Particle Flow Performance in CMS

Colin Bernet (IPNL)
for the CMS Collaboration

CLIC Workshop, CERN
25th of January, 2018

Menu

Ingrédients

CMS, tracks and clusters

La Recette

hadrons, photons, electrons, muons

Plat Principal

jets, MET, taus, leptons (and dealing with pileup)

Desserts

the future: HGCal, deep learning

Jinst

Published by IOP Publishing for Sissa Medialab

Received: *June 15, 2017*
Accepted: *September 11, 2017*
Published: *October 6, 2017*

Particle-flow reconstruction and global event description with the CMS detector

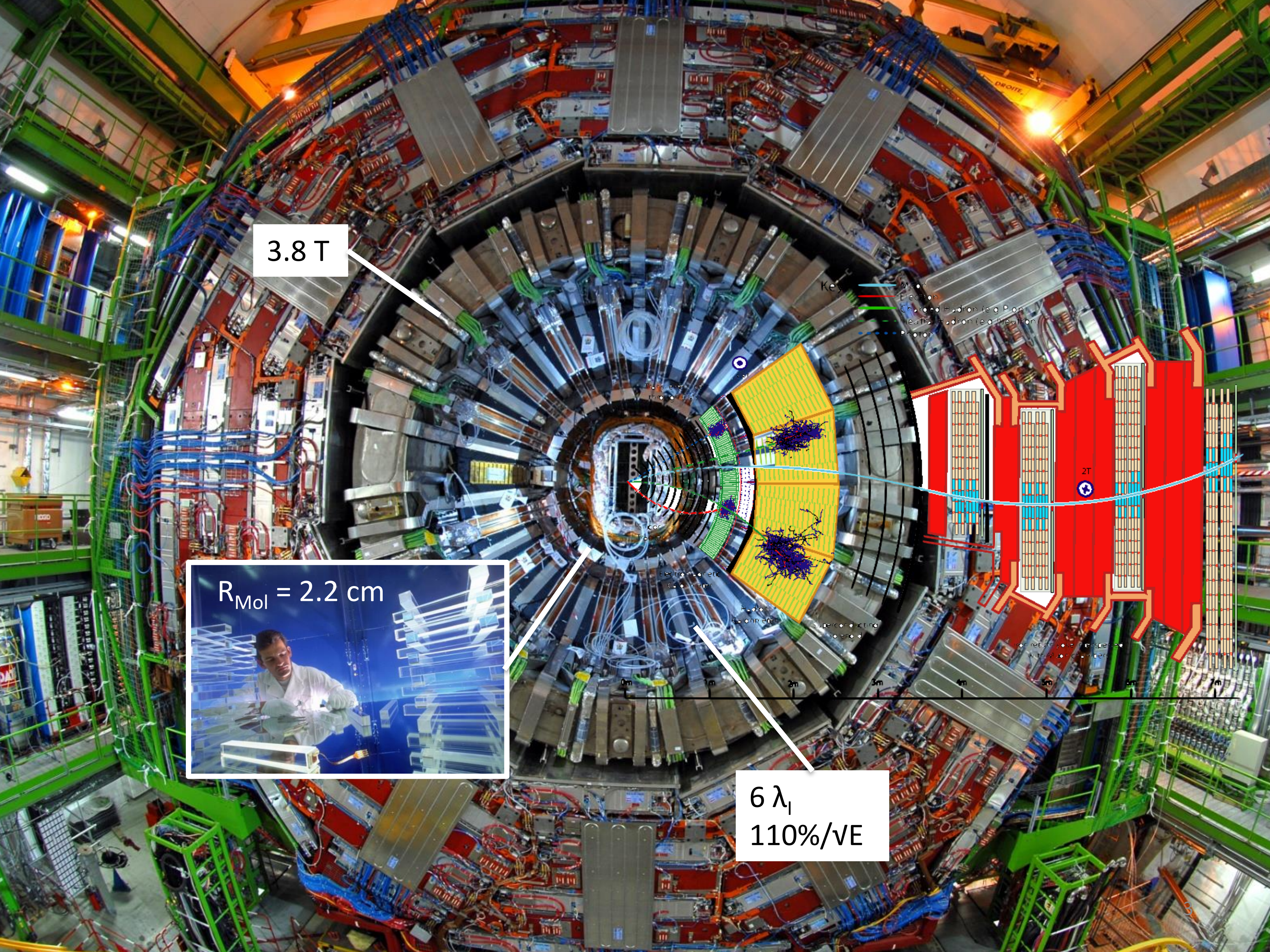
The CMS collaboration

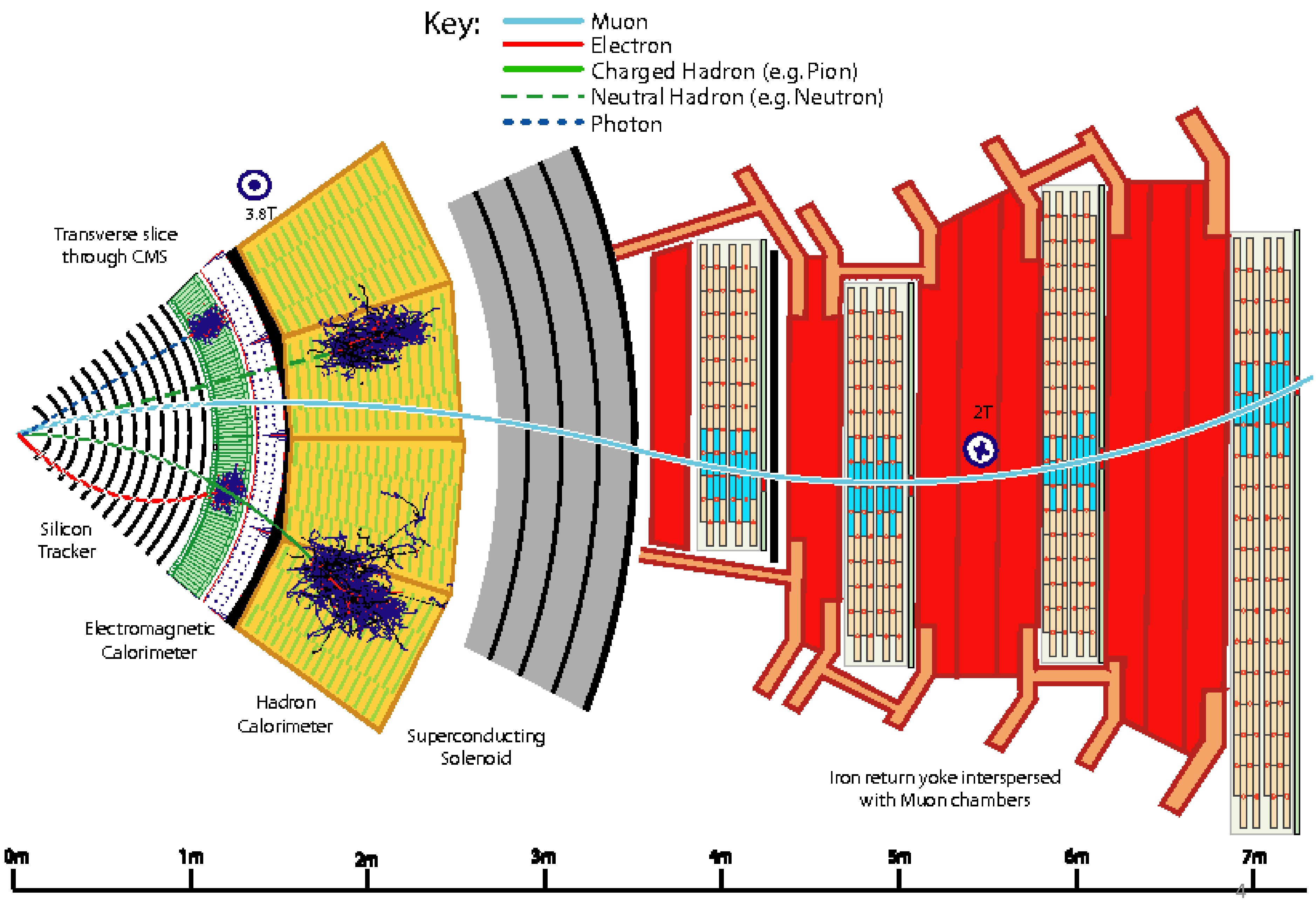
E-mail: cms-publication-committee-chair@cern.ch

Abstract: The CMS apparatus was identified, a few years before the start of the LHC operation at CERN, to feature properties well suited to particle-flow (PF) reconstruction: a highly-segmented tracker, a fine-grained electromagnetic calorimeter, a hermetic hadron calorimeter, a strong magnetic field, and an excellent muon spectrometer. A fully-fledged PF reconstruction algorithm tuned to the CMS detector was therefore developed and has been consistently used in physics analyses for the first time at a hadron collider. For each collision, the comprehensive list of final-state particles identified and reconstructed by the algorithm provides a global event description that leads to unprecedented CMS performance for jet and hadronic τ decay reconstruction, missing transverse momentum determination, and electron and muon identification. This approach also allows particles from pileup interactions to be identified and enables efficient pileup mitigation methods. The data collected by CMS at a centre-of-mass energy of 8 TeV show excellent agreement with the simulation and confirm the superior PF performance at least up to an average of 20 pileup interactions.

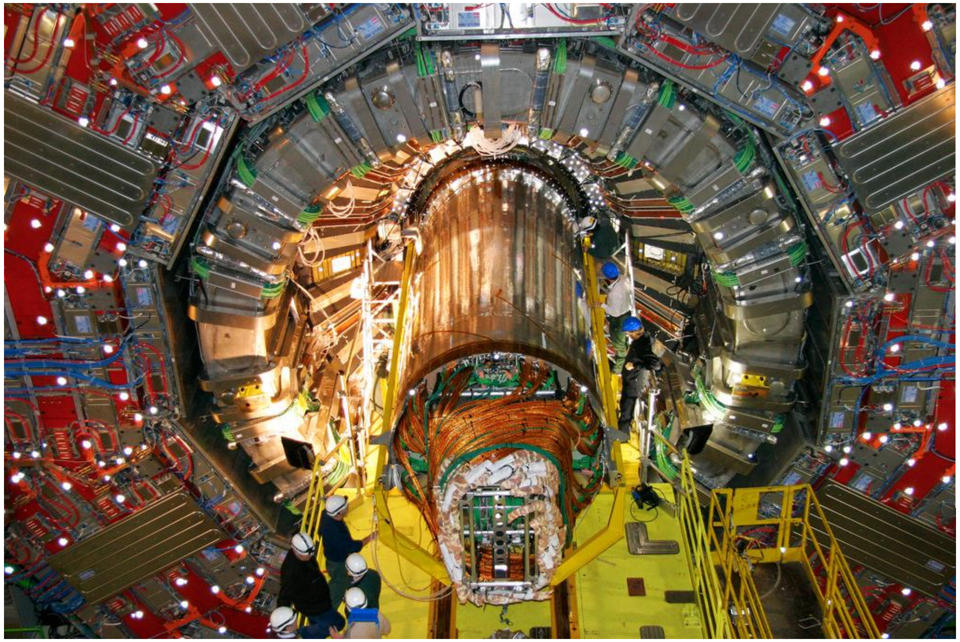
Keywords: Large detector systems for particle and astroparticle physics; Particle identification methods

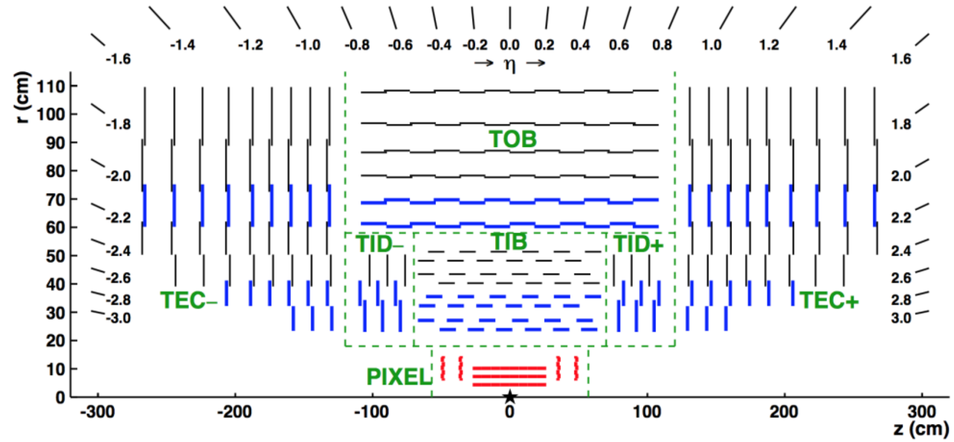
ArXiv ePrint: 1706.04965

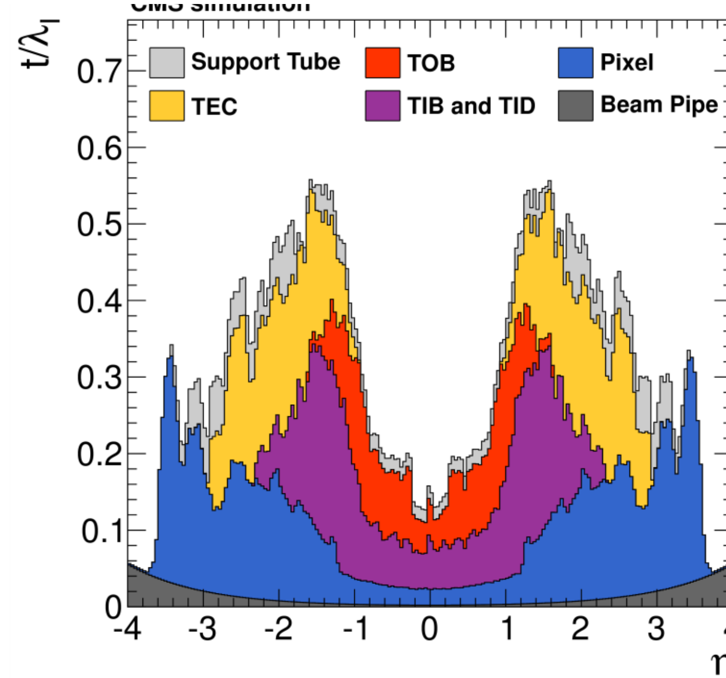

3.8 T
R_Mol = 2.2 cm
6 λ_I
110%/√E


Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon
3.8T
Transverse slice through CMS
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
2T
Iron return yoke interspersed with Muon chambers
0m
1m
2m
3m
4m
5m
6m
7m

The CMS Tracker: Big! (and Thick)



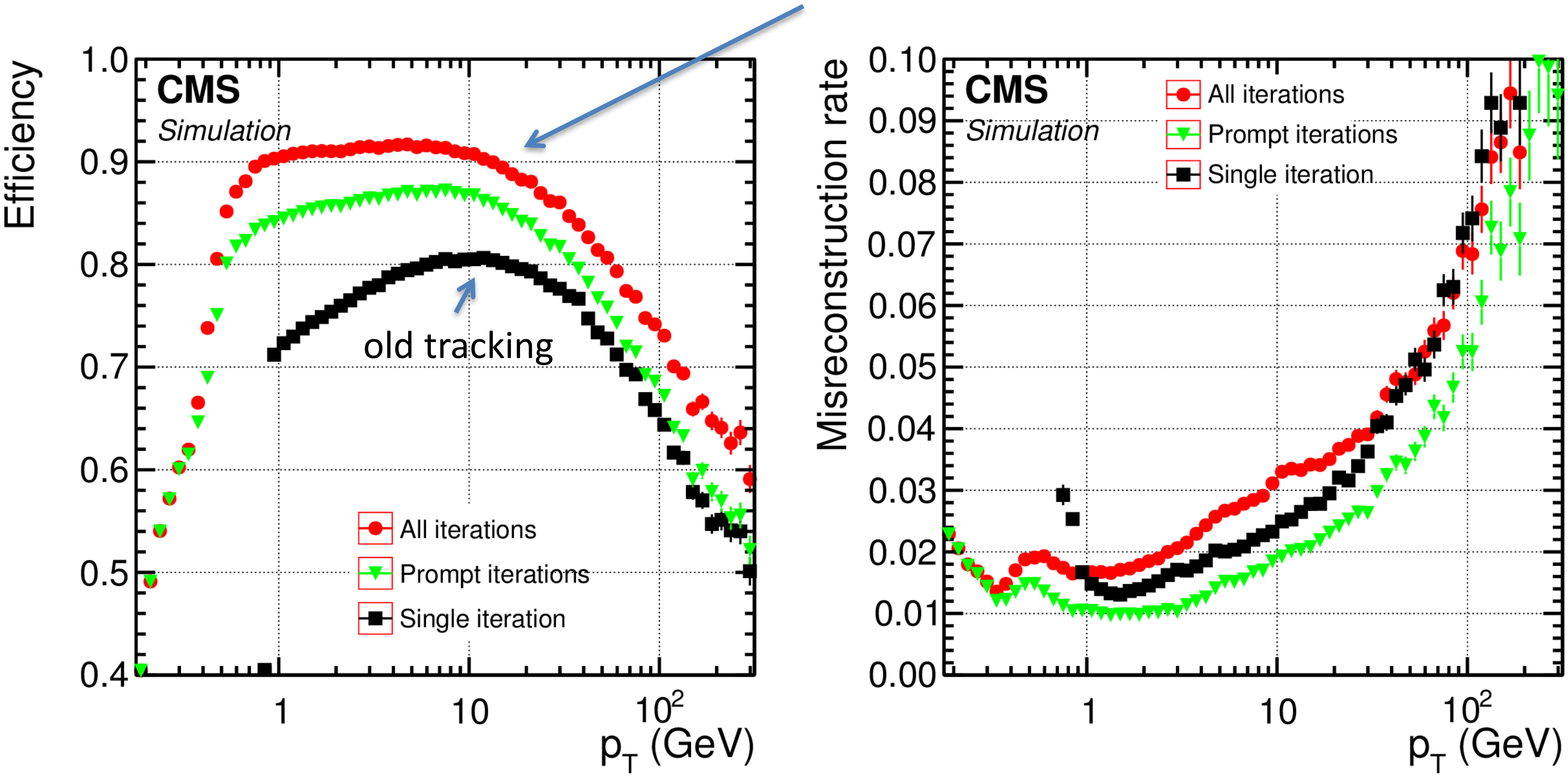




Hadrons: nuclear interactions

e/γ: bremsstrahlung, conversions

Iterative tracking



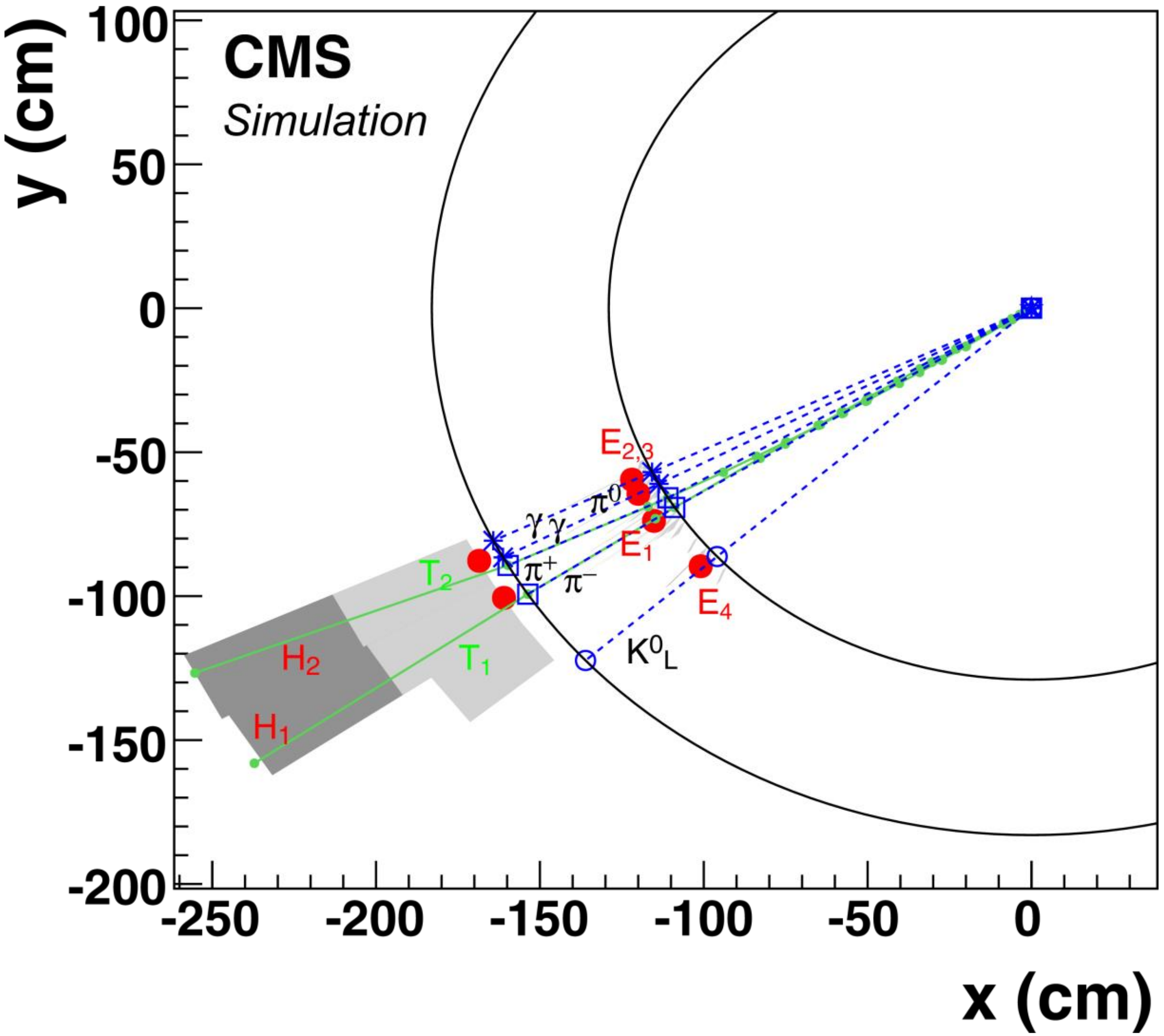
10 iterations:

1- Reconstruct easy tracks

2- Remove their hits

3- Reconstruct more difficult tracks

A Simple Jet

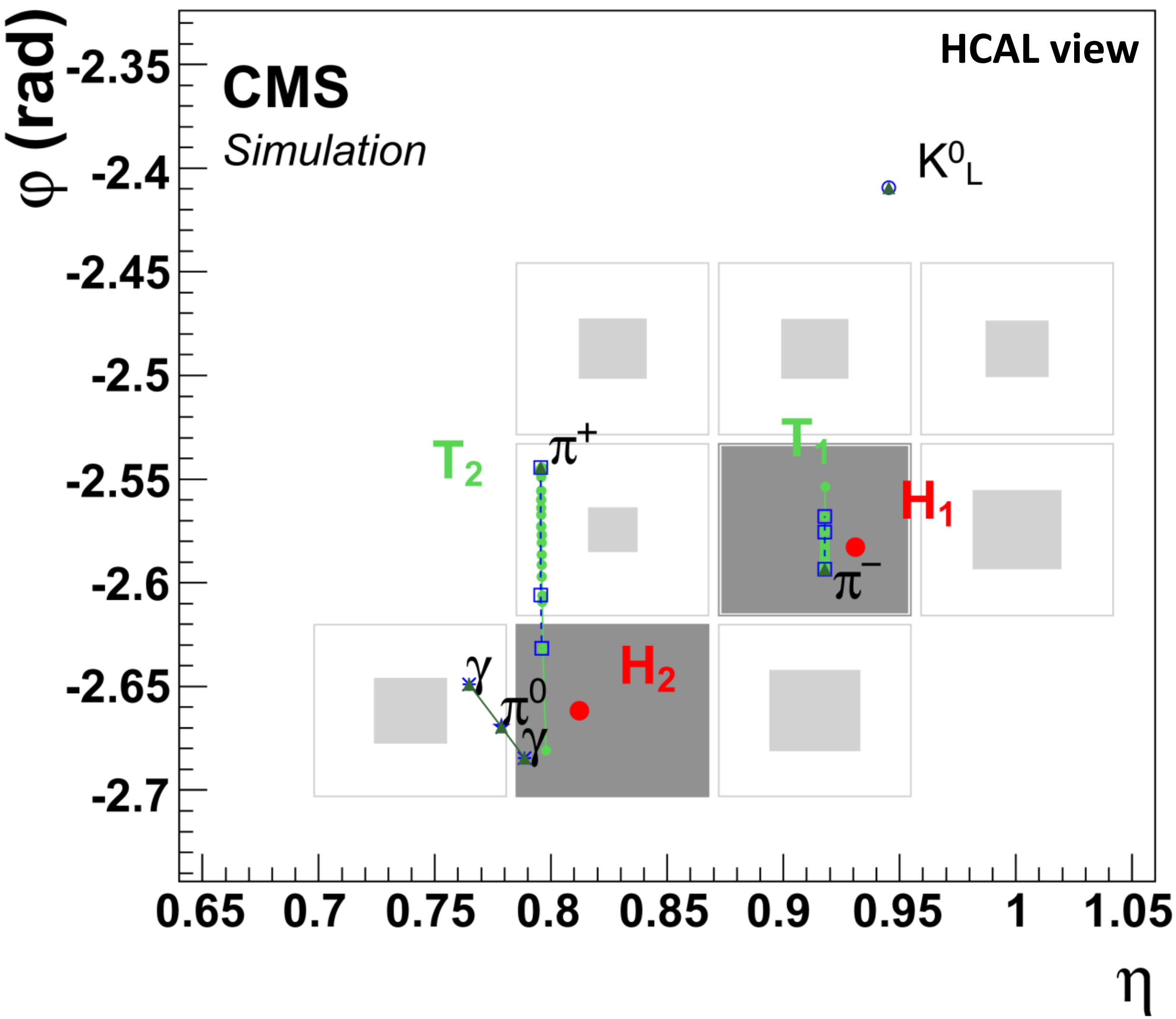


pT = 65 GeV
only 5 particles

Illustrate:

- calo clustering
- particle flow

Calorimeter Clustering

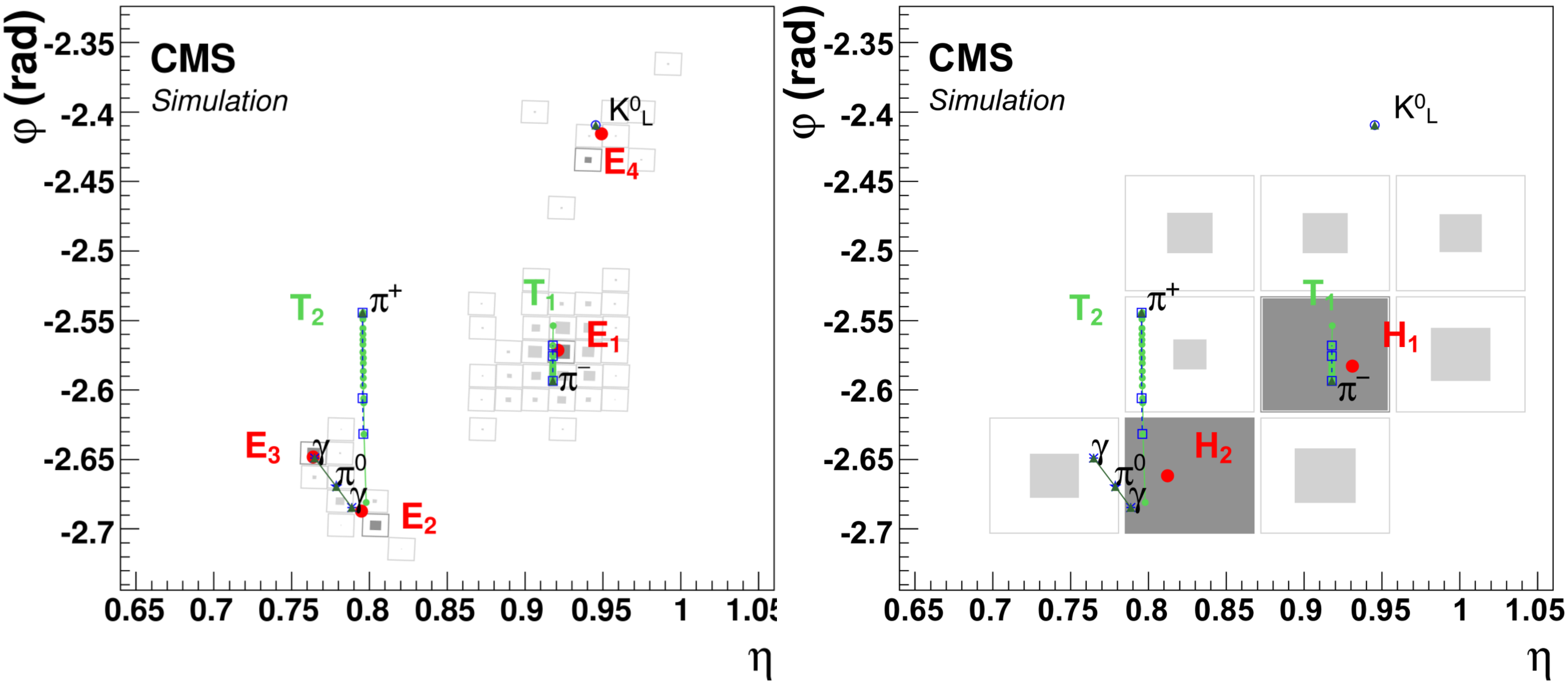


Seeds
- local energy maxima

Connected cells

Share energy
- iteratively
- assuming Gaussian shower profile

Particle Flow

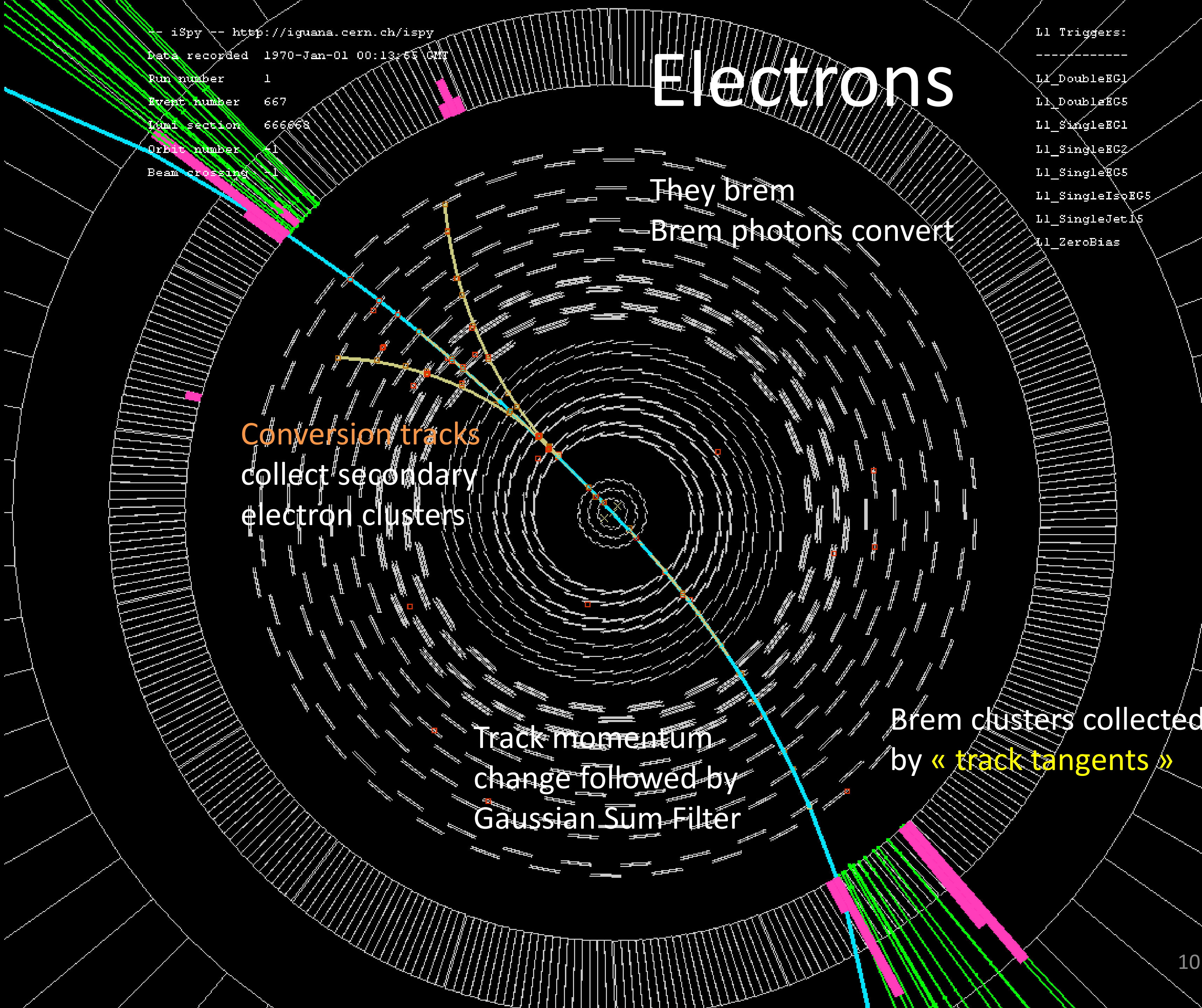


Clusters not linked to a track → photon (ECAL) or neutral hadron (HCAL)

Tracks → charged hadrons

- compatible energy in calos: energy from a fit of track and cluster measurements
- excess: additional neutrals
- deficit: muon, fake track

Electrons



They brem
Brem photons convert

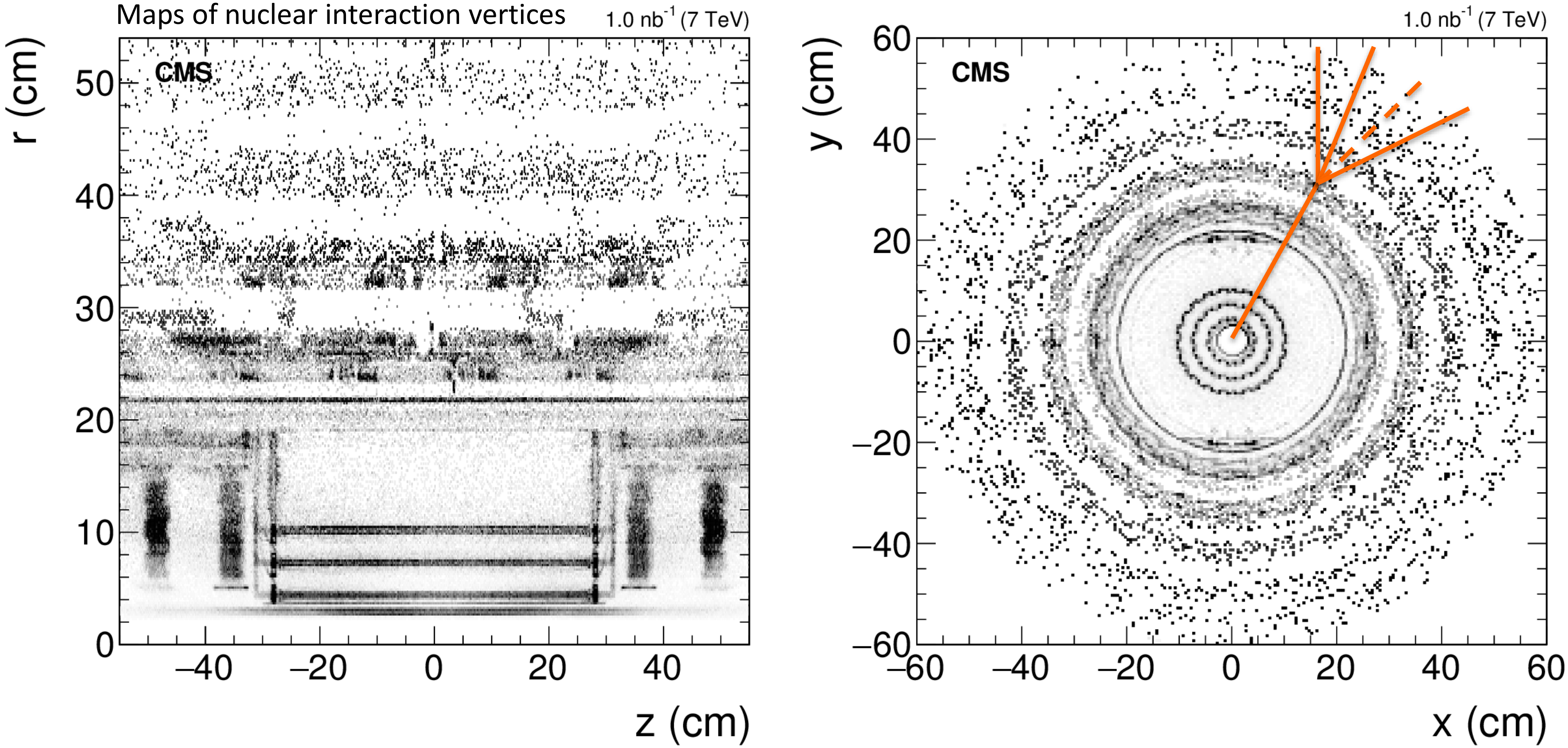
Conversion tracks collect secondary electron clusters

Track momentum change followed by Gaussian Sum Filter

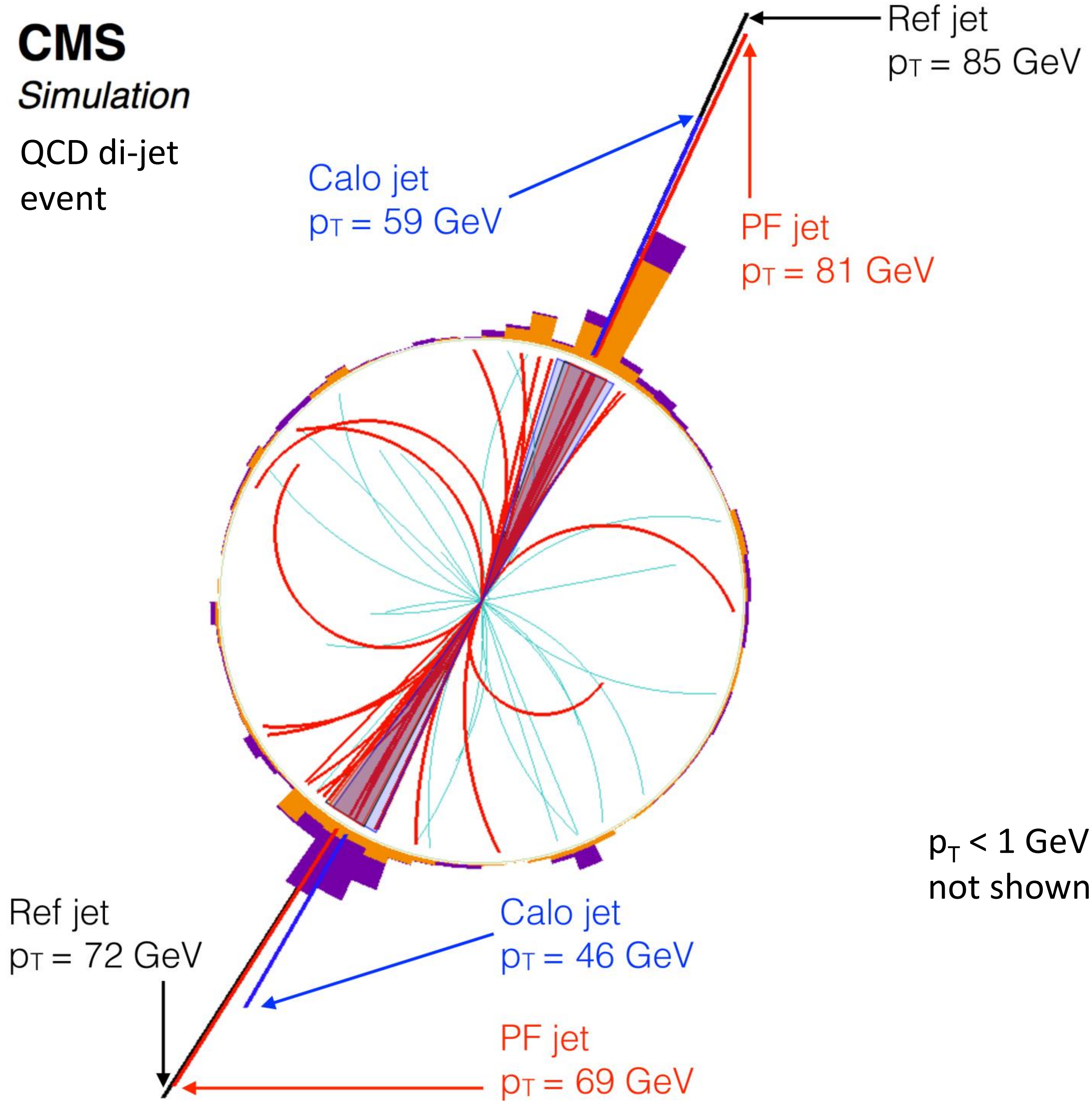
Brem clusters collected by « track tangents »

Nuclear Interactions

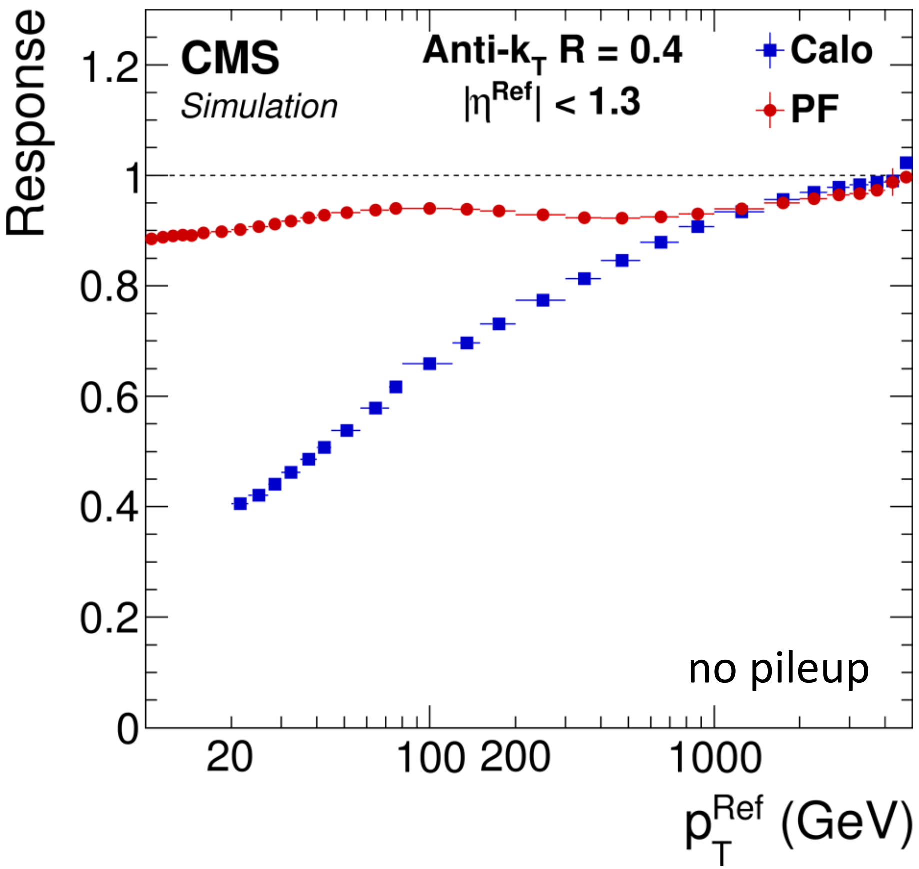
Maps of nuclear interaction vertices

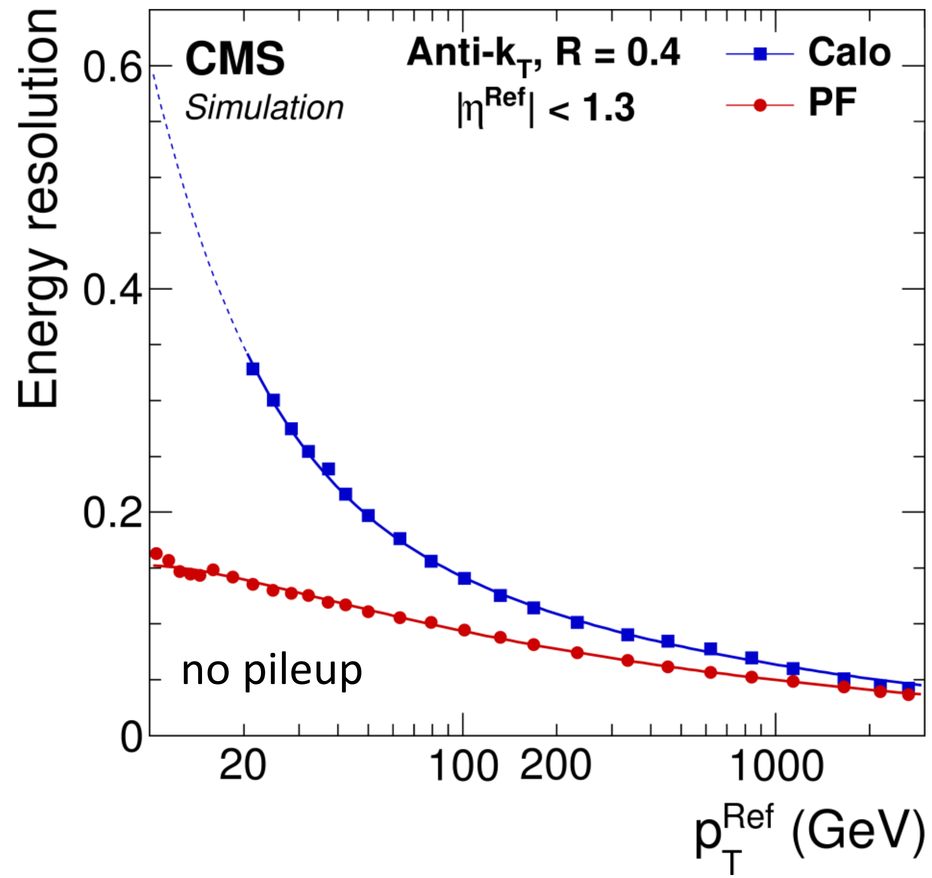


- Secondary vertices reconstructed from displaced tracks (+ 0 or 1 incoming track)
- Secondary charged particles reconstructed by PF → single charged hadron
- Secondary neutrals reconstructed by PF as usual
- Incoming track if any is discarded


CMS
Simulation
QCD di-jet
event
Ref jet
p_T = 85 GeV
Calo jet
p_T = 59 GeV
PF jet
p_T = 81 GeV
Ref jet
p_T = 72 GeV
Calo jet
p_T = 46 GeV
PF jet
p_T = 69 GeV
p_T < 1 GeV
not shown

Jet Response and Resolution

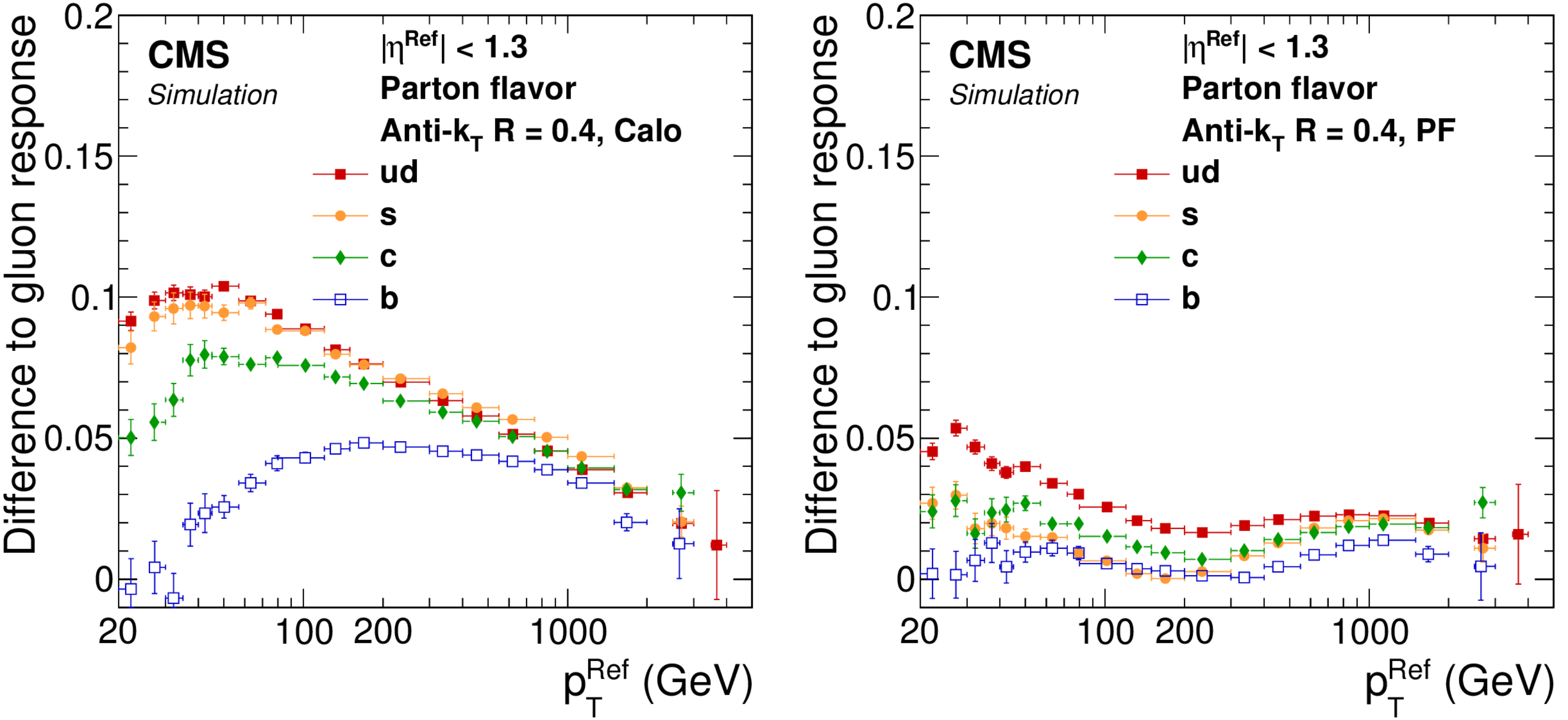




Response closer to unity
~linear

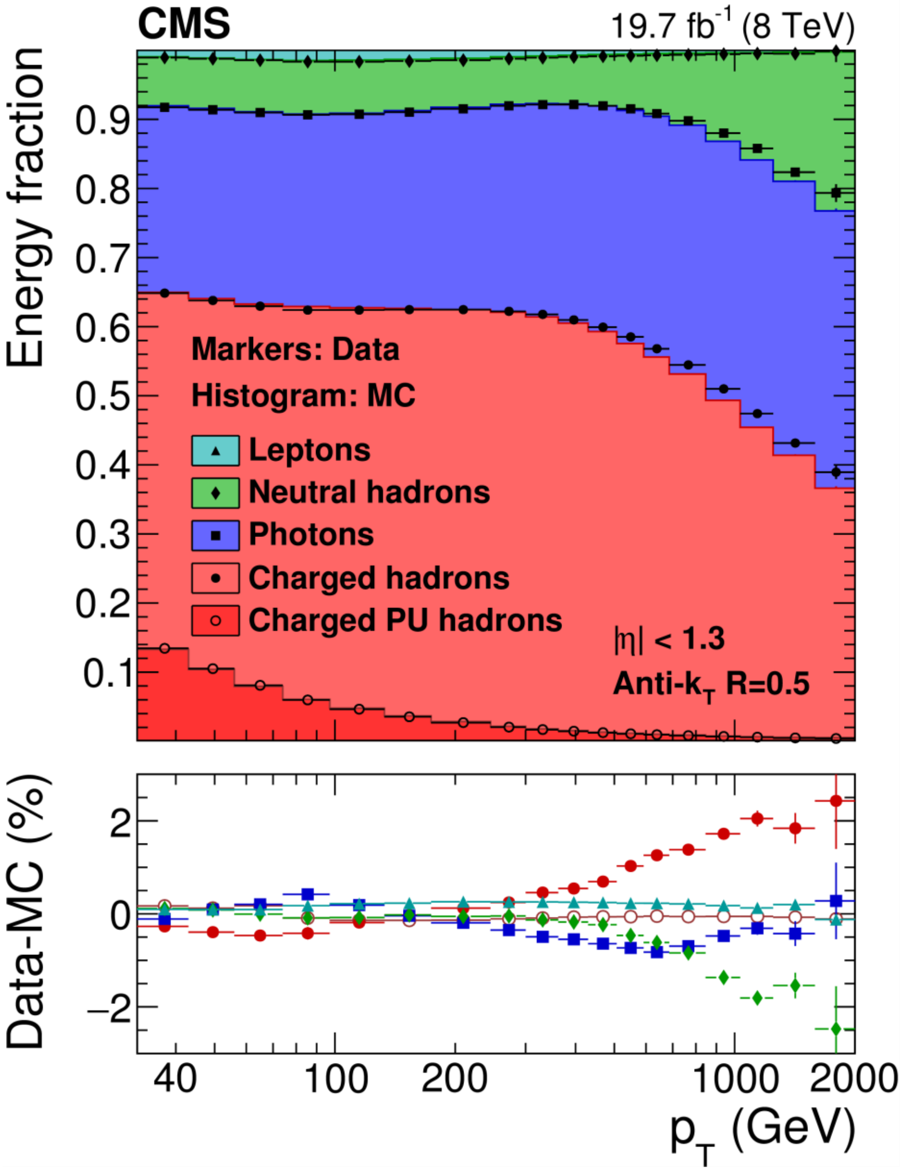
Resolution always better than for calo jets:
- PF works when particles are close
- charged hadron energy from fit of track and cluster measurements

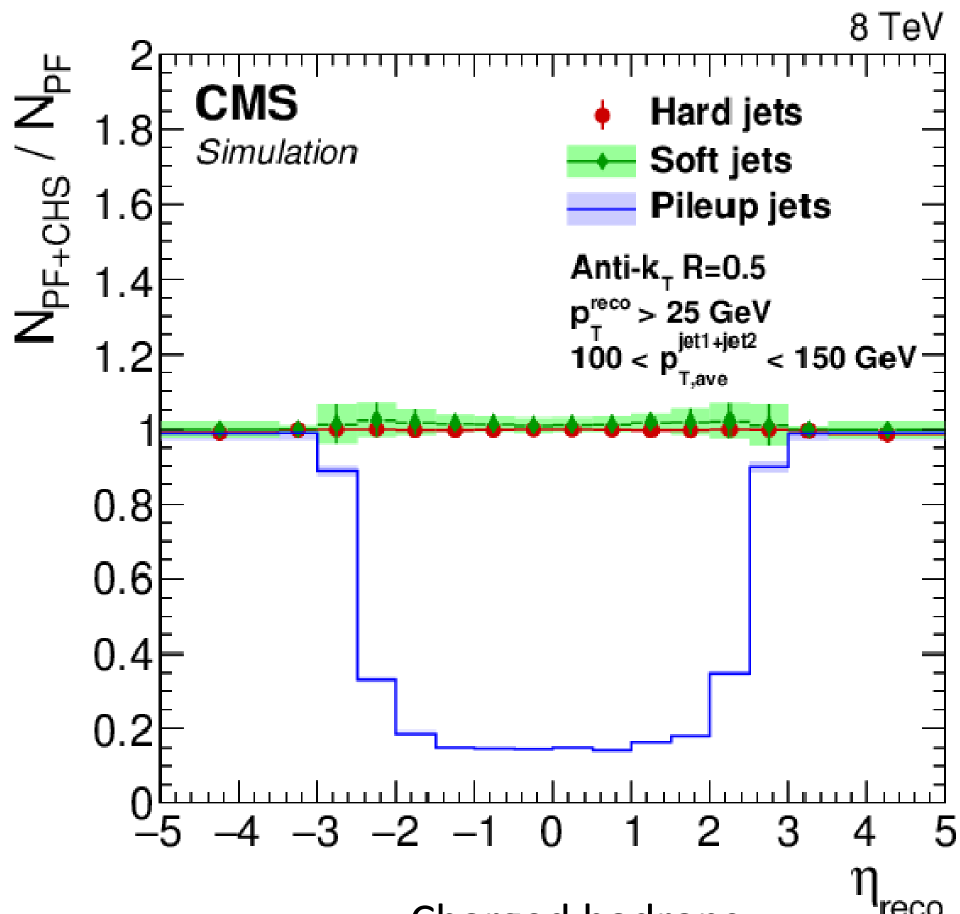
Response for different flavours



Sensitivity of the response to the parton flavour reduced
→ Jet energy scale systematic uncertainty reduced

Jet Composition, Pileup





Charged hadrons
from pileup vertices
removed before jet clustering

Particle Flow and Jets @HLT

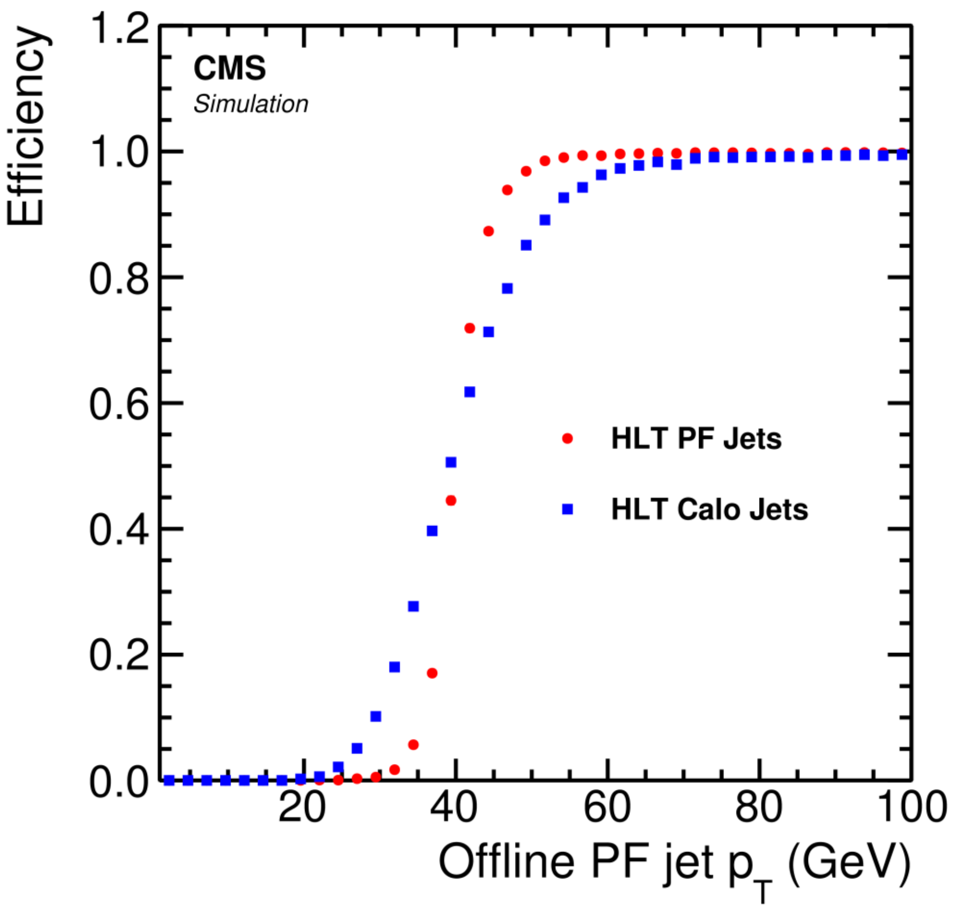
In the High-Level trigger, timing is crucial! (~140 ms / evt)

	Tracking	PF
Offline	600 ms	70 ms
HLT	60 ms	30 ms

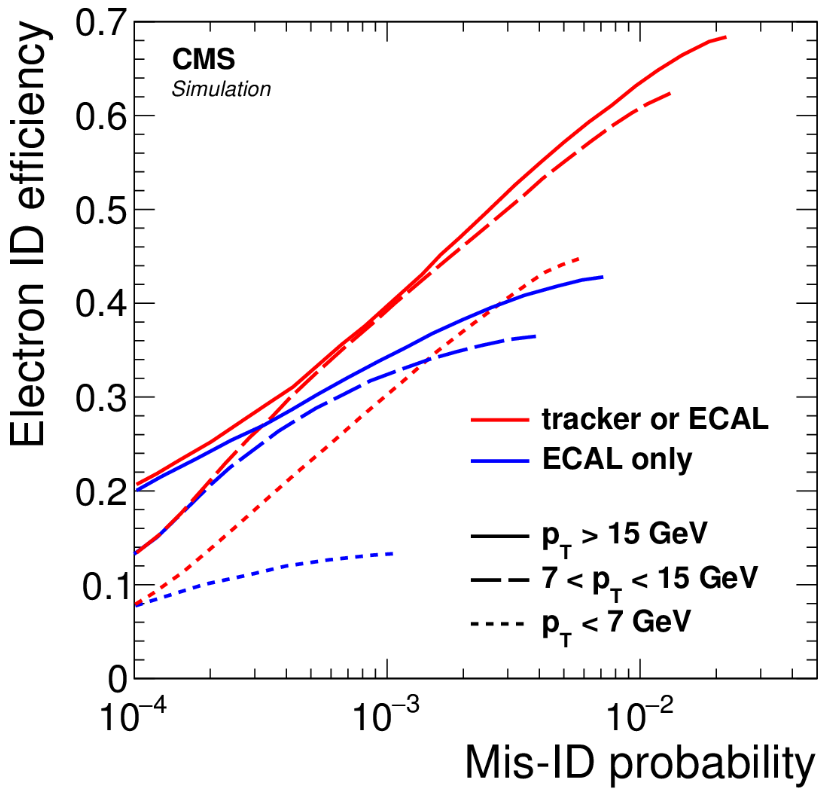
no pileup, assuming tracking and PF are performed for every event

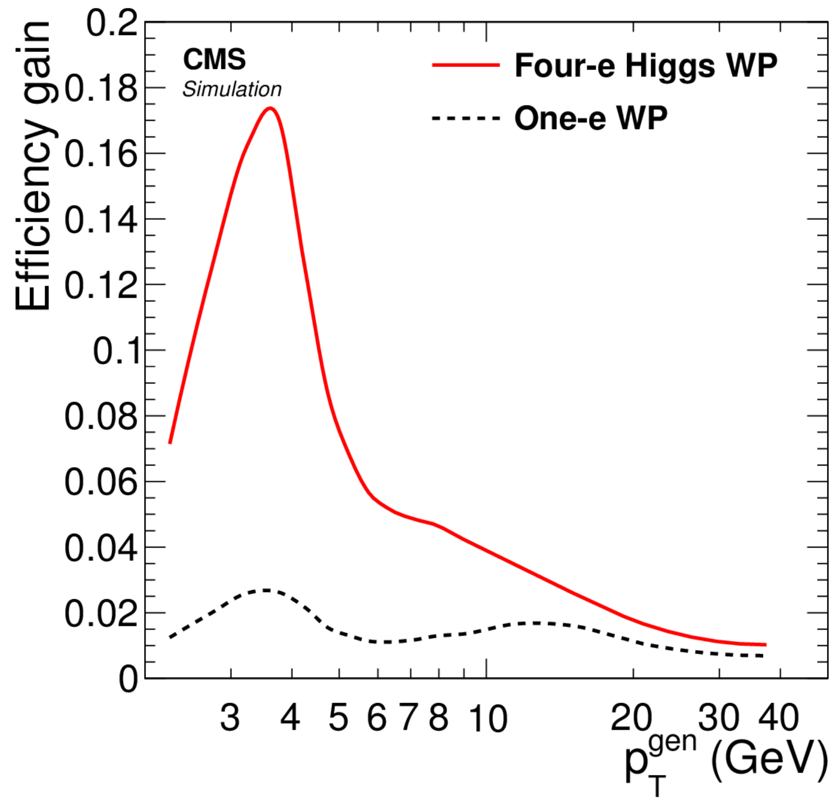
Fraction of total time at HLT @45 pileup:

	Tracking	PF
HLT	< 20%	< 10 %



Electrons



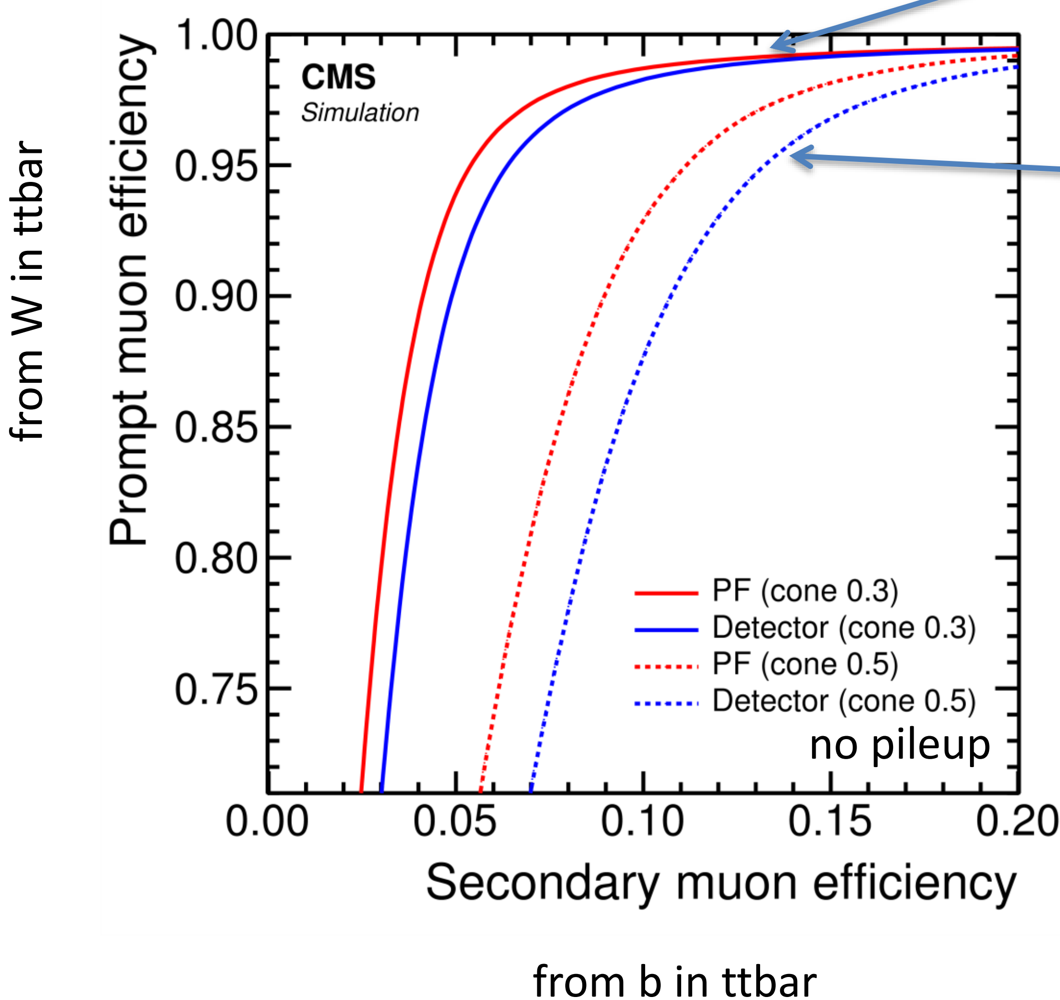


Electrons in jets
b tagging with electrons possible

Prompt electrons
Big efficiency gains at low pT:
+7% more H→ZZ→4e events

Isolation

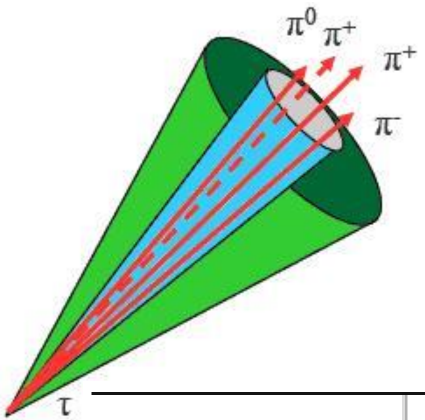
$$I_{\mathrm{PF}} = \frac{1}{p_{\mathrm{T}}}\left(\sum_{\mathrm{h}^{\pm}} p_{\mathrm{T}}^{\mathrm{h}^{\pm}} + \sum_{\gamma} p_{\mathrm{T}}^{\gamma} + \sum_{\mathrm{h}^0} p_{\mathrm{T}}^{\mathrm{h}^0}\right)$$



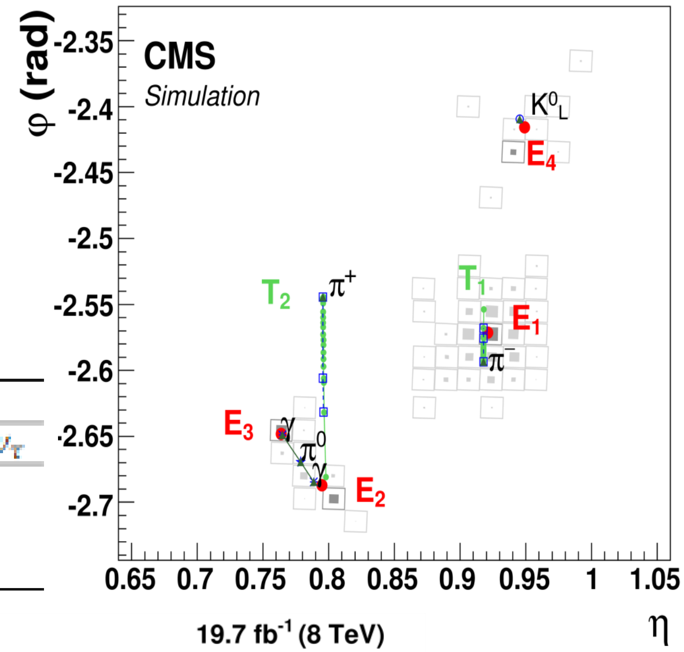
$$I_{\mathrm{det}} = \frac{1}{p_{\mathrm{T}}}\left(\sum_{\mathrm{tracks}} p_{\mathrm{T}}^{\mathrm{track}} + \sum_{\mathrm{ECAL}} E_{\mathrm{T}}^{\mathrm{ECAL}} + \sum_{\mathrm{HCAL}} E_{\mathrm{T}}^{\mathrm{HCAL}}\right)$$

Detector-based isolation:
- Charged-hadron energy double counted
- Cannot easily remove pileup calorimeter deposits

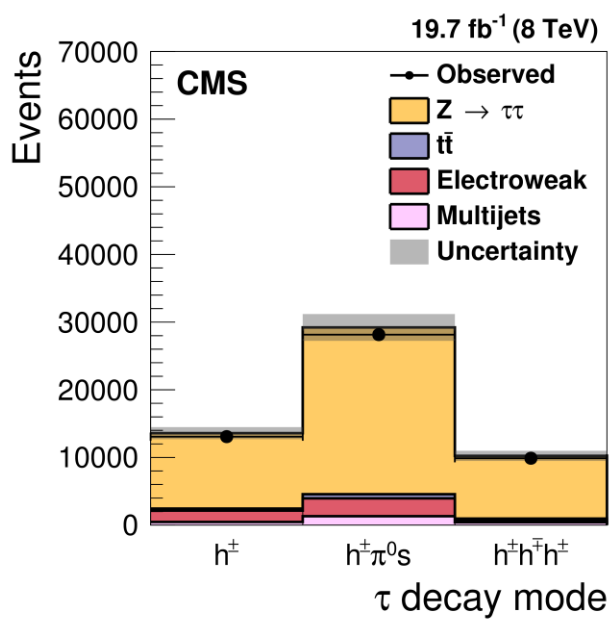
Tau Reconstruction & Identification

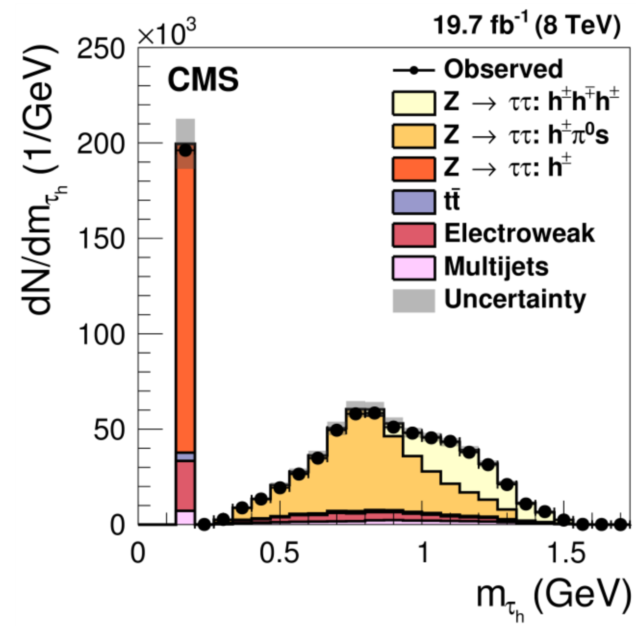


1 or 3 $\pi^{\pm}$
photons from π^0

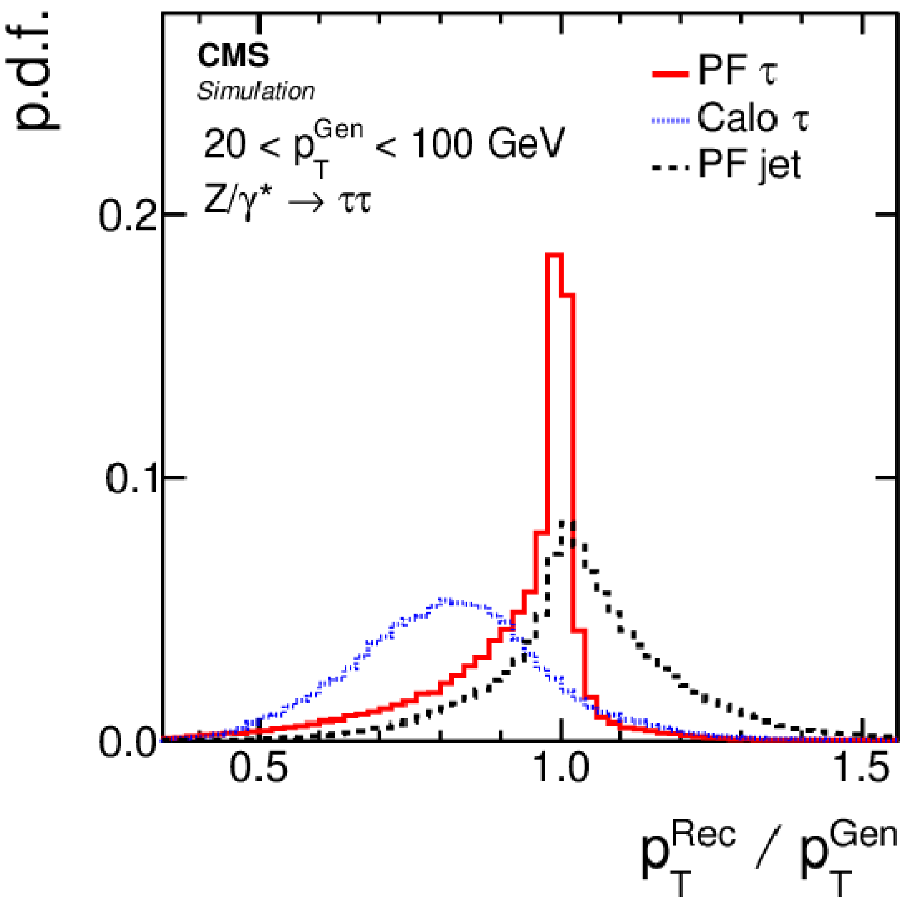


Reconstructed	Generated		
	$\tau^- \to h^- \nu_\tau$	$\tau^- \to h^- \geq 1\pi^0\,\nu_\tau$	$\tau^- \to h^- h^+ h^- \nu_\tau$
$\tau^- \to h^- \nu_\tau$	0.89	0.16	0.01
$\tau^- \to h^- \geq 1\pi^0\,\nu_\tau$	0.11	0.83	0.02
$\tau^- \to h^- h^+ h^- \nu_\tau$	0.00	0.01	0.97

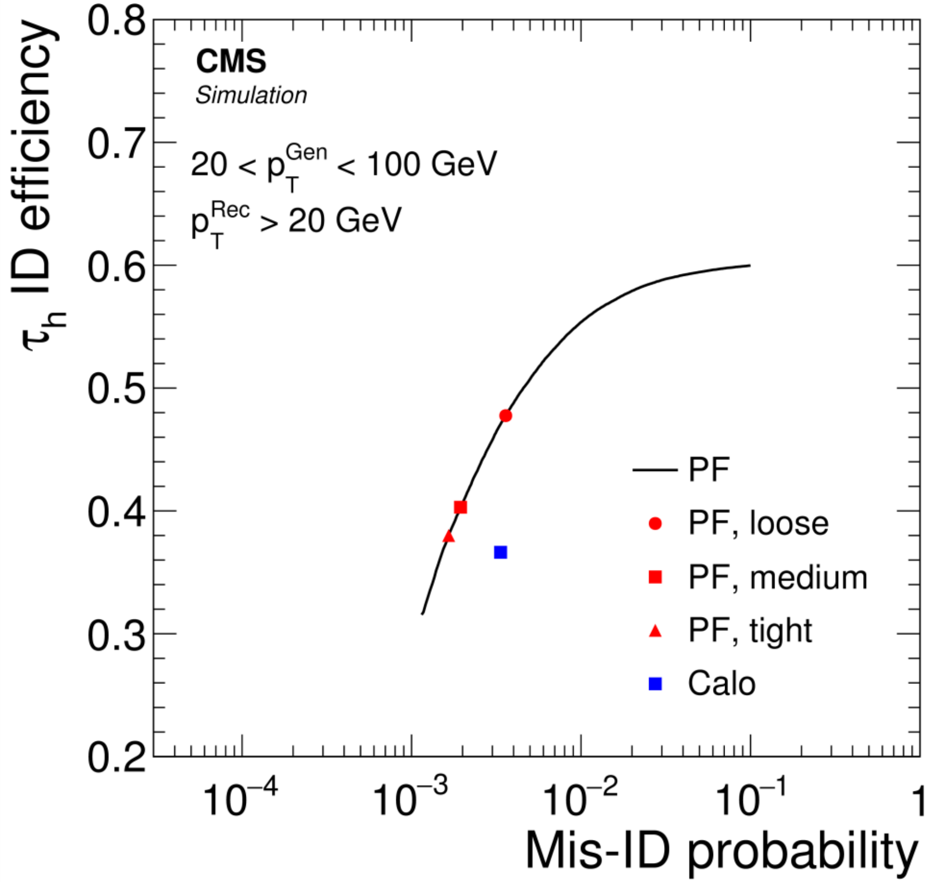




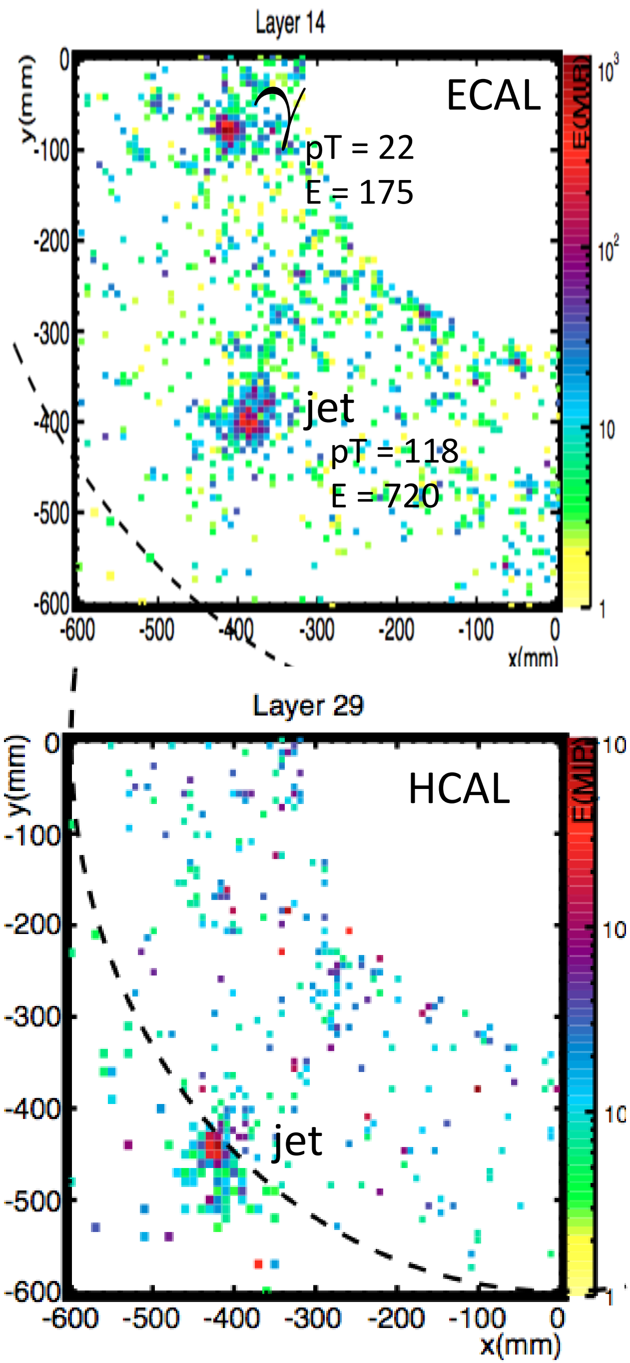
Tau Reconstruction & Identification



p_T resolution



identification performance



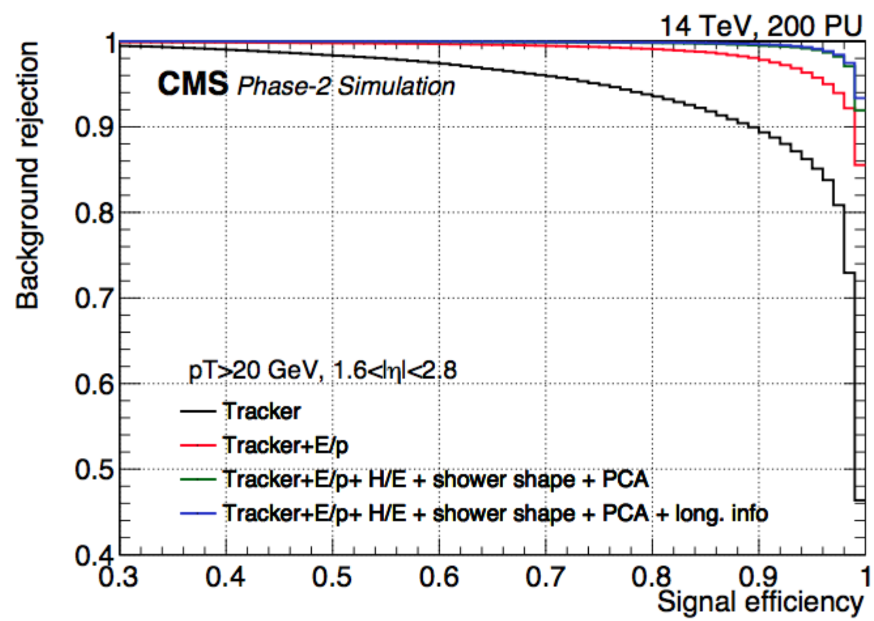
HGCAL

https://cds.cern.ch/record/2293646?ln=fr

200 pileup collisions / beam crossing
Boosted jets
Pandora and Arbor:
- aggregate many unrelated hits

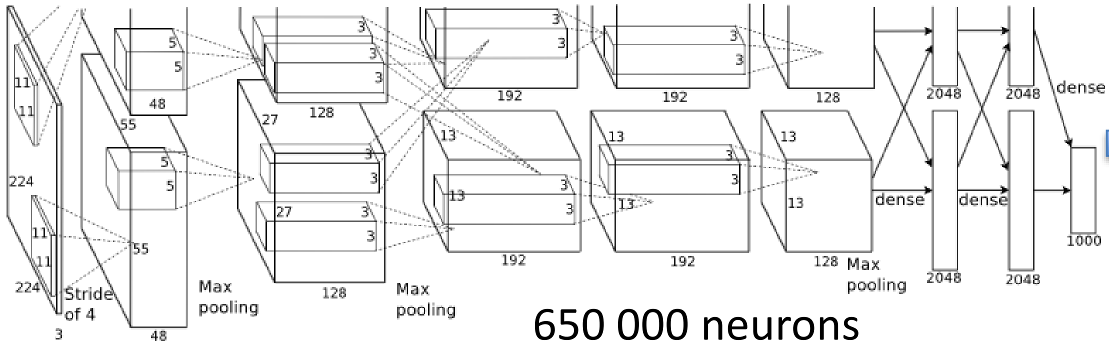
Tried a different approach:
- fast 2D clustering in each layer
- projective association into 3D clusters
- works only for electrons so far

Setting up a full PF algo will take time and effort



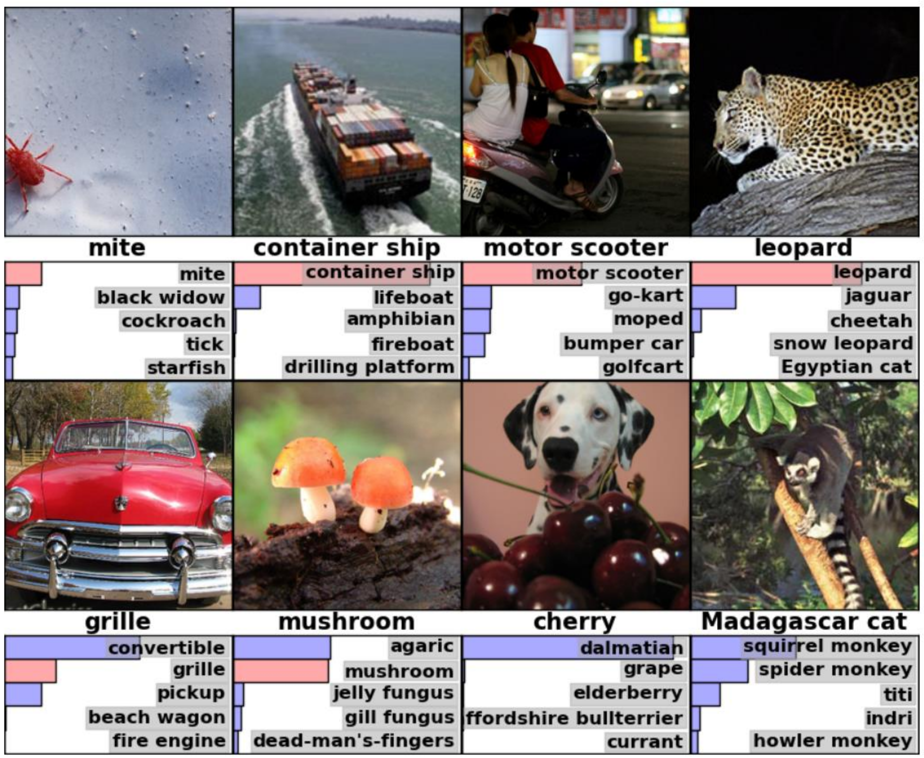
Deep Learning

Input:
raw image
(color levels
for each pixel)



650 000 neurons

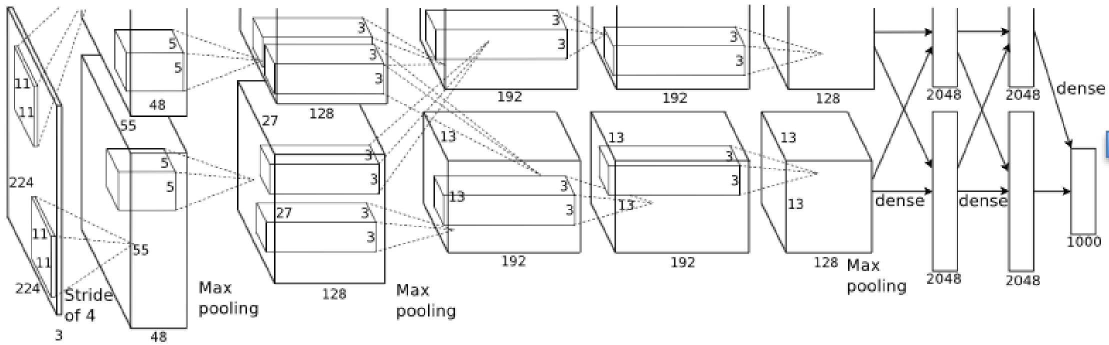
Output:
score for
each
category



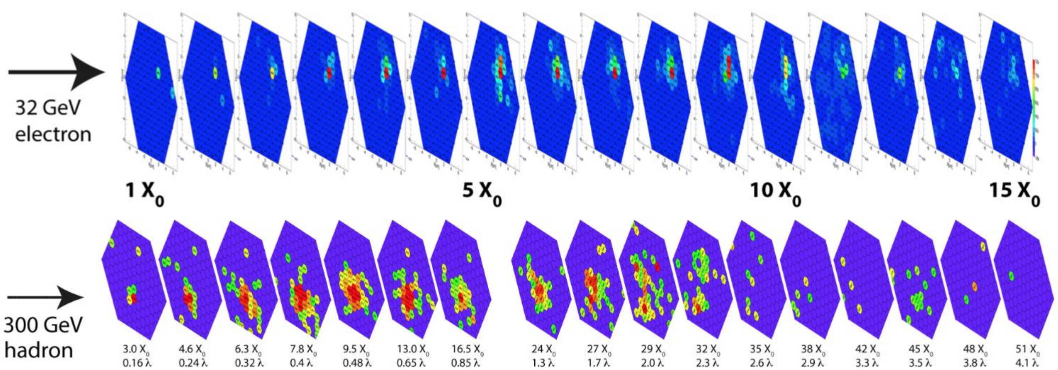
[AlexNet]()

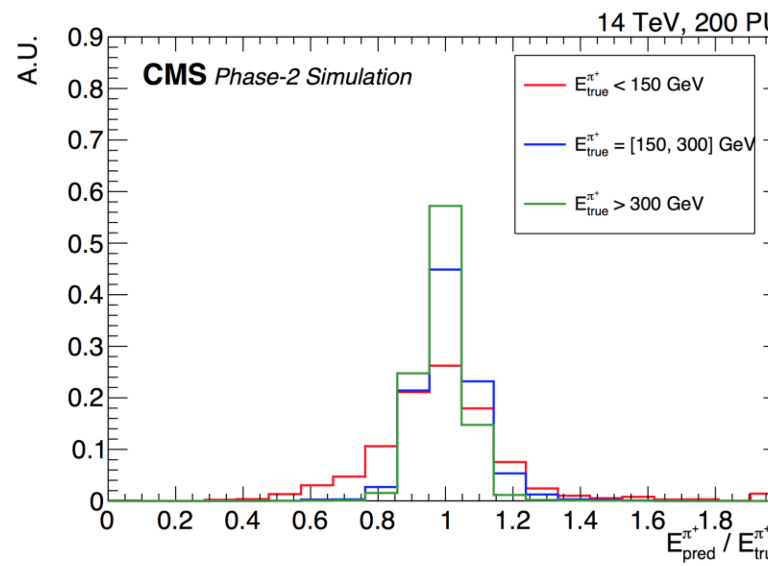
Deep Learning for HGCAL reconstruction

Input:
raw image



Output:
shower
energy

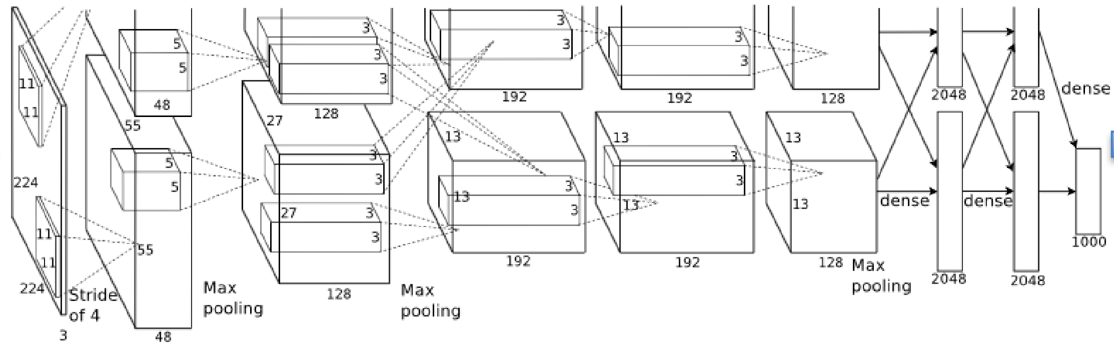




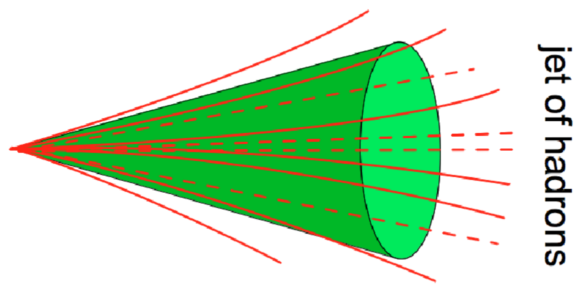
Deep Learning on PF output

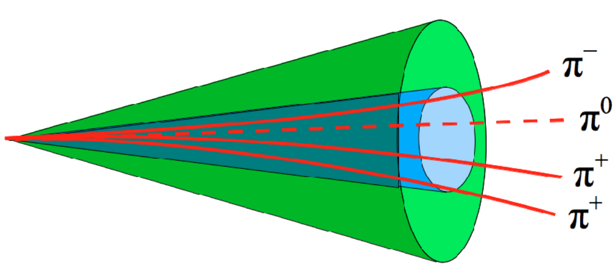
Input:
raw jet
(e.g. 1st 100 particles in jet)

or other collection of particles



Output:
score for each category
or regression value





e.g. :
- τ identification
- b jet tagging
- jet energy correction
- MET

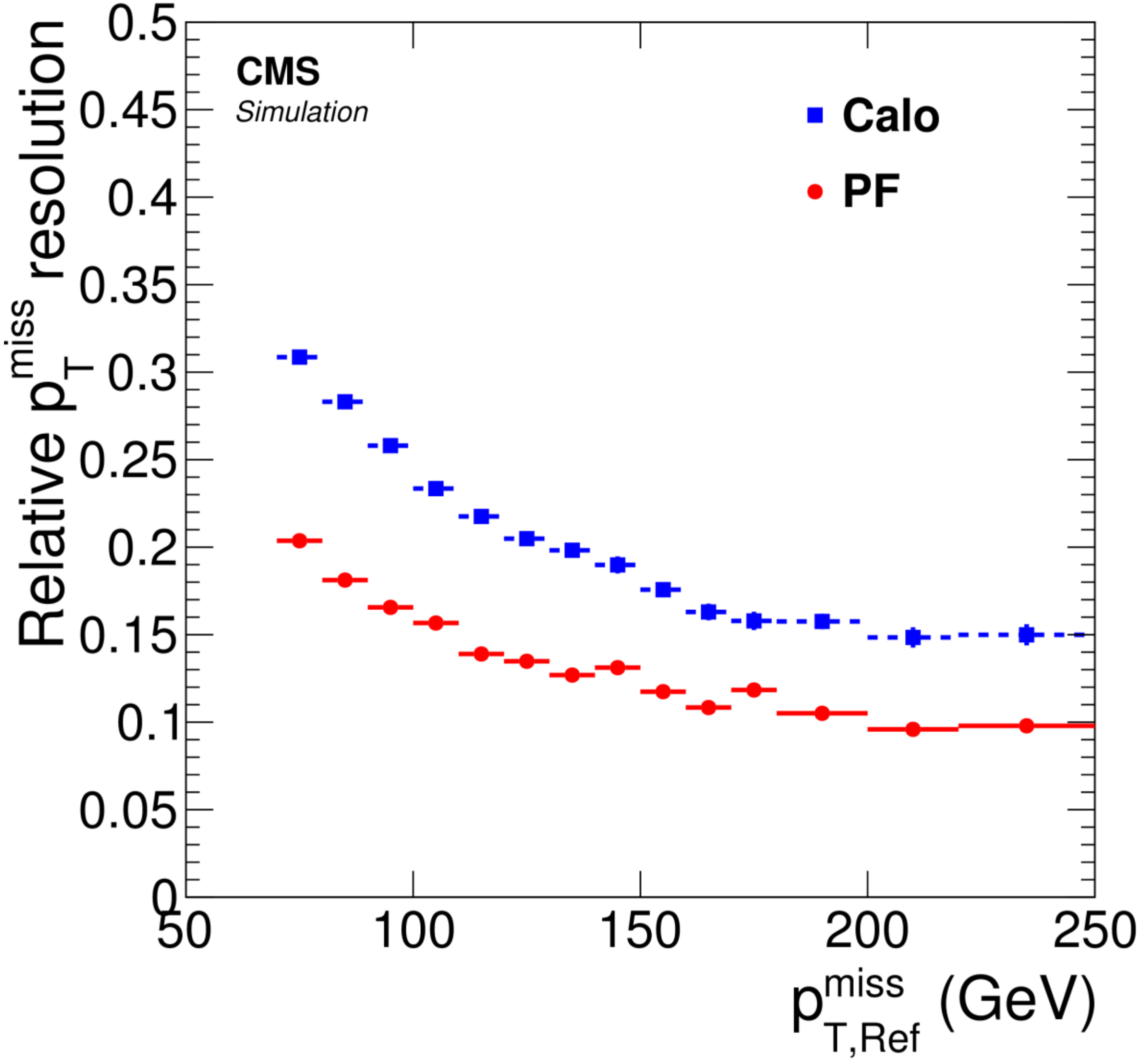
Our advantage:
Monte Carlo Simulation
- large number of events
- we know the truth (NN target)

Conclusion

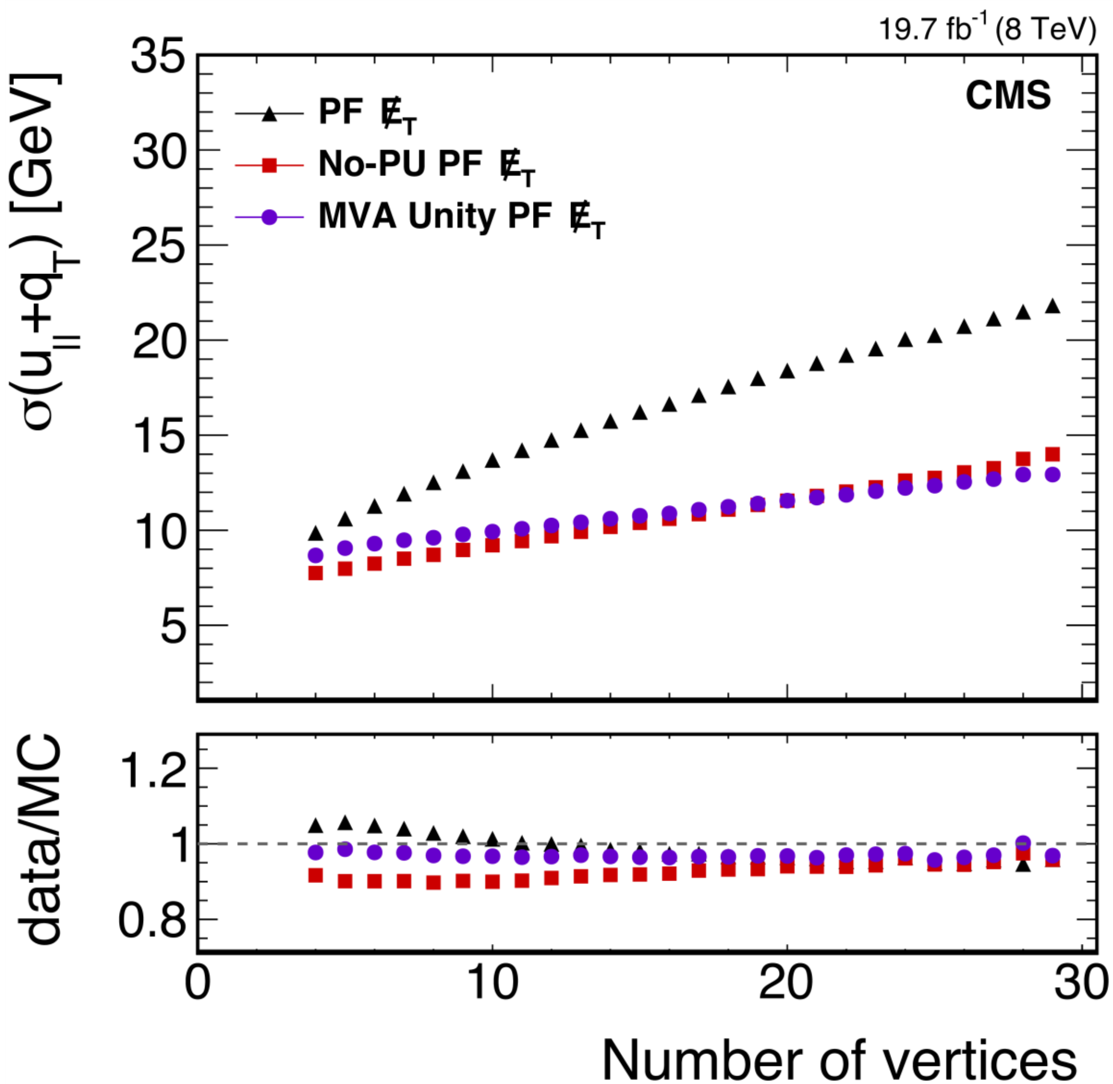
- The foundation of 99.5% of our analyses
  - all physics objects come from PF
- > 50 000 lines of code
  - to deal with the gory « details » :
    - fake tracks, muons, e/γ, secondary interactions, noise, …
- Algorithm unchanged since 2009
  - 0 → 45 pileup collisions
- But HL-LHC data will be a big challenge

Backup

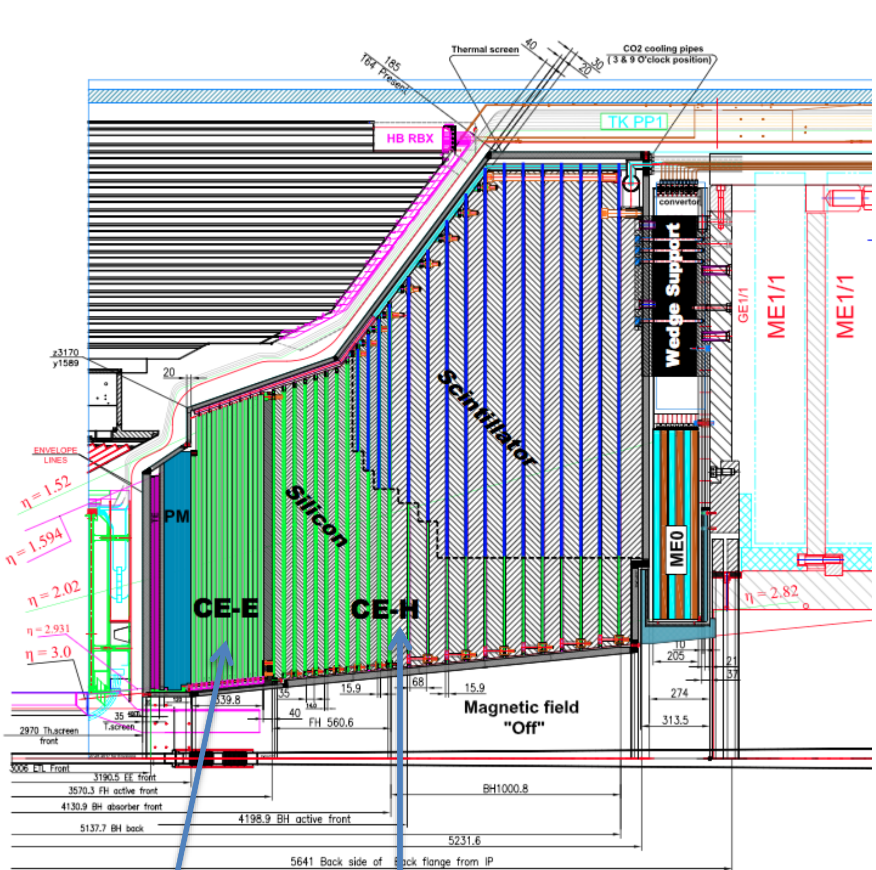
MET


CMS
Simulation
Calo
PF
Relative p_T^{miss} resolution
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
50
100
150
200
250
$p_{T,Ref}^{miss}$ (GeV)

MVA MET

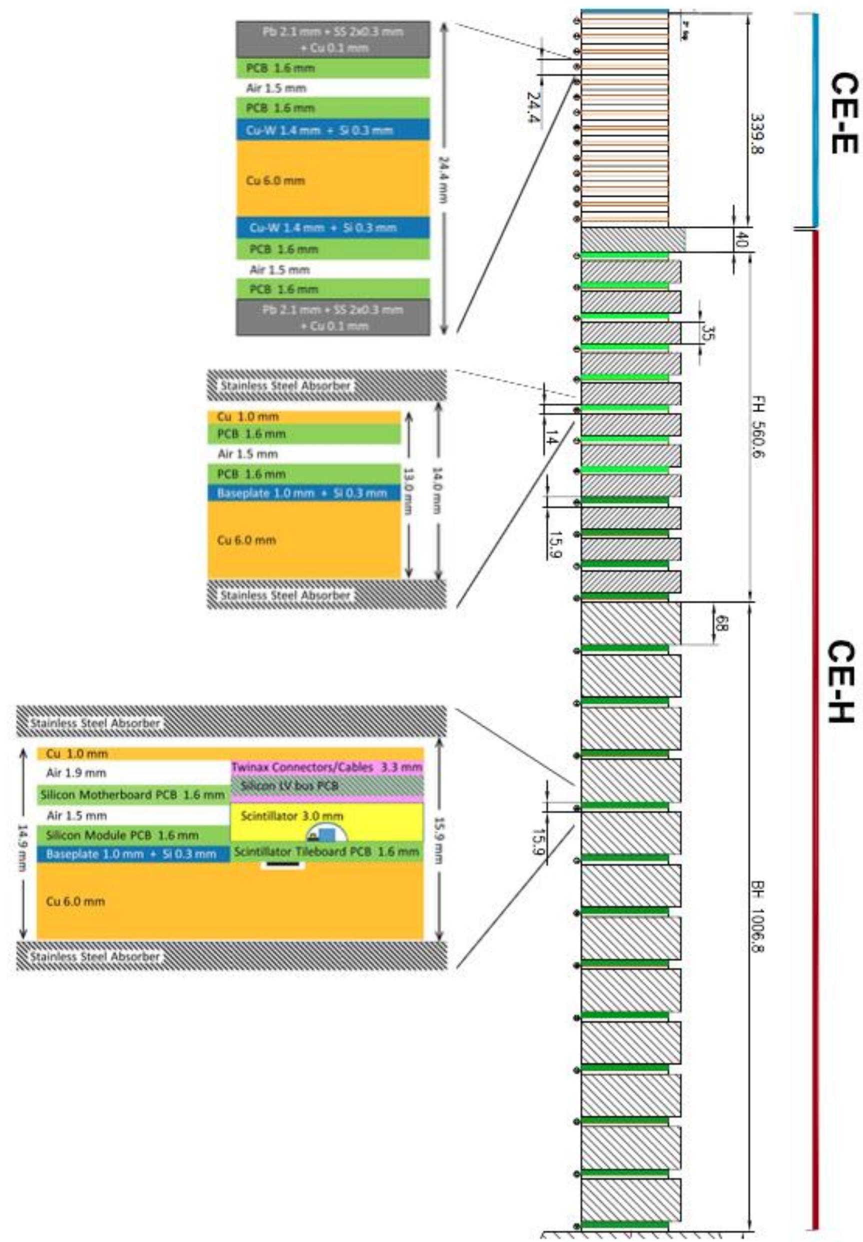

19.7 fb-1 (8 TeV)
CMS
PF $\not\!\!E_T$
No-PU PF $\not\!\!E_T$
MVA Unity PF $\not\!\!E_T$
σ(u|| + qT) [GeV]
data/MC
35
30
25
20
15
10
5
1.2
1
0.8
0
10
20
30
Number of vertices

The CMS HGCAL

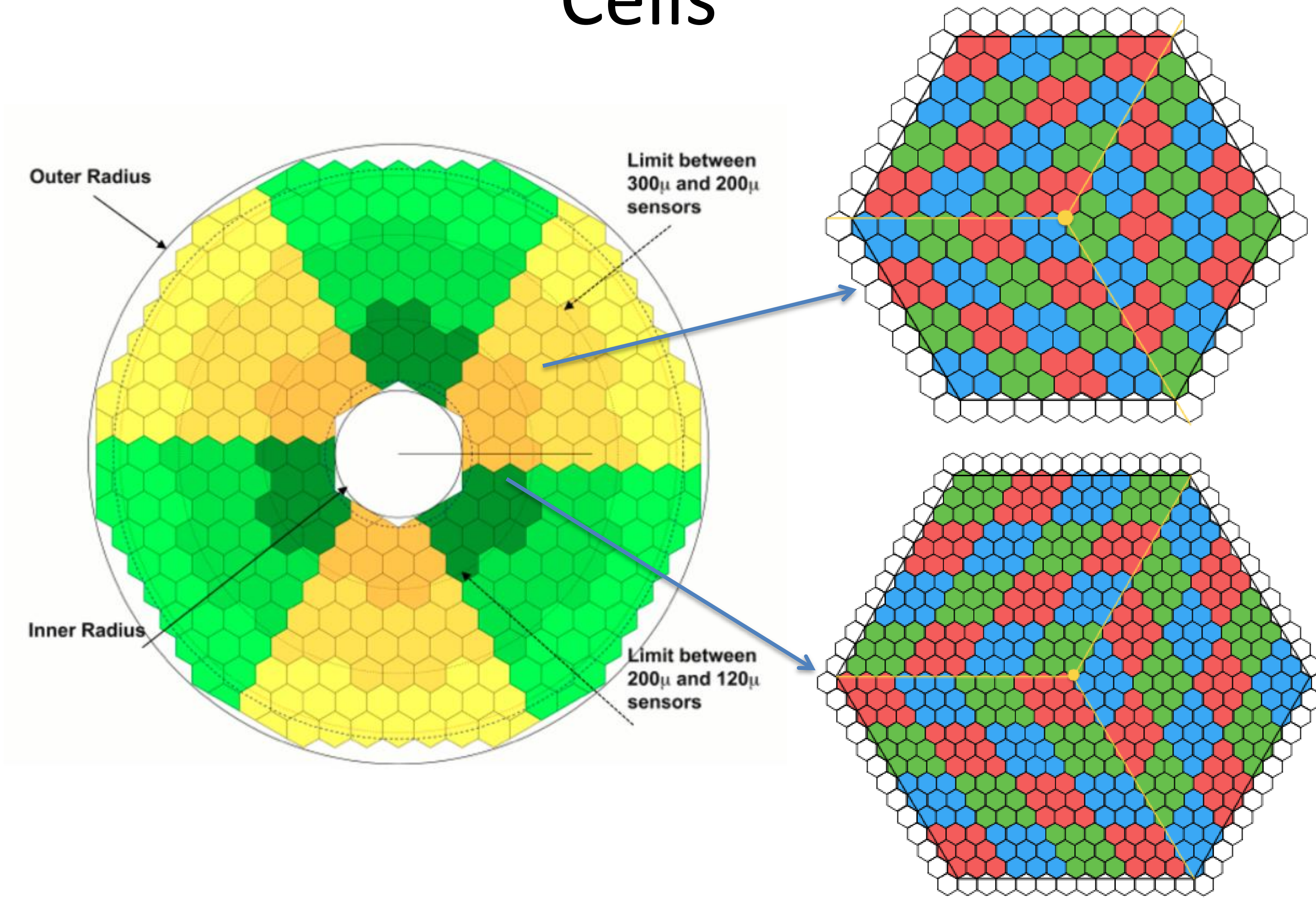


28 layers
26 X_0
1.7 λ_I

9 λ_I



Cells


Outer Radius
Limit between 300μ and 200μ sensors
Inner Radius
Limit between 200μ and 120μ sensors