PROCEEDINGS OF SCIENCE

AGKY Hadronization tuning with GENIE v3

Júlia Tena Vidal [a,*]**for the GENIE Collaboration**

[a] University of Liverpool, Dept. of Physics, Liverpool L69 7ZE, UK

E-mail: `j.tena-vidal@liverpool.ac.uk`

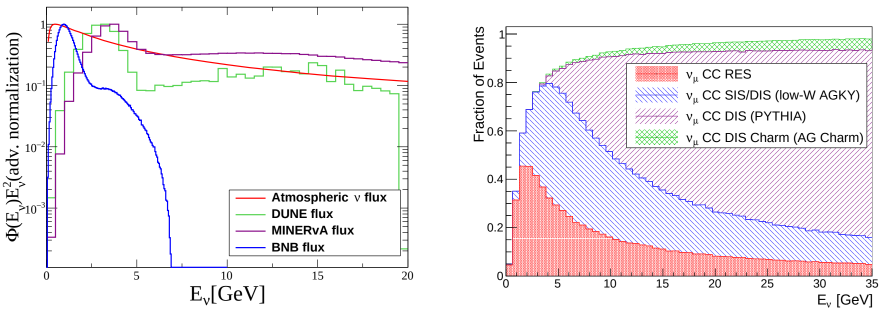
The next generation of neutrino oscillation experiments rely on the precise understanding of neutrino interactions in a wide energy range. The GENIE collaboration is constantly engaged in an effort to improve interaction models and fit them against available datasets. A lot of effort is going into pion producing processes, so far focusing on the resonant component of the pion production. This is not enough as pion production is entangled with hadronization models due to the interplay between deep inelastic scattering and resonant processes. In particular, the knowledge of the exact mixture of hadrons in showers affects the efficiency to distinguish between NC/CC events, the topological characterization, and impacts the estimation of backgrounds. The GENIE neutrino Monte Carlo [1] employs an effective low-mass hadronization model known as AGKY [2] whose validity spans from low to high W. At low invariant mass ($W < 2.3$ GeV/c^2), the model is based on the Koba-Nielsen-Olesen (KNO) scaling low while it gradually switches over to PYTHIA6 ($W > 3$ GeV/c^2) [3]. The default AGKY model parameters controlling hadronization at low invariant masses were extracted from some of the FNAL 15 ft and the BEBC analysis but PYTHIA has never been tuned to low energy neutrino hadronization data. Moreover, comparisons of the GENIE model against neutrino-induced hadron shower data exposed disagreements between different datasets, which further deteriorates at the PYTHIA region. The GENIE Collaboration addressed this issue by tuning the hadronization model against charged averaged multiplicity data on hydrogen and deuterium targets from bubble chamber experiments. All the experimental procedures followed in the original analysis have been taken into account in the simulation. The tune is performed with the Professor Framework [4] providing with a complete error estimation of the parameters and the correlation between the low-W AGKY parameters and PYTHIA parameters. In this talk, we focus on the discussion of the tuning procedure as well as the impact of the tune on other observables.

**** The 22nd International Workshop on Neutrinos from Accelerators (NuFact2021) ****
**** 6–11 Sep 2021 ****
**** Cagliari, Italy ****

---

[*] Speaker

PoS(NuFact2021)078



(a) **(b)**

Figure 1: Normalized neutrino fluxes are shown for the atmospheric neutrino flux at Kamioka, DUNE, MINERvA and BNB flux predictions (a). Breakdown of CC events as a function of the neutrino energy from ν_μ scattering on ^{40}Ar (b). The plot was obtained with GENIE v3.00.06 using tune `G18_02a_02_11a`. The main components are: RES, SIS/DIS and DIS charm production. DIS contributions are split according to the hadronization model used: low-W AGKY and PYTHIA.

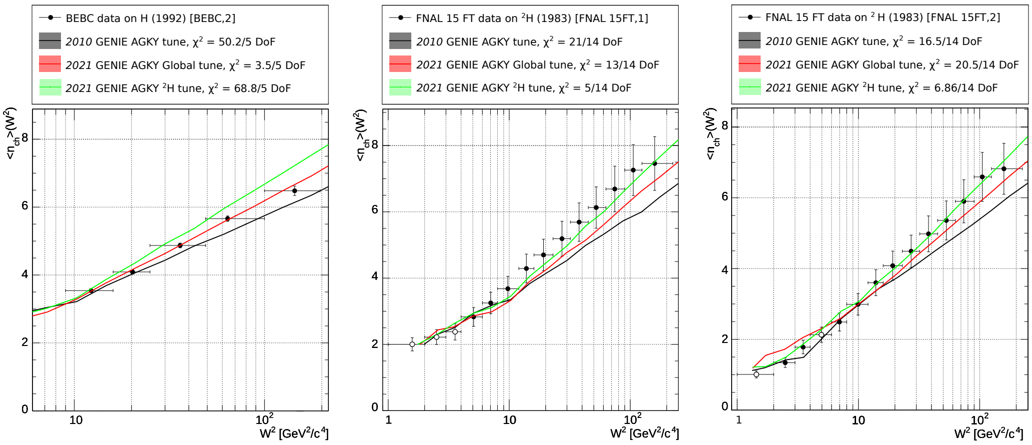
1. Introduction

Systematic uncertainties due to hadronization missmodelling can be relevant for neutrino experiments [5]. The next generation of neutrino oscillation experiments will rely on the precise understanding of neutrino interactions at the percent level. Hadronization is usually ignored in systematic uncertantiy studies but its understanding its crucial to describe pion production in neutrino experiments. In GENIE, the hadronization model is used to simulate multi-pion production events at the shallow-inelastic-scattering (SIS) and deep-inelastic-scattering (DIS) regions.

Hadronization models provide information of the final state hadrons after a DIS interaction before final-state-interactions (FSI). The knowledge of the exact mixture of hadrons in showers affects the efficiency to distinguish between Neutral-Current (NC) and charged-current (CC) events, the event topological characterization, impacts the estimation of backgrounds and the calorimetric energy reconstruction. Unfortunately, there is no unified model to describe exclusive hadronic multi-particle production over the energy range relevant for neutrino experiments. The GENIE neutrino Monte Carlo (MC) event generator [1] uses the AGKY hadronization model [2]. At low hadronic invariant mass, $W < 2.3$ GeV/c^2, the model is based on the KNO, while at high-W, $W > 3$ GeV/c^2, it is based on the PYTHIA MC [3]. In the $2.3 < W < 3$ GeV/c^2 region, the probability of using the low-W empirical model or PYTHIA changes linearly as a function of W.

Current and future experiments operate at high energies, where potential biases originating from hadronization mismodeling become important. For instance, DUNE expects about 45% of its events to come from SIS and DIS events. The neutrino energy dependence on the main inelastic components of the expected event rate for CC ν_μ-Ar^{40} scattering is shown in Fig. 1. Some relevant neutrino fluxes of interest are shown in Fig. 1 (a). It is seen that the contribution to the event rate from GeV neutrinos is mainly driven by CC RES events as well as SIS and DIS from the low-W AGKY model, whereas PYTHIA events dominate at high neutrino energies.

The AGKY model is an empirical model which needs to be tuned to neutrino data. This model was originally *tuned* and validated against bubble chamber data [2] The AGKY underlying models



(a) $\nu_\mu + p$ on hydrogen. **(b)** $\nu_\mu + p$ on deuterium. **(c)** $\nu_\mu + n$ on deuterium.

Figure 2: Comparison of averaged charged multiplicity GENIE predictions against data on hydrogen and deuterium from BEBC 7 ft and FNAL 12 ft bubble chambers.

were tuned separately: the low-W model was tuned using bubble chamber data from FNAL 15 ft and BEBC 7 ft. PYTHIA was not previously tuned against ν-scattering data. As a consequence, the current description of GENIE v3 of hadronization data is not satisfactory. Moreover, there is no information on the uncertainties due to hadronization modeling. In this work we present the first hadronization tuning against averaged charged multiplicity data from bubble chamber experiments in GENIE. The details on the exact analysis procedure can be found in Ref. [5].

2. AGKY hadronization tune

Two tunes on averaged charged multiplicity are discussed: the 2021 GENIE AGKY global tune on hydrogen and deuterium $\langle n_{ch} \rangle$ data and the 2021 GENIE AGKY ^2H tune which is performed using deuterium data only. The goal of these tunes is to improve the $\langle n_{ch} \rangle$ description for neutrino experiments and understand possible tensions between datasets. Other hadronization observables such as neutral pion multiplicity, dispersion or KNO distributions are not directly considered in the tune. In some cases, priors are imposed to preserve the agreement with some of these observables. The complete list of datasets used in the tune and the specifics of the analysis is specified in Ref. [5].

Comparisons of GENIE predictions against $\langle n_{ch} \rangle$ hadronization data are shown in Fig. 2. The main effect of the tunes is observed at the PYTHIA region: both tunes increase $\langle n_{ch} \rangle$. The best fit values for the parameters can be found in Ref. [5]. The global tune uncertainties are obtained with the profiling method. The relative error for Low-W parameters varies between 10-30%. For PYTHIA parameters, it is around 10%.

Tensions between hydrogen and deuterium samples are observed. In particular, the *2021* GENIE global tune under-predicts ^{2}H data whilst the *2021* GENIE ^{2}H tune over-predicts H data. This is highlighted in Tab. 1. After the global tune, the total agreement for hydrogen and deuterium datasets is improved. The contribution to the total χ^2 from H and ^{2}H samples is similar. Alternatively, even though the ^{2}H only goodness of fit is much better, $\chi^2_{2010(^2\mathrm{H})} = 37/52$ DoF, the agreement of the ^{2}H tune with H data is lost. This highlights the existing tensions between H and ^{2}H datasets.

Datasets	χ^2_{2010}	$\chi^2_{2021(\text{Global})}$	$\chi^2_{2021(^2\text{H})}$	DoF
All Data in tune	486	242	410	109
^{2}H Data in tune	230	105	37	52
H Data in tune	256	138	374	57

Table 1: Total χ^2 calculated with the datasets included in each fit: *2010* GENIE, χ^2_{2010}, *2021* GENIE, $\chi^2_{2021(\text{Global})}$, and *2021* GENIE, $\chi^2_{2021(^2H)}$.

3. Conclusions

In this proceedings we summarize the first GENIE tune [5] of the AGKY model [2]. The main analysis goal was to improve the GENIE agreement with neutrino charged averaged multiplicity data and provide with the first data driven constraints on hadronization parameters. The main effect of the tune is the increase of the averaged charged multiplicity for $W^2 > 10$ GeV$^2/c^4$, modeled with PYTHIA. The low-W region is also affected but constraints due to energy, momentum, charge, baryon number and strangeness conservation laws reduce the available phase space and the effect of the tuning procedure.

Tensions between hydrogen and deuterium data are observed and two separate tunes are performed: a global and a deuterium only tune. In particular, the *2021* AGKY global tune prediction under-predicts the deuterium data at the PYTHIA region whereas the deuterium only tune over-predicts the hydrogen data. Despite the tensions, the global tune shows a better agreement with the charged averaged multiplicity data and provides the first data driven analysis of this kind using neutrino interactions. This statistical analysis can be a useful input for systematic studies of modern neutrino experiments.

References

[1] C. Andreopoulos et al., *The GENIE Neutrino Monte Carlo Generator*, *Nucl. Instrum. Meth. A* **614** (2010) 87 [`0905.2517`].

[2] T. Yang, C. Andreopoulos, H. Gallagher, K. Hoffmann and P. Kehayias, *A hadronization model for few-GeV neutrino interactions*, *Eur. Phys. J. C* **63** (2009) 1 [`0904.4043`].

[3] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [`hep-ph/0603175`].

[4] A. Buckley, H. Hoeth, H. Lacker, H. Schulz and J.E. von Seggern, *Systematic event generator tuning for the LHC*, *Eur. Phys. J. C* **65** (2010) 331 [`0907.2973`].

[5] J. Tena-Vidal, C. Andreopoulos, C. Barry, S. Dennis, S. Dytman, H. Gallagher et al., *Hadronization model tuning in genie v3*, 2021.