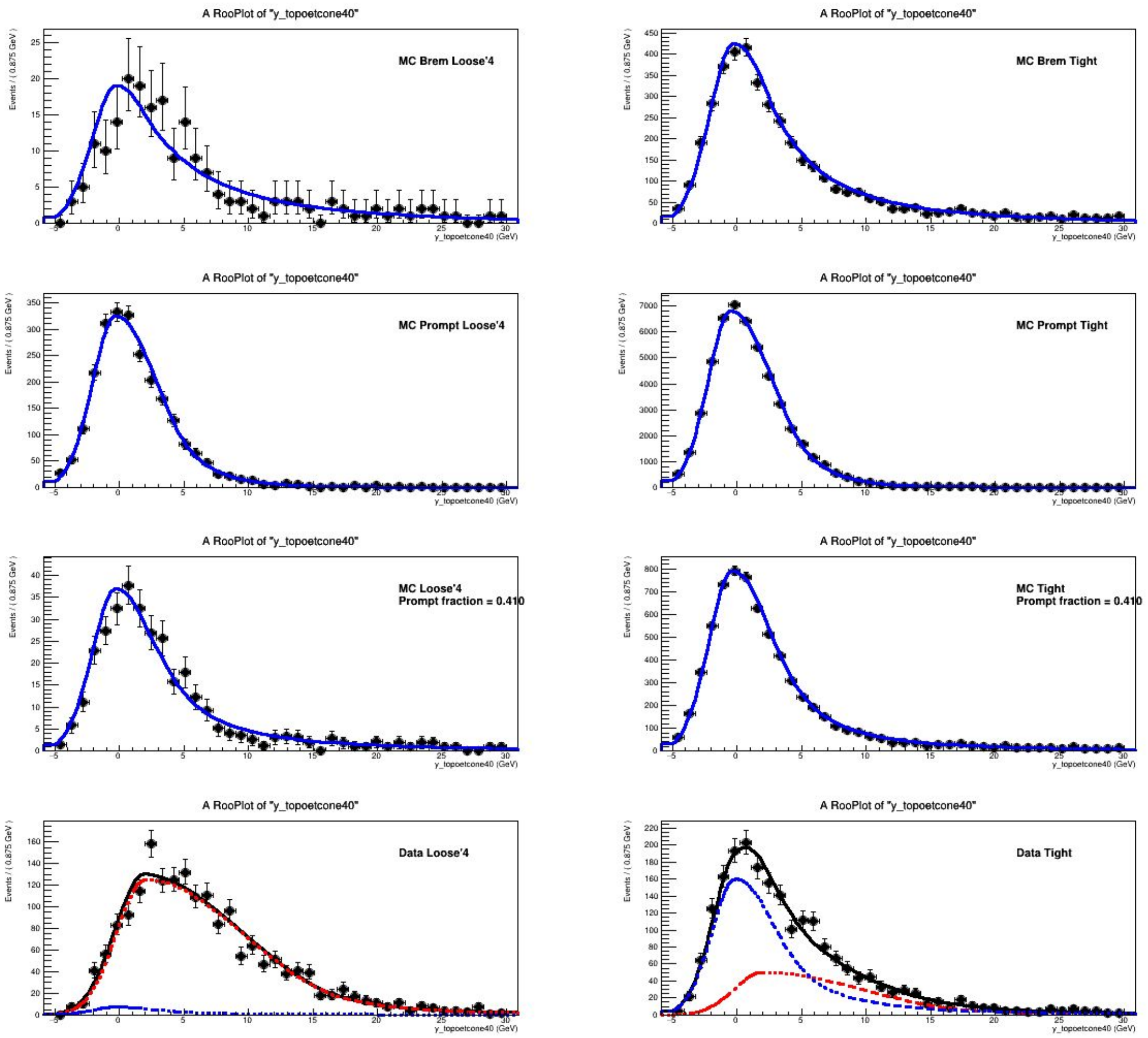
Mean compatible with zero
RMS not compatible with one



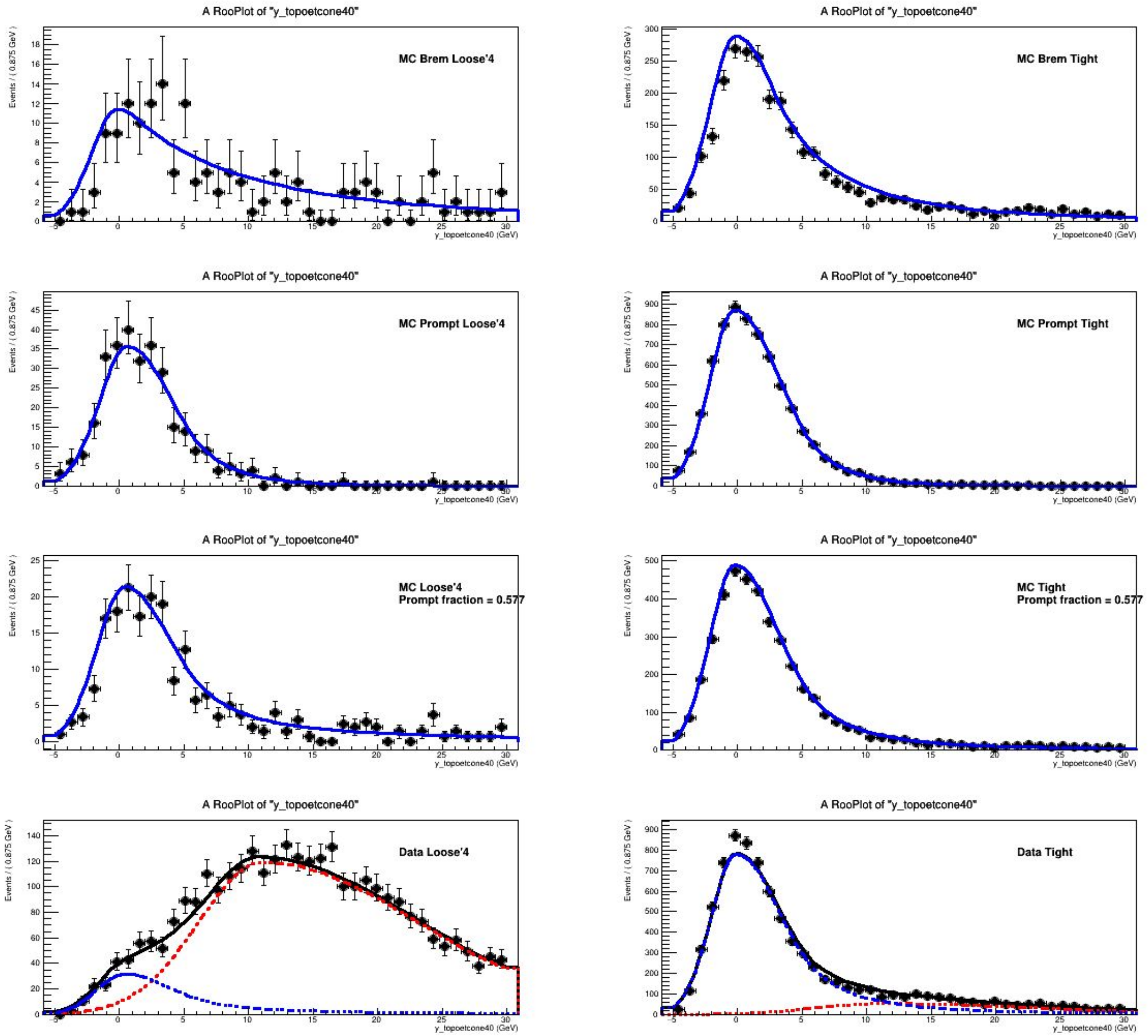
Mean = -0.0343635 +/- 0.0324723
RMS = 1.11752 +/- 0.0232902

Combined Fit Example

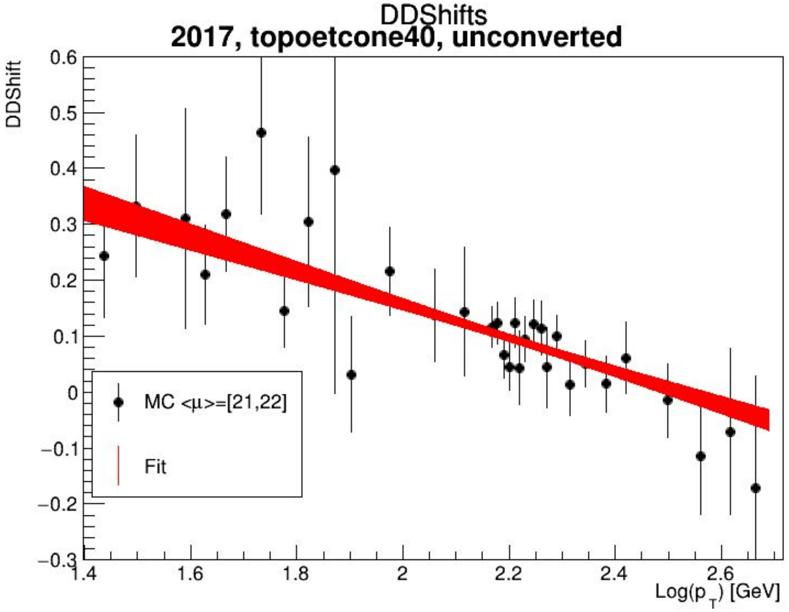
40-45 GeV



210-231 GeV



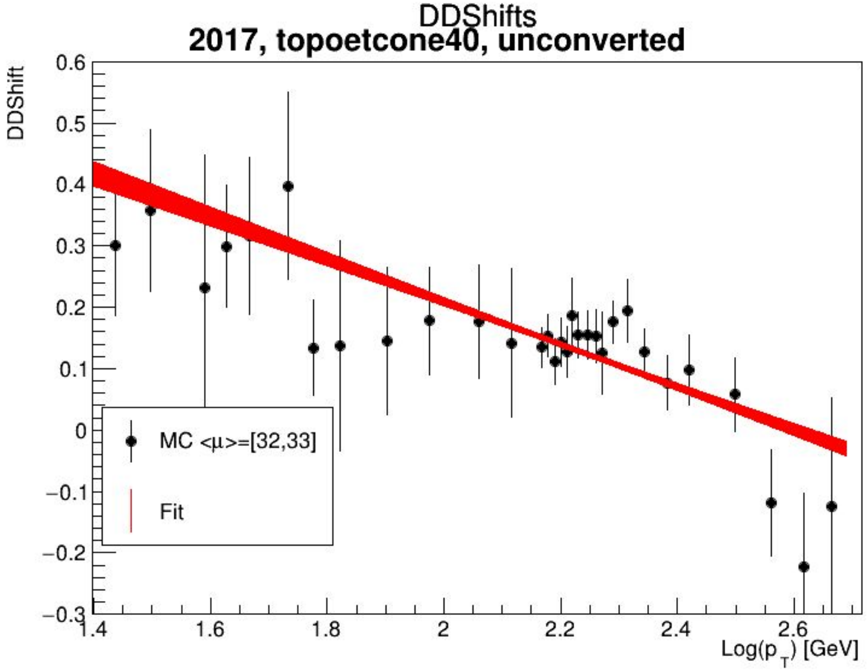
Data Driven Shifts

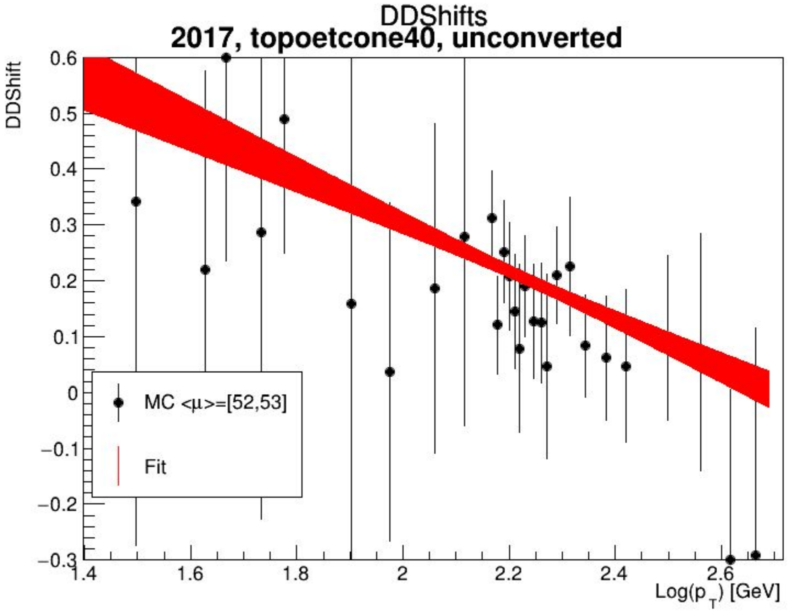


<mu> = [21-22]

Observation: DD shifts in general decrease with pT,

<mu> = [32-33]

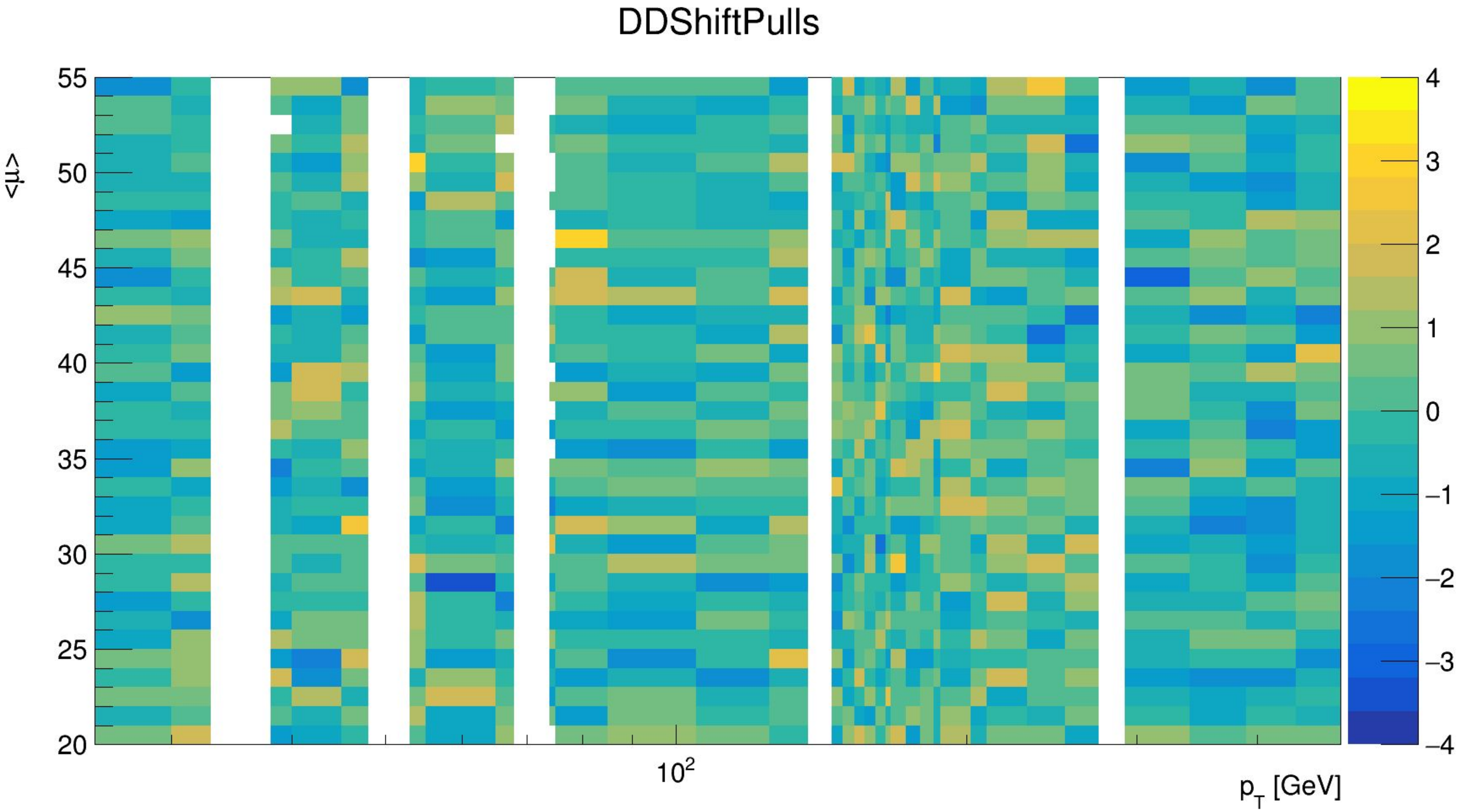




<mu> = [52-53]

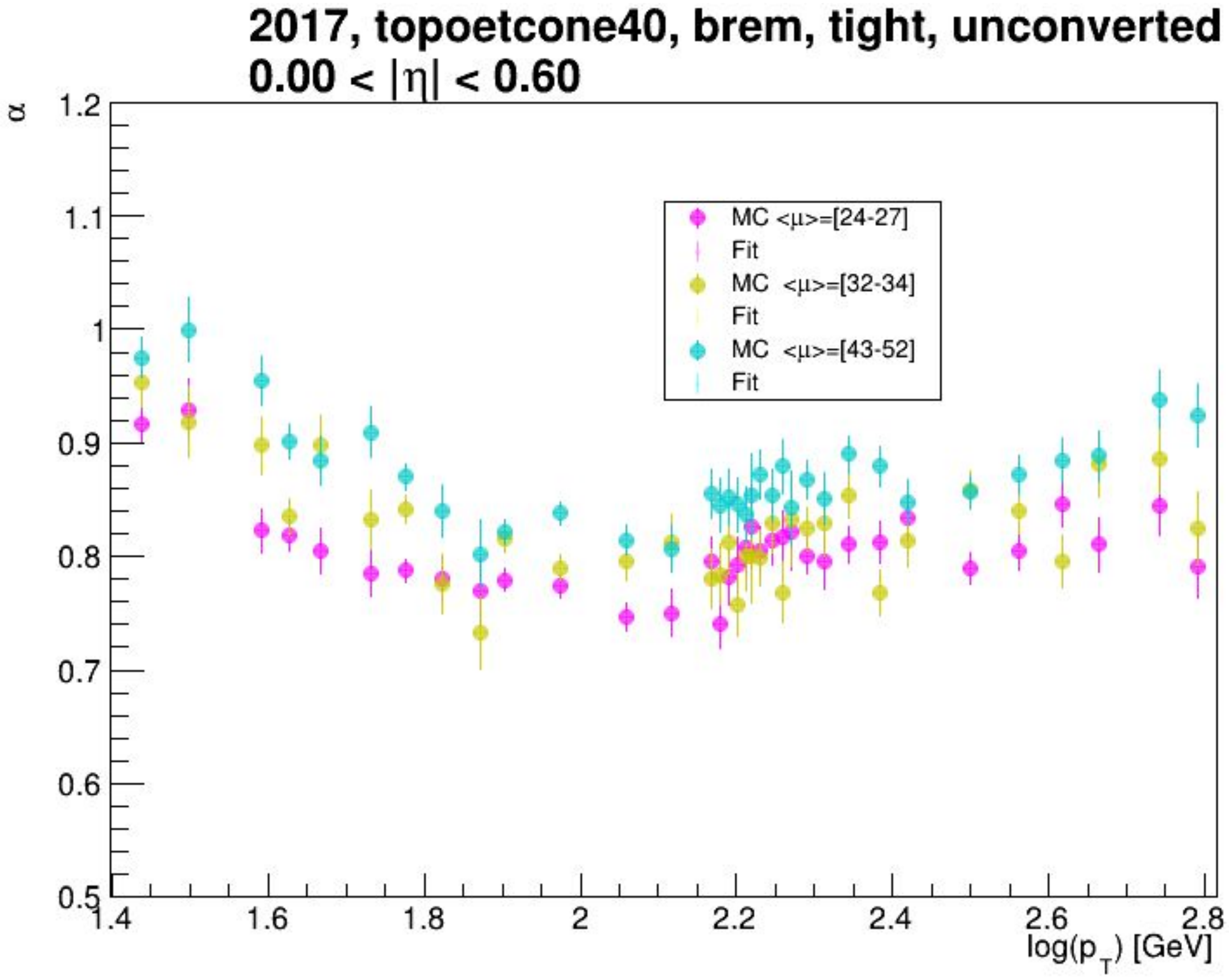
At the level of reconstructed variables (isolation energy), MC and data does not match since simulating calorimeter in MC is really hard.

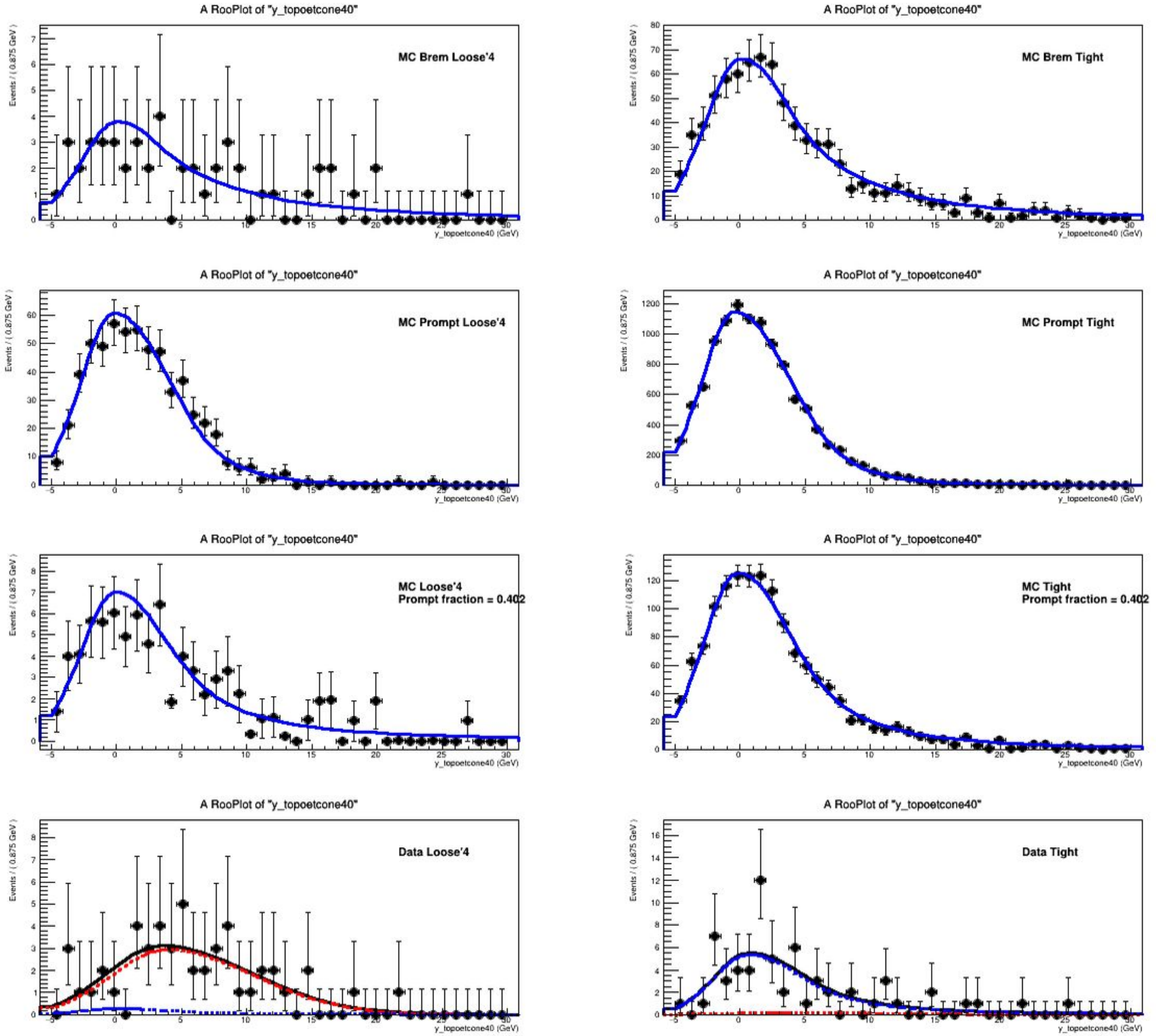
Data Driven Shifts



Conclusions

- In general: Add a few more parameters to the analytical functions to obtain better p-value.
- Specifically: Use thinner bins in the MC Brem tight region to have a better profile of the points trends for alpha distribution.
- Enlarge bins in <mu> for low-statistics bins to reduce large errors in DDshifts.
- Use the PDG procedure to rescale the confidence intervals on the DDshifts.


2017, topoetcone40, brem, tight, unconverted
0.00 < |η| < 0.60
α
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
log(p_T) [GeV]
MC <μ>=[24-27]
Fit
MC <μ>=[32-34]
Fit
MC <μ>=[43-52]
Fit



A RooPlot of "y_topoetcone40"
MC Brem Loose'4
MC Brem Tight
MC Prompt Loose'4
MC Prompt Tight
MC Loose'4
Prompt fraction = 0.402
MC Tight
Prompt fraction = 0.402
Data Loose'4
Data Tight
Events / (0.875 GeV)
y_topoetcone40 (GeV)

http://laconga.redclara.net

contacto@laconga.redclara.net

lacongaphysics

Latin American alliance for Capacity buildiNG in Advanced **physics**

LA-CoNGA **physics**

Cofinanciado por el programa Erasmus+ de la Unión Europea