MicroBooNE Analysis Tools High Level Summary

Eric and Herb
Collaboration Meeting
29,30-May-2014

Outline

- Perfunctory reco/sim summary

- MCC5

- Administrivia, Tools, Grids, Frameworks, Data

- Leonidas's Neutrino2014 poster plot: rubberstamping

Simulation and Reconstruction

□ Please see the summaries from Wednesday's wonderful presentations, which were themselves summaries: docdb 3477, 3480.

- There is a formidable set of improvements to the simulation: wire noise, wire signal responses, ROIs, PMT simulation, trigger simulation, timing service.

- There is a formidable set of reconstruction improvements: tracking efficiency, MCS, cluster– merging and shower creation, shower/track differences.

   ○ There is a new, positive trend toward using 2 planes to determine/check activity on the third and to aid the 3D deco.

MCC

- Monte Carlo Challenge 5
  - mid-July

  - **start practicing runs now**
    - new geometry
      - implies new photon look-up library needed, meaning that grid run must start now.
    - The CRY+Genie 1-big-window jobs push up to and past the 4 GB limit
      - requesting LArSoft valgrind/profiling help
        - We've secured 2+ wks of one person's effort. We're Meeting Monday, 3-June with LArSoft team and ART leads to detail the problem and the needed solution.

Monte Carlo Production

- We welcome Kazu as new MC Production co-convener (with Yun-Tse), and thank soon-to-be-outgoing former co-convener Ben Carls for all his great service.
  - Yay, Ben! It has not been easy at times.
  - Kazu will begin to institute a Double Chooz-like MC Production system (ala the data production system I showed in the DAQ talk yesterday), with project.py at its core and a status-keeping database.
  - Agrees to more pre-MCC testing, with which we all agree is needed.

Grid jobs

- Grid Tools
  - FermiGrid
    - Our 500 slots will get us through MCC5
    - Those slots have been sufficient for the Analysis Tools team, generally.

  - OSG (for MCC 5+epsilon)
    - Yun-Tse has successfully run jobs on OSG
      - ~30,000 cores at IU, Nebraska, SLAC…
      - project.py (our job submitter) uses a new jobsub to do this.
      - However, there are 2-3 GByte upper limits on OSG, so we have to get our memory footprint under control before we can use this resource
      - Fermicloud (10? cores) will allow 8 GB jobs.

Frameworks 1

- LArSoft
  - Project is pushing for unit tests and continuous integration (CI)
    - CI == builds triggered by pushes
    - Ruth is asking for summer help on unit tests
  - Considering builds of more than the develop branch
    - ubdev ?
  - Librarians to be instituted
  - We need to be careful about duties required of our appointed librarians
  - Documentation needs to be stepped up
    - A big push from the LArSoft project for this
      - Wants a (how-it-works) .tex file in git that is built to a pdf continuously
    - As well as from MicroBooNE Spokes
      - We need a detailed and descriptive tech note per algo/module
      - especially in this cross-experiment software, to take proper credit
  - We will soon resurrect the uboonecode nightly build

Frameworks 2

- LArLight

  - Lightweight, largely analysis modules

  - The favored environment of most of the Shower researchers, at least

  - We see this as a powerful alternative to LArSoft in many situations, and support its continued use

    - …As long as Kazu is happy to maintain it and keep data product compatibility with the LArSoft recob::dataProducts, which he has proclaimed he is.

dCache

- Fast networked non-permanent disk storage used by SAM, and allegedly available for our everyday lat usage too.

    - 4 PB shared by everybody. No quotas.

    - /pnfs/uboone/scratch/users

        - Can be written to fine.
        - Can not, it turns out, be read back from, except with 'ifdh cp' which our project.py is not yet using. (known — *if unanticipated* — problem for all of IF right now)
        - A bare ROOT session can not read files from there either, even using root://pnfs/uboone/scratch/users/echurch/file.root for the file path (this is xrootd), though we're told this should in fact work.

SAM

- SAM overview.

  - SAM database stores metadata about microboone files.
    - Writing.
  - SAM station manages data analysis projects (schedules files for delivery to worker nodes).
    - Reading.

  - Integrates with other elements of Fermilab data handling system.
    - Data tape storage (enstore).
    - File Transfer Service (FTS) for uploading data into enstore.
    - File Delivery (ifdh).

- SAM status.

  - Integrated with microboone production tools (project.py).
  - MCC4 files are stored in sam.

Workshops

- FIFE workshop.
  - June 16-17(am).
  - We will let you know if we think you should go; let us know if you would like to go
  - Will cover common infrastructure and tools topics among IF experiments
    - https://indico.fnal.gov/conferenceDisplay.py?confId=8406

- Larsoft workshop.
  - June 17(pm)-18.
  - Focus on automatic testing and validation procedures ("continuous integration"), as well as algorithm development.
    - unsexy topics, but useful for unit-test creation for assurance that our code is unperturbed by other check-ins
  - All LArSoft "authors" encouraged to go
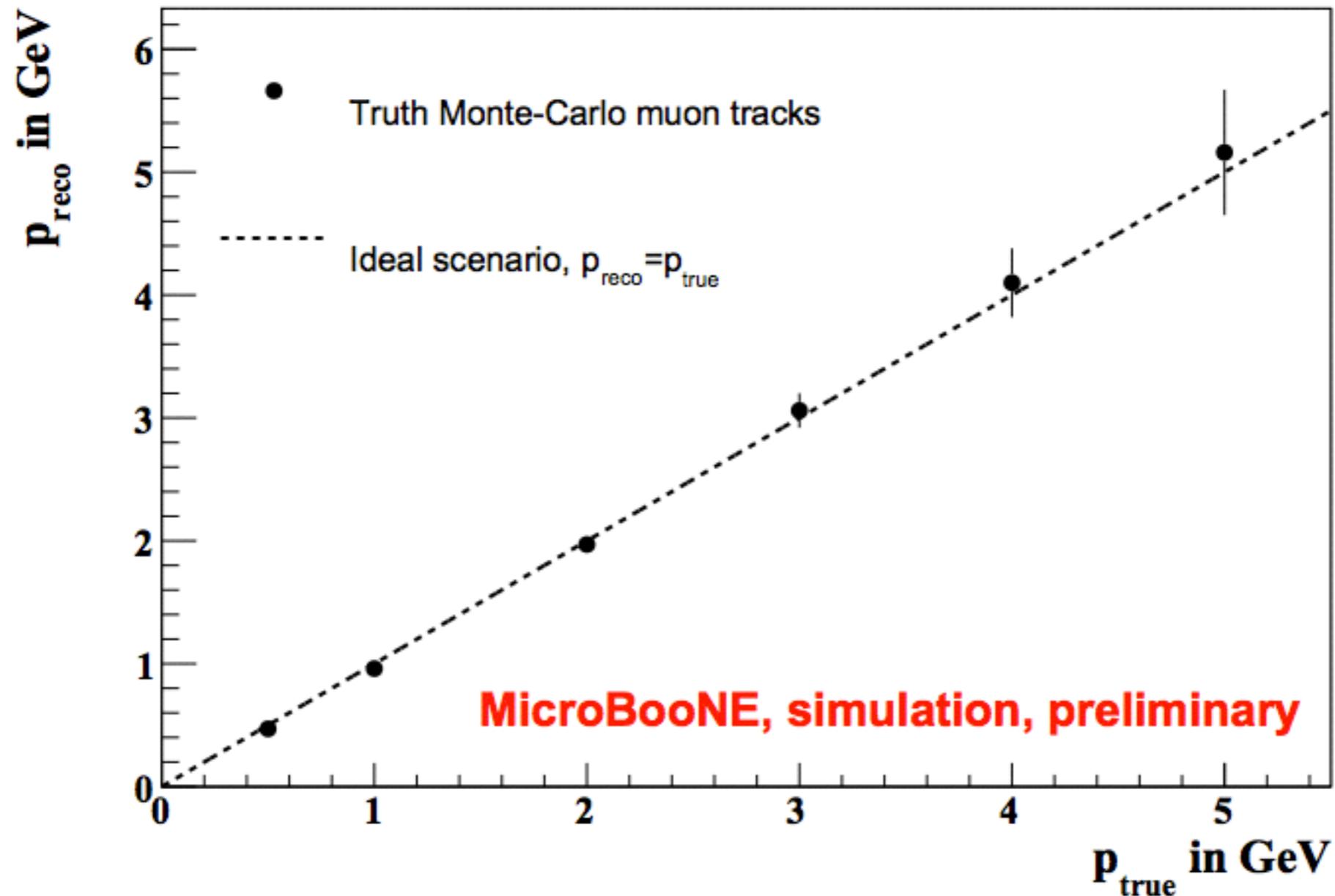 - No indico agenda yet, but there is a doodle poll.
    - http://doodle.com/t7uuey9pk8zp6kaq

Regular AT meetings

○ Usual 9:00, Thursdays in Libra (9SW)
○ Wednesdays lunch non-meeting in WH10
○ Tuesday Simulation Mtgs, 12XO

Leonidas's Neutrino2014 poster plot

Linearity plot demonstrating, the performance of the likelihood algorithm to measure particle momenta through multiple scattering. The plot shows the true momenta (x axis) versus the reconstructed one (y axis) and it was prepared with Monte-Carlo muons generated upstream of the TPC. Truth MC information of particle tracks and energies was used in the analysis. That means, strictly speaking, that the plot serves as the "proof of principle" of the likelihood approach to MCS.