

Software Coordinators Report

ILD Software and Analysis Meeting

11.05.22

Frank Gaede, DESY



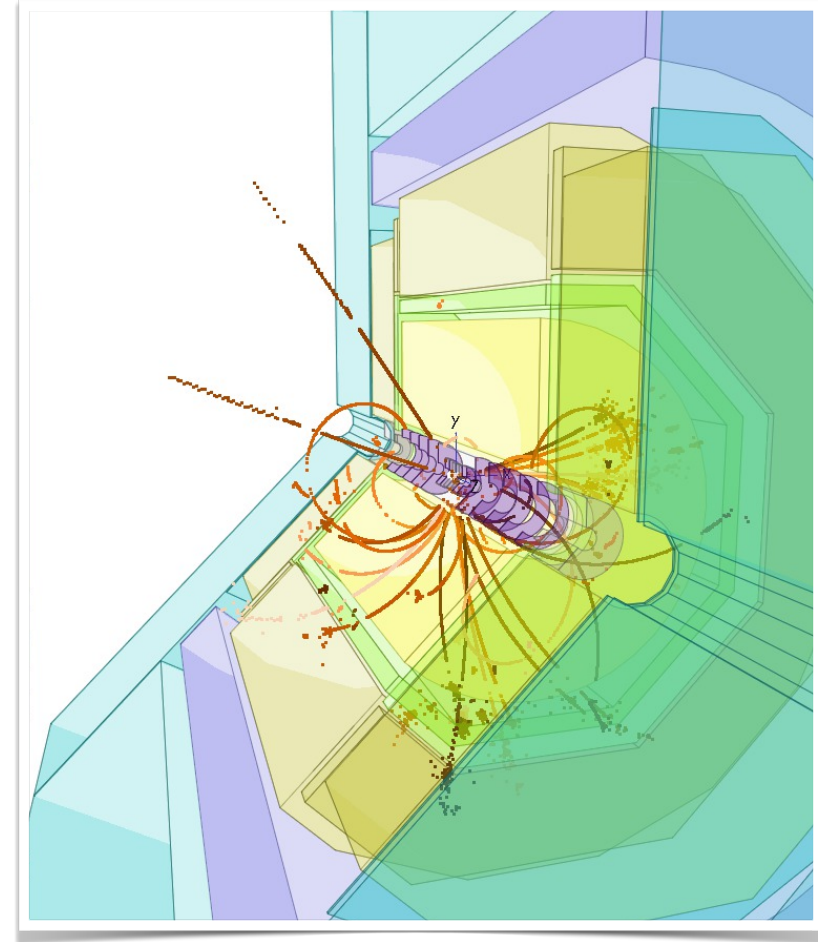
Outline



- Generator
- Simulation
- Reconstruction
- Monte Carlo Production

**some highlights from ECFA Higgs factory
topical meeting at DESY last week**

SW convenors meeting last Friday was cancelled



Generator

J.Tian, M.Berggren

- started to prepare generation of Adrian's request for 2f-hadronic at 500 with the mc2020 chain
 - ~900 files, each with 100k events, i.e. 90 million events
 - 300 MB/file -> total of 300 GB for generated files
 - more for DSTs
- details of exact production version to be used for reconstruction to be discussed

- recently generated new sample at 550 GeV for H-self-coupling study
 - see talk by Julie Torndal today

ECFA Higgs Factories: 1st Topical Meeting on Reconstruction

4./5. may at DESY

Tracking: Tracking I
Convener: Gerardo Ganis (CERN)

09:10 **MarlinTrk and Tracking in ILD**
Speaker: Frank-Dieter Gaede (Deutsches Elektronen-Synchrotron DESY)
gaede_ild_tracking...

09:40 **Conformal Tracking (SiD, CLIC, CLD)**
Speaker: Erica Brondolin (CERN)
20220504_Conform...

10:05 **Special Tracking Gaussian Sum Filter@LHC**
Speaker: Ying An (Peking University (CN))
ECFA_workshop.pdf

Tracking: Tracking II
Convener: Patrizia Azzi (INFN Padova (IT))

11:00 **Tracking for CEPC**
Speaker: Chengdong Fu
tracking_cepcc_ECFA...

11:30 **ACTS**
Speaker: Paul Gessinger (CERN)
2022-05-04-ecfa-hig...

12:00 **Tracking for FCC IDEA**
Speaker: Nicola De Filippis (Politecnico di Torino)
ECFAWorkshop_Tra...

14:55 **PFA or no PFA**
Speaker: Jean-Claude Brient (CERN)
What is PFA 020520...

Particle Flow: Particle Flow I
Convener: Frank-Dieter Gaede (Deutsches Elektronen-Synchrotron DESY)

16:00 **Particle Flow: Pandora**
Speaker: John Stuart Marshall (CERN)
Pandora_ECFA_04.0...

16:30 **Particle Flow with CMS**
Speaker: Kenneth Long (Massachusetts Institute of Technology)
2022_05_04_ECFA_...

Particle Flow: Particle Flow II
Convener: Gerardo Ganis (CERN)

09:00 **Particle Flow: Arbor**
Speakers: , Manqi Ruan (China University of Science and Technology)
Arbor-ECFA 4.pdf

09:30 **Energy Flow in ATLAS**
Speaker: Mark Hodgkinson (CERN)
PFlow_ECFA_ATLAS...

Particle ID: Particle ID II
Convener: Andre Sailer (CERN)

11:00 **Flavour Tagging in ILD**
Speaker: Adrian Irlles (IFIC CSIC/UV)
202205104_ILDFlav...

11:35 **Particle Identification with Gaseous Tracking and Fast Timing**
Speaker: Ulrich Einhaus (CERN)
2022_05_05_ECFA_...

12:00 **Flavour Tagging from CMS to FCC**
Speaker: Loukas Gouskos (CERN)
lg-jettagging-lhc2fc...

- organized by WG2 of the ECFA Higgs study group
- take stock of reconstruction tools and algorithms
 - what is currently used
 - under development
 - common activities

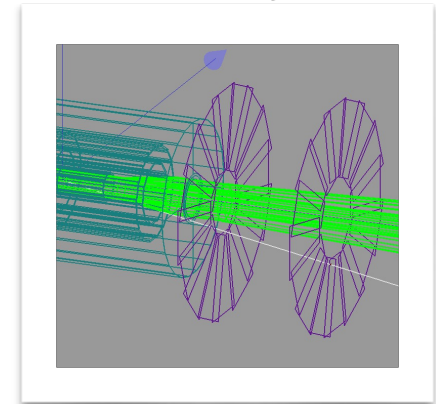
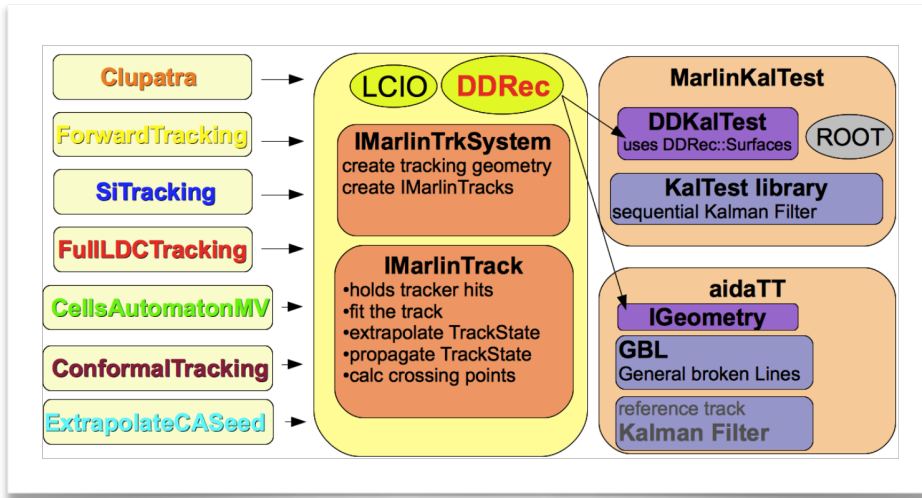
<https://indico.cern.ch/event/1124095/>



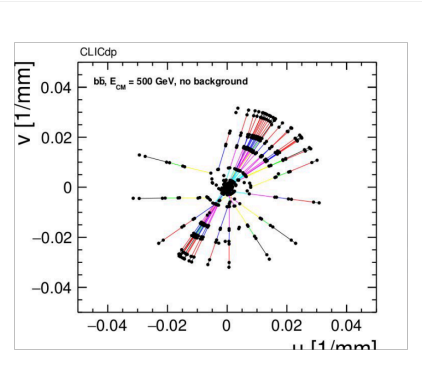
Tracking at LCs

ILD, CLIC, SiD, ...

- large production quality code base for track finding and fitting exists in iLCSoft
- needs to be preserved when moving to Key4hep
- possible w/ MarlinWrapper



plan to add ACTS to MarlinTrk



Proposed e⁺e⁻ collider detectors

E_{CM} up to 3 TeV
3.5 - 5 T solenoids

ILC: SiD

E_{CM} up to 365 GeV
2 - 3 T solenoids

CLIC: CLICdet

FCC-ee: CLD

CEPC: Baseline w/TPC

ILC: ILD

FCC-ee & CEPC: IDEA

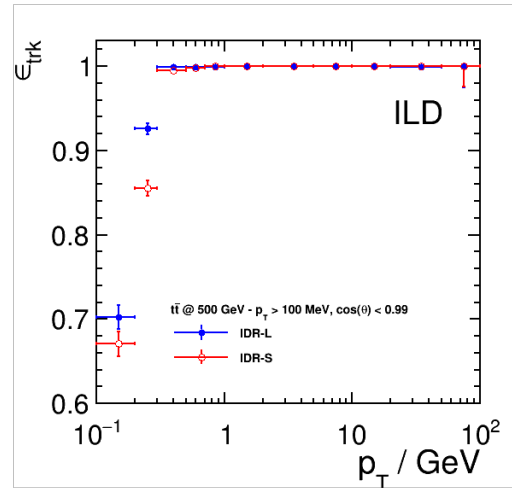
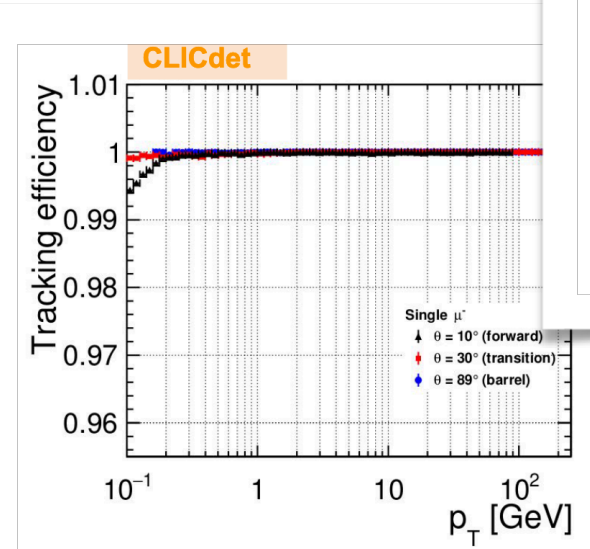
CEPC: w/FST

All-silicon trackers:

- compactness
- validated technology
- stability
- calibration
- redundancy

Conformal Tracking:

- Flexible (different geometries, ...)
- Robust (different beam-backgrounds, ...)



Other Tracking

at LHC, CEPC

C.Fu, Y. An, N. de Filipips



- @LHC use of Gaussian sum filters for e-reco
 - do we really need this at ILD ?
- @CEPC: started to investigate porting of existing parred code to key4hep
- @IDEA: need tracking code for a drift chamber
 - not (yet) available in iLCsoft, MarlinTrk, key4hep

Gaussian sum filter track

ECAL-Driven

1. Mustache SCs are used to extrapolate the collision vertex where there is no bremsstrahlung
2. The first matched hit is used to form a new trajectory starting from the beamspot

Tracker-Driven

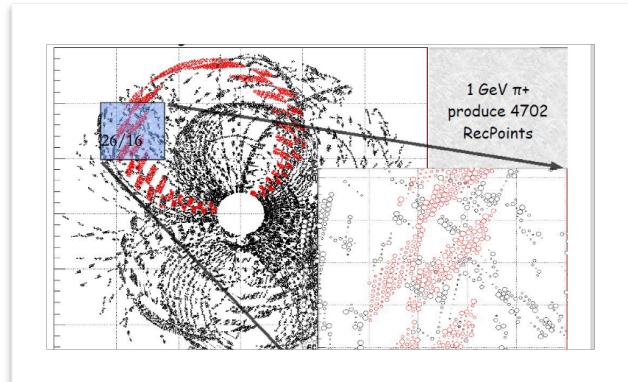
- KF algorithm collects hits if $r_{th} < E/p < 1$ ($r_{th} = 0.65$ or 0.75 for $2 < p_T < 6$ GeV or $p_T \geq 6$ GeV)
- Refit the KF tracks with a small number of hits or a large χ^2_{KF} by GSF
- N_{hits} , the χ^2_{KF} , the χ^2_{GSF} , and the geometrical and energy matching of the ECAL and tracker are used in a multivariate (MVA) method

Complement at low p_T

Track fitting – implementation aspects

What do we need to do?

- pass measurement points with their proper description
 - 3D (2D) point (pixel)
 - 1D point (strip)
- Drift distance



Migration/Implementation

SiliconTracking → ForwardTracking → TrackSubset → Clupatra → FullILDCTracking

- Migrated full tracking chain from Marlin
 - Switch class to GaudiAlgorithm
 - Switch data model to EDM4hep (first realized)

Selected external package → Official upgrade

1st migration → CEPCSW upgrade → 2nd migration

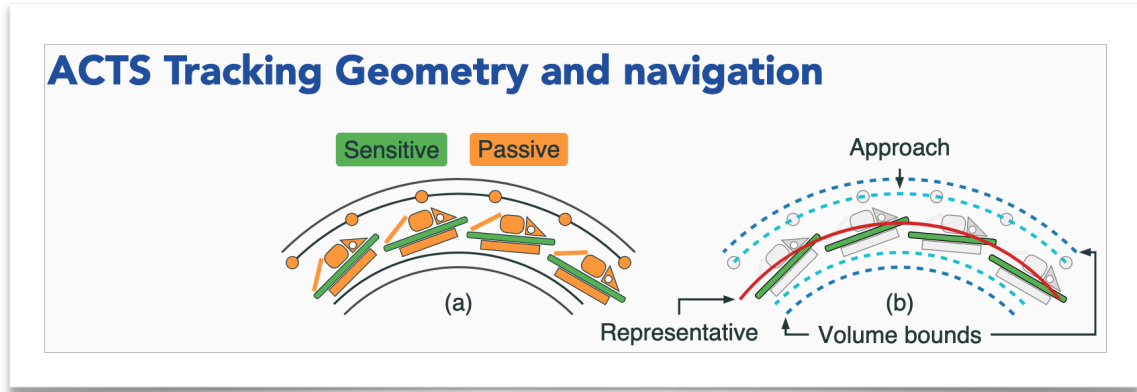
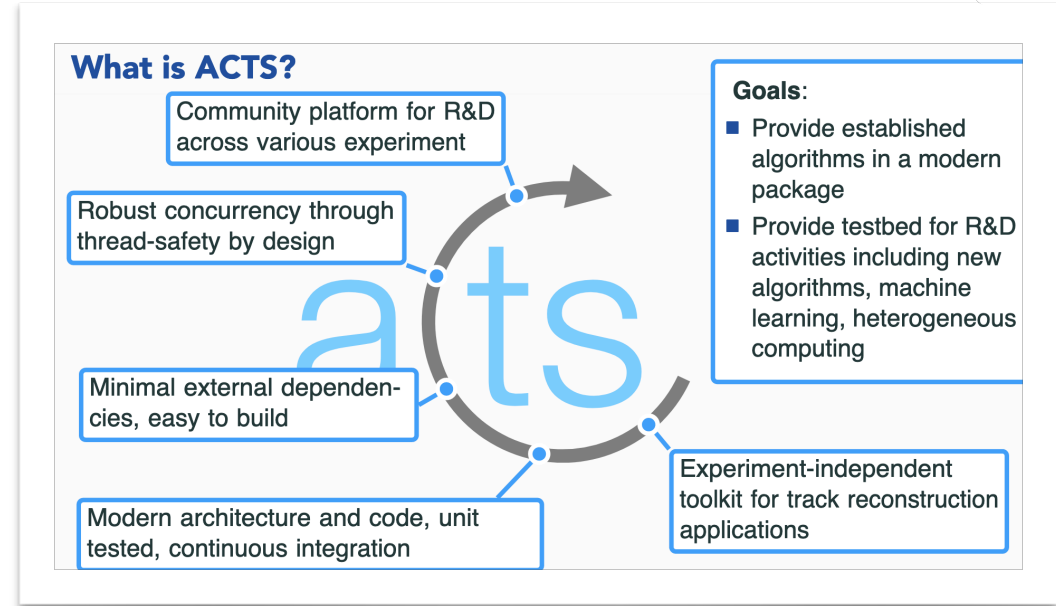
extra merge not friendly

- Another implement way is in considering, to test coversion cost
 - Create a GaudiAlgorithm to covert data model and call
 - ✓ EDM4hep → LCIO → call event loop function → LCIO → EDM4hep
- Key4hep (best) in plan
 - CallingAlg to call prepared API
 - event model support with same code is important

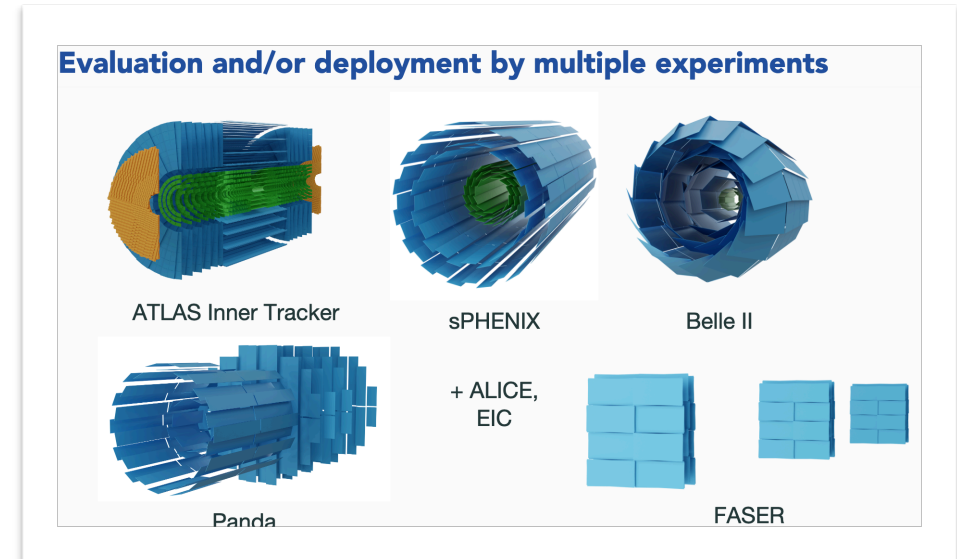
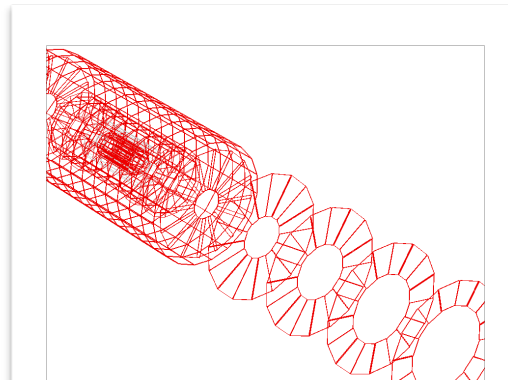
ACTS

the new community tracking toolkit

- based on ATLAS tracking code
- much faster filters and navigation
- need to investigate the use of ACTS in ILD tracking
- biggest question: can we re-use our *ddrec::surface* geometry for this ?



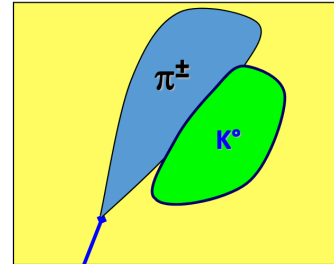
work on DD4hep geometry for ACTS has started...



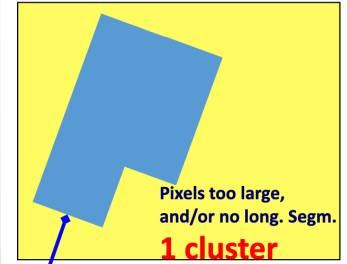
PFA and fast simulation

- reminder that PFA cannot be developed and understood w/ fast parameterised (smearing) simulations such as Delphes
- **confusion in imaging calorimeters is important**
- some attempt mad in SGV to parametrise this...

What is PFA As compared to Eflow



EFLOW Energy balance



PFA Topological analysis



Balance between Cluster Energy and track Momentum. IF $E_{cluster} > P + 3\sigma_E$ (or equivalent using $E_{cluster}$ and P)

The reconstruction of neutral(s)

Imaging calorimeter tells from the image topological analysis **The reconstruction of neutral(s) depends only of imaging capability**

π^0 spoils only neutral Energy estimation

Method of neutral(s) reconstruction in DELPHES

UCL Université catholique de Louvain

DELPHES 3.42.0

Particle-Flow

DELPHES 3.42.0

- Given charged track hitting calorimeter cell:
 - is deposit more compatible with charged only or charged + neutral hypothesis?
 - how to assign momenta to resulting components?
- We have two measurements (E_{cal}, σ_{cal}) and (E_{trk}, σ_{trk})
- Define $E_{neutral} = E_{cal} - E_{trk}$

Algorithm:

- If $E_{neutral} / (\sigma_{cal}^2 + \sigma_{trk}^2) > S$:
→ create PF-neutral particle + PF-track
- Else:
→ create PF-track and rescale momentum by combined calo+trk measurement

- EM (had) deposit 100% in ECAL (HCAL)
- No propagation in calorimeters
- No clustering (topological) clustering, exploiting pre-defined grid

it is NOT PFA

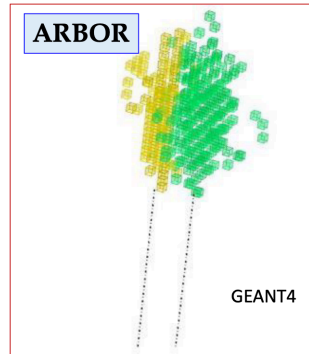
Why it is important :
For DELPHES, the creation/identification of neutral particle(s) will depend of the calorimeter energy resolution (σ_{calo})

It has NOTHING TO DO with PFA where the energy resolution does NOT play a role , at least at the first order *

It introduces spurious correlations between creation and reconstruction

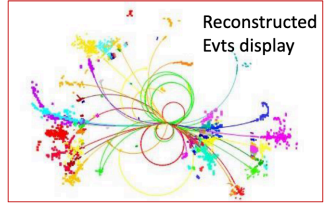
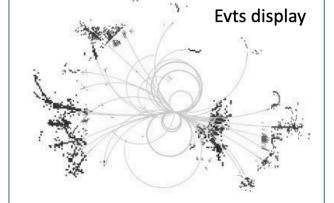
The performances on physics, with a dedicated detector for the Higgs Factory, can't be estimated with DELPHES as it is today

PFA is a software working on the CALO image identification



19GeV π^0 , the 2 photons are well reconstructed by ARBOR
From Yuqiao SHEN, Manqi RUAN, Article in EJPJ-D-20-00564

PANDORA



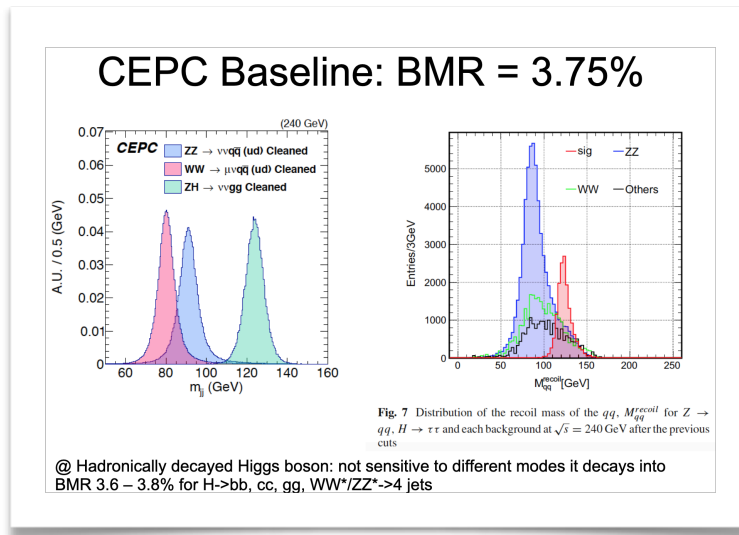
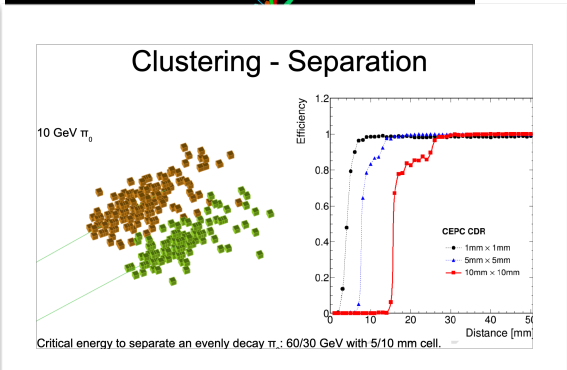
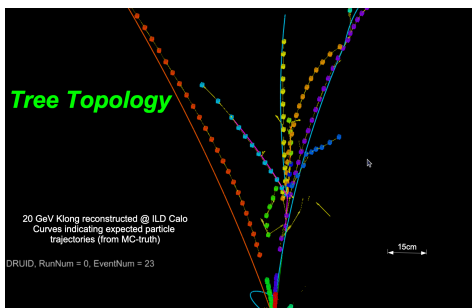
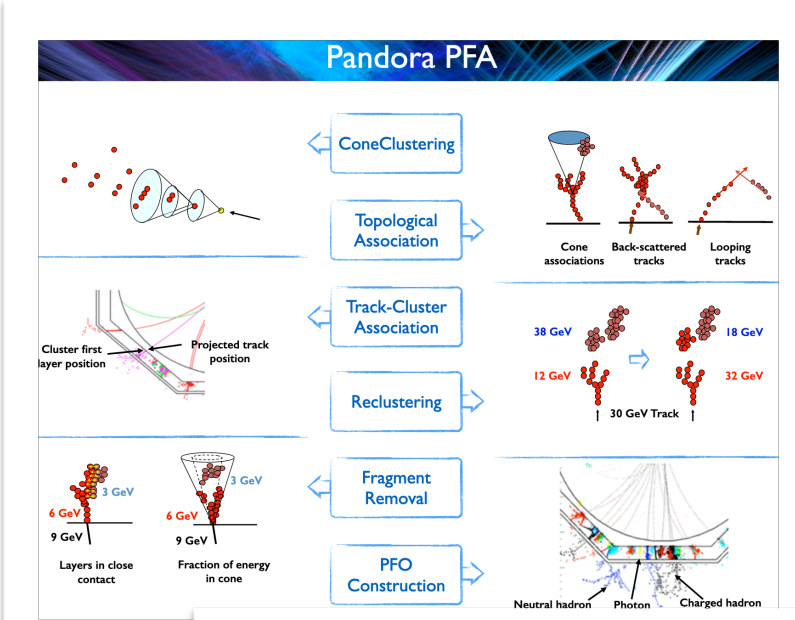
4 jets at CLIC e+e- interaction
The Pandora Software Development Kit for Pattern Recognition
J. S. Marshall, M. A. Thomson, arXiv:150605348



PFA at ILC, CLIC & CEPC

PandoraPFA and Arbor

- linear collider and CEPC have sophisticated PFA algorithms
- set the standard for JER w/ highly granular calorimeters at Higgs factories



Jet Energy Resolution

- Initial motivation for fine granularity particle flow:
Jet energy resolution: $\sigma_E/E < 3.5\%$
- **Benchmark** performance using jet energy resolution in Z decays to light quarks.
- Use total energy to avoid complications of jet finding and no backgrounds included.
- Performance with full GEANT4 simulations:

E_j	$RMS_{90}(E_j) / \text{mean}_{90}(E_j)$
20 GeV	5.4%
45 GeV	3.6%
100 GeV	2.9%
180 GeV	2.9%
250 GeV	2.9%

$\frac{RMS_{90}(E_j)}{\text{Mean}_{90}(E_j)} = \frac{RMS_{90}(E_{ij})}{\text{Mean}_{90}(E_{ij})} \sqrt{2}$ IDR-L
ILD Interim Design Report

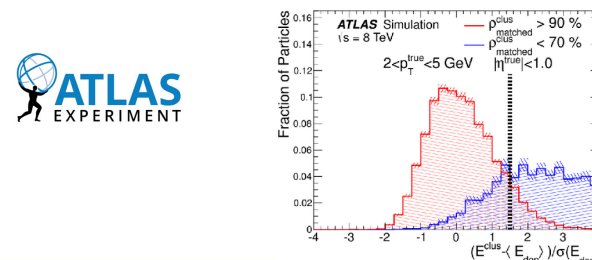
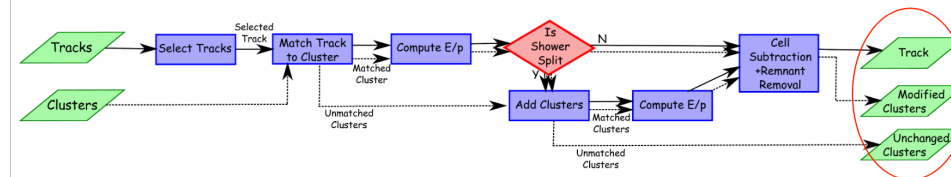


PFA at LHC

ATLAS and CMS

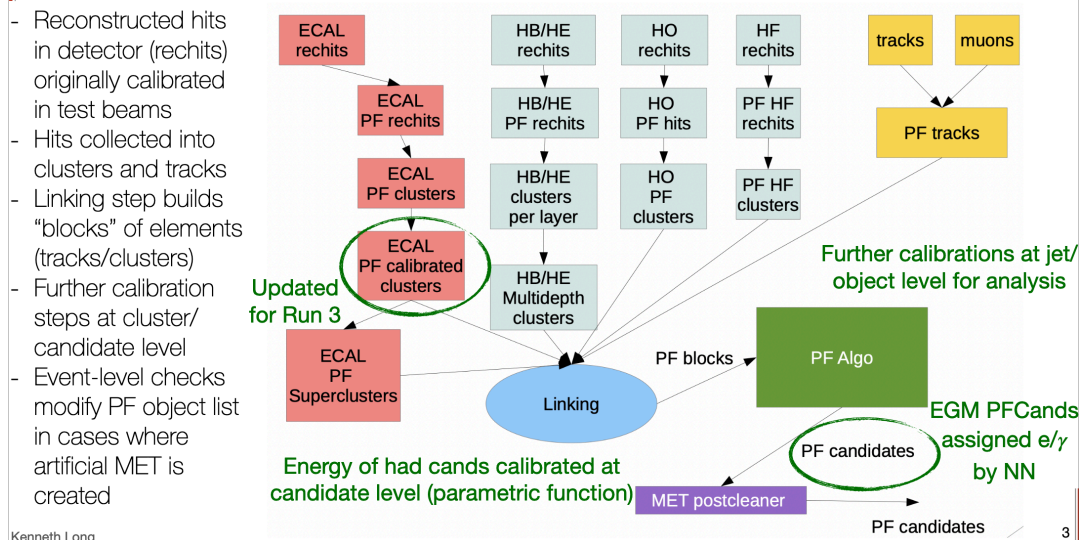
- also at LHC Particle(Energy) Flow algorithms used
- yet much coarser calorimeters (except HGCal)
- not clear that we can benefit from these algorithms

Particle Flow for Jets



► A key part to the improvement in the resolution is the successful removal of remaining noise in the calorimeter

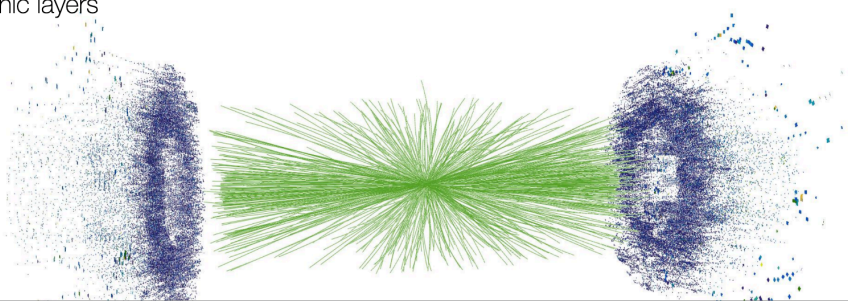
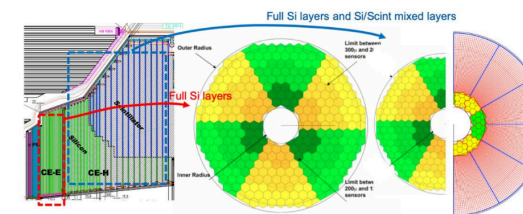
Quick overview of CMS particle flow



Kenneth Long

CMS upgrade and particle flow

- The high granularity calorimeter (HGCal) is one of the ways CMS will **combat the challenges of high pileup**
- Sampling calorimeter with very high longitudinal granularity
- ~6 M hexagonal silicon sensors, ~1 cm²
- ~14 layers with plastic scintillator, ~240k channels
- 28 EM layers + 22 hadronic layers
- "Imaging" calorimeter that presents unique challenges wrt current CMS calorimeters

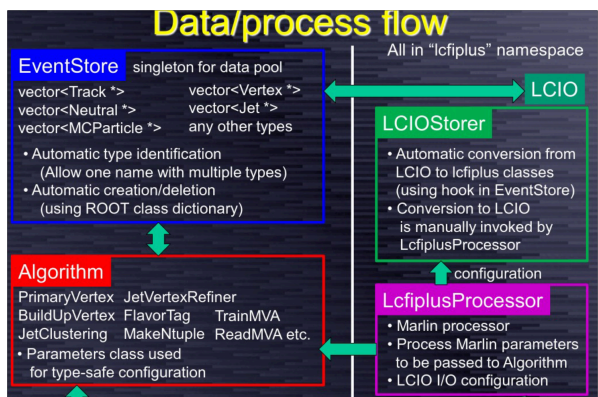




Particle ID

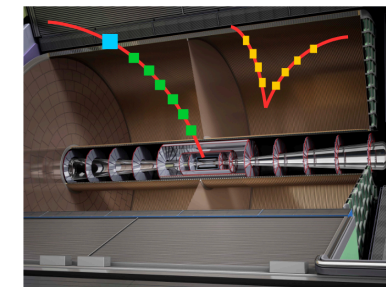
HLR tools developed in ILD

- very nice talks on ILD HLR tools
 - dE/dx and TOF based PID
 - LCFIPlus flavor tagging



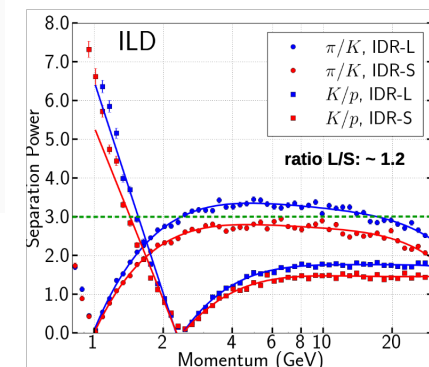
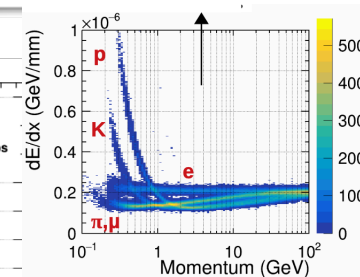
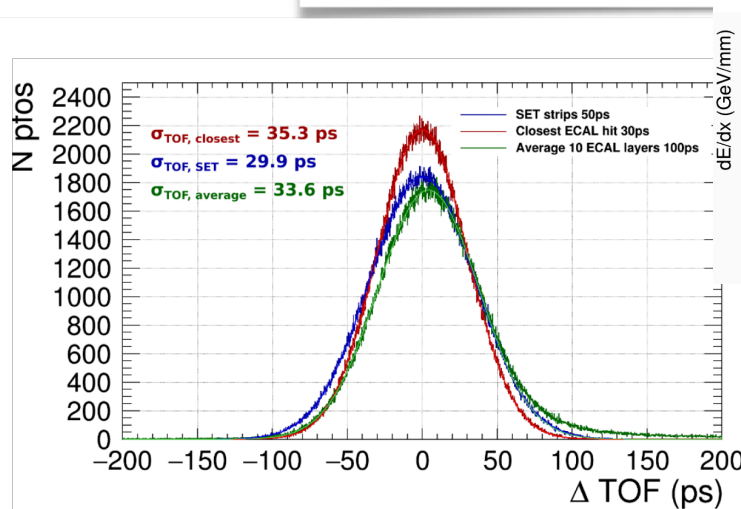
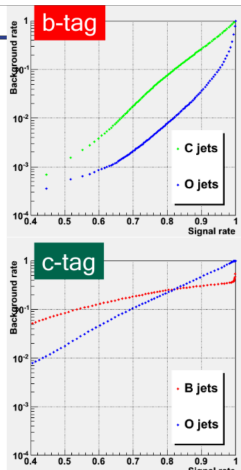
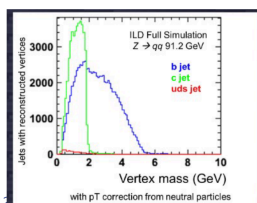
Particle Identification at ILD

- Pandora particle flow: electrons / muons / charged hadrons / photons / neutral hadrons
- **Time projection chamber dE/dx:** (electrons / muons /) pions / kaons / protons
- **Time of flight:** (electrons / muons /) pions / kaons / protons
- **Appearing double tracks = V0 finding:** K_S^0 , Λ^0 , photon conversion
- Low momentum muon/pion separation
- Shower shapes analyser (adds to particle flow)
- π^0 , τ , J/ψ finding, e.a.
- Combined in LikelihoodPID



Flavor tagging

- ▶ Boosted decision trees
 - Number of vertex and type used for categorization
 - Vertex mass
 - Impact parameter
 - ~20 related variables
- ▶ Training performed with dedicated "calibration" samples and with physics samples



Summary

personal impression

- very interesting workshop
 - great to have (some) participants in the room again
- good overview on reconstruction algorithms for future Higgs factories
 - a significant fraction provided by the iLCSoft tools developed in the community over the last decade(s)
- implication for ILD
 - need to continue to work w/ the larger future collider community
 - continue the transition to key4hep while preserving all the existing tools and knowledge
- major next steps:
 - ACTS integration and LCIO->EDM4hep transition

