SiD PFA and Simulations

S. R. Magill ANL
N. Graf SLAC
SiD/ALCPG 10/24/07

- Where we Stand
  - Simulation Status
  - PFA Performance
- Recent R&D Efforts
  - Standard Software Tools
  - PFA Development Tools
- Future Developments
  - Timetable
  - Studies/Manpower Needed
- Summary

Current Simulation Status - Goals

- Enable full studies of ILC physics to optimize detector design and eventual physics output
  - Use realistic detector geometries
  - Full simulation (in combination with fast parameterized MCs)
  - Full reconstruction
    - Simulate benchmark physics processes on different full detector designs.
    - Encourage development of realistic analysis algorithms
    - See how these algorithms work with full detector simulations
- Facilitate contribution from physicists in different locations with various amounts of time available (normally not much!)
  - Software should be easy to install, learn, use
    - Goal is to allow software to be installed from CD or web with no external dependencies
    - Support via web based forums, tutorials, meetings.

Improved Detector Simulations

- The full simulation package slic reads in geometries in lcdd, which is a low-level format that targets Geant4 primitives.
  - Detectors of arbitrarily complex shape and readout can be simulated using only xml file as input.
- However, it would be extremely tedious to generate these files.
- Would also not provide a connection to the reconstruction, nor to the event display.
- Prefer (but not required) to define geometries using a “compact” description.
- Small Java program for converting from compact description to a variety of other formats.
  - GeomConverter.

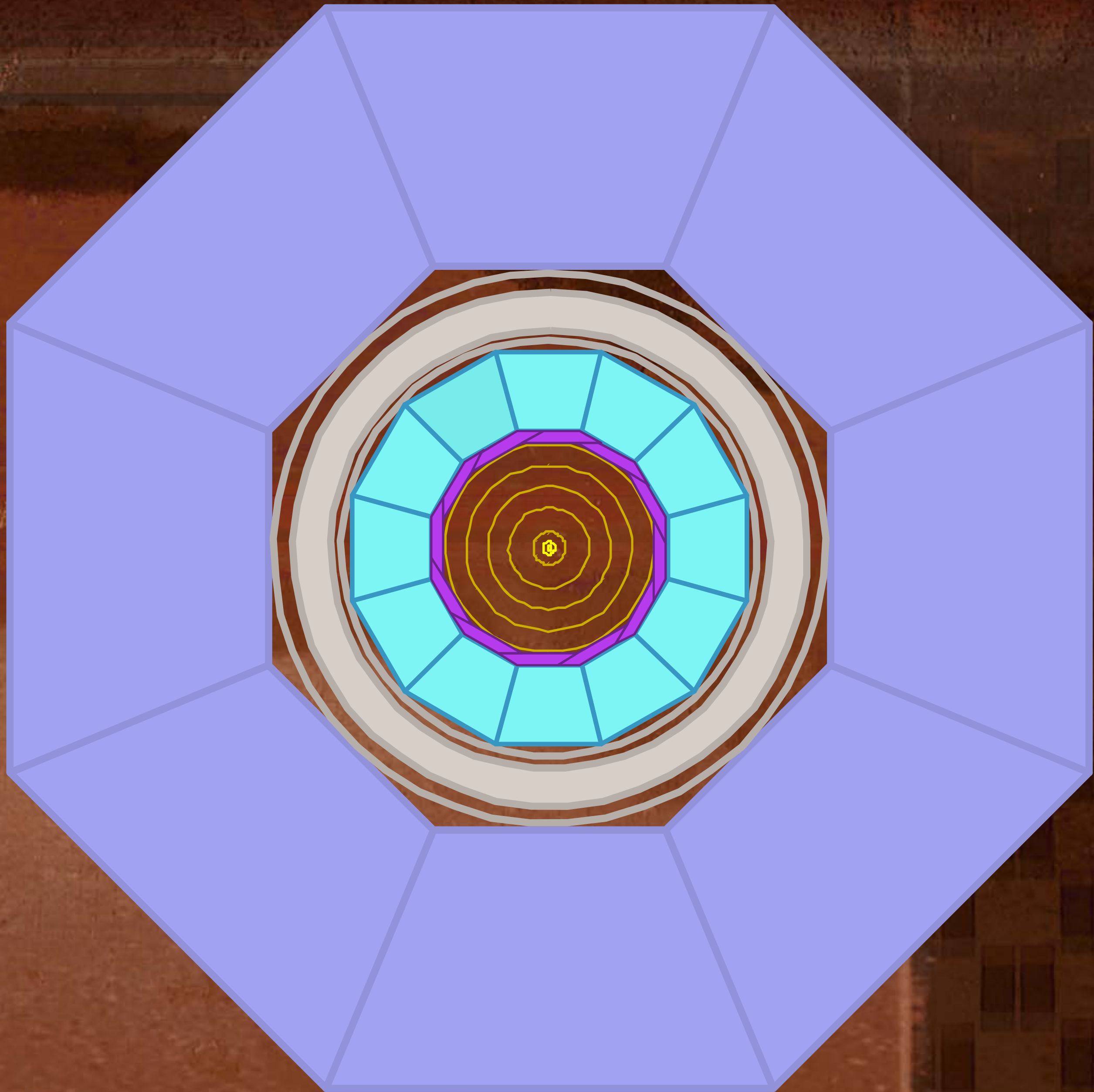
sid01_polyhedra

Dodecagonal, overlapping stave EMCal

Dodecagonal, wedge HCal

Cylindrical Solenoid with substructure

Octagonal, wedge Muon



Detector Variants

- Runtime XML format allows variations in detector geometries to be easily set up and studied:
  - Stainless Steel vs. Tungsten HCal sampling material
  - RPC vs. GEM vs. Scintillator readout
  - Layering (radii, number, composition)
  - Readout segmentation (size, projective vs. nonprojective)
  - Tracking detector technologies & topologies
    - TPC, Silicon microstrip, SIT, SET
    - “Wedding Cake” Nested Tracker vs. Barrel + Cap
  - Field strength
  - Far forward MDI variants (0, 2, 14, 20 mr)
- Prepared for Detector Optimization

“Signal” and Diagnostic Samples

- Have generated canonical data samples and have processed them through full detector simulations.
- simple single particles: γ, μ, e, $\pi^{+/-}$, n, …
- composite single particles: π^0,ρ, K^0_S ,τ, ψ, Z, …
- Z Pole events: comparison to SLD/LEP
- WW, ZZ, tt, qq, tau pairs, mu pairs, Zγ, Zh:
- Web accessible:

http://www.lcsim.org/datasets/ftp.html

and Backgrounds

- Cain (to be done) & GuineaPig pairs and photons.
  - Add crossing angle, converted to stdhep
- Muons and other backgrounds from upstream collimators & converted to stdhep.
- γγ→ hadrons generated as part of the “$2ab^{-1}$ SM sample.”
- All events then capable of being processed through full detector simulation.
- Additive at the detector hit level, with time offsets, using LCIO utilities.

Simulation Summary

- ALCPG Sim/Reco team supports an ambitious detector simulation effort.
- Goal is flexibility and interoperability, not technology or concept limited.
- Provides full data samples for ILC physics studies.
  - Stdhep and LCIO files available on the web.
- Provides a complete and flexible detector simulation package capable of simulating arbitrarily complex detectors with runtime detector description.
- Reconstruction & analysis framework exists, core functionality available, individual particle reconstruction template developed, various analysis algorithms implemented.
- Need to iterate and apply to various detector designs.

"Benchmark Processes" for PFA Development

e+e- -> ZZ -> qq + νν @ 500 GeV

- Development of PFAs on ~120 GeV jets – most common ILC jets
- Unambiguous dijet mass allows PFA performance to be evaluated w/o jet combination confusion
- PFA performance at constant mass, different jet E (compare to ZPole)
- dE/E, dθ/θ -> dM/M characterization with jet E

e+e- -> ZH

e+e- -> ZZ -> qqqq @ 500 GeV

- 4 jets - same jet E, but filling more of detector
- Same PFA performance as above?
- Use for detector parameter evaluations (B-field, IR, granularity, etc.)

e+e- -> tt @ 500 GeV

- Lower E jets, but 6 – fuller detector

e+e- -> qq @ 500 GeV

- 250 GeV jets – challenge for PFA, not physics

-> 1 TeV?

PFA Performance

PFA developers meeting weekly – Wednesdays AM

Ron Cassell, Dhiman Chakraborty, Mat Charles, Ray Cowan, Norman Graf, Guilherme Lima, Steve Magill, Jose Repond, Marcel Stanitzki, Andy White, Lei Xia, Vishnu Zutshi

Manpower estimate ~3-4 FTE (ANL,NIU TestBeam Commitments, e.g. Lei is 100% testbeam now)

Topics normally discussed :

1) PFA Performance

2) Common development software tools

3) Comparison of algorithms and detectors

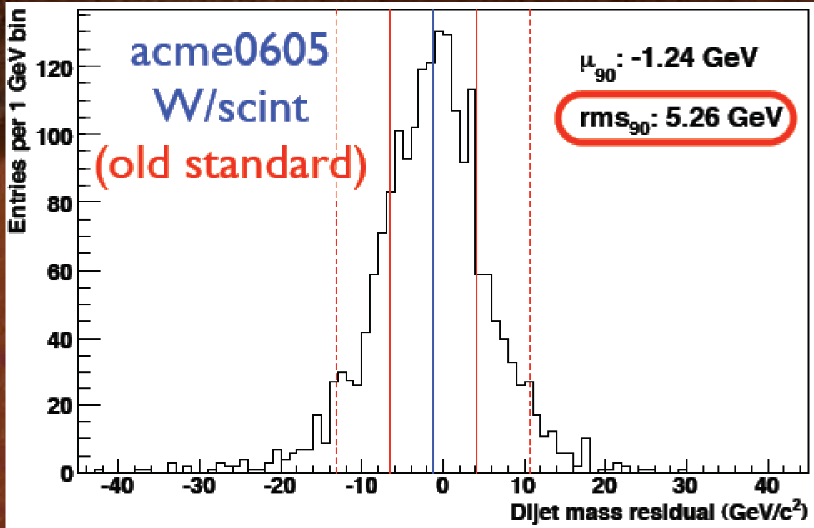
4) Definition/implementation of standard input/output

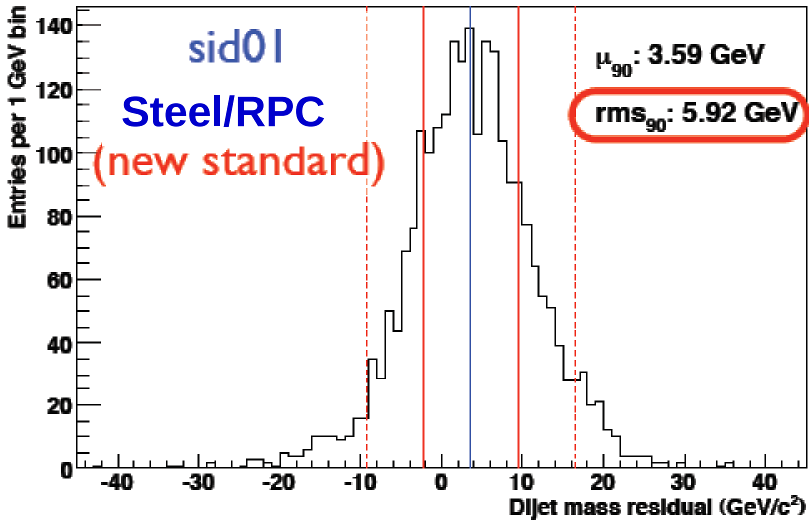
5) Timescales for performance measures

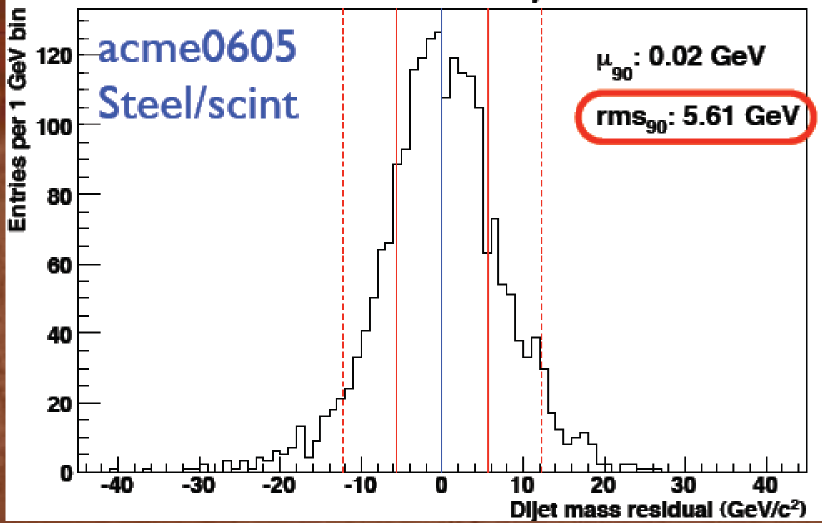
Structured Clustering Algorithm   Mat Charles Iowa

- 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

$e^+e^- \to Z(\nu\nu)\, Z(qq)$, q=u,d,s @ $\sqrt{s}$ = 500 GeV







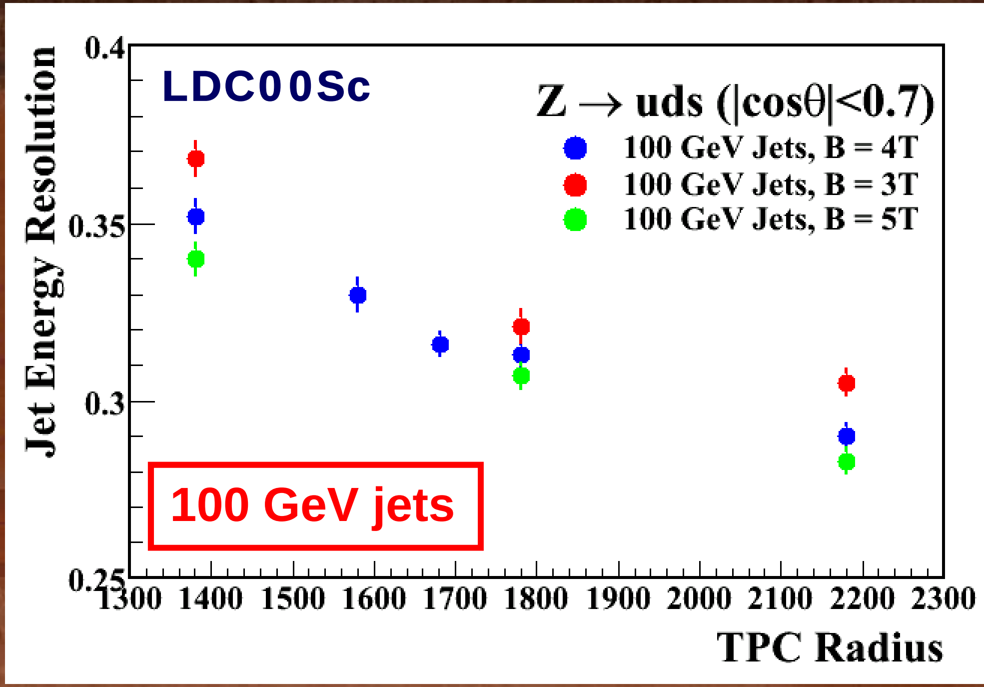
Something changed here & I lost ground.

It's not obvious what -- material details, extra layer in ECAL, artifact of algorithm, ...

- requiring primary quarks have |cos(theta)|<0.8
- reconstructing dijet invariant mass from reconstructed particles
- quoting residual = (true mass of Z - reconstructed mass of Z)

PANDORA PFA for SiD

Marcel Stanitzki RAL
Ray Cowan MIT



Original results from Mark Thomson (PANDORA Author)

Indicates that PFA results can reach ~34%/√E for an ~SiD-size detector

Run PANDORA on real SiD?

LDC00	->	LDC00Sc
B-field 4 T		4-6 T
Tracker radius 1.7 m		“SiD-ish”
Barrel length 2.7 m		“
Si/W 40 layer ECAL 10X10 mm pads		30 layers
Steel/Scintillator 40 layer analog HCAL 3X3 cm pads		

Results of PANDORA on Scaled LDC

From Marcel

Configuration	n/sqrt(E)	Jet energy
LDC00Sc	30.5	45
LDC00Sc 5T	31.2	45
LDC00Sc 30 layer ECAL	32.4	45
LDC00Sc Sid-ish 4T	32.6	45
LDC00Sc Sid-ish 5T	32.0	45
LDC00Sc Sid-ish 6T	33.8	45
LDC00Sc	36.7	100
LDC00Sc Sid-ish 4T	42.7	100
LDC00Sc Sid-ish 5T	41.0	100
LDC00Sc Sid-ish 6T	39.8	100

Errors ± 0.2-0.3

100 GeV Numbers very preliminary

Energy sum of reconstructed particles for $e^+e^- \to qq$ @ 91, 200 GeV

Recent (and some not so) R&D Efforts – Reconstruction/PFA Development Tools

Towards Like Comparison of PFA Results

- Detector Model starting point - realistic hits
  - Digisim package (NIU)
  - Random/correlated noise, energy thresholds, timing
- Standard Detector Calibration (SLAC)
  - ECAL (analog) and HCAL (digital)
- Perfect PFA definition (SLAC+All)
  - Standard realistic cheated Tracks, cheated clusters
- PFA Development Tools (All)
  - Template (ANL+All)
  - Cluster Algorithm standardization/comparison (SLAC)
  - PFA RP comparisons (SLAC)
    - Purity and efficiency of particle ID

Standard Perfect PFA (Perfect Reconstructed Particles)

Takes generated and simulated MC objects, applies rules to define what a particular detector should be able to detect, forms a list of the perfect reconstructed particles, perfect tracks, and perfect calorimeter clusters.

Complicated examples :
-> charged particle interactions/decays before cal
-> photon conversions
-> backscattered particles

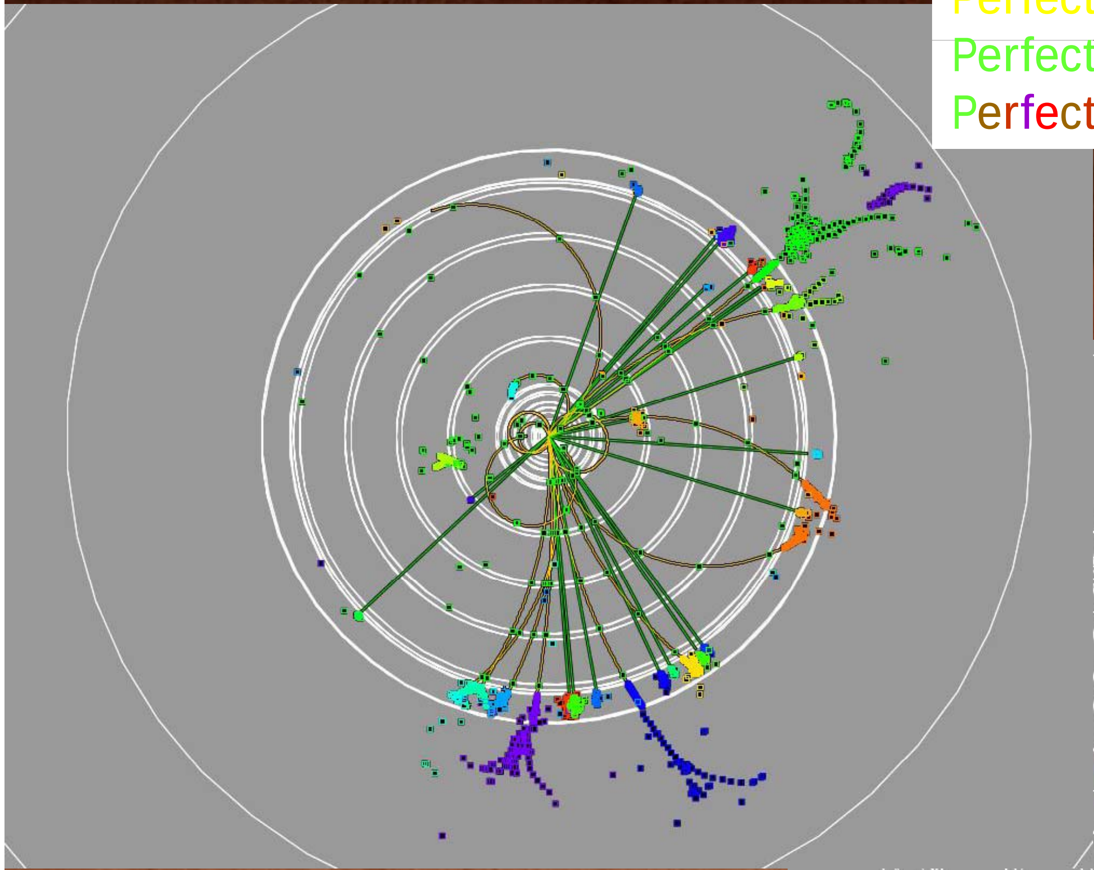
Critical for comparisons when perfect (cheated) tracks are used
Extremely useful for debugging PFA

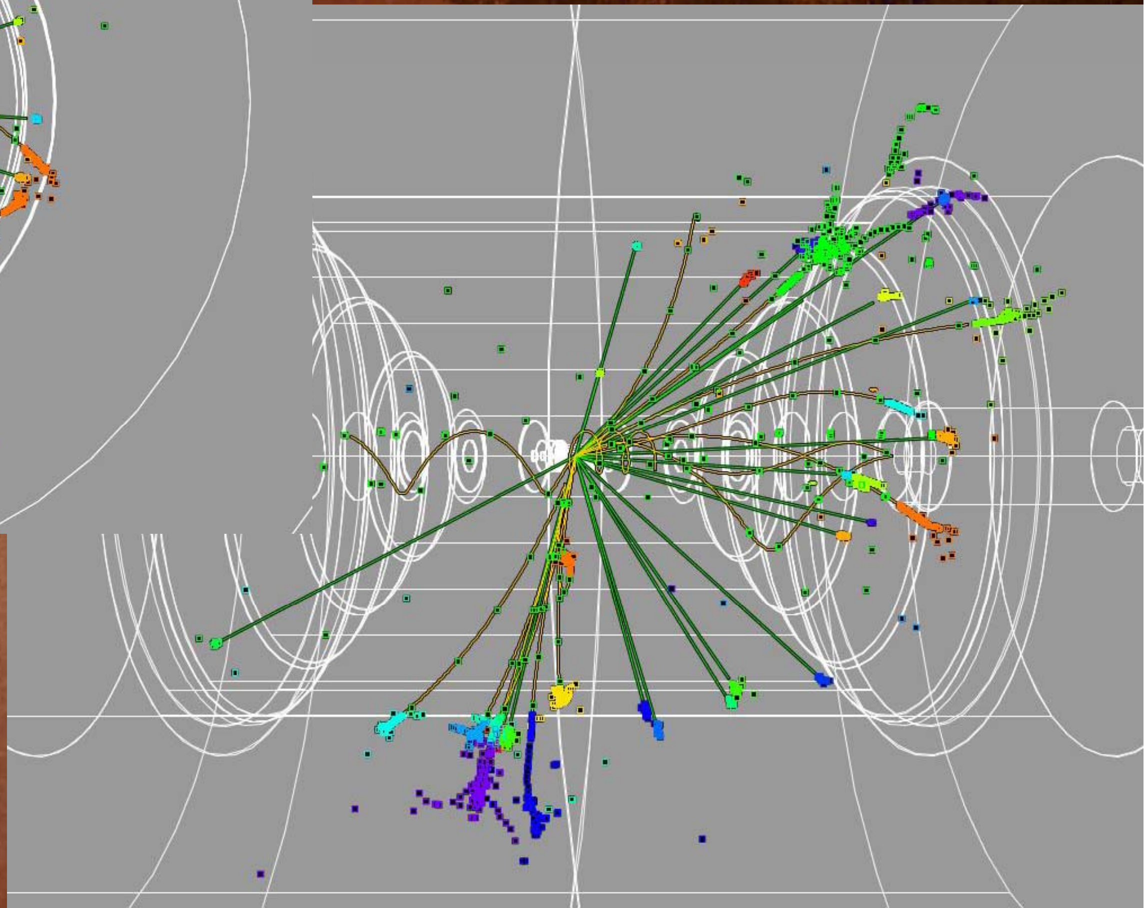
Standard Detector Calibration

Default detector calibration done with single particles
Basic Clusters contain calibrated energies – analog in ECAL and digital in HCAL
Standard for all SiD variants with analog ECAL, digital HCAL
Checked with Perfect PFA particles

Perfect PFA

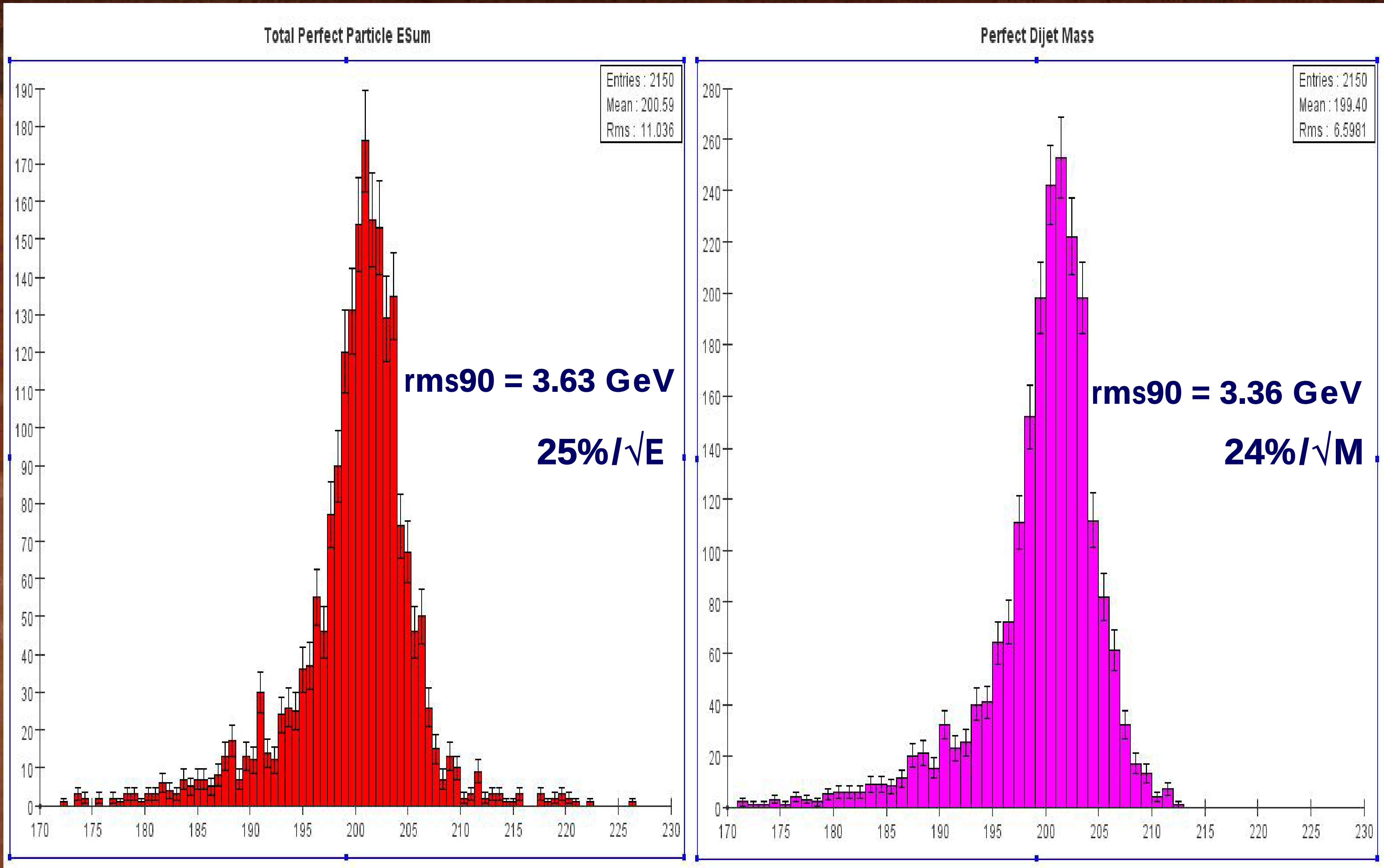
Perfect Tracks
Perfect Neutrals (photons, neutral hadrons)
Perfect Cal Clusters





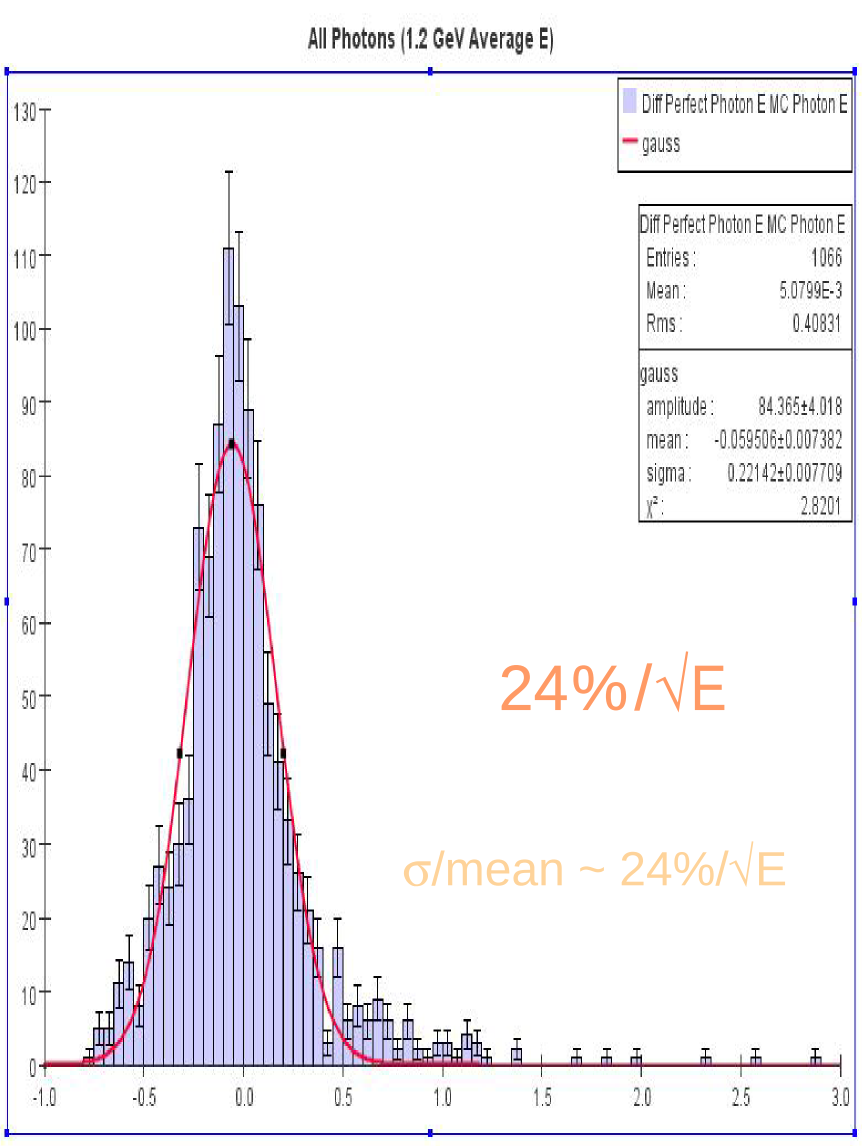
SiD (SS/RPC)
e^+e^- -> $Z(\nu\nu)$ Z(qq) @ 500 GeV

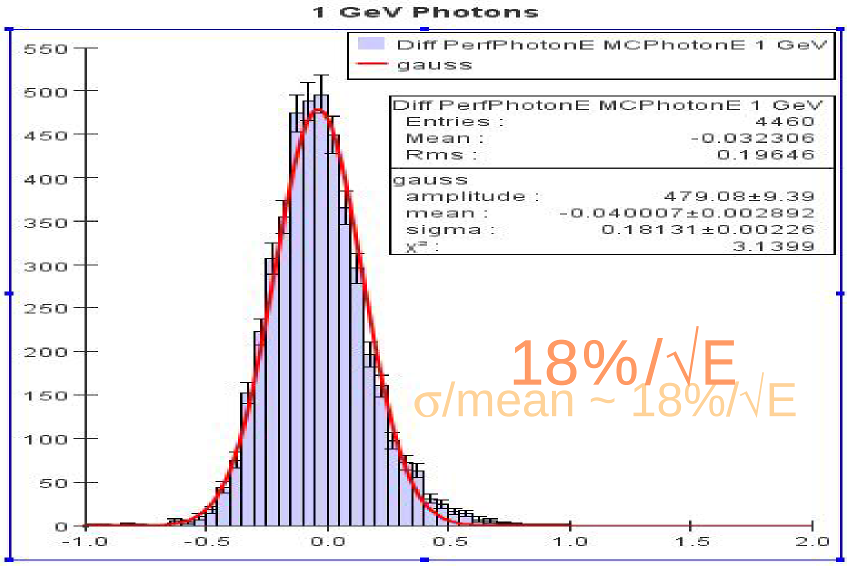
Perfect PFA – SiD01 e^+e^- -> qq @ 200 GeV

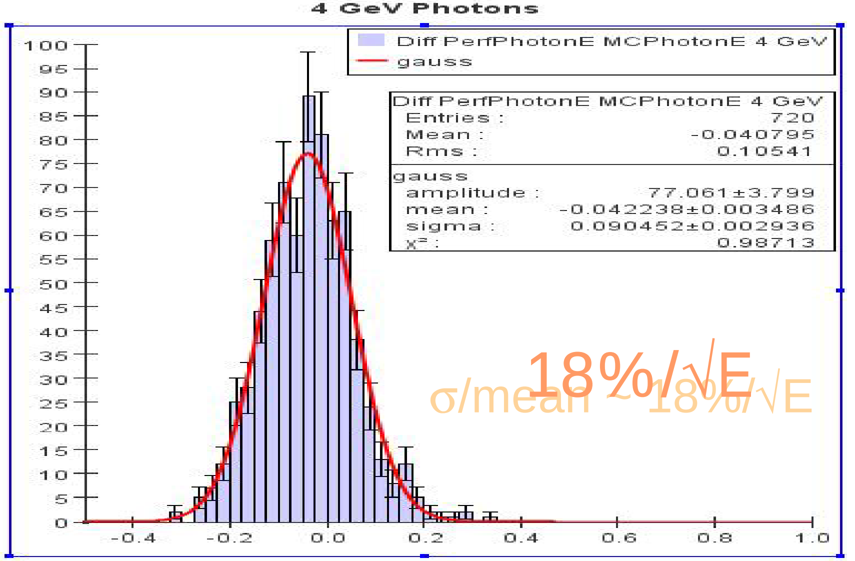


Detector Calibration Check

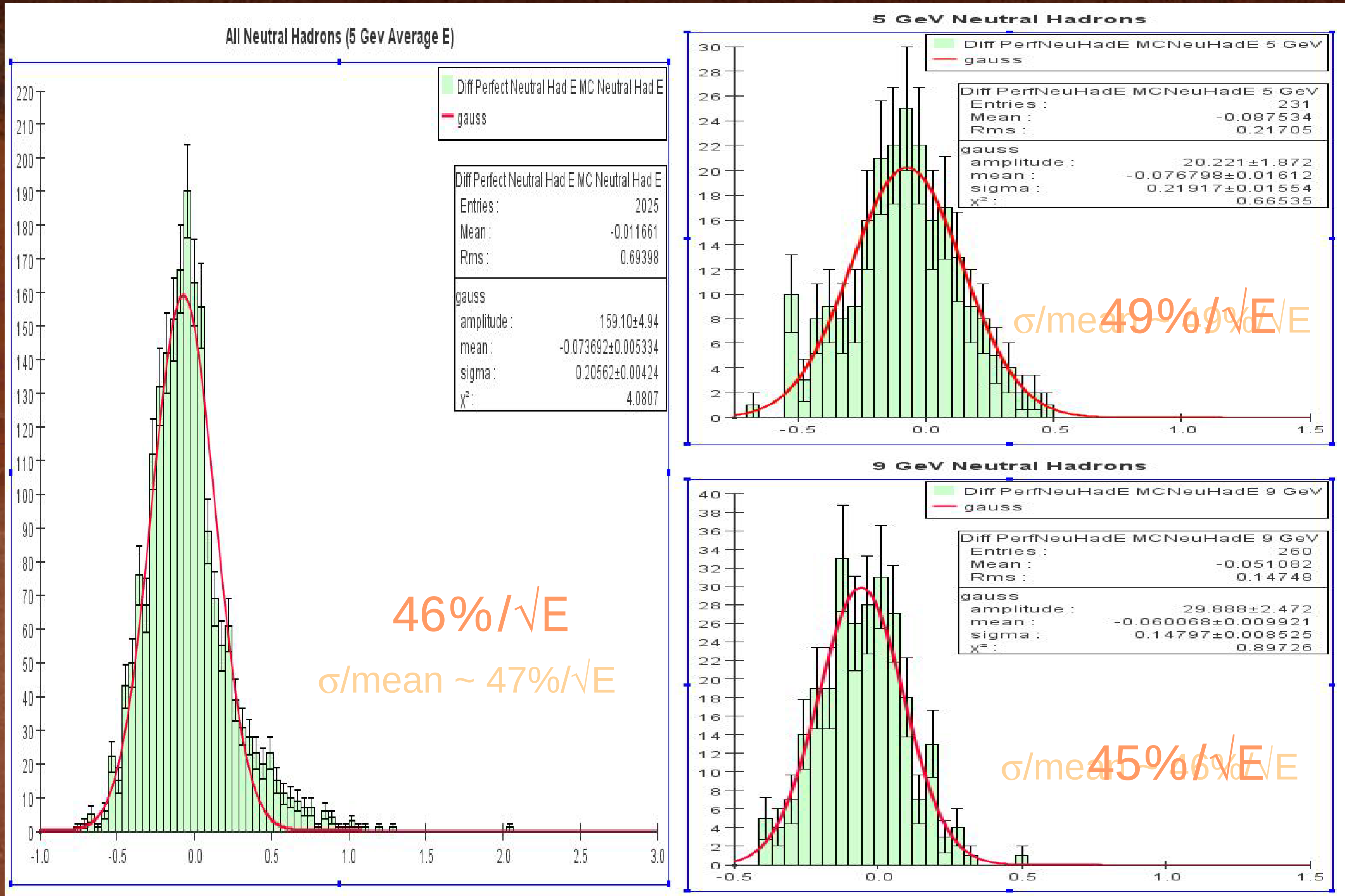
Photons from Perfect PFA (ZPole events in ACME0605 W/Scin HCAL)



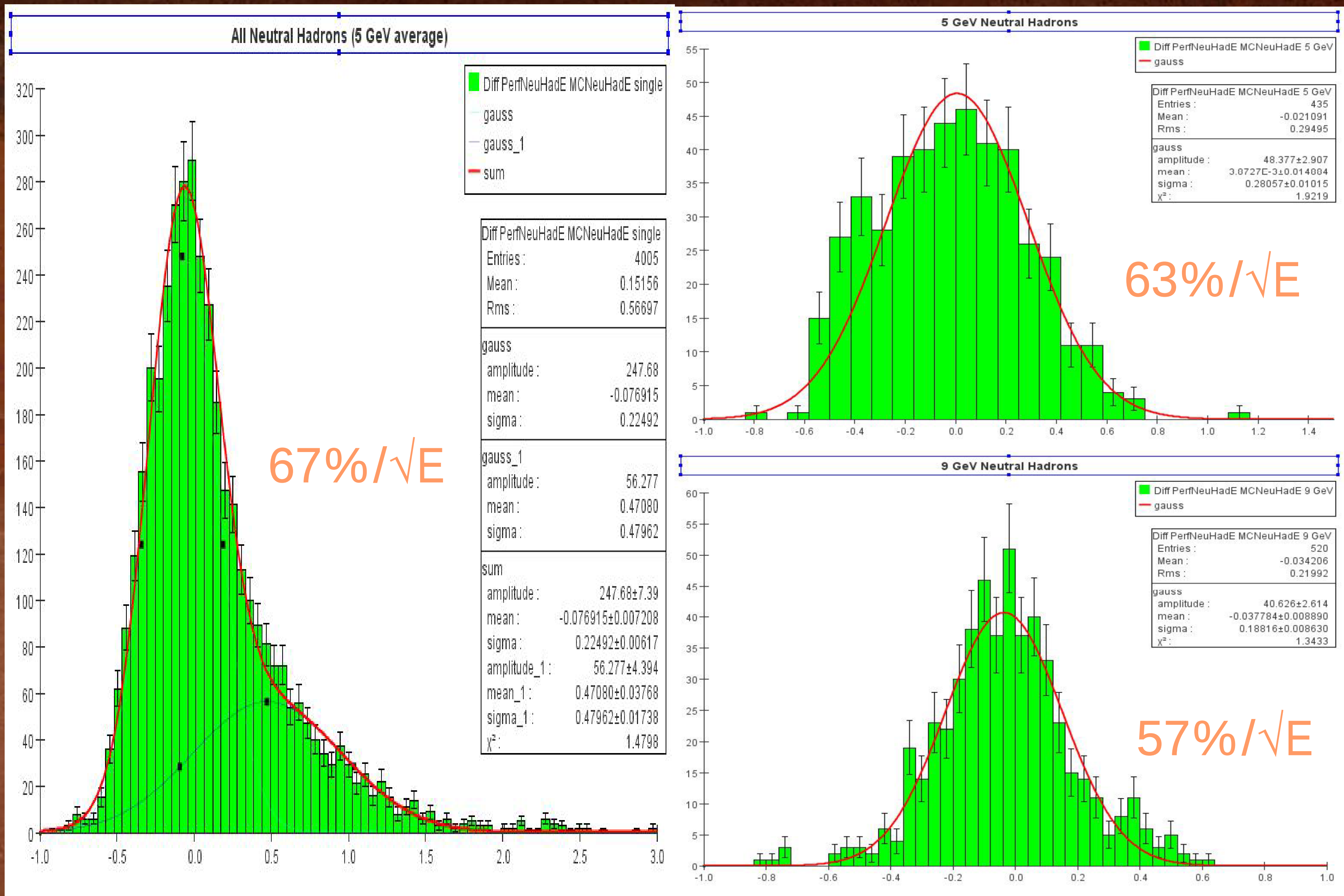




Neutral Hadrons from Perfect PFA (ZPole events in ACME0605 W/Scin HCAL)



Neutral Hadrons from Perfect PFA (ZPole events in ACME0605 SS/RPC HCAL)



PFA Reconstructed Particle Comparison Tool

From Ron Cassell, ran on Mat's PFA algorithm

- Rather than show plots, I'll explain in words what was done and Mat can object when I get it wrong.
- Run PFA, find ~6.5% deltaM/M. For cal hit assignments found:
- Photons: eff = 63%, pur = 83%
- Nhad: eff = 82%, pur = 27%
- Chhad: eff = 58%, pur = 92%
- Replacing the photon finder, the photon eff and purity were both ~ 85%. Not surprisingly, the mass width didn't change, since 40% of the tracks were being measured with the calorimeter!
- So it looked easy! The charged track association was very poor, so fixing that should be a big help

Towards Future Developments

Continue development of multiple PFA algorithms for SiD Optimization

Standardized on the following items :

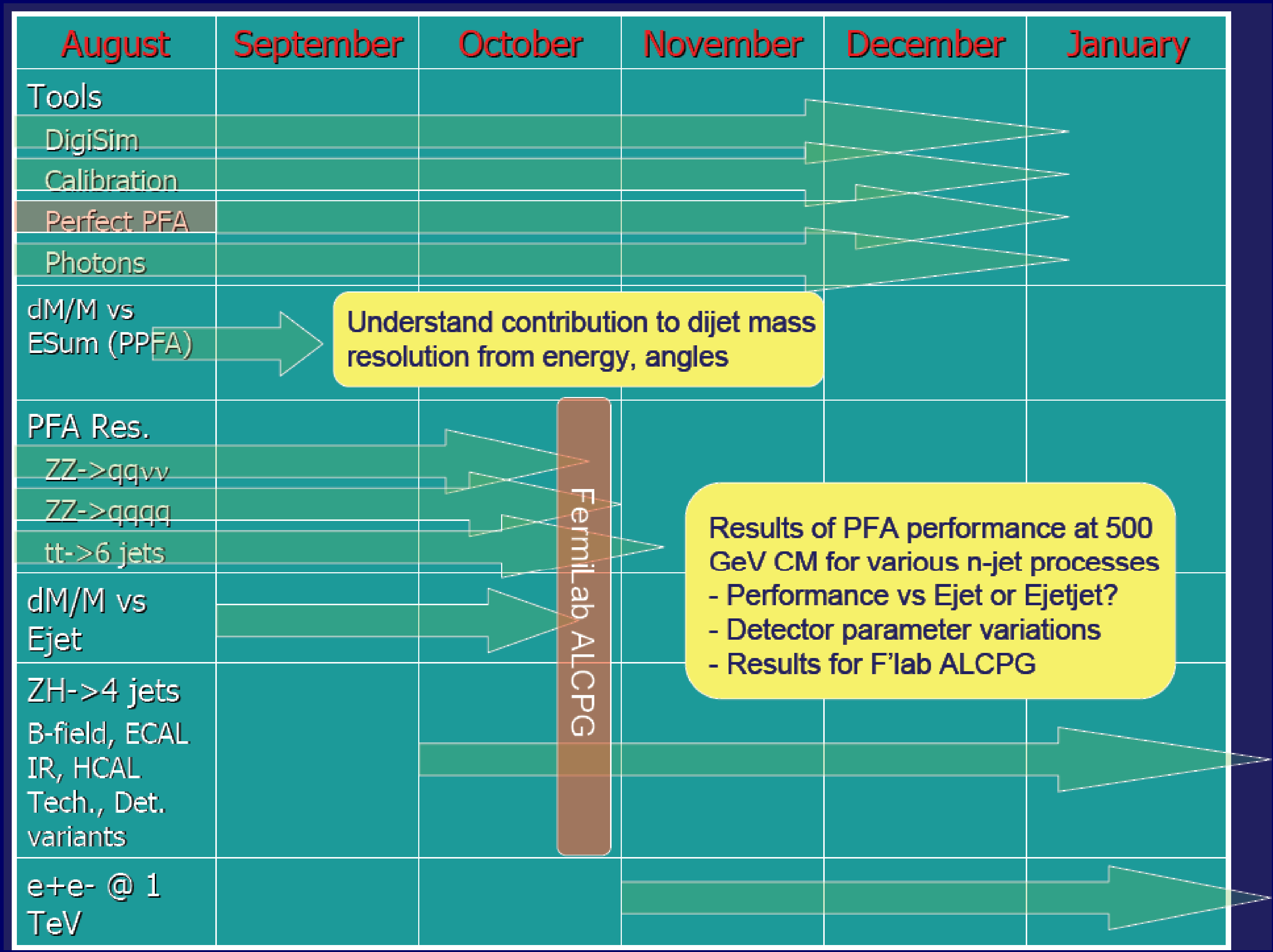
- Common calorimeter (and tracker) hit digitization
- Detector Calibration
- Perfect PFA Definition
- Reconstructed Particle IO format (PFA output)

Chose a common PFA Benchmark to make like comparisons possible
e+e- -> Z($\nu\nu$) Z(qq) @ 500 GeV

- PFA comparisons with total particle invariant mass
- Dijet mass comparisons for physics performance

Defined some milestones and plans for continued development – goal is several PFA alternatives for use in SiD Optimization

PFA Development Timeline



Manpower/Studies Needed

Attracting new help in the PFA development area :

A full PFA is a complicated analysis method involving many variables which are sometimes correlated, requiring time-consuming systematic studies to optimize its performance – however breaks in the correlations can be identified and exploited by modularizing pieces of the PFA.

The PFA Template is a modular structure which recognizes that these separations exist and is, therefore, conducive to incremental work projects and R&D studies on small pieces of the whole analysis.
- ideal for getting started in PFAs, undergrad projects, physicist part-time involvement

Some Studies of this type that are needed :

- Photon-finding in a dense fine-grained ECAL -> π^0 ID (Kansas)
- H-Matrix training and optimization for low, high energy photons
- Use of cluster pointing in matching, especially neutral hadron fragments
- Neural net cluster shape analysis
- PFAs with real reconstructed tracks

You and your student can do this!