Factorization at the LHC: From parton distribution functions to initial state jets

PHYSICAL REVIEW D **81,** 094035 (2010)

Factorization at the LHC: From parton distribution functions to initial state jets

Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn

Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 20 November 2009; published 28 May 2010)

We study proton-(anti)proton collisions at the LHC or Tevatron in the presence of experimental restrictions on the hadronic final state and for generic parton momentum fractions. At the scale Q of the hard interaction, factorization does not yield standard parton distribution functions (PDFs) for the initial state. The measurement restricting the hadronic final state introduces a new scale $\mu_B \ll Q$ and probes the proton prior to the hard collision. This corresponds to evaluating the PDFs at the scale μ_B. After the proton is probed, the incoming hard parton is contained in an initial-state jet, and the hard collision occurs between partons inside these jets rather than inside protons. The proper description of such initial-state jets requires ''beam functions''. At the scale μ_B, the beam function factorizes into a convolution of calculable Wilson coefficients and PDFs. Below μ_B, the initial-state evolution is described by the usual PDF evolution which changes x, while above μ_B it is governed by a different renormalization group evolution that sums double logarithms of μ_B/Q and leaves x fixed. As an example, we prove a factorization theorem for ''isolated Drell-Yan'', $pp \to X\ell^+\ell^-$ where X is restricted to have no central jets. We comment on the extension to cases where the hadronic final state contains a certain number of isolated central jets.

DOI: 10.1103/PhysRevD.81.094035 PACS numbers: 12.39.St, 12.38.Cy, 13.85.Qk, 13.87.Ce

I. INTRODUCTION

Factorization is one of the most basic concepts for understanding data from the Tevatron at Fermilab and the CERN Large Hadron Collider (LHC). For a review of factorization see Ref. [1]. Typically, factorization is viewed as the statement that the cross section can be computed through a product of probability functions, namely, parton distribution functions (PDFs), describing the probability to extract a quark or gluon from the protons in the initial state, a perturbative cross section for the hard scattering, and a probabilistic description of the final state by a parton shower Monte Carlo or otherwise. This factorization is of key importance in the program to search for new physics, as new physics is primarily a short-distance modification of the hard scattering that must be distinguished from the array of QCD interactions in the initial and final states. Factorization is also necessary for controlling QCD effects. For example, the momentum distributions of the colliding partons in the protons are nonperturbative, but factorization can imply that these are described by universal distributions which have been measured in earlier experiments.

As the primary goal of the experiments at the LHC or Tevatron is to probe the physics of the hard interaction, measurements often impose restrictions on the hadronic final state, requiring a certain number of hard leptons or jets in the final state [2–5]. For example, a typical new physics search looking for missing transverse energy may also require a minimum number of jets with p_T above some threshold. To identify the new physics and determine the masses of new-physics particles, one has to reconstruct decay chains with a certain number of jets and leptons in the final state.

Any theoretical prediction for pp or $p\bar{p}$ collisions, whether analytic or via Monte Carlo generators, depends on factorization. However, for the majority of processes of interest at hadron colliders where one distinguishes properties of the hadronic final state, so far no rigorous field-theoretic derivation of a factorization theorem to all orders in perturbation theory exists. The most well-known factorization theorem is

$$d\sigma = \sum_{i,j} d\sigma_{ij}^{\text{part}} \otimes f_i(\xi_a) \otimes f_j(\xi_b), \tag{1}$$

where f_i and f_j are the standard PDFs for partons $i, j = \{g, u, \bar{u}, d, \ldots\}$ carrying momentum fractions ξ_a and ξ_b (which we use as our PDF x-variables), and $d\sigma_{ij}^{\text{part}}$ is the partonic cross section to scatter i and j calculated in fixed-order perturbation theory. In Eq. (1), the hadronic final state is treated as fully inclusive. Hence, in the presence of experimental restrictions that make a process less inclusive, Eq. (1) is *a priori* not applicable. At best, an additional resummation of large phase-space logarithms must be carried out by a further factorization of $d\sigma_{ij}^{\text{part}}$, while at worst, additional nonperturbative information beyond that contained in the PDFs is required or there is no factorization.

Factorization theorems for threshold resummation in hadron-hadron collisions are a well-studied case where Eq. (1) can be extended to sum large phase-space logarithms [6–14]. The corresponding formalism however requires the limit $x \to 1$, and hence is not directly relevant at the LHC, where the cross section for most measurements is dominated by the region x far from one [15].

Our goal is to study factorization for a situation where the hard interaction occurs between partons with generic

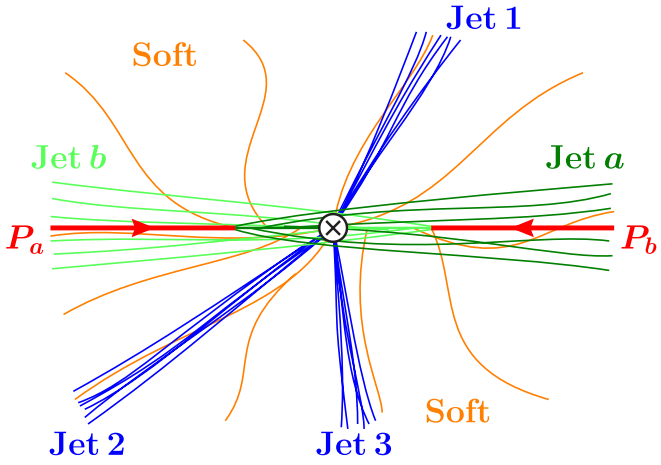


FIG. 1 (color online). A typical event with jet production at the LHC.

momentum fractions, away from the limit $x \to 1$, and where the hadronic final state is measured and restricted by constraints on certain kinematic variables. These restrictions allow one to probe more details about the final state and may be used experimentally to isolate central hard jets or leptons or to control backgrounds.

A typical event at the LHC with three high-p_T jets is illustrated in Fig. 1. There are several complications one has to face when trying to derive a factorization theorem in this situation. First, experimentally the number and properties of the final-state jets are determined with a jet algorithm. Second, to enhance the ratio of signal over background, the experimental analyses have to apply kinematic selection cuts. Third, in addition to the jets produced by the hard interaction, there is soft radiation everywhere (which is part of what is sometimes called the "underlying event"). Fourth, a (large) fraction of the total energy in the final state is deposited near the beam axes at high rapidities. An important component of this radiation can contribute to measurements, and when it does, it cannot be neglected in the factorization. In this paper we focus on the last three items. Methods for including jet algorithms in factorization have been studied in Refs. [8,16,17].

To allow a clean theoretical description, the observables used to constrain the events must be chosen carefully such that they are infrared safe and sensitive to emissions everywhere in phase space. Observables satisfying these criteria for hadron colliders have been classified and studied in Refs. [18,19], and are referred to as global event shapes. (Issues related to nonglobal observables have been discussed, for example, in Refs. [20–23].) For our analysis we use a very simple example of such an observable, constructed as follows. We define two hemispheres, a and b, orthogonal to the beam axis and two unit lightlike vectors n_a and n_b along the beam axis pointing into each hemisphere. Taking the beam axis along the z direction, hemisphere a is defined as $z > 0$ with $n_a^\mu = (1, 0, 0, 1)$, and hemisphere b as $z < 0$ with $n_b^\mu = (1, 0, 0, -1)$. We now divide the total momentum p_X of the hadronic final state into the contributions from particles in each hemisphere, $p_X = p_{X_a} + p_{X_b}$. Next, we remove the momenta p_J of all jets (defined by an appropriate jet algorithm) in each hemisphere. Of the remaining hemisphere momenta, we measure the components B_a^+ and B_b^+ defined by

$$B_a^+ = n_a \cdot \Big(p_{X_a} - \sum_{J \in a} p_J \Big), \tag{2}$$

and analogously for B_b^+. Because of the dot product with n_a or n_b, energetic particles near the beam axes only give small contributions to B_a^+ or B_b^+. In particular, any contributions from particles at very large rapidities outside the detector reach, including the remnant of unscattered partons in the proton, are negligible. All observed particles contribute either to B_a^+, B_b^+, or a jet momentum, so we are ensured that we cover all of phase space. Demanding that $B_{a,b}^+$ are small restricts the radiation between central jets, only allowing highly energetic particles either within these jets or inside jets along the beam directions labeled "Jet a" and "Jet b" in Fig. 1. Hence, measuring and constraining $B_{a,b}^+$ provides a theoretically clean method to control the remaining particles in the hadronic final state. This ensures that observables based on the large momenta of hard jets or leptons are clean, safe from uncontrolled hadronic effects.

In this paper, we consider the simplest situation where the above setup can be realized, allowing us to explore the implications of restrictions on the hadronic final state. We prove a factorization theorem for Drell-Yan production $pp \to X\ell^+\ell^-$ where X is allowed to have hard jets close to the beam, but no hard central jets. We call this "isolated Drell-Yan". Our proof of factorization uses the soft-collinear effective theory (SCET) [24–27] plus additional arguments to rule out possible Glauber effects based in part on Refs. [28,29]. Although we focus our discussion on Drell-Yan, our factorization theorem applies to processes $pp \to XL$, were the lepton pair is replaced by other non-strongly interacting particles, such as Higgs or Z' decaying nonhadronically. Though our analysis is only rigorous for $pp \to XL$, we also briefly discuss what the extended factorization formula may look like for processes with additional identified jets in the final state.

Our main result is to show that process-independent "beam functions", $B_i(t, x)$ with $i = \{g, u, \bar{u}, d, \ldots\}$, are required to properly describe the initial state. For the usual PDFs in Drell-Yan production appearing in Eq. (1), the hadronic final-state X is treated fully inclusively, and the effects of initial- and final-state soft radiation cancel out [1]. With restrictions on X, the effects of soft radiation can no longer cancel. Generically, by restricting X one performs an indirect measurement of the proton prior to the hard collision. At this point, the proton is resolved into a colliding hard parton inside a cloud of collinear and soft radiation. The proper description of this initial-state jet is given by a beam function in conjunction with an appro-

priate soft function describing the soft radiation in the event.

One might worry that the collision of partons inside initial-state jets rather than partons inside protons could drastically change the physical picture. Although the changes are not as dramatic, they have important implications. The beam function can be computed in an operator product expansion, giving

$$B_i(t, \xi, \mu_B) = \delta(t) f_i(\xi, \mu_B) + \mathcal{O}[\alpha_s(\mu_B)], \qquad (3)$$

where μ_B is an intermediate perturbative scale and t is an invariant-mass variable closely related to the off-shellness of the colliding parton (and the Mandelstam variable t). Thus, the beam functions reduce to standard PDFs at leading order. For what we call the gluon beam function, this was already found in Ref. [30], where the same matrix element of gluon fields appeared in their computation of $\gamma p \to J/\psi X$ using SCET.

Equation (3) implies that the momentum fractions $\xi_{a,b}$ are determined by PDFs evaluated at the scale $\mu_B \ll Q$, which is parametrically smaller than the scale Q of the partonic hard interaction. The renormalization group evolution (RGE) for the initial state now proceeds in two stages. For scales $\mu < \mu_B$, the RGE is given by the standard PDF evolution [31–35], which sums single logarithms, mixes the PDFs, and redistributes the momentum fractions in the proton to lower x values. For scales $\mu > \mu_B$, the jetlike structure of the initial state becomes relevant and its evolution is properly described by the RGE of the beam function. In contrast to the PDF, the evolution of the beam function is independent of x, does not involve any mixing between parton species, and sums Sudakov double logarithms. In addition to the change in evolution, the transition from PDFs to beam functions at the scale μ_B also involves explicit $\alpha_s(\mu_B)$ corrections as indicated in Eq. (3). These include mixing effects, such as a gluon from the proton pair-producing a quark that goes on to initiate the hard interaction and an antiquark that is radiated into the final state. For our observables such fluctuations are not fully accounted for by the PDF evolution. These beam effects must be taken into account, which can be done by perturbative calculations. The standard PDFs are still sufficient to describe the nonperturbative information required for the initial state.

One should ask whether the description of the initial state by beam functions, as well as their interplay with the soft radiation, are properly captured by current Monte Carlo event generators used to simulate events at the LHC and Tevatron, such as Pythia [36,37] and Herwig [38,39]. In these programs the corresponding effects should be described at leading order by the initial-state parton shower in conjunction with models for the underlying event [40–43]. The experimental implications and reliability of these QCD Monte Carlo models have been studied extensively [44–46]. We will see that the initial-state parton shower is in fact closer to factorization with beam functions than to the inclusive factorization formula in Eq. (1). In particular, the physical picture of off-shell partons that arises from the factorization with beam functions has a nice correspondence with the picture adopted for initial-state parton showers a long time ago [47,48]. There are also differences. Our analysis is based solely on QCD soft-collinear factorization, whereas the initial-state parton shower is partly based on the picture arising from small-x physics or semihard QCD [49]. For the parton distributions our formalism applies in a situation that is intermediate between the case of very small x, where a resummation of $\ln x$ becomes important, and the case $x \to 1$, where threshold resummation in $\ln(1-x)$ becomes important. Numerically, our results apply for the dominant region of x values that are of interest at the LHC. Experimentally, measurements of the isolated Drell-Yan cross section provide a simple observable that can rigorously test the accuracy of the initial-state shower in Monte Carlo programs, by contrasting it with the analytic results reported here.

In Sec. II, we discuss our main results and explain various aspects of the factorization with beam functions. The goal of this section is to give a thorough discussion of the physical picture behind our results which is nontechnical and accessible to nonexpert readers. In Sec. III, we elaborate on the field-theoretic definition and properties of the beam functions and their relation to the PDFs. We quote explicit results for the quark beam function at one loop, the derivation of which will be given in a separate publication [50]. In Sec. IV, we derive in detail the factorization theorem for isolated $pp \to XL$ using SCET, and apply it to the case of Drell-Yan. Readers not interested in the technical details can freely skip this section. Plots of the isolated Drell-Yan cross section are given in Sec. V. We conclude in Sec. VI.

II. FACTORIZATION WITH BEAM FUNCTIONS

This section provides an extensive discussion of how factorization with beam functions works, including the necessary kinematic definitions for the variables that constrain the hadronic final state. In the interest of avoiding technical details, we only discuss the physics contained in the factorization theorems. Readers interested in the field-theoretic definitions for the beam functions are referred to Sec. III, while those interested in the derivation of the factorization theorem in SCET and explicit definitions for all its ingredients are referred to Sec. IV.

In Sec. II A, we review the factorization theorems for inclusive Drell-Yan and threshold Drell-Yan, and then explain the factorization theorem for our isolated Drell-Yan process. We use a simple setup where measurements on the final-state hadrons use hemispheres orthogonal to the beam. These observables are generalized in Sec. II B to uniformly account for measurements that sample over a

wide variety of boosts between the hadronic and partonic center-of-mass frames. We explain the relation between beam functions and parton distribution functions in Sec. II C. We compare the beam-function renormalization group evolution to initial-state parton showers in Sec. II D. In Sec. II E, we show how the various pieces in the factorization theorem arise from the point of view of a fixed-order calculation. In Sec. II F, we compare the structure of large logarithms and their resummation for the different factorization theorems. This yields an independent argument for the necessity of beam functions and provides a road map for incorporating beam functions in other isolated processes. Finally in Sec. II G, we comment on the extension of the factorization with beam functions to the case where one has two or more isolated jets in the final state.

A. Drell-Yan factorization theorems

To describe the Drell-Yan process $pp \to X\ell^+\ell^-$ or $p\bar{p} \to X\ell^+\ell^-$, we take

$$P_a^\mu + P_b^\mu = p_X^\mu + q^\mu, \tag{4}$$

where $P_{a,b}^\mu$ are the incoming (anti)proton momenta, $E_{\rm cm} = \sqrt{(P_a + P_b)^2}$ is the total center-of-mass energy, and q^μ is the total momentum of the $\ell^+\ell^-$ pair. We also define

$$\begin{aligned} \tau &= \frac{q^2}{E_{\rm cm}^2}, \qquad Y = \frac{1}{2}\ln\frac{P_b\cdot q}{P_a\cdot q}, \\ x_a &= \sqrt{\tau}e^{Y}, \qquad x_b = \sqrt{\tau}e^{-Y}, \end{aligned} \tag{5}$$

where Y is the total rapidity of the leptons with respect to the beam axis, and x_a and x_b are in one-to-one correspondence with τ and Y. Their kinematic limits are

$$\begin{aligned} &0 \le \tau \le 1, \qquad 2|Y| \le -\ln\tau, \\ &\tau \le x_a \le 1, \qquad \tau \le x_b \le 1. \end{aligned} \tag{6}$$

The invariant mass of the hadronic final state is bounded by

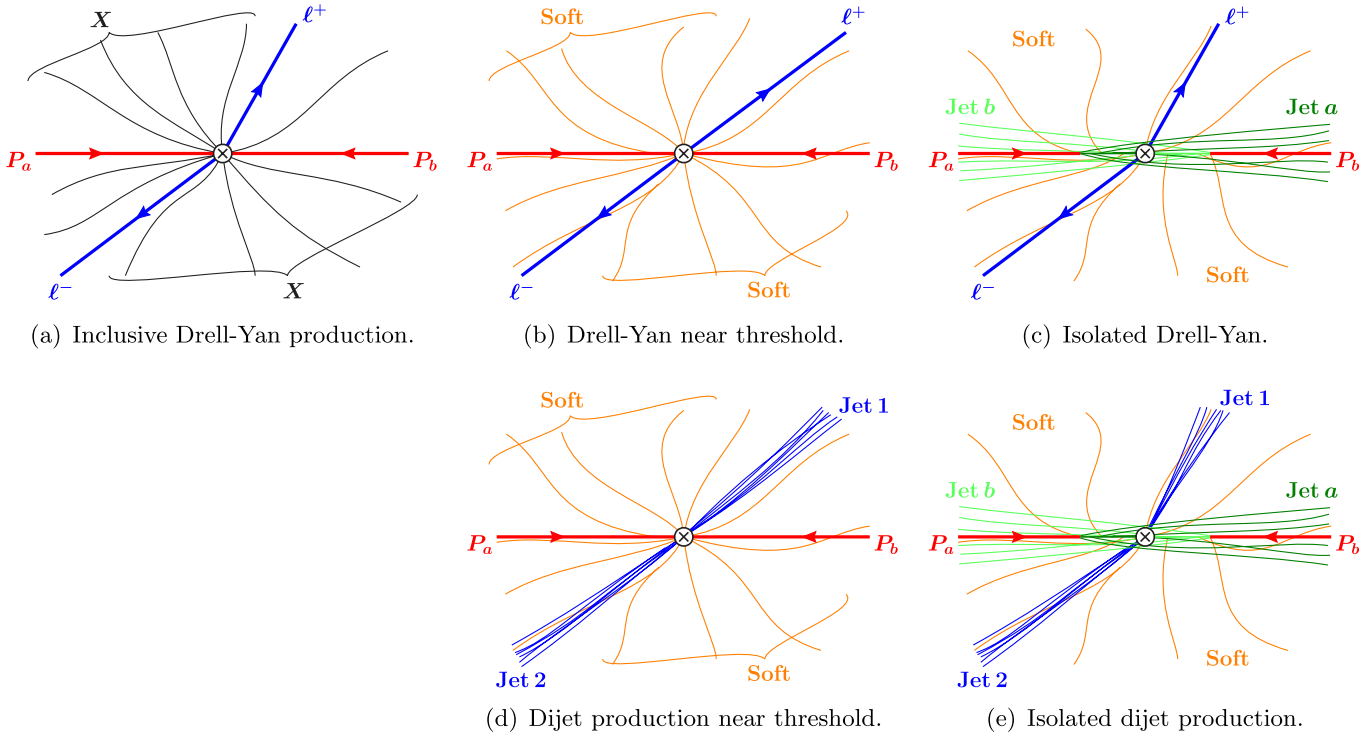
$$m_X^2 = p_X^2 \le E_{\rm cm}^2(1-\sqrt{\tau})^2. \tag{7}$$

In Drell-Yan

$$Q = \sqrt{q^2} \gg \Lambda_{\rm QCD} \tag{8}$$

plays the role of the hard interaction scale. In general, for factorization to be valid at some leading level of approximation with a perturbative computation of the hard scattering, the measured observable must be infrared safe and insensitive to the details of the hadronic final state.

For inclusive Drell-Yan, illustrated in Fig. 2(a), one sums over all hadronic final states X allowed by Eq. (7) without imposing any cuts. Hence, the measurement is insensitive to any details of X because one sums over all possibilities. In this situation there is a rigorous derivation



(a) Inclusive Drell-Yan production. (b) Drell-Yan near threshold. (c) Isolated Drell-Yan.

(d) Dijet production near threshold. (e) Isolated dijet production.

FIG. 2 (color online). Different final-state configurations for pp collisions. The top row corresponds to Drell-Yan factorization theorems for the (a) inclusive, (b) threshold, and (c) isolated cases. The bottom row shows the corresponding pictures with the lepton pair replaced by dijets.

of the classic factorization theorem [28,51,52]

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2dY} = \sum_{i,j}\int\frac{d\xi_a}{\xi_a}\frac{d\xi_b}{\xi_b}H_{ij}^{\rm incl}\left(\frac{x_a}{\xi_a},\frac{x_b}{\xi_b},q^2,\mu\right) \times f_i(\xi_a,\mu)f_j(\xi_b,\mu)\left[1+\mathcal{O}\left(\frac{\Lambda_{\rm QCD}}{Q}\right)\right], \quad (9)$$

where $\sigma_0 = 4\pi\alpha_{\rm em}^2/(3N_cE_{\rm cm}^2q^2)$, and the integration limits are $x_a \le \xi_a \le 1$ and $x_b \le \xi_b \le 1$. The sum is over partons i, $j = \{g, u, \bar{u}, d, \ldots\}$, and $f_i(\xi_a)$ is the parton distribution function for finding parton i inside the proton with light-cone momentum fraction ξ_a along the proton direction. Note that $\xi_{a,b}$ are partonic variables, whereas $x_{a,b}$ are leptonic, and the two are only equal at tree level. The inclusive hard function $H_{ij}^{\rm incl}$ can be computed in fixed-order perturbative QCD as the partonic cross section to scatter partons i and j [corresponding to $d\sigma_{ij}^{\rm part}$ in Eq. (1)] and is known to two loops [53–57].

For threshold Drell-Yan, one imposes strong restrictions to only allow soft hadronic final states with $m_X \ll Q$, as illustrated in Fig. 2(b). Using Eq. (7), this can be ensured by forcing $(1-\sqrt{\tau})^2 \ll \tau$, so that one is close to the threshold $\tau \to 1$. In this case, there are large double logarithms that are not accounted for by the parton distributions. Furthermore, since

$$1 \ge \xi_{a,b} \ge x_{a,b} \ge \tau \to 1, \quad (10)$$

a single parton in each proton carries almost all of the energy, $\xi_{a,b} \to 1$. The partonic analog of τ is the variable

$$z = \frac{q^2}{\xi_a\xi_bE_{\rm cm}^2} = \frac{\tau}{\xi_a\xi_b} \le 1, \quad (11)$$

and $\tau \to 1$ implies the partonic threshold limit $z \to 1$. As Eq. (6) forces $Y \to 0$ for $\tau \to 1$, it is convenient to integrate over Y and consider the $\tau \to 1$ limit for $d\sigma/dq^2$. The relevant factorization theorem in this limit is [6,7]

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2} = \sum_{ij}H_{ij}(q^2,\mu)\int\frac{d\xi_a}{\xi_a}\frac{d\xi_b}{\xi_b}f_i(\xi_a,\mu)f_j(\xi_b,\mu) \times QS_{\rm thr}\left[Q\left(1-\frac{\tau}{\xi_a\xi_b}\right),\mu\right] \times\left[1+\mathcal{O}\left(\frac{\Lambda_{\rm QCD}}{Q},1-\tau\right)\right], \quad (12)$$

where we view Eq. (12) as a hadronic factorization theorem in its own right, rather than simply a refactorization of $H_{ij}^{\rm incl}$ in Eq. (9). This Drell-Yan threshold limit has been studied extensively [12,13,58–62]. Factorization theorems of this type are the basis for the resummation of large logarithms in near-threshold situations. In contrast to Eq. (9), the sum in Eq. (12) only includes the dominant $q\bar{q}$ terms for various flavors, $ij = \{u\bar{u}, \bar{u}u, d\bar{d}, \ldots\}$. Other combinations are power-suppressed and only appear at $\mathcal{O}(1-\tau)$ or higher. The threshold hard function $H_{ij} \sim |C_iC_j^*|$ is given by the square of Wilson coefficients in SCET, and can be computed from the timelike quark form factor. The threshold Drell-Yan soft function $S_{\rm thr}$ is defined by a matrix element of Wilson lines and contains both perturbative and nonperturbative physics. If it is treated purely in perturbation theory at the soft scale $Q(1-\tau)$, there are in principle additional power corrections of $\mathcal{O}[\Lambda_{\rm QCD}/Q(1-\tau)]$ in Eq. (12) [63].

Our goal is to describe the isolated Drell-Yan process shown in Fig. 2(c). Here, the colliding partons in the hard interaction are far from threshold as in the inclusive case, but we impose a constraint that does not allow central jets. Soft radiation still occurs everywhere, including the central region. Away from threshold, the hard interaction only carries away a fraction of the total energy in the collision. The majority of the remaining energy stays near the beam. The colliding partons emit collinear radiation along the beams that can be observed in the final state, shown by the green lines labeled "Jet a" and "Jet b" in Fig. 2(c). This radiation cannot be neglected in the factorization theorem and necessitates the beam functions. In the threshold case, these jets are not allowed by the limit $\tau \to 1$, which forces all available energy into the leptons and leaves only soft hadronic radiation.[1] In the inclusive case there are no restrictions on additional hard emissions, in which case initial-state radiation is included in the partonic cross section in $H_{ij}^{\rm incl}$.

Also shown in Fig. 2(c) is the fact that the leptons in isolated Drell-Yan need not be back-to-back, though they are still back-to-back in the transverse plane [see Sec. IV B]. In this regard, isolated Drell-Yan is in-between the threshold case, where the leptons are fully back-to-back with $Y \approx 0$, and the inclusive case, where they are unrestricted.

In Figs. 2(d) and 2(e) we show analogs of threshold Drell-Yan and isolated Drell-Yan where the leptons are replaced by final-state jets. We will discuss the extension to jets in Sec. II G below.

To formulate isolated Drell-Yan we must first discuss how to veto hard emissions in the central region. For this purpose, it is important to use an observable that covers the full phase pace. Jet algorithms are good tools to identify jets, but not necessarily to veto them. Imagine we use a jet algorithm and require that it does not find any jets in the

[1]Note that the proof of factorization for the partonic cross section in the partonic threshold limit $z \to 1$ is not sufficient to establish the factorization of the hadronic cross section, unless one takes the limit $\tau \to 1$. The hadronic factorization theorem assumes that all real radiation is soft with only virtual hard radiation in the hard function. The weaker limit $z \to 1$ still allows the incoming partons to emit energetic real radiation that cannot be described by the threshold soft function. Only the $\tau \to 1$ limit forces the radiation to be soft. This point is not related to whether or not the threshold terms happen to dominate numerically away from $\tau \to 1$ due to the shape of the PDFs or other reasons.

central region. Although this procedure covers the full phase space, the restrictions it imposes on the final state depend in detail on the algorithm and its criteria to decide if something is considered a jet or not. It is very hard to incorporate such restrictions into explicit theoretical calculations, and, in particular, into a rigorous factorization theorem. Even if possible in principle, the resulting beam and soft functions would be very complicated objects, and it would be difficult to systematically resum the large logarithms arising at higher orders from the phase-space restrictions. Therefore, to achieve the best theoretical precision, it is important to implement the central jet veto using an inclusive kinematic variable. This allows us to derive a factorization theorem with analytically manageable ingredients, which can then be used to sum large phase-space logarithms.

We will consider a simple kinematic variable that fulfills the above criteria, leaving the discussion of more sophisticated generalizations to the next subsection. The key variables for the isolated Drell-Yan process are shown in Fig. 3. The proton momenta P_a^μ and P_b^μ are used to define lightlike vectors n_a^μ and n_b^μ,

$$P_a^\mu = \frac{E_{\rm cm}}{2} n_a^\mu, \qquad P_b^\mu = \frac{E_{\rm cm}}{2} n_b^\mu, \tag{13}$$

where the protons are massless and $n_a^2 = 0$, $n_b^2 = 0$, and $n_a \cdot n_b = 2$. Using the beam axis, we define two hemispheres a and b opposite to the incoming protons. We then divide up the total hadronic momentum as

$$p_X^\mu = B_a^\mu + B_b^\mu, \tag{14}$$

where $B_a^\mu = p_{X_a}^\mu$ and $B_b^\mu = p_{X_b}^\mu$ are the total final-state hadronic momenta in hemispheres a and b. Of these, we consider the components

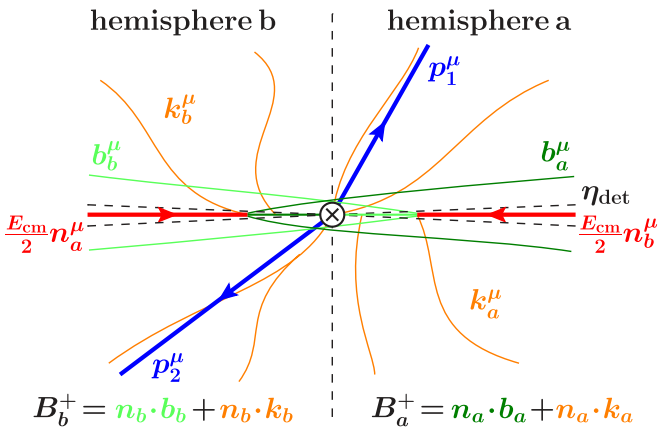


FIG. 3 (color online). Definition of hemispheres and kinematic variables for isolated Drell-Yan.

$$\begin{aligned} B_a^+ &= n_a \cdot B_a = B_a^0 (1 + \tanh y_a) e^{-2y_a}, \\ B_b^+ &= n_b \cdot B_b = B_b^0 (1 + \tanh y_b) e^{-2y_b}, \end{aligned} \tag{15}$$

where $B_{a,b}^0$ are the energy components and $y_{a,b}$ are the total rapidities of $B_{a,b}^\mu$ with respect to the forward direction $n_{a,b}$ for each hemisphere. Here, $\lim_{y\to\infty}(1 + \tanh y) = 2$ and $1 + \tanh y \geq 1.8$ for $y \geq 1$, so $B_{a,b}^+$ scale exponentially with the rapidities $y_{a,b}$.

In terms of the measured particle momenta p_k in hemisphere a,

$$B_a^+ = \sum_{k\in a} n_a \cdot p_k = \sum_{k\in a} E_k (1 + \tanh \eta_k) e^{-2\eta_k}. \tag{16}$$

Here, E_k and η_k are the experimentally measured energy and pseudorapidity with respect to $\vec{n}_a$, and we neglect the masses of final-state hadrons. An analogous formula applies for B_b^+. Hence, B_a^+ and B_b^+ receive large contributions from energetic particles in the central region, while contributions from particles in the forward region are suppressed. Thus, requiring small $B_{a,b}^+ \ll Q$ is an effective way to restrict the energetic radiation in each hemisphere as a smooth function of rapidity, allowing forward jets and disallowing central jets. At the same time, soft radiation with energies $\ll Q$ is measured, but not tightly constrained.

As an example, consider the cut

$$B_{a,b}^+ \leq Q e^{-2y_{\rm cut}}. \tag{17}$$

This constraint vetoes any events with a combined energy deposit of more than $Q/2$ per hemisphere in the central rapidity region $|y| \leq y_{\rm cut}$. In the smaller region $|y| \leq y_{\rm cut} - 1$, the energy allowed by Eq. (17) is reduced by a factor of $e^2 \simeq 7$, essentially vetoing any jets there. In the larger region $|y| \leq y_{\rm cut} + 1$, it is increased by the same factor, so beyond $y_{\rm cut} + 1$ the hadronic final state is essentially unconstrained. Thus, a typical experimental value might be $y_{\rm cut} = 2$, which vetoes energetic jets in the central region $|y| \leq 1$. The precise value of the cut on $B_{a,b}^+$ will of course depend on the requirements of the experimental analyses.

Note that the variable B_a^+ is similar to the total transverse energy in hemisphere a, defined as

$$E_{Ta} = \sum_{k\in a} \frac{E_k}{\cosh \eta_k} = \sum_{k\in a} E_k (1 + \tanh \eta_k) e^{-\eta_k}. \tag{18}$$

B_a^+ has two advantages over E_{Ta}. First, the exponential sensitivity to rapidity is much stronger for B_a^+, which means it provides a stronger restriction on jets in the central region and at the same time is less sensitive to jets in the forward region. Second, since B_a^+ is a specific four-momentum component and linear in four-momentum,

$(p_1 + p_2)^+ = p_1^+ + p_2^+$, it is much simpler to work with and to incorporate into the factorization theorem. It is clear that the isolated Drell-Yan factorization theorem discussed here can be extended to observables with other exponents, $e^{-a\eta_k}$, much like the angularity event shapes in e^+e^- [64].

One should ask, down to what values can $B^+_{a,b}$ be reliably measured experimentally? In principle, particles at any rapidity contribute to $B^+_{a,b}$, but the detectors only have coverage up to a maximum pseudorapidity $\eta_{\rm det}$, as indicated in Fig. 3. For the hadron calorimeters at the LHC $\eta_{\rm det} \simeq 5$ and at the Tevatron $\eta_{\rm det} \simeq 4$. In the hadronic center-of-mass frame, the unscattered partons inside the proton have plus components of $\mathcal{O}(\Lambda^2_{\rm QCD}/E_{\rm cm})$, so any contributions from the unmeasured proton remnants are always negligible. The question then is, what is the maximal contribution to $B^+_{a,b}$ from initial-state radiation that is missed as it is outside the detector? In the extreme scenario where all proton energy is deposited right outside $\eta_{\rm det}$, we would have $B^+_{a,b} = 14\ {\rm TeV} e^{-10} = 0.6\ {\rm GeV}$ at the LHC and $B^+_{a,b} = 2\ {\rm TeV} e^{-8} = 0.7\ {\rm GeV}$ at the Tevatron. In more realistic scenarios, the contribution from such radiation is suppressed by at least another factor of 10 or more. Therefore, the finite detector range is clearly not an issue for measuring values $B^+_{a,b} \gtrsim 2\ {\rm GeV}$, and the relevant limitation will be the experimental resolution in $B^+_{a,b}$.

The factorization theorem for isolated Drell-Yan, which we prove in Sec. IV, reads

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dY dB_a^+ dB_b^+} = \sum_{ij} H_{ij}(q^2,\mu) \int dk_a^+ dk_b^+ q^2 B_i[\omega_a(B_a^+ - k_a^+), x_a, \mu] B_j[\omega_b(B_b^+ - k_b^+), x_b, \mu] S_{\rm ihemi}(k_a^+, k_b^+, \mu) \times \left\{1 + \mathcal{O}\left[\frac{\Lambda_{\rm QCD}}{Q}, \frac{\omega_{a,b} B^+_{a,b}}{Q^2}\right]\right\}. \tag{19}$$

The physical interpretation of Eq. (19) is that we take partons i and j out of the initial-state jets B_i, B_j and hard-scatter them to final-state particles with H_{ij}, while including $S_{\rm ihemi}$ to describe the accompanying soft radiation. The hard function H_{ij} is identical to the one in the threshold factorization theorem in Eq. (12), and the sum in Eq. (19) is again only over $ij = \{u\bar{u}, \bar{u}u, d\bar{d}, \ldots\}$. The quark and antiquark beam functions B_q and $B_{\bar{q}}$ describe the effects of the incoming jets and have replaced the PDFs. The variables $\omega_{a,b} = x_{a,b} E_{\rm cm}$. The hard partons are taken from initial-state jets rather than protons, so unlike in the threshold case the gluon PDF now contributes via the beam functions. We will see how this works in more detail in Sec. II C. Finally, $S_{\rm ihemi}$ is the initial-state hemisphere soft function.

The kinematic variables in Eq. (19) are displayed in Fig. 3. The soft function depends on the momenta $k_a^+ = n_a \cdot k_a$ and $k_b^+ = n_b \cdot k_b$ of soft particles in hemispheres a and b, respectively. Much like PDFs, the beam functions $B_i(t_a, x_a, \mu)$ and $B_j(t_b, x_b, \mu)$ depend on the momentum fractions x_a and x_b of the active partons i and j participating in the hard collision. In addition, they depend on invariant-mass variables

$$t_a = \omega_a b_a^+ \geq 0, \qquad t_b = \omega_b b_b^+ \geq 0, \tag{20}$$

where $\omega_{a,b} = x_{a,b} E_{\rm cm}$ are the hard momentum components and $b_a^+ = n_a \cdot b_a$. The momentum b_a^μ is defined as the total momentum of the energetic particles radiated into hemisphere a, as shown in Fig. 3, and similarly for b_b^+. (The kinematics are shown in more detail in Fig. 13.) Before the hard interaction, the momentum of the active quark can be written as

$$\omega_a \frac{n_a^\mu}{2} - b_a^+ \frac{n_b^\mu}{2} - b_{a\perp}^\mu. \tag{21}$$

The first term is its hard momentum along the proton direction, and the last two terms are from the momentum it lost to radiation, where $b_{a\perp}^2 = -\vec{b}_{aT}^2$ contains the transverse components. The quark's spacelike invariant mass is $-\omega_a b_a^+ - \vec{b}_{aT}^2 = -t_a - \vec{b}_{aT}^2$. The beam function B_i for hemisphere a depends on $t_a = \omega_a b_a^+ = x_a E_{\rm cm} b_a^+$, which is the negative of the quark's transverse virtuality. (When the distinction is unimportant we will usually refer to t simply as the quark's virtuality.) By momentum conservation $b_a^+ = B_a^+ - k_a^+$, leading to the convolution of the beam and soft functions as shown in Eq. (19). Physically, the reason we have to subtract the soft momentum from B_a^+ is that the beam function only properly describes the collinear radiation, while the soft radiation must be described by the soft function. An analogous discussion applies to B_j and t_b for hemisphere b. The convolutions in the factorization theorem thus encode the cross talk between the soft radiation and energetic collinear radiation from the beams.

By measuring and constraining B_a^+ we essentially measure the virtuality of the hard parton in the initial state. As the proton cannot contain partons with virtualities larger than $\Lambda^2_{\rm QCD}$, the initial state at that point must be described as an incoming jet containing the hard off-shell parton. This is the reason why beam functions describing these initial-state jets must appear in Eq. (19). It also follows that since $t \gg \Lambda^2_{\rm QCD}$ we can calculate the beam functions perturbatively in terms of PDFs, which we discuss further in Sec. II C.

It is convenient to consider a cumulant cross section, including all events with $B^+_{a,b}$ up to some specified value, as in Eq. (17). Integrating Eq. (19) over $0 \le B^+_{a,b} \le B^+_{\rm max}$ we obtain

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dY}(B^+_{\rm max}) = \sum_{ij} H_{ij}(q^2,\mu) \int dk^+_a\, dk^+_b\, \tilde{B}_i[\omega_a(B^+_{\rm max}-k^+_a), x_a, \mu]\tilde{B}_j[\omega_b(B^+_{\rm max}-k^+_b), x_b, \mu] S_{\rm ihemi}(k^+_a, k^+_b, \mu) \times \left\{1 + \mathcal{O}\left[\frac{\Lambda_{\rm QCD}}{Q}, \frac{\omega_{a,b} B^+_{\rm max}}{Q^2}\right]\right\}, \tag{22}$$

where the soft function $S_{\rm ihemi}$ is the same as in Eq. (19), and we defined the integrated beam function

$$\tilde{B}_i(t_{\rm max}, x, \mu) = \int dt B_i(t, x, \mu)\theta(t_{\rm max} - t). \tag{23}$$

The cut $B^+_{a,b} \le B^+_{\rm max}$ implies the limit $b^+_{a,b} \le B^+_{\rm max} - k^+_{a,b}$ and $t_{a,b} \le \omega_{a,b}(B^+_{\rm max} - k^+_{a,b})$, leading to the convolutions in Eq. (22).

The factorization theorem Eq. (19) and its integrated version Eq. (22) are valid in the limit $t_{a,b}/Q^2 \simeq B^+_{a,b}/Q \equiv \lambda^2 \ll 1$, and receive power corrections of $\mathcal{O}(\lambda^2)$. Thus, for $B^+_{\rm max} = Qe^{-2y_{\rm cut}}$ with $y_{\rm cut} = 1$, we expect the power corrections not to exceed $e^{-2} \sim 10\%$. This is not a fundamental limitation, because the power corrections can be computed in SCET if necessary. If the soft function is treated purely perturbatively, there are additional power corrections of $\mathcal{O}(\Lambda_{\rm QCD}/B^+_{a,b})$, which account for soft singularities as $B^+_{a,b} \to 0$. The variables $B^+_{a,b}$ are infrared safe with respect to collinear splittings [65].

The hard function receives perturbative α_s corrections at the hard scale $\mu_H \simeq Q$, the beam functions have α_s corrections at the intermediate beam scale $\mu_B^2 \simeq t_{\rm max} \simeq QB^+_{\rm max}$, and the soft function at the soft scale $\mu_S \simeq B^+_{\rm max}$. For example, for $Q \simeq 1$ TeV and $y_{\rm cut} = 2$ we have $\mu_B \simeq 140$ GeV and $\mu_S \simeq 20$ GeV. Even with a very small $Q \simeq 100$ GeV, perhaps for Higgs production, $\mu_B \simeq 14$ GeV and $\mu_S \simeq 2$ GeV are still perturbative (although at this point nonperturbative contributions $\sim\Lambda_{\rm QCD}/\mu_S$ to the soft function might no longer be small and may be incorporated with the methods in Refs. [66,67]). In fixed-order perturbation theory, the cross section contains large single and double logarithms, $\ln(B^+_{\rm max}/Q) \simeq -4$ and $\ln^2(B^+_{\rm max}/Q) \simeq 16$, invalidating a fixed-order perturbative expansion. The factorization theorem allows us to systematically resum these logarithms to all orders in perturbation theory, which is discussed in more detail in Sec. II F.

The factorization theorem Eq. (19) also applies to other nonhadronic final states such as $Z' \to \ell^+\ell^-$, or Higgs production with $H \to \gamma\gamma$ or $H \to ZZ^* \to 4\ell$. In each case, q^2 and Y are the total nonhadronic invariant mass and rapidity, and central jets are vetoed with a cut on $B^+_{a,b}$. The only dependence on the process is in the hard function, which must be replaced appropriately and can be taken directly from the corresponding threshold factorization theorem. One may also consider W production with $W \to \ell\bar{\nu}$, with an appropriate replacement of q^2 and Y with the charged lepton's rapidity. For a light Higgs with $Q \sim m_H$, the isolated Drell-Yan factorization theorem applies to Higgs production through gluon fusion $gg \to H$ and Higgs-strahlung $q\bar{q} \to VH$, which are the dominant production channels at the LHC and Tevatron, respectively.[2] For a generic process $pp \to XL$, the sum over $ij = \{gg, u\bar{u}, \bar{u}u, d\bar{d}, \ldots\}$ includes a gluon-gluon contribution, but still no cross terms between different parton types, and there will be two independent soft functions $S^{q\bar{q}}_{\rm ihemi}$ and $S^{gg}_{\rm ihemi}$. [As shown in Sec. IV, only the $q\bar{q}$ soft function contributes to isolated Drell-Yan, so the labels were omitted in Eq. (19).] Indeed, the gluon-gluon contribution involving the gluon beam and soft functions, B_g and $S^{gg}_{\rm ihemi}$, gives the dominant contribution in the case of Higgs production.

With the above physical picture, we can understand why the gluon beam function appeared in $\gamma p \to J/\psi X$ in the analysis of Ref. [30] in the limit where $E_{J/\psi} \to E_\gamma$. Taking p_X as the total momentum of final-state hadrons other than the J/ψ, one has $n \cdot p_X \sim E_{\rm cm}(1 - E_{J/\psi}/E_\gamma)$, where n is the proton direction. For $E_{J/\psi}$ close to E_γ, energetic radiation in the final state is restricted to a jet close to the n direction. Just as for our $B^+_{a,b}$, the measurement of $E_{J/\psi}$ probes the radiation emitted by the colliding gluon in the initial state. Thus, the proton is broken apart prior to the hard collision, and the gluon beam function is required to describe the initial state.

B. Generalized observables

The factorization theorem in Eq. (19) applies for $t_a \ll q^2$ and $t_b \ll q^2$. This includes the situation where in the hadronic center-of-mass frame there is a numerically significant asymmetry $\omega_a = x_a E_{\rm cm} > \omega_b = x_b E_{\rm cm}$. This means that the boost between the hadronic and partonic center-of-mass frames, given by the leptonic $Y = \ln\sqrt{\omega_a/\omega_b} = \ln\sqrt{x_a/x_b}$, is significantly different from zero. We explore the implications of this here.

If there is no hierarchy, $\omega_a \approx \omega_b \approx \sqrt{\omega_a\omega_b} = Q$, corresponding to $Y \approx 0$, we can define a simple variable to constrain both hemispheres simultaneously,

[2] In vector-boson fusion and associated production $gg \to t\bar{t}H$, the situation is more complicated and one has to explicitly consider the process $pp \to Xjjh$ with two forward (top) jets.

$$\hat{B} = \frac{B_a^+ + B_b^+}{Q}. \tag{24}$$

From Eq. (19), this gives

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dY d\hat{B}} = \sum_{ij} H_{ij}(q^2,\mu)\int dt_a dt_b B_i(t_a,x_a,\mu) \times B_j(t_b,x_b,\mu) Q S_B\Big(Q\hat{B} - \frac{t_a}{\omega_a} - \frac{t_b}{\omega_b},\mu\Big), \tag{25}$$

where the soft function is defined as

$$S_B(k^+,\mu) = \int dk_a^+ dk_b^+ S_{\rm ihemi}(k_a^+,k_b^+,\mu)\delta(k^+ - k_a^+ - k_b^+). \tag{26}$$

The advantage of using $\hat{B}$ is that the soft function now only depends on the single variable $k^+ = k_a^+ + k_b^+$, much like the soft function for thrust in e^+e^- collisions.

If we have a hierarchy $\omega_b < Q < \omega_a$, the final state has a substantial boost in the n_a direction, as shown in Fig. 4. In this case, the energetic radiation will generically be much closer to the beam axis in hemisphere a than in hemisphere b. To take this into account, it is natural to impose different cuts on B_a^+ and B_b^+. Using the boost-invariant combinations $\omega_a B_a^+/q^2$ and $\omega_b B_b^+/q^2$ to define the cut, we obtain

$$\frac{\omega_a B_a^+}{q^2} = \frac{B_a^+}{\omega_b} \le e^{-2y_{\rm cut}}, \qquad \frac{\omega_b B_b^+}{q^2} = \frac{B_b^+}{\omega_a} \le e^{-2y_{\rm cut}}, \tag{27}$$

so B_a^+ has a tighter constraint than B_b^+, as desired. If we simply replace $\hat{B}$ by $B_a^+/\omega_b + B_b^+/\omega_a$, the soft function analogous to S_B in Eq. (26) will depend on the combination $(\omega_a k_a^+ + \omega_b k_b^+)/Q^2$.

However, we should also adjust the hemispheres themselves to take into account the significant boost of the partonic center-of-mass frame. We therefore define a generalized hemisphere a as $y > Y$ and hemisphere b as $y < Y$, as shown in Fig. 4. The corresponding total hemisphere momenta are denoted as $B_{a,b}^+(Y)$ and the soft hemisphere momenta as $k_{a,b}^+(Y)$. The original definitions in Fig. 3 correspond to $B_{a,b}^+(0) \equiv B_{a,b}^+$ and $k_{a,b}^+(0) \equiv k_{a,b}^+$. The generalization of $\hat{B}$ is given by the boost-invariant combination

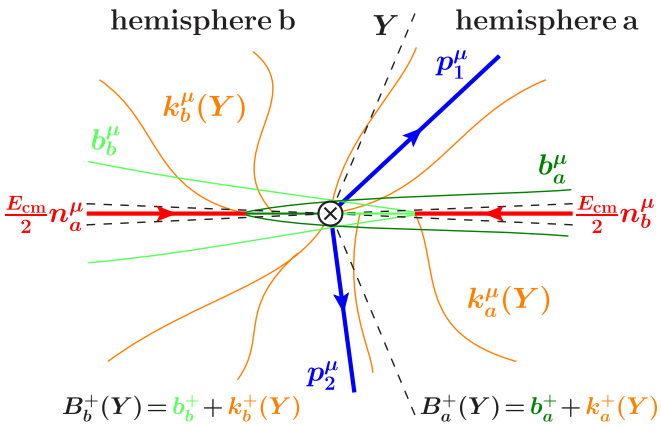


FIG. 4 (color online). Generalized definition of hemispheres. The total rapidity of the leptons is Y, $b_{a,b}^+ = n_{a,b}\cdot b_{a,b}$, and $k_{a,b}^+(Y) = n_{a,b}\cdot k_{a,b}(Y)$.

$$\tau_B = \frac{\omega_a B_a^+(Y) + \omega_b B_b^+(Y)}{q^2}. \tag{28}$$

With the generalized definition of the hemispheres, $B_{a,b}^+(Y)$ and $\omega_{a,b}$ transform under a boost by y in the n_a direction as

$$\begin{aligned} B_a^+(Y) &\to B_a^{+\prime}(Y+y) = e^{-y}B_a^+(Y),\\ B_b^+(Y) &\to B_b^{+\prime}(Y+y) = e^{y}B_b^+(Y),\\ \omega_a &\to \omega_a' = e^y\omega_a,\\ \omega_b &\to \omega_b' = e^{-y}\omega_b. \end{aligned} \tag{29}$$

Thus, boosting by $y = -Y$ from the hadronic to the partonic center-of-mass frame gives

$$\tau_B = \frac{\omega_a' B_a^{+\prime}(0) + \omega_b' B_b^{+\prime}(0)}{q^2} = \frac{B_a^{+\prime}(0) + B_b^{+\prime}(0)}{Q}. \tag{30}$$

In the partonic center-of-mass frame we have $\omega_a' = \omega_b' = Q$, so there is no hierarchy. Correspondingly, the generalized hemispheres in this frame are again perpendicular to the beam axis, so Eq. (30) has the same form as $\hat{B}$.

Note that for $e^+e^- \to$ jets, one can use the thrust axis to define two hemispheres with $n_{a,b}$ analogous to our case. In the 2-jet limit, thrust is then given by $1 - T = (Qn_a\cdot p_{X_a} + Qn_b\cdot p_{X_b})/Q^2$. Hence, we can think of τ_B as the analog of thrust for incoming jets. For this reason we will call τ_B the "beam thrust".

In analogy to Eqs. (17) and (27), we define the cutoff on τ_B by

$$\tau_B \le e^{-2y_B^{\rm cut}}. \tag{31}$$

For $\tau_B \to 0$ or equivalently $y_B^{\rm cut} \to \infty$ the jets along the beam axes become pencil-like, while for generic $y_B^{\rm cut}$ we allow energetic particles up to rapidities $y \lesssim y_B^{\rm cut}$ (with y measured in the partonic center-of-mass frame).

The beam functions are boost-invariant along the beam axis, so the different hemisphere definitions do not affect them. The soft function is boost-invariant up to the hemisphere definition, which defines its arguments $k_{a,b}^+$. Hence, boosting by $-Y$ we have $S_{\rm ihemi}[e^Y k_a^+, e^{-Y}k_b^+; Y] = S_{\rm ihemi}[k_a^+, k_b^+; 0] = S_{\rm ihemi}(k_a^+, k_b^+)$, where the third argument denotes the definition of the hemispheres. This implies that the soft function for τ_B is the same as in Eq. (26). The factorization theorem for τ_B following from Eq. (19) is

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dY d\tau_B} = \sum_{ij} H_{ij}(q^2,\mu)\int dt_a dt_b B_i(t_a,x_a,\mu) \times B_j(t_b,x_b,\mu) Q S_B\Big(Q\tau_B - \frac{t_a+t_b}{Q},\mu\Big). \tag{32}$$

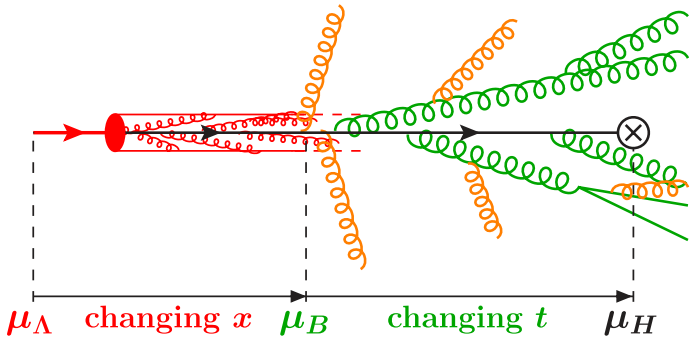


FIG. 5 (color online). Evolution of the initial state. Starting from the low scale μ_Λ, the incoming proton is described by the x-dependent evolution of the PDFs, which redistributes the total momentum of the proton between its constituents. At the scale μ_B, the proton is probed by measuring the radiation in the final state and breaks apart. This is the scale where the PDFs are evaluated and the x-dependent evolution stops. Above μ_B, the proton has ceased to exist, and the initial state behaves like an incoming jet, whose evolution is governed by the virtuality t of the off-shell spacelike parton that eventually enters the hard interaction at the scale μ_H.

Integrating over $0 \le \tau_B \le \exp(-2y_B^{\rm cut})$ we obtain

$$\frac{d\sigma}{dq^2 dY}(y_B^{\rm cut}) = \int_0^{\exp(-2y_B^{\rm cut})} d\tau_B \frac{d\sigma}{dq^2 dY d\tau_B}. \quad (33)$$

We will use Eqs. (32) and (33) to show plots of our results in Sec. V.

C. Relating beam functions and PDFs

The beam functions can be related to the PDFs by performing an operator product expansion, because $t_{a,b} \gg \Lambda^2_{\rm QCD}$.[3] This yields the factorization formula

$$B_i(t, x, \mu) = \sum_j \int_x^1 \frac{d\xi}{\xi} \mathcal{I}_{ij}\Big(t, \frac{x}{\xi}, \mu\Big) f_j(\xi, \mu) \times \Big[1 + \mathcal{O}\Big(\frac{\Lambda^2_{\rm QCD}}{t}\Big)\Big], \quad (34)$$

where we sum over partons $j = \{g, u, \bar{u}, d, \ldots\}$, $\mathcal{I}_{ij}$ are perturbatively calculable Wilson coefficients, and f_j is the standard PDF for parton j. The $\mathcal{O}(\Lambda^2_{\rm QCD}/t)$ power corrections in Eq. (34) involve proton structure functions at subleading twist. Further mathematical details on Eq. (34) are discussed in Sec. III, whereas here we focus on the physical ramifications.

The interpretation of Eq. (34) is illustrated in Fig. 5. At a hadronic scale $\mu_\Lambda \sim 1$ GeV, the initial conditions for the PDFs f_j can be specified, and one has the standard DGLAP evolution up to the scale μ_B,

$$\mu \frac{d}{d\mu} f_j(\xi, \mu) = \sum_{j'} \int \frac{d\xi'}{\xi'} P_{jj'}\Big(\frac{\xi}{\xi'}, \mu\Big) f_{j'}(\xi', \mu). \quad (35)$$

[3]A detailed discussion of the appropriate operator product expansion is given in Ref. [50].

The anomalous dimensions $P_{jj'}$ are the standard QCD splitting functions for quarks, antiquarks, and gluons (including the color factors and coupling constant). Equation (34) applies at the scale $\mu = \mu_B$, since this is the scale at which a measurement on the proton is performed by observing the soft and collinear radiation contributing to $B^+_{a,b}$. At this scale, a parton j with momentum fraction ξ is taken out of the incoming proton according to the probability distribution $f_j(\xi, \mu)$. As the parton continues to propagate and evolve with $\mu > \mu_B$, it is modified by virtual radiation and by the emission of real radiation, which forms a jet. The evolution in this region no longer depends on ξ, but instead on the virtuality t. This evolution occurs with fixed x and fixed parton type i, via the beam-function RGE

$$\mu \frac{d}{d\mu} B_i(t, x, \mu) = \int dt' \gamma_B^i(t - t', \mu) B_i(t', x, \mu). \quad (36)$$

This result for initial-state jet evolution has the same structure as the evolution for final-state jets. In fact, the anomalous dimension γ_B^q is identical to that for the quark jet function to all orders in perturbation theory [50]. We discuss this correspondence further in Sec. III.

The effect of initial-state real and virtual radiation is described by the perturbatively calculable Wilson coefficients $\mathcal{I}_{ij}(t, x/\xi, \mu)$ at the scale $\mu = \mu_B$. They encode several physical effects. The virtual loop corrections contribute to the $\mathcal{I}_{ii}$ and modify the effective strength of the various partons. If the radiation is real, it has physical timelike momentum. Hence, it pushes the active parton in the jet off shell with spacelike virtuality $-t < 0$ and reduces its light-cone momentum fraction from ξ to x.

In addition, the real radiation can change the identity of the colliding parton, giving rise to the sum over j in Eq. (34). For example, an incoming quark can radiate an

energetic gluon which enters the hard interaction, while the quark itself goes into the final state. This gives a contribution of the quark PDF to the gluon beam function through $\mathcal{I}_{gq}$. Similarly, an incoming gluon can pair-produce, with the quark participating in the hard interaction and the antiquark going into the final state. This gives a contribution of the gluon PDF to the quark beam function through $\mathcal{I}_{qg}$. There are also of course real radiation contributions to the diagonal terms, $\mathcal{I}_{qq}$ and $\mathcal{I}_{gg}$, where the parton in the PDF and the parton participating in the hard interaction have the same identity.

At lowest order in perturbation theory, the parton taken out of the proton directly enters the hard interaction without emitting radiation,

$$\mathcal{I}_{ij}^{\text{tree}}\left(t, \frac{x}{\xi}, \mu\right) = \delta_{ij}\delta(t)\delta\left(1 - \frac{x}{\xi}\right). \tag{37}$$

Thus at tree level, the beam function reduces to the PDF

$$B_i^{\text{tree}}(t, x, \mu) = \delta(t) f_i(x, \mu). \tag{38}$$

Beyond tree level, $\mathcal{I}_{ij}(t, x/\xi, \mu)$ can be determined perturbatively as discussed in more detail in Sec. III, where we give precise field-theoretic definitions of the beam functions and quote the one-loop results for $\mathcal{I}_{qq}$ and $\mathcal{I}_{qg}$.

Interestingly, in the threshold factorization theorem Eq. (12), cross terms between quark and gluon PDFs are power suppressed, so the gluon PDF does not contribute at leading order. In the inclusive case Eq. (9), such cross terms are leading order in the power counting. For isolated Drell-Yan, there are no cross terms between quark and gluon beam functions, but there are leading-order cross terms between different PDFs, which appear via the contributions of different PDFs to a given beam function in Eq. (34). Thus, the isolated case is again in-between the inclusive and threshold cases.

D. Comparison with initial-state parton shower

The physical situation associated with the beam evolution has an interesting correspondence with that of initial-state parton showers. As pictured in the region between μ_B and μ_H in Fig. 5, the parton in the beam function evolves forward in time while emitting a shower of radiation into the final state governed by the anomalous dimension $\gamma_B^i(t - t', \mu)$ in Eq. (36). This equation has no parton mixing. Each emission by the radiating parton increases the magnitude of its spacelike virtuality $-t < 0$, pushing it further off-shell in a spacelike direction. At the time the parton is annihilated in the hard collision, it has evolved to some t with $|t| \ll q^2$, so the large momentum transfer q^2 guarantees that no partons in the final state are spacelike. This description agrees quite well with the physical picture associated with the evolution of the primary parton in an initial-state parton shower, as summarized in Ref. [36].

Differences in the description arise when one considers the initial-state parton shower in more detail (for simplicity we focus on the so-called longitudinal evolution). The shower is based on the evolution equation for the PDFs in Eq. (35). An evolution forward in time is not practical because of the lack of prior knowledge of the scale of the hard interaction, so the shower uses backward evolution starting at a given partonic hard scale Q [47]. Knowing the identity of the final parton i, the shower evolves based on the probability $d\mathcal{P}_i/dt$ that parton i is unresolved into parton j via the splitting $j \to ij'$ at an earlier (lower) scale t. The evolution equation is [36]

$$\frac{d\mathcal{P}_i(x, t_{\text{max}}, t)}{dt} = \left[\sum_{jj'} \int_x^{z_{\text{max}}} \frac{dz}{z} P_{j\to ij'}(z, t) \frac{f_j(x/z, t)}{f_i(x, t)}\right] \times \frac{1}{t} \mathcal{P}_i(x, t_{\text{max}}, t), \tag{39}$$

where $\mathcal{P}_i(x, t_{\text{max}}, t)$ is the shower Sudakov exponential, which is interpreted as the probability for no emissions to occur between the initial value t_{max} and t. The evolution variable t, which determines the scale of the splitting, is usually chosen as the virtuality or transverse momentum of the parton.

The mixing of partons in the PDF evolution influences the shower. In particular, the evolution kernel depends on the PDF $f_j(x/z, t)$, which determines the number density of partons of type j at the scale t, and inversely on the PDF $f_i(x, t)$. Thus, unlike in the beam evolution in Eq. (36), the shower evolution in Eq. (39) still knows the identity of the initial-state hadron. Double logarithms in the initial-state parton shower are generated in $q \to qg$ and $g \to gg$ splittings because of the soft-gluon singularity $\sim 1/(1 - z)$ in the splitting functions. This singularity is regulated [36] by the upper cutoff $z_{\text{max}} = x/(x + x_\epsilon)$, where x_ϵ provides a lower cutoff on the gluon energy in the rest frame of the hard scattering, $E_g \geq x_\epsilon \gamma E_{\text{cm}}/2 \simeq 2$ GeV (where γ is the boost factor of the hard scattering). Hence, one logarithm, $\ln x_\epsilon$, is generated by the z integration, and one logarithm, $\ln t$, by the collinear $1/t$ singularity. In contrast, the beam function contains double logarithms $\ln^2 t$ similar to a final-state parton shower, where the z integration yields a kernel $\sim(\ln t)/t$ that produces a double logarithm $\ln^2 t$ via the t evolution.

The above comparison is very rough. For example, the influence of soft radiation on both the shower and on the isolated factorization theorem was not compared and is likely to be important. Furthermore, the goal of the shower is to provide a universal method for populating fully exclusive final states, while the beam function applies for a more inclusive situation with a particular measurement. Note that just the presence of mixing in the initial-state

parton shower and absence of mixing in the beam-function evolution does not imply an inconsistency. For example, it is well-known that the final-state parton shower reproduces the correct double logarithms for e^+e^- event shapes [68], even though there is no parton mixing in the evolution of the corresponding hard, jet, and soft functions. In the future it would be interesting to test in detail the correspondence between the double logarithms generated by the initial-state parton shower and those predicted by our factorization theorem for the isolated Drell-Yan process.

E. Relation to fixed-order calculation

The factorization theorem for the cross section in Eq. (19) and the factorization for the beam function in Eq. (34) together allow us to describe in more detail how various Feynman diagrams that would appear in a fixed-order calculation contribute to the cross section in our kinematic region. Various examples are shown in Fig. 6.

In Fig. 6(a), we have the tree-level $q\bar{q}$ annihilation producing a γ or Z, which involves the tree-level $\mathcal{O}(\alpha_s^0)$ hard function, beam functions, and soft function, denoted

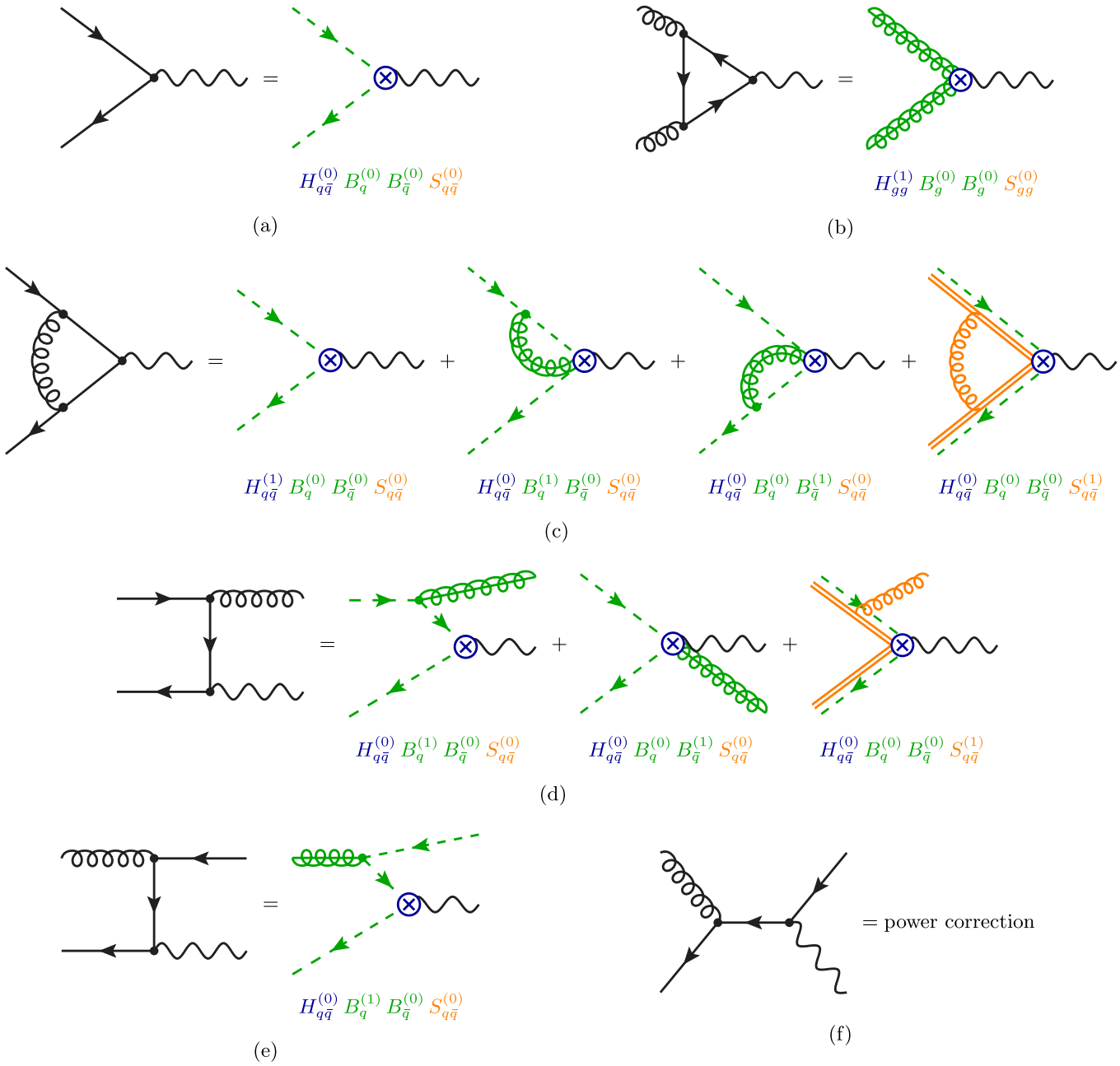


FIG. 6 (color online). Factorization for isolated Drell-Yan in pictures. The left-hand side of each equality are graphs in QCD, while the right-hand side shows the sum of the corresponding SCET diagrams. Dashed lines are collinear quarks, and springs with a line through them are collinear gluons. The double lines denote soft Wilson lines, and the gluons attached to them are soft.

by a superscript (0) in the figure. In Fig. 6(b), initial-state gluons couple to a quark loop (e.g. a top quark), which subsequently annihilates into a γ, Z, or Higgs. The quarks in this loop are far off shell, so they can be integrated out and appear as one-loop corrections, $H_{gg}^{(1)}$, to the hard coefficient in the factorization theorem. Other possibilities for this graph are power suppressed.

The situation for the vertex correction in Fig. 6(c) is more involved. If the gluon in the loop is hard, all particles in the loop are far off shell and can be integrated out, giving the one-loop hard function $H_{q\bar{q}}^{(1)}$ shown as the first term on the right-hand side. In the second term, the gluon is collinear to the incoming quark beam and gives a virtual one-loop contribution to the quark beam function, $B_q^{(1)}$. The third term is the analog of the second, but now with the gluon collinear to the incoming antiquark. Finally in the fourth term, the gluon is soft, communicating between the incoming collinear beams. Here, the eikonal approximation holds for describing the quark propagators. The generalization of this to all orders in α_s leads to the fact that the soft function is a matrix element of Wilson lines. Although a single loop graph contributes in several different places in the factorization theorem, all of these contributions have a precise separation in SCET. We will use this separation in Sec. IV to prove the isolated Drell-Yan factorization theorem.

An interesting contribution occurs in Fig. 6(d), where a gluon is radiated into the final state. Because of the kinematic restrictions in isolated Drell-Yan, this gluon can only be collinear to the incoming quark, collinear to the incoming antiquark, or soft, and these three possibilities are represented by the diagrams on the right-hand side of the equality. In the first case, we have a real-emission correction to the quark beam function, $B_q^{(1)}$. In the second case, the intermediate quark is far off shell and can be integrated out, and the gluon collinear to the antiquark arises from a collinear Wilson line contribution in $B_{\bar{q}}^{(1)}$. The third case gives a real-emission correction to the soft function, $S_{q\bar{q}}^{(1)}$. The full-theory graph in Fig. 6(d) has a t-channel singularity. An important fact about the isolated Drell-Yan factorization theorem is that it fully captures the dominant parts of this singularity, and allows a simple framework for a resummation of higher order α_s corrections enhanced by large double logarithms due to this singularity. For threshold Drell-Yan, the kinematic restrictions are stronger and only allow the third graph with soft initial-state radiation. In inclusive Drell-Yan, the gluon is treated as hard, and the graph in Fig. 6(d) only corrects $H_{q\bar{q}}^{\text{incl}}$, without providing a framework for summing the large double logarithms that appear when we make a global measurement of the radiation in each hemisphere defined by the beams.

The situation is a bit simpler for Figs. 6(e) and 6(f). In Fig. 6(e), the incoming collinear gluon from the PDF pair-produces a quark and antiquark both collinear to this beam direction, and the quark enters the hard interaction. Therefore, this is a one-loop correction to the quark beam function, $B_q^{(1)}$, proportional to the gluon PDF f_g. The beam functions again allow us to resum the possibly large logarithms due to this t-channel singularity. Other possibilities for the final-state antiquark in Fig. 6(e) lead to power-suppressed contributions. Similarly, the s-channel graph in Fig. 6(f), which has the same initial and final states as Fig. 6(e), has no leading-power contribution and only contributes to Eq. (19) in the power-suppressed terms. The same is also true for Drell-Yan in the threshold region. Only inclusive Drell-Yan receives a leading-order hard contribution from the s-channel graph, which is then treated as of the same size as the t-channel graphs.

F. Renormalization group evolution

In this subsection, we discuss and compare the structure of large logarithms in the cross sections for inclusive, threshold, and isolated Drell-Yan. These large logarithms may be summed using the renormalization group evolution of the individual functions appearing in the factorization theorems. In fact, the structure of large logarithms in the differential $B_{a,b}^{+}$ cross section allows us to infer the necessity of the beam functions in the isolated factorization theorem. This procedure provides a method of determining whether beam functions enter for other observables or processes than those studied here. The consistency of the RGE was used to provide a similar consistency check in Ref. [69] when deriving a new factorization theorem for the invariant-mass distribution of jets initiated by a massive quark in e^+e^- collisions. In that case, the RGE consistency provided important constraints on the structure of the factorization theorem at scales below the heavy-quark mass.

In inclusive Drell-Yan, the hard functions H_{ij}^{incl} are sensitive to the scale $\mu_H \simeq Q$ of the hard interaction, and the proton mass defines a low scale $\mu_\Lambda \simeq 1$ GeV $\gtrsim \Lambda_{\text{QCD}}$ (which is still large enough so perturbation theory can be applied for the PDF evolution). The measurement of q^2 and Y in this case does not introduce additional scales, and thus does not influence the structure of the logarithms. Thus, we have the hierarchy $\mu_\Lambda \ll \mu_H$, and the large logarithms are $L = \ln(\mu_\Lambda/\mu_H)$. Here, only single-logarithmic series, $(\alpha_s L)^k$, are generated at higher orders in perturbation theory. The logarithms are factorized as $\ln(\mu/\mu_H) + \ln(\mu_\Lambda/\mu)$ in the factorization theorem in Eq. (9) and may then be resummed. The general form of the running is pictured in Fig. 7(a). The logarithms $\ln(\mu_\Lambda/\mu)$ are summed by evolving the PDFs $f_i(\xi_a, \mu)$ and $f_j(\xi_b, \mu)$ from μ_Λ up to the common scale μ. The inclusive hard function, $H^{\text{incl}}(x_a/\xi_a, x_b/\xi_b, q^2, \mu)$, is evolved from μ_H down to μ, summing the logarithms $\ln(\mu/\mu_H)$. The choice of μ is arbitrary. Taking $\mu \simeq \mu_H$ corresponds to only running the PDFs up, while for $\mu \simeq \mu_\Lambda$ only H^{incl} runs

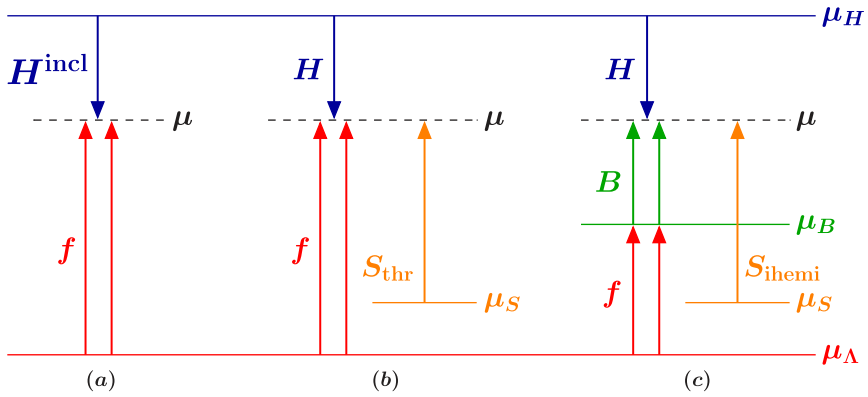


FIG. 7 (color online). RGE running for different Drell-Yan scenarios. Case (a) corresponds to the inclusive case. Case (b) corresponds to the threshold case, where the kinematics forces all hadrons in the final state to be soft. Case (c) corresponds to the isolated case. Here, the PDFs freeze out at the intermediate beam scale μ_B, above which they are replaced by beam functions.

down. The equivalence of these two choices implies that $H^{\rm incl}$ must be convoluted with the two PDFs and exhibit a factorized structure for logarithms in the a and b variables.

For threshold Drell-Yan, the kinematic restrictions only allow soft radiation in the final state. This induces additional large logarithms $\ln(1-\tau)$. These can be written in terms of a ratio of scales $\ln(\mu_S/\mu_H)$, where the soft scale $\mu_S \simeq Q(1-\tau)$ is another important scale in the analysis. The logarithms $L = \ln(\mu_S/\mu_H)$ appear as double-logarithmic series $(\alpha_s L^2)^k$ in the cross section. In the threshold factorization theorem in Eq. (12), these double logarithms can be summed by evolving the PDFs and the threshold soft and hard functions, $S_{\rm thr}$ and H, to a common scale μ, as shown in Fig. 7(b). Since $\xi_{a,b} \to 1$, the logarithms $\ln(1-\xi_a)$ and $\ln(1-\xi_b)$ are also large. The RGE for the PDFs must be expanded, and the result sums a double-logarithmic series of $\ln^2(1-\xi)$ terms. The threshold soft function sums double logarithms $\ln^2(\mu/\mu_S)$ between μ_S and μ, while the threshold hard function sums double logarithms $\ln^2(\mu/\mu_H)$ between μ_H and μ. The evolution equations are

$$\begin{aligned}
\mu\frac{d}{d\mu}H(q^2,\mu) &= \gamma_H(q^2,\mu)H(q^2,\mu),\\
\mu\frac{d}{d\mu}f_i(\xi,\mu) &= \int\frac{d\xi'}{\xi'}P_{ii}^{\rm expanded}\Big(\frac{\xi}{\xi'},\mu\Big)f_i(\xi',\mu),\\
\mu\frac{d}{d\mu}S_{\rm thr}(k,\mu) &= \int dk_s'\,\gamma_{S_{\rm thr}}(k-k',\mu)S_{\rm thr}(k',\mu).
\end{aligned}\tag{40}$$

The consistency of the RGE at the scale μ shown in Fig. 7(b) implies that the double logarithms in f_i, f_j, and $S_{\rm thr}$ combine in such a way that the RGE of the convolution $f_i f_j \otimes S_{\rm thr}$ is identical to that of H, and hence only depends on q^2.

For isolated Drell-Yan, the kinematic restrictions allow both soft and collinear initial-state radiation, and induce an invariant-mass scale for each beam function, $\mu_B^2 \simeq x_a E_{\rm cm} B_a^+$ and $\mu_B^2 \simeq x_b E_{\rm cm} B_b^+$, and a soft scale $\mu_S \simeq B_{a,b}^+$. For simplicity, we use a common scale μ_B for both beam functions in our discussion here. (Since the evolution of the two beam functions is independent, one can just as easily implement two independent beam scales.) As we saw in Sec. II A, at partonic center-of-mass energies of a hundred GeV to a few TeV there is a large hierarchy between the different scales, $\mu_\Lambda \ll \mu_S \ll \mu_B \ll \mu_H$, and correspondingly large double and single logarithms of the ratios of these scales. The RGE running for this case is shown in Fig. 7(c). Here, the PDFs are not restricted to their endpoints, so their evolution is given by Eq. (35), which involves the unexpanded and nondiagonal $P_{ij}(\xi/\xi')$ and sums single logarithms, $(\alpha_s L)^k$. For each f_j this evolution joins at $\mu = \mu_B$ with the Wilson coefficients $\mathcal{I}_{ij}$ in the beam-function factorization $B_i = \mathcal{I}_{ij} \otimes f_j$ of Eq. (34). The $\mathcal{I}_{ij}$ cancel the ξ-dependent evolution of f_j, and turn it into the t-dependent evolution of B_i, which sums a double-logarithmic series. The objects meeting at the common scale μ in Fig. 7(c) are the hard function, which is identical to the threshold case in Eq. (40), and the beam and soft functions,

$$\begin{aligned}
\mu\frac{d}{d\mu}B_i(t,x,\mu) &= \int dt'\gamma_B^i(t-t',\mu)B_i(t',x,\mu),\\
\mu\frac{d}{d\mu}S_{\rm ihemi}(k_a^+,k_b^+,\mu) &= \int dk_a' dk_b' S_{\rm ihemi}(k_a',k_b',\mu)\\
&\quad\times\gamma_{S_{\rm ihemi}}(k_a^+-k_a',k_b^+-k_b',\mu).
\end{aligned}\tag{41}$$

The consistency of the RGE at μ now implies that the double-logarithmic running in the different variables for B_i, B_j, and $S_{\rm ihemi}$ cancels such that the convolution $B_i B_j \otimes S_{\rm ihemi}$ has an RGE identical to H, which only depends on q^2. (A detailed discussion of this consistency can be found in Ref. [70] for the analogous case of two jet functions and

the final-state hemisphere soft function, $JJ \otimes S_{\rm hemi}$, and in Ref. [50] for the case discussed here.) It is important that this cancellation would not be possible if we tried to replace B_i by f_i in the isolated factorization theorem. Given the type of double logarithms in the cross section, the single logarithms summed by the PDFs at generic x cannot combine with the double logarithms in $S_{\rm ihemi}$ to give a result in agreement with the double logarithms in H. Thus, the structure of double logarithms necessitates the presence of beam functions in the isolated factorization theorem.

By the same argument we can conclude that for all processes involving a threshold-type hard function H with double logarithms, and with $x_{a,b}$ away from one, the description of the initial-state radiation will require beam functions B_i. This includes all situations where H is the square of Wilson coefficients of SCET operators, $H = \sum_k |C_k|^2$ (for example when the energetic partons in the hard collision all have distinct collinear directions). In particular, the theoretical description of any threshold process with $x \to 1$ can be extended to a factorization theorem for the respective isolated case with x away from one. This is achieved by adding variables $B^+_{a,b}$, replacing the PDFs by beam functions, and replacing the threshold soft function by an appropriate soft function for the isolated case.

Thus, beam functions are quite prevalent for cross sections that one may wish to study at the LHC. In situations where the hadronic final state is constrained with variables that are more complicated than $B^+_{a,b}$, one generically expects to find different beam functions and different soft functions encoding these constraints. This extension is analogous to how the choice of jet algorithm modifies the definition of the jet and soft functions for central jets produced by the hard collision [17]. Even with this generalization, the beam and soft functions will both sum double-logarithmic series, and we expect that the factorization relating the beam function to the PDFs will carry through, just with different coefficients $\mathcal{I}_{ij}$.

G. Extension to final-state jets

In Fig. 2, where we show the types of hadronic final states for inclusive, endpoint, and isolated Drell-Yan, we also included analogs where the lepton pair is replaced by two jets. Figure 2(d) shows the threshold dijet production process studied in Ref. [8], which is the generalization of the Drell-Yan threshold process in Fig. 2(b). Figure 2(e) shows the isolated dijet production process, which is the generalization of isolated Drell-Yan in Fig. 2(c). The goal of this subsection is to give a rough idea of how the isolated factorization theorem will look when it is extended to include final-state jets. Recall that our proof of factorization is only for the Drell-Yan case. We stress that the factorization formula for the isolated dijet case discussed here expresses our expectations and has not been rigorously derived.

Final-state jets are identified by a jet algorithm as more-or-less isolated groups of energetic particles within a cone[4] of some radius $R = [(\Delta\phi)^2 + (\Delta\eta)^2]^{1/2}$. For a dijet event, the jet algorithm allows us to define the total jet momenta P_1^μ and P_2^μ.[5] Given these, we let y_1 and y_2 be the rapidities of the two jets relative to the beam axis and define $\Delta y = y_1 - y_2$. The invariant masses of the jets are denoted as $M^2_{1,2} = P^2_{1,2}$. Two analogs for the q^2-variable of Drell-Yan are

$$\begin{aligned} M^2_{JJ} &= (P_1 + P_2)^2, \\ m^2_{JJ} &= 2P_1 \cdot P_2 = M^2_{JJ} - M^2_1 - M^2_2, \end{aligned} \tag{42}$$

where M^2_{JJ} is the total invariant mass of the two jets, and m^2_{JJ} is their total invariant mass minus their individual invariant masses. The corresponding analogs of the Drell-Yan τ-variable are then

$$\tau_J = \frac{M^2_{JJ}}{E^2_{\rm cm}}, \qquad \Delta\tau_J = \frac{m^2_{JJ}}{E^2_{\rm cm}}. \tag{43}$$

Identifying the two jets already restricts the hadronic final state, which means there is no analog of the inclusive Drell-Yan factorization theorem. In the threshold case, we take the limit $\tau_J \to 1$, which ensures that the final state consists of two back-to-back jets plus soft radiation and no additional energetic jets, as shown in Fig. 2(d). The limit $\Delta\tau_J \to 1$ is even more restrictive, since it also forces the two jets to have very small invariant masses, essentially behaving like massless particles for the factorization theorem.

The threshold factorization for m^2_{JJ} was considered in Ref. [8],[6,7]

[4]Here ϕ and η denote the azimuthal angle and pseudorapidity. Although our notation corresponds to a cone algorithm, at the level of our discussion we may equally well substitute a k_T algorithm.

[5]We take $P^\mu_{1,2}$ to be the momenta of the two hardest jets found by the jet algorithm. We will see below that the final-state restrictions considered here eliminate the possibility of having other hard central jets.

[6]To the best of our knowledge a proof of the decoupling of Glauber gluons does not exist for threshold dijet production in hadronic collisions, so there is no complete proof of Eq. (44).

[7]When replacing m^2_{JJ} by M^2_{JJ} in Eq. (44), the main difference is that now $\rho_i = 2k^0_i/M_{JJ}$. In this case, the threshold limit $\tau_J \to 1$ alone does not constrain the jet invariant masses, M^2_i, to be small. Since $M^2_i \sim R^2 M^2_{JJ}$, they are constrained to be small by jet algorithms with $R^2 \ll 1$. This induces complications in deriving an all-orders factorization theorem, but still suffices to imply that the factorization formula in Eq. (44) with the replacement $m^2_{JJ} \to M^2_{JJ}$ will sum the next-to-leading logarithms [8].

$$\frac{d\sigma}{dm_{JJ}^2 d(\Delta y)} = \frac{1}{E_{\rm cm}^2} H^{IL}(m_{JJ}^2, \Delta y) \int d\xi_a d\xi_b d\rho_1 d\rho_2 f(\xi_a) f(\xi_b) J_{\rm thr}(m_{JJ}, \rho_1, R) J_{\rm thr}(m_{JJ}, \rho_2, R) \times S_{\rm 2jthr}^{LI}[m_{JJ}(1 - \Delta\tau_J - (1 - \xi_a) - (1 - \xi_b) - \rho_1 - \rho_2), \Delta y] \left[1 + \mathcal{O}\left(\frac{\Lambda_{\rm QCD}}{m_{JJ}}, R, 1 - \Delta\tau_J\right)\right], \quad (44)$$

where one sums over the color basis IL, and for simplicity the dependence on flavor labels and μ of the various functions have been suppressed.[8] The first argument of the soft function $S_{\rm 2jthr}^{LI}$ is the energy it radiates outside the jet cones. The jet function $J_{\rm thr}$ depends on $\rho_i = 2k_i^0/m_{JJ} + M_i^2/m_{JJ}^2$, where k_i^0 is the energy of particles it radiates outside its cone. The threshold limit $\Delta\tau_J \to 1$ forces $\xi_{a,b} \to 1$ and $\rho_{1,2} \to 0$, so we have PDFs, jet functions, and soft functions that all correspond to the threshold limit and contain double logarithms.[9] When these functions are convoluted they consistently reproduce the double logarithms encoded in the renormalization group evolution of H^{IL}, which is shown in Fig. 8(a). Compared to threshold Drell-Yan as described in Sec. II F, there are extra convolutions for the jets and a more sophisticated soft function that is a matrix in color space.

To extend the isolated Drell-Yan factorization theorem to the dijet case, we need to define analogs of the $B_{a,b}^+$ variables in Sec. II A that can constrain the final state in an appropriate manner. We first define lightlike vectors along each jet direction, $n_1^\mu = (1, \vec{n}_1)$ and $n_2^\mu = (1, \vec{n}_2)$ with $\vec{n}_{1,2} = \vec{P}_{1,2}/|\vec{P}_{1,2}|$, and corresponding lightlike vectors in the opposite directions, $\bar{n}_{1,2}^\mu = (1, -\vec{n}_{1,2})$. Next, the final-state particles are divided into four categories $\{a, b, 1, 2\}$ as follows. We define $\mathcal{R}_1$ and $\mathcal{R}_2$ as all particles that have been grouped into jets 1 and 2 by the jet algorithm. The remaining particles not grouped into either of $\mathcal{R}_{1,2}$ are divided into the two hemispheres a and b as before, which defines $\mathcal{R}_{a,b}$. We then define plus momenta

$$P_1^+ = \sum_{k\in\mathcal{R}_1} n_1 \cdot p_k, \qquad P_2^+ = \sum_{k\in\mathcal{R}_2} n_2 \cdot p_k, \qquad B_a^+ = \sum_{k\in\mathcal{R}_a} n_a \cdot p_k, \qquad B_b^+ = \sum_{k\in\mathcal{R}_b} n_b \cdot p_k. \quad (45)$$

This definition of $B_{a,b}^+$ is identical to Eq. (2). Since the jet algorithm is used for the grouping, these variables are infrared safe. The union of the four categories covers all of phase space, so the measurement of all the momenta in Eq. (45) defines a global observable sensitive to all radiation in the event. Just as in our Drell-Yan discussion, the definition of the variables $B_{a,b}^+$ ensures that radiation outside the reach of the detector can safely be ignored.

For the isolated dijet limit we demand that $P_{1,2}^+/M_{JJ} \ll 1$ and $B_{a,b}^+/M_{JJ} \ll 1$. In addition, we constrain y_1 and y_2 such that the jets lie in the central region sufficiently separated from the beam directions. The condition on $P_{1,2}^+$ ensures that the jet regions $\mathcal{R}_{1,2}$ only contain energetic radiation along the direction of their jet plus soft radiation. The condition on $B_{a,b}^+$ has a similar effect as for isolated Drell-Yan. It ensures that there is only soft and no energetic radiation in the central region apart from the two jets. Thus, we have exactly two isolated central jets.

Since each category $\{a, b, 1, 2\}$ predominantly contains all the corresponding collinear particles, this division of phase space mainly affects how the soft radiation is associated to each jet. In analogy to isolated Drell-Yan, we divide the total soft momentum as $k = k_a + k_b + k_1 + k_2$, where each k_i is the total momentum of soft particles in $\mathcal{R}_i$, and we define $k_i^+ = n_i \cdot k_i$. The corresponding isolated dijet soft function, $S_{\not<2\rm j}^{LI}(k_a^+, k_b^+, k_1^+, k_2^+, y_1, y_2)$, depends on all four directions n_1, n_2, n_a, n_b and hence on the rapidities y_1 and y_2. It now contains both incoming and outgoing soft Wilson lines and is a matrix in color space, where we can use the same color basis $\{LI\}$ as in the threshold case. The soft function itself of course differs from the threshold case.

The total jet momenta can now be written as

$$P_1^\mu = \omega_1 \frac{n_1^\mu}{2} + q_1^+ \frac{\bar{n}_1^\mu}{2} + q_\perp^\mu + k_1^\mu, \quad (46)$$

and similarly for P_2^μ. The first three terms on the right-hand side are the contributions from the energetic radiation in the jet, with $\omega_{1,2} \sim M_{JJ}$ and $q_1^+, q_\perp \ll M_{JJ}$. Expanding in the small components of the beam and jet momenta, the hard momentum components in the beams and jets have to satisfy

$$\omega_a \frac{n_a^\mu}{2} + \omega_b \frac{n_b^\mu}{2} = \omega_1 \frac{n_1^\mu}{2} + \omega_2 \frac{n_2^\mu}{2}. \quad (47)$$

Thus, in this limit the two jets are massless and back-to-back in the transverse plane, but need not be back-to-back in three dimensions. In terms of $y_{1,2}$ and $M_{JJ}^2 = \omega_a \omega_b = \omega_1 \omega_2 n_1 \cdot n_2/2$, we then have

[8] We also made a redefinition so that the PDFs f depend on light-cone momentum fractions rather than fixed energy as in Ref. [8], absorbing the difference into the hard functions H_{IL}.

[9] We follow the SCET definition of soft functions as matrix elements of eikonal Wilson lines without subtractions, so $S_{\rm 2jthr}$ has double logarithms. In SCET the jet functions have subtractions.

$$\omega_a = M_{JJ} e^{(y_1+y_2)/2} \equiv x_a E_{\rm cm},$$
$$\omega_b = M_{JJ} e^{-(y_1+y_2)/2} \equiv x_b E_{\rm cm}, \qquad (48)$$
$$\omega_{1,2} = M_{JJ} \frac{\cosh y_{1,2}}{\cosh(\Delta y/2)}.$$

The collinear radiation in the jets is described by jet functions that depend on the invariant-mass variables $\omega_1 q_1^+$ and $\omega_2 q_2^+$. By momentum conservation we have $q_{1,2}^+ = P_{1,2}^+ - k_{1,2}^+$, so the jet functions are convoluted with the soft function through $k_{1,2}^+$. The subtraction of $k_{1,2}^+$ is necessary to remove the plus momentum of soft particles in the jet, since the momentum distribution of these particles is properly described by the soft function not by the jet function. Just as for isolated Drell-Yan, the collinear initial-state radiation in the beams is described by beam functions, which depend on the invariant-mass variable $t_a = \omega_a(B_a^+ - k_a^+)$ and momentum fraction x_a, and similarly for hemisphere b. Since the jets are well-separated from the beams, removing the particles in $\mathcal{R}_{1,2}$ from the hemispheres $\mathcal{R}_{a,b}$ mainly affects the soft radiation and not the energetic partons collinear to the beams. Therefore up to power corrections, we expect the same inclusive beam functions as before.

From the above discussion it is natural to suppose that the factorization theorem for isolated dijet production will be

$$\frac{d\sigma}{dM_{JJ}^2 dy_1 dy_2 dB_a^+ dB_b^+ dP_1^+ dP_2^+} = H^{IL}(M_{JJ}^2, y_1, y_2) \int dk_a^+ dk_b^+ dk_1^+ dk_2^+ J_{\not<}[\omega_1(P_1^+ - k_1^+)] J_{\not<}[\omega_2(P_2^+ - k_2^+)]$$
$$\times B[\omega_a(B_a^+ - k_a^+), x_a] B[\omega_b(B_b^+ - k_b^+), x_b] S_{\not<2\rm j}^{LI}(k_a^+, k_b^+, k_1^+, k_2^+, y_1, y_2)$$
$$\times \left\{1 + \mathcal{O}\left[\frac{\Lambda_{\rm QCD}}{M_{JJ}}, \frac{\omega_{a,b} B_{a,b}^+}{M_{JJ}^2}, \frac{\omega_{1,2} P_{1,2}^+}{M_{JJ}^2}\right]\right\}, \qquad (49)$$

where we again suppressed flavor labels and μ dependence. The hard function $H^{IL}(M_{JJ}, y_i)$ is precisely the threshold hard function, and we sum over the same color basis $\{IL\}$. The subscript $\not<$ on the soft and jet functions denotes the fact that their plus momenta depend on the regions $\mathcal{R}_i$, which in turn depend on y_i.

The consistency of the RGE for the isolated dijet factorization theorem, shown in Fig. 8(b), again provides important constraints on its structure. Each of the functions $J_{\not<}$, B, and $S_{\not<2\rm j}^{LI}$ includes a series of double logarithms, and when these functions are convoluted over the k_i^+ variables at a common scale μ, these different series have to collapse to precisely the double-logarithmic series of the hard function H^{IL}. The RGE for the hard function H^{IL} is a matrix equation in color space, but has no convolutions of kinematic variables. We expect that this equivalence will occur in the same manner as it does for the isolated Drell-Yan case.

Key missing ingredients in providing a rigorous derivation of Eq. (49) include (i) providing a mathematically rigorous treatment of the separation of jets and beams in the factorization, and (ii) determining the role of Glauber degrees of freedom, that in principle may couple the final-state jets and spoil factorization. It should be evident that if such a proof becomes available, it will be straightforward to generalize the above discussion to the case where we produce N isolated jets rather than just two.

III. THE BEAM FUNCTION

In this section, we discuss the properties of the beam function in more detail. We present its definition and relation to the standard PDF, as well as its renormalization

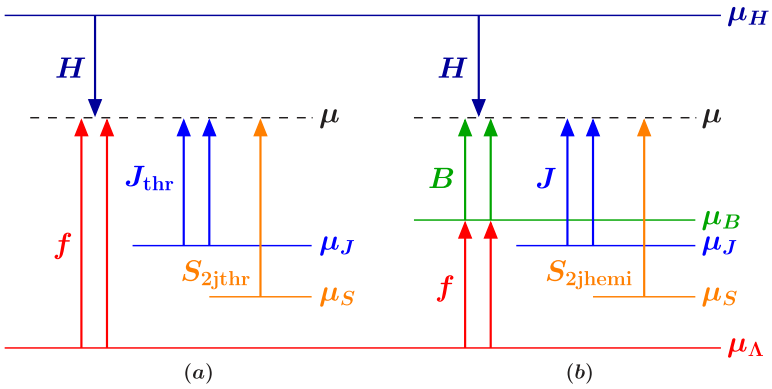


FIG. 8 (color online). RGE running for dijet production for (a) the threshold situation and (b) the isolated situation.

group evolution. We will display explicit results for the quark beam function at one loop (leaving a detailed derivation to a dedicated publication [50]). The comparison of effects in the beam functions and PDFs are illustrated with plots.

The quark, antiquark, and gluon beam functions are defined in SCET as

$$\begin{aligned}
B_q(\omega b^+, \omega/P^-, \mu) &= \frac{\theta(\omega)}{\omega} \int \frac{dy^-}{4\pi} e^{ib^+y^-/2} \Big\langle p_n(P^-) \Big| \bar\chi_n\Big(y^- \frac{n}{2}\Big) \delta(\omega - \bar{\mathcal{P}}_n) \frac{\bar{n}\!\!\!/}{2} \chi_n(0) \Big| p_n(P^-) \Big\rangle, \\
B_{\bar q}(\omega b^+, \omega/P^-, \mu) &= \frac{\theta(\omega)}{\omega} \int \frac{dy^-}{4\pi} e^{ib^+y^-/2} \Big\langle p_n(P^-) \Big| \mathrm{tr}_{\mathrm{spin}}\Big[\frac{\bar{n}\!\!\!/}{2} \chi_n\Big(y^- \frac{n}{2}\Big) \delta(\omega - \bar{\mathcal{P}}_n) \bar\chi_n(0)\Big] \Big| p_n(P^-) \Big\rangle, \\
B_g(\omega b^+, \omega/P^-, \mu) &= -\theta(\omega) \int \frac{dy^-}{4\pi} e^{ib^+y^-/2} \Big\langle p_n(P^-) \Big| \mathcal{B}^c_{n\perp\mu}\Big(y^- \frac{n}{2}\Big) \delta(\omega - \bar{\mathcal{P}}_n) \mathcal{B}^{\mu c}_{n\perp}(0) \Big| p_n(P^-) \Big\rangle.
\end{aligned} \tag{50}$$

We will briefly explain the relevant notation. (A more detailed overview of SCET and the definitions of the objects in Eq. (50) are given in Sec. IV A.) As before, $n^\mu = (1, \vec{n})$ and $\bar{n}^\mu = (1, -\vec{n})$ are lightlike vectors, $n^2 = \bar{n}^2 = 0$, $n \cdot \bar{n} = 2$, where $\vec{n}$ is a unit three-vector in the direction of the proton. The proton states $|p_n(P^-)\rangle$ have lightlike momentum $P^\mu = P^- n^\mu/2$, and the matrix elements are always implicitly averaged over the proton spin. The SCET fields for collinear quarks and gluons, $\chi_n(y)$ and $\mathcal{B}^\mu_{n\perp}(y)$, are composite fields containing Wilson lines of collinear gluons [see Eq. (70)]. Matrix elements with these fields include so-called zero-bin subtractions [71], which effectively divide by a matrix element of Wilson lines [72]. At lowest order in the strong coupling, the fields describe an energetic quark or gluon moving in the n direction with momentum $p^- n^\mu/2 + k^\mu$ with $k \ll p^-$. The momentum operator $\bar{\mathcal{P}}_n$ picks out the large light-cone component p^- of all particles annihilated by χ_n or $\mathcal{B}^\mu_{n\perp}$. Thus, when these fields annihilate the incoming colliding parton, the δ function in Eq. (50) sets ω equal to the p^- of that parton. Therefore, $x = \omega/P^-$ is the fraction of the proton's light-cone momentum that is carried by the parton into the hard collision. At the time of the collision, this parton is propagating in an initial-state jet rather than the proton, which is encoded by the dependence of the beam functions on the variable $b^+ = -k^+$. Here, $k^+ = n \cdot k$ is the small component of the incoming collinear parton's momentum. The variable $t = \omega b^+ \sim -p^2$ measures the parton's virtuality, where $t > 0$, because the parton is spacelike. As we already saw in Eq. (19), the beam functions are convoluted with the soft function through b^+.

The beam-function definitions in Eq. (50) can be compared with those of the standard quark, antiquark, and gluon PDFs in SCET [73],

$$\begin{aligned}
f_q(\omega'/P^-, \mu) &= \theta(\omega') \Big\langle p_n(P^-) \Big| \bar\chi_n(0) \delta(\omega' - \bar{\mathcal{P}}_n) \frac{\bar{n}\!\!\!/}{2} \chi_n(0) \Big| p_n(P^-) \Big\rangle, \\
f_{\bar q}(\omega'/P^-, \mu) &= \theta(\omega') \Big\langle p_n(P^-) \Big| \mathrm{tr}_{\mathrm{spin}}\Big[\frac{\bar{n}\!\!\!/}{2} \chi_n(0) \delta(\omega' - \bar{\mathcal{P}}_n) \bar\chi_n(0)\Big] \Big| p_n(P^-) \Big\rangle, \\
f_g(\omega'/P^-, \mu) &= -\theta(\omega') \omega' \langle p_n(P^-)| \mathcal{B}^c_{n\perp\mu}(0) \delta(\omega' - \bar{\mathcal{P}}_n) \mathcal{B}^{\mu c}_{n\perp}(0) |p_n(P^-)\rangle.
\end{aligned} \tag{51}$$

The f_i depend on the analogous light-cone momentum fraction $\xi = \omega'/P^-$. As discussed in the beginning of Sec. II C, ξ can be interpreted as the momentum fraction of the hard parton when it is taken out of the proton and due to perturbative corrections generally differs from $x = \omega/P^-$ appearing in the beam functions. A more common and equivalent definition of the PDFs is in terms of QCD fields. For example, for the quark PDF,

$$\begin{aligned}
f_q(\omega'/P^-, \mu) = \theta(\omega') \int \frac{dy^+}{4\pi} e^{-i\omega' y^+/2} \Big\langle p_n(P^-) \Big| \bar\psi\Big(y^+ \frac{\bar{n}}{2}\Big) \\
\times \frac{\bar{n}\!\!\!/}{2} W_{\bar n}\Big(y^+ \frac{\bar{n}}{2}, 0\Big) \psi(0) \Big| p_n(P^-) \Big\rangle.
\end{aligned} \tag{52}$$

Equation (51) is essentially the Fourier transform of Eq. (52), where the SCET fields are written in momentum space with respect to ω', while the QCD fields are separated along the $\bar{n}$ direction between 0 and $y^+\bar{n}/2$. The lightlike Wilson line $W_{\bar n}(y^+\bar{n}/2, 0)$ is required to render the product of the quark fields at different space-time points gauge invariant, and the corresponding Wilson lines in Eq. (51) are those hidden in the definitions of χ_n and $\mathcal{B}^\mu_{n\perp}$.

In the beam functions in Eq. (50), the fields are in addition separated along the n direction, with a large separation $y^- \gg y^+$ corresponding to the small momentum $b^+ \ll \omega$. This y^- separation is formulated with a gauge-invariant multipole expansion of fields in SCET. The possible gauge transformations in the effective theory are divided into global, collinear, and soft, and it is the coupling to soft gluons and the corresponding soft gauge transformations that are relevant for making the y^- sepa-

ration gauge invariant. The collinear fields in Eq. (50) are the ones that occur after making a field redefinition to decouple soft gluons into the soft function, and the resulting collinear fields no longer transform under soft gauge transformations. Hence, the SCET definitions in Eq. (50) are gauge invariant.

Note that formulating equivalent definitions of the beam functions directly in QCD is more challenging. It seems to require QCD fields that are simultaneously separated in the n and $\bar{n}$ directions, and *a priori* it is not clear how to obtain an unambiguous gauge-invariant expression in this case, because Wilson lines connecting the fields along different paths are not equivalent. For the beam functions, which one might think of as b^+-dependent PDFs, this problem is solved in SCET, because the effective theory distinguishes the large and small momentum components with the multipole expansion, resolving the ambiguity.

For $t = \omega b^+ \gg \Lambda^2_{\rm QCD}$, or equivalently $y^- \ll \omega/\Lambda^2_{\rm QCD}$, the beam functions can be related to the PDFs by performing an operator product expansion in $\Lambda^2_{\rm QCD}/t \ll 1$. This leads to the factorized form

$$B_i(t, x, \mu) = \sum_j \int_x^1 \frac{d\xi}{\xi} \mathcal{I}_{ij}\left(t, \frac{x}{\xi}, \mu\right) f_j(\xi, \mu) \times \left[1 + \mathcal{O}\left(\frac{\Lambda^2_{\rm QCD}}{t}\right)\right], \tag{53}$$

where $j = \{g, u, \bar{u}, d, \ldots\}$ and $\mathcal{I}_{ij}$ is a perturbatively calculable Wilson coefficient. The physical interpretation of this equation was discussed in Sec. II C. For B_g, the equivalent of the matching expression in Eq. (53) was derived in Ref. [30] for the $\mathcal{I}_{gg}$ term using a moment-space OPE to match $\mathrm{SCET_I}$ onto $\mathrm{SCET_{II}}$. Ref. [30] considered this matching at the level of the matrix element defining B_g, without the accompanying physical picture advocated here that implies that beam functions will occur in a wide variety of interesting processes. The mixing contributions were missed in their analysis, but the extension of their proof to the general case is straightforward [50].

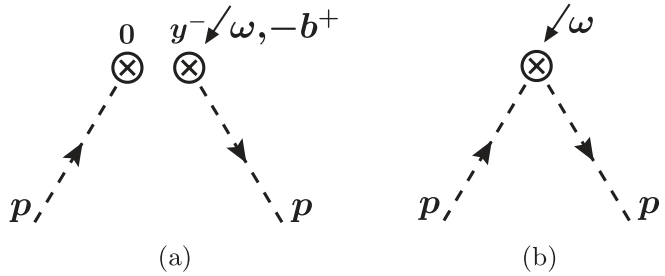


FIG. 9. Tree-level diagrams for the quark beam function (a) and quark PDF (b).

The coefficients $\mathcal{I}_{ij}(t, x/\xi, \mu)$ can be determined perturbatively by computing both sides of Eq. (53) with the proton states in the definitions of B_i and f_j replaced by quark and gluon states. The tree-level diagram for the quark beam function is shown in Fig. 9(a) and for the PDF in Fig. 9(b). They give

$$B^{\rm tree}_{q/q}(t, x, \mu) = \delta(t)\delta(1 - x), \qquad f^{\rm tree}_{q/q}(\xi, \mu) = \delta(1 - \xi), \tag{54}$$

from which we deduce $\mathcal{I}^{\rm tree}_{qq}(t, x/\xi, \mu) = \delta(t)\delta(1 - x/\xi)$. In general, we have

$$\mathcal{I}^{\rm tree}_{ij}\left(t, \frac{x}{\xi}, \mu\right) = \delta_{ij}\delta(t)\delta\left(1 - \frac{x}{\xi}\right), \tag{55}$$

so the tree-level beam functions reduce to the PDFs

$$B^{\rm tree}_i(t, x, \mu) = \delta(t) f_i(x, \mu). \tag{56}$$

The one-loop coefficients for the quark beam function are determined from the diagrams in Fig. 10 together with the corresponding diagrams for the PDFs. The beam functions and PDFs are renormalized using dimensional regularization with $\overline{\rm MS}$. From the first four diagrams in Fig. 10 we find the one-loop correction to the quark-quark coefficient (here $z = x/\xi$)

$$\mathcal{I}^{\rm 1loop}_{qq}(t, z, \mu) = \frac{\alpha_s(\mu) C_F}{2\pi}\theta(z)\left\{\frac{2}{\mu^2}\left[\frac{\theta(t/\mu^2)\ln(t/\mu^2)}{t/\mu^2}\right]_+ \delta(1 - z) + \frac{1}{\mu^2}\left[\frac{\theta(t/\mu^2)}{t/\mu^2}\right]_+ \left[\left[\theta(1 - z)\frac{1 + z^2}{1 - z}\right]_+ - \frac{3}{2}\delta(1 - z)\right] + \delta(t)\left[\left[\frac{\theta(1 - z)\ln(1 - z)}{1 - z}\right]_+ (1 + z^2) - \frac{\pi^2}{6}\delta(1 - z) + \theta(1 - z)\left(1 - z - \frac{1 + z^2}{1 - z}\ln z\right)\right]\right\}. \tag{57}$$

The last two diagrams in Fig. 10 determine the one-loop contribution of the gluon PDF to the quark beam function,

$$\mathcal{I}^{\rm 1loop}_{qg}(t, z, \mu) = \frac{\alpha_s(\mu) T_F}{2\pi}\theta(z)\theta(1 - z)\left\{\left[\frac{1}{\mu^2}\left[\frac{\theta(t/\mu^2)}{t/\mu^2}\right]_+ + \delta(t)\ln\frac{1 - z}{z}\right][z^2 + (1 - z)^2] + \delta(t) 2z(1 - z)\right\}. \tag{58}$$

At two loops, $\mathcal{I}_{q\bar{q}}(t, z, \mu)$ will start to contribute as well. The plus distributions are defined as

$$[\theta(x)g(x)]_+ = \lim_{\epsilon\to 0}\frac{d}{dx}[\theta(x-\epsilon)G(x)] \quad \text{with}$$
$$G(x) = \int_1^x dx' g(x'), \tag{59}$$

satisfying the boundary condition $\int_0^1 dx[\theta(x)g(x)]_+ = 0$. In particular,

$$\int_x^\infty \frac{dz}{z}[\theta(1-z)g(1-z)]_+ f\left(\frac{x}{z}\right)$$
$$= \int_x^1 dz\, g(1-z)\left[\frac{1}{z}f\left(\frac{x}{z}\right) - f(x)\right] + f(x)G(1-x),$$
$$\frac{1}{\mu^2}\int_{-\infty}^{t_{\max}} dt\left[\frac{\theta(t/\mu^2)\ln^n(t/\mu^2)}{t/\mu^2}\right]_+ = \frac{1}{n+1}\ln^{n+1}\frac{t_{\max}}{\mu^2}. \tag{60}$$

The infrared (IR) divergences in the diagrams in Fig. 10 precisely cancel those in the PDF calculation as they must, so the matching coefficients in Eqs. (57) and (58) are IR finite and independent of the IR regulator. The ultraviolet (UV) divergences in the diagrams determine the one-loop RGE and anomalous dimension of the quark beam function, which in $\overline{\text{MS}}$ are

$$\mu\frac{d}{d\mu}B_q(t, x, \mu) = \int dt' \gamma_B^q(t-t', \mu)B_q(t', x, \mu),$$
$$\gamma_B^q(t, \mu) = -2\Gamma_{\text{cusp}}[\alpha_s(\mu)]\frac{1}{\mu^2}\left[\frac{\theta(t/\mu^2)}{t/\mu^2}\right]_+ + \frac{\alpha_s(\mu)}{4\pi}6C_F\delta(t). \tag{61}$$

Here, $\Gamma_{\text{cusp}}[\alpha_s(\mu)]$ is the cusp anomalous dimension [74], and the coefficient of the plus function is equal to Γ_{cusp} to all orders in perturbation theory. The noncusp term of the anomalous dimension is equal to that for the quark jet function at one loop, and in Ref. [50] we prove that the anomalous dimensions for the quark beam and jet functions are identical to all orders in α_s. As stated before, the RGE in Eq. (61) does not change x. Also, the mixing graphs in Figs. 10(e) and 10(f) have no UV divergences and hence the gluon beam function does not mix into the quark beam functions under renormalization. Equation (61) leads to the physical picture discussed in Sec. II C. The RGE has a solution [70,75,76], which can be written as [67]

$$B_q(t, x, \mu) = \int dt' B_q(t-t', x, \mu_0)U_B(t', \mu_0, \mu),$$
$$U_B(t, \mu_0, \mu) = e^{K_B}\frac{e^{-\gamma_E\eta_B}}{\Gamma(1+\eta_B)}\left\{\frac{\eta_B}{\mu_0^2}\left[\frac{\theta(t/\mu_0^2)]}{(t/\mu_0^2)^{1-\eta_B}}\right]_+ + \delta(t)\right\}, \tag{62}$$

with the plus distribution defined according to Eq. (59). Furthermore, $K_B \equiv K_B(\mu_0, \mu)$ and $\eta_B \equiv \eta_B(\mu_0, \mu)$ are

$$K_B(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)}\frac{d\alpha_s}{\beta(\alpha_s)} \times\left[4\Gamma_{\text{cusp}}(\alpha_s)\int_{\alpha_s(\mu_0)}^{\alpha_s}\frac{d\alpha_s'}{\beta(\alpha_s')} + \gamma_B^q(\alpha_s)\right],$$
$$\eta_B(\mu_0, \mu) = -2\int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)}\frac{d\alpha_s}{\beta(\alpha_s)}\Gamma_{\text{cusp}}(\alpha_s), \tag{63}$$

where $\beta(\alpha_s)$ is the QCD β function and $\gamma_B^q(\alpha_s)$ is the coefficient of $\delta(t)$ in $\gamma_B^q(t, \mu)$ in Eq. (61).

Using the above results for the quark beam function, we can see explicitly that when we integrate over $0 \le t \le t_{\max}$ to get the beam-function $\tilde{B}_q(t_{\max}, x, \mu)$ in Eq. (23), the result contains double and single logarithms of $t_{\max}/\mu^2$,

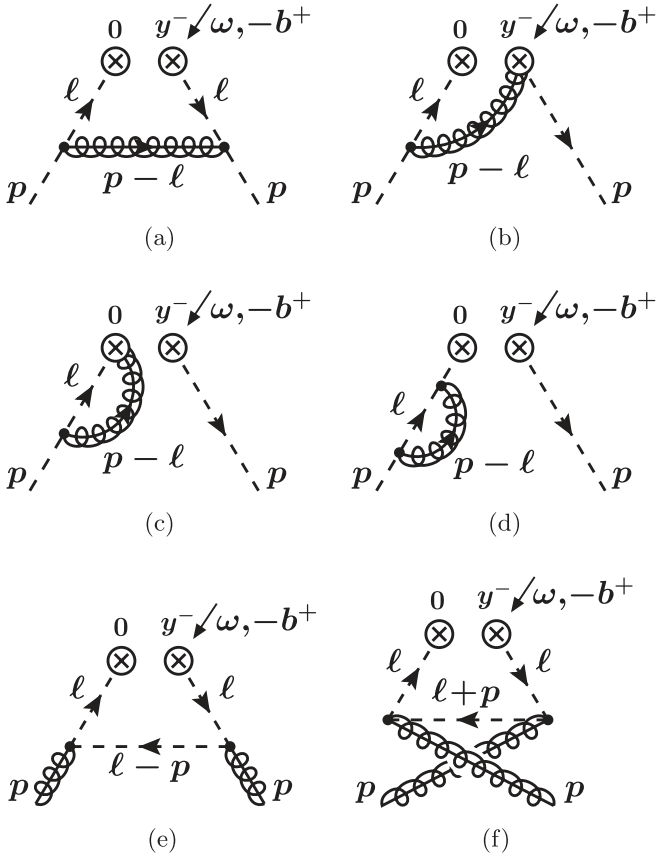


FIG. 10. One-loop diagrams for the quark beam function. Graphs (a) and (b) correspond to real gluon emission, while (c) and (d) are virtual corrections. Graphs (e) and (f) determine the contribution of the gluon PDF to the quark beam function.

$$\tilde{B}_q(t_{\max}, x, \mu) = \theta(t_{\max}) f_q(x, \mu) + \theta(t_{\max}) \frac{\alpha_s(\mu)}{2\pi} \Bigg\{ C_F \Bigg(\ln^2 \frac{t_{\max}}{\mu^2} - \frac{3}{2} \ln \frac{t_{\max}}{\mu^2} \Bigg) f_q(x, \mu) + \ln \frac{t_{\max}}{\mu^2} \int_x^1 \frac{dz}{z} \Bigg\{ C_F \Bigg[\theta(1-z) \frac{1+z^2}{1-z} \Bigg]_+ f_q\Big(\frac{x}{z}, \mu\Big) + T_F[z^2 + (1-z)^2] f_g\Big(\frac{x}{z}, \mu\Big) \Bigg\} + \cdots \Bigg\}. \quad (64)$$

The ellipses denote x-dependent terms that have no $\ln(t_{\max}/\mu^2)$. Equation (64) shows that the natural scale for the beam function is $\mu = \mu_B \sim t_{\max}$. The logarithms of $t_{\max}/\mu^2$ are summed by solving the beam function's RGE in Eq. (61). From Eq. (64) we can see how the matching coefficients $\mathcal{I}_{ij}$ convert the PDF running into the beam function running at one loop. Expanding Eq. (61) to $\mathcal{O}(\alpha_s)$, the integrated beam function satisfies

$$\mu \frac{d}{d\mu} \tilde{B}_q(t, x, \mu) = \frac{\alpha_s(\mu) C_F}{\pi} \Big(\frac{3}{2} - 2 \ln \frac{t}{\mu^2} \Big) \tilde{B}_q^{\text{tree}}(t, x, \mu) + \cdots. \quad (65)$$

Taking the derivative of Eq. (64) with respect to μ, the first term in curly brackets proportional to $f_q(x, \mu)$ reproduces the overall factor in Eq. (65), while the terms in the second line precisely cancel the μ dependence of the tree-level term $f_q(x, \mu)$. Thus, even though at tree level $\tilde{B}_q(t, x, \mu) = \delta(t) f_q(x, \mu)$, the running of $\tilde{B}_q$ does not depend on x.

To illustrate the difference between the beam functions and the PDFs, we may compare the PDFs and $\tilde{B}_q(t_{\max}, x, \mu)$ for fixed $t_{\max}$ as a function of x. For $t_{\max}$, following the discussion in Sec. II A, we take

$$t_{\max} = \omega B_{\max}^+ = Q^2 e^{-2y_{\text{cut}}} = (x E_{\text{cm}})^2 e^{-2y_{\text{cut}}}, \quad (66)$$

where the hard scale is taken as $Q^2 = (x E_{\text{cm}})^2$. By default we use $E_{\text{cm}} = 7$ TeV and $y_{\text{cut}} = 2$. For the parton distributions we use the NLO results of MSTW2008 [77].

The effect of the large logarithms contained in the beam function is illustrated in Fig. 11, where we show the u-quark beam-function $x\tilde{B}_u(t_{\max}, x, \mu_H)$ at NLO for two different values of $t_{\max}$, along with the u-quark PDF $x f_u(x, \mu_H)$, which is equal to the beam function at tree level. All solid central curves are evaluated at the common hard scale $\mu = Q = x E_{\text{cm}}$, and the surrounding bands correspond to varying the scale by a factor of 2. The lower green curve and band show the PDF (tree-level beam function). Including the one-loop matching corrections the beam functions are given for $y_{\text{cut}} = 1$ by the middle blue band and for $y_{\text{cut}} = 2$ by the upper orange band. Clearly, with this scale choice, the beam functions receive large α_s corrections with a dramatically increased scale dependence compared to the PDF, which is caused by the large logarithms of $t_{\max}/\mu^2$ in Eq. (64).

To eliminate the large logarithms in the matching, we have to compute the beam functions at the scale $\mu_B^2 \simeq t_{\max}$, which we show in Fig. 12. Let us first consider the up-quark beam function $\tilde{B}_u$ shown in the top row. The top left panel shows $x\tilde{B}_u(t_{\max}, x, \mu_B)$. The green (light) band shows the PDF $x f_u(x, \mu_H)$ evaluated at the hard scale $\mu_H = x E_{\text{cm}}$, varying the scale by a factor of 2 (which for the top left panel is identical to the corresponding band in Fig. 11). The one-loop beam function evaluated at $\mu_B^2 \simeq t_{\max}$ is shown by the orange (medium) bands. Here, the maximum scale variation for $\mu_B^2 \in [t_{\max}/2, 2t_{\max}]$ is not obtained at the edges of this region, but is closely approximated by taking $\mu_B = \{0.7, 2.0\}\sqrt{t_{\max}}$. The solid line shows the corresponding central value at $\mu_B = 1.4\sqrt{t_{\max}}$. For comparison, the dotted line shows the tree-level result, i.e. the PDF at the scale μ_B. The top right panel shows the same curves as the top left panel, but normalized to the tree-level beam function. The plot shows that the beam function can be reliably calculated at $\mu = \mu_B$. Using $\mu_B^2 \simeq t_{\max}$, the shift from the dotted to solid line is now of reasonable size, and the scale uncertainties are now similar in size to those of the PDFs. We also see that evaluating the PDF at the beam scale rather than the hard scale has a significant effect. The difference between evaluating the PDF at μ_H and μ_B is a $+30\%$ (-20%) correction at large (small) x, while the α_s corrections from the beam function at μ_B are only $\sim 10\%$. (The beam function curves increase for $x \gtrsim 0.4$ due to the threshold term $[\ln(1-z)/(1-z)]_+$ in $\mathcal{I}_{qq}$, but this region only has a small total contribution as can be seen on the left panel.) Since the residual μ_B dependence is only canceled by other contributions in the factorization theorem, these plots do not determine the overall size of the α_s correc-

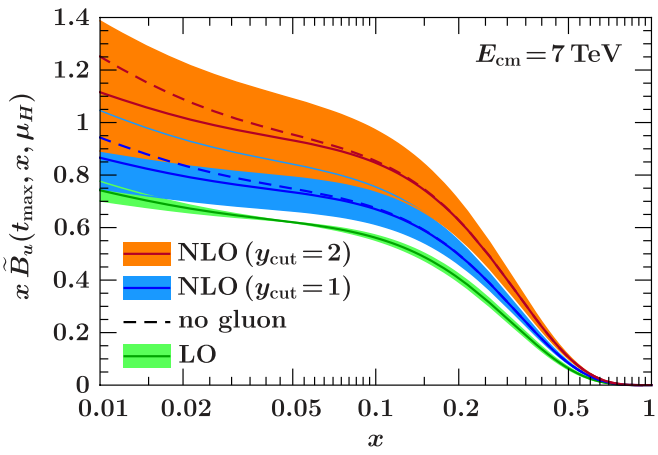


FIG. 11 (color online). The u-quark PDF (lower green band) and u-quark beam functions for $y_{\text{cut}} = 1$ (middle blue band) and $y_{\text{cut}} = 2$ upper orange band. All functions are evaluated at the same hard scale, with the bands showing the scale variation by factors of two.

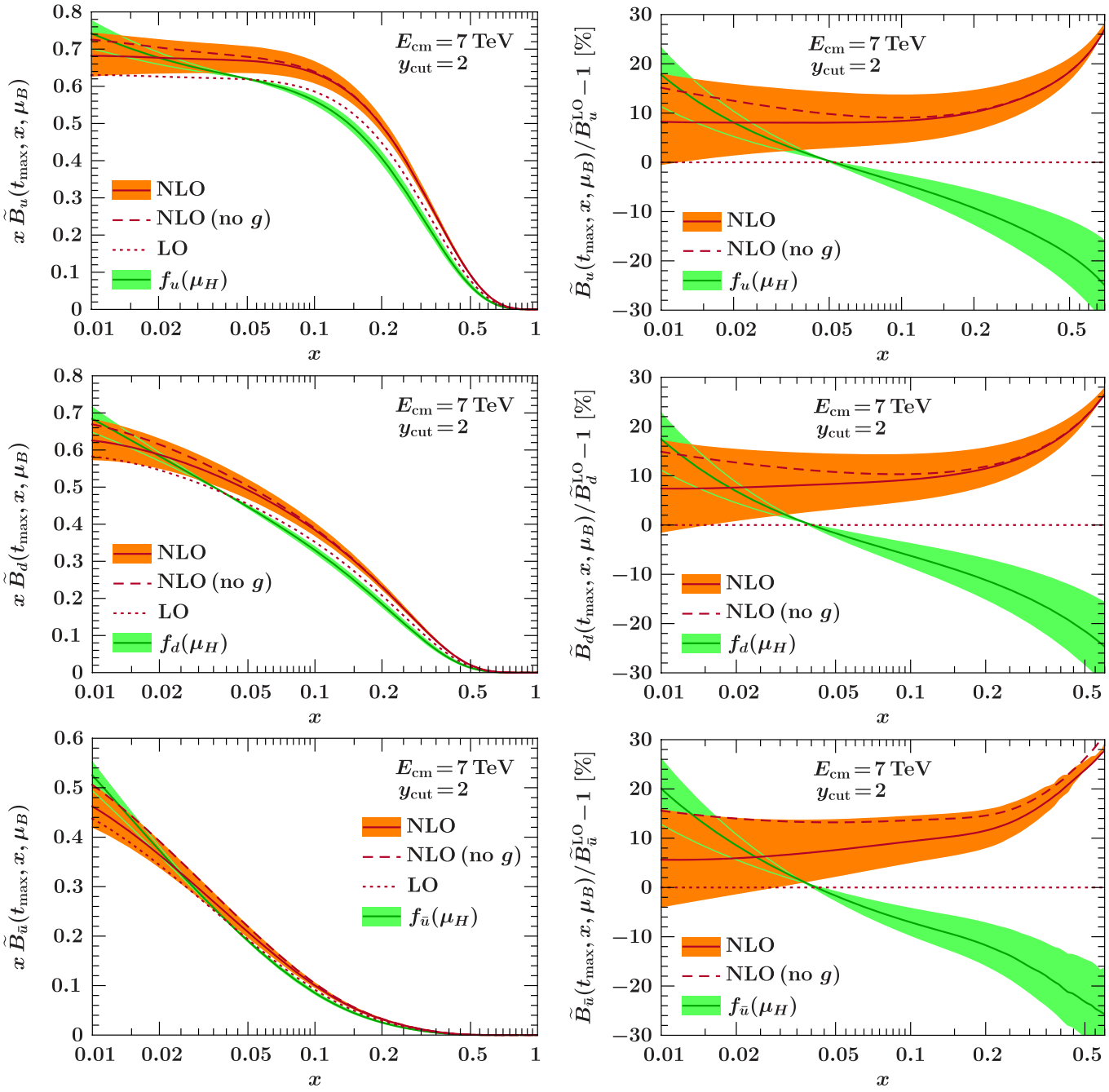


FIG. 12 (color online). The u (top row), d (middle row), and $\bar u$ (bottom row) beam functions at the beam scale $\mu_B^2 \simeq t_{\rm max}$ for $E_{\rm cm} = 7$ TeV and $t_{\rm max}$ with $y_{\rm cut} = 2$ [see Eq. (66)]. The panels on the left show the functions times x. The right panels show the relative differences compared to the respective tree-level (LO) beam functions given by the dotted lines on the left. The bands show the scale uncertainties as explained in the text.

tions, nor their uncertainty. These questions are addressed by plots of the full cross section in Sec. VI.

In the central and bottom panels of Fig. 12 we show analogous plots for the down-quark beam function, $\tilde B_d$, and the antiup-quark, $\tilde B_{\bar u}$. While the absolute size of the functions for different flavors in the left panels are quite different, the relative corrections shown in the right panel are very similar.

Figure 12 also contains dashed lines, which show how the solid lines are modified if we remove the gluon contribution $\mathcal{I}_{qg}$ from the one-loop beam function. The gluon contribution to the quark beam functions becomes noticeable at $x < 0.1$, increasing to about -5% at $x = 0.01$, while for the antiquark beam function it is important in the entire x range. This is expected, since the antiquark PDF is much smaller, so the gluon PDF can have a bigger

impact. The gluon contribution is always negative and partially compensates the quark matching correction. Even at $x = 0.01$ there is no indication that treating the logarithms $\ln x$ in fixed-order perturbation theory causes any problems. The contribution from the $\ln z$ term in Eq. (57) is of similar size as other contributions and within the perturbative uncertainties. It has roughly the same size as the gluon contribution.

IV. ISOLATED FACTORIZATION THEOREM

In this section, we derive the isolated factorization theorem in Eq. (19). Our analysis is based on factorization in SCET, which rigorously and systematically separates hard, soft and collinear contributions [26,27,73]. We make use of a setup with $\mathrm{SCET_I}$ and $\mathrm{SCET_{II}}$ [78], carrying out the factorization in two stages at the scales Q^2 and $\omega_{a,b}B^+_{a,b}$ respectively. We have an $\mathrm{SCET_I}$ analysis to factorize initial-state jets from soft radiation. The initial-state jets described by beam functions in $\mathrm{SCET_I}$ are then matched onto initial-state PDFs with lower off-shellness for the collinear particles in $\mathrm{SCET_{II}}$. In this section, we carry out the $\mathrm{SCET_I}$ computation, while the matching onto $\mathrm{SCET_{II}}$ was discussed in Sec. III. Our analysis below uses similar tools as used in the derivation of the factorization theorem for hemisphere invariant masses for $e^+e^- \to$ dijets in Ref. [69], but differs significantly due to the kinematics, and the fact that we have initial-state rather than final-state jets and a further matching onto $\mathrm{SCET_{II}}$. The soft dynamics of $e^+e^- \to$ dijets was studied earlier in SCET in Refs. [79,80]. We start with a brief overview of the necessary SCET ingredients in Sec. IV A and describe the relevant kinematics in Sec. IV B. We derive the factorization theorem for isolated $pp \to XL$ in Sec. IV C, including arguments to rule out contributions from so-called Glauber degrees of freedom. Finally in Sec. IV D, we apply the factorization theorem to $pp \to X\ell^+\ell^-$ and quote final results for the beam thrust cross section with one-loop corrections and logarithmic resummation.

A. SCET

Soft-collinear effective theory is an effective field theory of QCD that describes the interactions of collinear and soft particles [24–27]. Collinear particles are characterized by having large energy and small invariant mass. To separate the large and small momentum components, it is convenient to use light-cone coordinates. We define two light-cone vectors

$$n^\mu = (1, \vec{n}), \qquad \bar{n}^\mu = (1, -\vec{n}), \tag{67}$$

with $n^2 = \bar{n}^2 = 0$, $n \cdot \bar{n} = 2$, and $\vec{n}$ is a unit three-vector. Any four-momentum p can then be decomposed as

$$p^\mu = \bar{n} \cdot p \frac{n^\mu}{2} + n \cdot p \frac{\bar{n}^\mu}{2} + p^\mu_{n\perp}. \tag{68}$$

Choosing $\vec{n}$ close to the direction of a collinear particle, its momentum p scales as $(n \cdot p, \bar{n} \cdot p, p_{n\perp}) \sim \bar{n} \cdot p(\lambda^2, 1, \lambda)$, with $\lambda \ll 1$ a small parameter. For example, for a jet of collinear particles in the $\vec{n}$ direction with total momentum p_X, $\bar{n} \cdot p_X \simeq 2E_X$ corresponds to the large energy of the jet, while $n \cdot p_X \simeq p_X^2/E_X \ll E_X$, so $\lambda^2 \simeq p_X^2/E_X^2 \ll 1$.

To construct the fields of the effective theory, the momentum is written as

$$p^\mu = \tilde{p}^\mu + k^\mu = \bar{n} \cdot \tilde{p} \frac{n^\mu}{2} + \tilde{p}^\mu_{n\perp} + k^\mu, \tag{69}$$

where $\bar{n} \cdot \tilde{p} \sim Q$ and $\tilde{p}_{n\perp} \sim \lambda Q$ are the large momentum components, where Q is the scale of the hard interaction, while $k \sim \lambda^2 Q$ is a small residual momentum. The effective theory expansion is in powers of the small parameter λ.

The SCET fields for n-collinear quarks and gluons, $\xi_{n,\tilde{p}}(x)$ and $A_{n,\tilde{p}}(x)$, are labeled by the collinear direction n and their large momentum $\tilde{p}$. They are written in position space with respect to the residual momentum and in momentum space with respect to the large momentum components. Frequently, we will only keep the label n denoting the collinear direction, while the momentum labels are summed over and suppressed. Derivatives acting on the fields pick out the residual momentum dependence, $\mathrm{i}\partial^\mu \sim k \sim \lambda^2 Q$. The large label momentum is obtained from the momentum operator $\mathcal{P}^\mu_n$, e.g. $\mathcal{P}^\mu_n \xi_{n,\tilde{p}} = \tilde{p}^\mu \xi_{n,\tilde{p}}$. If there are several fields, $\mathcal{P}_n$ returns the sum of the label momenta of all n-collinear fields. For convenience, we define $\bar{\mathcal{P}}_n = \bar{n} \cdot \mathcal{P}_n$, which picks out the large minus component.

Collinear operators are constructed out of products of fields and Wilson lines that are invariant under collinear gauge transformations [25,26]. The smallest building blocks are collinearly gauge-invariant quark and gluon fields, defined as

$$\begin{aligned} \chi_{n,\omega}(x) &= [\delta(\omega - \bar{\mathcal{P}}_n) W_n^\dagger(x) \xi_n(x)], \\ \mathcal{B}^\mu_{n,\omega\perp}(x) &= \frac{1}{g}[\delta(\omega + \bar{\mathcal{P}}_n) W_n^\dagger(x) \mathrm{i}D^\mu_{n\perp} W_n(x)], \end{aligned} \tag{70}$$

where

$$\mathrm{i}D^\mu_{n\perp} = \mathcal{P}^\mu_{n\perp} + gA^\mu_{n\perp} \tag{71}$$

is the collinear covariant derivative and

$$W_n(x) = \Big[\sum_{\text{perms}} \exp\Big(\frac{-g}{\bar{\mathcal{P}}_n} \bar{n} \cdot A_n(x)\Big)\Big]. \tag{72}$$

The label operators in Eqs. (70) and (72) only act inside the square brackets. Here, $W_n(x)$ is a Wilson line of n-collinear gluons in label momentum space. It sums up arbitrary emissions of n-collinear gluons from an n-collinear quark or gluon, which are $\mathcal{O}(1)$ in the power counting. Since $W_n(x)$ is localized with respect to the residual position x, we can treat $\chi_{n,\omega}(x)$ and $\mathcal{B}^\mu_{n,\omega}(x)$ as local quark and gluon fields. The label momentum ω is treated as a continuous variable, which is why we use a δ-function operator in

Eq. (70). It is set equal to the sum of the minus label momenta of all fields that the δ function acts on, including those in the Wilson lines, while the label momenta of the individual fields are summed over.

In general, the effective theory can contain several collinear sectors, each containing collinear fields along a different collinear direction. To have a well-defined power expansion in this case, the different collinear directions n_i have to be well-separated [73],

$$n_i \cdot n_j \gg \lambda^2 \quad \text{for } i \neq j, \tag{73}$$

which is simply the requirement that different collinear sectors are distinct and do not overlap. For $pp \to X\ell^+\ell^-$, we need two collinear sectors, n_a and n_b, along the directions of the two beams. We use a bar to denote the conjugate lightlike vector, so $n_i \cdot \bar{n}_i = 2$. As the beams are back-to-back, we have $n_a \sim \bar{n}_b$, so $n_a \cdot n_b \sim 2$ and Eq. (73) is easily satisfied.

Particles that exchange large momentum of $\mathcal{O}(Q)$ between collinear particles moving in different directions have to be off shell by an amount of $\mathcal{O}(n_i \cdot n_j Q^2)$. These modes can be integrated out of the theory at the hard scale Q by matching full QCD onto SCET, which yields the hard function. The effective theory below the scale Q then splits into several distinct collinear sectors, where particles in the same collinear sector can still interact with each other, while at leading order in the power-counting particles from different collinear sectors can only interact by the exchange of soft particles. This means that before and after the hard interaction takes place, the jets described by the different collinear sectors evolve independently from each other with only soft but no hard interactions between them.

The soft degrees of freedom, responsible for the radiation between collinear jets, are described in the effective theory by soft[10] quark and gluon fields, $q_s(x)$ and $A_s(x)$, which only have residual soft momentum dependence $\mathrm{i}\partial^\mu \sim \lambda^2 Q$. They couple to the collinear sectors via the soft covariant derivative

$$\mathrm{i}\, D_s^\mu = \mathrm{i}\partial^\mu + g A_s^\mu \tag{74}$$

acting on the collinear fields. At leading order in λ, n-collinear particles only couple to the $n \cdot A_s$ component of soft gluons, so the leading-order n-collinear Lagrangian only depends on $n \cdot D_s$. For n-collinear quarks [25,26]

$$\mathcal{L}_n = \bar{\xi}_n \Big(\mathrm{i} n \cdot D_s + g n \cdot A_n + \mathrm{i} \not{D}_{n\perp} W_n \frac{1}{\bar{\mathcal{P}}_n} W_n^\dagger \mathrm{i} \not{D}_{n\perp} \Big) \frac{\bar{\not{n}}}{2} \xi_n. \tag{75}$$

[10]In some situations it is necessary to distinguish two types of soft sectors, referred to as soft and ultrasoft in the SCET literature. In this paper we only need what are usually called ultrasoft particles, so we will simply refer to these as soft.

The leading-order n-collinear Lagrangian for gluons is given in Ref. [27].

The coupling of soft gluons to collinear particles can be removed at leading order by defining new collinear fields [27]

$$\begin{aligned} \chi_{n,\omega}^{(0)}(x) &= Y_n^\dagger(x) \chi_{n,\omega}(x), \\ \mathcal{B}_{n,\omega\perp}^{\mu(0)}(x) &= Y_n^\dagger(x) \mathcal{B}_{n,\omega\perp}^{\mu}(x) Y_n(x) = \mathcal{B}_{n,\omega\perp}^{\mu d}(x) \mathcal{Y}_n^{dc}(x) T^c, \end{aligned} \tag{76}$$

where $Y_n(x)$ and $\mathcal{Y}_n(x)$ are soft Wilson lines in the fundamental and adjoint representations,

$$\begin{aligned} Y_n(x) &= P \exp\Big[\mathrm{i} g \int_{-\infty}^0 ds\; n \cdot A_s(x + sn) \Big], \\ T^c \mathcal{Y}_n^{cd}(x) &= Y_n(x) T^d Y_n^\dagger(x). \end{aligned} \tag{77}$$

The symbol P in Eq. (77) denotes the path ordering of the color generators along the integration path. The integral limits in Eq. (77) with the reference point at $-\infty$ are the natural choice for incoming particles [81]. The final results are always independent of the choice of reference point, and with the above choice the interpolating fields for the incoming proton states do not introduce additional Wilson lines [82].

After the field redefinition in Eq. (76), the leading-order SCET Lagrangian separates into the sum of independent n_i-collinear and soft Lagrangians,

$$\mathcal{L}_{\text{SCET}} = \sum_{n_i} \mathcal{L}_{n_i}^{(0)} + \mathcal{L}_s + \cdots, \tag{78}$$

with no interactions between any of the collinear and soft sectors. The ellipses denote terms that are subleading in the power counting. This decoupling is what will allow us to factorize the cross section into separate beam and soft functions. The field redefinition in Eq. (76) introduces soft Wilson lines in the operators, which because of Eq. (78) can be factored out of the matrix element and will make up the soft function.

B. Kinematics

Before deriving the factorization theorem, we discuss the relevant kinematics, as illustrated in Fig. 13. As already mentioned, we introduce a separate set of collinear fields for each of the beams, with the light-cone vectors n_a and n_b aligned with the beam directions. To derive the factorization theorem we work in the center-of-mass frame of the hadronic collision, so the momenta of the incoming protons are (neglecting the proton mass)

$$P_a^\mu = E_{\text{cm}} \frac{n_a^\mu}{2}, \qquad P_b^\mu = E_{\text{cm}} \frac{n_b^\mu}{2}, \tag{79}$$

with $\vec{n}_a = -\vec{n}_b$. In particular, $n_b = \bar{n}_a$ and $n_a \cdot n_b = 2$. We will mostly keep the dependence on the two beam directions explicit, but one should keep in mind that n_a and n_b are related.

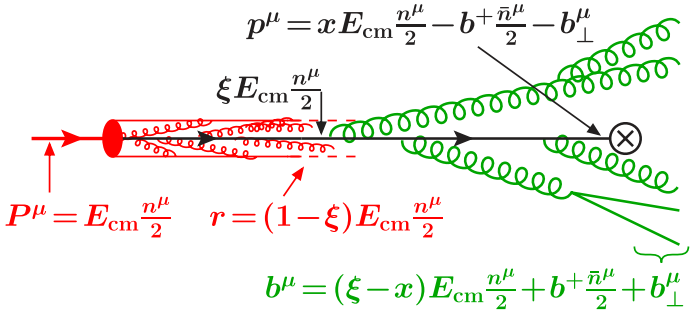


FIG. 13 (color online). Definition of the different collinear momenta related to the incoming beams. The soft radiation is not shown.

The collinear fields in the n_a and n_b directions describe the interactions within each of the beams before and after the collision, and are also responsible for initiating the hard interaction. We define the momenta of the spacelike off-shell partons that go into the hard interaction as

$$\begin{aligned} p_a^\mu &= x_a E_{\rm cm} \frac{n_a^\mu}{2} - b_a^+ \frac{\bar n_a^\mu}{2} - b_{a\perp}^\mu, \\ p_b^\mu &= x_b E_{\rm cm} \frac{n_b^\mu}{2} - b_b^+ \frac{\bar n_b^\mu}{2} - b_{b\perp}^\mu, \end{aligned} \tag{80}$$

where x_a and x_b are the light-cone momentum fractions at which the beam functions will be evaluated. The power-counting parameters for the collinear sectors are

$$\lambda_a^2 \sim \frac{b_a^+}{x_a E_{\rm cm}}, \qquad \lambda_b^2 \sim \frac{b_b^+}{x_b E_{\rm cm}}, \tag{81}$$

where the relevant momenta are those of the off-shell partons in Eq. (80), because these are the momenta carried by the n_a- and n_b-collinear fields.

We write the momentum of the incoming partons that are taken out of the proton as

$$\xi_a E_{\rm cm} \frac{n_a^\mu}{2} + \mathcal{O}(\Lambda_{\rm QCD}), \qquad \xi_b E_{\rm cm} \frac{n_b^\mu}{2} + \mathcal{O}(\Lambda_{\rm QCD}), \tag{82}$$

which defines the light-cone momentum fractions $\xi_{a,b}$ at which the PDFs are evaluated. The typical $\perp$ -momenta of partons in the proton are $\mathcal{O}(\Lambda_{\rm QCD})$, while the small plus components are $\mathcal{O}(\Lambda_{\rm QCD}^2/E_{\rm cm})$. These momenta are much smaller than any soft or residual momenta in $\rm SCET_I$ and are expanded, which precisely corresponds to the OPE for the beam functions in Eq. (34) when matching them onto $\rm SCET_{II}$.

The momentum of the final-state remnant of the proton is thus given by

$$r_a^\mu = (1 - \xi_a) E_{\rm cm} \frac{n_a^\mu}{2}, \tag{83}$$

while the remnant of the initial-state jet radiated into the final state by the beam function has momentum

$$b_a^\mu = (\xi_a - x_a) E_{\rm cm} \frac{n_a^\mu}{2} + b_a^+ \frac{\bar n_a^\mu}{2} + b_{a\perp}^\mu, \tag{84}$$

and similarly for the n_b direction. The total n_a-collinear momentum in the final state is the sum of Eqs. (83) and (84), or equivalently, the difference between the proton momentum and Eq. (80),

$$b_a^\mu + r_a^\mu = P_a^\mu - p_a^\mu = (1 - x_a) E_{\rm cm} \frac{n_a^\mu}{2} + b_a^+ \frac{\bar n_a^\mu}{2} + b_{a\perp}^\mu. \tag{85}$$

In addition to the collinear momenta, we define k_s^μ as the total four-momentum of the soft radiation in the final state. Hence, the total hadronic momentum in the final state is given by

$$p_X^\mu = (P_a^\mu - p_a^\mu) + (P_b^\mu - p_b^\mu) + k_s^\mu, \tag{86}$$

and we can write total momentum conservation $P_a^\mu + P_b^\mu = p_X^\mu + q^\mu$ as

$$p_a^\mu + p_b^\mu = q^\mu + k_s^\mu, \tag{87}$$

where q^μ is the total leptonic momentum.

The collinear and soft momenta, b_a^μ, b_b^μ, k_s^μ are not experimentally measurable quantities. Instead, the experiments can only measure hadronic quantities, such as the hemisphere momenta $B_a^+ = n_a \cdot B_a$ and $B_b^+ = n_b \cdot B_b$ introduced in Sec. II A. Splitting the total soft momentum into its contributions from each hemisphere, $k_s^\mu = k_a^\mu + k_b^\mu$ as shown in Fig. 3, we then have

$$B_a^\mu = b_a^\mu + r_a^\mu + k_a^\mu, \qquad B_b^\mu = b_b^\mu + r_b^\mu + k_b^\mu, \tag{88}$$

and defining $k_a^+ = n_a \cdot k_a$, $k_b^+ = n_b \cdot k_b$, we get

$$\begin{aligned} B_a^+ &= n_a \cdot B_a = b_a^+ + k_a^+, \\ B_b^+ &= n_b \cdot B_b = b_b^+ + k_b^+. \end{aligned} \tag{89}$$

In particular, the remnant momenta $r_{a,b}^\mu$ do not contribute to $B_{a,b}^+$. A physical argument for this was discussed in Sec. II A.

Next, we decompose the total leptonic momentum as

$$q^\mu = q^- \frac{n_a^\mu}{2} + q^+ \frac{n_b^\mu}{2} + q_\perp^\mu, \tag{90}$$

where $q_\perp^\mu$ contains the two components of q^μ transverse to the beam direction. Taking the z-axis along the $\vec{n}_a$ beam direction, we have

$$q^\pm = q^0 \mp q_z, \qquad q_\perp^\mu = (0, \vec{q}_T, 0), \tag{91}$$

where $\vec{q}_T = (q_x, q_y)$ is a two-vector in the transverse x-y-plane. The total leptonic invariant mass and rapidity are

$$q^2 = q^+ q^- + q_\perp^2 = q^+ q^- - \vec{q}_T^2, \qquad Y = \frac{1}{2} \ln \frac{q^-}{q^+}, \tag{92}$$

with

$$\begin{aligned} q^\mp &= e^{\pm Y} \sqrt{q^2 + \vec{q}_T^2}, \\ d^4q &= \frac{1}{2} dq^+ dq^- d^2\vec{q}_T = \frac{1}{2} dq^2 dY d^2\vec{q}_T. \end{aligned} \tag{93}$$

As we will see in the next subsection, the derivation of the factorization theorem requires us to be insensitive to the transverse components $\vec{q}_T$ such that we can freely integrate over them. Therefore, we have to expand the kinematics in the limit $\vec{q}_T = 0$. This expansion is justified because from Eq. (87) we have

$$q_\perp^\mu = -p_{X\perp}^\mu = -b_{a\perp}^\mu - b_{b\perp}^\mu - k_{s\perp}^\mu \sim \lambda Q. \tag{94}$$

A parametrically large $q_\perp^\mu \sim Q$ would require a separate jet at large $p_T \sim Q$ to balance the transverse momentum, which is not allowed in our setup. After integrating over the lepton phase space this expansion incurs power corrections of order $q_\perp^2 \sim \lambda^2 Q^2$. The kinematics of the hard matrix element in the factorization theorem is then given by the tree-level partonic kinematics, with the partonic momentum conservation

$$x_a E_{\rm cm} \frac{n_a^\mu}{2} + x_b E_{\rm cm} \frac{n_b^\mu}{2} = q = q^- \frac{n_a^\mu}{2} + q^+ \frac{n_b^\mu}{2}, \tag{95}$$

which implies

$$\begin{aligned} x_a E_{\rm cm} &= q^- = \sqrt{q^2} e^Y, \qquad x_b E_{\rm cm} = q^+ = \sqrt{q^2} e^{-Y}, \\ q^2 &= q^+ q^- = x_a x_b E_{\rm cm}^2, \qquad Y = \frac{1}{2} \ln \frac{q^-}{q^+} = \frac{1}{2} \ln \frac{x_a}{x_b}. \end{aligned} \tag{96}$$

Equations (94) and (95) imply that parametrically the leptons are back-to-back in the transverse plane. Since q^+ and q^- can differ substantially, the leptons do not need to be back-to-back in three dimensions.

C. Derivation of isolated factorization theorem

We now proceed to derive the isolated factorization theorem for generic processes $pp \to XL$, where the hadronic final state X has a restriction on the hemisphere momenta $B_{a,b}^+$. The derivation is carried out using SCET without Glauber degrees of freedom. The proof that Glauber effects are not required is given at the end of this subsection.

1. Cross section in QCD

We will generically refer to properties of L as "leptonic", even though L can contain any nonstrongly interacting particles. We only consider processes where the hard interaction couples the strong and electroweak sectors through one two-particle QCD current. (This includes, for example, Drell-Yan or Higgs production through gluon fusion with the Higgs decaying nonhadronically, but does not include electroweak Higgs production via vector-boson fusion.) Then, at leading order in the electroweak interactions, we can factorize the full-theory matrix element into its leptonic and hadronic parts

$$\mathcal{M}(pp \to XL) = \sum_J L_J \langle X|J|pp\rangle. \tag{97}$$

The sum runs over all relevant color-singlet two-particle QCD currents J, and L_J contains the corresponding electroweak matrix element, including the electroweak propagator coupling to J. For example, for Drell-Yan with $L = \ell^+\ell^-$, the relevant currents are

$$J_{Vf}^\mu = \bar{q}_f \gamma^\mu q_f, \qquad J_{Af}^\mu = \bar{q}_f \gamma^\mu \gamma_5 q_f, \tag{98}$$

so in this case the sum over J in Eq. (97) includes the sums over the two Dirac structures, the vector index μ, and the quark flavor $f = \{u, d, \ldots\}$. The corresponding L_{Vf}^μ and L_{Af}^μ are given below in Eq. (137).

The cross section for some hadronic observable O in the center-of-mass frame of the collision, averaged over proton spins, is

$$\begin{aligned} \frac{d\sigma}{dq^2 dY dO} = \frac{1}{2E_{\rm cm}^2} \int \frac{d^2\vec{q}_T}{2(2\pi)^4} \int d\Phi_L (2\pi)^4 \delta^4(q - p_L) \\ \times \frac{1}{4} \sum_{\rm spins} \sum_X |\mathcal{M}(pp \to XL)|^2 \delta[O - f_O(X)] \\ \times (2\pi)^4 \delta^4(P_a + P_b - q - p_X). \end{aligned} \tag{99}$$

Here, $P_{a,b}$ are the incoming proton momenta, p_X and p_L are the total hadronic and leptonic momenta, $d\Phi_L$ denotes the leptonic phase space, and the phase-space integrations for the hadronic final states are included in the sum over X. The last δ function is overall momentum conservation. The function $f_O(X)$ inside the second δ function returns the value of the hadronic observable O for a given hadronic state X, so the δ function picks out all final states that contribute to a certain value of O. The $\delta^4(q - p_L)$ under the leptonic phase-space integral defines the measured q as the total leptonic momentum. Expanding this δ function for $\vec{q}_T = 0$, the leptonic part does not depend on $\vec{q}_T$ at

leading order, and using Eq. (97), we can rewrite Eq. (99) as

$$\frac{d\sigma}{dq^2 dY dO} = \frac{1}{2E_{\rm cm}^2}\sum_{J,J'} L_{JJ'}(q^2, Y) W_{JJ'}(q^2, Y, O). \quad (100)$$

The leptonic tensor is defined as

$$L_{JJ'}(q^2, Y) = \int d\Phi_L L_J^\dagger L_{J'} (2\pi)^4 \delta^4\Big(q^- \frac{n_a}{2} + q^+ \frac{n_b}{2} - p_L\Big), \quad (101)$$

where $q^\pm = \sqrt{q^2} e^{\mp Y}$. The hadronic tensor contains the square of the hadronic matrix element

$$W_{JJ'}(q^2, Y, O) = \int \frac{d^2\vec{q}_T}{2(2\pi)^4} \sum_X \langle pp|J^\dagger(0)|X\rangle\langle X|J'(0)|pp\rangle \times (2\pi)^4 \delta^4(P_a + P_b - q - p_X) \times \delta[O - f_O(X)], \quad (102)$$

where as in Sec. III we keep the average over proton spins implicit in the matrix element. Since $W_{JJ'}$ is integrated over $\vec{q}_T$, it can only depend on q^2 and Y, as well as the hadronic observable O.

We are interested in the hadronic observables $B_a^+ = n_a \cdot B_a$ and $B_b^+ = n_b \cdot B_b$. The hemisphere hadronic momenta $B_{a,b}^\mu(X)$ can be obtained from the states $|X\rangle$ using the hemisphere momentum operators $\hat{p}_{a,b}^\mu$

$$\hat{p}_a^\mu |X\rangle = B_a^\mu(X)|X\rangle, \qquad \hat{p}_b^\mu |X\rangle = B_b^\mu(X)|X\rangle. \quad (103)$$

A field-theoretic definition of $\hat{p}_{a,b}^\mu$ in terms of the energy-momentum tensor of the field theory was given in Ref. [17]. The hadronic tensor for $O \equiv \{B_a^+, B_b^+\}$ is

$$W_{JJ'}(q^2, Y, B_a^+, B_b^+) = \int \frac{d^2\vec{q}_T}{2(2\pi)^4} \int d^4x e^{-iq\cdot x} \sum_X \langle pp|J^\dagger(x)|X\rangle\langle X|J'(0)|pp\rangle \delta[B_a^+ - n_a \cdot B_a(X)]\delta[B_b^+ - n_b \cdot B_b(X)]$$
$$= \int \frac{dx^+ dx^-}{(4\pi)^2} e^{-i(q^+x^- + q^-x^+)/2} \Big\langle pp \Big| J^\dagger\Big(x^- \frac{n_a}{2} + x^+ \frac{n_b}{2}\Big)\delta(B_a^+ - n_a \cdot \hat{p}_a)\delta(B_b^+ - n_b \cdot \hat{p}_b) J'(0) \Big| pp\Big\rangle. \quad (104)$$

In the first line we used momentum conservation to shift the position of $J^\dagger$, and in the second line we performed the integral over $\vec{q}_T$, which sets $\vec{x}_T$ to zero. We also used Eq. (103) to eliminate the explicit dependence on X, allowing us to carry out the sum over all states X. The restriction on the states X is now implicit through the operator δ functions inside the matrix element.

2. Matching QCD onto SCET

In the next step, we match the QCD currents J onto SCET currents by integrating out fluctuations at the hard scale Q. At leading order in the power counting, the matching takes the form

$$J(x) = \sum_{n_1,n_2} \int d\omega_1 d\omega_2 e^{-i(\tilde{b}_1 + \tilde{b}_2)\cdot x} \times \Big[\sum_q C_{Jq\bar{q}}^{\alpha\beta}(\tilde{b}_1, \tilde{b}_2) O_{q\bar{q}}^{\alpha\beta}(\tilde{b}_1, \tilde{b}_2; x) + C_{Jgg}^{\mu\nu}(\tilde{b}_1, \tilde{b}_2) O_{gg\mu\nu}(\tilde{b}_1, \tilde{b}_2; x)\Big], \quad (105)$$

where α, β are spinor indices, μ, ν are vector indices, and the sum over q runs over all quark flavors $\{u, d, \ldots\}$. The Wilson coefficients and operators depend on the large label momenta

$$\tilde{b}_1^\mu = \omega_1 \frac{n_1^\mu}{2}, \qquad \tilde{b}_2^\mu = \omega_2 \frac{n_2^\mu}{2}. \quad (106)$$

They will eventually be set to either $q^- n_a^\mu/2$ or $q^+ n_b^\mu/2$ by momentum conservation, but at this point are unspecified, and the sums and integrals over n_1, n_2 and ω_1, ω_2 in Eq. (105) run over all sets of distinct collinear directions and large label momenta. On the right-hand side of Eq. (105), the full x dependence of the current is separated into the x dependence appearing in the overall phase factor with large label momenta and the residual x dependence of the SCET operators.

The SCET operators $O_{q\bar{q}}^{\alpha\beta}(x)$ and $O_{gg}^{\mu\nu}(x)$ are constructed out of the collinear fields in Eq. (70). At leading order in the power counting they contain one field for each collinear direction. Since the QCD currents are color singlets, the leading operators that can contribute are

$$O_{q\bar{q}}^{\alpha\beta}(\tilde{b}_1, \tilde{b}_2; x) = \bar{\chi}_{n_1,-\omega_1}^{\alpha j}(x) \chi_{n_2,\omega_2}^{\beta j}(x),$$
$$O_{gg}^{\mu\nu}(\tilde{b}_1, \tilde{b}_2; x) = \sqrt{\omega_1\omega_2}\, \mathcal{B}_{n_1,-\omega_1\perp}^{\mu c}(x) \mathcal{B}_{n_2,-\omega_2\perp}^{\nu c}(x), \quad (107)$$

where j and c are color indices in the fundamental and adjoint representations. We included appropriate minus signs on the labels, such that we always have $\omega_{1,2} > 0$ for incoming particles. Here, $\chi \equiv \chi_q$ is a quark field of

flavor q, which for simplicity we keep implicit in our notation. Note that the entire spin and flavor structure of the current J is hidden in the label J on the matching coefficients in Eq. (105). The gluon operator is symmetric under interchanging both $\mu \leftrightarrow \nu$ and $\tilde{b}_1 \leftrightarrow \tilde{b}_2$, so its matching coefficient must have the same symmetry,

$$C^{\nu\mu}_{Jgg}(\tilde{b}_2, \tilde{b}_1) = C^{\mu\nu}_{Jgg}(\tilde{b}_1, \tilde{b}_2). \tag{108}$$

We define the conjugate quark operator and matching coefficient with the usual factors of γ^0, i.e.,

$$\begin{aligned} O^{\dagger\beta\alpha}_{q\bar{q}}(\tilde{b}_1, \tilde{b}_2) &= \bar{\chi}^{\beta j}_{n_2,\omega_2}(x)\chi^{\alpha j}_{n_1,-\omega_1}(x), \\ \bar{C}^{\beta\alpha}_{Jq\bar{q}}(\tilde{b}_1, \tilde{b}_2) &= [\gamma^0 C^{\dagger}_{Jq\bar{q}}(\tilde{b}_1, \tilde{b}_2)\gamma^0]^{\beta\alpha}. \end{aligned} \tag{109}$$

The matching coefficients are obtained by computing the renormalized matrix elements $\langle 0|\ldots|q\bar{q}\rangle$ and $\langle 0|\ldots|gg\rangle$ on both sides of Eq. (105) and comparing the results. In pure dimensional regularization for UV and IR divergences all loop graphs in SCET are scaleless and vanish, which means the UV and IR divergences in the bare matrix elements precisely cancel each other. The renormalized matrix elements of the right-hand side of Eq. (105) are then given by their tree-level expressions plus pure $1/\epsilon$ IR divergences, which cancel against those of the full-theory matrix elements $\langle 0|J|q\bar{q}\rangle$ and $\langle 0|J|gg\rangle$ of the left-hand side. Hence, the matching coefficients in $\overline{\text{MS}}$ are given in terms of the IR-finite parts of the renormalized full-theory matrix elements computed in pure dimensional regularization.

3. Soft-collinear factorization

The field redefinitions in Eq. (76) introduce soft Wilson lines into the operators in Eq. (107),

$$\begin{aligned} O^{\alpha\beta}_{q\bar{q}}(x) &= \bar{\chi}^{(0)\alpha j}_{n_1,-\omega_1}(x)T[Y^{\dagger}_{n_1}(x)Y_{n_2}(x)]^{jk}\chi^{(0)\beta k}_{n_2,\omega_2}(x), \\ O^{\mu\nu}_{gg}(x) &= \sqrt{\omega_1\omega_2}\mathcal{B}^{(0)\mu c}_{n_1,-\omega_1\perp}(x)T[\mathcal{Y}^{\dagger}_{n_1}(x)\mathcal{Y}_{n_2}(x)]^{cd} \\ &\quad \times \mathcal{B}^{(0)\nu d}_{n_2,-\omega_2\perp}(x). \end{aligned} \tag{110}$$

The time ordering is required to ensure the proper ordering of the soft-gluon fields inside the Wilson lines. It only affects the ordering of the field operators, while the ordering of the color generators is still determined by the (anti)path ordering of the Wilson lines. In the remainder, we use these redefined fields and drop the (0) superscript for convenience.

Since the momentum operator is linear in the Lagrangian, Eq. (78) allows us to write the hemisphere momentum operators as the sum of independent operators acting in the separate collinear and soft sectors,

$$\begin{aligned} \hat{p}_a &= \hat{p}_{a,n_a} + \hat{p}_{a,n_b} + \hat{p}_{a,s}, \\ \hat{p}_b &= \hat{p}_{b,n_a} + \hat{p}_{b,n_b} + \hat{p}_{b,s}. \end{aligned} \tag{111}$$

The n_a (n_b) collinear sector cannot contribute momentum in the n_b (n_a) hemisphere. Thus, $\hat{p}_{a,n_b} = \hat{p}_{b,n_a} = 0$, while $\hat{p}_{a,n_a} = \hat{p}_{n_a}$ and $\hat{p}_{b,n_b} = \hat{p}_{n_b}$ reduce to the total momentum operators for each of the collinear sectors. For the soft sector, the distinction between the two hemisphere operators is important. We can now write

$$\begin{aligned} \delta(B_a^+ - n_a\cdot\hat{p}_a) &= \int db_a^+ dk_a^+ \delta(B_a^+ - b_a^+ - k_a^+) \\ &\quad \times \delta(b_a^+ - n_a\cdot\hat{p}_{n_a})\delta(k_a^+ - n_a\cdot\hat{p}_{a,s}), \\ \delta(B_b^+ - n_b\cdot\hat{p}_b) &= \int db_b^+ dk_b^+ \delta(B_b^+ - b_b^+ - k_b^+) \\ &\quad \times \delta(b_b^+ - n_b\cdot\hat{p}_{n_b})\delta(k_b^+ - n_b\cdot\hat{p}_{b,s}). \end{aligned} \tag{112}$$

Using Eq. (105) in the hadronic tensor in Eq. (104), the forward matrix element of the product of currents turns into the forward matrix element of the product of the operators in Eq. (110). Since the Lagrangian in Eq. (78) contains no interactions between the collinear and soft sectors after the field redefinition, we can use Eq. (112) to factorize the resulting matrix element into a product of independent n_a-collinear, n_b-collinear, and soft matrix elements.

We first look at the contribution from $O_{q\bar{q}}$. The x integral of the forward matrix element of $O_{q\bar{q}}$ becomes

$$\begin{aligned} &\int \frac{dx^+ dx^-}{(4\pi)^2} e^{-i(q^+x^- + q^-x^+)/2} e^{i(\tilde{b}_1+\tilde{b}_2)\cdot x}\langle p_{n_a}p_{n_b}|O^{\dagger\beta\alpha}_{q\bar{q}}(x)\delta(B_a^+ - n_a\cdot\hat{p}_a)\delta(B_b^+ - n_b\cdot\hat{p}_b)O^{\alpha'\beta'}_{q\bar{q}}(0)|p_{n_a}p_{n_b}\rangle \\ &= \int \frac{dx^+ dx^-}{(4\pi)^2} e^{-i(q^+x^- + q^-x^+)/2} \int db_a^+ db_b^+ dk_a^+ dk_b^+ \delta(B_a^+ - b_a^+ - k_a^+)\delta(B_b^+ - b_b^+ - k_b^+) \\ &\quad \times \int d\omega_a d\omega_b e^{i(\omega_a x^+ + \omega_b x^-)/2}\{\delta_{n_2 n_a}\delta(\omega_2 - \omega_a)\delta_{n_2' n_a}\delta(\omega_2' - \omega_a)\delta_{n_1 n_b}\delta(\omega_1 - \omega_b)\delta_{n_1' n_b}\delta(\omega_1' - \omega_b) \\ &\quad \times \theta(\omega_a)\langle p_{n_a}|\bar{\chi}^{\beta k}_{n_a}(x)\delta(b_a^+ - n_a\cdot\hat{p}_{n_a})\delta(\omega_a - \bar{\mathcal{P}}_{n_a})\chi^{\beta' k'}_{n_a}(0)|p_{n_a}\rangle \\ &\quad \times \theta(\omega_b)\langle p_{n_b}|\chi^{\alpha j}_{n_b}(x)\delta(b_b^+ - n_b\cdot\hat{p}_{n_b})\delta(\omega_b - \bar{\mathcal{P}}_{n_b})\bar{\chi}^{\alpha' j'}_{n_b}(0)|p_{n_b}\rangle \\ &\quad \times \langle 0|\bar{T}[Y^{\dagger}_{n_a}(x)Y_{n_b}(x)]^{kj}\delta(k_a^+ - n_a\cdot\hat{p}_{a,s})\delta(k_b^+ - n_b\cdot\hat{p}_{b,s})T[Y^{\dagger}_{n_b}(0)Y_{n_a}(0)]^{j'k'}|0\rangle + (a\leftrightarrow b)\}. \end{aligned} \tag{113}$$

Here, $|p_{n_a}\rangle$ and $|p_{n_b}\rangle$ are the proton states with momenta $P^\mu_{a,b} = E_{\rm cm} n^\mu_{a,b}/2$ as in Eq. (79). The two terms in brackets in Eq. (113) arise from the different ways of matching up the fields with the external proton states. The restriction to have positive labels ω requires the fields in $O_{q\bar{q}}$ to be matched with the incoming proton states and the fields in $O^\dagger_{q\bar{q}}$ with the outgoing proton states. In principle, there are two more ways to match the fields and external states, yielding matrix elements with the structure $\langle p|\chi\chi|p\rangle$ and $\langle p|\bar{\chi}\,\bar{\chi}\,|p\rangle$, which vanish due to quark flavor number conservation in QCD. For the same reason, in the full product $(\sum_q O^\dagger_{q\bar{q}})(\sum_{q'} O_{q'\bar{q}'})$ only the flavor-diagonal term with $q = q'$ survives.

We abbreviate the collinear and soft matrix elements in the last three lines of Eq. (113) as $M_{\omega_a}(x^-)$, $M_{\omega_b}(x^+)$, $M_s(x^+, x^-)$. The collinear matrix elements only depend on one light-cone coordinate because the label momenta $\omega_{a,b}$ are defined to be continuous. We could have also started with discrete label momenta, $\tilde{\omega}_{a,b}$, and then convert to continuous labels by absorbing the residual $k^-_{n_a}$ dependence as follows:

$$\begin{aligned}
&\sum_{\tilde{\omega}_a} e^{i\tilde{\omega}_a x^+/2} M_{\tilde{\omega}_a}(x^+, x^-) \\
&= \sum_{\tilde{\omega}_a} \int dk^-_a\, e^{i(\tilde{\omega}_a + k^-_{n_a})x^+/2} M_{\tilde{\omega}_a + k^-_{n_a}}(x^-) \\
&= \int d\omega_a e^{i\omega_a x^+/2} M_{\omega_a}(x^-),
\end{aligned} \tag{114}$$

and analogously for $M_{\omega_b}(x^+)$. In the second step we used that by reparametrization invariance the Fourier-transformed matrix element can only depend on the linear combination $\tilde{\omega}_a + k^-_{n_a} = \omega_a$.

As an aside, note that in the well-studied case where the collinear matrix elements are between vacuum states, giving rise to jet functions, the distinction between discrete and continuous labels is not as relevant. In that case, the SCET Feynman rules imply that the collinear matrix elements do not depend on the residual k^- (and $k_\perp$) components, and therefore the label momenta can be treated in either way. In our case, momentum conservation with the external state forces the collinear matrix elements to depend on k^-. Therefore, the only way to eliminate the residual k^- dependence is to absorb it into continuous ω labels. One can easily see this already at tree level. Replacing the proton states by quark states with momentum $p = \tilde{p} + p_r$, we get

$$\begin{aligned}
&\int \frac{dx^+ dx^-}{(4\pi)^2} e^{-i(k^- x^+ + k^+ x^-)/2} \\
&\quad \times \langle q(p)|\bar{\chi}_n(x^+, x^-)\delta_{\tilde{\omega},\bar{\mathcal{P}}_n}\chi_n(0)|q(p)\rangle \\
&= \bar{u}u\,\delta_{\tilde{\omega},\tilde{p}^-}\delta(k^- - p_r^-)\delta(k^+ - p_r^+),
\end{aligned} \tag{115}$$

and the label and residual minus momenta are combined using $\delta_{\tilde{\omega},\tilde{p}^-}\delta(k^- - p_r^-) = \delta(\omega - p^-)$. Our continuous ω is physical and corresponds to the momentum fraction of the quark in the proton.

Returning to our discussion, to perform the x integral in Eq. (113), we take the residual Fourier transforms of the matrix elements,

$$\begin{aligned}
M_{\omega_a}(x^-) &= \int \frac{dk^+}{2\pi} e^{ik^+ x^-/2} \tilde{M}_{\omega_a}(k^+), \\
M_{\omega_b}(x^+) &= \int \frac{dk^-}{2\pi} e^{ik^- x^+/2} \tilde{M}_{\omega_b}(k^-), \\
M_s(x^-, x^+) &= \int \frac{dk_s^+ dk_s^-}{(2\pi)^2} e^{i(k_s^+ x^+ + k_s^- x^+)/2} \tilde{M}_s(k_s^+, k_s^-).
\end{aligned} \tag{116}$$

Just as $x^\pm$, the residual momenta $k^\pm$ and $k_s^\pm$ here are all defined with respect to the common $n = n_a$. The x integral in Eq. (113) now becomes

$$\begin{aligned}
&\int \frac{dx^+ dx^-}{(4\pi)^2} e^{i(\omega_a - q^-)x^+/2} e^{i(\omega_b - q^+)x^-/2} M_{\omega_a}(x^-) M_{\omega_b}(x^+) M_s(x^+, x^-) \\
&= \int \frac{dk^+}{2\pi} \frac{dk^-}{2\pi} \frac{dk_s^+ dk_s^-}{(2\pi)^2} \tilde{M}_{\omega_a}(k^+)\tilde{M}_{\omega_b}(k^-)\tilde{M}_s(k_s^+, k_s^-)\delta(\omega_a - q^- + k^- + k_s^-)\delta(\omega_b - q^+ + k^+ + k_s^+) \\
&= \delta(\omega_a - q^-)\delta(\omega_b - q^+) M_{\omega_a}(0) M_{\omega_b}(0) M_s(0).
\end{aligned} \tag{117}$$

In the last step we expanded $q^\pm - k^\pm - k_s^\pm = q^\pm[1 + \mathcal{O}(\lambda^2)]$. The remaining residual integrations are then simply the Fourier transforms of the matrix elements at $x = 0$.

Note that without the integration over $\vec{q}_T$ in the hadronic tensor Eq. (104), the currents would depend on $x_\perp$, which would require us to include perpendicular components $b_{a,b\perp}$ in the label momenta, and the soft matrix element would depend on $x_\perp$, too. (The residual $k_\perp$ dependence in the collinear matrix elements can again be absorbed into continuous $b_{a,b\perp}$.) The corresponding $x_\perp$ integration in Eq. (117) would yield an additional δ function $\delta^2(\vec{b}_{a\perp} + \vec{b}_{b\perp} + \vec{q}_T - \vec{k}_{s\perp})$. Integrating over $\vec{q}_T$ effectively eliminates this δ function, which would otherwise force us to introduce an explicit dependence on $b_{a,b\perp}$ in the beam functions. If one considers the q_T spectrum of the dileptons for $q_T^2 \ll q^2$, our analysis here provides a starting point but

requires further study. One cannot just use p_T in place of $B_{a,b}^+$ with our arguments to impose an analogous restriction on the final state, because at $\mathcal{O}(\alpha_s^2)$ one can have two jets at high $\vec{p}_T$ that still have small total $\vec{p}_T$.

The n_a-collinear matrix element now reduces to the quark beam functions defined in Eq. (50),

$$\begin{aligned}
M_{\omega_a}(0) &= \theta(\omega_a)\langle p_{n_a}|\bar{\chi}_{n_a}^{\beta k}(0)\delta(b_a^+ - n_a\cdot\hat{p}_{n_a})\delta(\omega_a - \bar{\mathcal{P}}_{n_a})\chi_{n_a}^{\beta' k'}(0)|p_{n_a}\rangle \\
&= \frac{\not{n}_a^{\beta'\beta}}{4}\frac{\delta^{k'k}}{N_c}\theta(\omega_a)\left\langle p_{n_a}\left|\,\bar{\chi}_{n_a}(0)\delta(b_a^+ - n_a\cdot\hat{p}_{n_a})\delta(\omega_a - \bar{\mathcal{P}}_{n_a})\frac{\not{\bar{n}}_a}{2}\chi_{n_a}(0)\,\right|p_{n_a}\right\rangle \\
&= \frac{\not{n}_a^{\beta'\beta}}{4}\frac{\delta^{k'k}}{N_c}\theta(\omega_a)\int\frac{dy^-}{4\pi}e^{ib_a^+y^-/2}\left\langle p_{n_a}\left|\,e^{-i\hat{p}_{n_a}^+y^-/2}e^{i\hat{p}_{n_a}^+y^-/2}\bar{\chi}_{n_a}(0)e^{-i\hat{p}_{n_a}^+y^-/2}\delta(\omega_a - \bar{\mathcal{P}}_{n_a})\frac{\not{\bar{n}}_a}{2}\chi_{n_a}(0)\,\right|p_{n_a}\right\rangle \\
&= \frac{\not{n}_a^{\beta'\beta}}{4}\frac{\delta^{k'k}}{N_c}\omega_a B_q(\omega_a b_a^+, \omega_a/P_a^-).
\end{aligned} \tag{118}$$

We abbreviated $\hat{p}_{n_a}^+ = n_a\cdot\hat{p}_{n_a}$, and in the last step we used $e^{i\hat{p}_n^+y^-/2}\bar{\chi}_n(0)e^{-i\hat{p}_n^+y^-/2} = \bar{\chi}_n(y^-n/2)$ and $\hat{p}_n^+|p_n\rangle = 0$. Similarly, for the antiquark beam function we have

$$M_{\omega_b}(0) = \theta(\omega_b)\langle p_{n_b}|\chi_{n_b}^{\alpha j}(x)\delta(b_b^+ - n_b\cdot\hat{p}_b)\delta(\omega_b - \bar{\mathcal{P}}_{n_b})\bar{\chi}_{n_b}^{\alpha' j'}(0)|p_{n_b}\rangle = \frac{\not{n}_b^{\alpha\alpha'}}{4}\frac{\delta^{jj'}}{N_c}\omega_b B_{\bar{q}}(\omega_b b_b^+, \omega_b/P_b^-). \tag{119}$$

Since the collinear matrix elements are color diagonal, the soft matrix element reduces to an overall color-singlet trace, which defines the $q\bar{q}$ incoming hemisphere soft function,

$$S_{\text{ihemi}}^{q\bar{q}}(k_a^+, k_b^+) = \frac{1}{N_c}\,\text{tr}\langle 0|\bar{T}[Y_{n_a}^\dagger(0)Y_{n_b}(0)]\delta(k_a^+ - n_a\cdot\hat{p}_{a,s})\delta(k_b^+ - n_b\cdot\hat{p}_{b,s})T[Y_{n_b}^\dagger(0)Y_{n_a}(0)]|0\rangle. \tag{120}$$

The trace is over color and the factor of $1/N_c$ is included by convention, such that at tree level we have $S_{\text{ihemi}}^{q\bar{q},\text{tree}}(k_a^+, k_b^+) = \delta(k_a^+)\delta(k_b^+)$. The soft matrix element in the second term of Eq. (113) with $a \leftrightarrow b$ interchanged is equal to the above one due to charge conjugation invariance of QCD. Under charge conjugation, the Wilson lines transform as $\mathrm{C}^{-1}Y_n^{ij}\mathrm{C} = T[Y_n^{\dagger ji}]$. The explicit time ordering is required because the fields in Y_n are time-ordered by default, and charge conjugation only changes the ordering of the color generators but not of the field operators. For us this is not relevant, because the ordering of the fields is determined by the overall (anti-)time ordering in the matrix element. Thus, for the soft matrix element with $a \leftrightarrow b$ interchanged, we find

$$\begin{aligned}
&\text{tr}\langle 0|\bar{T}[Y_{n_b}^\dagger Y_{n_a}]\delta(k_a^+ - n_a\cdot\hat{p}_{a,s})\delta(k_b^+ - n_b\cdot\hat{p}_{b,s})T[Y_{n_a}^\dagger Y_{n_b}]|0\rangle \\
&\overset{\mathrm{C}}{=} \text{tr}\langle 0|\bar{T}[Y_{n_b}^T Y_{n_a}^{\dagger T}]\delta(k_a^+ - n_a\cdot\hat{p}_{a,s})\delta(k_b^+ - n_b\cdot\hat{p}_{b,s})T[Y_{n_a}^T Y_{n_b}^{\dagger T}]|0\rangle = S_{\text{ihemi}}^{q\bar{q}}(k_a^+, k_b^+),
\end{aligned} \tag{121}$$

where the transpose refers to the color indices. In the last step we used $\text{tr}[A^TB^TC^TD^T] = \text{tr}[BADC]$ and the fact that the fields in $Y_{n_b}^\dagger$ and Y_{n_a} are spacelike separated and thus commute. Under parity, we have $\mathrm{P}^{-1}Y_{n_a}\mathrm{P} = Y_{n_b}$ and $\mathrm{P}^{-1}n_a\cdot\hat{p}_{a,s}\mathrm{P} = n_b\cdot\hat{p}_{b,s}$. Therefore, CP invariance implies that $S_{\text{ihemi}}^{q\bar{q}}$ is symmetric in its arguments,

$$S_{\text{ihemi}}^{q\bar{q}}(k_a^+, k_b^+) \overset{\text{CP}}{=} \frac{1}{N_c}\,\text{tr}\langle 0|\bar{T}[Y_{n_a}^\dagger Y_{n_b}]\delta(k_a^+ - n_b\cdot\hat{p}_{b,s})\delta(k_b^+ - n_a\cdot\hat{p}_{a,s})T[Y_{n_b}^\dagger Y_{n_a}]|0\rangle = S_{\text{ihemi}}^{q\bar{q}}(k_b^+, k_a^+). \tag{122}$$

Having worked out the different terms in Eq. (113), we are ready to include the remaining pieces from Eqs. (105) and (104). The $q\bar{q}$ contribution to the hadronic tensor becomes

$$
\begin{aligned}
W_{JJ'q\bar{q}}(q^2, Y, B_a^+, B_b^+) &= \int d\omega_a d\omega_b \delta(\omega_a - q^-)\delta(\omega_b - q^+) \sum_{n_1,n_2,n_1',n_2'} \int d\omega_1 d\omega_2 d\omega_1' d\omega_2' \bar{C}_{Jq\bar{q}}^{\beta\alpha}(\tilde{b}_1, \tilde{b}_2) C_{J'q\bar{q}}^{\alpha'\beta'}(\tilde{b}_1', \tilde{b}_2') \\
&\quad \times \Big\{ \delta_{n_2 n_a}\delta(\omega_2 - \omega_a)\delta_{n_2' n_a}\delta(\omega_2' - \omega_a)\delta_{n_1 n_b}\delta(\omega_1 - \omega_b)\delta_{n_1' n_b}\delta(\omega_1' - \omega_b) \frac{\not{n}_a^{\beta'\beta}}{4} \frac{\not{n}_{n_b}^{\alpha\alpha'}}{4} \frac{1}{N_c} \\
&\quad \times \int dk_a^+ dk_b^+ q^2 B_q[x_a E_{\rm cm}(B_a^+ - k_a^+), x_a] B_{\bar{q}}[x_b E_{\rm cm}(B_b^+ - k_b^+), x_b] S_{\rm ihemi}^{q\bar{q}}(k_a^+, k_b^+) + (a \leftrightarrow b) \Big\} \\
&= H_{JJ'q\bar{q}}(\tilde{b}_a, \tilde{b}_b) \int dk_a^+ dk_b^+ q^2 B_q[x_a E_{\rm cm}(B_a^+ - k_a^+), x_a] B_{\bar{q}}[x_b E_{\rm cm}(B_b^+ - k_b^+), x_b] \\
&\quad \times S_{\rm ihemi}^{q\bar{q}}(k_a^+, k_b^+) + (q \leftrightarrow \bar{q}).
\end{aligned} \tag{123}
$$

All label sums and integrations from Eq. (105) eliminate the label δ's from Eq. (113). In the second step we defined

$$
\tilde{b}_a^\mu = x_a E_{\rm cm} \frac{n_a^\mu}{2}, \qquad \tilde{b}_b^\mu = x_b E_{\rm cm} \frac{n_b^\mu}{2}, \qquad x_a \equiv \frac{\omega_a}{E_{\rm cm}} = \frac{q^-}{E_{\rm cm}} = \frac{\sqrt{q^2} e^Y}{E_{\rm cm}}, \qquad x_b \equiv \frac{\omega_b}{E_{\rm cm}} = \frac{q^+}{E_{\rm cm}} = \frac{\sqrt{q^2} e^{-Y}}{E_{\rm cm}}, \tag{124}
$$

as in Eq. (96), and introduced the hard functions

$$
H_{JJ'q\bar{q}}(\tilde{b}_a, \tilde{b}_b) = \frac{1}{N_c} \frac{1}{4} \, {\rm tr}_{\rm spins}\Big[\frac{\not{n}_a}{2} \bar{C}_{Jq\bar{q}}(\tilde{b}_b, \tilde{b}_a) \frac{\not{n}_b}{2} C_{J'q\bar{q}}(\tilde{b}_b, \tilde{b}_a) \Big], \qquad H_{JJ'\bar{q}q}(\tilde{b}_a, \tilde{b}_b) = H_{JJ'q\bar{q}}(\tilde{b}_b, \tilde{b}_a). \tag{125}
$$

Equation (123) is the final factorized result for the $O_{q\bar{q}}$ contribution to the hadronic tensor.

Repeating the same steps for O_{gg}, we obtain for the forward matrix element

$$
\begin{aligned}
&\int \frac{dx^+ dx^-}{(4\pi)^2} e^{-{\rm i}(q^+x^- + q^-x^+)/2} e^{{\rm i}(\tilde{b}_1 + \tilde{b}_2)\cdot x} \langle p_{n_a} p_{n_b} | O_{gg}^{\dagger\nu\mu}(x) \delta(B_a^+ - n_a \cdot \hat{p}_{a,s}) \delta(B_b^+ - n_b \cdot \hat{p}_{b,s}) O_{gg}^{\mu'\nu'}(0) | p_{n_a} p_{n_b} \rangle \\
&= \int d\omega_a d\omega_b \delta(\omega_a - q^-)\delta(\omega_b - q^+) \int db_a^+ db_b^+ dk_a^+ dk_b^+ \delta(B_a^+ - b_a^+ - k_a^+)\delta(B_b^+ - b_b^+ - k_b^+) \\
&\quad \times [\delta_{n_1 n_a}\delta(\omega_1 - \omega_a)\delta_{n_2 n_b}\delta(\omega_2 - \omega_b) + (a \leftrightarrow b)][\delta_{n_1' n_a}\delta(\omega_1' - \omega_a)\delta_{n_2' n_b}\delta(\omega_2' - \omega_b) + (a \leftrightarrow b)] \\
&\quad \times \omega_a \theta(\omega_a) \langle p_{n_a} | \mathcal{B}_{n_a\perp}^{\mu c}(0) \delta(b_a^+ - n_a \cdot \hat{p}_{n_a}) \delta(\omega_a - \bar{\mathcal{P}}_{n_a}) \mathcal{B}_{n_a\perp}^{\mu' c'}(0) | p_{n_a} \rangle \\
&\quad \times \omega_b \theta(\omega_b) \langle p_{n_b} | \mathcal{B}_{n_b\perp}^{\nu d}(0) \delta(b_b^+ - n_b \cdot \hat{p}_{n_b}) \delta(\omega_b - \bar{\mathcal{P}}_{n_b}) \mathcal{B}_{n_b\perp}^{\nu' d'}(0) | p_{n_b} \rangle \\
&\quad \times \langle 0 | \bar{T}[\mathcal{Y}_{n_a}^\dagger(0) \mathcal{Y}_{n_b}(0)]^{cd} \delta(k_a^+ - n_a \cdot \hat{p}_{a,s}) \delta(k_b^+ - n_b \cdot \hat{p}_{b,s}) T[\mathcal{Y}_{n_b}^\dagger(0) \mathcal{Y}_{n_a}(0)]^{d'c'} | 0 \rangle,
\end{aligned} \tag{126}
$$

where we already performed the integral over x. The four terms in the third line correspond to the four different ways to match up the gluon fields with the incoming proton states. The collinear matrix elements reduce to the gluon beam function defined in Eq. (50),

$$
\omega_a \theta(\omega_a) \langle p_{n_a} | \mathcal{B}_{n_a\perp}^{\mu c}(0) \delta(b_a^+ - n_a \cdot \hat{p}_{n_a}) \delta(\omega_a - \bar{\mathcal{P}}_{n_a}) \mathcal{B}_{n_a\perp}^{\mu' c'}(0) | p_{n_a} \rangle = \frac{g_\perp^{\mu\mu'}}{2} \frac{\delta^{cc'}}{N_c^2 - 1} \omega_a B_g(\omega_a b_a^+, \omega_a / P_a^-). \tag{127}
$$

Including the color traces from the beam functions, the soft matrix element defines the gluonic incoming hemisphere soft function,

$$
S_{\rm ihemi}^{gg}(k_a^+, k_b^+) = \frac{1}{N_c^2 - 1} \langle 0 | \, {\rm tr}_{\rm color}\{ \bar{T}[\mathcal{Y}_{n_a}^\dagger(0) \mathcal{Y}_{n_b}(0)] \delta(k_a^+ - n_a \cdot \hat{p}_{a,s}) \delta(k_b^+ - n_b \cdot \hat{p}_{b,s}) T[\mathcal{Y}_{n_b}^\dagger(0) \mathcal{Y}_{n_a}(0)] \} | 0 \rangle, \tag{128}
$$

where the normalization is again convention. Putting everything together, the gluon contribution to the hadronic tensor becomes

$$
W_{JJ'gg}(q^2, Y, B_a^+, B_b^+) = H_{JJ'gg}(\tilde{b}_a, \tilde{b}_b) \int dk_a^+ dk_b^+ q^2 B_g[x_a E_{\rm cm}(B_a^+ - k_a^+), x_a] B_g[x_b E_{\rm cm}(B_b^+ - k_b^+), x_b] S_{\rm ihemi}^{gg}(k_a^+, k_b^+), \tag{129}
$$

with the hard function

$$H_{JJ'gg}(\tilde{b}_a,\tilde{b}_b)=\frac{1}{N_c^2-1}\frac{1}{2}(g_{\perp\mu\mu'}g_{\perp\nu\nu'}+g_{\perp\mu\nu'}g_{\perp\nu\mu'})C^{\dagger\nu\mu}_{Jgg}(\tilde{b}_a,\tilde{b}_b)C^{\mu'\nu'}_{J'gg}(\tilde{b}_a,\tilde{b}_b). \quad (130)$$

Here we have used the symmetry of the Wilson coefficients in Eq. (108) to simplify the four terms that arise from interchanging $a\leftrightarrow b$ in Eq. (126).

To obtain the full result for the hadronic tensor all we have to do now is to add up the contributions from the different quark flavors and the gluon,

$$W_{JJ'}(q^2,Y,B_a^+,B_b^+)=\sum_q W_{JJ'q\bar{q}}(q^2,Y,B_a^+,B_b^+)+W_{JJ'gg}(q^2,Y,B_a^+,B_b^+). \quad (131)$$

Inserting this back into Eq. (100), the final result for the factorized cross section becomes

$$\frac{d\sigma}{dq^2dYdB_a^+dB_b^+}=\sum_{ij}H_{ij}(q^2,Y)\int dk_a^+dk_b^+q^2\times B_i[x_aE_{\rm cm}(B_a^+-k_a^+),x_a]\times B_j[x_bE_{\rm cm}(B_b^+-k_b^+),x_b]\times S^{ij}_{\rm ihemi}(k_a^+,k_b^+), \quad (132)$$

with $x_{a,b}E_{\rm cm}=\sqrt{q^2}e^{\pm Y}$ as in Eqs. (96) and (124) and the hard function

$$H_{ij}(q^2,Y)=\frac{1}{2E_{\rm cm}^2}\sum_{J,J'}L_{JJ'}(q^2,Y)\times H_{JJ'ij}\left(x_aE_{\rm cm}\frac{n_a}{2},x_bE_{\rm cm}\frac{n_b}{2}\right). \quad (133)$$

The sum in Eq. (132) runs over parton species $ij=\{gg,u\bar{u},\bar{u}u,d\bar{d},\ldots\}$, where B_i is the beam function for parton i in beam a and B_j for parton j in beam b. Equation (132) is the final factorization theorem for the isolated $pp\to XL$ and $p\bar{p}\to XL$ processes. In Sec. IV D below we will apply it to the case of Drell-Yan, which will yield Eq. (19).

The beam functions in Eq. (132) are universal and take into account collinear radiation for isolated processes with x away from one. Since the soft function only depends on the color representation, but not on the specific quark flavor, there are only two independent soft functions $S^{q\bar{q}}_{\rm ihemi}$ and $S^{gg}_{\rm ihemi}$. In the sum over ij in Eq. (132), there are no mixed terms with ij corresponding to beam functions of two different quark flavors. Likewise, there are no mixed terms with quark and gluon beam functions. For example, a graph like Fig. 6(e) is part of the $ij=q\bar{q}$ term in the sum. Thus, cross terms between quark and gluon PDFs only appear via the contributions of different PDFs to a given beam function, as shown in Eq. (34).

The only process dependence in Eq. (132) arises through the hard functions $H_{ij}(q^2,Y)$, and one can study any desired leptonic observables by inserting the appropriate projections in the leptonic phase-space integrations inside $L_{JJ'}(q^2,Y)$. Since the hard function $H_{JJ'ij}$ corresponds to the partonic matrix element $\langle ij|J^\dagger|0\rangle\langle 0|J'|ij\rangle$ and $L_{JJ'}$ is given by the square of the relevant electroweak matrix elements $L_J^\dagger L_{J'}$, $H_{ij}(q^2,Y)$ can be determined from calculations of the partonic cross section $ij\to L$. Furthermore, $H_{ij}(q^2,Y)$ is identical to the hard function in threshold factorization theorems and hence in many cases is known from existing computations.

4. Cancellation of glauber gluons

In the above derivation we have implicitly assumed that contributions from Glauber gluons cancel in the final cross section, so that we do not need Glauber interactions in the effective theory. To complete the proof of factorization, we now argue that this is indeed the case.

In principle, Glauber interactions add an additional term $\mathcal{L}_G$ to the SCET Lagrangian

$$\mathcal{L}_{\rm SCET}=\mathcal{L}_{n_a}(\chi_{n_a},A_s)+\mathcal{L}_{n_b}(\chi_{n_b},A_s)+\mathcal{L}_s(A_s)+\mathcal{L}_G(A_G,\chi_{n_a},\chi_{n_b},A_s). \quad (134)$$

Glauber interactions in SCET have been considered in Refs. [83,84], but we will not require an explicit construction of $\mathcal{L}_G$ here. Our arguments will be based on the one hand, on the consistency with processes where it has been proven that Glauber interactions cancel, and on the other hand on systematic scale separation in the language of effective field theory. The scale separation is valid independently of whether it leads to a factorization into simple matrix elements, or whether it leads to a nonfactorizable matrix element with complicated dynamics.

The possible danger of the Glauber modes comes from the fact that they couple the two collinear sectors n_a and n_b with momentum scaling $Q(\lambda^2,1,\lambda)$ and $Q(1,\lambda^2,\lambda)$. With $\mathcal{L}_G$, there will still be interactions between soft and collinear modes present in the Lagrangian even after the field redefinition, so we cannot *a priori* factorize the full matrix element into independent soft and collinear matrix elements. Therefore, we have to revisit each step in our derivation with $\mathcal{L}_G$ in mind.

Our argument will be divided into three steps: (i) above the scale μ_B, (ii) at the scale μ_B, and (iii) below the scale μ_B. For (i) and (ii) we have to consider Glauber modes with momentum scaling $Q(\lambda^2,\lambda^2,\lambda)$, which we call G_1 modes. Since they have virtuality $\sim\mu_B^2$ they are integrated out at this scale. Any residual effects of Glauber interactions below μ_B could occur from modes with momentum

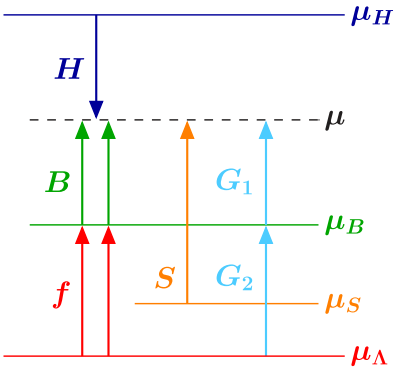


FIG. 14 (color online). RGE running including potential Glauber modes.

scaling $Q(\lambda^4, \lambda^4, \lambda^2)$, which we call G_2. These modes are illustrated in Fig. 14.

Above the hard scale $\mu_H \simeq Q$, we have full QCD and no distinction between different modes is required, so in step (i) we are concerned with contributions of G_1 in the region $\mu_H > \mu > \mu_B$. At the scale μ_H, we integrate out hard modes with virtualities Q^2 or higher by matching the QCD currents onto SCET currents. For our process, the leading operators are given in Eq. (107), which contain only one field per collinear direction. For the theory in Eq. (134), all other possible operators are power suppressed. The matching onto these currents is valid at an operator level and can be performed with quark and gluon states. It is independent of the hadronic matrix element we are going to take later on. The key point is that the exact same matching calculation and resulting Wilson coefficients C occur for threshold Drell-Yan and $e^+e^- \to 2$ jets. For these cases it is known [1,85] that G_1 modes do not affect the matching of the hard function $H \sim |C^2|$ at μ_H or the running of H in the region $\mu_H > \mu > \mu_B$ shown in Figs. 7(b) or 7(c). The hard function H gives a complete description of the physics down to the scale μ_B whether or not the modes in the SCET matrix elements factorize further. In Fig. 14, this corresponds to taking the scale $\mu = \mu_B$. Therefore, the G_1 modes can give neither large $\ln(\mu_B/\mu_H)$ terms nor finite contributions above μ_B.

In step (ii), we integrate out modes with virtualities $Q^2\lambda^2$ at the scale μ_B, which may involve matrix elements with G_1 modes exchanged. This matching affects the n_a-collinear, n_b-collinear, and G_1 modes, whose momentum scaling below μ_B changes to $Q(\lambda^4, 1, \lambda^2)$, $Q(1, \lambda^4, \lambda^2)$, and $Q(\lambda^4, \lambda^4, \lambda^2)$, respectively. Here we consider the theory right above μ_B including G_1 modes, leaving the discussion of the theory just below μ_B and G_2 modes to step (iii). Thus, we have to consider the matrix element of the composite operator

$$[\bar{\chi}_{n_a}\chi_{n_b}](x^+, x^-)\delta(B_a^+ - n_a\cdot\hat{p}_a)\delta(B_b^+ - n_b\cdot\hat{p}_b) \times \delta(\omega_a - \bar{\mathcal{P}}_{n_a})\delta(\omega_b - \bar{\mathcal{P}}_{n_b})[\bar{\chi}_{n_b}\chi_{n_a}](0), \tag{135}$$

where we suppressed all spin and color indices for simplicity, and these collinear fields still couple to soft fields in their Lagrangians. Since μ_B is a perturbative scale, we can carry out the matching onto the theory below μ_B at the operator level and do not yet have to consider proton states. Since the Glauber gluons are spacelike, they cannot cross the final-state cut indicated by the δ functions and only appear in virtual subdiagrams. We can therefore make a correspondence with the calculation in step (i) as follows. For any given final state with collinear and soft particles, the SCET computation for (ii) is identical to the SCET computation carried out for the matching in step (i) but using this particular choice of external states.[11] Since that SCET computation cannot induce any dependence on G_1 in step (i), there can also be no contributions from G_1 for the forward matrix-element computation here. The result of the step (ii) matching is thus given by a Wilson coefficient times an operator of the form

$$\int dk_a^+ dk_b^+ C(x^+, x^-, B_a^+ - k_a^+, B_b^+ - k_b^+) \times \bar{\chi}'_{n_a}(0)\bar{T}[Y^\dagger_{n_a}Y_{n_b}]\chi'_{n_b}(0)\delta(k_a^+ - n_a\cdot\hat{p}_a) \times \delta(k_b^+ - n_b\cdot\hat{p}_b)\delta(\omega_a - \bar{\mathcal{P}}_{n_a})\delta(\omega_b - \bar{\mathcal{P}}_{n_b}) \times \bar{\chi}'_{n_b}(0)T[Y^\dagger_{n_b}Y_{n_a}]\chi'_{n_a}(0), \tag{136}$$

where the primed collinear fields have scaling $Q(\lambda^4, 1, \lambda^2)$ and $Q(1, \lambda^4, \lambda^2)$, and the soft fields in the Y Wilson lines have scaling $Q(\lambda^2, \lambda^2, \lambda^2)$.

For step (iii) below μ_B, we have to consider the $\langle pp|\cdots|pp\rangle$ matrix element of Eq. (136) and possible contributions from G_2 Glauber gluons, which can now also connect to spectator lines in the proton (which are primed collinear modes). The G_2 gluons may spoil the factorization of the two collinear sectors. To argue that this is not the case, we rely heavily on the original proof of the cancellation of Glauber gluons for inclusive Drell-Yan in Ref. [28]. By construction, for our observables the $k^+_{a,b}$ variables in Eq. (136) are of $\mathcal{O}(Q\lambda^2)$ and thus only get contributions from the soft gluons. Hence, we are fully inclusive in the Hilbert space of the primed collinear fields. Therefore, the G_2 modes as well as possible "ultrasoft" $Q(\lambda^4, \lambda^4, \lambda^4)$ gluons cancel in the sum over states, just as in the inclusive case. This discussion for the cancellation of G_2 modes is identical to Ref. [29], where arguments were presented for the cancellation of G_2 gluons up to the scale induced by the measurement on the final state, which in our case is μ_B.

Physically, one could imagine that Glauber modes kick the spectators in the proton remnant such that they can contribute to $B^+_{a,b}$. The above arguments show that this is

[11] In practice one would never make such a complicated choice, but if one does, it must give the same result as picking a minimal state for the matching.

not the case, so that our treatment of the proton and its remnant in the derivation of the factorization is correct.

Note that the above arguments do not suffice to show that Glauber interactions cancel when there are additional hard central jets in the final state.

D. Final results for Drell-Yan

In this subsection, we present the final results for the isolated Drell-Yan cross section. Our discussion is split into four parts: the leptonic tensor, the hard function, the soft function, and the final cross section for beam thrust.

1. The leptonic tensor

To give an explicit example, we now apply the final factorization result in Eq. (132) to the Drell-Yan process with $L = \ell^+\ell^-$. The relevant QCD currents are the vector and axial-vector currents J^μ_{hf} with $h = \{V, A\}$, already given in Eq. (98). The corresponding leptonic contributions are

$$\begin{aligned} L^\mu_{Vf}(p_1, p_2) &= \frac{4\pi\alpha_{\rm em}}{q^2}\Big[-Q_f \bar u(p_2)\gamma^\mu v(p_1) \\ &\quad + \frac{v_f}{1 - m_Z^2/q^2}\bar u(p_2)\gamma^\mu(v_\ell - a_\ell\gamma_5)v(p_1)\Big], \\ L^\mu_{Af}(p_1, p_2) &= \frac{4\pi\alpha_{\rm em}}{q^2}\,\frac{-a_f}{1 - m_Z^2/q^2} \\ &\quad \times \bar u(p_2)\gamma^\mu(v_\ell - a_\ell\gamma_5)v(p_1), \end{aligned} \tag{137}$$

where in this subsection $p_1 = p_{\ell^+}$ and $p_2 = p_{\ell^-}$ are the lepton momenta, Q_f is the quark charge (in units of $|e|$), and $v_{\ell,f}$, $a_{\ell,f}$ are the standard vector and axial couplings of the leptons and quarks of flavor f to the Z boson. We will include the width of the Z later in Eq. (147).

The leptonic phase-space integral is

$$\begin{aligned} &\int \frac{d^4p_1 d^4p_2}{(2\pi)^2}\,\delta(p_1^2)\delta(p_2^2)\delta^4\Big(q^-\frac{n_a}{2} + q^+\frac{n_b}{2} - p_1 - p_2\Big) \\ &= \frac{1}{32\pi^2}\int \frac{d\Delta y d\varphi}{1 + \cosh\Delta y}, \end{aligned} \tag{138}$$

where φ is the azimuthal angle of the leptons in the transverse plane and Δy is the rapidity difference of the two leptons:

$$y_i = \frac{1}{2}\ln\frac{n_b\cdot p_i}{n_a\cdot p_i}, \qquad \Delta y = y_1 - y_2. \tag{139}$$

Since we expanded $\vec q_T = 0$, the leptons are back-to-back in the transverse plane, which implies that at the order we are working

$$\begin{aligned} &\frac{p_1^+}{p_2^-} = \frac{p_2^+}{p_1^-} = \frac{q^+}{q^-}, \qquad Y = \frac{1}{2}(y_1 + y_2), \\ &\vec p_{1T} = -\vec p_{2T}, \qquad \vec p_{1T}^{\,2} = \vec p_{2T}^{\,2} = \frac{q^2}{2(1 + \cosh\Delta y)}. \end{aligned} \tag{140}$$

Thus, the leptonic kinematics is described by the four independent variables $\{q^2, Y, \Delta y, \varphi\}$, with $\{Y, \Delta y\}$ being equivalent to $\{y_1, y_2\}$. For simplicity, we assume that we do not distinguish the two leptons, as one would, for example, by measuring their rapidities y_i or transverse momenta p_{iT}. We can then integrate over $0 \le \varphi \le 2\pi$ and $-\infty < \Delta y < \infty$ in Eq. (138), giving an overall factor of 4π. The leptonic tensor, Eq. (101), now becomes

$$\begin{aligned} L^{\mu\nu}_{hh'ff'}(q^2, Y) &= \frac{1}{32\pi^2}\int \frac{d\Delta y d\varphi}{1 + \cosh\Delta y}\sum_{\rm spins} L^{\dagger\mu}_{hf}(p_1, p_2) \\ &\quad \times L^\nu_{h'f'}(p_1, p_2) \\ &= \frac{8\pi\alpha_{\rm em}^2}{3q^2}\Big(\frac{q^\mu q^\nu}{q^2} - g^{\mu\nu}\Big)L_{hh'ff'}(q^2), \end{aligned} \tag{141}$$

where

$$\begin{aligned} L_{VVff'}(q^2) &= Q_f Q_{f'} - \frac{(Q_f v_{f'} + v_f Q_{f'})v_\ell}{1 - m_Z^2/q^2} \\ &\quad + \frac{v_f v_{f'}(v_\ell^2 + a_\ell^2)}{(1 - m_Z^2/q^2)^2}, \\ L_{AAff'}(q^2) &= \frac{a_f a_{f'}(v_\ell^2 + a_\ell^2)}{(1 - m_Z^2/q^2)^2}, \\ L_{AVff'}(q^2) &= \frac{-a_f}{1 - m_Z^2/q^2}\Big[-Q_{f'}v_\ell + \frac{v_{f'}(v_\ell^2 + a_\ell^2)}{1 - m_Z^2/q^2}\Big] \\ &= L_{VAf'f}(q^2). \end{aligned} \tag{142}$$

2. The hard function

Using parity and charge conjugation invariance of QCD, the matching coefficients for the vector and axial-vector QCD currents can be written as

$$\begin{aligned} C^{\mu\alpha\beta}_{Vfq\bar q}(\tilde b_a, \tilde b_b) &= C_{Vfq}(q^2)(\gamma_\perp^\mu)^{\alpha\beta}, \\ C^{\mu\alpha\beta}_{Afq\bar q}(\tilde b_a, \tilde b_b) &= C_{Afq}(q^2)(\gamma_\perp^\mu\gamma_5)^{\alpha\beta}, \\ C^{\mu\rho\sigma}_{Afgg}(\tilde b_a, \tilde b_b) &= C_{Ag}(q^2)(\tilde b_a + \tilde b_b)^\mu i\epsilon^{\rho\sigma}{}_{\lambda\kappa}\tilde b_a^\lambda \tilde b_b^\kappa. \end{aligned} \tag{143}$$

By Lorentz invariance (or reparametrization invariance of $n_{a,b}$ and $\bar n_{a,b}$ [86]), the scalar coefficients can only depend on $\tilde b_a\cdot\tilde b_b = x_a x_b E_{\rm cm}^2 = q^2$. In principle, parity and charge conjugation would also allow the Dirac structures $(\tilde b_a - \tilde b_b)^\mu\delta^{\alpha\beta}$ and $(\tilde b_a - \tilde b_b)^\mu(\gamma_5)^{\alpha\beta}$. However, as the vector and axial-vector currents are chiral even and the matching from QCD conserves chirality for massless quarks, these cannot be generated. For the gluon operator, the symmetry of the Wilson coefficient [see Eq. (108)] requires it to be proportional to $q^\mu = \tilde b_a^\mu + \tilde b_b^\mu$. Current conservation for the vector current requires $q_\mu C^\mu_{Vfq\bar q} = 0$, which eliminates this term. Thus, as expected, the only contribution for the gluon operator is due to the axial anomaly, coming from the diagram in Fig. 6(b). Since we neglect the lepton masses, $q_\mu L^\mu_{Af} = 0$, and thus C_{Afgg} does

not survive the contraction of the leptonic and hadronic tensors for $L = \ell^+\ell^-$. Hence, the gluon beam functions do not contribute to Drell-Yan, and the gluon PDF only appears through its contribution to the quark beam functions. Inserting Eq. (143) into the general expression for the hard function in Eq. (125), we obtain

$$\begin{aligned} H^{\mu\nu}_{hh'ff'q\bar{q}}(\tilde{b}_a, \tilde{b}_b) &= -\frac{1}{2N_c}\Big[g^{\mu\nu} - \frac{1}{2}(n_a^\mu n_b^\nu + n_a^\nu n_b^\mu)\Big] C^*_{hfq}(q^2) C_{h'f'q}(q^2) \qquad (\text{for } hh' = \{VV, AA\}), \\ H^{\mu\nu}_{hh'ff'q\bar{q}}(\tilde{b}_a, \tilde{b}_b) &= \frac{1}{4N_c} \mathrm{i}\epsilon^{\mu\nu}{}_{\lambda\kappa} n_a^\lambda n_b^\kappa C^*_{hfq}(q^2) C_{h'f'q}(q^2) \qquad (\text{for } hh' = \{VA, AV\}). \end{aligned} \tag{144}$$

At one loop, the vector and axial-vector coefficients are equal and diagonal in flavor and the SCET matching computation was performed in Refs. [80,87], in agreement with the one-loop form factors

$$\begin{aligned} C_{Vfq}(q^2) &= C_{Afq}(q^2) = \delta_{fq} C(q^2), \\ C(q^2, \mu) &= 1 + \frac{\alpha_s(\mu) C_F}{4\pi}\Big[-\ln^2\Big(\frac{-q^2 - \mathrm{i}0}{\mu^2}\Big) \\ &\quad + 3\ln\Big(\frac{-q^2 - \mathrm{i}0}{\mu^2}\Big) - 8 + \frac{\pi^2}{6}\Big]. \end{aligned} \tag{145}$$

The vector current coefficient at two loops was obtained in Refs. [12,88] from the known two-loop quark form factor [89–92]. Starting at three loops, it can have a contribution that is not diagonal in flavor, i.e., is not proportional to δ_{fq}. The axial-vector coefficient can also receive additional diagonal and nondiagonal contributions starting at two loops from the axial anomaly [93–95]. The anomaly contributions cancel in the final result in the sum over f as long as one sums over massless quark doublets. Therefore, they will cancel when the hard matching scale is much larger than the top-quark mass, in which case the top quark can be treated as massless. On the other hand, they have to be taken into account when the matching scale is below the top-quark mass, in which case the top quark is integrated out during the matching step and its mass cannot be neglected.

Combining Eqs. (144) and (145) with Eqs. (141) and (142), the coefficients $H_{ij}(q^2, Y)$ in Eq. (133) become

$$\begin{aligned} &\frac{1}{2E_{\mathrm{cm}}^2} \frac{8\pi\alpha_{\mathrm{em}}^2}{3q^2} \frac{1}{N_c} \sum_{ff'} [L_{VVff'}(q^2) C^*_{Vfq}(q^2) C_{Vf'q}(q^2) \\ &\quad + L_{AAff'}(q^2) C^*_{Afq}(q^2) C_{Af'q}(q^2)] \equiv \sigma_0 H_{q\bar{q}}(q^2, \mu), \end{aligned} \tag{146}$$

where at one loop

$$\sigma_0 = \frac{4\pi\alpha_{\mathrm{em}}^2}{3N_c E_{\mathrm{cm}}^2 q^2}, \qquad H_{q\bar{q}}(q^2, \mu) = H_{\bar{q}q}(q^2, \mu) = \Big[Q_q^2 + \frac{(v_q^2 + a_q^2)(v_\ell^2 + a_\ell^2) - 2Q_q v_q v_\ell (1 - m_Z^2/q^2)}{(1 - m_Z^2/q^2)^2 + m_Z^2\Gamma_Z^2/q^4}\Big] |C(q^2, \mu)|^2, \tag{147}$$

with $|C(q^2, \mu)|^2$ given by Eq. (145), and where we also included the nonzero width of the Z. The RGE for the hard function $H_{q\bar{q}}(q^2, \mu)$ is

$$\begin{aligned} \mu \frac{dH_{q\bar{q}}(q^2, \mu)}{d\mu} &= \gamma_H(q^2, \mu) H_{q\bar{q}}(q^2, \mu), \\ \gamma_H(q^2, \mu) &= 2\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)] \ln\frac{q^2}{\mu^2} + \gamma_H[\alpha_s(\mu)], \end{aligned} \tag{148}$$

where Γ_{cusp} is the universal cusp anomalous dimension [74], and the one-loop noncusp term is $\gamma_H[\alpha_s(\mu)] = -3\alpha_s(\mu) C_F/\pi$ [87]. The solution of Eq. (148) has the standard form

$$\begin{aligned} H_{q\bar{q}}(q^2, \mu) &= H_{q\bar{q}}(q^2, \mu_0) U_H(q^2, \mu_0, \mu), \\ U_H(q^2, \mu_0, \mu) &= e^{K_H(\mu_0, \mu)} \Big(\frac{q^2}{\mu_0^2}\Big)^{\eta_H(\mu_0, \mu)}, \end{aligned} \tag{149}$$

where $K_H(\mu_0, \mu)$ and $\eta_H(\mu_0, \mu)$ are analogous to Eq. (63),

$$\begin{aligned} K_H(\mu_0, \mu) &= \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)} \Big[-4\Gamma_{\mathrm{cusp}}(\alpha_s) \int_{\alpha_s(\mu_0)}^{\alpha_s} \frac{d\alpha_s'}{\beta(\alpha_s')} \\ &\quad + \gamma_H(\alpha_s)\Big], \\ \eta_H(\mu_0, \mu) &= 2\int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)} \Gamma_{\mathrm{cusp}}(\alpha_s). \end{aligned} \tag{150}$$

Together, Eqs. (149) and (150) sum the large logarithms occurring in isolated Drell-Yan between the scales μ_H and μ_B. Electroweak corrections to the hard function $H_{q\bar{q}}(q^2, \mu)$ can be included using the results of Refs. [14,96,97].

3. The $q\bar{q}$ soft function

The incoming hemisphere soft function contains incoming Wilson lines stretching from $-\infty$ to 0 along n_a and n_b. Under time reversal, each incoming Wilson line transforms into a corresponding outgoing Wilson line stretching from 0 to ∞ along the opposite direction,

$$\mathrm{T}^{-1} Y_{n_a} \mathrm{T} = \bar{P} \exp\Big[-ig \int_0^\infty ds\; n_b \cdot A_s(s n_b) \Big] = \tilde{Y}_{n_b}, \tag{151}$$

where $\bar{P}$ denotes anti-path ordering. Since T itself does not affect the original ordering of the field operators, time ordering turns into anti-time ordering and vice versa. In addition $\mathrm{T} n_a \cdot \hat{p}_{a,s} \mathrm{T}^{-1} = n_b \cdot \hat{p}_{b,s}$. Therefore, time-reversal invariance implies

$$\begin{aligned} S^{q\bar{q}}_{\mathrm{ihemi}}(k_a^+, k_b^+) &\overset{\mathrm{T}}{=} \frac{1}{N_c} \operatorname{tr}\langle 0|T[\tilde{Y}^\dagger_{n_b} \tilde{Y}_{n_a}] \delta(k_a^+ - n_b \cdot \hat{p}_{b,s}) \\ &\quad \times \delta(k_b^+ - n_a \cdot \hat{p}_{a,s}) \bar{T}[\tilde{Y}^\dagger_{n_a} \tilde{Y}_{n_b}]|0\rangle^* \\ &= \frac{1}{N_c} \operatorname{tr}\langle 0|T[\tilde{Y}^\dagger_{n_a} \tilde{Y}_{n_b}] \delta(k_a^+ - n_a \cdot \hat{p}_{a,s}) \\ &\quad \times \delta(k_b^+ - n_b \cdot \hat{p}_{b,s}) \bar{T}[\tilde{Y}^\dagger_{n_b} \tilde{Y}_{n_a}]|0\rangle. \end{aligned} \tag{152}$$

In the second step, the complex conjugation has no effect since the matrix element is real, and we used parity to switch $n_{b,a}$ back to $n_{a,b}$. For comparison, the hemisphere soft function with outgoing Wilson appearing in the double-differential hemisphere invariant-mass distribution in $e^+e^- \to 2$ jets [69,70,98–101] is

$$\begin{aligned} S^{q\bar{q}}_{\mathrm{hemi}}(k_a^+, k_b^+) &= \frac{1}{N_c} \operatorname{tr}\langle 0|\bar{T}[\tilde{Y}^\dagger_{n_a} \tilde{Y}_{n_b}] \delta(k_a^+ - n_a \cdot \hat{p}_{a,s}) \\ &\quad \times \delta(k_b^+ - n_b \cdot \hat{p}_{b,s}) T[\tilde{Y}^\dagger_{n_b} \tilde{Y}_{n_a}]|0\rangle. \end{aligned} \tag{153}$$

This is almost the same as Eq. (152), the only difference being the opposite time ordering. Thus, S_{ihemi} and S_{hemi} are equal at one loop, where the time ordering is still irrelevant. Beyond one loop, S_{ihemi} and S_{hemi} may in general be different. However, since the beam and jet functions have the same anomalous dimension, the combined anomalous dimension of the hard and beam functions in isolated Drell-Yan agrees with that of the hard and jet functions for the e^+e^- hemisphere invariant-mass distribution. The consistency of the RGE in both cases then requires that S_{ihemi} and S_{hemi} have the same anomalous dimension to all orders in perturbation theory. In addition, the purely virtual contributions, obtained by inserting the vacuum state, are the same in both cases,

$$\begin{aligned} S^{q\bar{q},\mathrm{virtual}}_{\mathrm{ihemi}}(k_a^+, k_b^+) &= \frac{1}{N_c} \delta(k_a^+) \delta(k_b^+) \operatorname{tr}|\langle 0|T[\tilde{Y}^\dagger_{n_a} \tilde{Y}_{n_b}]|0\rangle|^2 \\ &= S^{q\bar{q},\mathrm{virtual}}_{\mathrm{hemi}}(k_a^+, k_b^+). \end{aligned} \tag{154}$$

Hence, S_{ihemi} and S_{hemi} can only differ by finite real-emission corrections at each order in perturbation theory.

Using the one-loop results for $S^{q\bar{q}}_{\mathrm{hemi}}$ from Refs. [70,100], we have

$$\begin{aligned} S^{q\bar{q}}_{\mathrm{ihemi}}(k_a^+, k_b^+) &= \delta(k_a^+)\delta(k_b^+) + \delta(k_a^+) S^{1\mathrm{loop}}(k_b^+) \\ &\quad + S^{1\mathrm{loop}}(k_a^+)\delta(k_b^+), \\ S^{1\mathrm{loop}}(k^+) &= \frac{\alpha_s(\mu) C_F}{4\pi} \Big\{ -\frac{8}{\mu} \Big[\frac{\theta(k^+/\mu) \ln(k^+/\mu)}{k^+/\mu} \Big]_+ \\ &\quad + \frac{\pi^2}{6} \delta(k^+) \Big\}. \end{aligned} \tag{155}$$

The plus distribution is defined in Eq. (59). The one-loop soft function for beam thrust in Eq. (26) then becomes $S_B(k^+, \mu) = \delta(k^+) + 2S^{1\mathrm{loop}}(k^+)$.

4. Final cross section for beam thrust

The differential cross section for beam thrust in Eq. (32) including the RGE running is

$$\begin{aligned} \frac{d\sigma}{dq^2 dY d\tau_B} &= \sigma_0 \sum_{ij} H_{ij}(q^2, \mu_H) U_H(q^2, \mu_H, \mu_S) \\ &\quad \times \int dt_a dt_b Q S_B\Big(Q\tau_B - \frac{t_a + t_b}{Q}, \mu_S \Big) \\ &\quad \times \int dt'_a B_i(t_a - t'_a, x_a, \mu_B) U_B(t'_a, \mu_B, \mu_S) \\ &\quad \times \int dt'_b B_j(t_b - t'_b, x_b, \mu_B) U_B(t'_b, \mu_B, \mu_S). \end{aligned} \tag{156}$$

For simplicity, we evolve the hard and beam functions from their respective hard and beam scales, μ_H and μ_B, down to the common scale $\mu = \mu_S$ of the soft function. In this way, we do not need to consider the running of the soft function separately. Different choices for μ are all equivalent, as we discussed in Sec. II F. At LL, we include the one-loop cusp anomalous dimension in the evolution kernels U_H and U_B, and at NLL we include the two-loop cusp and one-loop noncusp anomalous dimensions. In both cases we use the LO results as initial conditions.

We also consider the fixed-order α_s expansion. To our knowledge $d\sigma/dq^2 dY d\tau_B$ has not been considered in perturbation theory in full QCD even at one loop. To obtain an expression for $d\sigma/dq^2 dY d\tau_B$ at NLO in α_s and leading order in the power counting, we drop the evolution factors U_H and U_B and expand all functions to NLO at a common scale μ. From the above NLO results for the hard and soft functions and the NLO results for the beam functions from Sec. III, we find

$$\begin{aligned} \frac{d\sigma}{dq^2 dY d\tau_B} &= \sigma_0 \sum_{i,j} \Big[Q_i^2 + \frac{(v_i^2 + a_i^2)(v_\ell^2 + a_\ell^2) - 2Q_i v_i v_\ell (1 - m_Z^2/q^2)}{(1 - m_Z^2/q^2)^2 + m_Z^2 \Gamma_Z^2/q^4} \Big] \\ &\quad \times \int \frac{d\xi_a}{\xi_a} \frac{d\xi_b}{\xi_b} C_{ij}\Big(\frac{x_a}{\xi_a}, \frac{x_b}{\xi_b}, q^2, \tau_B, \mu \Big) f_{i/a}(\xi_a, \mu) f_{j/b}(\xi_b, \mu). \end{aligned} \tag{157}$$

Here, $f_{i/a}(\xi_a, \mu)$ and $f_{j/b}(\xi_b, \mu)$ are the PDFs for parton i in proton a and parton j in (anti-)proton b. At tree level, the nonzero coefficients are

$$C^{\text{tree}}_{q\bar{q}}(z_a, z_b, q^2, \tau_B, \mu) = C^{\text{tree}}_{\bar{q}q}(z_a, z_b, q^2, \tau_B, \mu) = \delta(\tau_B)\delta(1 - z_a)\delta(1 - z_b). \tag{158}$$

At one loop, we obtain

$$\begin{aligned}
C^{\text{1loop}}_{q\bar{q}}(z_a, z_b, q^2, \tau_B, \mu) &= \frac{\alpha_s(\mu)C_F}{2\pi}\,\delta(1 - z_a)\theta(z_b)\Bigg\{\Bigg[-2\Bigg[\frac{\theta(\tau_B)\ln\tau_B}{\tau_B}\Bigg]_+ - \frac{3}{2}\Bigg[\frac{\theta(\tau_B)}{\tau_B}\Bigg]_+ - \delta(\tau_B)\Bigg(4 - \frac{\pi^2}{2}\Bigg)\Bigg]\delta(1 - z_b) \\
&\quad + \Bigg[\Bigg[\frac{\theta(\tau_B)}{\tau_B}\Bigg]_+ + \delta(\tau_B)\ln\frac{q^2}{\mu^2}\Bigg]\Bigg[\theta(1 - z_b)\frac{1 + z_b^2}{1 - z_b}\Bigg]_+ \\
&\quad + \delta(\tau_B)\Bigg[\Bigg[\frac{\theta(1 - z_b)\ln(1 - z_b)}{1 - z_b}\Bigg]_+(1 + z_b^2) + \theta(1 - z_b)\Bigg(1 - z_b - \frac{1 + z_b^2}{1 - z_b}\ln z_b\Bigg)\Bigg]\Bigg\} + (z_a \leftrightarrow z_b), \\
C^{\text{1loop}}_{\bar{q}q}(z_a, z_b, q^2, \tau_B, \mu) &= C^{\text{1loop}}_{q\bar{q}}(z_a, z_b, q^2, \tau_B, \mu), \\
C^{\text{1loop}}_{qg}(z_a, z_b, q^2, \tau_B, \mu) &= \frac{\alpha_s(\mu)T_F}{2\pi}\,\delta(1 - z_a)\theta(z_b)\theta(1 - z_b)\Bigg\{\Bigg[\Bigg[\frac{\theta(\tau_B)}{\tau_B}\Bigg]_+ + \delta(\tau_B)\ln\frac{q^2}{\mu^2}\Bigg][z_b^2 + (1 - z_b)^2] \\
&\quad + \delta(\tau_B)\Bigg[\ln\frac{1 - z_b}{z_b}[z_b^2 + (1 - z_b)^2] + 2z_b(1 - z_b)\Bigg]\Bigg\}, \\
C^{\text{1loop}}_{\bar{q}g}(z_a, z_b, q^2, \tau_B, \mu) &= C^{\text{1loop}}_{qg}(z_a, z_b, q^2, \tau_B, \mu), \\
C^{\text{1loop}}_{gq}(z_a, z_b, q^2, \tau_B, \mu) &= C^{\text{1loop}}_{g\bar{q}}(z_a, z_b, q^2, \tau_B, \mu) = C^{\text{1loop}}_{qg}(z_b, z_a, q^2, \tau_B, \mu).
\end{aligned} \tag{159}$$

The coefficient C_{gg} only starts to contribute at two loops. The single logarithms of q^2/μ^2 are multiplied by the QCD splitting kernels and are resummed by the PDFs. Thus, in fixed-order perturbation theory the PDFs should be evaluated at the hard scale $\mu = Q$, such that there are no large logarithms when integrating over $0 \le \tau_B \lesssim 1$. However, if the integration is restricted to $\tau_B \le \tau_B^{\text{cut}} \ll 1$, the plus distributions in τ_B produce large logarithms $\ln^2\tau_B^{\text{cut}}$ and $\ln\tau_B^{\text{cut}}$, which make a fixed-order expansion unreliable. These are precisely the logarithms that are resummed by the combined RGE of hard, jet, and soft functions in Eq. (156).

V. ISOLATED DRELL-YAN CROSS SECTION

In this section we illustrate our results with plots of the isolated Drell-Yan cross section. Rather than considering the cross section as a function of two variables, B_a^+ and B_b^+, we use the beam thrust τ_B defined in Sec. II B. We consider both the differential cross-section $d\sigma/dQdYd\tau_B$ as a function of τ_B, see Eq. (32), as well as the cross section integrated over $0 \le \tau_B \le \exp(-2y_B^{\text{cut}})$ as a function of y_B^{cut}, see Eqs. (31) and (33). We fix $Q = \sqrt{q^2}$ to a few representative values. We also restrict our discussion to back-to-back leptons, $Y = 0$, since measuring τ_B mainly affects the normalization and not the shape of the rapidity distribution.

For our cross-section predictions, we always use the NLO PDFs from MSTW2008 [77], with the corresponding $\alpha_s(m_Z) = 0.1202$ and two-loop five-flavor running for $\alpha_s(\mu)$.

We show results for the cross section both in the fixed-order expansion at LO and NLO [see Eq. (157)] and in the resummed expansions at LL and NLL [see Eq. (156)]. For the resummed results we always choose the hard, beam, and soft scales as $\mu_H = \mu$, $\mu_B = \mu\sqrt{\tau_B}$, $\mu_S = \mu\tau_B$, and for the fixed-order results we use a common fixed scale μ. The central values correspond to $\mu = Q$, and the bands show the scale variation obtained by varying $\mu = 2Q$ and $\mu = Q/2$. Since our purpose here is to illustrate the main features of the factorized cross section with beam functions, we will limit ourselves to the LL and NLL results without also including additional fixed-order corrections to the hard, beam, and soft functions. A complete analysis combining both NLO corrections and NNLL resummation, and a more detailed analysis of scale uncertainties, is left for future work.

In Fig. 15, we show results for pp collisions at the LHC with $E_{\text{cm}} = 7$ TeV and two different values $Q = 100$ GeV and $Q = 1$ TeV. For $Y = 0$ this corresponds to $x_a = x_b = 0.014$ and $x_a = x_b = 0.14$ respectively, representing two typical values for the isolated factorization theorem. In Fig. 16, we show results for $p\bar{p}$ collisions at the Tevatron for $Q = 500$ GeV corresponding to $x_a = x_b = 0.26$.

The top row in Fig. 15 depicts the LHC cross section as a function of τ_B at several different orders. At tree level, there is only a δ function and no radiation. The green

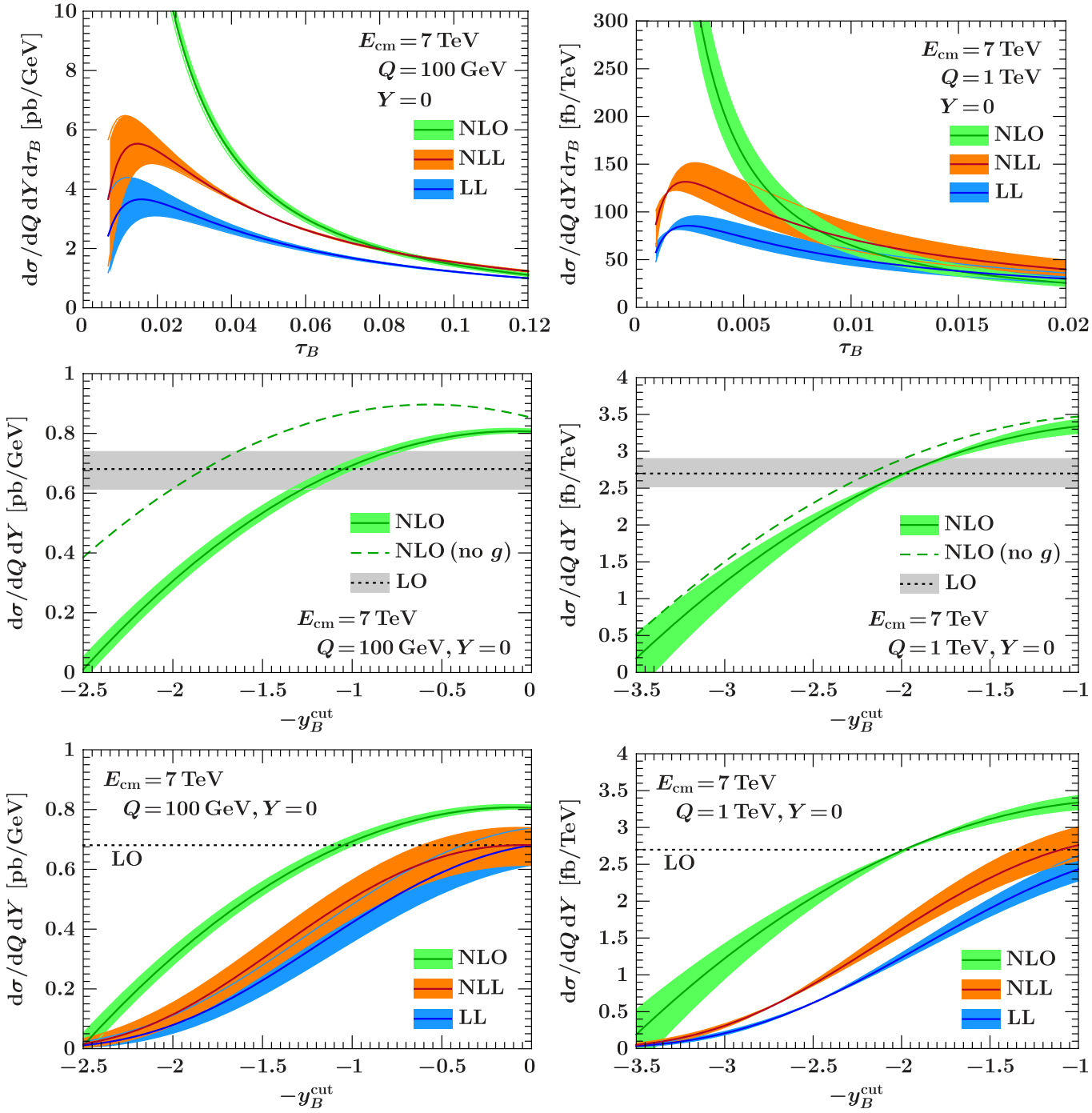


FIG. 15 (color online). Cross sections for beam thrust at the LHC with $E_{\rm cm} = 7$ TeV at $Y = 0$ and $Q = 100$ GeV (left column) and $Q = 1$ TeV (right column). Top row: The cross-section differential in τ_B at NLO, LL, and NLL. Middle and bottom rows: The cross section integrated up to $\tau_B \leq \exp(-2y_B^{\rm cut})$ at LO, NLO, LL, and NLL.

(light) curve and band show the NLO-expanded cross section, which grows for decreasing τ_B, showing that the radiation is peaked in the forward direction. As $\tau_B \to 0$ it exhibits the expected singular behavior, and this IR singularity is canceled in the integrated cross section by a corresponding δ function at $\tau_B = 0$. The blue (dark) and orange (medium) curves are the resummed results at LL and NLL, respectively. As expected, the resummation has a large effect at small τ_B and effectively regulates the IR singularity in the NLO result. The curves are not plotted for $\tau_B \leq 0.007$ in the left panel and $\tau_B \leq 0.001$ in the right panel, because at this point the soft scale drops below 1 GeV. Near this cutoff, the soft function becomes nonperturbative and so our purely perturbative results should

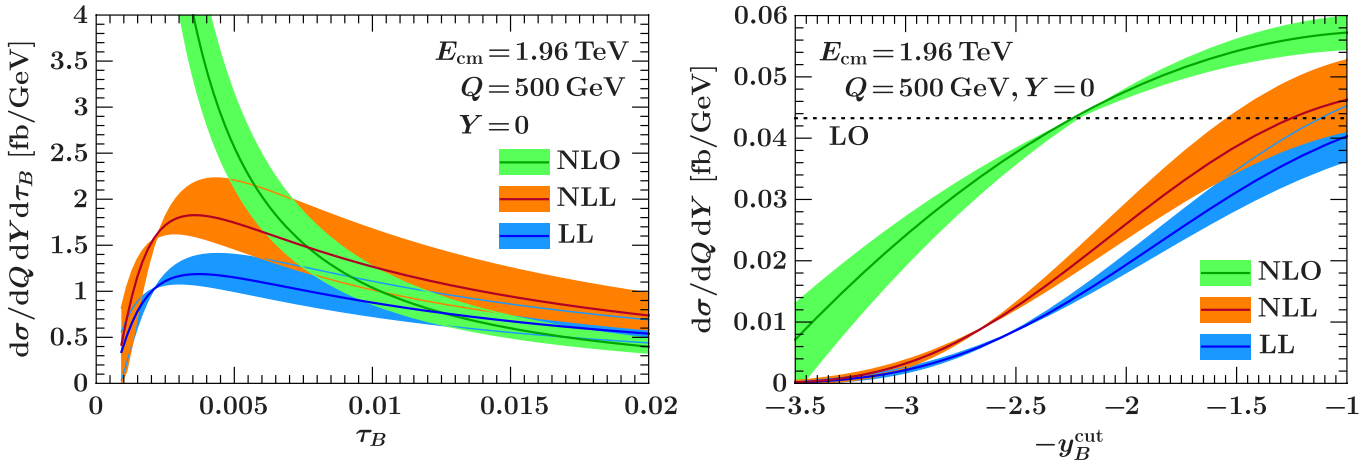


FIG. 16 (color online). The cross section for beam thrust at the Tevatron at $Y = 0$ and $Q = 500$ GeV. Left panel: The cross-section differential in τ_B at NLO, LL, and NLL. Right panel: The cross section integrated up to $\tau_B \leq \exp(-2y_B^{\rm cut})$ at LO, NLO, LL, and NLL.

not be taken too seriously. Nevertheless, it is interesting to see that the resummed results show a characteristic turnover that one would also expect from nonperturbative corrections. These nonperturbative corrections scale as $\Lambda_{\rm QCD}/(Q\tau_B)$ and hence become less relevant than the perturbative corrections for $\tau_B \gtrsim 0.02$ and $\tau_B \gtrsim 0.002$ in the left and right panels, respectively.

In the middle row of Fig. 15, we compare our results expanded to fixed-order at LO (gray horizontal band) and NLO (green band with solid central line) as a function of $y_B^{\rm cut}$. Since we plot the integrated cross section, the tree-level δ function gives a constant contribution in $y_B^{\rm cut}$. As expected, the NLO corrections to the LO result become very large as we move away from $y_B^{\rm cut} = 0$ because of the large double logarithms in the fixed-order expressions. Hence, a fixed-order expansion is not reliable here. The green dashed line illustrates what happens if we take the NLO result given by the solid line but exclude the contribution of the gluon PDF to the quark beam function. The gluon contribution has a bigger effect at smaller Q, because it corresponds to smaller x. Overall, it reduces the integrated cross section, which is the same effect we already observed in the integrated beam function in Sec. III. The plots in the bottom row of Fig. 15 show the LL, NLL, and NLO results, which are the integrated versions of the curves in the top row. Here, larger $y_B^{\rm cut}$ towards the left of the plot corresponds to smaller τ_B. A large $y_B^{\rm cut}$ implies a stronger constraint on the final state and hence a smaller cross section. In this region, the logarithmic resummation is necessary and suppresses the cross section. For $y_B^{\rm cut} \to 0$ the LL and NLL approach the LO result as they must.

Figure 16 shows the corresponding plots for the Tevatron. The left panel and right panel are equivalent to the top row and bottom row of Fig. 15, respectively, showing the differential and integrated cross sections. The plots show similar features overall, though the overall scale uncertainties are slightly larger here, because the PDFs are evaluated at a larger x. In the Tevatron case, even though we are at larger x compared to the right panels in Fig. 15, the cross section is larger due to the presence of valence antiquarks (note the different units in the two plots).

VI. CONCLUSIONS

Experimental measurements at the LHC or Tevatron are typically characterized by two conditions. First, the dominant part of the cross section arises from parton momentum fractions x away from one, and second, to probe the hard scattering the measurements impose restrictions on the final state to identify and isolate hard leptons or jets.

Factorization is required to separate the perturbatively calculable pieces from the nonperturbative parton distribution functions, and is a key ingredient for the resummation of large logarithms that occur due to phase-space restrictions. The most well-known factorization theorem for inclusive Drell-Yan applies for generic momentum fractions, but the hadronic final state is completely summed over, only subject to overall momentum conservation. This requires an inclusive experimental measurement, with only mild restrictions on the final state. On the other hand, threshold factorization theorems for Drell-Yan or dijet production take into account phase-space restrictions and resum resulting large logarithms, but they are only valid in the limit $x \to 1$. Thus each of these cases satisfies only one of the above experimental conditions.

In this paper we have studied factorization for generic x and with explicit restrictions on the hadronic final state. We considered the simplest situation, namely, Drell-Yan production $pp \to X\ell^+\ell^-$ at generic x with a restriction on X that vetoes hard central jets, which we call isolated Drell-Yan. The restriction on the hadronic final state is implemented by dividing the total hadronic momentum into two hemispheres and requiring the components B_a^+ and B_b^+ of the resulting hemisphere momenta to be small. For this

situation, we prove a factorization theorem for the cross-section differential in the hadronic variables B_a^+ and B_b^+ (as well as the leptonic phase-space variables). It allows us to systematically resum large double logarithms of $B_{a,b}^+/Q$ arising from the phase-space restrictions. The factorization theorem also applies to other isolated processes of the form $pp \to XL$, like $H \to \gamma\gamma$ or $H \to 4\ell$.

The main conclusion from our analysis is that PDFs alone are insufficient to properly describe the initial state in the collision. The restriction on the hadronic final state effectively probes the proton prior to the hard interaction by constraining the virtuality of the colliding hard parton. At that point, the parton cannot be confined to the proton anymore and must properly be described as part of an incoming initial-state jet. This description is given by quark and gluon beam functions, which replace the PDFs in the factorization theorem. In addition to the parton's momentum fraction, the beam functions depend on the virtuality of the colliding parton and live at an intermediate beam scale μ_B of the order of this virtuality. At the scale $\mu = \mu_B$, the beam functions can be matched onto PDFs evaluated at this μ times perturbatively calculable corrections. In this way, our factorization theorem unambiguously determines the proper scale at which the PDFs must be evaluated. This scale μ_B is much smaller than the hard scale of the partonic collision. As a result, the evolution of the initial state below μ_B is governed by the PDFs, while above μ_B it is governed by the beam function, whose evolution depends on the parton's virtuality rather than momentum fraction. Other differences compared to the PDF evolution are that the beam-function evolution sums double logarithms and does not involve parton mixing.

In a Monte Carlo setting the corresponding physical effects should be described by the initial-state parton shower. Our factorization theorem thus provides a way to explicitly check whether the initial-state parton shower resums the correct double logarithms, which is left for future work. We believe that experimental measurements of the isolated Drell-Yan spectrum will provide a direct method for testing the initial-state shower in Monte Carlo. This spectrum combined with the results reported here, is therefore useful for tuning the Monte Carlo with early LHC data.

Even though our derivation is only rigorous for the specific case of isolated Drell-Yan, we argued that the necessity for beam functions is more general, essentially applying to any process where the hadronic final state is restricted in a similar way. For example, the consistency of the renormalization group evolution implies that the description for any threshold process can be extended to a respective isolated case by supplementing it with corresponding variables $B_{a,b}^+$, replacing the PDFs by beam functions, and replacing the threshold soft function by an appropriate isolated soft function.

We briefly discussed the extension of isolated Drell-Yan to isolated dijet production. We also pointed out that if the hadronic final state is constrained by different global variables, one can expect to find different beam functions to encode these constraints, which will then sum double logarithms in these variables.

Ultimately, we hope that this type of factorization theorem with beam functions will bridge the gap between experimentally realistic cuts for LHC measurements and systematically improvable theoretical results that go beyond fixed-order calculations.

ACKNOWLEDGMENTS

We thank Matthew Schwartz for urging us to keep it simple. We also thank Carola Berger, Zoltan Ligeti, Andre Hoang, Aneesh Manohar, Ira Rothstein, Kerstin Tackmann, and Jesse Thaler for comments on the manuscript. I. S. and F. T. thank the MPI Munich and F. T. thanks the CERN theory group for hospitality during part of this work. This work was supported in part by the Office of Nuclear Physics of the U.S. Department of Energy under the Contract DE-FG02-94ER40818, and by Friedrich Wilhelm Bessel funding from the Alexander von Humboldt foundation.

[1] J. C. Collins, D. E. Soper, and G. Sterman, Adv. Ser. Dir. High Energy Phys. **5**, 1 (1988).
[2] ATLAS Collaboration, CERN Report No. CERN-LHCC-99-15, 1999.
[3] ATLAS Collaboration, CERN Report No. CERN-LHCC-99-14, 1999.
[4] G. L. Bayatian *et al.* (CMS Collaboration), CERN Report No. CERN-LHCC-2006-001.
[5] G. L. Bayatian *et al.* (CMS Collaboration), J. Phys. G **34**, 995 (2007).
[6] G. Sterman, Nucl. Phys. **B281**, 310 (1987).
[7] S. Catani and L. Trentadue, Nucl. Phys. **B327**, 323 (1989).
[8] N. Kidonakis, G. Oderda, and G. Sterman, Nucl. Phys. **B525**, 299 (1998).
[9] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, Nucl. Phys. **B529**, 424 (1998).
[10] E. Laenen, G. Oderda, and G. Sterman, Phys. Lett. B **438**, 173 (1998).
[11] S. Catani, D. de Florian, M. Grazzini, and P. Nason, J. High Energy Phys. 07 (2003) 028.
[12] A. Idilbi, X. dong Ji, and F. Yuan, Nucl. Phys. **B753**, 42 (2006).

[13] T. Becher, M. Neubert, and G. Xu, J. High Energy Phys. 07 (2008) 030.
[14] J.-y. Chiu, R. Kelley, and A. V. Manohar, Phys. Rev. D **78**, 073006 (2008).
[15] J. M. Campbell, J. W. Huston, and W. J. Stirling, Rep. Prog. Phys. **70**, 89 (2007).
[16] M. Trott, Phys. Rev. D **75**, 054011 (2007).
[17] C. W. Bauer, A. Hornig, and F. J. Tackmann, Phys. Rev. D **79**, 114013 (2009).
[18] A. Banfi, G. P. Salam, and G. Zanderighi, J. High Energy Phys. 08 (2004) 062.
[19] A. Banfi *et al.*, arXiv:hep-ph/0508096.
[20] M. Dasgupta and G. P. Salam, Phys. Lett. B **512**, 323 (2001).
[21] C. F. Berger, T. Kucs, and G. Sterman, Phys. Rev. D **65**, 094031 (2002).
[22] R. B. Appleby and M. H. Seymour, J. High Energy Phys. 09 (2003) 056.
[23] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, J. High Energy Phys. 08 (2006) 059.
[24] C. W. Bauer, S. Fleming, and M. E. Luke, Phys. Rev. D **63**, 014006 (2000).
[25] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, Phys. Rev. D **63**, 114020 (2001).
[26] C. W. Bauer and I. W. Stewart, Phys. Lett. B **516**, 134 (2001).
[27] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002).
[28] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B308**, 833 (1988).
[29] S. M. Aybat and G. Sterman, Phys. Lett. B **671**, 46 (2009).
[30] S. Fleming, A. K. Leibovich, and T. Mehen, Phys. Rev. D **74**, 114004 (2006).
[31] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).
[32] H. Georgi and H. D. Politzer, Phys. Rev. D **9**, 416 (1974).
[33] D. J. Gross and F. Wilczek, Phys. Rev. D **9**, 980 (1974).
[34] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
[35] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
[36] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.
[37] T. Sjöstrand, S. Mrenna, and P. Skands, Comput. Phys. Commun. **178**, 852 (2008).
[38] G. Corcella *et al.*, J. High Energy Phys. 01 (2001) 010.
[39] M. Bahr *et al.*, Eur. Phys. J. C **58**, 639 (2008).
[40] T. Sjöstrand and P. Z. Skands, J. High Energy Phys. 03 (2004) 053.
[41] T. Sjöstrand and P. Z. Skands, Eur. Phys. J. C **39**, 129 (2005).
[42] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. C **72**, 637 (1996).
[43] M. Bahr, S. Gieseke, and M. H. Seymour, J. High Energy Phys. 07 (2008) 076.
[44] A. A. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **65**, 092002 (2002).
[45] D. E. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **70**, 072002 (2004).
[46] D. Kar, FERMILAB Report No. FERMILAB-THESIS-2008-54, 2008.
[47] T. Sjöstrand, Phys. Lett. B **157**, 321 (1985).
[48] M. Bengtsson, T. Sjöstrand, and M. van Zijl, Z. Phys. C **32**, 67 (1986).
[49] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Phys. Rep. **100**, 1 (1983).
[50] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, arXiv:1002.2213.
[51] G. T. Bodwin, Phys. Rev. D **31**, 2616 (1985).
[52] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B261**, 104 (1985).
[53] G. Altarelli, R. K. Ellis, and G. Martinelli, Nucl. Phys. **B157**, 461 (1979).
[54] R. Hamberg, W. L. van Neerven, and T. Matsuura, Nucl. Phys. **B359**, 343 (1991).
[55] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002).
[56] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. Lett. **91**, 182002 (2003).
[57] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. D **69**, 094008 (2004).
[58] L. Magnea, Nucl. Phys. **B349**, 703 (1991).
[59] G. P. Korchemsky and G. Marchesini, Nucl. Phys. **B406**, 225 (1993).
[60] S. Catani, M. L. Mangano, P. Nason, and L. Trentadue, Nucl. Phys. **B478**, 273 (1996).
[61] A. V. Belitsky, Phys. Lett. B **442**, 307 (1998).
[62] S. Moch and A. Vogt, Phys. Lett. B **631**, 48 (2005).
[63] G. P. Korchemsky, arXiv:hep-ph/9610207.
[64] C. F. Berger, T. Kucs, and G. Sterman, Phys. Rev. D **68**, 014012 (2003).
[65] G. Sterman, Phys. Rev. D **17**, 2789 (1978).
[66] A. H. Hoang and I. W. Stewart, Phys. Lett. B **660**, 483 (2008).
[67] Z. Ligeti, I. W. Stewart, and F. J. Tackmann, Phys. Rev. D **78**, 114014 (2008).
[68] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, Nucl. Phys. **B407**, 3 (1993).
[69] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, Phys. Rev. D **77**, 074010 (2008).
[70] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, Phys. Rev. D **77**, 114003 (2008).
[71] A. V. Manohar and I. W. Stewart, Phys. Rev. D **76**, 074002 (2007).
[72] C. Lee and G. Sterman, Phys. Rev. D **75**, 014022 (2007).
[73] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. D **66**, 014017 (2002).
[74] G. P. Korchemsky and A. V. Radyushkin, Nucl. Phys. **B283**, 342 (1987).
[75] C. Balzereit, T. Mannel, and W. Kilian, Phys. Rev. D **58**, 114029 (1998).
[76] M. Neubert, Eur. Phys. J. C **40**, 165 (2005).
[77] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009).
[78] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. D **67**, 071502(R) (2003).
[79] C. W. Bauer, A. V. Manohar, and M. B. Wise, Phys. Rev. Lett. **91**, 122001 (2003).
[80] C. W. Bauer, C. Lee, A. V. Manohar, and M. B. Wise, Phys. Rev. D **70**, 034014 (2004).
[81] J. Chay, C. Kim, Y. G. Kim, and J.-P. Lee, Phys. Rev. D **71**, 056001 (2005).

[82] C. M. Arnesen, J. Kundu, and I. W. Stewart, Phys. Rev. D **72**, 114002 (2005).
[83] A. Idilbi and A. Majumder, Phys. Rev. D **80**, 054022 (2009).
[84] J. F. Donoghue and D. Wyler, arXiv:0908.4559.
[85] J. C. Collins and G. Sterman, Nucl. Phys. **B185**, 172 (1981).
[86] A. V. Manohar, T. Mehen, D. Pirjol, and I. W. Stewart, Phys. Lett. B **539**, 59 (2002).
[87] A. V. Manohar, Phys. Rev. D **68**, 114019 (2003).
[88] T. Becher, M. Neubert, and B. D. Pecjak, J. High Energy Phys. 01 (2007) 076.
[89] G. Kramer and B. Lampe, Z. Phys. C **34**, 497 (1987); **42**, 504(E) (1989).
[90] T. Matsuura and W. L. van Neerven, Z. Phys. C **38**, 623 (1988).
[91] T. Matsuura, S. C. van der Marck, and W. L. van Neerven, Nucl. Phys. **B319**, 570 (1989).
[92] T. Gehrmann, T. Huber, and D. Maitre, Phys. Lett. B **622**, 295 (2005).
[93] B. A. Kniehl and J. H. Kuhn, Phys. Lett. B **224**, 229 (1989).
[94] B. A. Kniehl and J. H. Kuhn, Nucl. Phys. **B329**, 547 (1990).
[95] W. Bernreuther *et al.*, Nucl. Phys. **B723**, 91 (2005).
[96] J.-y. Chiu, F. Golf, R. Kelley, and A. V. Manohar, Phys. Rev. D **77**, 053004 (2008).
[97] J.-y. Chiu, A. Fuhrer, R. Kelley, and A. V. Manohar, Phys. Rev. D **80**, 094013 (2009).
[98] G. P. Korchemsky and G. Sterman, Nucl. Phys. **B555**, 335 (1999).
[99] G. P. Korchemsky and S. Tafat, J. High Energy Phys. 10 (2000) 010.
[100] M. D. Schwartz, Phys. Rev. D **77**, 014026 (2008).
[101] A. H. Hoang and S. Kluth, arXiv:0806.3852.