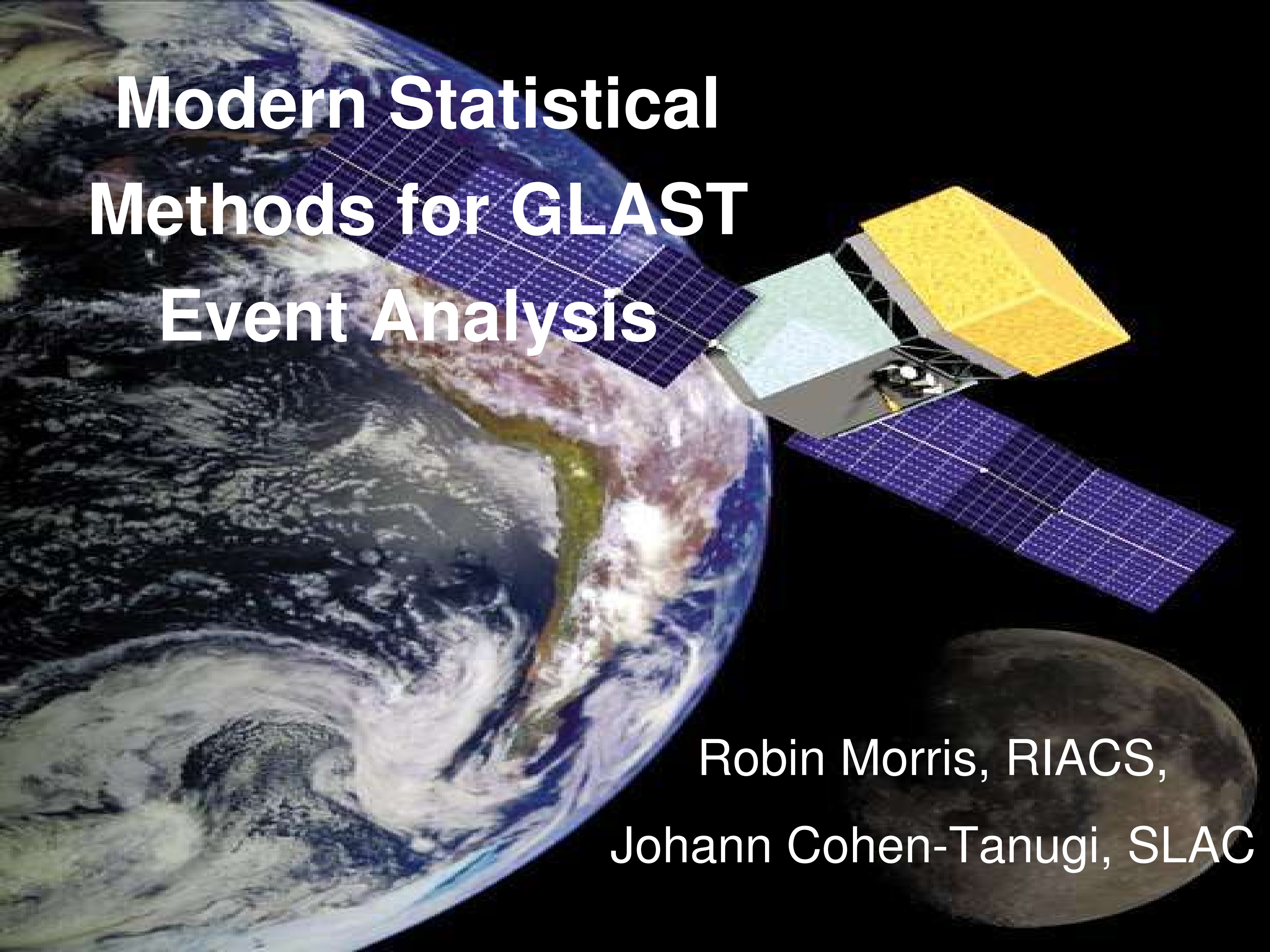
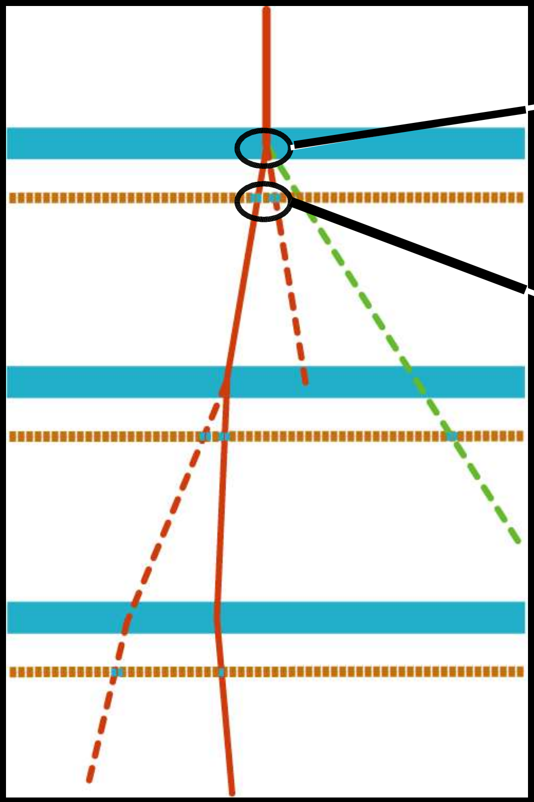


Modern Statistical Methods for GLAST Event Analysis

Robin Morris, RIACS,
Johann Cohen-Tanugi, SLAC

Goal:

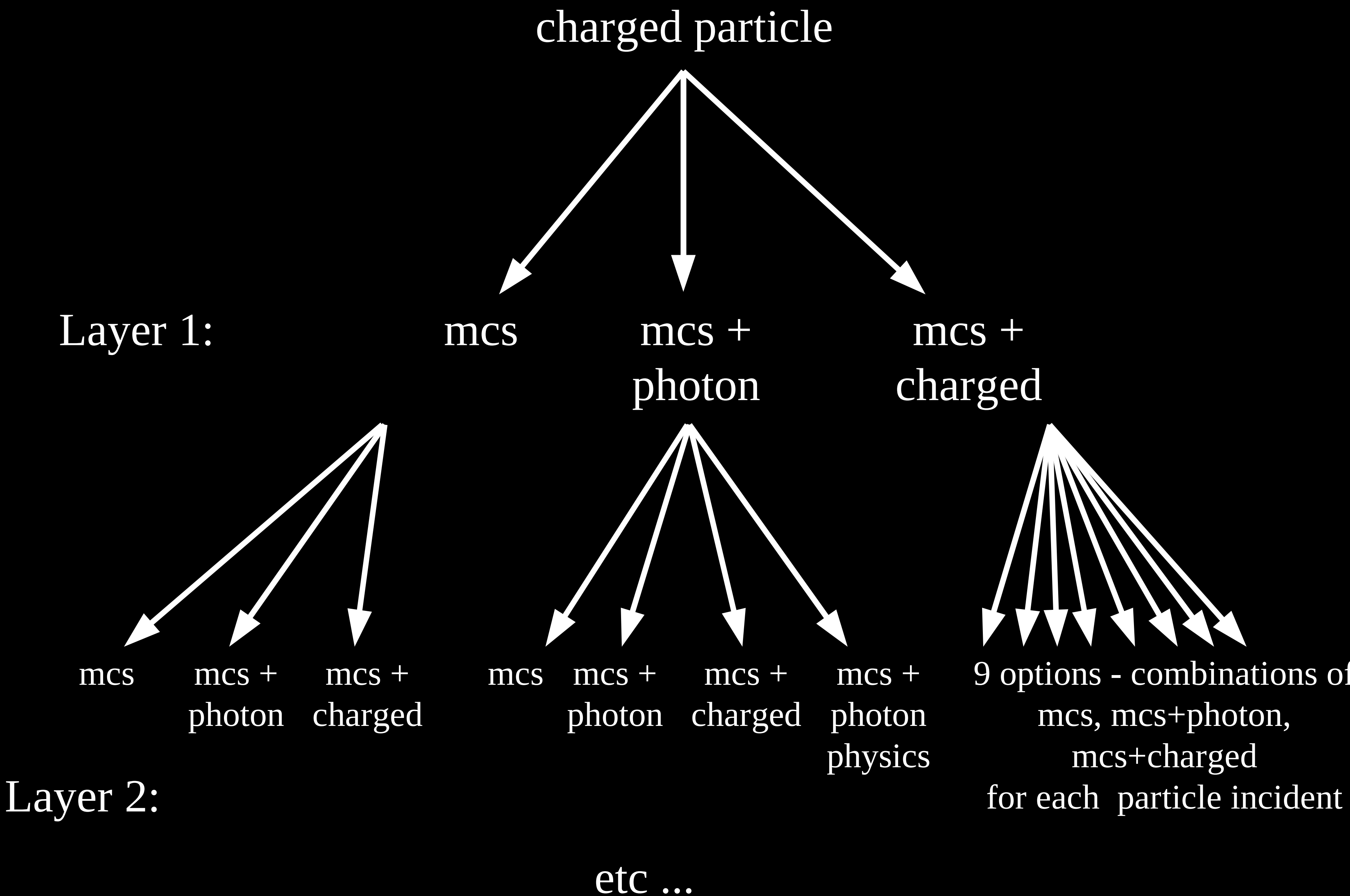
- To perform a probabilistic reconstruction that correctly accounts for the possible types of event that may have caused the detector response



Generate hypotheses for the physics processes that occurred in the tungsten foils

By looking at the patterns of microstrips that fired

Then compute the relative probability of each hypothesis, using the known probability of each physics process

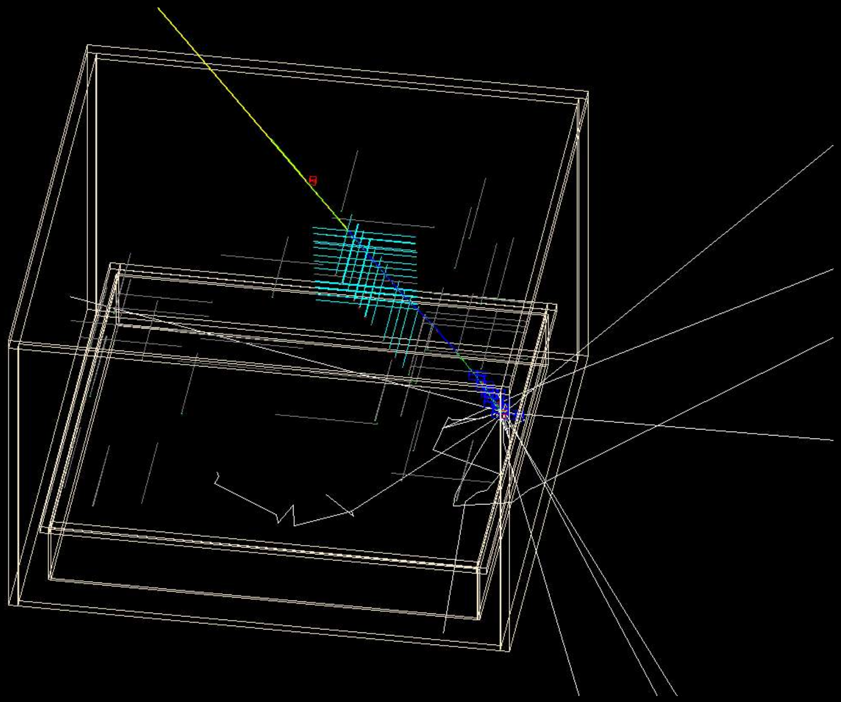

charged particle
Layer 1:
mcs
mcs + photon
mcs + charged
Layer 2:
mcs
mcs + photon
mcs + charged
mcs
mcs + photon
mcs + charged
mcs + photon physics
9 options - combinations of mcs, mcs+photon, mcs+charged for each particle incident
etc ...

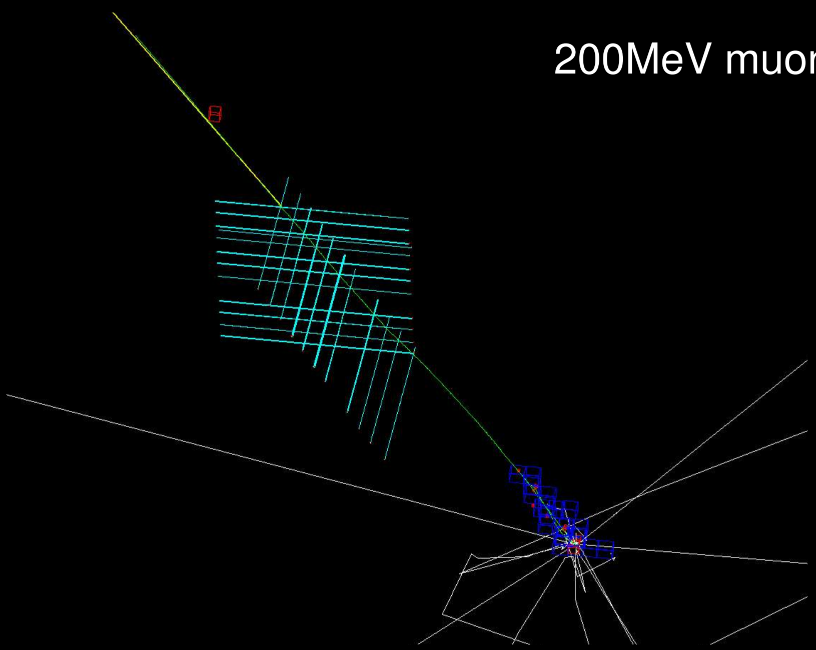
Controlling the Complexity

- use prior information about the physical processes
  - eg: the probability of a secondary charged particle being produced when a muon is incident is very small
- look at the data at each layer
  - eg: if there is only only one hit, there is very little chance that a secondary charged particle was produced
  - (strictly, this is looking at the data twice)
- proceed sequentially, and prune the tree as you go

Statistical Estimation

- Estimate: initial direction and energy of a particle
- Given: a set of microstrips that fired

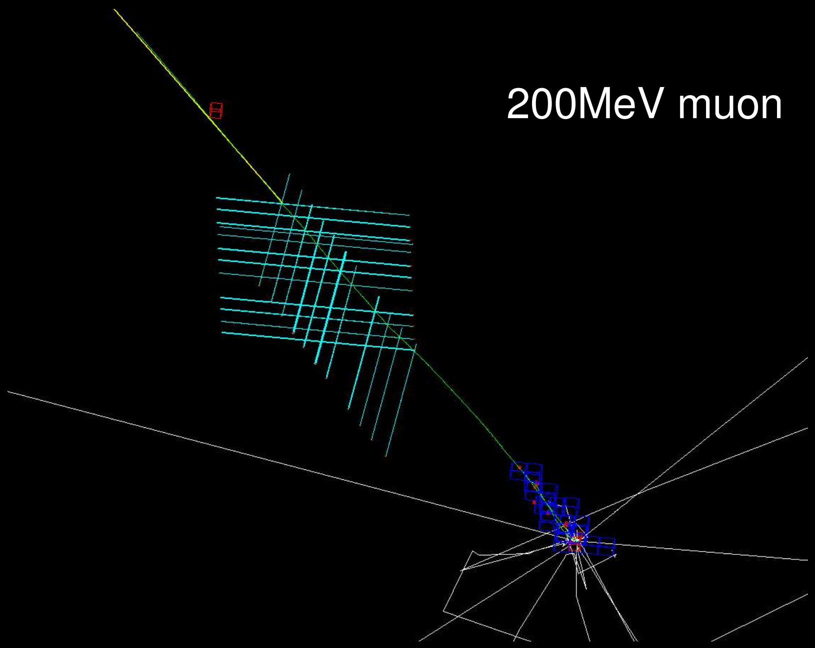




- Evaluate $p(\theta, \phi, E \mid s_1, s_2, \ldots, s_N)$

Statistical Estimation

- Evaluate $p(\theta, \phi, E \mid s_1, s_2, \ldots, s_N)$



- To do so, need to introduce auxilliary variables
- scattering angles at each layer
- energy loss at each layer

$$p(\theta, \phi, E, \theta_1, \theta_2, \ldots, \theta_k, \delta E_1, \delta E_2, \ldots, \delta E_k \mid s_1, s_2, \ldots, s_N)$$

Statistical Estimation

$p(\theta, \phi, E, \theta_1, \ldots, \theta_k, \delta E_1, \ldots, \delta E_k \mid s_1, s_2, \ldots, s_N)\ \alpha$

$p(s_1, s_2, \ldots, s_N \mid \theta, \phi, E, \theta_1, \ldots, \theta_k, \delta E_1, \ldots, \delta E_k)\ *$

Likelihood - does the hypothesized event described by $\theta, \phi, E, \theta_1, \ldots, \theta_k, \delta E_1, \ldots, \delta E_k$ trigger the strips that fired in the event?

$p(\theta, \phi, E, \theta_1, \theta_2, \ldots, \theta_k, \delta E_1, \delta E_2, \ldots, \delta E_k)$

Priors - the physics enters here. Decompose as

$p(\theta_1, \delta E_1 \mid E)$ - scattering at the first layer

$p(\theta_2, \delta E_2 \mid E, \delta E_1)$ - scattering at the second layer

$p(\theta_3, \delta E_3 \mid E, \delta E_1, \delta E_2)$ - scattering at the third layer

Marginalization to $p(\theta, \phi, E \mid s_1, \ldots, s_N)$ is done numerically.

Statistical Estimation

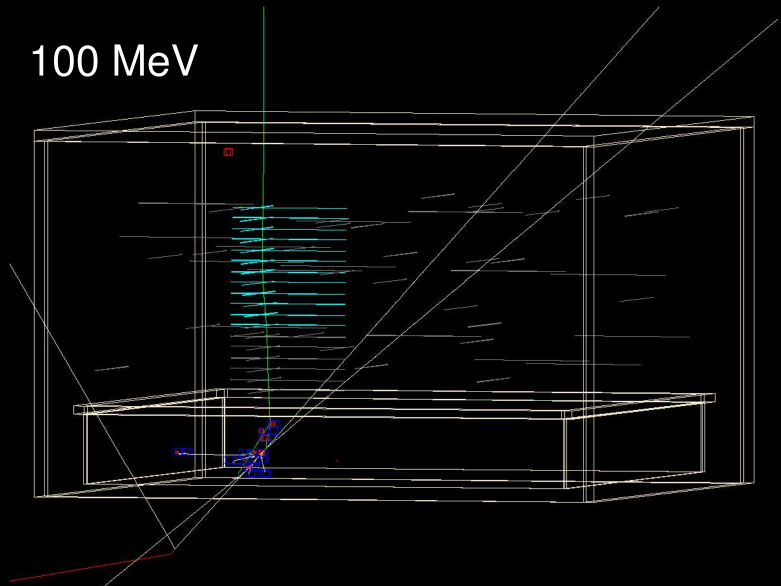
- Markov Chain Monte Carlo allows samples to be drawn from complex probability distributions
- Once we have samples from a distribution, we can compute means, variances etc

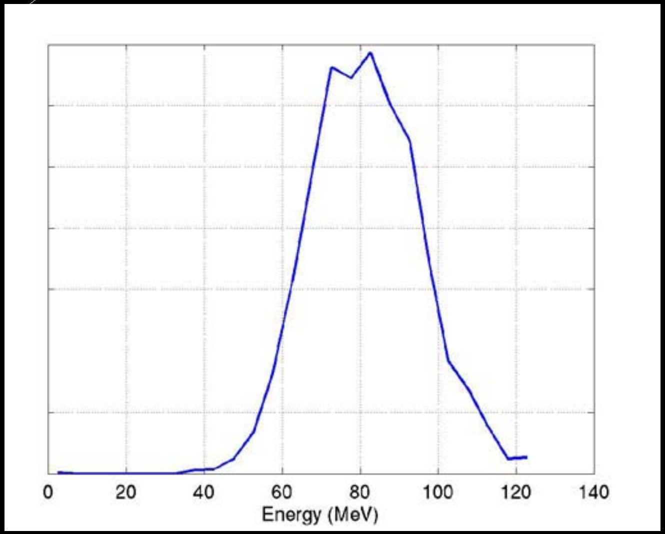
Algorithm:

1. initialize x_0
2. propose a change to x' using some proposal distribution $\pi(x'; x_i)$
3. accept the change with probability p_{acc} where
   $p_{acc} = p(x')\ \pi(x_i; x';) / p(x_i)\ \pi(x'; x_i)$
4. if accepted, $x_{i+1} = x'$, otherwise $x_{i+l} = x_i$
5. goto 2

- The samples $\{x_i\}$ form a Markov chain with limiting distribution p(x)
- Marginalization is performed simply by throwing away elements of (vector) x that we're not interested in

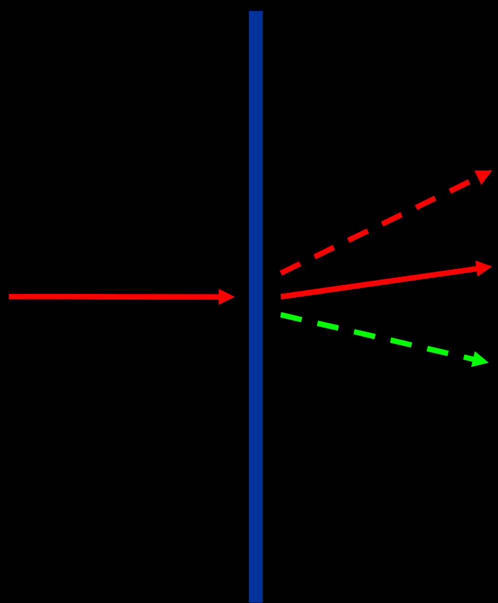
Inference for Muons





- prior for secondary production is ~0
- proposals change the
  - energy
  - azimuth
  - elevation
  - scattering angles
  - energy losses
- The scattering angle distribution is ~Gaussian with log(σ) α log(E). So in essence we're trying to estimate the variance of a Gaussian from a small number of samples (~12)

But we're really interested in electrons



For a 100MeV electron:

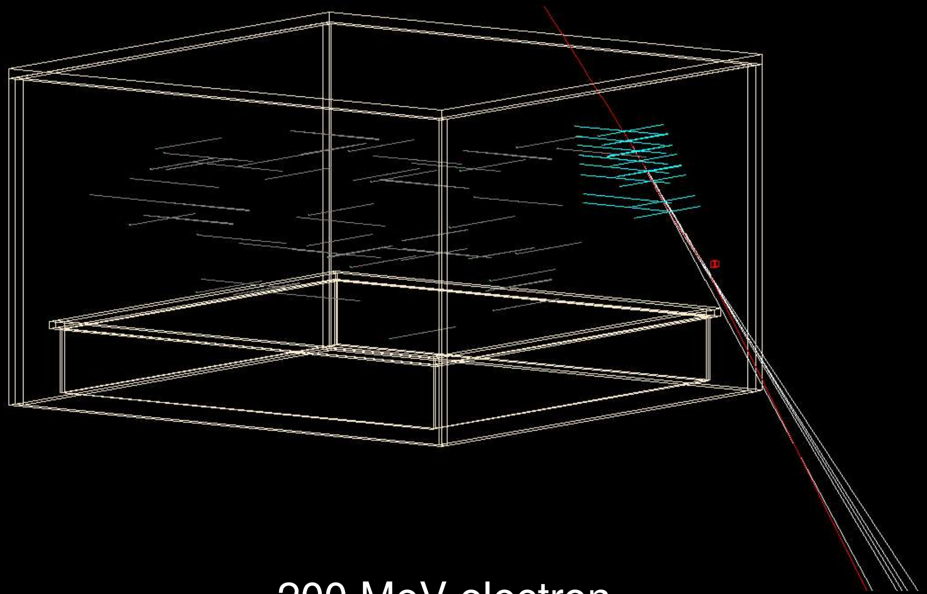
- MSC only: ~74% of the time
- MSC + gamma: ~25% of the time
- MSC + charged particle: ~1% of the time

- gamma production: the electron loses energy which typically is not recorded anywhere in the tracker
- charged particle production: there will be multiple strips that fire, causing ambiguity in the trajectory of the electron (and hence in the scattering angles)

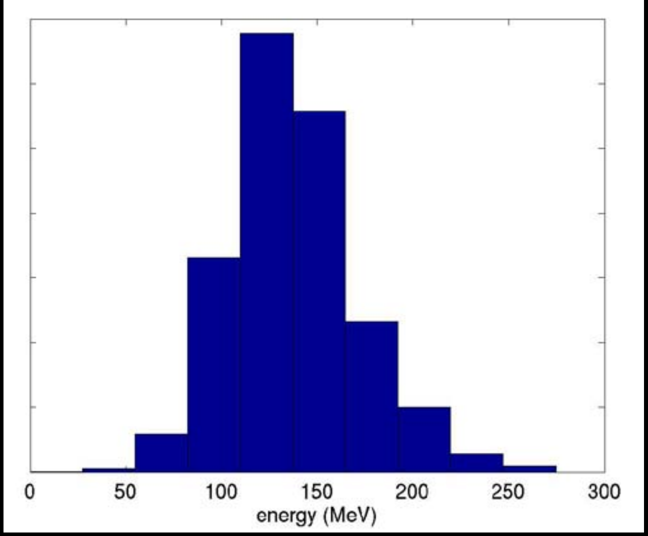
For now, ignore charged particle production

- Energy loss is from two sources
  - multiple scattering - always happens - Landau distribution
  - Bremsstrahlung - happens with prob p_b distributed as 1/e
- Total energy loss distribution is a mixture
  - first component is just multiple scattering
  - second component is convolution of Landau and 1/e distributions
- This has a long tail - up to the energy of the incident electron

Analysis of Electron Events



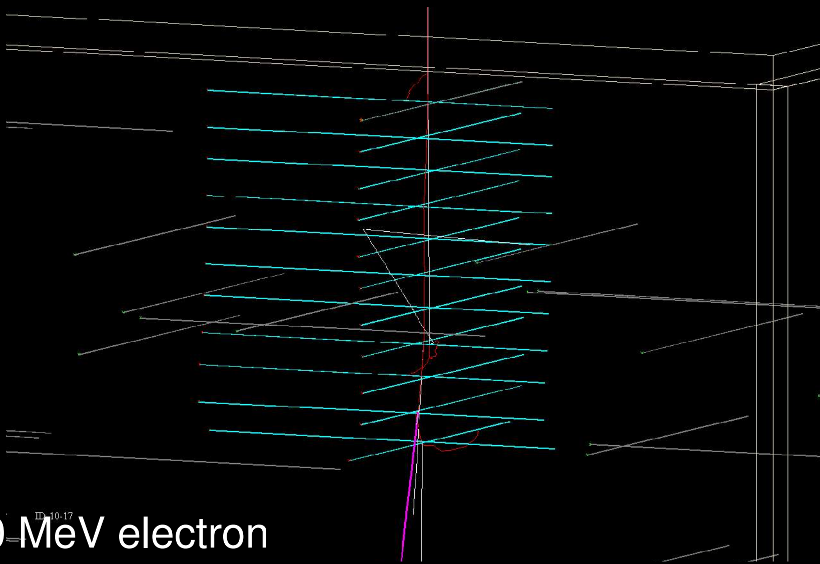
200 MeV electron



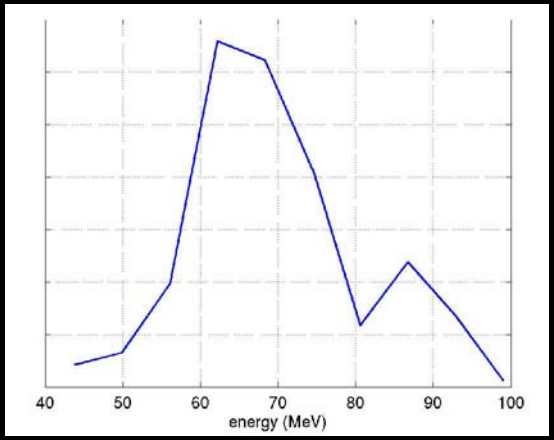
event energy

- The event does not hit the calorimeter; the only energy estimate available is from the tracker
- The event only hits 9 layers; there is not very much information available

Analysis of Electron Events

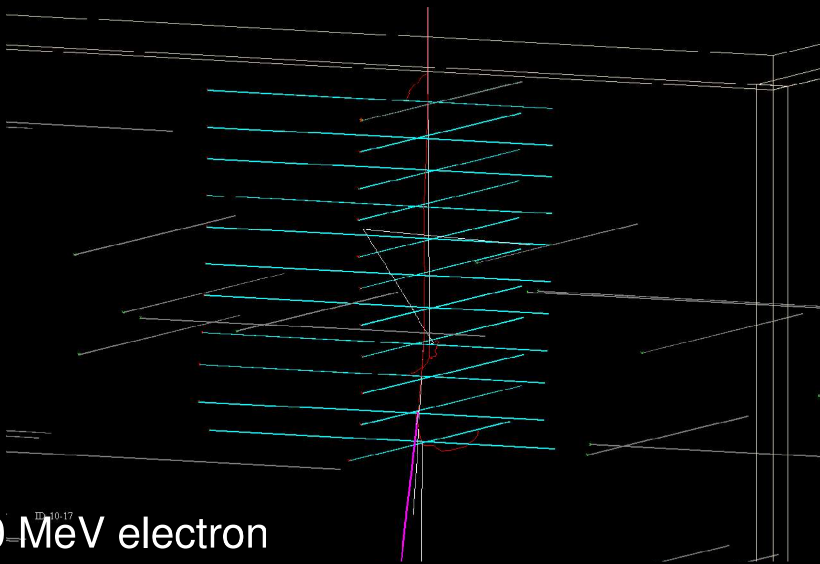


100 MeV electron

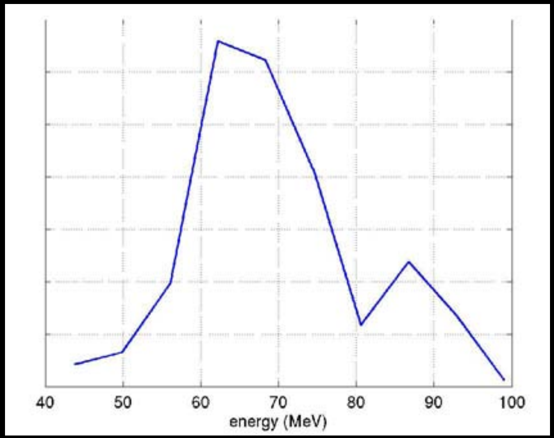


event energy

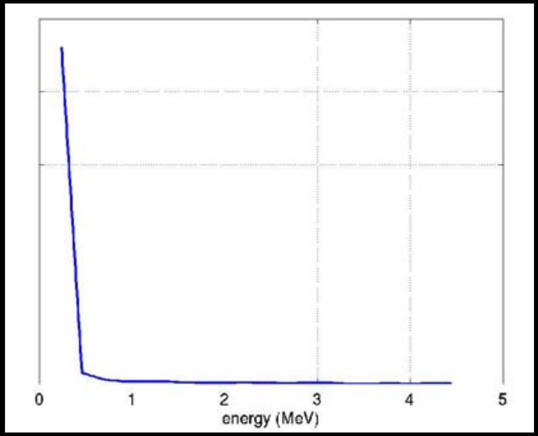
Analysis of Electron Events



100 MeV electron

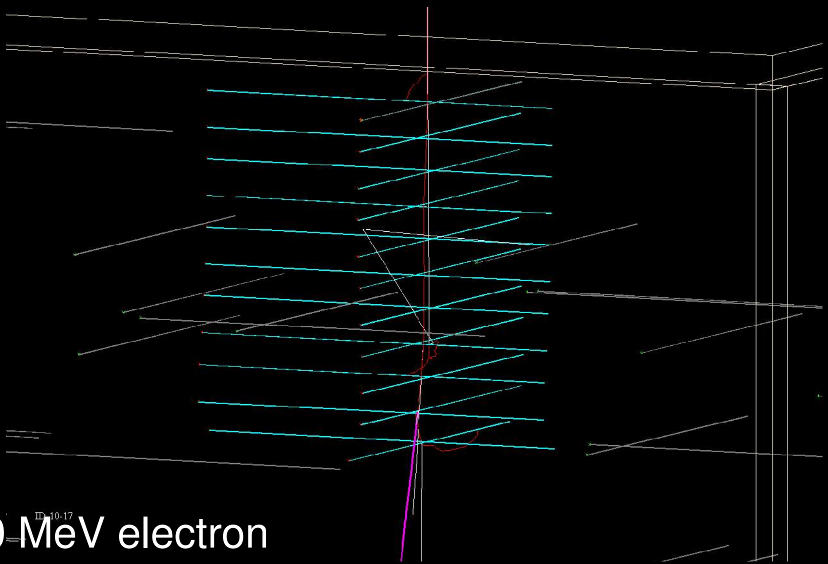


event energy

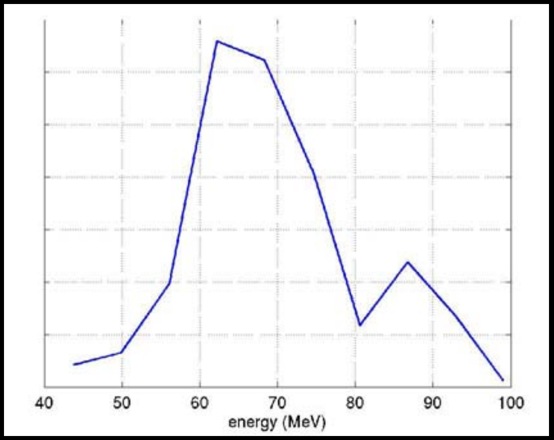


energy loss, layer 9

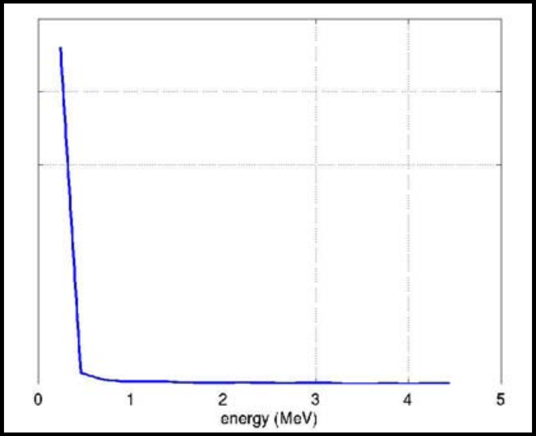
Analysis of Electron Events



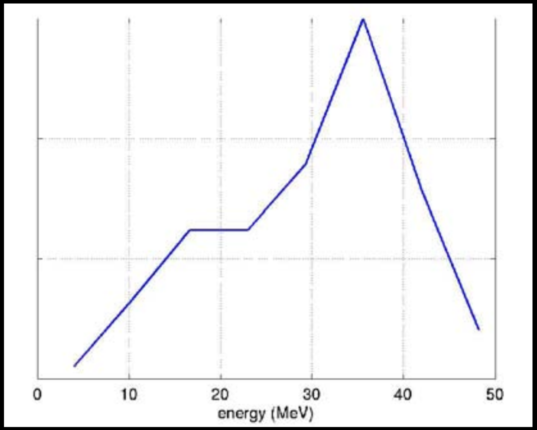
100 MeV electron



event energy

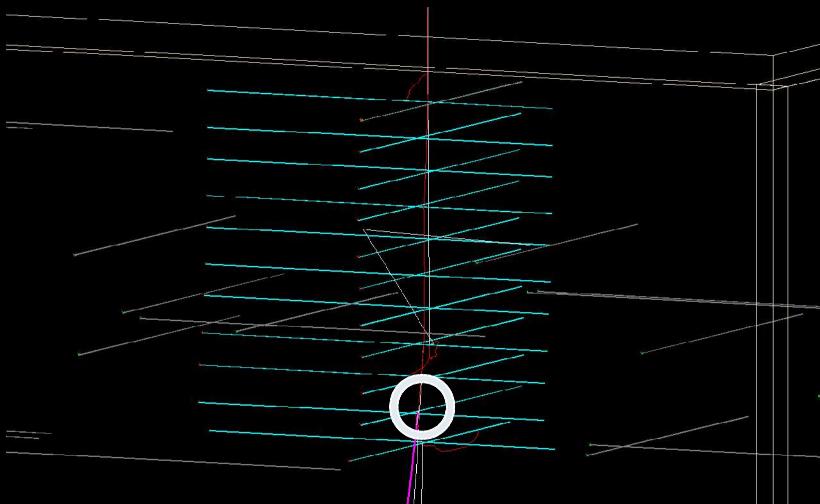


energy loss, layer 9

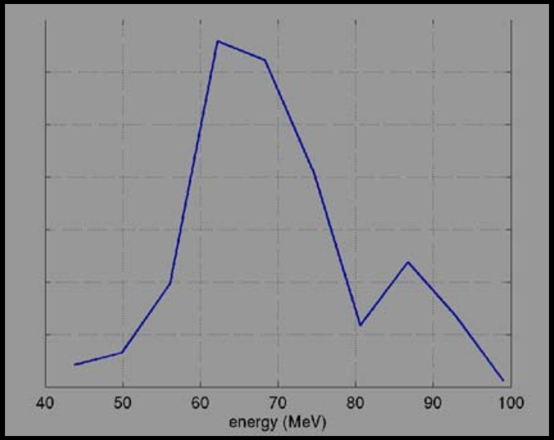


energy loss, layer 10

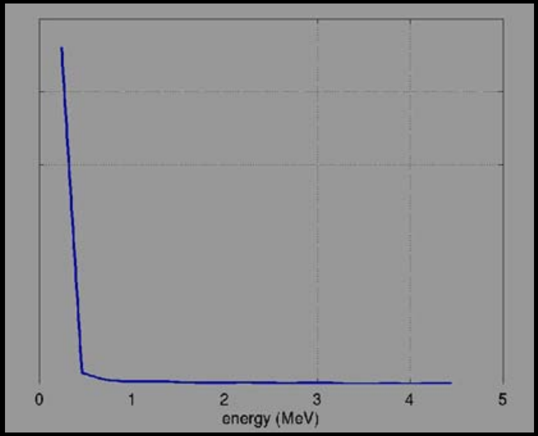
Analysis of Electron Events



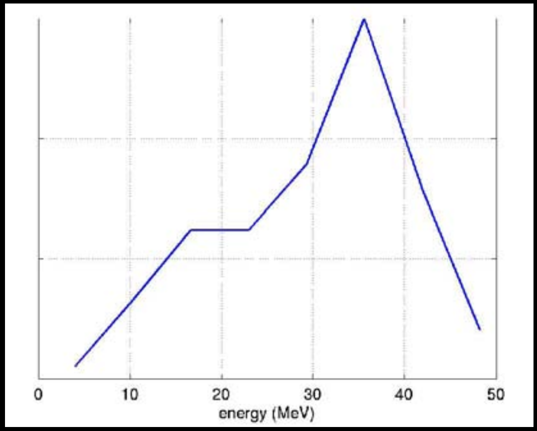
80 MeV photon produced



event energy



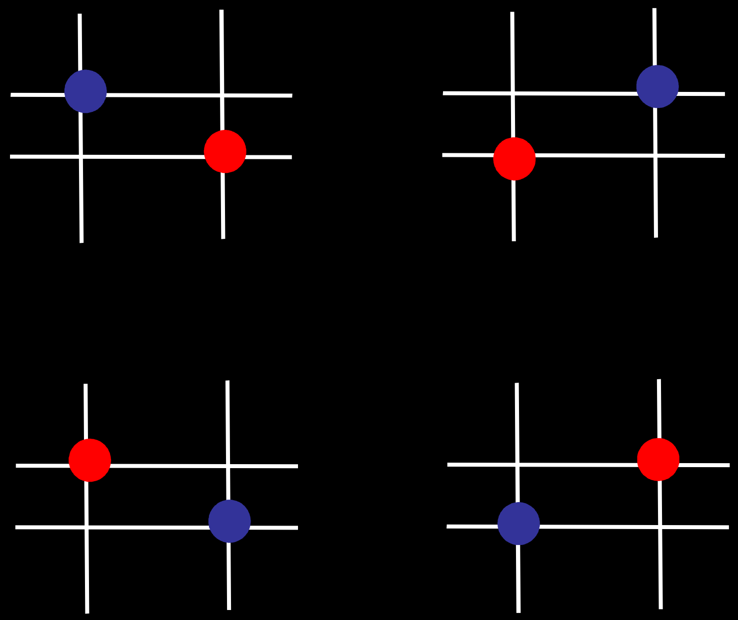
energy loss, layer 9



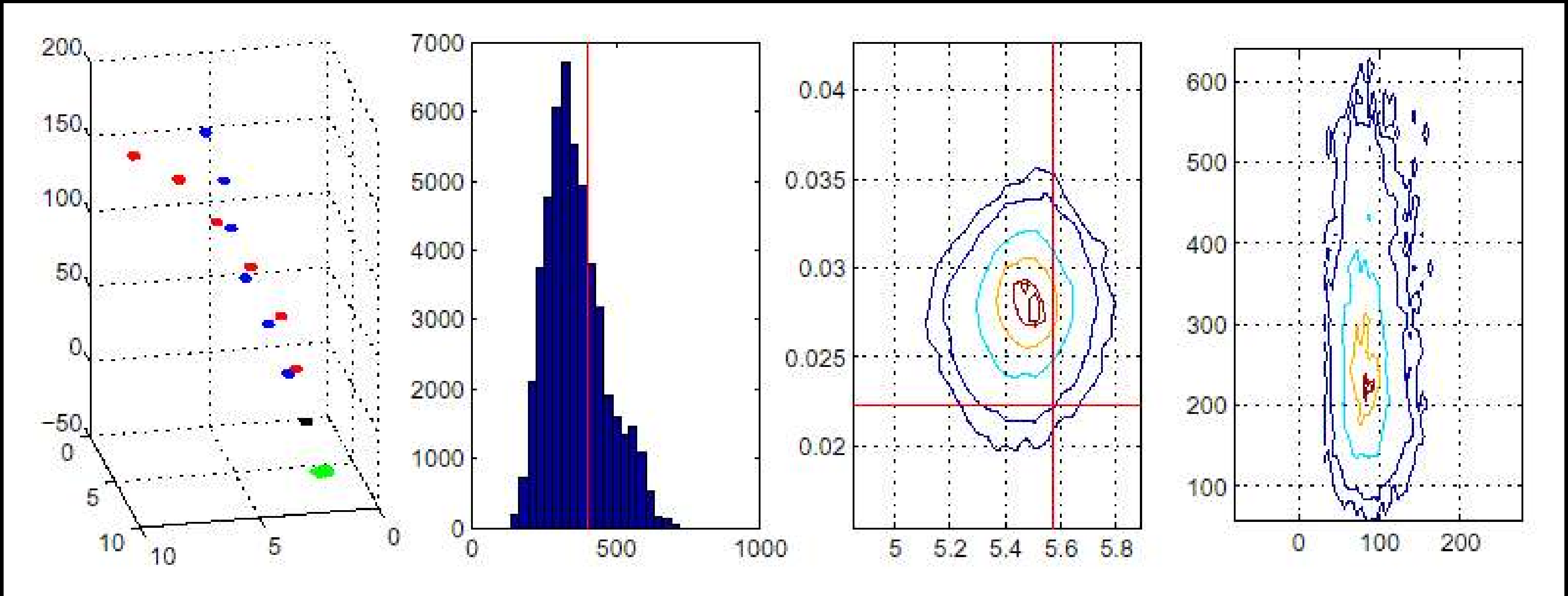
energy loss, layer 10

Gamma Events

- Necessarily more complex as at least two charged particles are present
- Need to consider the possible permutations of the trajectories



Gamma Results



- This event was known not to have any gamma or charged secondaries
- The best geometrical permutation accounts for over 95% of the probability
- The energy is slightly underestimated, especially the energy of the high-energy particle
- The psf is well determined

Current Work

- Including more physics for electrons interacting with the LAT
  - production of secondary charged particles
  - explicit representation of secondary photons
- Importance sampling for estimating the relative probabilities of different hypotheses
- Background rejection by analysing events as cosmic rays
- Extensive testing against the current event analysis methodology

Acknowledgements

- Funding is provided by the NASA Applied Information Systems Program through grant number NNG05GC80G