Z(→e^+e^-)+jets analysis: Status, plans, & validation of 3.1.X samples

Bryan Dahmes[1], Christos Lazaridis[2], Kalanand Mishra[3], Lovedeep Saini[4]

[1]Minnesota, [2]Wisconsin, [3]Fermilab, [4]Panjab

V+Jets, August 7, 2009

Overview

- What we learned from Summer08 (2.2.X) analysis ?
- Consolidation on $Z \rightarrow e^+e^-$ reconstruction
- Background subtraction
- Logistics, systematics, action items ....
- Overall plan for 3.1.X
- Status of Summer09 production
- Validation of Summer09 $Z(\rightarrow e^+e^-)$ pre-production samples.

What we learned from Summer08 (2.2.X) analysis

[V+jets, Ewk-electron W/Z inclusive, & JetMET: Zee+jet]

✦ $Z \rightarrow e^+e^-$ reconstruction:

- Can use “standard” electron selection, id, isolation ....
- Efficiency: The Z+jet reconstruction efficiency depends on jet p_T. A *tag-and-probe* tool developed. Also, normalize to $Z \rightarrow e^+e^-$ xsec.

✦ Jets

- Jets in the endcaps: Can use them.
- Efficiency: Not an issue (~100 %), but there is a caveat !
- Resolution: Need to apply unsmearing. Presently available from MC, will obtain from data eventually (using γ/Z+jet events)
- JEC: available from MC, but data-driven correction obtained
- Id: needed to remove fake jets
- Jet p_T cut: need to apply $p_T > 25$ GeV/c. (explained later)

✦ Residual background subtraction works well

- Small background (< 10%). ABCD method works well (@% level).
- Can increase yield by selecting $|m_{ee} - 91.2| < 20$ GeV/c^2.

Converging on baseline event selection

✓ **Use standard *Egamma* POG electron Id**
✓ **Use standard EWK electron isolation**

✦ $Z \to e^+e^-$ reconstruction

- $|m_{ee} - M_Z| < 20$ GeV/c^2
- Electrons
  - super cluster matched to track (*gsfElectrons*)
  - $p_T > 20$ GeV/c
  - within ECAL and tracker acceptance: $|\eta| < 1.4442$ (barrel) OR $1.56 < |\eta| < 2.5$ (endcaps)
  - loose isolation
  - "Robust Loose" electron id (or, just $\Delta\eta$, $\sigma_{i\eta i\eta}$ part of it)

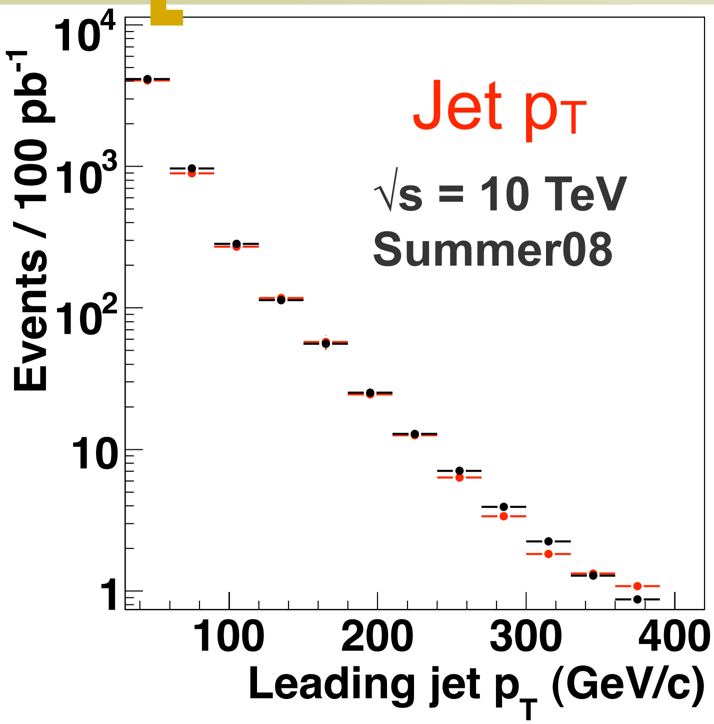
EWK-09-004

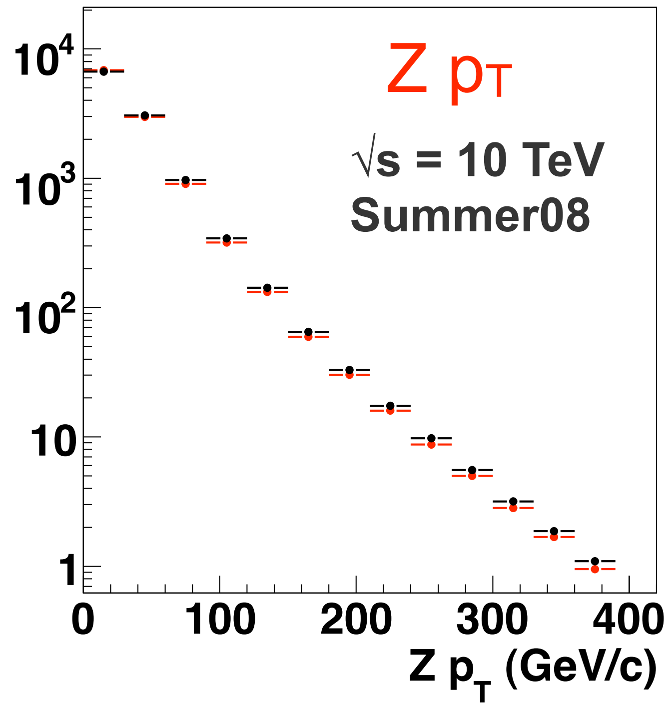
✦ Jets (CaloJets)

- in the barrel or endcaps: $|\eta_{Jet}| < 3$
- $p_T > 25$ GeV/c (on *corrected* CaloJets)

Currently using *iterativeCone5CaloJets*, but JetMET encourages using *SISCone* → we will move to latter

How much data we need for a 1st pass analysis





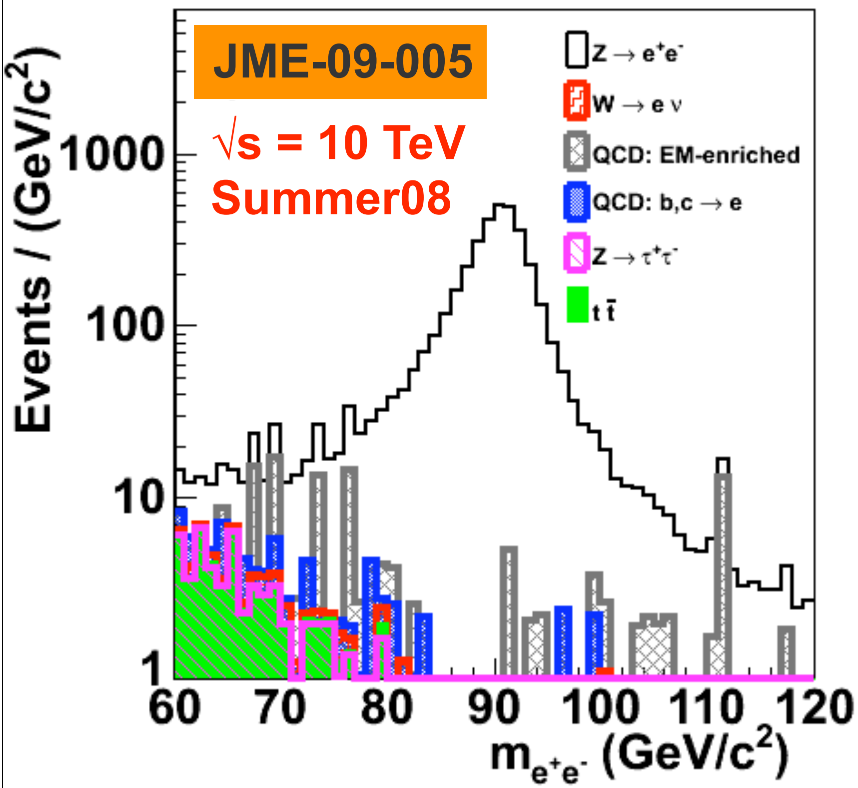
At $\sqrt{s}$ = 10 TeV,
$\int L\,dt$ = 100 pb^{-1}

- ✦data up to p_T = 300 GeV/c.
- ✦Total events ≈ 5000.

What will change at $\sqrt{s}$ = 7 TeV ?

- •The inclusive W/Z cross-section will go down by a factor ≈ 2.
- •We can expect ≈ 2000−2500 Zee+jet(s) events.
- •For an early measurement of this cross section we will only need ≈ 1000 events (our p_T reach will then be limited to about 200 GeV/c).
- •Remember that even with 50 pb^{-1}, the measurement will be systematics limited !!!

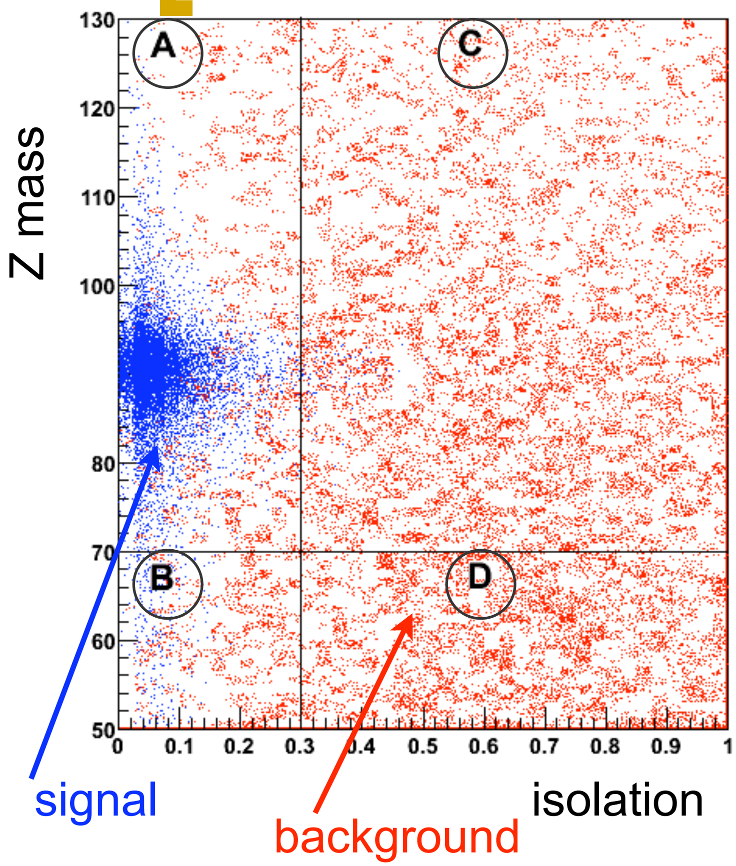
Background subtraction



- ✦Monte Carlo predicts *O*(few %) background contamination.
- ✦Somewhat dependent on jet multiplicity.
- ✦Background contribution is small
  - •if we make, say, 20% error in background estimation, it will still be just ~2% additional systematic uncertainty
    - ➡Not a big deal as we'll see shortly
  - •This means that we should not spend too much time on backgrounds
- ✦The ABCD method seems to be working fine for our purpose
  - •estimates signal with < 1% bias
  - •will benefit from the inclusive W/Z analysis experience

A simple 1-dimensional background subtraction is not optimal, because:
1. Z/γ* lineshape is not modeled easily (ISR, resolution, Z−γ interference, ...)
2. Background shape from Monte Carlo not very reliable

Background subtraction: ABCD method



8 variables: #signal events in A,B,C,D and #background events in A,B,C,D.

And **8** equations to solve:

$$N_A^{BG} = \frac{N_B^{BG} N_C^{BG}}{N_D^{BG}}$$ (assumption: the two variables are uncorrelated for backgd)

$$N_X^{OBS} = N_X^{BG} + N_X^{SIG}$$ (universal truth)

$$f_{XA} = \frac{N_X^{SIG}}{N_A^{SIG}}$$ (Obtained by tightening the cuts in data ➙ an assumption involved)

Studied performed so far in MC (where we can directly compare with MC truth) shows a bias at <1% level in signal estimation in region A.

What has not been done yet ?

The method has been tried and tested with isolation vs Z mass ➙ these were found uncorrelated for background. The same two variables may not be uncorrelated in data. We need to try different pair of variables.

Factors affecting the Z+jet(s) cross-section measurement

Physics Object

- Close coordination with inclusive $Z \rightarrow e^+e^-$ analysis
- Jet energy scale uncertainty
- Jet Id, resolution smearing efficiency
- Focus on data-driven systematic uncertainty

Logistics

- Skimming and processing of primary dataset
- Machinery to go through the whole analysis chain

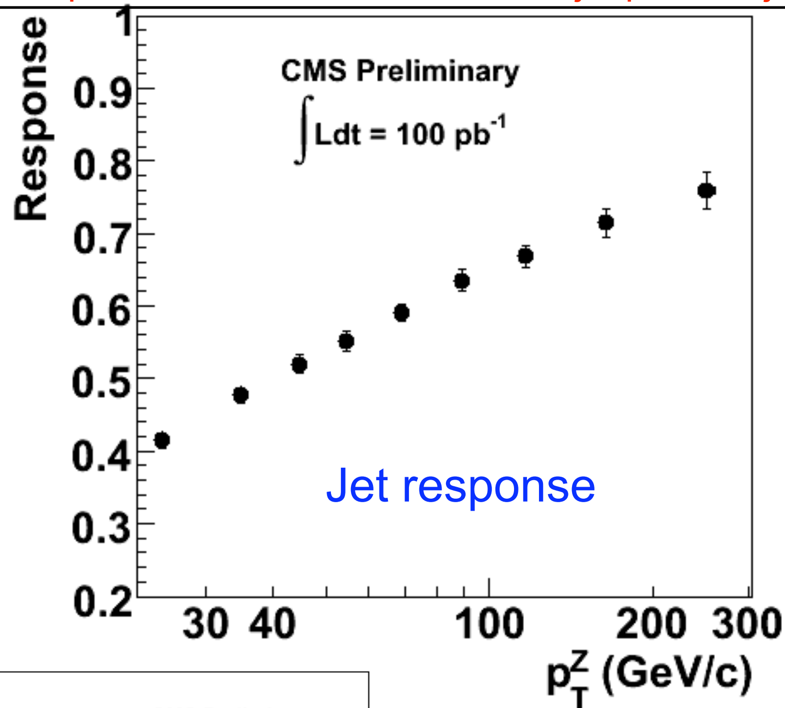
Comparison with theory

- Neglected so far
- Need strategy: share expertise/resources with QCD/EWK
- This part of the analysis can proceed w/o real data

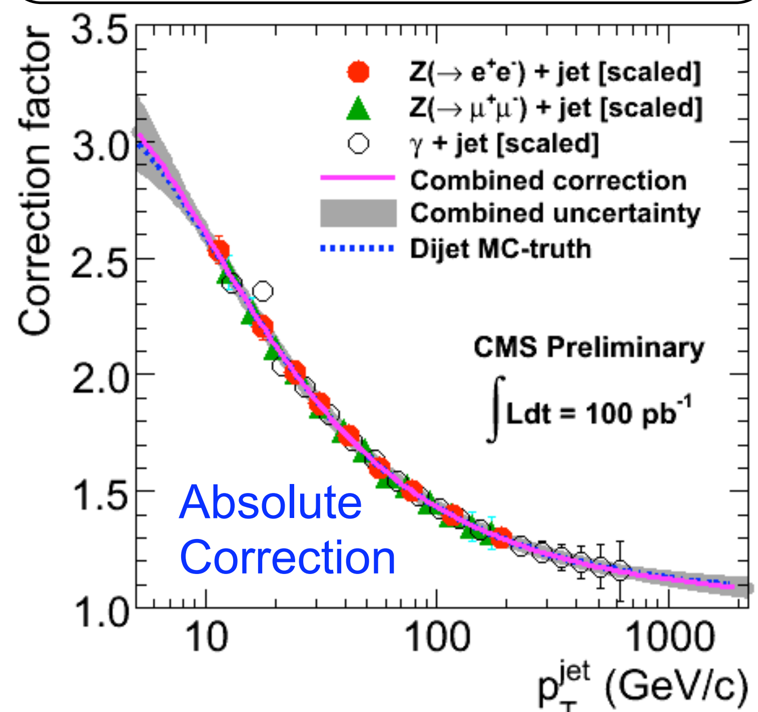
I will go over some of these points in the following slides

Jet energy scale (JES) correction

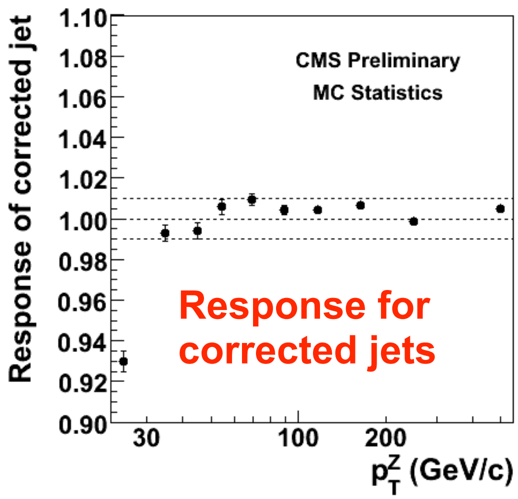
Jet response = calorimeter measured jet p_T / true jet p_T



Inverse of the jet response is the correction



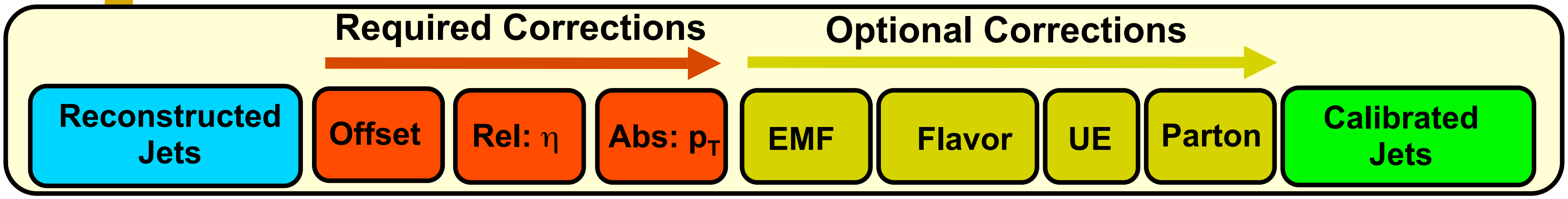
The correction is valid for the quark-gluon mixture in QCD dijet events.



JME-09-005

<u>Conclusion:</u> The correction procedure works well for $p_T > 30$ GeV/c. Work is under way to understand/correct the bias at low p_T. We possibly cannot go below 25 GeV jet p_T in this analysis.

CMS plans for Jet Energy Correction



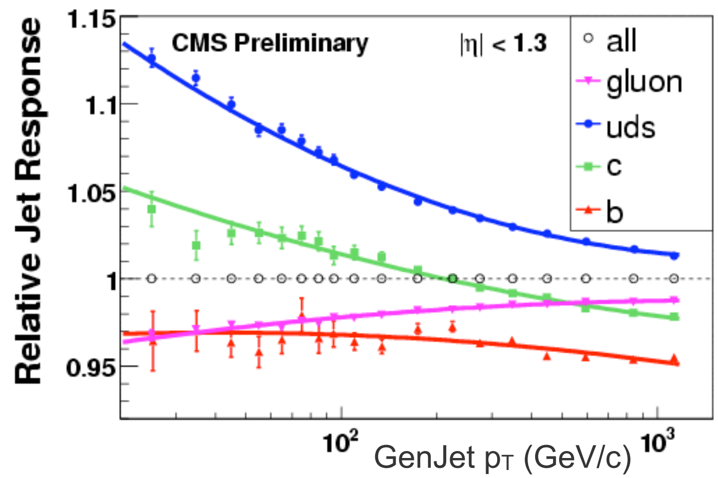
- CMS plans require the following jet energy corrections for most analysis
  - ➡ **Offset**: correction for pile-up and noise
  - ➡ **Relative**: correction for jet response vs. η relative to barrel
  - ➡ **Absolute**: correction for jet response vs. P_T in barrel
- The absolute correction will be measured ***in situ*** using p_T balance
  - ➡ In the three processes γ+jet, $Z(\rightarrow \mu^+\mu^-)$+jet, and $Z(\rightarrow e^+e^-)$+jet
- In $Z (\rightarrow e^+ e^-)$ + jets analysis we will also need to apply flavor correction
  - ➡ **Flavor**: Z + jet events are rich in quarks. Quark jets have larger response than gluon jets which dominate QCD sample (the default flavor composition for CMS absolute correction).

JME-07-002

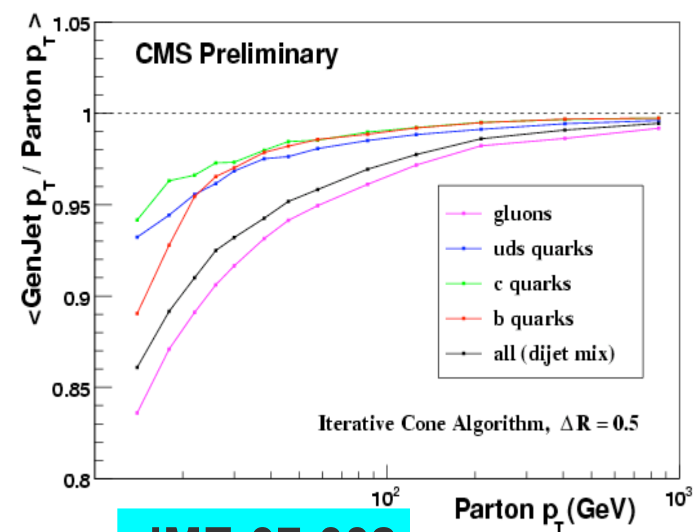
Flavor dependence and JES uncertainty

- γ/Z + jet events are rich in quarks
  - ➡ Quark jets have larger response than gluon jets which dominate QCD sample
- γ/Z + jet p_T balance corrects back to the parton level.
  - ➡ But the absolute correction is defined to be to the particle jet level

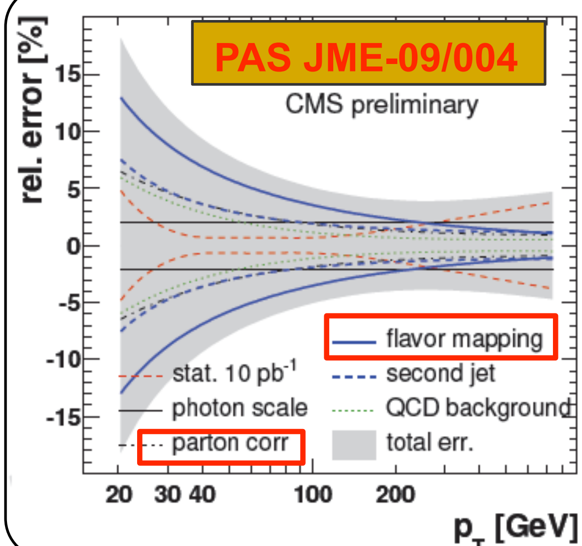
Flavor Corrections



Parton Corrections



JME-07-002

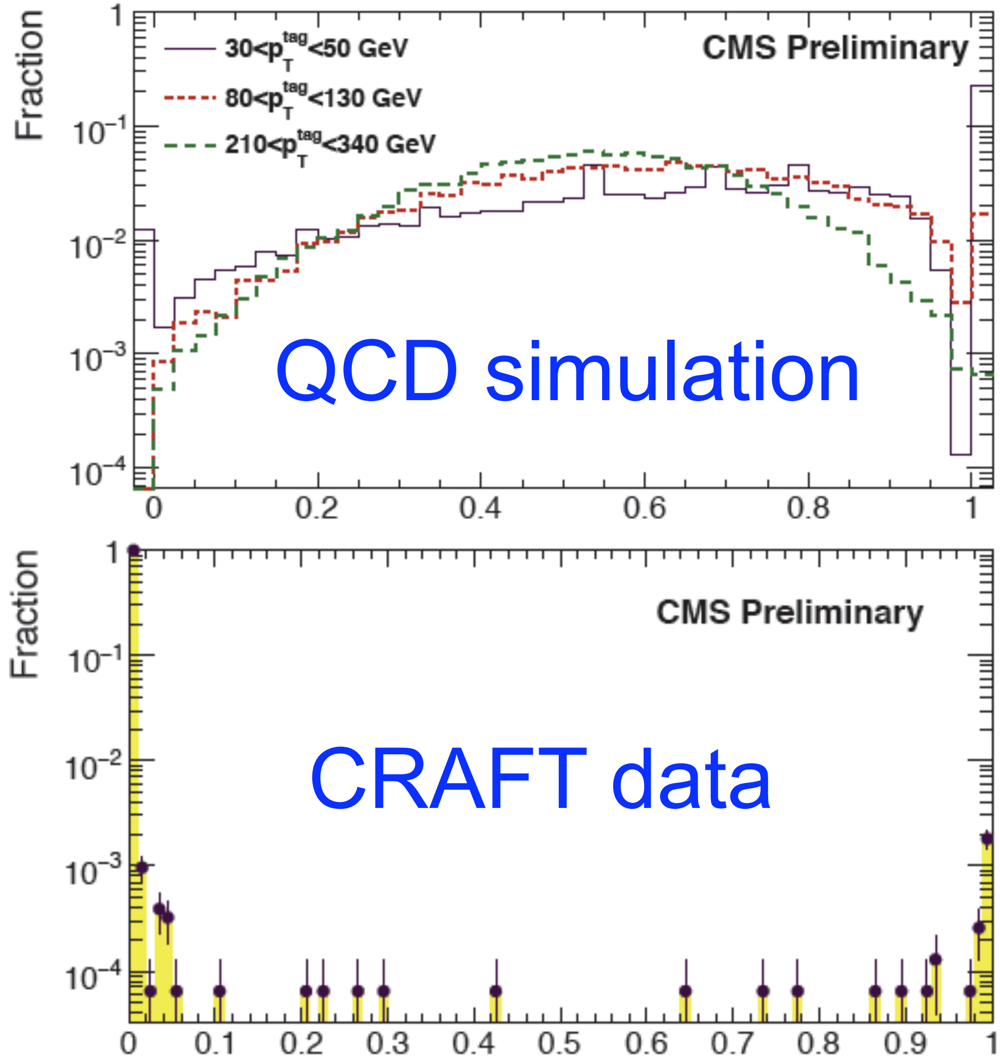


Current estimates of JEC uncertainty

JetMET expects a 10% uncertainty in the jet energy scale in the beginning of data taking !

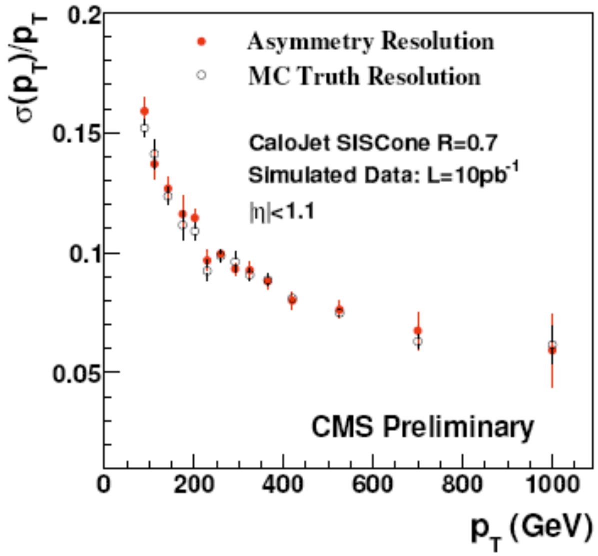
Jet Id

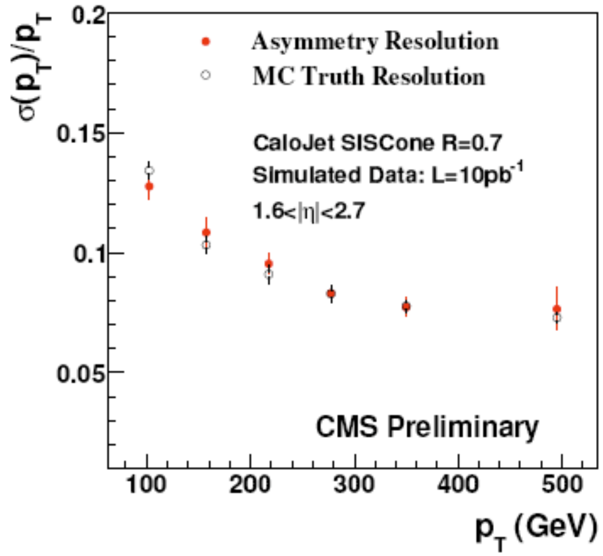
- We will be swamped by fake jets especially at low p_T.
- Given that we may start up with ~7 pileup events makes it worse.
- Studies, based on CRAFT data and MC simulation, indicate that the electromagnetic energy fraction (EMF or f_{EM}) is a powerful criterion to reject fake jets.
- Other quantities are being examined too:
  - #CaloTowers
  - #CaloTowers containing 90% of the jet energy
  - RMS of the E_T weighted ϕ distribution
  - fraction of energy from the hottest HPD/ RBX



JME-09-008

Jet resolution and reconstruction efficiency

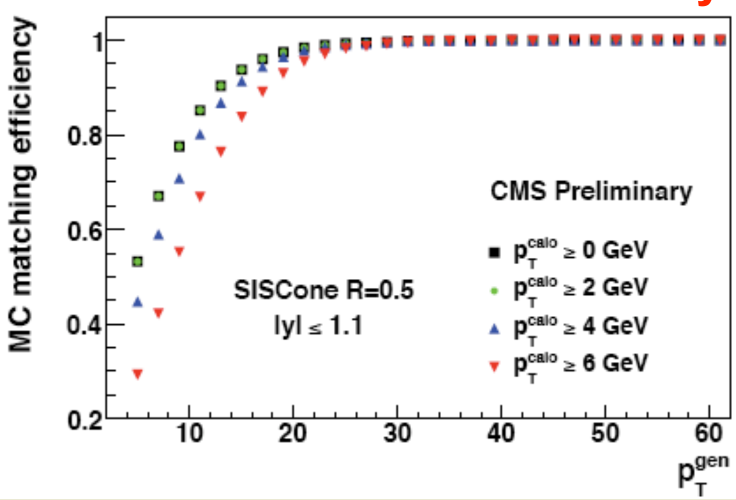




Jet p_T resolution

- Jet resolution improves at higher p_T.
- Resolution is slightly better in endcap than in the barrel !
- Can be measured in data using p_T balance in the QCD dijet events (asymmetry method).

Jet reconstruction efficiency



MC truth matching efficiency for several choices of CaloJet p_T threshold. The matching criteria is $\Delta R < 0.5$.

- The jet reconstruction efficiency is practically 100% for jets with $p_T > 20$ GeV/c.
- BUT this does not let us off the hook for systematic uncertainty:
  - -A 15 GeV jet can easily fluctuate to 30 GeV and efficiency at 15 GeV ≠ 100%.

JME-09-007

Action items: things to do

✦To cancel lots of Z-related systematic uncertainties (including lumi & Z efficiency), normalize Zee+jets x-sec w.r.t. inclusive Zee x-section. D0 paper arXiv: 0903.1748

✦We need to apply jet quality cuts, flavor correction, resolution

- At low p_T, fake jets can swamp the real jets
- In the past, EMF cut has been very effective: ECAL noise peaks at EMF=1, HCAL noise peaks at EMF=0
- Need to apply JES & flavor correction from "data", resolution unsmearing.
- We will get some help from JetMET, but we should be ready ...

✦Plan for data processing & skim: (can piggyback on JetMET skim for a while)

- this analysis will use single electron trigger: HLT_Ele15_SC10_LW_L1R, rate = 2.33 Hz
- we'll need to skim these events from the primary dataset: PD = EleGamma, rate = 23.19 Hz
- I have successfully tested a machinery to filter events using trigger bit. Working on which event content to keep/drop.

✦So far, we haven't thought much about comparison with theory

- The most obvious application of our measurement is in understanding the standard candles, tuning of MC, & MET calibration
- Need to compare with theory (pQCD LO, NLO, and NNLO) prediction.
- Need to repeat analysis in bins of jet multiplicity (maybe, also N+1 to N jet ratio).

Status and validation of 3.1.X samples

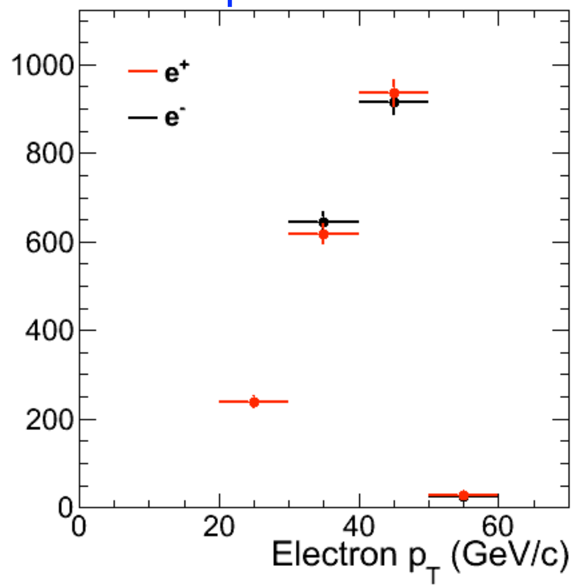
Status of Summer09 production

- The Zee+jets sample are produced, though not all available in DBS yet.
- Requested by JetMET physics object group for calibration purposes.
- Generated in 10 $\hat{p}_T$ bins: **0-15, 15-20, 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-inf**.

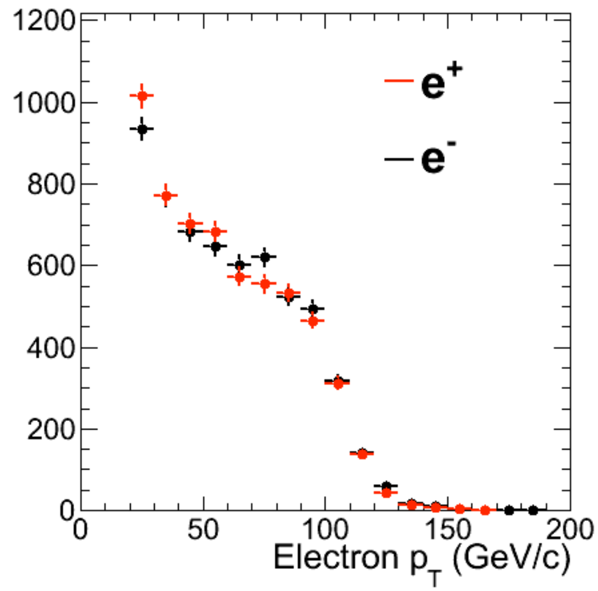
1006010	/ZeeJet_Pt0to15/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	200000	219730	100.00	0.059	IN2P3 (0.0)	T2_FR_IPHC (20.2) T2_BE_UCL (26.9) T2_FR_CCIN2P3 (22.2) T2_FR_GRIF_LLR (20.2) T2_CN_Beijing (20.2)	2009-07-29	done	No
1006070	/ZeeJet_Pt120to170/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	116496	100.00	0.051	RAL (100.0)	T2_EE_Estonia (32.7) T2_UK_London_Brunel (21.3) T3_UK_London_RHUL (1.8) T2_UK_SGrid_RALPP (20.1) T2_UK_London_IC (35.1) T3_UK_ScotGrid_GLA (5.5)	2009-07-30	done	No
1006020	/ZeeJet_Pt15to20/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	200000	220000	100.00	0.071	IN2P3 (0.0)	T2_FR_IPHC (30.9) T2_FR_CCIN2P3 (26.2) T2_BE_UCL (12.4) T2_FR_GRIF_LLR (3.5) T2_CN_Beijing (37.0)	2009-07-29	done	No
1006080	/ZeeJet_Pt170to230/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	131000	100.00	0.061	RAL (100.0)	T2_EE_Estonia (21.1) T2_UK_London_Brunel (0.4) T3_UK_London_RHUL (9.5) T2_UK_SGrid_RALPP (34.9) T2_UK_London_IC (49.7) T3_UK_ScotGrid_GLA (6.3)	2009-07-30	done	No
1006030	/ZeeJet_Pt20to30/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	150000	162570	100.00	0.055	PIC (0.0)	T2_ES_CIEMAT (96.7) T2_PT_LIP_Lisbon (11.7)	2009-08-02	done	No
1006090	/ZeeJet_Pt230to300/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	108542	100.00	0.053	ASGC (95.9)	T2_IN_TIFR (12.6)	2009-07-29	done	No
1006100	/ZeeJet_Pt300toInf/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	106355	100.00	0.054	ASGC (87.8)	T2_IN_TIFR (18.6)	2009-07-29	done	No
1006040	/ZeeJet_Pt30to50/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	150000	156285	100.00	0.056	PIC (0.0)	T2_PT_LIP_Lisbon (15.7) T2_ES_CIEMAT (88.5)	2009-08-02	done	No
1006050	/ZeeJet_Pt50to80/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	110000	100.00	0.042	PIC (100.0)		2009-08-01	done	No
1006060	/ZeeJet_Pt80to120/Summer09-MC_31X_V3-v1/GEN-SIM-RECO	100000	110587	100.00	0.046	RAL (100.0)	T2_UK_SGrid_Bristol (8.9) T2_EE_Estonia (25.6) T2_UK_London_Brunel (24.0) T3_UK_London_RHUL (1.2) T2_UK_SGrid_RALPP (12.2) T2_UK_London_IC (28.8) T3_UK_ScotGrid_GLA (8.9)	2009-07-31	done	No

Validation plots with pre-prod sample: Electron E_T

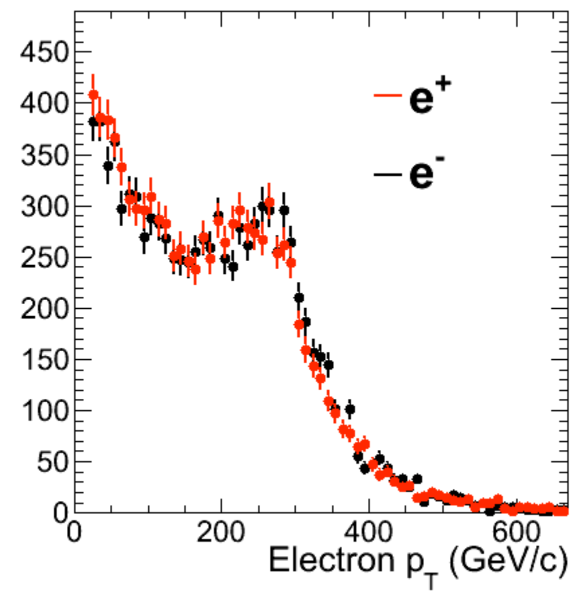
0 < pt < 15



80 < pt < 120



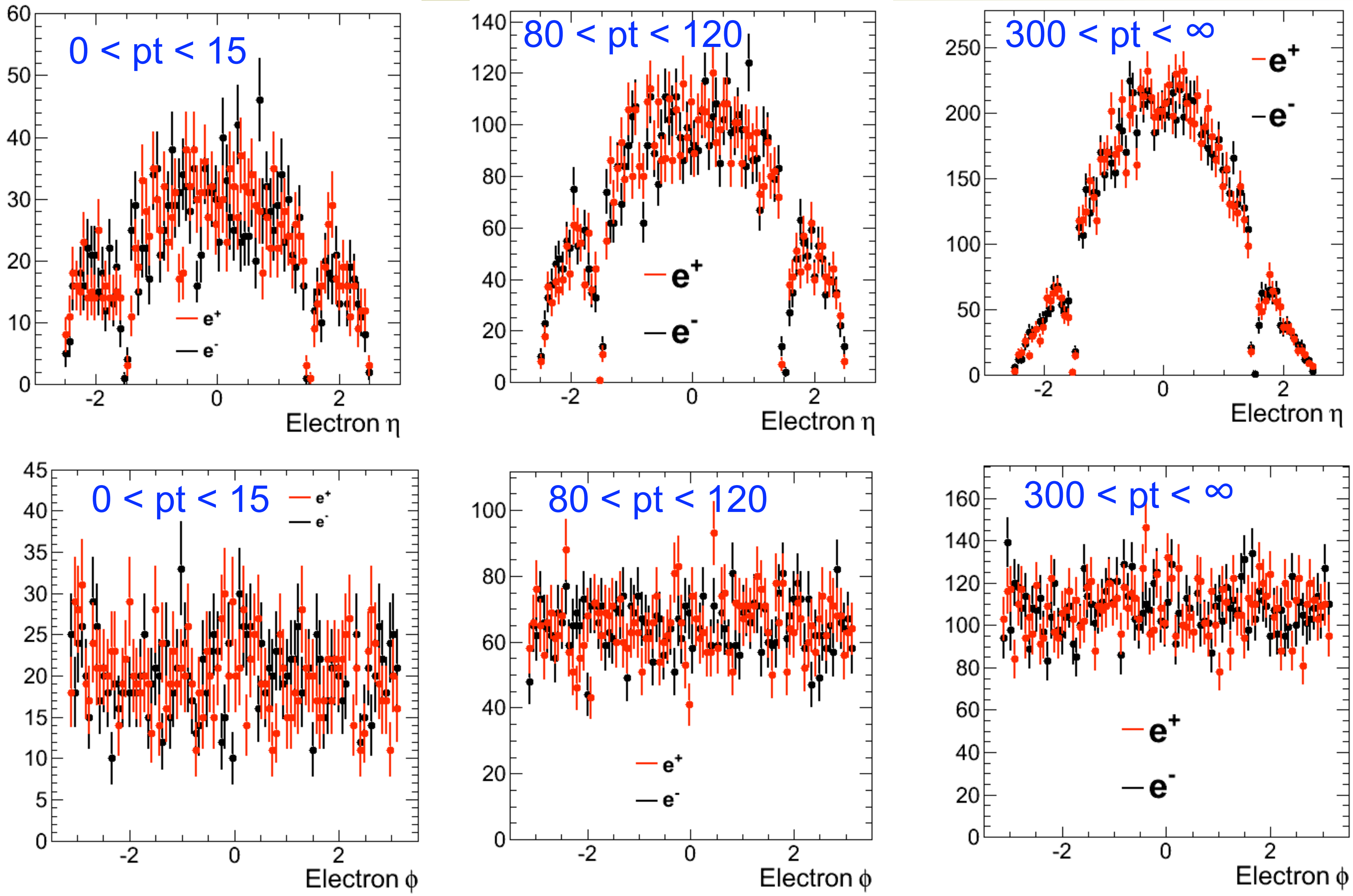
300 < pt < ∞



Electrons selection:

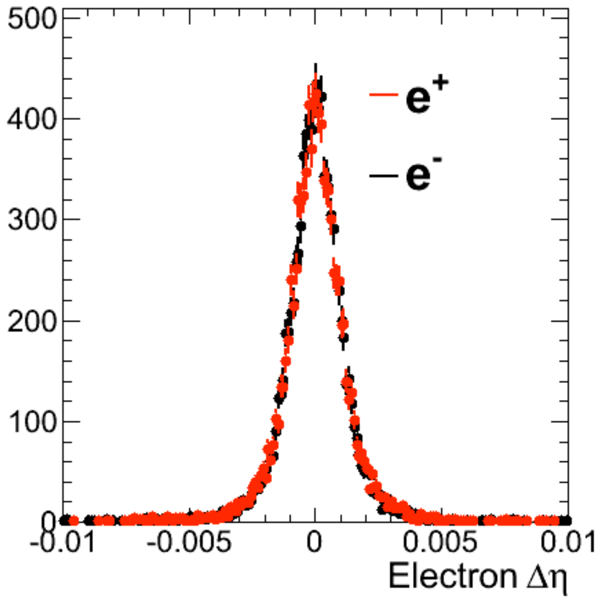
- super cluster matched to track (*gsfElectrons*)
- p_T > 20 GeV/c
- within ECAL and tracker acceptance:
  $|\eta| < 1.4442$ (barrel) OR $1.56 < |\eta| < 2.5$ (endcaps)

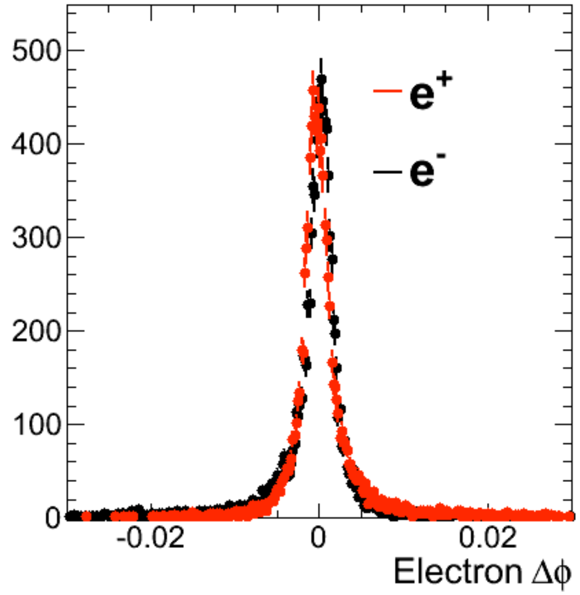
Electron eta, phi

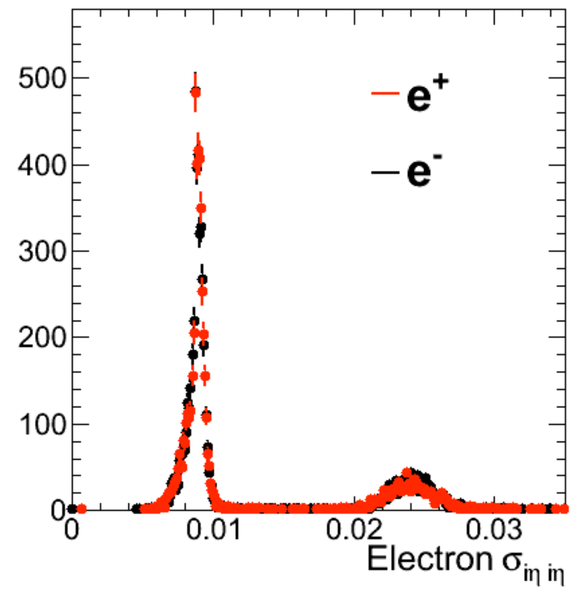


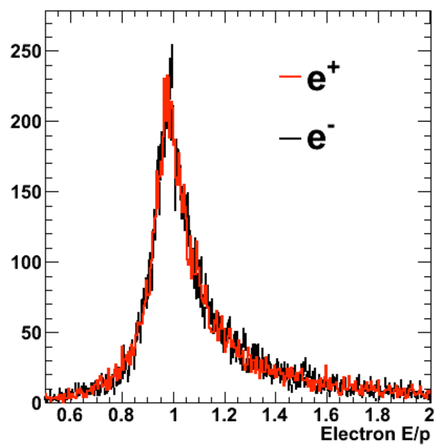
Electron id variables

all plots: $300 < p_t < \infty$









Definition

Electron identification variables:

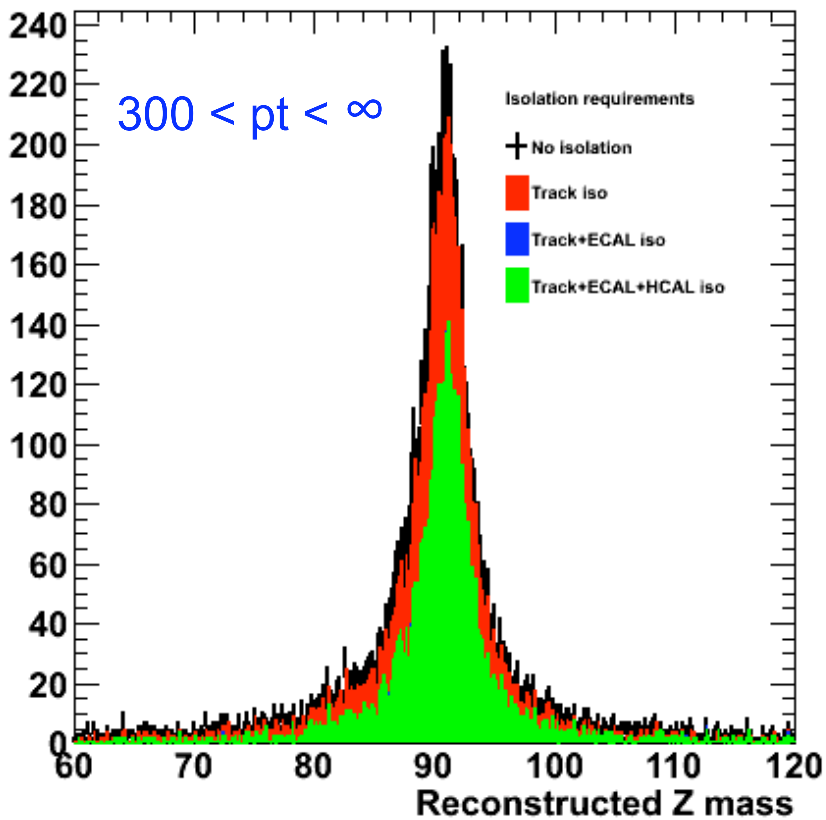
$\Delta\eta$: difference between η of electron supercluster and electron track

$\sigma_{i\eta i\eta}$: electron super cluster resolution

$\Delta\Phi$: difference between Φ of electron supercluster and electron track

E/p : ratio of electron supercluster energy to track momentum

Reconstructed Z mass



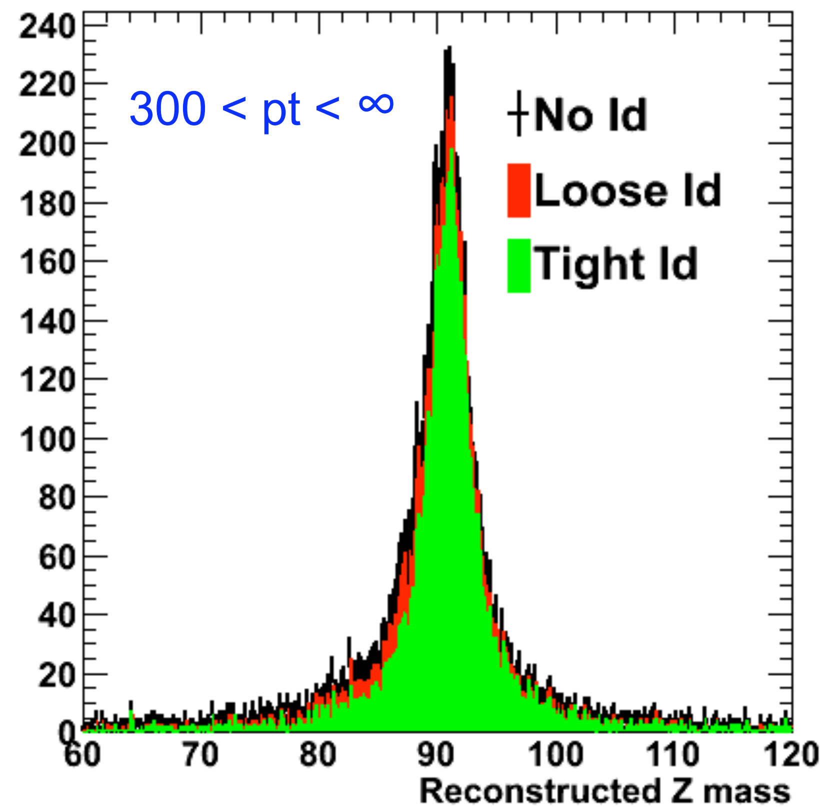


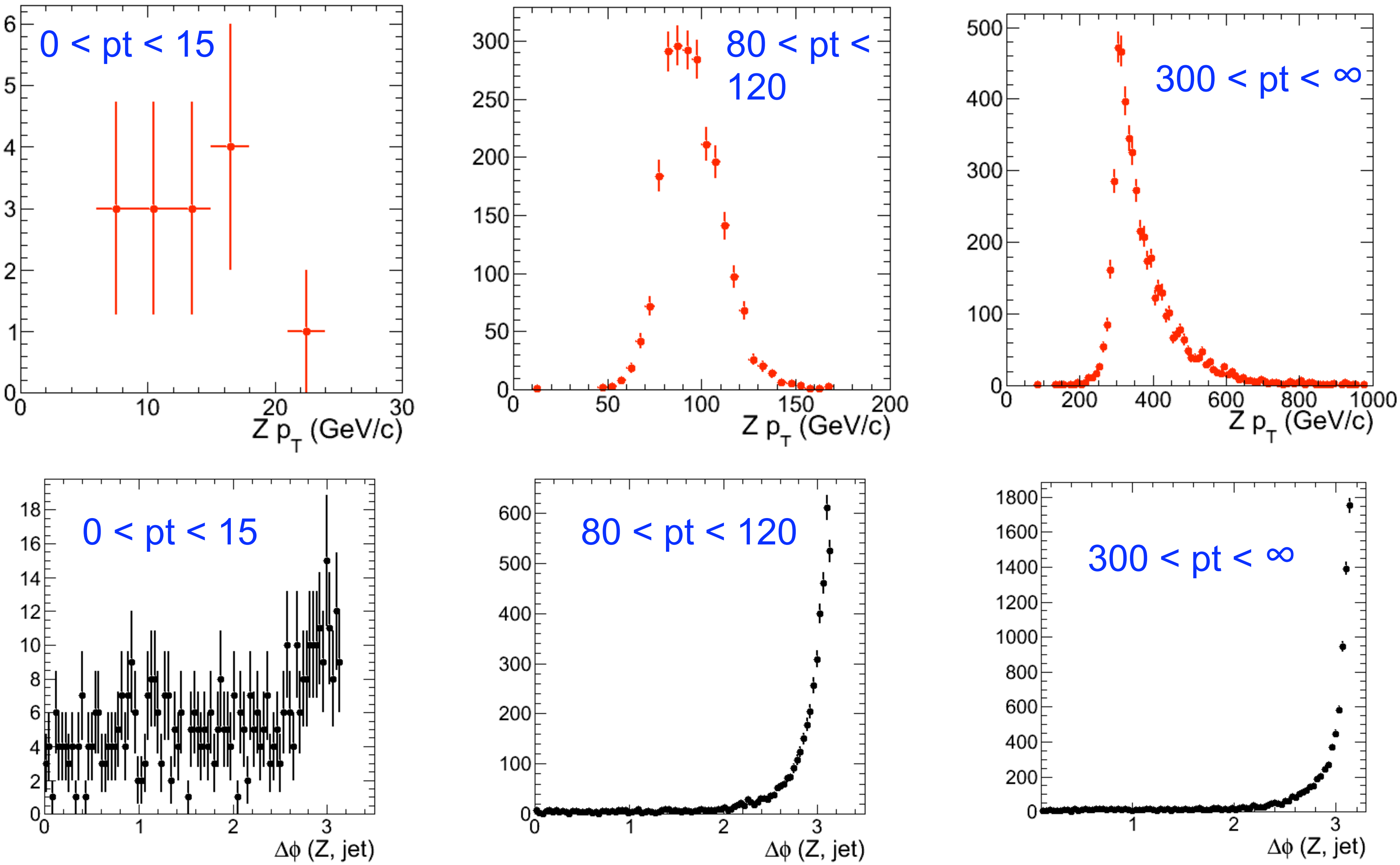
Table 1: Isolation criteria for $\gamma^*/Z \rightarrow e^+e^-$ candidates.

	Track	Ecal	Hcal
Barrel	7.2	5.7	8.1
Endcap	5.1	5.0	3.4

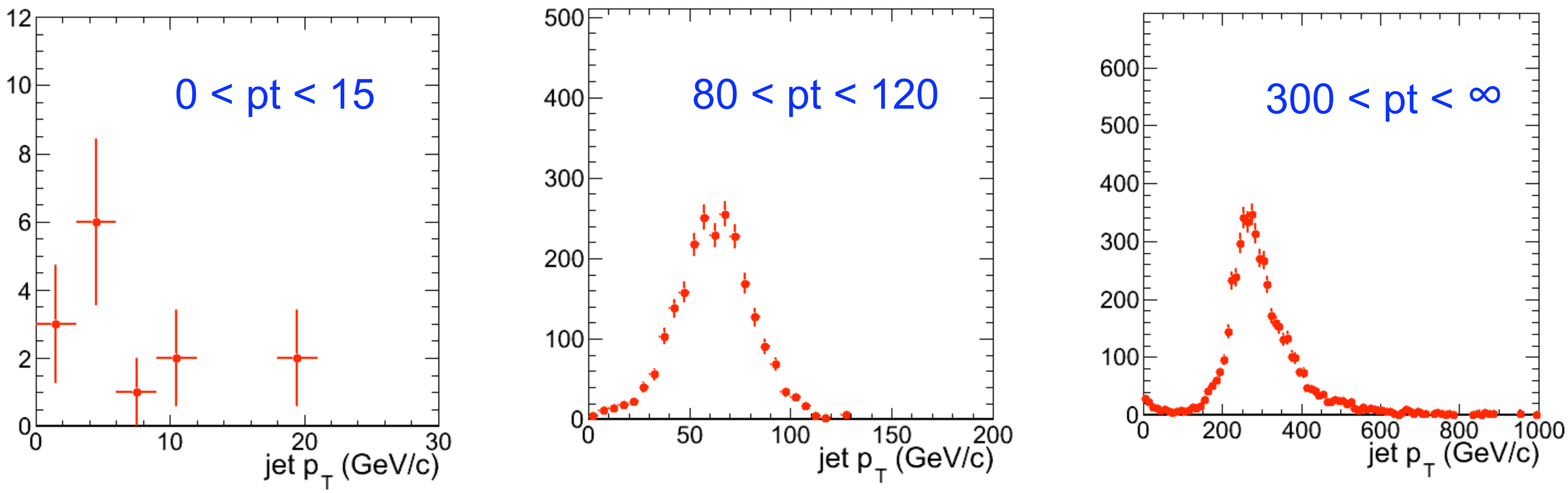
Standard *Egamma* electron ID:

https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideCutBasedElectronID

Z p_T and $\Delta\phi$(Z, jet) for Z+1jet events



CaloJet p_T



Summary of validation exercise

✦Physics validation of pre-production ZeeJet sample looks good.
✦Hope to have the Summer09 production samples available at Tier-2 sites very soon. The production is complete !!!
✦Can aim at performing the complete analysis chain by October.

End of Talk ! Thank You !

Backup slides

$Z \to e^+e^-$: electron isolation criteria

Definition

Track isolation: $\sum_{0.015<\Delta R<0.3,\ p_T^{Track}>1.0} \left(p_T^{\text{track}}\right)$

ECAL & HCAL isolation: $\sum_{\Delta R<0.4} \left(E_T^{\text{RecHits}}\right)$

Veto

ECAL:

- Inner cone radius
  barrel: $\Delta R > 0.045$, endcaps: $\Delta R > 0.070$
- E_T threshold
  barrel: $E_T > 0.08$, endcaps: $E_T > 0.3$
- "Jurassic footprint removal"
  $0.02 < |\Delta\eta| < 0.5$ (both barrel and endcaps).

HCAL: None

These are default *Egamma* definitions

Cuts

Table 1: Isolation criteria for $\gamma^*/Z \to e^+e^-$ candidates.

	Track	Ecal	Hcal
Barrel	7.2	5.7	8.1
Endcap	5.1	5.0	3.4

EWK-09-004

$Z \rightarrow e^+e^-$: electron id

Definition

The *EGamma* POG defines the following four electron identification variables to discriminate between real and fake electrons:

$\Delta\eta$: difference between η of electron supercluster and electron track
$\sigma_{i\eta i\eta}$: electron super cluster resolution
$\Delta\Phi$: difference between Φ of electron supercluster and electron track
H/E : ratio of hadronic to electromagnetic component of the energy

Cuts

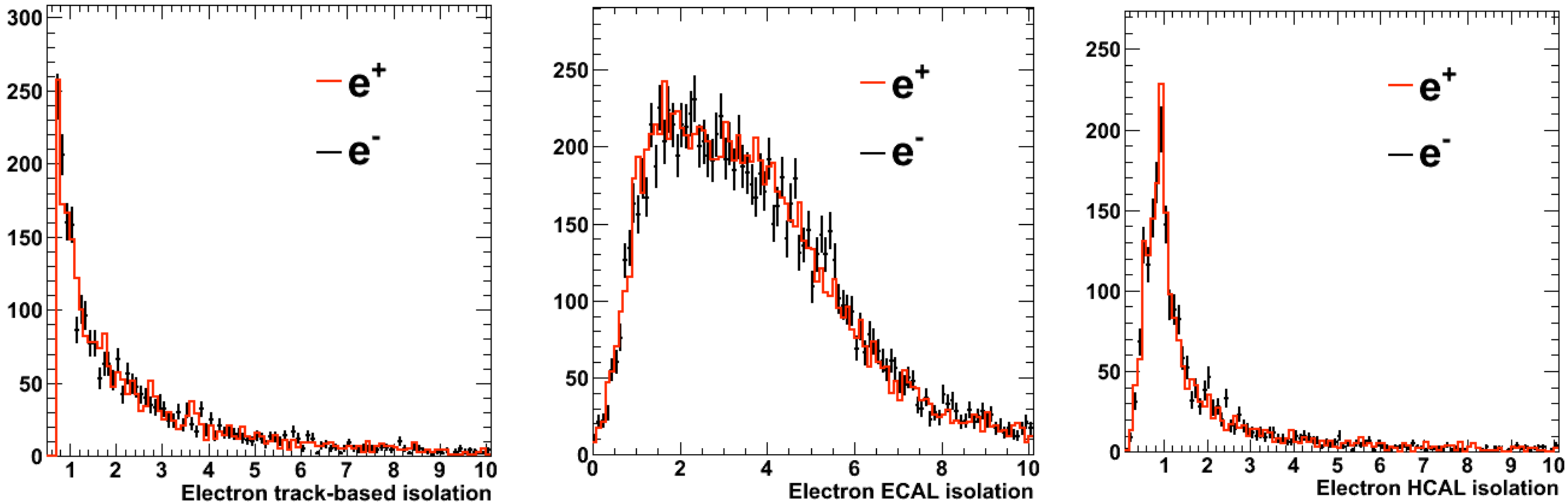
Variable	barrel	endcap
$\lvert\Delta\eta_{in}\rvert$	0.0077	0.0100
$\sigma_{i\eta i\eta}$	0.0132	0.027
$\lvert\Delta\varphi_{in}\rvert$	0.058	0.042
H/E	0.075	0.083

Standard *Egamma* POG electron ID
"Robust Loose"

https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideCutBasedElectronID

Electron isolation variables

all plots: 300 < pt < ∞



These are default *Egamma* definitions