# Data Preservation and Neutrinos

Mike Kordosky

William & Mary

May 17, 2011

Data Preservation Workshop
May 16-18,
Fermilab

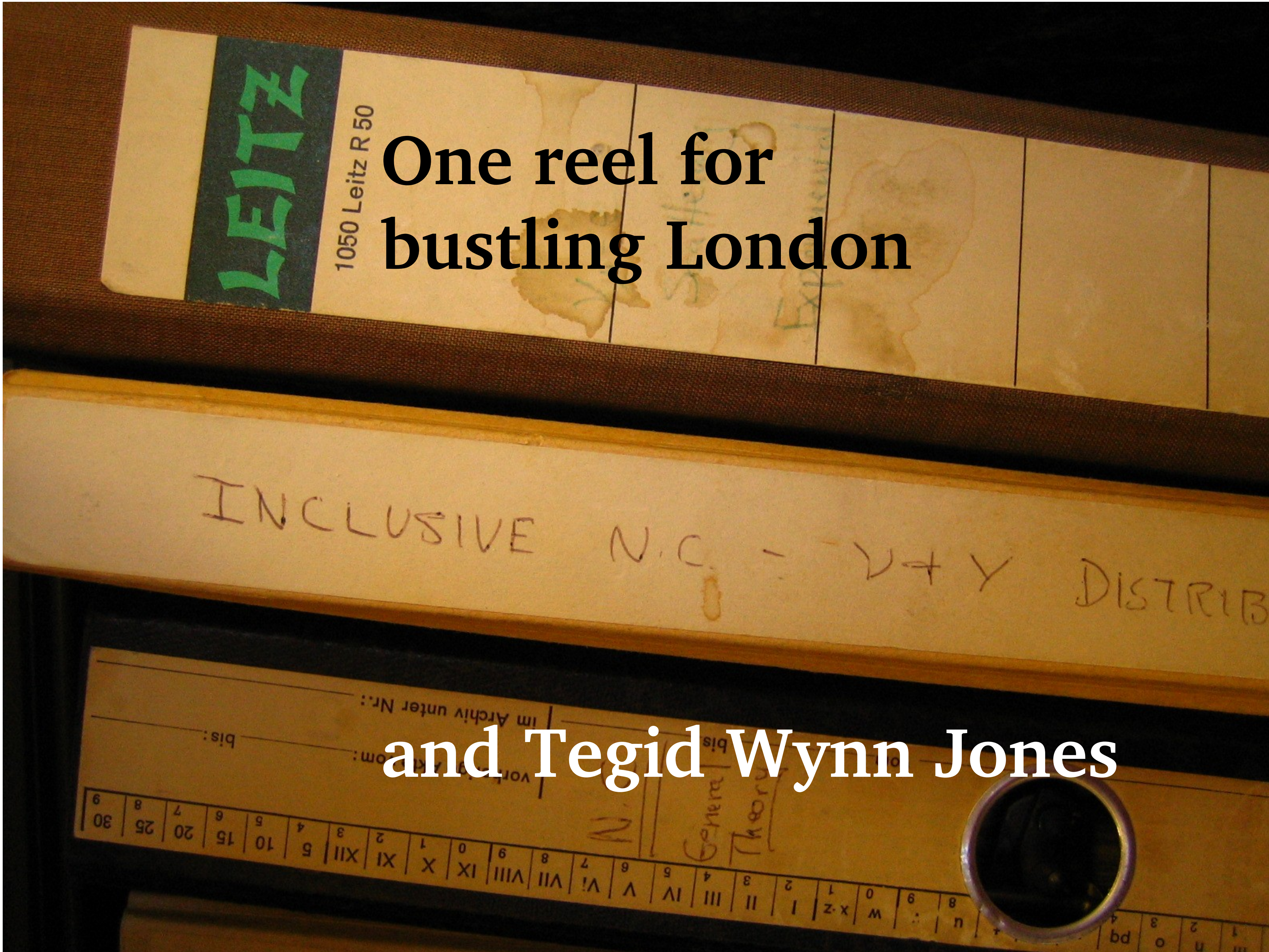

One reel for
ancient Aachen
and the famous king

# One reel for Don at Oxford


in his halls of stone

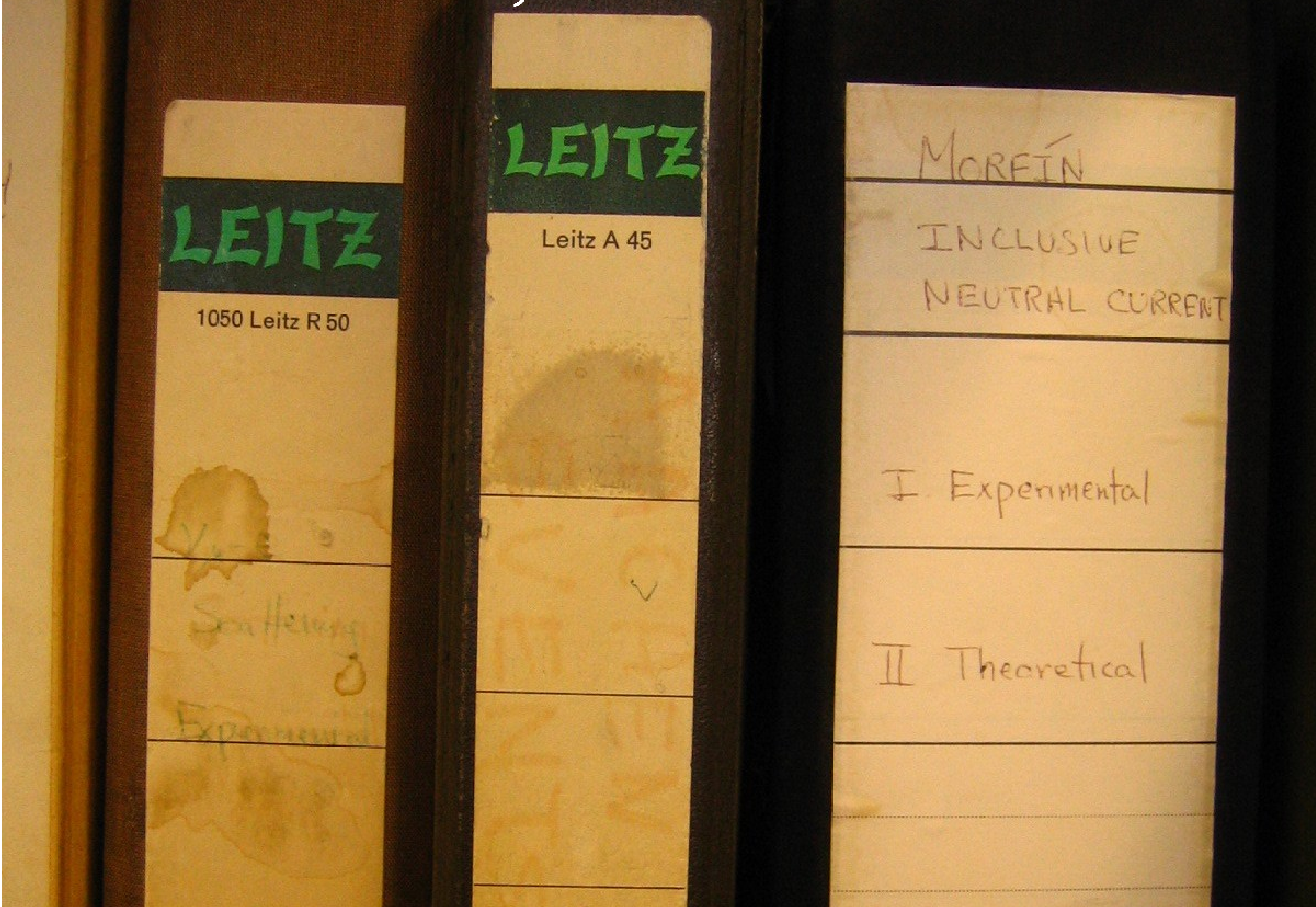
# One reel for bustling London

and Tegid Wynn Jones



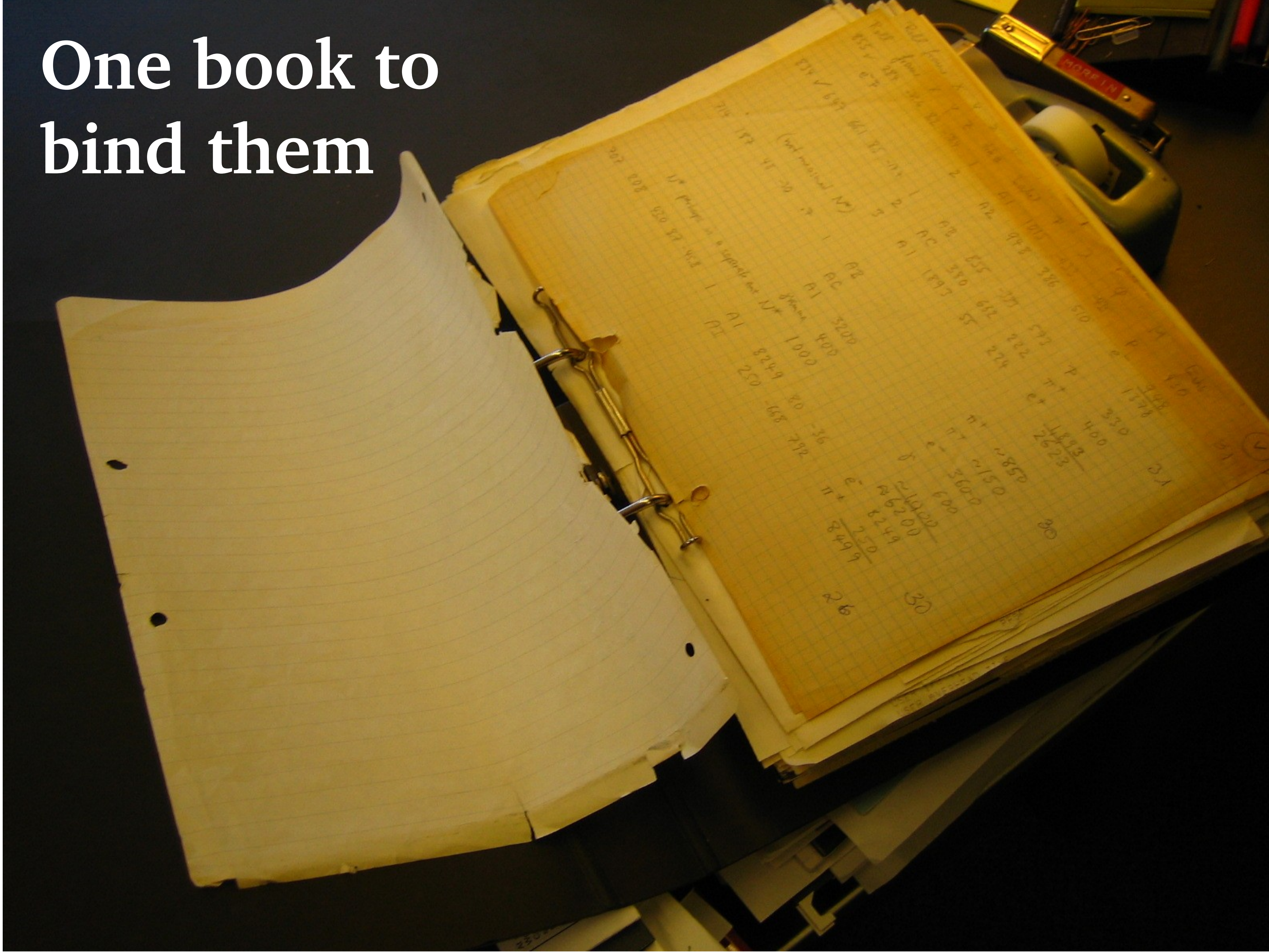

and one reel for the Italians
GARGAMELLE
USERS' HANDBOOK
CERN
down in Milano.
Detectors
1015 Plastic-Color

# All from CERN, where the chamber lies

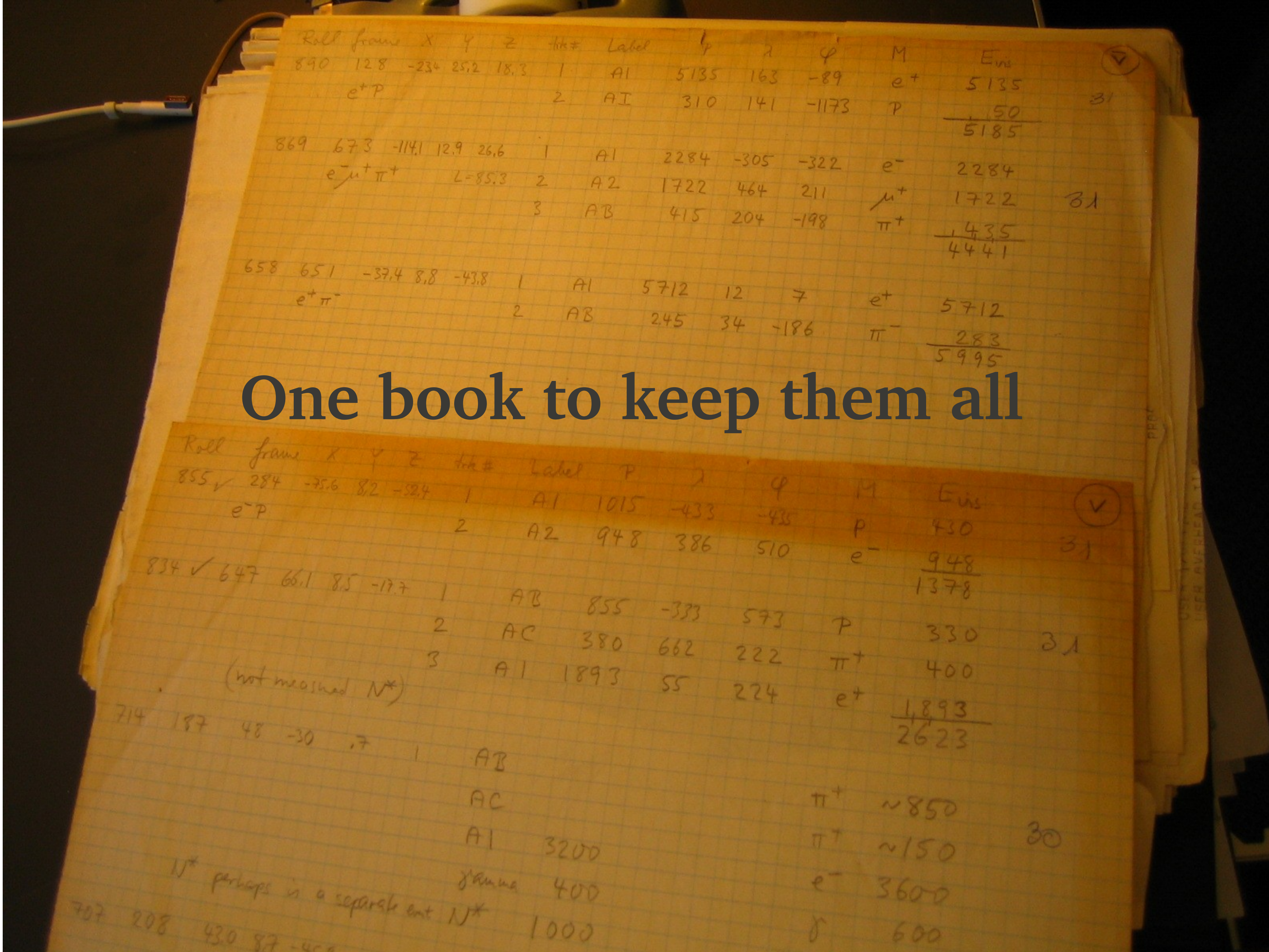



One book to rule them all
LEITZ
1050 Leitz R 50
LEITZ
Leitz A 45
MORFÍN
INCLUSIVE NEUTRAL CURRENT
I Experimental
II Theoretical

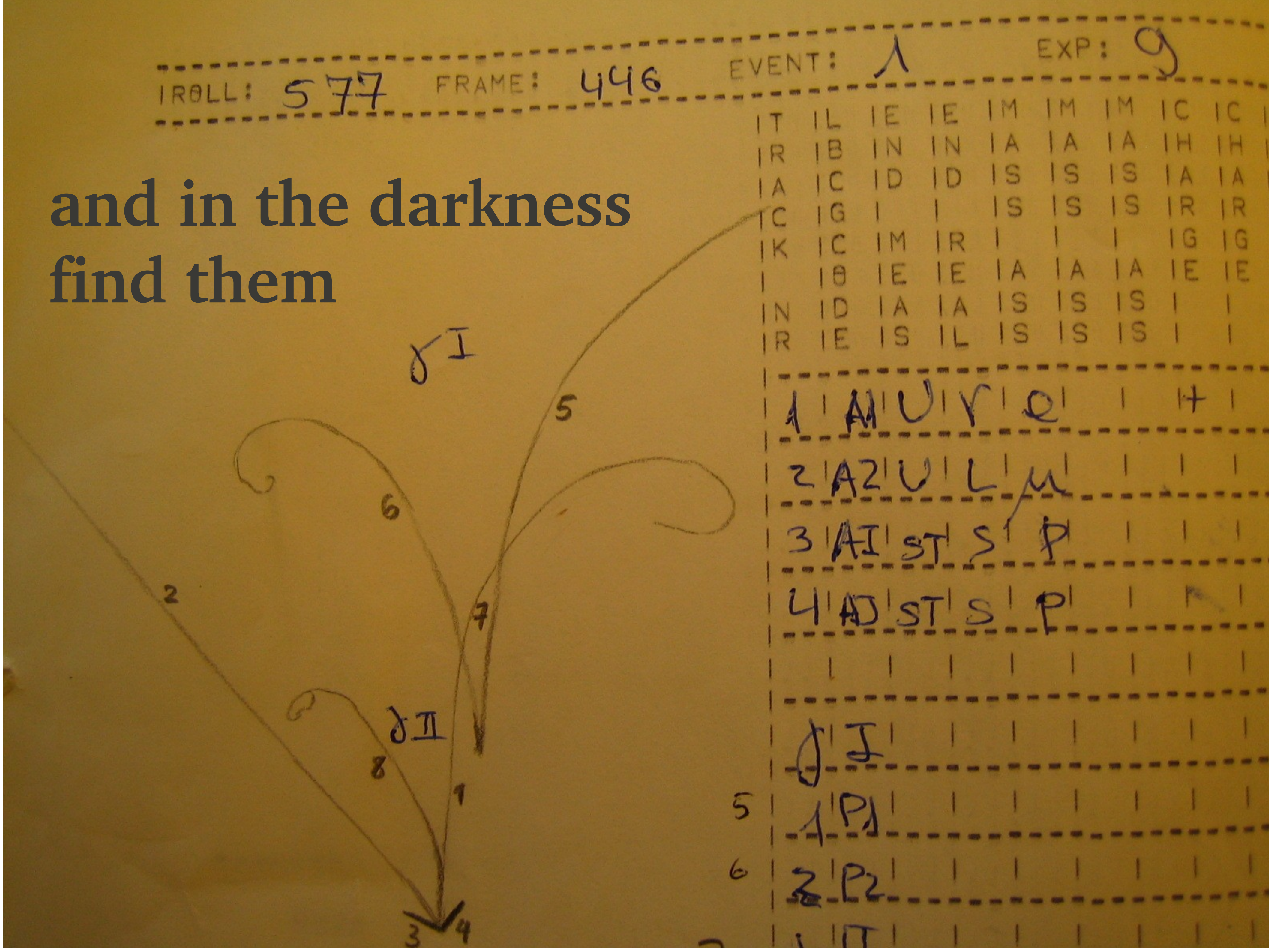
# One book to bind them



# One book to keep them all



## and in the darkness find them



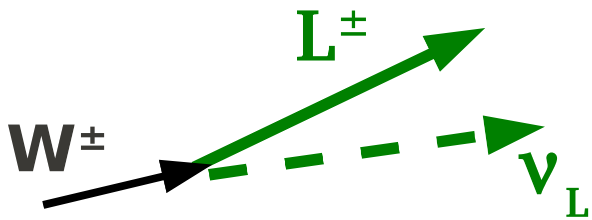
Far away, across the sea,
where neutrinos fly.

Track 1 has been
measured twice,
once as regular electron ⊕
from HV and as
γ - track (Q1)
(Look at difference
in momentum) !!!

# Plan for this talk

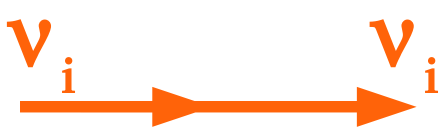
- Glorious neutrino bubble chamber era produced unique datasets that have been reanalyzed for various purposes over the years
- Data preservation model linked to the “computing” model
  - film shipped to collaborating institutions for scanning & reconstruction
  - paper records and summary tables assembled
  - Much of the raw data still exists
- We are now in the neutrino oscillations era. How will we preserve our data?

# Oscillation Formalism

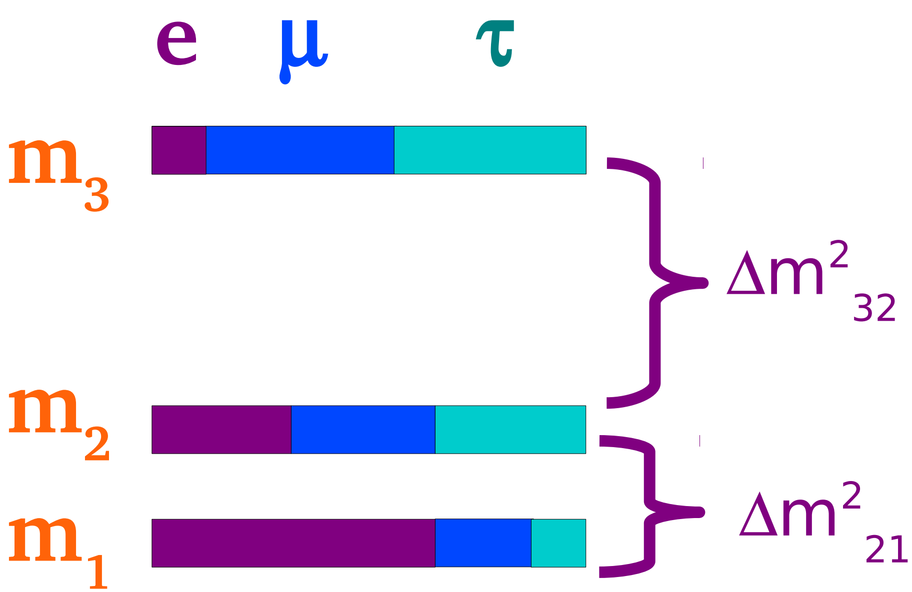



## Flavor

## Mass




$E^2=p^2+m_i^2$




PMNS
Mixing Matrix

$$U=\begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix}$$

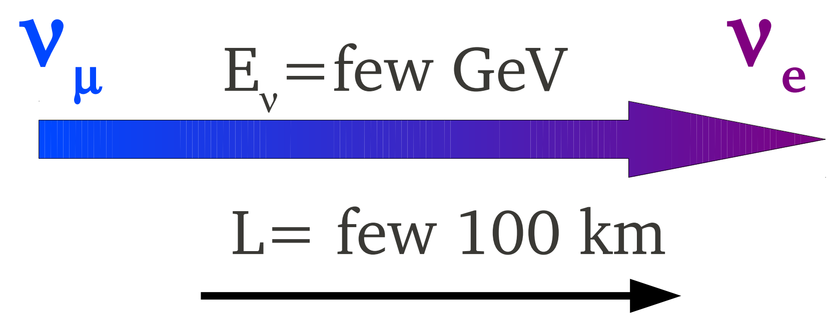
Propagation (vacuum)

flavor mass states

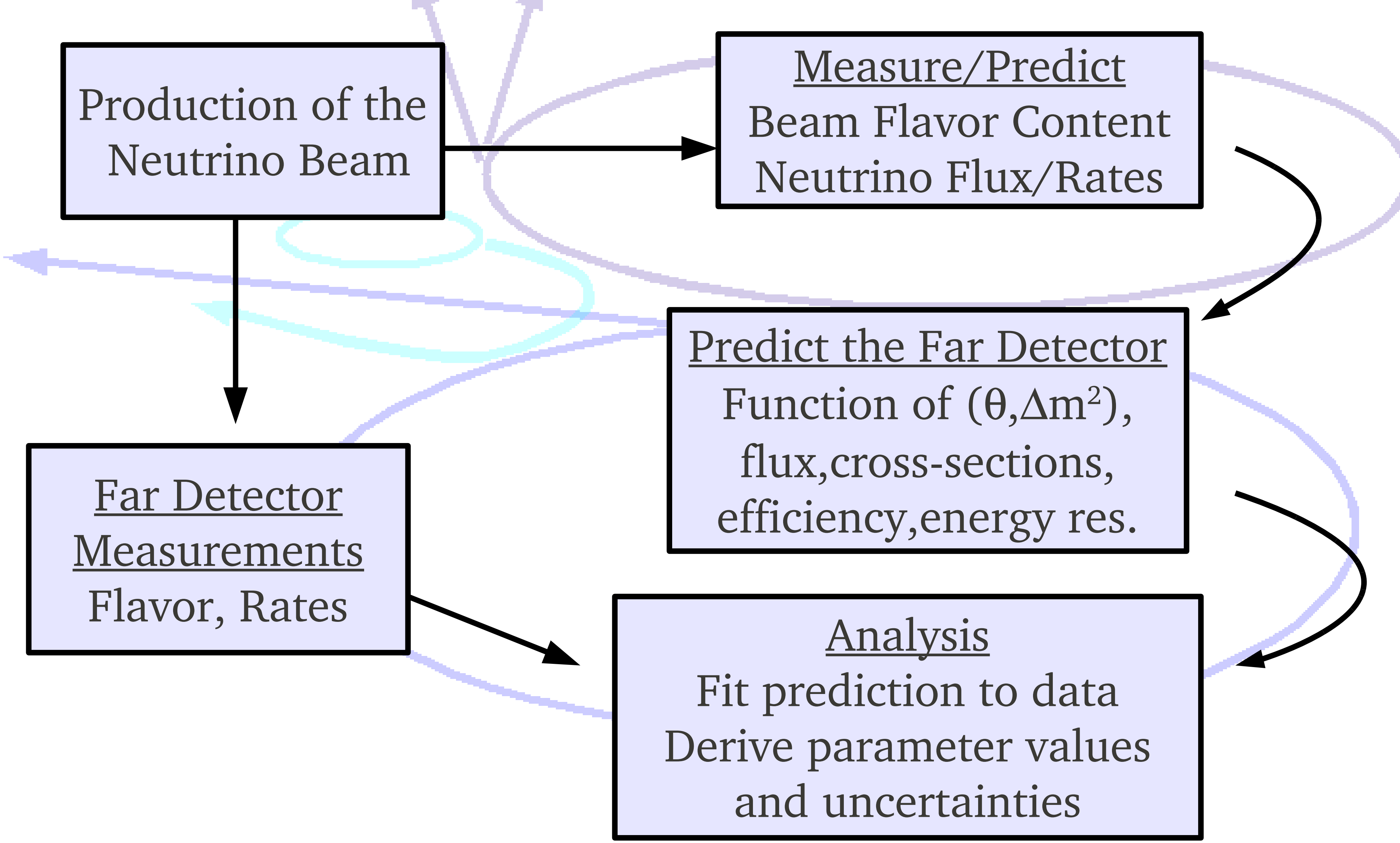
$$|\nu_\alpha(L)\rangle=\sum_j U^*_{\alpha j}|\nu_j\rangle \exp(-i L m_j^2/2p)$$

(L=distance)

## Flavor change

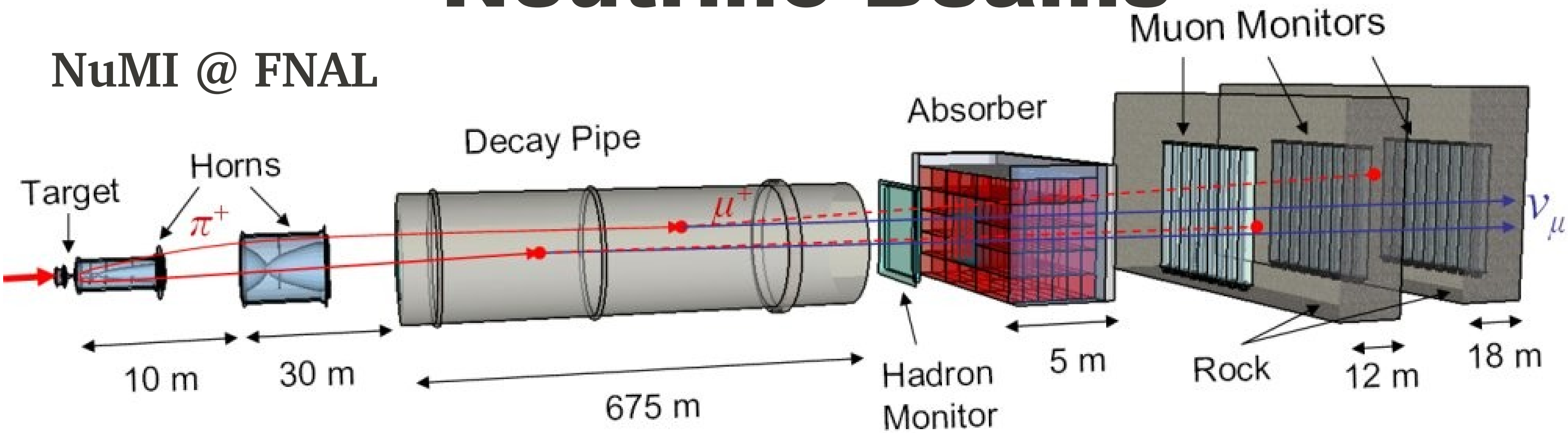


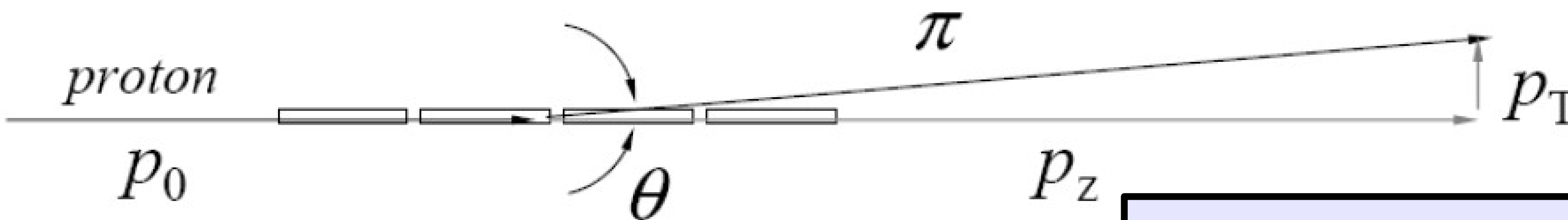
# An Accelerator Experiment In One Slide



# Neutrino Beams

## NuMI @ FNAL




- $\pi$,K production off a solid target
- Wide range of pt, pz
- Cross-sections not well known

<u>Important decay modes</u>

$\pi^+, K^+ \rightarrow \mu^+ \nu_\mu$

$K^+ \rightarrow \pi^0 \mu^+ \nu_\mu,\ \pi^0 e^+ \nu_e$

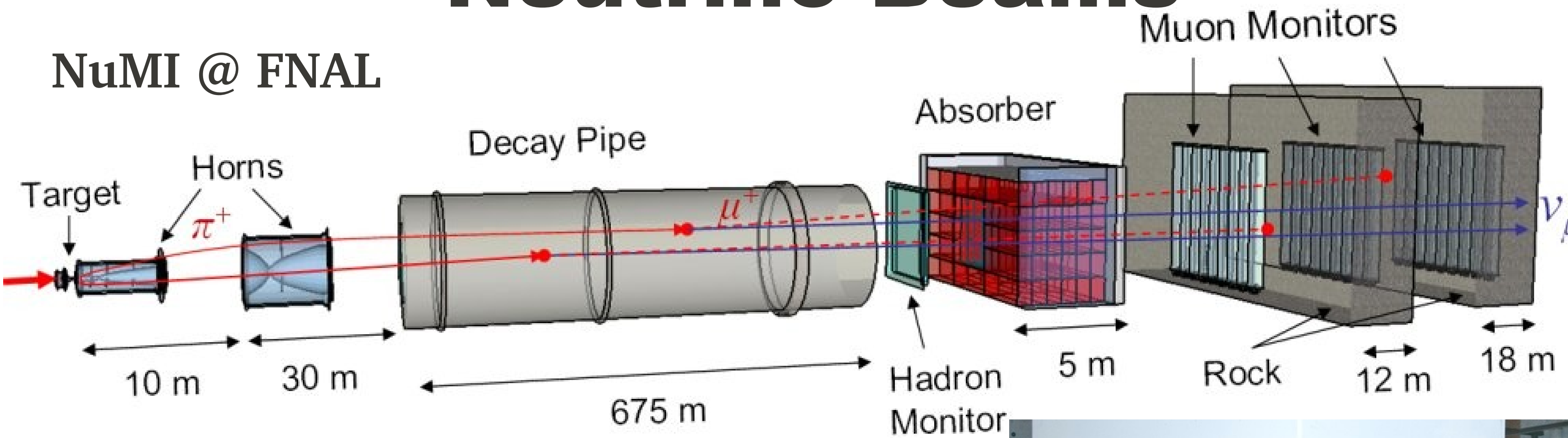
$K^0_L \rightarrow \pi^- e^+ \nu_e + CC$

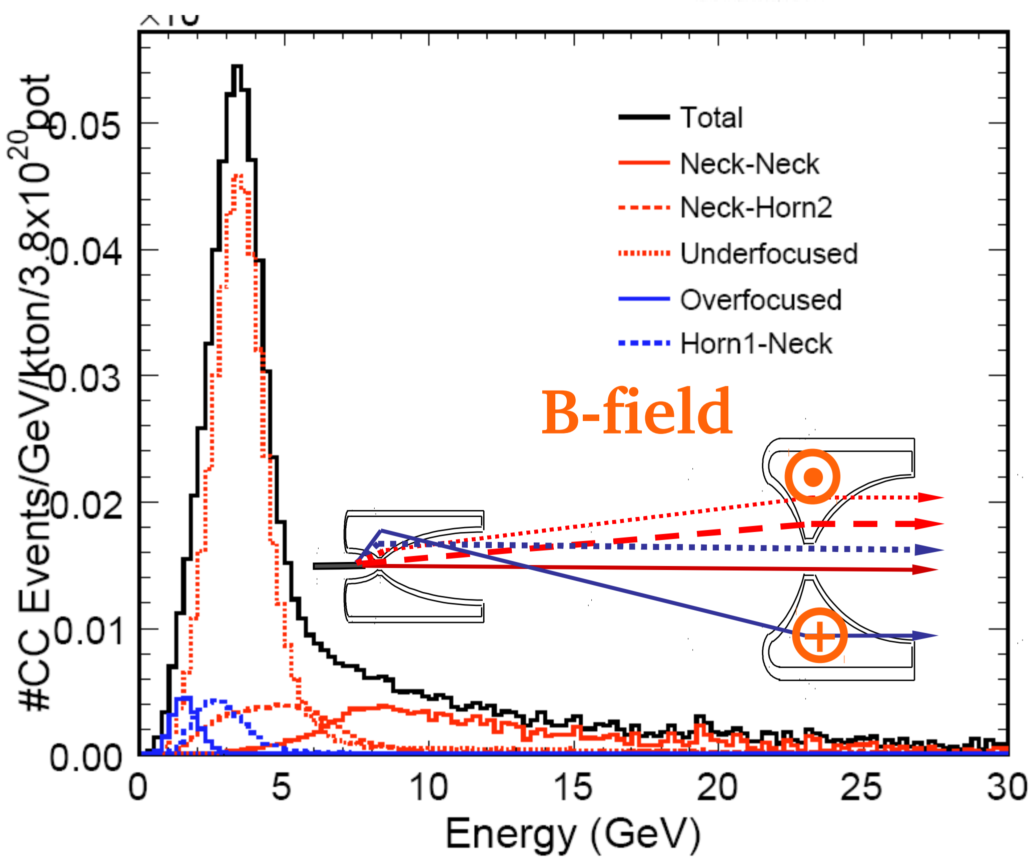
$\rightarrow \pi^- \mu^+ \nu_\mu + CC$

$\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$

# Neutrino Beams

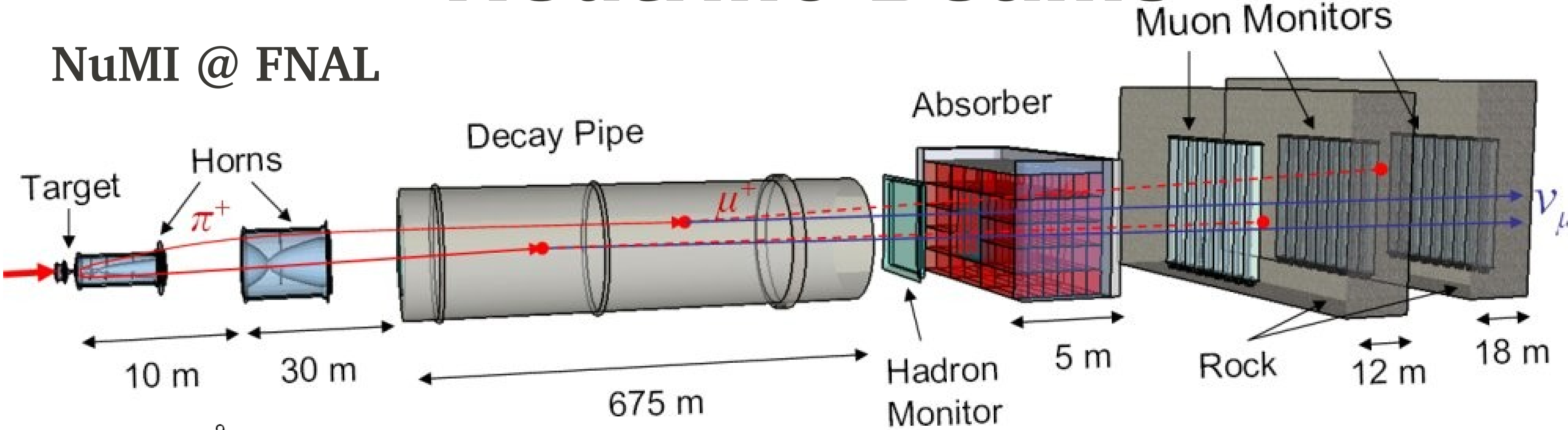
NuMI @ FNAL

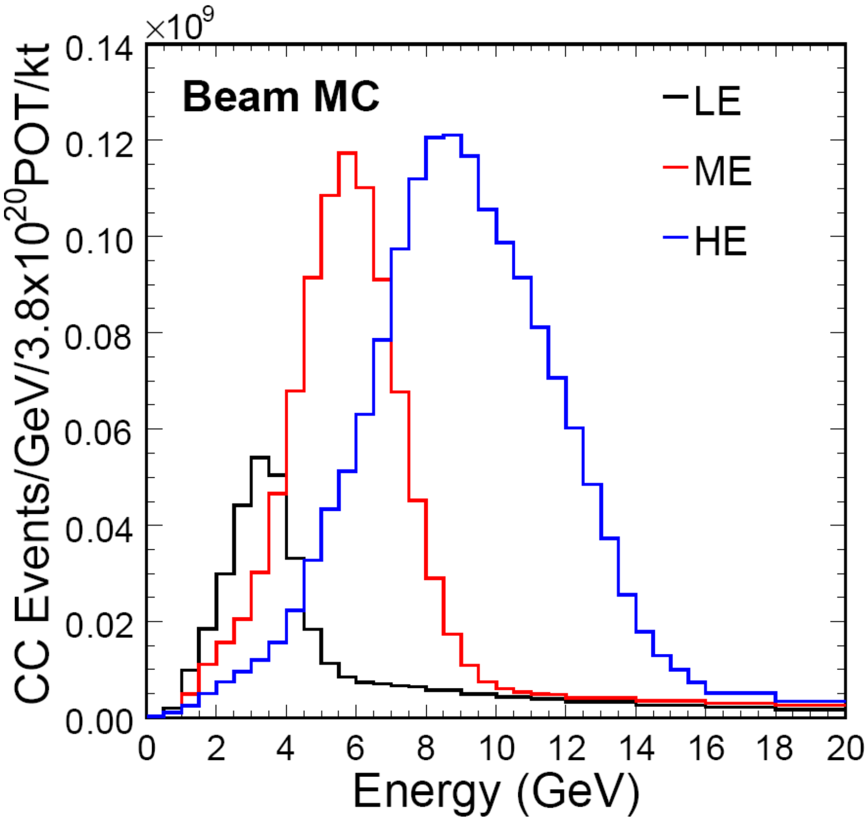


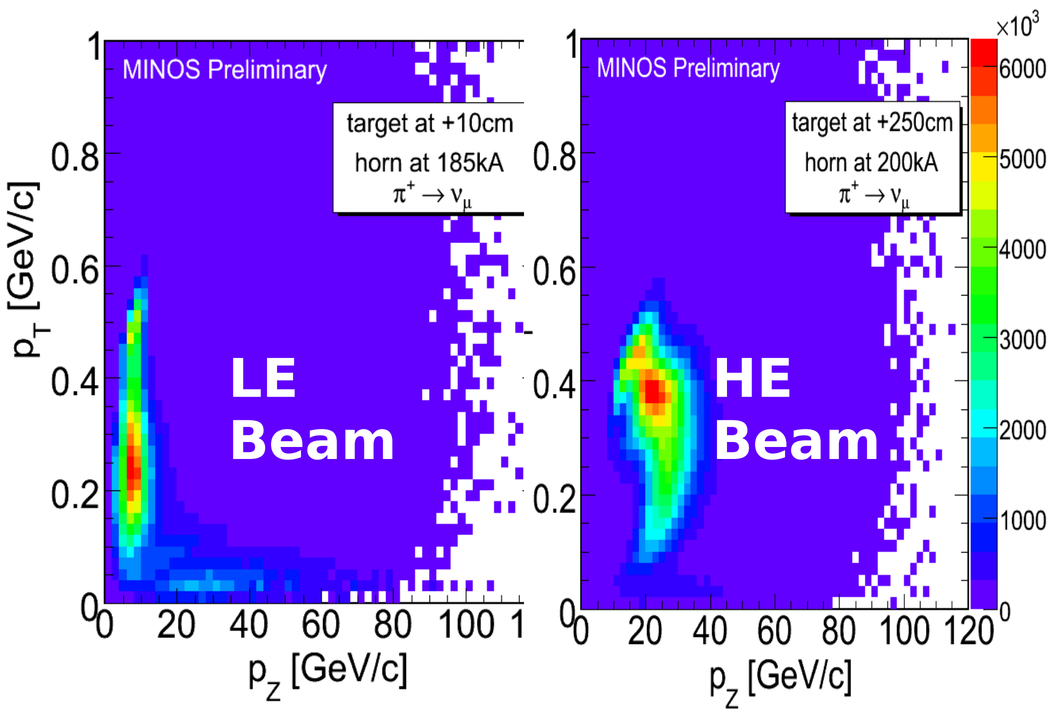




# Neutrino Beams

NuMI @ FNAL



# Muon neutrino disappearance

**PMNS Mixing Matrix**

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix}$$

**Propagation (vacuum)**

flavor mass states

$$|\nu_\alpha(L)> = \sum_j U^*_{\alpha j} |\nu_j> \exp(-i L m_j^2 / 2p)$$

(L=distance)

$$P(\nu_\mu \to \nu_\mu) = 1 - 4|U_{\mu 3}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{31} - 4|U_{\mu 3}|^2 |U_{\mu 2}|^2 \sin^2\Delta_{32} - 4|U_{\mu 2}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{21}$$

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

Characteristic oscillatory behavior depends on $\Delta m^2$ and L/E

"Survival Probability"
For a neutrino of energy E a distance L from the source

# Muon neutrino disappearance

**PMNS Mixing Matrix**

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix}$$

$$\mathbf{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13} e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} P(\nu_\mu \rightarrow \nu_\mu) = 1 &- 4|U_{\mu 3}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{31} \\ &- 4|U_{\mu 3}|^2 |U_{\mu 2}|^2 \sin^2\Delta_{32} \\ &- 4|U_{\mu 2}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{21} \end{aligned}$$

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

# Muon neutrino disappearance

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

"Atmospheric oscillations"
$L/E \sim 500$ km/GeV

Limiting Case
$\theta_{13} = 0$
$|\Delta_{21}| << |\Delta_{31}| \approx |\Delta_{32}|$

$$\left.\begin{aligned} P(\nu_\mu \to \nu_\mu) = 1 &- 4|U_{\mu3}|^2|U_{\mu1}|^2 \sin^2\Delta_{31} \\ &- 4|U_{\mu3}|^2|U_{\mu2}|^2 \sin^2\Delta_{32} \\ &- 4|U_{\mu2}|^2|U_{\mu1}|^2 \sin^2\Delta_{21} \end{aligned}\right\} \approx 1 - \sin^2 2\theta \sin^2(|\Delta m^2| L/4E)$$

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13} e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

# Muon neutrino disappearance

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

"Atmospheric oscillations"
L/E~500 km/GeV

Limiting Case
$\theta_{13} = 0$
$|\Delta_{21}| << |\Delta_{31}| \approx |\Delta_{32}|$

$$\left.\begin{aligned} P(\nu_\mu \to \nu_\mu) = 1 &- 4|U_{\mu 3}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{31} \\ &- 4|U_{\mu 3}|^2 |U_{\mu 2}|^2 \sin^2\Delta_{32} \\ &- 4|U_{\mu 2}|^2 |U_{\mu 1}|^2 \sin^2\Delta_{21} \end{aligned}\right\} \approx 1 - \sin^2 2\theta \sin^2(|\Delta m^2| L/4E)$$

$|\Delta m^2| \approx |\Delta m^2_{32}|$

$\theta \approx \theta_{23}$

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13}e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

# Muon neutrino disappearance

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

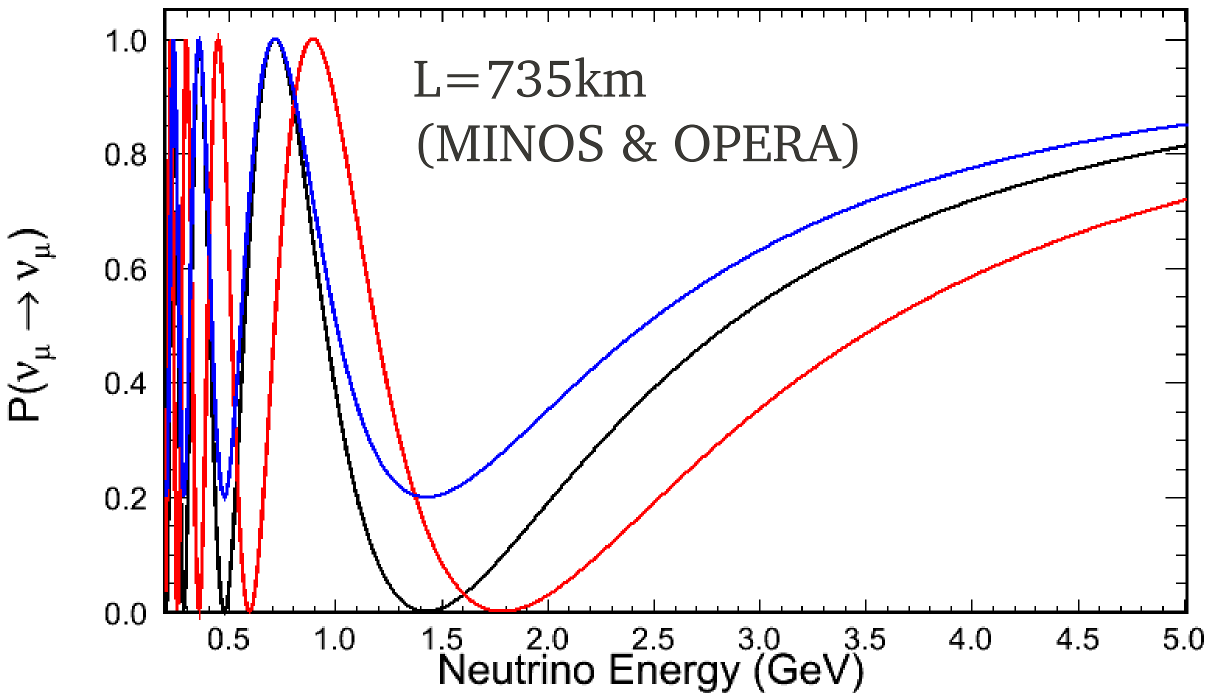
"Atmospheric oscillations"
L/E~500 km/GeV

<u>Limiting Case</u>
$\theta_{13} = 0$
$|\Delta_{21}| << |\Delta_{31}| \approx |\Delta_{32}|$

$$\left.\begin{aligned} P(\nu_\mu \to \nu_\mu) = 1 &- 4|U_{\mu 3}|^2|U_{\mu 1}|^2 \sin^2\Delta_{31} \\ &- 4|U_{\mu 3}|^2|U_{\mu 2}|^2 \sin^2\Delta_{32} \\ &- 4|U_{\mu 2}|^2|U_{\mu 1}|^2 \sin^2\Delta_{21} \end{aligned}\right\} \approx 1 - \sin^2 2\theta \sin^2(|\Delta m^2| L/4E)$$

$|\Delta m^2| \to \approx |\Delta m^2_{32}|$

$\theta \to \approx \theta_{23}$




$|\Delta m^2| = 2.5\times10^{-3}$ eV$^2$, $\sin^2 2\theta = 0.8$

$|\Delta m^2| = 2.5\times10^{-3}$ eV$^2$, $\sin^2 2\theta = 1.0$

$|\Delta m^2| = 3.0\times10^{-3}$ eV$^2$, $\sin^2 2\theta = 1.0$

# Muon neutrino disappearance

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

"Atmospheric oscillations"
$L/E \sim 500$ km/GeV

<u>Limiting Case</u>

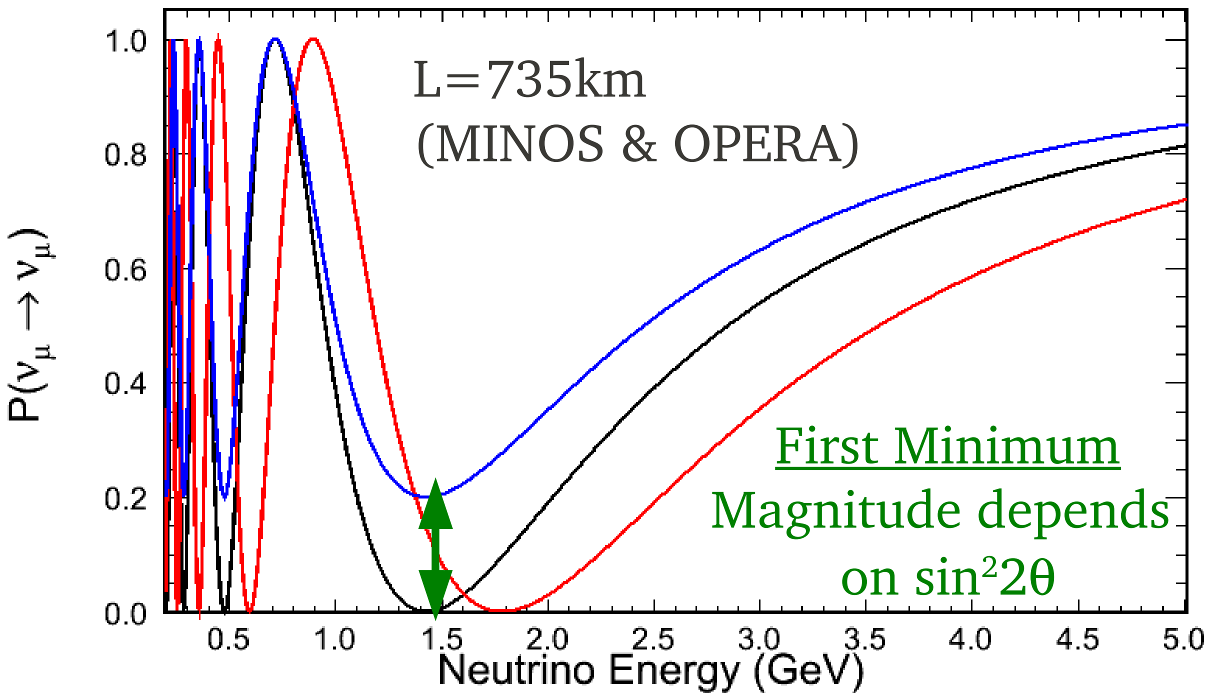
$\theta_{13} = 0$

$|\Delta_{21}| << |\Delta_{31}| \approx |\Delta_{32}|$

$$\left.\begin{aligned} P(\nu_\mu \rightarrow \nu_\mu) = 1 - & 4|U_{\mu 3}|^2|U_{\mu 1}|^2 \sin^2\Delta_{31} \\ - & 4|U_{\mu 3}|^2|U_{\mu 2}|^2 \sin^2\Delta_{32} \\ - & 4|U_{\mu 2}|^2|U_{\mu 1}|^2 \sin^2\Delta_{21} \end{aligned}\right\} \approx 1 - \sin^2 2\theta \sin^2(|\Delta m^2| L/4E)$$

$|\Delta m^2| \rightarrow \approx |\Delta m^2_{32}|$

$\theta \rightarrow \approx \theta_{23}$




$|\Delta m^2| = 2.5\text{x}10^{-3}$ eV$^2$, $\sin^2 2\theta = 0.8$

$|\Delta m^2| = 2.5\text{x}10^{-3}$ eV$^2$, $\sin^2 2\theta = 1.0$

$|\Delta m^2| = 3.0\text{x}10^{-3}$ eV$^2$, $\sin^2 2\theta = 1.0$

# Muon neutrino disappearance

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E}$$

"Atmospheric oscillations"
L/E~500 km/GeV

Limiting Case

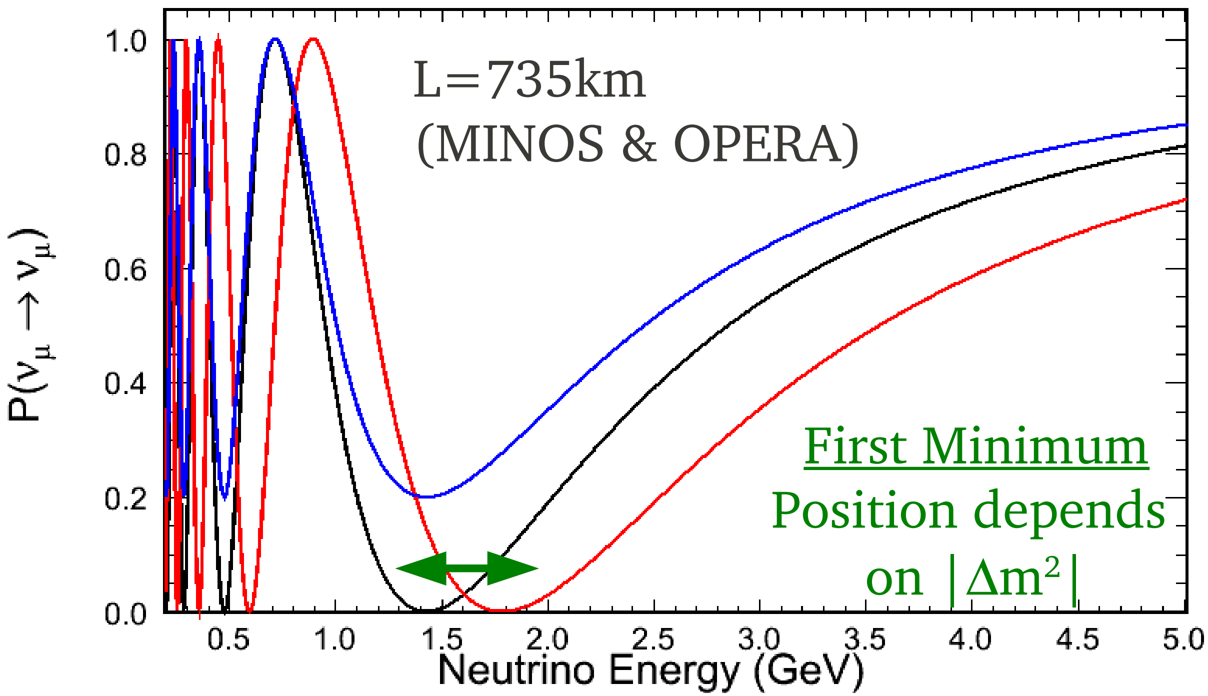
$\theta_{13} = 0$

$|\Delta_{21}| << |\Delta_{31}| \approx |\Delta_{32}|$

$$\left.\begin{aligned} P(\nu_\mu \to \nu_\mu) = 1 - &4|U_{\mu3}|^2|U_{\mu1}|^2 \sin^2\Delta_{31} \\ - &4|U_{\mu3}|^2|U_{\mu2}|^2 \sin^2\Delta_{32} \\ - &4|U_{\mu2}|^2|U_{\mu1}|^2 \sin^2\Delta_{21} \end{aligned}\right\} \approx 1 - \sin^2 2\theta \sin^2(|\Delta m^2| L/4E)$$

$|\Delta m^2| \to \approx |\Delta m^2_{32}|$

$\theta \to \approx \theta_{23}$




$|\Delta m^2| = 2.5\text{x}10^{-3}\ \text{eV}^2, \sin^2 2\theta = 0.8$

$|\Delta m^2| = 2.5\text{x}10^{-3}\ \text{eV}^2, \sin^2 2\theta = 1.0$

$|\Delta m^2| = 3.0\text{x}10^{-3}\ \text{eV}^2, \sin^2 2\theta = 1.0$

# Challenge: Predict the Far Detector

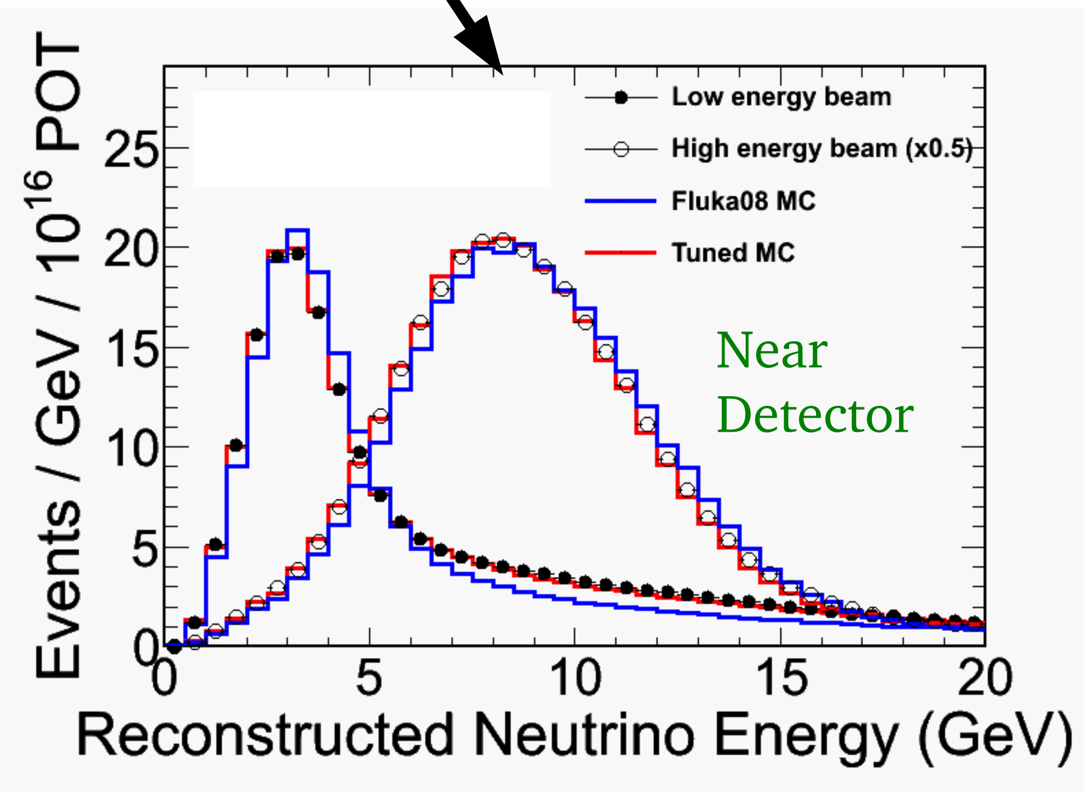
- You want to know neutrino interaction rates as a function of energy for all flavors.
- A priori, $dN/dp_T dp_z$ off your target material for $\pi$,K is not known very accurately.
- Your target is thick. Tertiary interactions occur in it, horns, beampipe, etc. Target may wear out over time.
- Your focusing system has uncertainties.
- The weak interaction forces you to use neutrino-nucleus interactions. Significant uncertainties in exclusive and inclusive CC and NC cross-sections for few-GeV $E_\nu$.
- $\nu_\mu$,$\nu_e$-CC & NC extrapolate differently.
- Some backgrounds may oscillate.

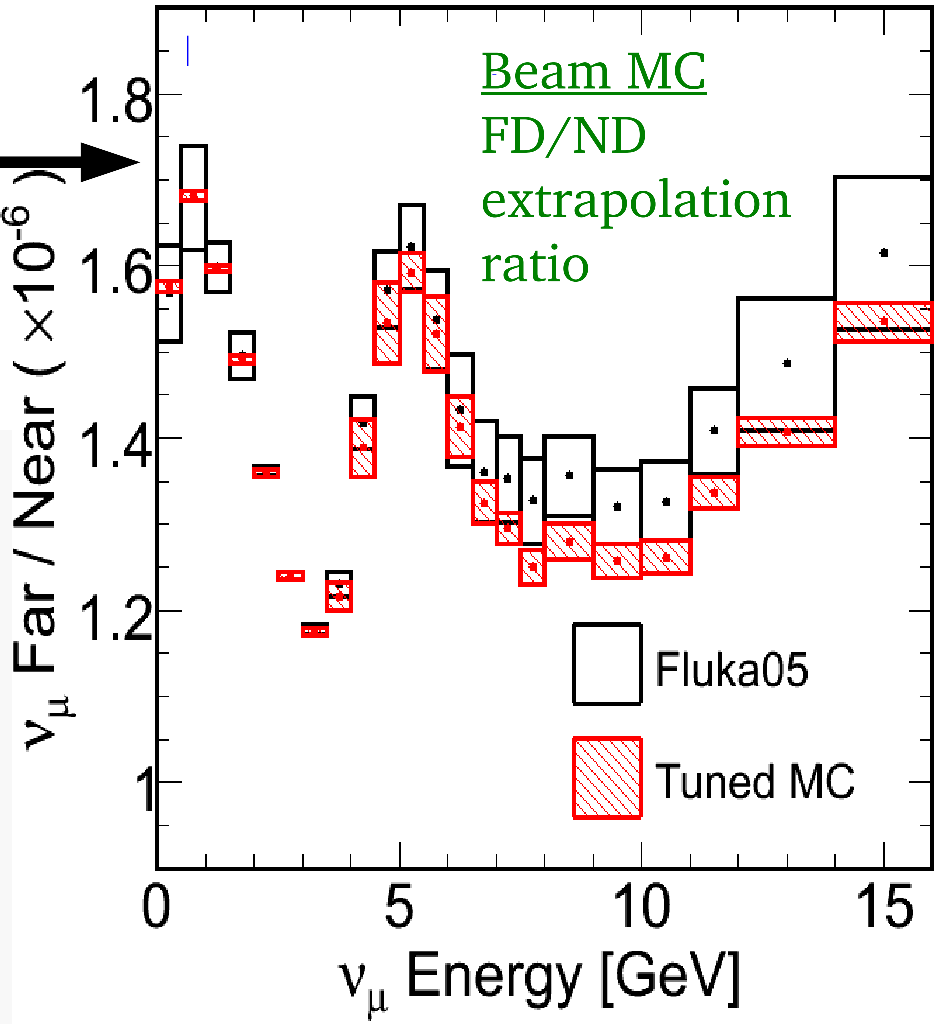
# Case study: MINOS $\nu_\mu \rightarrow \nu_\mu$

- MINOS has a ND and a FD made of the same materials and with very similar acceptance
- The ND is used to measure the inclusive $\nu_\mu$–CC rate close to the source.
- The ND measurement is corrected for acceptance and impurity using the MC
- It is "extrapolated" to the far detector using a 2 dimensional Near Energy vs. Far Energy Matrix
- This yields a no-oscillations prediction of the $\nu_\mu$–CC rate in bins of energy. The prediction is used to correct the FD MC which is then fit to the data.

# Case study: MINOS $\nu_\mu \rightarrow \nu_\mu$

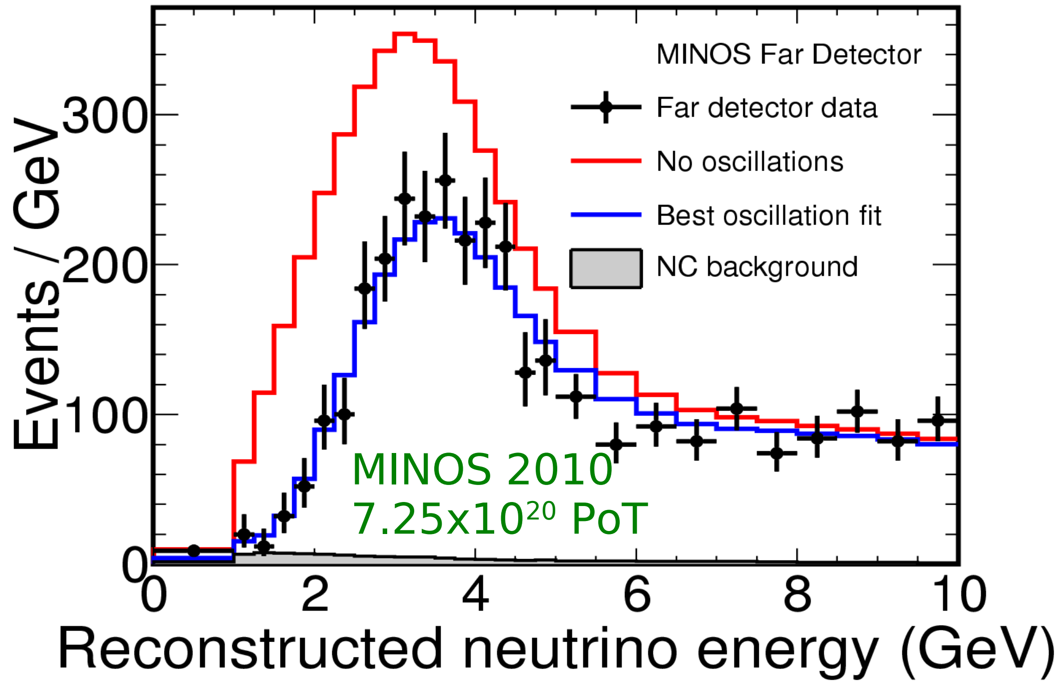
- Extrapolation improved by tuning beam MC to data in multiple focusing configurations

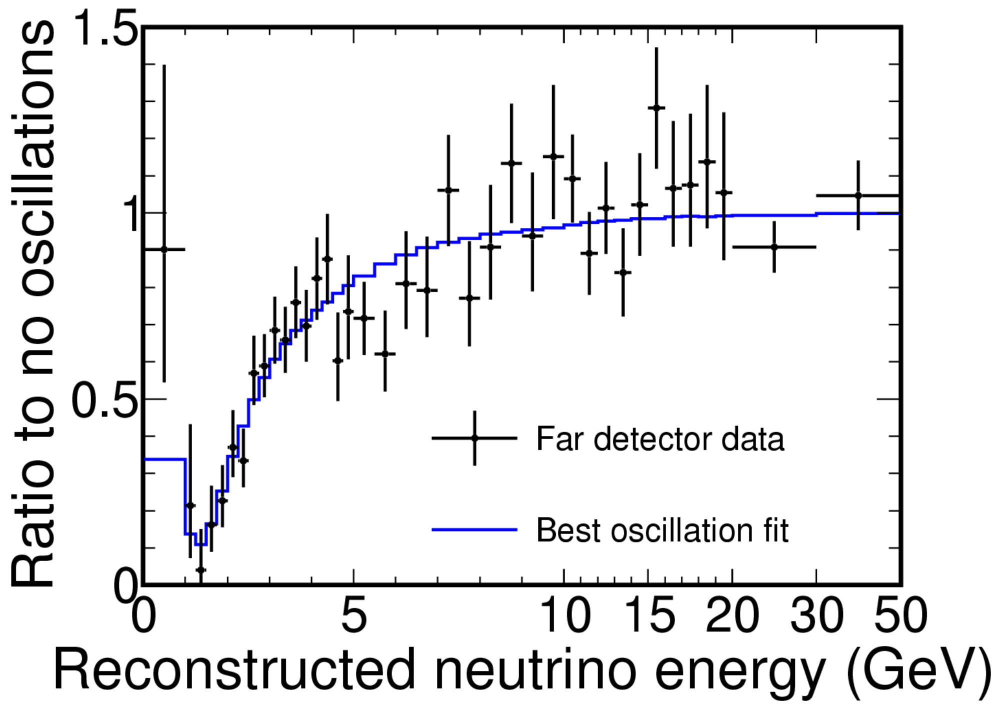


# Case study: MINOS $\nu_\mu \rightarrow \nu_\mu$

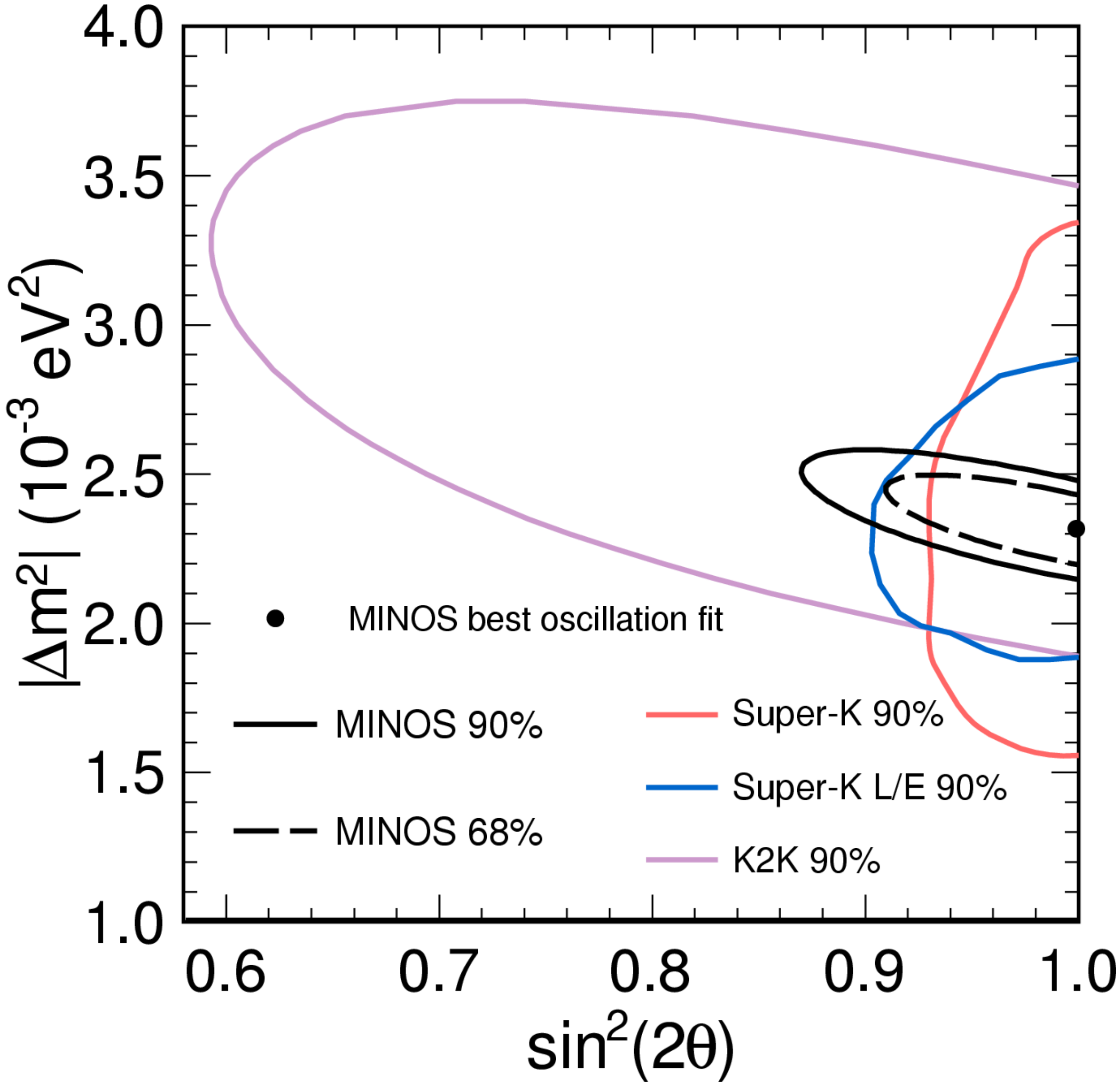
arXiv: 1103.0340v1




- At the FD, events are selected with 88.7% efficiency and 98.3% purity
- Analysis includes interactions inside the FD and muons from interactions in the surrounding rock.
- FD prediction fit to FD data, varying oscillation parameters
- Overall normalization, NC normalization and muon and shower energy scales included as nuisance parameters
- Fit is done in five energy resolution bins, which provides additional sensitivity.

# Case study: MINOS $\nu_\mu \rightarrow \nu_\mu$




2010 $7.25 \times 10^{20}$ PoT

**Fit result**

$|\Delta m^2| = 2.32 \, ^{+0.12}_{-0.08} \times 10^{-3}$ eV$^2$

$\sin^2 2\theta > 0.90$ (90% CL)

Consistent with maximal mixing.

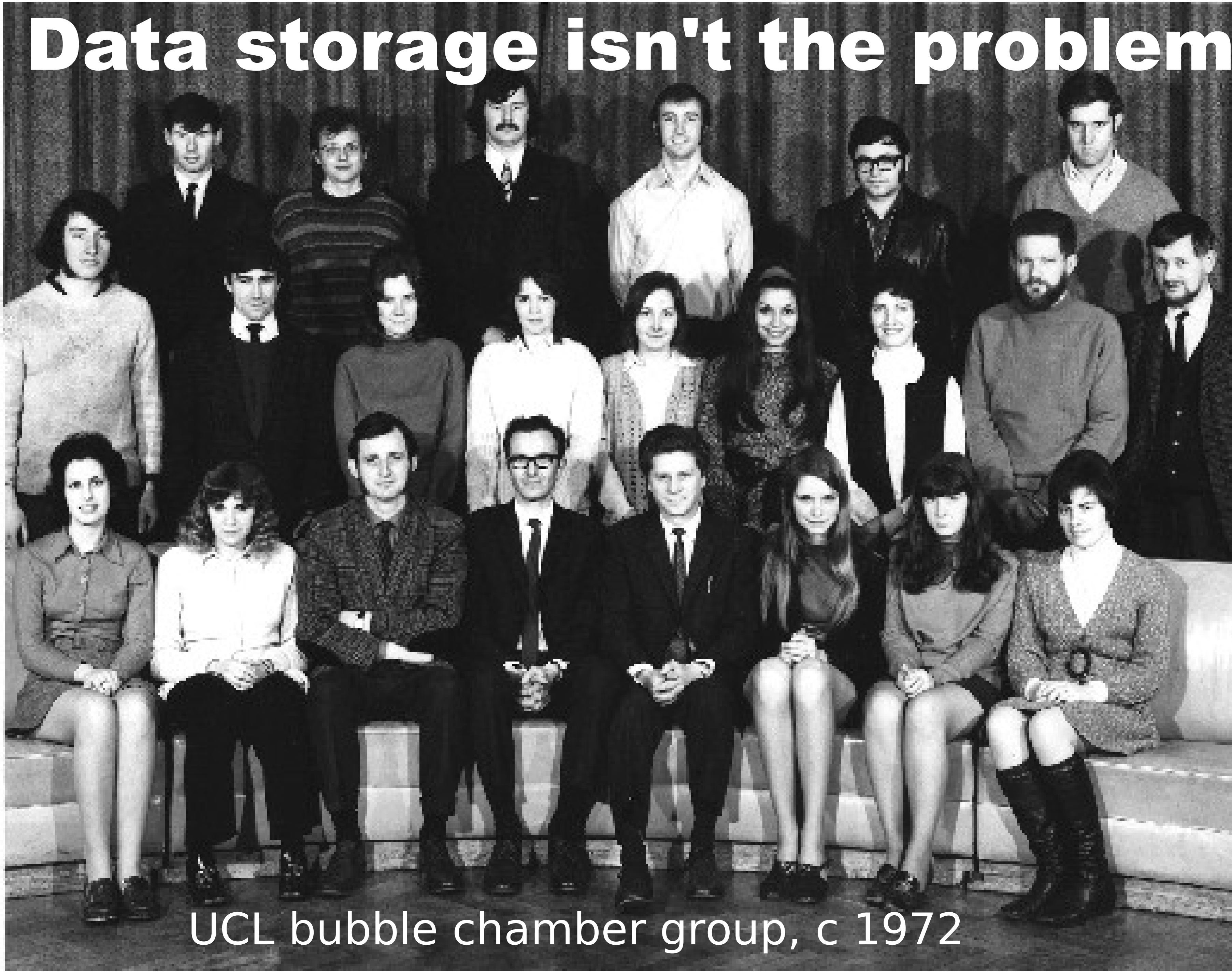
This is currently the best measurement of $|\Delta m^2|$ and the best accelerator measurement of $\sin^2 2\theta$

# How are data saved?

- Raw data in ROOT format
  - Very loose "trigger" keeps ~100% of neutrino interactions
  - Some online summarizing of calibrations, etc, done by event builder/trigger farm.
- Main files are standard ntuples "sntp" produced by reconstruction program
  - ROOT, readable in framework and via MakeClass
    - Minerva and Argoneut have used the latter successfully
  - Hit level information. Could almost redo reconstruction (need geometry, Bfield, etc)
  - MC truth included.
- I suppose that the data will be readable indefinitely.

# What resources are used?

- Enstore tape archive: total 600T
  - Raw Data: 10T   Reco data: 200T   Reco MC: 300T
- Single analysis footprint: ($\nu_\mu$ disappearance)
  - sntp reco data: 2.0T   sntp reco MC: 10T
- Generation:
  - GEANT3/NEUGEN3 based program → But, ROOT TVirtualMC
  - $1x10^{20}$ POT in LE takes 6,000,000 SpecInt2000 days (no reco)
- Database MySQL: ~2T
  - mostly backups.  Working DB is ~100G
  - Bigger concern: long term support.


Data storage isn't the problem
UCL bubble chamber group, c 1972

# Data storage isn't the problem

- The really irreplaceable thing is the knowledge locked up inside the heads of the collaboration
- Much is highly technical and known only by a few people
- In principal anything can be written down.
  - But, it takes time and is no fun.
  - People don't fully realize what they know.
- A re-analysis is likely done by a much smaller group than the original collaboration.
- These issues limit the scope of any re-analysis.

# What makes sense for MINOS?

- The most likely re-analysis would aim to use the data to constrain a new oscillation model
  - e.g, sterile neutrinos, non-standard interactions...
- A basic analysis would need:
  - Predicted distributions of $\nu_\mu$,$\nu_e$,$\nu_\tau$, NC in bins of true energy at the FD
    - The ND is used to make the predictions. Assume no-oscillations.
  - The detector response and efficiency for each category. Likely encoded in matrices.
  - Error matrices or shifted spectra for systematics studies.
- Having code to do the fitting and especially the Feldman-Cousins contour calculations would be useful but not essential.
- Some minimal documentation of what's what.

# How To Do A Little Better?

- In reality, rather than distributions, you'd save simple analysis ntuples.
  - MC truth kinematic & generator information
  - stdhep event record
  - neutrino's parent
  - reconstructed muon and shower energies
  - embedded weights for standard uncertainties
- Would allow you to do unbinned fits, calculate new systematic uncertainties.

# Looking for trouble?

- Maybe you want to change the event selection or look for a new process ($\nu_\tau$ appearance?)
- This requires data from the near detector.
- It may also require:
  - Hit level information (sntp files)
  - evaluation of new systematics
  - new MC for the ND and FD
  - new event generation
  - new reconstruction

  } = offline framework + DB
- Currently, this would probably require redoing calibrations.

# Summary

- Neutrino bubble chambers → unique datasets
  - long shelf life, reanalysis not uncommon
- Global neutrino oscillations program underway. Future experiments will be more sensitive, at least to standard oscillations.
  - Oscillations depend on L/E, and older experiments at different L,E may become relevant again.
  - Some experiments have unique capabilities which may not be present in future efforts.
- With a little work I think it's possible to save enough information from current experiments to permit some reanalysis.
  - The trouble comes if you want to change the event selection or reconstruction.