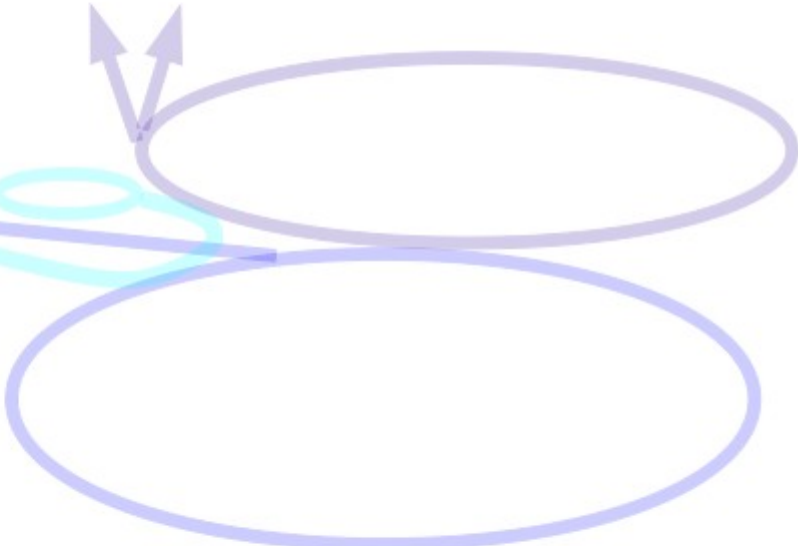
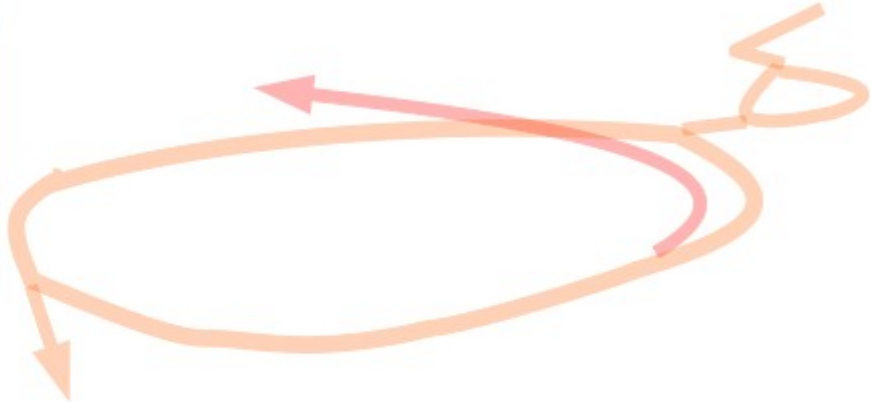


Data Preservation and Neutrinos



Data Preservation Workshop
May 16-18,
Fermilab



Mike Kordosky

William & Mary
May 17, 2011

A black and white photograph of a large, multi-story building complex with a prominent central tower or spire, situated along a riverbank. The buildings are reflected in the calm water of the river. The photograph is mounted in a window of a light-colored metal filing cabinet. To the left of the window, a blue folder is partially visible with the text "BORGHESE" and "BORGHESE" printed on it. To the right, another blue folder is visible with the number "172" printed on it. The filing cabinet has two drawers below the window, each with a silver handle and a small label. The top of the cabinet has a spiral-bound notebook and a small yellow object. The background shows a brightly lit office space with windows and ceiling lights.

One reel for
ancient Aachen
and the famous king

One reel for Don at Oxford



in his halls of stone

LEITZ

1050 Leitz R 50

One reel for
bustling London

INCLUSIVE N.C. - V + Y DISTRIB

and Tegid Wynn Jones

Im Archiv unter Nr.:
bis: vom: bis: Theoria

9	8	7	6	5	4	3	2	1	0	IX	VIII	VII	VI	V	IV	III	II	I	X-Z	W	U	8	9
30	25	20	15	10	5	XII	XI	X	0	IX	VIII	VII	VI	V	IV	III	II	I	X-Z	W	U	8	9

and one reel for the Italians

GARGAMELLE
USERS' HANDBOOK

CERN

down in Milano.

Detectors

1015 Plastic-Color

1015

All from CERN, where the chamber lies

LEITZ

1050 Leitz R 50

Yes
Scattering
Experimental

LEITZ

Leitz A 45

v

MOREIN

INCLUSIVE

NEUTRAL CURRENT

I. Experimental

II Theoretical

One book to

LEITZ

1050 Leitz R 50

Scattering
Experimental

LEITZ

Leitz A 45

rule
them all

MOREIN

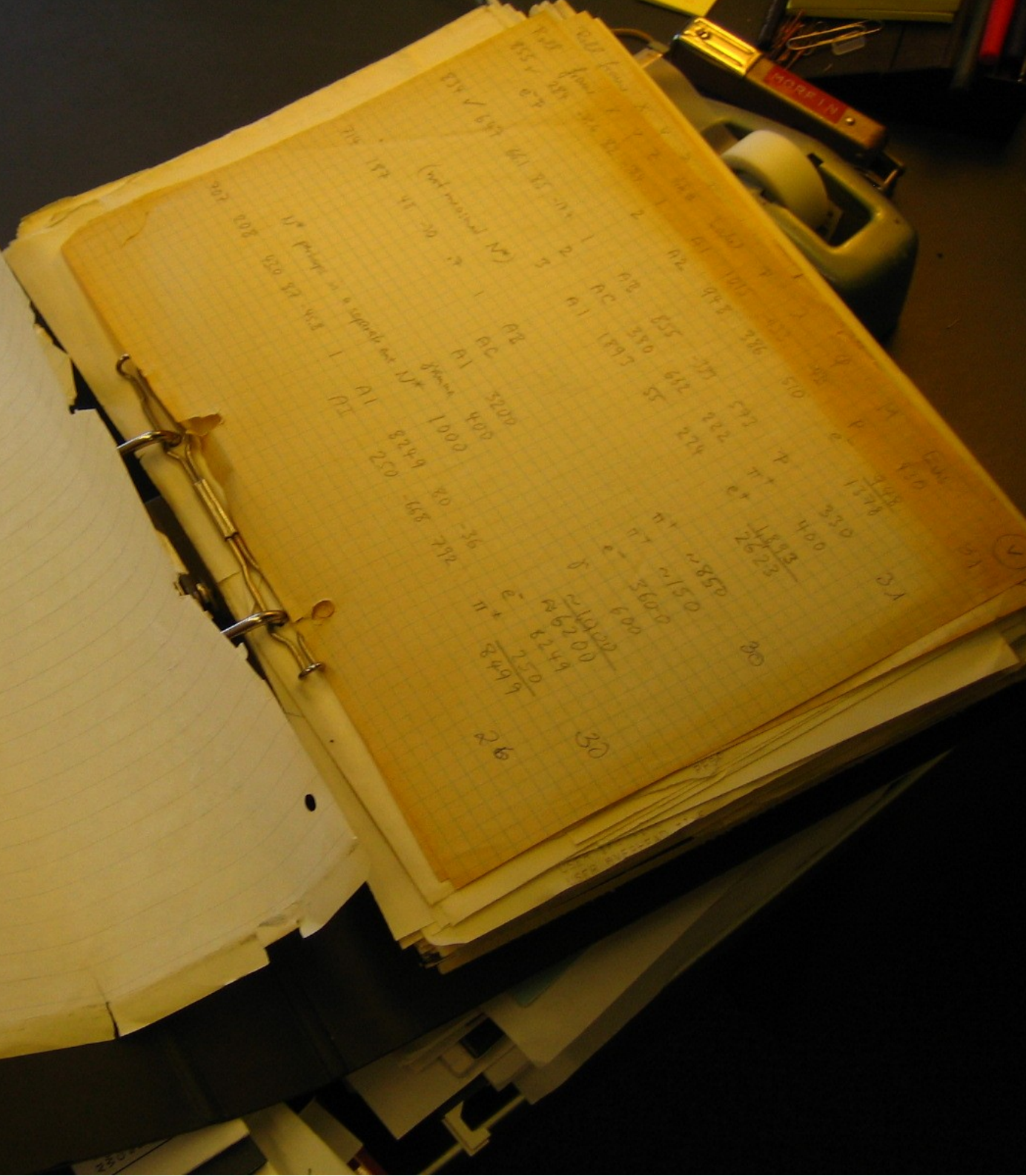
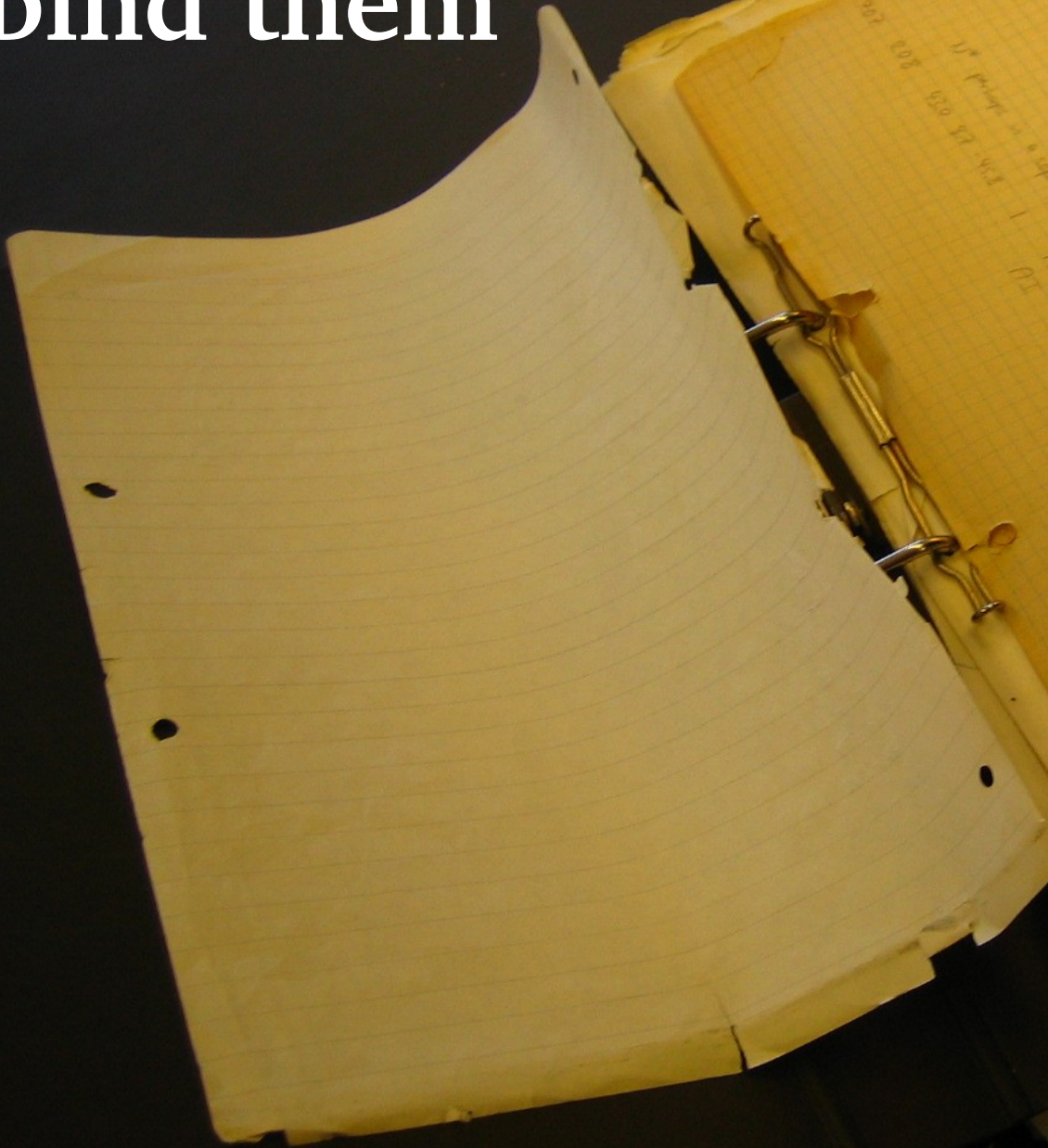
INCLUSIVE

NEUTRAL CURRENT

I. Experimental

II Theoretical

One book to
bind them



One book to keep them all

Roll frame	X	Y	Z	hit#	Label	p_x	p_y	p_z	M	E _{int}		
890	128	-234	25.2	18.3	1	A1	5135	163	-89	e ⁺	5135	31
e ⁺ P					2	A1	310	141	-1173	P	50	
											5185	
869	67.3	-114.1	12.9	26.6	1	A1	2284	-305	-322	e ⁻	2284	31
e ⁻ μ^+ π^+			L=85.3		2	A2	1722	464	211	μ^+	1722	
					3	AB	415	204	-198	π^+	435	
											4441	
658	65.1	-37.4	8.8	-43.8	1	A1	5712	12	7	e ⁺	5712	
e ⁺ π^-					2	AB	245	34	-186	π^-	283	
											5995	

Roll frame	X	Y	Z	hit#	Label	p_x	p_y	p_z	M	E _{int}		
855 ✓	284	-75.6	8.2	-52.4	1	A1	1015	-433	-435	P	430	31
e ⁻ P					2	A2	948	386	510	e ⁻	948	
											1378	
834 ✓	647	66.1	8.5	-17.7	1	AB	855	-333	573	P	330	31
					2	AC	380	662	222	π^+	400	
					3	A1	1893	55	224	e ⁺	1893	
											2623	

(not measured N*)

714	187	48	-30	.7	1	AB						
						AC				π^+	~850	
						A1	3200			π^+	~150	30
						gamma	400			e ⁻	3600	
						N*	1000			δ	600	

N* perhaps is a separate ent N*

707 208 450 87 -400

IROLL: 577

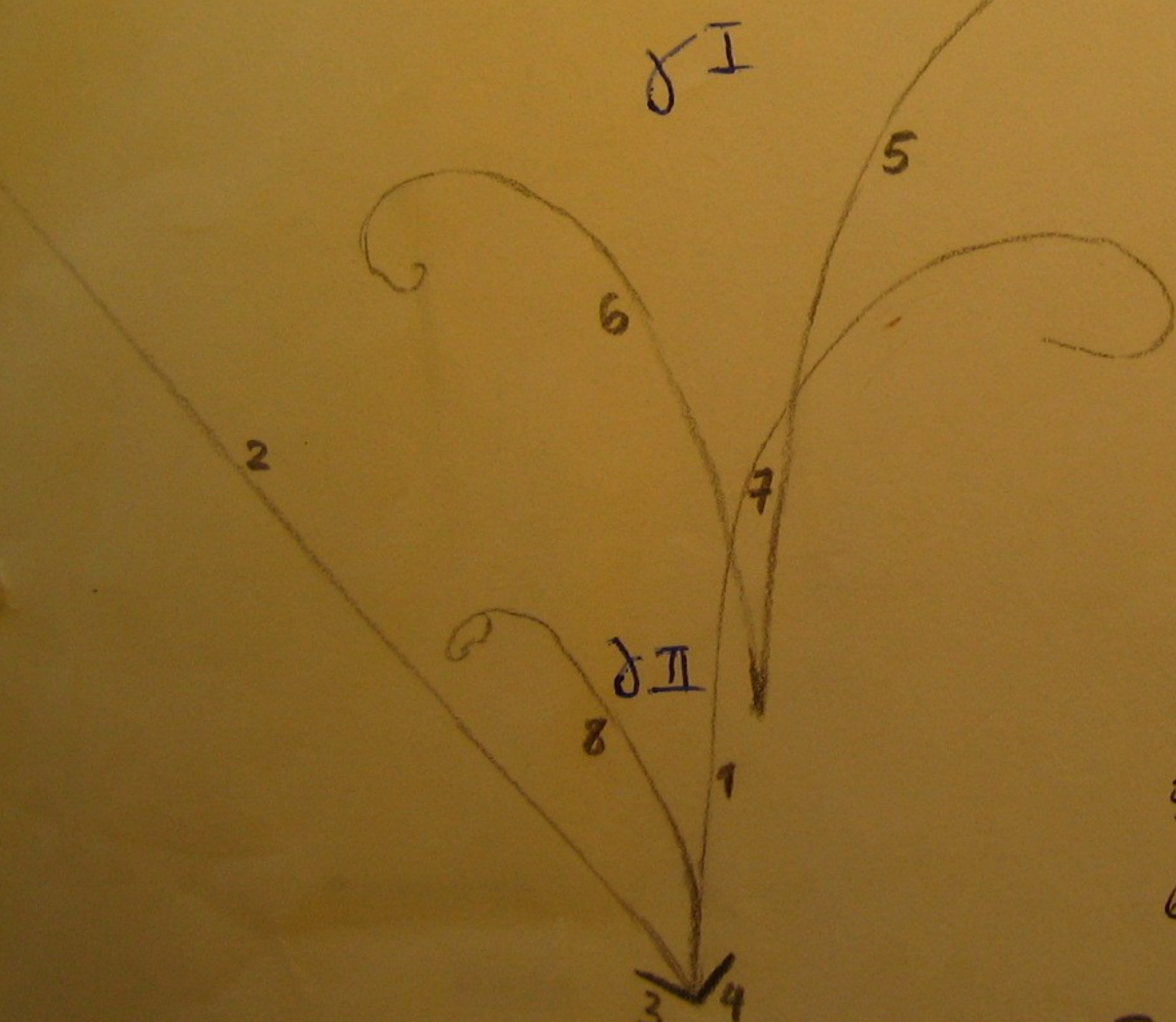
FRAME: 446

EVENT: λ

EXP: 9

and in the darkness
find them

IT	IL	IE	IE	IM	IM	IM	IC	IC
IR	IB	IN	IN	IA	IA	IA	IH	IH
IA	IC	ID	ID	IS	IS	IS	IA	IA
IC	IG	I	I	IS	IS	IS	IR	IR
IK	IC	IM	IR	I	I	I	IG	IG
I	IB	IE	IE	IA	IA	IA	IE	IE
IN	ID	IA	IA	IS	IS	IS	I	I
IR	IE	IS	IL	IS	IS	IS	I	I



1	A	U	V	Q	I	H
2	A	Z	U	L	M	
3	A	I	S	S	P	
4	A	D	S	S	P	

	δI					
5	1 P1					
6	2 P2					
	IT					

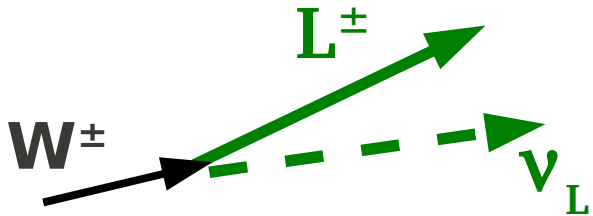
Far away, across the sea,
where neutrinos fly.

Track 1 has been
measured twice,
once as regular electron⁺
from HV and as
 γ -track (91)
(look at difference
in momentum) !!

Plan for this talk

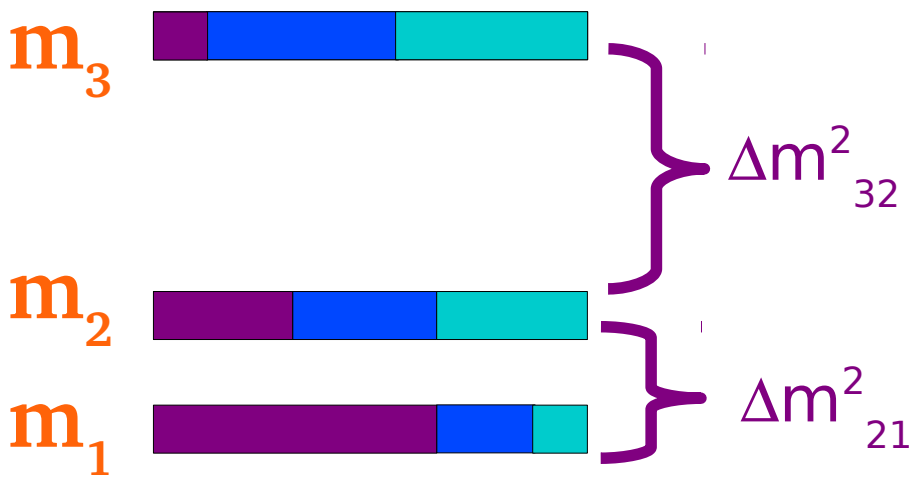
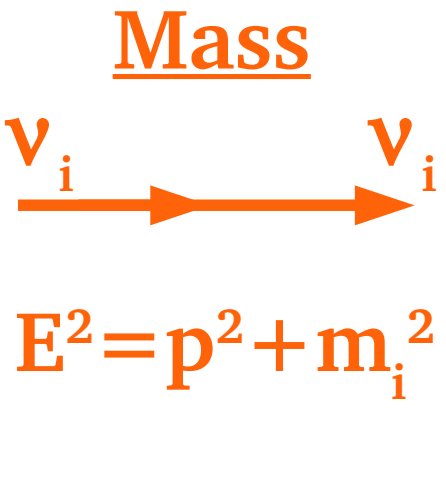
- Glorious neutrino bubble chamber era produced unique datasets that have been reanalyzed for various purposes over the years
- Data preservation model linked to the “computing” model
 - film shipped to collaborating institutions for scanning & reconstruction
 - paper records and summary tables assembled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

Oscillation Formalism



Flavor

e μ τ



PMNS
Mixing Matrix

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{bmatrix}$$

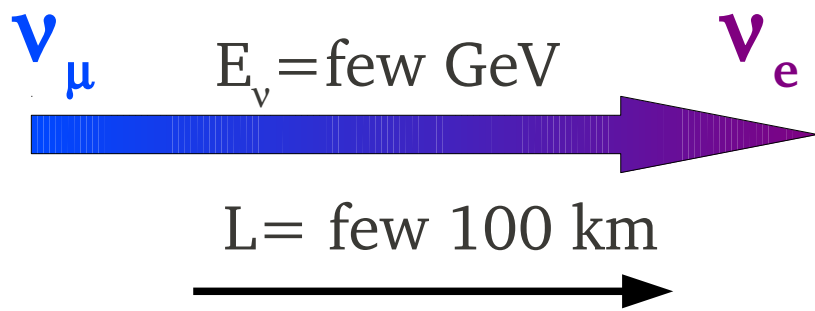
Propagation (vacuum)

flavor mass states

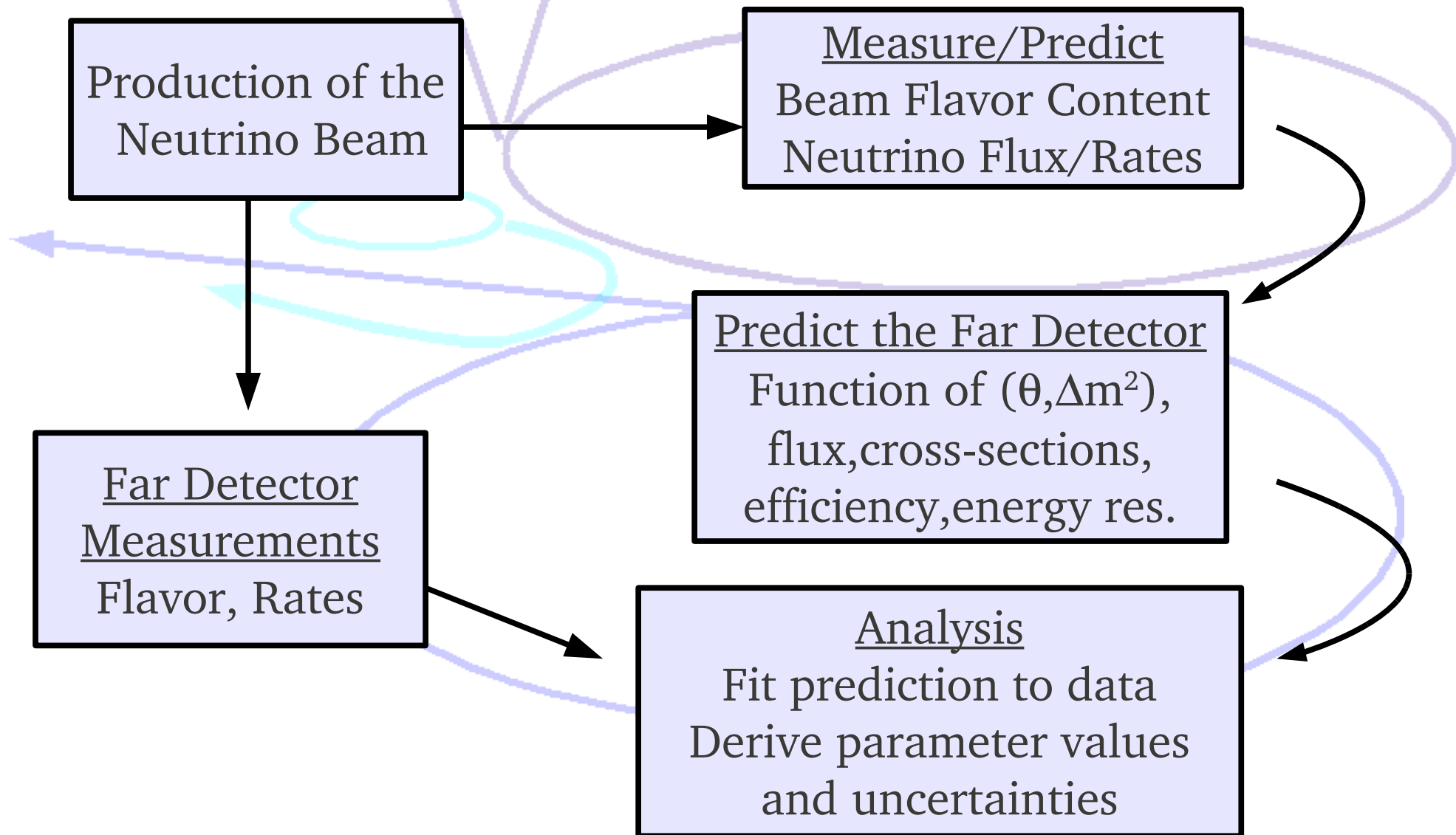
$$|\nu_\alpha(L)\rangle = \sum_j U_{\alpha j}^* |\nu_j\rangle \exp(-i L m_j^2 / 2p)$$

(L=distance)

Flavor change

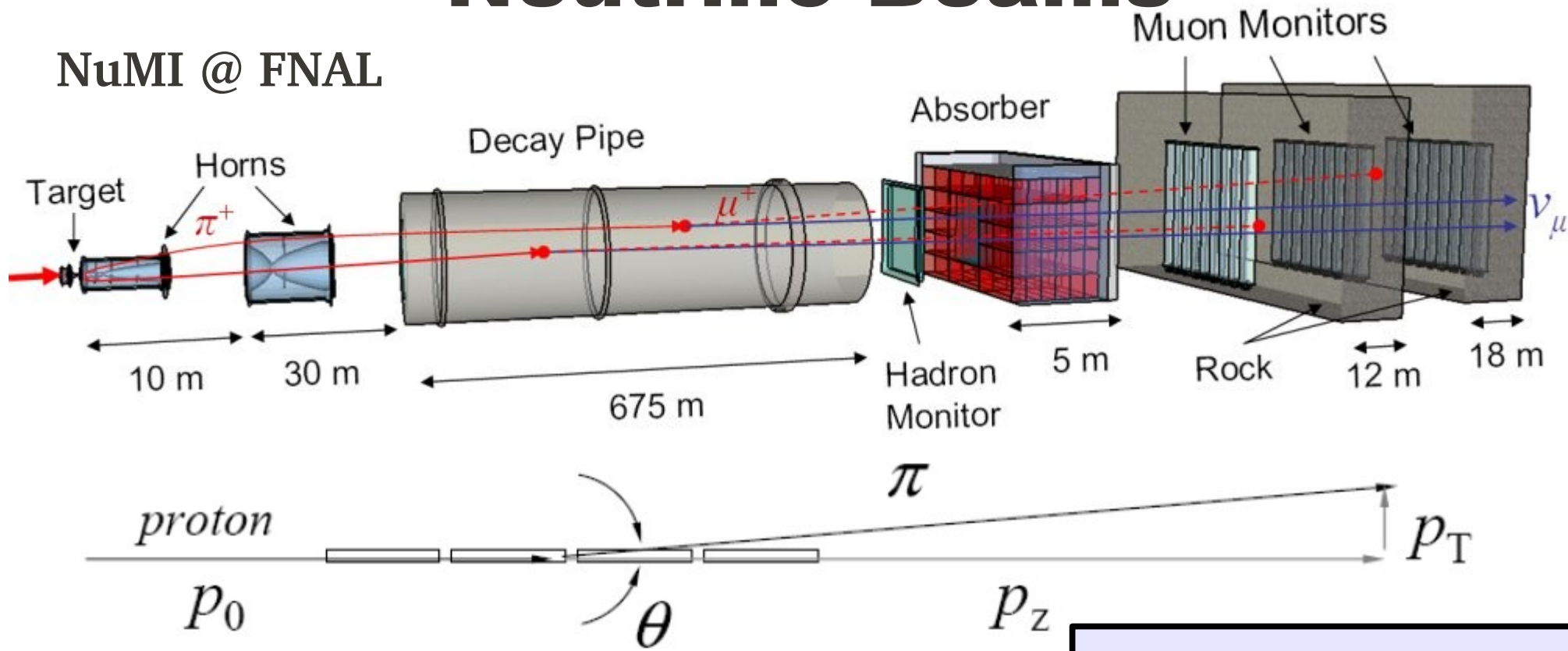


An Accelerator Experiment In One Slide



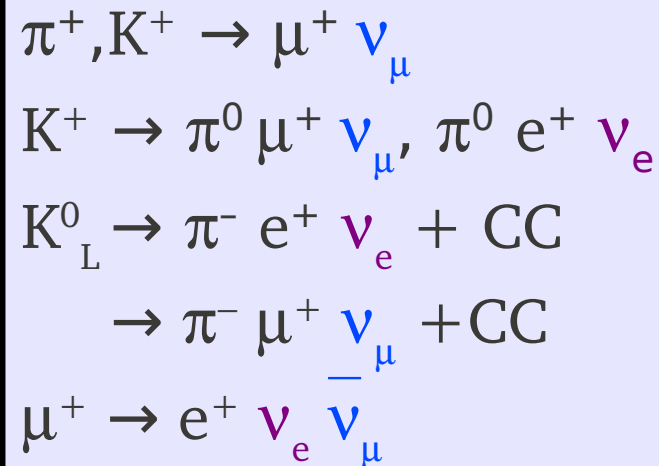
Neutrino Beams

NuMI @ FNAL



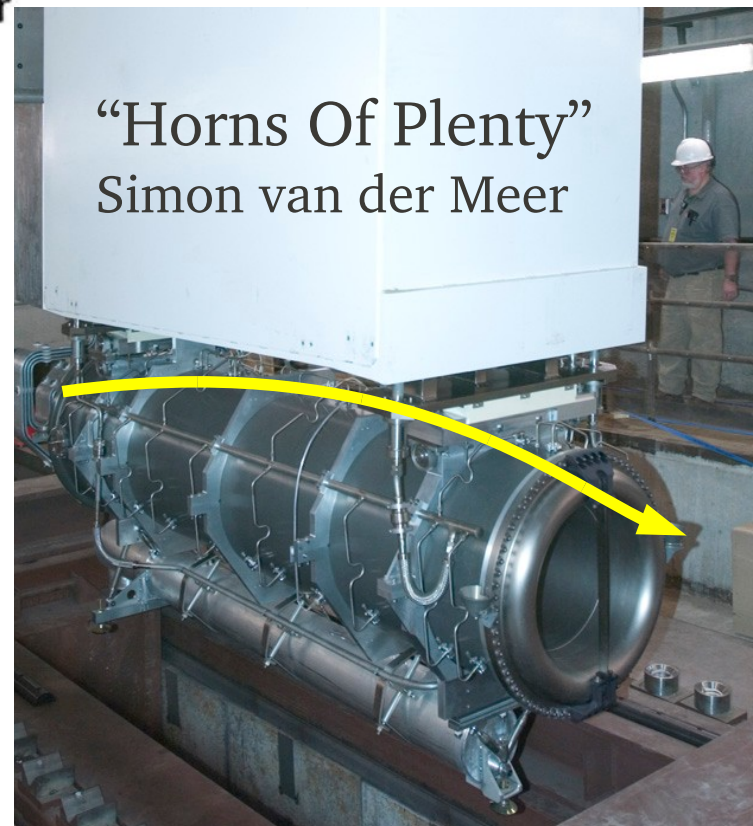
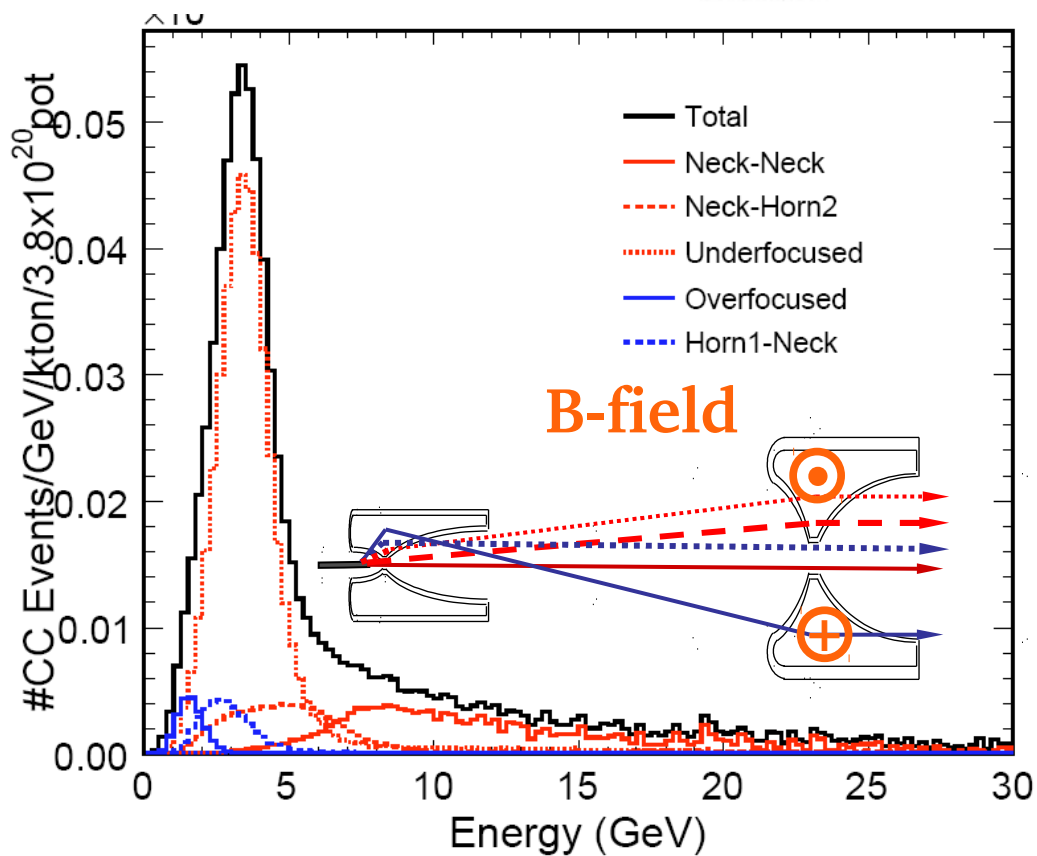
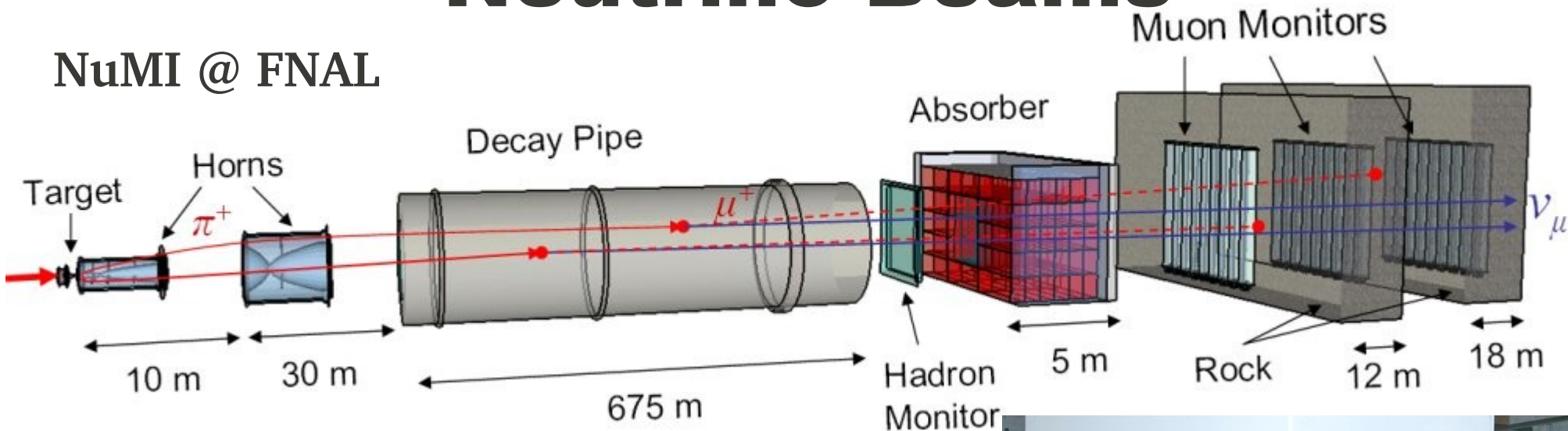
- π, K production off a solid target
- Wide range of p_t, p_z
- Cross-sections not well known

Important decay modes



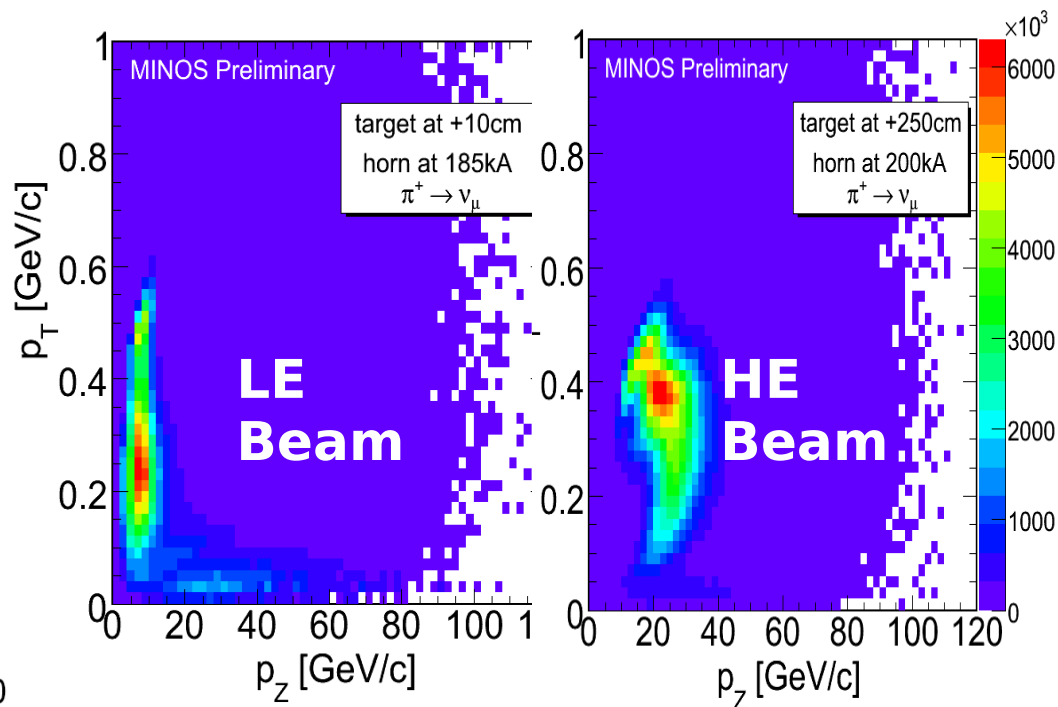
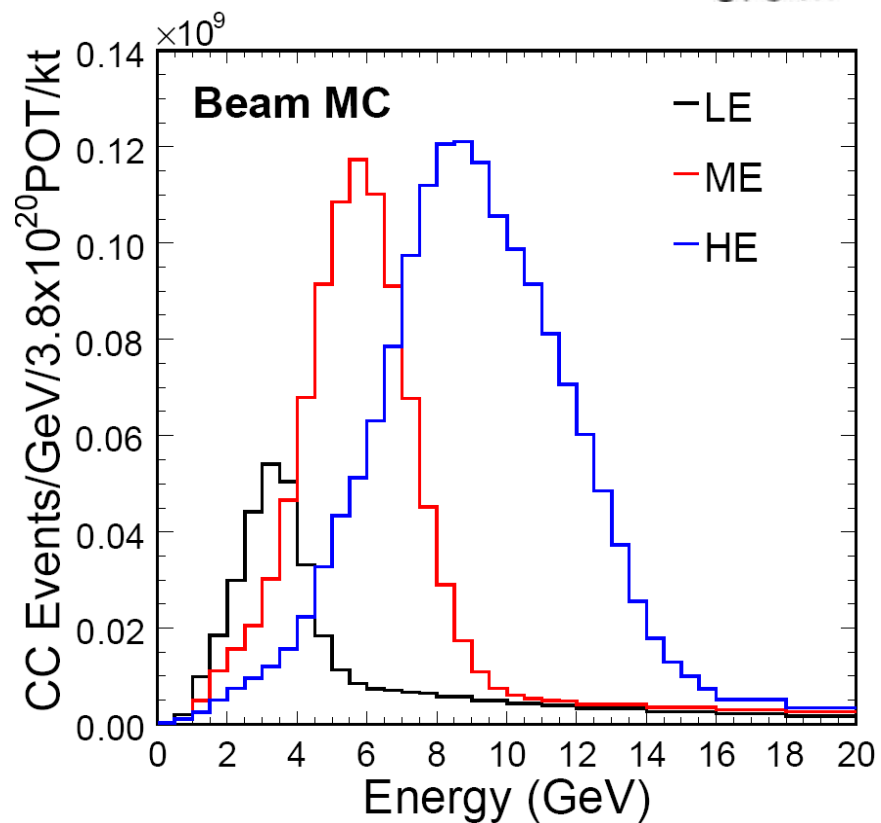
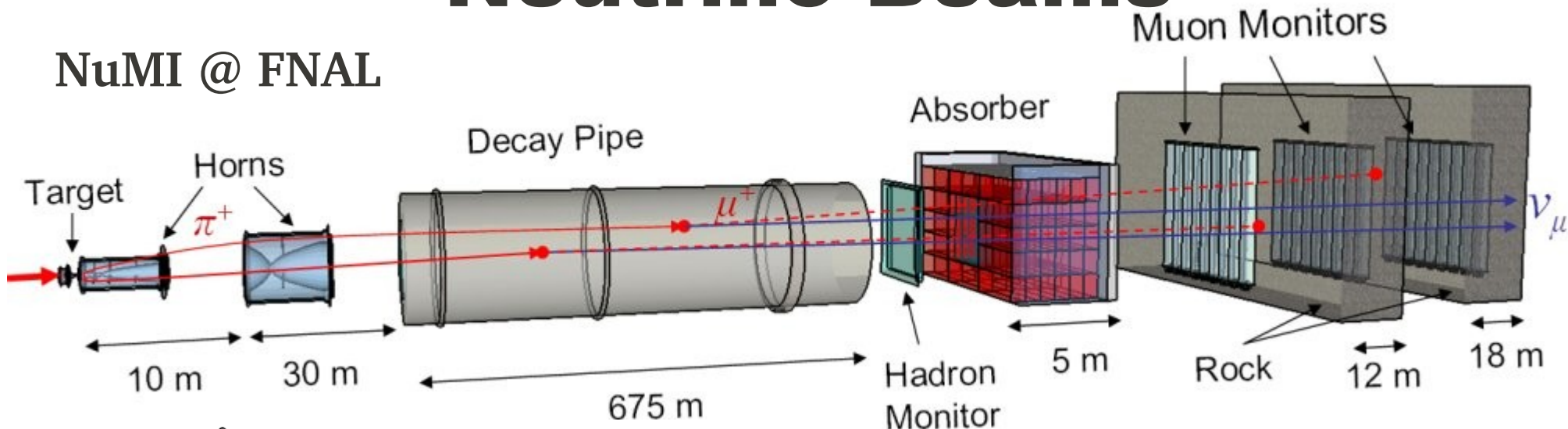
Neutrino Beams

NuMI @ FNAL



Neutrino Beams

NuMI @ FNAL



Muon neutrino disappearance

PMNS Mixing Matrix

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{bmatrix}$$

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)

$$P(\nu_\mu \rightarrow \nu_\mu) = 1 - 4 |U_{\mu3}|^2 |U_{\mu1}|^2 \sin^2 \Delta_{31} \\ - 4 |U_{\mu3}|^2 |U_{\mu2}|^2 \sin^2 \Delta_{32} \\ - 4 |U_{\mu2}|^2 |U_{\mu1}|^2 \sin^2 \Delta_{21}$$

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

Characteristic oscillatory behavior
depends on Δm^2 and L/E

“Survival Probability”

For a neutrino of energy E a distance L from the source

Muon neutrino disappearance

PMNS Mixing Matrix

$$U = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{bmatrix}$$

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

$$P(\nu_{\mu} \rightarrow \nu_{\mu}) = 1 - 4|U_{\mu3}|^2|U_{\mu1}|^2 \sin^2 \Delta_{31} \\ - 4|U_{\mu3}|^2|U_{\mu2}|^2 \sin^2 \Delta_{32} \\ - 4|U_{\mu2}|^2|U_{\mu1}|^2 \sin^2 \Delta_{21}$$

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

Muon neutrino disappearance

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

“Atmospheric oscillations”

$L/E \sim 500 \text{ km/GeV}$

Limiting Case

$$\theta_{13} = 0$$

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

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

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

Muon neutrino disappearance

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

“Atmospheric oscillations”

$L/E \sim 500 \text{ km/GeV}$

Limiting Case

$$\theta_{13} = 0$$

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

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

$$\begin{array}{l} \swarrow \\ \approx |\Delta m_{32}^2| \\ \searrow \\ \approx \theta_{23} \end{array}$$

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

Muon neutrino disappearance

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

“Atmospheric oscillations”

$L/E \sim 500 \text{ km/GeV}$

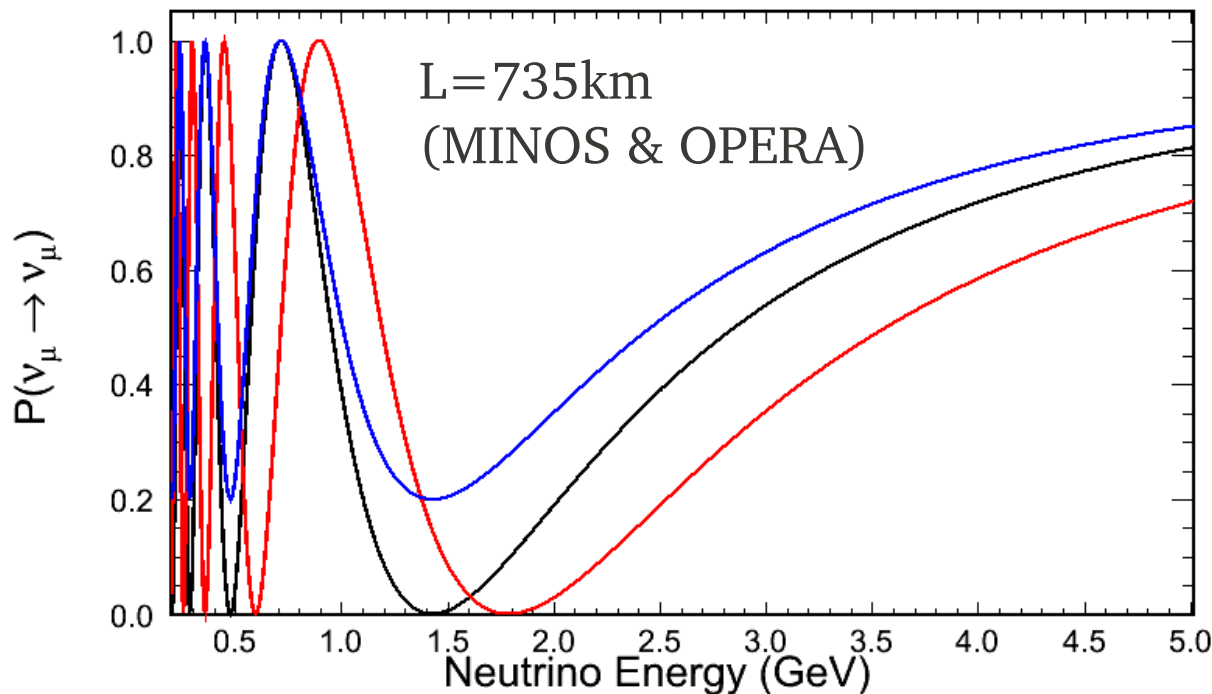
Limiting Case

$$\theta_{13} = 0$$

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

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

$$\begin{aligned} &\approx |\Delta m^2_{32}| \\ &\approx \theta_{23} \end{aligned}$$



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

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

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

Muon neutrino disappearance

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

“Atmospheric oscillations”

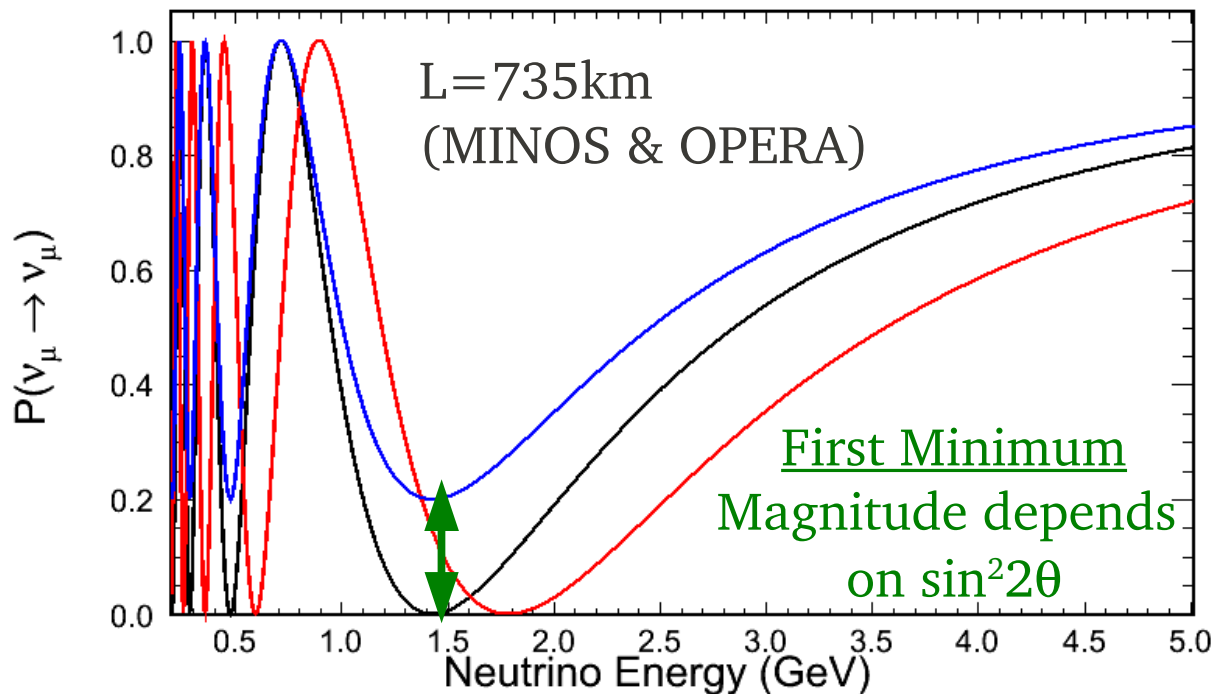
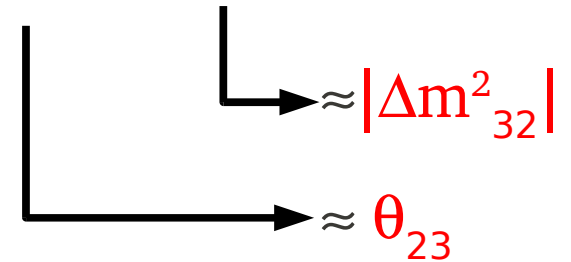
$L/E \sim 500 \text{ km/GeV}$

Limiting Case

$$\theta_{13} = 0$$

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

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



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

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

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

Muon neutrino disappearance

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

“Atmospheric oscillations”

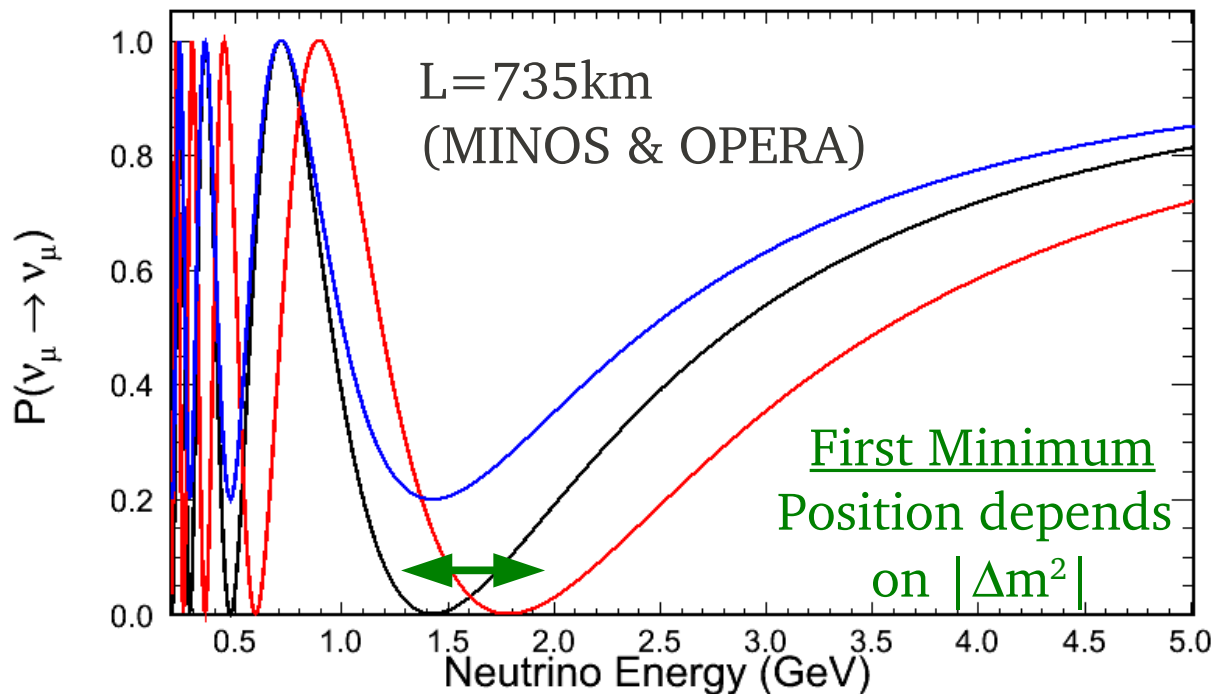
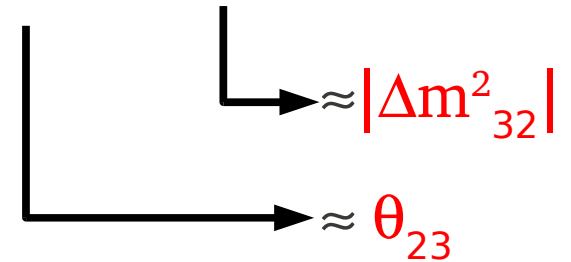
$L/E \sim 500 \text{ km/GeV}$

Limiting Case

$$\theta_{13} = 0$$

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

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



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

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

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

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 π, 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_ν .
- ν_μ, ν_e -CC & NC extrapolate differently.
- Some backgrounds may oscillate.

Case study: MINOS

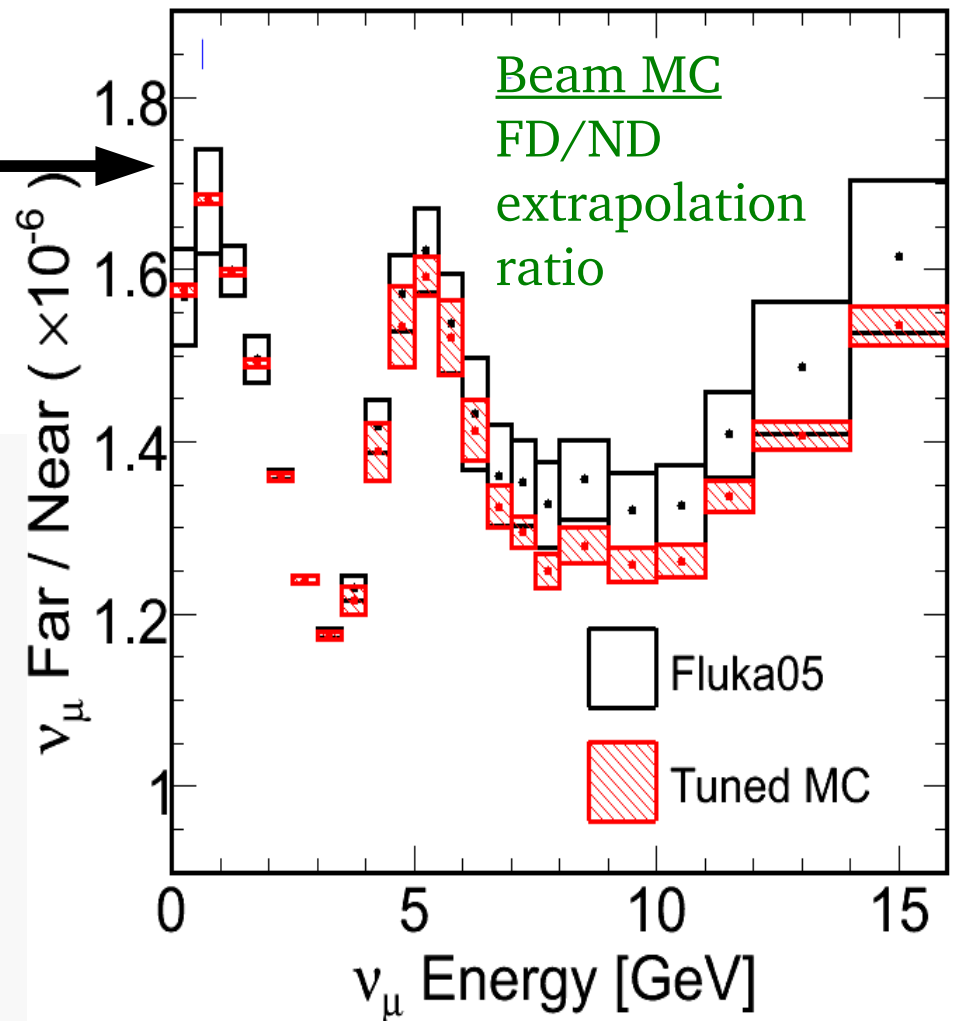
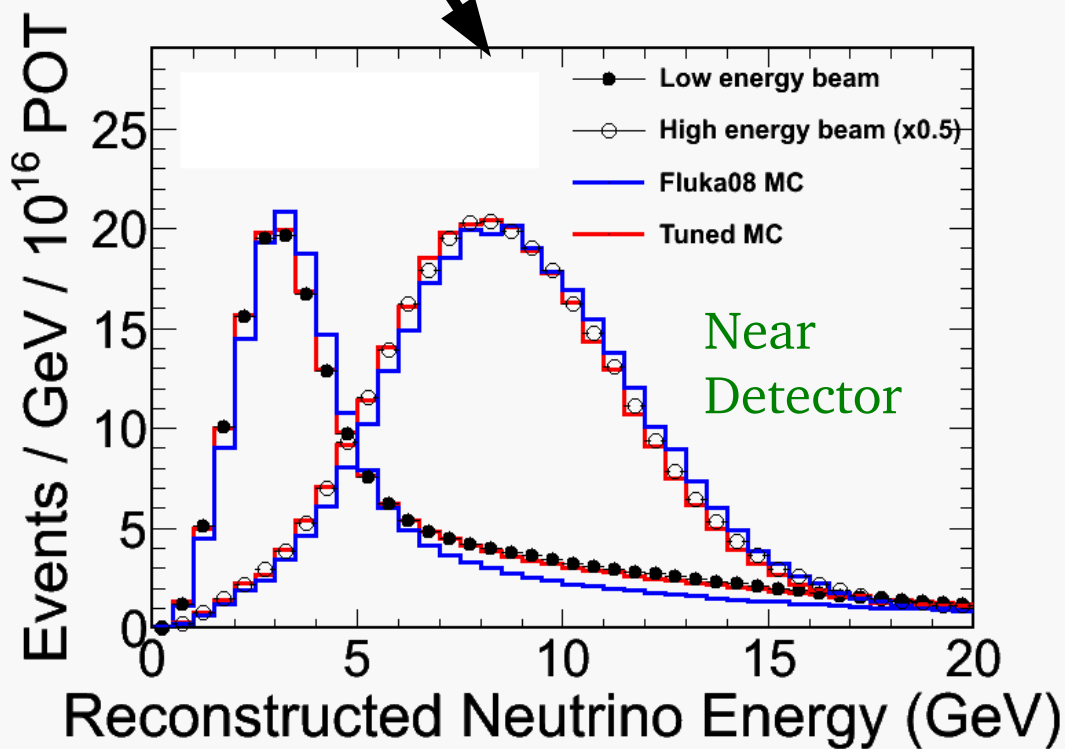
$$\nu_{\mu} \rightarrow \nu_{\mu}$$

- 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 ν_{μ} -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 ν_{μ} -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

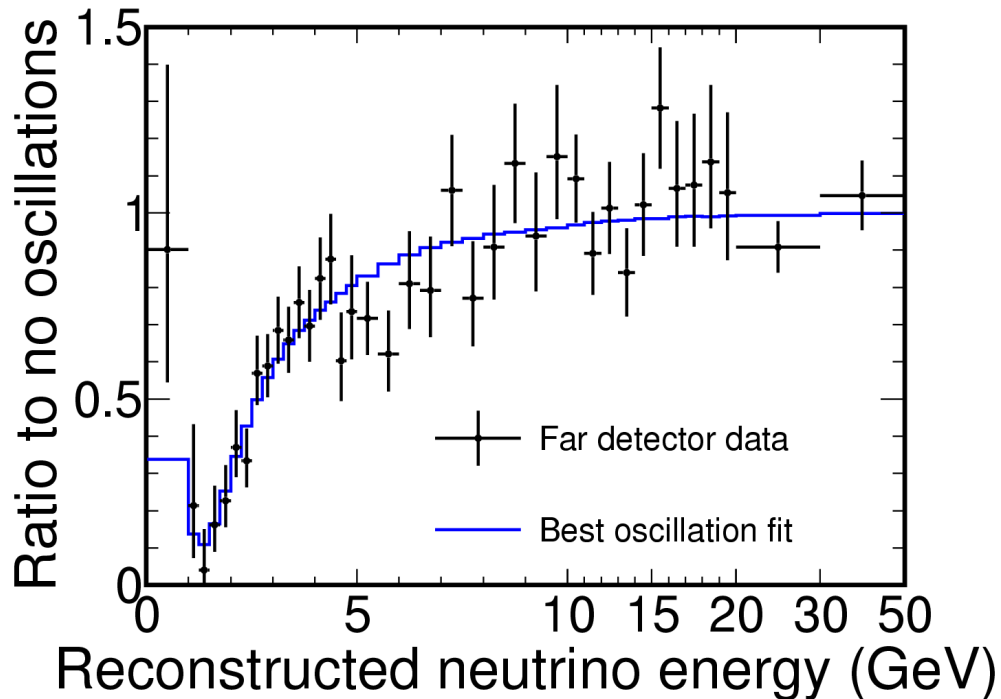
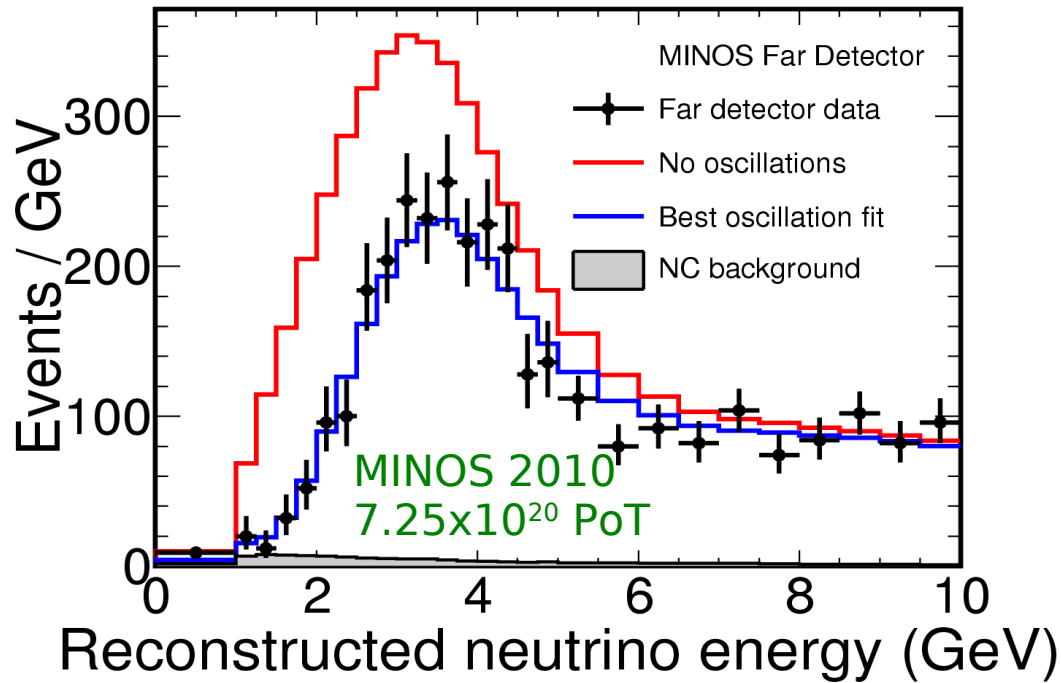


Case study:

MINOS $\nu_{\mu} \rightarrow \nu_{\mu}$

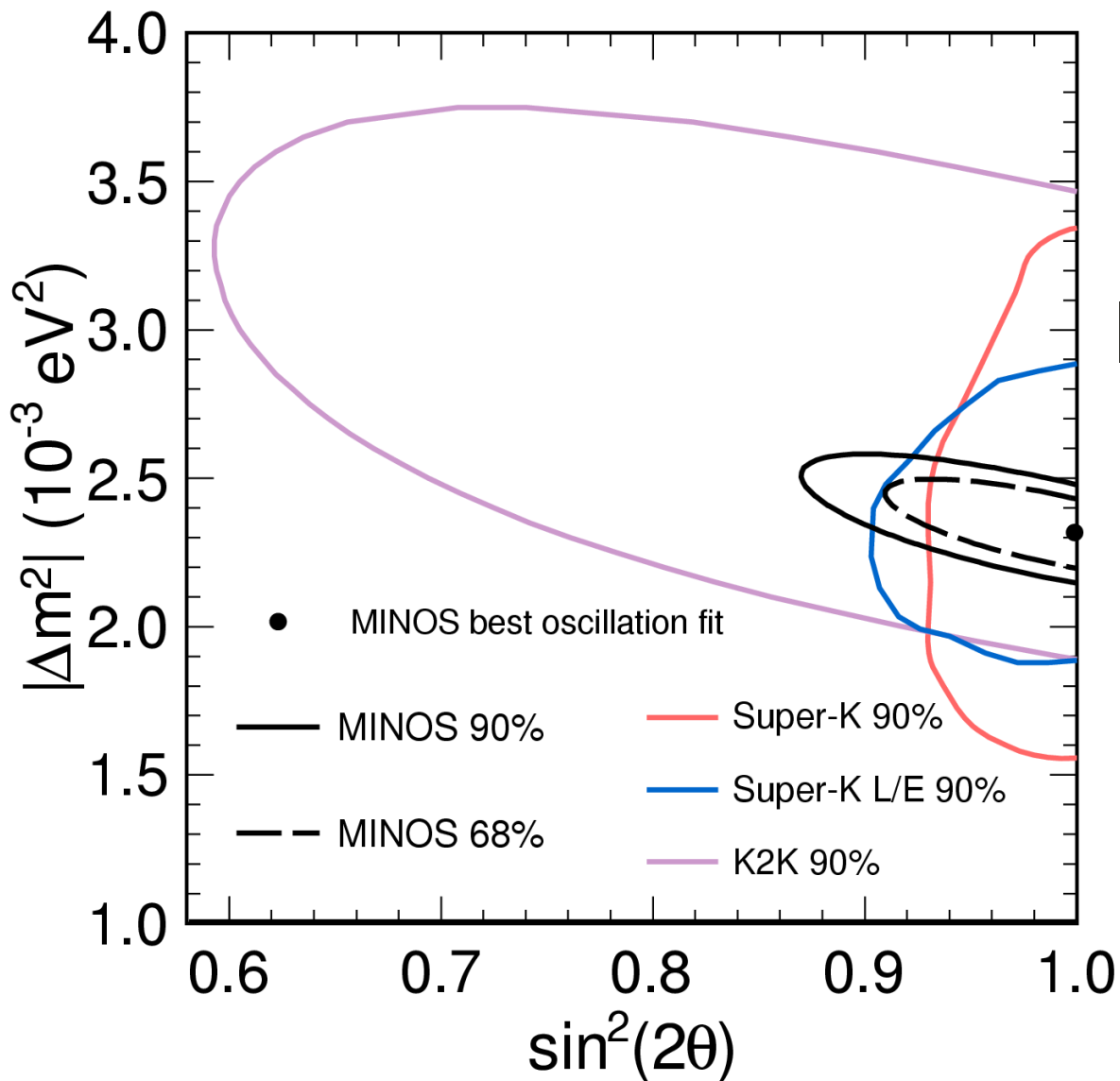
arXiv: 1103.0340v1

- 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 7.25×10^{20} PoT

Fit result

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

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

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 $\sim 100\%$ 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: (ν_{μ} disappearance)
 - sntp reco data: 2.0T sntp reco MC: 10T
- Generation:
 - GEANT3/NEUGEN3 based program → But, ROOT TVirtualMC
 - 1×10^{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 ν_μ, ν_e, ν_τ , 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 (ν_τ 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.