**DEUTSCHES ELEKTRONEN-SYNCHROTRON**
**Ein Forschungszentrum der Helmholtz-Gemeinschaft**

DESY 19-135
NIKHEF 2019-021
arXiv:1908.00985
August 2019

# Generalized Threshold Factorization with Full Collinear Dynamics

G. Lustermans

*Institute for Theoretical Physics Amsterdam and Delta Institute for Theoretical Physics, University of Amsterdam, The Netherlands*

*and*

*Nikhef Theory Group, Amsterdam, The Netherlands*

J. K. L. Michel, F. J. Tackmann

*Deutsches Elektronen-Synchrotron DESY, Hamburg*

ISSN 0418-9833

**NOTKESTRASSE 85 - 22607 HAMBURG**

DESY behält sich alle Rechte für den Fall der Schutzrechtserteilung und für die wirtschaftliche Verwertung der in diesem Bericht enthaltenen Informationen vor.

DESY reserves all rights for commercial use of information included in this report, especially in case of filing application for or grant of patents.

To be sure that your reports and preprints are promptly included in the HEP literature database
send them to (if possible by air mail):

| | |
|---|---|
| DESY<br>Zentralbibliothek<br>Notkestraße 85<br>22607 Hamburg<br>Germany | DESY<br>Bibliothek<br>Platanenallee 6<br>15738 Zeuthen<br>Germany |

DESY 19-135
NIKHEF 2019-021

# Generalized Threshold Factorization with Full Collinear Dynamics

Gillian Lustermans,[1,2,*] Johannes K. L. Michel,[3,†] and Frank J. Tackmann,[3,‡]

[1] *Institute for Theoretical Physics Amsterdam and Delta Institute for Theoretical Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands*
[2] *Nikhef, Theory Group, Science Park 105, 1098 XG, Amsterdam, The Netherlands*
[3] *Theory Group, Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany*

(Dated: August 2, 2019)

Soft threshold factorization has been used extensively to study hadronic collisions. It is derived in the limit where the momentum fractions $x_{a,b}$ of both incoming partons approach $x_{a,b} \to 1$. We present a generalized threshold factorization theorem for color-singlet processes, which holds in the weaker limit of only $x_a \to 1$ for generic $x_b$ (or vice versa), corresponding to the limit of large rapidity but generic invariant mass of the produced color singlet. It encodes the complete soft and/or collinear singular structure in the partonic momentum fractions to all orders in perturbation theory, including in particular flavor-nondiagonal partonic channels at leading power. It provides a more powerful approximation than the classic soft threshold limit, capturing a much larger set of contributions. We demonstrate this explicitly for the $Z$ and Higgs rapidity spectrum to NNLO, and we use it to predict a nontrivial set of its N$^3$LO contributions. Our factorization theorem provides the relevant resummation of large-$x$ logarithms in the rapidity spectrum required for resummation-improved PDF fits. One of our factorization ingredients is a new beam function closely related to the $N$-jettiness beam function. As a byproduct, we identify the correct soft threshold factorization for rapidity spectra among the differing results in the literature.

arXiv:1908.00985v1 [hep-ph] 2 Aug 2019

## I. INTRODUCTION

Color-singlet processes play a central role in the LHC physics program. The $pp \to Z, W$ Drell-Yan processes are precision benchmarks, providing determinations of electroweak parameters and important inputs for fits of parton distribution functions (PDFs). Higgs and diboson processes provide strong sensitivity to possible contributions beyond the Standard Model.

We consider the production of a generic color-singlet final state $L$ together with hadronic radiation $X$,

$$p(P_a^\mu) + p(P_b^\mu) \to L(q^\mu) + X(P_X^\mu)\,, \tag{1}$$

at hadronic center-of-mass energy $E_{\mathrm{cm}}^2 = (P_a + P_b)^2$. The key observables characterizing $L$ are its total invariant mass $Q \equiv \sqrt{q^2}$, rapidity $Y$, and transverse momentum $q_T \equiv |\vec{q}_T|$. We define the momentum fractions

$$x_a = \frac{Q}{E_{\mathrm{cm}}} e^{+Y}\,, \quad x_b = \frac{Q}{E_{\mathrm{cm}}} e^{-Y}\,, \quad \tau = x_a x_b\,, \tag{2}$$

which are equivalent to $Q$ and $Y$. The cross section differential in $x_{a,b}$ is given by [1–3]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_a \mathrm{d}x_b} = \int \frac{\mathrm{d}z_a}{z_a} \frac{\mathrm{d}z_b}{z_b}\, \hat{\sigma}_{ij}(z_a, z_b)\, f_i\Big(\frac{x_a}{z_a}\Big)\, f_j\Big(\frac{x_b}{z_b}\Big)\,, \tag{3}$$

where $\hat{\sigma}_{ij}(z_a, z_b)$ denotes the perturbatively calculable partonic cross section and $f_{i,j}(x)$ are the standard PDFs. We always implicitly sum over parton indices $i, j$, and keep the dependence on renormalization scales $\mu$ implicit.

In the soft threshold limit $\tau \to 1$, which implies that both $x_{a,b} \to 1$, Eq. (3) factorizes further [4–9],

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_a \mathrm{d}x_b} = H_{ij}(Q^2) \int \mathrm{d}k^-\, \mathrm{d}k^+ S(k^-, k^+) \times f_i^{\mathrm{thr}}\Big[x_a\Big(1 + \frac{k^-}{Q e^{+Y}}\Big)\Big] f_j^{\mathrm{thr}}\Big[x_b\Big(1 + \frac{k^+}{Q e^{-Y}}\Big)\Big]. \tag{4}$$

In this limit, the hadronic final state is forced to be soft, and is described by the soft function $S$, which encodes soft-gluon emissions from the colliding hard partons. Furthermore, only the hard Born processes, e.g. $q\bar{q} \to Z$ or $gg \to H$, contribute. They are encoded in the hard function $H_{ij}$, including the Born-like virtual corrections. Any nondiagonal partonic channels like $qg \to Lq$ vanish for $\tau \to 1$. The threshold PDF $f_i^{\mathrm{thr}}(x)$ encodes the extraction of parton $i$ from the proton for $x \to 1$. At partonic level, Eq. (4) implies that for $z \equiv z_a z_b \to 1$, which requires both $z_{a,b} \to 1$,

$$\hat{\sigma}_{ij}(z_a, z_b) = H_{ij}\, \hat{S}(z_a, z_b)\,, \tag{5}$$

up to power corrections in $1 - z$.

While taking $\tau \to 1$ forces $z \to 1$, even for typical LHC values of $\tau \ll 1$, the $z \sim 1$ region often numerically dominates the cross section. Soft threshold factorization has thus been widely used for decades. It enables the all-order resummation of the leading terms in $1 - z$, see e.g. Refs. [4–27]. The resummation at next-to-leading power (NLP) in $1 - z$ has also received recent interest [28–30]. Another important use is to approximate the fixed-order cross section by expanding in $1 - z$, e.g. at N$^3$LO [31–38].

In this letter, we derive the factorization that generalizes Eqs. (4) and (5) to the weaker limit where only one of $x_{a,b}$ (or $z_{a,b}$) approaches 1 while keeping the exact

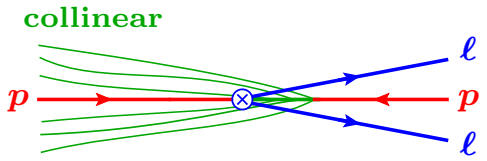


FIG. 1. Illustration of Drell-Yan at large dilepton rapidity.

dependence on the other variable. This corresponds to the kinematic limit $|Y| \to Y_{\max} = \ln(E_{\rm cm}/Q)$ for generic (including small) $Q$ values.

## II. GENERALIZED THRESHOLD FACTORIZATION

We use light-cone coordinates $p^\mu \equiv (n\cdot p, \bar n\cdot p, p_\perp) \equiv (p^+, p^-, p_\perp)$ with respect to lightlike vectors $n^\mu \equiv (1,\hat z)$ and $\bar n^\mu \equiv (1,-\hat z)$ along the beam axis $\hat z$. We first consider the observables $q^\mp$ instead of $Q$ and $Y$, with corresponding momentum fractions

$$x_\mp \equiv \frac{q^\mp}{P_{a,b}^\mp} = \frac{\sqrt{Q^2+q_T^2}}{E_{\rm cm}}\, e^{\pm Y}\,. \tag{6}$$

We consider the generalized threshold limit

$$\lambda_{\rm QCD}^2 \ll \lambda^2 \sim 1-x_- \ll 1 \quad \text{for generic } x_+\,, \tag{7}$$

where $\lambda_{\rm QCD} \equiv \Lambda_{\rm QCD}/Q$ and $\lambda$ are power-counting parameters. In this limit, illustrated in Fig. 1, $L$ has large $Y$ while the emissions in $X$ become collimated in the opposite direction with typical momenta

$$p_X^\mu \sim (q^+, P_a^- - q^-, p_{X\perp}) \sim (q^+, \lambda^2 q^-, \lambda\sqrt{q^+q^-})\,. \tag{8}$$

In this situation, the following factorization theorem holds to leading power in $1-x_-$,

$$\frac{{\rm d}\sigma}{{\rm d}x_-{\rm d}x_+} = H_{ij}(q^+q^-)\int {\rm d}t\, f_i^{\rm thr}\Big[x_-\Big(1+\frac{t}{q^+q^-}\Big)\Big] \times B_j(t,x_+)\,. \tag{9}$$

Here, $B_j(t,x)$ is the inclusive beam function that also appears in the factorization for $N$-jettiness [39, 40]. It depends on the transverse virtuality $t$ and momentum fraction $x$ of the colliding parton $j$. Since $t \sim p_X^2 \sim \lambda^2 Q^2 \gg \Lambda_{\rm QCD}^2$, it can be calculated perturbatively in terms of standard PDFs as [39, 41]

$$B_j(t,x) = \int\frac{{\rm d}z}{z}\,\mathcal{I}_{jk}(t,z)\, f_k\Big(\frac{x}{z}\Big)\,. \tag{10}$$

The matching coefficients $\mathcal{I}_{jk}$ are known to $\mathcal{O}(\alpha_s^2)$ [41–44], with progress at $\mathcal{O}(\alpha_s^3)$ [45, 46].

To derive Eq. (9), we use Soft-Collinear Effective Theory (SCET) [47–51], which is the effective theory of QCD in the limit $\lambda \ll 1$. The derivation proceeds in a standard fashion [39], with more details given in [52]. The key elements are the necessary degrees of freedom (modes) in the effective theory,

$$p_{\bar n} \sim Q\left(1,\lambda^2,\lambda\right), \quad P_{\bar n} \sim Q\left(1,\lambda_{\rm QCD}^2,\lambda_{\rm QCD}\right), \tag{11}$$
$$P_n \sim Q\left(\lambda_{\rm QCD}^2,1,\lambda_{\rm QCD}\right), \quad P_s \sim Q\left(\frac{\lambda_{\rm QCD}^2}{\lambda^2},\lambda^2,\lambda_{\rm QCD}\right).$$

The $p_{\bar n}$ collinear modes describe the QCD final state at the scale $\lambda Q$, which due to Eq. (8) is $\bar n$-collinear. The $P_{n,\bar n}$ collinear modes describe the PDFs at the scale $\lambda_{\rm QCD}Q$. The $P_s$ modes describe the soft interactions between the $p_{\bar n}$ and $P_n$ modes. Possible ultrasoft/Glauber modes $P_{us,G} \sim Q(\lambda_{\rm QCD}^2, \lambda_{\rm QCD}^2, \lambda_{\rm QCD}^{(2)})$ cancel as in Eq. (3) because the measurement is the same and its scale $\lambda Q \gg \lambda_{\rm QCD}Q$.

The hard function $H_{ij}$ in Eq. (9) is the same as in Eq. (4). It arises from matching the electroweak current for $L$ onto the corresponding SCET current. There are no interactions between the modes in the leading-power SCET Lagrangian, so the complete matrix element factorizes into separate ones in each sector. The matrix element of the combined $p_{\bar n}$ and $P_{\bar n}$ modes yields $B_j(t,x)$, and their separation leads to Eq. (10) [39, 41]. The combined matrix element of the $P_n$ and $P_s$ modes yields $f_i^{\rm thr}$ [53, 54]. The convolution structure in Eq. (9) follows from momentum conservation [52].

Next, we consider also measuring $\vec q_T$. From Eq. (8), it follows that generically $q_T \sim p_{X\perp} \sim \lambda Q$, so the $q_T$ dependence is entirely described by the $p_{\bar n}$ modes, which yields the factorization theorem

$$\frac{{\rm d}\sigma}{{\rm d}x_-{\rm d}x_+\,{\rm d}\vec q_T} = H_{ij}(q^+q^-)\int {\rm d}t\, f_i^{\rm thr}\Big[x_-\Big(1+\frac{t}{q^+q^-}\Big)\Big] \times B_j(t,\vec q_T,x_+)\,. \tag{12}$$

Here, $B_j(t,\vec k_T,x)$ is the double-differential beam function [55, 56] that also occurs in the joint resummation of $q_T$ and 0-jettiness [57, 58].

We can now change variables in Eq. (12) to $x_{a,b}$, using Eq. (6) and expanding in $\lambda$, which yields

$$\frac{{\rm d}\sigma}{{\rm d}x_a{\rm d}x_b\,{\rm d}\vec q_T} = H_{ij}(Q^2)\int {\rm d}t\, f_i^{\rm thr}\Big[x_a\Big(1+\frac{q_T^2}{2Q^2}+\frac{t}{Q^2}\Big)\Big] \times B_j(t,\vec q_T,x_b)\,. \tag{13}$$

Crucially, when expanding $x_- = x_a[1+q_T^2/(2Q^2)+\mathcal{O}(\lambda^4)]$, we have to keep the $q_T^2/(2Q^2) \sim \lambda^2$ term in the PDF argument because it is of the same order as $t/Q^2 \sim \lambda^2$. Integrating Eq. (13) over $\vec q_T$, we obtain

$$\frac{{\rm d}\sigma}{{\rm d}x_a{\rm d}x_b} = H_{ij}(Q^2)\int {\rm d}\tilde t\, f_i^{\rm thr}\Big[x_a\Big(1+\frac{\tilde t}{Q^2}\Big)\Big]\tilde B_j(\tilde t,x_b)\,. \tag{14}$$

The factorization theorems in Eqs. (13) and (14) hold at leading power in the generalized threshold limit $\lambda^2 \sim 1-x_a \ll 1$ for generic $x_b$. They are our key new results.

In Eq. (14) we changed variables to $\tilde{t} = t + q_T^2/2$, and defined the new modified beam function

$$\tilde{B}_j(\tilde{t}, x) = \int \mathrm{d}^2\vec{k}_T \, B_j\Big(\tilde{t} - \frac{k_T^2}{2}, \vec{k}_T, x\Big) \,. \tag{15}$$

It has the same $\mu$ evolution as $B_j(t,x)$ but different constant terms. It obeys a matching relation analogous to Eq. (10). Using the known results for $B_j(t, \vec{k}_T, x)$ [55, 56], we have calculated its matching coefficients $\tilde{\mathcal{I}}_{jk}(\tilde{t}, z)$ to $\mathcal{O}(\alpha_s^2)$ for $j = q$ and $\mathcal{O}(\alpha_s)$ for $j = g$ [52].

The factorization structure in Eqs. (9) and (14) turns out to be analogous to deep-inelastic scattering (DIS) at large Bjorken $x$ [4, 5, 53, 54, 59–61], which factorizes as $H_{ij} f_i^{\mathrm{thr}} \otimes J_j$, with the jet function $J_j$ describing collimated *final-state* radiation. For Eqs. (9) and (14) to be consistent, the $\mu$ dependence of the functions must cancel between them, and it does so in the same way as for DIS. The beam function is known to have the same $\mu$ evolution as the jet function [41]. The existence of the above factorization theorems provides an independent confirmation of this. The key differences to DIS are the additional dependence on $x_b$ and the nontrivial $q_T$ dependence. The latter bears some resemblance to different 1-jettiness definitions in exclusive DIS [62].

Since Eq. (14) is valid for $x_a \to 1$ and arbitrary $x_b$, it must contain the soft threshold factorization in Eq. (4) for $x_b \to 1$ as a special case. This implies [52]

$$\tilde{B}_j(\omega k^-, x_b) = \int \frac{\mathrm{d}k^+}{\omega} \, S(k^-, k^+) f_j^{\mathrm{thr}}\Big[x_b\Big(1 + \frac{k^+}{\omega}\Big)\Big] \tag{16}$$

to leading power in $1 - x_b$, and identically for $B_j$.

We can now combine Eq. (14) with the analogous result in the opposite limit $x_b \to 1$ for generic $x_a$ by adding the two and subtracting their overlap, which is precisely given by the soft limit. This yields our main result, the generalized threshold factorization theorem

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_a \mathrm{d}x_b} = H_{ij}\Big[f_i^{\mathrm{thr}} \otimes \tilde{B}_j + \tilde{B}_i \otimes f_j^{\mathrm{thr}} - S \otimes f_i^{\mathrm{thr}} f_j^{\mathrm{thr}}\Big] \,, \tag{17}$$

with the convolutions as in Eqs. (4) and (14). Analogous results hold for $x_\pm$ and differential in $\vec{q}_T$. Replacing $f_i^{\mathrm{thr}}[x(1+1-z)]$ by $f_i(x/z)/z$, which is justified at leading power in $1 - z$, and comparing to Eq. (3), we obtain the corresponding partonic factorization theorem

$$\begin{aligned}\hat{\sigma}_{ij}(z_a, z_b) = H_{k\ell}\big[&\delta_{ki}\, \hat{\mathcal{I}}_{\ell j}(z_a, z_b) + \hat{\mathcal{I}}_{ki}(z_b, z_a)\, \delta_{\ell j} \\ &- \delta_{ki}\, \delta_{\ell j}\, \hat{S}(z_a, z_b)\big] \,,\end{aligned} \tag{18}$$

where we changed variables to $z_{a,b}$ and defined

$$\begin{aligned}\hat{\mathcal{I}}_{ij}(z_a, z_b) &\equiv Q^2\, \tilde{\mathcal{I}}_{ij}\big[Q^2(1 - z_a), z_b\big] \,, \\ \hat{S}(z_a, z_b) &\equiv Q^2\, S\big[Qe^Y(1 - z_a), Qe^{-Y}(1 - z_b)\big] \,.\end{aligned} \tag{19}$$

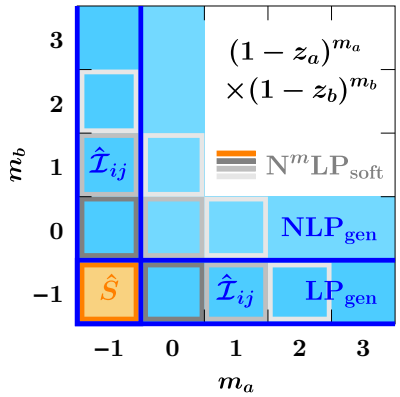


FIG. 2. Terms in the partonic cross section $\hat{\sigma}(z_a, z_b)$ captured by the soft and generalized threshold expansions.

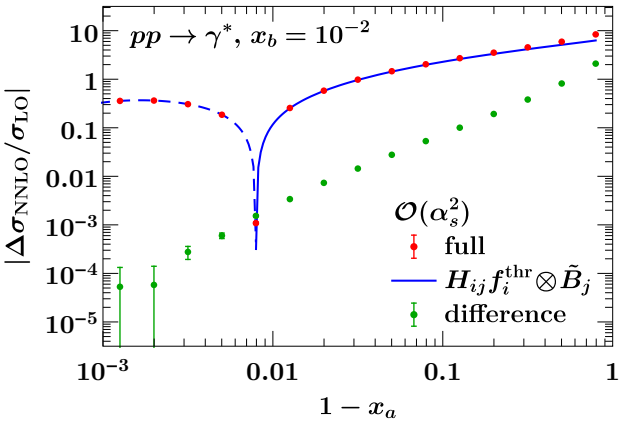


FIG. 3. Validation of the $\mathcal{O}(\alpha_s^2)$ contribution to $\mathrm{d}\sigma/\mathrm{d}x_a\mathrm{d}x_b$ predicted for $x_a \to 1$ by Eq. (14) (blue) against the full result from `Vrap` (red). Their difference (green) vanishes like a power as $1 - x_a \to 0$, as it must. The error bars indicate the integration uncertainties.

As illustrated in Fig. 2, $\hat{\mathcal{I}}_{ij}(z_a, z_b)$ captures all terms in $\hat{\sigma}(z_a, z_b)$ that are singular as $z_a \to 1$ with their exact dependence on $z_b$, and vice versa, with the overlap given by $\hat{S}(z_a, z_b)$. This includes flavor-nondiagonal contributions, e.g., $ij = qg$ with $k\ell = q\bar{q}$ or $gg$ in Eq. (18).

## III. VALIDATION

A nontrivial validation is to check Eq. (17) against the full fixed-order result. First, Eq. (16) implies

$$\hat{\mathcal{I}}_{ij}(z_a, z_b) = \delta_{ij}\, \hat{S}(z_a, z_b)\, \big[1 + \mathcal{O}(1 - z_b)\big] \,, \tag{20}$$

which we verified analytically to $\mathcal{O}(\alpha_s^2)$. It then suffices to check that Eq. (14) reproduces the fixed-order result for $x_a \to 1$ (or $z_a \to 1$) for generic $x_b$ (or $z_b$).

At NLO, we can check analytically for Drell-Yan and $gg \to H$ against the results of Refs. [63, 64], which are given as distributions in $z = z_a z_b$ and a variable

$y(z_a, z_b)$. Their translation to $(z_a, z_b)$ is nontrivial [52]. We find complete agreement for all partonic channels. We also combined all singular terms from Eq. (18) with the regular terms from Refs. [63, 64] to construct the full $\hat{\sigma}_{ij}(z_a, z_b)$ at NLO. The result is given in [52] and agrees with Refs. [6, 65, 66]. [1].

At NNLO, we numerically validate our own implementation of Eq. (14) in `SCETlib` [67] against `Vrap` [68]. We use flat PDFs, $f_i^{\text{thr}}(x) = f_i(x) = \theta(1-x)$, which amounts to taking cumulant integrals of the partonic cross section and provides the strongest possible numerical check. In Fig. 3, we compare the $\mathcal{O}(\alpha_s^2)$ contribution for Drell-Yan as a function of $1 - x_a$ at fixed $x_b = 10^{-2}$. We find perfect agreement. The breakdown into partonic channels is given in [52]. We also find similar agreement for other $x_b$ and for $pp \to Z/\gamma^*$ on the resonance.

## IV. ILLUSTRATIVE APPLICATIONS

The immediate question that arises is how well the generalized threshold limit approximates the full fixed-order result for physical PDFs, particularly in comparison to the soft limit. We use the `MMHT2014nnlo68cl` [69] PDFs, `Vrap` [68] to obtain the full NNLO result[2] and `SCETlib` [67] to implement Eq. (18). In Fig. 4 we compare the $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ contributions to the Drell-Yan rapidity spectrum at $Q = m_Z$, separated into quark channels ($q\bar{q} + qq'$) and channels involving a gluon ($qg + gq + gg$). Analogous results at $\mu = Q/2$ and for $gg \to H$ are provided in [52]. The generalized threshold limit approximates the full result well for all channels and all $Y$. As expected, it works particularly well toward large $Y$. It works significantly better than the soft limit, which only provides a poor approximation for the $q\bar{q}$ channel and none for the others. Currently, all ingredients are available to perform the resummation in the generalized threshold limit to N$^3$LL, which we leave to future work. We expect it to be much more powerful for improving the precision of perturbative predictions than the soft threshold resummation.

Next, we may ask how well the power expansion around the generalized threshold limit works and how it relates to the soft expansion. Consider the double expansion in $1 - z_a$ and $1 - z_b$, illustrated in Fig. 2,

$$\hat{\sigma}_{ij}(z_a, z_b) = \sum_{m_a, m_b} \hat{\sigma}_{ij}^{(m_a, m_b)}(z_a, z_b)\,, \tag{21}$$

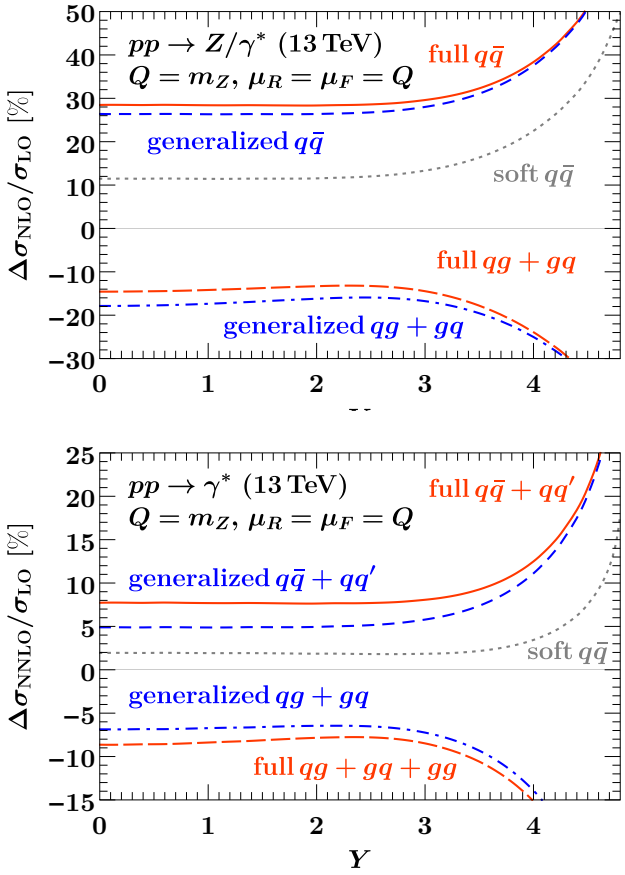


FIG. 4. The $\mathcal{O}(\alpha_s)$ (top) and $\mathcal{O}(\alpha_s^2)$ (bottom) contributions to $\sigma \equiv \mathrm{d}\sigma/\mathrm{d}Q\mathrm{d}Y$ normalized to the LO result. Shown are the full result (red), the generalized threshold approximation (blue), and the soft threshold approximation (gray).

where $\hat{\sigma}_{ij}^{(m_a,m_b)}(z_a, z_b) \sim (1-z_a)^{m_a}(1-z_b)^{m_b}$. Expanding around the soft $z = z_a z_b \to 1$ limit corresponds to counting powers of $(1-z)^{m_a+m_b+1}$. The leading-power result in Eq. (5) gives the $m_a = m_b = -1$ term. At the $m$th order, N$^m$LP$_{\text{soft}}$, we keep all terms with $m_a + m_b + 2 \leq m$. At leading power in the generalized expansion, Eq. (18) includes all terms with $\min\{m_a, m_b\} = -1$. Similarly, at the $m$th order, N$^m$LP$_{\text{gen}}$, we keep all terms with $\min\{m_a, m_b\} = m - 1$, so the *missing* corrections at N$^m$LP$_{\text{gen}}$ are $\mathcal{O}[(1-z_a)^m(1-z_b)^m]$.

In Fig. 5, we show the deviation from the exact result at various orders in both expansions for Drell-Yan at NLO, where we have full analytic control. Analogous results for $gg \to H$ are provided in [52]. The generalized expansion performs significantly better than the soft one for both partonic channels. We expect this to hold in general, since expanding a two-dimensional function along a one-dimensional boundary is superior to expanding it in a single point on that boundary.

In fact, as seen in Fig. 2, each order in the generalized expansion fully contains two orders in the soft expansion, and in particular, the LP$_{\text{gen}}$ result Eq. (18) contains the

[1] As discussed further in [52], several soft threshold factorizations differential in rapidity [19–23] differ from Eq. (4) and do not reproduce the correct soft limit already at NLO.

[2] The public `Vrap 0.9` assumes $f_q(x) = f_{\bar{q}}(x)$ for $q = s, c, b$. We modified it to allow for different sea quark and antiquark PDFs.

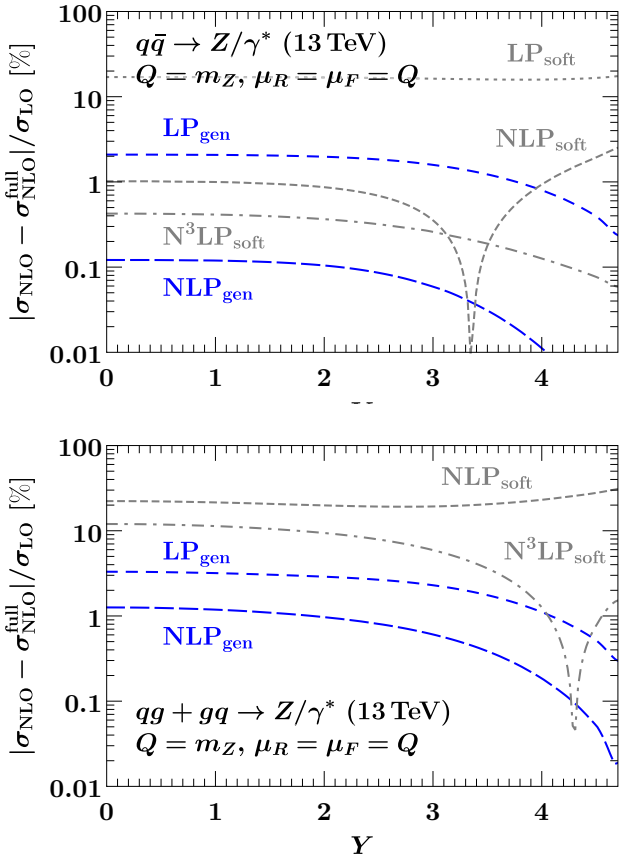


FIG. 5. Convergence of the generalized (blue) and soft (gray) threshold expansions. Shown are the deviations from the full NLO result for the $q\bar{q}$ channel (top) and the $qg + gq$ channel (bottom) normalized to the LO result.

entire NLP$_{\rm soft}$ contribution. This does not mean it can be used to perform the NLP$_{\rm soft}$ resummation because the $\mu$ evolution of $B_i(t,x)$ does not predict its $x$ dependence. It does, however, show that $H_{ij}$ factorizes for all partonic channels and reduces the problem to deriving the NLP$_{\rm soft}$ factorization for Eq. (16).

At N$^3$LO, Eq. (18) predicts a highly nontrivial set of terms for any color-singlet process, since all terms $\sim \mathcal{L}_n(1-z_a)$ in $\hat{\mathcal{I}}_{ij}(z_a,z_b)$ are known from its $\mu$ evolution, where $\mathcal{L}_n(y) \equiv [\ln^n(y)/y]_+$. To illustrate this, the coefficient of $\alpha_s^3/(4\pi)^3$ in $\hat{\sigma}_{ij}(\mu = Q)$ with $n \geq 3$ is

$$
\begin{aligned}
\hat{\sigma}_{ij}^{(3)} &= \mathcal{L}_5(1-z_a) H_{ij}^{(0)} \delta(1-z_b) \frac{(\Gamma_0^i)^3}{8} \\
&\quad + \mathcal{L}_4(1-z_a) H_{ik}^{(0)} \biggl[ \delta_{kj}\delta(1-z_b) \Bigl( -\frac{2\beta_0}{3} - \frac{\gamma_{B\,0}^i}{2} \Bigr) \\
&\quad + P_{kj}^{(0)}(z_b) \biggr] \frac{5}{8} (\Gamma_0^i)^2 + \mathcal{L}_3(1-z_a) \biggl\{ H_{ij}^{(1)} \delta(1-z_b) \frac{\Gamma_0^i}{2} \\
&\quad + H_{ik}^{(0)} \biggl[ \delta_{kj}\delta(1-z_b) \Bigl( \Gamma_1^i - \frac{\pi^2}{6} (\Gamma_0^i)^2 + \frac{\beta_0^2}{3} \\
&\quad + \frac{(\gamma_{B\,0}^i)^2}{4} + \frac{5}{6}\beta_0 \gamma_{B\,0}^i \Bigr) - P_{kj}^{(0)}(z_b) \Bigl( \frac{5\beta_0}{3} + \gamma_{B\,0}^i \Bigr) \\
&\quad + \bigl( P_{k\ell}^{(0)} \otimes P_{\ell j}^{(0)} \bigr)(z_b) + \tilde{I}_{kj}^{(1)}(z_b) \frac{\Gamma_0^i}{2} \biggr] \biggr\} \Gamma_0^i + \cdots .
\end{aligned}
\tag{22}
$$

The required ingredients are given in [52]. The extension down to $\mathcal{L}_0(1-z_a)$ and to the full $z_a$ dependence for $z_b \to 1$ is straightforward [70]. The $\delta(1-z_a)$ coefficient requires the still unknown $\mathcal{O}(\alpha_s^3)$ finite terms of the beam function. For Drell-Yan, Eq. (22) significantly extends the current knowledge at $\mathcal{O}(\alpha_s^3)$ [33], providing the full $z_b$ dependence for all partonic channels. For $gg \to H$, the extension to $\mathcal{L}_0(1-z_a)$ would also provide additional information beyond what is currently known [38].

## V. SUMMARY

We introduced a generalized threshold factorization, which is much more powerful than the often used soft threshold, thus opening the door to numerous applications to improve theoretical predictions for collider processes. It describes all kinematic limits in $(x_a, x_b)$ or $(Q, Y)$, including in particular $|Y| \to Y_{\rm max}$ at generic $Q$, which is directly accessible at the LHC. It enables the corresponding threshold resummation where only one PDF is probed at large $x$, which is not captured by the soft limit. At the partonic level, it captures all singularities of $\hat{\sigma}_{ij}(z_a, z_b)$, including nondiagonal partonic channels. It can be used to predict a rich set of terms at higher fixed order or to resum them to all orders. It is the weakest known limit in which the process-dependent virtual corrections ($H_{ij}$) factorize.

While we only considered color-singlet processes here, the same methods can be used to generalize the soft threshold factorization in other situations, such as for processes with heavy particles, jets, or identified hadrons in the final state.

*Acknowledgments* We thank M. Beneke, G. Billis, A. Broggio, G. Das, M. Diehl, L. Dixon, M. Ebert, S. Forte, A. Kulesza, S. Marzani, B. Mistlberger, F. Ringer, D. Scott, D. Soper, and W. Waalewijn for discussions. We thank M. Stahlhofen for providing the results of Ref. [56], and M. Prakash and V. Ravindran for help in comparing with Refs. [6, 66]. The authors thank each other's institutions for hospitality. This work was supported in part by the ERC grant ERC-STG-2015-677323 and the D-ITP consortium, a program of the Netherlands Organization for Scientific Research (NWO) that is funded by the Dutch Ministry of Education, Culture and Science (OCW).

* g.h.h.lustermans@uva.nl
† johannes.michel@desy.de
‡ frank.tackmann@desy.de

[1] G. T. Bodwin, Phys. Rev. **D31**, 2616 (1985), [Erratum: Phys. Rev.D34,3932(1986)].
[2] J. C. Collins, D. E. Soper, and G. F. Sterman, Nucl. Phys. **B261**, 104 (1985).
[3] J. C. Collins, D. E. Soper, and G. F. Sterman, Nucl. Phys. **B308**, 833 (1988).
[4] G. F. Sterman, Nucl. Phys. **B281**, 310 (1987).
[5] S. Catani and L. Trentadue, Nucl. Phys. **B327**, 323 (1989).
[6] V. Ravindran, J. Smith, and W. L. van Neerven, Nucl. Phys. **B767**, 100 (2007), arXiv:hep-ph/0608308.
[7] D. Westmark and J. F. Owens, Phys. Rev. **D95**, 056024 (2017), arXiv:1701.06716 [hep-ph].
[8] P. Banerjee, G. Das, P. K. Dhani, and V. Ravindran, Phys. Rev. **D97**, 054024 (2018), arXiv:1708.05706 [hep-ph].
[9] P. Banerjee, G. Das, P. K. Dhani, and V. Ravindran, Phys. Rev. **D98**, 054018 (2018), arXiv:1805.01186 [hep-ph].
[10] D. Appell, G. F. Sterman, and P. B. Mackenzie, Nucl. Phys. **B309**, 259 (1988).
[11] L. Magnea, Nucl. Phys. **B349**, 703 (1991).
[12] G. P. Korchemsky and G. Marchesini, Nucl. Phys. **B406**, 225 (1993), arXiv:hep-ph/9210281.
[13] H. Contopanagos, E. Laenen, and G. F. Sterman, Nucl. Phys. **B484**, 303 (1997), arXiv:hep-ph/9604313.
[14] S. Catani, M. L. Mangano, P. Nason, and L. Trentadue, Nucl. Phys. **B478**, 273 (1996), arXiv:hep-ph/9604351.
[15] A. V. Belitsky, Phys. Lett. **B442**, 307 (1998), arXiv:hep-ph/9808389.
[16] S. Moch and A. Vogt, Phys. Lett. **B631**, 48 (2005), arXiv:hep-ph/0508265.
[17] E. Laenen and L. Magnea, Phys. Lett. **B632**, 270 (2006), arXiv:hep-ph/0508284.
[18] A. Idilbi, X.-d. Ji, and F. Yuan, Nucl. Phys. **B753**, 42 (2006), arXiv:hep-ph/0605068.
[19] A. Mukherjee and W. Vogelsang, Phys. Rev. **D73**, 074005 (2006), arXiv:hep-ph/0601162.
[20] P. Bolzoni, Phys. Lett. **B643**, 325 (2006), arXiv:hep-ph/0609073.
[21] T. Becher, M. Neubert, and G. Xu, JHEP **07**, 030 (2008), arXiv:0710.0680 [hep-ph].
[22] M. Bonvini, S. Forte, and G. Ridolfi, Nucl. Phys. **B847**, 93 (2011), arXiv:1009.5691 [hep-ph].
[23] M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, M. Ubiali, R. D. Ball, V. Bertone, S. Carrazza, and N. P. Hartland, JHEP **09**, 191 (2015), arXiv:1507.01006 [hep-ph].
[24] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, Eur. Phys. J. **C73**, 2480 (2013), arXiv:1304.0790 [hep-ph].
[25] M. Bonvini and S. Marzani, JHEP **09**, 007 (2014), arXiv:1405.3654 [hep-ph].
[26] T. Schmidt and M. Spira, Phys. Rev. **D93**, 014022 (2016), arXiv:1509.00195 [hep-ph].
[27] A. A H, A. Chakraborty, G. Das, P. Mukherjee, and V. Ravindran, (2019), arXiv:1905.03771 [hep-ph].
[28] D. Bonocore, E. Laenen, L. Magnea, L. Vernazza, and C. D. White, JHEP **12**, 121 (2016), arXiv:1610.06842 [hep-ph].
[29] V. Del Duca, E. Laenen, L. Magnea, L. Vernazza, and C. D. White, JHEP **11**, 057 (2017), arXiv:1706.04018 [hep-ph].
[30] M. Beneke, A. Broggio, M. Garny, S. Jaskiewicz, R. Szafron, L. Vernazza, and J. Wang, JHEP **03**, 043 (2019), arXiv:1809.10631 [hep-ph].
[31] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, and B. Mistlberger, Phys. Lett. **B737**, 325 (2014), arXiv:1403.4616 [hep-ph].
[32] T. Ahmed, M. Mahakhud, N. Rana, and V. Ravindran, Phys. Rev. Lett. **113**, 112002 (2014), arXiv:1404.0366 [hep-ph].
[33] T. Ahmed, M. K. Mandal, N. Rana, and V. Ravindran, Phys. Rev. Lett. **113**, 212003 (2014), arXiv:1404.6504 [hep-ph].
[34] D. de Florian, J. Mazzitelli, S. Moch, and A. Vogt, JHEP **10**, 176 (2014), arXiv:1408.6277 [hep-ph].
[35] Y. Li, A. von Manteuffel, R. M. Schabinger, and H. X. Zhu, Phys. Rev. **D91**, 036008 (2015), arXiv:1412.2771 [hep-ph].
[36] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, JHEP **05**, 058 (2016), arXiv:1602.00695 [hep-ph].
[37] F. Dulat, B. Mistlberger, and A. Pelloni, JHEP **01**, 145 (2018), arXiv:1710.03016 [hep-ph].
[38] F. Dulat, B. Mistlberger, and A. Pelloni, Phys. Rev. **D99**, 034004 (2019), arXiv:1810.09462 [hep-ph].
[39] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. **D81**, 094035 (2010), arXiv:0910.0467 [hep-ph].
[40] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. Lett. **105**, 092002 (2010), arXiv:1004.2489 [hep-ph].
[41] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **09**, 005 (2010), arXiv:1002.2213 [hep-ph].
[42] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **04**, 092 (2011), arXiv:1012.4480 [hep-ph].
[43] J. R. Gaunt, M. Stahlhofen, and F. J. Tackmann, JHEP **04**, 113 (2014), arXiv:1401.5478 [hep-ph].
[44] J. Gaunt, M. Stahlhofen, and F. J. Tackmann, JHEP **08**, 020 (2014), arXiv:1405.1044 [hep-ph].
[45] K. Melnikov, R. Rietkerk, L. Tancredi, and C. Wever, JHEP **02**, 159 (2019), arXiv:1809.06300 [hep-ph].
[46] K. Melnikov, R. Rietkerk, L. Tancredi, and C. Wever, JHEP **06**, 033 (2019), arXiv:1904.02433 [hep-ph].
[47] C. W. Bauer, S. Fleming, and M. E. Luke, Phys. Rev. **D63**, 014006 (2000), arXiv:hep-ph/0005275.
[48] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, Phys. Rev. **D63**, 114020 (2001), arXiv:hep-ph/0011336.
[49] C. W. Bauer and I. W. Stewart, Phys. Lett. **B516**, 134 (2001), arXiv:hep-ph/0107001.
[50] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. **D65**, 054022 (2002), arXiv:hep-ph/0109045.
[51] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. **D66**, 014017 (2002), arXiv:hep-ph/0202088.
[52] See Supplemental Material at the end of this preprint.
[53] S. Fleming and O. Z. Labun, Phys. Rev. **D91**, 094011 (2015), arXiv:1210.1508 [hep-ph].
[54] A. H. Hoang, P. Pietrulewicz, and D. Samitz, Phys. Rev. **D93**, 034034 (2016), arXiv:1508.04323 [hep-ph].
[55] A. Jain, M. Procura, and W. J. Waalewijn, JHEP **04**, 132 (2012), arXiv:1110.0839 [hep-ph].
[56] J. R. Gaunt and M. Stahlhofen, JHEP **12**, 146 (2014), arXiv:1409.8281 [hep-ph].
[57] M. Procura, W. J. Waalewijn, and L. Zeune, JHEP **02**, 117 (2015), arXiv:1410.6483 [hep-ph].
[58] G. Lustermans, J. K. L. Michel, F. J. Tackmann, and

W. J. Waalewijn, JHEP **03**, 124 (2019), arXiv:1901.03331 [hep-ph].
[59] A. V. Manohar, Phys. Rev. **D68**, 114019 (2003), arXiv:hep-ph/0309176.
[60] T. Becher, M. Neubert, and B. D. Pecjak, JHEP **01**, 076 (2007), arXiv:hep-ph/0607228.
[61] J. Chay and C. Kim, JHEP **09**, 126 (2013), arXiv:1303.1637 [hep-ph].
[62] D. Kang, C. Lee, and I. W. Stewart, Phys. Rev. **D88**, 054004 (2013), arXiv:1303.6952 [hep-ph].
[63] C. Anastasiou, L. J. Dixon, and K. Melnikov, Nucl. Phys. Proc. Suppl. **116**, 193 (2003), hep-ph/0211141.
[64] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. Lett. **91**, 182002 (2003), arXiv:hep-ph/0306192.
[65] J. Kubar, M. Le Bellac, J. L. Meunier, and G. Plaut, Nucl. Phys. **B175**, 251 (1980).
[66] P. Mathews, V. Ravindran, K. Sridhar, and W. L. van Neerven, Nucl. Phys. **B713**, 333 (2005), arXiv:hep-ph/0411018.
[67] M. A. Ebert, J. K. L. Michel, F. J. Tackmann, *et al.*, DESY-17-099 (2018), webpage: `http://scetlib.desy.de`.
[68] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. **D69**, 094008 (2004), arXiv:hep-ph/0312266.
[69] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Eur. Phys. J. **C75**, 204 (2015), arXiv:1412.3989 [hep-ph].
[70] G. Billis, M. A. Ebert, J. K. L. Michel, and F. J. Tackmann, (in preparation).
[71] C. W. Bauer, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Phys. Rev. **D85**, 074006 (2012), arXiv:1106.6047 [hep-ph].
[72] A. V. Manohar, T. Mehen, D. Pirjol, and I. W. Stewart, Phys. Lett. **B539**, 59 (2002), arXiv:hep-ph/0204229.
[73] R. V. Harlander, S. Liebler, and H. Mantler, Comput. Phys. Commun. **184**, 1605 (2013), arXiv:1212.3249 [hep-ph].
[74] R. V. Harlander, S. Liebler, and H. Mantler, Comput. Phys. Commun. **212**, 239 (2017), arXiv:1605.03190 [hep-ph].
[75] E. Laenen and G. F. Sterman, in *The Fermilab Meeting DPF 92. Proceedings, 7th Meeting of the American Physical Society, Division of Particles and Fields, Batavia, USA, November 10-14, 1992. Vol. 1, 2* (1992) pp. 987–989.
[76] M. A. Ebert, J. K. L. Michel, and F. J. Tackmann, JHEP **05**, 088 (2017), arXiv:1702.00794 [hep-ph].
[77] M. Höschele, J. Hoff, A. Pak, M. Steinhauser, and T. Ueda, Comput. Phys. Commun. **185**, 528 (2014), arXiv:1307.6925 [hep-ph].
[78] G. P. Korchemsky and A. V. Radyushkin, Nucl. Phys. **B283**, 342 (1987).
[79] C. Duhr and F. Dulat, (2019), arXiv:1904.07279 [hep-th].

| Mode | Lab frame $(+,-,\perp)$ | Leptonic ($\hat{Y}=0$) frame $(+,-,\perp)$ |
|---|---|---|
| $p_{\bar{n}}$ | $(q^+, \lambda^2 q^-, \lambda\sqrt{q^- q^+})$ | $Q\,(1, \lambda^2, \lambda)$ |
| $P_{\bar{n}}$ | $\Big(q^+, \frac{\Lambda^2_{\rm QCD}}{q^+}, \Lambda_{\rm QCD}\Big)$ | $Q\,(1, \lambda^2_{\rm QCD}, \lambda_{\rm QCD})$ |
| $P_n$ | $\Big(\frac{\Lambda^2_{\rm QCD}}{q^-}, q^-, \Lambda_{\rm QCD}\Big)$ | $Q\,(\lambda^2_{\rm QCD}, 1, \lambda_{\rm QCD})$ |
| $P_s$ | $\Big(\frac{1}{\lambda^2}\,\frac{\Lambda^2_{\rm QCD}}{q^-}, \lambda^2 q^-, \Lambda_{\rm QCD}\Big)$ | $Q\,\Big(\frac{\lambda^2_{\rm QCD}}{\lambda^2}, \lambda^2, \lambda_{\rm QCD}\Big)$ |
| $P_{us}$ | $\Big(\frac{\Lambda^2_{\rm QCD}}{q^-}, \frac{\Lambda^2_{\rm QCD}}{q^+}, \frac{\Lambda^2_{\rm QCD}}{\sqrt{q^+ q^-}}\Big)$ | $Q\,(\lambda^2_{\rm QCD}, \lambda^2_{\rm QCD}, \lambda^2_{\rm QCD})$ |
| $P_G$ | $\Big(\frac{\Lambda^2_{\rm QCD}}{q^-}, \frac{\Lambda^2_{\rm QCD}}{q^+}, \Lambda_{\rm QCD}\Big)$ | $Q\,(\lambda^2_{\rm QCD}, \lambda^2_{\rm QCD}, \lambda_{\rm QCD})$ |

TABLE S1. Relevant EFT modes in the limit $\lambda_{\rm QCD} \sim \lambda^2 \sim 1 - x^- \ll 1$ in the lab (hadronic center-of-mass) frame and the leptonic frame where $\hat{Y} = 0$. In the right column we used that in the leptonic frame, $q^\pm \to \hat{q}^\pm = \sqrt{q^+ q^-} \sim Q$.

## SUPPLEMENTAL MATERIAL

### A. Factorization theorem

Here we give some more details on the derivation of the factorization theorem in Eq. (12), which underlies all other factorization theorems, and which we repeat here for easy reference,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_-\,\mathrm{d}x_+\,\mathrm{d}\vec{q}_T} = H_{ij}(q^+q^-)\int \mathrm{d}t\, f_i^{\rm thr}\Big[x_-\Big(1+\frac{t}{q^+q^-}\Big)\Big] B_j(t,\vec{q}_T,x_+)\,. \tag{S1}$$

As in the main text, we always implicitly sum over repeated flavor indices $i, j, k$. Equation (S1) is valid up to power corrections in $\lambda^2$ in the generalized threshold limit

$$\lambda^2 \sim 1 - x_- \ll 1 \qquad \text{for generic } x_+\,, \qquad \lambda_{\rm QCD} \equiv \Lambda_{\rm QCD}/Q \sim \lambda^2 \ll \lambda\,. \tag{S2}$$

We require $\lambda_{\rm QCD} \ll \lambda$ for reasons that will be apparent soon. Without loss of generality we can then consider $\lambda_{\rm QCD} \sim \lambda^2$. This relation is to be interpreted as follows: First, in our context, $\lambda_{\rm QCD}$ denotes the scale of the PDFs, which is generically allowed to be as large as $\lambda^2$ and does not necessarily have to be nonperturbative. If it happens to be a perturbative scale, then the physics below $\lambda_{\rm QCD}$ is simply described by the PDF evolution. Conversely, it also means that $\lambda^2$ is in principle allowed to be as small as $\lambda_{\rm QCD}$ including being nonperturbative, i.e., it is only relevant that $\lambda \gg \lambda_{\rm QCD}$ is perturbative.

The key step in deriving Eq. (S1) is to identify the relevant degrees of freedom (modes) in the effective field theory (EFT) that describe the physical situation. They are defined by the relative scaling of their light-cone momentum components and are summarized in Table S1. We note that rather than matching QCD directly onto these modes, one may also perform a multi-stage matching, as was done for endpoint DIS in Ref. [54], which yields the same end result.

The $p_{\bar{n}}$ modes describe the hadronic final state of the collision. Their scaling is determined by the fact that in the limit of Eq. (S2), there is only $p_{\bar{n}}^- \sim \lambda^2 E_{\rm cm} \sim \lambda^2 q^-$ minus momentum available. On the other hand, their plus momentum is unconstrained, which means it has generic scaling set by the hard interaction, $p_{\bar{n}}^+ \sim \xi_b E_{\rm cm} \sim x_b E_{\rm cm} \sim q^+$. Since $p_{\bar{n}}^2 \sim \lambda^2 q^+ q^- \sim \lambda^2 Q^2 \gg \Lambda^2_{\rm QCD}$, the $p_{\bar{n}}$ modes are perturbative. Therefore, they describe the perturbative QCD final state produced in the partonic collision in addition to $L$ (but excluding the beam remnant). The $P_n$ and $P_{\bar{n}}$ modes describe the incoming protons, or more precisely, the partons in the proton with the typical momentum fractions required to produce the hard final state. This means their scaling is determined by $P_n^- \sim q^-$ and $P_{\bar{n}}^+ \sim \xi_b E_{\rm cm} \sim x_b E_{\rm cm} \sim q^+$ and $P_n^2 \sim P_{\bar{n}}^2 \sim \Lambda^2_{\rm QCD}$.

The soft $P_s$ modes describe the interactions between the $p_{\bar{n}}$ and $P_n$ modes. Their scaling is thus determined by $P_s^- \sim p_{\bar{n}}^- \sim \lambda^2 Q$ and $P_{s\perp} \sim P_{n\perp} \sim \Lambda_{\rm QCD}$ or equivalently $P_s^2 \sim P_n^2 \sim \Lambda^2_{\rm QCD}$. Hence, they keep the $p_{\bar{n}}$ modes on shell and have a $\mathrm{SCET_I}$-like relation to them. Their interactions with the $p_{\bar{n}}$ modes in the leading-power SCET Lagrangian

are decoupled and moved into soft Wilson lines in the SCET current via the BPS field redefinition. At the same time, the $P_s$ modes have a SCET$_{\rm II}$-like relation to the $P_n$ modes, i.e., they have the same virtuality but are parametrically separated in rapidity. Hence, their interactions with the $P_n$ modes, which take the $P_n$ modes off shell, are described by soft Wilson lines in the SCET current that are directly produced during the matching onto SCET. The distinction of the $P_s$ modes relies on $\lambda_{\rm QCD} \ll \lambda$, while for $\lambda_{\rm QCD} \sim \lambda$, they would become degenerate with the $p_{\bar n}$ and $P_{\bar n}$ modes.

The power counting and the relations between the modes are easiest in the leptonic frame, which is the frame where the color-singlet final state has total rapidity $\hat{Y} = 0$. Boosting from the lab frame to the leptonic frame by $Y$, we have $\hat{q}^\pm = q^\pm e^{\pm Y} = \sqrt{q^+ q^-} \sim Q$. In the leptonic frame, the $p_{\bar n}$ modes are genuinely $\bar n$-collinear with $p_{\bar n}^- \sim \lambda^2 p_{\bar n}^+$, and the soft modes are homogeneous, $P_s \sim \lambda^2 Q \sim \lambda_{\rm QCD} Q$. By contrast, in the lab frame we must separately keep track of $q^+$ and $q^-$, i.e., we cannot count them as $q^+ \sim q^-$, because we want to take the limit of large $q^-$ for generic $q^+$. As a result, the $p_{\bar n}$ modes do not necessarily appear to be $\bar n$-collinear in the lab frame because $q^+ \sim \lambda^2 q^-$ is allowed. However, the key requirement for their factorization is that they are collinear *relative* to the soft modes, which in the lab frame are boosted in the $n$-collinear direction and become $n$-collinear-soft (csoft) modes [57, 71].

Finally, the ultrasoft (usoft) $P_{us}$ and Glauber $P_G$ modes describe the interactions between the $P_n$ and $P_{\bar n}$ modes that are possible without pushing either of them off shell, which requires $P_{us}^- \sim P_{\bar n}^-$ and $P_{us}^+ \sim P_n^+$. The $\perp$ component of the usoft modes is fixed by requiring them to be on-shell modes, $P_{us\perp}^2 \sim P_{us}^+ P_{us}^-$. The corresponding Glauber modes are allowed to be off shell, so their $\perp$ component can be as large as $P_{G\perp} \sim P_{n\perp} \sim P_{\bar n\perp} \sim \Lambda_{\rm QCD}$. The effects of the usoft and Glauber modes cancel, so we do not need to consider them further. This directly follows from the collinear factorization theorem [1–3] because in the limit we consider, the measurement is still fully inclusive over perpendicular momenta at the scale $\lambda_{\rm QCD} Q$. This is another reason why we require $\lambda_{\rm QCD} \ll \lambda$. Note also that there is only a single collinear sector at the perturbative $\lambda Q$ scale, so there are no perturbative Glauber modes with scaling $Q(\lambda^2, \lambda^2, \lambda)$ that could spoil factorization.

Since there are no interactions between the modes in the leading-power SCET Lagrangian, the cross section factorizes into separate forward matrix elements in each sector. The detailed derivation closely follows Ref. [39], with the matrix element of the combined $p_{\bar n}$ and $P_{\bar n}$ modes giving $B_j(t, x)$ and the combined matrix element of the $P_n$ and $P_s$ modes giving $f_i^{\rm thr}$. The factorization of the threshold PDF into separate collinear and csoft nonperturbative matrix elements is discussed in Refs. [53, 54]. It is not needed for our purposes. The arguments of the functions and their convolution structure in Eq. (S1) follow from overall momentum conservation among all sectors, which at leading power in $\lambda$ must hold separately for the large (label) and small (residual) momenta carried by the matrix elements. Denoting them as $\omega_{n,\bar n} \sim Q$, $k_{\bar n\perp} \sim \lambda Q$, and $k_{s,\bar n} \sim \lambda^2 Q$, it takes the form

$$\delta(\omega_{\bar n} - q^+)\, \delta(\omega_n - q^-)\, \delta(\vec{k}_{\bar n\perp} + \vec{q}_T)\, \delta(k_s^- - k_{\bar n}^-)\,. \tag{S3}$$

The first three $\delta$ functions set the $x_\pm$ and $\vec{q}_T$ arguments of $f_i^{\rm thr}$ and $B_j$. The analogous $\delta(k_s^+ - k_{\bar n}^+)$ disappeared by absorbing the $k_{\bar n}^+$ dependence into $\omega_{\bar n}$ [39]. By combining the $P_n$ and $P_s$ modes, the threshold PDF depends on $(\omega_n + k_s^-)/P_a^- = x_-(1 + k_s^-/q^-)$. Here, we do have to keep track of the momentum of the $P_s$ modes $k_s^- \sim \lambda^2 Q$, as it is much larger than the typical residual minus momentum $k_n^- \sim \lambda_{\rm QCD}^2 Q$ of the $P_n$ modes that is absorbed into $\omega_n$. The $\delta(k_s^- - k_{\bar n}^-)$ then yields the convolution in $t = q^+ k_{\bar n}^-$ in Eq. (S1). While Eq. (S1) is formally derived in the leptonic frame, it has exactly the same form in the lab frame. This is because the measured observables $x_\pm$ and $\vec{q}_T$ are boost invariant along the beam axis, and reparametrization invariance (RPI) [72] forces all functions to only depend on boost-invariant quantities.

To see that the generalized threshold limit contains the soft threshold limit, note that the limits $x_- \to 1$ and $x_+ \to 1$ commute, so taking one limit after the other is equivalent to taking $x_-, x_+ \to 1$ simultaneously. To see this, consider the hierarchy $\lambda_-^2 \sim 1 - x_- \ll \lambda_+^2 \sim 1 - x_+ \ll 1$, which we can interpret as taking $x_+ \to 1$ after having already taken $x_- \to 1$. In this limit, the $p_{\bar n}$ modes factorize into perturbative $\bar n$-csoft modes $p_{\bar n,cs} \sim Q(\lambda_+^2, \lambda_-^2, \lambda_+\lambda_-)$ and $\bar n$-csoft modes $P_{\bar n,cs} \sim Q(\lambda_+^2, \lambda_{\rm QCD}^2/\lambda_+^2, \lambda_{\rm QCD})$. Including $\lambda_+$, the condition $\lambda_{\rm QCD} \ll \lambda$ becomes $\lambda_{\rm QCD} \ll \lambda_-\lambda_+$ and without loss of generality we can consider $\lambda_{\rm QCD} \sim \lambda_-^2 \lambda_+^2$. This also implies that the $P_s$ modes now get boosted in the $n$ direction and become $n$-csoft modes, $P_s \equiv P_{n,cs}$. The beam function matching onto PDFs now takes the form

$$B_j(\omega_{\bar n} k^-, \vec{k}_T, x_+) = \int \frac{{\rm d}k^+}{\omega_{\bar n}}\, \mathcal{S}(k^-, k^+, \vec{k}_T)\, f_j^{\rm thr}\Big[x_+\Big(1 + \frac{k^+}{\omega_{\bar n}}\Big)\Big]\, \Big[1 + \mathcal{O}(\lambda_+)\Big]\,, \tag{S4}$$

where the combined $P_{\bar n,cs}$ and $P_{\bar n}$ modes yield the threshold PDF, and $\mathcal{S}(k^-, k^+, \vec{k}_T)$ is the matrix element of the $p_{\bar n,cs}$ modes. It has the same Wilson line structure as the soft function appearing in Eq. (4), and thus upon integration over $\vec{k}_T$ becomes equal to it to all orders by reparametrization invariance. We now have $k_T^2/Q^2 \sim \lambda_-^2\lambda_+^2 \ll t/Q^2 \sim \lambda_-^2$, so $B_j$ and $\tilde{B}_j$ become the same and integrating Eq. (S4) over $\vec{k}_T$ yields Eq. (16) for either of them.

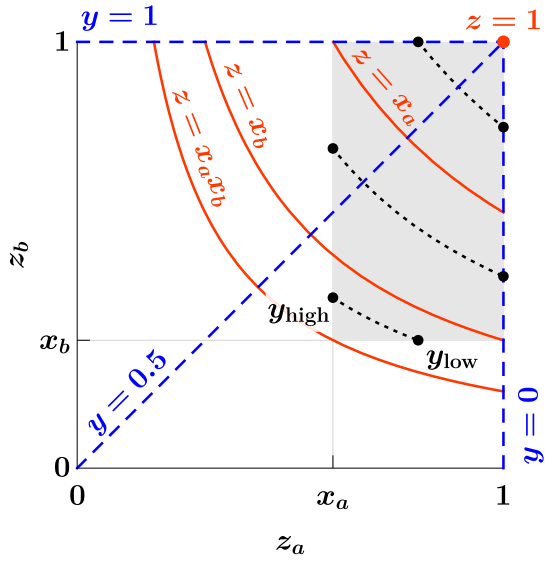


FIG. S1. The $(z_a, z_b)$ plane as parametrized by $(z, y)$. The gray area shows the integration region $x_a \le z_a \le 1$ and $x_b \le z_b \le 1$ (for the case of $x_a > x_b$) used to derive the distribution identities. The dotted black lines indicate integration paths over $y$ at representative fixed values of $z$. The solid red lines indicate the edge cases in $z$. The dashed blue lines are lines of constant $y$.

### B. Plus distribution identities

In Refs. [63, 64, 68], the partonic cross section for $(Q, Y)$ is given in terms of partonic variables $(z, y)$, defined as

$$z = z_a z_b\,, \qquad y = \frac{z_b(1-z_a^2)}{(1-z_a z_b)(z_a+z_b)}\,, \qquad 1-y = \frac{z_a(1-z_b^2)}{(1-z_a z_b)(z_a+z_b)} = y|_{a\leftrightarrow b}\,,$$
$$z_a = \sqrt{\frac{z[1-y(1-z)]}{z+y(1-z)}}\,, \qquad z_b = z_a|_{y\leftrightarrow 1-y}\,, \qquad \frac{\mathrm{d}z\,\mathrm{d}y}{\mathrm{d}z_a \mathrm{d}z_b} = \frac{2[1-y(1-z)][1-(1-y)(1-z)]}{1-z^2}\,, \tag{S5}$$

where $z_{a,b}$ are defined by Eq. (3), and the integration limits $0 \le z_{a,b} \le 1$ correspond to $0 \le z \le 1$ and $0 \le y \le 1$. In the following, we derive relations between plus distributions in $(z, y)$ and $(z_a, z_b)$.

In general, the plus distributions are uniquely defined by their functional form in the bulk, i.e. for $z_{a,b} < 1$ away from any singularity, and their integrals (against unit test functions) over arbitrary integration regions that include the singularities. First, the functional form in the bulk is easily obtained simply by plugging in Eq. (S5). Next, the correct boundary terms at $z_a = 1$ as a function of $z_b$, at $z_b = 1$ as a function of $z_a$, and at $z_a = z_b = 1$ are determined by comparing integrals over the integration region $x_a \le z_a(z,y) \le 1$ and $x_b \le z_b(z,y) \le 1$ for generic $x_a$ and $x_b$. This integration region is indicated by the gray box in the $(z_a, z_b)$ plane in Fig. S1. It is sufficiently general to fix all boundary terms and precisely corresponds to the relevant integration region for the physical cross section in $(Q, Y)$. In terms of $(z, y)$, the integration region is given by

$$\begin{aligned}\theta\big[z_a(z,y) \ge x_a\big]\,\theta\big[z_b(z,y) \ge x_b\big] &= \theta\big[x_a x_b \le z < \min\{x_a, x_b\}\big]\,\theta\big[y_{\rm low}(z) \le y \le y_{\rm high}(z)\big] + \theta\big[\max\{x_a, x_b\} \le z\big] \\ &\quad + \theta\big[x_b \le z < x_a\big]\,\theta\big[y \le y_{\rm high}(z)\big] + \theta\big(x_a \le z < x_b\big)\,\theta\big[y_{\rm low}(z) \le y\big]\,,\end{aligned} \tag{S6}$$

where the integration bounds in $y$, also illustrated in Fig. S1, are given by

$$y_{\rm high}(z) = \frac{z(1-x_a^2)}{(1-z)(x_a^2+z)}\,, \qquad y_{\rm low}(z) = \frac{x_b^2 - z^2}{(1-z)(z+x_b^2)}\,. \tag{S7}$$

In Table S2, we collect the distribution identities at leading power in $1-z_a$ and arbitrary $z_b$ that are required for validating Eq. (14) at NLO, where we denote the plus distributions as

$$\mathcal{L}_n(x) = \Big[\frac{\theta(x)\ln^n(x)}{x}\Big]_+ = \lim_{\epsilon\to 0}\frac{\mathrm{d}}{\mathrm{d}x}\Big[\theta(x-\epsilon)\frac{\ln^{n+1}x}{n+1}\Big] \quad \text{with} \quad \mathcal{L}_n(x>0) = \frac{\ln^n(x)}{x}\,, \qquad \int_0^1 \mathrm{d}x\,\mathcal{L}_n(x) = 0\,. \tag{S8}$$

| $\mathrm{d}z\,\mathrm{d}y\,\mathrm{LHS}$ | $=\mathrm{d}z_a\mathrm{d}z_b\,\mathrm{RHS}$ |
|---|---|
| $f(z)\,\delta(y)$ | $f(z_b)\,\delta(1-z_a)$ |
| $r(z)\,\mathcal{L}_0(y)$ | $r(z_b)\Big[\mathcal{L}_0(1-z_a)+\delta(1-z_a)\,\ln\frac{2z_b}{(1+z_b)(1-z_b)}\Big]+\mathcal{O}(1)$ |
| $f(z)\,\delta(1-y)$ | $f(z_a)\delta(1-z_b)$ |
| $r(z)\,\mathcal{L}_0(1-y)$ | $\mathcal{O}(1)$ |
| $r(z)\,\mathcal{O}[y^0(1-y)^0]$ | $\mathcal{O}(1)$ |
| $\mathcal{L}_0(1-z)\,\big[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)\big]$ | $-\mathcal{L}_1(1-z_a)\,\delta(1-z_b)+\mathcal{L}_0(1-z_a)\,\mathcal{L}_0(1-z_b)-\delta(1-z_a)\,\mathcal{L}_1(1-z_b)$ $+\frac{\pi^2}{6}\delta(1-z_a)\,\delta(1-z_b)+\delta(1-z_a)\,\frac{1}{1-z_b}\ln\frac{2z_b}{1+z_b}+\mathcal{O}(1)$ |

TABLE S2. Translation identities of two-dimensional plus distributions between the $(z,y)$ and $(z_a,z_b)$ parametrizations. Here, $f(x)$ is an arbitrary function of $x$, potentially distribution-valued for $x\to 1$, while $r(x)=\mathcal{O}[(1-x)^0]$ has at most an integrable singularity for $x\to 1$. When indicated, the relations receive power corrections in $1-z_a$ starting at $\mathcal{O}(1)\equiv\mathcal{O}[(1-z_a)^0]$. Overall factors of $\theta(1-z)\,\theta(y)\,\theta(1-y)=\theta(1-z_a)\,\theta(1-z_b)$ are understood on both sides.

The relations are derived by integrating each structure in the left column in terms of $(z,y)$ over the region in Eq. (S6), expanding the result to leading power in $1-x_a$, and comparing to the corresponding (straightforward) integral over the same region in terms of $z_{a,b}$ of each structure in the right column.

The last entry in Table S2 has the most intricate structure. Its exact integral without any expansion is given by

$$\begin{aligned}&\int\mathrm{d}z\mathrm{d}y\,\theta\big[z_a(z,y)-x_a\big]\,\theta\big[z_b(z,y)-x_b\big]\,\mathcal{L}_0(1-z)\Big[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)\Big]\\&\qquad=F_a(x_a)+F_b(x_b)-F_a(x_ax_b)-F_b(x_ax_b)\,,\\\text{with}\qquad&F_{a,b}(z)=-\ln(1-z)\ln(-z+\mathrm{i}0)-\mathrm{Li}_2\Big(\frac{1-z}{1+x_{a,b}}\Big)-\mathrm{Li}_2\Big(\frac{1-z}{1-x_{a,b}}-\mathrm{i}0\Big)-\mathrm{Li}_2(z)\,,\end{aligned}\tag{S9}$$

where the imaginary parts from the branch cuts cancel between the different terms. Matching this with the exact distribution in the bulk, we obtain the distributional identity

$$\begin{aligned}&\mathrm{d}z\,\mathrm{d}y\,\mathcal{L}_0(1-z)\,\big[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)\big]\\&\quad=\mathrm{d}z_a\mathrm{d}z_b\bigg[\frac{\pi^2}{6}\,\delta(1-z_a)\,\delta(1-z_b)-\mathcal{L}_1(1-z_a)\,\delta(1-z_b)+\mathcal{L}_0(1-z_a)\,\mathcal{L}_0(1-z_b)-\delta(1-z_a)\,\mathcal{L}_1(1-z_b)\\&\qquad+\delta(1-z_a)\,\frac{1}{1-z_b}\ln\frac{2z_b}{1+z_b}+\delta(1-z_b)\,\frac{1}{1-z_a}\ln\frac{2z_a}{1+z_a}+\frac{1}{(1+z_a)(1+z_b)}\bigg]\,.\end{aligned}\tag{S10}$$

Expanding the right-hand side to leading power in $1-z_a$ yields the result given in the last line of Table S2. We note that Eq. (S10) is the two-dimensional analog of a typical distribution identity like

$$\Big[\frac{1+z^2}{1-z}\Big]_+=2\mathcal{L}_0(1-z)+\frac{3}{2}\delta(1-z)-(1+z)\,.\tag{S11}$$

Namely, it expresses plus distributions of a function $y(z_a,z_b)$ in terms of simpler plus distributions of $1-z_{a,b}$ plus regular terms. Moving the regular terms out of the plus distribution incurs additional boundary terms.

A key property of the left-hand side of Eq. (S10) is that it vanishes when integrated over all of $y$. It is instructive to see how this is reproduced by the right-hand side by projecting onto $z$ via $\int\mathrm{d}z_a\mathrm{d}z_b\,\delta(z-z_az_b)$. The only nontrivial projection integral that is required is

$$\int\mathrm{d}z_a\mathrm{d}z_b\,\delta(z-z_az_b)\,\mathcal{L}_0(1-z_a)\mathcal{L}_0(1-z_b)=2\mathcal{L}_1(1-z)-\frac{\pi^2}{6}\delta(1-z)-\frac{\ln z}{1-z}\,.\tag{S12}$$

These terms precisely cancel the contributions from the other terms upon projecting onto $z$.

### C. Exact NLO partonic cross sections in terms of $(z_a, z_b)$

Since Eq. (18) captures the full singularity structure as $z_a\to 1$ and/or $z_b\to 1$, the power corrections to it are of relative $\mathcal{O}[(1-z_a)(1-z_b)]$ and are fully integrable. Hence, we can construct the exact partonic cross section in terms

of $(z_a, z_b)$ from the results in terms of $(z, y)$ from Refs. [63, 64] as

$$\hat{\sigma}_{ij}(z_a,z_b) = H_{ik}\hat{\mathcal{I}}_{kj}(z_a,z_b) + H_{kj}\hat{\mathcal{I}}_{ki}(z_b,z_a) - H_{ij}\hat{S}(z_a,z_b) + \left[\frac{\mathrm{d}z\,\mathrm{d}y}{\mathrm{d}z_a\mathrm{d}z_b}\,\hat{\sigma}_{ij}\big[z(z_a,z_b),y(z_a,z_b)\big] - H_{ik}\hat{\mathcal{I}}_{kj}(z_a,z_b) - H_{kj}\hat{\mathcal{I}}_{ki}(z_b,z_a) + H_{ij}\hat{S}(z_a,z_b)\right]_{z_{a,b}<1}. \tag{S13}$$

Here, the term in square brackets is evaluated in the bulk, away from any singularities, so we can simply plug in Eq. (S5). In the following, we collect the resulting expressions for the Drell-Yan and $gg \to H$ cross sections to NLO written explicitly in terms of $(z_a, z_b)$.

### 1. *Results for Drell-Yan*

The Born cross section for Drell-Yan production, $q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-$, is given by

$$\sigma^{\mathrm{DY}}_{B,q} = \frac{4\pi\alpha^2_{\mathrm{em}}}{3N_c Q^2}\left[Q_q^2 + \frac{(v_q^2+a_q^2)(v_\ell^2+a_\ell^2) - 2Q_q v_q v_\ell (1-m_Z^2/Q^2)}{(1-m_Z^2/Q^2)^2 + m_Z^2\Gamma_Z^2/Q^4}\right], \tag{S14}$$

with $N_c$ the number of colors, $Q_q$ the quark charge in units of $|e|$, $v_{\ell,q}$ and $a_{\ell,q}$ the standard electroweak vector and axial couplings of the leptons and quarks, and $m_Z$ and $\Gamma_Z$ the mass and width of the $Z$ boson. To restrict to $q\bar{q} \to \gamma^* \to \ell^+\ell^-$, only the first term $\propto Q_q^2$ is kept. The complete LO cross section is given by

$$\frac{\mathrm{d}\sigma^{\mathrm{DY}}_{\mathrm{LO}}}{\mathrm{d}Q\,\mathrm{d}Y} = \frac{2Q}{E^2_{\mathrm{cm}}}\frac{\mathrm{d}\sigma^{\mathrm{DY}}_{\mathrm{LO}}}{\mathrm{d}x_a\mathrm{d}x_b} = \frac{2Q}{E^2_{\mathrm{cm}}}\sum_q \sigma^{\mathrm{DY}}_{B,q}\left[f_q(x_a)\,f_{\bar{q}}(x_b) + f_{\bar{q}}(x_a)\,f_q(x_b)\right], \tag{S15}$$

where the sum runs over $q = \{u, d, c, s, b\}$. We expand the partonic cross section for Drell-Yan as

$$\hat{\sigma}_{ij}(z_a,z_b,Q,\mu) = \sum_{n=0}^{\infty}\Big[\frac{\alpha_s(\mu)}{4\pi}\Big]^n \sigma^{(n)}_{ij}(z_a,z_b,Q,\mu)\,. \tag{S16}$$

The LO result corresponding to Eq. (S15) is given by $\hat{\sigma}^{(0)}_{q\bar{q}}(z_a,z_b,Q,\mu) = \sigma^{\mathrm{DY}}_{B,q}\,\delta(1-z_a)\,\delta(1-z_b)$. Writing $\bar{z}_a \equiv 1-z_a$ and $\bar{z}_b \equiv 1-z_b$ for short, the NLO coefficient for the $q\bar{q}$ channel is given by

$$\begin{aligned}\frac{1}{\sigma^{\mathrm{DY}}_{B,q}}\frac{\hat{\sigma}^{(1)}_{q\bar{q}}(z_a,z_b,Q,\mu)}{C_F} &= \delta(\bar{z}_a)\delta(\bar{z}_b)(2\pi^2-16) + 4\mathcal{L}_1(\bar{z}_a)\delta(\bar{z}_b) + 4\mathcal{L}_0(\bar{z}_a)\mathcal{L}_0(\bar{z}_b) + 4\delta(\bar{z}_a)\mathcal{L}_1(\bar{z}_b)\\ &+ \Big\{-2(1+z_b)\mathcal{L}_0(\bar{z}_a) + \delta(\bar{z}_a)\Big[2\bar{z}_b - 4(1+z_b)\ln\bar{z}_b - \frac{2(1+z_b^2)\ln z_b}{\bar{z}_b}\\ &+ \frac{4}{\bar{z}_b}\ln\frac{2z_b}{1+z_b} + 2(1+z_b)\ln\frac{1-z_b^2}{2z_b}\Big] - 4\ln\frac{\mu}{Q}\,\delta(\bar{z}_a)\Big[2\mathcal{L}_0(\bar{z}_b) + \frac{3}{2}\delta(\bar{z}_b) - (1+z_b)\Big]\\ &+ \frac{2(z_a^2+z_b^2)[(1+z_a)^2 + z_a z_b(3+2z_a+z_a z_b)]}{(1+z_a)(1+z_b)(z_a+z_b)^2} + (z_a \leftrightarrow z_b)\Big\},\end{aligned} \tag{S17}$$

where $(z_a \leftrightarrow z_b)$ indicates all previous expressions in the curly brackets repeated with $z_a$ and $z_b$ interchanged. For the $qg$ channel we have

$$\begin{aligned}\frac{1}{\sigma^{\mathrm{DY}}_{B,q}}\frac{\hat{\sigma}^{(1)}_{qg}(z_a,z_b,Q,\mu)}{T_F} &= 2(z_b^2+\bar{z}_b^2)\mathcal{L}_0(\bar{z}_a) + \delta(\bar{z}_a)\Big[2(z_b^2+\bar{z}_b^2)\ln\frac{2\bar{z}_b}{1+z_b} + 4z_b\bar{z}_b\Big] - 4\ln\frac{\mu}{Q}\,\delta(\bar{z}_a)\big(z_b^2+\bar{z}_b^2\big)\\ &+ \frac{1}{(1+z_a)(z_a+z_b)^3}\Big[-4z_a^5z_b^3 - 4z_a^4z_b^2(-1+z_b+2z_b^2) + 2z_a^3(1+4z_b^2+2z_b^3-4z_b^4-4z_b^5)\\ &+ 2z_a^2z_b(1+4z_b+8z_b^2-8z_b^3-4z_b^4) + 2z_az_b^2(1+4z_b-2z_b^2-4z_b^3) - 2z_b^3(1-2z_b+2z_b^2)\Big].\end{aligned} \tag{S18}$$

The $gq$ channel is given by $\sigma^{(1)}_{gq}(Q,\mu,z_a,z_b) = \sigma^{(1)}_{qg}(Q,\mu,z_b,z_a)$. The results for $q \leftrightarrow \bar{q}$ are identical.

### 2. Results for gluon-fusion Higgs production

For gluon-fusion Higgs production, $gg \to H$, we use the effective Lagrangian in the limit $m_H^2 \ll 4m_t^2$,

$$\mathcal{L}_{\text{eff}}(m_H) = -\frac{C_t}{12\pi v}\,\alpha_s\,G^a_{\mu\nu}G^{a,\mu\nu}H\,, \qquad C_t = 1 + \frac{\alpha_s}{4\pi}(5C_A - 3C_F) + \mathcal{O}(\alpha_s^2)\,. \tag{S19}$$

As in Ref. [63], the Wilson coefficient $C_t$ is always perturbatively expanded against other fixed-order contributions. The Born cross section and LO rapidity spectrum are given by

$$\sigma_B^{ggH} = \frac{1}{72\pi v^2(N_c^2-1)}\,, \qquad \frac{\mathrm{d}\sigma_{\text{LO}}^{ggH}}{\mathrm{d}Y} = x_a x_b\,\sigma_B^{ggH}\,\alpha_s^2\,f_g(x_a)\,f_g(x_b)\,. \tag{S20}$$

We write the partonic cross section as

$$\hat\sigma_{ij}(z_a,z_b,m_t,m_H,\mu) \equiv \sigma_B^{ggH}\,\alpha_s^2(\mu)\,|C_t(m_t,\mu)|^2\sum_{n=0}^{\infty}\Big[\frac{\alpha_s(\mu)}{4\pi}\Big]^n\hat\eta_{ij}^{(n)}(z_a,z_b,m_H,\mu)\,. \tag{S21}$$

The NLO coefficient function for the $gg$ channel is given by

$$\begin{aligned}
\frac{\hat\eta_{gg}^{(1)}(z_a,z_b,m_H,\mu)}{C_A} &= 2\pi^2\delta(\bar z_a)\delta(\bar z_b) + 4\mathcal{L}_1(\bar z_a)\delta(\bar z_b) + 4\delta(\bar z_a)\mathcal{L}_1(\bar z_b) + 4\mathcal{L}_0(\bar z_a)\mathcal{L}_0(\bar z_b) \\
&\quad + \bigg\{4\mathcal{L}_0(\bar z_a)\Big[\frac{1}{z_b} - 2 + z_b - z_b^2\Big] + 4\delta(\bar z_a)\frac{1}{\bar z_b}\Big[\frac{\ln(2\bar z_b)}{z_b} - 3\ln\bar z_b + 2\ln\frac{1+z_b}{2} - \frac{\ln(1+z_b)}{z_b} \\
&\qquad - z_b(3 - 2z_b + z_b^2)\ln\frac{1+z_b}{2\bar z_b}\Big] - 4\ln\frac{\mu}{m_H}\delta(\bar z_a)\Big[2\mathcal{L}_0(\bar z_b)\frac{(1-z_b+z_b^2)^2}{z_b}\Big] \\
&\qquad + \frac{4z_b}{z_a(1+z_a)(1+z_b)(z_a+z_b)^4}\Big[2z_b^2 + z_b^3 + 3z_a^6z_b^4 + 2z_a^5z_b^3(5+5z_b+2z_b^2) \\
&\qquad + z_a^4z_b^2(16 + 17z_b + 12z_b^2 + 6z_b^3 + 2z_b^4) + z_a^3z_b(5 + 22z_b + 12z_b^2 + 8z_b^3 + 8z_b^4 + 2z_b^5) \\
&\qquad + z_a^2(3 + 2z_b^2 + 7z_b^3 + 2z_b^4 + 4z_b^5 + 2z_b^6) + z_az_b(4 + z_b + z_b^2 + z_b^3 + z_b^5)\Big] + (z_a \leftrightarrow z_b)\bigg\}\,.
\end{aligned} \tag{S22}$$

For the $gq$ and $qg$ channels we find, with $\eta_{qg}^{(1)}(z_a,z_b,m_H,\mu) = \eta_{gq}^{(1)}(z_b,z_a,m_H,\mu)$,

$$\begin{aligned}
\frac{\hat\eta_{gq}^{(1)}(z_a,z_b,m_H,\mu)}{C_F} &= 2\mathcal{L}_0(\bar z_a)\,\frac{2-2z_b+z_b^2}{z_b} + 2\delta(\bar z_a)\Big[z_b + \frac{2-2z_b+z_b^2}{z_b}\Big(\ln\frac{2\bar z_b}{1+z_b} - 2\ln\frac{\mu}{m_H}\Big)\Big] \\
&\quad + \frac{2}{(1+z_a)z_b(z_a+z_b)^3}\Big[z_a^3(2-2z_b+z_b^2) + z_a^2(4 - 2z_b - 4z_b^2 + 7z_b^3 - 2z_b^5) \\
&\quad + z_az_b(4 - 4z_b + 4z_b^2 + z_b^3 - 2z_b^4) - z_b^2(-2 + 2z_b - 2z_b^2 + z_b^3)\Big]\,.
\end{aligned} \tag{S23}$$

The $q\bar q$ channel is fully regular and given by

$$\frac{\hat\eta_{q\bar q}^{(1)}(z_a,z_b,m_H,\mu)}{C_F} = \frac{N_c^2-1}{N_c}\,\frac{4(1+z_az_b)(z_a^4z_b^2 + z_a^2z_b^4 - 4z_a^2z_b^2 + z_a^2 + z_b^2)}{(z_a+z_b)^4}\,. \tag{S24}$$

The $N_c$-dependent prefactor accounts for the different color average compared to the Born cross section.

The above expressions are in full agreement with Refs. [6, 66], which in turn agree with the earliest result for the NLO Drell-Yan rapidity spectrum in Ref. [65]. In Refs. [6, 66], the cross section was also parametrized in terms of $x_{a,b}$ and $z_{a,b}$, but all subtractions were written out in full at the level of the hadronic cross section.

In Ref. [64], only the sum of the $gq$ and $qg$ coefficient functions for Higgs productions was given. The separation of the singular terms into the two channels is unique because only the $gq$ ($qg$) channel can be singular as $y \to 0$ ($y \to 1$). We determined the separation of the regular terms by comparing to Ref. [6]. To the best of our knowledge, this is the first time that the explicit analytic agreement between the independent NLO calculations in terms of $(z_a, z_b)$ and $(z, y)$ has been established.

Finally, we have also compared our numerical implementation of Eq. (S16)-(S24) with the rapidity spectra obtained from `Vrap 0.9` [68] for Drell-Yan and from `SusHi 1.7.0` [73, 74] for Higgs production, finding excellent agreement. Since `Vrap 0.9` implements the $(z, y)$ parametrization, this effectively confirms the distributional identities in Table S2 numerically with physical PDFs as test functions.

### D. Comment on rapidity-dependent soft threshold results in the literature

As we have discussed, the soft threshold limit is fully contained in the generalized threshold limit. Our results thus provide an independent confirmation that Eq. (4) is the correct soft threshold factorization for the cross section differential in both $Q$ and $Y$, or equivalently Eq. (5) for the two-dimensional partonic cross section in $(z_a, z_b)$.

Several results in the literature [19–23] considering the soft threshold factorization differential in rapidity differ from Eq. (4). The difference is manifest already at fixed NLO in the term $\mathcal{L}_0(1-z)[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)]$, which appears in the flavor-diagonal partonic cross sections. The distributional identity in Eq. (S10) unambiguously shows that this term has a double singularity in the limit $z_a \to 1$ and $z_b \to 1$, which means it contributes a priori at leading power in the soft limit $z = z_a z_b \to 1$, i.e., it contributes to the $m_a = m_b = -1$ term in Eq. (21). This can already be seen just by considering the distribution in the bulk since

$$\mathrm{d}z\,\mathrm{d}y\,\frac{1}{1-z}\Big(\frac{1}{y}+\frac{1}{1-y}\Big) = \mathrm{d}z_a\mathrm{d}z_b\,\frac{1}{(1-z_a)(1-z_b)}\big[1+\mathcal{O}(1-z_a,1-z_b)\big]\,. \tag{S25}$$

Moreover, Eq. (S10) shows that it contributes at leading-logarithmic order. The soft function in Eqs. (4) and (5) precisely contains the leading-power contribution of this term, which is given by the first four terms on the right-hand side of Eq. (S10).

By contrast, this term and analogous ones at higher order are missing in the leading-power resummed results in Refs. [19–23]. There, it is effectively argued that the contribution of such terms to the rapidity spectrum is power-suppressed in $1-z$, leading to the incorrect conclusion that the rapidity dependence in the soft threshold limit can be included simply by taking $\hat{\sigma}_{ij}(z_a, z_b)$ to be $\hat{\sigma}_{ij}(z)\,[\delta(y)+\delta(1-y)]/2$ or, depending on the reference, $\hat{\sigma}_{ij}(z)\,\delta(y-1/2)$, where $\hat{\sigma}_{ij}(z)$ is the inclusive, rapidity-integrated, partonic cross section. In the following, we give a critical appraisal of the arguments used to support this conclusion and show why they are flawed.

This replacement first appeared in Ref. [75], where it was conjectured to provide an approximation to the threshold-resummed rapidity spectrum at small $Y$. The phenomenological impact of the correct convolution structure on PDF determinations relying on soft threshold resummation was discussed in Ref. [7]. A detailed numerical study of the difference at the level of the resummed Drell-Yan rapidity spectrum was performed in Ref. [9].

*Argument based on PDF momentum fractions* What makes the $\mathcal{L}_0(1-z)[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)]$ term subtle is that it vanishes upon integration over $y$, so it drops out in the inclusive cross section. This fact alone is of course insufficient to argue that it is power suppressed at each point in the spectrum. It simply means that different leading-power terms conspire to cancel upon integration, which is clear in terms of $(z_a, z_b)$, as discussed below Eq. (S10).

The argument in Ref. [21] rests on the observation that the PDF arguments, $x_a/z_a(z,y)$ and $x_b/z_b(z,y)$, in the two-dimensional convolution integral are independent of $y$ at $z=1$, from which it is concluded that the $y$ dependence of the PDF arguments is power suppressed in $1-z$ and can be dropped. If this is done, the $y$ integral becomes unconstrained and can be carried out freely, which eliminates this term. More generally, one could then replace $\hat{\sigma}_{ij}(z,y) = \hat{\sigma}_{ij}(z)\,\delta(y-1/2)[1+\mathcal{O}(1-z)]$ underneath the convolution integral.

However, a closer inspection of the PDF arguments in terms of $(z,y)$ reveals that

$$\frac{x_a}{z_a(z,y)} = x_a\Big\{1+y(1-z)+\mathcal{O}[(1-z)^2]\Big\}\,, \qquad \frac{x_b}{z_b(z,y)} = x_b\Big\{1+(1-y)(1-z)+\mathcal{O}[(1-z)^2]\Big\}\,. \tag{S26}$$

Hence, the $y$ dependence is not power suppressed but multiplies the *leading* dependence of the PDF arguments on $z$ itself, and so it cannot be dropped. This becomes even clearer in terms of $(z_a, z_b)$,

$$y(1-z) = 1-z_a+\mathcal{O}[(1-z)^2]\,, \qquad (1-y)(1-z) = 1-z_b+\mathcal{O}[(1-z)^2]\,, \tag{S27}$$

which are precisely the $\mathcal{O}(\lambda^2)$ convolution variables $k^{\mp}/(Qe^{\pm Y})$ in Eq. (4).

To illustrate explicitly that the $y$ dependence has a leading-power effect, consider the hadronic soft threshold limit $1-x_a \sim 1-x_b \ll 1$ and a simple toy PDF with a power-law behavior near the endpoint, with $\alpha > 0$,

$$f(x) \equiv \theta(1-x)(1-x)^{\alpha}\,. \tag{S28}$$

Using Eq. (S10), it is straightforward to show that $\mathcal{L}_0(1-z)[\mathcal{L}_0(y)+\mathcal{L}_0(1-y)]$ gives rise to double logarithms of $1-x_{a,b}$, but performing the integral directly in terms of $(z,y)$ is tedious, essentially as tedious as deriving Eq. (S10) itself. Instead, to disprove the above argument and show that the $y$ dependence is not power suppressed, it suffices to consider two terms that have the same $y$ integral,

$$A(z,y) \equiv \mathcal{L}_0(1-z)\,\frac{\delta(y)+\delta(1-y)}{2}\,, \qquad B(z,y) \equiv \mathcal{L}_0(1-z)\,\delta\Big(y-\frac{1}{2}\Big)\,, \tag{S29}$$

and show that they give different results at leading power, while the above argument would imply that they do not. Convolving $A(z,y)$ and $B(z,y)$ against the toy PDFs over the domain shown in Fig. S1 yields

$$\begin{aligned}\int \mathrm{d}z\,\mathrm{d}y\,A(z,y)\,f\Big[\frac{x_a}{z_a(z,y)}\Big]f\Big[\frac{x_b}{z_b(z,y)}\Big] &= f(x_a)f(x_b)\Big[\frac{1}{2}\ln\big(1-x_a\big)+\frac{1}{2}\ln\big(1-x_b\big)-H_\alpha+\mathcal{O}(1{-}x_a,1{-}x_b)\Big]\,,\\ \int \mathrm{d}z\,\mathrm{d}y\,B(z,y)\,f\Big[\frac{x_a}{z_a(z,y)}\Big]f\Big[\frac{x_b}{z_b(z,y)}\Big] &= f(x_a)f(x_b)\Big[\ln\big(1-\max\{x_a,x_b\}\big)-H_\alpha+\mathcal{O}(1{-}x_a,1{-}x_b)\Big]\,,\end{aligned} \tag{S30}$$

where $H_\alpha$ is the harmonic number. To evaluate the integrals it is convenient to already expand at integrand level, e.g. $1-x_a/z_a = 1-x_a+y(1-z)$ up to higher powers in $1-z$ and $1-x_a$. The maximum in the second case arises because the integration region in $z$ along fixed $y=1/2$ is cut off by the square of the larger of the two momentum fractions. (The order of expanding in $1-x_a$ and $1-x_b$ also needs to be picked accordingly.) Clearly, the two results only coincide for $Y=0$, where $x_a=x_b$. Away from $Y=0$, the logarithmic dependence on $x_{a,b}$ and thus on $Y$ differs at leading power.

*Fourier-transform argument* An alternative line of argument [19, 20, 22] relies on taking the Fourier transform of the partonic cross section to also argue that the $y$ dependence is trivial. A first step is to change variables from $y$ to

$$u \equiv \frac{1}{2}\ln\frac{z_a}{z_b}\,, \qquad -u_{\max}\le u\le u_{\max}\,, \qquad u_{\max}\equiv \ln\frac{1}{\sqrt{z}}\,. \tag{S31}$$

(Note that the variables $u$ and $y$ are precisely interchanged in the notation used in Ref. [22].) One then considers the Fourier transform of the partonic cross section $C(z,u)$ with respect to $u$,

$$\tilde{C}(z,M) \equiv \int \mathrm{d}u\, e^{\mathrm{i}Mu}\,C(z,u) \stackrel{?}{=} \int \mathrm{d}u\,C(z,u)\big[1+\mathcal{O}(1-z)\big]\,. \tag{S32}$$

The second equality, which is in question, is based on observing that $C(z,u)$ only has support on an interval bounded by $u_{\max}\sim 1-z$, and concluding that the Fourier kernel can be expanded in $u\sim 1-z$ as $e^{\mathrm{i}Mu}\stackrel{?}{=}1+\mathcal{O}(1-z)$, and so $\tilde{C}(z,M)$ is independent of $M$ at leading power. Thus, taking the inverse Fourier transform, the partonic cross section may be approximated as

$$C(z,u) = \int\frac{\mathrm{d}M}{2\pi}\,e^{-\mathrm{i}Mu}\,\tilde{C}(z,M) \stackrel{?}{=} \delta(u)\int \mathrm{d}u'\,C(z,u')\big[1+\mathcal{O}(1-z)\big]\,. \tag{S33}$$

This argument is flawed because in order to satisfy the Fourier inversion theorem, one must count $M\sim(1-z)^{-1}$ if one wants to count $u\sim 1-z$. In particular, one is not allowed to count $M\sim 1$ when taking the limit $z\to 1$ (or equivalently $N\to\infty$ for the Mellin conjugate $N$ of $z$). This is essential because $C(z,u)$ contains distributional terms in $u$ that cancel the suppression by the integration domain.

To disprove Eq. (S33), it again suffices to consider $A(z,y)$ and $B(z,y)$ defined in Eq. (S29). Changing variables to $u$, we have

$$\mathrm{d}y\,A(z,y) = \mathrm{d}u\,\mathcal{L}_0(1-z)\,\frac{\delta(u+u_{\max})+\delta(u-u_{\max})}{2}\,, \qquad \mathrm{d}y\,B(z,y) = \mathrm{d}u\,\mathcal{L}_0(1-z)\,\delta(u)\,. \tag{S34}$$

Both terms satisfy the assumptions of the above argument, i.e., they only have support for $|u|\le u_{\max}$. Changing variables back to $y$, Eq. (S33) would imply that up to power corrections in $1-z$,

$$A(z,y) = \mathcal{L}_0(1-z)\,\frac{\delta(y)+\delta(1-y)}{2} \stackrel{?}{=} \mathcal{L}_0(1-z)\,\delta\Big(y-\frac{1}{2}\Big) = B(z,y)\,. \tag{S35}$$

In fact, the overall factor found in Ref. [22] is $\delta(y-1/2)$, while it is $[\delta(y)+\delta(1-y)]/2$ in Ref. [21], and the above argument was used in Ref. [22] to argue that the two are equivalent. As a distributional identity, this is obviously incorrect. The only thing that is equal between $A(z,y)$ and $B(z,y)$ are their $y$ integrals, and as demonstrated before, this is insufficient because the $y$ dependence of the PDF arguments is a leading-power effect and cannot be neglected.

*Summary* For $pp$ production processes in general, to correctly describe the soft threshold limit of differential observables that are sensitive to the total rapidity of the Born system, one must maintain the two-dimensional dependence on $z_a$ and $z_b$ in the convolutions against the PDFs. Equivalently, in Mellin space one must maintain two Mellin fractions $N_a$ and $N_b$ as in the original Ref. [5]. In terms of the Mellin conjugate $N$ of $z$ and a Fourier conjugate $M$ of another variable like $u$, one has to keep the dependence on $M\sim|N_a-N_b|\sim N$. In particular, reducing the two-dimensional convolution structure to one dimension – such that the rapidity dependence is only carried by the luminosity function – amounts to making an additional assumption that is not justified by taking the soft limit.

### E. Perturbative ingredients

Here, we collect the required perturbative ingredients to evaluate the generalized threshold factorization theorem to two loops as well as the highest few logarithmic terms at three loops.

#### 1. Hard functions

The hard function for Drell-Yan production, $q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-$, is given by

$$H_{ij}^{\rm DY}(Q^2,\mu) = \sum_q \sigma_{B,q}^{\rm DY} \left(\delta_{iq}\delta_{j\bar{q}} + \delta_{i\bar{q}}\delta_{jq}\right) |C_{q\bar{q}}^V(Q^2,\mu)|^2 \,, \tag{S36}$$

where the sum runs over $q = \{u,d,c,s,b\}$, $\sigma_{B,q}^{\rm DY}$ is the Born cross section given in Eq. (S14), and $C_{q\bar{q}}^V$ is the Wilson coefficient from matching the QCD quark vector current onto SCET. In principle, we also need the matching coefficient $C_{q\bar{q}}^A$ for the axial-vector current, which differs from $C_{q\bar{q}}^V$ starting at $\mathcal{O}(\alpha_s^2)$ by small singlet corrections due to the large mass splitting between bottom and top quarks. We neglect these terms, as is done in `Vrap 0.9`, and use $H_{ij}^{\rm DY} \propto |C_{q\bar{q}}^V|^2$ throughout. The hard function for gluon-fusion Higgs production in the limit $m_H^2 \ll 4m_t^2$ reads

$$H_{ij}^{ggH}(m_t,m_H,\mu) = \sigma_B^{ggH} \left|\alpha_s(\mu)\, C_t(m_t,\mu)\right|^2 \delta_{ig}\delta_{jg} \left|C_{gg}(m_H,\mu)\right|^2 , \tag{S37}$$

where $\sigma_B^{ggH}$ and $C_t$ are given in Eqs. (S19) and (S20). The Wilson coefficient $C_{gg}$ arises from matching the $gg \to H$ operator in Eq. (S19) onto SCET. The Wilson coefficients contain the IR-finite virtual corrections to the Born process. They are normalized as $C = 1 + \mathcal{O}(\alpha_s)$ and can be found in Ref. [76] in our notation.

In the main text, we also refer to the fixed-order expansion of the hard function, where the coefficients $H_{ij}^{(n)}$ include all prefactors that are present at Born level,

$$H_{ij}(Q^2,\mu) = \sum_{n=0}^{\infty} \Big[\frac{\alpha_s(\mu)}{4\pi}\Big]^n H_{ij}^{(n)}(Q^2,\mu) \,. \tag{S38}$$

#### 2. Beam functions

For $t \gg \Lambda_{\rm QCD}^2$, the modified beam function defined by Eq. (15) can be matched onto PDFs as

$$\tilde{B}_i(t,x,\mu) = \int \frac{{\rm d}z}{z}\, \tilde{\mathcal{I}}_{ij}(t,z,\mu)\, f_j\Big(\frac{x}{z},\mu\Big) \Big[1 + \mathcal{O}\Big(\frac{\Lambda_{\rm QCD}^2}{t}\Big)\Big] \,, \tag{S39}$$

which directly follows from the analogous matching relations for the inclusive and double-differential beam functions, $B_i(t,x,\mu)$ and $B_i(t,\vec{k}_T,x,\mu)$, with matching coefficients $\mathcal{I}_{ij}(t,z,\mu)$ and $\mathcal{I}_{ij}(t,\vec{k}_T,z,\mu)$ [39, 41, 55].

The modified beam function satisfies the same RGE as the other beam functions,

$$\mu\frac{\rm d}{{\rm d}\mu}\tilde{B}_i(t,x,\mu) = \int {\rm d}t'\, \gamma_B^i(t-t',\mu)\, \tilde{B}_i(t',x,\mu) \,, \quad \gamma_B^i(t,\mu) = -2\Gamma_{\rm cusp}^i\big[\alpha_s(\mu)\big]\, \mathcal{L}_0(t,\mu^2) + \gamma_B^i\big[\alpha_s(\mu)\big]\delta(t) \,, \tag{S40}$$

because the RGE does not change the $\vec{k}_T$ dependence [55]. The matching coefficients satisfy the RGE [41]

$$\mu\frac{\rm d}{{\rm d}\mu}\tilde{\mathcal{I}}_{ij}(t,z,\mu) = \int {\rm d}t' \int \frac{{\rm d}z'}{z'}\, \tilde{\mathcal{I}}_{ik}\Big(t-t',\frac{z}{z'},\mu\Big) \Big\{\delta_{kj}\delta(1-z')\,\gamma_B^i(t',\mu) - 2\delta(t')\,P_{kj}\big[\alpha_s(\mu),z'\big]\Big\} \,, \tag{S41}$$

where $P_{ij}(\alpha_s,z)$ is the PDF anomalous dimension. We define the perturbative expansion of $\tilde{\mathcal{I}}_{ij}$ as

$$\tilde{\mathcal{I}}_{ij}(t,z,\mu) = \sum_{n=0}^{\infty} \Big[\frac{\alpha_s(\mu)}{4\pi}\Big]^n \tilde{\mathcal{I}}_{ij}^{(n)}(t,z,\mu) \,. \tag{S42}$$

Solving Eq. (S41) order by order in $\alpha_s$, we obtain a general expression for $\tilde{\mathcal{I}}_{ij}^{(n)}(t,z,\mu)$

$$
\begin{aligned}
\tilde{\mathcal{I}}_{ij}^{(0)}(t,z,\mu) &= \delta(t)\,\delta_{ij}\delta(1-z)\,,\\
\tilde{\mathcal{I}}_{ij}^{(1)}(t,z,\mu) &= \mathcal{L}_1(t,\mu^2)\,\Gamma_0^i\,\delta_{ij}\delta(1-z) + \mathcal{L}_0(t,\mu^2)\Big[P_{ij}^{(0)}(z) - \frac{\gamma_{B\,0}^i}{2}\,\delta_{ij}\delta(1-z)\Big] + \delta(t)\,\tilde{I}_{ij}^{(1)}(z)\,,\\
\tilde{\mathcal{I}}_{ij}^{(2)}(t,z,\mu) &= \mathcal{L}_3(t,\mu^2)\,\frac{(\Gamma_0^i)^2}{2}\,\delta_{ij}\delta(1-z)\\
&\quad + \mathcal{L}_2(t,\mu^2)\,\frac{\Gamma_0^i}{2}\Big[-\Big(\beta_0+\frac{3}{2}\gamma_{B\,0}^i\Big)\delta_{ij}\delta(1-z) + 3P_{ij}^{(0)}(z)\Big]\\
&\quad + \mathcal{L}_1(t,\mu^2)\Big[\Big(-\frac{\pi^2}{6}(\Gamma_0^i)^2 + \frac{\beta_0}{2}\gamma_{B\,0}^i + \frac{(\gamma_{B\,0}^i)^2}{4} + \Gamma_1^i\Big)\delta_{ij}\delta(1-z) - \big(\beta_0+\gamma_{B\,0}^i\big)P_{ij}^{(0)}(z)\\
&\qquad + \big(P_{ik}^{(0)}\otimes P_{kj}^{(0)}\big)(z) + \Gamma_0^i\tilde{I}_{ij}^{(1)}(z)\Big]\\
&\quad + \mathcal{L}_0(t,\mu^2)\Big[\Big(\zeta_3(\Gamma_0^i)^2 + \frac{\pi^2}{12}\Gamma_0^i\gamma_{B\,0}^i - \frac{\gamma_{B\,1}^i}{2}\Big)\delta_{ij}\delta(1-z) - \frac{\pi^2}{6}\Gamma_0^i P_{ij}^{(0)}(z) + P_{ij}^{(1)}(z)\\
&\qquad - \Big(\beta_0+\frac{\gamma_{B\,0}^i}{2}\Big)\tilde{I}_{ij}^{(1)}(z) + \big(\tilde{I}_{ik}^{(1)}\otimes P_{kj}^{(0)}\big)(z)\Big]\\
&\quad + \delta(t)\,\tilde{I}_{ij}^{(2)}(z)\,,\\
\tilde{\mathcal{I}}_{ij}^{(3)}(t,z,\mu) &= \mathcal{L}_5(t,\mu^2)\,\frac{(\Gamma_0^i)^3}{8}\delta_{ij}\delta(1-z)\\
&\quad + \mathcal{L}_4(t,\mu^2)\,\frac{5}{8}(\Gamma_0^i)^2\Big[-\Big(\frac{2}{3}\beta_0+\frac{\gamma_{B\,0}^i}{2}\Big)\delta_{ij}\delta(1-z) + P_{ij}^{(0)}(z)\Big]\\
&\quad + \mathcal{L}_3(t,\mu^2)\,\Gamma_0^i\Big[\Big(-\frac{\pi^2}{6}(\Gamma_0^i)^2 + \frac{\beta_0^2}{3} + \frac{5}{6}\beta_0\gamma_{B\,0}^i + \frac{(\gamma_{B\,0}^i)^2}{4} + \Gamma_1^i\Big)\delta_{ij}\delta(1-z) - \Big(\frac{5}{3}\beta_0+\gamma_{B\,0}^i\Big)P_{ij}^{(0)}(z)\\
&\qquad + \big(P_{ik}^{(0)}\otimes P_{kj}^{(0)}\big)(z) + \frac{\Gamma_0^i}{2}\,\tilde{I}_{ij}^{(1)}(z)\Big]\\
&\quad + \cdots + \delta(t)\,\tilde{I}_{ij}^{(3)}(z)\,,
\end{aligned}
\tag{S43}
$$

where $\mathcal{L}_n(t,\mu^2) \equiv (1/\mu^2)\mathcal{L}_n(t/\mu^2)$, $\tilde{I}_{ij}^{(n)}(z)$ is the $\mathcal{O}(\alpha_s^n)$ boundary term that is not predicted by the RGE, and the ellipses in the three-loop result indicate the terms proportional to $\mathcal{L}_{0,1,2}(t,\mu^2)$. We also introduced the shorthand $(g\otimes h)(z) \equiv \int_z^1 \mathrm{d}z'/z'\, g(z')\,h(z/z')$ for the Mellin convolution of two functions of $z$. In practice, we use the `MT` package [77] to evaluate them analytically. Expanding the three-loop expression for $\tilde{\mathcal{I}}_{ij}(t,z,\mu)$ against the hard function yields Eq. (22) in the main text.

The anomalous dimension coefficients in Eq. (S43) are as follows. The QCD $\beta$ function coefficients are

$$
\beta(\alpha_s) = -2\alpha_s\sum_{n=0}^{\infty}\beta_n\Big(\frac{\alpha_s}{4\pi}\Big)^{n+1}\,,\qquad \beta_0 = \frac{11}{3}\,C_A - \frac{4}{3}\,T_F\,n_f\,,\qquad \beta_1 = \frac{34}{3}\,C_A^2 - 2T_F\,n_f\Big(\frac{10}{3}\,C_A + 2C_F\Big)\,. \tag{S44}
$$

The anomalous dimensions are expanded as

$$
\Gamma_{\rm cusp}^i(\alpha_s) = \sum_{n=0}^{\infty}\Gamma_n^i\Big(\frac{\alpha_s}{4\pi}\Big)^{n+1}\,,\qquad \gamma_B^i(\alpha_s) = \sum_{n=0}^{\infty}\gamma_{B\,n}^i\Big(\frac{\alpha_s}{4\pi}\Big)^{n+1}\,,\qquad P_{ij}(\alpha_s,z) = \sum_{n=0}^{\infty}P_{ij}^{(n)}(z)\Big(\frac{\alpha_s}{4\pi}\Big)^{n+1}\,. \tag{S45}
$$

The coefficients of the cusp anomalous dimension to two loops are [78]

$$
\Gamma_0^i = 4C_i\,,\qquad \Gamma_1^i = 4C_i\Big[C_A\Big(\frac{67}{9}-\frac{\pi^2}{3}\Big) - \frac{20}{9}\,T_F\,n_f\Big] = \frac{4}{3}C_i\big[(4-\pi^2)C_A + 5\beta_0\big]\,,
$$

where $C_q = C_F$ and $C_g = C_A$. The beam function anomalous dimension coefficients are [41–44]

$$
\begin{aligned}
\gamma_{B\,0}^q &= 6C_F\,,\qquad \gamma_{B\,1}^q = C_F\Big[\Big(\frac{146}{9}-80\zeta_3\Big)C_A + (3-4\pi^2+48\zeta_3)C_F + \Big(\frac{121}{9}+\frac{2\pi^2}{3}\Big)\beta_0\Big]\,,\\
\gamma_{B\,0}^g &= 2\beta_0\,,\qquad \gamma_{B\,1}^g = \Big(\frac{182}{9}-32\zeta_3\Big)C_A^2 + \Big(\frac{94}{9}-\frac{2\pi^2}{3}\Big)C_A\,\beta_0 + 2\beta_1.
\end{aligned}
\tag{S46}
$$

The one-loop PDF anomalous dimensions are

$$P^{(0)}_{q_iq_j}(z) = P^{(0)}_{\bar{q}_i\bar{q}_j}(z) = 2C_F\,\delta_{ij}\,\theta(z)P_{qq}(z)\,, \qquad P^{(0)}_{q_ig}(z) = P^{(0)}_{\bar{q}_ig}(z) = 2T_F\,\theta(z)P_{qg}(z)\,,$$
$$P^{(0)}_{gg}(z) = 2C_A\,\theta(z)P_{gg}(z) + \beta_0\,\delta(1-z)\,, \qquad P^{(0)}_{gq_i}(z) = P^{(0)}_{g\bar{q}_i}(z) = 2C_F\,\theta(z)P_{gq}(z)\,, \tag{S47}$$

with the standard color-stripped splitting functions,

$$P_{qq}(z) = \mathcal{L}_0(1-z)(1+z^2) + \frac{3}{2}\delta(1-z)\,, \qquad P_{qg}(z) = \theta(1-z)\big[1-2z(1-z)\big]\,,$$
$$P_{gg}(z) = 2\mathcal{L}_0(1-z)\frac{(1-z+z^2)^2}{z}\,, \qquad P_{gq}(z) = \theta(1-z)\frac{1+(1-z)^2}{z}\,. \tag{S48}$$

### *3. Calculation of beam function boundary terms*

The structure of Eq. (S43) is exactly the same as for the inclusive beam function $\mathcal{I}_{ij}(t,z,\mu)$ [43, 44], except for the different boundary terms $I^{(n)}_{ij}(z) \neq \tilde{I}^{(n)}_{ij}(z)$. The definition in Eq. (15) implies for the matching coefficients

$$\tilde{\mathcal{I}}_{ij}(t,z,\mu) = \int \mathrm{d}^2\vec{k}_T\,\mathcal{I}_{ij}\Big(t - \frac{k_T^2}{2}, \vec{k}_T, z, \mu\Big)\,, \tag{S49}$$

which we use to calculate $\tilde{I}^{(n)}_{ij}(z)$ to the extent that the double-differential matching coefficients are known, i.e., to one loop for $i = g$ [55] and two loops for $i = q$ [55, 56]. Note that for $i = g$, the integral over all $\vec{k}_T$ leaves behind only the polarization-independent piece of the double-differential gluon beam function. We have also verified that the $\mu$-dependent pieces $\propto \mathcal{L}_n(t,\mu)$ obtained from Eq. (S49) agree with the RGE prediction Eq. (S43), i.e., we have explicitly checked that the projection and the RGE commute.

At one loop, we find that the following simple relation holds for all partonic channels,

$$\tilde{\mathcal{I}}^{(1)}_{ij}(t,z,\mu) = \mathcal{I}^{(1)}_{ij}(t,z,\mu) + \delta(t)\,P^{(0)}_{ij}(z)\ln\frac{2z}{1+z}\,, \tag{S50}$$

where $\mathcal{I}^{(1)}_{ij}$ is the one-loop matching coefficient for the inclusive beam function. Explicitly, the one-loop finite terms of the modified beam function are given by

$$\tilde{I}^{(1)}_{q_iq_j} = \tilde{I}^{(1)}_{\bar{q}_i\bar{q}_j} \equiv \delta_{ij}\tilde{I}^{(1)}_{qq}(z) = 2C_F\,\delta_{ij}\,\theta(z)\Big[\mathcal{L}_1(1-z)(1+z^2) - \frac{\pi^2}{6}\delta(1-z) + \theta(1-z)(1-z) + P_{qq}(z)\ln\frac{2}{1+z}\Big]\,,$$
$$\tilde{I}^{(1)}_{q_ig} = \tilde{I}^{(1)}_{\bar{q}_ig} = \tilde{I}^{(1)}_{qg}(z) = 2T_F\,\theta(z)\Big[P_{qg}(z)\ln\frac{2(1-z)}{1+z} + \theta(1-z)\,2z(1-z)\Big]\,,$$
$$\tilde{I}^{(1)}_{gg}(z) = 2C_A\,\theta(z)\Big[\mathcal{L}_1(1-z)\frac{2(1-z+z^2)^2}{z} - \frac{\pi^2}{6}\delta(1-z) + P_{gg}(z)\ln\frac{2}{1+z}\Big]\,,$$
$$\tilde{I}^{(1)}_{gq_i} = \tilde{I}^{(1)}_{g\bar{q}_j} = \tilde{I}^{(1)}_{gq}(z) = 2C_F\,\theta(z)\Big[P_{gq}(z)\ln\frac{2(1-z)}{1+z} + \theta(1-z)\,z\Big]\,. \tag{S51}$$

We decompose the two-loop quark finite terms $\tilde{I}^{(2)}_{ij}(z)$ by their flavor structure as

$$\tilde{I}^{(2)}_{q_iq_j}(z) = \tilde{I}^{(2)}_{\bar{q}_i\bar{q}_j}(z) = \delta_{ij}\tilde{I}^{(2)}_{qqV}(z) + \tilde{I}^{(2)}_{qqS}(z)\,, \qquad \tilde{I}^{(2)}_{q_i\bar{q}_j}(z) = \tilde{I}^{(2)}_{\bar{q}_iq_j}(z) = \delta_{ij}\tilde{I}^{(2)}_{q\bar{q}V}(z) + \tilde{I}^{(2)}_{qqS}(z)\,,$$
$$\tilde{I}^{(2)}_{q_ig}(z) = \tilde{I}^{(2)}_{\bar{q}_ig}(z) = \tilde{I}^{(2)}_{qg}(z)\,. \tag{S52}$$

As was done for $I^{(2)}_{ij}(z)$ in Refs. [43, 44], we find it convenient to pull common rational factors out of recurring terms with transcendental weight three ($\tilde{S}_3, \tilde{T}_3, \tilde{U}_3, \tilde{V}_3, \tilde{R}_{...}$), and group terms of lower transcendental weight separately by

color factor and flavor structure ($\tilde{C}_{...}$). We also pull out a conventional factor of four:

$$
\begin{aligned}
\tilde{I}^{(2)}_{qqV}(z) &= 4C_F^2\Big\{D_{qqV,C_F}(z) - \frac{2}{1-z}\tilde{T}_3(z) + \frac{1+z^2}{1-z}\big[\tilde{V}_3(z) - 2\tilde{U}_3(z)\big] + \tilde{C}_{qqV,C_F}(z)\Big\} \\
&\quad + 4C_F C_A\Big\{D_{qqV,C_A}(z) + \frac{1+z^2}{1-z}\big[\tilde{U}_3(z) + \tilde{R}_{qqV}(z)\big] + \tilde{C}_{qqV,C_A}(z)\Big\} + 4C_F\beta_0\Big[D_{qqV,\beta_0}(z) + \tilde{C}_{qqV,\beta_0}(z)\Big]\,, \\
\tilde{I}^{(2)}_{q\bar{q}V}(z) &= 4C_F(2C_F - C_A)\Big[\frac{1+z^2}{1+z}\tilde{S}_3(z) + \tilde{C}_{q\bar{q}V}(z)\Big]\,, \\
\tilde{I}^{(2)}_{qqS}(z) &= 4C_F T_F\Big[-2(1+z)\,\tilde{T}_3(z) + \tilde{C}_{qqS}(z)\Big]\,, \\
\tilde{I}^{(2)}_{qg}(z) &= 4T_F C_F\Big\{-2(1-z)^2\,\tilde{T}_3(z) + P_{qg}(z)\big[\tilde{V}_3(z) + \tilde{R}_{qg,C_F}(z)\big] + \tilde{C}_{qg,C_F}(z)\Big\} \\
&\quad + 4T_F C_A\Big\{-2(1+4z)\,\tilde{T}_3(z) - P_{qg}(z)\big[\tilde{U}_3(z) + \tilde{R}_{qg,C_A}(z)\big] + P_{qg}(-z)\,\tilde{S}_3(z) + \tilde{C}_{qg,C_A}(z)\Big\}\,.
\end{aligned}
\tag{S53}
$$

Here, overall factors of $\theta(z)\,\theta(1-z)$ are understood, but omitted for brevity. The $D_{qqV,...}(z)$ contain all distributional terms in $1-z$ and are the same as for the standard inclusive beam function [43],

$$
\begin{aligned}
D_{qqV,C_F}(z) &= (1+z^2)\Big[\mathcal{L}_3(1-z) - \frac{5\pi^2}{6}\mathcal{L}_1(1-z) + 4\zeta_3\mathcal{L}_0(1-z)\Big] + \frac{7\pi^4}{120}\delta(1-z)\,, \\
D_{qqV,C_A}(z) &= (1+z^2)\Big[\Big(\frac{2}{3} - \frac{\pi^2}{6}\Big)\mathcal{L}_1(1-z) + \Big(-\frac{8}{9} + \frac{7\zeta_3}{2}\Big)\mathcal{L}_0(1-z)\Big] + \Big(\frac{52}{27} - \frac{\pi^2}{6} - \frac{\pi^4}{36}\Big)\delta(1-z)\,, \\
D_{qqV,\beta_0}(z) &= (1+z^2)\Big[-\frac{1}{4}\mathcal{L}_2(1-z) + \frac{5}{6}\mathcal{L}_1(1-z) + \Big(-\frac{7}{9} + \frac{\pi^2}{12}\Big)\mathcal{L}_0(1-z)\Big] + \Big(\frac{41}{27} - \frac{5\pi^2}{24} - \frac{5\zeta_3}{6}\Big)\delta(1-z)\,.
\end{aligned}
\tag{S54}
$$

All remaining terms in Eq. (S53) are integrable for $z \to 1$. Their full expressions are lengthy and are available from the authors upon request. As an example of the structures that occur, we give

$$
\begin{aligned}
\tilde{S}_3(z) &= 2G\big(-1,-1,-\tfrac{1}{2};z\big) - 3G\big(-1,0,-\tfrac{1}{2};z\big) - 3G\big(0,-1,-\tfrac{1}{2};z\big) + 4G\big(0,0,-\tfrac{1}{2};z\big) \\
&\quad - 2H(-1,-1,0;z) + 2H(-1,0,-1;z) - 3H(-1,0,0;z) - 2H(-1,1,0;z) + H(0,-1,0;z) \\
&\quad - H(0,0,-1;z) - H(0,0,1;z) - 2H(1,-1,0;z) - \frac{\ln^3(z)}{12} - 2H(-1,0;z)\ln(1-z) \\
&\quad + \ln(2)\Big[-\frac{\pi^2}{6} + G\big(-1,-\tfrac{1}{2};z\big) - G\big(0,-\tfrac{1}{2};z\big) - 2H(-1,0;z) + \frac{\ln^2(z)}{2}\Big] \\
&\quad + \ln(1+z)\Big[-\frac{5\pi^2}{12} - G\big(-1,-\tfrac{1}{2};z\big) + G\big(0,-\tfrac{1}{2};z\big)\Big] \\
&\quad + \ln(z)\Big[\frac{\pi^2}{4} + 2G\big(-1,-\tfrac{1}{2};z\big) - 2G\big(0,-\tfrac{1}{2};z\big) + 2H(-1,0,z)\Big] + \frac{9\zeta_3}{4}\,.
\end{aligned}
\tag{S55}
$$

Here, we have used the recent `PolyLogTools` package [79] to convert all polylogarithms of rational functions of $z$ to standard harmonic polylogarithms $H(a_1,\ldots,a_n;z)$ as well as multiple polylogarithms $G(a_1,\ldots,a_n;z)$ of $z \in [0,1]$. The latter are as defined in Ref. [79], and for all $a_i = 0, \pm 1$ reduce to standard harmonic polylogarithms up to a sign. We find no evidence for a simple generalization of the one-loop relation Eq. (S50) at two loops. Finally, we note that the two-loop $\tilde{I}^{(2)}_{ij}(z)$ for the modified beam function are substantially more complicated than those for the standard inclusive beam function. For example, the latter does not involve polylogarithms with fractional weights, which only arise from the projection integral in Eq. (S49). We expect that similarly the three-loop cross section in terms of $(q^+, q^-, \vec{q}_T)$ will have a much simpler structure than in terms of $(Q, Y, \vec{q}_T)$.

### F. Breakdown of NNLO validation into partonic channels

Here, we provide the breakdown of the numerical validation of Eq. (14) at NNLO in Fig. 3 into individual partonic channels. Following `Vrap` [68], we take the $ij = q\bar{q}$ channel to include all topologies where $i$ and $j$ are part of the same quark line. The leading-power limit of these diagrams corresponds to the $qqV$ beam function matching coefficient in the decomposition in Eq. (S52). In addition, the $q\bar{q}$ channel also includes purely nonsingular contributions with topologies $q\bar{q} \to g \to q\bar{q}V$. We then take the $qq'$ channel to include the remaining quark-initiated processes, which

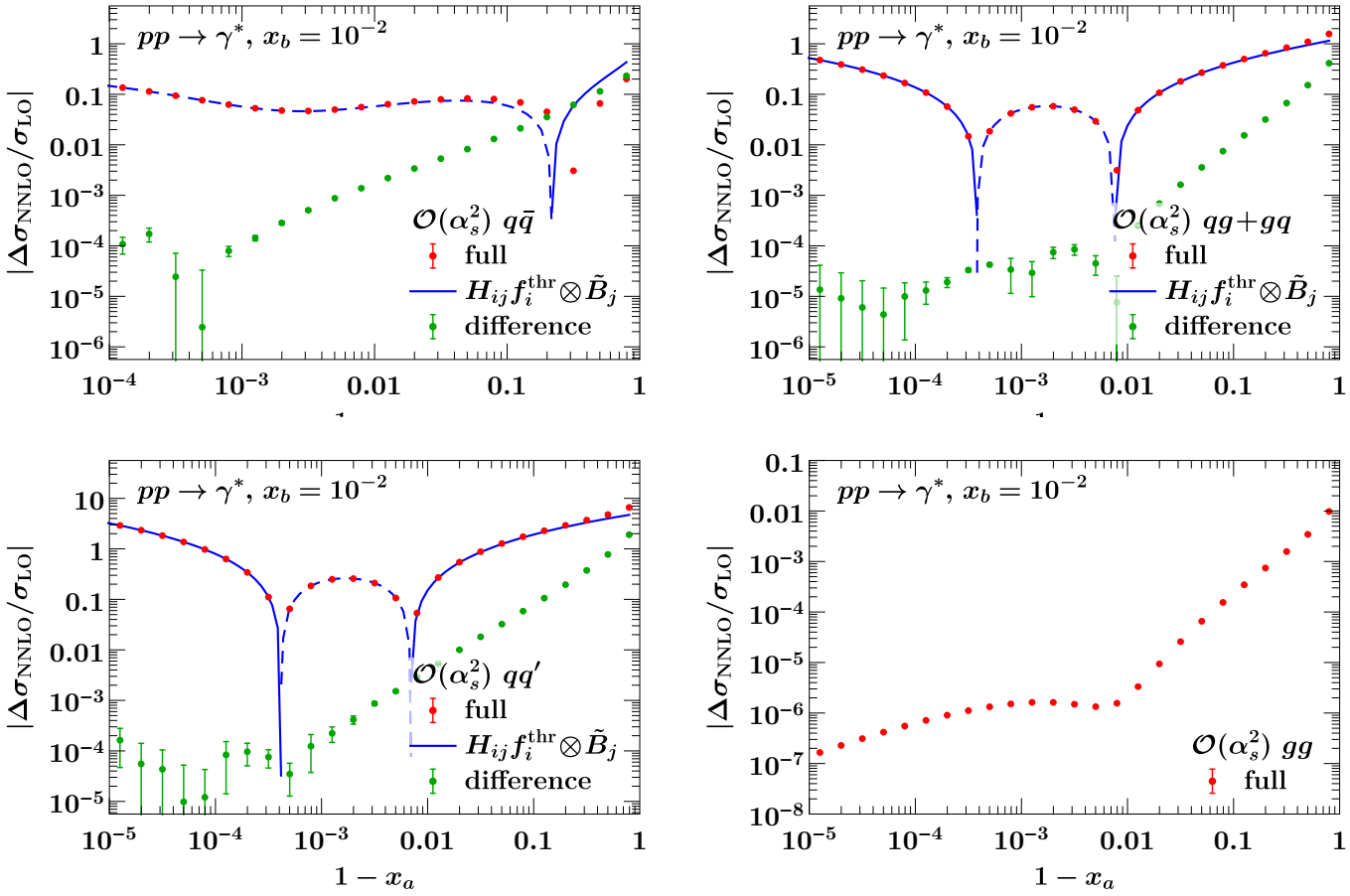


FIG. S2. Breakdown of Fig. 3 into partonic channels. Shown are the $\mathcal{O}(\alpha_s^2)$ contribution to $\mathrm{d}\sigma/\mathrm{d}x_a\mathrm{d}x_b$ predicted for $x_a \to 1$ by Eq. (14) (blue), the full result from `Vrap` (red). In all cases, their difference (green) vanishes like a power as $1 - x_a \to 0$, as it must. The $gg$ channel is power suppressed, so its full result by itself vanishes like a power. The error bars indicate the integration uncertainties. Dashing in the blue line indicates a negative result.

at leading power reduces to the sum of the $qqS$ and $q\bar{q}V$ beam function contributions in Eq. (S52). The $qg$ channel maps onto the $qg$ beam function contribution at leading power, while the $gq$ and $gg$ channels are purely nonsingular.

The results are shown in Fig. S2. In all cases, the prediction of Eq. (14) is in excellent agreement with the singular limit of the full calculation, with their difference vanishing as a power of $1 - x_a$ as it should. The excellent numerical stability of `Vrap` for the nondiagonal channels allows us to extend the check down to $1 - x_a = 10^{-5}$, where it becomes limited by MC statistics. For the $q\bar{q}$ channel, we start to see a systematic deviation at the $10^{-4}$ level below $1 - x_a \lesssim 10^{-4}$. We observe a similar deviation already at NLO, where the partonic cross sections agree analytically. We thus attribute this to a systematic effect in the PDF integrations in `Vrap`.

### G. Results for Drell-Yan at $\mu = Q/2$ and for gluon-fusion Higgs production

Here, we provide additional results for the generalized threshold expansion at fixed order. The analogs of Figs. 4 and 5 for the Drell-Yan rapidity spectrum at a different scale choice $\mu = Q/2$ are shown in Fig. S3, and for the $gg \to H$ NLO rapidity spectrum at $\mu = m_H$ and $\mu = m_H/2$ in Fig. S4.

When performing any threshold expansion for different scale choices, there are several options how to treat the terms in the partonic cross section that are predicted by the running of the PDFs or $\alpha_s$. One option is to expand these terms to the working order in the threshold expansion. This ensures that the partonic cross section has homogeneous power counting at any scale, but leaves the running of the PDFs and the coupling uncanceled beyond the working order. Another option is to threshold-expand the partonic cross section at a given reference scale and treat the running exactly. This leads to a privileged scale where the expansion was performed, but ensures the cancellation of

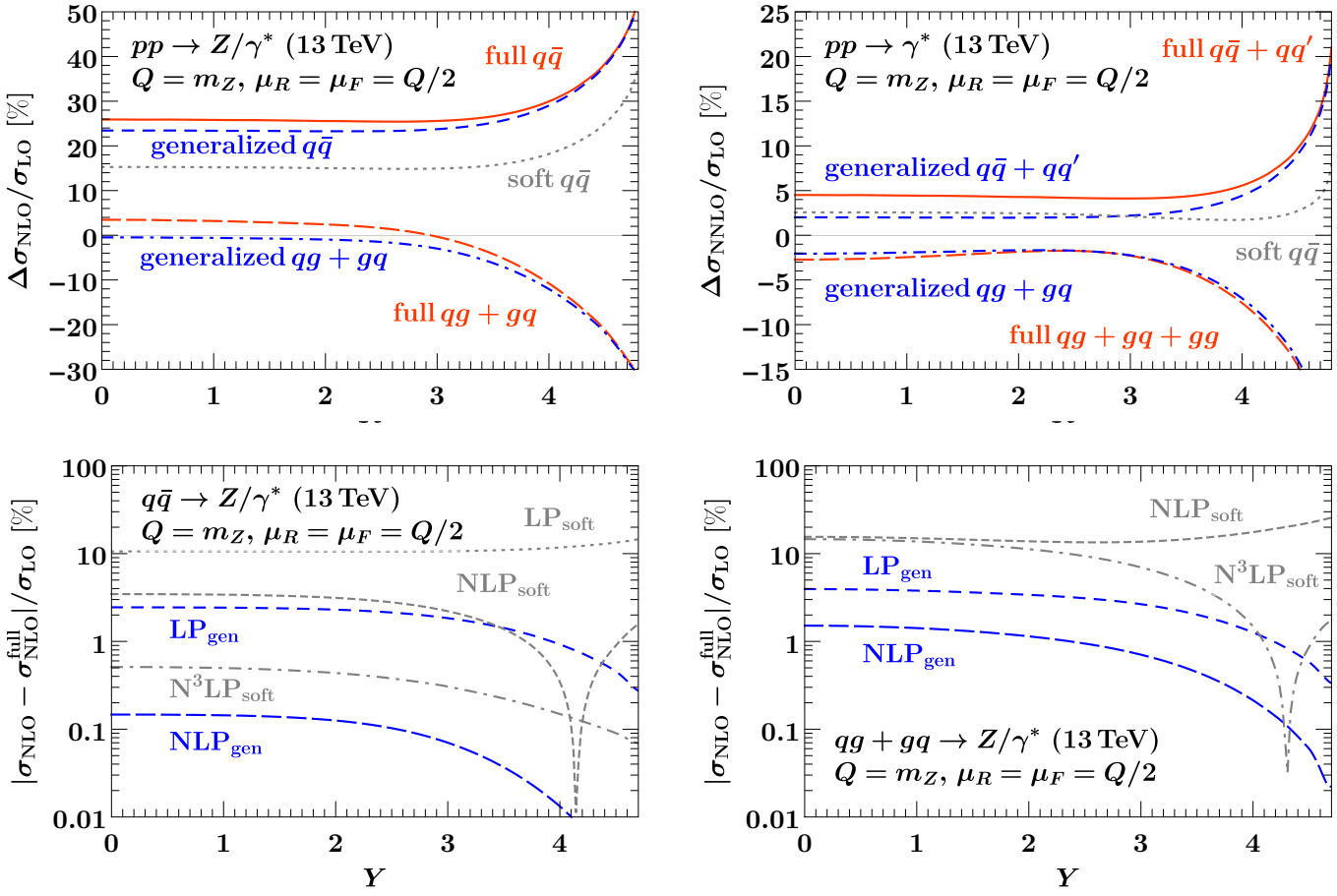


FIG. S3. Top row: Generalized threshold approximation of the $\mathcal{O}(\alpha_s)$ (top left) and $\mathcal{O}(\alpha_s^2)$ contribution (top right) to the Drell-Yan rapidity spectrum $\sigma \equiv \mathrm{d}\sigma/(\mathrm{d}Q\mathrm{d}Y)$ at $\mu = Q/2$ normalized to the LO result. This is the analog of Fig. 4.
Bottom row: Convergence of the generalized and soft threshold expansions for the $q\bar{q}$ (bottom left) and $qg + gq$ channels (bottom right) for $\mu = Q/2$. This is the analog of Fig. 5.

$\alpha_s$ and PDF running to all powers (up to higher orders in $\alpha_s$). In the following results, we choose the first option for definiteness. The difference between the two approaches could serve as a way to estimate the size of power corrections.

For Drell-Yan at $\mu = Q/2$, the generalized threshold expansion performs similarly well as for $\mu = Q$ in the main text, and again much better than the soft expansion. For $gg \to H$, the generalized threshold expansion again performs in a manner clearly superior to the soft one. Here, the increment from the leading-power soft to the leading-power generalized approximation of the $gg$ channel at NLO is roughly comparable to the piece still missing to the full result; either contribution amounts to $\mathcal{O}(20\%)$ in units of the Born cross section. This is consistent with the expectation that for a gluon-induced process, hard central radiation plays a larger role than for Drell-Yan. The shape of the NLO contributions at large $Y$ is well captured by the leading-power generalized approximation for both the $gg$ and $gq + qg$ channel. The leading-power soft approximation for the $gg$ channel (on close inspection) turns out to be off at large $Y$, and in both channels there is barely any convergence beyond leading power in the soft expansion at any $Y$.

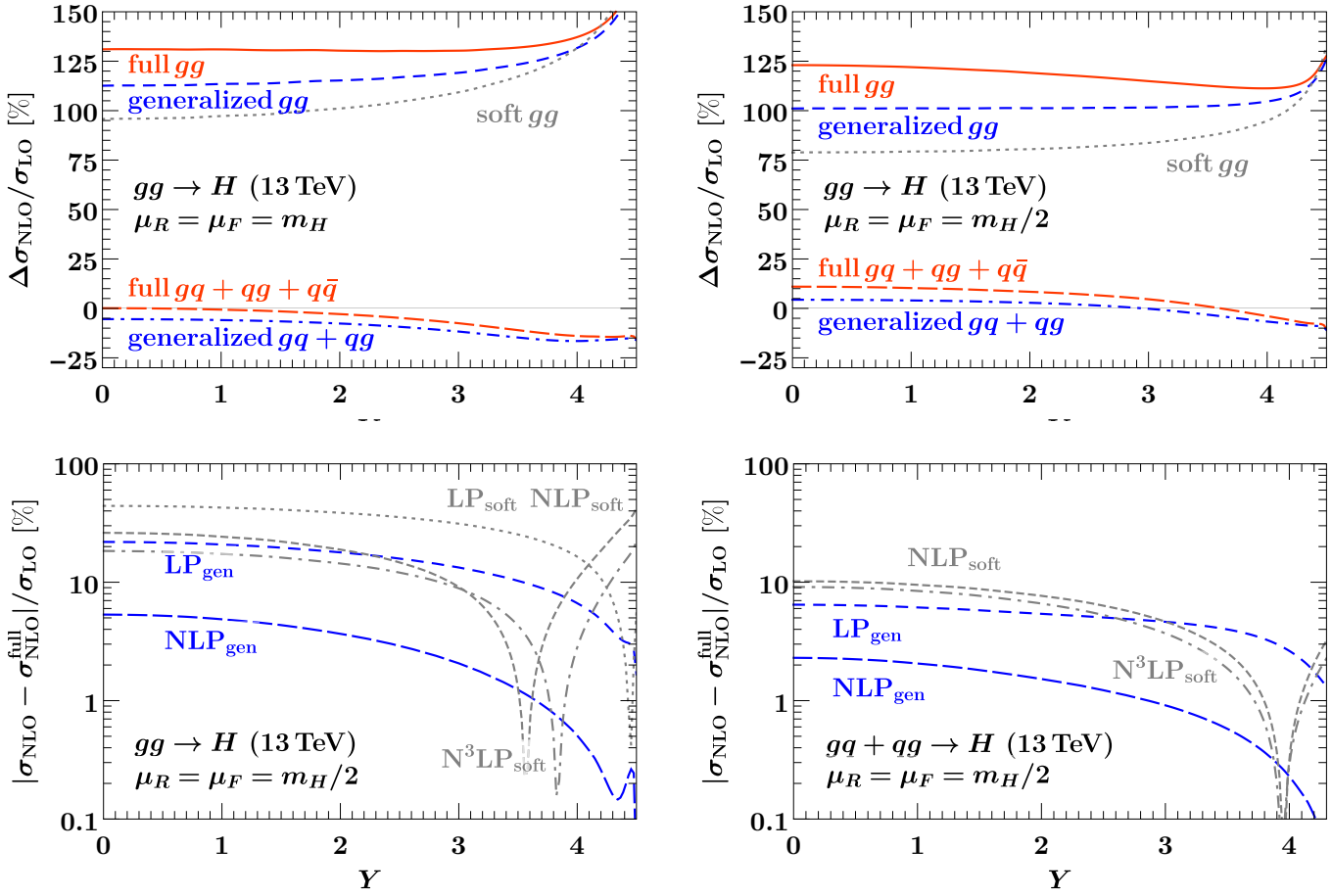


FIG. S4. Top row: Generalized threshold approximation of the $\mathcal{O}(\alpha_s)$ contribution to the $gg \to H$ rapidity spectrum $\sigma \equiv \mathrm{d}\sigma/\mathrm{d}Y$ at $\mu = m_H$ (top left) and $\mu = m_H/2$ (top right) normalized to the LO result. This is the analog of Fig. 4.
Bottom row: Convergence of the generalized and soft threshold expansions for the $gg$ (bottom left) and $gq + qg$ channels (bottom right) for $\mu = m_H/2$. This is the analog of Fig. 5.