Challenges of Single Particle Reconstruction in Hadronic Environments

Lindsey Gray (FNAL)

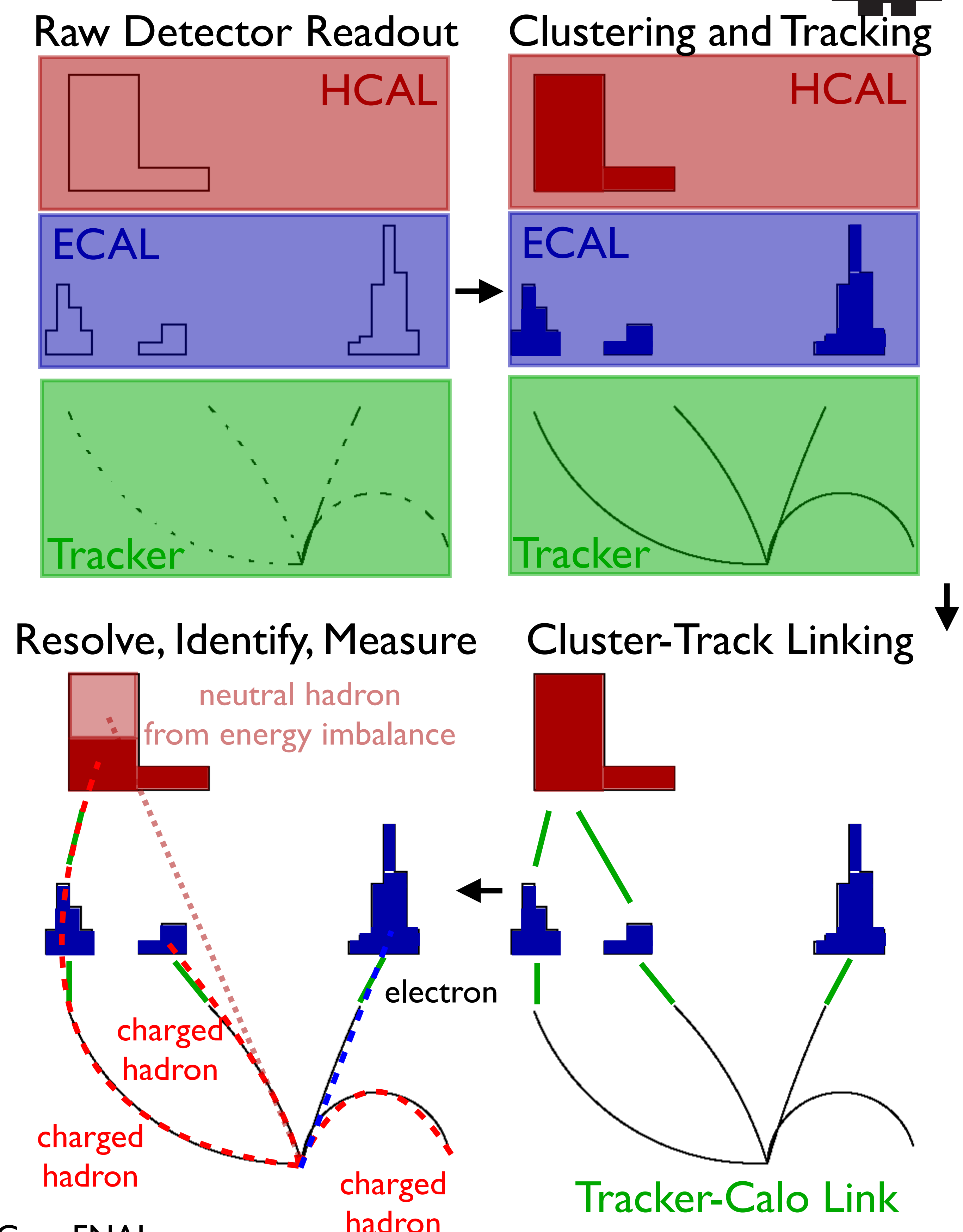
on behalf of all event reconstruction enthusiasts

11 August, 2014 - Qui-Nhon

Rencontres du Vietnam: Physics at LHC and Beyond

What is Single Particle Reconstruction?

- A reconstruction that yields an *unambiguous* list of *identified* final state particles
  - *Cluster* detector hits together in each detector
  - *Link* tracking data to calorimetric deposits
    - ~60% of particles in jets are charged hadrons
    - 30% γ, 10% neutral hadrons
    - Augment calorimeter response with tracking
  - Use of *all* detector information to *measure* and *identify* all particles in a collision
    - Optimized use of all information critical to performance
- This technique is colloquially known as "Particle Flow"



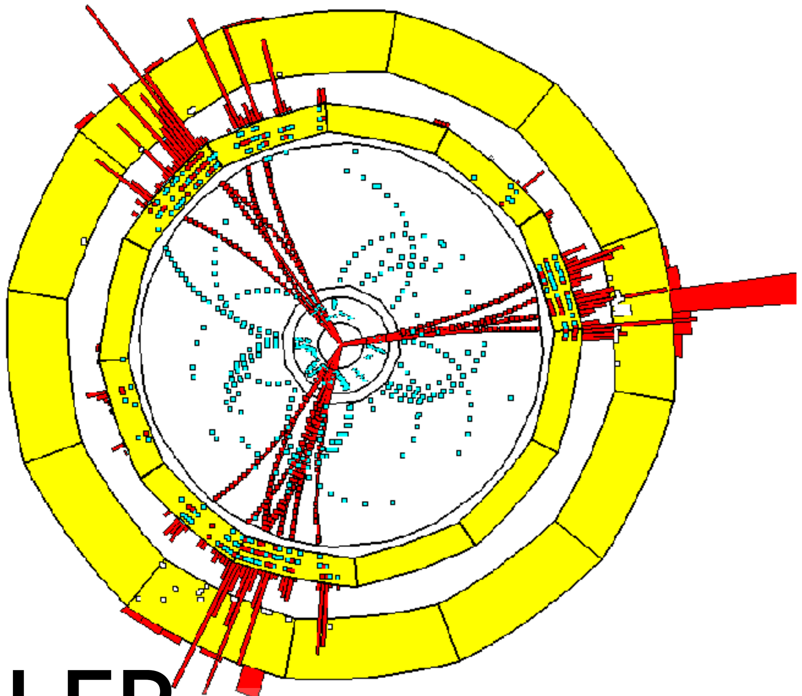
Lindsey Gray, FNAL

Varieties of Hadronic Environments

Event Complexity

Jets and low luminosity proton-proton collisions

Multiple overlaid proton-proton or heavy ion events

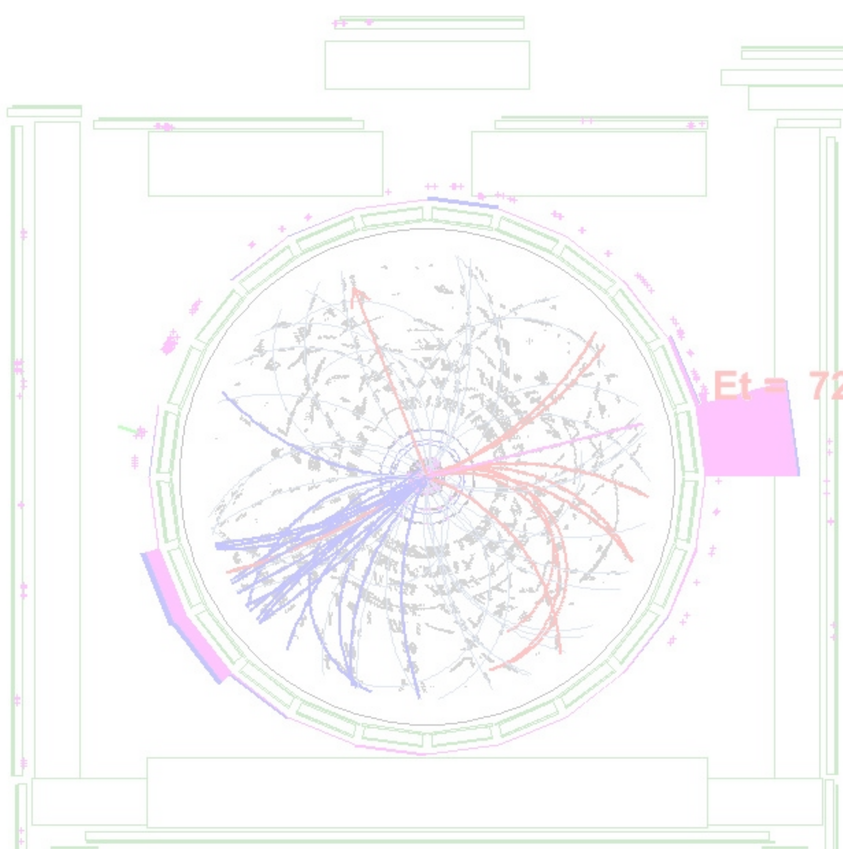


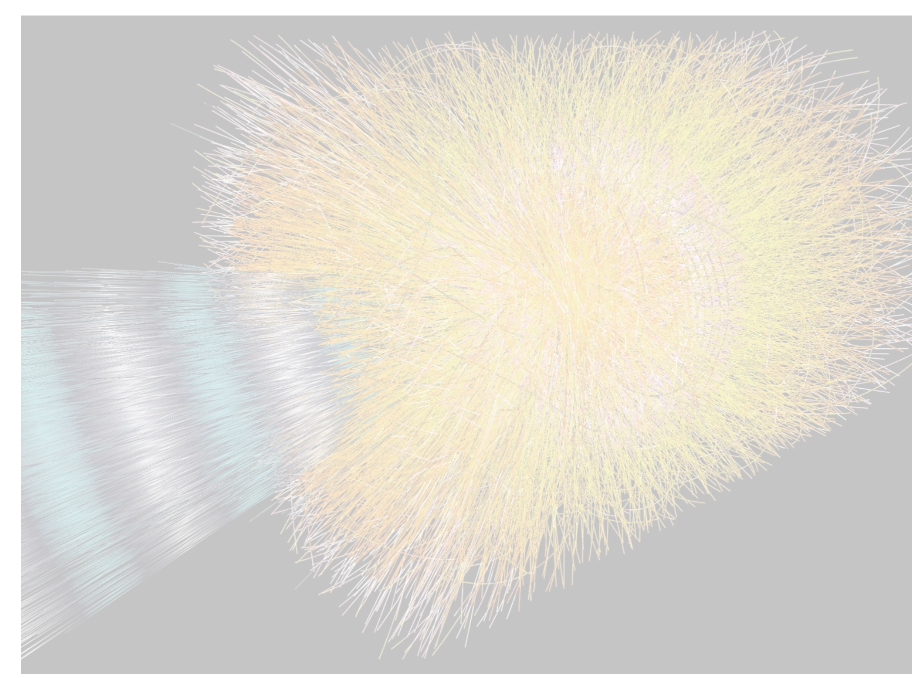
LEP

The simplest hadronic environments are jets at lepton colliders.

There's no color charge *at all* in the initial state and thus no initial state soft QCD radiation, or other 'clutter'.

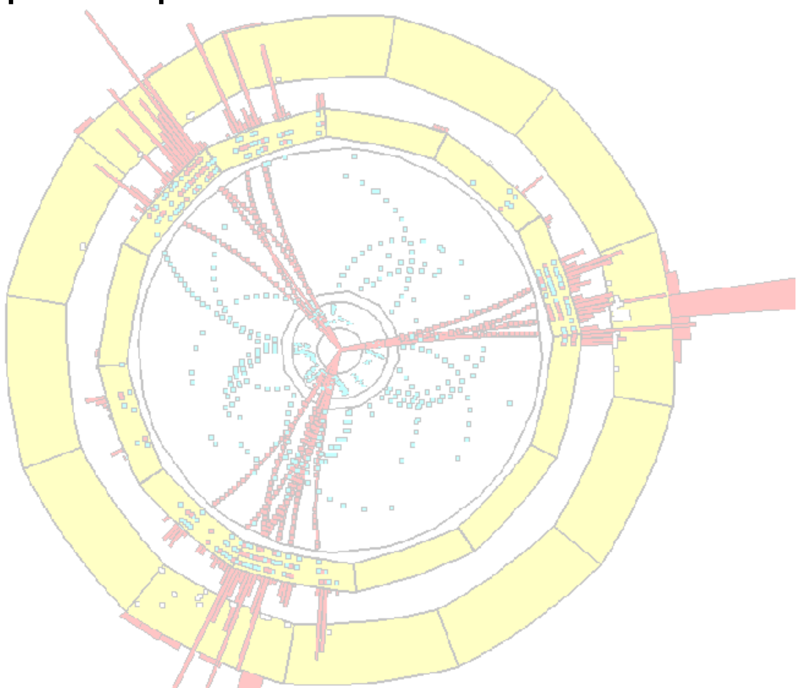
High p_T Jets still contain multiple tracks and have large calorimeter cluster multiplicity.





Varieties of Hadronic Environments

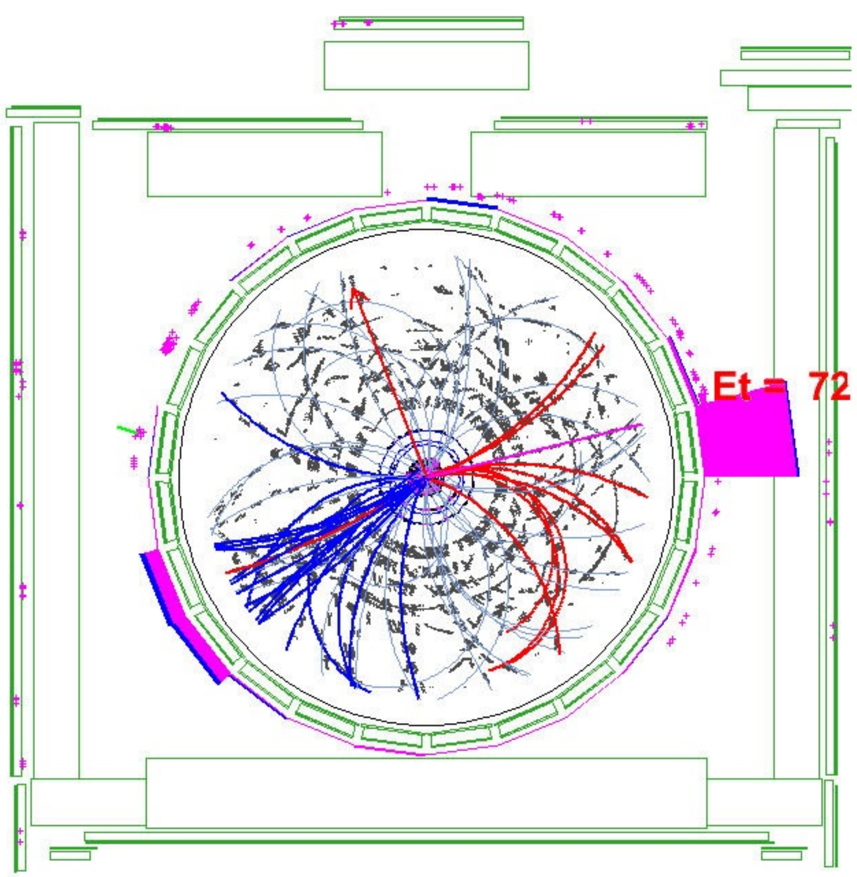
Jets and low luminosity proton-proton collisions — Event Complexity → Multiple overlaid proton-proton or heavy ion events



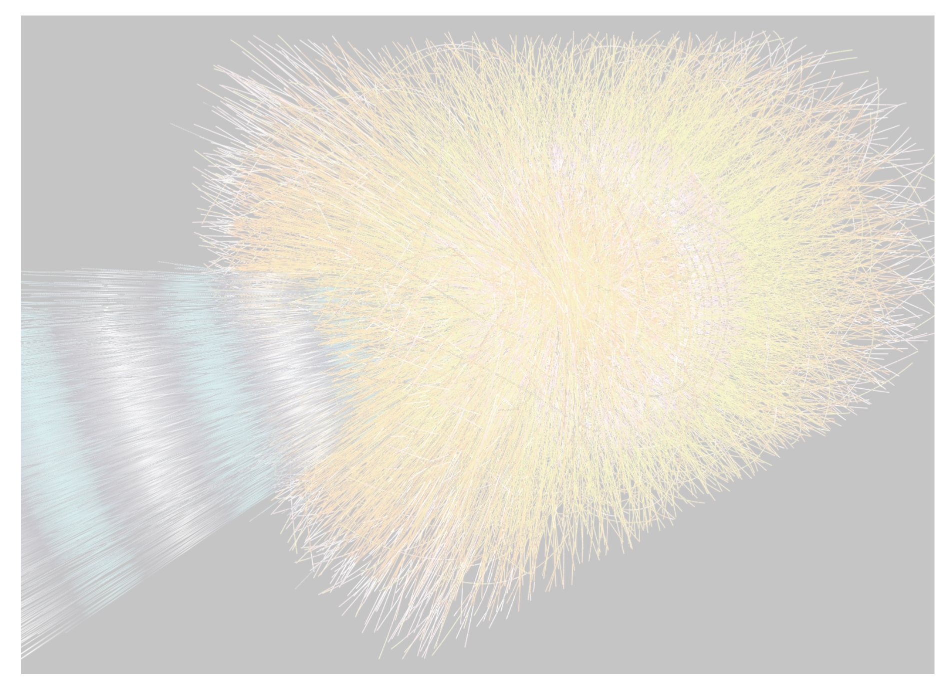
Hadron colliders bring additional complications, initial state color-charge exchange produces large amounts of soft hadronic radiation.

The number of charged and neutral particles increases due to proton breakup. More tracks and calorimeter deposits add complication to understanding each event.

Lepton efficiency and purity suffer from increased rate of hadrons, more "fakes" are created.



Tevatron

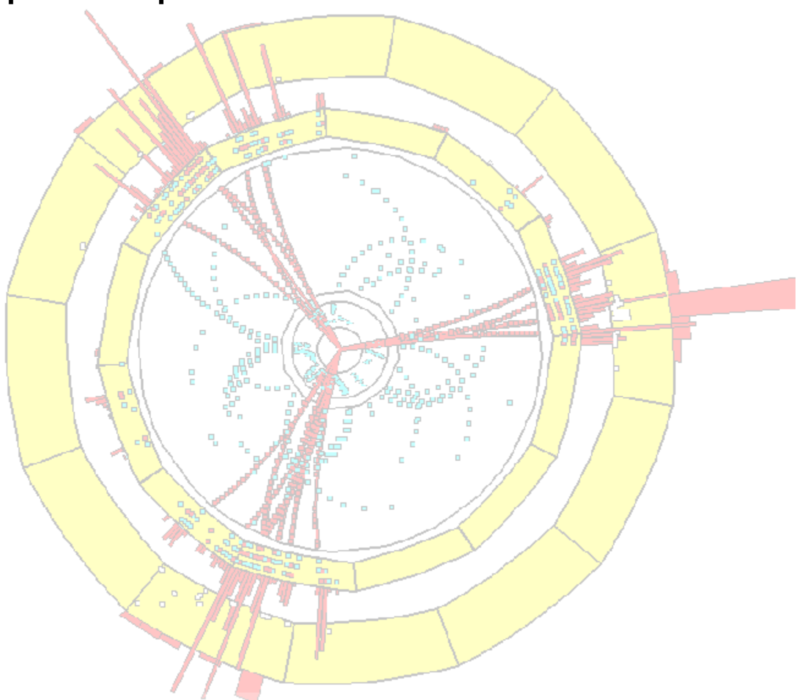


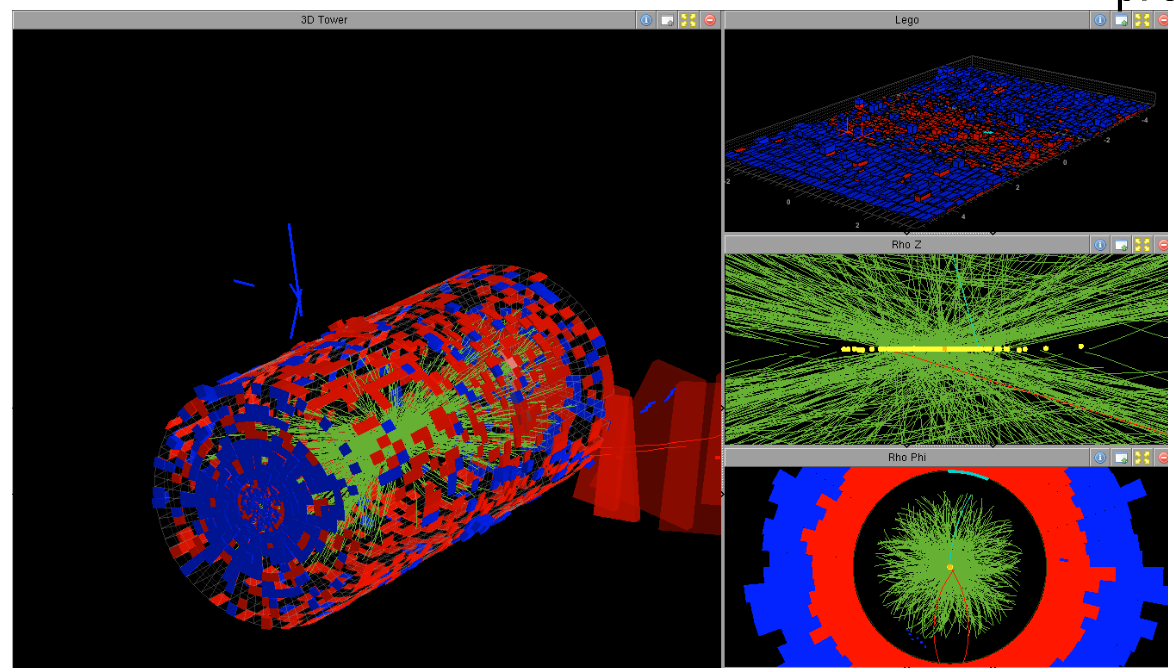
Varieties of Hadronic Environments

Jets and low luminosity proton-proton collisions

Event Complexity →

Multiple overlaid proton-proton or heavy ion events

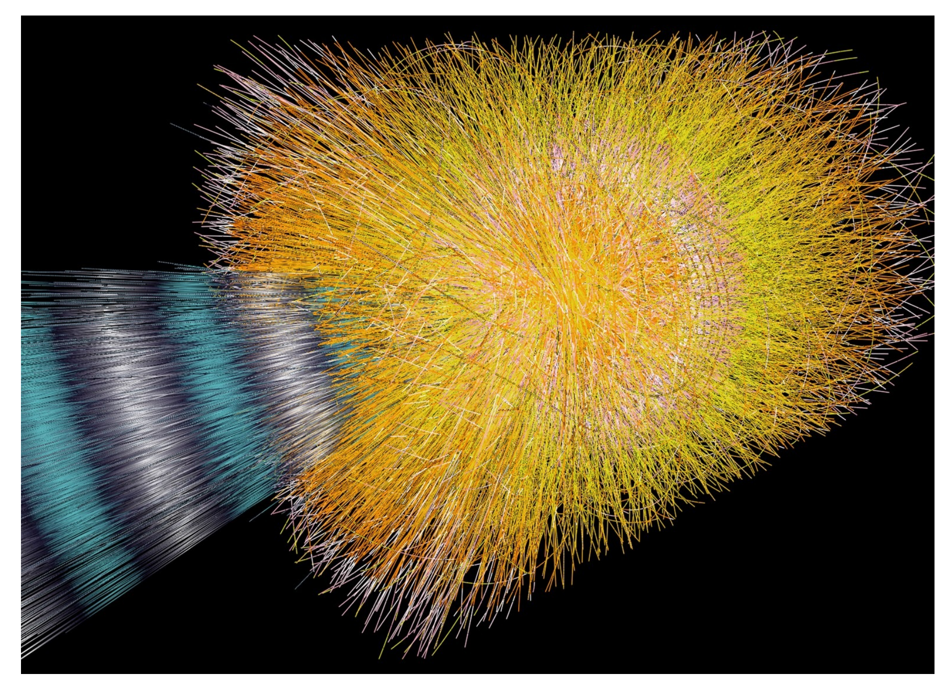




LHC pp*76 @ CMS

Large bunch occupancy / high luminosity in a hadron collider creates multiple collisions in each beam crossing, or ‘pileup’. Likewise, heavy-ion events demonstrate similar behavior concentrated at one originating point.

Each additional interaction combinatorially increases the complexity of an event. These events can result in hundreds or thousands of tracks, multiple TeV of energy in calorimeters.

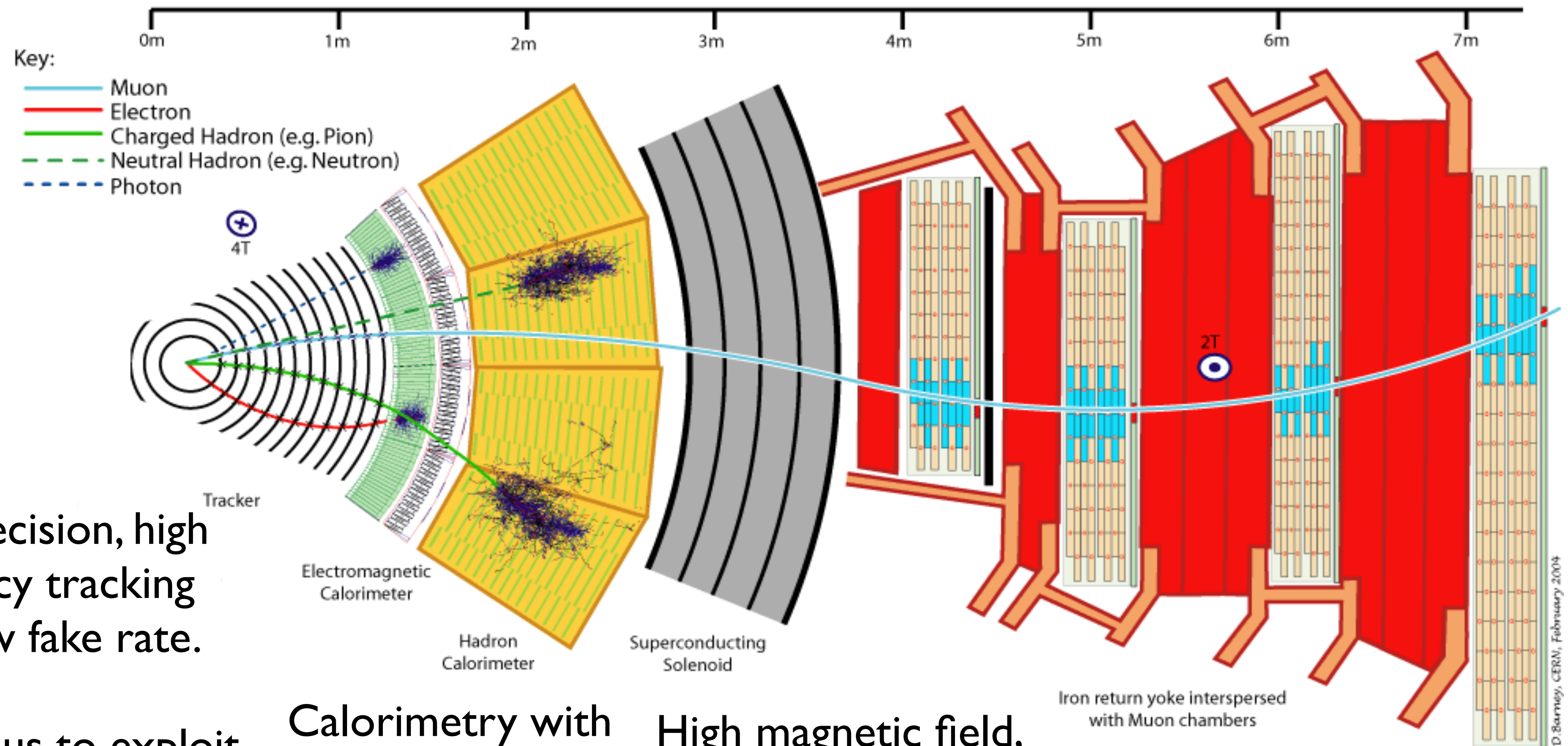


LHC PbPb @ ALICE

What’s The Challenge?

- Hadronic environments are naturally busy
  - Color-charged initial and final states produce jets composed of many particles and additional soft radiation
    - Calorimetric information can be ambiguous in this case
  - Accurately determining the event content requires careful accounting of detector effects
- Pileup and high particle density reduce effectiveness of individual detectors
  - Pileup can enhance existing detector peculiarities or dynamically reduce efficiencies due to readout or trigger limitations
  - Still possible to cross-reference detectors and resolve particles
  - Accurate matching of tracks to calorimetric information worsens due to combinatorics and increased noise
  - Residual energy from pileup can bias single particle response
- The challenge is *maintaining* the single particle response as the event complexity increases
  - Bonus: Make it work in the trigger!

Detector Design And Particle Flow



High precision, high efficiency tracking with low fake rate.

Large radius to exploit magnetic field.

→ Identify and measure charged particles with high efficiency

Calorimetry with good coverage and and fine transverse segmentation, at least

High magnetic field, improves effective granularity.

→ Separates charged particles from other charged particles and neutrals.

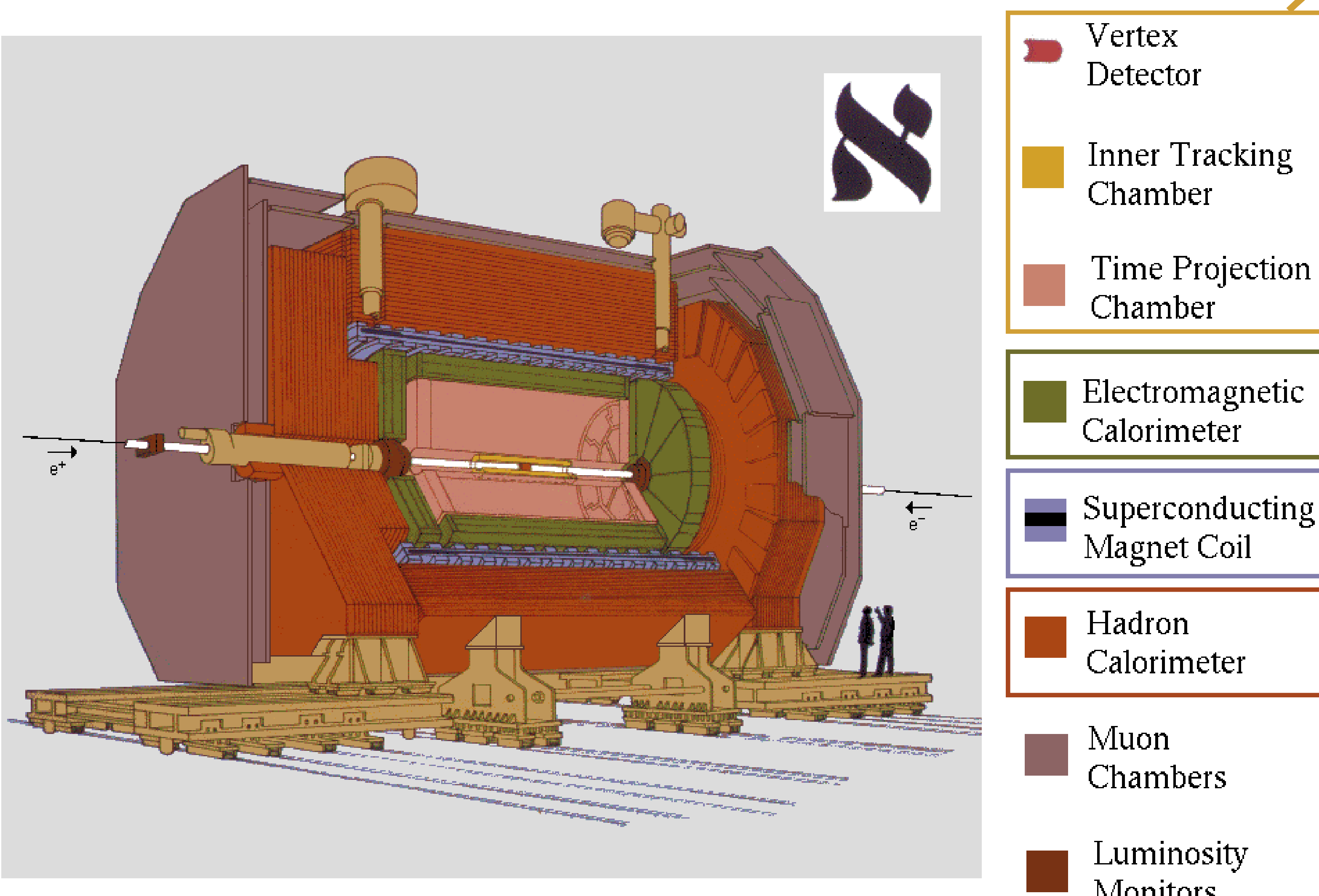
Efficient muon tagging system.

→ Tag muon tracks and change interpretation of nearby calorimeter deposits.

Case Study: The ALEPH Detector

Large number of hits, O(20), redundancy of tracking measurement.
fake track rate << 1%, efficiency ~100%,
~1% χ_0 material in front of calorimeters

VDET - two layers of silicon strips
ITC - 3D readout drift chamber
TPC - 3D readout time projection chamber



3x3cm transverse segmentation
4/9/9 χ_0 longitudinal segmentation
σ_E ~ 20%/sqrt(E)

ECAL - 22 χ_0 of lead interleaved with wire chambers for sampling before solenoid

Solenoid - 1.5 T,
$\sigma(1/p_T) \sim 6\times10^{-4}$ GeV^{-1}

HCAL - 7.2 λ_0 of iron + readout tubes located after solenoid

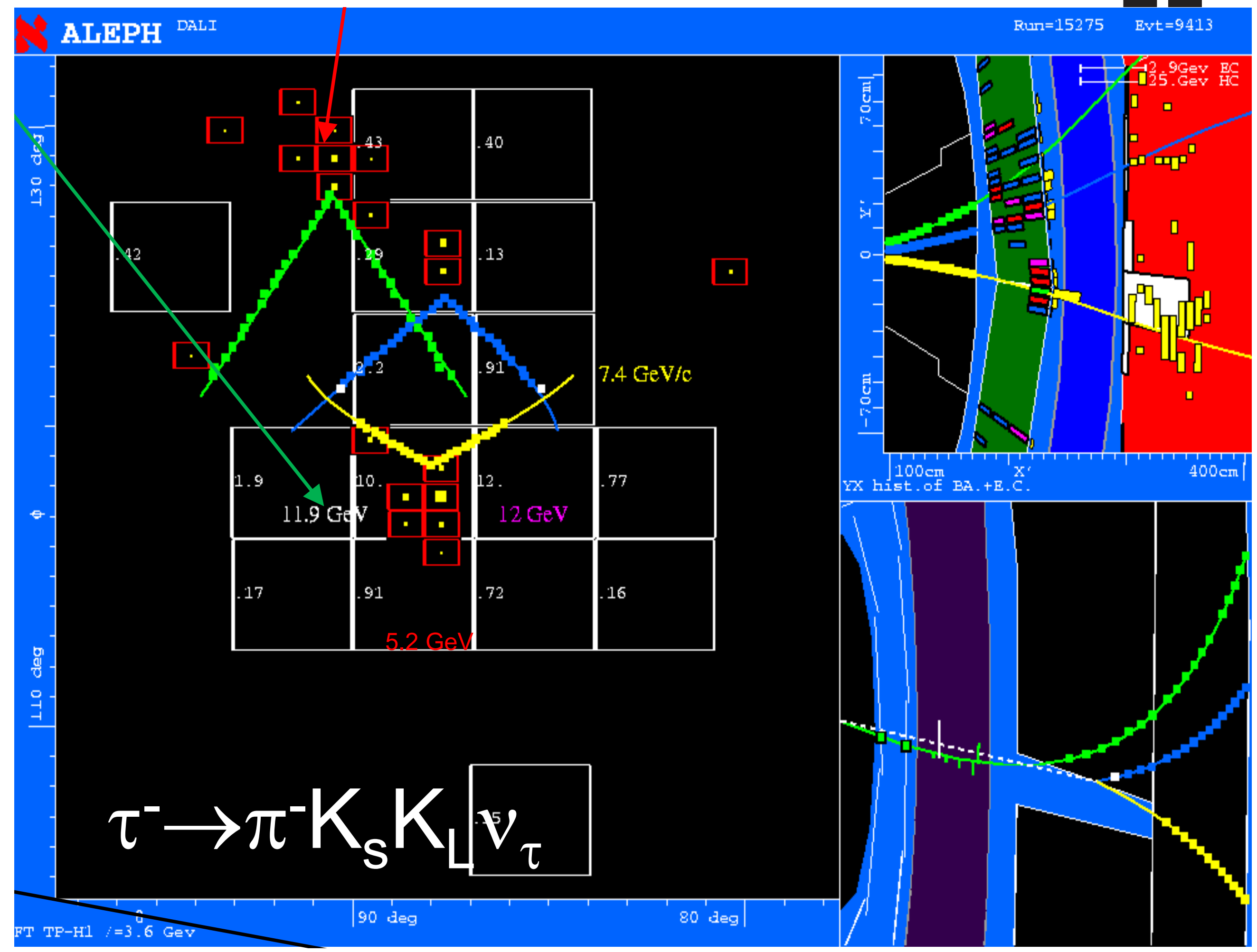
15x15cm transverse segmentation
No longitudinal segmentation
(Digital per-drift-tube 2D readout)
σ_E ~ 100%/sqrt(E)

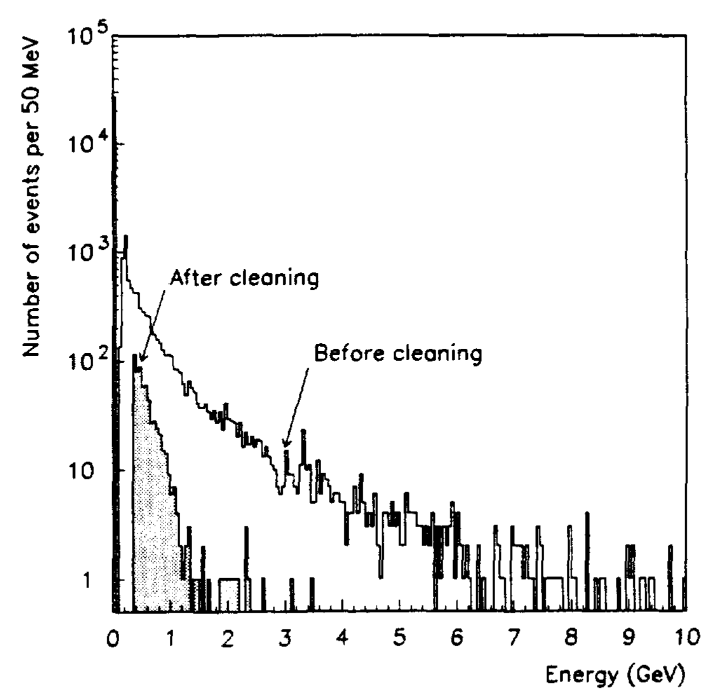
Summary: ALEPH is a detector *built* for particle flow
- Amazing tracking, even within jets, great p_T resolution
- Granular ECAL, multiple readout depths, can separate showers and follow shower development
- HCAL not the best but only used to measure 10% of total energy

Lindsey Gray, FNAL

Building The PFlow Algorithm at ALEPH

- Step 1, make an event display that lets you:
  - *See* exactly how tracks impact the calorimeters
    - Critical in designing track-cluster linking
  - *Confirm* that you are able to resolve particles that overlap in the detector
    - Used to determine handling of calorimetric excesses, splitting of charged hadrons
- Step 2, develop cleaning and calibration algorithms:
  - *Compare* redundant measurements
  - *Remove* anomalous detector signals
  - Create a *sane* environment free of spurious clusters and linking
- Step 3, understand the tails of reconstructed quantities:
  - Reconstruction *failures* almost always show up in the *tails* of distributions, few events to study for large gain!
  - Study tails to iterate *rapidly* on reconstruction improvements, create additional cleaning



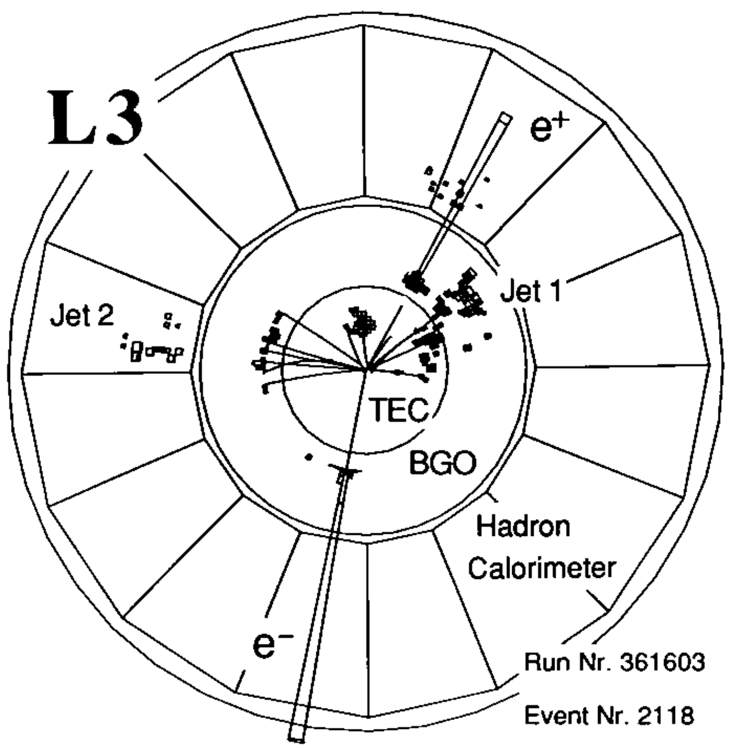


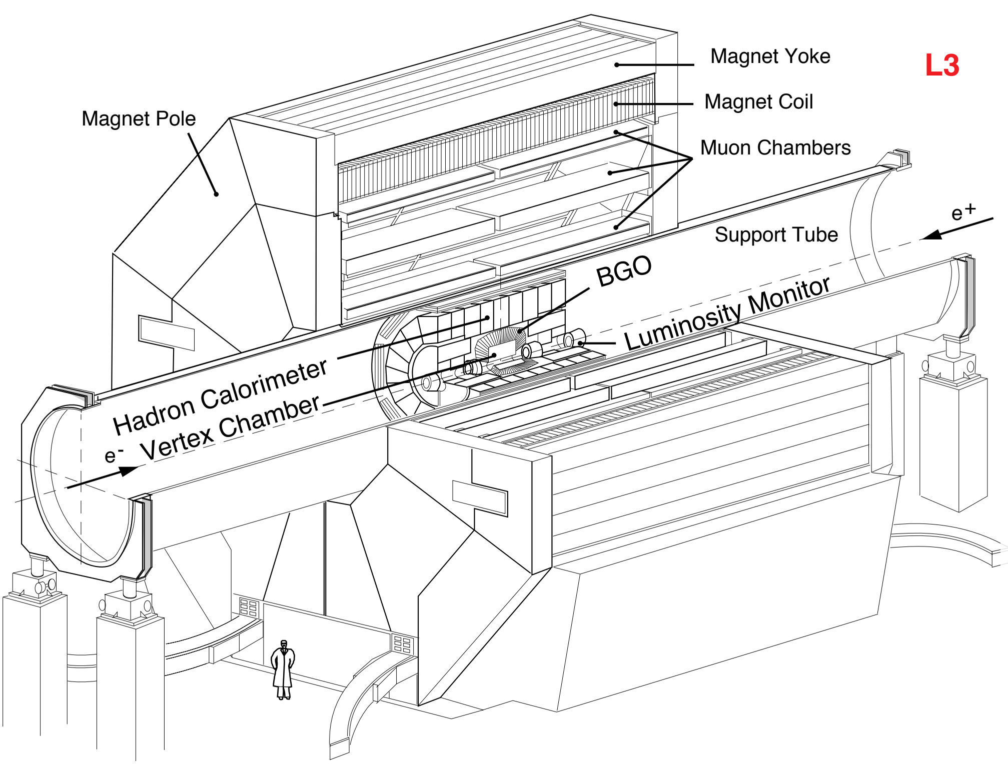
NB: These are three **very** big steps each with a host of devils and details to work through to arrive at the final product.

ALEPH: PFlow Deployment & Lessons

- After development & final tuning (5 months total!):
  - Particle Flow based analysis outperformed other ALEPH H+νν searches by a factor of <u>two</u>
  - Became the standard reconstruction in ALEPH until the end of LEP
- Yielded lessons, guidance in PFlow oriented detector design
  - Keep the individual sub-detector responses *simple*
    - *Little* material in front of calorimeters: one particle, one cluster
    - Makes combining, debugging detector information *easier*
  - Fine(-ish) granularity 2D or, better, 3D *only helps*
    - Improves efficiency and purity of tagging particles
    - Favored *over* resolution in hadronic environments (must balance)
  - Large magnetic fields *increase* effective granularity
    - Separation of charged particles from neutrals critical to accurate identification
  - Redundant measurements (like cathode vs. anode readout) in each detector are *critical*
    - Make creating a sane detector environment easy and less prone to over-cleaning

PFlow Contemporaries at LEP: L3

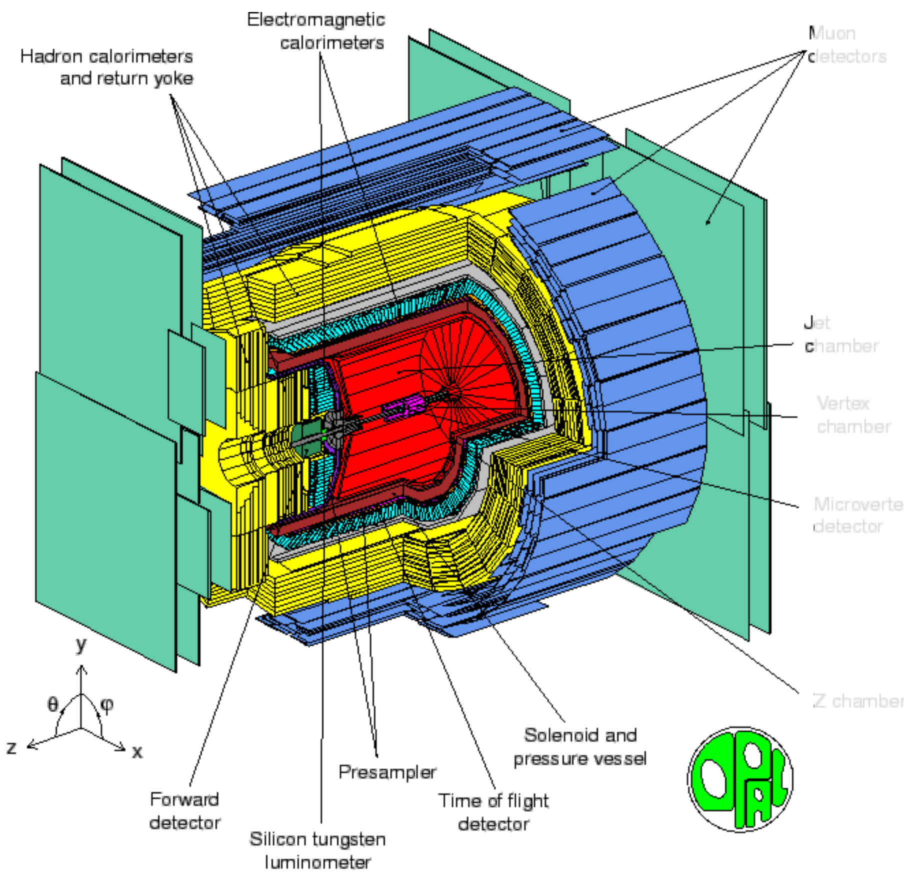


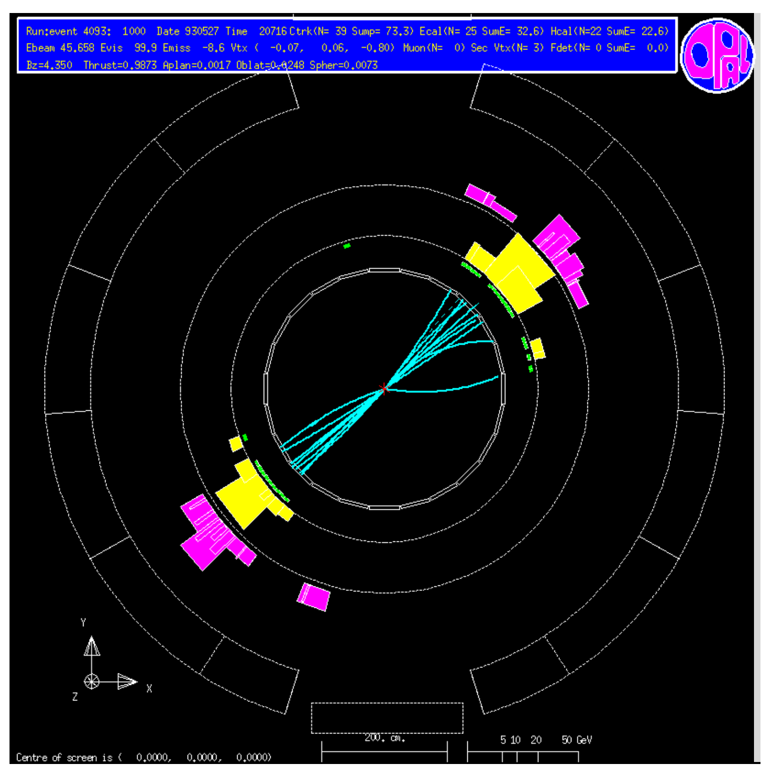


- Simple detector design
  - pro: all detectors within magnet
  - con: b-field is only 0.5 T
- 'Semi'-3D Tracking - lever arm of 0.317m
  - Transverse measurement from silicon vertex detector + "Time Expansion Chamber"
  - Z measurement using dedicated tracking chamber at large radius
- BGO Crystal ECAL, $\sigma(E)/E \sim 2\%$
  - No longitudinal segmentation, 21 X_0
  - 2x2cm front transverse segmentation
- Uranium + Wire Chamber HCAL $\sigma(E)/E \sim 55\%$
  - 53-58 layers in towers, ~6x6cm transverse size
- Track-cluster matching used, but limited in effectiveness by too-compact detector design

PFlow Contemporaries at LEP: OPAL

- Simple detector design
  - Only tracker within magnet
  - b-field is 0.44 T
- 3D Tracking - lever arm of 1.85m
  - Most of measurement comes from radially partitioned drift chamber
  - Silicon vertex detector for the innermost hits
  - Particle ID from dE/dx in jet chamber
  - 2 χ_0 in tracker alone

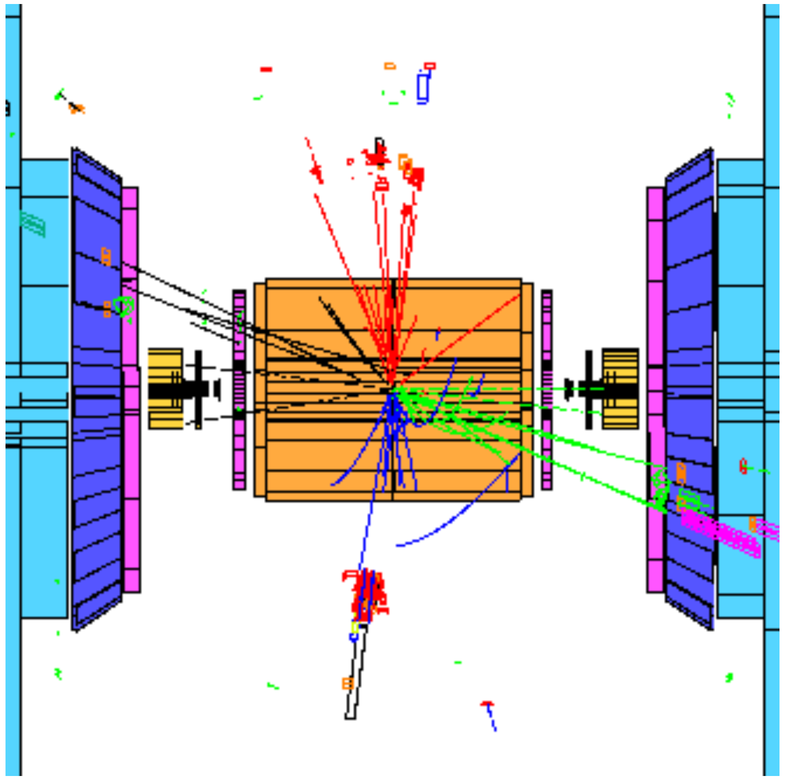


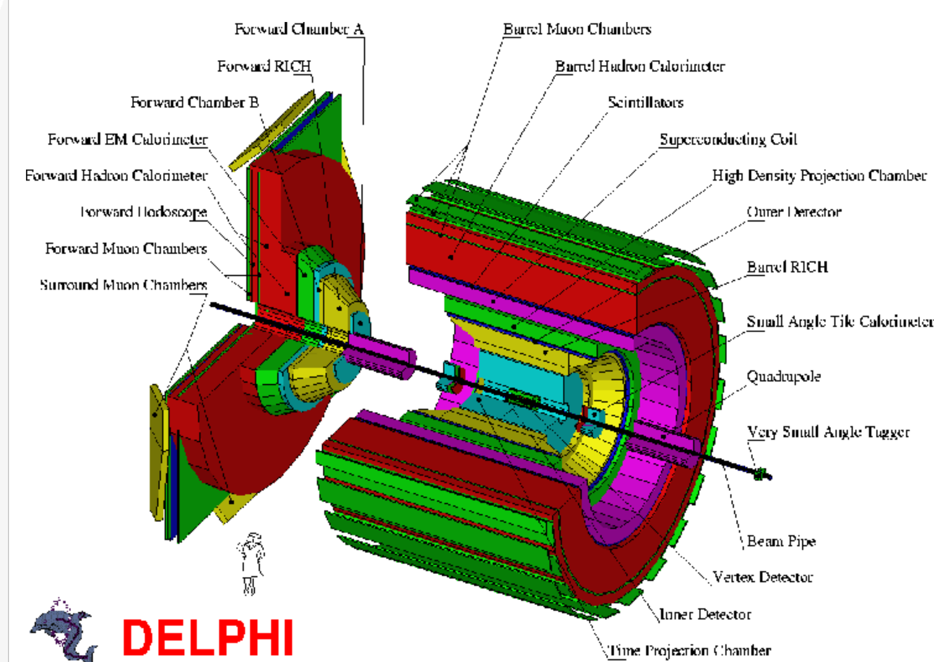


- Lead glass ECAL, $\sigma(E)/E \sim 5\%$
  - Poor spatial resolution
  - Augmented with fine-granularity pre-sampler
- Uranium + Wire Chamber HCAL $\sigma(E)/E \sim 120\%$
  - 9 layers sampled with streamer tubes
  - 2.2 λ_0 before HCAL!
  - Very hard to disambiguate particles in jets!
- Reconstruction achieved using tracks and calorimetry, but not in a combined way a la Particle Flow

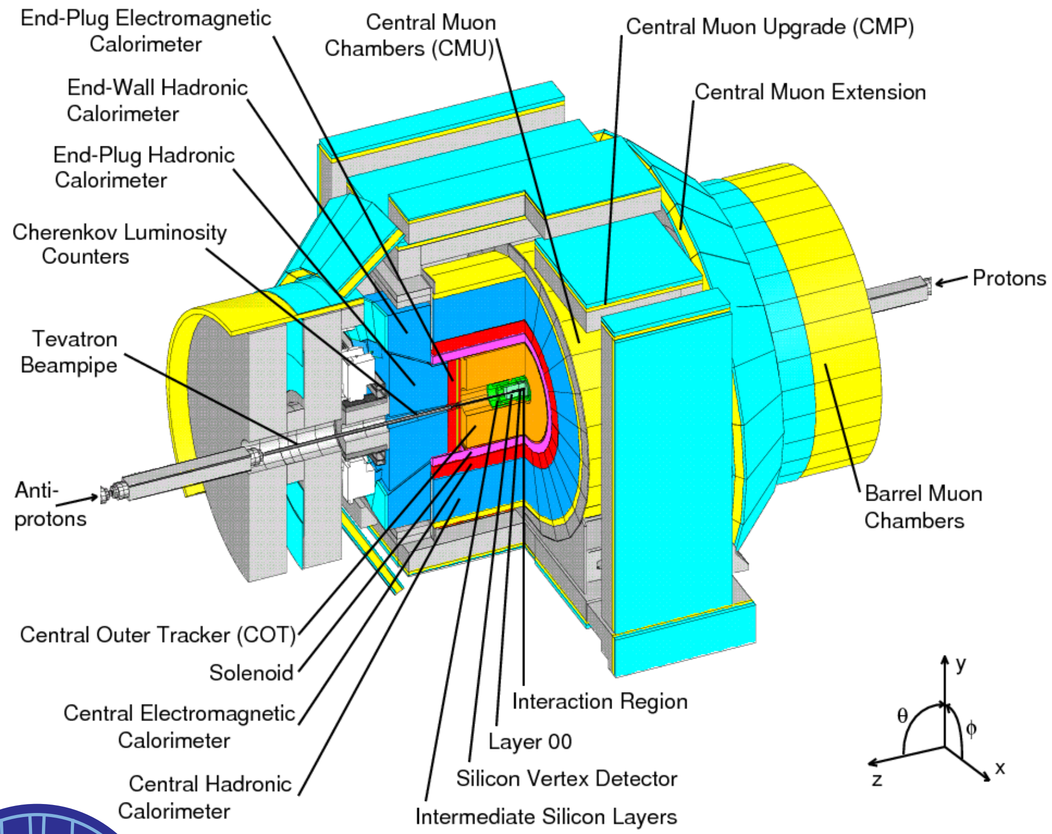
PFlow Contemporaries LEP: DELPHI

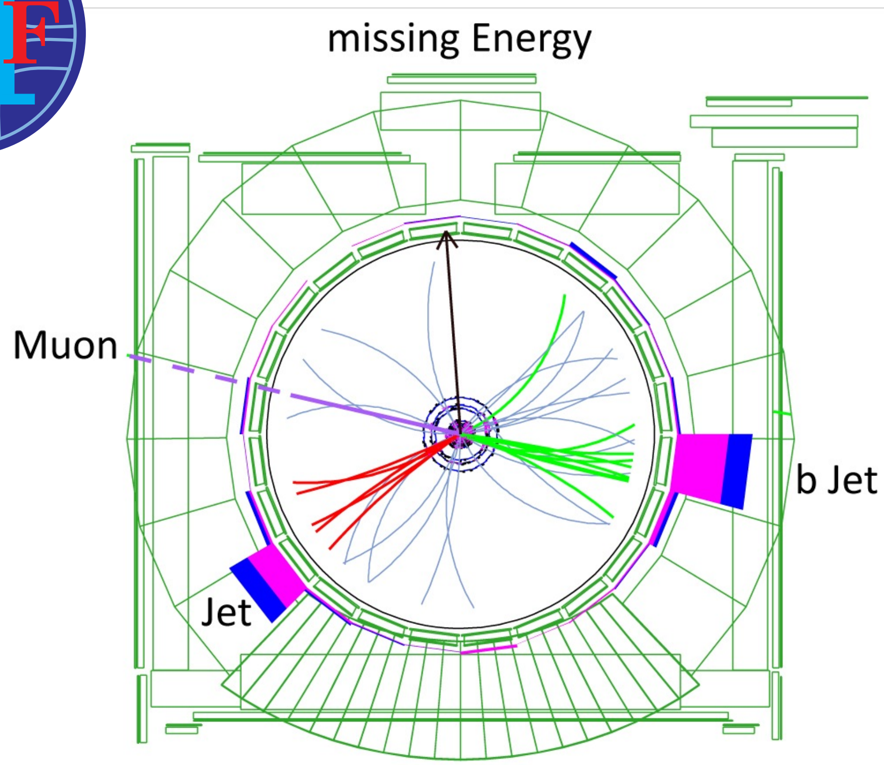
- Incredibly complex detector design!
  - ECAL, RICH, TPC all within magnetic field
  - Very focused on reconstruction of b-jets
  - b-field is 1.23 T
- 3D Tracking - lever arm of 1.22m, heavy tracker
  - Maximum 16 hits along tracks in TPC
  - RICH sandwiched between TPC and outer drift tubes
    - Allows for particle ID in addition to extensive tracking
- Lead interleaved with TPCs ECAL, $\sigma(E)/E \sim 6.4\%$
  - 9 Longitudinal samples
  - $0.1\theta \times 1\varphi$ transverse segmentation
- Iron + Streamer Tube HCAL $\sigma(E)/E \sim 17\%$ @ 45.6 GeV
  - 20 layers with rectangular shape
  - Embedded in field return
- Extensive and complex reconstruction algorithms to stitch multiple tracking detectors and calorimeters together, some ideas PFlow-like
- Above reconstruction not used centrally due to large variation of matching and vertex requires across analyses.





PFlow Contemporaries at The Tevatron

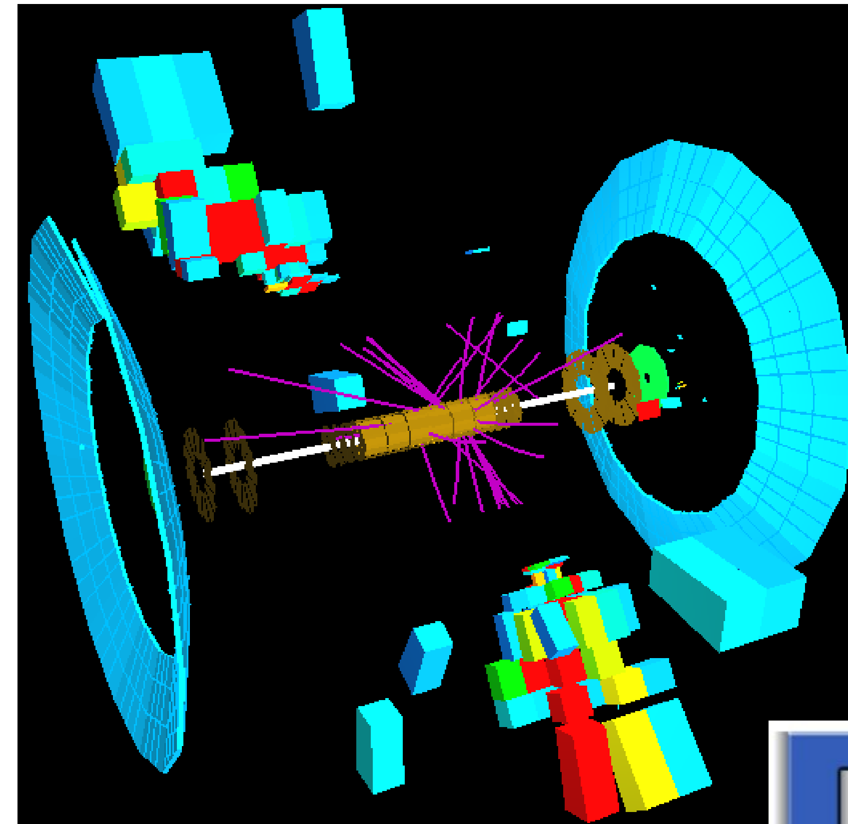


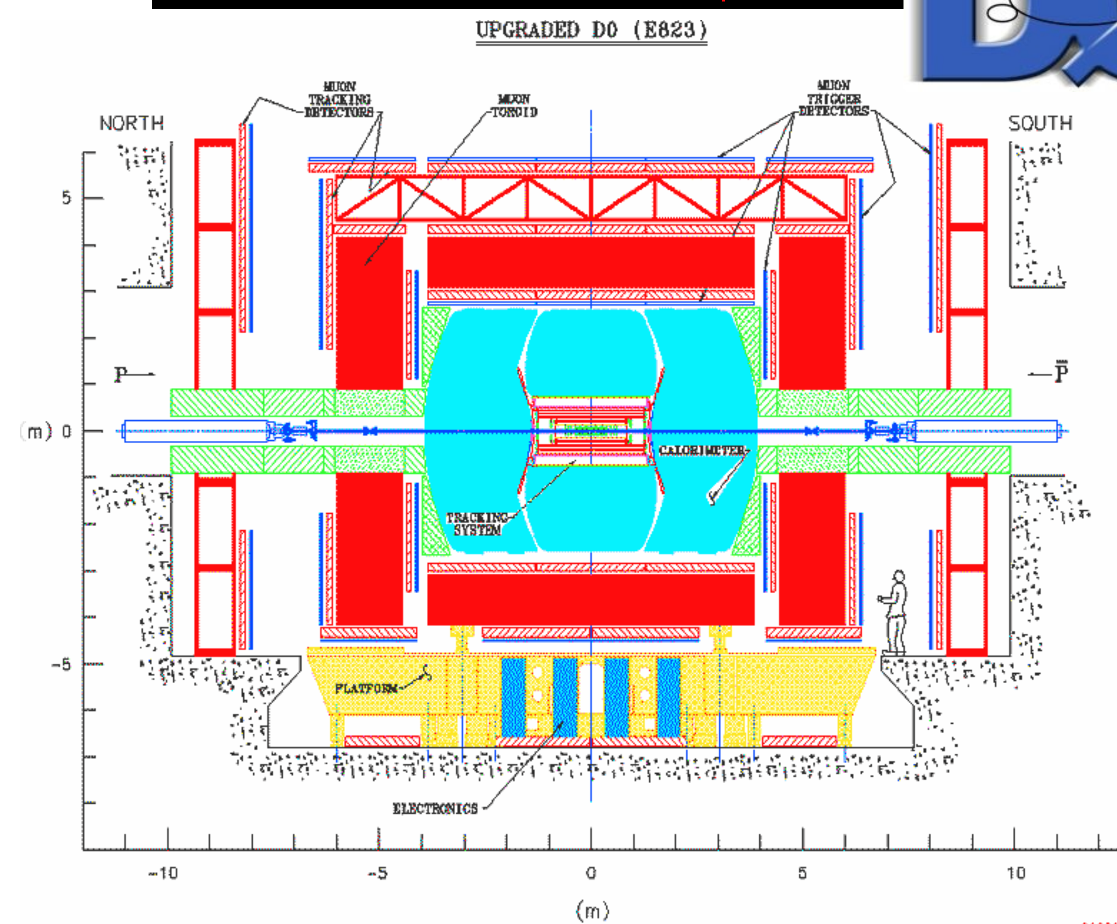


- Simple detector design
  - Only tracker within magnetic field
  - Calorimeters have coarse transverse and longitudinal segmentation
  - b-field is 1.4 T
- 3D Tracking, worse resolution in η/z, 1.3m lever arm
  - Large wire multiplicity drift chamber
  - First large deployment of modern silicon technologies in vertex detector and inner silicon strip layers
- Lead-Scintillator sampling ECAL, σ(E)/E ~ 13.5%
  - 15 degree φ segmentation, 0.11 η
  - Instrumented with shower maximum detectors for electron / photon tagging
- Steel-Scintillator HCAL σ(E)/E ~ 50%
  - Same segmentation as ECAL, all layers ganged together in readout
- Dedicated reconstructions for all particle types with heavy use of wire tracker, electrons measured in ECAL confirmed in tracker, for instance.

PFlow Contemporaries at The Tevatron

- Simple detector design
  - Only tracker within magnetic field
  - 2 T solenoid + 1.9 T endcap toroids
- 3D Tracking - lever arm of 0.5m
  - Silicon and scintillating fiber tracker
  - Very dense tracker, 5 X_0
- Liquid Argon + Uranium Calorimeter sampling calorimeter
  - ECAL σ(E)/E ~ 15%, four layers
  - HCAL σ(E)/E ~ 45%, four layers
  - Compensating calorimeter, transverse granularity not very fine given tracker radius
- Dedicated reconstructions for different particles, analyses

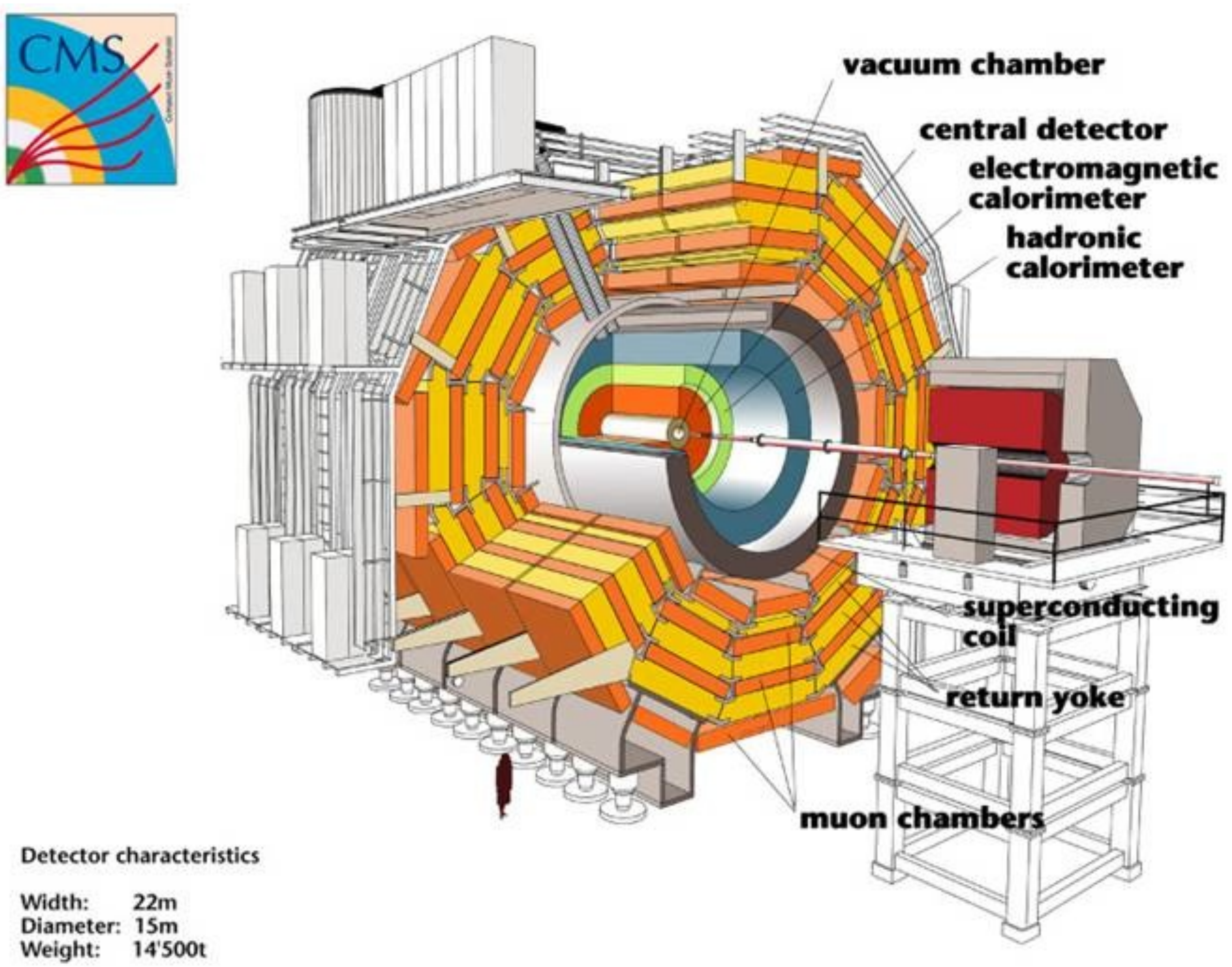




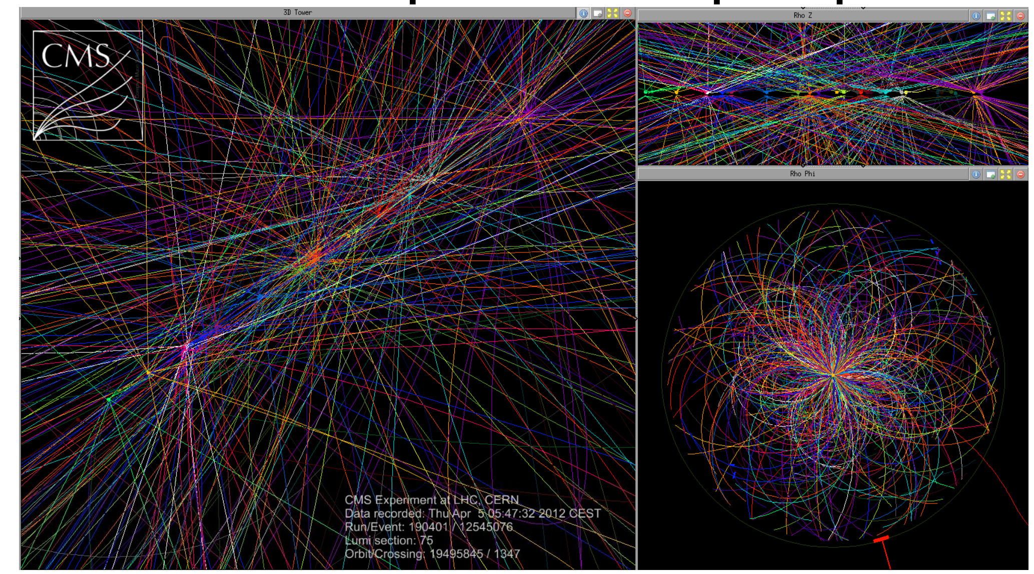
Particle Flow at CMS : History

CMS is an *edge* case for particle flow at first look:

4 T magnetic field enclosing Tracker, ECAL, HCAL
ECAL has fine transverse segmentation, $\sigma(E)/E$ ~3%
HCAL similar to that at ALEPH
but…
Tracker is not *fully* 3D and hit count is much lower
No redundant measurements in calorimeters
Tracker contains up to $2\chi_0$ of material before ECAL!



Not to mention the problem of pileup:



… but the ECAL granularity + 4 T B-field help to mitigate this. Let's see if the beneficial points of CMS outweigh the possible issues!

Meeting these challenges required:

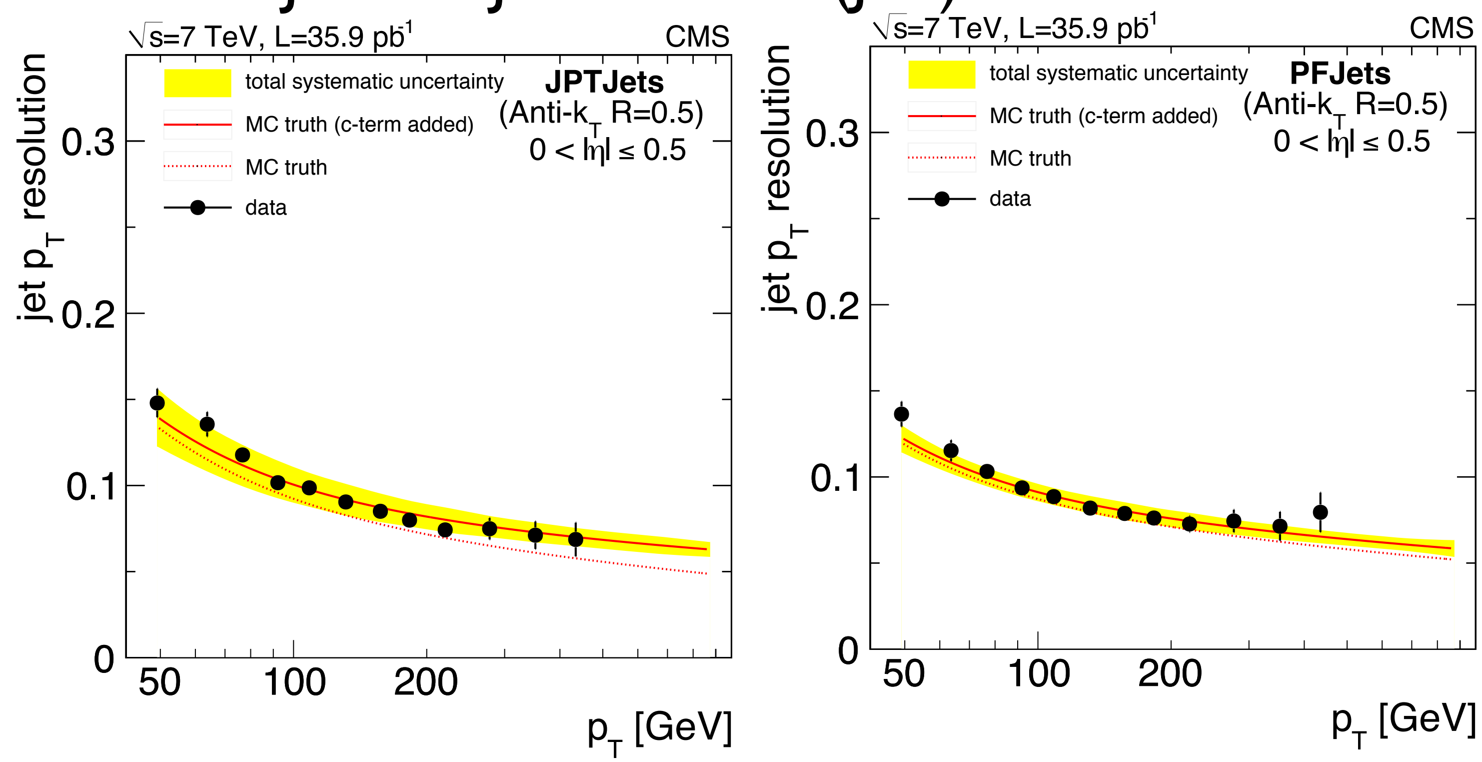
- Developing a custom particle flow algorithm for CMS
- Engineering iterative tracking techniques
- Use of GSF tracking to attain high quality electron tracks
- Developing cleaning against anomalous ECAL signals, readout noise, etc…
- Careful tuning to control fake rates and preserve jet and missing energy response

… and in the end these challenges were met.

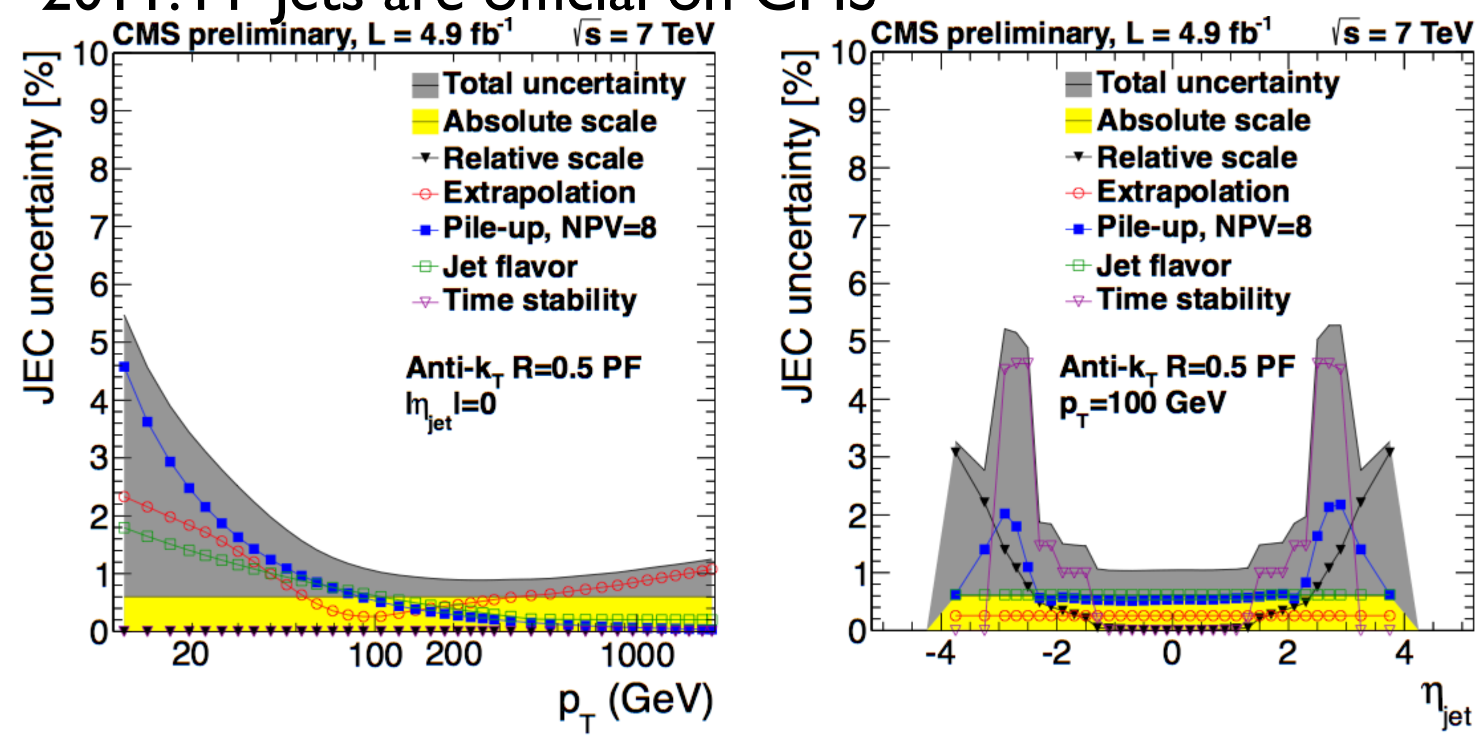
Jets + Tracks & Particle Flow

- However, this did not come without competition!
- Jets + Tracks, or JPT, jets were a competitor to particle flow in early data taking
  - Basic concept is to replace average calorimeter particle response in a jet with tracker measurement
  - Subject to problems of large calorimetric fluctuations at small jet momentum
- By end of 2010 and through 2011 particle flow based jets are used in most (> 90%) analyses
  - as well as particle flow based isolation, missing E_T, and leptons

2010: PFJets vs. "Jets+Tracks" (JPT)

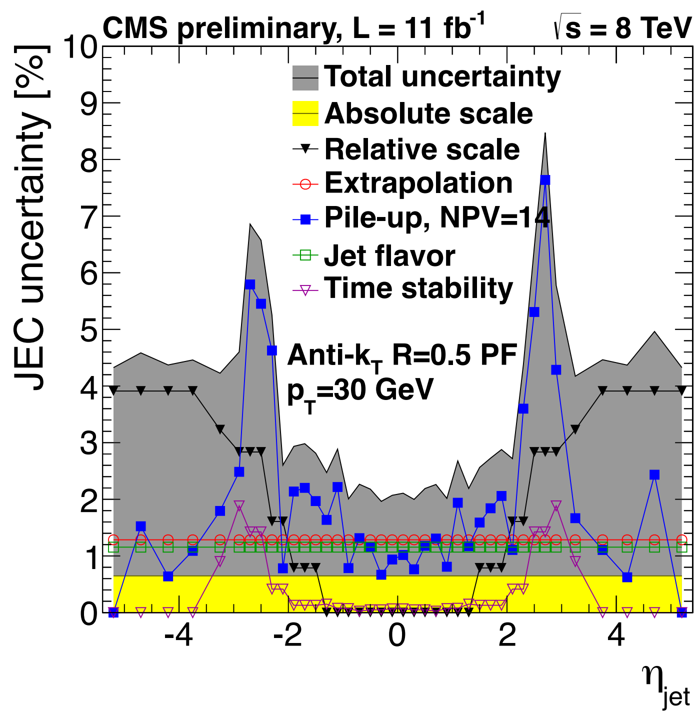


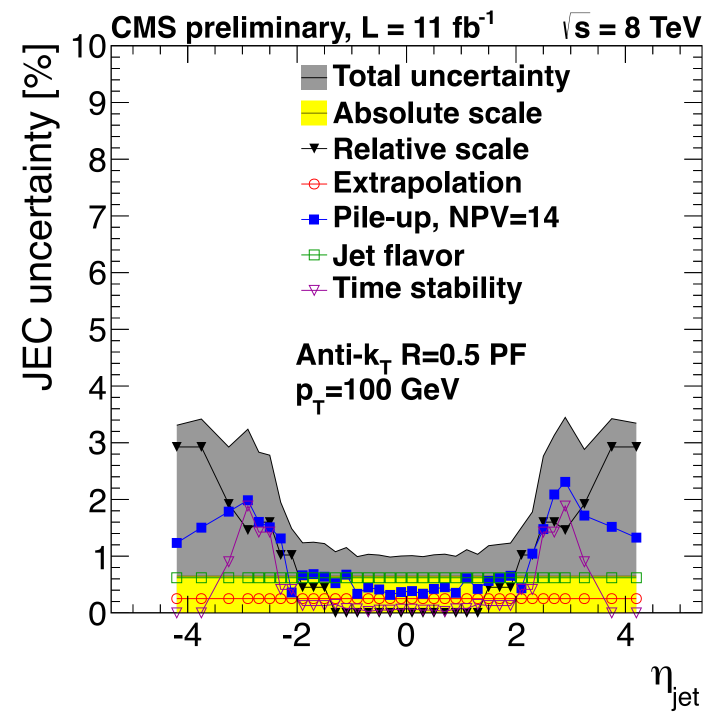
2011: PF Jets are official on CMS

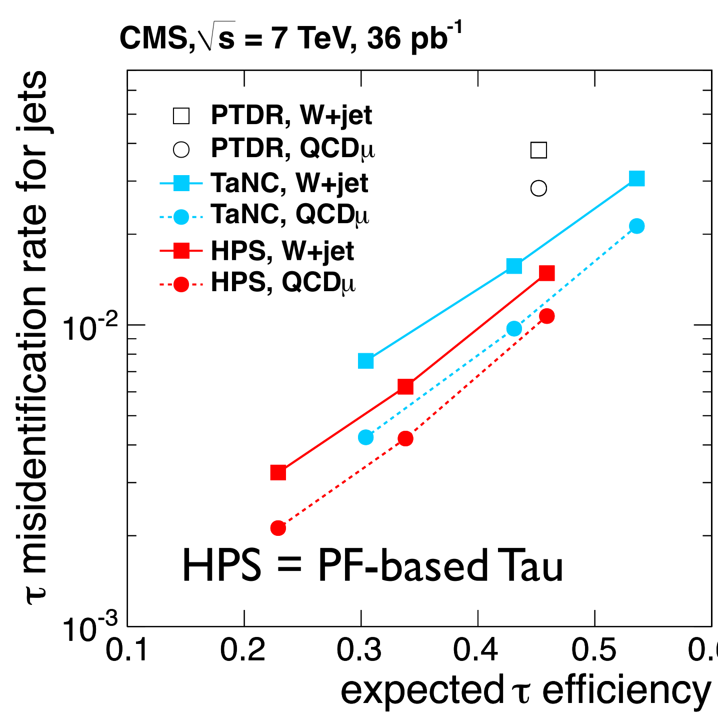


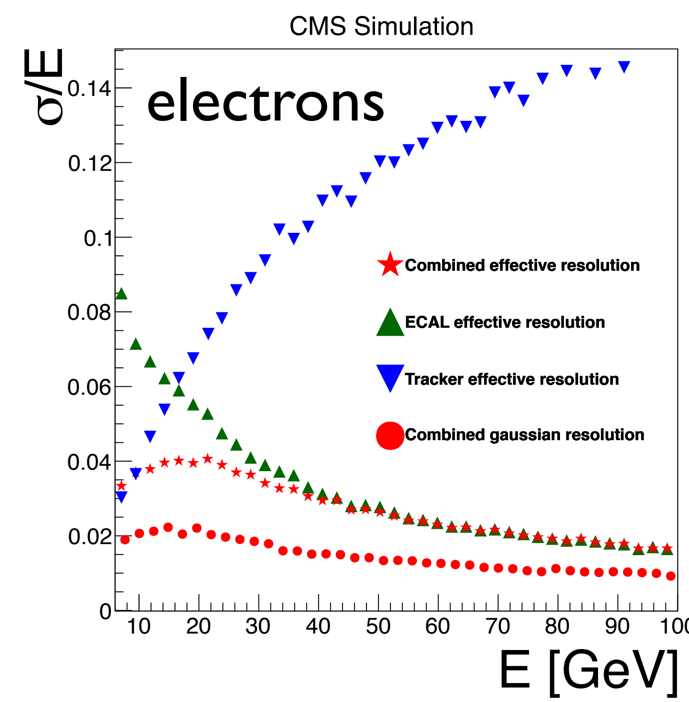
Today: Particle Flow at CMS

- Particle Flow remained a *common* reconstruction during 8 TeV
  - PFlow objects are often used ‘out of the box’
    - Reconstructed Taus are a textbook example
  - Few cases of other reconstructions being used
    - Photons / electrons used dedicated reconstructions that were reconciled post-hoc with particle flow
    - but… still use PFlow-based isolation, and PFlow ideas used in dedicated reconstructions
- The same people from ALEPH were critical in the success of Particle Flow at CMS
- With the start of Run2 all primary analysis objects on CMS will be unified under Particle Flow



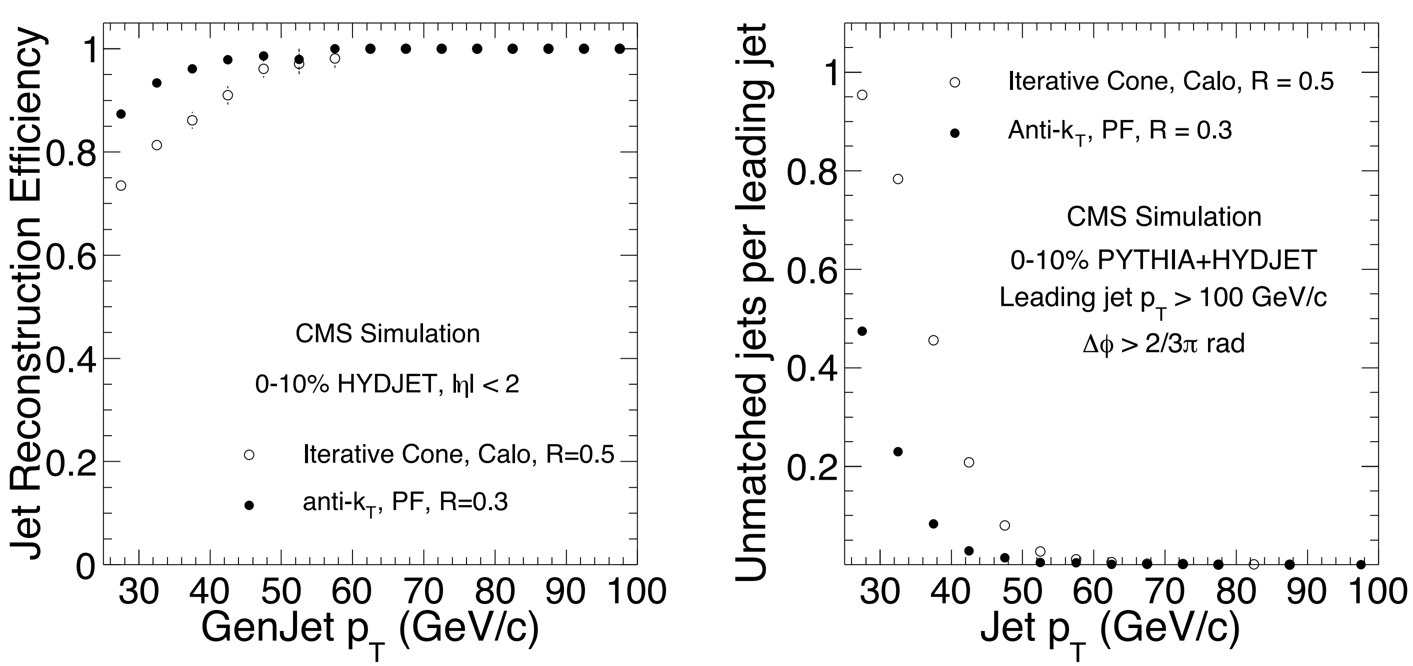




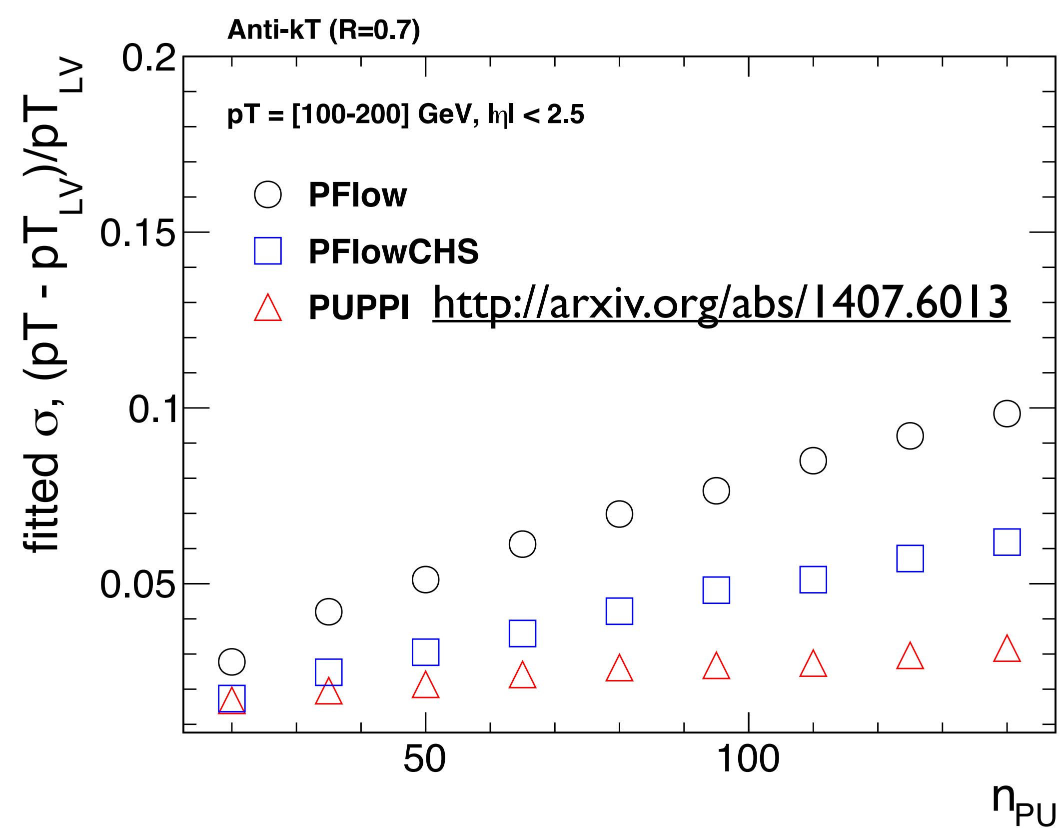


Lessons Learned from CMS

- You can do particle flow without redundant detectors, it is just *painful* and requires care
- Heavy trackers generate difficult problems for particle flow
  - Nuclear interactions, EM interactions in the tracker
    - hadrons, electrons, photons all shower *in the tracker* at CMS
  - Can be overcome with a careful approach to tracking
    - GSF tracking, iterative tracking
- Pileup and heavy-ion environments provide considerable challenges
  - Properly tuned, the <u>same</u> particle flow algorithm can handle these harsh environments
  - Furthermore, particle flow in high pileup environments can be *exploited* using algorithms such as PUPPI to compute the per-particle pileup probability



PF Jets in Heavy Ion Events



http://arxiv.org/abs/1407.6013

Event Reconstruction at ATLAS

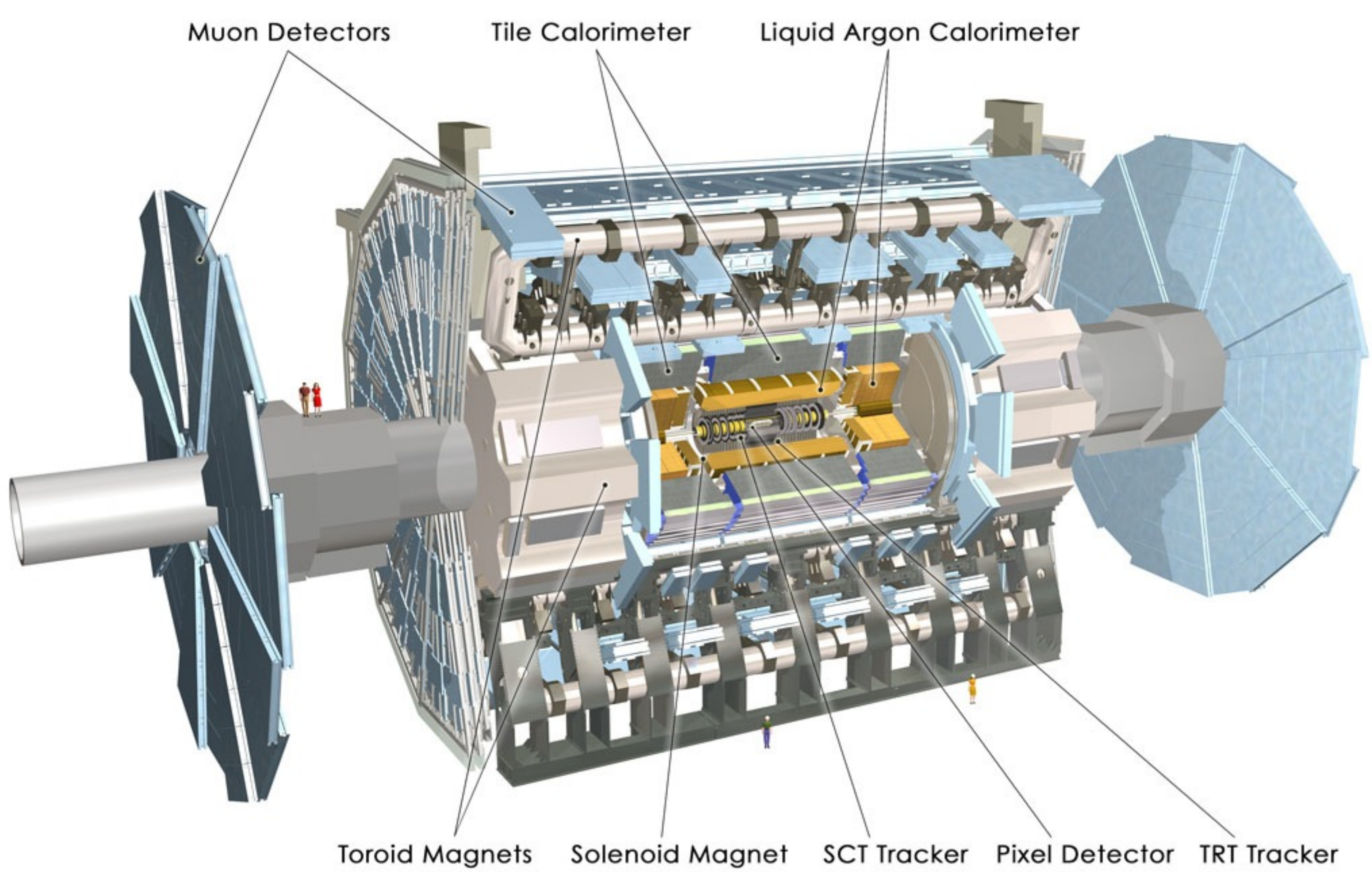
2 T magnetic field enclosing only Tracker, 1.1m lever arm
ECAL: Liquid Argon + Lead, $\sigma(E)/E \sim 10\%$

- very fine transverse segmentation with some longitudinal, up to 2 χ_0 in Tracker!,

HCAL: Scintillator/Steel or LAr/Cu/W $\sigma(E)/E \sim 55\%$

- worse transverse granularity than CMS, but farther from beam line

Aside from issues of tracker density, it could be a decent candidate for using particle flow.
Similar case to CMS.



Pileup in the calorimeter is dealt with through pile-up suppressed topological clusters, simple accounting for additional pileup activity as noise.

Excellent calorimeters and smaller B-field indicate that Particle Flow techniques likely bring little benefit.

However, there has been talk recently of…

Event Reconstruction at ATLAS

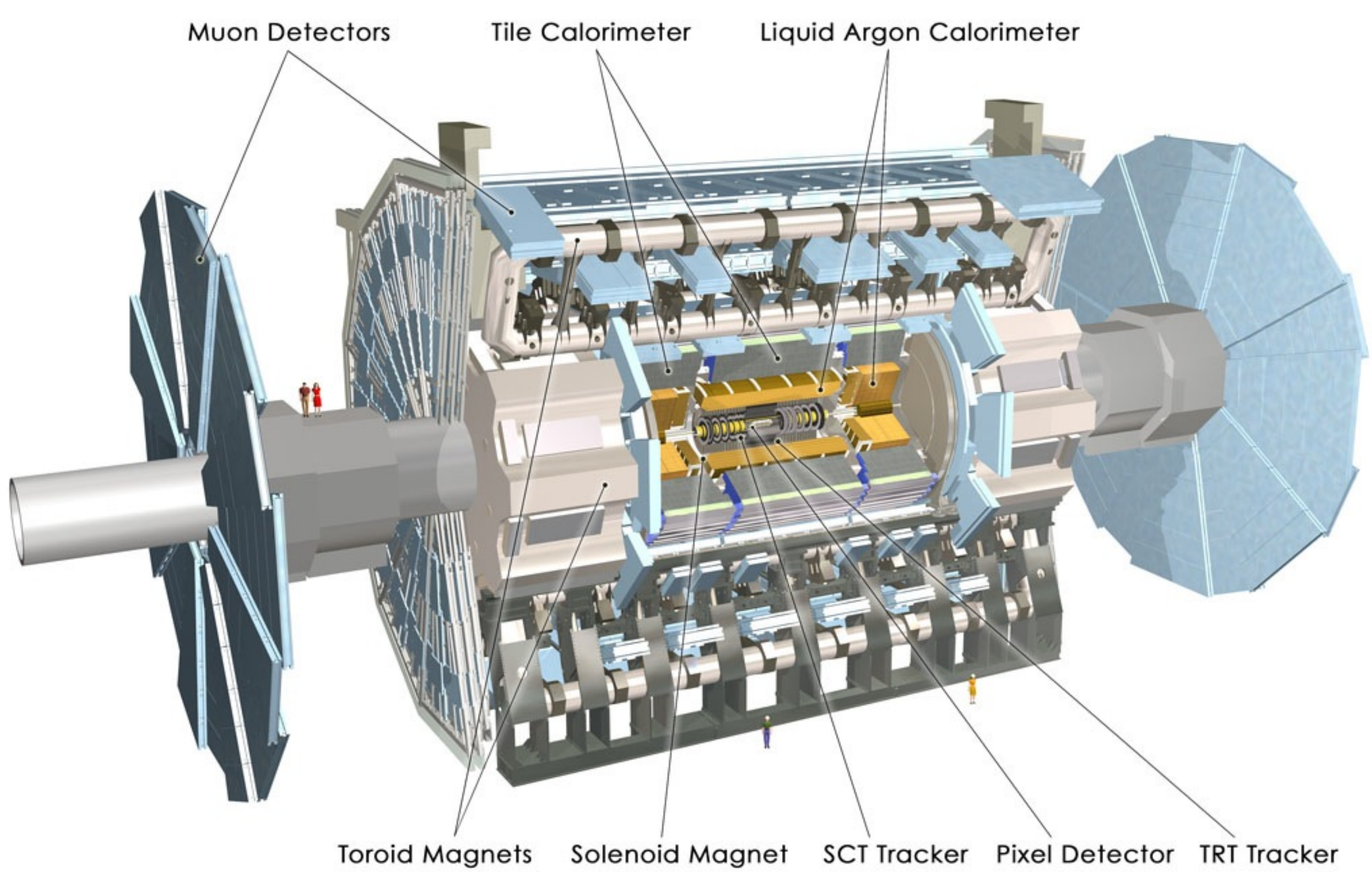
2 T magnetic field enclosing only Tracker, 1.1m lever arm
ECAL: Liquid Argon + Lead, σ(E)/E ~10%

- very fine transverse segmentation with some longitudinal, up to 2 χ_0 in Tracker!,

HCAL: Scintillator/Steel or LAr/Cu/W σ(E)/E ~55%

- worse transverse granularity than CMS, but farther from beam line

Aside from issues of tracker density, it could be a decent candidate for using particle flow.
Similar case to CMS.



Pileup in the calorimeter is dealt with through pile-up suppressed topological clusters, simple accounting for additional pileup activity as noise.

Excellent calorimeters and smaller B-field indicate that Particle Flow techniques likely bring little benefit.

However, there has been talk recently of…

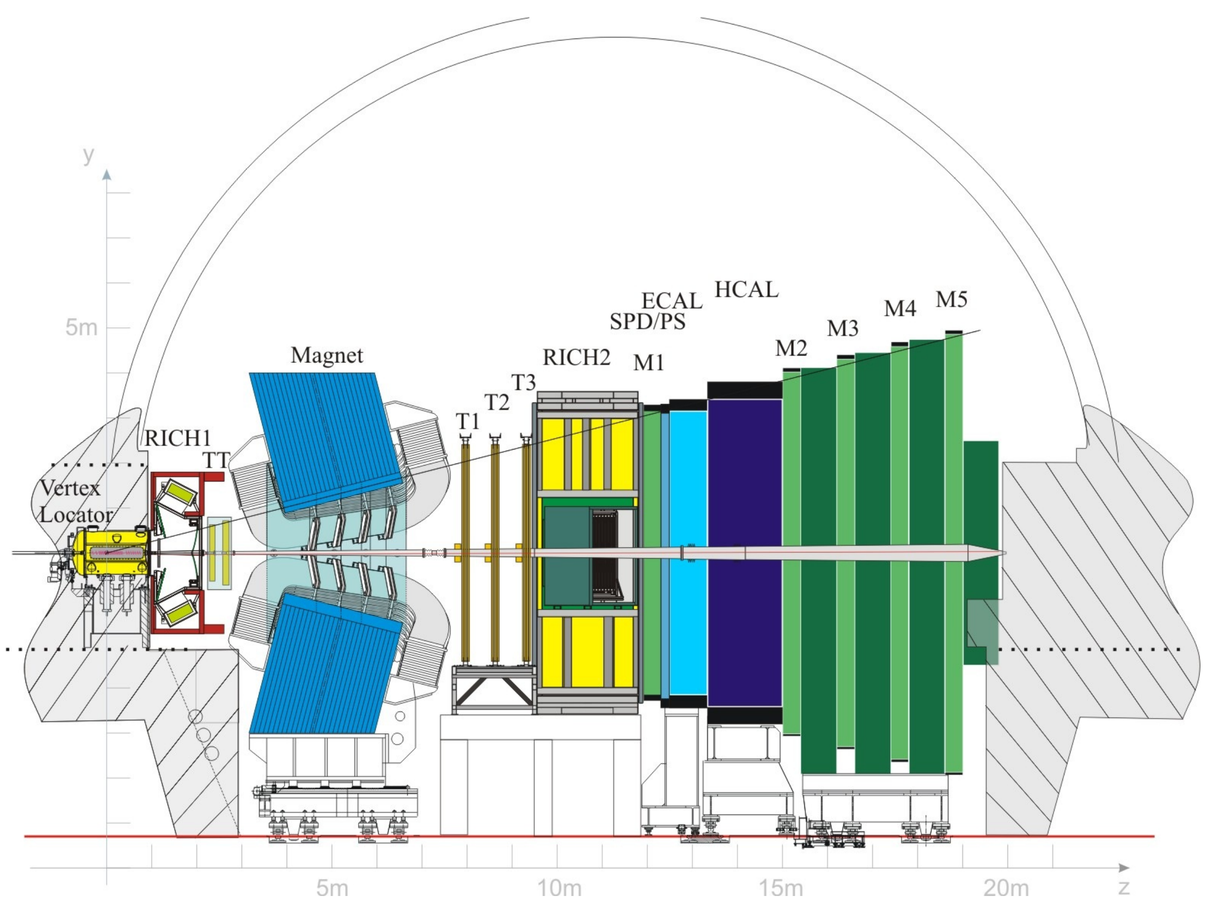
Particle Flow at ATLAS

Particularly with interest in improving the response of low-p_T jets, pileup mitigation.

In talking to developers, ideas seem to be along the lines of track-assisted clustering in 3D, and usual strategies of track-cluster energy balancing.

Following the usual PFlow maxim, "exploit everything you can use for maximum benefit".

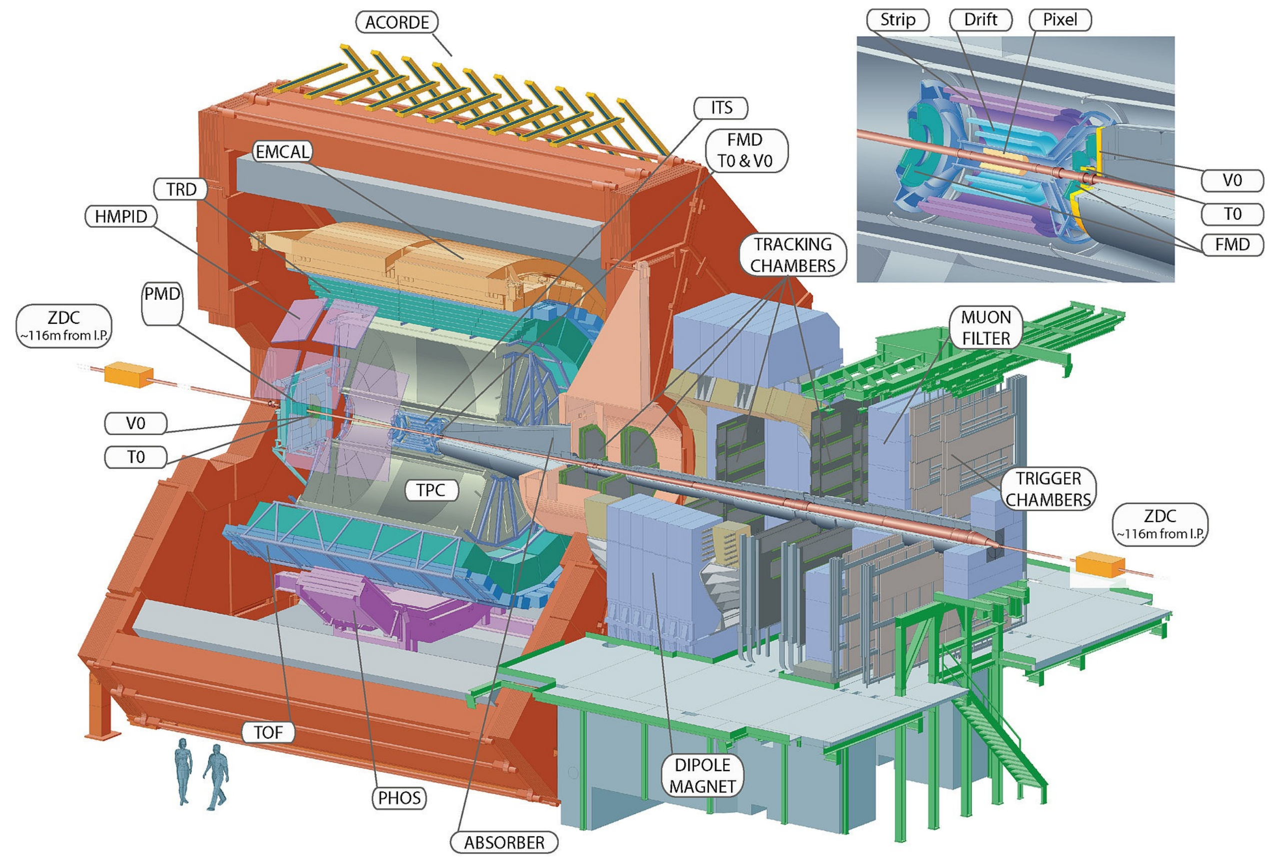
LHCb and ALICE



- Specialized detector (forward spectrometer!)
  - Multi-technology, extensive tracking system
  - RICH, many pixel layers close to beam for particle ID
  - Dipole B-field with B*dL = 4 Tm, similar to CMS
- 3D Tracking
  - Specialized, precise pixel detector (VELO)
  - Multiple tracking layers interleaved between RICH devices and magnets, extensive muon tracking for b-physics
  - 13m long tracking region with multiple gaps
- 'Shashlik'-Type ECAL, σ(E)/E ~ 10%
  - 4x4cm at high η down to 12x12cm at low η
  - Scintillator + lead pre-sampler
- Steel-Scintillator HCAL σ(E)/E ~ 70%
  - 13x13cm - 26x26cm
- 'Particle Flow'-ish algorithm employed in jet reconstruction. EM clusters not near a track are photons, HCAL clusters not near a track are neutral hadrons

LHCb and ALICE

- Purpose specific detector design for ion collisions
  - All detectors except muon system within magnet except forward muon taggers.
  - Forward muon spectrometer, RICH, transition radiation detector
  - Uses L3 magnet, 0.5 T + 3 Tm dipole for muon spectrometer
- Large TPC (2.5m outer radius) + Muon Spectrometer
  - Up to 159 hits on a track to deal with track density
  - Strip, pixel, drift chamber inner detectors
- Two ECALs
  - Lead tungstate central detector, similar to CMS
  - Shashlik ECAL covers until $|\eta| \sim 1.0$
- No HCAL, ~1 λ_0 in EMCAL
- Particle ID achieved using TPC, RICH, TRD, and Muon system
- Jet reconstruction using tracks + EMCAL, neutral hadrons accounted for with calibration.



Lindsey Gray, FNAL

Particle Flow at the HL-LHC (CMS)

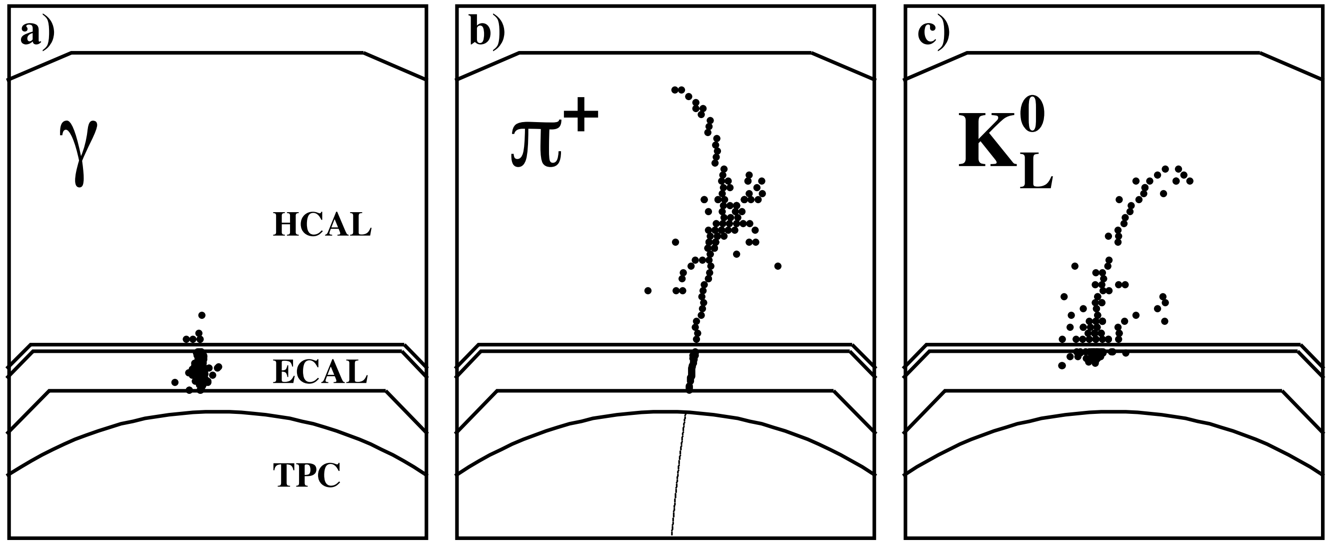
- At least at CMS, PFlow considered critical for maintaining a rich SLHC physics program
- Particle Flow arguments are at the core of many upgrade design decisions and proposals
  - Barrel essentially staying the same, particle density increase is not too large at 140 PU
    - Adding longitudinal segmentation to HCAL to help find low energy pileup deposits
  - Endcap proposals both favoring granularity over resolution to deal with large energy density in forward region
    - 1x1cm Shashlik-style calorimeter with resolution near that of present CMS ECAL Endcap (~5%)
    - High-granularity calorimeter option motivated by ILC designs with ~60 total layers and 1cm transverse granularity for first 40 layers
    - Latter has much worse resolution, but provides excellent tagging of charged hadrons in dense environments
- The particle flow concept has scaled well from LEP to the LHC, up to 20-40 pile up events, but the question still remains if it scales to the most dense environments.

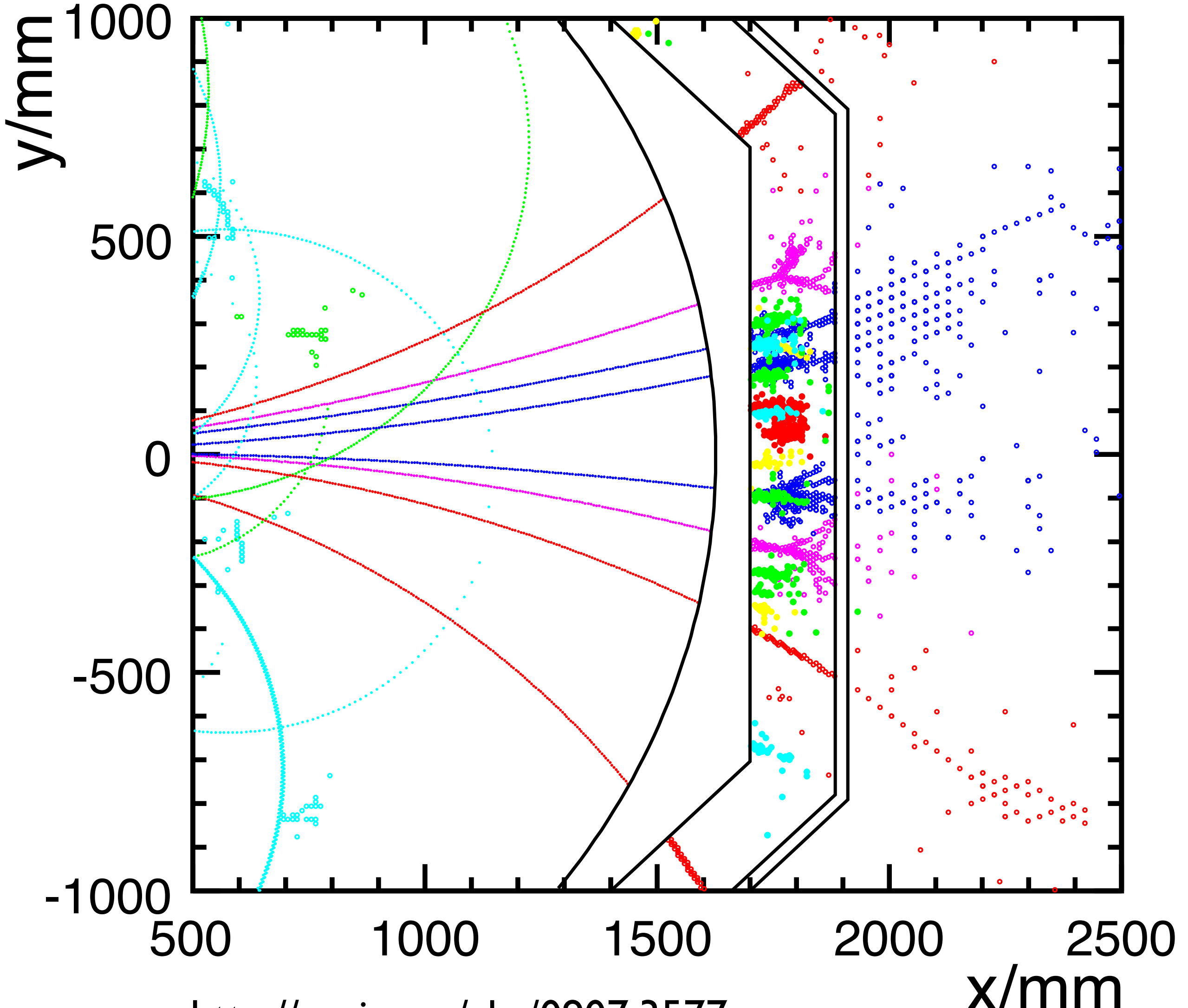
Variations on the PFlow Theme

- The guidelines provided by the particle flow concept provide many ‘knobs’ to turn
  - Both algorithmically and in detector design
  - e.g. : Granularity, absorber type/thickness, clustering principles, etc.
- Particle Flow motivates a rich R&D program for developing precision detectors
- The main avenues of research have been towards exploiting high-granularity calorimeters in the context of the ILC
  - 30 layers in ECAL, 30 layers in HCAL
  - O(cm) transverse segmentation, millions of cells
  - Gives unprecedented accuracy in determining particle multiplicity within a calorimeter
  - Furthermore, these detectors naturally have redundant energy measurements and cross checks (hit density, pulse heights, etc.)

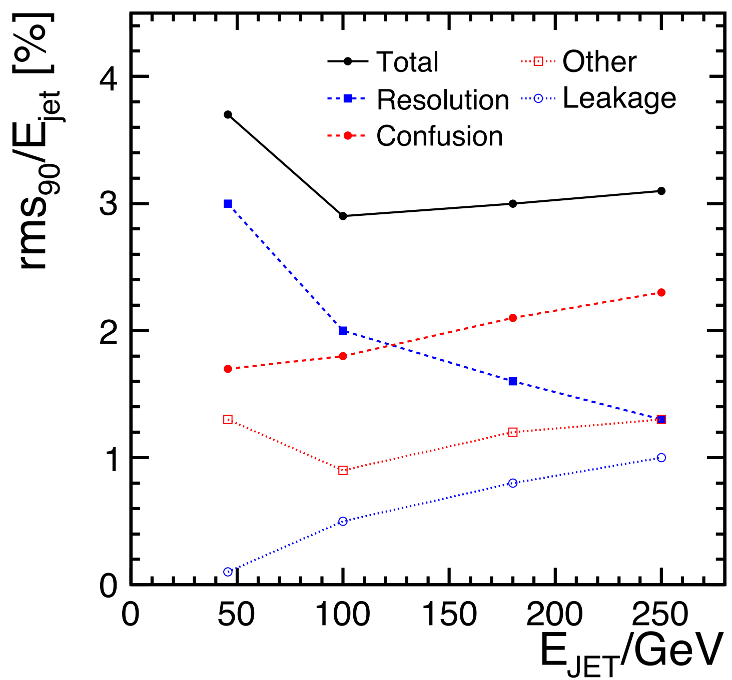
PandoraPFA High Granularity Algorithms

- Developed in the context of the ILD concept detector
  - Ambitious goal of < 4% jet energy resolution
    - To separate boosted hadronic Ws and Zs
    - goal *surpassed*
  - Exploits large TPC with high B-field to separate charged particles from neutrals
  - Exploits high-granularity calorimetry and tracking to tease apart completely overlapping clusters and power robust ID



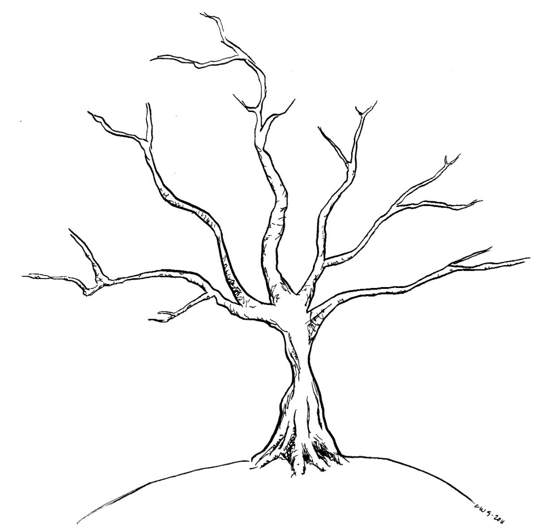


http://arxiv.org/abs/0907.3577

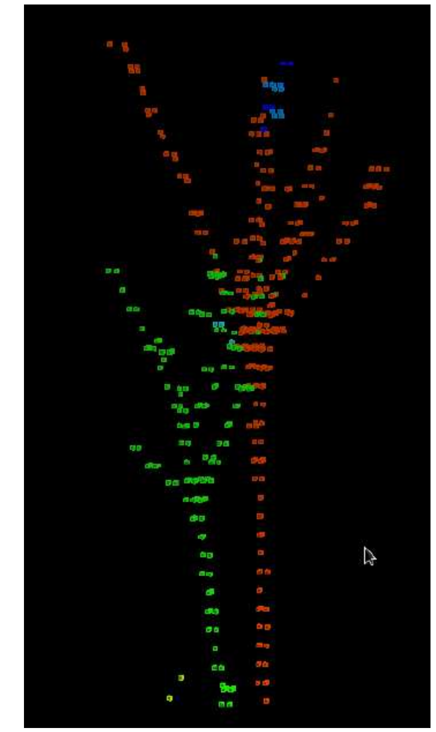


The Arbor Algorithm & Shower Growth

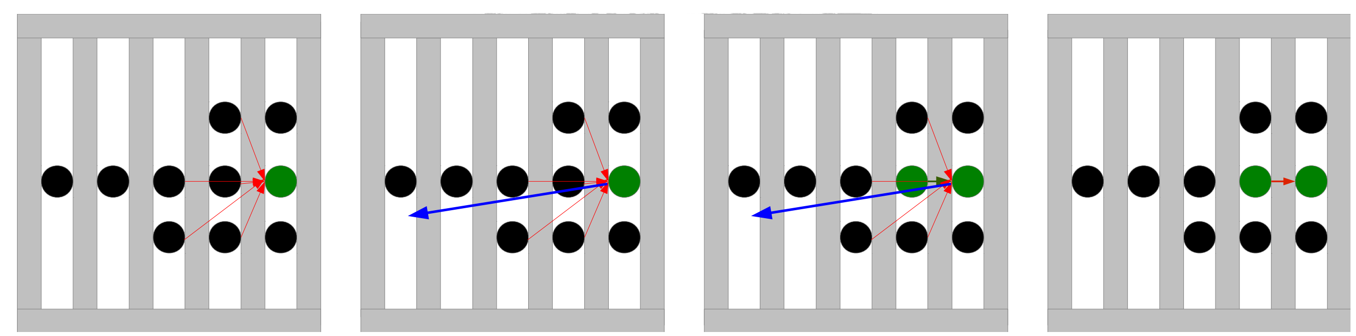
- A new twist on present particle flow concepts
  - Adds a priori knowledge that all showers in a calorimeter originate from a single point
    - i.e. showers look like trees



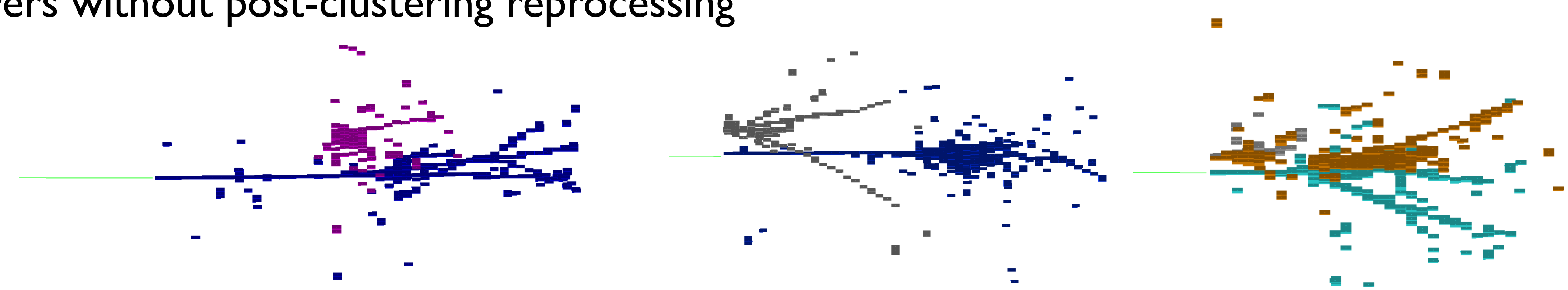
~



- Instead of starting from the calorimeter front…
  - Start from the back, and for each calorimeter hit determine most likely hit before it to connect to until ‘root’ is found
  - Principle can be applied to lower-granularity calorimeters as well, and nuclear interactions within a tracker…



- Allows for reconstruction of very close-by showers without post-clustering reprocessing



10 GeV Neutral Hadron + 30 GeV Charged Hadron, 5cm separation

http://arxiv.org/abs/1403.4784

Lindsey Gray, FNAL

Conclusions

- The success of the Particle Flow concept in describing hadronic environments is apparent
  - The degree of success of a given reconstruction is highly dependent on detector design
  - High-field, large-R tracker based detectors see significant gains in particle reconstruction efficiency and purity
    - Additionally, large gains for global objects like missing energy and jets
    - Due to improved reconstruction of each constituent particle
- Techniques developed with this concept in mind scale well to higher energies and energy-dense environments
  - From LEP to the LHC, the idea scaled with the detector
  - In high pileup and heavy ion collisions, provides benefits or even new techniques!
- Particle Flow is deeply rooted in the development of future detectors and the concept is evolving and being approached from multiple angle
  - Many years of interesting physics and detector research still to come!