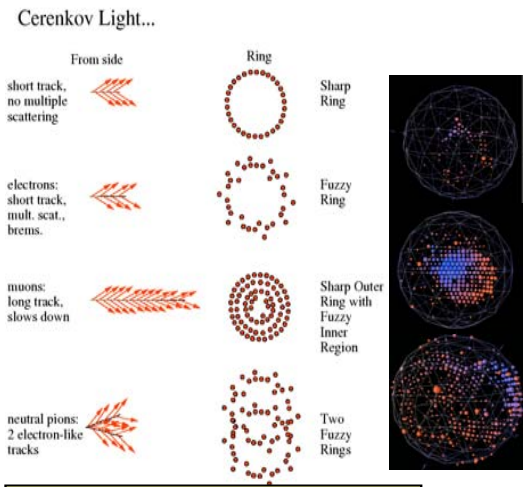
MiniBooNE Event Reconstruction and Particle Identification

Hai-Jun Yang (for MiniBooNE Collaboration), University of Michigan, Ann Arbor

APS April Meeting, Jacksonville, Florida, April 14-17, 2007

Goal: The MiniBooNE experiment is designed to confirm or refute $\nu_\mu \rightarrow \nu_e$ oscillation signal seen by LSND



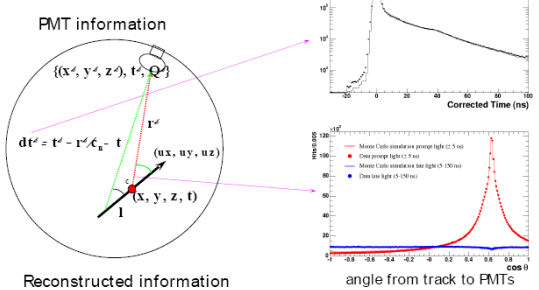
Event Reconstruction

An event consists of a set of charge, time, and spatial information for each PMT.

$\{(x^k, y^k, z^k), t^k, Q^k\}$; k = 1, 2, …, N PMT hits

Many primary and reconstructed variables:

- coordinate distribution
- track length
- time distribution
- event topology
- energy

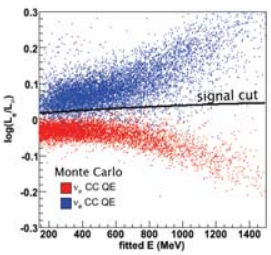


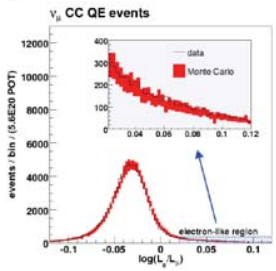
Track Max-Log-Likelihood

- Fit events with 3 hypotheses: e,μ,π^0
- Ratios of max-log-likelihoods provide particle identification information
- Staged background rejection

μ or e?

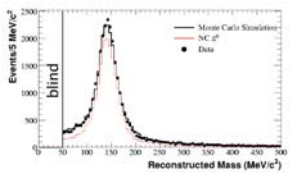
- positive (negative) $\log(L_e/L_\mu)$ favors the e (μ) hypothesis
- easier to distinguish at higher energies
- >90% efficient for signal
- check simulation with tagged μ sample

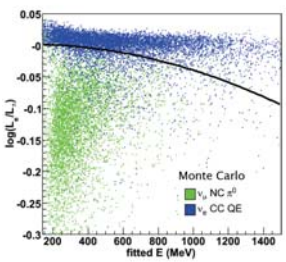


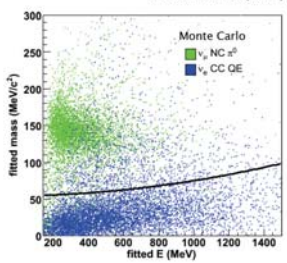


π^0 or e?

- two attacks: $\log(L_e/L_\pi)$ and $M_{\gamma\gamma}$
- $M_{\gamma\gamma}$ from mass-unconstrained fit
- π^0s have $M_{\gamma\gamma} \sim M_{\pi^0}$





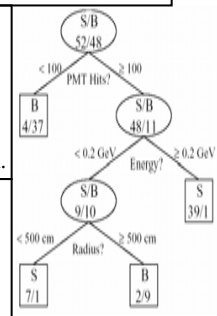


Boosted Decision Trees (BDT)

- Reconstruct O(300) physical/topological variables
- Combine/optimize 172 good variables using BDT to build a powerful PID variable.
- Global background rejection, e or non-e ?

➔ How to build a decision tree ?

For each node, try to find the best variable and splitting point which gives the best separation based on Gini index.
Gini_node = Weight_total*P*(1-P), P is weighted purity
Criterion = Gini_father – Gini_left_son – Gini_right_son
Variable is selected as splitter by maximizing the criterion.

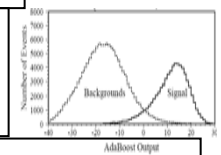


➔ How to boost the decision trees ?

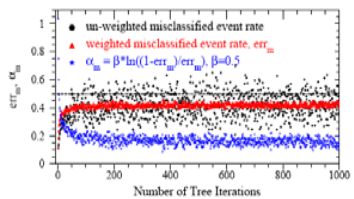
Weights of misclassified events in current tree are increased, the next tree is built using the same events but with new weights. Typically, one may build few hundred to thousand trees.

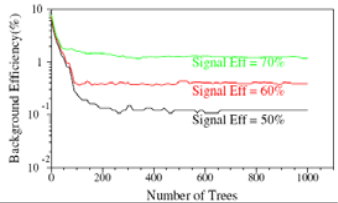
➔ How to calculate the event score ?

For a given event, if it lands on the signal leaf in one tree, it is given a score of 1, otherwise, -1. The sum of scores from all trees is the final score of the event.



➔The advantage of using boosted decision trees is that it combines many decision trees, "weak" classifiers, to make a powerful classifier. The performance of boosted decision trees is stable after a few hundred tree iterations.





1. B.P. Roe, H.J. Yang et.al., physics/0408124, NIM A543 (2005) 577-584.
2. H.J. Yang, B.P. Roe, J. Zhu, physics/0508045, NIM A555(2005) 370-385.
3. H.J. Yang, B.P. Roe, J. Zhu, physics/0610276, Accepted by NIMA (2007).
4. Y. Liu, I. Stancu, physics/0611267