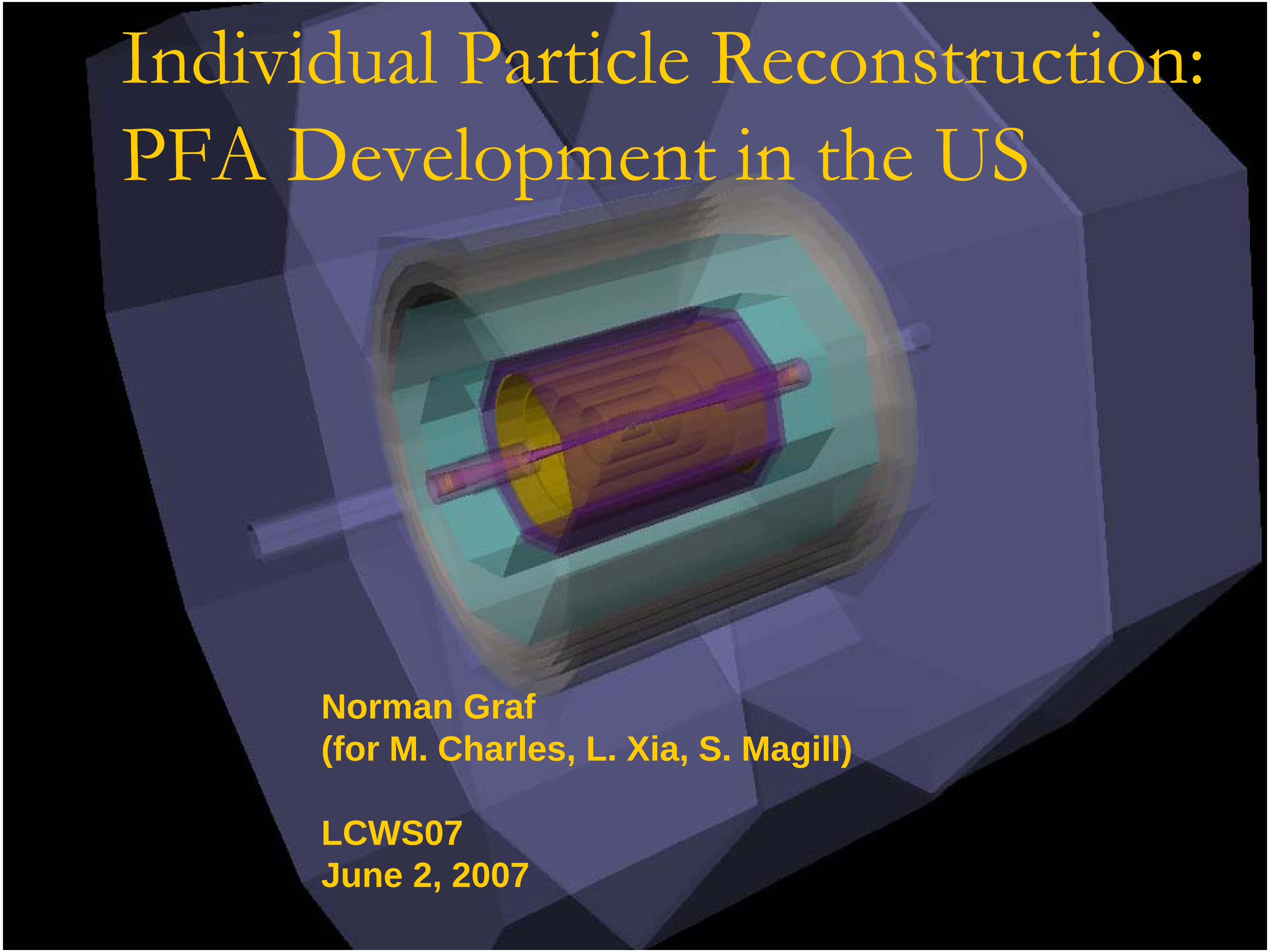
Individual Particle Reconstruction: PFA Development in the US



Norman Graf
(for M. Charles, L. Xia, S. Magill)

LCWS07
June 2, 2007

Overview

- PFA reconstruction is complex & you have to get many individual steps right:
  - track finding, fitting & extrapolation
  - track-cluster matching
  - MIP identification
  - photon identification
  - hadronic shower clustering
  - handling of displaced secondaries
  - calibration of photons
  - calibration of neutral hadrons
  - E/p cut (including calibration)
- PFA development isn't about finding the "magic bullet" perfect algorithm. It's about iteratively
  - finding the worst problems that are limiting performance
  - fixing them
  - hopefully seeing things improve a little
  - finding the next worst problems
- Interplay between detector and algorithm makes it tricky to really tune detector design
  - you need a really good (mistake-free) algorithm
  - fair comparison means equal tuning on different detectors
- But we can say: you can do at least this well with this detector.

Analysis Tools

- Common input data samples
  - Single particles for tuning detector response
  - Dijet samples (uds) @ 91, 200, 500 GeV cms
  - $e^+e^- \rightarrow ZZ \rightarrow (\nu\nu)(qq)$
- Common detector simulations
- Provide a number of QA tools to assure that some common tasks are handled in the same way.
- Calorimeter Calibrations
  - Sampling fractions & common energy corrections
- Perfect PFA performance
  - Common definitions of “final state” particles
  - Based on Generator or Simulated Particles?
  - Standard cheated tracks, cheated clusters

Analyses

- Presenting three different analyses
  - Mat Charles (U. Iowa)
  - Lei Xia (ANL)
  - Steve Magill (ANL), Norman Graf (SLAC)

- Detectors are variants of SiD concept.
  - But framework supports essentially arbitrary detectors and plan is to explore larger phase space.

- More details in slides presented at recent SiD workshop at FNAL.

Algorithm Description I

- Main philosophy is to tackle the (relatively) easy problems first.
- Step 1: Find photons, remove their hits.
  - Tight clustering
  - Apply shower size, shape, position cuts (very soft photons fail these)
  - Make sure that they aren't connected to a charged track
- Step 2: Identify MIPs/track segments in calorimeters. Identify dense clumps of hits.
  - These are the building blocks for hadronic showers
  - Pretty easy to define & find
- Step 3: Reconstruct skeleton hadronic showers
  - Coarse clustering to find shower components (track segments, clumps) that are nearby
  - Use geometrical information in likelihood selector to see if pairs of components are connected
  - Build topologically connected skeletons
  - If >1 track connected to a skeleton, go back and cut links to separate
  - Muons and electrons implicitly included in this step too
- Step 4: Flesh out showers with nearby hits
  - Proximity-based clustering with 3cm threshold
- Step 5: Identify charged primaries, neutral primaries, soft photons, fragments
  - Extrapolate tracks to clusters to find charged primaries
  - Look at size, pointing, position to discriminate between other cases
  - Merge fragments into nearest primary
  - Use E/p veto on track-cluster matching to reject mistakes (inefficient but mostly unbiased)
  - Use calibration to get mass for neutrals & for charged clusters without a track match (calibrations for EM, hadronic showers provided by Ron Cassell)
- Known issues & planned improvements:
  - Still some cases when multiple tracks get assigned to a single cluster
  - Punch-through (muons and energetic/late-showering hadrons) confuses E/p cut
  - Improve photon reconstruction & ID
  - Improve shower likelihood (more geometry input)
  - Use real tracking when available
  - No real charged PID done at this point

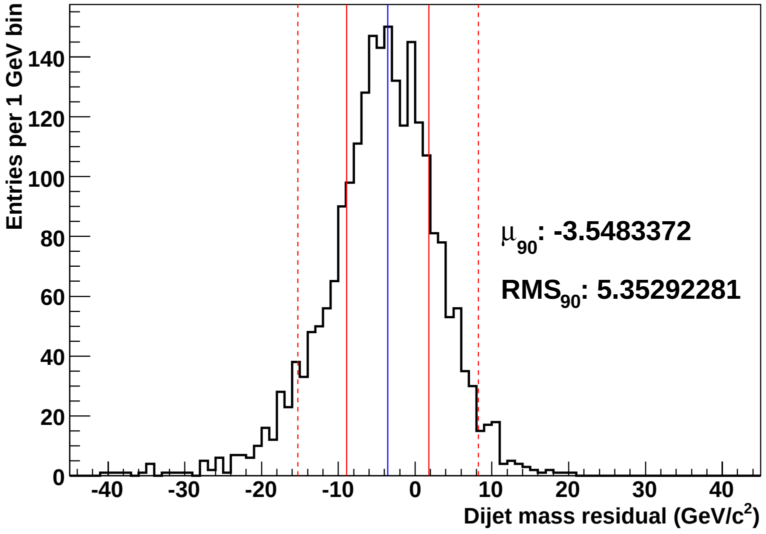
Current Performance

- Looking at: $e^+ e^- \rightarrow$ Z1 ($\nu\nu$) Z2 (qq) for q=u,d,s at $\sqrt{s}$=500 GeV
- requiring primary quarks have |cos(theta)|<0.8
- reconstructing dijet invariant mass, i.e. mass of Z2
- quoting residual = (true mass of Z2 - reconstructed mass of Z2)

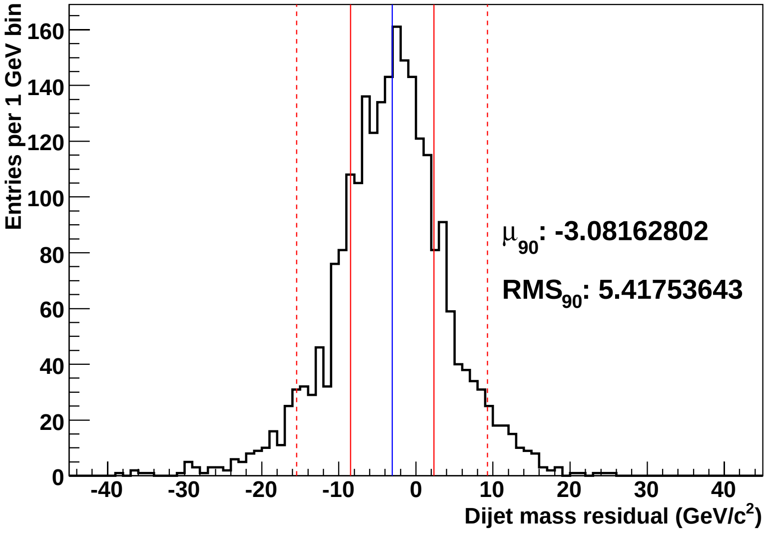
M. Charles

M. Charles

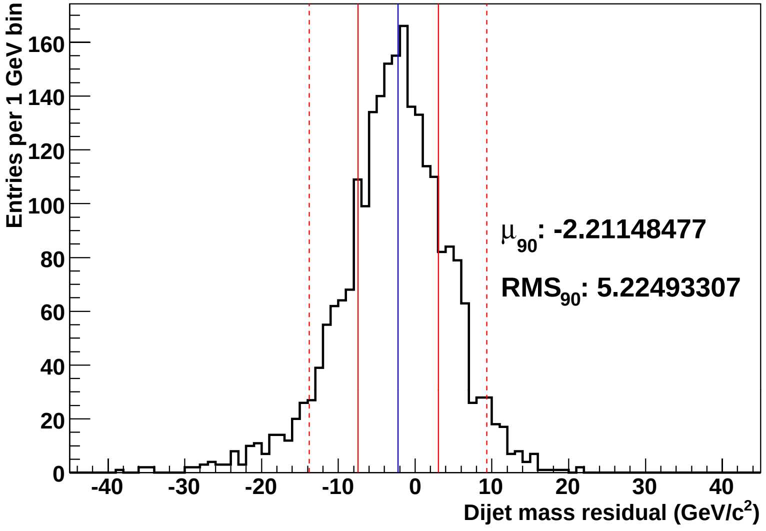
SiD W/Scin HCAL



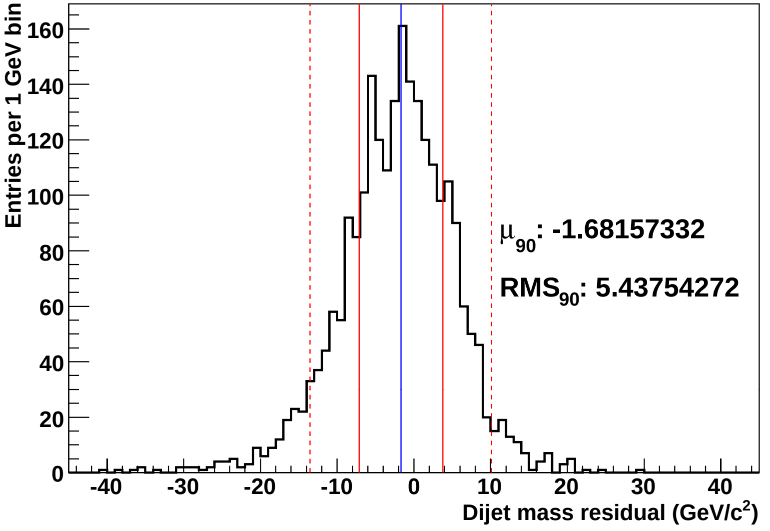
SiD W/RPC HCAL



SiD SS/Scin HCAL



SiD SS/RPC HCAL



Current Performance

M. Charles

	rms90	mean90
W/Scint	5.4 GeV	-3.5 GeV
W/RPC	5.4 GeV	-3.1 GeV
SS/Scint	5.2 GeV	-2.2 GeV
SS/RPC	5.4 GeV	-1.7 GeV

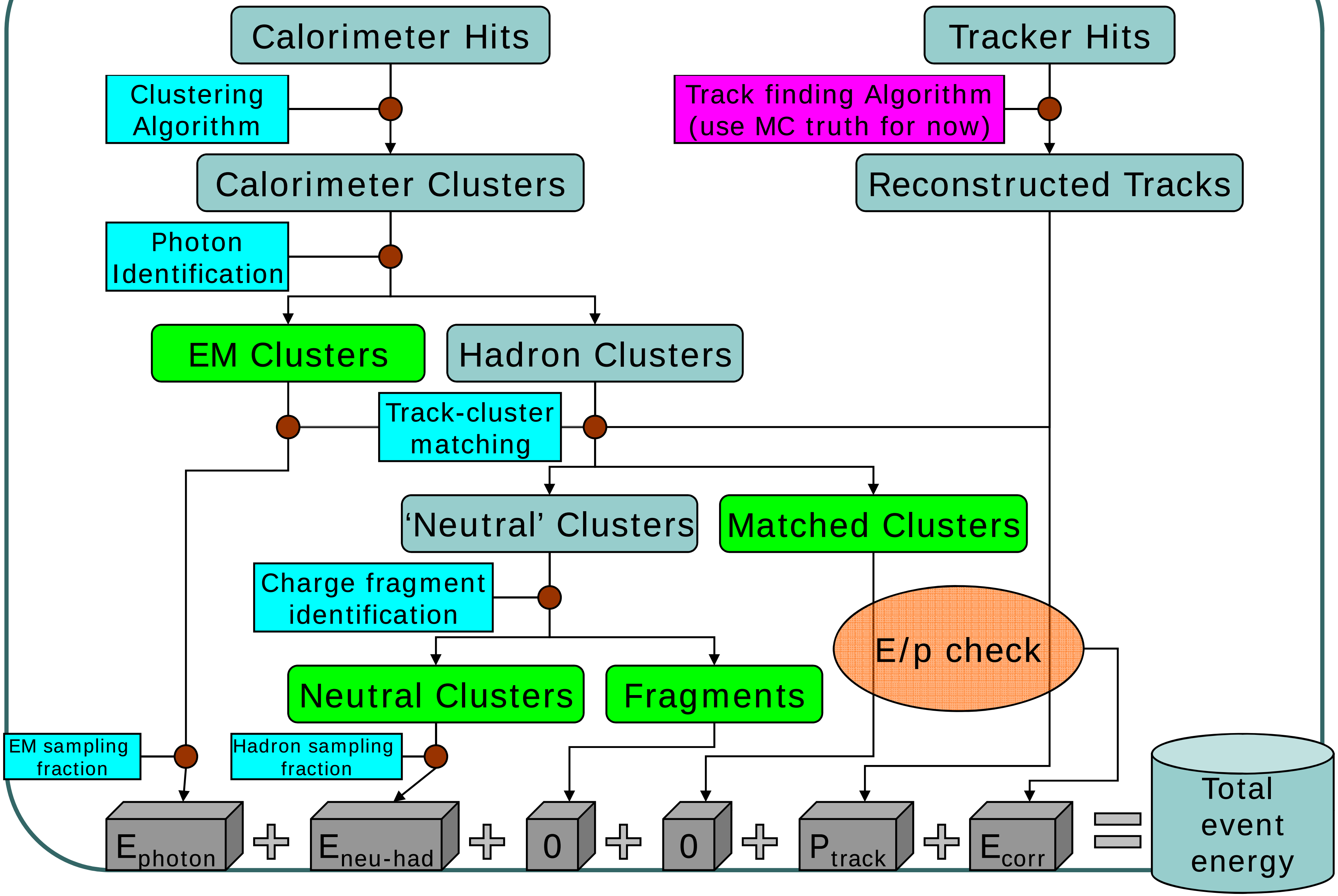
- Little discriminations between designs at this point
- Working to improve performance…
- … but not actually too far from “glass ceiling” of 4.1 GeV (for W/RPC).
- To understand/approach/move beyond that, need to think more broadly:
  - Is assumed tracking performance realistic? Too pessimistic? Too optimistic?
  - Can calibration be improved?
  - Can we put in more information? E.g. pick up low-pt tracks that are being ignored
  - Is this event type (with boosted jets) representative?
  - What physics models are being used for the showers? (lcphys vs lhep vs…)

Performance Caveats

- Numbers depend on things that are not really algorithm specific:
  - detector
  - physics event type
  - Assumptions about tracking & track extrapolation
  - polar angle & acceptance
  - calibration
  - how physics quantity (e.g. dijet mass) is measured
  - how figure of merit (e.g. rms90) is computed
- Some things lower the ceiling rapidly. For example, with $e^+e^- \to ZZ \to (\nu\nu)(qq)$ in |cos(theta)|<0.8 for W/RPC SiD detector, dijet mass resolution (rms90) is:
  - 0.0 GeV if completely cheating
  - 0.5 GeV dropping missed particles (using GenFinalStateParticles)
  - 1.2 GeV dropping missed particles (using SimFinalStateParticles with cuts)
  - 2.9 GeV including resolution of neutrals using Ron's Z-pole calibration
  - 3.2 GeV requiring a track for charged particles (else treated as neutral hadrons)
  - 3.4 GeV requiring that tracks can be extrapolated to ECAL surface
  - 4.1 GeV requiring cluster be within 25mm (depending how track extrapolation is done)
  - 5.1 GeV if using naive helical track extrapolation instead
- So different assumptions about tracking, calibration, etc can have a big impact.
  - … unless completely confusion-dominated

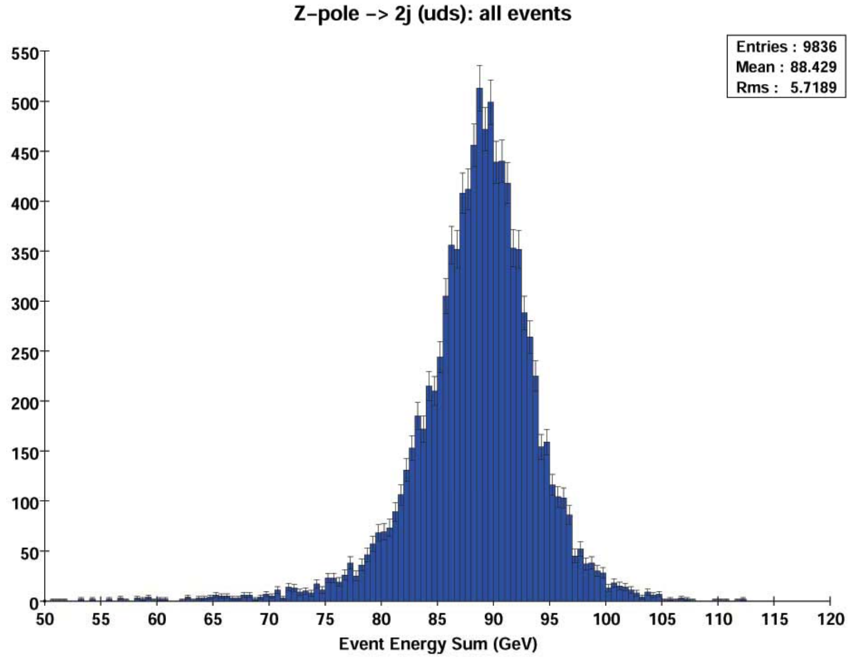
M. Charles

Algorithm description II



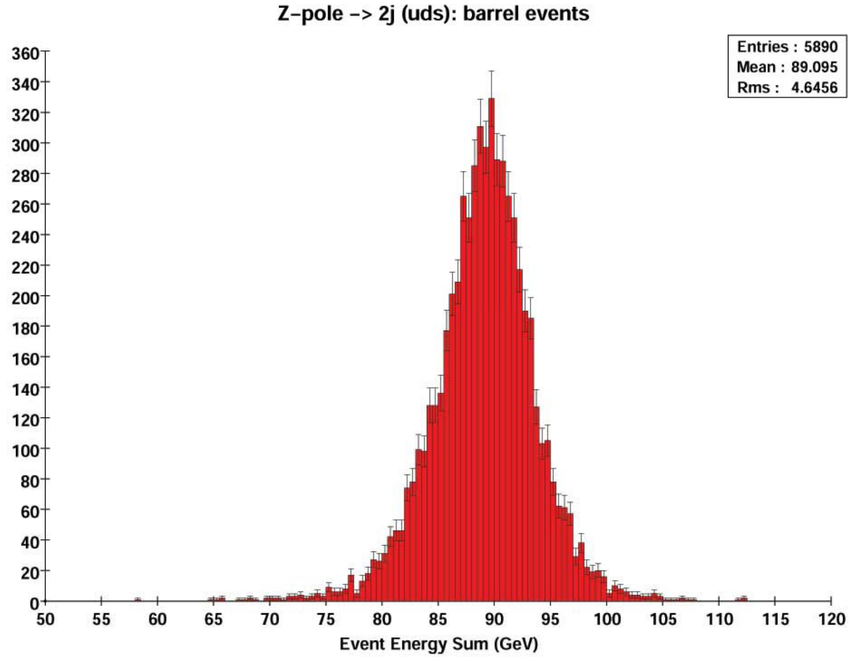
L. Xia

Current PFA Z-pole performance



All events, no cut

Mean 88.43 GeV
RMS 5.718 GeV
RMS90 3.600 GeV
[38.2 %/sqrt(E)]



Barrel events (cos(theta[Q]) < 1/sqrt(2))

Mean 89.10 GeV
RMS 4.646 GeV
RMS90 3.283 GeV
[34.7 %/sqrt(E)]

L. Xia

Progress on PFA performance at Z-pole

improved/tuned algorithms

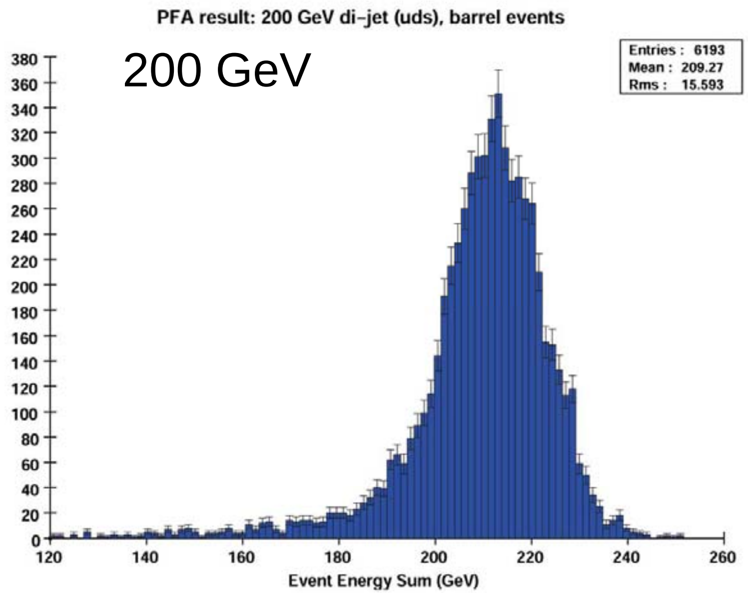
- Clustering algorithm
- Track extrapolation
- Track/cluster matching
- Photon finder
- E/P correction
- Fragment identification

Still need improvement

- Clustering algorithm
- Photon finder
- Fragment attachment
- …(more to come)

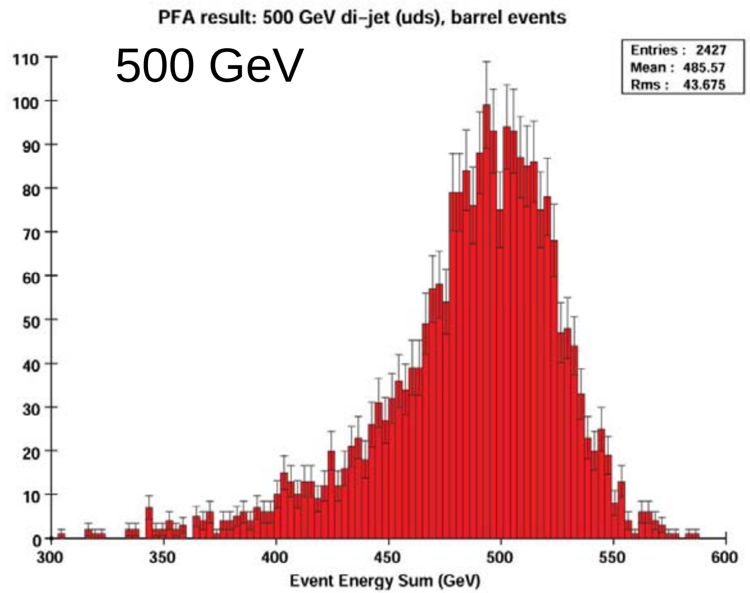
L. Xia

Using Z-pole tuned PFA at higher energies



Barrel event (cos(theta[Q]) < 1/sqrt(2))

Mean 209.3 GeV
RMS 15.6 GeV
RMS90 9.12 GeV
[62.6%/sqrt(E)]



Barrel event (cos(theta[Q]) < 1/sqrt(2))

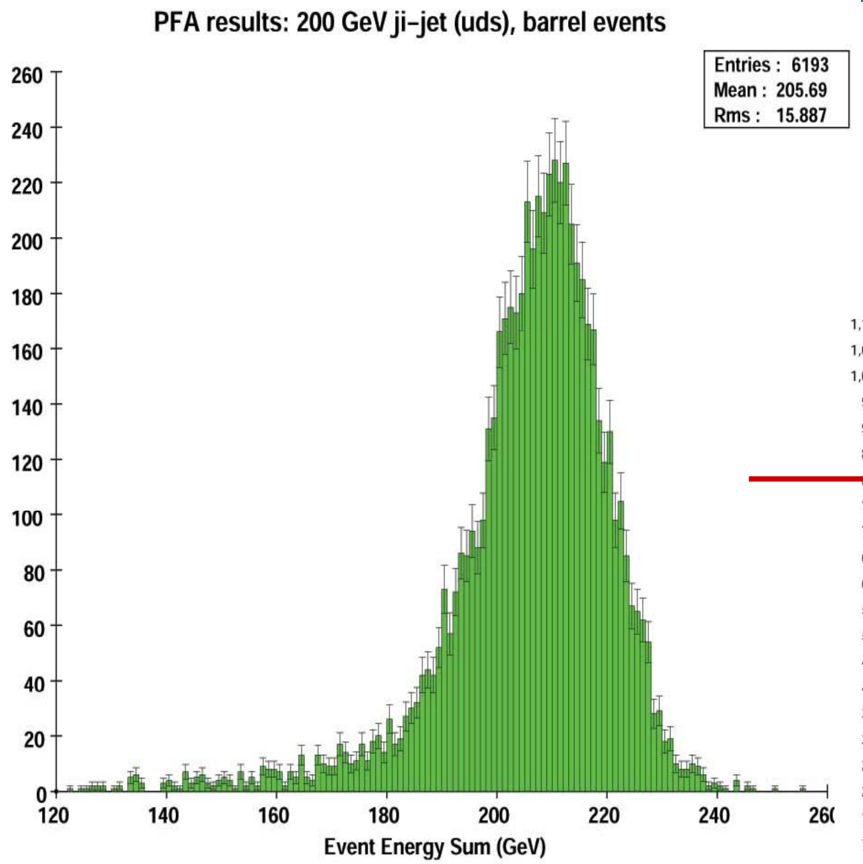
Mean 485.6 GeV
RMS 43.7 GeV
RMS90 27.6 GeV
[124.%/sqrt(E)]

Not good yet – but algorithms not tuned at these energy
A lot of improvement expected, clearly still a lot of work to be done!

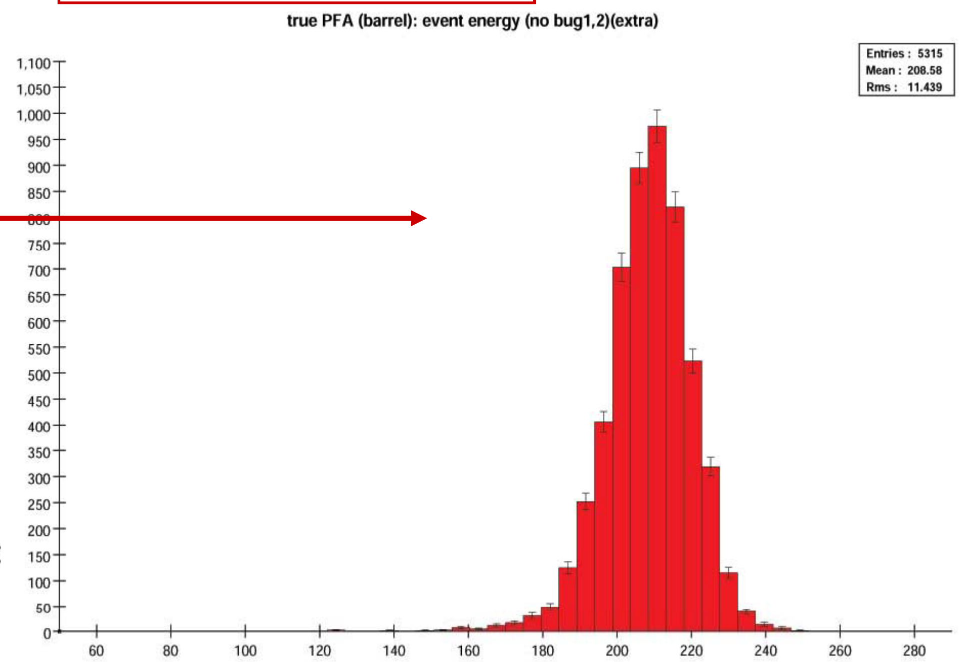
L. Xia

Shower leakage: di-jet at 200 GeV

L. Xia



Removing events with shower leakage

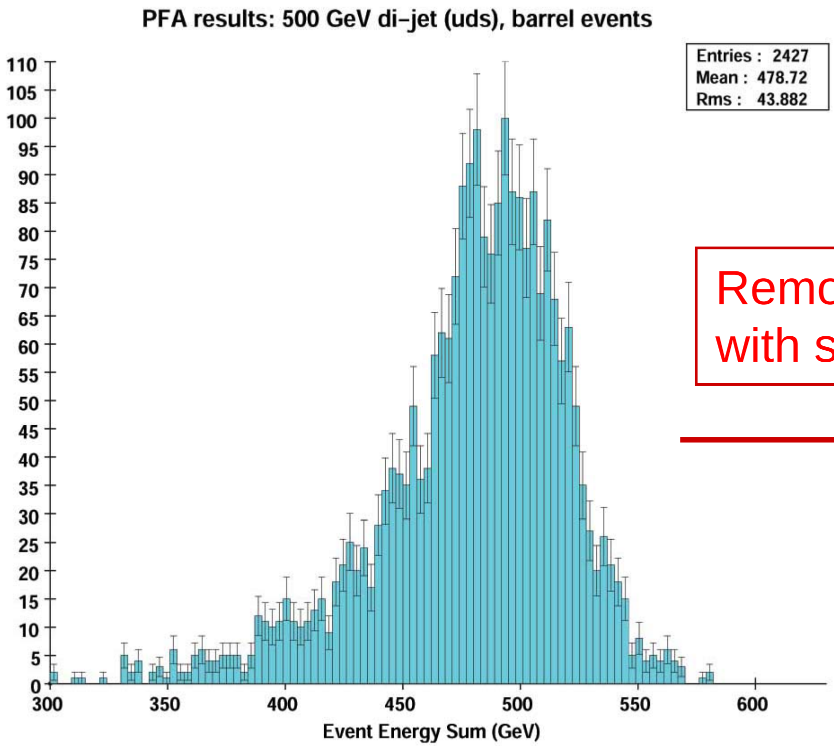


RMS = 15.89 GeV
RMS90 = 9.632 GeV
[66.7%/sqrt(E)]

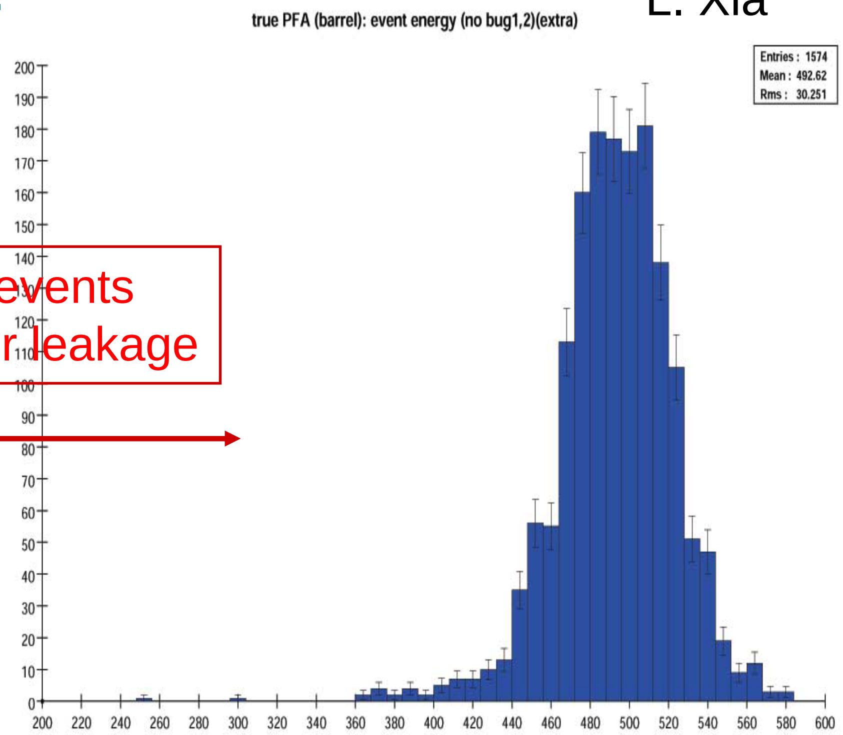
RMS = 11.44 GeV
RMS90 = 8.45 GeV
[~59%/sqrt(E)]

Shower leakage: di-jet at 500 GeV

L. Xia



Removing events
with shower leakage



RMS = 43.88 GeV
RMS90 = 28.11 GeV
[127.%/sqrt(E)]

RMS = 30.25 GeV
RMS90 = 21.4 GeV
[~97%/sqrt(E)]

- Shower leakage affect PFA performance at high energy
- Events with heavy shower leakage could be identified by hits in the muon detectors
- Use hits in the muon detectors to estimate shower leakage?

Re-writing PFA according to lcsim template

- Motivation
  - Facilitate exchange with other PFA efforts
  - Check my algorithm from head to toe
  - Write intermediate lcio output file to save running time on the rest of the PFA (do not repeat clustering for each run)
- Current status
  - Program re-writing is done
  - Algorithm is fully modular
  - Followed lcsim template convention as closely as I can
    - However, used some extensions of standard interface
  - Some issues exist
    - Z-pole result is still different from old algorithm, but the difference is very small now
    - Some problems with the intermediate lcio file

Re-writing PFA according to lcsim template

- Current status (continue)
  - Program performance
    - Overall running time is actually longer
      - 10k Z-pole events: 10hr => 14hr
    - If intermediate lcio successful, can save ~90% running time (by not repeating clustering each time)
  - Will upload to lcsim cvs, after solving some obvious issues

Future plans

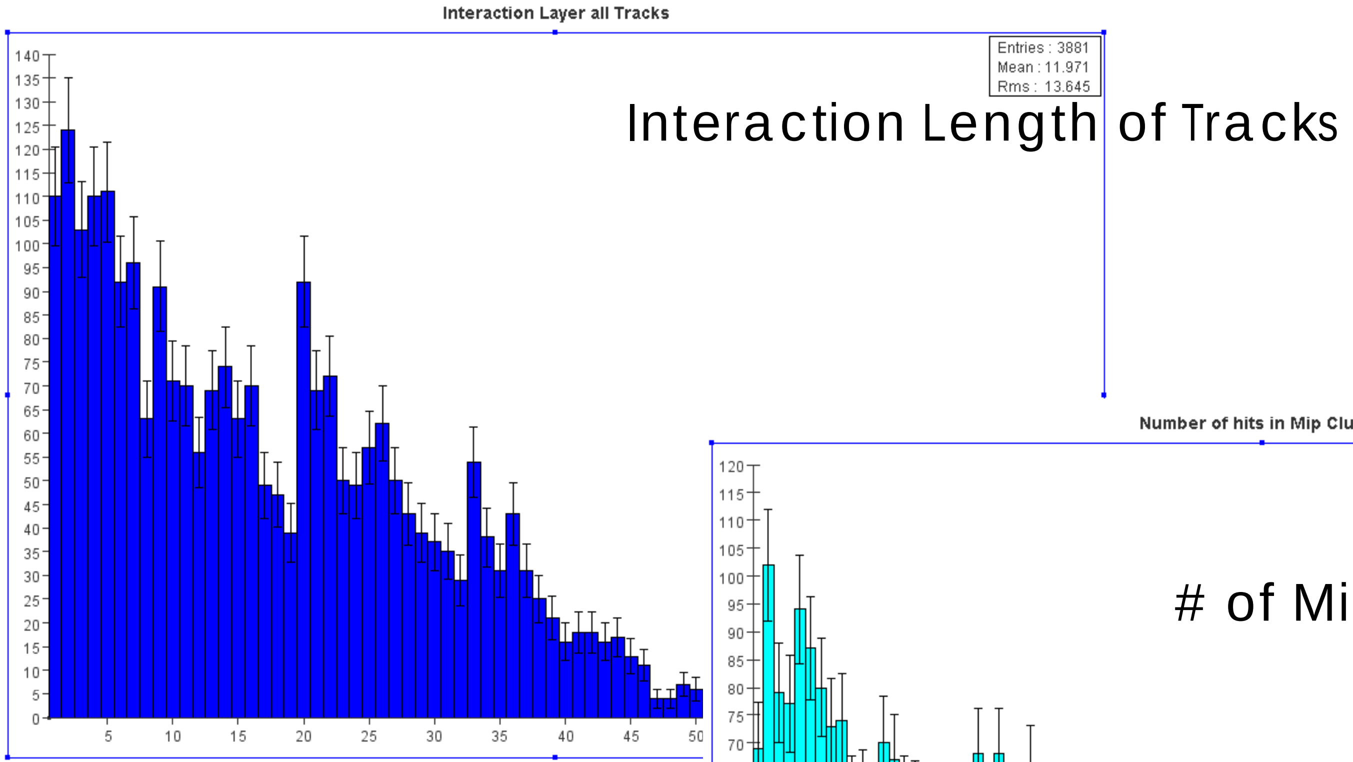
- Finish up PFA tuning at Z-pole
- Concentrate on performance improvement at higher energies
- Shower leakage study using detector models with extended HCal
- Detector performance study with fully developed PFA

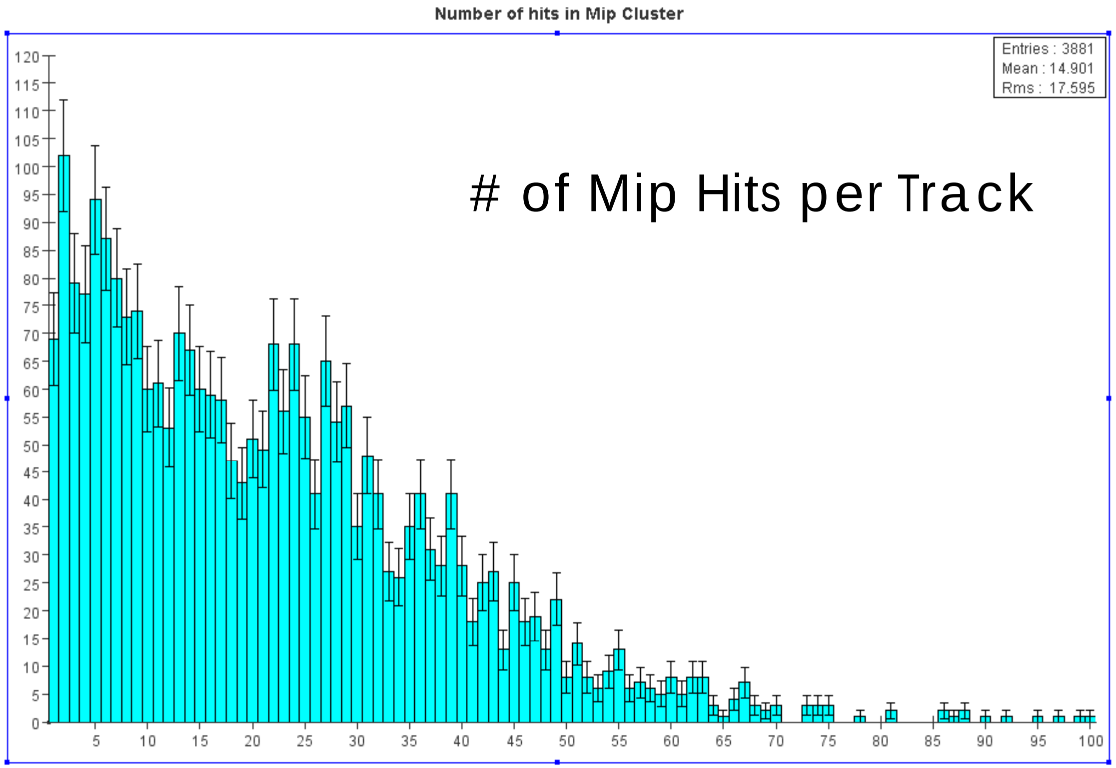
Algorithm Description III

- *Track-linked mip segments (ANL)*
  - find mip hits on extrapolated tracks, determine layer of first interaction based solely on cell density (no clustering of hits) ($\rightarrow \mu$ candidates)
- *Photon Finder (SLAC)*
  - use analytic longitudinal H-matrix fit to layer E profile with ECAL clusters as input ($\rightarrow \gamma$, π^0, $e^{+/-}$ candidates)
- *Track-linked EM and HAD clusters (ANL, SLAC)*
  - substitute for Cal objects (mips + non-EM ECAL shower clusters + HCAL calorimeter hits (or clusters))
  - reconstruct linked mip segments + clusters iterated in E/p
  - Analog or digital techniques in HCAL ($\rightarrow \pi^{+/-}$ candidates)
- Neutral Finder algorithm (SLAC, ANL)
  - cluster remaining CAL cells, merge, cut fragments ($\rightarrow$ n, K^0_L candidates)
- Jet algorithm
  - Reconstructed Particles used as input to jet algorithm, further analysis

Track-Linked mip segments

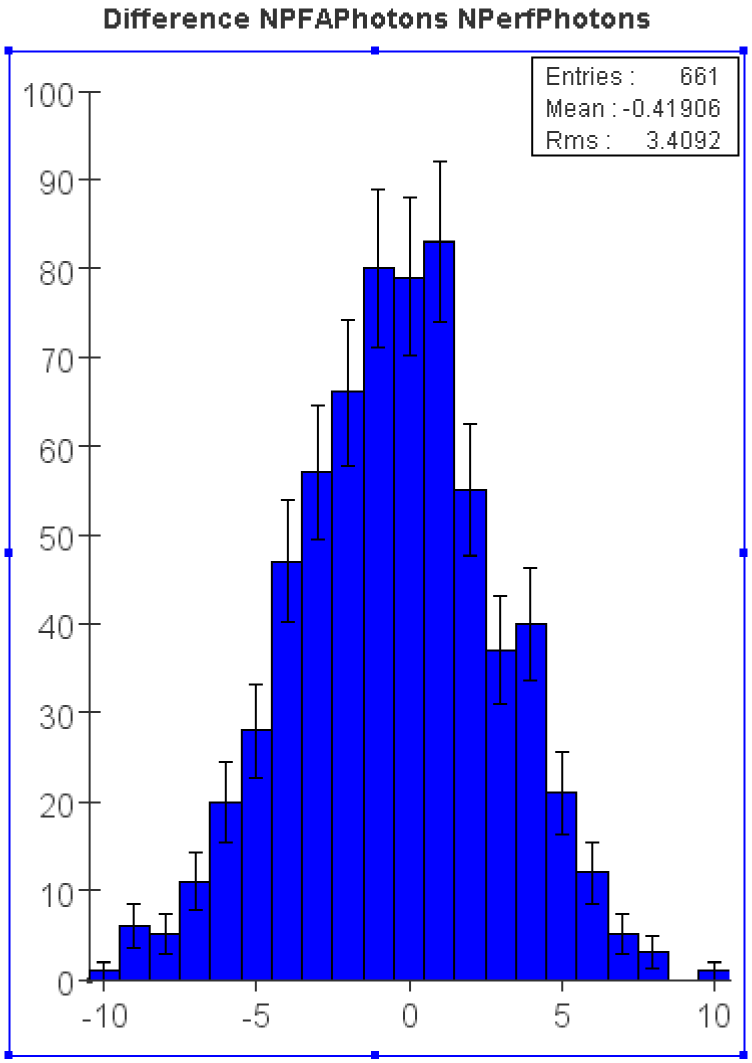
S. Magill



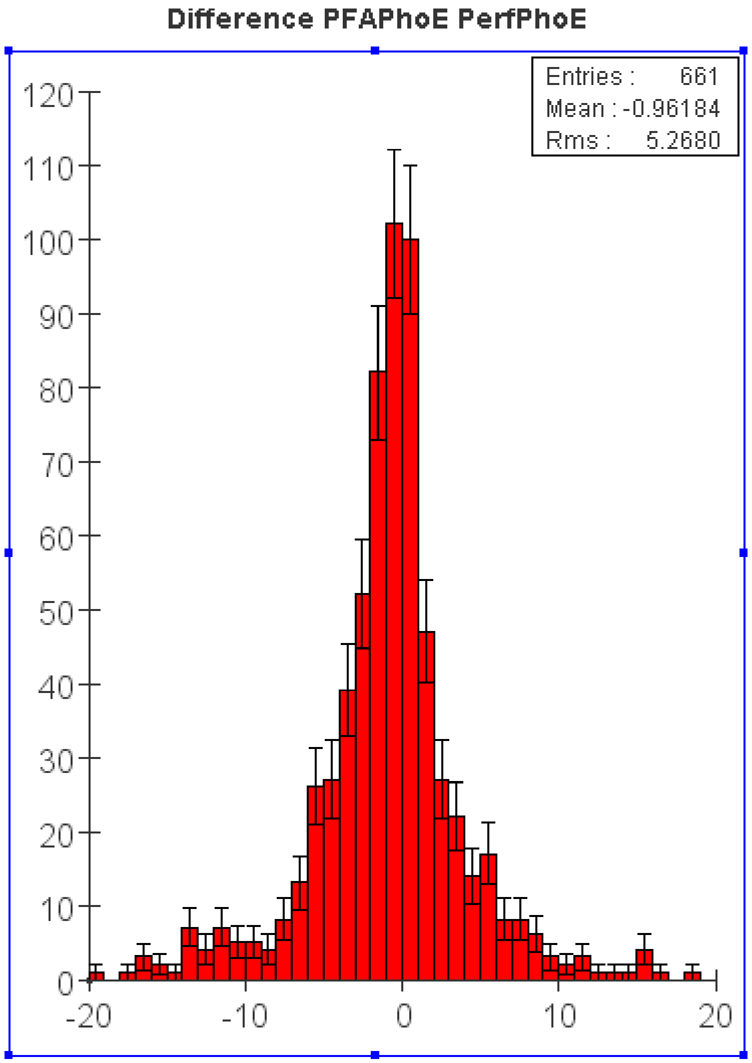


Photon Finding

S. Magill



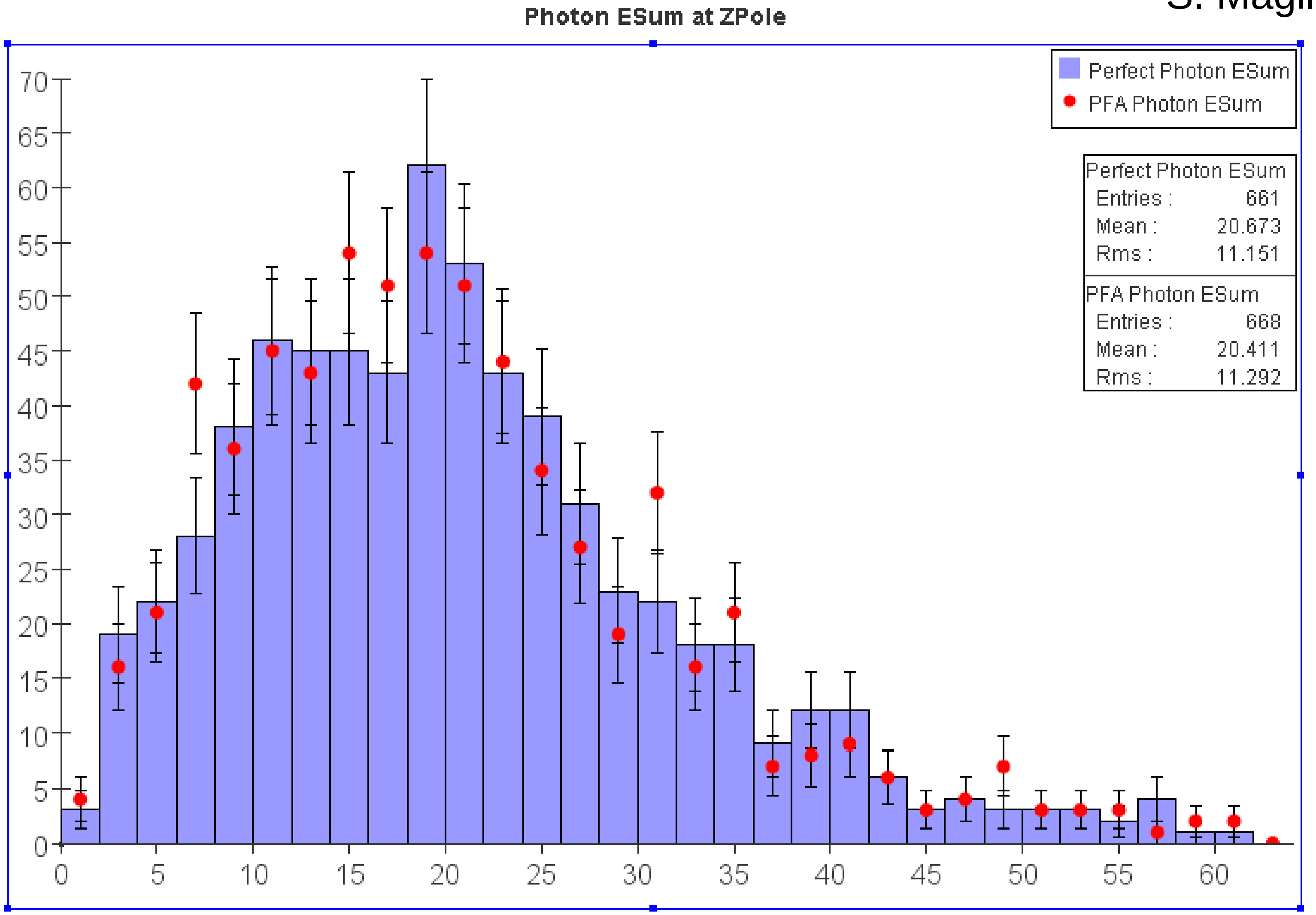
δNumber



δEnergy

Photon Finding

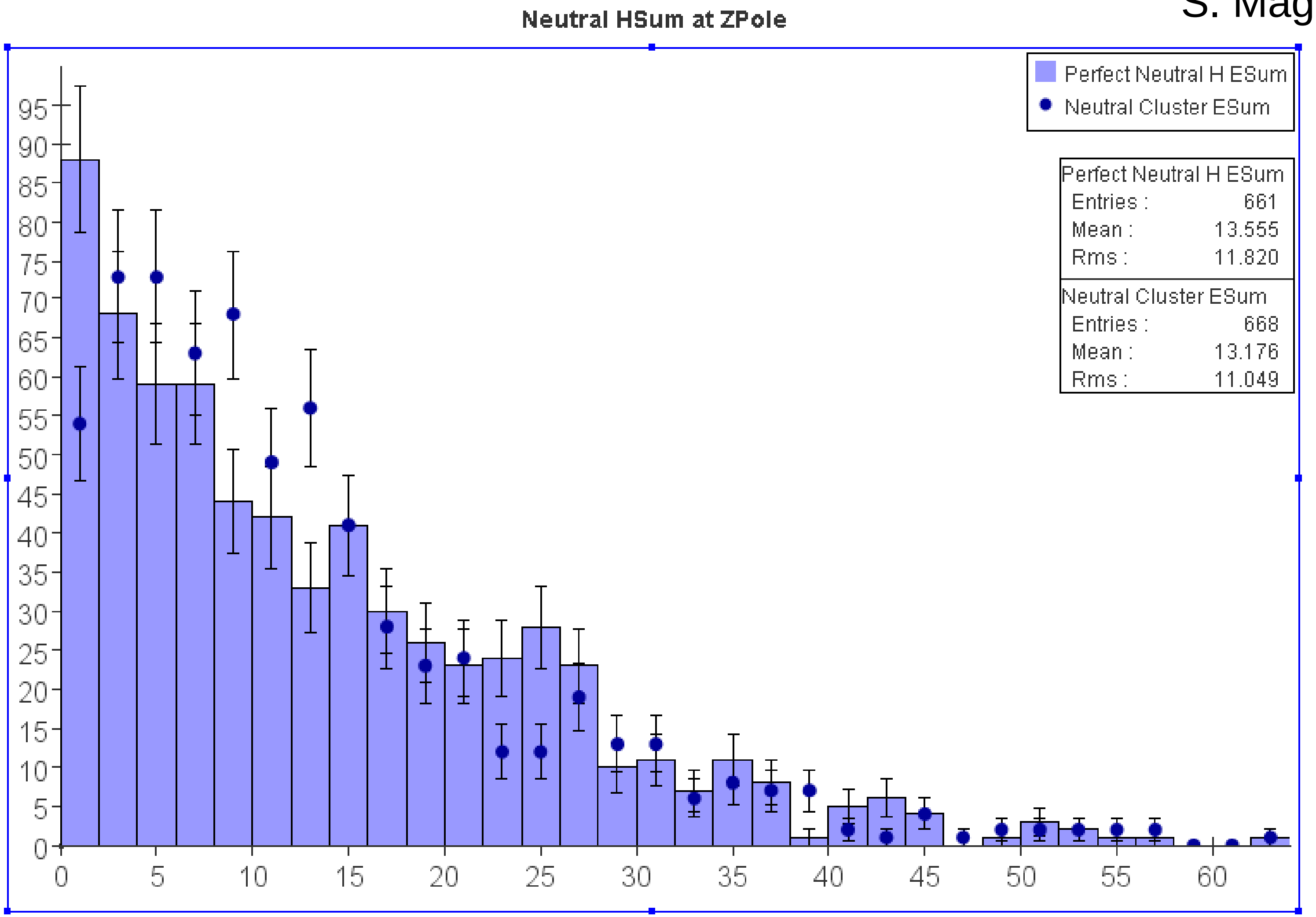
S. Magill



Photon Energy Sum @ Z Pole

Neutral Hadrons

S. Magill



Neutral Hadron Energy Sum @ Z Pole

Reconstruction Framework

- Analyses shown here done within the general ALCPG simulation & reconstruction environment.
- Framework exists for the full reconstruction chain which allows modular implementation of most aspects of the analysis.
- Interfaces allow different clustering algorithms to be swapped in and alternate strategies to be studied.
- Goal is to facilitate cooperative development and reduce time & effort between having an idea and seeing the results.

Summary

- Individual Particle Reconstruction algorithms being developed with minimal coupling to specific detector designs.
  - Will allow full phase space of detector designs to be studied in a common framework.
- Finishing development of common infrastructure tools
  - Calibration method for detector models
  - Perfect PFA prescription
- Released Reconstruction Template
  - Enables e.g. Cluster algorithm substitution, CAL hit/cluster accounting
  - Migrating individual analyses into this framework
- Optimization & Standardization of reconstructors
  - Photon & muon finders fairly mature, close to release
- Analysis emphasis on dijet invariant mass resolution in physics events
  - Currently $e^+e^- \to ZZ \to (\nu\nu)(qq)$ (No jet combinatorics, uds) (2)
  - Results soon from $e^+e^- \to ZZ \to (qq)(qq)$ & $e^+e^- \to ZZ\nu\nu, WW\nu\nu$ (4)
  - $\to tt$ (6)
  - $\to tth$ (8)
-
- Plan to release "canned" physics analyses to reduce systematic uncertainties in e.g. jet-finding, combinatorics, constrained fits, …
- Closing in on detector optimization using results.

Additional Information

- lcsim.org - http://www.lcsim.org
- ILC Forum - http://forum.linearcollider.org

- Wiki - http://confluence.slac.stanford.edu/display/ilc/Home
- org.lcsim - http://www.lcsim.org/software/lcsim
- Software Index - http://www.lcsim.org/software
- Detectors - http://www.lcsim.org/detectors

- LCIO - http://lcio.desy.de
- SLIC - http://www.lcsim.org/software/slic
- LCDD - http://www.lcsim.org/software/lcdd
- JAS3 - http://jas.freehep.org/jas3
- AIDA - http://aida.freehep.org
- WIRED - http://wired.freehep.org