Factorization of power corrections in the Drell-Yan process in EFT

Matthew Inglis-Whalen [*], Michael Luke [†], Jyotirmoy Roy [‡] and Aris Spourdalakis [§]

Department of Physics, University of Toronto, Toronto, Ontario M5S 1A7, Canada

We examine the quark-induced Drell-Yan process at next-to-leading power (NLP) in soft-collinear effective theory. Using an approach with no explicit soft or collinear modes, we discuss the factorization of the differential cross section in the small-q_T hierarchy with $q^2 \gg q_T^2 \gg \Lambda_{\text{QCD}}^2$. We show that the cross section may be written in terms of matrix elements of power-suppressed operators $T_{(i,j)}$, which contribute to $O(q_T^2/q^2)$ coefficients of the usual parton distribution functions. We derive a factorization for this observable at NLP which allows the large logarithms in each of the relevant factors to be resummed. We discuss the cancellation of rapidity divergences and the overlap subtractions required to eliminate double counting at next-to-leading power.

I. INTRODUCTION

The Drell-Yan (DY) process $N_1(P_1)N_2(P_2) \rightarrow \gamma^*(q) + X \rightarrow (\ell\bar{\ell}) + X$ has been extensively studied in perturbative QCD [1–3]. In the limiting case that the transverse momentum q_T of the lepton pair is parametrically larger than Λ_{QCD} and smaller than its invariant mass $\sqrt{q^2}$, the cross section may be written as [4]

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dy dq_T^2} = \sum_{a,b} \int \frac{dz_1 dz_2}{z_1 z_2} C_{ff}^{ab}(z_1, z_2, q^2, q_T^2)$$
$$\times f_{a/N_1}\left(\frac{\xi_1}{z_1}\right) f_{b/N_2}\left(\frac{\xi_2}{z_2}\right) + O\left(\frac{\Lambda_{\text{QCD}}^2}{q_T^2}\right), \quad (1)$$

where q^μ is the four-momentum of the lepton pair, $\xi_1 \equiv q^-/P_1^-$, $\xi_2 \equiv q^+/P_2^+$, $y = \log(q^-/q^+)/2$, and P_1^- and P_2^+ are the large light-cone components of the incoming hadron momenta. The sum is over parton types a, b, and the f_i are the usual parton distribution functions (PDFs). In this paper we only study the quark-induced process for a single flavor of quark, so we define $C_{ff} \equiv C_{ff}^{q\bar{q}}$.

The coefficient function C_{ff} may be expanded in powers of q_T^2/q^2,

$$C_{ff}(z_1, z_2, q^2, q_T^2) = C_{ff}^{(0)}(z_1, z_2, q^2, q_T^2)$$
$$+ \frac{1}{q^2} C_{ff}^{(2)}(z_1, z_2, q^2, q_T^2) + \cdots. \quad (2)$$

where each subsequent term is suppressed by increasing powers of q_T^2/q^2. Since they depend on two parametrically different scales q^2 and q_T^2, the fixed-order perturbative expansions for each $C_{ff}^{(n)}$ contain large logarithms of q_T^2/q^2 which can spoil the behavior of perturbation theory and need to be resummed. The resummation of the leading power (LP) term $C_{ff}^{(0)}$ has been extensively studied in the literature, using both perturbative QCD techniques [4–15] and effective field theory methods [16–19]. Factorization theorems allow $C_{ff}^{(0)}$ to be written as a product of separate terms depending on distinct scales, each of which may be resummed to arbitrary order using a variety of renormalization group (RG) or related techniques. The most recent analyses achieve a resummation up to $\text{N}^3\text{LL} + \text{NNLO}$ order [20–26]. However, much less is known about the factorization and resummation properties of the first power correction $C_{ff}^{(2)}$. C_{ff} has been computed in QCD at fixed order in perturbation theory up to N^2LO [27,28], but an all-orders RG resummation at next-to-leading power (NLP) has not been performed.

Soft-collinear effective theory (SCET) [29–35] is an effective field theory (EFT) that provides a systematic framework in which to study power corrections in hard QCD processes. There has been much recent work studying power corrections to various processes, with applications including beam thrust [36], Drell-Yan production near threshold [37], threshold Higgs production from gluon fusion [38], Higgs production and decay [39], the

energy-energy correlator in $\mathcal{N} = 4$ Supersymmetric Yang-Mills [40], and Higgs to diphoton decays [41–43]. Power corrections have also been studied using non-EFT QCD techniques [44–50].

The DY process at small q_T^2 is typically referred to as a $SCET_{II}$ process, characterized by collinear and soft modes in the EFT, and exhibiting rapidity logarithms in matrix elements. Rapidity logarithms are large logarithms in matrix elements which arise in SCET due to divergences in individual diagrams at large values of the rapidity of one of the particles. These divergences cancel between graphs with different modes, but the final result contains large finite logarithms of the hard scale of the scattering which cannot be resummed using usual RG techniques. Rapidity divergences require an additional regulator beyond dimensional regularization, and various techniques have been successfully employed to handle the rapidity resummation, including off-the-light-cone techniques [4], the rapidity renormalization group [51], the collinear anomaly framework [17], the exponential regulator [52], and the recently proposed pure rapidity regulator [40,53]. The latter regulator has recently been used [53] to calculate the small-q_T DY cross section by expanding the QCD graphs in the soft and collinear limits, where it correctly treats the power-law rapidity divergences arising at NLP. The connection between rapidity renormalization in $SCET_{II}$ and the usual renormalization group equation (RGE) in $SCET_I$ was discussed in [54].

In this paper we study power corrections to DY production using the version of SCET developed in [55–57]. In this approach the degrees of freedom in the EFT are not analyzed using the method of regions [58] in which they are explicitly separated into soft, collinear, ultrasoft, and possibly additional modes. Instead, states are separated into distinct sectors, where the relative invariant mass of particles within each sector is less than the renormalization scale μ of the EFT, but the relative invariant mass of different sectors is larger than the renormalization scale. As with the mode expansion, particles of the same type but in different sectors are described by different fields; however, interactions within a sector are described by QCD, while interactions between sectors are mediated via the external current, which is expanded in inverse powers of the hard matching scale. Factorization of different modes (soft-collinear, ultrasoft-collinear, and others) does not occur explicitly in the Lagrangian since different modes in a given sector are described by the same fields, but instead arises through the usual EFT process of integrating out degrees of freedom and matching onto a new EFT at appropriate threshold scales.

This reduces the number of separate fields in the Lagrangian and therefore simplifies the formalism, both conceptually and practically. One immediate feature is that subleading terms in the effective Lagrangian coupling different modes and violating manifest factorization are not present in this approach. In addition, rather than

deriving a factorization theorem in terms of jet and soft functions which are individually well-defined and renormalized at the appropriate scale, the rate is simply expressed in terms of bilocal products of operators in the EFT which may be run both in the renormalization scale μ as well as in the rapidity scale ν. Similar to the situation at LP discussed in [57], we show here that the DY cross section naturally factorizes into hard matching coefficients, rapidity evolution factors, soft matching coefficients, and parton distribution functions, and give expressions for the first three quantities up to NLP at one loop. The complete resummation of rapidity logarithms is left for a future work.

Consistency of this theory requires that double counting of degrees of freedom between the two sectors is consistently subtracted, similar to the usual zero-bin subtraction [59] in SCET. This procedure of overlap subtraction is necessary for the theory to be well-defined and is implicit in all matrix elements. Furthermore, as discussed in detail in [57], the scheme dependence of this subtraction allows rapidity logarithms to be summed using techniques similar to [51,53] without having manifest factorization of soft and collinear modes in the effective Lagrangian. At subleading powers this subtraction is nontrivial, requiring contributions from multiple operators as well as subleading corrections to the leading power subtraction. While these subtractions vanish using an appropriately chosen regulator, the interplay of these subtraction terms explains patterns of rapidity divergence cancellation between different operators, similar to the nontrivial cancellations of rapidity and endpoint divergences at NLP seen in other approaches [41–43].

QCD proofs of factorization in hard scattering processes require that the effects of the exchange of soft gluons in the Glauber regime relevant to small angle parton scattering cancel in the relevant observable [1–3,60,61]. Glauber modes have been the subject of much recent interest in SCET [62], and a consistent treatment of gluons in the Glauber regime has been shown to be necessary to ensure that operator statements in SCET are independent of the external states [63]. Investigation of these effects in the formalism presented here are beyond the scope of this paper, but we will assume that gluons in the Glauber regime do not introduce factorization-violating effects in the context of this calculation.

In Sec. I A we sketch the ingredients of the calculation and the approach to factorization in this formalism. We present the one-loop calculations of the various pieces in Sec. II and compare our fixed-order results with the unsummed QCD result. In Sec. III we consider the cross section with no rapidity regulator to demonstrate the cancellation of rapidity divergences between different operators and their respective overlap subtractions across different regions of phase space. We present our conclusions in Sec. IV. A few details of plus distributions used here are given in the appendixes, as well as a comparison to a recent one-loop analysis [53] of power corrections to the DY process.

A. Factorization

In the SCET formalism introduced in [55–57] there are no explicit soft, collinear, or ultrasoft modes, so factorization does not arise explicitly from a Lagrangian mode expansion, but instead by integrating out ultraviolet degrees of freedom at the relevant matching scales. In this section we briefly review the approach of [57] to DY scattering and introduce its extension to subleading power. Precise definitions of quantities appearing in this section will be given in Sec. II.

The cross section for the electromagnetic Drell-Yan production process, $N_1 N_2 \to \gamma^* + X \to (\ell\bar{\ell}) + X$, is given in QCD by

$$d\sigma = \frac{4\pi\alpha^2}{3q^2 s} \frac{d^4 q}{(2\pi)^4} \int d^d x\, e^{-iq\cdot x}(-g_{\mu\nu})$$
$$\times \langle N_{12}|J_{\text{QCD}}^{\mu\dagger}(x)J_{\text{QCD}}^\nu(0)|N_{12}\rangle, \qquad (3)$$

where $q^2 = (p_\ell + p_{\bar\ell})^2$ is the invariant mass of the lepton pair, $s = (P_1 + P_2)^2$ is the invariant mass of the incoming hadrons, the initial hadronic state is $|N_{12}\rangle = |N_1(P_1)N_2(P_2)\rangle$, and the vector QCD current J_{QCD}^μ is

$$J_{\text{QCD}}^\mu(x) = \bar\psi(x)\gamma^\mu\psi(x) \qquad (4)$$

for a single flavor of light quark. The extension to electroweak currents is straightforward [4,17].

For $q_T^2 \ll q^2$, perturbative corrections to the cross section in Eq. (3) contain powers of logarithms of q_T^2/q^2, which can spoil the apparent convergence of perturbation theory. SCET provides a systematic approach to resumming these terms. At the renormalization scale $\mu = \mu_H \sim \sqrt{q^2} \gg q_T$, hard interactions are integrated out of the theory and QCD is matched onto SCET. In the formalism used here, SCET consists of two decoupled QCD sectors, denoted by the lightlike vectors n^μ and $\bar{n}^\mu$, with total momenta p_n^μ and $p_{\bar n}^\mu$; the sectors are distinguished by the power counting

$$p_n^2, p_{\bar n}^2 \ll q^2, \qquad p_n \cdot p_{\bar n} \sim q^2. \qquad (5)$$

Interactions between the sectors are mediated by the external current J_{SCET}^μ, which is written as a sum of operators of increasing dimension,[1]

$$J_{\text{SCET}}^\mu(x) = \sum_i \frac{1}{q_L^{[i]}} C_2^{(i)}(\mu)O_2^{(i)\mu}(x,\mu), \qquad (6)$$

where an operator $O_2^{(i)}$ has mass dimension $[i]$ in excess of the leading-power operator $O_2^{(0)}$. We have defined

---

[1]Subleading operators are also labeled by continuous indices, so the discrete sums over operators also include integrals, which we neglect for simplicity in this section.

$q_L^2 \equiv q^+ q^-$, and for brevity we will not explicitly include the μ dependence of operators in subsequent equations unless required for clarity. It is convenient to expand in inverse powers of q_L^2 rather than $q^2 = q_L^2 - q_T^2$ so that the hard scale of the EFT is independent of the infrared (IR) scale q_T^2. This expansion has been performed up to $O(1/q_L^2)$ [55,64–67], the details of which are summarized in Sec. II A. The SCET expansion for the differential cross section is then given in SCET by

$$\frac{d\sigma}{dq^2 dy dq_T^2} = \frac{4\pi\alpha^2}{3q^2 s}(-g_{\mu\nu})\int \frac{d\Omega_T}{2}\int \frac{d^d x}{2(2\pi)^d}\sum_{ij}\frac{H_{(i,j)}(\mu)}{q_L^{[i]+[j]}}$$
$$\times e^{-iq\cdot x}\langle N_{12}|O_2^{(i)\mu\dagger}(x)O_2^{(j)\nu}(0)|N_{12}\rangle, \qquad (7)$$

where $H_{(i,j)}(\mu) \equiv C_2^{(i)\dagger}(\mu)C_2^{(j)}(\mu)$ and the final angular integral $d\Omega_T$ corresponds to the angular integral in the transverse momentum $\mathbf{q}_T$. Since we have not subdivided the degrees of freedom of SCET into separate soft and collinear modes, there is no expansion of the SCET Lagrangian beyond that in Eq. (6); in particular, there are no power corrections arising from soft-collinear mixing terms in the Lagrangian [34,35,68–71]. This simplifies the analysis of power corrections considerably.

While matrix elements of the operator products in Eq. (7) may be directly evaluated between partons in perturbation theory, it is convenient to perform a Fierz rearrangement to write the operator product as a convolution of transverse momentum dependent distribution operators (whose hadronic matrix elements are generally referred to as TMDPDFs), one in the n sector and one in the $\bar{n}$ sector. This is a standard procedure at leading power [2,17]; at subleading powers a similar procedure may be used to express the basis of operator products as convolutions of power-suppressed distribution operators,

$$\int \frac{d^d x}{2(2\pi)^d}(-g_{\mu\nu})e^{-iq\cdot x}O_2^{(i)\mu\dagger}(x)O_2^{(j)\nu}(0)$$
$$= \sum_{k,\ell}\frac{1}{N_c}K_{(k,\ell)}^{(i,j)}T_{(k,\ell)}(q^-, q^+, \mathbf{q}_T)$$
$$+ \text{spin dependent}, \qquad (8)$$

where each $T_{(i,j)}$ relevant to this calculation will be defined explicitly in Sec. II. Rewriting the operator products in terms of the operators $T_{(i,j)}$ is simply a change of operator basis, and not a matching condition or expansion in SCET, and so introduces no new perturbative corrections. Typically in SCET this Fierz rearrangement is performed to write the operator product in a form that manifestly factorizes into jet and soft functions; since this factorization is not needed here this change of basis is not strictly necessary, but it is included here for easier comparison with other approaches.

At $O(\alpha_s)$ matrix elements of the $T_{(i,j)}$'s at small q_T are insensitive to the cutoff scale q_L and so running the scattering operators $O_{2,(i)}$ from $\mu_H \sim q_L$ to $\mu_S \sim q_T$ sums the usual renormalization group logarithms of q_L^2/q_T^2 in the rate. If $q_T \sim \Lambda_{\mathrm{QCD}}$, matrix elements of each $T_{(i,j)}$ are nonperturbative quantities which would have to be either modeled or extracted from experiment. In the scaling of interest here, $q_T \gg \Lambda_{\mathrm{QCD}}$, each $T_{(i,j)}$ may be further expanded in powers $\Lambda_{\mathrm{QCD}}^2/q_T^2$, allowing the operator product in Eq. (7) to be matched onto the usual light-cone distribution operators whose hadronic matrix elements are the parton distribution functions. This expansion corresponds to matching SCET onto a soft theory of completely decoupled sectors of QCD at the scale $\mu_S \sim q_T$, and at leading twist takes the form

$$
T_{(k,\ell)}(q^-,q^+,\mathbf{q}_T,\mu_S) \to \int \frac{dz_1}{z_1}\frac{dz_2}{z_2} C_{S,(k,\ell)}(z_1,z_2,\mathbf{q}_T,\mu_S)
$$
$$
\times O_q\left(\frac{q^-}{z_1},\mu_S\right) O_{\bar{q}}\left(\frac{q^+}{z_2},\mu_S\right), \quad (9)
$$

where the various $C_{S,(k,\ell)}$ are matching coefficients and the hadronic matrix elements of the light-cone quark and antiquark distribution operators $O_{q,\bar{q}}$ are the usual spin-averaged parton distribution functions

$$
f_{q/N_1}(\zeta_1) = \langle N_1(P_1)|O_q(\zeta_1 P_1^-)|N_1(P_1)\rangle,
$$
$$
f_{\bar{q}/N_2}(\zeta_2) = \langle N_2(P_2)|O_{\bar{q}}(\zeta_2 P_2^+)|N_2(P_2)\rangle, \quad (10)
$$

with $P_i^- \equiv P_i \cdot \bar{n}$ and $P_i^+ \equiv P_i \cdot n$. Combining these matching steps gives an expression for the DY cross section for a single quark flavor of the form

$$
\frac{d\sigma}{dq^2 dy dq_T^2} = \sigma_0 \int \frac{dz_1}{z_1}\frac{dz_2}{z_2} C_{ff}(z_1,z_2,q_L^2,q_T^2)
$$
$$
\times f_{q/N_1}\left(\frac{\xi_1}{z_1}\right) f_{\bar{q}/N_2}\left(\frac{\xi_2}{z_2}\right) + \cdots, \quad (11)
$$

where $\sigma_0 = 4\pi\alpha^2/(3N_c q^2 s)$, $\xi_1 = q^-/P_1^-$, $\xi_2 = q^+/P_2^+$, and C_{ff} has the partially factorized form

$$
C_{ff}(z_1,z_2,q_L^2,q_T^2) = \int \frac{d\Omega_T}{2} \sum_{ijk\ell} K_{(k,\ell)}^{(i,j)} \left(\frac{1}{q_L}\right)^{[i]+[j]}
$$
$$
\times H_{(i,j)}(\mu_S) C_{S,(k,\ell)}(z_1,z_2,\mathbf{q}_T,\mu_S).
$$
$$
(12)
$$

However, in this form the matching coefficients C_S still contain large logarithms of q_T^2/q_L^2 which are not resummed by the usual renormalization group evolution. These rapidity logarithms arise because the graphs renormalizing matrix elements of $T_{(i,j)}$ in SCET are separately divergent

in each sector, even in d dimensions, and the divergences only cancel in the sum. These graphs therefore require the introduction of an additional regulator beyond dimensional regularization, and the rapidity divergences are reflected in logarithms of the (scheme-dependent) rapidity scale. While a number of regulators have been used at leading power [51–53,72,73], the "pure rapidity regulator" introduced in [53] is particularly convenient for studying power corrections, as it properly regulates the power divergences in phase space integrals arising at NLP.

In this paper we use a version of the pure rapidity regulator appropriate for our formalism which introduces separate scheme dependence for the n and $\bar{n}$ sectors, denoted by the parameters ν_n and $\nu_{\bar{n}}$. Rapidity logarithms are summed by running the operators $T_{(i,j)}$ from $\nu_{n,\bar{n}}^H = q_L$ to $\nu_{n,\bar{n}}^S = \mu \sim q_T$. Under rapidity renormalization the operators $T_{(i,j)}$ can mix, leading generically to rapidity renormalization group running of the form

$$
T_{(i,j)}(q^-,q^+,\mathbf{q}_T,\mu_S,\nu_{n,\bar{n}}^H)
$$
$$
= \sum_{k,\ell} \int \frac{d\omega_1 d\omega_2}{\omega_1 \omega_2} d^2\mathbf{p}_T V_{(i,j),(k,\ell)}(\omega_1,\omega_2,\mathbf{p}_T,\mu_S,\nu_{n,\bar{n}}^H,\nu_{n,\bar{n}}^S)
$$
$$
\times T_{(k,\ell)}\left(\frac{q^-}{\omega_1},\frac{q^+}{\omega_2},\mathbf{q}_T - \mathbf{p}_T,\mu_S,\nu_{n,\bar{n}}^S\right), \quad (13)
$$

where by $\nu_{n,\bar{n}}$ we denote depends on both ν_n and $\nu_{\bar{n}}$ separately, and the large logarithms of $\nu_{n,\bar{n}}^H/\nu_{n,\bar{n}}^S$ have been resummed in the rapidity evolution factors $V_{(i,j),(k,\ell)}$. Combining all these steps gives the DY cross section in Eq. (11), where C_{ff} now has the fully factorized form

$$
C_{ff}(z_1,z_2,q_L^2,q_T^2)
$$
$$
= \int \frac{d\Omega_T}{2} \int \frac{d\omega_1 d\omega_2}{\omega_1 \omega_2} \sum_{ijkk'\ell\ell'} \frac{H_{(i,j)}(\mu_S) K_{(k,\ell)}^{(i,j)}}{q_L^{[i]+[j]}}
$$
$$
\times \int d^2\mathbf{p}_T V_{(k,\ell),(k',\ell')}(\omega_1,\omega_2,\mathbf{p}_T,\mu_S,\nu_{n,\bar{n}}^H,\nu_{n,\bar{n}}^S)
$$
$$
\times C_{S,(k',\ell')}\left(\frac{z_1}{\omega_1},\frac{z_2}{\omega_2},\mathbf{q}_T - \mathbf{p}_T,\mu_S,\nu_{n,\bar{n}}^S\right). \quad (14)
$$

In this paper the fixed-order $O(\alpha_s)$ contributions to each of the factors in Eq. (14) which are required to determine the fixed-order cross section at NLP are calculated. The $O(\alpha_s)$ anomalous dimensions of the relevant hard matching coefficients may be found in the literature [56,65], and here we also calculate the $O(\alpha_s)$ off-diagonal entries for the rapidity evolution kernels $\gamma_{(k,\ell),(0,0)}$ which mix the various subleading operators $T_{(k,\ell)}$ into the leading operator $T_{(0,0)}$ with an NLP coefficient. The calculation of the one-loop entries which mix the subleading operators among themselves is left for future work. Additionally, in most phenomenological applications, q_T resummation is

performed for the Fourier conjugate of q_T (b space); here we will work in q_T space, where the SCET operators we are using are defined. Fourier transforming our results to b space may be useful for future applications.

II. NLP OPERATOR PRODUCTS IN SCET

In this section the $O(\alpha_s)$ ingredients that contribute to C_{ff} at next to leading power in SCET are calculated. We begin by summarizing the hard-scale matching of the QCD current onto SCET scattering operators, and then proceed by Fierz rearranging products of these scattering operators into a smaller basis of operators. The matrix elements of these operators are calculated using the pure rapidity regulator, and the final result is compared to the corresponding fixed order result from QCD.

A. Hard-scale matching

The invariant mass of the lepton pair is $q^2 = q^+ q^- - q_T^2 \equiv q_L^2 - q_T^2$, where $q_L^2 \gg q_T^2 \gg \Lambda_{\rm QCD}^2$ and $q^+ \equiv q \cdot n$ and $q^- = q \cdot \bar{n}$ are the large light-cone components of the external current defined in terms of the lightlike vectors $n^\mu = (1, 0, 0, 1)$ and $\bar{n}^\mu = (1, 0, 0, -1)$. This defines the relevant scales for this process. The incoming state consists of two hadrons; the invariant mass of partons in the same hadron is of order $\Lambda_{\rm QCD}$, while the invariant mass of partons in different hadrons is of order q_L. Therefore, at a hard scale $\mu_H \sim q_L$ partons in different hadrons are above one another's cutoff, and QCD is matched onto an EFT in which direct interactions between the sectors have been integrated out. In the SCET formalism used in this paper, SCET consists of decoupled copies of QCD for each sector which only mutually interact via the external electromagnetic current Eq. (6). Only quark and antiquark PDFs are considered in this paper. Gluon PDFs may be included in the same formalism, and the relevant hard-scattering operators are listed in Appendix A, but the calculation for incoming gluons is beyond the scope of this work. We work in a reference frame where the incoming hard quark is in the n sector and the antiquark is in the $\bar{n}$ sector.

The matching of the external vector current from QCD to SCET at subleading power has been considered in a number of papers [37,38,40,55,56,74] and is obtained by expanding QCD amplitudes in powers of $p_i \cdot n / q \cdot n$ for particles in the n sector and $p_i \cdot \bar{n} / q \cdot \bar{n}$ for particles in the $\bar{n}$ sector. In addition to the analogs of operators considered in [56] for two incoming partons, there are also operators suppressed by single powers of the net transverse momentum $p_{n_i,T}$ in either sector (which were eliminated by a choice of reference frame in [56]) as well as corrections to the multipole expansion of the energy-momentum conserving delta functions.

The SCET current has the expansion Eq. (6). The corresponding scattering operators are constructed from the field building blocks [56,75]

$$\bar{\chi}_{\bar{n}}(x) = \bar{\psi}_{\bar{n}}(x) \bar{W}_{\bar{n}}(x) P_n,$$

$$\chi_n(x) = W_n^\dagger(x) P_n \psi_n(x),$$

$$\mathcal{B}_{\bar{n}}^{\mu_1 \cdots \mu_N}(x) = \bar{W}_{\bar{n}}^\dagger(x) i D_{\bar{n}}^{\mu_1}(x) \cdots i D_{\bar{n}}^{\mu_N}(x) \bar{W}_{\bar{n}}(x),$$

$$\mathcal{B}_n^{\dagger \mu_1 \cdots \mu_N}(x) = (-1)^N \bar{W}_n^\dagger(x) i \bar{D}_n^{\mu_1}(x) \cdots i \bar{D}_n^{\mu_N}(x) \bar{W}_n(x), \quad (15)$$

where we note that $(\mathcal{B}^{\mu_1 \cdots \mu_N})^\dagger = \mathcal{B}^{\dagger \mu_N \cdots \mu_1}$. The incoming Wilson lines $\bar{W}$ are defined as

$$\bar{W}_n^\dagger(x) = \bar{\mathcal{P}} \exp\left(-ig \int_{-\infty}^0 ds \, \bar{n} \cdot A_n(x + \bar{n}s) e^{s0^+} \right),$$

$$\bar{W}_{\bar{n}}(x) = \mathcal{P} \exp\left(ig \int_{-\infty}^0 ds \, n \cdot A_{\bar{n}}(x + ns) e^{s0^+} \right). \quad (16)$$

We use the conventions

$$i D_{\bar{n}}^\mu(x) = i \partial^\mu + g A_{\bar{n}}^\mu(x), \quad i \bar{D}_n^\mu(x) = i \bar{\partial}^\mu - g A_n^\mu(x), \quad (17)$$

and it is convenient to define the four-vectors introduced in [55]

$$\eta^\mu = \sqrt{\frac{\bar{n} \cdot q}{n \cdot q}} n^\mu, \qquad \bar{\eta}^\mu = \sqrt{\frac{n \cdot q}{\bar{n} \cdot q}} \bar{n}^\mu, \quad (18)$$

which are invariant under the boost reparametrization $n^\mu \to e^y n^\mu$, $\bar{n}^\mu \to e^{-y} \bar{n}^\mu$.

At leading power there is a single scattering operator,

$$O_2^{(0)\mu}(x) = [\bar{\chi}_{\bar{n}}(x_{\bar{n}})] \gamma^\mu [\chi_n(x_n)], \quad (19)$$

where

$$x_n^\mu \equiv x^+ \frac{\bar{n}^\mu}{2} + x_\perp^\mu, \quad \text{and} \quad x_{\bar{n}}^\mu \equiv x^- \frac{n^\mu}{2} + x_\perp^\mu. \quad (20)$$

Note that the fields in the operator are multipole expanded; this is necessary for the energy-momentum conserving delta functions to preserve the correct power counting. For example, if p_n^μ and $p_{\bar{n}}^\mu$ are momenta in the n and $\bar{n}$ sectors, respectively, we have the expansion

$$\delta(p_n^- + p_{\bar{n}}^- - q^-) = \delta(p_{\bar{n}}^- - q^-) + p_n^- \delta'(p_{\bar{n}}^- - q^-) + \cdots \quad (21)$$

and similarly for the n components. Performing this expansion up to $O(1/q_L^2)$ gives

$$[\bar{\chi}_{\bar{n}}(x)] \gamma^\mu [\chi_n(x)] = O_2^{(0)\mu}(x)$$
$$+ \frac{1}{q_L^2} (O_2^{(2\delta^+)\mu}(x) + O_2^{(2\delta^-)\mu}(x)), \quad (22)$$

where

$$O_2^{(2\delta^+)\mu}(x) = \frac{1}{2} q^- q^+ x^- [\bar{\chi}_{\bar{n}}(x_{\bar{n}})] \gamma^\mu [n \cdot \partial \chi_n(x_n)],$$

$$O_2^{(2\delta^-)\mu}(x) = \frac{1}{2} q^- q^+ x^+ [\bar{n} \cdot \partial \bar{\chi}_{\bar{n}}(x_{\bar{n}})] \gamma^\mu [\chi_n(x_n)]. \quad (23)$$

Power counting the multipole-expanded operators is not immediately obvious. In $O_2^{(2\delta^+)}$, for example, $q^+ x^-$ is of order 1 since $x^- \sim \partial/\partial q^+$, whereas $q^- p_n^+ \sim O(p_n^- p_n^+) \sim O(p_{n\perp}^2)$; thus, matrix elements of the operators in Eq. (23) are $O(1/q_L^2)$ relative to leading power. Since we are working up to $1/q_L^2$ suppression, the contributions from higher multipole expansions in the fields are only included for the leading power operator $O_2^{(0)}$.

At $O(1/q_L)$, there are two operators suppressed by a single perpendicular derivative,

$$O_2^{(1_{\perp n})\mu}(x) = [\bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma^\mu \frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp[-i\partial^\alpha \chi_n(x_n)],$$

$$O_2^{(1_{\perp \bar{n}})\mu}(x) = [-i\partial^\alpha \bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma_\alpha^\perp \frac{\slashed{n}}{2}\gamma^\mu[\chi_n(x_n)]. \quad (24)$$

These were not required in [55,56] since they could be removed by a suitable choice of reference frame, while here the presence of initial-state radiation prevents such a choice.

Finally, there are several operators containing factors of $\mathcal{B}_{n,\bar{n}}$ whose matrix elements begin at $O(g_s)$. These operators are labeled by a continuous parameter t which parametrizes the separation of fields along the light cone [67]. We define the dimensionless parameter $\hat{t} \equiv q^- t$ if the shift occurs in the n sector, and by $\hat{t} \equiv q^+ t$ if the shift occurs in the $\bar{n}$ sector. We define the A-type operators in which a gluon is emitted at leading order in the n sector,

$$O_2^{(1A_1)\mu}(x,\hat{t}) = [\bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma_\alpha^\perp \frac{\slashed{\bar{n}}}{2}\gamma^\mu[\mathcal{B}_n^{\dagger\alpha}(x_n - \bar{n}t)\chi_n(x_n)],$$

$$O_2^{(1A_2)\mu}(x,\hat{t}) = -[\bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma^\mu \frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp[\mathcal{B}_n^{\dagger\alpha}(x_n - \bar{n}t)\chi_n(x_n)],$$

$$O_2^{(2A_1)\mu}(x,\hat{t}) = -2\pi i\theta(\hat{t}) \otimes [\bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma_\alpha^\perp\gamma_\beta^\perp\gamma^\mu$$
$$\times [\mathcal{B}_n^{\dagger\alpha\beta}(x_n - \bar{n}t)\chi_n(x_n)], \quad (25)$$

and the corresponding B-type operators where the gluon is emitted in the $\bar{n}$ sector,

$$O_2^{(1B_1)\mu}(x,\hat{t}) = -[\bar{\chi}_{\bar{n}}(x_{\bar{n}})\mathcal{B}_{\bar{n}}^\alpha(x_{\bar{n}} - nt)]\gamma^\mu \frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp[\chi_n(x_n)],$$

$$O_2^{(1B_2)\mu}(x,\hat{t}) = [\bar{\chi}_{\bar{n}}(x_{\bar{n}})\mathcal{B}_{\bar{n}}^\alpha(x_{\bar{n}} - nt)]\gamma_\alpha^\perp \frac{\slashed{\bar{n}}}{2}\gamma^\mu[\chi_n(x_n)],$$

$$O_2^{(2B_1)\mu}(x,\hat{t}) = -2\pi i\theta(\hat{t}) \otimes [\bar{\chi}_{\bar{n}}(x_{\bar{n}})\mathcal{B}_{\bar{n}}^{\alpha\beta}(x_{\bar{n}} - nt)]$$
$$\times \gamma^\mu\gamma_\alpha^\perp\gamma_\beta^\perp[\chi_n(x_n)]. \quad (26)$$

Following [67], it is convenient to work with the Fourier-transformed operators

$$O_2^{(i)}(x,u) = \int \frac{d\hat{t}}{2\pi} e^{-iu\hat{t}} O_2^{(i)}(x,\hat{t}),$$

$$C_2^{(i)}(x,u) = \int d\hat{t} e^{iu\hat{t}} C_2^{(i)}(x,\hat{t}). \quad (27)$$

We have also defined the convolutions in $\hat{t}$ space in these definitions as

$$f(\hat{t}) \otimes g(\hat{t}) = \int \frac{dxdy}{2\pi} f(x)g(y)\delta(\hat{t} - x - y). \quad (28)$$

Note that the one-gluon matrix element of $O_2^{(2A_1)}(x,u)$ is proportional to $\delta(u + k^-)/u$, where k^μ is the gluon momentum. If the convolution with $\theta(\hat{t})$ had not been included in its definition (as was the case in [56]), the matrix element of the operator would instead be proportional to $\delta(u + k^-)$, and the operator would have a factor of $1/u$ in its Wilson coefficient. This is inconvenient because in the DY process studied here, this factor of $1/u \sim 1/k^-$ corresponds to a rapidity divergence, and rapidity renormalizing operator products such as $O_2^{(2A1)\dagger}O_2^{(0)}$ without the factor of $1/u$ in its matrix element would then give rise to an unregulated rapidity divergence in the final integral over u.[2] These are similar to the endpoint divergences which have been previously noted at NLP in SCET, in particular in b-mediated $h \to \gamma\gamma$ decay [41–43]. With the definition given here—which is similar to the modification of SCET operators proposed in [78]—the u integral does not introduce any additional singularities, and thus all rapidity divergences are correctly regulated by the pure rapidity regulator. We illustrate this with an example in Sec. II D 3.

Since SCET currents and their products contain operators with zero, one, or two factors of u at this order, we use the notation $\{u\}$ to denote the dependence of a quantity on any number of u's, as well as $\int d\{u\}$ to indicate integration over any number of u's (including zero). The expansion of the SCET current may therefore be written

$$J_{\text{SCET}}^\mu(x) = \sum_i \int d\{u\} \frac{1}{q_L^{[i]}} C_2^{(i)}(\{u\}) O_2^{(i)\mu}(x,\{u\})$$

$$= C_2^{(0)}\left[O_2^{(0)\mu}(x) + \frac{1}{q_L}(O_2^{(1_{\perp n})\mu}(x) + O_2^{(1_{\perp\bar{n}})\mu}(x)) + \frac{1}{q_L^2}(O_2^{(2\delta^+)\mu}(x) + O_2^{(2\delta^-)\mu}(x))\right]$$

$$+ \frac{1}{q_L}\sum_i \int du C_2^{(1i)}(u) O_2^{(1i)\mu}(x,u) + \frac{1}{q_L^2}\sum_i \int du C_2^{(2i)}(u) O_2^{(2i)\mu}(x,u) + O\left(\frac{1}{q_L^3}\right), \quad (29)$$

---

[2]The importance of having a finite integral over convolution variables was stressed in [74,76,77].

where on the first line the sum is over all operators $i = 0, 1_{\perp n}, \ldots, 2B_1$, while in the last two lines the sums are over the operators of the appropriate dimension whose coefficients are not fixed by reparametrization or translation invariance. The operators $O_2^{(1_{\perp n, \bar{n}})}$ are related to $O_2^{(0)}$ through reparametrization invariance (RPI) [65,70], and so to all orders in α_s we have the equalities $C_2^{(0)} = C_2^{(1_{\perp n})} = C_2^{(1_{\perp \bar{n}})}$, while the translation invariance of QCD ensures that $C_2^{(0)} = C_2^{(2\delta^+)} = C_2^{(2\delta^-)}$. The normalizations of all operators have been chosen so that their tree-level matching coefficient is unity in u space,

$$C_2^{(i)}(\mu, \{u\}) = 1 + O(\alpha_s). \tag{30}$$

The one-gluon matrix elements of the operators $O_2^{(i)}(\mu, \{u\})$ are given in Appendix A.

There are additional operators not included in Eqs. (25) and (26) that are part of the general SCET current expansion [56], but which do not contribute to C_{ff} at the order (in α_s, q_T^2/q^2, or $\Lambda_{\mathrm{QCD}}^2/q^2$) to which we are working, or which contribute only to the gluon-initiated Drell-Yan subprocess. These operators do not mix under renormalization at one loop with the operators considered here, and so are not included in this analysis, though we list them in Appendix A for completeness.

B. Renormalization group running

The anomalous dimensions of all the required matching coefficients $C_2^{(i)}(\{u\})$ have been calculated previously in [56,64,65]. They obey the integro-differential equation

$$\frac{d}{d \log \mu} C_2^{(i)}(\mu, \{u\}) = \int d\{v\} \gamma_2^{(i)}(\{u\}, \{v\}) C_2^{(i)}(\mu, \{v\}), \tag{31}$$

where the kernels $\gamma_2^{(i)}$ have the form

$$\gamma_2^{(i)}(\{u\}, \{v\}) = \Gamma_{\mathrm{cusp}}^{(i)}[\alpha_s] \log \frac{-q_L^2 - i0^+}{\mu^2} \delta(\{u\} - \{v\})$$
$$+ \gamma_{\mathrm{non-cusp}}^{(i)}(\{u\}, \{v\}). \tag{32}$$

Working in the leading-log approximation only the cusp anomalous dimension is required. The one-loop cusp anomalous dimension is universal,

$$\Gamma_{\mathrm{cusp}}^{(0)} = \Gamma_{\mathrm{cusp}}^{(1A_i)} = \Gamma_{\mathrm{cusp}}^{(2A_i)} = \Gamma_{\mathrm{cusp}}^{(1B_i)} = \Gamma_{\mathrm{cusp}}^{(2B_i)} = \frac{\alpha_s C_F}{\pi}. \tag{33}$$

With the definition $H_{(i,j)}(\{u\}) = C_2^{(i)\dagger}(\{u\}) C_2^{(j)}(\{u\})$, the leading-log running of the hard functions is determined by the RGE,

$$\frac{d}{d \log \mu} H_{(i,j)}(\mu, \{u\}) = \left(2\Gamma_{\mathrm{cusp}} \log \frac{q_L^2}{\mu^2} \right) H_{(i,j)}(\mu, \{u\}). \tag{34}$$

This gives the leading log (LL) unitary evolution for all $H_{(i,j)}$,

$$H_{(i,j)}(\mu, \{u\}) = U_H^{\mathrm{LL}}(\mu, \mu_H) H_{(i,j)}(\mu_H, \{u\}), \tag{35}$$

where, with $\beta[\alpha_s] \equiv d\alpha_s/d \log \mu = -\beta_0 \alpha_s^2/2\pi + \cdots$,

$$\log U_H^{\mathrm{LL}}(\mu, \mu_H)$$
$$= -4 \int_{\alpha_s(\mu_H)}^{\alpha_s(\mu)} \frac{d\alpha}{\beta[\alpha]} \Gamma_{\mathrm{cusp}}[\alpha] \int_{\alpha(q_L)}^{\alpha} \frac{d\alpha'}{\beta[\alpha']}$$
$$= \frac{16\pi C_F}{\beta_0^2} \left(\frac{1}{\alpha_s(\mu_H)} - \frac{1}{\alpha(\mu)} - \frac{1}{\alpha(q_L)} \log \frac{\alpha(\mu)}{\alpha(\mu_H)} \right). \tag{36}$$

Beyond LL there will be operator mixing, and the solution to the RGE will be more involved. This sums the RG logarithms of μ_H/μ in the hard functions.

C. $T_{(i,j)}$ definitions

The differential cross section for DY production is given in terms of hadronic matrix elements of products of operators $O_2^{(i)\dagger}(x) O_2^{(j)}(0)$ in Eq. (7). Matrix elements of these operator products may be evaluated between partons in perturbation theory to calculate the matching conditions onto light-cone distribution operators (whose matrix elements are the usual PDFs); however, it is convenient to perform a Fierz rearrangement for each operator product to write it as the product of factors in the n and $\bar{n}$ sectors, corresponding to the convolution of generalized transverse momentum dependent distribution operators. At leading power, this gives

$$\int \frac{d^d x}{2(2\pi)^d} (-g_{\mu\nu}) e^{-iq \cdot x} O_2^{(0)\mu\dagger}(x) O_2^{(0)\nu}(0)$$
$$= \frac{1}{N_c} \int \frac{d^d x}{2(2\pi)^d} e^{-iq \cdot x} \Phi_n^{(0)}(x_n) \Phi_{\bar{n}}^{(0)}(x_{\bar{n}})$$
$$\equiv \frac{1}{N_c} T_{(0,0)}(q^-, q^+, \mathbf{q}_T). \tag{37}$$

The leading power position space distribution operators are defined as

$$\Phi_n(x_n) \equiv \bar{\chi}_n(x_n) \frac{\bar{\slashed{n}}}{2} \chi_n(0),$$
$$\Phi_{\bar{n}}(x_{\bar{n}}) \equiv \bar{\chi}_{\bar{n}}(0) \frac{\slashed{n}}{2} \chi_{\bar{n}}(x_{\bar{n}}), \tag{38}$$

where x_n and $x_{\bar{n}}$ are defined in (20), and thus consist of quark fields separated in the transverse direction by $x_\perp^\mu$ as well as along the light cone.

Products of power-suppressed operators may similarly be written as convolutions of higher dimension operators,

$$T_{(i,j)}(q,\{u\}) = \int \frac{d^d x}{2(2\pi)^d} e^{-iq\cdot x} \Phi_n^{(i)}(x_n,\{u\}) \Phi_{\bar{n}}^{(j)}(x_{\bar{n}},\{u\}),$$

(39)

where we define the relevant subleading transverse momentum dependent light-cone distribution operators $\Phi_n^{(i)}(x,\{t\})$ as

$$\Phi_n^{(2_1)}(x_n,\hat{t}) \equiv (i\partial^\mu \bar{\chi}_n(x_n)) \frac{\slashed{\bar{n}}}{2} \gamma_\mu^\perp \gamma_\nu^\perp \mathcal{B}_n^{\dagger\nu}(-\bar{n}t)\chi_n(0),$$

$$\Phi_n^{(2_2)}(x_n,\hat{t}_1,\hat{t}_2) \equiv -\bar{\chi}_n(x_n)\mathcal{B}_n^\mu(x_n - \bar{n}t_1) \frac{\slashed{\bar{n}}}{2} \gamma_\mu^\perp \gamma_\nu^\perp$$
$$\times \mathcal{B}_n^{\dagger\nu}(-\bar{n}t_2)\chi_n(0),$$

$$\Phi_n^{(2_3)}(x_n,\hat{t}) \equiv 2\pi i\theta(\hat{t}) \otimes \bar{\chi}_n(x_n)\mathcal{B}_n^{\mu\nu}(x_n - \bar{n}t)$$
$$\times \frac{\slashed{\bar{n}}}{2} \gamma_\nu^\perp \gamma_\mu^\perp \chi_n(0),$$

$$\Phi_n^{(2_4)}(x_n) \equiv q^+ q^- \frac{x^-}{2} (n\cdot\partial\bar{\chi}_n(x_n)) \frac{\slashed{\bar{n}}}{2} \chi_n(0). \quad (40)$$

The corresponding $\bar{n}$-sector operators $\Phi_{\bar{n}}^{(i)}$ are found by taking the Hermitian conjugate and changing $n \leftrightarrow \bar{n}$. The u-space Fourier conjugates of these building blocks are defined by the transformation in Eq. (27) for shifts relative to the origin (since these shifts come from an operator), and by the conjugate transformation for shifts relative to x_n (since these shifts come from the conjugated operator).

Thus, for example,

$$T_{(0,2_1)}(q,u) = \int \frac{d^d x}{2(2\pi)^d} e^{-iq\cdot x} \Phi_n^{(0)}(x_n) \Phi_{\bar{n}}^{(2_1)}(x_{\bar{n}},u)$$

$$= \int \frac{d^d x}{2(2\pi)^d} \frac{d\hat{t}}{2\pi} e^{-iu\hat{t}} e^{-iq\cdot x} \left[\bar{\chi}_n(x_n) \frac{\slashed{\bar{n}}}{2} \chi_n(0)\right]$$

$$\times \left[\bar{\chi}_{\bar{n}}(0)\mathcal{B}_{\bar{n}}^\nu(-nt)\gamma_\nu^\perp \gamma_\mu^\perp \frac{\slashed{n}}{2} (-i\partial^\mu \chi_{\bar{n}}(x_{\bar{n}}))\right].$$

(41)

In general, we can write

$$\int \frac{d^d x}{2(2\pi)^d} (-g_{\mu\nu}) e^{-iq\cdot x} O_2^{(i)\mu\dagger}(x) O_2^{(j)\nu}(0)$$

$$= \sum_{k,\ell} \frac{1}{N_c} K_{(k,\ell)}^{(i,j)} T_{(k,\ell)} + \text{spin dependent}, \quad (42)$$

where the only nonzero elements of K which are relevant at this order are

$$K_{(2_1,0)}^{(1\perp n,1A_1)} = K_{(2_1,0)}^{(1A_1,1\perp n)} = K_{(2_2,0)}^{(1A_2,1A_1)} = K_{(2_2,0)}^{(1A_1,1A_2)}$$

$$= K_{(0,2_1)}^{(1\perp\bar{n},1B_1)} = K_{(0,2_1)}^{(1B_1,1\perp\bar{n})} = K_{(0,2_2)}^{(1B_2,1B_1)}$$

$$= K_{(0,2_2)}^{(1B_1,1B_2)} = K_{(2_3,0)}^{(2A_1,0)} = K_{(2_3,0)}^{(0,2A_1)}$$

$$= K_{(2_4,0)}^{(2\delta^+,0)} = K_{(0,2_3)}^{(0,2B_1)} = K_{(0,2_3)}^{(2B_1,0)}$$

$$= K_{(0,2_4)}^{(2\delta^-,0)} = 1. \quad (43)$$

D. Matrix elements of operator products

Individual n- and $\bar{n}$-sector graphs contributing to the matrix elements of each $T_{(i,j)}$ are rapidity divergent and require a regulator to give finite results. We use a version of the pure rapidity regulator introduced in [53]. As discussed in that reference, other commonly used rapidity regulators such as the δ regulator [72] or the η regulator [79] are not suitable for handling the power-law rapidity divergences that arise at NLP. An explicit example of the δ regulator failing to regulate rapidity divergences at NLP is given in Sec. II D 3.

In what follows we define the pure rapidity regulator by modifying the integration measure of n-sector and $\bar{n}$-sector particles as

$$d^d k_n \to w_n^2 \left(\frac{q_L^2}{\nu_n^2}\right)^{\eta_n/2} \left(\frac{q^-}{q^+} \frac{k_n^+}{k_n^-}\right)^{\eta_n/2} d^d k_n,$$

$$d^d k_{\bar{n}} \to w_{\bar{n}}^2 \left(\frac{q_L^2}{\nu_{\bar{n}}^2}\right)^{\eta_{\bar{n}}/2} \left(\frac{q^+}{q^-} \frac{k_{\bar{n}}^-}{k_{\bar{n}}^+}\right)^{\eta_{\bar{n}}/2} d^d k_{\bar{n}}. \quad (44)$$

This regulator has the distinct advantage that—as in dimensional regularization—scaleless integrals vanish, and as a result all overlap integrals evaluate to zero. This greatly simplifies the calculation since, as is discussed in detail in Sec. III, in a scheme where overlap integrals do not vanish, the overlap subtraction procedure must be carried out to subleading powers.

The regulator in Eq. (44) is slightly modified from the form presented in [53]: the factors of $q^\pm$ ensure boost invariance, as in [40], the dimensionless parameter υ has been replaced by the equivalent dimensionful parameters ν_i, and distinct parameters η_i, ν_i, and w_i have been introduced for each sector, since the fields in the n and $\bar{n}$ sectors are independent. As discussed in [57], rapidity logarithms in SCET correspond to a scheme dependence in defining the sum of individually rapidity divergent graphs in the n and $\bar{n}$ sectors. Regulating both sectors (and their corresponding overlap graphs) in the same way and then removing the regulator is equivalent to naïvely adding the individual graphs together before performing the loop integrals, and reproduces the rapidity logarithms of QCD. Since QCD has no rapidity divergences, rapidity divergences cancel for this choice, which corresponds to choosing the parameters $\eta_n = -\eta_{\bar{n}}$, $\nu_n \nu_{\bar{n}} = q_L^2$, and

$w_n = w_{\bar{n}}$. This was explicitly demonstrated up to NLP in [53]: these authors showed that if QCD diagrams are first rapidity regulated and then expanded in the n-collinear, $\bar{n}$-collinear, and soft limits, the leading and subleading power matrix elements reproduce the rapidity-finite QCD results expanded to the same order. A similar cancellation of rapidity divergences will be shown here.

Using different rapidity regulator parameters in the two sectors moves the rapidity logarithm of q_L^2 into the Wilson coefficients of the EFT and allows the scheme dependence of the resulting graphs to be exploited to sum the corresponding rapidity logs. The corresponding rapidity divergences correspond to $1/\eta_i$ singularities which are canceled by introducing the appropriate counterterms into the EFT, and rapidity logarithms are then summed using rapidity renormalization group (RRG) techniques similar to [79]. The bookkeeping constants w_i are taken to formally obey the RRG equation

$$\frac{dw_i}{d\log\nu_i} = \frac{\eta_i}{2}w_i, \qquad (45)$$

which cancels the scheme dependence in the measure, keeping the bare theory ν_i independent and allowing techniques analogous to those in dimensional regularization to be used to extract the rapidity anomalous dimensions. As in [53,79], these bookkeeping constants are set to unity at the end of calculation. Rapidity logarithms are minimized by the appropriate choice of the dimensionless parameters $\nu_{n,\bar{n}}$.

As noted in [57], choosing $\nu_n\nu_{\bar{n}} \neq q_L^2$ requires rapidity counterterms for each $T_{(i,j)}$ which are sensitive to the scale q_T. Scale sensitivities in the counterterm generate the same scale dependence in the Wilson coefficient through the RGE, and since Wilson coefficients in an EFT must be independent of infrared physics, this adds the constraint that the theory must first be evolved to $\mu \sim q_T$ before running in rapidity. This will be discussed in more detail in Sec. II F.

At LP the only operator is $T_{(0,0)}$, so its divergences are absorbed by the renormalization constant $Z_{(0,0),(0,0)}$,

$$T_{(0,0)}^B(q^-, q^+, \mathbf{q}_T) = \int \frac{d\omega_1}{\omega_1}\frac{d\omega_2}{\omega_2}d^{d-2}\mathbf{p}_T$$
$$\times Z_{(0,0),(0,0)}(\omega_1, \omega_2, \mathbf{p}_T)$$
$$\times T_{(0,0)}\left(\frac{q^-}{\omega_1}, \frac{q^+}{\omega_2}, \mathbf{q}_T - \mathbf{p}_T\right), \qquad (46)$$

where $T_{(i,j)}^B$ and $T_{(i,j)}$ denote the bare and renormalized operators, respectively, and the integral corresponds to summing over the infinite set of operators $T_{(0,0)}(k^-, k^+, \mathbf{k_T})$. At subleading powers the various operators may mix with one another, so we have the general relation

$$T_{(i,j)}^B(q^-, q^+, \mathbf{q}_T, \{u\})$$
$$= \sum_{(k,l)}\int \frac{d\omega_1}{\omega_1}\frac{d\omega_2}{\omega_2}d^{d-2}\mathbf{p}_T d\{v\}$$
$$\times Z_{(i,j),(k,l)}(\omega_1, \omega_2, \mathbf{p}_T, \{v\})$$
$$\times T_{(k,l)}\left(\frac{q^+}{\omega_2}, \frac{q^-}{\omega_1}, \mathbf{q}_T - \mathbf{p}_T, \{u - v\}\right), \qquad (47)$$

where the sum over operators includes each subleading $T_{(i,j)}$ as well as the leading operator $T_{(0,0)}$ with a power-suppressed coefficient, as will be discussed in the following sections.

1. Leading power example

The leading power calculation of DY production in this formalism was presented in [57] using the δ regulator; we repeat the calculation here with the pure rapidity regulator. At leading power, there is a single bilocal operator contributing to the rate,

$$T_{(0,0)}(q^-, q^+, \mathbf{q}_T)$$
$$= \int \frac{d^d x}{2(2\pi)^d}e^{-iq\cdot x}\left[\bar{\chi}_n(x_n)\frac{\not{\bar{n}}}{2}\chi_n(0)\right]\left[\bar{\chi}_{\bar{n}}(0)\frac{\not{n}}{2}\chi_{\bar{n}}(x_{\bar{n}})\right]. \qquad (48)$$

With incoming quark and antiquark states $q(p_1)$ and $\bar{q}(p_2)$ the tree-level matrix element of this operator is

$$\frac{1}{4}\sum_{\text{spins}}\langle p_1^n p_2^{\bar{n}}|T_{(0,0)}|p_1^n p_2^{\bar{n}}\rangle$$
$$= \frac{1}{4}\sum_{\text{spins}}\int \frac{d^d x}{2(2\pi)^d}e^{-iq\cdot x}\langle p_1^n|\left[\bar{\chi}_n(x_n)|0\rangle\frac{\not{\bar{n}}}{2}\langle 0|\chi_n(0)\right]|p_1^n\rangle$$
$$\times \langle p_2^{\bar{n}}|\left[\bar{\chi}_{\bar{n}}(0)|0\rangle\frac{\not{n}}{2}\langle 0|\chi_{\bar{n}}(x_{\bar{n}})\right]|p_2^{\bar{n}}\rangle$$
$$= \delta(\bar{z}_1)\delta(\bar{z}_2)\delta^{d-2}(\mathbf{q}_T) \equiv \delta_1\delta_2\boldsymbol{\delta}_T, \qquad (49)$$

where the superscripts n and $\bar{n}$ in Eq. (49) denote the sector of the corresponding parton. We also use the notation

$$z_1 \equiv \frac{q^-}{p_1^-}, \qquad z_2 \equiv \frac{q^+}{p_2^+}, \qquad \bar{z}_i \equiv 1 - z_i, \qquad (50)$$

and

$$\delta_i \equiv \delta(\bar{z}_i), \qquad \delta_i' \equiv \delta'(\bar{z}_i), \qquad \boldsymbol{\delta}_T \equiv \delta^{d-2}(\mathbf{q}_T). \qquad (51)$$

At $O(\alpha_s)$, the matrix element corresponding to the emission of a real n-sector gluon is given by the three n-sector graphs shown of Fig. 1. Denoting the spin-averaged one-loop matrix element of $T_{(0,0)}$ by $\mathcal{M}_{(0,0)}$, we write
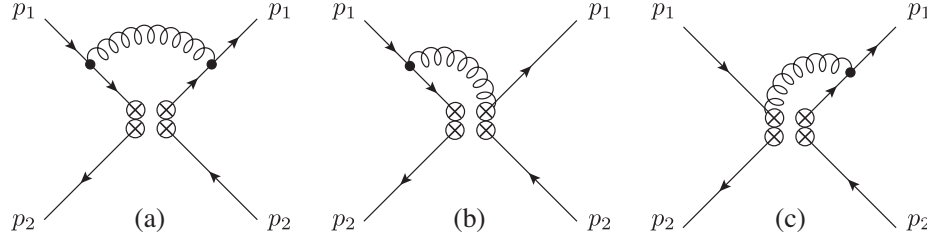
FIG. 1.   Nonvanishing graphs in the n sector contributing to the matrix element $\langle p_1^n p_2^{\bar{n}} | T_{(0,0)} | p_1^n p_2^{\bar{n}} \rangle$ at $O(\alpha_s)$.

$$\mathcal{M}_{(0,0)} = \mathcal{M}_{(0,0)}^n + \mathcal{M}_{(0,0)}^{\bar{n}} - \mathcal{M}_{(0,0)}^O, \qquad (52)$$

where the superscripts n and $\bar{n}$ denote the $O(\alpha_s)$ contribution from a gluon in the corresponding sector and the O superscript denotes the overlap subtraction. Since these matrix elements correspond to a matching calculation at the scale $\mu_S \sim q_T \gg \Lambda_{\mathrm{QCD}}$ we use the initial-state kinematics $p_1^+ = p_{1\perp} = 0 = p_{2\perp} = p_2^-$, and we obtain

$$\mathcal{M}_{(0,0)}^n = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} w_n^2 \left(\frac{q_L^2}{\nu_n^2} \frac{q^-}{q^+} \frac{k^+}{k^-} \right)^{\eta_n/2}$$
$$\times \delta(k^2) \delta(p_1^- - q^- - k^-) \delta(p_2^+ - q^+) \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)$$
$$\times \mathrm{Tr} \left[\frac{\not{p}_1}{2} \left(\frac{2p_1^\alpha - \gamma^\alpha \not{k}}{-2p_1 \cdot k} + \frac{\bar{n}^\alpha}{k^-} \right) \frac{\not{\bar{n}}}{2} \right.$$
$$\times \left. \left(\frac{2p_{1\alpha} - \not{k}\gamma_\alpha}{-2p_1 \cdot k} - \frac{\bar{n}_\alpha}{-k^-} \right) \right] \mathrm{Tr} \left[\frac{\not{p}_2}{2} \frac{\not{n}}{2} \right], \qquad (53)$$

which evaluates to

$$\mathcal{M}_{(0,0)}^n = \frac{\bar{\alpha}}{\pi} f_\epsilon w_n^2 \delta_2 \frac{(\mu^2)^{-\eta_n/2}}{(\mathbf{q}_T^2)^{1-\eta_n/2}} \left(\frac{z_1 \mu}{\nu_n} \right)^{\eta_n}$$
$$\times \frac{(2 - 2\bar{z}_1 + (1-\epsilon)\bar{z}_1^2)}{\bar{z}_1^{1+\eta_n}}, \qquad (54)$$

where

$$\bar{\alpha} \equiv \frac{\alpha_s C_F}{2\pi}, \qquad f_\epsilon \equiv (\pi \mu^2 e^\gamma)^\epsilon, \qquad (55)$$

and we work in $d = 4 - 2\epsilon$ dimensions.

To extract the singularity structure of this matrix element at $\bar{z}_1 = 0$ we use the distributional identity

$$\frac{\theta(\bar{z}_1)}{\bar{z}_1^{1+\eta}} = -\frac{\delta(\bar{z}_1)}{\eta} + \left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_+ + \cdots \qquad (56)$$

for scalars (see Appendix B for definitions) as well as the identity [19]

$$\frac{(\mu^2)^{-\eta/2}}{(\mathbf{q}_T^2)^{1-\eta/2}} = \mu^{-2\epsilon} \frac{S_{d-2}}{2} \left(\frac{\delta_T}{\frac{\eta}{2} - \epsilon} + \mathcal{L}_{0T} + \frac{\eta}{2} \mathcal{L}_{1T} + \cdots \right) \qquad (57)$$

for vectors in $(d-2)$ dimensions, where the $\mathcal{L}_{nT}$ are vector plus distributions [19,51] defined in Appendix B 2,

$$\mathcal{L}_{nT} \equiv \mathcal{L}_n(\mathbf{q}_T, \mu) = \frac{2\mu^{2\epsilon}}{S_{d-2}} \left[\frac{\log^n \frac{\mathbf{q}_T^2}{\mu^2} \theta(\mathbf{q}_T^2)}{\mathbf{q}_T^2} \right]_+^{\mu^2}, \qquad (58)$$

and $S_{d-2} = 2\pi^{\frac{d-2}{2}} / \Gamma(\frac{d-2}{2})$. Upon expanding first in η_n then in ϵ, the n-sector contribution to the matrix element from a single real emission is

$$\mathcal{M}_{(0,0)}^n = \frac{\bar{\alpha}}{2} w_n^2 \delta_2 \left[\frac{4}{\eta_n} \delta_1 \left(\frac{\delta_T}{\epsilon} - \mathcal{L}_{0T} \right) \right.$$
$$+ \delta_1 \delta_T \left(\frac{2}{\epsilon^2} + \frac{3 - 2\log \frac{\nu_n^2}{\mu^2}}{\epsilon} \right)$$
$$+ 2 \left(\mathcal{L}_{0T} - \frac{\delta_T}{\epsilon} \right) \left[\frac{1 + z_1^2}{\bar{z}_1} \right]_+ \delta_T$$
$$+ 2\bar{z}_1 \delta_T - \zeta_2 \delta_1 \delta_T - 2\delta_1 \mathcal{L}_{1T}$$
$$\left. - \delta_1 \mathcal{L}_{0T} \left(3 - 2\log \frac{\nu_n^2}{\mu^2} \right) \right]. \qquad (59)$$

The $\bar{n}$-sector contribution is obtained under the replacements $z_1 \to z_2$, $w_n \to w_{\bar{n}}$, $\nu_n \to \nu_{\bar{n}}$, and $\eta_n \to \eta_{\bar{n}}$.

As in [55], the overlap between the two sectors is obtained by taking the opposite-sector gluon limit of the n- and $\bar{n}$-sector graphs. As detailed in Sec. III, the subtraction prescription corresponds to subtracting half the wrong sector limit for the gluon of each sector,[3] which we denote

$$\mathcal{M}_{(0,0)}^O = \frac{1}{2} (\mathcal{M}_{(0,0)}^{n \to \bar{n}} + \mathcal{M}_{(0,0)}^{\bar{n} \to n}). \qquad (60)$$

For example, the wrong sector limit of Eq. (53) is

---

[3]In [55], the limits from either sector were equal.

$$
\mathcal{M}_{(0,0)}^{n\to\bar{n}} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} w_n^2 \left(\frac{q_L^2}{\nu_n^2}\frac{q^-}{q^+}\frac{k^+}{k^-}\right)^{\eta_n/2}
$$
$$
\times \delta(k^2)\delta(p_1^- - q^-)\delta(p_2^+ - q^+)\delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)
$$
$$
\times \mathrm{Tr}\left[\frac{\slashed{p}_1}{2}\left(\frac{n^\alpha}{-k^+} + \frac{\bar{n}^\alpha}{k^-}\right)\frac{\slashed{\bar{n}}}{2}\left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-}\right)\right]
$$
$$
\times \mathrm{Tr}\left[\frac{\slashed{p}_2}{2}\frac{\slashed{n}}{2}\right] + \cdots, \tag{61}
$$

where the dots indicate terms suppressed by powers of $1/q_L$ relative to leading power. Integrating with respect to $\mathbf{k}_T$ and then k^+, we find the $\bar{n}$ limit of the n-sector graphs is

$$
\mathcal{M}_{(0,0)}^{n\to\bar{n}} = \frac{2\bar{\alpha}}{\pi} f_\epsilon w_n^2 \delta_1 \delta_2 \left(\frac{q_L^2}{\nu_n^2}\frac{q^-}{q^+}\right)^{\eta_n/2} \frac{1}{(q_T^2)^{1-\eta_n/2}}
$$
$$
\times \int_0^\infty \frac{dk^-}{(k^-)^{1+\eta_n}}, \tag{62}
$$

which is a scaleless divergence and vanishes in this regularization scheme. The overlap subtraction between the two sectors is therefore zero when using the pure rapidity regulator at $O(\alpha_s)$, and this remains true beyond leading power.

Summing the contributions from each sector and subtracting off the (vanishing) overlap, we find the $O(\alpha_s)$ contribution to the matrix element of $T_{(0,0)}$,

$$
\mathcal{M}_{(0,0)} = \bar{\alpha}\left[2\left(\frac{w_n^2}{\eta_n} + \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}}\right)\delta_1\delta_2\left(\frac{\boldsymbol{\delta}_T}{\epsilon} - \mathcal{L}_{0T}\right)\right.
$$
$$
+ \delta_1\delta_2\boldsymbol{\delta}_T\left(\frac{2}{\epsilon^2} + \frac{3 - 2\log\frac{\nu_n\nu_{\bar{n}}}{\mu^2}}{\epsilon}\right)
$$
$$
+ \left(\mathcal{L}_{0T} - \frac{\boldsymbol{\delta}_T}{\epsilon}\right)\left(\delta_2\left[\frac{1+z_1^2}{\bar{z}_1}\right]_+ + \delta_1\left[\frac{1+z_2^2}{\bar{z}_2}\right]_+\right)
$$
$$
- \delta_1\delta_2\left[2\mathcal{L}_{1T} + \mathcal{L}_{0T}\left(3 - 2\log\frac{\nu_n\nu_{\bar{n}}}{\mu^2}\right)\right]
$$
$$
\left. + (\bar{z}_1\delta_2 + \bar{z}_2\delta_1 - \delta_1\delta_2\zeta_2)\boldsymbol{\delta}_T\right], \tag{63}
$$

where we have set $w_{n,\bar{n}} = 1$ for all η_i-independent terms. As discussed earlier, the rapidity divergences appear as the η-divergent terms in the first line. The rapidity-finiteness of the full theory is reflected in the fact that setting $w_n = w_{\bar{n}}$ and $\eta_n = -\eta_{\bar{n}}$ gives a total rate that is free from η poles, which is to be expected since this scheme corresponds to regulating the n and $\bar{n}$ sectors identically; using different schemes for the two sectors spoils the cancellation of rapidity divergences between the sectors. However, resumming the rapidity logarithms requires keeping the $w_{n,\bar{n}}$ and $\eta_{n,\bar{n}}$ scheme dependence. This introduces explicit rapidity divergences in the matrix elements which require rapidity counterterms, from which the RRG may be derived.

In the scheme where $\nu_n\nu_{\bar{n}} = q_L^2$, the purely ϵ-divergent terms (ultraviolet divergences) in the first line of Eq. (63) are canceled by the renormalization constant $Z_{2,(0)}$ for $O_2^{(0)}$

$$
Z_{2,(0)} = 1 + \frac{\bar{\alpha}}{2}\left(\frac{2}{\epsilon^2} + \frac{3 - 2\log\frac{q_L^2}{\mu^2}}{\epsilon}\right), \tag{64}
$$

which follows the product of renormalized operators $O_{2,(0)}^{\dagger\mu}O_{2,(0)}^{\dagger\nu}$ through the Fierz rearrangement. Since $Z_{2,(0)}$ depends only on $\log(q_L^2/\mu^2)$, the scheme $\nu_n\nu_{\bar{n}} = q_L^2$ is enforced at $\mu \sim q_L$ and throughout the μ running when $\mu > q_T$. As we later discuss in Sec. II F, when $\mu \sim q_T$ then q_T is no longer an infrared scale, and then $\nu_{n,\bar{n}}$ can be evolved with the RRG, allowing for the resummation of rapidity logarithms.

The IR divergent terms in the third line of Eq. (63) are the Altarelli-Parisi splitting functions and are reproduced by the infrared divergences in the light-cone distribution operators in the low-energy theory. The remaining divergences are rapidity divergences, and are absorbed by the counterterm in Eq. (46), where

$$
Z_{(0,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T)
$$
$$
= \delta_1\delta_2\left(\boldsymbol{\delta}_T + 2\bar{\alpha}\left(\frac{w_n^2}{\eta_n} + \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}}\right)\right.
$$
$$
\left. \times\left(\frac{\boldsymbol{\delta}_T}{\epsilon} - \mathcal{L}_{0T}\right) + 2\bar{\alpha}\frac{\boldsymbol{\delta}_T}{\epsilon}\log\frac{q_L^2}{\nu_n\nu_{\bar{n}}}\right). \tag{65}
$$

Using the running of the fictional coupling $w_{n,\bar{n}}$ in Eq. (45), we can obtain the rapidity anomalous dimension and rapidity evolution equation for $T_{(0,0)}$, which we further discuss in Sec. II F.

Note that for $\eta_n = -\eta_{\bar{n}}$ and $\nu_{n,\bar{n}} = q_L^2$ there are no additional ultraviolet (UV) divergences in the matrix element of $T_{(0,0)}$ beyond the renormalization of $O_2^{(0)}(\mu)$, indicating that the phase space integral in SCET is UV finite at $O(\alpha_s)$. Similarly, one-loop matrix elements at NLP will also be found to be UV finite. Additional UV divergences in matrix elements of the $T_{(i,j)}$'s would indicate phase space integrals which were sensitive to the UV scale q_L, in which case the RG running of the corresponding $T_{(i,j)}$ would not simply be given by the running of its constituent SCET operators, but would have additional contributions. It is possible that this could complicate the RG running of the $H_{(i,j)}$'s at higher orders in α_s, where the final state phase space can include multiple gluons with individually large $\mathbf{k}_T$ which largely cancel to contribute at small q_T, but this would not affect the one-loop running or the form of the factorization Eq. (14).

Subtracting the counterterms yields the renormalized matrix element

$$\langle p_1^n p_2^{\bar{n}} | T_{(0,0)} \left(q^-, q^+, \mathbf{q}_T, \frac{\mu}{q_T}, \frac{\nu_{n,\bar{n}}}{\mu} \right) | p_1^n p_2^{\bar{n}} \rangle_{1-\text{gluon}}$$

$$= \bar{\alpha} \left(\left(\mathcal{L}_{0T} - \frac{\delta_T}{\epsilon} \right) \left(\delta_2 \left[\frac{1 + z_1^2}{\bar{z}_1} \right]_+ + \delta_1 \left[\frac{1 + z_2^2}{\bar{z}_2} \right]_+ \right) \right.$$

$$- \delta_1 \delta_2 \left[2 \mathcal{L}_{1T} + \mathcal{L}_{0T} \left(3 - 2 \log \frac{\nu_n \nu_{\bar{n}}}{\mu^2} \right) \right]$$

$$\left. + (\bar{z}_1 \delta_2 + \bar{z}_2 \delta_1 - \delta_1 \delta_2 \zeta_2) \delta_T \right), \tag{66}$$

Here, we use the scalar distributional identity

$$\frac{\theta(\bar{z}_1)}{\bar{z}_1^{2+\eta}} = \frac{\delta'(\bar{z}_1)}{\eta} - \delta(\bar{z}_1) + \left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_{++} + \cdots, \tag{68}$$

where the double-plus distribution [53] is defined in Appendix B. We also use the usual expansion

$$\frac{(\mu^2)^{-\eta/2}}{(q_T^2)^{-\eta/2}} = 1 - \frac{\eta}{2} \log \frac{\mu^2}{q_T^2} + \cdots. \tag{69}$$

Equation (67) is finite as $\epsilon \to 0$. Expanding in η_n, we find the bare matrix element of $T_{(2_4,0)}$,

$$\mathcal{M}_{(2_4,0)}^n = \frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_2' \left((\delta_1 + \delta_1') \left(-\frac{2 w_n^2}{\eta_n} + \log \frac{\nu_n^2}{q_T^2} \right) \right.$$

$$\left. - 2 \left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_{++} + 2 \left[\frac{1}{\bar{z}_1} \right]_+ - 1 \right). \tag{70}$$

The $1/\eta_n$ rapidity divergence in (70) is similar in form to that found in the study of NLP jet and soft functions in [36,40,74,80]. The divergence is independent of $\mathbf{q}_T$ and may be absorbed through mixing of $T_{(2_4,0)}$ with the leading-power operator $T_{(0,0)}$, as in Eq. (47), with

$$Z_{(2_4,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T)$$

$$= -2 \frac{\bar{\alpha}}{\pi} \frac{w_n^2}{\eta_n} \omega_1 \omega_2 (\delta(\bar{\omega}_1) + \delta'(\bar{\omega}_1)) \delta'(\bar{\omega}_2). \tag{71}$$

This rapidity renormalization factor is suppressed by one power of q_T^2 relative to the leading term $Z_{(0,0),(0,0)}$ in (65) since it does not contain a factor of $\delta(q_T^2)$, and so the

which, with the replacement $\nu_n \nu_{\bar{n}} = \nu^2$, also reproduces the result in [57].

2. Next-to-leading power example: $T_{(2_4,0)}$

Since $O(\Lambda_{\text{QCD}})$ contributions are not considered, there is no 0-gluon contribution to the matrix element of $T_{(2_4,0)}$, and there is also no $\bar{n}$-sector contribution at $O(\alpha_s)$. Thus at first nontrivial order the graphs $T_{(2_4,0)}$ are those shown in Fig. 1, yielding

$$\mathcal{M}_{(2_4,0)}^n = 2\pi g^2 C_F q^+ q^- \int \frac{d^d k}{(2\pi)^d} w_n^2 \left(\frac{q_L^2}{\nu_n^2} \frac{q^- k^+}{q^+ k^-} \right)^{\eta_n/2} k^+ \delta(k^2) \delta(p_1^- - q^- - k^-) \delta'(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T) \text{Tr} \left[\frac{\not{p}_1}{2} \left(\frac{2 p_1^\alpha - \gamma^\alpha \not{k}}{-2 p_1 \cdot k} + \frac{\bar{n}^\alpha}{k^-} \right) \frac{\not{\bar{n}}}{2} \left(\frac{2 p_{1\alpha} - \not{k} \gamma_\alpha}{-2 p_1 \cdot k} - \frac{\bar{n}_\alpha}{-k^-} \right) \right] \text{Tr} \left[\frac{\not{p}_2}{2} \frac{\not{n}}{2} \right]$$

$$= -\frac{\bar{\alpha}}{\pi} f_\epsilon z_1 z_2 w_n^2 \delta_2' \left(\frac{z_1 q_T}{\nu_n} \right)^{\eta_n} \frac{(2 - 2\bar{z}_1 + (1 - \epsilon)\bar{z}_1^2)}{\bar{z}_1^{2 + \eta_n}}. \tag{67}$$

mixing is consistent with power counting. Equivalently, the divergence may be absorbed by $O(1)$ mixing with the NLP operator

$$\int d^{d-2} \mathbf{p}_T T_{(0,0)}(q^-, q^+, \mathbf{p}_T). \tag{72}$$

This is similar to the cumulant operators introduced in [36,40,74,80] except that Eq. (72) has no upper cutoff $|p_T| < \Lambda$ on the integral.

3. Next-to-leading power example: $T_{(0,2_3)}$

$T_{(0,2_3)}$ provides an example of a matrix element with nontrivial u dependence. Calculating its spin-averaged matrix element, the only contribution comes from an $\bar{n}$-sector gluon, and we find

$$\mathcal{M}_{(0,2_3)}^{\bar{n}} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} w_{\bar{n}}^2 \left(\frac{q_L^2}{\nu_{\bar{n}}^2} \frac{q^+ k^-}{q^- k^+} \right)^{\eta_{\bar{n}}/2} \delta(k^2)$$

$$\times \delta(p_1^- - q^- - k^-) \delta(p_2^+ - q^+) \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)$$

$$\times \text{Tr} \left[\frac{\not{p}_1}{2} \frac{\not{n}}{2} \right] \text{Tr} \left[\frac{1}{u} \bar{v}(p_2) \left(\frac{2 p_{2\alpha} - \gamma_\alpha \not{k}}{-2 p_2 \cdot k} - \frac{\bar{n}_\alpha}{-k^-} \right) \right.$$

$$\left. \times \frac{\not{n}}{2} \gamma_\nu^\perp \gamma_\mu^\perp \Delta^{\alpha\mu}(k) k^\nu v(p_2) \delta \left(u + \frac{k^+}{q^+} \right) \right]. \tag{73}$$

After using $\delta(u + k^+/q^+)/u = -q^+ \delta(u + k^+/q^+)/k^+$ and integrating over the gluon's phase space, the bare matrix element is

$$\mathcal{M}_{(0,2_3)}^{\bar{n}} = -\frac{\bar{\alpha}}{\pi} f_\epsilon w_{\bar{n}}^2 z_1 z_2 \left(\frac{z_2 q_T}{\nu_{\bar{n}}} \right)^{\eta_{\bar{n}}} \frac{1}{\bar{z}_2^{2-\eta_{\bar{n}}}} \delta_1 \delta \left(u + \frac{\bar{z}_2}{z_2} \right). \tag{74}$$

Using distributional identities to extract the pole structure of Eq. (74), we expand to find

$$
\mathcal{M}^{\bar{n}}_{(0,2_3)} = -\frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_1 \left(\delta_2' \left(\frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} - \frac{1}{2} \log \frac{\nu_{\bar{n}}^2}{q_T^2} \right) \right.
$$
$$
\left. + \left[\frac{\theta(\bar{z}_2)}{\bar{z}_2^2} \right]_{++} \right) \delta \left(u + \frac{\bar{z}_2}{z_2} \right),
\qquad (75)
$$

where the rapidity divergence is absorbed by Eq. (47) with the renormalization constant

$$
Z_{(0,2_3),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u)
$$
$$
= -\bar{\alpha}\, \tilde{\theta}(q_T^2) \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} \omega_1 \omega_2 \delta(\bar{\omega}_1) \delta'(\bar{\omega}_2) \delta \left(u + \frac{\bar{\omega}_2}{\omega_2} \right). \quad (76)
$$

Note that if the operator definition of $O_2^{(2B_1)}(x, \hat{t})$ had not included the $\theta(\hat{t})$ convolution discussed in Sec. II A, then its u-space matching coefficient would instead be $C_2^{(2B1)} = 1/u$, and the corresponding expression in (74) would contain only delta functions and single plus distributions in $\bar{z}_2$,

$$
\mathcal{M} \sim \frac{1}{\bar{z}_2^{1-\eta_{\bar{n}}}} \delta \left(u + \frac{\bar{z}_2}{z_2} \right)
$$
$$
= \left(\frac{\delta(\bar{z}_2)}{\eta_{\bar{n}}} + \left[\frac{\theta(\bar{z}_2)}{\bar{z}_2} \right]_+ + O(\eta_{\bar{n}}) \right) \delta \left(u + \frac{\bar{z}_2}{z_2} \right). \quad (77)
$$

Multiplying this by the Wilson coefficient $\sim 1/u$ and integrating over u would then give an unregulated divergence at $\bar{z}_2 = 0$. Instead, keeping the singular $1/u$ dependence in the matrix element of the operator gives the properly regulated result in Eqs. (75) and (76) in terms of δ' and double-plus distributions.

Finally, we can also demonstrate here that the δ regulator does not regulate all rapidity divergences at NLP. Replacing the previous pure rapidity regulator in Eqs. (73) and (74) with the δ regulator, the same expressions read

$$
\mathcal{M}^{\bar{n}\delta}_{(0,2_3)} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} \mathrm{Tr} \left[\frac{\rlap{/}{p}_1}{2} \frac{\rlap{/}{n}}{2} \right] \delta \left(u + \frac{k^+}{q^+} \right) \delta(k^2)
$$
$$
\times \delta(p_1^- - q^- - k^-) \delta(p_2^+ - q^+) \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)
$$
$$
\times \mathrm{Tr} \left[\frac{1}{u} \bar{v}(p_2) \left(\frac{2 p_{2\alpha} - \gamma_\alpha \rlap{/}{k}}{-2 p_2 \cdot k} - \frac{\bar{n}_\alpha}{-k^- - \delta_{\bar{n}}} \right) \right.
$$
$$
\left. \times \frac{\rlap{/}{n}}{2} \gamma_\nu^\perp \gamma_\mu^\perp \left(g^{\alpha\mu} - \frac{n^\alpha k^\mu}{k^+ + \delta_{\bar{n}}} \right) k^\nu v(p_2) \right]
$$
$$
= -\frac{\bar{\alpha}}{\pi} f_\epsilon w_{\bar{n}}^2 z_1 z_2 \frac{1}{\bar{z}_2 (\bar{z}_2 + \delta_{\bar{n}}/p_2^+)} \delta_1 \delta \left(u + \frac{\bar{z}_2}{z_2} \right), \quad (78)
$$

which contains an uncontrolled rapidity divergence when integrated over $\bar{z}_2$. Since the unregulated divergence does not originate from a Wilson line propagator, any regulator

that only modifies the definition of a Wilson line, such as the η regulator of [51], will suffer from similar problems.

4. One-loop results

As shown in previous examples, matrix elements of the $T_{(i,j)}$'s are rapidity divergent and require subtractions via rapidity counterterms proportional to the leading order operator $T_{(0,0)}$. The renormalization constants for the rest of the subleading $T_{(i,j)}$'s are found to be

$$
Z_{(2_1,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u) = \frac{\bar{\alpha}}{\pi} \frac{w_n^2}{\eta_n} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u),
$$
$$
Z_{(0,2_1),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u) = \frac{\bar{\alpha}}{\pi} \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u), \quad (79)
$$

and

$$
Z_{(2_2,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u_1, u_2)
$$
$$
= -\frac{\bar{\alpha}}{\pi} \frac{w_n^2}{\eta_n} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u_1) \delta(u_2),
$$
$$
Z_{(0,2_2),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u_1, u_2)
$$
$$
= -\frac{\bar{\alpha}}{\pi} \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u_1) \delta(u_2), \quad (80)
$$

for the operators $T_{(2_1,0)}$ through $T_{(0,2_2)}$, and

$$
Z_{(2_3,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u)
$$
$$
= -\frac{\bar{\alpha}}{\pi} \frac{w_n^2}{\eta_n} \omega_1 \omega_2 \delta'(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta \left(u + \frac{\bar{\omega}_1}{\omega_1} \right),
$$
$$
Z_{(0,2_3),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, u)
$$
$$
= -\frac{\bar{\alpha}}{\pi} \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} \omega_1 \omega_2 \delta(\bar{\omega}_1) \delta'(\bar{\omega}_2) \delta \left(u + \frac{\bar{\omega}_2}{\omega_2} \right),
$$
$$
Z_{(2_4,0),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T)
$$
$$
= -2\frac{\bar{\alpha}}{\pi} \frac{w_n^2}{\eta_n} \omega_1 \omega_2 (\delta(\bar{\omega}_1) + \delta'(\bar{\omega}_1)) \delta'(\bar{\omega}_2),
$$
$$
Z_{(0,2_4),(0,0)}(\omega_1, \omega_2, \mathbf{q}_T)
$$
$$
= -2\frac{\bar{\alpha}}{\pi} \frac{w_{\bar{n}}^2}{\eta_{\bar{n}}} \omega_1 \omega_2 (\delta(\bar{\omega}_2) + \delta'(\bar{\omega}_2)) \delta'(\bar{\omega}_1), \quad (81)
$$

for the remaining operators $T_{(2_3,0)}$ through $T_{(0,2_4)}$.

In contrast to the leading power operator, the matrix elements of the power suppressed operators $T_{(i,j)}$ are individually rapidity divergent even when setting $w_n = w_{\bar{n}} = 1$ and $\eta_n = -\eta_{\bar{n}}$. Nevertheless, these divergences cancel pairwise between $T_{(2_1,0)}$ and $T_{(0,2_1)}$, and $T_{(2_2,0)}$ and $T_{(0,2_2)}$. The divergences also cancel in the sum over the four operators in Eq. (81) when weighted and integrated against their appropriate prefactor $H_{(i,j)}(\{u\}) K^{(i,j)}_{(k,\ell)}$.

The cancellation of rapidity divergences in the total rate reflects the rapidity finiteness of the total NLP cross section in SCET and is a nontrivial check on the validity of the EFT expansion. In Sec. III B we will show that this cancellation can be understood without an explicit regulator, in which case the correct treatment of overlap subtraction graphs, which vanished here when using the pure rapidity regulator, is critical.

The hard-scale matching coefficients of all the subleading operators $T_{(i,j)}$ have the same LL anomalous dimensions [56], so these cancellations are manifestly maintained to all orders in the leading-log approximation. Since the rate must be finite beyond leading logarithms, finiteness of the theory will place constraints on rapidity divergences which are beyond the scope of this paper.

As in Eq. (9), the $T_{(i,j)}$ operators are matched onto a theory solely consisting of light-cone distribution operators, defined as

$$O_q(\ell^-) = \frac{1}{2\pi} \int d\xi e^{-i\xi\ell^-} \bar{\psi}_n(\bar{n}\xi) \frac{\slashed{\bar{n}}}{2} W(\bar{n}\xi, 0) \psi_n(0),$$

$$O_{\bar{q}}(\ell^+) = \frac{1}{2\pi} \int d\xi e^{-i\xi\ell^+} \bar{\psi}_{\bar{n}}(0) \frac{\slashed{n}}{2} W(0, n\xi) \psi_{\bar{n}}(n\xi). \quad (82)$$

Since the renormalized partonic matrix element of the product of these soft theory operators is [64,81]

$$\langle p_1^n p_2^{\bar{n}} | O_q(q^-) O_{\bar{q}}(q^+) | p_1^n p_2^{\bar{n}} \rangle$$
$$= \left(\delta_1 - \frac{\bar{\alpha}}{\epsilon} \left[\frac{1+z_1^2}{\bar{z}_1} \right]_+ \right) \left(\delta_2 - \frac{\bar{\alpha}}{\epsilon} \left[\frac{1+z_2^2}{\bar{z}_2} \right]_+ \right) + \cdots, \quad (83)$$

and since these IR divergences are precisely reproduced in the renormalized matrix element of $T_{(0,0)}$ [see Eq. (63)], the leading-power soft matching coefficient is then

$$C_{S,(0,0)} \left(z_1, z_2, \mathbf{q}_T, \frac{\mu}{q_T}, \frac{\nu_{n,\bar{n}}}{\mu} \right)$$
$$= \delta_1 \delta_2 \boldsymbol{\delta}_T + \bar{\alpha} \left\{ -\delta_1 \delta_2 \left[2\mathcal{L}_{1T} + \mathcal{L}_{0T} \left(3 - 2\log\frac{\nu_n \nu_{\bar{n}}}{\mu^2} \right) \right] \right.$$
$$+ \left(\delta_2 \left[\frac{1+z_1^2}{\bar{z}_1} \right]_+ + \delta_1 \left[\frac{1+z_2^2}{\bar{z}_2} \right]_+ \right) \mathcal{L}_{0T}$$
$$\left. + (\bar{z}_1\delta_2 + \bar{z}_2\delta_1 - \delta_1\delta_2\zeta_2)\boldsymbol{\delta}_T \right\}. \quad (84)$$

This also provides the fixed order expansion of $V_{(0,0),(0,0)}$,

$$V_{(0,0),(0,0)} \left(z_1, z_2, \mathbf{q}_T, \frac{\mu}{q_T}, \frac{q_L}{\nu_{n,\bar{n}}} \right)$$
$$= \delta_1 \delta_2 \boldsymbol{\delta}_T + 2\bar{\alpha}\mathcal{L}_{0T} \log\frac{q_L^2}{\nu_n \nu_{\bar{n}}} + \cdots, \quad (85)$$

where higher order terms can be generated using the running in Sec. II F.

At subleading power the renormalized matrix elements of $T_{(i,j)}$ begin at $O(\alpha_s)$ and thus match onto the tree-level term $\delta(\bar{z}_1)\delta(\bar{z}_2)$ of Eq. (83). The renormalized matrix elements of $T_{(i,j)}$ are thus equal to the soft matching coefficients $C_{S,(i,j)}$. Suppressing their scale dependence, the first four NLP soft matching coefficients are

$$C_{S,(2_1,0)}(z_1, z_2, \mathbf{q}_T) = -\frac{\bar{\alpha}}{\pi} \delta_2 \left(\left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_+ + \frac{1}{2} \delta_1 \log\frac{\nu_n^2}{q_T^2} \right)$$
$$\times \delta\left(u + \frac{\bar{z}_1}{z_1} \right),$$

$$C_{S,(0,2_1)}(z_1, z_2, \mathbf{q}_T) = -\frac{\bar{\alpha}}{\pi} \delta_1 \left(\left[\frac{\theta(\bar{z}_2)}{\bar{z}_2} \right]_+ + \frac{1}{2} \delta_2 \log\frac{\nu_{\bar{n}}^2}{q_T^2} \right)$$
$$\times \delta\left(u + \frac{\bar{z}_2}{z_2} \right),$$

$$C_{S,(2_2,0)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} \delta_2 \left(\left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_+ + \frac{1}{2} \delta_1 \log\frac{\nu_n^2}{q_T^2} \right)$$
$$\times \delta\left(u_1 + \frac{\bar{z}_1}{z_1} \right) \delta(u_1 - u_2),$$

$$C_{S,(0,2_2)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} \delta_1 \left(\left[\frac{\theta(\bar{z}_2)}{\bar{z}_2} \right]_+ + \frac{1}{2} \delta_2 \log\frac{\nu_{\bar{n}}^2}{q_T^2} \right)$$
$$\times \delta\left(u_2 + \frac{\bar{z}_2}{z_2} \right) \delta(u_1 - u_2), \quad (86)$$

and the remaining four matching coefficients are

$$C_{S,(2_3,0)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_2 \delta\left(u + \frac{\bar{z}_1}{z_1} \right)$$
$$\times \left(\frac{1}{2} \delta_1' \log\frac{\nu_n^2}{q_T^2} - \left[\frac{\theta(\bar{z}_1)}{\bar{z}_1^2} \right]_{++} \right),$$

$$C_{S,(0,2_3)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_1 \delta\left(u + \frac{\bar{z}_2}{z_2} \right)$$
$$\times \left(\frac{1}{2} \delta_2' \log\frac{\nu_{\bar{n}}^2}{q_T^2} - \left[\frac{\theta(\bar{z}_2)}{\bar{z}_2^2} \right]_{++} \right),$$

$$C_{S,(2_4,0)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_2' \left((\delta(\bar{z}_1) + \delta'(\bar{z}_1)) \log\frac{\nu_n^2}{q_T^2} \right.$$
$$\left. - 2\left[\frac{\theta(\bar{z}_1)}{\bar{z}_1^2} \right]_{++} + 2\left[\frac{\theta(\bar{z}_1)}{\bar{z}_1} \right]_+ - 1 \right),$$

$$C_{S,(0,2_4)}(z_1, z_2, \mathbf{q}_T) = \frac{\bar{\alpha}}{\pi} z_1 z_2 \delta_1' \left((\delta(\bar{z}_2) + \delta'(\bar{z}_2)) \log\frac{\nu_{\bar{n}}^2}{q_T^2} \right.$$
$$\left. - 2\left[\frac{\theta(\bar{z}_2)}{\bar{z}_2^2} \right]_{++} + 2\left[\frac{\theta(\bar{z}_2)}{\bar{z}_1} \right]_+ - 1 \right). \quad (87)$$

Matching QCD onto SCET at $\mu = q_L$ and $\nu_{n,\bar{n}} = q_L$, these matrix elements have large logarithms of q_L^2/q_T^2.

We will discuss the resummation of these logarithms using the rapidity renormalization group in Sec. II F.

E. C_{ff} at fixed order

It is useful at this stage to check the fixed order results for C_{ff} by comparing with the corresponding QCD calculation. At leading power, the $O(\alpha_s)$ expression for $C_{ff}^{(0)}$ in SCET is given by $C_{S(0,0)}$ in Eq. (84) with $\nu_{n,\bar n} = q_L$ and multiplied by the hard function $H_{(0,0)} = C_2^{(0)\dagger} C_2^{(0)}$. After integrating $d\Omega_T$, this gives the one-loop expression

$$
\begin{aligned}
&C_{ff}^{(0)}(z_1, z_2, q_L^2, q_T^2) \\
&= \bar\alpha \Bigg\{ \delta_1 \delta_2 \delta(q_T^2) \left(-\log^2 \frac{q_L^2}{\mu^2} + 3\log \frac{q_L^2}{\mu^2} - 8 + 7\zeta_2 \right) \\
&\quad + \left[\frac{1}{q_T^2} \right]_+^{\mu^2} \left(\delta_1 \left[\frac{1+z_2^2}{\bar z_2} \right]_+ + \delta_2 \left[\frac{1+z_1^2}{\bar z_1} \right]_+ \right) \\
&\quad - \delta_1 \delta_2 \left[2 \left[\frac{\log \frac{q_T^2}{\mu^2}}{q_T^2} \right]_+^{\mu^2} + \left[\frac{1}{q_T^2} \right]_+^{\mu^2} \left(3 - 2\log \frac{q_L^2}{\mu^2} \right) \right] \\
&\quad + \delta(q_T^2)(\bar z_1 \delta_2 + \bar z_2 \delta_1 - \zeta_2 \delta_1 \delta_2) \Bigg\}.
\end{aligned} \tag{88}
$$

At NLP, adding up the contributions from the unsummed matching coefficients in Eqs. (86) and (87) with $\nu_{n,\bar n} = q_L$ weighted by the corresponding coefficients $H_{(i,j)} K_{(k,\ell)}^{(i,j)}$ gives the $O(\alpha_s)$ coefficient function

$$
\begin{aligned}
C_{ff}^{(2)}(z_1, z_2, \mathbf{q}_T) &= \bar\alpha z_1 z_2 \Bigg[\left(2\log \frac{\hat s}{q_T^2} - 3 \right) \delta_1' \delta_2' \\
&\quad + \left(2\log \frac{\hat s}{q_T^2} + 1 \right) (\delta_1' \delta_2 + \delta_1 \delta_2') + 4\delta_1 \delta_2 \\
&\quad - \delta_1' \left[\frac{2 - 2\bar z_2 + \bar z_2^2}{\bar z_2^2} \right]_{++} - \delta_2' \left[\frac{2 - 2\bar z_1 + \bar z_1^2}{\bar z_1^2} \right]_{++} \\
&\quad - 2 \left(\delta_1 \left[\frac{1}{\bar z_2^2} \right]_{++} + \delta_2 \left[\frac{1}{\bar z_1^2} \right]_{++} \right) \Bigg],
\end{aligned} \tag{89}
$$

where, along with $q_L^2 = z_1 z_2 \hat s$, we have used the identities

$$
\begin{aligned}
\left[\frac{\theta(\bar z)}{\bar z} \right]_{++} &= \left[\frac{\theta(\bar z)}{\bar z} \right]_+ + \delta'(\bar z), \\
[\theta(\bar z)]_{++} &= 1 + \frac{1}{2}\delta'(\bar z) - \delta(\bar z).
\end{aligned} \tag{90}
$$

These results may be compared with the direct QCD calculation. C_{ff} is determined in QCD by the partonic rate

$$
\begin{aligned}
\mathcal{R}_{\mathrm{QCD}} &= -\int \frac{d^d x}{(2\pi)^d} \langle p_1 p_2 | \bar\psi(x) \gamma^\mu \psi(x) \bar\psi(0) \gamma_\mu \psi(0) | p_1 p_2 \rangle \\
&= \frac{1}{2} \int \frac{d^d x}{(2\pi)^d} \langle p_1 p_2 | \bar\psi(x) \gamma^\mu \psi(0) \bar\psi(0) \gamma_\mu \psi(x) | p_1 p_2 \rangle.
\end{aligned} \tag{91}
$$

The single gluon real emission contribution evaluates to

$$
\begin{aligned}
\mathcal{R}_{\mathrm{QCD}}^{1g} &= \frac{\bar\alpha}{\pi} f_\epsilon \frac{\delta(\bar z_1 \bar z_2 \hat s - \mathbf{q}_T^2)}{\bar z_1 \bar z_2} [2 - 2(\bar z_1 + \bar z_2) \\
&\quad + (\bar z_1^2 + \bar z_2^2) - \epsilon(\bar z_1 + \bar z_2)^2].
\end{aligned} \tag{92}
$$

Expanding (92) in powers of q_T^2/q_L^2,

$$
\mathcal{R}_{\mathrm{QCD}}^{1g} = \mathcal{R}_{\mathrm{QCD}}^{(0)1g} + \mathcal{R}_{\mathrm{QCD}}^{(2)1g} + \cdots, \tag{93}
$$

is straightforward away from $\bar z_1 = \bar z_2 = 0$,

$$
\begin{aligned}
\mathcal{R}_{\mathrm{QCD}}^{1g}\big|_{\bar z_1 \neq 0} &= \frac{\bar\alpha}{\pi} \frac{f_\epsilon}{\mathbf{q}_T^2} \Bigg[\delta(\bar z_2) \left(\frac{2 - 2\bar z_1 + \bar z_1^2}{\bar z_1} \right) \\
&\quad - 2\frac{q_T^2}{\hat s} \delta_2 \left(\frac{1}{\bar z_1^2} \right) - \frac{q_T^2}{\hat s} \delta_2' \left(\frac{2 - 2\bar z_1 + \bar z_1^2}{\bar z_1^2} \right) \\
&\quad + O\left(\frac{q_T^4}{\hat s^2} \right) \Bigg]
\end{aligned}
$$

(with a similar result for $\bar z_2 \neq 0$), but care is required at the singular points. At leading power, $\mathcal{R}$ may be written

$$
\mathcal{R}_{\mathrm{QCD}}^{(0)1g} = \frac{\bar\alpha f_\epsilon}{\pi \mathbf{q}_T^2} (A^{(0)} \delta_1 \delta_2 + \delta_2 [f_n(\bar z_1)]_+ + \delta_1 [f_{\bar n}(\bar z_2)]_+), \tag{94}
$$

where, from (94),

$$
f_{n,\bar n}(\bar z) = \left(\frac{2 - 2\bar z + \bar z^2}{\bar z} \right) \tag{95}
$$

and $A^{(0)}$ is determined from the integrated rate

$$
A = \frac{q_T^2}{\frac{\bar\alpha}{\pi} f_\epsilon} \int_0^1 d\bar z_1 d\bar z_2 \mathcal{R}_{\mathrm{QCD}}^{1g} = A^{(0)} + \frac{q_T^2}{\hat s} A^{(2)} + \cdots. \tag{96}
$$

Similarly at NLP the rate has the general form

$$
\begin{aligned}
\mathcal{R}_{\mathrm{QCD}}^{(2)1g} &= \frac{\bar\alpha f_\epsilon}{\pi \hat s} (A^{(2)} \delta_1 \delta_2 + B^{(2)} \delta_1' \delta_2' + C^{(2)} (\delta_1' \delta_2 + \delta_1 \delta_2') \\
&\quad + \delta_2' [g_n(\bar z_1)]_{++} + \delta_2 [h_n(\bar z_1)]_{++} + \delta_1' [g_{\bar n}(\bar z_2)]_{++} \\
&\quad + \delta_1 [h_{\bar n}(\bar z_2)]_{++}),
\end{aligned} \tag{97}
$$

where

$$g_{n,\bar{n}}(\bar{z}) = -\left(\frac{2 - 2\bar{z} + \bar{z}^2}{\bar{z}^2}\right),$$

$$h_{n,\bar{n}}(\bar{z}) = -2\left(\frac{1}{\bar{z}^2}\right), \tag{98}$$

and the constants B and C are given by the appropriate moments of the rate,

$$B = \frac{q_T^2}{\frac{\bar{\alpha}}{\pi} f_\epsilon} \int_0^1 d\bar{z}_1 d\bar{z}_2 \bar{z}_1 \bar{z}_2 \mathcal{R}_{\text{QCD}}^{1g} = B^{(0)} + \frac{q_T^2}{\hat{s}} B^{(2)} + \cdots,$$

$$C = -\frac{q_T^2}{\frac{\bar{\alpha}}{\pi} f_\epsilon} \int_0^1 d\bar{z}_1 d\bar{z}_2 \bar{z}_1 \mathcal{R}_{\text{QCD}}^{1g} = -\frac{q_T^2}{\frac{\bar{\alpha}}{\pi} f_\epsilon} \int_0^1 d\bar{z}_1 d\bar{z}_2 \bar{z}_2 \mathcal{R}_{\text{QCD}}^{1g}$$

$$= C^{(0)} + \frac{q_T^2}{\hat{s}} C^{(2)} + \cdots. \tag{99}$$

The integrals in Eqs. (96) and (99) give the endpoint constants

$$A^{(0)} = 2 \log \frac{\hat{s}}{q_T^2} - 3 - \epsilon,$$

$$A^{(2)} = 4,$$

$$B^{(2)} = 2 \log \frac{\hat{s}}{q_T^2} - 3,$$

$$C^{(2)} = 2 \log \frac{\hat{s}}{q_T^2} + 1, \tag{100}$$

where we drop the ϵ dependence in the NLP terms since, unlike the LP rate, the NLP rate contains no infrared divergences stemming from a $1/q_T^2$ prefactor.

At leading power, applying (B23) and (B24) gives

$$\mathcal{R}_{\text{QCD}}^{(0)1g} = \frac{\bar{\alpha} f_\epsilon}{\pi \mathbf{q}_T^2} \left\{ \left(2 \log \frac{\hat{s}}{q_T^2} - 3 - \epsilon\right) \delta_1 \delta_2 + \delta_2 \left[\frac{2 - 2\bar{z}_1 + \bar{z}_1^2}{\bar{z}_1}\right]_+ + \delta_1 \left[\frac{2 - 2\bar{z}_2 + \bar{z}_2^2}{\bar{z}_1}\right]_+ \right\}$$

$$= \bar{\alpha} \left[\delta_1 \delta_2 \boldsymbol{\delta}_T \left(\frac{2}{\epsilon^2} + \frac{3 - 2 \log \frac{\hat{s}}{\mu^2}}{\epsilon}\right) + \left(\mathcal{L}_{0T} - \frac{\boldsymbol{\delta}_T}{\epsilon}\right) \left(\delta_2 \left[\frac{1 + z_1^2}{\bar{z}_1}\right]_+ + \delta_1 \left[\frac{1 + z_2^2}{\bar{z}_2}\right]_+ \right) \right.$$

$$\left. - \delta_1 \delta_2 \left[2\mathcal{L}_{1T} + \mathcal{L}_{0T} \left(3 - 2 \log \frac{\hat{s}}{\mu^2}\right)\right] + (\bar{z}_1 \delta_2 + \bar{z}_2 \delta_1 - \delta_1 \delta_2 \zeta_2) \boldsymbol{\delta}_T \right]. \tag{101}$$

The LP vertex correction gives an additional contribution

$$\mathcal{R}_{\text{QCD}}^{(0)\text{virt}} = \bar{\alpha} \delta_1 \delta_2 \boldsymbol{\delta}_T \left[-\left(\frac{2}{\epsilon^2} + \frac{3 - 2 \log \frac{\hat{s}}{\mu^2}}{\epsilon}\right) - \log^2 \frac{q_L^2}{\mu^2} + 3 \log \frac{q_L^2}{\mu^2} - 8 + 7\zeta_2\right]. \tag{102}$$

Combining the finite pieces of Eqs. (102) and (101) and then integrating $d\Omega_T$ reproduces the SCET result for $C_{ff}^{(0)}$ in (88). The remaining divergent terms are equal to the infrared divergences of the light-cone distribution operator matrix elements and thus cancel in the matching onto the soft theory.

QCD virtual corrections do not contribute to the NLP coefficient function at $O(\alpha_s)$, and so $C_{ff}^{(2)}$ is determined from Eqs. (97)–(100). After integrating $d\Omega_T$, this gives

$$C_{ff}^{(2)}(z_1, z_2, q_L^2, q_T^2) = \bar{\alpha} z_1 z_2 \left[4\delta_1 \delta_2 + \left(2 \log \frac{\hat{s}}{q_T^2} - 3\right) \delta_1' \delta_2' + \left(2 \log \frac{\hat{s}}{q_T^2} + 1\right) (\delta_1' \delta_2 + \delta_1 \delta_2') \right.$$

$$\left. - \left(\delta_1' \left[\frac{2 - 2\bar{z}_2 + \bar{z}_2^2}{\bar{z}_2^2}\right]_{++} + \delta_2' \left[\frac{2 - 2\bar{z}_1 + \bar{z}_1^2}{\bar{z}_1^2}\right]_{++} \right) - 2\left(\delta_1 \left[\frac{1}{\bar{z}_2^2}\right]_{++} + \delta_2 \left[\frac{1}{\bar{z}_1^2}\right]_{++} \right) \right], \tag{103}$$

in agreement with the SCET result in Eq. (89). Thus, the SCET result and the expanded QCD result agree to NLP, as required.

Our fixed order results may also be compared with those obtained in [53].[4] In that reference, the DY rate was determined up to NLP by expanding the QCD matrix element in the n collinear, $\bar{n}$ collinear, and soft limits, regulating the ensuing rapidity divergences, and combining

the results. The results in that reference are also in agreement with the expanded QCD results in this section, but are presented in different variables, which makes the comparison more involved. We have checked that our results are in agreement with theirs; details of this comparison are given in Appendix C.

F. Rapidity running

Rapidity logarithms arise in this formalism as a scheme dependence in summing together the individually divergent

---

[4]Similar results, integrated over rapidity, were presented in [47].

contributions from the n and $\bar{n}$ sectors to a given matrix element. It was argued in [57] that in this formalism rapidity renormalization should be performed at the matching scale onto the light-cone distribution operators in order to ensure that Wilson coefficients in SCET are independent of infrared physics.

As discussed in [57], the rapidity regulators in the two sectors are fixed by matching at the hard scale from QCD onto SCET by the requirement that when $\mu \gg \mu_S$ the Wilson coefficients of SCET are independent of infrared energy scales of order μ_S. In the rapidity regularization scheme used here, this corresponds to choosing $\nu_n \nu_{\bar{n}} = q_L^2$, which, as discussed in Sec. II D, corresponds to using the same rapidity regulator in the n and $\bar{n}$ sectors, and is required for the rapidity divergences to cancel in the EFT. The necessity of this choice can be seen from the q_T dependence of the leading-power matrix element of Eq. (66), which contains the term

$$\langle p_1^n p_2^{\bar{n}} | T_{(0,0)}(q^-, q^+, \mathbf{q}_T, \nu_{n,\bar{n}}) | p_1^n p_2^{\bar{n}} \rangle_{1-\text{gluon}}$$
$$= 2\bar{\alpha} \delta_1 \delta_2 \mathcal{L}_{0T} \log \frac{\nu_n \nu_{\bar{n}}}{\mu^2} + \cdots. \qquad (104)$$

Since physical quantities are independent of the rapidity regulator, any variation in $\nu_n \nu_{\bar{n}}$ in the matrix element of $T_{(0,0)}$ must be compensated by a Wilson coefficient proportional to $\mathcal{L}_{0T}$ in the EFT. This variation would then introduce nonanalytic dependence on the IR scale q_T into the effective Lagrangian through the Wilson coefficient, which is inconsistent with factorization of hard and soft scales.

However, at the soft scale $\mu \sim q_T$ where SCET operators are matched onto light-cone distribution operators, the scale q_T is no longer an infrared scale in the EFT, and the Wilson coefficients are free to have nonanalytic dependence on q_T. The operators $T_{(i,j)}$ may therefore be run in $\nu_{n,\bar{n}}$ to minimize rapidity logarithms in the matching coefficients C_S in Eqs. (84), (86), and (87). These operators obey the RRG equation

$$\frac{d}{d \log \nu_{n,\bar{n}}} T_{(i,j)}(q^-, q^+, \mathbf{q}_T, \nu_{n,\bar{n}})$$
$$= \sum_{k,\ell} (\gamma^{n,\bar{n}}_{(i,j),(k,\ell)} * T_{(k,\ell)})(q^-, q^+, \mathbf{q}_T, \nu_{n,\bar{n}}), \qquad (105)$$

where $\gamma^{n,\bar{n}}$ is the rapidity anomalous dimension for each sector, and we define the convolution $*$ by

$$(f * g)(\lambda_1, \lambda_2, \mathbf{k}_T) \equiv \int \frac{d\omega_1}{\omega_1} \frac{d\omega_2}{\omega_2} d^{d-2} \mathbf{p}_T$$
$$\times f(\omega_1, \omega_2, \mathbf{p}_T) g\left(\frac{\lambda_1}{\omega_1}, \frac{\lambda_2}{\omega_2}, \mathbf{k}_T - \mathbf{p}_T\right). \qquad (106)$$

The solution to Eq. (105) can be written in the form of Eq. (13),

$$T_{(i,j)}(q^-, q^+, \mathbf{q}_T, \nu_{n,\bar{n}} = q_L)$$
$$= \sum_{k,\ell} (V_{(i,j)(k,\ell)}(q_L, \nu_{n,\bar{n}}) * T_{(k,\ell)}(\nu_{n,\bar{n}}))(q^-, q^+, \mathbf{q}_T). \qquad (107)$$

The explicit form of this solution to the RRG in momentum space can be found using the techniques in [19].

From the counterterm definitions in Eqs. (46) and (47) relating the bare and renormalized operators, and using the fact that the bare operators are independent of the parameters $\nu_{n,\bar{n}}$ (as guaranteed by the fictional coupling $w_{n,\bar{n}}$), the rapidity anomalous dimensions for the operators $T_{(i,j)}$ may be calculated in terms of the renormalization constants as

$$\gamma^{n,\bar{n}}_{(i,j),(k,\ell)} = -\sum_{\kappa,\lambda} Z^{-1}_{(i,j),(\kappa,\lambda)} * \frac{d}{d \log \nu_{n,\bar{n}}} Z_{(\kappa,\lambda),(k,\ell)}. \qquad (108)$$

Here, the inverse counterterm satisfies the relation

$$\sum_{\kappa,\lambda} (Z^{-1}_{(i,j)(\kappa,\lambda)} * Z_{(\kappa,\lambda)(k,\ell)})(\omega_1, \omega_2, \mathbf{q}_T)$$
$$= \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(\mathbf{q}_T) \delta_{ik} \delta_{j\ell}. \qquad (109)$$

At leading power the rapidity anomalous dimension of $T_{(0,0)}$ is calculated from the renormalization constant in Eq. (65), which gives

$$\gamma^n_{(0,0),(0,0)} = \gamma^{\bar{n}}_{(0,0),(0,0)} = 2\bar{\alpha} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \mathcal{L}_{0T}. \qquad (110)$$

The leading-power operator $T_{(0,0)}$ thus obeys the RRG equation

$$\frac{d}{d \log \nu_{n,\bar{n}}} T_{(0,0)}(\omega_1, \omega_2, \mathbf{q}_T, \nu_{n,\bar{n}})$$
$$= 2\bar{\alpha} \int d^2 \mathbf{p}_T \mathcal{L}_{0T}(\mathbf{q}_T - \mathbf{p}_T, \mu) T_{(0,0)}(\omega_1, \omega_2, \mathbf{p}_T, \nu_{n,\bar{n}}), \qquad (111)$$

similar to the results in [57], and with all the complications of running and scale setting of vector distributions described in [19]. Symmetrically, this RRG equation begins in the scheme $\nu_{n,\bar{n}} = q_T$ where the logarithms of $T_{(0,0)}$ are minimized, and runs up to the scheme $\nu_{n,\bar{n}} = q_L$ which, as we have argued, reproduces the QCD result.

At subleading power, the rapidity mixing of each $T_{(i,j)}$ with the leading order $T_{(0,0)}$ may easily be read off from Eqs. (79)–(81):

$$\gamma^n_{(2_1,0),(0,0)} = \gamma^{\bar{n}}_{(0,2_1),(0,0)} = -\frac{\bar{\alpha}}{\pi} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u),$$

$$\gamma^n_{(2_2,0),(0,0)} = \gamma^{\bar{n}}_{(0,2_2),(0,0)} = \frac{\bar{\alpha}}{\pi} \delta(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta(u_1) \delta(u_2),$$

$$\gamma^n_{(2_3,0),(0,0)} = \frac{\bar{\alpha}}{\pi} \omega_1 \omega_2 \delta'(\bar{\omega}_1) \delta(\bar{\omega}_2) \delta\left(u + \frac{\bar{\omega}_1}{\omega_1}\right),$$

$$\gamma^{\bar{n}}_{(0,2_3),(0,0)} = \frac{\bar{\alpha}}{\pi} \omega_1 \omega_2 \delta(\bar{\omega}_1) \delta'(\bar{\omega}_2) \delta\left(u + \frac{\bar{\omega}_2}{\omega_2}\right),$$

$$\gamma^n_{(2_4,0),(0,0)} = 2\frac{\bar{\alpha}}{\pi} \omega_1 \omega_2 \delta'(\bar{\omega}_1)(\delta(\bar{\omega}_2) + \delta'(\bar{\omega}_2)),$$

$$\gamma^{\bar{n}}_{(0,2_4),(0,0)} = 2\frac{\bar{\alpha}}{\pi} \omega_1 \omega_2 (\delta(\bar{\omega}_1) + \delta'(\bar{\omega}_1)) \delta'(\bar{\omega}_2). \tag{112}$$

As noted in [38], since each subleading $T_{(i,j)}$ only has a nonvanishing matrix element beginning at $O(\alpha_s)$, calculating the complete rapidity renormalization for each $T_{(i,j)}$ requires calculating matrix elements at $O(\alpha_s^2)$. There will be some constraints on these rapidity anomalous dimensions because of μ independence of the final result [51], as discussed in this formalism in [57], but the full calculation is beyond the scope of this paper and will be the subject of future work.

III. OVERLAP SUBTRACTIONS AT NLP

As discussed in [55,57], in this formulation of SCET it is necessary to subtract the double-counting of low-energy degrees of freedom which are simultaneously below the cutoff of both the n and $\bar{n}$ sectors, analogous to zero-bin subtraction in SCET [59]. Rapidity logarithms in this formulation of SCET arise from the scheme dependence in summing the individually rapidity divergent diagrams in each sector and subtracting the corresponding overlap.

In this paper we have used a rapidity renormalization scheme in which overlap subtraction graphs vanish; while this is convenient for calculations, it obscures the cancellations that occur between different operators in different regions of phase space which are required to obtain a rapidity-finite result. In this section we generalize the overlap subtraction prescription to NLP and repeat the calculations without a rapidity regulator in order to explicitly show the cancellation of rapidity divergences due to the overlap subtraction, similar to what was done at LP in [57].

At LP, the zero-bin prescription of [59] has been shown to be equivalent to the nonperturbative subtraction definition of dividing the naïve matrix element by a vacuum expectation value of Wilson lines [82–84]. This equivalence also holds for the overlap prescription of [55,56]. At subleading power, however, this simple prescription does not hold: matrix elements of the NLP operators $T_{(i,j)}$ begin at $O(\alpha_s)$, and thus dividing by a vacuum expectation value of the form $(1 + O(\alpha_s))$ does not provide the necessary $O(\alpha_s)$ subtraction to regulate their matrix elements. Calculations of probabilities in the effective theory

therefore require a systematic way to implement the necessary overlap subtraction. In this section we describe a simple diagram-based prescription to perform the overlap subtraction at subleading powers, and illustrate in the case of DY at NLP that it is required to obtain the correct, finite, result. This allows us to extend the LP discussion of [57] on the relationship between scheme dependence and rapidity logarithms up to NLP. We show that the previous observation in Sec. II D 4—that at NLP rapidity divergences do not cancel for matrix elements of individual operators, but instead cancel between distinct operators—occurs because different linear combinations of operators are required to reproduce the correct rate in different regions of phase space.

Consider, for example, the process in Fig. 1 in which a gluon is produced in DY annihilation in addition to the lepton pair. In SCET this corresponds to two distinct processes in which the gluon is emitted in the n sector or the $\bar{n}$ sector. At NLP, the first receives contributions from the $T_{(2_i,0)}$ operators while the second receives contributions from the corresponding $T_{(0,2_i)}$ operators. Since in loop graphs all momenta are integrated over, the first class of operators will give nonvanishing spurious contributions in the momentum region described by the second, and vice versa. Thus, the overlap subtraction procedure at NLP necessarily involves cancellations between different operators, and the subtraction required in order to avoid overcounting in each is found by taking the wrong limit of matrix elements in the other sector. In the symmetric process that we are examining in this paper, this may be achieved by subtracting one-half of each of the wrong limits from each sector. Schematically, we have the prescription

$$P_{\text{SCET}} = P_n + P_{\bar{n}} - \frac{1}{2}(P_{n\to\bar{n}} + P_{\bar{n}\to n}), \tag{113}$$

where P_i is the probability to produce a gluon in the i sector and the subscripts $i \to j$ denote the wrong sector limits.

The power counting of these subtractions follows the power counting of the limit in which the gluon is taken. An n-sector gluon has the scaling $k_n^-/q^- \sim O(1)$, $k_n^+/q^+ \sim O(q_T^2/q_L^2)$, while its wrong-sector limit has the scaling $k_{n\to\bar{n}}^-/q^- \sim O(q_T^2/q_L^2)$ and $k_{n\to\bar{n}}^+/q^+ \sim O(1)$. This definition of overlap subtraction ensures that probabilities in QCD are properly reproduced to the appropriate order by SCET in all regions of phase space. This prescription is inherently perturbative, and further work is required to determine an operator definition of overlap subtraction which correctly reproduces QCD probabilities at both leading and next-to-leading power.

In the next subsection we review the discussion of overlap subtraction at LP presented in [57] using the prescription (113). We then demonstrate that the same prescription may be used to calculate the NLP coefficient

function $C_{ff}^{(2)}$, and discuss the nature of the overlap subtraction in various regions of phase space.

A. Overlap subtraction and scheme dependence at LP

The DY cross section at LP is determined by the spin-averaged matrix element of $T_{(0,0)}$, which takes the general form

$$\mathcal{M}_{(0,0)} = \frac{\bar{\alpha} f_\epsilon}{\pi q_T^2} (A_{(0,0)} \delta_1 \delta_2 + \delta_2 [f_n(\bar{z}_1)]_+ + \delta_1 [f_{\bar{n}}(\bar{z}_2)]_+),$$

(114)

where as before we define $\delta_1 \equiv \delta(\bar{z}_1)$, $\delta_2 \equiv \delta(\bar{z}_2)$. Away from the singular point $\bar{z}_1 = \bar{z}_2 = 0$ the unregulated n- and $\bar{n}$-sector contributions to the matrix element of $T_{(0,0)}$ are determined by the graphs in Fig. 1 and their $\bar{n}$-sector equivalents, and are given by Eq. (54) (and the corresponding expression in the $\bar{n}$ sector) with $\omega_n = 1$ and $\eta_n = 0$. This immediately gives the functions $f_n(\bar{z}_1)$ and $f_{\bar{n}}(\bar{z}_2)$ in (95) which describe the spectrum away from the endpoint. Since each $f_{n,\bar{n}}$ only receives contributions from a single sector, there is no overcounting, and these expressions are finite and well-defined without a rapidity regulator.

The constant $A_{(0,0)}$ may be most simply obtained by integrating the rate over $\bar{z}_1$ and $\bar{z}_2$, which receives contributions from both sectors. Adding these contributions overcounts the probability of producing a gluon that lies below the cutoff of both sectors and so must be subtracted using the overlap prescription (113), given by taking the wrong limit of the matrix elements in each sector. These are given by Eq. (61) with the rapidity regulator set to unity,

$$\mathcal{M}_{(0,0)}^{n\to\bar{n}} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} \delta(p_1^- - q^-) \delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T) \delta(k^2) \mathrm{Tr}\left[\frac{\not{p}_1}{2} \left(\frac{n^\alpha}{-k^+} + \frac{\bar{n}^\alpha}{k^-} \right) \right.$$

$$\times \left. \frac{\not{\bar{n}}}{2} \left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-} \right) \right] \mathrm{Tr}\left[\frac{\not{p}_2}{2} \frac{\not{n}}{2} \right] + \cdots, \quad (115)$$

and the corresponding (and identical) wrong limit $\mathcal{M}_{(0,0)}^{\bar{n}\to n}$ of the $\bar{n}$ matrix element. The dots indicate terms suppressed by powers of $1/q_L$ relative to leading power, which do not contribute at LP but which will be important at NLP. By integrating these graphs with respect to $\bar{z}_1$ and $\bar{z}_2$ before integrating over the gluon momentum the contributions to the endpoint constant $A_{(0,0)}$ from each sector and their wrong limit subtractions can be obtained. As discussed in [57], because the individual graphs each have rapidity divergences, the ordering of integration is important; the sum is defined here by performing the $\bar{z}_1$, $\bar{z}_2$, $\mathbf{k}_T$, and k^+

integrals, leaving only a single rapidity-divergent k^- integral[5]

$$A_{(0,0)} = \int_0^\infty \frac{dk^-}{k^-} \left[\theta(p_1^- - k^-) A_{(0,0)}^n(k^-) \right.$$

$$+ \theta\left(k^- - \frac{q_T^2}{p_2^+} \right) A_{(0,0)}^{\bar{n}}(k^-)$$

$$\left. - \frac{1}{2} \left(A_{(0,0)}^{n\to\bar{n}}(k^-) + A_{(0,0)}^{\bar{n}\to n}(k^-) \right) \right], \quad (116)$$

where

$$A_{(0,0)}^n(k^-) = 2 - 2\left(\frac{k^-}{p_1^-} \right) + (1 - \epsilon) \left(\frac{k^-}{p_1^-} \right)^2,$$

$$A_{(0,0)}^{\bar{n}}(k^-) = 2 - 2\left(\frac{q_T^2}{k^- p_2^+} \right) + (1 - \epsilon) \left(\frac{q_T^2}{k^- p_2^+} \right)^2,$$

$$A_{(0,0)}^{n\to\bar{n}}(k^-) = A_{(0,0)}^{\bar{n}\to n}(k^-) = 2. \quad (117)$$

Physically, regions of phase space where $k^- \sim O(q^-)$ are properly described in the EFT by n-sector gluons. Regions where $k^+ = k_T^2/k^- \sim O(q^+)$ give spurious contributions in the n sector, producing the unphysical divergence in $A_{(0,0)}^n(k^-)$ as $k^- \to 0$. Similarly, the divergence in the $\bar{n}$ sector as $k^- \to \infty$ corresponds to the large k^- region which is not properly described by the $\bar{n}$ sector. Both of these spurious divergent contributions are cancelled by the overlap terms, leaving the finite result

$$A_{(0,0)} = 2 \log \frac{\hat{s}}{q_T^2} - 3 - \epsilon. \quad (118)$$

This is the same endpoint constant we determined from QCD in (100), and so we find the same LP coefficient function $C_{ff}^{(0)}$. Equivalently, in Eq. (117), the constant terms in $A_n^{(0,0)}$ and $A_{\bar{n}}^{(0,0)}$ are common to both sectors, and so the double-counting is removed by subtracting the overlap on the third line.

As discussed in [57], however, the $\hat{s}$ dependence in $A_{(0,0)}$ is actually a scheme dependence in the EFT, which allows rapidity divergences to be resummed in SCET. Since each integral represents the momentum of a distinct particle in each sector, the momentum in each integral can be independently rescaled, which changes the term in the rapidity logarithm. For example, rescaling $k^- \to k^- \zeta^2/\hat{s}$ in the $A_{\bar{n}}$ integral of Eq. (117) gives the manifestly scheme-dependent result

---

[5]This is equivalent to the prescription in [72] of adding the integrands together before performing any loop integrals.

$$A_{(0,0)}(\zeta) = 2\log\frac{\zeta^2}{q_T^2} - 3 - \epsilon. \qquad (119)$$

In the remainder of this section we will demonstrate a similar origin of rapidity logarithms at NLP.

B. Overlap subtraction and scheme dependence at NLP

At NLP the overlap subtraction follows the same procedure as at LP, but here more terms are kept in the wrong limit expansion of each operator's matrix elements. The NLP cancellation of divergences is also slightly more involved, since rapidity divergences cancel between different operators, as may be seen in Eqs. (79)–(81). Similar cancellations between different operators in SCET were also discussed in detail in [43]. There, the endpoint divergences are regulated by explicit hard cutoffs and expressed in a refactorized form that makes obvious the cancellation between different NLP operators contributing to the observable. Overcounting of hard regions arises from the convolutional structure of the operators with a hard cutoff, and thus an "infinity-bin" prescription, distinct from the usual zero-bin prescription, is introduced to correct for this double counting. In this section we will show that the same overlap subtraction required to correct for overcounting in the soft region also properly regulates endpoint divergences. This uniform treatment of divergences is possible because all spurious terms have a common origin, arising from an overcounting of probabilities induced by wrong limit contributions in each individual sector.

The operator products $T_{(2_1,0)}$ through $T_{(0,2_2)}$ come from products of scattering operators $O_{2,(1_i)}^{\dagger\mu}O_{2,(1_j)}^{\nu}$ whose definitions pick out the longitudinal Lorentz structure $\bar{n}^\mu n^\nu$ or $\bar{n}^\nu n^\mu$, while the remaining operators $T_{(2_3,0)}$ through $T_{(0,2_4)}$, along with the leading order $T_{(0,0)}$, come from products of operators that are proportional to $g_\perp^{\mu\nu}$. It is therefore convenient to classify each $T_{(i,j)}$ according to its Lorentz structure, as either transverse or longitudinal. We consider these two classes of operators in turn.

1. Longitudinal class

From Eqs. (79) and (80), matrix elements of $T_{(2_1,0)}$ and $T_{(0,2_1)}$ are individually rapidity divergent, but the divergences cancel in the sum (and hence in the cross section, since their Wilson coefficients are equal). The same is true for $T_{(2_2,0)}$ and $T_{(0,2_2)}$, and in both cases the cancellation may be understood by examining the unregulated diagrams and corresponding overlaps, as in the previous section.

Taking $T_{(2_1,0)}$ as an example, its unregulated spin-averaged matrix element is

$$\mathcal{M}_{(2_1,0)}^n = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d} \delta(p_1^- - q^- - k^-)\delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)\delta(k^2)\delta\left(u + \frac{k^-}{q^-}\right)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\left(\frac{2p_1^\alpha - \gamma^\alpha\not{k}}{-2p_1\cdot k} + \frac{\bar{n}^\alpha}{k^-}\right)\frac{\not{\bar{n}}}{2}\not{k}_\perp\gamma_\mu^\perp\bar{\Delta}^{\alpha\mu}(k)\right]$$

$$= -\frac{\bar{\alpha}}{\pi}\delta_2\int_0^\infty \frac{dk^-}{k^-}p_1^-\delta(p_1^- - q^- - k^-)\delta\left(u + \frac{k^-}{q^-}\right)$$

$$= -\frac{\bar{\alpha}\,\delta_2}{\pi\,\bar{z}_1}\delta\left(u + \frac{\bar{z}_1}{z_1}\right), \qquad (120)$$

where $\bar{\Delta}$ is defined in Appendix A and its wrong-sector limit is

$$\mathcal{M}_{(2_1,0)}^{n\to\bar{n}} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d}\delta(p_1^- - q^-)\delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)\delta(k^2)\delta\left(u + \frac{k^-}{q^-}\right)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\left(\frac{n^\alpha}{-k^+} + \frac{\bar{n}^\alpha}{k^-}\right)\frac{\not{\bar{n}}}{2}\not{k}_\perp\gamma_\mu^\perp\bar{\Delta}^{\alpha\mu}(k)\right]$$

$$= -\frac{\bar{\alpha}}{\pi}\delta(\bar{z}_1)\delta(\bar{z}_2)\int_0^\infty \frac{dk^-}{k^-}\delta\left(u + \frac{k^-}{q^-}\right). \qquad (121)$$

Away from $\bar{z}_1 = \bar{z}_2 = 0$ the overlap does not contribute and (120) gives a well-defined result; however, it is rapidity divergent at $\bar{z}_1 = 0$. Following the LP approach, the matrix element may be written in the general form

$$\mathcal{M}_{(2_1,0)} = \frac{\bar{\alpha}}{\pi}\delta_2\left(A_{(2_1,0)}\delta_1\delta(u) - \left[\frac{1}{\bar{z}_1}\right]_+\delta\left(u + \frac{\bar{z}_1}{z_1}\right)\right), \quad (122)$$

in accordance with Eq. (86). The constant $A_{(2_1,0)}$ is determined by integrating with respect to u, $\bar{z}_1$, and $\bar{z}_2$, which gives

$$A_{(2_1,0)} = \int_0^\infty \frac{dk^-}{k^-}\left[\theta(p_1^- - k^-)A_{(2_1,0)}^n(k^-)\right.$$

$$\left. - \frac{1}{2}A_{(2_1,0)}^{n\to\bar{n}}(k^-)\right], \qquad (123)$$

where

$$A_{(2_1,0)}^n(k^-) = A_{(2_1,0)}^{n\to\bar{n}}(k^-) = -1. \qquad (124)$$

The integral in Eq. (123) is divergent: matrix elements of $T_{(2_1,0)}$ alone are not rapidity finite, in agreement with the result (79) using the pure rapidity regular. This is to be expected, since gluons in both the n and $\bar{n}$ sectors are required to reproduce the QCD rate, and the corresponding $\bar{n}$-sector gluon is emitted from the operator $T_{(0,2_1)}$.

Including this operator and its corresponding subtraction gives

$$
\begin{aligned}
\mathcal{M}_{(2_1,0)} &+ \mathcal{M}_{(0,2_1)} \\
&= \frac{\bar{\alpha}}{\pi} \left\{ A_{2_1} \delta_1 \delta_2 \delta(u) - \delta_2 \left[\frac{1}{\bar{z}_1} \right]_+ \delta\left(u + \frac{\bar{z}_1}{z_1} \right) \right. \\
&\left. - \delta_1 \left[\frac{1}{\bar{z}_2} \right]_+ \delta\left(u + \frac{\bar{z}_2}{z_2} \right) \right\},
\end{aligned}
\tag{125}
$$

where

$$
\begin{aligned}
A_{2_1} = \int_0^\infty \frac{dk^-}{k^-} & \left[\theta(p_1^- - k^-) A_{(2_1,0)}^n(k^-) \right. \\
& + \theta\left(k^- - \frac{q_T^2}{p_2^+} \right) A_{(0,2_1)}^{\bar{n}}(k^-) \\
& \left. - \frac{1}{2} (A_{(2_1,0)}^{n\to\bar{n}}(k^-) + A_{(0,2_1)}^{\bar{n}\to n}(k^-)) \right]
\end{aligned}
\tag{126}
$$

and

$$
A_{(0,2_1)}^{\bar{n}}(k^-) = A_{(0,2_1)}^{\bar{n}\to n}(k^-) = -1.
\tag{127}
$$

The integral in Eq. (126) is finite; as at LP, the spurious divergences from the n sector as $k^- \to 0$ and the $\bar{n}$ sector as $k^- \to \infty$ have been canceled by the overlap subtraction to give the finite result

$$
A_{2_1} = -\log\frac{\hat{s}}{q_T^2},
\tag{128}
$$

which, by a similar rescaling argument as at leading power, gives a scheme-dependent rapidity logarithm reproducing that in Eq. (86).

A similar argument holds for $T_{(2_2,0)}$ and $T_{(0,2_2)}$. Explicitly, we find

$$
\begin{aligned}
\mathcal{M}_{(2_2,0)} + \mathcal{M}_{(0,2_2)} &= \frac{\bar{\alpha}}{\pi} \left\{ A_{2_2} \delta_1 \delta_2 \delta(u_1) \right. \\
&+ \delta_2 \left[\frac{1}{\bar{z}_1} \right]_+ \delta\left(u_1 + \frac{\bar{z}_1}{z_1} \right) \\
&\left. + \delta_1 \left[\frac{1}{\bar{z}_2} \right]_+ \delta\left(u_1 + \frac{\bar{z}_2}{z_2} \right) \right\} \delta(u_1 - u_2),
\end{aligned}
\tag{129}
$$

where

$$
\begin{aligned}
A_{(2_2,0)}^n(k^-) &= A_{(2_2,0)}^{n\to\bar{n}}(k^-) = A_{(0,2_2)}^{\bar{n}}(k^-) \\
&= A_{(0,2_2)}^{\bar{n}\to n}(k^-) = 1,
\end{aligned}
\tag{130}
$$

and so

$$
A_{2_2} = \log\frac{\hat{s}}{q_T^2},
\tag{131}
$$

again in agreement with Eq. (86). The total fixed-order contribution to the cross section therefore cancels between the four longitudinal operators.

2. Transverse class

Matrix elements of the transverse class of operators $T_{(2_3,0)}$ through $T_{(0,2_4)}$ are more complicated because they originate from operator products having the same Lorentz structure as those that produce the leading-power operator $T_{(0,0)}$, and power corrections to the overlap subtraction of $T_{(0,0)}$ must also be included to achieve a rapidity-finite combination. Thus, while in the longitudinal case rapidity divergences canceled between the corresponding n- and $\bar{n}$- sector operators, here they only cancel in the particular linear combination of transverse operators that contribute to the DY cross section.

The contribution of the transverse operators to the coefficient function $C_{ff}^{(2)}$ is calculated from Eq. (12) and has the general form

$$
\begin{aligned}
C_{ff}^{(2)T} &= \bar{\alpha} z_1 z_2 (A_T^{(2)} \delta_1 \delta_2 + B_T^{(2)} \delta_1' \delta_2' + C_T^{(2)} (\delta_1' \delta_2 + \delta_1 \delta_2')) \\
&+ \delta_2' [g_n^T(\bar{z}_1)]_{++} + \delta_2 [h_n^T(\bar{z}_1)]_{++} \\
&+ \delta_1' [g_{\bar{n}}^T(\bar{z}_2)]_{++} + \delta_1 [h_{\bar{n}}^T(\bar{z}_2)]_{++}).
\end{aligned}
\tag{132}
$$

Away from the endpoint $\bar{z}_1 = \bar{z}_2 = 0$ there are no rapidity divergences, so the contribution from each operator to $g_{n,\bar{n}}^T$ and $h_{n,\bar{n}}^T$ are the same as in Eqs. (86) and (87). After summing and integrating over u's, these combine to give the functions $g_{n,\bar{n}}^T$, $h_{n,\bar{n}}^T$:

$$
g_{n,\bar{n}}^T(\bar{z}) = -\left(\frac{2 - 2\bar{z} + \bar{z}^2}{\bar{z}^2} \right),
$$

$$
h_{n,\bar{n}}^T(\bar{z}) = -\frac{2}{\bar{z}^2}.
\tag{133}
$$

The endpoint region is overcounted in the sum of the two sectors and must be compensated by subtracting away half the wrong limit of each sector. In contrast with the previous cases, the power counting of the required overlap subtractions is more subtle because the overlap graphs must subtract not only logarithmic, but also linear rapidity divergences.

Consider first the various contributions to $A_T^{(2)}$, which are found by integrating unweighted matrix elements over $\{u\}$, $\bar{z}_1$, and $\bar{z}_2$. The naïve contributions from $T_{(2_3,0)}$ and $T_{(0,2_3)}$ are calculated to be

$$
\int_0^\infty \frac{dk^-}{k^-} \left[\theta(p_1^- - k^-) A_{(2_3,0)}^n(k^-) + \theta\left(k^- - \frac{q_T^2}{p_2^+} \right) A_{(0,2_3)}^{\bar{n}}(k^-) \right],
\tag{134}
$$

where

$$A^n_{(2_3,0)}(k^-) = -2\left(\frac{p_1^-}{k^-}\right),$$

$$A^{\bar{n}}_{(0,2_3)}(k^-) = -2\left(\frac{k^- p_2^+}{q_T^2}\right). \tag{135}$$

The integral in (134) is rapidity divergent. Both terms in Eq. (135) are $O(1)$ in their respective correct regions, $k^- \sim O(q^-)$ in the n sector and $k^+ = q_T^2/k^- \sim O(q^+)$ in the $\bar{n}$ sector, but are enhanced and give rise to linear rapidity divergences in the regions where this correct momentum scaling is no longer valid. The contributions to C_{ff} from these spurious regions are subtracted away by the overlap. There are two sources of overlap subtraction for $A_T^{(2)}$: the wrong limits $A^{n\to\bar{n}}_{(2_3,0)}$ and $A^{\bar{n}\to n}_{(0,2_3)}$, and also the subleading wrong limits from the leading power operator $T_{(0,0)}$.

Expanding the Feynman diagrams of $T_{(2_3,0)}$ and $T_{(0,2_3)}$ in their wrong limits gives the same functions as in Eq. (135) but are integrated over the region $0 < k^- < \infty$. Explicitly, there are two nonvanishing terms from the wrong limit of the $T_{(2_3,0)}$ matrix element,

$$\mathcal{M}^{n\to\bar{n},\mathrm{I}}_{(2_3,0)} = g^2 C_F \int \frac{d^d k}{(2\pi)^d} \delta(p_1^- - q^-)\delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)2\pi\delta(k^2)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\frac{1}{u}k^\mu\bar{\Delta}(k)^{\alpha\nu}\frac{\not{\bar{n}}}{2}\gamma_\nu^\perp\gamma_\mu^\perp\right.$$

$$\times \left.\left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-}\right)\delta\left(u + \frac{k^-}{q^-}\right)\right]$$

$$= \frac{\bar{\alpha}}{\pi}\delta_1\delta_2 \int_0^\infty \frac{dk^-}{k^-}\frac{1}{u}\delta\left(u + \frac{k^-}{q^-}\right) \tag{136}$$

and

$$\mathcal{M}^{n\to\bar{n},\mathrm{II}}_{(2_3,0)} = g^2 C_F \int \frac{d^d k}{(2\pi)^d}(-k^-)\delta'(p_1^- - q^-)\delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)2\pi\delta(k^2)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\frac{1}{u}k^\mu\bar{\Delta}(k)^{\alpha\nu}\frac{\not{\bar{n}}}{2}\gamma_\nu^\perp\gamma_\mu^\perp\right.$$

$$\times \left.\left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-}\right)\delta\left(u + \frac{k^-}{q^-}\right)\right]$$

$$= -\frac{\bar{\alpha}}{\pi}\delta'(\bar{z}_1)\delta(\bar{z}_2)\int_0^\infty \frac{dk^-}{p_1^-}\frac{1}{u}\delta\left(u + \frac{k^-}{q^-}\right). \tag{137}$$

The wrong-limit expansion is truncated after the terms reach an $O(q_T^2/q_L^2)$ suppression relative to the leading-

power operator in the wrong-limit momentum scaling $p_1^- \sim O(q^-)$ and $k^+, p_2^+ \sim O(q^+)$. The subtraction term in Eq. (136) contributes to $A_T^{(2)}$ while the term in Eq. (137) contributes to $C_T^{(2)}$.

Similarly, expanding the $\bar{n}$- and n-sector graphs of $T_{(0,0)}$ up to NLP gives the $O(1/k^-)$ term in $A^n_{(2_3,0)}$ and the $O(k^-)$ term in $A^{\bar{n}}_{(0,2_3)}$ in Eq. (135), respectively, which are again integrated over all values of k^-. Explicitly, expanding $\mathcal{M}^n_{(0,0)}$ gives two contributions to the subleading overlap,

$$\mathcal{M}^{n\to\bar{n},\mathrm{NLP}_1}_{(0,0)} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d}(-k^-)\delta'(p_1^- - q^-)$$

$$\times \delta(p_2^+ - q^+)\delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)\delta(k^2)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\left(\frac{n^\alpha}{-k^+} + \frac{\bar{n}^\alpha}{k^-}\right)\frac{\not{\bar{n}}}{2}\left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-}\right)\right]$$

$$= -2\delta'(\bar{z}_1)\delta(\bar{z}_2)\frac{\bar{\alpha}}{\pi}\int_0^\infty \frac{dk^-}{k^-}\frac{k^- p_2^+}{q_T^2} \tag{138}$$

and

$$\mathcal{M}^{n\to\bar{n},\mathrm{NLP}_2}_{(0,0)} = -2\pi g^2 C_F \int \frac{d^d k}{(2\pi)^d}\delta(p_1^- - q^-)\delta(p_2^+ - q^+)$$

$$\times \delta^{d-2}(\mathbf{q}_T + \mathbf{k}_T)2\pi\delta(k^2)\mathrm{Tr}\left[\frac{\not{p}_2}{2}\frac{\not{n}}{2}\right]$$

$$\times \mathrm{Tr}\left[\frac{\not{p}_1}{2}\left(\frac{n^\alpha}{-k^+} + \frac{\bar{n}^\alpha}{k^-}\right)\frac{\not{\bar{n}}}{2}\left(\frac{\not{k}_\perp\gamma_\perp^\mu}{p_1^- k^+}\Delta_{\alpha\mu}(k)\right)\right.$$

$$\left.+ \frac{\not{p}_1}{2}\left(\Delta^{\alpha\mu}(k)\frac{\gamma_\mu^\perp\not{k}_\perp}{p_1^- k^+}\right)\frac{\not{\bar{n}}}{2}\left(\frac{n_\alpha}{-k^+} - \frac{\bar{n}_\alpha}{-k^-}\right)\right]$$

$$= -2\frac{\bar{\alpha}}{\pi}\delta(\bar{z}_1)\delta(\bar{z}_2)\int_0^\infty \frac{dk^-}{k^-}\frac{k^- p_2^+}{q_T^2}, \tag{139}$$

the first of which comes from higher corrections to the momentum-conserving delta function, while the second comes from higher corrections to the quark propagator expansions. Only the second term contributes here; the first contributes to $C_T^{(2)}$.

Putting these together, we obtain the expression for $A_T^{(2)}$,

$$A_T^{(2)} = \int_0^\infty \frac{dk^-}{k^-}\left[\theta(p_1^- - k^-)A^n_{(2_3,0)}(k^-)\right.$$

$$+ \theta\left(k^- - \frac{q_T^2}{p_2^+}\right)A^{\bar{n}}_{(0,2_3)}(k^-)$$

$$- \frac{1}{2}(A^{n\to\bar{n},\mathrm{NLP}}_{(0,0)}(k^-) + A^{\bar{n}\to n,\mathrm{NLP}}_{(0,0)}(k^-))$$

$$\left.- \frac{1}{2}(A^{n\to\bar{n}}_{(2_3,0)}(k^-) + A^{\bar{n}\to n}_{(0,2_3)}(k^-))\right], \tag{140}$$

where the contributions from $T_{(2_4,0)}$ and $T_{(0,2_4)}$ all vanish, and explicitly

$$A_{(2_3,0)}^{n\to\bar{n}}(k^-) = A_{(0,0)}^{\bar{n}\to n,\text{NLP}}(k^-) = -2\left(\frac{p_1^-}{k^-}\right),$$

$$A_{(0,2_3)}^{\bar{n}\to n}(k^-) = A_{(0,0)}^{n\to\bar{n},\text{NLP}}(k^-) = -2\left(\frac{k^- p_2^+}{q_T^2}\right). \tag{141}$$

This gives the finite result

$$A_T^{(2)} = 4. \tag{142}$$

Next consider the contributions to the endpoint constant $C_T^{(2)}$, which are obtained by integrating the various matrix elements weighted with $\bar{z}_1$ (or equivalently $\bar{z}_2$). The naïve contributions to the $\bar{z}_1$ moment give

$$\int_0^\infty \frac{dk^-}{k^-}\left[\theta(p_1^- - k^-)C_{(2_3,0)}^n(k^-)\right.$$
$$\left. + \theta\left(k^- - \frac{q_T^2}{p_2^+}\right)C_{(0,2_4)}^{\bar{n}}(k^-)\right], \tag{143}$$

where

$$C_{(2_3,0)}^n(k^-) = 2,$$

$$C_{(0,2_4)}^{\bar{n}}(k^-) = -2\left(\frac{k^- p_2^+}{q_T^2}\right) + 2 - \left(\frac{q_T^2}{k^- p_2^+}\right). \tag{144}$$

This is again rapidity divergent: $C_{(2_3,0)}^n(k^-)$ gives a logarithmically divergent contribution as $k^- \to 0$, while $C_{(0,2_4)}^{\bar{n}}(k^-)$ gives contributions that are both logarithmically and linearly divergent as $k^- \to \infty$. As with $A_T^{(2)}$, taking the wrong limit of the Feynman diagrams contributing to Eq. (144) gives the $k^- \to 0$ and $k^- \to \infty$ expansions of these terms. For example, the wrong-limit expansion of $\mathcal{M}_{(0,2_4)}^{\bar{n}}$ gives three terms which correspond almost exactly to the overlaps of $T_{(0,0)}$ in Eqs. (115), (138), and (139), except they have a different momentum-conserving delta function structure. These give the contributions

$$C_{(0,2_4)}^{\bar{n}\to n}(k^-) = -2\left(\frac{k^- p_2^+}{q_T^2}\right) + 2, \tag{145}$$

while from the $T_{(2_3,0)}$ overlap given in Eq. (137) is the contribution

$$C_{(2_3,0)}^{n\to\bar{n}}(k^-) = 2. \tag{146}$$

The overlap term from $T_{(0,2_4)}$ contains two terms: the leading term is proportional to k^- and cancels a linear rapidity divergence, while the $O(1)$ term contributes to the cancellation of a logarithmic divergence.

We also have the contribution from the NLP overlap of $T_{(0,0)}$ in Eq. (138),

$$C_{(0,0)}^{n\to\bar{n},\text{NLP}}(k^-) = -2\left(\frac{k^- p_2^+}{q_T^2}\right), \tag{147}$$

with the sum of all contributions giving the result

$$C_T^{(2)} = \int_0^\infty \frac{dk^-}{k^-}\left[\theta(p_1^- - k^-)C_{(2_3,0)}^n(k^-)\right.$$
$$+ \theta\left(k^- - \frac{q_T^2}{p_2^+}\right)C_{(0,2_4)}^{\bar{n}}(k^-) - \frac{1}{2}(C_{(0,0)}^{n\to\bar{n},\text{NLP}})(k^-))$$
$$\left. - \frac{1}{2}(C_{(2_3,0)}^{n\to\bar{n}}(k^-) + C_{(0,2_4)}^{\bar{n}\to n}(k^-))\right]$$
$$= 2\log\frac{\hat{s}}{q_T^2} + 1. \tag{148}$$

Once again there is a precise interplay between naïve matrix elements and overlap subtractions required to obtain the same finite result as using the pure rapidity regulator. Rescaling the integrals for $C_{(0,2_4)}^{\bar{n}}(k^-)$, $C_{(0,2_4)}^{\bar{n}\to n}(k^-)$, and $C_{(0,0)}^{n\to\bar{n},\text{NLP}}(k^-)$ as $k^- \to k^-\zeta^2/\hat{s}$ replaces the $\hat{s}$ dependence in the result of Eq. (148) with ζ^2 scheme dependence. This correlation between the rescaling of individual integrals is necessary to maintain a finite result and is a general feature of power-law divergences.

Finally, the endpoint constant B_T is found by weighting the integrals by $\bar{z}_1\bar{z}_2$, giving

$$B_T^{(2)} = \int_0^\infty \frac{dk^-}{k^-}\left[\theta(p_1^- - k^-)B_{(2_4,0)}^n(k^-)\right.$$
$$+ \theta\left(k^- - \frac{q_T^2}{p_2^+}\right)B_{(0,2_4)}^{\bar{n}}(k^-)$$
$$\left. - \frac{1}{2}(B_{(2_4,0)}^{n\to\bar{n}}(k^-) + B_{(0,2_4)}^{\bar{n}\to n}(k^-))\right], \tag{149}$$

where explicitly

$$B_{(2_4,0)}^n(k^-) = 2 - 2\left(\frac{k^-}{p_1^-}\right) + \left(\frac{k^-}{p_1^-}\right)^2,$$

$$B_{(0,2_4)}^{\bar{n}}(k^-) = 2 - 2\left(\frac{q_T^2}{k^- p_2^+}\right) + \left(\frac{q_T^2}{k^- p_2^+}\right)^2, \tag{150}$$

and the overlap terms cancel just the logarithmic divergences,

$$B_{(2_4,0)}^{n\to\bar{n}}(k^-) = B_{(0,2_4)}^{\bar{n}\to n}(k^-) = 2. \tag{151}$$

This gives

$$B_T^{(2)} = 2\log\frac{\hat{s}}{q_T^2} - 3, \qquad (152)$$

where rescaling the second line of Eq. (150) as $k^- \to k^- \zeta^2/\hat{s}$ replaces the $\hat{s}$ dependence in Eq. (152) with ζ^2 scheme dependence.

This concludes the calculation of all the endpoint constants $A_{L,T}$, $B_{L,T}$, and $C_{L,T}$. In each case, these constants agree with those of QCD as calculated in Sec. II E. We have thus demonstrated an overlap subtraction prescription that allows us to properly calculate probabilities at NLP without an explicit rapidity regulator, providing a nontrivial cross-check of our results using different rapidity regularization schemes.

IV. CONCLUSION

In this paper we have shown that factorization of the Drell-Yan production cross section into hard matching coefficients, rapidity evolution factors, soft matching coefficients, and PDFs occurs naturally in a formulation of SCET in which the low energy degrees of freedom are not separated into distinct fields for each mode relevant to the process. The DY rate is given by the matrix element of the nonlocal product of two external currents in SCET. Usually in SCET observables are factorized into jet and soft factors which are separately renormalized and run to the appropriate scales; here, the EFT is first run in μ down to the soft matching scale $\mu \sim q_T$, at which point the product of currents is renormalized in rapidity space. After resumming the rapidity logs at the soft matching scale, the operator products are then matched onto a product of light-cone distribution operators, whose hadronic matrix elements are the usual PDFs. At $O(\alpha_s)$, our EFT cross section reproduces the fixed-order QCD cross section at NLP, as well as the equivalent fixed-order cross section calculated using the pure rapidity regulator in [53]. Off-diagonal rapidity anomalous dimensions were calculated and rapidity divergences were shown to cancel in the cross section. The resummation of rapidity logarithms at NLP requires the complete rapidity anomalous dimension matrix for the subleading operators $T_{(i,j)}$, which is beyond the scope of this paper, and will be the subject of future work.

The factorization and resummation of the DY process is particularly simple in this approach: it does not depend on proving factorization at a given order in the SCET expansion or in the leading-log approximation, but instead is a straightforward consequence of the usual EFT approach of matching and running. By not explicitly factorizing modes in the Lagrangian, the complication of power corrections coupling different modes in the Lagrangian is avoided, as is the necessity to refactorize

the result to make individual jet and soft functions well-defined. Divergences analogous to the endpoint divergences arising at NLP in other approaches arise, but are regulated by the rapidity regulator and systematically canceled by the same overlap subtraction procedure required to avoid double counting at leading power.

Rapidity divergences were considered in detail, and the cancellation of rapidity divergences in the rate was shown in two ways. Using the pure rapidity regulator, it was shown that all rapidity poles canceled between the different linear combinations of subleading operators arising in the expression for the differential rate, as was found in previous analyses [40–43]. In Sec. III it was shown that even without an explicit rapidity regulator, rapidity divergences in the DY cross section cancel between particular linear combinations of operators, and that these linear combinations could be understood by requiring that SCET reproduce the correct differential rate in different regions of phase space. A consistent treatment of subleading overlap subtractions from the leading order operator was shown to be necessary for this cancellation.

APPENDIX A: SUMMARY OF MATRIX ELEMENTS OF HARD SCATTERING OPERATORS

In the following equations we list all relevant u-space matrix elements of scattering operators which contribute to the quark-induced DY process through the emission of an n-sector gluon. We use the soft-scale matching kinematics $p_1^+ = p_{1\perp} = 0 = p_2^- = p_{2\perp}$, and define the noncommon factor $\mathcal{A}_{n,\bar{n}}^{(i)}$ of these matrix elements through the relation

$$\int \frac{d^d x}{2(2\pi)^d} e^{-iq\cdot x} \langle k^{n,\bar{n}} | O_2^{(i)\mu}(x, \{u\}) | p_1^n p_2^{\bar{n}} \rangle$$
$$\equiv g\bar{v}(p_2^{\bar{n}}) T^a \mathcal{A}_{n,\bar{n}}^{(i)} u(p_1^n) \epsilon_\nu^*, \qquad (A1)$$

where $\mathcal{A}$ is tensor valued with implied Lorentz indices μ and ν. We find for the n-gluon emissions

$$\mathcal{A}_n^{(0)} = -P_n\gamma^\mu P_n\left(\frac{2p_1^\nu - \slashed{k}\gamma^\nu}{-2p_1\cdot k} - \frac{\bar{n}^\nu}{-k^-}\right)\delta_n^-\delta_n^+\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_n^{(1_{\perp n})} = -P_n\gamma^\mu\frac{\slashed{\bar{n}}}{2}\slashed{k}_\perp\left(\frac{2p_1^\nu - \slashed{k}\gamma^\nu}{-2p_1\cdot k} - \frac{\bar{n}^\nu}{-k^-}\right)\delta_n^-\delta_n^+\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_n^{(1A_1)} = \gamma_\alpha^\perp\frac{\slashed{\bar{n}}}{2}\gamma^\mu P_n\bar{\Delta}^{\nu\alpha}(k)\delta_n^-\delta_n^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^-),$$

$$\mathcal{A}_n^{(1A_2)} = -P_n\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp\bar{\Delta}^{\nu\alpha}(k)\delta_n^-\delta_n^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^-),$$

$$\mathcal{A}_n^{(2\delta^+)} = q^+q^-P_n\gamma^\mu P_n\left(\frac{2p_1^\nu - \slashed{k}\gamma^\nu}{-2p_1\cdot k} - \frac{\bar{n}^\nu}{-k^-}\right)k^+\delta_n^-\delta_n^{+\prime}\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_n^{(2A_1)} = -\frac{1}{u}\gamma_\alpha^\perp\gamma_\beta^\perp P_n\gamma^\mu P_n\bar{\Delta}^{\nu\alpha}(k)k^\beta\delta_n^-\delta_n^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^-),$$
$$(A2)$$

while for the $\bar{n}$-gluon emissions we find

$$\mathcal{A}_{\bar{n}}^{(0)} = \left(\frac{2p_2^\nu - \gamma^\nu\slashed{k}}{-2p_2\cdot k} - \frac{n^\nu}{-k^+}\right)P_n\gamma^\mu P_n\delta_{\bar{n}}^-\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_{\bar{n}}^{(1_{\perp\bar{n}})} = \left(\frac{2p_2^\nu - \gamma^\nu\slashed{k}}{-2p_2\cdot k} - \frac{n^\nu}{-k^+}\right)\slashed{k}_\perp\frac{\slashed{n}}{2}\gamma^\mu P_n\delta_{\bar{n}}^-\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_{\bar{n}}^{(1B_1)} = -P_n\gamma^\mu\frac{\slashed{n}}{2}\gamma_\alpha^\perp\Delta^{\nu\alpha}(k)\delta_{\bar{n}}^-\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^+),$$

$$\mathcal{A}_{\bar{n}}^{(1B_2)} = \gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu P_n\Delta^{\nu\alpha}(k)\delta_{\bar{n}}^-\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^+),$$

$$\mathcal{A}_{\bar{n}}^{(2\delta^-)} = q^+q^-\left(\frac{n^\nu}{-k^+} - \frac{2p_2^\nu - \gamma^\nu\slashed{k}}{-2p_2\cdot k}\right)P_n\gamma^\mu P_n k^-\delta_{\bar{n}}^{-\prime}\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp,$$

$$\mathcal{A}_{\bar{n}}^{(2B_1)} = \frac{1}{u}P_n\gamma^\mu P_n\gamma_\alpha^\perp\gamma_\beta^\perp k^\alpha\bar{\Delta}^{\nu\beta}(k)\delta_{\bar{n}}^-\delta_{\bar{n}}^+\boldsymbol{\delta}_\perp\delta(u+\hat{k}^+). \quad (A3)$$

The one-gluon matrix elements of the scattering operators defined in Eqs. (A2) and (A3) use the following definitions:

$$\bar{\Delta}^{\alpha\mu}(k) = g^{\alpha\mu} - \frac{\bar{n}^\alpha k^\mu}{\bar{n}\cdot k}, \qquad \Delta^{\alpha\mu}(k) = g^{\alpha\mu} - \frac{n^\alpha k^\mu}{n\cdot k}. \quad (A4)$$

These are common structures associated with the covariant derivative. We also define the dimensionless quantities

$$\hat{\ell}^- = \frac{\ell^-}{q^-}, \qquad \hat{\ell}^+ = \frac{\ell^+}{q^+}, \quad (A5)$$

and we have used the shorthand notation $\delta_n^- = \delta(p_1^- - k^- - q^-)$, $\delta_n^+ = \delta(p_2^+ - q^+)$, $\delta_{\bar{n}}^- = \delta(p_1^- - q^-)$, $\delta_{\bar{n}}^+ = \delta(p_2^+ - k^+ - q^+)$, and $\boldsymbol{\delta}_\perp = \delta^{(d-2)}(k_\perp + q_\perp)$.

There are additional operators that are present from the hard-scale matching [56,65,85–87], but that do not contribute to the quark-initiated DY process to the order at which we are working. Up to a $1/q_L^2$ suppression, these include an operator with two perpendicular derivatives

$$O_2^{(2\perp\perp)\mu}(x) = [i\partial^\alpha\bar{\chi}_{\bar{n}}(x_{\bar{n}})]\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\beta^\perp[i\partial^\beta\chi_n(x_n)], \quad (A6)$$

the A-type operators

$$O_2^{(2A_2)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[\bar{\chi}_{\bar{n}}(x)]\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\beta^\perp$$
$$\times[\mathcal{B}_n^{\dagger\alpha\beta}(x)\chi_n(x-\bar{n}t)],$$

$$O_2^{(2A_3)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[\bar{\chi}_{\bar{n}}(x)]\gamma_\beta^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp$$
$$\times[i\partial^\alpha\mathcal{B}_n^{\dagger\beta}(x)\chi_n(x-\bar{n}t)],$$

$$O_2^{(2A_4)\mu}(x,\hat{t}) = -2\pi i\theta(\hat{t})\otimes[i\partial^\beta\bar{\chi}_{\bar{n}}(x-nt)]\gamma_\beta^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\alpha^\perp$$
$$\times[\mathcal{B}_n^{\dagger\alpha}(x)\chi_n(x)],$$

$$O_2^{(2A_5)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[i\partial^\alpha\bar{\chi}_{\bar{n}}(x)]\gamma^\mu\{\gamma_\alpha^\perp,\gamma_\beta^\perp\}$$
$$\times[\mathcal{B}_n^{\dagger\beta}(x-\bar{n}t)\chi_n(x)], \quad (A7)$$

and the corresponding B-type operators

$$O_2^{(2B_2)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[\bar{\chi}_{\bar{n}}(x-nt)\mathcal{B}_{\bar{n}}^{\alpha\beta}(x)]$$
$$\times\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\beta^\perp[\chi_n(x)],$$

$$O_2^{(2B_3)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[i\partial^\alpha\bar{\chi}_{\bar{n}}(x-nt)\mathcal{B}_{\bar{n}}^\beta(x)]$$
$$\times\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\beta^\perp[\chi_n(x)],$$

$$O_2^{(2B_4)\mu}(x,\hat{t}) = -2\pi i\theta(\hat{t})\otimes[\bar{\chi}_{\bar{n}}(x)\mathcal{B}_{\bar{n}}^\alpha(x-\bar{n}t)]$$
$$\times\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\frac{\slashed{\bar{n}}}{2}\gamma_\beta^\perp[i\partial^\beta\chi_n(x)],$$

$$O_2^{(2B_5)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[\bar{\chi}_{\bar{n}}(x)\mathcal{B}_{\bar{n}}^\alpha(x-\bar{n}t)]$$
$$\times\gamma^\mu\{\gamma_\alpha^\perp,\gamma_\beta^\perp\}[i\partial^\beta\chi_n(x)]. \quad (A8)$$

There are also the C-type operators, which are only relevant for gluon-induced Drell-Yan

$$O_2^{(1C_1)\mu}(x,\hat{t}) = -2\pi i\theta(\hat{t})\otimes[\mathcal{B}_n^{\alpha cc'}(x)]$$
$$\times\left[\bar{\chi}_{\bar{n}}^c(x)\gamma^\mu\frac{\slashed{n}}{2}\gamma_\alpha^\perp\chi_{\bar{n}}^{c'}(x-nt)\right],$$

$$O_2^{(1C_2)\mu}(x,\hat{t}) = 2\pi i\theta(\hat{t})\otimes[\mathcal{B}_n^{\alpha cc'}(x)]$$
$$\times\left[\bar{\chi}_{\bar{n}}^c(x-nt)\gamma_\alpha^\perp\frac{\slashed{n}}{2}\gamma^\mu\chi_{\bar{n}}^{c'}(x)\right]. \quad (A9)$$

APPENDIX B: PLUS DISTRIBUTION IDENTITIES

1. Single variable plus distributions

The familiar plus distribution may be written as

$$[\theta(x)f(x)]_+ = \lim_{\beta \to 0}[F(\beta) - F(1)]\delta(x - \beta)$$
$$+ \theta(x - \beta)f(x), \tag{B1}$$

where $f(x) = dF(x)/dx$, and has the properties

$$\int_0^1 dx[\theta(x)f(x)]_+ = 0,$$
$$[\theta(x)f(x)]_+ = f(x), \qquad x > 0. \tag{B2}$$

Rearranging these equations gives the differential relation

$$\frac{d}{dx}[\theta(x)F(x)] = \left[\theta(x)\frac{dF(x)}{dx}\right]_+ + \delta(x)F(1), \tag{B3}$$

which is useful for expanding rapidity divergent integrals in terms of plus functions. For example, taking $f(x) = x^{-1-\eta}$ gives $F(x) = -x^{-\eta}/\eta$, and so

$$\frac{d}{dx}\left[\frac{-\theta(x)}{\eta x^\eta}\right] = \frac{\theta(x)}{x^{1+\eta}} - \frac{\delta(x)}{\eta x^\eta} = \left[\frac{\theta(x)}{x^{1+\eta}}\right]_+ - \frac{\delta(x)}{\eta}. \tag{B4}$$

The factor of $\delta(x)x^{-\eta}$ vanishes by analytic continuation, so expanding about $\eta = 0$ gives

$$\frac{\theta(x)}{x^{1+\eta}} = -\frac{\delta(x)}{\eta} + \left[\frac{\theta(x)}{x} - \eta\frac{\log(x)\theta(x)}{x} + \cdots\right]_+$$
$$= -\frac{\delta(x)}{\eta} + \left[\frac{\theta(x)}{x}\right]_+ - \eta\left[\frac{\log(x)\theta(x)}{x}\right]_+ + \cdots. \tag{B5}$$

Matrix elements at next-to-leading power involve higher-order poles that are more singular than the usual plus distributions. As in [53], we define double-plus distributions that satisfy

$$\int_0^1 dx[\theta(x)f(x)]_{++} = 0,$$
$$\int_0^1 dxx[\theta(x)f(x)]_{++} = 0,$$
$$[\theta(x)f(x)]_{++} = f(x), \qquad x > 0. \tag{B6}$$

They are related to the single-plus distributions by

$$[\theta(x)f(x)]_{++} - [\theta(x)f(x)]_+$$
$$= \lim_{\beta \to 0}\delta'(x - \beta)\int_\beta^1 dy(y - \beta)f(y). \tag{B7}$$

For example, taking $f(x) = x^{-2-\eta}$, then $F(x) = -x^{-1-\eta}/(1 + \eta)$, and we obtain

$$\frac{\theta(x)}{x^{2+\eta}} = \left[\frac{\theta(x)}{x^{2+\eta}}\right]_+ - \frac{\delta(x)}{1 + \eta}. \tag{B8}$$

Since $\left[\frac{\theta(x)}{x^2}\right]_+$ is not well-defined, we convert to a double-plus distribution before expanding in η,

$$\left[\frac{\theta(x)}{x^{2+\eta}}\right]_+ = \left[\frac{\theta(x)}{x^{2+\eta}}\right]_{++} - \delta'(x)\int_0^1 dx\frac{1}{x^{1+\eta}}$$
$$= \left[\frac{\theta(x)}{x^{2+\eta}}\right]_{++} + \frac{\delta'(x)}{\eta} \tag{B9}$$

to obtain the expansion

$$\frac{1}{x^{2+\eta}} = \frac{\delta'(x)}{\eta} - \frac{\delta(x)}{1+\eta} + \left[\frac{\theta(x)}{x^2}\right]_{++} - \eta\left[\frac{\log(x)\theta(x)}{x^2}\right]_{++} \tag{B10}$$

as in Eq. (2.40) of [53].

2. Vector plus distributions

The same techniques may be applied to divergent vector-valued functions. Since our operators $T_{(i,j)}$ live in $d \neq 4$ spacetime dimensions, we define the vector plus distribution (also known as the ξ^2 distribution) by the relations

$$\int_{q_T^2 < \xi^2} d^{d-2}\mathbf{q}_T[\theta(q_T^2)(f(\mathbf{q}_T)]_+^{\xi^2} = 0,$$
$$[\theta(q_T^2)f(\mathbf{q}_T)]_+^{\xi^2} = f(\mathbf{q}_T), \qquad q_T^2 > 0. \tag{B11}$$

When $f(\mathbf{q}_T) = f(q_T^2)$ is a rotationally symmetric function, we have

$$\int d^{d-2}\mathbf{q}_T f(q_T^2) = \int J_{q_T}^\epsilon f(q_T^2)dq_T^2, \tag{B12}$$

where

$$J_{q_T}^\epsilon \equiv \frac{S_{2-2\epsilon}}{2}q_T^{-2\epsilon} \tag{B13}$$

and $S_{d-2} = 2\pi^{\frac{d-2}{2}}/\Gamma(\frac{d-2}{2})$, e.g., $S_2 = 2\pi$. We also note that a $(d-2)$-dimensional delta function at the origin may be written as

$$\delta(\mathbf{q}_T) = \frac{\delta(q_T^2)}{J_{q_T}^\epsilon}. \tag{B14}$$

Therefore if

$$g(q_T^2) = \int_{\mathbf{p}_T^2 < q_T^2} d^{d-2}\mathbf{p}_T f(\mathbf{p}_T) \tag{B15}$$

for some rotationally invariant function $f(\mathbf{p}_T)$, then

$$f(q_T^2) = \frac{1}{J_{q_T}^\epsilon} \frac{d}{dq_T^2} g(q_T^2), \tag{B16}$$

which is useful for converting between distributions and their cumulants.

The vector plus distribution may be written as the limit

$$[f(\mathbf{q}_T)]_+^{\xi^2} = \lim_{\beta \to 0} A(\beta, \epsilon, \xi) \delta(\mathbf{q}_T) + \theta(q_T^2 - \beta^2) f(\mathbf{q}_T), \tag{B17}$$

where, from (B11),

$$A(\beta, \epsilon, \xi) = - \int_{\beta^2 < q_T^2 < \xi^2} d^{d-2} \mathbf{q}_T f(\mathbf{q}_T). \tag{B18}$$

For example, we have the explicit form of the ξ^2 distribution

$$\left[\frac{1}{\mathbf{q}_T^2}\right]_+^{\xi^2} = \xi^{-2\epsilon} \left(1 - \left(\frac{\beta^2}{\xi^2}\right)^{-\epsilon}\right) \frac{S_{d-2}}{2\epsilon} \delta(\mathbf{q}_T) + \frac{\theta(q_T^2 - \beta^2)}{\mathbf{q}_T^2} \tag{B19}$$

(where the limit $\beta \to 0$ is implicit). We can also derive the analog of (B3) [19],

$$\frac{1}{J_{q_T}^\epsilon} \frac{d}{dq_T^2} \theta(q_T^2) F(q_T^2) = \left[\theta(q_T^2) \frac{1}{J_{q_T}^\epsilon} \frac{dF(q_T^2)}{dq_T^2}\right]_+^{\xi^2} + \delta(\mathbf{q}_T) F(\xi^2). \tag{B20}$$

As in Appendix B 1, we may then derive the expansion

$$\frac{(\nu^2)^{-\eta/2}}{(\mathbf{q}_T^2)^{1-\eta/2}} = J_\xi^\epsilon \left(\frac{\nu^2}{\xi^2}\right)^{-\eta/2} \frac{\delta(\mathbf{q}_T)}{\frac{\eta}{2} - \epsilon} + \left[\frac{\theta(q_T^2)}{q_T^2}\right]_+^{\xi^2} + \frac{\eta}{2} \left[\frac{\log \frac{q_T^2}{\nu^2} \theta(q_T^2)}{q_T^2}\right]_+^{\xi^2} + \cdots. \tag{B21}$$

The choice of ξ in these identities is entirely arbitrary. However, since each diagram comes with an overall $\mu^{2\epsilon}$, and since these identities put all the ϵ dependence into the delta-function prefactor $J_\xi^\epsilon \propto \xi^{-2\epsilon}$, the canonical choice that avoids spurious logarithms is $\xi = \mu$.

It is convenient to rescale the vector plus distributions to have the same scaling dimensions and π counting as $\delta(\mathbf{q}_T)$. Borrowing from the generalized-log notation of [19], we define

$$\mathcal{L}_n(\mathbf{q}_T, \mu) = \frac{1}{J_\mu^\epsilon} \left[\frac{\log^n \frac{q_T^2}{\mu^2} \theta(q_T^2)}{\mathbf{q}_T^2}\right]_+^{\mu^2}. \tag{B22}$$

With these definitions, and taking $\nu = \mu = \xi$, we have

$$\frac{(\mu^2)^{-\eta/2}}{(\mathbf{q}_T^2)^{1-\eta/2}} = J_\mu^\epsilon \left(\frac{\delta(\mathbf{q}_T)}{\frac{\eta}{2} - \epsilon} + \mathcal{L}_0^{(0)}(\mathbf{q}_T, \mu) + \frac{\eta}{2} \mathcal{L}_1^{(0)}(\mathbf{q}_T, \mu) + \cdots\right). \tag{B23}$$

Finally, we also need the identity

$$\frac{\log \frac{q_T^2}{\mu^2}}{\mathbf{q}_T^2} = J_\mu^\epsilon \left(-\frac{\delta(\mathbf{q}_T)}{\epsilon^2} + \mathcal{L}_1^{(0)}(\mathbf{q}_T, \mu)\right), \tag{B24}$$

which appears in the context of calculations without a regulator.

APPENDIX C: FIXED-ORDER COMPARISON

In this section we compare our results to that of [53]. In that reference, the QCD cross section for the process $N_1 N_2 \to V + X$ up to NLP is decomposed into a sum of convolutions of coefficient functions multiplied by PDFs and their first derivatives, so that

$$\frac{1}{\sigma_0} \frac{d\sigma}{dq^2 dy d^2 \mathbf{q}_T} = \int \frac{dz_a dz_b}{z_a z_b} \left[C_{f_q f_{\bar{q}}}^{(0)}(z_a, z_b, q^2, q_T^2) f\left(\frac{x_a}{z_a}\right) f\left(\frac{x_b}{z_b}\right) + \frac{1}{q^2} C_{f_q f_{\bar{q}}}^{(2)}(z_a, z_b, q^2, q_T^2) f\left(\frac{x_a}{z_a}\right) f\left(\frac{x_b}{z_b}\right)\right.$$
$$+ \frac{1}{q^2} C_{f'_q f_{\bar{q}}}^{(2)}(z_a, z_b, q^2, q_T^2) \frac{x_a}{z_a} f'\left(\frac{x_a}{z_a}\right) f\left(\frac{x_b}{z_b}\right) + \frac{1}{q^2} C_{f_q f'_{\bar{q}}}^{(2)}(z_a, z_b, q^2, q_T^2) f\left(\frac{x_a}{z_a}\right) \frac{x_b}{z_b} f'\left(\frac{x_b}{z_b}\right)$$
$$\left. + \frac{1}{q^2} C_{f'_q f'_{\bar{q}}}^{(2)}(z_a, z_b, q^2, q_T^2) \frac{x_a}{z_a} f'\left(\frac{x_a}{z_a}\right) \frac{x_b}{z_b} f'\left(\frac{x_b}{z_b}\right)\right], \tag{C1}$$

where at one-loop

$$C_{f_q f_{\bar{q}}}^{(0)} = \bar{\alpha} \left\{\delta(\bar{z}_a) \delta(\bar{z}_b) \delta(q_T^2) \left(-\log^2 \frac{q^2}{\mu^2} + 3 \log \frac{q^2}{\mu^2} - 8 + 7\zeta_2\right) + \left[\frac{1}{q_T^2}\right]_+^{\mu^2} \left(\delta(\bar{z}_a) \left[\frac{1 + z_b^2}{\bar{z}_b}\right]_+ + \delta(\bar{z}_b) \left[\frac{1 + z_a^2}{\bar{z}_a}\right]_+\right)\right.$$
$$\left. - \delta(\bar{z}_a) \delta(\bar{z}_b) \left(2 \left[\frac{\log q_T^2/q^2}{q_T^2}\right]_+^{\mu^2} + 3 \left[\frac{1}{q_T^2}\right]_+^{\mu^2}\right) + \delta(q_T^2)(\bar{z}_a \delta(\bar{z}_b) + \bar{z}_b \delta(\bar{z}_a) - \zeta_2 \delta(\bar{z}_a) \delta(\bar{z}_b))\right\}, \tag{C2}$$

at LP, and

$$C^{(2)}_{f_q f_{\bar{q}}} = \bar{\alpha}\left[-4\delta(\bar{z}_a)\delta(\bar{z}_b) - \delta(\bar{z}_a)\frac{1 + z_b^2 - 4z_b^3}{z_b} - \frac{1 + z_a^2 - 4z_a^3}{z_a}\delta(\bar{z}_b)\right], \quad \text{(C3)}$$

$$C^{(2)}_{f'_q f_{\bar{q}}} = \bar{\alpha}\left[\left(-\log\frac{q^2}{q_T^2} - 1\right)\delta(\bar{z}_a)\delta(\bar{z}_b) + \delta(\bar{z}_a)\left(\frac{1 + 3z_b + 2z_b^2}{2z_b} - \left[\frac{1}{\bar{z}_b}\right]_+\right) - \left(\frac{1 + z_a + 2z_a^3}{2z_a} + \left[\frac{1}{\bar{z}_a}\right]_+\right)\delta(\bar{z}_b)\right], \quad \text{(C4)}$$

$$C^{(2)}_{f_q f'_{\bar{q}}} = \bar{\alpha}\left[\left(-\log\frac{q^2}{q_T^2} - 1\right)\delta(\bar{z}_a)\delta(\bar{z}_b) - \delta(\bar{z}_a)\left(\frac{1 + z_b + 2z_b^3}{2z_b} + \left[\frac{1}{\bar{z}_b}\right]_+\right) + \left(\frac{1 + 3z_a + 2z_a^2}{2z_a} - \left[\frac{1}{\bar{z}_a}\right]_+\right)\delta(\bar{z}_b)\right], \quad \text{(C5)}$$

$$C^{(2)}_{f'_q f'_{\bar{q}}} = \bar{\alpha}\left[\left(2\log\frac{q^2}{q_T^2} + 4\right)\delta(\bar{z}_a)\delta(\bar{z}_b) - \delta(\bar{z}_a)\left(\frac{1 - 2z_b - z_b^2}{2z_b} + 2\left[\frac{1}{\bar{z}_b}\right]_+\right) + \left(\frac{1 - 2z_a - z_a^2}{2z_a} + 2\left[\frac{1}{\bar{z}_a}\right]_+\right)\delta(\bar{z}_b)\right] \quad \text{(C6)}$$

at NLP.

Since the $x_{a,b}$ in Eq. (C1) differ from the $\xi_{1,2}$ used in Eq. (11) at $O(q_T^2/q_L^2)$, and since our results are expressed entirely in terms of PDFs instead of PDFs and their first derivatives, the results in Eqs. (C2) and (C6) are related to Eqs. (88) and (89) by a change of variables, integration by parts, and a few distributional identities. Working in the hadronic center-of-mass frame for simplicity, where $P_1^- = \sqrt{s} = P_2^+$, the variables $x_{a,b}$ may be written in terms of $\xi_{1,2}$ as

$$x_a = \xi_1\left(1 - \frac{1}{2}\frac{q_T^2}{q_L^2} + \cdots\right),$$

$$x_b = \xi_2\left(1 - \frac{1}{2}\frac{q_T^2}{q_L^2} + \cdots\right). \quad \text{(C7)}$$

Expanding (C1)–(C6) up to $O(q_T^2/q_L^2)$ gives

$$\frac{1}{\sigma_0}\frac{d\sigma}{dq^2 dy d^2\mathbf{q}_T} = \sigma_0 \int \frac{dz_1}{z_1}\frac{dz_2}{z_2}\left[C^{(0)}_{f_q f_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2)f\left(\frac{\xi_1}{z_1}\right)f\left(\frac{\xi_2}{z_2}\right)\right.$$

$$+ \frac{1}{q_L^2}\left(C^{(2)}_{f_q f_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2) + \delta C^{(0)}_{f_q f_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2)\right)f\left(\frac{\xi_1}{z_1}\right)f\left(\frac{\xi_2}{z_2}\right)$$

$$+ \frac{1}{q_L^2}\left(C^{(2)}_{f'_q f_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2) - \frac{1}{2}\frac{q_T^2}{q_L^2}C^{(0)}_{ff}(z_1, z_2)\right)\frac{\xi_1}{z_1}f'\left(\frac{\xi_1}{z_1}\right)f\left(\frac{\xi_2}{z_2}\right)$$

$$+ \frac{1}{q_L^2}\left(C^{(2)}_{f_q f'_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2) - \frac{1}{2}\frac{q_T^2}{q_L^2}C^{(0)}_{ff}(z_1, z_2)\right)f\left(\frac{\xi_1}{z_1}\right)\frac{\xi_2}{z_2}f'\left(\frac{\xi_2}{z_2}\right)$$

$$\left.+ \frac{1}{q_L^2}C^{(2)}_{f'_q f'_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2)\frac{\xi_1}{z_1}f'\left(\frac{\xi_1}{z_1}\right)\frac{\xi_2}{z_2}f'\left(\frac{\xi_2}{z_2}\right)\right], \quad \text{(C8)}$$

where $C^{(0)}_{f_q f_{\bar{q}}}(z_1, z_2, q_L^2, q_T^2) = C^{(0)}_{ff}(z_1, z_2, q_L^2; q_T^2)$ in Eq. (88), and

$$C^{(2)}_{f_q f_{\bar{q}}} + \delta C^{(0)}_{f_q f_{\bar{q}}} = \bar{\alpha}\left[-6\delta_1\delta_2 - \delta_1\frac{1 + z_2^2 - 4z_2^3}{z_2} - \frac{1 + z_1^2 - 4z_1^3}{z_1}\delta_2\right], \quad \text{(C9)}$$

$$C^{(2)}_{f'_q f_{\bar{q}}} - \frac{1}{2}\frac{q_T^2}{q_L^2}C^{(0)}_{f_q f_{\bar{q}}} = \bar{\alpha}\left[\left(-2\log\frac{q_L^2}{q_T^2} - 1\right)\delta_1\delta_2 + \delta_1\left(\frac{1 + 4z_2 + 3z_2^2}{2z_2} - 2\left[\frac{1}{\bar{z}_2}\right]_+\right) - \left(\frac{1 - z_1^2 + 2z_1^3}{2z_1} + 2\left[\frac{1}{\bar{z}_1}\right]_+\right)\delta_2\right], \quad \text{(C10)}$$

$$C^{(2)}_{f_q f'_{\bar{q}}} - \frac{1}{2}\frac{q_T^2}{q_L^2}C^{(0)}_{f_q f_{\bar{q}}} = \bar{\alpha}\left[\left(-2\log\frac{q_L^2}{q_T^2} - 1\right)\delta_1\delta_2\right.$$
$$-\delta_1\left(\frac{1-z_2^2+2z_2^3}{2z_2} + 2\left[\frac{1}{\bar{z}_2}\right]_+\right)$$
$$\left.+\left(\frac{1+4z_1+3z_1^2}{2z_1} - 2\left[\frac{1}{\bar{z}_1}\right]_+\right)\delta_2\right],$$

(C11)

$$C^{(2)}_{f'_q f'_{\bar{q}}} = \bar{\alpha}\left[\left(2\log\frac{q_L^2}{q_T^2} + 4\right)\delta_1\delta_2\right.$$
$$-\delta_1\left(\frac{1-2z_2-z_2^2}{2z_2} + 2\left[\frac{1}{\bar{z}_2}\right]_+\right)$$
$$\left.+\left(\frac{1-2z_1-z_1^2}{2z_1} + 2\left[\frac{1}{\bar{z}_1}\right]_+\right)\delta_2\right].$$

(C12)

Finally, the comparison is completed by applying the following integration by parts identities, valid when $f(x/z) = 0$ for $x \geq z$,

$$\int\frac{dz}{z}\delta(\bar{z})\frac{x}{z}f'\left(\frac{\xi}{z}\right) = \int\frac{dz}{z}[-z\delta'(\bar{z})]f\left(\frac{\xi}{z}\right),$$
$$\int\frac{dz}{z}z^n\frac{x}{z}f'\left(\frac{\xi}{z}\right) = \int\frac{dz}{z}[nz^n - z\delta(\bar{z})]f\left(\frac{\xi}{z}\right),$$
$$\int\frac{dz}{z}\left[\frac{1}{\bar{z}}\right]_+\frac{x}{z}f'\left(\frac{\xi}{z}\right) = \int\frac{dz}{z}\left(z\left[\frac{1}{\bar{z}^2}\right]_{++} + z\delta'(\bar{z})\right.$$
$$\left.- z\delta(\bar{z})\right)f\left(\frac{\xi}{z}\right).$$

(C13)

These identities transform the coefficient functions in Eq. (C9) from acting on derivatives of PDFs to the equivalent form of coefficient functions acting only on PDFs, and in doing so reproduces the NLP coefficient function in Eq. (89).

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Factorization of hard processes in QCD, Adv. Ser. Dir. High Energy Phys. **5**, 1 (1989).

[2] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, Cambridge, England, 1996).

[3] J. Collins, *Foundations of Perturbative QCD*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, Cambridge, England, 2011).

[4] J. Collins, D. E. Soper, and G. Sterman, Transverse momentum distribution in Drell-Yan pair and W and Z boson production, Nucl. Phys. **B250**, 199 (1985).

[5] Y. L. Dokshitzer, D. Diakonov, and S. I. Troian, Hard processes in quantum chromodynamics, Phys. Rep. **58**, 269 (1980).

[6] Y. L. Dokshitzer, D. Diakonov, and S. I. Troian, On the transverse momentum distribution of massive lepton pairs, Phys. Lett. **79B**, 269 (1978).

[7] G. Parisi and R. Petronzio, Small transverse momentum distributions in hard processes, Nucl. Phys. **B154**, 427 (1979).

[8] G. Curci, M. Greco, and Y. Srivastava, QCD jets from coherent states, Nucl. Phys. **B159**, 451 (1979).

[9] R. K. Ellis and S. Veseli, W and Z transverse momentum distributions: Resummation in q_T space, Nucl. Phys. **B511**, 649 (1998).

[10] S. Frixione, P. Nason, and G. Ridolfi, Problems in the resummation of soft gluon effects in the transverse momentum distributions of massive vector bosons in hadronic collisions, Nucl. Phys. **B542**, 311 (1999).

[11] C. T. H. Davies and W. J. Stirling, Nonleading corrections to the Drell-Yan cross-section at small transverse momentum, Nucl. Phys. **B244**, 337 (1984).

[12] C. T. H. Davies, B. R. Webber, and W. J. Stirling, Drell-Yan cross sections at small transverse momentum, Nucl. Phys. **B256**, 413 (1984).

[13] D. de Florian and M. Grazzini, Next-to-Next-to-Leading Logarithmic Corrections at Small Transverse Momentum in Hadronic Collisions, Phys. Rev. Lett. **85**, 4678 (2000).

[14] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini, Production of Drell-Yan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy, Phys. Lett. B **696**, 207 (2011).

[15] S. Catani, D. de Florian, G. Ferrera, and M. Grazzini, Vector boson production at hadron colliders: transverse-momentum resummation and leptonic decay, J. High Energy Phys. 12 (2015) 047.

[16] T. Becher, M. Neubert, and B. D. Pecjak, Factorization and momentum-space resummation in deep-inelastic scattering, J. High Energy Phys. 01 (2007) 076.

[17] T. Becher and M. Neubert, Drell-Yan production at small q_T, transverse parton distributions and the collinear anomaly, Eur. Phys. J. C **71**, 1665 (2011).

[18] M. G. Echevarria, A. Idilbi, and I. Scimemi, Factorization theorem for Drell-Yan at low q_T and transverse momentum distributions on-the-light-cone, J. High Energy Phys. 07 (2012) 002.

[19] M. A. Ebert and F. J. Tackmann, Resummation of transverse momentum distributions in distribution space, J. High Energy Phys. 02 (2017) 110.

[20] W. Bizon, P. F. Monni, E. Re, L. Rottoli, and P. Torrielli, Momentum-space resummation for transverse observables and the Higgs $p_\perp$ at N³LL + NNLO, J. High Energy Phys. 02 (2018) 108.

[21] V. Bertone, I. Scimemi, and A. Vladimirov, Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production, J. High Energy Phys. 06 (2019) 028.

[22] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza, and M. Radici, Transverse-momentum-dependent parton distributions up to N³LL from Drell-Yan data, J. High Energy Phys. 07 (2020) 117.

[23] M. A. Ebert, J. K. L. Michel, I. W. Stewart, and F. J. Tackmann, Drell-Yan q_T resummation of fiducial power corrections at N³LL, J. High Energy Phys. 04 (2021) 102.

[24] T. Becher and T. Neumann, Fiducial q_T resummation of color-singlet processes at N³LL + NNLO, J. High Energy Phys. 03 (2021) 199.

[25] S. Camarda, L. Cieri, and G. Ferrera, Drell-Yan lepton-pair production: q_T resummation at N³LL accuracy and fiducial cross sections at N³LO, arXiv:2103.04974.

[26] E. Re, L. Rottoli, and P. Torrielli, Fiducial Higgs and Drell-Yan distributions at N³LL′ + NNLO with RadISH, arXiv:2104.07509.

[27] R. K. Ellis, G. Martinelli, and R. Petronzio, Lepton pair production at large transverse momentum in second order QCD, Nucl. Phys. **B211,** 106 (1983).

[28] R. J. Gonsalves, J. Pawlowski, and C.-F. Wai, QCD radiative corrections to electroweak boson production at large transverse momentum in hadron collisions, Phys. Rev. D **40,** 2245 (1989).

[29] C. W. Bauer, S. Fleming, and M. E. Luke, Summing Sudakov logarithms in $B \to X_s \gamma$ in effective field theory, Phys. Rev. D **63,** 014006 (2000).

[30] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, An effective field theory for collinear and soft gluons: Heavy to light decays, Phys. Rev. D **63,** 114020 (2001).

[31] C. W. Bauer and I. W. Stewart, Invariant operators in collinear effective theory, Phys. Lett. B **516,** 134 (2001).

[32] C. W. Bauer, D. Pirjol, and I. W. Stewart, Soft collinear factorization in effective field theory, Phys. Rev. D **65,** 054022 (2002).

[33] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, Hard scattering factorization from effective field theory, Phys. Rev. D **66,** 014017 (2002).

[34] M. Beneke, A. P. Chapovsky, M. Diehl, and T. Feldmann, Soft collinear effective theory and heavy to light currents beyond leading power, Nucl. Phys. **B643,** 431 (2002).

[35] M. Beneke and T. Feldmann, Multipole-expanded soft collinear effective theory with non-Abelian gauge symmetry, Phys. Lett. B **553,** 267 (2003).

[36] I. Moult, I. W. Stewart, G. Vita, and H. X. Zhu, First subleading power resummation for event shapes, J. High Energy Phys. 08 (2018) 013.

[37] M. Beneke, A. Broggio, M. Garny, S. Jaskiewicz, R. Szafron, L. Vernazza, and J. Wang, Leading-logarithmic threshold resummation of the Drell-Yan process at next-to-leading power, J. High Energy Phys. 03 (2019) 043.

[38] M. Beneke, M. Garny, S. Jaskiewicz, R. Szafron, L. Vernazza, and J. Wang, Leading-logarithmic threshold resummation of Higgs production in gluon fusion at next-to-leading power, J. High Energy Phys. 01 (2020) 094.

[39] A. Bhattacharya, I. Moult, I. W. Stewart, and G. Vita, Helicity methods for high multiplicity subleading soft and collinear limits, J. High Energy Phys. 05 (2019) 192.

[40] I. Moult, G. Vita, and K. Yan, Subleading power resummation of rapidity logarithms: The energy-energy correlator in $\mathcal{N} = 4$ SYM, J. High Energy Phys. 07 (2020) 005.

[41] Z. L. Liu and M. Neubert, Factorization at subleading power and endpoint-divergent convolutions in $h \to \gamma\gamma$ decay, J. High Energy Phys. 04 (**2020**) 033.

[42] Z. L. Liu, B. Mecaj, M. Neubert, and X. Wang, Factorization at subleading power and endpoint divergences in soft-collinear effective theory, Phys. Rev. D **104,** 014004 (2021).

[43] Z. L. Liu, B. Mecaj, M. Neubert, and X. Wang, Factorization at subleading power and endpoint divergences in $h \to \gamma\gamma$ decay. Part II. Renormalization and scale evolution, J. High Energy Phys. 01 (2021) 077.

[44] M. Kramer, E. Laenen, and M. Spira, Soft gluon radiation in Higgs boson production at the LHC, Nucl. Phys. **B511,** 523 (1998).

[45] A. A. Penin, High-energy limit of quantum electrodynamics beyond sudakov approximation, Phys. Lett. B **745,** 69 (2015).

[46] N. Bahjat-Abbas, D. Bonocore, J. Sinninghe Damsté, E. Laenen, L. Magnea, L. Vernazza, and C. D. White, Diagrammatic resummation of leading-logarithmic threshold effects at next-to-leading power, J. High Energy Phys. 11 (2019) 002.

[47] L. Cieri, C. Oleari, and M. Rocco, Higher-order power corrections in a transverse-momentum cut for colour-singlet production at NLO, Eur. Phys. J. C **79,** 852 (2019).

[48] M. van Beekveld, E. Laenen, J. S. Damst, and L. Vernazza, Next-to-leading power threshold corrections for finite order and resummed colour-singlet cross sections, J. High Energy Phys. 05 (2021) 114.

[49] C. Oleari and M. Rocco, Power corrections in a transverse-momentum cut for vector-boson production at NNLO: The qg-initiated real-virtual contribution, Eur. Phys. J. C **81,** 183 (2021).

[50] R. Boughezal, A. Isgrò, and F. Petriello, Next-to-leading power corrections to $V + 1$ jet production in N-jettiness subtraction, Phys. Rev. D **101,** 016005 (2020).

[51] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, A formalism for the systematic treatment of rapidity logarithms in quantum field theory, J. High Energy Phys. 05 (2012) 084.

[52] Y. Li, D. Neill, and H. X. Zhu, An exponential regulator for rapidity divergences, Nucl. Phys. **B960,** 115 (2020).

[53] M. A. Ebert, I. Moult, I. W. Stewart, F. J. Tackmann, G. Vita, and H. X. Zhu, Subleading power rapidity divergences and power corrections for q_T, J. High Energy Phys. 04 (2019) 123.

[54] C. W. Bauer, A. V. Manohar, and P. F. Monni, Disentangling observable dependence in SCET$_I$ and SCET$_{II}$ anomalous dimensions: angularities at two loops, J. High Energy Phys. 07 (2021) 214.

[55] R. Goerke and M. Luke, Power counting and modes in SCET, J. High Energy Phys. 02 (2018) 147.

[56] R. Goerke and M. Inglis-Whalen, Renormalization of dijet operators at order $1/Q^2$ in soft-collinear effective theory, J. High Energy Phys. 05 (**2018**) 023.

[57] M. Inglis-Whalen, M. Luke, and A. Spourdalakis, Rapidity logarithms in SCET without modes, Nucl. Phys. **A1014,** 122260 (2021).

[58] M. Beneke and V. A. Smirnov, Asymptotic expansion of Feynman integrals near threshold, Nucl. Phys. **B522,** 321 (1998).

[59] A. V. Manohar and I. W. Stewart, The zero-bin and mode factorization in quantum field theory, Phys. Rev. D **76,** 074002 (2007).

[60] J. C. Collins and A. Metz, Universality of Soft and Collinear Factors in Hard-Scattering Factorization, Phys. Rev. Lett. **93,** 252001 (2004).

[61] M. Diehl, J. R. Gaunt, D. Ostermeier, P. Plößl, and A. Schäfer, Cancellation of Glauber gluon exchange in the double Drell-Yan process, J. High Energy Phys. 01 (2016) 076.

[62] I. Z. Rothstein and I. W. Stewart, An effective field theory for forward scattering and factorization violation, J. High Energy Phys. 08 (2016) 025.

[63] C. W. Bauer, B. O. Lange, and G. Ovanesyan, On glauber modes in soft-collinear effective theory, J. High Energy Phys. 07 (2011) 077.

[64] A. V. Manohar, Deep inelastic scattering as $x \to 1$ using soft collinear effective theory, Phys. Rev. D **68,** 114019 (2003).

[65] S. M. Freedman and R. Goerke, Renormalization of subleading dijet operators in soft-collinear effective theory, Phys. Rev. D **90,** 114010 (2014).

[66] K. S. M. Lee and I. W. Stewart, Factorization for power corrections to $B \to X_s\gamma$ and $B \to X_u\ell\bar{\nu}$, Nucl. Phys. **B721,** 325 (2005).

[67] R. J. Hill, T. Becher, S. J. Lee, and M. Neubert, Sudakov resummation for subleading SCET currents and heavy-to-light form-factors, J. High Energy Phys. 07 (2004) 081.

[68] D. Pirjol and I. W. Stewart, A complete basis for power suppressed collinear ultrasoft operators, Phys. Rev. D **67,** 094005 (2003).

[69] J. Chay and C. Kim, Collinear effective theory at subleading order and its application to heavy—light currents, Phys. Rev. D **65,** 114016 (2002).

[70] A. V. Manohar, T. Mehen, D. Pirjol, and I. W. Stewart, Reparameterization invariance for collinear operators, Phys. Lett. B **539,** 59 (2002).

[71] C. W. Bauer, D. Pirjol, and I. W. Stewart, On power suppressed operators and gauge invariance in SCET, Phys. Rev. D **68,** 034021 (2003).

[72] J.-y. Chiu, A. Fuhrer, A. H. Hoang, R. Kelley, and A. V. Manohar, Soft-collinear factorization and zero-bin subtractions, Phys. Rev. D **79,** 053007 (2009).

[73] V. A. Smirnov, Asymptotic expansions of two loop Feynman diagrams in the Sudakov limit, Phys. Lett. B **404,** 101 (1997).

[74] I. Moult, I. W. Stewart, and G. Vita, Subleading power factorization with radiative functions, J. High Energy Phys. 11 (2019) 153.

[75] D. W. Kolodrubetz, I. Moult, and I. W. Stewart, Building blocks for subleading helicity operators, J. High Energy Phys. 05 (2016) 139.

[76] I. Moult, I. W. Stewart, G. Vita, and H. X. Zhu, The soft quark Sudakov, J. High Energy Phys. 05 (2020) 089.

[77] M. Beneke, A. Broggio, S. Jaskiewicz, and L. Vernazza, Threshold factorization of the Drell-Yan process at next-to-leading power, J. High Energy Phys. 07 (2020) 078.

[78] M. Beneke, M. Garny, R. Szafron, and J. Wang, Violation of the Kluberg-Stern-Zuber theorem in SCET, J. High Energy Phys. 09 (2019) 101.

[79] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, The Rapidity Renormalization Group, Phys. Rev. Lett. **108,** 151601 (2012).

[80] G. Paz, Subleading jet functions in inclusive B decays, J. High Energy Phys. 06 (**2009**) 083.

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

[82] C. Lee and G. F. Sterman, Momentum flow correlations from event shapes: Factorized soft gluons and soft-collinear effective theory, Phys. Rev. D **75,** 014022 (2007).

[83] A. Idilbi and T. Mehen, On the equivalence of soft and zero-bin subtractions, Phys. Rev. D **75,** 114017 (2007).

[84] A. Idilbi and T. Mehen, Demonstration of the equivalence of soft and zero-bin subtractions, Phys. Rev. D **76,** 094015 (2007).

[85] S. M. Freedman and M. Luke, SCET, QCD and Wilson lines, Phys. Rev. D **85,** 014003 (2012).

[86] S. M. Freedman, Subleading corrections to thrust using effective field theory, arXiv:1303.1558.

[87] I. Feige, D. W. Kolodrubetz, I. Moult, and I. W. Stewart, A complete basis of helicity operators for subleading factorization, J. High Energy Phys. 11 (2017) 142.