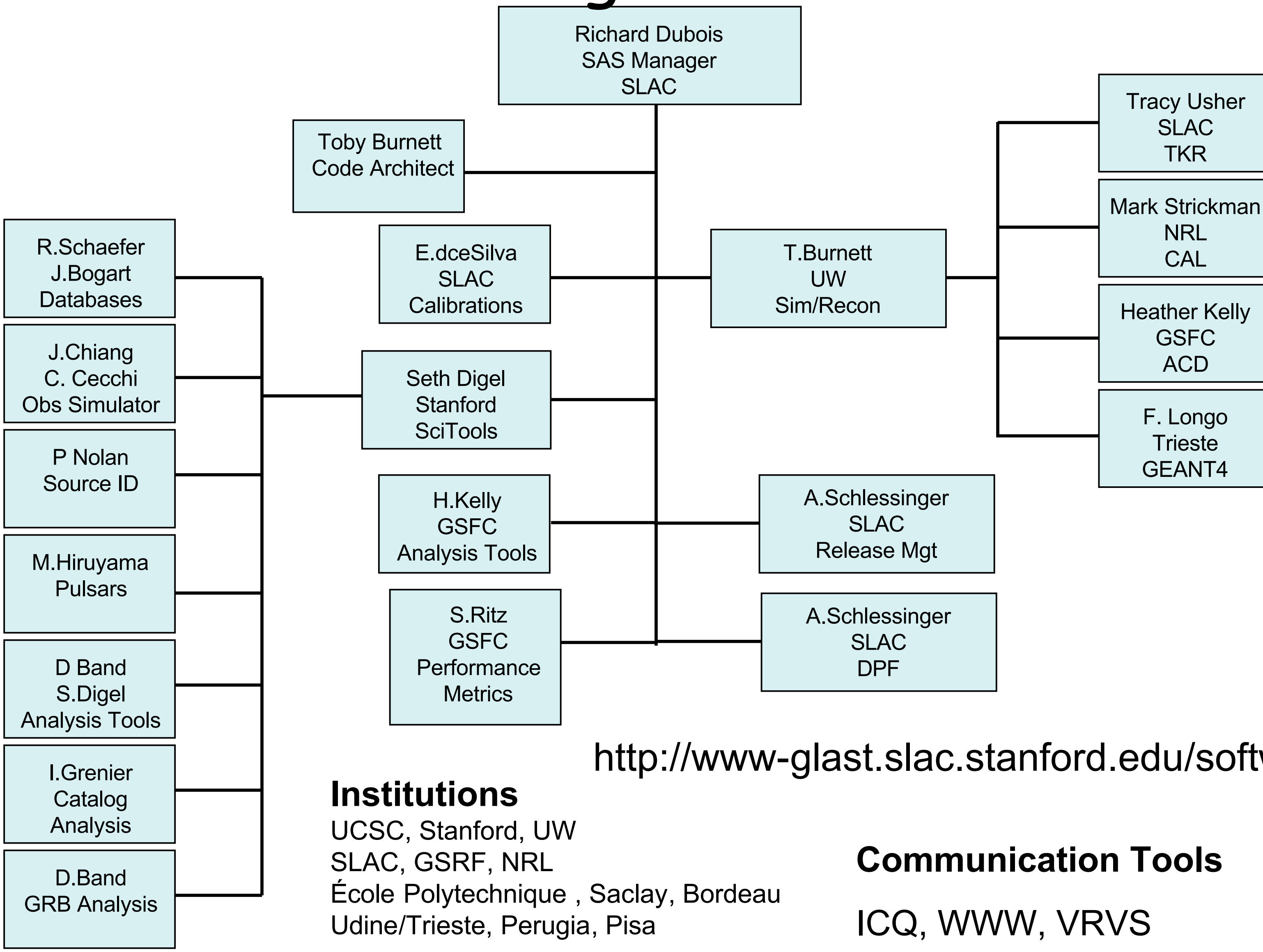
LAT Reconstruction (and simulation)

Toby Burnett
University of Washington

Outline

- An overview of the Science Analysis Software group
- History of GLAST software – back to 1990!
  - GLAST has been driven by software
- Some details of sim/recon
- Plans

SAS organization



http://www-glast.slac.stanford.edu/software/

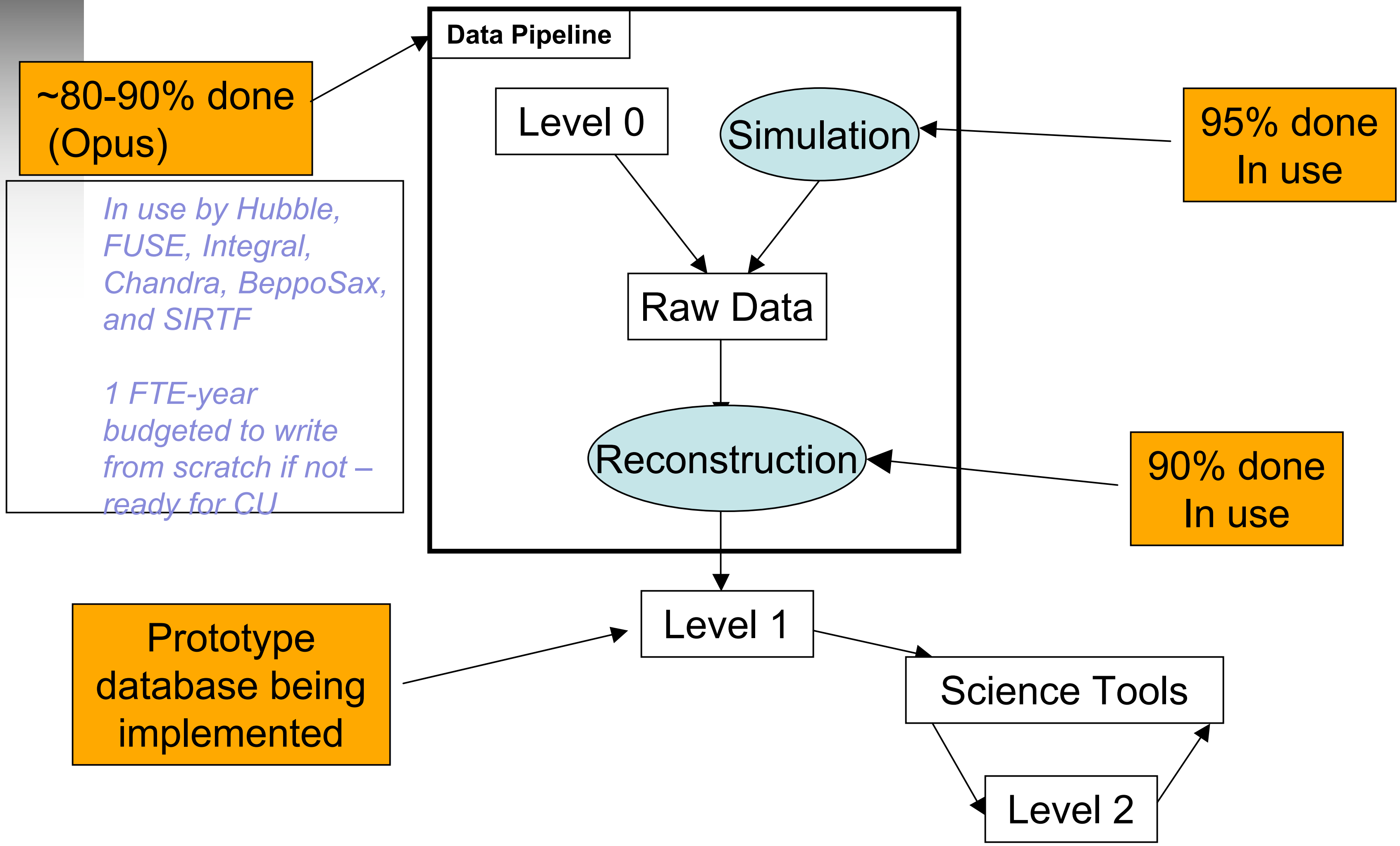
Institutions

UCSC, Stanford, UW
SLAC, GSRF, NRL
École Polytechnique , Saclay, Bordeau
Udine/Trieste, Perugia, Pisa

Communication Tools

ICQ, WWW, VRVS

Processing flow, current status



Birth of GLAST code:
CERN Canteen #1, summer 1990

Legacies of that meeting

- Object-oriented design for new simulation toolkit, named *Gismo*
  - Code is organized into *classes*
  - Data hiding in objects
- Detector description (geometry, materials, active regions) accessible in same form to both *simulation* and *reconstruction*
- Interactive application
  - Combines simulation and reconstruction in one package
  - Choice of source parameters on the fly
- Integrated 3-D display with GUI controls
  - Interactive control over display of:
    - geometry: what is where
    - particles: where they go, what happens to them
    - detector response: how the active regions respond to deposited ionization (and is it in the right place?)
    - reconstruction: how well does the pattern recognition and fitting represent the input response?
- Easy transition to batch mode, tools to generate n-tuple summaries

The rest is history

- 1992:
  - Bill Atwood and Peter Michelson consider a modern design for the just-launched CGRO/EGRET; Bill starts using the toolkit to test designs.
  - Basic attributes of the current design emerge quickly:
    - Si strip tracker/converter, converters just above strips
    - Segmented ACD, *not* in the trigger!
    - Onboard level 1 trigger, software filter
    - Segmented CAL.
    - Large aspect ratio for good FOV, modular design (originally 7x7)
    - Basic scale (1.8 m square, <10 Rad Len CsI) set by Delta II launch capability
- 1994:
  - Toby Burnett joins, takes over top-level design
  - Bill and Peter get NASA's attention with mission concept study
    - All the basic performance parameters based on simulations

History, cont.

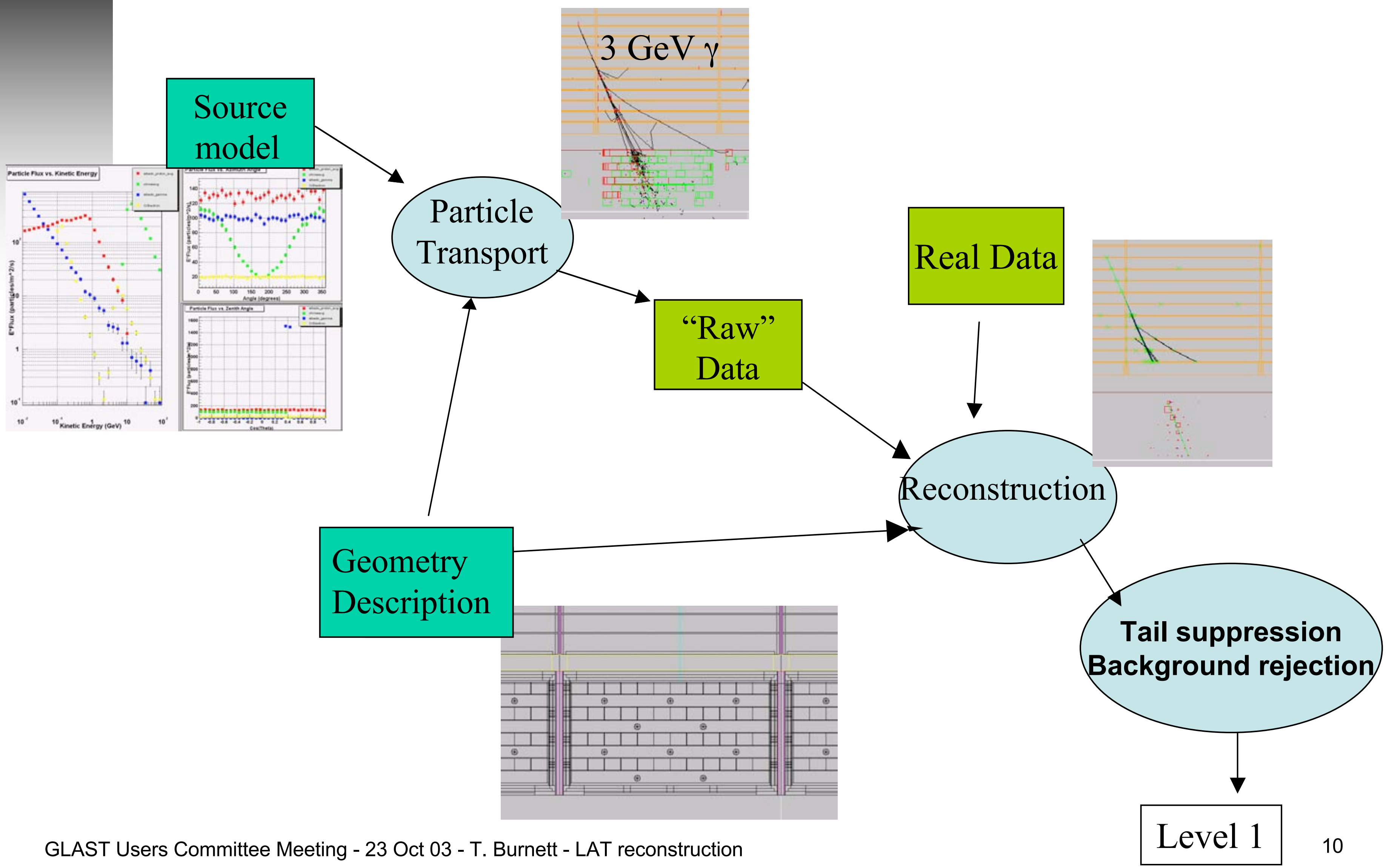
- 1995-1998
  - Gradual increase in collaboration size, UCSC and SLAC
  - Start using Kalman filter for track fitting
  - Steve Ritz joins
  - Beam tests validate simulation
- 1999
  - (Dec) AO response submitted following extensive simulations
- 2000
  - (Feb) LAT selected
- 2001
  - Adopt the present infrastructure (all supported elsewhere)
    - Source management – cvs, repository at SLAC
    - Package management/ build system – CMT
    - Execution framework – Gaudi
      - Component model with Abstract interfaces
      - Dynamic loading of components at runtime
    - Support only linux/gcc and Windows/Developer Studio
  - Define xml-based geometry data base
  - Switch from Gismo to Geant4
- 2002
  - (Jan) PDR baseline
  - Choose the name “Gleam” to represent the combined sim/recon executable [**GL**ast **E**vent **A**nalysis **M**achine]

Currently

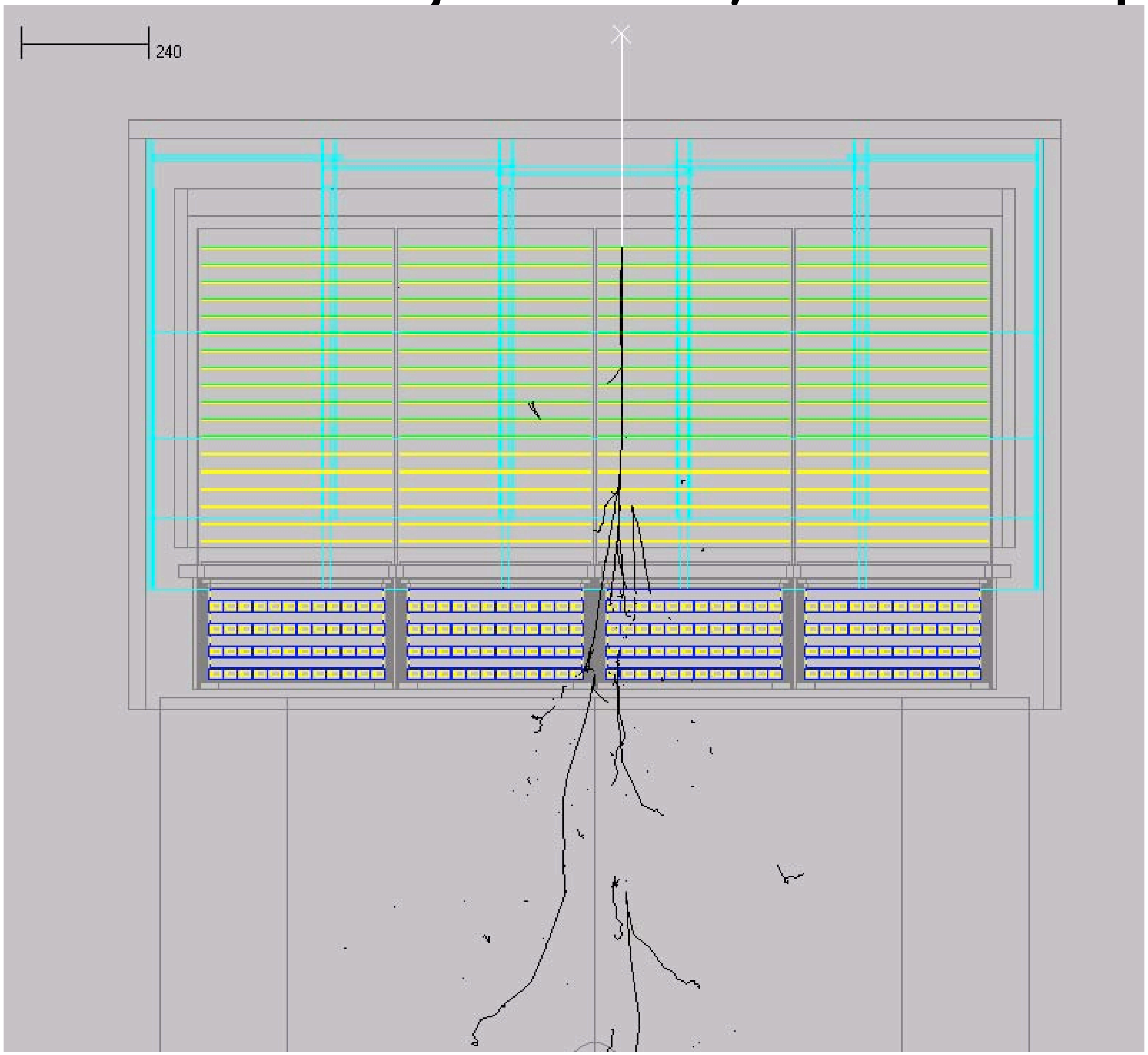
- Testing new Background model based on AMS Shuttle observations.
- Code from onboard filter incorporated into analysis for evaluation and testing
- Preparing for Data Challenge 1.

Bottom line: modeling and reconstruction software has driven the development of GLAST, not lagged behind hardware development

A more detailed cartoon



Details: generate, fit 1 GeV photon

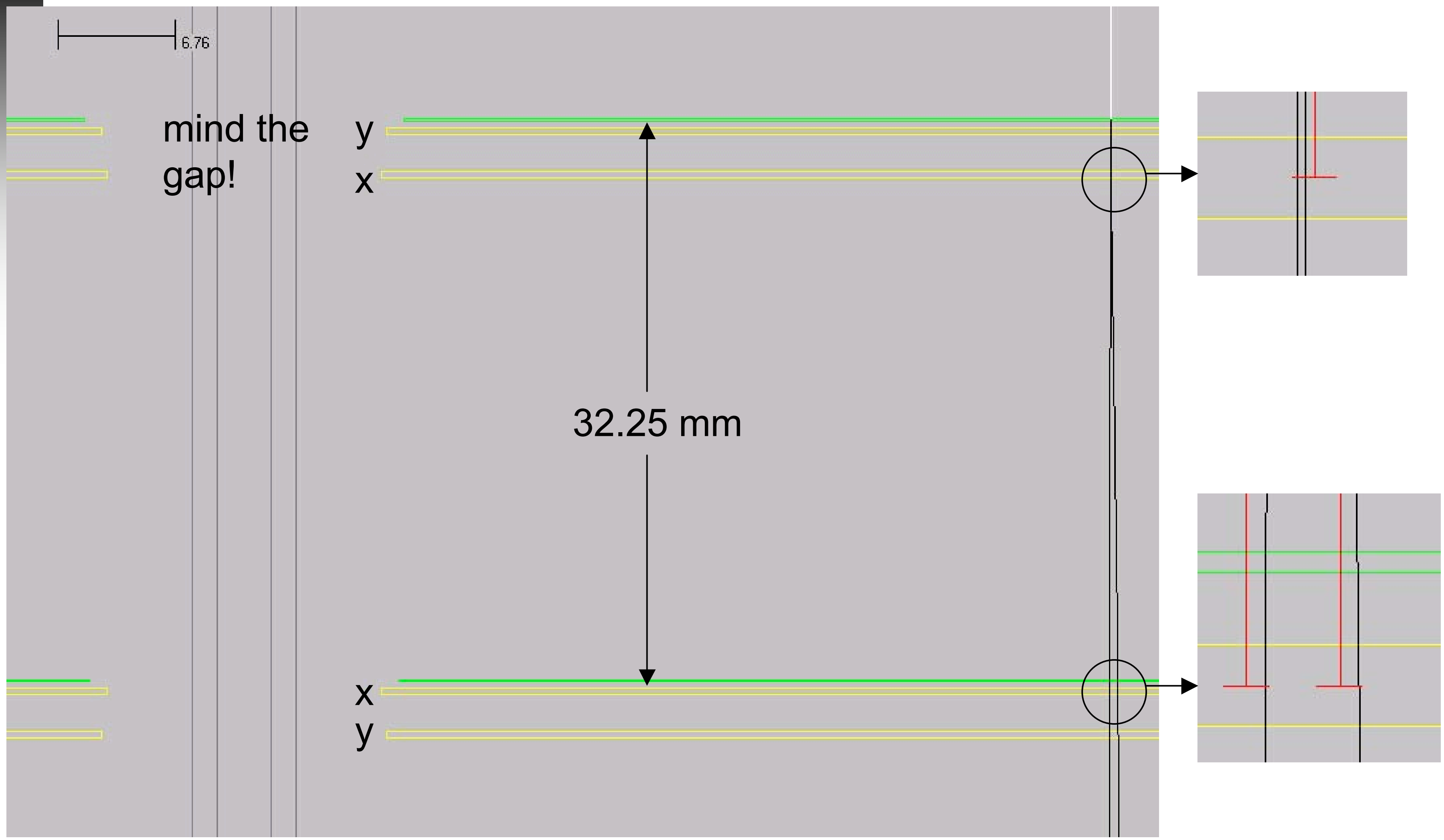


aqua: ACD tiles
yellow: Si
green: W

only charged tracks shown

no detector response or recon shown

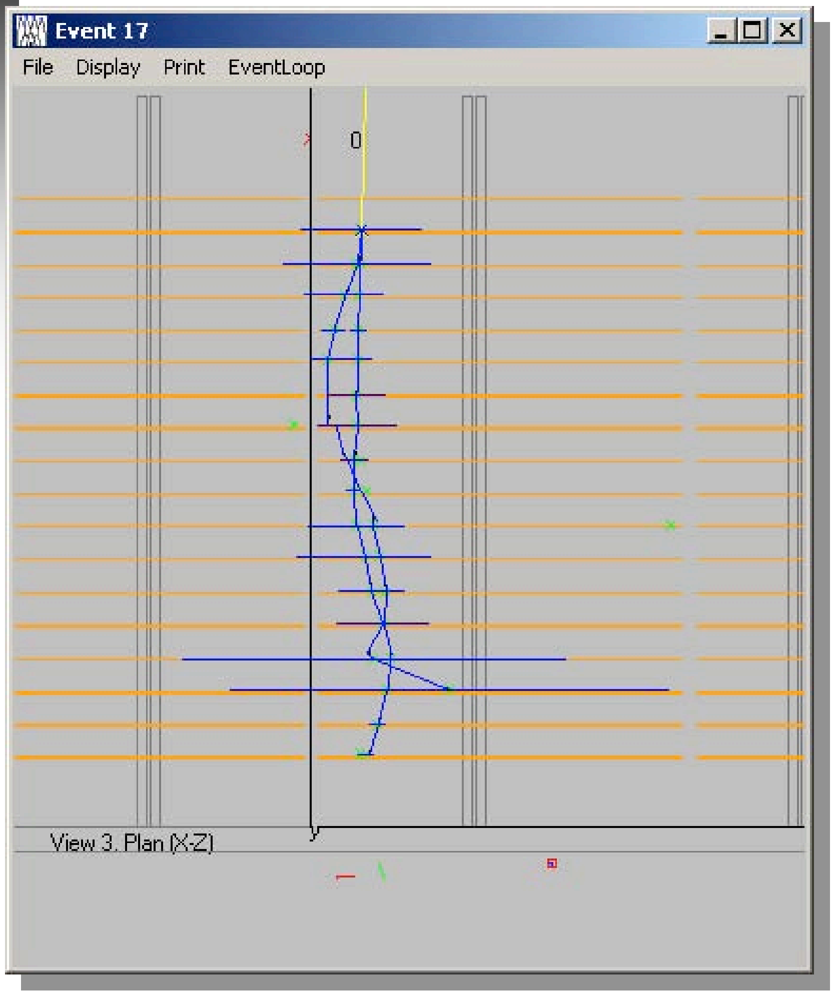
Zoom in to the conversion

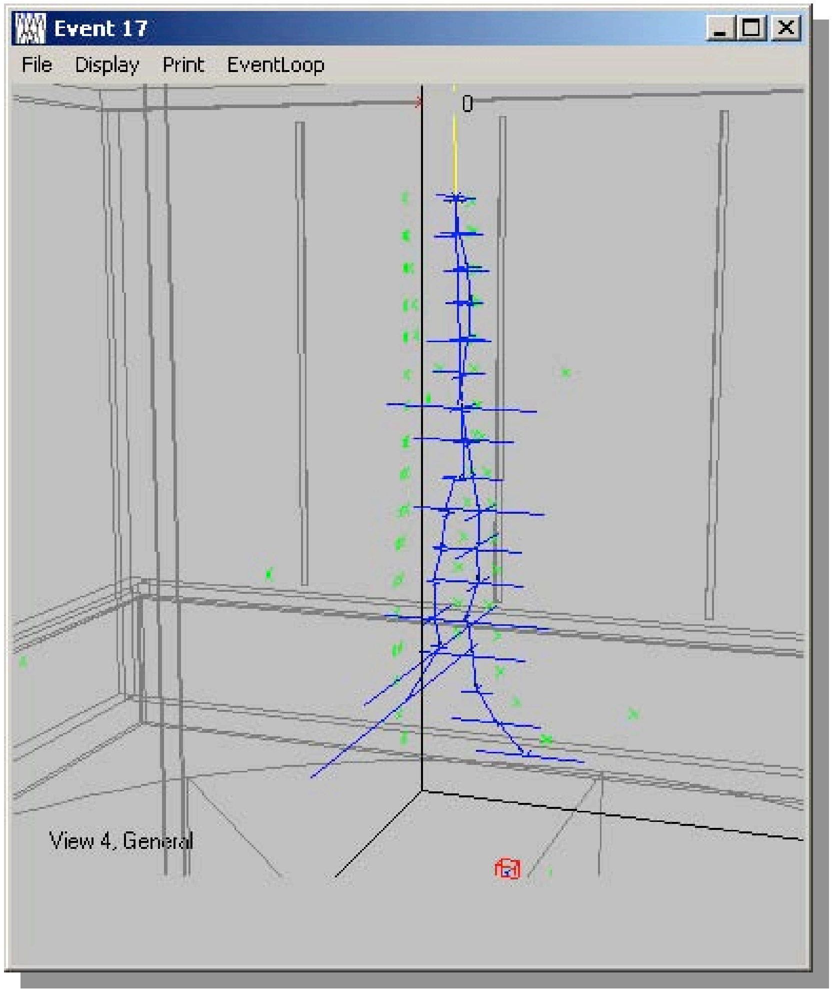


Angular resolution and track fitting

- Intrinsic limits (projected)
  - multiple scattering: in 1.25% RL (1/2 a "thin" layer):
    1.5 mrad * (1 GeV / p)
  - pitch: 2 mrad for one layer.
  - [for the astro guys: 1 deg = 17 mrad]
- Naïve fitting strategies
  - Low energy: use only first two layers, since next conversion layer adds error to subsequent layer measurements
  - High energy: simple least squares fit
- Better way: Kalman filter
  - designed to combine "process" and "measurement" noise.
  - Equivalent to the naïve limit, but interpolates properly in-between
  - Implies that there is a measurement of the energy/momentum, at each plane
  - Even for low energy, follows each track.

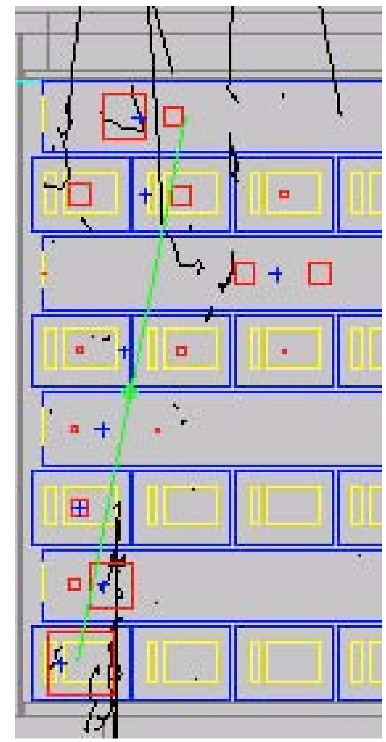
Example: 100 MeV gamma





The Calorimeter

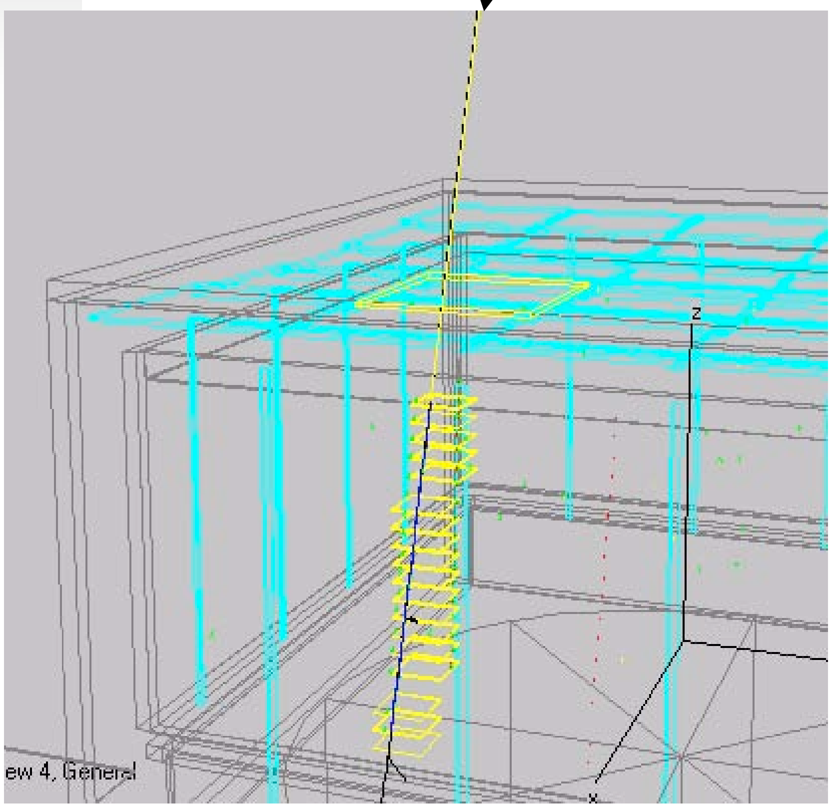
- 4 layers of 18% RL start shower early, but absorb energy (only 380 MeV in the CsI here
- Large gap between modules
- Reconstruction is done iteratively with tracker, two passes
  1. Preliminary measurement with basic clustering algorithm, predicts energy and direction
  2. Tracking uses this, and estimates energy in tracker (using observed MS, counting hits)
  3. Calorimeter refines measurement with track direction(s)
- – High energies: use shower shape to correct for leakage



mind this gap!

The ACD

– Extrapolate to plane of each hit tile, measure (signed) distance to edge of the tile
– Reject incoming charged particles if inside

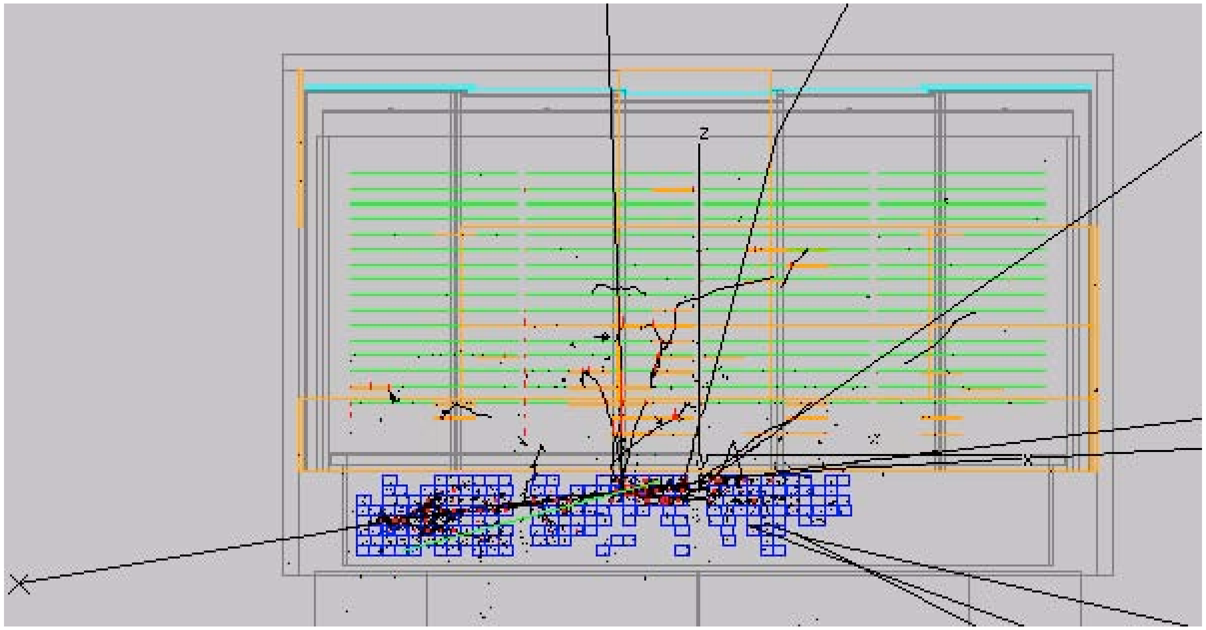


simulated muon, showing (in yellow) the tile and Si wafers

MC track, fit, and projected direction all collinear

Background rejection

- Requirements:
  - Onboard: filter factor of ~100. (for downlink)
  - Ground: need another factor of 100 (for science)
  - Simulation: create events that find all the “holes”
- Ground Strategy
  - Generate useful discrimination variables
  - Apply cuts (or classification trees)



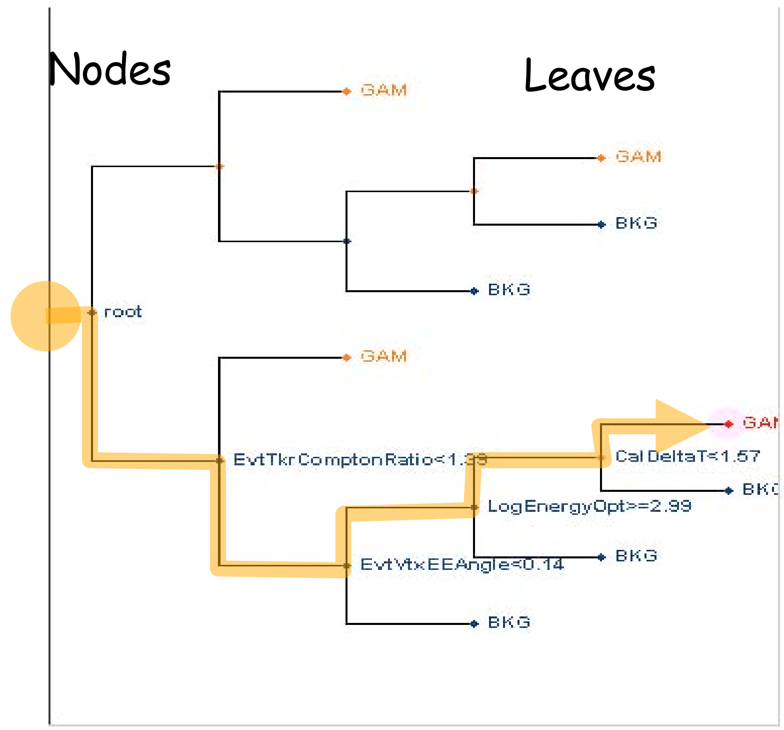
Classification

- What is it?
  - A new category (for us) of analysis: depends on application of *Classification* and *Regression* trees
  - Common use in “soft” sciences, found by Bill Atwood.
  - A systematic way to find optimal regions in multidimensional parameter space to separate populations: result is expressed as a tree of cuts in the space
- Where do we use it?
  - Determine if **energy is well measured** (important for track fit)
  - Choose **vertex or single track** gamma direction estimate
  - Assess probability that an event is in the **PSF core** distribution
  - Predict the PSF itself
  - Assess probability that an event is really a **gamma ray** (vs. background)
- How are the trees generated?
  - Generate large samples of gammas and background for training, using the commercial tool Insightful Miner
  - Output in the form of XML trees is used by recon software.

Classification Primer from W. Atwood

Origin: Social Sciences - 1963

How a CT works is simple:
A series of "cuts" parse the data into a "tree" like structure, where final nodes (leaves) are "pure"



A "traditional analysis" is just ONE path through such a tree.

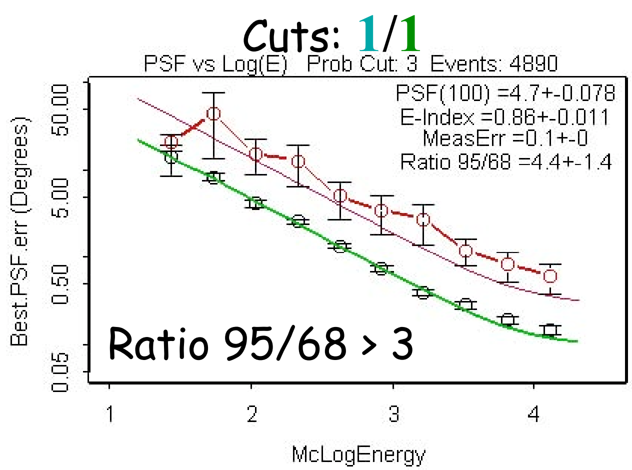
Tree are *much* more efficient!

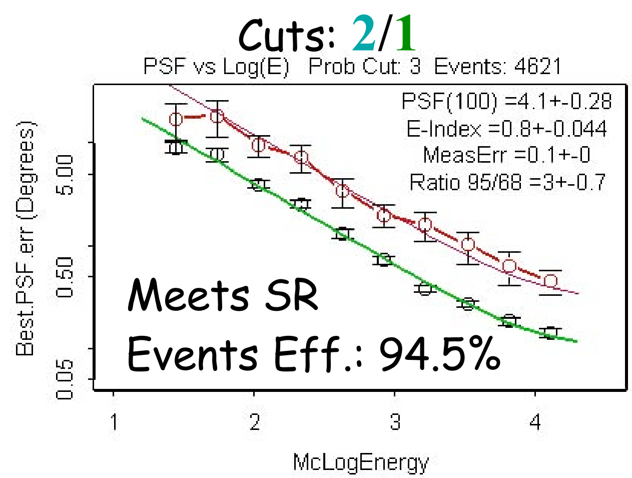
Mechanism of tree generation less subject to "investigator basis."

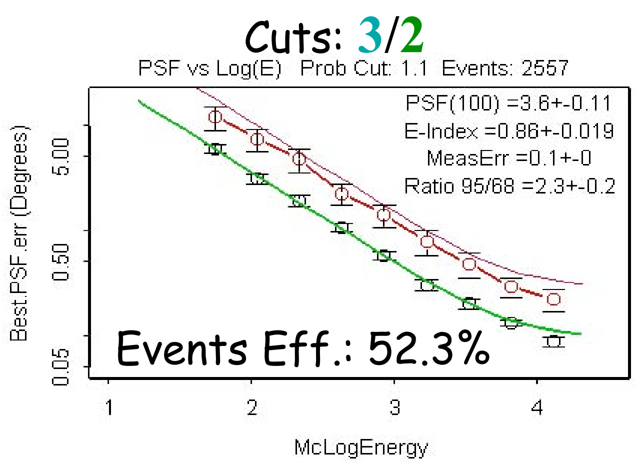
STATISTICALLY HONEST!

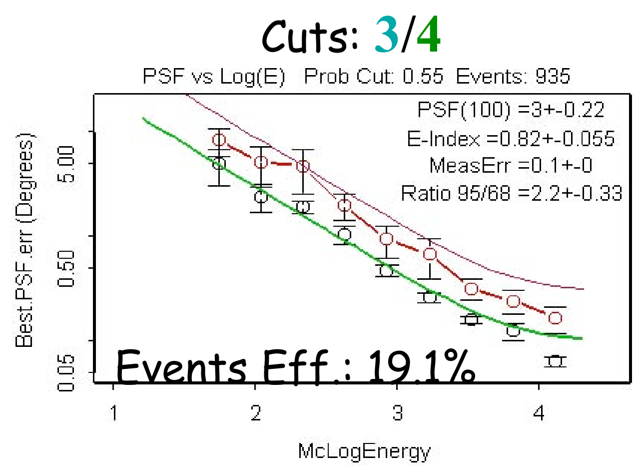
Bottom Line—characterize response

- Use MC events passing classification cuts to characterize PSF and A_{eff}
- Example of current status: PSF integrated over angles, as function of energy, with different selections on core probability









Data Challenges

- Now traditional in HEP experiments
  - exercise the full analysis chain with simulated data, including hidden signals
  - challenge the collaboration to find the signals, using tools
- Doing planning now
  - Fall 2003 - DC1
    - 1 day's data through full instrument simulation and first look at Science Tools
    - Launched at Sept collaboration meeting
    - Simulation challenge: needs ~500 CPU weeks for background.First use of pipeline
  - Fall 2004 – DC2
    - 1 month's background/1 year signal
    - Test more Science Tools; improved Pipeline
  - Spring 2006 – DC3
    - run up to flight – test it all!

Other reviews, tests coming

- Peer review in Feb '04, CDR to be done with Mission Ground Systems Review in May 04
- EM and beam test for practice and tuning of sim/recon as part of development plan
- Mission Ground Tests in late '04 and '05 to exercise the movement of bits from the MOC to the SSC; with the DC's

Summary

- Sim/Recon has played a vital part in the definition of GLAST
- With the LAT design now final, the geometry description is approaching a faithful summary
- Algorithms for reconstruction and classification continue to be improved
- Serious testing, including the pipeline, is about to start with DC1
- We are optimistic about the LAT IOC Ground Systems CDR, scheduled for 2/2004, with Peer Review in 11/2003
- Variations on the geometry, but using the same software is ready to support the current Engineering Module (EM) and Calibration Unit (CU) for 2005 beam test.
- Expect to be ready far in advance for Mission Ground Tests in late '04 and '05 to exercise the movement of bits from the MOC to the SSC; with the DC's

Possible User interaction

- All code is documented and publicly available (but will not be supported by GSSC)
- Possible use
  - Monte Carlo studies
  - Refitting with different algorithms
  - Event display with level 1 data (using FRED)