Determination of K^0_S Fragmentation Functions including BESIII Measurements and using Neural Networks

Maryam Soleymaninia[1] ,* Hadi Hashamipour[2,1] ,† Maral Salajegheh[3] ,‡ Hamzeh Khanpour[4,5,1] ,§ Hubert Spiesberger[6],¶ and Ulf-G. Meißner[3,7,8] **

[1] *School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM), P.O.Box 19395-5531, Tehran, Iran.*
[2] *Istituto Nazionale di Fisica Nucleare, Gruppo collegato di Cosenza, I-87036 Arcavacata di Rende, Cosenza, Italy.*
[3] *Helmholtz-Institut für Strahlen-und Kernphysik and Bethe Center for Theoretical Physics, Universität Bonn, D-53115 Bonn, Germany.*
[4] *AGH University, Faculty of Physics and Applied Computer Science, Al. Mickiewicza 30, 30-055 Kraków, Poland.*
[5] *Department of Physics, University of Science and Technology of Mazandaran, P.O.Box 48518-78195, Behshahr, Iran.*
[6] *PRISMA$^+$ Cluster of Excellence, Institut für Physik, Johannes-Gutenberg-Universität, Staudinger Weg 7, D-55099 Mainz, Germany*
[7] *Institute for Advanced Simulation (IAS-4), Forschungszentrum Jülich, D-52425 Jülich, Germany.*
[8] *Tbilisi State University, 0186 Tbilisi, Georgia.*

(Dated: June 1, 2024)

In this study, we revisit the extraction of parton-to-K^0_S hadron fragmentation functions, named FF24-K^0_S, focusing on both next-to-leading-order and next-to-next-to-leading-order accuracy in perturbative QCD. Our approach involves the analysis of single inclusive electron-positron annihilation (SIA) data. The two key improvements are, on the one hand, the incorporation of the latest experimental data from the BESIII experiment and, on the other hand, the adoption of Neural Networks in the fitting procedure. To address experimental uncertainties, the Monte Carlo method is employed. Our investigation also explores the impact of hadron mass corrections on the description of SIA data, spanning a broad kinematic regime with a particular emphasis on the range of small z values. The theory prediction for K^0_S production at both NLO and NNLO accuracy exhibits good agreement with experimental data within their respective uncertainties.

CONTENTS

I. INTRODUCTION

Fragmentation functions (FFs) play a crucial role in the computation of scattering cross sections that involve detected final-state hadrons. When there is a high energy scale, QCD factorization allows us to disentangle the physics of a parton transitioning to a colorless hadron from the hard scattering process that generated the parton [1]. Parton-to-hadron FFs precisely depict this transition and illustrate how QCD final states are nonperturbatively created through hadronization, providing insight into the dynamics of the strong interaction. Fragmentation functions are process-independent quantities and can be determined from data through a comprehensive QCD analysis. The variation of the FFs with energy can be perturbatively calculated, comprising an expansion in the strong coupling, and is presently understood up to next-to-next-to-leading order (NNLO) accuracy [2, 3]. Due to the significance of FFs, there has been considerable theoretical attention towards their global analysis, resulting in the development of multiple new sets for various hadron FFs [4–7].

Fragmentation functions have a large diversity of applications. Their knowledge is essential in numerous hard scattering processes at both existing and upcoming experiments, such as at the Large Hadron Collider (LHC), the Future Circular Collider (FCC) [8, 9], and the Electron-Ion Collider (EIC) [10]. Particle production measurements provide valuable insight into the dynamics of QCD interactions at low momentum transfer. Accurate modeling of these interactions is crucial for understanding and constraining the impact of the underlying event in collisions at high transverse momentum investigated at the LHC. Other applications include the exami-

* Maryam_Soleymaninia@ipm.ir
† hadi.hashamipour@lnf.infn.it
‡ Maral@hiskp.uni-bonn.de
§ Hamzeh.Khanpour@cern.ch
¶ spiesber@uni-mainz.de
** meissner@hiskp.uni-bonn.de

nation of the spin structure of the nucleon [11]. Moreover, FFs play a pivotal role in investigating hadron production rates in scattering processess involving heavy nuclei [12].

The distribution of hadrons containing strange quarks has been documented at various center-of-mass energies, both at the LHC [13, 14] and at the Tevatron [15]. Kaons are particularly important. FFs for K^0_S production have been calculated by some of the authors of the present paper [16], denoted SAK20. Other research groups, such as BKK96 [17], BS [18], AKK05 [19], AKK08 [20], DSS17 [21], and NNFF [22] have also investigated K^0_S FFs. Details of these analyses are discussed elsewhere, and we direct the reader to our previous paper [16] for a review.

In our previous investigations, referred to as SAK20, we incorporated experimental data from the single inclusive electron-positron annihilation process (SIA) to determine FFs for K^0_S production. We provided FF sets for K^0_S, along with their associated uncertainties determined through the Hessian method, at both next-to-leading order (NLO) and NNLO precision. Additionally, in [16] we had examined the effect of hadron mass corrections as well.

Recently, the BESIII detector has performed measurements of the normalized differential cross sections for inclusive K^0_S production in e^+e^- annihilation, covering six energies from 2.2324 to 3.6710 GeV with relative hadron energies, z, ranging from 0.2 to 0.9 [23]. Since this energy range is not well-covered by previous experimental data, it will provide valuable input for a QCD analysis to determine the kaon FFs. This new dataset motivates us to revisit and update our previous analysis of K^0_S FFs [16], now incorporating the BESIII measurements. Our new FF sets, FF24-K^0_S, at both NLO and NNLO, are publicly available in the LHAPDF format.

In recent years, machine learning (ML) has gained increasing popularity in various domains of particle physics, particularly in collider physics. One promising application of ML methods is their role in advancing our understanding of non-perturbative quantities related to nucleons, such as parton distribution functions (PDFs) and FFs [22, 24, 25]. Given this trend, we chose to leverage artificial neural networks (NN) for the extraction of FFs for K^0_S hadrons in this work. Utilizing neural networks is expected to minimize the bias in the parameterization of the FFs. In addition, we use the Monte Carlo sampling method as a robust statistical approach to account for uncertainties in experimental data and derive probability density distributions of FFs from the data. To facilitate this analysis, we utilize the publicly available code MontBlanc, accessible from [26].

The organization of this paper is as follows: In Sec. II we present and discuss the experimental observable for K^0_S production in the SIA process. The theoretical framework employed in this work as well as the hadron mass corrections are presented in Sec. III. Sec. IV deals with the χ^2 minimization and the methodology for the calculation of FFs uncertainties. Our main results from this study are presented and discussed in detail in Sec. V. Finally, Sec. VI provides our summary and conclusions.

II. EXPERIMENTAL DATA

This analysis relies on experimental measurements of the K^0_S production cross sections in the SIA process. Various collaborations have reported SIA data for the cross section of K^0_S hadrons in the final state, considering the longitudinal momentum fraction z at different energy scales. The majority of differential cross sections have been normalized to the total cross section.

In this analysis, we incorporate a comprehensive set of experimental data from various collaborations conducted at CERN, DESY, SLAC, KEK, and BESIII. Our dataset encompasses untagged data from the TASSO collaboration at $\sqrt{s}$ = 14, 22, and 34 GeV [27], as well as at $\sqrt{s}$ = 14.8, 21.5, 34.5, 35, and 42.6 GeV [28]. Additionally, we include measurements from the HRS [29], TPC [30], and MARK II [31] collaborations at $\sqrt{s}$ = 29 GeV. Our dataset also includes data from the CELLO collaboration at $\sqrt{s}$ = 35 GeV [32] and the TOPAZ collaboration at $\sqrt{s}$ = 58 GeV [33].

Included in our analysis are also datasets measured at $\sqrt{s} = M_Z$ from ALEPH [34], DELPHI [35], and OPAL [36], as well as untagged and tagged quark data (light quarks (u, d, s) and heavy quarks c and b) from the SLD [37] collaboration.

The primary motivation for this analysis is the incorporation of the most recent measurements of normalized differential cross sections for inclusive K^0_S production in SIA processes reported by the BESIII collaboration [23]. Their dataset spans six center-of-mass energies ranging from $\sqrt{s}$ = 2.2324 to 3.6710 GeV and features a z coverage from 0.2 to 0.9. Collected with the BESIII detector at BEPCII, these new measurements offer crucial input for fitting FFs in the region with $\sqrt{s} < 10$ GeV, where precision SIA data have been scarce.

As Fig. 3 of the BESIII publication [23] shows, the comparison between the K^0_S data and different theoretical predictions indicates a discrepancy, likely attributed to the fact that existing FFs had been determined from SIA data at high energies and have to be extrapolated to the lower energy scales of the BESIII measurement. One can conclude that BESIII data should be used to improve the fits of FFs at the low energy scales.

Kinematic cuts are implemented to ensure that the selected data points are in a regime where perturbative fixed-order predictions hold reliable. The kinematic coverage of the datasets applied in this analysis is shown in Fig.1. We adjust the minimal value of z at $z_{\min} = 0.013$ for all experiments. For most of the experiments, the upper limit $z_{\max}$ is set at 0.9, except for two datasets from BESIII: the data at the energies $\sqrt{s}$ = 2.2324 and 2.4 GeV are cut at $z_{\max} = 0.8$. Due to outliers observed in the final data points of these datasets, their inclusion in the fit leads to an increased χ^2 value without substantially altering the resulting FFs. Hence, there is no

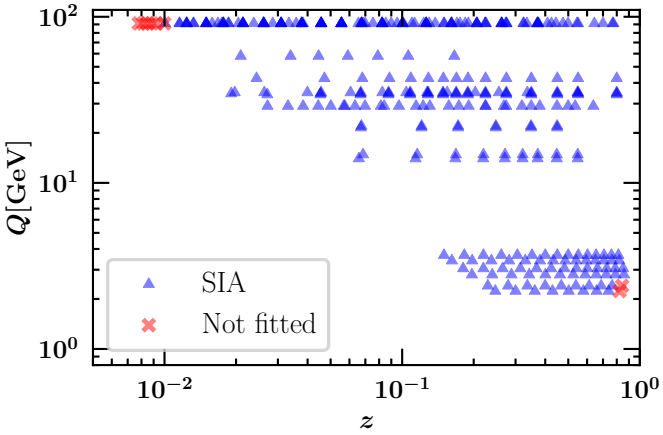

Figure 1: Kinematic range of the experimental SIA data in the (z, Q) plane used to determine the K^0_S FFs.

compelling rationale for integrating them into the analysis. Overall, we omit just 8 data points across all the SIA datasets used in this analysis and find that $N_{\rm dat} = 346$ data points can be used.

It is worth mentioning that, unlike in our previous analysis [16], we chose not to incorporate the experimental data from `HRS` [29] and `DELPHI` at 183 and 189 GeV [38] in this study. This decision was made since we could not find a satisfactory fit for these particular datasets. Specifically, their inclusion resulted in a noticeable decline in the overall quality of the fit, with χ^2 values per data point exceeding 7 for `HRS`, approximately 7 for `DELPHI` at 183 GeV, and surpassing 17 at 189 GeV. This discrepancy is evident when comparing the high χ^2 values per data point observed in our previous analysis and in `AKK08` [20] to the even poorer performance in the current analysis. We have identified the source of this discrepancy as an apparent inconsistency between these three datasets and the others, particularly with BESIII.

III. THEORETICAL SETUP

The theoretical setup of our analysis closely follows that illustrated in Sec. III of Ref. [16]. In this section, we recapitulate the theoretical understanding of standard collinear factorization, comprising perturbative and non-perturbative components of the cross-section for the SIA process. We also review the time-like DGLAP evolution of FFs, alongside detailing numerical computations of SIA cross-sections, including theoretical settings and physical parameters.

The total cross-section for e^+e^- annihilation in the production of hadrons ($e^+e^- \to h + X$) can be expressed in terms of the convolution of coefficient functions and FFs

$$\frac{1}{\sigma_{tot}}\frac{d\sigma^h}{dz} = \frac{1}{\sigma_{tot}} C_i\left(x, \alpha_s(\mu), \frac{Q^2}{\mu^2}\right) \otimes D_i^h\left(\frac{z}{x}, \mu^2\right) \quad (1)$$

The coefficient functions $C_i(x, \alpha_s(\mu), Q^2/\mu^2)$ are currently computed up to $\mathcal{O}(\alpha_s^2)$ in perturbative QCD [3, 39] and $D_i^h(z/x, \mu^2)$ denotes the FFs. The parton FFs at different energy scales are connected through the time-like DGLAP evolution equations [40], given by:

$$\frac{\partial D_i^h(z,\mu^2)}{\partial \ln \mu^2} = P_{ji}\left(x, \alpha_s(\mu^2)\right) \otimes D_j^h\left(\frac{z}{x}, \mu^2\right). \quad (2)$$

In this equation, $P_{ji}(x, \alpha_s(\mu^2))$ represent the time-like splitting functions and $\otimes$ denotes the convolution integral.

When incorporating the mass of the produced hadrons into the kinematic calculations, hadron mass corrections modify the FFs. In the context of the SIA process, these corrections are very large and therefore indispensable, in particular at low values of z. Hadron mass effects are introduced by employing light-cone coordinates and can be taken into account by using a modified scaling variable η instead of z:

$$\eta = \frac{z}{2}\left(1 + \sqrt{1 - \frac{4m_h^2}{sz^2}}\right), \quad (3)$$

where m_h is the hadron mass. With this change of variables we can express the cross section as

$$\frac{d\sigma}{dz} = \frac{1}{1 - m_h^2/s\eta^2} \sum_a \int_\eta^1 \frac{dx_a}{x_a} \frac{d\hat{\sigma}_a}{dx_a} D_a^h\left(\frac{\eta}{x_a}, \mu\right), \quad (4)$$

Here, $d\hat{\sigma}_a/dx_a$ denotes the differential cross section of the hard sub-process with parton a, calculated in pQCD, and a runs over all quark flavors and the gluon. We use Eq. (4) for the theory predictions of this work. In this study, we compute the K^0_S hadron FFs within the framework of the Zero-Mass Variable-Flavor-Number Scheme (ZM-VFNS), where all active flavors are treated as massless. However, the inclusion of heavy quark masses is necessary in order to calculate the number of active flavors taking into account the correct heavy-quark thresholds.

In this analysis, fixed values are adopted for the charm and bottom quark masses, setting $m_c = 1.51$ GeV and $m_b = 4.92$ GeV, respectively. The strong running coupling is calculated at two-loop order by setting $\alpha_s(m_Z = 91.1876\ \text{GeV}) = 0.118$.

To illustrate the need for hadron mass corrections, we present in Fig. 2 the ratio η/z as a function of z at five representative values of $\sqrt{s}$. This figure indicates that the deviation of the momentum fraction available to the hadron (η) from z becomes significant when z and/or $\sqrt{s}$ decrease. We see that at $\sqrt{s} = 2.2324$ GeV and $\sqrt{s} = 3.4$ GeV this correction is sizeable in the full range of z values. The corrections are larger than 20% for $z < 0.6$ at $\sqrt{s} = 2.2324$ GeV, $z < 0.3$ at $\sqrt{s} = 3.4$ GeV, $z < 0.1$ at

$\sqrt{s} = 14$ GeV, $z < 0.04$ at $\sqrt{s} = 29$ GeV, and $z < 0.015$ at $\sqrt{s} = 91.2$ GeV. Corrections exceeding 50% can be observed around $z < 0.4$ at $\sqrt{s} = 2.2324$ GeV, $z < 0.3$ at $\sqrt{s} = 3.44$ GeV, $z < 0.06$ at $\sqrt{s} = 14$ GeV, $z < 0.03$ at $\sqrt{s} = 29$ GeV, and $z < 0.01$ at $\sqrt{s} = 91.2$ GeV. Since we include the new `BESIII` data, which have been measured from 2.2324 to 3.6710 GeV, the hadron mass corrections are essential in the kinematic region covered by the data and significantly improves the description of the data used in this analysis.

For the cross sections, Eq. (4), our investigations show that at the energy scale of 91.2 GeV, the hadron mass correction is approximately $\pm 4\,\%$. At lower values of $\sqrt{s} = 29$ GeV, relevant for the TPC data, this correction varies approximately in the range between $\sim -20\,\%$ and $\sim +6\,\%$ depending on the value of z, while for the case of the TASSO 14 data, the correction is between $-20\,\%$ and $+12\,\%$. Our finding is that for energies of 3.4 GeV and 2.2324 GeV, mass corrections become more important and reach $\sim \pm 45\,\%$ at $\sqrt{s} = 3.4$ GeV and $\sim +35\,\%$ to $\sim -50\,\%$ at $\sqrt{s} = 2.2324$ GeV for BESIII. All in all, hadron mass corrections prove to be significant in mid-range energies and indispensable for low-energy experiments conducted at BESIII.

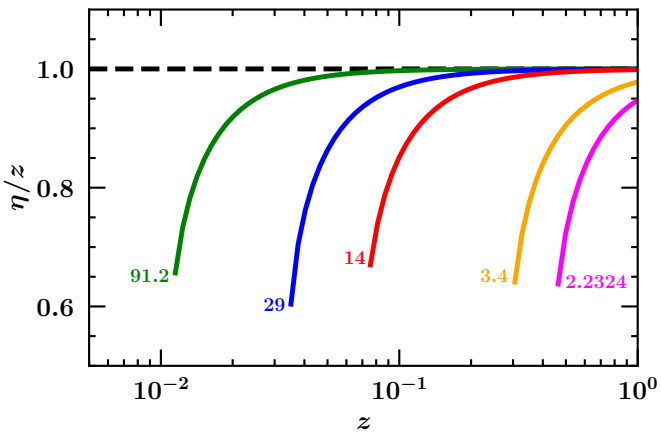

Figure 2: Ratio η/z as a function of z at five representative values of $\sqrt{s}$.

As outlined in our previous publication [16], the analysis of inclusive SIA data only allows one to determine the combined quark and antiquark FFs, i.e. for the quark combinations $q^+ = q + \bar{q}$. SIA cross section data are available for tagged light, charm, and bottom quarks. Therefore we can adopt a flavor decomposition approach and, keeping in mind the flavor content of K^0_S, we chose the following combinations to be fitted:

$$D^{K^0_S}_{u^+},\ D^{K^0_S}_{d^+},\ D^{K^0_S}_{s^+},\ D^{K^0_S}_{c^+},\ D^{K^0_S}_{b^+},\ D^{K^0_S}_{g}. \tag{5}$$

In the next section we describe the optimization methodology and the method of FFs determination.

IV. OPTIMIZATION METHODOLOGY AND UNCERTAINTY

The determination of unknown parameters or functions typically involves employing a maximum log-likelihood method. Assuming small uncertainties and therefore an approximately linear parameter dependence, the problem can be reduced to minimizing an appropriately chosen χ^2. In this work we use neural networks to describe the theoretical predictions and choose to minimize the following χ^2 function:

$$\chi^{2(k)} \equiv \left(\mathbf{T}(\boldsymbol{\theta}^{(k)}) - \boldsymbol{x}^{(k)}\right)^{\mathrm{T}} \cdot \mathbf{C}^{-1} \cdot \left(\mathbf{T}(\boldsymbol{\theta}^{(k)}) - \boldsymbol{x}^{(k)}\right) . \tag{6}$$

Here, k is the replica number, $\boldsymbol{x}^{(k)}$ denotes pseudo-data for replica k which are obtained by randomly shifting the central values proportional to their uncertainties, and $\boldsymbol{\theta}$ represents the set of weights and biases in the neural network, $\mathbf{C}$ represents the covariance matrix including all uncertainties and (possible) correlations of the measured data $\boldsymbol{x}$, and finally $\mathbf{T}$ represents the theoretical prediction. We use the open source framework `MontBlanc` in all the analyses performed in this work. In this framework, minimization of χ^2 is performed by the `Ceres Solver` open source code [41]. The NN parameterization is provided by the `NNAD` code [42]. The core of our theoretical predictions relies on the `APFEL++` package [43]. Using `NNAD` an analytic expression for χ^2 is formed and `Ceres Solver` is used to solve the corresponding minimization problem using trust region methods. The FFs for each parton flavor are expressed through a NN model at the initial scale $Q_0 = 5$ GeV as:

$$z D_i^{h^+}(z, Q_0) = \left(N_i(z;\boldsymbol{\theta}) - N_i(1;\boldsymbol{\theta})\right)^2 . \tag{7}$$

$N_i(z;\boldsymbol{\theta})$ denotes the output generated by the neural network. We opt not to incorporate pre-factors like $z^\alpha(1-z)^\beta$ to regulate the behavior at low and high values of z. However, it is important to note that the NN output is subtracted at $z = 1$, ensuring that the fragmentation functions vanish at this specific kinematic point. Moreover, the resulting output is squared to guarantee that FFs remain non-negative.

The neural network is a type of multi-layer feed-forward perceptron [44]. It comprises nodes organized into sequential layers, where the output $\xi_i^{(l)}$ of the i-th node in the l-th layer is,

$$\xi_i^{(l)} = g\left(\sum_j \omega_{ij}^{(l)} \xi_j^{(l-1)} + \theta_i^{(l)}\right) . \tag{8}$$

The function g is called *activation function*, which is a sigmoid function for hidden layers and linear for the output layer. The parameters $\left\{\omega_{ij}^{(l)}, \theta_i^{(l)}\right\}$, denote *weights* and *biases* respectively, and are determined by minimization of the χ^2 function.

We use a simple but effective method in this study. Its effectiveness is based on the so-called universal approximation theorem [45], which states that a simple feed-forward NN like the one we are using can represent any function in any specified range. Our NN setup has just one hidden layer, and we pick 25 nodes for it. We have one node in the input layer, which is the value of z, and 6 output nodes, one for each combination of FFs in Eq. (5). This architecture is therefore denoted as (1-25-6) and has 175 weight and 31 bias parameters, i.e. 206 parameters in total to be fitted to the data. Fewer nodes could work just as well. We have tried a NN setup with two hidden layers, (1-9-9-6), to see if this would change our results. Unsurprisingly, our tests showed that the results are consistent, as other research supports this idea [46].

We use cross-validation to avoid overfitting our FFs. For each data replica, the datasets with more than 10 points are randomly split into training and validation subsets, each containing half of the points. We only use the training group to perform the fit. If there are 10 or fewer points, we use all of them for training. During the training step, the χ^2-value of the validation set is monitored. The fit is stopped when χ^2 of the validation set reaches its minimum. Replicas whose total χ^2 per point is larger than three are discarded. We are therefore convinced that our results do indeed describe the data, rather than the specific setup of the NN. Finally, we decided to use 100 replicas for our analysis, as it was shown in a previous study to be enough to represent the central values and uncertainties of the FFs [47].

The Monte Carlo approach has become a prevalent technique in various QCD analyses for error propagation, as evidenced by its application in studies such as [25, 48–51]. Here, this method involves estimating the posterior probability distribution of NN parameters through fitting multiple times. Each fit is conducted independently using a pseudo dataset, known as a replica, resulting in an optimal set of parameters and therefore FFs. Subsequently, the results of all fits collectively encapsulate the probability distribution of the data in FFs. Therefore, we define both the central value as the mean of the replicas, and the uncertainties of FFs as the standard deviation over the replica set.

V. RESULTS AND DISCUSSION

The main results and findings of this study are presented and discussed in detail in this section. First, we evaluate the quality of the FF24-K_S^0 fit by examining the total and individual values of χ^2 per dataset. Second, we perform a comparison between the analyzed experimental data and the predictions obtained using our K_S^0 FFs for all the datasets included in this analysis. The kinematical cuts are also discussed in this section. Finally, we compare our FFs with other determinations available in the literature and illustrate the origin of some specific features of our new FFs.

The experimental data used in this analysis are collected in Table I. The table includes the reference for each dataset and the associated energy scale. Additionally, details about the kinematic cuts for both small and large regions of z are provided. The last two columns show the χ^2 values per data point for each dataset at both NLO and NNLO accuracy. The table also presents the total χ^2 values for all the SIA datasets.

The analysis includes a total of 346 data points after the application of the kinematical cuts described above. The overall χ^2 per data point is 0.91 for the NLO and 0.87 for the NNLO analysis, showcasing a highly effective description of the entire dataset for both perturbative orders.

A detailed examination of the entries in Table I reveals a consistent fit quality and satisfactory description for each individual dataset. It is noteworthy that the incorporation of NNLO corrections results in a reduction of χ^2 per data point for most of the datasets, contributing to a further enhancement of the overall fitting quality indicated by the total χ^2.

In the following, we compare the theoretical predictions calculated using the FF24-K_S^0 FFs with the analyzed experimental data for all the datasets, focusing on NNLO accuracy. In Figs. 3, 4, and 5, we present the predictions for the normalized differential cross section at NNLO accuracy as a function of z, compared with the SIA data. The upper panels of the figures show the actual cross sections from experiment (black points) and theory (red points). In the lower panels of Figs. 3-5, the black points show the ratios of data over theory for each dataset, allowing for a better comparison of the central values of our theoretical predictions at NNLO accuracy with experimental data. The error bars of the black (red) points are obtained by normalizing the experimental (theoretical) uncertainties to the central values of our theory prediction.

Some of the data points have been excluded from the fits, as explained earlier. In the figures, these points are separated by a blue dotted line. Specifically, we have omitted two data points from ALEPH and four data points from DELPHI in the small z region due to their tension with the rest of the data. Additionally, we observed outliers in the last data points of BESIII2.2324 and BESIII2.4, which deviate from the overall trend. Consequently, we decided to exclude them from our analysis as well. All data points from the remaining datasets are included in our analysis. The χ^2 values listed in Table I demonstrate that our predictions are in good agreement with the data. A detailed comparison of the K_S^0 production cross section calculated at NNLO accuracy and the newly added BESIII data are presented in Fig. 5. As can clearly be seen, very nice agreement is achieved, which is consistent with the χ^2 values presented in Table I. The incorporation of BESIII data contributes significantly to the precision and reliability of our new NNLO FF set.

It is worth noting that when datasets are consistent with each other, one would typically expect the predic-

Experiment	Reference	$\sqrt{s}$ GeV	z_{min}	z_{max}	$\chi^2_{\texttt{NLO}}$/#data	$\chi^2_{\texttt{NNLO}}$/#data
BESIII	[23]	2.2324	0.013	0.8	0.92	0.94
BESIII	[23]	2.4	0.013	0.8	0.14	0.20
BESIII	[23]	2.8	0.013	0.9	2.55	2.18
BESIII	[23]	3.05	0.013	0.9	1.37	1.39
BESIII	[23]	3.4	0.013	0.9	0.84	0.85
BESIII	[23]	3.6710	0.013	0.9	0.55	0.56
TASSO	[27]	14	0.013	0.9	0.68	0.63
TASSO	[28]	14.8	0.013	0.9	1.60	1.55
TASSO	[28]	21.5	0.013	0.9	1.19	1.07
TASSO	[27]	22	0.013	0.9	1.27	1.17
TPC	[30]	29	0.013	0.9	0.23	0.22
MARKII	[31]	29	0.013	0.9	0.37	0.33
TASSO	[27]	34	0.013	0.9	2.04	1.83
TASSO	[28]	34.5	0.013	0.9	1.27	1.21
TASSO	[28]	35	0.013	0.9	0.92	0.85
CELLO	[32]	35	0.013	0.9	0.40	0.35
TASSO	[28]	42.6	0.013	0.9	1.19	1.15
TOPAZ	[33]	58	0.013	0.9	0.34	0.32
ALEPH	[34]	91.2	0.013	0.9	0.35	0.32
DELPHI	[35]	91.2	0.013	0.9	0.70	0.73
OPAL	[36]	91.2	0.013	0.9	0.90	0.88
SLD total	[37]	91.2	0.013	0.9	0.93	0.93
SLD uds	[37]	91.2	0.013	0.9	0.65	0.73
SLD charm	[37]	91.2	0.013	0.9	0.64	0.66
SLD bottom	[37]	91.2	0.013	0.9	1.64	1.46
				Total #data		
Total χ^2/#data				346	0.91	0.87

Table I: The list of input datasets included in the analysis of K^0_S FFs at NLO and NNLO accuracy. For each dataset, we provide the name of the experiment, the corresponding published reference, the center-of-mass energy $\sqrt{s}$, and the value of χ^2 per data point for the individual dataset at NLO and NNLO accuracy. The total value of χ^2 per data point is shown in the last line.

tions to have uncertainties of a similar magnitude as those of the experimental data. However, if there are inconsistencies in the data, uncertainties may decrease to account for this discrepancy. This phenomenon, where the combination of multiple datasets leads to a reduction in the overall uncertainty, is evident in Figs. 3, 4, and 5. Similar observations can be made in Figure 2 of [24] and in our previous analysis using MontBlanc [4].

Now we present our K^0_S FFs, denoted as FF24-K^0_S, obtained from our baseline NLO and NNLO QCD fits and compare them with some other FF parametrizations available in the literature. Such comparisons are shown in Figs. 6 and 7. For the NLO analysis, we compare with SAK20 [16], AKK08 [20], and DSS17 [21]. Results from SAK20 [16] and NNFF1.0 [22] are compared with our NNLO fit results.

All four FF sets depicted in Fig. 6 include a comparable array of SIA datasets. However, our new FF24-K^0_S set uniquely integrates the BESIII measurements for the first time. Additionally, the AKK08 FF encompasses K^0_S hadron production measurements from $pp(\bar{p})$ collisions. With the exception of DSS17, hadron mass corrections are accounted for in the other three FF sets. Both the FF24-K^0_S and SAK20 FFs are computed up to NNLO ac-

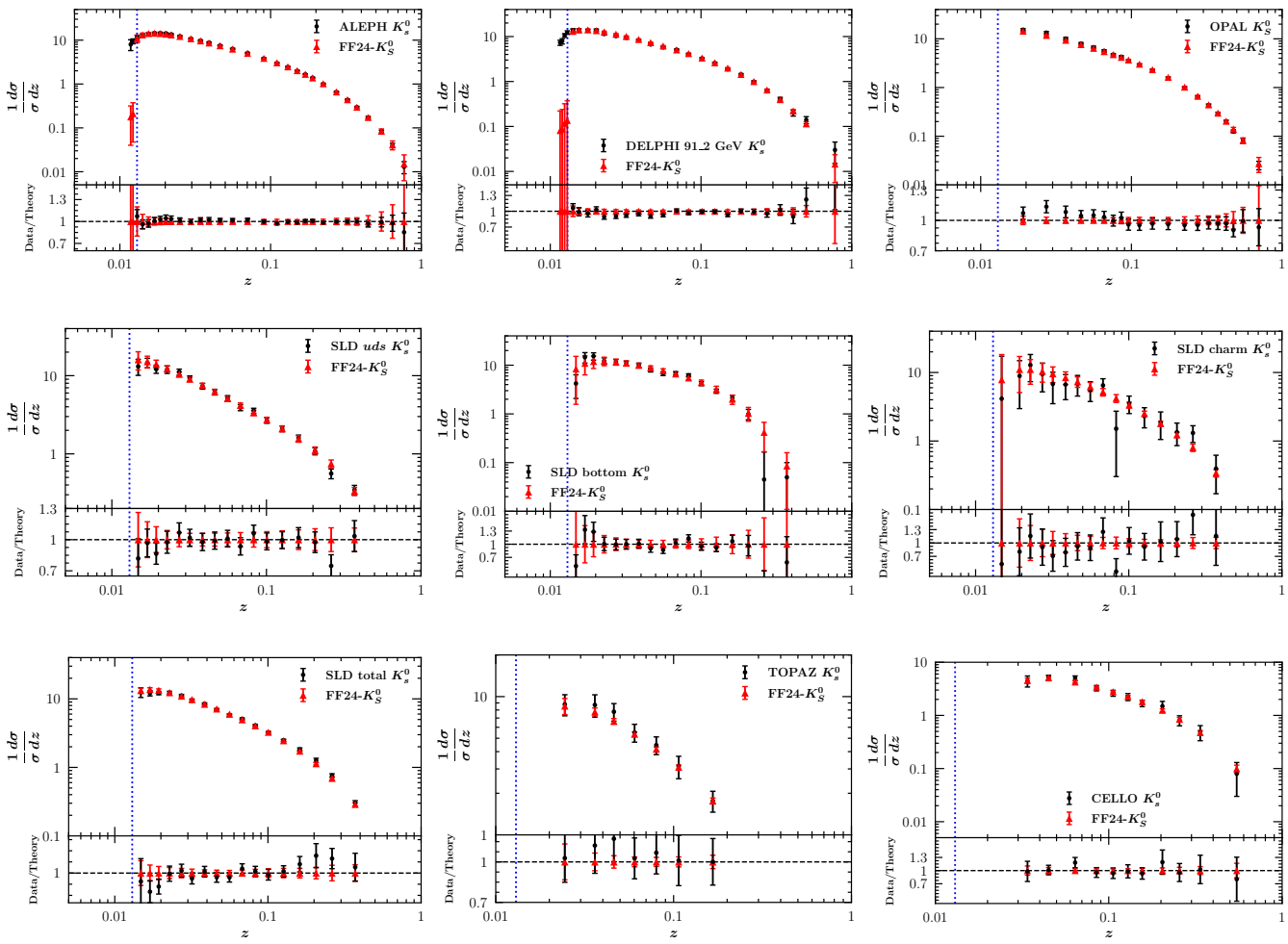

Figure 3: Comparison between the K_S^0 production cross section calculated at NNLO accuracy and experimental data from different experiments, including ALEPH, DELPHI91.2, OPAL, SLD_{uds}, SLD_b, SLD_c, $\text{SLD}_{\text{total}}$, TOPAZ, and CELLO. The lower panels show the data/theory ratios.

curacy. In contrast, the AKK08 FFs are determined only up to NLO accuracy and available only without an uncertainty estimate. DSS17 [21] also presented their FFs at NLO accuracy along with uncertainty bands extracted from SIA, lepton-nucleon DIS, and proton-proton collision data. The minimum value for z is considered to be 0.05 in both the SAK20 and AKK08 analyses. However, in the FF24-K_S^0 analysis, we incorporated data points also for smaller z, down to $z \geq 0.013$. This difference could result in some variations among the four FF analyses for all flavors, in particular in the small-z region, $z < 0.05$.

Considering the comparisons presented in Fig. 6, several observations are noteworthy. It is evident that the central values of these analyses exhibit different behaviors, particularly for the gluon and u^+ FFs, even within regions defined by the same kinematical cuts. Despite these differences, some similarities emerge, such as at large z values for the b^+, c^+, s^+, and d^+ FFs. In the small z region, where more data points are included in the FF24-K_S^0 analysis, the FFs are generally larger than those of the SAK20 FF, except for the gluon FF.

Regarding uncertainty bands, there are some comments need to be made as well. As shown in Fig. 6, the uncertainty bands for FF24-K_S^0 FFs are broader compared to those of the SAK20 analysis across nearly all the z region. This difference primarily arises from our utilization of the Monte Carlo method to accommodate error bands, whereas the SAK20 analysis employs the standard Hessian method. Notably, the uncertainty bands for the new FF24-K_S^0 FFs are wide over regions at small values of z due to the insufficient coverage of SIA data points in those areas. These wide uncertainty bands are particu-

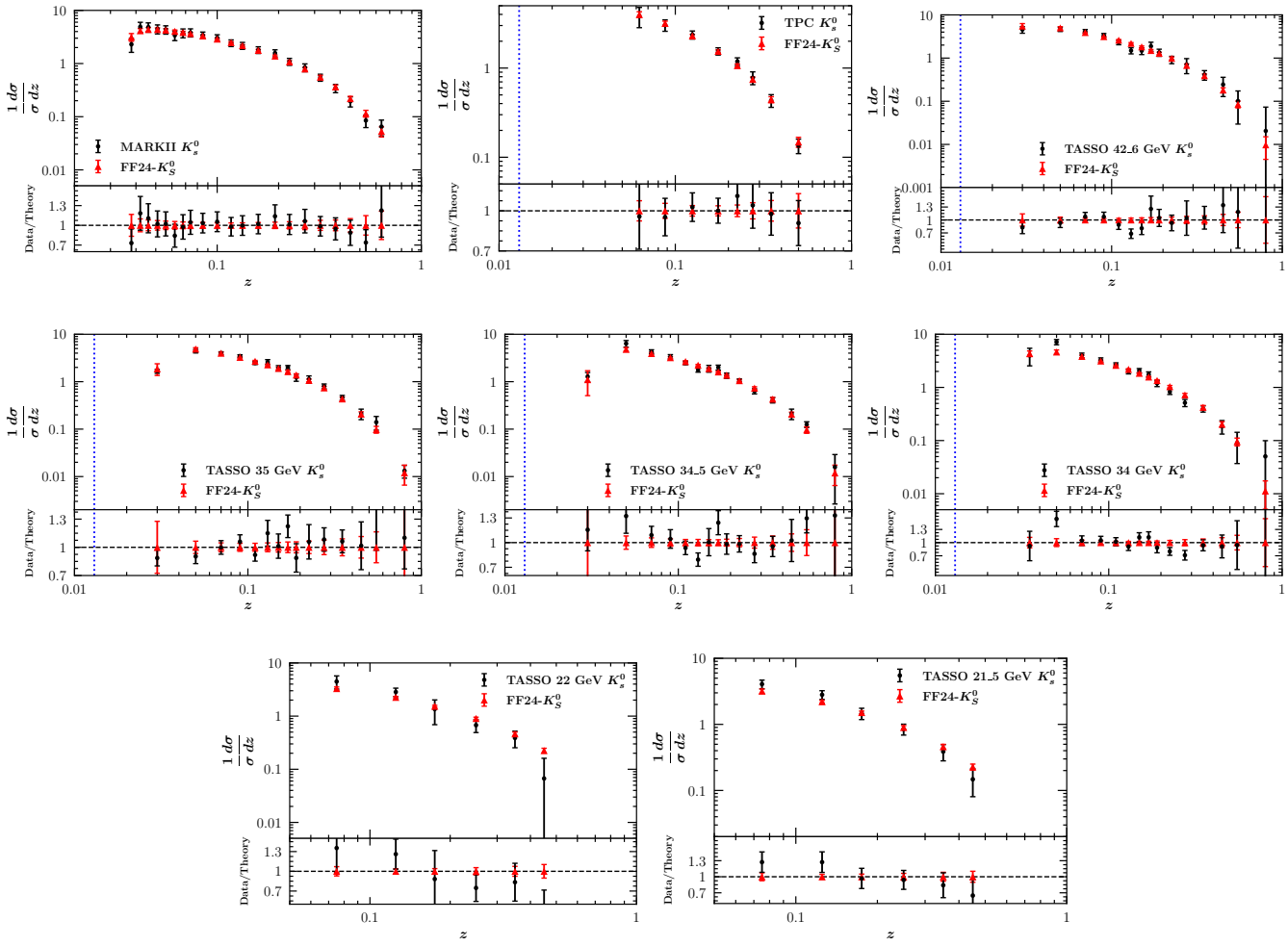

Figure 4: Same as Fig. 3 but this time for some other selected experimental data, including `MARKII`, `TPC`, `TASSO42.6`, `TASSO35`, `TASSO34.5`, `TASSO34`, `TASSO22`, `TASSO21.5`.

larly pronounced for the u^+ and s^+ FFs. Upon comparison with `DSS17`, it becomes evident that their analysis yields larger error bands across nearly all FFs, except for the d^+ FFs.

We now shift our focus to the comparison of the `FF24-`K^0_S at NNLO accuracy with those of `SAK20` [16] and `NNFF1.0` [22]. These comparisons are illustrated in Fig. 7. Considering the central values of these analyses, one can observe different behaviors for all parton species, even within regions defined by the same kinematical cuts. For the case of b^+, c^+, and s^+ FFs, the central values of `FF24-`K^0_S are smaller than those of `SAK20` [16] and `NNFF1.0` [22]. However, for the u^+ and gluon FFs, although they are larger in comparison to `SAK20`, they are still smaller than those of `NNFF1.0`. In terms of uncertainty bands, it is noticeable that the error bands of `FF24-`K^0_S are smaller than those of `NNFF1.0` for almost all parton species and across all regions of z, particularly for the small-z region $z < 0.1$. When comparing with `SAK20`, it can be concluded that smaller error bands are achieved, except for the case of d^+ and s^+ FFs.

In order to see how the `BESIII` data for K^0_S production affect the determination of FFs, we compare in Fig. 8 data with predictions obtained from both the previous `SAK20` FFs and the new `FF24-`K^0_S FFs. We observe large differences between data and `SAK20` predictions over the full range of z and for all energies. The theory predictions based on `SAK20` FFs are not in good agreement with the `BESIII` data. In particular, the increase of cross section predictions from `SAK20` towards low values of z is not seen in the data. We believe that the main reason for this difference is that the `SAK20` FFs had been extracted from high-energy SIA data and the predictions using QCD-backward evolution of FFs from high to low

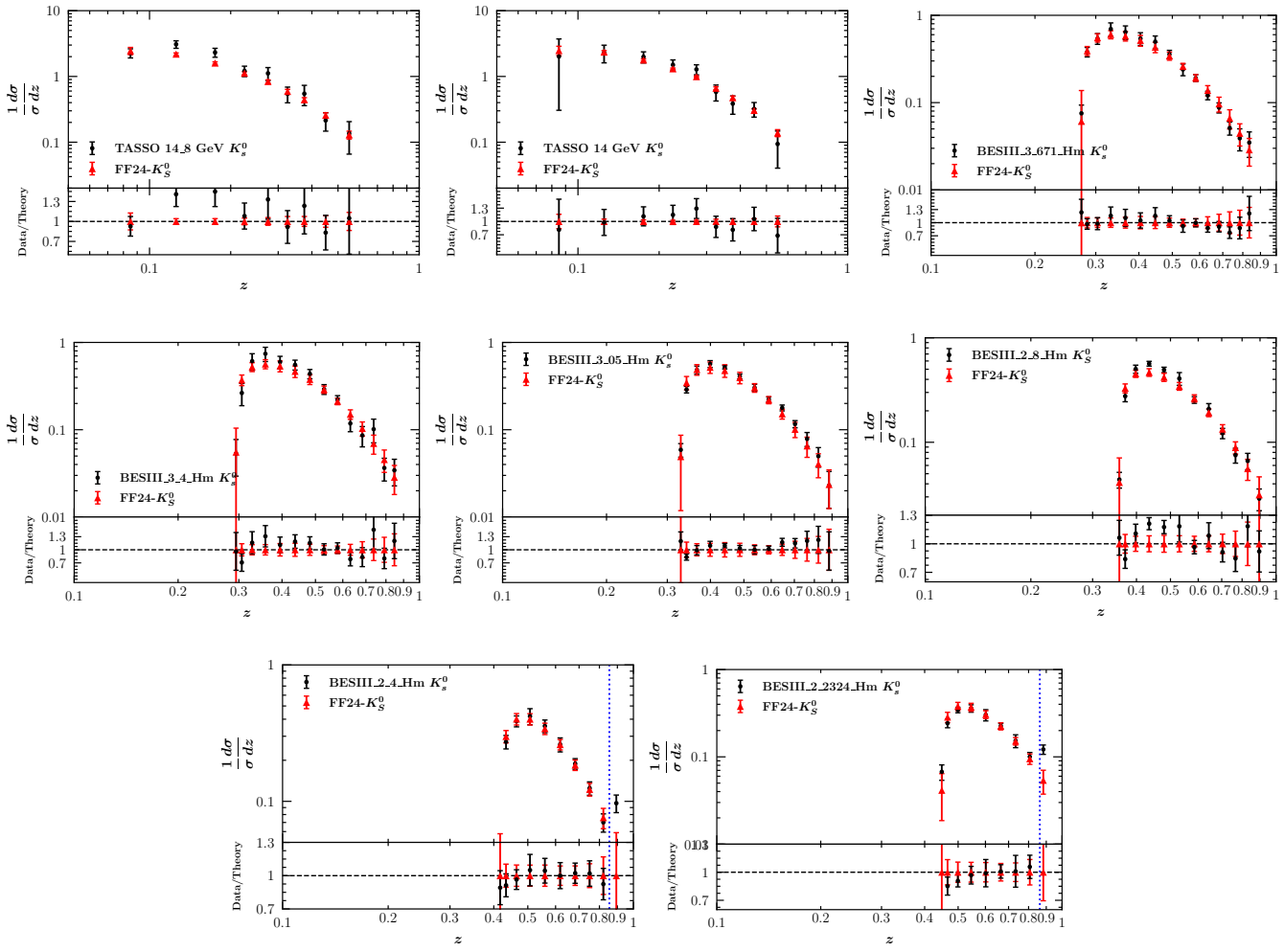

Figure 5: Same as Fig. 3 but for some other selected experimental data including `TASSO14.8`, `TASSO14`, `BESIII3.671`, `BESIII3.4`, `BESIII3.05`, `BESIII2.8`, `BESIII2.4` and `BESIII2.2324`.

energies do not match the experimental results. We conclude that including `BESIII` data improve the fit for FFs in particular at low z.

In summary, our approach highlights two key improvements: First, the integration of the latest experimental data from `BESIII`, which enriches the dataset and enhances the reliability of our results. Second, the adoption of neural networks in the fitting procedure introduces a novel technique to refine the extraction of FFs. This innovative combination of advanced methodologies contributes to a more accurate determination of K^0_S FFs, offering improved precision and reliability for the prediction of measurements at future high-energy scattering experiments.

VI. SUMMARY AND OUTLOOK

In summary, the main goal of this paper is to introduce new sets of fragmentation functions for K^0_S, entitled `FF24-`K^0_S, obtained from a QCD analysis including NNLO corrections. Our analysis relies on extensive experimental data, incorporating measurements of K^0_S production in the SIA process and including the most recent experimental data reported by the `BESIII` Collaboration.

The `BESIII` measurement covers collision energies spanning from $\sqrt{s} = 2.2324$ to $\sqrt{s} = 3.6710$ GeV below the range used in earlier FF fits. The findings presented in this study hold particular significance as they contribute valuable insight, especially in this new energy range below $\sqrt{s} < 10$ GeV where precise e^+e^- annihilation data have been scarce before. By analyzing normalized differential cross sections across this energy range,

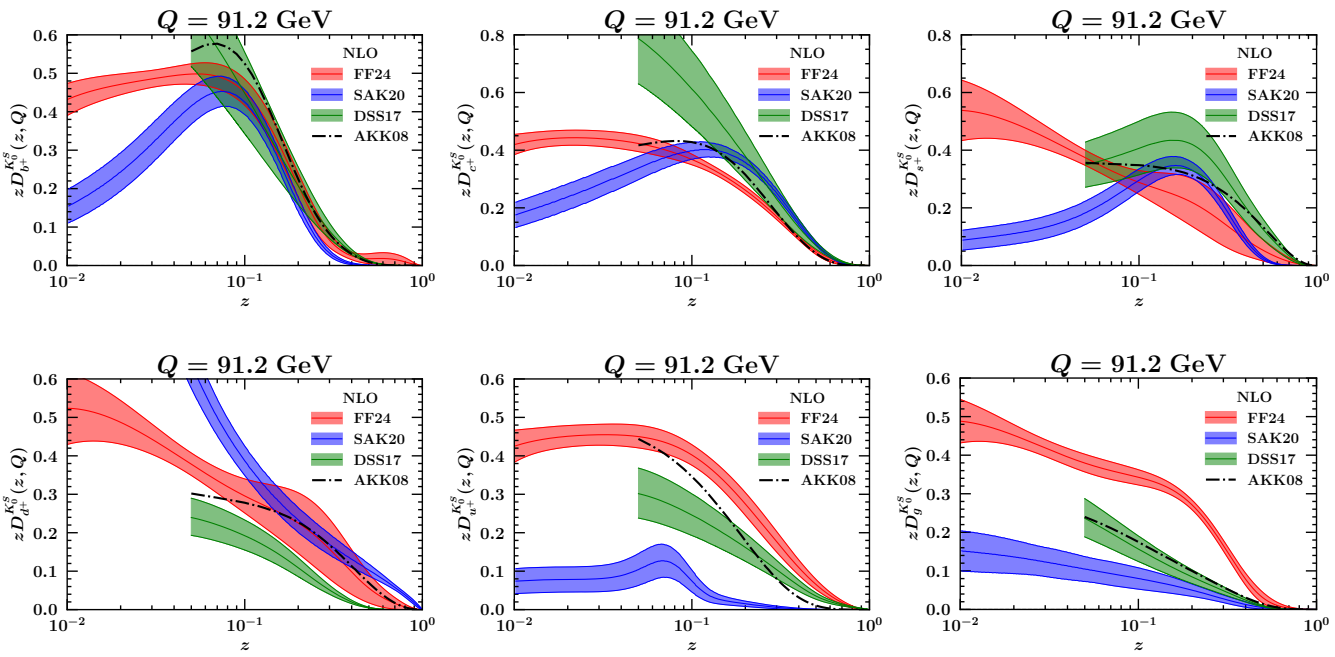

Figure 6: The FF24-K_S^0 at NLO accuracy obtained for various partons at $\sqrt{s} = 91.2$ GeV. The shaded bands represent 1σ uncertainty estimates derived from the Monte Carlo method. Results from SAK20 [16], DSS17 [21], and AKK08 [20] at NLO accuracy, along with their 1σ error bands, are also included for comparison.

we have successfully addressed critical gaps in our understanding of K_S^0 hadron production.

Incorporating hadron mass corrections particularly in datasets with low values of energy and z, is crucial. In our analyses, we have applied K_S^0 mass corrections which allowed us to incorporate low-z data points, excluding only 8 data points with $z_{\min} = 0.013$. While small-z resummation is not explicitly addressed in our present analysis, we acknowledge its importance and plan to explore it in future investigations.

We have utilized neural networks to parameterize the fragmentation functions, aiming to minimize theoretical bias. The NN parameters are tuned to fit the SIA data. Our approach used the recent publicly available package MontBlanc for FF parametrization, evolution, and SIA production cross-section calculation. Moreover, we employ a Monte Carlo sampling method to propagate experimental uncertainties into the fitted FFs and determine uncertainties for both FFs and associated observables.

Both the overall χ^2 and the χ^2 values for individual datasets used in the fitting procedure are satisfactory. This is evident in the excellent agreement observed between the experimental data and the corresponding theoretical predictions. However, upon comparison with two other determinations, namely SAK20 and AKK08, some notable differences were identified in the resulting FF24-K_S^0 set. In this work, we focus exclusively on SIA observables to accurately determine the quark FFs of K_S^0. However, these measurements provide limited information about the gluon-to-hadron FFs since they are only accessible starting at NLO. Consequently, significant differences arise between our findings and those of other results available in the literature, such as AKK08, where K_S^0 production was investigated using also collider data. To gain a better understanding of the gluonic contribution, we plan to extend our analysis to include data from the production of K_S^0 particles in proton collisions. Previous studies suggest that this will be particularly important for a better determination of the role of gluons in the overall picture.

The resulting FF24-K_S^0 sets at NLO and NNLO accuracy, along with their associated uncertainties, are made available in the standard LHAPDF format for broader utilization in the scientific community[52].

ACKNOWLEDGMENTS

The authors gratefully acknowledge many helpful discussions and comments by Valerio Bertone that elevated and enriched the paper significantly. Hamzeh Khanpour, Maryam Soleymaninia and Hadi Hashamipour thank the School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM) for financial sup-

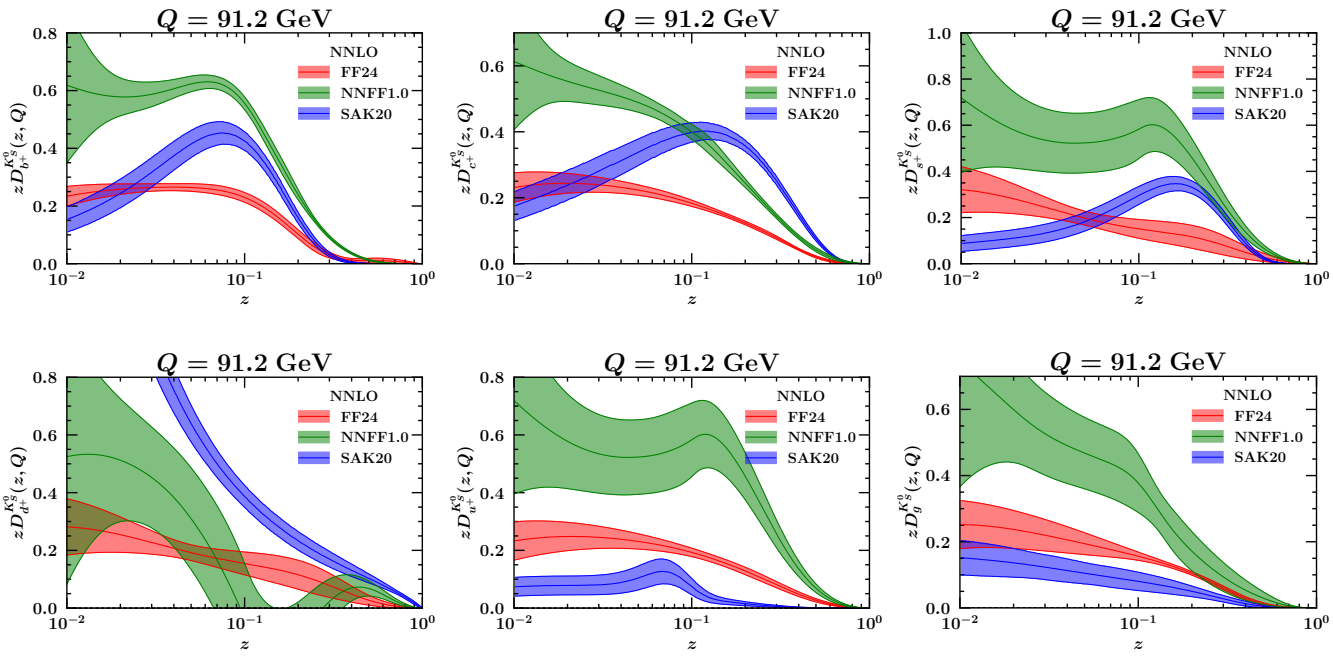

Figure 7: Same as Fig. 6, but this time for the `FF24`-K_S^0 at NNLO accuracy. Results from `SAK20` [16] and `NNFF1.0` [22], along with their error bands at NNLO accuracy, are also included for comparison.

port of this project. Hamzeh Khanpour also appreciates the financial support from NAWA under grant number BPN/ULM/2023/1/00160 and from the IDUB programme at the AGH University. This work was also supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the funds provided to the Sino-German Collaborative Research Center TRR110 "Symmetries and the Emergence of Structure in QCD" (DFG Project-ID 196253076 - TRR 110). The work of UGM was also supported in part by the Chinese Academy of Sciences (CAS) President's International Fellowship Initiative (PIFI) (Grant No. 2018DM0034).

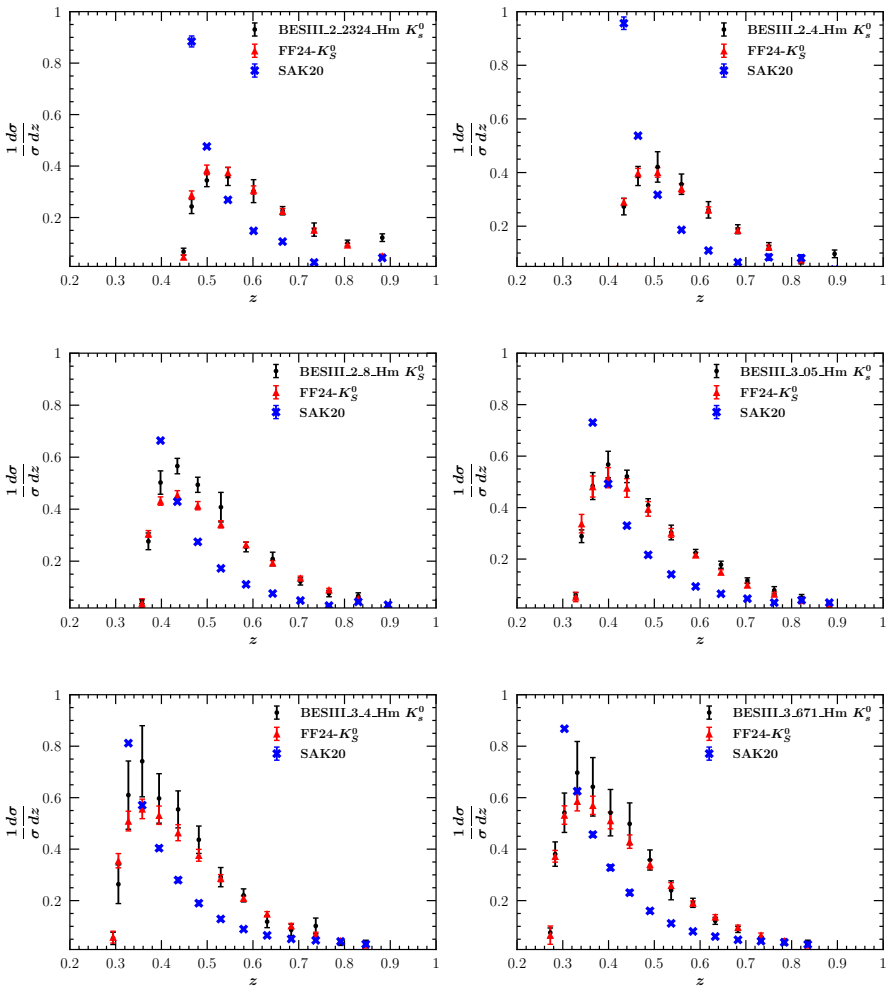

Figure 8: Comparison between the predictions for cross sections of K_S^0 production from `SAK20` [16] and the present analysis at NLO accuracy and the measurements from the `BESIII` experiment.

[1] J.C. Collins, D.E. Soper, G.F. Sterman, Adv. Ser. Dir. High Energy Phys. 5, 1 (1989).

[2] A. A. Almasy, S. Moch and A. Vogt, "On the Next-to-Next-to-Leading Order Evolution of Flavour-Singlet Fragmentation Functions," Nucl. Phys. B **854**, 133-152 (2012) doi:10.1016/j.nuclphysb.2011.08.028 [arXiv:1107.2263 [hep-ph]].

[3] A. Mitov, S. Moch and A. Vogt, "Next-to-Next-to-Leading Order Evolution of Non-Singlet Fragmentation Functions," Phys. Lett. B **638**, 61-67 (2006) doi:10.1016/j.physletb.2006.05.005 [arXiv:hep-ph/0604053 [hep-ph]].

[4] M. Soleymaninia, H. Hashamipour and H. Khanpour, "Neural network QCD analysis of charged hadron fragmentation functions in the presence of SIDIS data," Phys. Rev. D **105**, no.11, 114018 (2022) doi:10.1103/PhysRevD.105.114018 [arXiv:2202.10779 [hep-ph]].

[5] R. Abdul Khalek *et al.* [MAP (Multi-dimensional Analyses of Partonic distributions)], "Pion and kaon fragmentation functions at next-to-next-to-leading order," Phys. Lett. B **834**, 137456 (2022) doi:10.1016/j.physletb.2022.137456 [arXiv:2204.10331 [hep-ph]].

[6] I. Borsa, R. Sassot, D. de Florian, M. Stratmann and W. Vogelsang, "Towards a Global QCD Analysis of Fragmentation Functions at Next-to-Next-to-Leading Order Accuracy," Phys. Rev. Lett. **129**, no.1, 012002 (2022) doi:10.1103/PhysRevLett.129.012002 [arXiv:2202.05060 [hep-ph]].

[7] E. Moffat *et al.* [Jefferson Lab Angular Momentum (JAM)], "Simultaneous Monte Carlo analysis of parton densities and fragmentation functions," Phys. Rev. D **104**, no.1, 016015 (2021) doi:10.1103/PhysRevD.104.016015 [arXiv:2101.04664 [hep-ph]].

[8] A. Abada *et al.* [FCC], "FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1," Eur. Phys. J. C **79**, no.6, 474 (2019) doi:10.1140/epjc/s10052-019-6904-3

[9] A. Abada *et al.* [FCC], "HE-LHC: The High-Energy Large Hadron Collider: Future Circular Collider Conceptual Design Report Volume 4," Eur. Phys. J. ST **228**, no.5, 1109-1382 (2019) doi:10.1140/epjst/e2019-900088-6

[10] R. Abdul Khalek, A. Accardi, J. Adam, D. Adamiak, W. Akers, M. Albaladejo, A. Al-bataineh, M. G. Alexeev, F. Ameli and P. Antonioli, *et al.* "Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report," Nucl. Phys. A **1026**, 122447 (2022) doi:10.1016/j.nuclphysa.2022.122447 [arXiv:2103.05419 [physics.ins-det]].

[11] I. Borsa, G. Lucero, R. Sassot, E. C. Aschenauer and A. S. Nunes, "Revisiting helicity parton distributions at a future electron-ion collider," Phys. Rev. D **102**, no.9, 094018 (2020) doi:10.1103/PhysRevD.102.094018 [arXiv:2007.08300 [hep-ph]].

[12] M. Klasen and H. Paukkunen, "Nuclear PDFs After the First Decade of LHC Data," doi:10.1146/annurev-nucl-102122-022747 [arXiv:2311.00450 [hep-ph]].

[13] G. Aad *et al.* [ATLAS], "Measurement of K^0_S and Λ^0 production in $t\bar{t}$ dileptonic events in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector," Eur. Phys. J. C **79**, no.12, 1017 (2019) doi:10.1140/epjc/s10052-019-7512-y [arXiv:1907.10862 [hep-ex]].

[14] G. Aad *et al.* [ATLAS], "Kshort and Λ production in pp interactions at $\sqrt{s}$ = 0.9 and 7 TeV measured with the ATLAS detector at the LHC," Phys. Rev. D **85**, 012001 (2012) doi:10.1103/PhysRevD.85.012001 [arXiv:1111.1297 [hep-ex]].

[15] D. Acosta *et al.* [CDF], "K^0_S and Λ^0 production studies in $p\bar{p}$ collisions at $\sqrt{s}$ = 1800-GeV and 630-GeV," Phys. Rev. D **72**, 052001 (2005) doi:10.1103/PhysRevD.72.052001 [arXiv:hep-ex/0504048 [hep-ex]].

[16] M. Soleymaninia, H. Abdolmaleki and H. Khanpour, "First NNLO fragmentation functions of K^0_S and $\Lambda/\bar{\Lambda}$ and their uncertainties in the presence of hadron mass corrections," Phys. Rev. D **102**, no.11, 114029 (2020) doi:10.1103/PhysRevD.102.114029 [arXiv:2009.08139 [hep-ph]].

[17] J. Binnewies, B. A. Kniehl and G. Kramer, "Neutral kaon production in e^+e^-, ep and $p\bar{p}$ collisions at next-to-leading order," Phys. Rev. D **53**, 3573 (1996), [hep-ph/9506437].

[18] C. Bourrely and J. Soffer, "Statistical approach for unpolarized fragmentation functions for the octet baryons," Phys. Rev. D **68**, 014003 (2003), [hep-ph/0305070].

[19] S. Albino, B. A. Kniehl and G. Kramer, "Fragmentation functions for K0(S) and Lambda with complete quark flavor separation," Nucl. Phys. B **734**, 50 (2006), [hep-ph/0510173].

[20] S. Albino, B. A. Kniehl and G. Kramer, "AKK Update: Improvements from New Theoretical Input and Experimental Data," Nucl. Phys. B **803**, 42 (2008), [arXiv:0803.2768 [hep-ph]].

[21] D. de Florian, M. Epele, R. J. Hernandez-Pinto, R. Sassot and M. Stratmann, Phys. Rev. D **95**, no.9, 094019 (2017) doi:10.1103/PhysRevD.95.094019 [arXiv:1702.06353 [hep-ph]].

[22] V. Bertone *et al.* [NNPDF], "A determination of the fragmentation functions of pions, kaons, and protons with faithful uncertainties," Eur. Phys. J. C **77**, no.8, 516 (2017).

[23] M. Ablikim *et al.* [BESIII], "Measurements of Normalized Differential Cross Sections of Inclusive π0 and KS0 Production in e+e- Annihilation at Energies from 2.2324 to 3.6710 GeV," Phys. Rev. Lett. **130**, no.23, 231901 (2023) doi:10.1103/PhysRevLett.130.231901 [arXiv:2211.11253 [hep-ex]].

[24] R. A. Khalek, V. Bertone and E. R. Nocera, "Determination of unpolarized pion fragmentation functions using semi-inclusive deep-inelastic-scattering data," Phys. Rev. D **104**, no.3, 034007 (2021).

[25] R. D. Ball *et al.* [NNPDF], "Parton distributions from high-precision collider data," Eur. Phys. J. C **77**, no.10, 663 (2017).

[26] https://github.com/MapCollaboration/MontBlanc

[27] M. Althoff *et al.* [TASSO Collaboration], "A Detailed Study of Strange Particle Production in e^+e^- Annihilation at High-energy," Z. Phys. C **27**, 27 (1985).

[28] W. Braunschweig *et al.* [TASSO Collaboration], "Strange Meson Production in e^+e^- Annihilation," Z. Phys. C **47**, 167 (1990).

[29] M. Derrick *et al.*, "Hadron Production in e^+e^- Annihilation at $\sqrt{s} = 29$-GeV," Phys. Rev. D **35**, 2639 (1987).

[30] H. Aihara *et al.* [TPC/Two Gamma Collaboration], "K^{*0} and K^0_s meson production in e^+e^- annihilations at 29-GeV," Phys. Rev. Lett. **53**, 2378 (1984).

[31] H. Schellman *et al.*, "Measurement of $K^{\pm}$ and K^0 Inclusive Rates in e^+e^- Annihilation at 29-GeV," Phys. Rev. D **31**, 3013 (1985).

[32] H. J. Behrend *et al.* [CELLO Collaboration], "Inclusive Strange Particle Production in e^+e^- Annihilation," Z. Phys. C **46**, 397 (1990).

[33] R. Itoh *et al.* [TOPAZ Collaboration], "Measurement of inclusive particle spectra and test of MLLA prediction in e+ e- annihilation at s**(1/2) = 58-GeV," Phys. Lett. B **345**, 335 (1995), [hep-ex/9412015].

[34] R. Barate *et al.* [ALEPH Collaboration], "Studies of quantum chromodynamics with the ALEPH detector," Phys. Rept. **294**, 1 (1998).

[35] P. Abreu *et al.* [DELPHI Collaboration], "Production characteristics of K0 and light meson resonances in hadronic decays of the Z0," Z. Phys. C **65**, 587 (1995).

[36] G. Abbiendi *et al.* [OPAL Collaboration], "Leading particle production in light flavor jets," Eur. Phys. J. C **16**, 407 (2000), [hep-ex/0001054].

[37] K. Abe *et al.* [SLD Collaboration], "Production of pi+, K+, K0, K*0, phi, p and Lambda0 in hadronic Z0 decays," Phys. Rev. D **59**, 052001 (1999), [hep-ex/9805029].

[38] P. Abreu *et al.* [DELPHI Collaboration], "Charged and identified particles in the hadronic decay of W bosons and in e+ e- —> q anti-q from 130-GeV to 200-GeV," Eur. Phys. J. C **18**, 203 (2000) Erratum: [Eur. Phys. J. C **25**, 493 (2002)], [hep-ex/0103031].

[39] A. Mitov and S. O. Moch, "QCD Corrections to Semi-Inclusive Hadron Production in Electron-Positron Annihilation at Two Loops," Nucl. Phys. B **751**, 18-52 (2006) doi:10.1016/j.nuclphysb.2006.05.018 [arXiv:hep-ph/0604160 [hep-ph]].

[40] 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**, 641-653 (1977)

[41] S. Agarwal, K. Mierle, and Others, "Ceres solver." `http://ceres-solver.org`.

[42] https://github.com/rabah-khalek/NNAD

[43] V. Bertone, APFEL++: A new PDF evolution library in C++, PoS **DIS2017**, 201 (2018).

[44] S. Forte, L. Garrido, J. I. Latorre and A. Piccione, JHEP **05**, 062 (2002) doi:10.1088/1126-6708/2002/05/062 [arXiv:hep-ph/0204232 [hep-ph]].

[45] B. C. Csáji et al., Approximation with artificial neural networks, MSc Thesis, Faculty of Sciences, Etvs Lornd University, Hungary (2001).

[46] R. A. Khalek, [arXiv:2110.01924 [hep-ph]].

[47] M. Soleymaninia, H. Hashamipour, H. Khanpour and H. Spiesberger, Nucl. Phys. A **1029**, 122564 (2023) doi:10.1016/j.nuclphysa.2022.122564 [arXiv:2202.05586 [hep-ph]].

[48] H. Moutarde, P. Sznajder and J. Wagner, "Unbiased determination of DVCS Compton Form Factors," Eur. Phys. J. C **79**, no.7, 614 (2019).

[49] 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**, no.11, 114004 (2016).

[50] R. Abdul Khalek *et al.* [NNPDF], "Nuclear parton distributions from lepton-nucleus scattering and the impact of an electron-ion collider," Eur. Phys. J. C **79**, no.6, 471 (2019).

[51] N. Sato *et al.* [Jefferson Lab Angular Momentum], "Iterative Monte Carlo analysis of spin-dependent parton distributions," Phys. Rev. D **93**, no.7, 074005 (2016).

[52] https://github.com/hashamipour/FF24_K0S_LHAPDF