Improving Jet Resolution with Calorimeter Segmentation with Data

ATLAS

David López Mateos, November 13, 2008

Where we Stand and Long-Term Plans

Event Selection: Truth to Data-Driven

Understanding Low-Response Jets

Event Selection: η region

Numerical Inversion with New Event Selection

Conclusions and Plans

Where We Stand Now and Future Plans

- Proof of Principle is complete: this is a technique that is worth developing

- Now we need a more long-term plan:

1. Redefine event selection so that it works on a data-driven approach
2. Define p_T and η regions (and data sets) for full study
3. Redo pre-calibration steps vs. p_T^{reco}
4. Complete the four-layer approach study
5. Perform a 6-layer study
6. Try to see what can be understood from a principal-component analysis

Where We Stand Now and Future Plans

Today

1. Redefine event selection so that it works on a data-driven approach

   Address problem of low-response (fake jets), define uniform eta region

2. Define p_T and η regions (and data sets) for full study

   20 GeV < p_T < 150 GeV (early data)+one high-p_T bin. MC08 J2 and J3 datasets. For high-p_T bin: J5? What eta regions?

By the end of the year

3. Redo pre-calibration steps vs. p_T^{reco}

   Numerical inversion

4. Complete the four-layer approach study

Early next year

5. Perform a 6-layer study

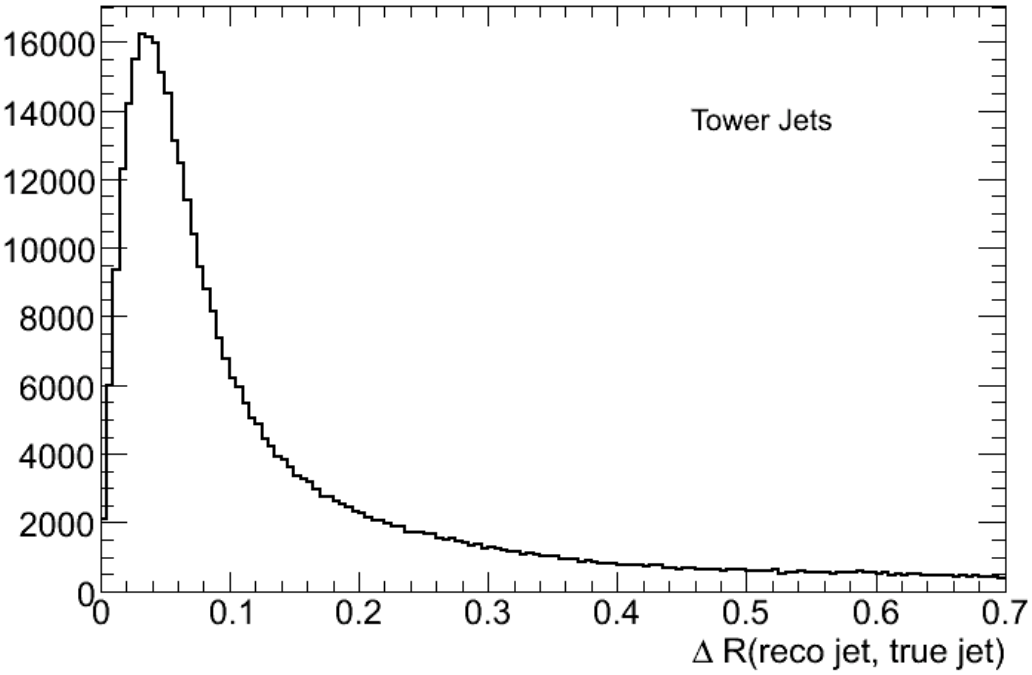
   Use all 6 layers. It could be better, but also more complicated

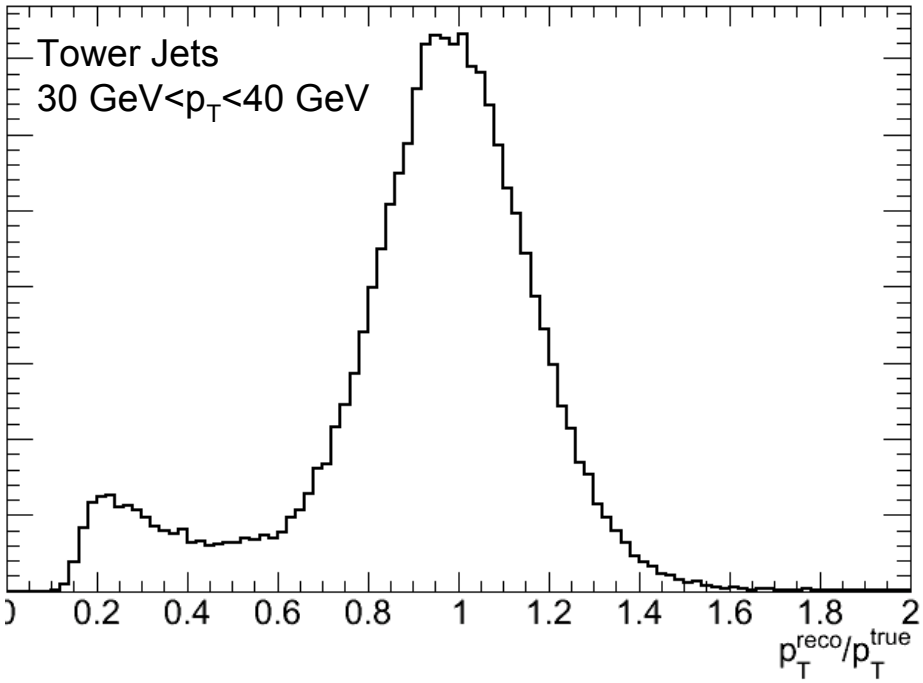
6. Try to see what can be understood from a principal-component analysis

   Most complicated analysis. Might not be very useful with first data, but could also point to physically meaningful quantities

Event Selection: Truth to Data-Driven

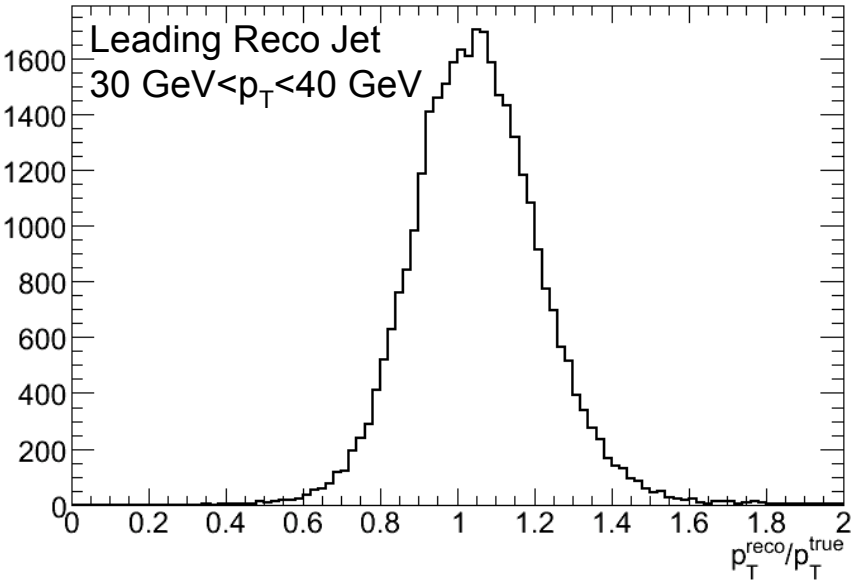
- For the event selection so far we used jets with a truth jet matched within a 0.2 radius
- This throws away ~20% of the jets, but it is necessary not to get low-response jets (probably fakes)

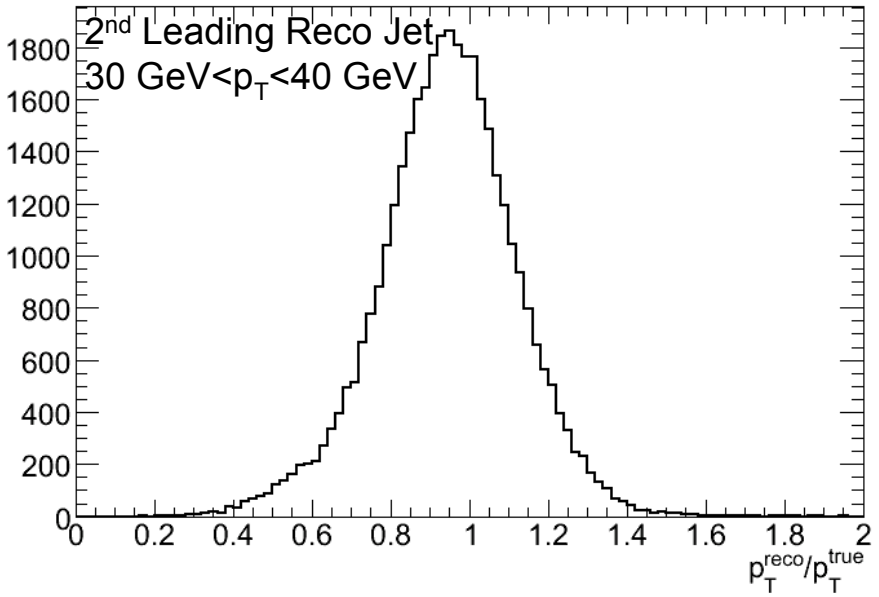


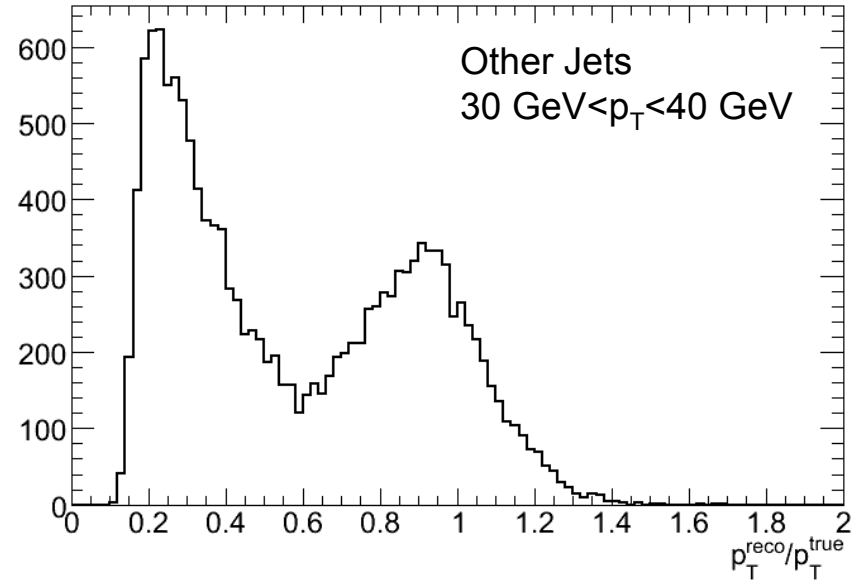


- But we cannot do this with data. So what will be the event selection with data?

Where Are Low-Response Jets Coming From?





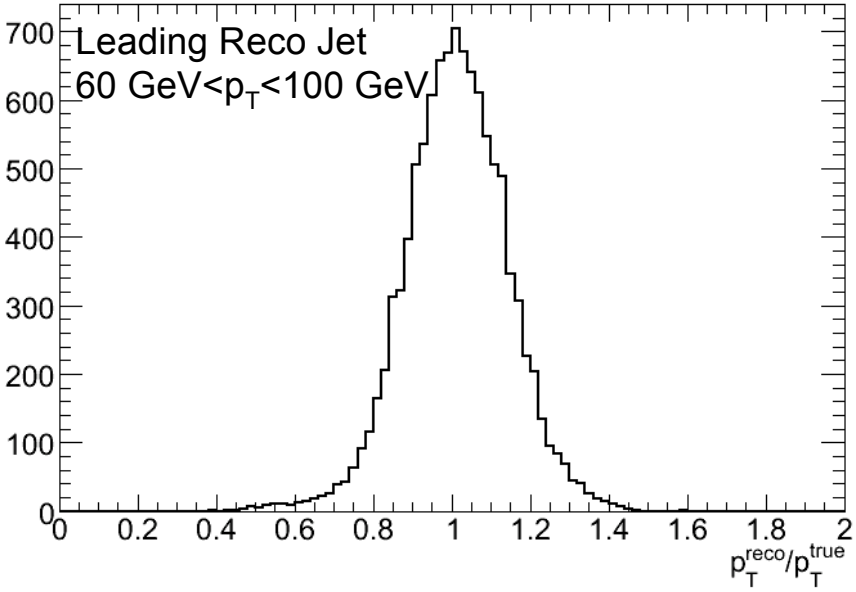


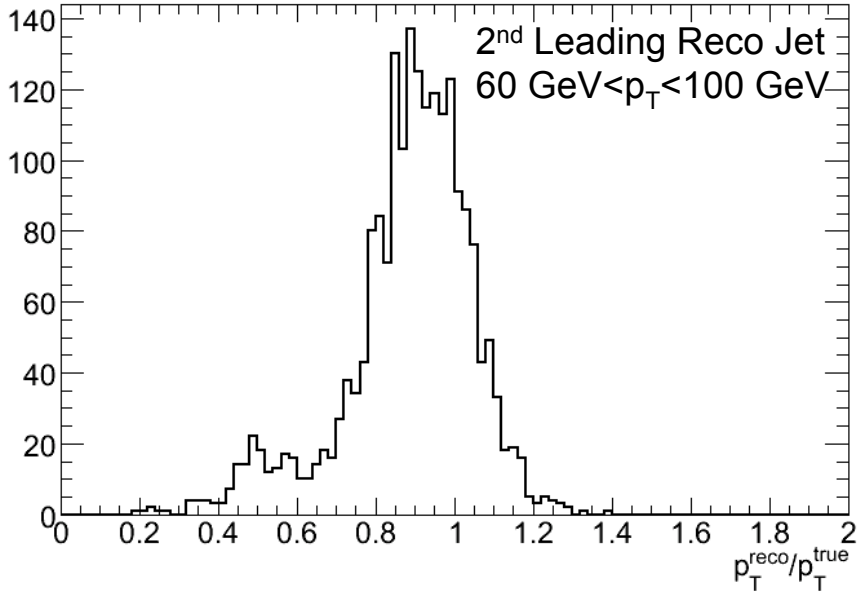
This problem seems to not be present if we choose the leading jet, or the 2nd leading jet

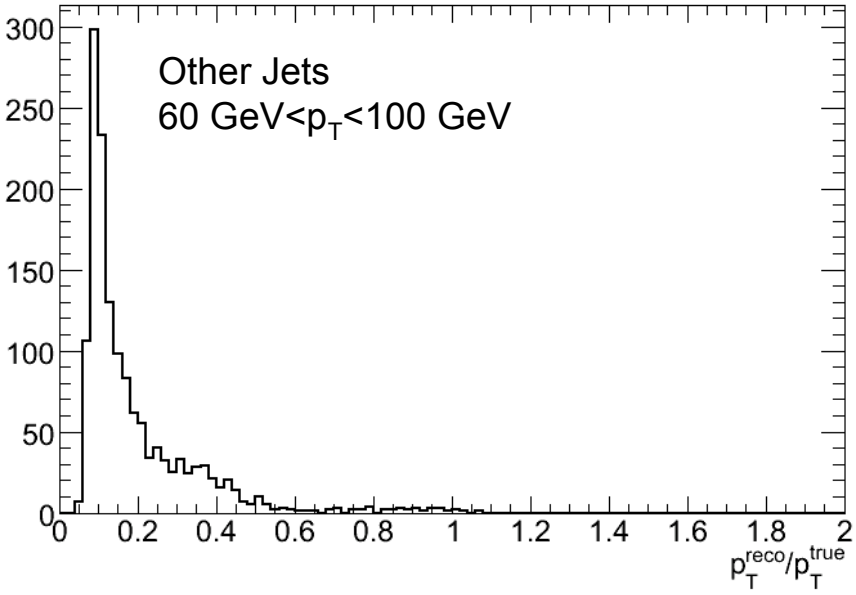
Can this problem disappear completely with a cut on p_T^{reco}?

Otherwise: other jets are ~50% of all jets in the J2 sample

Where Are Low-Response Jets Coming From?





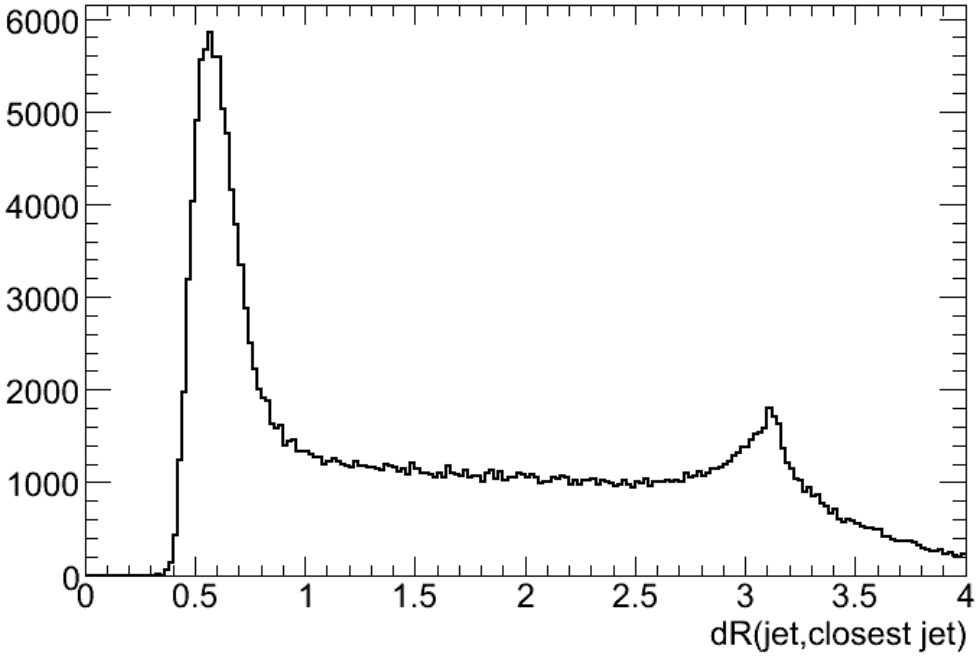


These jets can have a reco p_T of over 20 GeV…

The 2nd leading jet might also be a bit vulnerable to these “fake jets”

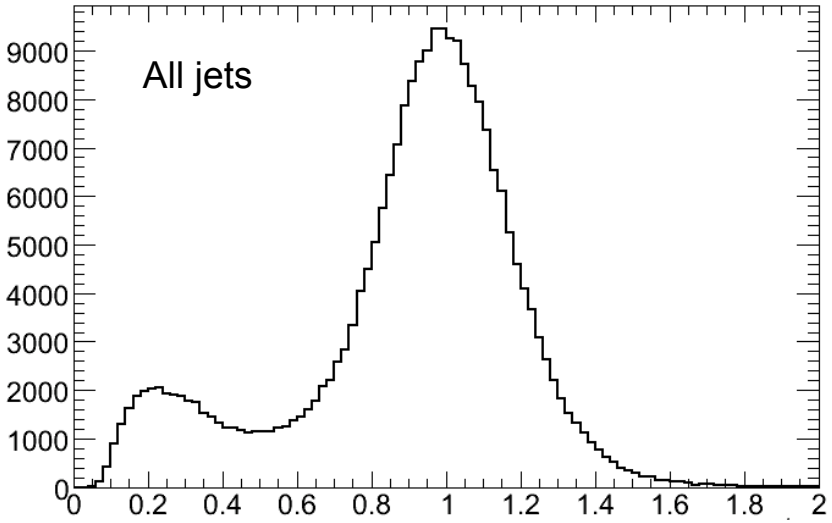
Are These Split-Merge Jets?

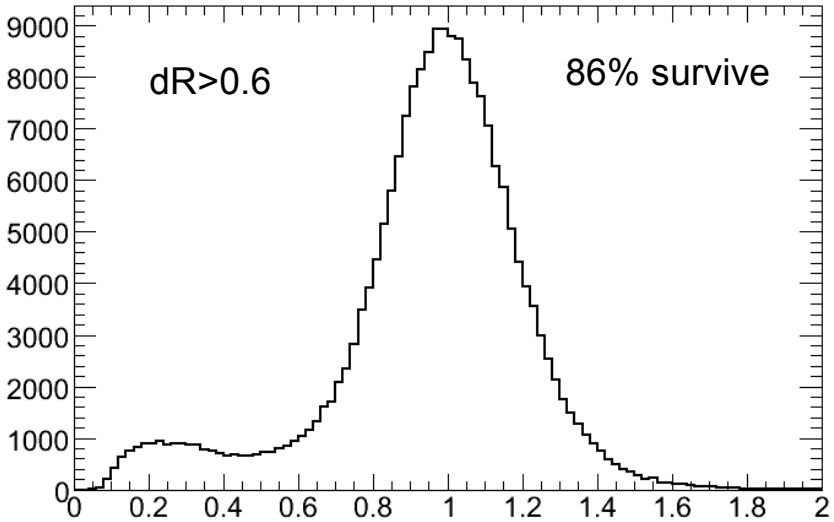
If so, jet isolation cut could help.
Look at jets matched to a truth jet above 20 GeV

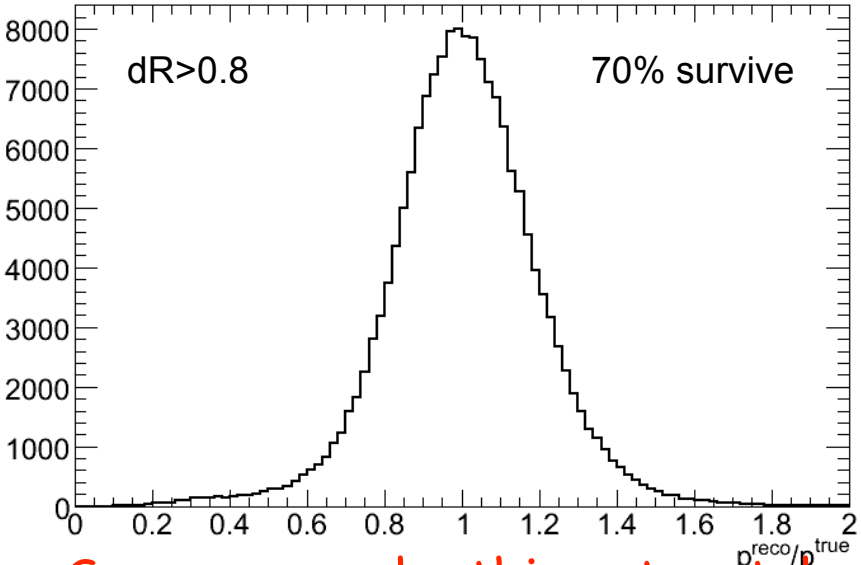


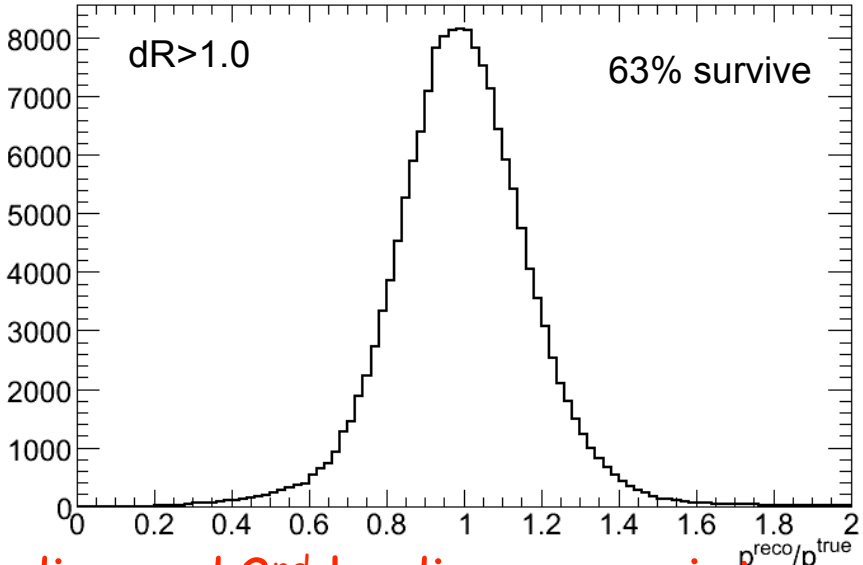
Clearly, very close-by jets can be the problem
Try different cuts, and look at the response, also check what percentage of jets survive

Are These Split-Merge Jets?





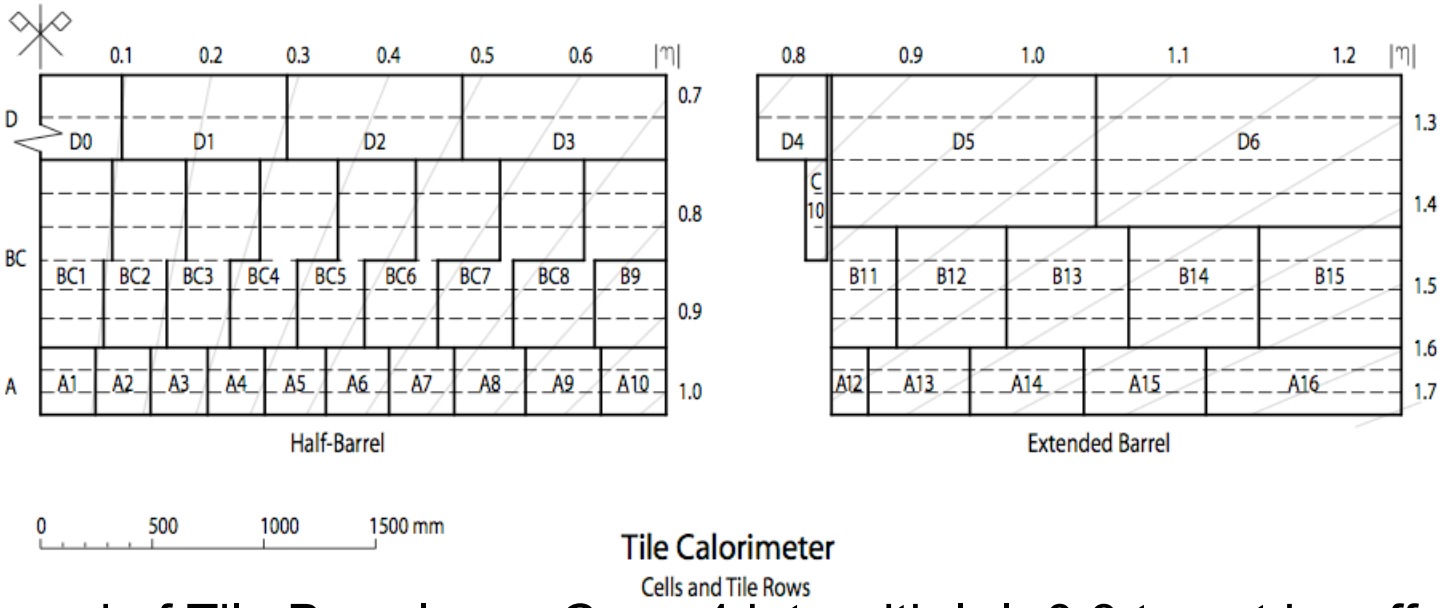




So we can make this cut or take only leading and 2nd leading, or a mixture.
For the rest of today I use dR>1.0 for the event selection

Event Selection: η region

- Eventually, this will be done in small η bins, say 0.1
- But with the Monte Carlo, statistics are not high enough for such fine binning
- We need uniform section of the Calorimeter: particularly important since layer definition changes from one part to another



η~0.7: end of Tile Barrel: use Cone 4 jets with |η|<0.3 to not be affected by this

η~1.1: end of EM Barrel: use Cone 4 jets with |η|<0.7 to not be affected by this

⇒I will look at both, see if I can merge them or not

Numerical Inversion

Reminder: A method to apply a Monte Carlo-derived correction to data

In MC you have: $R(p_T^{\text{true}})$ p_T^{true}

In data you have only: p_T^{reco}

And it is hard to calculate $R(p_T^{\text{reco}})$ in MC because the distribution is not gaussian in bins of p_T^{reco}

$\Rightarrow$So we use $R(p_T^{\text{true}})$ to transform the p_T^{true} axis in p_T^{reco}

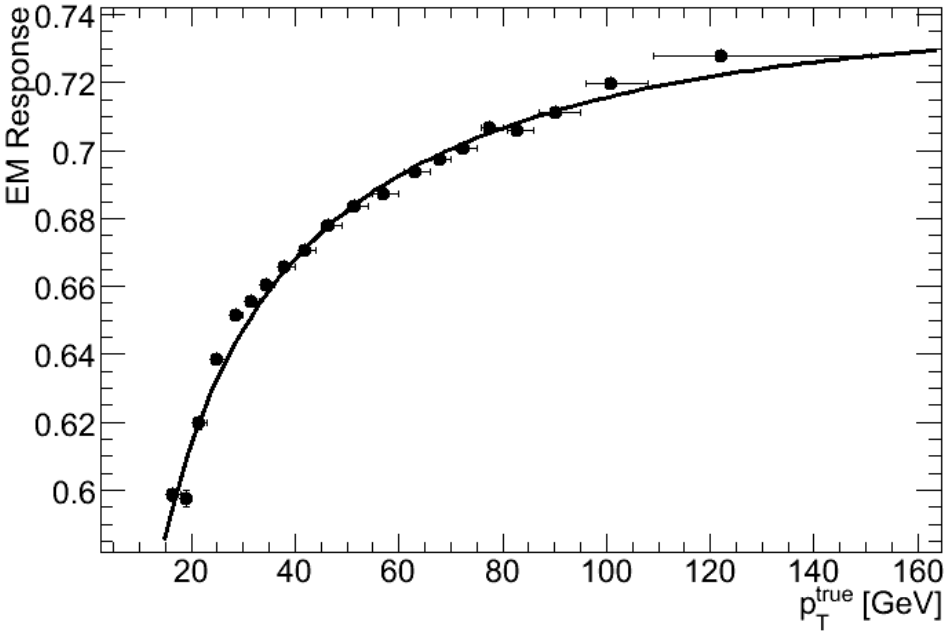
Note: distributions are still gaussian because the events in the different bins are not being reshuffled

Then you can apply the correction as a function of p_T^{reco}

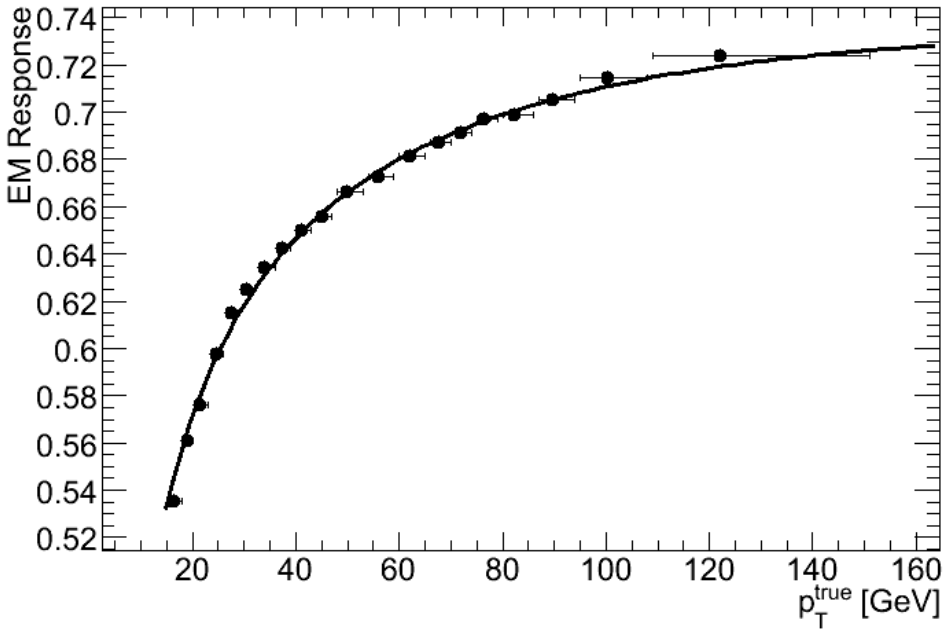
Numerical Inversion: Step 1

Calculation of: $R(p_T^{\text{true}})$

Tower Jets



Topo Jets



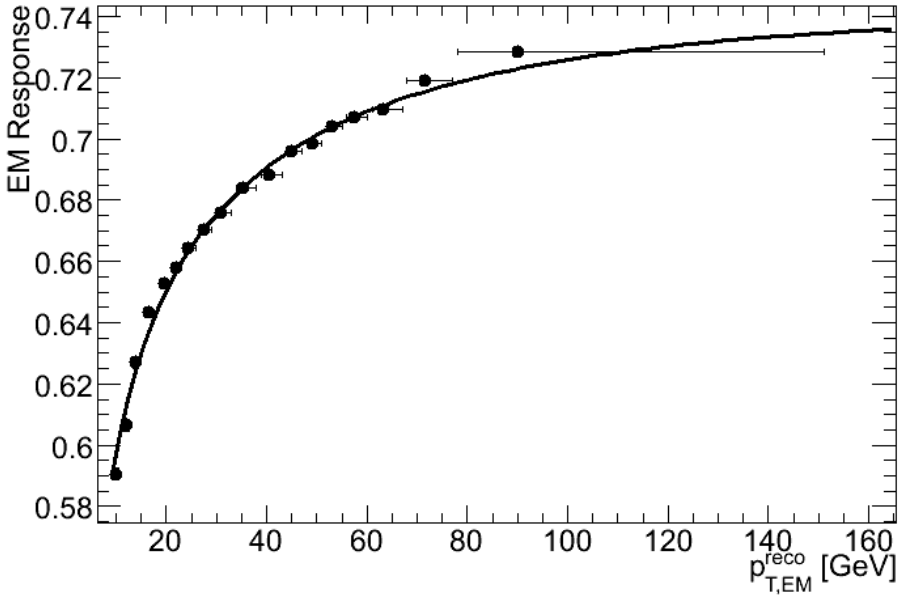
Fits to $a+b*\log(p_T)+c*(\log(p_T))^2$

$|\eta|<0.3$

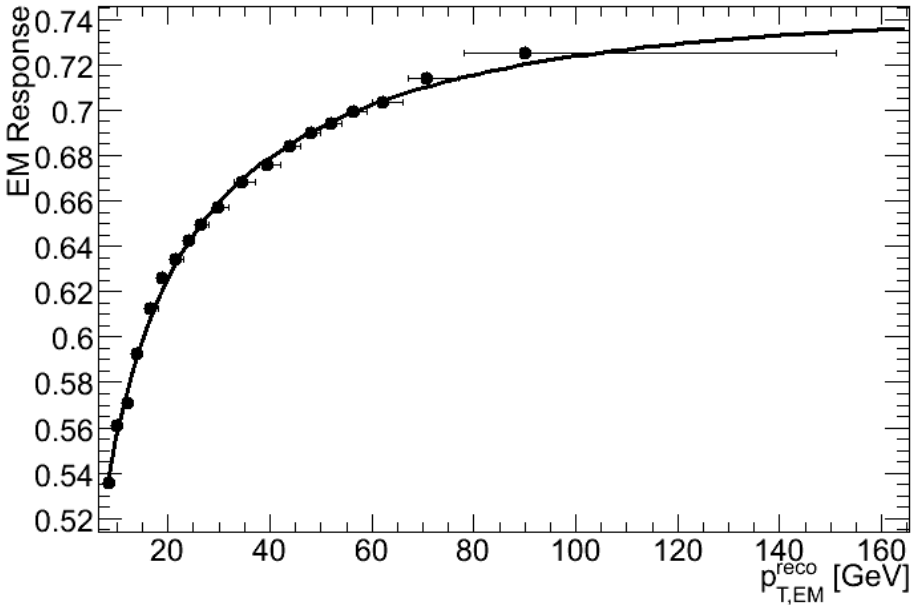
Numerical Inversion: Step 2

Calculation of: $R(p_T^{\mathrm{reco}})$

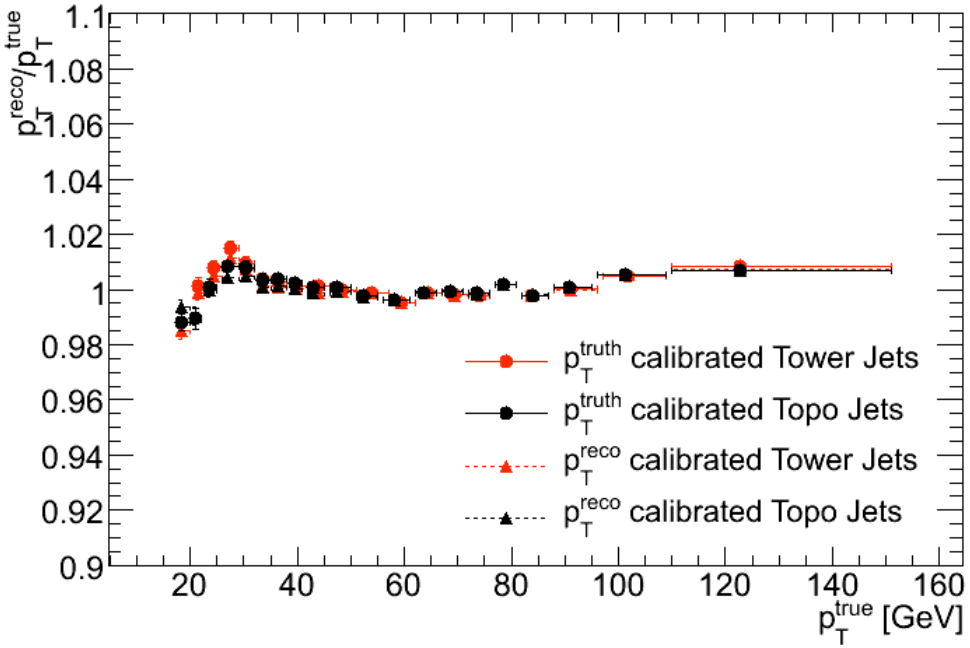
Tower Jets

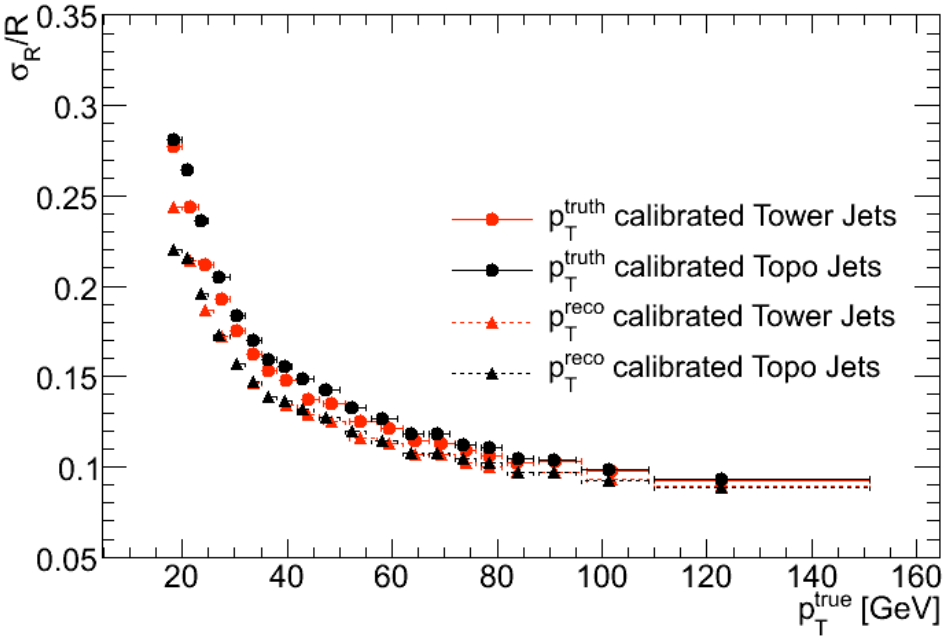


Topo Jets



Numerical Inversion: Results





The response is virtually the same with a truth-correction and a numerical inversion correction

<u>The resolution is much better!</u> (and the differences in resolution between Tower and Topo Jets disappear (!?))

~10% improvement in resolution (recall ~7% from f_3). Compare to total ~20% improvement with H1

Conclusions and Plans

- Now starting to perform study more systematically
- It should be complete with 4 layers before the end of the year
- Event selection is now clearer
- Numerical inversion works very well, even when starting from the EM scale. Amazing improvement in resolution