Improving Prompt EM Energy Component of Jet Energy Resolution with π^0 Fitting

Brian van Doren
University of Kansas

LCWS11
29 Sep 2011

Outline

- Previous Single π^0 Study
- Multiple π^0's Using Truth Information
  - Z^0 Study
- Reconstruction without Truth Information
  - Procedure
  - Matching Algorithms
  - Performance
- Conclusion and Future Work

The Software Environment

- Generation: π^0 4-Vectors towards the barrel of ILD_00
  45° < θ < 135°

- Simulation: MOKKA – Geant4
  ilcsoft v01-09

- Reconstruction: Marlin framework
  - 1) Pandora Particle Flow Analysis
    - Reconstruction of 4-vectors of all visible particles
    - Identification of particle (photon, electron, neutron, etc...)
  - 2) pi^0 mass constrained fitting using MarlinKinFit
    - Implemented as a Pandora algorithm

Mass Constrained Fit

- A quick reminder...

Given process $\pi^0 \rightarrow \gamma_1 + \gamma_2$

We apply mass of π^0 as constraint C. Then minimize S by adjusting x^f subject to C.

$$C = (p_{\gamma 1} + p_{\gamma 2})^2 - m_{\pi^0}^2 = 0 \qquad S = \sum \left(\frac{x_i^{(m)} - x_i^{(f)}}{\sigma_i} \right)^2$$

Our case using E, θ, ϕ

$$S = \left(\frac{E_1^{(m)} - E_1^{(f)}}{\sigma_{E1}} \right)^2 + \left(\frac{\theta_1^{(m)} - \theta_1^{(f)}}{\sigma_{\theta 1}} \right)^2 + \left(\frac{\phi_1^{(m)} - \phi_1^{(f)}}{\sigma_{\phi 1}} \right)^2 + \left(\frac{E_2^{(m)} - E_2^{(f)}}{\sigma_{E2}} \right)^2 + \left(\frac{\theta_2^{(m)} - \theta_2^{(f)}}{\sigma_{\theta 2}} \right)^2 + \left(\frac{\phi_2^{(m)} - \phi_2^{(f)}}{\sigma_{\phi 2}} \right)^2$$

$$C = (p_{\gamma 1}^{(0)} + p_{\gamma 2}^{(0)})^2 - (p_{\gamma 1}^{(1)} + p_{\gamma 2}^{(1)})^2 - (p_{\gamma 1}^{(2)} + p_{\gamma 2}^{(2)})^2 - (p_{\gamma 1}^{(3)} + p_{\gamma 2}^{(3)})^2 - m_{\pi^0}^2 = 0$$

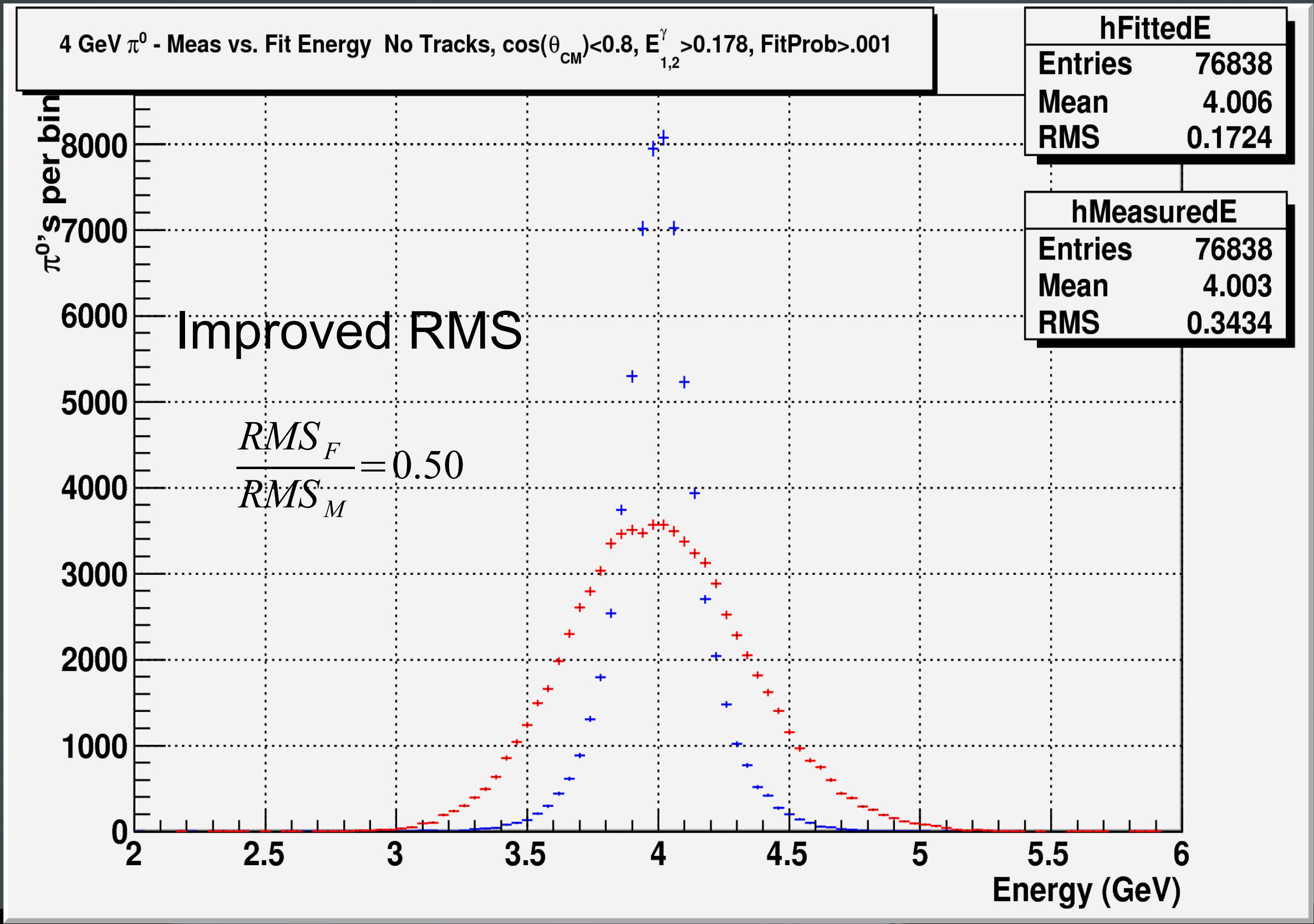
4.0 GeV π^0 Mass Constrained Fits

- Single π^0 in ILD_00 towards the barrel
- Efficiency of $\cos(\theta_{CM})$ cut: 84%

Relative to $\cos(\theta_{CM})$ cut

No Tracks	92%
Fit Prob >.001	98%
Low E Cut	99%
Combined	91%

Overall efficiency is 77%



Fitting Multiple π^0's

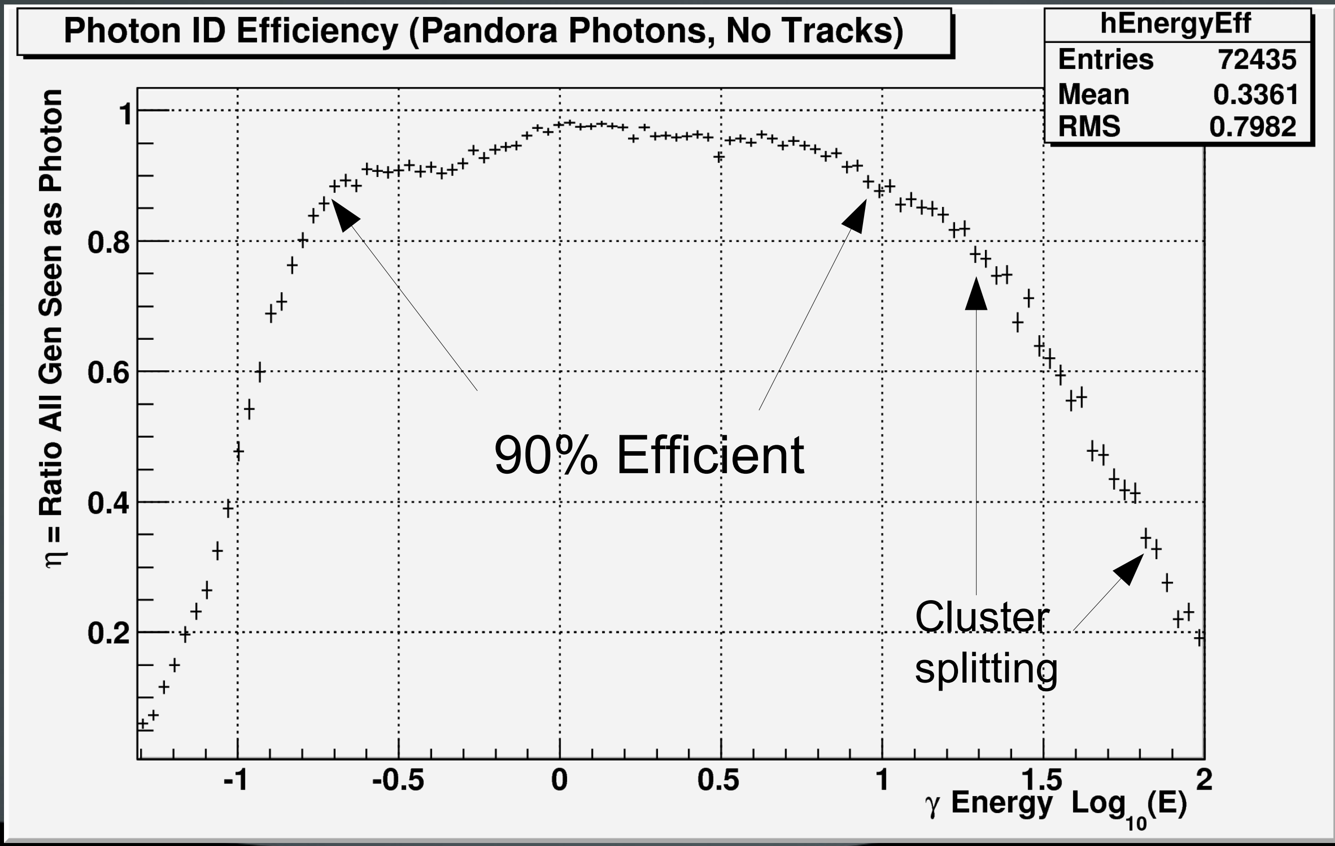
- What happens if we apply this to a more realistic situation? How well can we do?
- Consider 91.2 GeV Z^0 -> q q-bar, q = uds
- Extract and simulate the prompt π^0's and apply fitting procedure using truth information
- Reconstruction uses improved center of gravity position estimate
- Only match photons energy greater than 50 MeV
- Require 95% of energy deposited in barrel
- No tracks in the event

Fitting Multiple π^0's

- Overall efficiency of correctly detecting photons

$$\eta = \frac{\text{single PFO identified as photon}}{\text{all photon events with no tracks}}$$

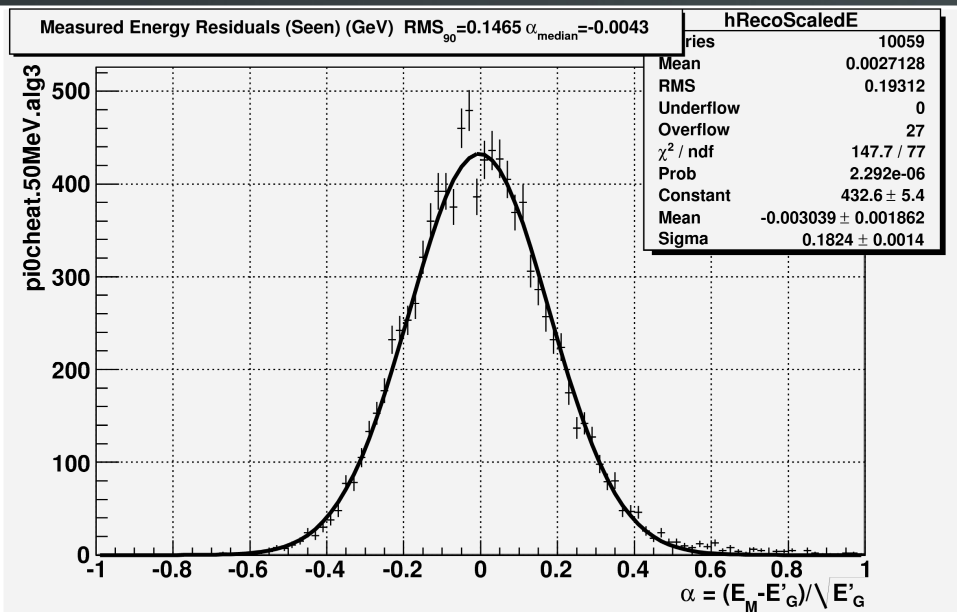
~90% Efficiency between 180 MeV < E < 10 GeV

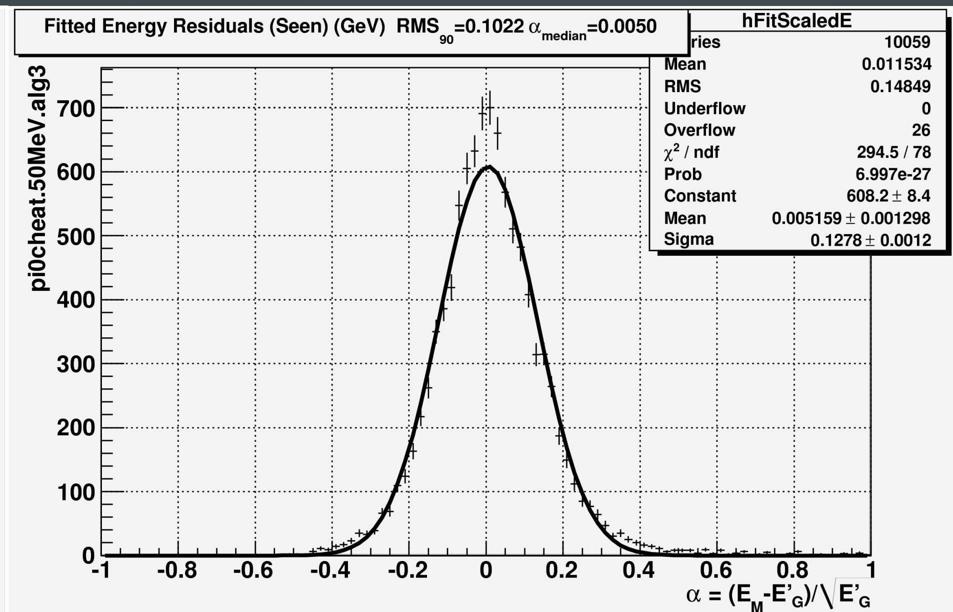


Fitting Multiple π^0's

Results of procedure on 91.2 GeV Z^0 -> q q-bar

(π^0 contribution only, 95% energy in barrel, 50 MeV minimum energy, no tracks)

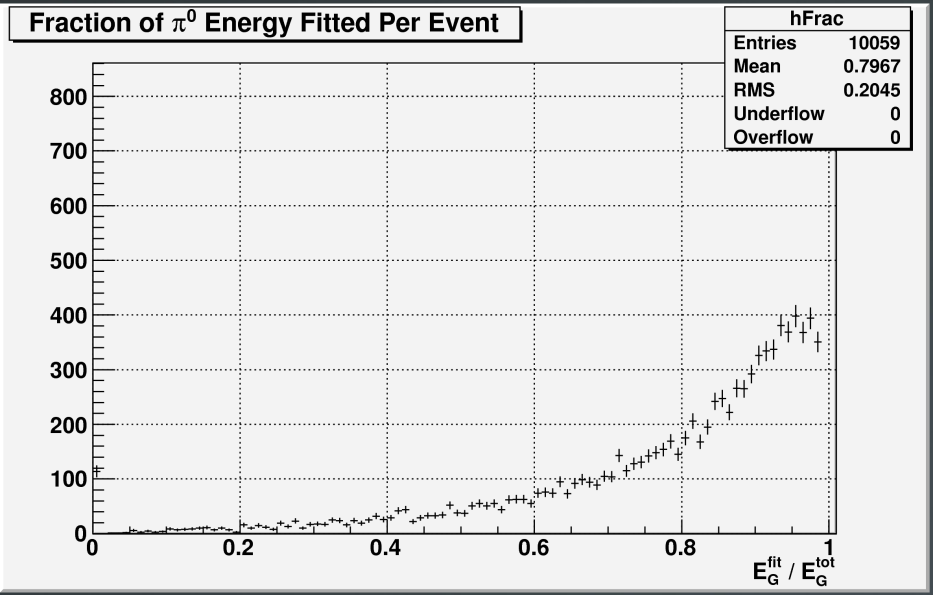


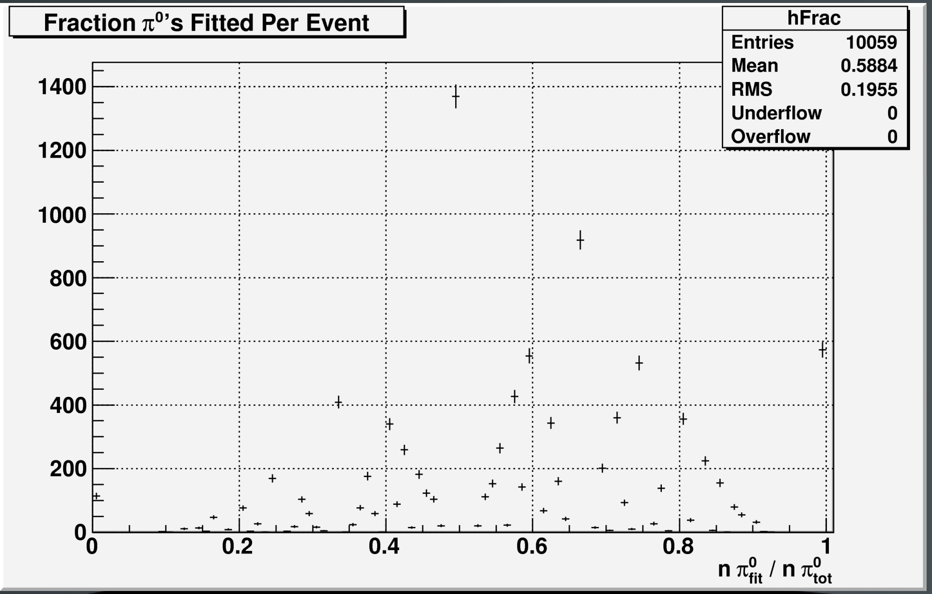


Improvement in α: .182 -> .128

(Using truth information)

Fitting Multiple π^0's





Fraction of overall energy that is fitted is 79%. Some loss due to 1% probability cut and also undetected photons.

Fraction of π^0's fitted (59%) suggests fitting favors the higher energy pions, this is likely due to lost low energy photons

Fitting Multiple π^0's

- Exploration of matching procedures that do **not** use truth information
- The challenge: Enumerate over all potential event solutions and determine the "best"
- Some basic restrictions:

  Minimum photon energy 50 MeV

  95% of energy deposited in barrel

  Accept potential fits with greater than 1% fit probability

  - No tracks

Fitting Multiple π^0's

- Photon Matching Procedure

  1 Perform kinematic fits on all photon pairs

  2 Remove fits where fit probability is less than 1%

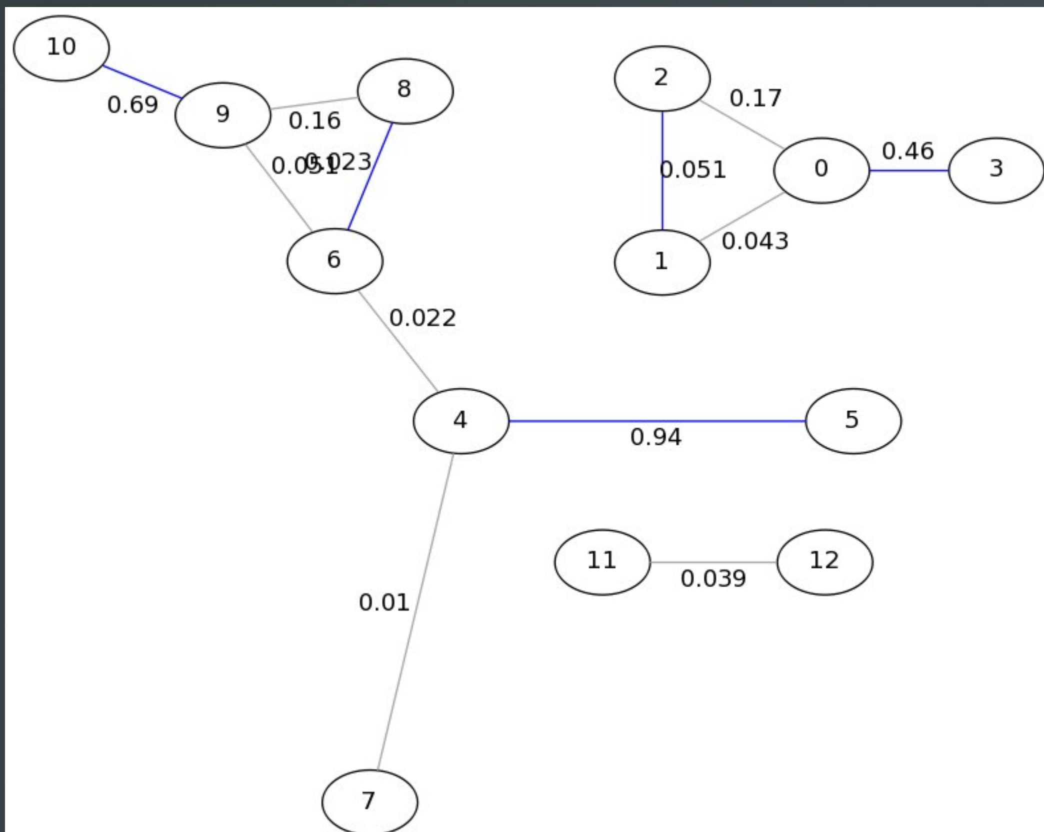
  3 Generate all potential solutions by combining remaining pairs such that each photon is used at most once

  4 “Score” each solution and pick the best

Fitting Multiple π^0's

- Photon Matching Procedure

The collection of all >1% pairs can be modeled as a graph with vertices and edges



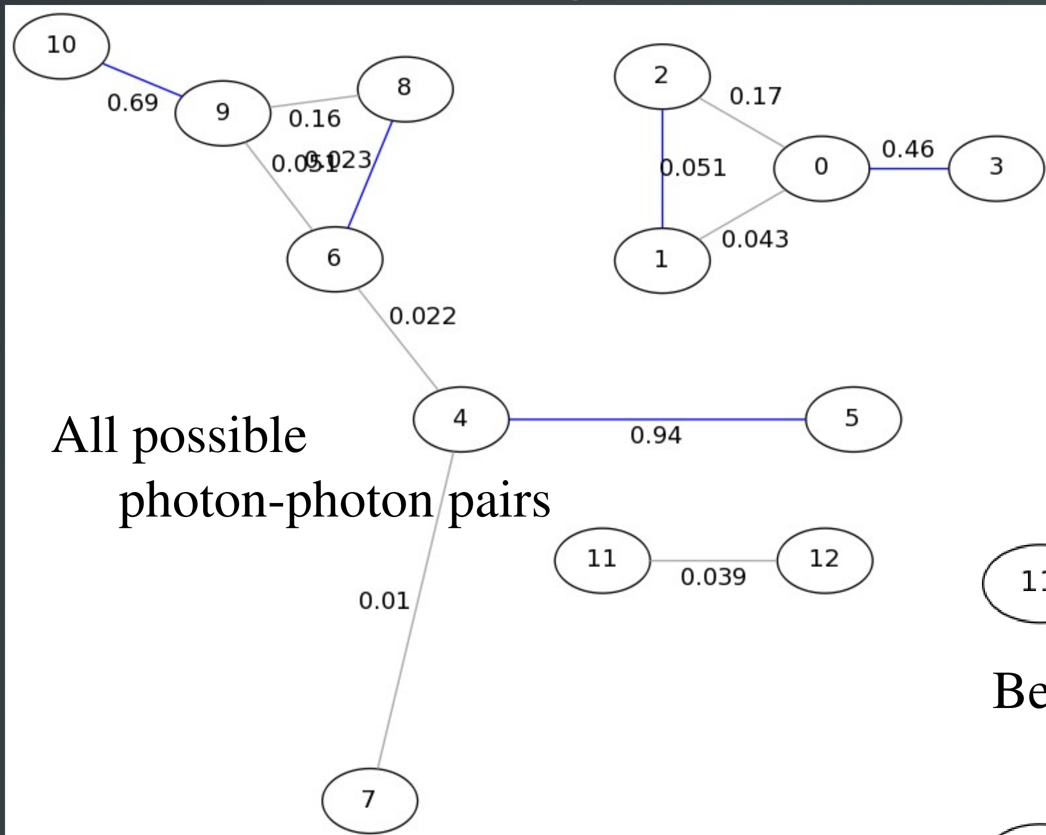
Vertices are photons

Edges represent fit probability between the photons

(correct edges are blue)

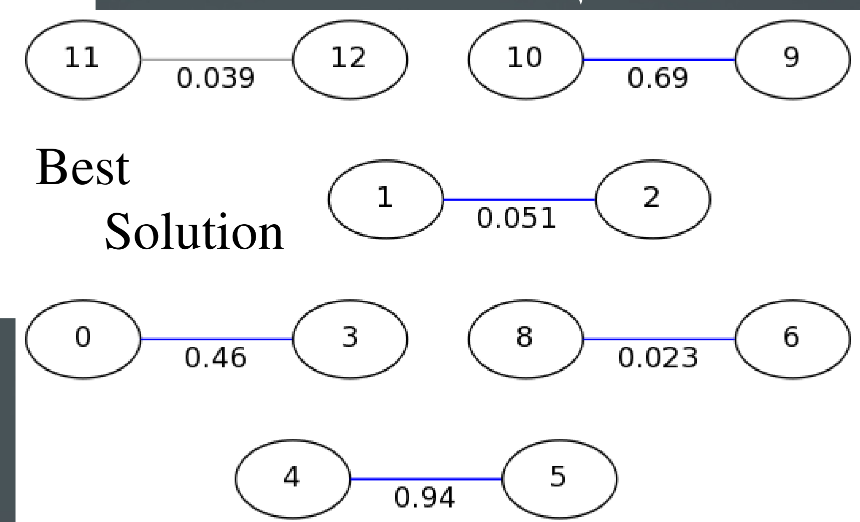
Fitting Multiple π^0's

- Photon Matching Procedure



Evaluate all solutions

blue = correct



- Final solution uses each photon at most once

Fitting Multiple π^0's

- Several ways to approach scoring of the solutions:

  Evaluated functions involving: fit probability, number of fits, overall χ^2.

  Best scoring method so far is to consider solutions with maximal fits and the lowest total χ^2

  Example:

  Solution a: 6 Fits, $\chi^2/6 = 5/6$

  Solution b: 7 Fits, $\chi^2/7= 8.2/7$

  Solution c: 7 Fits, $\chi^2/7 = 14/7$

  Best solution is "b"

Fitting Multiple π^0's

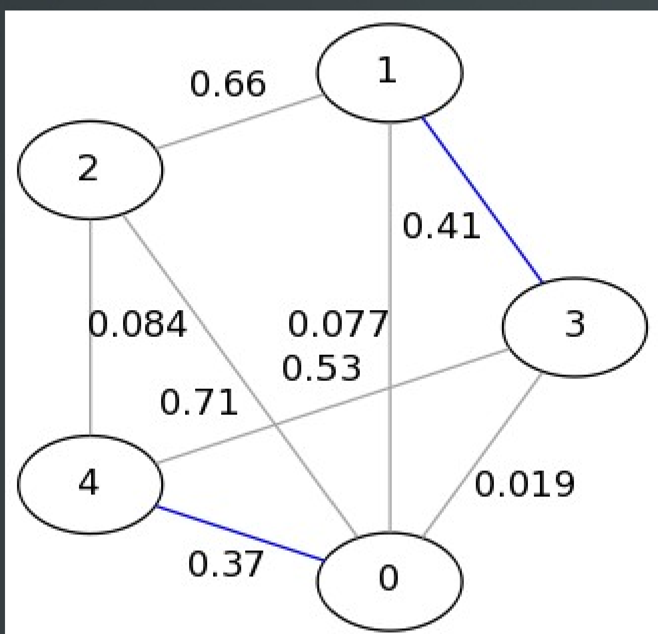
- How does this scale with high multiplicity? (i.e. many vertices and edges)
- We use the matching algorithm Blossom V.
  - Finds **perfect match** with minimum cost (χ^2)
  - For n vertices and m edges, worst case complexity is $O(n^3m)$ but on average is better than this
  - Most graphs require modification to guarantee perfect match exists

Vladimir Kolmogorov. Blossom V: A new implementation of a minimum cost perfect matching algorithm. Mathematical Programming Computation (MPC), July 2009, 1(1):43-67.

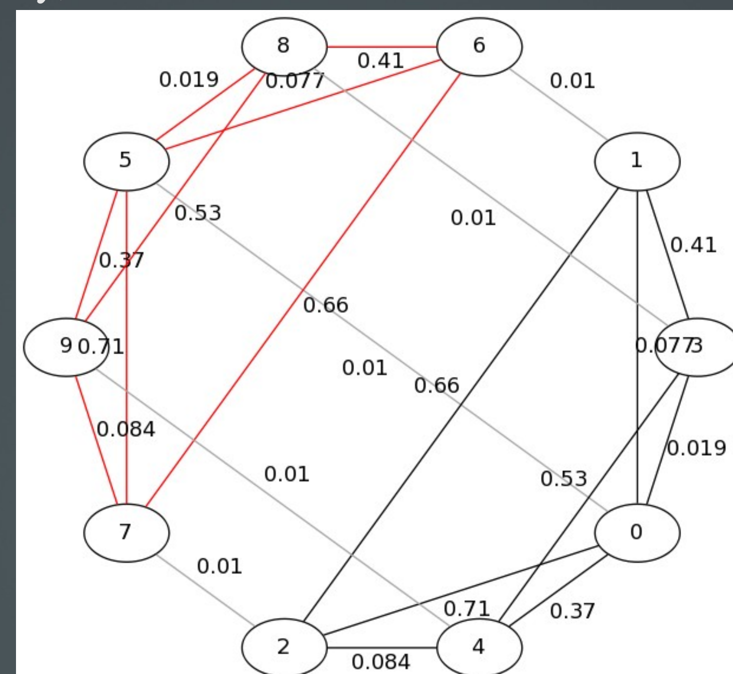
Fitting Multiple π^0's

- Modification to guarantee perfect match
  - Perfect match: Solution uses each vertex exactly once.
  - Most graphs from detector data do not allow this
  - Modify by duplicating graph and linking each vertex with its duplicate

G. Schäfer. Weighted matchings in general graphs. Master's thesis, Fachbereich Informatik, Universität des Saarlandes, Saarbrücken, Germany, 2000.

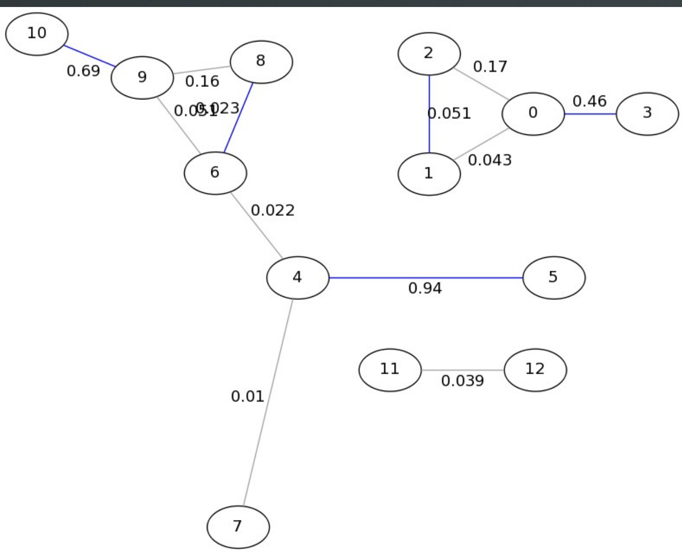


This allows photons to remain unmatched if necessary

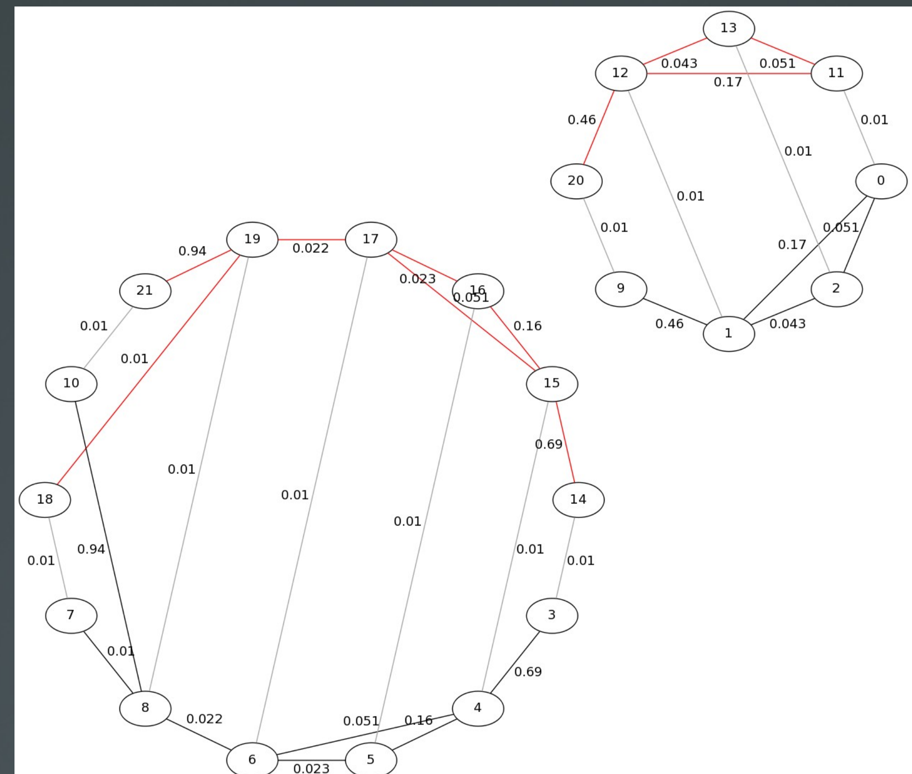


Fitting Multiple π^0's

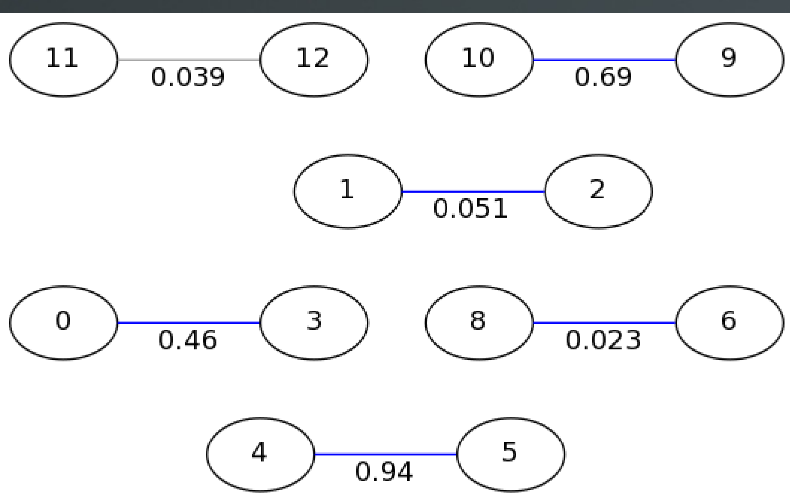
- The complete process



Modify graph



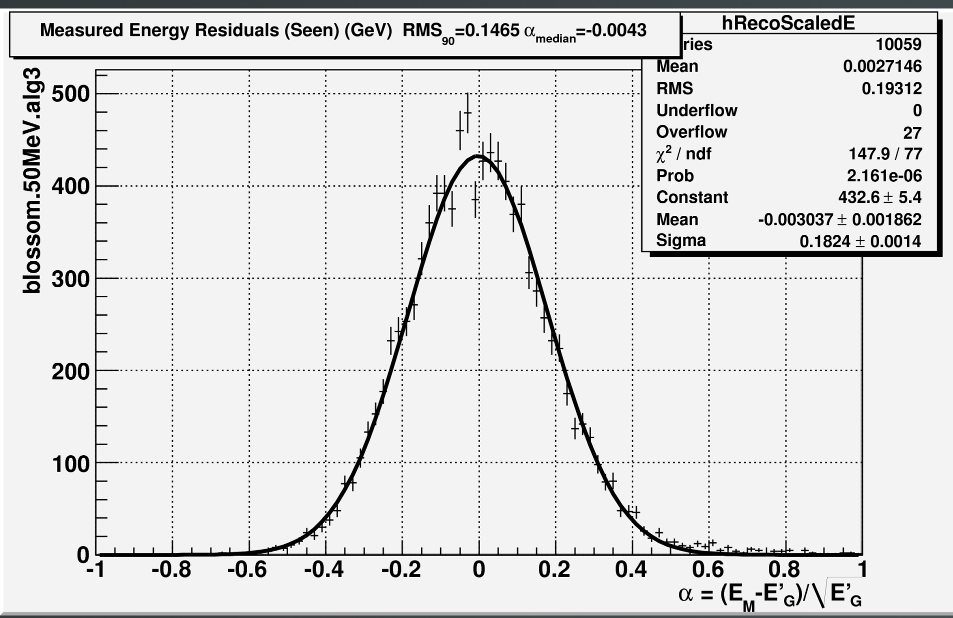
Best solution



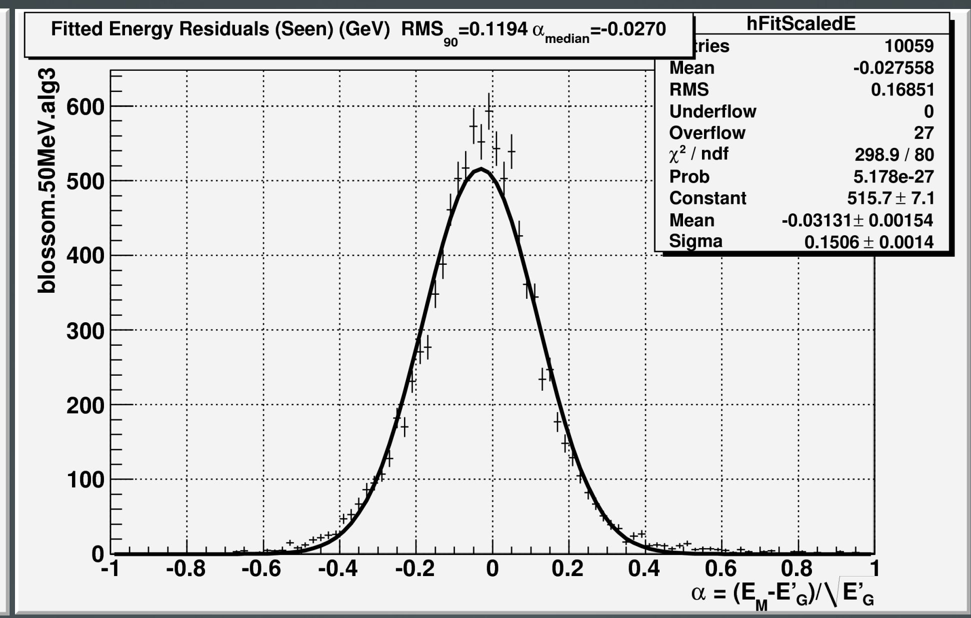
Fitting Multiple π^0's

- Fitting 91.2 GeV Z^0 using only π^0 photons

Reconstructed



Fitted: Blossom5, Max Fits, Min χ^2



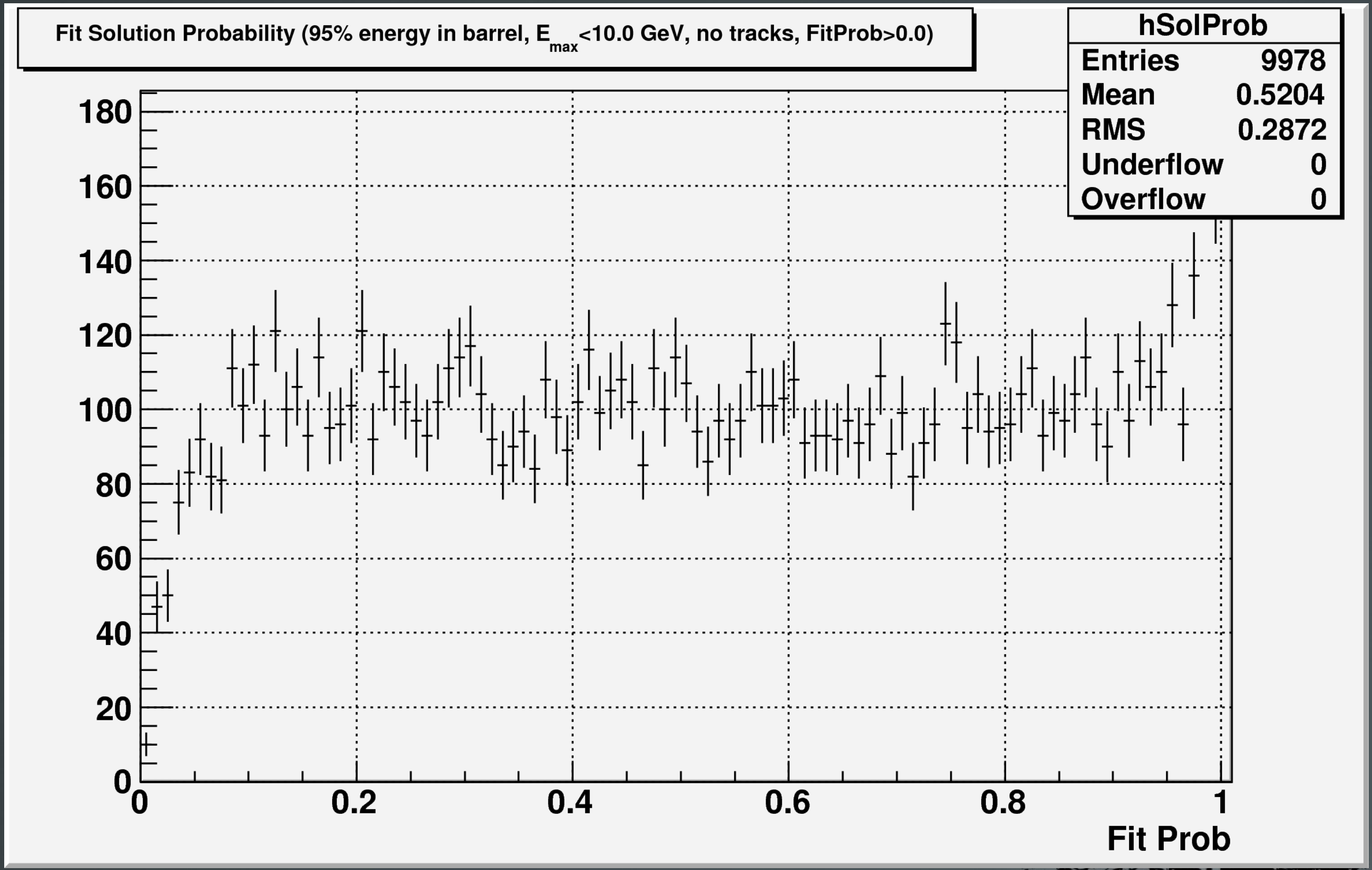
$\alpha = .182$ -> $\alpha = .151$

(much better than ALCPG11 numbers)

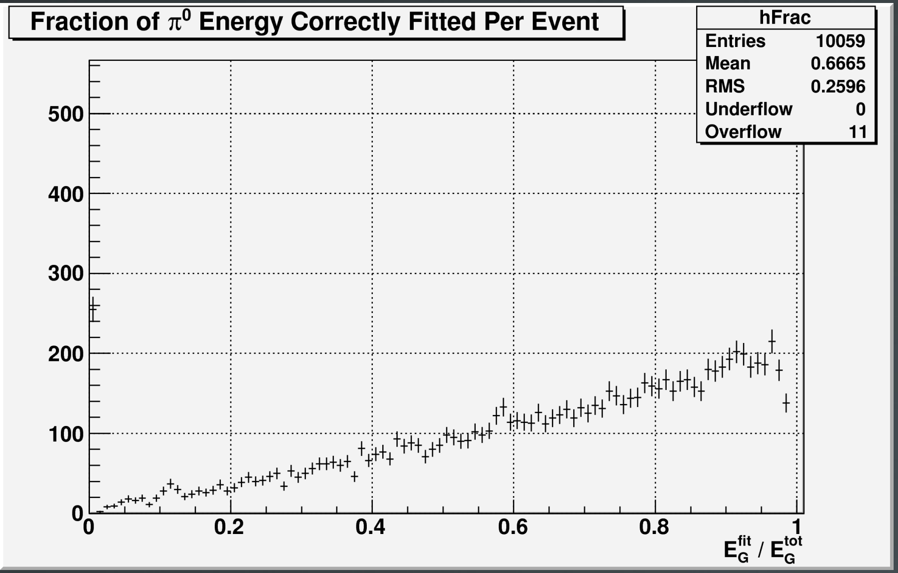
(recall best possible $\alpha = .128$)

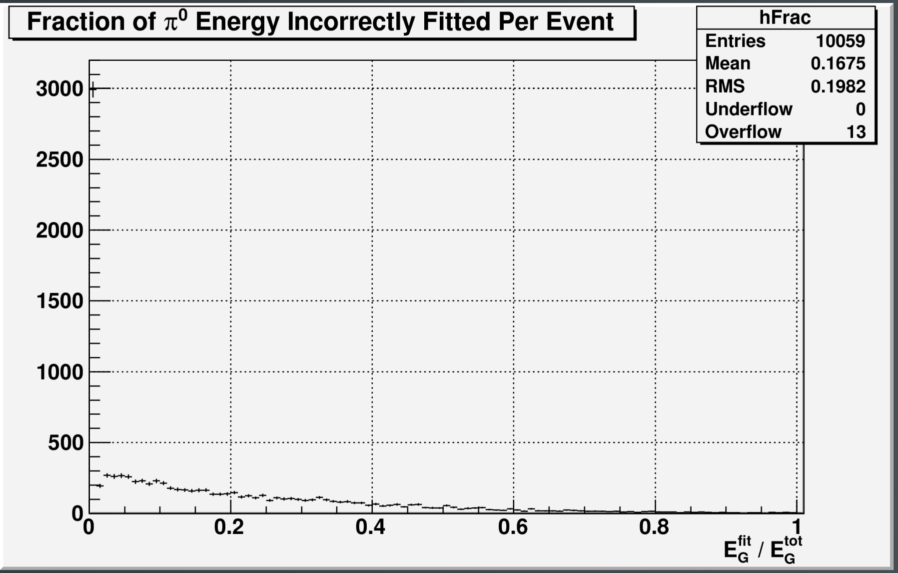
Fitting Multiple π^0's

- Overall solution probability is nearly flat, similar to when truth information is used.



Fitting Multiple π^0's





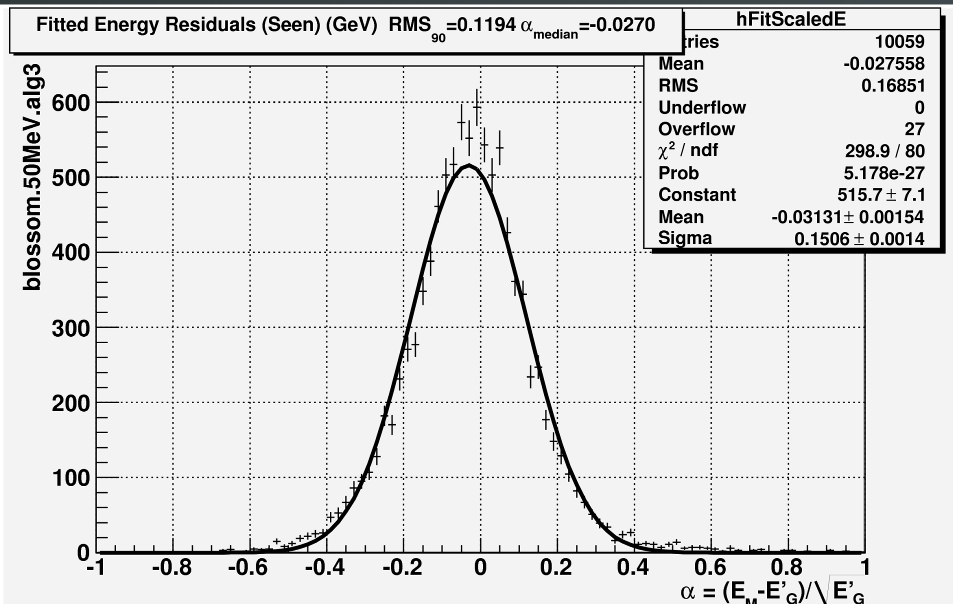
Fraction of overall energy that is correctly fitted is 67% (compared to 79% when cheating) while 17% is incorrectly fit.

What is the impact of incorrect fits?

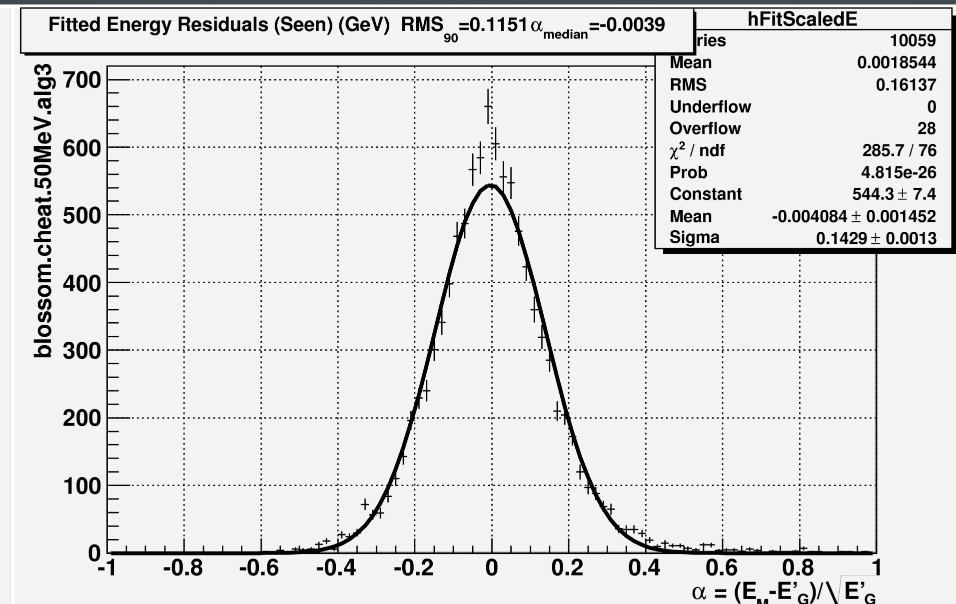
Fitting Multiple π^0's

- What is the impact of incorrect fits?

Blossom5, Max Fits, Min χ^2



Blossom5, Max Fits, Min χ^2, Remove incorrect fits



Primary impact is some worsening of resolution and a small bias in the energy.

Fitting Multiple π^0's

- Tuning the Algorithm for 91.2 GeV Z^0
  - To minimize fitted sigma, studied range of values for the following and found optimal:
    - Fit Probability Cut = 0.01
    - Single Photon Chi2 = 6.6348 (p = 0.01)
    - Minimum Photon Energy = ~50 MeV

      This is in region where photon detection is not efficient, but benefits still exist by contributing to overall matching solution.

Summary

- On an individual basis, mass constrained fitting can greatly improve energy resolution of a neutral pion

  17.2% to 8.7% at 4 GeV

- Application to multiple π^0's from Z^0 decay in ILD_00 sees significant improvement in energy resolution
  - From 18.2% down to 15.1% (compare to cheating 12.8%) using shower CoG cluster position estimate
- Further Study:

  Use additional information to inform the matching process
  Removal of incorrect fits
  Explore higher energies
  Cluster splits

Back up slides

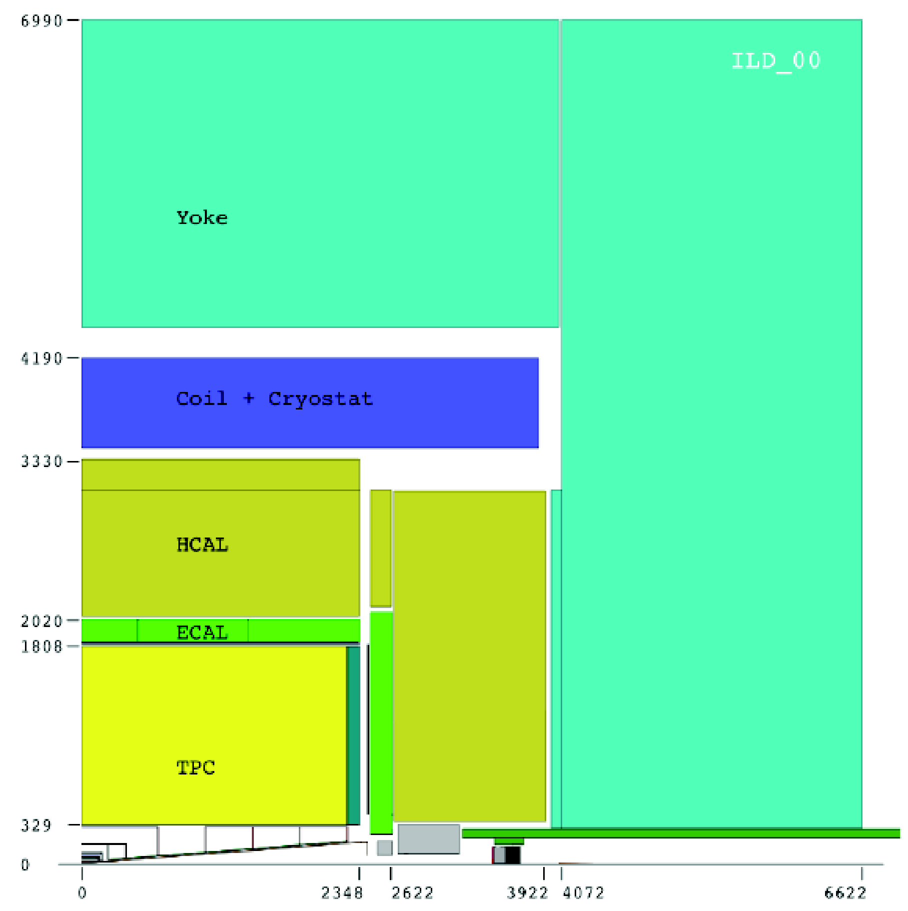
International Large Detector (ILD)

- Detector concept being studied for the International Linear Collider (electron-positron).

ECAL

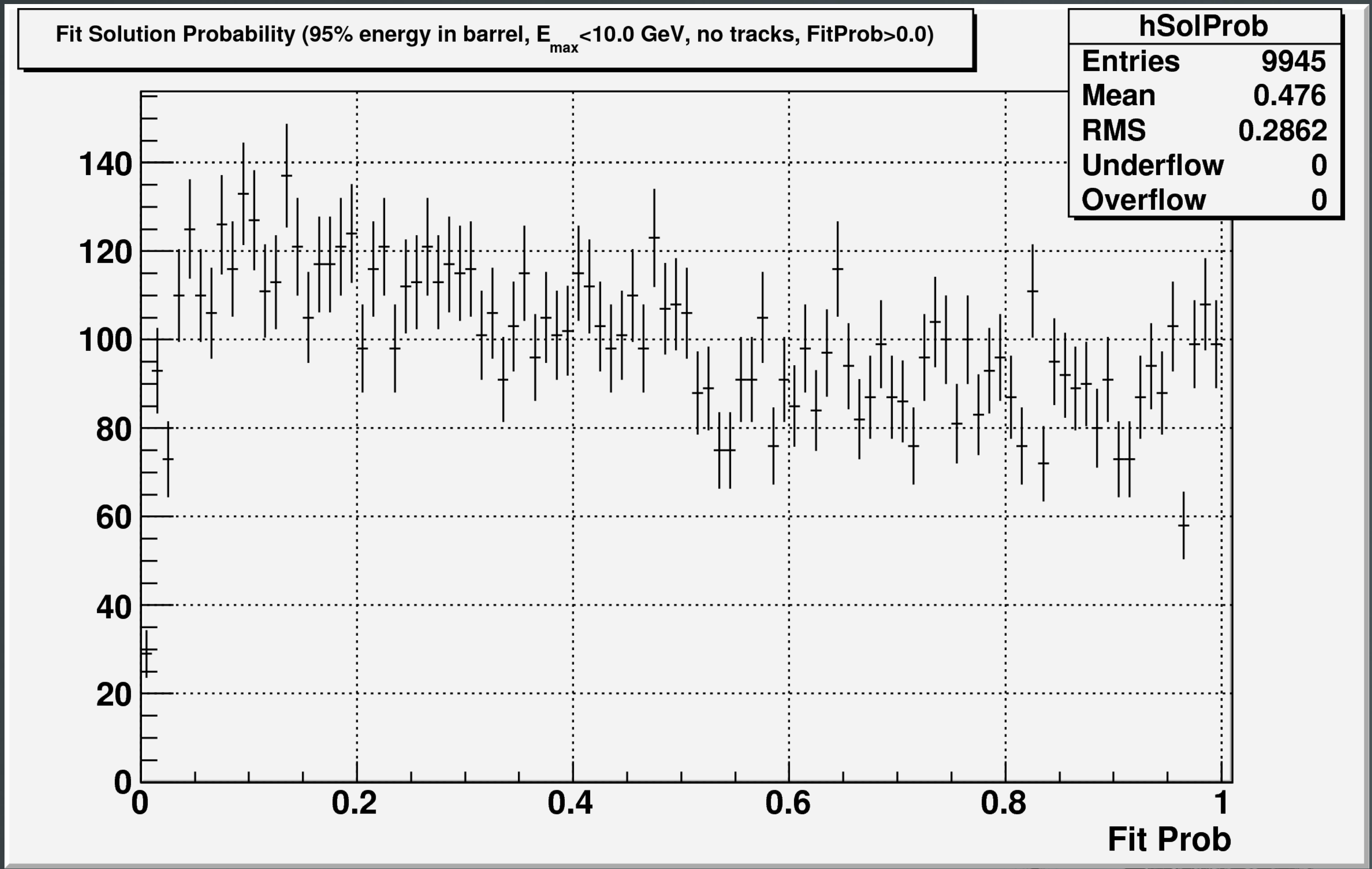
- 20+9 Layers Si-W
- Active layer segmented into 5mm x 5mm "*highly granular*"
- Typical photon uncertainties

$$\sigma_E = 16\% \sqrt{E}$$
$$\sigma_\phi = 1.2\, mrad\, @\, 1\, GeV$$
$$\sigma_\theta = \text{similar, but } \theta \text{ dependence}$$

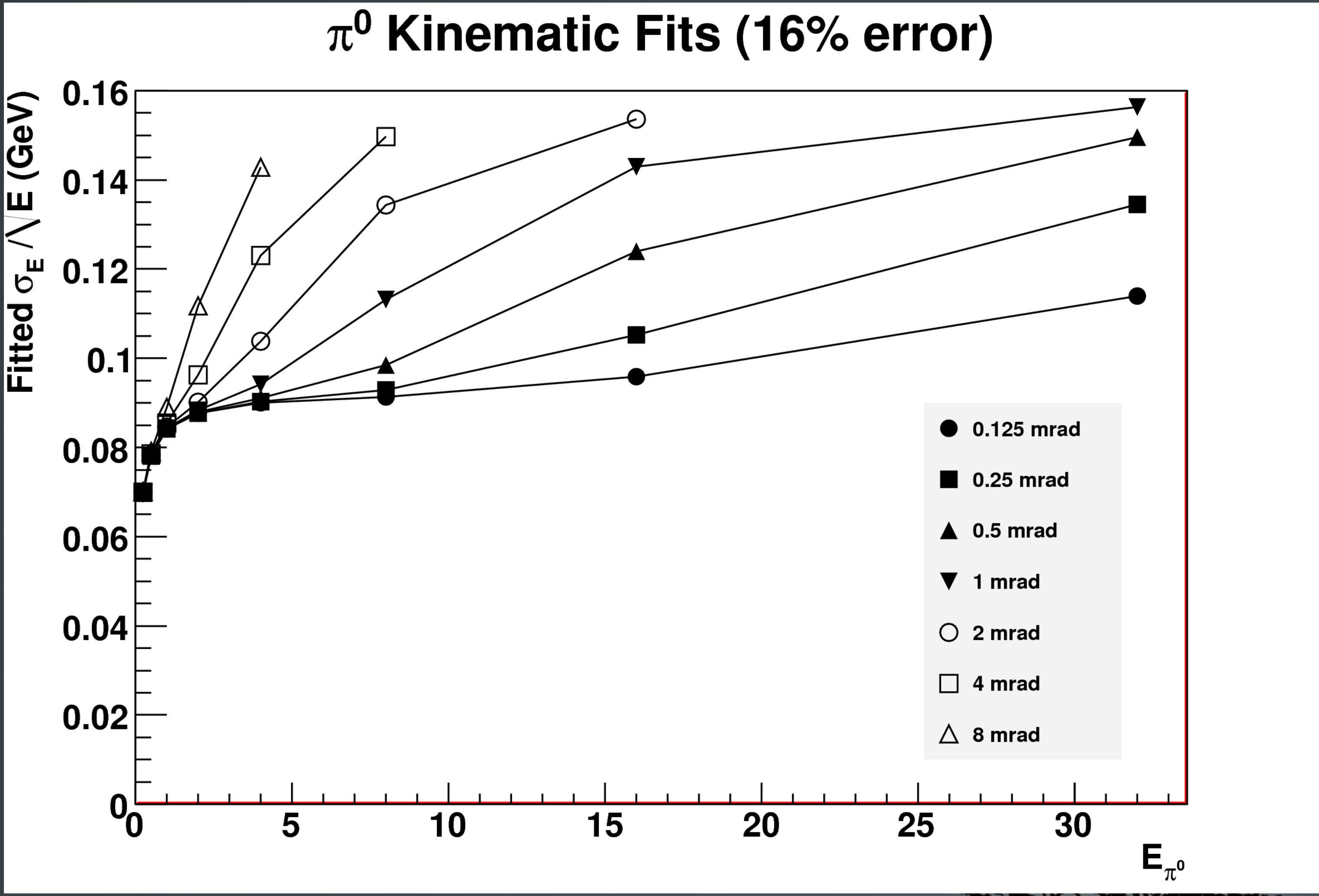


Fitting Multiple π^0's

Overall solution probability is reasonably flat when using MC truth information

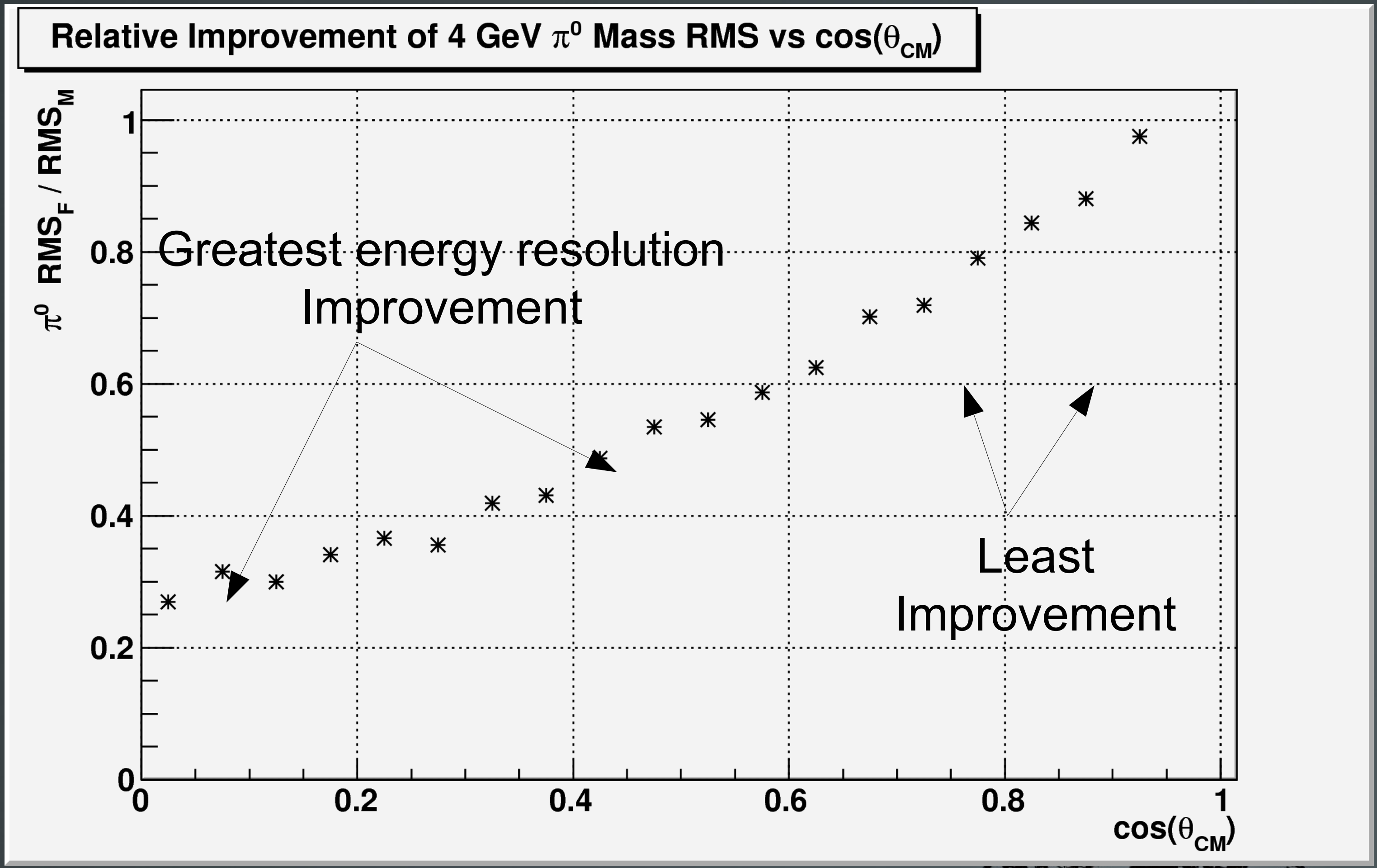


Fit performance by Energy and Angular Uncertainty



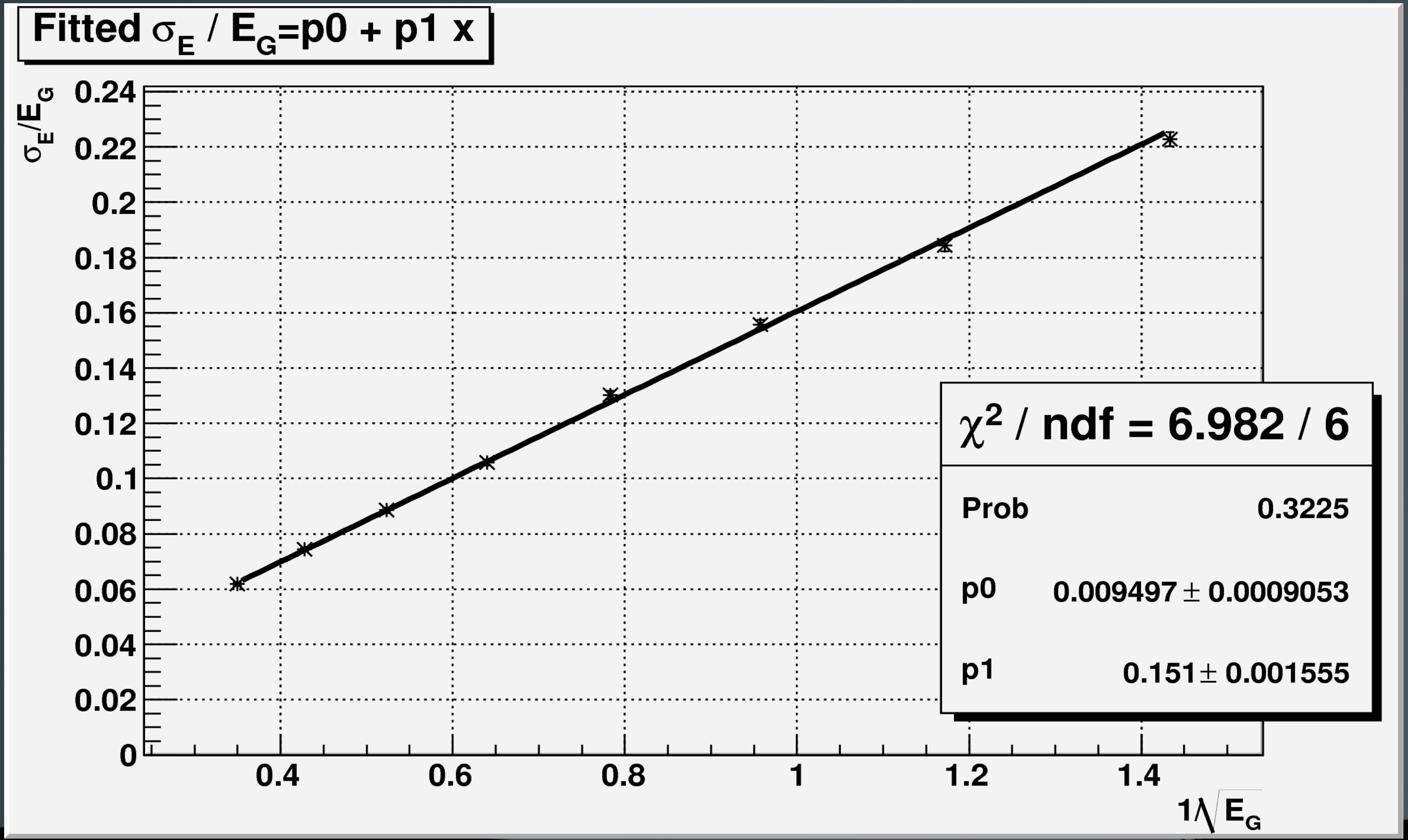
4.0 GeV π^0 Mass Constrained Fits

- Greatest improvement with symmetric decays.



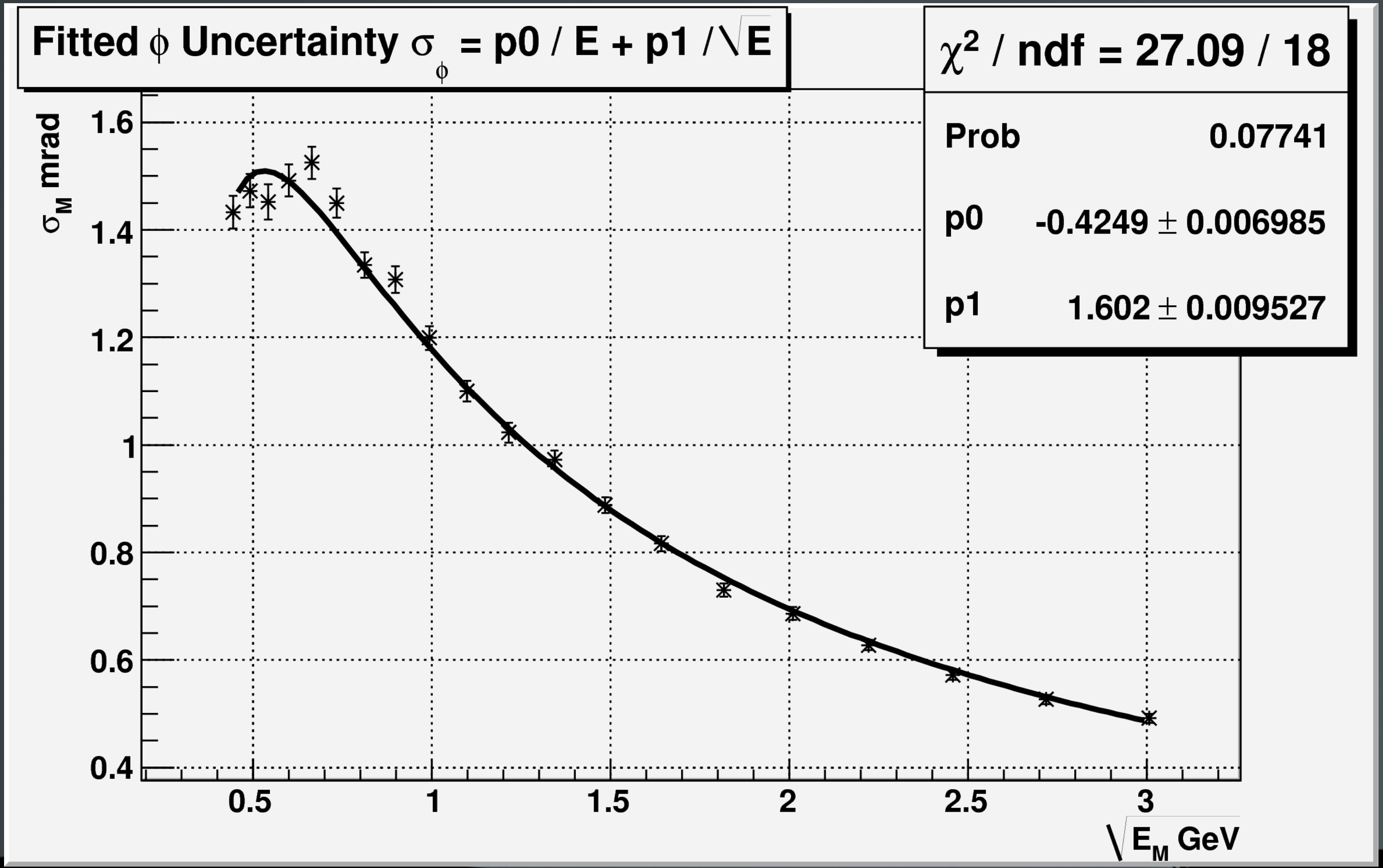
Software: Simulation and Reconstruction

- Uncertainty Modeling: Accuracy important for kinematic fits.
- Energy Uncertainty as function of Energy $\frac{\sigma_E}{E} = \frac{.151}{\sqrt{E}} + 0.0095$



Software: Simulation and Reconstruction

- Uncertainty Modeling: Phi
  - "Turns over" or "flattens out" at low energies



Software: Simulation and Reconstruction

- Uncertainty Modeling: Theta

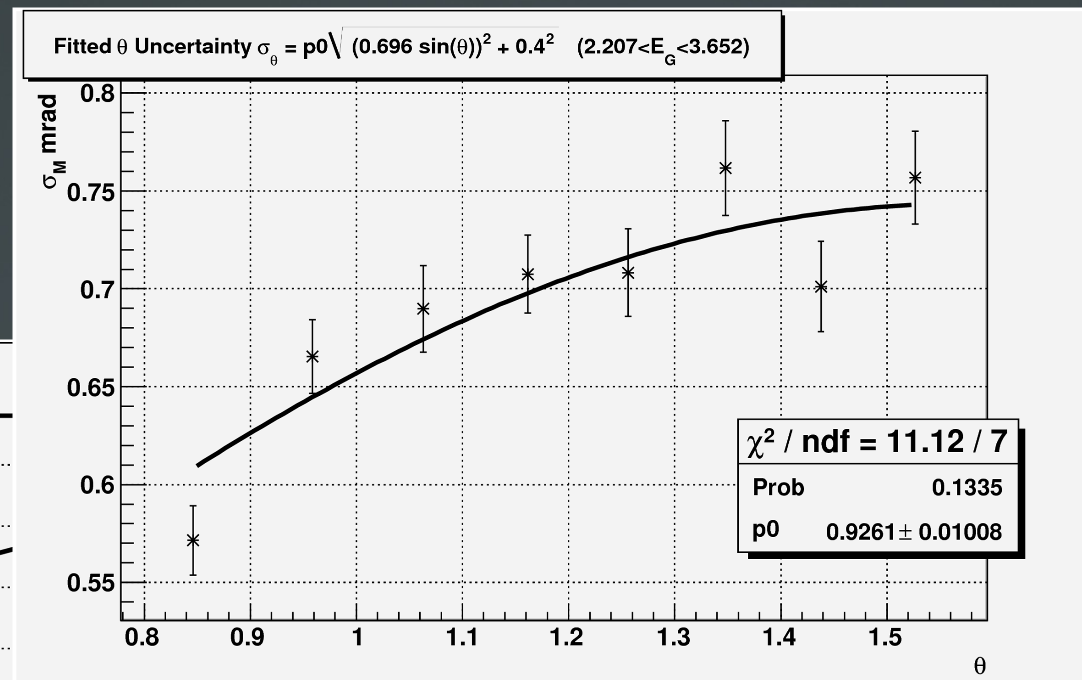
Want smooth function

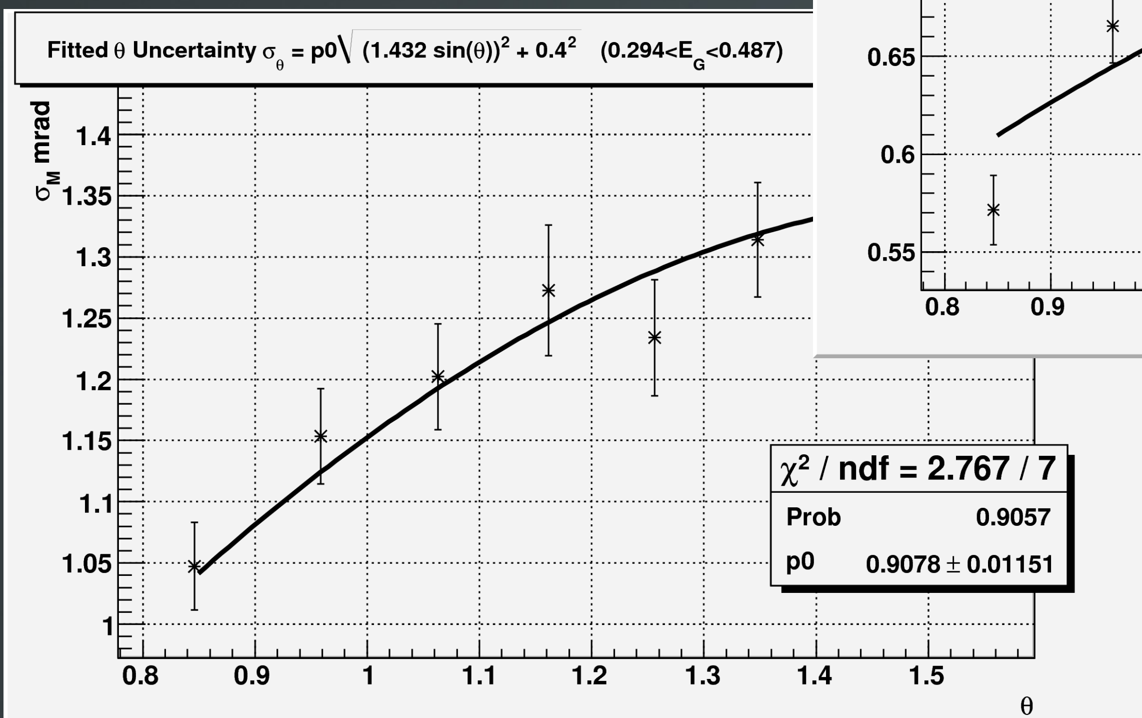
Hypothesis: $\sigma_\theta \rightarrow \sigma_\phi$ as $\theta \rightarrow \pi/2$

$\sigma_\theta \rightarrow 0$ as $\theta \rightarrow 0$

Try: $\sigma_\theta^2 = 0.91^2[(\sigma_\phi^* \sin(\theta))^2 + 0.4^2]$

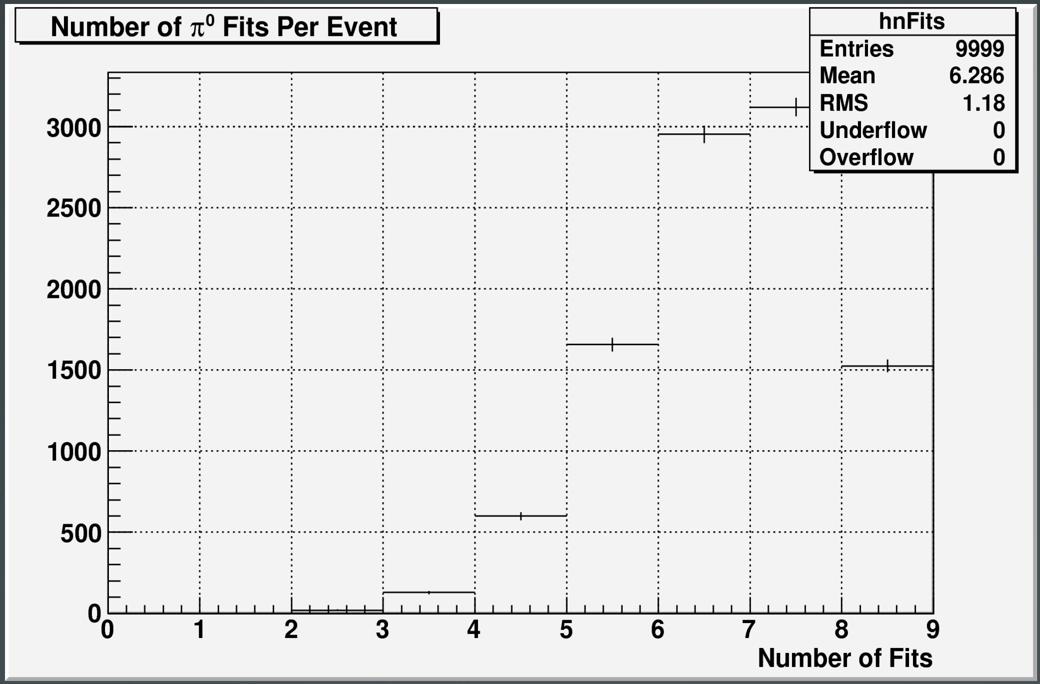
$$\sigma_\phi^* = \sqrt{\sigma_\phi^2 - 0.4^2}$$





Fitting Multiple π^0's

Using truth information, matching is about 80% efficient for 8 pi0's at 4 GeV

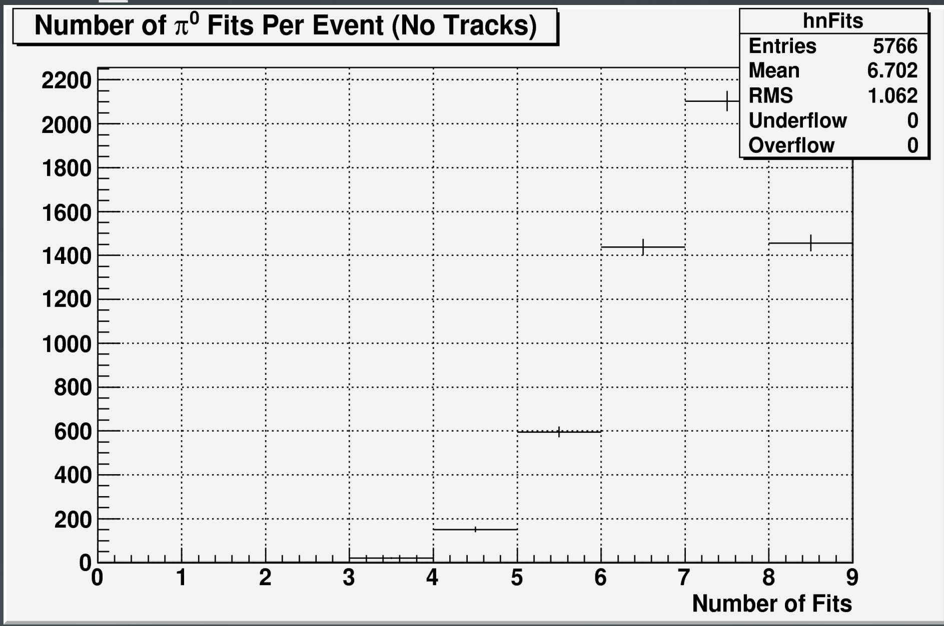


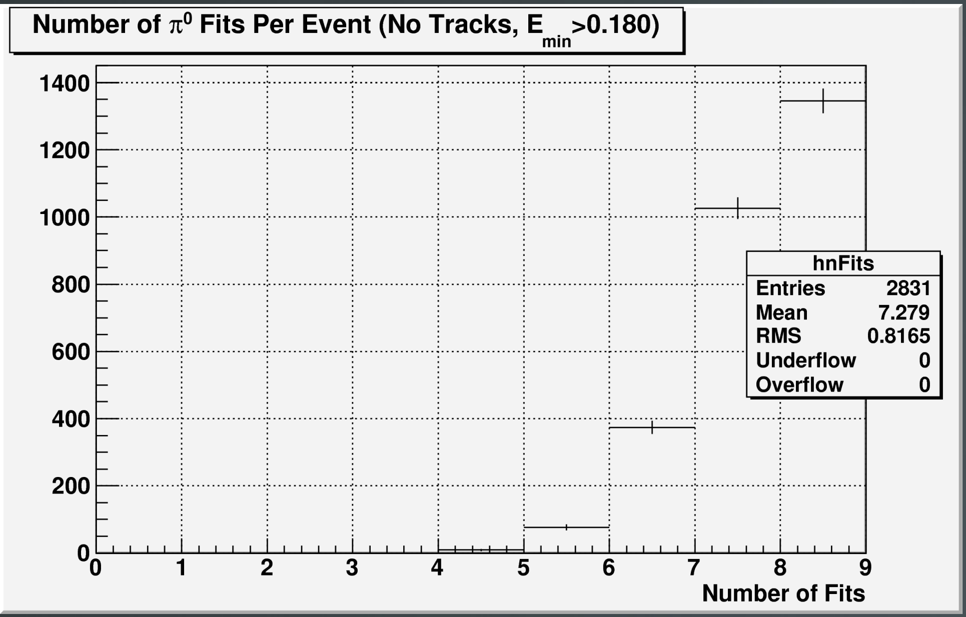
Why is it not 100%?

- e^+e^- pair production
- low energy photon cut (180 MeV)
- Base 1% fit probability cut

Fitting Multiple π^0's

Removing events with tracks increases efficiency to ~84%



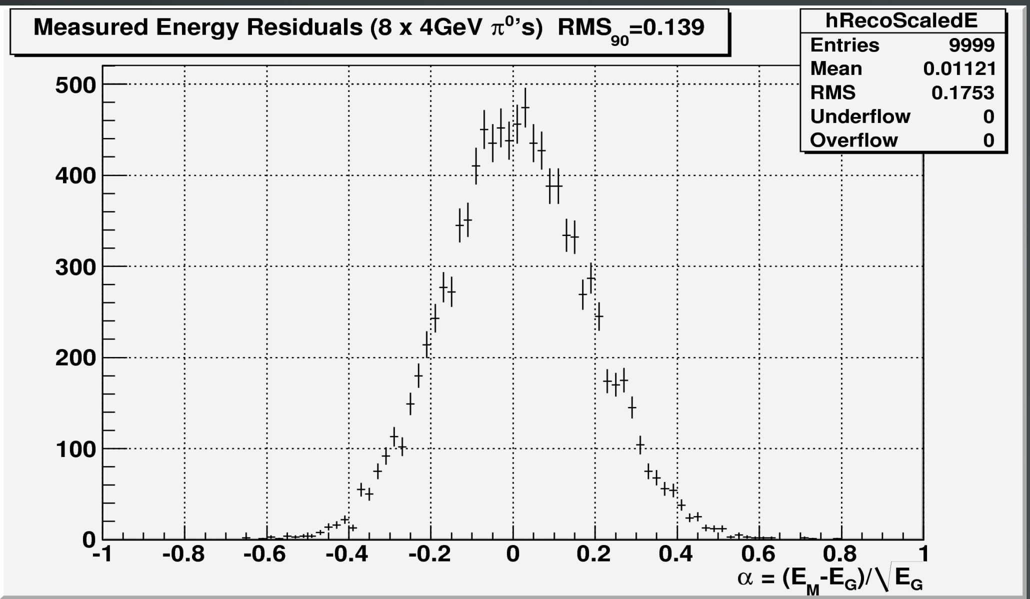


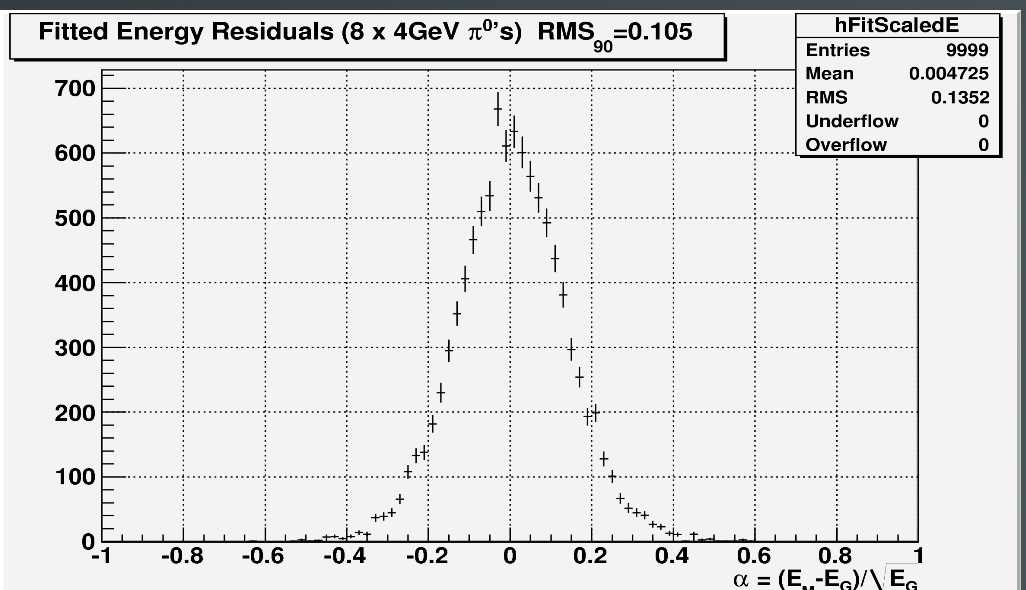
Additionally, removing events with photons below 180 MeV results in ~91% efficiency

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Consistent with binomial distribution where $p = .99^8$ suggesting 1% cut responsible for remainder

Fitting Multiple π^0's





At 80% matching efficiency the energy uncertainty (RMS) improves

$\alpha = 17.5\%$

to

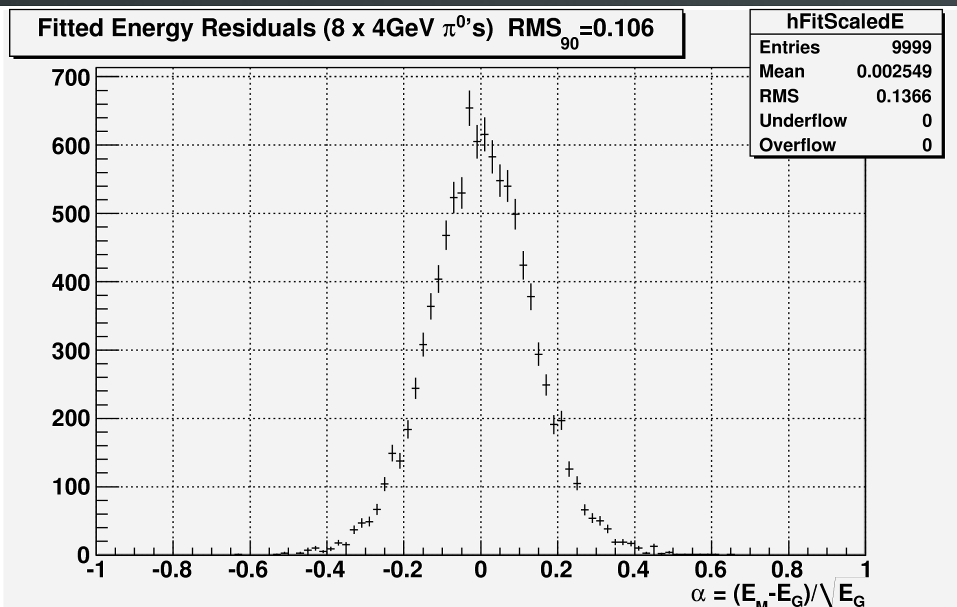
$\alpha = 13.5\%$

$$\frac{\alpha}{\sqrt{E}} = \frac{\Delta E}{E}$$

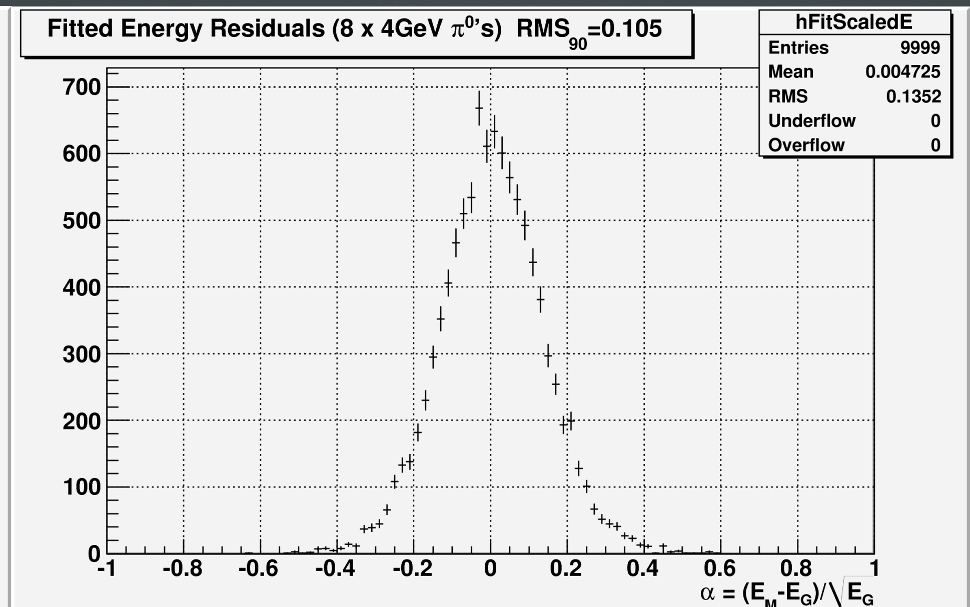
Fitting Multiple π^0's

- Comparison to truth information (8 x 4 GeV π^0's)

Max Fits, Min χ^2



Truth Information



Performance is nearly identical
(for this situation)

$\alpha = .137$ vs. $\alpha = .135$

Fitting Multiple π^0's

- How do these efficiencies vary with multiplicity and energy?

4 GeV π^0's

# of π^0's per event	2	4	8	16	32
% π^0's Fit	79	79	80	78	77.7
Unfitted α	.179	.176	.175	.180	.175
Fitted α	.137	.137	.135	.137	.139

8 π^0's per event

Angular resolution limits high energy fits

Energy (GeV)	4	8	16	32
% π^0's Fit	80	78.3	63.6	46.3
Unfitted α	.175	.179	.189	20.8
Fitted α	.135	.162	.197	20.8

$$\frac{\alpha}{\sqrt{E}} = \frac{\Delta E}{E}$$

Fitting Multiple π^0's

- How does this method compare to using truth information?

4 GeV π^0's

# of π^0's	2	4	8	16
% π^0's Fit	79	79.5	79.3	74.7
% Correct	79	79	78.0	72.9
Cheating	79	79	80	78

8 π^0's per event

Energy (GeV)	4	8	16
% π^0's Fit	79.3	79.0	66.3
% Correct	78	77.9	63.6
Cheating	.80	78.3	63.6

Pretty good!