OUTP-15-AAA

Neural network determination of nuclear PDFs in lead

Carlota A. Casas[1], Stefano Carrazza[2], Alberto Guffanti[3], Nathan P. Hartland[1] and Juan Rojo[1]

[1] *Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, University of Oxford, OX1 3NP Oxford, UK*
[2] *Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*
[3] *Niels Bohr International Academy and Discovery Center, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*

Abstract

The determination of the nuclear modifications of the free nucleon PDFs inside lead nuclei is an important ingredient for the proton-lead and lead-lead heavy ion program at the Large Hadron Collider (LHC). As compared to the PDFs of free nucleons, inside nuclei PDFs are modified by a variety of effects from shadowing, the EMC effect, anti-shadowing and Fermi motion. In this work we extend the NNPDF fitting methodology, already successfully applied for unpolarized and polarized PDF fits, to the determination of the nuclear parton distributions of lead. The basic strategy is to use existing nuclear PDF sets, in particular EPS09 and DSSZ12, to convert all available neutral-current DIS $l^{\pm}N$ scattering data into lead structure functions, and fit these using the same theory as those of the free nucleon case. Our fits are performed both at NLO and NNLO, and account for heavy quark schemes using the FONLL general-mass scheme. Our results indicate that current determinations of nPDFs substantially underestimate the uncertainty in the nuclear gluon modifications, and that the total quarks are reasonable well known. The present works provides a suitable baseline to further study the impact of LHC pPb data into nuclear PDF analysis by means of the Bayesian reweighting method.

Contents

1 Introduction

At the LHC, the precision determination of the parton distributions of the proton is an important ingredient for many analysis, from Higgs boson characterization to New Physics searches and precision Standard Model measurements. In addition to the proton-proton program, the LHC is also running a very exciting program proton-lead and lead-lead collisions. The main physics goals of this program is to understand the quark-gluon plasma created in the hot and dense collisions, as well as to characterize the initial state of cold nuclear matter. To carry out this program, the determination of the the modifications of the PDFs of nucleons inside a lead nucleus is of paramount importance.

Nuclear parton distributions are determined mostly from neutral-current and charged-current deep-inelastic structure functions on nuclear targets. In addition, data from other experiments like Drell-Yan production at RHIC provide additional information. Various groups provide public sets of nuclear parton distributions, such as EPS09 [1], DSSZ08 [2]. Since the experimental information on nuclear collisions is much more scarce than the the proton case, nuclear PDF uncertainties are affected by substantial uncertainties.

In a series of papers [3–13], the NNPDF collaboration has introduced a methodology aimed at reducing as much as possible this procedural uncertainty.

As a byproduct of this work, we present for the first time a nuclear PDF fit at NNLO, accounting for heavy quark mass effects using the FONLL general-mass scheme. Although both NNLO effects and heavy quark mass effects are presumably smaller than current experimental uncertainties, it does not cost anything to do a proper job since we have all the machinery ready.

Now we want to apply it to the case of nuclear PDFs ....

There are two main motivations to provide a set of lead PDFs using the NNPDF framework. First one Second, it provides a suitable starting point to study the impact of LHC pPb into the lead nuclear PDFs, using for example the Bayesian reweighting technique [14,15]. A large number of observables in pPb collisions have been advocated to reduce the large uncertainties in nuclear PDFs including low mass Drell-Yan and inclusive hadroproduction [16], dijet production [17, 18], isolated photon production [19] or even top quark pair producition [20] Some of these measurements have already become available at Run I, and more will be so at the upcoming Run II heavy ion program. The availability of nNNPDFs thus provides an alternative method to quantify the impact of available LHC pPb measurements.

2 Experimental data

In this section we summarize the data sets that have been used in the present determination of the parton distribution of lead nuclei. We also discuss the kinematical cuts and the treatment of experimental uncertainties. Then we present the basic strategy that we adopt in this work for the conversion of the experimental data for the ratio of structure functions of different nuclei, $F_2^{A_1}$ and $F_2^{A_2}$, into a common observable which is chosen to be the ratio of lead F_2^{Pb} and deuteron F_2^d structure functions.

2.1 Nuclear DIS structure functions data

In this first study of nuclear PDFs with the NNPDF methodology we restrict ourselves to fixed-target neutral current lepton-nuclei deep-inelastic scattering structure functions. In particular, we include the the NMC [21–24], SLAC139 [25] and EMC [26] measurements of ratios of nuclear structure functions, defined as

$$\frac{F_2^{A_1}(x,Q^2)}{F_2^{A_2}(x,Q^2)} \,. \tag{1}$$

The different types of nuclei A_1 and A_2 that are used in each of these experiments is summarized in Table 1, as well as the number of data points that survive our baseline kinematical cuts and the corresponding publication reference. In terms of neutral current DIS nuclear data, the experiments included in Table 1 are the same of the EPS09 analysis.

In this fit we use the standard kinematical cuts of the NNPDF unpolarized fits, namely $Q^2 \geq Q^2_{\rm min} = 3.5$ GeV^2 and $W^2 \geq W^2_{\rm min} = 12.5$ GeV^2, which are chosen to minimize the impact of low scale non-perturbative corrections and of higher twists. Note that this cut is slightly more stringent that the corresponding kinematical cut $Q^2 \geq 1.69$ used in EPS09 [1].

Since the experimental covariance matrix is not available for any of the nuclear DIS experiments summarized in Table 1, we add in quadrature the statistical and systematic errors for each data point, assuming that the systematic errors are fully uncorrelated

$$\sigma_i^{\rm exp,tot} = \left(\sigma_i^{\rm exp,stat} + \sigma_i^{\rm exp,sys}\right)^2 \,. \tag{2}$$

Note that since measurements are performed in terms of ratios of structure functions between different nuclei, Eq. (1), common multiplicative systematic uncertainties such as luminosity will cancel in the ratio. As we discuss in the next section, we will add to Eq. (2) another contribution due to the theoretical uncertainty in the conversion from Eq. (1) to the ratio of lead over deuteron structure functions

$$\frac{F_2^{Pb}(x,Q^2)}{F_2^{d}(x,Q^2)} \,, \tag{3}$$

which is the quantity which will be used for the nuclear PDF fits presented in this work.

In Fig. 1 we show the kinematical coverage in the (x,Q^2) plane of the DIS nuclear data included in the present analysis, and summarized in Table 1. Only those points that satisfy the baseline kinematical cuts $Q^2 \geq Q^2_{\rm min} = 3.5$ GeV^2 and $W^2 \geq W^2_{\rm min} = 12.5$ GeV^2 are shown in this figure.

2.2 Conversion of nuclear data to lead structure functions

In this section we explain the strategy that we use in this work based on the conversion of experimental data for the ratio of structure functions of different nuclei, $F_2^{A_1}$ and $F_2^{A_2}$, into

Experiment	A_1/A_2	$N_{\rm dat}$	Reference
SLAC E-139	He(4)/D	2	[25]
NMC 95, re.	He/D	13	[21]
NMC 95	Li(6)/D	12	[22]
NMC 95, Q^2 dependence	Li/D	113	[22]
SLAC E-139	Be(9)/D	2	[25]
NMC 96	Be/C	15	[23]
CERN EMC	C(12)/D	9	[26]
SLAC E-139	C/D	1	[25]
NMC 95, NMC 95, re.	C/D	13	[21, 22]
NMC 95, Q^2 dependence	C/D	125	[22]
NMC 95, re.	C/Li	10	[21]
SLAC E-139	Al(27)/D	2	[25]
NMC 96	Al/C	15	[23]
SLAC E-139	Ca(40)/D	1	[25]
NMC 95, re.	Ca/D	13	[21]
NMC 95, re.	Ca/Li	10	[21]
NMC 96	Ca/C	15	[23]
SLAC E-139	Fe(56)/D	4	[25]
NMC 96	Fe/C	15	[23]
CERN EMC	Cu(64)/D	19	[26]
SLAC E-139	Ag(108)/D	1	[25]
CERN EMC	Sn(117)/C	8	[26]
NMC 96	Sn/C	15	[23]
NMC 96, Q^2 dependence	Sn/C	121	[24]
SLAC E-139	Au(197)/D	2	[25]
NMC 96	Pb/C	15	[23]
Total		571	

Table 1: Data sets included in the present analysis. In the second column we indicate the nuclei A_1 and A_2 which have been used in the measurement, see Eq. (1), where when needed we indicate the atomic mass number in parentheses. In the third column the indicate the number of data points that survive the baseline kinematical cuts. In the last column we indicate the corresponding publication reference.

a common observable which is chosen to be the ratio of lead F_2^{Pb} and deuteron F_2^d structure functions. The main advantage of this method is that it is not necessary anymore to determine from the data the dependence of the nuclear modification ratios $R_i(x,Q^2,A)$ on the atomic mass number A. Therefore, we can simply fit the parton distributions of nucleons in lead nuclei using exactly the same technology as in the case of the unpolarized global NNPDF fits. After presenting the relevant conversion formulae, we discuss how theoretical uncertainties are propagated from the input nuclear PDFs used to the experimental data, and show representative results for a variety of experiments.

The basic idea is the following. Available nuclear DIS structure function data is presented as ratios of structure functions measured on different nuclear targets

$$\frac{F_2^{A_1}(x,Q^2)}{F_2^{A_2}(x,Q^2)}\,, \tag{4}$$

and we would like to express these measurements in terms of a common ratio of the lead structure functions over the deuteron structure functions,

$$\frac{F_2^{\rm Pb}(x,Q^2)}{F_2^{\rm d}(x,Q^2)}\,. \tag{5}$$

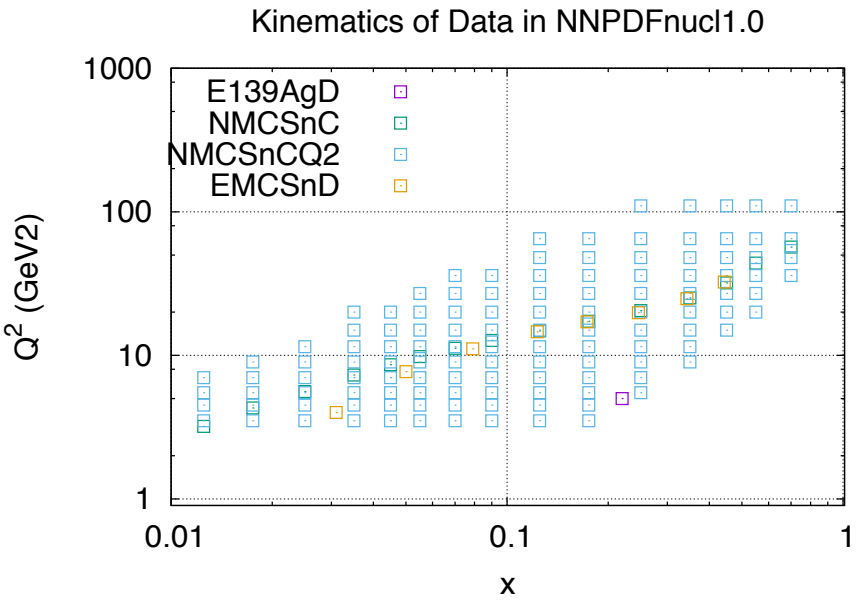

Figure 1: Kinematical coverage in the (x, Q^2) plane of the DIS nuclear data included in the present analysis, and summarized in Table 1. Only those points that satisfy the baseline kinematical cuts $Q^2 \geq Q^2_{\rm min} = 3.5$ GeV^2 and $W^2 \geq W^2_{\rm min} = 12.5$ GeV^2 are shown in this figure.

To achieve this, we use the results of a previous nuclear fit to write this conversion as follows

$$\left(\frac{F_2^{\rm Pb}(x,Q^2)}{F_2^{\rm d}(x,Q^2)}\right)_{\rm exp+th} = \left(\frac{F_2^{A_1}(x,Q^2)}{F_2^{A_2}(x,Q^2)}\right)_{\rm th} \cdot \mathcal{C}_{\rm th}(x,Q^2,A_1,A_2) \ , \tag{6}$$

where the theoretical conversion factor $\mathcal{C}_{\rm th}(x,Q^2,A_1,A_2)$ is evaluated using an existing nuclear PDF analysis as follows

$$\mathcal{C}_{\rm th}(x,Q^2,A_1,A_2) \equiv \left(\frac{F_2^{A_2}(x,Q^2)}{F_2^{A_1}(x,Q^2)}\right)_{\rm th} \cdot \left(\frac{F_2^{\rm Pb}(x,Q^2)}{F_2^{\rm d}(x,Q^2)}\right)_{\rm th} \ . \tag{7}$$

The advantage of using the conversion factor Eq. (7) is that a substantial fraction of the nuclear PDF uncertainties (as well as many other theory uncertainties such as higher perturbative orders) cancel in the ratio of nuclear structure functions, specially if A_1 is not to different to A_2. Therefore, by ensuring that the theory uncertainty associated to the conversion factor is not much larger than the experimental uncertainty, we can reliably use the ratio of lead to deuteron structure functions Eq. (6) as a good proxy of the original experimental measurements. In addition, since the conversion factor is expressed in terms of nuclear structure functions (which are fitted) rather than in individual nuclear PDFs, we expect them to be reasonably independent of the specific choice of input nuclear PDF fit.

In the calculation of the conversion factor Eq. (7) we use either the EPS09 or the DSSZ NLO nuclear PDF fits. For the free nucleon PDF we use the MSTW08 NLO set, consistently with the original nuclear PDF fits, though we have checked that the dependence on this choice is really mild due to the cancellations in the ratios. In the calculation of the structure function of deuterium $F_2^{\rm d}$ we assume isospin symmetry and neglect any nuclear effect, which are known to be negligible as compared to the typical uncertainties in nuclear PDF fits.[1]. The structure

[1]The impact of nuclear effects in unpolarized proton PDF analysis as been discussed for example in Refs. [27–29]

functions in Eq. (7) are computed using the original EPS09 fitting code, in a zero-mass variable-flavor number scheme at NLO, and no target mass corrections are included. Note that the actual fit settings used to extract the lead nuclear structure functions from Eq. (6) will be different, but the approximations enumerate above are enough for the computation of the conversion factor.

The conversion factor Eq. (7) is affected by the PDF uncertainties of the nuclear fit used in its computation. In both EPS09 and DSSZ, PDF uncertainties are estimated with the Hessian method [30]. Therefore, we can use the usual expressions for linear error propagation in the Hessian approach to compute the nuclear PDF uncertainties associated to Eq. (7). On the other hand, we neglect the free proton PDF uncertainties from the MSTW08 set, which cancel to a very good approximation in the ratios that define the conversion factor and that are much smaller than nuclear PDF uncertainties.

In the Hessian approach, the PDF uncertainties in a given cross-section $\mathcal{O} = \mathcal{O}[f]$ can be computed as follows

$$(\Delta\mathcal{O})^2 = \frac{1}{4}\sum_{k=1}^{N_{\rm eig}} \left(\mathcal{O}[S_k^+] - \mathcal{O}[S_k^-]\right)^2 , \tag{8}$$

where $\mathcal{O}[S_k^+]$ and $\mathcal{O}[S_k^-]$ denote the values of the quantity $\mathcal{O}$, computed by the nPDF error sets S_k^+ and S_k^- and S_0 is the central PDF set, and $N_{\rm eig}$ is the number of asymmetric eigenvectors of this Hessian PDF set.

Using Eq. (8), we can compute the nuclear PDF uncertainty associated to the conversion factor as follows

$$(\Delta\mathcal{C})^2 = \frac{1}{4}\sum_{k=1}^{N_{\rm eig}} \left(\mathcal{C}[S_k^+] - \mathcal{C}[S_k^-]\right)^2 = \frac{1}{4}\sum_{k=1}^{N_{\rm eig}} \left(\frac{F_2^{A_2}(S_k^+)}{F_2^{A_1}(S_k^+)}\frac{F_2^{\rm Pb}(S_k^+)}{F_2^{\rm d}(S_0)} - \frac{F_2^{A_2}(S_k^-)}{F_2^{A_1}(S_k^-)}\frac{F_2^{\rm Pb}(S_k^-)}{F_2^{\rm d}(S_0)}\right)^2 , \tag{9}$$

where note that the deuteron structure functions are always calculated with the central set S_0, since nuclear PDF uncertainties are ignored in this case. The correlation between the same eigenvector variations between different nuclei ensures that the total PDF uncertainty Eq. (9) associated to the conversion factor will be typically rather smaller than the corresponding experimental uncertainties, guaranteeing the robustness of our strategy.

In Fig. 2 we illustrate the typical size of the PDF uncertainties associated with the conversion factor, computed using Eq. (9), with the corresponding experimental uncertainties for a number of representative datasets. We show the results when the conversion factor is computed with EPS09 and with DSSZ, as percentage compared to the central value of the conversion factor. As can be seen, for most of the points in x and Q^2 the PDFs uncertainties in the conversion factor are smaller than the total experimental uncertainty. Therefore, by cutting those data points for which the theory error is too large, we can end up with a reduced nuclear data set expressed in terms of $F_2^{\rm Pb}/F_2^{\rm d}$.

Therefore, in this work we will keep only those data points for the experiments listed in Table 1 which satisfy the condition that

$$\frac{\Delta\mathcal{C}}{\mathcal{C}} \leq K_{\rm th}\cdot\frac{\sigma^{\rm exp,tot}}{\left(F_2^{A_1}/F_2^{A_1}\right)_{\rm exp}} , \tag{10}$$

where $\sigma^{\rm exp,tot}$ is the total experimental uncertainty, Eq. (2), and $\left(F_2^{A_1}/F_2^{A_1}\right)_{\rm exp}$ is the corresponding central value of the original experimental measurement. The value of $K_{\rm th}$ cannot be much larger than 1, else our nuclear PDF fit would be biased by the choice of input nuclear

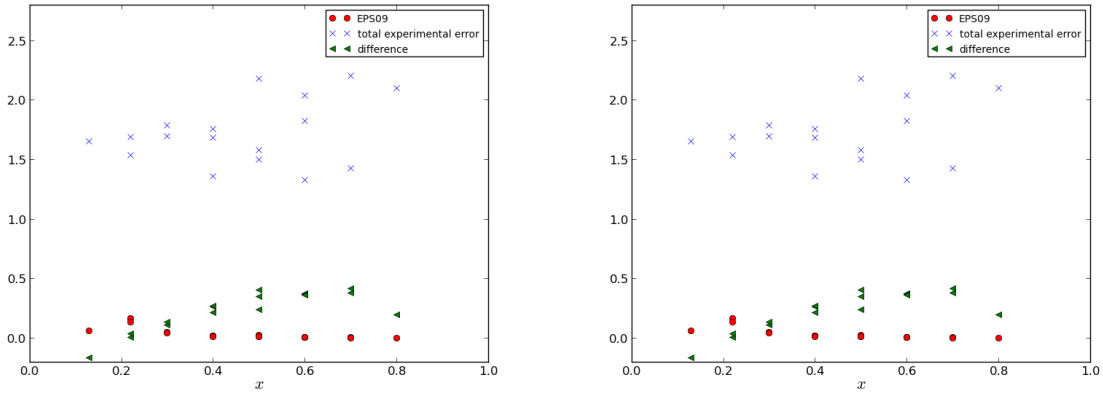

Figure 2: Left plot: comparison of the magnitude of the total experimental uncertainty in percentage (blue crosses) with the size of the PDF uncertainty in the conversion factor, in percent computed with respect to the central value, using EPS09 as input PDF set, for the E139 experiment. Right plot: the same comparison now using DSSZ as input in the calculation of the conversion factors.

PDF set. In this work we will explore different choices for $K_{\rm th}$ in the range $[1.0, 1.5]$, and study the stability if the fit results with respect to this choice.

For the data points that satisfy the condition Eq. (10) we add the theoretical uncertainty in the conversion factor Eq. (9) as an additional systematic uncertainty. This time however this new theory systematic is treated to be as fully correlated among all the data bins of a given experiment, since it is obtained from the same underlying set of nuclear PDFs. Therefore, the total experimental uncertainty taking into account the theoretical uncertainty in the conversion factor will be

$$\sigma_i^{\rm exp,tot} = \left(\sigma_i^{\rm exp,stat,2} + \sigma_i^{\rm exp,sys,2} + \sigma_i^{\rm th,2}\right)^{1/2} , \tag{11}$$

where the theoretical error is defined as

$$\sigma_i^{\rm th} \equiv \frac{\Delta\mathcal{C}}{\mathcal{C}} \cdot \left(\frac{F_2^{\rm Pb}(x,Q^2)}{F_2^{\rm d}(x,Q^2)}\right)_{\rm exp+th} \tag{12}$$

and where note that the statistical and systematic experimental uncertainties have also been corrected with the same conversion factor as the central experimental data.

In Fig. 3 we show representative subset of the NMC data on ratios of nuclear structure functions, after the conversion to the common ratio $F_2^{\rm Pb}/F_2^{\rm d}$ using the strategy discussed in the text. In the upper plots the conversion has been performed using EPS09, while in the lower plots it has been done using DSSZ. In both cases, the corresponding theoretical uncertainty from the conversion factor has been added in quadrature to the total experimental uncertainty, as indicated in Eq. (11).

3 Fitting methodology

In this section we discuss the fitting methodology, including, the sum rules, the flavor decomposition, positivity, the parameterization of nuclear PDFs as ratios and similar

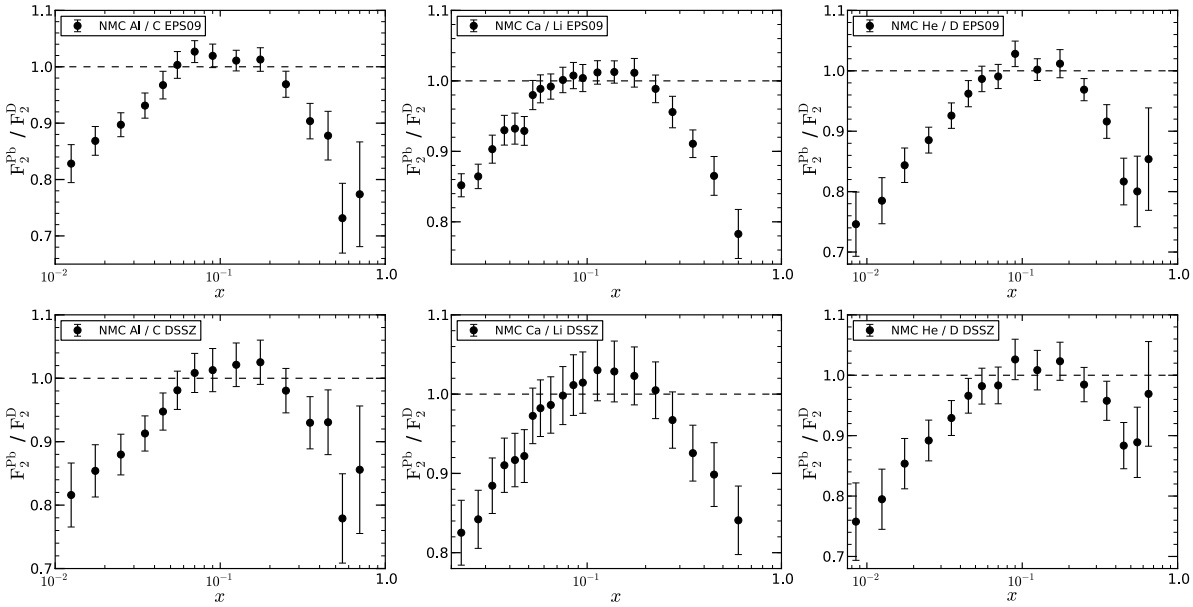

Figure 3: A representative subset of the NMC data on ratios of nuclear structure functions, after the conversion to the common ratio $F_2^{\rm Pb}/F_2^{\rm d}$ using the strategy discussed in the text. In the upper plots the conversion has been performed using EPS09, while in the lower plots it has been done using DSSZ. In both cases, the corresponding theoretical uncertainty from the conversion factor has been added in quadrature to the total experimental uncertainty.

3.1 Nuclear PDF parametrizations

In this work we want to determine the PDFs of bound nucleons that for a lead nucleus Since nuclear-induced modifications of the free-nucleon PDF are known to be moderate, it is advantageous from the methodological point of view to express the lead PDFs $q_i^{\rm Pb}(x,Q^2)$ as follows

$$q_i^{\rm Pb}(x,Q^2) \equiv R_i(x,Q^2)\cdot q_i^{\rm p}(x,Q^2)\,, \quad i=-n_f,\ldots,n_f\,, \tag{13}$$

where $q_i^{\rm p}(x,Q^2)$ are the free proton PDFs and $R_i^{\rm Pb}(x,Q^2)$ is the nuclear modification factor of lead nuclei that we want to extract from the data. In the absence of nuclear effects $R_i^{\rm Pb}(x,Q^2)=1$. The number of active quark flavors in the variable-flavor-number scheme that we use is denoted by n_f, and the flavor index $i=0$ corresponds to the gluon PDF. Thanks to the conversion strategy presented in Sect. 2.2, we express all available DIS nuclear data in terms of lead structure functions, and therefore the fitting is simplified since we don't need to extract the dependence of the nuclear modification factor $R_i^{\rm A}(x,Q^2)$ on the atomic mass number A.

The nuclear modification factors of lead $R_i^{\rm Pb}(x,Q^2)$ are parametrized at the input evolution scale $Q_0=1$ GeV with a feed-forward artificial neural network, as customary in the NNPDF fits

$$R_i^{\rm Pb}(x,Q_0^2) = {\rm NN}_i(x)\,. \tag{14}$$

with no further preprocessing terms. This is another advantage of parametrising of the lead PDFs as ratios with respect the proton PDFs, Eq. (13): no preprocessing is required in $R_i^{\rm Pb}(x,Q_0^2)$, since to first approximation the leading physical behavior, such as the vanishing of the PDFs in the elastic limit $x\to 0$ or their power-like rise at small-x is already present in the baseline proton PDFs used. The architecture of the neural network in Eq. (14) is 2-5-3-1, with the two input neurons taking as input x and $\ln x$ respectively. All layers use a sigmoid activation function

except for the final layer which is based on a linear activation function, to avoid bounding the output of the neural network.

Once the fit determines the shape of $R_i^{\rm Pb}(x,Q_0^2)$, it is used to reconstruct the lead PDFs at the input evolution scale

$$q_i^{\rm Pb}(x,Q_0^2) \equiv R_i(x,Q_0^2)\cdot q_i^{\rm p}(x,Q_0^2)\,, \tag{15}$$

which is then evolved upwards in Q^2 using the NLO or NNLO DGLAP evolution equations using the `APFEL` program [31]. At scales other than Q_0 the nuclear modification factor will thus defined as

$$R_i(x,Q^2) = \frac{q_i^{\rm Pb}(x,Q^2)}{q_i^{\rm p}(x,Q^2)}\,, \tag{16}$$

where of course exactly the same theory settings such as solution of the DGLAP equations and the values of the heavy quark masses must be used consistently in the evolution of the lead and of the proton PDFs.

3.2 Flavor decomposition and theoretical settings

We will parametrize the nuclear modification factors below the charm threshold, so in principle we should parametrize seven R_i, three for each flavor of quark, three for the antiquarks, and the gluon. Heavy quarks are generated radiately using the DGLAP evolution equations. However, we are fitting only a single observable

$$\frac{F_2^{\rm Pb}(x,Q^2)}{F_2^{\rm d}(x,Q^2)} \tag{17}$$

so we cannot determine independently all R_i, but need to take some flavor assumptions. To see which PDF combinations can be probed by the data Eq. (17) in the fit, let us write the leading order expression at the input parametrization scale, and we have

$$\left.\frac{F_2^{\rm Pb}(x,Q_0^2)}{F_2^{\rm d}(x,Q_0^2)}\right|_{\rm LO} = \frac{\left((3Z+A)(u+\bar{u})^{\rm Pb}/A + (4A-3Z)(d+\bar{d})^{\rm Pb}/A + (s+\bar{s})^{\rm Pb}\right)/9}{\left(5(u+\bar{u}+d+\bar{d})^{\rm p} + 2(s+\bar{s})^{\rm p}\right)/18}\,, \tag{18}$$

where Z and A are the atomic and mass numbers of lead respectively, and where the deuteron structure function has been expressed in terms of the free proton PDFs neglecting nuclear effects for the $A=2$ case.

It is easy to see that Eq. (18) in the case $Z\simeq A/2$ simplifies substantially to

$$\left.\frac{F_2^{\rm Pb}(x,Q_0^2)}{F_2^{\rm d}(x,Q_0^2)}\right|_{\rm LO} = \frac{\left(5(u+\bar{u}+d+\bar{d})^{\rm Pb} + 2(s+\bar{s})^{\rm Pb}\right)/9}{\left(5(u+\bar{u}+d+\bar{d})^{\rm p} + 2(s+\bar{s})^{\rm p}\right)}\,, \tag{19}$$

and we see therefore that at leading order, fitting only fixed target neutral current DIS nuclear observables, we are only sensitive to a single quark combination

$$\Sigma^{\rm d}(x,Q_0) \equiv \left(5(u+\bar{u}+d+\bar{d})^{\rm p} + 2(s+\bar{s})^{\rm p}\right)(x,Q_0^2)\,, \tag{20}$$

which is the sum of quark PDFs, weighted by their electric charge, for NC DIS off an isoscalar target. In reality we are also sensitive to $T_3 \equiv (u+\bar{u}-d-\bar{d})$, the difference between the up-type and down-type quarks, since lead is not exactly an isoscalar nucleus, but as we will show below the constraints on T_3, and thus on flavor separation are really weak. At NLO and NNLO, we also have indirect constraints on the gluon nuclear modifications both from the higher order

corrections and from the DGLAP evolution effects. As expected, these constraints are weak in NC fixed-target DIS resulting in large uncertainties for the gluon nuclear modifications in lead.

Taking into account the above discussion, in this work we will parametrize using neural networks two nuclear modification factors

- The quark singlet $R_\sigma(x,Q_0) \equiv \Sigma^{\rm Pb}(x,Q_0)/\Sigma^{\rm p}(x,Q_0)$,
- The gluon $R_g(x,Q_0) \equiv g^{\rm Pb}(x,Q_0)/g^{\rm p}(x,Q_0)$,

and for cross-checks we will also attempt to extract the nuclear modification factor of the isotriplet PDF $R_{T_3}(x,Q_0) \equiv T_3^{\rm Pb}(x,Q_0)/T_3^{\rm p}(x,Q_0)$.

In this work, structure functions are computed using the FONLL general-mass variable-flavor-number (GM-VFN) scheme that includes the effect of the charm mass into the massless calculation. In particular we will use FONLL-B for the NLO fits and FONLL-C for the NNLO fits. PDF evolution and structure functions will be computed at NLO and at NNLO using the `APFEL` program. Unless otherwise specified, the theoretical settings are the same as in the NNPDF3.0 paper, in particular the fit uses values of the heavy quark (pole) masses of $m_c = 1.275$ GeV and $m_b = 4.18$ GeV. The maximum number of active quark flavors allowed in the structure functions is $n_f = 5$. We would like to emphasize that this is the first nuclear fit that has been performed up to NNLO accuracy, and also the first fit that includes heavy quark mass effects in the DIS coefficient functions. While these two effects are likely to be smaller than the typical nuclear PDF uncertainties, specially in the kinematic region covered by available data, it is important to derive for instance consistent predictions at NNLO for observables in pPb collisions.

3.3 Fitting strategy

The fitting strategy used in this works follows closely that used in the NNPDF3.0 proton PDF analysis [32]. We thus discuss here those ingredients that are different as compared to the NNPDF3.0 case. To begin with, the fraction of points used in the cross-validation algorithm has been reduced from 50% to 20%. Implementing cross-validation is particularly important in this case since we basically have a single observable, the ratio $F_2^{\rm Pb}/F_2^{\rm d}$, and thus the possibility of over-fitting should be carefully avoided. Using a 20% fraction in the validation set is enough to obtain a representative sample of the dataset without removing too much experimental information from the training set, given that in the nuclear case the dataset is much smaller as compared to the proton PDF fits.

The χ^2 minimized by the Genetic Algorithm is based on the t_0 method [9, 33], in order to avoid the D'Agostini bias in the presence of multiplicative correlated systematic uncertainties. While as discussed in Sect. 2, since the experimental covariance matrix is not available, we add statistical and systematic uncertainties in quadrature, we still have a correlated multiplicative systematic which comes from the PDF uncertainty in the conversion factor, which must thus be treated in the fit using the t_0 prescription.

4 Results

Now we turn to discuss the results of our nuclear PDF fits. First of all we show our main results, using EPS09 in the determination of the conversion factors, and then study the dependence on the input nuclear PDF set used by comparing with using DSSZ as input. We then to discuss the implications for the LHC heavy ion program, by comparing our determination of the lead PDFs with other recent nuclear PDF analysis available.

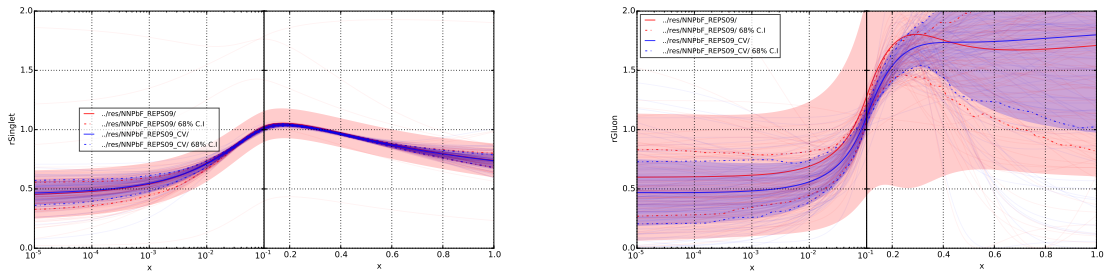

Figure 4: The nuclear modifications factors of lead for the quark singlet $R_\Sigma(x, Q_0)$ (left plot) and for the gluon $R_g(x, Q_0)$ (right plot) at the input parametrization scale of Q_0=1 GeV for the NLO fit The baseline proton PDF is NNPDF3.0 NLO and the EPS09 set has been used for the computation of the conversion factor. The solid band indicates the one-sigma PDF uncertainty in the nuclear modification factors.

4.1 NNPDFnucl1.0

In Fig. 4 we show the nuclear modifications factors of lead for the quark singlet $R_\Sigma(x, Q_0)$ (left plot) and for the gluon $R_g(x, Q_0)$ (right plot) at the input parametrization scale of Q_0=1 GeV for the NLO fit The baseline proton PDF is NNPDF3.0 NLO and the EPS09 set has been used for the computation of the conversion factor. The solid band indicates the one-sigma PDF uncertainty in the nuclear modification factors.

4.2 Dependence of the nuclear PDF set used for the conversion

One critical aspect of our approach is the use of a nuclear PDF set to perform the conversion of the nuclear DIS data into a common observable $F_2^{\rm Pb}/F_2^{\rm d}$. For our approach to be sound, the determination of the nuclear modification factors $R_i^{\rm Pb}$ should be relatively independent of the input PDF set used. In this section we compare the results of fits using EPS09 in the conversion with those of DSSZ, and show that a reasonable agreement is found. We also discuss how to combine into a single set the fits obtained using the two nuclear PDF sets as input.

In Fig. 5 we compare nuclear modification factors obtained using EPS09 as input with those obtained using DSSZ as input. The solid bands correspond to the one-sigma PDF uncertainty, while the dot-dashed lines indicate the 68% confidence level intervals. As can be seen, the difference between the two determinations are always smaller than the PDF uncertainties. As expected there are larger differences in the case of the gluon than for the singlet, though interestingly the qualitative behavior of the gluon modification factor is the same in the two cases.

Note that in the case of DSSZ the number of fitted data points is smaller than in the case of EPS09, since the typically ;larger values of the PDF uncertainties sin the former lead to the cut Eq. to be applied more restrictively. As expected, this leads to somewhat larger PDF uncertainties, specially at small-x. We also note that the nuclear modification factors are relatively Gaussian in the data region, while as expected the distribution becomes non-Gaussian in the extrapolation region at small-x, where no experimental constraints are available.

MC-PDFs.

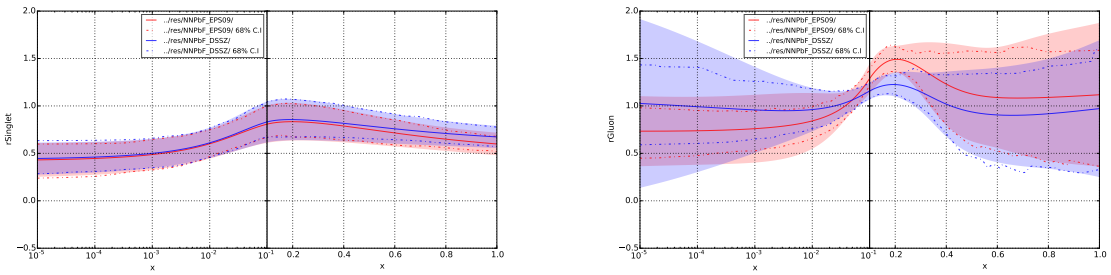

Figure 5: Same as Fig. 4, now comparing the nuclear modification factors obtained using EPS09 as input with those obtained using DSSZ as input. The solid bands correspond to the one-sigma PDF uncertainty, while the dot-dashed lines indicate the 68% confidence level intervals.

5 Summary and outlook

In this work we have presented the first determination of nuclear PDFs using the NNPDF methodology.

J. R. is supported by an STFC Rutherford Fellowship ST/K005227/1. N. H. and J. R. are supported by an European Research Council Starting Grant "PDF4BSM".

References

[1] K. J. Eskola, H. Paukkunen, and C. A. Salgado, *EPS09 - a New Generation of NLO and LO Nuclear Parton Distribution Functions*, arXiv:0902.4154.

[2] D. de Florian, R. Sassot, P. Zurita, and M. Stratmann, *Global Analysis of Nuclear Parton Distributions*, *Phys.Rev.* **D85** (2012) 074028, [arXiv:1112.6324].

[3] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione, *Neural network parametrization of deep-inelastic structure functions*, *JHEP* **05** (2002) 062, [hep-ph/0204232].

[4] **The NNPDF** Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, *Unbiased determination of the proton structure function f2(p) with estimation*, *JHEP* **03** (2005) 080, [hep-ph/0501067].

[5] **The NNPDF** Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, *Neural network determination of parton distributions: The nonsinglet case*, *JHEP* **03** (2007) 039, [hep-ph/0701127].

[6] **The NNPDF** Collaboration, R. D. Ball et al., *A determination of parton distributions with faithful uncertainty estimation*, *Nucl. Phys.* **B809** (2009) 1–63, [arXiv:0808.1231].

[7] **The NNPDF** Collaboration, J. Rojo et al., *Update on Neural Network Parton Distributions: NNPDF1.1*, arXiv:0811.2288.

[8] **The NNPDF** Collaboration, R. D. Ball et al., *Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering*, *Nucl. Phys.* **B823** (2009) 195–233, [arXiv:0906.1958].

[9] **The NNPDF** Collaboration, R. D. Ball et al., *Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties*, *JHEP* **05** (2010) 075, [arXiv:0912.2276].

[10] **The NNPDF** Collaboration, R. D. Ball et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, *Nucl. Phys.* **B838** (2010) 136–206, [arXiv:1002.4407].

[11] **The NNPDF** Collaboration, R. D. Ball et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, *Nucl. Phys.* **B849** (2011) 296–363, [arXiv:1101.1300].

[12] **The NNPDF** Collaboration, R. D. Ball et al., *Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO*, *Nucl.Phys.* **B855** (2012) 153–221, [arXiv:1107.2652].

[13] **NNPDF** Collaboration, R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, et al., *Parton distributions with LHC data*, *Nucl.Phys.* **B867** (2013) 244–289, [arXiv:1207.1303].

[14] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al., *Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data*, *Nucl.Phys.* **B855** (2012) 608–638, [arXiv:1108.1758].

[15] **The NNPDF** Collaboration, R. D. Ball et al., *Reweighting NNPDFs: the W lepton asymmetry*, *Nucl. Phys.* **B849** (2011) 112–143, [arXiv:1012.0836].

[16] N. Armesto, J. Rojo, C. A. Salgado, and P. Zurita, *Bayesian reweighting of nuclear PDFs and constraints from proton-lead collisions at the LHC*, *JHEP* **1311** (2013) 015, [arXiv:1309.5371].

[17] H. Paukkunen, K. J. Eskola, and C. Salgado, *Dijets in p + Pb collisions and their quantitative constraints for nuclear PDFs*, *Nucl.Phys.* **A931** (2014) 331–336, [arXiv:1408.4563].

[18] K. J. Eskola, H. Paukkunen, and C. A. Salgado, *A perturbative QCD study of dijets in p+Pb collisions at the LHC*, *JHEP* **1310** (2013) 213, [arXiv:1308.6733].

[19] I. Helenius, K. J. Eskola, and H. Paukkunen, *Probing the small-x nuclear gluon distributions with isolated photons at forward rapidities in p+Pb collisions at the LHC*, *JHEP* **1409** (2014) 138, [arXiv:1406.1689].

[20] D. d'Enterria, K. Krajczr, and H. Paukkunen, *Top-quark production in protonnucleus and nucleusnucleus collisions at LHC energies and beyond*, *Phys.Lett.* **B746** (2015) 64–72, [arXiv:1501.05879].

[21] **New Muon** Collaboration, P. Amaudruz et al., *A Reevaluation of the nuclear structure function ratios for D, He, Li-6, C and Ca*, *Nucl.Phys.* **B441** (1995) 3–11, [hep-ph/9503291].

[22] **New Muon** Collaboration, M. Arneodo et al., *The Structure Function ratios F2(li) / F2(D) and F2(C) / F2(D) at small x*, *Nucl.Phys.* **B441** (1995) 12–30, [hep-ex/9504002].

[23] **New Muon** Collaboration, M. Arneodo et al., *The A dependence of the nuclear structure function ratios*, *Nucl.Phys.* **B481** (1996) 3–22.

[24] **New Muon** Collaboration, M. Arneodo et al., *The Q**2 dependence of the structure function ratio F2 Sn / F2 C and the difference R Sn - R C in deep inelastic muon scattering*, *Nucl.Phys.* **B481** (1996) 23–39.

[25] J. Gomez et al., *Measurement of the a dependence of deep-inelastic electron scattering*, *Phys. Rev.* **D49** (1994) 4348.

[26] **European Muon** Collaboration, J. Ashman et al., *A Measurement of the ratio of the nucleon structure function in copper and deuterium*, *Z.Phys.* **C57** (1993) 211–218.

[27] **The NNPDF** Collaboration, R. D. Ball et al., *Theoretical issues in PDF determination and associated uncertainties*, *Phys.Lett.* **B723** (2013) 330–339, [arXiv:1303.1189].

[28] J. Owens, A. Accardi, and W. Melnitchouk, *Global parton distributions with nuclear and finite-Q^2 corrections*, arXiv:1212.1702.

[29] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, arXiv:1412.3989.

[30] J. Pumplin et al., *Uncertainties of predictions from parton distribution functions. 2. The Hessian method*, *Phys. Rev.* **D65** (2001) 014013, [hep-ph/0101032].

[31] V. Bertone, S. Carrazza, and J. Rojo, *APFEL: A PDF Evolution Library with QED corrections*, *Comput.Phys.Commun.* **185** (2014) 1647–1668, [arXiv:1310.1394].

[32] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **1504** (2015) 040, [arXiv:1410.8849].

[33] R. D. Ball, S. Carrazza, L. Del Debbio, S. Forte, J. Gao, et al., *Parton Distribution Benchmarking with LHC Data*, *JHEP* **1304** (2013) 125, [arXiv:1211.5142].