

First simultaneous global QCD analysis of dihadron fragmentation functions and transversity parton distribution functions

C. Cocuzza¹,¹ A. Metz,¹ D. Pitonyak,² A. Prokudin,^{3,4} N. Sato,⁴ and R. Seidl⁵

(Jefferson Lab Angular Momentum (JAM) Collaboration)

¹*Department of Physics, SERC, Temple University, Philadelphia, Pennsylvania 19122, USA*

²*Department of Physics, Lebanon Valley College, Annville, Pennsylvania 17003, USA*

³*Division of Science, Penn State University Berks, Reading, Pennsylvania 19610, USA*

⁴*Jefferson Lab, Newport News, Virginia 23606, USA*

⁵*RIKEN BNL Research Center, Upton, New York 11973, USA*



(Received 5 September 2023; accepted 12 December 2023; published 27 February 2024)

We perform a comprehensive study within quantum chromodynamics (QCD) of dihadron observables in electron-positron annihilation, semi-inclusive deep-inelastic scattering, and proton-proton collisions, including recent cross section data from Belle and azimuthal asymmetries from STAR. We extract simultaneously for the first time $\pi^+\pi^-$ dihadron fragmentation functions (DiFFs) and the nucleon transversity distributions for up and down quarks as well as antiquarks. For the transversity distributions we impose their small- x asymptotic behavior and the Soffer bound. In addition, we utilize a new definition of DiFFs that has a number density interpretation to then calculate expectation values for the dihadron invariant mass and momentum fraction. Furthermore, we investigate the compatibility of our transversity results with those from single-hadron fragmentation (from a transverse momentum dependent/collinear twist-3 framework) and the nucleon tensor charges computed in lattice QCD. We find a universal nature to all of this available information. Future measurements of dihadron production can significantly further this research, especially, as we show, those that are sensitive to the region of large parton momentum fractions.

DOI: [10.1103/PhysRevD.109.034024](https://doi.org/10.1103/PhysRevD.109.034024)

I. INTRODUCTION

Phenomenological analyses of hadrons, at the level of quarks and gluons (partons), based on high-energy collision measurements rely on two key ingredients: parton distribution functions (PDFs) and fragmentation functions (FFs). The former provides insight into the momentum-space partonic structure of the incoming nucleon while the latter encodes how a parton forms multiple (colorless) particles in the final state. With regard to FFs, there are two common situations: experiments tagging on one hadron or on a hadron pair (dihadron) within a parton-initiated jet. The single-hadron unpolarized collinear FF $D_1^{h/i}(z)$ (h denoting the hadron type, i the parton, and z the momentum fraction carried by the hadron) has been studied in observables described by leading-twist collinear factorization and extracted by numerous groups over the last nearly four

decades [1–36]. When transverse momentum dependent (TMD) observables are considered, then, for unpolarized hadrons, not only is $D_1^{h/i}(z, z^2 \vec{k}_T^2)$ relevant ($\vec{k}_T$ giving the transverse momentum of the parton relative to the hadron), but also the chiral-odd Collins TMD FF $H_1^{\perp h/i}(z, z^2 \vec{k}_T^2)$ [37]. Both $D_1^{h/i}(z, z^2 \vec{k}_T^2)$ [38–45] and $H_1^{\perp h/i}(z, z^2 \vec{k}_T^2)$ [46–54] have been extracted using processes described by leading-power TMD factorization. We refer to Ref. [55] for a review on FFs.

On the dihadron side, the FFs depend on more variables [56,57], which can be chosen to be the dihadron total momentum fraction z , the relative momentum fraction ζ , the relative transverse momentum $\vec{K}_T$ (whose magnitude can be related to the invariant mass M_h of the dihadron [56,57]), along with the parton transverse momentum $\vec{k}_T$. Consequently, even after integrating over $\vec{k}_T$, two dihadron FFs (DiFFs) remain that can be studied within leading-twist collinear factorization [56–61]: $D_1^{h_1 h_2 / i}(z, M_h)$ and $H_1^{\triangleleft h_1 h_2 / i}(z, M_h)$, with the latter being chiral-odd and sometimes referred to as the interference FF (IFF). (Although both D_1 and $H_1^{\triangleleft}$ are DiFFs, we will often use the term IFF

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

to distinguish $H_1^{\triangleleft}$ from D_1 .) Unlike the single-hadron unpolarized (collinear and TMD) FFs, only one group has extracted these DiFFs [62,63], and only one measurement exists ($d\sigma/dz dM_h$ cross section from Belle for $e^+e^- \rightarrow (h_1 h_2)X$ [64], which postdates Refs. [62,63]) that is directly sensitive to $D_1^{h_1 h_2/i}(z, M_h)$. We note that a new definition of DiFFs has recently been introduced [65] that, in particular, allows $D_1^{h_1 h_2/i}(z, M_h)$ to retain a number density interpretation, from which one can meaningfully calculate expectation values.

With regard to PDFs, at leading twist the longitudinal momentum structure of the nucleon is characterized by the spin-averaged PDF $f_1(x)$, the helicity PDF $g_1(x)$, and the transversity PDF $h_1(x)$, with x being the momentum fraction carried by the parton. The transversity PDF quantifies the degree of transverse polarization of quarks within a transversely polarized nucleon. It is also used to calculate the tensor charges of the nucleon:

$$\delta u = \int_0^1 dx h_1^{u_v}(x; \mu), \quad \delta d = \int_0^1 dx h_1^{d_v}(x; \mu), \quad (1)$$

where $h_1^{q_v} \equiv h_1^q - h_1^{\bar{q}}$ are the valence distributions, and μ is the renormalization scale. In addition to phenomenological analyses of experimental data [46–54,63,66–69], the tensor charges can be found through *ab initio* computations in lattice QCD (LQCD) [70–81] and model calculations [82–93].

The chiral-odd nature of transversity makes it difficult to extract (relative to the spin-averaged and helicity PDFs) since it must couple to another chiral-odd function. Three possible candidates are the aforementioned Collins TMD FF $H_1^{\perp h/i}(z, z^2 \vec{k}_T^2)$ and IFF $H_1^{\triangleleft h_1 h_2/i}(z, M_h)$ as well as chiral-odd higher-twist (twist-3) collinear FFs [53,54,94–96]. Analyses that extract $h_1(x)$, or its TMD version $h_1(x, \vec{k}_T^2)$, have been performed using single-hadron TMD processes [46–52], both single-hadron TMD and collinear twist-3 observables [53,54], or dihadron reactions [63,66,68,69]. Specifically, transverse single-spin asymmetries can be investigated in electron-positron (e^+e^-) annihilation, semi-inclusive deep-inelastic scattering (SIDIS), and proton-proton (pp) collisions with transversely polarized nucleons and/or partons involved where either a single hadron or dihadron is detected in the final state. For the single-hadron case, simultaneous extractions of $h_1(x, \vec{k}_T^2)$ and $H_1^{\perp h/i}(z, z^2 \vec{k}_T^2)$ in e^+e^- and SIDIS have been carried out [46–52]. Moreover, global analyses that also included pp data and further extracted the chiral-odd collinear twist-3 FF $\tilde{H}^{h/i}(z)$ and Sivers TMD PDF $f_{1T}^{\perp}(x, \vec{k}_T^2)$ have been performed (JAM3D) [53,54]. For the dihadron case, $h_1(x)$, $H_1^{\triangleleft h_1 h_2/i}(z, M_h)$, and $D_1^{h_1 h_2/i}(z, M_h)$ have all been

extracted separately [62,63,66,68,69], but no group has considered a simultaneous extraction of the three functions using all available dihadron measurements/kinematic binnings.

Given the above discussion, the motivation for this work is the following:

- (1) Perform the first simultaneous global QCD analysis (JAMDiFF) of the $\pi^+\pi^-$ DiFFs and transversity PDFs (for up and down quarks and antiquarks) from e^+e^- annihilation, SIDIS, and pp data.
- (2) Include, for the first time, the Belle cross section data [64], the latest pp dihadron measurements from STAR [97], and all kinematic variable binnings for the relevant processes under consideration, making this the most comprehensive study of dihadron observables to date.
- (3) Examine the compatibility between phenomenological results for $h_1(x)$ and/or the tensor charges based on the dihadron approach, those in the TMD/collinear twist-3 framework, and LQCD computations by implementing theoretical constraints at small x [98] and large x [99] (where experimental data is absent) in order to meaningfully calculate the integrals in Eq. (1) and include LQCD data into the analysis.

Connected to point (3), the ability to probe certain low-energy beyond the Standard Model (BSM) physics relies on knowledge of the isovector tensor charge, $g_T = \delta u - \delta d$, as well as the individual quark tensor charges δu and δd (see, e.g., Refs. [100–108]). Therefore, it is crucial to have precise values for δu , δd , and g_T and to test how compatible they are between different approaches used for their determination. To highlight the importance of this issue, we have assembled essential results and discussion regarding our extraction of transversity and calculation of the tensor charges, as well as a comparison to other phenomenological work and lattice QCD, in Ref. [109], with some supplemental details on that aspect of the analysis provided here alongside points (1) and (2).

We organize the paper as follows. In Sec. II we summarize the factorization formulas for all relevant observables, while in Sec. III we discuss the parametrization choices for the DiFFs, IFFs, and transversity PDFs as well as review the Bayesian methodology employed for the analysis. In Sec. IV we compare our theoretical calculations to the data and discuss the quality of the fit. In Sec. V we show and discuss the extracted DiFFs and IFFs as well as calculate expectation values for z and M_h . In Sec. VI we present the extracted transversity PDFs and tensor charges, comparing to findings from other phenomenological analyses and LQCD, and providing additional information beyond Ref. [109]. We also elaborate on the importance of high- x measurements in further testing the compatibility between phenomenology and LQCD results for the tensor charges. Finally, in Sec. VII we summarize our results and

discuss future directions for this analysis, including the role of new experimental data.

II. EXPERIMENTAL OBSERVABLES USED IN THE GLOBAL ANALYSIS

In this analysis we study $\pi^+\pi^-$ dihadron production in e^+e^- annihilation, SIDIS, and proton-proton collisions with the theoretical formulas given at leading order (LO) in the strong coupling α_s using leading-twist collinear factorization. Since we only consider $\pi^+\pi^-$ pairs, we will drop the superscript from our notation for the DiFFs. In the subsections below we summarize each of the three processes and the relevant equations. We note that the formulas in this section use a new definition of the DiFFs that has a number density interpretation [65].

A. Dihadron production in electron-positron annihilation

In the e^+e^- annihilation process, $e^+e^- \rightarrow (\pi^+\pi^-)X$, an electron and positron annihilate to form an inclusive spectrum of $\pi^+\pi^-$ pairs with invariant mass M_h and fractional energy z . The center of mass (c.m.) energy is denoted by $\sqrt{s}$. The cross section for this process is given by [65]

$$\frac{d\sigma}{dz dM_h} = \frac{4\pi N_c \alpha_{\text{em}}^2}{3s} \sum_q \bar{e}_q^2 D_1^q(z, M_h), \quad (2)$$

where the sum is over quarks and antiquarks, α_{em} is the fine structure constant, and the hard scale of the process is set to be $\mu = \sqrt{s}$. In addition, one has

$$\begin{aligned} \bar{e}_q^2 = e_q^2 + (1 + a_Z^2)(1 + a_q^2) \frac{G_F^2 M_Z^4 s^2}{128\pi^2 \alpha_{\text{em}}^2 [(s - M_Z^2) + \Gamma_Z^2 M_Z^2]} \\ - a_Z e_q a_q \frac{G_F M_Z^2 s (s - M_Z^2)}{4\sqrt{2}\pi \alpha_{\text{em}} [(s - M_Z^2) + \Gamma_Z^2 M_Z^2]}, \end{aligned} \quad (3)$$

where e_q is the charge of the quark of flavor q in units of the elementary charge, G_F is the Fermi constant, and M_Z and Γ_Z are the mass and decay width of the Z boson, respectively. Furthermore, $a_Z = -1 + 4\sin^2 \theta_W$, and $a_q = 1 - 8/3\sin^2 \theta_W$ for $q = u, c, \bar{u}, \bar{c}$ or $a_q = -1 + 4/3\sin^2 \theta_W$ for $q = d, s, b, \bar{d}, \bar{s}, \bar{b}$, where θ_W is the weak mixing angle. As we will argue below [see Eq. (13) and Appendix B], there are five independent DiFFs ($D_1^u, D_1^s, D_1^c, D_1^b$, and D_1^q), and this observable is only capable of constraining one of them (which we choose to be D_1^u). Thus, we supplement the experimental data from Belle with information from PYTHIA [110], which will be discussed in more detail in Sec. IV A.

We also consider the process $e^+e^- \rightarrow (\pi^+\pi^-)(\pi^+\pi^-)X$ where two dihadron pairs are detected, with the second having invariant mass $\bar{M}_h$ and fractional energy $\bar{z}$. By

measuring an azimuthal correlation of two hadron pairs detected in opposite hemispheres, one obtains an observable that is sensitive to the IFF $H_1^{\leftarrow,q}(z, M_h)$. This modulation (known as the Artru-Collins asymmetry [111] and denoted as a_{12R} by Belle [112]) is given by [62,63,111,113–115]

$$\begin{aligned} A^{e^+e^-}(z, M_h, \bar{z}, \bar{M}_h) \\ = \frac{\sin^2 \theta \sum_q e_q^2 H_1^{\leftarrow,q}(z, M_h) H_1^{\leftarrow,\bar{q}}(\bar{z}, \bar{M}_h)}{(1 + \cos^2 \theta) \sum_q e_q^2 D_1^q(z, M_h) D_1^{\bar{q}}(\bar{z}, \bar{M}_h)}, \end{aligned} \quad (4)$$

where θ is defined as the polar angle between the beam axis and the reference axis in the c.m. system [113]. As we will argue later [see Eq. (18) and Appendix B], there is only one IFF that is needed for phenomenology, which we choose to be $H_1^{\leftarrow,u}$. Thus, only one observable is needed to constrain the IFFs. However, one can see from Eq. (4) that $A^{e^+e^-}$ is proportional to $[H_1^{\leftarrow,u}]^2$. Consequently, the asymmetry cannot uniquely determine the sign of $H_1^{\leftarrow,u}$, and it must be fixed by hand. Given the sign of the SIDIS asymmetries from experimental measurements (see the following section) and assuming that the transversity up quark PDF must be positive (as is found in all phenomenological, model, and lattice QCD studies) leads to the conclusion that $H_1^{\leftarrow,u}$ must be negative (which is also supported by model calculations for $H_1^{\leftarrow}$ [116]). In the analysis we therefore choose $H_1^{\leftarrow,u}$ to be negative.

B. Dihadron production in SIDIS

The SIDIS process is given by $\ell N \rightarrow \ell'(\pi^+\pi^-)X$, where a lepton with 4-momentum k scatters off a transversely polarized proton ($N = p$) or deuteron ($N = D$) with 4-momentum P and is detected with 4-momentum k' along with a $\pi^+\pi^-$ pair with 4-momentum P_h . The 4-momentum transfer squared is given by $Q^2 \equiv -(k - k')^2$, and we define the Bjorken scaling variable $x_{\text{bj}} \equiv Q^2/(2P \cdot q)$ and the inelasticity $y \equiv P \cdot q/P \cdot k$. We denote the virtual-photon 4-momentum and relative 4-momentum between the two hadrons (divided by 2) by q and R , respectively. The transverse component $\vec{R}_T$ of R makes an azimuthal angle ϕ_R about the virtual-photon direction, and the transverse nucleon spin $\vec{S}_T$ makes an azimuthal angle ϕ_S . COMPASS [117] defines the angle ϕ_{RS} as $\phi_{RS} \equiv \phi_R + \phi_S - \pi$, while HERMES [118] defines it as $\phi_{RS} \equiv \phi_R + \phi_S$.

The asymmetry (usually denoted as $A_{UT}^{\sin(\phi_R + \phi_S) \sin \theta}$ [118] or $A_{UT}^{\sin(\phi_R + \phi_S)}$ [117]) can be written as [57,63,66,119,120]

$$A_{UT}^{\text{SIDIS}} = c(y) \frac{\sum_q e_q^2 h_1^q(x) H_1^{\leftarrow,q}(z, M_h)}{\sum_q e_q^2 f_1^q(x) D_1^q(z, M_h)}. \quad (5)$$

The factor $c(y)$ is given by

$$c(y) = \begin{cases} 1 & (\text{COMPASS}), \\ -\frac{1-y}{1-y+\frac{y^2}{2}} & (\text{HERMES}). \end{cases}$$

The opposite sign between the asymmetries for COMPASS and HERMES comes from the different definitions for the angle ϕ_{RS} mentioned above. The hard scale is set to be $\mu = Q$, and the unpolarized PDF f_1 is taken from Ref. [36] (we utilize only the mean value).

Both COMPASS and HERMES measured the asymmetry for proton targets, while COMPASS also used a deuteron target. For the deuteron target we neglect nuclear corrections and take the (unpolarized and transversity) PDFs in Eq. (5) to be the average of proton and neutron PDFs, using isospin symmetry to relate the neutron PDFs to those of the proton. Due to the symmetry of the $\pi^+\pi^-$ IFFs [see Eq. (18)], one has in the numerator for a proton target

$$\begin{aligned} \sum_q e_q^2 h_1^q H_1^{\lessdot,q} &= \frac{1}{9} H_1^{\lessdot,u} [4(h_1^u - h_1^{\bar{u}}) - (h_1^d - h_1^{\bar{d}})] \\ &= \frac{1}{9} H_1^{\lessdot,u} [4h_1^{u_v} - h_1^{d_v}] \approx \frac{4}{9} H_1^{\lessdot,u} h_1^{u_v}, \end{aligned} \quad (6)$$

while for the deuteron one has

$$\begin{aligned} \sum_q e_q^2 h_1^{q/D} H_1^{\lessdot,q} &= \frac{1}{6} H_1^{\lessdot,u} [(h_1^u - h_1^{\bar{u}}) + (h_1^d - h_1^{\bar{d}})] \\ &= \frac{1}{6} H_1^{\lessdot,u} [h_1^{u_v} + h_1^{d_v}], \end{aligned} \quad (7)$$

where $h_1^{q/D}$ are the deuteron transversity PDFs. From this one immediately sees that the SIDIS data is sensitive only to the valence transversity PDFs [63,66,69], with the proton asymmetry primarily sensitive to u_v and the deuteron asymmetry equally sensitive to u_v and d_v .

C. Dihadron production in proton-proton collisions

In the process $p^\uparrow p \rightarrow (\pi^+\pi^-)X$, a transversely polarized proton with 4-momentum P_A collides with an unpolarized proton with 4-momentum P_B producing an outgoing dihadron pair with 4-momentum P_h (with P_{hT} the magnitude of the transverse component) and pseudorapidity η . The measured asymmetry (denoted as A_{UT} by STAR [121]) is given by [122]

$$A_{UT}^{pp} = \frac{\mathcal{H}(M_h, P_{hT}, \eta)}{\mathcal{D}(M_h, P_{hT}, \eta)}, \quad (8)$$

where the numerator and denominator read [123]

$$\mathcal{H}(M_h, P_{hT}, \eta) = 2P_{hT} \sum_i \sum_{a,b,c,d} \int_{x_a^{\min}}^1 dx_a \int_{x_b^{\min}}^1 \frac{dx_b}{z} h_1^a(x_a) f_1^b(x_b) \frac{d\Delta\hat{\sigma}_{a^\dagger b \rightarrow c^\dagger d}}{d\hat{t}} H_1^{\lessdot,c}(z, M_h), \quad (9)$$

$$\mathcal{D}(M_h, P_{hT}, \eta) = 2P_{hT} \sum_i \sum_{a,b,c,d} \int_{x_a^{\min}}^1 dx_a \int_{x_b^{\min}}^1 \frac{dx_b}{z} f_1^a(x_a) f_1^b(x_b) \frac{d\hat{\sigma}_{ab \rightarrow cd}}{d\hat{t}} D_1^c(z, M_h), \quad (10)$$

with x_a (x_b) the momentum fraction of the parton coming from the proton with momentum P_A (P_B). We note that A_{UT}^{pp} is sensitive to both the quark and antiquark transversity PDFs. The sum $\sum_i$ is over all partonic subprocesses, with $\sum_{a,b,c,d}$ summing over all possible parton flavor combinations for a given channel. The relevant Mandelstam variables are $s = (P_A + P_B)^2$, $t = (P_A - P_h)^2$, and $u = (P_B - P_h)^2$. The total momentum fraction of the dihadron is fixed at LO to be

$$z = \frac{P_{hT}}{\sqrt{s}} \left(\frac{x_a e^{-\eta} + x_b e^{\eta}}{x_a x_b} \right). \quad (11)$$

The hard scale is set to be $\mu = P_{hT}$. The hard (perturbative) partonic cross sections $d\hat{\sigma}$ and $d\Delta\hat{\sigma}$ [123] depend on $\hat{s} = x_a x_b s$, $\hat{t} = x_a s/z$, and $\hat{u} = x_b u/z$. The explicit

expressions for $d\Delta\hat{\sigma}$ can be found in Appendix A. The limits on the integration are

$$x_a^{\min} = \frac{P_{hT} e^{\eta}}{\sqrt{s} - P_{hT} e^{-\eta}}, \quad x_b^{\min} = \frac{x_a P_{hT} e^{-\eta}}{x_a \sqrt{s} - P_{hT} e^{\eta}}. \quad (12)$$

III. PHENOMENOLOGICAL METHODOLOGY

In this section we discuss the parametrization of the DiFFs, IFFs, and transversity PDFs. All functions are evolved using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [124–126] appropriate for the transversity PDFs [127–130] and DiFFs/IFFs [65] with LO splitting functions. We mention that evolution equations were previously derived for “extended” DiFFs (at LO) in Ref. [131] (and commented on in Ref. [114])

and for collinear DiFFs [functions of (z_1, z_2)] in Refs. [132–135] at LO and recently in Refs. [136,137] at next-to-leading order (NLO). The strong coupling α_s is evolved at leading-logarithmic accuracy with the boundary condition $\alpha_s(M_Z) = 0.118$. We use the zero mass variable flavor scheme for DGLAP evolution. The values of the heavy quark mass thresholds for the evolution are taken from the Particle Data Group as $m_c = 1.28$ GeV and $m_b = 4.18$ GeV in the $\overline{\text{MS}}$ scheme [138].

A. DiFF parametrization

The DiFF D_1 for $\pi^+\pi^-$ production satisfies (see Appendix B and Ref. [62])

$$\begin{aligned} D_1^u &= D_1^d = D_1^{\bar{u}} = D_1^{\bar{d}}, \\ D_1^s &= D_1^{\bar{s}}, \quad D_1^c = D_1^{\bar{c}}, \quad D_1^b = D_1^{\bar{b}}, \end{aligned} \quad (13)$$

which means, including the gluon, there are five independent D_1 functions to be fitted. D_1^u , D_1^s , and D_1^g are parametrized at the input scale $\mu_0 = 1$ GeV, while D_1^c and D_1^b are parametrized at $\mu = m_c$ and $\mu = m_b$, respectively. The primary challenge with extracting DiFFs is their dependence on both z and M_h . Here we choose to discretize the M_h dependence while using an explicit functional form for the z dependence. Such a choice has the computational advantage of being convertible to Mellin space. For D_1^u , we choose the following M_h grid:

$$\begin{aligned} \mathbf{M}_h^u &= [2m_\pi, 0.40, 0.50, 0.70, 0.75, 0.80, 0.90, 1.00, 1.20, \\ &\quad \times 1.30, 1.40, 1.60, 1.80, 2.00] \text{ GeV}. \end{aligned}$$

We note that the grid is not uniform and is instead chosen in a way to best describe the detailed resonance structure of the e^+e^- cross section [64]. For s , c , b and g we choose grids that are less dense:

$$\begin{aligned} \mathbf{M}_h^s &= \mathbf{M}_h^c = [2m_\pi, 0.50, 0.75, 1.00, 1.20, 1.60, 2.00] \text{ GeV}, \\ \mathbf{M}_h^b &= [2m_\pi, 0.70, 1.00, 1.40, 2.00] \text{ GeV}, \\ \mathbf{M}_h^g &= [2m_\pi, 0.70, 1.40, 2.00] \text{ GeV}. \end{aligned} \quad (14)$$

We find that these grids, and the parameters associated with them [see Eq. (15)], are necessary and sufficient to describe the Belle and PYTHIA data. At each value of M_h on the grid, denoted by $M_h^{i,j}$, the z dependence is parametrized as

$$D_1^i(z, M_h^{i,j}) = \sum_{k=1,2,3} \frac{N_{jk}^i}{\mathcal{M}_{jk}^i} z^{\alpha_{jk}^i} (1-z)^{\beta_{jk}^i}, \quad (15)$$

where $i = u, s, c, b, g$, $\mathbf{a} \equiv \{N_{jk}^i, \alpha_{jk}^i, \beta_{jk}^i\}$ is the set of parameters to be inferred, and $\mathcal{M}_{jk}^i = B[\alpha_{jk}^i + 1, \beta_{jk}^i + 1]$, where $B(a, b)$ is the beta function, normalizes the function to the first moment to largely decorrelate the normalization

and shape parameters. For the up quark, it is necessary to include terms up to $k = 3$ in order to describe the Belle cross section data. The other functions will be constrained by PYTHIA-generated data (see Sec. IV A) and we find that only the $k = 1$ term is necessary to describe them. This leads to $14 \times 9 = 126$ parameters for D_1^u , $7 \times 3 = 21$ parameters each for D_1^s and D_1^c , $5 \times 3 = 15$ parameters for D_1^b , and $4 \times 3 = 12$ parameters for D_1^g , for a total of 195 free parameters. The parameters are generally restricted within the ranges $0 < N_{jk}^i < 15$, $-2 < \alpha_{jk}^i < 10$, and $0 < \beta_{jk}^i < 10$, but, in practice, these ranges may be restricted further depending on i, j , and k . These ranges are chosen to give the fit sufficient but not unnecessary flexibility.

The functional form in Eq. (15) is easily converted into Mellin space at each $M_h^{i,j}$. The inverse Mellin transform requires a complex contour integration [see Eq. (33)], which is performed by discretizing the contour. To obtain the function at any value of M_h , we interpolate both the real and imaginary components of the function in Mellin space using the values calculated at $M_h^{i,j}$. Once interpolated, the real and imaginary components are recombined and the inverse Mellin transform is performed to obtain the function in momentum space at any value of M_h .

We contrast our choice of parametrization with that of Ref. [62], which used a considerably more complicated functional form for u , s , and c with different functions for the continuum and three resonant channels. The bottom quark DiFF was not included in the analysis of Ref. [62], and the gluon DiFF was only generated perturbatively. In Ref. [68] three choices for the gluon DiFF at the input scale were used: $D_1^g = 0$, $D_1^g = D_1^u/4$, and $D_1^g = D_1^u$. Our results in Fig. 11 suggest that none of these scenarios seem to be correct across all z and M_h . Overall, our parametrization has the advantages of requiring fewer assumptions about the functional forms, reducing complicated correlations between parameters, and being convertible to Mellin space. The latter two features make the analysis far more computationally efficient.

We also enforce the positivity bound [57],

$$D_1^i(z, M_h; \mu) > 0, \quad (16)$$

approximately on each Monte Carlo replica by using a Bayesian prior (see Sec. III D) that in effect imposes a penalty on the χ^2 function when the bounds are violated [139]. For each replica and at each step of the χ^2 minimization we first calculate D_1 at the input scale $\mu_0 = 1$ GeV at 300 points spaced linearly in the (z, M_h) plane, with $0.2 < z < 1$ and $2m_\pi < M_h < 2.0$ GeV, noting that if positivity is enforced at the input scale, then it will automatically hold at larger scales for any function that evolves through the DGLAP equation [140,141]. We repeat

this process for u, s, c, b , and g . Any DiFFs that are negative contribute to the overall χ^2 through

$$\Delta_{D_1 < 0}^2 = w^2 \left(\sum_{z, M_h} \sum_i \Theta(-D_1^i(z, M_h; \mu_0)) |D_1^i(z, M_h; \mu_0)| \right)^2, \quad (17)$$

where the sum $\sum_{z, M_h}$ represents summing over the 300 (z, M_h) points, and $\Theta(X)$ is the step function. If the contribution is nonzero, then it is proportional to the size of the violation. The weight w is chosen to be 3 so that the initial contribution to the χ^2 is generally $\mathcal{O}(1000)$, although the size of the initial contribution can vary significantly depending on the starting parameters. This ensures that the contribution is non-negligible but that it also does not dominate the entire χ^2 function which otherwise is also $\mathcal{O}(1000)$. This choice ensures that the violation of the bound is small.

B. IFF parametrization

The IFF $H_1^{\triangleleft}$ for $\pi^+\pi^-$ production satisfies (see Appendix B and Ref. [62])

$$H_1^{\triangleleft, u} = -H_1^{\triangleleft, d} = -H_1^{\triangleleft, \bar{u}} = H_1^{\triangleleft, \bar{d}},$$

$$H_1^{\triangleleft, s} = H_1^{\triangleleft, \bar{s}} = H_1^{\triangleleft, c} = H_1^{\triangleleft, \bar{c}} = H_1^{\triangleleft, b} = H_1^{\triangleleft, \bar{b}} = 0. \quad (18)$$

Consequently, there is just a single independent IFF (since the transversity cannot couple to the IFF for gluons at LO), which we choose to be $H_1^{\triangleleft, u}$. The IFF $H_1^{\triangleleft, u}$ is parametrized similarly to D_1^i . Since the relevant data for $H_1^{\triangleleft, u}$ is comparatively sparse and has larger errors, far fewer parameters are needed. Thus, we are able to choose a less dense M_h grid,

$$\mathbf{M}_h^u = [2m_\pi, 0.50, 0.70, 0.85, 1.00, 1.20, 1.60, 2.00] \text{ GeV},$$

and at each $\mathbf{M}_h^{u,j}$, the z dependence is given by

$$H_1^{\triangleleft, u}(z, \mathbf{M}_h^{u,j}) = \sum_{k=1,2} \frac{N_{jk}^u}{\mathcal{M}_{jk}^u} z^{\alpha_{jk}^u} (1-z)^{\beta_{jk}^u}, \quad (19)$$

at the input scale $\mu_0 = 1$ GeV. This leads to a total of $8 \times 6 = 48$ free parameters for the IFF. The parameters are restricted within the ranges $-1 < N_{jk}^u < 0$ [see discussion below Eq. (4)], $-1 < \alpha_{jk}^u < 5$, and $0 < \beta_{jk}^u < 10$ for all j and k . As with D_1^i , this form is easily converted into Mellin space.

We also enforce the positivity bound [57]

$$|H_1^{\triangleleft, i}(z, M_h; \mu)| < D_1^i(z, M_h; \mu), \quad (20)$$

which, due to the symmetry relations Eqs. (13) and (18), only needs to be applied to $i = u$. Since we have fixed D_1^u to be positive and $H_1^{\triangleleft, u}$ to be negative, we only need to check the condition $D_1^u > -H_1^{\triangleleft, u}$. This is implemented similarly to the positivity constraint on D_1 in Eq. (16), with a Bayesian prior (see Sec. III D) that in effect leads to a χ^2 penalty equal to

$$\Delta_{|H_1^{\triangleleft, u}| > D_1^u}^2 = w^2 \left(\sum_{z, M_h} \Theta(-[H_1^{\triangleleft, u}(z, M_h; \mu_0) + D_1^u(z, M_h; \mu_0)]) |H_1^{\triangleleft, u}(z, M_h; \mu_0) + D_1^u(z, M_h; \mu_0)| \right)^2, \quad (21)$$

with w again chosen to be 3 [see discussion below Eq. (17)].

C. Transversity PDF parametrization

The transversity PDFs are parametrized at the input scale $\mu_0 = 1$ GeV using the form

$$h_1^i(x) = \frac{N^i}{\mathcal{M}^i} x^{\alpha^i} (1-x)^{\beta^i} (1 + \gamma^i \sqrt{x} + \delta^i x), \quad (22)$$

normalized to the first moment $\mathcal{M}^i = \text{B}[\alpha^i + 1, \beta^i + 1] + \gamma^i \text{B}[\alpha^i + \frac{3}{2}, \beta^i + 1] + \delta^i \text{B}[\alpha^i + 2, \beta^i + 1]$. We choose to parametrize the valence distributions $h_1^{u_v}$ and $h_1^{d_v}$, as well as the antiquark distributions $h_1^{\bar{u}} = -h_1^{\bar{d}}$. Since we only have three unique observables to constrain the transversity PDFs (proton SIDIS, deuteron SIDIS, and pp collisions), we choose the relation between the antiquarks based on predictions from the large- N_c limit [142]. We note that only one previous phenomenological analysis has included antiquark transversity PDFs, which was an exploratory study in Ref. [54] that found $h_1^{\bar{u}}$ and $h_1^{\bar{d}}$ to be small and consistent with zero (although had set $h_1^{\bar{u}} = h_1^{\bar{d}}$, at variance with the large- N_c limit). The N parameter is restricted between $0 < N^i < 1$ for $i = u$ [see discussion below Eq. (4)] and $-1 < N^i < 1$ for $i = d, \bar{u}$, while the other parameters are always restricted between $0.09 < \alpha^i < 0.26$ [see discussion below surrounding Eq. (23)], $0 < \beta^i < 20$, $-20 < \gamma^i < 20$, and $-20 < \delta^i < 20$. We have tested the model dependence of our input transversity functions by adding a second shape of the form in Eq. (22) to all three quark flavors. We find that the results only change marginally, indicating that the experimental data do not call for a more expressive form for these PDFs. We caution, however, that this test is not exhaustive of all possible functional forms one could choose for the transversity functions. We have a total of $3 \times 5 = 15$ parameters for the transversity PDFs. Combined with the 195 parameters for D_1 , the 48

parameters for $H_1^\triangleleft$, and 7 normalization parameters [see Eq. (28) below], we end up with a total of 265 fitted parameters.

We also place a constraint on the small- x behavior of the transversity PDFs [governed by the α^i parameter in Eq. (22)], where no experimental data is available. Theoretical calculations have placed limits on this parameter as $x \rightarrow 0$ (ignoring saturation effects) [98]:

$$\alpha^i \xrightarrow{x \rightarrow 0} 1 - 2\sqrt{\frac{\alpha_s N_c}{2\pi}}. \quad (23)$$

We apply this limit to both the valence quarks and antiquarks, and there is a roughly 50% uncertainty on this value from $1/N_c$ and NLO corrections [143]. At the input scale, we calculate $\alpha^i \rightarrow 0.17$, and so we limit the α^i parameter to the range $0.085 \leq \alpha^i \leq 0.255$ for all quark flavors. Strictly speaking, this limit only applies as $x \rightarrow 0$, while our approach places a limit on the entire range of x . We find, however, that limiting α^i as such has no impact on the resulting transversity PDFs in the measured region or on our ability to describe the experimental data. Thus, this simplified approach is sufficient to capture the $x \rightarrow 0$ behavior while not affecting results at moderate or high x .

We also enforce the Soffer bound [99] on the transversity PDFs, given by

$$|h_1^i(x; \mu)| \leq \frac{1}{2}[f_1^i(x; \mu) + g_1^i(x; \mu)], \quad (24)$$

where f_1^i and g_1^i are taken from Ref. [36]. This is again enforced through a Bayesian prior (see Sec. III D) that in effect leads to an χ^2 penalty, similar to Eqs. (17) and (21):

$$\Delta_{\text{SB}}^2 = w^2 \left(\sum_{\pm} \sum_x \sum_i \Theta(-F_{\pm}^i(x; \mu_0)) |F_{\pm}^i(x; \mu_0)| \right)^2, \quad (25)$$

where $i = u, d, \bar{u}, \bar{d}$ and $F_{\pm}^i(x; \mu) \equiv \frac{1}{2}[f_1^i(x; \mu) + g_1^i(x; \mu)] \pm h_1^i(x; \mu)$. The sum over $\pm$ is due to the fact that the

positivity bound applies to the absolute value of h_1^i . The sum $\sum_x$ represents summing over 100 points in x linearly spaced in the range $0.001 < x < 0.99$. The weight w is chosen to be 10 [see discussion below Eq. (17)]. In order to account for the errors on f_1 and g_1 , we add their 1σ errors in quadrature, add this error to the mean of the Soffer bound, and use these values in the fit.

D. Bayesian analysis

Our Bayesian analysis consists of sampling the posterior distribution given by

$$\mathcal{P}(\mathbf{a}|\text{data}) \propto \mathcal{L}(\mathbf{a}, \text{data})\pi(\mathbf{a}), \quad (26)$$

with a likelihood function of Gaussian form,

$$\mathcal{L}(\mathbf{a}, \text{data}) = \exp\left(-\frac{1}{2}\chi^2(\mathbf{a}, \text{data})\right), \quad (27)$$

and a prior function $\pi(\mathbf{a})$. The χ^2 function in Eq. (27) is defined for each replica as

$$\chi^2(\mathbf{a}) = \sum_{e,i} \left(\frac{d_{e,i} - \sum_k r_{e,k} \beta_{e,i}^k - T_{e,i}(\mathbf{a})/N_e}{\alpha_{e,i}} \right)^2, \quad (28)$$

where $d_{e,i}$ is the data point i from experimental dataset e , and $T_{e,i}$ is the corresponding theoretical value. Note that the PYTHIA and LQCD data are not included in this sum. All uncorrelated uncertainties are added in quadrature and labeled by $\alpha_{e,i}$, while $\beta_{e,i}^k$ represents the k th source of point-to-point correlated systematic uncertainties for the i th data point weighted by $r_{e,k}$. The latter are optimized per values of the parameters $\mathbf{a}$ via $\partial\chi^2/\partial r_{e,k} = -2r_{e,k}$, which can be solved in a closed form. We include normalization parameters N_e for each dataset e as part of the posterior distribution per data set.

The prior $\pi(\mathbf{a})$ for each replica is given by

$$\begin{aligned} \pi(\mathbf{a}) = & \prod_l \Theta((a_l - a_l^{\min})(a_l^{\max} - a_l)) \prod_f \prod_i \exp\left\{-\frac{1}{2}\left(\frac{d_{f,i} - T_{f,i}(\mathbf{a})}{\alpha_{f,i}}\right)^2\right\} \\ & \times \prod_e \prod_k \exp\left\{-\frac{1}{2}r_{e,k}^2\right\} \prod_e \exp\left\{-\frac{1}{2}\left(\frac{1 - N_e}{\delta N_e}\right)^2\right\} \exp\left\{-\frac{1}{2}\Delta_{D_1 < 0}^2\right\} \exp\left\{-\frac{1}{2}\Delta_{|H_1^\triangleleft| > D_1}^2\right\} \exp\left\{-\frac{1}{2}\Delta_{\text{SB}}^2\right\}, \end{aligned} \quad (29)$$

where $\prod_l$ is a product over the parameters, and $\prod_f$ is a product over the LQCD and PYTHIA datasets. The step function $\Theta(X)$ forces the parameters to be within the chosen range between $a_l^{\min}$ and $a_l^{\max}$, which also automatically enforces the small- x constraint Eq. (23). A Gaussian penalty controlled by the experimentally quoted normalization

uncertainties δN_e is applied when fitting N_e . The final three exponentials utilize Eqs. (17), (21), and (25).

The posterior distribution is sampled via data resampling, whereby multiple maximum likelihood optimizations of $\mathcal{P}(\mathbf{a}|\text{data})$ are carried out by adding Gaussian noise with width $\alpha_{e,i}$ to each data point across all data sets

(including the PYTHIA and LQCD priors). The resulting ensemble of replicas $\{\mathbf{a}_k; k = 1, \dots, n\}$ is then used to obtain statistical estimators for a given observable (which are functions of DiFFs/PDFs) $\mathcal{O}(\mathbf{a})$, such as the mean and variance,

$$\mathbb{E}[\mathcal{O}] = \frac{1}{n} \sum_k \mathcal{O}(\mathbf{a}_k), \quad \mathbb{V}[\mathcal{O}] = \frac{1}{n} \sum_k [\mathcal{O}(\mathbf{a}_k) - \mathbb{E}[\mathcal{O}]]^2. \quad (30)$$

The agreement between data and theory is assessed by using the “reduced” χ^2 which is defined as

$$\chi_{\text{red}}^2 \equiv \frac{1}{N_{\text{dat}}} \sum_{e,i} \left(\frac{d_{e,i} - \mathbb{E}[\sum_k r_{e,k} \beta_{e,i}^k + T_{e,i}/N_e]}{\alpha_{e,i}} \right)^2, \quad (31)$$

where N_{dat} is the number of data points under consideration and $\mathbb{E}[\dots]$ represents the mean theory as defined in Eq. (30). We note that in Eq. (31) we also consider the LQCD and PYTHIA-generated data in order to quantify their compatibility with the experimental data.

E. Mellin space techniques

We have chosen parametrizations for all the functions in a way that can be converted to Mellin space, which has two

advantages. First, as is well known, it allows us to solve the DGLAP evolution equations analytically rather than using an iterative solution as is required in momentum space. We now demonstrate the second advantage of this approach when calculating the pp asymmetry [144]. One can relate the DiFFs, IFFs, and transversity PDFs in momentum space and Mellin space through

$$\begin{aligned} h_1(x; \mu^2) &= \frac{1}{2\pi i} \int dN x^{-N} \mathbf{h}_1(N; \mu^2), \\ H_1^{\leftarrow}(z, M_h; \mu^2) &= \frac{1}{2\pi i} \int dM z^{-M} \mathbf{H}_1^{\leftarrow}(M, M_h; \mu^2), \\ D_1(z, M_h; \mu^2) &= \frac{1}{2\pi i} \int dM z^{-M} \mathbf{D}_1(M, M_h; \mu^2), \end{aligned} \quad (32)$$

where $\mathbf{F}(N; \mu^2)$ is the N th moment of $F(x; \mu^2)$ and $\mathbf{F}(M, M_h; \mu^2)$ is the M th moment of $F(z, M_h; \mu^2)$. The integration is over a contour in the complex N plane that intersects the real axis to the right of the rightmost poles of $\mathbf{F}(N; \mu^2)$. Then, the numerator Eq. (9) and denominator Eq. (10) of the pp asymmetry can be written as

$$\mathcal{H}(M_h, P_{hT}, \eta) = \frac{2P_{hT}}{(2\pi i)^2} \sum_i \sum_{a,b,c} \int dN \mathbf{h}_1^a(N) \int dM \mathbf{H}_1^{\leftarrow,c}(M, M_h) \left[\int_{x_a^{\min}}^1 dx_a \int_{x_b^{\min}}^1 \frac{dx_b}{z} x_a^{-N} z^{-M} f_1^b(x_b) \frac{d\Delta \hat{\sigma}_{a^\dagger b \rightarrow c^\dagger d}}{d\hat{t}} \right], \quad (33)$$

$$\mathcal{D}(M_h, P_{hT}, \eta) = \frac{2P_{hT}}{2\pi i} \sum_i \sum_{a,b,c} \int dM \mathbf{D}_1^{\leftarrow,c}(M, M_h) \left[\int_{x_a^{\min}}^1 dx_a \int_{x_b^{\min}}^1 \frac{dx_b}{z} z^{-M} f_1^a(x_a) f_1^b(x_b) \frac{d\hat{\sigma}_{ab \rightarrow cd}}{d\hat{t}} \right]. \quad (34)$$

Since we do not fit f_1 and instead take it from Ref. [36], the quantities in brackets are independent of any fitting parameters and so only need to be calculated once prior to the actual fit. Thus, the integrals over x_a and x_b need not be calculated at every step of the fitting procedure, massively reducing the computation time, especially for the double convolution in the numerator. We take advantage of this computational efficiency, allowing us to perform a simultaneous fit of the DiFFs, IFFs, and transversity PDFs.

IV. GLOBAL ANALYSIS DATA AND QUALITY OF FIT

In this section we discuss the experimental and LQCD results that enter the analysis as well as the PYTHIA-generated data.

A. PYTHIA-generated data

This analysis must be supplemented by PYTHIA [145,146] data due to the fact that there are five independent D_1 functions (see Sec. III A), but only one experimental observable (e^+e^- cross section $d\sigma/dz dM_h$) currently available to constrain them. (Although D_1 appears in the denominators of the SIDIS and pp asymmetries, those observables are being used to constrain the transversity PDFs and offer no significant sensitivity to D_1 .) Thus, we will discuss how we generate the PYTHIA data and how it is used to help constrain $D_1^i(z, M_h)$.

Our goal with the PYTHIA-generated data is to provide reasonable constraints on the D_1 functions for the strange, charm, bottom, and gluon in the absence of experimental data beyond that from Belle. To do this, we first generate data at the Belle energy $\sqrt{s} = 10.58$ GeV for the ratio

TABLE I. Data from experiment and LQCD used in this analysis for the DiFFs D_1 and $H_1^\perp$ and the transversity PDF h_1 .

Collaboration	References	Observable	Process	Nonperturbative function(s)
Belle	[64]	$d\sigma/dz dM_h$	$e^+e^- \rightarrow (\pi^+\pi^-)X$	D_1
Belle	[112]	$A^{e^+e^-}$	$e^+e^- \rightarrow (\pi^+\pi^-)(\pi^+\pi^-)X$	$D_1, H_1^\perp$
HERMES	[118]	A_{UT}^{SIDIS}	$ep^\uparrow \rightarrow e'(\pi^+\pi^-)X$	$D_1, H_1^\perp, h_1$
COMPASS	[117]	A_{UT}^{SIDIS}	$\mu\{p, D\}^\uparrow \rightarrow \mu'(\pi^+\pi^-)X$	$D_1, H_1^\perp, h_1$
STAR	[97,121]	A_{UT}^{pp}	$p^\uparrow p \rightarrow (\pi^+\pi^-)X$	$D_1, H_1^\perp, h_1$
ETMC	[77]	$\delta u, \delta d$	LQCD	h_1
PNDME	[71]	$\delta u, \delta d$	LQCD	h_1

$\sigma^q/\sigma^{\text{tot}}$ with $q = s, c, b$, where $\sigma^{\text{tot}} \equiv \sum_q \sigma^q$ with the sum over all quark flavors and σ^q is the flavor-dependent cross section. This data provides priors that constrain the strange, charm, and bottom DiFFs. In order to provide some constraint on D_1^q , we repeat this process at five energy scales, evenly spaced between Belle and large electron positron (LEP) energies,

$$\sqrt{s} = [10.58, 30.73, 50.88, 71.04, 91.19] \text{ GeV}.$$

This strategy allows us to gain sensitivity to D_1^q scaling violations. We do not include the bottom quark contribution at the lowest Belle energy $\sqrt{s} = 10.58$ GeV but include it at all energies above this, due to the fact that our massless quark formalism would be highly inadequate at such low energies comparable to $2m_b$.

This leads to a total of 14 generated datasets from PYTHIA, with each one covering the same (z, M_h) region as the Belle data. The same cuts used on the real Belle data are used on the PYTHIA data. The cut in Eq. (35) below is used with $\sqrt{s} = 10.58$ GeV regardless of the value of $\sqrt{s}$, as there is no benefit within our analysis of going beyond the kinematic region of the Belle data.

The PYTHIA [145,146] data can be generated with arbitrarily high statistics, so we neglect the tiny statistical error. In order to propagate model uncertainties in PYTHIA we use a collection of tunes (“PYTHIA 6 default,” “PYTHIA 6 ALEPH,” “PYTHIA 6 LEP/Tevatron,” “PYTHIA 8”) to estimate the systematic uncertainties on the flavor-dependent cross sections. This list of tunes is similar to that used in Ref. [64]. The tunes are summarized in Table IV in Appendix C. The resulting spread of the flavor-dependent cross sections is used to estimate the errors due to the tunes, with the mean taken as the central value and the minimum and maximum defining the uncorrelated systematic error. We choose to fit the ratio $\sigma^q/\sigma^{\text{tot}}$ as it leads to smaller variance between the different tunes compared to taking the absolute cross sections σ^q . Taking such a ratio should also help to cancel NLO and thrust-cut effects. The χ^2_{red} for the PYTHIA data is shown in Table V in Appendix C, along with comparisons between the PYTHIA-generated data and theory calculations.

B. Experimental data and lattice QCD results

A summary of the experimental data and lattice results included in our analysis is given in Table I, where we also indicate which fitted nonperturbative functions enter the observables. The kinematic coverage of the measurements is shown in Fig. 1. We use the $\pi^+\pi^-$ production in e^+e^- annihilation cross section data from Belle [64] at $\sqrt{s} = 10.58$ GeV. (We note that there is a typo in the Belle publication regarding the units of the cross section.

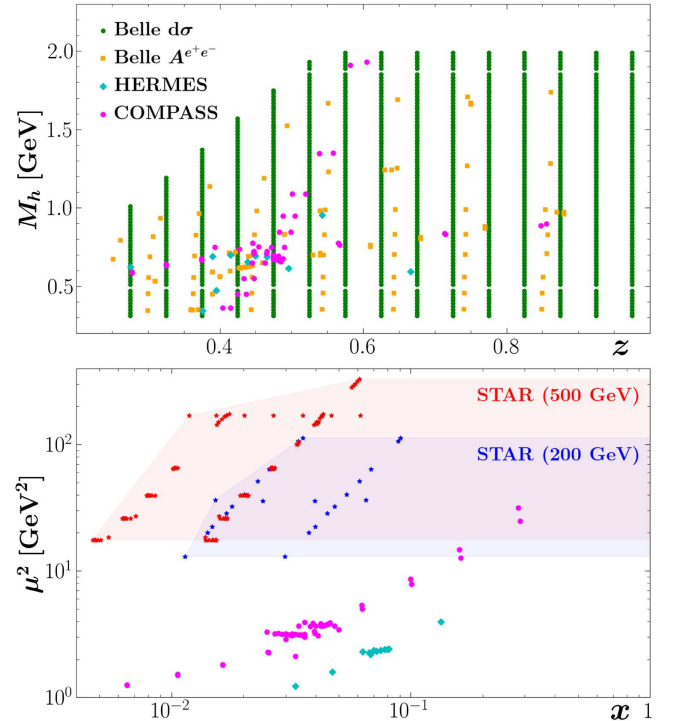


FIG. 1. Kinematic coverage of the datasets included in this analysis. The top panel shows the e^+e^- annihilation and SIDIS data as a function of z and M_h . The bottom panel shows the SIDIS and pp data as a function of x and μ^2 . The variable x represents x_{bj} for SIDIS and x_a for pp collisions, while the scale μ^2 represents Q^2 for SIDIS and P_{hT}^2 for pp collisions. For STAR, the solid points are x_a^{min} and the light shaded region is to indicate that the x_a -integration extends up to $x_a = 1$.

The units are labeled as $\mu\text{b}/\text{GeV}$ but the cross section is actually shown in units of nb/GeV .) Since we can only extract information on the IFF up to $M_h = 2.0$ GeV (see below), we place a cut of $M_h < 2.0$ GeV on the cross section data. The data are provided in bins of M_h , with width $\Delta M_h = 0.02$ GeV, and bins of z , with width $\Delta z = 0.05$. We evaluate the cross section at the mean of each M_h bin $\langle M_h \rangle$ and z bin $\langle z \rangle$. In order to avoid the sharp kaon and D^0 resonances, we also cut out all data points with $\langle M_h \rangle = 0.49$ GeV and $\langle M_h \rangle = 1.87$ GeV. The kinematic limit of the Belle data is given by $M_h < \frac{\sqrt{s}}{2} z$, which is reached in the bins of z below $z = 0.45$. We place a cut of

$$\frac{2\langle M_h \rangle}{\langle z \rangle \sqrt{s}} < 0.7 \quad (35)$$

in order to avoid the region near the kinematic limit. We also cut out the lowest z bin $0.20 < z < 0.25$ as we find difficulty in describing the data in this bin, which may be an indication of the need to resum small- z logarithms. In total, with these cuts we include 1,094 of the 1,468 data points provided by Belle, and the kinematic ranges are $0.25 < z < 1.0$ and $0.3 < M_h < 2.0$ GeV. The data has an overall normalization uncertainty of 1.6%, and the systematic errors are uncorrelated point-to-point [64].

We include the e^+e^- annihilation Artru-Collins asymmetry data provided by Belle [112]. The same cuts on M_h and z discussed above are applied to both hadron pairs. The data is available in three binnings: (z, M_h) , $(M_h, \overline{M}_h)$, and $(z, \overline{z})$. We include all three binnings in our analysis and evaluate the theory formulas using the mean values $\langle M_h \rangle$, $\langle z \rangle$, $\langle \overline{M}_h \rangle$, $\langle \overline{z} \rangle$. The systematic errors on the data are uncorrelated point-to-point and there is no overall normalization error [112]. We note that in all Belle measurements in this analysis a cut is placed on the thrust T [147] of $T > 0.8$. In our LO formalism, it is not possible to take this thrust cut into account because $T < 1$ can only be generated via NLO effects. We expect that the errors from this omission largely cancel out in the ratio of $H_1^</math>/ D_1 which appears in all of the asymmetries, thus reducing the impact on the extraction. However, since the extraction of D_1 relies on cross section data rather than asymmetries, its extraction could be noticeably influenced by NLO corrections.$

The transversity PDFs are constrained by SIDIS asymmetry and pp asymmetry data. SIDIS data is available from HERMES [118] and COMPASS [117]. We use all three available binnings in x_{bj} , z , and M_h , and, as for the e^+e^- annihilation data, apply the cuts $\langle z \rangle > 0.25$ and $\langle M_h \rangle < 2.0$ GeV. We evaluate the theory formulas using the mean values $\langle x_{bj} \rangle$, $\langle z \rangle$, $\langle M_h \rangle$, $\langle y \rangle$. The HERMES data has an overall normalization uncertainty of 8.1%, while the COMPASS data has 2.2% for proton and 5.4% for deuterium. The systematic errors on the HERMES data

are correlated point-to-point [118], while those on the COMPASS data are uncorrelated [117].

The pp collision data is provided by STAR, at both $\sqrt{s} = 200$ GeV [121] and $\sqrt{s} = 500$ GeV [97]. We apply the cut $\langle M_h \rangle < 2.0$ GeV to all of the data. The $\sqrt{s} = 200$ GeV data is provided with three different upper cuts (0.2, 0.3, 0.4) on the opening angle R of the pion pair, with 0.3 treated as the default. This cut is used to filter out pion pairs that do not originate from a single parton. We use the data corresponding to $R < 0.3$ and have verified that the changes to the resulting functions, by instead using data from the other cuts, are negligible compared to the statistical uncertainties. The $\sqrt{s} = 500$ GeV data is provided with an opening angle of $R < 0.7$. A larger opening angle cut is acceptable here, as the increased energy means that gluon radiation is occurring at wider angles, allowing the dihadron pair to still be considered as originating from a single parton even with a larger R -cut value. The data is provided binned in P_{hT} , M_h , and η , with the results (often) provided for both $\eta > 0$ and $\eta < 0$ when binned in P_{hT} or M_h . We include all binnings and take the central values $\langle P_{hT} \rangle$, $\langle M_h \rangle$, and $\langle \eta \rangle$ when evaluating our theory formulas. All systematic errors are uncorrelated, and the 200 and 500 GeV data have normalization uncertainties of 4.8% and 4.5%, respectively [97,121].

We also consider the inclusion of LQCD data on tensor charges in the fit and treat them as Bayesian priors (see Sec. III D). We restrict ourselves to results at the physical pion mass with $2 + 1 + 1$ flavors, where calculations are available from ETMC [77] and PNDME [71] on δu , δd , and g_T . We choose to fit δu and δd rather than g_T in order to provide flavor separation. The reported uncertainties are treated as uncorrelated.

C. Data vs theory

We now discuss the results of our global analysis, where we consider three different scenarios. The first includes all experimental and LQCD data discussed above and will be referred to as “JAMDIFF (w/ LQCD).” The “JAMDIFF (no LQCD)” fit removes the LQCD data. The “JAMDIFF (SIDIS only)” fit excludes all of the STAR pp and LQCD data, and also sets the antiquark transversity PDFs to zero as they cannot be constrained in this scenario. The χ^2_{red} for the three scenarios is summarized in Table II. In the plots that follow comparing our theory results with the experimental measurements, we show the JAMDIFF (w/ LQCD) fit. We will reserve a discussion about our calculated tensor charges and comparison to those from LQCD computations (and other values from phenomenology) to Sec. VI. We reemphasize that we have performed a simultaneous global analysis of DiFFs/IFFs and transversity PDFs, where, unlike previous work [63,66,68,69], the parameters for the DiFFs are not fixed (from a fit of only e^+e^- annihilation) but allowed to be free along with the transversity PDF parameters. We have also studied an

TABLE II. Summary of χ^2_{red} values for the different fit configurations defined in Sec. IV C. Note main results are set in bold.

Experiment	Binning	N_{dat}	χ^2_{red}		
			JAMDiFF		
			(w/ LQCD)	(no LQCD)	(SIDIS only)
Belle (cross section) [64]	z, M_h	1094	1.01	1.01	1.01
	z, M_h	55	1.27	1.24	1.28
Belle (Artru-Collins) [112]	$M_h, \bar{M}_h$	64	0.60	0.60	0.60
	$z, \bar{z}$	64	0.42	0.42	0.41
	x_{bj}	4	1.77	1.70	1.67
HERMES [118]	M_h	4	0.41	0.42	0.47
	z	4	1.20	1.17	1.13
	x_{bj}	9	1.98	0.65	0.59
COMPASS (p) [117]	M_h	10	0.92	0.94	0.93
	z	7	0.77	0.60	0.63
	x_{bj}	9	1.37	1.42	1.22
COMPASS (D) [117]	M_h	10	0.45	0.37	0.38
	z	7	0.50	0.46	0.46
	x_{bj}	9	1.37	1.42	1.22
STAR [121] $\sqrt{s} = 200$ GeV $R < 0.3$	$M_h, \eta < 0$	5	2.57	2.56	
	$M_h, \eta > 0$	5	1.34	1.55	
	$P_{hT}, \eta < 0$	5	0.98	1.00	
	$P_{hT}, \eta > 0$	5	1.73	1.74	
	η	4	0.52	1.46	
STAR [97] $\sqrt{s} = 500$ GeV $R < 0.7$	$M_h, \eta < 0$	32	1.30	1.10	
	$M_h, \eta > 0$	32	0.81	0.78	
	$P_{hT}, \eta > 0$	35	1.09	1.07	
	η	7	2.97	1.83	
ETMC δu [77]		1	0.71		
ETMC δd [77]		1	1.02		
PNDME δu [71]		1	8.68		
PNDME δd [71]		1	0.04		
Total χ^2_{red} (N_{dat})			1.01 (1475)	0.98 (1471)	0.96 (1341)

exhaustive set of available data on dihadron observables, which includes, for the first time, the Belle cross section [64], the $\sqrt{s} = 500$ GeV measurements from STAR [97], and all kinematic variable binnings for the relevant processes under consideration, amassing 1471 experimental data points. We are able to describe all of the experimental data very well.

The results for the Belle cross section are shown in Fig. 2. We find that our parametrization discussed in Sec. III A is able to account for the resonance structure in the data and the general behavior of the measurement across the kinematic range in z and M_h , with $\chi^2_{\text{red}} = 1.01$. There are some slight discrepancies in the lowest z bin at small M_h . The fitted normalization N_e [see Eq. (28)] for this dataset is 1.00 (1). In Figs. 3–5 we display the Belle Artru-Collins asymmetry results. Again our theory curves are able to describe the data very well across all kinematics, with $\chi^2_{\text{red}} = 1.27, 0.60, 0.42$ for the $(z, M_h), (M_h, \bar{M}_h), (z, \bar{z})$

binnings, respectively. The only discrepancy is in the $0.77 < M_h < 0.90$ GeV bin at high z , which is what causes the χ^2_{red} for the (z, M_h) binning to be much larger than the other two. There is no fitted normalization for this dataset. As seen in Table II, the χ^2_{red} for the Belle data (both the cross section and Artru-Collins asymmetry) is nearly identical across the three fit configurations.

The results for the SIDIS asymmetry are shown in Fig. 6. The fitted normalizations N_e [see Eq. (28)] are 0.98(4) for HERMES, 1.014(6) for COMPASS proton, and 1.007(8) for COMPASS deuteron. The theory calculations are generally in agreement with the data for all three projections (x_{bj}, M_h, z) . However, we focus the reader's attention on the highest x_{bj} COMPASS proton point (data/curve in red in the left panel of Fig. 6). The trend is for JAMDiFF (w/ LQCD) to increase at larger x_{bj} more rapidly than the COMPASS proton data. The JAMDiFF (no LQCD) fit follows the trend of the data much better (see Fig. 19 in

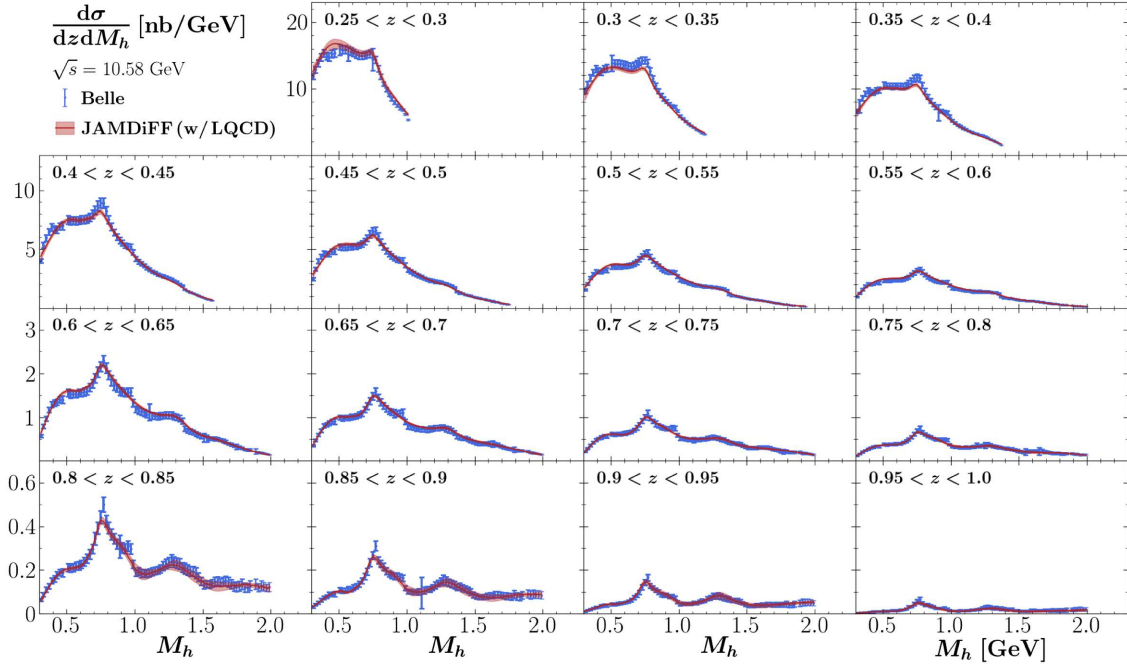


FIG. 2. $d\sigma/dz dM_h$ cross section data from Belle [64] at $\sqrt{s} = 10.58$ GeV (blue circles) plotted as a function of M_h against the mean JAMDiFF (w/ LQCD) result [using Eq. (2)] with 1σ uncertainty (red lines with bands). Each panel is a different bin of z .

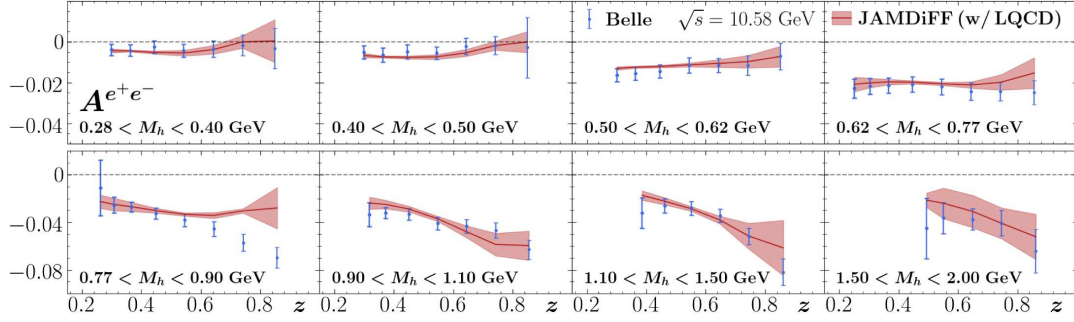


FIG. 3. Artru-Collins asymmetry data from Belle [112] binned in (z, M_h) at $\sqrt{s} = 10.58$ GeV (blue circles) plotted as a function of z against the mean JAMDiFF (w/ LQCD) result [using Eq. (4)] with 1σ uncertainty (red lines with bands). The different panels show different bins of M_h .

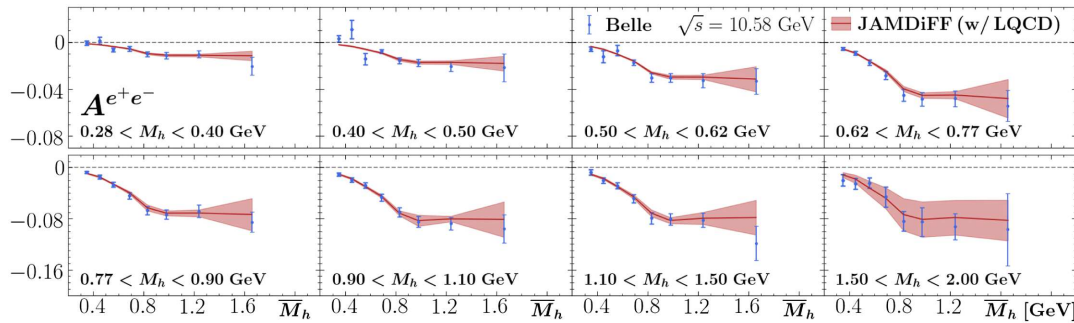


FIG. 4. Artru-Collins asymmetry data from Belle [112] binned in $(M_h, \bar{M}_h)$ at $\sqrt{s} = 10.58$ GeV (blue circles) plotted as a function of $\bar{M}_h$ against the mean JAMDiFF (w/ LQCD) result [using Eq. (4)] with 1σ uncertainty (red lines with bands). The different panels show different bins of M_h .

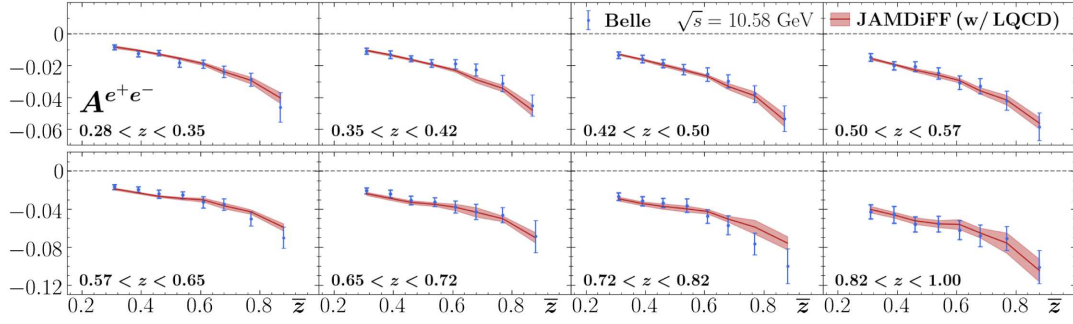


FIG. 5. Artru-Collins asymmetry data from Belle [112] binned in $(z, \bar{z})$ at $\sqrt{s} = 10.58$ GeV (blue circles) plotted as a function of $\bar{z}$ against the mean JAMDiFF (w/ LQCD) result [using Eq. (4)] with 1σ uncertainty (red lines with bands). The different panels show different bins of z .

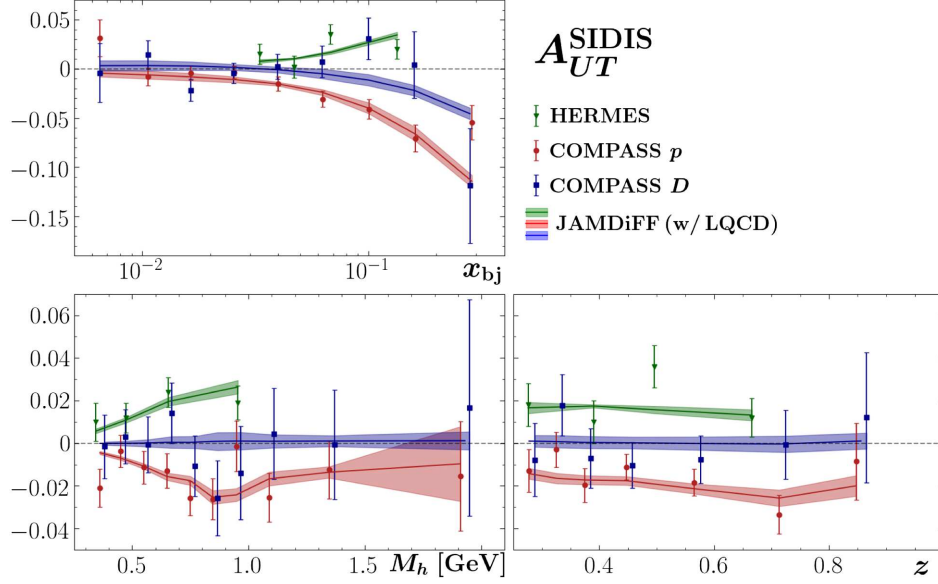


FIG. 6. SIDIS asymmetry data from HERMES [118] (green triangles) and COMPASS [117] (red circles for proton, blue squares for deuteron) plotted against the mean JAMDiFF (w/ LQCD) result [using Eq. (5)] with 1σ uncertainty (colored lines with bands). The data is binned in x_{bj} (top left), M_h (bottom left), and z (bottom right). We note that the asymmetries from HERMES and COMPASS are defined such that they have opposite signs [see Eq. (5) and the surrounding discussion].

Sec. VI B), which is apparent in the reduction of χ^2_{red} for the COMPASS proton x_{bj} binning from 1.98 to 0.65 when the lattice data is removed. In Sec. VI B we will elaborate on the importance of high- x experimental measurements in testing the compatibility between phenomenology and LQCD results for the tensor charges. We lastly remark that the SIDIS only fit has χ^2_{red} values similar to the analyses that included pp data, demonstrating clear compatibility between the two reactions.

The results for the pp asymmetry are shown in Figs. 7 and 10 for the STAR $\sqrt{s} = 200$ GeV data and in Figs. 8–10 for the $\sqrt{s} = 500$ GeV data. The fitted normalizations N_e [see Eq. (28)] are 1.00(3) for the 200 GeV STAR data and 1.14(6) for the 500 GeV STAR data. As with the e^+e^- and

SIDIS observables, we again find generally good agreement with the pp reaction across all kinematic binnings. The one discrepancy of note is with the η binning for the 500 GeV STAR data at the highest η value (see Fig. 10) that has more sensitivity to larger x . The JAMDiFF (w/ LQCD) fit overshoots that point and has a trend of continuing to increase with increasing η . Data at larger η are needed in order to test this behavior. The JAMDiFF (no LQCD) fit still increases with η but falls below the JAMDiFF (w/ LQCD) curve (see Fig. 19). This accounts for the improvement in the χ^2_{red} for the 500 GeV STAR data in the η binning, going from 2.97 to 1.83, once the LQCD data is removed. Interestingly, the 200 GeV STAR data for the η binning improves when LQCD data is included, going from 1.46 to

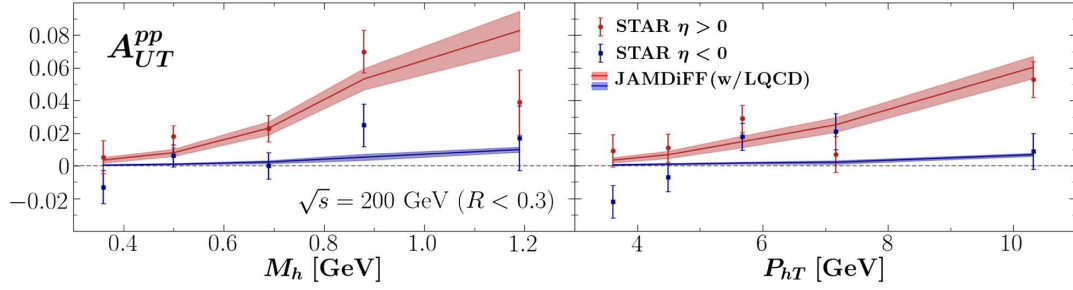


FIG. 7. Proton-proton asymmetry data from STAR at $\sqrt{s} = 200$ GeV [121] with opening angle cut $R < 0.3$ plotted against the mean JAMDiFF (w/LQCD) result [using Eq. (8)] with 1σ uncertainty (colored lines with bands). The left panel shows the results binned in M_h while the right panel shows them binned in P_{hT} , for both $\eta > 0$ (red circles) and $\eta < 0$ (blue squares).

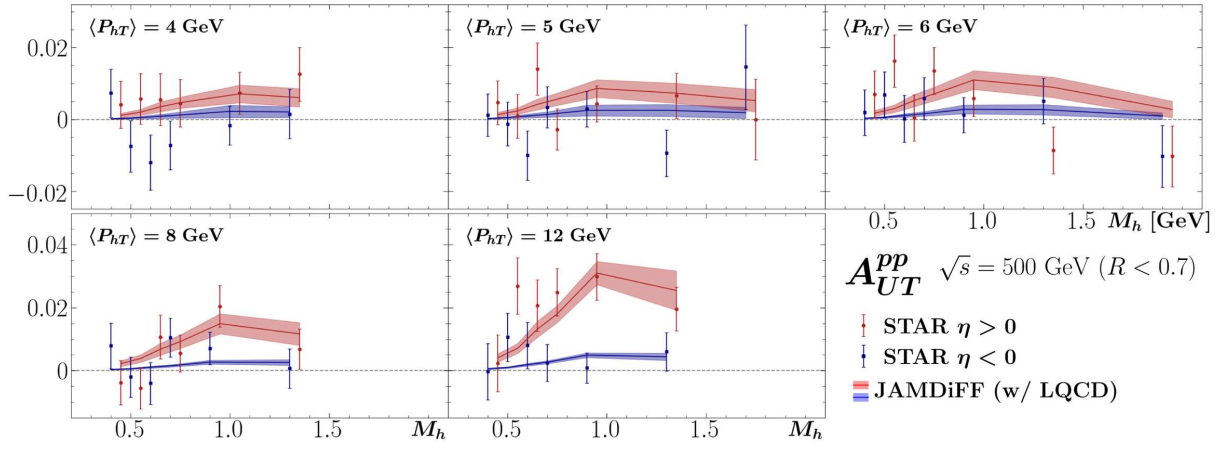


FIG. 8. Proton-proton asymmetry data from STAR at $\sqrt{s} = 500$ GeV [97] with opening angle cut $R < 0.7$ (red circles for $\eta > 0$, blue squares for $\eta < 0$) plotted as a function of M_h against the mean JAMDiFF (w/LQCD) result [using Eq. (8)] with 1σ uncertainty (colored lines with bands). The different panels show different bins of P_{hT} .

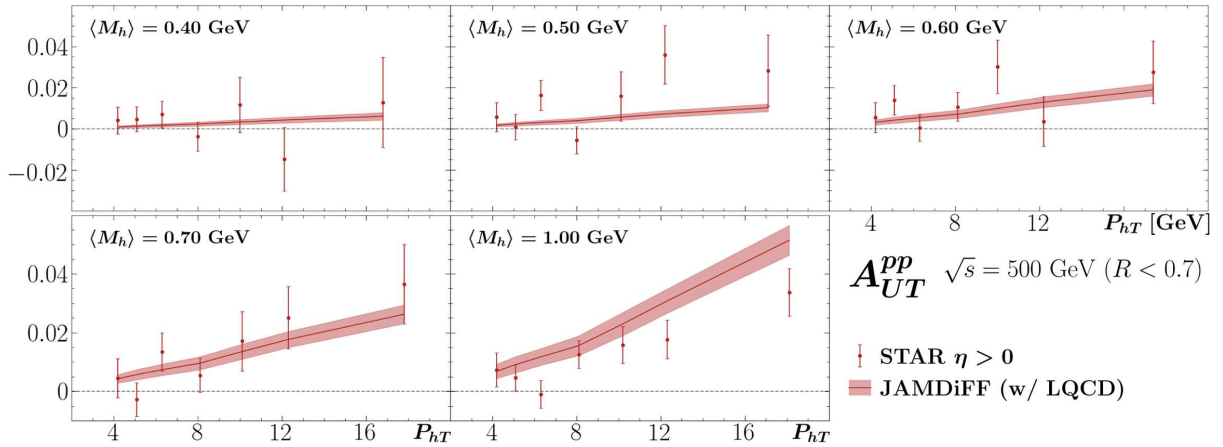


FIG. 9. Proton-proton asymmetry data from STAR at $\sqrt{s} = 500$ GeV [97] with opening angle cut $R < 0.7$ and $\eta > 0$ (red circles) plotted as a function of P_{hT} against the mean JAMDiFF (w/LQCD) result [using Eq. (8)] with 1σ uncertainty (red lines with bands). The different panels show different bins of M_h .

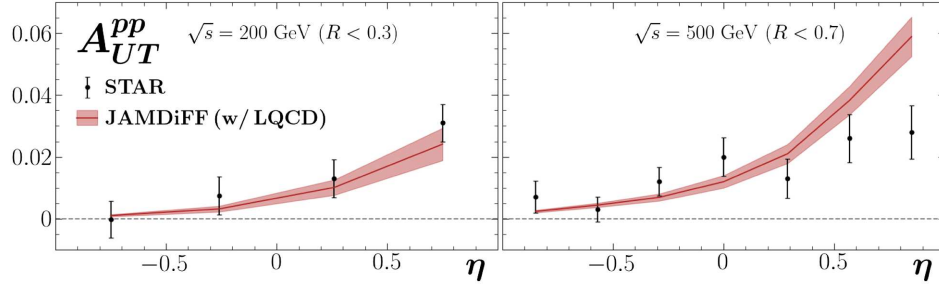


FIG. 10. Proton-proton asymmetry data from STAR (black circles) plotted as a function of η against the mean JAMDiFF (w/ LQCD) result [using Eq. (8)] with 1σ uncertainty (red lines with bands). The left panel shows the data at $\sqrt{s} = 200$ GeV [121] with opening angle cut $R < 0.3$, $\langle P_{hT} \rangle = 6.0$ GeV, $\langle M_h \rangle = 0.6$ GeV. The right panel shows the data at $\sqrt{s} = 500$ GeV [97] with opening angle cut $R < 0.7$, $\langle P_{hT} \rangle = 13$ GeV, and $\langle M_h \rangle = 1$ GeV.

0.52. So it seems that, within our analysis, there are competing trends at larger η in pp that measurements at more forward rapidities may be able to resolve.

V. EXTRACTED DIHADRON FRAGMENTATION FUNCTIONS

In this subsection we discuss our extracted DiFFs. The results shown below for the no LQCD and w/ LQCD fits are produced from over 900 replicas each. Due to our new definition [65] we do not make direct comparisons to the results of Refs. [62,63]. We start with $D_1^i(z, M_h)$, shown in Figs. 11 and 12 at the scale $\mu^2 = 100$ GeV² (approximately the scale of the Belle experiment). We generally find strong constraints on all of the quark DiFFs, but very weak constraints on the gluon DiFF. This is to be expected, as we have no observables sensitive to the gluon DiFF at LO

and instead depend entirely on evolution to constrain it. (As alluded to previously, the pp asymmetry cannot help, as it is mainly sensitive to the transversity PDFs and $H_1^\perp$ in the numerator.) In the near future, data on the unpolarized pp cross section from STAR will be available [148], providing stronger constraints on D_1^g . In the Radici, Bacchetta analysis [68] (which we will refer to as RB18), they considered three scenarios for the gluon DiFF at the input scale: $D_1^g = 0$, $D_1^g = D_1^u/4$, and $D_1^g = D_1^u$, but it is clear from Fig. 11 that none of these scenarios holds across all z and M_h . The extraction and error quantification for D_1^g shown here should provide a more realistic uncertainty propagation on the extracted transversity PDFs. For $D_1^i(z, M_h)$, one sees the resonance structure in the e^+e^- cross section data causes similar behavior in the up quark DiFF as a function of M_h . The quark DiFFs generally are larger at smaller z and M_h

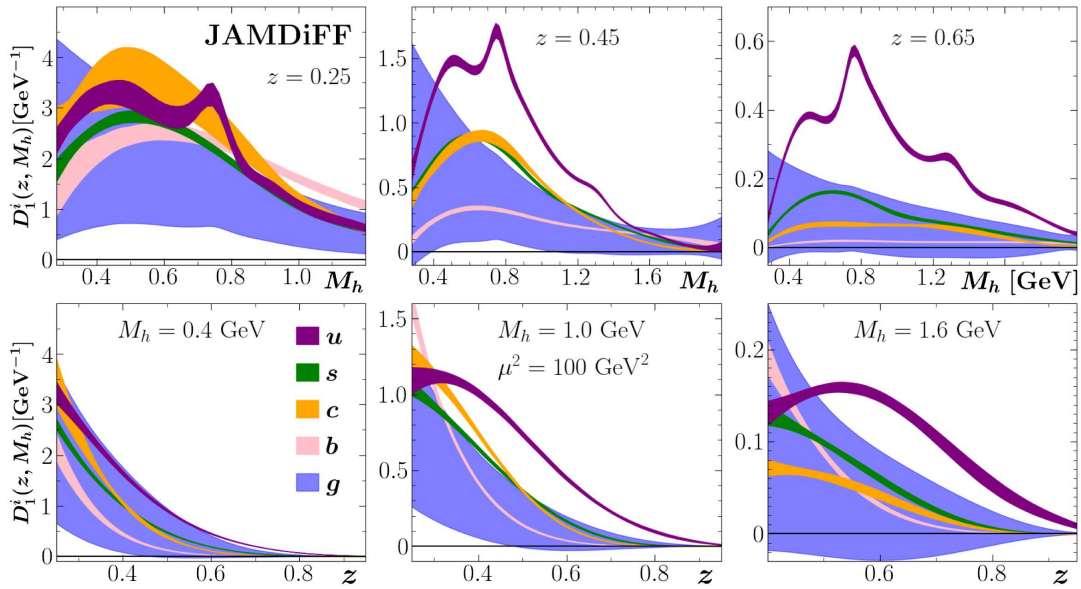


FIG. 11. DiFFs $D_1^i(z, M_h)$ from the JAMDiFF (w/ LQCD) fit plotted as a function of M_h with $z = 0.25, 0.45, 0.65$ (top row) and as a function of z with $M_h = 0.4, 1.0, 1.6$ GeV (bottom row) at the scale $\mu^2 = 100$ GeV². The up, strange, charm, bottom, and gluon are shown in purple, green, orange, pink, and blue, respectively, with 1σ uncertainty.

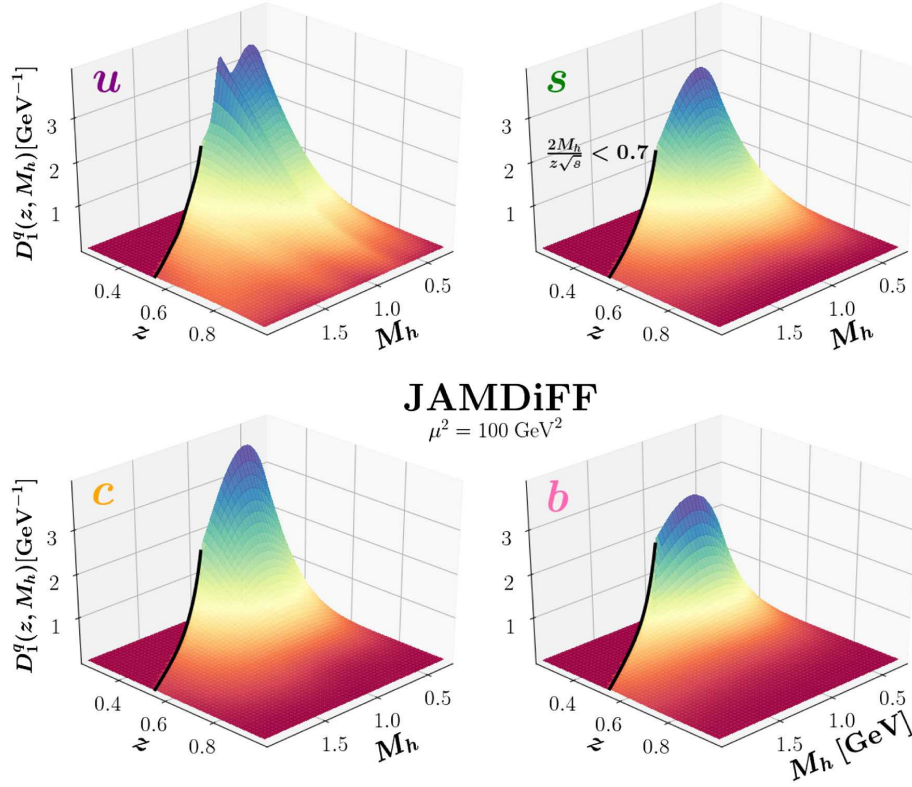


FIG. 12. DiFFs D_1^q from the JAMDiFF (w/ LQCD) fit plotted as a function of M_h and z at the scale $\mu^2 = 100 \text{ GeV}^2$. Only the mean values are shown. The black line represents the cut Eq. (35), beyond which we do not show the DiFFs.

and decrease as either of those variables increases. We expect such a behavior since the phase space to produce two hadrons from the same parton-initiated jet becomes smaller at larger z and/or M_h . Note that the different definition we use for the DiFFs does not allow for a direct quantitative comparison to be made to the extraction of $D_1^i(z, M_h)$ in Ref. [62]. Nevertheless, the qualitative features of the DiFFs as a function of z and M_h can be compared, and we find similarities to the behavior seen in Figs. 2 and 3 of Ref. [62] (with the caveat that the functions are plotted at different energy scales). In Fig. 12 we present the z and M_h dependence of D_1^q collectively in a three-dimensional plot to emphasize, given our new definition of DiFFs [65], it is a joint number density in those variables (only quarks are shown since the gluon has a significant uncertainty).

To that point, we are now able to calculate expectation values associated with dihadron fragmentation. For example, the average value of M_h for a single $\pi^+\pi^-$ pair with a given z formed from the fragmentation of a parton i is

$$\langle M_h | z \rangle^i = \frac{\int dM_h M_h D_1^i(z, M_h)}{\int dM_h D_1^i(z, M_h)}. \quad (36)$$

Similarly, the average value of z for a single $\pi^+\pi^-$ pair with a given M_h formed from the fragmentation of a parton i is

$$\langle z | M_h \rangle^i = \frac{\int dz z D_1^i(z, M_h)}{\int dz D_1^i(z, M_h)}. \quad (37)$$

These quantities are shown in Fig. 13. The general feature is that the average $M_h(z)$ increases as $z(M_h)$ becomes larger. This can perhaps be understood by the fact that the number of dihadrons produced is greater at smaller z and decreases with increasing z (see Fig. 11). We notice from the left panel of Fig. 13 that the $u, d = u, s$, and c quarks typically fragment into $\pi^+\pi^-$ pairs with similar invariant masses when the dihadron carries a small to moderate momentum fraction z . As $z \rightarrow 1$, there is a tendency for the c quark to produce a heavier $\pi^+\pi^-$ pair than the u, d , and s quarks. When more $\pi^+\pi^-$ pairs are produced (smaller z), we expect on average they will have a smaller mass, in which case the mass of the quark flavor becomes less relevant. As less $\pi^+\pi^-$ pairs are produced as one nears threshold ($z \rightarrow 1$), then the mass of the quark more directly correlates to the mass of the dihadron. The mass of the b quark being much heavier than the others, it separates itself to yield heavier dihadrons even at small z . The large uncertainties for the gluon expectation values are a reflection of the large uncertainties on D_1^g . The right panel of Fig. 13 displays a clear hierarchy of the average dihadron momentum fractions, with z decreasing from the lightest (u, d, s quarks) to the heaviest (c, b quark). This pattern

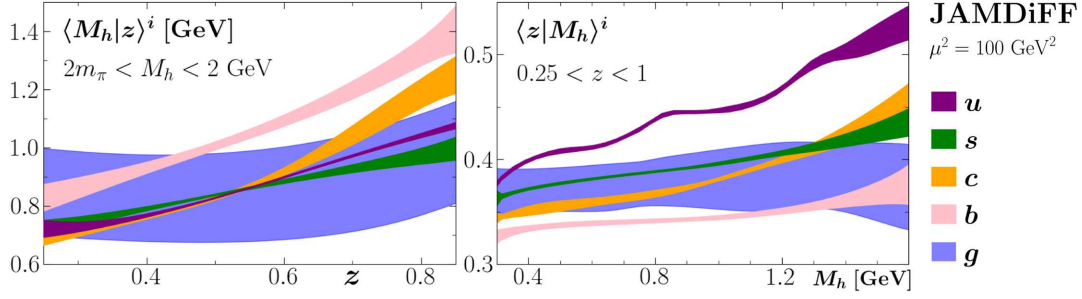


FIG. 13. Left panel: average value $\langle M_h | z \rangle^i$ with 1σ uncertainty for the ensemble of all $\pi^+\pi^-$ pairs formed from the fragmentation of a parton i , computed using $D_1^i(z, M_h)$ at $\mu^2 = 100$ GeV² [see Eq. (36)], with M_h integrated from $2m_\pi$ to 2 GeV. Right panel: similar to the left panel but for $\langle z | M_h \rangle^i$ [see Eq. (37)], with z integrated from 0.25 to 1.

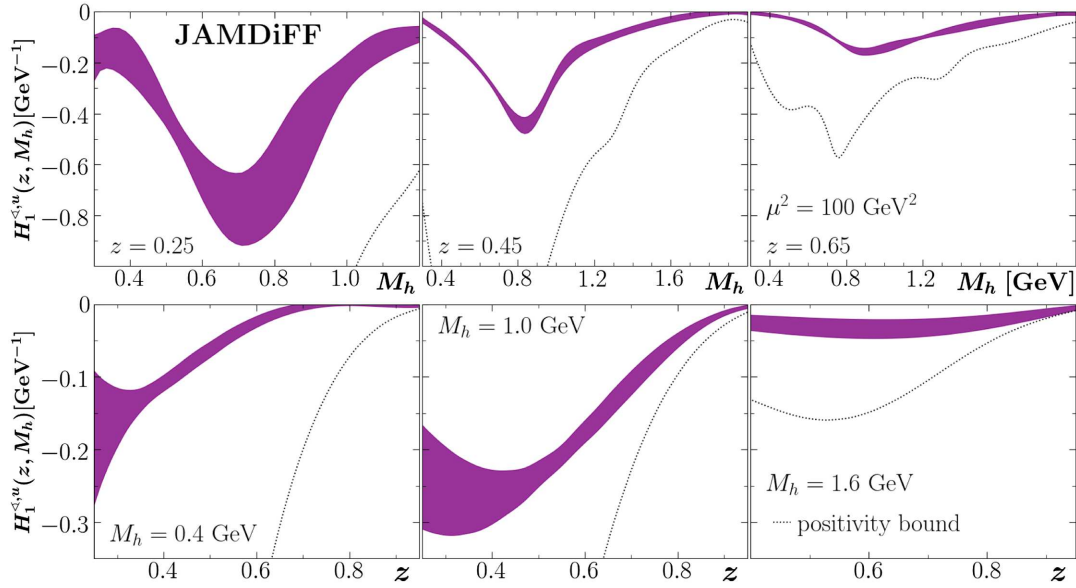


FIG. 14. IFF $H_1^{\leftarrow,u}$ from the JAMDiFF (w/ LQCD) fit plotted at 1σ uncertainty as a function of M_h with $z = 0.25, 0.45, 0.65$ (top row) and as a function of z with $M_h = 0.4, 1.0, 1.6$ GeV (bottom row) at the scale $\mu^2 = 100$ GeV². The dashed black lines represent the positivity bound $|H_1^{\leftarrow,u}| < D_1^u$.

continues until around m_c , where $\pi^+\pi^-$ pairs produced from the c quark start to carry slightly larger momentum fractions that are comparable to dihadrons produced by the s quark. The trend in the plot aligns with the idea that for a given M_h , $\pi^+\pi^-$ pairs need to carry smaller momentum fractions if they arise from the fragmentation of a heavier quark. The gluon produces dihadron pairs with the same z (≈ 0.35 – 0.4) independent of the invariant mass.

In Figs. 14 and 15 we show the IFF at the scale $\mu^2 = 100$ GeV². As mentioned in Sec. II A, the sign of $H_1^{\leftarrow,u}$ cannot be fixed by the experimental data alone, and we have thus forced $H_1^{\leftarrow,u}$ to be negative to be consistent with the expected sign of $h_1^{u_v}$. We generally find that the IFF is not constrained by the positivity bound $|H_1^{\leftarrow,u}| < D_1^u$, except at large M_h and z where its magnitude begins to be

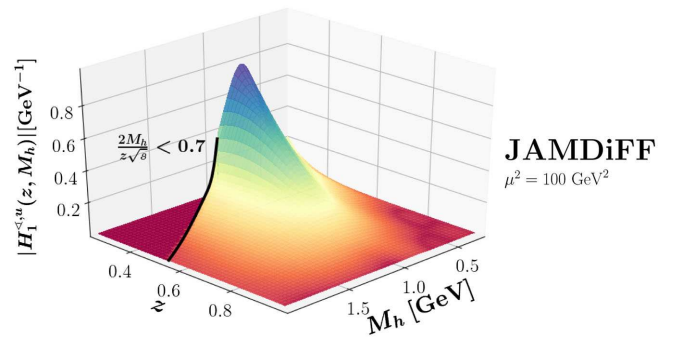


FIG. 15. IFF $|H_1^{\leftarrow,u}|$ from the JAMDiFF (w/ LQCD) fit plotted as a function of M_h and z at the scale $\mu^2 = 100$ GeV². Only the mean values are shown. The black line represents the cut Eq. (35), beyond which we do not show the IFF.

limited by the bound. The IFF decreases with increasing z and peaks at $M_h \approx 0.8$ GeV.

VI. EXTRACTED TRANSVERSITY DISTRIBUTIONS AND TENSOR CHARGES

A. Main results

As discussed in Sec. IV B, we perform three different analyses for the transversity PDFs to assess the impact of experimental datasets and LQCD on the results. In Fig. 16 we show the JAMDiFF (no LQCD) and SIDIS only results for the transversity PDFs and the tensor charges. The two fits agree within errors, with the JAMDiFF (no LQCD) fit preferring a slightly larger u_v distribution and the SIDIS only fit preferring a more negative d_v . These results translate directly into the tensor charge values, with the SIDIS only result being slightly smaller for δu and more negative for δd .

From the SIDIS only fit one can easily understand the signs of the extracted transversity PDFs. Referring to Fig. 6, one sees that the proton asymmetry from COMPASS is negative. In Eq. (6), we showed that the proton asymmetry is dominated by $h_1^{u_v}$. Since $H_1^{\leq, u}$ is (chosen to be) negative (see Fig. 14), $h_1^{u_v}$ must be positive to lead to an overall negative asymmetry. A similar argument applies to the HERMES proton asymmetry, except that the negative sign in Eq. (5) makes the asymmetry positive. For the deuteron asymmetry from COMPASS one sees that it is largely consistent with zero. From Eq. (7) the asymmetry is proportional to the sum of $h_1^{u_v}$ and $h_1^{d_v}$. Since the proton asymmetry fixes $h_1^{u_v}$ to be positive, the deuteron asymmetry then fixes $h_1^{d_v}$ to be negative so that the sum largely cancels. This qualitative finding nicely agrees with a model-independent QCD analysis for large N_c [142]. Taking the SIDIS results for

$h_1^{u_v}$ and $h_1^{d_v}$ and neglecting antiquarks, one can also understand the sign of the pp asymmetry. The measurements are at midrapidity where $\hat{s} \approx -\hat{t} \approx -\hat{u}$. The $q^\uparrow q(q') \rightarrow q^\uparrow q(q')$ and $q^\uparrow g \rightarrow q^\uparrow g$ channels give the main contributions and their hard factors in this region are negative (see Appendix A). The numerator of the asymmetry for these channels involves the combinations [cf. Eq. (9)] $H_1^{u, \leq} [f_1^{u(q')} h_1^u - f_1^{d(q')} h_1^d] < 0$ and $f_1^g H_1^{u, \leq} [h_1^u - h_1^d] < 0$, respectively. Thus, when multiplied by the negative hard factors, one ultimately obtains a positive asymmetry.

In Fig. 17 we compare our results with and without LQCD for the transversity PDFs to those from Radici and Bacchetta [68] (whose analysis did not consider the inclusion of lattice data). We also compare to a version of JAM3D that has been slightly updated from Ref. [54] (see the footnote of [149]) that we will refer to as JAM3D*. For the no LQCD results we agree with RB18 within errors, but with a larger $h_1^{u_v}$ in the region $0.04 \lesssim x \lesssim 0.3$. The overall smaller errors on our analysis can be partially attributed to the inclusion of all three SIDIS binnings (x_{bj} , z , M_h), while the RB18 analysis only included the x_{bj} binning. We note that the inclusion of the small- x constraint [Eq. (23)] and antiquarks in this analysis (neither of which were included in RB18) have no significant impact on the valence transversity distributions in the measured region. For details on the comparison to JAM3D*, see Ref. [109].

Within our analysis, the increase in $h_1^{u_v}$ in the $x \gtrsim 0.3$ region when LQCD is included is a consequence of the LQCD results for δu being larger than the result of the fit without LQCD (see Fig. 2 of Ref. [109]). The experimental data provides strong constraints in the $0.005 \lesssim x \lesssim 0.2$ region (see Fig. 1), while the small- x constraint of Eq. (23) prevents $h_1^{u_v}$ from becoming significantly larger at $x \lesssim 0.005$. Similarly, the Soffer bound does not allow

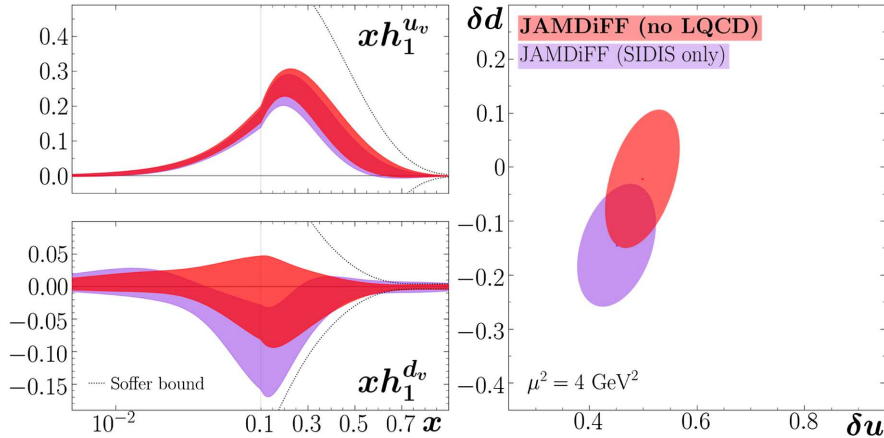


FIG. 16. Left panel: transversity PDFs $xh_1^{u_v}$ and $xh_1^{d_v}$ plotted as a function of x at the scale $\mu^2 = 4$ GeV². We show results for the JAMDiFF (no LQCD) fit (red) and SIDIS only fit (violet) with 1σ uncertainty. The Soffer bound is indicated by the dashed black lines. Right panel: the tensor charges δu and δd at the scale $\mu^2 = 4$ GeV² for the JAMDiFF (no LQCD) and SIDIS only fits with 1σ uncertainty.

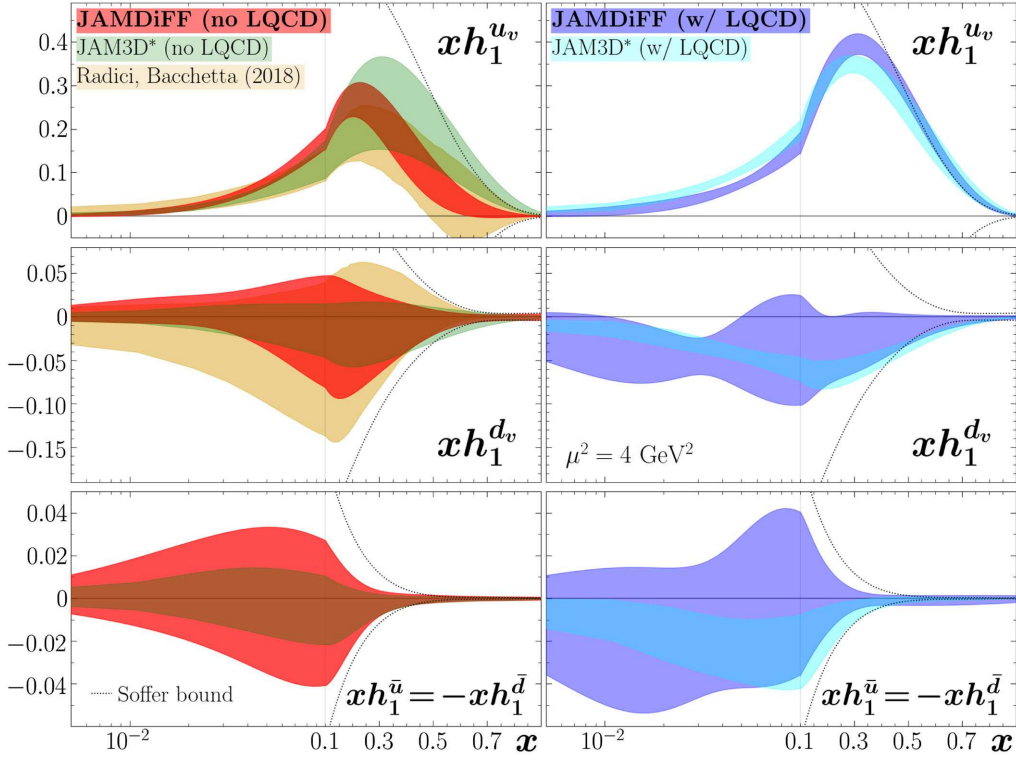


FIG. 17. Transversity PDFs $xh_1^{u_v}$ (top row), $xh_1^{d_v}$ (middle row) and $xh_1^{\bar{u}} = -xh_1^{\bar{d}}$ (bottom row) plotted as a function of x at the scale $\mu^2 = 4 \text{ GeV}^2$. Our results (JAMDiFF) are shown at 1σ both with (blue) and without (red) LQCD included in the fit and are compared to those from JAM3D* [54,149] at 1σ with (cyan) and without (green) LQCD and RB18 [68] (gold, 90% CL). The Soffer bound is indicated by the dashed black lines. Note that for JAMDiFF and JAM3D* the relation $h_1^{\bar{d}} = -h_1^{\bar{u}}$ from the large- N_c limit is enforced, while for RB18 the antiquarks are not fitted.

the transversity PDF to get large at very high x . Thus, in order to increase δu , the best option for the fit is to increase $h_1^{u_v}$ in the $x \approx 0.3$ region, although this leads to a slight deterioration in the description of the SIDIS data (see Table II), especially the COMPASS x_{bj} binning. The lack of overlap within errors between the no LQCD and w/ LQCD results will be discussed in detail in Sec. VIB. For $h_1^{d_v}$, the largest change occurs below $x \approx 0.05$, where, after the inclusion of the LQCD data, it now tends to be negative. This is a consequence of the LQCD result for δd being more negative than the result of the fit without LQCD.

In Fig. 17 we also show in the bottom row the JAMDiFF (no LQCD) result for the antiquark transversity distributions and compare to JAM3D*, where both analyses assume $h_1^{\bar{u}} = -h_1^{\bar{d}}$ (see Sec. IIIC). The two results are in agreement. The Soffer bound forces the antiquarks to be very small above $x \gtrsim 0.3$. Below that region, they still remain small and consistent with zero. The smaller uncertainties on the JAM3D* result are partly due to the less flexible parametrization used in that analysis. The inclusion of the LQCD measurements causes $h_1^{\bar{u}}$ to become slightly negative at small x , though the result is still compatible with zero. For JAM3D*, $h_1^{\bar{u}}$ is negative up to $x \approx 0.2$. While the

observation of nonzero antiquark transversity PDFs is interesting, it would be premature to assign any significance to this result.

We now move on to the tensor charges. We refer the reader to Ref. [109] for a detailed discussion of the resulting tensor charges from the no LQCD and with LQCD fits. In Fig. 18 we present a full comparison of the JAMDiFF (w/ LQCD) fit to recent LQCD results [70–73,75–81]. The phenomenological results of JAMDiFF (w/ LQCD) and JAM3D* (w/ LQCD) are generally on the lower end for g_T when considering all LQCD values, but nevertheless are in good agreement with ETMC, NME, Mainz, and LHPC. They are in reasonable agreement with PNDME, RQCD, and JLQCD, and the worst agreement is found with QCDSF/UKQCD/CSSM, PACS, and χ QCD. We thus find that our extracted g_T is compatible with most LQCD calculations. Our tensor charge results for the JAMDiFF (w/ LQCD), JAMDiFF (no LQCD), and JAMDiFF (SIDIS only) fits are summarized in Table III.

B. Further exploring compatibility between lattice QCD and experimental data

In this section we explore further the compatibility of the experimental measurements and LQCD data. We begin by

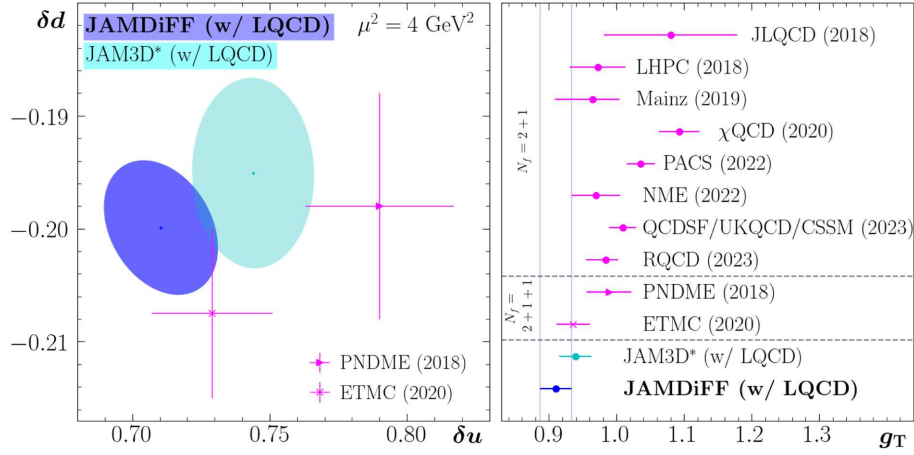


FIG. 18. The tensor charges δu , δd , and g_T at the scale $\mu^2 = 4 \text{ GeV}^2$. We show results at 1σ for the JAMDiFF (w/ LQCD) fit (blue) and compare to the JAM3D* [54,149] result with lattice QCD (cyan, 1σ) and LQCD results [70–73,75–81] (magenta points). The LQCD results are organized by N_f with $N_f = 2 + 1 + 1$ or $N_f = 2 + 1$.

discussing the resulting χ^2_{red} values for the LQCD data (see Table II). While our description of the ETMC and PNDME δd are very good, the PNDME δu has a large χ^2_{red} of 8.68 even after its inclusion in the fit. Following the methodology of NNPDF4.0 [150] (see Sec. 4.2.3), we test the compatibility of the LQCD δu data with the experimental data by reweighting its contribution to the total χ^2 , so that it is comparable to the χ^2 for the rest of the experimental data, and rerunning the fit. Specifically, we weight both ETMC and PNDME in a single fit, with weighting factor $w = (1475 + 2978)/2 = 2226.5$ for each. [The numerator is the total number of points in the fit (1475 from experimental + LQCD data, 2978 from PYTHIA data) and the denominator is the number of weighted points (the two δu points from ETMC and PNDME).] Upon doing so, we find a similar χ^2_{red} of 0.86 for ETMC δu , and a significantly lower χ^2_{red} of 2.26 for PNDME δu . Thus, the ability for the fit to describe the PNDME δu point is restricted by its tension with the ETMC result combined with the fact that the ETMC point lies closer to the JAMDiFF (no LQCD) result. The χ^2_{red} for the STAR data remains nearly identical, while the χ^2_{red} for SIDIS increases from 1.04 to 1.09. While the increases in the experimental χ^2_{red} are small, they are significant enough to prevent a better description of the PNDME δu point due to its small weight in the fit. We note

that in this fit the u_v distribution naturally becomes a bit larger (but still overlaps within errors with the w/ LQCD fit), while the rest of the PDFs and DiFFs remain stable. We will discuss in detail below the sources of the increases in the χ^2_{red} of the experimental data when LQCD information is included.

Before doing so, we mention an alternative way of treating the LQCD data, which takes into account the tension between the ETMC and PNDME δu measurements in a more conservative manner. We label this method as “JAMDiFF (w/ LQCD Flat Sampling).” Instead of fitting the ETMC and PNDME points directly, we construct a range for δu that covers both points including their 1σ errors: $0.707 < \delta u < 0.817$. We then generate two points randomly (flat sampling) within this range with the 1σ error taken to be $(0.817 - 0.707)/2 = 0.055$. We repeat this process for δd . Upon including these new points in the fit, we find $\delta u = 0.62(5)$, $\delta d = -0.195(11)$, and $g_T = 0.82(5)$. The result for δd unsurprisingly agrees with that of the w/ LQCD fit (see Table III). The result for δu , on the other hand, lies in between the no LQCD and w/ LQCD fits (1.4σ larger than the former and 1.7σ smaller than the latter) in both its central values and in the size of its errors. This is to be expected given the larger errors on the generated δu data, which have less of an impact on the transversity PDFs and cause the extracted result for δu to also have larger errors. Ultimately, we find a 1.9σ discrepancy between the extracted δu and the constructed data point $\delta u = 0.762(55)$, again demonstrating reasonable agreement between experiment and LQCD. We note that this method of fitting the LQCD data is more lenient with regard to testing the compatibility between experiment and LQCD. Our “standard” method of fitting ETMC and PNDME separately is a much more stringent test of their compatibility with experimental data and still successful.

TABLE III. Summary of tensor charge results. The errors correspond to 1σ . Note main results are set in bold.

Fit	δu	δd	g_T
JAMDiFF (w/ LQCD)	0.71(2)	−0.200(6)	0.91(2)
JAMDiFF (no LQCD)	0.50(7)	−0.02(13)	0.52(12)
JAMDiFF (SIDIS only)	0.45(7)	−0.15(13)	0.60(11)

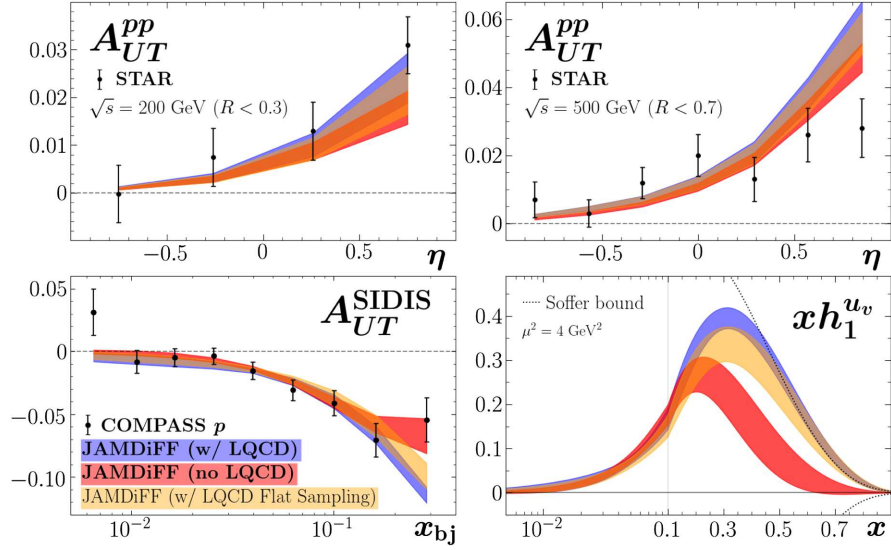


FIG. 19. Results for three different fits: JAMDiFF (w/ LQCD) (blue), JAMDiFF (no LQCD) (red), and the JAMDiFF (w/ LQCD Flat Sampling) (orange). Top left: the pp asymmetry data from STAR [121] at $\sqrt{s} = 200$ GeV binned in η (black circles) plotted against the results with 1σ uncertainty bands. Top right: same as the top left, except for the STAR data at $\sqrt{s} = 500$ GeV [97]. Bottom left: same as the top left, except for the SIDIS asymmetry data from COMPASS [117] on a proton target binned in x_{bj} . Bottom right: transversity PDF $xh_1^{u_v}$ plotted as a function of x at the scale $\mu^2 = 4$ GeV 2 with 1σ uncertainty bands. The Soffer bound is indicated by the dashed black lines.

From Table II we see that the largest changes in the χ^2_{red} going from the no LQCD to w/ LQCD fits are seen in the COMPASS proton measurement (increasing from 0.65 to 1.98), the STAR $\sqrt{s} = 500$ GeV data binned in η (increasing from 1.83 to 2.97), and the STAR $\sqrt{s} = 200$ GeV data binned in η (decreasing from 1.46 to 0.52). Thus, in Fig. 19 we show these three datasets, as well as $h_1^{u_v}$, for three different fits: no LQCD, w/ LQCD, and w/ LQCD (flat sampling). We see that the changes (in absolute terms) between the theory calculations from the fits occur almost entirely at larger x for SIDIS and larger η for pp (which likewise corresponds to probing transversity in the higher- x region). In this regime the observables are determined primarily by the $h_1^{u_v}$ distribution, and so we see a direct correspondence between the magnitude of $h_1^{u_v}$ at higher x and the magnitude of the asymmetries.

With the small- x constraint [Eq. (23)] and strong constraints from experimental data in the region $0.005 \lesssim x \lesssim 0.2$, the fit has no choice but to increase the size of $h_1^{u_v}$ in the large $x > 0.2$ region in its attempt to accommodate the δu data from LQCD. Within our framework, this puts it in tension with a few experimental data points, most notably the highest- x COMPASS proton point and the highest- η STAR $\sqrt{s} = 500$ GeV point (but in better agreement with the highest- η STAR $\sqrt{s} = 200$ GeV point). Therefore, in order to further test the compatibility between LQCD and experimental data, it is of vital importance to have more measurements at larger x (for SIDIS) and more forward rapidity (for pp). We see that, in the no LQCD fit, $h_1^{u_v}$ has a

maximum and then begins to decrease around where the x coverage of the experimental data ends ($x \approx 0.3$) due to the highest- x COMPASS data point. Within our parametrization, the PDFs fall off smoothly and monotonically as $x \rightarrow 1$, and this drives the behavior (and uncertainty) of $h_1^{u_v}$ in the unmeasured ($x > 0.3$) region (i.e., $h_1^{u_v}$ cannot suddenly increase at larger $x > 0.3$ once it begins its downturn). The fit with LQCD included has additional constraints at larger x due to the fact that one integrates from $x \in [0, 1]$ to calculate the tensor charges. As mentioned previously, this causes $h_1^{u_v}$ to now peak at slightly higher $x \approx 0.35$ in order to accommodate both LQCD and experimental data. The Soffer bound forces the with LQCD $h_1^{u_v}$ to then decrease shortly after $x = 0.35$ and the PDF again falls off smoothly and monotonically as $x \rightarrow 1$. The LQCD data influences the behavior and uncertainty (making it smaller) in the large- x region in order for $h_1^{u_v}$ to give the appropriate value for δu . This in part helps to explain why, in our analysis, there is not overlap in the $h_1^{u_v}$ and δu error bands between the JAMDiFF (w/ LQCD) and JAMDiFF (no LQCD) fits.

In addition, we have seen in Fig. 19 that the LQCD data and STAR $\sqrt{s} = 200$ GeV data have a preference for a larger $h_1^{u_v}$ at large x , while the COMPASS proton data and STAR $\sqrt{s} = 500$ GeV data prefer a smaller $h_1^{u_v}$. In such a situation where there are competing preferences, and we compare analyses containing different subsets of the data, the choice of likelihood function $\mathcal{L}$ and prior π do not guarantee that the fits overlap within statistical uncertainties

(see, e.g., Ref. [151]). Before drawing a conclusion about the compatibility between LQCD tensor charges and experimental data, one needs first to include both in the analysis. One should only be concerned if the description of the lattice data remains poor even after its inclusion and/or if the description of the experimental data suffers significantly. Indeed, we have performed several comprehensive tests that show we are able to describe the LQCD results while not significantly sacrificing the description of the experimental data, showing their compatibility.

VII. SUMMARY AND OUTLOOK

We have presented the most comprehensive study of dihadron observables (JAMDiFF) to date by performing, for the first time, a simultaneous global QCD analysis of the $\pi^+\pi^-$ DiFFs/IFFs and transversity PDFs from e^+e^- annihilation, SIDIS, and pp collisions while including the Belle cross section data [64], the latest pp dihadron measurements from STAR [97], and all kinematic variable binnings for the relevant processes under consideration. Our extracted transversity PDFs and DiFFs at any kinematics can be found in a `github` library [152] and a `google colab` notebook [153] that provide a platform to independently access our results. We utilized a parametrization for $D_1(z, M_h)$ [and $H_1^\perp(z, M_h)$] which allowed an accurate description of the $d\sigma/dz dM_h$ data from Belle and events generated by PYTHIA while remaining computationally efficient. Studying multiple center-of-mass energies with the latter allowed for the first extraction of the gluon $D_1(z, M_h)$. From the extracted quark and gluon $D_1(z, M_h)$, where we use a definition that now has a number density interpretation [65], we made the first calculations of expectation values for the $\pi^+\pi^-$ dihadron invariant mass (as a function of z) and momentum fraction (as a function of M_h), studying also how these vary with parton flavor.

Furthermore, we have tested the compatibility of different techniques used for determining the tensor charges of the nucleon. By incorporating theory constraints on the transversity PDFs at small x and large x we were able to meaningfully include LQCD results for δu and δd from ETMC [74] and PNDME [71] into our analysis. We found compatibility with this data while maintaining a very good description of the experimental measurements. In addition, our results for δu , δd and the x dependence of transversity match closely to those from the single-hadron TMD/collinear twist-3 analysis of JAM3D [54,149]. We have thus demonstrated, for the first time, the universal nature of all available information on the transversity PDFs and tensor charges of the nucleon. We also note that we extracted information about the antiquark transversity PDFs, which had not been done previously for dihadron production.

There are several future directions for extending this analysis. More data is expected from STAR on not only the dihadron azimuthal asymmetry but also the unpolarized cross section. The latter will be important for providing the first experimental constraint on the gluon $D_1(z, M_h)$. We also anticipate measurements of dihadron multiplicities in SIDIS to offer further experimental quark flavor separation of $D_1(z, M_h)$. Furthermore, it is mandatory to extend the theoretical framework to NLO. Generally, NLO corrections to cross sections can be large, and they will certainly modify our quantitative results. However, we expect that our main qualitative findings about the transversity PDFs and the tensor charges will not be affected, as the relevant observables are asymmetries for which typically significant cancellations of higher-order corrections occur. Nevertheless, a NLO analysis is needed to provide a definitive answer to that question. Overall, future measurements at STAR, Jefferson Lab, and the Electron-Ion Collider can provide crucial information to help reduce uncertainties on the transversity PDFs and tensor charges [154,155]. In particular, we identified the large- x (in the case of SIDIS) and forward rapidity (in the case of pp) regions as key to further test the compatibility between LQCD results for the tensor charges and experiment. Measurements in the small- x region will also help test the small- x asymptotic behavior of transversity imposed in our analysis and give further insight on the antiquark transversity PDFs. Moreover, dihadron, single-hadron TMD/collinear twist-3 observables, and LQCD (tensor charge and x -dependent transversity computations [77,156]) should eventually be fit simultaneously in a “universal” analysis.

ACKNOWLEDGMENTS

We thank Wally Melnitchouk for helpful discussions, Yuri Kovchegov for useful exchanges on the small- x behavior of transversity, and Anna Martin and Gunar Schnell for clarification on the COMPASS and HERMES data, respectively. This work was supported by the National Science Foundation under Grants No. PHY-2110472 (C. C. and A. M.), No. PHY-2011763 and No. PHY-2308567 (D. P.), and No. PHY-2012002, No. PHY-2310031, No. PHY-2335114 (A. P.), and the U.S. Department of Energy Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab (A. P. and N. S.). The work of N. S. was supported by the DOE, Office of Science, Office of Nuclear Physics in the Early Career Program. The work of C. C., A. M., and A. P. was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, within the framework of the TMD Topical Collaboration, and by Temple University (C. C. and A. P.).

APPENDIX A: PARTONIC CROSS SECTIONS IN PROTON-PROTON COLLISIONS

Here we list the partonic cross sections that enter Eq. (9). The unpolarized partonic cross sections match exactly those from the Appendix of Ref. [123]. For the transversely polarized partonic cross sections, we use results that are similar to those of Ref. [123], except with an opposite overall sign on the four channels involving only quarks. Furthermore, because we define proton A to be the transversely polarized proton, rather than proton B , we interchange $\hat{u} \leftrightarrow \hat{t}$ relative to Ref. [123]. This leads to the following for the partonic cross sections involving quark flavors q, q' :

$$\frac{d\Delta\hat{\sigma}_{q^\dagger q \rightarrow q^\dagger q}}{d\hat{t}} = \frac{8\pi\alpha_s^2}{27\hat{s}^2} \frac{\hat{s}(3\hat{u} - \hat{t})}{\hat{t}^2}, \quad (\text{A1})$$

$$\frac{d\Delta\hat{\sigma}_{q^\dagger q' \rightarrow q^\dagger q'}}{d\hat{t}} = \frac{d\Delta\hat{\sigma}_{q^\dagger \bar{q}' \rightarrow q^\dagger \bar{q}'}}{d\hat{t}} = \frac{8\pi\alpha_s^2}{9\hat{s}^2} \frac{\hat{u}\hat{s}}{\hat{t}^2}, \quad (\text{A2})$$

$$\frac{d\Delta\hat{\sigma}_{q^\dagger \bar{q} \rightarrow q^\dagger \bar{q}}}{d\hat{t}} = \frac{8\pi\alpha_s^2}{27\hat{s}^2} \frac{\hat{u}(3\hat{s} - \hat{t})}{\hat{t}^2}, \quad (\text{A3})$$

$$\frac{d\Delta\hat{\sigma}_{\bar{q}^\dagger q \rightarrow \bar{q}^\dagger q}}{d\hat{t}} = \frac{8\pi\alpha_s^2}{27\hat{s}^2}, \quad (\text{A4})$$

$$\frac{d\Delta\hat{\sigma}_{q^\dagger g \rightarrow q^\dagger g}}{d\hat{t}} = -\frac{8\pi\alpha_s^2}{9\hat{s}^2} \left(1 - \frac{9}{4} \frac{\hat{u}\hat{s}}{\hat{t}^2}\right). \quad (\text{A5})$$

The hard factors involving antiquark fragmentation are identical to the hard factors of the corresponding charge-conjugated partonic processes. We note these are also consistent with the partonic cross sections associated with the “derivative term” of the fragmentation part of A_N in $p^\uparrow p \rightarrow hX$ [94].

APPENDIX B: SYMMETRY OF $\pi^+\pi^-$ DiFFs

We provide details on deriving the symmetry relations for $\pi^+\pi^-$ pairs that allow us to reduce the number of DiFFs to be fitted, which is an important aspect of the analysis. For this discussion, it is necessary to introduce the notation $D_1^{h_1 h_2}$ and $H_1^{\triangleleft, h_1 h_2}$, where $h_1 h_2$ indicates the dihadron pair. First, under the interchange of h_1 and h_2 one has the relations

$$\begin{aligned} D_1^{\pi^+\pi^-/q} &= D_1^{\pi^-\pi^+/q}, \\ H_1^{\triangleleft, \pi^+\pi^-/q} &= -H_1^{\triangleleft, \pi^-\pi^+/q}, \end{aligned} \quad (\text{B1})$$

for all quarks q . The sign change for $H_1^{\triangleleft}$ occurs due to the fact that interchanging the hadrons switches the sign of $\vec{R}_T$, which appears in the prefactor of $H_1^{\triangleleft}$ in its correlator definition, but the correlator itself is independent of the

ordering of h_1 and h_2 [56,65]. From charge-conjugation symmetry, one also has the relations

$$\begin{aligned} D_1^{\pi^+\pi^-/q} &= D_1^{\pi^-\pi^+/\bar{q}}, \\ H_1^{\triangleleft, \pi^+\pi^-/q} &= H_1^{\triangleleft, \pi^-\pi^+/\bar{q}}, \end{aligned} \quad (\text{B2})$$

for all quarks q . Combining Eqs. (B1) and (B2) to eliminate the $\pi^-\pi^+$ DiFFs, one has

$$\begin{aligned} D_1^{\pi^+\pi^-/q} &= D_1^{\pi^+\pi^-/\bar{q}}, \\ H_1^{\triangleleft, \pi^+\pi^-/q} &= -H_1^{\triangleleft, \pi^+\pi^-/\bar{q}}. \end{aligned} \quad (\text{B3})$$

From isospin symmetry (which holds to sufficient accuracy for this discussion), one also has the relations for the light quarks

$$\begin{aligned} D_1^{\pi^+\pi^-/u} &= D_1^{\pi^-\pi^+/d}, & D_1^{\pi^+\pi^-/\bar{u}} &= D_1^{\pi^-\pi^+/\bar{d}}, \\ H_1^{\triangleleft, \pi^+\pi^-/u} &= H_1^{\triangleleft, \pi^-\pi^+/d}, & H_1^{\triangleleft, \pi^+\pi^-/\bar{u}} &= H_1^{\triangleleft, \pi^-\pi^+/\bar{d}}. \end{aligned} \quad (\text{B4})$$

Using Eqs. (B1) and (B4), one has

$$\begin{aligned} D_1^{\pi^+\pi^-/u} &= D_1^{\pi^+\pi^-/d}, & D_1^{\pi^+\pi^-/\bar{u}} &= D_1^{\pi^+\pi^-/\bar{d}}, \\ H_1^{\triangleleft, \pi^+\pi^-/u} &= -H_1^{\triangleleft, \pi^+\pi^-/d}, & H_1^{\triangleleft, \pi^+\pi^-/\bar{u}} &= -H_1^{\triangleleft, \pi^+\pi^-/\bar{d}}. \end{aligned} \quad (\text{B5})$$

Combining the above with Eq. (B3) leads to the final relations for the light quarks

$$\begin{aligned} D_1^{\pi^+\pi^-/u} &= D_1^{\pi^+\pi^-/d} = D_1^{\pi^+\pi^-/\bar{u}} = D_1^{\pi^+\pi^-/\bar{d}}, \\ H_1^{\triangleleft, \pi^+\pi^-/u} &= -H_1^{\triangleleft, \pi^+\pi^-/d} = -H_1^{\triangleleft, \pi^+\pi^-/\bar{u}} = H_1^{\triangleleft, \pi^+\pi^-/\bar{d}}. \end{aligned} \quad (\text{B6})$$

For the heavier quarks (s, c , and b), which are not valence quarks of $\pi^+\pi^-$, one can assume

$$H_1^{\triangleleft, \pi^+\pi^-/q} = H_1^{\triangleleft, \pi^+\pi^-/\bar{q}}, \quad (q = s, c, b) \quad (\text{B7})$$

from which it follows, using Eq. (B3), that

$$H_1^{\triangleleft, \pi^+\pi^-/q} = -H_1^{\triangleleft, \pi^+\pi^-/q} = 0, \quad (q = s, c, b, \bar{s}, \bar{c}, \bar{b}). \quad (\text{B8})$$

In summary, one has the following symmetry relations (see also Ref. [62]) for the $\pi^+\pi^-$ DiFFs:

$$\begin{aligned} D_1^u &= D_1^d = D_1^{\bar{u}} = D_1^{\bar{d}}, \\ D_1^s &= D_1^{\bar{s}}, & D_1^c &= D_1^{\bar{c}}, & D_1^b &= D_1^{\bar{b}}, \end{aligned} \quad (\text{B9})$$

TABLE IV. Summary of the different PYTHIA 6 tunes used to generate the PYTHIA data. The “PYTHIA 6 default” tune uses the corresponding default parameters shown above. The “PYTHIA 6 Aleph” and “PYTHIA 6 LEP/Tev.” tunes modify those parameters as shown in their respective columns (if the row is blank then the default value is used). For the “PYTHIA 8” tune (not shown in the table) we use all default parameters.

Parameter	PYTHIA 6 default	PYTHIA 6 ALEPH	PYTHIA 6 LEP/Tevatron
PARJ(1)	0.1	0.106	0.073
PARJ(2)	0.3	0.285	0.2
PARJ(3)	0.4	0.71	0.94
PARJ(4)	0.05	0.05	0.032
PARJ(11)	0.5	0.55	0.31
PARJ(12)	0.6	0.47	0.4
PARJ(13)	0.75	0.65	0.54
PARJ(14)	0	0.02	
PARJ(15)	0	0.04	
PARJ(16)	0	0.02	
PARJ(17)	0	0.2	
PARJ(19)	1	0.57	
PARJ(21)	0.36	0.37	0.325
PARJ(25)	1		0.63
PARJ(26)	0.4	0.27	0.12
PARJ(33)	0.8	0.8	0.8
PARJ(41)	0.3	0.4	0.5
PARJ(42)	0.58	0.796	0.6
PARJ(45)	0.5		
PARJ(46)	1		
PARJ(47)	1		
PARJ(54)	−0.05	−0.04	−0.05
PARJ(55)	−0.005	−0.0035	−0.005
PARJ(81)	0.29	0.292	0.29
PARJ(82)	1	1.57	1.65
MSTJ(11)	4	3	5
MSTJ(12)	2	3	
MSTJ(26)	2	2	2
MSTJ(45)	5		
MSTJ(107)	0	0	0

and $\pi^+\pi^-$ IFFs

$$H_1^{\leftarrow,u} = -H_1^{\leftarrow,d} = -H_1^{\leftarrow,\bar{u}} = H_1^{\leftarrow,\bar{d}},$$

$$H_1^{\leftarrow,s} = H_1^{\leftarrow,\bar{s}} = H_1^{\leftarrow,c} = H_1^{\leftarrow,\bar{c}} = H_1^{\leftarrow,b} = H_1^{\leftarrow,\bar{b}} = 0, \quad (\text{B10})$$

where we have again dropped the $\pi^+\pi^-$ superscript.

APPENDIX C: COMPARISON TO PYTHIA-GENERATED DATA

Here we provide more information related to the data generated from PYTHIA [145,146] used in the analysis of the DiFF D_1 (see Sec. IV A). In Table IV we give the details of the different PYTHIA tunes used to estimate the systematic errors on the generated data. In Table V we present the χ^2_{red} of those data. Figures 20–24 below show the results for the five different energies used.

TABLE V. Summary of χ^2_{red} values from the JAMDiFF (w/ LQCD) fit for the PYTHIA datasets where the observable is $\sigma^q/\sigma^{\text{tot}}$.

$\sqrt{s}$	q	N_{dat}	χ^2_{red}
10.58	s	200	4.00
10.58	c	198	1.60
30.73	s	229	0.37
30.73	c	225	0.66
30.73	b	199	0.26
50.88	s	223	0.57
50.88	c	222	0.70
50.88	b	203	0.22
71.04	s	223	0.58
71.04	c	217	0.55
71.04	b	203	0.29
91.19	s	223	0.59
91.19	c	209	0.67
91.19	b	204	0.34
Total		2978	0.80

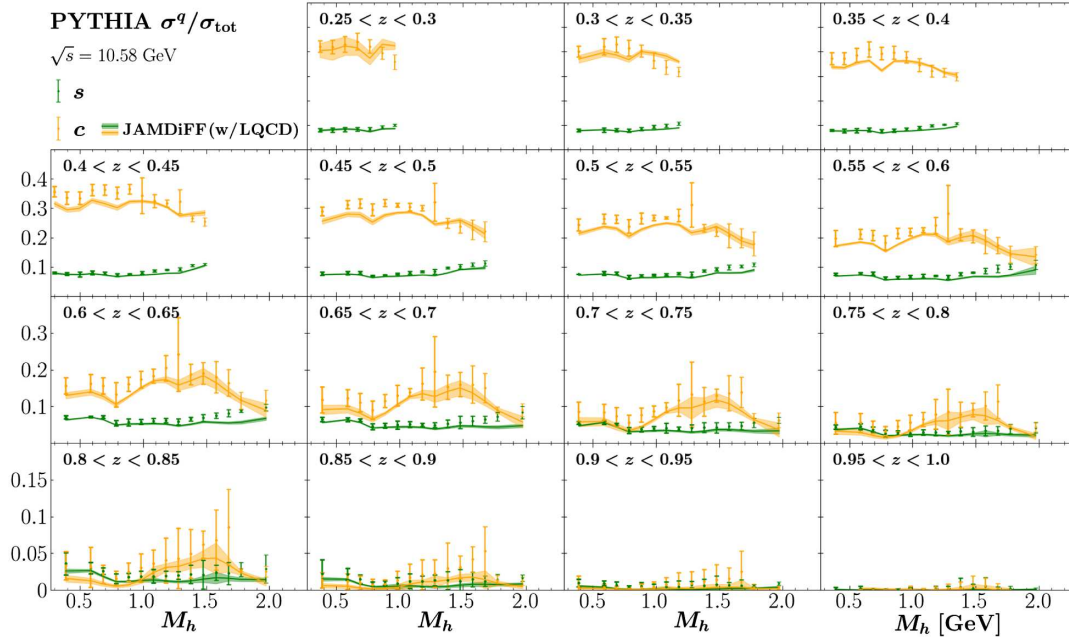


FIG. 20. PYTHIA-generated data at $\sqrt{s} = 10.58 \text{ GeV}$. The strange (green points) and charm (orange points) cross section ratios are plotted as a function of M_h and compared to the mean JAMDiFF (w/ LQCD) results with 1σ uncertainty bands (colored bands). The different panels show different bins of z .

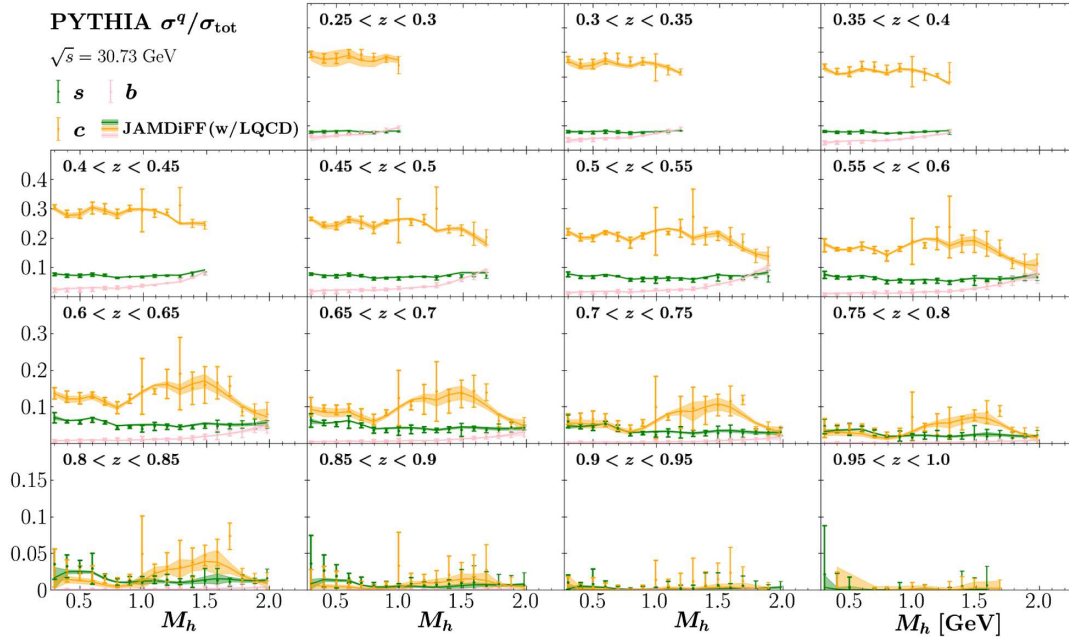
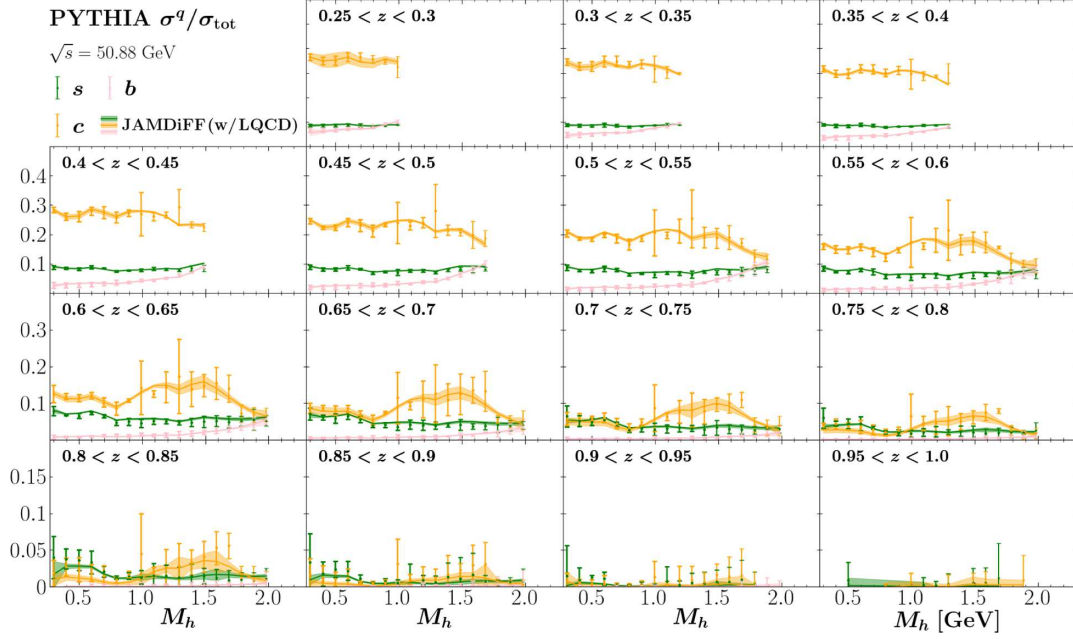
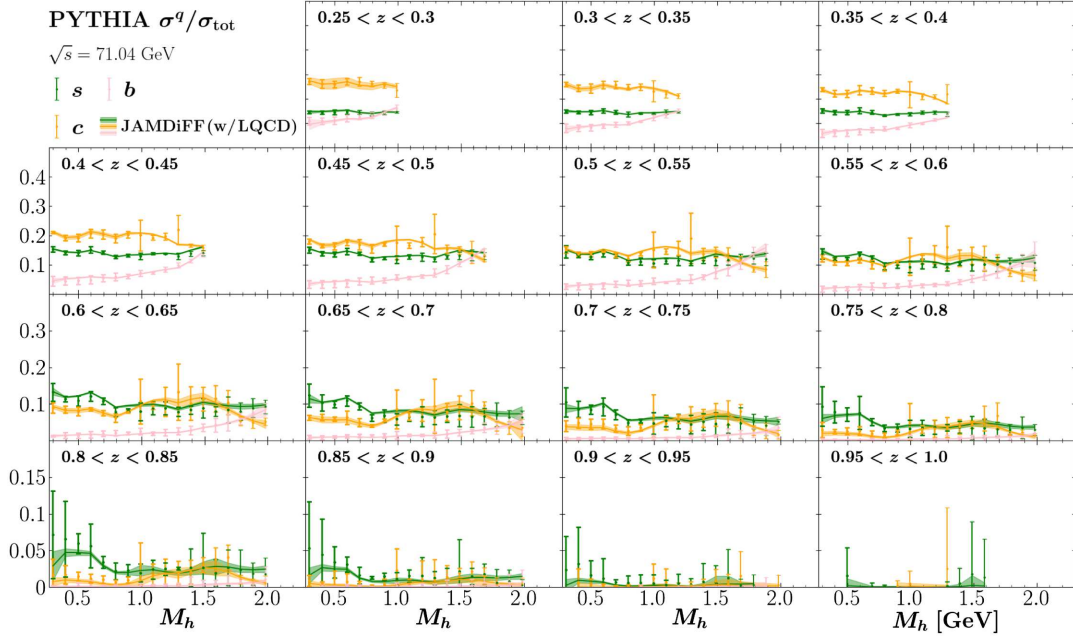


FIG. 21. PYTHIA-generated data at $\sqrt{s} = 30.73 \text{ GeV}$. The strange (green points), charm (orange points), and bottom (pink points) cross section ratios are plotted as a function of M_h and compared to the mean JAMDiFF (w/ LQCD) results with 1σ uncertainty bands (colored bands). The different panels show different bins of z .

FIG. 22. PYTHIA-generated data. Same as Fig. 21 but for $\sqrt{s} = 50.88$ GeV.FIG. 23. PYTHIA-generated data. Same as Fig. 21 but for $\sqrt{s} = 71.04$ GeV.

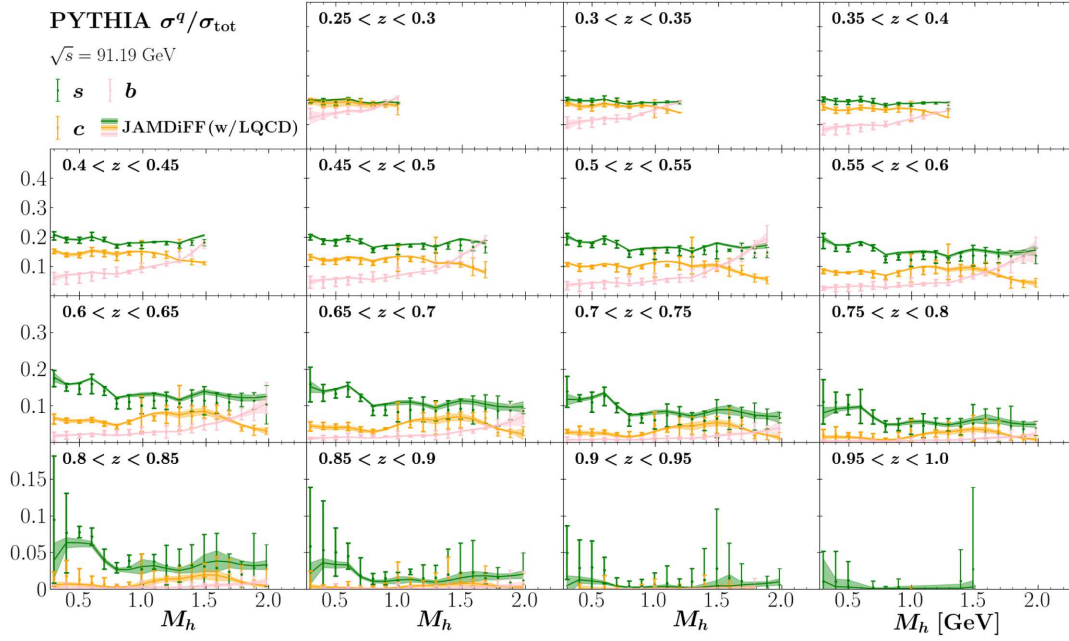


FIG. 24. PYTHIA-generated data. Same as Fig. 21 but for $\sqrt{s} = 91.19$ GeV.

-
- [1] J. J. Aubert *et al.* (European Muon Collaboration), *Phys. Lett.* **160B**, 417 (1985).
 - [2] M. Arneodo *et al.* (European Muon Collaboration), *Nucl. Phys.* **B321**, 541 (1989).
 - [3] P. Chiappetta, M. Greco, J. P. Guillet, S. Rolli, and M. Werlen, *Nucl. Phys.* **B412**, 3 (1994).
 - [4] J. Binnewies, B. A. Kniehl, and G. Kramer, *Z. Phys. C* **65**, 471 (1995).
 - [5] J. Binnewies, B. A. Kniehl, and G. Kramer, *Phys. Rev. D* **52**, 4947 (1995).
 - [6] J. Binnewies, B. A. Kniehl, and G. Kramer, *Phys. Rev. D* **53**, 3573 (1996).
 - [7] D. de Florian, M. Stratmann, and W. Vogelsang, *Phys. Rev. D* **57**, 5811 (1998).
 - [8] S. Kretzer, *Phys. Rev. D* **62**, 054001 (2000).
 - [9] L. Bourhis, M. Fontannaz, J. P. Guillet, and M. Werlen, *Eur. Phys. J. C* **19**, 89 (2001).
 - [10] B. A. Kniehl, G. Kramer, and B. Potter, *Nucl. Phys.* **B582**, 514 (2000).
 - [11] S. Kretzer, E. Leader, and E. Christova, *Eur. Phys. J. C* **22**, 269 (2001).
 - [12] C. Bourrely and J. Soffer, *Phys. Rev. D* **68**, 014003 (2003).
 - [13] M. Hirai, S. Kumano, T. H. Nagai, and K. Sudoh, *Phys. Rev. D* **75**, 094009 (2007).
 - [14] D. de Florian, R. Sassot, and M. Stratmann, *Phys. Rev. D* **75**, 114010 (2007).
 - [15] D. de Florian, R. Sassot, and M. Stratmann, *Phys. Rev. D* **76**, 074033 (2007).
 - [16] S. Albino, B. A. Kniehl, and G. Kramer, *Nucl. Phys.* **B803**, 42 (2008).
 - [17] C. A. Aidala, F. Ellinghaus, R. Sassot, J. P. Seele, and M. Stratmann, *Phys. Rev. D* **83**, 034002 (2011).
 - [18] D. de Florian, R. Sassot, M. Epele, R. J. Hernández-Pinto, and M. Stratmann, *Phys. Rev. D* **91**, 014035 (2015).
 - [19] E. Leader, A. V. Sidorov, and D. B. Stamenov, *Phys. Rev. D* **93**, 074026 (2016).
 - [20] D. P. Anderle, F. Ringer, and M. Stratmann, *Phys. Rev. D* **92**, 114017 (2015).
 - [21] C. Adolph *et al.* (COMPASS Collaboration), *Phys. Lett. B* **764**, 1 (2017).
 - [22] N. Sato, J. J. Ethier, W. Melnitchouk, M. Hirai, S. Kumano, and A. Accardi, *Phys. Rev. D* **94**, 114004 (2016).
 - [23] D. de Florian, M. Epele, R. J. Hernandez-Pinto, R. Sassot, and M. Stratmann, *Phys. Rev. D* **95**, 094019 (2017).
 - [24] J. J. Ethier, N. Sato, and W. Melnitchouk, *Phys. Rev. Lett.* **119**, 132001 (2017).
 - [25] V. Bertone, S. Carrazza, N. P. Hartland, E. R. Nocera, and J. Rojo (NNPDF Collaboration), *Eur. Phys. J. C* **77**, 516 (2017).
 - [26] M. Soleymaninia, M. Goharipour, and H. Khanpour, *Phys. Rev. D* **98**, 074002 (2018).

- [27] V. Bertone, N. P. Hartland, E. R. Nocera, J. Rojo, and L. Rottoli (NNPDF Collaboration), *Eur. Phys. J. C* **78**, 651 (2018).
- [28] M. Soleymaninia, M. Goharipour, H. Khanpour, and H. Spiesberger, *Phys. Rev. D* **103**, 054045 (2021).
- [29] E. Moffat, W. Melnitchouk, T. C. Rogers, and N. Sato (Jefferson Lab Angular Momentum (JAM) Collaboration), *Phys. Rev. D* **104**, 016015 (2021).
- [30] R. A. Khalek, V. Bertone, and E. R. Nocera (MAP (Multi-dimensional Analyses of Partonic distributions) Collaboration), *Phys. Rev. D* **104**, 034007 (2021).
- [31] H. Abdolmaleki, M. Soleymaninia, H. Khanpour, S. Amoroso, F. Giuli, A. Glazov, A. Luszczak, F. Olness, and O. Zenaiev (xfitter Developers' Team), *Phys. Rev. D* **104**, 056019 (2021).
- [32] I. Borsa, D. de Florian, R. Sassot, and M. Stratmann, *Phys. Rev. D* **105**, L031502 (2022).
- [33] I. Borsa, R. Sassot, D. de Florian, M. Stratmann, and W. Vogelsang, *Phys. Rev. Lett.* **129**, 012002 (2022).
- [34] M. Soleymaninia, H. Hashamipour, and H. Khanpour, *Phys. Rev. D* **105**, 114018 (2022).
- [35] R. Abdul Khalek, V. Bertone, A. Khoukli, and E. R. Nocera, *Phys. Lett. B* **834**, 137456 (2022).
- [36] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato (Jefferson Lab Angular Momentum (JAM) Collaboration), *Phys. Rev. D* **106**, L031502 (2022).
- [37] J. C. Collins, *Nucl. Phys.* **B396**, 161 (1993).
- [38] M. Anselmino, M. Boglione, U. D'Alesio, A. Kotzinian, F. Murgia, and A. Prokudin, *Phys. Rev. D* **71**, 074006 (2005).
- [39] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, *J. High Energy Phys.* **04** (2014) 005.
- [40] A. Signori, A. Bacchetta, M. Radici, and G. Schnell, *J. High Energy Phys.* **11** (2013) 194.
- [41] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, *J. High Energy Phys.* **06** (2017) 081; **06** (2019) 51(E).
- [42] I. Scimemi and A. Vladimirov, *J. High Energy Phys.* **06** (2020) 137.
- [43] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Piacenza, M. Radici, and A. Signori (MAP (Multi-dimensional Analyses of Partonic Distributions) Collaboration), *J. High Energy Phys.* **10** (2022) 127.
- [44] M. Boglione, J. O. Gonzalez-Hernandez, and A. Simonelli, *Phys. Rev. D* **106**, 074024 (2022).
- [45] M. Boglione and A. Simonelli, *J. High Energy Phys.* **09** (2023) 006.
- [46] M. Anselmino, M. Boglione, U. D'Alesio, A. Kotzinian, F. Murgia, A. Prokudin, and C. Turk, *Phys. Rev. D* **75**, 054032 (2007).
- [47] M. Anselmino, M. Boglione, U. D'Alesio, A. Kotzinian, F. Murgia, A. Prokudin, and S. Melis, *Nucl. Phys. B, Proc. Suppl.* **191**, 98 (2009).
- [48] M. Anselmino, M. Boglione, U. D'Alesio, S. Melis, F. Murgia, and A. Prokudin, *Phys. Rev. D* **87**, 094019 (2013).
- [49] M. Anselmino, M. Boglione, U. D'Alesio, J. O. Gonzalez Hernandez, S. Melis, F. Murgia, and A. Prokudin, *Phys. Rev. D* **92**, 114023 (2015).
- [50] Z.-B. Kang, A. Prokudin, P. Sun, and F. Yuan, *Phys. Rev. D* **93**, 014009 (2016).
- [51] H.-W. Lin, W. Melnitchouk, A. Prokudin, N. Sato, and H. Shows, *Phys. Rev. Lett.* **120**, 152502 (2018).
- [52] U. D'Alesio, C. Flore, and A. Prokudin, *Phys. Lett. B* **803**, 135347 (2020).
- [53] J. Cammarota, L. Gamberg, Z.-B. Kang, J. A. Miller, D. Pitonyak, A. Prokudin, T. C. Rogers, and N. Sato (Jefferson Lab Angular Momentum Collaboration), *Phys. Rev. D* **102**, 054002 (2020).
- [54] L. Gamberg, M. Malda, J. A. Miller, D. Pitonyak, A. Prokudin, and N. Sato (Jefferson Lab Angular Momentum (JAM) Collaboration), *Phys. Rev. D* **106**, 034014 (2022).
- [55] A. Metz and A. Vossen, *Prog. Part. Nucl. Phys.* **91**, 136 (2016).
- [56] A. Bianconi, S. Boffi, R. Jakob, and M. Radici, *Phys. Rev. D* **62**, 034008 (2000).
- [57] A. Bacchetta and M. Radici, *Phys. Rev. D* **67**, 094002 (2003).
- [58] J. C. Collins, S. F. Heppelmann, and G. A. Ladinsky, *Nucl. Phys.* **B420**, 565 (1994).
- [59] J. C. Collins and G. A. Ladinsky, *arXiv:hep-ph/9411444*.
- [60] R. L. Jaffe, X.-m. Jin, and J. Tang, *Phys. Rev. Lett.* **80**, 1166 (1998).
- [61] R. L. Jaffe, X.-m. Jin, and J.-a. Tang, *Phys. Rev. D* **57**, 5920 (1998).
- [62] A. Courtoy, A. Bacchetta, M. Radici, and A. Bianconi, *Phys. Rev. D* **85**, 114023 (2012).
- [63] M. Radici, A. Courtoy, A. Bacchetta, and M. Guagnelli, *J. High Energy Phys.* **05** (2015) 123.
- [64] R. Seidl *et al.* (Belle Collaboration), *Phys. Rev. D* **96**, 032005 (2017).
- [65] D. Pitonyak, C. Cocuzza, A. Metz, A. Prokudin, and N. Sato, *Phys. Rev. Lett.* **132**, 011902 (2024).
- [66] A. Bacchetta, A. Courtoy, and M. Radici, *J. High Energy Phys.* **03** (2013) 119.
- [67] G. R. Goldstein, J. O. Gonzalez Hernandez, and S. Liuti, *arXiv:1401.0438*.
- [68] M. Radici and A. Bacchetta, *Phys. Rev. Lett.* **120**, 192001 (2018).
- [69] J. Benel, A. Courtoy, and R. Ferro-Hernandez, *Eur. Phys. J. C* **80**, 465 (2020).
- [70] R. Gupta, Y.-C. Jang, B. Yoon, H.-W. Lin, V. Cirigliano, and T. Bhattacharya, *Phys. Rev. D* **98**, 034503 (2018).
- [71] R. Gupta, B. Yoon, T. Bhattacharya, V. Cirigliano, Y.-C. Jang, and H.-W. Lin, *Phys. Rev. D* **98**, 091501(R) (2018).
- [72] N. Yamanaka, S. Hashimoto, T. Kaneko, and H. Ohki (JLQCD Collaboration), *Phys. Rev. D* **98**, 054516 (2018).
- [73] N. Hasan, J. Green, S. Meinel, M. Engelhardt, S. Krieg, J. Negele, A. Pochinsky, and S. Syritsyn, *Phys. Rev. D* **99**, 114505 (2019).
- [74] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco, *Phys. Rev. D* **102**, 054517 (2020).
- [75] T. Harris, G. von Hippel, P. Jannackar, H. B. Meyer, K. Ottnad, J. Wilhelm, H. Wittig, and L. Wrang, *Phys. Rev. D* **100**, 034513 (2019).

- [76] D. Horkel, Y. Bi, M. Constantinou, T. Draper, J. Liang, K.-F. Liu, Z. Liu, and Y.-B. Yang (χ QCD Collaboration), *Phys. Rev. D* **101**, 094501 (2020).
- [77] C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, and F. Manigrasso, *Phys. Rev. D* **104**, 054503 (2021).
- [78] S. Park, R. Gupta, B. Yoon, S. Mondal, T. Bhattacharya, Y.-C. Jang, B. Joó, and F. Winter (Nucleon Matrix Elements (NME) Collaboration), *Phys. Rev. D* **105**, 054505 (2022).
- [79] R. Tsuji, N. Tsukamoto, Y. Aoki, K.-I. Ishikawa, Y. Kuramashi, S. Sasaki, E. Shintani, and T. Yamazaki (PACS Collaboration), *Phys. Rev. D* **106**, 094505 (2022).
- [80] G. S. Bali, S. Collins, S. Heybrock, M. Löffler, R. Rödl, W. Söldner, and S. Weishäupl, *Phys. Rev. D* **108**, 034512 (2023).
- [81] R. E. Smail *et al.*, *Phys. Rev. D* **108**, 094511 (2023).
- [82] H.-x. He and X.-D. Ji, *Phys. Rev. D* **52**, 2960 (1995).
- [83] V. Barone, T. Calarco, and A. Drago, *Phys. Lett. B* **390**, 287 (1997).
- [84] P. Schweitzer, D. Urbano, M. V. Polyakov, C. Weiss, P. V. Pobylitsa, and K. Goeke, *Phys. Rev. D* **64**, 034013 (2001).
- [85] L. P. Gamberg and G. R. Goldstein, *Phys. Rev. Lett.* **87**, 242001 (2001).
- [86] B. Pasquini, M. Pincetti, and S. Boffi, *Phys. Rev. D* **72**, 094029 (2005).
- [87] M. Wakamatsu, *Phys. Lett. B* **653**, 398 (2007).
- [88] C. Lorce, *Phys. Rev. D* **79**, 074027 (2009).
- [89] N. Yamanaka, T. M. Doi, S. Imai, and H. Suganuma, *Phys. Rev. D* **88**, 074036 (2013).
- [90] M. Pitschmann, C.-Y. Seng, C. D. Roberts, and S. M. Schmidt, *Phys. Rev. D* **91**, 074004 (2015).
- [91] S.-S. Xu, C. Chen, I. C. Cloet, C. D. Roberts, J. Segovia, and H.-S. Zong, *Phys. Rev. D* **92**, 114034 (2015).
- [92] Q.-W. Wang, S.-X. Qin, C. D. Roberts, and S. M. Schmidt, *Phys. Rev. D* **98**, 054019 (2018).
- [93] L. Liu, L. Chang, and Y.-X. Liu, *Phys. Rev. D* **99**, 074013 (2019).
- [94] A. Metz and D. Pitonyak, *Phys. Lett. B* **723**, 365 (2013); **762**, 549(E) (2016).
- [95] K. Kanazawa, Y. Koike, A. Metz, and D. Pitonyak, *Phys. Rev. D* **89**, 111501(R) (2014).
- [96] L. Gamberg, Z.-B. Kang, D. Pitonyak, and A. Prokudin, *Phys. Lett. B* **770**, 242 (2017).
- [97] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Lett. B* **780**, 332 (2018).
- [98] Y. V. Kovchegov and M. D. Sievert, *Phys. Rev. D* **99**, 054033 (2019).
- [99] J. Soffer, *Phys. Rev. Lett.* **74**, 1292 (1995).
- [100] P. Herczeg, *Prog. Part. Nucl. Phys.* **46**, 413 (2001).
- [101] J. Erler and M. J. Ramsey-Musolf, *Prog. Part. Nucl. Phys.* **54**, 351 (2005).
- [102] N. Severijns, M. Beck, and O. Naviliat-Cuncic, *Rev. Mod. Phys.* **78**, 991 (2006).
- [103] V. Cirigliano, S. Gardner, and B. Holstein, *Prog. Part. Nucl. Phys.* **71**, 93 (2013).
- [104] A. Courtoy, S. Baeßler, M. González-Alonso, and S. Liuti, *Phys. Rev. Lett.* **115**, 162001 (2015).
- [105] M. González-Alonso, O. Naviliat-Cuncic, and N. Severijns, *Prog. Part. Nucl. Phys.* **104**, 165 (2019).
- [106] M. Pospelov and A. Ritz, *Ann. Phys. (Amsterdam)* **318**, 119 (2005).
- [107] N. Yamanaka, B. K. Sahoo, N. Yoshinaga, T. Sato, K. Asahi, and B. P. Das, *Eur. Phys. J. A* **53**, 54 (2017).
- [108] T. Liu, Z. Zhao, and H. Gao, *Phys. Rev. D* **97**, 074018 (2018).
- [109] C. Cocuzza, A. Metz, D. Pitonyak, A. Prokudin, N. Sato, and R. Seidl (Jefferson Lab Angular Momentum (JAM) Collaboration), companion Letter, *Phys. Rev. Lett.* **132**, 091901 (2024).
- [110] T. Sjostrand, L. Lonnblad, and S. Mrenna, *arXiv:hep-ph/0108264*.
- [111] X. Artru and J. C. Collins, *Z. Phys. C* **69**, 277 (1996).
- [112] A. Vossen *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **107**, 072004 (2011).
- [113] D. Boer, R. Jakob, and M. Radici, *Phys. Rev. D* **67**, 094003 (2003); **98**, 039902(E) (2018).
- [114] A. Bacchetta, F. A. Ceccopieri, A. Mukherjee, and M. Radici, *Phys. Rev. D* **79**, 034029 (2009).
- [115] H. H. Matevosyan, A. Bacchetta, D. Boer, A. Courtoy, A. Kotzinian, M. Radici, and A. W. Thomas, *Phys. Rev. D* **97**, 074019 (2018).
- [116] H. H. Matevosyan, A. Kotzinian, and A. W. Thomas, *Phys. Rev. D* **97**, 014019 (2018).
- [117] The COMPASS Collaboration, *Phys. Lett. B* **845**, 138155 (2023).
- [118] HERMES Collaboration, *J. High Energy Phys.* **06** (2008) 017.
- [119] A. Bacchetta and M. Radici, *Phys. Rev. D* **69**, 074026 (2004).
- [120] A. Bacchetta, A. Courtoy, and M. Radici, *Phys. Rev. Lett.* **107**, 012001 (2011).
- [121] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **115**, 242501 (2015).
- [122] M. Radici, A. M. Ricci, A. Bacchetta, and A. Mukherjee, *Phys. Rev. D* **94**, 034012 (2016).
- [123] A. Bacchetta and M. Radici, *Phys. Rev. D* **70**, 094032 (2004).
- [124] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 438 (1972).
- [125] Y. L. Dokshitzer, *Sov. Phys. JETP* **46**, 641 (1977).
- [126] G. Altarelli and G. Parisi, *Nucl. Phys.* **B126**, 298 (1977).
- [127] F. Baldracchini, N. S. Craigie, V. Roberto, and M. Socolovsky, *Fortschr. Phys.* **30**, 505 (1981).
- [128] X. Artru and M. Mekhfi, *Z. Phys. C* **45**, 669 (1990).
- [129] J. Blumlein, *Eur. Phys. J. C* **20**, 683 (2001).
- [130] M. Stratmann and W. Vogelsang, *Phys. Rev. D* **65**, 057502 (2002).
- [131] F. A. Ceccopieri, M. Radici, and A. Bacchetta, *Phys. Lett. B* **650**, 81 (2007).
- [132] U. P. Sukhatme, K. E. Lassila, and R. Orava, *Phys. Rev. D* **25**, 2975 (1982).
- [133] D. de Florian and L. Vanni, *Phys. Lett. B* **578**, 139 (2004).
- [134] A. Majumder and X.-N. Wang, *Phys. Rev. D* **72**, 034007 (2005).
- [135] A. Majumder and X.-N. Wang, *Phys. Rev. D* **70**, 014007 (2004).
- [136] H. Chen, M. Jaarsma, Y. Li, I. Moul, W. J. Waalewijn, and H. X. Zhu, *J. High Energy Phys.* **07** (2023) 185.

- [137] H. Chen, M. Jaarsma, Y. Li, I. Moulton, W. J. Waalewijn, and H. X. Zhu, [arXiv:2210.10061](#).
- [138] M. Tanabashi *et al.* (Particle Data Group), *Phys. Rev. D* **98**, 030001 (2018).
- [139] R. D. Ball, S. Forte, A. Guffanti, E. R. Nocera, G. Ridolfi, and J. Rojo (NNPDF Collaboration), *Nucl. Phys.* **B874**, 36 (2013).
- [140] G. Altarelli, S. Forte, and G. Ridolfi, *Nucl. Phys.* **B534**, 277 (1998).
- [141] S. Forte, G. Altarelli, and G. Ridolfi, *Nucl. Phys. B, Proc. Suppl.* **74**, 138 (1999).
- [142] P. V. Pobylitsa, [arXiv:hep-ph/0301236](#).
- [143] Y. V. Kovchegov (private communication).
- [144] M. Stratmann and W. Vogelsang, *Phys. Rev. D* **64**, 114007 (2001).
- [145] T. Sjostrand, L. Lonnblad, S. Mrenna, and P. Z. Skands, [arXiv:hep-ph/0308153](#).
- [146] C. Bierlich *et al.*, *SciPost Phys. Codeb.* **2022**, 8 (2022).
- [147] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, *Phys. Lett.* **12**, 57 (1964).
- [148] B. Pokhrel and B. Surrow (private communication).
- [149] In order to align with the methodology of JAMDiFF, we show here results from the JAM3D analysis that are slightly updated from Ref. [54]: antiquark transversity PDFs are now included (with $h_1^d = -h_1^u$), the small- x constraint Eq. (23) is imposed, and, for the fit with LQCD, δu and δd from ETMC and PNDME are used (instead of only the g_T data point from ETMC).
- [150] R. D. Ball *et al.* (NNPDF Collaboration), *Eur. Phys. J. C* **82**, 428 (2022).
- [151] D. S. Sivia, *Proceedings for Maximum Entropy and Bayesian Methods!* (NMB Printers, Port Elizabeth, South Africa, 1996), pp. 131–137.
- [152] C. Cocuzza, A. Metz, D. Pitonyak, A. Prokudin, N. Sato, and R. Seidl, Github repository, https://github.com/prokudin/JAMDiFF_library (2023).
- [153] C. Cocuzza, A. Metz, D. Pitonyak, A. Prokudin, N. Sato, and R. Seidl, Google colab notebook, https://colab.research.google.com/github/prokudin/JAMDiFF_library/blob/main/JAMDiFF_Library.ipynb (2023).
- [154] L. Gamberg, Z.-B. Kang, D. Pitonyak, A. Prokudin, N. Sato, and R. Seidl, *Phys. Lett. B* **816**, 136255 (2021).
- [155] R. Abdul Khalek *et al.*, *Nucl. Phys.* **A1026**, 122447 (2022).
- [156] C. Egerer *et al.* (HadStruc Collaboration), *Phys. Rev. D* **105**, 034507 (2022).