Near Detector Analysis tools Report

Léna Osu
GUNDAM Working Group

Oscillation Analysis tools
Friday, 13th June 2025
IRN

Léna OSU

Main GUNDAM Analyzer

Quick overview

During this presentation, we will talk about:

1. **Systematics in T2K**
   1.1 T2K experiment
   1.2 Flux
   1.3 Cross-section - neutrino interaction modes
   1.4 Detector

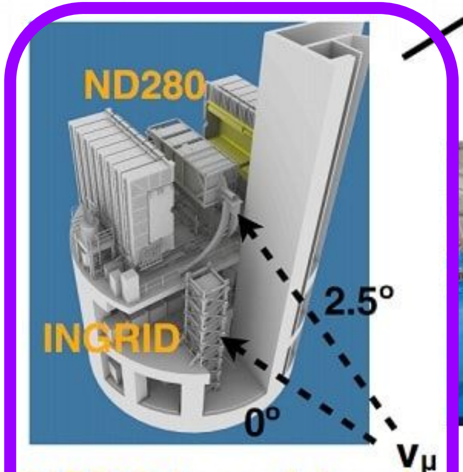
2. **GUNDAM fitter**
   2.1 ND fit goals
   2.2 GUNDAM framework

3. **More recent analysis status and novelties**
   3.1 New 4π selection
   3.2 New RPA/FSI cross-section parameters implementation

T2K Experiment

- Study neutrino oscillation
- Measure U_e appearance probability
- ND280 : sees the un -oscillated flux
- SK : measures the effects of the oscillation

Super-Kamiokande



INGRID (on-axis) and ND280 (off-axis)

Neutrino beam created at J-PARC main ring

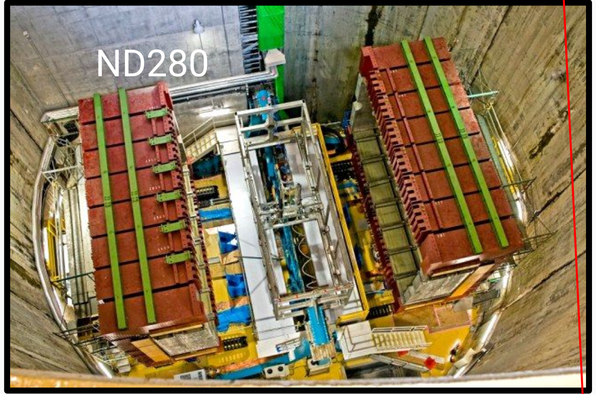
Concept of a neutrino oscillation baseline experiment

Cross-section (neutrino→target probability)

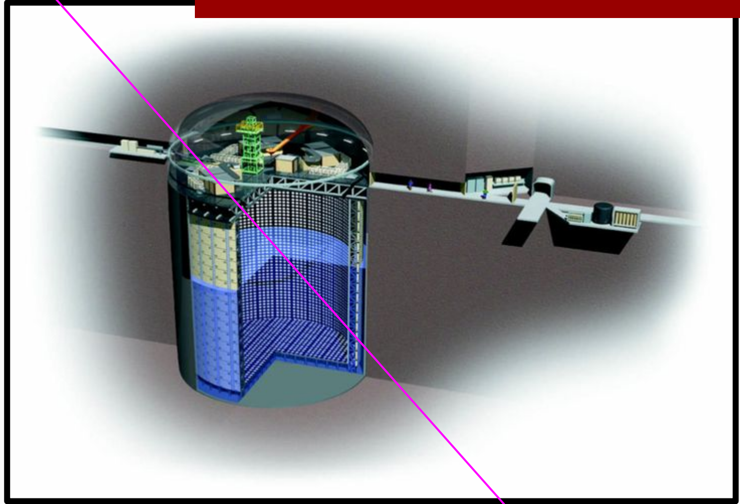
$$\frac{N^{far}_{events}(\vec{x})}{N^{near}_{events}(\vec{x})} = \frac{\sigma(E_\nu,\vec{x})\otimes\Phi(E_\nu)\otimes D^{far}(\vec{x})\otimes P_{osc}(E_\nu)}{\sigma(E_\nu,\vec{x})\otimes\Phi(E_\nu)\otimes D^{near}(\vec{x})}$$

Flux (energy spectrum/rate of produced neutrinos)

Detector (see/reconstruct interactions)



ND **tuned models** propagated to far detector(best possible prediction)



- different detector effects from ND to SK
- systematics uncertainties

$P_{osc}(E_\nu)$ is more precise when E_ν is well-reconstructed

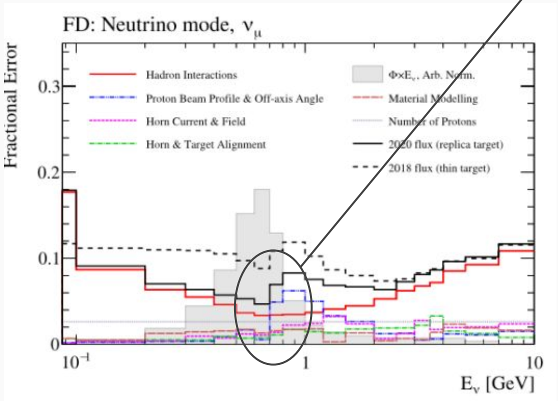
Focus on Near detector !

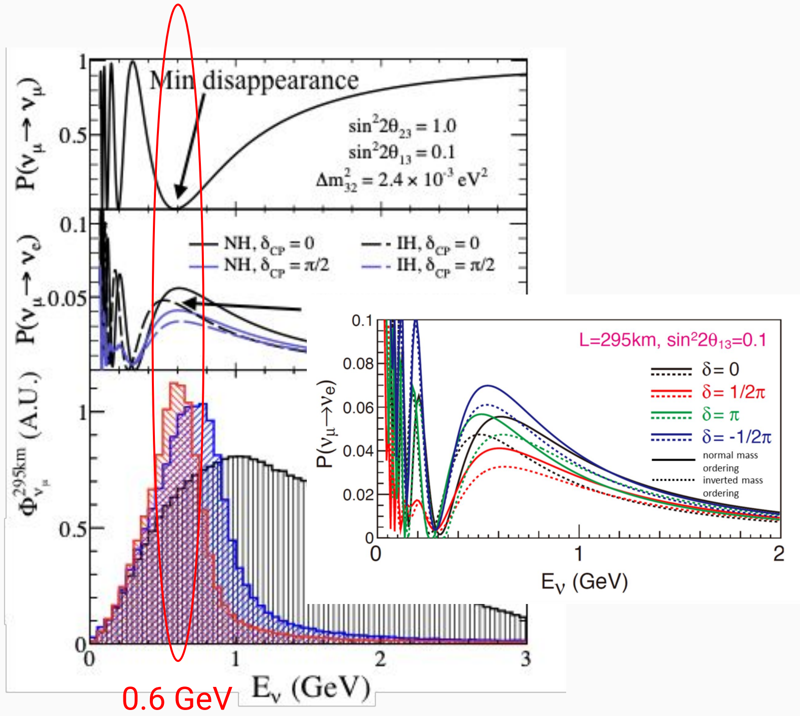
Beam

Flux

$$N^{near}_{events}(\vec{x}) = \sigma(E_\nu,\vec{x}) \otimes \boxed{\Phi(E_\nu)} \otimes D^{near}(\vec{x})$$

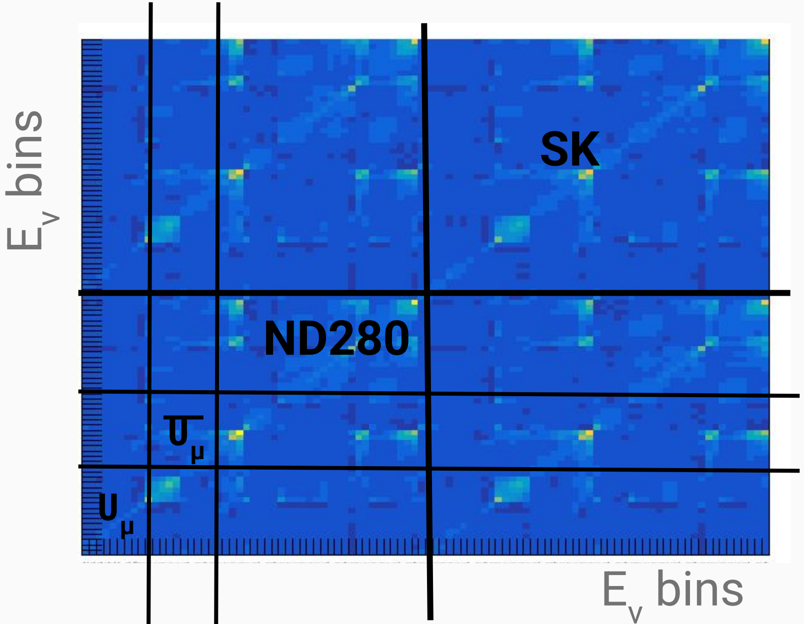
- **protons interact with a graphite target** to produce hadrons (mainly pions and kaons)
- Detectors 2.5° off the direction of the beam **centered around 0.6 GeV.**
- Off-axis method reduce mean energy to maximize oscillation detection probability for 295 km baseline
- Flux model tuned against external data (NA61) - remain **~5% uncertainties**





Flux systematics parameterisation

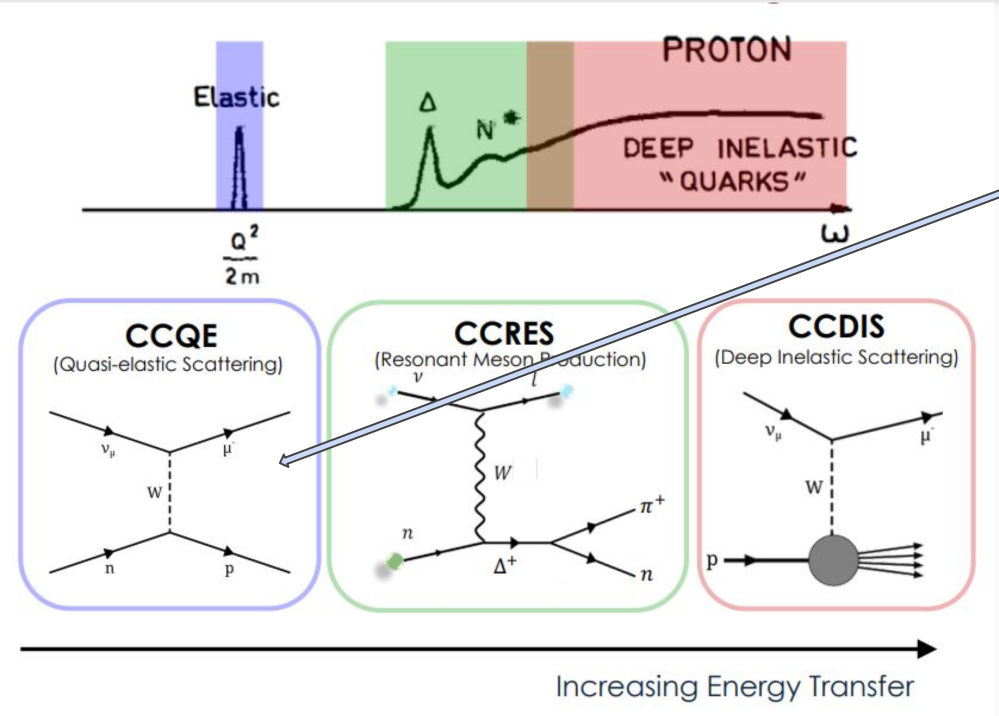
- Uncertainty on the flux reflects errors on hadrons productions or beam alignment variations for instance
- Flux is binned in energy bins and neutrino types (numu, antinumu, nue, antinue)
- In T2K, flux systematics are described as normalization parameters affecting the bin content (in true neutrino energy) according the prior covariance matrix provided by beam experts



Flux covariance matrix

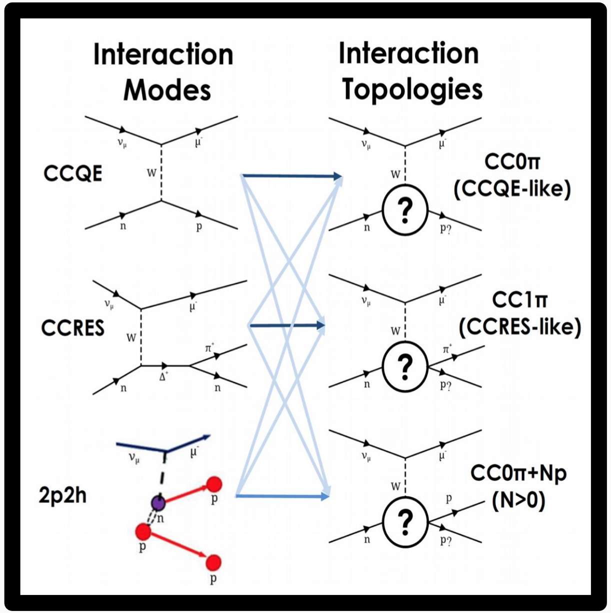
impact ND and SK predictions

neutrino interaction mode



CCQE interactions are the dominant ones at T2K energies.

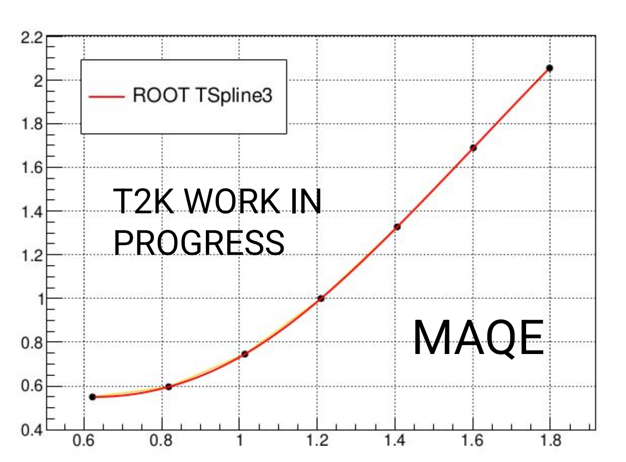
Ideally we want to select pure CCQE events, but nuclear effects play an important role and are source of systematics effects

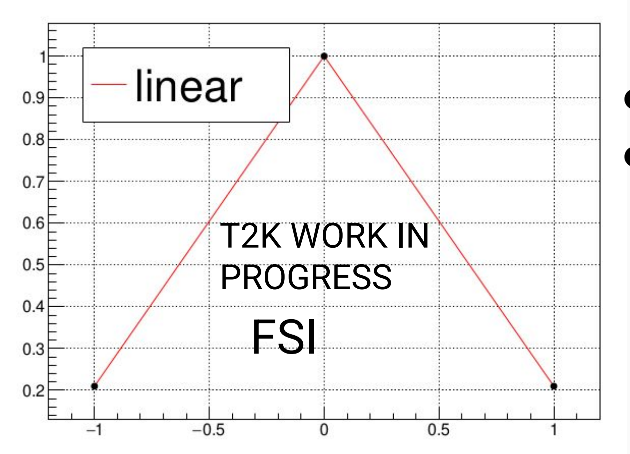


- At T2K energies → multiple interaction contributing
- Each interaction has uncertainties + affected by nuclear effects
- series of parameters to describe and constrain these uncertainties for each interaction channel

Cross-section systematics parameterisation

- **1D event by event response function ≈ spline** : polynomial and continuous function used for flexible interpolation of knots for which we attribute a weight
- Cross-section uncertainties reflect the lack of knowledge on interaction models and data experimental limits to constrain those models
  - Uncertainties are modeled by systematics as MAQE (CCQE axial mass) or FSI (Final State Interactions) → affect each MC event

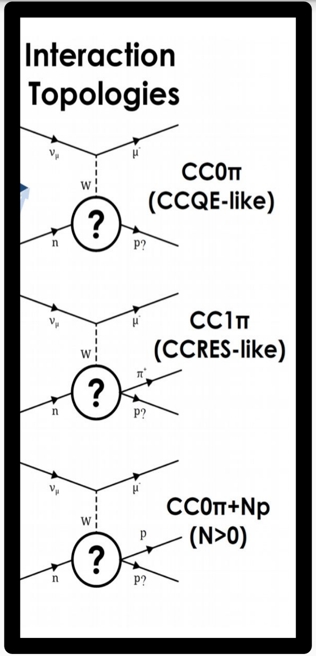




- *MAQE affects Q^2*
- *FSI models what happen to the particle after the neutrino interaction*

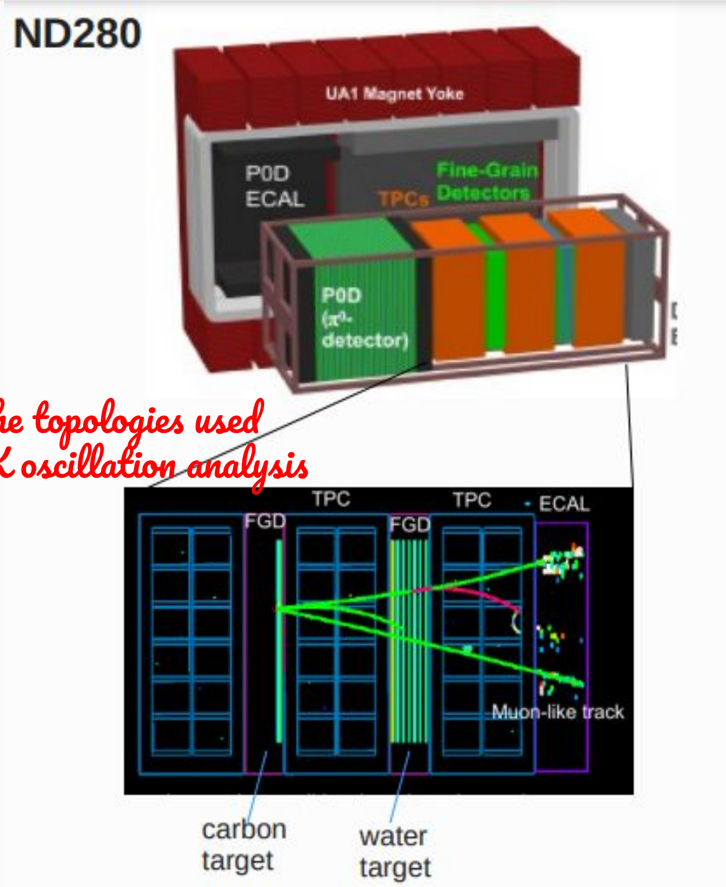
101 cross-section parameters in total in current analysis

Event selection in the detector



- interaction topologies defined on **the final state particles that exit the nucleus**
- Each topology has a **dominant interaction mode but contaminated by others**

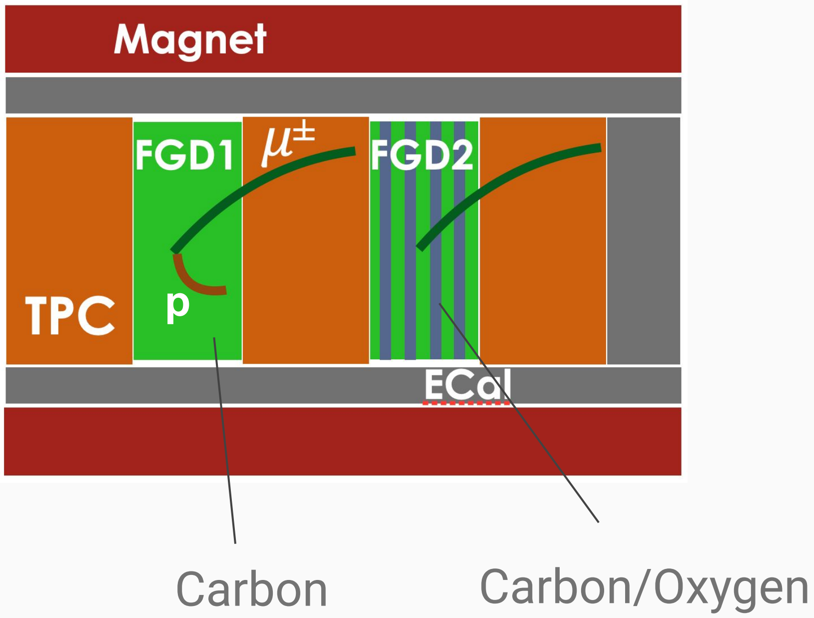
Those are the topologies used for the T2K oscillation analysis



- In the detector, neutrinos interact with bound nucleons inside nuclei
- Full tracking and particle reconstruction (FGDs, TPCs)
- Most exploited observables are muon angle/momentum
- Still using pre-upgrade era data for the analysis

Detector systematic uncertainties

- Each **detector** uncertainty is represented by a parameter in the fit
  → model how detector effects impact MC predictions
- Typical detector systematics : Track efficiency, matching efficiency, secondary interaction, target systematics



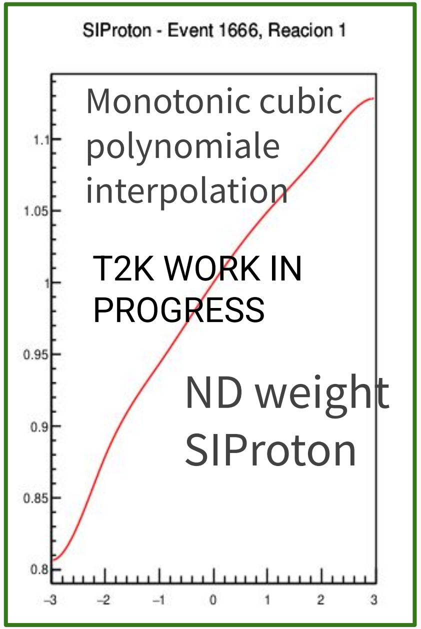
Systematic examples

Secondary Interaction = uncertainty on the probability of a proton/pion produced at the neutrino interaction vertex to undergo a secondary interaction

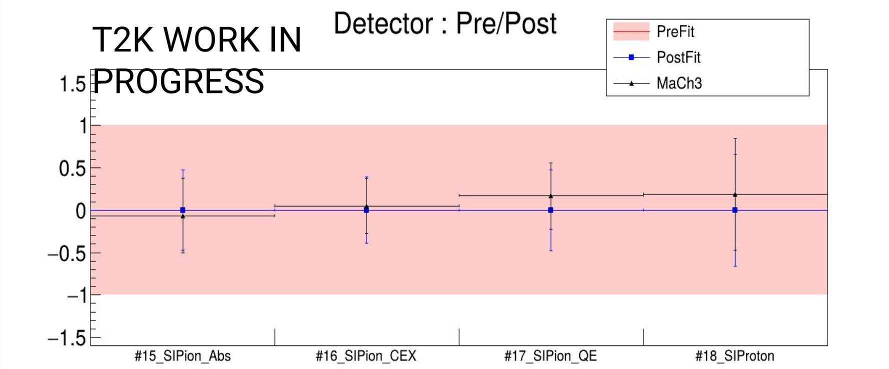
Detector systematics parameterisation

Two kind of detector systematics now:

- Weight-like physical parameters described with response functions (parameterised via splines - 81 parameters)
- Variation-like allowing event migration (encoded via a covariance matrix - 571 bins) describing the bin content variation



We can now constrain the original detector systematics
(instead of an overall effect on the event rate)

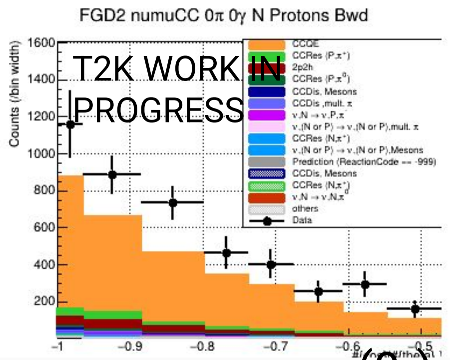


ND fit goals and frameworks used

- ND fit used to **constrain the flux and Cross-section models** and provide a tuned model for oscillation analysis
  - **Ingredients of ND fit** : samples, uncertainty models (detector, Cross-section and flux)

Frequentist approach

INPUTS
samples
systematics
(uncertainty models)



cos(θ_μ)

NEW fitter !

GUNDAM

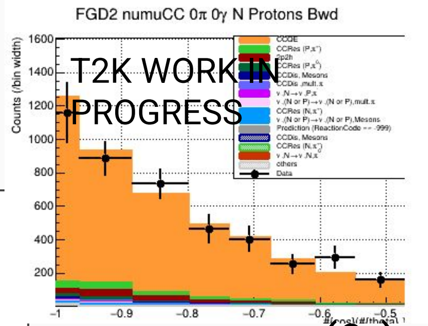
Gradient descent fit using Minuit2

GUNDAM

***G**eneralized and **U**nified **N**eutrino **D**ata **A**nalysis **M**ethods*
→ Find the link towards the gundam github page here

OUTPUTS

Event rates
LLH scans
postFit pars values
& errors



cos(θ_μ)

Minimization methods:

1. Simplex (solid but slow)
2. Migrad (used by ROOT by default)
3. Hesse (more accurate error estimation)

GUNDAM

GUNDAM :

① Datasets (select Asimov/data mode, implement inputs files, POT, Flux, xsec and Detector weights)

Samples (32 samples + binning 4140 bins)

842 parameters (xsec, detector and flux parameters, covariance matrices)

② propagator file (inputs files, samples, parameters)

PREFIT

③ main configuration file (call the propagator yaml file)

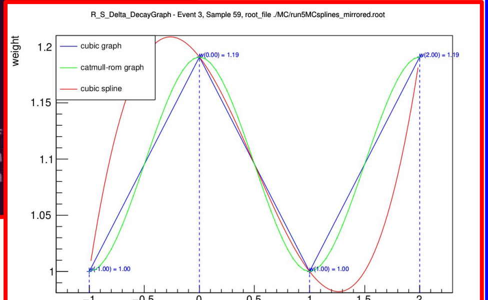
POSTFIT

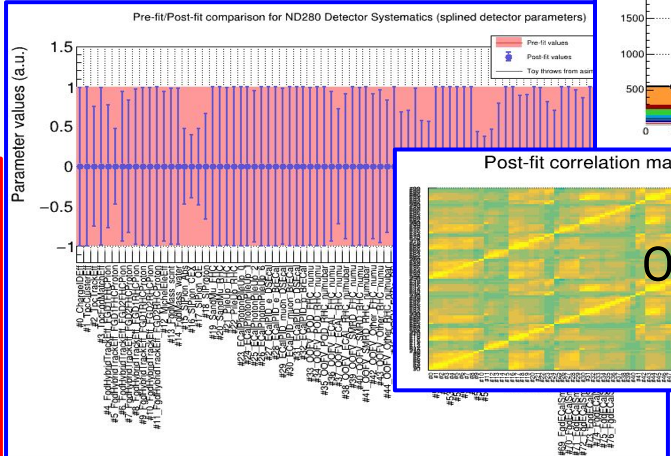
④ Fit 24h for an asimov fit

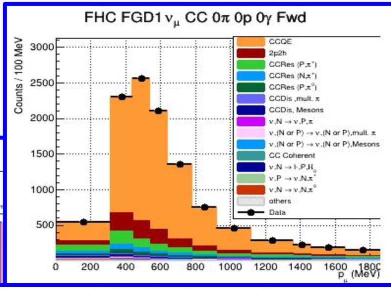
⑤ Outputs

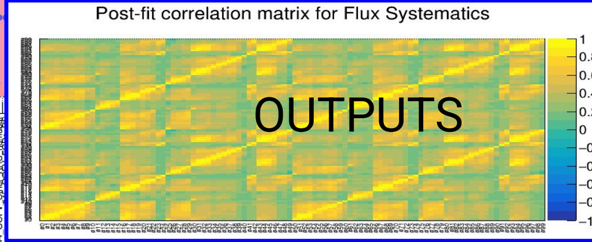
```
# bellow this version, some features are missing
minGundamVersion: 1.8.0

# folder where output files are writen
outputFolder: "./outputs/fit"

fitterEngineConfig:

  # sigma variations
  enableParamVariations: true
  allParamVariations: [-3, -1, 0, 1, 3]

  # effective when using --pca option
  pcaDeltaChi2Threshold: 1

  propagatorConfig: "./inputs/propagator/conf
  minimizerConfig: "./inputs/fitter/configMin
  scanConfig: "./inputs/propagator/output/con
~
```

INPUTS









OUTPUTS

Barlow-Beeston LLH (likelihood)

What's a log likelihood ?

- Quantifies the compatibility between observed data and predictive model
- GUNDAM uses MIGRAD to minimize -log(L) to estimate model parameters
- Barlow-Beeston formula (handles MC statistical error) :

2D binning in P_μ and $\cos\theta_\mu$

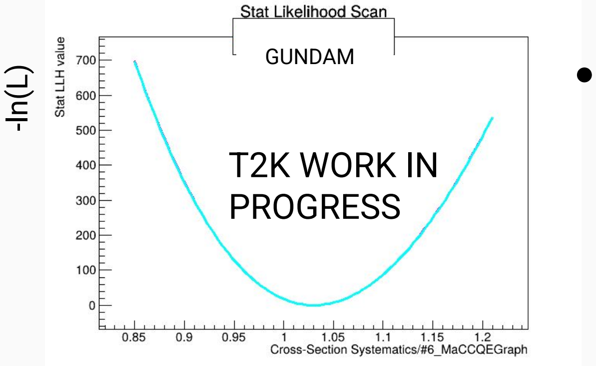
$$-2\log(L_{stat}) = 2\sum_{j}^{samples}\sum_{i}^{bins}\left[N_{MC} - N_{Data} + N_{Data}log(\frac{N_{Data}}{N_{MC}}) + \frac{(\beta_i - 1)^2}{2\sigma_{\beta_i}^2}\right]$$

$L_{stat} = L_{flux} \times L_{cross} \times L_{ND280}$

$N_{MC} = \beta_i N_{MC}^{generated}$

scaling parameters β_i

relative statistical uncertainty

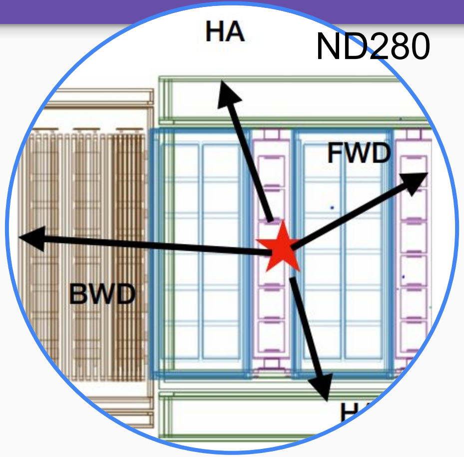


- LLH scan = vary a parameter while keeping the others fixed

→observe how the likelihood responds
→ identify the minimum
→ estimate sensitivity of the parameter

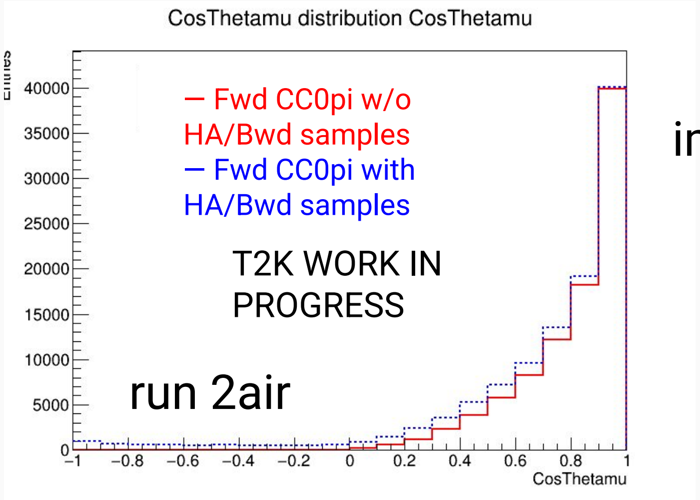
New selection at ND280

- Improved acceptance (new 4π selection)
    - **from 22 to 32 samples** (no lepton going forward only)
- Expanded phase space in a previously poorly explored region (gain ~ 13% statistics thanks to High-angle and backward going tracks)



Thanks to new 4pi selection we *expect* to :

1. detect more pions (RES & DIS interactions)
2. more protons (CCQE/2p2h contraints improvement)
3. lower energy muon spectrum

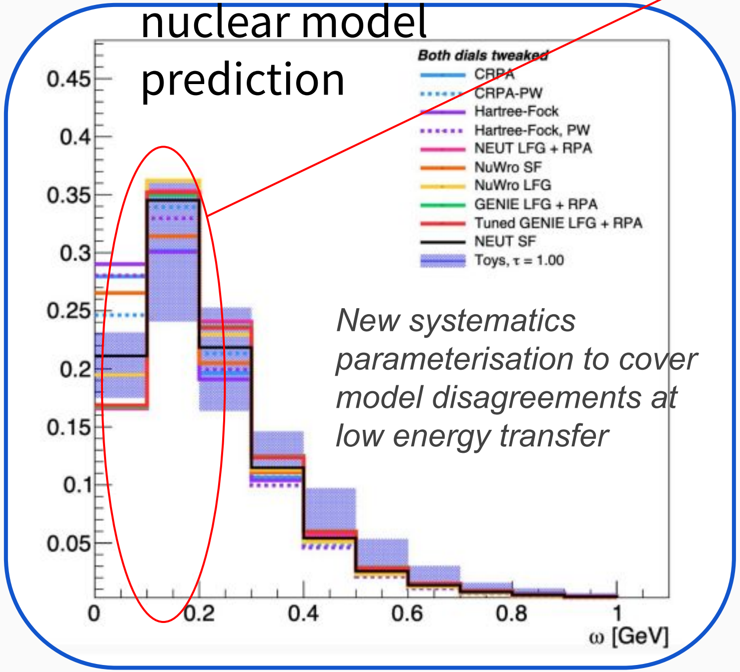


improved efficiency

+10 samples implemented in GUNDAM

New RPA/FSI strength cross-section parameters

- **Nuclear model disagree in the low Q2 region (< 0.4 GeV)**
- **Need to add more freedom to our current model**



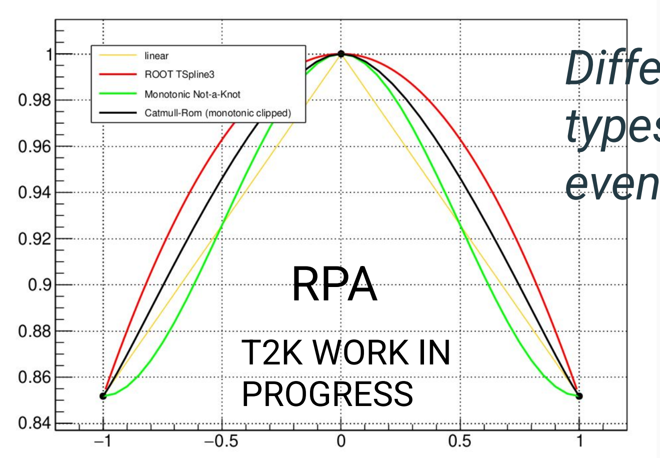
24 new parameters introduced !

- RPA (*Random Phase Approximation*) to investigate neutrino-nucleus collective interactions properties
- FSI (*Final State Interaction)* or Optical potential to dig into interaction between residual nucleus and incoming particles

Dealing with new RPA/FSI cross-section parameters

- **Mirroring** is a technique where the effect of a parameter variation is symmetrically extended to the negative side
  → symmetric response
- Allows a smoother interpolation of model prediction across the parameter range
- Different **interpolation** types exist : linear or cubic (no cusp for cubics)
- Helps GUNDAM to converge when statistical fluctuations bring the best-fit value in an unphysical parameter region

RPA and FSI parameters struggle with linear interpolation

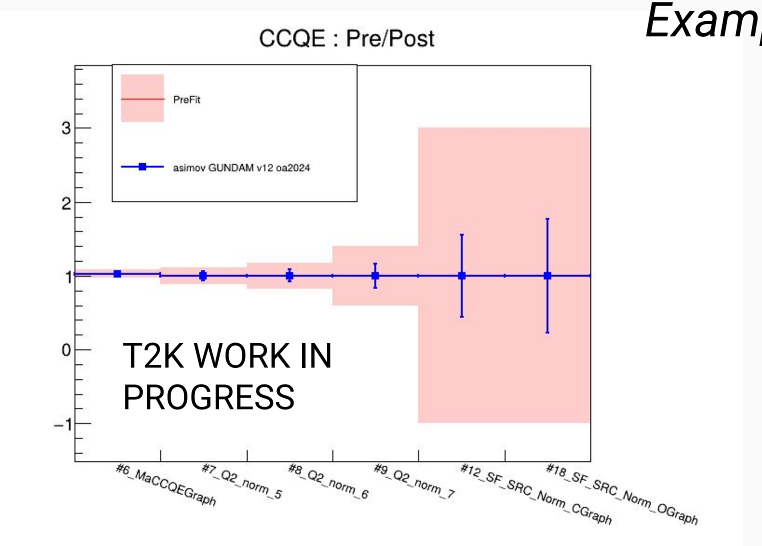


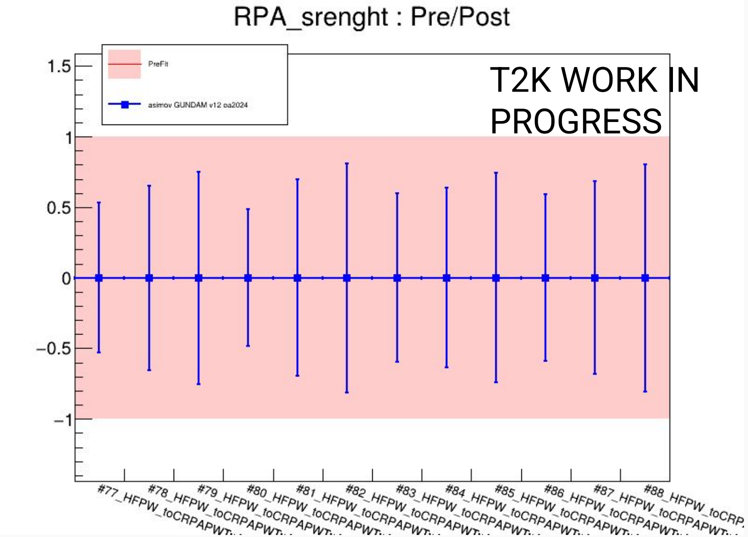
Different interpolation types used for a same event/parameter

Latest Asimov fit results

Asimov fit : predict the expected performance of the analysis (sensitivity and parameter constraints)
→ a “perfect” fit used to study the effect of the model (no statistical fluctuations)

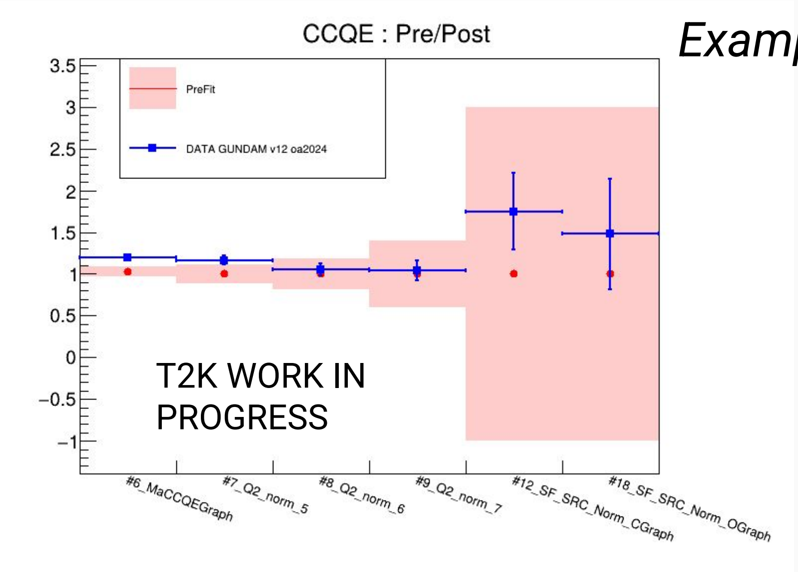
Examples



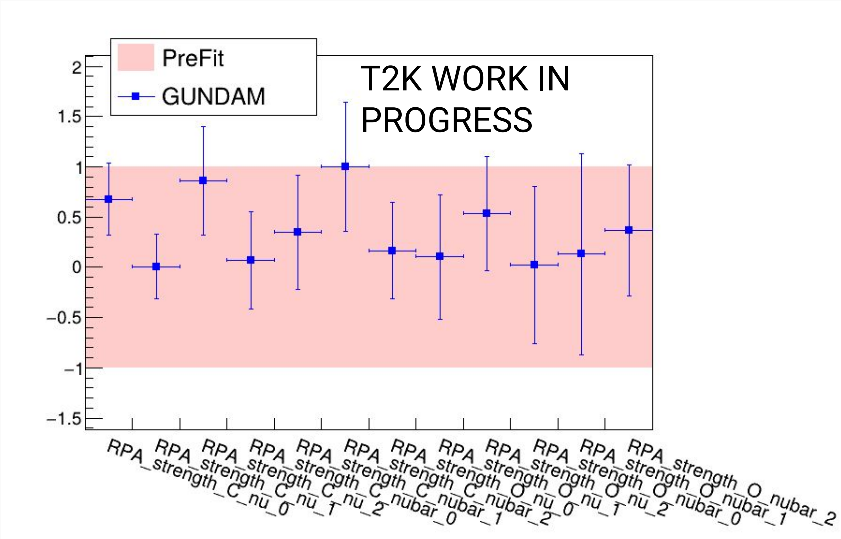


Some CCQE and splined RPA parameters to illustrate systematic constraints (asimov prefit/postfit comparison)

Datafit = performed on the actual experimental data and used to constrain systematic uncertainties to provide final results



Examples

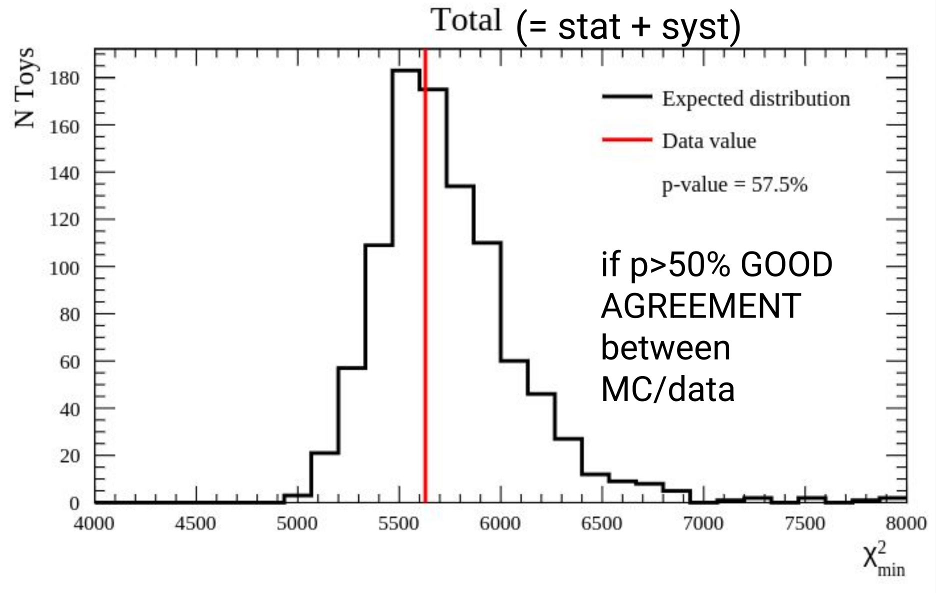


Some CCQE and splined RPA parameters to illustrate systematic constraints (data prefit/postfit comparison)

Conclusion

- ND280 role : caracterize neutrino flux before oscillation
  - constrain neutrino-nucleus interaction models
- Systematics :
  - Flux
  - Cross-section
  - Detector
- Novelties for the new analysis :
  - GUNDAM
  - Improved selection
  - Detector systematic treatment
- P-value studies
  →shows the distribution of postfit llh under the assumption the MC model is correct

Example from previous OA2021 analysis



Back-up

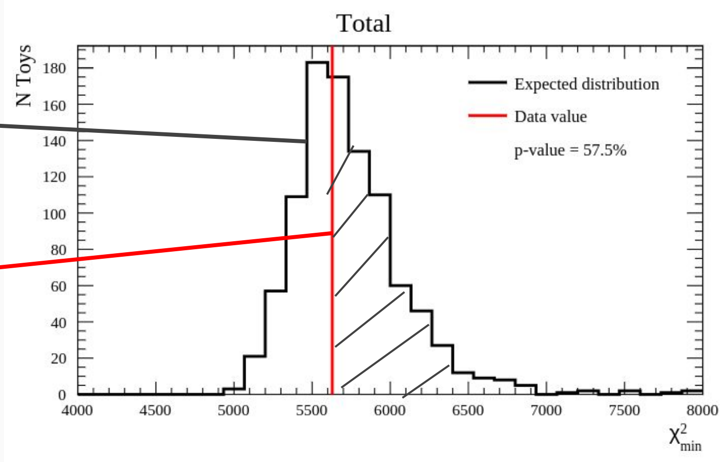
Back-up

P-value sudies

- Measure the level of agreement between the data and our MC model
- p-value should be > 50% if the model provided is robust and a good description of the data
- One can generate a set of toys starting from the MC : vary all systematics and bin contents to account for statistical fluctuations
- p-value = integral beyond the red line (actual data)/total integral of the black curve

total postfit loglh histogram

postfit loglh from actual data



RPA/FSI cross-section parameters investigations

The RPA and FSI parameters are designed to interpolate between two models:
→ 0 corresponds to the nominal model,
→ 1 corresponds to a model with enhanced RPA or FSI effects.

However, when using linear interpolation, a **problem** arises:
→ The model response has a cusp (non-smooth behavior) at parameter value 0, especially for some observables.

-->The complication around 0 comes from the fact that the lower bound of the parameter is also fixed at 0, so the fit cannot explore negative values to smooth out the cusp, making the behavior near 0 more difficult to handle.

→ This can create issues during the fit, particularly in data fits, where the optimizer may struggle or bias the parameter estimation around this point.

→ in order to solve this issue : we use mirroring + smoother interpolation type (cubics instead of linear)

RPA

→ Adjusts predictions in nuclear response functions
→ low Q2 = bigger wavelength and this gives less localised interaction
→ introduce collective behaviors in the nucleus
→ long distance effect more important (more that just the first nucleon neighbor)

FSI or optical potential = complex potential used to model how particles, such as neutrons or protons, interact with nuclei (scattering or absorption, nuclear density interactions)
It represents the average interaction between a projectile particle (like a nucleon or pion) and a target nucleus

Nuclear effects

Fermi motion: nucleons are not at rest inside the nucleus
With typical Fermi momenta up to ~200 MeV/c
This motion introduces uncertainty to the initial state → biases the reconstructed neutrino energy.

2p2h : These events mimic CCQE topologies (muon + protons, no pion), but the underlying interaction is different → also biases the reconstructed neutrino energy
→Leads to extra nucleons in the final state.

FSI (Final State Interactions): outgoing particles re-interact inside the nucleus. They can:

→ Be absorbed (e.g. pion absorption → makes a RES event look like CCQE)

→ Produce new particles (secondary interactions)

FSI blur the connection between the true interaction and what we finally observe.

--> Detectors only see the final state particles that escape the nucleus, not the original interaction.

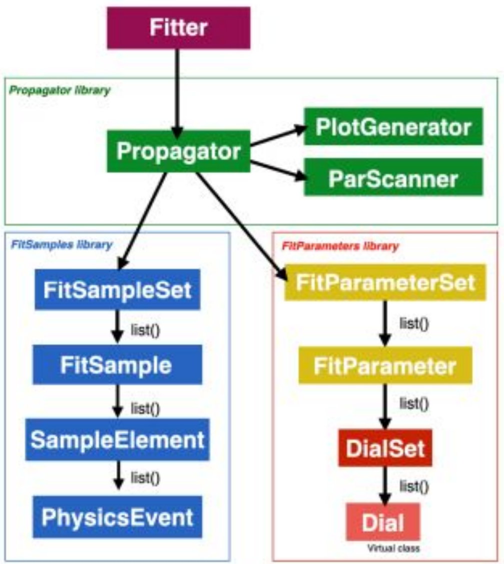
These nuclear effects introduce model-dependent uncertainties in neutrino energy reconstruction and cross-section measurements

GUNDAM : **G**eneralized and **U**nified **N**eutrino **D**ata **A**nalysis **M**ethods

- Fitter framework for current and next statistical analysis of T2K
- Designed to host multiple analysis using JSON/YAML configuration files
- C++ code based on ROOT

A new way to perform statistical analysis

- Better traceability and validation of the output → share inputs easily with other people



Flexible data handling structure
- Generic sample definition
- Parameter definition with multiple options
  - Covariance matrix
  - Apply conditions
  - Per dataset response function

Config & usage by a single gear: Propagator
- Initialise samples & parameters
- Fast propagation engine (optimised & parallelised)
- Embeds diagnostic tools (figure generators)

Adrien Blanchet pdf