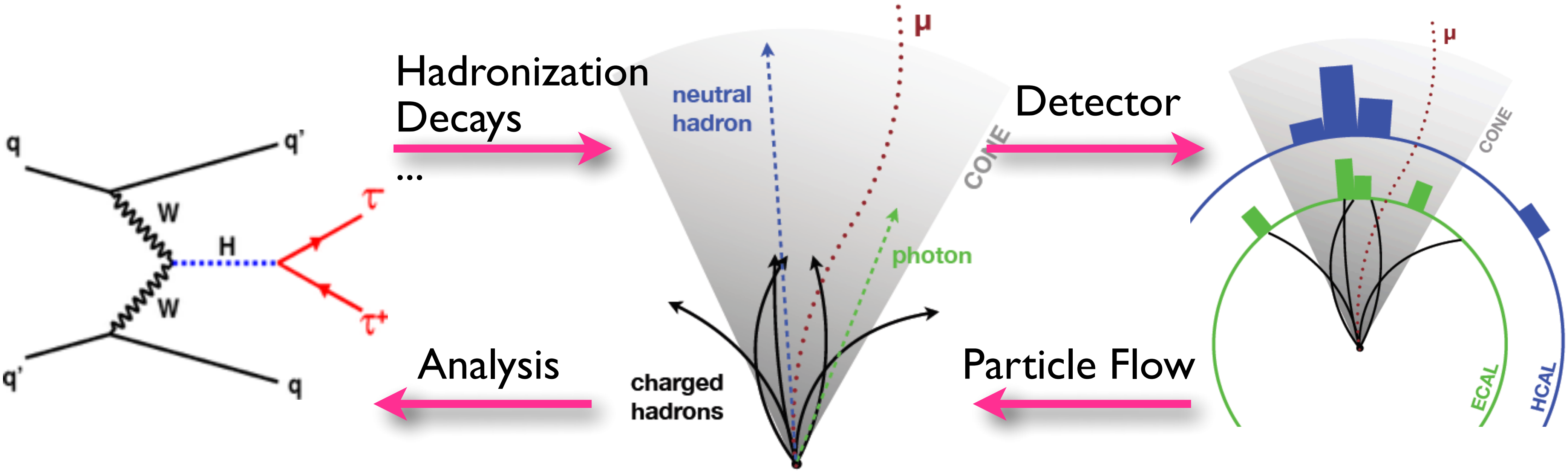
The CMS Particle Flow Algorithm

Florian Beaudette

Including material from:
CMS-PAS-PFT-10-002
CMS-PAS-PFT-10-003
CMS-PAS-PFT-09-001
CMS-DP-2012/012
CMS-PAS-JME-12-002
CMS-PAS-JME-13-003

PFA Mini-Workshop 20/03/2017

Overview : the Particle Flow algorithm



Hadronization
Decays
...

Detector

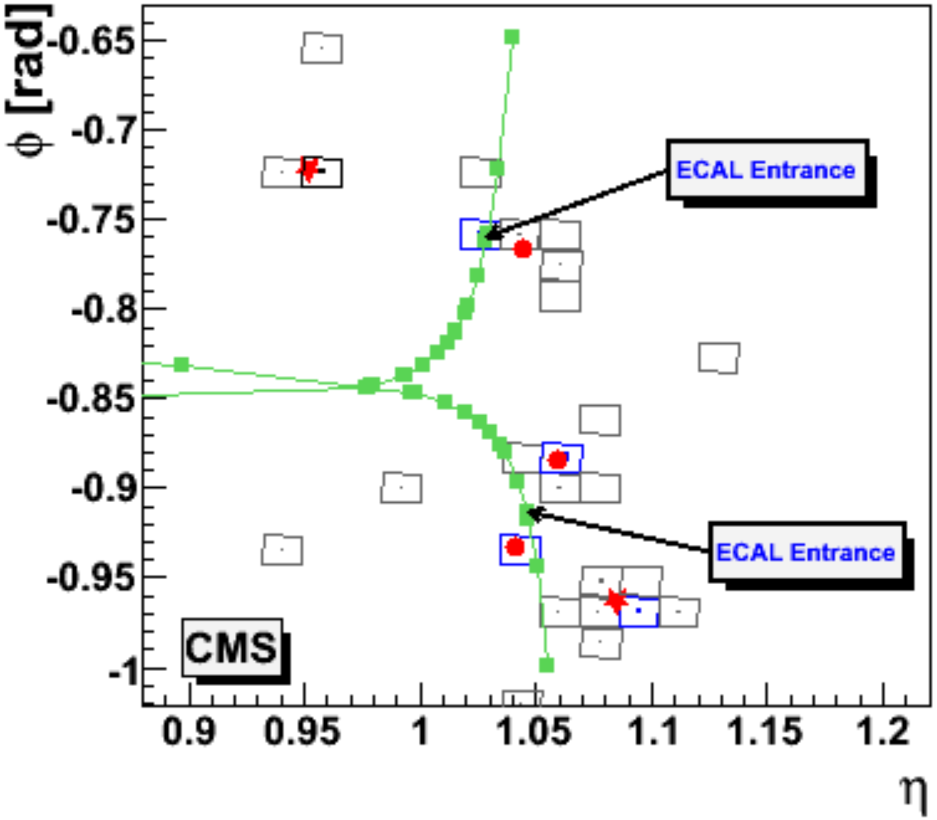
Particle Flow

Analysis

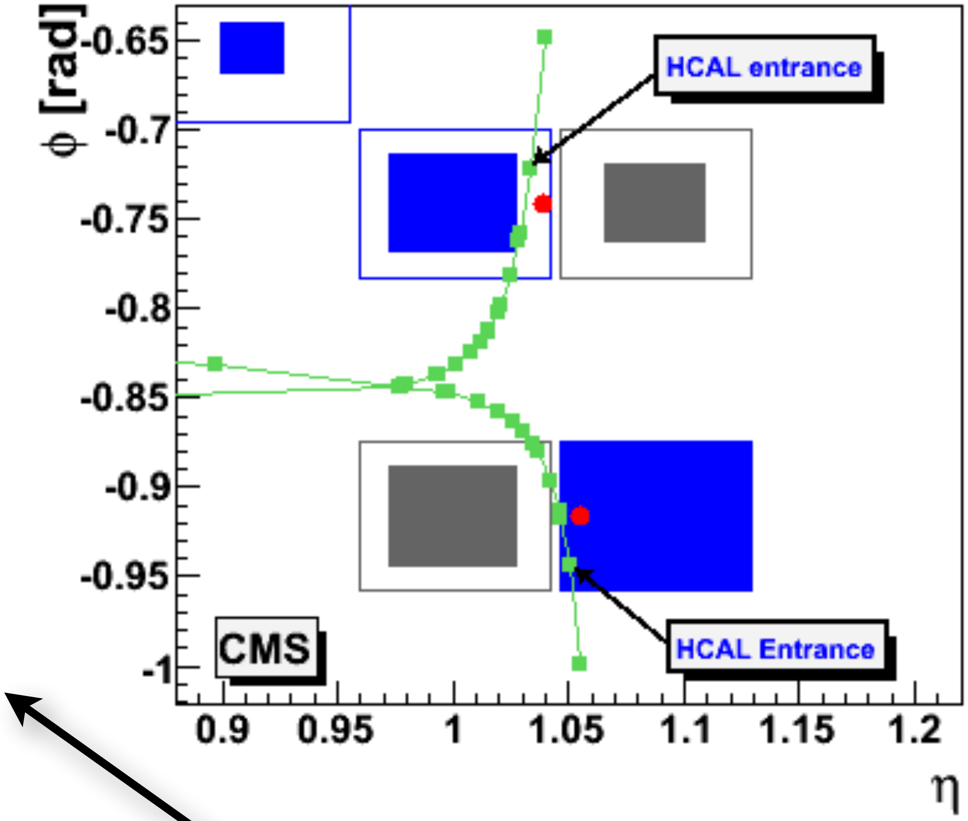
The list of individual particles is then used to build jets, to determine the missing transverse energy, to reconstruct and identify taus from their decay products, to tag b jets ...

Track-cluster link

ECAL surface



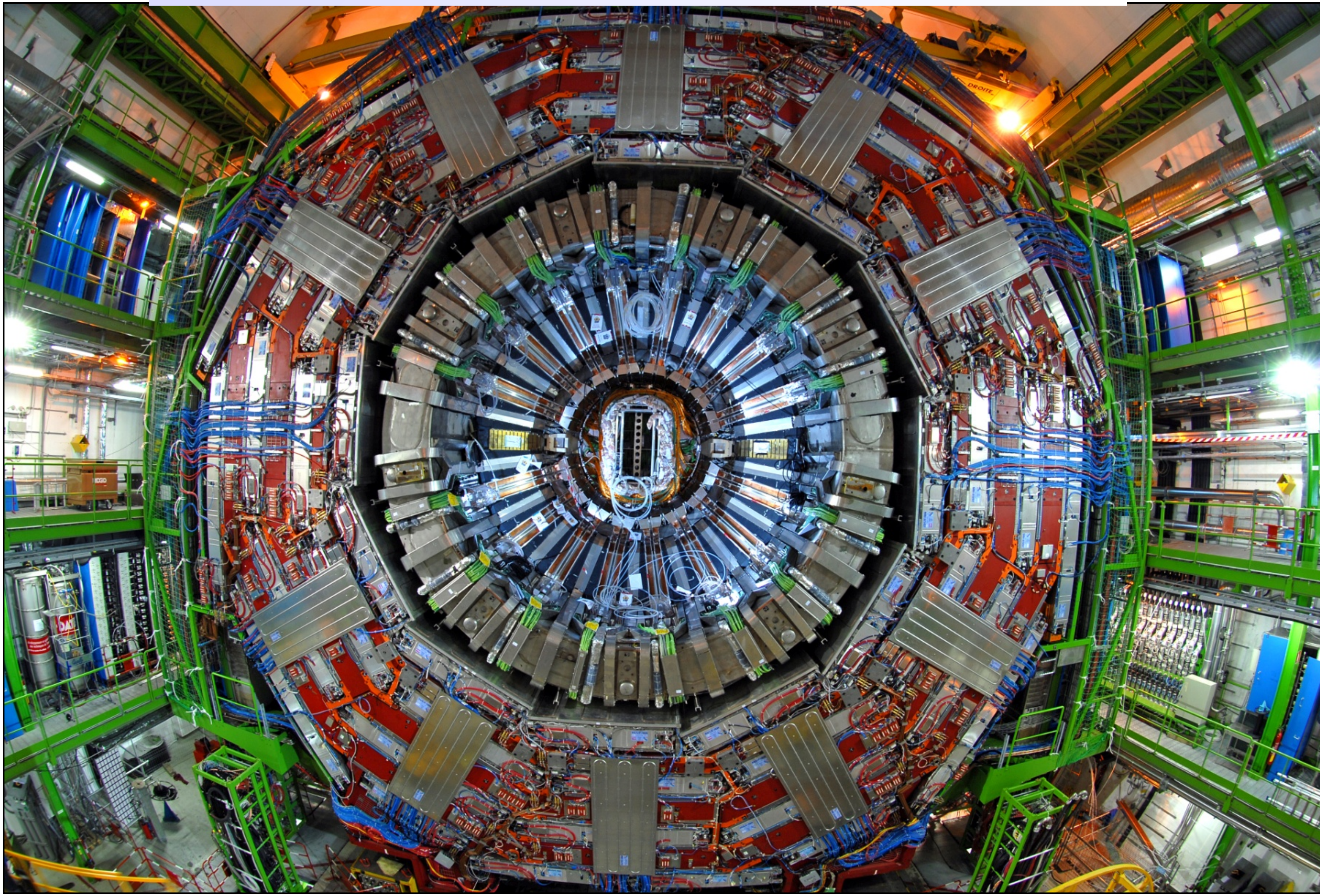
HCAL surface



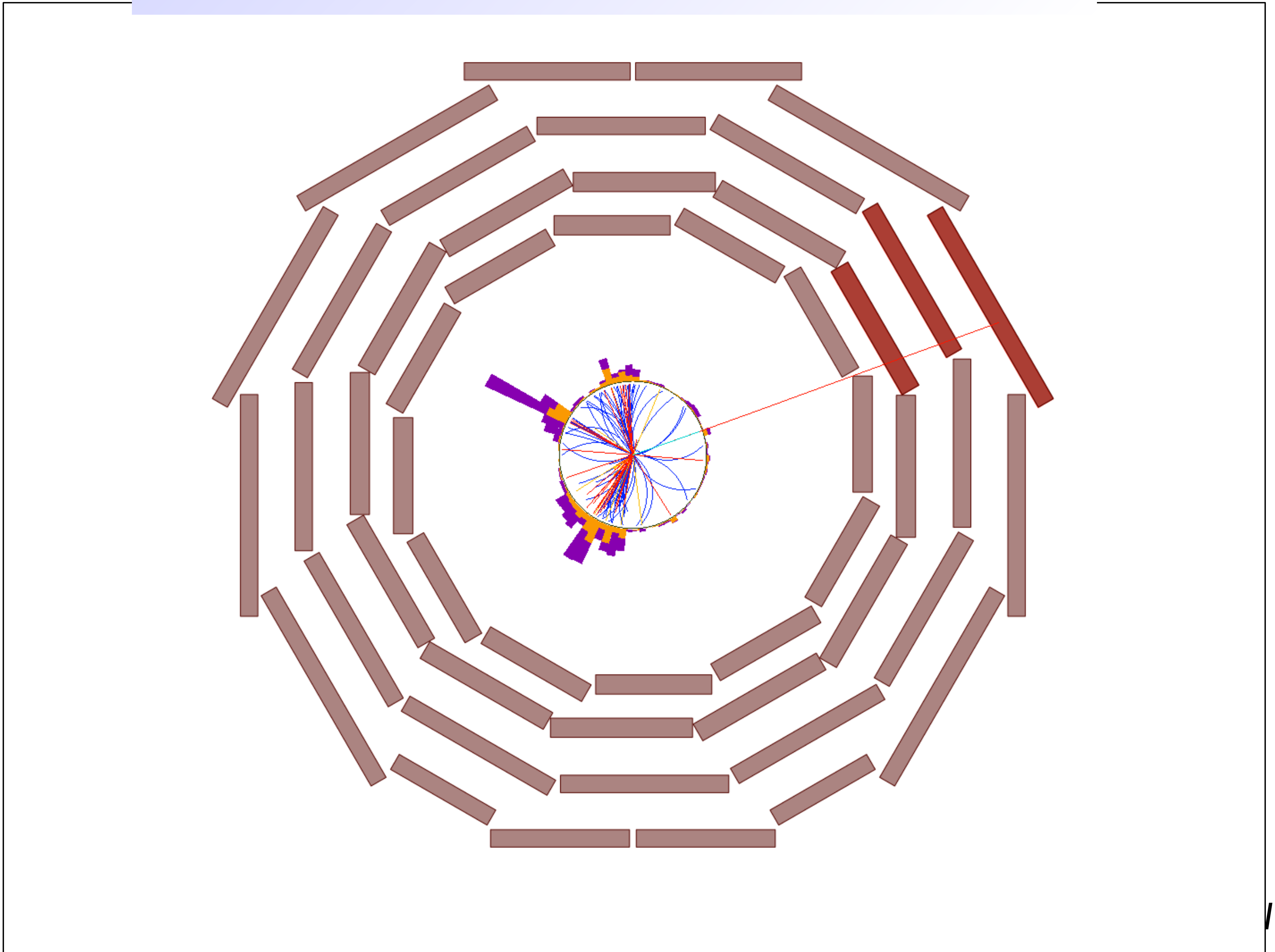
★ Two photons (ECAL clusters not linked to a track) plus a π^+ and π^-

Major role of the clustering algorithm !

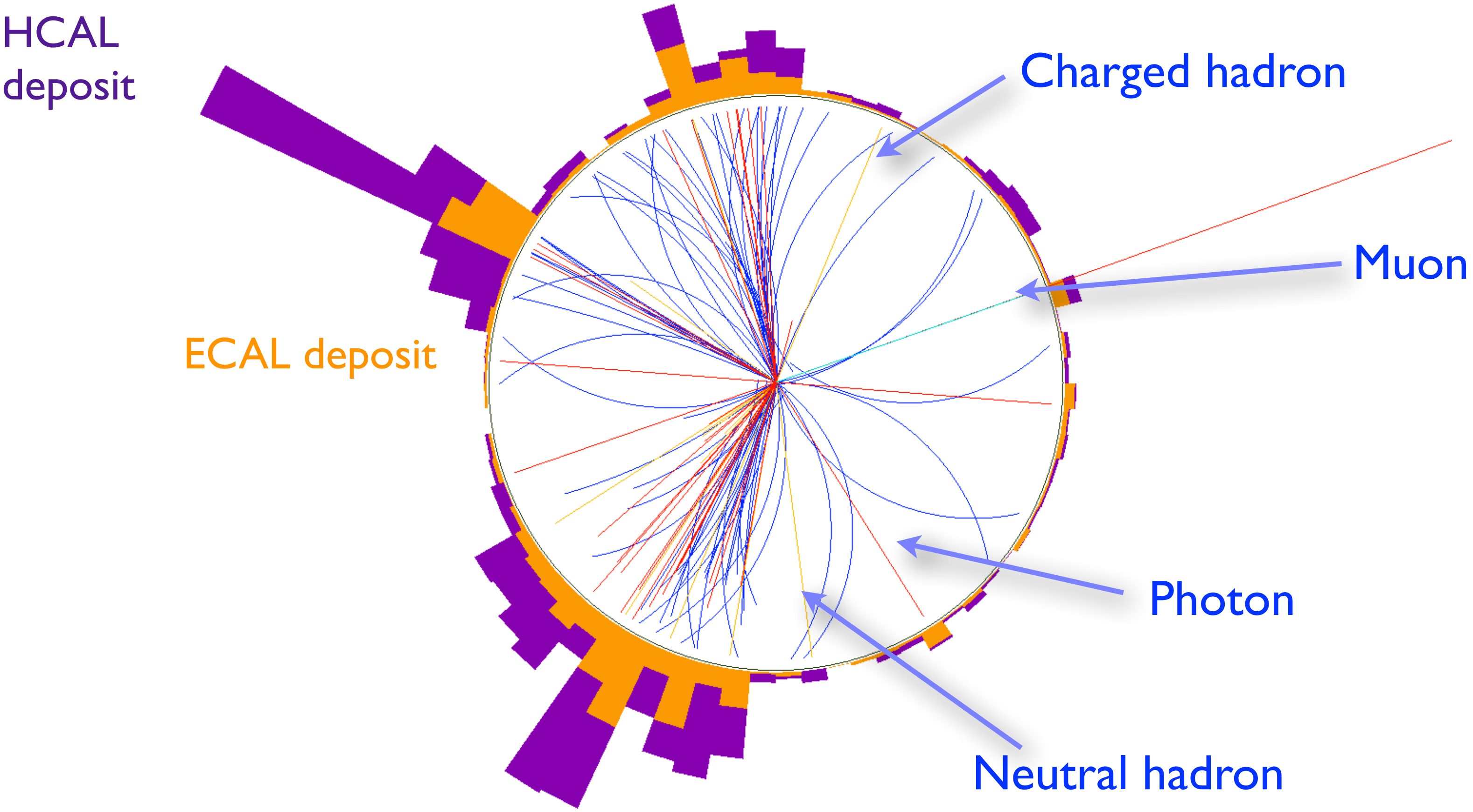
The CMS detector



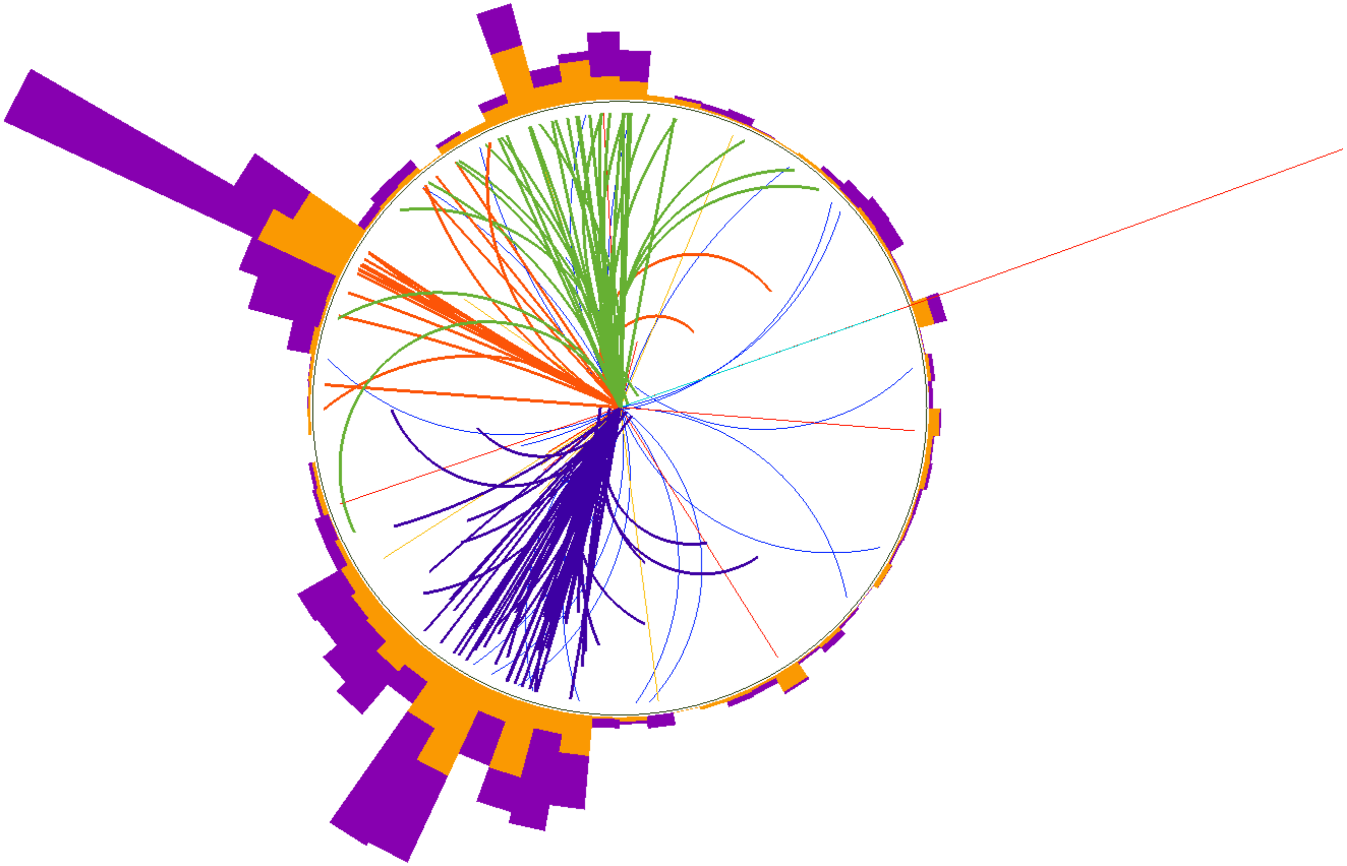
The CMS detector



Zooming in

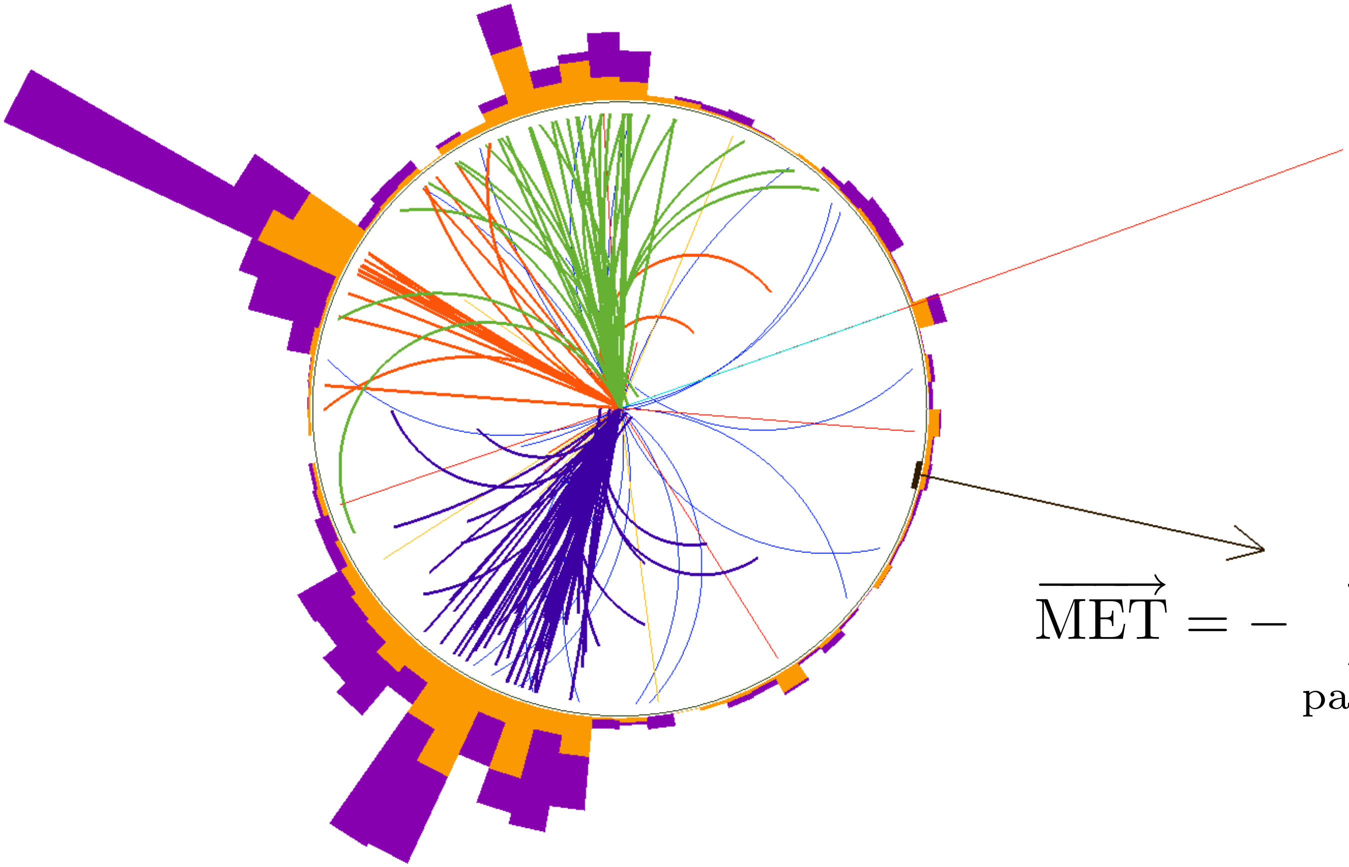

HCAL deposit
ECAL deposit
Charged hadron
Muon
Photon
Neutral hadron

Particle jets



Missing transverse momentum

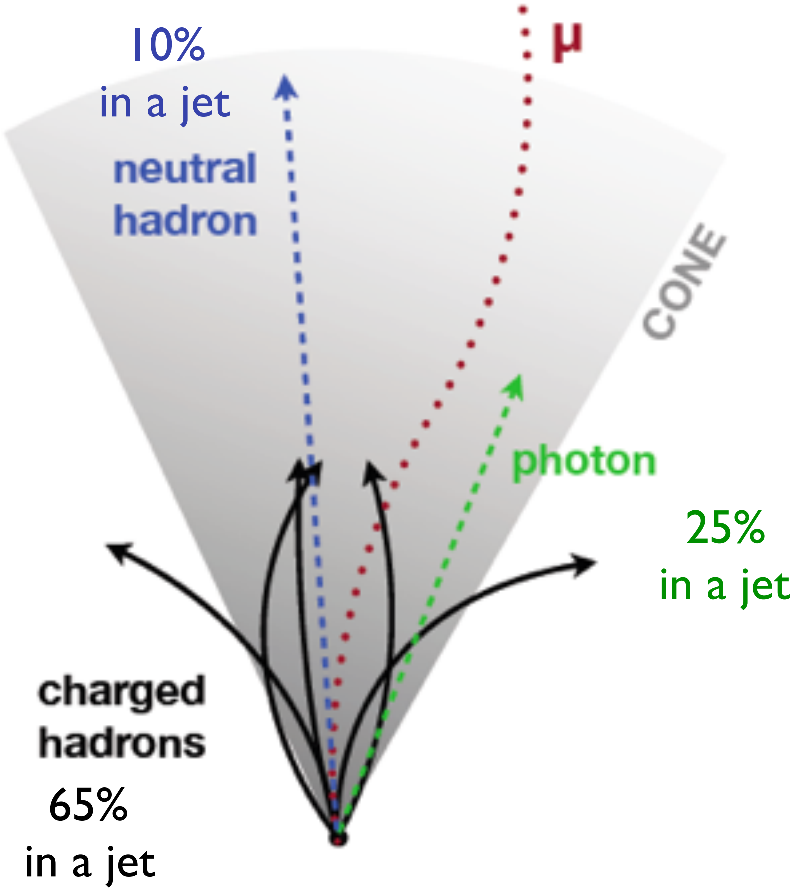
and not "energy" !



$$\overrightarrow{\mathrm{MET}} = -\sum_{\mathrm{particles}} \vec{p}_{\mathrm{T}}$$

For the first time at a hadron collider !

Expectations for the jet energy response & resolution



About 90% of the jet energy is carried by charged-hadrons and photons

Even in high-pT jets, the average p_T of the stable particles remains usually soft

As a result, 90% of the energy is measured with a high precision:

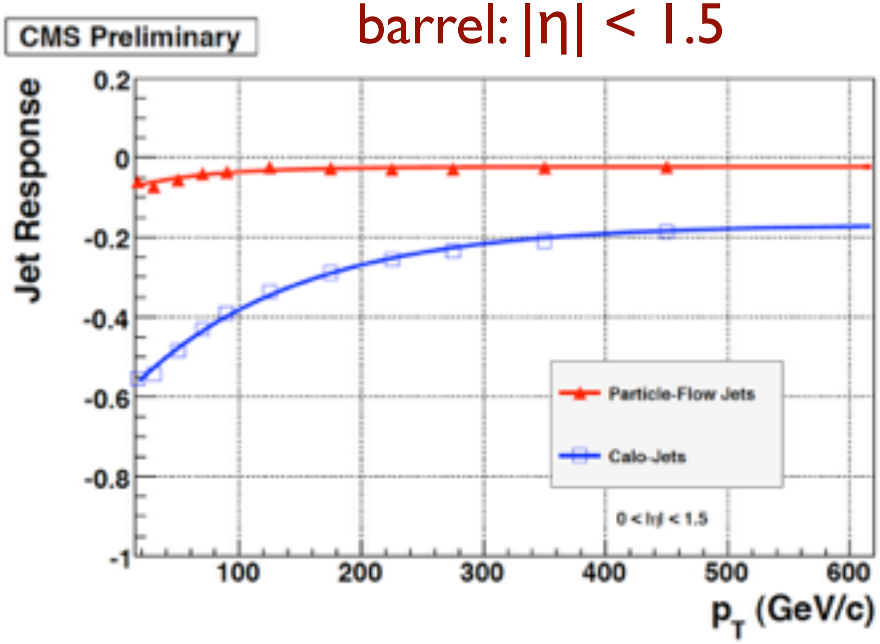
- Tracking: resolution < 1%
- Photons: $2.7\%/\sqrt{E} \oplus 0.5\%$ barrel
  $5.7\%/\sqrt{E} \oplus 0.5\%$ endcaps

The remaining 10% are measured with the HCAL resolution

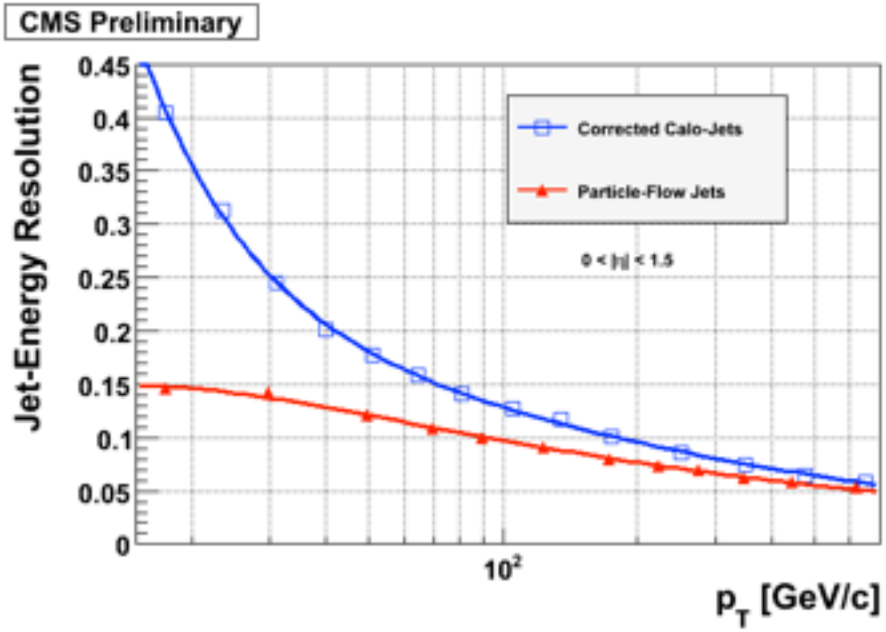
- Hadrons: $120\% / \sqrt{E}$

Jet energy response & resolution

simulated QCD-multijets events
barrel: $|\eta| < 1.5$



95-97% of the p_T reconstructed, over the whole range

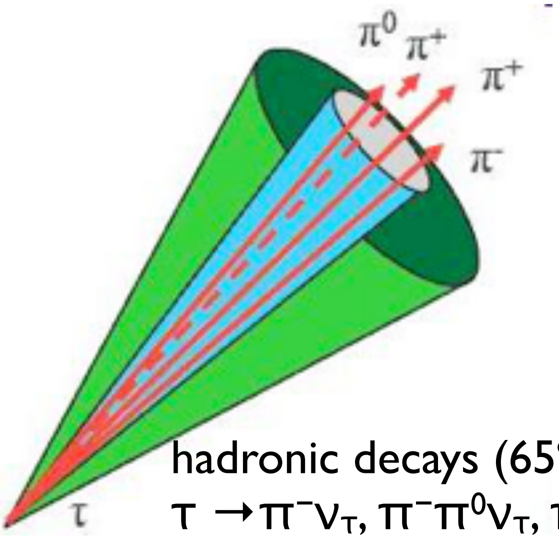


Very large improvement at low p_T, thanks to the tracks

Not mentioning:

- the gain in angular resolution (factor 2-3)
- much reduced dependency on the jet parton flavour (< 2% for jets p_T > 20 GeV) instead of ~10%

Tau lepton reconstruction

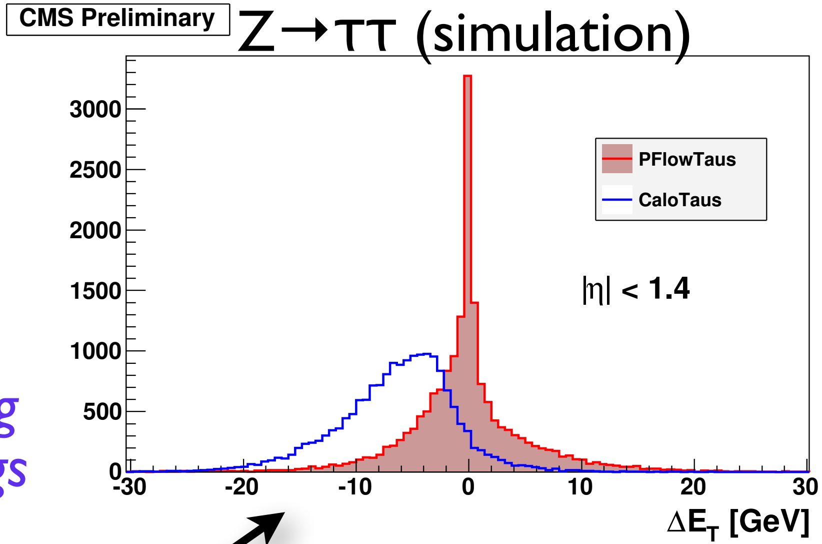


hadronic decays (65%)

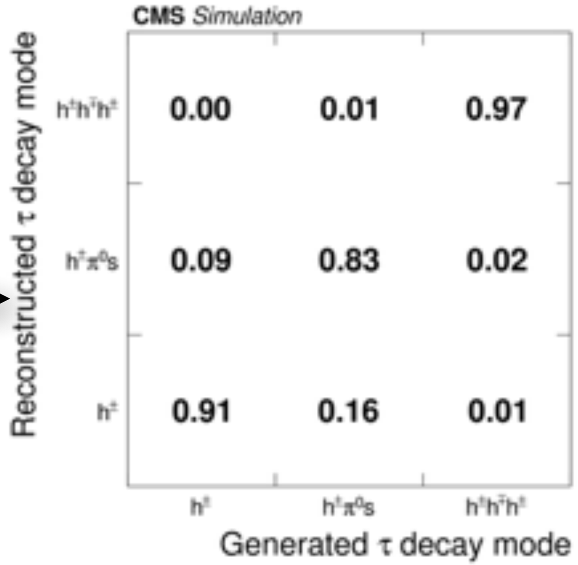
$\tau \rightarrow \pi^-\nu_\tau, \pi^-\pi^0\nu_\tau, \pi^-(n\pi^0)\nu_\tau$ *1-prong*

$\tau \rightarrow \pi^+\pi^-\pi^-(n\pi^0)\ \nu_\tau$ *3-prongs*

(Reconstructed - simulated) visible energy



- Immediate and important benefit of using self-calibrated particles
- Access to the τ-decay channel

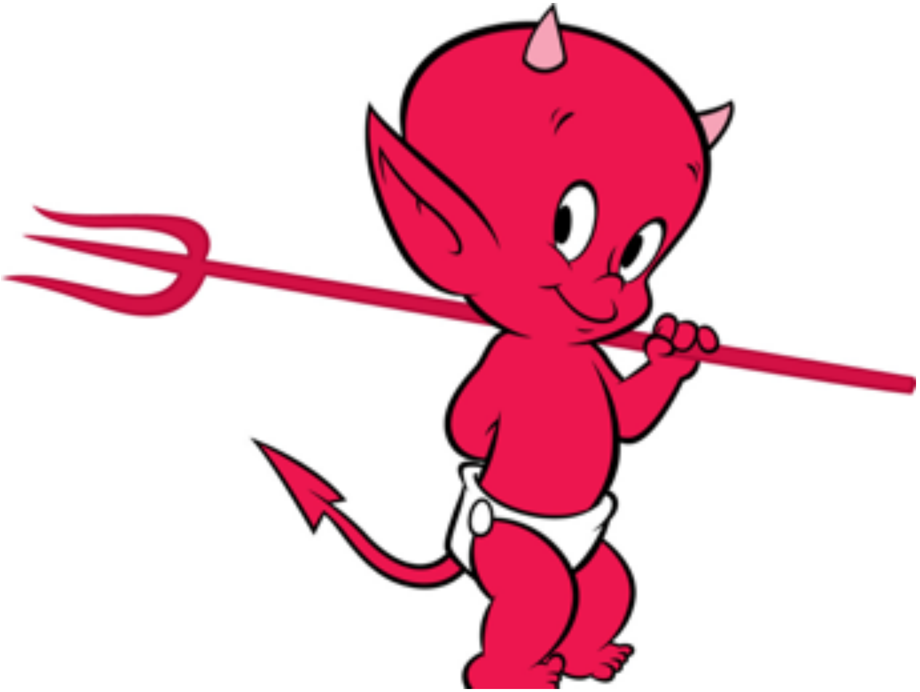


CMS *Simulation*

Reconstructed τ decay mode	$h^\pm$	$h^\pm\pi^0$s	$h^\pm h^\mp h^\pm$
$h^\pm h^\mp h^\pm$	0.00	0.01	0.97
$h^\pm\pi^0$s	0.09	0.83	0.02
$h^\pm$	0.91	0.16	0.01

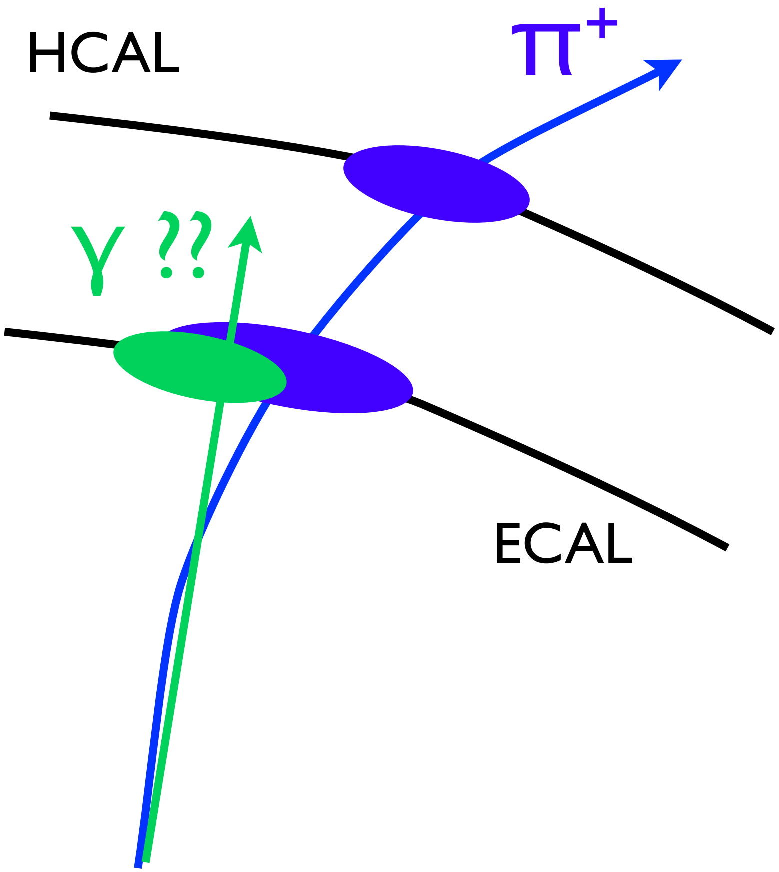
Generated τ decay mode

A bit more into the details



Charged hadrons, overlapping neutrals

How to deal with neutrals overlapping with charged hadrons ?
→ exploit redundancy !



For each HCAL cluster, compare:

- Sum of track momenta p
- Calorimeter energy E
    - Linked to the tracks
    - Calibrated for hadrons
      $a(E)E_{ECAL}+b(E)E_{HCAL}+\text{offset}$

E and p compatible

- Charged hadrons

$E > p + 120\% \sqrt{p}$

- Charged hadrons +
  Photon / neutral hadron

Precedence is given to photons as they are more abundant than neutral hadrons

$E \ll p$

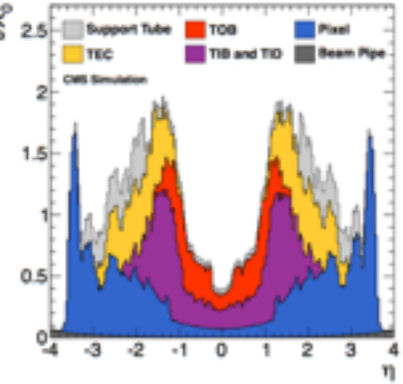
- Need attention …
  Rare: muon, fake track

Florian Beaudette – LLR

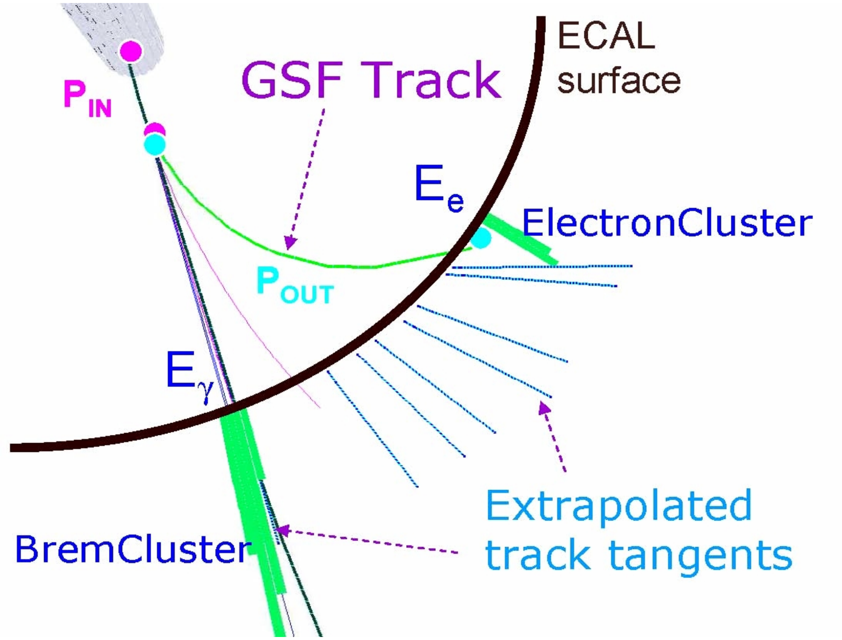
Electrons

The PFlow needs to reconstruct all the electrons isolated or not, at all p_Ts

- if not identified, the energy of a Bremsstrahlung photon cluster is counted twice
- in a jet, the electron should not swallow the entire jet



Material: up to $2X_0$ in front of the ECAL
Magnetic field: 3.8T



The electron track reconstruction is able to follow the electron up to the ECAL entrance, a Gaussian Sum Filter fit is then carried out

For each tracker layer where most of the material is located: mimic a Brem emission (tangent straight-line extrapolation)

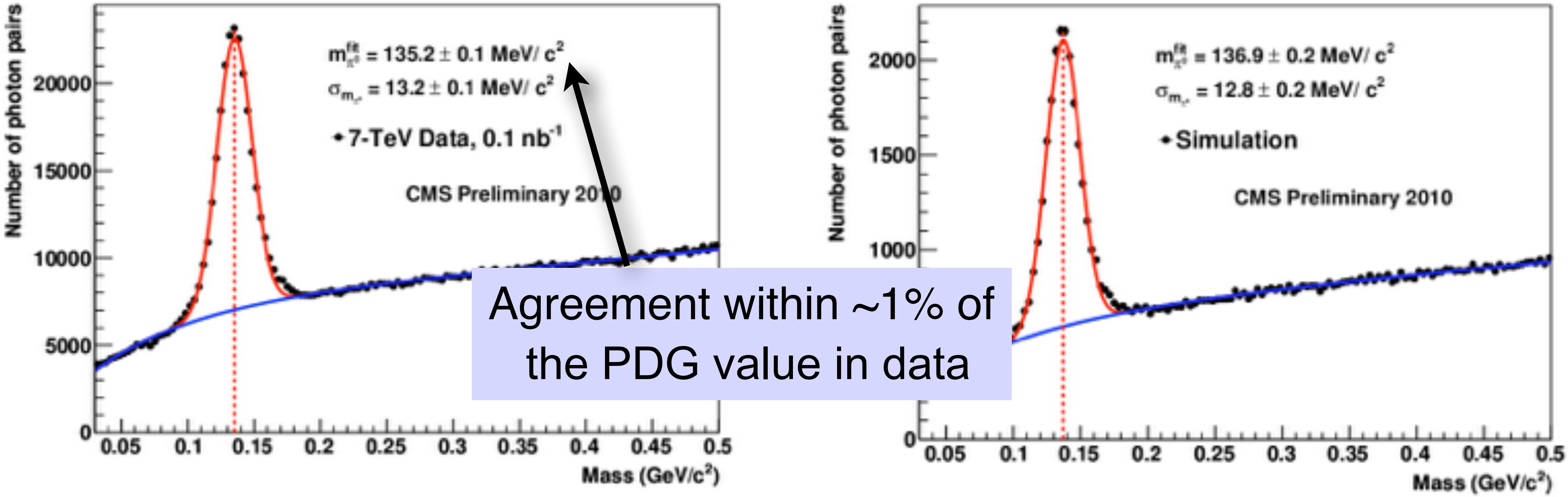
Catch the electron cluster with a helix extrapolation from the outermost state of the track

Then, specific procedures to identify clusters from converted Brems

Finally, a loose electron-ID is applied

Photons

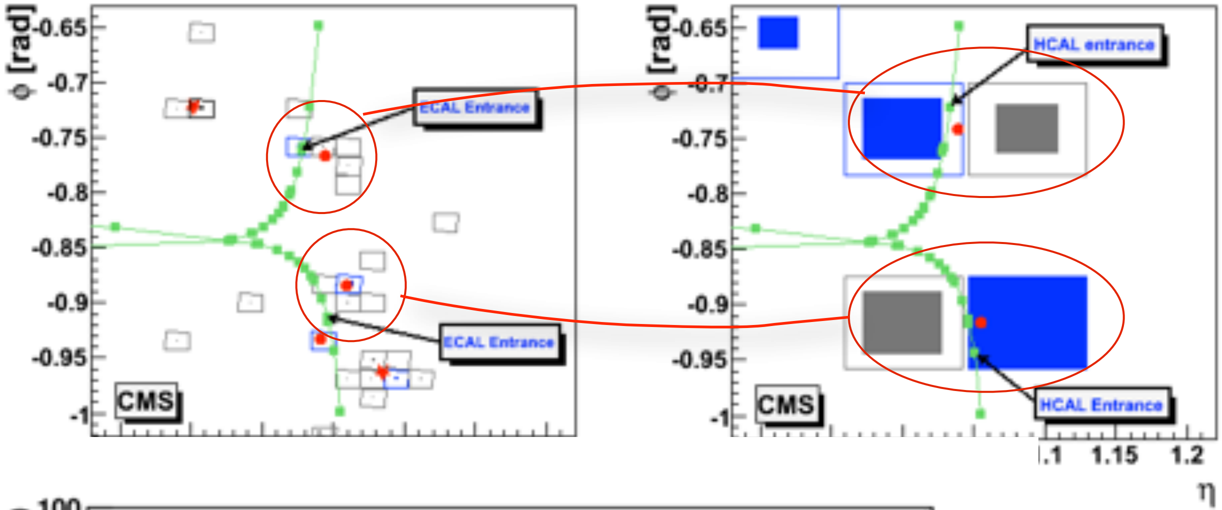
Commissioning with π⁰→γγ in 2010



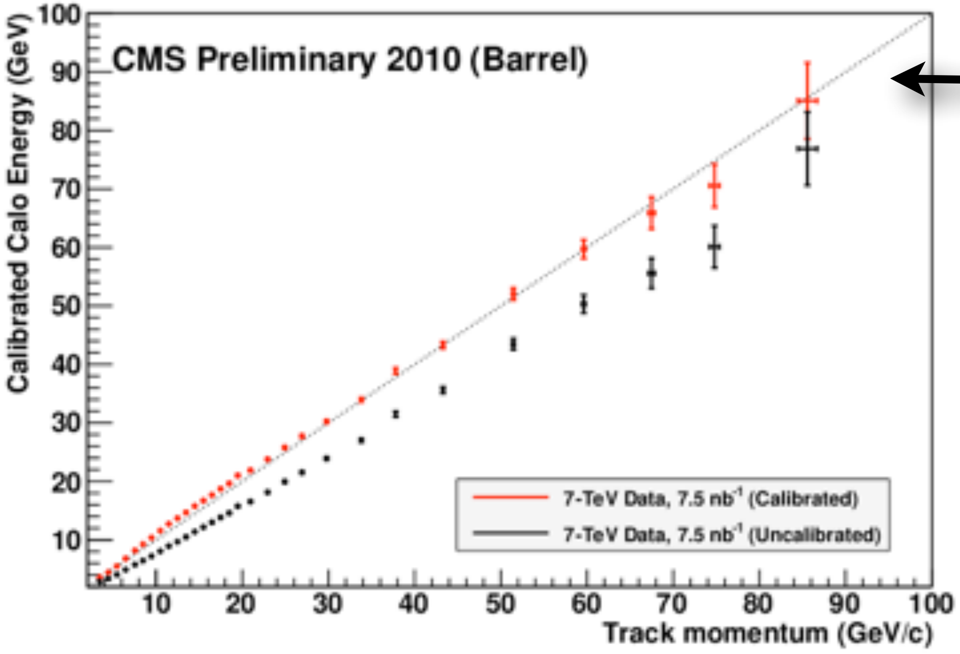
Agreement within ~1% of the PDG value in data

Demonstrates the suitability of the absolute ECAL calibration
(has improved since)

Calorimeter response to hadrons



Obtained with charged hadrons

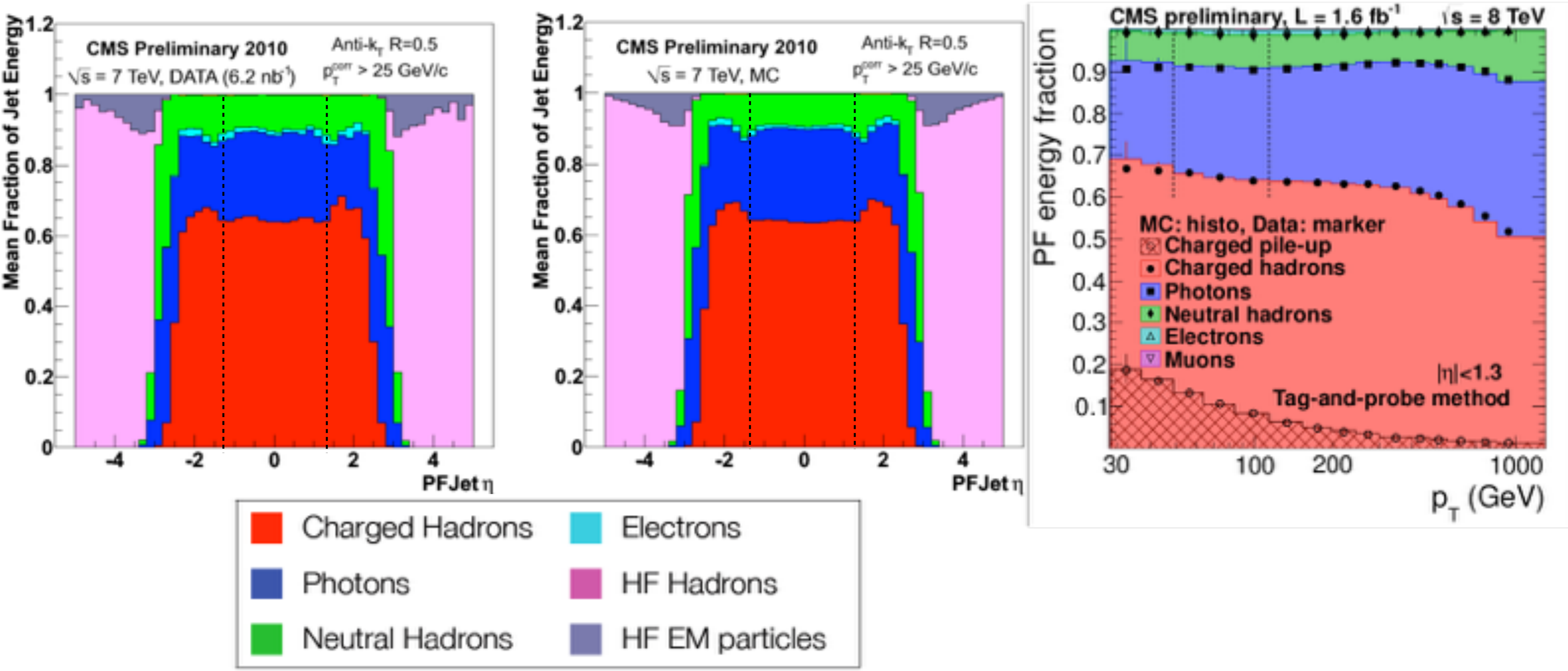


Calorimeter response is important for neutral hadrons

The calorimeter response to hadrons is well simulated.
The hadron response in the calorimeters is adequate at the 5% level (2010 result)
(similar agreement in the end-caps)

Currently around 2%

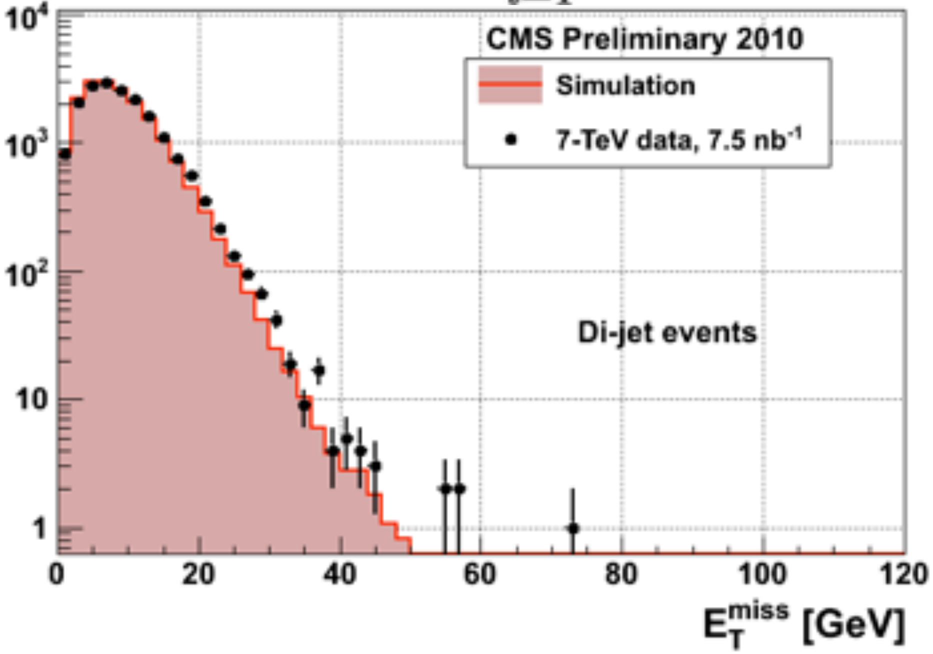
Jet composition

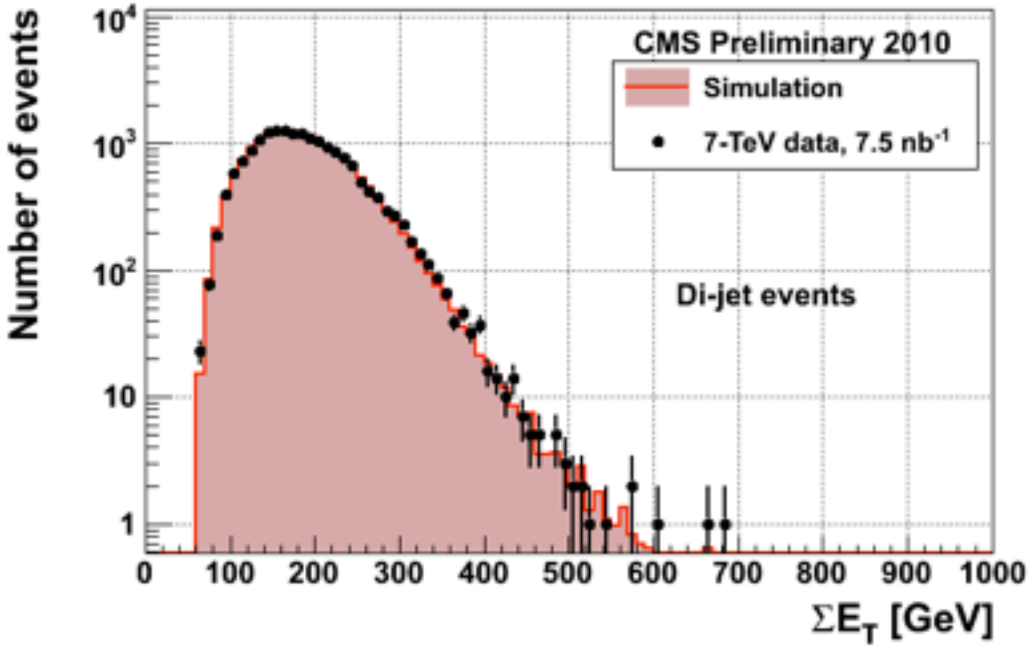


The agreement of the first days is confirmed with high statistics, even in presence of pile-up

Missing transverse momentum

$$\overrightarrow{E_T^{miss}} = -\sum_{i=1}^{N_{particles}} \overrightarrow{E_T}^{i}$$





Agreement over 3 orders of magnitude for the E_T^{miss}

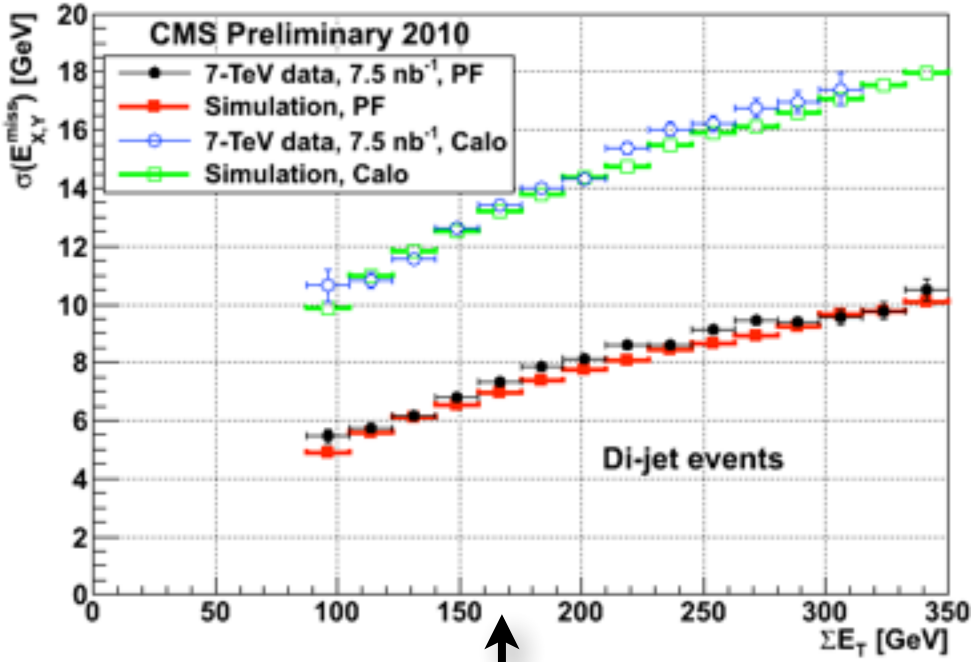
Even more challenging: the scalar sum of the particle p_T were no cancellation can occur !

All together, a remarkable agreement is obtained on these quantities known to be difficult to reproduce at hadron colliders

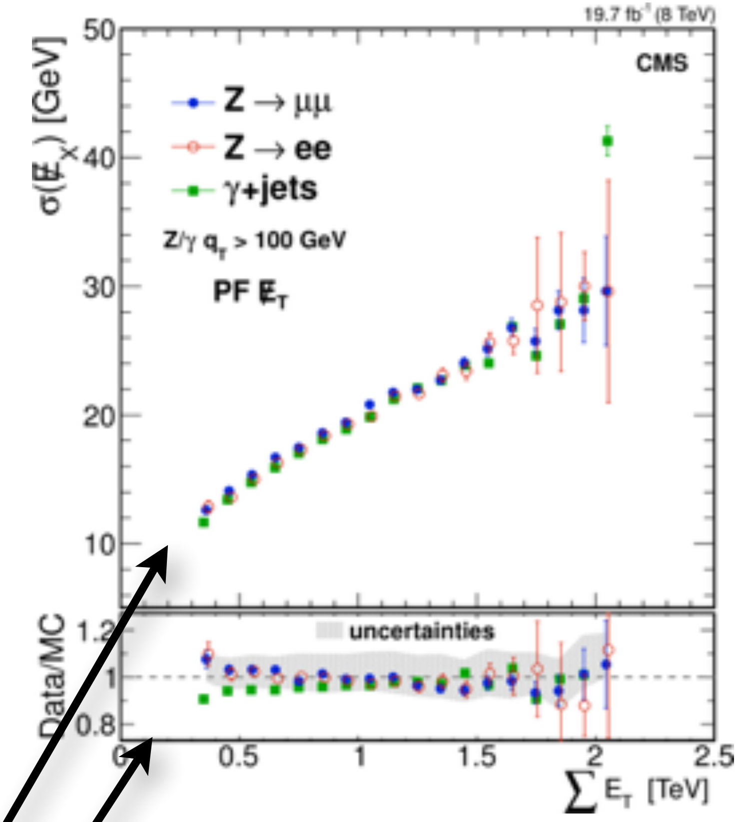
- robustness of the algorithm
- precise simulation of the detectors

Missing transverse momentum resolution

jet $p_T > 20$ GeV



jet $p_T > 100$ GeV



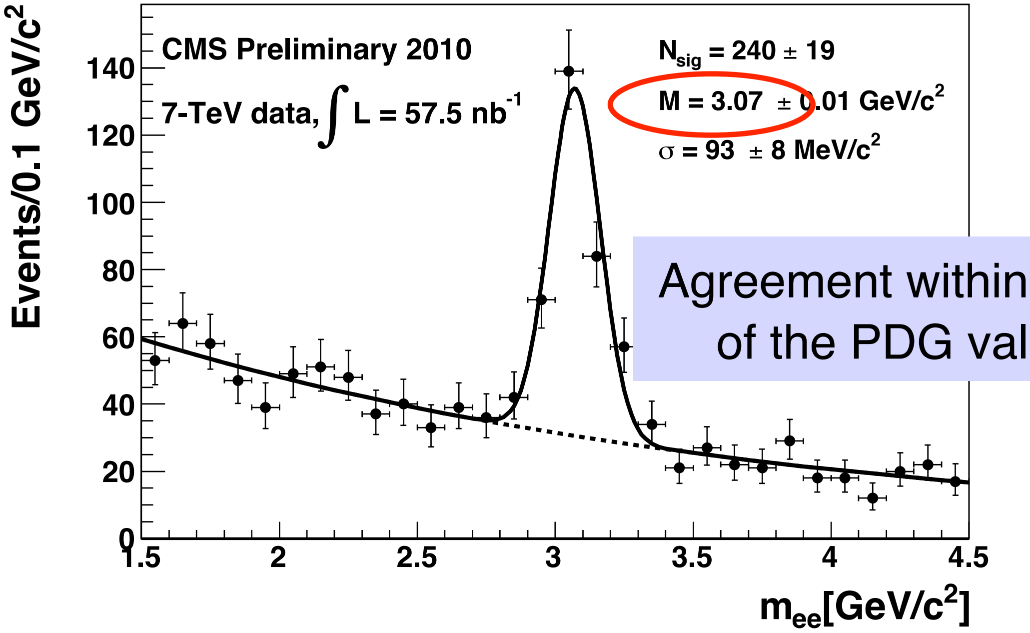
The E_T^{miss} resolution is improved by a factor of ~2 wrt. the calorimeter-based E_T^{miss}

Nowadays, boosted Z's and γ+jets are used, the data/simulation agreement is impressive

Electrons & Photons

From 2010 .... to ...2013

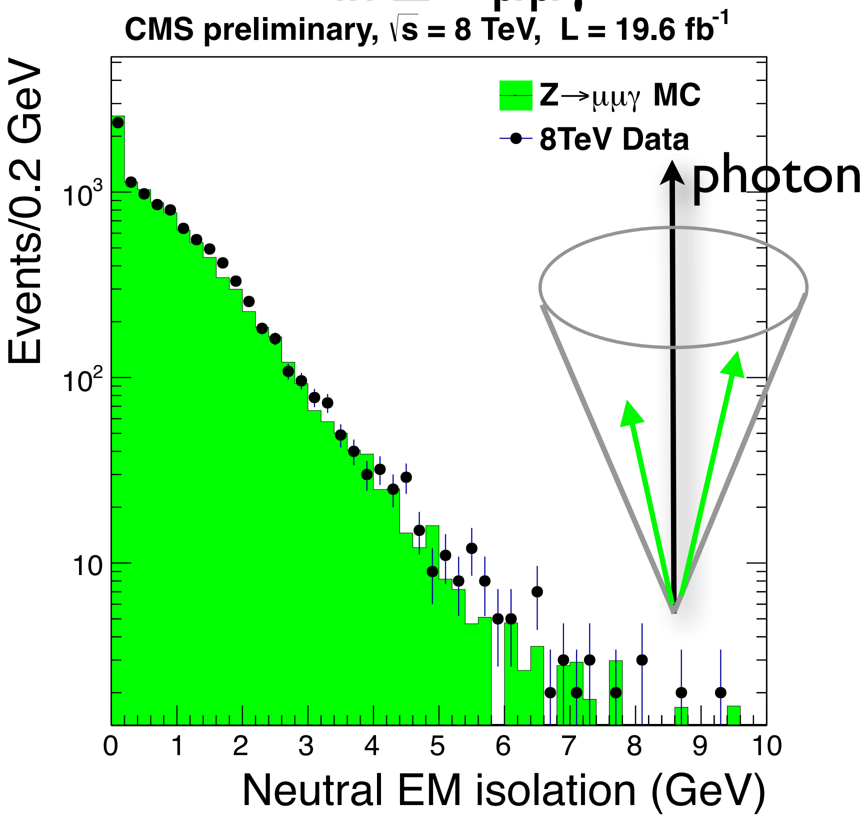
$J/\Psi \rightarrow ee$



Agreement within ~1% of the PDG value

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

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



With large statistics, rare and clean events can be used to carry out the detailed performance studies required by the high precision analyses done nowadays

Effects on physics analyses

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

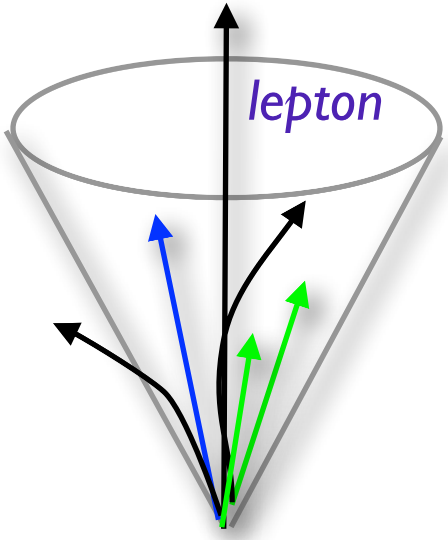
- improved performance, pile-up control

Physics analyses

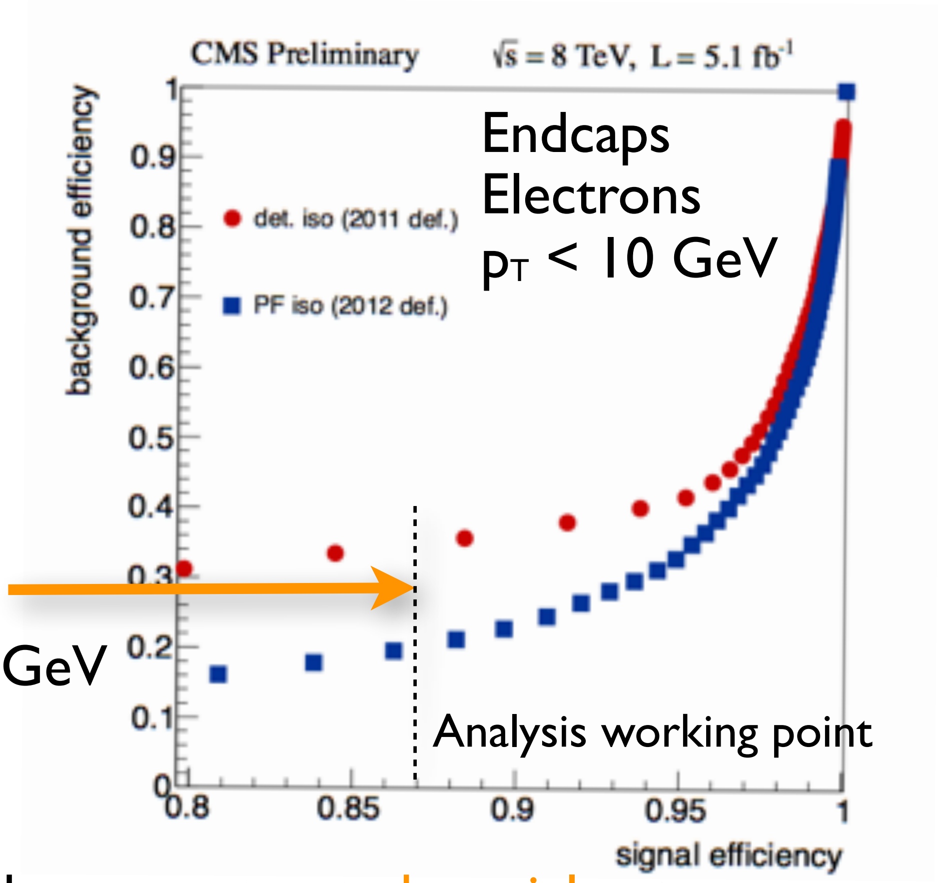
- Better trigger for jets, MET, taus (PF@HLT)
- FSR photon recovery in HZZ
- embedding in H→ττ
- jet substructure

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



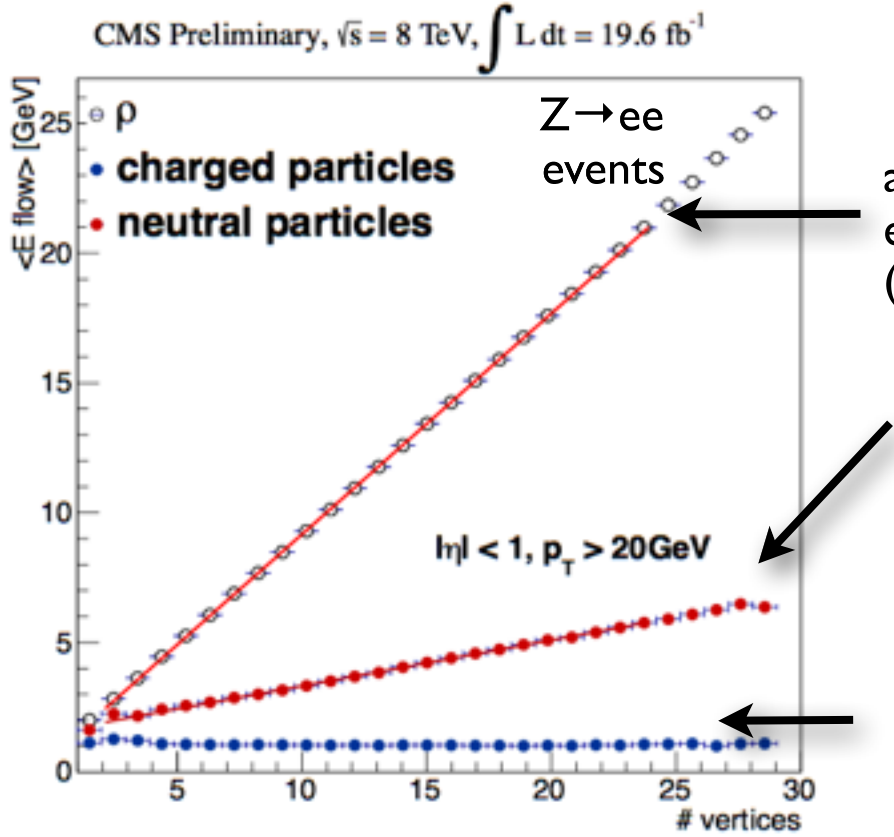
Endcaps
Electrons
$p_T < 10$ GeV

Analysis working point

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

Lepton/Photon isolation : pile-up mitigation



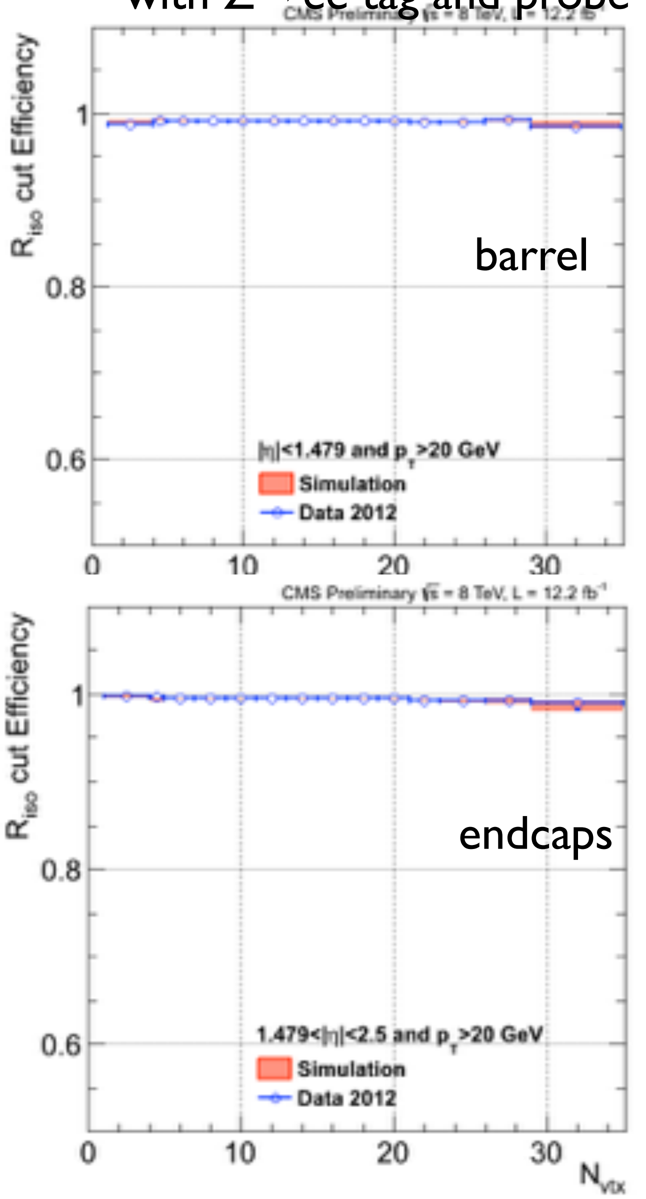
Z→ee events

average event energy density (fastjet)

Neutral particles around leptons

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

Isolation efficiency measurement with Z→ee tag and probe

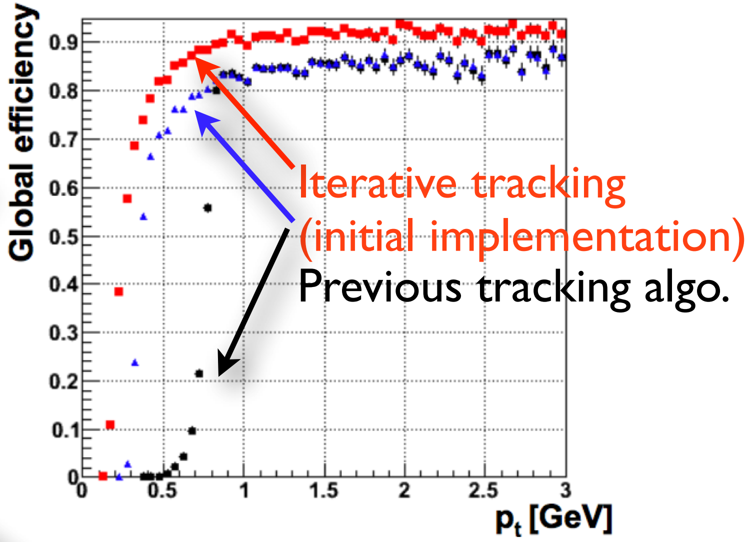


- No correction needed for the charged hadrons (vertex constraint)
- For the neutrals: the PU contribution in the cone is estimated (proportional to the energy density) and subtracted

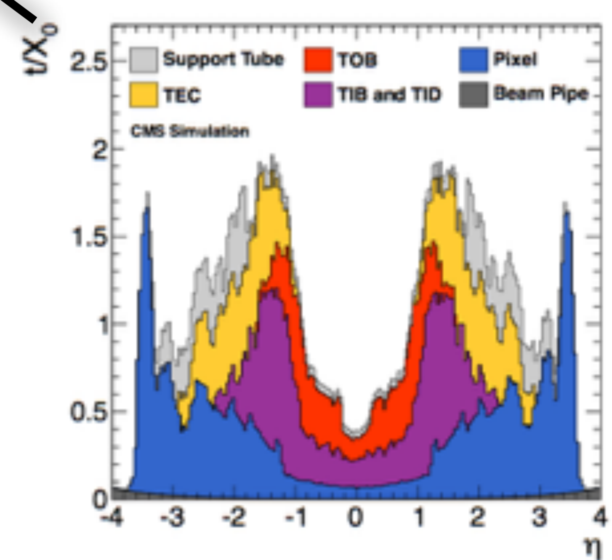
Ingredients for a good Particle Flow

- Separate neutrals *from* charged hadrons
  - Field integral (BxR)
  - Calorimeter granularity
- Efficient tracking
- Minimize material before calorimeters
- Clever algorithm to compensate for detector imperfections

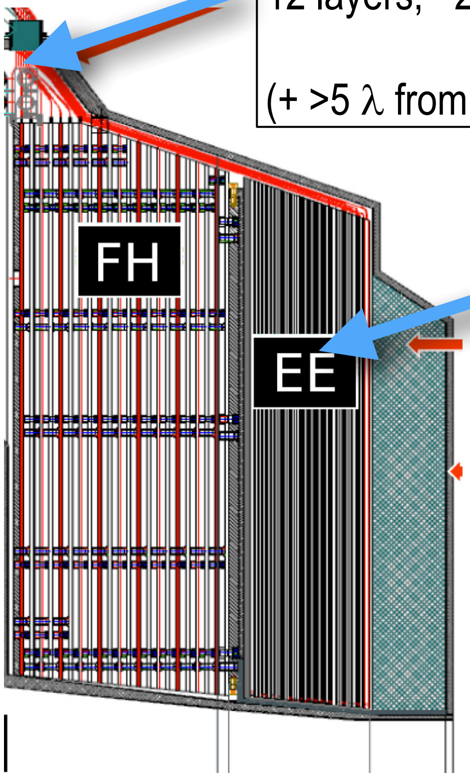
CMS: 4.9 T.m
(x~2 existing and
past detectors)



Iterative tracking
(initial implementation)
Previous tracking algo.



HGCAL



HGC-HCAL: Si+Brass or Steal
12 layers, ~26 X_0 (>3.5 λ)

(+ >5 λ from BH)

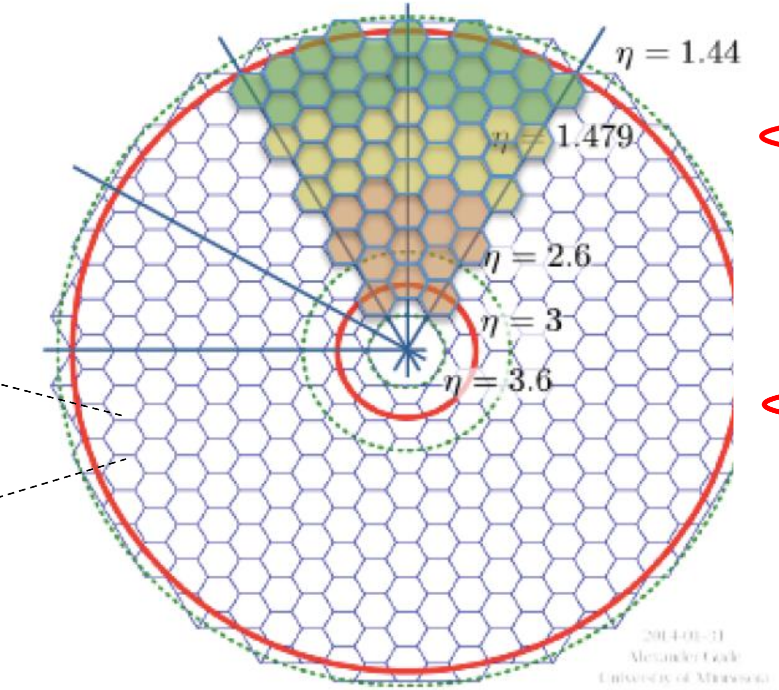
Operation at -30°C via CO_2 Cooling
(to mitigate Si leakage current)

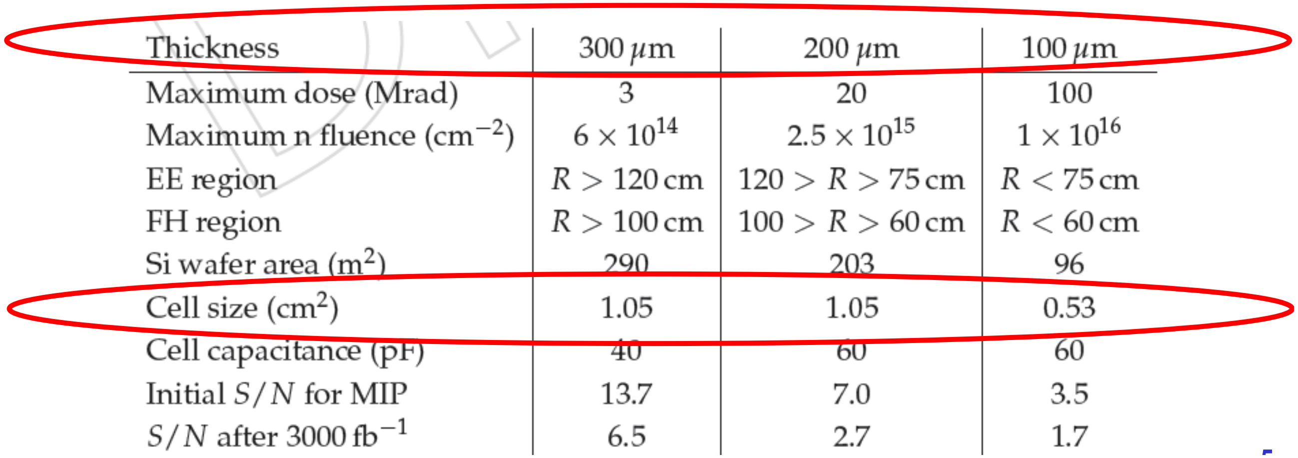
HGC-ECAL: Si+W/Cu
28 layers, ~26 X_0 (1.5 λ)
10 x 0.65 X_0 +
10 x 0.88 X_0 +
8 x 1.26 X_0

	EE	FH	Total
Area of silicon (m^2)	380	209	589
Channels	4.3M	1.8M	6.1M
Detector modules	13.9k	7.6k	21.5k
Weight (one endcap) (tonnes)	16.2	36.5	52.7
Number of Si planes	28	12	40

The entire current ECAL has O(80k) channels

Geometry, characteristics





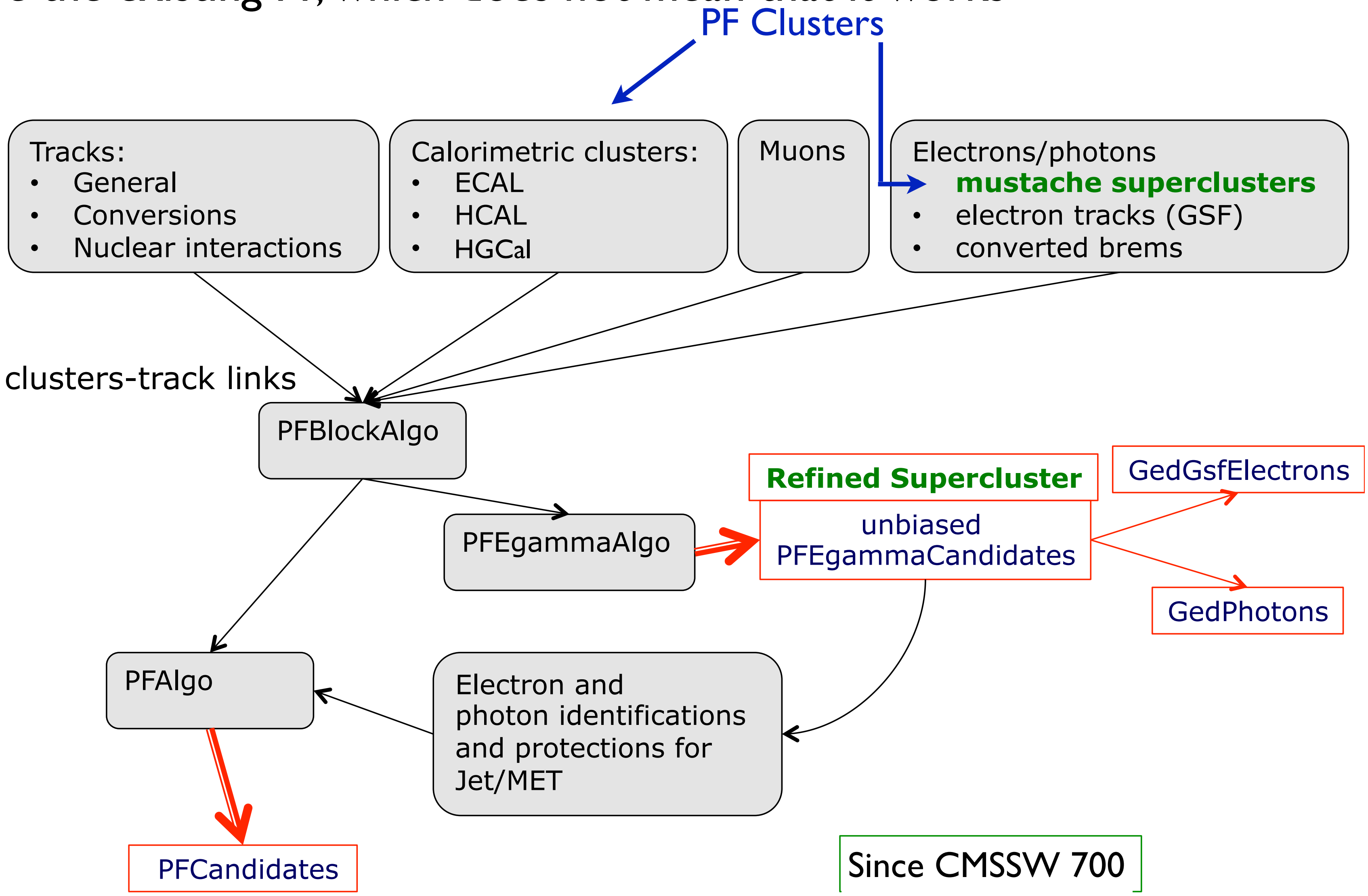
Thickness	300 μm	200 μm	100 μm
Maximum dose (Mrad)	3	20	100
Maximum n fluence (cm^{-2})	6×10^{14}	2.5×10^{15}	1×10^{16}
EE region	$R > 120$ cm	$120 > R > 75$ cm	$R < 75$ cm
FH region	$R > 100$ cm	$100 > R > 60$ cm	$R < 60$ cm
Si wafer area (m^2)	290	203	96
Cell size (cm^2)	1.05	1.05	0.53
Cell capacitance (pF)	40	60	60
Initial S/N for MIP	13.7	7.0	3.5
S/N after 3000 fb^{-1}	6.5	2.7	1.7

It has many common aspects with a tracker (silicon, high granularity) but

- we want to measure energies, so we need a clustering algorithm
- the size of clusters is large, the clusters can start at any depth
- the showers fluctuate much more than a track does
- the multiplicity of hits for showers is much larger than for tracks

Status of the PF with HGCAL

We have the existing PF, which does not mean that it works



PF Clusters

Tracks:
- General
- Conversions
- Nuclear interactions

Calorimetric clusters:
- ECAL
- HCAL
- HGCal

Muons

Electrons/photons
- **mustache superclusters**
- electron tracks (GSF)
- converted brems

clusters-track links

PFBlockAlgo

PFEgammaAlgo

Refined Supercluster

unbiased PFEgammaCandidates

GedGsfElectrons

GedPhotons

PFAlgo

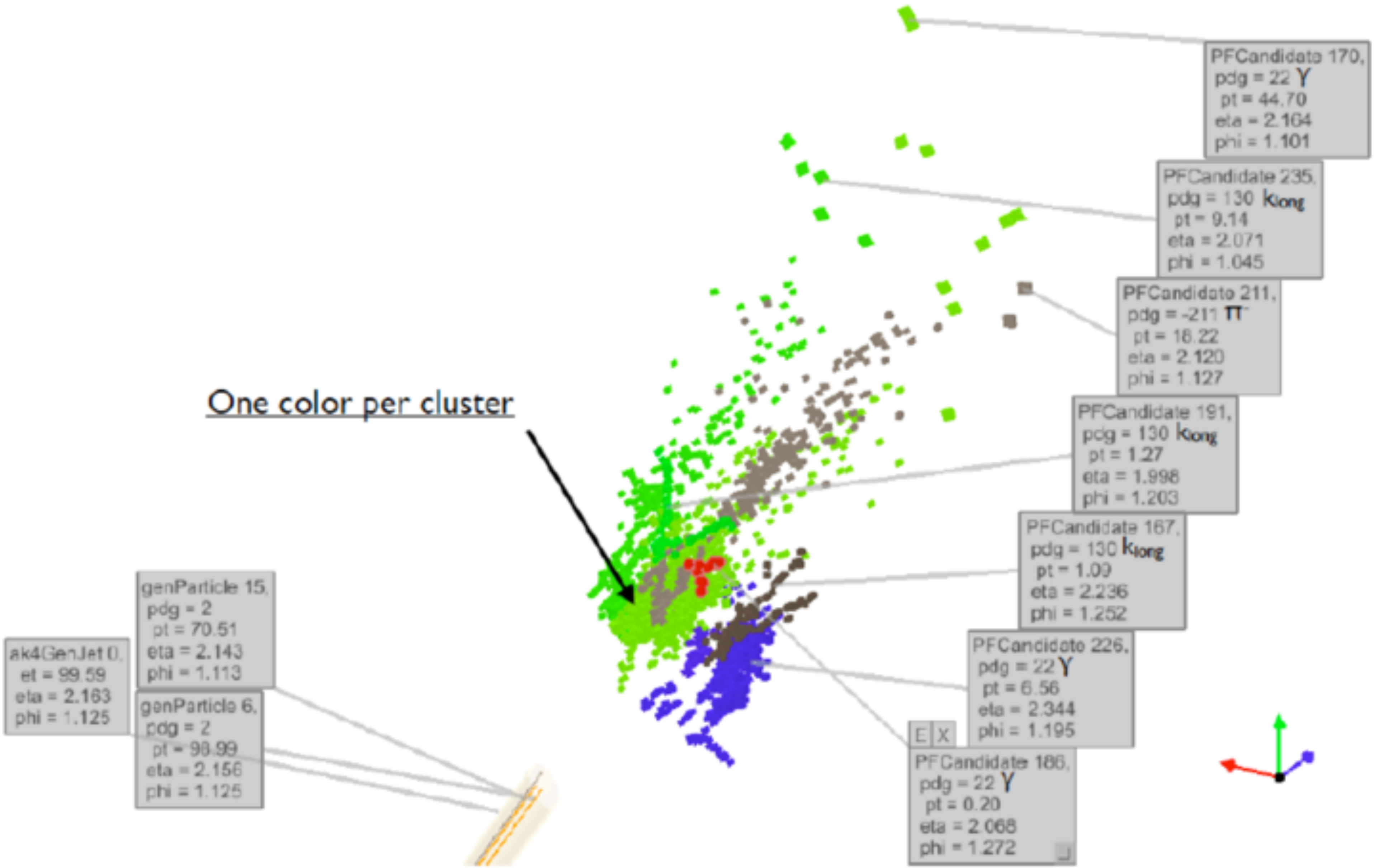
Electron and photon identifications and protections for Jet/MET

PFCandidates

Since CMSSW 700

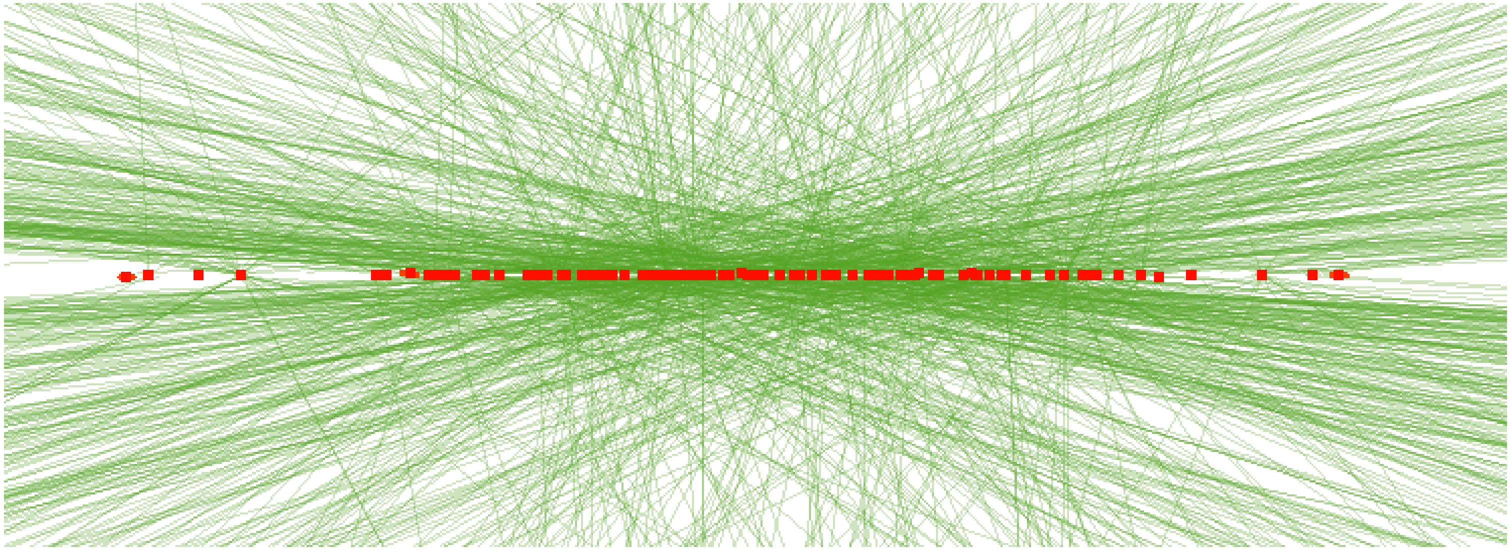
“The PF development can only really start once the clustering is hammered down”

Jet tomography


One color per cluster

PU & Timing

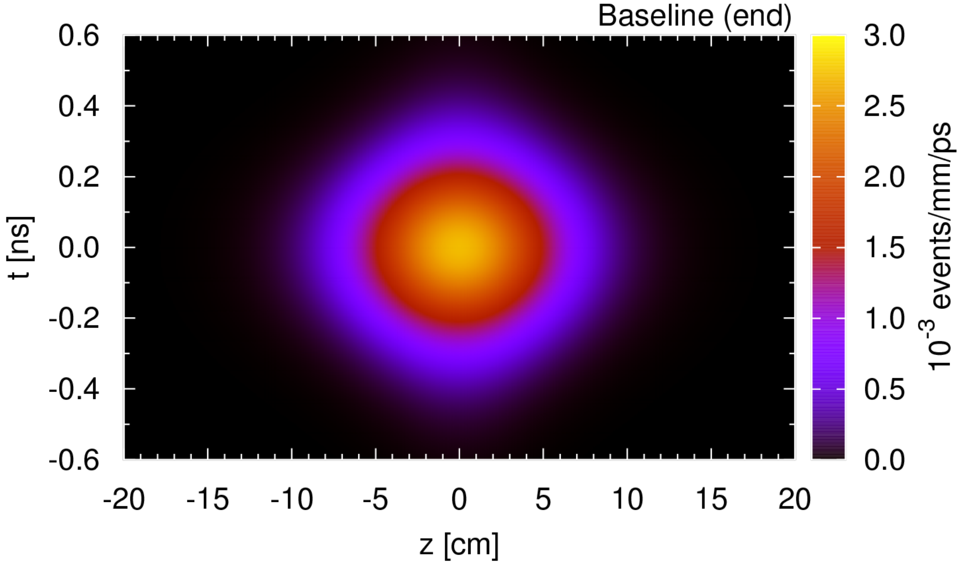
Collisions every 25ns induce out-of-time Pile-Up
Tackled through fast detector response and fast shaping



140 PU
on top of
a physics
event

Interactions are spread over space and time over 100-200 ps

→Use precise timing to disentangle overlapping vertices



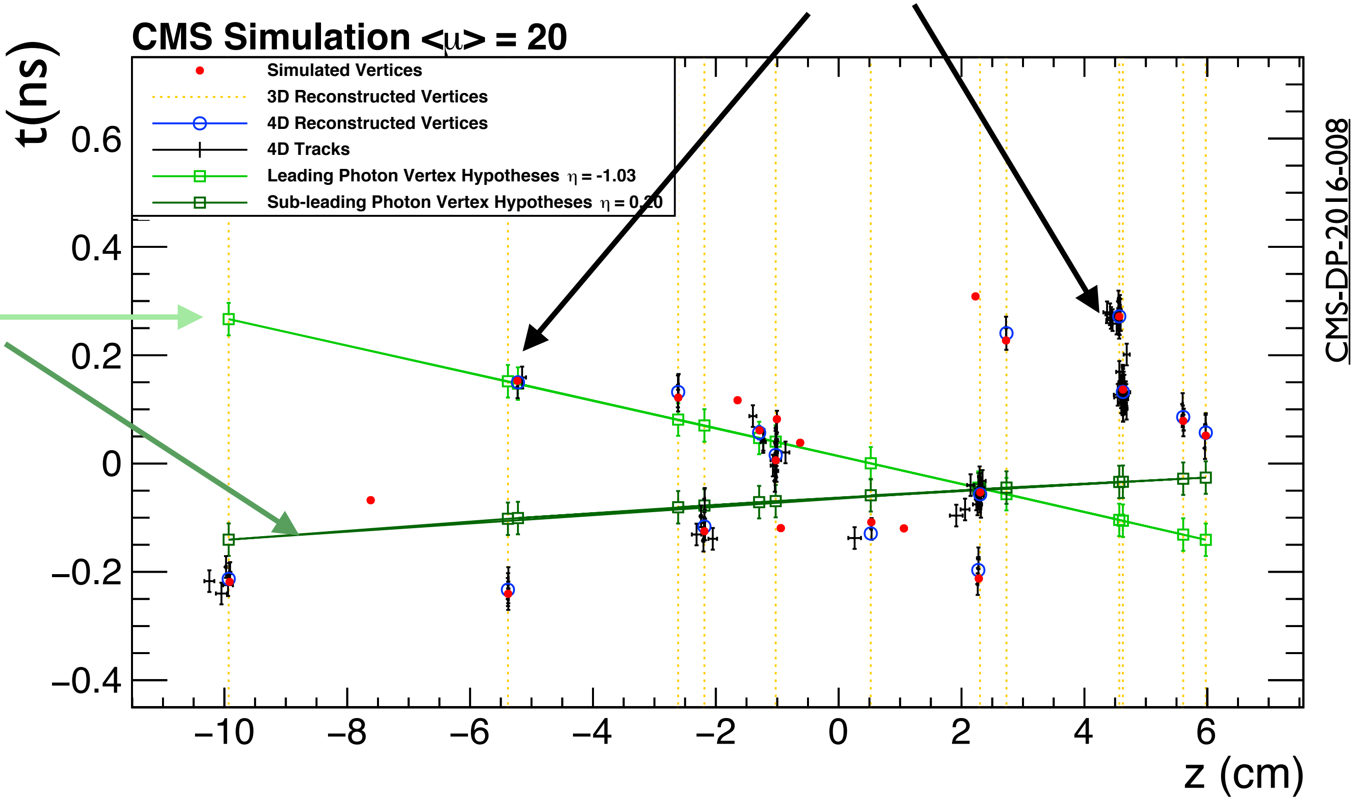
Timing

A cell resolution of 50ps provides a 20ps resolution for the clusters for E>10 GeV

Associating a track to a cluster provides timing information to the track
→4D vertexing can be carried out.
→PU mitigation

Reconstructed 3D/4D vertices

The timing information is also available for photons →(z,t) constraint that can be correlated with the 4D vertexing



Conclusion

The Particle Flow algorithm has been deployed in CMS

The concept is simple, the implementation required some work...

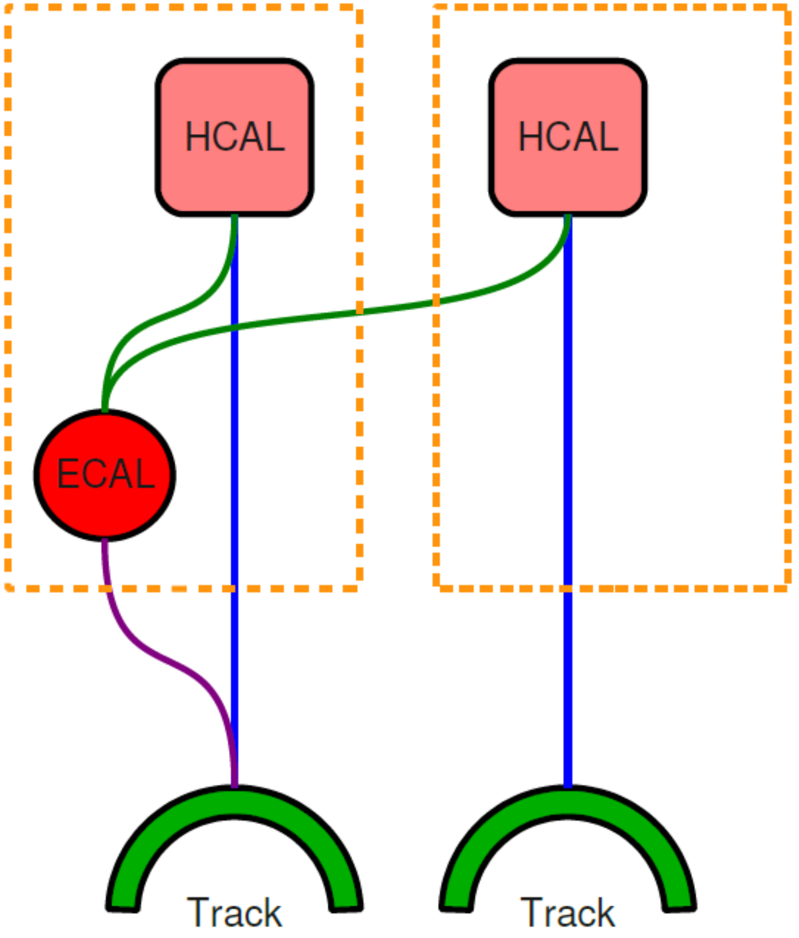
It has significantly improved the performance on jets, MET, taus, lepton isolation..

The PF has not only proven to work well with high PU, it is considered as the only way to preserve or even hopefully improve the performance of CMS
→Global Event Description

The future HGCAL detector will use a GED approach.
The new capabilities of the detector (granularity, timing) will certainly open up new possibilities for the PF and they will have to optimally exploited to mitigate the high PU.

Backup

Charged+neutrals: E > p



Significant excess of energy in the calorimeters:

$E > p + 120\% \sqrt{E}$

Charged hadrons [p_i]

Neutrals:

E from ECAL or HCAL only:

HCAL	h^0	[$E - p$]
ECAL	γ	[$E_{ECAL} - p/b$]

E from ECAL and HCAL:

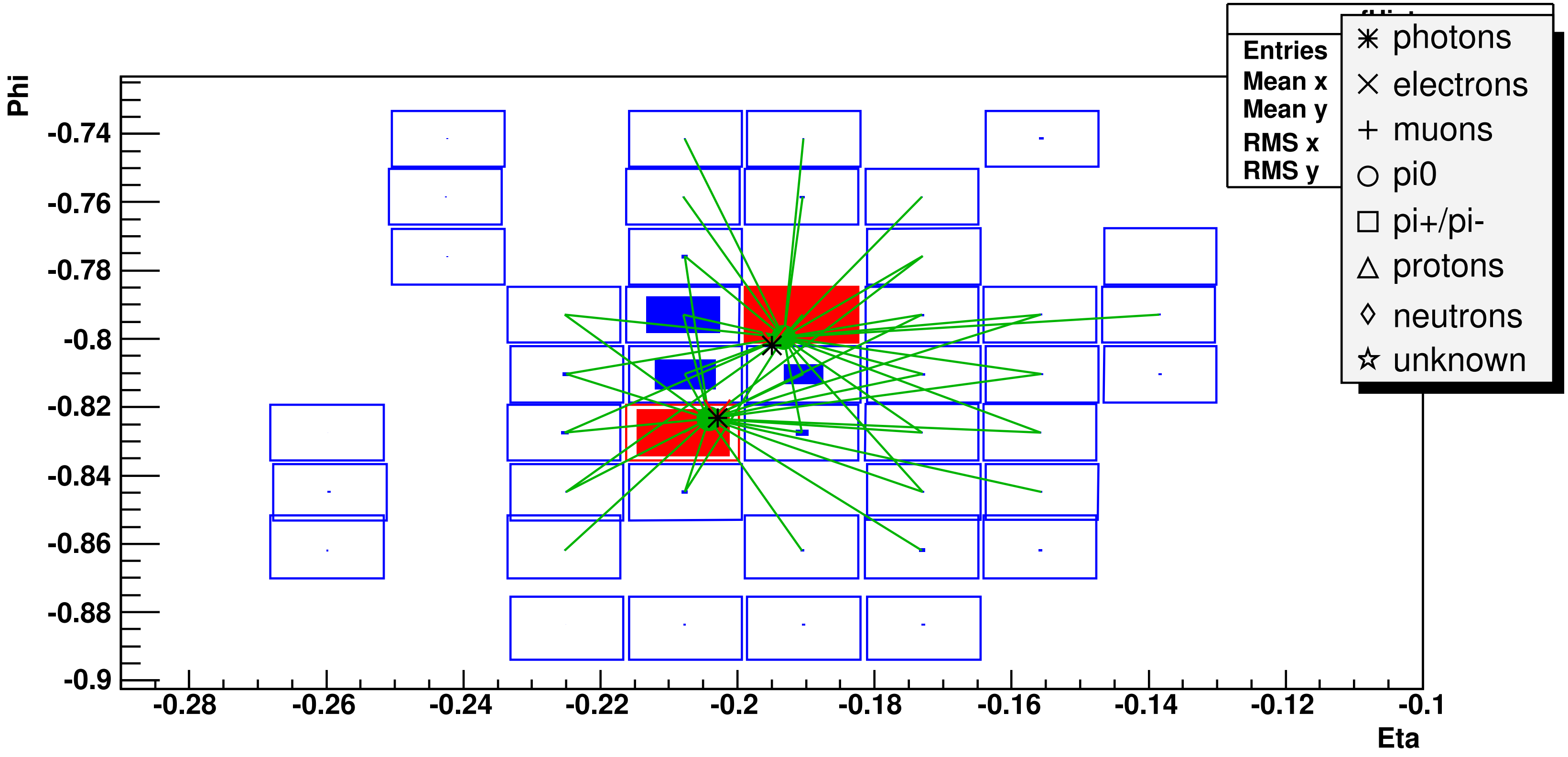
$E\text{-}p > E_{ECAL}$?

γ	[E_{ECAL}]
h^0	with the rest

Else:

γ [$(E - p) / b$]

Clustering


Entries
Mean x
Mean y
RMS x
RMS y
✳ photons
× electrons
+ muons
○ pi0
□ pi+/pi-
△ protons
◊ neutrons
☆ unknown
Phi
-0.74
-0.76
-0.78
-0.8
-0.82
-0.84
-0.86
-0.88
-0.9
-0.28
-0.26
-0.24
-0.22
-0.2
-0.18
-0.16
-0.14
-0.12
-0.1
Eta