Low-energy reconstruction methods & performances in ice

João Pedro A. M. de André
for the IceCube-PINGU Collaboration

Penn State Univesity

20 September 2014

Overview

1. Reconstruction method: HybridReco/MultiNest
   - used in our *Letter of Intent*
   - uses “hybrid” particle hypothesis and MultiNest as “minimizer”

2. Particle IDentification in PINGU
   - used in our *Letter of Intent*

Why are we using MultiNest to reconstruct data?

- Reconstructing data means finding which particle(s) best matches our data
  - This typically is a minimization/maximization problem

- To solve this problem the most used minimizers are MIGRAD and SIMPLEX
  - "*This [MIGRAD] is the best minimizer for nearly all functions.*"

    Minuit Users guide, ROOT
  - "*Its [MIGRAD] main weakness is that it depends heavily on knowledge of the first derivatives, and fails miserably if they are inaccurate*"

    Minuit Users guide, ROOT

- Unfortunately our likelihood space is **not** a smooth distribution, neither MIGRAD nor SIMPLEX work well with it

- We want a minimizer that works better for our likelihood space

  $\Rightarrow$ use MultiNest for minimization!

The MultiNest algorithm

A full description is presented in a paper by F. Feroz et al. (arXiv:0809.3437).

- The MultiNest algorithm searches for the maximum in a multidimensional likelihood by scanning the likelihood space.

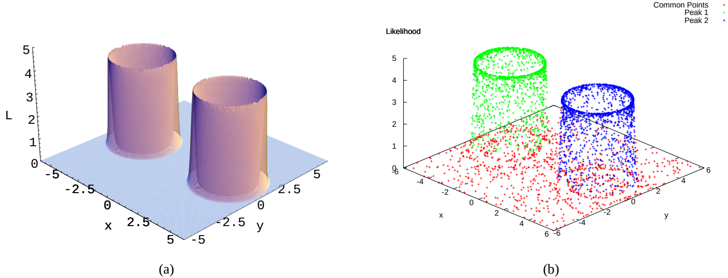
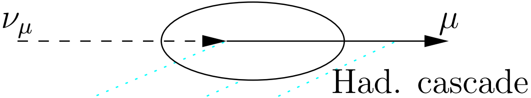


Figure 6. Toy model 2: (a) two-dimensional plot of the likelihood function defined in Eqs. (32) and (33); (b) dots denoting the points with the lowest likelihood at successive iterations of the MultiNest algorithm. Different colours denote points assigned to different isolated modes as the algorithm progresses.

(Figure extracted from arXiv:0809.3437)

The HybridReco/MultiNest hypothesis

- Goal: reconstruct ν_μ CC (DIS) interactions (total 8 parameters)



4 parameters: vertex(3), time

3 parameters: direction(2), energy

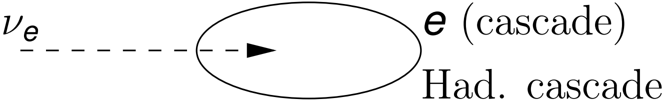
3 parameters: direction(2), energy/length

assume same direction for μ and cascade

- To calculate likelihood of physics hypothesis use millipede
  - relies on tables for μ and E.M. cascade light expectation
  - μ is segmented according to μ-table requirements
    - tables are made for a given track hypothesis/length
    - recent creation of new improved tables
  - after fit E.M. cascade energy converted to Had. cascade energy
- Currently being used in several DeepCore analysis and **PINGU**
- Can run in *every* event on the sample

Reconstructing ν_e CC events as ν_μ CC events

- ν_e CC interaction produces superposing cascades

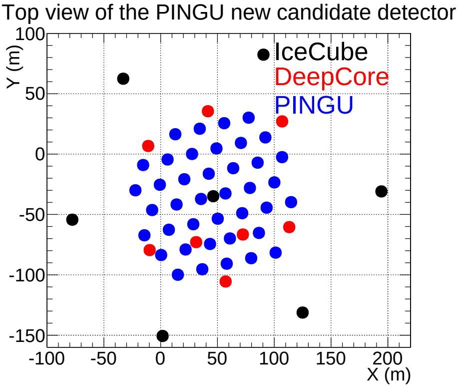


- Fitting a “cascade+μ” hypothesis
  $\Rightarrow$ reconstructed μ should be 0 GeV

- A ν_e CC final state cascade has an EM component and a Hadronic component
  $\Rightarrow$ reconstructed ν_e energy should be higher than “truth”

Datasets used for studies

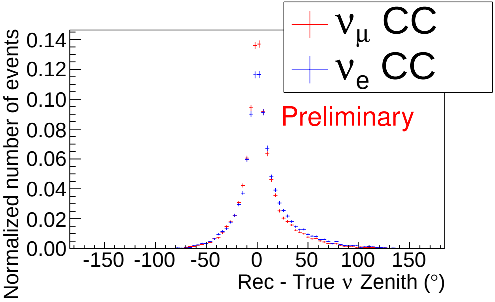
- Using the latest considered geometry

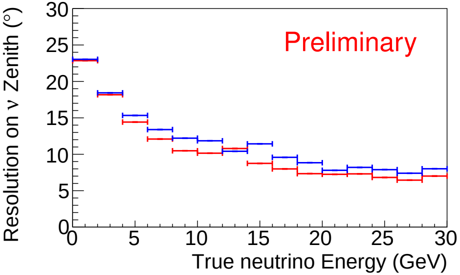


- ν events simulated with the latest IceCube tools
  - E_ν spectra of $E^{-1} \rightarrow$ optimized for parametric analyses
- Using slightly older event selection in respect to presentation by Andreas Gross on “Muon rejection in ice”
  - atmospheric μ rejection should not strongly affect ν resolutions

Resolutions @PINGU: ν Zenith

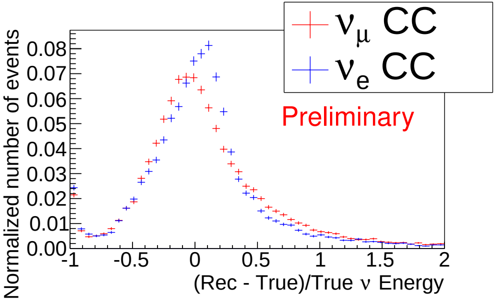
- ν_e and ν_μ resolutions very similar
  - timing information is very useful to pin direction
- With improved tables, ν_μ CC reconstruct better than ν_e CC
- Resolutions of $\sim 15^\circ$ at 5 GeV and below 10° above 16 GeV

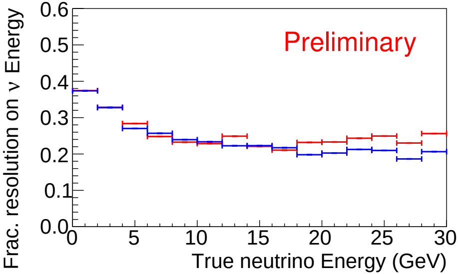




Resolutions @PINGU: E_ν

- As expected, ν_e CC energy reconstructed higher than truth
- ν_e and ν_μ resolutions very similar
- Peak at -1 related to noise-only events
  - should be removed by the new version of the rejection cuts
- Resolutions of $\sim$ 25% above 6 GeV



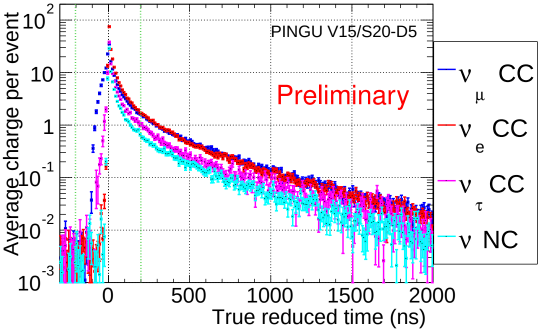


Particle IDentification

- The main goal is to separate cascades (ν_e CC) events from tracks (ν_μ CC) events
- At lower energies no separation by single variable
  $\rightarrow$ use multivariate analysis (TMVA) with 6 variables
- HybridReco/MultiNest provides some variables
  1. Reconstructed track (μ) length
  2. Reconstructed $\frac{E_\mu}{E_\nu} = 1 - Y$ (Y is Bjorken-y parameter)
  3. LLH difference between best fit and cascade only hypothesis
- Other variables by looking at hit timing (see next slide)

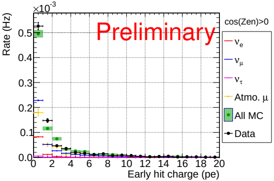
Variables used for PID – “reduced time” variables

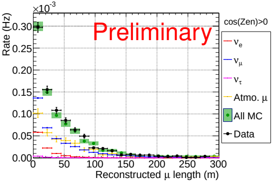
$$t_{hit}^{reduced} := t_{hit} - t_{iter} - ||\vec{x}_{hit} - \vec{x}_{iter}||/c_{ice}$$

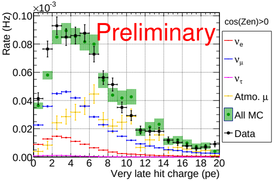


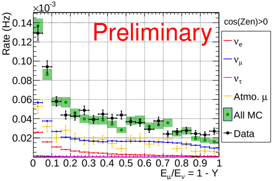
- Charge in reduced time slices $[-200, -6]$ns (early), $[200, 20000]$ns (very late)
- Time required to have more integrated charge than 10% of the total charge

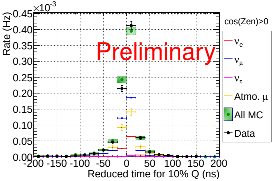
PID variables agreement – looking at DeepCore data

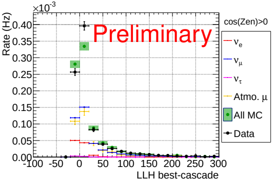


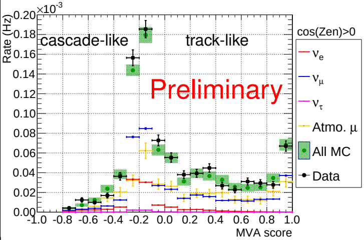








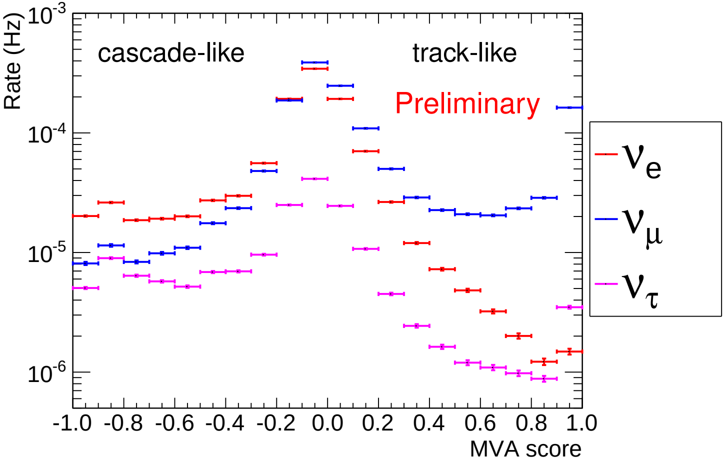




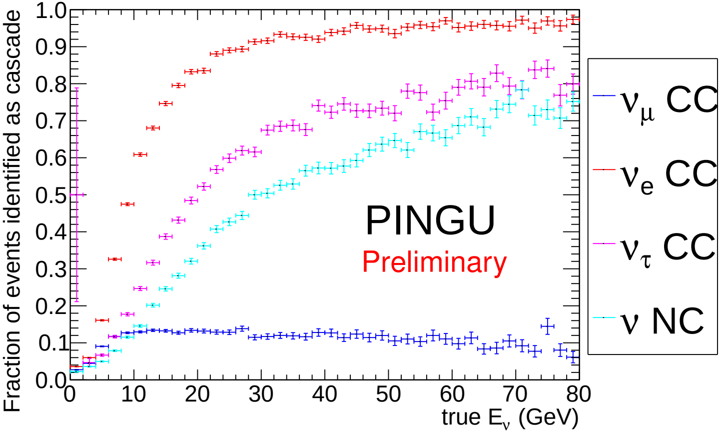
- Good agreement between data and MC for input variables and MVA score

Results from MVA training

- Still need to train and check PID for new geometry
  $\Rightarrow$ using geometry from LoI on following plot!



PID efficiency energy dependence



- At low-energy low distinction power between ν_μ CC and ν_e CC
- At higher energy “easier” PID: there is a clear μ track

Summary

- Currently using HybridReco/MultiNest to reconstruct PINGU events
  - Reconstruction does not require specific event selection
    - ★ increased sample size
    - ★ evaluating quality criteria to further improve resolutions
  - Zenith resolution better of 15° at 5 GeV and below 10° above 16 GeV
  - Resolution of 25% above 6 GeV

- Particle IDentification performing well
  - Currently using MVA, verified good agreement between data/MC of inputs and MVA distribution
  - 2-3 main samples: “cascade-like”, “track-like” (and “unknowns”)
  - At high energy high efficiency of particle identification
  - However it doesn’t really work below 5-10 GeV