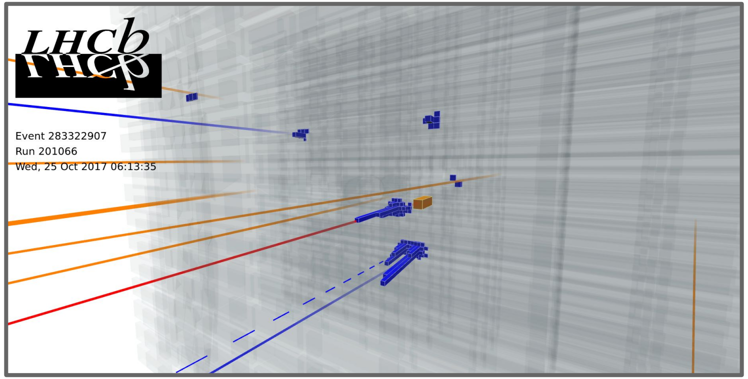
Role of HCAL in current PID and impact of its removal



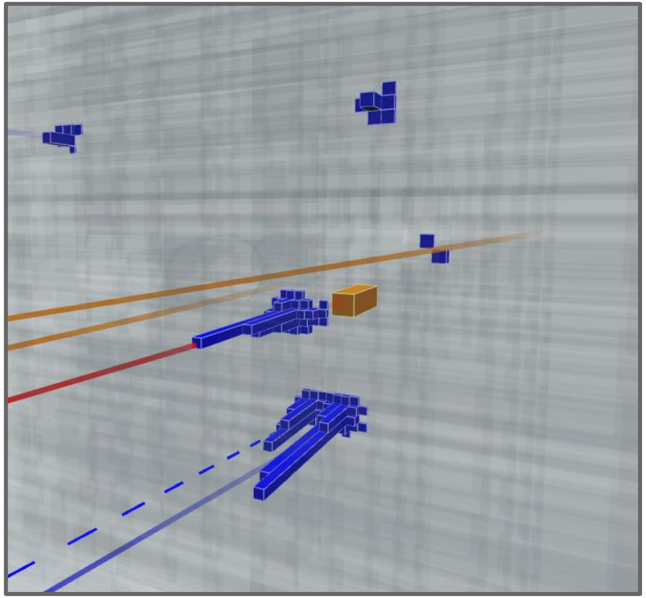
Maarten van Veghel

Nikhef

Charged PID with Hcal

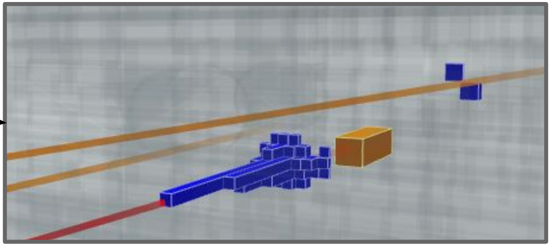
Two types of PID contributions

- **Electron** *hadron* separation
  - Small energy deposit for electron, already lost most of its energy in Ecal
  - ***Will focus on electrons***
- **Muon** *hadron* separation
  - Lack of energy deposit for muon, only small ionisation energy losses
- **Example shown** with (run 2 data)
  - Hadron, a **kaon** (red line), with energy deposits both in Ecal and Hcal
  - **Electron** with two energy deposits in Ecal (track (blue line) + bremsstrahlung (dashed blue line))



Hcal reconstruction (for Run 3)

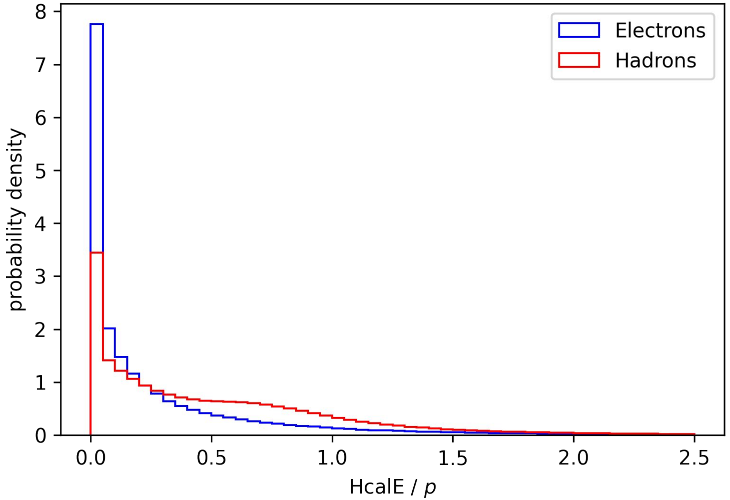
- As *(additional) information for tracks (charged PID)*
  - There is no separate / independent clustering for the Hcal!
- One algorithm, **one output**, for the whole Hcal
  - Sum of energies of cells intersecting track extrapolation (line)
    - Referred to as *CaloHcalE* in *ProtoParticle*
    - Typically matches just **one cell,** due to Hcal cell sizes
  - **Simple but very effective!**
    - *The Ecal version gave us the most performant electron-hadron PID feature in Run 1/2*

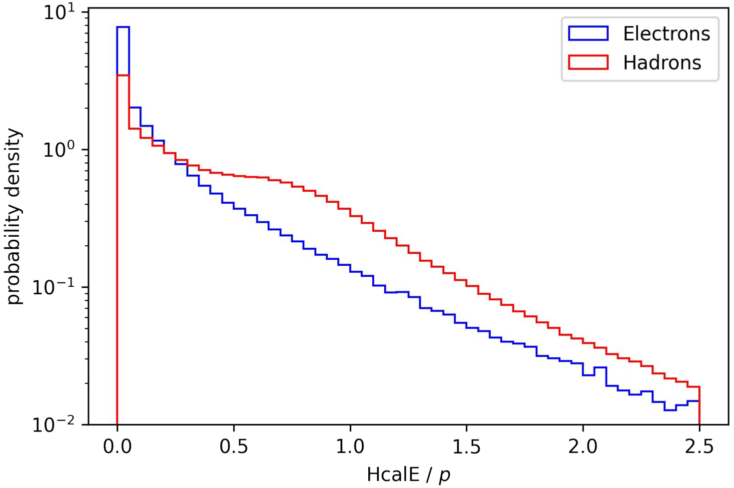


```
// loop over input tracks
for ( auto const& trackincalo : tracksincalo.scalar() ) {
  auto track     = trackincalo.from();
  auto ref_state = track.state( state_loc.value() );
  if ( !propagateToCalo( calo_state, ref_state, calo_front ) ) continue;
  float energy = getEnergy( calo_state, digits, calo, calo_zsize, m_nplanes );
  // save result for this index in tracks
  output_table.add( track, energy );
  // statistics
  m_energy += energy;
}
```

How do Hcal energies typically look like?

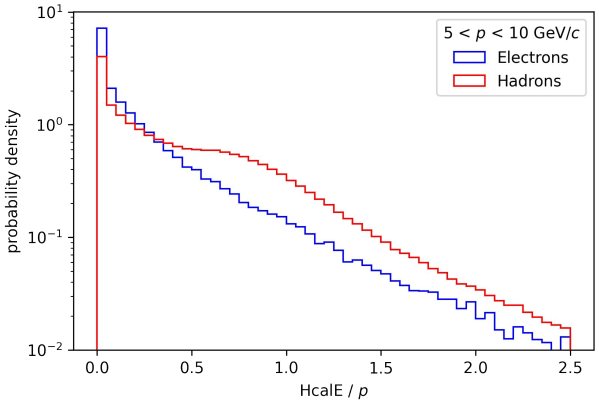
- ***Normalized energies, to track momentum***
  - *as it is highly correlated, and we don't want to select on momenta, but PID features*
- Electrons clearly deposit less than hadrons
- **Likelihood ratios of about 2 to 5**
  - Examples shown
    - from all tracks in *B→J/ψ(→ee)K* simulation; same plot: linear (left), log (right)

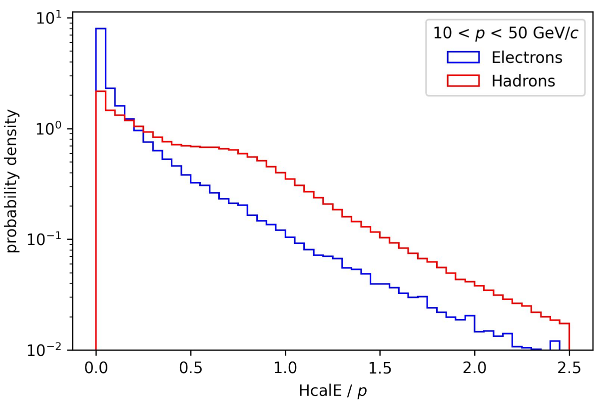


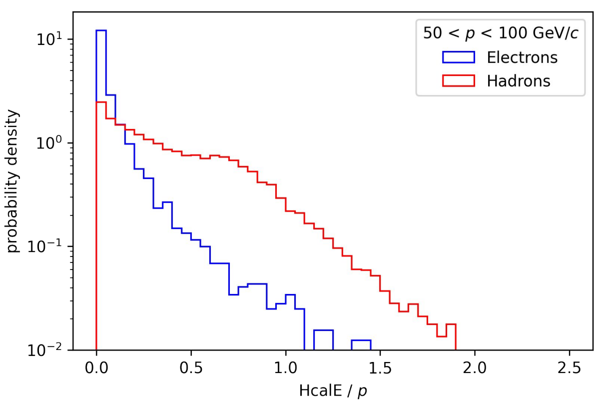


Hcal PID at different momenta

- **At higher momenta, performance increases**
    - At > 50 GeV/c up to order of magnitude false positive rate reduction!
- Suggesting overlap with other deposits mostly low momentum

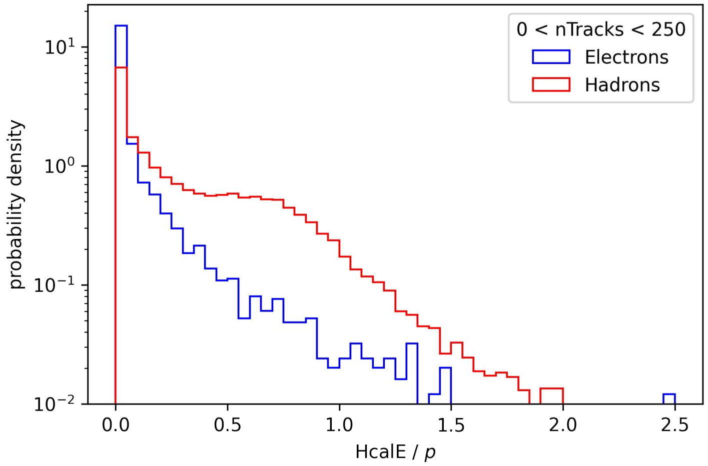


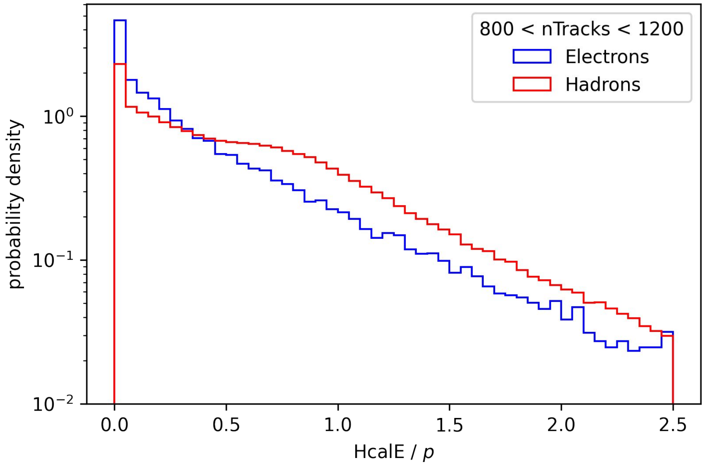




How does Hcal PID depend on occupancy?

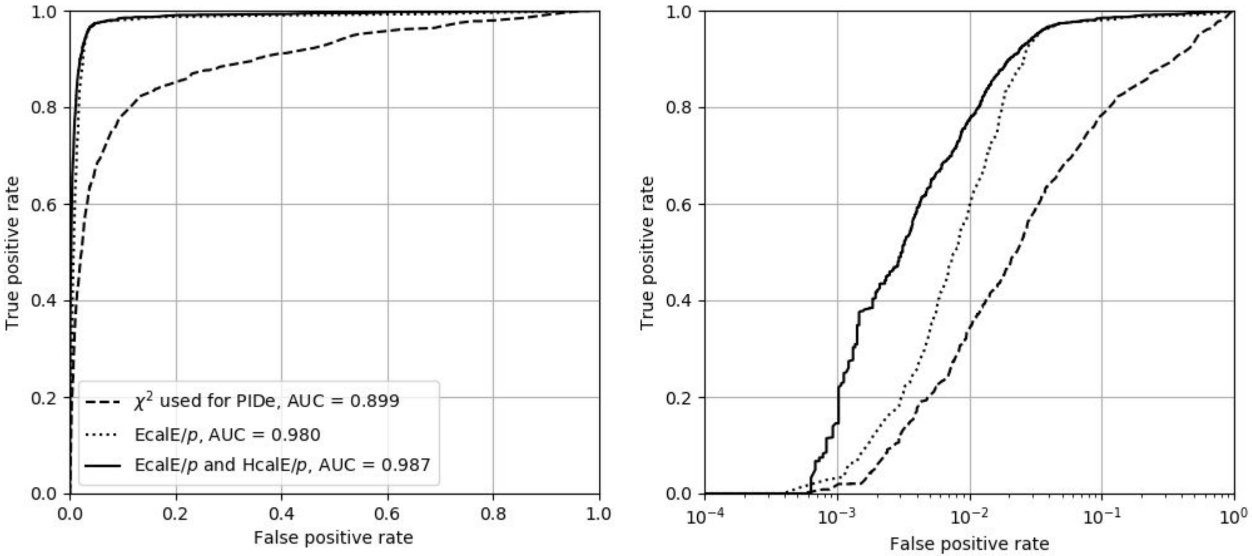
- Consistent picture as with momentum
  - Low momentum deposit overlap diluting performance
- Suggesting the **PID performance** scales with
  - (track) **momentum times inverse of occupancy**





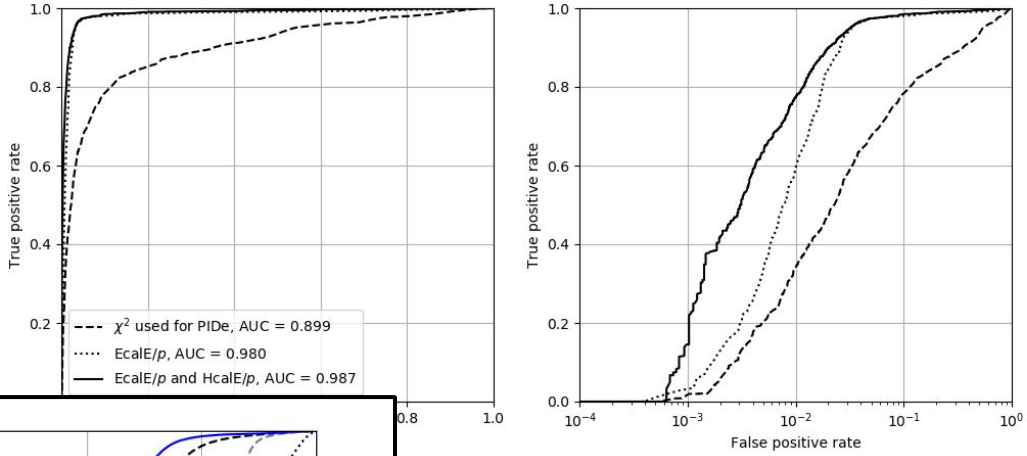
Hcal PID for electrons in comparison to Ecal

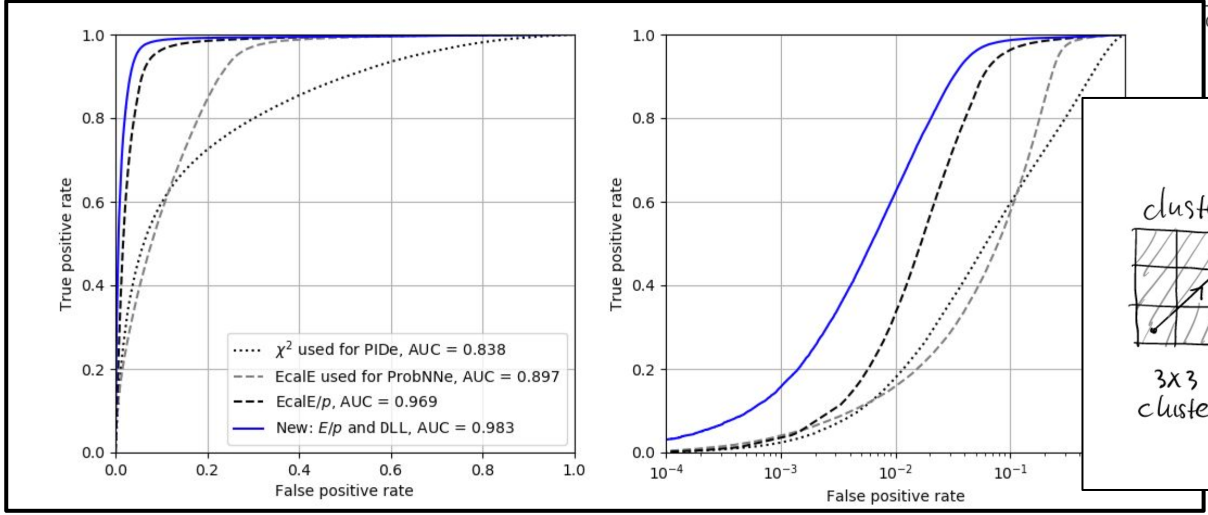
- **Correlations** with other PID **not taken into account with DLL sums**
  - Better to see / check it combined, using ML
    - In *GradientBoostingClassifier* from *Sklearn*, essentially electron versus pion
- Given a **factor 2-5 reduction in false positive rate**, consistent with what we saw earlier (note this is mostly low momentum)

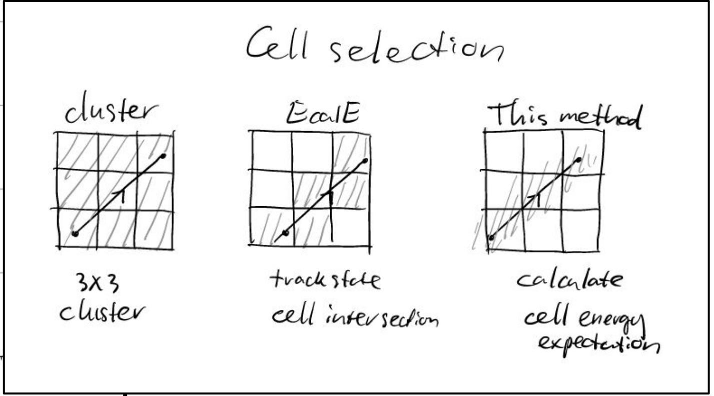


Hcal PID w.r.t better Ecal occupancy handling

- **Ecal PIDe is using new cell selection method**
  - More in backup
- ***More is gained*** *by* ***improving occupancy handling in Ecal*** *than adding Hcal info!*



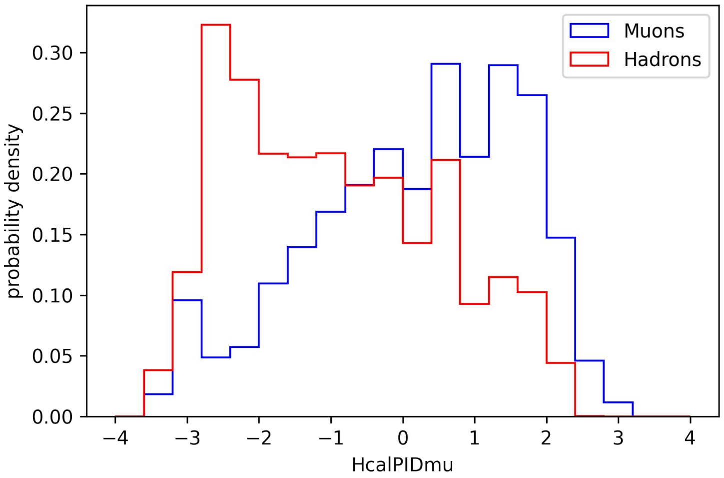


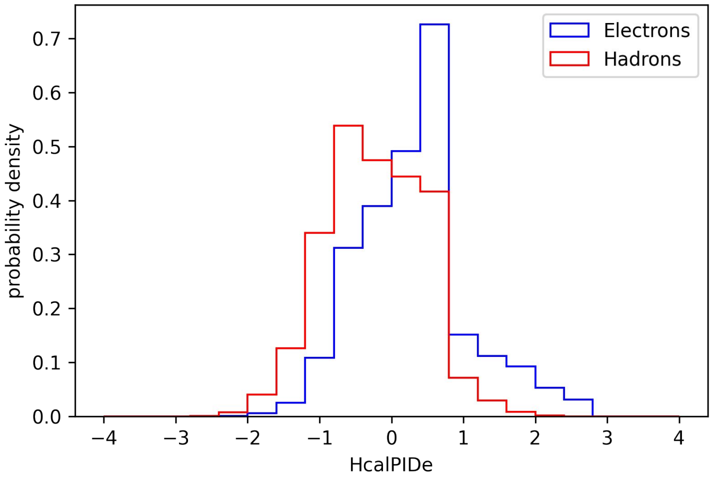


And muon PID?

- Also based on *HcalE / p*
- **Muon Hcal PID tends to be a bit more performant than for electrons**
  - Muons deposit even less energy than electrons
- Same occupancy and momentum dependencies / issues, very similar to electrons

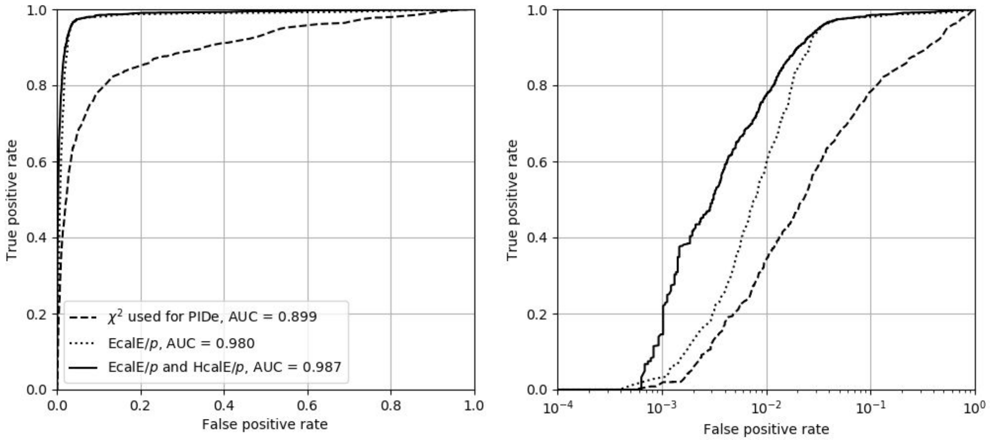
- **Just a thought / question**, but also for muons, **higher granularity in muon systems is better?** (than putting that money in Hcal?), especially considering decay in flight (kink detecting)?





Summary

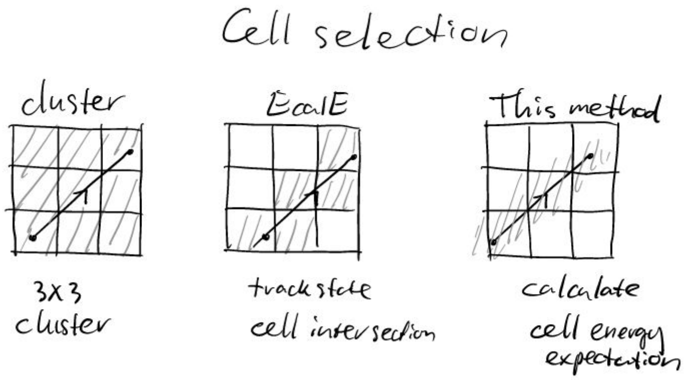
- **Charged PID** from Hcal in the form of Hcal energy over momentum
  - Energy determined from cells intersecting track extrapolation
- **Both electron and muon PID** (with respect to hadrons)
- Performance scaling with momentum and inverse of occupancy
  - **Higher luminosity clearly decreasing performance**
- **Hcal adding typically 2 - 5 false positive rate reduction**
  - **Can be overcome with better granularity (treatment) in Ecal** (?)

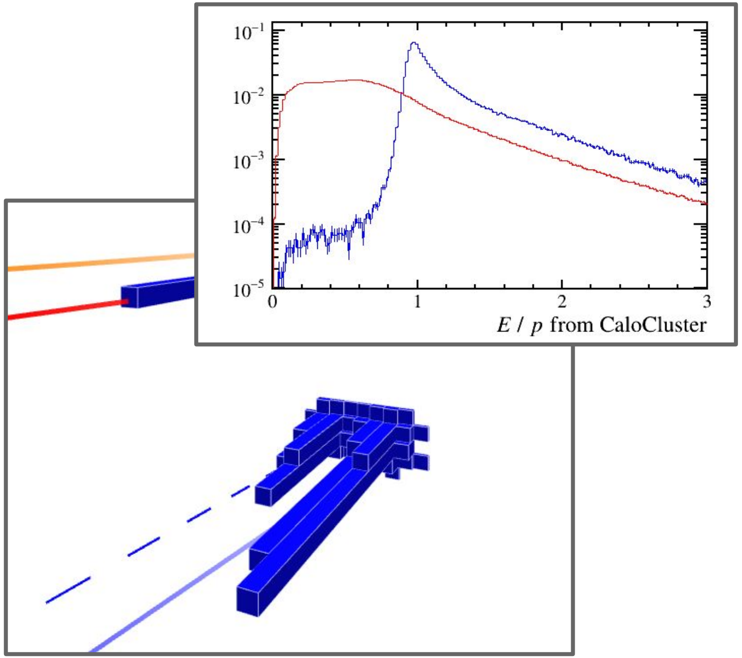


Backup

Electron PID from electron Ecal shower

- PID for electrons based on energy deposit in cells directly related to track (not brem)
- What options did we have (Run 1/2)?
  - 3x3 cluster (track-cluster matching)
  - EcalE method: track state - cell intersection
- New, main Run 3 method
  - cell selection based on energy expectation per cell

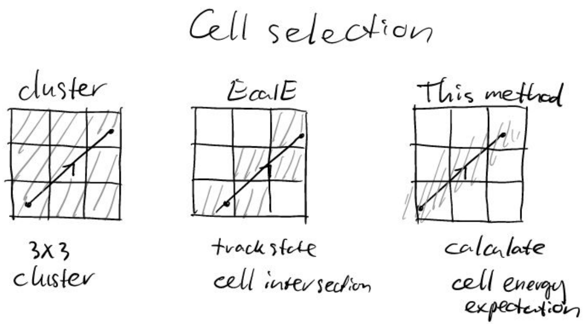


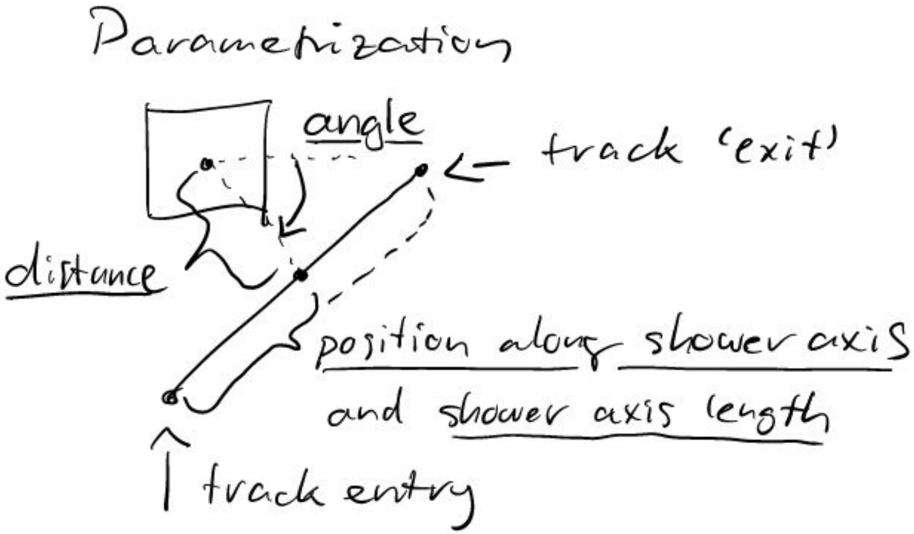


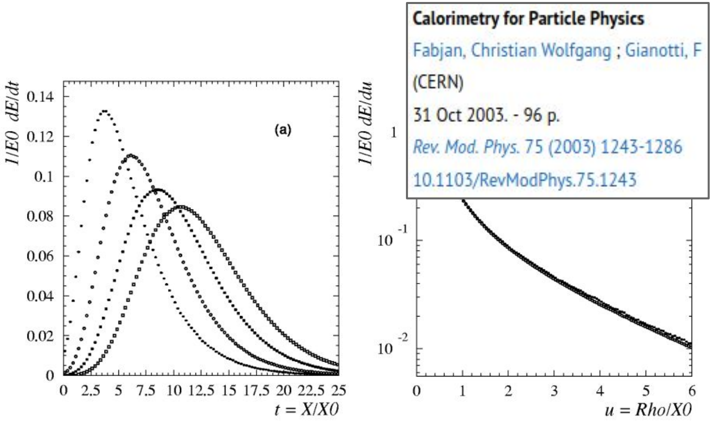
- Can we be **more selective in the cells we select?** And can we extract more per cell information?

How to estimate cell energy?

- Use first principle electron shower profiles
- Generate showers with Monte Carlo with said distributions
- Parametrize results based on track and cell parameters

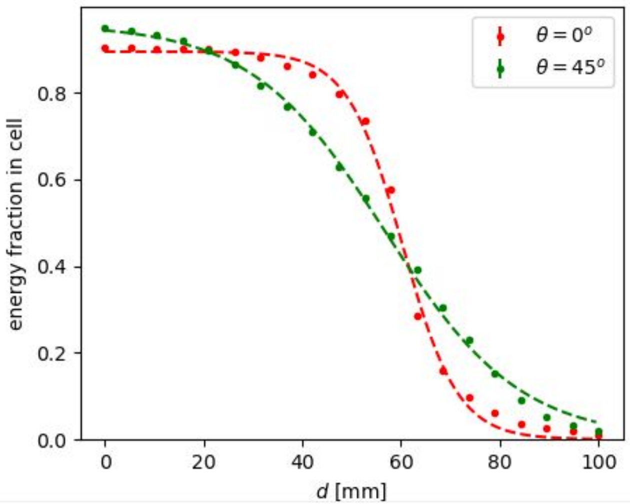


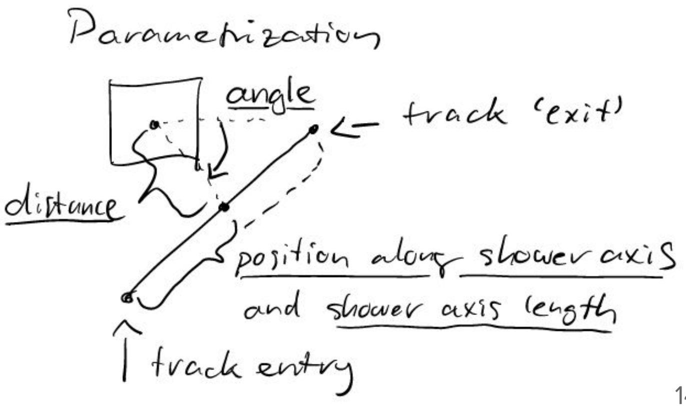




Parametrizing cell energy

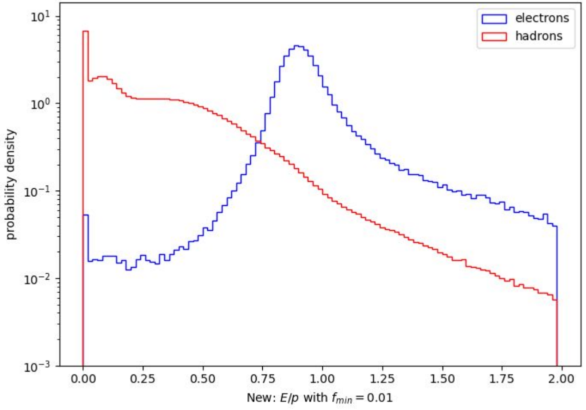
- **Shower parameters**
  - closest distance to shower axis (d)
  - angle of closest distance vector in xy plane (theta)
  - length in xy plane of shower axis (txy = sqrt(tx^2+ty^2))
  - position along shower axis (lbar)
- **Distance strongest effect**, **parametrize by sigmoid** / logistic function
- Bin sigmoid parameters in other shower parameters (lookup table, basically)
- Store in TH3, so can also do trilinear interpolation

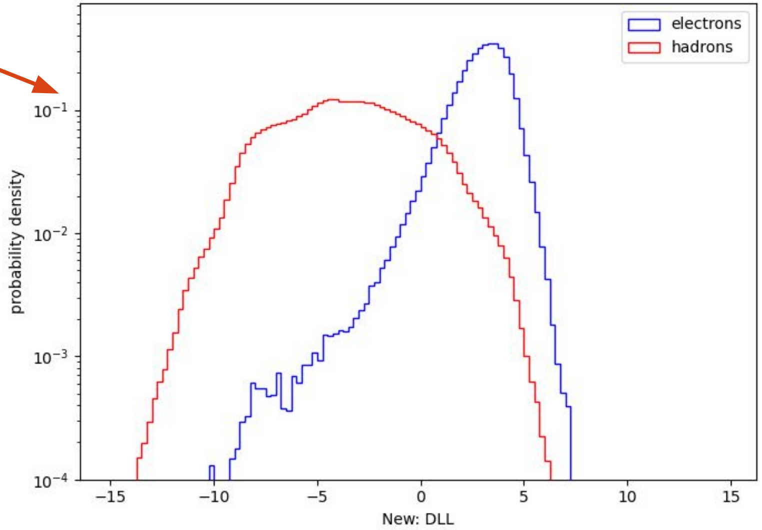


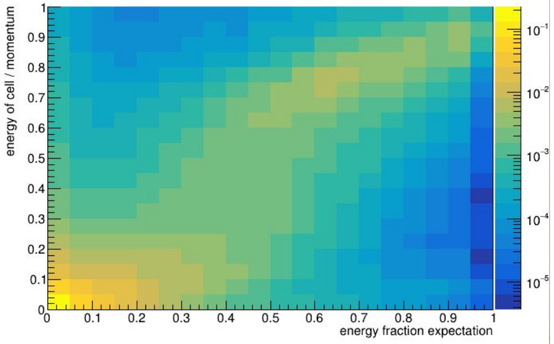


Using energy expectation for PID

- Build more **selective** (in cell choice) **E/p**
  - Total energy of cells with minimum of energy / momentum fraction of 10%
- Can we squeeze more info out of it per cell?
  - **Construct likelihood ratio: summed, per cell, delta-log-likelihood (DLL)**
  - DLL parametrized/conditional per:
    - expected energy fraction
    - momentum
  - Parametrization based on LHCb full MC







Performance

- In *GradientBoostingClassifier* from *Sklearn*, essentially electron versus pion (and a bit of kaon)
- **Both new variables individually outperform current best variable** (EcalE/p)
- Both contain different information as well, **combination of E/p and DLL works well**
- **Reduction of false positives with a factor of about 2 - 5** (w.r.t EcalE/p)