A look inside the particle identification of MiniBooNE

Alexis Aguilar-Arevalo and Kendall Mahn, Columbia University
Teppei Katori, Indiana University
Serge Ouedraogo, Lousiana State University
for the MiniBooNE Collaboration

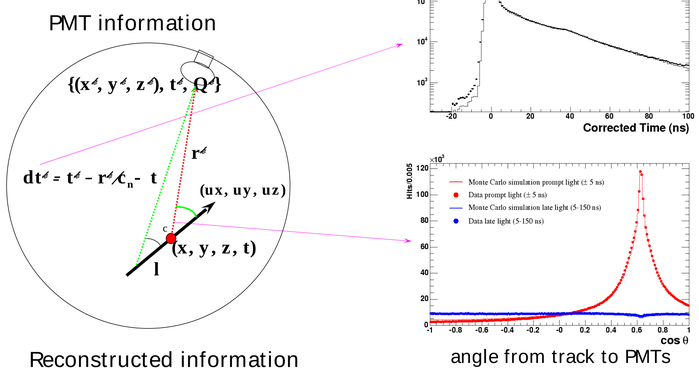
Event Reconstruction

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



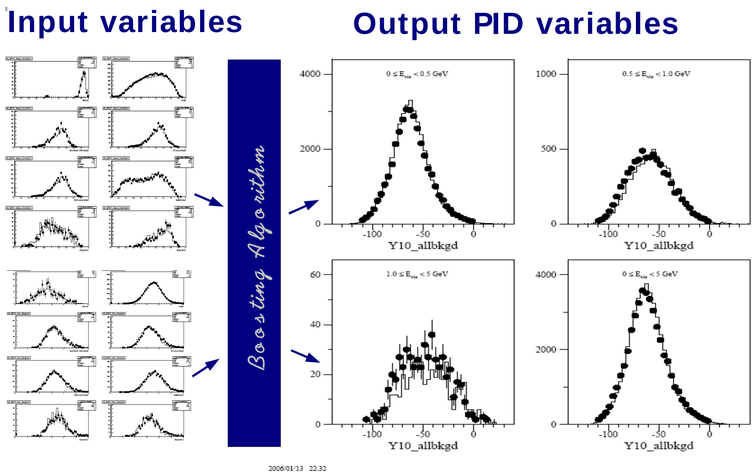
Particle Identification (PID)

Designed to classify signal and background.

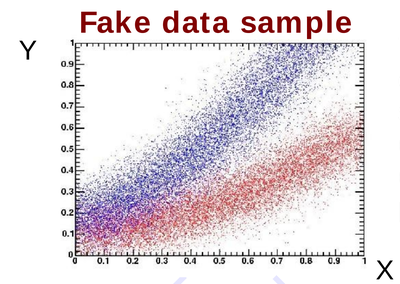
Signal = oscillation ν_e CCQE events

Background = everything else (intrinsic ν_e, misID)

All variables are required to have good agreement between Data and MC simulation within statistical and systematic errors.

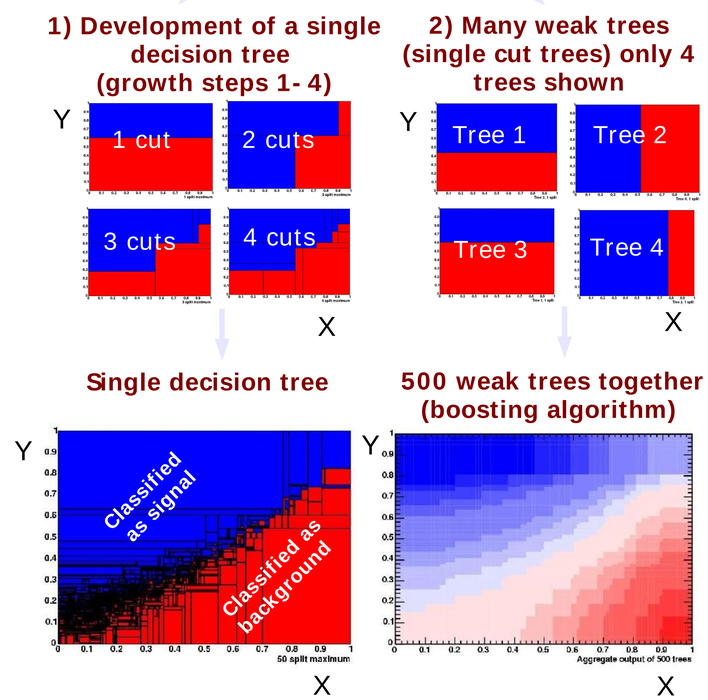


Simple Example



The goal of the classifier is to separate blue (signal) and red (background) populations.

Two ways to use decision trees. 1) Multiple cuts on X and Y in a big tree, 2) Many weak trees (single-cut trees) combined



Degrading of classifier performance by demanding perfect training separation is called "overtraining"

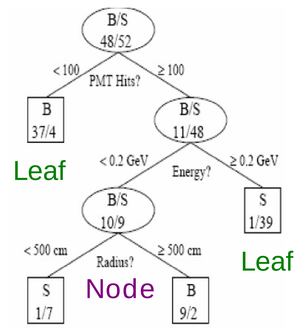
Boosting Algorithm has all the advantages of single decision trees, and less susceptibility to overtraining.

MiniBooNE collaboration gratefully acknowledges support from NSF, DOE, and the help of Fermilab

National Science Foundation · ENERGY · Fermilab

Decision Trees

Single Decision Tree

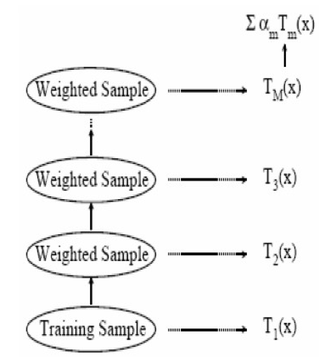


- Go through all PID variables and find best variable to separate signal from background and cut on it.
- For each of the two subsets repeat the process.
- This forking decision pattern is called a tree.
- Split points are called nodes.
- Ending nodes are called leaves.

Boosted Decision Trees

- Given a training sample, boosting increases the weights of misclassified events (background witch is classified as signal, or vice versa), such that they have a higher chance of being correctly classified in subsequent trees.
- Trees with more misclassified events are also weighted, having a lower weight than trees with fewer misclassified events.
- Build many trees (~1000) and do a weighted sum of event scores from all trees (score is 1 if signal leaf, -1 if background leaf)

Boosted Decision Trees



Nucl. Inst. & Meth. A, 543(2005)577
Nucl. Inst. & Meth. A, 555(2005)370

Summary and Outlook

- MIniBooNE is looking for ν_μ to ν_e oscillations with a probability of ~0.25%
- A powerful particle identification is essential to achieve this goal. Boosted decision trees provide a way to effectively combine all the information available in each event.
- NuMI and high energy MiniBooNE events provide useful checks of this particle identification on electron neutrino events.
- On track for a result as soon as this summer.

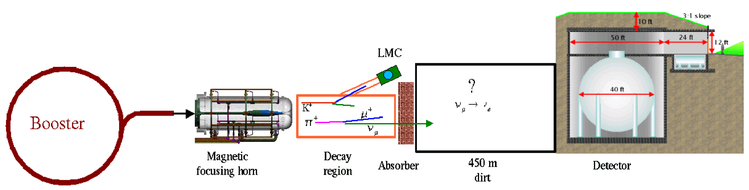
Introduction

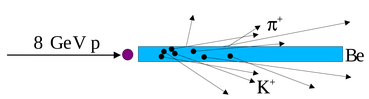
The MiniBooNE experiment is designed to address the unconfirmed oscillation signal seen by the LSND experiment at Los Alamos (a 0.25% oscillation probability with a 3.3σ statistical significance). Searching for the appearance of such a small fraction of electron neutrinos in a beam of muon neutrinos is a difficult task. In this poster we give a few details about the reconstruction and particle identification methods used in the experiment.

Important steps since last review

August 2005: Publication of HARP thin target analysis.
January 2006: Switch horn polarity to begin antineutrino running.
March 2006: Booster delivers 7.2 x 10^{20} POT to MiniBooNE. ~5000 events are observed in the antineutrino mode configuration.
Nov05- May 2006: Made substantial progress in refining Optical Model through *in situ* calibrations.

Neutrino source



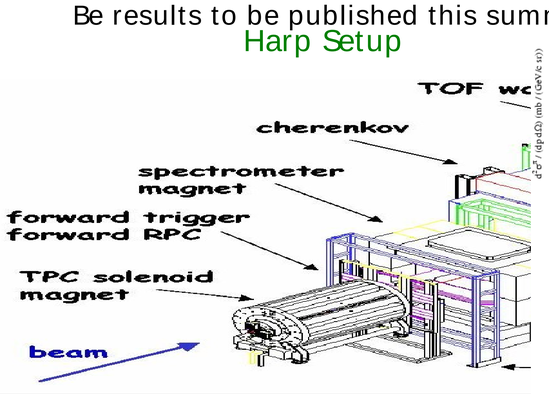


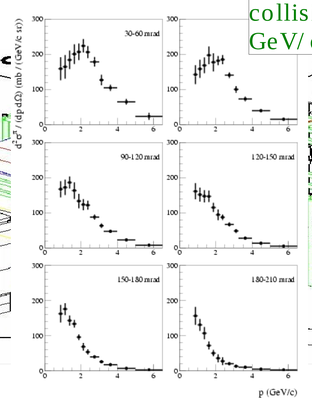
$\pi^+ \to \mu^+ + \nu_\mu$	source of ν_μ
$\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e$	source of ν_e background
$K^+ \to \mu^+ + \nu_\mu$	source of ν_μ
$K^+ \to e^+ + \pi^0 + \nu_e$	source of ν_e background
$K^0_L \to e^+ + \pi^- + \nu_e$	source of ν_e background

- Knowing the production cross sections for mesons produced at the target is critical for determing the neutrino flux.
- In summer of 2002, the cross sections for an 8 GeV proton beam on Be were measured by the HARP experiment at CERN.

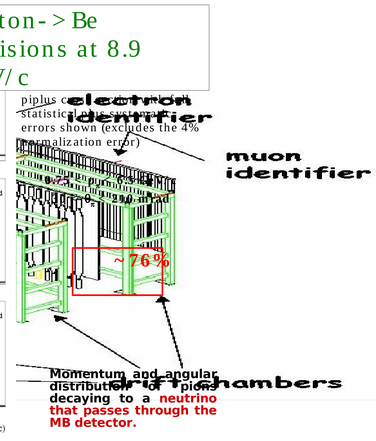
HARP has published its first results on Al **(Nucl.Phys.A752:24-33,2005).**
Be results to be published this summer.

Harp Setup

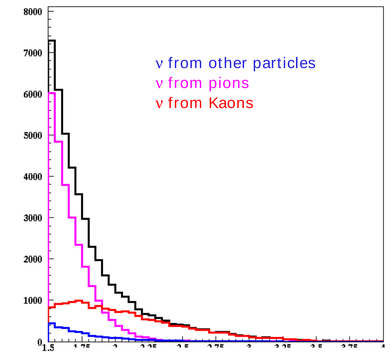




proton->Be collisions at 8.9 GeV/c



High energy data sample



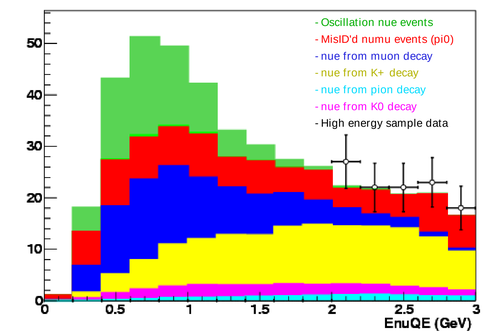
The highest energy events are primarily from kaon decays.

Two samples at high energy provide normalization and shape information for kaon backgrounds to oscillation analysis:

- ν_e events passing oscillation event selection cuts
- ν_μ induced CCQE, CCπ+events

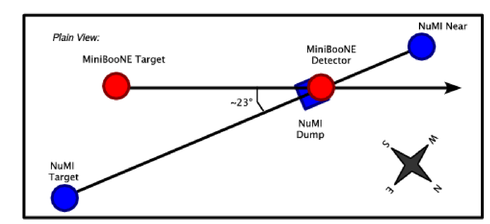
Boosting-based PID selects electron-like neutrino events

- Oscillation events
- Intrinsic beam ν_e s
- Mis-identified ν_μ interactions

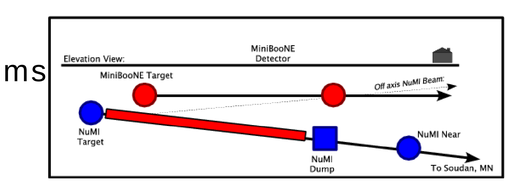


Events from the NuMI Beamline

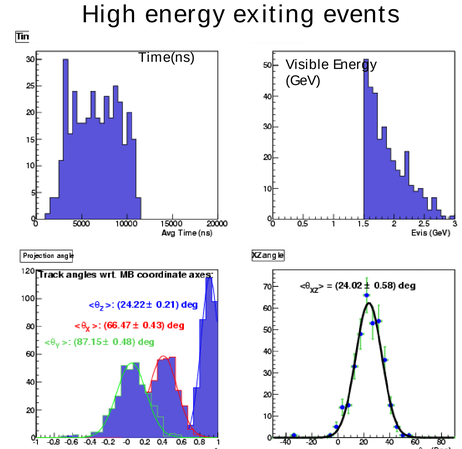
- NuMI beam events provide an useful calibration sample
- Collected over 70K candidates before shutdown



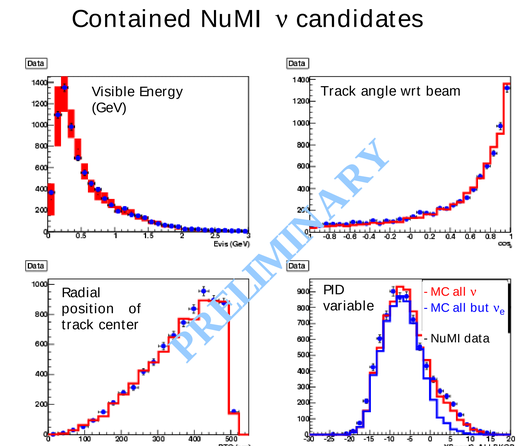
- World's first off-axis neutrino beam
- Important tests of PID algorithms on ν_e of relevant energies



High energy exiting events



Direction cosines of reconstructed tracks wrt to MiniBooNE coordiates axes

Track angle in x-z plane showing correlation of x,z dir

Contained NuMI ν candidates