Particle Flow

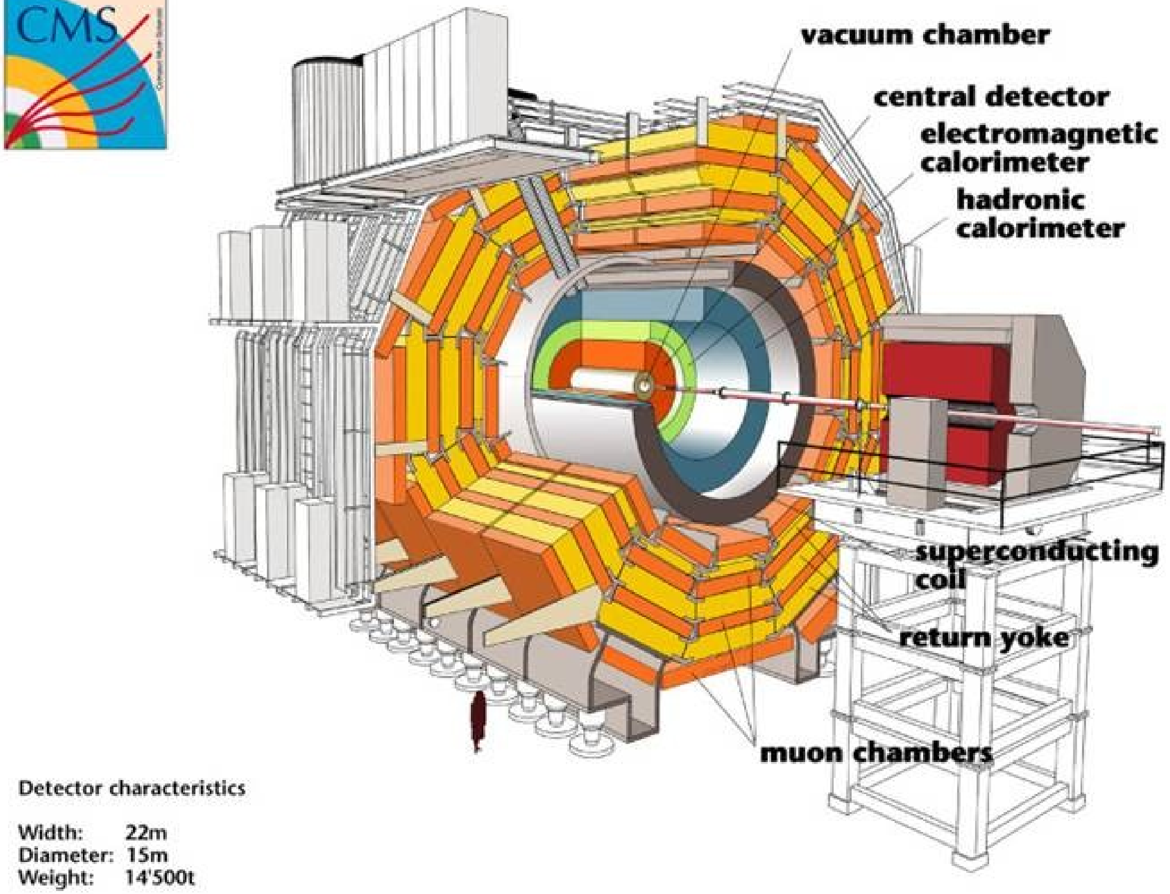
(Slides for morning discussion on particle flow)

Satyaki Bhattacharya
Saha Institute for Nuclear Physics

Plan of the talk

- CMS detector
- Particle flow overview
- Particle flow steps
  - PF elements
  - PF linking and blocks
  - PF reconstruction of candidates
- Performance

The CMS Detector

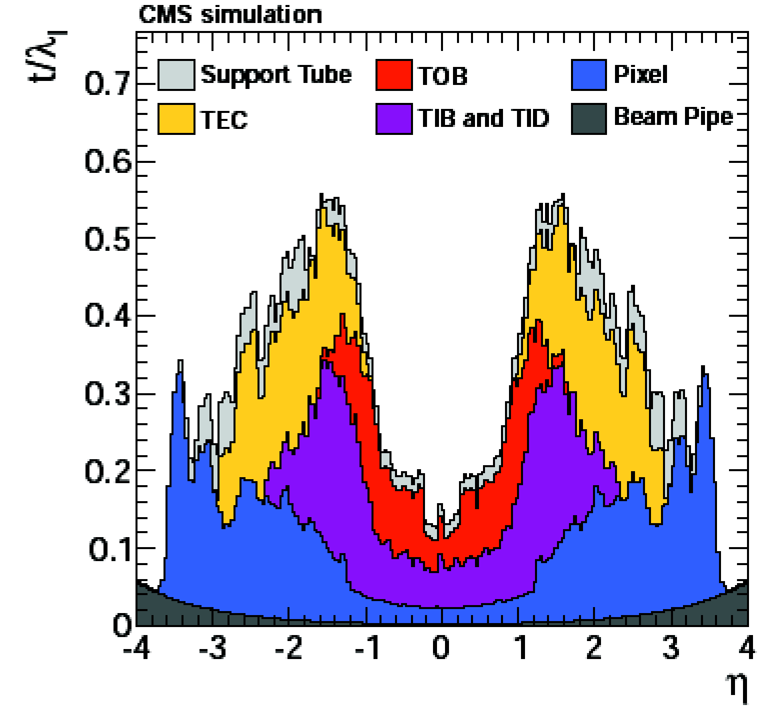


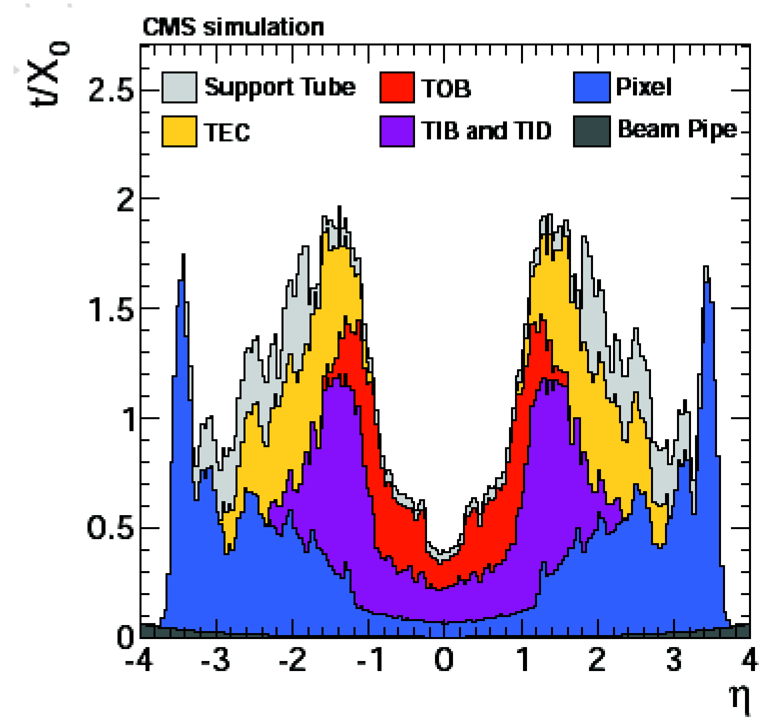
CMS design goals

- Muon chambers (and tracker):
  - Good muon id over a wide pT range in |eta|<2.5,
  - ~1% @ 100 GeV dimuon mass resolution
  - Charge id upto p < 1 TeV
- Tracker:
  - Good charged particle momentum resolution and reconstruction efficiency
  - Tau, b jets tagging --> pixel layers close to interaction
- EM calorimeter
  - Dielectron mass resolution ~1% @ 100 GeV
  - Coverage |eta| < 2.5
  - Pi0 rejection, isolation at high luminosities
- Hadron calorimeter
  - Good missing ET and dijet mass resolution
  - Hermitic coverage |eta| < 5, good dijet mass resolution, |dEta|X|dPhi| = (0.1 X0.1)

Magnet, tracker

- Magnet: 3.8 Tesla, 3.18 meter free bore radius
- HCAL, ECAL inside, no showering in magnet before calorimetry
- 66M 100X150 micron pixels and 9.6M 80 to 180 micron pitch strips within 1.2 m radius.
- Capable of closely spaced tracks within a jet
- At |eta| = 1.5 probability of a photon converting is 85%
- And a hadron doing a nuclear interaction is 20%
- Major challange to overcome for particle flow





ECAL

- Fine-grained (0.0175X0.0175 in etaXphi), clearly separated energy deposits from particles in a jet up to jet pT of the order of a TeV
- excellent resolution: ~0.3% for high pT photons and electrons

$$\frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%$$

HCAL

- good geometric coverage ($|\eta|<5$)
- can separate charged and neutral hardon energy deposits up to a jet pT of 200-300 GeV
- modest resolution: ~9% for high pT jets
- fine-grained, clearly separated energy deposits from particles in a jet up to jet pT of the order of a TeV

$$\frac{110\%}{\sqrt{E}} \oplus 9\%$$

Muon chambers

almost perfect identification of muons

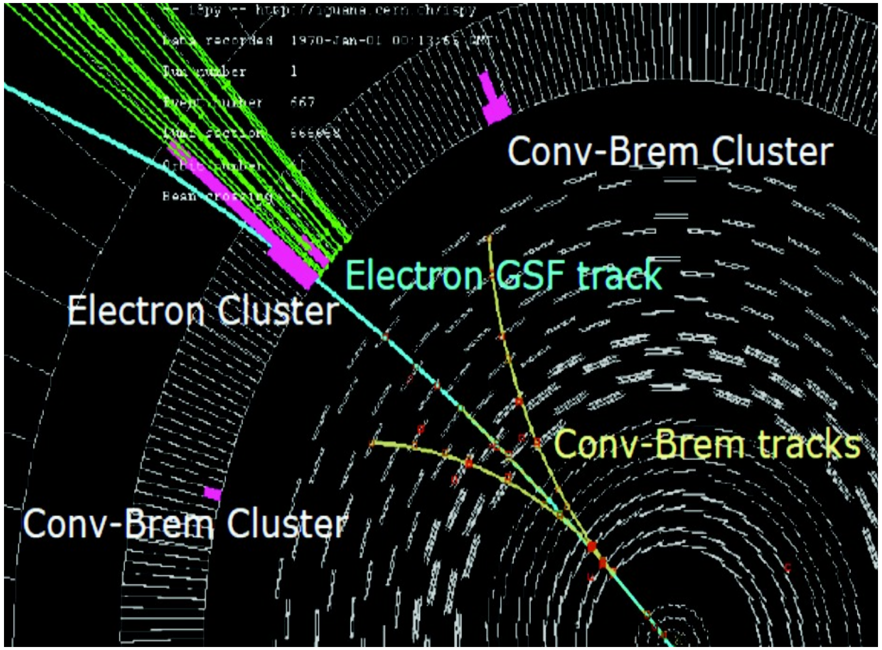
versatile muon tracking

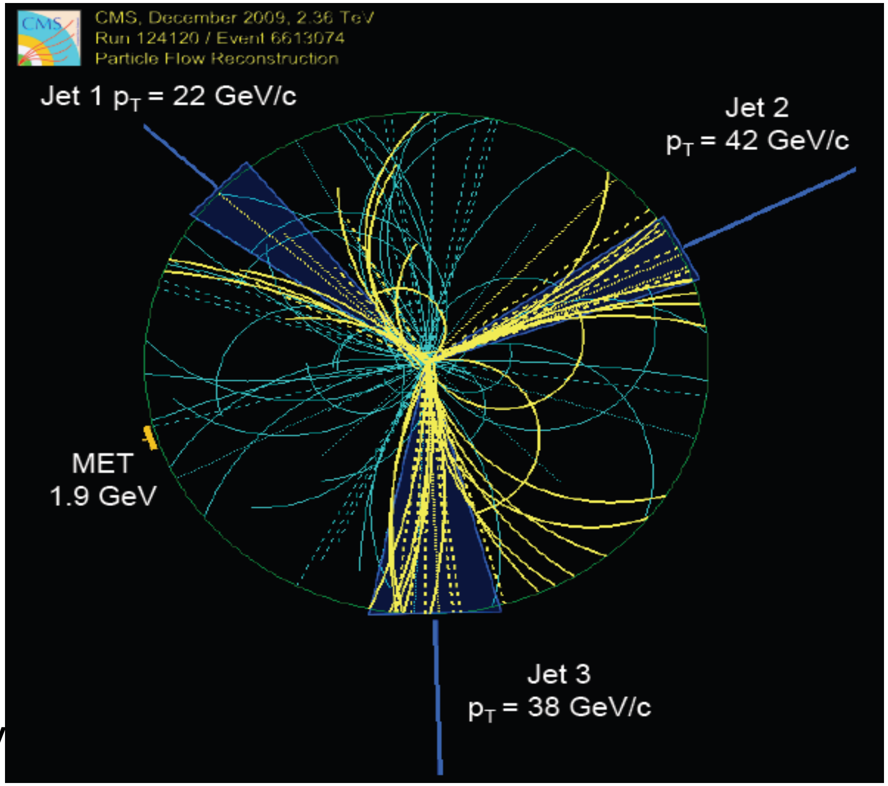
good dimuon mass resolution (~ 1% at 100 GeV)

unambiguous charge determination of muons with momentum < 1 TeV

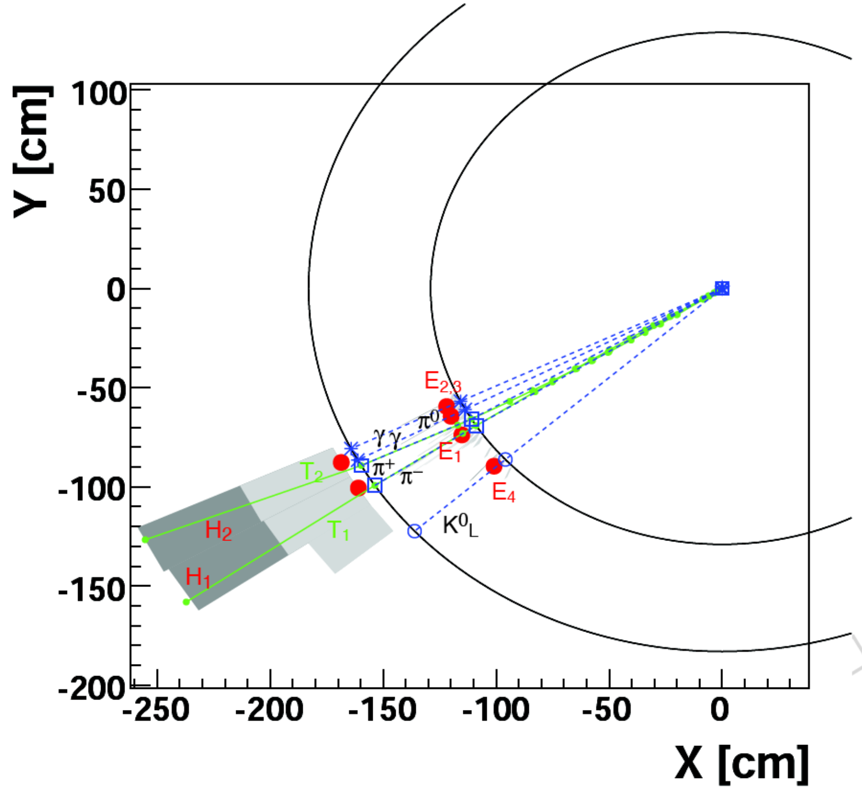
Particle Flow

- Reconstruct all stable particles in CMS detector by linking responses of subdetectors
  - Photon, electron, muon, charged and neutral hadrons
  - Resulting list of particles can be used as if they came from a MC generator
  - Composite objects like jets, taus, MET can be reconstructed from the “PF candidates”

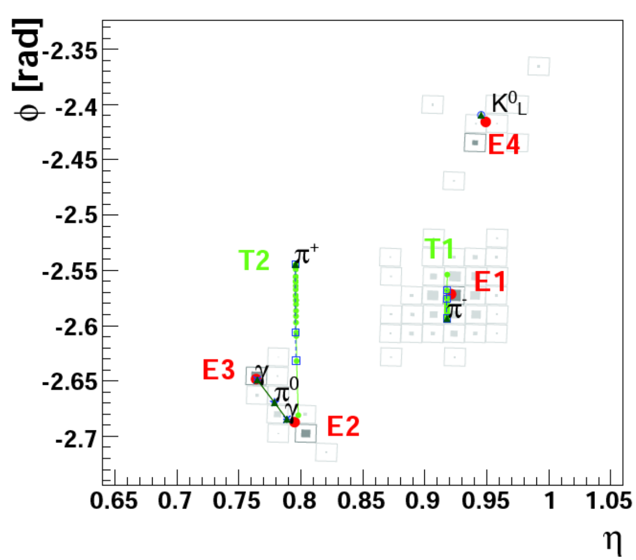




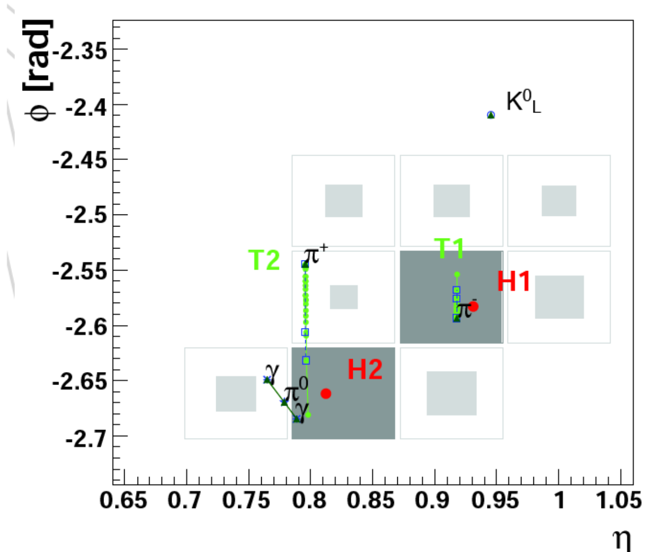
A simple example



(a) The (x, y) view



(b) The (η, φ) view on ECAL



(c) The (η, φ) view on HCAL

Particle Flow in CMS: History

(from a slide by Albert de Roeck, 2016)

- CMS was not designed having PF explicitly in mind.
- Interest started to develop ~ 2007 realizing the power of the tracker and ECAL (granularity) and the gain of PF
- The HCAL resolution in CMS is modest (2x worse than eg in ATLAS), hence important to reduce the impact of pure calo measurements.
- Particle Flow evolved with the years and was already validated on first MB data in 2009/2010 for the initial analyses. Now more than 90% of the analyses in CMS use full or partial PF (especially for jets and MET). We call it the "Global Event Description"
- Also used in Heavy Ion collisions analysis and at high pile-up! Planned for HL-LHC running with ~140 PU
- Used in part in the trigger (High-Level Trigger) eg jets, taus

PF elements

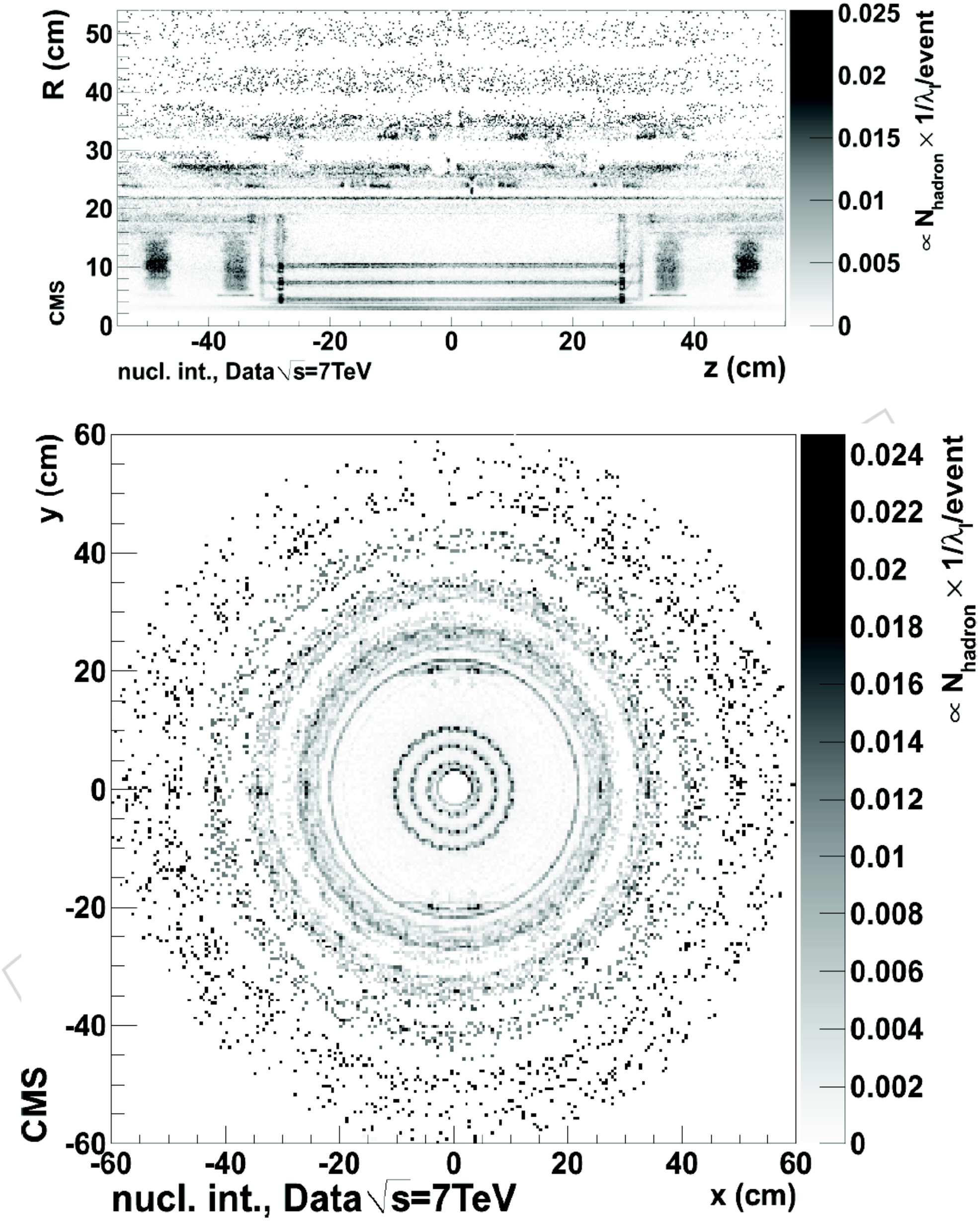
- MuCh tracks
- Tracker tracks
- ECAL clusters
- ES clusters
- HCAL clusters

Traditional tracks

- Kalman filter based tracking
- Seed: two consecutive hits in three pixel layers
- >= 8 hits, at most one missing hit on the way
- Each hit contributes < 30% to track fit chi2
- XY-dca few mm
- PT > 0.9 GeV

interactions in tracker

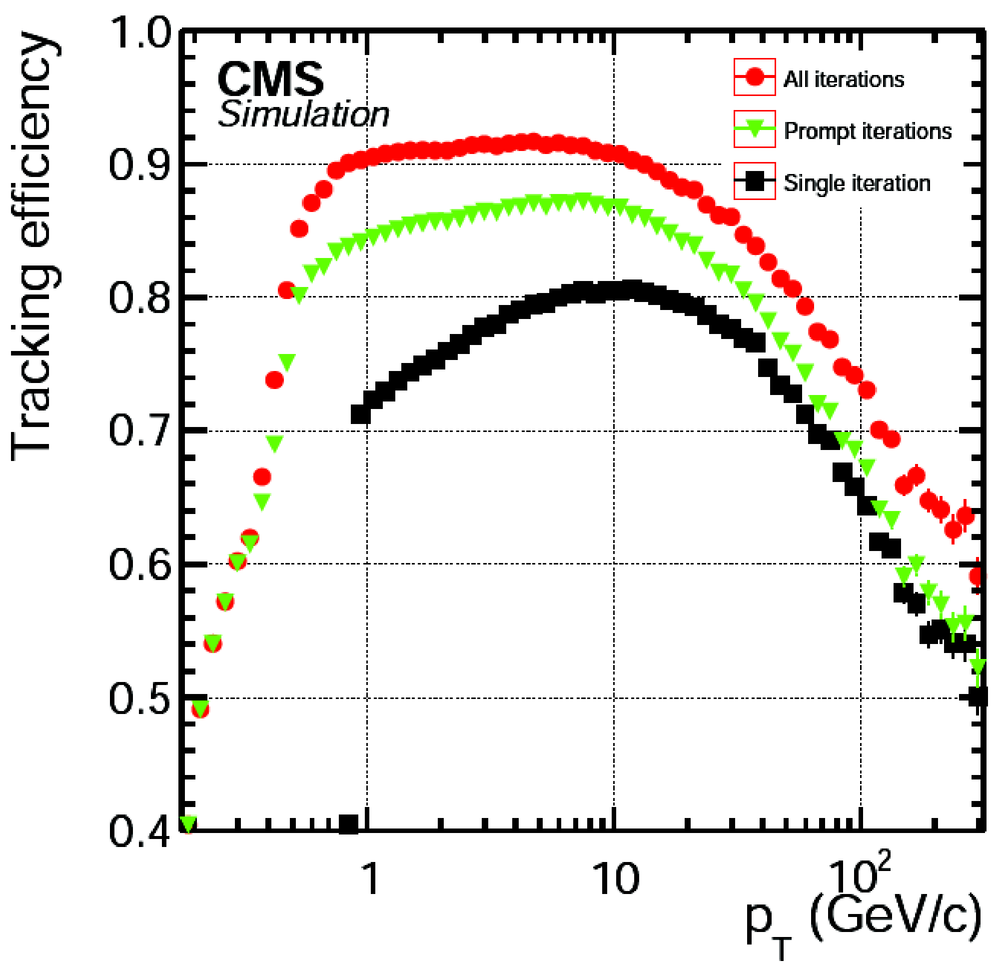
- Nuclear interactions can either produce a kink or give secondaries
- 2/3 of secondaries charged
- Will give displaced tracks
- Conversion of photons also will give displaced tracks

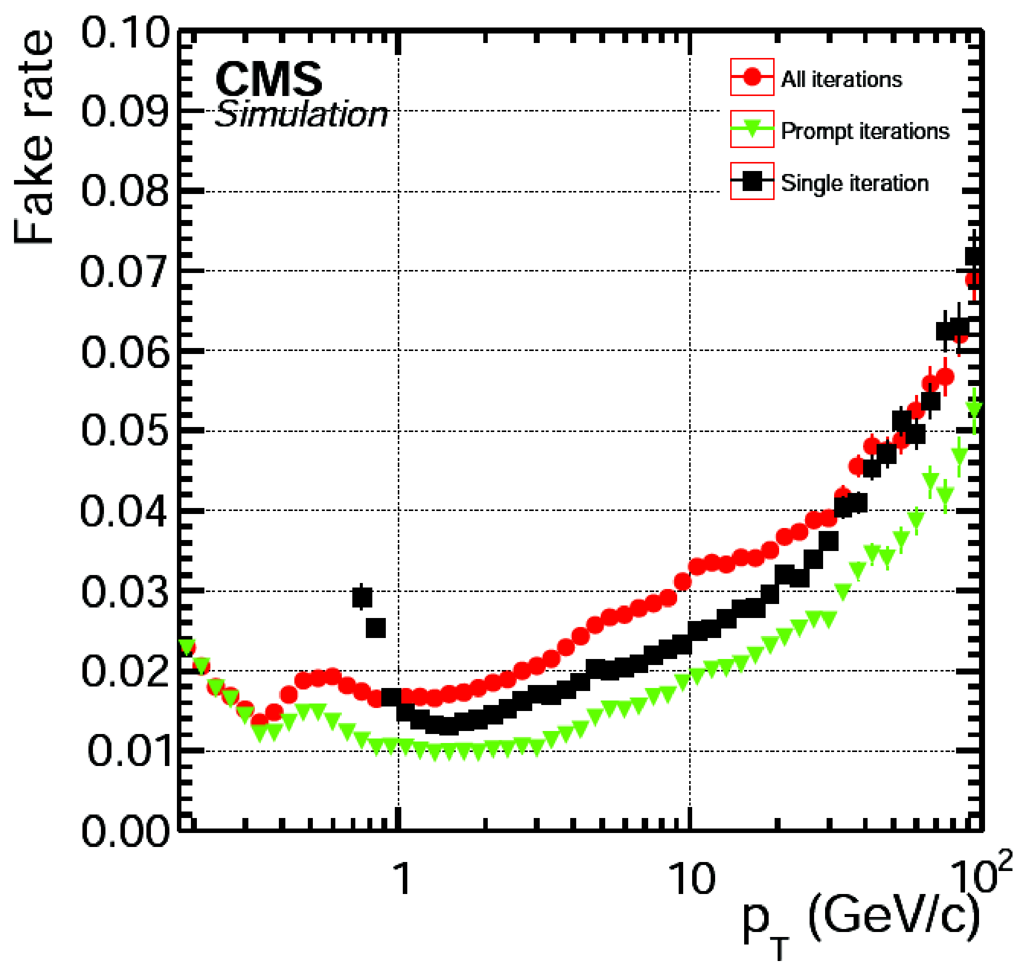


performance

- Fake rate few percent, efficiency for pi+- > 1 GeV 70-80%
- Probability of nuclear interaction before 8 hits 15-30% (loss of track)
- Tracking efficiency falls rapidly at higher pT
  - Limited by strip pitch for overlapping particles
  - Important loss for boosted and collimated jets
- About 2/3 of jet energy from charged tracks
- Inefficiency of 15% would increase neutrals by 10%, could worsen the energy resolution by 50%
- Also will bias the jet direction (shifted cluster position)

Efficiency and fake rate





Iterative tracking

- Loosen pT threshold, require fewer hits --> recover half of the tracks.
- Fake rate increases 5 times with pT threshold lowered to 300 MeV
- + require 5 hits --> fake rate 80%
- Solution: iterative tracking
- Start with tight tracks. Remove used hits. Relax criteria, do tracking with the remaining hits

Iterative tracking

- Loosen pT threshold, require fewer hits --> recover half of the tracks.
- Fake rate increases 5 times with pT threshold lowered to 300 MeV
- + require 5 hits --> fake rate 80%
- Solution: iterative tracking
- Start with tight tracks. Remove used hits. Relax criteria, do tracking with the remaining hits

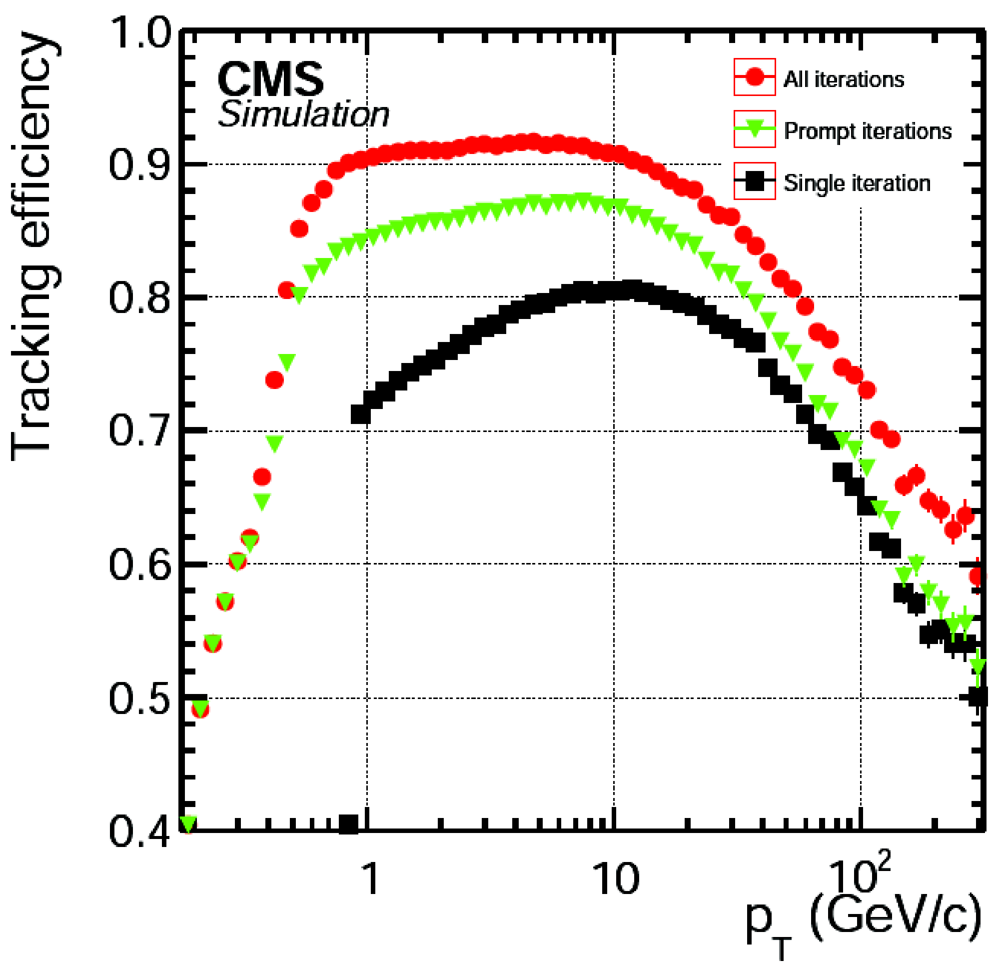
Iterative tracking

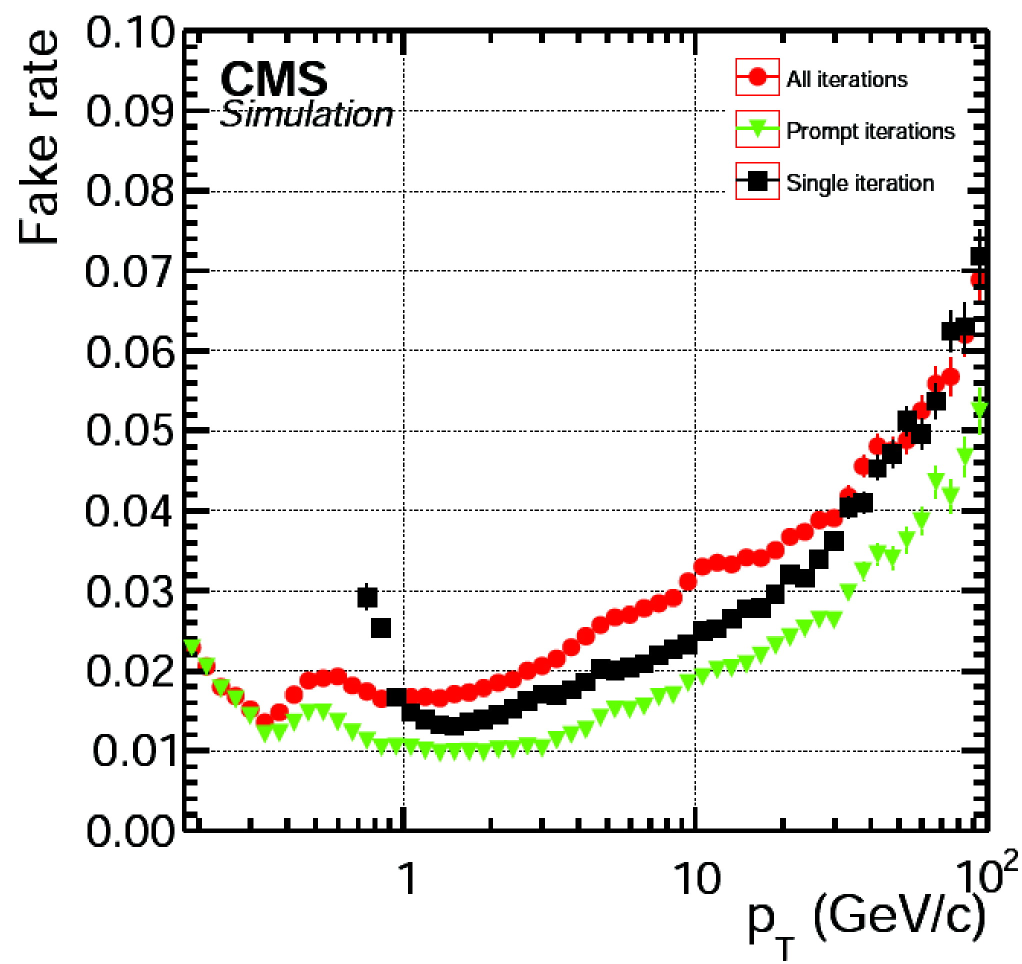
- Seeded with pixel triplet for prompt, high pT
- Pixel triplet for displaced R<5cm
- Pixel triplet for prompt, low pT
- Pixel pair for recovering high pT
- pixel+strip triplet for displaced R<7cm
- Pixel+strip pair for displaced R<25cm
- Pixel+strip pair for displaced R<60cm
- pixel+strip pair for very high pT inside high pT jets
- Muon tagged tracks for recvering muons
- Muon chamber for recovering muons

Iterative tracking performance

- First 3 iterations removes 40%(20%) hits from pixel(strips) retaining 80% efficiency
- Steps with >= pixel hit (0-3,7) recover 50% of the tracks
- Lowers pT threshold from 900 MeV to 200 MeV
- Steps 5,6 recover neuclear interaction tracks, another 5% but adds 1% in fake rate
- Twice faster
- Fakes still an issue, addressed in later stages

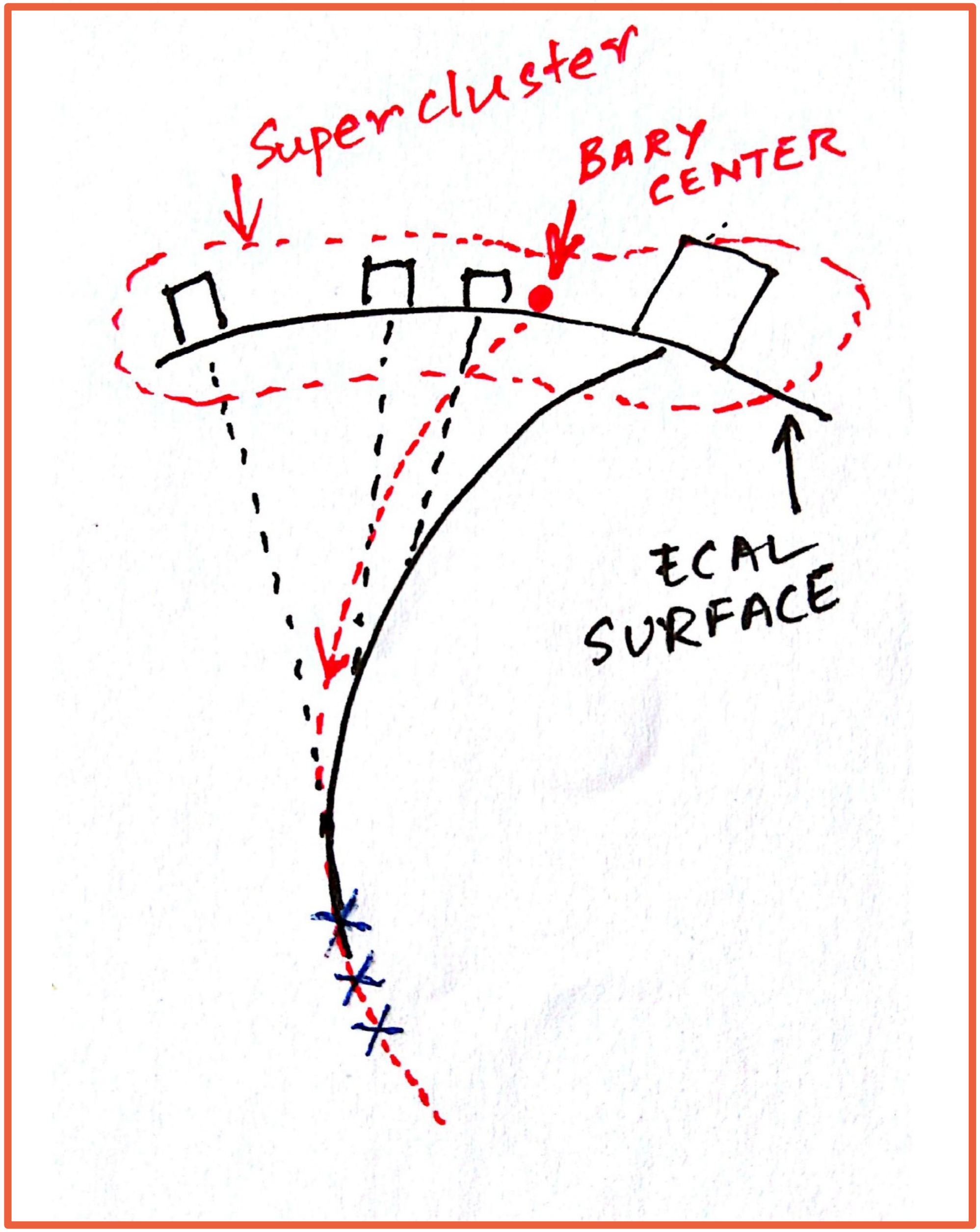
Efficiency and fake rate





Original electron tracking

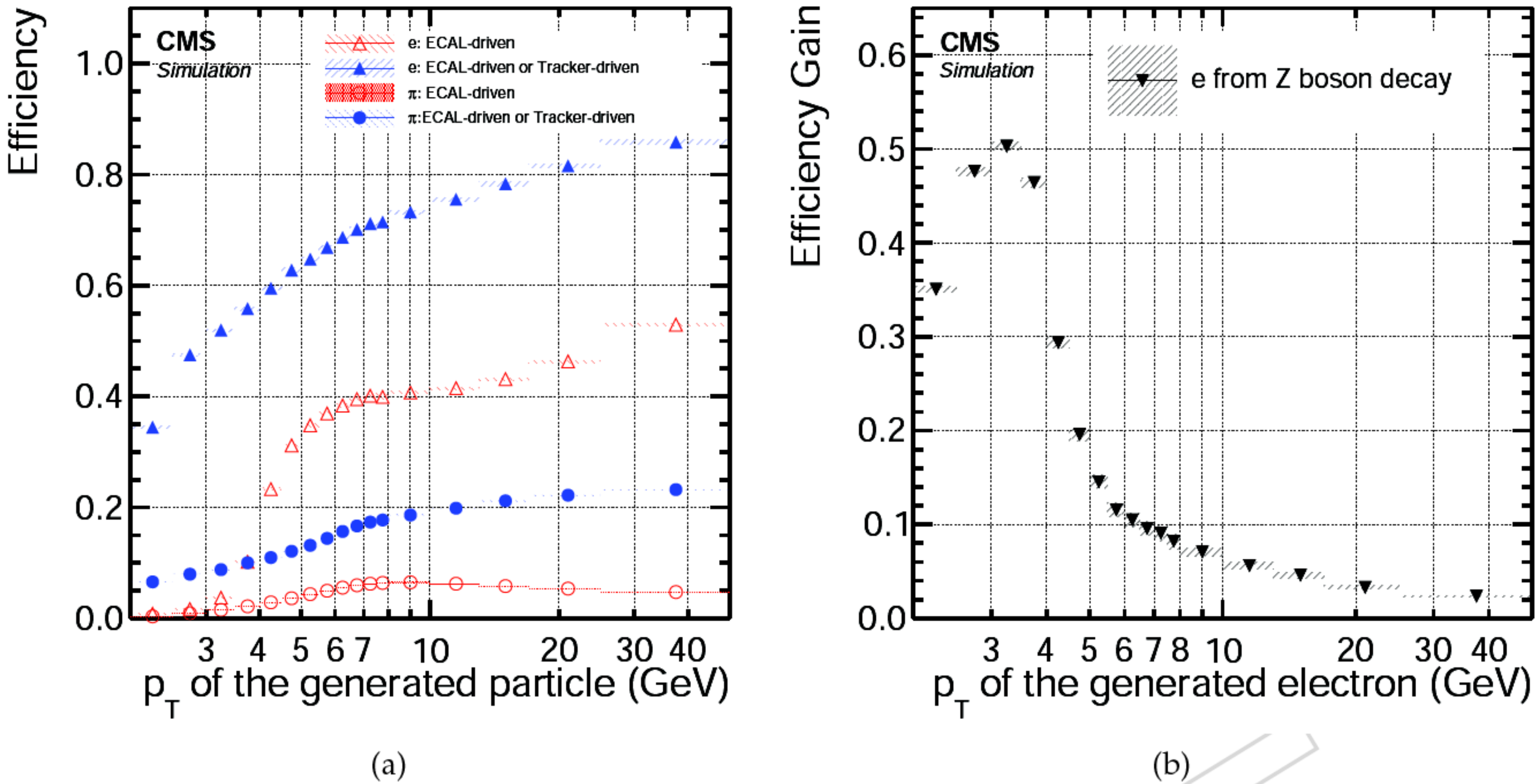
- ECAL seeded
- Simple rule for bremming electrons: baricenter of all the ECAL clusters is on the original helix of the electron
- Success depends on superclustering
- Inefficient for electrons inside jets and many compatible seeds
- Biases baricenter for low pT electrons



Tracker seeded tracking

- Use iterative tracks as potential candidates
- Utilize brem information
- Unified list of tracker and ECAL seeded tracks is made
- Use gaussian sum filter to refit tracks (takes care of non-gaussian energy losses in tracker material)
- More on electron reco later...

Electron seeding efficiency



CMS simulation: a) b jets in 80 to 170 GeV b) efficiency gain in electrons from Z

Muon tracks

- Use muon chamber tracks and tracker tracks. Precise momentum from tracker tracks
- Stand alone Muons: tracks reconstructed from muon chamber hits (and track segments) only
- Global Muons: Done by matching stand alone muon track parameters with tracker track parameters on a surface where both are propagated
- Tracker Muon: consider all tracker tracks with p>2.5 GeV and pT > 0.5 GeV. Propagate to muon system. If one matching segment found it is tracker muon

(pull< 4 or dx < 3 cm in local x co-ordinate)

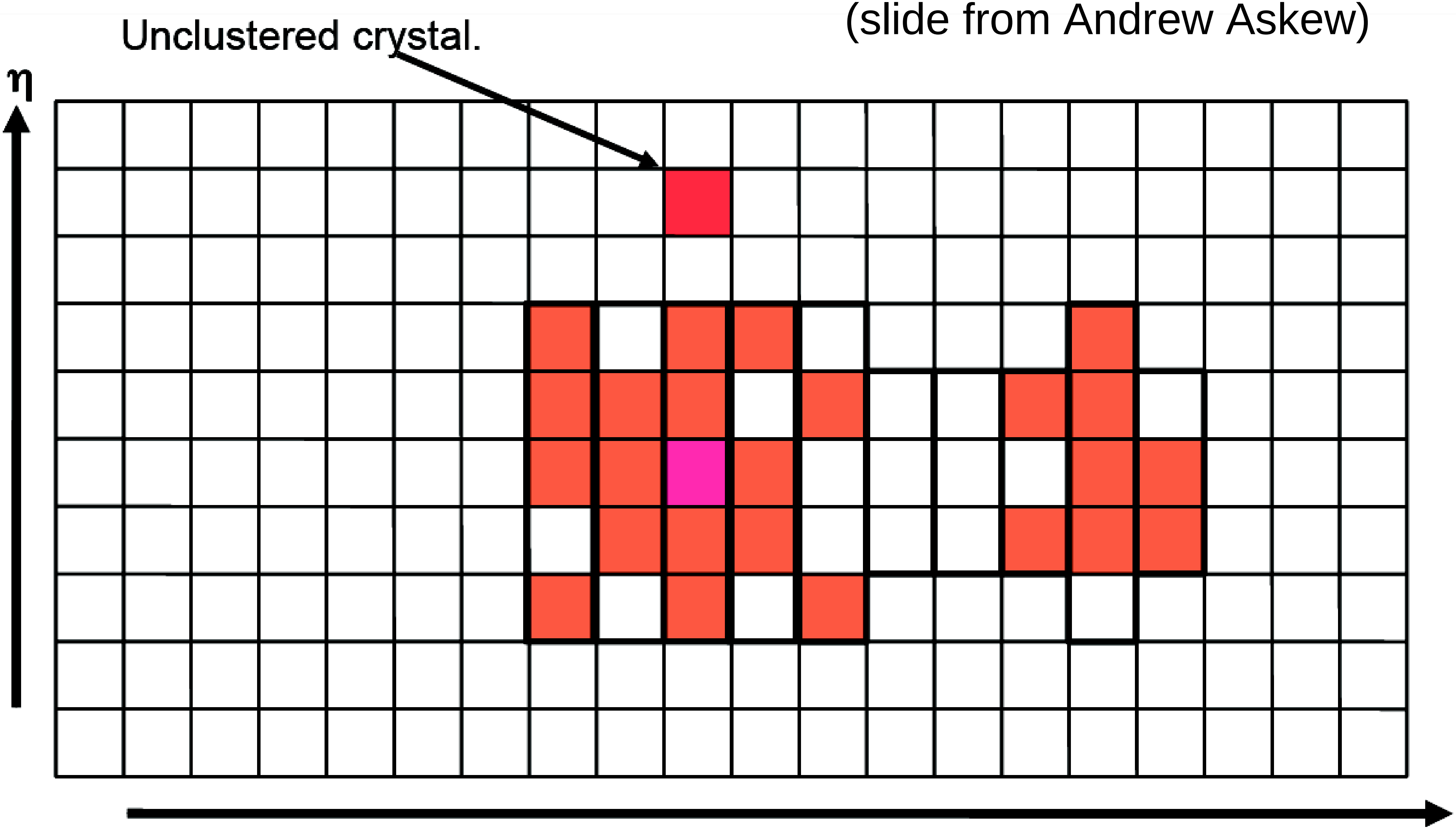
- Global + Tracker muon reconstruction efficiency is ~ 99%

Calo Clusters: Original clustering

- Contiguous set of cells around local maxima is calorimeter cluster.
- Dedicated superclustering in ECAL to recover brem and conversions in tracker
- Hybrid in barrel and multi5x5 in endcap

Hybrid algorithm

(slide from Andrew Askew)



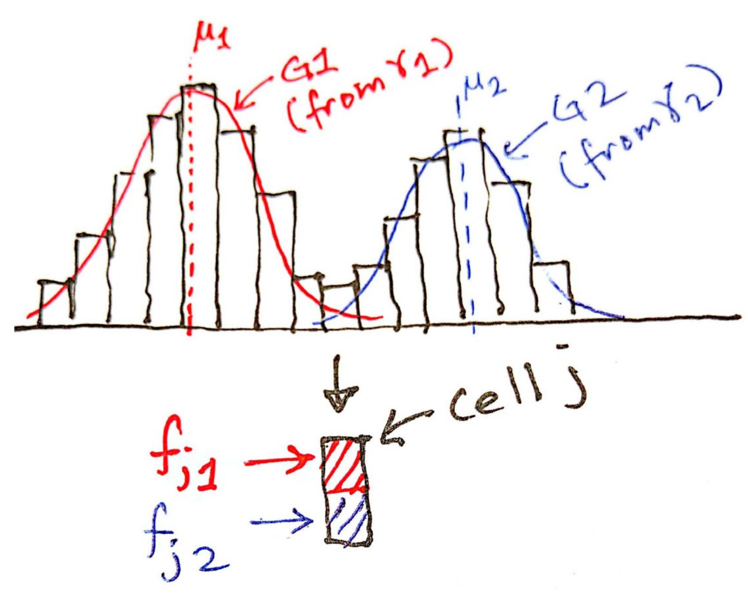
Then take steps in negative ϕ.

PF: topological clustering

- PF uses a different approach:
  - Identify local maxima. Use expectation-maximization algorithm in eta-phi space, gaussian mixture model.
- E-step:
  - missing data is f_{ij}.
  - Update f_{ij} for present value

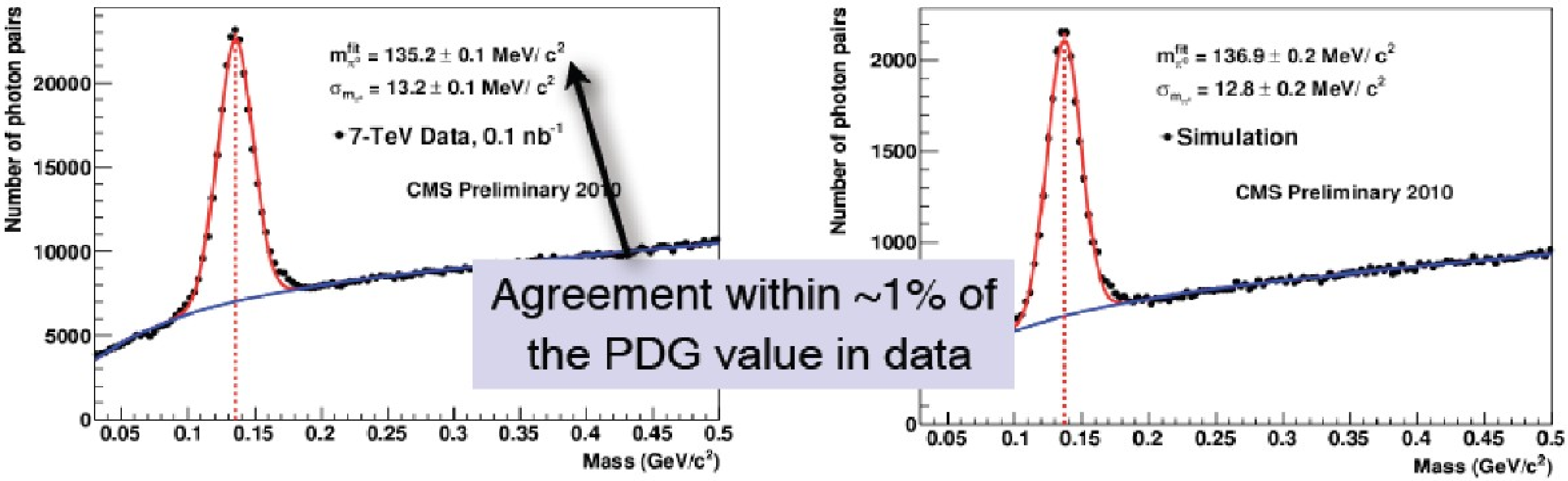
  of parameters $\mathbf{A}_i$ and $\boldsymbol{\mu}_i$
- M-step: analytical likelihood

  maximization of the parameters



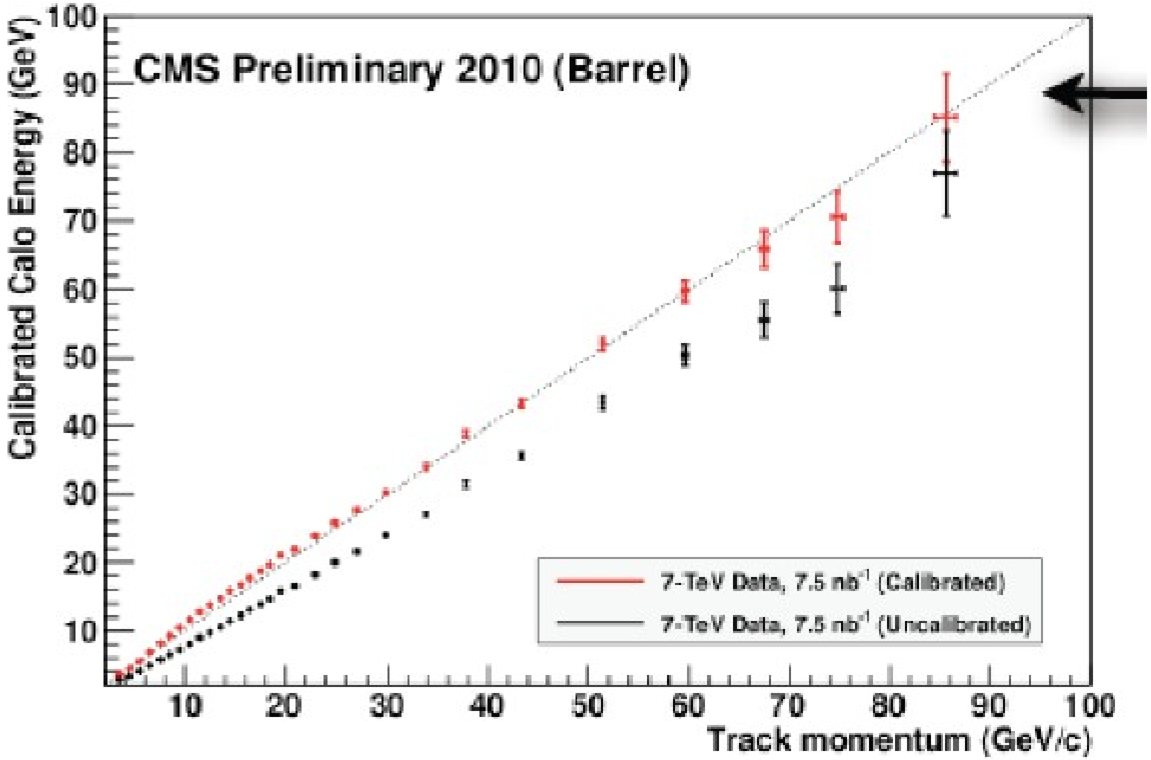
ECAL calibration

Commissioning with $\pi^0 \rightarrow \gamma\gamma$ in 2010



- Stable absolute ECAL calibration

HCAL calibration



- Calorimeter response important for neutral hadrons
- Present calibrated response at the level of 2%

linking and candidate identification

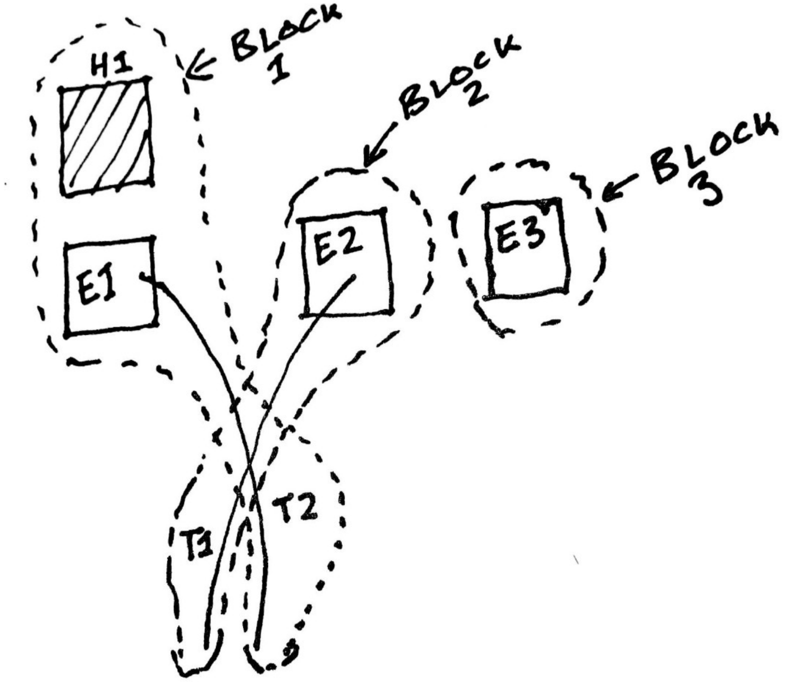
Linking algorithm

- Link elements to blocks
- Blocks = collection of elements linked directly or indirectly
- Purity of linking: all linked elements belongs to one particle
  - limited by granularity
- Efficiency of linking: probability to find all links due to a particle
  - limited by material present in front of a detector element
- Every pair is checked for link.Pair up pf elements based on proximity
- Grows quadratically with n (problem for high pileup, heavy ion)
- Dichotomic sorting with k dimensional tree for linear growth with n

J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching", *Commun. ACM* 18 (September, 1975) 509–517, doi:10.1145/361002.361007.

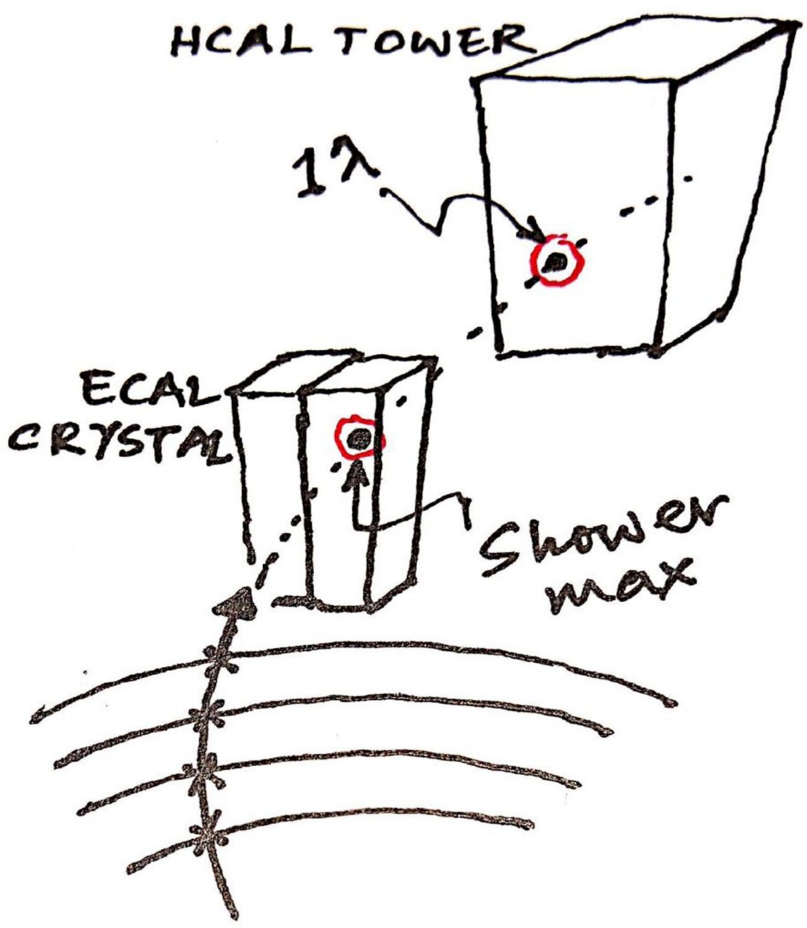
Linking contd...

- After sorting the link algorithm produces blocks
- Blocks are typically upto 3 elements
- Smallness of the blocks ensures performance of algorithm independent of event complexity
- Jets much more complex than the simple example essentially has the same energy response



Linking tracks to clusters

- A track is linked to a calorimeter or preshower as follows:
- From the last hit, extrapolate the track to the expected ECAL shower maximum of an electron
- To 1 interaction length depth in HCAL
- To the two layers of preshower
- Link if the extrapolated point falls within cluster area
- Link distance = distance in eta-phi between extrapolated track and cluster position

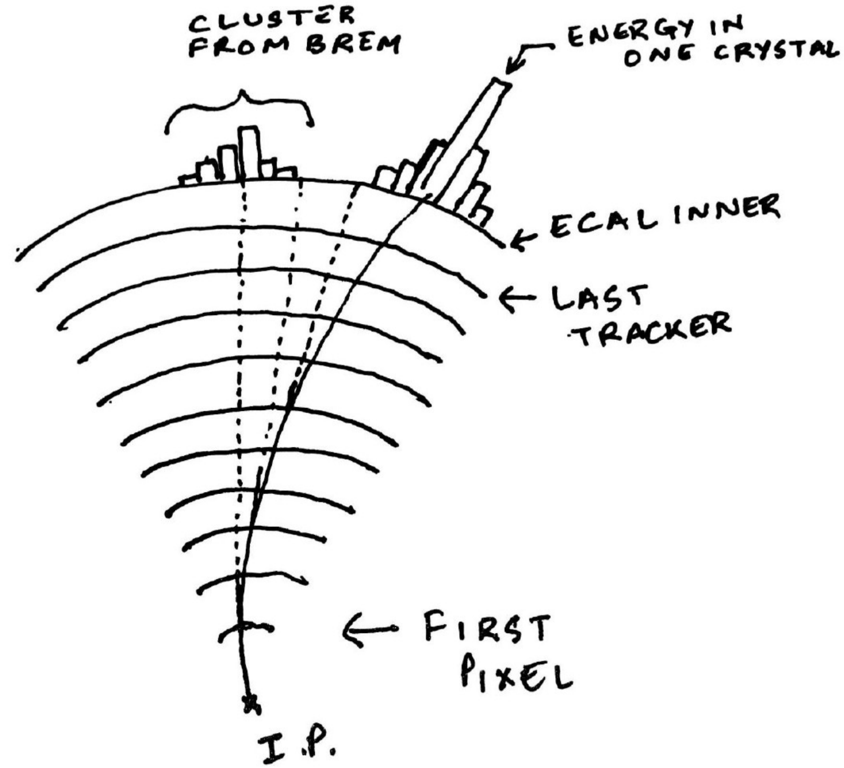


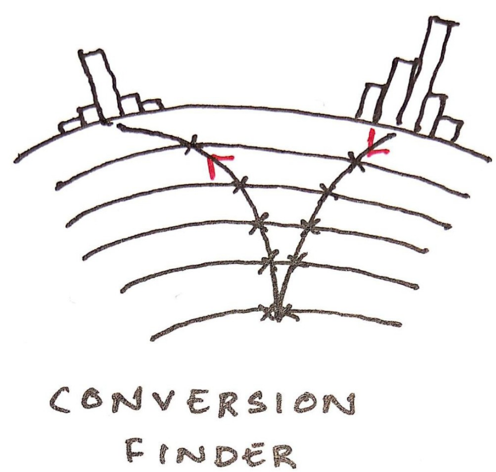
Linking calorimeter clusters

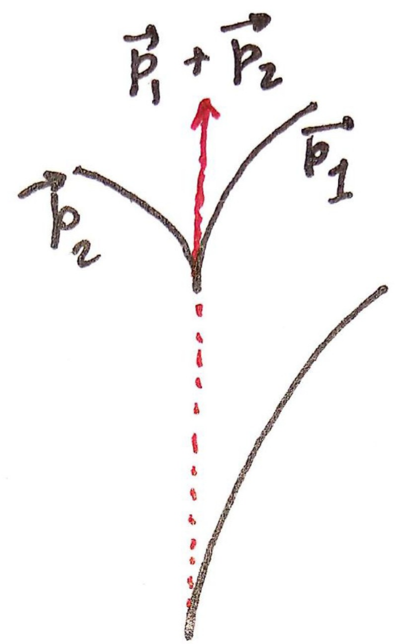
- If the position of the higher granular calorimeter cluster falls inside the cluster area of lower granular calorimeter, link the two
- Multiple links – keep the one with shortest distance

Linking: brem and conversion

- If extrapolated tangent to a track falls inside a ECAL cluster, link it.
- For conversion dedicated conversion track finder links track pairs compatible with a conversion
- If the resultant of two conversion tracks is tangent to another track then link







Linking muon tracks and tracker tracks

- Nothing special is done in particle flow

From blocks to candidates

- Linking produces blocks
- In each blocks candidates are searched in a sequence
- Identify muons and remove corresponding tracker tracks and HCAL,ECAL deposits
- Then electron and photons are identified and corresponding elements are removed
- Remaining elements are examined for charged hadrons, neutral hadrons and non-prompt photons from fragmentation and decays in jets

Muon identification

- Sum of ET of calo deposits and pT of tracks in tracker in 0.3 isolation cone < 10% pT of muon
  - Gives extremely high purity
- Care is needed for muons inside jets (heavy flavour decay)
  - Charged hadron track identified as muon will give spurious neutral hadron
  - Failing to remove muon track will tend to give spurious charged hadrons
  - tighter muon selection is applied (atleast three track segments, calo deposits compatible with muons)

Electrons and photons

- Candidates are
  - GSF track with a cluster
  - Topological ECAL cluster
- All other clusters linked to the candidate by track tangent are added to the candidate
- All tracks linked to the candidate are added to the candidate
- Pions removed by (ECAL energy/track pT) and ECAL energy/HCAL energy)
- GSF electron candidates are then passed through a BDT with 14 inputs (track quality, radiated energy, energy to momentum ratio, HCAL deposit...)
- Clusters without GSF track are prompt photons if they have desired shower shape and are loosely isolated

Hadrons and photons

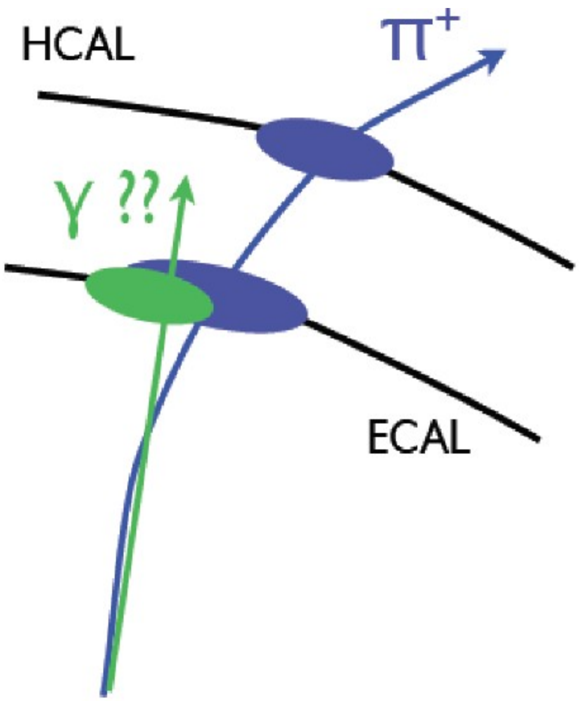
- Stable charged: K+-, pi+-, protons
- Stable neutrals: neutron, K^0_L
- Photons from pi^0's, eta, fragmentation
- charged hadrons: require tracks to have pT uncertainty smaller than the linked cluster energy uncertainty to control fakes at high pT
- 0.2% tracks are rejected of which 90% are indeed fakes, remaining 0.02% get reconstructed as photons or neutrals
- Hadron id starts from HCAL clusters linked to a track.
- In case of many clusters to one track, take the nearest.

Charged Hadrons first

- Sum of track momenta, sum(p) is compared to sum of calibrated ECAL + HCAL (E+H) energies.
- True energy for calibration: E+H> sum(p)? (E+H):sum(p)
- if E+H<<sum(p) use a fake track removal procedure
  - 0.03% of tracks affected
- Remaining tracks are charged hadrons. p is taken from track and then recomputed as a weighted sum of calo energy and track p

Then photons and neutrals

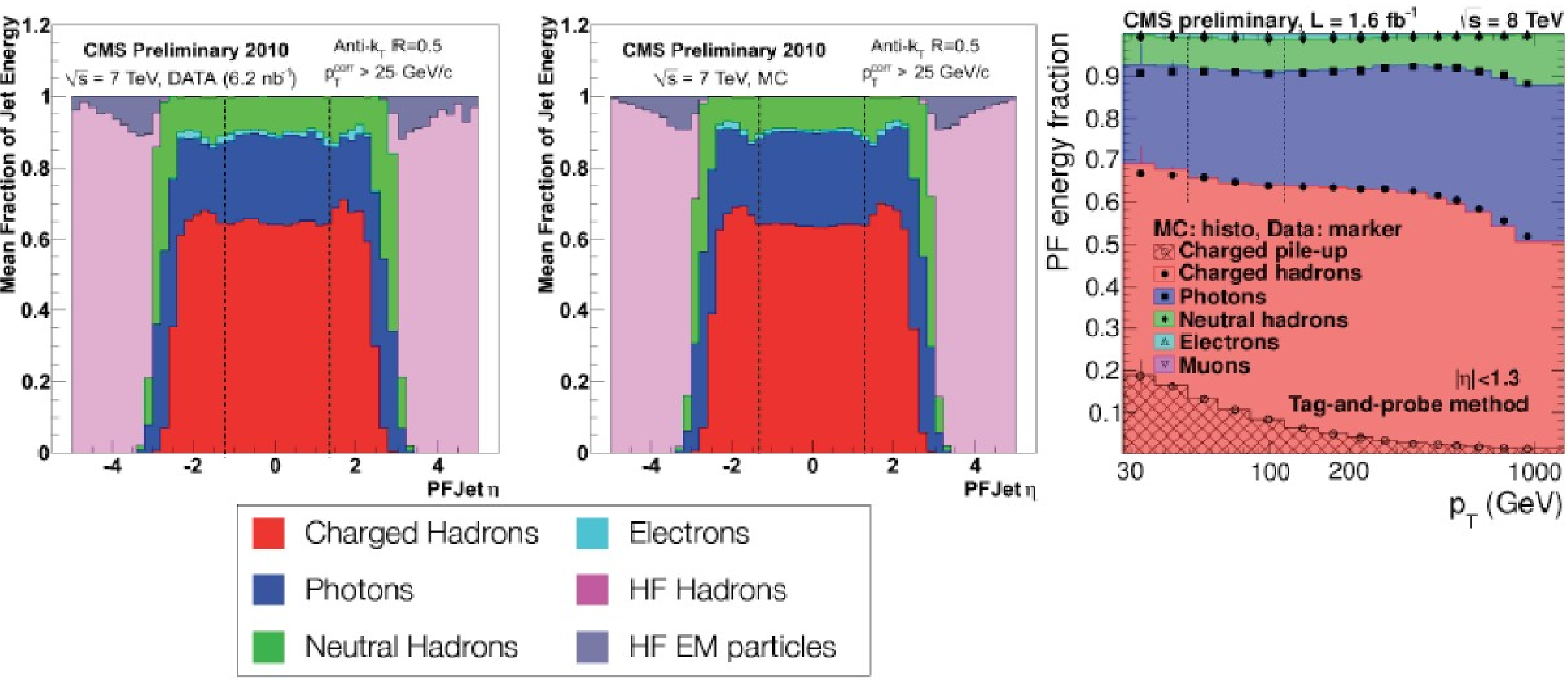
- If E+H >> sum(p) assume due to photons and neutrals
- Start with photons (25% of jet energy in ECAL vs 3% by neutral hadron)
- If ((E+H) -sum(p) <= E) create photon
- Else create a photon from E and a neutral from remaining excess.
- Remaining ECAL elements are photons and HCAL elements neutral hads. (within tracker coverage)
- Beyond tracker coverage ECAL cluster linked to a HCAL cluster treated as hadron and only ECAL cluster as photon



Performance of particle flow

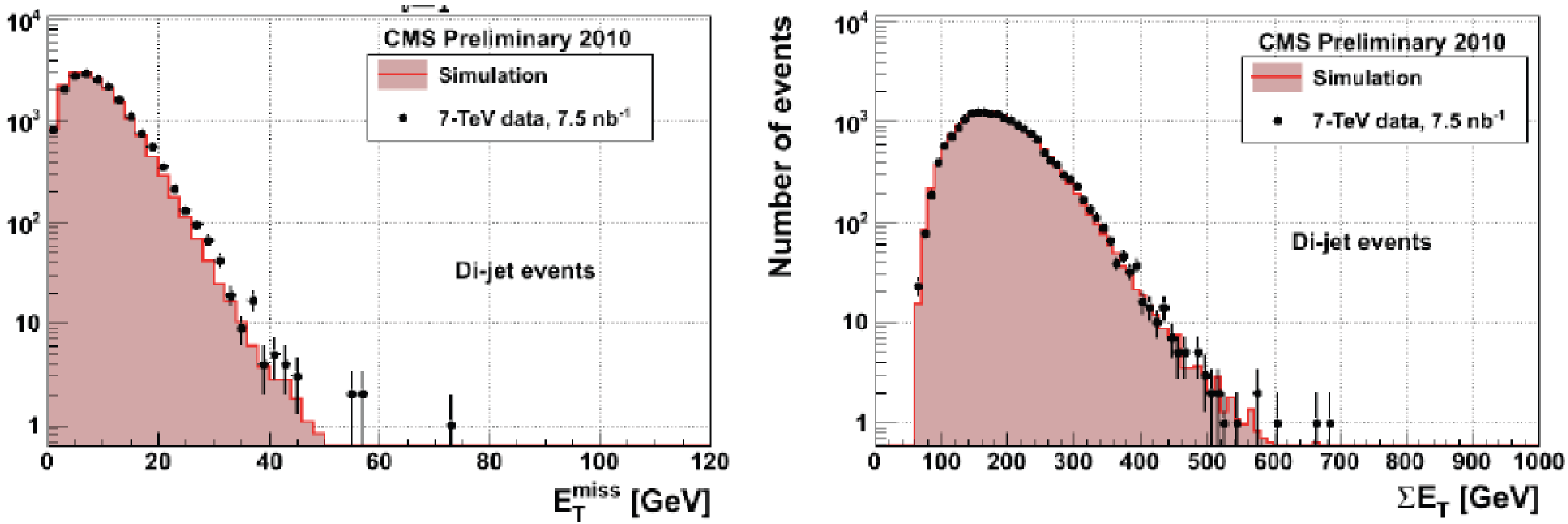
taken from 2016 slides of Albert de Roeck

Jet composition



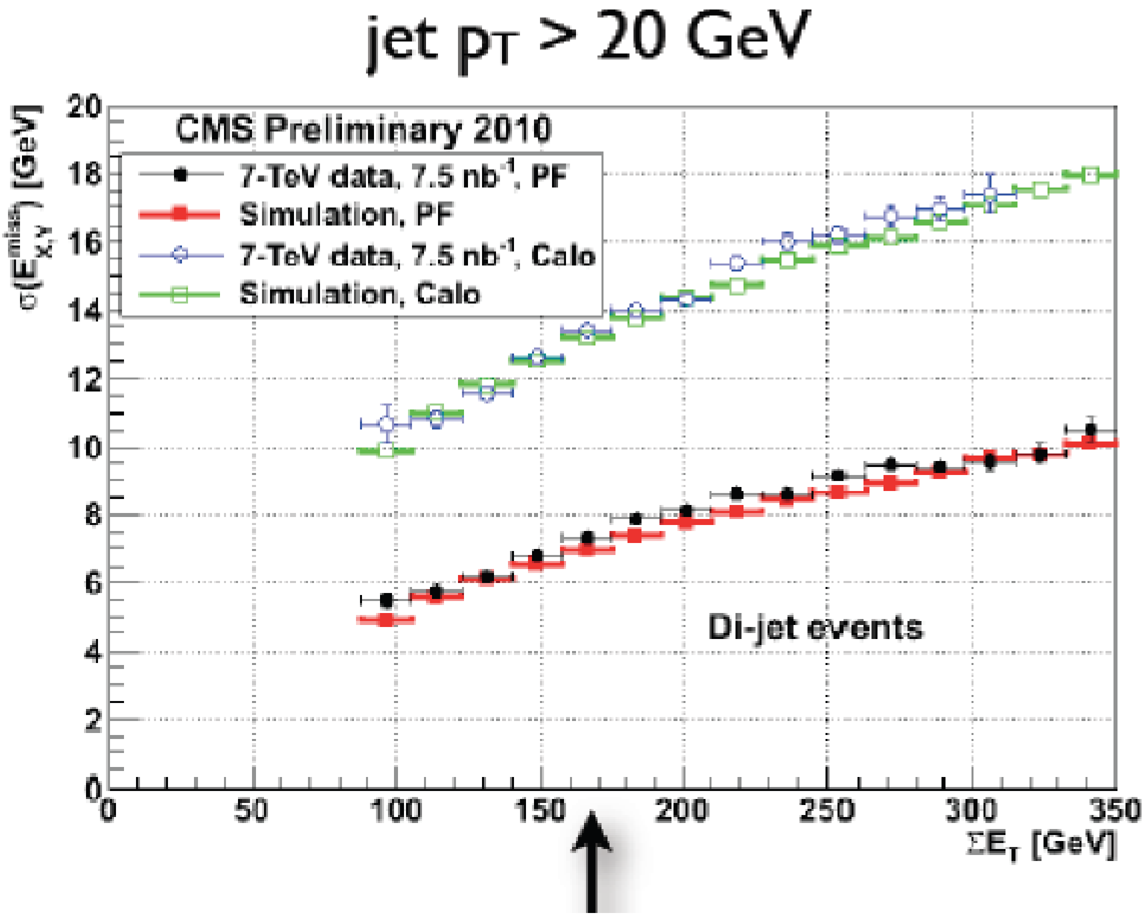
The agreement of the first days is confirmed with high statistics, even in presence of pile-up

MET



- Agreement over 3 orders of magnitude in scalar and vector sums of PF candidate momenta

MET improvement

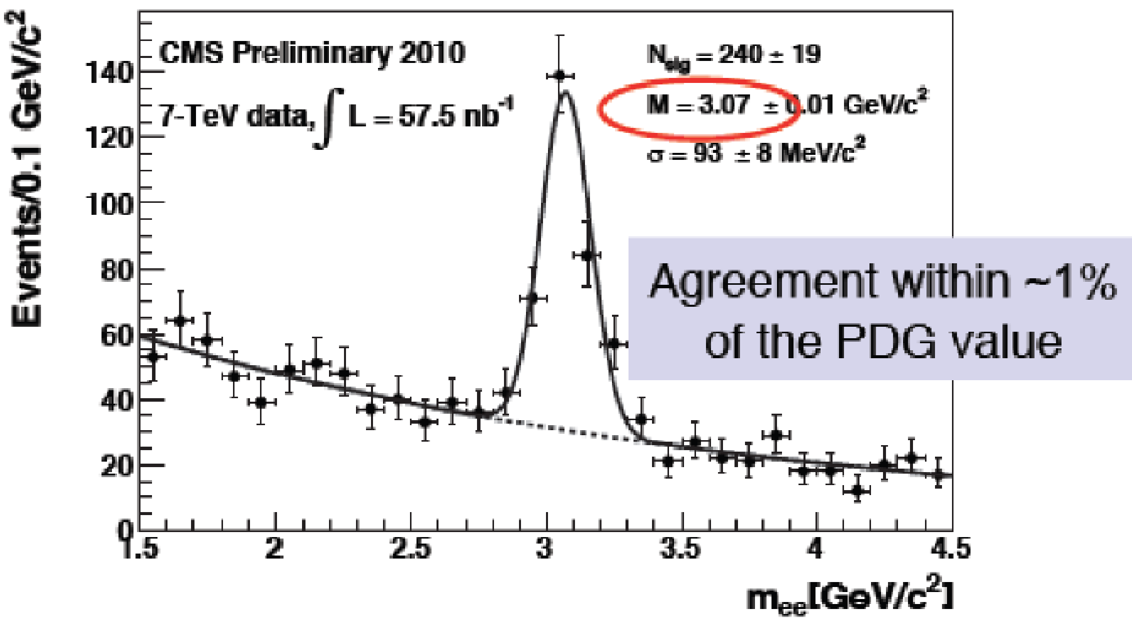


- Improvement by a factor 2!

Photons and electrons

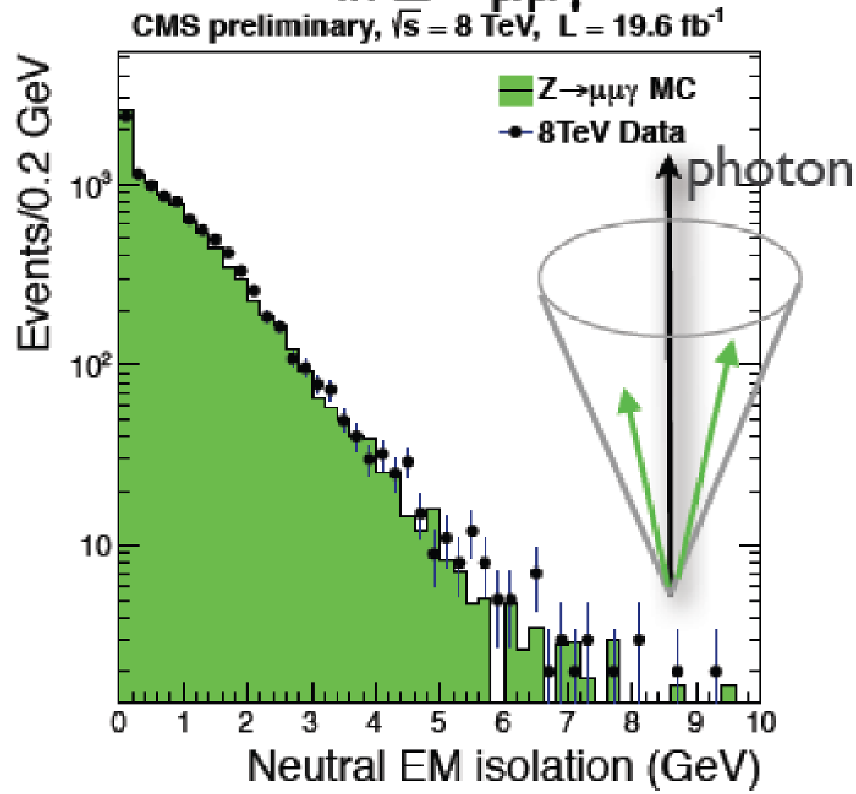
From 2010 .... to ...2013

$J/\Psi \rightarrow ee$



With a simple selection a nice $J/\Psi \rightarrow ee$ peak is obtained

Photon isolation around photons in $Z \rightarrow \mu\mu\gamma$



Effect on physics

Jets

- energy resolution / 2
- angular resolution / 3
- Flavour dependence of response / 3
- Systematic error on JES / 2
- « electron in jet » b tagging
- quark-gluon jet tagging

MET

- resolution / 2
- less tails

τ

- jet fake rate / 3 @ same eff.
- energy resolution / 4
- decay mode

Electrons

- down to pT = 3 GeV
- in jets

μ

- 4% more efficient ID @ same bkg rate
- better momentum assignment at high p_T

e, μ, τ, γ isolation

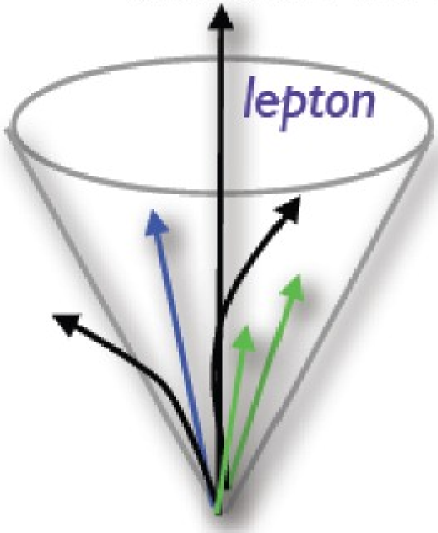
- improved performance, pile-up control

Physics analyses

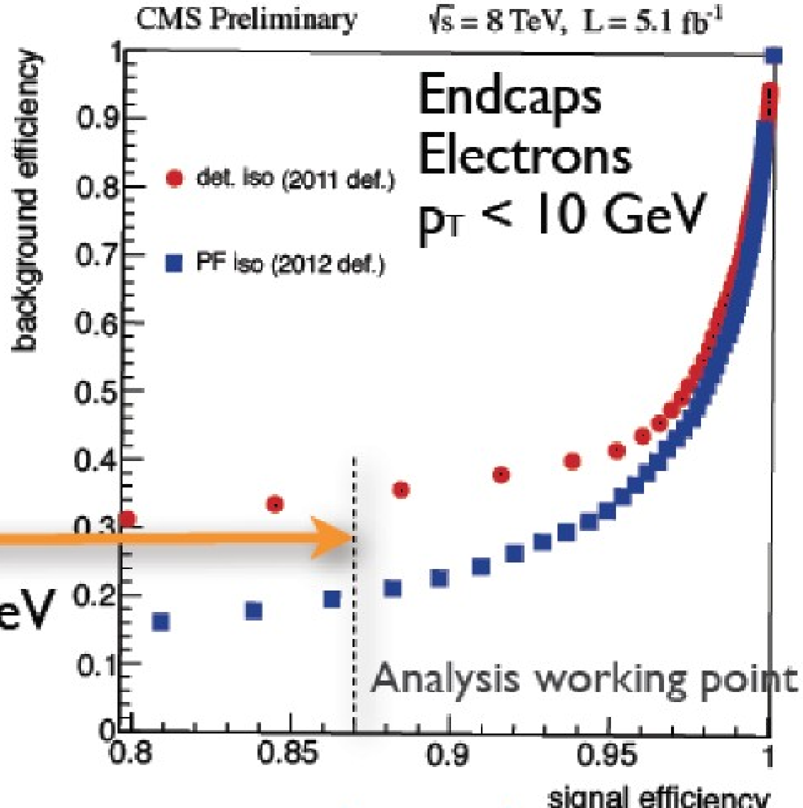
- Better trigger for jets, MET, taus (PF@HLT)
- FSR photon recovery in HZZ
- embedding in $H\rightarrow\tau\tau$
- jet substructure

PF isolation

The "classic" method to compute the lepton/photon isolation was to sum the energy deposits in the tracker, the ECAL and the HCAL



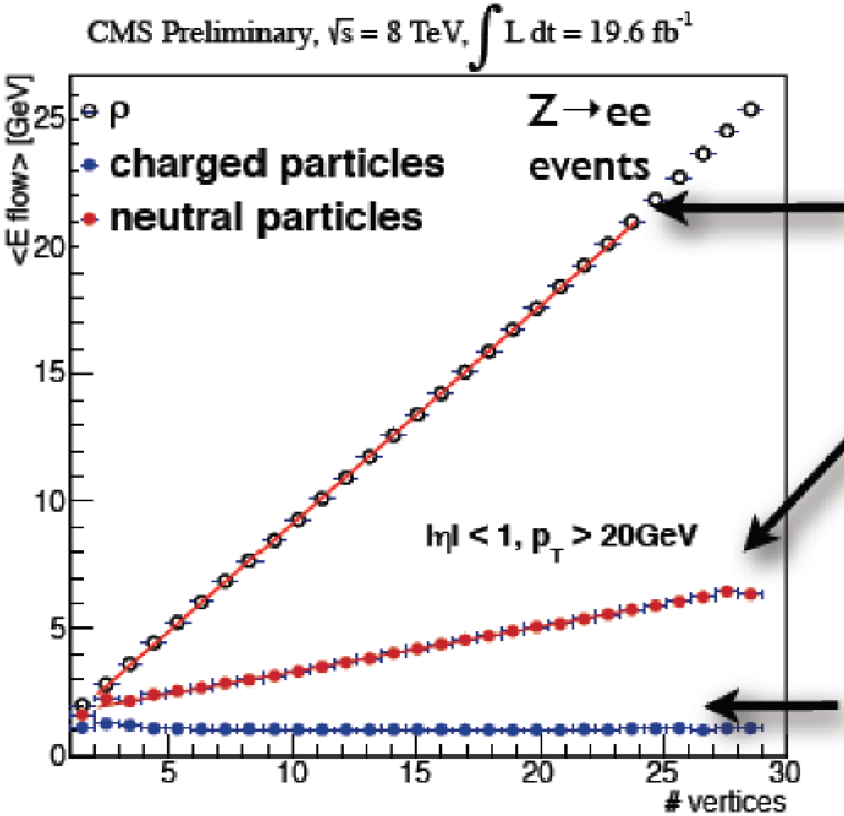
Background efficiency divided by a factor of 2
Similar gain for p_T < 20 GeV



With the Particle Flow it is natural to use the reconstructed particles, to compute the momentum carried by charged hadrons/photons/neutral hadrons in a cone centered on the lepton/photon

- The object footprint is automatically removed by the PF
- No double counting of track and calorimeter energy deposits for charged particles

Isolation and pileup mitigation

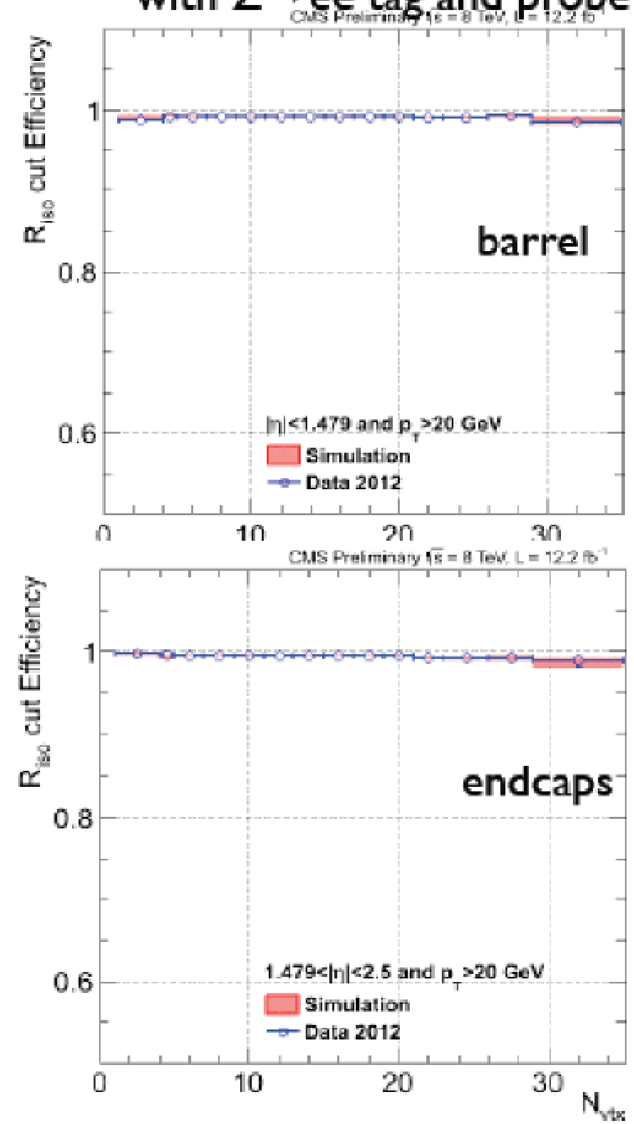


average event energy density (fastjet)

Neutral particles around leptons

Charged hadrons (required to come from the primary vertex) around leptons

Isolation efficiency measureme with Z→ee tag and probe



- No correction needed for the charged hadrons (vertex constraint)
- For the neutrals: the PU contribution in the cone is estimated (proportional to the energy density) and subtracted

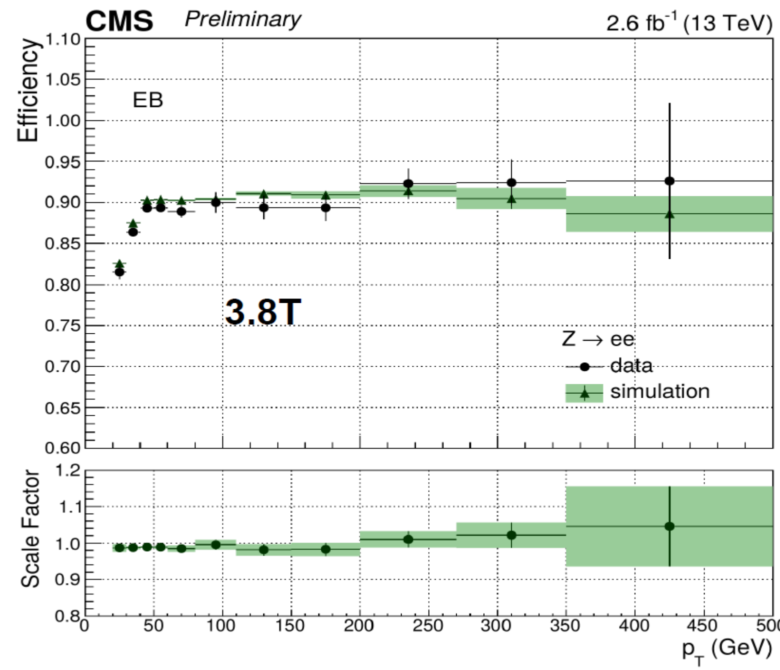
Summary

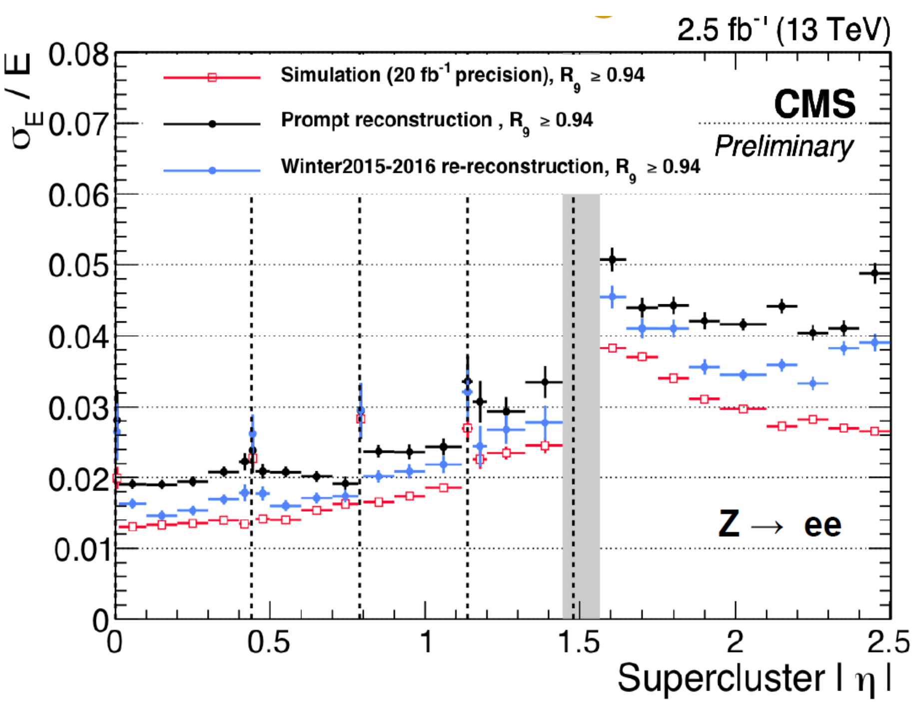
- Algorithm established to be working in CMS
- Exploits power of granularity of the CMS detector
- Implemented and commissioned in run 1
- Significant improvements in jets, MET, tau, lepton isolation
- Most analyses use PF objects
- Some triggers also use PF
- Not only withstands pile-up, it is the way forward to maintain same performance in coming runs and for HL-LHC

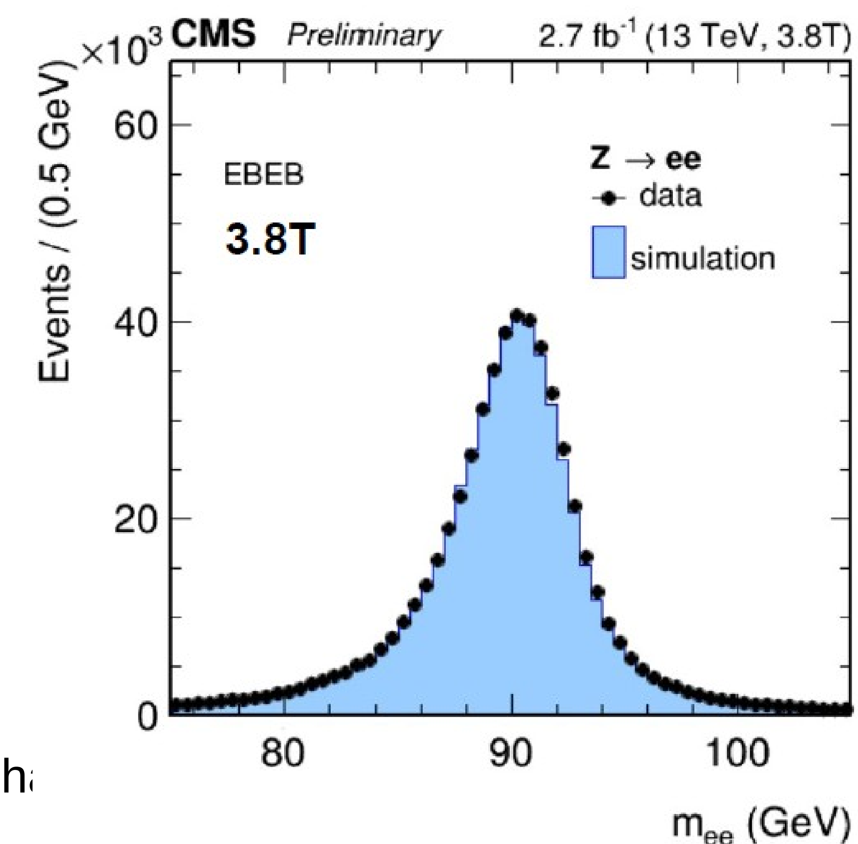
backup

Electrons and photons

- Photon identification efficiency ~ 90%
- Photon energy resolution ~ 1% from Z to ee data

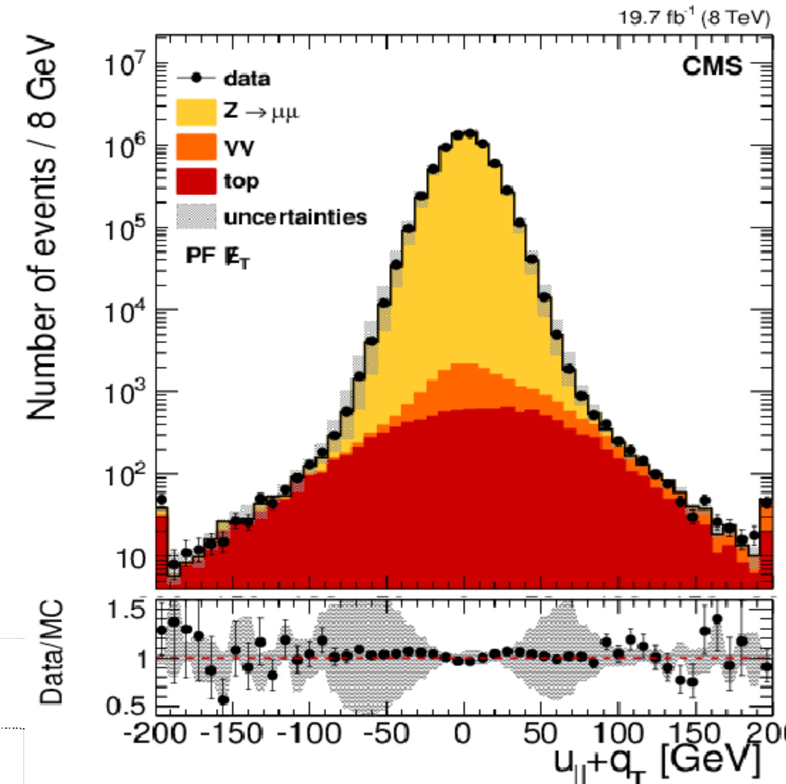


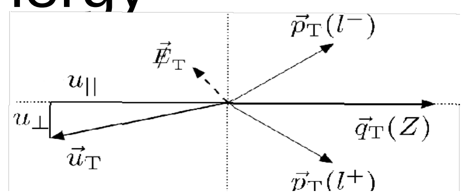


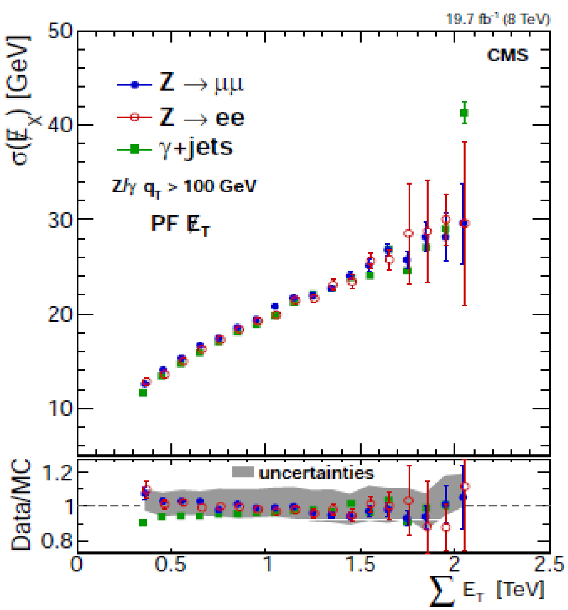


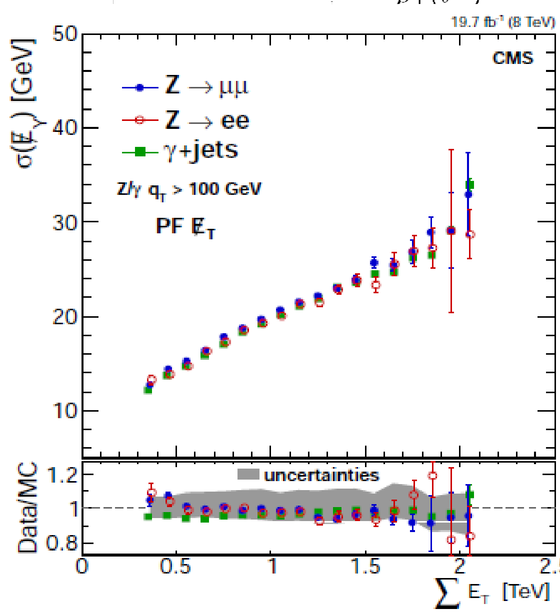
MET ($\not{E}_T$)

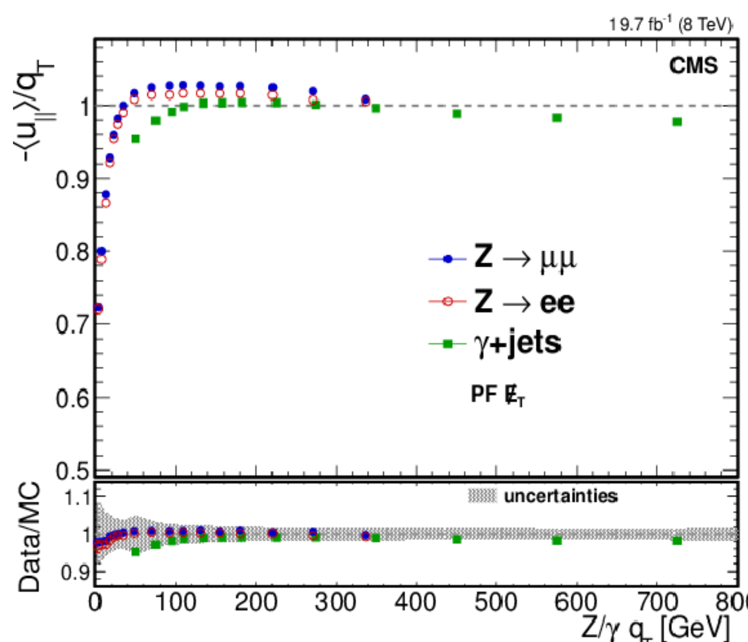
- Constructed from PF candidates
- Correted for various detector effects
- Dominated by jet energy resolution





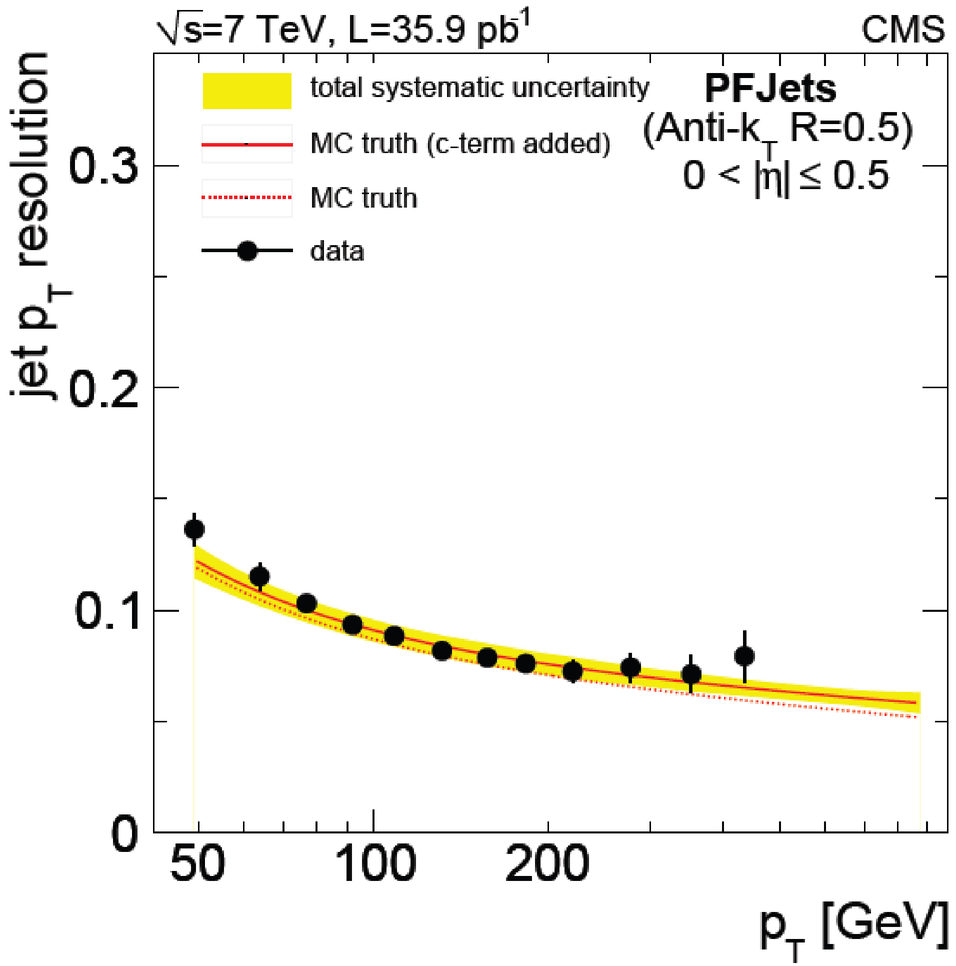






Jets

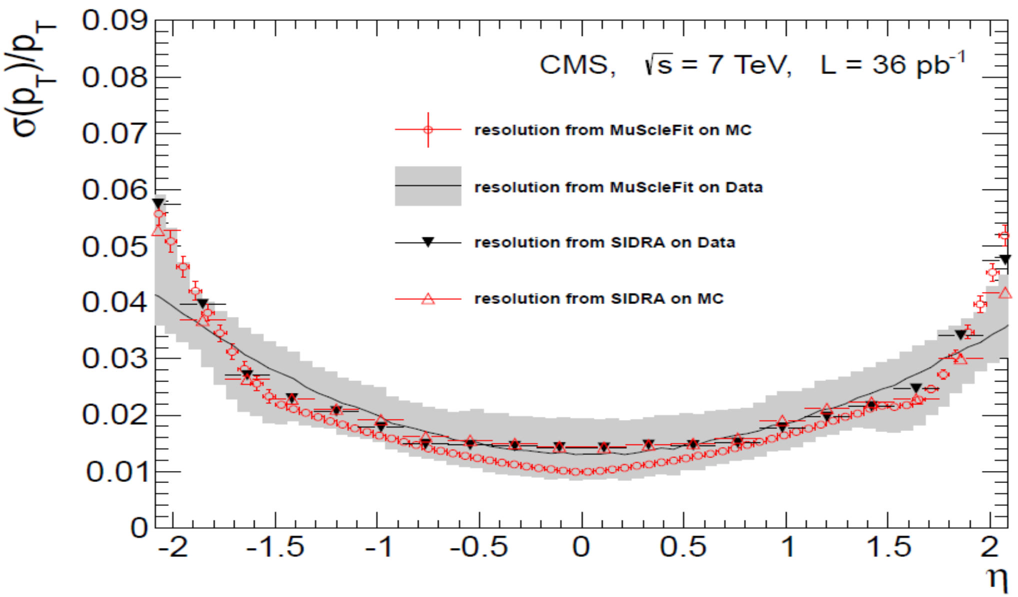
- Anti-KT with distance parameter 0.5
- CALO, JPT, PF
- PF jets clustered from PF candidate particles
- Resolution measured from MC and various energy balancing methods



2011 JINST 6 P11002

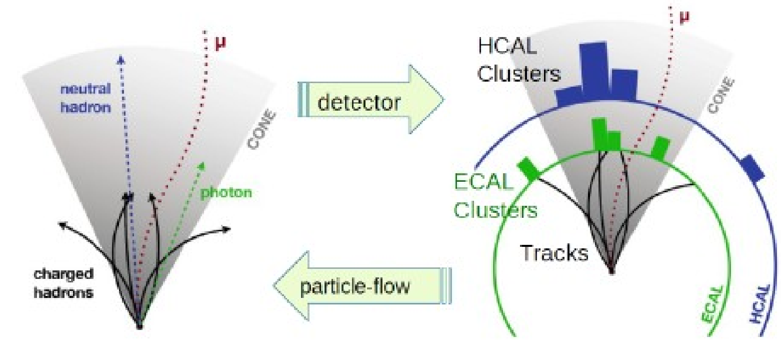
Muons

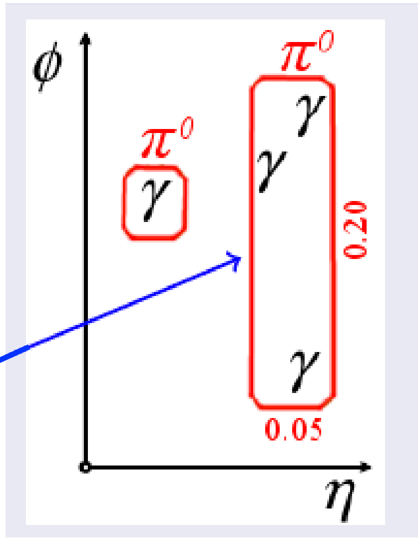
- 1-6% relative momentum resolution for pT<100GeV
- > 10% at a TeV
- > 1% hadron to muon fake probability
- Single muon trigger rates (much) better than 90% above a few GeV



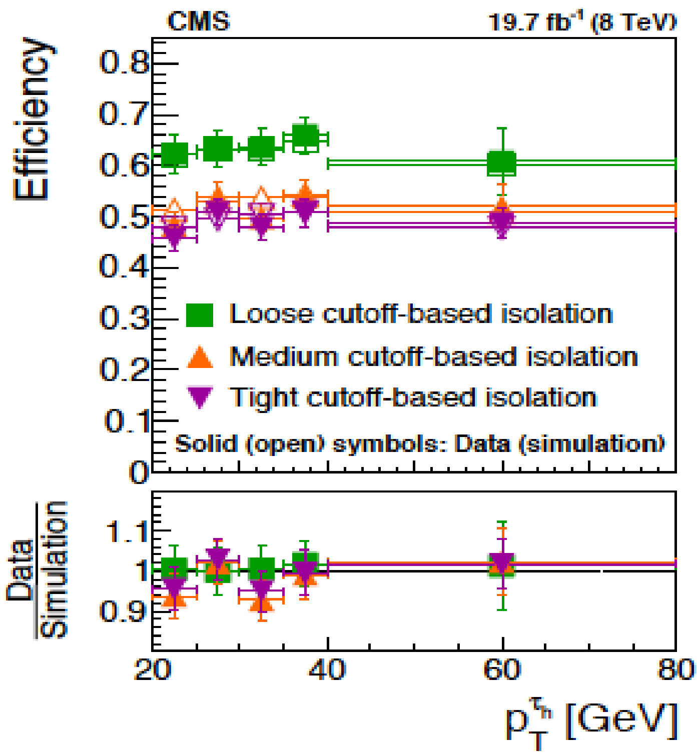
Taus: the HPS algorithm

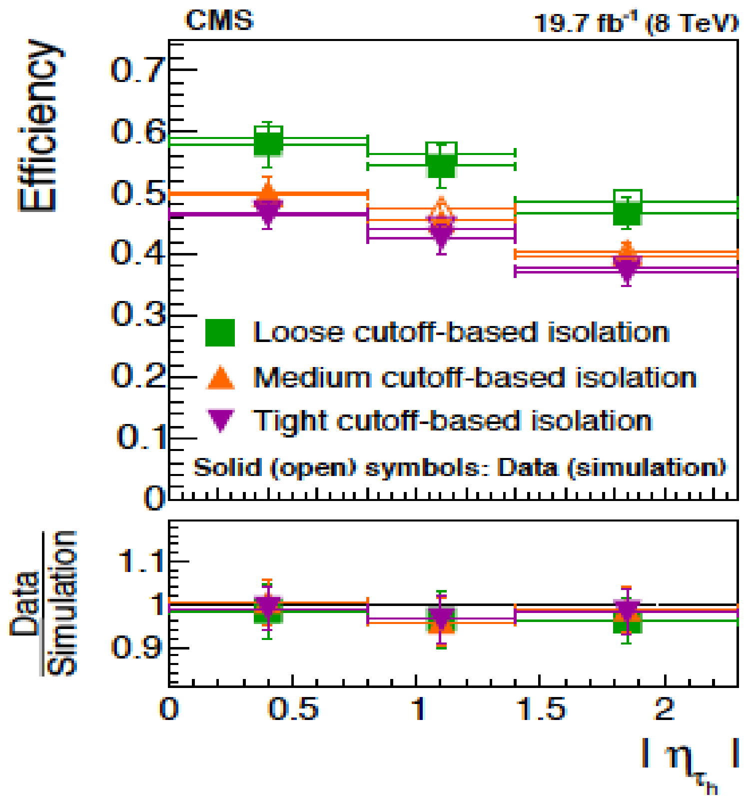
- charged hadrons reconstructed using PF algorithm
  - π0's are reconstructed in ECAL as strips
- Strips:
  - π0 -> γγ
  - Photon conversion in the tracker material
  - electron tracks bending in the magnetic field: broadening of the signal in the azimuthal direction
  - A strip of 0.05 in η and 0.2 in φ is built
  - Mass is required to be consistent with π0





Tau efficiency





b-tagging efficiency

- The impact parameter (IP) of the track wrt the primary vertex is used to distinguish the decay product of the b hadron from the prompt tracks
- Algorithms:
  - Track counting: sorts tracks in a jet by decreasing value of IP significance
  - Jet probability (JP): uses estimate of the likelihood that all the tracks associated to the jet come from primary vertex
  - Jet B probability (JBP): same as JP, in addition, it gives more weight to the tracks with high IP significance