New jet treatment

Thibault Guillemin,

on behalf of the horns task force

D0-France

May 31, 2011

The starting point

To reach a satisfactory data/MC agreement, because of our ZB overlay procedure, all jets in the simulation can not be treated in a uniform way

- In general, what we want to do is to have a **coherent treatment of zero-bias (ZB) jets and hard-scatter (HS) jets in the simulation**
- We hope at the end:

- to reach a reasonable data-MC agreement for good jets with p_T >15 GeV, also in the ICR

- to be able to believe (to some extent) in the sentence:
"the remaining effects are due to physics mismodellings"

<u>The principle</u>

- Apply the ICD correction to data

- MC:

• identify ZB jets and HS jets

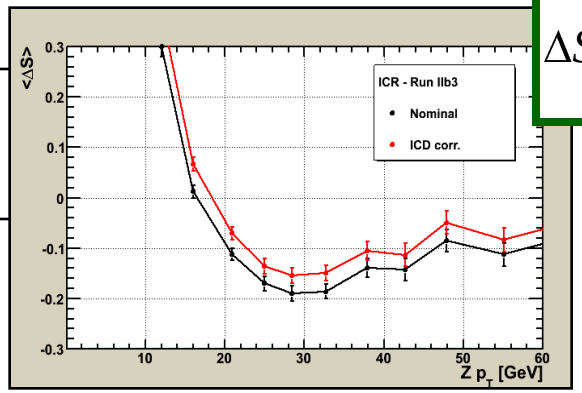
• apply the data JES (including the ICD correction) to ZB jets

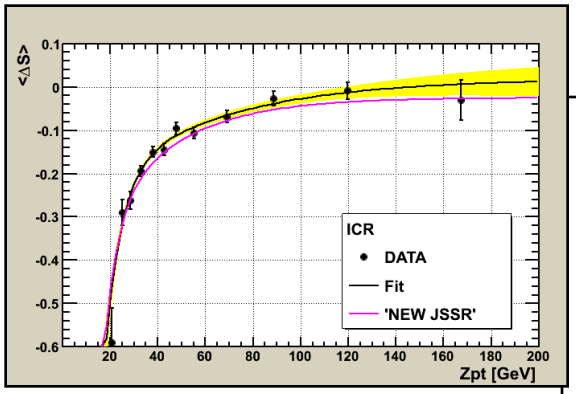
• apply MC JES + JSSR to HS jets

Step by step…

Step 1: is the ΔS variable used for the JSSR parameter determination significantly affected?

$$\Delta S = \frac{p_T^{jet} - p_T^Z}{p_T^Z}$$





Step 2: determine a new set of JSSR parameters

➔ **JSSR combo**:

MC Run IIb1 + MC Run IIb2 vs Run IIb1-2-3 data

+ apply the ICD correction to the data

Step 3: define a method to tag the ZB jets in the simulation

➔ two categories of jet are obtained:

- HS jets
- ZB jets

Step 4: validate the full procedure in analyses

Shifting: is there a bias at low Z p_T?

Question in the air since a long time…

- At low Z p_T, the expected gaussian behaviour of the ΔS distribution is completely distorted by the jet turn-on

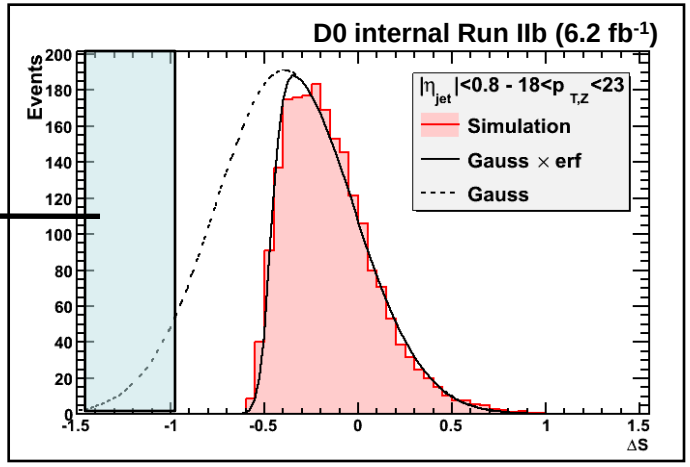
➔ the gaussian mean estimation may become biased

➔ **a lot** of attempts were tried to overcome this difficulty (some simple ones, some elaborate ones) but the "pathology" proved to be very stable…

Only one method changed significantly the picture:

to replace the mean of the gaussian fit by the function average over the physical region $\Delta S>-1$

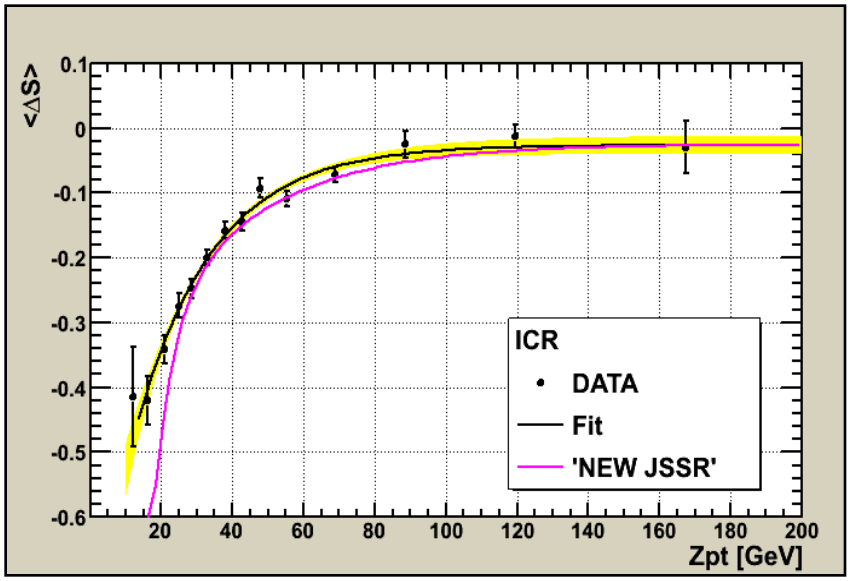
non-physical region

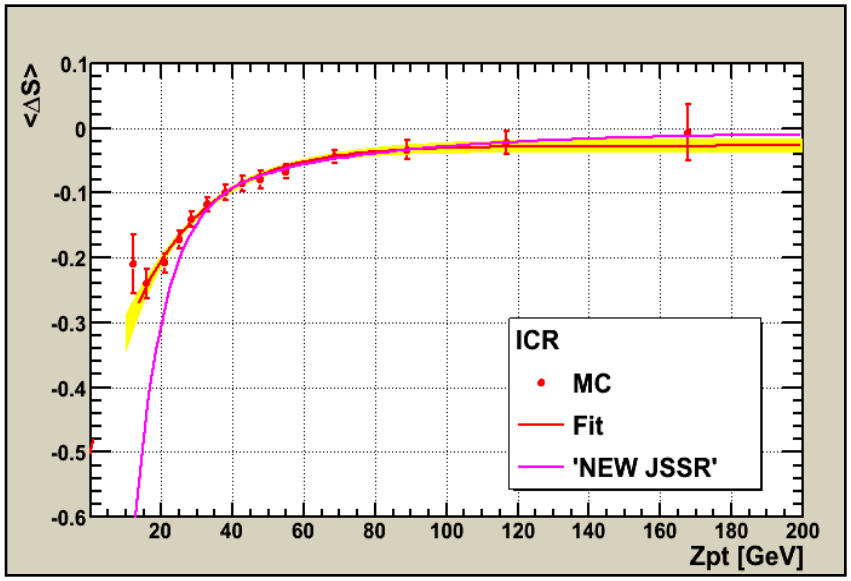


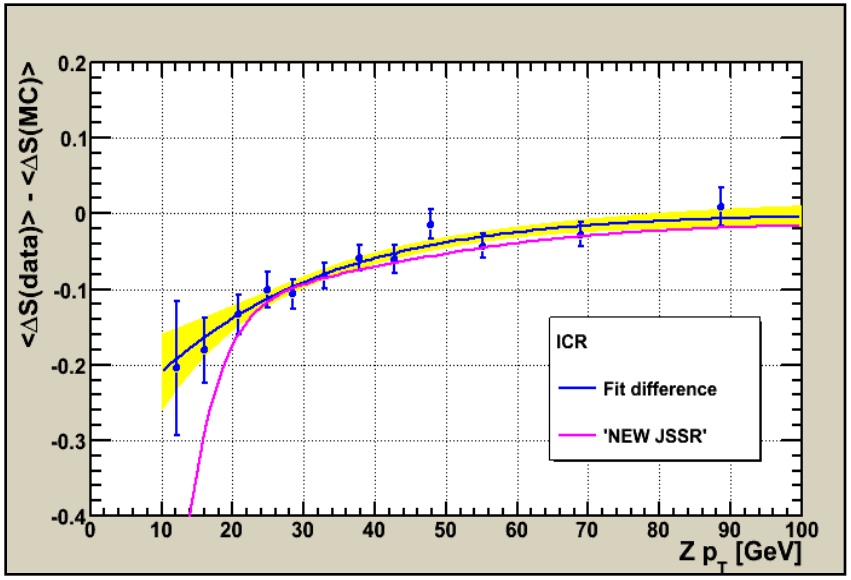
Instead of the gaussian mean, we use the estimator E:

$$E = \frac{\int_{-1}^{+\infty} x f(x)\,dx}{\int_{-1}^{+\infty} f(x)\,dx}$$

ICR shifting results







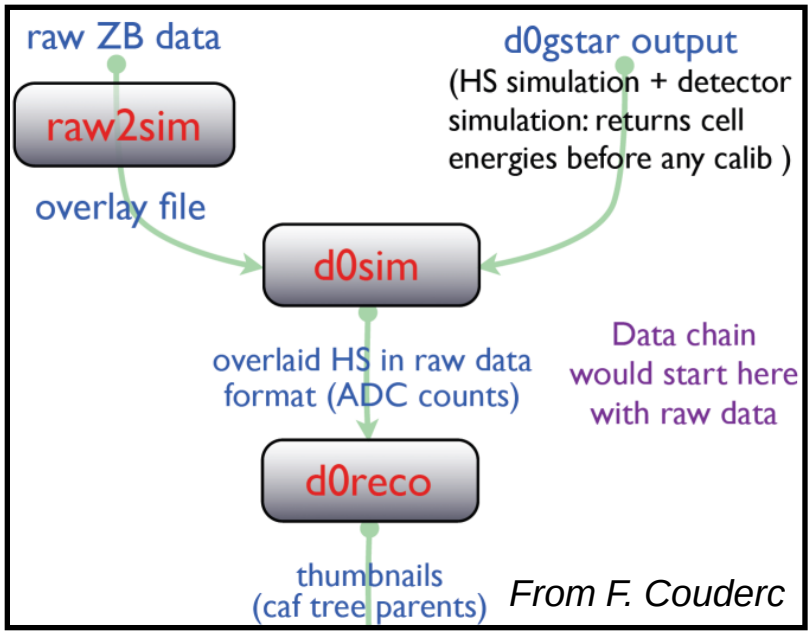
Significant difference w.r.t “NEW” JSSR

- below 25 GeV: change mainly due to the new estimator
- above 25 GeV: change mainly due to the ICD boost

N.B: the difference between the north and the south regions was studied but no significant difference has been observed

How do we identify ZB jets?

Overlay chain



After d0reco, we need to use the true MC particles to identify ZB jets

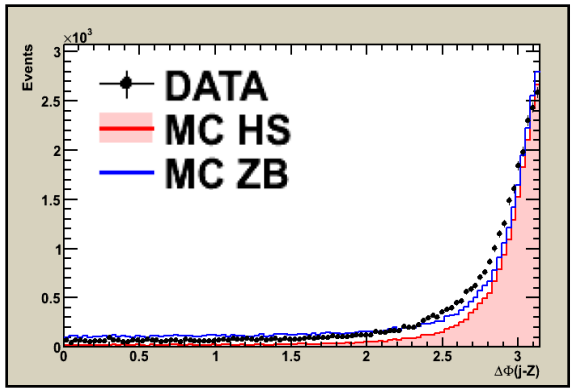
➔ the variable used is the following:

$$\frac{\sum_{\textit{true particles in } 0.5} p_T}{\textit{raw jet } p_T}$$

How does this quantity look like?

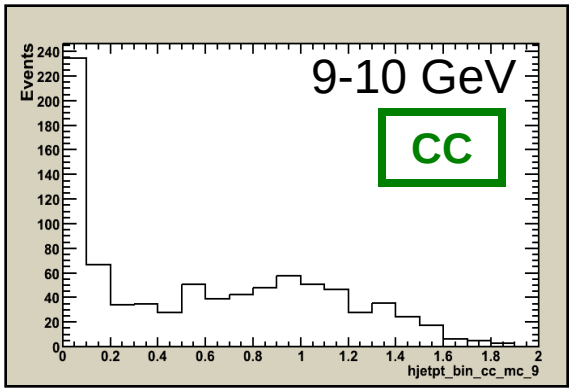
In the next slide, look at two phase space regions in *Z*+1 jet events:

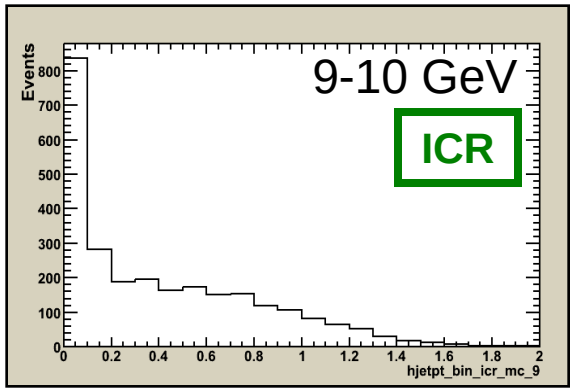
- $\Delta\Phi$(Z-jet)<π/2 (dominated by ZB jets)
- $\Delta\Phi$(Z-jet)>2.8 (dominated by HS jets)

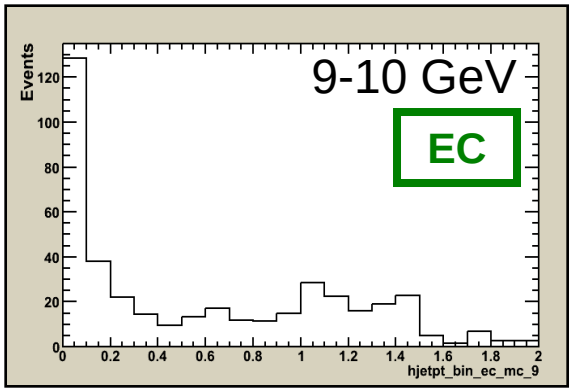


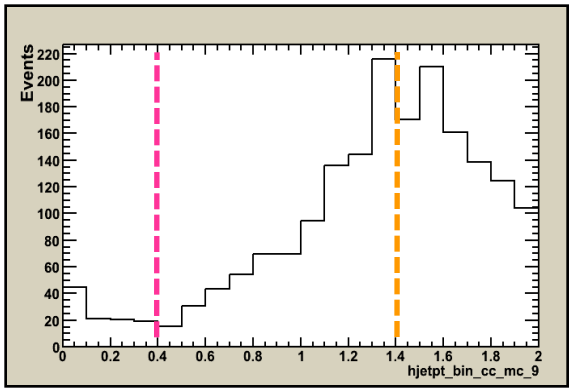
ZB jet tagging: single raw p_T bin distributions

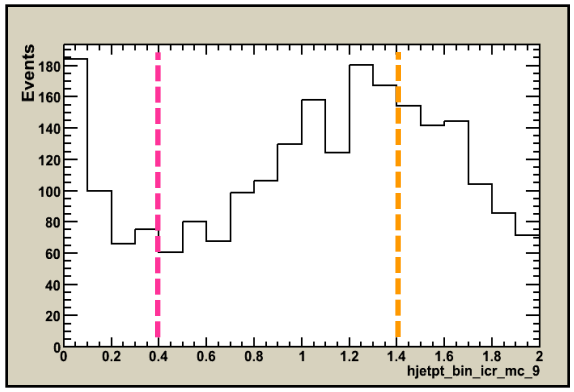
$\Delta\Phi<\pi/2$

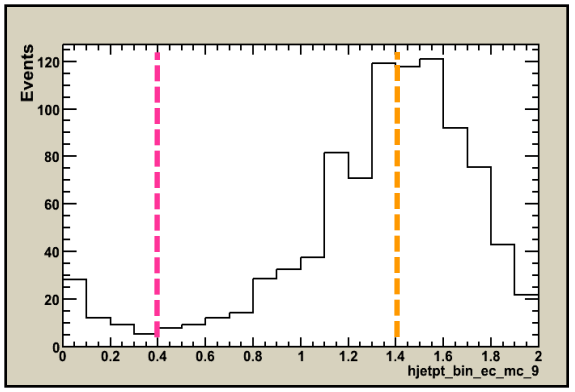












$\Delta\Phi>2.8$

Summary for the ZB jet tagging

- Jets with no true particles in the cone are pure ZB jets
- A cut at 0.4 seems reasonable to define HS jets
- The response for HS jets is peaking in all regions at 1.4
- **We have a large fraction of "mixed" jets between 0 and 0.4 (in particular in the ICR):**

these are jets whose energy is coming from both the hard scatter (or the underlying event) and the overlay.

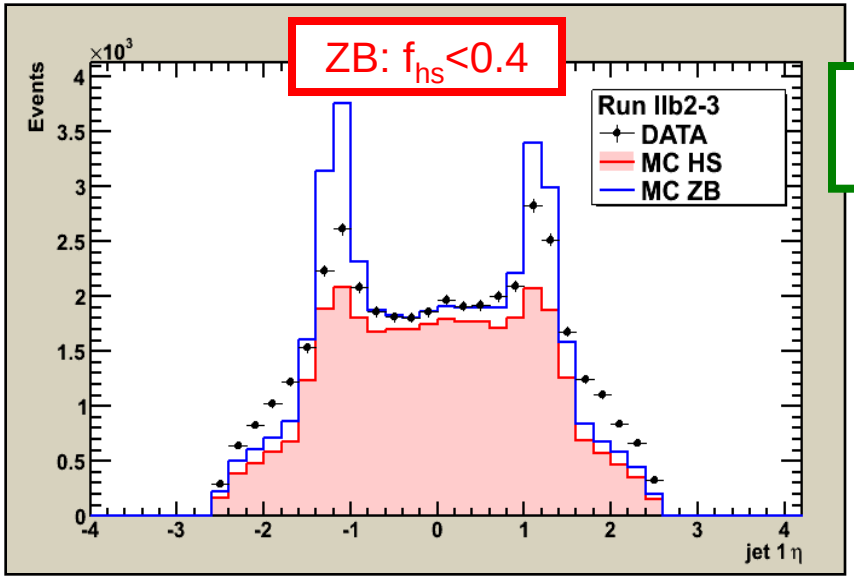
If we apply the data JES to these jets, then their energy will be overestimated

➔ they should be calibrated in a specific way

Ad hoc calibration for mixed jets: see backup slides

Test of the full prescription

Look at 15 GeV good jets to be most sensitive to the problem



jet 1 CC normalization

Clear excess of ZB jets in the ICR

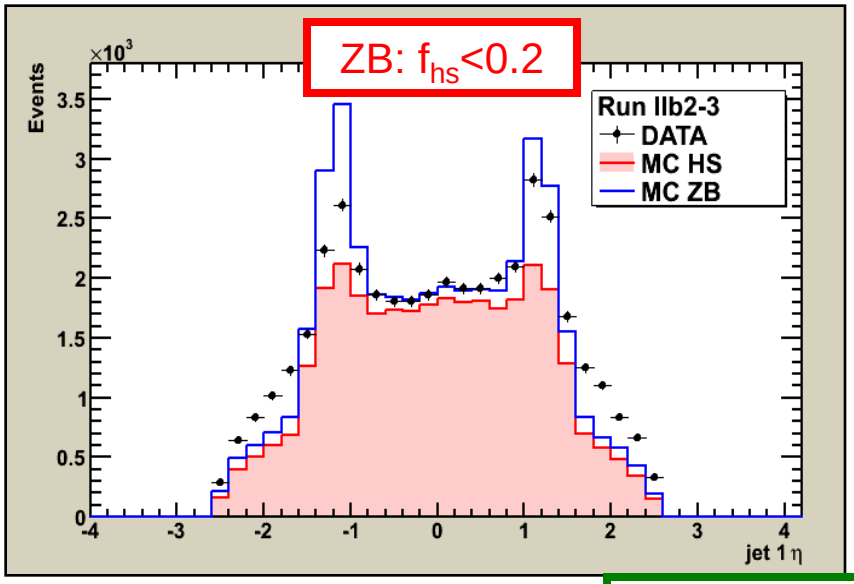
(the ZB contribution is still bigger if we apply the data JES to the mixed jets)

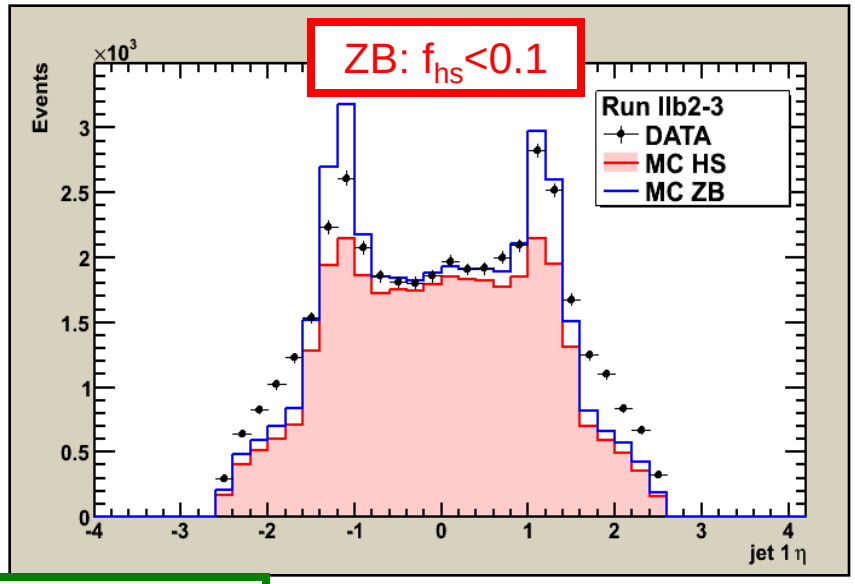
N.B: the ZB histogram corresponds to the ZB and the mixed jets (but the 2 types of jets are treated differently)

We need to reduce in some way the overestimated number of mixed jets in the simulation.

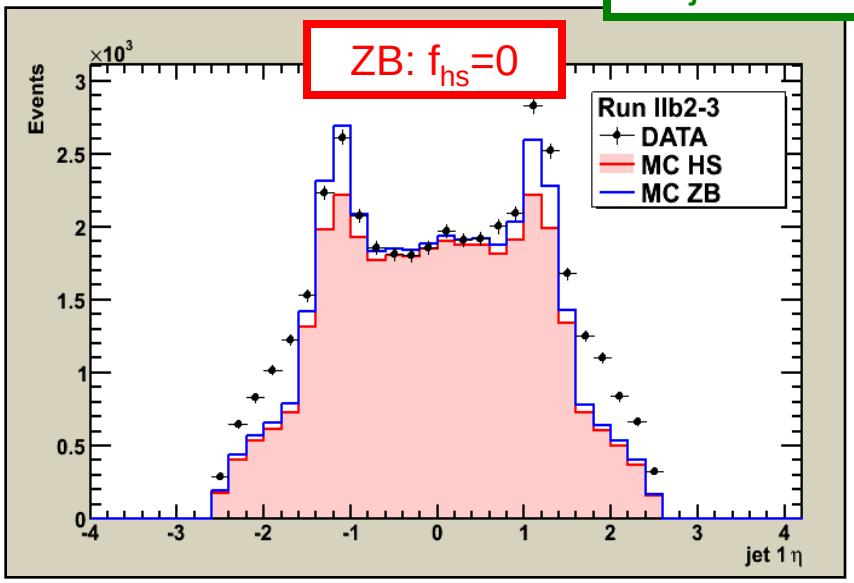
But which room for manoeuvre do we have?

Using various f_{HS} thresholds to remove mixed jets





jet 1 CC normalization

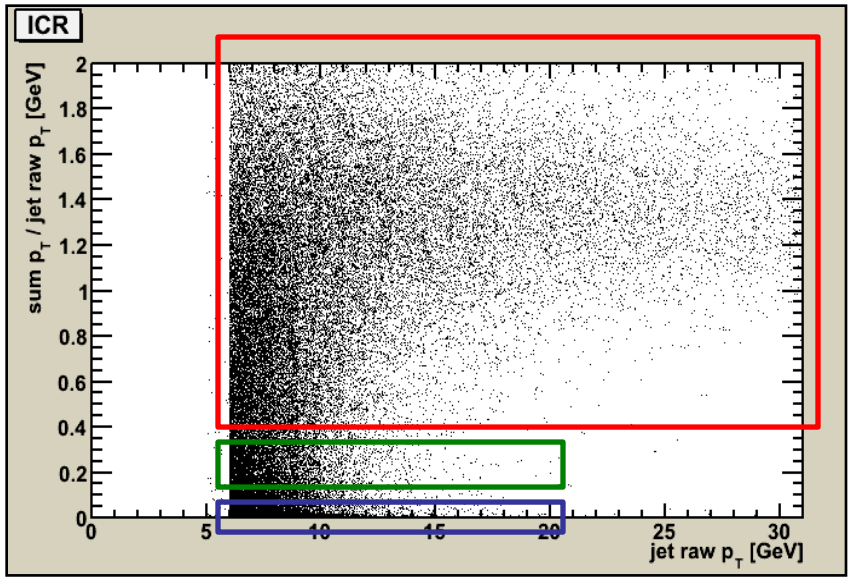


Let's try to find the best ad hoc procedure by varying the f_{HS} threshold

➔ no choice!

All the mixed jets with $0<f_{HS}<0.4$ have to be removed: we have to kill the promotion effect for $0<f_{HS}<0.4$

New jet treatment



HS jets ➔ JSSR-combo

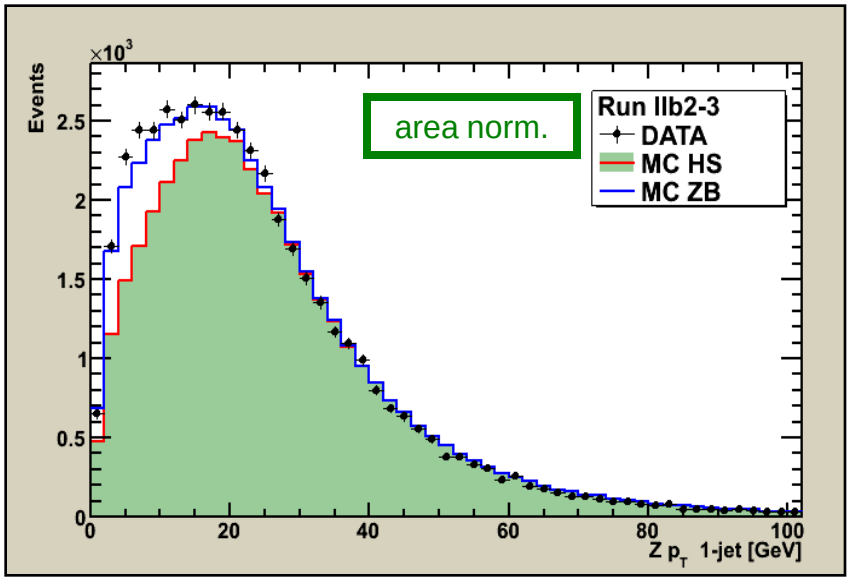
Mixed jets ➔ removed

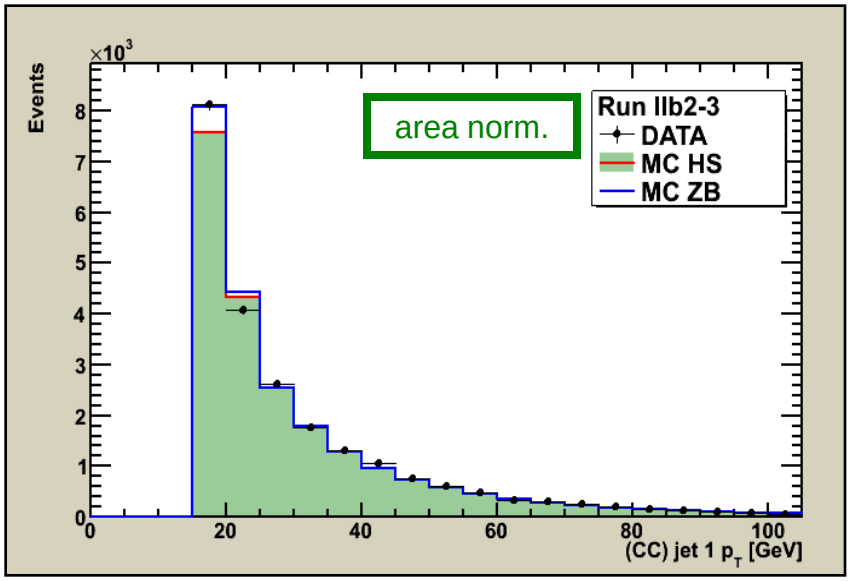
Pure ZB jets ➔ data JES

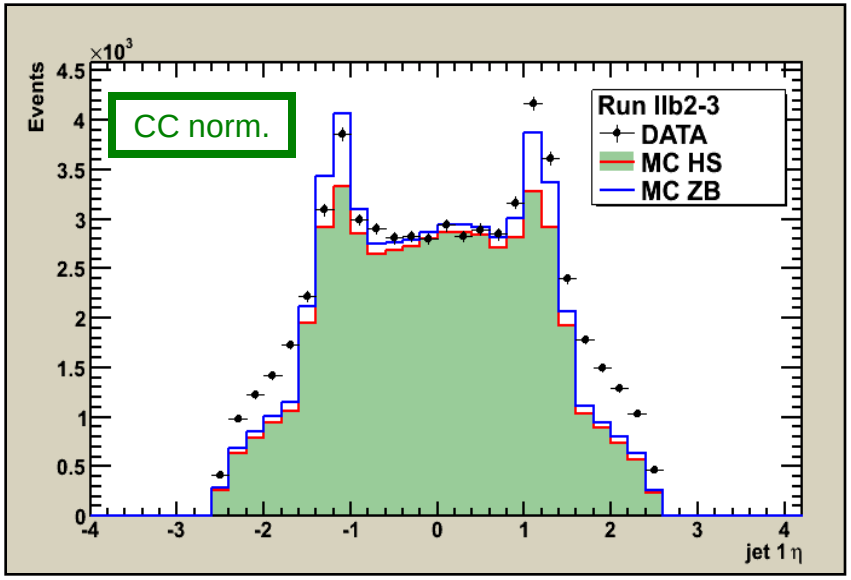
This prescription has been successfully tested for the standard jet configurations in $Z \rightarrow ee$ and in $Z \rightarrow \mu\mu$ events:

- good jet 15 GeV
- good jet 20 GeV
- VC jet 15 GeV
- VC jet 20 GeV

≥ 1 jet (good, 15 GeV)







- The size of the horns is correctly described w.r.t. the central region
- We see the usual need for an η reweighting in the EC

Distributions for 2-jet events: see backup slides

Conclusion

- The phenomenon of jet promotion (jets receiving energy from both the overlay and the hard scatter and thus passing above the reconstruction threshold) is strongly overestimated in the simulation.

To handle this effect in the best way we can at the post d0reco level, the ad hoc procedure we propose is to **remove all mixed jets with a low hard-scatter energy fraction.**

This new jet treatment defines two jet categories:

• **ZB jets** (f_{HS}=0): data JES (with ICD correction), no jet ID correction, not taken into account in the VC/taggability weight computation

• **HS jets** (f_{HS}≥0.4): MC JES, JSSR, jet ID correction, taken into account in the VC/taggability weight computation

- A **full set of JSSR parameters for 6.2 fb^{-1} of Run IIb data with the ICD correction applied** has been determined

➔ new procedure to get a better extraction of the shifting parameters at low p_T

- Everything is implemented and will be used in many analyses for the Summer results

Backup

Ad hoc treatment for mixed jets

A hard scatter fraction per jet can be defined as:

$$f_{HS} = \frac{\sum_{true\ particles\ in\ 0.5} p_T \ / raw\, p_T}{1.4}$$

The energy correction is then computed according to this fraction (only for jets with f_{HS}<0.4):

$$k_{mixed} = (1 - f_{HS}) \times k_{DATA} + f_{HS} \times k_{SSR}$$

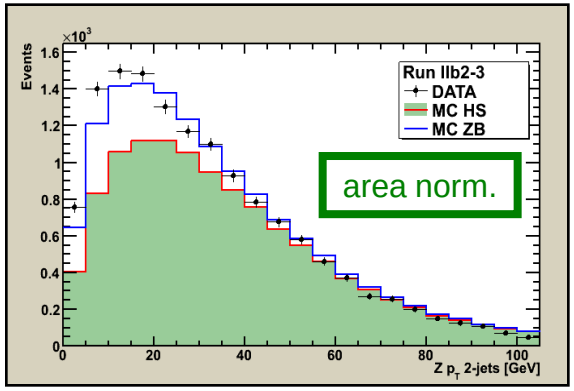
k_{data} includes the ICD correction

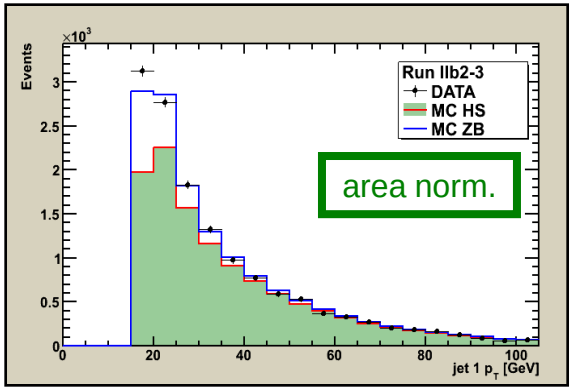
k_{data} and k_{ssr} are estimated at the jet raw energy. In principle, they should be estimated at $(1-f_{HS})\times E_{raw}$ and $f_{HS}\times E_{raw}$, which can be as low as 2 and 4 GeV for a jet near the threshold…

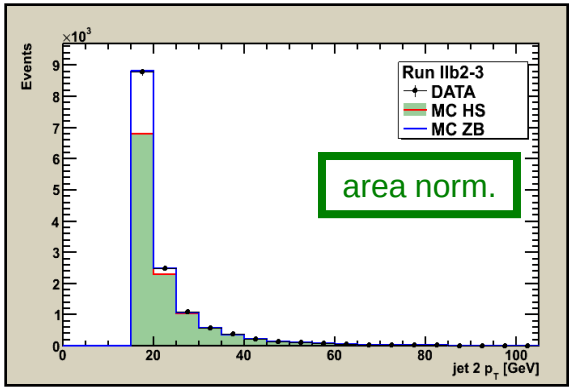
Now we have 3 types of jets in the MC:

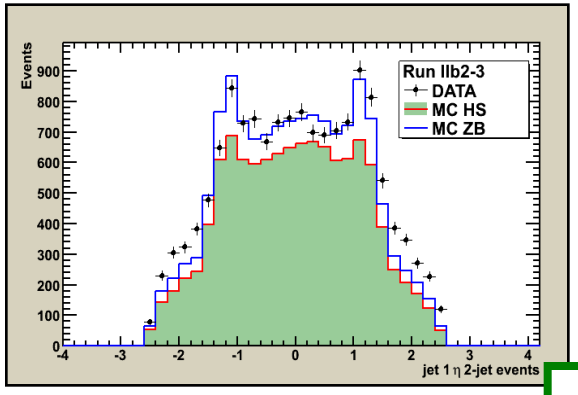
- ZB: $\sum p_{T,true} = 0$ → **k_{data}**
- mixed: $0 < \sum p_{T,true} / p_{T,raw} < 0.4$ → **k_{mixed}**
- HS: $\sum p_{T,true} / p_{T,raw} > 0.4$ → **k_{SSR}**

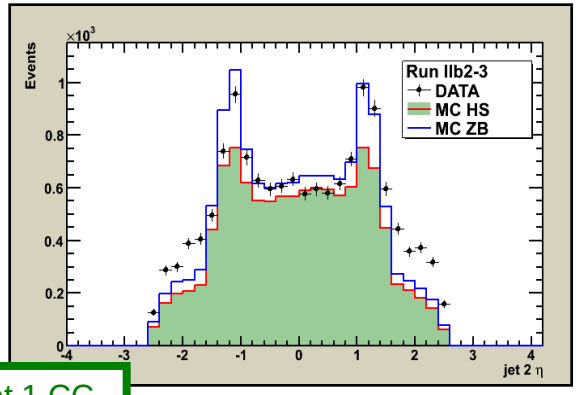
≥ 2 jets (good, 15 GeV)











jet 1 CC norm.

The size of the horns is correctly described w.r.t. the central region