Analysis of nuclear fragmentation functions for pions with A and ν dependence

Mengyang Li,[1,*] Zijian Ye,[2,3,†] Jun Gao,[1,4,‡] XiaoMin Shen,[3,5,§] Hongxi Xing,[2,4,6,‖] and Yuxiang Zhao[3,4,5,7,¶]

[1]*State Key Laboratory of Dark Matter Physics, Shanghai Key Laboratory for Particle Physics and Cosmology, Key Laboratory for Particle Astrophysics and Cosmology (MOE), School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*
[2]*State Key Laboratory of Nuclear Physics and Technology, Institute of Quantum Matter, South China Normal University, Guangzhou 510006, China*
[3]*Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, Gansu 730000, China*
[4]*Southern Center for Nuclear-Science Theory (SCNT), Institute of Modern Physics, Chinese Academy of Sciences, Huizhou 516000, China*
[5]*University of Chinese Academy of Sciences, Beijing 100049, China*
[6]*Guangdong Basic Research Center of Excellence for Structure and Fundamental Interactions of Matter, Guangdong Provincial Key Laboratory of Nuclear Science, Guangzhou 510006, China*
[7]*Institute of Particle Physics, Central China Normal University, Wuhan 480079, China*

(Received 9 July 2026; accepted 31 August 2026; published 23 September 2026)

We present a QCD analysis of pion nuclear fragmentation functions (nFFs), which encode nuclear modifications to hadronization in high-energy nuclear collisions. Within this framework, vacuum fragmentation functions and their nuclear modifications are extracted simultaneously. The nuclear effects are parametrized as functions of the mass number A, the energy of the fragmenting parton in the target rest frame ν, and the hadron energy fraction z, allowing their dependence on these variables to be quantified. Our analysis includes semi-inclusive deep-inelastic scattering data on nuclear targets, with kinematic cuts chosen to ensure the applicability of perturbative QCD and collinear factorization. The resulting fit provides a good description of most datasets, with the nFFs well constrained in the energy fraction range $z \in [0.2, 0.7]$. Additionally, with our new nFFs, we present next-to-leading order predictions in pp and pA collisions, which show reasonable agreement with ALICE data within the current experimental uncertainties.

DOI: 10.1103/wdbl-ncxk

I. INTRODUCTION

Understanding hadronization dynamics is essential for exploring color confinement in QCD. Fragmentation functions (FFs) [1,2] provide the standard nonperturbative framework for describing hadron production in the final state, encoding the transition of partons from hard scattering processes into observed hadrons [3–5] in single-inclusive e^+e^- annihilation (SIA), semi-inclusive deep-inelastic scattering (SIDIS), and proton-proton (pp) collisions. Similar to parton distribution functions (PDFs) [6–9], these functions have been proven to be universal and cannot be computed perturbatively. A reliable way is to extract various FFs from world data through global QCD analysis.

While PDFs and FFs provide a robust framework for factorized cross-section calculations in the vacuum, their application to nuclear environments requires additional considerations. When hard scattering processes occur in nuclei rather than free nucleons, the initial-state parton distributions are modified and described by nuclear parton distribution functions (nPDFs) [10–18]. These nPDFs successfully describe inclusive observables such as nuclear deep-inelastic scattering (DIS) and Drell-Yan production. However, they alone cannot account for the observed hadron-species-dependent differences in production processes when comparing heavy nuclear targets with light nuclei or proton (p) targets [19–24].

Even though such differences can be attributed to a variety of conceivable mechanisms beyond initial-state modifications of parton densities [25], including, for

*Contact author: limengyang@sjtu.edu.cn
†Contact author: yezj@m.scnu.edu.cn
‡Contact author: jung49@sjtu.edu.cn
§Contact author: xiaominshen@impcas.ac.cn
‖Contact author: hxing@m.scnu.edu.cn
¶Contact author: yxzhao@impcas.ac.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

instance, parton energy loss caused by medium-induced multiple gluon emission [26–29], it has been proposed that QCD factorization can be extended to incorporate final-state nuclear effects through the introduction of nuclear fragmentation functions (nFFs) with an explicit dependence on nuclear mass number A [30]. The idea was further explored in Ref. [31], and has been extended to the three-dimensional imaging of PDFs and FFs in nuclei [32–34].

Although such approaches based on A-dependent nFFs provide a successful description of nuclear modifications across a broad range of data, they do not explicitly account for additional kinematic variables that may influence the hadronization process in nuclear environments. In particular, recent high-precision measurements from HERMES [35,36] and CLAS [37] suggest that the strength of nuclear modifications exhibits a nontrivial dependence on the virtual-photon energy ν in the target rest frame, indicating that the dynamics of in-medium hadronization cannot be fully captured by A dependence. The role of ν in nuclear hadronization has also been investigated within a TMD-based model [38], in which the ν dependence is incorporated through the treatment of medium-induced parton energy loss and prehadron absorption. These experimental and phenomenological observations motivate the present study. In this work, we introduce a parametrization of nuclear modifications that depends on A and ν, and simultaneously extract the vacuum FFs and their nuclear corrections. The implementation of the Hessian method in our study further improves the reliability of the extraction of nFFs by constructing uncertainty eigenvector sets for error quantification. In addition, updated nPDF sets, including nCTEQ15WZ [39], nNNPDF3.0 [40], and EPPS21 [41], are used to perform systematic comparisons and assess the associated theoretical uncertainties. Heavy-quark-tagged SIA data are also included to improve constraints on heavy-flavor fragmentation.

The paper is organized as follows. Section II describes the experimental datasets used in this analysis, including SIDIS measurements from HERMES and CLAS and heavy-flavor-tagged SIA data from SLD. Section III presents the theoretical framework for vacuum FFs and nFFs, including details of the theoretical calculations and the definition of the goodness-of-fit function. The extraction of vacuum FFs and the corresponding predictions compared with data are presented in Sec. IV, which constrain the theoretical conditions for the subsequent determination of nFFs and define the corresponding default setup. Based on this, the extracted nFFs and their corresponding comparisons with data are discussed in Sec. V. Predictions for π production in pp and proton-lead (pPb) collisions at ALICE are presented in Sec. VI. Finally, Sec. VII summarizes our main findings.

II. EXPERIMENTAL DATASETS FITTED

We begin by introducing the key observables in nuclear SIDIS, i.e., $\ell + A \to \ell' + h + X$. The four-momenta of the target nucleon, exchanged virtual photon, incoming lepton, and produced hadron are denoted by P, q, ℓ, and P_h, respectively. The differential multiplicity of identified hadrons is defined as

$$\frac{\mathrm{d}M^h(x,z,Q^2)}{\mathrm{d}z} = \frac{\mathrm{d}^3\sigma^h(x,z,Q^2)/\mathrm{d}x\,\mathrm{d}Q^2\mathrm{d}z}{\mathrm{d}^2\sigma^{\mathrm{DIS}}(x,Q^2)/\mathrm{d}x\,\mathrm{d}Q^2}, \tag{1}$$

where $x = Q^2/(2P\cdot q)$ is the Bjorken variable, $z = (P\cdot P_h)/(P\cdot q)$ is the hadron energy fraction, and $Q^2 = -q^2$ denotes the virtuality of the exchanged photon. In this process, we expect the nuclear modification from initial-state nPDFs to be largely canceled, therefore providing a golden channel for the global extraction of nFFs.

To quantify nuclear modifications, the multiplicity ratio R_A^h is introduced as the ratio of the differential hadron multiplicity for a nuclear target with mass number A to that for a deuterium (D) target. Since the virtual-photon energy ν in the target rest frame provides an additional kinematic variable in nuclear SIDIS, we keep it explicitly in the nuclear multiplicity ratio in order to test possible ν-dependent effects. The multiplicity ratio is then defined as

$$R_A^h(x,z,Q^2,\nu) = \frac{\mathrm{d}M_A^h(x,z,Q^2,\nu)/\mathrm{d}z}{\mathrm{d}M_D^h(x,z,Q^2,\nu)/\mathrm{d}z}, \tag{2}$$

where $\mathrm{d}M_A^h/\mathrm{d}z$ and $\mathrm{d}M_D^h/\mathrm{d}z$ denote the corresponding differential multiplicities for the nuclear target with mass number A and the D target, respectively.

In Table I, we present a comprehensive compilation of the experimental data used in this analysis, including SIDIS measurements from HERMES [35,36,42,43] and CLAS [37] and heavy-flavor-tagged SIA measurements from SLD [44]. The HERMES data span multiple years and include measurements on p, helium (He), neon (Ne), krypton (Kr), xenon (Xe), and D targets, with final-state hadrons π^+, π^-, and π^0. These data are presented in various kinematic bins of z, Q^2, and ν, and are reported either as the double ratio R_A^h or as the normalized differential multiplicity $dN_h/(dzdQ^2)$. The data cover a wide range of nuclear targets, providing valuable constraints on the A dependence of nFFs. It should be noted that the HERMES data reported in 2011 [36] are not statistically independent of earlier HERMES measurements [35]. The 2011 HERMES results present nuclear multiplicity ratios in two-dimensional kinematic bins, while the earlier results were mostly shown as functions of a single kinematic variable. Therefore, including the two HERMES measurements in the fit simultaneously can introduce a certain degree of double counting. Nevertheless, in order to make

TABLE I. Summary of experimental datasets included in the analysis, together with the corresponding correlated uncertainties.

Number	Experiment	Year	$\sqrt{s}$ (GeV)	Target (A)	Particle	Binning	Obs	Normalization uncertainty (%)	N_{pt}
1	HERMES [43]	2001	27.5	p	π^0	z	$\frac{1}{N_{DIS}}\frac{dN^{\pi^0}}{dz}$	9	10
2	HERMES [35]	2007	27.6	He	$\pi^{\pm}$	z, ν	R_A^h	3	16
3				He	$\pi^{\pm}$	z, Q^2	R_A^h	3	20
4				He, Ne, Kr, Xe	π^+	z	R_A^h	3	24
5				He, Ne, Kr, Xe	π^-	z	R_A^h	3	24
6				He, Ne, Kr, Xe	π^0	z	R_A^h	3	24
7	HERMES [36]	2011	27.6	Ne, Kr, Xe	π^+	z, Q^2	R_A^h	3	60
8				Ne, Kr, Xe	π^-	z, Q^2	R_A^h	3	60
9				Ne, Kr, Xe	π^+	z, ν	R_A^h	3	51
10				Ne, Kr, Xe	π^-	z, ν	R_A^h	3	51
11	HERMES [42]	2013	27.6	p	π^+	z, Q^2	$\frac{1}{N_{DIS}}\frac{dN^{\pi^+}}{dzdQ^2}$	0	25
12				p	π^-	z, Q^2	$\frac{1}{N_{DIS}}\frac{dN^{\pi^-}}{dzdQ^2}$	0	25
13				D	π^+	z, Q^2	$\frac{1}{N_{DIS}}\frac{dN^{\pi^+}}{dzdQ^2}$	0	25
14				D	π^-	z, Q^2	$\frac{1}{N_{DIS}}\frac{dN^{\pi^-}}{dzdQ^2}$	0	25
15	CLAS [37]	2021	5.0	C, Fe, Pb	π^+	z, Q^2, ν	R_A^h	3	117
16				C, Fe, Pb	π^-	z, Q^2, ν	R_A^h	3	114
								SIDIS total:	671
17	SLD [44]	2003	91.2	SIA c-tagged	$\pi^{\pm}$	z	$\frac{1}{N_{evt}}\frac{dN^{\pi^{\pm}}}{dz}$	1	21
18				SIA b-tagged	$\pi^{\pm}$	z	$\frac{1}{N_{evt}}\frac{dN^{\pi^{\pm}}}{dz}$	1	21
								SIA total:	42
								Total:	713

maximal use of the available experimental information, we include both measurements in the present analysis. We have checked that the extracted nFFs using alternative selections of the potentially overlapping HERMES data are similar to those obtained in our default fit. Their uncertainty bands are also nearly unchanged in the data-constrained region, indicating that the impact of potential double counting is negligible. We note that several statistically correlated HERMES datasets have also been fitted simultaneously in global QCD analyses [31]. More recently, the CLAS Collaboration has published high-precision measurements on carbon (C), iron (Fe), and lead (Pb) targets, covering a broad range of z, Q^2, and ν [37]. These data extend the nuclear coverage to heavier nuclei and are reported as R_A^h in multiple ν bins, enabling a detailed study of nuclear medium effects across different energy regimes. The combined dataset, as summarized in Table I, offers a robust and diverse set of observables for constraining nFFs. To further constrain the vacuum FFs, we also include SLD SIA data [44] on charm- and bottom-tagged events, which provide direct constraints on the heavy flavor FFs.

We apply kinematic cuts to ensure the reliability of the QCD collinear factorization framework. For both SIA and SIDIS data sets, we require $z > 0.1$. To further suppress regions where power corrections may become important, we impose an additional requirement on the hadron energy E_h, which is evaluated in the center-of-mass frame for SIA and in the Breit frame for SIDIS, by requiring the corresponding hadron energy scale, $zQ/2$, to satisfy $zQ/2 > 0.19$ GeV. This cutoff was selected from a stability test of the vacuum-FF fit, since lowering it to 0.10 GeV significantly worsens the fit quality. The value was set slightly below 0.20 GeV to retain data points at this boundary.

III. FRAMEWORK FOR MEDIUM-MODIFIED FRAGMENTATION FUNCTIONS

A. QCD factorization and vacuum FFs

In the collinear factorization framework, the cross sections can be factorized into perturbatively calculable short-distance partonic cross sections, and nonperturbative distribution functions [3]. For example, the SIDIS differential cross section can be written as

$$\frac{d^3\sigma^{\ell p\to \ell hX}}{\mathrm{d}x\mathrm{d}y\mathrm{d}z} = \frac{2\pi\alpha_{\mathrm{em}}^2}{Q^2}\left[\frac{1+(1-y)^2}{y}F_T^h(x,z,Q^2) + \frac{2(1-y)}{y}F_L^h(x,z,Q^2)\right], \tag{3}$$

where F_T^h and F_L^h are the semi-inclusive structure functions, and $y = (P \cdot q)/(P \cdot \ell)$ is the inelasticity.

At next-to-leading order (NLO), the SIDIS structure functions can be expressed as,

$$\begin{aligned} F_T^h(x,z,Q^2) = \sum_q e_q^2 \Big[& f_1^{q/p}(x,Q^2) D_q^h(z,Q^2) \\ & + \frac{\alpha_s(Q^2)}{2\pi} (f_1^{q/p} \otimes C_1^{qq} \otimes D_q^h \\ & + f_1^{q/p} \otimes C_1^{gq} \otimes D_g^h \\ & + f_1^{g/p} \otimes C_1^{qg} \otimes D_q^h) \Big], \end{aligned} \tag{4}$$

$$\begin{aligned} F_L^h(x,z,Q^2) = & \frac{\alpha_s(Q^2)}{2\pi} \sum_q e_q^2 [f_1^{q/p} \otimes C_L^{qq} \otimes D_q^h \\ & + f_1^{q/p} \otimes C_L^{gq} \otimes D_g^h \\ & + f_1^{g/p} \otimes C_L^{qg} \otimes D_q^h], \end{aligned} \tag{5}$$

where $f_1^{i/p}$ denotes the PDFs, while FFs (D_i^h) describe the nonperturbative hadronization of a final-state parton of flavor i into an observed hadron h. The functions $C_{1(L)}^{ij}$ are the NLO perturbatively calculable SIDIS coefficient functions [45–49], and the next-to-next-to-leading order (NNLO) coefficient functions can be found in [50–52]. Approximate NNLO and N^{3}LO (next-to-next-to-next-to-leading order) structure functions have also been obtained through expansions of threshold resummation expressions [53–55]. More recently, complete NNLO and N^{3}LO calculations for the SIDIS process have become available [56,57]. The symbol $\otimes$ denotes the standard convolution integral defined as

$$[f \otimes g](z) = \int_0^1 dx \int_0^1 dy f(x) g(y) \delta(z - xy). \tag{6}$$

Focusing on the FFs, D_i^h are universal within the framework of QCD factorization in the vacuum and cannot be computed perturbatively, and can be extracted from world data through global QCD analysis, similar to PDFs [6–9]. Such analyses include hadron production in SIA, SIDIS, and pp collisions, with energy scale satisfying $Q^2 \gg \Lambda_{\rm QCD}^2$. The evolution of FFs D_i^h with respect to the scale Q^2 is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation [58–61],

$$\frac{\partial}{\partial \ln Q^2} D_i^h(z,Q^2) = \sum_j [P_{ji} \otimes D_j^h](z,Q^2), \tag{7}$$

where P_{ji} are the timelike splitting functions, describing the probability for a parton i to split into a parton j. These timelike splitting functions are currently known up to $\mathcal{O}(\alpha_s^3)$ [62–67], providing the theoretical basis for high-precision global analyses of FFs. It should be noted that, although timelike and spacelike splitting functions (the latter relevant for PDFs) are identical at leading order, they differ beyond NLO [68,69].

B. Nuclear FFs and parametrization

Assuming that the QCD collinear factorization for inclusive production is valid also for collisions involving nuclei, one can naturally define fragmentation functions $D_{i/A}^h(z,Q^2)$ for nuclear environment characterized by mass number A [30,31]. In these factorization approaches, $D_{i/A}^h(z,Q^2)$ parametrize nuclear modifications solely through their dependence on A as,

$$D_{i/A}^h(z,Q_0^2) = \int_z^1 \frac{dy}{y} W_i^h(y,A,Q_0^2) D_i^h\left(\frac{z}{y}, Q_0^2\right), \tag{8}$$

where the subscript i labels different parton flavors, the superscript h labels different hadron species, and $W_i^h(y,A,Q_0^2)$ encodes the nuclear modification dependence derived from fixed-target hadron-nucleus data [10,11], and has a complex parametrization. A similar reasoning is applied with great phenomenological success in analyses of nPDFs that account for medium-induced effects in the initial state [10–12].

However, such models may not fully capture the possible dependence of nuclear modifications on the virtual-photon energy ν in the target rest frame, which is directly accessible in eA measurements. Recent precision measurements [36,37] provide data over a broad range of ν, offering an opportunity to test whether additional energy dependence is favored by the data. Such a dependence may be connected to the formation time of hadrons. At larger ν, the struck parton carries more energy, which may modify the space-time development of hadronization and thereby change the strength of final-state interactions in the nuclear medium.

Motivated by this possibility, we introduce a phenomenological dependence on both A and ν into the nFF framework, and examine whether the available SIDIS data prefer such an extension over an A-only description. Within the collinear factorization framework for nuclear systems, we propose a model that decomposes the nFFs $D_{i/A}^h(z,Q^2,\nu)$ into vacuum FFs $D_i^h(z,Q^2)$ and a modified term $D_i'^h(z,Q^2)$,

$$D_{i/A}^h(z,Q^2,\nu) = D_i^h(z,Q^2) - \mathcal{F}_A(A,\nu) \times D_i'^h(z,Q^2), \tag{9}$$

where the $D_i^h(z,Q^2)$, $D_i'^h(z,Q^2)$, and $\mathcal{F}_A$ are constrained simultaneously in this work. In this parametrization, $D_i'^h(z,Q^2)$ can be either positive or negative in principle, allowing the model to describe both suppression and enhancement effects. The minus sign convention is chosen

because our analysis focuses on the medium to large z region where experimental data predominantly show suppression effects. Consequently, the fitted $D_i'^h(z,Q^2)$ is expected to be positive.

To parametrize the $D^h_{i/A}$ at the initial scale Q_0, we adopt the following functional forms:

$$zD_i^h(z,Q_0) = z^{\alpha_i^h}(1-z)^{\beta_i^h}\exp\left(\sum_{n=0}^{k} a_{i,n}^h z^{n/2}\right),$$
$$zD_i'^h(z,Q_0) = z^{\alpha_i'^h}(1-z)^{\beta_i'^h}\exp\left(\sum_{n=0}^{k} a_{i,n}'^h z^{n/2}\right), \quad (10)$$

where α, β, a are the fitted parameters. By varying $k = 0, 1, 2\cdots$, one can freely choose different parametrizations. In practice, we increase the value of k until no discernible improvement in fit quality can be obtained. In this framework, we assume that the factorization of the cross section used in vacuum in Eqs. (4) and (5) remains applicable to nuclear targets, with the vacuum FFs replaced by the nFFs. Accordingly, both $D_i^h(z,Q^2)$ and $D_i'^h(z,Q^2)$ should obey the DGLAP evolution equations.

The parametrization for π is summarized in Table II. By assuming charge-conjugation symmetry, nFFs for π^- can be obtained from π^+. To reduce the number of free parameters, we impose flavor symmetries between favored and unfavored light (anti)quark FFs, as indicated by the equality signs in the table, where favored u and $\bar{d}$ quarks share the same values of α, β, a_0, and a_1 at the initial scale. Identical distributions are also imposed among the unfavored (anti)quarks $\bar{u}$, d, s, and $\bar{s}$. Nuclear modifications are assumed to apply only to light quarks, with similar flavor symmetries adopted in light quark contributions.

Finally, $\mathcal{F}_A(A,\nu)$ encodes the nuclear modifications, incorporating both A and ν dependence, via,

$$\mathcal{F}_A(A,\nu) = \frac{f(A)}{(\nu/\nu_0)^\delta}. \quad (11)$$

TABLE II. Summary of the parameterization for π^+ vacuum FFs and nuclear modification terms. Check marks under each parameter mean that the parameter is free to vary.

	Flavor	a_0	α	β	a_1
$D_i^h(z,Q^2)$	$u=\bar{d}$	✓	✓	✓	✓
	$d=\bar{u}$	✓	✓	✓	✓
	$s=\bar{s}$	$=a_{0,d}$	$=\alpha_d$	$=\beta_d$	$=a_{1,d}$
	$c=\bar{c}$	✓	✓	✓	not applicable
	$b=\bar{b}$	✓	✓	✓	not applicable
	g	✓	✓	✓	not applicable
$D_i'^h(z,Q^2)$	$u=\bar{d}$	✓	✓	✓	✓
	$d=\bar{u}$	✓	✓	✓	✓
	$s=\bar{s}$	$=a'_{0,d}$	$=\alpha'_d$	$=\beta'_d$	$=a'_{1,d}$

Here, $\nu_0 = 1$ GeV is introduced as a reference energy scale to ensure that $\mathcal{F}_A(A,\nu)$ is dimensionless. The function $f(A)$ parametrizes the nuclear dependence, with an independent parameter assigned to each nucleus, and will be discussed in Sec. V B. The exponent δ controls the strength of the ν dependence. In this work, we consider three representative choices, $\delta = 0$, 0.5, and 1. The case $\delta = 0$ corresponds to an A-dependent but ν-independent modification, and is included as a reference scenario. For $\delta > 0$, the modification factor $\mathcal{F}_A(A,\nu)$ decreases with increasing ν and vanishes in the limit $\nu \to \infty$, so that $D^h_{i/A}$ approaches the vacuum FF D_i^h.

C. Theoretical computations

The theoretical computation techniques have been summarized in the previous NPC23 studies [70–72]. Here we briefly recall the main ingredients. In this work, the main analysis is carried out at NLO accuracy, while NNLO results are used in Sec. IV as a benchmark to assess higher-order effects and perturbative stability. For the NLO analysis, the FFs are evolved using the two-loop timelike splitting kernels, which were calculated in Ref. [73]. For the NNLO comparison, the evolution is performed with the three-loop timelike splitting kernels [62–65]. The corresponding NLO and NNLO DGLAP evolutions are carried out with HOPPET [74,75]. Theoretical calculations of differential cross sections are carried out up to NNLO in QCD using the FMNLO program [70,76], which can generate and store interpolation tables of the coefficient functions, ensuring fast convolution with arbitrary FFs without repeating the calculations. Furthermore, in this work, the dependence on ν is taken into account, and the corresponding interpolation grid construction is extended accordingly to incorporate this dependence, ensuring consistency in the nFFs analysis.

We adopt a zero-mass variable flavor number scheme, in which heavy quarks FFs are nonzero but do not evolve until the mass thresholds are reached with maximum of $n_f = 5$, specifically at $m_c = 1.4$ GeV and $m_b = 4.5$ GeV for charm and bottom quarks. The strong coupling constant is consistently taken as $\alpha_S(M_Z) = 0.118$ throughout the calculations. For theoretical predictions of hadron production in SIA with heavy-flavor-tagged events, there are ambiguities on matching theoretical predictions to the experimental measurements beyond NLO, for example, in the treatment of contributions from gluon splitting into heavy quarks. In our calculations, we only include contributions from Feynman diagrams in which the specified heavy quark couples directly to the Z boson or photon, both at NLO and NNLO.

The central values for the renormalization and fragmentation scales ($\mu_{R,0}$ and $\mu_{D,0}$) are set to the momentum transfer Q for both SIA and SIDIS. The factorization scale ($\mu_{F,0}$) of initial hadrons for SIDIS is also set to Q.

D. Goodness of fit function and the covariance matrix

The agreement between the data points D_k and the corresponding theoretical predictions T_k is quantified by the χ^2 function [77],

$$\chi^2(\{a\},\{\lambda\}) = \sum_{k=1}^{N_{\rm pt}} \frac{1}{s_k^2}\left(D_k - T_k - \sum_{\alpha=1}^{N_\lambda}\beta_{k,\alpha}\lambda_\alpha\right)^2 + \sum_{\alpha=1}^{N_\lambda}\lambda_\alpha^2, \tag{12}$$

where $\{a\}$ are parameters of FFs, the nuisance parameters $\{\lambda\}$ describe sources of correlated errors, which are assumed to follow standard normal distributions, and s_k represents the total uncorrelated systematic and statistical errors. $\beta_{k,\alpha}$ quantifies the sensitivity of the kth measurement to the αth correlated error source, while λ_α is the corresponding dimensionless nuisance parameter. In our case, the correlated errors can be either the normalization uncertainties from the measurements, listed in the penultimate column of Table I, or the theoretical uncertainties estimated with scale variations, following Ref. [76]. Here, N_λ denotes the total number of independent correlated uncertainty sources included in the fit.

The best-fit fragmentation parameters are determined by minimizing the χ^2 and then further validated through a series of profile scans on each of those parameters. These parameter space scans are conducted using the MINUIT program [78]. We apply a tolerance criterion of $\Delta\chi^2 \sim \chi^2/N_{\rm pt_{max}}$ to determine parameter uncertainties. Additionally, we employ the iterative Hessian approach [79] to generate error sets of FFs, which can be used to propagate parameter uncertainties to physical observables.

IV. ANALYSIS OF VACUUM FRAGMENTATION FUNCTIONS

This section presents an extraction of vacuum FFs, which provides the default setup for the subsequent determination of nFFs. The extraction is performed through a comprehensive study of HERMES data and heavy-flavor-tagged measurements from SLD, as summarized in Table III. The HERMES experiment provides SIDIS measurements of hadron production with high precision for charged and neutral π over a broad kinematic range, allowing for detailed constraints on vacuum FFs. Specifically, possible nuclear modifications in the D target data are assumed to be small throughout this analysis, allowing these data to be used as constraints on the vacuum FFs. Complementary constraints on heavy flavor FFs are provided by SLD data, where heavy-flavor-tagged measurements offer direct sensitivity to c and b quark FFs.

The robustness of our results is systematically evaluated by varying several theoretical setups, and the comparative results are also summarized in Table III, which provides the $\chi^2/N_{\rm pt}$ values for each dataset and for the total fit under different theoretical setups. We begin by comparing perturbative calculations at NLO and NNLO in QCD to evaluate the impact of higher-order corrections and the stability of the results. The NNLO fits lead to larger χ^2 values than the corresponding NLO fits for the present dataset. We note that, a bin-by-bin inspection suggests that this difference is mainly driven by the low-Q^2 HERMES SIDIS data. In this region, the NNLO predictions have difficulty reproducing the observed Q^2 dependence, while the corresponding scale variation uncertainties are not reduced compared to NLO. We therefore use the NLO setup as the default choice in the following analysis. The dependence on the initial scale is then examined through comparisons between fits performed with $Q_0 = 1.0$ GeV

TABLE III. Summary of $\chi^2/N_{\rm pt}$ values for different theoretical setups in the vacuum FFs fits.

					With theoretical uncertainties (χ^2/N_{pt})				Without theoretical uncertainties (χ^2/N_{pt})			
					$Q_0 = 1.0$ GeV		$Q_0 = 1.3$ GeV		$Q_0 = 1.0$ GeV		$Q_0 = 1.3$ GeV	
Experiment	Target (A)	Particle	Binning	N_{pt}	NLO	NNLO	NLO	NNLO	NLO	NNLO	NLO	NNLO
HERMES [43]	p	π^0	z	10	0.56	1.48	0.50	1.59	2.51	3.18	2.93	3.92
HERMES [42]	p	π^+	z, Q^2	25	0.88	1.48	0.64	1.33	1.16	3.73	0.87	2.94
	p	π^-	z, Q^2	25	0.74	0.99	0.48	0.83	0.91	4.76	0.55	3.89
	D	π^+	z, Q^2	25	1.13	2.33	0.87	2.19	1.29	3.60	1.04	3.05
	D	π^-	z, Q^2	25	1.30	1.84	1.25	1.82	1.36	4.11	1.25	3.52
			SIDIS total:	110	0.97	1.64	0.78	1.55	1.30	3.97	1.11	3.40
SLD [44]	SIA c-tagged	$\pi^\pm$	z	21	0.77	0.61	0.84	0.61	0.73	3.00	0.81	2.55
	SIA b-tagged	$\pi^\pm$	z	21	0.84	0.78	0.92	0.81	0.81	0.89	0.91	0.90
			SIA total:	42	0.80	0.69	0.88	0.71	0.77	1.95	0.86	1.73
			Total:	152	0.92	1.38	0.81	1.31	1.15	3.41	1.04	2.94

and $Q_0 = 1.3$ GeV. The latter choice generally leads to an improved description of the data, in particular for the SIDIS measurements. This can be attributed to the positive-definite constraint on the input distributions, because at the lower initial scale DGLAP evolution may favor slightly negative values in certain z regions to improve the fit, which is excluded by construction, whereas the higher Q_0 allows a better overall χ^2 within the positive-definite framework. Furthermore, the role of theoretical uncertainties introduced through the covariance matrix (see Sec. III D) is investigated by comparing fits with and without their inclusion. While including theoretical uncertainties leads to reduced χ^2 values, the reduction does not correspond to a significant improvement in the fit quality. Given the current experimental precision and the already satisfactory fit quality, the inclusion of such uncertainties is therefore not required in the present analysis. Based on these comparisons, we identify a default configuration characterized by $Q_0 = 1.3$ GeV, NLO perturbative calculations, and the exclusion of theoretical uncertainties in the fitting procedure. This setup provides a stable and balanced description of the data across different datasets. For this default choice, the total χ^2 amounts to 158.2 for a total of 152 data points, corresponding to $\chi^2/N_{\rm pt} = 1.04$, indicating good overall agreement between theory and experiment. The corresponding $\chi^2/N_{\rm pt}$ values are 1.11 and 0.86 for the SIDIS and SIA datasets, respectively.

We now turn to the details of the vacuum FFs extraction. Section IV A focuses on comparisons between theoretical predictions and experimental data for SIDIS and SIA observables, while Sec. IV B presents the extracted vacuum FFs and investigates their dependence on different theoretical configurations.

A. Comparison with experimental data

Figure 1 compares the HERMES measurements [42] with theoretical predictions for $\pi^{\pm}$ production in SIDIS off p and D targets under several theoretical setups considered in this work. To facilitate a direct comparison among the different theoretical predictions and the experimental data, both the experimental data and the theoretical predictions are divided by the default prediction in each kinematic bin. The experimental uncertainties are small, remaining at the

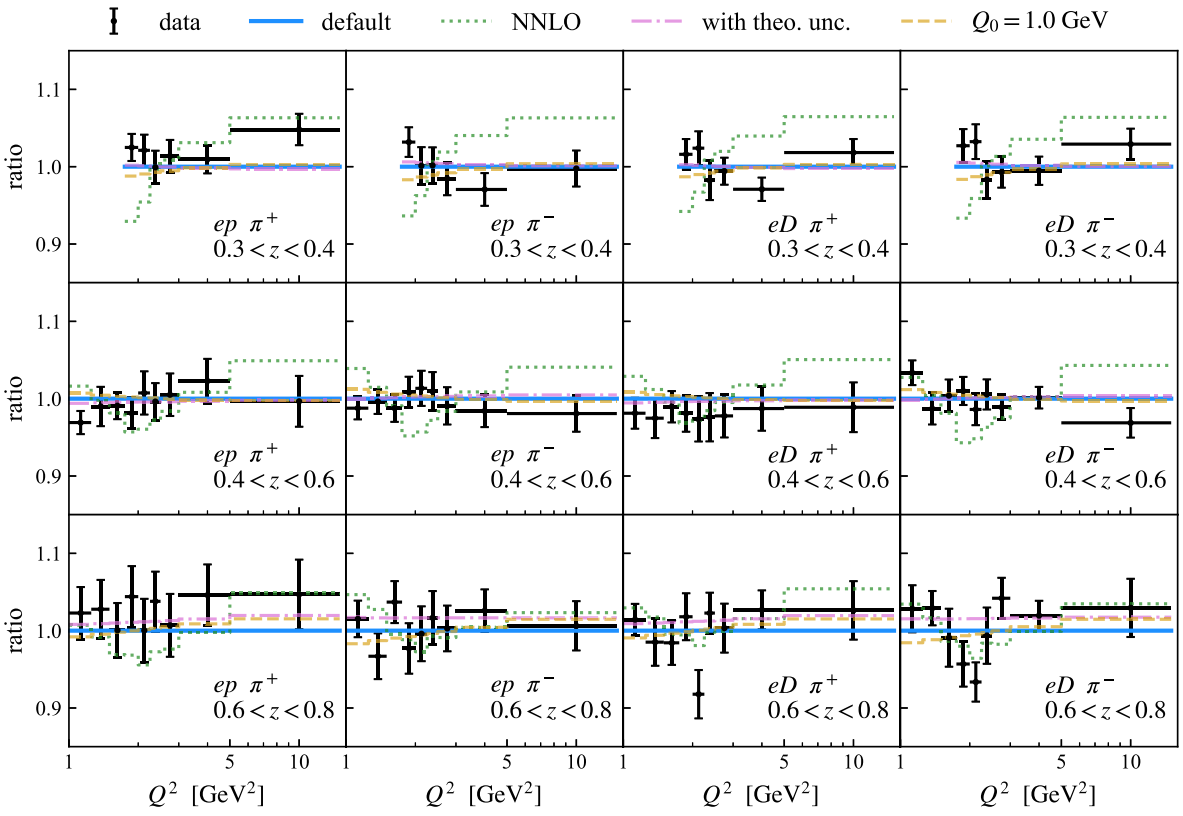


FIG. 1. Comparison of HERMES measurements [42] and theoretical predictions for π^+ and π^- production in SIDIS off p and D targets, shown in bins of z as functions of Q^2. All results are presented as ratios to the default prediction, defined as the setup with $Q_0 = 1.3$ GeV, NLO, and the exclusion of theoretical uncertainties. The black points with error bars represent the experimental data divided by the default prediction. The solid blue line corresponds to the default prediction and is therefore equal to unity by construction. The dotted green line shows the NNLO prediction, the dash-dotted pink line shows the result obtained when theoretical uncertainties are included in the fit, and the dashed orange line corresponds to the initial scale choice $Q_0 = 1.0$ GeV.

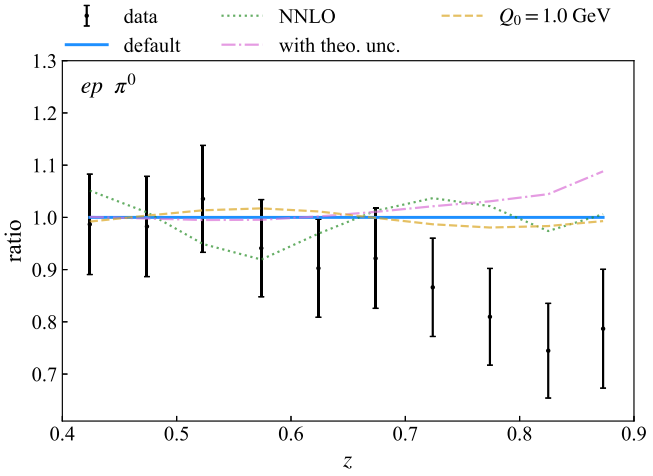


FIG. 2. Similar to Fig. 1, but for π^0 production in HERMES experiments [43], shown as a function of z.

level of $2\% - 4\%$ for $z < 0.6$ and increasing moderately to below 5% in the range $0.6 < z < 0.8$. This high precision provides stringent constraints on the extracted vacuum FFs. The default theoretical prediction describes the p data well across the full kinematic range. A mild deviation is observed in the π^+ channel for $0.3 < z < 0.4$, where the data lie approximately 5% above the central prediction. Although this deviation exceeds the experimental uncertainties, its impact on the fit quality remains limited because only a small number of data points are affected. For the D target, the data show some mild fluctuations relative to the default prediction. Since these deviations remain compatible with the experimental uncertainties, the default prediction still provides a consistent description of the data. The combined fit quality for p and D datasets demonstrates the stability of the extraction and supports the reliability of the resulting vacuum FFs. The impact of the theoretical setup is then assessed by comparing the different predictions. In the low and intermediate z region ($z < 0.6$), the default setup, the fit including theoretical uncertainties, and the initial scale choice $Q_0 = 1.0$ GeV give nearly identical predictions. This indicates that the extracted vacuum FFs are stable with respect to these variations. A different pattern becomes visible at larger z ($0.6 < z < 0.8$), where the $Q_0 = 1.0$ GeV setup lies slightly below the default at low Q^2 and slightly above it at higher Q^2, while the setup including theoretical uncertainties shows a mild upward shift over the full Q^2 range. The NNLO calculation shows qualitatively different behavior across all z regions. In the region $0.3 < z < 0.4$, the NNLO prediction evolves from a suppression at low Q^2 (approximately 5% below the default) to an enhancement at high Q^2.

For π^0 production, as shown in Fig. 2, the agreement between theory and data remains good, with a mild overestimation observed in the high-z region ($z > 0.7$). This feature persists across all theoretical configurations considered, indicating that it is not driven by perturbative order or initial scale choices. Such deviations may reflect limitations of the current theoretical framework in describing π^0 production at high z, potentially requiring modifications of FFs near $z \to 1$. Nevertheless, given the relatively large uncertainties and the limited statistical precision of the π^0 data, this deviation does not significantly impact the total fit and remains consistent with the overall description of the data.

Figure 3 presents the comparison between theory and data for charm tagged and bottom tagged $\pi^{\pm}$ production in SIA at the Z pole, based on measurements from the SLD Collaboration. These data provide clean probes of heavy quark fragmentation due to their high flavor purity and well-controlled kinematics. For clarity of presentation, the error bars of the last three data points in charm tagged and the last two data points in bottom tagged are not displayed in the figure, as the experiment uncertainties become large after normalization. The default theoretical prediction describes the data well across most of the z range. Although

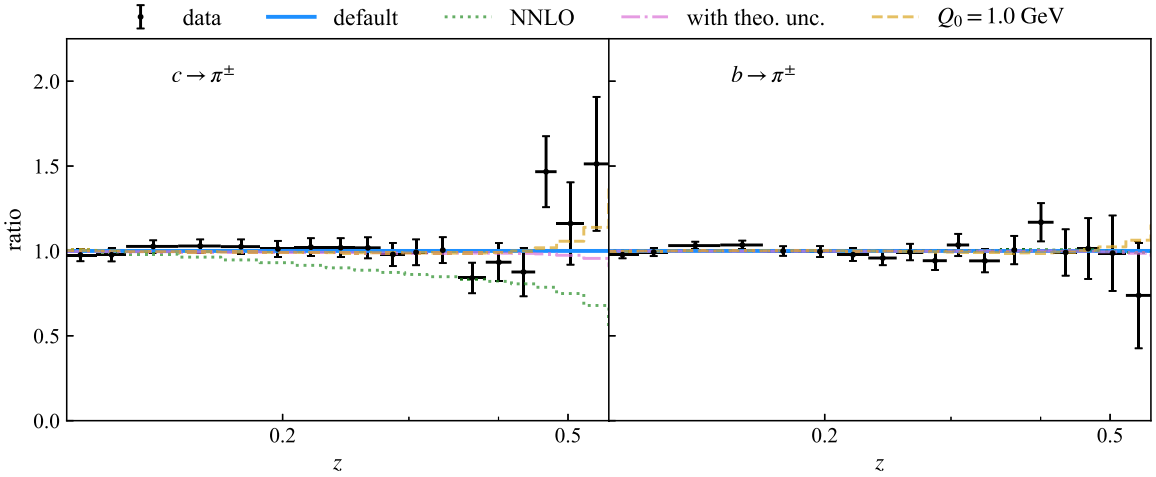


FIG. 3. Similar to Fig. 1, but for $\pi^{\pm}$ production in SIA at SLD [44], shown as a function of z.

mild deviations are observed at large z ($z \gtrsim 0.5$), where the last data points slightly exceed the experimental uncertainty, this behavior remains compatible with the data within the corresponding large experimental uncertainties. The predictions obtained using the NNLO setup and those obtained with $Q_0 = 1.0$ GeV are close to the default results over the full z range, except for the charm-tagged data, where the NNLO predictions are found to lie below the default results.

B. Extracted vacuum fragmentation functions

The extracted default vacuum FFs for gluon and quarks are shown in Fig. 4 at scales $Q = 5$ GeV and $Q = 100$ GeV. To assess the impact of different theoretical assumptions, we compare results obtained under several configurations, including NNLO corrections, a lower initial scale ($Q_0 = 1.0$ GeV), and the inclusion of theoretical uncertainties. In addition, a comparison with the NPC23 FFs [80], which are constrained by high-energy data, is included as a reference to assess the consistency of our results. It should be noted that the experimental data included in our analysis are primarily restricted to the region $z > 0.3$.

For the light quarks u and d, all configurations show good agreement for $z > 0.2$. Differences appear mainly at low z, where the NNLO result is slightly suppressed, while the $Q_0 = 1.0$ GeV configuration shows a mild enhancement. The inclusion of theoretical uncertainties has a negligible impact on the central values. Compared to our extraction, the NPC23 distributions are generally lower in the low z region. The heavy quark c and b FFs exhibit a high degree of consistency across all configurations and with NPC23. This robustness reflects the strong constraints provided by the SLD heavy-flavor-tagged measurements at the Z pole, which effectively fix the normalization and shape of heavy quark fragmentation. For the gluon distribution, the configurations considered in this work exhibit very similar overall shapes. The low z region ($z < 0.3$) shows a decrease, whereas a pronounced local maximum is observed around $z \approx 0.5$. Despite this similarity in shape, noticeable differences in magnitude arise across different configurations. The $Q_0 = 1.0$ GeV setup leads to a significant enhancement in the low z region, while the NNLO calculation produces a systematically suppressed distribution over the full z range. The NPC23 result exhibits a decrease over the full z range and is significantly larger than our result at low z, while becoming smaller around $z = 0.3$.

V. ANALYSIS OF NUCLEAR FRAGMENTATION FUNCTIONS

Based on the investigation of theoretical setups, we proceed to the extraction of nFFs. Compared to the vacuum FFs, the determination of nFFs is primarily based on SIDIS data from the HERMES and CLAS experiments. The datasets include a variety of nuclear targets, such as He, Ne, Kr, Xe, C, and Pb, spanning a broad range of atomic mass numbers, while the measurements are presented in various kinematic bins, including z, (z, ν), (z, Q^2), and (z, ν, Q^2), providing multidimensional constraints on the hadronization process.

We adopt the default theoretical setup established in the previous section at NLO perturbative accuracy, without theoretical uncertainties. The initial scale is set to $Q_0 = 1.3$ GeV. Within this framework, vacuum FFs and nuclear modifications are fitted simultaneously, with the ν dependence incorporated as described in Sec. III. To study

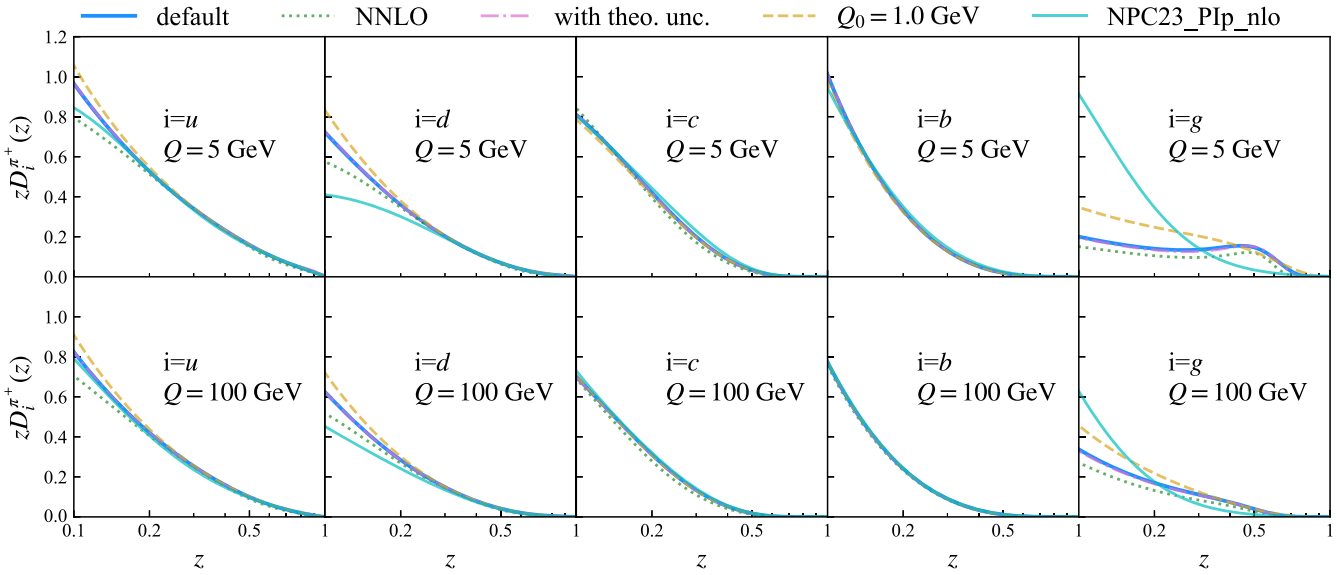


FIG. 4. Comparison of the extracted vacuum FFs for π^+ with the NPC23 [80] for different parton flavors at $Q = 5$ GeV and $Q = 100$ GeV.

the impact of the ν dependence, we consider three representative choices of $\delta = 0$, 0.5, 1 in Eq. (11). The sensitivity of the extracted nFFs to the assumed ν dependence is quantified by comparing the fit quality and the resulting variations across these scenarios.

To further assess the uncertainties arising from initial-state nuclear effects, we supplement the default nPDF set nCTEQ15WZ [39] with two recent global nPDF analyses: nNNPDF3.0 [40] and EPPS21 [41]. By repeating the nFFs extraction procedure with these three nPDF sets, we examine the stability of the extracted nFFs with respect to the choice of nPDFs.

A. Comparison with nuclear data

Table IV summarizes the quality of the fits, quantified by $\chi^2/N_{\rm pt}$, for different choices of nPDF sets and ν dependence. Since the dependence on the nPDF input is found to be weak, we first discuss this point to justify the use of the nCTEQ15WZ results in the following figures. The $\chi^2/N_{\rm pt}$ values for each individual dataset show a similar level of agreement across nCTEQ15WZ, nNNPDF, and EPPS21, indicating that the extraction of nFFs is not strongly affected by the specific implementation of initial state nuclear effects. To further illustrate this feature, a dedicated benchmark study is presented in the Appendix, where the default nFFs extracted using the nCTEQ15WZ set are kept fixed while the SIDIS cross sections are recalculated with nNNPDF3.0 and EPPS21. This weak dependence may be attributed to the use of hadron multiplicities, defined in Eq. (1) as the ratio of the semi-inclusive hadron-production cross section to the inclusive DIS cross section. Initial-state nuclear effects from nPDFs enter both the numerator and denominator of this ratio and tend to cancel. For clarity in the following comparison, only the results obtained with the nCTEQ15WZ are shown in the figures. Having established that the nPDF-set dependence is small, we now focus on the sensitivity to the assumed ν dependence. The results show that the $\delta = 1$ scheme generally leads to a significantly worse description of the data across all nPDF sets, as indicated by the larger $\chi^2/N_{\rm pt}$ values. In contrast, the $\delta = 0$ and $\delta = 0.5$ schemes both improve the description relative to $\delta = 1$ and give comparable fit qualities. The total χ^2 values alone therefore do not establish a clear preference between these two scenarios. A more differential comparison is provided by the HERMES [36] measurements, where the ν dependence of R_A^h can be examined in fixed z regions, as discussed below.

Figure 5 presents the HERMES [36] measurements of the nuclear modification ratio R_A^h as a function of ν for π^+ and π^- production on Ne, Kr, and Xe targets. The results are shown in two representative z regions, and the three ν-dependent parametrizations are displayed separately. This representation provides a direct test of the assumed

TABLE IV. Per-dataset and total $\chi^2/N_{\rm pt}$ values for all datasets using different nuclear PDF parametrizations (nCTEQ15WZ, nNNPDF, EPPS21) with various $\delta = 0$, $\delta = 0.5$, $\delta = 1$ schemes.

					nCTEQ ($\chi^2/N_{\rm pt}$)			nNNPDF ($\chi^2/N_{\rm pt}$)			EPPS21 ($\chi^2/N_{\rm pt}$)		
Exp	Target (A)	Particle	Binning	$N_{\rm pt}$	$\delta=0$	$\delta=0.5$	$\delta=1$	$\delta=0$	$\delta=0.5$	$\delta=1$	$\delta=0$	$\delta=0.5$	$\delta=1$
HERMES [43]	p	π^0	z	10	2.30	2.57	2.53	2.22	2.55	2.50	2.23	2.56	2.52
HERMES [35]	He	$\pi^\pm$	z, ν	16	0.23	0.25	0.25	0.21	0.24	0.24	0.23	0.26	0.26
	He	$\pi^\pm$	z, Q^2	20	0.20	0.25	0.25	0.20	0.25	0.25	0.21	0.25	0.25
	He, Ne, Kr, Xe	π^+	z	24	0.31	0.61	1.44	0.42	0.60	1.35	0.33	0.59	1.44
	He, Ne, Kr, Xe	π^-	z	24	0.17	0.44	0.99	0.50	0.75	1.40	0.18	0.44	1.05
	He, Ne, Kr, Xe	π^0	z	24	0.28	0.54	1.14	0.28	0.47	1.10	0.28	0.51	1.16
HERMES [36]	Ne, Kr, Xe	π^+	z, Q^2	60	0.49	2.37	2.44	0.65	1.83	1.76	0.44	1.96	2.01
	Ne, Kr, Xe	π^-	z, Q^2	60	0.70	2.29	2.41	0.76	2.22	2.38	0.63	2.03	2.11
	Ne, Kr, Xe	π^+	z, ν	51	2.47	1.11	2.41	1.88	1.02	2.84	2.13	1.03	2.60
	Ne, Kr, Xe	π^-	z, ν	51	2.27	0.96	1.40	2.52	1.09	1.60	2.06	0.86	1.47
HERMES [42]	p	π^+	z, Q^2	25	1.17	1.20	1.32	1.08	1.03	1.11	1.13	1.09	1.19
	p	π^-	z, Q^2	25	0.83	0.77	0.90	0.81	0.96	1.14	0.84	0.88	1.04
	D	π^+	z, Q^2	25	1.25	1.21	1.28	1.36	1.16	1.16	1.33	1.20	1.24
	D	π^-	z, Q^2	25	1.33	1.16	1.19	1.17	1.19	1.28	1.32	1.19	1.24
CLAS [37]	C, Fe, Pb	π^+	z, Q^2, ν	117	0.83	0.85	1.47	0.81	0.84	1.56	0.82	0.83	1.45
	C, Fe, Pb	π^-	z, Q^2, ν	114	1.99	1.62	1.60	2.04	1.65	1.58	2.02	1.66	1.64
			SIDIS total:	671	1.19	1.27	1.60	1.20	1.23	1.61	1.15	1.20	1.56
SLD [44]	c-tagged	$\pi^\pm$	z	21	0.68	0.87	0.90	0.72	0.88	0.92	0.67	0.85	0.89
	b-tagged	$\pi^\pm$	z	21	0.81	0.89	0.90	0.83	0.90	0.90	0.81	0.89	0.90
			SIA total:	42	0.75	0.88	0.90	0.78	0.89	0.91	0.74	0.87	0.90
			Total:	713	1.17	1.24	1.56	1.17	1.21	1.57	1.12	1.18	1.52

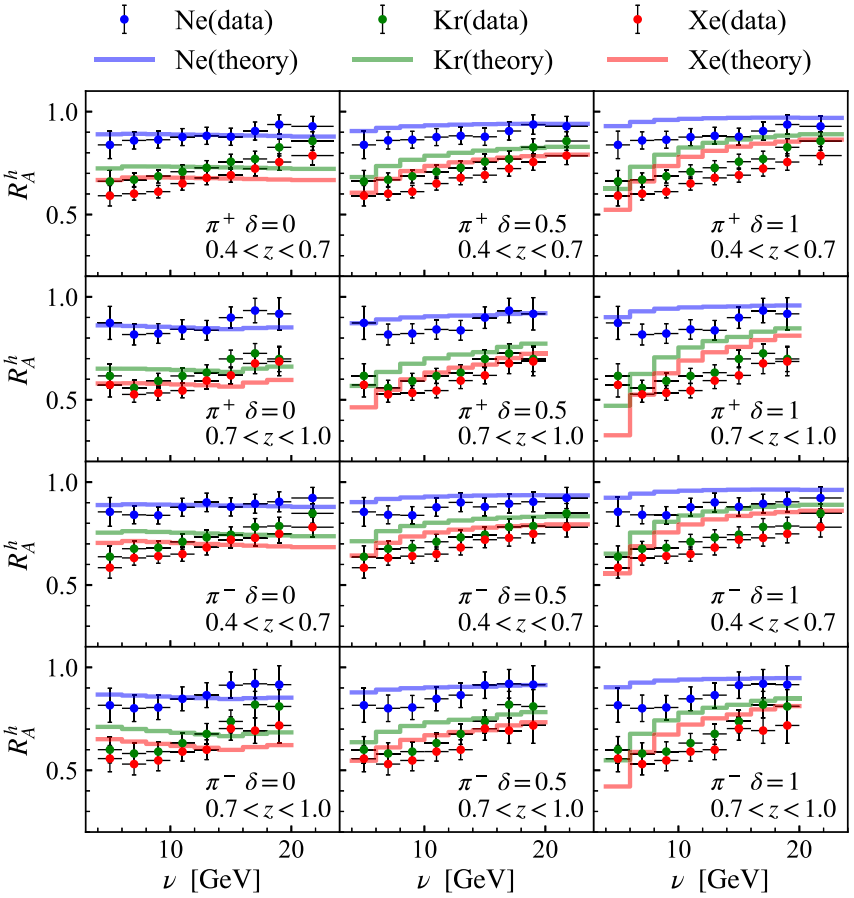


FIG. 5. Comparison of the measured and predicted nuclear modification with respect to z and ν for π^+ and π^- in HERMES experiments [36].

ν dependence in the nFFs. A clear increase of R_A^h with ν is observed, especially for the heavier Kr and Xe and in the larger z region. This behavior indicates that the nuclear modification becomes weaker as the energy transferred to the fragmenting parton increases. The $\delta = 0$ scheme, which contains no explicit ν dependence in the nuclear modification, therefore does not capture the observed rise as well and tends to give a flatter behavior. In contrast, the $\delta = 1$ scheme generates a stronger increase with ν, but it generally overestimates the data, in particular for heavier nuclei and at larger z. The $\delta = 0.5$ scheme provides the most balanced description. It captures the increasing trend with ν while keeping the overall magnitude of R_A^h close to the data, although localized deviations remain. The persistence of these deviations illustrates the complexity of simultaneously describing the detailed A and ν dependences across different nuclear targets and kinematic regions. This challenge is also evident in the TMD-based study of Ref. [38], where nuclear SIDIS is described through medium-induced parton energy loss and prehadron absorption.

Figure 6 shows the HERMES [36] measurements as a function of Q^2 for π^+ and π^- production on Ne, Kr, and Xe targets in different z bins. The data show the expected nuclear mass dependence, with the suppression becoming stronger from Ne to Kr and Xe. The overall dependence on Q^2 is relatively mild in all three z intervals, while the suppression becomes more pronounced as z increases. For π^+ production, both the $\delta = 0$ and $\delta = 0.5$ schemes capture the overall features of the data, considering the relatively large experimental uncertainties, especially in the large-z region and at high Q^2. The $\delta = 0.5$ predictions are slightly higher in some bins, particularly for heavier nuclei and at larger z, while the $\delta = 1$ scheme tends to overestimate the data. For π^- production, similar features are observed. The data show a weak Q^2 dependence, while the suppression becomes stronger from Ne to Kr and Xe. Compared with the π^+ case, the π^- data exhibit more visible variations, especially in the intermediate and large z intervals. The $\delta = 0$ and $\delta = 0.5$ schemes both follow the main trends of the measurements, while the $\delta = 1$ scheme tends to give higher R_A^h values, especially for the heavier targets.

Figure 7 presents the 2007 HERMES [35] measurements of the nuclear modification ratio R_A^h for π^+, π^-, and π^0 on He, Ne, Kr, and Xe targets. For each nuclear target, the z distributions of π^+, π^-, and π^0 are largely similar in shape.

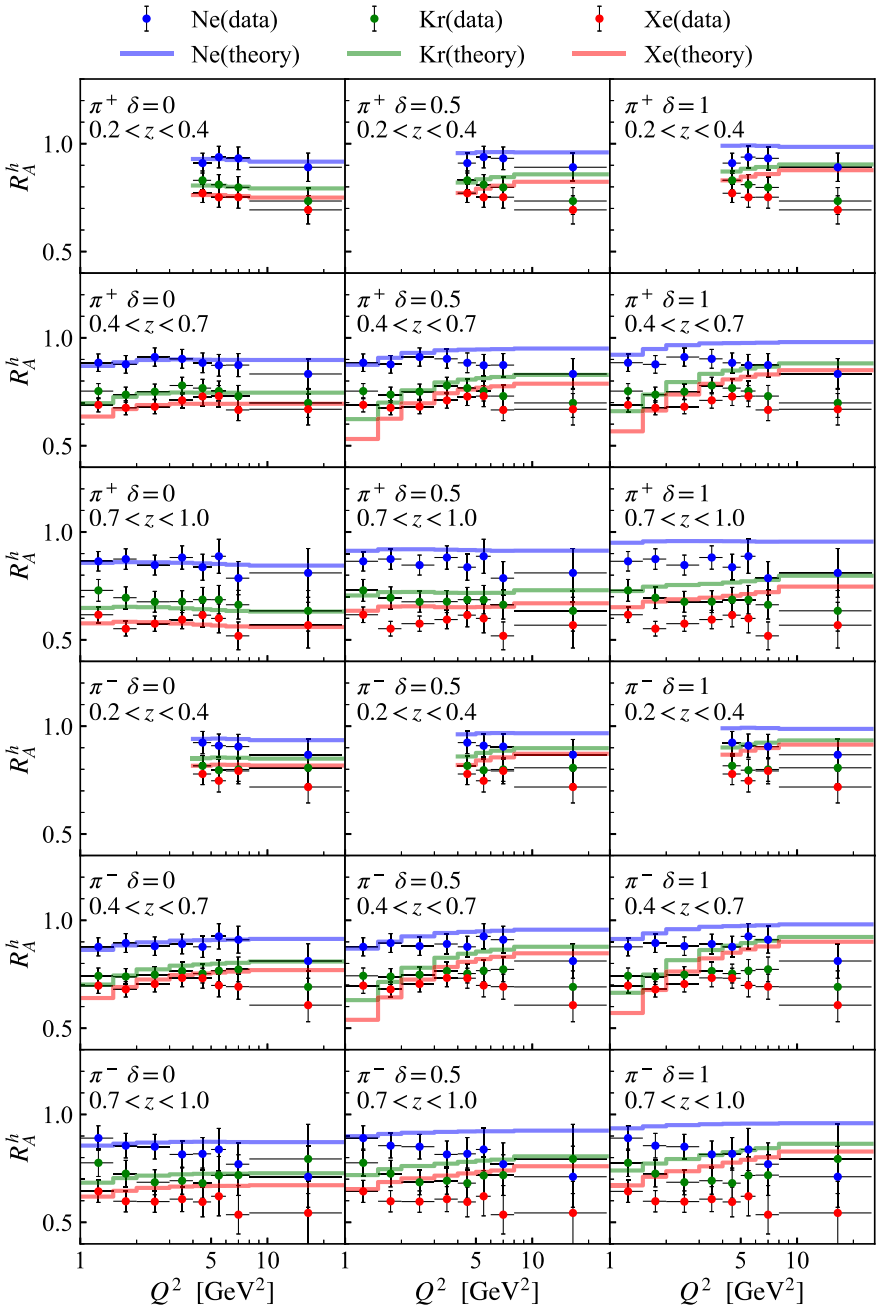


FIG. 6. Similar to Fig. 5, but shown with respect to z and Q^2.

The He data remain close to unity over the measured z range, showing that nuclear effects are weak for the lightest target. For Ne, Kr, and Xe, a clear suppression pattern develops with increasing z, becoming stronger for heavier nuclei, which provides important experimental input for the determination of the A dependence in the present analysis. From the figure, we observe that all three ν schemes provide a satisfactory description of the data, supporting a stable and consistent extraction.

Figure 8 shows the nuclear modification ratio R_A^h for π^+ and π^- measured by the CLAS Collaboration. A clear overall trend is observed, with R_A^h decreasing as a function of z. For π^+ production, the Q^2 dependence can be seen by comparing the z distributions across different Q^2 bins within the same ν interval, with higher Q^2 values leading to an upward shift of the R_A^h distributions. This effect becomes less pronounced for heavier nuclei, suggesting a reduced sensitivity to Q^2 as the mass number A increases. The ν dependence is inferred by comparing the z distributions across different ν bins within the same Q^2 interval. In the low-z region, the data exhibit only a weak dependence on ν, while at larger z an enhancement with increasing ν becomes visible. A comparison among different ν schemes shows that the $\delta = 0$ scheme, which gives the

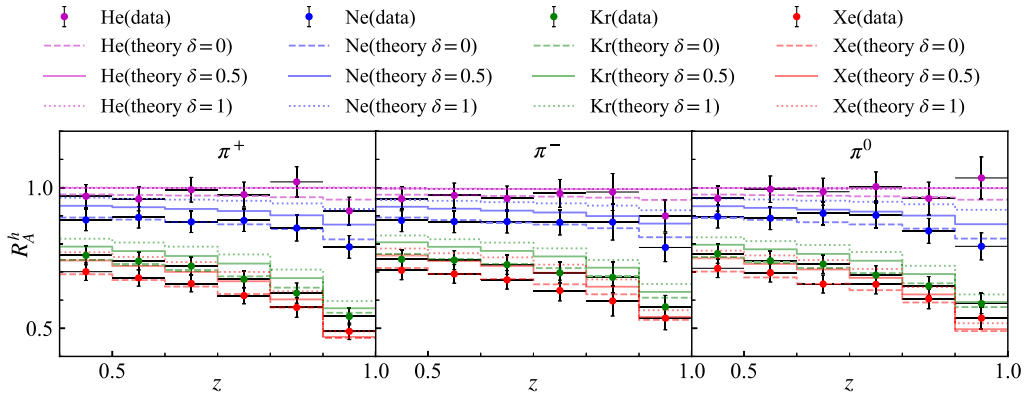


FIG. 7. Comparison of the measured and predicted nuclear modification with respect to z for π production in HERMES experiments [35]. Theoretical predictions are obtained for different ν dependences and curves.

same nuclear modification at fixed z and Q^2 when only ν is varied, slightly deviates from the data in the lowest ν interval but performs reasonably well elsewhere. The predictions of the $\delta = 0.5$ scheme provide a good overall description of the data, with minor overestimation observed at higher ν. Meanwhile, the $\delta = 1$ scheme tends to overestimate the data at intermediate and large ν, consistent with its larger χ^2 value reported in Table IV. For π^- production, however, more pronounced structures emerge at larger z ($z > 0.5$). In particular, peaklike features develop in the $3.2 < \nu < 3.7$ GeV range, becoming more visible at $3.7 < \nu < 4.2$ GeV range. These structures are most prominent in the lower Q^2 bins and are accompanied by increasing experimental uncertainties. Despite these localized features, the predictions of the $\delta = 0.5$ scheme provide a satisfactory description of the π^- data across $2.2 < \nu < 3.2$ GeV and $3.2 < \nu < 3.7$ GeV regions. In the higher-ν range ($3.7 < \nu < 4.2$ GeV), all three ν parametrizations show a relatively poor description of the data. However, given the relatively large experimental uncertainties in this region, the $\chi^2/N_{\rm pt}$ values of CLAS π^- remain within a reasonable range.

B. Extracted nuclear fragmentation functions

In this subsection, we first present the extracted nFFs obtained within the $\delta = 0.5$ scheme, which incorporates the dependence of both A and ν. Figure 9 summarizes the results at the scales $Q = 5$ GeV for u quarks, d quarks, and gluons on C, Kr, and Pb targets, which allows for a direct comparison of different flavors at fixed A, as well as nFFs with increasing A for a given flavor. For the default $\delta = 0.5$ scheme, each panel shows the results at three representative values of the energy transfer, $\nu = 2$ GeV, $\nu = 15$ GeV, and $\nu \to \infty$. In this scheme, the $\nu \to \infty$ limit corresponds to the vacuum FFs because $\mathcal{F}_A(A, \nu)$ vanishes for $\delta > 0$. For comparison, the ν-independent result obtained with the $\delta = 0$ scheme is also shown.

In the region $z > 0.1$, the nuclear modifications for all parton species are reasonably well constrained by the available experimental data. For u and d quarks, the curves for $\nu = 2$ GeV, $\nu = 15$ GeV, and $\nu \to \infty$ show a clear separation, with the prediction of vacuum FFs part lying above the $\nu = 15$ GeV result, which in turn remains above the $\nu = 2$ GeV curve. This behavior follows naturally from the ν-dependent structure of the parametrization in the $\delta = 0.5$ scheme, where smaller values of ν lead to stronger suppression for the same A. The difference becomes more pronounced with increasing mass number A. The nuclear suppression of the u quark is found to reach about 20% for a C nucleus at $\nu = 2$ GeV, and increases to nearly 50% for Pb. A similar trend is observed for the d quark, where the suppression is about 20% for C at $\nu = 2$ GeV and rises to nearly 30% for Pb. This indicates that the sensitivity of nuclear medium effects to the variable ν is enhanced in heavier nuclei. For gluons, the curves for $\nu = 2$ GeV, $\nu = 15$ GeV, and $\nu \to \infty$ are very close to each other. This behavior arises because the gluon nuclear modifications are set to zero at the initial scale and are generated only through QCD evolution.

As discussed in Sec. III B, the normalization of nuclear modification is characterized by the parameter a_0 in Eq. (10) and the function $f(A)$ in Eq. (11), where the data points of $f(A)$ are determined independently for each mass number A within the framework. To investigate this A dependence, we perform independent fits to the extracted $f(A)$ data points using several parametrizations employed in studies of nPDFs and nFFs [81,82],

$$f(A) = aA^b - a, \quad f(A) = aA^b \ln A, \quad f(A) = a \ln A, \tag{13}$$

where a and b are free parameters to be fitted, and all parametrizations are constructed to satisfy the physical constraint $f(A = 1) = 0$. The corresponding results for the $\delta = 0.5$ scheme, normalized by the parameter a to compare the A-dependent shape, are shown in the left panels of Fig. 10. For comparison, the results obtained with the $\delta = 0$

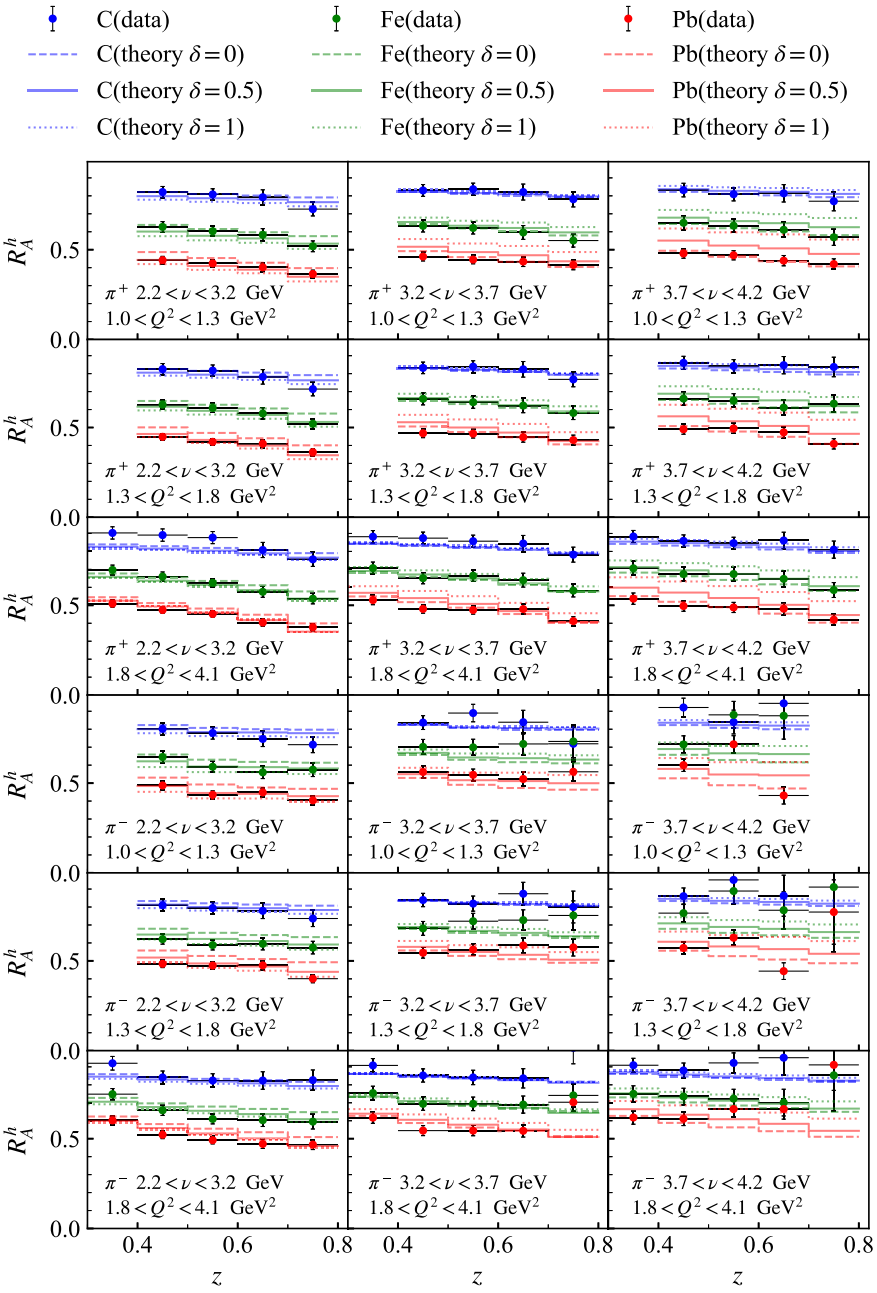


FIG. 8. Comparison of the measured and predicted nuclear modification with respect to z and ν for π^+ and π^- in CLAS experiments [37]. Theoretical predictions are shown for different assumptions on the ν dependence. The C, Fe, and Pb targets are shown in blue, green, and red, respectively. The $\delta = 0.5$, $\delta = 0$, and $\delta = 1$ schemes are represented by solid, dashed, and dotted lines, respectively.

scheme are displayed in the right panels. The extracted $f(A)$ data points are obtained from separate nFF fits performed with the three different nPDF inputs, and are shown with error bars in different colors for comparison.

Consistent with our previous findings, the results show that the normalized $f(A)$ exhibits a similar behavior across different nPDFs, indicating that the determination of nuclear modifications is insensitive to the choice of nPDFs. A comparison between the $\delta = 0.5$ and the $\delta = 0$ schemes reveals that the latter exhibits noticeably larger uncertainties in the extracted $f(A)$ values, as reflected by the wider error bars. This leads to a reduced χ^2 in the corresponding fits. Such behavior can be understood from the fact that the $\delta = 0$ scheme does not include any explicit ν dependence, thereby shifting more of the variation into the A dependence and resulting in a stronger, but less constrained, A-dependent behavior. For the parametrization $f_A = aA^b - a$, shown in Fig. 10, the fit provides a good description of the extracted $f(A)$ values. The exponent b is found to lie in the range 0.35–0.45, with

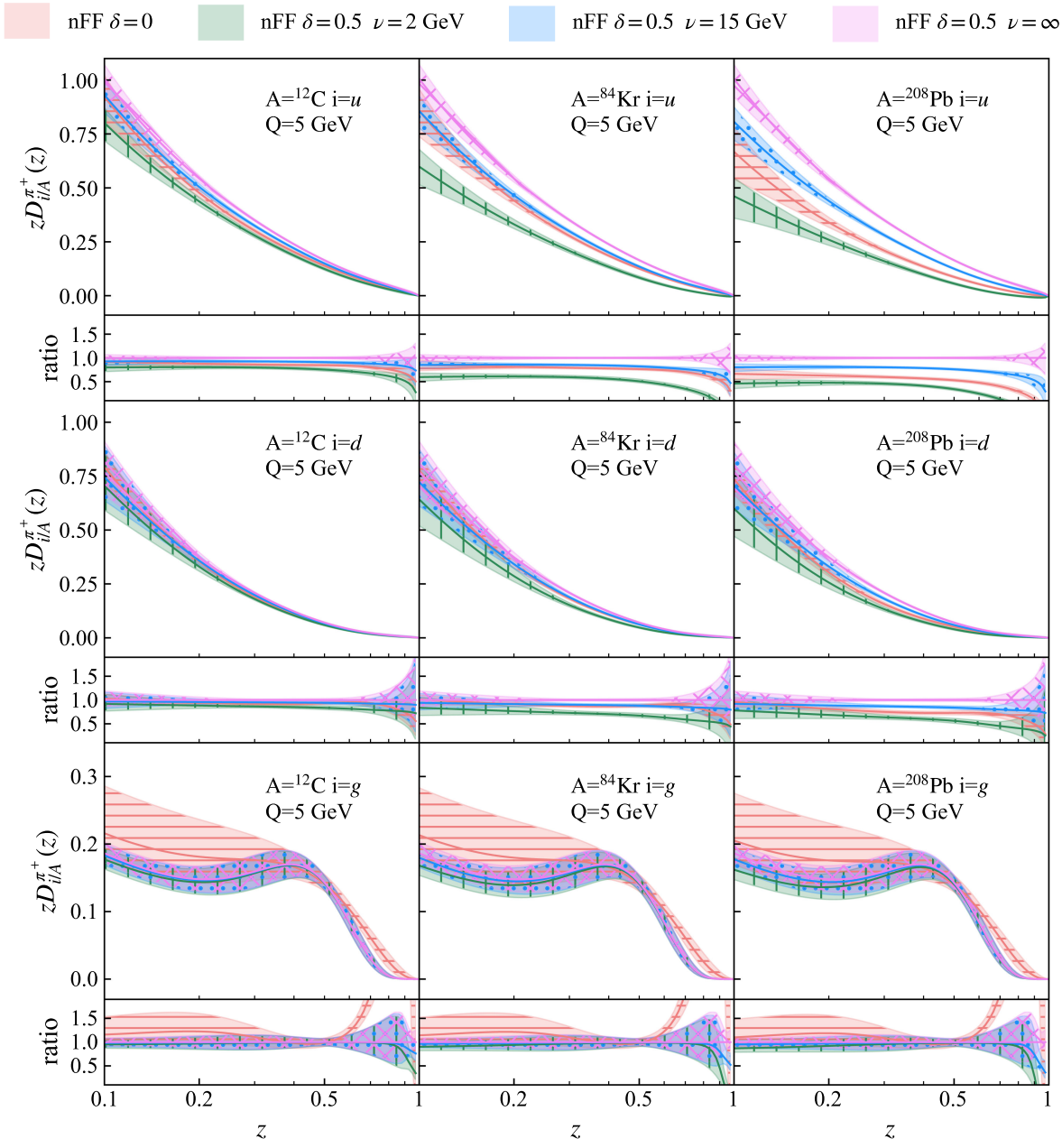


FIG. 9. Extracted nFFs for π^+ obtained with the nCTEQ15WZ nPDFs for u quarks, d quarks, and gluons on C, Kr, and Pb targets at $Q = 5$ GeV, and ratios normalized to the case with $\delta = 0.5$ and $\nu = \infty$. The results are shown for the default $\delta = 0.5$ scheme at $\nu = 2$ GeV, $\nu = 15$ GeV, and $\nu \to \infty$, together with the ν-independent result from the $\delta = 0$ scheme. The Hessian uncertainty bands are also included.

little dependence on the choice of nPDF set or ν scheme. This range is consistent with the expected scaling behavior between $A^{1/3}$ and $A^{2/3}$, as predicted by the multiparton scattering mechanism [82]. The alternative parameterizations, $aA^b \ln A$ and $a \ln A$, shown in Fig. 10, also provide a satisfactory description of the extracted $f(A)$ values. These forms introduce additional flexibility through logarithmic dependence while preserving the constraint $f(1) = 0$.

VI. PREDICTIONS FOR PA COLLISIONS

In this section, we provide predictions for hadron production in pp and pA collisions. Recent measurements by the ALICE Collaboration [83] at the LHC provide high-precision $\gamma^{\rm iso}$-tagged fragmentation functions in pp and pPb collisions at a center-of-mass energy of $\sqrt{s_{\rm NN}} = 5.02$ TeV. The measurement considers isolated photons in the range $|\eta| < 0.67$ and $12 < p_{\rm T}^{\gamma} < 40$ GeV,

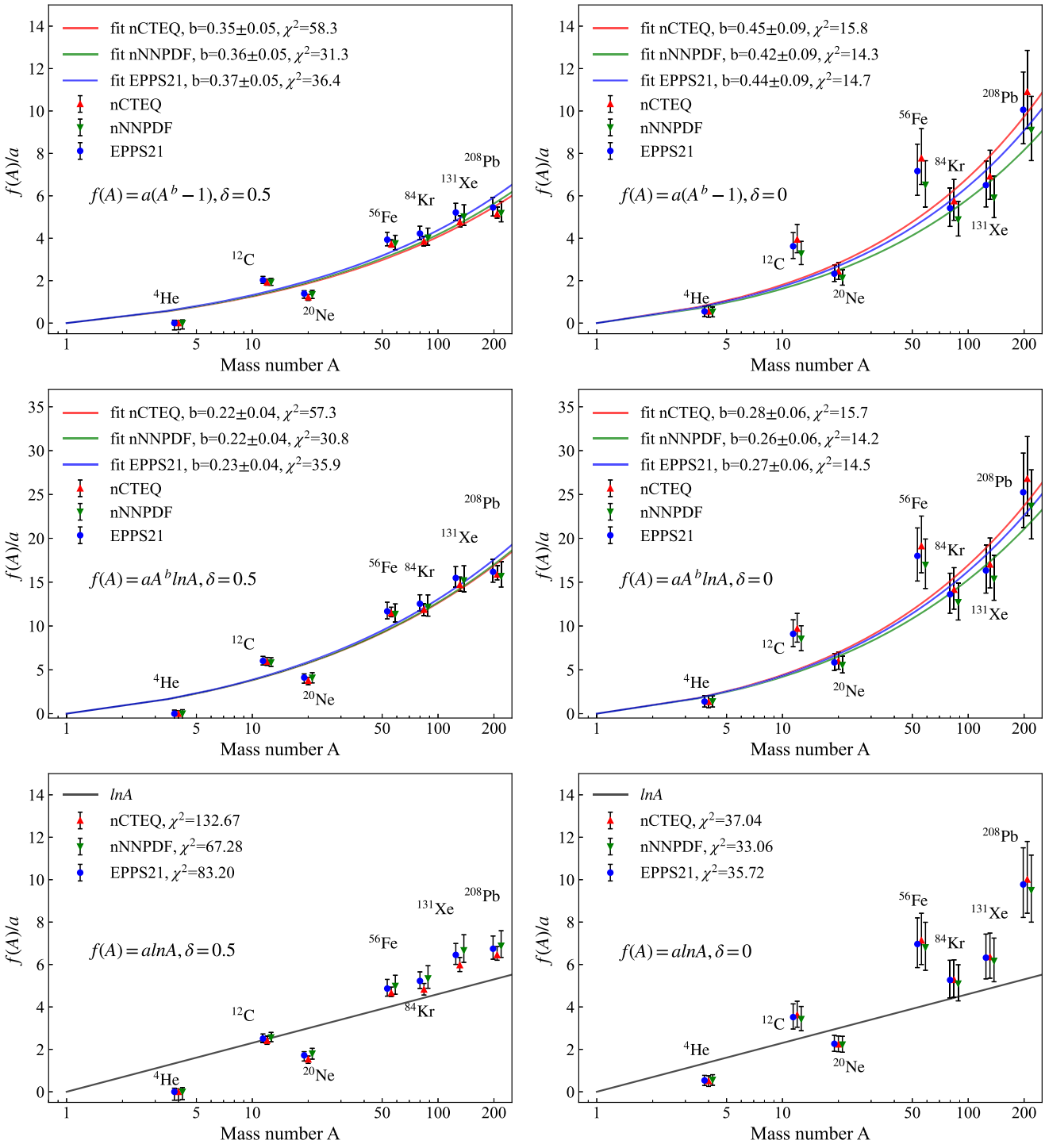


FIG. 10. A dependence of the function $f(A)$ for the $\delta = 0.5$ and $\delta = 0$ schemes. The extracted $f(A)$ points are obtained from separate nFF fits using different nPDF inputs and are compared with several functional forms normalized by the parameter a to show the A-dependent shape.

and reports the associated yield of charged hadrons within $|\eta| < 0.80$ and $0.5 < p_{\mathrm{T}}^h < 10$ GeV. The γ^{iso}-tagged fragmentation functions are measured in different forms with respect to $z_{\mathrm{T}} = p_{\mathrm{T}}^h / p_{\mathrm{T}}^\gamma$. This definition relates the transverse momentum of the produced hadron to that of the isolated photon and, owing to the momentum balance between the photon and the recoil parton, allows z_{T} to serve as a good approximation of the hadron energy fraction z in FFs.

Furthermore, as in SIDIS where ν represents the virtual-photon energy in the target rest frame, a corresponding quantity can be introduced for isolated photon-hadron

production in pPb collisions. For the partonic subprocess $a + b \to \gamma + c$, the final-state parton c fragments into the observed hadron. We define the analog of ν as the energy of this fragmenting parton in the rest frame of the Pb nucleus,

$$\nu_c^{\mathrm{Pb}} \equiv p_c^{\mu} u_{\mu}^{\mathrm{Pb}} = p_{T,c} \cosh(y_c - Y_{\mathrm{Pb}}), \tag{14}$$

where p_c^{μ} is the four-momentum of the fragmenting parton, and u_{Pb}^{μ} is the four-velocity of the Pb nucleus. In the last equality, $p_{T,c}$ and y_c denote the transverse momentum and rapidity of the fragmenting parton, while Y_{Pb} is the rapidity of the Pb beam. Unlike fixed-target SIDIS measurements, where the nucleus is at rest and ν is typically below 50 GeV, the corresponding fragmenting parton energy in pPb collisions is strongly enhanced by the Lorentz boost of the colliding nucleus and can easily reach several hundred GeV or even TeV scale. For the ALICE kinematics considered here, a fragmenting parton with $p_{T,c} \simeq p_T^{\gamma} =$ 12–40 GeV produced near midrapidity corresponds to $\nu_c^{\mathrm{Pb}} \sim 20$–$70$ TeV in the Pb rest frame. As a consequence, γ^{iso}-tagged fragmentation functions in pp and pPb collisions provide a useful and independent probe of possible ν-dependent nuclear modifications in the fragmentation process.

Figure 11 shows the ALICE measurements of γ^{iso}-tagged fragmentation functions in pp collisions at $\sqrt{s} = 5.02$ TeV, together with our NLO vacuum FFs prediction. It should be noted that the experimental measurements are for charged hadrons, whereas our calculation includes only the $\pi^{\pm}$ contribution. The contributions from $K^{\pm}$ and $p/\bar{p}$, which are expected to amount to roughly 20% of the π contribution, are not included. Therefore, the comparison should be interpreted as a π-only approximation to the charged-hadron observable. Despite this limitation, the prediction gives a reasonable description of the measured γ^{iso}-tagged fragmentation functions and provides a useful baseline for the study of nuclear effects in pPb collisions.

Using the same π-only setup, Fig. 12 presents pPb-to-pp nuclear modification ratio for the γ^{iso}-tagged fragmentation functions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. The three panels correspond to the predictions obtained with the nFFs extracted using the nCTEQ15WZ, EPPS21, and nNNPDF nPDF sets, respectively. In each case, the $\delta = 0$, 0.5, and 1 schemes are also shown to illustrate the sensitivity to the assumed ν dependence. For pPb collisions, the theoretical predictions show a distinct pattern, where the $\delta = 0.5$ and $\delta = 1$ schemes are very similar over the full z_{T} range and remain close to unity, while the $\delta = 0$ scheme tends to give smaller nuclear modification ratios. This difference originates from the fact that the $\delta = 0$ scheme does not include ν dependence in the fit, leading to a situation where the

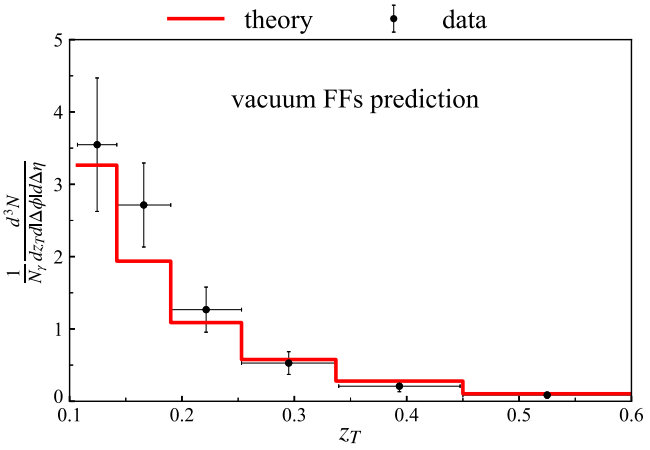


FIG. 11. Comparison of the $\pi^{\pm}$ NLO prediction obtained with the default vacuum FFs and ALICE measurements for γ^{iso}-tagged fragmentation functions of a charged hadron in pp collisions at $\sqrt{s} = 5.02$ TeV.

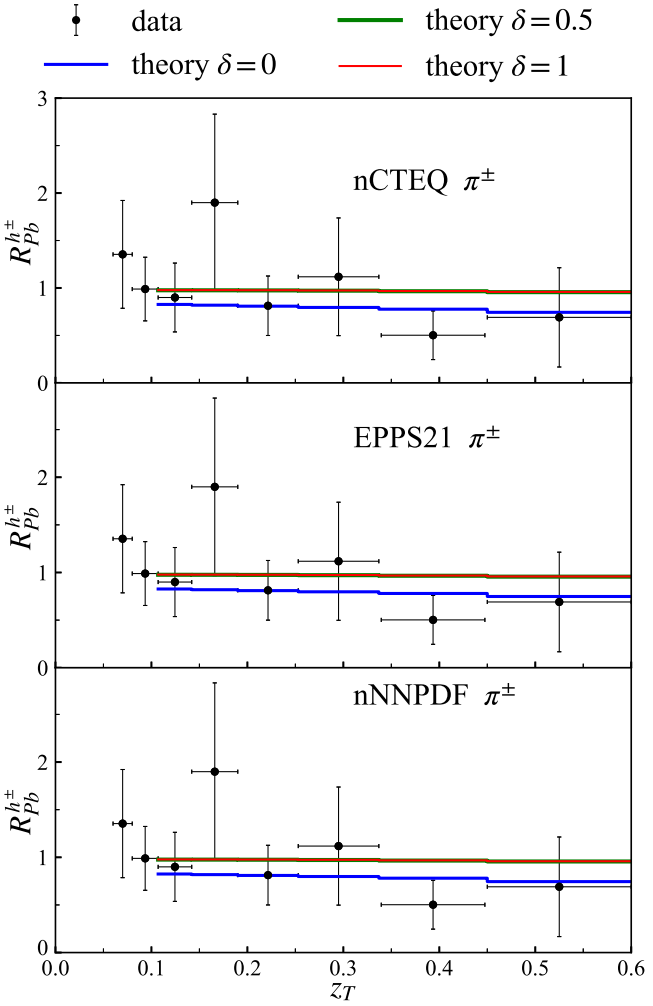


FIG. 12. pPb-to-pp nuclear modification ratio for γ^{iso}-tagged fragmentation functions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, compared with the ALICE measurements. The three panels show predictions obtained with nFFs extracted using the nCTEQ15WZ, EPPS21, and nNNPDF nPDF inputs, respectively.

nuclear suppression is carried over to pA collisions, without allowing it to vary with the relevant energy scale ν. As discussed above, this fragmenting parton energy scale in pPb is much larger than the virtual-photon energy ν in the target rest frame. In contrast, the $\delta = 0.5$ and $\delta = 1$ schemes incorporate the ν-dependent modification, which allows the nuclear effects to respond to the appropriate ν scale in pPb collisions. Despite these differences, all three schemes provide a reasonable description of the experimental data within the relatively large experimental uncertainties, which do not allow a clear discrimination among them. The results also show only a weak dependence on the choice of nPDF set for all δ schemes.

Overall, within the π-only setup, the present results are compatible with the γ^{iso}-tagged fragmentation functions measurement in pp collisions and with the corresponding pPb-to-pp nuclear modification ratio. More precise measurements from LHC experiments such as ALICE, CMS, and ATLAS would be valuable for further constraining nuclear fragmentation effects and improving the discrimination among different δ schemes.

VII. DISCUSSION AND CONCLUSIONS

Understanding hadronization in the nuclear environment remains an important challenge in QCD. Collinear QCD factorization with modified nFFs offers an approach to describe it. In this framework, nFFs serve as nonperturbative inputs that effectively encode the medium-induced modifications of the hadronization process inside nuclei. In this work, we have presented an extraction of nFFs for π, based on a comprehensive analysis of SIDIS measurements on nuclear targets. The analysis has been performed at NLO accuracy in perturbative QCD with uncertainties quantified using the Hessian method. In the kinematic region $z \sim (0.2, 0.8)$, where precise measurements are available, the extracted nFFs are well constrained. As for the quality of the fit, reasonable χ^2 values are obtained for most datasets from various observables. With the inclusion of ν-dependent nuclear measurements in the fit of π nFFs, we are able to explore different possible forms of ν-dependence of the extracted nFFs.

Our work introduces several methodological aspects in the treatment of the fit, apart from the selections of kinematics. In previous nFF analyses, the vacuum FFs were typically fixed to existing vacuum FFs obtained mainly from global fits to high-energy experimental data, while only the nuclear modifications were fitted. However, such an approach may not fully account for the fact that SIDIS measurements on nuclear targets are predominantly performed in a relatively low Q region, where the existing vacuum FFs may not provide a sufficiently accurate description. Motivated by this consideration, we adopt a framework in which vacuum FFs and nuclear modifications are extracted simultaneously, allowing the vacuum FFs to be constrained by low-energy experimental data relevant for the determination of nFFs. To establish a reliable fitting setup for the nFFs extraction, we first performed a series of independent fits of vacuum FFs using SIDIS data on p and D targets together with heavy-flavor-tagged SIA measurements. Within these studies, we investigated the stability of the extracted vacuum FFs under different theoretical setups, including variations of the perturbative accuracy between NLO and NNLO, different choices of the initial scale $Q_0 = 1.0$ and 1.3 GeV, as well as the impact of including theoretical uncertainties in the fit. Based on these studies, we selected the default setup that provides the most stable and reliable baseline for the subsequent extraction of nuclear modifications. The vacuum FFs and nuclear modifications were then fitted simultaneously within the selected theoretical setup. The dependence of the nuclear effects on ν was first explored through different parametrization scenarios. Independent fits were performed for the $\delta = 0$, $\delta = 0.5$, and $\delta = 1$ schemes introduced in Eq. (11). Comparisons among these scenarios indicate that the $\delta = 1$ scheme provides a poorer description of the data, while the $\delta = 0$ and $\delta = 0.5$ schemes give comparable global fit qualities. In the HERMES comparison as a function of ν in fixed z regions, the $\delta = 0.5$ scheme gives a more balanced description of the observed increasing trend of R_A^h with ν. At $Q = 5$ GeV and $\nu = 2$ GeV, the nuclear modification of $\delta = 0.5$ scheme reaches up to about 20% for C and up to nearly 50% for Pb, with a weaker effect observed for d quarks compared to u quarks. Furthermore, the sensitivity of the extracted nFFs to the choice of initial-state nPDFs has also been examined using several nPDF sets, where the results show only a weak dependence on the nPDF input. The nuclear-mass dependence of the extracted modifications has been investigated through independent fits of the function $f(A)$ in Eq. (11) for different nuclear targets. The results indicate a power-like behavior with an effective exponent in the range 0.35–0.45, which provides a good description of the data. This behavior is consistent with expectations from multiple parton scattering mechanisms. As an application, the extracted nFFs have been used to provide predictions for γ^{iso}-tagged fragmentation functions in pp and pPb collisions at LHC energies. Within the π-only approximation used in this work, the comparison with ALICE data shows a reasonable agreement within relatively large uncertainties. This suggests that such observables can provide useful complementary constraints on nuclear modifications in high-energy pA collisions. More precise measurements from LHC experiments would be highly valuable for further constraining nuclear fragmentation effects and improving the sensitivity to different implementations of the ν dependence.

This work is a continuation of the previous NPC FFs analysis with extension to nFFs. Taking into account the new constraints provided by SIDIS data with ν information, the present analysis establishes a framework for future

studies of nFFs and hadronization in nuclear environments. The extracted FFs are publicly available in LHAPDF format.

ACKNOWLEDGMENTS

The work of J. G. is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 12275173, Shanghai Municipal Education Commission under Grant No. 2024AIZD007, and open fund of Key Laboratory of Atomic and Subatomic Structure and Quantum Control (Ministry of Education). H. X. is supported by the NSFC under Grants No. 12525508, No. 12475139. Y. Z. is supported by the NSFC under Grant No. U2032105 and the CAS Project for Young Scientists in Basic Research No. YSBR-117.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX: IMPACT OF NPDFS

We conducted a benchmark study to quantify the impact of different nPDF sets on the theoretical predictions. This step is important in order to disentangle possible uncertainties associated with the initial-state nuclear structure from those originating from the final-state fragmentation process. Since both nPDFs and nFFs enter the SIDIS cross sections simultaneously within the collinear factorization framework, it is necessary to verify that the extracted nuclear modifications are not biased by the particular choice of nPDF parametrization. In this benchmark, the baseline nuclear fragmentation functions were first extracted using the nCTEQ15WZ nPDF set. The extracted nFFs were then kept fixed, while the SIDIS cross sections were recalculated using alternative nPDF sets, namely EPPS21 and nNNPDF. In this way, the role of the initial-state nuclear input can be isolated, allowing a direct assessment of how the nPDF choice propagates into the final observable predictions.

For clarity of presentation, representative HERMES and CLAS datasets are shown in Figs. 13 and 14. As can be seen from these figures, replacing the nPDF input from nCTEQ15WZ to EPPS21 or nNNPDF while keeping the nFFs fixed leads only to very small changes in the resulting predictions. The overall agreement with the experimental data remains essentially unchanged for all considered nPDF sets. Overall, these results further prove that the extracted nFFs exhibit good stability against variations of the nPDF input. The residual dependence on the choice of nPDF set is significantly smaller than the differences associated with the various δ-scaling scenarios considered in the main analysis.

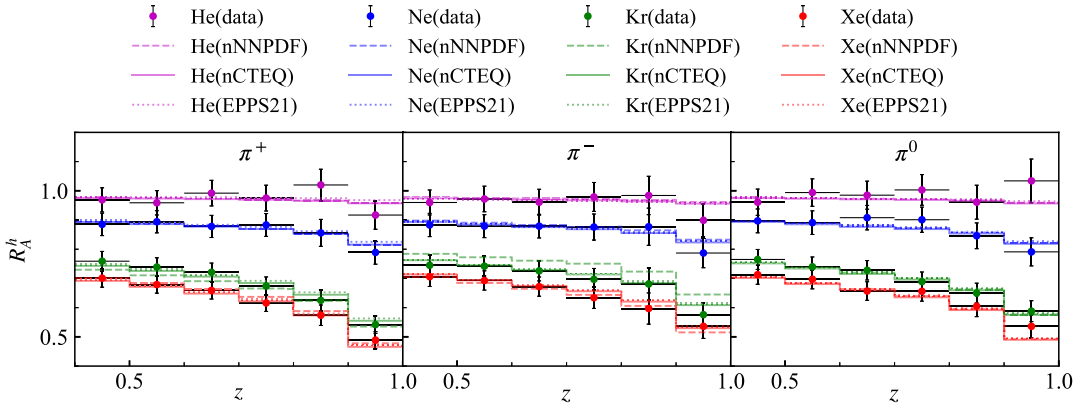


FIG. 13. Benchmark comparison of measured and predicted nuclear modifications for pion production in HERMES experiments [35] using different nPDF inputs while keeping the extracted nFFs fixed.

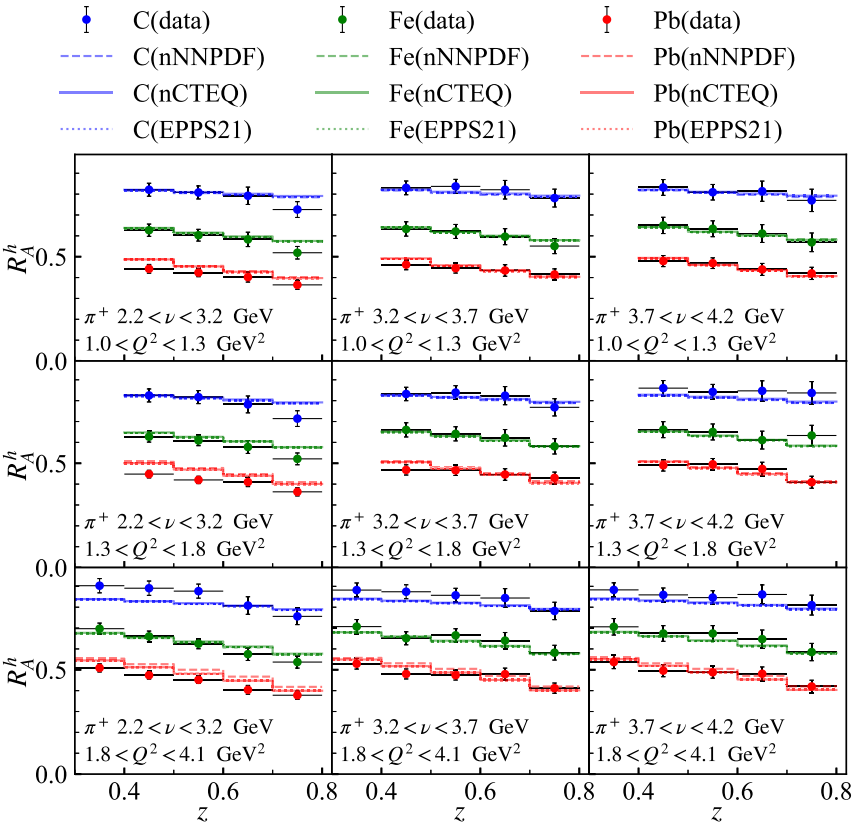


FIG. 14. Similar to Fig. 13, but for the CLAS measurements [37] shown as functions of z and ν.

[1] R. D. Field and R. P. Feynman, A parametrization of the properties of quark jets, Nucl. Phys. **B136**, 1 (1978).

[2] A. Metz and A. Vossen, Parton fragmentation functions, Prog. Part. Nucl. Phys. **91**, 136 (2016).

[3] J. C. Collins, D. E. Soper, and G. F. Sterman, Factorization of hard processes in QCD, Nucl. Phys. **B308**, 833 (1989).

[4] J. C. Collins and D. E. Soper, Back-to-back jets in QCD, Nucl. Phys. **B193**, 381 (1981); **B213**, 545(E) (1983).

[5] J. C. Collins and D. E. Soper, Parton distribution and decay functions, Nucl. Phys. **B194**, 445 (1982).

[6] S. Forte, Parton distributions at the dawn of the LHC, Acta Phys. Pol. B **41**, 2859 (2010).

[7] S. Forte and G. Watt, Progress in the determination of the partonic structure of the proton, Annu. Rev. Nucl. Part. Sci. **63**, 291 (2013).

[8] J. Rojo *et al.*, The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II, J. Phys. G **42**, 103103 (2015).

[9] K. Kovařík, P. M. Nadolsky, and D. E. Soper, Hadronic structure in high-energy collisions, Rev. Mod. Phys. **92**, 045003 (2020).

[10] D. de Florian and R. Sassot, Nuclear parton distributions at next-to-leading order, Phys. Rev. D **69**, 074028 (2004).

[11] M. Hirai, S. Kumano, and T. H. Nagai, Determination of nuclear parton distribution functions and their uncertainties in next-to-leading order, Phys. Rev. C **76**, 065207 (2007).

[12] K. J. Eskola, H. Paukkunen, and C. A. Salgado, Nuclear PDFs at NLO—status report and review of the EPS09 results, Nucl. Phys. **A855**, 150 (2011).

[13] M. Klasen and H. Paukkunen, Nuclear parton distribution functions after the first decade of LHC data, Annu. Rev. Nucl. Part. Sci. **74**, 49 (2024).

[14] P. Duwentäster, T. Ježo, M. Klasen, K. Kovařík, A. Kusina, K. F. Muzakka, F. I. Olness, R. Ruiz, I. Schienbein, and J. Y. Yu, Impact of heavy quark and quarkonium data on nuclear gluon PDFs, Phys. Rev. D **105**, 114043 (2022).

[15] N. Derakhshanian, P. Risse, T. Jezo, M. Klasen, K. Kovarik, and A. Kusina, Determination of nuclear PDFs using Markov Chain Monte Carlo methods, arXiv:2603.13150.

[16] M. Klasen, nCTEQ global analysis of nuclear PDFs, Proc. Sci. DIS2025 (**2025**) 036 [arXiv:2510.05880].

[17] M. Klasen, Global analysis of nuclear parton distribution functions, Proc. Sci. DIS2024 (**2025**) 002 [arXiv:2407.10719].

[18] A. V. Kotikov and A. V. Lipatov, A simple model for nuclear modification of parton distribution functions, Chin. Phys. C **50**, 034109 (2026).

[19] M. R. Adams *et al.* (E665 Collaboration), Scaled energy (z) distributions of charged hadrons observed in deep inelastic muon scattering at 490-GeV from xenon and deuterium targets, Phys. Rev. D **50**, 1836 (1994).
[20] A. Airapetian *et al.* (HERMES Collaboration), Transverse momentum broadening of hadrons produced in semi-inclusive deep-inelastic scattering on nuclei, Phys. Lett. B **684**, 114 (2010).
[21] S. S. Adler *et al.* (PHENIX Collaboration), Centrality dependence of π^0 and η production at large transverse momentum in $\sqrt{s_{NN}} = 200$-GeV d + Au collisions, Phys. Rev. Lett. **98**, 172302 (2007).
[22] J. Adams *et al.* (STAR Collaboration), Pion, kaon, proton and anti-proton transverse momentum distributions from $p + p$ and d + Au collisions at $\sqrt{s_{NN}} = 200$ GeV, Phys. Lett. B **616**, 8 (2005).
[23] J. Adams *et al.* (STAR Collaboration), Identified hadron spectra at large transverse momentum in p + p and d + Au collisions at $\sqrt{s_{NN}} = 200$-GeV, Phys. Lett. B **637**, 161 (2006).
[24] O. Grebenyuk, Neutral meson production in d + Au and p + p collisions at 200 GeV in STAR, Ph.D. thesis, Utrecht University, 2007.
[25] F. Arleo (Medium-Modified Collaboration), Fragmentation functions, Eur. Phys. J. C **61**, 603 (2009).
[26] R. Baier, D. Schiff, and B. G. Zakharov, Energy loss in perturbative QCD, Annu. Rev. Nucl. Part. Sci. **50**, 37 (2000).
[27] A. Kovner and U. A. Wiedemann, Gluon radiation and parton energy loss, arXiv:hep-ph/0304151.
[28] M. Gyulassy, I. Vitev, X.-N. Wang, and B.-W. Zhang, *Jet Quenching and Radiative Energy Loss in Dense Nuclear Matter* (World Scientific, Singapore, 2004), pp. 123–191.
[29] A. Majumder, A comparative study of jet-quenching schemes, J. Phys. G **34**, S377 (2007).
[30] R. Sassot, M. Stratmann, and P. Zurita, Fragmentations functions in nuclear media, Phys. Rev. D **81**, 054001 (2010).
[31] M. Doradau, R. T. Martinez, R. Sassot, and M. Stratmann, Pion nuclear fragmentation functions revisited, Phys. Rev. D **111**, 034045 (2025).
[32] M. Alrashed, D. Anderle, Z.-B. Kang, J. Terry, and H. Xing, Three-dimensional imaging in nuclei, Phys. Rev. Lett. **129**, 242001 (2022).
[33] M. Alrashed, Z.-B. Kang, J. Terry, H. Xing, and C. Zhang, Nuclear modified transverse momentum dependent parton distribution and fragmentation functions, arXiv:2312.09226.
[34] P. C. Barry, L. Gamberg, W. Melnitchouk, E. Moffat, D. Pitonyak, A. Prokudin, and N. Sato (Jefferson Lab Angular Momentum (JAM) Collaboration), Tomography of pions and protons via transverse momentum dependent distributions, Phys. Rev. D **108**, L091504 (2023).
[35] A. Airapetian *et al.* (HERMES Collaboration), Hadronization in semi-inclusive deep-inelastic scattering on nuclei, Nucl. Phys. **B780**, 1 (2007).
[36] A. Airapetian *et al.* (HERMES Collaboration), Multidimensional study of hadronization in nuclei, Eur. Phys. J. A **47**, 113 (2011).
[37] S. Moran *et al.* (CLAS Collaboration), Measurement of charged-pion production in deep-inelastic scattering off nuclei with the CLAS detector, Phys. Rev. C **105**, 015201 (2022).
[38] F. Ceccopieri and R. Dupré, Transverse-momentum dependent model for hadronization in deep-inelastic lepton-nucleus scattering off heavy nuclei, Phys. Rev. C **110**, 045202 (2024).
[39] A. Kusina *et al.*, Impact of LHC vector boson production in heavy ion collisions on strange PDFs, Eur. Phys. J. C **80**, 968 (2020).
[40] R. Abdul Khalek, R. Gauld, T. Giani, E. R. Nocera, T. R. Rabemananjara, and J. Rojo, nNNPDF3.0: Evidence for a modified partonic structure in heavy nuclei, Eur. Phys. J. C **82**, 507 (2022).
[41] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, EPPS21: A global QCD analysis of nuclear PDFs, Eur. Phys. J. C **82**, 413 (2022).
[42] A. Airapetian *et al.* (HERMES Collaboration), Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron, Phys. Rev. D **87**, 074029 (2013).
[43] A. Airapetian *et al.* (HERMES Collaboration), Multiplicity of charged and neutral pions in deep inelastic scattering of 27.5-GeV positrons on hydrogen, Eur. Phys. J. C **21**, 599 (2001).
[44] K. Abe *et al.* (SLD Collaboration), Production of π^+, π^-, K^+, K^-, p and $\bar{\text{p}}$ in light (uds), c and b jets from Z^0 decays, Phys. Rev. D **69**, 072003 (2004).
[45] P. Nason and B. R. Webber, Scaling violation in e^+e^- fragmentation functions: QCD evolution, hadronization and heavy quark mass effects, Nucl. Phys. **B421**, 473 (1994); **B480**, 755(E) (1996).
[46] W. Furmanski and R. Petronzio, Lepton—hadron processes beyond leading order in quantum chromodynamics, Z. Phys. C **11**, 293 (1982).
[47] D. Graudenz, One particle inclusive processes in deeply inelastic lepton—nucleon scattering, Nucl. Phys. **B432**, 351 (1994).
[48] D. de Florian, M. Stratmann, and W. Vogelsang, QCD analysis of unpolarized and polarized lambda baryon production in leading and next-to-leading order, Phys. Rev. D **57**, 5811 (1998).
[49] D. de Florian and Y. Rotstein Habarnau, Polarized semi-inclusive electroweak structure functions at next-to-leading-order, Eur. Phys. J. C **73**, 2356 (2013).
[50] S. Goyal, S.-O. Moch, V. Pathak, N. Rana, and V. Ravindran, Next-to-next-to-leading order QCD corrections to semi-inclusive deep-inelastic scattering, Phys. Rev. Lett. **132**, 251902 (2024).
[51] L. Bonino, T. Gehrmann, and G. Stagnitto, Semi-inclusive deep-inelastic scattering at next-to-next-to-leading order in QCD, Phys. Rev. Lett. **132**, 251901 (2024).
[52] J. Gao, H. T. Li, H. X. Zhu, and Y. J. Zhu, Transeverse-momentum subtraction for semi-inclusive deep-inelastic scattering, arXiv:2602.06364.
[53] M. Abele, D. de Florian, and W. Vogelsang, Approximate NNLO QCD corrections to semi-inclusive DIS, Phys. Rev. D **104**, 094046 (2021).
[54] M. Abele, D. de Florian, and W. Vogelsang, Threshold resummation at N^3LL accuracy and approximate N^3LO

corrections to semi-inclusive DIS, Phys. Rev. D **106**, 014015 (2022).
[55] C.-Q. He, H. Xing, T.-Z. Yang, and H. X. Zhu, Single-inclusive hadron production in electron-positron annihilation at next-to-next-to-next-to-leading order in QCD, Phys. Rev. Lett. **135**, 101901 (2025).
[56] L. Dong, S. Fang, J. Gao, H. T. Li, D. Y. Shao, and Y. J. Zhu, Next-to-next-to-leading-order QCD corrections to hadron production in deep-inelastic scattering at finite transverse momentum, Phys. Rev. Lett. **137**, 081902 (2026).
[57] L. Dong, S. Fang, J. Gao, H. T. Li, D. Y. Shao, H. X. Zhu, and Y. J. Zhu, Two-dimensional transverse-momentum subtraction and semi-inclusive deep-inelastic scattering at N^3LO in QCD, arXiv:2603.29673.
[58] V. N. Gribov and L. N. Lipatov, Deep inelastic e p scattering in perturbation theory, Sov. J. Nucl. Phys. **15**, 438 (1972).
[59] L. N. Lipatov, The parton model and perturbation theory, Yad. Fiz. **20**, 181 (1974).
[60] G. Altarelli and G. Parisi, Asymptotic freedom in parton language, Nucl. Phys. **B126**, 298 (1977).
[61] Y. L. Dokshitzer, Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics, Sov. Phys. JETP **46**, 641 (1977).
[62] A. Mitov, S. Moch, and A. Vogt, Next-to-next-to-leading order evolution of non-singlet fragmentation functions, Phys. Lett. B **638**, 61 (2006).
[63] S. Moch and A. Vogt, On third-order timelike splitting functions and top-mediated Higgs decay into hadrons, Phys. Lett. B **659**, 290 (2008).
[64] A. A. Almasy, S. Moch, and A. Vogt, On the next-to-next-to-leading order evolution of flavour-singlet fragmentation functions, Nucl. Phys. **B854**, 133 (2012).
[65] H. Chen, T.-Z. Yang, H. X. Zhu, and Y. J. Zhu, Analytic continuation and reciprocity relation for collinear splitting in QCD, Chin. Phys. C **45**, 043101 (2021).
[66] M. A. Ebert, B. Mistlberger, and G. Vita, TMD fragmentation functions at N^3LO, J. High Energy Phys. 07 (2021) 121.
[67] M.-x. Luo, T.-Z. Yang, H. X. Zhu, and Y. J. Zhu, Unpolarized quark and gluon TMD PDFs and FFs at N^3LO, J. High Energy Phys. 06 (2021) 115.
[68] G. Curci, W. Furmanski, and R. Petronzio, Evolution of parton densities beyond leading order: The nonsinglet case, Nucl. Phys. **B175**, 27 (1980).
[69] W. Furmanski and R. Petronzio, Singlet parton densities beyond leading order, Phys. Lett. **97B**, 437 (1980).
[70] J. Gao, C. Liu, X. Shen, H. Xing, and Y. Zhao, Global analysis of fragmentation functions to charged hadrons with high-precision data from the LHC, Phys. Rev. D **110**, 114019 (2024).
[71] J. Gao, C. Liu, M. Li, X. Shen, H. Xing, Y. Zhao, and Y. Zhou, Global analysis of fragmentation functions to light neutral hadrons, Phys. Rev. D **112**, 054045 (2025).
[72] J. Gao, X. Shen, H. Xing, Y. Zhao, and B. Zhou, Fragmentation functions of charged hadrons at next-to-next-to-leading order and constraints on the proton parton distribution functions, Phys. Rev. Lett. **135**, 041902 (2025).
[73] M. Stratmann and W. Vogelsang, Next-to-leading order evolution of polarized and unpolarized fragmentation functions, Nucl. Phys. **B496**, 41 (1997).
[74] G. Salam and J. Rojo, A higher order perturbative parton evolution toolkit (HOPPET), Comput. Phys. Commun. **180**, 120 (2009).
[75] G. Salam and J. Rojo, The HOPPET NNLO parton evolution package, arXiv:0807.0198.
[76] C. Liu, X. Shen, B. Zhou, and J. Gao, Automated calculation of jet fragmentation at NLO in QCD, J. High Energy Phys. 09 (2023) 108.
[77] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, New generation of parton distributions with uncertainties from global QCD analysis, J. High Energy Phys. 07 (2002) 012.
[78] F. James and M. Roos, MINUIT—A system for function minimization and analysis of the parameter errors and correlations, Comput. Phys. Commun. **10**, 343 (1975).
[79] J. Pumplin, D. R. Stump, and W. K. Tung, Multivariate fitting and the error matrix in global analysis of data, Phys. Rev. D **65**, 014011 (2001).
[80] J. Gao, C. Liu, X. Shen, H. Xing, and Y. Zhao, Global analysis of fragmentation functions to charged hadrons with high-precision data from the LHC, Phys. Rev. D **110**, 114019 (2024).
[81] A. W. Denniston *et al.* (nCTEQ Collaboration), Modification of quark-gluon distributions in nuclei by correlated nucleon pairs, Phys. Rev. Lett. **133**, 152502 (2024).
[82] X.-N. Wang and X. Guo, Multiple parton scattering in nuclei: Parton energy loss, Nucl. Phys. **A696**, 788 (2001).
[83] S. Acharya *et al.* (ALICE Collaboration), Measurement of isolated photon-hadron correlations in $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV pp and p-Pb collisions, Phys. Rev. C **102**, 044908 (2020).