We would like to express our gratitude to the referee for the appreciation of our work and for providing constructive comments on our manuscript. The valuable feedback has been used to improve the quality and clarity of our work. In response to the suggestions, we address below each of the points raised by the referee and describe the actions that have been taken in the revised version of the manuscript. To facilitate the review process, in the revised version of the manuscript we have highlighted all the changes that have been made with respect to the original submission.

1. *Page 3, below (2.8). It is noted that nuclear corrections are not included when interpreting the neutrino structure function data in terms of proton PDFs. It would be good to provide some brief discussion of the size of these and the associated uncertainties.*

   In order to address the need for a comment regarding the sizes of the nuclear corrections in proton PDF fits, we have swapped the order of the last two sentences in that paragraph. The cited references, and the references therein, provide extensive discussions regarding the procedure to include nuclear corrections and their effects.

2. *Table 2.2. It is not clear to me at the point where these numbers are produced exactly how the cross section inputs corresponding to these numbers are calculated. So I think a reference forward to Section 2.5, where the theory settings are described, is needed. Although even then, for completeness giving the PDF set that is used and some uncertain ty on the event rates here would be useful.*

   The numbers in Table 2.2 are computed by integrating the event yields shown in Eq. 2.12 in which the differential cross sections are computed using the central value of the PDF4LHC21 set. For completeness, we have added a sentence above Eq. (2.16) that mention the specific PDF set used to compute these numbers, and where we also indicate that the calculation is performed at NLO in the QCD expansion. Although Section 2.5 indeed provides some descriptions regarding the computation of the differential cross sections (which were also described in Section 2.1), it mainly focuses on the construction of the pseudodata needed for the fit/profiling.

   This said, we would like to emphasize that the predicted event rates in Table 2.2 depend only mildly on the specific settings of the DIS cross-section entering the calculation. Furthermore, these event rates are only used to determine the expected bib-by-bin statistical uncertainty. Hence our results are relatively independent on the theoretical settings entering the calculation of the event yields in Table 2.2.

3. *Page 11. The discussion about consistency between the PDF set and theory settings used to produce the pseudodata and those entering the fit/profiling is in my view not correct or at least too strong. In particular, while it is perfectly reasonable to keep these the same there is definitely no requirement to, as is currently strongly implied in the discussion. In real PDF fits we often see that the fit quality for a given dataset does not follow textbook expectations, with $\chi^2/N \sim 1$. So some inconsistency between data and theory is often observed, rather than being artificial. Indeed, it is precisely because of this effect that tolerances (which are included in e.g. the PDF4LHC profiling) are included. In other words, one could perfectly reasonably generate pseudodata with a different PDF set, or different theory settings in order to emulate this inconsistency. One is free not to, but it should not be suggested that complete consistency is the only option here. It is a choice that is made, and not the only possible one.*

   We agree with the referee that, when generating pseudo-data, one may or may not want to assume the same underlying truth as in the prior of the corresponding PDF fit. Indeed, data inconsistencies are present in real data.

Our motivation to keep these settings consistent is that in this work we aim to investigate what is the PDF constraining power of the FPF measurements assuming that there are no inconsistencies, which is the best-case scenario. In the presence of inconsistencies, one finds two competing effects: a shift in the central value of the PDF and a reduction of the PDF error whose magnitude depends on the assumed inconsistencies. For example, if we generate the FPF pseudo-data far from the PDF4LHC21 prior, the main impact of the data will be to shift the central prediction, rather than to reduce the uncertainty.

Another reason to assume this consistency is that when carrying out Hessian profiling, one keeps the PDF parametrisation fixed. In the presence of inconsistencies, it may be that the best fit with FPF data requires a different parametrisation and hence a different PDF error analysis. So Hessian profiling is most reliable in the absence of data inconsistencies.

In the revised version of the paper, we have emphasized our two-fold region to assume full consistency between the prior PDF and the FPF pseudo-data: to quantify the PDF reach of the FPF pseudo-data in the most advantageous scenario, and to ensure the procedural validity of the Hessian profiling procedure.

This said, the suggestion from the referee to generate data with an "inconsistent" PDF set is valuable and it would be interesting to carry out this exercise in the future. To emphasize this, we have added a sentence stating that generating pseudo-data with a different PDF set from the prior would test the capabilities of the FPF to disentangle between different PDF sets among them.

4. *Page 15, and Fig. 3.3. Perhaps some explanation of why dV and to a lesser extend uV benefits from charge-lepton identification could be provided?*

   We have slightly rephrased the subsequent sentences to make it clear that the improvements seen in d_V and to a lesser extend in u_V can be understood by looking at analytical LO expressions of the structure functions in terms of the PDFs for ν and $\bar{\nu}$ provided in Sect. 2.

5. *Page 16, and Appendix A. The fact that the FASERν (and SND@LHC) projections lead to a very limited improvement on the PDF uncertainties is rather hidden in a paragraph here, and then in the appendix. In my view, this 'negative' result should be given more prominence. It after all motivates the improvements that might come with the FPF. I would suggest moving this to the main body of the text and starting with this as the first study.*

   The main reason to put the study of the constraints provided by FASERν in the Appendix was because the analysis in Section 3 is mainly dedicated to the FPF experiments. We nevertheless agree with the referee that such findings should be highlighted and used as a motivation for the FPF. We have moved App. A to the first subsection of Sect 3, to highlight that the ultimate constraining power of LHC neutrino experiments can only be realised with the FPF.

   Actually, this result is also relevant concerning another points raised by the referee below: the fact that the FASERν and SND@LHC data not not impact the PDFs indicates that the interpretation of their event rate measurements depends only loosely on the modelling of the deep-inelastic scattering interaction. Hence, the FASERν and SND@LHC data can be safely used to tune the incoming neutrino fluxes and learn for example about PDFs at very small-x, without compromising the validity of the interpretation. So this "null result" is also relevant in this context, and motivates the use of the current FASERν and SND@LHC data to validate models of charm production, forward light hadron production, and small-x QCD.

6. *Page 22. I am rather unsure about the approach for presenting results here, and in particular in showing numbers without including systematic errors, which are described as being 'optimistic'. Having zero systematic uncertainties is surely unrealistic, rather than optimistic, so I feel as though a clearer justification for this needs to be given. Even more importantly, the labelling of the result without systematic uncertainties as 'FPF' and those with as 'FPF*' is surely the wrong way round, given it implies that the case where the systematic uncertainties are not included is the default in some sense. So these should be swapped, and the rationale behind showing numbers without systematic uncertainties accounted for more clearly presented.*

   The referee is surely correct in that zero-systematic uncertainties is not realistic. However, the reason to also show results with statistical uncertainties only is twofold. First, it substantiates the claim that FPF measurements will be dominated by systematics, rather than statistical uncertainties, which is a very important and non-trivial consideration informing detector design. The FPF detectors have not been designed with "precision physics" as main goal, and showing that improving systematics has strong physics benefits is important in the context of these studies. Note that this is not the case in FASERν and SND@LHC, where the physics reach (at least concerning proton structure studies via neutrino DIS) is completely limited by the low statistics available.

   Second, and perhaps most importantly, we cannot claim to have a fully realistic modelling of systematic errors, and in particular we miss a detailed estimate of their correlation, which is known to be decisive in stablishing the PDF sensitivity. Such analysis can only be carried by the experimental collaborations themselves and be based on a complete detector simulation, and indeed our study provides a strong motivation for this. In this respect, it is also worth noting that the estimation of the systematic uncertainties in our analysis is very much conservative, and already the insights provided by our paper are helping the experimentalists to improve designs to reduce systematics. So there is a clear added value in separating the two scenarios.

   Following the referee's suggestion, the results which only include the statistical errors are now labelled "FPF$\star$" while the one that also account for the systematics are labelled "FPF". We have also extended the rationale for showing results in the statistics-only scenario, alongside the considerations described above.

7. *Section 4 and elsewhere. Given these are HL-LHC projections, somewhere these should be compared with the HL-LHC PDFs of Ref [34]. This would surely be the fairer comparison, or in any case will give a clearer picture of where things may stand.*

   The reason why we did not provide comparisons with the HL-LHC PDFs of Ref. [34] is because the pseudodata used to determine these PDFs were produced with the PDF4LHC15 set which does not account for the recent LHC measurements. The most consistent comparisons would be to re-do the analysis done in ref. [34] using PDF4LHC21 and then compare the resulting PDFs with our determination, something which is clearly outside the scope of this paper.

   This said, to answer the question from the referee, we have repeated the phenomenological study of Sect 4 comparing PDF4LHC15 with PDF4LHC15+HLHC PDFs for the same processes. Specifically, we have added to the paper an Appendix where the phenomenology study of Sect 4 is repeated now for the PDF4LHC15 with PDF4LHC15+HLHC PDFs. From this comparison, one can observe that .... ADD

   Furthermore, as mentioned in the paper, the PDF constraints from the FPF and from the HL-LHC are fully orthogonal and complementary, and the former have the very important property that possible contamination from new physics on the PDF fits can be completely ignored, since this process is driven by Q^2 values outside the possible presence of BSM physics. Specially should a large-mass anomaly be revealed at the HL-LHC, having the fully independent validation of the large-x PDFs provided by the FPF would be extremely valuable for its interpretation.

We now address the comments of referee B

1. *Page 3, lines 34-36 column 2 "since in general the strange and charm ... not expected to vanish". I agree but I suggest adding a reference or a comment supporting this statement.*

   We have added the following references which explore strange and charm PDF asymmetries [1, 2]. We have also added a reference to the recent NNPDF study of the intrinsic charm asymmetry in the proton, showing that these effects are potentially sizable and could hence be relevant for the interpretation of the FPF measurements.

2. *p. 4 l. 30-35 col.2 "Also, to identify... event yields." Why a cut on the hadronic energy should properly simulate a cut on the number of charged tracks emerging from the interaction vertex?*

   The charged track multiplicity is expected to grow with the hadronic system's invariant mass. We make this clear and cite a neutrino-Hydrogen interaction multiplicity study [3] as well as a FASER paper [4] which also point this out. We also append this sentence:

   "..., as the charged track multiplicity is expected to grow with the invariant mass of the hadronic final state W [3, 4]."

3. *Table 2.1 What is the meaning of the asterisk in "FLArE (*)"? I cannot see it referred to in the caption*

   The meaning of the asterisk was to be clear that this row represents 2 proposed detectors, as opposed to the remaining rows which are each one detector. We make this clear by appending the end of the caption:

   "..., which we denote for the two detectors as as FLArE(*)."

4. *p.5 l. 38 col.2 "is made of thin sensitive layers" What is the material/detector the layers are made of?*

   The target layers are interleaved with tungsten and emulsion. We make this clear by ammending this sentence:

   "...sensitive layers of emulsion..."

5. *p.5 l.49 col 1 and col 2 ("AdvSND" and "FLArE") Can you put a reference about these experiments? Did you take the info from ref [10]?*

   We have added the original FLArE experiment proposal paper as well as the FPF whitepaper (ref[10]) to the FLArE section. For AdvSND, we have added included the reference to the FPF whitepaper, which is currently the most detailed description of this proposed experiment.

6. *p.6 l. 50-53 col.2 "Here we neglect efficiency .. simulation." I am aware that this might be difficult to simulate at this stage but I think this effect is non-negligible due to misidentification (a large background could enter this sample). Can you comment further on it? In addition, some guesses on misidentification capabilities can be drawn by the current FASER and SND@CERN capabilities.*

   Indeed, uncertainties due to misidentification are in general not expected to be negligible. While there are no detailed studies of charm tagging efficiencies at the FPF, as we point out in the paper, charm tagging can be done by through multiple methods, including reconstructing the topology of D-meson decays, and through dimuon events. With multiple techniques at hand, one should be able to determine in-situ quite well the D-meson tagging efficiencies, for example comparing the dimuon selection method with the charm tagging via the emulsion detector. Furthermore, the availability of different experiments based on complementary techniques will help to cross-calibrate these effects.

It is worth pointing out that our framework is flexible enough to accomodate variations in the modelling of systematic uncertainties, and in particular there is no conceptual issue preventing anyone to repeat our analysis for different treatments of efficiencies and other related experimental parameters.

7. *p.9 l. 41-43 col.2 "We note that ... in our estimation." (see also the conclusions p. 27 l. 52-55 col.1) I am puzzled by this statement. It is well known that flux systematics in Faser and SND@CERN play a prominent role and set the normalization (and shape?) of eqn 2.12. How can you ignore such an important effect in your analyses? Even if the PDF constraints are marginally affected by the normalization of the neutrino flux, some of your considerations about the on-axis (faser and faser2) versus the off-axis experiments (SND@CERN and its upgrades) may be affected by flux uncertainties. Can you comment on it in the paper?*

   Indeed, the large neutrino flux uncertainties are well established, and the shape and normalization of the flux can vary widely between different models, although the neutrino muon component which is mostly relevant for neutrino DIS experiments is the one which is known the best.

   It is important to note that neutrino measurements actually constrain the product of flux and cross-section - each of these components brings an uncertainty with it, and a full analysis with real data would constrain them simultaneously. We wish to motivate this joint analysis by calculating the impact that FPF data can bring to a PDF fit. In this sense, by taking the flux to be known we can understand the full reach of FPF data on cross-section measurements, analogous to what was done in 2309.10417. Moreover, projections of FPF data on flux measurements has shown that HL-LHC data can bring flux uncertainties to a sub-percent level.

   We would like to raise two more considerations which justify, in this work, to assume that the incoming neutrino fluxes are known:

   - As shown in appendix A, now moved to Sect 3, FASERν and SND@LHC data from run III do not have sensitivity on the modelling of the neutrino-nucleus interaction cross-section. This means that these measurements can be safely used, before the start of the FPF, to cross-calibrate incoming neutrino fluxes and markedly improve their predictions for subsequent analysis. Just due to this, by the time the FPF starts data taking the neutrino fluxes will be known more precisely than now.
   - The flux and the neutrino DIS cross-section have very different kinematic dependences. The incoming nuetrino flux is mostly fixed from the (E_ν, y_ν) dependence of the event rates, while the PDFs in the DIS cross-section are fixed by the (x, Q^2) dependence. By measuring event rates differentially in the three kinematic variables (E_ν, x, Q^2), and also accessing the y_ν dependence of the incoming neutrinos, one can efficiency disentangle the fluxes from the DIS cross-section. Of course, one needs to explicitely demonstrate this via a dedicated analysis, but there is no conceptual reason preventing us from this goal.

   We thank the referee for this important point and add the following paragraph after we introduce the flux model in the middle of section 2.3.

"As pointed out in Ref. [5] there are notable neutrino flux uncertainties, as various event generators do not agree on the forward parent hadron spectra. If the spread of various generators' predictions was taken as a means of flux uncertainty, corresponding to a $\lesssim 50\%$ uncertainty on the interacting muon neutrino spectrum, this would be a significant systematic if left unresolved. However, it is noteworthy that many existing predictions are yet to be tuned for the purposes of experiments such as those planned for the FPF. Nevertheless, there are projections of FPF measurements which would reduce this uncertainty to the sub-percent level already in the context of the contemporary predictions, based on parametrizing their expected correlations [6], as well as efforts to describe the uncertainty in a data-driven way while improving the modelling of forward hadronization [7]. However, it is important to note that forward neutrino experiments actually constrain the product of flux and cross-section, and one must be assumed to measure the other. In a full analysis, the flux and cross-section would be constrained simultaneously in a joint measurement, utilizing their different kinematic dependences on x, Q^2, E_ν and neutrino rapidity η_ν. In our study, we aim to understand the full impact of FPF data on the PDF fit, thus motivating this future joint measurement. To this aim, we take the neutrino flux to be known and focus on the irreducible systematics associated with event reconstruction. With this assumption, we will show that Run 3 measruements will not be sufficient to impact PDF fits. Instead, Run 3 measurements could be used to calibrate incoming neutrino fluxes, effectively reducing the large uncertainties by the time FPF data is collected in the future. The expected reduction of FPF neutrino flux uncertainties further justifies our choice to take the neutrino flux as known.

8. *p.11 l. 33-37 col.1 Are you sure that an assumption of $f_{corr} = 0.5$ is realistic? The estimates made in ref.[34] are for LHC experiments, not for fixed-target experiments like those of the FPF.*

   In the absence of full detector simulation and detailed correlation model, it is not possible to answer this point conclusively. What we can say is that we have repeated our analysis for $f_{corr} = 1$ and find that our main qualitative conclusions remain unchanged.

9. *Fig. 2.4 In the bottom-left plot, I see a sharp increase in the fractional error at about x= 4 e-2. Why?*

   The points near x=4e-2 with large uncertainties have small E_h, and are approaching the $E_h > 100$ GeV threshold that we set in Table 2.1. When we fluctuate the data according to our definition of uncertainty (Eq. 2.19 with E_h instead of E_ℓ), these points drop below the acceptance and thus contribute to a fluctuation in this bin. We have added to the revised version of the paper a short discussion to explain this observation.

10. *p.19 l.50-53 col 1 and in general along the paper: I cannot see a discussion of the impact of neutral-current events at the FPF. Can you further comment on it?*

    There are expected to be roughly as many neutral current neutrino scattering events as charged current events. However, due to the lack of information on both the incoming and outgoing neutrino, the full event cannot be reconstructed, only the total hadronic energy can be measured. One could try to use the total NC event rate as a means of constraining the integrated PDF, however we expect that this would negligibly improve the impact on the PDF constraints as compared to fully reconstructed CC events. Indeed, if anything, NC scattering brings in less sensitivity to PDFs than CC events, as can be seen from the LO decomposition of the corresponding structure functions.

    We have added to Sect 2 a short discussion on why it is very likely that NC events do not provide useful PDF information: first of all, event kinematics cannot be reconstructed on an event by event basis as in the CC case, second, that systematics will be large since are dominated by the hadronic final state, and third, that NC scattering in general provides less complete information on the PDF flavour decomposition as compared to CC scattering.

References

[1] R. S. Sufian, T. Liu, G. F. de Téramond, H. G. Dosch, S. J. Brodsky, A. Deur, M. T. Islam, and B.-Q. Ma, "Nonperturbative strange-quark sea from lattice QCD, light-front holography, and meson-baryon fluctuation models," *Phys. Rev. D* **98** no. 11, (2018) 114004, arXiv:1809.04975 [hep-ph].

[2] R. S. Sufian, T. Liu, A. Alexandru, S. J. Brodsky, G. F. de Téramond, H. G. Dosch, T. Draper, K.-F. Liu, and Y.-B. Yang, "Constraints on charm-anticharm asymmetry in the nucleon from lattice QCD," *Phys. Lett. B* **808** (2020) 135633, arXiv:2003.01078 [hep-lat].

[3] **Aachen-Bonn-CERN-Munich-Oxford** Collaboration, P. Allen *et al.*, "Multiplicity Distributions in Neutrino - Hydrogen Interactions," *Nucl. Phys. B* **181** (1981) 385–402.

[4] **FASER** Collaboration, H. Abreu *et al.*, "Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC," *Eur. Phys. J. C* **80** no. 1, (2020) 61, arXiv:1908.02310 [hep-ex].

[5] F. Kling and L. J. Nevay, "Forward neutrino fluxes at the LHC," *Phys. Rev. D* **104** no. 11, (2021) 113008, arXiv:2105.08270 [hep-ph].

[6] F. Kling, T. Mäkelä, and S. Trojanowski, "Investigating the fluxes and physics potential of LHC neutrino experiments," *Phys. Rev. D* **108** no. 9, (2023) 095020, arXiv:2309.10417 [hep-ph].

[7] M. Fieg, F. Kling, H. Schulz, and T. Sjöstrand, "Tuning Pythia for Forward Physics Experiments," arXiv:2309.08604 [hep-ph].