Particle Flow & DIS Electron Finding in ePIC

Derek Anderson &
Daniel Brandenburg (speaker)

ePIC Collaboration Meeting

Warsaw, Poland 07/26 – 07/29

Outline: What's In this talk?

- Particle Flow (Squad Leader: Derek Anderson)
  - What is PF?
  - Charge & Goals
  - Survey of PF in other Exps
  - Plan & Next Steps

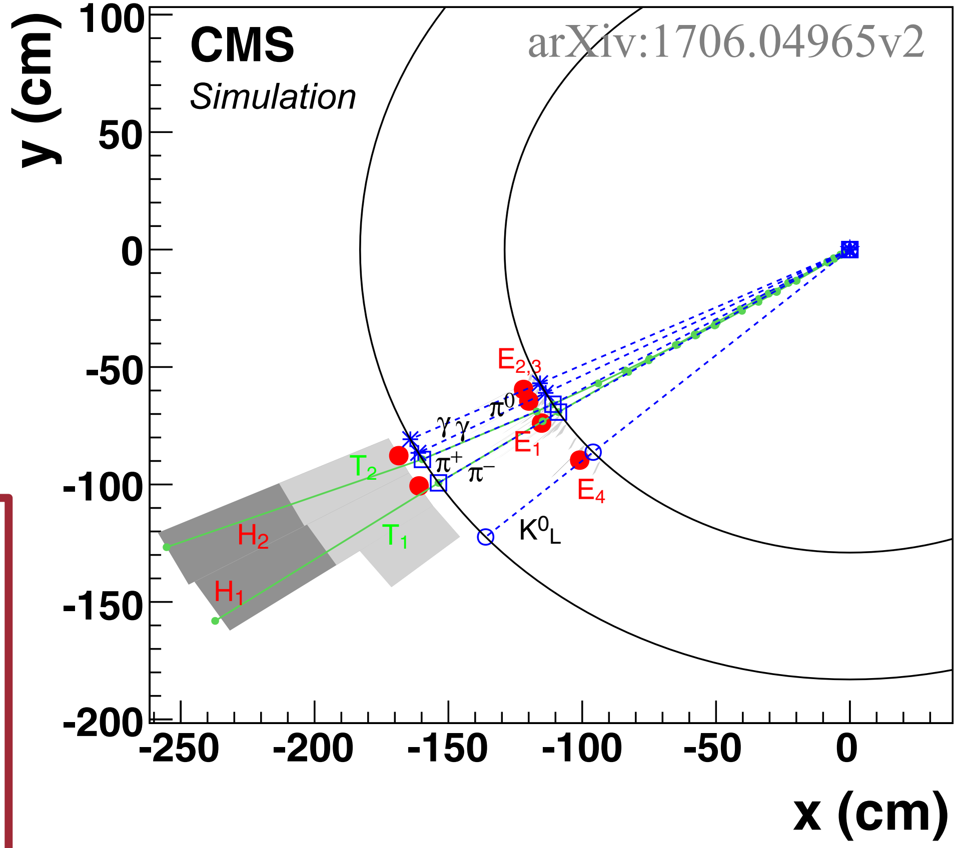
- **NB: Physics Analysis and C/S Coordinators identified 4 priorities for reconstruction software:**
  - Vertexing and PID
  - Low Q2 Tagger
  - **Electron Finder**
  - **Particle Flow**

- DIS Electron Finder (Squad Leader: JDB)
  - Charge & Goals
  - Electron Finder Progress
  - Next steps

Particle Flow 101

PF is a *holistic* reconstruction approach based on physics objects (vs. detector info)

- Combines info from sub-detectors
- Simplifies analysis, at cost of reconstruction complexity



Generic PF Recipe (simplified)

1. Local Reco (CALO Clusters, Tracking, etc.)
2. "Link" info from various sub-systems
3. Identify "particles" from linked info
4. Post-processing & cleaning

Particle Flow task squad: Charge & Goals

- PF Squad Charge:
  - *"improve jet reconstruction using particle flow information"*

- **2 distinct regions for PF at ePIC**
  - **Barrel/Backward:** JER set by tracker + EMCal
    - ☞ Need tracks to deconvolve clusters for neutrals
  - **Forward:** JER can be improved by combining track + calorimeter information
    - ☞ Need to separate overlapping clusters

- **2 initial tasks:**
  - 👍 Survey existing implementations
  - ☞ Explore necessity of custom approach in barrel/backwards

Particle Flow: Survey

Input from several collaborators with experience at other experiments (Brian P, Mathew N, Antonio S. etc.)

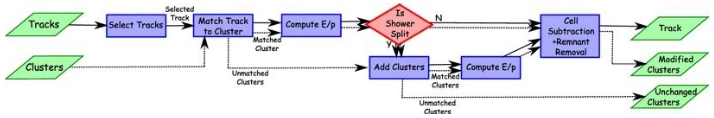
- ALEPH and DELPHI implemented the first "modern" PF algorithms
  - ☞ Both make use of PID capabilities

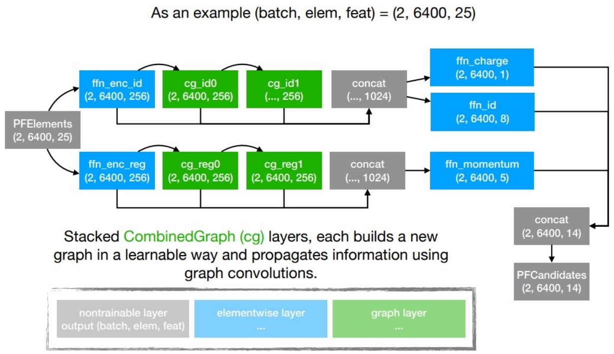
An ML alternative to existing algorithm is being explored at CMS

- Ref.s:
  - › EPJC **81**, 381 (2021)
  - › JP:CS **2438**, 012100 (2023)
- Could ease computational requirements

PandoraPFA: a *very* sophisticated PF algorithm for high granularity calorimeters

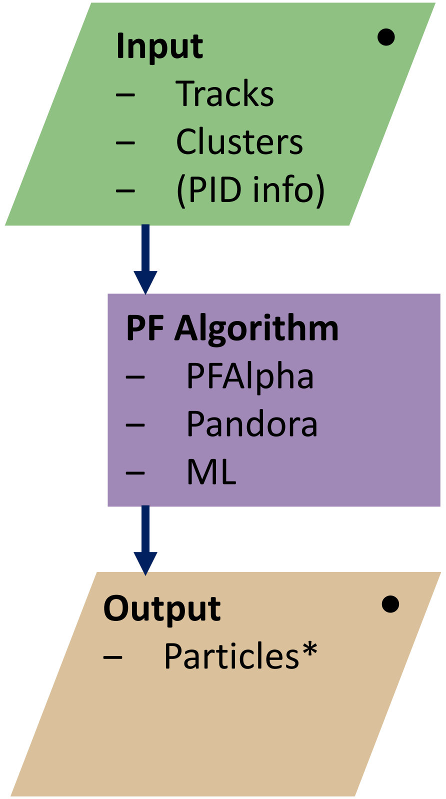
- Part of AIDAsoft
- Has produced an extensive detector-agnostic implementation
- Currently deployed at MicroBooNE





- PF Used in many Experiments(links below):
  - CELLO (PETRA)
  - ALEPH (LEP)
  - DELPHI (LEP)
  - H1 (HERA)
  - D0 (Tevatron)
  - CDF (Tevatron)
  - Pandora (ILC/CLIC, MicroBooNE)
  - ATLAS (LHC)
  - CMS (LHC)
  - sPHENIX (RHIC)

Particle Flow Discussion | Algorithm Strawman



* = ReconstructedParticle object
● = fixed constraint

- Infrastructure to do "bare-bones" PF *largely* exists
  - ⇒ Could implement a basic "alpha" **(PFAlpha)** algorithm
- **Rationale:**
  - Motivate and test development of necessary software
  - Serve as baseline to compare refinements against
  - Allow analyzers to quickly start working w/ output
- Development then proceeds with testing more refined approaches, e.g.
  - Such as PandoraPFA
  - ML-based models
  - Etc.

Particle Flow Discussion | Algorithm Strawman

PFAlpha:

1) Project tracks through calos
2) Associate all calo clusters within a cone of size R around the track
3) Sum all calo energy in cone and subtract expected track energy from sum
4) **Return**
   - Tracks
   - Subtracted clusters
   - Unassociated clusters

- Infrastructure to do "bare-bones" PF *largely* exists
  - ⇒ Could implement a basic "alpha" **(PFAlpha)** algorithm
- **Rationale:**
  - Motivate and test development of necessary software
  - Serve as baseline to compare refinements against
  - Allow analyzers to quickly start working w/ output
- Development then proceeds with testing more refined approaches, e.g.
  - Such as PandoraPFA
  - ML-based models
  - Etc.

Particle Flow Discussion | To-Do

Missing Infrastructure (Major):

- PF Framework
  - Factories
  - Algorithm + configuration files
- Improved track-cluster associator
  - Extend to include Hcals
  - However, truth-based implementation may work for interim

Missing Infrastructure (Minor):

- PFObject Visualizer:
  - Plugin (or service?) to visualize clusters, tracks, etc.
  - Crucial for debugging
- Downstream analysis:
  - Code to look at impact of changes
  - Existing jet benchmarks are good starting place

Open Questions:

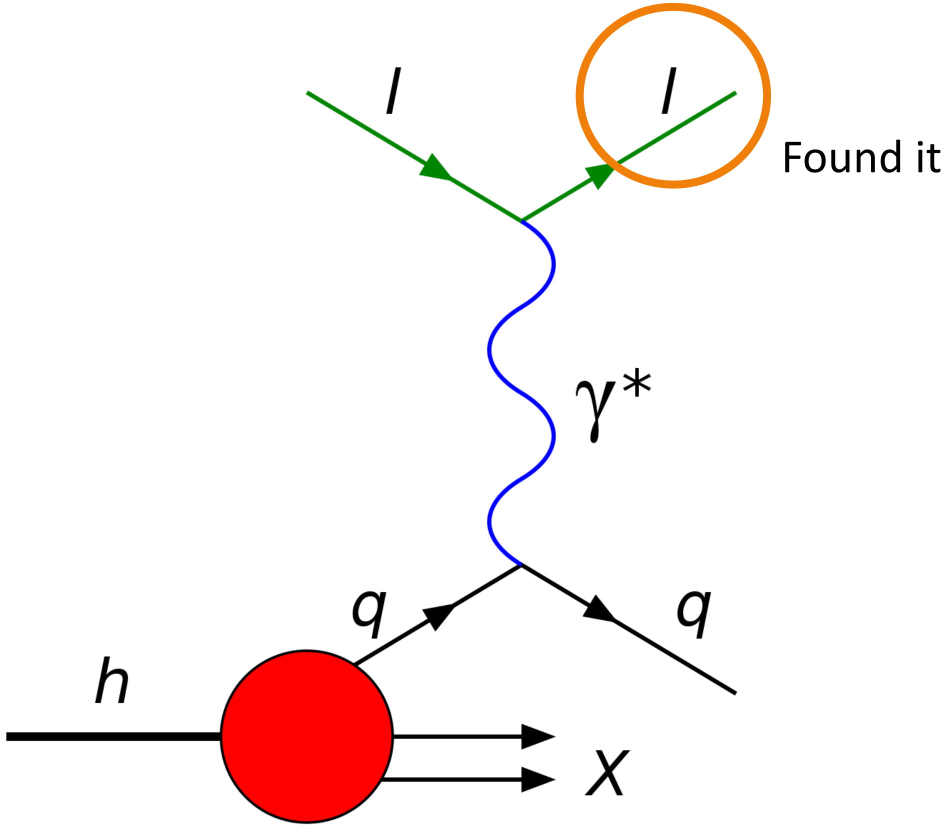
- Does implemented cluster splitting work in non-enabled* detectors?
- How well do existing MC-cluster associations work?
  - Currently handled by MatchClusters algorithm
  - Would a separate MC-cluster associator be better?

Major = necessary for implementation
Minor = can be pursued in parallel with implementation
Yellow = connection with other groups
* = existing implementation enabled for central ECals and ECalLumiSpec (not enabled for Imaging/SciFi)

DIS Electron Finding

DIS Electron Finder

- Charge: *Developing an efficient and accurate algorithm for identifying electrons and identifying the scattered electron of the DIS process*



Major Goals:
- Develop unified electron identification
- Implement DIS lepton finder algorithms

Realistic DIS lepton finding is crucial for many benchmarks and analyses needed to inform detector design

DIS Electron Finder | Approaches

- Taking two “parallel” approaches

“Truth”

Use truth associations

- Particle-to-Cluster associations utilize “truth” information
- Select electron using E/p utilize momentum from tracking and energy from reco Cluster

- + Similarly “truth” informed DIS lepton finder

“Reco”

- Track projections to Calorimeters
- Track to Cluster matching
- Electron identification utilizing full PID capabilities
- …

- + Realistic DIS lepton finder

DIS Electron Finder | Progress Timeline

✓ Truth level Particle <-> Cluster associations (S/C team, esp. Wouter)
- PR #666 (merged on June 24)

✓ Electron Identification framework strawman
- Provides "ReconstructedElectron" collection
- Currently implements E/p cuts utilizing ECAL info
- Uses "Truth" associations
- PR #751 (merged July 8th)

⇒ Track Projection Factory (Tyler Kutz)
- Provides track projections to common surfaces (each ECAL / HCAL etc.)
- EDM4EIC updates to accommodate projections
- PR in preparation

⇒ Track-to-cluster matching (some prototype work from Nicho Schmidt)
- Provides processor for matching (see PR#606)
- Still significant work needed to utilize projections + implement as factory

DIS Lepton Finder Algorithms

- Goal: identify the DIS lepton using only final state information

In case of one electron anything will work. In case of multiple electrons even a simple rule, such as take electon with "maximal" energy/*pz* /etc. will work for 50+% cases if only 2 electrons are present. Basically, implemented in Rivet.

Challenges:

- How to select **the** DIS electron when there are multiple candidates? **Essentially check the hadronic final state kinematics is consistent with the kinematics of each electron, i.e. combine multiple measurements [1].**
- How to associate the semi-hard radiation from the electron with the electron? **Try to cluster** *some* **energy from the calorimeter deposits with the electron. Cluster the photons from the interaction point if they are angularly close to the electron?**

Credit: Andrii Verbytskyi

DIS Lepton Finder Algorithms

- Goal: identify the DIS lepton using only final state information

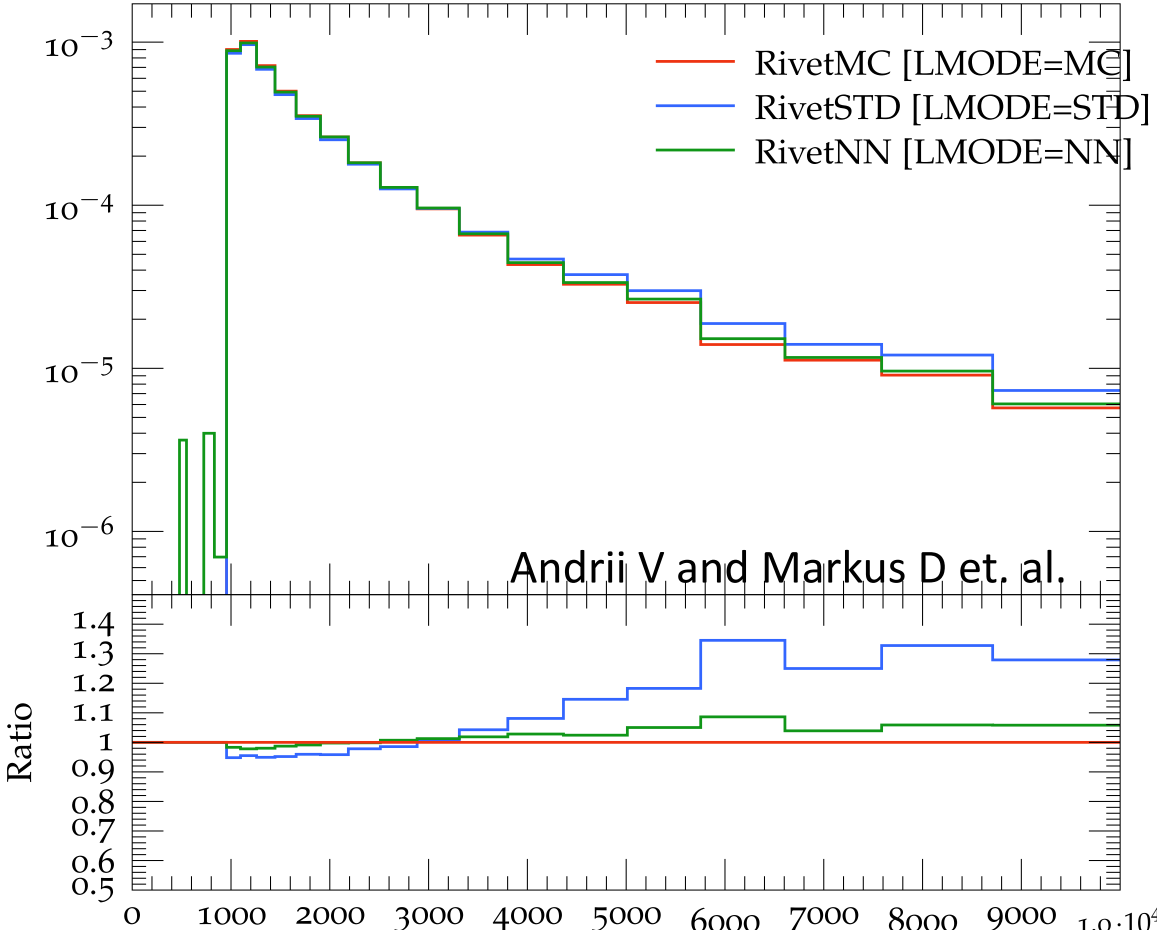
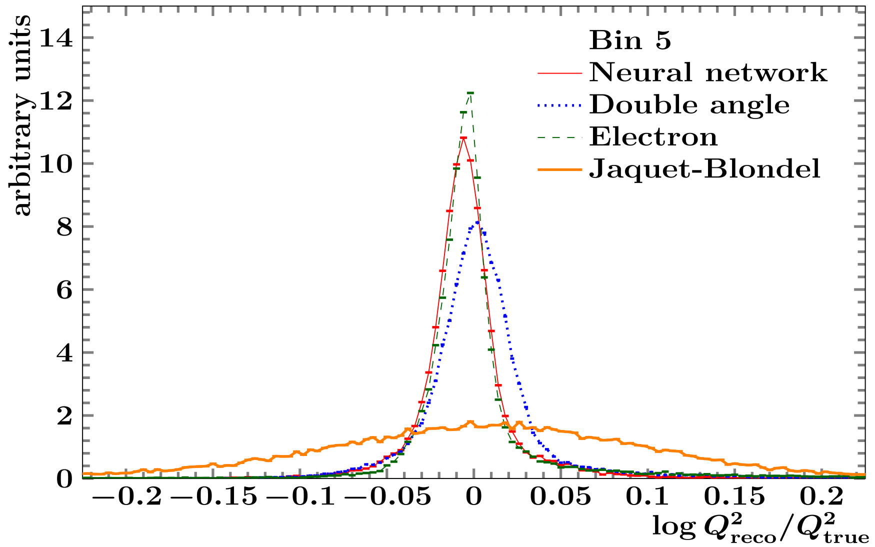


Figure: the true Q^2 from the full MC event record, the standard Rivet algorithm from the final state (STD) and the NN algorithm from the final state. The cases fro multiple electrons.



[1] Diefenthaler, Markus and Farhat, Abdullah and Verbytskyi, Andrii and Xu, Yuesheng, Deeply learning deep inelastic scattering kinematics. Eur. Phys. J. C **82**, 1064 (2022). arXiv:2108.11638.

Summary

- Particle Flow (contact: Derek Anderson dmawxc@iastate.edu)
    - Survey of PF in existing / planned experiments
    - Plan and Strawman of “alpha” version of ePIC’s PF implementation
    - Next steps (Volunteers needed)
        - Validation of existing cluster-splitting
        - Validation of existing MC-cluster associations
        - Development of cluster/track visualizers
        - Extending track-cluster associations to Hcal
        - Implementation of PF algorithm + factories
- DIS Electron Finder (contact: JDB Brandenburg.89@osu.edu)
    - Lots of progress in last 2 months – initial framework outlined
    - Next steps (Volunteers needed)
        - Study and optimize E/p cuts for electron ID
        - Validate existing Electron Finder purity / efficiency
        - Study optimal use of HCAL info for e.g. pion rejection
        - Implement track-matching using track projections
    - Stay tuned, next major milestones will be accomplished in August

Particle Flow @ ALEPH & DELPHI

- ALEPH and DELPHI implemented the first "modern" PF algorithms
  - ☞ Both make use of PID capabilities

ALEPH:

- Associate tracks to clusters and iteratively clean track+cluster objects
- **Algorithm:**
  1) Project tracks & associate them to clusters
  2) Identify $e^{\pm}$ & remove them
  3) Identity π^0/γ & remove them
  4) Identify $\mu^{\pm}$ & remove them
  5) Do track-cluster subtraction
  6) Any remaining calorimeter energies are flagged as h^0

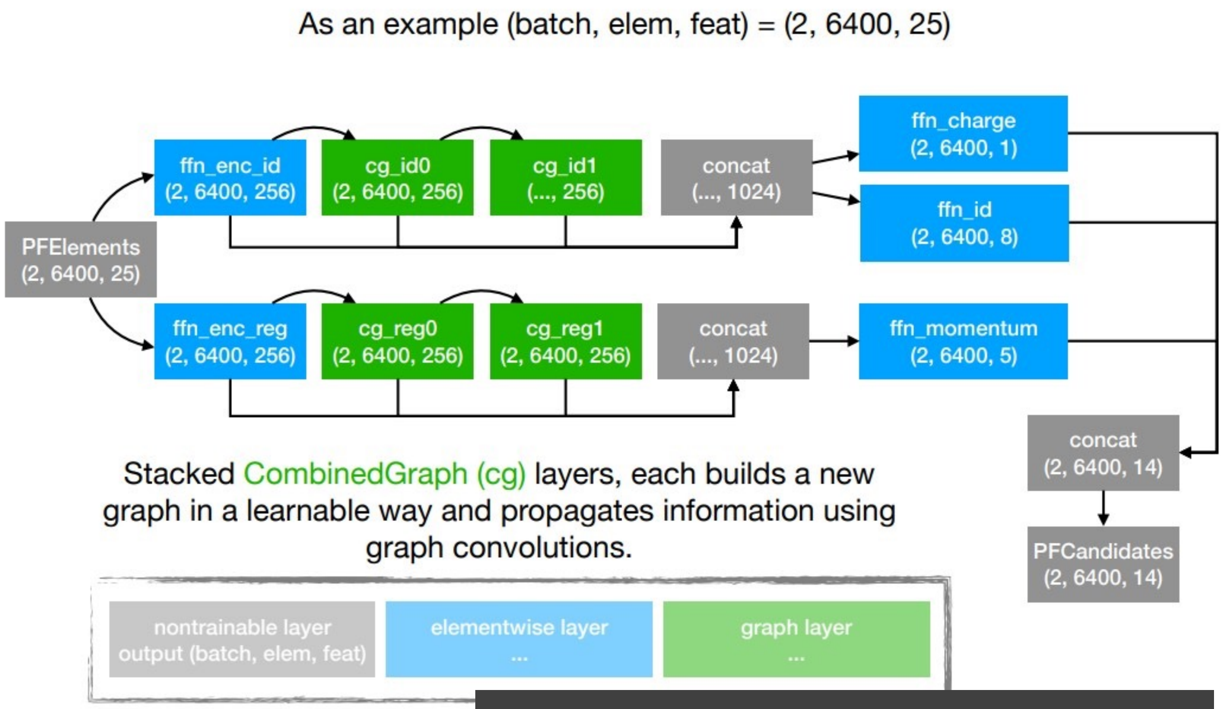
DELPHI:

- Very similar, but slightly different order of operations
- **Algorithm:**
  1) Identify $e^{\pm}/\gamma$ & remove them
  2) Extrapolate tracks through HPC (EMCal) + HCal
  3) Any clusters "close" to extrapolated tracks are associated with track and removed
  4) Any remaining clusters are flagged as h^0

Particle Flow Advances

An ML alternative to existing algorithm is being explored at CMS
- Ref.s:
  - EPJC **81**, 381 (2021)
  - JP:CS **2438**, 012100 (2023)
- Could ease computational requirements

Implementation:
1) Extensive track and calorimeter information is fed to a GNN model
2) GNN converts track/calorimeter hits/cells into connected graphs
3) Graphs are then regressed to particles



CMS [JP:CS **2438**, 012100 (2023)]

Particle Flow Advances

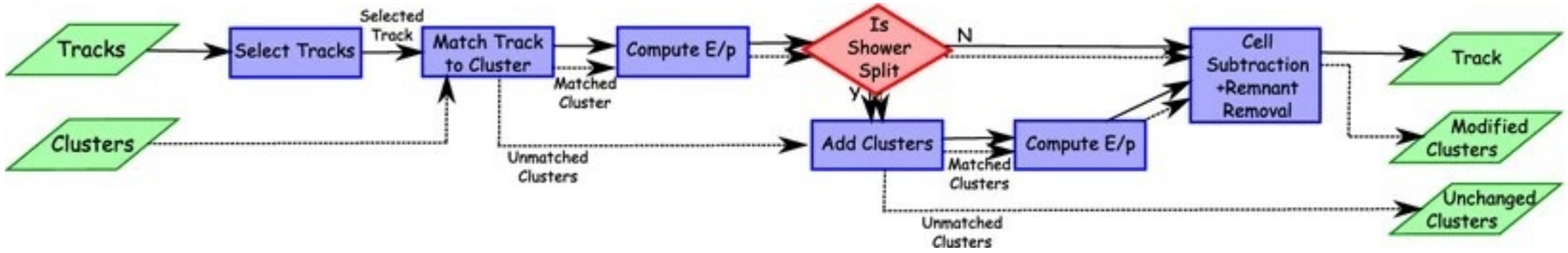
- ILC/CLIC (esp. CALICE) efforts have focused on PF
  - Produced many algorithms and approaches
  - ☞ **PandoraPFA**

- **PandoraPFA:** a *very* sophisticated PF algorithm for high granularity calorimeters
  - Part of AIDAsoft
  - Has produced an extensive detector-agnostic implementation
  - Currently deployed at MicroBooNE

Algorithm: 8 stages in total

1) Select tracks for analysis
2) Select calorimeter cells & cluster based on geometry
3) Recluster cells into cones around track+EMCal projections:
   i. 1st identify and remove possible γ clusters
   ii. Then cluster remaining cells
4) Recluster non-γ clusters based on topology
5) Attempt to split overlapping clusters
6) Apply more sophisticated γ-ID algorithm to separate γ from h^0
7) Neutral fragments from $h^{\pm}$ are identified and removed
8) **Return:** "PF Objects"
   - Matched track+cluster objects with rudimentary PID applied

Particle Flow Survey | ATLAS



ATLAS [arXiv:1703.10485]

- ATLAS makes use of a sophisticated variation on the "cluster – track" idea
  - ☞ Very similar (but still distinct) to what's being utilized at sPHENIX

ATLAS:

- **Algorithm:**
  1) Match tracks to clusters
  2) Determine if cluster is split
     a) **If yes**, then add more clusters to track+cluster object
     b) **Otherwise** move on
  3) Subtract expected track energy cell-by-cell from clusters
  4) **Return:**
     - Tracks
     - Matched clusters w/ nonzero energy after subtraction
     - Unmatched clusters

Electron Finder 07-10-2023

- Progress in June (for July Sim Campaign)
  - PR #666: Provides association containers + truth associations
    merged into main on June 24
  - PR #751: Implement basic electron finding with truth cluster matching
  - Work in parallel with the RECO approach (see tasks below)
- Major Tasks (July):
  - Utilize the July sim campaign output:
    - Implement a processor to test DIS lepton finder
    - Check purity of selected electrons
  - Track Projection Factory: provide track projections at relevant detectors (Tyler Kutz)
    - ✓ Needed data structure identified
    - ✓ Prototype factory in progress
  - Track Match Factory: Matching of projecting tracks to clusters (volunteer?)
    - Nicholas Schmidt already has some code (processor) to study track matching
    - Provides a starting point for factory
  - Study of E/p cuts to implement (volunteer potentially identified, discussing next steps)
  - Study HCAL info for hadron rejection / electron id
- Plans for July sim campaign
  - Utilize “ReconstructedElectrons” to test-drive DIS lepton finder (should be in EICRecon for Aug)
  - Continue work towards fully RECO level (complete track matching / compare to truth level matching)

Truth approach

- PR #751 Add reconstructed electron factory, algorithm utilizing E/p cut
  - https://github.com/eic/EICrecon/pull/751
  - ReconstructedElectrons Factory
  - Input:

```
        {"MCParticles", "ReconstructedChargedParticles", "ReconstructedChargedParticleAssociations",
        "EcalBarrelScFiClusterAssociations",
        "EcalEndcapNClusterAssociations",
        "EcalEndcapPClusterAssociations",
        "EcalEndcapPInsertClusterAssociations",
        "EcalLumiSpecClusterAssociations",
        },
```

  - Output: “ReconstructedElectrons”
- Utilizes the ElectronReconstruction Algorithm
  - Any track with an ECAL match
  - Accept if 0.9 < E/p < 1.2 (needs to be studied and optimized)
  - TODO: use HCAL
  - TODO: handle multiple matches
- This is meant to be initial skeleton – keep same structure for RECO approach