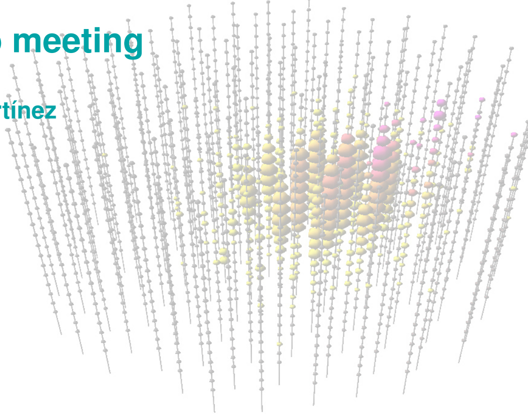
Genie re-weighting for oscillations

Simulations group meeting

Santiago Peña Martínez

APC - KM3NeT



September 25, 2024

Using GENIE for cross section systematics

- GENIE knows the accuracy of the parameters given as input to the simulation and it is able to propagate neutrino interaction uncertainties.
- Systematic errors can be computed on an event reweighing strategy. Therefore evaluate the systematics due to a given simulation input parameter, there is no need to repeat the simulation but just reweight the generated events.
- For each physics related quantity input P, GENIE introduces a systematic parameter x_P. Tweaking this systematic parameter modifies the corresponding physics parameter P as follows:

$$P \rightarrow P' = P \cdot (1 + x_p \cdot \frac{\delta P}{P})$$

where δP is the estimated standard deviation of P. Setting the $x_P = 0$ corresponds to the nominal value of parameter.

Using GENIE for cross section systematics

- The neutrino event weight w_{Sys} to account for change in physics parameters controlling the cross-section is computed as:

$$w_{sys} = \frac{d^n\sigma'_\nu / dK^n}{d^n\sigma_\nu / dK^n},$$

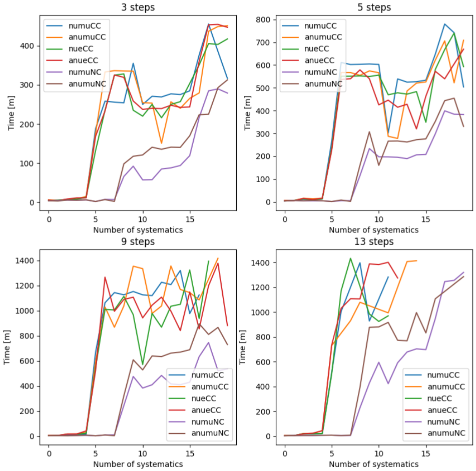
  where $d^n\sigma_\nu / dK^n$ is the nominal cross-section, $d^n\sigma'_\nu / dK^n$ is the cross-section with the modified systematic parameter and $\{K^n\}^3$ is the kinematical space factor.
- The modified global event weight will be $w'_{evt} = w_{evt} \cdot w_{sys}$.
- This weight can be incorporated into fitting software via template fittings.

Cross section weight for the mass processing

- For each event we would like to have different $w_{sys} = w_{sys}(P, x_P)$.
- gSeaGen already has some tools to compute the reweighting for a given set of systematics which can be put in place for the mass processing.
- But, there are $\sim$ 20 cross section systematics which could be computed. Additionally we would like to have various values for x_P (ideally 12 + 1 which will is 6 steps to each side of the nominal plus nominal).

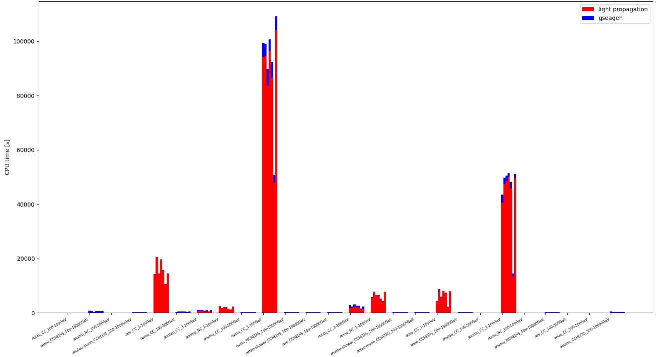
Preliminary benchmarking of the reweighting

- The generation without reweighting the systematics takes about 5 minutes for each case.
- In the most time consuming scenario an increase of about $\sim$ 300 times of CPU-time is foreseen.



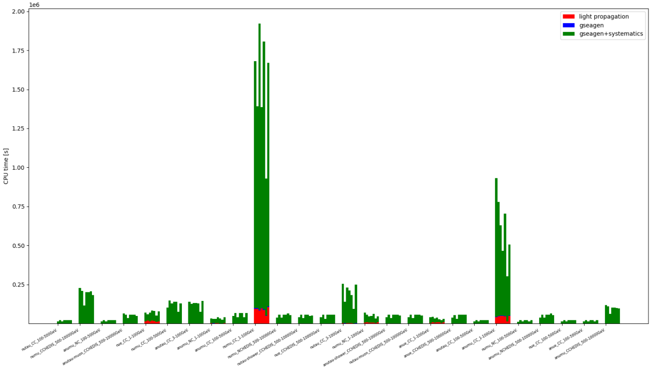
Current mass processing CPU-time contributions

- Check the CPU time for 7 runs of ORCA6.
- Each bar in the plot represents a run. To have an idea, the number of events generated for ν_μ CC 1-100 GeV is around 10^6 and for $\bar{\nu}_\mu$ CC 1-100 GeV around $5 \cdot 10^5$ for a 6 hour run.
- Current bottleneck comes from light propagation by km3sim, generation step by gSeaGen is very small in comparison.



CPU-time contributions including Xsections

- gSeaGen contribution scaled by 300 times just to have a feeling on the previous plot.
- Generation of events becomes the bottleneck by far. Not feasable.



How to solve this problem

1. Run full mass-processing pipeline.
2. Identify events which passed reconstruction at generation level.
3. Reweight only selected events.

- For a 6 hour run $\sim 2 \cdot 10^6$ neutrino events are generated, $\sim 10^3$ events are selected and reconstructed. Reduces number of events to compute the reweighting by a factor of 10^3.
- Working on tool to make this possible and to integrate it into the mass-processing pipeline.

Status of the tool

- Due to RH migrations I have been using a local installation of GENIE.
- Uses many classes already present in gSeaGen that I just copy pasted, should be easily integrated to gSeaGen once tool is done.
- Uses same concept of xml file to store systematics to implement already used in gSeaGen.
- At the moment it can cycle a GHEP (GENIE native) file and compute the reweighting given the xml input file.

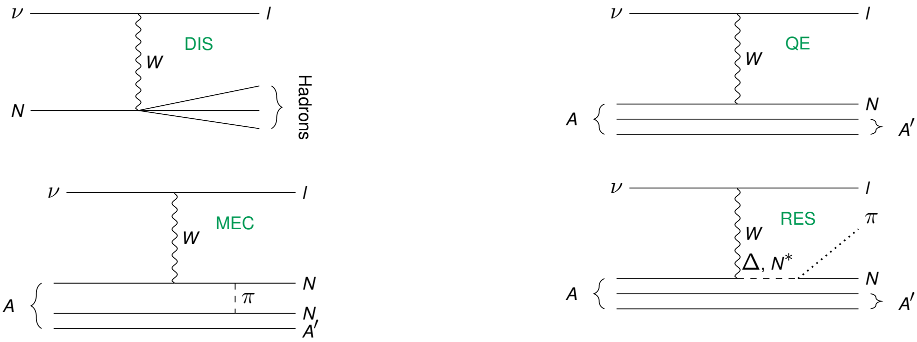
TODO:

- Write weights into friend tree.
- From seeds stored in offline files MultiHead make script which makes a new generation file.
- Read reconstruction file and new generation file to pair which events pass selection and apply re-weight only on these.

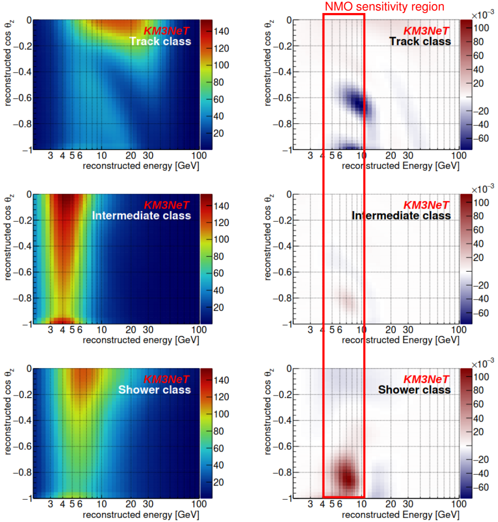
Backup

Context

- DIS (Deep Inelastic Scattering) interactions are relatively well known from perturbative QCD.
- There are significant difficulties in the modeling of QE (Quasi-Elastic) , RES (Resonance Scattering), and MEC (Meson Exchange) channels due to poor knowledge of the nucleus and low statistics from experiments. The cross section uncertainties in these channels need to be considered.



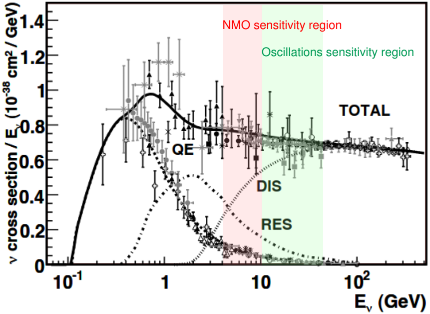
ORCA NMO sensitivity paper

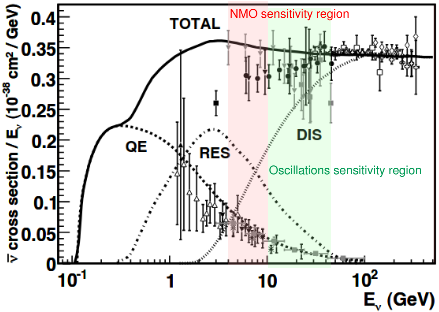


- Most sensitivity to NMO comes from the shower class.
- NMO sensitivity lies at the region of energies of 4-10 GeV.
- This energy range corresponds to the transition from QE, RES scatterings to DIS.

(Anti-)Neutrino cross sections

- Significant portion of events in the NMO sensitivity region comming from non-DIS events.
- For the 10-40 GeV energy range for ν_μ disappearance channel, most of the events are DIS.



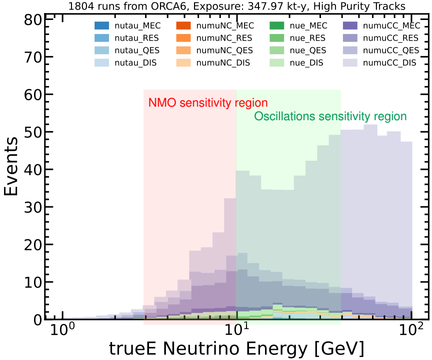


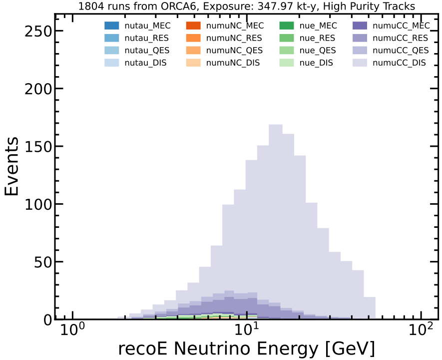
A view on our recent dataset

- I took all ORCA6 runs from the Neutrino2024 analysis which makes a sample of $\sim$ 398 kt-y and looked at the distribution of neutrino events before SWIM for each scattering channel for main analysis classes.
- Selections are the same as in the analysis except for the “anti-sparks” cut.

High Purity Tracks

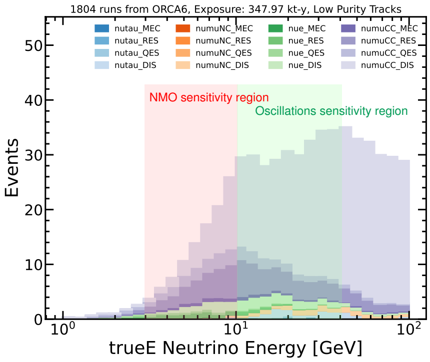
- The event sample consists mostly of $\nu_\mu + \bar{\nu}_\mu$ CC DIS events.
- Some non negligible contribution from $\nu_\mu + \bar{\nu}_\mu$ CC RES.
- Event sample from the NMO sensitive region consists of roughly half of events coming from $\nu_\mu + \bar{\nu}_\mu$ CC RES. Also a considerable contribution of $\nu_\mu + \bar{\nu}_\mu$ CC RES in the oscillations region.

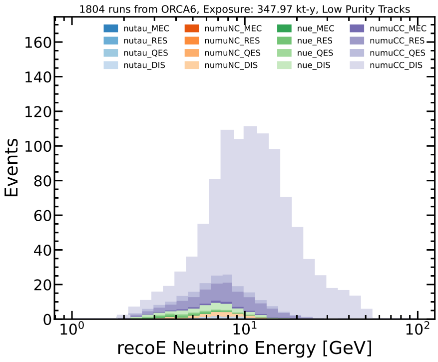




Low Purity Tracks

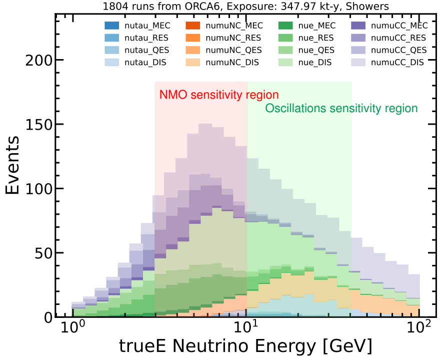
- Same comments as the High Purity Tracks class.

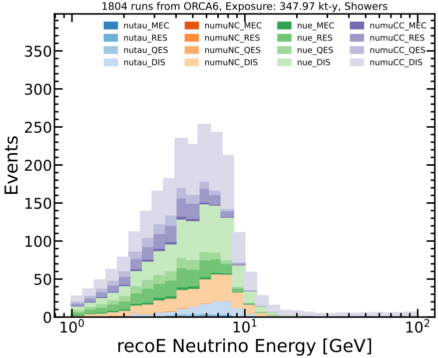




Showers

- A bit less than half of the events in the NMO sensitive region correspond to events coming from $\nu + \bar{\nu}$ CC RES.
- Contributions from $\nu + \bar{\nu}$ CC QE are non negligible.





All neutrino events

- Big contribution in general from non-DIS events for NMO measurement. For Δm^2_{31}, θ_{23} measurement mostly DIS events but a good amount of RES events are present. MEC events are a general minority.
- $\nu_\tau + \bar{\nu}_\tau$ CC events come from oscillations, they have higher energies therefore the scattering channel is DIS.
- For $\nu + \bar{\nu}$ NC only (almost) DIS. This is because π^0 is neutral and nucleons from RES or QE need energies of order of GeV to produce Cherenkov light making these channels difficult to trigger the detector.