

What is a parton shower?

Zoltán Nagy*

DESY, Notkestrasse 85, 22607 Hamburg, Germany

Davison E. Soper†

Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403-5203, USA



(Received 9 November 2017; published 25 July 2018)

We consider idealized parton shower event generators that treat parton spin and color exactly, leaving aside the choice of practical approximations for spin and color. We investigate how the structure of such a parton shower generator is related to the structure of QCD. We argue that a parton shower with splitting functions proportional to α_s can be viewed not just as a model, but as the lowest order approximation to a shower that is defined at any perturbative order. To support this argument, we present a formulation for a parton shower at order α_s^k for any k . Since some of the input functions needed are specified by their properties but not calculated, this formulation does not provide a useful recipe for an order α_s^k parton shower algorithm. However, in this formulation we see how the operators that generate the shower are related to operators that specify the infrared singularities of QCD.

DOI: [10.1103/PhysRevD.98.014034](https://doi.org/10.1103/PhysRevD.98.014034)

I. INTRODUCTION

Parton shower event generators for hadron collisions, such as HERWIG [1], PYTHIA [2], and SHERPA [3], perform calculations of cross sections according to an approximation to the standard model or its possible extensions. They are essential for the analysis of experiments at the Large Hadron Collider. The main ideas behind these generators were developed in the 1980s [4–6]. There has been extensive development of the algorithms since then [7–16]. The successor programs [1–3,17–23], are quite sophisticated. Useful reviews of the field can be found in [24,25]. One of the available successor programs is our own, DEDUCTOR [15,21,26–33]. This paper concerns the perturbative part of these parton shower generators, leaving aside models for the underlying event and hadronization. Furthermore, we consider an idealized version of a parton shower generator in which one accounts exactly for spin and color. Approximations for spin and color are a separate issue, which we do not discuss here.

Our aim in this paper is to investigate how the structure of a shower that treats spin and color exactly is related to the structure of QCD. In particular, we ask whether a

shower with splitting functions proportional to α_s can be the leading order approximation to something that is defined at any order in α_s . We find an affirmative answer to this question. Specifically, we find that there is a construction for defining a parton shower that generalizes current showers at any order of perturbation theory. We find also that the problem of relating the structure of a parton shower to the structure of QCD is not as straightforward as one might have naively guessed. First, the construction makes use of functions analogous to the Catani-Seymour dipole splitting functions [34] that specify the infrared behavior of QCD, but beyond leading order the shower splitting functions are not related to the functions that specify the infrared behavior of QCD by anything so simple as just changing their sign. Second, the formulas for the shower automatically includes factors that sum threshold logarithms [35–80]. These factors are not included in current parton shower generators at even leading order, except for DEDUCTOR [32,33]. Third, the formulas automatically include matching of the parton shower to a perturbative calculation beyond leading order of the hard scattering that starts the shower. This is fairly straightforward [81–103] for a leading order shower, but not for a shower beyond leading order.

We believe that is important to understand that a lowest order parton shower generator can represent the lowest order in a systematically improvable approximation. However, the construction in this paper is not a useful recipe for actually creating a parton shower algorithm beyond the leading order: some of the components of the recipe are specified by their properties but not explicitly

*Zoltan.Nagy@desy.de

†soper@uoregon.edu

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

constructed. A complete construction will require specifying such choices as a shower ordering variable and momentum mappings. Such a specification will require considerable effort, which lies beyond the scope of this paper.

The papers [104–108] present treatments of a parton shower at order α_s^2 . Reference [106] attempts to extend the dipole splittings often used in a leading order shower to a higher order analogue for the case of e^+e^- annihilation. This approach is similar in spirit to what we do in this paper. An alternative approach [104,105,107,108], concentrates on the next-to-leading order (NLO) Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) kernel for the evolution of parton distribution functions. Although the parton evolution kernels play a role in our formalism, it is not a central role.

II. OVERVIEW

The construction of a parton shower at any perturbative order, presented in Sec. VI, is rather abstract. In this section, we attempt to provide an overview of what the later mathematics is intending to do.

In our view, it is most useful to think of a parton shower algorithm as beginning with the theorem [109] that allows us to write a cross section for an infrared safe observable as a convolution of a hard scattering factor with parton distribution functions. Then the parton shower fills in more detail by using the renormalization group. The parton shower develops with decreasing values of a parameter that is a measure of the hardness of interactions.¹ The essential insight is that the scattering process appears differently depending on the hardness scale at which one examines it. At the hardest scale, the scale of the hard interaction, there are just a few partons (typically quarks and gluons). Then, as the hardness scale at which we examine the process decreases, these partons split, making more partons in a parton shower.² At any stage, a certain amount of structure has emerged, while softer structure remains unresolved.

In this paper, we start with the principle that a parton shower should fully reflect the infrared singularity structure of Feynman graphs for QCD and also the role of parton distribution functions in absorbing initial state singularities. Thus, we start with the infrared sensitive operator associated with the parton distributions and with a perturbative operator $\mathcal{D}(\mu^2)$ that represents the infrared singularities of QCD Feynman diagrams. We connect $\mathcal{D}(\mu^2)$ to both the shower splitting kernels and to the subtractions [110–117] needed to calculate a perturbative cross section beyond the leading order (LO). We work at arbitrary perturbative order.

¹HERWIG then rearranges the ordering of splittings in its shower so that larger angle splittings come first.

²Thus, with respect to initial state partons, the shower evolution starts from the hard interaction and moves backward in time to softer initial state interactions.

That is, we consider a hard scattering cross section calculated, with subtractions, at N^k LO and a parton shower with N^{k-1} LO splitting functions. (This counts a LO shower as having splitting functions proportional to α_s .)

The construction that we present is based on the operator $\mathcal{D}(\mu^2)$. This operator is to contain the infrared singularity structure of Feynman graphs for QCD. There is no unique recipe for constructing the α_s^n contribution, $\mathcal{D}^{(n)}(\mu^2)$, to $\mathcal{D}(\mu^2)$. As described in Sec. V, one needs to specify a definition of hardness associated with the integrations in graphs, one needs a momentum mapping, and one needs to specify the form of the functions used as one moves away from the strict soft and collinear limits. At first order, we have made these choices, so that $d\mathcal{D}^{(1)}(\mu^2)/d\log(\mu^2)$ is part of DEDUCTOR. At higher orders, we do not attempt to construct the $\mathcal{D}^{(n)}(\mu^2)$. Rather, we provide formulas for what to do once one has $\mathcal{D}^{(n)}(\mu^2)$ for $n \leq k$.

The formalism uses another operator $\mathcal{V}(\mu^2)$. This operator is obtained from $\mathcal{D}(\mu^2)$ and factors associated with the parton distributions but it is obtained by integrating over all of the parton splitting variables, so that it is infrared finite. In a standard first order parton shower, the Sudakov exponent is quite directly related to the part of $\mathcal{V}(\mu^2)$ that comes from one real parton emission. There is some freedom in setting the color and spin structure of $\mathcal{V}(\mu^2)$. Thus we leave $\mathcal{V}(\mu^2)$ partly unspecified.

The parton shower defined here needs to respect the structure of quantum field theory. Thus it includes quantum interference and maintains an exact accounting for the quantum spins and colors of the partons in the shower. The formulation is based on what we call the statistical space, introduced in Ref. [15]. It consists of states that describe the momenta, flavors, colors, and spins of any number of partons. The colors and spins are treated as fully quantum mechanical. This means that the statistical states are density matrices in the quantum color and spin space. We review the statistical space in Sec. IV.

Using the statistical space, we maintain an exact accounting for the quantum spins and colors. It is not known how to make this practical in computer code, particularly for color. Thus one needs separate approximations, such as the LC + approximation for color that is used in DEDUCTOR. We view the issue of approximations for spin and color as separate from the construction of the shower with exact spin and color. We do not discuss spin and color approximations in this paper.

The final result of the construction, presented in Eq. (134), is

$$\begin{aligned} \sigma[J] = & (1|\mathcal{O}_J\mathcal{U}(\mu_f^2, \mu_H^2)\mathcal{U}_V(\mu_f^2, \mu_H^2)\mathcal{F}(\mu_H^2)|\rho_H) \\ & + \mathcal{O}(\alpha_s^{k+B+1}) + \mathcal{O}(\mu_f^2/Q[J]^2). \end{aligned} \quad (1)$$

Here $|\rho_H\rangle$ is a statistical state representing the hard scattering, calculated at order α_s^{k+B} , where α_s^B is the order

of the Born hard scattering process. The hard scattering statistical state $|\rho_H\rangle$ includes subtractions, as in a normal N^kLO perturbative calculation. Then $\mathcal{F}(\mu_H^2)$ is an operator on the statistical space that multiplies by appropriate parton distribution functions and a parton luminosity factor.

The next operator in Eq. (1),

$$\mathcal{U}_{\mathcal{V}}(\mu_F^2, \mu_H^2) = \mathbb{T} \exp \left(\int_{\mu_F^2}^{\mu_H^2} \frac{d\mu^2}{\mu^2} \mathcal{S}_{\mathcal{V}}(\mu^2) \right), \quad (2)$$

is a process independent operator on the statistical space that leaves the number of partons unchanged and provides perturbative corrections needed to keep the measured cross section correct to order α_s^{k+B} . This factor also sums threshold logarithms associated with the hard scattering statistical state. The threshold logarithms are an essential part of the construction. As we will discuss, they are included in DEDUCTOR, but they are not part of other leading order shower generators. We presented an earlier formulation of threshold summation in a leading order shower in Ref. [32]. This formulation turned out to have certain flaws. In a companion paper [33], we exhibit the practical effects of the threshold summation according to Eq. (1) (but with $|\rho_H\rangle$ evaluated at leading order only).

The next operator in Eq. (1),

$$\mathcal{U}(\mu_F^2, \mu_H^2) = \mathbb{T} \exp \left(\int_{\mu_F^2}^{\mu_H^2} \frac{d\mu^2}{\mu^2} \mathcal{S}(\mu^2) \right), \quad (3)$$

is a process independent operator that creates more partons in a parton shower. In the first order case, $k = 1$, this is a rather standard parton shower if we average over spins and take the leading color approximation. In general, the splitting generator $\mathcal{S}(\mu^2)$ consists of terms that are of order α_s^n with $1 \leq n \leq k$. The shower starts at a hardness scale μ_H^2 appropriate for the hard scattering and ends at smaller hardness scale μ_F^2 that should be large enough so that perturbation theory at this scale can still be trusted.

The final operator in Eq. (1), \mathcal{O}_J , specifies the infrared safe measurement that one wants to make on the parton state after the shower. The hardness scale associated with this measurement is $Q[J]^2$. Finally, $(1|$ is an instruction to integrate over all of the parton variables.

The cross section $\sigma[J]$ is then correct to order α_s^{k+B} and includes a version of the cross section beyond this order, within the approximations of a parton shower. Notice that the property that the cross section including showering, $\sigma[J]$, is correct to order α_s^{k+B} means that the shower is matched to an order α_s^{k+B} perturbative calculation of $\sigma[J]$. This matching is an intrinsic part of the shower formulation.

We discuss a very general formulation of parton showers. However, we want to keep the notation simple, so, without loss of generality, we use Higgs boson production as an

TABLE I. New and old notation.

New	Old
$\mathcal{S}_{\text{pert}}^{(1,0)}$	$\mathcal{H}_I^{\text{pert}}$
$\mathcal{S}_{\text{pert}}^{(0,1)}$	$-\mathcal{S}^{\text{pert}}$
$[\mathcal{F} \circ \bar{\mathcal{S}}^{(1,0)}] \mathcal{F}^{-1}$	\mathcal{V}
$\mathcal{S}^{(1)} - \mathcal{S}_{i\bar{x}}^{(0,1)}$	$\mathcal{H}_I - \mathcal{V}$
$\mathcal{S}_{\mathcal{V}}^{(1)} + \mathcal{S}_{i\bar{x}}^{(0,1)}$	$\mathcal{V} - \mathcal{S}$

example. We use five flavors of quarks. In practical applications, one uses a variable flavor number scheme in which nonzero values of m_b and m_c appear. However, this creates complications, especially if we want to work at an arbitrary order of perturbation theory. Thus in this paper we set $m_b = m_c = 0$.

We have found it useful to change some of the notation that we used in our previous papers in order to address a much more general problem. We hope that this does not cause confusion. We provide a translation in Table I.

Following this brief overview, we include a brief Sec. III on factorization, which plays an important role in the conceptual development. Then we devote Sec. IV to partons and the spin and color density operator, which we use to define the statistical space. This leads us describe the perturbative cross section and the infrared sensitive operator $\mathcal{D}(\mu^2)$ in Sec. V. Then in Sec. VI we manipulate the perturbative cross section to define the parton shower. Section VII presents a summary and outlook.

There are four appendices. Appendix A presents a toy model for the operators used in the construction. We hope that this concrete model will prove instructive. Appendix B discusses the definition of parton distribution functions needed for a shower. Appendix C discusses how certain scale parameters can be chosen dynamically instead of statically, as in the main text. Appendix D discusses the relation of $\overline{\text{MS}}$ renormalization to the definition of the parton shower.

III. FACTORIZATION

We consider a hard scattering process in the collisions of two high energy hadrons, A and B. The hadrons carry momenta P_A and P_B . The hadron energies are high enough that we can simplify the equations describing the collision kinematics by treating the colliding hadrons as being massless. Then with a suitable choice of reference frame, the hadron momenta are

$$P_A = (P_A^+, 0, \mathbf{0}), \quad P_B = (0, P_B^-, \mathbf{0}). \quad (4)$$

Here we use momentum components $(p^+, p^-, \mathbf{p}_\perp)$ with $p^\pm = (p^0 \pm p^3)/\sqrt{2}$. We then imagine a parton level process in which a parton from hadron A, with flavor a

and momentum $p_a = \eta_a P_A$ collides with a parton from hadron B, with flavor b and momentum $p_b = \eta_b P_B$.

We will be interested in an inclusive cross section to create some hard state, for instance, a Z boson plus possibly jets, or just jets. We will use the production of a Higgs boson, $A + B \rightarrow H + \text{QCD partons}$ as our principle example. At the Born level, it is produced via the partonic process $g + g \rightarrow H$. We treat the Higgs boson as being stable and on shell. We denote the momentum of the Higgs boson by

$$p_H = \left(e^{y_H} \sqrt{(m_H^2 + \mathbf{p}_{H,\perp}^2)/2}, e^{-y_H} \sqrt{(m_H^2 + \mathbf{p}_{H,\perp}^2)/2}, \mathbf{p}_{H,\perp} \right). \quad (5)$$

The collision also produces QCD partons with flavors f_i and momenta p_i , with $i = 1, \dots, m$. In this paper, we consider the QCD partons to be massless. Each final state parton has rapidity y_i and transverse momentum $\mathbf{p}_{i,\perp}$, so that the components of its momentum are

$$p_i = \left(e^{y_i} \sqrt{\mathbf{p}_{i,\perp}^2/2}, e^{-y_i} \sqrt{\mathbf{p}_{i,\perp}^2/2}, \mathbf{p}_{i,\perp} \right). \quad (6)$$

It is up to us to decide what we want to measure about the final state of our process. We can consider many cases at once by simply saying that we are interested in a cross section $\sigma[J]$ to measure an observable quantity J , leaving the definition of J unspecified. We will see in the following subsection how $\sigma[J]$ can be specified for a general observable J . Then parton distribution functions relate $\sigma[J]$ to an analogous cross section $\hat{\sigma}[J]$ for the collision of two partons. In its briefest form, the relation is

$$\sigma[J] \approx \sum_{a,b} \int d\eta_a \int d\eta_b f_{a/A}(\eta_a, \mu^2) f_{b/B}(\eta_b, \mu^2) \hat{\sigma}[J]. \quad (7)$$

A. Infrared safety

We demand that the observable J be infrared safe. To specify what that means, we write Eq. (7) in more detail:

$$\begin{aligned} \sigma[J] = & \int dy_H \frac{d\sigma_0}{dy_H} J_0(p_H) \\ & + \int dy_H dy_1 d\mathbf{p}_{1,\perp} \frac{d\sigma_1}{dy_H dy_1 d\mathbf{p}_{1,\perp}} J_1(p_H, p_1) \\ & + \frac{1}{2!} \int dy_H dy_1 dy_2 d\mathbf{p}_{1,\perp} d\mathbf{p}_{2,\perp} \\ & \times \frac{d\sigma_2}{dy_H dy_1 dy_2 d\mathbf{p}_{1,\perp} d\mathbf{p}_{2,\perp}} J_2(p_H, p_1, p_2) + \dots \quad (8) \end{aligned}$$

Here we start with the cross section to produce the Higgs boson plus m partons with momenta

$$\{p\}_m = \{p_H, p_1, \dots, p_m\}. \quad (9)$$

We multiply the cross section by a function $J_m(\{p\}_m)$ that specifies the measurement that we want to make on the final state partons. These functions are taken to be symmetric under interchange of the QCD momentum arguments $\{p_1, \dots, p_m\}$. Accordingly, we divide by the number $m!$ of permutations of the QCD parton labels. We integrate over the momenta of the final state partons. The transverse momentum of the Higgs boson and the needed momentum fractions for the incoming partons are determined by momentum conservation. Finally, we sum over the number m of final state QCD partons.

Infrared safety is a property of the functions J_m that relates each function $J_{m+1}(\{p\}_{m+1})$ to the function $J_m(\{p\}_m)$ with one fewer parton. There are two requirements needed for J to be infrared safe.

First, consider the limit in which partons $m+1$ and m become collinear:

$$p_{m+1} \rightarrow z\tilde{p}_m, \quad p_m \rightarrow (1-z)\tilde{p}_m. \quad (10)$$

Here \tilde{p}_m is a lightlike momentum and $0 < z < 1$. We can concentrate on just partons with labels $m+1$ and m because the functions J are assumed to be symmetric under interchange of the parton labels. In order for J to be infrared safe, we demand that

$$\begin{aligned} J_{m+1}(\{p_H, p_1, \dots, p_{m-1}, p_m, p_{m+1}\}) \\ \rightarrow J_m(\{p_H, p_1, \dots, p_{m-1}, \tilde{p}_m\}) \quad (11) \end{aligned}$$

in the collinear limit (10).

Second, consider also the limit in which parton $m+1$ becomes collinear to one of the beams,

$$p_{m+1} \rightarrow \xi P_A \quad (12)$$

or

$$p_{m+1} \rightarrow \xi P_B. \quad (13)$$

Here $0 \leq \xi$. When $\xi = 0$, parton $m+1$ is simply becoming infinitely soft. In order for J to be infrared safe, we demand that

$$J_{m+1}(\{p_H, p_1, \dots, p_m, p_{m+1}\}) \rightarrow J_m(\{p_H, p_1, \dots, p_m\}) \quad (14)$$

in either limit (12) or (13).

Briefly, then, infrared safety means that the result of the measurement is not sensitive to whether or not one parton splits into two almost collinear partons and it is not sensitive to any partons that have very small momenta transverse to the beam directions.

B. A more quantitative view of infrared safety

We now discuss infrared safety a little more quantitatively. Consider, as above, two final state partons that are nearly collinear. This is modeled in a parton shower algorithm as a splitting of a parton with momentum \tilde{p}_m into two partons with momenta p_m and p_{m+1} . We can measure how close we are to the collinear limit by calculating³

$$\mu_{\text{split}}^2 = \frac{\sqrt{Q_H^2}}{E_m + E_{m+1}} (p_m + p_{m+1})^2. \quad (15)$$

Here Q_H is a momentum vector that describes the hard scattering, for instance the momentum p_H of the produced Higgs boson in our example. We define the parton energies in the rest frame of Q_H . The limit expressed in Eq. (10) is $\mu_{\text{split}}^2 \rightarrow 0$.

Alternatively, we can consider a splitting of an initial state parton as modeled in a parton shower algorithm. It suffices to consider the splitting of an initial state parton in hadron A. The initial state parton with momentum \tilde{p}_a becomes a new initial state parton with momentum p_a and a new final state parton with momentum p_{m+1} .⁴ We can measure how close we are to the collinear limit by calculating

$$\mu_{\text{split}}^2 = -\frac{\sqrt{Q_H^2}}{E_a - E_{m+1}} (p_a - p_{m+1})^2. \quad (16)$$

Again, the limit expressed in Eq. (12) is $\mu_{\text{split}}^2 \rightarrow 0$.

We now suppose that we measure the cross section $\sigma[J]$ corresponding to an infrared safe measurement function J , using Eq. (8). When $\mu_{\text{split}}^2 = 0$ in either example, application of Eq. (8) gives a cross section that we can call $\sigma_0[J]$. In applying Eq. (8), we can use the term with $m+1$ final state partons, with parton $m+1$ exactly collinear with either parton m or parton “a.” Equivalently, we can use the m parton state before the splitting. Because of Eq. (11) or Eq. (14), the result is exactly the same. Now, when μ_{split}^2 is small but not zero, we get a slightly different result, $\sigma[J]$. Let $\delta\sigma[J] = \sigma_0[J] - \sigma[J]$. We assume that $J_{m+1}(\{p\}_{m+1})$ is a smooth function of the parton momenta, at least near the soft or collinear limits. Then we will have $\delta\sigma[J] \rightarrow 0$ as $\mu_{\text{split}}^2 \rightarrow 0$. A typical case is $\delta\sigma[J] \propto \mu_{\text{split}}^2$ as $\mu_{\text{split}}^2 \rightarrow 0$. Then we can define a scale $Q^2[J]$ that is characteristic of the observable by

$$\left\langle \frac{\delta\sigma[J]}{\mu_{\text{split}}^2} \right\rangle = \frac{\sigma[J]}{Q^2[J]}. \quad (17)$$

The ratio in Eq. (17) can be sensitive to the parton configuration, so we average over all configurations with the same μ_{split}^2 . The scale $Q^2[J]$ measures how sensitive the cross section is to parton splittings.

An example may be helpful. Let $\sigma[J]$ measure the one-jet-inclusive cross section for jet transverse momentum P_J at small rapidity, using the anti- k_T jet algorithm with radius parameter R . Consider a very narrow jet in which one parton splits into two, one of which is soft. Then a simple estimate using the definitions above is

$$Q^2[J] \approx \frac{\sqrt{Q_H^2} (RP_J)^2}{P_J N}. \quad (18)$$

Here N measures how fast the jet cross section falls with increasing P_J :

$$N = \left| \frac{P_J}{\sigma(P_J)} \frac{d\sigma(P_J)}{dP_J} \right|. \quad (19)$$

This is a fairly large number, so that the effective scale $Q^2[J]$ for the jet measurement is smaller than P_J^2 . Additionally, R is typically chosen to be less than one. The smaller R is, the smaller $Q^2[J]$ is.

We see that we can understand infrared safety in terms of measurements that we might make in a parton shower simulation of a high energy scattering event. As the shower progresses, there are splittings. The corresponding values of μ_{split}^2 get smaller and smaller. If we measure $\sigma[J]$ at each stage of the shower, we will see that $\sigma[J]$ changes as the shower develops, but the changes get smaller and smaller according to

$$\delta\sigma[J] \sim \frac{\mu_{\text{split}}^2}{Q^2[J]} \sigma[J]. \quad (20)$$

This leads us to a limit on the possible accuracy of a perturbative approach to calculating $\sigma[J]$. When the parton shower has evolved to a 1 GeV scale, then we have reached the limits of perturbation theory. The fractional uncertainty associated with one more splitting is then

$$\delta\sigma[J] \sim \frac{(1 \text{ GeV})^2}{Q^2[J]} \sigma[J]. \quad (21)$$

The accuracy of the perturbative calculation can never be better than this.

C. Factorization

We can now state how one calculates the cross section for whatever observable J we want—as long as J is infrared

³One could choose other hardness measures μ^2 . This one is based on the parton shower formulation in [21,29].

⁴This is in the “backwards evolution” picture. Going forward in time, the parton with momentum p_a splits.

safe. The formula we use was stated in Eq. (7) and we restate it here in a slightly more detailed form [109]:

$$\sigma[J] = \sum_{a,b} \int d\eta_a \int d\eta_b f_{a/A}(\eta_a, \mu^2) f_{b/B}(\eta_b, \mu^2) \times \hat{\sigma}_{a,b}[\eta_a, \eta_b, \mu^2; J] + \mathcal{O}([m/Q]^n). \quad (22)$$

The intuitive basis for this is very simple. The factor $f_{a/A}(\eta_a, \mu^2) d\eta_a$ represents the probability to find a parton of flavor a in a hadron of flavor A . For the other hadron, the corresponding probability is $f_{b/B}(\eta_b, \mu^2) d\eta_b$. Then $\hat{\sigma}[J]$ is the cross section to obtain the observable J from the scattering of these partons, as given in Eq. (8). Naturally, this parton level cross section depends on the parton variables a, b, η_a, η_b . Here the differential cross sections to produce m final state partons contain delta functions that relate the momentum fractions η_a and η_b to the final state parton momenta. The parton distributions depend on a scale μ . This is often called the factorization scale μ_F and distinguished from the argument of α_s and other running couplings, which is called the renormalization scale μ_R . In order to keep our notation simple, we set $\mu_F = \mu_R = \mu$.

The cross section $\hat{\sigma}[J]$ has a perturbative expansion in powers of $\alpha_s(\mu^2)$. That is

$$\begin{aligned} \hat{\sigma}_{a,b}[\eta_a, \eta_b, \mu^2; J] &= \hat{\sigma}_{a,b}^{(0)}[\eta_a, \eta_b, \mu^2; J] \\ &+ \left[\frac{\alpha_s(\mu^2)}{2\pi} \right] \hat{\sigma}_{a,b}^{(1)}[\eta_a, \eta_b, \mu^2; J] \\ &+ \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^2 \hat{\sigma}_{a,b}^{(2)}[\eta_a, \eta_b, \mu^2; J] + \dots \end{aligned} \quad (23)$$

Here we do not display the factors of α_s or α_{ew} that appear in the Born level cross section $\hat{\sigma}^{(0)}$. Perturbative calculations can be at lowest order (LO), corresponding to one term in the expansion, next-to-lowest order (NLO) with two terms, sometimes NNLO, and, in general, N^kLO.

One useful property is that the dependence of the calculated cross section on μ^2 diminishes as we go to higher orders. Indeed, the cross section in nature, $\sigma[J]$, does not depend on μ^2 . Thus if we calculate to order α_s^k , the derivative of the calculated cross section with respect to μ^2 will be of order α_s^{k+1} .

There is an error term in Eq. (22). No matter how many terms are included in $\hat{\sigma}$, there are contributions that are left out. These terms are suppressed by a power of $m \sim 1$ GeV divided by a large scale parameter Q that characterizes the hard scattering process to be measured. These contributions arise from the approximations needed to derive Eq. (22). For instance if a loop momentum l flows through the wave function of quarks in a proton, we have to neglect l compared to the hard momenta, say the transverse

momentum of an observed jet. Not much is known about the general form of the power corrections for hadron-hadron collisions. In the rest of this paper, we will assume that the power n in Eq. (22) is $n = 2$ and we use $\sqrt{Q^2[J]}$ from Eq. (20) for Q . However, even if we lack a good estimate of the power corrections, it is important that they are there. If Q is of order 100 GeV, then the power corrections are completely negligible. However, if Q is of order 5 GeV, then we ought not to claim 1% accuracy in the calculation of $\sigma[J]$, no matter how many orders of perturbation theory we use.

IV. PARTONS AND THE DENSITY OPERATOR

We will describe perturbative calculations of cross sections, how these are connected to the parton shower description of these same cross sections, and how this is connected to factorization. We begin in this section with definitions that we need to describe the evolution of a parton shower. We follow the framework of Ref. [15].

A. Amplitudes and the density operator in spin and color

In a perturbative calculation of a cross section, one constructs an amplitude $|M(\{p, f\}_m)\rangle$. This amplitude depends on the momenta and flavors of two initial state partons, whatever outgoing electroweak partons there are, and m outgoing QCD partons. For our example of Higgs boson production, the momentum and flavor observables are

$$\{p, f\}_m = \{\eta_a, a, \eta_b, b, p_H, p_1, f_1, \dots, p_m, f_m\}. \quad (24)$$

The partons carry spin and color, so the amplitude is a vector in the partonic spin and color space for m final state QCD partons plus the two incoming partons, as indicated by the representation of the amplitude as a ket vector $|M(\{p, f\}_m)\rangle$. We use basis vectors $|\{s\}_m\rangle$ for the partonic spin space and basis vectors $|\{c\}_m\rangle$ for the partonic color space.⁵ To describe the evolution of a parton shower, it is useful to use quantum statistical mechanics, keeping the full quantum nature of the colors and spins. Thus we use the density operator in the color and spin space. The density operator is a linear combination of basis operators $|\{s, c\}_m\rangle\langle\{s', c'\}_m|$. Thus to describe the system at a certain stage of evolution, we use a function ρ of the number m of final state QCD partons and of the momenta and flavors $\{p, f\}_m$, where the value of ρ is a color-spin density operator. That is

⁵The spin basis vectors can be chosen in a very simple way, but it is not so trivial to choose useful basis vectors for the color space. The choice that we make for DEDUCTOR is specified in Ref. [15]. With this choice, the color basis vectors are not exactly normalized and the basis vectors for different colors are not exactly orthogonal.

$$\begin{aligned} \rho(\{p, f\}_m) &= \sum_{\{s, s', c, c'\}_m} \rho(\{p, f, s, s', c, c'\}_m) \\ &\times |\{s, c\}_m\rangle\langle\{s', c'\}_m|. \end{aligned} \quad (25)$$

The interpretation of this is that the differential probability dP for there to be m final state QCD partons with momenta and flavors $\{p, f\}_m$, times the expectation value of an operator O on the color and spin space, is

$$\begin{aligned} dP \times \langle O \rangle &= [d\{p\}_m] \sum_{\{s, s', c, c'\}_m} \rho(\{p, f, s, s', c, c'\}_m) \\ &\times \langle\{s', c'\}_m|O|\{s, c\}_m\rangle, \end{aligned} \quad (26)$$

where

$$\begin{aligned} [d\{p\}_m] &\equiv \frac{d^d p_H}{(2\pi)^d} 2\pi\delta_+(p_H^2 - m_H^2) \\ &\times \prod_{i=1}^m \left\{ \frac{d^d p_i}{(2\pi)^d} 2\pi\delta_+(p_i^2) \right\} d\eta_a d\eta_b \\ &\times (2\pi)^d \delta\left(p_a + p_b - p_H - \sum_{i=1}^m p_i\right). \end{aligned} \quad (27)$$

Here we use dimensional regularization with $d = 4 - 2\epsilon$.

The set of all such functions ρ constitutes a vector space, which we call the statistical space. We represent the vector ρ as a ket vector, $|\rho\rangle$. The rounded end of the ket is meant to distinguish a vector in the statistical space from a vector in the quantum spin \otimes color space.

Notice that we use the symbol ρ for four different but related concepts. First, for each choice of m and $\{p, f, s, s', c, c'\}_m$, $\rho(\{p, f, s, s', c, c'\}_m)$ is a complex number. Second, for each choice of m and $\{p, f\}_m$, $\rho(\{p, f\}_m)$ is a linear operator on the quantum spin \otimes color space for the $m + 2$ partons. Third, ρ is a linear map from m and $\{p, f\}_m$ to the space of operators on the quantum spin \otimes color space. Fourth, $|\rho\rangle$ is this linear map considered as an element of a vector space, the statistical space. This may seem complicated, but in the end we use almost entirely the statistical vectors $|\rho\rangle$. This then gives us what we think is a compact and powerful notation.

We can define basis vectors $|\{p, f, s, s', c, c'\}_m\rangle$ in the statistical space in such a way that

$$(\{p, f, s, s', c, c'\}_m|\rho) = \rho(\{p, f, s, s', c, c'\}_m). \quad (28)$$

The completeness relation for the basis vectors is

$$\begin{aligned} 1 &= \sum_m \frac{1}{m!} \int [d\{p\}_m] \sum_{\{f\}_m} \sum_{\{s, s', c, c'\}_m} |\{p, f, s, s', c, c'\}_m\rangle \\ &\times (\{p, f, s, s', c, c'\}_m|. \end{aligned} \quad (29)$$

B. Making an inclusive measurement

There is a special vector $(1|$ defined by

$$(1|\{p, f, s, s', c, c'\}_m) = \langle\{s', c'\}_m|\{s, c\}_m\rangle. \quad (30)$$

With this definition,

$$\begin{aligned} (1|\rho) &= \sum_m \frac{1}{m!} \int [d\{p\}_m] \\ &\times \sum_{\{f\}_m} \sum_{\{s, s', c, c'\}_m} \rho(\{p, f, s, s', c, c'\}_m) \\ &\times \langle\{s', c'\}_m|\{s, c\}_m\rangle \end{aligned} \quad (31)$$

is the total probability associated with the statistical state $|\rho\rangle$ as defined in Eq. (26) with $O = 1$.

With this notation, we begin with perturbatively calculated amplitudes $|M(\{p, f\}_m)\rangle$ for just a few partons. Thus we begin with a perturbatively calculated vector $|\rho\rangle$ in the statistical space. Then we use perturbative operations that are represented as linear operators on the statistical space. Similarly, the measurement J in Sec. III A is represented as a linear operator on the statistical space. Finally, multiplication by $(1|$ allows us to obtain the expectation value of the measurement operator.

C. Scales

The formalism also uses a reference vector Q_H and several scales: a renormalization scale μ_R^2 , a factorization scale μ_F^2 , and an ultraviolet cutoff scale μ_s^2 . For simplicity, all of these scales are set to a single scale μ^2 . The vector Q_H is used to set the value μ_H^2 of the common scale associated with the hard state $|\rho_H\rangle$, defined later in Eq. (58): $\mu_H^2 = Q_H^2$. A parton shower needs a measure of hardness of parton splittings. We also use Q_H as a vector to help define one possible measure of hardness, Λ^2 defined in Eq. (55). For this purpose, Q_H should be roughly in the direction of $p_a + p_b$ in an imagined initial hard scattering. The simplest way to define Q_H is to use the hard core part of the intended measurement. For instance, if we are looking for the cross section to produce a Higgs boson with rapidity near zero, we can take $Q_H^2 = m_H^2$ with zero rapidity and zero transverse part for Q_H . We will mention another, more dynamic, way to define Q_H later, in Appendix C.

D. Multiplying by parton distribution functions

In order to turn matrix elements into cross sections, we divide by a parton luminosity factor⁶ $n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b$. Here the color counting factor is $n_c(f) = 3$ for a quark flavor f and $n_c(f) = 8$ for $f = g$. The spin counting factor is $n_s(f) = 2$ for a quark flavor f and

⁶We have noticed that this factor was too small by a factor 2 in Eq. (3.15) of [15].

$n_s(f) = 2(1 - \epsilon)$ for $f = g$ when we work in $4 - 2\epsilon$ dimensions in order to regularize infrared divergences. Then we need to multiply by a parton distribution factor

$$F_{a,b}(\eta_a, \eta_b, \mu^2) = f_{a/A}(\eta_a, \mu^2) f_{b/B}(\eta_b, \mu^2). \quad (32)$$

The parton distribution functions here could be the five flavor $\overline{\text{MS}}$ parton distribution functions, or they could have a different definition. We combine the two parton distribution functions as one operator $\mathcal{F}(\mu^2)$ that acts on the statistical space:

$$\begin{aligned} \mathcal{F}(\mu^2) | \{p, f, s, s', c, c'\}_m \rangle \\ = \frac{F_{a,b}(\eta_a, \eta_b, \mu^2)}{n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b} | \{p, f, s, s', c, c'\}_m \rangle. \end{aligned} \quad (33)$$

We sometimes need a more general parton factor in which the parton distributions are convolved with a function that is also a matrix in the parton flavors. For instance, the evolution equation for the product of parton distribution functions is

$$\mu^2 \frac{d}{d\mu^2} F_{a,b}(\eta_a, \eta_b, \mu^2) = F'_{a,b}(\eta_a, \eta_b, \mu^2), \quad (34)$$

where the prime denotes differentiation and F' is given by the convolution product of F with an evolution kernel P ,

$$F' = F \circ P. \quad (35)$$

The precise definition is⁷

$$\begin{aligned} F'_{a,b}(\eta_a, \eta_b, \mu^2) \\ = \sum_{a',b'} \int_0^1 \frac{dz_a}{z_a} \int_0^1 \frac{dz_b}{z_b} F_{a',b'}(\eta_a/z_a, \eta_b/z_b, \mu^2) \\ \times P_{a,a',b,b'}(z_a, z_b, \mu^2). \end{aligned} \quad (36)$$

The evolution kernel for the product of parton distributions is the sum of parton evolution kernels for each of the two parton distribution functions:

$$\begin{aligned} P_{a,a',b,b'}(z_a, z_b, \mu^2) \\ = P_{a,a'}(z_a, \mu^2) \delta_{b,b'} \delta(z_b - 1) + \delta_{a,a'} \delta(z_a - 1) P_{b,b'}(z_b, \mu^2). \end{aligned} \quad (37)$$

This gives us an operator $\mathcal{F}'(\mu^2)$ defined by

⁷Based on the order of the flavor indices, it would be more conventional to write this as $P \circ F$, but we believe that the notation in Eq. (35) better expresses the physics in the context of this paper.

$$\begin{aligned} \mathcal{F}'(\mu^2) | \{p, f, s, s', c, c'\}_m \rangle \\ = \frac{F'_{a,b}(\eta_a, \eta_b, \mu^2)}{n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b} | \{p, f, s, s', c, c'\}_m \rangle. \end{aligned} \quad (38)$$

In place of $\mathcal{F}'(\mu^2)$, we use a notation that directly displays that $F'_{a,b}(\eta_a, \eta_b, \mu^2)$ is constructed according to Eq. (36),

$$\mathcal{F}'(\mu^2) = [\mathcal{F}(\mu^2) \circ \mathcal{P}(\mu^2)]. \quad (39)$$

The circle indicates the convolution and the square brackets $[\dots]$ indicate what is included in the convolution.

V. THE PERTURBATIVE CROSS SECTION

We now consider the cross section for some hard process. We can use any hard process that can lead to an infrared safe cross section, but the details of the notation depend on what hard process we consider. In order to keep the notation simple, we consider a specific process, the cross section make a Higgs boson plus QCD partons, calculated at $N^k\text{LO}$. We let J represent an infrared safe measurement of interest as in Sec. III A. The hardest scale associated with this measurement is μ_H^2 . The scale $Q^2[J]$ introduced in Eq. (17) could be much smaller, as long as it is large compared to 1 GeV^2 . In that case, the sort of fixed order calculation discussed in this section is not so useful because there are large logarithms, $\log(\mu_H^2/Q^2[J])$. Thus in this section, it may be helpful to imagine that $Q^2[J]$ is not very different from μ_H^2 .

A. The Born cross section

Let us begin at the Born level. We call the statistical state corresponding to the Born level matrix element $|\rho^{(0,0)}(\mu^2)\rangle$. The measurement J can be represented as an operator \mathcal{O}_J . To make a cross section, we need a luminosity factor and parton distribution functions. We define an operator $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$ as in Eq. (33) by

$$\begin{aligned} \mathcal{F}_{\overline{\text{MS}}}(\mu^2) | \{p, f, s, s', c, c'\}_m \rangle \\ = \frac{f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2) f_{b/B}^{\overline{\text{MS}}}(\eta_b, \mu^2)}{n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b} | \{p, f, s, s', c, c'\}_m \rangle. \end{aligned} \quad (40)$$

We use $\overline{\text{MS}}$ parton distributions for five flavors. Then the cross section for measurement operator \mathcal{O}_J is

$$\sigma[J] = (1 | \mathcal{F}_{\overline{\text{MS}}}(\mu^2) \mathcal{O}_J | \rho^{(0,0)}(\mu^2) \rangle). \quad (41)$$

B. The higher order cross section

Now we consider higher orders of perturbation theory. We define α_s and the parton distribution functions that we start with using $\overline{\text{MS}}$ renormalization, as reviewed in Appendix D. Including terms up to order α_s^k , the N^kLO cross section can be written as

$$\sigma[J] = (1 | [\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{Z}_F(\mu^2)] \mathcal{O}_J | \rho(\mu^2)) + \mathcal{O}(\alpha_s^{k+1}). \quad (42)$$

Here we convolve $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$ with the renormalization factor $\mathcal{Z}_F(\mu^2)$ for five flavor $\overline{\text{MS}}$ parton distribution functions.⁸ This factor is just 1 at order α_s^0 . Its higher order contributions are defined by working in $4 - 2\epsilon$ dimensions and consist of the $1/\epsilon^n$ pole terms needed to remove the ultraviolet poles from the renormalized operator that defines parton distributions. Equation (42) is a perturbative formula. We are to expand all of the factors up to the desired order α_s^k and neglect the remainder, indicated by the error estimate $\mathcal{O}(\alpha_s^{k+1})$.

The statistical state $|\rho(\mu^2)\rangle$ has a perturbative expansion

$$|\rho(\mu^2)\rangle = |\rho^{(0)}(\mu^2)\rangle + \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n |\rho^{(n)}(\mu^2)\rangle + \mathcal{O}(\alpha_s^{k+1}). \quad (43)$$

The order n contribution to the statistical state is a sum of terms,⁹

$$|\rho^{(n)}(\mu^2)\rangle = \sum_{n_r=0}^n \sum_{n_v=0}^n \theta(n_r + n_v = n) |\rho^{(n_r, n_v)}(\mu^2)\rangle. \quad (44)$$

In $|\rho^{(n_r, n_v)}(\mu^2)\rangle$, there are n_r final state partons and n_v virtual loops. This is for Higgs boson production as the Born process. If we had chosen two jet production as the Born level hard process, then there would be $2 + n_r$ partons in the final state.

The n_v virtual loops in $|\rho^{(n_r, n_v)}(\mu^2)\rangle$ can each produce $1/\epsilon$ and $1/\epsilon^2$ poles. The n_r partons in the final state can give soft and collinear singularities. The statistical state vector is singular when any of these partons become soft or

⁸This is for massless partons. Conceptually, $\mathcal{Z}_F(\mu^2)$ should be understood as the inverse, in the sense of convolutions, of the product of two parton-in-a-parton distribution functions with on-shell massless incoming parton states. Then this factor removes infrared poles from the cross section. At the level of bare operators, $f_{a/b}^{\text{bare}}(\xi) = \delta_{ab}\delta(1 - \xi)$. Convolving the renormalized parton-in-a-parton distribution functions with \mathcal{Z}_F gives the bare distribution functions, leading to Eq. (42). See Appendix D.

⁹It may be helpful to note that we define $|\rho^{(n)}(\mu^2)\rangle$ using on-shell matrix elements $M(\{p, f\}_m)$ and their complex conjugates, including the factors needed to make $M(\{p, f\}_m)$ into an S-matrix element. Then the right-hand side of Eq. (43) is invariant under changes of the renormalization scale μ^2 up to order α_s^{k+1} .

collinear with the beam directions or collinear with each other. In the case of a Born process that has final state QCD partons, there are also singularities when any of the n_r additional partons becomes collinear with starting final state partons. In these singular regions, the infrared safe measurement operator \mathcal{O}_J sees the partons that are collinear to a given direction as equivalent to a single parton and it does not see partons that are soft or collinear to the beam directions at all. Thus, it is as if we had a completely inclusive measurement as defined by left multiplying by $(1|$. Then we again get $1/\epsilon$ and $1/\epsilon^2$ poles. Many of the poles cancel between real and virtual graphs. There are, however, some poles associated with momenta that are collinear with the initial state parton momenta. These cancel the poles in $\mathcal{Z}_F(\mu^2)$. We are left with a finite result.

C. Introduction of the infrared sensitive operator $\mathcal{D}(\mu^2)$

The formula (42) is not completely practical because each term $|\rho^{(n_r, n_v)}(\mu^2)\rangle$ generates infrared singularities or poles. Only at the end of a calculation, which includes some complicated integrations that include the measurement function, do the poles produced by the infrared singularities cancel. To make this more practical, we define a certain operator $\mathcal{D}(\mu^2)$ and insert a factor $\mathcal{D}(\mu^2)\mathcal{D}^{-1}(\mu^2)$ into Eq. (42), giving

$$\sigma[J] = (1 | [\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{Z}_F(\mu^2)] \mathcal{D}(\mu^2) \mathcal{D}^{-1}(\mu^2) \mathcal{O}_J | \rho(\mu^2)) + \mathcal{O}(\alpha_s^{k+1}). \quad (45)$$

The operator $\mathcal{D}(\mu^2)$ depends on the dimensional regularization parameter ϵ , but we do not display this dependence explicitly. It depends on a second scale μ_s^2 along with μ^2 , but we set $\mu^2 = \mu_s^2$.

The idea behind $\mathcal{D}(\mu^2)$ is that a contribution to $|\rho(\mu^2)\rangle$ has poles from virtual loops and has singularities when some of its external lines become collinear or soft. This is illustrated in Fig. 1. It is simplest to think of the graphs depicted as being in a physical gauge. There are two hard subgraphs, represented as yellow blobs, one for the amplitude and one for the conjugate amplitude. The subgraphs can be tree graphs or can contain virtual loops. We suppose that everything inside the hard subgraphs is harder than a reference scale μ_{hard}^2 . That is, all of the internal propagators are far off shell. Two initial state lines and m final state parton lines emerge from the hard subgraph. Here $m = 3$. Then there are additional interactions. Some number n_r of additional partons are emitted and n_v parton lines are exchanged. Here $n_r = 1$ and $n_v = 1$. The external parton momenta are labeled $\{\hat{p}\}_{m+n_r}$. There is an infrared divergence when the virtual gluon becomes soft and there are singularities when the real gluon momentum becomes soft or collinear to the antiquark line while the virtual gluon momentum is becoming soft.

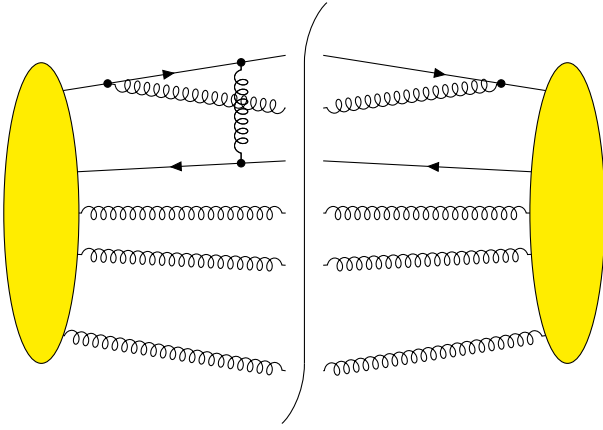


FIG. 1. An infrared singular diagram. Each yellow blob represents a graph, possibly with loops, in which everything is harder than the scale μ_{hard}^2 .

We want to capture the structure of these infrared singularities, as illustrated in Fig. 2. We note that near the singularities, the partons emerging from the hard subgraphs are nearly on shell and their momenta lie in the directions of the external parton momenta. The momenta carried by lines internal to the hard subgraphs are almost unchanged. Therefore, we can approximate the graph by letting the momenta $\{p\}_m$ of the partons emerging from the hard subgraphs be exactly on shell. Their momenta are given as functions of the momenta $\{\hat{p}\}_{m+n_R}$ of the external partons. Here, we need to define a momentum mapping $\{\hat{p}\}_{m+n_R} \rightarrow \{p\}_m$. Then we can approximate the original graph by a hard part $(\{p, f, s, s', c, c'\}_m | \rho_{\text{hard}}(\mu^2))$ and a singular factor. We call the singular factor $(\{\hat{p}, \hat{f}, \hat{s}, \hat{s}', \hat{c}, \hat{c}'\}_{m+n_R} | \mathcal{D}(\mu^2) | \{p, f, s, s', c, c'\}_m)$. The singular factor is derived just from the singular part of the graph. It is independent of what is in the hard part.

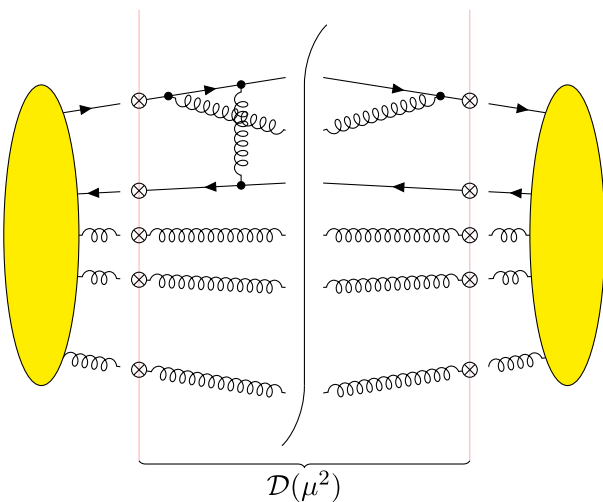


FIG. 2. The diagram in Fig. 1 after separating the hard diagram from $\mathcal{D}(\mu^2)$.

We thus assert that any such set of poles and singularities can be organized into a hard subgraph, $|\rho_{\text{hard}}(\mu^2)\rangle$, convolved with a singular factor:

$$\begin{aligned} & (\{\hat{p}, \hat{f}, \hat{s}, \hat{s}', \hat{c}, \hat{c}'\}_{m+n_R} | \rho(\mu^2)) \\ & \sim \frac{1}{m!} \int [d\{p\}_m] \sum_{\{f\}_m} \sum_{\{s, s', c, c'\}_m} \\ & \quad \times (\{\hat{p}, \hat{f}, \hat{s}, \hat{s}', \hat{c}, \hat{c}'\}_{m+n_R} | \mathcal{D}(\mu^2) | \{p, f, s, s', c, c'\}_m) \\ & \quad \times (\{p, f, s, s', c, c'\}_m | \rho_{\text{hard}}(\mu^2)). \end{aligned} \quad (46)$$

The division between singular and hard factors depends on the singularity to be examined. In the hard factor, the external momenta $\{p\}_m$ and any internal loop momenta are to be hard at some scale that we can call μ_{hard}^2 . This means that they are not closely collinear to each other or soft at scales softer than μ_{hard}^2 .

The singular factor is typically represented as separate factors labeled *soft* and *jet*, where *soft* can include Glauber exchanges and two of the jets are in the beam directions. However, we do not need to separate the singular factor into separate subfactors. An early and instructive analysis of the singularities of QCD was given by Libby and Sterman [118]. An extensive modern analysis can be found in Collins [119]. For one real gluon emission, an example of $\mathcal{D}(\mu^2)$ can be defined from the Catani-Seymour dipole splitting functions [34]. The operators $\mathcal{S}p$ of Catani, de Florian, and Rodrigo [120] are also closely related to $\mathcal{D}(\mu^2)$ for certain cases. At one loop, our version of $d\mathcal{D}(\mu^2)/d \log \mu^2$ is implemented in DEDUCTOR.

The operator $\mathcal{D}(\mu^2)$ has a perturbative expansion

$$\mathcal{D}(\mu^2) = 1 + \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \mathcal{D}^{(n)}(\mu^2) + \mathcal{O}(\alpha_s^{k+1}). \quad (47)$$

The order n contribution, $\mathcal{D}^{(n)}(\mu^2)$, is a sum of infrared sensitive operators,

$$\mathcal{D}^{(n)}(\mu^2) = \sum_{n_R=0}^n \sum_{n_V=0}^n \theta(n_R + n_V = n) \mathcal{D}^{(n_R, n_V)}(\mu^2). \quad (48)$$

Acting on a state $|\{p, f, s, s', c, c'\}_m\rangle$ with m final state QCD partons, $\mathcal{D}^{(n_R, n_V)}(\mu^2)$ produces a state with $m + n_R$ final state QCD partons with momenta $\{\hat{p}\}_{m+n_R}$. There are integrations over the loop momenta $\{\ell_1, \dots, \ell_{n_V}\}$ of n_V virtual loops.

The operator $\mathcal{D}(\mu^2)$ depends on a hardness scale μ_s^2 that defines an infrared sensitive region $R(\mu_s^2)$ in the space of the momenta $\{\hat{p}\}_{m+n_R}$ and $\{\ell_1, \dots, \ell_{n_V}\}$. We can think of μ_s^2 as being comparable to the scale μ_{hard}^2 of $|\rho_{\text{hard}}(\mu^2)\rangle$. The infrared sensitive region $R(\mu_s^2)$ surrounds the leading singularity, at which each of the momenta $\{\hat{p}\}_{m+n_R}$

$\{\ell_1, \dots, \ell_{n_v}\}$ is soft or collinear to one of the input momenta $\{p\}_m$. In the case of the output momenta $\{\hat{p}\}_{m+n_R}$, this means that, at the leading singularity, these momenta form m infinitely narrow jets with momenta $\{p\}_m$. Of course, if we look just a little bit away from the limit of infinitely narrow jets, we see that the jets can have subjects. The singularity structure of the subjects, including both real and virtual momenta, is included in $\mathcal{D}(\mu^2)$.

We need the infrared sensitive region $R(\mu_s^2)$ because, when we form $\mathcal{D}(\mu^2)$ by making approximations that apply near the leading singularity, we necessarily simplify the behavior away from this singularity. We introduce cuts such that $\mathcal{D}^{(n_R, n_v)}(\mu^2)$ gets contributions only from inside $R(\mu_s^2)$. The leading singularity is inside the region for any μ_s^2 , but for larger μ_s^2 , the region is larger. Of course, there is more than one way to introduce μ_s^2 .

For simplicity, we set the renormalization and factorization scale μ^2 equal to μ_s^2 .

We will say more about the $\mathcal{D}^{(n)}(\mu^2)$ later, although we do not construct them. For now we simply assume that they are available and investigate how they can be used to construct subtractions for a fixed order calculation and splitting functions for a parton shower. See Appendix A for an example of the operators $\mathcal{D}^{(n)}(\mu^2)$ in a toy model.

D. Subtractions for the perturbative cross section

Given $\mathcal{D}(\mu^2)$, we can construct

$$\mathcal{D}^{-1}(\mu^2) = 1 - \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \tilde{\mathcal{D}}^{(n)}(\mu^2) + \mathcal{O}(\alpha_s^{k+1}). \quad (49)$$

The perturbative coefficients $\tilde{\mathcal{D}}^{(n)}(\mu^2)$ are defined by

$$\mathcal{D}^{-1}(\mu^2)\mathcal{D}(\mu^2) = 1. \quad (50)$$

This gives, for instance,

$$\begin{aligned} \tilde{\mathcal{D}}^{(1)}(\mu^2) &= \mathcal{D}^{(1)}(\mu^2), \\ \tilde{\mathcal{D}}^{(2)}(\mu^2) &= \mathcal{D}^{(2)}(\mu^2) - \mathcal{D}^{(1)}(\mu^2)\mathcal{D}^{(1)}(\mu^2), \end{aligned} \quad (51)$$

or for the higher orders

$$\tilde{\mathcal{D}}^{(n)}(\mu^2) = \mathcal{D}^{(n)}(\mu^2) - \sum_{k=1}^{n-1} \mathcal{D}^{(k)}(\mu^2)\tilde{\mathcal{D}}^{(n-k)}(\mu^2). \quad (52)$$

The order n contribution, $\tilde{\mathcal{D}}^{(n)}(\mu^2)$, is a sum of operators,

$$\tilde{\mathcal{D}}^{(n)}(\mu^2) = \sum_{n_R=0}^n \sum_{n_V=0}^n \theta(n_R + n_V = n) \tilde{\mathcal{D}}^{(n_R, n_V)}(\mu^2). \quad (53)$$

Acting on a state $|\{p, f, s, s', c, c'\}_m\rangle$ with m partons, $\tilde{\mathcal{D}}^{(n_R, n_V)}(\mu^2)$ produces a state with $m + n_R$ partons while integrating over n_V virtual loops. One constructs $\tilde{\mathcal{D}}^{(n_R, n_V)}(\mu^2)$ using Eq. (50).

The operator $\mathcal{D}^{-1}(\mu^2)$ is very useful. In Eq. (45), we have the factor $\mathcal{D}^{-1}(\mu^2)\mathcal{O}_J|\rho(\mu^2)\rangle$. We are to expand this product in powers of α_s , keeping the terms up to order α_s^k . The factor $\mathcal{O}_J|\rho(\mu^2)\rangle$ has infrared singularities and poles, but the operator $\mathcal{D}^{-1}(\mu^2)$ removes them. The simple argument is that $\mathcal{D}(\mu^2)|\{p, f, s, s', c, c'\}_m\rangle$ contains the infrared singularities and poles produced by QCD from a hard parton state $|\{p, f, s, s', c, c'\}_m\rangle$ but, according to Eq. (50), when we apply $\mathcal{D}^{-1}(\mu^2)$ to this state, the poles and singular terms are cancelled. That is, $\mathcal{D}^{-1}(\mu^2)$ provides the subtraction terms that we need to remove the singularities and poles from an N^kLO perturbative calculation.

To understand what the operator $\mathcal{D}^{-1}(\mu^2)$ does, it is helpful to examine the familiar case of an NLO calculation. At this order, Eq. (45) becomes

$$\begin{aligned} \sigma[J] &= (1|[\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{Z}_F(\mu^2)]\mathcal{D}(\mu^2) \left\{ \mathcal{O}_J|\rho^{(0,0)}(\mu^2)\rangle \right. \\ &\quad + \frac{\alpha_s(\mu^2)}{2\pi} [\mathcal{O}_J|\rho^{(0,1)}(\mu^2)\rangle - \mathcal{D}^{(0,1)}(\mu^2)\mathcal{O}_J|\rho^{(0,0)}(\mu^2)\rangle] \\ &\quad \left. + \frac{\alpha_s(\mu^2)}{2\pi} [\mathcal{O}_J|\rho^{(1,0)}(\mu^2)\rangle - \mathcal{D}^{(1,0)}(\mu^2)\mathcal{O}_J|\rho^{(0,0)}(\mu^2)\rangle] \right\} \\ &\quad + \mathcal{O}(\alpha_s^2). \end{aligned} \quad (54)$$

We have defined $\mathcal{D}^{(0,1)}(\mu^2)$ so that it leaves the number of partons unchanged and so that it has infrared poles. Furthermore, the poles in $\mathcal{D}^{(0,1)}(\mu^2)|\rho^{(0,0)}\rangle$ should directly cancel those of $|\rho^{(0,1)}(\mu^2)\rangle$. We have defined $\mathcal{D}^{(1,0)}(\mu^2)$ so that, acting on the state $|\rho^{(0,0)}(\mu^2)\rangle$, it adds one parton and so that when this parton is soft or nearly collinear with one of the existing partons (in our example, the initial state partons) $\mathcal{D}^{(1,0)}(\mu^2)\mathcal{O}_J|\rho^{(0,0)}(\mu^2)\rangle$ approaches $\mathcal{O}_J|\rho^{(1,0)}(\mu^2)\rangle$. In a standard application, one performs the integrations over the momentum of the emitted parton numerically. The integrand in the subtraction cancels the integrand in $|\rho^{(1,0)}(\mu^2)\rangle$ in the infrared region, so that one obtains a convergent integration. Having subtracted the operators $\mathcal{D}^{(1,0)}$ and $\mathcal{D}^{(0,1)}$, we add them back as part of \mathcal{D} . Now, in a standard application, all of the integrations corresponding to the first line of Eq. (54) are performed analytically. All of the $1/\epsilon^2$ and $1/\epsilon$ poles cancel and we are left with a completely finite order α_s contribution to the cross section. Note that the contribution from the first line beyond just the parton distribution functions is infrared finite, but it is not zero. It forms a significant part of the NLO calculation. In Eq. (54), we have a product of one infrared finite object times another, each expanded to order α_s . In the customary NLO calculation, one expands the product to order α_s and drops the α_s^2 term, but one could keep the α_s^2 term if desired.

We note that if we had wanted to use $\mathcal{D}(\mu^2)$ for the single purpose of defining subtractions for the hard scattering, $|\rho(\mu^2)\rangle$, we could have used a fixed scale. We need an adjustable scale μ_s^2 to use $\mathcal{D}(\mu^2)$ in a shower algorithm

because the hardness scale of the shower changes as the shower progresses.

E. Properties of the infrared sensitive operator

We can now say a little more about the infrared sensitive operator $\mathcal{D}(\mu^2)$, without giving a detailed specification. This operator is decomposed into operators $\mathcal{D}^{(n_r, n_v)}(\mu^2)$ according to Eqs. (47) and (48). Acting on a state $|\{p, f, s, s', c, c'\}_m\rangle$ with m partons, $\mathcal{D}^{(n_r, n_v)}(\mu^2)$ produces a state with $m + n_r$ partons while adding n_v virtual loops. The resulting state can be expanded in basis states for $m + n_r$ partons, $|\{\hat{p}, \hat{f}, \hat{s}, \hat{s}', \hat{c}, \hat{c}'\}_{m+n_r}\rangle$. There is then an invertible mapping between the new momenta $\{\hat{p}\}_{m+n_r}$ and the starting momenta $\{p\}_m$ together with a set of splitting variables $\zeta_p(n_r)$. One can choose what this mapping is.

To make this a little more concrete, consider $\mathcal{D}^{(1,0)}(\mu^2)$ with one parton emitted. There is one term in the emission probability for each final state parton $l \in \{1, \dots, m\}$ and one for each initial state parton $l \in \{a, b\}$. We think of l as the emitting parton and let the emission probability be singular when \hat{p}_{m+1} becomes collinear with \hat{p}_l . The emission probabilities are also singular in the limit in which \hat{p}_{m+1} becomes soft, $\hat{p}_{m+1} \rightarrow 0$. DEDUCTOR uses something similar to the Catani-Seymour [34] dipole splitting functions to construct $\mathcal{D}^{(1,0)}(\mu^2)$. The splitting variables ζ_p are taken to be an azimuthal angle ϕ , a momentum fraction z , and a measure of the hardness of the splitting. In DEDUCTOR, the hardness variable is the virtuality of the splitting divided by the energy of the mother parton,

$$\Lambda^2 \equiv \frac{2\hat{p}_l \cdot \hat{p}_{m+1}}{2p_l \cdot Q_H} Q_H^2, \quad (55)$$

where the vector Q_H is defined globally as described in Sec. IV C. Alternatively, it can be defined dynamically as described in Appendix C.

There is freedom to choose the functional form of $\mathcal{D}^{(1,0)}(\mu^2)$ away from the limits of soft and collinear emissions.

There is also freedom to choose the momentum mapping. The simplest case is a splitting of a parton l into two partons l and $m + 1$. Then we cannot have p_l be the same as $\hat{p}_l + \hat{p}_{m+1}$ with $p_l^2 = \hat{p}_l^2 = \hat{p}_{m+1}^2 = 0$, so the momentum mapping has to take a some momentum from the other partons and supply it to $\hat{p}_l + \hat{p}_{m+1}$. In DEDUCTOR, we use a global mapping, taking a small amount of momentum from each of the other partons. A second possibility beyond a simple splitting is interference between emission of a gluon $m + 1$ from parton l_L in the ket state and emission from another parton l_R in the bra state. In this case, $\mathcal{D}^{(1,0)}(\mu^2)$ in DEDUCTOR is a linear combination of contributions that use the momentum mappings for a splitting of parton l_L and for a splitting of parton l_R . The coefficients in the linear combination are a ‘‘dipole partitioning’’ function A' that is specified in DEDUCTOR.

The operator $\mathcal{D}^{(1,0)}(\mu^2)$, acting on a state $|\{p, f, s, s', c, c'\}_m\rangle$, produces a state with one more parton, parton $m + 1$. It is crucial that there be an ultraviolet cutoff for \hat{p}_{m+1} . The cutoff is specified by a parameter that we call μ_s^2 . In DEDUCTOR, we use Λ^2 as given in Eq. (55) to define the cutoff. In $\mathcal{D}^{(1,0)}(\mu^2)$, we require

$$\Lambda^2 < \mu_s^2. \quad (56)$$

A similar cutoff applies inside the integration for a virtual loop in $\mathcal{D}^{(0,1)}(\mu^2)$. Defining this cutoff is more involved than we can review here. The calculations are described in Refs. [32,33].

VI. FROM THE PERTURBATIVE CROSS SECTION TO A PARTON SHOWER

In this section, we begin with Eq. (45) for the perturbative cross section. We set the scale to μ_H^2 , which we take to be equal to Q_H^2 . Here Q_H^2 is defined in Sec. IV C to be a fixed vector, although we can use a dynamical definition as described in Appendix C. We now seek a more powerful formulation that will enable us to use more general measurement operators \mathcal{O}_J for which a perturbative expansion of the cross section might contain large logarithms of the generic form $\alpha_s^n \log^{2n}(k^2/Q^2)$. Often, a parton shower can approximately sum such logarithms.

As just stated, we set the scale μ^2 in Eq. (45) to μ_H^2 . In our example in which the Born process is Higgs boson production, we might choose $\mu_H^2 = m_H^2$. This affects the scale at which α_s and the parton distribution functions are evaluated. It also affects the upper cutoff on the scale of emissions in the subtraction terms in $\mathcal{D}^{-1}(\mu_H^2)$ and in the parton shower that we will discuss below. However, in the multiparton matrix elements in $|\rho(\mu_H^2)\rangle$, the partons can have any momenta. Parton emissions with small transverse momenta need subtractions, but parton emissions with very large transverse momenta do not need subtractions. If we work at order α_s^k , then we can have up to k high transverse momentum jets in addition to the Higgs boson in $|\rho(\mu_H^2)\rangle$.¹⁰

A. Moving the measurement operator

The first step toward a more general formulation is to interchange the order of the measurement operator \mathcal{O}_J and the operators $\mathcal{D}^{-1}(\mu_H^2)$ and $\mathcal{D}(\mu_H^2)$. This does not change the result, since $\mathcal{D}\mathcal{D}^{-1} = 1$ and \mathcal{O}_J commutes with 1. With the \mathcal{O}_J moved, we have

¹⁰If we are limited to a $k = 1$ shower, then Higgs plus one jet is LO in $|\rho(\mu_H^2)\rangle$. Then one might want define a lower cutoff on the P_T of the jet and use a calculation with $p + p \rightarrow H + J$ as the Born process, calculated at NLO. In that case, one has two calculations and one may want to define a procedure to merge them. If a $k = 2$ shower is available, then $|\rho(\mu_H^2)\rangle$ can include inclusive Higgs production at NNLO and Higgs plus one jet at NLO. Then merging different calculations is less needed.

$$\sigma[J] = (1[|\mathcal{F}_{\overline{\text{MS}}}(\mu_{\text{H}}^2) \circ \mathcal{Z}_F(\mu_{\text{H}}^2)]\mathcal{O}_J\mathcal{D}(\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}). \quad (57)$$

Here we have denoted

$$|\rho_{\text{H}}) = \mathcal{D}^{-1}(\mu_{\text{H}}^2)|\rho(\mu_{\text{H}}^2)), \quad (58)$$

where the product is expanded to N^kLO using Eqs. (43), (44), (49), and (53):

$$|\rho_{\text{H}}) = \sum_{n=0}^k \left[\frac{\alpha_s(\mu_{\text{H}}^2)}{2\pi} \right]^n \sum_{n_{\text{R}}=0}^n \sum_{n_{\text{V}}=0}^n \theta(n_{\text{R}} + n_{\text{V}} = n) |\rho_{\text{H}}^{(n_{\text{R}}, n_{\text{V}})}) + \mathcal{O}(\alpha_s^{k+1}), \quad (59)$$

where

$$|\rho_{\text{H}}^{(n_{\text{R}}, n_{\text{V}})}) = |\rho^{(n_{\text{R}}, n_{\text{V}})}(\mu_{\text{H}}^2)) - \sum_{r=0}^{n_{\text{R}}} \sum_{l=0}^{n_{\text{V}}} \tilde{\mathcal{D}}^{(r,l)}(\mu_{\text{H}}^2) |\rho^{(n_{\text{R}}-r, n_{\text{V}}-l)}(\mu_{\text{H}}^2)). \quad (60)$$

These quantities $|\rho_{\text{H}}^{(n_{\text{R}}, n_{\text{V}})})$ are finite without dimensional regularization.

B. Introducing shower oriented parton distribution functions

We can do a little more by introducing an operator $\mathcal{F}(\mu_{\text{H}}^2)$ that multiplies by parton distribution functions $f_{a/A}(\eta_a, \mu^2) f_{b/B}(\eta_b, \mu^2)$ and a luminosity factor. However, these are not the five-flavor $\overline{\text{MS}}$ parton distribution functions that we used in $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$. Rather, they are adapted to the choice of the definition for $\mathcal{D}(\mu^2)$ that we use.

The shower oriented parton operator $\mathcal{F}(\mu^2)$ is related to $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$ by factor $\mathcal{K}(\mu^2)$,

$$\mathcal{F}_{\overline{\text{MS}}}(\mu^2) = [\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2)]. \quad (61)$$

This is a rather compact notation, so it is worthwhile to write it in more detail. The left-hand side is defined by

$$\mathcal{F}_{\overline{\text{MS}}}(\mu^2)|\{p, f, s, s', c, c'\}_m) = \frac{f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2) f_{b/B}^{\overline{\text{MS}}}(\eta_b, \mu^2)}{n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b} |\{p, f, s, s', c, c'\}_m). \quad (62)$$

The right-hand side is

$$[\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2)]|\{p, f, s, s', c, c'\}_m) = \sum_{a', b'} \int_0^1 \frac{dz_a}{z_a} \int_0^1 \frac{dz_b}{z_b} \frac{f_{a'/A}(\eta_a/z_a, \mu^2) f_{b'/B}(\eta_b/z_b, \mu^2)}{n_c(a)n_s(a)n_c(b)n_s(b)4p_a \cdot p_b} \times K_{a,a'}^{(a)}(z_a, \mu^2, \{p, f\}_m) K_{b,b'}^{(b)}(z_b, \mu^2, \{p, f\}_m) |\{p, f, s, s', c, c'\}_m). \quad (63)$$

Here we take the kernel K to be a product, so that each of the two parton distributions is transformed separately. We allow each kernel to depend on the momentum and flavor variables of the parton state to which $[\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2)]$ is applied. The kernels each have a perturbative expansion beginning with

$$K_{a,a'}^{(a)}(z, \mu^2, \{p, f\}_m) = \delta_{a,a'} \delta(1-z) + \frac{\alpha_s(\mu^2)}{2\pi} K_{a,a'}^{(a,1)}(z, \mu^2, \{p, f\}_m) + \mathcal{O}(\alpha_s^2). \quad (64)$$

The choice of $\mathcal{K}(\mu^2)$ defines the shower-oriented parton distribution functions. The evolution of these parton distribution functions needs to be matched to the parton splitting functions introduced in the following sections. In particular, the choice of $\mathcal{K}(\mu^2)$ is largely determined by

the definition of the cutoff μ_s^2 that we use for the shower. We provide an example in Appendix B.

The parton operators $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$ and $\mathcal{F}(\mu^2)$ obey evolution equations

$$\mu^2 \frac{d}{d\mu^2} \mathcal{F}(\mu^2) = [\mathcal{F}(\mu^2) \circ \mathcal{P}(\mu^2)],$$

$$\mu^2 \frac{d}{d\mu^2} \mathcal{F}_{\overline{\text{MS}}}(\mu^2) = [\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{P}_{\overline{\text{MS}}}(\mu^2)]. \quad (65)$$

Using Eq. (61), we see that the evolution kernels are related by

$$\mathcal{P}(\mu^2) = [\mathcal{K}(\mu^2) \circ \mathcal{P}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{K}^{-1}(\mu^2)] - \left[\left(\mu^2 \frac{d}{d\mu^2} \mathcal{K}(\mu^2) \right) \circ \mathcal{K}^{-1}(\mu^2) \right]. \quad (66)$$

With the transformation from $\mathcal{F}_{\overline{\text{MS}}}(\mu^2)$ to $\mathcal{F}(\mu^2)$, we write

$$[\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{Z}_F(\mu^2)] = [\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2) \circ \mathcal{Z}_F(\mu^2)]. \quad (67)$$

Thus our cross section is

$$\sigma[J] = (1|[\mathcal{F}(\mu_{\text{H}}^2) \circ \mathcal{K}(\mu_{\text{H}}^2) \circ \mathcal{Z}_F(\mu_{\text{H}}^2)] \mathcal{O}_J \mathcal{D}(\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}). \quad (68)$$

We will introduce $\mathcal{F}(\mu^2)$ into another place in the formalism shortly.

C. Changing the scale of the subtraction operators

Next, we would like to change the scale of the operators in Eq. (68) from a large scale μ_{H}^2 to something smaller. We let μ_1^2 be an ‘‘intermediate’’ size scale that is much smaller than the scale $Q^2[J]$ associated with the operator \mathcal{O}_J but is nevertheless large compared to 1 GeV^2 and is certainly large enough to allow the use of perturbation theory in $\alpha_s(\mu_1^2)$. We can change the parton factor to be evaluated at scale μ_1^2 because this factor is a renormalization group invariant:

$$[\mathcal{F}(\mu_{\text{H}}^2) \circ \mathcal{K}(\mu_{\text{H}}^2) \circ \mathcal{Z}_F(\mu_{\text{H}}^2)] = [\mathcal{F}(\mu_1^2) \circ \mathcal{K}(\mu_1^2) \circ \mathcal{Z}_F(\mu_1^2)]. \quad (69)$$

The operator $\mathcal{D}(\mu^2)$ is not invariant under changes of scale. However, we can write

$$\mathcal{D}(\mu_1^2) = \mathcal{D}(\mu_2^2) \mathcal{U}_{\text{pert}}(\mu_2^2, \mu_1^2), \quad (70)$$

where

$$\mathcal{U}_{\text{pert}}(\mu_2^2, \mu_1^2) = \mathcal{D}^{-1}(\mu_2^2) \mathcal{D}(\mu_1^2). \quad (71)$$

Here we note that $\mathcal{D}(\mu_1^2)$ generates $1/\epsilon$ poles and infrared singularities, but $\mathcal{D}^{-1}(\mu_2^2)$ provides the proper subtractions to remove the poles and infrared singularities when we expand the product of operators to a fixed order of perturbation theory. Thus we can evaluate $\mathcal{U}_{\text{pert}}(\mu_2^2, \mu_1^2)$ in four dimensions instead of $4 - 2\epsilon$ dimensions. The perturbative evolution operator $\mathcal{U}_{\text{pert}}(\mu^2, \mu'^2)$ obeys the differential equation

$$\mu^2 \frac{d}{d\mu^2} \mathcal{U}_{\text{pert}}(\mu^2, \mu'^2) = -S_{\text{pert}}(\mu^2) \mathcal{U}_{\text{pert}}(\mu^2, \mu'^2), \quad (72)$$

where

$$S_{\text{pert}}(\mu^2) = \mathcal{D}^{-1}(\mu^2) \mu^2 \frac{d}{d\mu^2} \mathcal{D}(\mu^2). \quad (73)$$

Since $\mathcal{U}_{\text{pert}}(\mu_2^2, \mu_1^2)$ is infrared finite, so is $S_{\text{pert}}(\mu^2)$. We can write the solution of Eq. (72) as

$$\mathcal{U}_{\text{pert}}(\mu^2, \mu'^2) = \mathbb{T} \exp \left(\int_{\mu'^2}^{\mu^2} \frac{d\mu^2}{\mu^2} S_{\text{pert}}(\mu^2) \right), \quad (74)$$

where \mathbb{T} indicates μ^2 ordering of the exponential with smaller μ^2 to the left. Working to order α_s with use of Eqs. (47) and (48), we have

$$S_{\text{pert}}(\mu^2) = \frac{\alpha_s(\mu^2)}{2\pi} S_{\text{pert}}^{(1,0)}(\mu^2) + \frac{\alpha_s(\mu^2)}{2\pi} S_{\text{pert}}^{(0,1)}(\mu^2) + \mathcal{O}(\alpha_s^2), \quad (75)$$

where $\alpha_s(\mu^2)$ is the running coupling in the four dimensional theory and

$$\begin{aligned} S_{\text{pert}}^{(1,0)}(\mu^2) &= \mu^2 \frac{d}{d\mu^2} \mathcal{D}^{(1,0)}(\mu^2), \\ S_{\text{pert}}^{(0,1)}(\mu^2) &= \mu^2 \frac{d}{d\mu^2} \mathcal{D}^{(0,1)}(\mu^2). \end{aligned} \quad (76)$$

Recall that $\mu^2 = \mu_s^2$. Thus $S_{\text{pert}}^{(1,0)}(\mu^2)$ is the derivative of (approximated) real emission graphs with respect to the ultraviolet cutoff that we impose. Similarly, $S_{\text{pert}}^{(0,1)}(\mu^2)$ is the derivative of approximated one loop virtual graphs with respect to the ultraviolet cutoff. The subscript ‘‘pert’’ emphasizes that only perturbative Feynman diagrams are used to obtain $S_{\text{pert}}^{(1,0)}(\mu^2)$ and $S_{\text{pert}}^{(0,1)}(\mu^2)$.

With these changes, we have

$$\begin{aligned} \sigma[J] &= (1|[\mathcal{F}(\mu_1^2) \circ \mathcal{K}(\mu_1^2) \circ \mathcal{Z}_F(\mu_1^2)] \\ &\quad \times \mathcal{O}_J \mathcal{D}(\mu_1^2) \mathcal{U}_{\text{pert}}(\mu_1^2, \mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}). \end{aligned} \quad (77)$$

Now we note that very soft or collinear splittings at scales much smaller than $Q[J]^2$ are not resolved by the measurement operator \mathcal{O}_J . The operator $\mathcal{D}(\mu_1^2)$ generates splittings at scales μ_1^2 and smaller. Since we have chosen $\mu_1^2 \ll Q[J]^2$, the operator \mathcal{O}_J commutes with $\mathcal{D}(\mu_1^2)$ to a good approximation, with an error of order $\mu_1^2/Q[J]^2$. Thus Eq. (68) can be written as

$$\begin{aligned} \sigma[J] &= (1|[\mathcal{F}(\mu_1^2) \circ \mathcal{K}(\mu_1^2) \circ \mathcal{Z}_F(\mu_1^2)] \\ &\quad \times \mathcal{D}(\mu_1^2) \mathcal{O}_J \mathcal{U}_{\text{pert}}(\mu_1^2, \mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) + \mathcal{O}(\mu_1^2/Q[J]^2). \end{aligned} \quad (78)$$

D. The inclusive infrared finite operator $\mathcal{V}(\mu^2)$

We now introduce an operator $\mathcal{X}(\mu^2)$ defined by

$$\mathcal{X}(\mu^2) = [\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2) \circ \mathcal{Z}_F(\mu^2)] \mathcal{D}(\mu^2) \mathcal{F}^{-1}(\mu^2). \quad (79)$$

Using $\mathcal{X}(\mu^2)$, Eq. (78) is more compact:

$$\sigma[J] = (1|\mathcal{X}(\mu_1^2)\mathcal{F}(\mu_1^2)\mathcal{O}_J\mathcal{U}_{\text{pert}}(\mu_1^2, \mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) \\ + \mathcal{O}(\mu_1^2/Q[J]^2). \quad (80)$$

The operator $\mathcal{X}(\mu^2)$ involves parton distribution functions and purely perturbative operators. If we evaluate the perturbative operators at order zero, we get simply $\mathcal{F}(\mu^2)\mathcal{F}^{-1}(\mu^2)$. Thus

$$\mathcal{X}(\mu^2) = 1 + \mathcal{O}(\alpha_s). \quad (81)$$

The operator $\mathcal{X}(\mu^2)$, when expanded in powers of α_s , creates partons, up to k partons at order α_s^k .

The operator $\mathcal{X}(\mu^2)$ is infrared sensitive. When we apply $\mathcal{X}(\mu^2)$ to a state $|\{p, f, s, s', c, c'\}_m\rangle$, we get a state $\mathcal{X}(\mu^2)|\{p, f, s, s', c, c'\}_m\rangle$ containing poles $1/\epsilon$ and singularities when the partons that $\mathcal{X}(\mu^2)$ creates become soft or collinear with other partons or with each other. In that sense, $\mathcal{X}(\mu^2)$ is like $\mathcal{D}(\mu^2)$. However, $\mathcal{X}(\mu^2)$ contains the parton factor $[\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2) \circ \mathcal{Z}_F(\mu^2)]$. This factor gives $\mathcal{X}(\mu^2)$ a property not shared by $\mathcal{D}(\mu^2)$. If we integrate over the momenta of the partons created by $\mathcal{X}(\mu^2)$ and sum over their colors and flavors by forming the inclusive sum $(1|\mathcal{X}(\mu^2)|\{p, f, s, s', c, c'\}_m)$, then the singularities cancel and we obtain a finite result.

In fact, we need to ensure that $(1|\mathcal{X}(\mu^2)|\{p, f, s, s', c, c'\}_m)$ is not only finite after dimensional regularization is removed but that it vanishes in the limit $\mu^2 \rightarrow 0$. For example, if this quantity arises from an integration

$$(1|\mathcal{X}(\mu^2)|\{p, f, s, s', c, c'\}_m) = \int_0^{\mu^2} \frac{dk^2}{k^2} G(k^2), \quad (82)$$

then, with subtractions included, $G(k^2)$ needs to be a smooth function that is well enough behaved for $k^2 \rightarrow 0$ that the integral is convergent. This property is needed later in Eq. (133).

Suppose for a moment that we worked in a modified theory, denoted by subscripts M, in which partons carried only momenta and flavors, but not color and spin. Then from the inclusive sum $(1|\mathcal{X}_M(\mu^2)|\{p, f\}_m)$ we could define another operator $\mathcal{V}_M(\mu^2)$ that leaves the number of partons, their momenta and flavors unchanged:

$$\mathcal{V}_M(\mu^2)|\{p, f\}_m) = \lambda(\{p, f\}_m)|\{p, f\}_m). \quad (83)$$

Then $\lambda(\{p, f\}_m, \mu^2) = (1|\mathcal{V}_M(\mu^2)|\{p, f\}_m)$. We define $\mathcal{V}_M(\mu^2)$ by Eq. (83) and

$$(1|\mathcal{V}_M(\mu^2)|\{p, f\}_m) = (1|\mathcal{X}_M(\mu^2)|\{p, f\}_m). \quad (84)$$

Now return to QCD. With spin and color, we can define an operator $\mathcal{V}(\mu^2)$ that satisfies $(1|\mathcal{V}(\mu^2) = (1|\mathcal{X}(\mu^2)$. However, its structure is more complex. The operator $\mathcal{X}(\mu^2)$ can be expanded in powers of α_s :

$$\mathcal{X}(\mu^2) = 1 + \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \mathcal{X}^{(n)}(\mu^2) + \mathcal{O}(\alpha_s^{k+1}). \quad (85)$$

The order n contribution, $\mathcal{X}^{(n)}(\mu^2)$, is a sum of infrared sensitive operators,

$$\mathcal{X}^{(n)}(\mu^2) = \sum_{n_R=0}^n \sum_{n_V=0}^n \theta(n_R + n_V = n) \mathcal{X}^{(n_R, n_V)}(\mu^2). \quad (86)$$

Acting on a state $|\{p, f, s, s', c, c'\}_m\rangle$ with m final state partons, $\mathcal{X}^{(n_R, n_V)}(\mu^2)$ produces a state with $m + n_R$ final state partons with momenta and flavors $\{\hat{p}, \hat{f}\}_{m+n_R}$. There are integrations over the loop momenta $\{\ell_1, \dots, \ell_{n_V}\}$ of n_V virtual loops.

We need to understand the color and spin structure of $\mathcal{X}^{(n_R, n_V)}(\mu^2)$. Suppose that we have constructed a basis of operators that act on the quantum spin \otimes color space and create a quantum spin \otimes color state for n_R more partons. We label the basis operators by an index i . A convenient choice would be

$$i = \{m, \{\hat{s}, \hat{c}\}_{m+n_R}, \{s, c\}_m\}. \quad (87)$$

Then we could let

$$\sigma_i^{(n_R)}|\{s', c'\}_{m'}\rangle \\ = \begin{cases} |\{\hat{s}, \hat{c}\}_{m+n_R}\rangle & m = m' \ \& \ \{s, c\}_m = \{s', c'\}_{m'} \\ 0 & \text{otherwise} \end{cases}. \quad (88)$$

Using these basis operators, we can expand $\mathcal{X}^{(n_R, n_V)}(\mu^2)$ as

$$\mathcal{X}^{(n_R, n_V)}(\mu^2) = \sum_{i,j} \mathcal{X}_{i,j}^{(n_R, n_V)}(\mu^2) \sigma_i^{(n_R)} \otimes \sigma_j^{(n_V)\dagger}. \quad (89)$$

Here $\mathcal{X}_{i,j}^{(n_R, n_V)}(\mu^2)$ is still an operator on the momentum and flavor part of the statistical space, which has basis vectors $|\{p, f\}_m\rangle$. In the case $n_R = 0$, this operator adds no partons and leaves the parton momenta and flavors $\{p, f\}_m$ unchanged.

Now we wish to define another operator $\mathcal{V}(\mu^2)$ with an expansion

$$\mathcal{V}(\mu^2) = 1 + \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \mathcal{V}^{(n)}(\mu^2) + \mathcal{O}(\alpha_s^{k+1}). \quad (90)$$

The order n contribution is to add no partons and leave the parton momenta and flavors $\{p, f\}_m$ unchanged, but it still can be a nontrivial operator on the spin \otimes color space

$$\mathcal{V}^{(n)}(\mu^2) = \sum_{i,j} \mathcal{V}_{i,j}^{(n)}(\mu^2) \sigma_i^{(0)} \otimes \sigma_j^{(0)\dagger}. \quad (91)$$

Here $\mathcal{V}_{i,j}^{(n)}(\mu^2)$ is still an operator on the momentum and flavor part of the statistical space. We want $\mathcal{V}(\mu^2)$ to be related to $\mathcal{X}(\mu^2)$ by

$$(1|\mathcal{V}(\mu^2) = (1|\mathcal{X}(\mu^2). \quad (92)$$

Thus we want

$$\begin{aligned} & \left(1\left|\sum_{i,j}\mathcal{V}_{i,j}^{(n)}(\mu^2)\sigma_i^{(0)}\otimes\sigma_j^{(0)\dagger}\right|\{p,f,s,s',c,c'\}_m\right) \\ &= \sum_{n_r=0}^n\sum_{i,j}(1|\mathcal{X}_{i,j}^{(n_r,n-n_r)}(\mu^2)\sigma_i^{(n_r)}\otimes\sigma_j^{(n_r)\dagger} \\ & \times |\{p,f,s,s',c,c'\}_m\rangle. \end{aligned} \quad (93)$$

This is the same as

$$\begin{aligned} & \sum_{i,j}\langle\{s',c'\}_m|\sigma_j^{(0)\dagger}\sigma_i^{(0)}|\{s,c\}_m\rangle(1|\mathcal{V}_{i,j}^{(n)}(\mu^2)|\{p,f\}_m) \\ &= \sum_{n_r=0}^n\sum_{i,j}\langle\{s',c'\}_m|\sigma_j^{(n_r)\dagger}\sigma_i^{(n_r)}|\{s,c\}_m\rangle \\ & \times (1|\mathcal{X}_{i,j}^{(n_r,n-n_r)}(\mu^2)|\{p,f\}_m). \end{aligned} \quad (94)$$

This needs to work for any choice of m -parton spin \otimes color states $\langle\{s',c'\}_m|$ and $|\{s,c\}_m\rangle$, so we need an identity of spin \otimes color operators,

$$\begin{aligned} & \sum_{i,j}\sigma_j^{(0)\dagger}\sigma_i^{(0)}(1|\mathcal{V}_{i,j}^{(n)}(\mu^2)|\{p,f\}_m) \\ &= \sum_{n_r=0}^n\sum_{i,j}\sigma_j^{(n_r)\dagger}\sigma_i^{(n_r)}(1|\mathcal{X}_{i,j}^{(n_r,n-n_r)}(\mu^2)|\{p,f\}_m). \end{aligned} \quad (95)$$

The right-hand side of Eq. (95) is an operator on the spin \otimes color space for m final state partons. On the left-hand side, the operators $\sigma_i^{(0)}$ form a basis for this space of operators, as do the operators $\sigma_j^{(0)\dagger}$, so the operators $\sigma_j^{(0)\dagger}\sigma_i^{(0)}$ span this space and are, in fact, over-complete. That is, one can always find coefficients $(1|\mathcal{V}_{i,j}^{(n)}(\mu^2)|\{p,f\}_m)$ so that we match the operator on the right-hand side. However, the choice is not unique. At order $n=1$, we have made a simple choice in DEDUCTOR. It is beyond our scope here to investigate what choices might be best at NLO, $n=2$.

Since $(1|\mathcal{X}(\mu^2)$ is infrared finite, Eq. (92) tells us that $\mathcal{V}(\mu^2)$ is infrared finite.

E. A more sophisticated shower evolution operator

Using Eq. (92), Eq. (80) becomes

$$\begin{aligned} \sigma[J] &= (1|\mathcal{V}(\mu_1^2)\mathcal{F}(\mu_1^2)\mathcal{O}_J\mathcal{U}_{\text{pert}}(\mu_1^2,\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) \\ & + \mathcal{O}(\mu_1^2/Q[J]^2). \end{aligned} \quad (96)$$

Since $\mathcal{V}(\mu_1^2)\mathcal{F}(\mu_1^2)$ does not change the number, momenta, or flavors of partons, it commutes with \mathcal{O}_J . Thus

$$\begin{aligned} \sigma[J] &= (1|\mathcal{O}_J\mathcal{V}(\mu_1^2)\mathcal{F}(\mu_1^2)\mathcal{U}_{\text{pert}}(\mu_1^2,\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) \\ & + \mathcal{O}(\mu_1^2/Q[J]^2). \end{aligned} \quad (97)$$

Now we can define the shower evolution operator that we need,

$$\mathcal{U}(\mu_2^2,\mu_1^2) = \mathcal{V}(\mu_2^2)\mathcal{F}(\mu_2^2)\mathcal{U}_{\text{pert}}(\mu_2^2,\mu_1^2)\mathcal{F}^{-1}(\mu_1^2)\mathcal{V}^{-1}(\mu_1^2). \quad (98)$$

With this definition, the cross section is

$$\begin{aligned} \sigma[J] &= (1|\mathcal{O}_J\mathcal{U}(\mu_1^2,\mu_{\text{H}}^2)\mathcal{V}(\mu_{\text{H}}^2)\mathcal{F}(\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) \\ & + \mathcal{O}(\mu_1^2/Q[J]^2). \end{aligned} \quad (99)$$

This moves \mathcal{F} next to $|\rho_{\text{H}}\rangle$ so that at the hard interaction we have the proper factors to make a cross section. It also moves \mathcal{V} next to $|\rho_{\text{H}}\rangle$. We will see later what the consequences of this are.

In Eq. (99), we use a scale μ_1^2 that was left undefined except that it should be small compared to $Q^2[J]$ (which was the scale of \mathcal{O}_J) and should be large enough to allow the use of perturbation theory with coupling $\alpha_s(\mu_1^2)$. Our cross section is independent of the value of μ_1^2 . Let us now fix on a standard choice near the lower end of this range. We take $\mu_1^2 \rightarrow \mu_{\text{f}}^2$, where μ_{f}^2 is on the order of 1 GeV². Then

$$\begin{aligned} \sigma[J] &= (1|\mathcal{O}_J\mathcal{U}(\mu_{\text{f}}^2,\mu_{\text{H}}^2)\mathcal{V}(\mu_{\text{H}}^2)\mathcal{F}(\mu_{\text{H}}^2)|\rho_{\text{H}}) + \mathcal{O}(\alpha_s^{k+1}) \\ & + \mathcal{O}(\mu_{\text{f}}^2/Q[J]^2). \end{aligned} \quad (100)$$

We can write $\mathcal{U}(\mu_2^2,\mu_1^2)$ in a simpler form. We note that, using Eqs. (71) and (79),

$$\begin{aligned} \mathcal{U}(\mu_2^2,\mu_1^2) &= \mathcal{V}(\mu_2^2)\mathcal{F}(\mu_2^2)\mathcal{D}^{-1}(\mu_2^2)\mathcal{D}(\mu_1^2)\mathcal{F}^{-1}(\mu_1^2)\mathcal{V}^{-1}(\mu_1^2) \\ &= \mathcal{V}(\mu_2^2)\mathcal{X}^{-1}(\mu_2^2)[\mathcal{F}(\mu_2^2)\circ\mathcal{K}(\mu_2^2)\circ\mathcal{Z}_F(\mu_2^2)] \\ & \times [\mathcal{F}(\mu_1^2)\circ\mathcal{K}(\mu_1^2)\circ\mathcal{Z}_F(\mu_1^2)]^{-1}\mathcal{X}(\mu_1^2)\mathcal{V}^{-1}(\mu_1^2). \end{aligned} \quad (101)$$

Since the operator $[\mathcal{F}(\mu^2)\circ\mathcal{K}(\mu^2)\circ\mathcal{Z}_F(\mu^2)]$ is independent of scale, this is

$$\mathcal{U}(\mu_2^2,\mu_1^2) = \mathcal{V}(\mu_2^2)\mathcal{X}^{-1}(\mu_2^2)\mathcal{X}(\mu_1^2)\mathcal{V}^{-1}(\mu_1^2). \quad (102)$$

F. Probability preservation in $\mathcal{U}(\mu_2^2,\mu_1^2)$

The operator $\mathcal{U}(\mu_2^2,\mu_1^2)$ has an important property, which we now derive. From Eq. (102), we have

$$(1|\mathcal{U}(\mu_2^2,\mu_1^2) = (1|\mathcal{V}(\mu_2^2)\mathcal{X}^{-1}(\mu_2^2)\mathcal{X}(\mu_1^2)\mathcal{V}^{-1}(\mu_1^2). \quad (103)$$

Then using Eq. (92) twice, we have

$$(1|\mathcal{U}(\mu_2^2, \mu_1^2) = (1|. \quad (104)$$

Multiplying any statistical state $|\rho\rangle$ by $(1|$ gives the total probability associated with that state. Thus Eq. (104) says that shower evolution as represented in $\mathcal{U}(\mu_2^2, \mu_1^2)$ is probability preserving. Current parton shower algorithms are typically constructed to have this property. Here probability preservation is a derived property.

G. Factorization

Let $Q^2[J]$ be the smallest scale associated with the measurement operator \mathcal{O}_J in Eq. (100), as discussed in Sec. III A. If $Q^2[J]$ is close to the scale μ_H^2 of the hard process with which we start the shower, then measuring $\sigma[J]$ does not make use of the full power of a parton shower. Suppose now that $Q^2[J] \gg 1 \text{ GeV}^2$ but that $Q^2[J] \ll \mu_H^2$. Then perturbation theory for $\sigma[J]$ is still applicable, but there may be large logarithms, $\log(\mu_H^2/Q^2[J])$ in the perturbative expansion of $\sigma[J]$. In many cases, a parton shower is useful for summing such logarithms. It takes a dedicated analysis to show that a given parton shower algorithm does sum the logarithms associated with a given operator \mathcal{O}_J , but there is at least a chance that if we use a parton shower we will do better than if we simply use fixed order perturbation theory. Thus we consider this sort of measurement operator and examine how factorization works when $Q^2[J] \ll \mu_H^2$.

We argued in Sec. III A that shower splittings at scale μ^2 can change the measurement by a fraction $\mu^2/Q^2[J]$. We can neglect these modifications as long as μ^2 is small enough, say

$$\mu^2 < \epsilon_s Q^2[J]. \quad (105)$$

We want ϵ_s to be small enough that we can regard fractional errors of order ϵ_s as negligible. However, we may want $\log(1/\epsilon_s)$ not to be large.

We can use our knowledge of the scale of \mathcal{O}_J by writing

$$\mathcal{U}(\mu_f^2, \mu^2) = \mathcal{U}(\mu_f^2, \epsilon_s Q^2[J]) \mathcal{U}(\epsilon_s Q^2[J], \mu^2). \quad (106)$$

Then writing

$$\mathcal{O}_J \mathcal{U}(\mu_f^2, \epsilon_s Q^2[J]) \approx \mathcal{U}(\mu_f^2, \epsilon_s Q^2[J]) \mathcal{O}_J \quad (107)$$

results in a negligible error. With this substitution, Eq. (100) becomes

$$\begin{aligned} \sigma[J] &= (1|\mathcal{U}(\mu_f^2, \epsilon_s Q^2[J]) \mathcal{O}_J \mathcal{U}(\epsilon_s Q^2[J], \mu_H^2) \mathcal{V}(\mu_H^2) \mathcal{F}(\mu_H^2)|\rho_H) \\ &\quad + \mathcal{O}(\alpha_s^{k+1}) + \mathcal{O}(\epsilon_s). \end{aligned} \quad (108)$$

Now, factorization for the cross section measured by \mathcal{O}_J requires that splittings at scales smaller than $\epsilon_s Q^2[J]$ not affect the cross section. Thus we need

$$\begin{aligned} \sigma[J] &= (1|\mathcal{O}_J \mathcal{U}(\epsilon_s Q^2[J], \mu_H^2) \mathcal{V}(\mu_H^2) \mathcal{F}(\mu_H^2)|\rho_H) + \mathcal{O}(\alpha_s^{k+1}) \\ &\quad + \mathcal{O}(\epsilon_s). \end{aligned} \quad (109)$$

This follows by using Eq. (104) to obtain $(1|\mathcal{U}(\mu_f^2, \epsilon_s Q^2[J]) = (1|$.

We should emphasize that in order to measure the cross section corresponding to the infrared safe operator \mathcal{O}_J with scale $Q^2[J]$, it is not necessary to cut off the shower at scale $\epsilon_s Q^2[J]$ as in Eq. (109). Rather, one simply runs the shower down to μ_f^2 and measures \mathcal{O}_J on the final state produced by the full shower, as in Eq. (100). When we do that, we are setting $\epsilon_s = \mu_f^2/Q^2[J]$, so the error estimate $\mathcal{O}(\epsilon_s)$ becomes $\mathcal{O}(\mu_f^2/Q^2[J])$.

H. The shower evolution equation

Using its definition Eq. (98), we see that the shower evolution operator $\mathcal{U}(\mu^2, \mu'^2)$ obeys an evolution equation of the form

$$\mu^2 \frac{d}{d\mu^2} \mathcal{U}(\mu^2, \mu'^2) = -\mathcal{S}(\mu^2) \mathcal{U}(\mu^2, \mu'^2). \quad (110)$$

Thus

$$\mathcal{U}(\mu^2, \mu'^2) = \mathbb{T} \exp \left(\int_{\mu'^2}^{\mu^2} \frac{d\mu^2}{\mu^2} \mathcal{S}(\mu^2) \right). \quad (111)$$

Since, according to Eq. (104), $(1|\mathcal{U}(\mu^2, \mu'^2) = (1|$, we have

$$(1|\mathcal{S}(\mu^2) = 0. \quad (112)$$

Using Eqs. (98) and (72), we see that the shower generator \mathcal{S} in Eq. (110) is

$$\begin{aligned} \mathcal{S}(\mu^2) &= \mathcal{V}(\mu^2) \mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}(\mu^2) \mathcal{F}^{-1}(\mu^2) \mathcal{V}^{-1}(\mu^2) \\ &\quad - \left(\mu^2 \frac{d}{d\mu^2} \mathcal{V}(\mu^2) \mathcal{F}(\mu^2) \right) \mathcal{F}^{-1}(\mu^2) \mathcal{V}^{-1}(\mu^2) \\ &= \mathcal{V}(\mu^2) \mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}(\mu^2) \mathcal{F}^{-1}(\mu^2) \mathcal{V}^{-1}(\mu^2) \\ &\quad - \mathcal{V}(\mu^2) \left(\mu^2 \frac{d}{d\mu^2} \mathcal{F}(\mu^2) \right) \mathcal{F}^{-1}(\mu^2) \mathcal{V}^{-1}(\mu^2) \\ &\quad - \left(\mu^2 \frac{d}{d\mu^2} \mathcal{V}(\mu^2) \right) \mathcal{V}^{-1}(\mu^2). \end{aligned} \quad (113)$$

It is convenient to define

$$\mathcal{S}_\mathcal{V}(\mu^2) = \mathcal{V}^{-1}(\mu^2) \mu^2 \frac{d}{d\mu^2} \mathcal{V}(\mu^2). \quad (114)$$

Also, we can use Eq. (65) for the evolution of $\mathcal{F}(\mu^2)$ and we can note that since $\mathcal{V}(\mu^2)$ does not change the number of partons or their momenta or flavors, $\mathcal{V}(\mu^2)$ commutes with $\mathcal{F}(\mu^2)$. Then

$$\begin{aligned} \mathcal{S}(\mu^2) &= \mathcal{V}(\mu^2) \mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}(\mu^2) \mathcal{F}^{-1}(\mu^2) \mathcal{V}^{-1}(\mu^2) \\ &\quad - [\mathcal{F}(\mu^2) \circ \mathcal{P}(\mu^2)] \mathcal{F}^{-1}(\mu^2) - \mathcal{V}(\mu^2) \mathcal{S}_{\mathcal{V}}(\mu^2) \mathcal{V}^{-1}(\mu^2). \end{aligned} \quad (115)$$

The operator $\mathcal{V}(\mu^2)$ here has a perturbative expansion beginning with

$$\mathcal{V}(\mu^2) = 1 + \frac{\alpha_s(\mu^2)}{2\pi} \mathcal{V}^{(1)}(\mu^2) + \dots \quad (116)$$

Then also

$$\mathcal{S}_{\mathcal{V}}(\mu^2) = \frac{\alpha_s(\mu^2)}{2\pi} \mathcal{S}_{\mathcal{V}}^{(1)}(\mu^2) + \dots \quad (117)$$

A sensible procedure for determining $\mathcal{S}(\mu^2)$ is to expand it perturbatively to whatever order is known, e.g., order α_s^k ,

$$\mathcal{S}(\mu^2) = \frac{\alpha_s(\mu^2)}{2\pi} \mathcal{S}^{(1)}(\mu^2) + \dots \quad (118)$$

It is of interest to see how this works out at order α_s . Since $\mathcal{S}_{\text{pert}}(\mu^2)$ and $\mathcal{S}_{\mathcal{V}}$ are already order α_s , we can simply replace $\mathcal{V}(\mu^2)$ by 1 in the first and third terms of Eq. (115). Then we have

$$\begin{aligned} \mathcal{S}^{(1)}(\mu^2) &= \mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}^{(1)}(\mu^2) \mathcal{F}^{-1}(\mu^2) \\ &\quad - [\mathcal{F}(\mu^2) \circ \mathcal{P}^{(1)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) - \mathcal{S}_{\mathcal{V}}^{(1)}(\mu^2). \end{aligned} \quad (119)$$

Let us multiply by $(1|$ and use $(1|\mathcal{S}^{(1)}(\mu^2) = 0$. In Ref. [32], we found that $(1|\mathcal{S}_{\text{pert}}^{(1,0)}(\mu^2)$ has a simple form,

$$\begin{aligned} (1|\mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}^{(1,0)}(\mu^2) \mathcal{F}^{-1}(\mu^2) \\ = (1|[\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2), \end{aligned} \quad (120)$$

where $\bar{\mathcal{S}}_{\text{pert}}^{(1,0)}(\mu^2)$ leaves the number of partons, their momenta and their flavors unchanged but has a nontrivial color structure.¹¹ We also divide $\mathcal{S}_{\text{pert}}^{(0,1)}(\mu^2)$ into two pieces

$$\mathcal{S}_{\text{pert}}^{(0,1)}(\mu^2) = \mathcal{S}_{i\pi}^{(0,1)}(\mu^2) + \mathcal{S}_{\text{Re}}^{(0,1)}(\mu^2). \quad (121)$$

Here $\mathcal{S}_{i\pi}^{(0,1)}(\mu^2)$ is the contribution from the virtual graphs that is proportional to $i\pi$, while $\mathcal{S}_{\text{Re}}^{(0,1)}(\mu^2)$ is the remaining part. We note that [32]

$$(1|\mathcal{S}_{i\pi}^{(0,1)}(\mu^2) = 0. \quad (122)$$

¹¹The operator $[\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2)$ was called \mathcal{V} in Ref. [32], but here we are letting \mathcal{V} denote a different operator.

This gives us

$$\begin{aligned} (1|\mathcal{S}_{\mathcal{V}}^{(1)}(\mu^2) &= (1| \{ [\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) + \mathcal{S}_{\text{Re}}^{(0,1)}(\mu^2) \\ &\quad - [\mathcal{F}(\mu^2) \circ \mathcal{P}^{(1)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) \}. \end{aligned} \quad (123)$$

All of the operators here leave the parton state unchanged except for being operators on the color and spin space. We define $\mathcal{S}_{\mathcal{V}}^{(1)}(\mu^2)$ to have the color and spin structure of the right-hand side of the equation, so that

$$\begin{aligned} \mathcal{S}_{\mathcal{V}}^{(1)}(\mu^2) &= [\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) + \mathcal{S}_{\text{Re}}^{(0,1)}(\mu^2) \\ &\quad - [\mathcal{F}(\mu^2) \circ \mathcal{P}^{(1)}(\mu^2)] \mathcal{F}^{-1}(\mu^2). \end{aligned} \quad (124)$$

This result also gives us¹²

$$\begin{aligned} \mathcal{S}^{(1)}(\mu^2) &= \mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}^{(1,0)}(\mu^2) \mathcal{F}^{-1}(\mu^2) \\ &\quad - [\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) + \mathcal{S}_{i\pi}^{(0,1)}(\mu^2). \end{aligned} \quad (125)$$

Compare this to the more general Eq. (D27) in Appendix D.

To use $\mathcal{S}^{(1)}(\mu^2)$, one solves Eq. (110) in the form

$$\begin{aligned} \mathcal{U}(\mu_2^2, \mu_1^2) &= \mathcal{N}(\mu_2^2, \mu_1^2) + \int_{\mu_2^2}^{\mu_1^2} \frac{d\mu^2}{\mu^2} \mathcal{U}(\mu_2^2, \mu^2) \mathcal{F}(\mu^2) \\ &\quad \times \frac{\alpha_s(\mu^2)}{2\pi} \mathcal{S}_{\text{pert}}^{(1,0)}(\mu^2) \mathcal{F}^{-1}(\mu^2) \mathcal{N}(\mu^2, \mu_1^2), \end{aligned} \quad (126)$$

where

$$\begin{aligned} \mathcal{N}(\mu_2^2, \mu_1^2) &= \mathbb{T} \exp \left(\int_{\mu_2^2}^{\mu_1^2} \frac{d\mu^2}{\mu^2} \frac{\alpha_s(\mu^2)}{2\pi} \right. \\ &\quad \left. \times \{ -[\mathcal{F}(\mu^2) \circ \bar{\mathcal{S}}^{(1,0)}(\mu^2)] \mathcal{F}^{-1}(\mu^2) + \mathcal{S}_{i\pi}^{(0,1)}(\mu^2) \} \right). \end{aligned} \quad (127)$$

Here the Sudakov factor \mathcal{N} is the exponential of the part of \mathcal{S} that does not change the number of partons or their momenta or flavors. Normally its spin and color structure is simplified and the $i\pi$ contribution is not included. The splitting operator $\mathcal{F}(\mu^2) \mathcal{S}_{\text{pert}}^{(1,0)}(\mu^2) \mathcal{F}^{-1}(\mu^2)$ adds one parton. Its spin and color structure is also normally simplified. Then Eq. (126) is implemented by solving it iteratively, so that there are some number of splittings interleaved with Sudakov factors.

¹²In Ref. [32], we neglected the $i\pi$ term and we averaged over spin. Then the right-hand side of Eq. (125) was called $\mathcal{H}_I - \mathcal{V}$.

I. The structure of $\mathcal{V}(\mu^2)$

In Eq. (100), there is a factor $\mathcal{V}(\mu_{\text{H}}^2)$ that multiplies $\mathcal{F}(\mu_{\text{H}}^2)$ and $|\rho_{\text{H}}\rangle$. We can write this operator as

$$\mathcal{V}(\mu_{\text{H}}^2) = \mathcal{V}(\mu_{\text{f}}^2)\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2), \quad (128)$$

where

$$\mathcal{U}_{\mathcal{V}}(\mu_2^2, \mu_1^2) = \mathcal{V}^{-1}(\mu_2^2)\mathcal{V}(\mu_1^2). \quad (129)$$

The operator $\mathcal{U}_{\mathcal{V}}(\mu^2, \mu'^2)$ obeys the evolution equation

$$\mu^2 \frac{d}{d\mu^2} \mathcal{U}_{\mathcal{V}}(\mu^2, \mu'^2) = -\mathcal{S}_{\mathcal{V}}(\mu^2)\mathcal{U}_{\mathcal{V}}(\mu^2, \mu'^2), \quad (130)$$

where $\mathcal{S}_{\mathcal{V}}(\mu^2)$ was defined in Eq. (114). Thus

$$\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2) = \mathbb{T} \exp \left(\int_{\mu_{\text{f}}^2}^{\mu_{\text{H}}^2} \frac{d\mu^2}{\mu^2} \mathcal{S}_{\mathcal{V}}(\mu^2) \right). \quad (131)$$

To use Eq. (131), we apply $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ to a statistical state $|\{p, f, s, s', c, c'\}_m\rangle$ that contributes to $|\rho_{\text{H}}\rangle$. We expand $\mathcal{S}_{\mathcal{V}}(\mu^2)$ in powers of $\alpha_s(\mu^2)$ with factors involving the running parton distributions $\mathcal{F}(\mu^2)$. The available scales other than μ^2 come from $|\rho_{\text{H}}\rangle$, so the relevant matrix elements involve μ^2/μ_{H}^2 . Since $\mathcal{V}(\mu^2)$ is an infrared finite operator, the perturbative coefficients in $\mathcal{S}_{\mathcal{V}}(\mu^2)$ will then be proportional to μ^2/μ_{H}^2 , possibly times logarithms of μ^2/μ_{H}^2 . Thus the low μ^2 end of the integration is power suppressed. However, it is important that there is a lower bound μ_{f}^2 on the integration. That is because the running coupling $\alpha_s(\mu^2)$ and the running parton distributions $\mathcal{F}(\mu^2)$ are not well defined for very small μ^2 .

We are left with a factor

$$\mathcal{V}(\mu_{\text{f}}^2) = 1 + \mathcal{O}(\alpha_s(\mu_{\text{f}}^2)). \quad (132)$$

We can expand this perturbatively in powers of $\alpha_s(\mu_{\text{f}}^2)$, using parton distributions $\mathcal{F}(\mu_{\text{f}}^2)$. The coefficients of $\alpha_s^n(\mu_{\text{f}}^2)$ for $n \geq 1$ are then proportional to $\mu_{\text{f}}^2/\mu_{\text{H}}^2$, possibly times logarithms of $\mu_{\text{f}}^2/\mu_{\text{H}}^2$. Since $\mu_{\text{f}}^2 \ll \mu_{\text{H}}^2$, we can safely neglect all of the higher order terms and simply replace

$$\mathcal{V}(\mu_{\text{f}}^2) \rightarrow 1. \quad (133)$$

Thus we make the replacement $\mathcal{V}(\mu_{\text{H}}^2) \rightarrow \mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ in Eq. (100), giving us

$$\begin{aligned} \sigma[J] &= (1|\mathcal{O}_J \mathcal{U}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)\mathcal{F}(\mu_{\text{H}}^2)|\rho_{\text{H}}\rangle \\ &\quad + \mathcal{O}(\alpha_s^{k+1}) + \mathcal{O}(\mu_{\text{f}}^2/Q[J]^2)). \end{aligned} \quad (134)$$

The factor $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ does two things. First, it provides perturbative corrections to the hard scattering state $|\rho_{\text{H}}\rangle$,

which we need in order to calculate the cross section correct to N^kLO . For this purpose, it would suffice to expand the exponential in $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ to the desired perturbative order. The second function of $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ is to sum threshold logarithms. For this purpose, it is important that $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ is an exponential.

To understand the relation of the operator $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ to threshold logarithms, it is instructive to look at it at order α_s with Λ ordering for the shower. It is structurally the same as the operator introduced in Ref. [32], which concerned the summation of threshold logarithms.¹³ The analysis in Ref. [32] simply averaged over spins, so we leave out spin here. The operator $\mathcal{S}_{\mathcal{V}}(\mu^2)$ in Ref. [32] contains several terms. Rather than listing them all, we simply recall the most important terms:

$$\begin{aligned} \mathcal{S}_{\mathcal{V}}(\mu^2)|\{p, f, c, c'\}_m\rangle &= \frac{\alpha_s}{2\pi} \int_{1/(1+\mu^2/\mu_{\text{H}}^2)}^1 dz \left[\left(1 - \frac{f_{a/A}(\eta_a/z, \mu^2)}{f_{a/A}(\eta_a, \mu^2)} \right) \frac{2C_a}{1-z} \right. \\ &\quad \left. + \left(1 - \frac{f_{b/B}(\eta_b/z, \mu^2)}{f_{b/B}(\eta_b, \mu^2)} \right) \frac{2C_b}{1-z} \right] [1 \otimes 1]|\{p, f, c, c'\}_m\rangle \\ &\quad + \dots, \end{aligned} \quad (135)$$

where $C_a = C_{\text{F}}$ when a is a quark flavor and $C_g = C_{\text{A}}$. The operator $[1 \otimes 1]$ is the unit operator in the color space. Note that when $\mu^2 \ll \mu_{\text{H}}^2$, the range of the z integration shrinks to just a tiny region near $z = 1$. Thus, this contribution might seem unimportant. However, in Eq. (135) there is a factor $1/(1-z)$. This multiplies a factor involving the parton distribution functions. The result is that $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$ is substantially different from 1 when the parton distribution functions are falling quickly as the momentum fraction grows. This gives a ‘‘threshold logarithm’’ effect that we can sum in a leading approximation by using $\mathcal{U}_{\mathcal{V}}(\mu_{\text{f}}^2, \mu_{\text{H}}^2)$.

The form of $\mathcal{S}_{\mathcal{V}}(\mu^2)$ depends on the definition of the infrared sensitive operator $\mathcal{D}(\mu^2)$. In particular, shower evolution uses parton distributions $\mathcal{F}(\mu^2)$ that are related to the $\overline{\text{MS}}$ parton distributions according to Eq. (61). This relation is discussed in Appendix B. Briefly, $\overline{\text{MS}}$ parton distributions are defined by removing ultraviolet divergences using $\overline{\text{MS}}$ renormalization, while the ultraviolet region of shower splittings is removed with a cutoff at scale μ_s^2 as defined in $\mathcal{D}(\mu^2)$. See Sec. V C. The evolution kernel $\mathcal{P}(\mu^2)$ for the parton distributions reflects the definition of $\mathcal{F}(\mu^2)$. This evolution kernel appears in $\mathcal{S}_{\mathcal{V}}(\mu^2)$, Eqs. (124) and (D26). Thus the choice of shower oriented parton

¹³In Ref. [32], an approximate version of $\mathcal{U}_{\mathcal{V}}(\mu_2^2, \mu_1^2)$ was used to sum threshold logarithms, but it appeared between each pair of parton splittings at scales μ_1^2 and μ_2^2 and also between the initial hard interaction and the first splitting. This caused problems, which were alleviated by inserting an artificial cut in $\mathcal{S}_{\mathcal{V}}(\mu^2)$.

distribution functions affects how the summation of threshold logarithms appears in the overall result (134) for the cross section: part of the summation of threshold logarithms appears as $\mathcal{U}_V(\mu_F^2, \mu_H^2)$ and part appears as the redefinition of the parton distributions, $\mathcal{F}(\mu^2)$ in Eq. (134). This is described in Sec. (9.4) of Ref. [32]. If one uses k_T ordering instead of Λ ordering for the shower, then the shower oriented parton distributions are very close to the $\overline{\text{MS}}$ parton distributions and all of the leading threshold logarithms appear in $\mathcal{U}_V(\mu_F^2, \mu_H^2)$. On the other hand, in the factorization scheme in Refs. [96,97], a different definition of factorization results in the leading threshold corrections being absorbed into the parton distribution functions.

VII. SUMMARY AND OUTLOOK

We began with an expression (42) for the cross section $\sigma[J]$ for an infrared safe measurement J calculated at $N^k\text{LO}$. The pieces in this expression are infrared divergent in four dimensions, so that they are defined by working in $4 - 2\epsilon$ dimensions. Integrating over the phase space that is unresolved by the measurement leads to some cancellations of poles $1/\epsilon$. Other poles cancel after factorization of initial state infrared sensitivity into parton distribution functions. This leaves a result that is finite in four dimensions, even though it consists of pieces that are divergent in four dimensions. We introduced an infrared sensitive operator $\mathcal{D}(\mu^2)$ and an inclusive infrared finite operator $\mathcal{V}(\mu^2)$ to help organize the cancellations.

After some analysis, we have represented the cross section as Eq. (134). Here the separate factors are all finite in four dimensions. If we expand this expression to order α_s^k , we have the same cross section that we started with except for a power suppressed contribution that we have dropped.

In Eq. (134), we have a hard scattering state $|\rho_H\rangle$ and a factor $\mathcal{F}(\mu_H^2)$ that supplies parton distribution functions and a parton luminosity factor so that if we trace over colors and spins, we have a differential cross section in the space of parton number, flavors, and momenta. Then we have a factor $\mathcal{U}_V(\mu_F^2, \mu_H^2)$ that supplies a summation of threshold logarithms associated with the hard state and also part of the $N^k\text{LO}$ perturbative corrections to the hard scattering cross section. Next, we have a complete parton shower generated by $\mathcal{U}(\mu_F^2, \mu_H^2)$. The parton shower operator preserves inclusive probabilities: $(1|\mathcal{U}(\mu_2^2, \mu_1^2) = (1|$. We have ended the shower at a scale μ_F^2 . After that, the factor $(1|\mathcal{O}_J$ represents the measurement that we want to make. We suppose that this is an infrared safe measurement that is not sensitive to soft or collinear parton splittings at the scale μ_F^2 or below. That means that the error in the calculation, estimated by $\mathcal{O}(\mu_F^2/Q[J]^2)$, is small. With such an infrared safe measurement, the result of the measurement is not sensitive to hadronization. If we wanted to use a measurement operator that is sensitive to hadronization, then we would need to include a model of hadronization

before the measurement operator. It is then less certain what a good error estimate is.

There is some temptation to imagine Eq. (134) as being simpler than it is. In our Higgs boson example, if we expand $(1|\mathcal{U}_V(\mu_F^2, \mu_H^2)\mathcal{F}(\mu_H^2)|\rho_H)$ to order α_s^k , it is the $N^k\text{LO}$ inclusive cross section to make a Higgs boson. The operator $\mathcal{U}(\mu_F^2, \mu_H^2)$ generates a probability preserving parton shower. Thus it might seem that one takes the hard scattering cross section and then distributes the probability across different final states according to what the shower generates. However, $\mathcal{U}_V(\mu_F^2, \mu_H^2)\mathcal{F}(\mu_H^2)|\rho_H)$ is not a cross section. It is a statistical state, representing different numbers of final state partons, which come with their own quantum color and spin states. The shower operator acts separately on each component of this statistical state. Then if we measure $\sigma[J]$ for an observable that is more complicated than just the inclusive measurement of a Higgs boson, the separate contributions are not sensible by themselves, but they sum to give $\sigma[J]$ correct to order α_s^k with only a power suppressed correction.

We have spoken of getting $\sigma[J]$ correct to order α_s^k , but, of course, that is not the point of a parton shower. In applying Eq. (134), one would evaluate the splitting function $\mathcal{S}(\mu^2)$ in the exponent of $\mathcal{U}(\mu_F^2, \mu_H^2)$ to order α_s^k , then retain $\mathcal{U}(\mu_F^2, \mu_H^2)$ as an exponential. When the desired measurement operator \mathcal{O}_J contains widely different scales, the cross section will contain large logarithms. Then $\mathcal{U}(\mu_F^2, \mu_H^2)$ has the potential to sum these logarithms. After all, it exponentiates the soft and collinear singularities of QCD at order α_s^k . Unfortunately, one needs to study the structure of $\mathcal{U}(\mu_F^2, \mu_H^2)$ as it relates to the structure of \mathcal{O}_J in order to check how well the shower does in summing the logarithms.

One can wonder whether the formalism of this paper is of any use for just a LO shower. We suggest that it is. If one averages over spins and makes the leading color approximation, the shower operator $\mathcal{U}(\mu_F^2, \mu_H^2)$ generates a rather conventional probability preserving dipole shower. With Λ^2 as the ordering variable, it is the leading color version of the shower in DEDUCTOR. One can choose other ordering variables. The operator $\mathcal{U}_V(\mu_F^2, \mu_H^2)$ generates threshold corrections, as described in Ref. [32]. These corrections are numerically important in some cases and could be included in standard parton shower programs.

In fact, Eq. (134) has been useful in improving DEDUCTOR. While working on our paper [32] on threshold corrections, we did not have Eq. (134). The result was a structure that had certain undesirable features that needed to be controlled by means of an *ad hoc* cutoff. The current more general formulation in Eq. (134) removes this problem, although it does not much change the numerical results. We present phenomenological results from the new version of DEDUCTOR in a separate paper [33].

The formalism is based on an operator $\mathcal{D}(\mu^2)$ that encodes the infrared structure of QCD starting with a state

with any number m of final state partons. If we know $\mathcal{D}(\mu^2)$ up to order α_s^k , then we generate, in an automatic way, the subtraction terms for an $N^k\text{LO}$ perturbative calculation. See the example of the toy model in Appendix A. This appears to us to be simpler than constructing the subtraction terms directly [110–117]. From $\mathcal{D}(\mu^2)$ at order α_s^k , we also generate, in an automatic way, the shower splitting kernels at order α_s^k .

The perturbative contributions to $\mathcal{D}(\mu^2)$ are not simple in full QCD. Furthermore, their form depends on choices like the momentum mapping scheme and the choice of a hardness ordering variable. At order α_s^1 , we have made the required choices, made suitable approximations, and calculated the corresponding splitting functions $\mathcal{S}(\mu^2)$ in Refs. [32,33]. Similarly, from $\mathcal{D}(\mu^2)$ we generate the inclusive infrared finite operator $\mathcal{V}(\mu^2)$. In general, there are some choices that one can make in defining $\mathcal{V}(\mu^2)$. At order α_s^1 , we have made the required choices, made suitable approximations, and calculated $\mathcal{S}_\mathcal{V}(\mu^2)$ in Ref. [32,33].

In the first order DEDUCTOR shower, we need only $d\mathcal{D}^{(1)}(\mu^2)/d\log(\mu^2)$, which is finite in four dimensions. For subtractions and matching to the shower, one needs the full integrated operator $\mathcal{D}^{(1)}(\mu^2)$. Techniques for this are described in Ref. [95] and the earlier papers [121–123].

We leave it to future work to make suitable choices for a momentum mapping scheme, a hardness ordering variable, and definitions away from the strict soft and collinear limits so as to construct order α_s^2 contributions to $\mathcal{D}(\mu^2)$. With a choice for color structure, we could then also construct $\mathcal{V}(\mu^2)$. We thus hope that the formalism presented in this paper might prove useful in developing a parton shower with order α_s^2 splitting functions.

We also hope that the formalism presented in this paper might provide support for the view that a parton shower is similar to a more straightforward perturbative calculation at $N^k\text{LO}$. In this view, the parton shower is an approximate way to calculate cross sections, but the approximation is systematically improvable by working at higher perturbative order. In a practical program, there may be further approximations with respect to color and spin. These need a separate justification.

ACKNOWLEDGMENTS

This work was supported in part by the United States Department of Energy under Grant No. DE-SC0011640. This project was begun while the authors were at the Munich Institute for Astro- and Particle Physics program “Challenges, Innovations and Developments in Precision Calculations for the LHC.” It was continued while the authors were at the Kavli Institute for Theoretical Physics at the University of California, Santa Barbara, program “LHC Run II and the Precision Frontier,” which was supported by the U.S. National Science Foundation under Grant No. NSF PHY11-25915. D.S. thanks the Erwin

Schrödinger Institute of the University of Vienna for its hospitality at its program “Challenges and Concepts for Field Theory and Applications in the Era of the LHC Run-2.” We thank the MIAPP, the KITP, and the ESI for providing stimulating research environments. We thank Tim Cohen, Marat Freytsis, and Jonathan Gaunt for helpful discussions.

APPENDIX A: A TOY MODEL FOR PARTON SHOWER OPERATORS

The construction in this paper relies on an operator \mathcal{D} that contains the infrared singularities of statistical states $|\rho\rangle$ that represent the color and spin density matrix elements of QCD. The operators \mathcal{D} can be used to construct both the infrared subtractions needed for a perturbative calculation of the cross section at $N^k\text{LO}$ and also the splitting kernels needed to construct a parton shower at the corresponding order, $N^{k-1}\text{LO}$. This construction has been quite abstract, especially since we lack an example of \mathcal{D} at $N^2\text{LO}$, which would correspond to an NLO shower.

In this Appendix, we illustrate some of the ideas of the paper using a toy model that provides a concrete example of $|\rho\rangle$ and \mathcal{D} at $N^2\text{LO}$. In this example, we construct the splitting kernel $\mathcal{S}_{\text{pert}}$ at NLO. The toy model is very simple. There are no parton distributions. The coupling α_s does not run. The momenta are one dimensional. There is no spin. There is quantum color, but the color structure is vastly simplified compared to what one has in real QCD.

1. Statistical states in the toy model

We use momentum states $\{p_1, p_2, \dots, p_m\}$ for m partons, with each p_i being a real number in the range $0 < p_i < \infty$. We use “color” represented by basis states labeled by an index pair (c_R, c_V) with $c_R \in \{0, 1, \dots, m\}$, $c_V \in \{0, 1, \dots\}$. The statistical states have the form

$$|\{p\}_m, (c_R, c_V)\rangle = |\{p_1, p_2, \dots, p_m\}, (c_R, c_V)\rangle. \quad (\text{A1})$$

These are defined to be invariant under permutations of the p_i . The Born level cross section is $|\{\}_0, (0, 0)\rangle$.

We will make use of color operators \mathcal{C}_R and \mathcal{C}_V that act on the space of statistical states. These have the form

$$\begin{aligned} \mathcal{C}_R &= 1 + \mathcal{A}_R^\dagger, \\ \mathcal{C}_V &= 1 + \mathcal{A}_V^\dagger, \end{aligned} \quad (\text{A2})$$

where \mathcal{A}_R^\dagger and \mathcal{A}_V^\dagger are raising operators:

$$\begin{aligned} \mathcal{A}_R^\dagger |\{p\}_m, (c_R, c_V)\rangle &= |\{p\}_m, (c_R + 1, c_V)\rangle, \\ \mathcal{A}_V^\dagger |\{p\}_m, (c_R, c_V)\rangle &= |\{p\}_m, (c_R, c_V + 1)\rangle. \end{aligned} \quad (\text{A3})$$

The inclusive probabilities corresponding to our statistical states are defined by

$$(1|\{p\}_m, (c_R, c_V)) = \left(\frac{1}{N_c^2}\right)^{c_R+c_V}. \quad (\text{A4})$$

Here $N_c = 3$ represents the number of colors. Thus keeping only $c_R = c_V = 0$ is analogous to the leading color approximation.

We include an operator representing an observable. The observable depends only on the parton momenta, not on their color state:

$$\mathcal{O}_J|\{p\}_m, (c_R, c_V) = \left(1 + \frac{p_1^2 + p_2^2 + \dots + p_m^2}{Q^2}\right)|\{p\}_m, (c_R, c_V). \quad (\text{A5})$$

2. Perturbative hard scattering states

We take the perturbative hard scattering states in the toy model to have the form

$$|\rho\rangle = |\rho^{(0,0)}\rangle + \alpha_s(|\rho^{(1,0)}\rangle + |\rho^{(0,1)}\rangle) + \alpha_s^2(|\rho^{(2,0)}\rangle + |\rho^{(1,1)}\rangle + |\rho^{(0,2)}\rangle) + \mathcal{O}(\alpha_s^3). \quad (\text{A6})$$

Here $|\rho^{(n_r, n_v)}\rangle$ represents a perturbative contribution with n_r real partons emitted and n_v virtual loops. The contributions are defined to be

$$\begin{aligned} |\rho^{(0,0)}\rangle &= |\{\}_0, (0, 0)\rangle, \\ |\rho^{(1,0)}\rangle &= \int_0^{Q^2} \frac{dp_1^2}{p_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{p_1^2 + Q^2} \mathcal{C}_R |\{p\}_1, (0, 0)\rangle, \\ |\rho^{(0,1)}\rangle &= - \int_0^\infty \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{k_1^2 + Q^2} \mathcal{C}_V |\{\}_0, (0, 0)\rangle, \\ |\rho^{(2,0)}\rangle &= \int_0^{Q^2} \frac{dp_2^2}{p_2^2} \left(\frac{p_2^2}{Q^2}\right)^\epsilon \int_0^{Q^2} \frac{dp_1^2}{p_1^2 + 2p_2^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{p_1^2 + p_2^2 + Q^2} \mathcal{C}_R^2 |\{p\}_2, (0, 0)\rangle, \\ |\rho^{(1,1)}\rangle &= - \left\{ \int_0^{Q^2} \frac{dp_1^2}{p_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \int_0^\infty \frac{dk_1^2}{k_1^2 + 2p_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{k_1^2 + p_1^2 + Q^2} \right. \\ &\quad \left. + \int_0^\infty \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \int_0^{Q^2} \frac{dp_1^2}{p_1^2 + 2k_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{p_1^2 + k_1^2 + Q^2} \right\} \mathcal{C}_R \mathcal{C}_V |\{p\}_1, (0, 0)\rangle, \\ |\rho^{(0,2)}\rangle &= \int_0^\infty \frac{dk_2^2}{k_2^2} \left(\frac{k_2^2}{Q^2}\right)^\epsilon \int_0^\infty \frac{dk_1^2}{k_1^2 + 2k_2^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \frac{Q^2}{k_1^2 + k_2^2 + Q^2} \mathcal{C}_V^2 |\{\}_0, (0, 0)\rangle. \end{aligned} \quad (\text{A7})$$

The integrals here are regularized in the infrared by factors $(p^2/Q^2)^\epsilon$. The individual contributions $|\rho^{(n_r, n_v)}\rangle$ contain singularities when parton momenta become small and $1/\epsilon$ poles that arise from integration from virtual parton momenta. Furthermore, the contributions with real parton emissions have different color states than the corresponding contributions with virtual loops. However, when we calculate the cross section $(1|\mathcal{O}_J|\rho\rangle$, we can use the fact that

$$(1|\mathcal{C}_R = (1|\mathcal{C}_V, \quad (\text{A8})$$

so that the color contributions from real emissions and virtual loops match. Then, in fact, there are *real – virtual*

cancellations, with the result that $(1|\mathcal{O}_J|\rho\rangle$ is infrared finite.

3. The infrared sensitive operator

The infrared structure of this is fairly simple and can be represented using the infrared sensitive operator $\mathcal{D}(\mu_s^2)$ with

$$\begin{aligned} \mathcal{D}(\mu_s^2) &= 1 + \alpha_s(\mathcal{D}^{(1,0)}(\mu_s^2) + \mathcal{D}^{(0,1)}(\mu_s^2)) \\ &\quad + \alpha_s^2(\mathcal{D}^{(2,0)}(\mu_s^2) + \mathcal{D}^{(1,1)}(\mu_s^2) + \mathcal{D}^{(0,2)}(\mu_s^2)) \\ &\quad + \mathcal{O}(\alpha_s^3), \end{aligned} \quad (\text{A9})$$

with

$$\begin{aligned}
\mathcal{D}^{(1,0)}(\mu_s^2)|\{p\}_m, (c_R, c_V) &= \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \mathcal{C}_R|\{p\}_{m+1}, (c_R, c_V), \\
\mathcal{D}^{(0,1)}(\mu_s^2)|\{p\}_m, (c_R, c_V) &= - \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \mathcal{C}_V|\{p\}_m, (c_R, c_V), \\
\mathcal{D}^{(2,0)}(\mu_s^2)|\{p\}_m, (c_R, c_V) &= \int_0^{\mu_s^2} \frac{dp_{m+2}^2}{p_{m+2}^2} \left(\frac{p_{m+2}^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2 + 2p_{m+2}^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \mathcal{C}_R^2|\{p\}_{m+2}, (c_R, c_V), \\
\mathcal{D}^{(1,1)}(\mu_s^2)|\{p\}_m, (c_R, c_V) &= - \left\{ \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2 + 2p_{m+1}^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \right. \\
&\quad \left. + \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2 + 2k_1^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \right\} \mathcal{C}_R \mathcal{C}_V|\{p\}_{m+1}, (c_R, c_V), \\
\mathcal{D}^{(0,2)}(\mu_s^2)|\{p\}_m, (c_R, c_V) &= \int_0^{\mu_s^2} \frac{dk_2^2}{k_2^2} \left(\frac{k_2^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2 + 2k_2^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \mathcal{C}_V^2|\{p\}_m, (c_R, c_V). \tag{A10}
\end{aligned}$$

From $\mathcal{D}(\mu_s^2)$ we can construct $\mathcal{D}^{-1}(\mu_s^2)$,

$$\begin{aligned}
\mathcal{D}^{-1}(\mu_s^2) &= 1 - \alpha_s(\mathcal{D}^{(1,0)}(\mu_s^2) + \mathcal{D}^{(0,1)}(\mu_s^2)) - \alpha_s^2(\mathcal{D}^{(2,0)}(\mu_s^2) + \mathcal{D}^{(1,1)}(\mu_s^2) + \mathcal{D}^{(0,2)}(\mu_s^2) - \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{D}^{(1,0)}(\mu_s^2) \\
&\quad - \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{D}^{(0,1)}(\mu_s^2) - \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{D}^{(1,0)}(\mu_s^2) - \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{D}^{(0,1)}(\mu_s^2)) + \mathcal{O}(\alpha_s^3). \tag{A11}
\end{aligned}$$

4. Subtractions for the hard scattering states

We can now construct the subtracted statistical state including the measurement operator,

$$|\hat{\rho}\rangle = \mathcal{D}^{-1}(\mu_s^2)\mathcal{O}_J|\rho\rangle. \tag{A12}$$

This has the expansion

$$|\hat{\rho}\rangle = |\rho^{(0,0)}\rangle + \alpha_s(|\hat{\rho}^{(1,0)}\rangle + |\hat{\rho}^{(0,1)}\rangle) + \alpha_s^2(|\hat{\rho}^{(2,0)}\rangle + |\hat{\rho}^{(1,1)}\rangle + |\hat{\rho}^{(0,2)}\rangle) + \mathcal{O}(\alpha_s^3). \tag{A13}$$

At first order in α_s , there are two terms. The first is

$$|\hat{\rho}^{(1,0)}\rangle = \mathcal{O}_J|\rho^{(1,0)}\rangle - \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle. \tag{A14}$$

This is

$$|\hat{\rho}^{(1,0)}\rangle = \int_0^\infty \frac{dp_1^2}{p_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \left\{ \theta(p_1^2 < Q^2) \frac{Q^2}{p_1^2 + Q^2} \left(1 + \frac{p_1^2}{Q^2}\right) - \theta(p_1^2 < \mu_s^2) \right\} \mathcal{C}_R|\{p\}_1, (0, 0). \tag{A15}$$

The first term is singular when $p_1^2 \rightarrow 0$, but the subtraction from $\mathcal{D}^{(1,0)}(\mu_s^2)$ eliminates the singularity. Then $(|\hat{\rho}^{(1,0)}\rangle)$ is finite at $\epsilon = 0$. The coefficient of α_s that corresponds to virtual graphs is

$$|\hat{\rho}^{(0,1)}\rangle = \mathcal{O}_J|\rho^{(0,1)}\rangle - \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle. \tag{A16}$$

This is

$$|\hat{\rho}^{(0,1)}\rangle = - \int_0^\infty \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \left\{ \frac{Q^2}{k_1^2 + Q^2} - \theta(k_1^2 < \mu_s^2) \right\} \mathcal{C}_V|\{p\}_0, (0, 0). \tag{A17}$$

The first term has a $1/\epsilon$ pole from $k_1^2 \rightarrow 0$, but the subtraction from $\mathcal{D}^{(0,1)}(\mu_s^2)$ eliminates the $k_1^2 \rightarrow 0$ singularity. Then $(|\hat{\rho}^{(0,1)}\rangle)$ is finite at $\epsilon = 0$.

At order α_s^2 , there are three terms in $|\hat{\rho}\rangle$. Let us look at the contribution from two real emissions:

$$|\hat{\rho}^{(2,0)}\rangle = \mathcal{O}_J|\rho^{(2,0)}\rangle - \mathcal{D}^{(2,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle - \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(1,0)}\rangle + \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle. \quad (\text{A18})$$

This is

$$\begin{aligned} |\hat{\rho}^{(2,0)}\rangle = & \int_0^\infty \frac{dp_2^2}{p_2^2} \left(\frac{p_2^2}{Q^2}\right)^\epsilon \int_0^\infty \frac{dp_1^2}{p_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \left\{ \frac{p_1^2}{p_1^2 + 2p_2^2} \left[\theta(p_2^2 < Q^2)\theta(p_1^2 < Q^2) \frac{Q^2}{p_1^2 + p_2^2 + Q^2} \left(1 + \frac{p_1^2 + p_2^2}{Q^2}\right) \right. \right. \\ & \left. \left. - \theta(p_2^2 < \mu_s^2)\theta(p_1^2 < \mu_s^2) \right] - \theta(p_2^2 < \mu_s^2) \left[\theta(p_1^2 < Q^2) \frac{Q^2}{p_1^2 + Q^2} \left(1 + \frac{p_1^2}{Q^2}\right) - \theta(p_1^2 < \mu_s^2) \right] \right\} \\ & \times \mathcal{C}_R^2|\{p\}_2, (0, 0)\rangle. \end{aligned} \quad (\text{A19})$$

Each term here exhibits infrared singularities, but the singularities cancel. Then $(1|\hat{\rho}^{(2,0)}\rangle)$ can be evaluated at $\epsilon = 0$. Specifically, for $p_2^2 \rightarrow 0$ at fixed p_1^2 , the first term cancels the third term and the second term cancels the fourth term. For $p_1^2 \rightarrow 0$ with fixed p_2^2 , the first two terms are nonsingular, while the third term cancels the fourth term. When $p_2^2 \rightarrow 0$ and $p_1^2 \rightarrow 0$, the first term cancels the second term and the third term cancels the fourth term.

The contribution to $|\hat{\rho}\rangle$ from two virtual emissions is similar. We have

$$|\hat{\rho}^{(0,2)}\rangle = \mathcal{O}_J|\rho^{(0,2)}\rangle - \mathcal{D}^{(0,2)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle - \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,1)}\rangle + \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle. \quad (\text{A20})$$

This is

$$\begin{aligned} |\hat{\rho}^{(0,2)}\rangle = & \int_0^\infty \frac{dk_2^2}{k_2^2} \left(\frac{k_2^2}{Q^2}\right)^\epsilon \int_0^\infty \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \left\{ \frac{k_1^2}{k_1^2 + 2k_2^2} \left[\frac{Q^2}{k_1^2 + k_2^2 + Q^2} - \theta(k_2^2 < \mu_s^2)\theta(k_1^2 < \mu_s^2) \right] \right. \\ & \left. - \theta(k_2^2 < \mu_s^2) \left[\frac{Q^2}{k_1^2 + Q^2} - \theta(k_1^2 < \mu_s^2) \right] \right\} \mathcal{C}_V^2|\{0\}_0, (0, 0)\rangle. \end{aligned} \quad (\text{A21})$$

Each term here exhibits $1/\epsilon$ poles, but the poles cancel. The pattern of cancellations is the same as for $|\hat{\rho}^{(2,0)}\rangle$.

The contribution to $|\hat{\rho}\rangle$ from one real emission and one virtual emission is a little more complicated. We have

$$\begin{aligned} |\hat{\rho}^{(1,1)}\rangle = & \mathcal{O}_J|\rho^{(1,1)}\rangle - \mathcal{D}^{(1,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle - \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,1)}\rangle - \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(1,0)}\rangle \\ & + \mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle + \mathcal{D}^{(0,1)}(\mu_s^2)\mathcal{D}^{(1,0)}(\mu_s^2)\mathcal{O}_J|\rho^{(0,0)}\rangle. \end{aligned} \quad (\text{A22})$$

We obtain

$$\begin{aligned} |\hat{\rho}^{(1,1)}\rangle = & \int_0^\infty \frac{dp_1^2}{p_1^2} \left(\frac{p_1^2}{Q^2}\right)^\epsilon \int_0^\infty \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \left\{ -\theta(p_1^2 < Q^2) \frac{Q^2}{k_1^2 + p_1^2 + Q^2} \left[\frac{k_1^2}{k_1^2 + 2p_1^2} + \frac{p_1^2}{p_1^2 + 2k_1^2} \right] \left(1 + \frac{p_1^2}{Q^2}\right) \right. \\ & + \theta(p_1^2 < \mu_s^2)\theta(k_1^2 < \mu_s^2) \left[\frac{k_1^2}{k_1^2 + 2p_1^2} + \frac{p_1^2}{p_1^2 + 2k_1^2} \right] + \theta(p_1^2 < \mu_s^2) \left[\frac{Q^2}{k_1^2 + Q^2} - 2\theta(k_1^2 < \mu_s^2) \right] \\ & \left. + \theta(p_1^2 < Q^2)\theta(k_1^2 < \mu_s^2) \frac{Q^2}{p_1^2 + Q^2} \left(1 + \frac{p_1^2}{Q^2}\right) \right\} \mathcal{C}_R\mathcal{C}_V|\{p\}_1, (0, 0)\rangle. \end{aligned} \quad (\text{A23})$$

Each term here exhibits $1/\epsilon$ poles and singularities but the poles and singularities cancel. Then $(1|\hat{\rho}^{(1,1)}\rangle)$ can be evaluated at $\epsilon = 0$. Specifically, expanding the square brackets gives us seven terms, T_i , $i \in \{1, \dots, 7\}$. For the singularity in the integrand at $k_1^2 \rightarrow 0$ with fixed p_1^2 , T_2 cancels T_7 and T_6 cancels $T_4 + T_5$. For the singularity at $p_1^2 \rightarrow 0$ with fixed k_1^2 , T_1 cancels T_5 and T_6 cancels

$T_3 + T_7$. For the singularity when $p_1^2 \rightarrow 0$ and $k_1^2 \rightarrow 0$, T_1 cancels T_3 , T_2 cancels T_4 , and T_6 cancels $T_5 + T_7$.

We thus see that $\mathcal{D}^{-1}(\mu_s^2)$ provides subtraction terms that cancel all of the singularities of $|\rho\rangle$ at order α_s and α_s^2 . If we had wanted to construct all of the subtraction terms directly, it would have been somewhat difficult. However, constructing $\mathcal{D}(\mu_s^2)$ for our toy model was quite

simple. Then obtaining the subtraction terms was automatic.

We also note that the perturbative states $|\rho^{(n_r, n_v)}\rangle$ have nontrivial color structures in our toy model. Then the operators $\mathcal{D}^{(n_r, n_v)}(\mu_s^2)$ must reflect this color structure. If they do not, the subtractions will not work.

5. The perturbative shower evolution operator

We are now in a position to construct the order α_s and α_s^2 terms in the perturbative shower generating operator, Eq. (73),

$$S_{\text{pert}}(\mu_s^2) = \mathcal{D}^{-1}(\mu_s^2) \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}(\mu_s^2). \quad (\text{A24})$$

The operator S_{pert} will have the expansion

$$\begin{aligned} \mathcal{S}_{\text{pert}}(\mu_s^2) &= \alpha_s (\mathcal{S}_{\text{pert}}^{(1,0)}(\mu_s^2) + \mathcal{S}_{\text{pert}}^{(0,1)}(\mu_s^2)) \\ &+ \alpha_s^2 (\mathcal{S}_{\text{pert}}^{(2,0)}(\mu_s^2) + \mathcal{S}_{\text{pert}}^{(1,1)}(\mu_s^2) + \mathcal{S}_{\text{pert}}^{(0,2)}(\mu_s^2)) \\ &+ \mathcal{O}(\alpha_s^3), \end{aligned} \quad (\text{A25})$$

where

$$\begin{aligned} \mathcal{S}_{\text{pert}}^{(1,0)}(\mu_s^2) &= \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,0)}(\mu_s^2), \\ \mathcal{S}_{\text{pert}}^{(0,1)}(\mu_s^2) &= \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,1)}(\mu_s^2), \\ \mathcal{S}_{\text{pert}}^{(2,0)}(\mu_s^2) &= \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(2,0)}(\mu_s^2) - \mathcal{D}^{(1,0)}(\mu_s^2) \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,0)}(\mu_s^2), \\ \mathcal{S}_{\text{pert}}^{(1,1)}(\mu_s^2) &= \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,1)}(\mu_s^2) - \mathcal{D}^{(1,0)}(\mu_s^2) \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,1)}(\mu_s^2) \\ &- \mathcal{D}^{(0,1)}(\mu_s^2) \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,0)}(\mu_s^2), \\ \mathcal{S}_{\text{pert}}^{(0,2)}(\mu_s^2) &= \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,2)}(\mu_s^2) - \mathcal{D}^{(0,1)}(\mu_s^2) \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,1)}(\mu_s^2). \end{aligned} \quad (\text{A26})$$

To use this, we first need the derivatives of $\mathcal{D}^{(n_r, n_v)}(\mu_s^2)$.

The derivatives of $\mathcal{D}^{(1,0)}(\mu_s^2)$ and $\mathcal{D}^{(0,1)}(\mu_s^2)$ are simple

$$\begin{aligned} \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,0)}(\mu_s^2) | \{P\}_m, (c_R, c_V) & \\ &= \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \mathcal{C}_R | \{P_1, \dots, P_m, \mu_s\}, (c_R, c_V), \\ \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,1)}(\mu_s^2) | \{P\}_m, (c_R, c_V) & \\ &= - \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \mathcal{C}_V | \{P\}_m, (c_R, c_V). \end{aligned} \quad (\text{A27})$$

Here in the state $|\{P_1, \dots, P_m, P_{m+1}\}, (c_R, c_V)\rangle$, we have substituted μ_s for the real number P_{m+1} .

The derivative of $\mathcal{D}^{(2,0)}(\mu_s^2)$ is a little more complicated. There are two terms. If we rename the integration variable and permute the arguments of the statistical state to match in the two terms, we obtain

$$\begin{aligned} \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(2,0)}(\mu_s^2) | \{P\}_m, (c_R, c_V) & \\ &= \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \int_0^{\mu_s^2} \frac{dP_{m+1}^2}{P_{m+1}^2} \left(\frac{P_{m+1}^2}{Q^2} \right)^\epsilon \\ &\times \left[\frac{P_{m+1}^2}{P_{m+1}^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2P_{m+1}^2} \right] \\ &\times \mathcal{C}_R^2 | \{P_1, \dots, P_m, P_{m+1}, \mu_s\}, (c_R, c_V). \end{aligned} \quad (\text{A28})$$

The derivative of $\mathcal{D}^{(0,2)}(\mu_s^2)$ is similar to the derivative of $\mathcal{D}^{(2,0)}(\mu_s^2)$:

$$\begin{aligned} \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(0,2)}(\mu_s^2) | \{P\}_m, (c_R, c_V) & \\ &= \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2} \right)^\epsilon \left[\frac{k_1^2}{k_1^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2k_1^2} \right] \\ &\times \mathcal{C}_V^2 | \{P\}_m, (c_R, c_V). \end{aligned} \quad (\text{A29})$$

The singularity at $k_1^2 \rightarrow 0$ gives us a $1/\epsilon$ pole after integration.

The derivative of $\mathcal{D}^{(1,1)}(\mu_s^2)$ is more complicated. After combining terms, we obtain

$$\begin{aligned} \mu_s^2 \frac{d}{d\mu_s^2} \mathcal{D}^{(1,1)}(\mu_s^2) | \{P\}_m, (c_R, c_V) & \\ &= - \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2} \right)^\epsilon \left[\frac{k_1^2}{k_1^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2k_1^2} \right] \\ &\times \mathcal{C}_R \mathcal{C}_V | \{P_1, \dots, P_m, \mu_s\}, (c_R, c_V) - \left(\frac{\mu_s^2}{Q^2} \right)^\epsilon \\ &\times \int_0^{\mu_s^2} \frac{dP_{m+1}^2}{P_{m+1}^2} \left(\frac{P_{m+1}^2}{Q^2} \right)^\epsilon \left[\frac{P_{m+1}^2}{P_{m+1}^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2P_{m+1}^2} \right] \\ &\times \mathcal{C}_R \mathcal{C}_V | \{P\}_{m+1}, (c_R, c_V). \end{aligned} \quad (\text{A30})$$

In the first term, there is a pole from the integration region $k_1^2 \rightarrow 0$, while in the second term there is a singularity at $P_{m+1}^2 \rightarrow 0$.

With the derivatives of $\mathcal{D}^{(n_r, n_v)}(\mu_s^2)$ at hand, it is straightforward to use Eq. (A26) to construct the shower kernel $\mathcal{S}_{\text{pert}}(\mu_s^2)$.

The shower kernel at order α_s is simple:

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(1,0)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= \mathcal{C}_R|\{p_1, \dots, p_m, \mu_s\}, (c_R, c_V)), \\ \mathcal{S}_{\text{pert}}^{(0,1)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= -\mathcal{C}_V|\{p\}_m, (c_R, c_V)).\end{aligned}\quad (\text{A31})$$

For the shower kernel for two real emissions, we find

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(2,0)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= \left(\frac{\mu_s^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \\ &\times \left[\frac{p_{m+1}^2}{p_{m+1}^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2p_{m+1}^2} - 1 \right] \\ &\times \mathcal{C}_R^2|\{p_1, \dots, p_m, p_{m+1}, \mu_s\}, (c_R, c_V)).\end{aligned}\quad (\text{A32})$$

The subtraction removes the infrared singularity at $p_{m+1}^2 \rightarrow 0$, allowing us to set $\epsilon \rightarrow 0$. This gives

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(2,0)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= \int_0^{\mu_s^2} dp_{m+1}^2 \left[\frac{1}{p_{m+1}^2 + 2\mu_s^2} - \frac{2}{\mu_s^2 + 2p_{m+1}^2} \right] \\ &\times \mathcal{C}_R^2|\{p_1, \dots, p_m, p_{m+1}, \mu_s\}, (c_R, c_V)).\end{aligned}\quad (\text{A33})$$

For the shower kernel for two virtual loops, we find a result that is analogous to what we found above. The subtraction removes the infrared singularity at $k_1^2 \rightarrow 0$ so that we can set $\epsilon \rightarrow 0$, giving

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(0,2)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= \int_0^{\mu_s^2} dk_1^2 \left[\frac{1}{k_1^2 + 2\mu_s^2} - \frac{2}{\mu_s^2 + 2k_1^2} \right] \mathcal{C}_V^2|\{p\}_m, (c_R, c_V)).\end{aligned}\quad (\text{A34})$$

For the shower kernel with one real emission and one virtual loop, we obtain

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(1,1)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= -\left(\frac{\mu_s^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dk_1^2}{k_1^2} \left(\frac{k_1^2}{Q^2}\right)^\epsilon \left[\frac{k_1^2}{k_1^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2k_1^2} - 1 \right] \\ &\times \mathcal{C}_R \mathcal{C}_V|\{p_1, \dots, p_m, \mu_s\}, (c_R, c_V)) \\ &- \left(\frac{\mu_s^2}{Q^2}\right)^\epsilon \int_0^{\mu_s^2} \frac{dp_{m+1}^2}{p_{m+1}^2} \left(\frac{p_{m+1}^2}{Q^2}\right)^\epsilon \\ &\times \left[\frac{p_{m+1}^2}{p_{m+1}^2 + 2\mu_s^2} + \frac{\mu_s^2}{\mu_s^2 + 2p_{m+1}^2} - 1 \right] \\ &\times \mathcal{C}_R \mathcal{C}_V|\{p\}_{m+1}, (c_R, c_V)).\end{aligned}\quad (\text{A35})$$

Again, the subtraction removes the infrared singularities. Then we can set $\epsilon \rightarrow 0$, giving

$$\begin{aligned}\mathcal{S}_{\text{pert}}^{(1,1)}(\mu_s^2)|\{p\}_m, (c_R, c_V)) &= -\int_0^{\mu_s^2} dk_1^2 \left[\frac{1}{k_1^2 + 2\mu_s^2} - \frac{2}{\mu_s^2 + 2k_1^2} \right] \\ &\times \mathcal{C}_R \mathcal{C}_V|\{p_1, \dots, p_m, \mu_s\}, (c_R, c_V)) \\ &- \int_0^{\mu_s^2} dp_{m+1}^2 \left[\frac{1}{p_{m+1}^2 + 2\mu_s^2} - \frac{2}{\mu_s^2 + 2p_{m+1}^2} \right] \\ &\times \mathcal{C}_R \mathcal{C}_V|\{p\}_{m+1}, (c_R, c_V)).\end{aligned}\quad (\text{A36})$$

Thus Eq. (A26) gives us a completely straightforward way to calculate the operators $\mathcal{S}_{\text{pert}}^{(n_r, n_v)}(\mu_s^2)$. All of these operators at order α_s^1 and α_s^2 are infrared finite.

6. The operator $\mathcal{V}(\mu_s^2)$

We also construct an operator $\mathcal{X}(\mu_s^2)$ in Eq. (79). In our toy model, which has no parton distributions, we have

$$\mathcal{X}(\mu_s^2) = \mathcal{D}(\mu_s^2). \quad (\text{A37})$$

Then we define an operator $\mathcal{V}(\mu_s^2)$ using Eq. (92),

$$(1|\mathcal{X}(\mu_s^2) = (1|\mathcal{V}(\mu_s^2). \quad (\text{A38})$$

In our toy model, at least up to the order that we have defined it, the relation $(1|\mathcal{C}_R = (1|\mathcal{C}_V$ gives us

$$(1|\mathcal{D}(\mu_s^2) = (1|. \quad (\text{A39})$$

Thus

$$\mathcal{V}(\mu_s^2) = 1. \quad (\text{A40})$$

Then according to the definition Eq. (98) (with $\mathcal{F} = 1$) we have

$$\mathcal{U}(\mu_2^2, \mu_1^2) = \mathcal{U}_{\text{pert}}(\mu_2^2, \mu_1^2). \quad (\text{A41})$$

The full shower operator \mathcal{U} is the same as $\mathcal{U}_{\text{pert}}$.

APPENDIX B: TRANSFORMATION TO SHOWER ORIENTED PARTON DISTRIBUTION FUNCTIONS

In Sec. VIB, we introduced shower oriented parton distribution functions that are adapted to the choice of the cutoff μ_s^2 used in the shower. The shower oriented parton distributions $f_{a/A}(\eta_a, \mu^2)$ are related to the five-flavor $\overline{\text{MS}}$ parton distribution functions $f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2)$ by means of a kernel K . Following the notation of Sec. VIB, the transformation of the parton distribution for hadron A is

$$f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2) = \sum_{a'} \int_0^1 \frac{dz}{z} K_{aa'}^{(a)}(z, \mu^2, \{p, f\}_m) f_{a'/A}(\eta_a/z, \mu^2). \quad (\text{B1})$$

The kernel K depends on the flavor indices a' and a , a momentum fraction variable z and the scale μ^2 . We also allow it to depend on the momenta and flavors $\{p, f\}_m$ of the partonic statistical state. Then K has a perturbative expansion beginning with

$$K_{aa'}^{(a)}(z, \mu^2, \{p, f\}_m) = \delta_{aa'} \delta(1-z) + \frac{\alpha_s(\mu^2)}{2\pi} K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m) + \mathcal{O}(\alpha_s^2). \quad (\text{B2})$$

In order to keep the presentation in this paper as simple as possible, we set the masses of all five quarks d,u,s,c,b to zero. In a more complete picture, one uses a variable flavor number scheme. Then, even with $\overline{\text{MS}}$ evolution, we change the renormalization scheme when the scale μ decreases past the bottom quark mass and then the charm quark mass. Thus the β function for α_s evolution and the parton evolution kernel change. For the shower oriented parton distributions $f_{a/A}(\eta_a, \mu^2)$ and $f_{b/B}(\eta_b, \mu^2)$, there is a different dependence on the c- and b-quark masses in the evolution kernels compared to what one uses in the variable flavor number version of $\overline{\text{MS}}$ evolution.

There are various possible versions of $K_{aa'}^{(a)}(z, \mu^2, \{p, f\}_m)$. The simplest is

$$\begin{aligned} K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m) &= \delta_{aa'} \left[\frac{2zC_a}{1-z} \log\left(\frac{Q_H^2}{(1-z)2p_a \cdot Q_H}\right) \right]_+ \\ &+ \sum_{\hat{a}} P_{aa'}^{\text{reg}}(z) \log\left(\frac{Q_H^2}{(1-z)2p_a \cdot Q_H}\right) \\ &+ \delta_{aa'} \delta(1-z) \gamma_a \log\left(\frac{Q_H^2}{2p_a \cdot Q_H}\right) - P_{aa'}^{(\epsilon)}(z). \end{aligned} \quad (\text{B3})$$

The first order $\overline{\text{MS}}$ DGLAP kernel is

$$P_{aa'}^{\overline{\text{MS}}}(z) = \delta_{aa'} \left[\frac{2zC_a}{1-z} \right]_+ + P_{aa'}^{\text{reg}}(z) + \delta_{aa'} \delta(1-z) \gamma_a. \quad (\text{B4})$$

Here $C_a = C_F$ and $\gamma_a = 3C_F/2$ when a is a quark flavor and $C_g = C_A$, $\gamma_g = 11C_A/6 - 2T_R n_f/3$. The functions $P_{aa'}^{\text{reg}}(z)$ and $P_{aa'}^{(\epsilon)}(z)$ are

$$\begin{aligned} P_{q\bar{q}}^{\text{reg}}(z) &= C_F(1-z), & P_{q\bar{q}}^{(\epsilon)}(z) &= C_F(1-z), \\ P_{g\bar{g}}^{\text{reg}}(z) &= 2C_A \left(\frac{1-z}{z} + z(1-z) \right), & P_{g\bar{g}}^{(\epsilon)}(z) &= 0, \\ P_{qg}^{\text{reg}}(z) &= T_R(1-2z(1-z)), & P_{qg}^{(\epsilon)}(z) &= T_R 2z(1-z), \\ P_{gq}^{\text{reg}}(z) &= C_F \left(z + 2 \frac{1-z}{z} \right), & P_{gq}^{(\epsilon)}(z) &= C_F z. \end{aligned} \quad (\text{B5})$$

The kernel $K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m)$ is a distribution in z with a singularity at $z \rightarrow 1$. The singularity is represented by the + prescription in the first term of Eq. (B3) and by the term proportional to $\delta(1-z)$. The coefficient of $\delta(1-z)$ is associated with how the virtual loop function $D^{(0,1)}(\mu^2)$ is treated. We have here taken a simple choice based on what is in Ref. [32], but other choices are possible.

The logarithms of $Q_H^2/[(1-z)2p_a \cdot Q_H]$ in Eq. (B3) come about as follows. We attempt to calculate $(1|\mathcal{X}(\mu^2)$ using the definition (79) of $\mathcal{X}(\mu^2)$. We look at emissions of a parton in the initial state. Call the virtuality associated with this splitting $|k^2| = 2\hat{p}_a \cdot \hat{p}_{m+1}$. We integrate over $|k^2|$ and over the momentum fraction z . We use the hardness variable Λ^2 , Eq. (55), used in DEDUCTOR. This means that there is an upper bound for the integration over $|k^2|$,

$$|k^2| < \frac{2p_a \cdot Q_H}{Q_H^2} \mu_s^2. \quad (\text{B6})$$

There is an infrared divergence coming from the $|k^2| \rightarrow 0$ limit of the integration. This divergence is regularized by integrating in $4 - 2\epsilon$ dimensions. Now, dimensional regularization effectively acts as an infrared cutoff on the transverse momentum

$$|k_{\perp}^2| = (1-z)|k^2|. \quad (\text{B7})$$

Thus we integrate over $|k_{\perp}^2|$ with an upper bound

$$|k_{\perp}^2| < \frac{2p_a \cdot Q_H}{Q_H^2} (1-z) \mu_s^2. \quad (\text{B8})$$

The $1/\epsilon$ pole produced by the integration over $|k_{\perp}^2|$ is removed by the factor $\mathcal{Z}_F(\mu^2)$ in $\mathcal{X}(\mu^2)$. This leaves us with a $\log(Q_H^2/[(1-z)2p_a \cdot Q_H])$, which multiplies the DGLAP splitting kernel $P_{aa'}(z)$. This remaining contribution does not have a $1/\epsilon$ pole. It has a term $\log(\mu^2/\mu_s^2)$, where μ_s^2 is the scale that defines the upper cutoff in the momentum integration and μ^2 is the renormalization scale. We do not see this logarithm because we set these scales equal to each other. The Feynman rules for the splitting functions have some explicit ϵ dependence, giving a function of the form $f(\epsilon)/\epsilon = f(0)/\epsilon + f'(0) + \mathcal{O}(\epsilon)$. The term $f'(0)$ gives us the contributions $P_{aa'}^{(\epsilon)}(z)$.

This calculation leaves us with an order α_s contribution to $(1|\mathcal{X}(\mu^2) = (1|\mathcal{V}(\mu^2)$ that we need to eliminate because it does not vanish in the limit $\mu^2 \rightarrow 0$. If we did not eliminate this term, we would lose Eq. (133). The offending contribution can be removed by the factor $\mathcal{K}(\mu^2)$ in $\mathcal{X}(\mu^2)$ if we choose the definition (B3).

We also note that if we use k_T^2 instead of Λ^2 , Eq. (55), as the shower hardness variable, then there is no $\log(1-z)$. Then by redefining μ_s^2 by a factor $Q_H^2/[2p_a \cdot Q_H]$, we obtain

$$K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m) = -P_{aa'}^{(e)}(z). \quad (\text{B9})$$

That is, the shower oriented parton distribution functions are close to the $\overline{\text{MS}}$ parton distribution functions. However, they are not quite equal. That is because for the shower oriented parton distribution functions we are imposing an ultraviolet cutoff with a theta function, while with the $\overline{\text{MS}}$ prescription we subtract a pole $1/\epsilon$.

We now turn to a construction that puts a different cut on parton splitting, leading to a less simple kernel $K_{aa'}^{(a)}(z, \mu^2, \{p, f\}_m)$. We retain the ultraviolet cut (B6) when μ_s^2 is not too small. However, when μ_s^2 is small, the upper bound Eq. (B8) can be very small indeed. We can relax this cut to

$$|k_T^2| < \max \left[\frac{2p_a \cdot Q_H}{Q_H^2} (1-z)\mu_s^2, m_\perp^2 \right], \quad (\text{B10})$$

where m_\perp^2 is, say, 1 GeV^2 . Then the matching kernel that defines the shower adapted parton distributions is

$$\begin{aligned} K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m) &= \delta_{a\hat{a}} \left[\frac{2zC_a}{1-z} \log \left(\min \left[\frac{Q_H^2}{(1-z)2p_a \cdot Q_H}, \frac{\mu_s^2}{m_\perp^2} \right] \right) \right]_+ \\ &+ \sum_{a'} P_{aa'}^{\text{reg}}(z) \log \left(\min \left[\frac{Q_H^2}{(1-z)2p_a \cdot Q_H}, \frac{\mu_s^2}{m_\perp^2} \right] \right) \\ &+ \delta_{a\hat{a}} \delta(1-z) \gamma_a \log \left(\min \left[\frac{Q_H^2}{2p_a \cdot Q_H}, \frac{\mu_s^2}{m_\perp^2} \right] \right) \\ &- P_{a\hat{a}}^{(e)}(z). \end{aligned} \quad (\text{B11})$$

When $m_\perp^2 \ll \mu_s^2$, this reduces to our previous definition. When $\mu_s^2 < m_\perp^2 Q_H^2/2p_a \cdot Q_H$, this becomes, after using Eq. (B4),

$$K_{aa'}^{(a,1)}(z, \mu^2, \{p, f\}_m) = \log \left(\frac{\mu_s^2}{m_\perp^2} \right) P_{a\hat{a}}^{\overline{\text{MS}}}(z) - P_{a\hat{a}}^{(e)}(z). \quad (\text{B12})$$

If we write Eq. (B1) to first order as

$$\begin{aligned} f_{a/A}(\eta_a, \mu^2) &= f_{a/A}^{\overline{\text{MS}}}(\eta_a/z, \mu^2) \\ &- \frac{\alpha_s}{2\pi} \sum_{a'} \int_0^1 \frac{dz}{z} K_{a,a'}^{(a,1)}(z, \mu^2, \{p, f\}_m) \\ &\times f_{a'/A}^{\overline{\text{MS}}}(\eta_a/z, m_\perp^2) \end{aligned} \quad (\text{B13})$$

and use Eq. (B12), recognizing that the $\overline{\text{MS}}$ kernel generates scale changes in $f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2)$, we find

$$\begin{aligned} f_{a/A}(\eta_a, \mu^2) &= f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2) - [f_{a/A}^{\overline{\text{MS}}}(\eta_a, \mu^2) - f_{a/A}^{\overline{\text{MS}}}(\eta_a, m_\perp^2)] \\ &+ \frac{\alpha_s}{2\pi} \sum_{a'} \int_0^1 \frac{dz}{z} P_{a\hat{a}}^{(e)}(z) f_{a'/A}^{\overline{\text{MS}}}(\eta_a/z, m_\perp^2) \\ &= f_{a/A}^{\overline{\text{MS}}}(\eta_a, m_\perp^2) \\ &+ \frac{\alpha_s}{2\pi} \sum_{a'} \int_0^1 \frac{dz}{z} P_{a\hat{a}}^{(e)}(z) f_{a'/A}^{\overline{\text{MS}}}(\eta_a/z, m_\perp^2). \end{aligned} \quad (\text{B14})$$

That is, with this definition, for small values of the scale μ^2 , the shower oriented parton distribution functions approximately equal the $\overline{\text{MS}}$ parton distribution functions at scale m_\perp^2 . However, at larger scales, the shower oriented parton distribution functions evolve differently from the $\overline{\text{MS}}$ ones. DEDUCTOR uses a definition similar to this, except with non-zero c and b quark masses and a corresponding variable flavor number scheme.¹⁴

APPENDIX C: CHOOSING Q_H AND μ^2 DYNAMICALLY

In Sec. IV C, we introduced a vector Q_H that is used to set scales and to help define one measure of hardness, Λ^2 , that can be used in the shower. [See Eq. (55).] We stated that one can use the intended measurement operator to set Q_H globally. There is another possibility: we can use the statistical state $|\rho_H\rangle$ that represents the hard scattering, Eq. (58). This state has the expansion

$$\begin{aligned} |\rho_H\rangle &= \sum_m \frac{1}{m!} \int [d\{p\}_m] \sum_{\{f\}_m} \sum_{\{s,s',c,c'\}_m} |\{p, f, s, s', c, c'\}_m\rangle \\ &\times (\{p, f, s, s', c, c'\}_m | \rho_H\rangle). \end{aligned} \quad (\text{C1})$$

Our example process in this paper is Higgs boson production. With this example, the Higgs boson momentum p_H is part of $\{p, f, s, s', c, c'\}_m$. We can set $Q_H^2 = m_H^2$ and set the rapidity of Q_H to be the rapidity of p_H while letting the transverse part of Q_H be zero. Another example would be jet production, for which the Born process is the production of two jets. In this case, we could use whatever infrared safe

¹⁴The contribution from $P_{a\hat{a}}^{(e)}(z)$ is ignored in DEDUCTOR.

jet algorithm we like to find the two highest P_T jets in $\{p\}_m$. Letting the momenta of these jets be P_1 and P_2 , we could set $Q_H^2 = (\vec{P}_{1,T}^2 + \vec{P}_{2,T}^2)/2$, set the rapidity of Q_H to the rapidity of $P_1 + P_2$, and set the transverse part of Q_H to zero.

This procedure requires some additional definitions since we have defined $|\rho_H\rangle = \mathcal{D}^{-1}(\mu_H^2)|\rho(\mu_H^2)\rangle$, so that we need Q_H and $\mu_H^2 = Q_H^2$ to create $|\rho_H\rangle$, but we cannot use these variables before they have been defined.

Let us see what is needed in the case of jet production, for which the number of final state partons at the Born level is $m = 2$. We content ourselves with what happens with an NLO hard cross section. We have

$$|\rho_H\rangle = \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^2 \left\{ |\rho^{(2,0)}\rangle + \frac{\alpha_s(\mu^2)}{2\pi} [|\rho^{(2,1)}(\mu^2)\rangle - \mathcal{D}^{(0,1)}(\mu^2)|\rho^{(2,0)}\rangle] + \frac{\alpha_s(\mu^2)}{2\pi} [|\rho^{(3,0)}\rangle - \mathcal{D}^{(1,0)}(\mu^2)|\rho^{(2,0)}\rangle] + \mathcal{O}(\alpha_s^2) \right\}. \quad (\text{C2})$$

We need to implement setting μ^2 to μ_H^2 . In the case of the argument of α_s , this is a simple replacement. For the Born statistical state $|\rho^{(2,0)}\rangle$, there is nothing further to do. In the terms $|\rho^{(2,1)}(\mu^2)\rangle$ and $\mathcal{D}^{(0,1)}(\mu^2)|\rho^{(2,0)}(\mu^2)\rangle$, there are cancelling poles. Then from $\mathcal{D}^{(0,1)}(\mu^2)$ there is a left over $\log \mu^2$ from the cutoff μ_s^2 in the integration over the virtual loop graph. We simply have to set $\mu^2 = \mu_H^2$ here. In the term $|\rho^{(3,0)}\rangle$, there is nothing further to do. In the term $\mathcal{D}^{(1,0)}(\mu^2)|\rho^{(2,0)}\rangle$, the momenta $\{p\}_3$ determine the momenta $\{p\}_2$ in $|\rho^{(2,0)}\rangle$ and also the splitting variables k^2, z, ϕ in $\mathcal{D}^{(1,0)}(\mu^2)$. Increasing the ultraviolet cutoff μ_s^2 provides an increased range for k^2 . As a result, the range in $\{p\}_3$ that is covered increases. Setting $\mu^2 = \mu_H^2$, we check whether the point $\{p\}_m$ is inside the allowed range. If it is, we multiply by 1. If it is not, we multiply by zero.

It seems clear that this procedure works beyond NLO, although it becomes more complicated. We work term by term in the expansion of $|\rho_H\rangle$. Once we have set Q_H and μ_H^2 , we set $\mu^2 = \mu_H^2$ in the argument of α_s and in all explicit logs that come from virtual loop integrals. We check whether the point $\{p\}_m$ is generated by a subtraction term with an ultraviolet cutoff μ_H^2 . If not, we omit this term. Our description here has been algorithmic. To formulate this in terms of operators on the statistical space requires additional notation, which we omit.

APPENDIX D: RENORMALIZATION

In this paper we use $\overline{\text{MS}}$ renormalization. In particular, this defines α_s and the $\overline{\text{MS}}$ parton distributions that we start with. In the parton shower algorithm that we obtain, each

element of the calculation is infrared finite and ultraviolet finite in four dimensions. However, part of the cancellation of infrared divergences is tied to the removal of ultraviolet divergences by renormalization. For this reason, the details of the ultraviolet renormalization scheme are significant. In this Appendix, we gather the most important formulas that we use, mostly following Ref. [119].

1. Renormalization of the QCD coupling

In the $\overline{\text{MS}}$ scheme, the renormalization of the coupling is¹⁵

$$\alpha_s^{\text{bare}} S_\epsilon = Z_\alpha(\mu^2) \mu^{2\epsilon} \alpha_s(\mu^2), \quad (\text{D1})$$

where μ is the renormalization scale, α_s^{bare} has mass dimension 2ϵ , and

$$S_\epsilon = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}. \quad (\text{D2})$$

The renormalization constant of the strong coupling is given as a sum,

$$Z_\alpha(\mu^2) = 1 + \sum_{n=1}^{\infty} \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \sum_{k=1}^n \frac{Z_\alpha^{[n,k]}}{e^k}. \quad (\text{D3})$$

The scale independent coefficients of the singularities, $Z_\alpha^{[n,k]}$, can be given in terms of the expansion parameters of the $\beta(\alpha_s)$ function by a recursion relation:

$$Z_\alpha^{[n,k+1]} = -\frac{1}{n} \sum_{l=k}^{n-1} (l+1) \beta_{n-l} Z_\alpha^{[l,k]}, \quad Z_\alpha^{[n,1]} = -\frac{1}{n} \beta_n. \quad (\text{D4})$$

The running coupling obeys the evolution equation

$$\begin{aligned} \mu^2 \frac{d\alpha_s(\mu^2)}{d\mu^2} &= -\alpha_s(\mu^2) (\epsilon + \beta(\alpha_s)) \\ &= -\alpha_s(\mu^2) \left(\epsilon + \sum_{n=1}^{\infty} \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \beta_n \right), \end{aligned} \quad (\text{D5})$$

where the first two β_i coefficients are

$$\begin{aligned} \beta_1 &= \frac{11C_A - 4T_R n_f}{6}, \\ \beta_2 &= \frac{17C_A^2 - 10C_A T_R n_f - 6C_F T_R n_f}{6}. \end{aligned} \quad (\text{D6})$$

¹⁵We find this definition useful for our purposes. Reference [119] uses a different strong coupling, $\tilde{\alpha}_s$, with $\alpha_s = S_\epsilon \tilde{\alpha}_s$.

2. Renormalization of the parton distribution functions

The $\overline{\text{MS}}$ parton distribution functions enter the cross section formula (42) in the form

$$\begin{aligned} & [\mathcal{F}_{\overline{\text{MS}}}(\mu^2) \circ \mathcal{Z}_F(\mu^2)] |\{p, f, s, s', c, c'\}_m\rangle \\ &= \sum_{a', b'} \int_0^1 \frac{dz_a}{z_a} \int_0^1 \frac{dz_b}{z_b} \frac{f_{a'/A}^{\overline{\text{MS}}}(\eta_a/z_a, \mu^2) f_{b'/B}^{\overline{\text{MS}}}(\eta_b/z_b, \mu^2)}{n_c(a) n_s(a) n_c(b) n_s(b) 4p_a \cdot p_b} \\ & \times Z_F(a, a'; z_a, \alpha_s(\mu^2)) Z_F(b, b'; z_b, \alpha_s(\mu^2)) \\ & \times |\{p, f, s, s', c, c'\}_m\rangle. \end{aligned} \quad (\text{D7})$$

The renormalization factor \mathcal{Z}_F is a product, so that each of the two parton distributions is transformed separately. The renormalization kernel Z_F relates the renormalized parton distribution to the bare parton distribution:

$$f_{a/A}^{\text{bare}}(\eta) = \sum_{a'} \int_0^1 \frac{dz}{z} Z_F(a, a'; z, \alpha_s(\mu^2)) f_{a'/A}^{\overline{\text{MS}}}(\eta/z, \mu^2). \quad (\text{D8})$$

The kernel has a perturbative expansion

$$\begin{aligned} Z_F(a, a'; z, \alpha_s(\mu^2)) &= \delta_{a,a'} \delta(1-z) \\ &+ \sum_{n=1}^{\infty} \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \sum_{k=1}^n \frac{Z_{a,a'}^{[n,k]}(z)}{e^k}. \end{aligned} \quad (\text{D9})$$

It follows from the requirement that $f_{a/A}^{\text{bare}}(\eta)$ is independent of μ^2 and $f_{a/A}^{\overline{\text{MS}}}(\eta, \mu^2)$ has no poles that the renormalized parton distribution function obeys the DGLAP evolution equation,

$$\mu^2 \frac{d f_{a/A}^{\overline{\text{MS}}}(\eta, \mu^2)}{d\mu^2} = \int_0^1 \frac{dz}{z} P_{a,a'}(z, \mu^2) f_{a'/A}^{\overline{\text{MS}}}(\eta/z, \mu^2), \quad (\text{D10})$$

where

$$P_{a,a'}(z, \mu^2) = \sum_{n=1}^{\infty} \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n P_{a,a'}^{(n)}(z) \quad (\text{D11})$$

with

$$P_{a,a'}^{(n)}(z) = n Z_{a,a'}^{[n,1]}(z). \quad (\text{D12})$$

The coefficients of $1/\epsilon$ to higher powers are then determined by the recursion relation

$$\begin{aligned} Z_{a,a'}^{[n,k+1]}(z) &= \frac{1}{n} \sum_{l=k}^{n-1} \int_z^1 \frac{dx}{x} \sum_c Z_{a,c}^{[l,k]}(z/x) \\ & \times [P_{c,a'}^{(n-l)}(x) - \delta_{c,a'} \delta(1-x) l \beta_{n-l}]. \end{aligned} \quad (\text{D13})$$

We can also write the evolution kernel as

$$\begin{aligned} P_{a,a'}(z, \mu^2) &= - \int_z^1 \frac{dx}{x} \sum_c Z_F^{-1}(a, c; z/x, \alpha_s(\mu^2)) \\ & \times \mu^2 \frac{d Z_F(c, a'; z, \alpha_s(\mu^2))}{d\mu^2}. \end{aligned} \quad (\text{D14})$$

The factors contain poles $1/\epsilon^k$, but the poles cancel.

3. Renormalization scale dependence

The physical states $|\rho(\mu^2)\rangle$ defined in Eq. (43) represent the quantum density operator of the partonic scattering. It is constructed from amplitudes $|M(\{p, f\}_m)\rangle$ and conjugate amplitudes $\langle M(\{p, f\}_m)|$. We include the proper Lehmann-Symanzik-Zimmermann (LSZ) factors for the incoming and outgoing partons so that the amplitudes are S-matrix elements in the renormalized theory. With massless partons and dimensional regularization in Feynman gauge, this amounts to multiplying by field strength renormalization factors $Z_\psi^{-1/2}$ or $Z_A^{-1/2}$ for each external leg of an amputated graph. Then $|\rho(\mu^2)\rangle$ is independent of the renormalization scale if calculated at all perturbative orders. When we calculate it up to order α_s^k as in Eq. (43), we have

$$\mu^2 \frac{d}{d\mu^2} |\rho(\mu^2)\rangle = \mathcal{O}(\alpha_s^{k+1}). \quad (\text{D15})$$

Let us consider next the infrared sensitive operator $\mathcal{D}(\mu^2)$. This operator depends on two scales, the renormalization scale μ^2 and shower scale μ_s^2 . We have set these scales to equal each other, but in this Appendix we highlight their separate roles. Thus we write $\mathcal{D}(\mu^2, \mu_s^2)$ with two arguments and let the operator with only one argument denote

$$\mathcal{D}(\mu^2) = \mathcal{D}(\mu^2, \mu^2). \quad (\text{D16})$$

For any basis state $|\{p, f, s, s', c, c'\}_m\rangle$, the state $\mathcal{D}(\mu^2, \mu_s^2) |\{p, f, s, s', c, c'\}_m\rangle$ is to be defined so as to have the properties of a physical statistical state $|\rho(\mu^2)\rangle$. It is constructed in the renormalized theory and has the proper field strength renormalization factors for its external partons so that $\mathcal{D}(\mu^2, \mu_s^2)$ is independent of the renormalization scale up to the order that we calculate. That is, if $\mathcal{D}(\mu^2, \mu_s^2)$ is calculated up to order α_s^k , we have

$$\mu^2 \frac{\partial}{\partial \mu^2} \mathcal{D}(\mu^2, \mu_s^2) = \mathcal{O}(\alpha_s^{k+1}). \quad (\text{D17})$$

In order to calculate the parton shower generators $\mathcal{S}_{\text{pert}}(\mu^2)$ and $\mathcal{S}(\mu^2)$, we need the total derivative of $\mathcal{D}(\mu^2)$ with respect to the common scale. Since $\mathcal{D}(\mu^2, \mu_s^2)$ is independent of μ^2 , this is actually the derivative with respect to μ_s^2 :

$$\begin{aligned} \frac{d}{d\mu^2} \mathcal{D}(\mu^2) &= \left. \frac{\partial}{\partial \mu^2} \mathcal{D}(\mu^2, \mu_s^2) \right|_{\mu_s^2=\mu^2} + \left. \frac{\partial}{\partial \mu_s^2} \mathcal{D}(\mu^2, \mu_s^2) \right|_{\mu_s^2=\mu^2} \\ &= \left. \frac{\partial}{\partial \mu_s^2} \mathcal{D}(\mu^2, \mu_s^2) \right|_{\mu_s^2=\mu^2} + \mathcal{O}(\alpha_s^{k+1}). \end{aligned} \quad (\text{D18})$$

The perturbative expansion of this is

$$\begin{aligned} \mu^2 \frac{d}{d\mu^2} \mathcal{D}(\mu^2) &= \sum_{n=1}^k \left[\frac{\alpha_s(\mu^2)}{2\pi} \right]^n \mu^2 \left. \frac{\partial}{\partial \mu_s^2} \mathcal{D}^{(n)}(\mu^2, \mu_s^2) \right|_{\mu_s^2=\mu^2} \\ &+ \mathcal{O}(\alpha_s^{k+1}). \end{aligned} \quad (\text{D19})$$

Using this, the generators of the perturbative shower operator can be obtained from Eqs. (73) and (49) as

$$\begin{aligned} \frac{1}{\mu^2} \mathcal{S}_{\text{pert}}^{(n)}(\mu^2) &= \left[\frac{\partial}{\partial \mu_s^2} \mathcal{D}^{(n)}(\mu^2, \mu_s^2) \right. \\ &\left. - \sum_{k=1}^{n-1} \tilde{\mathcal{D}}^{(k)}(\mu^2) \frac{\partial}{\partial \mu_s^2} \mathcal{D}^{(n-k)}(\mu^2, \mu_s^2) \right]_{\mu_s^2=\mu^2}. \end{aligned} \quad (\text{D20})$$

This simplifies the shower generators. The operators $\mathcal{D}^{(n)}(\mu^2, \mu_s^2)$ always contains a theta function like $\theta(\Lambda^2 < \mu_s^2)$ or the equivalent constraint for the loop contributions. The partial derivative turns one of these theta functions into a Dirac delta function.

The inclusive infrared finite operator $\mathcal{V}(\mu^2)$ is defined by the condition (92). This operator is derived from the infrared sensitive operator \mathcal{D} and, just like \mathcal{D} , depends on both the renormalization and shower scales. Thus we write $\mathcal{V}(\mu^2, \mu_s^2)$ with two arguments and let the operator with only one argument denote

$$\mathcal{V}(\mu^2) = \mathcal{V}(\mu^2, \mu^2). \quad (\text{D21})$$

Using Eqs. (79), (92), we have

$$(1|\mathcal{V}(\mu^2, \mu_s^2)\mathcal{F}(\mu^2) = (1|[\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2) \circ \mathcal{Z}_F(\mu^2)]\mathcal{D}(\mu^2, \mu_s^2)). \quad (\text{D22})$$

Recall that $[\mathcal{F}(\mu^2) \circ \mathcal{K}(\mu^2) \circ \mathcal{Z}_F(\mu^2)]$ is independent of μ^2 . We have just seen that $\mathcal{D}(\mu^2, \mu_s^2)$ is independent of μ^2 at fixed μ_s^2 . Recall from Sec. VID that we fix the color and spin content of $\mathcal{V}(\mu^2, \mu_s^2)$ to make Eq. (D22) work for

$(1|\mathcal{V}(\mu^2, \mu_s^2))$. Once we have done that for one choice of μ^2 at a given μ_s^2 , we can keep the same operator $\mathcal{V}(\mu^2, \mu_s^2)$ for other choices of μ^2 and Eq. (D22) will continue to hold. That is, we can define $\mathcal{V}(\mu^2, \mu_s^2)\mathcal{F}(\mu^2)$ so that it is independent of μ^2 up to whatever order we calculate:

$$\frac{\partial}{\partial \mu^2} [\mathcal{V}(\mu^2, \mu_s^2)\mathcal{F}(\mu^2)] = 0 + \mathcal{O}(\alpha_s^{k+1}). \quad (\text{D23})$$

This gives us

$$\begin{aligned} \mathcal{V}(\mu^2, \mu_s^2)^{-1} \frac{\partial}{\partial \mu^2} \mathcal{V}(\mu^2, \mu_s^2) \\ = - \left[\frac{d}{d\mu^2} \mathcal{F}(\mu^2) \right] \mathcal{F}^{-1}(\mu^2) + \mathcal{O}(\alpha_s^{k+1}). \end{aligned} \quad (\text{D24})$$

The generator $\mathcal{S}_{\mathcal{V}}(\mu^2)$ of $\mathcal{U}_{\mathcal{V}}(\mu_2^2, \mu_1^2)$ is defined in Eq. (114) as

$$\frac{1}{\mu^2} \mathcal{S}_{\mathcal{V}}(\mu^2) = \mathcal{V}^{-1}(\mu^2) \left[\frac{\partial}{\partial \mu^2} \mathcal{V}(\mu^2, \mu_s^2) + \frac{\partial}{\partial \mu_s^2} \mathcal{V}(\mu^2, \mu_s^2) \right]_{\mu_s^2=\mu^2}. \quad (\text{D25})$$

This gives us

$$\frac{1}{\mu^2} \mathcal{S}_{\mathcal{V}}(\mu^2) = \mathcal{V}^{-1}(\mu^2) \left. \frac{\partial}{\partial \mu_s^2} \mathcal{V}(\mu^2, \mu_s^2) \right|_{\mu_s^2=\mu^2} - \frac{d\mathcal{F}(\mu^2)}{d\mu^2} \mathcal{F}^{-1}(\mu^2). \quad (\text{D26})$$

Here the first term represents the evolution of the perturbative part of \mathcal{V} and the second term gives the evolution of the parton distribution functions. This generalizes Eq. (124) for the first order contribution to $\mathcal{S}_{\mathcal{V}}$.

Finally the generator of the probability preserving shower evolution operator is given in Eqs. (115) and (73). We can simplify this by using Eqs. (D18) and (D26)

$$\begin{aligned} \frac{1}{\mu^2} \mathcal{S}(\mu^2) &= \mathcal{V}(\mu^2)\mathcal{F}(\mu^2)\mathcal{D}^{-1}(\mu^2) \frac{\partial \mathcal{D}(\mu^2, \mu_s^2)}{\partial \mu_s^2} \\ &\times \mathcal{F}^{-1}(\mu^2)\mathcal{V}^{-1}(\mu^2) \Big|_{\mu_s^2=\mu^2} \\ &- \frac{\partial \mathcal{V}(\mu^2, \mu_s^2)}{\partial \mu_s^2} \mathcal{V}^{-1}(\mu^2) \Big|_{\mu_s^2=\mu^2}. \end{aligned} \quad (\text{D27})$$

- [1] M. Bahr *et al.*, Herwig++ physics and manual, *Eur. Phys. J. C* **58**, 639 (2008).
- [2] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P.Z. Skands, An introduction to PYTHIA 8.2, *Comput. Phys. Commun.* **191**, 159 (2015).
- [3] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, Event generation with SHERPA 1.1, *J. High Energy Phys.* **02** (2009) 007.
- [4] T. Sjöstrand, A model for initial state parton showers, *Phys. Lett. B* **157**, 321 (1985).
- [5] T.D. Gottschalk, Backwards evolved initial state parton showers, *Nucl. Phys.* **B277**, 700 (1986).
- [6] G. Marchesini and B. R. Webber, Simulation of QCD jets including soft gluon interference, *Nucl. Phys.* **B238**, 1 (1984).
- [7] G. Gustafson, Dual description of a confined color field, *Phys. Lett. B* **175**, 453 (1986).
- [8] R. K. Ellis, G. Marchesini, and B. R. Webber, Soft radiation in parton parton scattering, *Nucl. Phys.* **B286**, 643 (1987); Erratum **B294**, 1180 (1987).
- [9] G. Gustafson and U. Pettersson, Dipole formulation of QCD cascades, *Nucl. Phys.* **B306**, 746 (1988).
- [10] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour, and L. Stanco, HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons, *Comput. Phys. Commun.* **67**, 465 (1992).
- [11] T. Sjöstrand, High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4, *Comput. Phys. Commun.* **82**, 74 (1994).
- [12] T. Gleisberg, S. Hoeche, F. Krauss, A. Schaliche, S. Schumann, and J. C. Winter, SHERPA 1. Alpha: A proof of concept version, *J. High Energy Phys.* **02** (2004) 056.
- [13] T. Sjöstrand and P.Z. Skands, Transverse-momentum-ordered showers and interleaved multiple interactions, *Eur. Phys. J. C* **39**, 129 (2005).
- [14] Z. Nagy and D.E. Soper, A new parton shower algorithm: Shower evolution, matching at leading and next-to-leading order level, in *New trends in HERA physics, Proceedings of the Ringberg Workshop, Tegernsee, Germany, 2005*, edited by G. Grindhammer, B. A. Kniehl, G. Kramer, and W. Ochs (World Scientific, Singapore, 2006).
- [15] Z. Nagy and D.E. Soper, Parton showers with quantum interference, *J. High Energy Phys.* **09** (2007) 114.
- [16] C. W. Bauer, F. J. Tackmann, and J. Thaler, GenEvA. I. A new framework for event generation, *J. High Energy Phys.* **12** (2008) 010.
- [17] M. Dinsdale, M. Ternick, and S. Weinzierl, Parton showers from the dipole formalism, *Phys. Rev. D* **76**, 094003 (2007).
- [18] S. Schumann and F. Krauss, A parton shower algorithm based on Catani-Seymour dipole factorisation, *J. High Energy Phys.* **03** (2008) 038.
- [19] S. Platzer and S. Gieseke, Coherent parton showers with local recoils, *J. High Energy Phys.* **01** (2011) 024.
- [20] S. Platzer and S. Gieseke, Dipole showers and automated NLO matching in Herwig++, *Eur. Phys. J. C* **72**, 2187 (2012).
- [21] Z. Nagy and D.E. Soper, A parton shower based on factorization of the quantum density matrix, *J. High Energy Phys.* **06** (2014) 097.
- [22] S. Höche and S. Prestel, The midpoint between dipole and parton showers, *Eur. Phys. J. C* **75**, 461 (2015).
- [23] N. Fischer, S. Prestel, M. Ritzmann, and P. Skands, Vincia for hadron colliders, *Eur. Phys. J. C* **76**, 589 (2016).
- [24] A. Buckley *et al.*, General-purpose event generators for LHC physics, *Phys. Rep.* **504**, 145 (2011).
- [25] T. Sjöstrand, Status and developments of event generators, *Proc. Sci., LHCP2016* (2016) 007 [arXiv:1608.06425].
- [26] Z. Nagy and D.E. Soper, Parton showers with quantum interference: Leading color, spin averaged, *J. High Energy Phys.* **03** (2008) 030.
- [27] Z. Nagy and D.E. Soper, Parton showers with quantum interference: Leading color, with spin, *J. High Energy Phys.* **07** (2008) 025.
- [28] Z. Nagy and D.E. Soper, Parton shower evolution with subleading color, *J. High Energy Phys.* **06** (2012) 044.
- [29] Z. Nagy and D.E. Soper, Ordering variable for parton showers, *J. High Energy Phys.* **06** (2014) 178.
- [30] Z. Nagy and D.E. Soper, Parton distribution functions in the context of parton showers, *J. High Energy Phys.* **06** (2014) 179.
- [31] Z. Nagy and D.E. Soper, Effects of subleading color in a parton shower, *J. High Energy Phys.* **07** (2015) 119.
- [32] Z. Nagy and D.E. Soper, Summing threshold logs in a parton shower, *J. High Energy Phys.* **10** (2016) 019.
- [33] Z. Nagy and D.E. Soper, following paper, Jets and threshold summation in Deductor, *Phys. Rev. D* **98**, 014035 (2018).
- [34] S. Catani and M. H. Seymour, A general algorithm for calculating jet cross-sections in NLO QCD, *Nucl. Phys.* **B485**, 291 (1997); Erratum **B510**, 503 (1998).
- [35] G. F. Sterman, Summation of large corrections to short distance hadronic cross-sections, *Nucl. Phys.* **B281**, 310 (1987).
- [36] D. Appell, G. F. Sterman, and P. B. Mackenzie, Soft gluons and the normalization of the Drell-Yan cross-section, *Nucl. Phys.* **B309**, 259 (1988).
- [37] S. Catani and L. Trentadue, Resummation of the QCD perturbative series for hard processes, *Nucl. Phys.* **B327**, 323 (1989).
- [38] S. Catani, B. R. Webber, and G. Marchesini, QCD coherent branching and semiinclusive processes at large x , *Nucl. Phys.* **B349**, 635 (1991).
- [39] L. Magnea, All order summation and two loop results for the Drell-Yan cross-section, *Nucl. Phys.* **B349**, 703 (1991).
- [40] G. P. Korchemsky and G. Marchesini, Resummation of large infrared corrections using Wilson loops, *Phys. Lett. B* **313**, 433 (1993).
- [41] S. Catani, M. L. Mangano, P. Nason, and L. Trentadue, The resummation of soft gluons in hadronic collisions, *Nucl. Phys.* **B478**, 273 (1996).
- [42] H. Contopanagos, E. Laenen, and G. F. Sterman, Sudakov factorization and resummation, *Nucl. Phys.* **B484**, 303 (1997).
- [43] N. Kidonakis and G. F. Sterman, Resummation for QCD hard scattering, *Nucl. Phys.* **B505**, 321 (1997).

- [44] N. Kidonakis, G. Oderda, and G. F. Sterman, Threshold resummation for dijet cross-sections, *Nucl. Phys.* **B525**, 299 (1998).
- [45] N. Kidonakis, G. Oderda, and G. F. Sterman, Evolution of color exchange in QCD hard scattering, *Nucl. Phys.* **B531**, 365 (1998).
- [46] E. Laenen, G. Oderda, and G. F. Sterman, Resummation of threshold corrections for single particle inclusive cross-sections, *Phys. Lett. B* **438**, 173 (1998).
- [47] S. Catani, M. L. Mangano, and P. Nason, Sudakov resummation for prompt photon production in hadron collisions, *J. High Energy Phys.* **07** (1998) 024.
- [48] H. n. Li, Unification of the $k(T)$ and threshold resummations, *Phys. Lett. B* **454**, 328 (1999).
- [49] N. Kidonakis and J. F. Owens, Soft gluon resummation and NNLO corrections for direct photon production, *Phys. Rev. D* **61**, 094004 (2000).
- [50] G. F. Sterman and W. Vogelsang, Threshold resummation and rapidity dependence, *J. High Energy Phys.* **02** (2001) 016.
- [51] N. Kidonakis and J. F. Owens, Effects of higher order threshold corrections in high $E(T)$ jet production, *Phys. Rev. D* **63**, 054019 (2001).
- [52] E. Laenen, G. F. Sterman, and W. Vogelsang, Recoil and threshold corrections in short distance cross-sections, *Phys. Rev. D* **63**, 114018 (2001).
- [53] A. Kulesza, G. F. Sterman, and W. Vogelsang, Joint resummation in electroweak boson production, *Phys. Rev. D* **66**, 014011 (2002).
- [54] N. Kidonakis, Next-to-next-to-next-to-leading-order soft-gluon corrections in hard-scattering processes near threshold, *Phys. Rev. D* **73**, 034001 (2006).
- [55] A. Mukherjee and W. Vogelsang, Threshold resummation for W -boson production at RHIC, *Phys. Rev. D* **73**, 074005 (2006).
- [56] P. Bolzoni, Threshold resummation of Drell-Yan rapidity distributions, *Phys. Lett. B* **643**, 325 (2006).
- [57] D. de Florian and W. Vogelsang, Resummed cross-section for jet production at hadron colliders, *Phys. Rev. D* **76**, 074031 (2007).
- [58] V. Ravindran, J. Smith, and W. L. van Neerven, QCD threshold corrections to di-lepton and Higgs rapidity distributions beyond N^2 LO, *Nucl. Phys.* **B767**, 100 (2007).
- [59] V. Ravindran and J. Smith, Threshold corrections to rapidity distributions of Z and W^\pm bosons beyond N^2 LO at hadron colliders, *Phys. Rev. D* **76**, 114004 (2007).
- [60] M. Bonvini, S. Forte, and G. Ridolfi, Soft gluon resummation of Drell-Yan rapidity distributions: Theory and phenomenology, *Nucl. Phys.* **B847**, 93 (2011).
- [61] M. Bonvini, S. Forte, and G. Ridolfi, The Threshold region for Higgs Production in Gluon Fusion, *Phys. Rev. Lett.* **109**, 102002 (2012).
- [62] D. de Florian, P. Hinderer, A. Mukherjee, F. Ringer, and W. Vogelsang, Approximate Next-to-Next-to-Leading Order Corrections to Hadronic Jet Production, *Phys. Rev. Lett.* **112**, 082001 (2014).
- [63] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, Threshold resummation at N^3 LL accuracy and soft-virtual cross sections at N^3 LO, *Nucl. Phys.* **B888**, 75 (2014).
- [64] T. Ahmed, M. K. Mandal, N. Rana, and V. Ravindran, Rapidity Distributions in Drell-Yan and Higgs Productions at Threshold to Third Order in QCD, *Phys. Rev. Lett.* **113**, 212003 (2014).
- [65] M. Bonvini and S. Marzani, Resummed Higgs cross section at N^3 LL, *J. High Energy Phys.* **09** (2014) 007.
- [66] M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, M. Ubiali, R. D. Ball, V. Bertone, S. Carrazza, and N. P. Hartland, Parton distributions with threshold resummation, *J. High Energy Phys.* **09** (2015) 191.
- [67] A. V. Manohar, Deep inelastic scattering as $x \rightarrow 1$ using soft collinear effective theory, *Phys. Rev. D* **68**, 114019 (2003).
- [68] A. Idilbi and X. d. Ji, Threshold resummation for Drell-Yan process in soft-collinear effective theory, *Phys. Rev. D* **72**, 054016 (2005).
- [69] T. Becher and M. Neubert, Threshold Resummation in Momentum Space from Effective Field Theory, *Phys. Rev. Lett.* **97**, 082001 (2006).
- [70] T. Becher, M. Neubert, and B. D. Pecjak, Factorization and momentum-space resummation in deep-inelastic scattering, *J. High Energy Phys.* **01** (2007) 076.
- [71] T. Becher, M. Neubert, and G. Xu, Dynamical threshold enhancement and resummation in Drell-Yan production, *J. High Energy Phys.* **07** (2008) 030.
- [72] T. Becher and M. D. Schwartz, Direct photon production with effective field theory, *J. High Energy Phys.* **02** (2010) 040.
- [73] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Factorization at the LHC: From PDFs to initial state jets, *Phys. Rev. D* **81**, 094035 (2010).
- [74] M. Beneke, P. Falgari, and C. Schwinn, Threshold resummation for pair production of coloured heavy (s) particles at hadron colliders, *Nucl. Phys.* **B842**, 414 (2011).
- [75] C. W. Bauer, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Factorization and resummation for dijet invariant mass spectra, *Phys. Rev. D* **85**, 074006 (2012).
- [76] Y. Wang, C. S. Li, Z. L. Liu, and D. Y. Shao, Threshold resummation for $W^\pm Z$ and ZZ pair production at the LHC, *Phys. Rev. D* **90**, 034008 (2014).
- [77] G. Lusterians, W. J. Waalewijn, and L. Zeune, Joint transverse momentum and threshold resummation beyond NLL, *Phys. Lett. B* **762**, 447 (2016).
- [78] M. Bonvini, S. Forte, M. Ghezzi, and G. Ridolfi, Threshold resummation in SCET vs. perturbative QCD: An analytic comparison, *Nucl. Phys.* **B861**, 337 (2012).
- [79] G. Sterman and M. Zeng, Quantifying comparisons of threshold resummations, *J. High Energy Phys.* **05** (2014) 132.
- [80] M. Bonvini, S. Forte, G. Ridolfi, and L. Rottoli, Resummation prescriptions and ambiguities in SCET vs. direct QCD: Higgs production as a case study, *J. High Energy Phys.* **01** (2015) 046.
- [81] S. Frixione and B. R. Webber, Matching NLO QCD computations and parton shower simulations, *J. High Energy Phys.* **06** (2002) 029.

- [82] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, *J. High Energy Phys.* **11** (2004) 040.
- [83] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with Parton Shower simulations: The POWHEG method, *J. High Energy Phys.* **11** (2007) 070.
- [84] K. Hamilton and P. Nason, Improving NLO-parton shower matched simulations with higher order matrix elements, *J. High Energy Phys.* **06** (2010) 039.
- [85] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, *J. High Energy Phys.* **06** (2010) 043.
- [86] R. Frederix and S. Frixione, Merging meets matching in MC@NLO, *J. High Energy Phys.* **12** (2012) 061.
- [87] K. Hamilton, P. Nason, and G. Zanderighi, MINLO: Multi-scale improved NLO, *J. High Energy Phys.* **10** (2012) 155.
- [88] K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching, *J. High Energy Phys.* **05** (2013) 082.
- [89] S. Höche, F. Krauss, M. Schonherr, and F. Siegert, QCD matrix elements + parton showers: The NLO case, *J. High Energy Phys.* **04** (2013) 027.
- [90] L. Lönnblad and S. Prestel, Merging multi-leg NLO matrix elements with parton showers, *J. High Energy Phys.* **03** (2013) 166.
- [91] S. Plätzer, Controlling inclusive cross sections in parton shower + matrix element merging, *J. High Energy Phys.* **08** (2013) 114.
- [92] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, NNLOPS simulation of Higgs boson production, *J. High Energy Phys.* **10** (2013) 222.
- [93] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *J. High Energy Phys.* **07** (2014) 079.
- [94] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Matching fully differential NNLO calculations and parton showers, *J. High Energy Phys.* **06** (2014) 089.
- [95] M. Czakon, H. B. Hartanto, M. Kraus, and M. Worek, Matching the Nagy-Soper parton shower at next-to-leading order, *J. High Energy Phys.* **06** (2015) 033.
- [96] S. Jadach, W. Płaczek, S. Sapeta, A. Siódmok, and M. Skrzypek, Matching NLO QCD with parton shower in Monte Carlo scheme – the KrkNLO method, *J. High Energy Phys.* **10** (2015) 052.
- [97] S. Jadach, W. Płaczek, S. Sapeta, A. Siódmok, and M. Skrzypek, Parton distribution functions in Monte Carlo factorisation scheme, *Eur. Phys. J. C* **76**, 649 (2016).
- [98] R. Frederix and K. Hamilton, Extending the MINLO method, *J. High Energy Phys.* **05** (2016) 042.
- [99] S. Höche, Y. Li, and S. Prestel, Higgs-boson production through gluon fusion at NNLO QCD with parton showers, *Phys. Rev. D* **90**, 054011 (2014).
- [100] S. Höche, Y. Li, and S. Prestel, Drell-Yan lepton pair production at NNLO QCD with parton showers, *Phys. Rev. D* **91**, 074015 (2015).
- [101] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, QCD matrix elements + parton showers, *J. High Energy Phys.* **11** (2001) 063.
- [102] L. Lönnblad, Correcting the color dipole cascade model with fixed order matrix elements, *J. High Energy Phys.* **05** (2002) 046.
- [103] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, Matching matrix elements and shower evolution for top-quark production in hadronic collisions, *J. High Energy Phys.* **01** (2007) 013.
- [104] S. Jadach, A. Kusina, M. Skrzypek, and M. Slawinska, Two real parton contributions to non-singlet kernels for exclusive QCD DGLAP evolution, *J. High Energy Phys.* **08** (2011) 012.
- [105] S. Jadach, A. Kusina, W. Płaczek, and M. Skrzypek, NLO corrections in the initial-state parton shower Monte Carlo, *Acta Phys. Polon. B* **44**, 2179 (2013).
- [106] H. T. Li and P. Skands, A framework for second-order parton showers, *Phys. Lett. B* **771**, 59 (2017).
- [107] S. Höche and S. Prestel, Triple collinear emissions in parton showers, *Phys. Rev. D* **96**, 074017 (2017).
- [108] S. Höche, F. Krauss, and S. Prestel, Implementing NLO DGLAP evolution in parton showers, *J. High Energy Phys.* **10** (2017) 093.
- [109] J. C. Collins, D. E. Soper, and G. F. Sterman, Soft gluons and factorization, *Nucl. Phys.* **B308**, 833 (1988).
- [110] A. Daleo, A. Gehrmann-De Ridder, T. Gehrmann, and G. Luisoni, Antenna subtraction at NNLO with hadronic initial states: initial-final configurations, *J. High Energy Phys.* **01** (2010) 118.
- [111] R. Boughezal, A. Gehrmann-De Ridder, and M. Ritzmann, Antenna subtraction at NNLO with hadronic initial states: Double real radiation for initial-initial configurations with two quark flavours, *J. High Energy Phys.* **02** (2011) 098.
- [112] T. Gehrmann and P. F. Monni, Antenna subtraction at NNLO with hadronic initial states: Real-virtual initial-initial configurations, *J. High Energy Phys.* **12** (2011) 049.
- [113] A. Gehrmann-De Ridder, T. Gehrmann, and M. Ritzmann, Antenna subtraction at NNLO with hadronic initial states: double real initial-initial configurations, *J. High Energy Phys.* **10** (2012) 047.
- [114] J. Currie, E. W. N. Glover, and S. Wells, Infrared structure at NNLO using antenna subtraction, *J. High Energy Phys.* **04** (2013) 066.
- [115] G. Somogyi, Z. Trocsanyi, and V. Del Duca, A subtraction scheme for computing QCD jet cross sections at NNLO: Regularization of doubly-real emissions, *J. High Energy Phys.* **01** (2007) 070.
- [116] G. Somogyi and Z. Trocsanyi, A subtraction scheme for computing QCD jet cross sections at NNLO: Regularization of real-virtual emission, *J. High Energy Phys.* **01** (2007) 052.

- [117] G. Somogyi, A subtraction scheme for computing QCD jet cross sections at NNLO: integrating the doubly unresolved subtraction terms, *J. High Energy Phys.* **04** (2013) 010.
- [118] S. B. Libby and G. F. Sterman, Mass divergences in two particle inelastic scattering, *Phys. Rev. D* **18**, 4737 (1978).
- [119] J. Collins, *Foundations of Perturbative QCD* (Cambridge University Press, Cambridge, England, 2011).
- [120] S. Catani, D. de Florian, and G. Rodrigo, Space-like (versus time-like) collinear limits in QCD: Is factorization violated?, *J. High Energy Phys.* **07** (2012) 026.
- [121] C. H. Chung, M. Krämer, and T. Robens, An alternative subtraction scheme for next-to-leading order QCD calculations, *J. High Energy Phys.* **06** (2011) 144.
- [122] C. H. Chung and T. Robens, Nagy-Soper subtraction scheme for multiparton final states, *Phys. Rev. D* **87**, 074032 (2013).
- [123] G. Bevilacqua, M. Czakon, M. Kubocz, and M. Worek, Complete Nagy-Soper subtraction for next-to-leading order calculations in QCD, *J. High Energy Phys.* **10** (2013) 204.