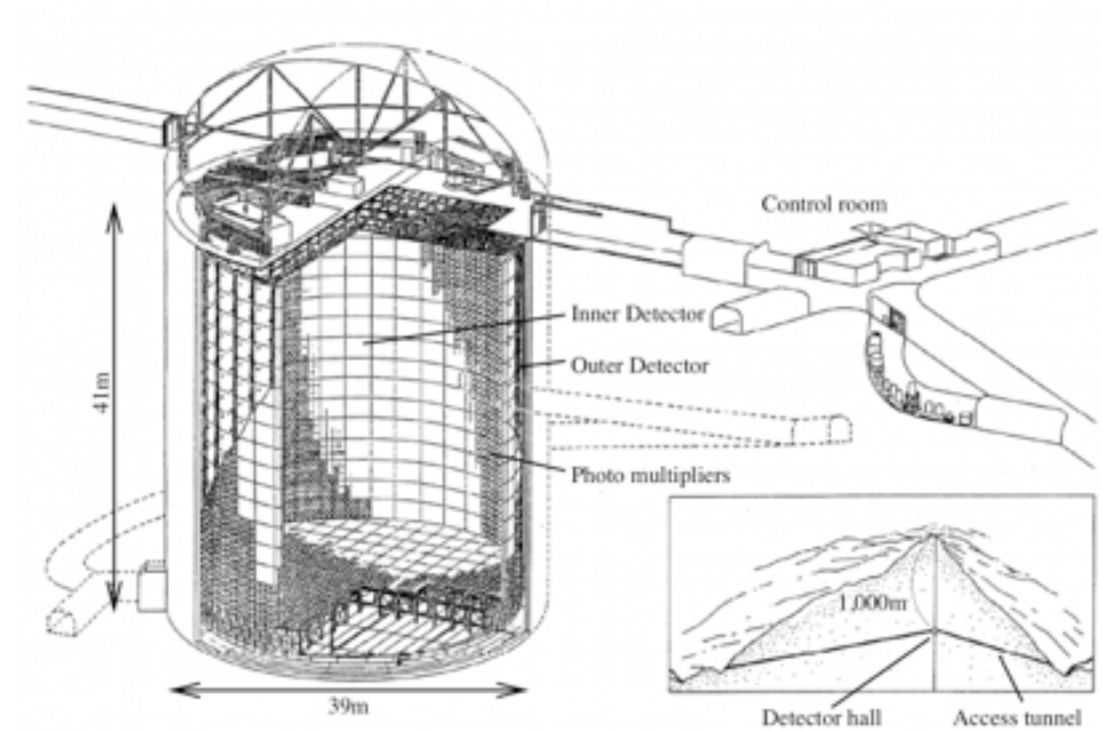
New Event Reconstruction Algorithm for Super-Kamiokande Water Cherenkov Detector

Shimpei Tobayama
University of British Columbia

WNPPC 2012
February 25, 2012

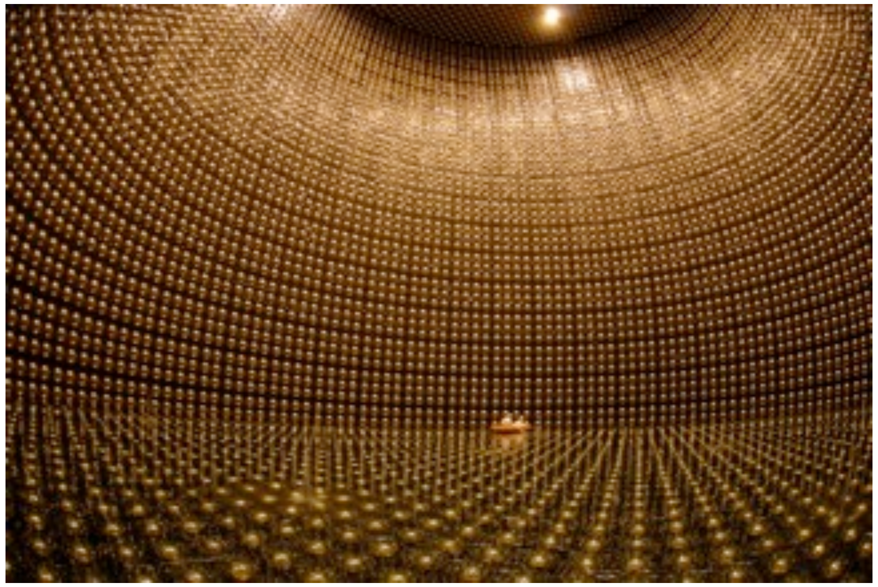
Super-Kamiokande Detector

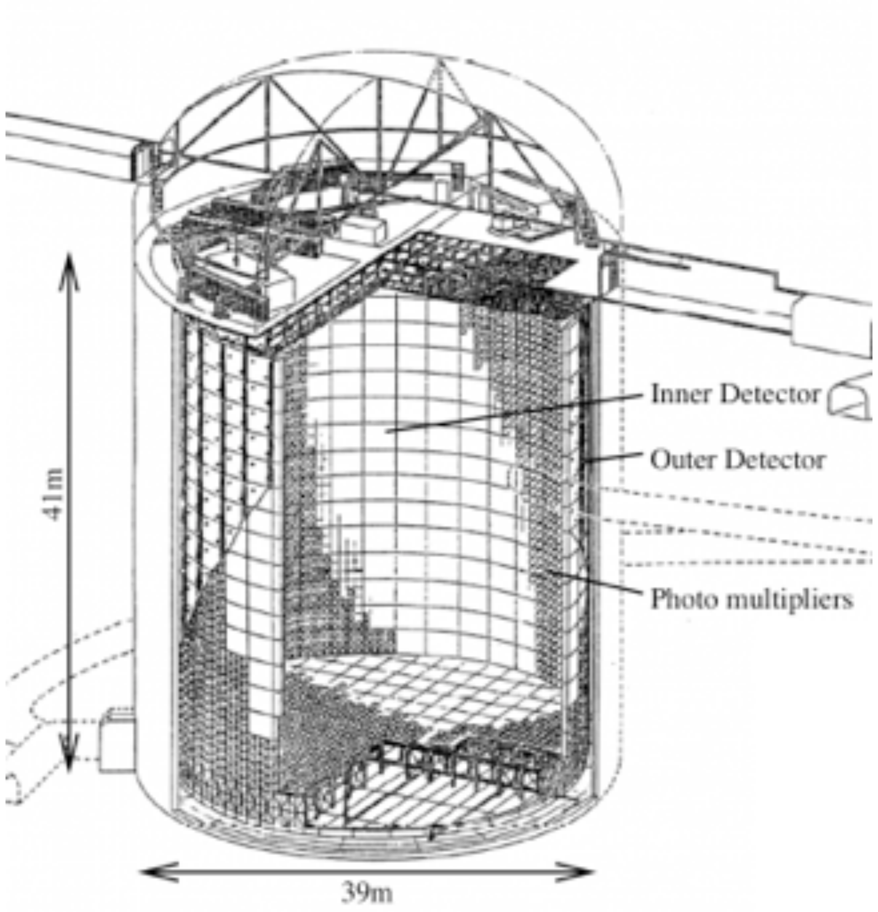
- World's largest water Cherenkov detector
- Kamioka mine, Gifu, Japan
- Objectives:
  - Proton decay search
  - Neutrino detection
    - Atmospheric, solar and supernova
    - Far detector for long baseline neutrino oscillation experiments
      - K2K(1999~2004), T2K(2009~)



Detector structure

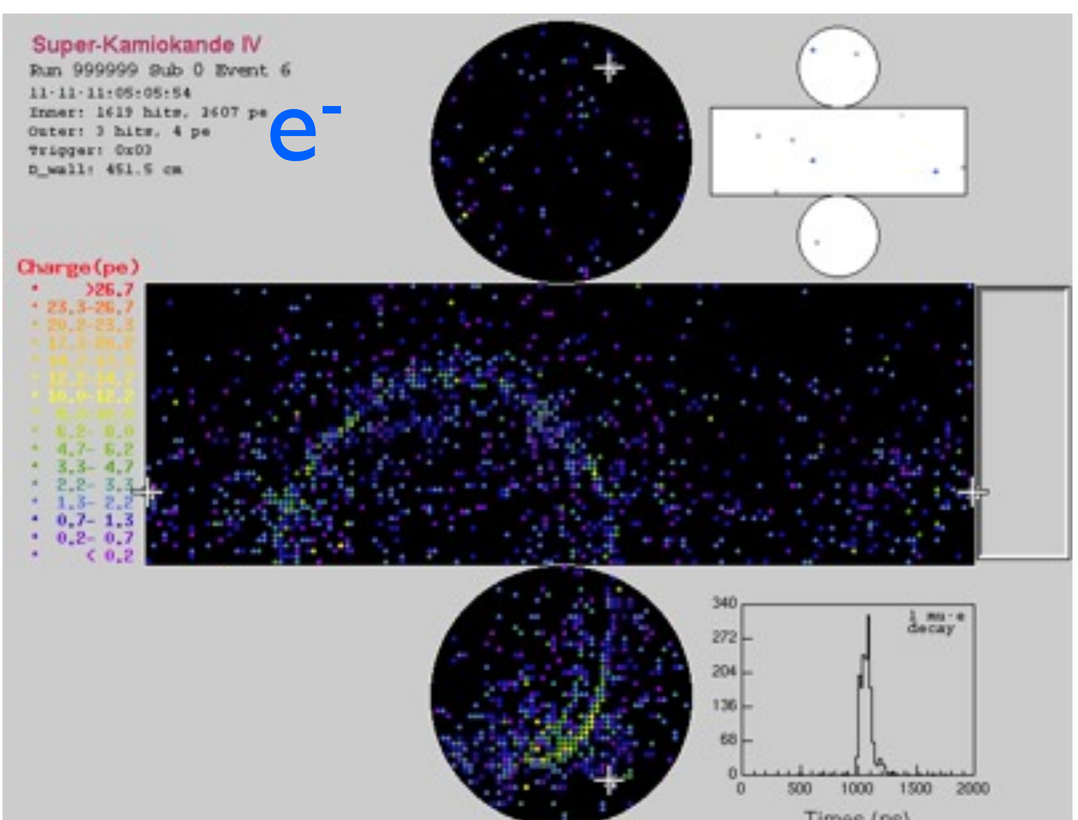
- Cylindrical tank, 50kt pure water
  - 22.5kt fiducial volume
- 11,146 inward-facing 20inch PMT's
  - Record Cherenkov photons emitted by charged particles propagating in water

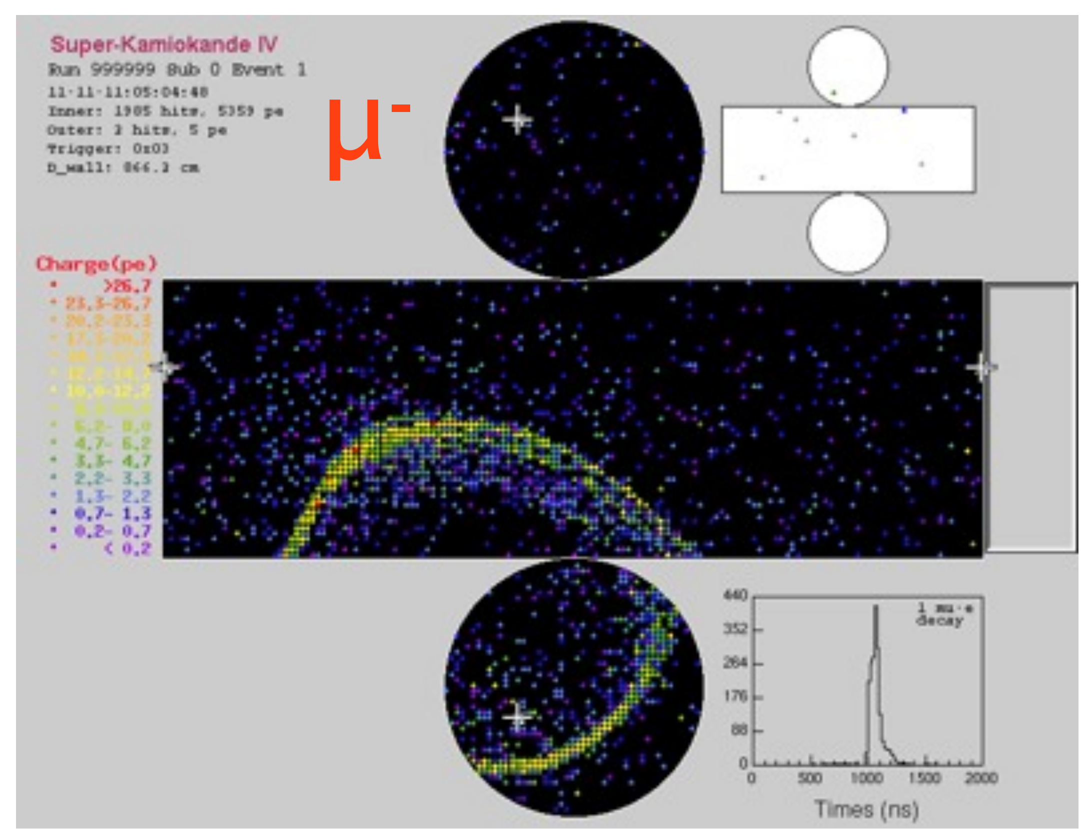




Cherenkov Ring Patterns

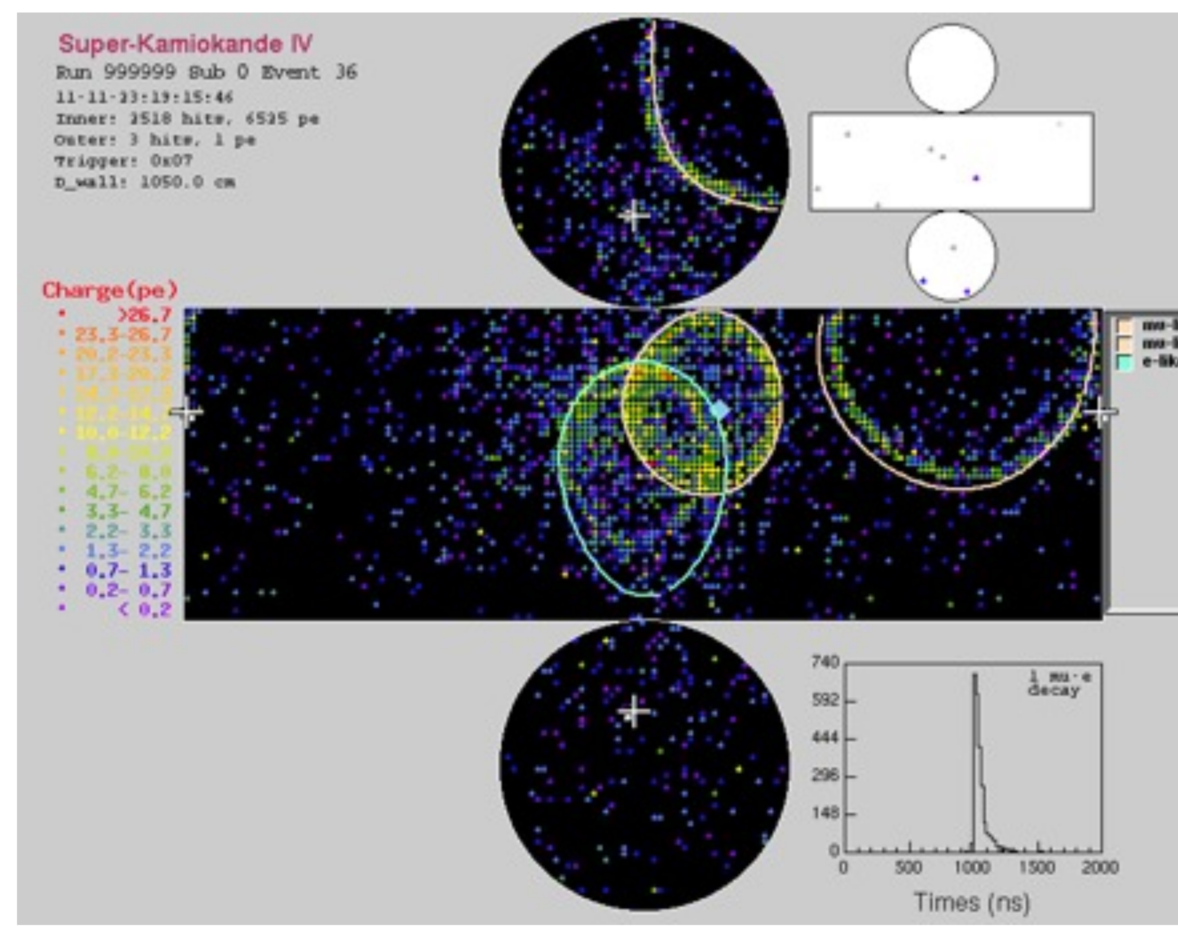
- Ring pattern differs between particle types:
  - e^-,γ: blurry ring (caused by EM shower)
  - μ^-: sharp ring
- The difference is used for particle identification





Event Reconstruction

- From charge and hit timing information from the PMT's, events are reconstructed
- The current reconstruction software (apfit) determines step-by-step:
  - Vertex position
  - Number of rings
  - Direction
  - Particle type
  - Momentum



A New Algorithm

- Simultaneous maximum likelihood fitting of track parameters
  - Similar methods are used in MiniBooNE
  - Correlations between parameters will be taken into account automatically
- Naturally extendable to multi-ring fit
  - Fitting π^+($\leftarrow$not done by current reconstruction software) and π^0
    - Main backgrounds for T2K

Maximum Likelihood Fit

- Find a set of parameters **x** that maximizes the Likelihood function:

i-th PMT's charge, time

$$L(\mathbf{x}) = \prod_i^{\mathrm{unhit}} P(i\mathrm{unhit}|\mathbf{x}) \prod_i^{\mathrm{hit}} P(i\mathrm{hit}|\mathbf{x}) f_q(q_i|\mathbf{x}) f_t(t_i|\mathbf{x})$$

Unhit probability   Hit probability   Charge likelihood   Time likelihood

- For single particle events, parameters are:
  - vertex, direction, momentum
- Parameters are fit simultaneously
- PID by compering the best-fit likelihood values of different hypotheses

Predicted Charge

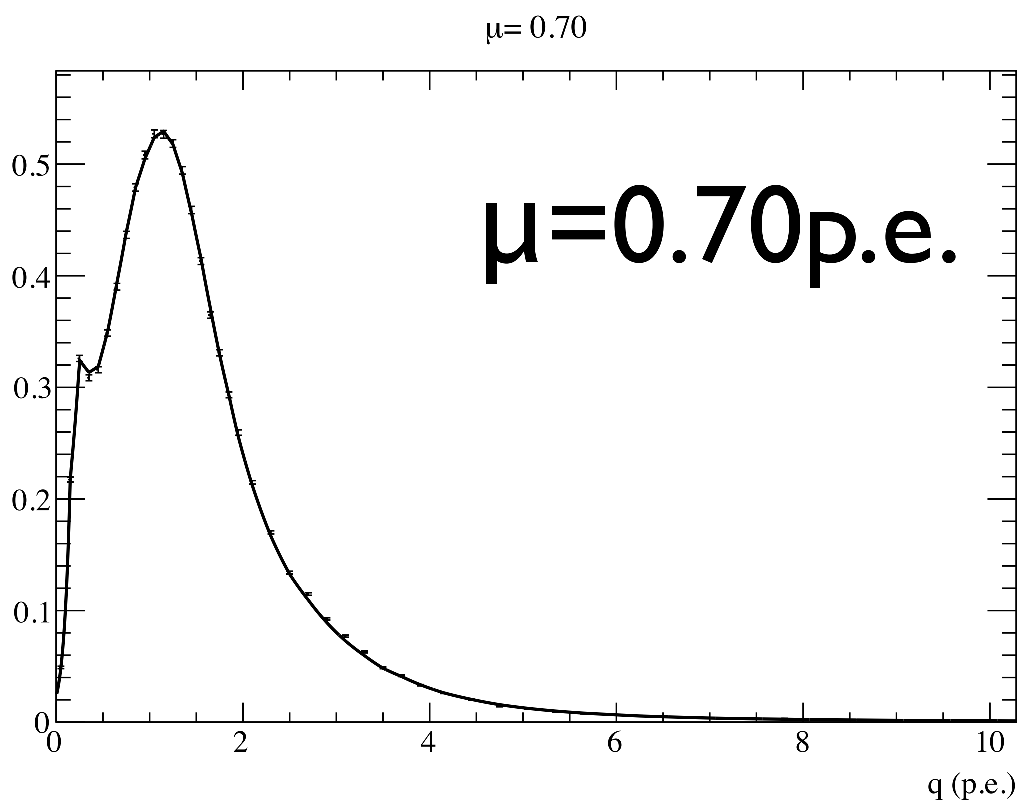
- For charge likelihood, we can decouple:
  - Particle & photon propagation in the detector
  - PMT response/properties of electronics
- Given track parameters **x**, calculate the expected number of photoelectrons liberated at each PMT:
  - Predicted charge μ_i
- Hit/unhit prob., charge PDF depends on **x** through μ_i

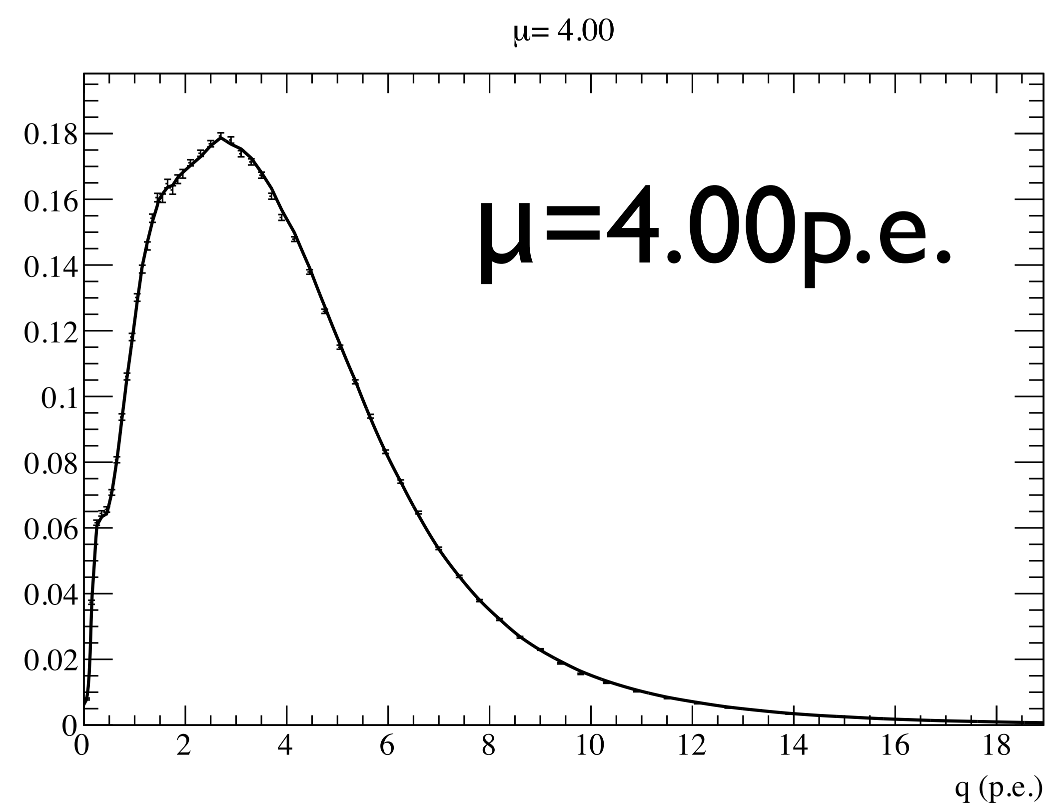
$$P(i\text{unhit}|\mu_i) = (1 + \text{corrections})e^{-\mu_i}$$

$$P(i\text{hit}|\mu_i) = 1 - P(i\text{unhit}|\mu_i)$$

Charge PDF $f_q(q|\mu)$

- Probability of observing charge q given mean μ
- Produced using detector MC
- Fit in order to smooth likelihood surface

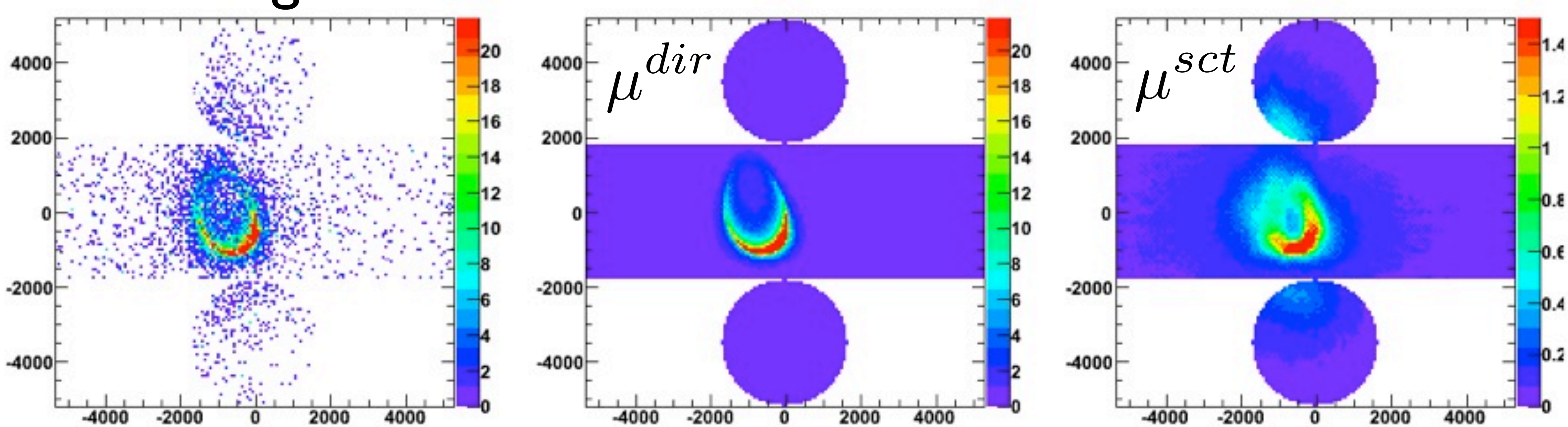




Direct & Scattered Light

- The task now is to calculate μ for each PMT
- Predicted charge is divided into: $\mu = \mu^{dir} + \mu^{sct}$
  - μ^{dir}: Direct light
  - μ^{sct}: Scattered light (includes reflections)

Hit PMT charge↓

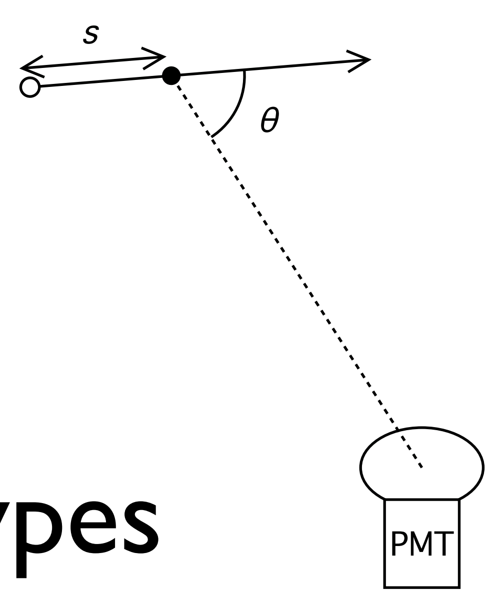


*scale is different

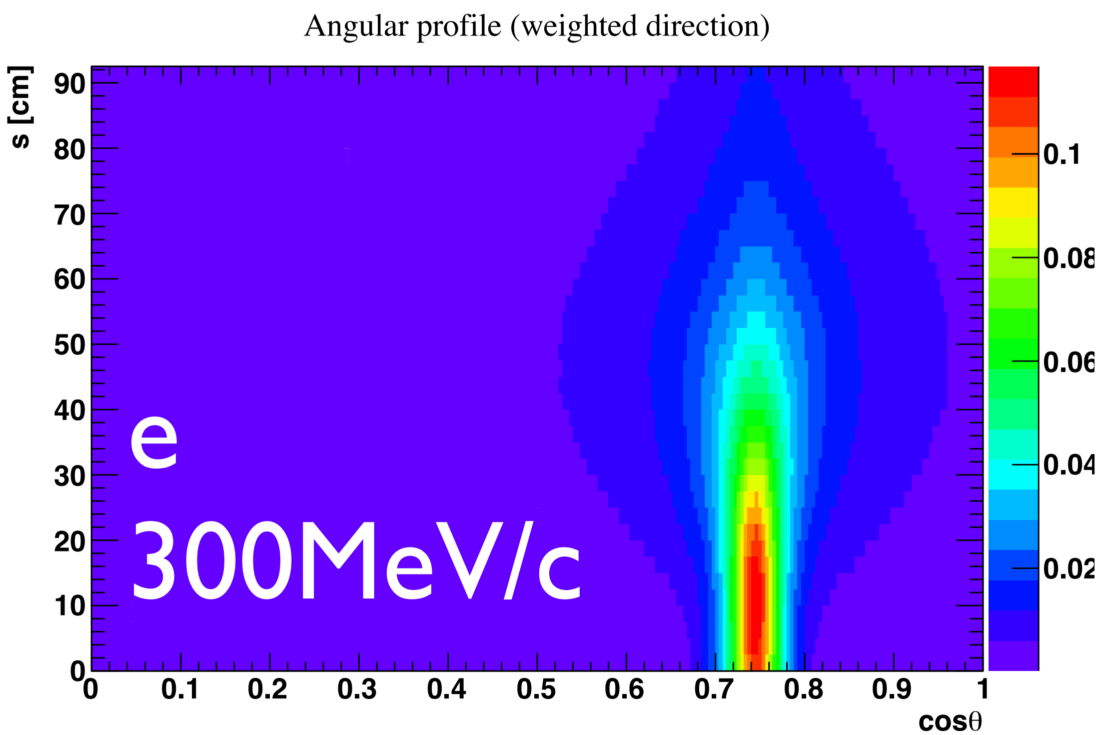
Cherenkov Emission Profile

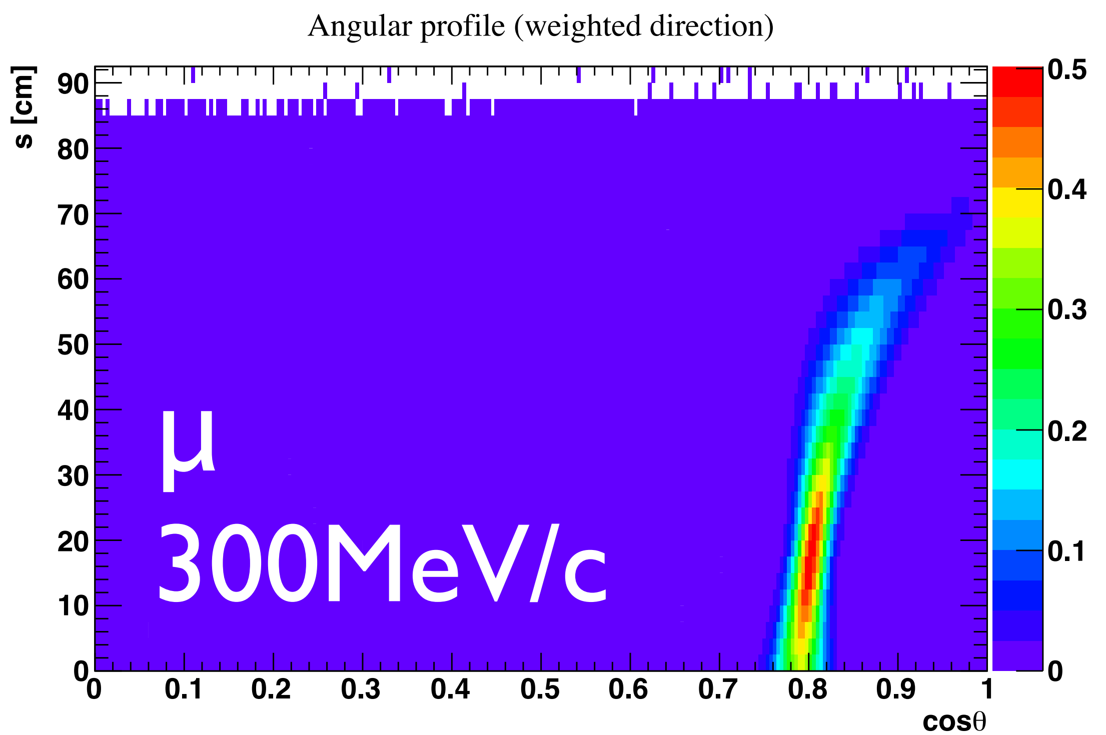
$$\mu^{\text{dir}} = \Phi(p) \int ds\, g(s, \cos\theta) \Omega(R) T(R) \epsilon(\eta)$$

↑also momentum-dependent



- Rate of photons emitted in direction θ
- Profiles are different between particle types
  - Differentiates particle types





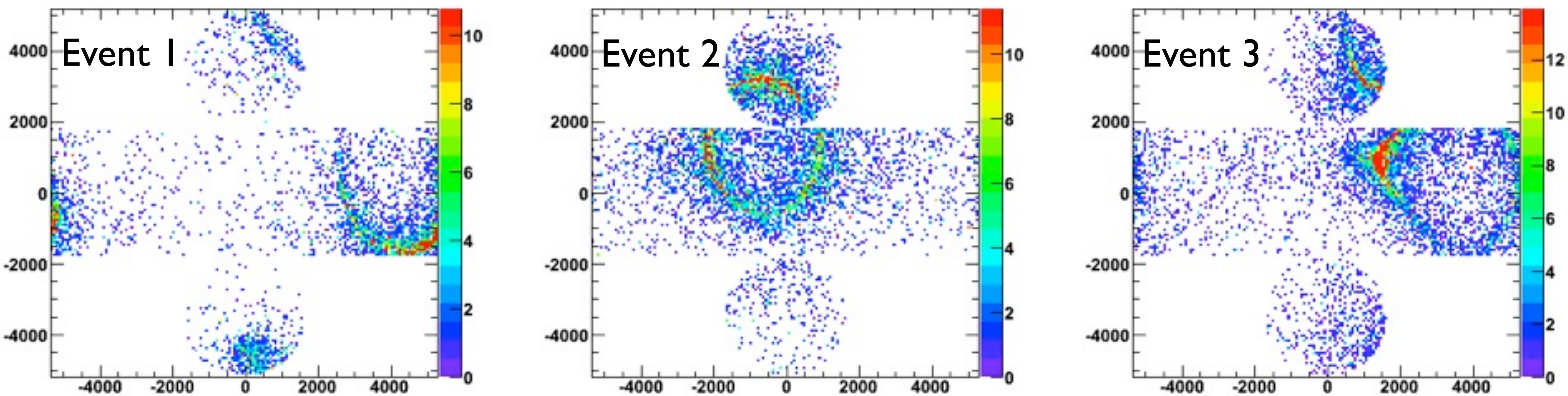
Fitter Development

- Single ring fitter is complete
  - Reconstructs single e/μ events
- Effort was made to reduce computation time
  - Predicted charge calculations/likelihood evaluations are done ~10^7 times per event
    - >800 MINUIT iterations/event x11k PMTs
    - Without care it will easily take minutes
  - 22s/event to run both e/μ hypotheses

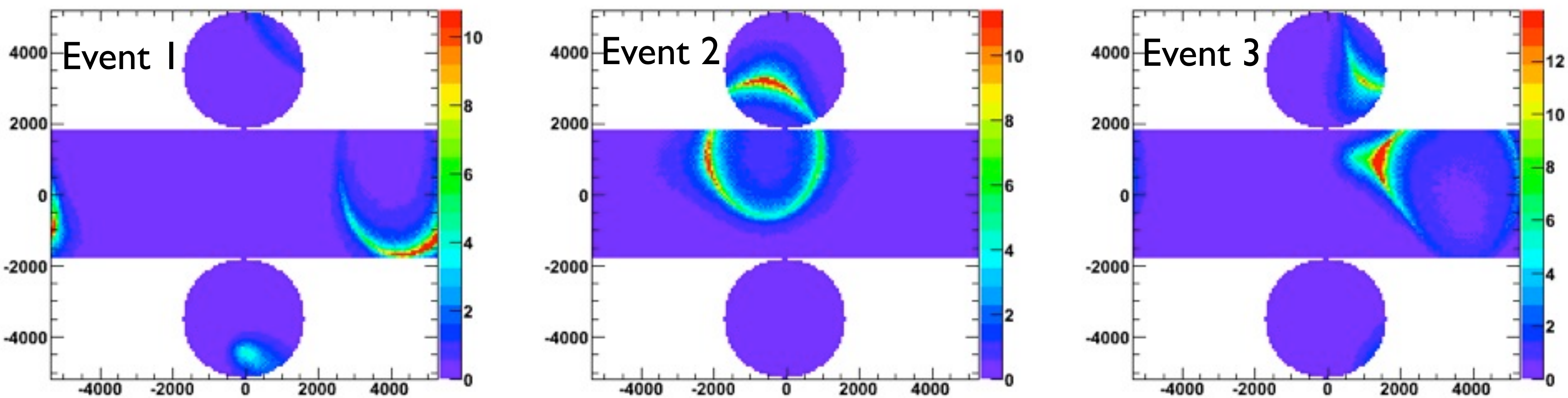
Charge Distributions: e^-

Example fit results

Hit charge information from PMT:



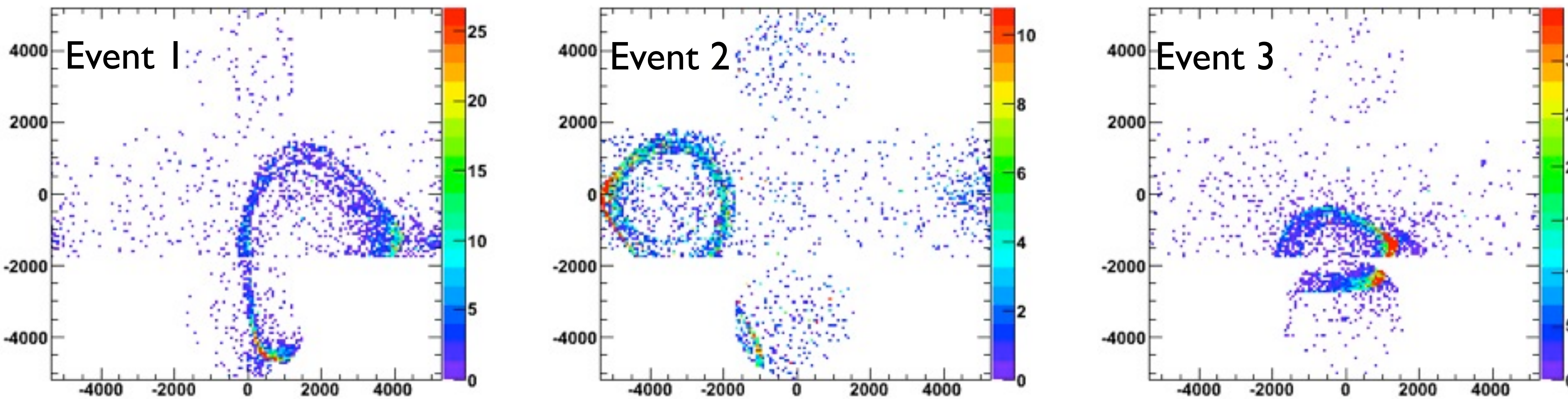
Predicted charge after fit:



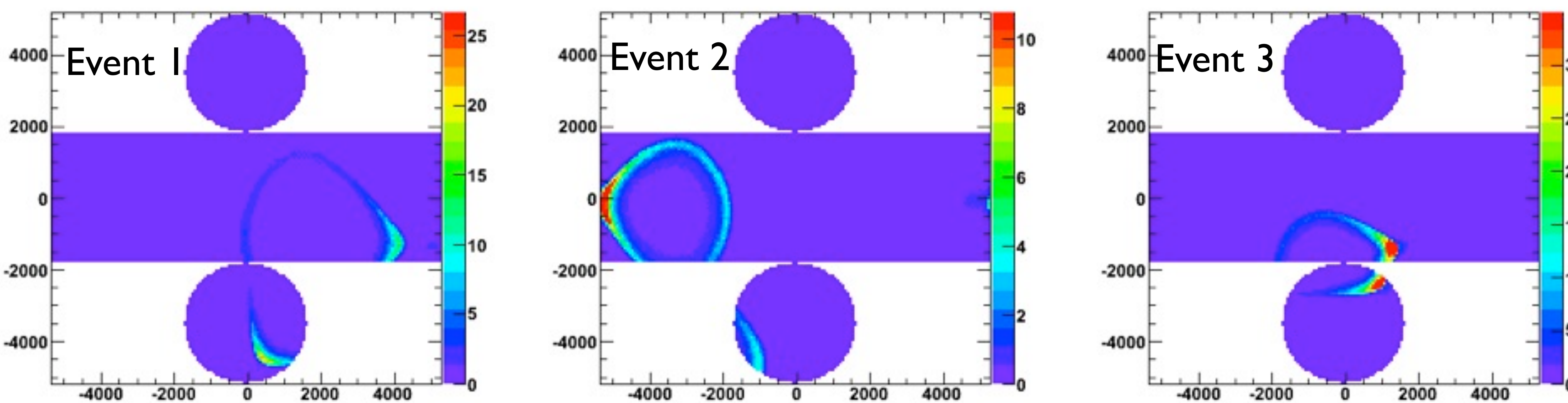
Charge Distributions: μ^-

Example fit results

Hit charge information from PMT:

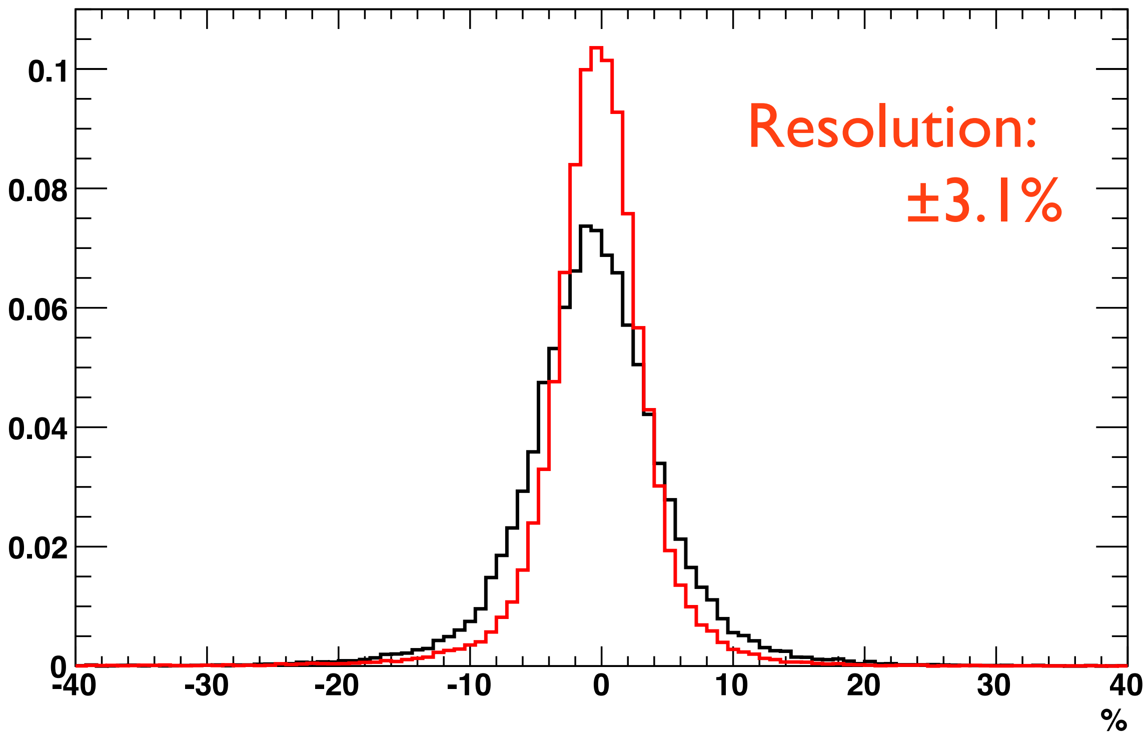


Predicted charge after fit:



Performance: Single e event

Tru/Rec momentum percentage difference

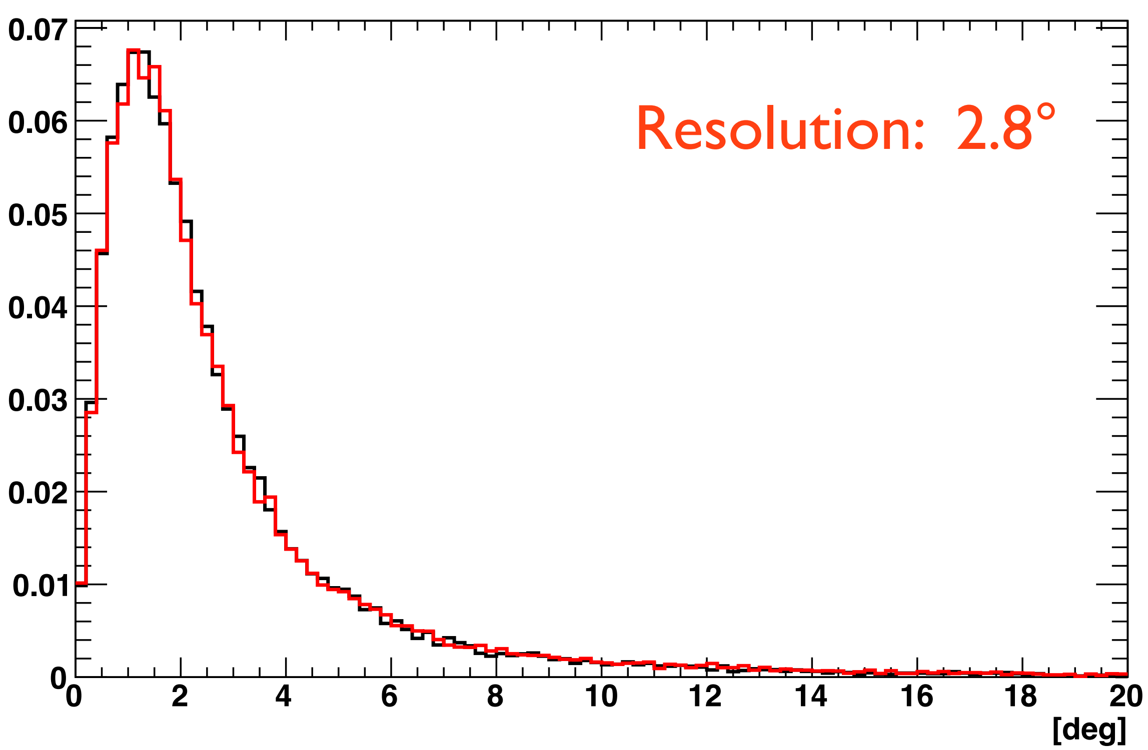


*Uniformly random p<1GeV/c
Fully contained, fiducial volume events

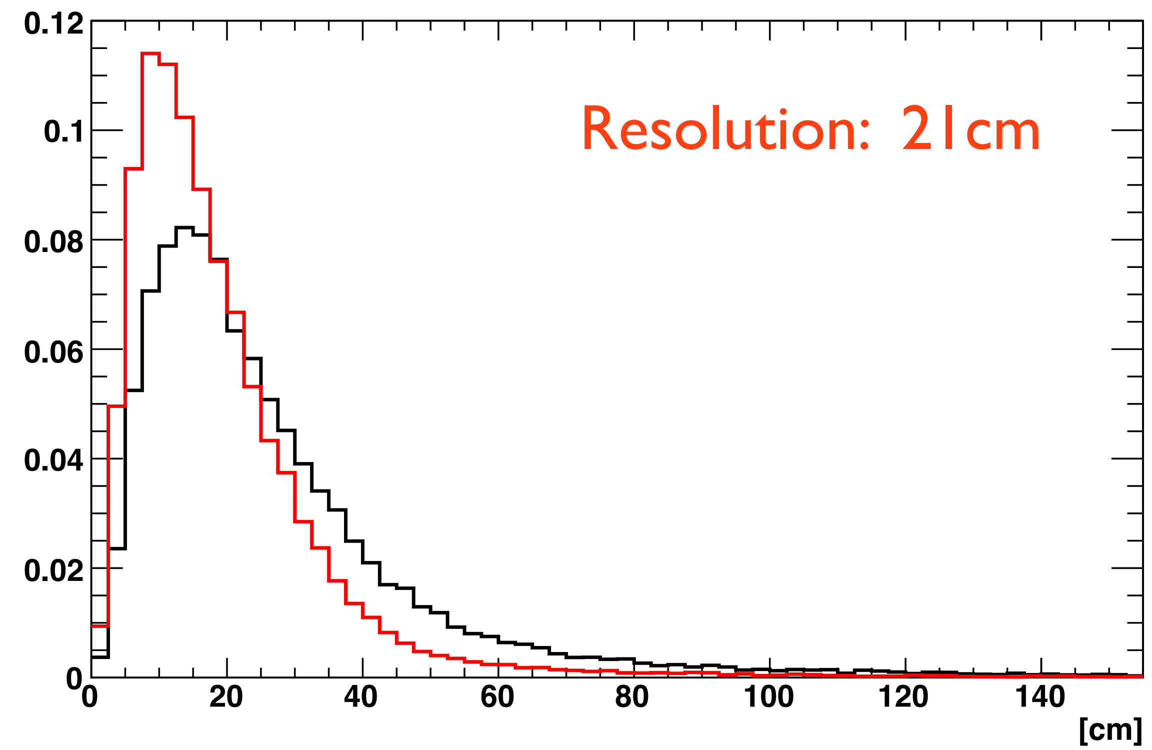
In plots: Existing fitter
New fitter

Significant improvement in momentum & vertex resolution

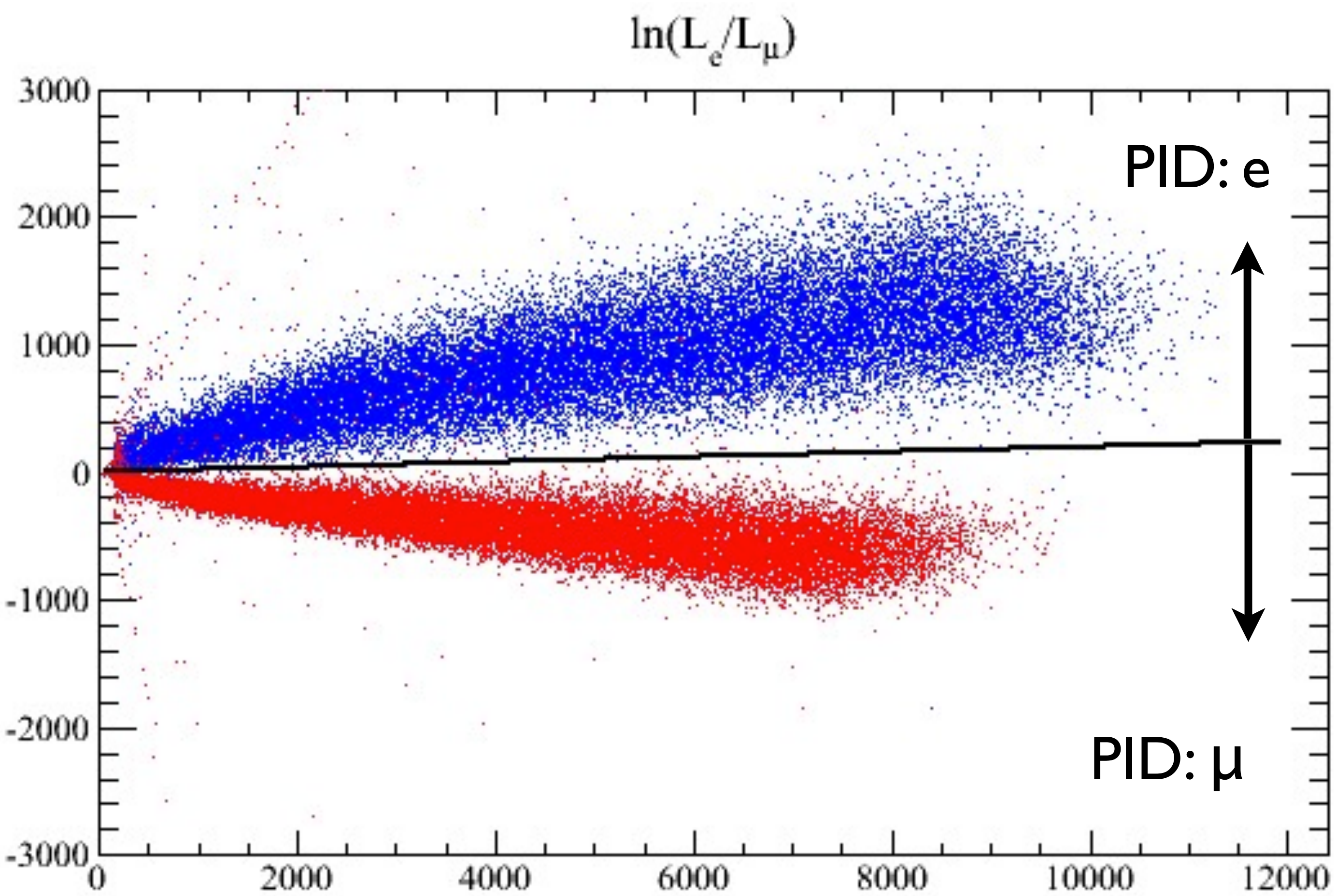
Angle between tru/rec direction



Distance between tru/rec vertex



Particle Identification



Electron evens

Miss ID rate: 0.20%*

Muon Events

Miss ID rate: 0.85%*

*Contains low energy events not relevant for T2K and atmospheric neutrino analysis
Such events are eliminated by selection cuts, and consequently miss ID rate decreases

- Good e/μ separation
- Miss ID is dominantly due to MINUIT minimization not converging properly
- Will improve significantly by optimizing seeding

Summary & Plans

- Single particle event reconstruction performance is improved compared to the current algorithm
  - Further performance improvement expected by considering wavelength dependencies, water quality variations in tank etc.
- Multi-ring fitters to be developed
  - Fitting π^+, π^0 tracks
- Systematic studies

Backup

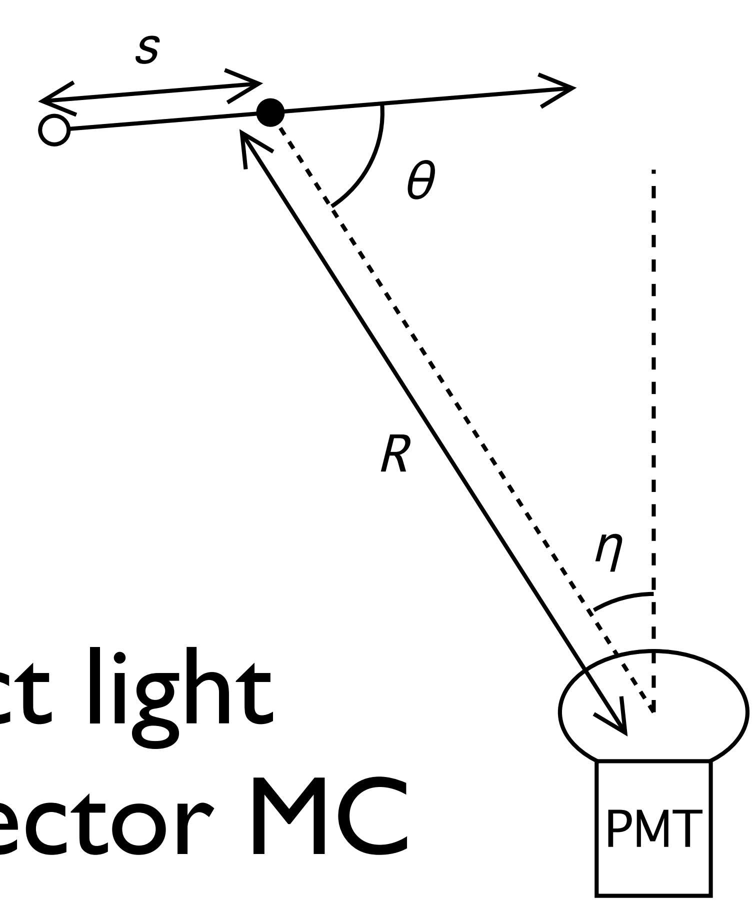
Calculating Predicted Charge

$$\mu^{\text{dir}} = \Phi(p)\int ds\, g(s,\cos\theta)\Omega(R)T(R)\epsilon(\eta)$$

- $\Phi(p)$: Light yield
- $g(s,\cos\theta)$: Cherenkov emission profile
- $\Omega(R)$: Solid angle subtended by PMT
- $T(R)$: Light transmission function
- $\epsilon(\eta)$: PMT angular acceptance

s: Distance particle traveled along the track

-R,θ,η are functions of s



$$\mu^{\text{sct}} = \Phi(p)\int ds\, g(s)\Omega(s)T(s)\epsilon(s)A(s)$$

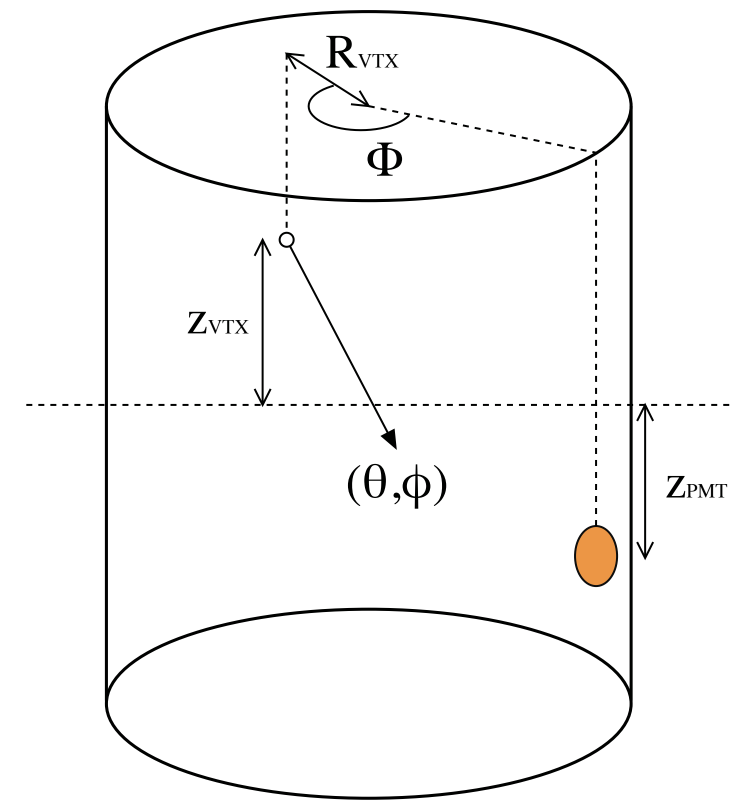
$A(s)$: Scattering table

-Ratio of scattered light to direct light

-Tabulated in advance using detector MC

Scattered Light

$$\mu^{\rm sct} = \Phi(p_0) \int ds \rho(s) \Omega(s) T(s) \epsilon(s) A(s)$$



Scattering table:

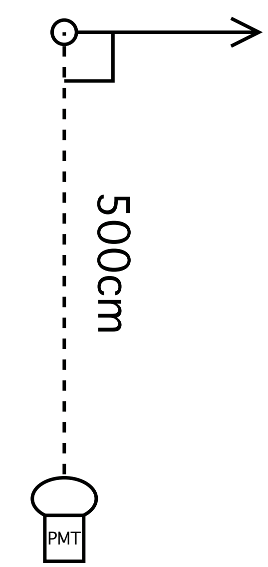
$$A(s) = A(z_{\rm PMT}, z_{\rm vtx}, r_{vtx}, \Phi, \Theta, \phi) = \frac{d\mu^{\rm sct}}{d\mu^{\rm dir}}$$

- Ratio of scattered light to direct light
- Using detector MC, shoot 3MeV/c electrons(point-like Cherenkov source) randomly, and fill a 6-D histogram
  - Multiple scattering was turned off
- Linearly interpolate at run time

Parabolic approximation

Integral is CPU-intensive

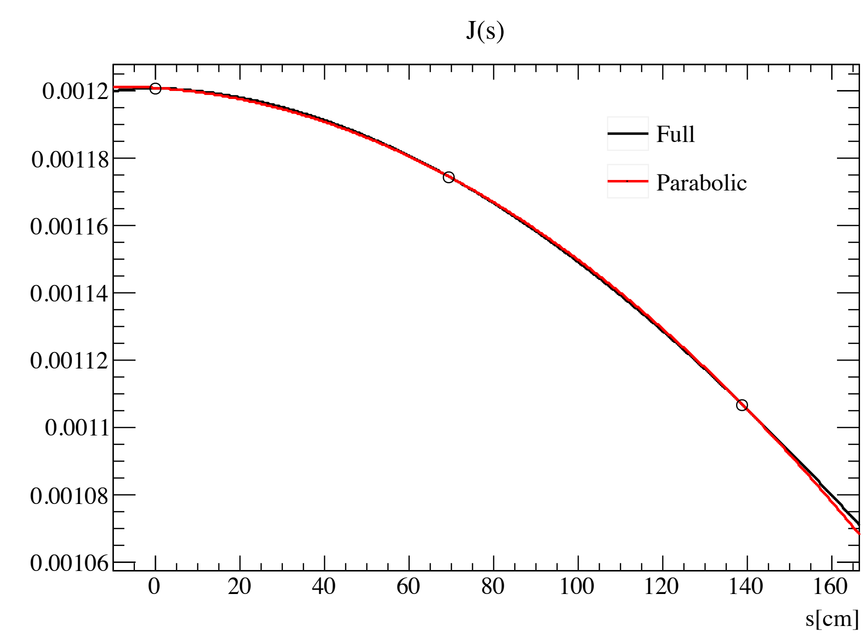
$$\mu^{\rm dir} = \Phi(p) \int ds g(s) \Omega(s) T(s) \epsilon(s)$$



Acceptance factors $J(s) = \Omega(s) T(s) \epsilon(s)$ vary slowly as a function of s

Approximate it by parabola, perform the integral in advance

$$\begin{aligned} \mu^{\rm dir} &= \Phi(p) \int ds g(s) (j_0 + j_1 s + j_2 s^2) \\ &= \Phi(p) (I_0 j_0 + I_1 j_1 + I_2 j_2) \end{aligned}$$

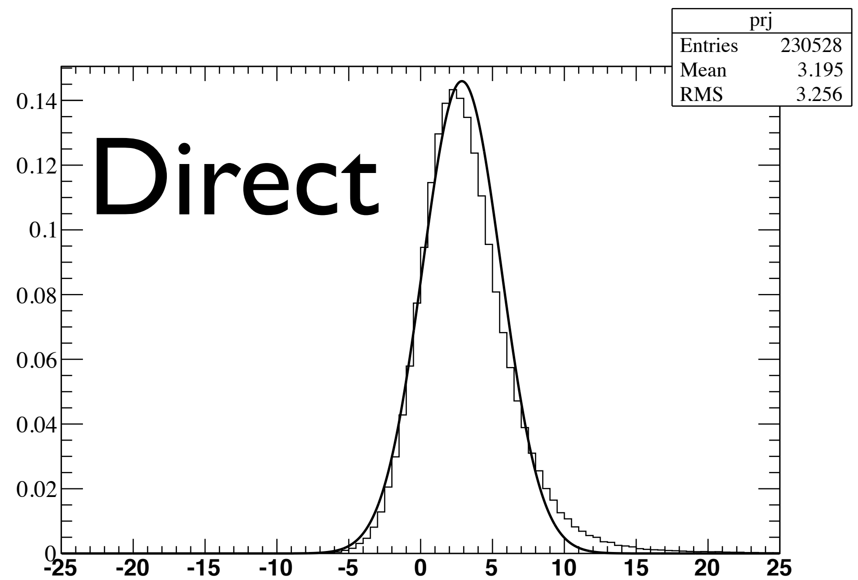


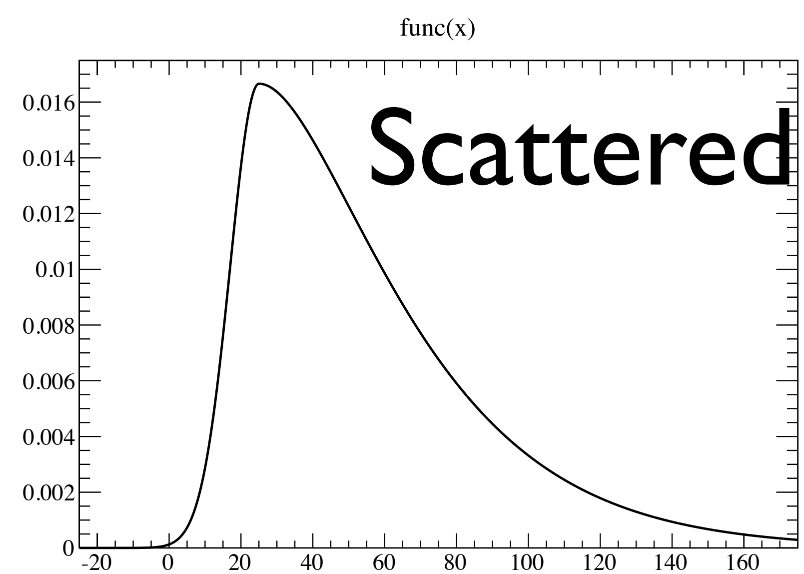
Determine j_n, read off the integral values I_n from table

Similar thing is done for scattered light

Time Likelihood

- Make primitive time likelihood functions separately for direct and scattered light





- Based on the relative strength of direct/scattered light, time PDF is constructed on the fly:

$$f_t(t_c) = w \cdot f_t^{\mathrm{dir}}(t_c) + (1 - w) \cdot f_t^{\mathrm{sct}}(t_c)$$