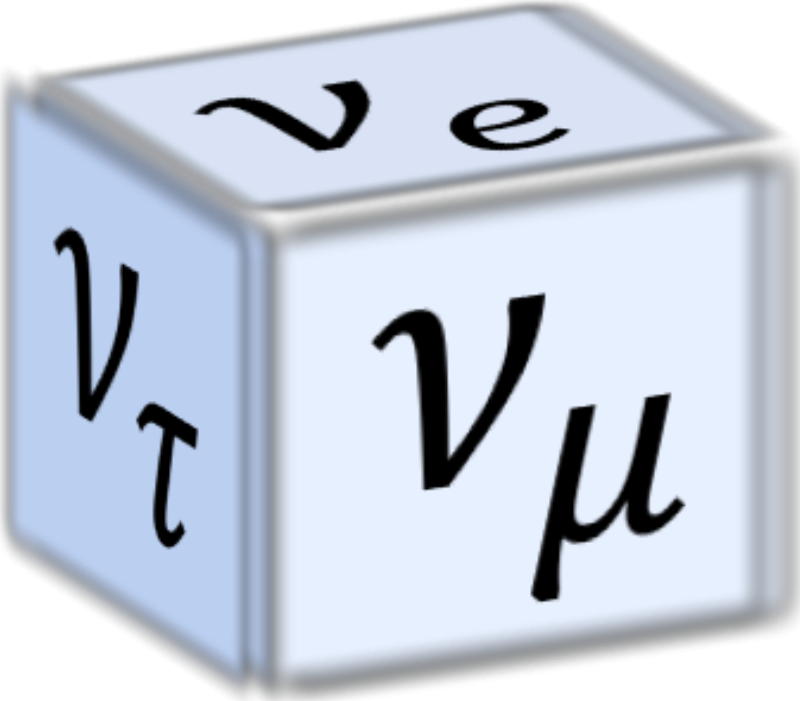
“Neutrino Summary” of PHYSTATnu

Kevin McFarland
University of Rochester

25 January 2019

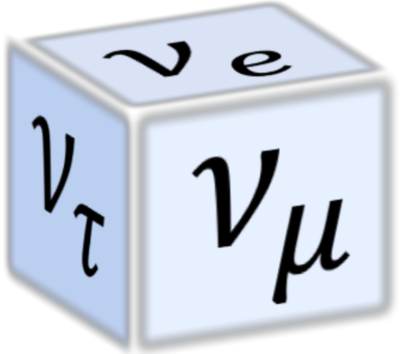


Summary::Summary

1. Preliminaries
2. Problems in neutrino physics
3. Extracting results from data
4. Confessions of neutrino physicists
5. Learning to live without square wheels
6. Parting Thoughts

“Who am I? Why am I here?”

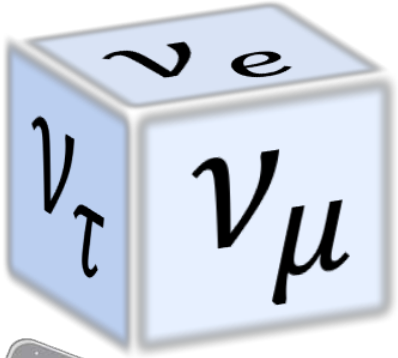
- Admiral James Stockdale, Candidate for Vice President of USA, 1992



- I’ve done neutrino experiments for a while, although unlike Prof. Blondel, this does not require a reference to my thesis experiment.
  - (If you are curious, I found that the lightest meson decays to a pair of the lightest charged fermions, with 9 events on a predicted background of 1. I measured a branching ratio and didn’t worry if that constituted a “discovery”.)
- My first neutrino adventure was “high energy” deep inelastic neutrino scattering and studies of neutrino neutral current couplings.
- I currently work on accelerator neutrino oscillation experiments.
- I am particularly interested in the uncertainties in neutrino interactions and how they affect those experiments.

“Who am I? Why am I here?” (cont’d)

- Admiral James Stockdale, Candidate for Vice President of USA, 1992



- The second question may be more difficult to answer.
- My attitude toward thinking carefully about statistical issues in neutrino measurements has largely been to favor “just in time” delivery of insight.

“Who am I? Why am I here?” (cont’d)

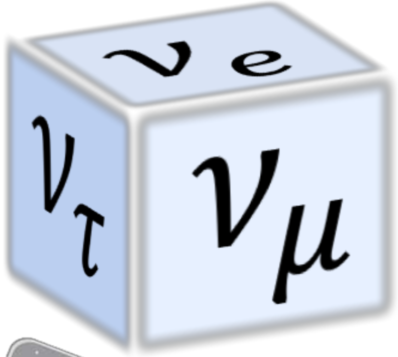
- Admiral James Stockdale, Candidate for Vice President of USA, 1992

or

I confess that when it comes to statistics or lobster,
I’m the one on the right.

"Who am I? Why am I here?" (cont'd)

- Admiral James Stockdale, Candidate for Vice President of USA, 1992

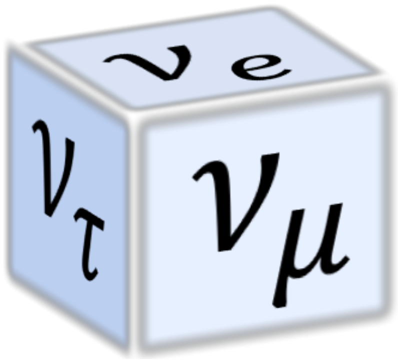


- The second question may be more difficult to answer.
- My attitude toward thinking carefully about statistical issues in neutrino measurements has largely been to favor "just in time" delivery of insight.

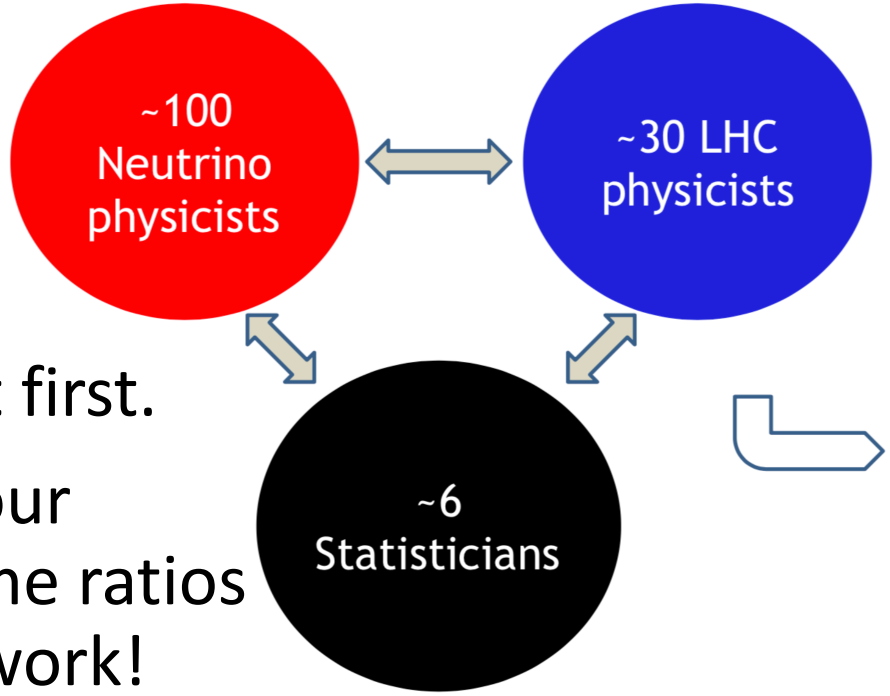
- However…
  … there are a number of clear problems in my area of interest that have been bothering me. For me, one in particular.
  - The interplay between flux and interaction models and near detector constraints in oscillation experiments.

- It's probably just about time.

Then, am I in the right place?



- We learned about the demographics of our meeting from Olaf's introduction.
- The asymmetry worried me at first.
- However I then realized that our population is nearly in the same ratios as quantities relevant for my work!



Aiming at a fruitful exchange between the three communities!

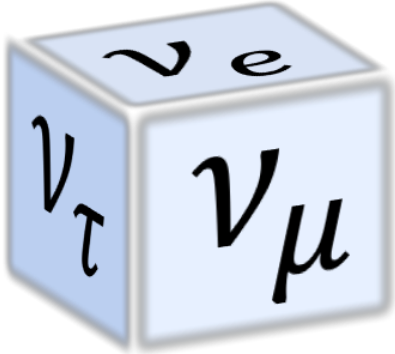
$$(m_\Delta - m_N) : m_\mu : E_{\text{"binding"}}$$

Energy to excite a nucleon

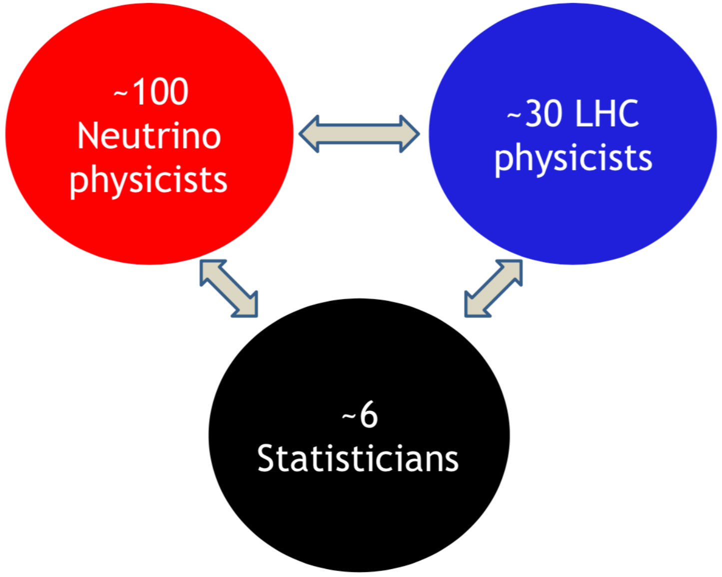
Muon mass

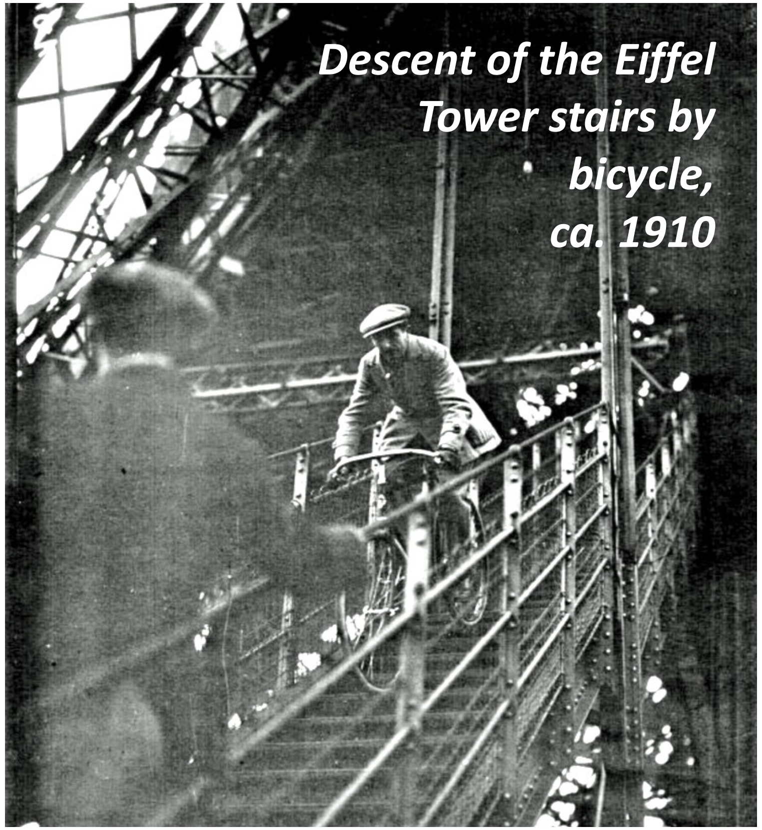
Removal energy of a nucleon in a nucleus

Then, am I in the right place? (cont'd)



- Sometimes I refer to this in the context of neutrino-nucleon interaction physics as "a failed multiscale problem".





Descent of the Eiffel Tower stairs by bicycle, ca. 1910

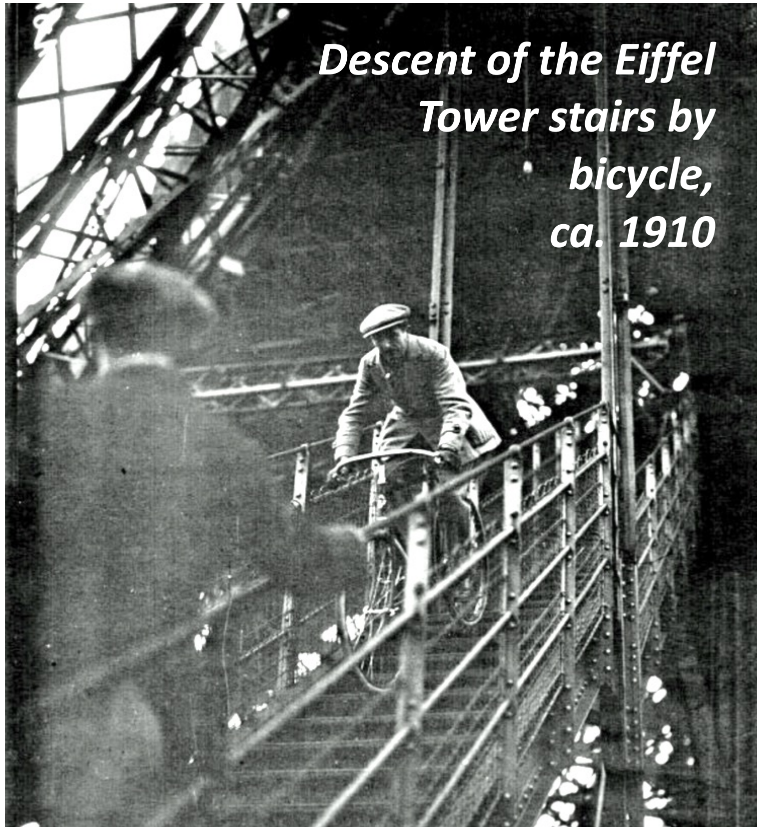
$$(m_\Delta - m_N) : m_\mu : E_{\text{"binding"}}$$

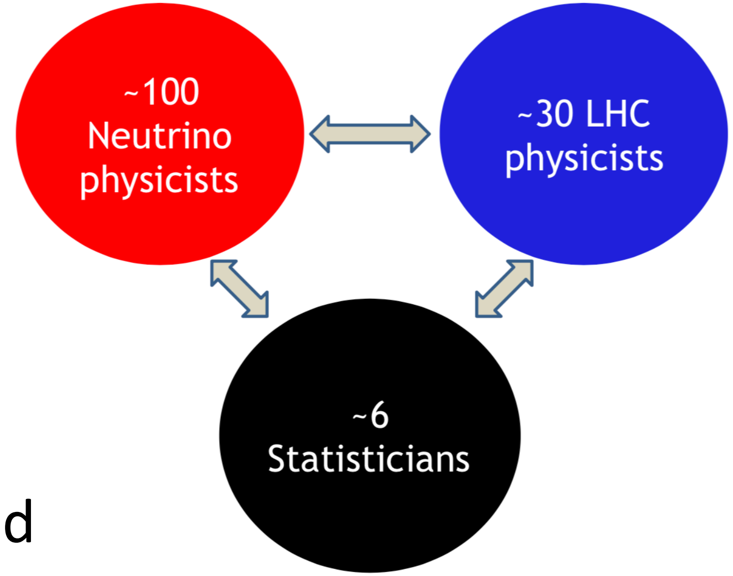
- Consider a bicycle rider at right, descending the stairs of the Eiffel Tower.
- A bicycle wheel is ~1m in diameter
- If steps were ~1cm or ~100m in height, we could perfectly predict the cyclist's trajectory
- *Since wheel size is close to step size, all our theoretical tools allow us to predict is that the outcome is going to be painful.*

Then, am I in the right place? (cont'd)

- Sometimes I refer to this in neutrino-nucleon interaction physics as "a failed multiscale problem".

$$(m_{\Delta} - m_N) : m_{\mu} : E_{\text{"binding"}}$$





- Similarly, it's true that the trajectory of our meeting wasn't perfectly predictable.
- This is a good thing! We didn't spend three days in the decoupling limit!
  - *That said, I think we can agree that there were a few painful parts to our valuable discussions.*

Then, am I in the right place? (cont'd)

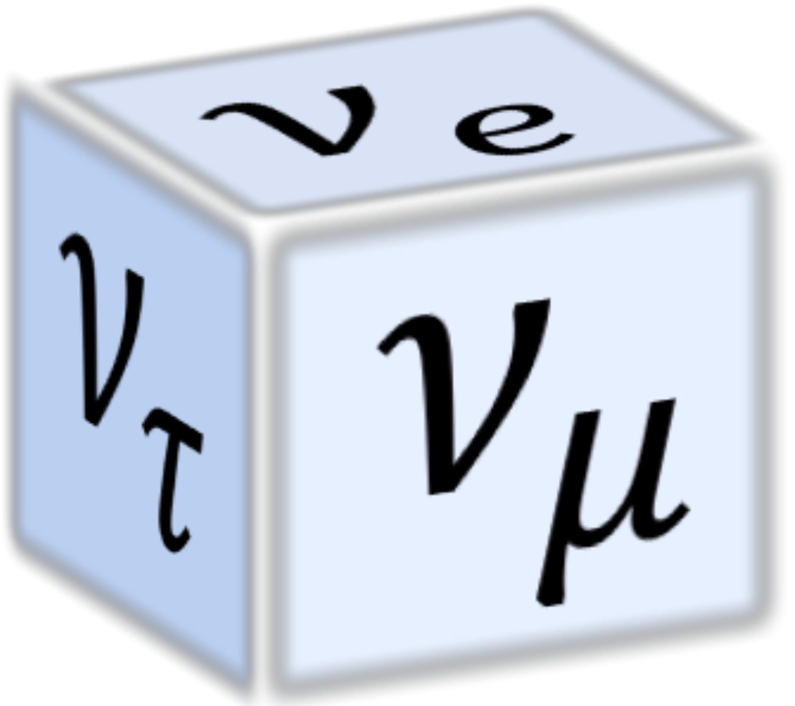
- This is a good thing! We didn't spend three days in the decoupling limit!

- *That said, I think we can agree that there were a few painful parts to our valuable discussions.*

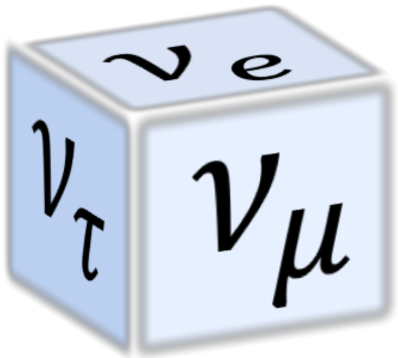
And not only your summary speakers having to scrap most of what they wrote in advance…

Statistical Problems of Neutrino Physicists

“I want you to be open and honest and not to leave any hairs on the couch”



Neutrino Physics is a Vast Field with Many Interesting Goals



- We heard about accelerator and natural neutrinos of different energies, neutrinos from supernovae and reactors, and single and double beta decay.

Supernova neutrino signal detection in NOvA
Speaker: Andrey Sheshukov (JINR)

Short-Baseline experiments
Speaker: Prof. Georgia Karagiorgi (CU)

Reactor experiments
Speaker: Chao Zhang (BNL)

Atmospheric neutrinos
Speaker: Christophe Bronner (University of Tokyo)

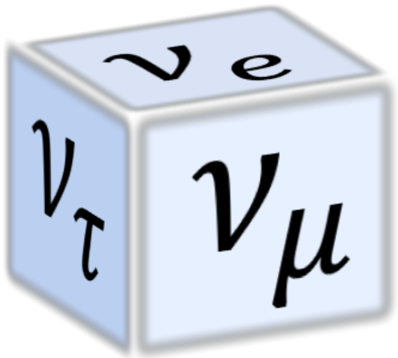
Long-baseline experiments
Speaker: Alexander Himmel (Fermilab)

Cosmological (high energy) neutrinos
Speaker: Tim Ruhe (TU Dortmund)

Neutrino-less double-beta decay experiments
Speaker: Dr Matteo Agostini (TUM)

Direct neutrino mass measurement
Speaker: Thierry Lasserre (CEA)

Neutrino Physics is a Vast Field with Many Interesting Goals



- We heard about accelerator and natural neutrinos of different energies, neutrinos from supernovae and reactors, and single and double beta decay.
  - "All happy families are alike; each unhappy family is unhappy in its own way."
    - Tolstoy, *Anna Karenina*

Supernova neutrino signal detection in NOvA
Speaker: Andrey Sheshukov (JINR)

Short-Baseline experiments
Speaker: Prof. Georgia Karagiorgi (CU)

Reactor experiments
Speaker: Chao Zhang (BNL)

Atmospheric neutrinos
Speaker: Christophe Bronner (University of Tokyo)

Long-baseline experiments
Speaker: Alexander Himmel (Fermilab)

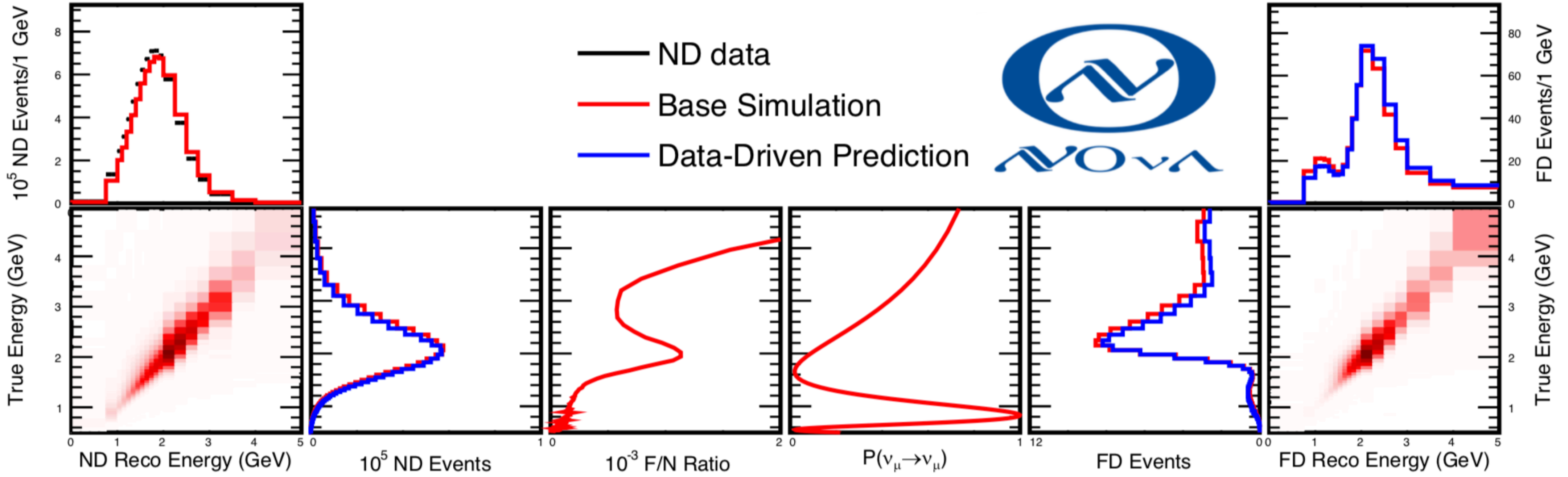
Cosmological (high energy) neutrinos
Speaker: Tim Ruhe (TU Dortmund)

Direct neutrino mass measurement
Speaker: Thierry Lasserre (CEA)

Neutrino-less double-beta decay experiments
Speaker: Dr Matteo Agostini (TUM)

Some neutrino measurements are incredibly complex

Long-baseline experiments
Speaker: Alexander Himmel (Fermilab)

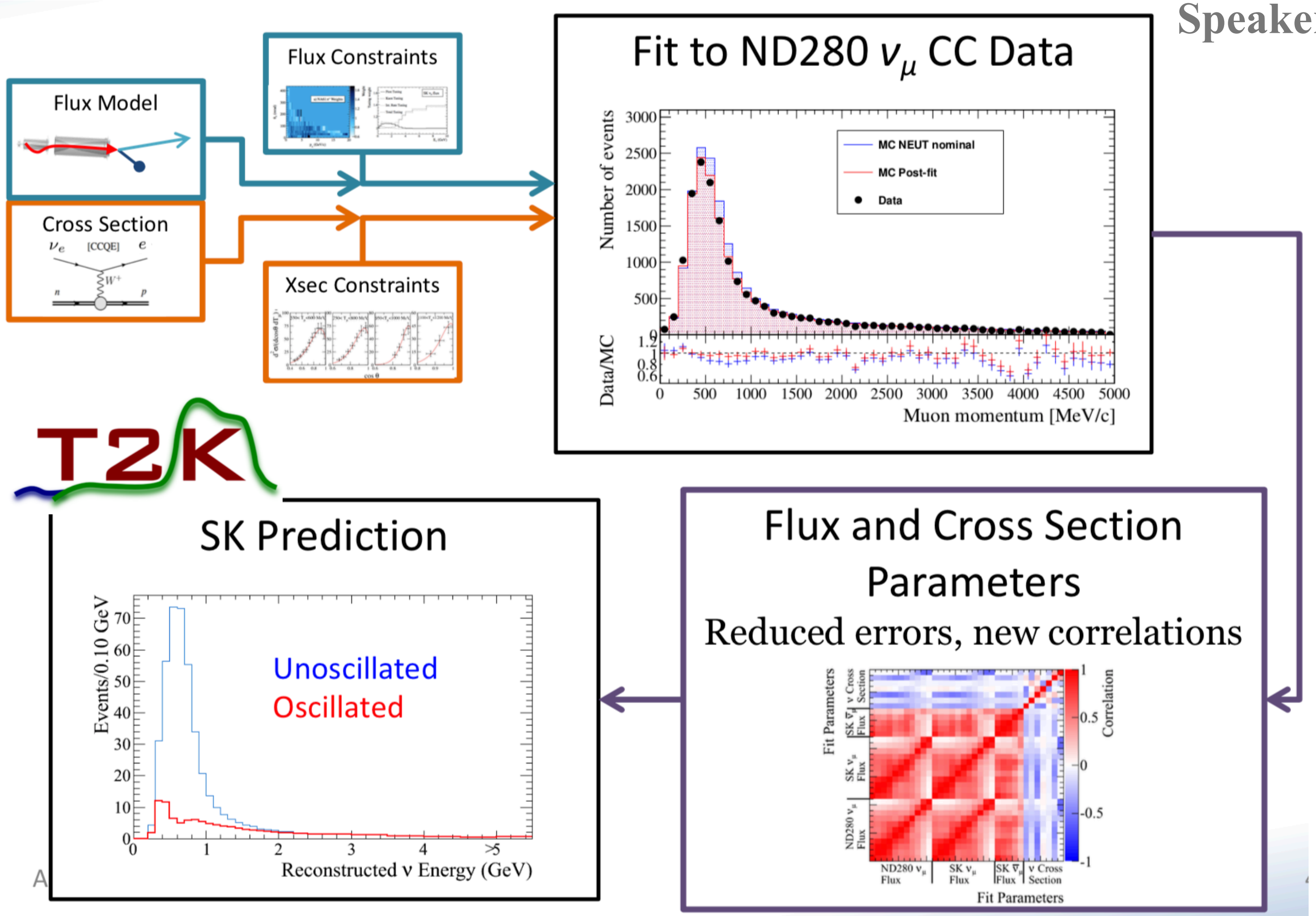


- Use the ND ν_μ sample to predict the FD ν_μ sample.
- Reco vs True energy uses interaction model, detector model.
- F/N Ratio includes flux simulation

Both have a lot of detail, and uncertainty, underneath.

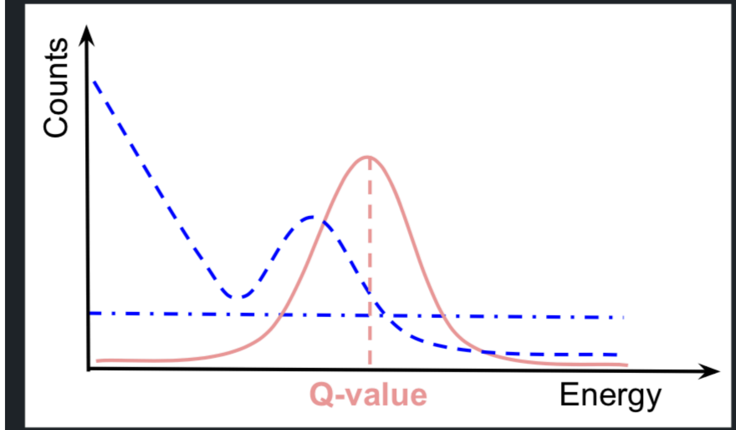
Some neutrino measurements are incredibly complex

Long-baseline experiments
Speaker: Alexander Himmel (Fermilab)



- Similar inputs at T2K
- Different fit mechanics, and statistical treatment (profiling vs brute force marginalization)
- Both measurements have many (hundreds) of nuisance parameters in model to extract a small number of Parameters Of Interest.

Elements of these experiments are difficult to control



Signal:

- ➢ spread due to energy resolution
- ➢ peak at Q-value (often Gaussian)

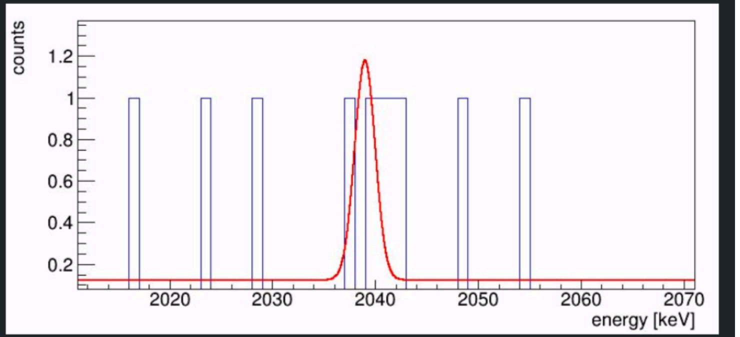
Background:

- ➢ typically flat if resolution is <0.1%
- ➢ not connected to physics mechanism generating the signal -> hard to model!

Statistical analysis

Search for a peak at fixed position

- ➢ λ_S signal expectation (typically very small)
- ➢ λ_B background expectation
- ➢ **no look elsewhere**
- ➢ background control sample (on/off problem)



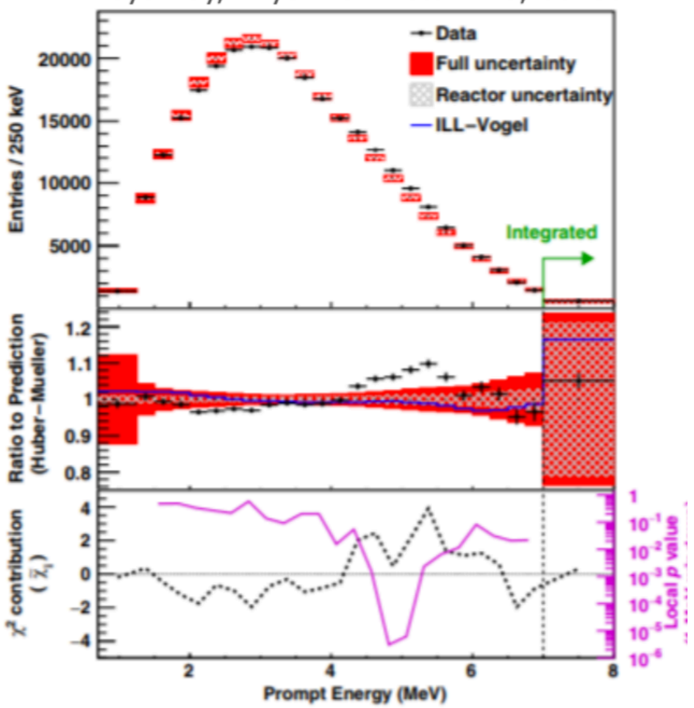
Neutrino-less double-beta decay experiments

Speaker: Dr Matteo Agostini (TUM)

- In case of $0\nu\beta\beta$ experiments, signal is well understood.
- Background, however, is not modeled a priori.
  - Problems with measuring in “deep Poisson” limit.
- Not unique situation to $0\nu\beta\beta$!
  - We saw many cases where data itself had to be used to draw inferences about un(der)known processes.

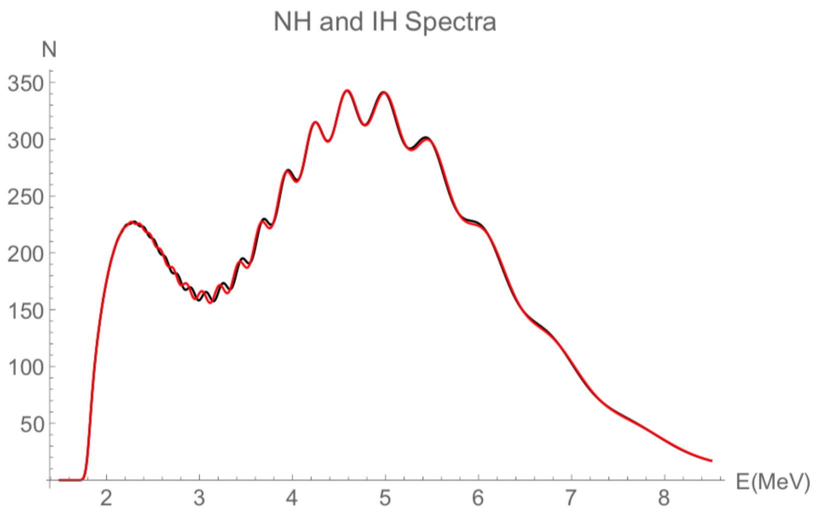
Poorly modeled inputs affect precision measurements

Daya Bay, Phys. Rev. Lett. **116**, 061801



Reactor experiments
Speaker: Chao Zhang (BNL)

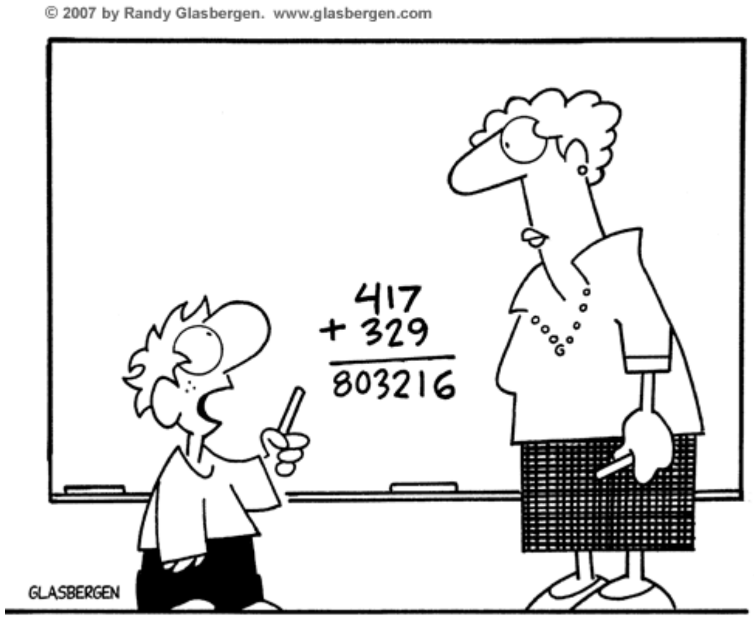
Uncertainty in the Reactor Neutrino Spectrum and the Mass Hierarchy
Speaker: Emilio Ciuffoli (IMP, CAS)



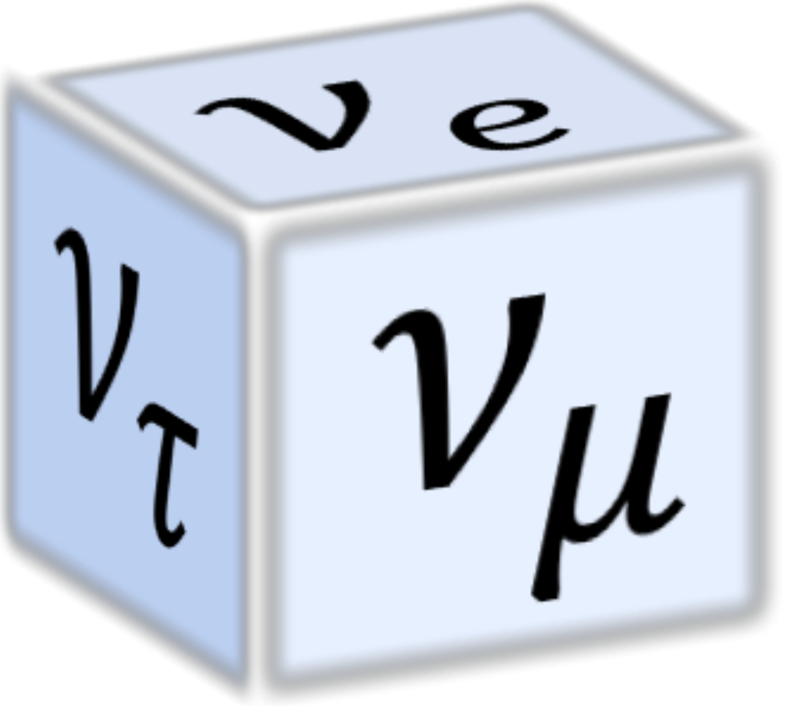
Expected spectra for normal and inverted hierarchy at 53km. Inverted hierarchy: Δm^2_{23} shifted (by $\simeq 1\sigma$'s). Finite energy resolution

- Precision measurement (JUNO) must survive a large discrepancy in its input.
  - ~5 MeV flux anomaly
- Nearly perfect parallels in T2K and NOvA in observations at the near detector that invalidate default interaction model.

Getting the Answer from the Data



My dad says persistence is the key to success, so I'm going to keep giving you the same wrong answer until it becomes the right answer."



Techniques for Interpreting Data

- We discussed unfolding and forward folding, signal processing, machine learning, combining results.
- Many interesting threads to pull on. I'll pick a few.

Uncertainty in the Reactor Neutrino Spectrum and the Mass Hierarchy
Speaker: Emilio Ciuffoli (IMP, CAS)

The neutrino experiment experience with unfolding
Speaker: Philip Andrew Rodrigues (University of Oxford

Statistical issues on the neutrino Mass Hierarchy determination
Speaker: Fatma Sawy (INFN Padova)

Look-elsewhere effect in neutrino oscillation searches
Speaker: Phillip Litchfield (Imperial College, London)

Review of Linear Algebra Applications in Some Recent Neutrino Experiments
Speaker: Xin Qian (BNL)

Simulating Light in Large Volume Detectors Using Metropolis Light Transport
Speaker: Gabriel Collin

Regularisation for T2K cross-section analyses
Speaker: Stephen Dolan (LLR / CEA Saclay)

The neutrino experiment experience [combining data]
Speaker: Prof. Constantinos Andreopoulos (Liverpool, STFC/RAL)

Reactor Anti-neutrino Data in Global Analyses
Speaker: Alvaro Hernandez Cabezudo (KIT)

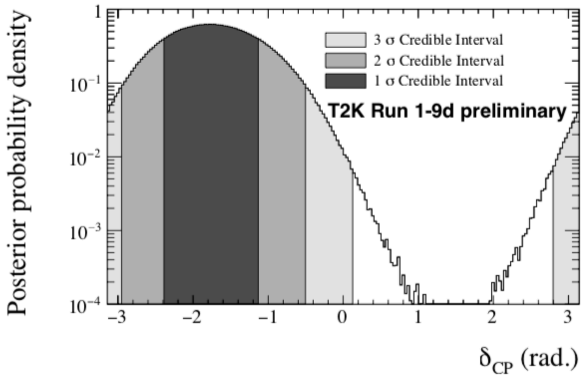
The neutrino experiment experience [machine learning]
Speaker: Saul Alonso Monsalve (University Carlos III)

Machine Learning methods for JUNO Experiment
Speaker: Yu Xu

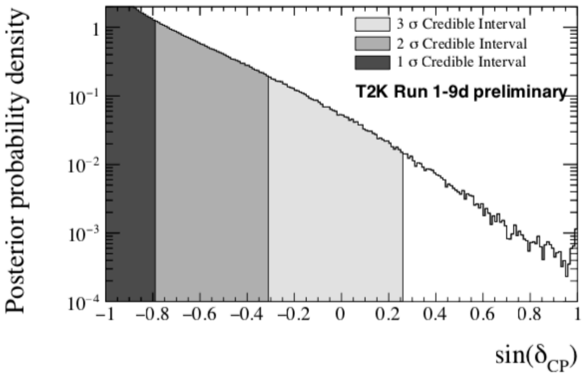
Fits to large data sets
Speaker: Dr Stefano Gariazzo (IFIC Valencia)

Multiple Statistical Methods and One Observable

Flat in δ



Flat in sin δ



	δ			sin(δ)	
1σ	-2.39	to -1.13	-1	to	-0.68 or -0.91
2σ	-2.95	to -0.50	-1	to	-0.19 or -0.48
3σ	2.80	to 0.13	-1	to	0.33 or 0.10

	sin(δ)		
1σ	-1	to	-0.79
2σ	-1	to	-0.32
3σ	-1	to	0.25

- most experiments quote results from multiple methods and give enough info to reproduce the analysis
- sensitivity always reported (sometimes also for Bayesian methods)
- blind analysis is almost the standard
- frequentist intervals still used as Bayesian intervals (even when Bayesian interval is available)
- sensitivity computed for the no signal hypothesis, more interesting to quote discovery power

[Phys. Rev. Lett. 120, 132502 (2018)]

Statistical Method in the last PRL of the MAJORANA DEMONSTRATOR	$T_{1/2}$ lower limit 90% prob [10^{25} yr]	$T_{1/2}$ lower limit sensitivity [10^{25} yr]
Counting (FC)	1.6	
Unbinned likelihood fit (FC)	1.9	2.1
Unbinned likelihood fit & CLs	1.5	1.4
Bayesian flat prior	1.6	
Bayesian Jeffreys prior	2.6	

Long-baseline experiments
Speaker: Alexander Himmel (Fermilab)

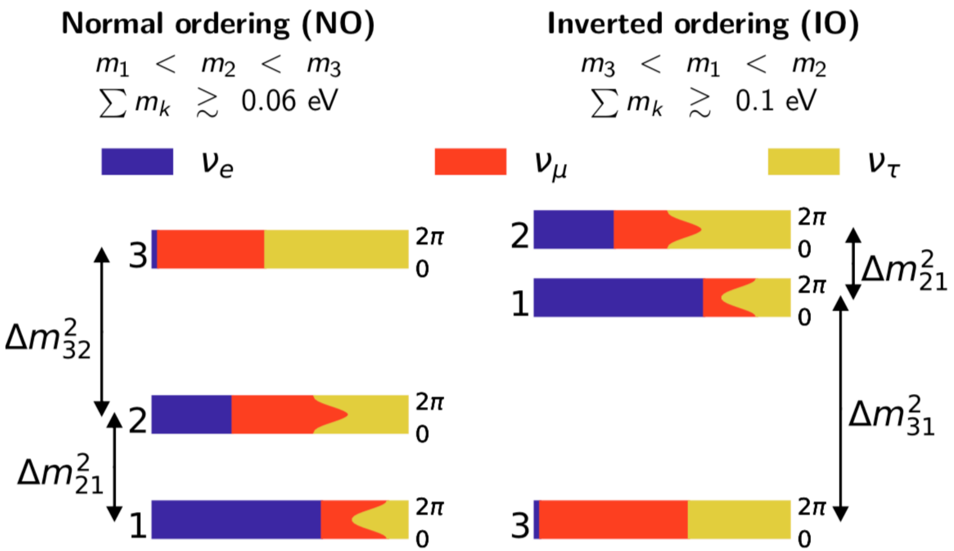
- We've explored many examples where the same data, as expected, gives different answers with different statistical approaches to interval estimation.
- Some subfields, $0\nu\beta\beta$, have an "industry standard" to facilitate comparisons.

Neutrino-less double-beta decay experiments
Speaker: Dr Matteo Agostini (TUM)

Multiple Statistical Methods and One Observable

Fits to large data sets
Speaker: Dr Stefano Gariazzo (IFIC Valencia)

- Many pieces of marginal evidence related to mass ordering.
- Why such differences? According to Phill, logarithmic prior in masses strongly favors “normal” ordering.



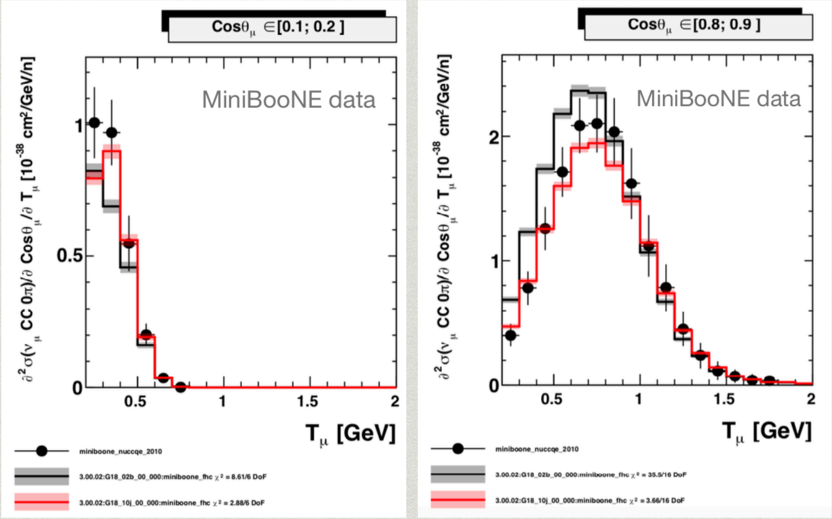
Can current data tell us the neutrino mass ordering?

1. [Hannestad, Schwetz, 2016]: extremely weak (2:1, 3:2) preference for NO (cosmology + [Bergstrom et al., 2015] neutrino oscillation fit) Bayesian approach;
2. [Gerbino et al, 2016]: extremely weak (up to 3:2) preference for NO (cosmology only). Bayesian approach;
3. [Simpson et al., 2017]: strong preference for NO (cosmological limits on $\sum m_\nu$ + constraints on Δm^2_{21} and $|\Delta m^2_{31}|$) Bayesian approach;
4. [Schwetz et al., 2017], “Comment on ...”[Simpson et al., 2017]: effect of prior?
5. [Capozzi et al., 2017]: 2σ preference for NO (cosmology + [Capozzi et al., 2016, updated 2017] neutrino oscillation fit) frequentist approach;
6. [Caldwell et al., 2017] very mild indication for NO (cosmology + neutrinoless double-beta decay + [Esteban et al., 2016] readapted oscillation results) Bayesian approach;
7. [Wang, Xia, 2017]: Bayes factor NO vs IO is not informative (cosmology only).

Combining Datasets

The neutrino experiment experience [combining data]
Speaker: Prof. Constantinos Andreopoulos (Liverpool, STFC/RAL)

- **G18_02b:** An improved empirical model in GENIE3.
- **G18_10j:** A more theory-based model configuration in GENIE3.

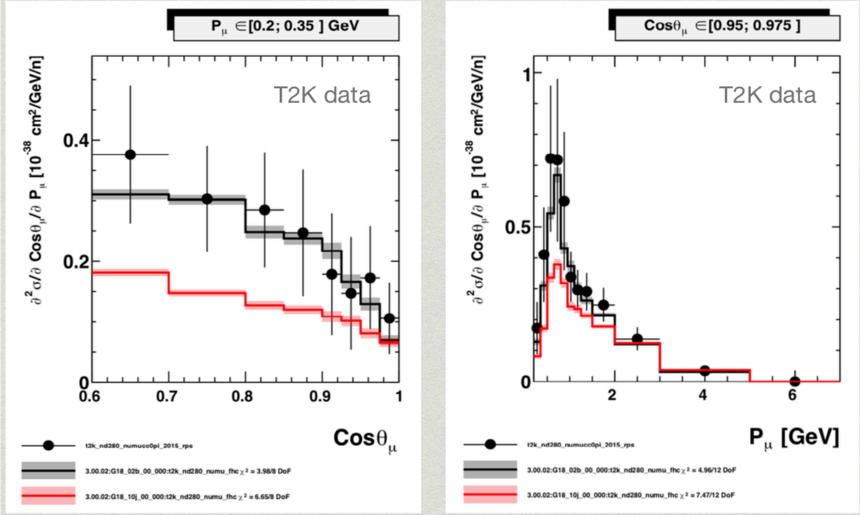


GENIE tune	χ²/ndf
G18_02b_00_000	330 / 137
G18_10j_00_000	63.7 / 137

MiniBooNE data

prefers G18_10j

- **G18_02b:** An improved empirical model in GENIE3.
- **G18_10j:** A more theory-based model configuration in GENIE3.



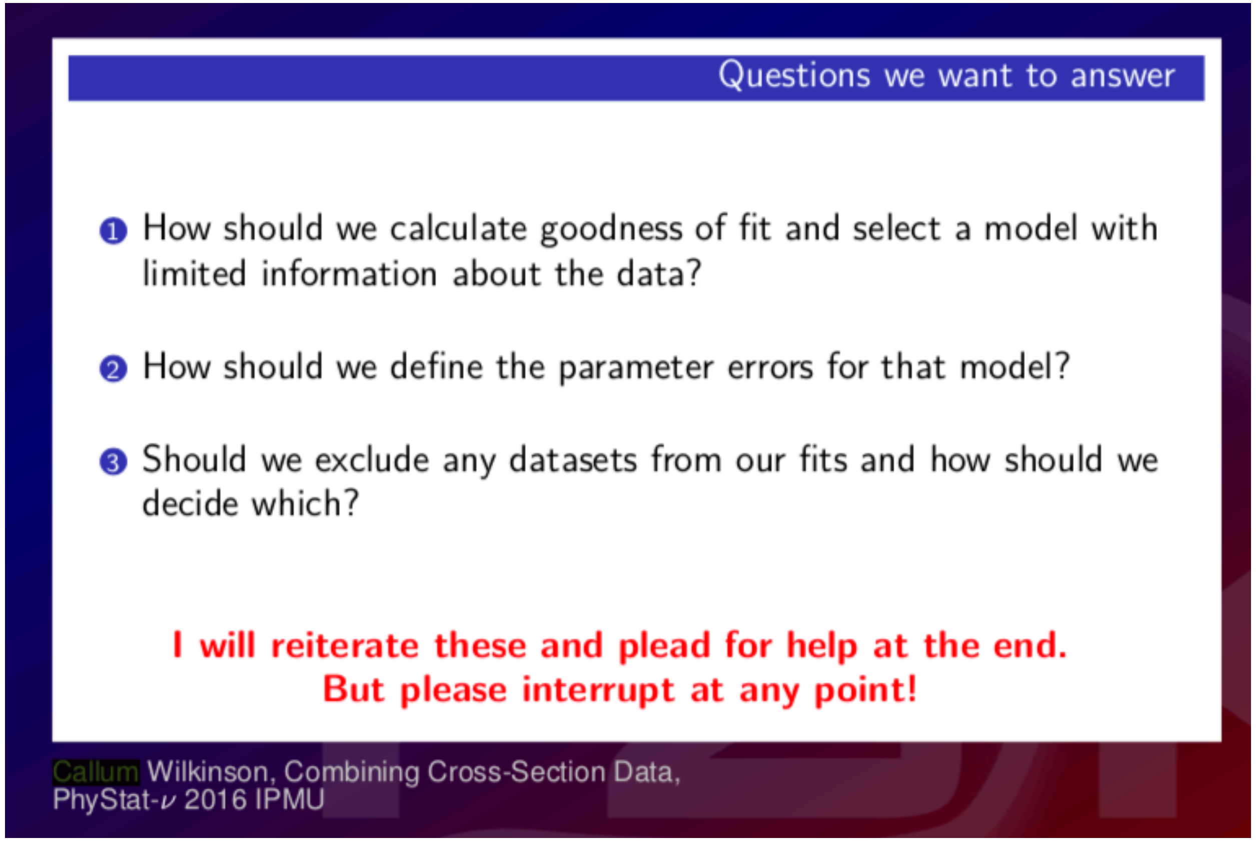
GENIE tune	χ²/ndf
G18_02b_00_000	73.9 / 80
G18_10j_00_000	80.4 / 80

T2K data

prefers G18_02b

- Tension between recent datasets, MiniBooNE and T2K $CC0\pi$ data
- One way that GENIE, in producing a reference interaction model, addresses these tensions is with "partial tunes" on consistent subsets of data.

Combining Datasets

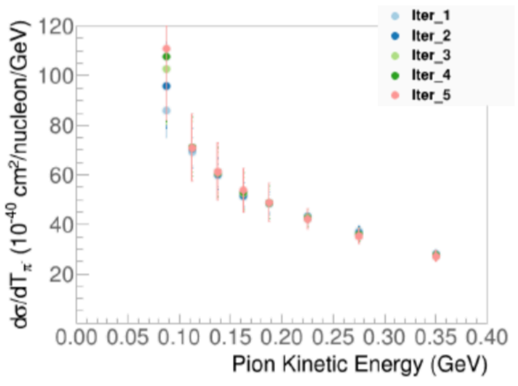


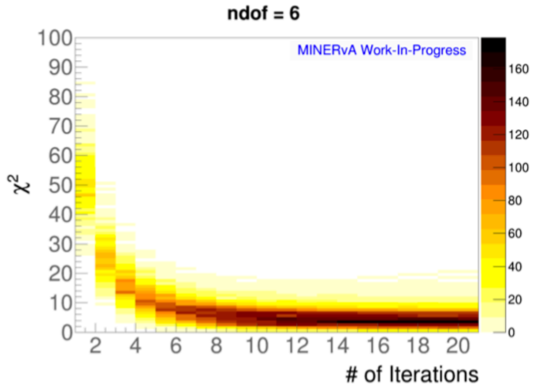
Overview of statistics for neutrino physics
Speaker: Yoshi Uchida (Imperial College London)

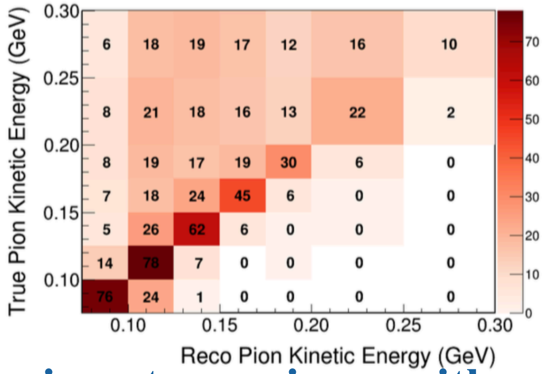
- But recall (via Yoshi from a previous PHYSTATnu discussion) that some of this tension comes from the fact that covariance information is missing from the MiniBooNE result.
  - Was too difficult to evaluate reliably in the many bins of the analysis in the presence of the detector systematics.
- Argues, in part, for the importance of archival quality data that can be tested against new hypotheses.

Unfolding into "true" space

1. When does the unfolded distribution closely approximate the (warped) truth?
2. When is the $\chi^2/\text{dof} \sim 1$, averaged over many Poisson throws?
3. In this case, $\chi^2/\text{dof} \sim 1$ with bins removed





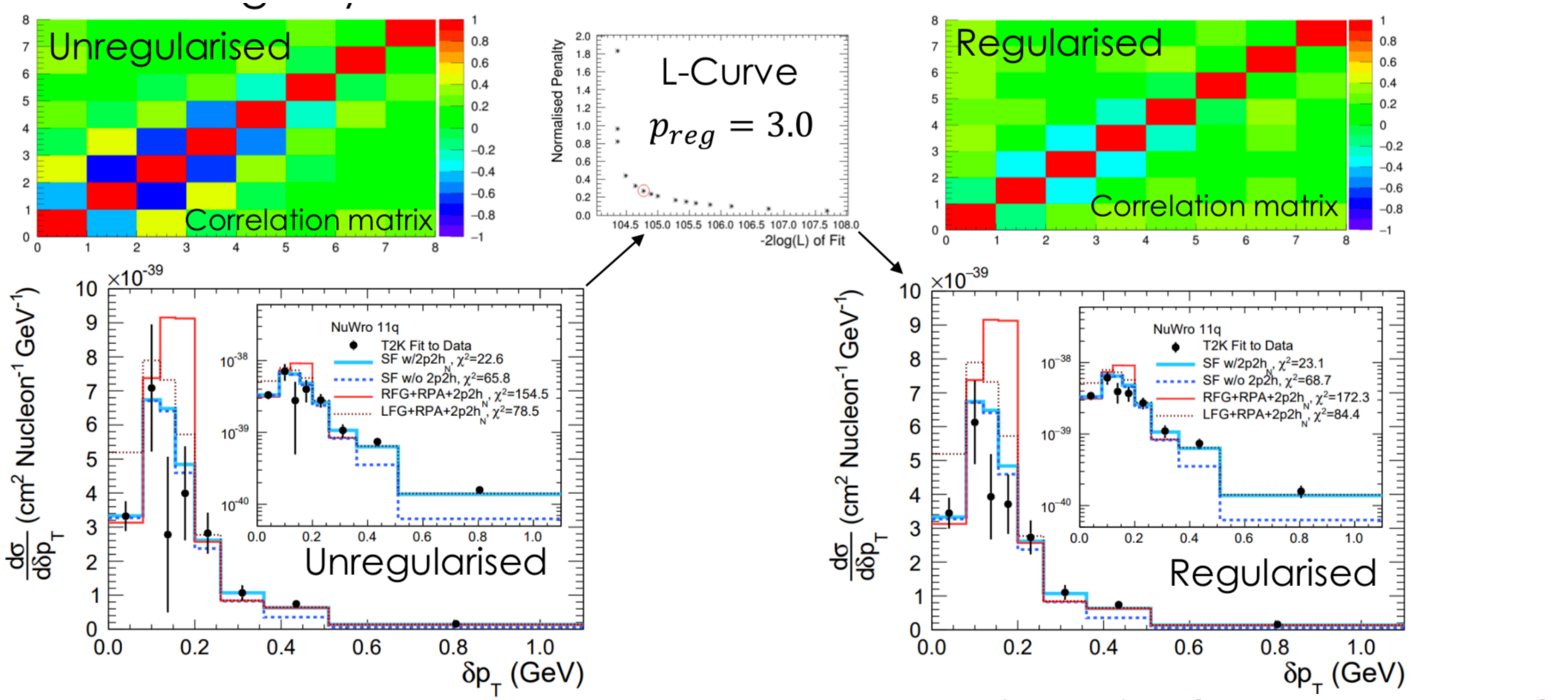


- Unfolding from reconstructed to true variables is an important tool to enable comparison with theory, particularly for spectra, interaction cross-sections, etc.
- Many well known difficulties.
- MINERvA described a particular regularized unfolding (expectation-maximization, truncated interactions, d'Agostini)
- Checks live in the smeared "reco" space, and can't validate "true" space accuracy.

The neutrino experiment experience with unfolding

Speaker: Philip Andrew Rodrigues (University of Oxford

Unfolding into “true” space (cont’d)



Regularisation for T2K cross-section analyses
Speaker: Stephen Dolan (LLR / CEA Saclay)

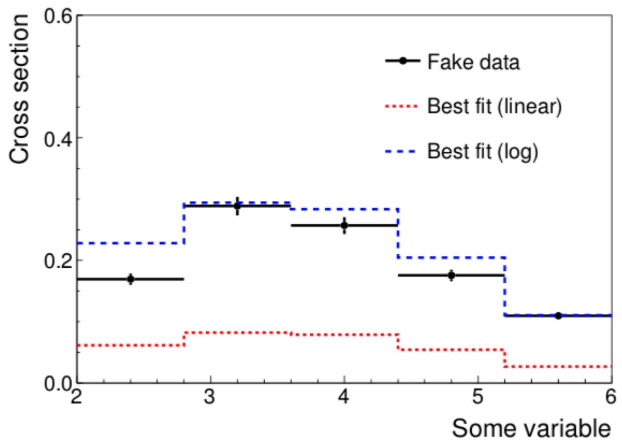
Response Matrix Utils (Lukas Koch)

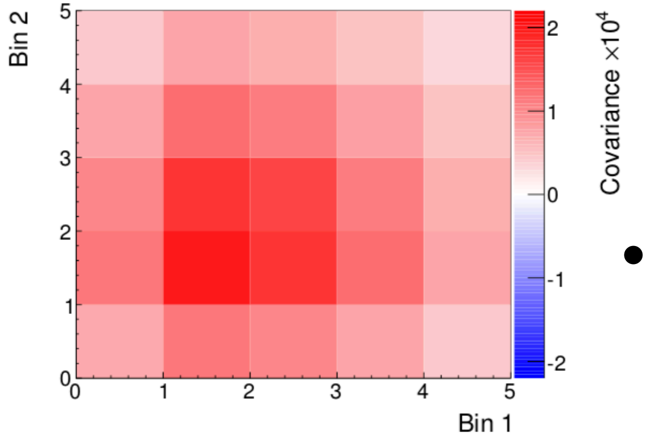
- A tool for forward folding analyses
  - Builds response matrix
  - Tests model dependence
  - Evaluates uncertainties
  - Compare model to data (likelihoods, p-values, MCMC)
- More information: https://remu.readthedocs.io/

- T2K also produces unregularized unfolding, and can compare against models in parallel regularized versions.
  - This may provide a better range of descriptions.
- Also investigating prodcution of reco results and response matrix.
  - Technically difficult, but it may be the best solution if it could be made practical.

Solving "Peelle's Pertinent Puzzle" three times fast

Fits to strongly-correlated data 2





$$\chi^2 = (\mathbf{D}-\mathbf{M})^T \mathbf{C}^{-1}(\mathbf{D}-\mathbf{M})$$

$$C_{ij} = \sum_{\text{universe } k} \left(y_i^{(k)} - y_i^*\right)\left(y_j^{(k)} - y_j^*\right)$$

The neutrino experiment experience with unfolding
Speaker: Philip Andrew Rodrigues (University of Oxford

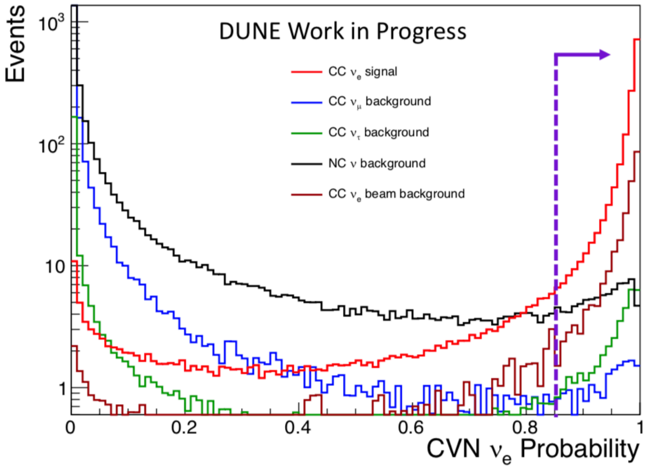
- Phil Rodrigues added the bonus topic to his talk on unfolding… how to fit to strongly correlated data.
- MINERvA has had some success with log-normal uncertainties (many uncertainties, like flux, are largely multiplicative).
- Many high statistics neutrino datasets have large correlations, so problem merits some attention.

"Box-Cox transformation for resolving the Peelle's Pertinent Puzzle in curve fitting", Oh and Seo 2004

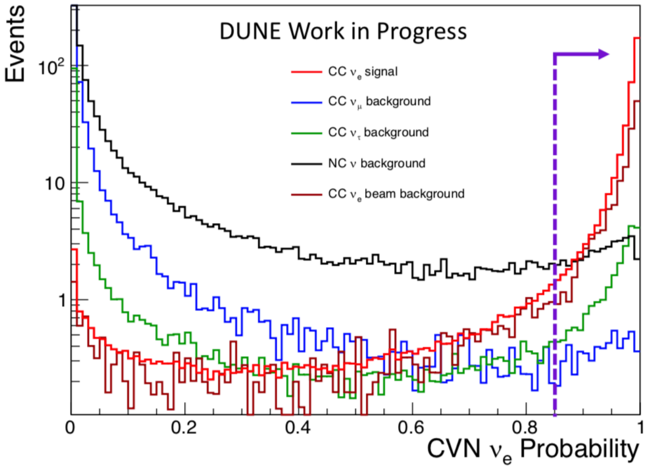
Machine Learning, and fact-checking our robot overlords

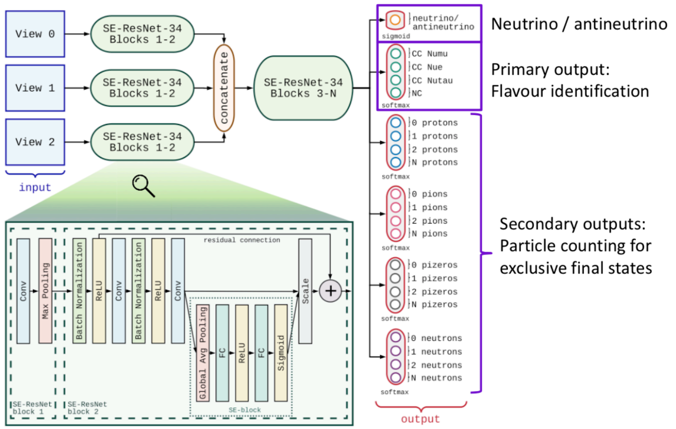
- Electron neutrino probability spectra from the DUNE CVN.
  - Curves combine neutrinos and antineutrinos.

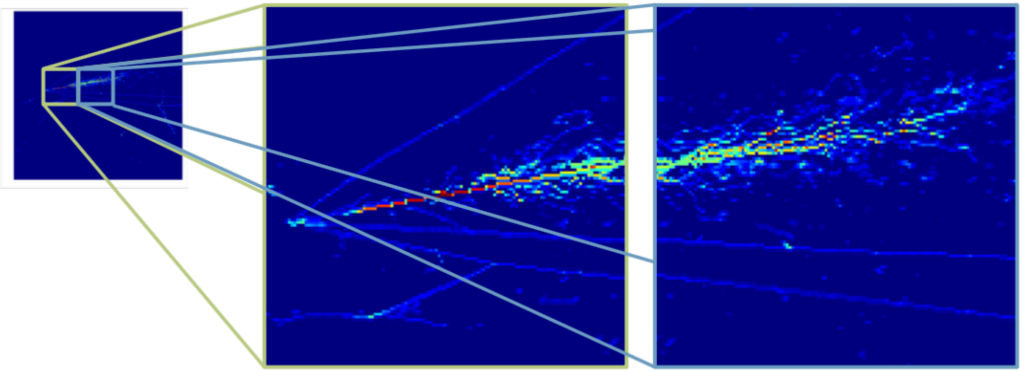
Neutrino beam



Antineutrino beam



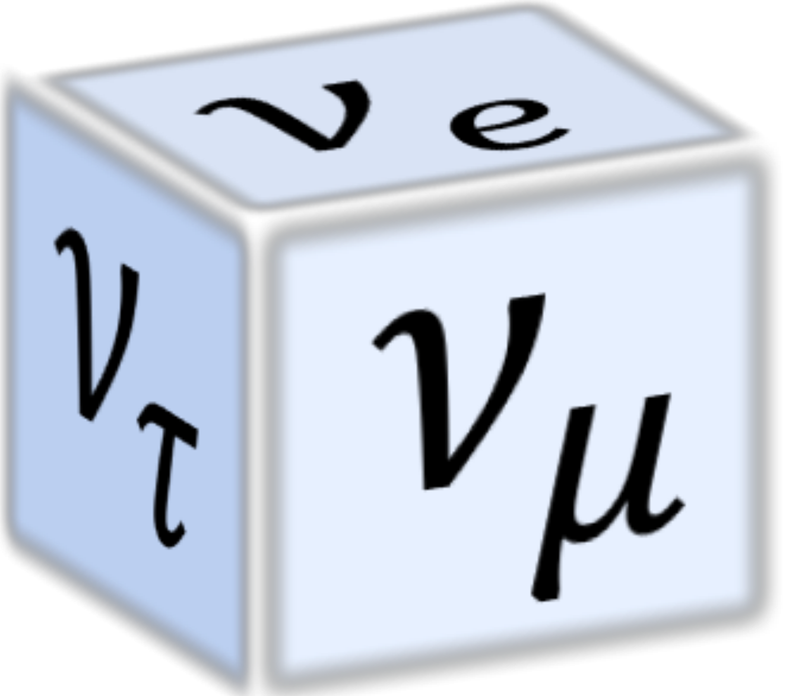




The neutrino experiment experience [machine learning]
Speaker: Saul Alonso Monsalve (University Carlos III)

- Have demonstrated gains in efficiency and purity .
- Less easy to demonstrate independence of interaction and detector model. NOvA, MINERvA are working on this.

Confessions of neutrino physicists



Confession? What confession?

- Many neutrino experiments are looking at very complicated systems, even when the underlying neutrino physics is simple.
- In many cases, neutrino data has badly outstripped our ability to model it.

- Should we be worried today? Maybe…

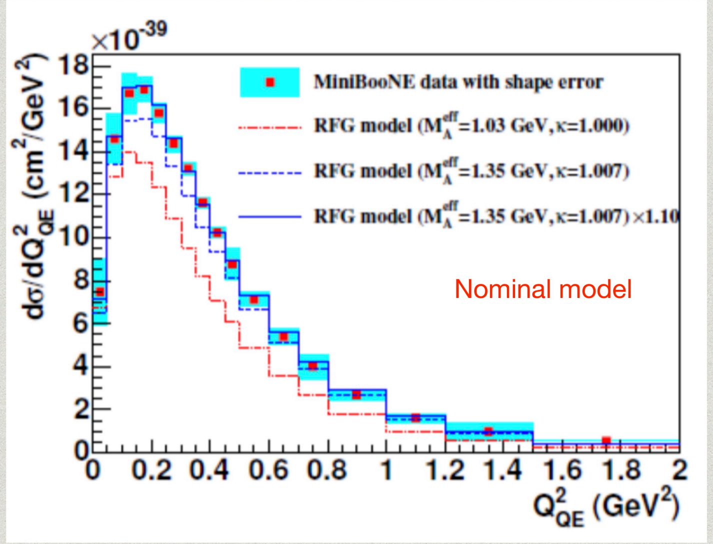
Canonical cautionary tale: MiniBooNE

Experiments confronting data/MC discrepancies

Experiments <u>need</u> a model that describes their data

However, often, **data/MC agreements are handled in a non-satisfactory way**

- Overemphasising own data - breaking consistency with other neutrino data
- Largely ignoring complementary constraints from charged-lepton and hadron scattering



A typical (and conveniently old and non-controversial) example comes from the MiniBooNE experiment:

Tweaking axial form factor parameter

- Axial mass 1.03 -> 1.35 GeV
- Not consistent with bubble chamber results

Tweaking Pauli blocking

- Not consistent with textbook physics

Good description of own data.
But wrong physics!

22

- MiniBooNE observed a discrepancy in its “CCQE” (charged current quasielastic) events vs Q^2.
  - Attributed to axial form factor and Pauli blocking, just an event distortion in Q^2.
  - We understand now this is, at least in part, due to multinucleon production with a different energy-momentum transfer relationship.
- Burying the difference in form factor means misreconstructing E_ν.

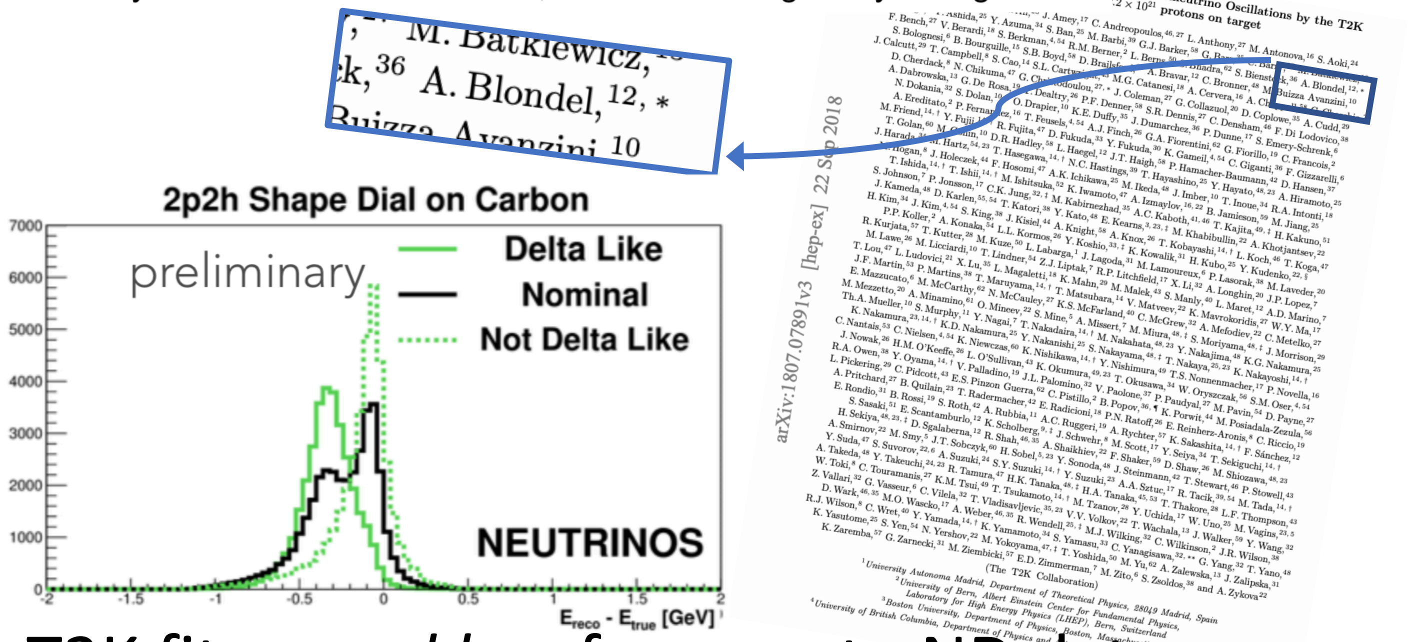
The neutrino experiment experience [combining data]
Speaker: Prof. Constantinos Andreopoulos (Liverpool, STFC/RAL)

Have we learned from this?

- Yes, we have learned.
- Our models for neutrino interactions have improved their description of the data.
- We include more possibilities in trying to describe differences between data and interaction models.

- No, we haven't learned.

It is however not recommended (i.e. should be forbidden really) to fit some data with a convenient but arbitrary or unsure or model-dependent function (i.e. fit looks good) and act as if the error matrix of the fit represents the uncertainty on the fit data. It does not, -- and this can go very wrong!



- T2K fits an *ad hoc* $f_{\text{“delta like”}}$ to ND data.

Have we learned from this?

"Lord, help me to be pure. But not yet."

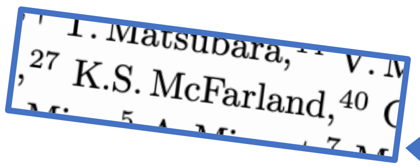
St. Augustine

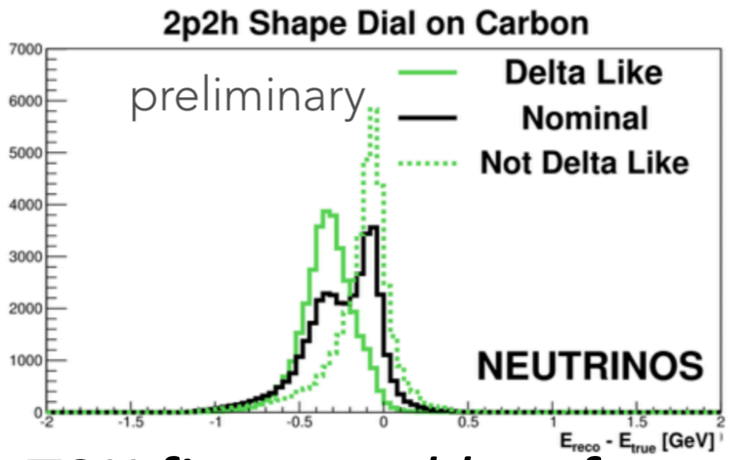
"Yeah, I want to do right, but not right now."

David Rawlings/
Gillian Welch
"Look at Miss Ohio"

- No, we haven't learned.

It is however not recommended (i.e. should be forbidden really) to fit some data with a convenient but arbitrary or unsure or model-dependent function (i.e. fit looks good) and act as if the error matrix of the fit represents the uncertainty on the fit data. It does not, -- and this can go very wrong!





- T2K fits an *ad hoc* $f_{\text{“delta like”}}$ to ND data.

What should we have learned about constraining *ad hoc* models?

- Clearly this is problematic.
  - Data is used to establish the need for, and the form of, the parameterized “fix” to the interaction model.
  - Then the same data is fit to establish the parameters of the “fix”.
  - Easy to show that this is a potential problem. It risks “overfitting” by using the same data to chose the parametrization and the parameters.
- Is there ever a role for *ad hoc* parameters, and, if so, what is it?
  - The *ad hoc* parameters are sometimes all that we have where there is no theoretical guidance about how to reproduce features in data.
  - If we can avoid the “overfitting” trap, then these *ad hoc* models can tell us something about the effect of deficient models.
  - Right now, “overfitting” is avoided by human intelligence. Need better standards for establishing.

Uncertainty estimation by survey of theory models

- A technique used to evaluate theoretical uncertainties is to survey available models that describe datasets.
  - At hadron colliders in the early Tevatron days, this was a stalwart method for PDF uncertainties.
  - Today, in oscillation experiments, this is common for evaluating nuclear initial state and final state models.
  - Also a common test of machine learning performance.

- Dave Soper compared this method to attempting to measure the width of a valley…

… by studying the variance of the position of sheep grazing in it.

Uncertainty estimation by survey of theory models

- This has an obvious and fatal failure mode.
- Dave Soper compared this method to attempting to measure the width of a valley…

… by studying the variance of the position of sheep grazing in it.

Uncertainty estimation by survey of theory models

- This has an obvious and fatal failure mode.
  - Sheep read each others' papers.
- It's just wrong.
- But we continue to do it because often there are not straightforward alternatives.
  - In the PDF community, this was addressed by fitters explicitly producing uncertainties as an output.

- Dave Soper compared this method to attempting to measure the width of a valley…

… by studying the variance of the position of sheep grazing in it.

Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- So who has read the latest T2K-NOvA joint oscillation result paper?

Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- So who has read the latest T2K-NOvA joint oscillation result paper?
  - Anyone?

Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- So who has read the latest T2K-NOvA joint oscillation result paper?
  - Anyone?
  - Anyone?

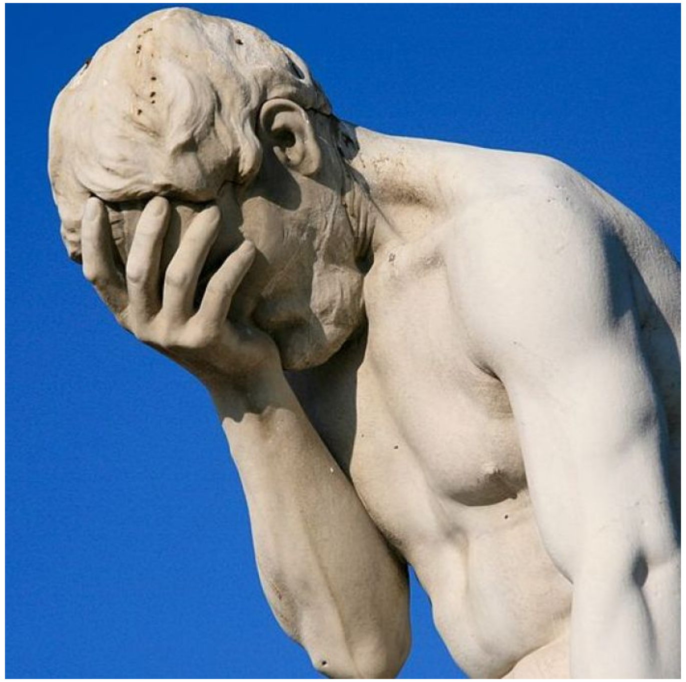
Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- Oh wait, there is no T2K-NOvA joint oscillation paper.

Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- Oh wait, there is no T2K-NOvA joint oscillation paper.

- But a quick inspirehep.net search revealed 11 papers since 2016 combining T2K and NOvA and others with over 500 citations.

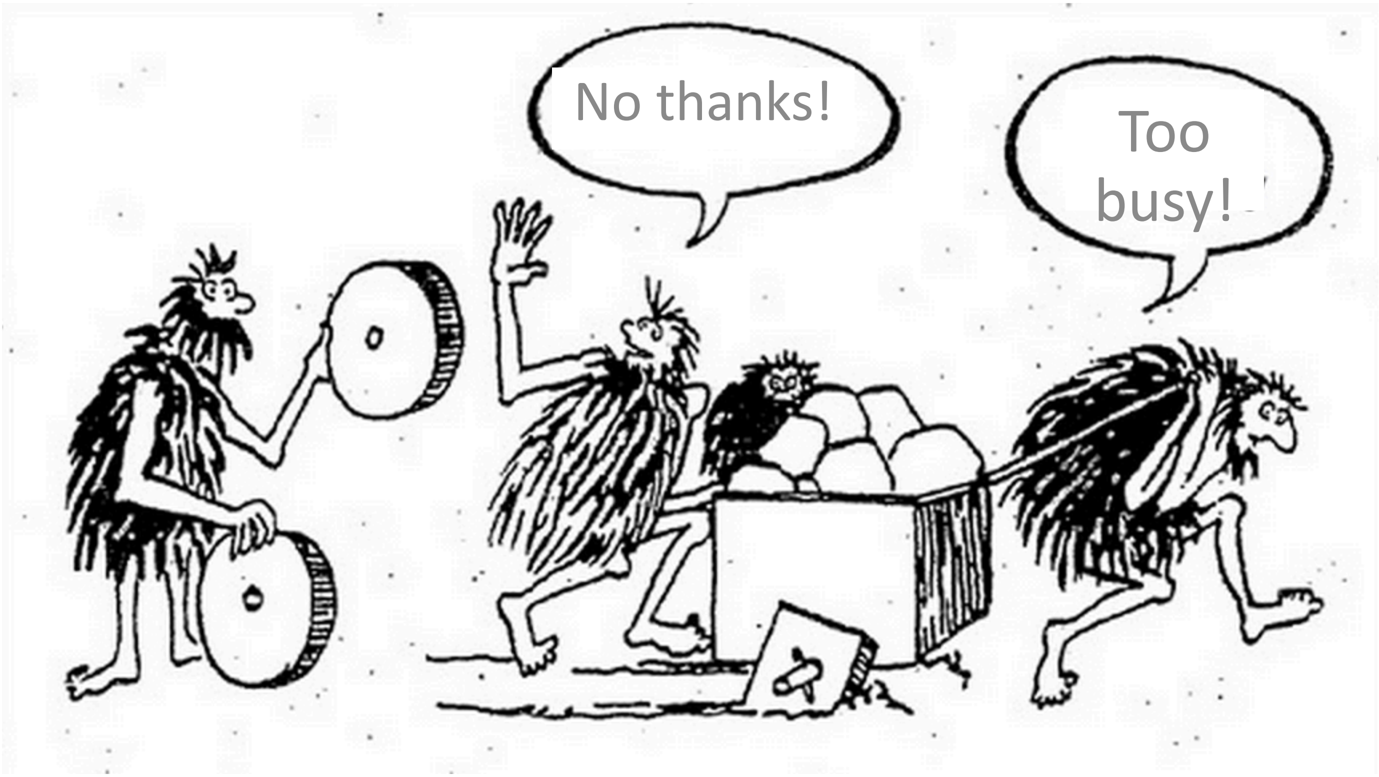


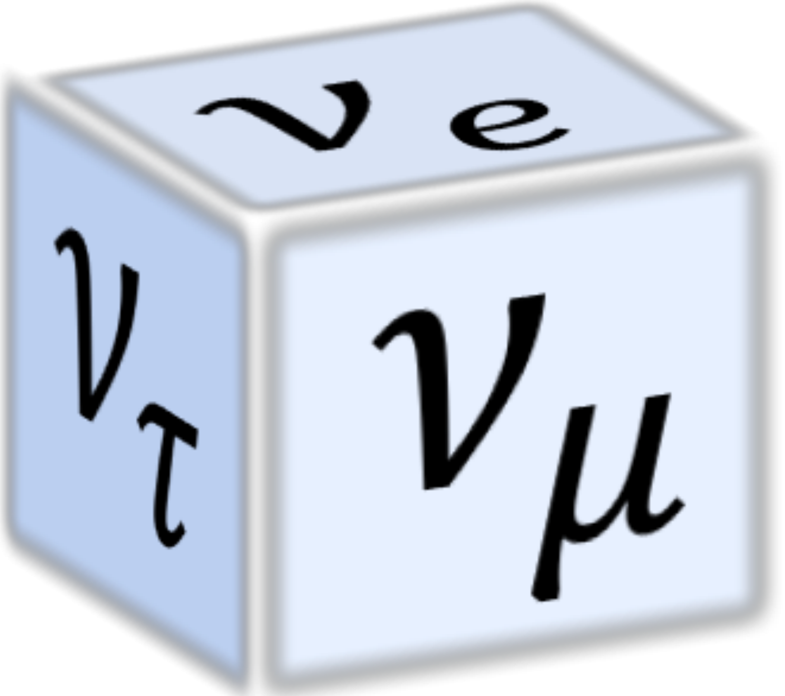
Underprepared for combinations

- The collider community has a long, and even glorious, history of anticipating the need to combine results across experiments. And then organizing to do so.
- Oh wait, there is no T2K-NOvA joint oscillation paper.

- Generally, neutrino experiments…
  - haven't tackled all the problems of developing globally compatible models.
  - haven't invested in the required infrastructure.
  - haven't set up results to use tools developed in the collider community.
- There is obvious motivation to learn from the collider experience.

Maybe someone has already solved our problems?





Homemade solutions are not always the best solutions

- Many "best practices" to learn from the LHC experience, and from other fields. The wealth of new data in neutrino physics can benefit.

Model building, analysing and treatment of systematic uncertainties in modern HEP experiments
Speaker: Wouter Verkerke (Nikhef)

Practical experiences with ATLAS/CMS Higgs combine procedures
Speaker: Nicholas Wardle (Imperial College)

Efficient Neutrino Oscillation Parameter Inference with Gaussian Processes
Speaker: Bannanje Nitish Nayak (Irvine)

The collider experience with unfolding
Speaker: Stefan Schmitt (DESY)

Introduction to Unfolding
Speaker: Mikael Kuusela (CMU)

A new unified perspective on the problem of limited Monte Carlo for likelihood calculations
Speaker: Thorsten Glüsenkamp (Universität Erlangen-Nürnberg)

Statistical Models with Uncertain Error Parameters
Speaker: Glen Cowan (RHUL)

The collider experience [machine learning]
Speaker: Maurizio Pierini (CERN)

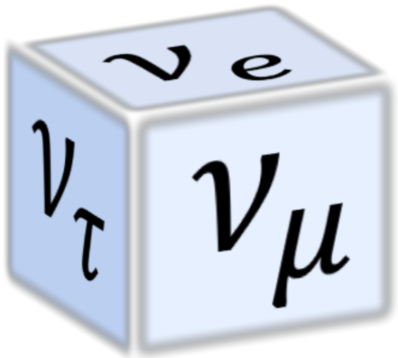
Theoretical aspects of Machine Learning
Speaker: Chad Shafer (CMU)

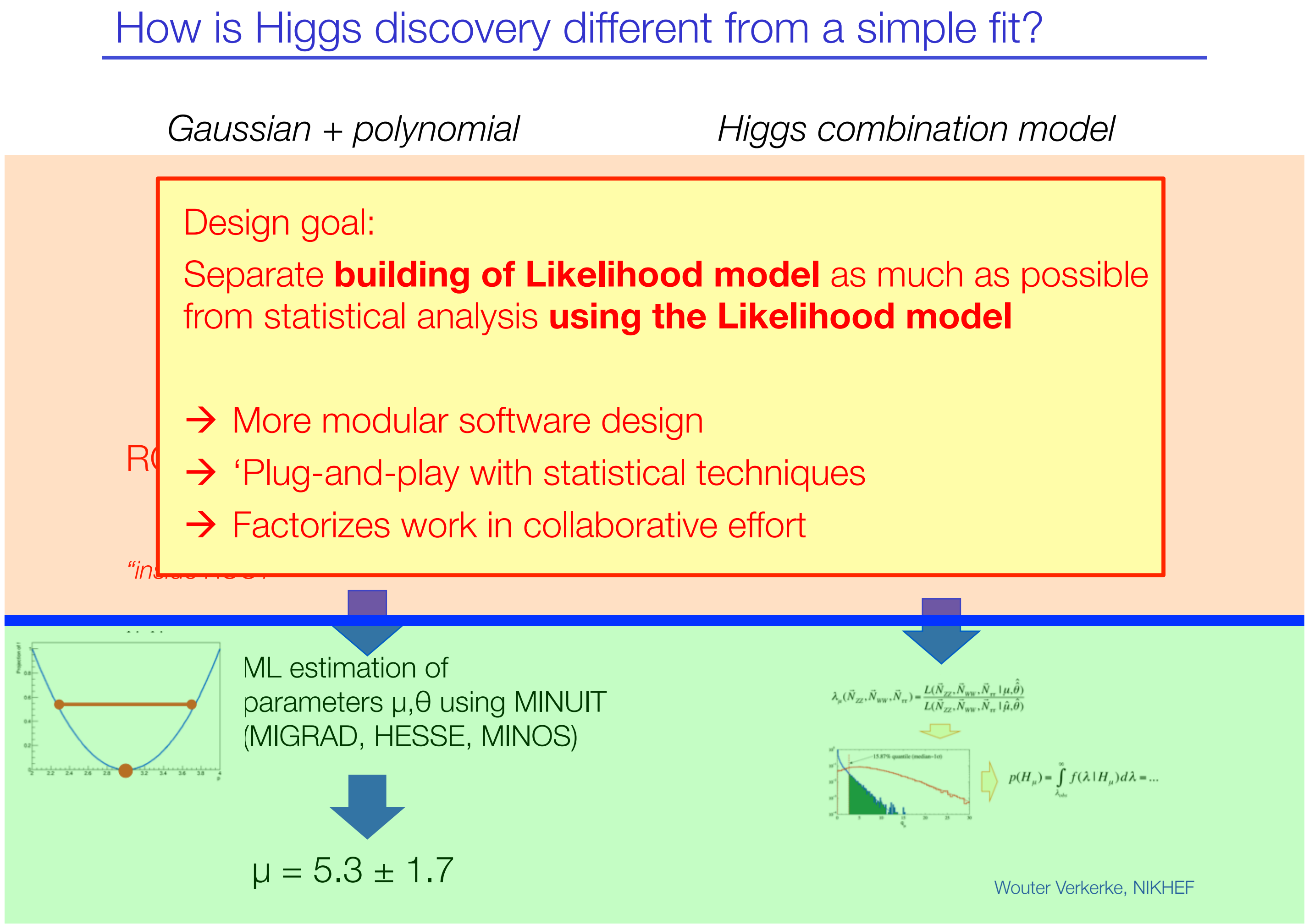
Introduction to machine learning
Speaker: Michael Aaron Kagan (SLAC)

20 Years of Feldman-Cousins
Speakers: Robert Cousins (UCLA)

A GPU-based framework for multi-variate analysis in particle physic
Speaker: Xuefeng Ding (GSSI)

One likelihood to rule them all?





- LHC Higgs combination benefitted from robust approaches in the analysis that aided the ultimate goal of interpreting the data.
- Obvious application for us: many neutrino experiments contribute to measurements of few parameters.

Model building, analysing and treatment of systematic uncertainties in modern HEP experiments
Speaker: Wouter Verkerke (Nikhef)

A problem we face daily and almost never address… Uncertainties of errors

Outline

Using measurements with "known" systematic errors:

Least Squares (BLUE)

Allowing for uncertainties in the systematic errors

Estimates of sys errors ~ Gamma

Single-measurement model

Asymptotics, Bartlett correction

Curve fitting, averages

Confidence intervals
Goodness-of-fit
Sensitivity to outliers

Discussion and conclusions

Details in: G. Cowan, *Statistical Models with Uncertain Error Parameters*, arXiv:1809.05778 [physics.data-an]

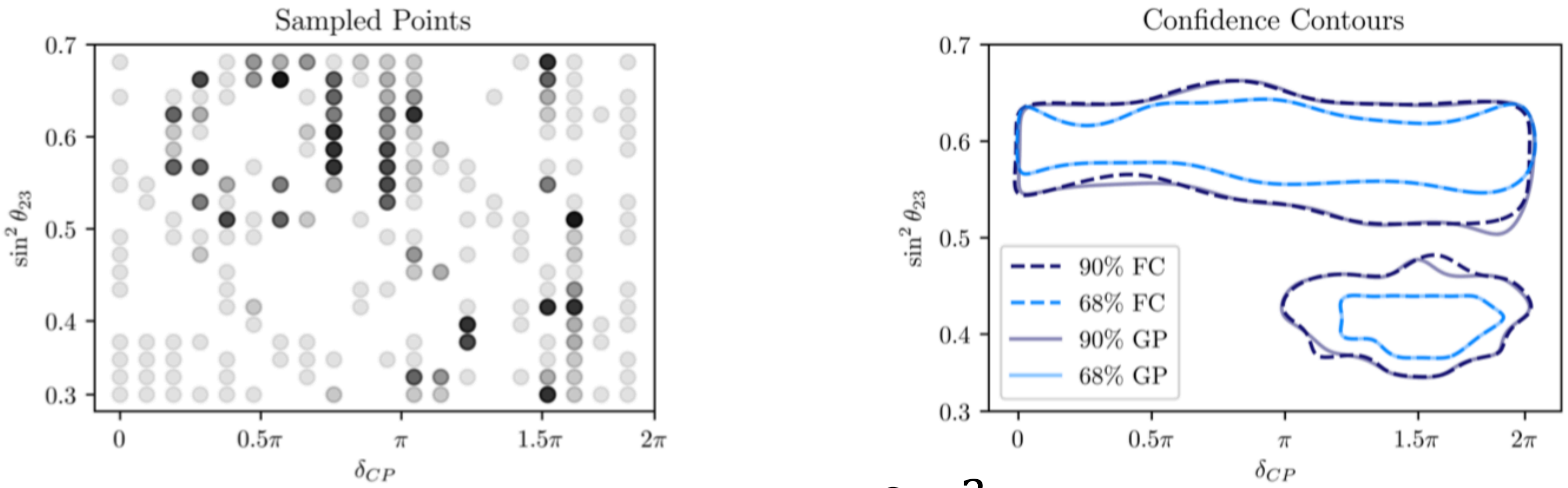
G. Cowan PHYSTAT-ν 24 Jan 2019 / Statistical Models with Uncertain Error Parameters

Statistical Models with Uncertain Error Parameters
Speaker: Glen Cowan (RHUL)

- Generally, we take the "conservative" approach of inflating uncertainties.
- But we learned that there exist rigorous methods that give proper coverage, combination of constraints, etc.
  - And we learned that the "conservative" method does not.

Overcoming computational morass of Feldman-Cousins

- "Real" data similar to latest best-fit estimate from NOvA. ($sin^2\theta_{23} = 0.56$, $\Delta m^2_{32} = 2.44 \times 10^{-3} eV^2$, $\delta_{CP} = 1.5\pi$)
- $sin^2\theta_{23} - \delta_{CP}$ 68% and 90% CI for NH after 5 iterations



Efficient Neutrino Oscillation Parameter Inference with Gaussian Processes
Speaker: Bannanje Nitish Nayak (Irvine)

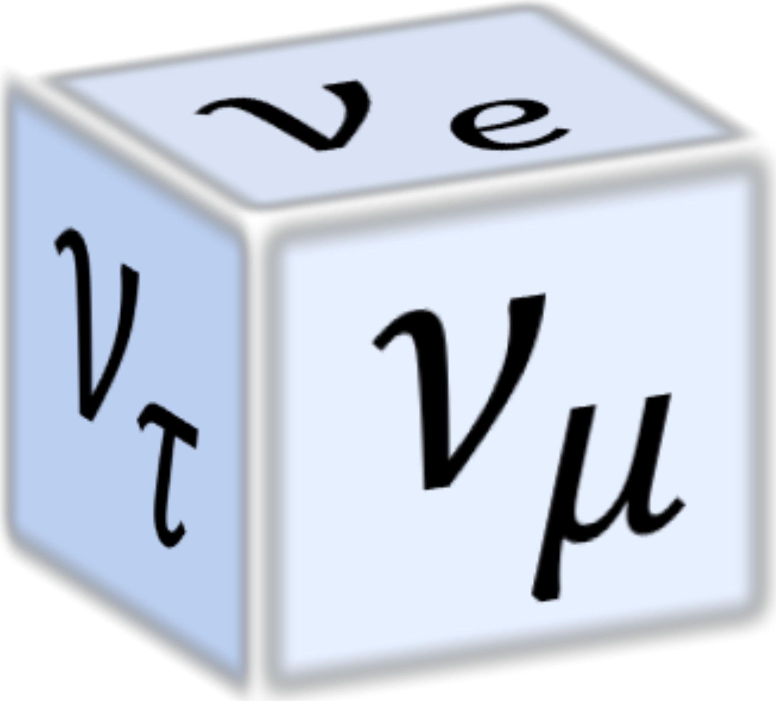
- Computing problem of scanning $\delta\chi^2$ across a vast space practically limits ability of T2K (probably also NOvA?) to quickly update results.
- Gaussian Process method (developed for LHC dijets backgrounds by M. Frate, K. Cranmer, et al.) provides a significant reduction in CPU demand.

Most inadequate part of the summary?

- I've picked out only a few examples, and feel like I have barely scratched at the surface of this vast tool kit.
- In the introduction, an interface with "statistics committees" was suggested.
- Some neutrino experiments are too small for this, but the idea is sound.

"Hey, could we have some volunteers to stay and help clean up?"

Parting words of “wisdom”



From confessions to reparations

- It doesn't do a whole lot of good to confess the sins of our community without concrete plans to fix them.
- I am struck, but not surprised, and how many tools have been left for us by the LHC community, and I think in many domains the neutrino experiments have already benefitted from these.
  - But we are leaving a lot of applicable work on the table.
- Some of these problems require creative solutions.
  - The flux and interaction degeneracy (*as noted by Costas in his talk*).
  - Specific modeling deficiencies that leave us unable to benefit from control samples.
  - Our community failure to take responsibility for data combinations.

Community standards

- Some of our failures are simple inability to communicate and agree upon best practices.
- Pressures faced by experiments at different parts of their life cycle are different.
  - Advocacy and review phase has different political demands…
    … then the publication and legacy stage.
- How can we constructively address this problem?

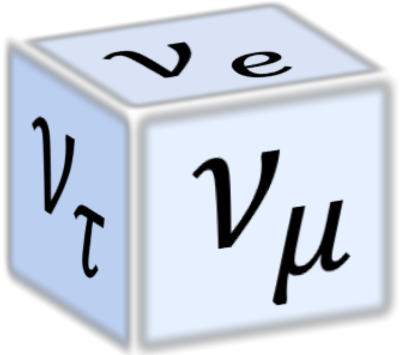
Education

- Much of what we've learned about this week is not widely taught in schools, lectures, etc.
- I know this community has worked extensively on outreach,.
- It's probably time for the neutrino experiments and laboratories to reach out reciprocally.
  - We could use the energy we save from not riding on our favorite square wheels.

Acknowledgments

- Thank you to the organizers for putting together a productive set of discussions.
- Thanks to Alain Blondel, Yoshi Uchida, and Tom Junk for their summaries and for making my job both easier and less mission critical.
- Thank you to the speakers for providing such helpful and easy to navigate material.
  - And a special thank you to speakers who posted slides in advance!

Acknowledgments



- Thank you to the organizers for putting together a productive set of discussions.
- Thanks to Alain Blondel, Yoshi Uchida, and Tom Junk for their summaries and for making my job both easier and less mission critical.
- Thank you to the speakers for providing such helpful and easy to navigate material.
  - And a special thank you to speakers who posted slides in advance!
- And thanks to all of you for your kind attention.