ATMPD comments for draft v1.3

M. Miura

Please add the following authors to the author list.
F.J.P. Soler (Glasgow) Added.
B. Cortez (UCI) Added.
T. V. Ngoc (Kyoto) Added.
P.Govindaraj (Warsaw) Added.
B. R. Smithers (TRIUMF) Added.

S. Miki

- line141: When mentioning Cowan & Reines as an example, would it be better to refer to νe,ν¯e\nu_e, \bar{\nu}_eve,ν¯e rather than νμ,ν¯μ\nu_\mu, \bar{\nu}_\muνμ,ν¯μ?
  Done. Changed to nu_e.

- line169: When indicating a range, use a double hyphen (e.g., 0.1--1).
  Done.

- Table1: The notation for "Dates" should also use a double hyphen, as mentioned above.
  Done.

- line215: Define the abbreviation for "the outer detector (OD)" somewhere, as "OD" is not defined anywhere in the text.
  Done.

- line232: The fraction of Gd captures in SK6 is less than half, so saying "the majority of neutrons" might be an overstatement. "A large fraction of neutrons" might be more appropriate, as this phrase is sometimes used even when the fraction is below half.
  Done. Changed to “a large fraction of neutrons”

- lines249, 256: Use a double hyphen for range notation.
  Done.

- line272: Ensure that numbers and units do not get separated by a line break.
  Done.

- line281: Is the fiducial volume 27.2 kton?
  Done. Changed to 23.2 kton.

- line283: "Corner-clipping cosmic-ray muons" in SK typically refer to through-going muons grazing the boundary between the top and barrel. Since the background for FC events consists of stopping muons, "cosmic-ray muons that stop in the detector" might be a clearer phrase.
  Done. Changed to “cosmic-ray muons stopping in the detector.”

- line339 (and many others): "Geant3" should be written as "GEANT3" officially, while "Geant4" remains as is.
  Done.

- Figure 5 caption: The last phrase, "neutrino flux uncertainties," should be singular ("a neutrino flux uncertainty") or rephrased as "an uncertainty of absolute neutrino flux normalization" since it only refers to absolute normalization uncertainty and not uncertainties in the flavor ratio or energy dependence.
  Done. Changed to “accounting for the uncertainty in absolute neutrino flux normalization.”

- lines409, 410: Ensure that numbers and units do not get separated by a line break.
  Done.

- line436: "-100" is not in math mode, making the minus sign appear incorrect.
  Done. Changed to math mode.

- line438: "Subtracted from NHits" might be better phrased as "minus NHits."
  Done.

- line451: The explanation is unclear. Would a phrase like "The distance from the assumed vertex to the tank wall in the mean hit direction" be more appropriate?
  Done. Changed to “The shorter of the radial and vertical distances from the \textit{assumed} signal vertex to the tank wall, calculated along the average direction of the unit vectors connecting the vertex to each hit PMT.”

As a comment to this change, yeah, you're right, I myself has always had the "distance in mean hit direction" definition in mind for the whole time but checking the actual implementation in the program I realized I haven't updated the old definition which was used in Akutsu-san's thesis (see nToWall definition in his thesis, or this source code). I will update the manual, or its misleading variable name.

- Figure 8 caption: "-0.5" is not in math mode, making the minus sign appear incorrect.
  Done.

- Figures 8, 9: The thin pink lines above and below the red line—are they MC statistical errors? If so, it should be explicitly stated.
  Done.

- line666: "Visible neutrino energy" could be ambiguous (whether it means "visible energy" or "neutrino energy"). For consistency with other sections, "neutrino event visible energy" might be a better term.
  Done. Changed to "neutrino event visible energy"

- line733: Does "multiplicities are expected to linearly increase" refer to a simulation result rather than a physical expectation? If so, is "expected to" still the right phrase in this context?
  For now, leave as is. While we have not validated the full connection between the visible energy and the total neutron production, at least there is some vague connection in kinematic variables to believe that the average (n,g) multiplicity should linearly increase with visible energy. For example, <Q2> should be linearly dependent on the visible part of the energy (e.g., E_lepton in Q2~E_\nu E_lepton (1-cos\theta) for CCQE) and similarly hadronic invariant mass W2 is linearly dependent on Q2. Each hadron's momentum linearly determines the number of nuclear interactions it undergoes, and this in turn determines the number of total secondary neutrons (most of them being from neutron evaporation)
- Table 3 caption: In the fourth line, "<N>_total" should be "<N>_overall" instead.
  Done.

- line1011: Use a double hyphen for range notation.
  Done.

S. Abe

- L130: Might better to cite Minerva as well here in this context.
  Done.

- L184: It might be better to quickly summarizes SNO and KamLAND results.
  For now, leave as is, to keep the introduction short

- L296: I think this is very natural and it's not necessary to mention in this context. I could not understand what you really want to say here (I understand coverage Q^2 depend on Evis, but not sure we need this info here). Candidates of appendix.
  For now, leave as is. This is just to show that our horizontal axis is a proxy for q2, just like transverse muon momentum was used in T2K/MINERvA analyses.

- L308: Honda et al. should be italic.
  Done.

- L310: People who are not familiar this analysis might be curious this point. It's better to mention how tau. is treated.
  Changed to: ν_τ and $\bar{\nu}_\tau$ interactions were neglected as their overall presence in data is expected to be below 1\%.

- L322: I prefer having this paragraph in the same one as before. It's a bit unclear which part (Fsiorsi) you are talking abut.
  Done.

- Sec 3C: I had a impression that there are too many paragraph here a bit. For example boundary between neut and geant. is not clean at a first glance. I prefer less paragraphs or having subsub section.
  For now, leave as is.

- L489: italic
  Done.

- Fig 11: (I'm saying this by a quick read through. Sorry if I miss something). The capture ratio modification of Ref. 52 is discussed here, but it looks the MC is not used in your analysis finally (instead of it you introduced a correction factor). If this is correct, I suggest the following modifications. In the current version, the discussion of Ref 52. looks to be very main of you analysis and to be used in your analysis. 1. Move discussion of Ref 52 (L568-588) at the last of this section. I prefer to introduce the correction factor, which is used in your analysis. 2. Reduce contents of the discussion of Ref 52,

just introduce that the cause of this discrepancy is almost understood by Ref 52.

I think this kind of information is discussed and shown in Ref. 52. It might be better to avoid repeatingly show the similar things in different paper. (We might have different info but it looks you showed the similar things with Fig.4 of Ref. 52. My concern is that this might be violating license or something.
Removed Fig 11, and reduced contents already introduced in Ref 52.

- Fig 12: "Average" is unclear. Is it from AmBe? If so, please explicitly say so.
  Added in caption: triangle markers: using the overall average signal efficiency and false positive rate obtained from the given simulation) as a function of visible energy, for baseline and test simulations of atmospheric neutrino events.

- L636-643: I'm not familiar with this analysis. I had a bit difficulty to understand what you want to say in Fig. 12. Your point might be - "Average" does not reproduce the "true" - "GAM" quality has less NEUT version dependence. I'm wondering why "Average" is only shown for 5.1.4. Naively it's much straightforward to have it for the baseline MC. Isn't it? At least, I suggest to clarify the 2nd points, in the text.
  Removed baseline MC in Figure 12.

- L743-758: This is general trend of FSI and SI, which might be natural for many readers. I prefer to have this paragraph before showing Fig. 16.
  Done.

- L762-763: I think these explanations ("nu events" and "n signals") are better to mentioned in Table. III caption. That's much reader friendly.
  Done.

- Fig 18: Candidates of appendix.
  Done. Moved to appendix.

- L791: In where? Do you mean "Average" method in Fig. 12? if so please explicitly say so with the hyperlink.
  Moved to appendix.

- Fig 21: I felt a bit curious about introducing the LE multiplicity suddenly. The slope fit is still fine, but the way LE is determined seems really arbitrary. It may not always be necessary to have a reason when taking it to these quantitative discussions, but it felt a bit forced. I'm not sure what other people (including referee) think about this parameter.
  Moved to appendix.

- L843: It looks like "valley" only for single ring. Others looks flat or something. Furthermore, you assumed flat distribution for this region in LE multiplicity. I felt a bit strange with this word.
  Done. Added “of the single-ring sample” at the end of the sentence.

- Eq 6: Sorry if you mentioned somewhere. But it's better to clarify near the equation.
  Added: The goodness of fit between model predictions and data for the entire visible energy range of [0.03, 10] GeV

- Fig 22: should be "GENIE INCL" for consistent notation.
  This is clarified in the caption. For now, leave as is.

- L916-927: I felt that the chi2 has higher-level info than slope and LE multiplicity. I understand that the slope and LE is prepared as a simple parameterization, but it's really arbitrary. Probably we don't have a way to justify this method. This function is not used for further analysis and you showed chi2, which is much better than slope and LE. So I personally think we can omit this discussion. I prefer to remove or move it to appendix.
  Done. Moved to appendix.

- Fig 24: The boundary between solid and hatched areas are difficult to see. Please modify the design.
  Removed figure and replaced it with a table.

- Fig 24: It's a bit hard to see the difference especially for GENIE INCL choices. I prefer to have this with a TABLE, which might be more suitable than figure for this kind of quantitative disucussion.
  Removed figure and replaced it with a table.

- Fig 24: "normalization constraint" might be enough.
  Done.

- L946-949: Needs citation.
  Done.

- L958: It's a bit strange wording. The point is that it assummes constant Fermi momentum and applied step-function. I don't think "strict" is appropriate expression.
  This is an expression I think is used quite often, e.g., here, here, here, here. I guess this is okay as long as we have the following explanation “which forbids all nucleon scattering below the predefined oxygen Fermi momentum of 225 MeV/c”.

- L965-968: We may need citation.
  Added: In contrast, the Liège INC model applies a probabilistic approach for nucleon collisions throughout the cascade (except for the initial collision), accounting for nucleon holes in the surrounding phase space volume \cite{incl4, incl_pb}.

- L975-983: This sentense is a bit long and does not read well. We may also need clarifying citation about this point.
  Rephrased: In the 0.2--1GeV/c range, the number of outgoing neutrons predicted by the Liège INC and Geant4 Bertini cascade models is lower than that from the \texttt{GENIE hA}, \texttt{hN}, and \texttt{NEUT} nucleon FSI models, as illustrated in Figure\ref{fig:out_n_mom}. One contributing factor may be the inclusion of light cluster formation—such as deuterons, tritons, and alpha particles—among cascade products, which reduces the number of isolated neutrons contributing to the observable signal. Ref.\ \cite{ershova} also identifies cluster formation as a major mechanism behind the reduction of outgoing protons in the Liège INC model. Additional in-medium effects, considered either explicitly or implicitly in the two models—such as short-range nucleon repulsion and local nuclear density reduction after each collision—further reduce the number of nucleon-nucleon collisions during the cascade.

- Summary: If we can have future prospects more concretely, that enhances significance of your paper. Currently you only said that you compare data and various MCs (and a general trend of slope LE multipliicty, which is natural without looking at SK data).

  Your conclusion can go further. I think this data indicates out issue of G3 and G4 BERT (not G4 BERT_PC). It should be emphasized. And also please think about adding a messsage (statement) to the community.
  For example,
  - What's the future prospects for SK? Can we use neutron for physics analysis? Can we go with these model prediction level? What's the remaining tasks?
  - Will we still use SK-IV default, G3 etc??
  - What's the status and future of this neutron/evaporation community?
  - Connection between other nu experiments SNO, Minerva, KL, etc. you mentioned at the beginning. I think it's better to emphasize about BERT_PC and BERT for these commnunity here.
  - What about non-neutrino experiments situation? For example E525.

  Added:The observed discrepancy with the evaporation model \cite{dostrovsky} commonly used in Bertini-based cascade models aligns with previous studies, including neutron measurements from SK neutral-current quasielastic selections \cite{sakai} and secondary gamma-ray measurements using Germanium detectors with a neutron beam on a water target \cite{e525}. Our data also supports the selection of specific models, such as \texttt{GENIE INCL} and \texttt{G4 BERT_PC}, which reduce the uncertainty in total neutron production from atmospheric neutrino events to around 10\%. These models may help explain the deficits in total neutron capture signals reported by T2K \cite{t2k_neutron_2019} and SNO \cite{sno_neutron_2019}, and offer insights into the low-energy neutron signal deficits observed in MINERvA \cite{mv_neutron_2019}. Lastly, this work will improve the use of neutron tagging for identifying atmospheric antineutrino events at SK, enhancing sensitivity to neutrino mass ordering and searches for proton decays and the diffuse supernova neutrino background.