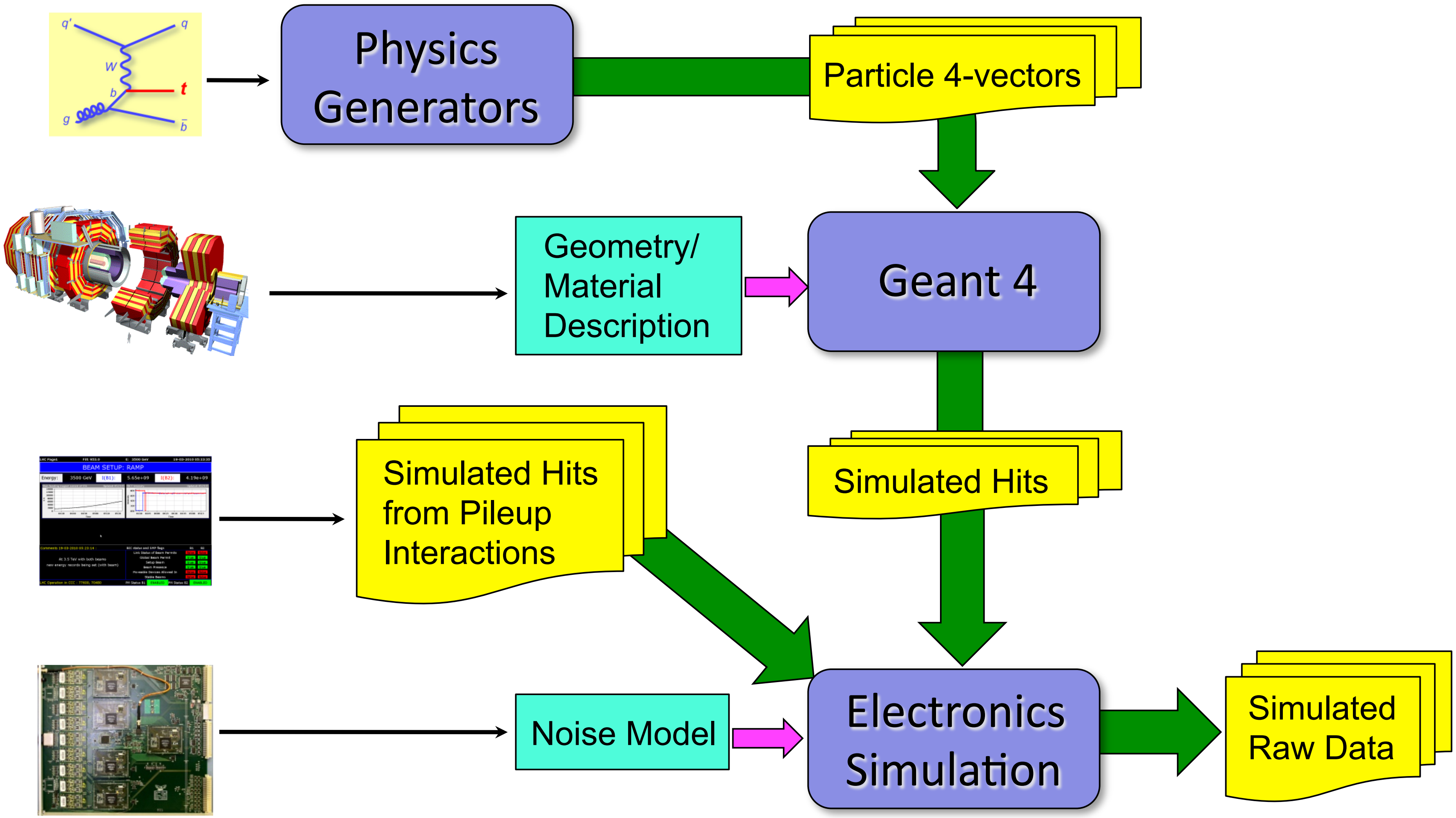
Validation and Tuning of the CMS Full Simulation

Mike Hildreth

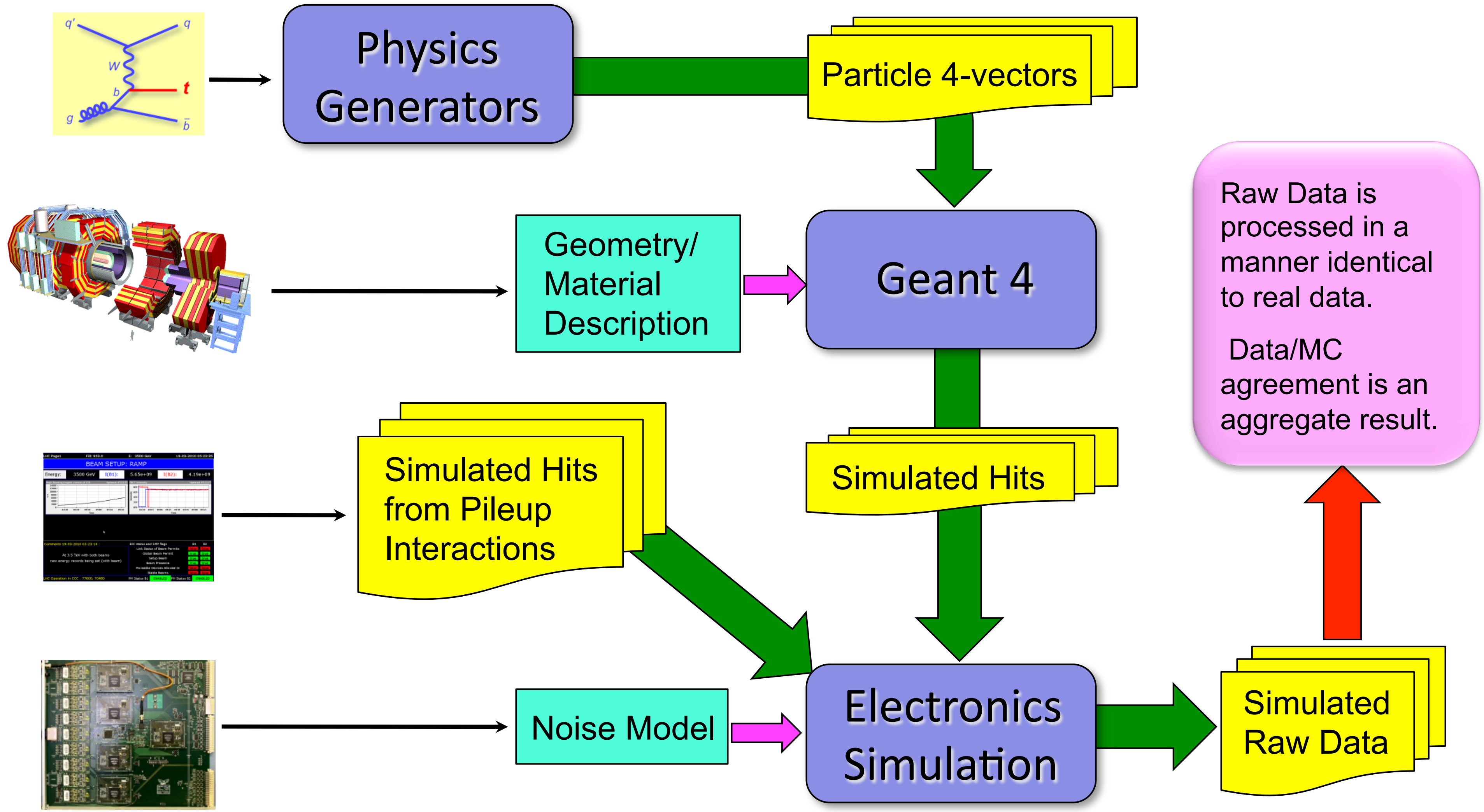
Université de Notre Dame du Lac & Fermilab

Representing the CMS Collaboration

Overview of CMS Simulation

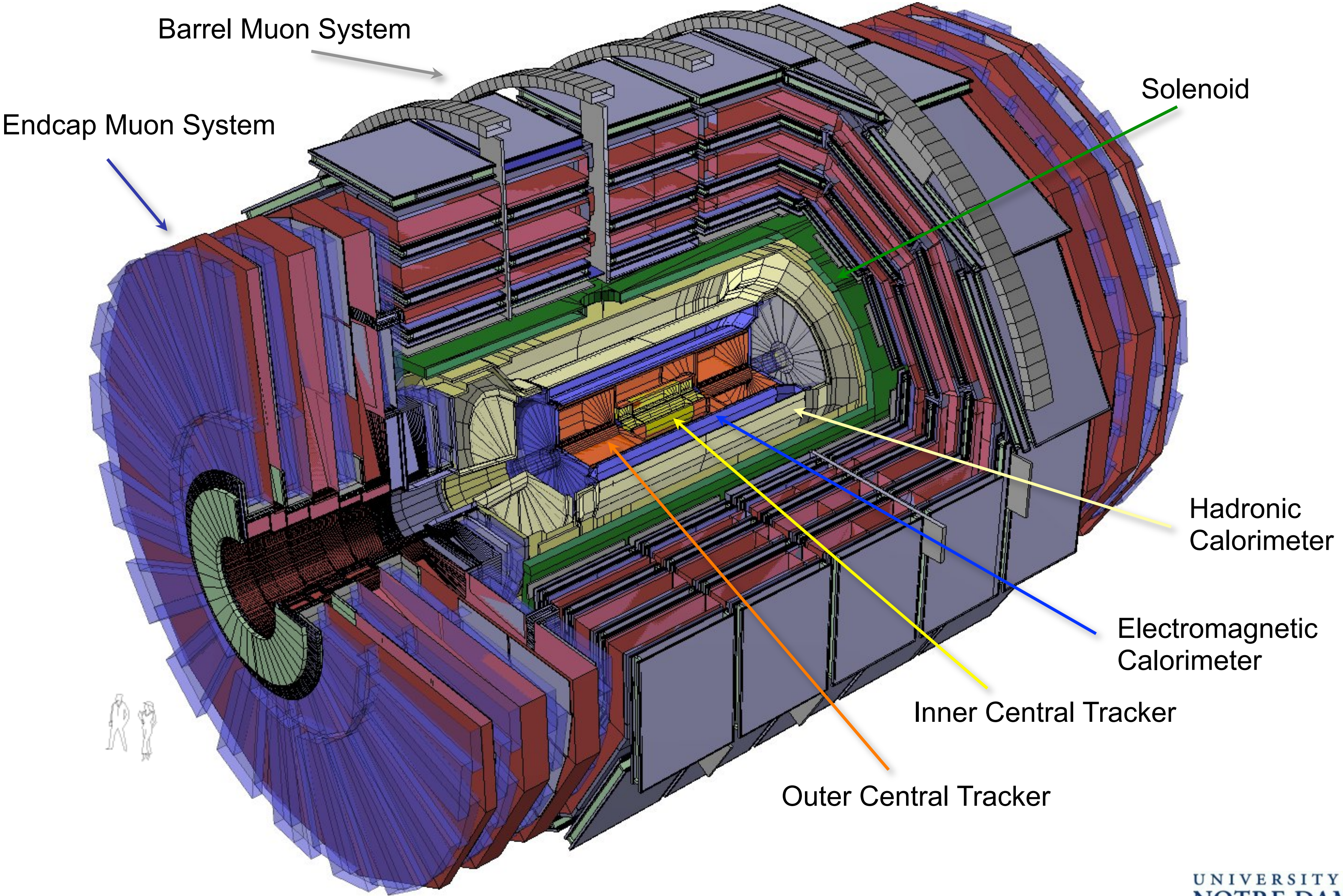


Overview of CMS Simulation



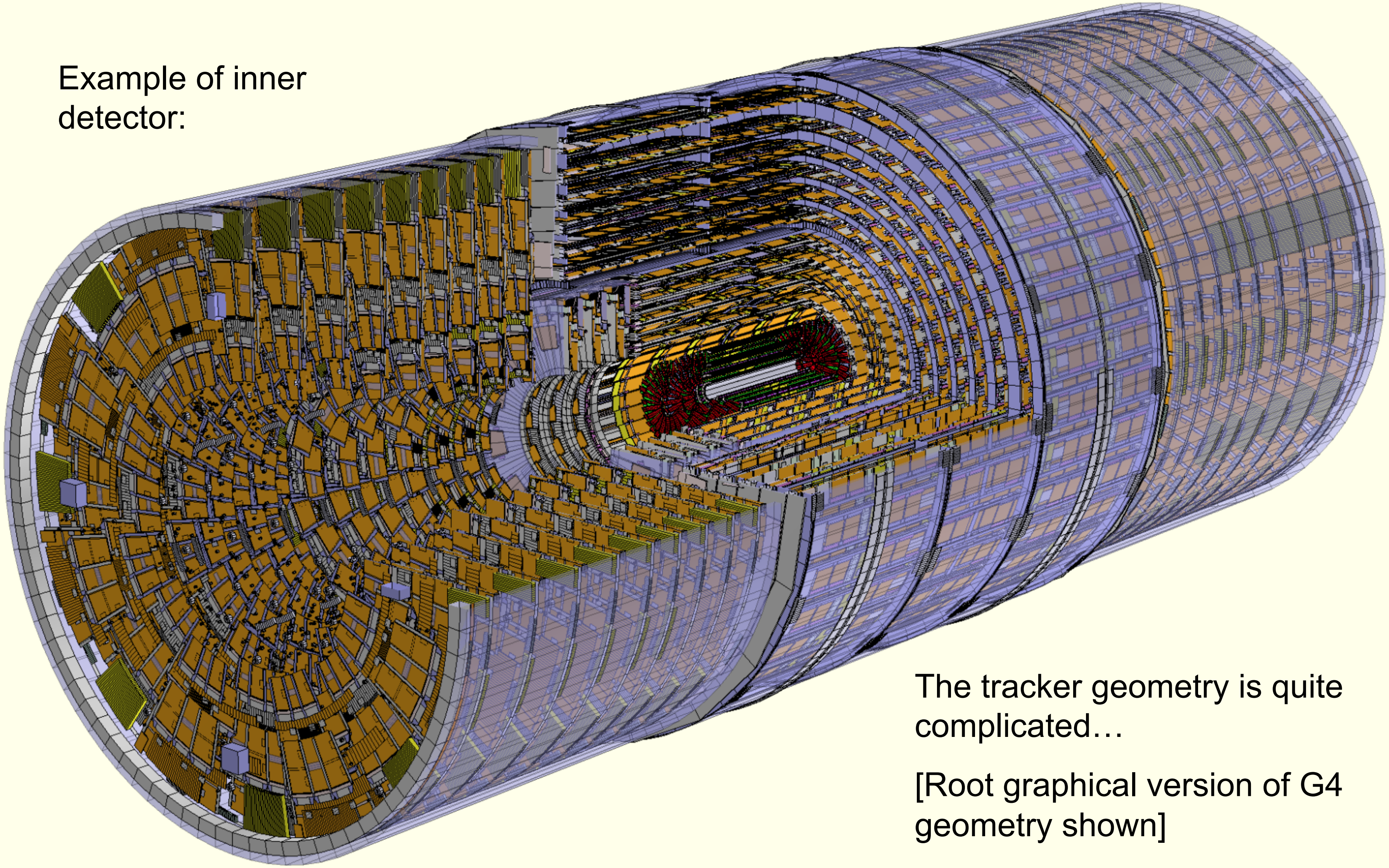
UNIVERSITY OF NOTRE DAME
Fermilab

Overview of CMS



Overview of CMS

Example of inner detector:



The tracker geometry is quite complicated…

[Root graphical version of G4 geometry shown]

Overview

- Status of Tracker Geometry & Material Description
  - Material studies
    - photon conversions/nuclear interactions
  - Tracker dE/dx results
  - Track distributions

  } low energy nuclear and EM modeling, accuracy of material specification, low p_T generator physics

- Calorimeter Modeling
  - Electron bremsstrahlung
  - Jet and Missing Energy studies

  } material specification, physics models in particle showers, calorimeter noise models

- Muon System
  - Hit patterns and isolation variables

  } material specification, neutron transport, shower models

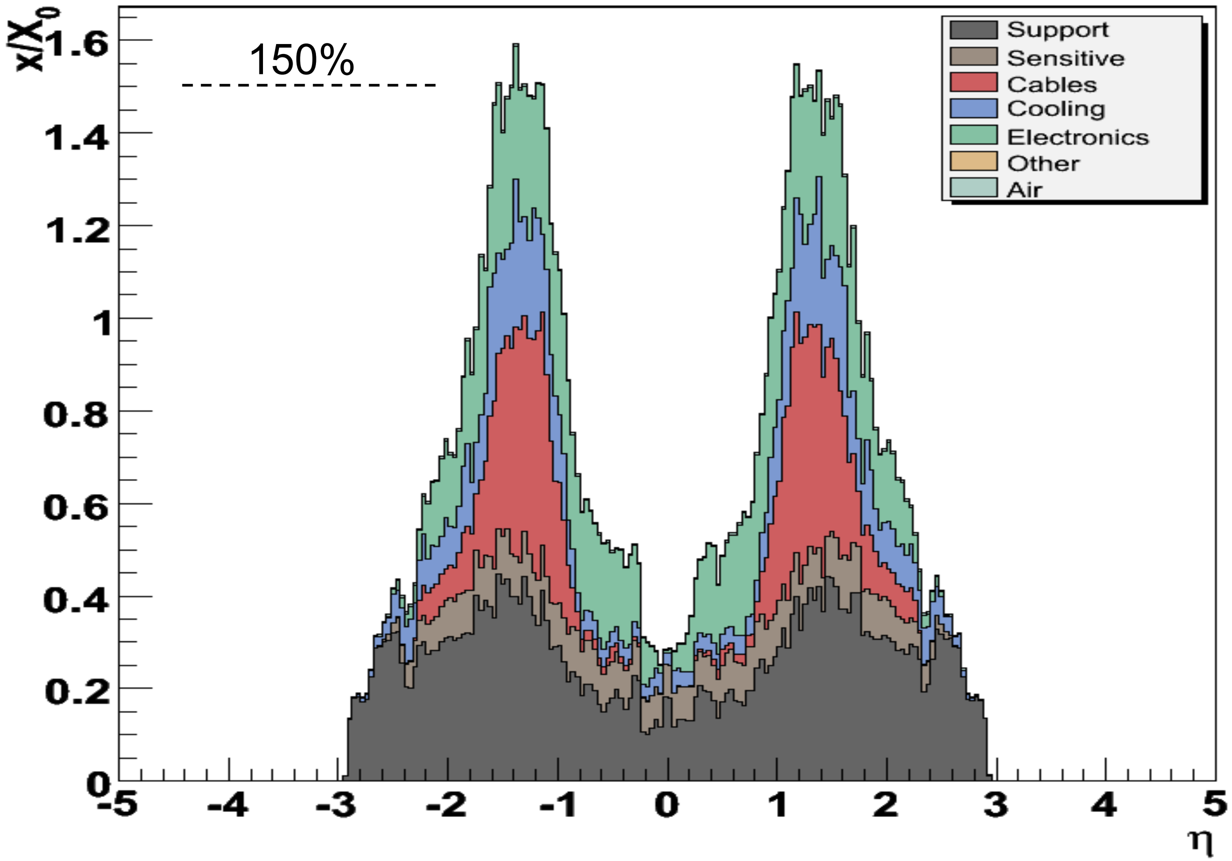
- Future Prospects

For other results, see:

Validation of Geant4 Physics Models with LHC Collision Data (PS08-1-170), Sunanda Banerjee

Detector Material Budget

Material distribution in current CMS Tracker (estimated):



- Very large photon conversion probability
- large effects of multiple-scattering
- must test with data to validate simulation

potentially large physics consequences

UNIVERSITY OF NOTRE DAME

Fermilab

Detector Material Studies

Reconstruction of Photon Conversions and Nuclear Interactions allow a mapping of the material distribution in the detector

Reminder - Photon conversion probability in a thin cylindrical shell:

$$dN_{conv} = N_{\gamma}(R,\theta,\phi) \cdot R^2 \sin\theta \, d\theta \, d\phi \frac{P}{X_0} dR \qquad N_{\gamma}(R,\theta,\phi) \propto 1/R^2 \sin\theta$$

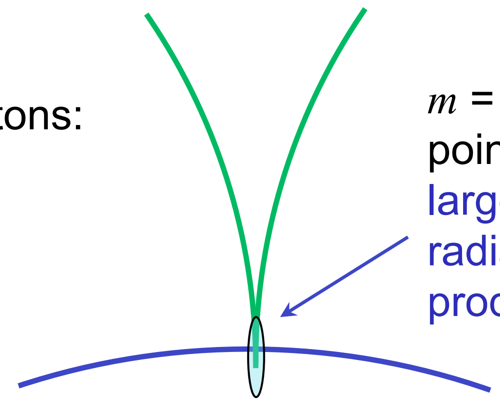
For Nuclear Interactions:

– swap P(photons) ~7/9 to P = 1, $X_0 \rightarrow \lambda_0$

(But, X_0 and λ_0 are sensitive to different physics)

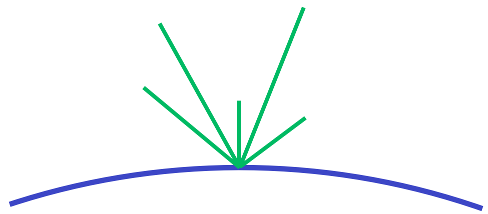
- Different reconstruction characteristics:

Photons:



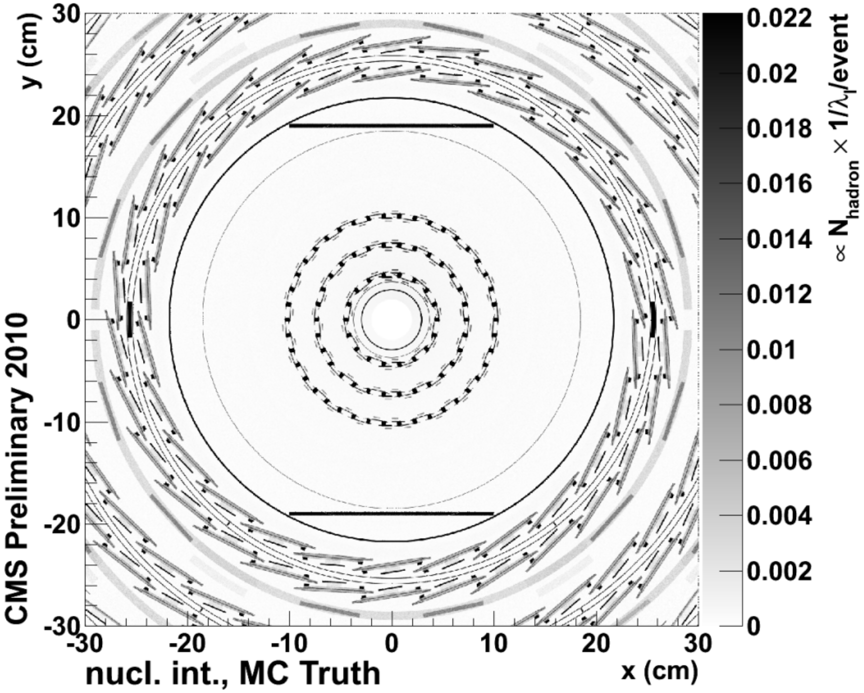
m = 0: tracks parallel at point of production ⇒ larger uncertainty in radial position of production point

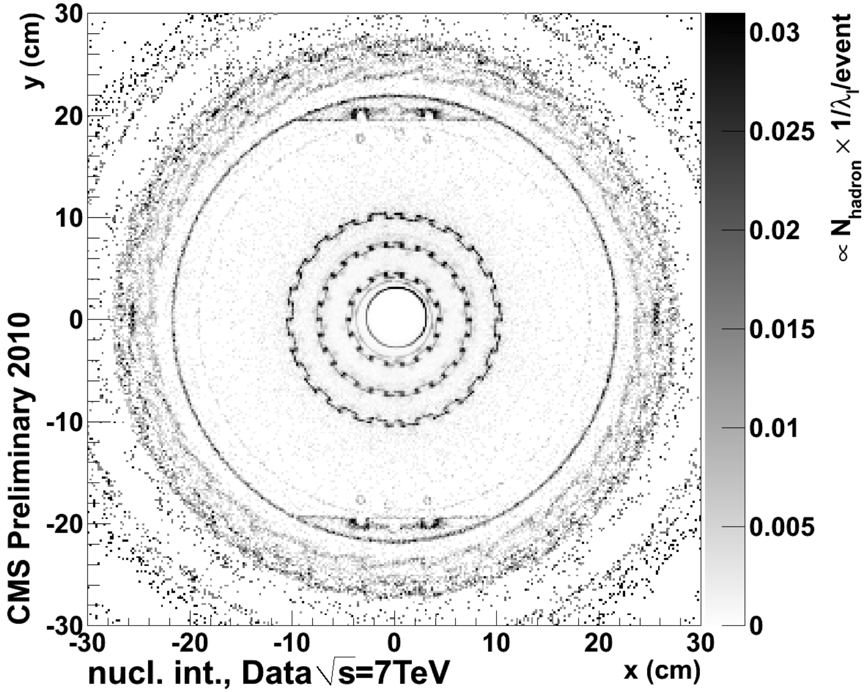
Nuclear interactions:



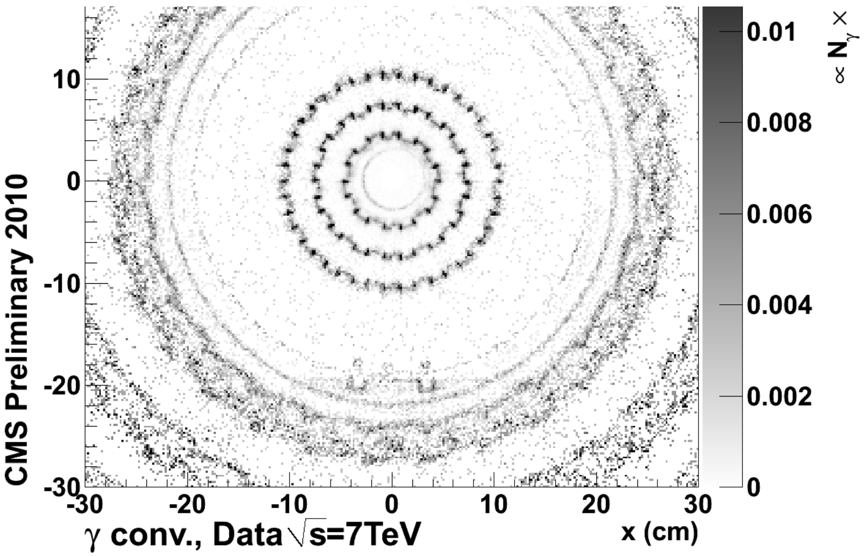
Good vertex resolution, many soft tracks with large impact parameters

Some examples:





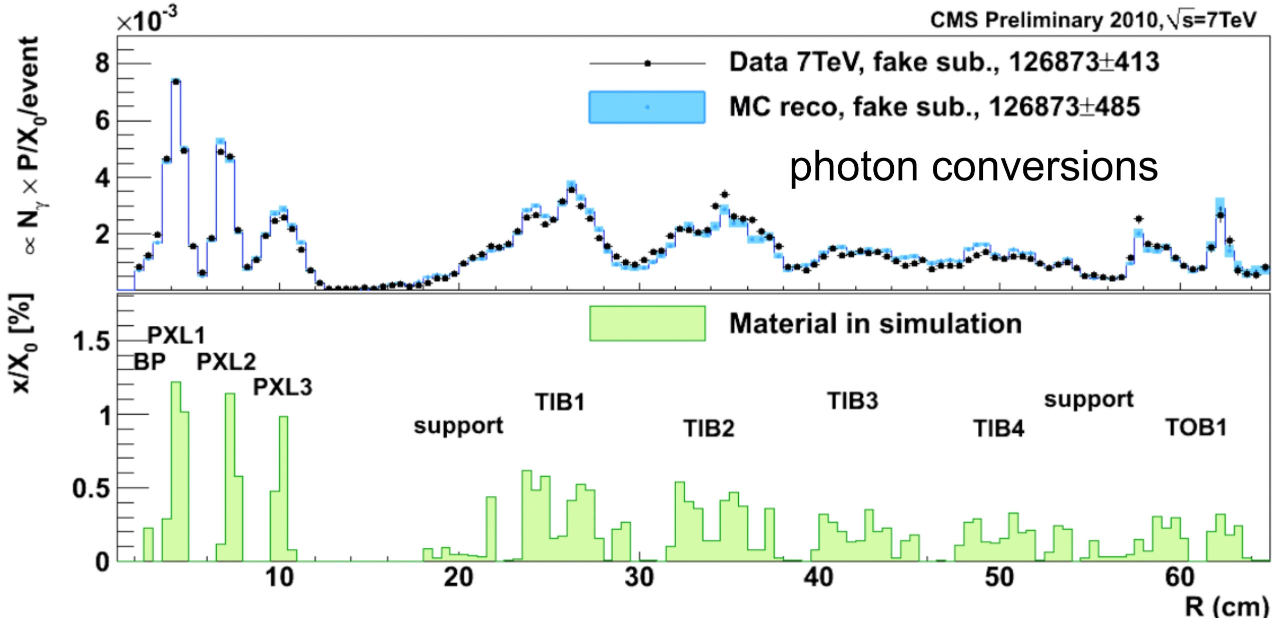
Note the superior position resolution of the nuclear interaction data



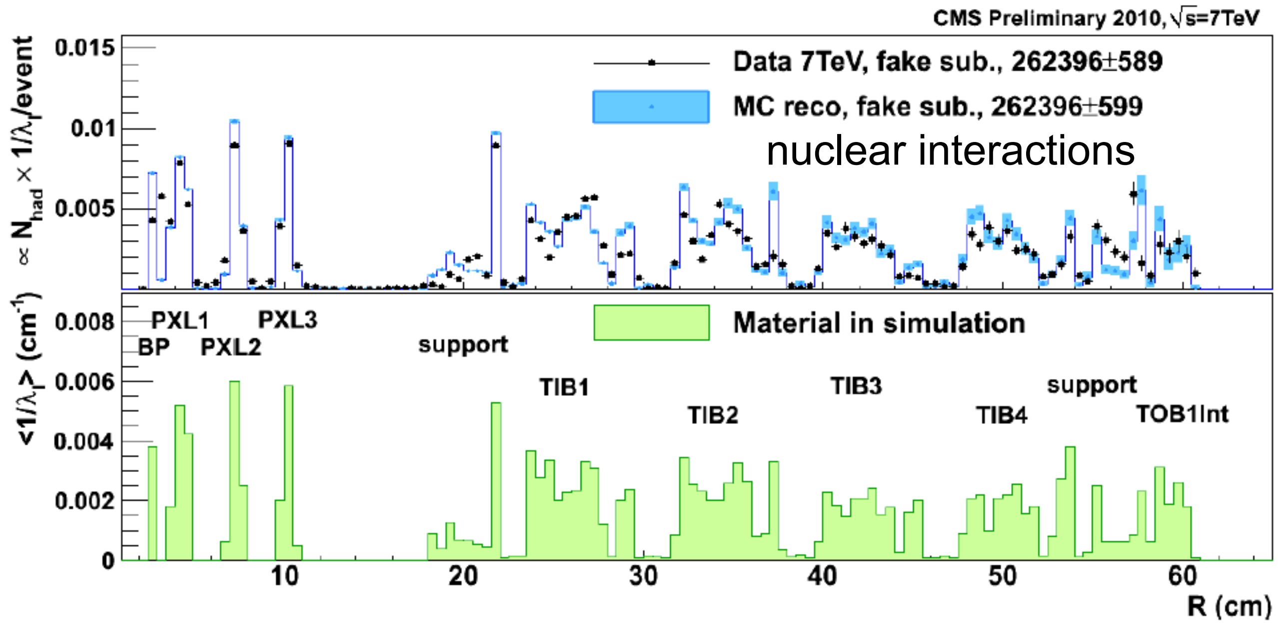
The beampipe isn't centered!

CMS PAS: TRK-10-003

Extracting the material budget

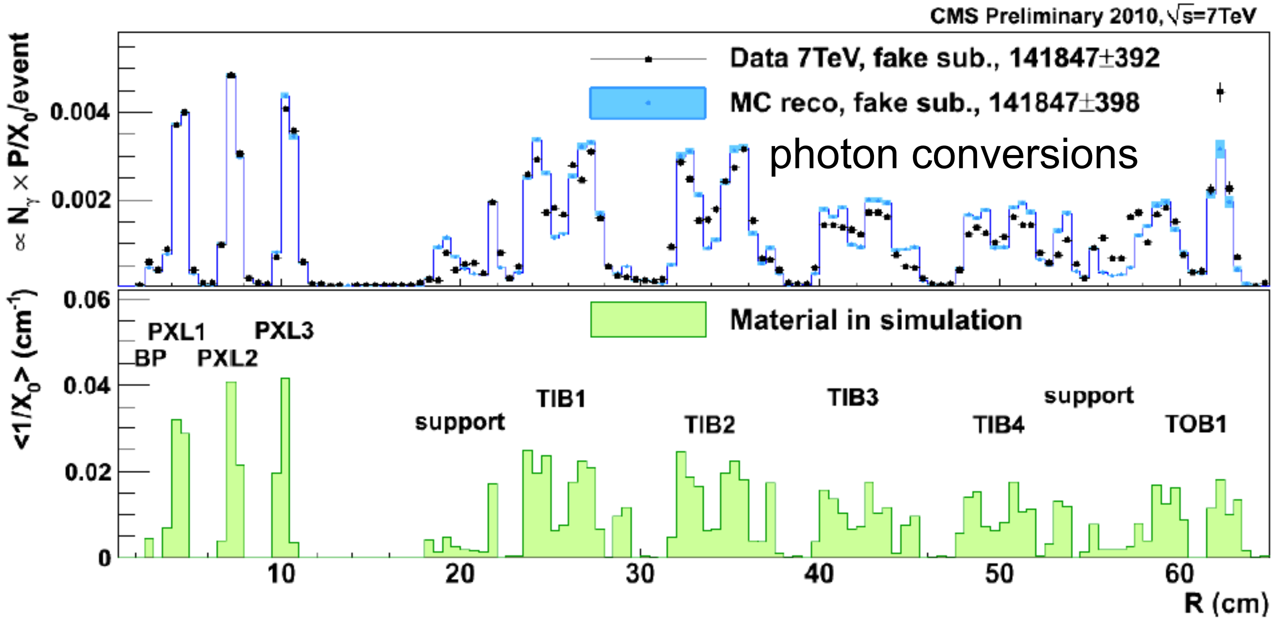


can unfold this distribution using estimates of the photon position resolution

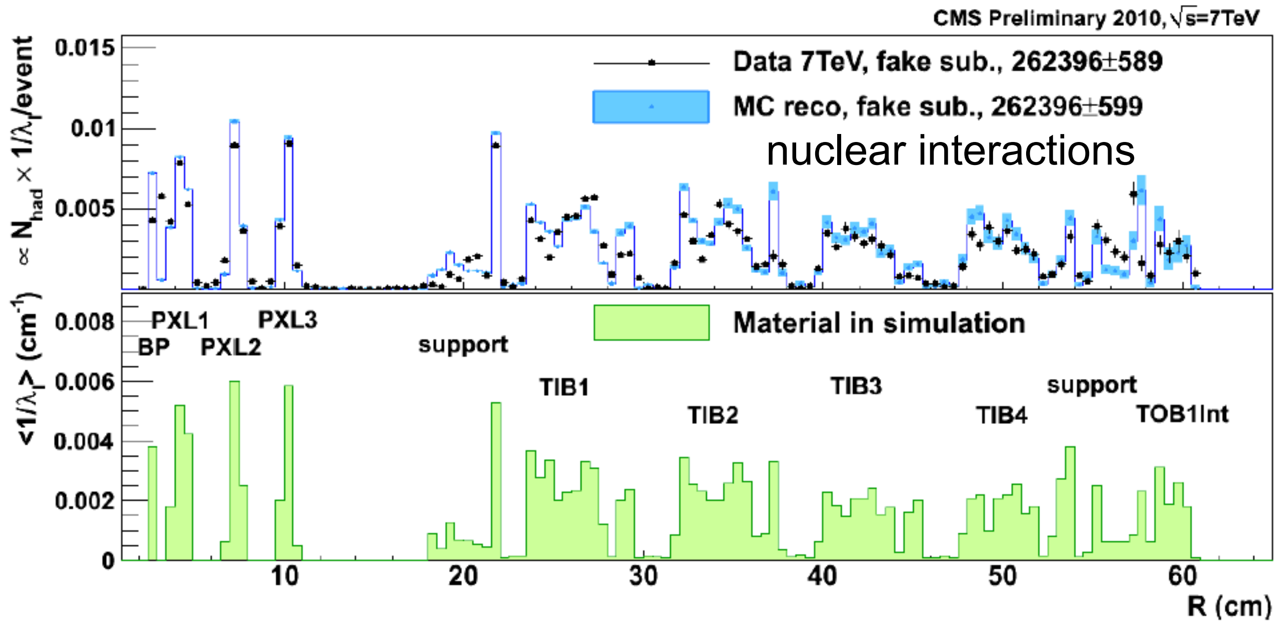


UNIVERSITY OF NOTRE DAME
Fermilab

Extracting the material budget



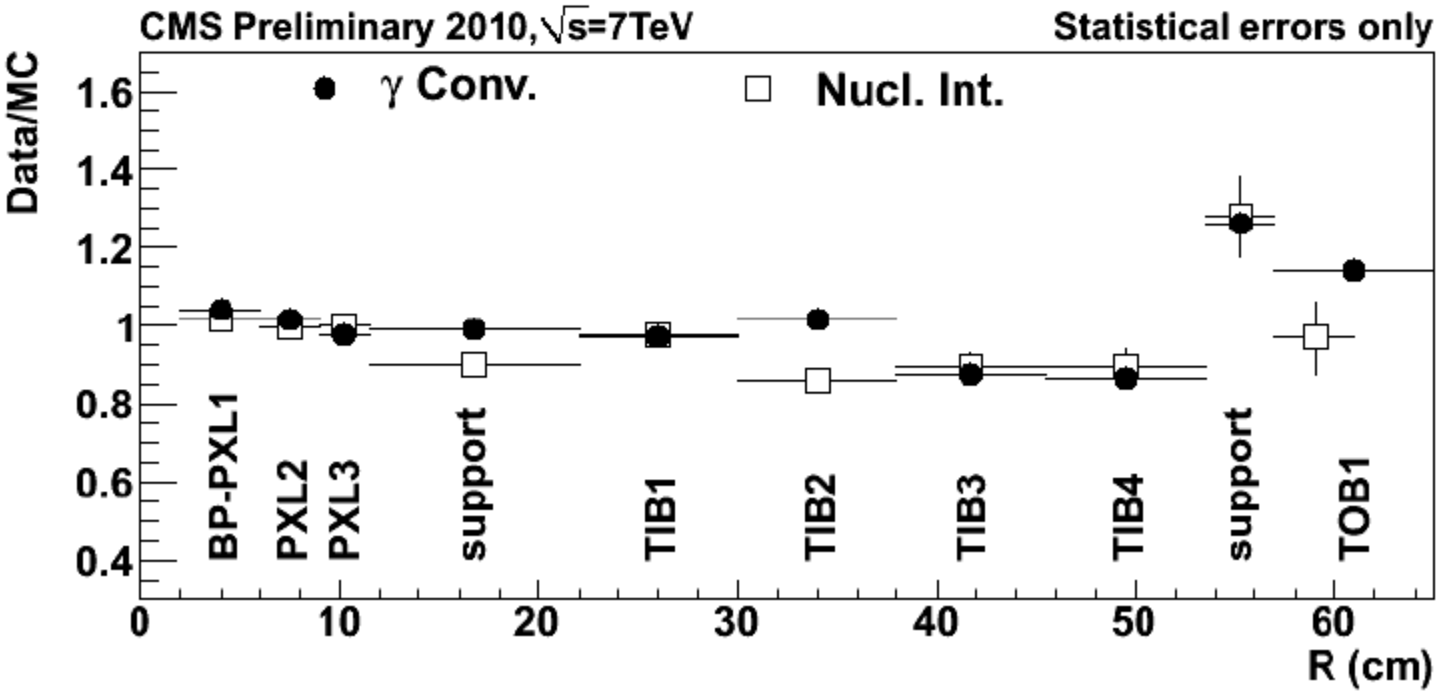
can unfold this distribution using estimates of the photon position resolution



astonishingly good agreement between data and simulation

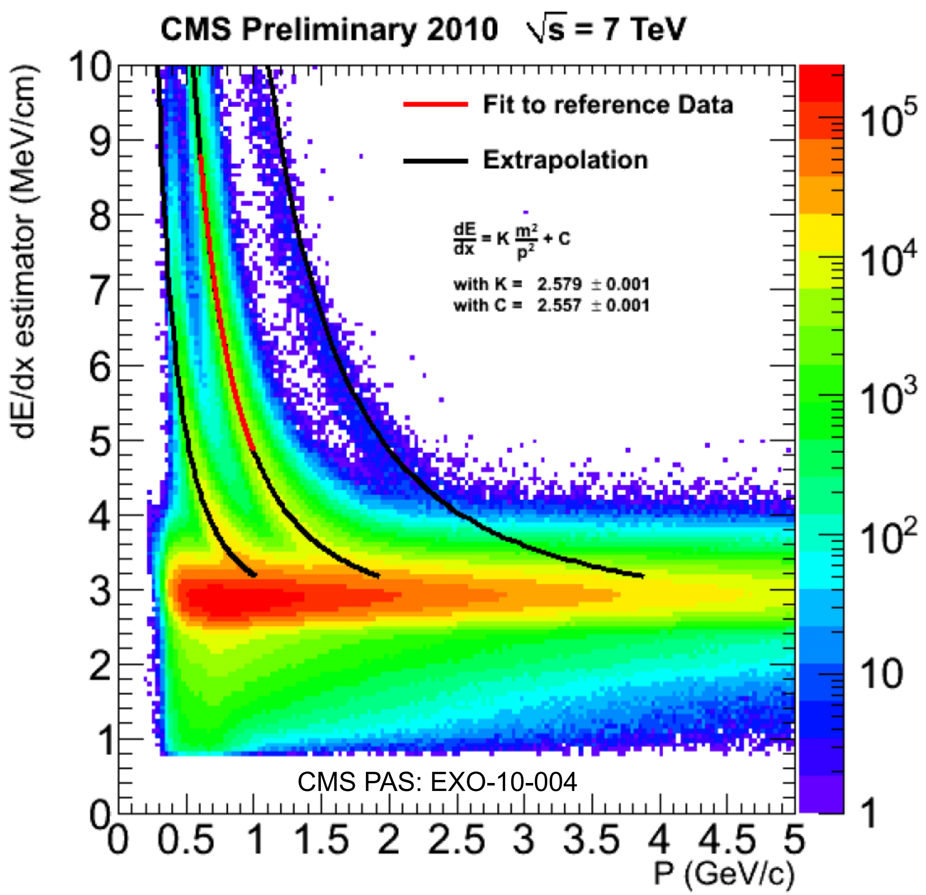
UNIVERSITY OF NOTRE DAME
Fermilab

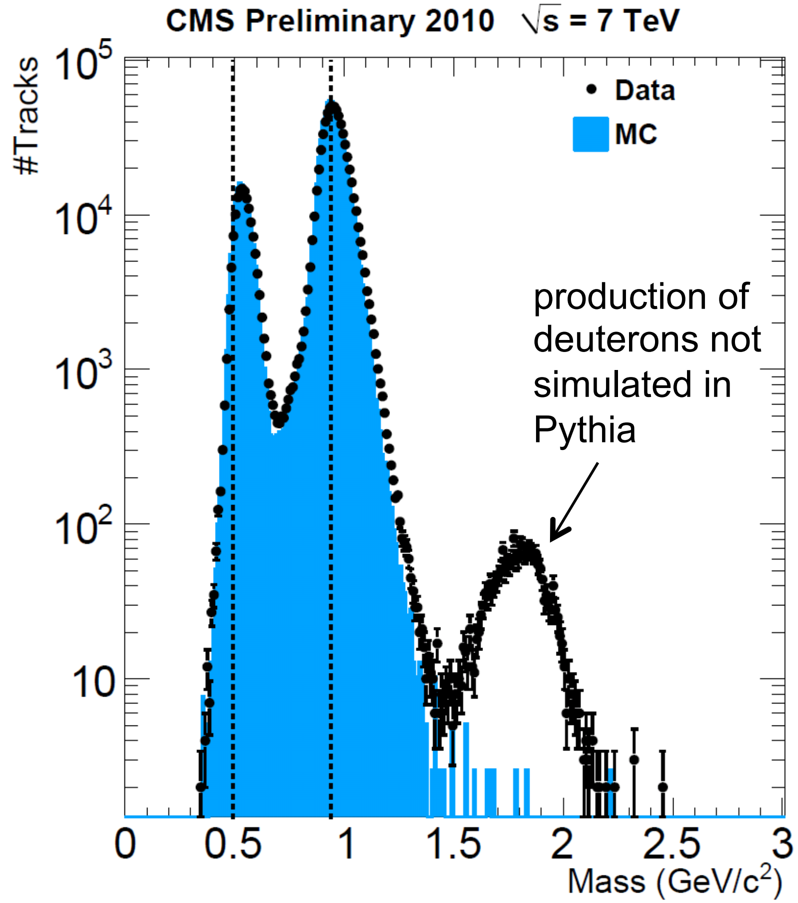
Extracting the material budget



- Other methods also employed:
  - track multiple scattering, momentum scale, etc.
- Agreement between photon conversions and nuclear interactions on the location and composition of materials gives us good confidence that the simulation geometry is an accurate representation of the real detector
- Uncertainties in the amount of material and its distribution are estimated to of order 5% (CMS PAS: TRK-10-003)

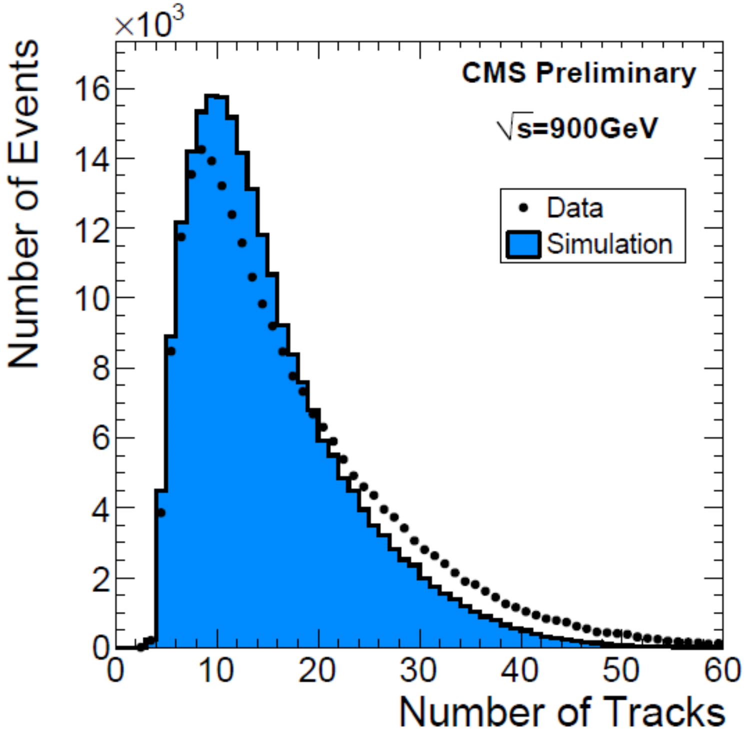
Tracker dE/dx Simulation

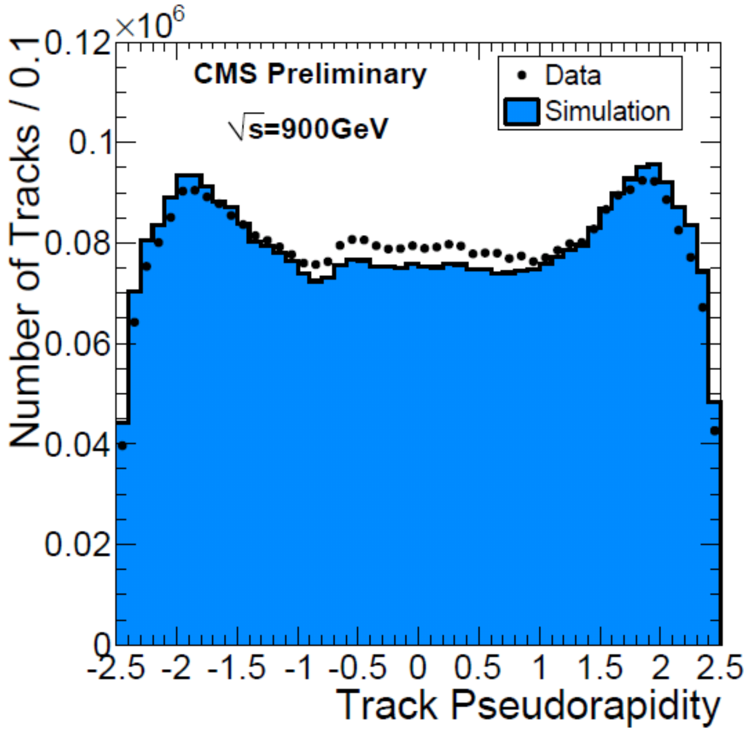




- Signal simulation in tracker includes charge propagation, charge collection efficiencies, saturation effects, and tracker noise modeling
  - tuned on cosmic data and early collisions
- detailed test of Geant4 descriptions of energy loss mechanisms in tracker material

Charged particle multiplicity

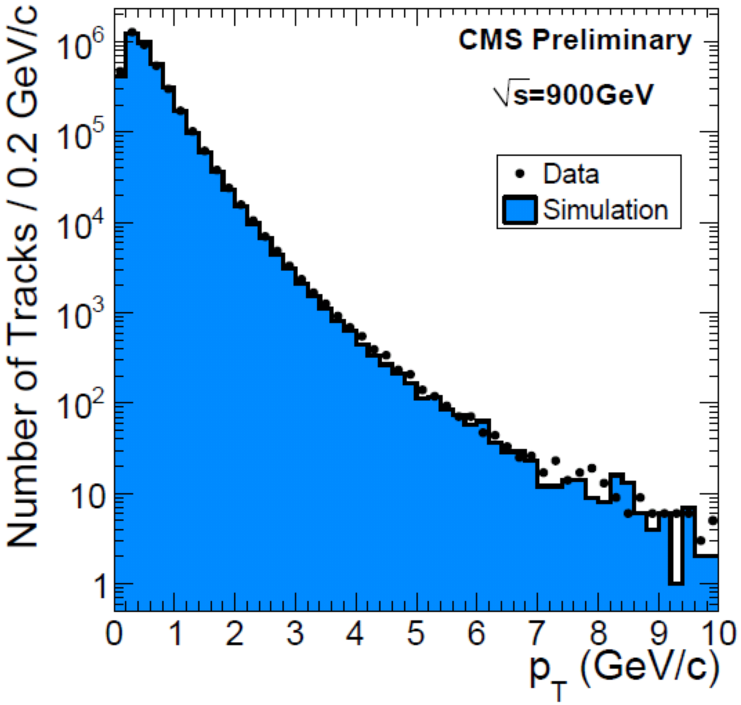


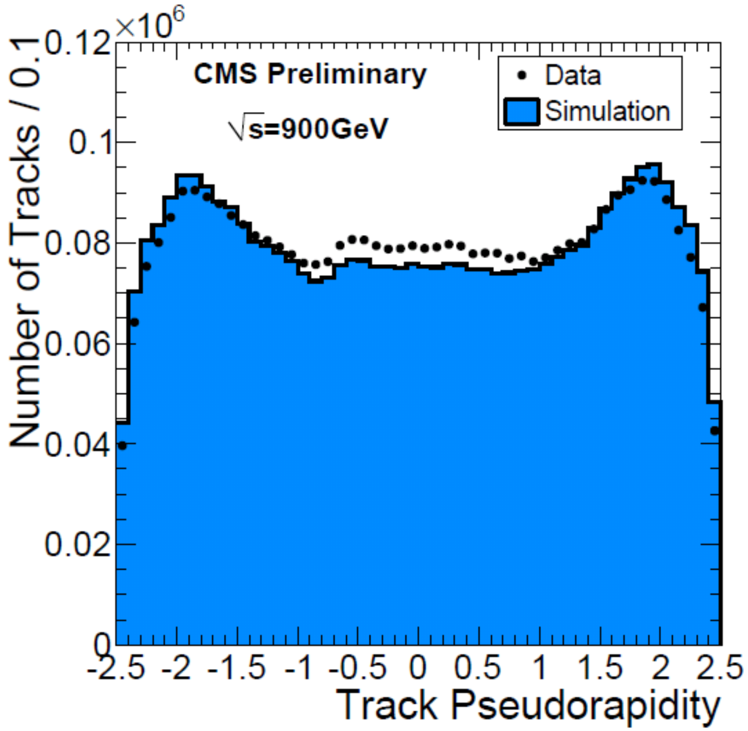


- Minimum Bias events
- Original Pythia 6.4 tunes largely divergent from Data distributions (tune D6T)
  - charged particle multiplicity very different
  - Surprising, given previous Tevatron studies

CMS PAS: TRK-10-001

Charged particle multiplicity

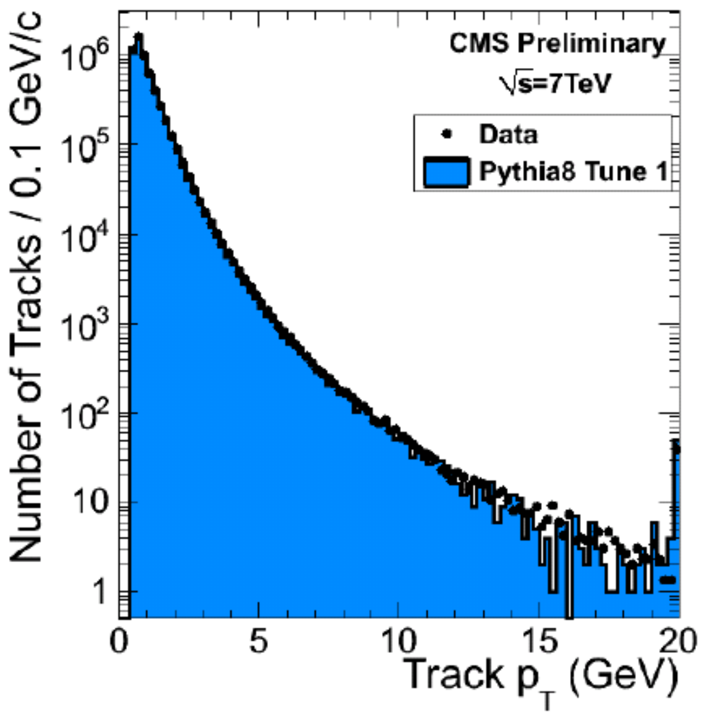


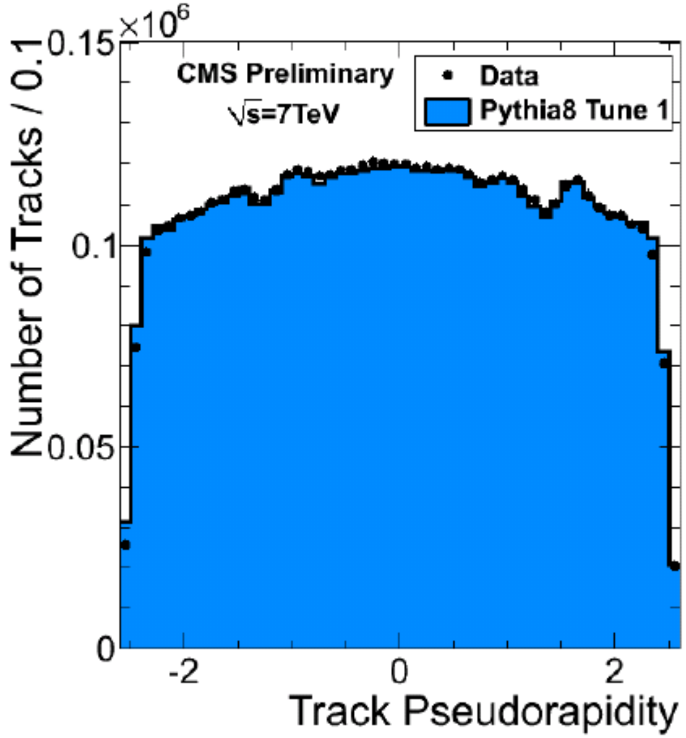


- Minimum Bias events
- Original Pythia 6.4 tunes largely divergent from Data distributions (tune D6T)
  - charged particle multiplicity very different
  - Surprising, given previous Tevatron studies

CMS PAS: TRK-10-001

Charged particle multiplicity



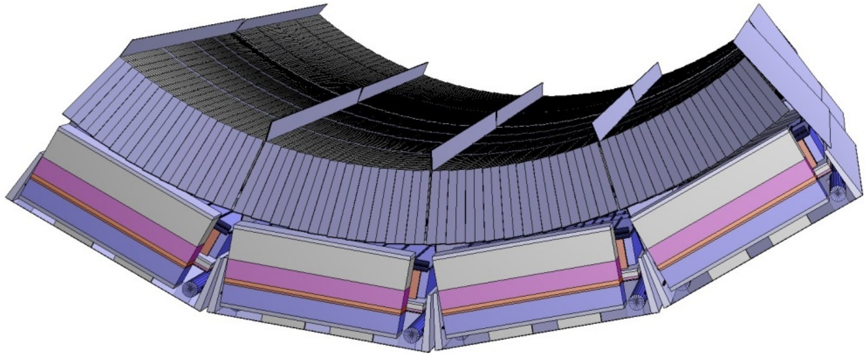


- Minimum Bias events
- Original Pythia 6.4 tunes largely divergent from Data distributions (tune D6T)
  - charged particle multiplicity very different
  - Surprising, given previous Tevatron studies
- New Pythia 8, Tune 1 gives much better agreement
  - new: hard scattering in diffractive interactions
    - relative increase in population of high-p_T, "forward" regions

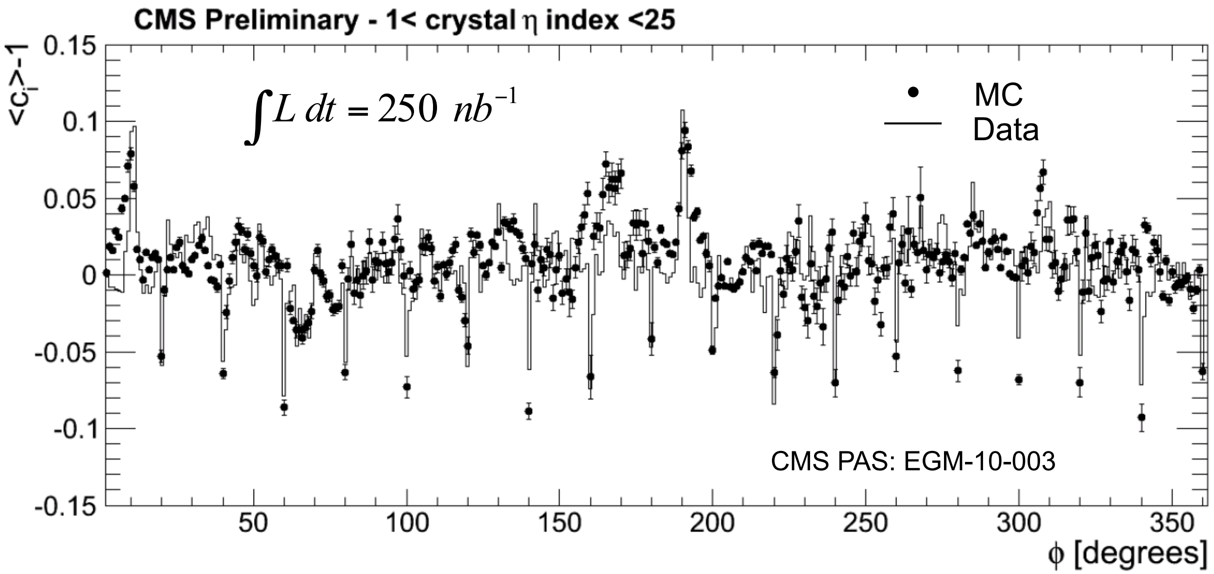
CMS PAS: TRK-10-005

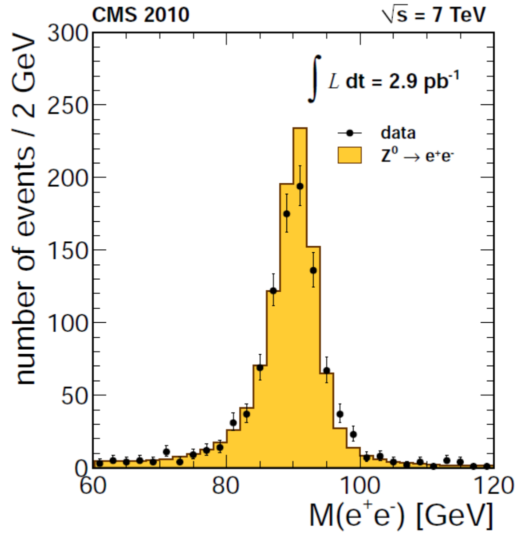
Electromagnetic Calorimeter Calibration

- Based on expected ϕ symmetry of energy deposition in minbias events
- Non-uniformity of response correction caused by inter-module gaps and different distributions of material in front of the calorimeter



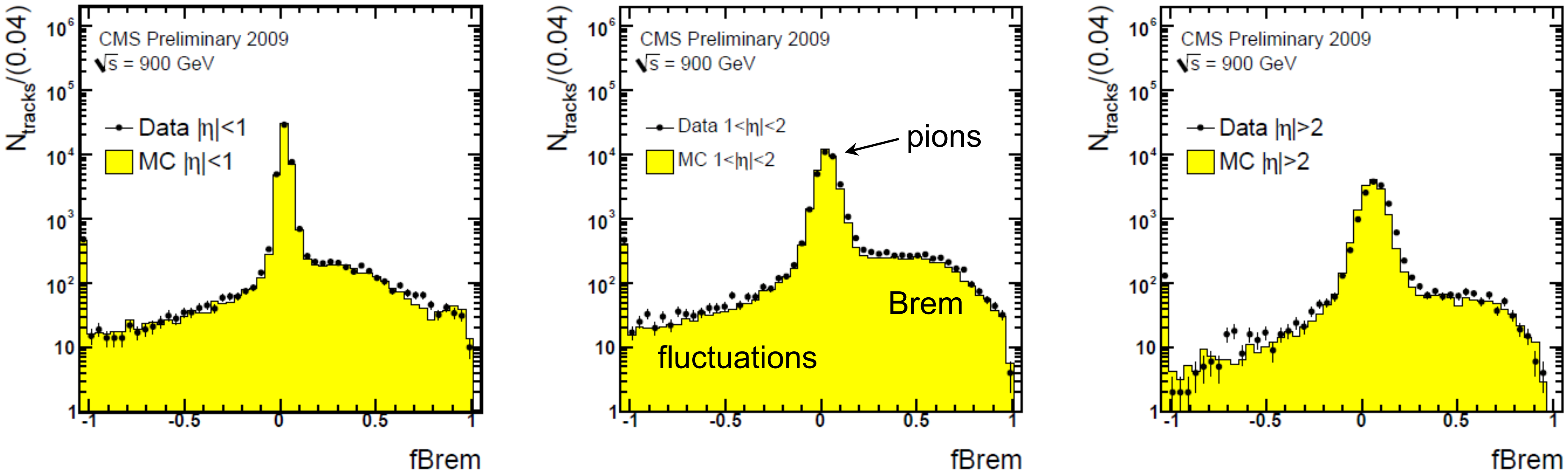
- Here: response correction for each calorimeter module:





Electromagnetic Interactions

- Fraction of energy loss in the Ecal for different ranges in η: $f_{\mathrm{Brem}} = (p_{\mathrm{in}} - p_{\mathrm{out}})/p_{\mathrm{in}}$
  - inclusive distributions based on high-purity track selection
    - Gaussian Sum Filter track fit to account for energy loss



- MC Minbias events; again remarkable MC/Data agreement

CMS PAS: EGM-10-001

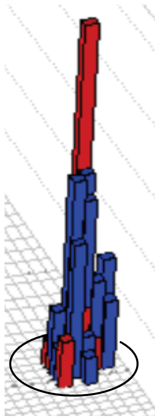
- depends on accurate modeling of:
  - material distributions
  - showers
  - correct distribution of particle types in Data and MC

Jet-Finding at CMS

• Calorimeter Jets

Jets clustered from ECAL and HCAL deposits (Calorimeter Towers)

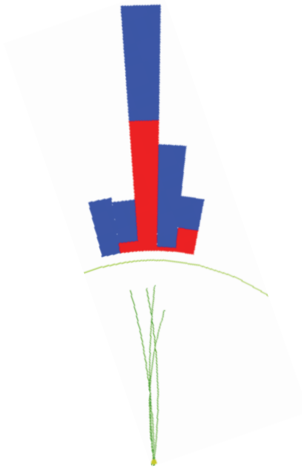
Correspondingly: Calo MET



• Jet-Plus-Track Jets (JPT)

Subtract average calorimeter response from CaloJet and replace it with the track measurement

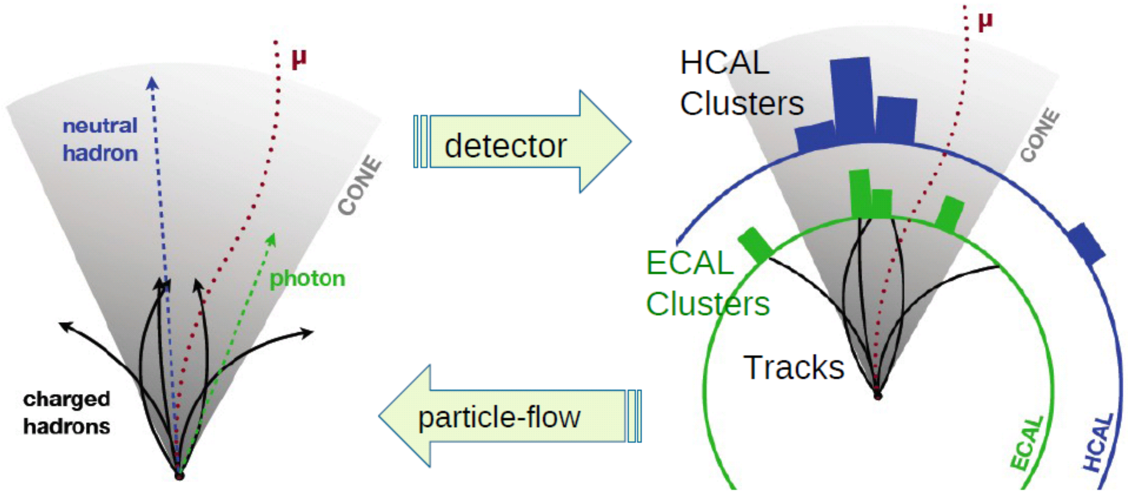
Correspondingly: Tc MET



• Particle Flow Jets (PF)

Cluster derived Particle Flow objects: unique list of calibrated "particles" representing "generator level" information

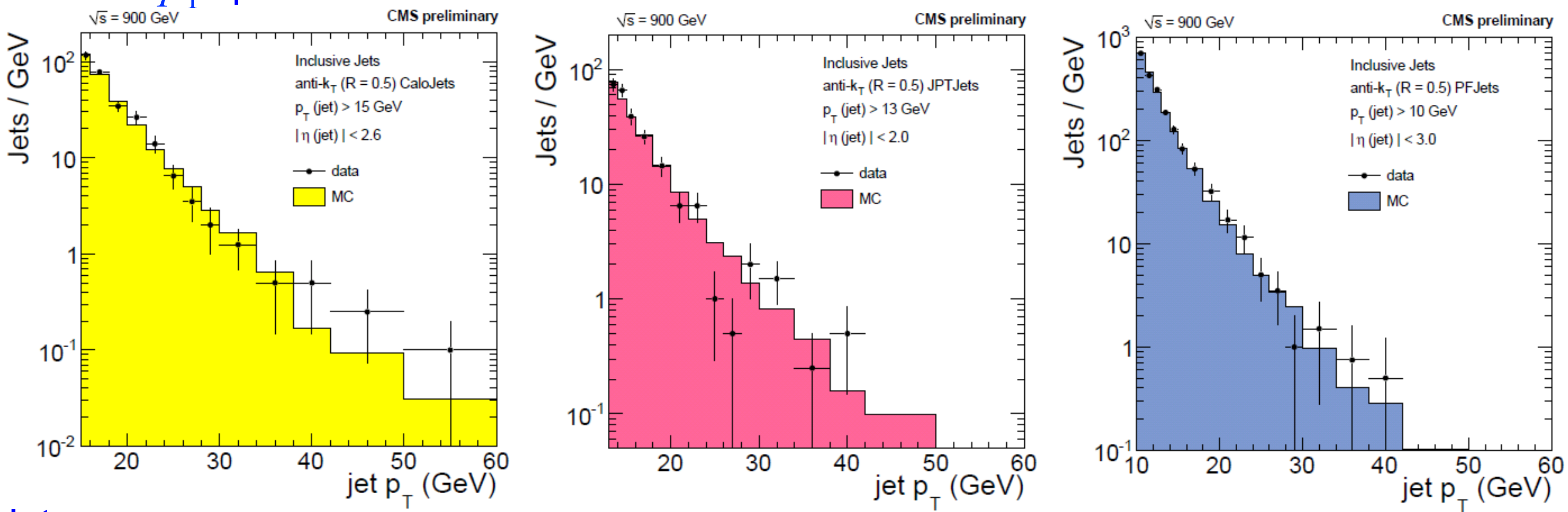
Correspondingly: PFMET



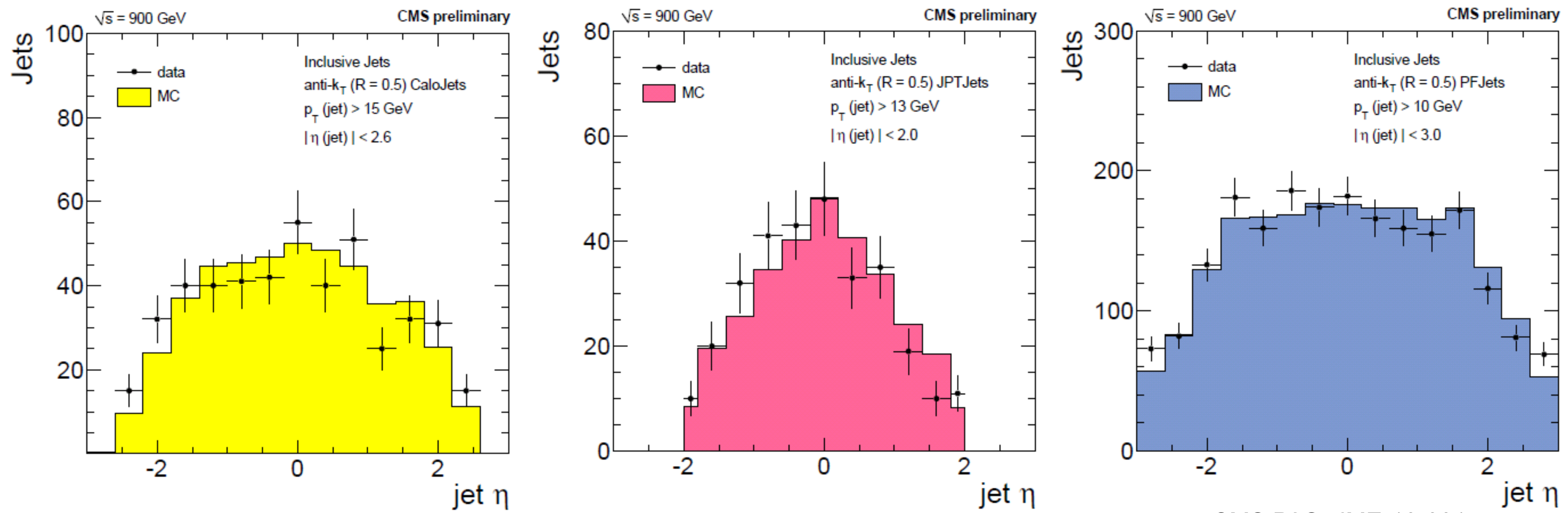
Default jet clustering algorithm: Anti-k_T with R = 0.5

Jet Variables

- Inclusive p_T spectrum

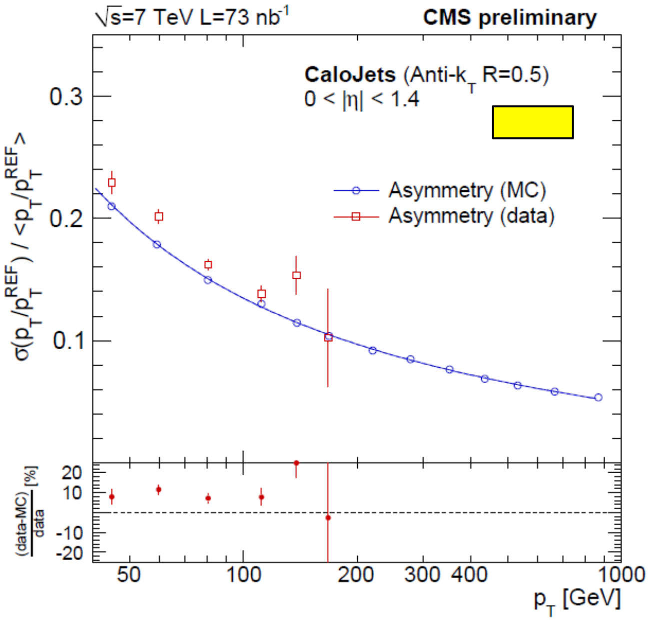


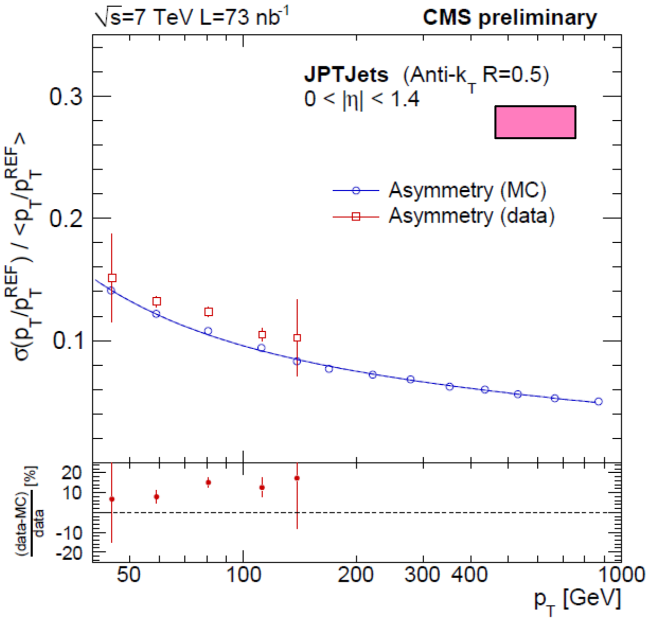
- Jet η



CMS PAS: JME-10-001

Jet Resolutions

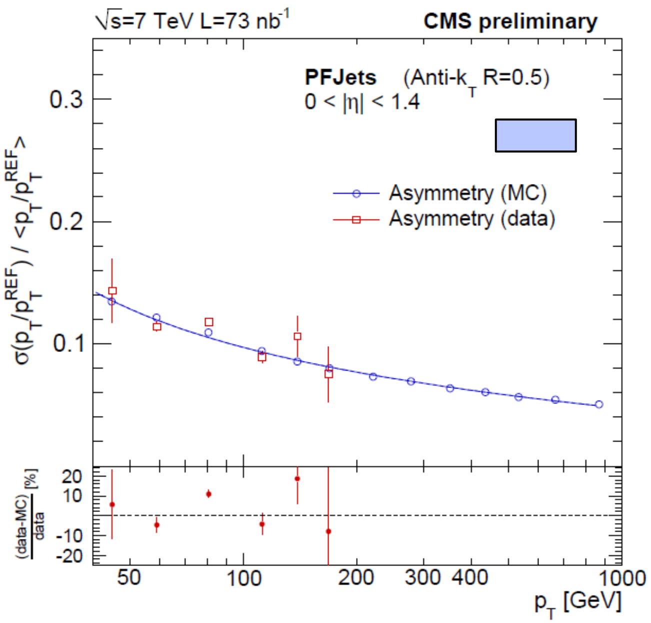




Di-jet Asymmetry Method:

$$A = \frac{p_T^{jet1} - p_T^{jet2}}{p_T^{jet1} + p_T^{jet2}}$$

$$\Rightarrow \frac{\sigma(p_T)}{p_T} = \sqrt{2}\sigma_A$$



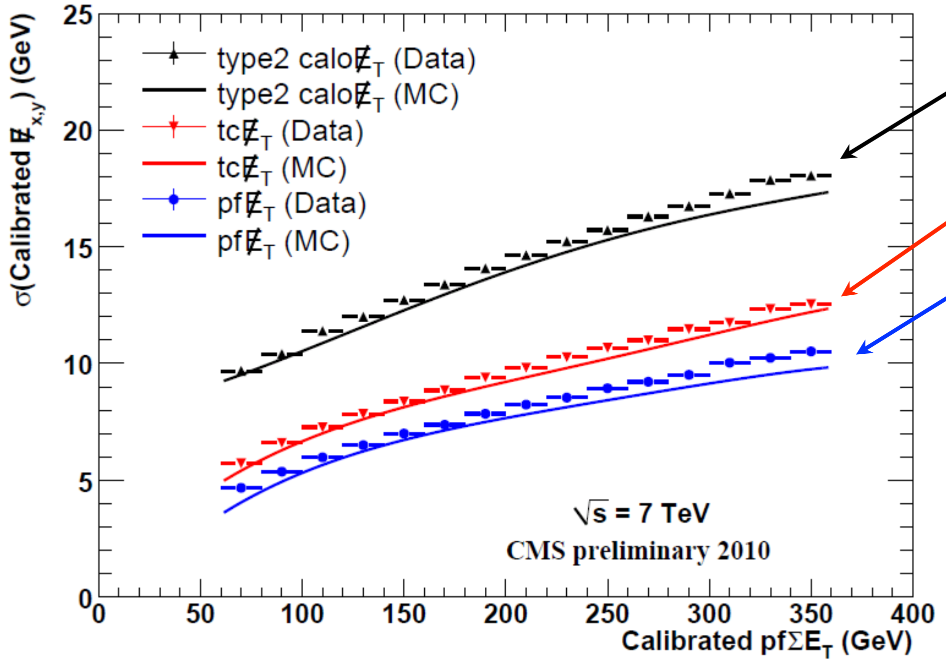
- Resolution as a function of average p_T
  - extrapolated to zero additional activity in each bin by measuring σ_A at decreasing values of the third jet p_T^{jet3}
  - same treatment applied to QCD MC
  - within 10% agreement for all three jet algorithms
  - validates combination of generators, material description, shower models

CMS PAS: JME-10-003

UNIVERSITY OF NOTRE DAME
Fermilab

Missing E_T: Resolution Studies

- MET provides a stringent test of noise simulation, showering, and resolution modeling ⇒ all elements have to be correct



Calorimeter MET

Track-corrected MET

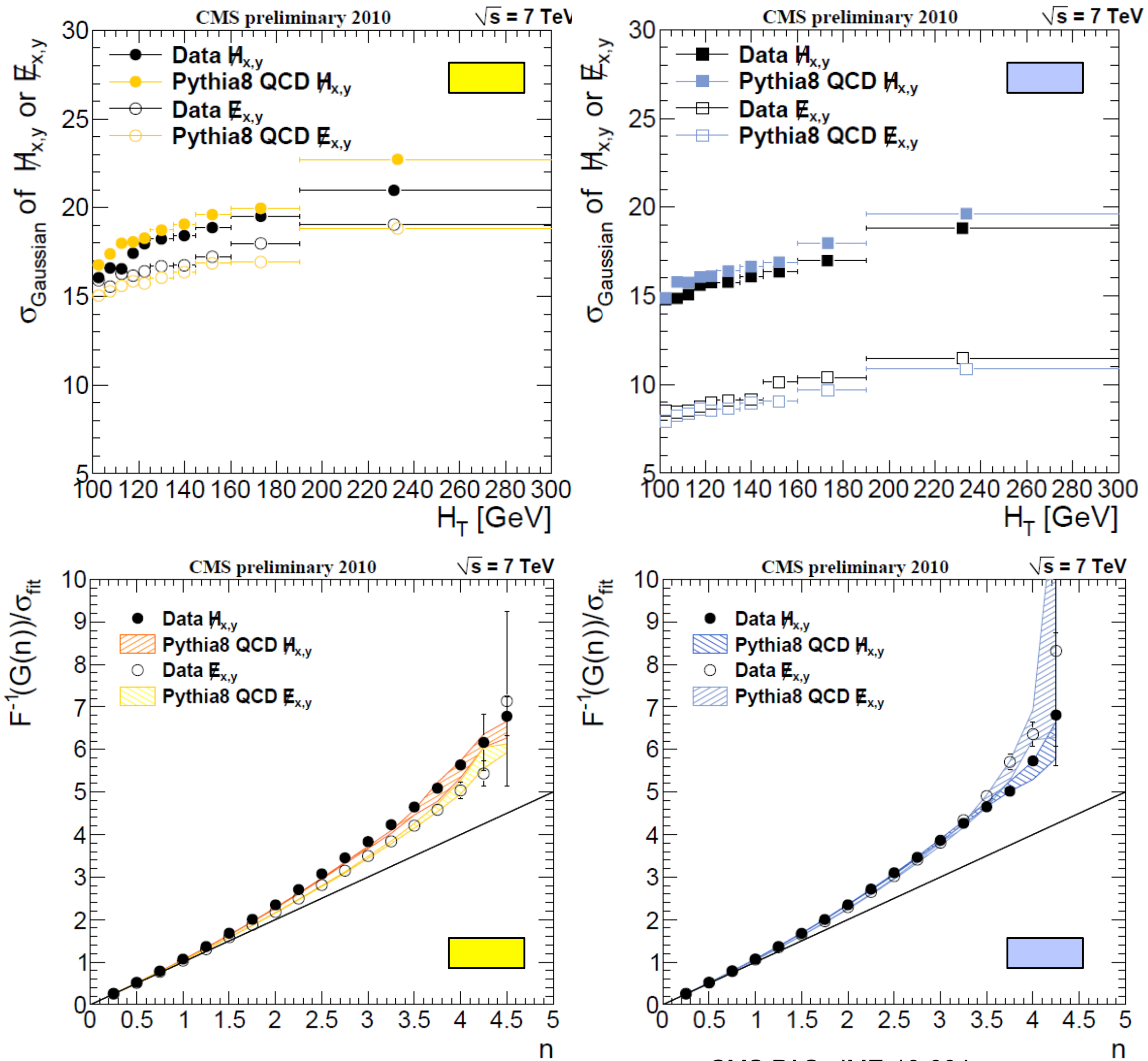
Particle Flow MET

(Jet Energy Scale corrections are applied to all jets with p_T>30 GeV; remaining unclustered energy is corrected using a scale derived from the hadronic recoil opposite Z → ee events)

- Here, $\not{E}_T$ resolution is measured using the width of the $\not{E}_x$, $\not{E}_y$ distributions
  - overall $\not{E}_T$ calibration from transverse energy balance in γ+jet events
  - at least two jets of p_T > 25 GeV required
  - identical MC/Data treatment

CMS PAS: JME-10-004

Comparisons of Calorimeter Resolutions



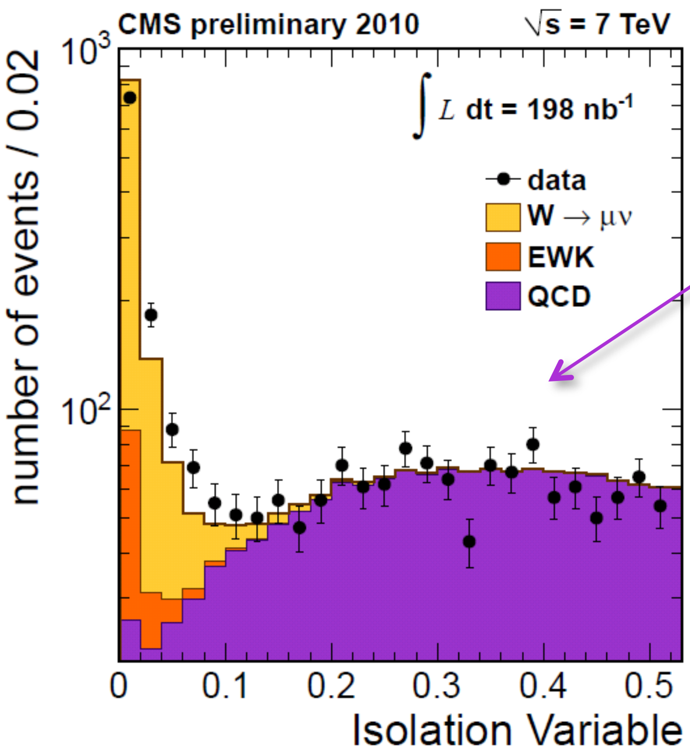
CMS PAS: JME-10-004

- Resolution of MET and H_T (total jet transverse energy) for Calorimeter Jets and Particle Flow Jets
  - H_T potentially more robust
  - multi-jet events
    - leading jet p_T > 40 GeV
  - width of central gaussian in $\not{H}_x, \not{H}_y$ and $\not{E}_x, \not{E}_y$
- Characterization of the non-Gaussian nature of the tails:
  - important for searches
  - plot shows width of data distributions (in sigma) containing nσ of a gaussian
  - deviation from gaussian form outside of 2σ

UNIVERSITY OF NOTRE DAME
Fermilab

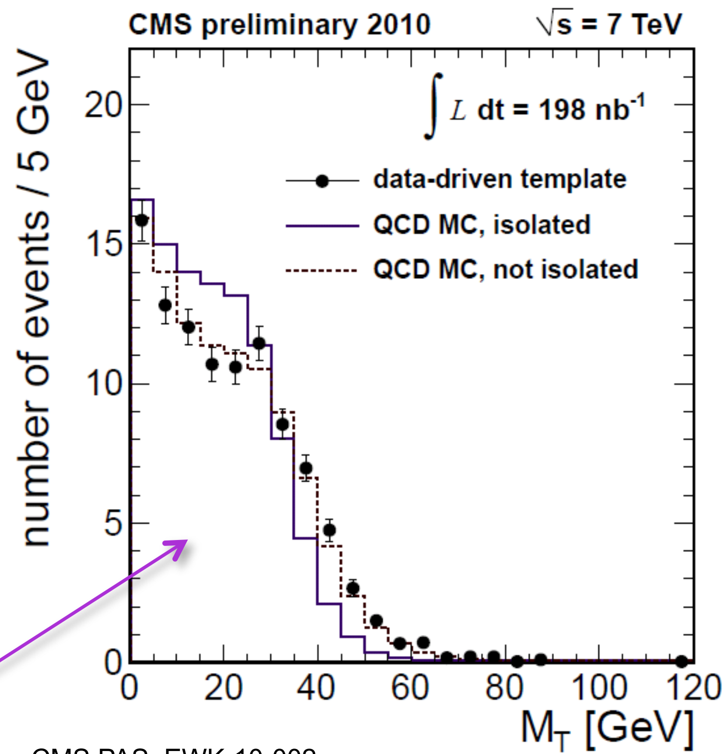
Muon System

- Energy deposition characterized by proper modeling of the absorber interaction lengths
  - punch-through, decays in flight account for much of the fakes
    - isolation variable critical to differentiate signal from QCD
  - future: significant backgrounds from neutron interactions



Sum of track, Ecal, Hcal energy in cone of $\Delta R < 0.3$ around muon

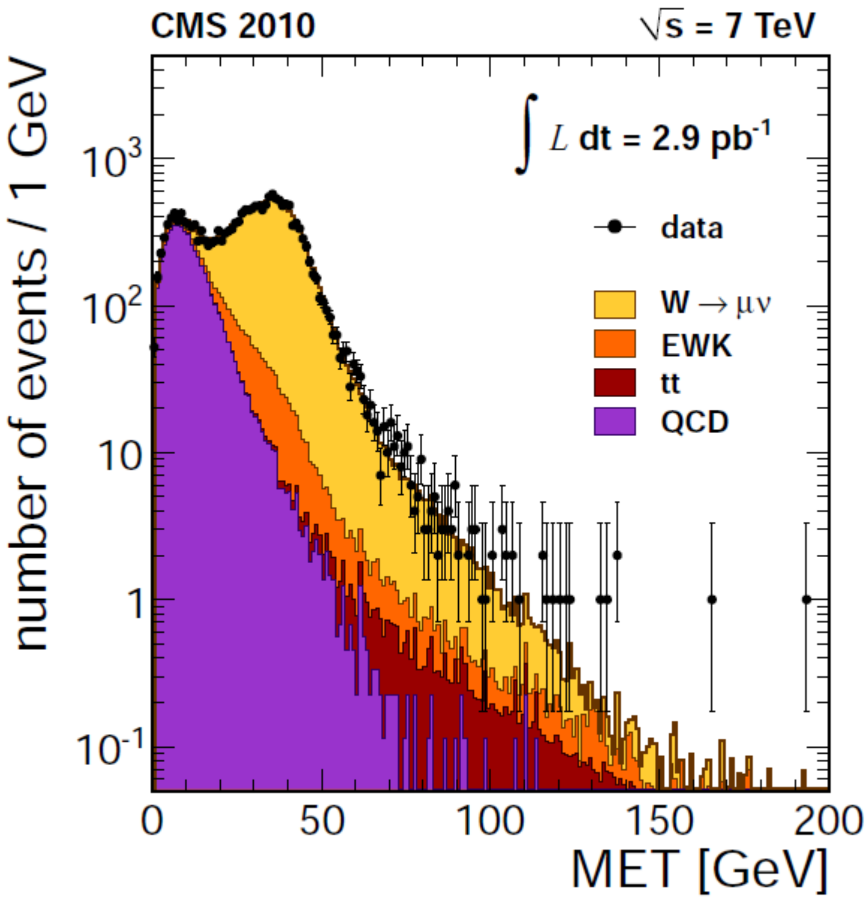
Comparison of isolation-inverted data sample with expectations from QCD MC

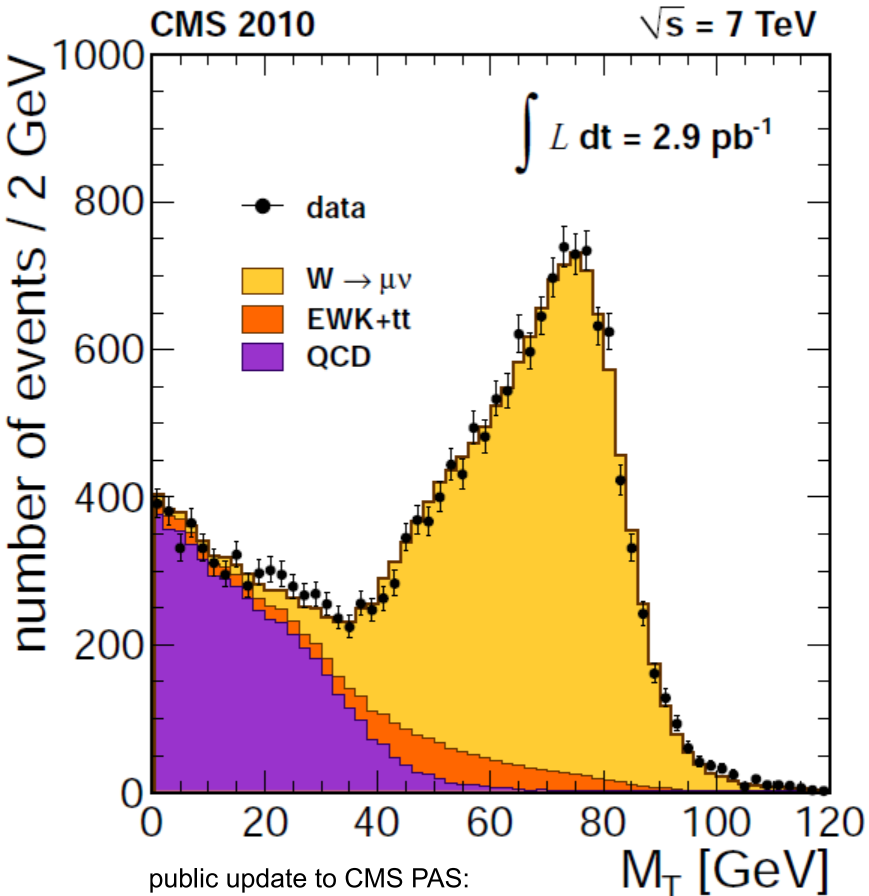


CMS PAS: EWK-10-002

Electroweak Distributions

- MET also well-modeled in $W \rightarrow \mu\nu$ events





public update to CMS PAS: EWK-10-002

- Here: combination of intrnisic resolution and generator models

Conclusions

- “Tuning” of CMS Simulation has been a multi-year process
  - based on Test Beam data, extensive Cosmic running
  - improved by comparisons using collision data
    - no substitute for the real thing…
- “Validation” ongoing
  - continual refinement as the dataset grows
    - higher statistics comparisons possible for a growing number of studies
- Current (excellent) level of Data/MC agreement is a product of a huge amount of work over many years by many people
  - not an accident!