Analysis Methods in Long-Baseline Accelerator-Based Neutrino Experiments

Thomas R. Junk

Fermilab

January 12, 2016

Brookhaven National Laboratory HEP Seminar

Open Questions in Neutrino Physics

Are Neutrinos Majorana or Dirac? Can they be a mixture?

Is there CP-violation in the neutrino sector?

What is the value of δ_{CP}?

What can we learn of leptogenesis from measurements we can make at accessible energy scales?

Which ordering, Normal or Inverted, characterizes the neutrino masses?

Are there light sterile neutrinos?

Is the 3x3 PMNS matrix unitary?

Are there heavy neutrinos?

Open Questions in Neutrino Physics

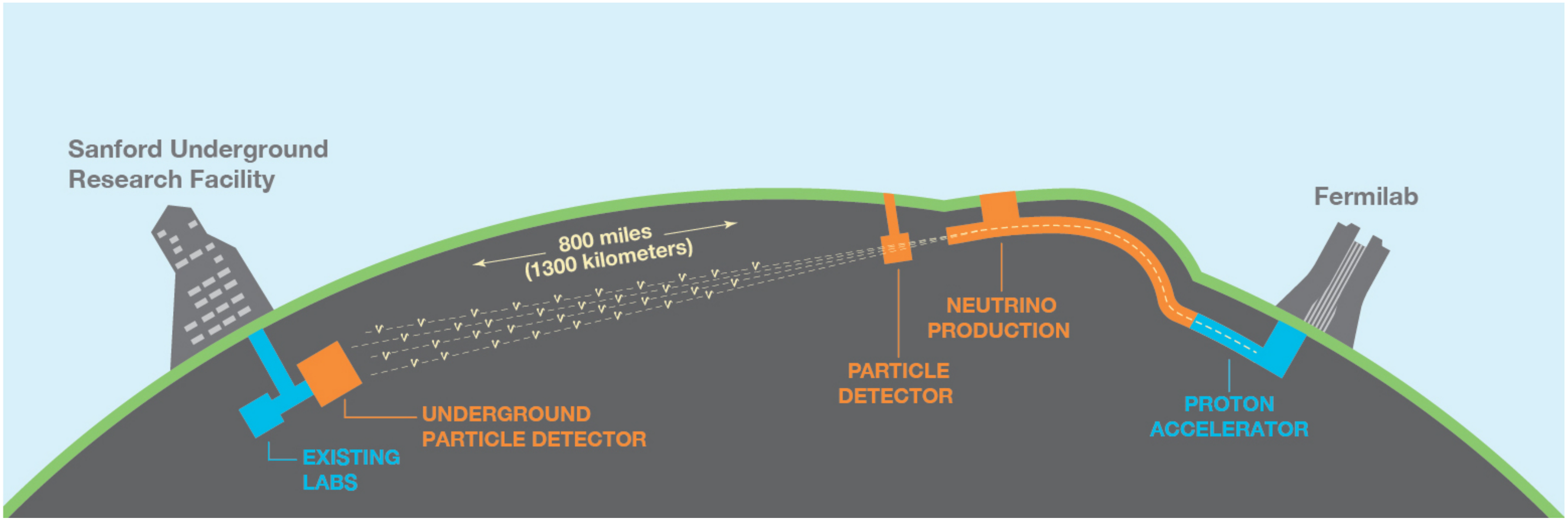
Do neutrinos get their masses from the Higgs mechanism, or some other means?

What can we learn about supernovae from neutrinos we can detect?

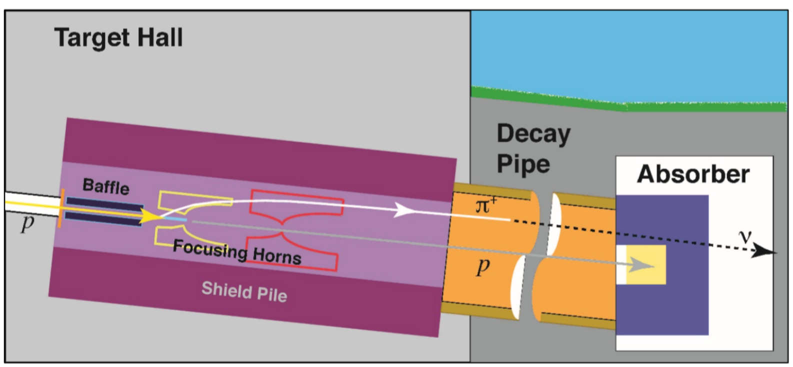
Neutrino astronomy is brand new, and requires ambitious experiments to perform. What can we learn of the cosmos from neutrinos?

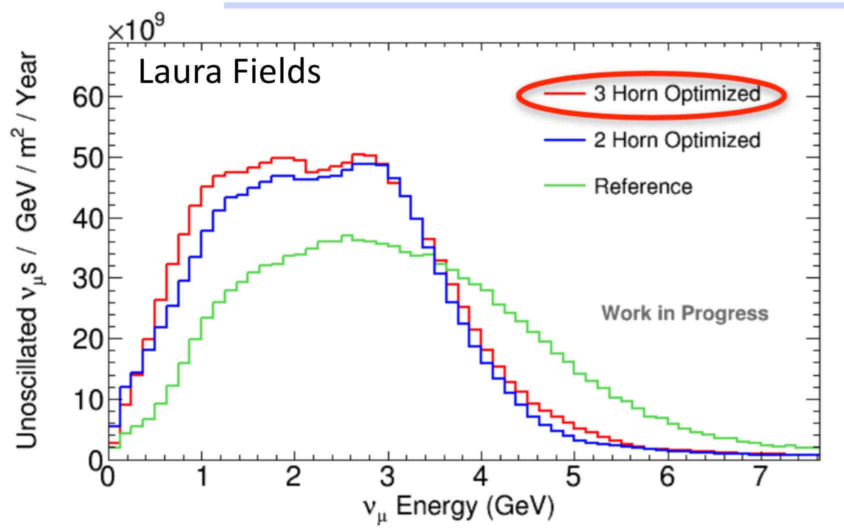
Not neutrino physics but you get these too: Are nucleons unstable?
Do neutrons oscillate into their own antiparticles?

DUNE: Deep Underground Neutrino Experiment



Cartoon of Neutrino Beamline from DUNE CD-1R CDR





Current best optimization has three focusing horns
Horn current sign selects neutrinos or antineutrinos

The DUNE Near Detector

Magnetized Spectrometer

Straw-Tube Tracker

Lead-Scintillator ECAL

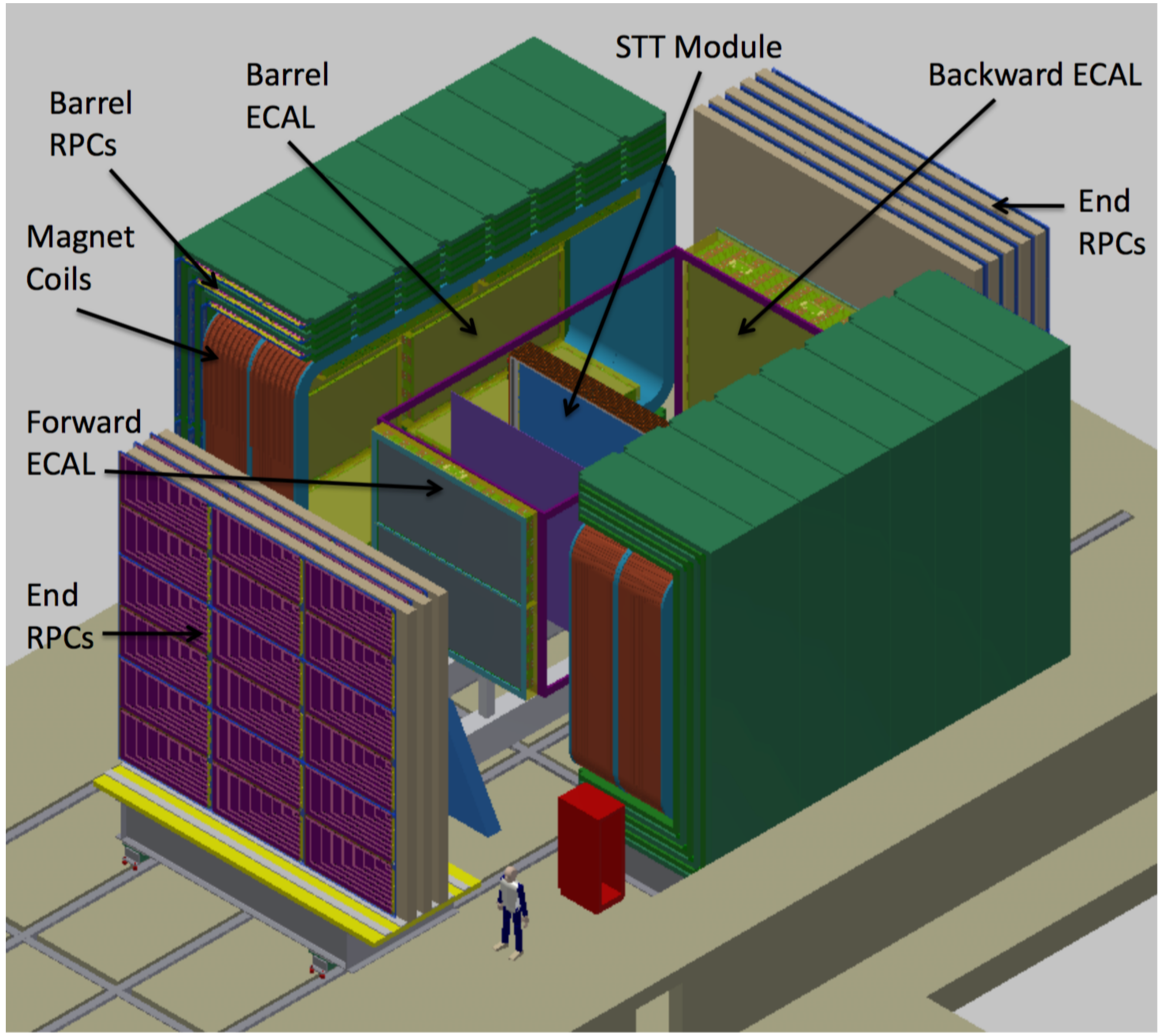
Resistive-Plate Chambers
for muon ID

Two competing alternatives:

Liquid Argon TPC (like the Far Detector)

Gaseous Argon TPC

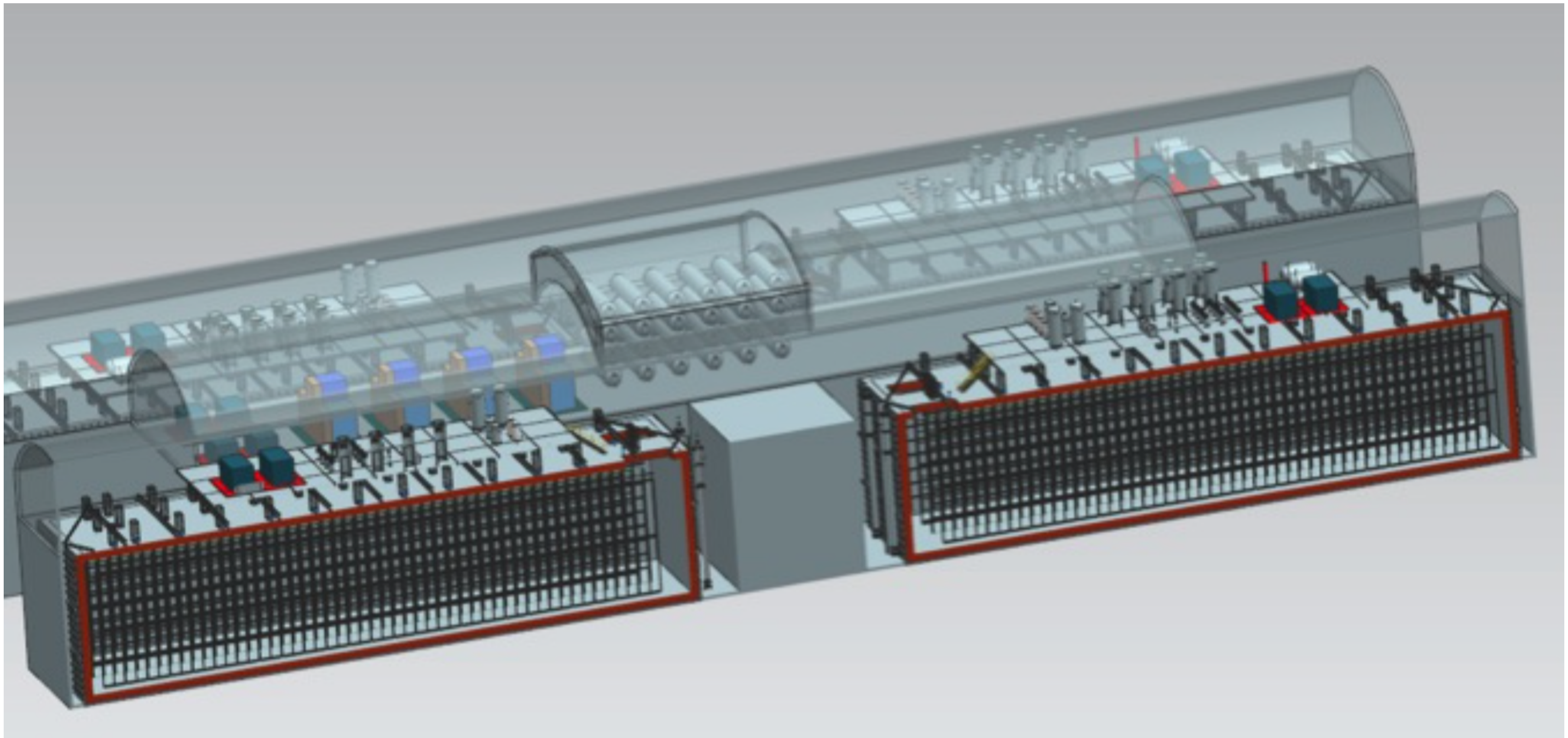
Reference Design – Fine-Grained Tracker



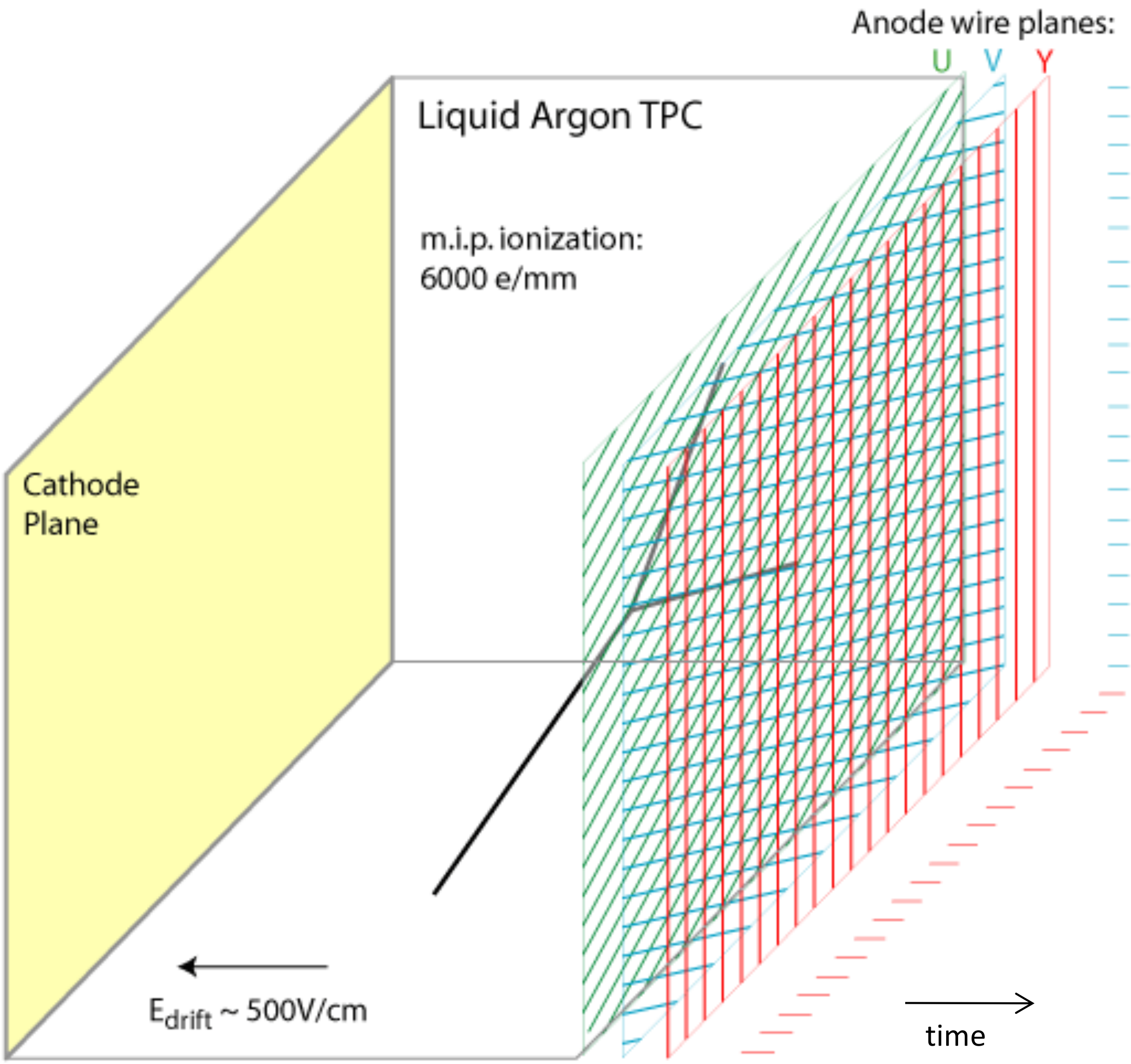
The DUNE Far Detector

- Liquid-Argon TPC
- Four modules with 10 kton fiducial mass each: 40 kton total. 70 kton of liquid argon
- Located at the 4850' Level of SURF, 1300 km from Fermilab
- Reduced cosmic-ray rate at depth enables a rich physics program

  Proton Decay, SNB, Atmospherics, rare events
- First module will be a single-phase detector like ICARUS and MicroBooNE
- Subsequent modules may be single-phase or dual-phase detectors
- Start installing Module 1 in 2021, Ready to run in 2024



Liquid-Argon Time-Projection Chambers (Single-Phase)



Charged particle leaves a trail of ionized Argon atoms and drifting electrons.

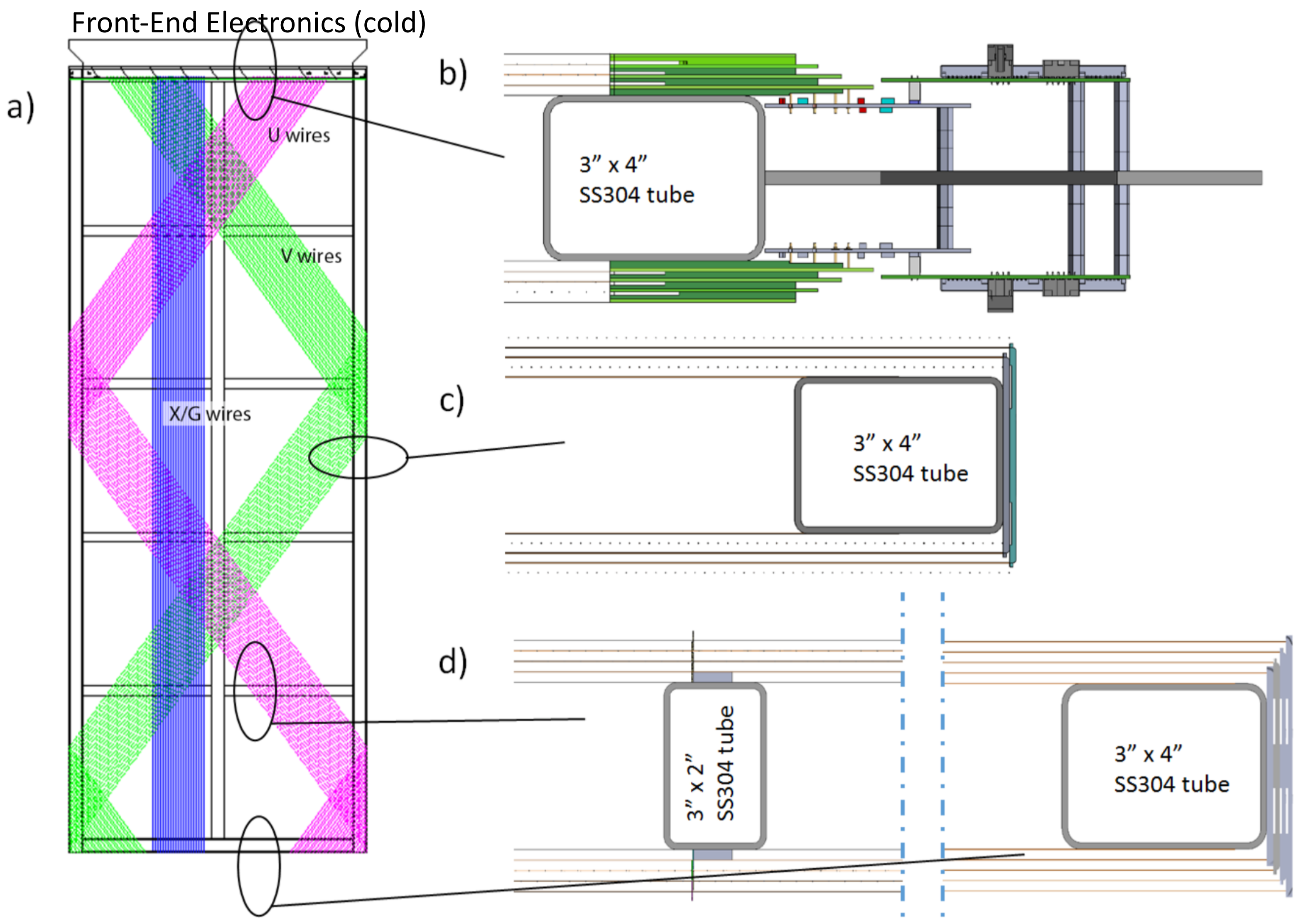
Some recombination at the ionization point, some attach to impurities during the drift.

Signals picked up on induction-plane wires and collection-plane wires.

Two-dimensional views of a three-dimensional interaction

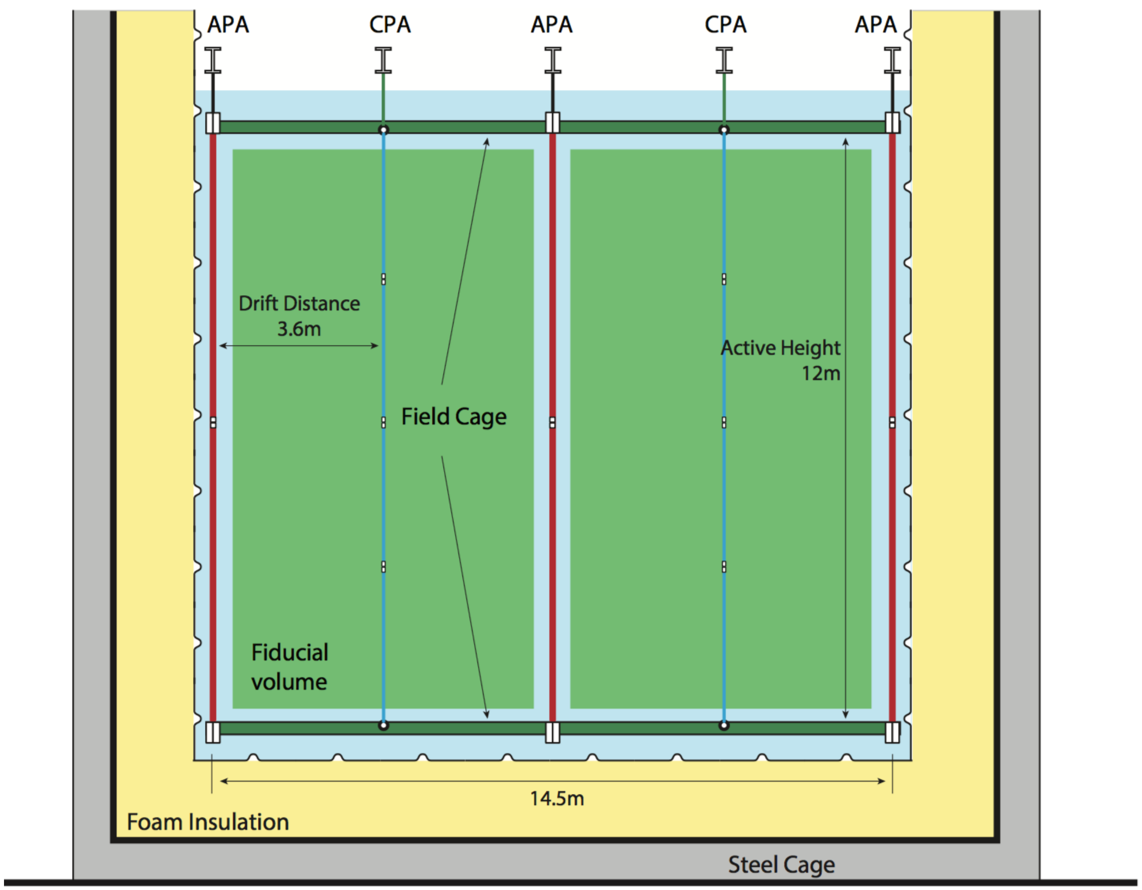
The DUNE Single-Phase Far Detector Anode Plane Assemblies

Wrapped *U* and *V* induction-plane wires provide for front-and-back readout with electronics only needed on one end.



2560 readout wires per APA

The DUNE Single-Phase Far Detector



3.6 meter drift
Two APA's stacked vertically, each 6 m tall.
Drift field: 500 V/cm: 180 kV on the cathode
Anodes are grounded
150 APA's per 10 kt module

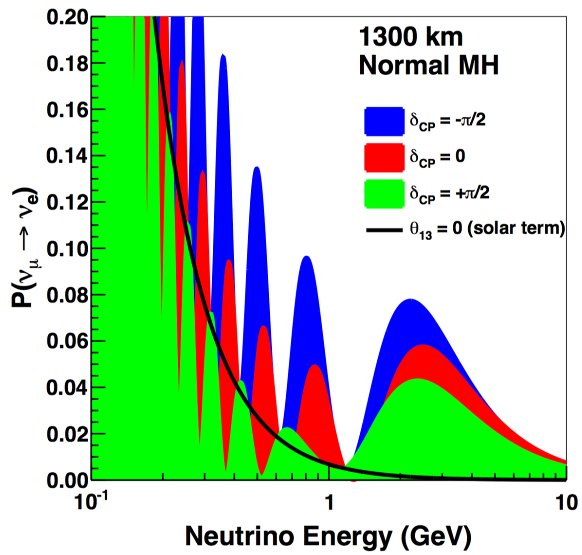
4.7 mm wire pitch
384,000 readout wires per 10 kt module
500 MHz digitization (12-bits) for each readout wire.

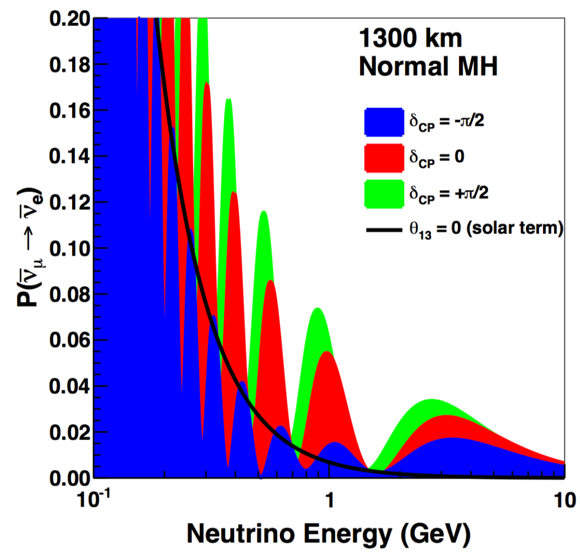
Photon detectors inside APA's

Not magnetized

Beam perpendicular to electric field

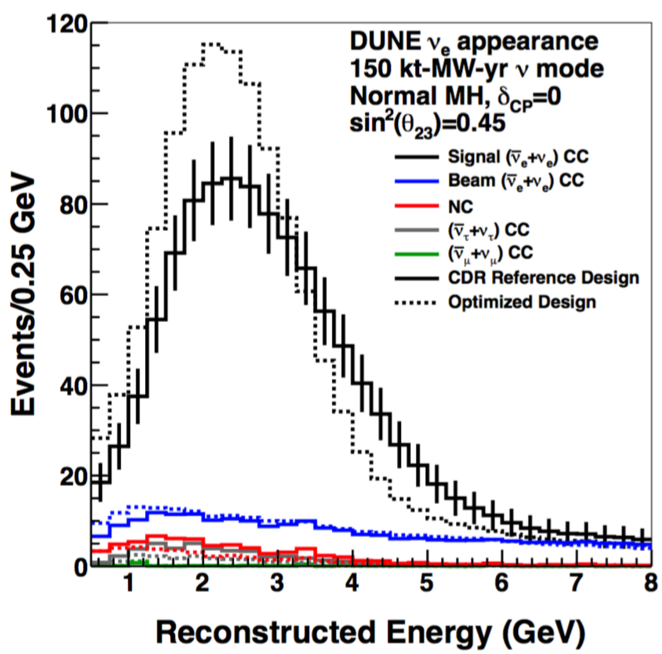
Sensitivity of Observable Distributions to Oscillation Parameters

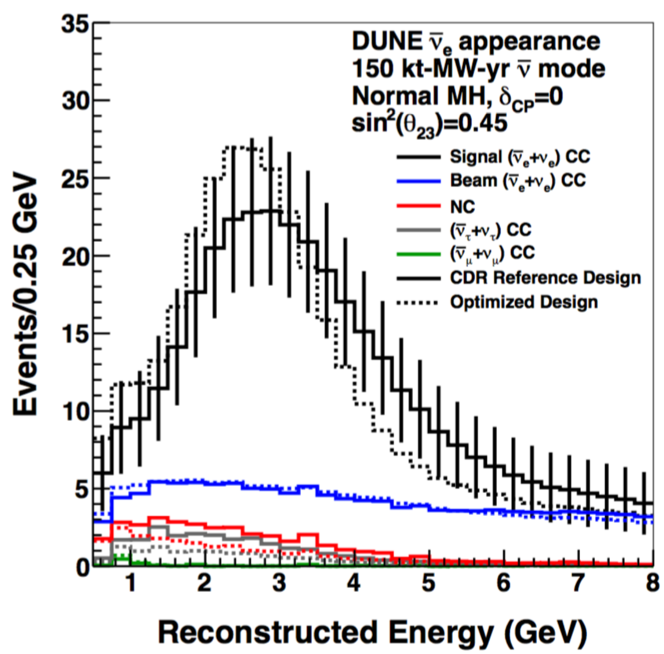




Oscillation Probabilities depend on L/E, and matter.

With L fixed, E needs to be measured in a broad-band beam.





Expected reconstructed energy distributions in $\boldsymbol{\nu}_e$CC-selected events in DUNE .

Assumes 99% rejection of NC events. 10x background (90% rejection) would reduce the sensitivity by an enormous amount.

Source: DUNE CD1-R CDR

A Physics Program Beyond Oscillations

Neutrino cross section measurements on Argon, Carbon, Oxygen, Iron, electrons
- inclusive, and
- exclusive, by final state

Beam Flux determination

Nuclear physics effects:
- Multiparticle interactions in the nucleus
- Final-State Interactions
- Hadron formation and absorption

Search for Exotic new particles

Other exotica:
- CPT violation
- Neutrino Decay
- Violation of the rules of Quantum Mechanics
- Anomalous neutrino magnetic moments

Ambiguities in Interpreting Individual Events

- Misidentification rates are non-zero: The big ones of concern:
  - electron and a photon from π^0 decay
  - muons and charged pions
  - charged pions and protons
  - Short track vs. no track
  - Missing energy – neutrons, neutrinos, and escaping photons

- Causes misclassification of events among ν_μCC, NC, and ν_eCC

- Energy resolution is not perfect – the measurements of an event do not determine the energy unambiguously

- For atmospheric neutrinos – the direction measurement is also not perfect affects L measurement

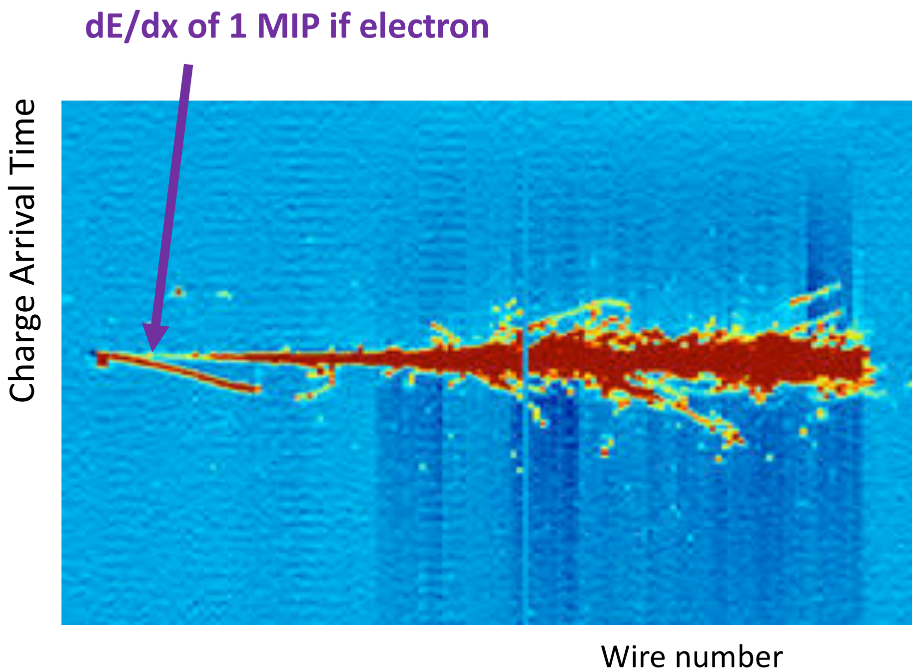
- Detector inhomogeneity and anisotropy contribute, but these effects are small isochronous tracks wire-wrapping ambiguities

- Two approaches to handling uncertainty:
  - Best-fit and local uncertainties: "Profile"
  - Integrate over them: "Marginalize"

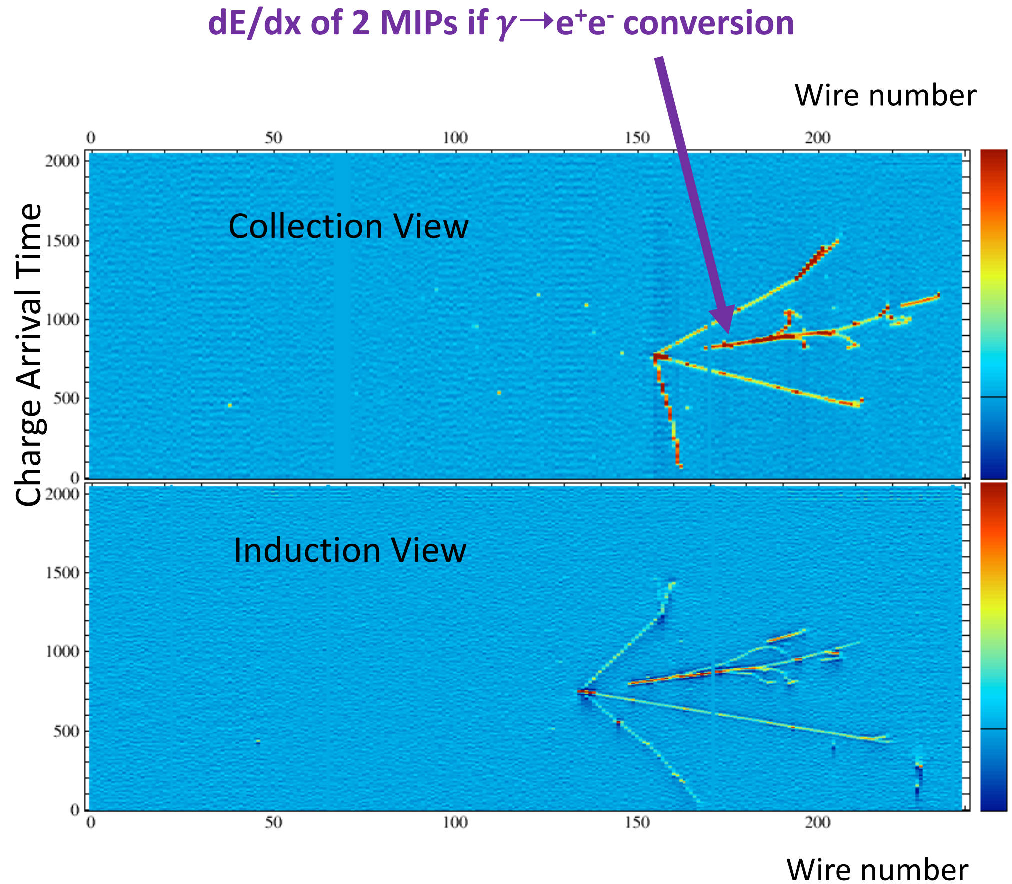
Event Classification

- Different categories of events have different signal and background expectations.
- Example: The dE/dx of the initial part of the electron shower is needed to distinguish between 1-MIP and 2-MIP signals.
  - The first 2.5 cm is the most important part (5 hits in the best case, fewer in views with wires not perpendicular to the track)
  - Ideally you should know where the primary vertex is so you can start the 2.5 cm search.
  - Events with additional hadronic recoil (protons, charged pions) have a well-defined primary vertex
  - Events with just one EM shower and nothing else are more ambiguous: half a pizero, or an electron?
  - More NC background expected in the shower+nothing sample.

Example Events in ArgoNeuT with EM showers



A ν_eCC candidate with a recoiling proton



An event with a $\pi \rightarrow \gamma\gamma$ candidate in which only one γ conversion is visible

Two discriminants here – displaced vertex topology, and dE/dx.
Can use the data to calibrate one by using the other to purify a sample of electrons or photons.
Need a model of the correlations however.

If all you have is the shower and no other tracks from the primary vertex, it is harder to look at the first 2.5 cm

Separation of Electrons and Photons from Ionization Alone

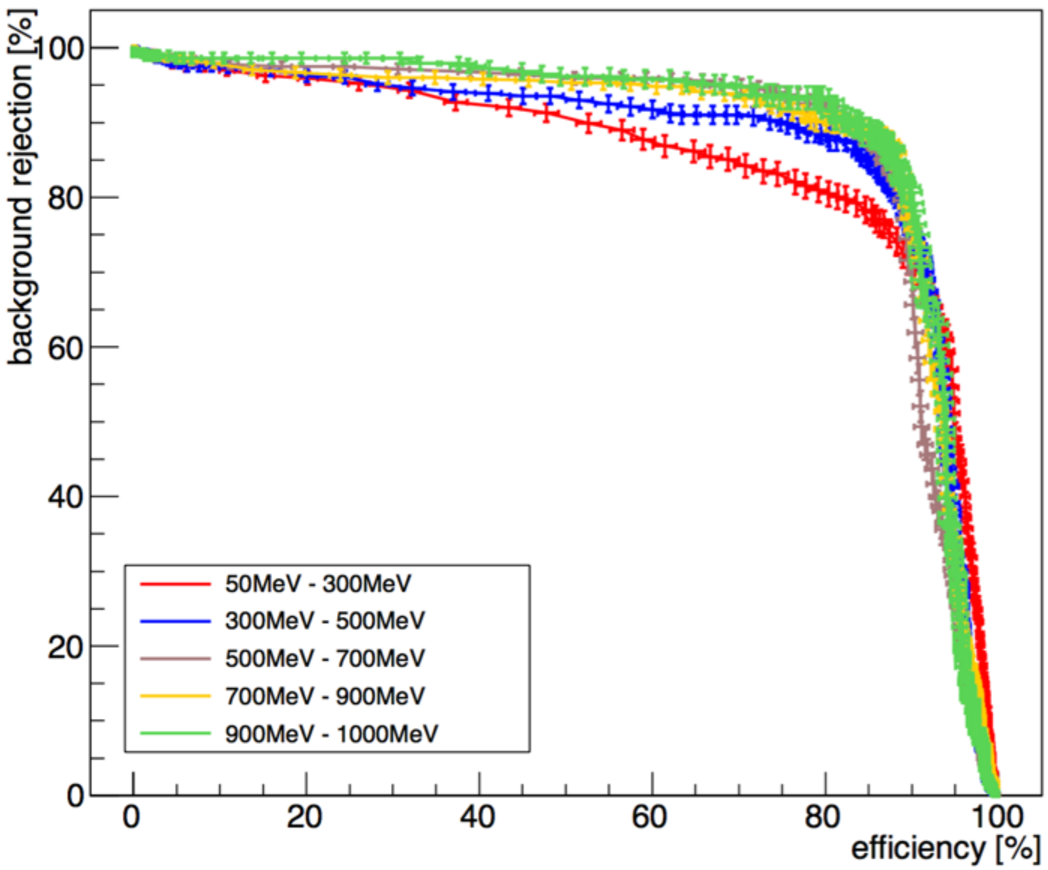


Figure 2.10: Peformance of the $e-\gamma$ separation algorithm using the ionization density of the initial part of the EM shower, as functions of the energy of the electron or photon. Only events with well-reconstructed initial portions of the shower are included (see text).

D. Stefan

DUNE
CDR: Annex 4C

90% background rejection, just from dE/dx

We expect to get another factor of 10 from topology

Event Classification

- Another example – events with a good electron candidate, but also a π^0 candidate.
  - Rejecting events with π^0 candidates in them reduces signal efficiency.
  - Keeping them lets in more backgrounds and poorer energy resolution (shower overlaps and combinatoric mistakes)
  - Put these events in a separate category!
  - Three showers vs. two. Suppose we only get one leg of the π^0 but the other looks very electron-like. Give it the right weight!
- Other examples –
  - Energy resolution of partially contained events
  - Particle ID of exiting or interacting particles
  - Particle ID in a dense environment
    - Collider analyses use lepton isolation in order to purify samples.
    - Neutrino experiments are desperate for every event

Event Classification

- ν_μ **CC**
  1. 1-track 0π (μ^- only)
  2. 2-track 0π (μ^- + nucleon)
  3. N-track 0π (μ^- + (>1) nucleons)
  4. 3-track Δ-enhanced (μ^- + π^+ + p, with $W_{reco} \approx 1.2$ GeV)
  5. $1\pi^\pm$ (μ^- + $1\pi^\pm$ + X)
  6. $1\pi^0$ (μ^- + $1\pi^0$ + X)
  7. $1\pi^\pm + 1\pi^0$ (μ^- + $1\pi^\pm$ + $1\pi^0$ + X)
  8. Other
- **Wrong-sign** ν_μ **CC**
  9. 0π (μ^+ + X)
  10. $1\pi^\pm$ (μ^+ + $\pi^\pm$ + X)
  11. $1\pi^0$ (μ^+ + π^0 + X)
  12. Other
- ν_e **CC**
  13. 0π (e^- + X)
  14. $1\pi^\pm$ (e^- + $\pi^\pm$ + X)
  15. $1\pi^0$ (e^- + π^0 + X)
  16. Other
- **Wrong-sign** ν_e **CC**
  17. Inclusive
- **NC**
  18. 0π (nucleon(s))
  19. $1\pi^\pm$ ($\pi^\pm$ + X)
  20. $1\pi^0$ (π^0 + X)
  21. Other
- ν**-e**
  22. $\nu_e + e^-$ elastic
  23. Inverse muon decay $\bar{\nu}_e + e^- \rightarrow \mu^- + \bar{\nu}_\mu$ (including the annihilation channel $\nu_\mu + e^- \rightarrow \mu^- + \nu_e$)

Current set of event categories (reconstructible) envisaged for the DUNE ND

FD similar.

Additional Benefits of Classifying Events into Many Categories

- Events are given the most statistical power they can have
  - Simply adding events of different kinds together in E_{reco} histograms stirs together background and signal that could be separated.

- Mitigates the effect of high-significance events near cuts.

  A re-calibration can cause important events to disappear from a cut-based analysis
  A multi-category analysis is more likely to keep the events.

- You can split the work up among members of a large team without stepping on toes.

  - You need to define boundaries of who gets which events of course, and keep the communication channels open

  - Some categories may not diagonalize nicely onto theses or publications however. Example: "Loose" vs. "Tight" electron selection categories provide rather nonphysical boundaries for publications.

Additional Benefits of Classifying Events in Many Categories

- Some event samples are control regions for other kinds of events
  - Example: ν_μCC+1π^0 provides a control sample for ν_eCC+1π^0 events (lepton universality) and even NC events with a π^0, though the NC events may differ in their nuclear effects.
  - Generator prediction of the fraction of events in which π^0's are produced needs to be validated/tuned
  - Larger Near Detector samples can be used to tune MC for the FD.
  - A choice – include constraints from other samples as external priors or fit constraints, or include them in the fits (or Bayesian integrals) for your results.

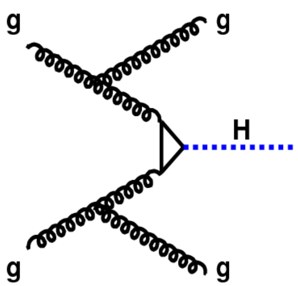
Pitfalls of Over-Classifying Data

Inclusive predictions almost always have smaller uncertainties that differential predictions.

Especially in the early days of an experiment when the theory is still being developed.

gg→H (+jets) with H→WW→lvlv

Anastasiou, Dissertori, Grazzini, Stöckli, and Webber JHEP 0908 (2009) 099



σ [fb]	LO (pdfs, α_s)	NLO (pdfs, α_s)	NNLO (pdfs, α_s)
0-jets	$3.452^{+7\%}_{-10\%}$	$2.883^{+4\%}_{-9\%}$	$2.707^{+5\%}_{-9\%}$
1-jet	$1.752^{+30\%}_{-26\%}$	$1.280^{+24\%}_{-23\%}$	$1.165^{+24\%}_{-22\%}$
$\geq$ 2-jets	$0.336^{+91\%}_{-44\%}$	$0.221^{+81\%}_{-42\%}$	$0.196^{+78\%}_{-41\%}$

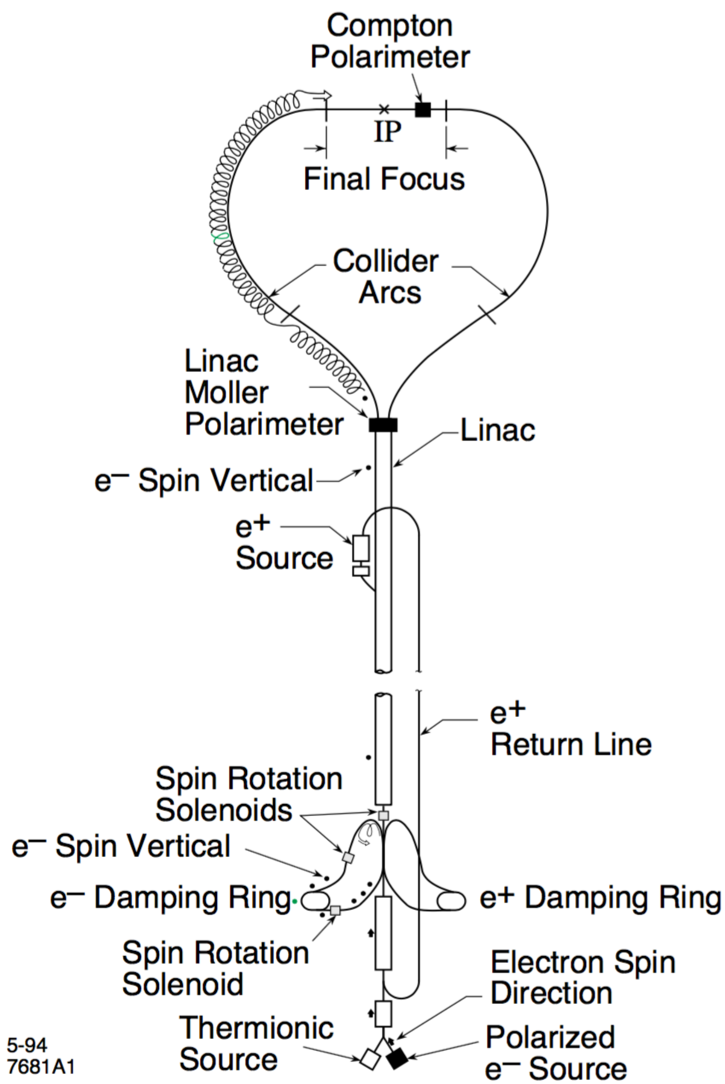
Table 2: Inclusive cross sections in the different jet bins.

m_H=160 GeV at the Tevatron

WWjj+X has a bigger background – ttbar! So experiments had to classify their data by #jets
ttbar background is very big at the LHC!
Needed good predictions for 0 and 1 jets, but only had inclusive predictions in 2009, and they weren't that high order.

Theory has since improved, necessary to get Higgs physics out of the LHC.
DIS Neutrino interactions may have a similar difficulty with pion production instead of jets.

Another Example: Compton and Møller Polarimetry at SLC



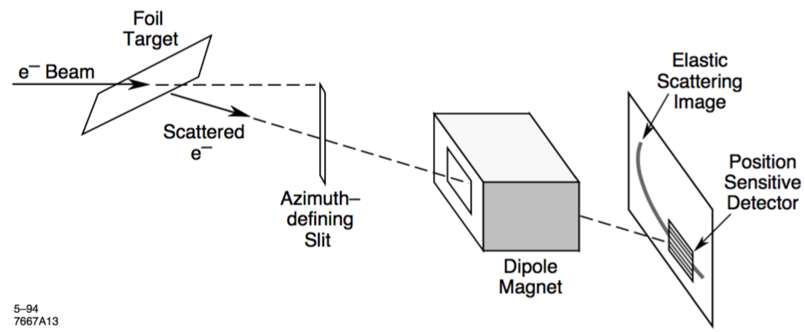
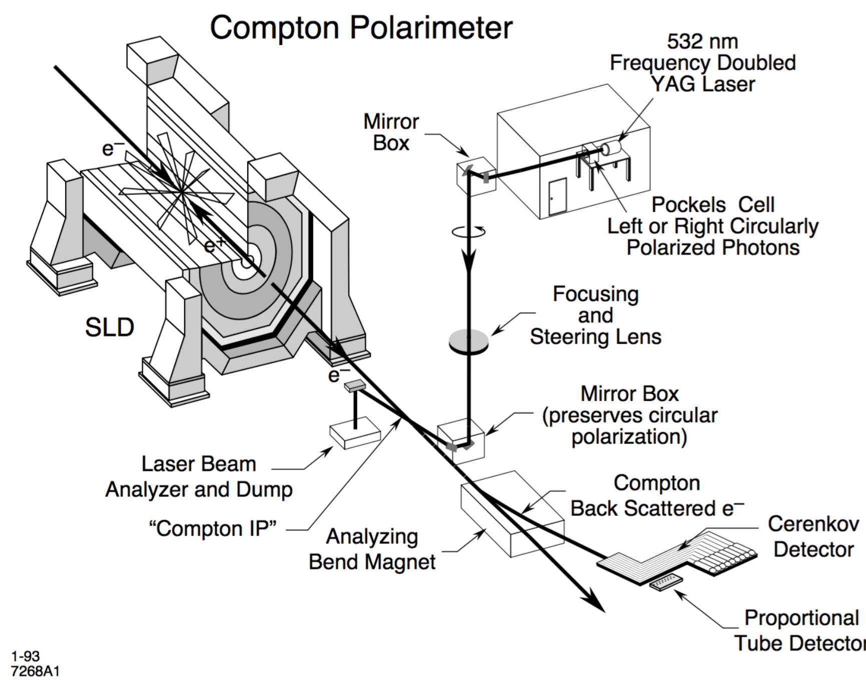


Figure 1. A schematic diagram of a single-arm Møller polarimeter.



Morris Swartz, Nucl.Instrum.Meth. A **363** (1995) 526-537

Møller Polarimetry at SLC and the Levchuk Effect

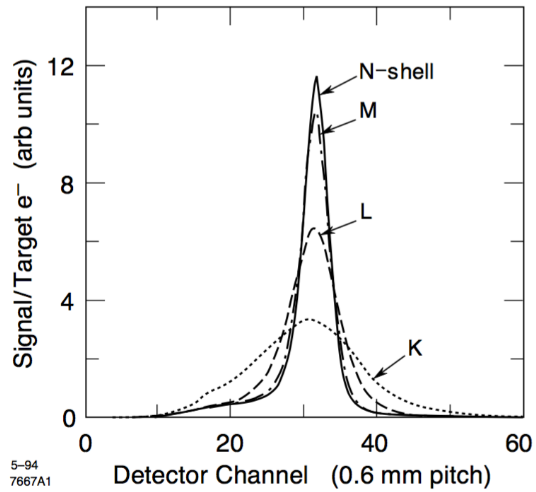
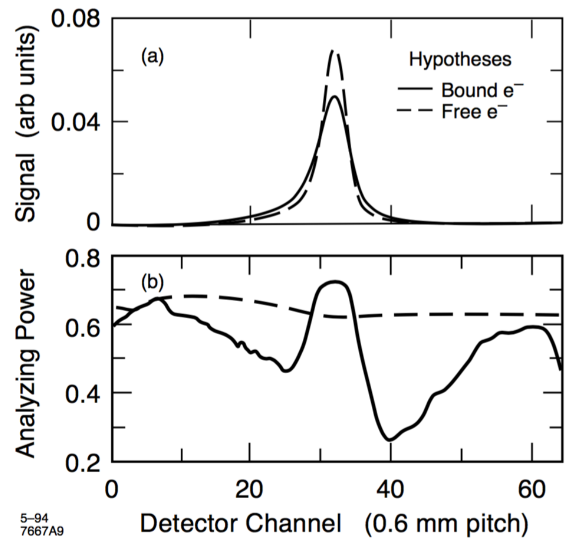
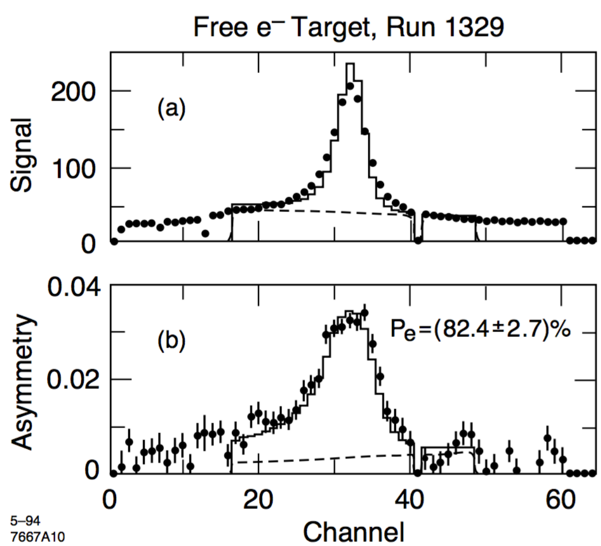
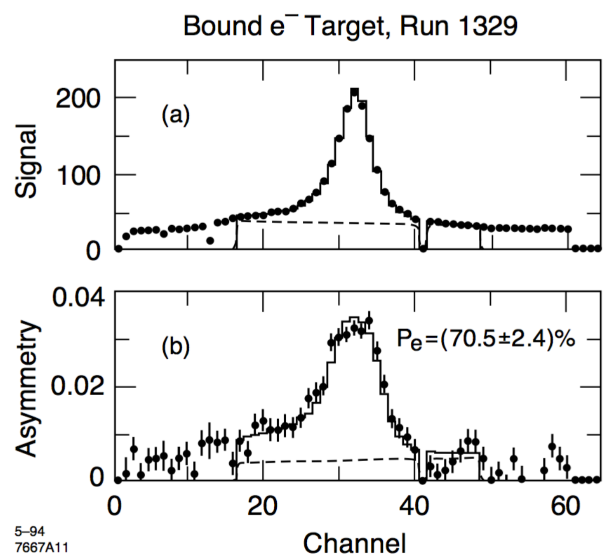


Figure 2. The simulated signal observed at the SLC Linac Møller detector per target electron for each of the atomic iron shells.







Lesson: Fit the wrong function, get the wrong answer!

Morris Swartz, Nucl.Instrum.Meth. A **363** (1995) 526-537

L.G. Levchuk, *Nucl. Instrum. Meth.* **A345**, 496 (1994).

Traditional MVA Techniques

- Uses: particle classification, event classification.

- If you're just trying to measure the fraction a process contributes to a mixture (signal + background), then the MVA output values (event scores) can be directly used in the statistical analysis. Each bin of each histogram is statistically independent of the others (but shares systematics of course)

- We're interested in the energy distribution of course.

- Boosted Decision Trees – See implementation in TMVA
- Random Forests
- Neural Networks – Classic, easy to train. Hard to pick input variables. Hard to justify network topology a priori.
- Genetic Algorithms

- Training a MVA may not optimize what we care about! We care about:
  - Expected p-value for ruling out $\delta_{\mathrm{CP}} = 0$.
  - Expected uncertainty on the measurement of δ_{CP}
  - Expected p-value for excluding the incorrect mass ordering
  - Expected p-value for discovering proton decay

- Instead, NN training typically minimizes $E = \sum(desired \;\; output - obtained \; output)^2$
- Boosted Decision Trees typically maximize the Gini Score in each leaf node.

- The expected p-values and uncertainties however are very expensive to compute, whereas the contribution of a training event to the NN weights is easy to back-propagate

Choosing Input Variables

- Lots of experience in collider analyses
- Variables usually chosen by hand to capitalize on signal characteristics
- But we really want to separate signals from backgrounds! And the background may have more distinctive characteristics than the signal. So some strong variables are background oriented.
- Sequential NN's trained to identify individual processes
- Analyses seeking rare processes usually have gross selection first, and then train the NN on the events passing the selection -- no need to make the NN work hard on events that are otherwise easy to classify. BDT's automate this process
- Some MVA methods perform less well when an irrelevant input variable is added to the list. BDT's are (mostly) immune to this.

Different Ways to Rank MVA Input Variables

"Importance" of input variables is not a well-defined idea. One can rank input variables in a variety of ways

- Correlation to target: Make a 2D plot of the variable vs. the desired NN output for the training sample. Calculate the coefficient of linear correlation.
  - Not too effective if you're looking for a rare process and the variable is great but only for a tiny fraction of the events. Good variable, but correlation to target may be skewed by the vast majority of don't-care events.
- Sensitivity with just that variable
  - Not too effective if variables work together synergistically
- N-1 sensitivity: Knock out the variable and re-compute the sensitivity
- Sum of squares of weights on the input layer of a NN (if applicable).

Genetic Optimization of Analysis Methods

Figures of Merit that we care about are often expensive to compute.

- Range of δ_{CP} with a minimum expected significance
- Expected significance of Mass Ordering sensitivity, assuming any value of other oscillation parameters

The space of optimizable parameters is often (very!) high-dimensional.
We'd like to explore it all but do not have the time.

- A population of MVA discriminant functions can be created randomly, or with some initial guidance from the analyzer (we're not dogmatic!)
- Figure of merit is computed for each member of the population.
- Next generation of MVA discriminant functions is generated from the previous one by selecting high-performing functions and varying features by a bit, and exchanging features between pairs.
- Features of a NN can include weights and also the structure (# layers, inputs, outputs, and unusual connections)
- Process is iterated and best functions retained.

S. Whiteson and D. Whiteson, Engineering Applications of Artificial Intelligence **22**, 1203 (2009).

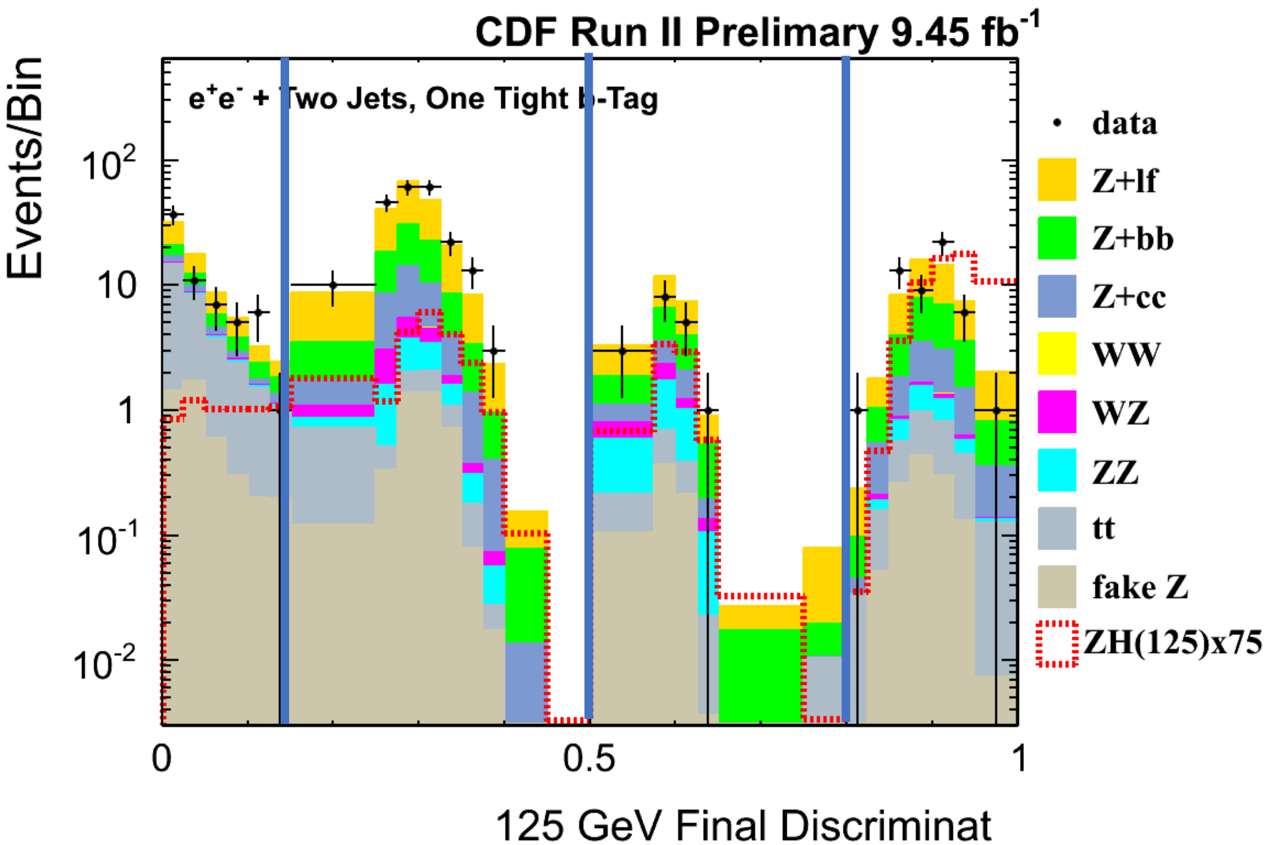
Software for NN training: *Kenneth O. Stanley & Risto Miikkulainen*
"Evolving Neural Networks Through Augmenting Topologies". Evolutionary Computation. ***10*** *(2): 99–127. (2002).*

A similar genetic algorithm is used to optimize the LBNF target, horn and decay pipe (L. Fields).

An Example of Cascaded Neural Networks: CDF ZH→llbb Search

Multiple sources of background, all larger than the signal
Analysis proceeds by a

- Preselection
- anti-ttbar Discriminant. What's left after a cut on that goes through
  - anti-Z+jets discriminant. " " " " "
    - anti-Diboson discriminant " " " "
      - Signal Discriminant trained on what's left



Pass all four distributions to the fitter for nuisance parameters and signal rate

This method is great if you want to purify a small sample of rare events with several sources of large, uncertain backgrounds (proton decay folks!)

Not so good if you want to do an inclusive analysis!
The properties of the selected events are highly sculpted!

What you get is what you set out to find.

Uncertainties on Reconstructed Quantities

- Frequently a track fit will also produce an error matrix. Also cluster energies can be provided with uncertainties.
- Cores and tails of distributions are frequently not represented well by measurement uncertainty estimates.
- A naive inclusion of the measurement uncertainty in a likelihood function can cause fit biases.
- Giovannni Punzi, "Comments on Likelihood Fits with Variable Resolutions", in the Proceedings of PHYSTAT2003, SLAC-R-703, available at http://www.slac.stanford.edu/econf/C030908/
- See Louis Lyons's description of the Punzi Effect https://indico.cern.ch/event/431039/
- Measurement uncertainties are really just additional observables, which can be used to help us classify events.
- Can be used in kinematic fits in order to compute other observables, but use only in order to place events in categories or bins whose contents are predicted using MC or data-driven models

Filling Histograms with Distributions for Each Event

The "ideogram" method. Or "kernel smearing".

Contents of each bin are no longer Poisson.

Results are highly correlated from one bin to the next.

Need a MC model to predict the shapes of such distributions – not easy to compute likelihoods for extracting model parameters

Not very popular in HEP, but it has been used.

Matrix-Element Techniques

For a particular physics process indexed i, the probability to measure observables $\boldsymbol{y}$ is

$$P_i(\boldsymbol{y},\boldsymbol{\theta}) = \frac{1}{\sigma_i(\boldsymbol{\theta})A_i(\boldsymbol{\theta})}\int_{X_i} W_i(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})\,\epsilon_i(\boldsymbol{x},\boldsymbol{\theta})|M_i(\boldsymbol{x},\boldsymbol{\theta})|^2 T_i(\boldsymbol{x},\boldsymbol{\theta})d\boldsymbol{x}$$

Where $\boldsymbol{\theta}$ are the model parameters, both physics and detector
$\boldsymbol{x}$ are the true kinematic variables ($\boldsymbol{y}$ are measured)
$W_i(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})$ are transfer functions for observing $\boldsymbol{y}$ given a true $\boldsymbol{x}$ (Detector model)
$M_i(\boldsymbol{x}, \boldsymbol{\theta})$ are quantum-mechanical matrix elements (theoretical)
$\epsilon_i(\boldsymbol{x}, \boldsymbol{\theta})$ are the efficiencies to reconstruct, identify, and select events, and
$T_i(\boldsymbol{x}, \boldsymbol{\theta})$ are factors that do not depend on $\boldsymbol{y}$, such as beam flux, parton density functions, oscillation parameters

The observables $\boldsymbol{y}$ are usually things like reconstructed lepton momenta, jets, etc. Not hits or ADC values.

See I. Volobouev, arXiv:1101.2259 for a good introduction.

Matrix-Element Techniques

Uses of $P_i(\boldsymbol{y},\ \boldsymbol{\theta})$:

- You can use Bayes' Theorem to derive $P_i(\boldsymbol{\theta},\boldsymbol{y})$ and extract the physics parameters $\boldsymbol{\theta}$. Used in top-quark mass measurements at D0 and CDF
- You can use it as an input to a MVA method to separate classes of events. A likelihood ratio of the P_i's is optimal if all the information is in the kinematics of $\boldsymbol{y}$

$$L_{\text{signal}} = \frac{P_{\text{signal}}(\boldsymbol{y},\boldsymbol{\theta})}{\sum_i P_i(\boldsymbol{y},\boldsymbol{\theta})}$$

- Discriminant variables are just functions of the observed $\boldsymbol{y}$. Events are still just counts in histograms.
- Example use in neutrino experiments: The momentum and angular distribution for photons from π^0 decay in NC events won't be the same as electrons in $\boldsymbol{\nu}_e$CC events, and other kinematic quantities such as missing p_T will be correlated.

Matrix-Element Techniques

- Collider use found that BDT's and NN's performed as well as or better than matrix-element
  techniques. Often particle-ID variables are strong inputs too, and sometimes these are put in separately assuming they factorize from kinematics

- See for example CDF's single-top
  observation: T. Aaltonen *et al.*, Phys.Rev. D **82** (2010) 112005.   The *P*'s are as before, and *b* is a b-tag discrimiant

$$EPD = \frac{b \cdot P_{\rm s}}{b \cdot (P_{\rm s} + P_{Wb\bar{b}} + P_{t\bar{t}}) + (1-b) \cdot (P_{Wc\bar{c}} + P_{Wcg} + P_{Wgg})}$$

- These methods may not be fully optimal,but they
  can be an important part of a larger effort.

Matrix elements do not have to be correct, they only have to be useful.

Collider people frequently used Leading-Order matrix elements to separate signals from backgrounds.
Need to calibrate performance with data or with a full MC however, using Feldman-Cousins or other techniques.

In a neutrino experiment, Signal and Background matrix elements may formally be nearly the same!
But backgrounds often involve misidentified particles.

MVA Techniques for an Oscillation Experiment

An issue: The efficiency as a function of neutrino energy almost certainly won't be a uniform function.

Higher-energy events are often easier to classify than lower-energy ones.

Lower-energy events carry more oscillation information.

Mismodeled sculpting of the reconstructed energy distribution by a cut on an MVA function can fake a different value of the oscillation parameters.

Convolutional Neural Networks

A great, new idea – getting traction in HEP.

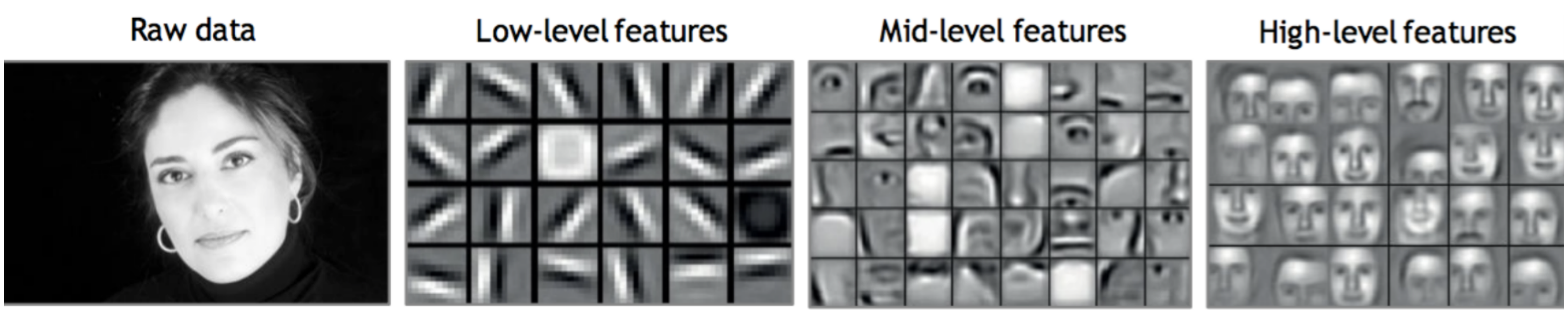
Useful in image analysis: Networks analyze patches of an image, and are scanned over the entire image, one bit at a time (the "convolution").

Don't need to process raw data to find tracks and showers: just feed raw (or better yet, noise-filtered, electronics-deconvoluted) data in as image data.

"shift-invariant" neural network – if the detector is homogeneous, the image processing should respond to a neutrino scatter anywhere in the detector the same way.

Instead of training a weight for every input pixel, try learning weights that describe kernel operations, convolving that kernel across the entire image to exaggerate useful features. Inspired by research showing that cells in the visual cortex are only responsive to small portions of the visual field.

Slide from A. Radovic

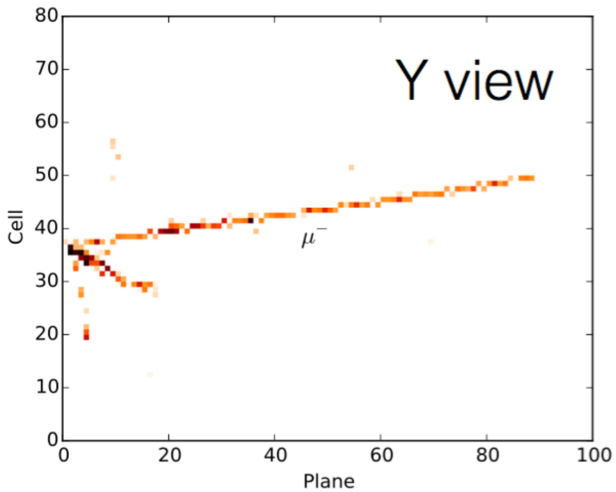


https://developer.nvidia.com/deep-learning-courses

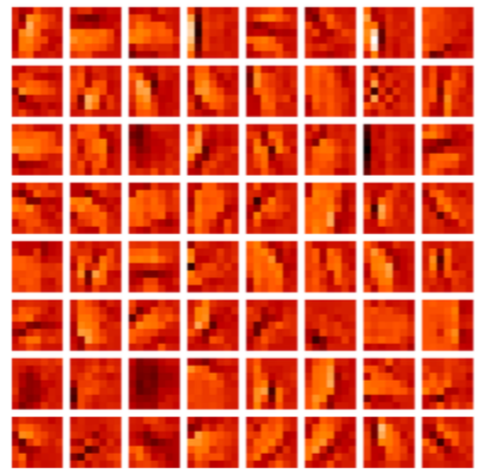
Example from NOvA

A. Radovic

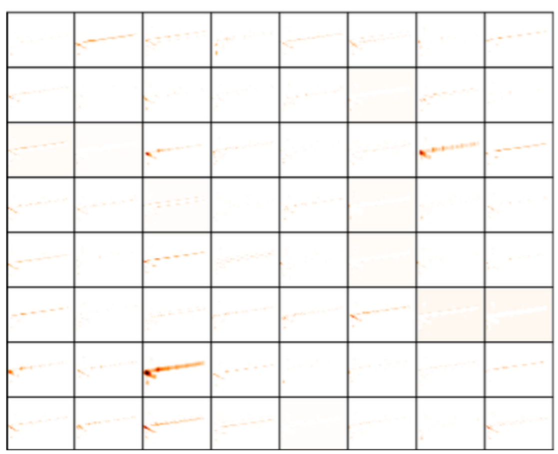
Example CVN Kernels In Action: First Convolution



X



=



Here the earliest convolutional layer in the network starts by pulling out primitive shapes and lines.

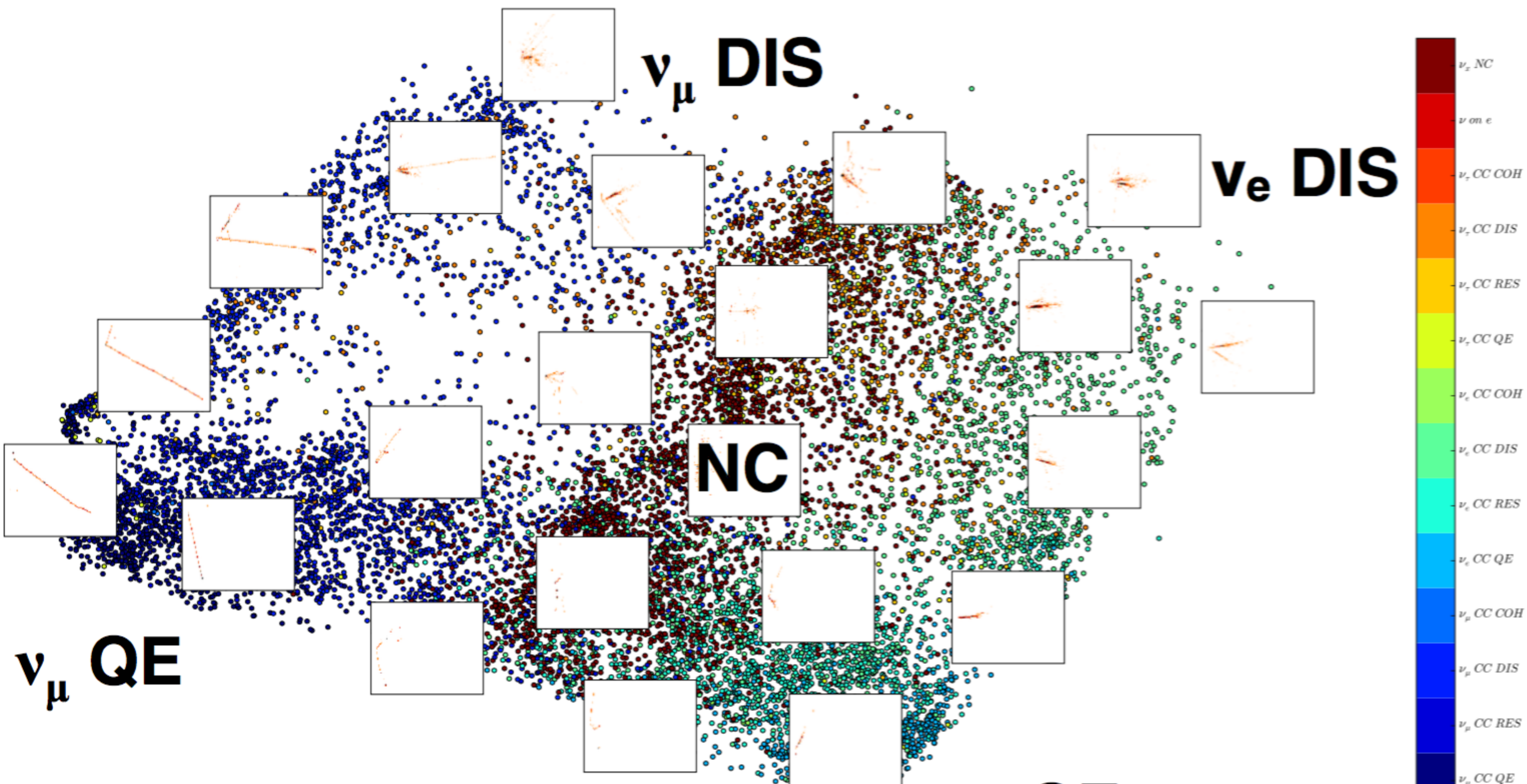
Already “showers” and “tracks” are starting to form.

34

Full Network based on GoogleNet (2014).

Example from NOvA

t-SNE Representation of Test Sample

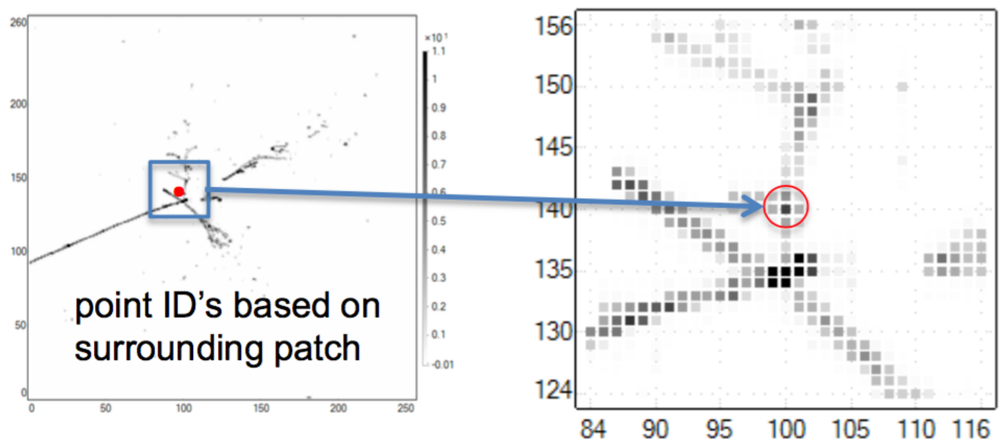


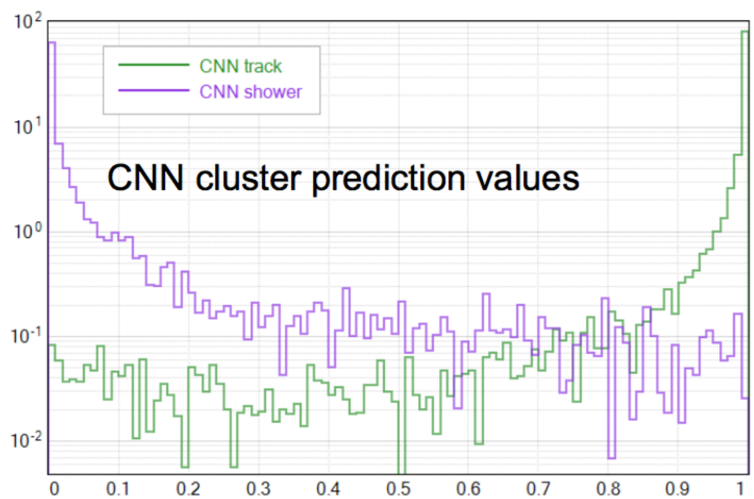
t-SNE projection of final features to 2D.
Truth labels, training sample subset.

Convolutional Neural Network Application in DUNE – Separating EM-like from Track-Like Pixels One at a Time

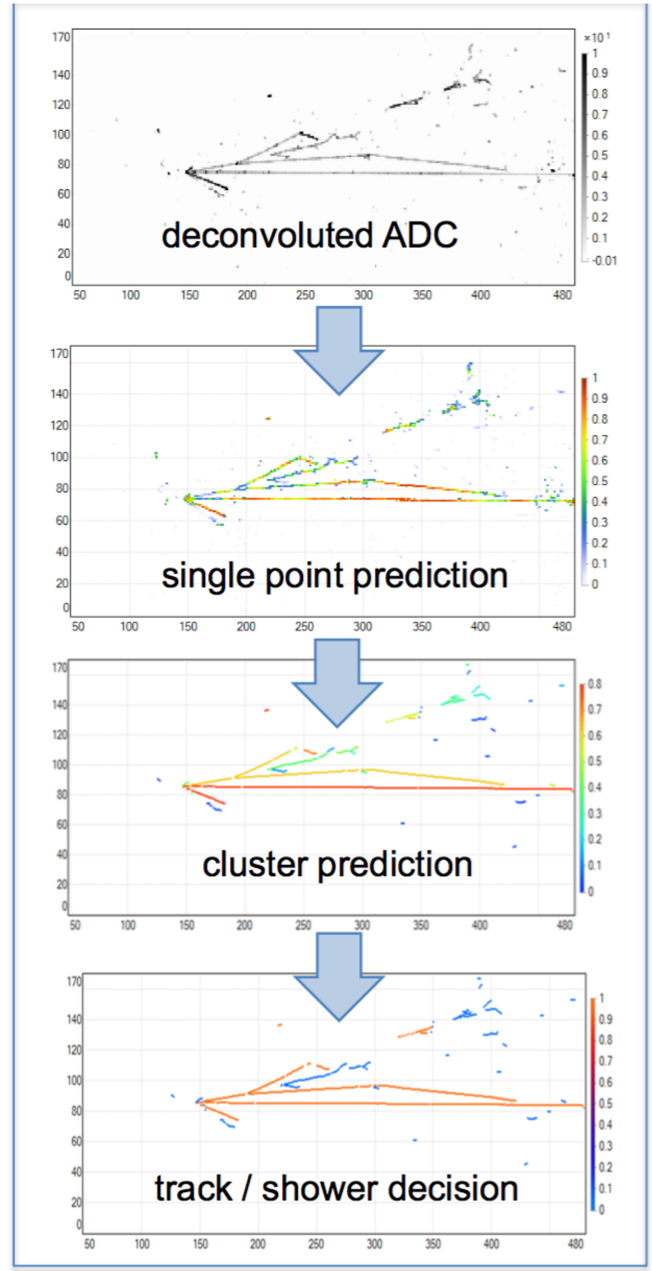
R. Sulej
D. Stefan
P. Płoński

EM-like / track-like identification flow



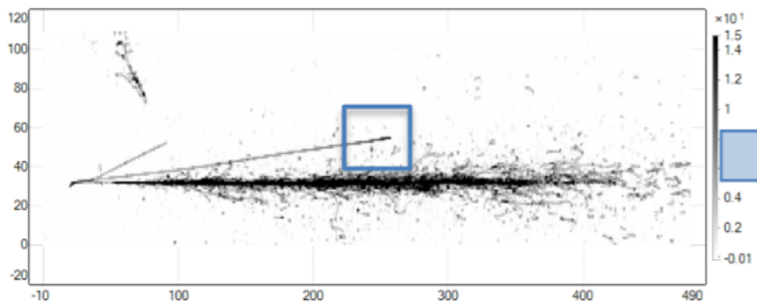


90.0% track / 90.0% EM correct cluster ID rate
(2GeV/c π^+ in ProtoDUNE)

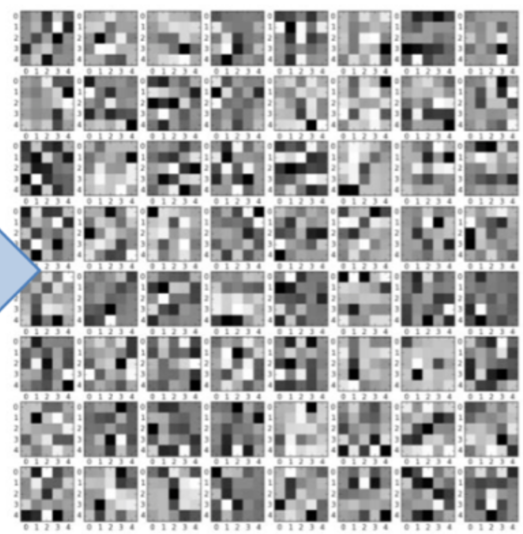


Convolutional layers

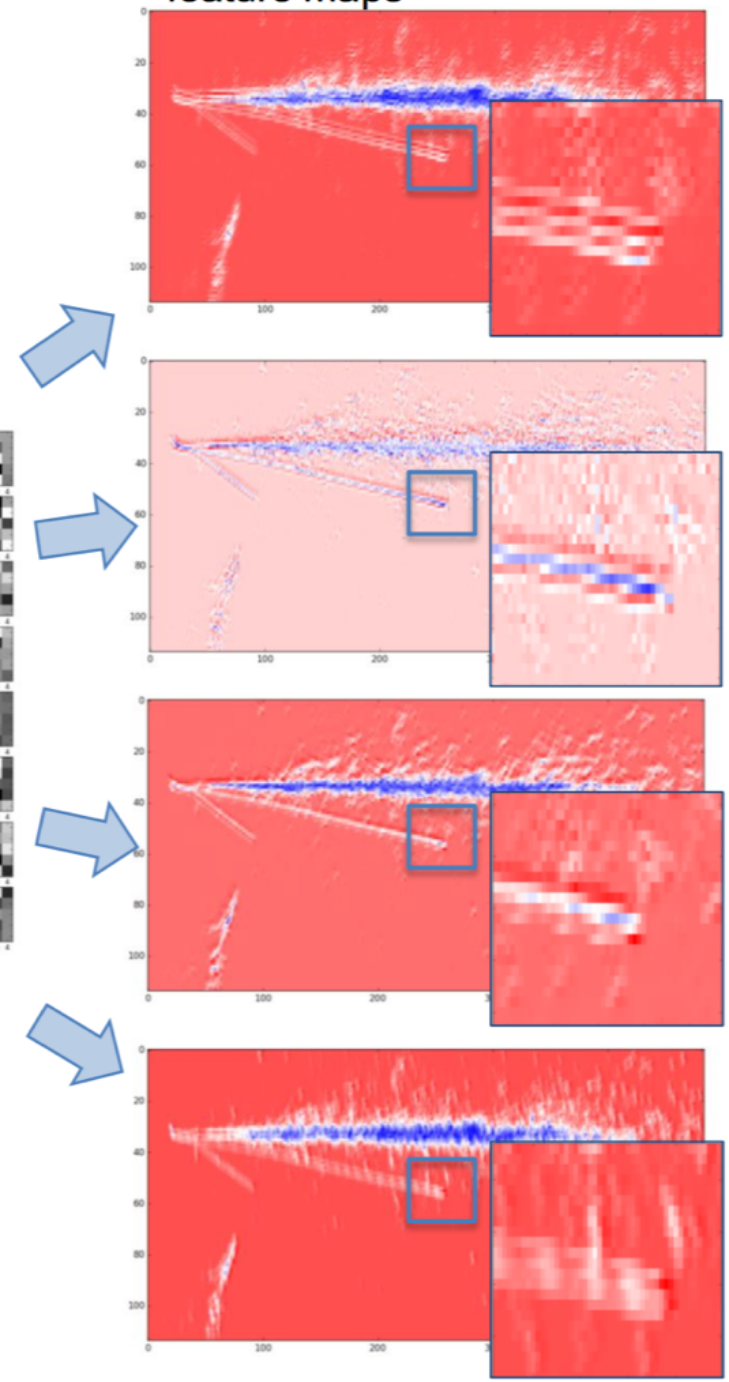
→ Sets of filters optimized during the training to act as a feature detectors



2D kernels / filters



feature maps



→ Output: feature maps, input to the standard (*dense*) MLP layers

→ Can think of this like a Library Matching of event parts

Training using Data

At CDF, the Monte Carlo models were not trusted to simulate misreconstructed objects reliably.

- lepton ID misassignment ("fakes")
- b-tag mistakes ("mistags")

Predictions are data-driven, or use a mixture of data and MC to predict background rates.

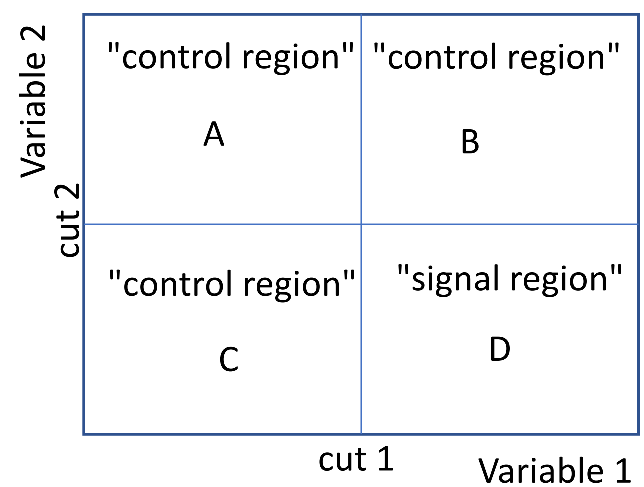
But how to train the MVA's?

- Invert a selection requirement to purify a sample in a chosen background and use those events.
- Must be careful not to use the same data events to train as to predict the backgrounds.
- Dataset size is often the dominant limitation.
- Events failing an selection requirement may be biased in other ways too.
  - Model biases using Monte Carlo and correct for them in the modeling of the background, OR
  - Invert another (hopefully independent!) selection requirement. Use the other two regions to measure the bias in the data ("ABCD" method)
  - Used in the modeling step, not necessarily for the training step

- Topic has come up in discussions of how to train deep-learning discriminants when real data are preferred over simulations.

- The problem of limited training example size may be worse than the problemcof training with a biased sample.

- No such thing as incorrect discriminants, just incorrect models.



Many Analyses All Doing the Same Thing

What if they produce different outcomes?

How best to divide the work among many collaborators.

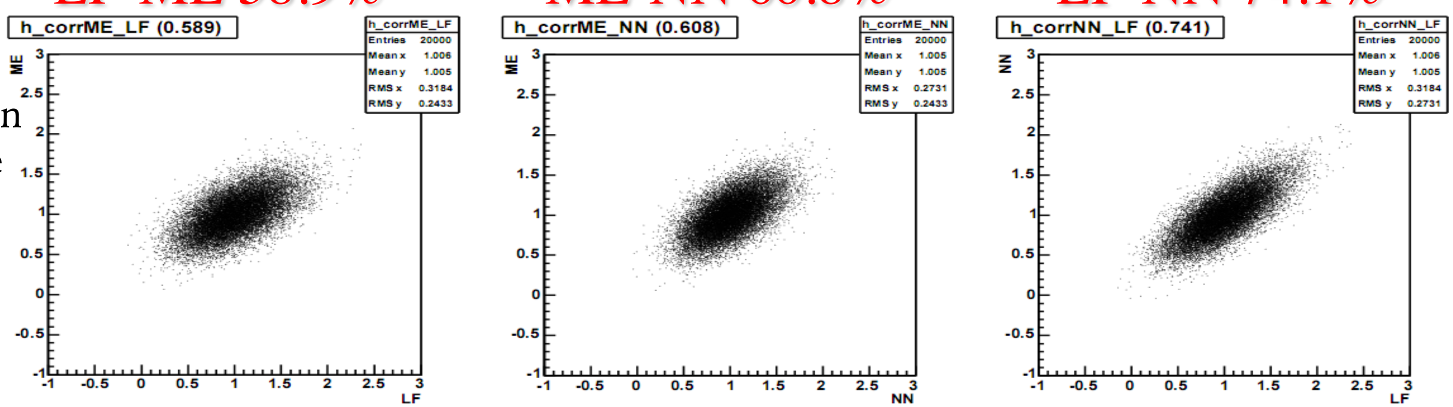
A common strategy: divide the event samples into non-overlapping pieces

A more difficult strategy: have many analysis teams select exactly the same events but apply different techniques to them. Example: single top combinations.

Uncoordinated teams usually select highly overlapping event samples but not exactly the same.
- Overlapping events in data are usually easy enough to identify and count
- Overlapping events in shared MC samples also aren't too hard
- But some predictions are mixtures of data and MC, are not individual events, or are extrapolations, and thus overlap is not the right word. Statistical correlation needs to be evaluated in repeated simulated experiments. Sampling finite event MC's with replacement is usually fine.

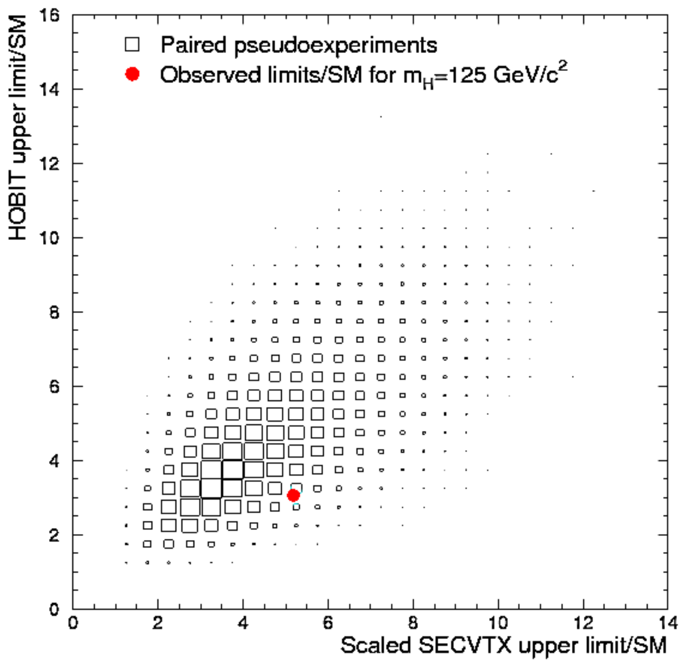
Example: CDF single-top cross section measurements. Three analysis groups, same events.

LF-ME 58.9% ME-NN 60.8% LF-NN 74.1%



Upgrading An Analysis Without Using Fresh Data

- A common situation at the end of an experiment – no opportunity to take more data.
- Also common if many groups want to analyze the same data to infer the same parameters but using different techniques.
- A cause of endless acrimony – especially if not blind.
- Example: T. Aaltonen *et al.*, Phys.Rev. D **87** (2013) 052008 -- updated Higgs boson search at CDF in the METbb channel, using a new b-tag.



Assumes SM Higgs boson present at m_H=125 GeV.

People will always want a *p*-value for the discrepancy between the results.

What is the test statistic? The ratio of the two results? The difference? The absolute value of the difference? How about the probability of the difference or bigger assuming one of the values?
What true value is assumed when generating pseudoexperiments?

If you must choose a method, how do you do it?

- Best expected performance, of course! Ignore the data result when looking at performance.
- But suppose an analysis team has many possible improvements, all of which improve the expected performance. It is still possible to cherry-pick those that give a preferred result in the data.
  More of a higher-order effect, but people work very hard.

An Example from NOvA: R. Patterson, FNAL JETP Seminar, Aug 6, 2016

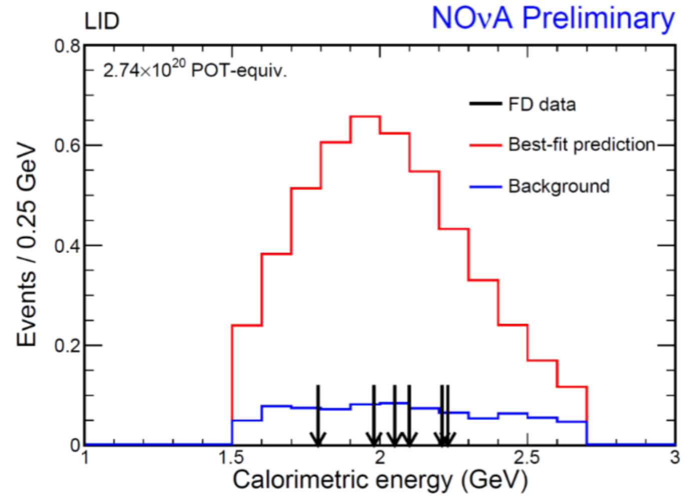
Both selectors have ~1 background event and ~6 signal events for a particular oscillation hypothesis. Sensitivities were indistinguishable. Thus, prior to unblinding, decided to **show both results** and to use the more traditional **LID technique** as the primary result where required.

Far Detector selected events

Likelihood ID

LID: 6 ν_e candidates

3.3σ significance for ν_e appearance



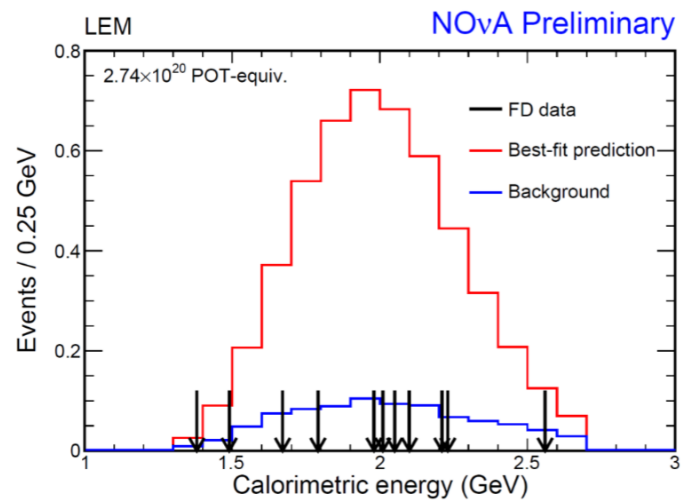
At right:
Calorimetric energy

Library Event Matching

LEM: 11 ν_e candidates

5.5σ significance for ν_e appearance

(All 6 LID events present in LEM set)



Ryan Patterson, Caltech 55 Fermilab JETP, August 6, 2015

Note: significances here are computed comparing data with background only (1 event). Both outcomes are much more likely if there is signal, and the significance of the difference is less.

Control Over Systematic Uncertainty

The traditional concern: underestimated uncertainties mean the result is not valid within its error bars

But: overestimated uncertainties can be just as bad in a subtle way:

Nuisance parameters are constrained in the oscillation fits.
A measured sample with an overestimated systematic uncertainty on its predictions ends up claiming it is more sensitive than it really is to a particular nuisance parameter. Fitting that constrains it too hard for use in other samples.

Missing uncertainties (unknown unknowns)

One nuisance parameter covering for another's misprediction -- observed discrepancy between data and MC may be misascribed to the wrong nuisance parameter.

MVA's: check inputs *and* outputs in control samples and in the signal sample

- Theoretical predictions, especially those that depend on the hadronic recoil (or lack thereof), are highly uncertain.
- Standard MVA techniques are really just functions of observables to compute more observable quantities that are more optimal and are thus not uncertain
- But they can amplify mismodeling, and thus uncertainties need to be propagated through standard MVA functions.

Using Data Control Samples to Model Difficult Backgrounds

ν_μCC events provide a fertile ground to study nuclear effects and hadronic recoil systems

Lepton universality (almost! The mass difference between e and μ has an impact on the cross sections. See for example Day and McFarland Phys.Rev. D **86** (2012) 053003))

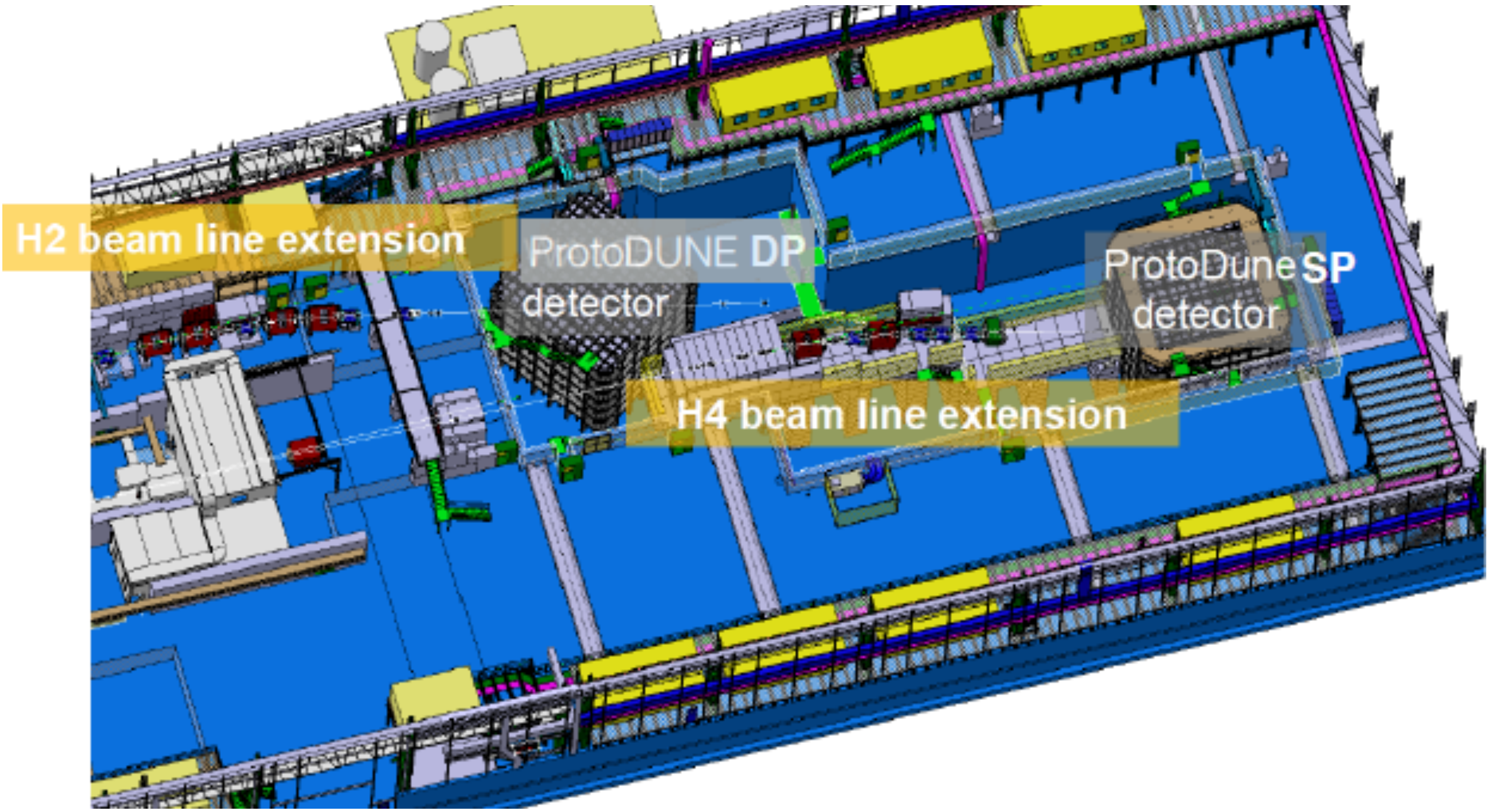
Neutral-Current event candidates on the other sides of cuts provide control samples for NC background rate measurements.

Rare-event searches like proton decays will have plenty of backgrounds from cosmics and atmospheric neutrinos – invert cuts and extrapolate measurements in to the signal samples.

A problem if the control samples have fewer events than the signal sample.

Prototyping and Test Beams

Single-Phase and Dual-Phase 6x6x6 meter detectors in charged-particle beams at the Prevessin site at CERN
Scheduled to run before LHC's Long Shutdown 2 starts (end 2018).
Collect about 2.5 PBytes of raw data each



Beam is mostly pions, with some protons, electrons, muons, and kaons. Energies from 0.5--10 GeV

Test-Beam Measurements

Constraints on physics quantities

- Electromagnetic Energy Scale
- Muon Energy Calibration -- range and multiple-scattering
- Hadron energy scale
- e/γ Separation
- π^0 identification and energy measurement
- Cross sections of charged pions and kaons on argon
- EM fraction in hadronic showers
- Recombination Angular Dependence
- Dependence on electric field of
  - Drift velocity
  - Recombination
  - Light Yield
  - Diffusion
  - Space-Charge Distortion

Some of these measurements may be the ultimate calibrations of these quantities for DUNE.

Detector Performance Measurements:

- Diffusion
- Electron Lifetime
- Detector Uniformity
- Detector Alignment
- Performance in gaps between APA's and for tracks crossing through APA's

Calibration of Reconstruction Performance

Energy Scale
Energy Resolution
Efficiency
Background

Common calibrations and systematics can be shared as long as the definitions of the objects are the same and the environment in which they are used is the same.

Collider examples: Jet energy scale, lepton ID efficiencies, lepton energy scales, tracking efficiency

Shared Systematic Uncertainty Prescriptions

works well for isolated leptons, but non-isolated leptons (i.e. in jets) have many sources heavy flavor, fakes, and gauge boson decay.

Neutrino examples: EM energy scale, electron ID, pizero ID, muon ID.

Summary

We are still in the very early years of getting liquid argon TPC detectors to work

We can take advantage of previous experiments' lessons: ICARUS, ArgoNeuT, LArIAT, MicroBooNE

Test-stands and prototypes let us know what to watch out for and what the expected performance is

Test-beam experiments may be our most important calibration sources

Analysis techniques brought over from collider experiments and non-HEP work benefit from years of study

New deep-learning techniques look promising!

Our understanding of systematics will drive our sensitivity in the long run.

The goal is to turn as many systematic uncertainties into statistical uncertainties so more data can beat them down.

Extras

MC Challenges

Colliders: Sneak an exotic particle or other physics process into the MC and see if your collaborators can discover what it is.

Simpler challenges like Banff Challenge 2 test if physicists can find bumps in histograms (much more limited question)

https://www-cdf.fnal.gov/~trj/

I am unaware of this sort of thing for DUNE. All MC is either unoscillated (pure $\boldsymbol{\nu}_{\mu}$ or pure $\boldsymbol{\nu}_{e}$).

Can generate samples with hidden MC truth values for the physics of the events and for the oscillation parameters.

Maybe even distort the nuclear modeling spectra.

Tools of the Trade

Incomplete List

Neutrino Experiments with Beams

- T2K
- NOvA
- MINOS
- MINERvA
- MiniBooNE
- MicroBooNE
- ICARUS (CNGS and future FNAL)
- SBND (future)
- DUNE (future – the expt I work on)
- OPERA
- LVD
- T2HK (future)
- LSND
- ANL and BNL bubble chamber experiments

Neutrino Observatories without Beam

- SuperK
- HyperK (future)
- SNO
- AMANDA/IceCube/DeepCore/PINGU
- ANTARES
- NESTOR
- IMB
- INO (future)

Reactor Neutrnio Experiments

- Daya Bay
- RENO
- Double Chooz
- KamLAND
- JUNO (future)

Gallium-Enhanced Neutrino Experiments

- GALLEX/GNO
- SAGE
- ANNIE

Neutrinoless Double-Beta Decay Experiments

- NEMO
- CUORE
- EXO
- GERDA
- MAJORANA
- KamLAND-Zen

Matrix-Element Techniques

- Calculate probability density of an event resulting from a given process

Phase space factor: Integrate over unknown or poorly measured quantities

Parton distribution functions

$$P(p_l^\mu, p_{j1}^\mu, p_{j2}^\mu) = \frac{1}{\sigma} \int d\rho_{j1} d\rho_{j2} dp_\nu^z \sum_{comb} \phi_4 \, | M(p_i^\mu) |^2 \frac{f(q_1) f(q_2)}{|q_1| |q_2|} W_{jet}(E_{jet}, E_{part})$$

Inputs: lepton and jet 4-vectors – no other information needed!

Matrix element: Different for each process. Leading order, obtained from MadGraph

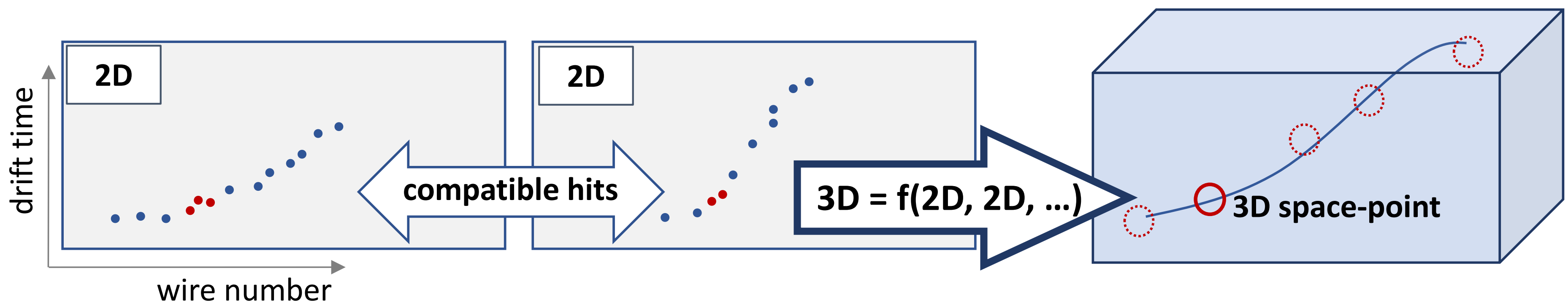
Transfer functions: Account for detector effects in measurement of jet energy

- The input variables are the same for all matrix elements – adding a new matrix element requires more calculation but does not use any different information from the data

Projection Matching Algorithm (PMA) 3D approach: what is different

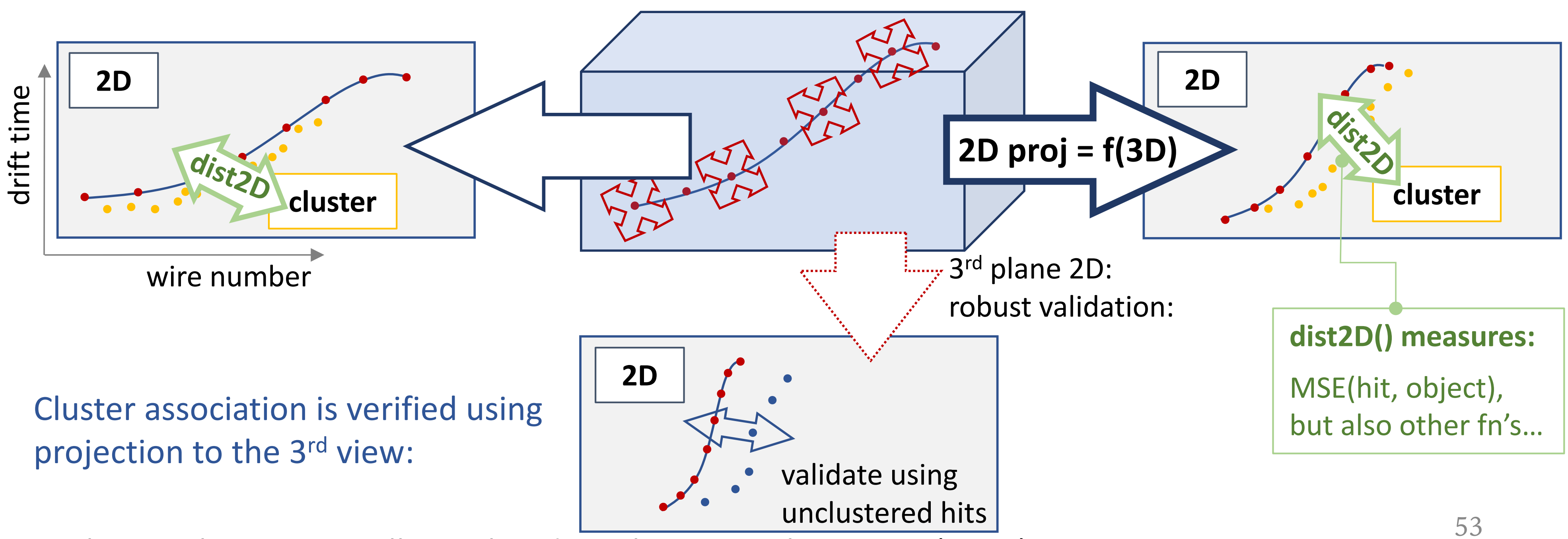
R. Sulej
D. Stefan

usual: *n* compatible 2D hits → 3D space-point → fit 3D trajectories to space-points



PMA: other way around: 3D object(s) ↔ hits or raw ADC in multiple 2D views

build 3D (***single track*** or ***full track structures***) to minimize distance the object's 2D projections to 2D hits



Cluster association is verified using projection to the 3rd view:

Earlier work: M. Antonello et al., Adv.High Energy Phys. 2013 (2013) 260820