Systematic Constraint Fitting with VALOR

Steve Dennis for the VALOR group

PhyStat-Nu 2016
Fermilab
September 19 2016

An Introduction to VALOR

VALOR is an experimental neutrino fitting package:

- It was originally designed for T2K oscillation analysis.
- We've branched out since, to fit arbitrary numbers of samples for arbitrary groups of detectors.
- Today, I'll be talking about our experience fitting near detector data to provide systematic constraints for a far detector oscillation fit.
    - The ND-only fitting work has primarily been applied to the proposed DUNE near detector designs.
    - There have also been VALOR simultaneous near-and-far fits for a proposed T2K Water Cerenkov near detector (TITUS).
- We can also fit oscillation parameters using far detector detector data simultaneously with near detector data.
    - I'll be sticking to systematic-only fits today. See Lorena's talk for oscillation fits.

Personnel

We currently have 12 members:

- University of Cambridge:
  - Lorena Escudero - Postdoc (VALOR thesis - T2K joint ν_e-ν_μ).
- University of Geneva:
  - Davide Sgalaberna - Postdoc (VALOR thesis - T2K joint ν-$\bar{\nu}$).
- University of Lancaster:
  - Tom Dealtry - Postdoc (VALOR thesis - T2K ν_μ disappearance).
- University of Liverpool:
  - **Costas Andreopoulos** - **Group founder and leader.**
  - Christopher Barry - PhD student.
  - Francis Bench - PhD student.
  - Steve Dennis - Postdoc (VALOR thesis - T2K $\overline{\nu_\mu}$ disappearance).
  - Rhiannon Jones - PhD student.
  - Marco Roda - Postdoc.
- University of Oxford:
  - Raj Shah - PhD student.
- University of Warwick:
  - Andy Chappell - PhD student.
  - Nick Grant - Postdoc.

Long Baseline Oscillation Experiments Extremely Briefly

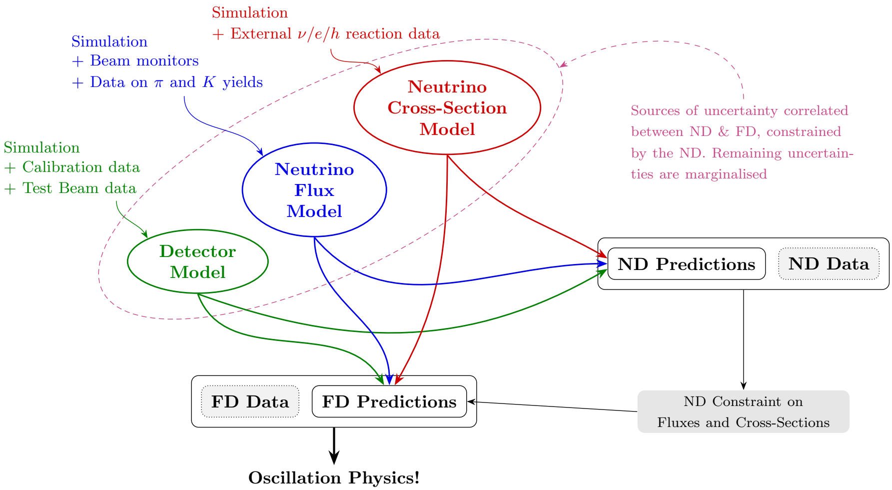
- We produce a neutrino beam of definite neutrino flavour.
- After a long distance (hundreds of km), we detect these neutrinos at our far detector, where the flavour composition has changed due to oscillations.
- The event rate we observe at the far detector is a product of oscillation probabilities P_{osc}, fluxes Φ, interaction cross-sections σ and detector acceptance ϵ:

  $N_{far} \propto P_{osc} * \epsilon * \Phi * \sigma$
- Critically, uncertainties on the fluxes and the cross-sections limit our resolution for precision oscillation measurements.

Objectives of a Neutrino LBL Near Detector Fit

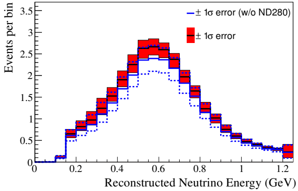
- The point of a near detector is to constrain systematic uncertainties for a far detector (oscillation) fit.
- The far detector event rate observed depends on the neutrino flux, interaction rates and detector uncertainties.
- In order to maximise oscillation sensitivity, we use a near detector's high statistics and lack of oscillations to obtain accurate measurements of the parameter values, their uncertainties and their correlations.
- We produce as output a vector of best-fit values and a covariance matrix with the relevant near detector-only uncertainties marginalised.
    - For a data fit, it is also important to provide a measure of goodness-of-fit, indicating the extent to which the data fits the model and parameterisation used.
- The current major use of the VALOR ND fitting machinery is comparing the effectiveness of different proposed near detector designs for the DUNE experiment.

VALOR Analysis Strategy for T2K-style Indirect Extrapolation



Next Generation Experiments

- In the next generation of experiments, neutrino physics will be producing long-baseline statistics at a level we've never seen before.
- We need new and better near detectors to achieve the level of systematics constraint required to make the most of these new event rates.
- Future near detectors must constrain our systematics to $\sim 2-3\%$.



(T2K run 1-7 constraint on NuE appearance using the ND280 fit: 12.6% without ND, 6.8% with ND)

Spectrum Prediction

The VALOR framework can fit together an arbitrary number of samples to determine the parameters of a physics hypothesis in the presence of systematic uncertainties.

Each sample corresponds to a given:

- d - detector or subdetector
- b - beam configuration (FHC, RHC)
- s - selection, i.e. final state topology
- r,t - **multi dimensional** kinematical reconstructed and true phase space

systematics

physics

true interaction mode

true kinematics

$$\underbrace{n^{pred}_{d;b;s}(r;\vec{\theta};\vec{f})}_{\text{Predicted number of events}} = \sum_{m} \sum_{t} \underbrace{P_{d;b;m}(t;\vec{\theta})}_{\text{Oscillation probabilities}} \underbrace{R_{d;b;s;m}(r,t;\vec{f})}_{\text{Systematic parameters variations}} \cdot \underbrace{T_{d;b;s;m}(r,t)}_{\text{Nominal unoscillated number of events}}$$

reconstructed kinematics

(slide from L. Escudero)

The Fit Statistic

- We fit N samples simultaneously. For the DUNE ND work, we're using 46 samples:
  - 23 selections in each of two beam modes, in a single detector.
- We evaluate a negative log-likelihood ratio (NLLR) in a detector d, beam b and selection s, as a function of parameter values $\vec{\theta}$ and $\vec{f}$, summed over each observable kinematic bin r:

$$-2\ln\ \lambda_{d;b;s}(\vec{\theta};\vec{f}) = 2\sum_r \left\{ \left(n^{pred}_{d;b;s}(r;\vec{\theta};\vec{f}) - n^{obs}_{d;b;s}(r) \right) + n^{obs}_{d;b;s}(r)\cdot \ln \frac{n^{obs}_{d;b;s}(r)}{n^{pred}_{d;b;s}(r;\vec{\theta};\vec{f})} \right\}$$

- We sum over an arbitrary number of samples, and can include a penalty term to reflect our prior knowledge of the parameters:

$$-2\ln\lambda(\vec{\theta};\vec{f}) = \sum_d \sum_b \sum_s -2\ln\lambda_{d;b;s}(\vec{\theta};\vec{f}) - 2\ln\lambda_{prior}$$

- The negative log-likelihood ratio formulated in this way tends to a χ^2 distribution
  - This is useful for constructing a goodness-of-fit test.

Example Selection Lists

Selection	Code	2016a	2016b	Comment
Undefined	-1	✓	✓	
$\nu + e^-$	25		✓	Neutrino + electron elastic scattering
Inverse Muon Decay	26		✓	$\bar{\nu}_e + e^- \rightarrow \mu^- + \bar{\nu}_\mu$ or $\nu_\mu + e^- \rightarrow \mu^- + \nu_e$ (IMD)
ν_μ CC 0π 1-track	1	✓	✓	μ^- only, excluding IDM selection above
ν_μ CC 0π 2-tracks	2	✓	✓	μ^- + nucleon
ν_μ CC 0π N-tracks	11		✓	μ^- + (>1) nucleons
ν_μ CC 3-track Δ-enhanced	12		✓	Δ-enhanced: $\mu^- + \pi^+ +$ p, with $W_{reco} \approx 1.2$ GeV
ν_μ CC $1\pi^\pm$	3	(✓)	✓	$\mu^- + 1\pi^\pm +$ X, excluding Δ-enhanced selection above
ν_μ CC $1\pi^0$	5	✓	✓	$\mu^- + 1\pi^0 +$ X
ν_μ CC $1\pi^\pm + 1\pi^0$	6	✓	✓	$\mu^- + 1\pi^\pm + 1\pi^0 +$ X
ν_μ CC other	7	✓	✓	μ^- + other, excluding selections above
Wrong-sign ν_μ CC inclusive	9	✓		Wrong-sign ν_μ CC inclusive - split in sub-samples in 2016b
Wrong-sign ν_μ CC 0π	13		✓	$\mu^+ +$ X
Wrong-sign ν_μ CC $1\pi^\pm$	14		✓	$\mu^+ + 1\pi^\pm +$ X
Wrong-sign ν_μ CC $1\pi^0$	15		✓	$\mu^+ + 1\pi^0 +$ X
Wrong-sign ν_μ CC other	16		✓	μ^+ + other, excluding selections above
ν_e CC inclusive	8	✓		ν_e CC inclusive - split in sub-samples in 2016b
ν_e CC 0π	17		✓	$e^- +$ X, excluding "$\nu + e^-$" selection above
ν_e CC $1\pi^\pm$	18		✓	$e^- + 1\pi^\pm +$ X
ν_e CC $1\pi^0$	19		✓	$e^- + 1\pi^0 +$ X
ν_e CC other	20		✓	e^- + other, excluding selections above
Wrong-sign ν_e CC inclusive	27		✓	Wrong-sign ν_e CC inclusive ($e^+ +$ X)
NC inclusive	10	✓		NC inclusive - split in sub-samples in 2016b
NC 0π	21		✓	nucleon(s)
NC $1\pi^\pm$	22		✓	$1\pi^\pm +$ X
NC $1\pi^0$	23		✓	$1\pi^0 +$ X
NC other	24		✓	other, excluding selections above
Total:		9	23	

Low statistics but valuable flux information. Selected with specific kinematics

Split in different primary vertex topologies. Same for nue and NC

Enough statistics? Otherwise will not be fitted separately

(slide from L. Escudero)

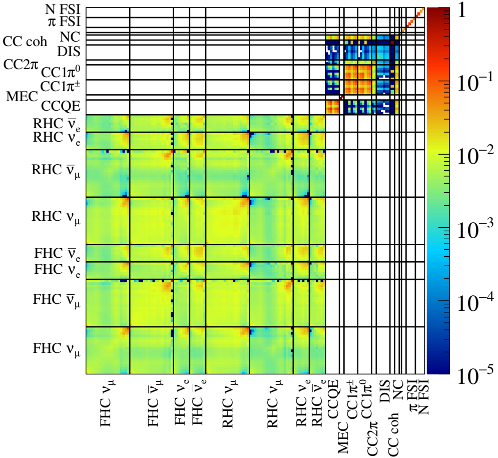
The Penalty Term

- In order to take into account our prior knowledge of the parameter values, we optionally apply a penalty term to the fit statistic as the parameters move further from their prior central values ($\vec{\theta_0}$ and $\vec{f_0}$).
- We use $\vec{\theta}$ to represent physics parameters of interest and $\vec{f}$ to represent systematic uncertainties.
- Each has a covariance matrix **C** whose inverse is used to calculate the penalty term:
$$-2\ln\ \lambda_{prior}(\vec{\theta};\vec{f}) = \left\{(\vec{\theta}-\vec{\theta}_0)^T\mathbf{C}_\theta^{-1}(\vec{\theta}-\vec{\theta}_0) + (\vec{f}-\vec{f}_0)^T\mathbf{C}_f^{-1}(\vec{f}-\vec{f}_0)\right\}$$
- For our near detector only fits, these matrices **C** form both an important input to the fit, and along with a new central values vector $\vec{f_0}$, the primary output (which can then be used as an input to an independent far detector fit).
- The use of a penalty term in this way assumes a Gaussian prior distribution.
    - The alternative is not to use the penalty term, instead using a statistic integrated over the prior distribution.
    - This allows non-Gaussian parameters naturally.

Parameterisation and Prior Uncertainties

- The neutrino flux uncertainties are binned according to true neutrino energy and initial neutrino flavour.
  - For the current generation of DUNE fits, these use 104 parameters each for the near and far detectors, and the prior uncertainties and correlations are provided by the DUNE beam task force (L. Fields).
- Our current neutrino interaction uncertainties are parameterised according to an effective model.
  - This was developed in-house by the VALOR group, and is an effective model fitting cross-section normalisations in primary vertex interaction type and kinematic bins.
  - As well as the binned normalisations, we apply Final State Interaction systematics using non-linear responses generated using GENIE.
  - All of the interaction systematics have their prior ranges calculated using the GENIE Monte Carlo Generator.
- Detector uncertainties are highly dependent on the detector technology and software, and for now we use a highly generic ad-hoc model as a very rough substitute for properly evaluated detector responses.

An example prior covariance matrix, used for a DUNE ND-only fit:



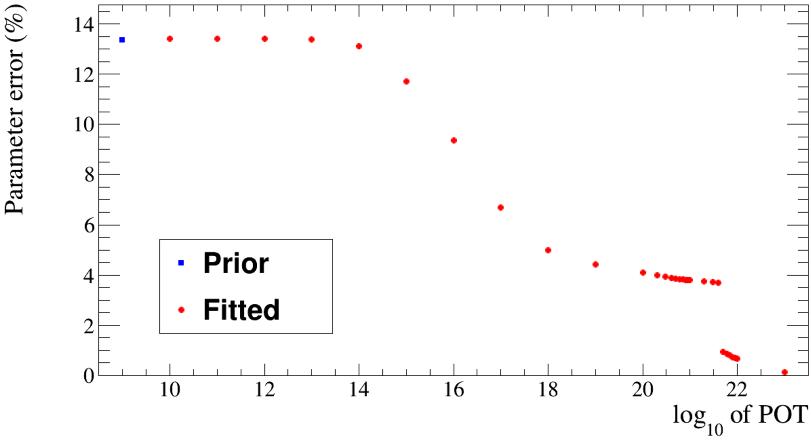
(showing only ND flux and interaction uncertainties)

Extremising the Statistic

- For any given dataset, we use the MIGRAD gradient descent algorithm from MINUIT to find the minimum in the parameter space.
  - Care must be taken to ensure that a global minimum is found.
  - For the fake-data fits we're using for simulated detectors at DUNE, local minima haven't been an issue... yet.
- Initially, we calculated the output uncertainty matrix using the HESSE routine from MINUIT.
  - This routine calculates the Hessian matrix: $\mathbf{H}_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$
  - The inverse of this matrix gives the covariance matrix.
- Unfortunately, we hit a bit of a snag with this algorithm:
  - With sufficiently high sensitivity, HESSE appears to get a little over-optimistic.
  - We see a break-down in the routine, giving incorrect results.
  - For LBL physics, this issue seems to be new to the next-generation experiments.
    - We don't see it occur at (eg) ND280 sensitivity levels.

Parameter error vs fitted POT (HESSE)

This is my favourite plot to show the error calculation failure state:

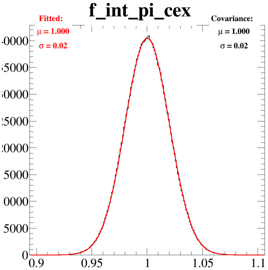


You can see a clear break when the error routine breaks down at $\sim 3 \times 10^{21}$ POT.

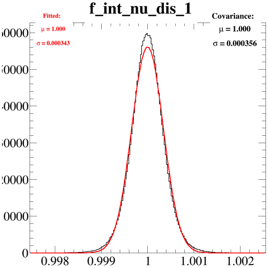
Our Current Approach

- First attempted alternative algorithms to calculating the Hessian matrix.
  - These did not yield reliably better results.
  - The method of inverting the Hessian matrix is also somewhat unsuitable to systematics which have an undefined derivative at the central value - which we should be able to handle.
    - This is something we expect to run into soon - my implementation of binned energy scale systematics looks generally Gaussian, but has a slight 'kink' at 1.
- The alternative we decided on:
  - Use a Markov Chain Monte Carlo to sample the parameter space according to the fit statistic.
  - Use the sampled positions to calculate the final covariance matrix.
- Given we have to do this anyway, maybe we should abandon the use of MINUIT to find the minimum and just do the MCMC?

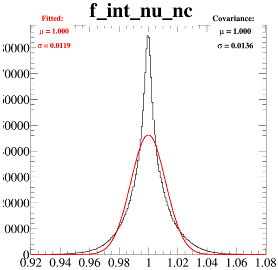
Example Parameter Distributions from the MCMC



Perfect
(22/156)



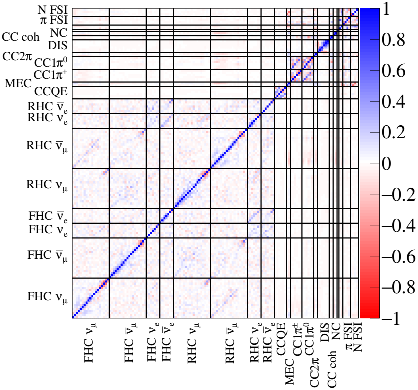
Acceptable
(131/156)



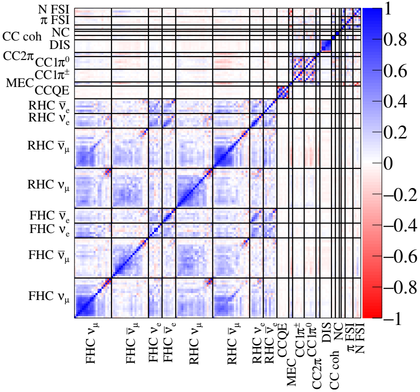
Not good
(3/156 - this is understood)

I'm using the similarity between the Gaussian fitted parameter shapes and the actual MCMC steps recorded to quantify the sufficiency of the number of steps and the mixing rate. Suggestions on other quantities to check and ways to appropriately validate correlations are welcome.

HESSE vs MCMC Correlations



HESSE

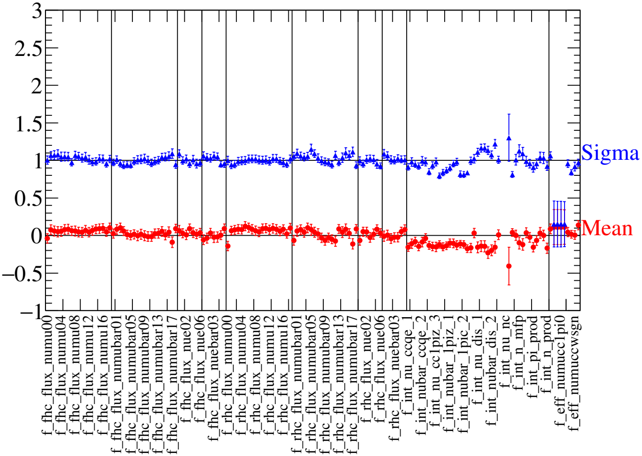


MCMC sampler

Note that HESSE believed it could break almost all the correlations between the parameters. The sampler proves that it can't.

An example of some of our bias studies

Fitted pulls for 500 toy experiments with randomised systematic values: Pull $= \frac{p_{\text{bestfit}} - p_{\text{nominal}}}{\sqrt{\sigma^2_{\text{prior}} - \sigma^2_{\text{postfit}}}}$



Fitted mean should be 0 or we're biased.

Fitted RMS should be 1 or our uncertainty estimates are wrong.

Method for Near to Far Flux Extrapolation

- Our near detector data obviously only directly constrains the flux at the near detector.
  - but of course our oscillation fits require uncertainties on the flux at the far detector.
- Beam flux groups provide full correlation matrices between the near and far detector flux parameters.
- At the moment, our regime for near-to-far flux extrapolation in the ND-only fits is:
  - Fit only the near detector data, but include the far detector flux parameters in the systematic error matrix.
  - The near detector data constrains the near detector flux parameters, and because of the near-far correlations, the fit statistic penalty term adds a weaker constraint on the far detector fluxes.
  - This seems to work relatively well - but the validity of indirect near-then-far fits as opposed to joint near-and-far fits must be checked.

Near-then-far vs simultaneous near-and-far

- Two paradigms:
  - Fitting near detector data to provide a covariance matrix propagated to the far detector, followed by an independent far detector fit.
  - Fitting near detector data simultaneously with far detector data to produce results solely in the parameters of interest, with the ND parameters marginalised away.
- The joint fits are naturally more justifiable.
  - but separating the two detector fits allows clearer understanding of the sources of systematic effects on the final results.
- Of course, the only way to judge if the independent fits are sufficient is to do both methods and compare the results.
- If we see significantly asymmetric or otherwise non-Gaussian parameters in the uncertainty parameterisation, this seriously erodes the viability of independent fits.

Future Near Detector Fitting

- There are increasing discussions of fits using multiple near detectors:
  - For example, DUNE may use hybrid or combination near detector technologies.
  - The current NuPrism proposal involves the same detector collecting data at various off-axis angles - acting in the fit as detectors in separate beam fluxes.
- The Short Baseline Neutrino program at Fermilab has a lot of exciting physics potential.
  - The search for a sterile neutrino signal requires the sample diversity and high-statistics robustness that a near detector fit apparatus provides.
  - It also requires a multiple-detector approach and a robust oscillation framework
  - In other words, they need something a lot like VALOR.

Conclusions

- We've developed an analysis for generic neutrino experiment fitting.
  - It's capable of fitting an arbitrary combination of detectors, beam types and topology selections.
- When using near detectors alone, we have a method for evaluating the constrained systematic parameter values, uncertainties and correlations that an effective far-detector only oscillation analysis requires.
- As well as near-detector only fits, we can produce oscillation fits simultaneously fitting near-and-far detector data.
- This work has found a home in the DUNE near-detector evaluation group, where it will form a key element of the upcoming DUNE TDRs.
- We believe that we can be of use to current and near-future experiments such as the Fermilab SBN program.

Backup