QCD analysis of pion fragmentation functions in the xFitter framework

Hamed Abdolmaleki,[1,*] Maryam Soleymaninia,[1,†] Hamzeh Khanpour,[1,2,3,‡] Simone Amoroso,[4,§] Francesco Giuli,[5,‖]
Alexander Glazov,[4,¶] Agnieszka Luszczak,[6,**] Fredrick Olness,[7,††] and Oleksandr Zenaiev[8,‡‡]

(xFITTER Developers' Team:)

[1]*School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5531, Tehran, Iran*
[2]*Department of Physics, University of Science and Technology of Mazandaran,
P.O. Box 48518-78195, Behshahr, Iran*
[3]*Department of Theoretical Physics, Maynooth University, Maynooth, Co. Kildare, Ireland*
[4]*Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany*
[5]*University of Rome Tor Vergata and INFN, Sezione di Roma 2,
Via della Ricerca Scientifica 1, 00133 Roma, Italy*
[6]*T. Kosciuszko Cracow University of Technology, PL-30-084 Cracow, Poland*
[7]*Department of Physics, Southern Methodist University, Dallas, Texas 75275-0175, USA*
[8]*CERN, CH-1211 Geneva 23, Switzerland*

We present the first open-source analysis of fragmentation functions (FFs) of charged pions (entitled IPM-xFITTER) performed at next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) accuracy in perturbative QCD using the xFITTER framework. This study incorporates a comprehensive and up-to-date set of pion production data from single-inclusive annihilation (SIA) processes, together with the most recent measurements of inclusive cross-sections of single pion by the BELLE collaboration. The determination of pion FFs along with their theoretical uncertainties is performed in the zero-mass variable-flavor number scheme (ZM-VFNS). We also present comparisons of our FFs set with recent fits from the literature. The resulting NLO and NNLO pion FFs provide valuable insights for applications in present and future high-energy analysis of pion final state processes.

I. INTRODUCTION

In the context of high-energy particle physics, Quantum chromodynamics (QCD) is a remarkably successful theory which governs the interactions of strongly interacting particles in hard scattering processes. Accurate predictions from the QCD theory are essential to provide a detailed description of the hadron structure. This information allows for incisive tests of the Standard Model (SM), and facilitates searches for new physics.

Although the nonperturbative character of QCD imposes formidable calculational challenges, perturbative QCD circumvents this difficulty by factorizing the physics into universal parton distributions functions (PDFs) and fragmentation functions (FFs) which encode the dynamical structure of the hadrons.

The PDFs and FFs must be determined phenomenologically by a global QCD analysis of a diverse collection of hard scattering measurements involving initial- and final-state hadrons.[1] The universal feature of the PDFs and FFs allows us to extract them in a global fit if we have sufficient data to constrain these distributions. Fortunately, we have extensive and precise measurements from a variety

[*]Abdolmaleki@ipm.ir
[†]Maryam_Soleymaninia@ipm.ir
[‡]Hamzeh.Khanpour@cern.ch
[§]Simone.Amoroso@cern.ch
[‖]francesco.giuli@roma2.infn.it
[¶]alexandre.glazov@desy.de
[**]agnieszka.luszczak@desy.de
[††]olness@smu.edu
[‡‡]oleksandr.zenaiev@cern.ch

[1]Note that in recent years, significant advancements in Lattice QCD have enabled these calculations to contribute to constraining the PDFs and FFs in global analyses. For recent reviews, see Refs. [1–3].

of facilities including PETRA, SLC, LEP, BELLE, *BABAR*, Tevatron, HERA, and the Large Hadron Collider (LHC).

In recent years, a large number of studies have performed global QCD analyses of experimental data to obtain precise phenomenological parametrizations of collinear (or integrated) FFs [4–7], as well as transverse momentum dependent FFs (TMD FFs) [8–10]. Additionally, a recent analysis included the effects of medium modified fragmentation functions (nFFs) [11]. Furthermore, a simultaneous QCD analysis of both PDFs and FFs has been performed by the JAM Collaboration using inclusive and semi-inclusive deep inelastic scattering (SIDIS), Drell-Yan, and single inclusive electron-positron annihilation (SIA) experimental data [6,12–14].

Recent determinations of the pion FFs include results by several groups such as HKNS [15], NNFF [4], JAM [6,13], SGKS [7], SGK [16], and other studies available in the literature including KKP [17,18], BKK [19], and KLC [20] These analyses differ in their choices of input functional form for the pion FFs, the fitting methodology, the method of error estimation, the input experimental datasets, and the order of the QCD perturbative expansion in the theory predictions.

The primary objective of the present study is to perform an extraction of the collinear FFs for the pion (in a vacuum) using an exhaustive set of SIA data along with the very recent pion production measurements by the BELLE collaboration (BELLE20) [21]. To help benchmark the impact of this new dataset we will be especially interested to compare the BELLE20 data with the previous 2013 BELLE measurement (BELLE13) [22] as well as with the *BABAR* [23] data (cf., Table I). Additionally, we will compare our analysis at NLO and NNLO accuracy.

xFITTER is an open-source package that provides a framework for the determination of the PDFs, and now FFs, for many different kinds of analyses in QCD. The current xFITTER version 2.1 offers an expanded set of tools and options, and incorporates experimental data from a wide range of experiments [24,25]. xFITTER can analyze this data up to next-to-next-to-leading-order (NNLO) in perturbation theory with a variety of theoretical calculations. While primarily based on the collinear factorization foundation, xFITTER also provides facilities for fits of dipole models and transverse-momentum dependent (TMD) PDFs.

xFITTER has been used for a wide range of analyses involving different processes; examples include: pion PDFs [26]; nuclear PDFs [27]; and as well as various other studies [24,25,28,29,30]. This work further extends the utility and capabilities of xFITTER to the case of unpolarized collinear FFs of hadrons analysis. The pion FFs analysis presented in this paper represents the first step of a broader program to demonstrate this open-source xFITTER tool which can serve as the basis for future analyses of both PDFs and FFs.

The structure of this paper is as follows. Section II reviews the theoretical formalism for e^+e^- annihilation into hadrons. The analysis framework and methodology for the global QCD analysis is presented in Sec. III. Here we also describe our methodology including our parametrization and the treatment of the FFs uncertainties. The SIA experimental datasets analyzed in this study are summarized in Sec. IV. Section V presents the details of our results, as well as comparisons other FFs available in the literature. Additionally, we discuss the impact of the new BELLE20 data on the extracted pion FFs and compare with results from the literature. Finally, we summarize our conclusions in Sec. VI and consider future applications that may prove fruitful for the xFITTER framework.

To facilitate further investigations of the pion FFs, both the FF grids and the code are publicly available on the xFITTER webpage (`xfitter.org`).

This project is a collaborative effort between the Institute for Research in Fundamental Sciences (IPM) and xFITTER; hence, we shall identify our fit results as IPM-xFITTER, or just IPMX for short.

II. THEORETICAL FRAMEWORK

The cross section for charged hadron production in an electron-positron annihilation process $(e^+e^- \to h^\pm X)$ is typically measured as a function of the scaling variable $z = 2E_h/\sqrt{s}$ at a given center of mass energy of $\sqrt{s} = Q$ with hadron energy E_h. The differential cross section is related to the hadronic fragmentation function $F_2^{h^\pm}$ as[2]

$$\frac{d\sigma^{h^\pm}}{dz}(z, Q) = \sigma_0 F_2^{h^\pm}(z, Q), \qquad (1)$$

where $\sigma_0 = 4\pi\alpha^2/Q^2$ and α is the electromagnetic coupling.

The factorization theorem allows us to write the nonperturbative hadronic fragmentation function $F_2^{h^\pm}$ as a convolution of a perturbative coefficient function C_i and a nonperturbative partonic fragmentation function $D_i^{h^\pm}$ given by: [31–33]

$$F_2^{h^\pm}(z, Q) = \sum_i C_i(z, \alpha_s(Q)) \otimes D_i^{h^\pm}(z, Q), \qquad (2)$$

where we sum over parton flavors $i = q, \bar{q}, g$. The coefficient functions C_i have been calculated up to the NNLO accuracy in the $\overline{MS}$ scheme [34,35]. The nonperturbative partonic FFs $D_i^{h^\pm}$ are universal and represent the number density for a hadron of type $h^\pm$ from parton i with momentum fraction z at scale Q. It is the universal property

---

[2]Here, we will follow the notation of Ref. [4], and the subscript on $F_2^{h^\pm}$ suggest an analogy with the F_2 DIS hadronic structure function.

of the FFs which will allow us to extract these quantities by parametrizing their functional form and fitting to experimental data.

To simplify the expansion of the hard scattering cross section, we choose the renormalization scale μ_R and the factorization scale μ_F equal to the center-of-mass energy[3]; thus, we have $\mu_R = \mu_F = \sqrt{s} \equiv Q$.

The scale dependence of the partonic FFs is described by the DGLAP evolution equations, [36–39]

$$\frac{dD_i^{h^\pm}(z, Q)}{d\ln(Q^2)} = [P_{ij} \otimes D_j^{h^\pm}](z, Q), \qquad (3)$$

where P_{ij} are the perturbative timelike splitting functions, and the convolution integral $\otimes$ is

$$[P \otimes D](z) = \int_z^1 \frac{dy}{y} P(y) D\left(\frac{z}{y}\right). \qquad (4)$$

We solve the integro-differential DGLAP equations directly in z space using the APFEL package [40] which provides NLO and NNLO accuracy.

Now that we have outlined the key elements of the SIA cross section calculation, we next examine the framework for our analysis, including the parametrization of the nonperturbative FFs.

III. ANALYSIS FRAMEWORK

We will obtain the fragmentation functions (FFs) by parametrizing their functional form in z and then performing a fit by minimizing a χ^2 function in comparison with experimental data. In the following, we detail the analysis framework including the parametrization form, the fitting procedure, and the uncertainty analysis.

A. FFs parametrization and assumptions

We parametrize the z dependence of the FFs at an initial scale $Q_0 = 5$ GeV which keeps us above m_b, and use the DGLAP equations to evolve to arbitrary Q scale. The flexible parametric form we use is

$$D_i^{\pi^\pm}(z, Q_0) = \frac{\mathcal{N}_i z^{\alpha_i}(1 - z)^{\beta_i}[1 + \gamma_i(1 - z)^{\delta_i}]}{B[2 + \alpha_i, \beta_i + 1] + \gamma_i B[2 + \alpha_i, \beta_i + \delta_i + 1]}, \qquad (5)$$

which has (maximally) five free parameters $\{\mathcal{N}_i, \alpha_i, \beta_i, \gamma_i, \delta_i\}$ per parton flavor. Here, $B[a, b]$ is the Euler beta function. For the charged pion FFs, we fit the flavor combinations $i = u^+, d^+, s^+, c^+, b^+$ and g. The beta

---

[3]To be more precise, in Eq. (2) the $\alpha_S(\mu_R)$ depends on the renormalization scale μ_R, and the partonic fragmentation function $D_i^{h^\pm}(z, \mu_F)$ depends on the factorizaton scale μ_F.

functions in the denominator of Eq. (5) simply ensures $\int_0^1 dz z D_i = \mathcal{N}_i$.

There are a number of constraints we can impose to reduce the number of free parameters of the fit. From the energy sum rule, we have the relation:

$$\sum_h \int_0^1 dz z D_i^h = \sum_h \mathcal{N}_i^h = 1, \qquad (6)$$

where h sums over all possible produced hadrons. For the pion FFs ($h = \pi^\pm$) this relation provides only an upper bound, but if we expect the lighter pions carry most of the parton momentum, then we have

$$\mathcal{N}_i^{\pi^\pm} < 1, \qquad (7)$$

where $i = g, q, \bar{q}$. Note, in Table II we report $\mathcal{N}_{u^+}$ where $u^+ = u + \bar{u}$, hence the limit on this quantity is $\mathcal{N}_{u^+} < 2$.

Thus, we will use four shape parameters $\{\alpha_i, \beta_i, \gamma_i, \delta_i\}$ together with the normalization parameter $\mathcal{N}_i$ to fit our FFs.

For the π^+ FFs, we assume isospin symmetry for the favored $(u, \bar{d})$ and unfavored $(\bar{u}, d)$ components [4,14,15]:

$$D_u^{\pi^+} = D_{\bar{d}}^{\pi^+}$$
$$D_{\bar{u}}^{\pi^+} = D_d^{\pi^+}. \qquad (8)$$

We can also use charge conjugation to relate the above π^+ FFs to the π^- FFs:

$$D_i^{\pi^+} = D_i^{\pi^-}, \qquad i = u^+, d^+, s^+, c^+, b^+, g. \qquad (9)$$

Additionally, for the s^+, c^+ and g flavors we impose the condition $\gamma_i = \delta_i = 0$.

We thus have a total of 19 free parameters to be determined by a standard χ^2 minimization strategy. We now discuss the details of the fit.

B. Fitting procedure

By comparing the theoretical and experimental measurements from our selected datasets, we determine the optimal values of the 19 independent fit parameters by minimizing the χ^2 function as implemented in the xFITTER framework: [24,41,42]

$$\chi^2 = \sum_i \frac{[d_i - t_i(1 - \sum_\alpha b_\alpha \delta_\alpha^i)]^2}{\delta_{i,\text{unc}}^2 t_i^2 + \delta_{i,\text{stat}}^2 d_i t_i} + \sum_\alpha b_\alpha^2, \qquad (10)$$

where the sum on i is over the number of data points. Here t_i is the theory prediction, d_i is the data measurement, b_α are the nuisance parameters, and $\delta_{i,\text{stat}}$ and $\delta_{i,\text{unc}}$ are the statistical and uncorrelated systematic uncertainties.

In general, for this analysis we will add the correlated and uncorrelated uncertainties in quadrature; while this

TABLE I.   The Single Inclusive electron-positron Annihilation (SIA) datasets used in the pion FFs analysis. The values of χ^2 per N data points for the individual SIA experiments are shown. The z range for each experiment is displayed in Fig. 8. The measured observable is also listed where $\sqrt{s}$ is the total CMS energy, $\beta = p_h/E_h$, and $z = 2E_h/\sqrt{s}$.

| | | | χ^2/number of points | | | | | |
Observable	Experiment	$\sqrt{s}$ [GeV]	Fit A (NLO)	Fit A (NNLO)	Fit B (NNLO)	Fit C (NNLO)	Fit D (NNLO)	Fit E (NNLO)
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}$	SLD	91.20	57/34	41/34	41/34	48/34	39/34	45/34
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}\big|_{\text{uds}}$	SLD$_{\text{uds}}$	91.20	66/34	52/34	56/34	44/34	43/34	45/34
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}\big|_{c}$	SLD$_c$	91.20	35/34	33/34	32/34	32/34	32/34	32/34
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}\big|_{b}$	SLD$_b$	91.20	25/34	24/34	24/34	24/34	23/34	24/34
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dp_h}$	OPAL	91.20	42/24	41/24	41/24	39/24	39/24	39/24
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dp_h}$	DELPHI	91.20	37/21	41/21	41/21	44/21	44/21	43/21
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dp_h}\big|_{\text{uds}}$	DELPH$_{\text{uds}}$	91.20	25/21	27/21	26/21	30/21	31/21	30/21
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dp_h}\big|_{b}$	DELPH$_b$	91.20	20/21	20/21	21/21	19/21	20/21	19/21
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}$	ALEPH	91.20	21/23	14/23	14/23	11/23	11/23	12/23
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}$	TASSO44	44.00	15/6	17/6	15/6	18/6	16/6	18/6
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}$	TASSO34	34.00	6.8/9	8.0/9	6.8/9	9.3/9	7.3/9	8.3/9
$\frac{1}{\beta\sigma_{\text{tot}}}\frac{d\sigma^h}{dz}$	TPC	29.00	6.3/13	11/13	11/13	11/13	7.1/13	9.2/13
$\frac{s}{\beta}\frac{d\sigma^h}{dz}$	TASSO22	22.00	5.7/8	5.5/8	5.6/8	6.1/8	5.9/8	5.8/8
$\frac{s}{\beta}\frac{d\sigma^h}{dz}$	TASSO14	14.00	11/9	11/9	11/9	9.9/9	11/9	9.8/9
$\frac{s}{\beta}\frac{d\sigma^h}{dz}$	TASSO12	12.00	1.4/4	1.4/4	1.3/4	0.96/4	1.4/4	1.1/4
$\frac{1}{\sigma_{\text{tot}}}\frac{d\sigma^h}{dp_h}$	*BABAR*	10.52	71/40	53/40	77/40	$\cdots$	$\cdots$	33/37
$\frac{d\sigma^h}{dz}$	BELLE13	10.54	21/70	14/70	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\frac{d\sigma^h}{dz}$	BELLE20	10.58	$\cdots$	$\cdots$	82/32	32/32	9.2/28	17/28
Correlated χ^2			11	9.4	8.4	16	9.4	12
Log penalty χ^2			+4.2	+3.0	+4.2	+7.7	+5.6	+6.8
Total χ^2/dof			480/386	427/386	518/348	404/308	357/304	410/341

allows a generous band of uncertainties on our FFs, we still find it challenging to describe all the data as we detail in Sec. V. In the case where an individual experiment provides multiple datasets (e.g., TASSO, DELPHI, and SLD *c.f.*, Table I), the overall normalization of these data points is correlated; we take this into account as a systematic error. The effect of this correlated systematic normalization will be evident in Sec. V B 5 where we display both the theory prediction as well as the "theory + shifts" showing the impact of the correlated normalization.

C. Uncertainty analysis

In addition to a central value of the FFs and PDFs, it is important to also evaluate the uncertainty band of these quantities. For the present study, we will apply the Hessian method to estimate the FFs uncertainty band [24,43,44]. The Hessian approach is based on a quadratic expansion of the χ^2 function about its global minimum and can be expressed as:

$$\chi^2 = \chi_0^2 + \Delta\chi^2 = \chi_0^2 + \sum_{i,j} H_{ij} y_i y_j, \qquad (11)$$

where i, j run over the 19 free parameters of the fit. Here, χ_0^2 is the global minimum, $\Delta\chi^2$ is the displacement from the minimum, and y_i are the parameter shifts about the minimum.

The Hessian matrix is constructed from the second derivatives of χ^2 and defined as $H_{ij} = \frac{1}{2}(\partial^2\chi^2)/(\partial y_j \partial y_j)$. The Hessian is a symmetric $n \times n$ matrix where $n = 19$ is the number of free parameters, and it will have n orthogonal eigenvectors v_i with eigenvalues ϵ_i.

We find it convenient to transform from the y_i parameter basis to the v_i eigenvector basis as the Hessian is diagonal in this basis. Additionally, we scale the eigenvectors by the square-root of their eigenvalue to introduce a scaled eigenvector $z_i = \sqrt{\epsilon_i} v_i$. Thus, we can express the χ^2 shift as:

$$\Delta\chi^2 = \sum_{i,j} H_{i,j} y_i y_j = \sum_i \epsilon_i v_i^2 = \sum_i z_i^2. \qquad (12)$$

Finally, we can relate the original fitting parameters y_i to the scaled eigenvectors z_i via the equation:

$$y_i = \sum_j V_{ij}^{-1} \frac{z_j}{\sqrt{\epsilon_j}}, \qquad (13)$$

where V_{ij} is the $n \times n$ matrix composed of the eigenvectors v_i which diagonalizes the Hessian as in $V_{ii'} H_{i'j'} V_{j'j}^{-1} = \delta_{ij} \epsilon_i$.

An important consideration when constructing the FFs uncertainty range is to decide upon the acceptable tolerance range $T = \Delta\chi^2$ over which we will allow the fit parameters to vary. For a simple case of a single parameter, the confidence level is 68% for $\Delta\chi^2 = 1$. The present case of the pion FFs involves combining large datasets across many different measurements with uncertainties from both experimental and theoretical sources that can only be approximated. Thus, determination of a $\Delta\chi^2$ tolerance criteria is not straightforward. For guidance on the choice of tolerance, we reviewed the analyses used in Refs. [15,41,45–47]. For the present analysis we will use $\Delta\chi^2 = 20$ which was chosen based on a number of factors. We shall present the details in the following sections, but the $\Delta\chi^2 = 20$ tolerance criteria gives us a reasonable overlap of the bands for our comparison of NLO and NNLO fits, our fits comparing BELLE13 and BELLE20, and also our fits using different scale choices. With a tolerance of $\Delta\chi^2 = 1$, the bands for the above comparisons would not overlap, and would clearly underestimate the true uncertainties.

IV. EXPERIMENTAL OBSERVABLES

The datasets used in the present pion FFs analyses are based on electron-positron SIA cross sections for pion ($\pi^\pm$) production. We incorporate SIA experimental data from an range of experiments including SLAC (*BABAR* [23], TPC [48], SLD [49]), CERN (ALEPH [50], OPAL [51] DELPHI [52]), KEK (BELLE [21,22]), and DESY (TASSO [53–55]). These datasets contain inclusive and flavor-tagged measurements of pion scaled energy (z) corresponding to the sum of light quarks (u, d, s) as well as individual charm (c) and bottom (b) quarks.

We list all the datasets used in our fits in Table I, and indicate the experiment and the observable. The bulk of our data points are measured by the LEP and SLD Collaborations ($\sqrt{s} = M_Z$) and the B-factory experiments ($\sqrt{s} \simeq 10$). The remaining measurements range across intermediate scales of energy.

In a manner similar to Ref. [4], we will impose z cuts on the datasets to avoid nonperturbative effects not described by our calculations; such effects could arise from contributions beyond the order of our NNLO calculation, or from nonfactorizable (e.g., higher twist) contributions. Specifically, for the experiments with $\sqrt{s} = M_Z$ we take data in the range $z \in [0.02, 0.9]$, and for all other sets we take $z \in [0.075, 0.9]$.

For all our fits we used $\alpha_s(M_Z^2) = 0.1185$ which is compatible with the world average data [47]. We also use $M_Z = 91.1876$ GeV, $\sin^2\theta_W = 0.231$, $Q_0 = 5$ GeV, and pole masses for the heavy quarks of $m_c = 1.43$ GeV and $m_b = 4.5$ GeV.

We now describe the specific features of the B-factory measurements from *BABAR* and BELLE which are taken at the lower scale $\sqrt{s} \simeq 10$ GeV. As this energy is below the B-meson pair production threshold, the effects of bottom quark production are not included for this data. These datasets are important for the fit as they have high precision and extend the momentum fraction z close to one.

A. Updated BELLE Data

The BELLE Collaboration has recently released a new measurement (BELLE20) [21] which we will compare to the previous dataset (BELLE13) [22].

The new dataset (BELLE20) has removed the backgrounds from Υ decays, τ production and two-photon processes, and also applied an updated initial-state radiative (ISR) correction. Additionally, the updated BELLE20 data contains all systematic uncertainties separated into correlated and uncorrelated contributions. For the present analysis, these uncertainties are added in quadrature; even with this generous error budget, we will find this data challenging to describe with our theoretical calculations. Furthermore, the number of equidistant bins in the range $0.075 < z < 0.9$ has decreased from 70 in BELLE13 measurements to 32 in BELLE20.

In the case of the BELLE13 experiment, the measurements have been provided in the form $d\sigma^{H^\pm}/dz$. To properly treat this dataset consistently with the other measurements included in our analysis, we multiply the BELLE13 measurements [22] by a correction factor as explained in Ref. [4] to compensate for the kinematic cut on radiative photon events which was applied to the BELLE13 data. In contrast, no such correction is needed for the BELLE20 data.

V. RESULTS

We now present the results of our fit for the charged pion FFs. Additionally, we will also compare our resulting FFs with recent results in the literature.

A. Pion fragmentation function fits

1. IPM-xFITTER fits

As we studied the influence of the various datasets, we found it especially useful to compare the effects of B-factory datasets from BELLE and *BABAR*. Thus, we perform five sets of fits which we summarize in the following. Note, we include all the data in Table I *except* what is indicated below.

Fit A: This fit focuses on the impact of the BELLE13 dataset. Thus, we *exclude* the BELLE20 data for a total of 386 points.

Fit B: This fit focuses on the impact of the BELLE20 dataset. Thus, we *exclude* the BELLE13 data for a total of 348 points.

Fit C: This fit focuses on the impact of the BELLE20 data *without* the *BABAR* set. Thus, we *exclude* the *BABAR* and BELLE13 data for a total of 308 points.

Fit D: This fit focuses on the impact of cutting the low-z BELLE20 data. Thus, we *exclude* the *BABAR* and BELLE13 data, and impose a $z > 0.2$ cut on the BELLE20 data, for a total of 304 points.

Fit E: This fit focuses on the impact of the BELLE20 and *BABAR* sets with cuts imposed to remove low-z data. Thus, we *exclude* the BELLE13 data, impose a $z > 0.2$ cut on the BELLE20 data and impose a $z > 0.1$ cut on the *BABAR* data for a total of 341 points.

Thus, in the above, Fit A will reflect the influence of the updated BELLE13 and Fit B will reflect BELLE20. We can discern the impact of the *BABAR* data by comparing Fits C and D with Fit E. And finally, we can see the impact of the low z cuts by comparing Fits D and E.

2. Fit parameters

In Table II we present the values of the 19 fitted parameters, Eq. (5), and their uncertainties for our fits as computed by Minuit [41].

Examining the relatively small uncertainties on the shape parameters (with the possible exception of α_g and α_{s^+}), this suggests that these parameters appear to be relatively well constrained by the SIA data.

Similarly, Table I displays the χ^2 values for each individual experiment and the number of degrees of freedom, as well as the totals. The normalizations are treated as a correlated systematic and are also reported [24].

3. NLO and NNLO comparison

Since xFITTER can perform both NLO and NNLO calculations, we will compare these for the case of Fit A. The quality of our fits can be surmised by examining the total χ^2 and χ^2 per degree of freedom (χ^2/dof) as displayed in Table I and Fig. 1. For Fit A we observe that the NNLO fit yields a substantially improved χ^2/dof of $427/386$ for NNLO *vs* $480/386$ for NLO representing an improvement of ~11%.

If we look in more detail, Table I also displays the χ^2 values for each individual experiment. While most of the datasets show improvement of the NNLO compared to the NLO results, there are some observables where the individual χ^2 does slightly increase such as the ones measured by the TASSO, TPC and DELPHI experiments; these increases are small compared to the overall improvement of the fit from $\chi^2/\text{dof} \sim 1.24$ to 1.11. A similar behavior is observed in other analyses such as Ref. [4].

TABLE II. The best fit parameters with uncertainties obtained for the pion FFs fits at our initial scale $Q_0^2 = 25$ GeV². For all fits we use $\alpha_s = 0.1185$.

| Parameters	Fit A (NLO)	Fit A (NNLO)	Fit B (NNLO)	Fit C (NNLO)	Fit D (NNLO)	Fit E (NNLO)
N_g	0.559 ± 0.021	0.564 ± 0.068	0.43 ± 0.13	0.346 ± 0.092	0.62 ± 0.14	0.448 ± 0.080
α_g	0.91 ± 0.25	0.22 ± 0.80	0.8 ± 3.0	0.26 ± 0.80	-0.27 ± 0.43	-0.10 ± 0.54
β_g	19.9 ± 2.6	17.3 ± 8.4	24 ± 32	9.0 ± 4.0	8.5 ± 3.2	9.6 ± 4.0
N_{u^+}	1.107 ± 0.021	1.075 ± 0.093	1.036 ± 0.082	1.399 ± 0.100	1.20 ± 0.12	1.276 ± 0.095
α_{u^+}	-0.648 ± 0.052	-0.69 ± 0.19	-0.87 ± 0.15	-1.098 ± 0.088	-0.99 ± 0.12	-1.02 ± 0.11
β_{u^+}	1.730 ± 0.044	1.84 ± 0.20	1.73 ± 0.10	1.61 ± 0.11	1.61 ± 0.12	1.64 ± 0.11
γ_{u^+}	7.37 ± 0.91	7.4 ± 2.6	4.0 ± 2.3	4.1 ± 1.0	4.3 ± 1.2	4.4 ± 1.1
δ_{u^+}	5.27 ± 0.29	6.1 ± 1.4	4.7 ± 1.1	4.34 ± 0.81	3.87 ± 0.79	4.58 ± 0.76
N_{s^+}	0.348 ± 0.046	0.45 ± 0.14	0.66 ± 0.19	0.22 ± 0.12	0.23 ± 0.10	0.29 ± 0.13
α_{s^+}	1.03 ± 0.34	0.75 ± 0.77	-0.11 ± 0.64	2.0 ± 1.7	1.9 ± 1.3	1.4 ± 1.1
β_{s^+}	9.89 ± 0.83	8.5 ± 2.2	7.8 ± 3.6	10.7 ± 3.7	11.3 ± 4.2	9.7 ± 2.7
N_{c^+}	0.754 ± 0.013	0.776 ± 0.055	0.827 ± 0.057	0.926 ± 0.086	0.777 ± 0.088	0.856 ± 0.070
α_{c^+}	-0.843 ± 0.037	-0.90 ± 0.15	-0.93 ± 0.13	-1.23 ± 0.13	-1.09 ± 0.15	-1.14 ± 0.14
β_{c^+}	4.99 ± 0.18	5.03 ± 0.65	5.13 ± 0.62	4.11 ± 0.58	4.33 ± 0.66	4.32 ± 0.57
N_{b^+}	0.711 ± 0.013	0.731 ± 0.047	0.778 ± 0.054	0.833 ± 0.054	0.722 ± 0.068	0.788 ± 0.051
α_{b^+}	-0.432 ± 0.050	-0.47 ± 0.19	-0.49 ± 0.16	-0.75 ± 0.13	-0.65 ± 0.17	-0.69 ± 0.15
β_{b^+}	4.05 ± 0.24	4.08 ± 0.78	4.03 ± 0.84	3.89 ± 0.68	3.90 ± 0.74	3.94 ± 0.69
γ_{b^+}	12.5 ± 1.7	12.7 ± 5.2	13.9 ± 6.8	8.9 ± 3.3	9.3 ± 3.8	9.1 ± 3.3
δ_{b^+}	8.12 ± 0.42	8.4 ± 1.6	8.4 ± 1.7	8.7 ± 1.7	8.5 ± 1.8	8.5 ± 1.7
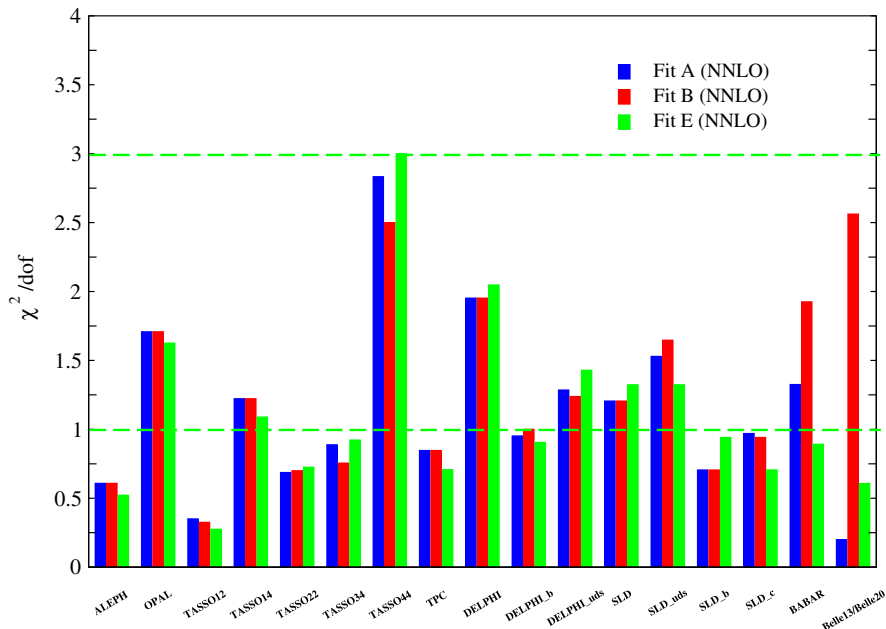
FIG. 1.   The χ^2/dof for each individual experiment for Fits A, B, and E. The data used in the Fit A are in blue, Fit B is red, and Fit E is green. The numerical results for all fits (A,B,C,D,E) are listed in Table I. For reference, (green) guide lines are shown for a χ^2/dof of 1 and 3.

Finally, if we examine the uncertainty bands for the FFs of Fig. 2, the *a posteriori* fact the NLO and NNLO bands overlap can be taken as an indication that perturbative uncertainties are under control and this additionally suggests that our choice of tolerance T is reasonable.

While we do not explicitly show details of the other NLO fits, the results are similar. Hence, the inclusion of the higher-order QCD corrections noticeably improves the quality of our fits, and this result is in agreement with other analyses in the literature [4,16,56].

Since the NNLO yields improved results over the NLO calculation, we will focus only on the NNLO fits for the following analyses.

B. Comparison of datasets

1. BELLE Data

To examine how well our fits describe the experimental data, we start by comparing Fit A which includes the BELLE13 data and Fit B which includes the recent BELLE20 data. This information is contained both in Table I and also graphically in Fig. 1.

Examining Table I, we see at NNLO order Fit A yields a χ^2/dof of $427/386 \sim 1.11$ as compared to the NNLO Fit B with $518/348 \sim 1.49$.

If we look at the χ^2 values for the individual experiments, we see that most of the values are comparable between the Fit A and Fit B results, with the exception of the *BABAR* data which contributes an extra ~ 24 units of χ^2 to Fit B. The remainder of the difference is of course due to

the BELLE data itself where the BELLE13 data (Fit A) gives $14/70 \sim 0.2$ *vs* the BELLE20 data (Fit B) with $82/32 \sim 2.6$.

In Fig. 3 we plot the FFs for both Fits A and B. Note that in our framework, the u^+ and d^+ distributions will be equivalent; hence, we display u^+ as well as the combination $d^+ + s^+$. We also display the uncertainty bands for the FFs computed with our tolerance of $T = 20$. We see that the $u^+ = d^+$, charm and bottom distributions show only moderate difference between the two fits, and Fit B yields a slightly softer distribution for u^+ and d^+. In contrast, the strange and gluon show larger difference which reflects, in part, their larger uncertainties. Compared to Fit A, Fit B moderately increases s^+, c^+, b^+ while decreasing g.

2. BELLE and BABAR comparison

Comparing Fits A and B above, we observed that the inclusion of the updated BELLE20 data caused a substantial increase in the *BABAR* χ^2. To explore if this is primarily a tension between the BELLE20 and *BABAR* datasets, we introduce Fit C which fits BELLE20 but excludes the *BABAR* data.

Examining the results of Fit C in Table I we see the quality of the BELLE20 data χ^2 improves from $82/32$ (Fit B) to $32/32$ (Fit C), and the overall χ^2 improves from $518/348 \sim 1.49$ to $404/308 \sim 1.31$.

Additionally, in Fig. 4 we display the comparisons of our A, B, and C fits with the data for BELLE13, BELLE20 and *BABAR*. For BELLE13 (Fit A), the theoretical predictions are entirely consistent with the experimental data, but this is
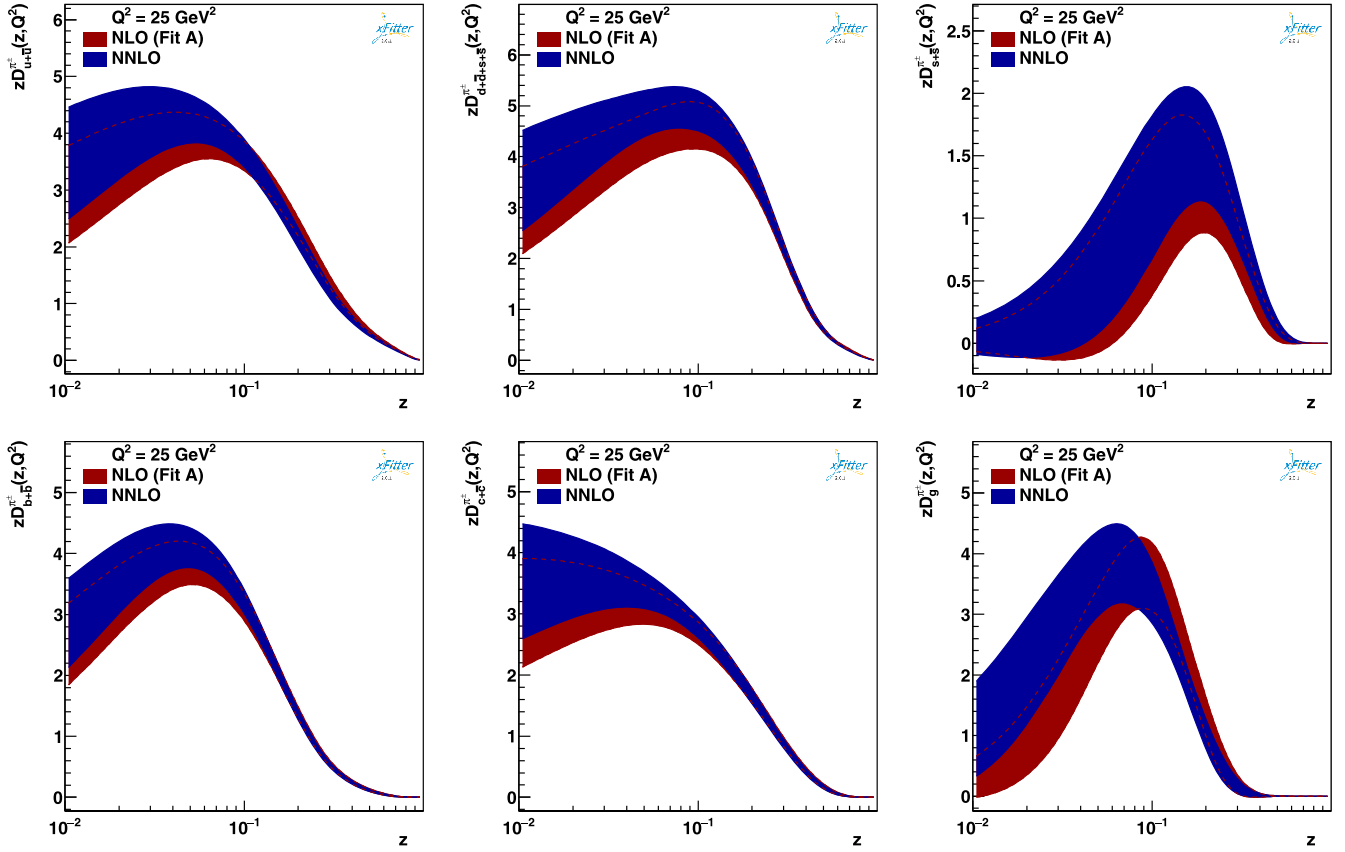
FIG. 2. A comparison of the NLO and NNLO results for Fit A at $Q_0^2 = 25$ GeV^2. As the results for up and down are identical, we choose to also display the $d^+ + s^+$ combination. Note that we are plotting the sum $\pi^+ + \pi^-$.

partly due to the larger uncertainties; note the vertical scale of the "theory/data" plot.

Comparing BELLE20 and *BABAR*, we see that the fits yield a good description of the data with the exception of the low-z region. Here, the BELLE20 and *BABAR* datasets appear to pull the fit in opposite directions; BELLE20 is above the theory, and *BABAR* is below the theory.

If we compare Fit B (with *BABAR*) to Fit C (without *BABAR*) we notice Fit C shifts toward the data throughout the z range, including the lower z region, thus reducing the χ^2 for BELLE20 from 82/32 (Fit B) to 32/32 (Fit C).

3. Low-z cuts

The deviations between theory and data in the low-z region, as observed in Fig. 4, warrants a closer examination. Thus, we introduce Fit D which imposes a $z > 0.2$ cut on the BELLE20 data again does not include the *BABAR* data. As summarized in Table I, the BELLE20 dataset is reduced from 32 to 28 points with a significantly improved χ^2 of 9.2/28, and the total χ^2 for Fit D improves to 357/304 = 1.17.

The encouraging results of imposing the low-z cut leads us to try Fit E which includes both the BELLE20 and *BABAR* data, but with low-z cuts imposed. Specifically, for

BELLE we impose $z > 0.2$ which removes 4 points, and for *BABAR* we impose $z > 0.1$ which removes 3 points. (Note, we could impose a more stringent cut of $z > 0.2$ on both datasets, but we shall find this more conservative choice already yields a notable improvement.)

We note that other analyses in the literature have also imposed kinematic cuts on the low z data. For example, the JAM19 focus was on SIDIS in the region of larger z; hence, they imposed a cut of $z \gtrsim 0.2$ on their datasets. Similarly, NNFF1 used a lower kinematic cut of $z_{\min} = 0.02$ for $Q = M_Z$ and 0.075 for $Q < M_Z$.

The result of these low-z cuts is dramatic. We obtain an improved 17/28 for BELLE20, 33/37 for *BABAR*, and a good $\chi^2 = 410/341 \sim 1.20$ for the total dataset of Fit E.

To see the impact of these fits in more detail, Fig. 5 displays the comparisons of our fits with the data for BELLE20 and *BABAR*. Examining the *BABAR* results, we see that by cutting out the lowest 3 points in z, the theoretical predictions are generally contained within the band of the experimental uncertainty.

Similarly, for BELLE20 we see that the predictions also are generally contained within the band of the experimental uncertainty. We also see that including the *BABAR* data (Fit E) increases the theory curve in the low-z region, consistent with the effects observed in Fig. 4.
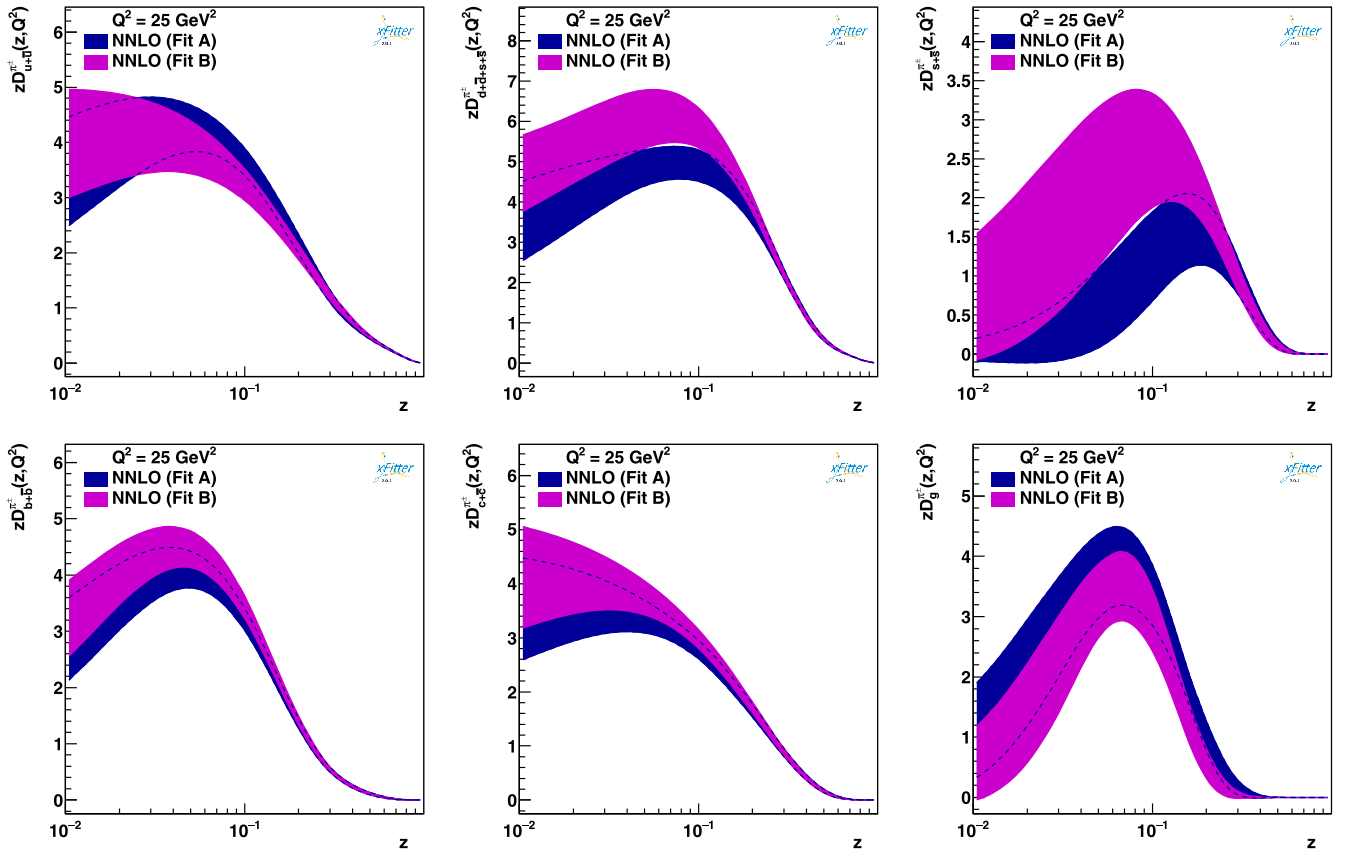
FIG. 3. A comparison of the NNLO results for Fit A (with BELLE13) and Fit B (with BELLE20) analyses at $Q_0^2 = 25$ GeV^2.

To see the impact of the low-z cuts on the FFs, in Fig. 6 we display the results of Fit B (no z cuts) and Fit E (with z cuts). The difference is dramatic and (with the exception of the heavy charm and bottom quarks) the uncertainty bands do not overlap throughout much of the z range.

While our tolerance criteria of $T = 20$ gave reasonable results for the comparisons of the NLO vs NNLO results of Fig. 2 and the comparison of BELLE13 and BELLE20 of

Figs. 3 and 6 underscores the fact that this does not represent the uncertainty due to the choice of data in the low z region; hence, this must be separately accounted for when computing the maximal FF uncertainty.

Regarding the impact of the z cuts on the shape of FFs, we see that this generally increases the FFs for up, down, charm and bottom while decreasing the strange and gluon FFs (which have comparably larger uncertainties).
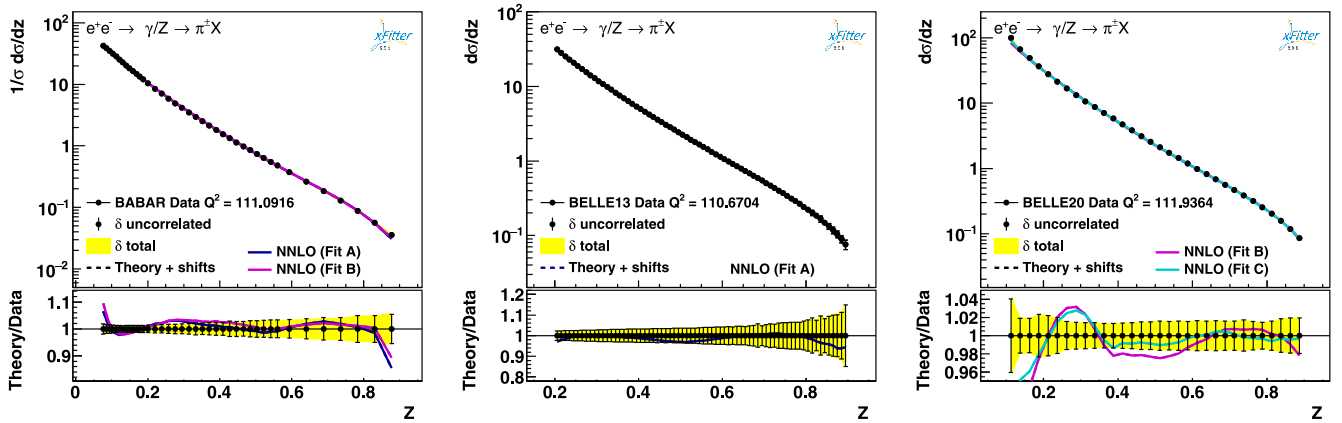
FIG. 4. A comparison of BELLE13, BELLE20, and *BABAR* for Fits A, B, and C. (Not all datasets are included in each fit.)
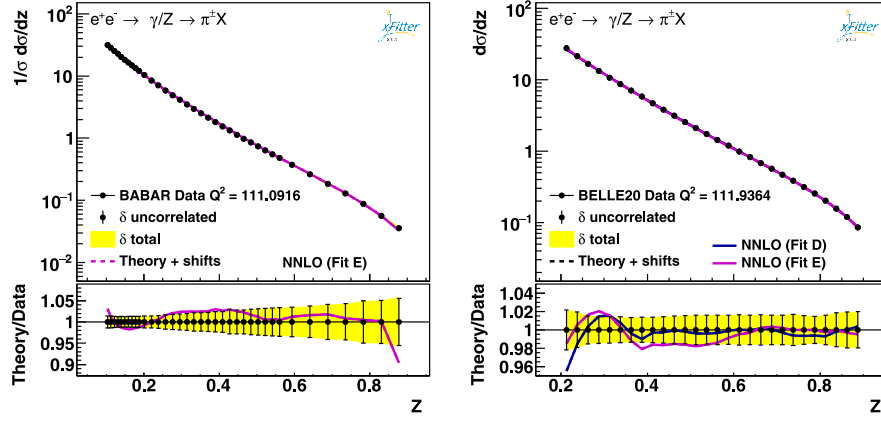
FIG. 5.   A comparison of BELLE20 and *BABAR* for Fits D and E. (*BABAR* is not included in Fit D.)

4. Scale variation

Since the results presented thus far have imposed the scale choice $\mu_F = \mu_R = Q$, we also want to study the effect of varying the perturbative scales, in part, to see if this could have any significant effect on the partonic FF $zD^{\pi^\pm}(z, Q)$ in the low-z region. In Fig. 7, we display the resulting FFs for Fit E with their uncertainties for the

variation of the scale choice as a function of z. We start with a factorization scale of $\mu = Q = 10$ GeV, and vary this up and down by a factor of 2 to display FFs for $\mu = Q = \{5, 10, 20\}$ GeV. Essentially, this reflects the influence of the DGLAP evolution in Q.

The variation of the uncertainty bands for the different scales is rather modest for all the flavors with the exception
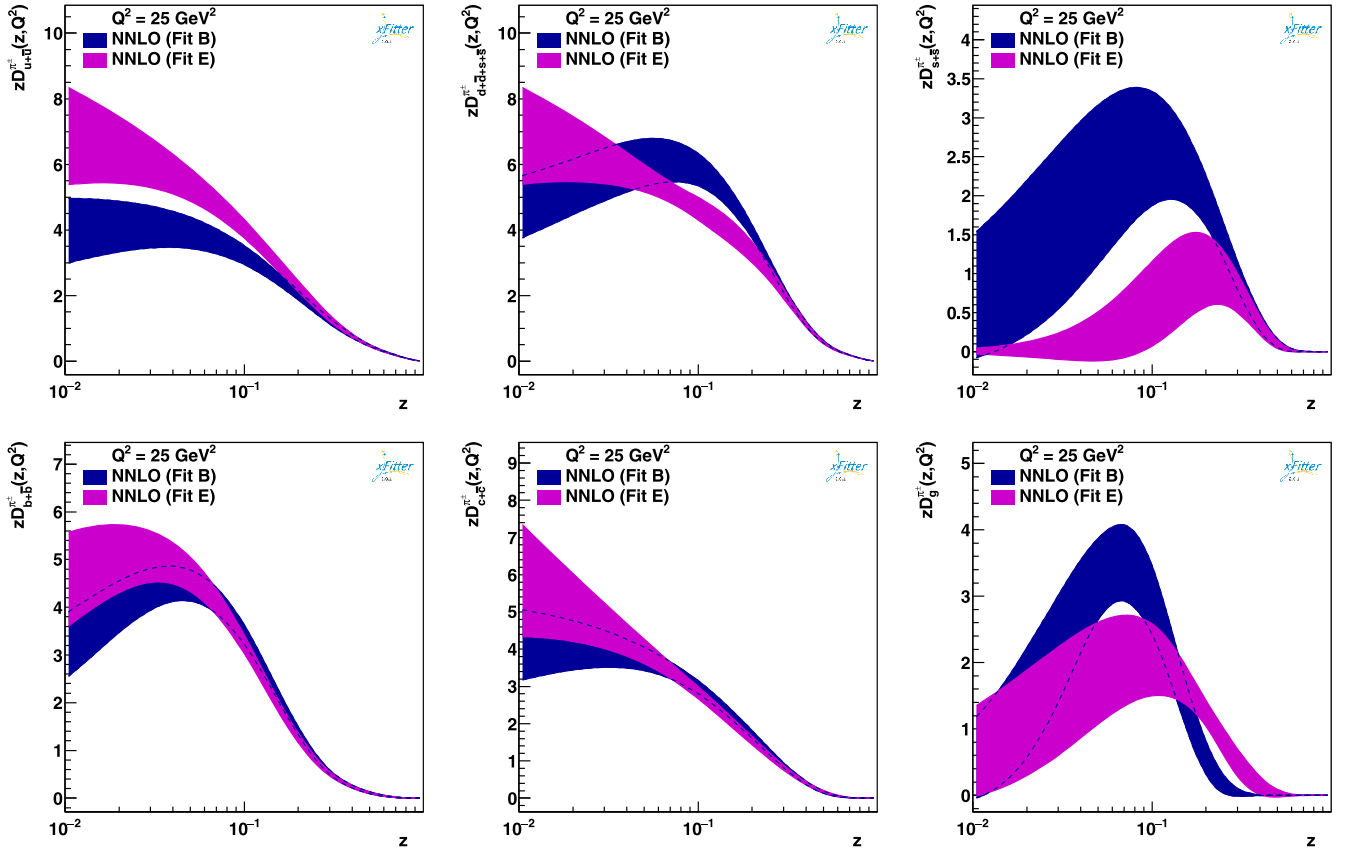
FIG. 6.   The comparison of the charged pion FFs determined for Fit B and Fit E at $Q_0^2 = 25$ GeV^2. as a function of z. Both fits include BELLE20 and *BABAR* data. Fit B has no cuts in z, while Fit E imposed a $z > 0.2$ cut on BELLE20 and a $z > 0.1$ cut on *BABAR*.
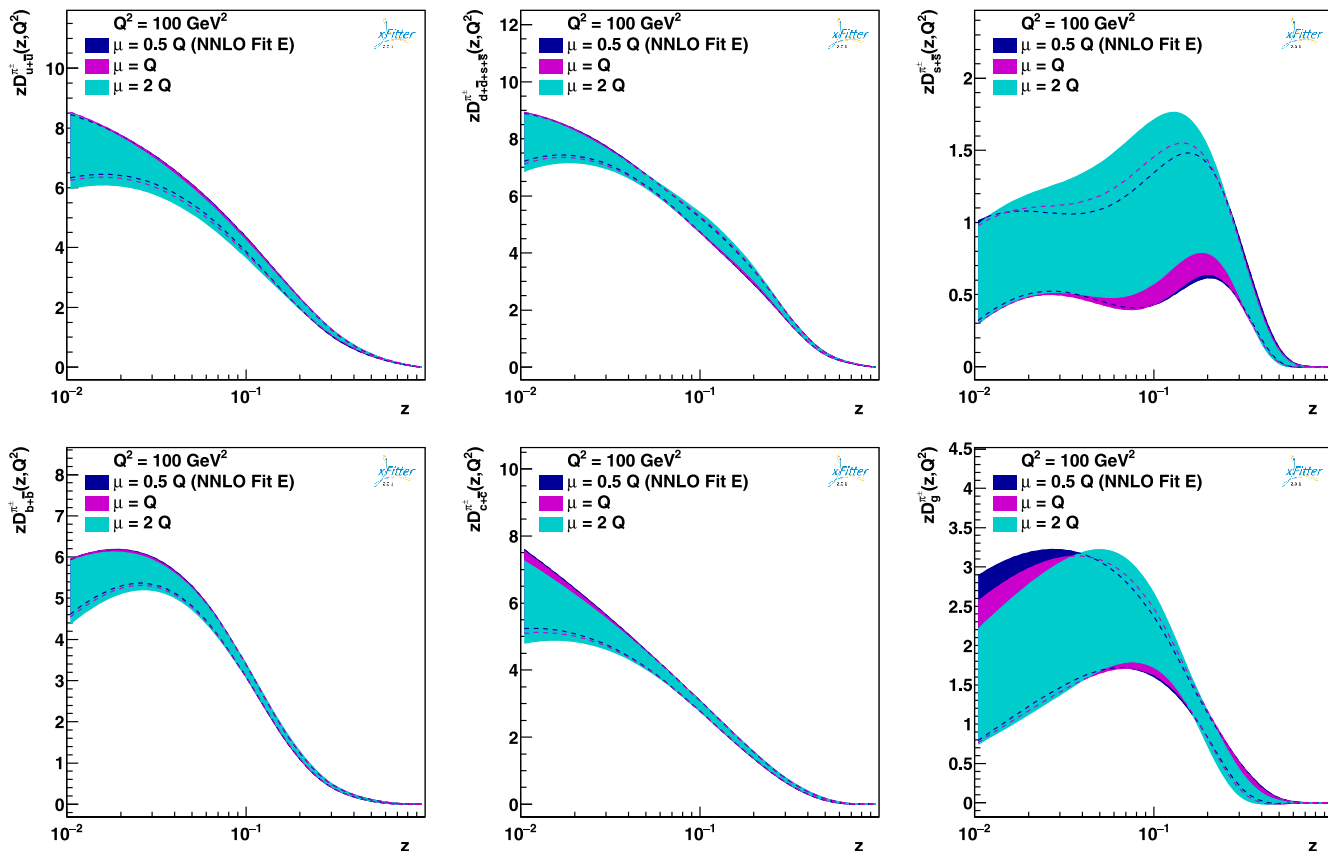
FIG. 7.   The NNLO charged pion FFs extracted of Fit E for different choices of the scale μ as function of z. We choose $Q^2 = 100$ GeV^2 and display the values $\mu = \{5, 10, 20\}$ GeV.

the strange FFs which shows a observable shift in the intermediate z region, and the gluon FFs which shows a shift in the low z region. This variation can provide an indication of higher order contributions which may not be included in our fixed-order perturbative calculation. For increasing scales, the FFs generally decrease in the small z region ($z \sim 0.01$), and this is most evident in the gluon FF.

5. Comparison with SIA data

To provide further details on the quality of the QCD fits, we now compare our theoretical predictions directly with the inclusive pion production data.

In Fig. 8 we display the theoretical predictions for Fit A, Fit B and Fit E at NNLO together with the inclusive pion production data and their uncertainties. The comparisons are given as a function of z for the various energies, and the measured z region for each experiment can be read from the plots. To provide detail, we show both the absolute distributions (upper panel) and the data over theory ratios (lower panel).

The results of TASSO, DELPHI, and SLD are split across multiple datasets (cf., Table I), but the experiments provide the correlated systematic uncertainty for the overall normalization of the data. For these datasets, the difference

between the "theory" and "theory + shifts" is due to this systematic normalization. For the other datasets of Fig. 8, there is no systematic normalizaton, so the "theory" and "theory + shifts" curves will coincide.

The theoretical predictions are generally in good agreement with the experimental measurements in the central z-range. We observe some deviations in the large-z region, but as the data uncertainties are typically larger here this does not significantly impact the χ^2 of the fits. In contrast, in the low-z region where the experimental uncertainties are typically smaller we do see some deviations of the theoretical predictions which, as discussed above, can substantially impact both the FFs (e.g., Fig. 6) and the χ^2 values (e.g., Table I). For the *BABAR* and BELLE20 plots, we note Fits A and B show deviations in the low z region, and this is in contrast to Fit E where these deviations are largely absent.

6. Theoretical effects at low-z

What could be the source of the large χ^2 contributions coming from the low-z region? One concern is that we could be missing important theoretical corrections such as nonfactorizable higher-twist contributions, or terms beyond the order or our perturbation theory, some of which can be

estimated using a resummation procedure. The original set of z cuts discussed in Sec. IV and outlined in Ref. [4] aim to minimize the nonperturbative contribution not described by our calculations.

Additionally, Ref. [57] performs an all-order resummation of logarithmically enhanced contributions at small momentum fraction z at NNLO logarithmic order, including the dependence on the factorization and renormalization scales. Specifically, they examine terms of the form $(\alpha_s^n \ln^k z)$ and resum these contributions in Mellins space to reduce the theoretical uncertainty in the low z region.

This analysis suggests that although the lowest few z data points of BELLE and *BABAR* may have modest corrections from resummation contributions, these alone are not sufficient to address the differences we observe with the current fits. Additionally, the fact that the BELLE and *BABAR* data seem to pull the theory in opposite directions further complicates the resolution of this situation and may
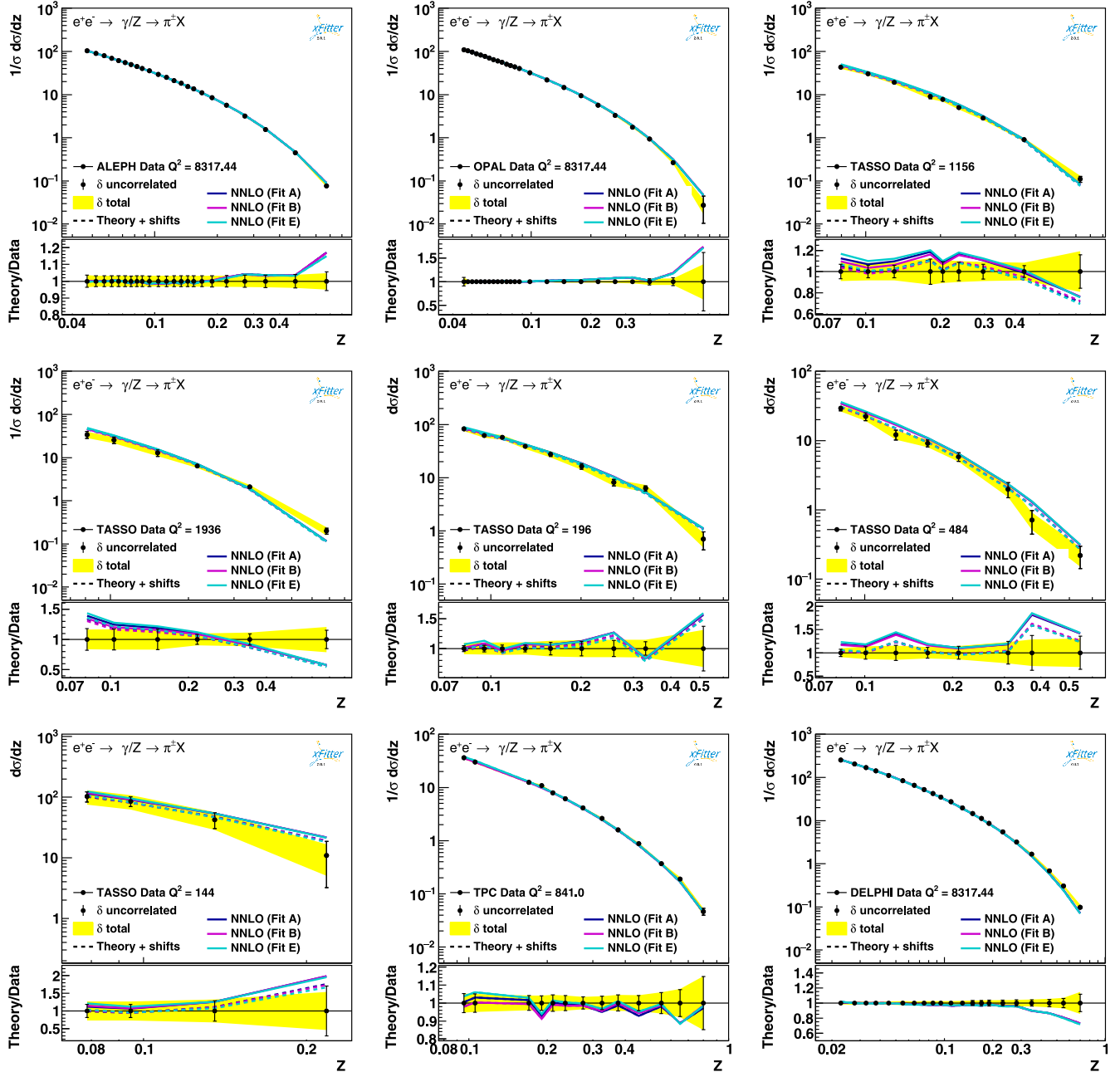
FIG. 8. (Continued).
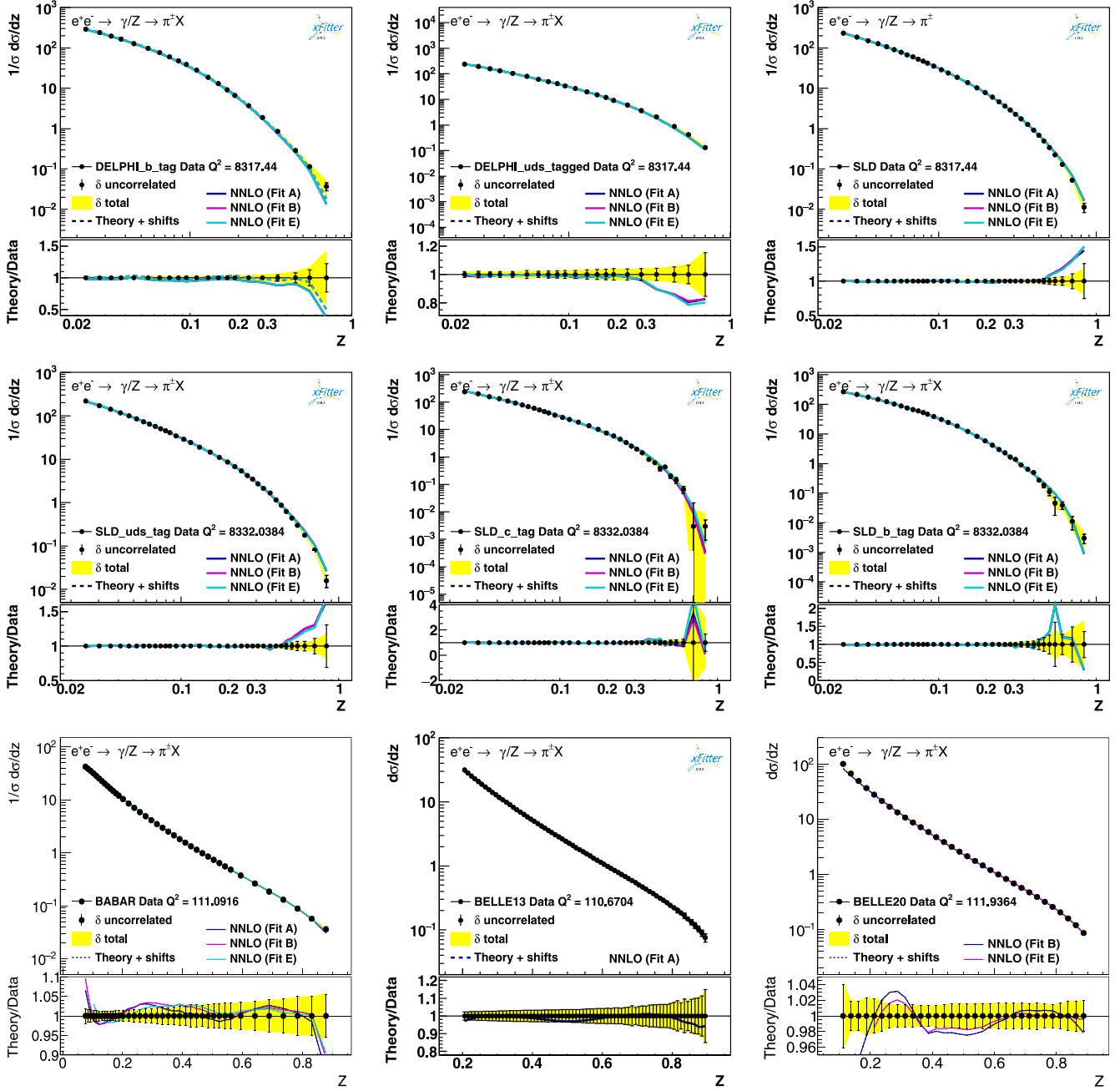
FIG. 8. The inclusive pion production as a function of z for fixed Q^2 values (displayed) for Fit A, Fit B and Fit E at NNLO. We compare with experimental data from ALEPH [50], OPAL [51], TASSO [53–55], TPC [48], DELPHI [52], SLD [49], *BABAR* [23] BELLE13 [22] and BELLE20 [21]. We show both the absolute distributions (upper panel) and the data over theory ratios (lower panel). The theory shifts arise from the correlated systematic overall normalization of the data points in the case where an individual experiment provides multiple datasets, such as TASSO, DELPHI, and SLD (cf., Table I).

point to nonperturbative effects not incorporated in our calculations.

In conclusion, while we generally obtain a good quality for our fits across most of the z range, the description of the data in the low-z region remains an unresolved puzzle, and we must leave this for future investigation.

C. Comparison to results from the literature

We now compare our results with recent analyses from the literature. While Fit E (IPM-xFITTER) is our preferred final fit, we also display Fit B to highlight the impact of the low z cuts. We compare with the results from the NNFF1 [4,58,59], JAM19 [13] and DSEHS14 [5] collaborations
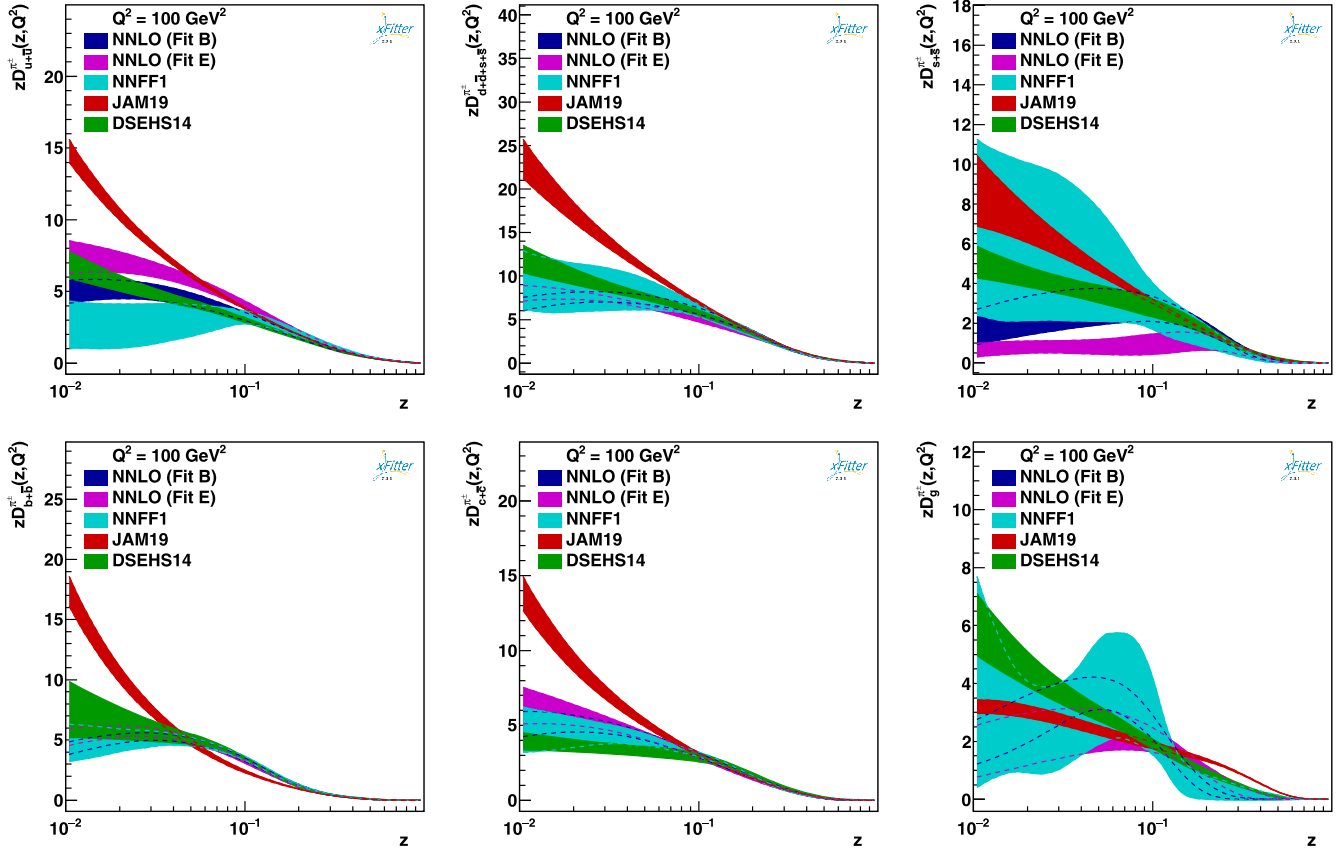
FIG. 9.   A comparison of our preferred Fit E [IPMX] as well as Fit B for charged pion FFs ($\pi^+ + \pi^-$) at NNLO with results from the literature at $Q^2 = 100$ GeV^2. We display NNFF1 [4] at NNLO, JAM19 [13] at NLO, DSEHS [5] at NLO, with their uncertainties at $Q^2 = 100$ GeV^2. Note, discretion is necessary when interpreting the very low z region as the extrapolation of the FF grids extends beyond the region fitted in the individual analyses. For example, the JAM19 focus was on SIDIS in the region $z \gtrsim 0.2$, and NNFF1 used a lower kinematic cut of $z_{\min} = 0.02$ for $Q = M_Z$ and 0.075 for $Q < M_Z$. While Fit E is our preferred fit, we also display Fit B to highlight the impact of the low z cuts.

which are computed at NLO (JAM19 and DSEHS14) and at NNLO (NNFF1). These results are displayed in Figs. 9 and 10 at $Q^2 = 100$ GeV^2 and $Q^2 = M_Z^2$, respectively.

Discretion is necessary when interpreting the very low z region as the extrapolation of the FF grids may extend beyond the region fitted in the individual analyses of Refs. [4,5,13]. For example, the JAM19 focus was on SIDIS in the region $z \gtrsim 0.2$, NNFF1 used a lower kinematic cut of $z_{\min} = 0.02$ for $Q = M_Z$ and 0.075 for $Q < M_Z$, and DSEHS14 also used various z cuts. Additionally, differences can arise due to the choice of the tolerance criteria as well as the method for computing the error bands such as the Hessian or Monte Carlo approach.

The DSEHS14 uses a combination of data from SIA, SIDIS, and hadron-hadron collisions in an NLO framework; additionally, our parametric form of Eq. (5) matches their initial FFs. The NNFF1 analysis is based on electron-positron SIA cross-sections for the sum of charged pion, charged kaon, and proton/antiproton production. The JAM analysis simultaneously fits both the PDFs and FFs using DIS, SIDIS, Drell-Yan and SIA data.

Comparing our up and down FFs with NNFF1 and DSEHS14, we see they are generally compatible at larger z, but differ in the low-z region; this is more pronounced for Fit E which imposes cuts on some of the low-z data. Similar conclusions apply to the case of the c^+ and b^+ heavy quarks. Comparing our gluon FF with NNFF1 and DSEHS14, again we see our FFs are generally compatible as our curves lie within the NNFF1 uncertainty band, and have overlap with DSEHS14 in the larger z region. For the strange quark there is more of a spread in results suggesting an overall increased uncertainty (note the vertical scale). This reflects, in part, the fact that our chosen dataset has minimal constraints on the strange FF.

In contrast, the above FFs generally have a different behavior as compared with the JAM19 analysis.[4] The JAM19 FFs have a much steeper slope at small z for the

_____

[4]While the JAM parametrizes their initial distributions as $z^a(1-z)^b(1 + c\sqrt{z} + dz)$, it is possible to match this to our functional form quite closely at Q_0; hence, this is not the source of the differences.
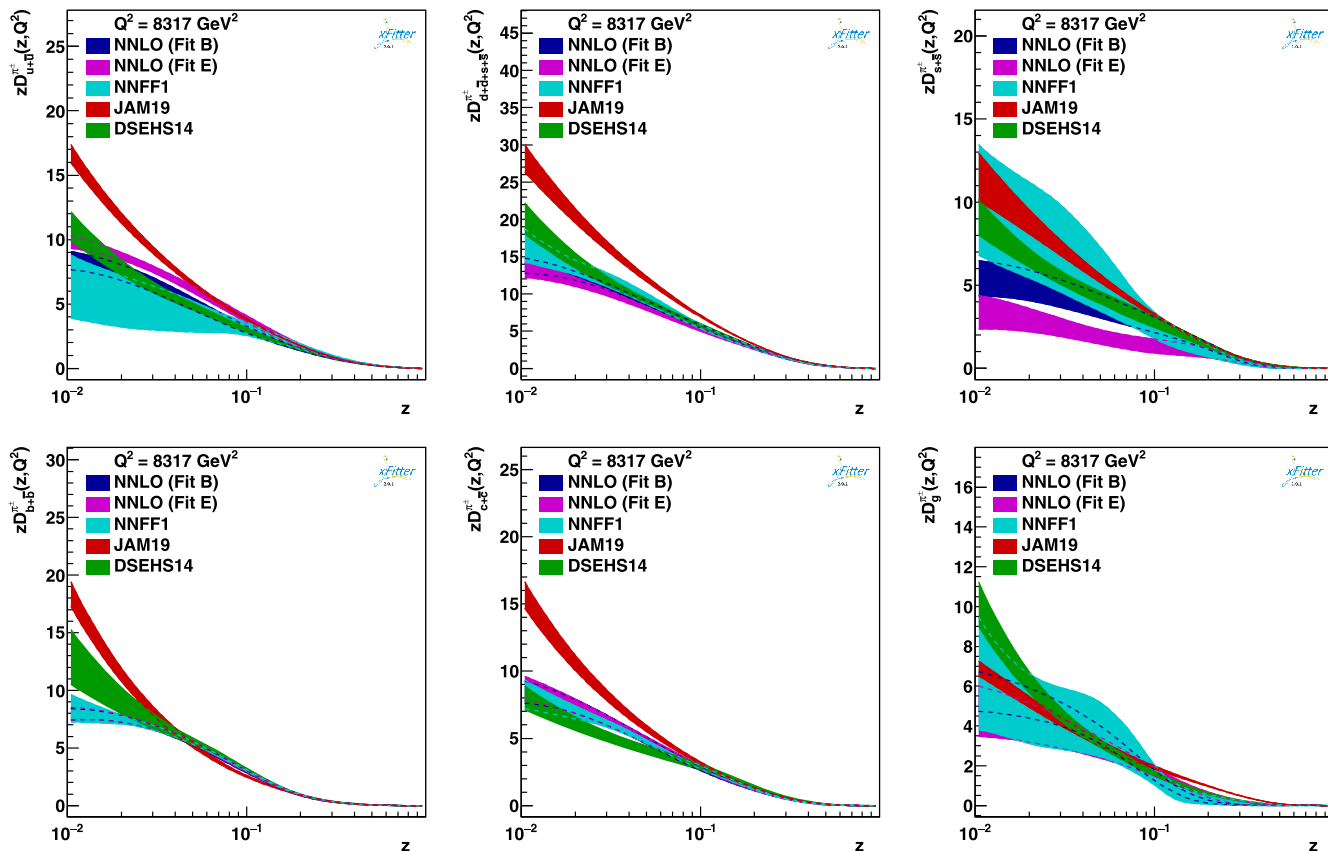
FIG. 10. Same as Fig. 9 but for $Q^2 = M_Z^2$.

quark flavors as compared to both our Fit B and Fit E, while the gluon generally lies above our curves for intermediate to larger z values. In addition, much smaller error bands are obtained in JAM19 analysis in respect to our results and the NNFF1 analysis. As noted above, the focus of the JAM19 study was on SIDIS in the region $z \gtrsim 0.2$, and in this region there is closer agreement among the FFs. The JAM19 study also performed a separately analysis without the SIDIS data (not shown here) and this resulted in a substantive shift of FFs. Specifically, this tends to increase the u^+ and decrease the g FFs in the larger z region, $z \gtrsim 0.2$. As the JAM19 results are obtain by a combined fit to both PDFs and FFs, it would be interesting to investigate further to see what difference might be separately attributable to this combined analysis and the choice of datasets.

VI. SUMMARY AND CONCLUSIONS

In this paper we have presented the first open source analysis of unpolarized pion FFs at NNLO accuracy in pQCD. Our analysis is based on a comprehensive dataset from SIA experiments, including the most recent measurement of single pion production by the BELLE collaboration [21].

This analysis includes several novel features including the use of the open source xFITTER framework to perform

the analysis, as well as to compute the Hessian uncertainties for the FFs. We have compared both NLO and NNLO results, and also studied the impact of the B-factory data.

The SIA experimental data are reasonably well described by our QCD fits, with the exception of the low-z region. This remains as an important unresolved issue and will clearly require additional study.

The comparisons of our pion FFs to the results in the literature are also informative. While some of the results agree within uncertainties, there are a number of areas where the FFs differ, and thus warrants further examination.

This paper facilitates future investigations as these results are available within the open-source xFITTER program, and also as pion FF grids formatted for the LHAPDF6 library. Therefore, this analysis serves as an important step toward a broader program to improve the determination of the nuclear PDFs and FFs of hadrons as we strive to better comprehend the full character of the QCD theory.

Note added.—As we completed this study, a related investigation was reported in Ref. [60] using a complementary neural-network approach with different datasets; this opens new avenues for future study.

---

[1] H.-W. Lin *et al.*, Parton distributions and lattice QCD calculations: A community white paper, Prog. Part. Nucl. Phys. **100,** 107 (2018).

[2] M. Constantinou *et al.*, Parton distributions and lattice QCD calculations: Toward 3D structure, arXiv:2006.08636.

[3] X. Ji, Y. Liu, Y.-S. Liu, J.-H. Zhang, and Y. Zhao, Large-momentum effective theory, arXiv:2004.03543.

[4] V. Bertone, S. Carrazza, N. P. Hartland, E. R. Nocera, and J. Rojo (NNPDF Collaboration), A determination of the fragmentation functions of pions, kaons, and protons with faithful uncertainties, Eur. Phys. J. C **77,** 516 (2017).

[5] D. de Florian, R. Sassot, M. Epele, R. J. Hernández-Pinto, and M. Stratmann, Parton-to-pion fragmentation reloaded, Phys. Rev. D **91,** 014035 (2015).

[6] N. Sato, J. J. Ethier, W. Melnitchouk, M. Hirai, S. Kumano, and A. Accardi, First Monte Carlo analysis of fragmentation functions from single-inclusive e^+e^- annihilation, Phys. Rev. D **94,** 114004 (2016).

[7] M. Soleymaninia, M. Goharipour, H. Khanpour, and H. Spiesberger, Simultaneous extraction of fragmentation functions of light charged hadrons with mass corrections, Phys. Rev. D **103,** 054045 (2021).

[8] I. Scimemi and A. Vladimirov, Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum, J. High Energy Phys. 06 (2020) 137.

[9] M. Soleymaninia and H. Khanpour, Transverse momentum dependent of charged pion, kaon, and proton/antiproton fragmentation functions from e^+e^- annihilation process, Phys. Rev. D **100,** 094033 (2019).

[10] H. Jung, S. T. Monfared, and T. Wening, Determination of collinear and TMD photon densities using the Parton Branching method, Phys. Lett. B **817,** 136299 (2021).

[11] P. Zurita, Medium modified fragmentation functions with open source xFITTER, arXiv:2101.01088.

[12] J. J. Ethier, N. Sato, and W. Melnitchouk, First Simultaneous Extraction of Spin-Dependent Parton Distributions and Fragmentation Functions from a Global QCD Analysis, Phys. Rev. Lett. **119,** 132001 (2017).

[13] N. Sato, C. Andres, J. J. Ethier, and W. Melnitchouk (JAM Collaboration), Strange quark suppression from a simultaneous Monte Carlo analysis of parton distributions and fragmentation functions, Phys. Rev. D **101,** 074020 (2020).

[14] E. Moffat, W. Melnitchouk, T. Rogers, and N. Sato, Simultaneous Monte Carlo analysis of parton densities and fragmentation functions, Phys. Rev. D **104,** 016015 (2021).

[15] M. Hirai, S. Kumano, T. H. Nagai, and K. Sudoh, Determination of fragmentation functions and their uncertainties, Phys. Rev. D **75,** 094009 (2007).

[16] M. Soleymaninia, M. Goharipour, and H. Khanpour, Impact of unidentified light charged hadron data on the determination of pion fragmentation functions, Phys. Rev. D **99,** 034024 (2019).

[17] B. A. Kniehl, G. Kramer, and B. Potter, Testing the universality of fragmentation functions, Nucl. Phys. **B597,** 337 (2001).

[18] B. A. Kniehl, G. Kramer, and B. Potter, Fragmentation functions for pions, kaons, and protons at next-to-leading order, Nucl. Phys. **B582,** 514 (2000).

[19] J. Binnewies, B. A. Kniehl, and G. Kramer, Next-to-leading order fragmentation functions for pions and kaons, Z. Phys. C **65,** 471 (1995).

[20] S. Kretzer, E. Leader, and E. Christova, Fragmentation functions from semiinclusive DIS pion production and implications for the polarized parton densities, Eur. Phys. J. C **22,** 269 (2001).

[21] R. Seidl *et al.* (Belle Collaboration), Update of inclusive cross sections of single and pairs of identified light charged hadrons, Phys. Rev. D **101,** 092004 (2020).

[22] M. Leitgab *et al.* (Belle Collaboration), Precision Measurement of Charged Pion and Kaon Differential Cross Sections in e^+e^- Annihilation at $\sqrt{s} = 10.52$ GeV, Phys. Rev. Lett. **111,** 062002 (2013).

[23] J. P. Lees *et al.* (*BABAR* Collaboration), Production of charged pions, kaons, and protons in e^+e^- annihilations into hadrons at $\sqrt{s} = 10.54$ GeV, Phys. Rev. D **88,** 032011 (2013).

[24] S. Alekhin *et al.*, HERAFitter, Eur. Phys. J. C **75,** 304 (2015).

[25] V. Bertone *et al.* (xFITTER Developers' Team Collaboration), xFITTER 2.0.0: An open source QCD fit framework, *Proc. Sci.*, DIS2017 (2018) 203 [arXiv:1709.01151].

[26] I. Novikov *et al.*, Parton distribution functions of the charged pion within the xFITTER framework, Phys. Rev. D **102,** 014040 (2020).

[27] M. Walt, I. Helenius, and W. Vogelsang, Open-source QCD analysis of nuclear parton distribution functions at NLO and NNLO, Phys. Rev. D **100,** 096015 (2019).

[28] H. Abdolmaleki *et al.* (xFITTER Developers' Team Collaboration), Probing the strange content of the proton with charm production in charged current at LHeC, Eur. Phys. J. C **79,** 864 (2019).

[29] E. Accomando *et al.*, PDF profiling using the forward-backward asymmetry in neutral current Drell-Yan production, J. High Energy Phys. 10 (2019) 176.

[30] H. Abdolmaleki and A. Khorramian, Parton distribution functions and constraints on the intrinsic charm content of the proton using the Brodsky-Hoyer-Peterson-Saka approach, Phys. Rev. D **99**, 116019 (2019).

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

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

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

[34] P. J. Rijken and W. L. van Neerven, Higher order QCD corrections to the transverse and longitudinal fragmentation functions in electron-positron annihilation, Nucl. Phys. **B487**, 233 (1997).

[35] A. Mitov and S.-O. Moch, QCD corrections to semi-inclusive hadron production in electron-positron annihilation at two loops, Nucl. Phys. **B751**, 18 (2006).

[36] V. N. Gribov and L. N. Lipatov, Deep inelastic e p scattering in perturbation theory, Sov. J. Nucl. Phys. **15**, 438 (1972).

[37] L. N. Lipatov, The parton model and perturbation theory, Yad. Fiz. **20**, 181 (1974).

[38] G. Altarelli and G. Parisi, Asymptotic freedom in parton language, Nucl. Phys. **B126**, 298 (1977).

[39] 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).

[40] V. Bertone, S. Carrazza, and J. Rojo, APFEL: A PDF evolution library with QED corrections, Comput. Phys. Commun. **185**, 1647 (2014).

[41] 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).

[42] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C **75**, 580 (2015).

[43] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, Uncertainties of predictions from parton distribution functions. 2. The Hessian method, Phys. Rev. D **65**, 014013 (2001).

[44] 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).

[45] W. T. Giele, S. A. Keller, and D. A. Kosower, Parton distribution function uncertainties, arXiv:hep-ph/0104052.

[46] K. Kovarik, I. Schienbein, F. I. Olness, J. Y. Yu, C. Keppel, J. G. Morfin, J. F. Owens, and T. Stavreva, Nuclear Corrections in Neutrino-Nucleus DIS and Their Compatibility with Global NPDF Analyses, Phys. Rev. Lett. **106**, 122301 (2011).

[47] P. A. Zyla *et al.* (Particle Data Group Collaboration), Review of particle physics, Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).

[48] H. Aihara *et al.* (TPC/Two Gamma Collaborations), Charged Hadron Inclusive Cross-Sections and Fractions in e^+e^- Annihiliation $\sqrt{s} = 29$ GeV, Phys. Rev. Lett. **61**, 1263 (1988).

[49] K. Abe *et al.* (SLD Collaboration), Production of π^+, π^-, K^+, K^-, p and $\bar{b}$ in light (uds), c and b jets from Z^0 decays, Phys. Rev. D **69**, 072003 (2004).

[50] D. Buskulic *et al.* (ALEPH Collaboration), Inclusive $\pi^{\pm}$, $K^{\pm}$ and $(p, \bar{p})$ differential cross-sections at the Z resonance, Z. Phys. C **66**, 355 (1995).

[51] R. Akers *et al.* (OPAL Collaboration), Measurement of the production rates of charged hadrons in e^+e^- annihilation at the Z^0, Z. Phys. C **63**, 181 (1994).

[52] P. Abreu *et al.* (DELPHI Collaboration), $\pi^{\pm}$, $K^{\pm}$, p and $\bar{p}$ production in $Z^0 \rightarrow q\bar{q}$, $Z^0 \rightarrow b\bar{b}$, $Z^0 \rightarrow u\bar{u}$, $d\bar{d}$, $s\bar{s}$, Eur. Phys. J. C **5**, 585 (1998).

[53] R. Brandelik *et al.* (TASSO Collaboration), Charged pion, kaon, proton and anti-proton production in high-energy e^+e^- annihilation, Phys. Lett. **94B**, 444 (1980).

[54] M. Althoff *et al.* (TASSO Collaboration), Charged hadron composition of the final state in e^+e^- annihilation at high-energies, Z. Phys. C **17**, 5 (1983).

[55] W. Braunschweig *et al.* (TASSO Collaboration), Pion, kaon and proton cross-sections in e^+e^- annihilation at 34 GeV and 44 GeV center-of-mass energy, Z. Phys. C **42**, 189 (1989).

[56] M. Soleymaninia, M. Goharipour, and H. Khanpour, First QCD analysis of charged hadron fragmentation functions and their uncertainties at next-to-next-to-leading order, Phys. Rev. D **98**, 074002 (2018).

[57] D. P. Anderle, T. Kaufmann, M. Stratmann, and F. Ringer, Fragmentation functions beyond fixed order accuracy, Phys. Rev. D **95**, 054003 (2017).

[58] E. R. Nocera, Towards a neural network determination of charged pion fragmentation functions, in *22nd International Symposium on Spin Physics* (2017).

[59] V. Bertone, S. Carrazza, E. R. Nocera, N. P. Hartland, and J. Rojo, Towards a neural network determination of pion fragmentation functions, in *Parton Radiation and Fragmentation from LHC to FCC-ee* (2017).

[60] R. Abdul Khalek, V. Bertone, and E. R. Nocera, A determination of unpolarised pion fragmentation functions using semi-inclusive deep-inelating-scattering data: MAPFF1.0, arXiv:2105.08725.