Appendix A LBNE Detector Simulation and Reconstruction

A 10-kt or larger LArTPC far detector fulfills the high-mass requirement for LBNE and provides excellent particle identification with high signal-selection efficiency ($\geq$ 80%) over a wide range of energies. The far detector is described in detail in the LBNE Conceptual Design Report Volume 1 [1] and briefly in Section 3.6 of this document. This appendix summarizes the status of the LBNE LArTPC simulation and reconstruction efforts and their expected performance.

A.1 Far Detector Simulation

A.1.1 Tools and Methods

In the full simulation of the far detector, neutrino interactions are simulated with Geant4 [2] using the LArSoft [3] package. LArSoft is being developed to provide an integrated, experiment-agnostic set of software tools to perform simulation, data reconstruction and analysis for LArTPC neutrino experiments. Individual experiments provide experiment-specific components including a detector geometry description and analysis code, and they contribute to the LArSoft software development itself.

LArSoft is based on *art* [4], an event-processing framework developed and supported by the Fermilab Scientific Computing Division. *Art* is designed to be shared by multiple experiments and is currently used by several intensity frontier experiments, including NOνA, Mu2e, MicroBooNE [5] and ArgoNeuT [6]. The last two have liquid argon TPC-based detectors and thus share many simulation and reconstruction requirements with LBNE. Reconstruction algorithms developed in LArSoft for the ArgoNeuT and MicroBooNE experiments can readily benefit LBNE. Examples of neutrino beam interactions in a LArTPC obtained from the LArSoft package using the MicroBooNE detector geometry are shown in Figure A.1.

The LBNE far detector geometries currently available in LArSoft are the LBNE 10-kt surface detector and the 34-kt underground detector. Also included is geometry for a 35-t prototype that LBNE has constructed at Fermilab*. The LBNE far detector geometry description is generated in a flexible way that allows the simulation of various detector design parameters such as the wire spacing and angles, drift distances, and materials. The photon-detector models are based on the design that uses acrylic bars coated with wavelength-shifting tetraphenyl butadiene (TPB), read out with silicon photomultiplier tubes (SiPMs).

Geant4 is used to simulate particles traveling through the active and inactive detector volumes

*One of the goals of the 35-t prototype is to test key elements of the TPC module design for the 10-kt and 34-kt detectors including the wrapped wire planes and drift distances.

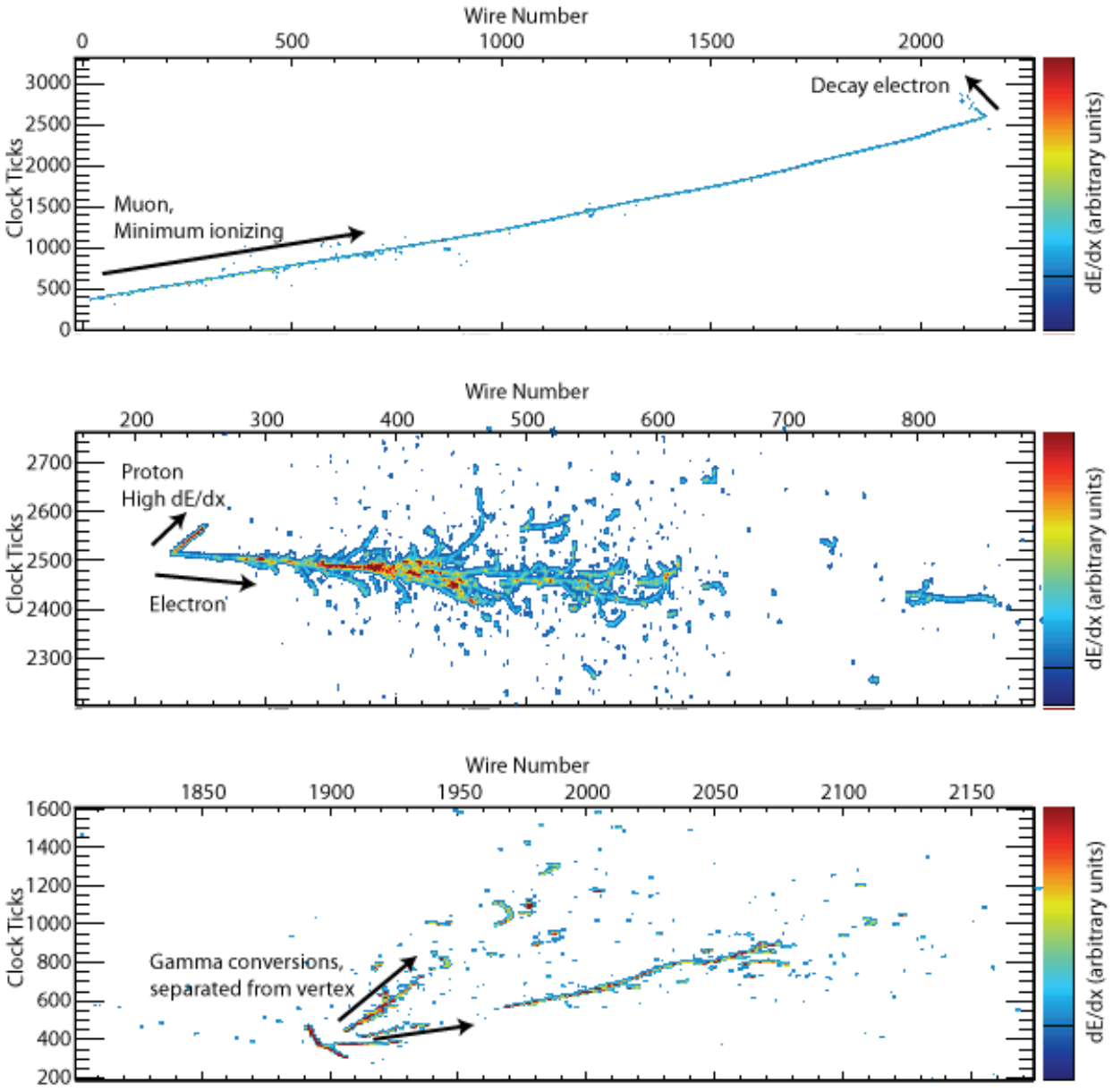

Figure A.1: Examples of neutrino beam interactions in a LArTPC obtained from a Geant4 simulation [3]. A ν_μ-CC interaction with a stopped μ followed by a decay Michel electron (top), a ν_e-CCQE interaction with a single electron and a proton (middle), and an NC interaction which produced a π^0 that then decayed into two γ's with separate conversion vertices (bottom).

and the surrounding materials such as the cryostat and rock. The tens of thousands of photons and electrons produced (by the ionization of the argon) per MeV deposited are simulated using a parameterization rather than a full Geant4 Monte Carlo, as tracking them individually would be prohibitive. The drifting electrons are modeled as many small clouds of charge that diffuse as they travel toward the collection wires. The response of the channels to the drifting electrons is parameterized as a function of drift time, with a separate response function for collection and induction wires. The signals on the induction-plane wires result from induced currents and are thus bipolar as a function of time as charge drifts past the wires, while the signals on the collection-plane wires are unipolar. The response functions include the expected response of the electronics. Noise is simulated using a spectrum measured in the ArgoNeuT detector. The decays of ^{39}Ar are included, but some work is required to make them more realistic.

For the 10-kt far detector, a 1.5-ms readout of the TPC signals at 2 MHz gives a simulated data volume of just under 2 GB per event. If the readout is extended to include the beam window, then in order to collect charge deposited by cosmic rays (which would otherwise be partially contained), a greater data volume will be required. To reduce the data volume and speed up the calculation, long strings of consecutive ADC counts below a settable threshold are suppressed in the readout. Huffman coding of the remaining data is included in the digitization [7].

The photon-detection system likewise requires a full Monte Carlo simulation. Photons propagating from the TPC to the acrylic bars have been fully simulated using Geant4, and their probabilities of striking each bar (as a function of the emission location and the position along the bar at which the photon strikes) have been computed. Smooth parameterizations of these functions are currently used in the simulation to compute the average number of photons expected to strike a bar (as a function of position along it). Given the current design of the optical detectors, approximately 2-3% of VUV (vacuum ultraviolet) photons produced uniformly in the fiducial detector volume strike the bars. This low number is largely due to the small fraction of the total area in contact with the argon that is represented by the bars, and the low reflectivity of the stainless steel cathode planes, the field cage and the CuBe wires.

A second function is used to parameterize the attenuation of light within the bar as a function of position along the bar. The total response of a SiPM to light produced in the detector is the product of the number of photons produced, the probability of the photons to survive propagation, the interaction with the wavelength shifter (commonly called *downconversion*), the attenuation in the bar, and the detection efficiency of the SiPM. This product is used as the mean of a Poisson distribution from which the number of photoelectrons is randomly drawn to simulate the measurement of the SiPM. Measured waveforms for cold SiPMs are used in simulating the digitized response. Measurements in prototype dewars will be used to normalize the yield for signals in the SiPMs as a function of the incident location of the VUV photon on the bar. The NEST [8] model, which describes the conversion of ionization energy into both electrons and photons in an anticorrelated manner, and which has been shown to model a large range of data from noble liquid detectors, is

currently being incorporated into the LBNE detector simulation.

A variety of event generators are available for use in the simulation. Neutrino hard-scattering interactions and subsequent nuclear breakup are simulated using GENIE [9], though the use of other generators is possible. Cosmic rays are simulated with CRY [10]. Single particles can be generated one at a time, and general text-file interfaces are available allowing arbitrary generators to be used without linking them with LArSoft.

Currently, samples of single electrons, muons, charged and neutral pions, protons and tau leptons have been generated and simulated using the 10-kt surface geometry and the 35-ton geometry, though without photon-detector simulation. These samples are being used to develop reconstruction algorithms.

Planned improvements to the simulation include creating an interface to a calibration database, updating the response functions with measured responses from MicroBooNE, which uses an electronics design very similar to that of LBNE, simulating the effects of space-charge buildup in the drift volume, and creating more detailed maps of the drift in the gaps between the APAs and the charge that is deposited between the wire planes.

A.1.2 Low-Energy Neutrino-Response Studies with LArSoft

Work is currently underway using the LArSoft simulation package to characterize low-energy response for realistic LBNE detector configurations. Figure A.2 shows a sample 20-MeV event in the LBNE 35-t prototype geometry simulated with LArSoft. So far, most studies have been done with the MicroBooNE geometry, with the results expected to be generally applicable to the larger LBNE detector. For a preliminary understanding of achievable energy resolution, isotropic and uniform monoenergetic electrons with energies of 5-50 MeV (which should approximate the ν_e-CC electron products) were simulated and reconstructed with the LArSoft package. The charge of reconstructed hits on the collection plane was used to reconstruct the energy of the primary electrons. (Induction-plane charge as well as track-length-based reconstruction were also considered, but with inferior results). Figure A.3 shows the results. A correction to compensate for loss of electrons during drift, $Q_{collection} = Q_{production} \times e^{-T_{\rm drift}/T_{\rm electron}}$ (where T_{drift} is the drift time of the ionization electrons, and $T_{electron}$ is the electron lifetime), using Monte Carlo truth to evaluate $T_{\rm drift}$, improved resolution significantly. This study indicated that photon time information will be valuable for low-energy event reconstruction. Some of the resolution was determined to be due to imperfect hit-finding by the nominal reconstruction software. A tuned hit-finding algorithm did somewhat better (Figure A.3), and further improvements for reconstruction algorithms optimized for low-energy events are expected.

Also under study is the potential for tagging ν_e-CC absorption events ($\nu_e + {}^{40}\mathrm{Ar} \rightarrow e^- + {}^{40}\mathrm{K}^*$) using the cascade of de-excitation γ rays, which should serve the dual purposes of rejecting background and isolating the CC component of the signal.

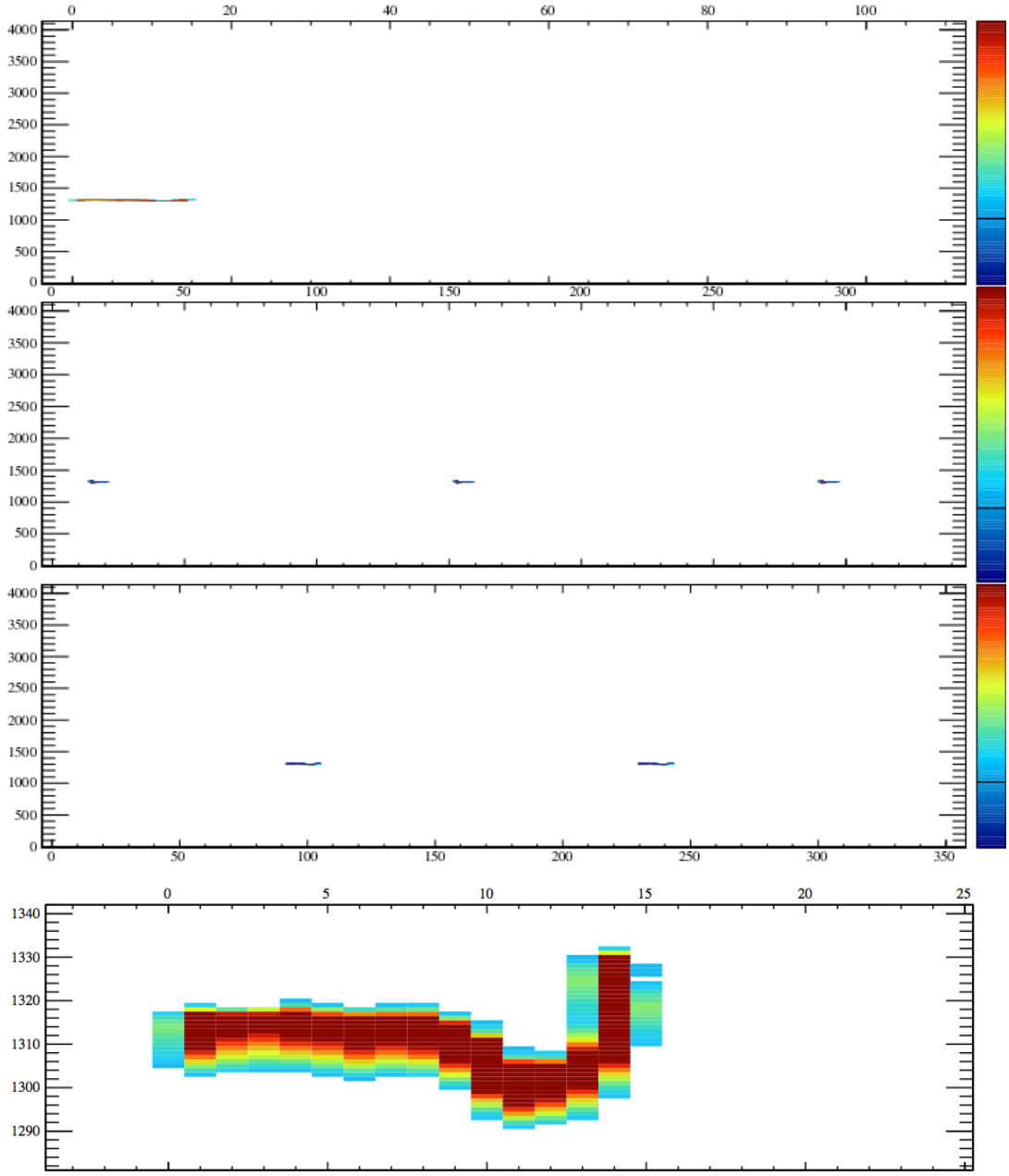

Figure A.2: Raw event display of a simulated 20-MeV event in the LBNE 35-t prototype; the top panel shows the collection plane, and the lower two panels show the induction planes (with multiple images due to wire wrapping). The bottom panel shows a zoom of the collection plane image.

A.2 Far Detector Reconstruction

The first stage of reconstruction of TPC data is unpacking and deconvoluting the electronics and field response of the wire planes. The deconvolution function includes a noise filter that currently is parameterized with ArgoNeuT's noise, but will be tuned for the eventual noise observed in the

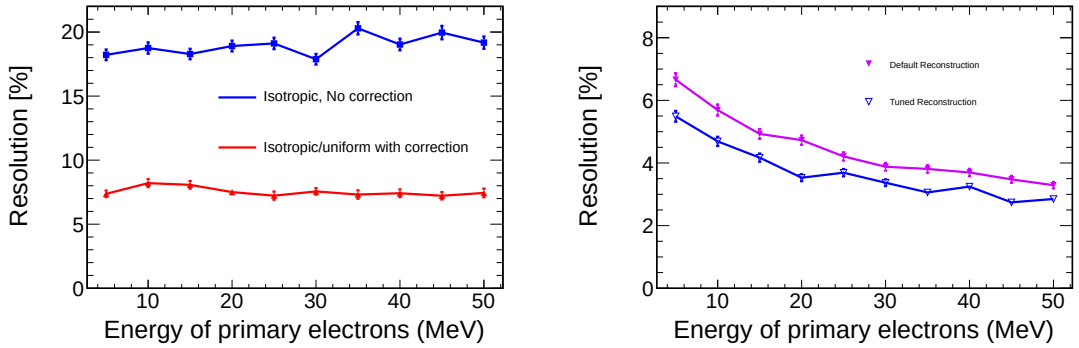

Figure A.3: Left: Comparison of energy resolution (defined as σ/E, where σ is the spread of the collection-plane-charge-based event energy E for a monoenergetic electron), with and without electron-lifetime correction, as a function of electron energy. The blue curve is the energy resolution of isotropic and uniform electrons without electron-lifetime correction. The red curve is the energy resolution with electron-lifetime correction based on MC truth. Right: Comparison of energy resolution before and after tuning the reconstruction algorithm (for fixed position/direction electron events).

LBNE detector. The deconvolution makes sharp, unipolar pulses from the bipolar induction-plane signals and also sharpens the response to collection-plane signals. Hits are then identified in the deconvoluted signals by fitting Gaussian functions, allowing for sums of several overlapping hits in each cluster. In LBNE, because of the large quantity of channels in the far detector, any inefficiency in CPU and memory is magnified. Improvements in the memory-usage efficiency relative to the ArgoNeuT and MicroBooNE implementations have been realized by rearrangement of the processing order and limiting the storage of the intermediate uncompressed raw data and the deconvoluted waveforms.

After signal deconvolution, line-finding and clustering based on a Hough transform in two dimensions is done using an algorithm called *fuzzy clustering* [11]. This clustering is performed separately on data from each induction plane. Since the hit data on LArTPCs are inherently 2D — wire number and arrival time of the charge — the location of the initial ionization point has a 2D ambiguity if the deposition time is unknown. For beam events, the t_0 is known, and thus only a 1D ambiguity remains; this 1D ambiguity is broken by angling the induction-plane wires relative to the collection-plane wires, in order to measure the y location of the hits for which t (thus x) and z are known. For (non-beam) cosmic-ray signals which arrive uniformly in time, the photon system provides t_0. After clustering, 3D track-fitting is performed using a Kalman filter [12]. Dedicated algorithms have been developed to optimize electromagnetic shower reconstruction and energy resolution.

LBNE poses a unique challenge for reconstruction because the induction-plane wires wrap around the edges of the APA frames. This introduces discrete ambiguities that are not present in other LArTPC designs. Whereas a hit on a collection-plane wire identifies uniquely the side of the APA from which it came, this is not known for a hit on an induction-plane wire. The angles between

the U and V plane wires are slightly different from 45° and from each other in order to break the ambiguities. A combinatoric issue arises, however, if many hits arrive on different wires at nearly the same time, for instance when a track, or even a track segment, propagates in a plane parallel to the wire planes (i.e., at constant drift distance). Showers will also contain many hits on different wires that arrive at similar times. Hits that arrive at different times can be clustered separately in the Z, U, and V views without ambiguity, while hits that arrive at similar times must be associated using a topological pattern-recognition technique. LBNE is developing a version of the fuzzy clustering tool for use as a pattern-recognition step to allow association of Z, U and V hits, a step that is needed to assign the correct y position to a track segment or portion of a cluster. This process is called *disambiguation* of the induction hits. Misassignment can affect particle-ID performance and reconstructed-energy resolution because fully contained tracks may appear partially contained and vice versa. After disambiguation has been performed, standard track, vertex and cluster reconstruction algorithms are applied.

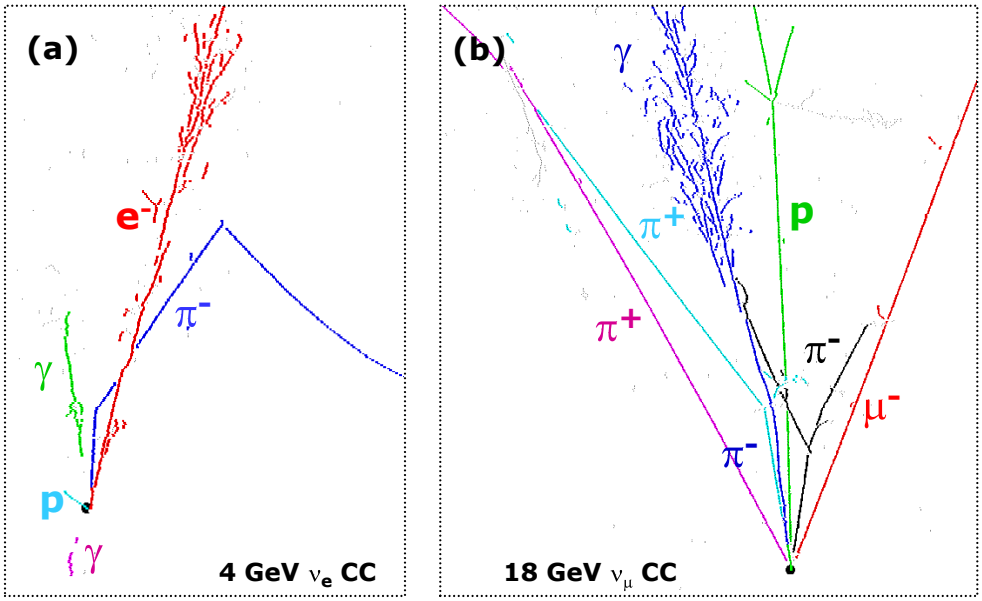

Figure A.4: PANDORA's 2D clusterings of hits created by the particles in two CC neutrino interactions in liquid argon. Panel (a) shows a 4-GeV ν_e interaction, and panel (b) shows an 18-GeV ν_μ interaction. The colors indicate the clusters into which PANDORA has divided the hits, and the particle labels are from the MC truth.

A promising suite of algorithms for event reconstruction is provided by the PANDORA toolkit [13], which provides a framework for reconstruction algorithms and visualization tools. Currently it is being used to develop pattern-recognition algorithms and to reconstruct primary vertices. PANDORA's pattern-recognition algorithm merges hits based on proximity and pointing to form 2D clusters. Vertices are then identified from the clusters that best connect to the same event. Clusters that best correspond to particles emitted from the primary vertex are identified in 2D. These

particle candidates are then used to seed 3D reconstructed particles, and a 3D primary vertex is identified. Examples of PANDORA's 2D clustering are shown in Figure A.4 for two simulated CC neutrino-scattering events. Figure A.5 shows the primary vertex spatial resolution in 3D with well-contained simulated beam-neutrino events, using the nominal LBNE spectrum and MicroBooNE geometry.

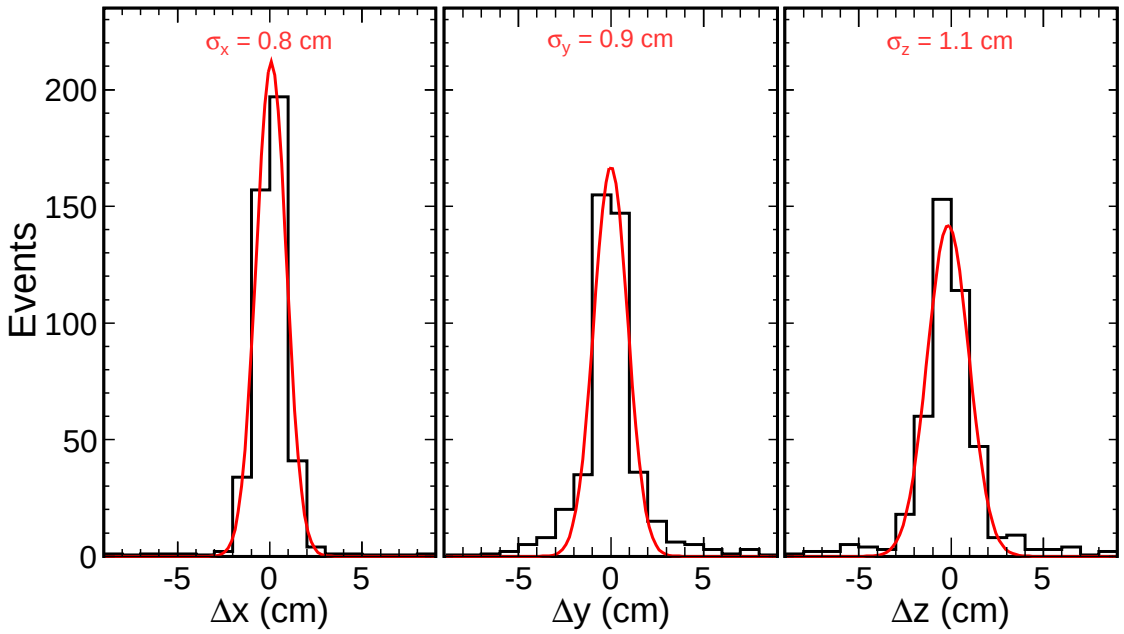

Figure A.5: Distributions of the residuals between the reconstructed and the Monte Carlo true locations of primary vertices in neutrino interactions in the MicroBooNE geometry using the LBNE beam spectrum. The x axis is oriented along the drift field, the y axis is parallel to the collection-plane wires, and the z axis points along the beam direction.

A.3 Fast Monte Carlo

The LBNE full Monte Carlo (MC) simulation will use a Geant4 simulation of the beamline to estimate the neutrino flux, a neutrino interaction generator (e.g., GENIE), and detailed detector simulation that mimics the real detector output for data events. Both data and MC will have the same reconstruction algorithms applied to produce quantities that will be used to analyze the data. The full MC detector simulation and reconstruction algorithms are still under development. Due to their detailed nature, these algorithms are CPU-intensive and time-consuming to run.

In parallel, a Fast Monte Carlo simulation has been developed and is available for use in place of the full MC to explore long-baseline physics analysis topics. A preliminary version of the Fast MC is currently available. Results from the latest detector simulations and advancements in reconstruction algorithms are actively being incorporated to improve the physics models and detector

parameterization. Because the Fast MC replaces CPU-intensive portions of the full MC simulation with a fast parameterized model, it offers a quick, dynamic alternative which is useful for trying out new ideas before implementing them in the full MC. This usefulness is expected to remain even after the full MC simulation is mature.

To accurately approximate a full MC simulation, the Fast MC combines the Geant4 LBNE beamline flux predictions, the GENIE event interaction generator, and a parameterized detector response that is used to simulate the measured (reconstructed) energy and momentum of each final-state particle. The simulated energy deposition of the particles in each interaction is then used to calculate reconstructed kinematic quantities (e.g., the neutrino energy), and classify the type of neutrino interaction, including backgrounds and misidentified interactions.

The Fast MC is designed primarily to perform detailed sensitivity studies that allow for the propagation of realistic systematic uncertainties. It incorporates effects due to choices of models and their uncertainties and design decisions and tolerances. The neutrino flux predictions, the neutrino-interaction cross-section models, and the uncertainties related to these are also incorporated. The parameterized detector response is informed by Geant4 simulations of particle trajectories in liquid argon, by studies of detector response simulation in MicroBooNE [5], results reported by the ICARUS Collaboration, and by the expected LBNE detector geometry. The realistic parameterization of reconstructed energy and angle resolution, missing energy, and detector and particle identification acceptances provide a simulation that respects the physics and kinematics of the interaction and allows for propagation of model changes to final-state reconstructed quantities.

Future efforts will allow for propagation of uncertainties in detector effects and of detector design choices. It should be noted that the same GENIE files generated for the Fast MC can be used as inputs for the full detector simulation and the results of the two simulations can be compared both on an event-by-event basis and in aggregate. Studies of this nature can be used to tune the Fast MC and to cross-check the full simulations.

In the current configuration of the Fast MC, GENIE generates interactions on ^{40}Ar nuclei with neutrinos selected from the energy spectra predicted by the Collaboration's Geant4 flux simulations (described in Section 3.4). For each interaction simulated in GENIE, a record of the interaction process, its initial kinematics, and the identity and four-momenta of the final-state particles is produced. The parameterized detector response applies spatial and energy/momentum smearing to each of the final-state particles based on the particle properties and encoded detector-response parameters. Detection thresholds are applied to determine if a final-state particle will deposit energy in the detector and if that energy deposition will allow for particle identification. The detector responses for neutrons and charged pions account for a variety of possible outcomes that describe the way these particles deposit energy in the detector. Neutral pions are decayed into two photons. Their conversion distance from the point of decay determines the starting position of the resulting electromagnetic showers. This distance is chosen from an exponential distribution with a characteristic length based on the radiation length of photons in liquid argon. Tau leptons are also decayed

by the Fast MC and their decay products are dealt with appropriately. The spatial extent of tracks and showers in liquid argon is simulated in Geant4 and encoded as a probability distribution function (PDF) or parameterization. Combined with vertex placement in a fiducial volume, the fraction of particle energy and/or track length visible in the detector is determined.

Once the Fast MC reconstructs the kinematics of the event (E_ν, E_{had}, Q^2, x, y, and so on), based on the smeared four-vectors of particles that are above detection threshold, it searches interaction final-state particle lists for lepton candidates to be used in event classification algorithms. The resulting classifications are used to isolate samples for the ν_e appearance and the ν_μ disappearance analyses which are in turn used to build energy spectra on an event-by-event basis.

Currently the classification algorithm categorizes each event as either ν_e-CC, ν_μ-CC, or NC. Events with a candidate muon are classified as ν_μ-CC. Events without a candidate muon, but with a candidate electron/positron are classified as ν_e-CC. Events without a candidate muon or a candidate electron/positron are classified as NC. A ν_τ-CC classification, which would identify ν_τ candidates is under development.

A muon candidate is defined as a MIP-like track that is greater than 2.0 m long, and is not consistent with the behavior of a charged pion. Charged pions will often *shower*, depositing a relatively large amount of energy in the detector at the end of its track, as compared to a muon. There are several situations in which a pion topology will be indistinguishable from a muon: (1) the pion stops at the end of its range without interacting, (2) the kinetic energy of the pion is sufficiently small when is showers, (3) the pion is absorbed cleanly by a nucleus with no hadronic debris, (4) the pion decays in flight, and (5) the track exists the detector. The 2.0-m cut was chosen because the probability of (1) or (2) is very small for pion tracks above this threshold.

An additional selection probability is enforced for low-energy tracks to simulate acceptance losses due to increased difficulty in particle identification for short tracks, especially in high-multiplicity events. (The falling edge of the selection probability is well below the energy required to generate a 2.0-m track, minimizing the effect of this criterion.)

An electron candidate is defined as the highest-momentum electromagnetic (EM) shower in an event that is not consistent with a photon. An EM shower is identified as a photon (1) if it converts 2.0 cm or more from the event vertex, (2) if it can be matched with another EM shower in the events to reconstruct the π^0 mass (135 ± 40 MeV), or (3) if dE/dx information from the first several planes of the track is more photon-like than e$\pm$-like. The latter is determined on a probabilistic basis as a function of EM-shower energy and hadronic-shower multiplicity. Signal and background efficiencies from the dE/dx e/γ discriminant are based on MicroBooNE simulations. Cut values are tuned to preserve 95% of the signal across all neutrino energies. As with muon candidates a low-energy selection probability is enforced to account for acceptance losses at low EM-shower energies, especially in high-multiplicity events. For the electron candidates this selection probability is tuned to agree with hand scan studies.

An event with no muon candidate and no electron candidate is assumed to be an NC interaction. Preliminary studies evaluating the use of transverse-momentum imbalance to identify ν_τ-CC interaction candidates have shown promising results for identifying NC candidates as well, and are likely to be included in the near future.

Currently no attempt is made to identify tau lepton candidates in order to isolate a ν_τ-CC sample. A preliminary algorithm to remove $\tau \to \mu + \nu + \nu$ and $\tau \to e + \nu + \nu$ backgrounds has recently been incorporated in the Fast MC. This algorithm may also prove useful for isolating a sample of ν_τ-CC interactions, in which the tau decays to a lepton. Development of an algorithm to identify taus that decay to hadrons is under discussion.

All of the selection criteria can easily be updated to reflect improved simulations or new understanding of particle-identification capabilities and analysis sample acceptances. Changes can also be made to investigate alternate analysis techniques, or more conservative or optimistic assumptions on signal acceptance and/or background-rejection rates. Furthermore, the information required to simulate effects related to particle identification is available in the Fast MC files and users are encouraged to construct and evaluate their own selection criteria.

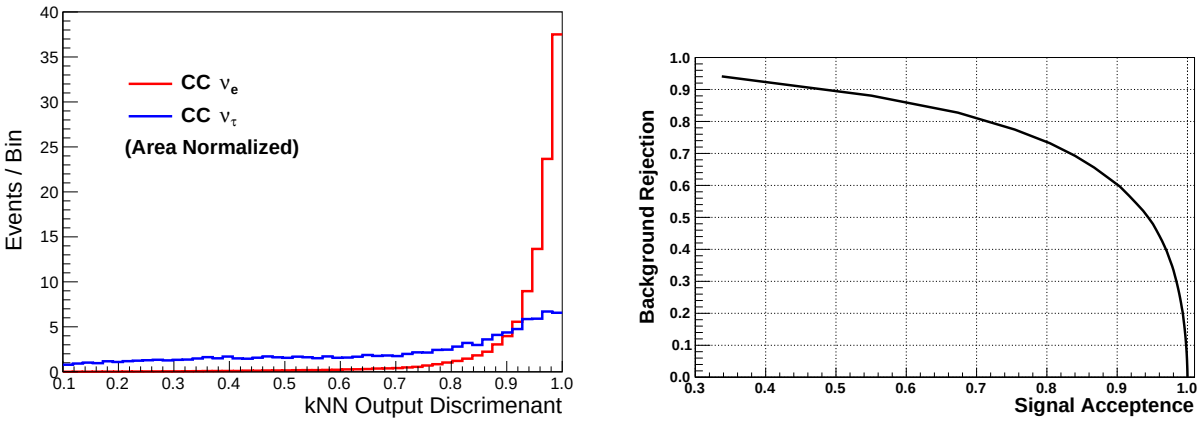

Figure A.6: The output discriminant of a kNN (left) created to remove ν_τ-CC-induced backgrounds from the $\nu_\mu \to \nu_e$ oscillation analysis sample. Signal events (red) tend toward high values, while the ν_τ-CC-induced background events (blue) are more evenly distributed. The fraction of ν_τ-CC-induced backgrounds removed from the $\nu_\mu \to \nu_e$ appearance candidate sample as a function of the corresponding signal efficiency (right). The curve is generated by varying the cut value on the kNN discriminant.

A preliminary algorithm for removing ν_τ-CC-induced backgrounds from from the ν_μ-CC and the ν_e-CC samples has been developed. It employs a k-Nearest Neighbor (kNN) machine-learning technique as implemented in the ROOT TMVA package. The inputs to the kNN are (1) the sum of the transverse momentum with respect to the incoming neutrino direction, (2) the reconstructed energy of the incoming neutrino, and (3) the reconstructed energy of the resulting hadronic shower. Figure A.6 (right) shows the distribution of the output discriminant for true ν_e-CC signal events, and for true ν_τ-CC-induced backgrounds. The algorithm is still being optimized but initial results are promising.

As can be seen in Figure A.6 (left), cuts on the discriminant that preserve 90% of the signal remove roughly 60% of the ν_τ-CC-induced background in the ν_e-CC sample. Similar results are expected for the ν_τ-CC-induced background in the ν_μ-CC sample.

A similar approach is being studied to isolate the ν_τ-CC sample for the ν_τ-CC appearance analysis. Current efforts are focused on identifying a set of reconstructed quantities that separate ν_τ-CC interactions from potential backgrounds. For leptonic decay channels the quantities used in the above kNN are prime candidates. Attempts to reconstruct a ρ mass from tracks originating at the vertex are expected to help to isolate hadronic τ decays. The parameterized pion response will allow for selection of high-energy charged pions produced in hadronic τ decays.

Figures A.7 and A.8 show the Fast MC reconstructed energy spectra of the signal and background for the ν_e appearance and the ν_μ disappearance samples, respectively. As an example of the cross-section and nuclear-effect systematics that can be studied, the black histograms and the bottom insert in each plot show the variation of the spectrum for each event type induced by changing the value of CC M_A^{res} by +1σ (+15%, 2014 GENIE official uncertainty). CC M_A^{res} is the axial mass parameter appearing in the axial form factor describing resonance production interactions in GENIE. This particular example demonstrates a spectral distortion that is not a simple normalization and is different for signal and for background. The effect of varying CC M_A^{res} on the $\nu_\mu \to \nu_e$ analysis sample exhibits a strong correlation with the changes induced in the $\nu_\mu \to \nu_\mu$ analysis sample.

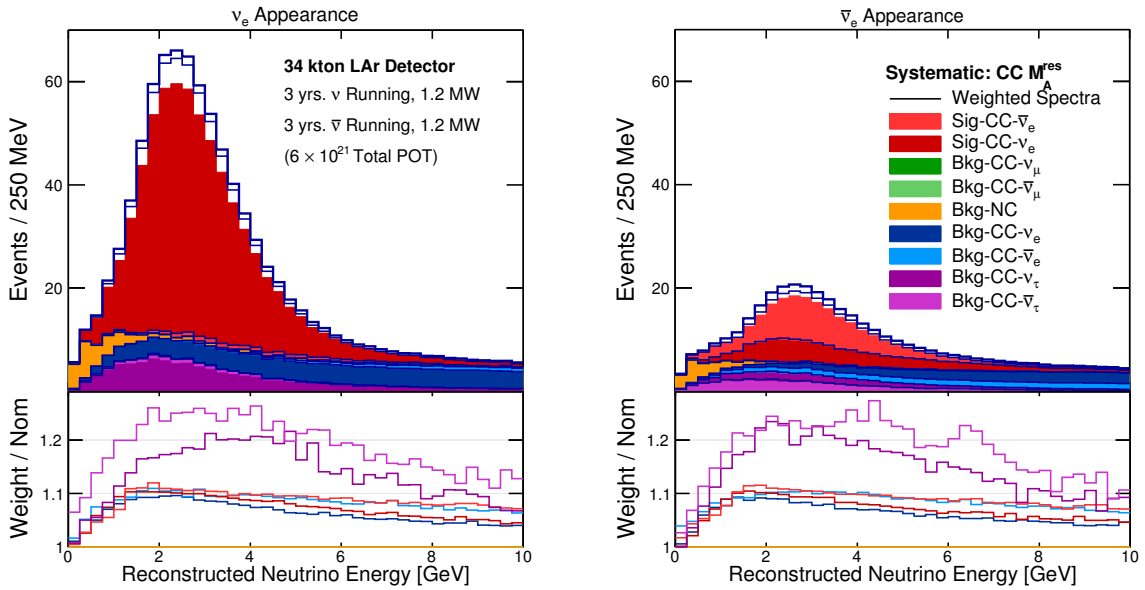

Figure A.7: The reconstructed energy distributions for the signals and backgrounds in the ν_e- (left) and $\overline{\nu}_e$ appearance (right) samples, as predicted by the Fast MC. The black histograms and bottom insert in each plot shows, for each event type, the variation in the spectrum that is induced by changing the value of CC M_A^{res} by +15%.

The left-hand plots of Figures A.9 and A.10 show the acceptance (efficiency) of the signal and the background for the Fast MC ν_e appearance and ν_μ disappearance selections, respectively. The

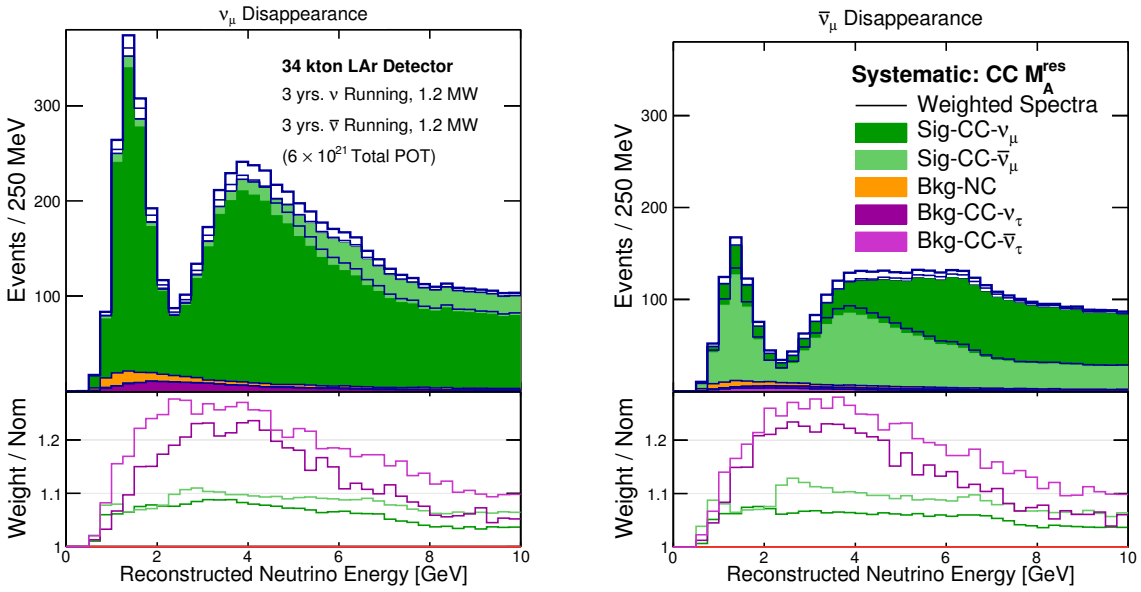

Figure A.8: The reconstructed energy distributions for the signals and backgrounds in the ν_μ (left) and $\overline{\nu}_\mu$ disappearance (right) samples, as predicted by the Fast MC. The black histograms and bottom insert in each plot shows, for each event type, the variation in the spectrum that is induced by changing the value of CC M_A^{res} by +15%.

effects of the low-energy selection probabilities induce the observed low-energy fall off in the ν_e appearance sample. On the other hand, the 2.0-m track length requirement is mainly responsible for the low-energy behavior in the ν_μ disappearance sample. The corresponding plots on the right-hand side show the relative fraction (purity) of the signal and each background sample for the Fast MC ν_e appearance and ν_μ disappearance selections. The increased wrong-sign contamination is evident in the $\overline{\nu}$ beam samples as compared to the ν beam samples. No attempt has been made to reduce the ν_τ background in these plots.

The output of the Fast MC is a file containing the information one would expect from a full MC simulation. There are truth level quantities that describe the generated event, and reconstructed quantities that are calculated from simulated observables. The latter mimic the information that is expected to be available from reconstructing data or full simulation and can be used in designing analyses aimed at measuring physics parameters. Analyses based on the simulated reconstruction produce event samples that can be used to estimate the sensitivity of LBNE to physics model parameters, specifically the parameters of the PMNS matrix, as a function of a variety of input parameters. Currently these studies are done using the GLoBES [14] software package. However, instead of constructing the event-rate spectra as a function of true neutrino energy from predictions of the flux and neutrino-interaction cross sections, they are built event-by-event from the Fast MC. Similarly, smearing functions that give the distribution of measured (reconstructed) neutrino energies as a function of the true neutrino energy are built event-by-event from the Fast MC, rather than estimated from external sources.

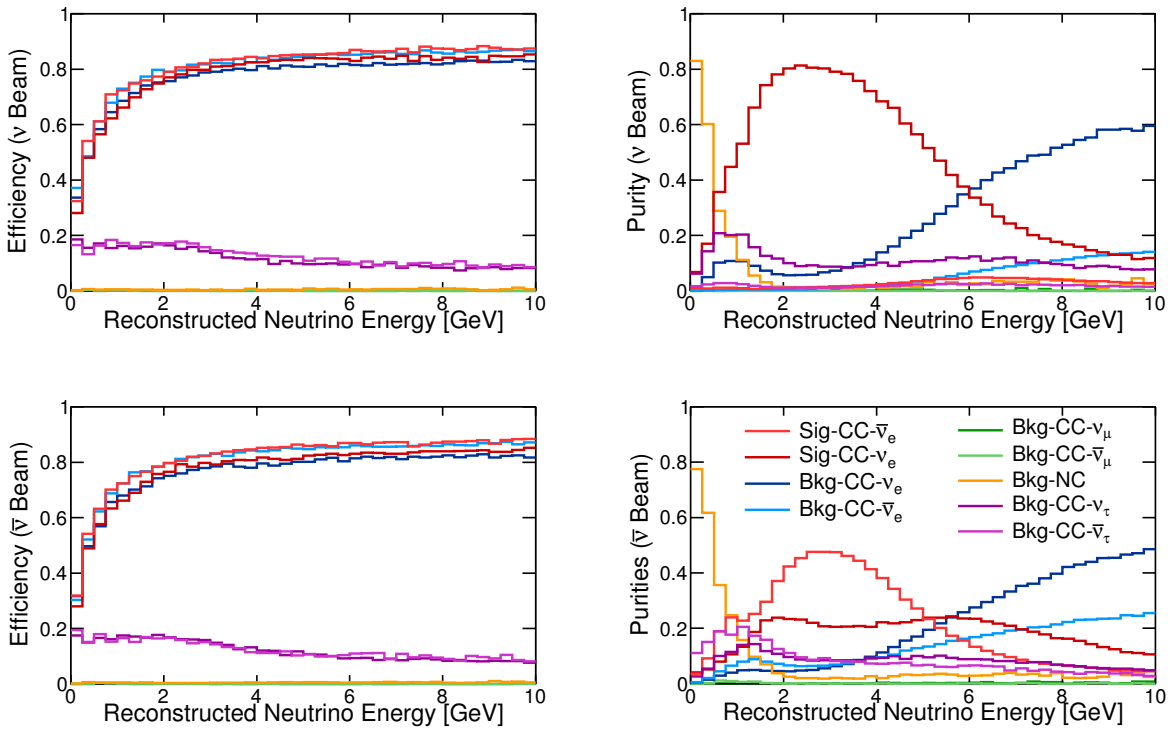

Figure A.9: The expected efficiencies and purities of selecting ν_e appearance events in a LArTPC, obtained from the Fast MC.

In addition to the usual GLoBES inputs the Fast MC can provide systematic uncertainty response functions, which encode the expected changes to the energy spectra when input model parameters are varied within their uncertainties. These response functions, along with an augmented version of GLoBES, can be used to propagate realistic systematic uncertainties in sensitivity studies.

The systematic uncertainty response functions are calculated from weights stored in the Fast MC output files. Each weight corresponds to the probability of producing the event with an alternate physics model relative to the model used. Currently the Fast MC generates weights for parameters in interaction models that can be reweighted in GENIE as well as a variety of parameters related to the neutrino flux. The flux parameters come in three varieties related to: changes to the beamline design, tolerances in the beamline design, and uncertainties in the physics models used in the simulations. The latter two contribute to systematic uncertainties while the first can be used to evaluate the impact of design optimizations.

Propagation of systematic uncertainties through LBNE sensitivity studies using the Fast MC will require inclusion of new algorithms and improvements to existing reweighting algorithms. This includes (1) the introduction of new models into GENIE, (2) adding to and improving the reweighting functions currently in GENIE, (3) constructing flux files that correspond to the changes in the three aforementioned categories, (4) implementing a system for reweighting individual events based on

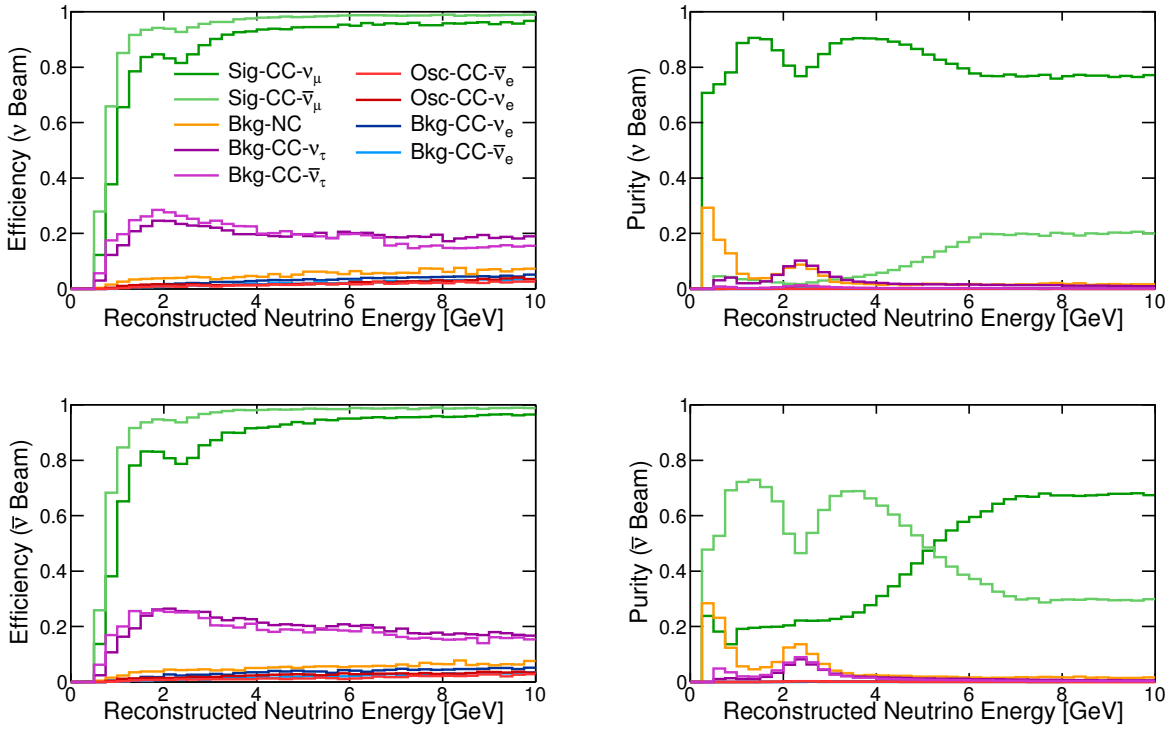

Figure A.10: The expected efficiencies and purities of selecting ν_μ disappearance events in a LArTPC obtained from the Fast MC.

changes to the models of hadronization from proton-target interactions, and (5) introducing detector parameterizations representing alternate detector designs, detector design tolerances, and model choices used in detector simulations.

The current focus of Fast MC studies is estimation of the effect of model uncertainties on sensitivity projections. This includes several steps, the first of which is to look at the changes in the analysis sample spectra induced by propagating individual systematic uncertainties. These studies are benchmarked by calculating the χ^2 between the nominal and altered spectra. In the second step, sensitivities are calculated for combined fits of the four main analysis samples ($\nu_\mu/\overline{\nu}_\mu$ disappearance, $\nu_e/\overline{\nu}_e$ appearance). These studies must be done carefully to allow for realistic constraints of systematic uncertainties across analysis samples within GLoBES. Input covariance matrices can also be used to enforce external constraints on the relations between sources of systematic uncertainty. The results of these studies will inform the investigators as to which model uncertainties cause significant degradation of the sensitivities and therefore must be constrained by other methods. Methods to constrain these parameters will be sought from currently running experiments, proposed intermediate experiments, and from the LBNE beam monitoring and the LBNE near detector. Estimates of these constraints can then be propagated to sensitivity calculations to estimate the degree to which they mitigate the decline in sensitivity.

Current studies focus on propagating uncertainties in flux and GENIE model parameters via reweighting techniques. A example study shown in Figure A.11 illustrates the effect of including the uncertainty on CC M_A^{res} in the calculation of sensitivity to CP violation. The sensitivity studies are performed for (1) a fit to the ν_e appearance sample (three years of ν-beam running), (2) a combined fit of the ν_e appearance sample and the $\overline{\nu}_e$ appearance sample (three years of ν-beam plus three years of $\overline{\nu}$-beam running), and (3) a combined fit of the $\nu_e/\overline{\nu}_e$ appearance samples along with the corresponding $\nu_\mu/\overline{\nu}_\mu$ disappearance samples. All three studies are done in two ways: with no allowance for non-oscillation parameter systematic variation, and with allowed 15% (width gaussian PDF) variations in CC M_A^{res}.

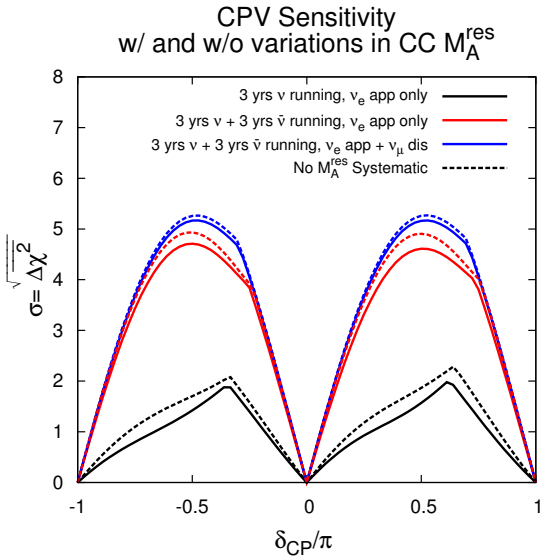

Figure A.11: The sensitivity to CP violation calculated using the energy spectra generated by the Fast MC. The sensitivities were generated with (solid) and without (dashed) allowed variations in the CC M_A^{res} resonance production model parameter in GENIE. The allowed variation degrades the sensitivity, however combined fits of multiple analysis samples provide additional constraints and reduce the impact.

As Figure A.11 shows, the inclusion of allowed variations in CC M_A^{res} degrades the sensitivity. However, combined fits of multiple analysis samples provide additional constraints and reduce the impact. The effect of these sample-to-sample constraints is dependent on the sample statistics, and the curves in Figure A.11 include the statistical limitations on sample-to-sample constraints from a six-year (three years ν + three years $\overline{\nu}$ running) exposure. However, the software also allows for the inclusion of other possible limitations on sample-to-sample constraints related to the relative lack of experimental constraints on cross-section ratios (i.e., $\sigma_{\nu_e}/\sigma_{\nu_\mu}$, $\sigma_{\nu_\tau}/\sigma_{\nu_\mu}$, and $\sigma_{\overline{\nu}}/\sigma_{\nu}$), as well as theoretical considerations.

The preliminary Fast MC spectra shown in Figures A.7 and A.8 were generated with a different beam configuration than the ones shown in Figures 4.2 and 4.3. Consequently, the sensitivities to CPV shown in Figure A.11 cannot be directly compared to the corresponding figures in Section 4.2.

However, both the Fast MC and the methods discussed in Section 4.2 have been used to generate comparable spectra and to perform a series of sensitivity studies. The two methods are consistent, except regarding known differences between the two simulations, e.g., the inclusion of ν_τ-CC-induced backgrounds. These differences are well understood, as are their impact on oscillation parameter sensitivities.

Eventually the Fast MC seeks to incorporate near detector and atmospheric-neutrino analyses and directly perform combined fits with the long-baseline neutrino analysis samples. These studies will provide the most accurate estimate of the ultimate sensitivity of LBNE, and provide a template for future data analysis procedures.

A.4 Simulation of Cosmic-Ray Background for a 10-kt Surface Detector

A preliminary study of the background events expected from cosmic rays in the 10-kt far detector installed near the surface at the Sanford Underground Research Facility is detailed in [15]. The study simulated cosmic-ray interactions in the far detector and focused on cosmic-ray induced events from neutrons and muons that mimic electron-neutrino interactions in the detector. These include electromagnetic cascades from knock-on electrons, muon bremsstrahlung, and hadronic cascades with electromagnetic components from photons and π^0's. The background from decays of neutral hadrons into electrons such as $K^0_L \to \pi e \nu$ were also studied. The energy of the cascades was required to be > 0.1 GeV.

These initial studies indicate that a combination of simple kinematic and beam timing cuts will help to significantly reduce the cosmic-ray background event rate in this far detector configuration. In particular:

1. Only electromagnetic cascades with energies greater than 0.25 GeV are considered background. For the neutrino oscillation sensitivity calculations, only neutrino energies ≥ 0.5 GeV are considered.

2. $e^\pm$ background candidates are tracked back to the parent muon; the distance between the muon track and the point-of-closest-approach (PoCA) to the muon track is required to be > 10 cm.

3. The vertex of the $e^\pm$ shower is required to be within the fiducial volume of the detector (defined as 30 cm from the edge of the active detector volume).

4. The $e^\pm$ cascade is required to be within a cone around the beam direction (determined from the angular distribution of the beam signal $e^\pm$ and the incoming neutrino beam).

5. It is assumed that EM showers initiated by γ's and $\pi^0 \to \gamma\gamma$ can be effectively distinguished from primary electron interactions using particle ID techniques such as dE/dX.

6. Events are timed with a precision of $\leq$1 µs using the photon-detection system, which limits background to events occurring within the 10 µs of the beam spill.

The result of applying these selection criteria to the electromagnetic showers initiated by cosmic rays is summarized in Table A.1 and Figure A.12. The background rates given in Table A.1 include the recalculation for the cosmic flux at 1,500 m above sea level, which was not included in the previous study [15] (and is not included in Figure A.12). In the table, the initial background event rate is calculated for one calendar year assuming a 1.4-ms drift time per beam pulse, a beam pulse every 1.33 seconds and 2×10^7 s/year of running. The expected event rate/yr after various selection criteria is applied from left to right in the table. The rates in all columns except the last are given for a time window of 1.4 ms, corresponding to the maximum electron drift time. The last column shows the rate reduction assuming an efficient photon-detection system. The first three rows show events with a muon in the detector where a PoCA cut (column 3) can be applied. The row labeled 'Missing μ' shows events without a muon in the detector; as there is no muon track, a PoCA cut can not be applied. The detector is assumed to be on the surface with three meters of rock overburden.

Table A.1: Cosmic-ray-induced background (at 1,500 m above sea level) to the beam ν_e-CC signal in the 10-kt detector.

Processes	**$E_e > 0.25$ GeV**	**PoCA $>$ 10 cm and D $>$ 30 cm**	**Beam angle**	**e/γ PID**	**Beam timing**
$\pi^0 \to \gamma \to e^{\pm}$	2.2×10^6	9.7×10^4	4.8×10^4	1.7×10^3	12
$\mu \to \gamma \to e^{\pm}$	7.1×10^6	12	0	0	< 0.003
Ext $\gamma \to e^{\pm}$	1.9×10^6	660	340	13	0.1
$\pi^0, K^0 \to e^{\pm}$	1.4×10^6	810	240	240	1.7
Missing μ	1.3×10^6	1.8×10^3	580	20	0.1
Atm n	2.9×10^6	1.6×10^4	6.5×10^2	240	1.7
Total	$\mathbf{1.1 \times 10^7}$	$\mathbf{1.2 \times 10^5}$	$\mathbf{5.6 \times 10^4}$	$\mathbf{2.2 \times 10^3}$	**16**

The dominant background is from $\pi^0 \to \gamma \to e^{\pm}$, which contributes 12 out of the 16 total events per year and comes from π^0's originating in cosmic showers. The study does not yet include specific π^0 reconstruction, only individual e/γ separation. More sophisticated reconstruction techniques should further reduce the π^0 background. The studies indicate that application of these selection criteria coupled with a more detailed background event reconstruction can potentially reduce the background from cosmic rays to a few events per year — mostly in the energy region < 1 GeV.

In Figure A.12, black-filled circles show events before any cuts are applied. The other point icons represent successively applied cuts in the order listed below and in the figure's legend:

1. Blue squares: PoCA to the muon track greater than 30 cm
2. Red triangles: angle with respect to the beam such that 99% of signal events are retained
3. Green triangles: application of energy-dependent e/γ discrimination
4. Magenta open circles: application of efficient photon detection, this allows the reduction of the time window from a maximum drift time of 1.4 ms down to a beam spill of 10 µs

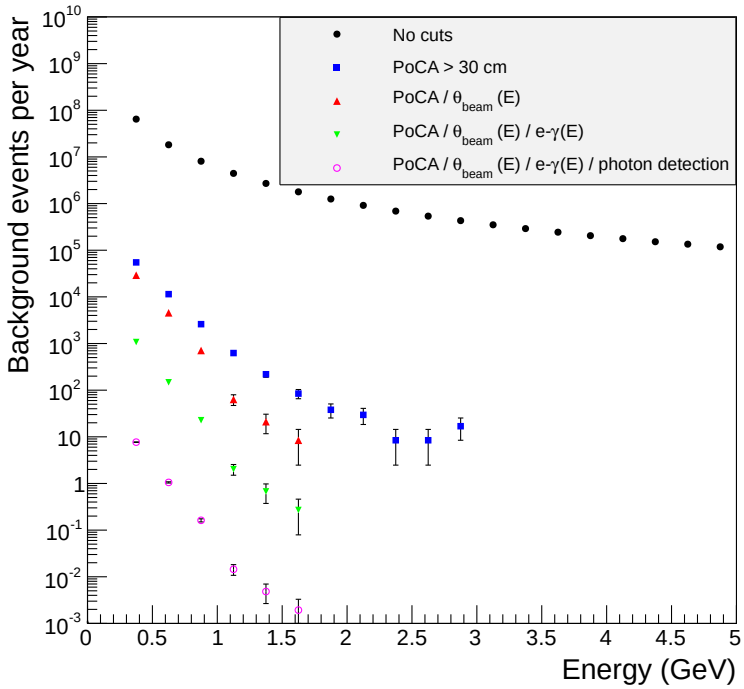

Figure A.12: Energy spectra of muon-induced background events for successively applied background rejection cuts. Simulations have been done for a muon spectrum at sea level. Correction for an altitude of 1,500 m above sea level has not been applied to the data.

References

1. **LBNE Project Management Team** , "LBNE Conceptual Design Report, Volume 1: The LBNE Project," LBNE-doc-5235, 2012. Cited in Section A (pg.213).
2. S. Agostinelli *et al.*, **GEANT4** , "GEANT4: A simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250–303. Cited in Section A.1.1 (pg.213).
3. E. D. Church, "LArSoft: A Software Package for Liquid Argon Time Projection Drift Chambers," arXiv:1311.6774 [physics.ins-det], 2013. Cited in Section A.1.1 (pg.213).
4. C. Green, J. Kowalkowski, M. Paterno, M. Fischler, L. Garren, *et al.*, "The art framework," *J.Phys.Conf.Ser.* **396** (2012) 022020. Cited in Section A.1.1 (pg.213).
5. T. Katori, **MicroBooNE Collaboration** , "MicroBooNE, A Liquid Argon Time Projection Chamber (LArTPC) Neutrino Experiment," *AIP Conf.Proc.* **1405** (2011) 250–255, arXiv:1107.5112 [hep-ex]. Cited in Sections A.1.1 (pg.213) and A.3 (pg.221).
6. M. Soderberg, **ArgoNeuT Collaboration** , "ArgoNeuT: A Liquid Argon Time Projection Chamber Test in the NuMI Beamline," FERMILAB-CONF-09-516-E, arXiv:0910.3433 [physics.ins-det], 2009. Cited in Section A.1.1 (pg.213).
7. D.Huffman, "A Method for the Construction of Minimum-Redundancy Codes," in *Proceedings of the IRE*. 1952. Cited in Section A.1.1 (pg.215).
8. M. Szydagis, N. Barry, K. Kazkaz, J. Mock, D. Stolp, *et al.*, "NEST: A Comprehensive Model for Scintillation Yield in Liquid Xenon," *JINST* **6** (2011) P10002, arXiv:1106.1613 [physics.ins-det]. Cited in Section A.1.1 (pg.215).
9. C. Andreopoulos, **GENIE Collaboration** , "The GENIE neutrino Monte Carlo generator," *Acta Phys.Polon.* **B40** (2009) 2461–2475. Cited in Section A.1.1 (pg.216).
10. C. Hagman, D. Lange, J. Verbeke, and D. Wright, "Cosmic-ray Shower Library (CRY)," Lawrence Livermore National Laboratory, UCRL-TM-229453, March, 2012. http://nuclear.llnl.gov/simulation/doc_cry_v1.7/cry.pdf. Cited in Section A.1.1 (pg.216).
11. R. P. Sandhir, S. Muhuri, and T. Nayak, "Dynamic Fuzzy c-Means (dFCM) Clustering and its Application to Calorimetric Data Reconstruction in High Energy Physics," *Nucl.Instrum.Meth.* **A681** (2012) 34–43, arXiv:1204.3459 [nucl-ex]. Cited in Section A.2 (pg.218).
12. R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering* **82** no. Series D, (1960) 35–45. Cited in Section A.2 (pg.218).
13. J. Marshall and M. Thomson, "The Pandora software development kit for particle flow calorimetry," *J.Phys.Conf.Ser.* **396** (2012) 022034. Cited in Section A.2 (pg.219).

14. P. Huber, M. Lindner, and W. Winter, "Simulation of long-baseline neutrino oscillation experiments with GLoBES (General Long Baseline Experiment Simulator)," *Comput.Phys.Commun.* **167** (2005) 195, arXiv:hep-ph/0407333 [hep-ph]. Cited in Section A.3 (pg.225).
15. E. Church *et al.*, "Muon-induced background for beam neutrinos at the surface," LBNE-doc-6232, October, 2012. Cited in Sections A.4 (pg.229) and A.4 (pg.230).