Collinearity criteria for transverse momentum dependent distributions in semi-inclusive deep-inelastic scattering

Mason Albright[a], Scott Dolan[a], Leonard Gamberg[d], W. Melnitchouk[e],
Daniel Pitonyak[d], Alexei Prokudin[d,e], Nobuo Sato[e], Zackary Scalyer[f]

[a]*College of Engineering, Penn State University, State College, Pennsylvania 16801, USA*
[b]*Dipartimento di Fisica Teorica, Università di Torino, Via P. Giuria 1, I-10125 Torino, Italy*
[c]*Dipartimento di Fisica, Università di Cagliari, Cittadella Universitaria, I-09042 Monserrato (CA), Italy*
[d]*Division of Science, Penn State University Berks, Reading, Pennsylvania 19610, USA*
[e]*Theory Center, Jefferson Lab, 12000 Jefferson Avenue, Newport News, Virginia 23606, USA*
[f]*Department of Mathematics and Statistics, Villanova University, Villanova, PA 19085, USA*

Abstract

We discuss the impact of data selection on the determination of nonperturbative transverse momentum dependence in semi-inclusive deep-inelastic scattering. In particular, we implement the recently introduced collinearity criteria that allow selection of data predominantly in the current fragmentation region, and apply this framework to pion and kaon multiplicity data from HERMES. We compare the resulting unpolarized transverse momentum dependent distributions with previous extractions, and discuss the potential impact of the selection criteria on future experiments.

Keywords: unpolarized TMDs, semi-inclusive deep-inelastic scattering, perturbative QCD
PACS: 12.38.-t, 12.38.Bx, 13.85Fb, 13.85.Ni, 13.87Fh

1. Introduction

Understanding the internal structure of hadrons has been a topic of intense research for over 50 years. From inclusive and semi-inclusive deep inelastic scattering (DIS) experiments, we know that hadrons have a complex internal structure composed of quarks, anti-quarks, and gluons (partons). In addition to the partons' collinear momentum, which is highly correlated with the direction of a fast-moving parent hadron, partons also possess intrinsic transverse motion and structure. Several types of high energy scattering experiments are sensitive to this intrinsic momentum structure: semi-inclusive deep-inelastic scattering (SIDIS) ($e\,N \rightarrow e'\,h\,X$) [2], electron-positron annihilation to almost back-to-back hadrons ($e^+e^- \rightarrow h_a\,h_b\,X$) [3], and Drell-Yan ($p\,p \rightarrow \ell^+\,\ell^-\,X$) or weak gauge boson production ($p\,p \rightarrow \{Z, W^+, W^-\}\,X$) [4]. To connect these measurements to a theoretical framework, one relies on QCD factorization theorems, where TMD factorization [5–7] describes these processes in terms of a collinear perturbative (hard) scattering cross section and nonperturbative transverse momentum dependent (TMD) parton distribution functions (PDFs) and fragmentation functions (FFs) (collectively called TMDs) [2, 8, 9].

A condition implicit in the proofs of TMD factorization in SIDIS, where final state hadrons are fragments of the struck quark, is the assumption of a clear separation in momentum of the struck quark from the target spectators. In this framework, the fragmentation of the quark (what about gluons?) into hadrons is assumed to be independent of the production of the quark [10, 11]. That is, fragmentation is described by a function of the momentum fraction z of the quark carried by the produced hadron, independent of the momentum fraction x of the parent hadron carried by the struck quark. By contrast, if the produced hadron moves in nearly the same direction as the target then the hadron is said to be in the target fragmentation region and the relevant factorization theorem uses fracture functions [11–13]. A clear distinction between the current and target regions requires large enough separation in the momentum of the current and target fragments, and it is convenient to use rapidity to delineate these regions. Some time ago

Berger [10, 14] provided a specific rapidity gap criterion to study the dynamics of quark fragmentation in the current region [10, 14]. In practice, however, the classification of distinct current, target, and central fragmentation regions is not always sharp [1, 10, 14–16].

Recently, it has been shown [1] that at moderate to small values of momentum transfer Q, estimating the adequacy of the current fragmentation criteria requires knowledge of intrinsic, nonperturbative properties of partons. This conclusion was reached after a close examination of the errors involved in TMD factorization, resulting in a more restrictive region where current fragmentation by itself is valid [1]. Additionally, it was pointed out that the applicability of TMD factorization in the current region not only depends on relatively distinct target and current rapidity regions, but also on the transverse momentum of the current hadron, P_{hT}. Indeed, recent work shows that the standard cuts [17, 29] applied in analysis of SIDIS experiments are insufficient to guarantee that SIDIS data are uniquely in the current region. While the standard cuts might suggest that certain kinematic bins are in the current regime, in practice the bins may actually consist of an admixture of current, central, and target fragmentation regions [1]. In principle, this mixing of hadrons produced from the target, central and current region affects the extraction of TMDs and interpretation of the fitted functions. Therefore, it is important to analyze the role that data cuts play in discriminating current from target and central fragmentation regions and the impact on the extraction of TMDs from future SIDIS data.

In this paper, we implement for the first time the collinearity criteria [1] in an extraction of TMD widths for unpolarized PDFs and FFs. In order to avoid issues with scale evolution, we focus only on HERMES multiplicity data in SIDIS. We use a simple analytical approximation for the solutions of the TMD evolution equations that is valid in the nonperturbative region, and extract the transverse momentum dependence of TMDs and the flavor dependence of their widths. We organize the discussion as follows: In Sec. 2 we summarize the theoretical formalism needed to analyze the data, including our model for the transverse momentum dependence. Next, in Sec. 3 we review the collinearity criteria used to isolate the current fragmentation region. We then provide details on our data selection and perform our fit of the HERMES multiplicities to extract the relevant TMD widths. We also compare our results with previous analyses and provide an interpretation of our results in the context of other extractions. Finally, in Sec. 4 we summarize our findings and discuss the impact of this analysis for ongoing and future experiments.

2. Theoretical Formalism

Consider the process of SIDIS from an unpolarized nucleon and deuteron target $N = p, n, d$,

$$\ell(l) + p(N) \rightarrow \ell'(l') + h(P_h) + X, \tag{1}$$

where the momenta of the particles are l, p and P_h for the lepton, target and produced hadron respectively. The standard Lorentz invariants for this reaction are

$$S = (l + P)^2, \qquad Q^2 = -q^2 = -(l - l')^2, \qquad x_B = \frac{Q^2}{2P \cdot q}, \qquad y = \frac{Q^2}{x_B S}, \qquad z_h = \frac{P \cdot P_h}{P \cdot q}. \tag{2}$$

One often uses a vector $\vec{q}_T \simeq -\vec{P}_{hT}/z_h$ since the power counting applied in proofs of TMD factorization [7] is carried out for the region, $q_T \ll Q$. For the situation where $q_T \ll Q$, one finds at leading-order [18]

$$\frac{d\sigma}{dx_B \, dQ^2 \, dz_h \, dP_{hT}^2} = \frac{2\pi^2 \alpha_{em}^2}{(x_B S)^2} \frac{1 + (1 - y)^2}{y^2} F_{UU}(x_B, z_h, P_{hT}^2, Q^2) + O(q_T/Q), \tag{3}$$

where TMD factorization yields for the unpolarized structure function[7],

$$F_{UU}(x, z, P_{hT}^2, Q^2) = \mathcal{H}(Q; \mu_Q) \sum_a e_a^2 \int d^2\vec{k}_\perp \, d^2\vec{p}_\perp \, \delta^{(2)}\left(\vec{P}_{hT} - z\vec{k}_\perp - \vec{p}_\perp\right) f_{a/N}(x, k_\perp; Q^2, \mu_Q) D_{h/a}(z, p_\perp; Q^2, \mu_Q). \tag{4}$$

The functions $f_{a/N}(x, k_\perp; Q^2, \mu_Q)$ and $D_{h/a}(z, p_\perp; Q^2, \mu_Q)$ are the unpolarized TMDs, for a parton of flavor a, on a target N, and a produced hadron h [7, 18]. The function $\mathcal{H}$ is the hard function, which at leading order in the strong

coupling is normalized as, $\mathcal{H} = 1$. In the γ^*-p center-of-mass frame, we have the following relations up to $O(P_{hT}/Q)$ corrections [18]

$$x \equiv \frac{k^+}{P^+} \approx x_B, \qquad z \equiv \frac{P_h^-}{p^-} \approx z_h. \tag{5}$$

Apart from kinematical variables and transverse momentum, the TMDs $f_{a/p}(x, k_\perp; \zeta_N, \mu_Q)$ and $D_{h/a}(z, p_\perp; \zeta_h, \mu_Q)$, depend on the scale of the reaction Q^2, through the renormalization scale, μ_Q, and rapidity scales ζ_N and ζ_h as a result of jointly solving the Collins-Soper (CS) in b_T space and renormalization group equations for the TMDs. The rapidity parameters can be set to $\zeta_N = \zeta_h = Q^2$ [7, 19].

In this analysis we study the current fragmentation region employing an analytical form of TMDs in the momentum space that enables us to speed up our numerical calculations by using an estimate for the solution of CS equations that is valid in nonperturbative region. Our analysis begins with the solution of CS equations given in Ref. [19] at a given fixed scale, $Q_0 = \mu_0$,

$$\tilde{f}_{a/N}(x, b_T; Q^2, \mu_Q) = \tilde{f}_{a/N}(x, b_T; Q_0^2, \mu_0)\, e^{-S(b_T, Q, \mu_Q, Q_0, \mu_0)/2}, \tag{6}$$

$$\tilde{D}_{h/a}(z, b_T; Q^2, \mu_Q) = \frac{1}{z^2} \tilde{D}_{h/a}(z, b_T; Q_0^2, \mu_0)\, e^{-S(b_T, Q, \mu_Q, Q_0, \mu_0)/2}, \tag{7}$$

where b_T is Fourier conjugate variable to $k_\perp$. Q_0 and μ_0 are chosen fixed reference scales (we will choose $\mu_0 = Q_0$). The evolution factor $S(b_T, Q, Q_0, \mu_0)$ is,

$$S(b_T, Q, \mu_Q, Q_0, \mu_0) = \tilde{K}(b_T, \mu_0) \ln \frac{Q^2}{Q_0^2} - \int_{\mu_0}^{\mu_Q} \frac{d\mu'}{\mu'} \left[-2\gamma_i(\alpha_s(\mu'); 1) + \ln \frac{Q^2}{\mu'^2} \gamma_K(\alpha_s(\mu')), \right] \tag{8}$$

where $\tilde{K}$ is the Collins-Soper evolution kernel, and γ_i and γ_K are anomalous dimensions of the TMD. They can be calculated perturbatively provided α_s is small enough. We will choose $\mu_Q = Q$.

The next step is to choose b_T parameterization of TMDs at the initial scale Q_0. Since low Q^2 data is known [20] to have an approximate Gaussian dependence, we will follow a very simple analytical form used in Refs [17, 21], and choose a Gaussian ansatz for b_T dependence (note that we also allow for a flavor dependence in the widths of the Gaussians):

$$\tilde{f}_{a/N}(x, b_T; Q_0^2, \mu_0) \simeq \tilde{f}_{a/N}(x; Q_0^2, \mu_0) e^{-b_T^2 \frac{\langle k_\perp^2 \rangle_a}{4}},$$

$$\tilde{D}_{h/a}(z, b_T; Q_0^2, \mu_0) \simeq \frac{1}{z^2} \tilde{D}_{h/a}(z; Q_0^2, \mu_0) e^{-b_T^2 \frac{\langle p_\perp^2 \rangle_a}{4z^2}}. \tag{9}$$

The parameterizations in Eqs. (15) are chosen in such a way that we can make a connection of $\tilde{f}_{a/N}(x; Q_0^2, \mu_0)$ and $\tilde{D}_{h/a}(z; Q_0^2, \mu_0)$ to the usual collinear PDF and FF functions. Indeed, using result of Ref. [22] we know that at leading order in α_s, the collinear integrated cross-section corresponds to the integral in P_{hT} of the TMD approximated cross-section. Using $S(b_T, Q_0, \mu_0, Q_0, \mu_0) = 0$ we obtain:

$$\tilde{f}_{a/N}(x; Q_0^2, \mu_0) = f_{a/N}(x, Q_0), \tag{10}$$

$$\tilde{D}_{h/a}(z; Q_0^2, \mu_0) = D_{h/a}(z, Q_0), \tag{11}$$

where $f_{a/p}(x, Q_0)$ and $D_{h/a}(z, Q_0)$ are the standard unpolarized collinear PDF and FF, respectively, for a parton of flavor a, at the initial scale.

Here it is important to note, that using a typical value of $\langle k_\perp^2 \rangle = 0.25$ (GeV^2) [23] for the width of TMDs at the initial scale Q_0, the dominant support in b_T space for the FT TMDs is at large values of $b_T \sim 3$ to 4 (GeV^{-1}) which correspond to ~ 1 fm of separation of quark fields. These distances are associated with nonperturbative QCD dynamics which we depict in Fig. 1 (see [26]).

3

2.1. Approximate Solution to the CS Equation in the nonperturbative regime

In this sub-section, we outline our approximate solution to the CS evolution kernel in the nonperturbative regime. The Collins-Soper evolution kernel obeys the following renormalization-group equation,

$$\frac{d\,\tilde{K}(b_T;\mu)}{d\ln\mu} = -\gamma_K(\alpha_s(\mu))\,, \tag{12}$$

where we use notations of Ref. [7]. Using the one loop result for γ_K for from Ref. [24], we have

$$\gamma_K(\alpha_s(\mu)) = 2C_F\frac{\alpha_s(\mu)}{\pi}\,, \tag{13}$$

$$\tilde{K}(b_T;\mu) = -\frac{\alpha_s(\mu)}{\pi}\left[\ln\frac{b_T^2\mu^2}{4} + 2\gamma_E\right] + O(\alpha_s(\mu)^2), \tag{14}$$

this result is applicable for perturbative distances and scales; for instance $\mu > 1$ (GeV) and at low $b_T < 1\,\text{GeV}^{-1} = 0.2$ fm. As we noted in Fig. 1, the distances where we need to calculate $\tilde{K}(b_T;\mu)$ are much larger than the perturbative distance scales, and thus the fixed-order approximation is not applicable anymore. Numerically it happens due to the growth of $\ln\left(b_T^2\mu^2/4\right)$, so that neglected higher orders become comparable to the one loop order. One of the methods to circumvent this problem is to introduce an intermediate scale $\mu_b \sim 1/b_T$ such that all logs are minimized, the Landau pole however will manifest itself in the growth and divergence of $\alpha_s(\mu)$. In order to deal with the Landau pole (or nonperturbative regime) one can introduce a b_* prescription [7] that will limit the region of b_T to a maximum value of $b_{max} \sim 1\,\text{GeV}^{-1}$ and introduce the non perturbative Collins-Soper kernel $g_K(b_T;b_{max}) \equiv -\tilde{K}(b_T;\mu_{b_*}) + \tilde{K}(b_*;\mu_{b_*})$. A great deal of successful phenomenology [CITE] was done using this procedure.
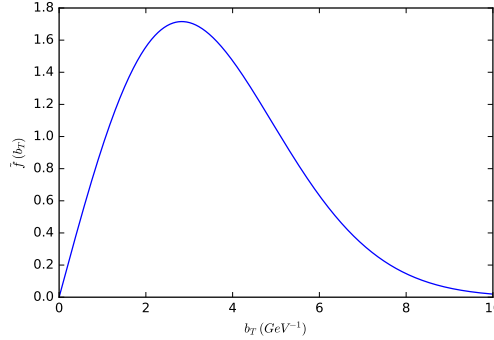

Figure 1: [Color online] Typical shape of non perturbative TMDs in Eq. (15).

In our particular situation, however, the functional form of Fourier transform (FT) TMDs are such that *only* large values of b_T dominate. Thus we *approximate* $K(b_T,\mu_0)$ by its large b_T asymptote which is approximately constant, and negative, as suggested in Ref. [19]. We will call this constant $g_{K_0} \equiv \lim_{b_T\to\infty}\tilde{K}(b_T,\mu_0)$, such that in this approximation our FT TMDs take the form,

$$\tilde{f}_{a/N}(x,b_T;Q^2,\mu_Q) \simeq \tilde{f}_{a/N}(x;Q_0)e^{-b_T^2\frac{\langle k_\perp^2\rangle a}{4}}\left(\frac{Q}{Q_0}\right)^{g_{K_0}}e^{-S_{pert}/2}\,,$$

$$\tilde{D}_{h/a}(z,b_T;Q^2,\mu_Q) \simeq \frac{1}{z^2}D_{h/a}(z;Q_0)e^{-b_T^2\frac{\langle p_\perp^2\rangle a}{4z^2}}\left(\frac{Q}{Q_0}\right)^{g_{K_0}}e^{-S_{pert}/2}\,, \tag{15}$$

where the perturbative part of the CSS exponent is

$$S_{pert} \equiv \int_{\mu_0}^{\mu_Q}\frac{d\mu'}{\mu'}\left[-2\gamma_i(\alpha_s(\mu');1) + \ln\frac{Q^2}{\mu'^2}\gamma_K(\alpha_s(\mu'))\right]\,. \tag{16}$$

4

Using the one loop results [24] for γ_i from Eq (14) and γ_K we have,

$$\gamma_i(\alpha_s(\mu); 1) = \frac{3C_F}{2} \frac{\alpha_s(\mu)}{\pi} \, , \tag{17}$$

and thus Eq. (16) is

$$S_{pert} = 2C_F \int_{\mu_0}^{\mu_Q} \frac{d\mu'}{\mu'} \frac{\alpha_s(\mu')}{\pi} \left[\ln \frac{Q^2}{\mu'^2} - \frac{3}{2} \right] \, , \tag{18}$$

where $C_F = 4/3$. In order to calculate analytically the integral from Eq. (18) we use the one loop result for the strong coupling constant,

$$\alpha_s(\mu') = \frac{1}{\beta_0 \ln Q^2/\Lambda^2} \, , \qquad \beta_0 = \frac{33 - 2n_f}{12\pi} \, , \tag{19}$$

where we will use $n_f = 3$ as we are interested in the region of low Q; and $\Lambda = 0.25$ (GeV), such that $\alpha_s(M_Z) = 0.118$ [25]. We obtain, see also Ref. [26],

$$S_{pert} = -\frac{2C_F}{\pi\beta_0} \left[\ln(Q/Q_0) - \ln(Q/\Lambda) \ln\left(\frac{\ln(Q/\Lambda)}{\ln(Q_0/\Lambda)} \right) + \frac{3}{4} \ln\left(\frac{\ln(Q/\Lambda)}{\ln(Q_0/\Lambda)} \right) \right] . \tag{20}$$

Our result suggests that in the region of applicability, TMD evolution amounts to a shift in the normalization in Q^2, and does not lead to the usual broadening of TMDs observed [7] at large Q^2. This result is consistent with the experimental findings [27] of the HERMES collaboration, where no significant broadening of multiplicities was observed as a function of Q^2. Notice that our approximation of $K(b_T, \mu_0)$ is quite crude and we do not expect the result to hold generally. In particular, it will fail at large Q^2 where low values of b_T dominate in the FT. As we stated earlier, this simplification is taken deliberately to be able to speed up the numerical computations.

Finally we arrive at the following parameterization of the unpolarized structure function from Eq. (4)

$$F_{UU}(x_B, z_h, P_{hT}^2, Q^2) = \sum_a e_a^2 f_{a/N}(x_B, Q_0) D_{h/a}(z_h, Q_0) \left(\frac{Q^2}{Q_0^2} \right)^{g_{K_0}} e^{-S_{pert}} \frac{e^{-P_{hT}^2/\langle P_{hT}^2 \rangle_a}}{\pi \langle P_{hT}^2 \rangle_a} \, , \tag{21}$$

where

$$\langle P_{hT}^2 \rangle_a = \langle p_\perp^2 \rangle_{a/N} + z_h^2 \langle k_\perp^2 \rangle_{h/a} \, , \tag{22}$$

results from carrying out the FT to momentum space. Notice that our result is similar to Refs. [17, 21], with the following important differences: We do not utilize DGLAP evolution for TMDs as was done in Ref. [17] and we do not freeze the scale as was done in Ref. [21]. We mention that a fit of HERMES [27] and COMPASS [28] multiplicities with full TMD evolution and next-to-leading logarithmic accuracy was published in Ref. [29]. In this study we do not aim at a complete fit of the data at different Q^2 and choose to fit only HERMES [27] multiplicities. Our approximate result, Eq. (21) is consistent with CS equations, provided that the scale Q is not drastically different from Q_0 and TMDs are dominated by the region of large b_T, so that $\tilde{K}(b_T, \mu_0)$ is approximately constant.

From this, we find our formula for the HERMES multiplicities [27]

$$M_N^h(x_B, Q^2, z_h, P_{hT}) = 2P_{hT} \frac{\pi \sum_a e_a^2 f_{a/N}(x_B, Q_0) D_{h/a}(z_h, Q_0)}{\sum_a e_a^2 f_{a/N}(x_B, Q^2)} \left(\frac{Q^2}{Q_0^2} \right)^{g_{K_0}} e^{-S_{pert}} \frac{e^{-P_{hT}^2/\langle P_{hT}^2 \rangle_a}}{\pi \langle P_{hT}^2 \rangle_a} \, . \tag{23}$$

3. Phenomenological Analysis and Discussion

3.1. Data Selection

Prior to performing fits of experimental measurements, one often needs to make cuts on the data to ensure that the fit is restricted only to the data that can be described by the theoretical framework. For example, the so called standard cut on the HERMES multiplicity data is [17]

$$z < 0.6 \qquad Q^2 > 1.69 \text{ GeV}^2 \qquad 0.2 < P_{hT} < 0.9 \text{ GeV} \qquad \text{(Torino cut)} . \tag{24}$$

However, as was argued in Ref. [1], Eq. (24) is not sufficient to guarantee that one stays strictly in the current fragmentation region. To fully understand the conditions under which the outgoing hadron is in the current fragmentation region, we need to know how accurately the factorization theorem holds, as a function of the kinematic variables. Errors in factorization correspond to the deviations of momenta from their limiting cases, notably of collinear momenta from their exactly collinear configurations. Generically these errors are suppressed by a power of m^2/Q^2 by taking the limit of $Q \to \infty$ with x_B and z_h fixed.[1] For collinear momenta, there are three parts to the relevant deviations. One concerns the initial and final quark momenta k_i and k_f (see further discussion on their estimates below). Related to this is that the target remnant should also be a momentum collinear to the target. A third component to the error in the fragmentation picture arises from the deviation of P_h from the exact collinear direction for the outgoing quark. All of these errors can be expressed and quantified in terms of rapidities. Thus, while data is normally presented with Q, x_B, z_h and $\vec{P}_{hT}$ being used as the independent variables, in our analysis we will also relate these variables to the set, Q, x_n, y_h and $\vec{P}_{hT}$.

Here it also proves convenient to carry out this analysis in the Breit frame. So, we make a Lorentz transformation to the Breit-frame of the γ^*-p system, where the virtual photon moves along the $-z$ direction, and in the collinear limit, the struck quark's 3-momentum is reversed in a hard collision. Further it is important to distinguish the Nachtmann variable Eq. (5) defined as $x_n = -q^+/P^+$, when the proton mass in not neglected; it approaches x_B as $M_p/Q \to 0$, and is related to x_B by

$$x_n \equiv \frac{2x_B}{1 + \sqrt{1 + 4x_B^2 M_p^2/Q^2}}, \qquad x_B = \frac{x_n}{1 - x_n^2 M_p^2/Q^2}. \tag{25}$$

Now, in light-front coordinates in the Breit frame, the virtual photon, proton, and produced hadron are given by,

$$q = \left(-x_n P^+, \frac{Q^2}{2x_n P^+}, \vec{0}_T\right) = \left(-\frac{Q}{\sqrt{2}}, \frac{Q}{\sqrt{2}}, \vec{0}_T\right), \quad P = \left(P^+, \frac{M_p^2}{2P^+}, \vec{0}_T\right) = \left(\frac{Q}{x_n \sqrt{2}}, \frac{x_n M_p^2}{Q \sqrt{2}}, \vec{0}_T\right) \tag{26}$$

$$P_h = \left(\frac{M_{hT}}{\sqrt{2}} e^{y_h}, \frac{M_{hT}}{\sqrt{2}} e^{-y_h}, \vec{P}_{hT}\right), \tag{27}$$

where $M_{hT} \equiv \sqrt{P_{hT}^2 + M_h^2}$ is the transverse mass, and y_h is the rapidity of the observed hadron, $y_h \equiv \frac{1}{2} \log(P_h^+/P_h^-)$.

In terms of y_h and P_{hT}, z_h is given by

$$z_h = \frac{M_{hT}}{Q} \left(1 - x_n^2 \frac{M_p^2}{Q^2}\right)^{-1} \left(e^{-y_h} + x_n^2 \frac{M_p^2}{Q^2} e^{y_h}\right), \tag{28}$$

where y_p is the proton rapidity, is $y_p = \ln\left(Q/x_n M_p\right)$. It is now straightforward to show that the hadron's rapidity is as a function of z_h and P_{hT} in the current fragmentation region,

$$y_h = \ln\left[\frac{Qz_h\left(Q^2 - x_n^2 M_p^2\right)}{2x_n^2 M_p^2 M_{hT}} - \frac{Q}{x_n M_p} \sqrt{\frac{z_h^2\left(Q^2 - x_n^2 M_p^2\right)^2}{4x_n^2 M_p^2 M_{hT}^2} - 1}\right], \tag{29}$$

with M_P (M_h) the mass of the proton (hadron).

In addition to the variables specifying observed hadrons, the partonic momenta for the struck and fragmenting quark momenta $k \equiv k_i$ and $p \equiv k_f$ light-cone momenta are parameterized as

$$k_i = \left(\frac{M_{iT}}{\sqrt{2}} e^{y_i}, -\frac{M_{iT}}{\sqrt{2}} e^{-y_i}, \vec{k}_T\right), \quad k_f = \left(\frac{M_{fT}}{\sqrt{2}} e^{y_f}, \frac{M_{fT}}{\sqrt{2}} e^{-y_f}, \vec{k}_T\right), \tag{30}$$

where $M_{(i/f)T}$ are the transverse masses of the quarks. The typical values of these quantities are crucial ingredients for an analysis of the errors in factorization formulas and hence for determining a characterization of the current

---

[1] In the discussion of power counting and TMD factorization m is represents a typical hadronic scale, Λ_{QCD}.

fragmentation region. The transverse masses depend on nonperturbative parameters such as k_T and the jet and remnant masses. Since such constraints are not well known, a conservative range of values was motivated by models and event generators that are fit to data (for further details see [1]). Additionally, since the parton-model approximation sets $k^+ = -q^+$ and $k_f^- = q^-$, the quark rapidities are approximately,

$$y_i = \ln \frac{Q}{M_{iT}}, \quad y_f = -\ln \frac{Q}{M_{fT}}, \tag{31}$$

which should be large (positive and negative, respectively) for TMD factorization to hold true [1, 7]. Thus, to quantify to what extent the final state hadron is produced in the current region one knows from power counting in TMD factorization [2], we have $P_h \cdot k_f \ll P_h \cdot k_i$.

Thus, the authors [1] identified the collinearity of the produced hadron in terms of the Lorentz invariant ratio $R(y_h, z_h, x_B, Q)$, (the so-called the collinearity) defined as

$$R(y_h, z_h, x_B, Q) \equiv \frac{P_h \cdot k_f}{P_h \cdot k_i}, \tag{32}$$

and from generic power counting it is clear that the current region is defined by

$$R(y_h, z_h, x_B, Q) \ll 1 : \text{ collinear to outgoing quark}. \tag{33}$$

Using Eqs. (27,30) and taking the average over the azimuthal angle of $\vec{k}_T$, we can express Eqs. (32) as,

$$R = \frac{M_{fT}}{M_{iT}} \frac{e^{y_f - y_h} + e^{y_h - y_f}}{e^{y_i - y_h} - e^{y_h - y_i}}, \tag{34}$$

However due to the dependence on the unobserved initial and final quark rapidities, y_i and y_f and their uncertainty in terms of their transverse masses M_{iT}, M_{fT}, using the collinearity as a discriminator suffers from some ambiguity.

The typical values of these quantities are crucial ingredients for an analysis of the errors in factorization formulas and hence for determining a characterization of the current fragmentation region. The transverse masses depend on nonperturbative parameters such as k_T and the jet and remnant masses. Since such constraints are not well known, a conservative range of values was motivated by models and event generators that are fit to data (for further details see [1]), where it was determined that the transverse masses that span roughly this range of values $M_{iT}^2 = M_{fT}^2 = 0.5 \pm 0.3 \text{ GeV}^2$

On the other hand, the (positive) rapidity of the target proton y_p and the (negative) rapidity of the produced pion y_h can be calculated directly from the kinematics of the process. Under a boost in the z direction, rapidity transforms additively [7]: $y \to y' = y + \psi$, this implies that the difference in rapidity is boost independent. We could then use $y_p - y_h$ as a boost independent measure of the total rapidity interval between the proton and the produced hadron.

In our following analysis, we will use a mix of the rapidity difference and R to analyze the HERMES data.

3.2. Filtering HERMES Data

The HERMES multiplicity data set [27] consists of 2660 points for $\pi^\pm$ and $K^\pm$ produced off hydrogen or deuterium targets and measured in 8 bins in the following kinematical region: $0.037 < x_B < 0.41$, $0.13 < z_h < 0.95$, $1.25 \text{ GeV}^2 < Q^2 < 9.22 \text{ GeV}^2$, and $0.06 \text{ GeV} < P_{hT} < 1.36 \text{ GeV}$. We use the data set where vector meson contributions were subtracted. We sum in quadrature statistical and systematic errors and we ignore correlations. In our numerical calculations we always use the average values of the kinematic variables in each bin.

The scatter plot of q_T^2/Q^2 vs $y_p - y_h$ for HERMES multiplicity data is shown in Fig. 2. In order to interpret the data in terms of TMD factorization, one must ensure applicability of factorization and make sure that the production mechanism corresponds to the current fragmentation. Another requirement is $q_T \ll Q$. One can see from Fig. 2 that

---

[2] The canonical partonic power counting for the initial and final quark light-cone momenta $k_i = \left(O(Q), O(m^2/Q), O(\mathbf{m}) \right)$, $k_f = \left(O(m^2/Q), O(Q), O(\mathbf{m}) \right)$; $|k_i^2| = k_f^2 = O(m^2)$ (Note that k_i is normally space-like.) For power counting purposes, m is to be understood as a combination of the small mass scales, $m \in \{\Lambda_{QCD}, M_p\}$.
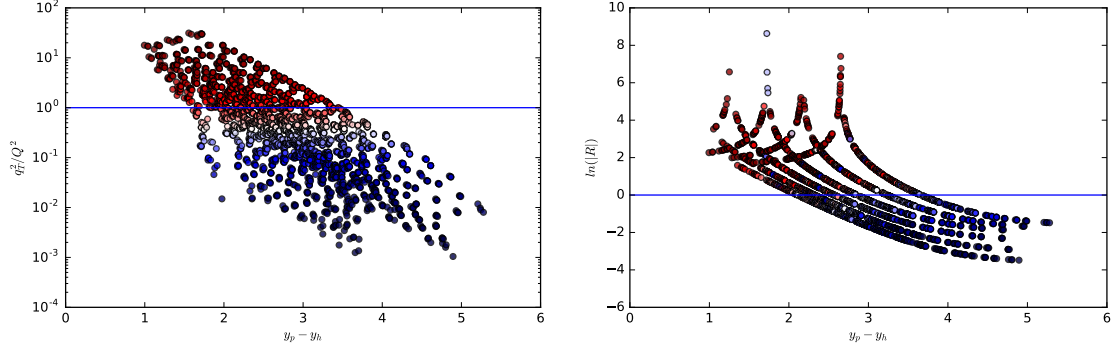
Figure 2: [Color online] **Left panel:** Scatter plot of q_T^2/Q^2 vs $y_p - y_h$ for HERMES multiplicity data. The solid line corresponds to $q_T^2/Q^2 = 1$. **Right panel:** Scatter plot of $\ln|R|$ vs $y_p - y_h$ for HERMES multiplicity data. Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.
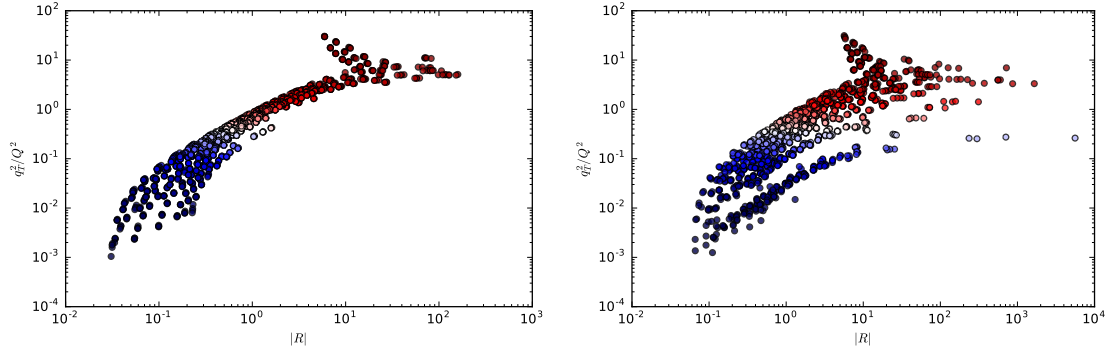

Figure 3: [Color online] **Left panel:** Scatter plot of q_T^2/Q^2 vs $|R|$ for $\pi^\pm$ HERMES multiplicity data . **Right panel:** Scatter plot of q_T^2/Q^2 vs $|R|$ for $K^\pm$ HERMES multiplicity data . Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

q_T^2/Q^2 for HERMES data can reach high values, and indeed 799 points are such that $q_T^2/Q^2 > 1$. As we mentioned, Ref. [1] proposes collinearity, R, as a filter to safely separate the "forward" and "backward" rapidity regions and ensure current fragmentation. One can also see from the right panel of Fig. 2 that values of $|R|$ HERMES are correlated with values of $y_p - y_h$. In particular, after $y_p - y_h$ reaches 4 units in rapidity, $|R|$ values become small. One can see from Fig. 4 that lower values of $|R|$ are correlated with higher values of $y_p - y_h$, one can also see that $|R|$ becomes very large for low values of $y_p - y_h$. The distinct families of curves in the right panel of Fig. 2 correspond to 8 bins in x_B of HERMES data. For a given bin in x_B all data for different hadron types, different targets, values of z_h, Q^2, and P_{hT} belong to a particular distinct curve. One can see that R is an exponential measure, so we will introduce an R-filter using $\ln R < A$ instead of $R < A$, where $A < 1$.

Figure 3 shows correlation of q_T^2/Q^2 and $|R|$ separately for pions (left panel) and Kaons (right panel). We can state that if one imposes $R < 1$-cut, then *simultaneously* q_T^2/Q^2 are cut to smaller values $q_T^2/Q^2 < 1$, see both panels of Fig. 3. It means that collinearity can simultaneously ensure that the particle is produced in current fragmentation region and that $q_T^2/Q^2 < 1$.

Figure 4 shows correlation of q_T^2/Q^2 and $y_p - y_h$ separately for pions (left panel) and Kaons (right panel). One can see from Fig. 4 only for large rapidity separation $y_p - y_h > 3.5$ the values of q_T^2/Q^2 become smaller than 1. One can also see that pions have larger maximal rapidity interval (~ 5.5) compared to Kaons (~ 4.5). It means that if we use $y_p - y_h > A$, where $A > 1$ is sufficiently large as a cut, Kaon TMD parameters will be determined worser for large values of A.

The exact starting value of q_T^2/Q^2 when the TMD factorization should become appropriate is not exactly known

8

and should be determined phenomenologically. In order to be able to describe SIDIS cross section in a wide region of q_T one should use the so-called $W + Y$ prescription, see for instance Ref. [22], however at present for SIDIS data there are difficulties in implementation of this prescription, see Ref. [30]. In this study we will use only TMD approximated cross-section as in Eq. (3) and consequently multiplicity as in Eq. (23).
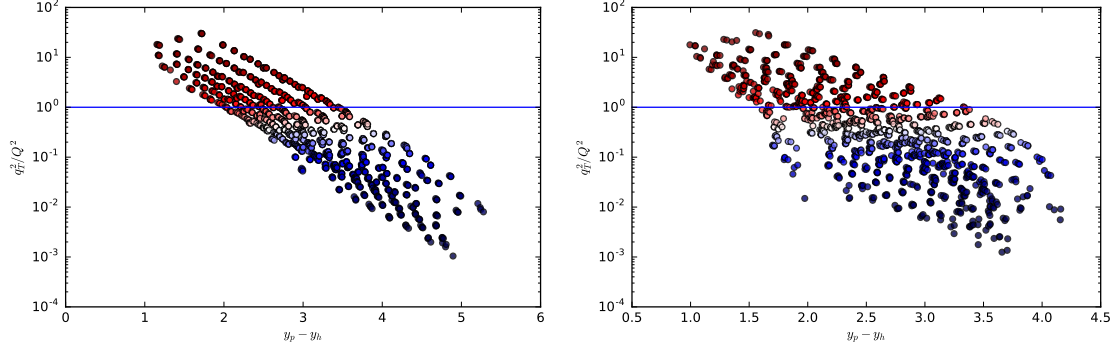

Figure 4: [Color online] **Left panel:** Scatter plot of q_T^2/Q^2 vs $y_p - y_h$ for $\pi^\pm$ HERMES multiplicity data . **Right panel:** Scatter plot of q_T^2/Q^2 vs $y_p - y_h$ for $K^\pm$ HERMES multiplicity data . The solid line corresponds to $R = 1$. Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

Let us discuss first the application of the "standard" cut from Eq. (24). As shown in Ref. [17] the data filtered by Eq. (24) can be described well in the framework of TMD parton model. The cut of Eq. (24) however does not necessarily ensure filtering the data for which TMD factorization is applicable, moreover, as pointed out in Ref. [1] it does not ensure filtering of the data in the current fragmentation region. In order to demonstrate it, let us plot q_T^2/Q^2 versus $y_p - y_h$ in Fig. 5. One can see from Fig. 5 that q_T^2/Q^2 can reach high values, even though one should ensure
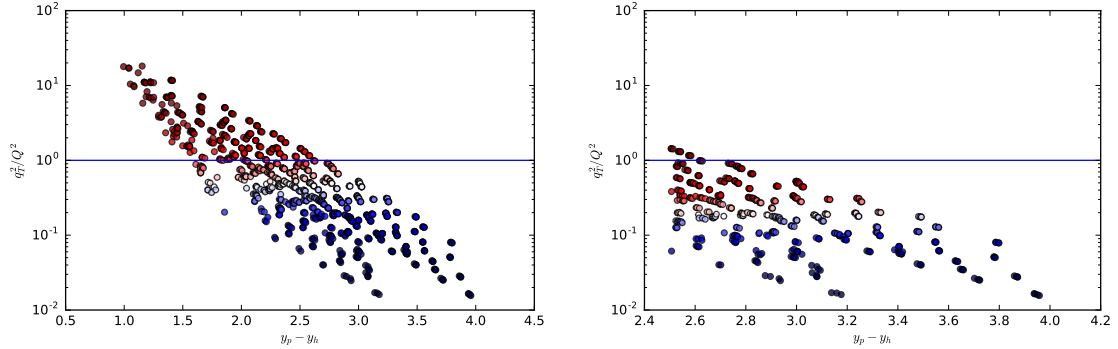

Figure 5: [Color online] **Left panel:** Scatter plot of q_T^2/Q^2 vs $y_p - y_h$ for HERMES multiplicity data filtered by the "standard" cut from Eq. (24). **Right panel:** Scatter plot of HERMES multiplicity data filtered by the "standard" cut from Eq. (35) and $y_p - y_h > 2.5$. The solid line corresponds to $q_T^2/Q^2 = 1$. Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

that $q_T^2/Q^2 \ll 1$ for the applicability of TMD factorization. In case of HERMES multiplicities, cuts from Eq. (24) result in total of 978 points and 292 points are such that $q_T^2/Q^2 > 1$. On the other hand, if we add $y_p - y_h > 2.5$ selection criteria to Eq. (24), then only 11 out of 473 points are such that $q_T^2/Q^2 > 1$. In order to study the influence of restriction of the rapidity interval on the results of the fit, we will use cuts from Eq. (24) and add $y_p - y_h > A$ selection criteria, where $A \in [1.25, 3.5]$:

$$0.2 < z_h < 0.6 \qquad Q^2 > 1.69 \, \text{GeV}^2 \qquad 0.2 < P_{hT} < 0.9 \, \text{GeV} \qquad y_p - y_h > A \qquad \text{(Torino + rapidity cut)} . \quad (35)$$

An example of such data selection for A = 2.5 is shown in the right panel of Fig. 5. Notice that we cut the values of the "padding" bin $z_h < 0.2$, Ref. [31].

9

The second set of cuts that we will explore will be motivated by ensuring applicability of TMD factorization together with selection of large rapidity interval, we will call it "rapidity" cut:

$$0.2 < z_h < 0.6 \qquad Q^2 > 1.69\,\text{GeV}^2 \qquad q_T^2/Q^2 < 0.15 \qquad y_p - y_h > A \qquad \text{(rapidity + q}_\text{T}\text{ cut)}. \qquad (36)$$

We will vary the value $A \in [1.25, 3.5]$ in order to test the sensitivity of the fit results to this value. Scatter plot of the filtered data by the rapidity cut from Eq. (36) is shown in Fig. 6.
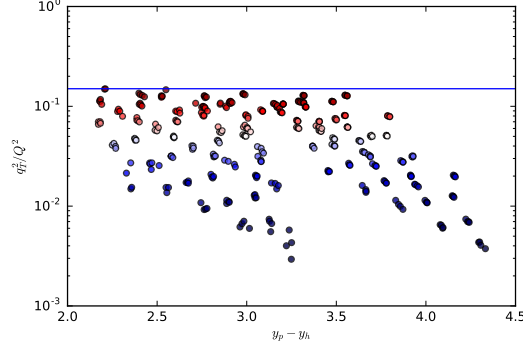
Figure 6: [Color online] Scatter plot of HERMES multiplicity data filtered by the rapidity cut from Eq. (36). The solid line corresponds to $q_T^2/Q^2 = 0.15$. Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

The third set of cuts that we will explore will include collinearity criteria from Ref. [1], we will discard any data for large values of $|R|$, we will call it R-cut:

$$0.2 < z_h < 0.6 \qquad Q^2 > 1.69\,\text{GeV}^2 \qquad \ln|R| < A \qquad \text{(R cut)}. \qquad (37)$$

We will vary the value $A \in [-0.5, -2.5]$ in order to test the sensitivity of the fit results to this value. The scatter plot of HERMES multiplicity data filtered by the R-cut from Eq. (37), $\ln|R| < -0.5$ and $\ln|R| < -2.5$, is shown in Fig. 7. One can see that $\ln|R| < -0.5$ ensures selection of $q_T^2/Q^2 < 1$. One can also see from Fig. 7, that filtering larger negative values of $\ln|R|$ correspond to simultaneous filtering out smaller values of $y_p - y_h$, and large values of q_T^2/Q^2 (compare left and right panels of Fig. 7).
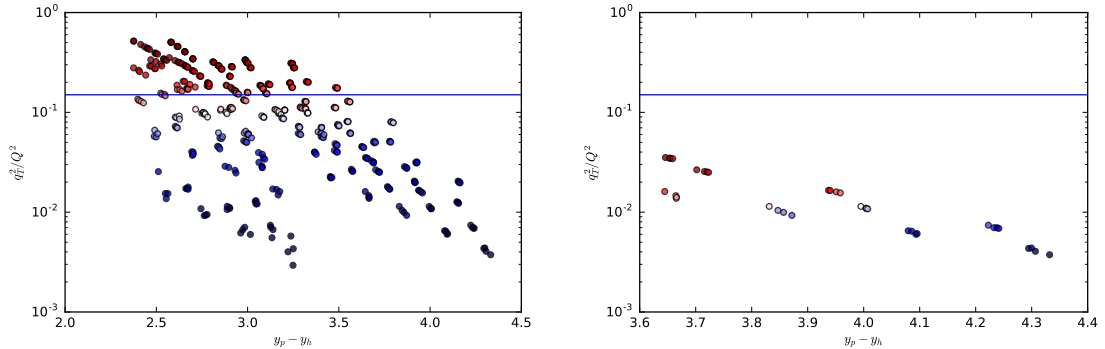

Figure 7: [Color online] **Left panel:** Scatter plot of HERMES multiplicity data filtered by the R-cut from Eq. (37), $\ln R < -0.5$. **Right panel:** Scatter plot of HERMES multiplicity data filtered by the R-cut from Eq. (37), $\ln R < -2.5$. The solid line corresponds to $q_T^2/Q^2 = 0.15$. Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

If one compares right panel of Fig. 5, Fig. 6, and Fig. 7, it is easy to see that data selections are quite similar, and thus we expect that fit results will be similar. We fit the (flavor-dependent) Gaussian widths to the the multiplicities using Eq. (23) (along with (22)). The details of this extraction and the results for our parameters will be discussed in the next section.

10

3.3. Fit and Results

We will use the following collinear functions in our analysis: PDFs are CJ15LO from Ref. [32], FFs are the DSS LO FF set from Ref. [33]. The value of initial scale Q_0 will always coincide with the lowest Q^2 cut-off in our cuts: $Q_0^2 = 1.69$ (GeV^2). We will make the following assumptions on the intrinsic non perturbative parameters: For the transverse-momentum widths $\langle k_\perp^2 \rangle_q$ of the TMD PDFs, two Gaussian widths are used, one for the valence type $\langle k_\perp^2 \rangle_{val}$ ($q = u, d$) and one for the sea-quark type $\langle k_\perp^2 \rangle_{sea}$ ($q = \bar{u}, \bar{d}, s, \bar{s}$) functions. Similarly, for the TMD FFs two Gaussian widths for $\langle p_\perp^2 \rangle_q$ are used, for the favored (such as u or $\bar{d}$ to π^+) and unfavored ($\bar{u}$ or d to π^+) type of FF. We have separate favored and unfavored widths for pion and Kaon fragmentation TMDs. Finally, we also fit g_{K_0} that encodes non perturbative TMD evolution. In total, we therefore have 7 parameters to be extracted from HERMES data. The theoretical expectation from the chiral soliton model [34] is to have a larger sea-quark width compared to valence quark width, so that $\langle k_\perp^2 \rangle_{sea} / \langle k_\perp^2 \rangle_{val} \sim 5$. Extraction of parameters and error analysis will be performed by using Monte Carlo techniques [35, 35] developed by Jefferson Lab JAM Collaboration [36]. Using the nested sampling MC algorithm [37–39], we compute the expectation value E[O] and variance V[O],

$$
\mathrm{E}[O] \;=\; \int d^n a\, \mathcal{P}(\boldsymbol{a}|\mathrm{data})\, O(\boldsymbol{a}) \;\simeq\; \sum_k w_k\, O(\boldsymbol{a}_k), \tag{38a}
$$

$$
\begin{aligned}
\mathrm{V}[O] \;&=\; \int d^n a\, \mathcal{P}(\boldsymbol{a}|\mathrm{data})(O(\boldsymbol{a}) - \mathrm{E}[O])^2 \\
&\simeq\; \sum_k w_k (O(\boldsymbol{a}_k) - \mathrm{E}[O])^2,
\end{aligned} \tag{38b}
$$

for each observable O (such as a TMD or a function of TMDs), which is a function of the n-dimensional vector parameters $\boldsymbol{a}$ with probability density $\mathcal{P}(\boldsymbol{a}|\mathrm{data})$ [40]. Using Bayes' theorem, the latter is given by

$$
\mathcal{P}(\boldsymbol{a}|\mathrm{data}) \;=\; \frac{1}{Z}\, \mathcal{L}(\mathrm{data}|\boldsymbol{a})\, \pi(\boldsymbol{a}), \tag{39}
$$

where $\pi(\boldsymbol{a})$ is the prior distribution for the vector parameters $\boldsymbol{a}$, and

$$
\mathcal{L}(\mathrm{data}|\boldsymbol{a}) \;=\; \exp\left[-\frac{1}{2}\chi^2(\boldsymbol{a})\right] \tag{40}
$$

is the likelihood function, with $Z = \int d^n a\, \mathcal{L}(\mathrm{data}|\boldsymbol{a})\, \pi(\boldsymbol{a})$ the Bayesian evidence parameter. Using a flat prior, the nested sampling algorithm constructs a set of MC samples $\{\boldsymbol{a}_k\}$ with weights $\{w_k\}$, which are then used to evaluate observables.

Resulting parameters of the fits with the new "standard" cut from Eq. (35) are shown in Fig. 8.

[Comments, discussions....]

Notice, that our results for the lowest $y_p - y_h$ cut are very close to the results of Ref. [17]: $\langle k_\perp^2 \rangle = 0.57$ GeV^2, $\langle p_\perp^2 \rangle = 0.12$ GeV^2. We will explore the reason for this similarity in Appendix A.

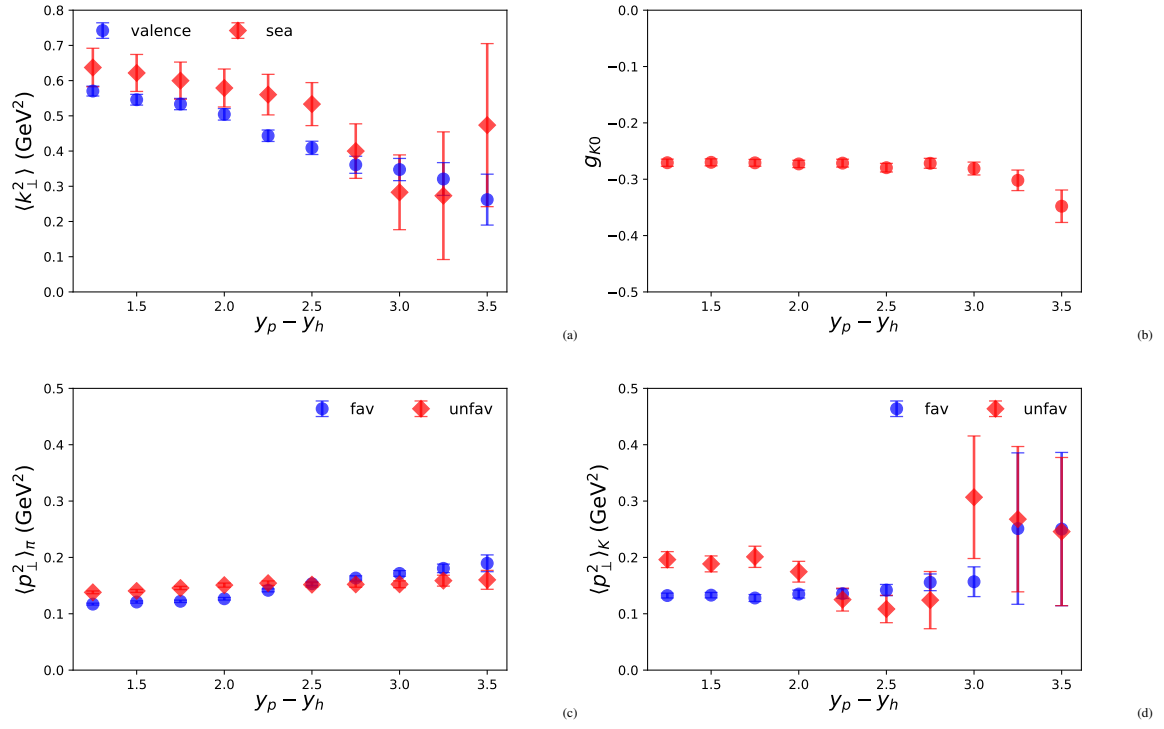Figure 8: Fitted parameters of the model with cuts from Eq. (35) as function of $y_p - y_h$. (a) $\langle k_\perp^2 \rangle_a$ for valence and sea quarks, (b) g_{K_0}, (c) $\langle p_\perp^2 \rangle_a$ for pion favored and unfavored fragmentation, (d) $\langle p_\perp^2 \rangle_a$ for Kaon favored and unfavored fragmentation.

Resulting parameters of the fits with the "rapidity" cut from Eq. (36) are shown in Fig. 9.
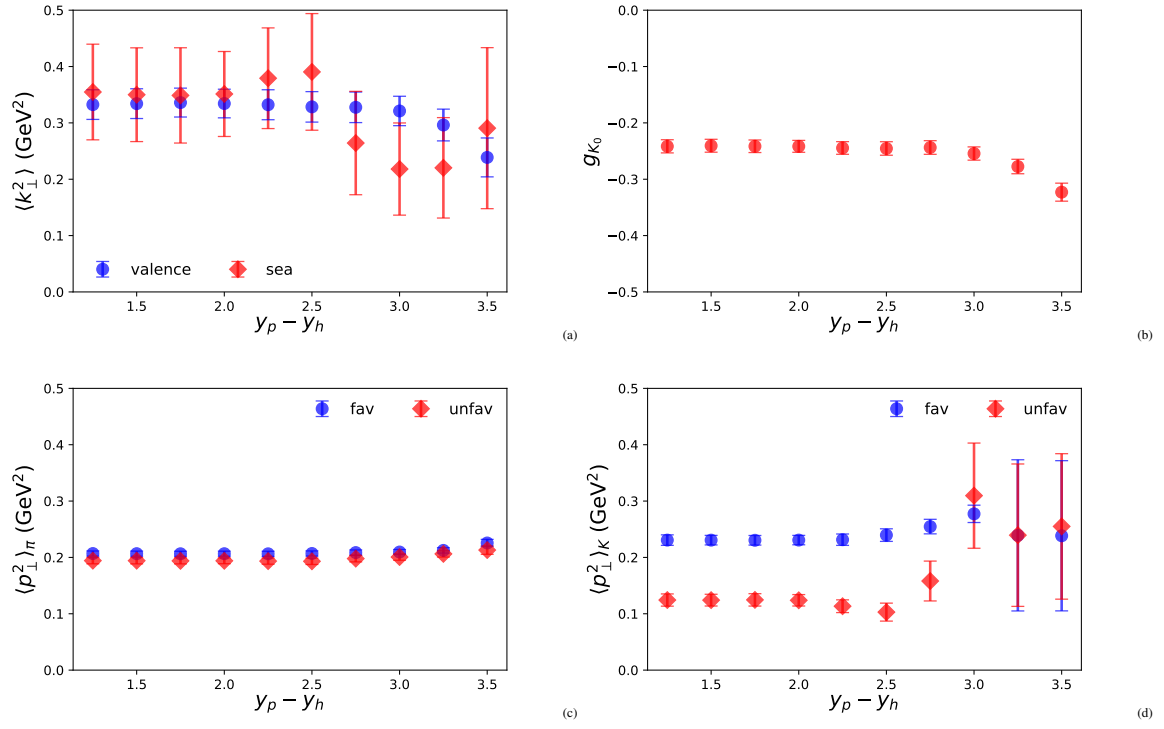
[Comments, discussions....]

Figure 9: Fitted parameters of the model with cuts from Eq. (36) as function of $y_p - y_h$. (a) $\langle k_\perp^2 \rangle_a$ for valence and sea quarks, (b) g_{K_0}, (c) $\langle p_\perp^2 \rangle_a$ for pion favored and unfavored fragmentation, (d) $\langle p_\perp^2 \rangle_a$ for Kaon favored and unfavored fragmentation.

Resulting parameters of the fits with the R-cut from Eq. (37) are shown in Fig. 10.
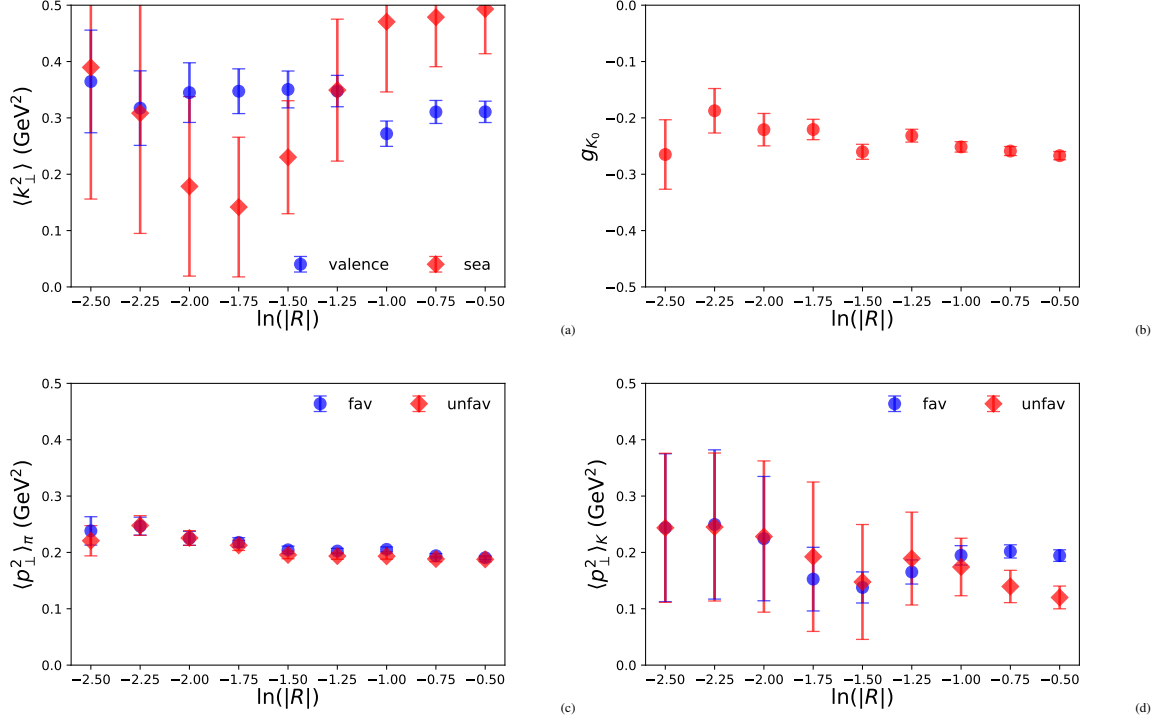
[Comments, discussions....]

13

Figure 10: Fitted parameters of the model with cuts from Eq. (37) as function of $\ln(|R|)$. (a) $\langle k_\perp^2 \rangle_a$ for valence and sea quarks, (b) g_{K_0}, (c) $\langle p_\perp^2 \rangle_a$ for pion favored and unfavored fragmentation, (d) $\langle p_\perp^2 \rangle_a$ for Kaon favored and unfavored fragmentation.

[Give details of fit: PDFs/FFs used, $\chi^2/d.o.f.$, error analysis, plots.] In the end, we find the following Gaussian widths: $\langle k_\perp^2 \rangle_{u_v} = 0.25\,\text{GeV}^2$, $\langle p_\perp^2 \rangle_{fav} = 0.17\,\text{GeV}^2, \ldots$. These widths are consistent with an earlier extraction using high-energy EMC data [CITE], where one would expect little impact from the R-cut given that the current and target fragmentation regions are well-separated in that case. Therefore, we see the reduction in especially $\langle k_\perp^2 \rangle$ found in Refs. [CITE] is most likely due to "contamination" from data in the target fragmentation region. Also, even though Ref. [CITE] considers HERMES and COMPASS data, the majority of the data points used in that analysis are from high-energy Z-boson production at Fermi-lab. Consequently, the widths found there are closer to the ones in this analysis since the "contaminated" HERMES and COMPASS data do not weigh heavily in the global analysis. [Elaborate more. Evolution effects?]

4. Conclusions

[Summary of results. Future work. Impact moving forward.]

Acknowledgments

14

Appendix A. Comparison of TMD evolution and Torino method

Ref. [17] uses usual DGLAP evolution for TMDs, such that

$$f_{a/N}(x, k_\perp; Q^2) = f_{a/N}(x; Q^2) \frac{e^{-k_\perp^2/\langle k_\perp^2 \rangle_a}}{\pi \langle k_\perp^2 \rangle_a},$$

$$D_{h/a}(z, k_\perp; Q^2) = \frac{1}{z^2} D_{h/a}(z; Q^2) \frac{e^{-k_\perp^2/\langle p_\perp^2 \rangle_a}}{\pi \langle p_\perp^2 \rangle_a}. \tag{A.1}$$

Using our formulas we obtain

$$f_{a/N}(x, k_\perp; Q^2) = f_{a/N}(x; Q_0^2) \left(\frac{Q}{Q_0}\right)^{g_{\kappa_0}} e^{-S_{pert}/2} \frac{e^{-k_\perp^2/\langle k_\perp^2 \rangle_a}}{\pi \langle k_\perp^2 \rangle_a},$$

$$D_{h/a}(z, k_\perp; Q^2) = \frac{1}{z^2} D_{h/a}(z; Q_0^2) \left(\frac{Q}{Q_0}\right)^{g_{\kappa_0}} e^{-S_{pert}/2} \frac{e^{-k_\perp^2/\langle p_\perp^2 \rangle_a}}{\pi \langle p_\perp^2 \rangle_a}. \tag{A.2}$$

A natural question is: if the widths extracted using Eqs. (A.1) and Eqs. (A.2) are similar, does it mean that functions of x, z and Q^2 are similar?

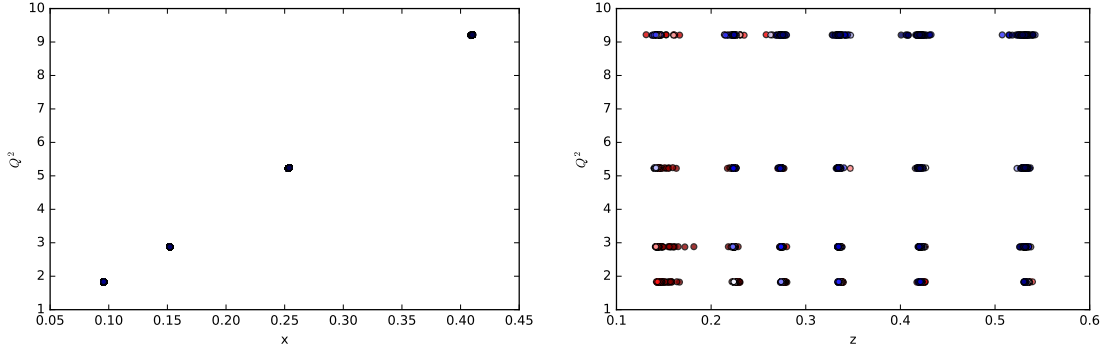In order to study it, let us first plot the rage of x, z and Q^2 for HERMES data in Fig. A.11

Figure A.11: [Color online] **Left panel:** Scatter plot of Q^2 vs x for $\pi^\pm$ HERMES multiplicity data . **Right panel:** Scatter plot of Q^2 vs z for $K^\pm$ HERMES multiplicity data . Red (deep blue) color in the plots corresponds to growing (diminishing) values of q_T^2/Q^2.

In order to achieve similar extractions, one would need to have Eqs. (A.1) and Eqs. (A.2) similar for corresponding values of x, z, Q^2.

We plot in Fig. A.12 up quark distributions for typical values of $Q^2 = 2, 3, 5, 9 \text{ GeV}^2$. One can see that indeed the values of distributions are quite similar in $x_B = 0.1, 0.15, 0.25, 0.4$ accordingly.

The same happens for fragmentation functions, especially at larger values of z, see Fig. A.13.

More precise data, in particular from Jefferson Lab 12 GeV upgrade will allow to appreciate the difference of the two approaches.

We conclude that even though our methods are very different in the shape, the results are similar for HERMES data.

[1] M. Boglione *et al.*, Phys. Lett. **B766**, 245 (2017), 1611.10329.
[2] A. Kotzinian, Nucl. Phys. **B441**, 234 (1995), hep-ph/9412283.
[3] D. Boer, R. Jakob, and P. J. Mulders, Nucl. Phys. **B504**, 345 (1997), hep-ph/9702281.
[4] R. D. Tangerman and P. J. Mulders, Phys. Rev. **D51**, 3357 (1995), hep-ph/9403227.
[5] J. C. Collins and D. E. Soper, Nucl. Phys. **B194**, 445 (1982).
[6] X.-d. Ji, J.-p. Ma, and F. c. g. g. h. Yuan, Phys. Rev. **D71**, 034005 (2005), hep-ph/0404183.
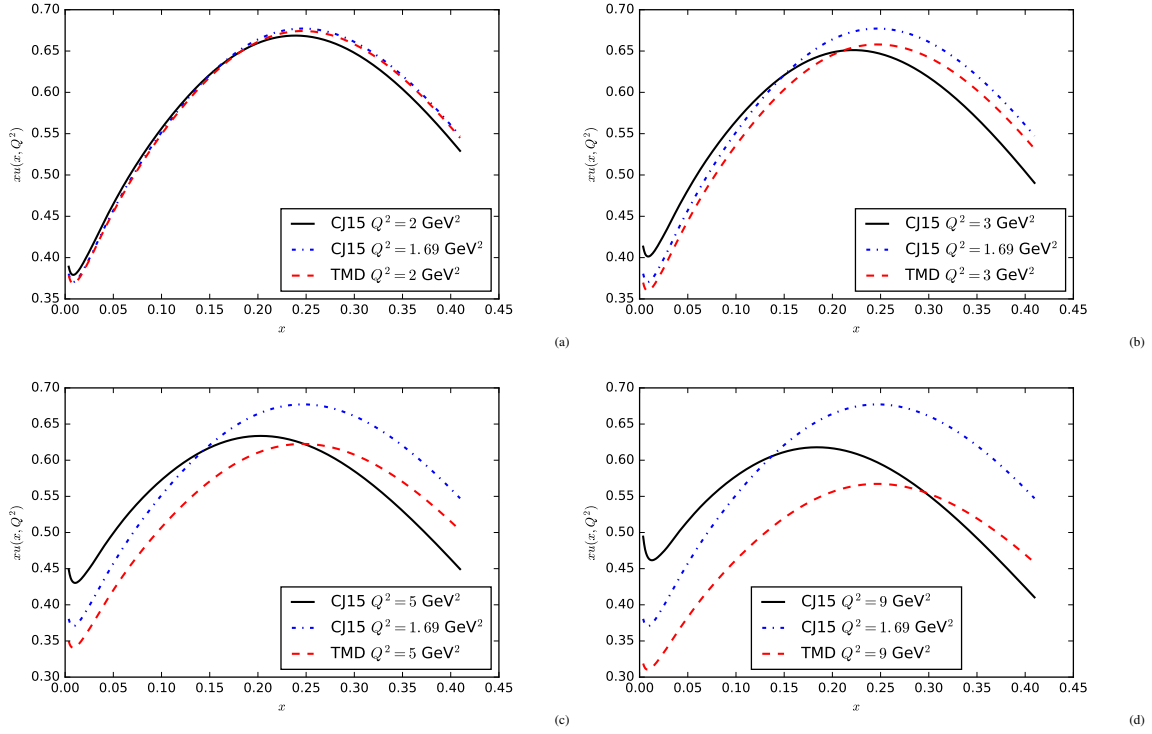[7] J. Collins, *Foundations of Perturbative QCD*, (Cambridge University Press, Cambridge, England, 2011).

Figure A.12: Up quark distributions at different values of Q^2 using DGLAP and Eqs. A.2.

[8] P. J. Mulders and R. D. Tangerman, Nucl. Phys. **B461**, 197 (1996), hep-ph/9510301.

[9] D. Boer and P. J. Mulders, Phys. Rev. **D57**, 5780 (1998), hep-ph/9711485.

[10] E. L. Berger, SEMIINCLUSIVE INELASTIC ELECTRON SCATTERING FROM NUCLEI, in *NPAS WORKSHOP ON ELECTRONU-CLEAR PHYSICS WITH INTERNAL TARGETS, SLAC, JANUARY 5-8, 1987*, pp. 82–91, 1987.

[11] L. Trentadue and G. Veneziano, Phys. Lett. **B323**, 201 (1994).

[12] M. Grazzini, L. Trentadue, and G. Veneziano, Nucl. Phys. **B519**, 394 (1998), hep-ph/9709452.

[13] M. Anselmino, V. Barone, and A. Kotzinian, Phys. Lett. **B699**, 108 (2011), 1102.4214.

[14] P. J. Mulders, AIP Conf. Proc. **588**, 75 (2001), hep-ph/0010199.

[15] S. J. Joosten, *Fragmentation and nucleon structure in semi-inclusive deep-inelastic scattering at the HERMES experiment*, PhD thesis, Illinois U., Urbana, 2013.

[16] J. Collins and T. C. Rogers, (2018), 1801.02704.

[17] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, JHEP **04**, 005 (2014), 1312.6261.

[18] A. Bacchetta *et al.*, JHEP **0702**, 093 (2007), hep-ph/0611265.

[19] J. Collins and T. Rogers, Phys. Rev. **D91**, 074020 (2015), 1412.3820.

[20] P. Schweitzer, T. Teckentrup, and A. Metz, Phys. Rev. **D81**, 094019 (2010), 1003.2190.

[21] A. Signori, A. Bacchetta, M. Radici, and G. Schnell, JHEP **11**, 194 (2013), 1309.3507.

[22] J. Collins *et al.*, Phys. Rev. **D94**, 034014 (2016), 1605.00671.

[23] M. Anselmino *et al.*, Phys.Rev. **D71**, 074006 (2005), hep-ph/0501196.

[24] S. Aybat and T. C. Rogers, Phys.Rev. **D83**, 114042 (2011), 1101.5057.

[25] S. Bethke, (2012), 1210.0325, [Nucl. Phys. Proc. Suppl.234,229(2013)].

[26] C. A. Aidala, B. Field, L. P. Gamberg, and T. C. Rogers, Phys. Rev. **D89**, 094002 (2014), 1401.2654.

[27] HERMES, A. Airapetian *et al.*, Phys. Rev. **D87**, 074029 (2013), 1212.5407.

[28] COMPASS, C. Adolph *et al.*, Eur. Phys. J. **C73**, 2531 (2013), 1305.7317, [Erratum: Eur. Phys. J.C75,no.2,94(2015)].

[29] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, JHEP **06**, 081 (2017), 1703.10157.

[30] J. O. Gonzalez-Hernandez, T. C. Rogers, N. Sato, and B. Wang, (2018), 1808.04396.

[31] G. Schnell, Private communications.

[32] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, Phys. Rev. **D93**, 114017 (2016), 1602.03154.

[33] D. de Florian, R. Sassot, and M. Stratmann, Phys. Rev. **D75**, 114010 (2007), hep-ph/0703242.

[34] P. Schweitzer, M. Strikman, and C. Weiss, JHEP **01**, 163 (2013), 1210.1267.

[35] N. Sato, W. Melnitchouk, S. E. Kuhn, J. J. Ethier, and A. Accardi, Phys. Rev. D **93**, 074005 (2016).

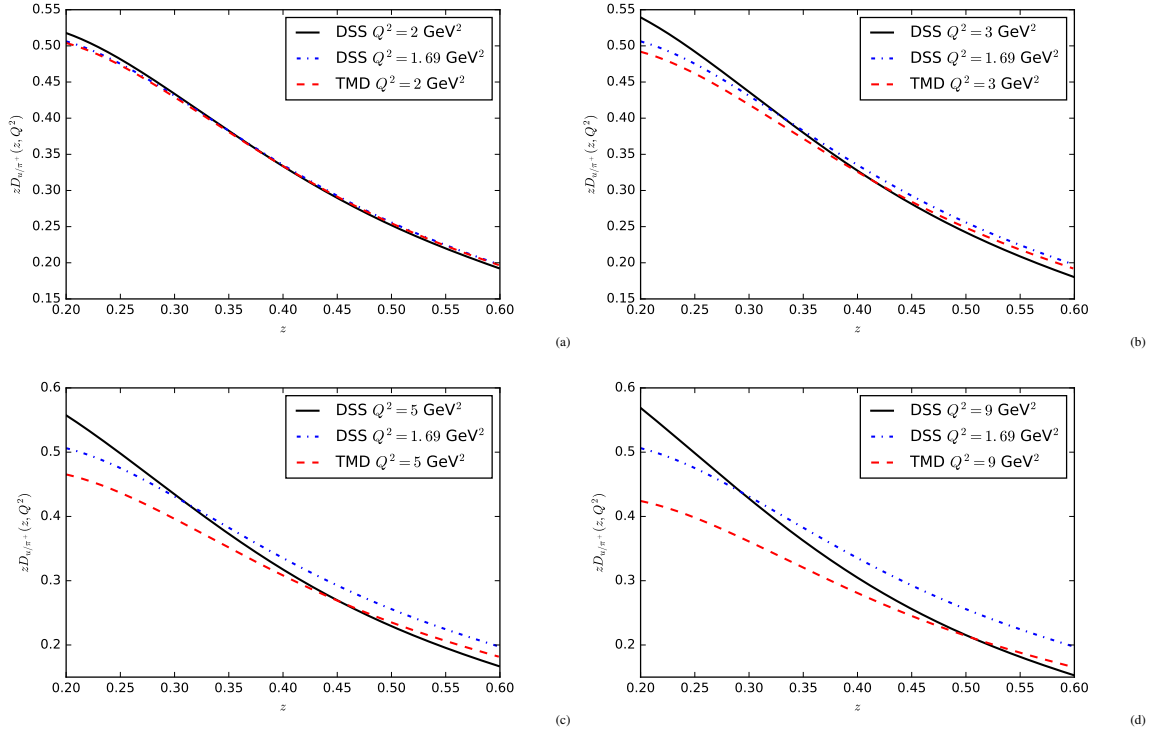[36] JAM Collaboration, https://www.jlab.org/theory/jam/.

Figure A.13: Up quark into π^+ fragmentation functions at different values of Q^2 using DGLAP and Eqs. A.2.

[37] J. Skilling, Nested sampling for general Bayesian computation, http://www.inference.phy.cam.ac.uk/bayesys, 2004.

[38] P. Mukherjee, D. Parkinson, and A. R. Liddle, Astrophys. J. **638**, L51 (2006), astro-ph/0508461.

[39] R. Shaw, M. Bridges, and M. P. Hobson, Mon. Not. Roy. Astron. Soc. **378**, 1365 (2007), astro-ph/0701867.

[40] N. Sato *et al.*, Phys. Rev. D **94**, 114004 (2016).