Resummation of Transverse Momentum Distributions in Distribution Space

Markus A. Ebert and Frank J. Tackmann

Theory Group, Deutsches Elektronen-Synchrotron (DESY),
D-22607 Hamburg, Germany

E-mail: markus.ebert@desy.de, frank.tackmann@desy.de

ABSTRACT: Differential spectra in observables that resolve additional soft or collinear QCD emissions exhibit Sudakov double logarithms in the form of logarithmic plus distributions. Important examples are the total transverse momentum q_T in color-singlet production, N-jettiness (with thrust or beam thrust as special cases), but also jet mass and more complicated jet substructure observables. The all-order logarithmic structure of such distributions is often fully encoded in differential equations, so-called (renormalization group) evolution equations. We introduce a well-defined technique of distributional scale setting, which allows one to treat logarithmic plus distributions like ordinary logarithms when solving these differential equations. In particular, this allows one (through canonical scale choices) to minimize logarithmic contributions in the boundary terms of the solution, and to obtain the full distributional logarithmic structure from the solution's evolution kernel directly in distribution space. We apply this technique to the q_T distribution, where the two-dimensional nature of convolutions leads to additional difficulties (compared to one-dimensional cases like thrust), and for which the resummation in distribution (or momentum) space has been a long-standing open question. For the first time, we show how to perform the RG evolution fully in momentum space, thereby directly resumming the logarithms $[\ln^n(q_T^2/Q^2)/q_T^2]_+$ appearing in the physical q_T distribution. The resummation accuracy is then solely determined by the perturbative expansion of the associated anomalous dimensions.

Contents

---

1   Introduction

An important class of differential observables at colliders are those that resolve additional soft or collinear QCD emissions on top of the underlying hard Born process. They typically exhibit Sudakov double logarithms of the form $\alpha_s^n \ln^m(Q/k)$ with $m \leq 2n$, where Q is the relevant hard scale of the problem and k the differential observable. The bulk of the cross section is usually contained in the regime $k \ll Q$, where the double logarithms can become large and eventually spoil the convergence of the fixed-order perturbative expansion. The resummation of the Sudakov logarithms to all orders in α_s becomes necessary to obtain a stable and reliable prediction in this regime.

    Resummation can be carried out either using Monte Carlo techniques such as parton showers, or analytically based on factorization theorems which can be derived diagrammatically or using effective field theories. All ingredients of the factorized cross section obey (renormalization group) evolution equations of the form

$$\mu\frac{\mathrm{d}F(k,\mu)}{\mathrm{d}\mu} = \gamma_F(\mu)F(k,\mu)\,, \tag{1.1}$$

where F represents an ingredient of the factorized cross section depending on the (momentum space) variable k and the (unphysical) scale μ, and γ_F is its anomalous dimension. eq. (1.1) encodes the logarithmic structure of F to all orders. Solving it allows to exponentiate all large logarithms $\ln(k/\mu)$,

$$F(k,\mu) = F(k,\mu_0) \exp\left[\int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_F(\mu')\right], \qquad (1.2)$$

provided that μ_0 is chosen of the order of k, $\mu_0 \sim k$. In this case, $F(k,\mu_0)$ is free of large logarithms and can be reliably calculated in fixed-order perturbation theory, whereas all logarithms $\ln(\mu/k)$ are explicitly exponentiated.

While this is straightforward for ordinary functions F such as hard functions describing virtual corrections, it becomes more involved when F contains (plus) distributions. These naturally arise for many observables in order to properly cancel infrared divergences between the different functions. Well-known examples are the total transverse momentum q_T in color-singlet production, N-jettiness (with thrust or beam thrust as special cases), but also jet mass and more complicated jet substructure observables. The simplest example to illustrate the complications in such cases is a pure δ-distribution, which typically arises as the LO boundary term. In this case, eq. (1.2) reads

$$F(k,\mu) = \delta(k) \exp\left[\int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_F(\mu')\right] \approx \delta(k) \exp\left[\ln\frac{\mu}{\mu_0} \gamma_F(\mu)\right], \qquad (1.3)$$

such that it is obviously not possible to choose $\mu_0 = k$ to fully resum all logarithms $\ln(k/\mu)$. Hence one needs a technique that correctly turns this expression into a sum of plus distributions $\left[\ln^n(k/\mu)/k\right]_+^{\mu}$ that are known to arise at higher orders.

In this paper, we derive a well-defined technique to solve RGEs in distribution space by introducing a distributional scale setting $\mu_0 = k|_+$. It allows to treat evolution equations such as eq. (1.1) like ordinary differential equations, or equivalently treats distributional logarithms like ordinary logarithms. Hence this technique allows to minimize logarithmic terms in the boundary condition $F(k,\mu_0)$ through distributional canonical scale setting, thereby fully resumming all logarithms directly in distribution space. We also show that this is fundamentally different from performing the resummation in conjugate space, which also turns distributions into ordinary functions, but induces subleading terms to all orders.

We then apply this technique to the resummation of the transverse momentum ($\vec{q}_T$) distribution in color-singlet production. The transverse-momentum spectrum is a key observable for many processes at hadron colliders. In Higgs production, it is one of the primary variables describing the production kinematics [1–8]. In Drell-Yan production it is an important benchmark observable, which has been measured to very high precision [9–13].

It has been a long-standing open question whether a direct resummation of the $\vec{q}_T$ distribution in momentum (or distribution) space is possible at all. Instead, the resummation is usually carried out in the conjugate Fourier ($\vec{b}_T$) space, which however fundamentally resums logarithms $\ln(b_T Q)$ rather than the logarithms $\ln(Q/q_T)$ appearing in the physical

q_T spectrum (with Q being the hard scale of the process). Previous attempts so far have either encountered spurious divergences in the resummed q_T spectrum, see e.g. refs. [14, 15], or the evolution is still partially performed in Fourier space, see e.g. refs. [16–20], while refs. [21–23] tried to obtain a closed form of the Fourier-resummed spectrum in momentum space. Recently, ref. [24] obtained the resummation in momentum space based on the coherent branching formalism [25, 26] rather than solving the evolution equations associated with a factorization theorem. They showed that the spurious divergences are avoided by expanding around the transverse momentum of the hardest emission instead of q_T. We briefly compare our findings to those of ref. [24] in sec. 7.

In this paper, we derive the solution to perform the RG evolution entirely in distribution (momentum) space, thus allowing for the explicit resummation of all logarithmic contributions $[\ln^n(q_T^2/Q^2)/q_T^2]_+$ appearing in the physical q_T distribution. We show that it intrinsically requires distributional scale setting due to the two-dimensional convolutions appearing for q_T, which is not the case for one-dimensional observables such as thrust. In particular, we find that this is the origin of the divergences observed in previous attempts of momentum-space resummation.

An advantage of performing the resummation via the solution of the q_T evolution equations is that the solution automatically applies to all orders in resummed perturbation theory, and the resummation accuracy is solely defined through the perturbative accuracy of the anomalous dimensions (as well as matching conditions). In particular, this allows one to completely avoid any discussions of how to consistently count explicit logarithms in the cross section (which has been part of the difficulties in previous attempts). We indeed find that apparent subleading terms in the spectrum become important to obtain a well-defined prediction. The resummation through RG evolution also provides a natural and convenient way to smoothly turn off the resummation and match to the fixed-order region at large q_T, as well as to estimate perturbative uncertainties through the variation of all the appearing resummation scales. Importantly, by performing the evolution in momentum space, these steps can be carried out directly at the level of the physical q_T spectrum. In this paper, our primary purpose will be to derive the all-order solution for the momentum-space evolution, while we leave its numerical implementation to future work.

The remainder of the paper is organized as follows. In sec. 2, we introduce the technique of distributional scale setting and discuss how it is used to solve distributional differential (RG) equations in distribution space. We also discuss why this is fundamentally different from performing the resummation in conjugate space. For simplicity, most of this general discussion is for one-dimensional observables, and then generalized to two-dimensional cases at the end. The remaining sections are then devoted to the $\vec{q}_T$ resummation. In sec. 3, we briefly review the relevant factorization theorem and evolution equations, and discuss the difficulties associated with the two-dimensional nature of $\vec{q}_T$. In sec. 4, we discuss the resummation of the rapidity anomalous dimension via the solution of its RGE. We also briefly discuss how its intrinsic nonperturbative contributions arise and can be handled in momentum space. In sec. 5, we derive the solution for the momentum-space evolution for the soft and beam functions. We then discuss how all pieces get assembled into the final resummed $\vec{q}_T$ spectrum in sec. 6 and provide some comparisons to existing approaches and

implementations in the literature in sec. 7. We conclude in sec. 8. In the appendices we collect many relevant and useful definitions and relations for plus distributions.

2 Scale setting in distribution space

A key feature of observables k that resolve the IR structure of soft and collinear emissions, including the transverse momentum q_T, is that the logarithmic structure of the differential spectrum in k is given in terms of plus and delta distributions, which encode the cancellation of the associated IR divergences between real and virtual contributions.

The correct scale setting to minimize logarithmic distributions is a crucial ingredient in the resummation of the differential distributions directly in distribution space. In particular, as we will see, for the q_T spectrum this becomes a necessity to obtain well-defined predictions.

In this section, we discuss on general grounds several aspects of solving distributional differential equations. In particular, we introduce a distributional scale setting to correctly minimize distributional logarithms in the RGE boundary conditions. We first discuss the simpler case of one-dimensional distributions, before generalizing it to the two-dimensional case relevant for q_T in sec. 2.7.

2.1 Toy example

To illustrate the problem we consider a toy function $F(k, \mu)$, which contains logarithms $\ln(k/\mu)$ and obeys the toy RGE

$$\mu \frac{\mathrm{d}F(k, \mu)}{\mathrm{d}\mu} = -\alpha_s \, F(k, \mu) \,. \tag{2.1}$$

(For the purposes of this toy example we neglect the running of α_s for simplicity.) Solving eq. (2.1) yields the formal solution

$$F(k, \mu) = F(k, \mu_0) \, U(\mu_0, \mu) \,, \qquad U(\mu_0, \mu) = \exp\left(\alpha_s \ln \frac{\mu_0}{\mu}\right) \,, \tag{2.2}$$

where $U(\mu_0, \mu)$ is the evolution kernel with $U(\mu_0, \mu_0) = 1$ and $F(k, \mu_0)$ is the boundary condition. The evolution kernel resums logarithms $\ln(\mu_0/\mu)$ and shifts the logarithms $\ln(k/\mu_0)$ of $F(k, \mu_0)$ into the logarithms $\ln(k/\mu)$ of $F(k, \mu)$. Hence, if $F(k, \mu_0)$ is known "exactly" at the starting scale μ_0, eq. (2.2) gives F at the different scale μ with all its logarithms $\ln(k/\mu)$ resummed. More precisely, eq. (2.2) determines $F(k, \mu)$ to the logarithmic accuracy to which both the boundary condition and the evolution kernel are known.

In some cases, the boundary condition is known (or assumed to be known) exactly. For example, the nonperturbative parton distribution functions (PDFs) are extracted from the experimental data at some reference scale μ_0 and are then evolved to an arbitrary scale μ using the DGLAP evolution equations.

In contrast, if the boundary condition $F(k, \mu_0)$ is calculated perturbatively, it must be calculated including all logarithms (to the desired logarithmic accuracy). Typically, this is achieved by choosing a particular starting scale μ_0 for which $F(k, \mu_0)$ is free of logarithms.

In the simplest case, k is a scalar quantity such that F is a regular function and the RGE eq. (2.1) is multiplicative, as is the case for example for a hard matching coefficient in SCET. The logarithmic structure of F is then

$$F(k, \mu_0) = 1 + \alpha_s \ln \frac{k}{\mu_0} + \frac{1}{2} \alpha_s^2 \ln^2 \frac{k}{\mu_0} + \cdots , \qquad (2.3)$$

where the ellipses denote higher-order terms $\sim \alpha_s^n \ln^n(k/\mu_0)$ as well as possible constant (nonlogarithmic) terms. In this case, all logarithms vanish at $\mu_0 = k$, and $F(k, \mu_0 = k)$ can be calculated in fixed-order perturbation theory,

$$F(k, \mu_0 = k) = 1 + \cdots , \qquad (2.4)$$

where the ellipses now only contain possible higher-order constant terms. Equation (2.2) then gives $F(k, \mu)$ at an arbitrary scale μ by

$$F(k, \mu) = (1 + \cdots) \, U(\mu_0 = k, \mu) , \qquad (2.5)$$

where the logarithms $\ln(k/\mu)$ in $F(k, \mu)$ are now directly and *fully* resummed by the evolution kernel $U(k, \mu)$. Hence, the crucial ingredient that allows the RGE to *predict* the logarithms $\ln(k/\mu)$, as opposed to merely shifting them from one scale to another, is the choice of μ_0 that eliminates all logarithms in the boundary condition $F(k, \mu_0)$.

In the more complicated cases we are interested in, $F(k, \mu)$ is a differential distribution involving plus and delta distributions in the observable k. Let us first consider a toy example where after transforming to an appropriate conjugate space, e.g. Fourier or Mellin space, the RG eq. (2.1) has an analogous form

$$\mu \frac{\mathrm{d}\tilde{F}(y, \mu)}{\mathrm{d}\mu} = -\alpha_s \, \tilde{F}(y, \mu) , \qquad \tilde{F}(y, \mu) = \tilde{F}(y, \mu_0) \, \tilde{U}(\mu_0, \mu) . \qquad (2.6)$$

Here, y is the conjugate variable to k and $\tilde{F}(y, \mu)$ is the corresponding transformed function in conjugate space. We stress that as long as μ_0 is kept symbolic, the formal RGE solution is equivalent in any space, i.e. it does not actually matter in which space the evolution kernel is determined, it might just be easier to find a concrete solution in a particular space. The differences arise entirely in how the boundary condition is chosen. The above discussion for the multiplicative example now applies to $\tilde{F}(y, \mu)$, which has the simple logarithmic structure

$$\tilde{F}(y, \mu_0) = 1 - \alpha_s \ln(y\mu_0) + \frac{1}{2} \alpha_s^2 \ln^2(y \, \mu_0) + \cdots . \qquad (2.7)$$

Choosing $\mu_0 = 1/y$, the conjugate boundary condition $\tilde{F}(y, \mu_0 = 1/y)$ is free of logarithms and can be calculated in fixed-order perturbation theory, while all logarithms $\ln(y\mu)$ in the conjugate function $\tilde{F}(y, \mu)$ are correctly predicted by $\tilde{U}(\mu_0 = 1/y, \mu)$. This procedure of solving the RGE with setting the starting scale (or more generally determining the boundary condition) intrinsically in conjugate space fundamentally resums the conjugate logarithms and therefore we will refer to it as *evolution or resummation in conjugate space.*

For the reasons discussed before, we are interested in directly resuming the (distributional) momentum-space logarithms of k, which we will refer to as *evolution/resummation in momentum (or distribution) space*. The first attempt would be to use the canonical scale choice from momentum space, $\mu_0 = k$, in the formal solution in conjugate space, eq. (2.6). Doing so, the required boundary condition is

$$\tilde{F}(y, \mu_0 = k) = 1 - \alpha_s \ln(yk) + \frac{1}{2}\alpha_s^2 \ln^2(yk) + \cdots . \tag{2.8}$$

However, using fixed-order perturbation theory amounts to truncating this series at some low order in α_s and neglecting all logarithmic terms at higher orders in α_s. This would only be sufficient if it were guaranteed that the neglected logarithms in the boundary condition do not affect the desired logarithmic accuracy in the final momentum-space distribution, which is far from obvious. Although naively, one might think that the y-integration in the inverse transform should be dominated by $y \sim 1/k$, where the neglected logarithms are small, the integration also includes the regions $y \to 0$ and $y \to \infty$, where the neglected logarithms get arbitrarily large. As we will see, this is the reason why this naive attempt in fact fails in case of q_T (leading to the aforementioned spurious divergences).

Instead, performing the evolution in momentum space requires to set the scales *and* determine the appropriate fixed-order boundary condition in momentum space. Before discussing the analogous exponential toy example in distribution space, we can illustrate the arising difficulties with the simple distribution

$$D_0(k, \mu) = \delta(k) + \alpha_s \mathcal{L}_0(k, \mu) , \tag{2.9}$$

where

$$\mathcal{L}_0(k, \mu) \equiv \frac{1}{\mu} \mathcal{L}_0\left(\frac{k}{\mu}\right) \equiv \left[\frac{\theta(k)}{k}\right]_+^{\mu} , \tag{2.10}$$

and $\mathcal{L}_0(x) = [\theta(x)/x]_+$ is the standard plus distribution (see appendix B for more details). Since $\mathcal{L}_0(k, \mu)$ encodes a logarithmic divergence, it counts as a single logarithm in logarithmic accuracy counting. The distribution $D_0(k, \mu)$ fulfills the differential equation

$$\mu \frac{dD_0(k, \mu)}{d\mu} = -\alpha_s \delta(k) , \tag{2.11}$$

which has the general solution

$$D_0(k, \mu) = D_0(k, \mu_0) + \alpha_s \delta(k) \ln \frac{\mu_0}{\mu} . \tag{2.12}$$

Using the full result for $D_0(k, \mu_0)$ from eq. (2.9), together with the distributional identity [see eq. (B.5)]

$$\delta(k) \ln \frac{\mu_0}{\mu} = \mathcal{L}_0(k, \mu) - \mathcal{L}_0(k, \mu_0) \equiv \left[\frac{\theta(k)}{k}\right]_+^{\mu} - \left[\frac{\theta(k)}{k}\right]_+^{\mu_0} , \tag{2.13}$$

eq. (2.12) gives

$$D_0(k, \mu) = \delta(k) + \alpha_s \mathcal{L}_0(k, \mu_0) + \alpha_s \left[\mathcal{L}_0(k, \mu) - \mathcal{L}_0(k, \mu_0)\right]$$
$$= \delta(k) + \alpha_s \mathcal{L}_0(k, \mu) . \tag{2.14}$$

Thus, the μ_0-dependent terms cancel and we reproduce the correct result for $D_0(k,\mu)$. This shows that the formal solution in eq. (2.12) is sufficient to shift the μ-dependence from μ_0 to μ. However, obtaining the correct result for $D_0(k,\mu)$ from it like this crucially relies on the fact that we used the full result for the boundary condition $D_0(k,\mu_0)$.

In reality, calculating the boundary condition in fixed-order perturbation theory would amount to truncating $D_0(k,\mu_0) = \delta(k)$ and neglecting the higher-order $\alpha_s \mathcal{L}_0(k,\mu_0)$ term.[1] Naively setting $\mu_0 = k$ to eliminate this term is clearly ill defined. Since μ_0 appears in the boundary condition of the distribution, this would lead to an ill-defined expression $\delta(k)\ln k$ [as can been seen e.g. from eq. (2.13)]. Instead, we need a distributional way of choosing "$\mu_0 = k$" that behaves analogously to the multiplicative case above. Namely, it should set the logarithmic distribution $\mathcal{L}_0(k,\mu_0)$ in the boundary condition to zero so fixed-order perturbation theory can be used for the boundary condition without having to neglect any higher-order logarithmic terms, and such that the RGE directly predicts the correct $\mathcal{L}_0(k,\mu)$. One of the key results of this paper is a generic method to do so, which works for arbitrary distributions and will be discussed next.

2.2 Distributional scale setting

A general plus distributions of a function $g(k)$ is defined through the conditions (see appendix B for details)

$$\Big[\theta(k)\,g(k,\mu)\Big]_+^\mu = \theta(k)\,g(k,\mu) \qquad \text{for } k \neq 0\,, \tag{2.15}$$

$$\int^\mu \mathrm{d}k\,\Big[\theta(k)\,g(k,\mu)\Big]_+^\mu = 0\,. \tag{2.16}$$

Here and in the following we always keep the lower integration limit implicit in the support of the distribution. An important example are the logarithmic distributions

$$\mathcal{L}_n(x) = \left[\frac{\theta(x)\ln^n x}{x}\right]_+\,, \qquad \mathcal{L}_n(k,\mu) \equiv \frac{1}{\mu}\mathcal{L}_n\Big(\frac{k}{\mu}\Big) \equiv \left[\frac{\theta(k)}{k}\ln^n\frac{k}{\mu}\right]_+^\mu\,. \tag{2.17}$$

For a general distribution $D(k,\mu)$ we define the distributional scale setting $\mu = k|_+$ as

$$\boxed{D(k,\mu = k|_+) \equiv \frac{\mathrm{d}}{\mathrm{d}k}\left[\int^k \mathrm{d}k'\,D(k',\mu = k)\right]\,.} \tag{2.18}$$

Here, the derivative acts on everything inside the square brackets, and $\mu = k$ is set normally in the integrand, which is well defined since D is evaluated at the integration variable k'.

The idea behind eq. (2.18) is that it turns logarithmic distributions into ordinary logarithms, which can then be eliminated by a normal scale choice. For example,

$$\mathcal{L}_0(k,\mu = k|_+) = \frac{\mathrm{d}}{\mathrm{d}k}\int^k \mathrm{d}k'\left[\frac{\theta(k')}{k'}\right]_+^{\mu=k} = \frac{\mathrm{d}}{\mathrm{d}k}\left[\theta(k)\ln\frac{k}{\mu}\bigg|_{\mu=k}\right] = \frac{\mathrm{d}}{\mathrm{d}k}\,0 = 0\,. \tag{2.19}$$

---

[1] This is one of the reasons for utilizing $\mathrm{N}^n\mathrm{LL}'$ counting in the resummation of distributions [27–30]. There, the boundary conditions are included to one higher fixed order, which remedies this issue because the $\mathcal{L}_0$-terms are now explicitly included in the boundary condition.

Similarly for the higher logarithmic distributions,

$$\mathcal{L}_n(k, \mu = k|_+) = \frac{\mathrm{d}}{\mathrm{d}k} \int^k \mathrm{d}k' \left[\frac{\theta(k')}{k'} \ln^n \frac{k'}{\mu} \right]_+^\mu \Bigg|_{\mu=k}$$

$$= \frac{\mathrm{d}}{\mathrm{d}k} \left[\frac{\theta(k)}{n+1} \ln^{n+1} \frac{k}{\mu} \Bigg|_{\mu=k} \right] = \frac{\mathrm{d}}{\mathrm{d}k} 0 = 0 \,. \qquad (2.20)$$

Note that although $\mathcal{L}_n(k, \mu) = [\ln^n(k/\mu)/k]_+^\mu$ contains a pure logarithm $\ln(k/\mu)$, the appearance of μ in the boundary formally prohibits to simply set "$\mu = k$" to extract the boundary term.[2] It is also easy to see that this generalizes to any distribution $[\theta(k)g(k)]_+^\mu$,

$$\left[g(k, \mu) \right]_+^\mu \Bigg|_{\mu=k|_+} = \frac{\mathrm{d}}{\mathrm{d}k} \int^k \mathrm{d}k' \left[\theta(k')g(k', k) \right]_+^k = \frac{\mathrm{d}}{\mathrm{d}k} 0 = 0 \,. \qquad (2.21)$$

The cumulant integral exactly vanishes by definition of the distribution.

On the other hand, any μ-independent constant terms (i.e. pure boundary terms) are unaffected since the operator $\frac{\mathrm{d}}{\mathrm{d}k}$ precisely inverts the cumulant $\int^k \mathrm{d}k'$. In particular,

$$\delta(k) \Big|_{\mu=k|_+} = \frac{\mathrm{d}}{\mathrm{d}k} \int^k \mathrm{d}k' \, \delta(k') = \frac{\mathrm{d}}{\mathrm{d}k} \theta(k) = \delta(k) \,. \qquad (2.22)$$

Here it is crucial to keep track of any $\theta(k)$ appearing in the cumulant to properly recover the distributional structure. For future convenience we use the notation

$$\boxed{D[k] \equiv D(k, \mu)\Big|_{\mu=k|_+}} \qquad (2.23)$$

to denote the pure μ-independent constant term, which will typically be $D[k] \sim \delta(k)$, but in general could also contain regular (integrable) functions of k.

Having a well-defined method for distributional scale setting allows us to easily solve distributional differential equations, as it allows us to essentialy treat distributions like ordinary logarithms. To see that eq. (2.18) has the desired properties in more nontrivial cases, consider the simple examples

$$\delta(k) \ln^{n+1} \frac{\mu_0}{\mu} \Bigg|_{\mu_0=k|_+} = (n+1)\mathcal{L}_n(k, \mu) \qquad (n \geq 0) \,,$$

$$(m+1)\mathcal{L}_m(k, \mu) \ln^n \frac{\mu_0}{\mu} \Bigg|_{\mu_0=k|_+} = (n+m+1)\mathcal{L}_{m+n}(k, \mu) \qquad (n \geq 0) \,,$$

$$\mathcal{L}_m(k, \mu_0) \ln^n \frac{\mu_0}{\mu} \Bigg|_{\mu_0=k|_+} = 0 \qquad (n \geq 0) \,. \qquad (2.24)$$

These relations are quite intuitive in that the distributional scale setting essentially moves any naively appearing $\ln k$ terms inside a suitably regulated logarithmic distribution. The appearing prefactors count the order of the distributional logarithm in each equation.

---

[2] Although this naive choice is mathematically ill-defined, it will actually yield the same result as the proper scale setting. It explicitly fails for $\mathcal{L}_0$, which does not contain an explicit logarithm.

More generally, we have

$$\delta(k)\, g(\mu_0, \mu)\Big|_{\mu_0 = k|_+} = \delta(k)\, g(\mu, \mu) + \left[\theta(k)\frac{\mathrm{d}g(k,\mu)}{\mathrm{d}k}\right]_+^\mu,$$

$$\left[\theta(k)\, f(k,\mu)\right]_+^\mu g(\mu_0, \mu)\Big|_{\mu_0 = k|_+} = \left[\theta(k)\frac{\mathrm{d}}{\mathrm{d}k}\left(g(k,\mu)\int_\mu^k \mathrm{d}k'\, f(k',\mu)\right)\right]_+^\mu,$$

$$\left[\theta(k)\, f(k,\mu_0)\right]_+^{\mu_0} g(\mu_0, \mu)\Big|_{\mu_0 = k|_+} = 0\,. \tag{2.25}$$

2.3 Integrating distributional differential equations

First, consider again the simple example in eq. (2.9). Setting $\mu_0 = k|_+$ now has the desired effect of being able to predict the complete logarithmic distribution from the general solution eq. (2.12) by a scale choice,

$$D_0(k,\mu) = D_0(k,\mu_0) + \alpha_s\delta(k)\ln\frac{\mu_0}{\mu}\bigg|_{\mu_0 = k|_+} = \big[\delta(k) + 0\big] + \alpha_s \mathcal{L}_0(k,\mu)\,. \tag{2.26}$$

Similarly, we can reproduce a $\mathcal{L}_n(k,\mu)$ for $n \geq 1$ from its μ dependence. Consider a distribution $D_n(k,\mu)$, which satisfies the differential equation

$$\mu\frac{\mathrm{d}D_n(k,\mu)}{\mathrm{d}\mu} = -n\,\alpha_s\mathcal{L}_{n-1}(k,\mu) = -n\,\alpha_s\left[\frac{\theta(k)}{k}\ln^{n-1}\frac{k}{\mu}\right]_+^\mu\,. \tag{2.27}$$

Integrating this from μ_0 to μ, we get the general solution

$$D_n(k,\mu) = D_n(k,\mu_0) - n\,\alpha_s\int_{\mu_0}^\mu \frac{\mathrm{d}\mu'}{\mu'}\left[\frac{\theta(k)}{k}\ln^{n-1}\frac{k}{\mu'}\right]_+^{\mu'}$$

$$= D_n(k,\mu_0) + \alpha_s\left[\frac{\theta(k)}{k}\left(\ln^n\frac{k}{\mu} - \ln^n\frac{k}{\mu_0}\right)\right]_+^{\mu_0} + \frac{\delta(k)}{n+1}\,\alpha_s\ln^{n+1}\frac{\mu}{\mu_0}\,. \tag{2.28}$$

Here we used eq. (B.5) to shift the boundary condition in the plus distribution from μ' to μ_0, such that the integral can be pulled inside the plus distribution and performed. (This step explicitly requires that μ_0 does not depend on k, as otherwise the boundary condition of the plus distribution would be changed.) Using eq. (2.24) to set $\mu_0 = k|_+$, we thus find

$$D_n(k,\mu) = D_n[k] + \alpha_s\mathcal{L}_n(k,\mu)\,. \tag{2.29}$$

In practice, we can also specialize eq. (2.18) to define an integral over an arbitrary distribution G with starting scale $\mu_0 = k|_+$ as

$$\boxed{\int_{k|_+}^\mu \frac{\mathrm{d}\mu'}{\mu'}\, G(k,\mu') \equiv \frac{\mathrm{d}}{\mathrm{d}k}\int^k \mathrm{d}k'\int_k^\mu \frac{\mathrm{d}\mu'}{\mu'}\, G(k',\mu')\,.} \tag{2.30}$$

This allows us to write a generic integral solution as

$$\mu\frac{\mathrm{d}D(k,\mu)}{\mathrm{d}\mu} = G(k,\mu) \qquad \Rightarrow \qquad D(k,\mu) = D[k] + \int_{k|_+}^\mu \frac{\mathrm{d}\mu'}{\mu'}\, G(k,\mu')\,. \tag{2.31}$$

Although this looks like an ordinary integral solution, the important difference lies in the lower integration limit $k|_+$ which enforces the distributional scale setting.

2.4 Toy example in distribution space

We can now discuss the exponential toy example in distribution space. Consider a distribution with the logarithmic structure

$$F(k,\mu) = \delta(k) + \alpha_s \mathcal{L}_0(k,\mu) + \alpha_s^2 \mathcal{L}_1(k,\mu) + \frac{1}{2}\alpha_s^3 \mathcal{L}_2(k,\mu) + \cdots, \qquad (2.32)$$

corresponding to an exponential in distribution space. It satisfies the differential equation

$$\mu \frac{\mathrm{d}F(k,\mu)}{\mathrm{d}\mu} = -\alpha_s F(k,\mu), \qquad (2.33)$$

whose general solution is easily seen to be

$$F(k,\mu) = F(k,\mu_0)\, U(\mu_0,\mu), \qquad U(\mu_0,\mu) = \exp\left(\alpha_s \ln \frac{\mu_0}{\mu}\right). \qquad (2.34)$$

Using the definition in eq. (2.18) to set $\mu_0 = k|_+$ we obtain

$$\begin{aligned}
F(k,\mu) &= F(k,\mu_0)\exp\left(\alpha_s \ln \frac{\mu_0}{\mu}\right)\bigg|_{\mu_0 = k|_+} = \frac{\mathrm{d}}{\mathrm{d}k}\left[\int^k \mathrm{d}k'\, F(k',k)\exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}k}\left[(1+\cdots)\theta(k)\exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right] \\
&= \delta(k)\exp\left(\alpha_s \ln \frac{\xi}{\mu}\right) + \left[\theta(k)\frac{\mathrm{d}}{\mathrm{d}k}\exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right]_+^{\xi} \\
&= \delta(k) + \alpha_s\left[\frac{\theta(k)}{k}\exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right]_+^{\mu}.
\end{aligned} \qquad (2.35)$$

The k'-integral in the first step only acts on $F(k')$ and eliminates all distributions leaving only constant boundary terms $1 + \cdots$. In the second step we used eq. (B.4) to take the derivative involving the $\theta(k)$, where ξ is arbitrary and cancels between the two terms. In the last step we chose $\xi = \mu$. Expanding the exponential in α_s, we can easily see that this reproduces the full distributional logarithmic structure of eq. (2.32)

$$\begin{aligned}
F(k,\mu) &= \delta(k) + \alpha_s\left[\frac{\theta(k)}{k}\left(1 + \alpha_s \ln \frac{k}{\mu} + \frac{1}{2}\alpha_s^2 \ln^2 \frac{k}{\mu} + \cdots\right)\right]_+^{\mu} \\
&= \delta(k) + \alpha_s \mathcal{L}_0(k,\mu) + \alpha_s^2 \mathcal{L}_1(k,\mu) + \frac{1}{2}\alpha_s^3 \mathcal{L}_2(k,\mu) + \cdots.
\end{aligned} \qquad (2.36)$$

With the general properties in eqs. (2.24) and (2.25), we can also simply use the fact that setting $\mu_0 = k|_+$ eliminates any contributions from distributional logarithms in the boundary condition $F(k,\mu_0)$. We can therefore directly plug in $F(k,\mu_0 = k|_+) = F[k] = (1 + \cdots)\delta(k)$ for the boundary condition to get

$$F(k,\mu) = F[k]\, U(\mu_0,\mu)\bigg|_{\mu_0 = k|_+} = (1+\cdots)\delta(k)\, U(\mu_0,\mu)\bigg|_{\mu_0 = k|_+}. \qquad (2.37)$$

Using eq. (2.25) we can see that this is identical to what we got in eq. (2.35),

$$\begin{aligned}
F(k,\mu) &= \delta(k)\, U(\mu_0,\mu)\bigg|_{\mu_0 = k|_+} \\
&= \delta(k) + \left[\theta(k)\frac{\mathrm{d}}{\mathrm{d}k}U(k,\mu)\right]_+^{\mu} = \delta(k) + \alpha_s\left[\frac{\theta(k)}{k}\exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right]_+^{\mu}.
\end{aligned} \qquad (2.38)$$

Alternative derivation Another way to derive the same solution without distributional scale setting is to start from the general Ansatz

$$F(k, \mu) = f_0(\mu)\, \delta(k) + \Big[\theta(k)\, f_1(k, \mu)\Big]_+^{\mu}.$$ (2.39)

Plugging this back into eq. (2.33) and using eq. (B.7) to take the μ derivative,

$$\mu \frac{\mathrm{d}F(k, \mu)}{\mathrm{d}\mu} = \delta(k)\left(\mu \frac{\mathrm{d}f_0(\mu)}{\mathrm{d}\mu} - \mu f_1(\mu, \mu)\right) + \Big[\theta(k)\, \mu \frac{\mathrm{d}f_1(k, \mu)}{\mathrm{d}\mu}\Big]_+^{\mu}$$
$$\overset{!}{=} -\alpha_s f_0(\mu)\, \delta(k) - \alpha_s \Big[\theta(k)\, f_1(k, \mu)\Big]_+^{\mu},$$ (2.40)

yields a coupled system of differential equations for f_0 and f_1,

$$\mu \frac{\mathrm{d}f_1(k, \mu)}{\mathrm{d}\mu} = -\alpha_s f_1(k, \mu)\,,$$
$$\mu \frac{\mathrm{d}f_0(\mu)}{\mathrm{d}\mu} = -\alpha_s f_0(\mu) + \mu f_1(\mu, \mu)\,.$$ (2.41)

The advantage is that these are now two ordinary RGEs, which can be solved straightforwardly without having to be careful about producing distributions from taking derivatives of $\theta(k)$. Solving for f_1 in terms of f_0,

$$f_1(k, \mu) = f_1(k, k) \exp\left(\alpha_s \ln \frac{k}{\mu}\right) = \left[\frac{\mathrm{d}f_0(k)}{\mathrm{d}k} + f_0(k) \frac{\alpha_s}{k}\right] \exp\left(\alpha_s \ln \frac{k}{\mu}\right)$$
$$= \frac{\mathrm{d}}{\mathrm{d}k} f_0(k) \exp\left(\alpha_s \ln \frac{k}{\mu}\right).$$ (2.42)

Plugging back into the Ansatz, we obtain the result

$$F(k, \mu) = f_0(\mu)\, \delta(k) + \left[\theta(k) \frac{\mathrm{d}}{\mathrm{d}k} f_0(k) \exp\left(\alpha_s \ln \frac{k}{\mu}\right)\right]_+^{\mu},$$ (2.43)

which for $f_0 = 1 + \cdots$ confirms the earlier result eq. (2.35). This form suggests to interpret $f_0(\mu)$ as the boundary term, as it completely predicts the (logarithmic) structure of F. Indeed, setting now $\mu = k|_+$ yields $F[k] = f_0(\xi)\, \delta(k) + [\theta(k) f_0'(k)]_+^{\xi}$, confirming that the boundary term can be obtained as the coefficient $f_0(\mu)$ of $\delta(k)$. Note that the essential element in this derivation is the same as above, namely that with the distributional canonical scale setting $F(k, \mu = k|_+) = F[k]$ reduces to a pure boundary term that is free of logarithmic distributions. In actual applications, it would be calculated in a fixed-order expansion in $\alpha_s(\mu)$.

This simple toy example illustrates that with the distributional scale setting, using the canonical scale choice $\mu_0 = k|_+$ as in eq. (2.37) becomes exactly analogous to using the ordinary canonical scale choice in the multiplicative case in eq. (2.5). In particular, the pure boundary condition $F[k]$ can now be calculated in fixed-order perturbation theory, while the RGE solution predicts the complete distributional logarithmic structure. It should be evident that this is true generically also for more complicated cases. We will encounter two more complicated examples when discussing the RGE for the rapidity anomalous dimension in sec. 4 and for the soft function in sec. 5.

2.5 Comparison to evolution in conjugate space

It is interesting to compare the result of the momentum-space evolution to the evolution in appropriate conjugate spaces, i.e. solving the RGEs in conjugate space with scale setting therein. In general, using equivalent boundary conditions in different spaces can lead to different predictions in physical space. To illustrate this, we consider again the simple example distributions D_n, satisfying

$$\mu \frac{\mathrm{d}D_n(k,\mu)}{\mathrm{d}\mu} = -\alpha_s \, n\mathcal{L}_{n-1}(k,\mu) \qquad (n \geq 1)\,, \tag{2.44}$$

and assume that only the LO boundary term is known from a fixed-order calculation,

$$D_n(k,\mu) = \delta(k) + \cdots\,, \tag{2.45}$$

where the ellipses denote the unknown higher order terms. Since all unknown logarithmic higher order terms vanish for $\mu = k|_+$, the required LO boundary term to solve eq. (2.44) distributionally is $D[k] = \delta(k)$. The solution then is (see sec. 2.3)

$$D_n(k,\mu) = \delta(k) + \alpha_s \mathcal{L}_n(k,\mu)\,. \tag{2.46}$$

In the following we compare this momentum-space result to the resummation in cumulant and Fourier space.

2.5.1 Cumulant space

Taking the cumulant, eqs. (2.44) and (2.45) become

$$\mu \frac{\mathrm{d}\bar{D}_n(k_{\mathrm{cut}},\mu)}{\mathrm{d}\mu} = -\alpha_s \, \theta(k_{\mathrm{cut}}) \ln^n \frac{k_{\mathrm{cut}}}{\mu}\,, \tag{2.47}$$

$$\bar{D}_n(k_{\mathrm{cut}},\mu) = \theta(k_{\mathrm{cut}}) + \cdots\,, \tag{2.48}$$

where the cumulant is defined as

$$\bar{D}_n(k_{\mathrm{cut}},\mu) = \int^{k_{\mathrm{cut}}} \mathrm{d}k' \, D_n(k',\mu)\,. \tag{2.49}$$

The solution to eq. (2.47) is easily obtained as

$$\bar{D}_n(k_{\mathrm{cut}},\mu) = \bar{D}_n(k_{\mathrm{cut}},\mu_0) - \alpha_s \, \theta(k_{\mathrm{cut}}) \int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \ln^n \frac{k_{\mathrm{cut}}}{\mu'}\,. \tag{2.50}$$

The important point is that all distributions in momentum space correspond to logarithms $\ln(k_{\mathrm{cut}}/\mu)$ in cumulant space, which can be fully resummed by choosing $\mu_0 = k_{\mathrm{cut}}$. With this choice, all logarithms in $\bar{D}_n(k_{\mathrm{cut}},\mu)$ are eliminated, and from eq. (2.48) it follows that the LO boundary condition is $\bar{D}_n(k_{\mathrm{cut}},k_{\mathrm{cut}}) = \theta(k_{\mathrm{cut}})$. The solution is thus given by

$$\bar{D}_n(k_{\mathrm{cut}},\mu) = \theta(k_{\mathrm{cut}}) + \frac{\alpha_s}{n+1}\theta(k_{\mathrm{cut}}) \ln^{n+1} \frac{k_{\mathrm{cut}}}{\mu}\,. \tag{2.51}$$

Transforming back to momentum space by taking the derivative with respect to $k_{\rm cut}$ yields

$$D_n(k,\mu) = \delta(k) + \alpha_s \left[\frac{\theta(k)}{k} \ln^n \frac{k}{\mu} \right]_+^\mu = \delta(k) + \alpha_s \mathcal{L}_n(k,\mu) \,, \tag{2.52}$$

which exactly reproduces the momentum-space solution in eq. (2.46). Hence, resummation in cumulant and momentum space to predict the logarithmic structure is equivalent for this example. This is of course not very surprising, considering the intimate relation between distribution and cumulant space.

2.5.2 Fourier space

Taking the Fourier transform of eqs. (2.44) and (2.45) using eq. (B.13), we find

$$\mu \frac{\mathrm{d}\tilde{D}_n(y,\mu)}{\mathrm{d}\mu} = -\alpha_s \sum_{k=0}^n (-1)^k \binom{n}{k} \ln^k \big(iy\mu e^{\gamma_E} \big) R_1^{(n-k)} \,, \tag{2.53}$$

$$\tilde{D}_n(y,\mu) = 1 + \cdots \,, \tag{2.54}$$

where the Fourier transform is defined as

$$\tilde{D}_n(y,\mu) = \int \mathrm{d}k\, e^{-iky} D_n(k,\mu) \,, \tag{2.55}$$

and the constant $R_1^{(n)}$ is defined as (see appendix B.2 for more details)

$$R_1^{(n)} = \left. \frac{\mathrm{d}^n}{\mathrm{d}a^n} e^{\gamma_E a} \Gamma(1+a) \right|_{a=0} \,. \tag{2.56}$$

The general solution is given by

$$\tilde{D}_n(y,\mu) = \tilde{D}_n(y,\mu_0) - \alpha_s \int_{\mu_0}^\mu \frac{\mathrm{d}\mu'}{\mu'} \sum_{k=0}^n (-1)^k \binom{n}{k} \ln^k \big(iy\mu' e^{\gamma_E} \big) R_1^{(n-k)} \,. \tag{2.57}$$

Since all logarithms depend on $iy\mu e^{\gamma_E}$, they are fully resummed by choosing $\mu_0 = -ie^{-\gamma_E}/y$,[3] which eliminates all logarithms in the boundary term. This allows us to use the known fixed-order boundary term eq. (2.54), $\tilde{D}_n(y,\mu_0) = 1$, and gives

$$\tilde{D}_n(y,\mu) = 1 - \alpha_s \frac{R_1^{(n+1)}}{n+1} + \alpha_s \mathrm{FT}[\mathcal{L}_n](y,\mu) \,. \tag{2.58}$$

Transforming back to momentum space thus yields

$$D_n(k,\mu) = \delta(k) + \alpha_s \left[\mathcal{L}_n(k,\mu) - \frac{R_1^{(n+1)}}{n+1} \delta(k) \right] \,. \tag{2.59}$$

---

[3]In principle, one is of course free to shift constants from the logarithms into subleading terms. This convention for the canonical logarithm in conjugate space is the most natural one, since at least at lowest order $n=0$ the pure plus distribution $\mathcal{L}_0(k,\mu)$ corresponds to a pure logarithm $\ln(iy\mu e^{\gamma_E})$, see table 1 in appendix B.

While the plus distribution matches the one in eq. (2.46), which is of course necessary to provide the correct μ-dependence, the solution contains an additional term $-\delta(k)R_1^{(n+1)}/(n+1)$ at $\mathcal{O}(\alpha_s)$.

This shows explicitly that the Fourier-space resummation, which fundamentally resums logarithms of the conjugate variable y, is not equivalent to the momentum-space resummation, as it induces an additional boundary term. This discrepancy is due to the fact that pure plus distributions $\mathcal{L}_n(k,\mu)$ do not correspond to pure powers of $\ln(iy\mu e^{\gamma_E})$ and vice versa.

In principle, the additional $\alpha_s\delta(k)$ term in eq. (2.59) would be compensated for if we were to include the boundary condition $\tilde{D}_n(y,\mu_0)$ in Fourier space to $\mathcal{O}(\alpha_s)$. However, for a general distribution, such spurious terms are generated to all orders in α_s. (We will see this explicitly in sec. 4 for the example of the rapidity anomalous dimension.) Thus in practice for the purposes of resummation, where the boundary condition is only calculated to some fixed order, the resummation in Fourier space intrinsically induces additional subleading terms to all orders in α_s, which is one the main motivations for performing the resummation directly in momentum space.

2.6 Implementation of scale variations and profile scales

For phenomenological applications it is necessary to also use noncanonical scale choices. First, varying the scales away from their strict canonical values is a standard and convenient way to probe the size of higher-order logarithms and thereby estimate perturbative uncertainties. Furthermore, it is important to be able to dynamically turn off the resummation in the fixed-order region of the distribution. A standard way to achieve this is to employ profile scales $\mu_0(k)$ [27, 28], which smoothly interpolate as a function of k between the strict canonical scale choice in the resummation region and a common fixed scale μ_{FO} in the fixed-order region. Profile scales are also used to implement dynamical scale variations to distinguish different sources of perturbative uncertainties, for example resummation and fixed-order uncertainties [29, 31, 32].

We discuss the implementation of scale variations and more generally profile scales within our framework using the exponential toy example in eq. (2.32) with the associated RGE solution in eq. (2.34):

$$F(k,\mu) = F^{\text{FO}}(k,\mu_0)U(\mu_0,\mu)\,, \qquad U(\mu_0,\mu) = \exp\Big(\alpha_s \ln\frac{\mu_0}{\mu}\Big)\,. \tag{2.60}$$

Here we have included the superscript "FO" to stress that the boundary condition is obtained from a fixed-order calculation, whereas U is the evolution kernel. In a typical application, one wants to perform the resummation using the canonical scale choice $\mu_0 = k|_+$ for $k \ll Q$, where the logarithms of k are large and should be resummed. On the other hand, for $k \sim Q$, one wants to turn off the resummation and recover the exact fixed-order result by taking $\mu_0 = \mu_{\text{FO}} = \mu$. Both requirements are fulfilled by choosing μ_0 to be a profile function $\mu_0(k)$ behaving as

$$\mu_0(k) \to k\,, \qquad k \ll Q \tag{2.61}$$

$$\mu_0(k) \to \mu_{\text{FO}} = \mu\,, \quad k \sim Q\,. \tag{2.62}$$

The profile furthermore smoothly interpolates between the two regimes to capture the turning-off of the resummation. Such a profile can be conveniently implemented by generalizing the distributional scale setting in eq. (2.18), $\mu_0 = k|_+$, to a generic function $\mu_0 = \mu_0(k)|_+$,

$$F(k,\mu) = \frac{\mathrm{d}}{\mathrm{d}k}\left[\int^k \mathrm{d}k'\, F^{\mathrm{FO}}(k',\mu_0(k))U(\mu_0(k),\mu)\right].$$ (2.63)

In the resummation regime, this reproduces exactly the canonical distributional scale setting $\mu_0 = k|_+$. In the fixed-order regime, $U = 1$ and μ_0 is independent of k, such that the derivative simply inverts the integral to reproduce $F(k,\mu) = F^{\mathrm{FO}}(k,\mu)$.

Furthermore, the profile should allow scale variations to probe subleading logarithms. To see in more detail how this works, consider the NLO boundary term

$$F^{\mathrm{FO}}(k,\mu_0) = \delta(k) + \alpha_s\left[f_1\delta(k) + \mathcal{L}_0(k,\mu_0)\right].$$ (2.64)

The solution with arbitrary profile $\mu_0(k)$ is then

$$\begin{aligned}
F(k,\mu) &= \frac{\mathrm{d}}{\mathrm{d}k}\left\{\theta(k)\left(1 + \alpha_s\left[f_1 + \ln\frac{k}{\mu_0(k)}\right]\right)U(\mu_0(k),\mu)\right\} \\
&= \delta(k)\left(1 + \alpha_s\left[f_1 + \ln\frac{\mu}{\mu_0(\mu)}\right]\right)U(\mu_0(\mu),\mu) \\
&\quad + \alpha_s\left[\frac{\theta(k)}{k}\left(1 + \alpha_s\left[f_1 + \ln\frac{k}{\mu_0(k)}\right]\frac{\mathrm{d}\ln\mu_0(k)}{\mathrm{d}\ln k}\right)U(\mu_0(k),\mu)\right]_+^\mu.
\end{aligned}$$ (2.65)

The canonical scale choice to predict all logarithmic terms to this order is $\mu_0(k) = k$, such that we get

$$F(k,\mu) = \left(1 + \alpha_s f_1\right)\left[\delta(k) + \alpha_s\mathcal{L}_0 + \alpha_s^2\mathcal{L}_1 + \cdots\right],$$ (2.66)

which is exactly the expected structure.

More general, we can vary the scale μ_0 around the canonical value k, $\mu_0(k) = a \cdot k$, to probe subleading logarithms in the resummation regime. With this choice, we obtain

$$F(k,\mu) = \left[1 + \alpha_s(f_1 - \ln a)\right]\left[\delta(k) + \alpha_s\mathcal{L}_0 + \alpha_s^2\mathcal{L}_1 + \cdots\right]e^{\alpha_s\ln a}.$$ (2.67)

This clearly probes subleading terms through the induced exponential. Note that the $-\ln a$ in the first term cancels the $\mathcal{O}(\alpha_s)$-term of the exponential to correctly reproduce the fixed-order content.

Lastly, we can test the fixed-order structure by varying μ_0 in the fixed-order regime. To fully turn off the evolution kernel in eq. (2.63), we choose $\mu_0(k) = \mu = a \cdot \mu_{\mathrm{FO}}$, yielding

$$\begin{aligned}
F(k, a\mu_{\mathrm{FO}}) &= \delta(k) + \alpha_s\left[f_1\delta(k) + \mathcal{L}_0(k, a\mu_{\mathrm{FO}})\right] = F^{\mathrm{FO}}(k, a\mu_{\mathrm{FO}}) \\
&= \delta(k) + \alpha_s\left[(f_1 - \ln a)\delta(k) + \mathcal{L}_0(k, \mu_{\mathrm{FO}})\right] = F^{\mathrm{FO}}(k, \mu_{\mathrm{FO}}) - \alpha_s\ln a\,\delta(k).
\end{aligned}$$ (2.68)

The effect of such a scale variation is to induce the term $\alpha_s\ln(a)\delta(k)$, which precisely probes the $\mathcal{O}(\alpha_s)$ boundary term. For $a = 1$, this obviously restores the FO content.

For a more general scale variation, one can also vary $\mu_0(k)$ and μ independently. We consider the example $\mu_0(k) = a \cdot \mu_{\mathrm{FO}}$, $\mu = \mu_{\mathrm{FO}}$, where we only vary μ_0. This gives

$$F(k, \mu_{\mathrm{FO}}) = \delta(k) + \alpha_s \big[(f_1 - \ln a)\delta(k) + \mathcal{L}_0(k, \mu_{\mathrm{FO}}) \big] e^{\alpha_s \ln a} \,. \tag{2.69}$$

The additional exponential probes higher order logarithms and is a remnant of the mismatch between the resummation scale μ_0 and the common scale μ.

Distributional scale setting thus allows for a straightforward implementation of profile scales, which can be used to smoothly transition from resummation to fixed-order regime and allow one to probe the typical size of subleading terms and thereby estimate perturbative uncertainties in both regimes. Of course, in actual applications one has to choose suitable profile functions and variations such that they produces reasonable uncertainty estimates.

2.7 Distributional scale setting in 2D

In this section we collect the corresponding formulas for distributional scale setting and solving distributional differential equations for the two-dimensional case. The derivations are almost identical to the one-dimensional case, and are not repeated here.

We consider the example of transverse momentum dependent functions, which will be used throughout the rest of this paper. Such functions inherently contain divergences as $1/p_T^2$ in the limit of small transverse momenta $\vec{p}_T$, which are regulated through plus distributions. These are defined through the conditions

$$\Big[f(\vec{p}_T, \mu) \Big]_+^\mu = f(\vec{p}_T, \mu) \qquad \text{for } |\vec{p}_T| > 0 \,, \tag{2.70}$$

$$\int_{|\vec{p}_T| \leq \mu} \mathrm{d}^2\vec{p}_T \, \Big[f(\vec{p}_T, \mu) \Big]_+^\mu = 0 \,, \tag{2.71}$$

see appendix C for more details. Important examples are the logarithmic distributions

$$\mathcal{L}_n(\vec{p}_T, \mu) \equiv \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \ln^n \frac{\vec{p}_T^2}{\mu^2} \right]_+^\mu \equiv \frac{1}{\pi\mu^2} \mathcal{L}_n\!\left(\frac{\vec{p}_T^2}{\mu^2} \right) \,, \tag{2.72}$$

which are the two-dimensional analog of the $\mathcal{L}_n(k, \mu)$ in eq. (2.17). As in the one-dimensional case, the boundary condition of the $\mathcal{L}_n(\vec{p}_T, \mu)$ encodes a logarithmic dependence $\ln^n(p_T/\mu)$, which can be seen by shifting their boundary condition [see eq. (C.4)],

$$\mathcal{L}_n(\vec{p}_T, \mu) = \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \ln^n \frac{\vec{p}_T^2}{\mu^2} \right]_+^{\mu_0} + \frac{\delta(\vec{p}_T)}{n+1} \ln^{n+1} \frac{\mu_0^2}{\mu^2} \tag{2.73}$$

For a general two-dimensional distribution $D(\vec{p}_T, \mu)$ we define the distributional scale setting $\mu = p_T|_+$ as

$$\boxed{ D(\vec{p}_T, \mu = p_T|_+) \equiv \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \left[\int_{|\vec{k}_T| \leq p_T} \mathrm{d}^2\vec{k}_T \, D(\vec{k}_T, \mu = p_T) \right] . } \tag{2.74}$$

The derivative acts on everything inside square brackets, and $\mu = p_T$ can be set normally in the integrand since the integral runs over the auxiliary vector $\vec{k}_T$. Like in the one-dimensional case, eq. (2.74) sets purely distributional terms to zero,

$$\mathcal{L}_n(\vec{p}_T, \mu = p_T|_+) = 0\,, \qquad \left[f(\vec{p}_T, \mu)\right]_+^\mu \Big|_{\mu = p_T|_+} = 0\,, \qquad (2.75)$$

since the cumulant integral exactly vanishes by definition of the plus distribution, whereas any μ-independent constant (boundary) terms are not affected at all. For convenience, we denote the pure μ-independent boundary terms as

$$D[\vec{p}_T] \equiv D(\vec{p}_T, \mu)\Big|_{\mu = p_T|_+}\,. \qquad (2.76)$$

Here, $D[\vec{p}_T]$ can only depend on the quantity $\vec{p}_T$. The simplest case is $D[\vec{p}_T] \sim \delta(\vec{p}_T)$, but in general it can also contain regular (integrable) functions of $\vec{p}_T$. We will see examples of more general boundary terms in secs. 4 and 5.

The distributional scale setting can be readily applied to solve differential equations with two-dimensional distributions. The solution to the differential equation

$$\mu \frac{\mathrm{d}D(\vec{p}_T, \mu)}{\mathrm{d}\mu} = G(\vec{p}_T, \mu) \qquad (2.77)$$

is given by

$$D(\vec{p}_T, \mu) = D[\vec{p}_T] + \int_{p_T|_+}^\mu \frac{\mathrm{d}\mu'}{\mu'} G(\vec{p}_T, \mu')\,, \qquad (2.78)$$

where the integral over the distribution $G(\vec{p}_T, \mu)$ starting at the canonical scale $\mu = p_T|_+$ is given by

$$\int_{p_T|_+}^\mu \frac{\mathrm{d}\mu'}{\mu'} G(\vec{p}_T, \mu') \equiv \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \int_{|\vec{k}_T| \leq p_T} \mathrm{d}^2\vec{k}_T \int_{p_T}^\mu \frac{\mathrm{d}\mu'}{\mu'} G(\vec{k}_T, \mu')\,, \qquad (2.79)$$

A list of useful integrals is given in appendix C.4.

We conclude this section by comparing the resummation in distribution space with resummation in Fourier space, analogous to sec. 2.5. Specifying to the example

$$\mu \frac{\mathrm{d}D_n(\vec{p}_T, \mu)}{\mathrm{d}\mu} = -2n\alpha_s \mathcal{L}_{n-1}(\vec{p}_T, \mu) \qquad (n \geq 1)\,, \qquad (2.80)$$

and ignoring α_s running, the distributional solution for the simplest boundary term $D_n[\vec{p}_T] = \delta(\vec{p}_T)$ is given by

$$D_n(\vec{p}_T, \mu) = \delta(\vec{p}_T) + \alpha_s \mathcal{L}_n(\vec{p}_T, \mu)\,. \qquad (2.81)$$

The solution obtained in cumulant space is again equivalent.

In contrast, solving the RGE with scale setting in Fourier space with the boundary condition $\tilde{D}_n(\vec{b}_T, \mu_0 = b_0/b_T) = 1$ yields

$$D_n(\vec{p}_T, \mu) = \delta(\vec{p}_T) + \alpha_s \left[\mathcal{L}_n(\vec{p}_T, \mu) - \frac{R_2^{(n+1)}}{n+1}\right]\,, \qquad (2.82)$$

where the constant is given by eq. (C.17),

$$R_2^{(n)} = \frac{\mathrm{d}^n}{\mathrm{d}a^n} e^{2\gamma_E a} \frac{\Gamma(1+a)}{\Gamma(1-a)}\bigg|_{a=0}. \tag{2.83}$$

As for the one-dimensional case, we find that performing the resummation in Fourier space adds an additional boundary term compared to the momentum-space resummation. In this specific example, it could be compensated for by including the boundary condition $\tilde{D}_n(\vec{b}_T, \mu)$ in Fourier space to $\mathcal{O}(\alpha_s)$. However, for a general distribution, such spurious term are generated to all orders in α_s. Thus in practice, the resummation in Fourier space induces additional subleading term to all orders in α_s as well.

3 Overview and complications in q_T resummation

Many of the difficulties in the resummation of the q_T spectrum in momentum space are due to the intrinsic two-dimensional nature of $\vec{q}_T$ and the involved convolutions, and are absent for one-dimensional variables like transverse energy or thrust. In this section, we explore in detail the underlying reason for this. After briefly reviewing q_T factorization and the relevant associated RG equations in sec. 3.1, we argue in sec. 3.2 that the appearing two-dimensional convolutions requires very careful scale setting, which turns out to be the crucial complication of q_T resummation in momentum space. In particular, we will see that setting the boundary scales to the overall q_T does not correctly resum all logarithms, as one might naively expect. This will be illustrated in secs. 3.3 and 3.4 by reproducing a spurious divergence, which is well-known in the literature and as we will show originates from this incorrect scale setting, or more generally from an incorrect treatment of the boundary condition. In the following secs. 4 and 5 we then show how the q_T resummation via RG evolution in momentum space can be carried out using the distributional scale setting of sec. 2.

3.1 Review of q_T factorization

We briefly review the factorization theorem for the transverse-momentum distribution and set up our notation. We follow the rapidity renormalization formalism of refs. [15, 33] using soft-collinear effective theory (SCET) [34–37]. There are several other formalisms known in the literature, in particular the original formulation by Collins, Soper, and Sterman (CSS) [38–40] and its applications and extensions [41–55], as well as other work in SCET [16–18, 56–58]. A recent review of transverse-momentum dependent factorization and evolution can be found in ref. [59].

We consider the measurement of the transverse momentum $\vec{q}_T$ of a color-singlet system X with invariant mass Q and total rapidity Y. In the limit of very small transverse momentum, $q_T \ll Q$, the cross section is dominated by soft and collinear gluon emissions from the incoming partons that recoil against the hard system X. The emissions cause large logarithms $\ln^m(q_T^2/Q^2)$, with a power $m \leq 2n-1$ at order α_s^n, which we like to resum to all orders to retain predictive power in the perturbative series at small q_T. In the limit

$q_T \ll Q$, the cross section can be factorized as [15]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} \equiv \frac{1}{q_T} \frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}q_T \mathrm{d}\phi} = \sigma_0 \, H(Q,\mu) \int \mathrm{d}^2\vec{k}_a \, \mathrm{d}^2\vec{k}_b \, \mathrm{d}^2\vec{k}_s \, \delta(\vec{q}_T - \vec{k}_a - \vec{k}_b - \vec{k}_s)$$
$$\times \, B_a(\omega_a, \vec{k}_a, \mu, \nu) \, B_b(\omega_b, \vec{k}_b, \mu, \nu) \, S(\vec{k}_s, \mu, \nu) \,, \qquad (3.1)$$

where

$$\omega_a = Q e^Y \,, \qquad \omega_b = Q e^{-Y} \,, \qquad (3.2)$$

and for simplicitly we have kept the sum over partonic channels (and helicities in case of incoming gluons) implicit. Equation (3.1) is valid up to power corrections in q_T/Q. We are only interested in the leading-power cross section, which contains all singular logarithms, and drop the power correction.

Here, σ_0 denotes the Born cross section and $H(Q,\mu)$ is the hard function containing virtual corrections to the partonic process. Following refs. [60, 61], the bare beam functions B_i are defined in SCET in terms of forward proton matrix elements of collinear quark and gluon operators,

$$B_q^{\mathrm{bare}}(\omega, \vec{p}_T) = \theta(\omega) \, \langle P(p_n)| \, \bar{\chi}_n \frac{\slashed{\bar{n}}}{2} \delta(\omega - \overline{\mathcal{P}}_n) \, \delta(\vec{p}_T - \mathcal{P}_{\perp n}) \chi_n \, |P(p_n)\rangle \,,$$
$$B_g^{\mu\nu \, \mathrm{bare}}(\omega, \vec{p}_T) = \omega\theta(\omega) \, \langle P(p_n)| \, \mathrm{Tr}\{\mathcal{B}_{n,\perp}^\mu \, \delta(\omega - \overline{\mathcal{P}}_n) \, \delta(\vec{p}_T - \mathcal{P}_{\perp n}) \, \mathcal{B}_{n,\perp}^\nu\} \, |P(p_n)\rangle \,. \qquad (3.3)$$

They encode the effects of collinear initial-state radiation and are equivalent to transverse-momentum dependent parton distributions. They depend on the flavor i and light-cone momentum $\omega = xp_n^-$ of the parton that enters the hard interaction, where p_n is the proton momentum. They also depend on the total transverse momentum $\vec{p}_T$ of collinear initial-state radiation that was emitted prior to the hard interaction. The soft function S measures the total transverse momentum originating from soft emissions and is defined as the vacuum matrix element

$$S_{gg}^{\mathrm{bare}}(\vec{p}_T) = \frac{1}{N_c^2 - 1} \, \langle 0| \, \mathrm{Tr}\{\bar{T}[S_{n\perp}^\dagger S_{\bar{n}\perp}] \, \delta(\vec{p}_T - \mathcal{P}_{\perp s}) \, T[S_{\bar{n}\perp}^\dagger S_{n\perp}]\} \, |0\rangle \,, \qquad (3.4)$$

$$S_{qq}^{\mathrm{bare}}(\vec{p}_T) = \frac{1}{N_c} \, \langle 0| \, \mathrm{Tr}\{\bar{T}[S_{n\perp}^\dagger S_{\bar{n}\perp}] \, \delta(\vec{p}_T - \mathcal{P}_{\perp s}) \, T[S_{\bar{n}\perp}^\dagger S_{n\perp}]\} \, |0\rangle \,. \qquad (3.5)$$

The $\mathcal{B}_{n,\perp}^\mu$ and χ_n are collinear gluon and quark fields in SCET, $\mathcal{P}$ is the momentum label operator, and the $S_{n,\perp}$ are soft Wilson lines along the light-cone direction n. For more details see refs. [15, 62].

The bare hard, beam, and soft functions are divergent quantities and require renormalization. The UV divergences are regulated as usual by dimensional regularization. The resulting renormalized functions, which appear in eq. (3.1), satisfy the renormalization group equations

$$\mu\frac{\mathrm{d}H(Q,\mu)}{\mathrm{d}\mu} = \gamma_H(Q,\mu) \, H(Q,\mu) \,, \qquad (3.6)$$

$$\mu\frac{\mathrm{d}B(\omega, \vec{p}_T, \mu, \nu)}{\mathrm{d}\mu} = \gamma_B(\omega, \mu, \nu) \, B(\omega, \vec{p}_T, \mu, \nu) \,, \qquad (3.7)$$

$$\mu\frac{\mathrm{d}S(\vec{p}_T, \mu, \nu)}{\mathrm{d}\mu} = \gamma_S(\mu, \nu) \, S(\vec{p}_T, \mu, \nu) \,, \qquad (3.8)$$

where the anomalous dimensions have the all-order structure

$$\gamma_H(Q,\mu) = 4\Gamma_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{Q}{\mu} + \gamma_H[\alpha_s(\mu)] \,, \qquad (3.9)$$

$$\gamma_B(\omega,\mu,\nu) = 2\Gamma_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{\nu}{\omega} + \gamma_B[\alpha_s(\mu)] \,, \qquad (3.10)$$

$$\gamma_S(\mu,\nu) = 4\Gamma_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{\mu}{\nu} + \gamma_S[\alpha_s(\mu)] \,. \qquad (3.11)$$

The μ-independence of the cross section implies the RG consistency relation

$$\gamma_H(\alpha_s) + 2\gamma_B(\alpha_s) + \gamma_S(\alpha_s) = 0 \,. \qquad (3.12)$$

The beam and soft functions furthermore depend on a rapidity scale ν, associated with an additional regulator required to regulate rapidity divergences. These arise because both soft and beam modes describe modes of virtuality $\mu^2 \sim q_T^2$, leading to an "overlap" of soft and collinear momentum regions which are not resolved by dimensional regularization There have been a variety of rapidity regulators suggested in the literature to deal with these divergences. In the original CSS approach a non-lightlike axial gauge was employed [38], whereas in recent work Wilson lines off the light cone are used [41, 63]. In the context of SCET, the utilized regulators include the analytic regulator acting on eikonal propagators [64–66], the η-regulator inserted into Wilson lines [15, 33], the δ-regulator adding a mass to eikonal propagators [56, 67], and the exponential regulator acting on the phase space [68]. For all of these approaches, the beam and soft functions are currently known to NNLO [62, 69–73]. These fixed-order ingredients are not necessary for the purpose of this paper, which only relies on the RGE structure of beam and soft functions.

For concreteness we employ the η-regulator together with the rapidity renormalization group [15, 33]. Our discussion can be applied similarly to other regulators. The crucial point is that the additional rapidity regulator induces an additional rapidity scale, here denoted as ν. In particular, ν is analogous to the scale ζ in the original CSS formulation ref. [38], see also ref. [59]. (In some formalisms or applications the rapidity scale is kept implicit and/or fixed to canonical values.)

The rapidity renormalization group equations are given by

$$\nu\frac{\mathrm{d}B(\omega,\vec{p}_T,\mu,\nu)}{\mathrm{d}\nu} = \int \mathrm{d}^2\vec{k}_T\,\gamma_{\nu,B}(\vec{k}_T,\mu)\,B(\omega,\vec{p}_T - \vec{k}_T,\mu,\nu) \,, \qquad (3.13)$$

$$\nu\frac{\mathrm{d}S(\vec{p}_T,\mu,\nu)}{\mathrm{d}\nu} = \int \mathrm{d}^2\vec{k}_T\,\gamma_{\nu,S}(\vec{k}_T,\mu)\,S(\vec{p}_T - \vec{k}_T,\mu,\nu) \,. \qquad (3.14)$$

The overall ν-independence of the cross section implies the consistency relation

$$\gamma_\nu(\vec{p}_T,\mu) \equiv \gamma_{\nu,S}(\vec{p}_T,\mu) = -2\gamma_{\nu,B}(\vec{p}_T,\mu) \,, \qquad (3.15)$$

which means that there is only one independent rapidity anomalous dimension, which we denote as γ_ν. The crucial difference to the μ-RGE is that the ν-RGE is intrinsically a convolution. The commutativity of $\mathrm{d}/\mathrm{d}\nu$ and $\mathrm{d}/\mathrm{d}\mu$ relates the μ and ν anomalous dimensions through

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\gamma_\nu(\vec{k}_T,\mu) = \nu\frac{\mathrm{d}}{\mathrm{d}\nu}\gamma_S(\mu,\nu)\delta(\vec{k}_T) = -4\Gamma_{\text{cusp}}[\alpha_s(\mu)]\delta(\vec{k}_T) \,. \qquad (3.16)$$

We stress that the singular logarithmic structure of the q_T spectrum is fully encoded in the RGE equations eqs. (3.6)–(3.16). While their precise definitions and derivation depend to some extent on the employed effective-field theory framework, equivalent evolution equations with the same momentum structure exist in all approaches. In particular, the original CSS formulation for q_T resummation is based on analogous evolution equations [38–40]. Also, the corresponding system of differential equations in Fourier space, which is often considered, is completely equivalent. The SCET framework with rapidity renormalization we use is convenient in that it makes the structure maximally general and explicit.

The nontrivial task at hand is to solve the RG equations with appropriate momentum-space boundary conditions in order to perform the resummation in momentum space. In other words, we want to predict the all-order distributional logarithmic structure in q_T from the RG evolution. As we will see, carrying out the RG evolution in momentum space is quite complicated due to the distributional nature coupled with the two-dimensional convolutions in $\vec{k}_T$. The correct solution for the momentum-space RG evolution is the main purpose of this paper.

An important comment concerns the definition of the formal resummation accuracy. In problems with Sudakov double logarithms, the cross section exponentiates into the form $\sigma \sim \exp[\alpha_s^n \ln^{n+1} + \alpha_s^n \ln^n + \alpha_s^n \ln^{n-1} + \cdots]$. Including the first set of terms in the exponent $\sim \alpha_s^n \ln^{n+1}$ corresponds to the leading-logarithmic (LL) order, including the next set of terms $\sim \alpha_s^n \ln^n$ corresponds to the next-to-leading logarithmic (NLL) order, and so forth. Alternatively, one can consider the logarithm of the cross section and count the corresponding terms in its α_s expansion. One issue with this way of counting logarithms is that it is intrinsically ambiguous, i.e., it is always possible to have slightly different definitions which agree to a given order, but differ by contributions that in one or the other definition are formally subleading. As we will see the q_T spectrum in momentum space does not exponentiate into a simple exponential, but rather it will have a generalized exponential structure in distribution space, which makes this counting of explicit logarithms even less well-defined. (We will see that this can even lead to producing spurious divergences in the resummed result from formally subleading terms.)

The exponential structure of the resummed q_T spectrum is fully encoded in its evolution equations, or equivalently the anomalous dimenions are in essence the generalized "logarithm" of the resummed cross section. Therefore, an easy, fully consistent, and unambiguous way to define the resummation order is strictly by the pure fixed-order expansions of the cusp and noncusp anomalous dimensions together with the fixed-order expansions of the *pure* boundary terms for the hard, beam, and soft functions. Since these fixed-order series are the fundamental input to the RG evolution, with everything else following from them, it makes sense to define the fundamental resummation accuracy solely by the perturbative accuracy at which these inputs are included. In particular, with this strict definition one is not allowed to disregard seemingly subleading logarithmic terms at intermediate stages. This also means that the various RG consistency relations should always be exactly fulfilled by the resummed result. Our goal is to derive the solution for the evolution and resummation in momentum space, i.e., with anomalous dimensions and boundary terms defined in momentum space, valid to in principle any order in this strict definition.

Note that given the resummed result at strict LL, NLL, etc. one can of course consider additional choices or further approximations to simplify the result and study their numerical impact, which however is not our concern here.

Notation: In the remainder of this paper, we denote the final transverse momentum of the produced color-singlet state X by $\vec{q}_T$, while the transverse-momentum argument of a specific function is typically denoted by $\vec{p}_T$, and the integration momenta in the convolution integrals are usually denoted by $\vec{k}_i$. Convolutions are often abbreviated as

$$(f \otimes g)(\vec{p}_T, \dots) \equiv \int d^2\vec{k}_T\, f(\vec{p}_T - \vec{k}_T, \dots)\, g(\vec{k}_T, \dots)\,, \tag{3.17}$$

where the ellipses denote possible additional variables. Multiple convolutions are abbreviated as

$$(f \otimes^n)(\vec{p}_T) \equiv \int d^2\vec{k}_1 \dots d^2\vec{k}_n\, f(\vec{k}_1) \dots f(\vec{k}_n)\, \delta(\vec{p}_T - \vec{k}_1 - \dots - \vec{k}_n)\,. \tag{3.18}$$

These formulas are also summarized in appendix A.

3.2   Implications of two-dimensional convolutions

The transverse momentum spectrum is generically given in terms of plus distributions, which are necessary to regulate its $1/q_T^2$ divergences. Furthermore, the factorized cross section and its RGEs involve two-dimensional convolutions, see eqs. (3.1) and (3.14). While the issue of scale setting with distributions has been addressed in sec. 2, we now discuss the implications of the two-dimensional convolutions.

To illustrate the arising subtleties, it is sufficient to focus on the rapidity RGE of the soft function, as evolving it to the beam scale $\nu = \nu_B \sim Q$ eliminates all rapidity logarithms in the beam functions entering eq. (3.1). A formal solution to eq. (3.14) is readily derived to be

$$S(\vec{p}_T, \mu, \nu_B) = \int d^2\vec{k}_T\, V(\vec{p}_T - \vec{k}_T, \mu, \nu_B, \nu_S)\, S(\vec{k}_T, \mu, \nu_S)\,, \tag{3.19}$$

where the rapidity evolution kernel V is given by

$$\begin{aligned}
V(\vec{p}_T, \mu, \nu_B, \nu_S) &= \delta(\vec{p}_T) + \sum_{n=1}^{\infty} \frac{1}{n!} \ln^n \frac{\nu_B}{\nu_S} (\gamma_\nu \otimes^n)(\vec{p}_T, \mu) \\
&= \delta(\vec{p}_T) + \ln\frac{\nu_B}{\nu_S}\gamma_\nu(\vec{p}_T, \mu) \\
&\quad + \frac{1}{2}\ln^2\frac{\nu_B}{\nu_S} \int d^2\vec{k}_1 \int d^2\vec{k}_2\, \gamma_\nu(\vec{k}_1, \mu)\, \gamma_\nu(\vec{k}_2, \mu)\, \delta(\vec{k}_1 + \vec{k}_2 - \vec{p}_T) \\
&\quad + \cdots\,.
\end{aligned} \tag{3.20}$$

Here $(\gamma_\nu \otimes^n)$ denotes n convolutions of γ_ν with itself, see eq. (3.18). V is an exponential in convolution space and eq. (3.20) can be derived equivalently in convolution space or Fourier space (see sec. 3.3 below). Taking the derivative with respect to ν, one can easily verify that it provides a solution for eq. (3.14).

The evolution kernel eq. (3.20) has a simple physical interpretation. Each factor of $\gamma_\nu(\vec{k}_i, \mu)$ corresponds to a real soft emission with momentum $\vec{k}_i$. The convolution integrals are the remaining transverse phase-space integrals, which are constrained such that the transverse momenta of all emissions sum up to the total $\vec{p}_T$. Each emission is dressed with a rapidity logarithm $\ln(\nu_B/\nu_S)$ to evolve in rapidity from the soft scale ν_S to the beam scale ν_B, which corresponds to the effective range in rapidity over which the soft emission has been integrated. The n emission term then scales with $\ln^n(\nu_B/\nu_S)$, which precisely builds up an exponential in convolution space.

To investigate the structure of V in more detail, we focus on the first nontrivial convolution. It involves integrating over two real emissions with momenta $\vec{k}_{1,2}$ such that $\vec{k}_1 + \vec{k}_2 = \vec{p}_T$. Figure 1 illustrates the momentum regions in $|\vec{k}_{1,2}|$ contributing to this integral. The region between the solid lines is allowed, while the gray region outside cannot fulfill the measurement constraint. The dashed lines correspond to a fixed angle between the two emissions of $\angle(\vec{k}_1, \vec{k}_2) = 90^\circ, 135^\circ$. The larger this angle is, the larger the allowed magnitudes $|\vec{k}_{1,2}|$ are. In the limit where the emissions are back-to-back, the magnitudes $|\vec{k}_{1,2}|$ can become infinitely large, as long as their difference still gives $\vec{p}_T$. This is the limit given by the two blue lines. Hence, the convolution integrals in eq. (3.20) in principle receive contributions from infinitely large momenta. Physically, the limit of both emissions having large $|\vec{k}_{1,2}| \sim Q$ should be power-suppressed in q_T/Q and hence not affect the singular logarithmic structure. However, in this limit the emissions are not correctly described anymore by the underlying soft expansion, which assumes $|\vec{k}_{1,2}| \sim p_T \ll Q$. Therefore, we can in principle receive spurious contributions to the integral from this region. On the other hand, there can also be relevant physical contributions from any intermediate region $p_T \ll |\vec{k}_{1,2}| \ll Q$ that must be correctly taken into account. It was already argued in ref. [74] (see also ref. [24]) that the very small $\vec{q}_T$ region can be influenced (or even be dominated) by such kinematic cancellations of harder emissions.

One might now be worried that the factorization theorem is intrinsically ill-defined, as it contains effects of arbitrarily hard emissions. This is unavoidable, as the soft approximation eliminates the phase-space constraints that would normally cut off such emissions, as already noticed in ref. [14]. However, all large rapidity logarithms should arise from the soft region where all emissions are of the order of the final q_T, $|\vec{k}_i| \sim q_T$. From the above observations, it is clear that the rapidity evolution kernel eq. (3.20) violates this requirement. This is perfectly fine because a priori eqs. (3.19) and (3.20) only shift logarithms $\ln(p_T/\nu_S)$ in $S(\vec{p}_T, \mu, \nu_S)$ into logarithms $\ln(p_T/\nu_B)$ in $S(\vec{p}_T, \mu, \nu_B)$. In particular, if $S(\vec{p}_T, \mu, \nu_S)$ is already resummed, eqs. (3.19) and (3.20) are valid solutions of the rapidity RGE.

The situation changes if we want to use the rapidity RGE to predict all logarithms of the soft function. Then we would like to start from a boundary condition for the soft function without any real emissions, and subsequently let the evolution add real emissions by convolving with γ_ν. Each such emission with momentum $\vec{k}_i$ should scale with a rapidity logarithm $\sim \ln(\nu_B/|\vec{k}_i|)$ to evolve in rapidity from its own emission scale to the beam scale. For the example of two emissions, we thus expect a contribution of the form

$$\sim \int d^2\vec{k}_1 \int d^2\vec{k}_2 \, \ln\frac{\nu_B}{|\vec{k}_1|}\gamma_\nu(\vec{k}_1, \mu) \ln\frac{\nu_B}{|\vec{k}_2|}\gamma_\nu(\vec{k}_2, \mu)\delta(\vec{p}_T - \vec{k}_1 - \vec{k}_2)\,. \qquad (3.21)$$
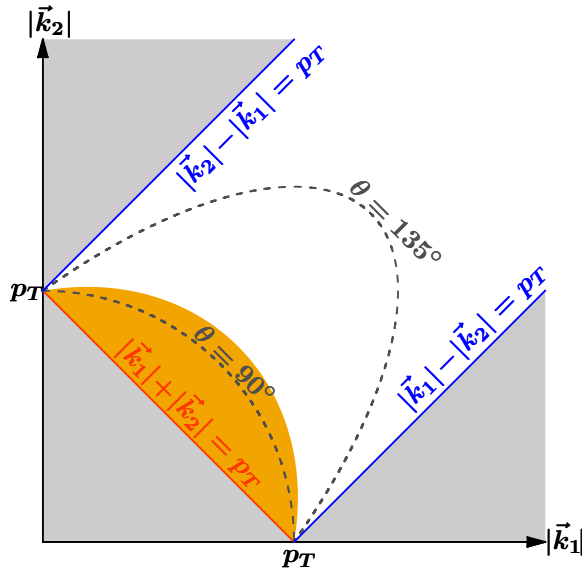
Figure 1. Illustration of the transverse momentum regions contributing to the convolution $(\gamma_\nu \otimes \gamma_\nu)(\vec{p}_T, \mu)$. The region between the solid blue and orange lines contributes to the convolution integral. The dashed lines correspond to a fixed angle $\theta = \angle(\vec{k}_1, \vec{k}_2) = 90, 135°$ between the two emissions. Canonical scaling $|\vec{k}_1| \sim |\vec{k}_2| \sim p_T$ is only fulfilled in the shaded orange region.

(For the moment we ignore that the rapidity logarithms need to be properly included into plus distributions, which will be taken care of in the final solution in sec. 5.) If the $\vec{k}_{1,2}$ integrals were dominated by $|\vec{k}_{1,2}| \sim p_T$, indicated by the shaded orange region in figure 1, both rapidity logarithms could be approximated as $\sim \ln(\nu_B/p_T)$, and a double logarithm $\ln^2(\nu_B/p_T)$ could be pulled out of the convolution integrals. This is precisely what happens in eq. (3.20). However, as explained above, the convolution also gets contributions from energetic emissions $|\vec{k}_{1,2}| \gg p_T$, which happen to be back-to-back such that $\vec{k}_1 + \vec{k}_2 = \vec{p}_T$. In this region, the rapidity logarithms $\ln(\nu_B/|\vec{k}_{1,2}|)$ in eq. (3.21) get smaller and eventually become irrelevant for $|\vec{k}_{1,2}| \sim Q$. Instead, the approximation $|\vec{k}_{1,2}| \sim p_T$ would result in a spuriously large rapidity logarithm $\ln^2(\nu_B/p_T)$, which would artificially enhance this otherwise suppressed phase-space region. To properly treat these emissions, it is necessary to keep the correct rapidity logarithms $\sim \ln(\nu_B/|\vec{k}_i|)$ in the convolutions. This is obviously not possible by simply choosing a fixed value for ν_S in eqs. (3.19) and (3.20). In particular, if we were to use the naive canonical choice $\nu_S \sim p_T$, we would artificially enhance unphysical contributions from energetic emissions. In sec. 3.4, this point will be stressed by showing that doing so actually produces a spurious divergence in the q_T-spectrum. The correct RG evolution to predict the resummed soft function is significantly more complicated than eq. (3.20), and in particular will rely on distributional scale setting. It will be derived in sec. 5.

So far we have focused on the role of the rapidity scale. The two-dimensional convolution structure also has implications for the scale μ entering the rapidity anomalous dimension and soft function. For this purpose, it suffices to consider eqs. (3.19) and (3.20),

which correctly evolve the (already resummed) soft function between two scales ν_S and ν_B. The final results in eqs. (3.19) and (3.20) after all integrations can only depend on p_T and thus can only contain logarithms $\ln(p_T/\mu)$. One might therefore expect that when performing the rapidity evolution at $\mu \sim p_T$, it should be sufficient to evaluate γ_ν at fixed order. However, what is important is that the rapidity anomalous dimension contains logarithms $\ln(|\vec{k}_i|/\mu)$ (in distributional form). These are only minimized by $\mu \sim p_T$ if the convolutions were dominated by $|\vec{k}_i| \sim p_T$. However, as argued above, since the convolution intrinsically probes momenta $|\vec{k}_i| \gg p_T$, setting $\mu \sim p_T$ can induce spurious logarithms $\ln(|\vec{k}_i|/p_T)$. Hence, the rapidity anomalous dimension γ_ν entering the convolutions should always be resummed in order to correctly describe emissions at any $|\vec{k}_i| \lesssim Q$. As we will see in sec. 4, the main effect of resumming $\gamma_\nu(\vec{k}_i, \mu)$ is to evaluate it at $\alpha_s(|\vec{k}_i|)$ rather than $\alpha_s(\mu)$. This suppresses the amplitude of energetic emissions, which is particularly important because there is no phase-space suppression due to the soft approximations, as discussed above.

At this point we can also discuss why these complications do not arise for scalar quantities like thrust or transverse energy. Taking transverse energy as the closest example[4] to q_T, the measurement constraint (for two emissions) changes to

$$\delta(\vec{k}_1 + \vec{k}_2 - \vec{p}_T) \quad \rightarrow \quad \delta(|\vec{k}_1| + |\vec{k}_2| - E_T). \tag{3.22}$$

This is clearly a much stronger constraint and forces both momenta to be of the order of the total transverse energy, $|\vec{k}_i| \sim E_T$. In figure 1, this corresponds to forcing the momenta to lie on the orange line, and it does not allow any contributions from the large momentum region. In particular, the rapidity logarithms $\ln(\nu_B/|\vec{k}_i|)$ are now well approximated by $\sim \ln(\nu_B/E_T)$. As we will see in the next subsection, this is precisely the reason why in the one-dimensional case the convolution structure of the RGE does not lead to a spurious singularity.

3.3 Rapidity evolution in Fourier space

To illustrate that the problems observed in the previous section are not an artifact of the momentum space approach, we now consider the resummation of the soft function in Fourier space. There, the rapidity RGE in eq. (3.14) becomes multiplicative,

$$\tilde{S}(\vec{b}_T, \mu, \nu) = \tilde{\gamma}_\nu(\vec{b}_T, \mu)\, \tilde{S}(\vec{b}_T, \mu, \nu), \tag{3.23}$$

where $\vec{b}_T$ is Fourier conjugate to $\vec{p}_T$ and the tilde denotes the Fourier transformed quantities. Equation (3.23) is easily solved by

$$\tilde{S}(\vec{b}_T, \mu, \nu) = \tilde{S}(\vec{b}_T, \mu, \nu_0) \exp\left[\ln \frac{\nu}{\nu_0} \tilde{\gamma}_\nu(\vec{b}_T, \mu)\right], \tag{3.24}$$

from which we recover the momentum space solution

$$S(\vec{p}_T, \mu, \nu) = \int \frac{\mathrm{d}^2 \vec{b}_T}{(2\pi)^2}\, e^{i\vec{p}_T \cdot \vec{b}_T} \tilde{S}(\vec{b}_T, \mu, \nu_0) \exp\left[\ln \frac{\nu}{\nu_0} \tilde{\gamma}_\nu(\vec{b}_T, \mu)\right]. \tag{3.25}$$

---

[4]The factorization of E_T is known to be affected by Glauber modes [75], but the E_T-distribution can still serve as an example for the different mathematical structure of the renormalization group equations.

For *arbitrary* ν_0, this is exactly equal to eqs. (3.19) and (3.20), and allows one to correctly shift logarithms from ν_0 to ν.

However, to instead predict the soft function including its logarithms, one needs to choose ν_0 suitably. Since the final $S(\vec{p}_T, \mu, \nu)$ result is known to contain logarithms $\ln(p_T/\nu)$, one might be tempted to set $\nu_0 = p_T$ and use the fixed-order boundary condition $\tilde{S}(\vec{b}_T, \mu, \nu_0) = 1 + \cdots$. However, we know that $\tilde{S}$ contains logarithms $\ln(\nu_0 b_T)$, and these are of course not eliminated by $\nu_0 \sim p_T$. Since the Fourier integral runs over all $\vec{b}_T$, they can in principle become relevant at small and large $\vec{b}_T$, in particular wherever the simple scaling $b_T \sim p_T^{-1}$ is violated. (This is the same situation as discussed in eq. (2.8).) This is analogous to setting ν_S in eq. (3.19) to the overall q_T and using the pure fixed-order boundary condition for $S(\vec{k}_T, \mu, \nu_S)$.

We will explicitly see in the next section that the $b_T \to 0$ region causes troubles if one were to set $\tilde{S}(\vec{b}_T, \mu, \nu_0) = 1$. In Fourier space, the problem is easily overcome by choosing $\nu_0 \sim 1/b_T$, which would allow one to evaluate the boundary condition in a pure fixed-order expansion. However, this corresponds to an intrinsic scale setting and thus resummation in Fourier space. The correct momentum-space analog corresponds to the discussion in eq. (3.21) and will be derived in sec. 5.

3.4 Illustration: effects from energetic emissions

In the sec. 3.2 we argued that the two-dimensional convolutions of γ_ν are intrinsically sensitive to large transverse momenta. As a result the formal solution eq. (3.20) does not allow one to correctly predict the all-order soft function because these energetic emissions are artificially enhanced by large rapidity logarithms. We will now explicitly demonstrate that this incorrect treatment is what causes a well-known spurious singularity in the evolution kernel.

We first calculate the rapidity evolution kernel V in eq. (3.20) using the fixed leading-order expression for γ_ν, which is given by

$$\gamma_\nu(\vec{p}_T, \mu) = 2\Gamma_{\text{cusp}}[\alpha_s(\mu)] \, \mathcal{L}_0(\vec{p}_T, \mu) \, . \tag{3.26}$$

Inserting this into eq. (3.20), the evolution kernel can be calculated either through Fourier transformation or by explicitly calculating $\mathcal{L}_0 \otimes^n$ [see eq. (C.24)]. Both techniques yield

$$V(\vec{p}_T, \mu, \nu, \nu_0) = \left[\delta(\vec{p}_T) + \omega_s \mathcal{L}^{\omega_s}(\vec{p}_T, \mu) \right] e^{-2\gamma_E \omega_s} \frac{\Gamma(1 - \omega_s)}{\Gamma(1 + \omega_s)} \, , \tag{3.27}$$

where

$$\omega_s = 2\Gamma_{\text{cusp}}[\alpha_s(\mu)] \ln \frac{\nu}{\nu_0} \, , \tag{3.28}$$

and the plus distribution is defined as [see eq. (C.8)]

$$\mathcal{L}^a(\vec{p}_T, \mu) \equiv \frac{1}{\pi\mu^2} \left[\left(\frac{\vec{p}_T^2}{\mu^2} \right)^{a-1} \right]_+^\mu \equiv \frac{1}{\pi\mu^2} \mathcal{L}^a(\vec{p}_T^2/\mu^2) \, . \tag{3.29}$$

The result in eq. (3.27) contains an explicit divergence at $\omega_s = 1$, which has been encountered before [14–17, 24]. Using what would seem to be the canonical scale choices,

$\mu = \nu_0 = q_T$ and $\nu = Q$, the divergence at $\omega_s = 1$ occurs when

$$\Gamma_{\text{cusp}}[\alpha_s(q_T)] = \ln^{-1}\frac{Q^2}{q_T^2}. \tag{3.30}$$

For illustration, for gluon-induced processes like $gg \to H$, where $\Gamma_{\text{cusp}} \sim C_A$, this happens at $q_T \approx 8\,\text{GeV}$ for $Q = 125\,\text{GeV}$ and $q_T \approx 27$ GeV for $Q = 1\,\text{TeV}$. For quark-induced process like Drell-Yan, where $\Gamma_{\text{cusp}} \sim C_F$, it occurs at $q_T \approx 2\,\text{GeV}$ for $Q = m_Z$ and $q_T \approx 4\,\text{GeV}$ for $Q = 1\,\text{TeV}$.

Clearly, the all-order q_T spectrum cannot contain such a singularity, especially since for large Q it happens at purely perturbative q_T. Its appearance for small enough q_T/Q respectively large enough $\ln(q_T/Q)$ indicates that the above naive attempt simply does not properly treat the resummation of logarithms. As we have shown, eq. (3.27) can be derived working entirely in momentum space and without any reference to Fourier space, which means that it is not related to a possibly ill-defined inverse Fourier transformation.

To show that the origin of this divergence is indeed due to contributions from large transverse momenta in the convolution, as argued earlier, we regulate the LO anomalous dimension in eq. (3.26) by introducing an explicit cut off Λ in momentum space,

$$\gamma_\nu(\vec{p}_T, \mu) \to \gamma_\nu^{(\Lambda)}(\vec{p}_T, \mu) \equiv \gamma_\nu(\vec{p}_T, \mu)\theta(\Lambda - p_T). \tag{3.31}$$

Since $\gamma_\nu(\vec{p}_T, \mu)$ corresponds to a single real emission with momentum $\vec{p}_T$, this is equivalent to cutting off emissions with large transverse momentum $|\vec{p}_T| > \Lambda$. Using eq. (3.31) in eq. (3.20), the rapidity evolution kernel becomes rather complicated and is most easily evaluated in Fourier space, where the regulated anomalous dimension eq. (3.31) reads

$$\tilde{\gamma}_\nu^{(\Lambda)}(\vec{b}_T, \mu) = 2\Gamma_{\text{cusp}}[\alpha_s(\mu)]\left[\ln\frac{\Lambda^2}{\mu^2} - \frac{b_T^2\Lambda^2}{4}\,{}_2F_3\left(-\frac{b_T^2\Lambda^2}{4}\right)\right], \tag{3.32}$$

where ${}_2F_3(x) = {}_2F_3(1,1;2,2,2;x)$ is a generalized hypergeometric function. The rapidity evolution kernel then is

$$\begin{aligned} V(\vec{p}_T, \mu, \nu, \nu_0) &= \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2}\, e^{i\vec{b}_T\cdot\vec{p}_T} \exp\left[\ln\frac{\nu}{\nu_0}\tilde{\gamma}_\nu^{(\Lambda)}(\vec{b}_T, \mu)\right] \\ &= \frac{1}{2\pi}\left(\frac{\Lambda^2}{\mu^2}\right)^{\omega_s}\int_0^\infty \mathrm{d}b_T\, b_T J_0(b_T p_T) \exp\left[-\omega_s\frac{b_T^2\Lambda^2}{4}\,{}_2F_3\left(-\frac{b_T^2\Lambda^2}{4}\right)\right] \\ &= \frac{2}{\pi\mu^2}\left(\frac{\Lambda^2}{\mu^2}\right)^{\omega_s-1}\int_0^\infty \mathrm{d}x\, x J_0\left(\frac{2xp_T}{\Lambda}\right)\exp\left[-\omega_s x^2\,{}_2F_3(-x^2)\right]. \end{aligned} \tag{3.33}$$

This is perfectly finite for $\omega_s = 1$. It shows that by cutting off energetic emissions the divergence disappears, and confirms our observation that it is precisely the large momentum region that is being treated incorrectly. Taking the $\Lambda \to \infty$ limit in eq. (3.33) is slightly nontrivial but reproduces eq. (3.27). In particular, from the expression in the last line of eq. (3.33) it is easy to see that for $\omega_s = 1$ the prefactor diverges for $\Lambda \to \infty$, while it does not for $\omega_s < 1$.

Interestingly, by choosing $\Lambda = p_T$, meaning that no single emission is allowed to be harder than the total $\vec{p}_T$, one obtains

$$V(\vec{p}_T, \mu, \nu, \nu_0) = \frac{2}{\pi\mu^2}\left(\frac{p_T^2}{\mu^2}\right)^{\omega_s-1} \int_0^\infty dx\, x J_0(2x) \exp\left[-\omega_s x^2 {}_2F_3(-x^2)\right], \qquad (3.34)$$

where the p_T-dependence is exactly the same as in the unregulated solution eq. (3.27), while the remaining integral is a function of ω_s that is finite for any ω_s (except for $\omega_s = 0$, for which the whole expression has to reproduce the boundary condition $V(\vec{p}_T, \mu, \nu, \nu) = (2\pi)^2\delta(\vec{p}_T)$). This stresses again that emissions much harder than the actual final state $\vec{p}_T$ cause the divergence.

To see more explicitly how the divergence in eq. (3.27) arises, we can focus on its $\delta(\vec{p}_T)$ term. Using the expression for $\mathcal{L}_0^{\otimes n}$ from eq. (C.24), we get

$$V(\vec{p}_T, \mu, \nu, \nu_0) = \delta(\vec{p}_T) + \sum_{n=1}^\infty \frac{\omega_s^n}{n!}(\mathcal{L}_0^{\otimes n})(\vec{p}_T, \mu) = \delta(\vec{p}_T)\sum_{n=0}^\infty \frac{\omega_s^n}{n!}(-1)^n R_2^{(n)} + \cdots$$

$$\sim \delta(\vec{p}_T)\sum_{n=0}^\infty \frac{\omega_s^n}{n!} \times n! + \cdots, \qquad (3.35)$$

where the ellipses denote the remaining distributions $\mathcal{L}_n$ that built up the $\mathcal{L}^{\omega_s}$ term in eq. (3.27). The $(-1)^n R_2^{(n)}$ is exactly the nth derivative of the ω_s-dependent factor in eq. (3.27), $e^{-2\gamma_E\omega_s}\Gamma(1-\omega_s)/\Gamma(1+\omega_s)$, see eq. (C.17). The last sum in the first line hence precisely leads to the $\delta(\vec{p}_T)$ piece of eq. (3.27), including its divergence at $\omega_s = 1$. In the last step we inserted the asymptotic behaviour $R_2^{(n)} \sim (-1)^n n!$ [see eq. (C.21)]. Each factor ω_s corresponds to a single γ_ν and hence a single real emission. Therefore, the $n!$ growth of the contribution from n emissions, which leads to the divergence for $\omega_s = 1$, simply reflects the number of combinations of n individual emissions $\vec{k}_i$ to yield an overall $\vec{p}_T = \sum_i \vec{k}_i$. Hence it is of a kinematic origin and intrinsic to the two-dimensional nature of $\vec{p}_T$.

To further illustrate this, we can show that this divergence does not appear for the one-dimensional case, where the convolution momentum is strictly limited to $k' \leq k$. The analogous one-dimensional problem (which would be relevant for E_T) is given by

$$\nu\frac{dF(k, \mu, \nu)}{d\nu} = \int dk'\, \gamma_{\nu,F}(k-k', \mu)F(k', \mu, \nu),$$

$$\gamma_{\nu,F}(k, \mu) = 2\Gamma_{\rm cusp}[\alpha_s(\mu)]\mathcal{L}_0(k, \mu). \qquad (3.36)$$

The corresponding evolution kernel is given by,

$$U(k, \mu, \nu, \nu_0) = \delta(k) + \sum_{n=1}^\infty \frac{\omega_s^n}{n!}(\mathcal{L}_0^{\otimes n})(k, \mu) = \left[\delta(k) + \omega_s\mathcal{L}^{\omega_s}(k, \mu)\right]\frac{e^{-\gamma_E\omega_s}}{\Gamma(1+\omega_s)}$$

$$= \delta(k)\sum_{n=0}^\infty \frac{\omega_s^n}{n!}(-1)^n \tilde{R}_1^{(n)} + \cdots = \delta(k)\frac{e^{-\gamma_E\omega_s}}{\Gamma(1+\omega_s)} + \cdots, \qquad (3.37)$$

where the ellipses in the second line denote plus distributions $\mathcal{L}_n(k, \mu)$, which we omitted for simplicity to focus on the $\delta(k)$ term. As before, each power of ω_s corresponds to a

single emission. The n-emission term comes with a coefficient $(-1)^n \tilde{R}_1^{(n)}$, which is the nth derivative of $e^{-\gamma_E \omega_s}/\Gamma(1 + \omega_s)$, see eq. (B.16). This function is well defined for all ω_s, and in particular its derivatives do not show the factorial growth that is present in the two-dimensional case $R_2^{(n)} \sim (-1)^n n!$.

Finally, for completeness we show how the divergence arises from the Fourier space calculation. In this case, the evolution kernel is given by the inverse Fourier transform

$$V(\vec{p}_T, \mu, \nu, \nu_0) = \int \frac{\mathrm{d}^2 \vec{b}_T}{(2\pi)^2} \, e^{i\vec{b}_T \cdot \vec{p}_T} \exp\left[\ln \frac{\nu}{\nu_0} \tilde{\gamma}_\nu(b_T, \mu)\right]. \tag{3.38}$$

The Fourier transform of the LO expression for γ_ν in eq. (3.26) is given by

$$\tilde{\gamma}_\nu(b_T, \mu) = \int \mathrm{d}^2 \vec{p}_T \, e^{-i\vec{b}_T \cdot \vec{p}_T} \gamma_\nu(\vec{p}_T, \mu) = -2\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)] \ln \frac{b_T^2 \mu^2}{b_0^2}, \tag{3.39}$$

such that

$$V(\vec{p}_T, \mu, \nu, \nu_0) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}b_T \, b_T \, J_0(b_T p_T) \exp\left(-\omega_s \ln \frac{b_T^2 \mu^2}{b_0^2}\right), \tag{3.40}$$

with ω_s as in eq. (3.28). The integral converges at large b_T. At small b_T the integrand behaves as

$$b_T \, J_0(b_T p_T) \exp\left(-\omega_s \ln \frac{b_T^2 \mu^2}{b_0^2}\right) \sim b_T^{1-2\omega_s}, \tag{3.41}$$

where we approximated $J_0(b_T p_T) \approx 1$ for $b_T \to 0$. Hence we find a logarithmic divergence from the small b_T limit when $\omega_s \to 1$. This is consistent with the previous calculation, since roughly $b_T \sim p_T^{-1}$, and thus we can again conclude that energetic emission cause the divergent term. To identify the factorial growth, we expand the exponential in the integrand and perform the integration in the small b_T region, where $J_0(b_T p_T) \sim 1$,

$$\begin{aligned}
V(\vec{p}_T, \mu, \nu, \nu_0) &= \frac{1}{2\pi} \int_0^\infty \mathrm{d}b_T \, b_T \, J_0(b_T p_T) \exp\left(-\omega_s \ln \frac{b_T^2 \mu^2}{b_0^2}\right) \\
&\sim \frac{1}{2\pi} \sum_{n=0}^\infty \frac{1}{n!} (-\omega_s)^n \int \mathrm{d}b_T \, b_T \ln^n \frac{b_T^2 \mu^2}{b_0^2} \\
&\sim \frac{b_0^2}{4\pi\mu^2} \sum_{n=0}^\infty \frac{1}{n!} \omega_s^n n! = \frac{b_0^2}{4\pi\mu^2} \frac{1}{1-\omega_s}.
\end{aligned} \tag{3.42}$$

Hence, the small-b_T region of the Fourier integral reproduces the $n!$ growth responsible for the first pole in $\omega_s = 1$ in eq. (3.27). Note that expanding the Bessel function J_0 to higher orders similarly produces the other poles at $\omega_s = 2, 3, \cdots$ of eq. (3.27).

In conclusion, the two-dimensional nature of the convolutions appearing in transverse momentum distributions pose a significant complication to performing the associated renormalization group evolution and resummation. The main problem is that these convolutions can intrinsically probe emissions at arbitrarily large momentum via kinematic cancellations. In principle, such kinematic cancellations among several hard emissions to produce a small

value of $\vec{q}_T$ is a physical effect which will be present in the full all-order result for the spectrum, However, in the above naive approach, the large-momentum emissions get artificially enhanced by fake rapidity logarithms. By explicitly cutting off such emissions, we have shown that they indeed produces unphysical contributions and are the origin of a spurious singularity in the rapidity evolution kernel. As discussed in sec. 3.2, the appearance of these energetic emissions requires a careful scale setting when performing the RG evolution for q_T-spectrum in momentum space, which will be the focus of the remainder of this paper.

4 Resummation of the rapidity anomalous dimension

In this section, we carry out the resummation of the rapidity anomalous dimension $\gamma_\nu(\vec{p}_T, \mu)$ in momentum space by solving its differential equation [see eq. (3.16)],

$$\mu \frac{\mathrm{d}\gamma_\nu(\vec{p}_T, \mu)}{\mathrm{d}\mu} = -4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)]\,\delta(\vec{p}_T)\,, \tag{4.1}$$

using the techniques introduced in sec. 2. Since eq. (4.1) encodes the consistency (i.e. exact path independence) between the μ and ν evolutions [15, 33], the solution of eq. (4.1) is an important ingredient in the full momentum-space resummation. In particular, as was discussed in sec. 3, since the two-dimensional convolutions are intrinsically sensitive to emissions at all momentum scales, one cannot naively use a fixed-order approximation for γ_ν even when the rapidity RGE is performed at $\mu \sim q_T$.

We will also use the resummed result for $\gamma_\nu(\vec{p}_T, \mu)$ as an example to study the differences to carrying out the resummation in Fourier space, the appearance of nonperturbative effects in the q_T spectrum, as well as the implementation of a profile scale and how it allows to probe subleading logarithms.

4.1 Resummation of γ_ν in closed form

We can solve eq. (4.1) distributionally by integrating it from arbitrary μ_0 to μ and then setting $\mu_0 = p_T|_+$ using eq. (2.74),

$$\gamma_\nu(\vec{p}_T, \mu) = \left[\gamma_\nu(\vec{p}_T, \mu_0) - \delta(\vec{p}_T)\int_{\mu_0}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\,4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')]\right]_{\mu_0 = p_T|_+}$$
$$= \frac{1}{2\pi p_T}\frac{\mathrm{d}}{\mathrm{d}p_T}\left\{\theta(p_T)\gamma_\nu[\alpha_s(p_T)] - \theta(p_T)\int_{p_T}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\,4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')]\right\}. \tag{4.2}$$

Here we used that the boundary term can only depend on $\alpha_s(p_T)$ and that by virtue of the cumulant must be proportional to $\theta(p_T)$, to write it as $\theta(p_T)\gamma_\nu[\alpha_s(p_T)]$. Evaluating the derivative [see eq. (C.7)] yields

$$\gamma_\nu(\vec{p}_T, \mu) = \left[\frac{4\Gamma_{\mathrm{cusp}}[\alpha_s(p_T)]}{2\pi p_T^2}\right]_+^{\mu} + \left[\frac{1}{2\pi p_T^2}\frac{\mathrm{d}\gamma_\nu[\alpha_s(p_T)]}{\mathrm{d}\ln p_T}\right]_+^{\xi} + \delta(\vec{p}_T)\gamma_\nu[\alpha_s(\xi)]\,. \tag{4.3}$$

The μ dependence is carried by the first distribution, which can easily be seen to fulfill eq. (4.1). The last two terms arise from the derivative acting on $\theta(p_T)\gamma_\nu[\alpha_s(p_T)]$ and are

hence independent of any scale except p_T. The ξ dependence exactly cancels between the two terms, but is necessary to introduce a plus distribution to regulate the $1/p_T^2$ divergence. The same result is also obtained by directly solving eq. (4.1) in cumulant space.

Let us briefly discuss the form of eq. (4.3). Recall that $\gamma_\nu(\vec{p}_T, \mu)$ corresponds to a single real emission with transverse momentum $\vec{p}_T$. The factor $1/p_T^2$ inside the plus distributions corresponds to the propagator associated with such an emission. The associated infrared singularity at $\vec{p}_T \to 0$ has to cancel against virtual corrections, which is encoded by the plus distribution that regulates the divergence. The effect of solving the RGE of γ_ν is to evaluate the anomalous dimension at $\alpha_s(p_T)$ rather than $\alpha_s(\mu)$. This is not very surprising, as we would expect the μ-RGE to resum virtual corrections to this single emission, which naturally pushes α_s to be evaluated at the emission scale rather than the overall scale μ.

The boundary term $\gamma_\nu[\alpha_s(\xi)]$ can be extracted by integrating eq. (4.3) up to $p_T \leq \mu$,

$$\theta(\mu)\gamma_\nu[\alpha_s(\mu)] = \int_{|\vec{p}_T| \leq \mu} \mathrm{d}^2\vec{p}_T \, \gamma_\nu(\vec{p}_T, \mu) \,. \tag{4.4}$$

It corresponds precisely to the noncusp piece of the anomalous dimensions. We expand it as

$$\gamma_\nu(\alpha_s) = \sum_{n=0}^{\infty} \gamma_{\nu\,n} \left(\frac{\alpha_s}{4\pi}\right)^{n+1} , \tag{4.5}$$

and the constants $\gamma_{\nu\,n}$ are the coefficients of $\delta(\vec{p}_T)$ in the fixed-order calculation of $\gamma_\nu(\vec{p}_T, \mu)$.

4.2 Iterative resummation of γ_ν

It is instructive to see how the resummed result for $\gamma_\nu(\vec{p}_T, \mu)$ arises order-by-order in perturbation theory by using a recurrence relation. To obtain the recurrence relation, we expand

$$\gamma_\nu(\vec{p}_T, \mu) = \sum_{n=0}^{\infty} \gamma_\nu^{(n)}(\vec{p}_T, \mu) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^{n+1} . \tag{4.6}$$

Differentiating this with respect to μ and plugging it back into eq. (4.1) gives the relation

$$\mu \frac{\mathrm{d}\gamma_\nu^{(n)}(\vec{p}_T, \mu)}{\mathrm{d}\mu} = -4\Gamma_n \delta(\vec{p}_T) + \sum_{m=0}^{n-1} 2(m+1)\beta_{n-m-1}\gamma_\nu^{(m)}(\vec{p}_T, \mu) \,, \tag{4.7}$$

where Γ_n and β_n are the coefficients of the cusp anomalous dimensions and beta function, see eq. (A.11). Applying the integration rule in eq. (2.79) yields the solution

$$\gamma_\nu^{(n)}(\vec{p}_T, \mu) = 2\Gamma_n \mathcal{L}_0(\vec{p}_T, \mu) + \sum_{m=0}^{n-1} 2(m+1)\beta_{n-m-1} \int_{\mu_0=p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'}\gamma_\nu^{(m)}(\vec{p}_T, \mu') + \gamma_{\nu\,n}\delta(\vec{p}_T) \,. \tag{4.8}$$

The first two terms follow from integrating the right hand side of eq. (4.7), leaving only a pure number times $\delta(\vec{p}_T)$ as possible boundary term. This shows explicitly that the noncusp coefficients $\gamma_{\nu\,n}$ determine the boundary condition, which can (and must) be determined in

fixed-order perturbation theory, while all plus distributions arise from the μ-evolution. Note that the remaining integral has to be evaluated with distributional scale setting according to eq. (2.79).

The first few terms following from eq. (4.8) are given by

$$
\begin{aligned}
\gamma_\nu^{(0)}(\vec{p}_T, \mu) &= 2\Gamma_0\,\mathcal{L}_0(\vec{p}_T, \mu)\,, \\
\gamma_\nu^{(1)}(\vec{p}_T, \mu) &= -2\beta_0\Gamma_0\,\mathcal{L}_1(\vec{p}_T, \mu) + 2\Gamma_1\,\mathcal{L}_0(\vec{p}_T, \mu) + \gamma_{\nu\,1}\,\delta(\vec{p}_T)\,, \\
\gamma_\nu^{(2)}(\vec{p}_T, \mu) &= 2\beta_0^2\Gamma_0\,\mathcal{L}_2(\vec{p}_T, \mu) - 2(\beta_1\Gamma_0 + 2\beta_0\Gamma_1)\,\mathcal{L}_1(\vec{p}_T, \mu) \\
&\quad + 2(\Gamma_2 - \beta_0\gamma_{\nu\,1})\,\mathcal{L}_0(\vec{p}_T, \mu) + \gamma_{\nu\,2}\,\delta(\vec{p}_T)\,, \\
\gamma_\nu^{(3)}(\vec{p}_T, \mu) &= -2\beta_0^3\Gamma_0\,\mathcal{L}_3(\vec{p}_T, \mu) + (5\beta_0\beta_1\Gamma_0 + 6\beta_0^2\Gamma_1)\,\mathcal{L}_2(\vec{p}_T, \mu) \\
&\quad - 2(\beta_2\Gamma_0 + 2\beta_1\Gamma_1 + 3\beta_0\Gamma_2 - 3\beta_0^2\gamma_{\nu\,1})\,\mathcal{L}_1(\vec{p}_T, \mu) \\
&\quad + (2\Gamma_3 - 2\beta_1\gamma_{\nu\,1} - 3\beta_0\gamma_{\nu\,2})\,\mathcal{L}_0(\vec{p}_T, \mu) + \gamma_{\nu\,3}\,\delta(\vec{p}_T)\,.
\end{aligned}
\tag{4.9}
$$

Here we already used $\gamma_{\nu\,0} = 0$ to simplify the expressions. One can easily check that the same result is obtained by expanding the full resummed result in eq. (4.3).

4.3 Comparison to resummation in Fourier space

The transverse momentum resummation is typically performed in Fourier (impact parameter) space. The rapidity anomalous dimension in Fourier space satisfies the corresponding differential equation

$$
\mu\frac{\mathrm{d}\tilde{\gamma}_\nu(\vec{b}_T, \mu)}{\mathrm{d}\mu} = -4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)]\,.
\tag{4.10}
$$

Here, $\tilde{\gamma}_\nu(\vec{b}_T, \mu)$ is the Fourier transform of $\gamma_\nu(\vec{p}_T, \mu)$. It naturally depends on $\ln(b_T^2\mu^2/b_0^2)$ with $b_0 = 2e^{-\gamma_E}$, such that the resummation with canonical scale choice $\mu_0 = b_0/b_T$ is

$$
\tilde{\gamma}_\nu(\vec{b}_T, \mu) = \tilde{\gamma}_\nu[\alpha_s(b_0/b_T)] - \int_{b_0/b_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'}\,4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')]\,,
\tag{4.11}
$$

where $\tilde{\gamma}_\nu[\alpha_s(b_0/b_T)]$ is the boundary term.

To compare eq. (4.11) to the momentum space solution in eq. (4.3), we need to take the inverse Fourier transform, which involves integrating $\tilde{\gamma}_\nu(\vec{b}_T, \mu)$ over the nonperturbative region $1/b_T \lesssim \Lambda_{\mathrm{QCD}}$. It is hence easier to compare at the level of the perturbative reexpansion of γ_ν. A recurrence relation similar to eq. (4.8) is easily derived in Fourier space,

$$
\tilde{\gamma}_\nu^{(n)}(\vec{b}_T, \mu) = -2\Gamma_n \ln\frac{b_T^2\mu^2}{b_0^2} + 2\sum_{m=0}^{n-1}(m+1)\beta_{n-m-1}\int_{b_0/b_T}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\,\gamma_\nu^{(m)}(\vec{b}_T, \mu') + \tilde{\gamma}_{\nu\,n}\,.
\tag{4.12}
$$

Calculating the first few terms and transforming them to momentum space, we obtain

$$
\begin{aligned}
\gamma_\nu^{(0)}(\vec{p}_T, \mu) &= 2\Gamma_0\mathcal{L}_0(\vec{p}_T, \mu)\,, \\
\gamma_\nu^{(1)}(\vec{p}_T, \mu) &= -2\beta_0\Gamma_0\mathcal{L}_1(\vec{p}_T, \mu) + 2\Gamma_1\mathcal{L}_0(\vec{p}_T, \mu) + \tilde{\gamma}_{\nu\,1}\delta(\vec{p}_T)\,, \\
\gamma_\nu^{(2)}(\vec{p}_T, \mu) &= 2\beta_0^2\Gamma_0\mathcal{L}_2(\vec{p}_T, \mu) - 2(\beta_1\Gamma_0 + 2\beta_0\Gamma_1)\,\mathcal{L}_1(\vec{p}_T, \mu)
\end{aligned}
$$

$$+ 2(\Gamma_2 - \beta_0\tilde{\gamma}_{\nu\,1})\,\mathcal{L}_0(\vec{p}_T,\mu) + \left(\tilde{\gamma}_{\nu\,2} + \frac{8}{3}\zeta_3\beta_0^2\Gamma_0\right)\delta(\vec{p}_T)\,,$$

$$\gamma_\nu^{(3)}(\vec{p}_T,\mu) = -2\beta_0^3\Gamma_0\,\mathcal{L}_3(\vec{p}_T,\mu) + (5\beta_0\beta_1\Gamma_0 + 6\beta_0^2\Gamma_1)\,\mathcal{L}_2(\vec{p}_T,\mu)$$
$$- 2(\beta_2\Gamma_0 + 2\beta_1\Gamma_1 + 3\beta_0\Gamma_2 - 3\beta_0^2\tilde{\gamma}_{\nu\,1})\,\mathcal{L}_1(\vec{p}_T,\mu)$$
$$+ (2\Gamma_3 - 2\beta_1\tilde{\gamma}_{\nu\,1} - 3\beta_0\tilde{\gamma}_{\nu\,2} - 8\zeta_3\beta_0^3\Gamma_0)\,\mathcal{L}_0(\vec{p}_T,\mu)$$
$$+ \left(\tilde{\gamma}_{\nu\,3} + \frac{20}{3}\zeta_3\beta_0\beta_1\Gamma_0 + 8\zeta_3\beta_0^2\Gamma_1\right)\delta(\vec{p}_T)\,. \tag{4.13}$$

The first two terms agree with eq. (4.9), and hence to this order we find identical noncusp constants, $\tilde{\gamma}_{\nu\,1} = \gamma_{\nu\,1}$. The differences compared to eq. (4.9) start at $\mathcal{O}(\alpha_s^3)$, where the b-space resummation induces additional terms. This means that the 3-loop and 4-loop noncusp constants are related by

$$\gamma_{\nu\,2} = \tilde{\gamma}_{\nu\,2} + \frac{8}{3}\zeta_3\beta_0^2\Gamma_0\,,$$
$$\gamma_{\nu\,3} = \tilde{\gamma}_{\nu\,3} + \frac{20}{3}\zeta_3\beta_0\beta_1\Gamma_0 + 8\zeta_3\beta_0^2\Gamma_1\,. \tag{4.14}$$

In addition, the $\mathcal{L}_0$ term in $\gamma_\nu^{(3)}$ differs by a contribution $8\zeta_3\beta_0^3\Gamma_0$, which is induced by the different boundary term $\tilde{\gamma}_{\nu\,2}$, which feeds into the logarithmic distribution at higher orders.

Similar additional terms are induced at each higher order. The reason is that pure b-space logarithms $\ln^n(b^2\mu^2/b_0^2)$ do not correspond to pure plus distributions $\mathcal{L}_{n-1}(\vec{p}_T,\mu)$, but also induce $\delta(\vec{p}_T)$ terms when Fourier transformed back to momentum space. The implications of this were discussed already in secs. 2.5 and 2.7. If γ_ν is calculated in full fixed order to α_s^n, the boundary terms can be extracted up to this order, which also takes into account the differences between $\gamma_{\nu\,n}$ and $\tilde{\gamma}_{\nu\,n}$ up to this order. However, for the unknown higher order terms one would set $\tilde{\gamma}_{\nu\,n} = 0$ in the b-space resummation (i.e. by setting the boundary condition in b-space) and $\gamma_{\nu\,n} = 0$ in the momentum-space resummation (i.e. by setting the boundary condition in momentum space). This means if one were to use the b-space resummation to obtain the momentum-space result this would induce additional boundary terms which then lead to additional subleading terms to all orders compared to the momentum-space resummation, and vice versa.

Even though the inverse Fourier transformation over the full b-space result is not possible, we can express the differences in closed form by relating eqs. (4.3) and (4.11) through

$$\gamma_\nu(\vec{p}_T,\mu) = \mathrm{FT}^{-1}[\tilde{\gamma}_\nu(\vec{b}_T,\mu)](\vec{p}_T) - \left[\frac{1}{2\pi p_T^2}\frac{\mathrm{d}\Delta\gamma_\nu[\alpha_s(p_T)]}{\mathrm{d}\ln p_T}\right]_+^\mu - \delta(\vec{p}_T)\,\Delta\gamma_\nu[\alpha_s(\mu)]\,, \tag{4.15}$$

where

$$\Delta\gamma_\nu[\alpha_s(\mu)] = \int_0^\infty \mathrm{d}b_T\,\mu\,J_1(b_T\mu)\,\tilde{\gamma}_\nu(|\vec{b}_T|,\mu) - \gamma_\nu[\alpha_s(\mu)]\,. \tag{4.16}$$

Here we used that $\tilde{\gamma}_\nu$ only depends on the magnitude of $\vec{b}_T$, $\tilde{\gamma}_\nu(\vec{b}_T,\mu) \equiv \tilde{\gamma}_\nu(|\vec{b}_T|,\mu)$. The relation is derived by relating the pure $\delta(\vec{p}_T)$ pieces via eq. (4.4) to each other. It holds in the sense that the fixed-order expansions of both sides match to all orders. The $\Delta\gamma_\nu$ contains

the differences in the boundary condition and the second term in eq. (4.15) explicitly shows the subleading logarithmic terms contained in the inverse Fourier transform of the b-space solution.

4.4 Turning off resummation using profiles

Having obtained the resummed rapidity anomalous dimension, we can use it as an illustrative example for the implementation of a profile scale to smoothly transition between resummation and fixed-order regimes, as discussed in sec. 2.6. Starting from the formal solution

$$\gamma_\nu(\vec{p}_T, \mu) = \gamma_\nu(\vec{p}_T, \mu_0) - \delta(\vec{p}_T) \int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \,, \tag{4.17}$$

the aim is to choose a profile scale $\mu_0(p_T)$ that smoothly transitions from the canonical scale choice $\mu_0 = p_T|_+$ in the resummation regime to the fixed scale $\mu_0 = \mu$ that turns off the resummation. Note that the profile function can also depend on further variables such as Q or the final $\vec{q}_T$. Since these dependences do not require distributional scale setting, we suppress them.

As discussed in sec. 2.6, the p_T-dependence of $\mu_0(p_T)$ requires distributional scale setting using eq. (2.74) generalized to a function $\mu_0(p_T)$,

$$\gamma_\nu(\vec{p}_T, \mu) = \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \int_{|\vec{k}_T| \leq p_T} \mathrm{d}^2\vec{k}_T \left\{ \gamma_\nu^{\mathrm{FO}}(\vec{k}_T, \mu_0(p_T)) - \delta(\vec{k}_T) \int_{\mu_0(p_T)}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \right\} \,. \tag{4.18}$$

Here the superscript "FO" makes explicit that the boundary term is obtained from a fixed-order calculation. Focusing for the moment on the second term, which predicts the logarithmic terms at higher orders in α_s, the derivative can be evaluated [see eq. (C.7)],

$$\gamma_\nu(\vec{p}_T, \mu) \supset \left[\frac{1}{2\pi p_T^2} \frac{\mathrm{d}\ln\mu_0(p_T)}{\mathrm{d}\ln p_T} 4\Gamma_{\mathrm{cusp}}[\mu_0(p_T)] \right]_+^{\mu} - \delta(\vec{p}_T) \int_{\mu_0(\mu)}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \,. \tag{4.19}$$

This has to be compared to the corresponding term in the canonical solution in eq. (4.3),

$$\gamma_\nu(\vec{p}_T, \mu) \supset \left[\frac{4\Gamma_{\mathrm{cusp}}[\alpha_s(p_T)]}{2\pi p_T^2} \right]_+^{\mu} \,. \tag{4.20}$$

This result is recovered using the canonical scale $\mu_0(p_T) = p_T$, for which the integral in eq. (4.19) vanishes. On the other hand, the resummation is turned off by choosing $\mu_0(p_T) = \mu$, for which both terms in eq. (4.19) vanish and only the fixed-order term in eq. (4.18) survives. In between, the shape of the derivative $\frac{\mathrm{d}\ln\mu_0(p_T)}{\mathrm{d}\ln p_T} \neq 1$ entering in the plus distribution controls how the plus distribution, which is entirely generated by the resummation, is being turned off until it vanishes in the fixed-order regime.

Another purpose of the profile scale is to probe the size of higher-order terms through scale variations. To show concretely how this works, consider choosing $\mu_0(p_T) = 2p_T$, for which eq. (4.19) becomes

$$\gamma_\nu(\vec{p}_T, \mu) \supset \left[\frac{4\Gamma_{\mathrm{cusp}}[\alpha_s(2p_T)]}{2\pi p_T^2} \right]_+^{\mu} - \delta(\vec{p}_T) \int_{2\mu}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \,. \tag{4.21}$$

We observe two effects. On the one hand, the plus distribution now has a different shape in $\vec{p}_T$ due to $\Gamma_{\mathrm{cusp}}[\alpha_s(2p_T)]$. This is reasonable, as all higher-order logarithms in γ_ν are resummed into the scale entering the cusp anomalous dimension Γ_{cusp}, so we expect a variation there. On the other hand, the $\delta(\vec{p}_T)$ term does not vanish anymore, and hence probes the constant fixed-order boundary terms.

In conclusion, the distributional scale setting allows one to use profile scales as usual, allowing to smoothly connect the resummation and fixed-order regimes as well as to implement profile scale variations.

Example at lowest order As an illustrative example, we consider resumming γ_ν at lowest order (i.e. LL including β_0 and Γ_0) with a profile scale in conjunction with the full $\mathcal{O}(\alpha_s)$ boundary term,

$$\gamma_\nu^{\mathrm{FO}}(\vec{p}_T, \mu) = 2\Gamma_0 \frac{\alpha_s(\mu)}{4\pi} \mathcal{L}_0(\vec{p}_T, \mu) \,. \tag{4.22}$$

(This combination corresponds to a partial NLL result and would typically not arise in practice, but it is useful for illustration.) From eq. (4.18), we obtain

$$\gamma_\nu^{\mathrm{(N)LL}}(\vec{p}_T, \mu) = \frac{\Gamma_0}{\pi} \left[\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \left\{ \alpha_s[\mu_0(p_T)] \ln \frac{p_T}{\mu_0(p_T)} - \int_{\mu_0(p_T)}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \alpha_s(\mu') \right\} \right]_+^{\mu}$$
$$+ \frac{\Gamma_0}{\pi} \delta(\vec{p}_T) \left\{ \alpha_s[\mu_0(\mu)] \ln \frac{\mu}{\mu_0(\mu)} - \int_{\mu_0(\mu)}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \alpha_s(\mu') \right\} \,. \tag{4.23}$$

The first term in each of the curly brackets is from the fixed-order boundary condition and the second term is from the LL μ evolution. Expanding the result in $\alpha_s(\mu)$, we obtain

$$\gamma_\nu^{\mathrm{(N)LL(0)}}(\vec{p}_T, \mu) = 2\Gamma_0 \, \mathcal{L}_0(\vec{p}_T, \mu) = \gamma_\nu^{(0)}(\vec{p}_T, \mu) \,,$$
$$\gamma_\nu^{\mathrm{(N)LL(1)}}(\vec{p}_T, \mu) = -2\beta_0 \Gamma_0 \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \ln \frac{\mu_0(p_T)^2}{\mu^2} \right]_+^{\mu} + 2\beta_0 \Gamma_0 \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \ln \frac{\mu_0(p_T)^2}{p_T^2} \frac{\mathrm{d}\ln \mu_0(p_T)}{\mathrm{d}\ln p_T} \right]_+^{\mu}$$
$$+ 4\Gamma_0 \beta_0 \ln^2 \frac{\mu}{\mu_0(\mu)} \delta(\vec{p}_T) \,. \tag{4.24}$$

As discussed in generality above, by including the full $\mathcal{O}(\alpha_s)$ boundary condition, the $\mathcal{O}(\alpha_s)$ result $\gamma_\nu^{(0)}$ is exactly reproduced independently of the choice of $\mu_0(p_T)$. The two-loop term has to be compared to the full fixed-order result from eq. (4.9),

$$\gamma_\nu^{(1)}(\vec{p}_T, \mu) = -2\beta_0 \Gamma_0 \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \ln \frac{p_T^2}{\mu^2} \right]_+^{\mu} + 2\Gamma_1 \frac{1}{\pi\mu^2} \left[\frac{\mu^2}{\vec{p}_T^2} \right]_+^{\mu} + \gamma_{\nu\,1} \delta(\vec{p}_T) \,. \tag{4.25}$$

With the fixed-order scale $\mu_0(p_T) = \mu$, the two-loop term $\gamma_\nu^{\mathrm{(N)LL(0)}}$ completely vanishes, leaving only the fixed-order term $\gamma_\nu^{(0)}$. With the canonical resummation scale $\mu_0(p_T) = p_T$, the first term in $\gamma_\nu^{\mathrm{(N)LL(0)}}$ exactly reproduces the leading-logarithmic two-loop term $\sim \beta_0 \Gamma_0 \mathcal{L}_1$, while the other terms vanish. Varying $\mu_0(p_T)$ induces higher-order terms with precisely the structure of the formally higher-order Γ_1 and $\gamma_{\nu\,1}$ terms. Furthermore, the p_T-dependence of the variation allows one to distinguish and separately probe the size of the higher-order logarithmic and higher fixed-order boundary pieces.

4.5 Nonperturbative modeling with the moment expansion

It is well known from the resummation in Fourier space that the rapidity evolution kernel becomes intrinsically nonperturbative at $1/b \ll \Lambda_{\rm QCD}$ [38–40]. In momentum space this corresponds to the fact that the resummed result for $\gamma_\nu(\vec{p}_T, \mu)$ in eq. (4.3) explicitly depends on $\alpha_s(p_T)$, which means that it becomes nonperturbative for $p_T \lesssim \Lambda_{\rm QCD}$. This nonperturbative region is necessarily probed in the rapidity RGE of the soft function, see e.g. eqs. (3.14) and (3.20), since $\gamma_\nu(\vec{k}_T, \mu)$ is integrated over all $\vec{k}_T$.

Although we will show in sec. 5.3 that the final result for the momentum convolutions is actually perturbative up to power corrections $\mathcal{O}(\Lambda_{\rm QCD}^2/p_T^2)$, it is still important to properly handle and isolate the nonperturbative contributions to γ_ν. To do so, we can write the true all-order result for γ_ν as

$$\gamma_\nu(\vec{p}_T, \mu) = \left[\frac{4\Gamma_{\rm cusp}[\alpha_s(\mu_0(p_T))]}{2\pi p_T^2}\right]_+^\mu + \left[\frac{1}{2\pi p_T^2}\frac{\mathrm{d}\gamma_\nu[\alpha_s(\mu_0(p_T))]}{\mathrm{d}\ln p_T}\right]_+^\xi + \delta(\vec{p}_T)\,\gamma_\nu[\alpha_s(\mu_0(\xi))]$$
$$+ \gamma_\nu^{(\rm np)}(\vec{p}_T)\,. \tag{4.26}$$

Here, we have replaced the fixed canonical scale $\mu_0 = p_T|_+$ in the resummed result in eq. (4.3) by a cutoff function $\mu_0(p_T)$, which is equal to $\mu_0(p_T) = p_T$ for $p_T \gtrsim \Lambda_{\rm QCD}$ but cuts off and approaches some constant (perturbative) value for $p_T \lesssim \Lambda_{\rm QCD}$. The difference to the true result is absorbed into the nonperturbative contribution $\gamma_\nu^{(\rm np)}(\vec{p}_T)$. Since the resummed perturbative result is now only ever evaluated at perturbative scales, all nonperturbative contributions must be contained in $\gamma_\nu^{(\rm np)}(\vec{p}_T)$. The precise choice of $\mu_0(p_T)$ here corresponds to a perturbative scheme dependence in the definition of $\gamma_\nu^{(\rm np)}(\vec{p}_T)$, which cancels on the right-hand sice in eq. (4.26) (and which we suppress in our notation). At the same time, $\gamma_\nu^{(\rm np)}(\vec{p}_T)$ is μ independent because the (perturbative) μ dependence is by definition carried by the resummed perturbative contributions.

Letting $\mu_0(p_T) = p_T$ in the perturbative regime, $\gamma_\nu^{(\rm np)}(\vec{p}_T)$ will have support of order $\Lambda_{\rm QCD}$. Hence we can expand it in its moments,

$$\gamma_\nu^{(\rm np)}(\vec{p}_T) = \left(\sum_{n=1}^{\infty} \Omega_n \Delta_{\vec{p}_T}^n\right)\delta(\vec{p}_T) \equiv \gamma_\nu^{(\rm np)}[\Delta_{\vec{p}_T}]\,\delta(\vec{p}_T)\,, \tag{4.27}$$

where $\Delta_{\vec{p}_T} = \partial^2/\partial^2 p_x + \partial^2/\partial^2 p_y$ is the Laplace operator. Since γ_ν is known to be azimuthaly symmetric, no other operators, such as e.g. linear operators, may arise. The associated Ω_n coefficients are nonperturbative parameters, which have to be extracted from data. Like $\gamma_\nu^{(\rm np)}(\vec{p}_T)$ itself, they are μ independent but scheme dependent. This is reminiscent of the moment expansions to treat the nonperturbative corrections in one-dimensional soft functions [27, 76].

The moment expansion eq. (4.27) becomes more intuitive upon Fourier transformation. In impact parameter space, it turns into a simple polynomial in $-b_T^2$,

$$\tilde{\gamma}_\nu^{(\rm np)}(\vec{b}_T) = \sum_{n=1}^{\infty} \Omega_n(-b_T^2)^n = \gamma_\nu^{(\rm np)}[-b_T^2]\,. \tag{4.28}$$

The analysis of ref. [77] found that the perturbative result for $\tilde{\gamma}_\nu(\vec{b}_T, \mu)$ has a leading renormalon contribution which scales as b_T^2. This is consistent with the fact that the first nonperturbative moment Ω_1 scales with $\Delta_{\vec{p}_T}$ or b_T^2, and the renormalon in the perturbative series should be precisely cancelled by a corresponding renormalon in Ω_1. We expect that the above cutoff definition can be used to provide a renormalon-free scheme definition for Ω_1 in momentum space and presumably $\gamma_\nu^{(\mathrm{np})}(\vec{p}_T)$ as a whole. It would be interesting to investigate this in more detail in the future.

To compare the above moment expansion to the literature, consider the formal solution of the rapidity RGE in eq. (3.14) in impact parameter space,

$$\tilde{S}(\vec{b}_T, \mu, \nu) = \tilde{S}(\vec{b}_T, \mu, \nu_0) \exp\left[\ln\frac{\nu}{\nu_0}\tilde{\gamma}_\nu(b_T, \mu)\right]. \tag{4.29}$$

Including both the perturbative and nonperturbative contributions to $\tilde{\gamma}_\nu$, this yields

$$\tilde{S}(\vec{b}_T, \mu, \nu) = \tilde{S}(\vec{b}_T, \mu, \nu_0) \exp\left[\ln\frac{\nu}{\nu_0}\tilde{\gamma}_\nu^{(\mathrm{pert})}(b_T, \mu)\right] \exp\left[-\Omega_1 \ln\frac{\nu}{\nu_0}b_T^2 + \cdots\right], \tag{4.30}$$

where $\tilde{\gamma}_\nu^{(\mathrm{pert})}(b_T, \mu)$ is the Fourier transform of the resummed perturbative contribution. Hence, the momentum-space resummation reproduces and confirms the often-used procedure in the literature to model nonperturbative effects using a Gaussian factor in impact parameter space, which is also motivated by renormalon analyses [77, 78]. In practice, a variety of nonperturbative models have been suggested, see e.g. refs. [79, 80] for recent studies. The above also confirms that the nonperturbative correction scales with a rapidity logarithm [39, 81].

5 Resummation of soft and beam functions

In the previous section we have solved the RGE for the rapidity anomalous dimension γ_ν, which already illustrated some of the key features of the distributional scale setting in a nontrivial setting. In this section, we will solve the RGEs for both beam and soft function and derive the correct momentum-space evolution to predict their complete distributional logarithmic structure. The implications on the full q_T spectrum are discussed in sec. 6.

5.1 Soft function

The soft function obeys the coupled system of renormalization group equations, eqs. (3.8) and (3.14),

$$\mu\frac{\mathrm{d}S(\vec{p}_T, \mu, \nu)}{\mathrm{d}\mu} = \gamma_S(\mu, \nu)S(\vec{p}_T, \mu, \nu), \tag{5.1}$$

$$\nu\frac{\mathrm{d}S(\vec{p}_T, \mu, \nu)}{\mathrm{d}\nu} = \int \mathrm{d}^2\vec{k}_T\, \gamma_\nu(\vec{k}_T, \mu)\, S(\vec{p}_T - \vec{k}_T, \mu, \nu), \tag{5.2}$$

where the ν-RGE involves a two-dimensional convolution, which is the main source of complications, as discussed in sec. 3.2.

– 37 –

5.1.1 Iterative solution

We first derive an iterative solution order by order in α_s. This requires the expansions

$$S(\vec{p}_T, \mu, \nu) = \sum_{n=0}^{\infty} S^{(n)}(\vec{p}_T, \mu, \nu) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^n, \tag{5.3}$$

$$\gamma_S(\mu, \nu) = \sum_{n=0}^{\infty} \left(4\Gamma_n \ln\frac{\mu}{\nu} + \gamma_{Sn}\right) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^{n+1}, \tag{5.4}$$

$$\gamma_\nu(\vec{p}_T, \mu) = \sum_{n=0}^{\infty} \gamma_\nu^{(n)}(\vec{p}_T, \mu) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^{n+1}, \tag{5.5}$$

where the $\gamma_\nu^{(n)}$ were derived in sec. 4.2. The two RGEs then become

$$\mu\frac{\mathrm{d}S^{(n)}(\vec{p}_T, \mu, \nu)}{\mathrm{d}\mu} = \sum_{m=0}^{n-1} \left(4\Gamma_{n-m-1}\ln\frac{\mu}{\nu} + \gamma_{Sn-m-1} + 2m\beta_{n-m-1}\right) S^{(m)}(\vec{p}_T, \mu, \nu), \tag{5.6}$$

$$\nu\frac{\mathrm{d}S^{(n)}(\vec{p}_T, \mu, \nu)}{\mathrm{d}\nu} = \sum_{m=0}^{n-1} \left(\gamma_\nu^{(n-m-1)} \otimes S^{(m)}\right)(\vec{p}_T, \mu, \nu). \tag{5.7}$$

These equations allow to determine the α_s^n coefficient $S^{(n)}$ from the lower order terms $S^{(m)}$ with $m < n$. Solving first the ν-RGE in eq. (5.7), we obtain the formal solution

$$S^{(n)}(\vec{p}_T, \mu, \nu) = S^{(n)}(\vec{p}_T, \mu, \nu_0) + \int_{\nu_0}^{\nu} \frac{\mathrm{d}\nu'}{\nu'} \sum_{m=0}^{n-1} \gamma_\nu^{(n-m-1)}(\vec{p}_T, \mu) \otimes S^{(m)}(\vec{p}_T, \mu, \nu'). \tag{5.8}$$

The missing boundary term $S^{(n)}(\vec{p}_T, \mu, \nu_0)$ is deduced from the μ-RGE in eq. (5.6),

$$S^{(n)}(\vec{p}_T, \mu, \nu_0) = S^{(n)}(\vec{p}_T, \mu_0, \nu_0) \tag{5.9}$$

$$+ \int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \sum_{m=0}^{n-1} \left(4\Gamma_{n-m-1}\ln\frac{\mu'}{\nu_0} + \gamma_{Sn-m-1} + 2m\beta_{n-m-1}\right) S^{(m)}(\vec{p}_T, \mu', \nu_0).$$

The remaining boundary term $S^{(n)}(\vec{p}_T, \mu_0, \nu_0)$ can only contain logarithms $\ln(\mu_0/p_T)$ or $\ln(\nu_0/p_T)$ or any combination of these. In particular, these logarithms can be hidden inside the boundary condition of plus distributions $\mathcal{L}_n(\vec{p}_T, \mu), \mathcal{L}_n(\vec{p}_T, \nu)$. To eliminate them, we apply eq. (2.74) to set $\mu_0 = \nu_0 = p_T|_+$. The remaining pure fixed-order boundary term must then be proportional to $\delta(\vec{p}_T)$, and we obtain

$$S^{(n)}(\vec{p}_T, \mu, \nu) = S_n\delta(\vec{p}_T)$$

$$+ \int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \sum_{m=0}^{n-1} \left(4\Gamma_{n-m-1}\ln\frac{\mu'}{\nu_0} + \gamma_{Sn-m-1} + 2m\beta_{n-m-1}\right) S^{(m)}(\vec{p}_T, \mu', \nu_0)\Big|_{\nu_0=p_T|_+}$$

$$+ \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu'}{\nu'} \sum_{m=0}^{n-1} \left(\gamma_\nu^{(n-m-1)} \otimes S^{(m)}\right)(\vec{p}_T, \mu, \nu'). \tag{5.10}$$

In the μ-integral, both distributional scale settings should be performed in one step, $\mu_0 = \nu_0 = p_T|_+$. By construction, this solution fulfills both RGEs and fully resums all logarithmic

distributions. We have explicitly verified this solution by predicting the structure of the soft function through $\mathcal{O}(\alpha_s^6)$ and comparing with the results obtained from impact parameter space, where the distributions become simple functions and hence the complication of distributional scale setting does not arise. As for γ_ν, the constant pieces differ in general, i.e. $S_n \neq \tilde{S}_n$, thereby also inducing different logarithms at higher orders. Upon making the appropriate identifications between $\tilde{S}_n$ and S_n, exact agreement between both iterative solutions is obtained. This shows that the distributional scale setting is well-defined for the soft function, even though it involves two distinct scales. For illustration, the fixed-order expansion of the soft function is given through $\mathcal{O}(\alpha_s^3)$ in appendix D, including the relation between the boundary terms S_n and $\tilde{S}_n$.

For completeness, we also give the iterative solution when first solving the μ-RGE and then the ν-RGE:

$$
\begin{aligned}
S^{(n)}(\vec{p}_T, \mu, \nu) = {} & S_n \delta(\vec{p}_T) \\
& + \int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \sum_{m=0}^{n-1} \Big(4\Gamma_{n-m-1} \ln \frac{\mu'}{\nu} + \gamma_{S\,n-m-1} + 2m\beta_{n-m-1} \Big) S^{(m)}(\vec{p}_T, \mu', \nu) \\
& + \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu'}{\nu'} \sum_{m=0}^{n-1} \big(\gamma_\nu^{(n-m-1)} \otimes S^{(m)} \big)(\vec{p}_T, p_T|_+, \nu') \,.
\end{aligned}
\tag{5.11}
$$

Equations (5.10) and (5.11) are useful to illustrate again the correct scale setting for distributions. Firstly, the lower order terms $S^{(m)}$ feeding into $S^{(n)}$ are themselves distributions with boundaries μ or ν. Hence integrating over μ and ν with starting scale p_T requires distributional scale setting, which we apply according to eq. (2.74). Secondly, we see that it is crucial to set $\nu_0 = p_T|_+$ at each order rather than keeping ν_0 arbitrary and setting $\nu_0 = p_T$ only at the very end. The reason lies in the convolution term in eqs. (5.10) and (5.11), $(\gamma_\nu^{(n-m-1)} \otimes S^{(m)})(\vec{p}_T, \mu, \nu')$: The convolution depends on whether the integrand contains a formal scale ν_0 or whether it is set to the convolution variable $\vec{k}_T$. In contrast, multiplicative RGEs are not sensitive to this, such that the scale can be set at the very end.

5.1.2 Solution in closed form

We now proceed to derive a closed form of the solution. We start by solving the μ-RGE in eqs. (3.8) and (3.11) at $\nu_0 = p_T|_+$. The solution is straightforwardly obtained using the technique of sec. 2.3. The formal solution

$$
S(\vec{p}_T, \mu, \nu_0) = S(\vec{p}_T, \mu_0, \nu_0) \exp\left[\int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_S(\mu', \nu_0) \right]
\tag{5.12}
$$

only requires to set $\mu_0 = \nu_0 = p_T|_+$, yielding[5]

$$
\begin{aligned}
S(\vec{p}_T, \mu, \nu_0 = p_T|_+) &= \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \theta(p_T) \mathcal{S}[\alpha_s(p_T)] \exp\left[\int_{p_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_S(\mu', p_T) \right] \tag{5.13} \\
&= \delta(\vec{p}_T) \mathcal{S}[\alpha_s(\mu)] + \left[\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \mathcal{S}[\alpha_s(p_T)] \exp\left\{ \int_{p_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_S(\mu', p_T) \right\} \right]_+^{\mu}.
\end{aligned}
$$

---

[5]To allow scale variations, one could of course choose profile functions $\mu_0 = \mu_0(p_T)|_+, \nu_0 = \nu_0(p_T)|_+$.

The plus distribution contains the exponentiated Sudakov double logarithm at $\nu_0 = p_T|_+$ as dictated by the RGE. The pure boundary term, which can only depend on α_s, is defined as

$$\theta(p_T)\mathcal{S}[\alpha_s(p_T)] = \int_{|\vec{k}_T| \leq p_T} \mathrm{d}^2\vec{k}_T \, S(\vec{k}_T, \mu = p_T, \nu = p_T) \,. \tag{5.14}$$

and has the perturbative expansion

$$\mathcal{S}(\alpha_s) = \sum_{n=0}^{\infty} S_n \left(\frac{\alpha_s}{4\pi}\right)^n \,. \tag{5.15}$$

The constants S_n can be obtained as coefficients of $\delta(\vec{p}_T)$ in the fixed-order calculation. More generally, the soft function at the canonical distributional scales is given in terms of this boundary term by

$$S(\vec{p}_T, p_T|_+, p_T|_+) = \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \theta(p_T)\mathcal{S}[\alpha_s(p_T)]$$

$$= \delta(\vec{p}_T)\mathcal{S}[\alpha_s(\xi)] + \left[\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T}\mathcal{S}[\alpha_s(p_T)]\right]_+^{\xi} \,, \tag{5.16}$$

where the ξ dependence exactly cancels.

In the next step, the ν-RGE in eq. (3.14) has to be solved, which is more complicated. Inspired by the iterative solution, we expand the soft function as

$$S(\vec{p}_T, \mu, \nu) = \sum_{n=0}^{\infty} S^{[n]}(\vec{p}_T, \mu, \nu) \,, \tag{5.17}$$

where the $S^{[n]}$ correspond to the soft function including n contributions from γ_ν. More intuitively, $S^{[n]}$ is the piece of the soft function originating from n real emissions. Since γ_ν describes a single real emission, the rapidity RGE becomes

$$\nu \frac{\mathrm{d}S(\vec{p}_T, \mu, \nu)}{\mathrm{d}\nu} = \sum_{n=0}^{\infty} \nu \frac{\mathrm{d}}{\mathrm{d}\nu} S^{[n]}(\vec{p}_T, \mu, \nu) = \sum_{n=0}^{\infty} (\gamma_\nu \otimes S^{[n]})(\vec{p}_T, \mu, \nu)$$

$$= \sum_{n=1}^{\infty} (\gamma_\nu \otimes S^{[n-1]})(\vec{p}_T, \mu, \nu) \,. \tag{5.18}$$

Hence the $S^{[n]}$ are determined through the recursive RGE

$$\nu \frac{\mathrm{d}}{\mathrm{d}\nu} S^{[n]}(\vec{p}_T, \mu, \nu) = (\gamma_\nu \otimes S^{[n-1]})(\vec{p}_T, \mu, \nu) \,, \quad \nu \frac{\mathrm{d}}{\mathrm{d}\nu} S^{[0]}(\vec{p}_T, \mu, \nu) = 0 \,. \tag{5.19}$$

The solution at order n is

$$S^{[n]}(\vec{p}_T, \mu, \nu) = S^{[n]}(\vec{p}_T, \mu, p_T|_+) + \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu'}{\nu'} (\gamma_\nu \otimes S^{[n-1]})(\vec{p}_T, \mu, \nu') \,, \tag{5.20}$$

where it is now quite intuitive that ν_0 should be set to $p_T|_+$ and not to the overall q_T. Iterating this, the only leftover boundary term is precisely $S(\vec{p}_T, \mu, p_T|_+) = \sum_n S^{[n]}(\vec{p}_T, \mu, p_T|_+)$,

and we can write the formal all-order solution as

$$S(\vec{p}_T, \mu, \nu) = S(\vec{p}_T, \mu, p_T|_+) + \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu_1}{\nu_1} \int \mathrm{d}^2\vec{k}_1 \, \gamma_\nu(\vec{p}_T - \vec{k}_1, \mu) S(\vec{k}_1, \mu, k_1|_+)$$

$$+ \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu_1}{\nu_1} \int \mathrm{d}^2\vec{k}_1 \, \gamma_\nu(\vec{p}_T - \vec{k}_1, \mu) \int_{k_1|_+}^{\nu_1} \frac{\mathrm{d}\nu_2}{\nu_2} \int \mathrm{d}^2\vec{k}_2 \, \gamma_\nu(\vec{k}_1 - \vec{k}_2, \mu) S(\vec{k}_2, \mu, k_2|_+)$$

$$+ \cdots$$

$$= S(\vec{p}_T, \mu, p_T|_+) + \sum_{n=1}^{\infty} \left[\prod_{i=1}^{n} \int_{k_{i-1}|_+}^{\nu_{i-1}} \frac{\mathrm{d}\nu_i}{\nu_i} \int \mathrm{d}^2\vec{k}_i \, \gamma_\nu(\vec{k}_{i-1} - \vec{k}_i, \mu) \right] S(\vec{k}_n, \mu, k_n|_+) \,,$$

$$(5.21)$$

where in the last line $k_0 \equiv p_T$ and $\nu_0 \equiv \nu$. All ν_i integrals have to be understood according to eq. (2.79).

From the explicit form of the first few terms, it is easy to see that eq. (5.21) fulfills the rapity RGE and that by choosing the overall $\nu = p_T|_+$ reproduces the (μ-evolved) boundary condition $S(\vec{p}_T, \mu, p_T|_+)$. Note that although both γ_ν and S will be probed at nonperturbative α_s in the convolutions of eq. (5.21), the result is actually perturbative up to corrections $\mathcal{O}(\Lambda_{\mathrm{QCD}}^2/p_T^2)$, as will be discussed in sec. 5.3. For practical purposes it is nevertheless necessary to use a nonperturbative modelling, as already discussed for γ_ν in sec. 4.5. To allow scale variations, one would replace all $k_i|_+$ by a profile functions $\nu_0(k_i)|_+$.

Together, eqs. (5.13) and (5.21) completely predict the logarithmic structure of the soft function to all orders. Crucially, the canonical-scale boundary condition for the solution is $S(\vec{p}_T, p_T|_+, p_T|_+)$, which can be reliably calculated in a pure fixed-order calculation. The iterative structure of the rapidity evolution eq. (5.21) ensures that the correct rapidity logarithms $\ln(\nu/|\vec{k}_i|)$ evolving each emission to the overall rapidity scale are resummed.

Once the soft function has been evolved to some scale ν_S using this method, it can be evolved further to a different overall scale ν using the simple kernel of eq. (3.20),

$$S(\vec{p}_T, \mu, \nu) = \int \mathrm{d}^2\vec{k}_T \, V_S(\vec{p}_T - \vec{k}_T, \mu, \nu, \nu_S) \, S(\vec{k}_T, \mu, \nu_S) \,,$$

$$V_S(\vec{p}_T, \mu, \nu, \nu_S) = \delta(\vec{p}_T) + \sum_{n=1}^{\infty} \frac{1}{n!} \ln^n \frac{\nu}{\nu_S} \, (\gamma_\nu \otimes^n)(\vec{p}_T, \mu) \,, \qquad (5.22)$$

where we introduce the notation V_S for the evolution kernel when used in this way for continuing the evolution. This simple structure immediately emerges from eq. (5.21) when replacing all $\nu_0 = k_i|_+$ by a common scale ν_S, such that all ν-integrals factor out of the convolutions. This shows that eq. (5.22) is indeed sufficient to continue the ν evolution in eq. (5.21) from the overall scale ν_S to a new overall scale ν.

5.1.3 Comparison to "naive" scale setting

Having obtained a solution correctly minimizing the boundary term distributionally, we can turn back to the discussion of sec. 3.2. There we had argued that the naive solution, eqs. (3.19) and (3.20), leads to wrong predictions, because energetic emissions are artificially enhanced by unphysical logarithms $\ln(Q/q_T)$. Instead, each convolution should be dressed

with its proper rapidity logarithm $\ln(Q/|\vec{k}_i|)$, which is precisely what is happening in eq. (5.21).

To see in more detail the different treatment of these rapidity logarithms, we now compare the iterative solution to the naive evolution kernel eq. (3.20). To do this, we simply assume the trivial boundary condition $S(\vec{k}_T, \mu, k_T|_+) = \delta(\vec{k}_T)$. This of course means that the following result is not the properly resummed soft function, but it allows to disentangle effects from rapidity and μ evolution, since the latter are now neglected in the boundary term. Furthermore, we neglect running of α_s for simplicity to keep the following formula as compact as possible, and work only to LL accuracy. In this toy model, the rapidity anomalous dimension is given by

$$\gamma_\nu(\vec{p}_T, \mu) = 2\Gamma_{\text{cusp}}\mathcal{L}_0(\vec{p}_T, \mu)\,, \tag{5.23}$$

where at LL we can ignore all constant pieces. To keep track of only the first two emissions, we evaluate the resummed toy soft function only up to $\mathcal{O}(\Gamma_{\text{cusp}}^2)$. We then obtain from eq. (5.21)

$$\begin{aligned}
S^{(\text{toy})}(\vec{p}_T, \mu, \nu) &= \delta(\vec{p}_T) + \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu_1}{\nu_1}\gamma_\nu(\vec{p}_T, \mu) \\
&+ \int_{p_T|_+}^{\nu} \frac{\mathrm{d}\nu_1}{\nu_1} \int \mathrm{d}^2\vec{k}_1\, \gamma_\nu(\vec{p}_T - \vec{k}_1, \mu) \int_{k_1|_+}^{\nu_1} \frac{\mathrm{d}\nu_2}{\nu_2}\gamma_\nu(\vec{k}_1, \mu) + \cdots \\
&= \delta(\vec{p}_T) - 2\frac{\Gamma_{\text{cusp}}}{\pi\mu^2}\left[\frac{\mu^2}{\vec{p}_T^2}\ln\frac{\vec{p}_T^2}{\mu\nu}\right]_+^\mu + \frac{2\Gamma_{\text{cusp}}^2}{\pi\mu^2}\left[\frac{\mu^2}{\vec{p}_T^2}\ln\frac{\vec{p}_T^2}{\mu^2}\ln\frac{\vec{p}_T^2}{\nu^2}\ln\frac{\vec{p}_T^2}{\mu\nu}\right]_+^\mu \\
&+ 4\zeta_3\Gamma_{\text{cusp}}^2\mathcal{L}_0(\vec{p}_T, \nu) + \mathcal{O}(\Gamma_{\text{cusp}}^3)\,. \tag{5.24}
\end{aligned}$$

For comparison, we calculate the soft function using the naive resummation eqs. (3.19) and (3.20), where the starting scale ν_0 is kept arbitrary. Similarly to above, we set the boundary term to $S(\vec{p}_T, \mu, \nu_0) = \delta(\vec{p}_T)$, and evaluate the resummed soft function only at LL without α_s running. We find

$$\begin{aligned}
S^{(\text{toy})}(\vec{p}_T, \mu, \nu) &= \delta(\vec{p}_T) + \ln\frac{\nu}{\nu_0}\gamma_\nu(\vec{p}_T, \mu) + \frac{1}{2}\ln^2\frac{\nu}{\nu_0}\int \mathrm{d}^2\vec{k}_T\, \gamma_\nu(\vec{p}_T - \vec{k}_T, \mu)\gamma_\nu(\vec{k}_T, \mu) \\
&= \delta(\vec{p}_T) + \frac{\Gamma_{\text{cusp}}}{\pi\mu^2}\ln\frac{\nu^2}{\nu_0^2}\left[\frac{\mu^2}{\vec{p}_T^2}\right]_+^\mu + \frac{\Gamma_{\text{cusp}}^2}{\pi\mu^2}\ln^2\frac{\nu^2}{\nu_0^2}\left[\frac{\mu^2}{\vec{p}_T^2}\ln\frac{\vec{p}_T^2}{\mu^2}\right]_+^\mu\,. \tag{5.25}
\end{aligned}$$

Comparing the two results eqs. (5.24) and (5.25), we observe a very different logarithmic structure. In particular, in the solution eq. (5.24) derived from the known exact solution, all rapidity logs are necessarily sensitive to the soft function $\vec{p}_T$. This becomes crucial when the soft function is inserted into further convolutions, for example to predict the higher order terms of $S^{(\text{toy})}$ itself.

5.2 Beam functions

The beam function RGE in eqs. (3.7) and (3.13)

$$\mu\frac{\mathrm{d}B(\omega,\vec{p}_T,\mu,\nu)}{\mathrm{d}\mu} = \gamma_B(\omega,\mu,\nu)\,B(\omega,\vec{p}_T,\mu,\nu)\,, \tag{5.26}$$

$$\nu\frac{\mathrm{d}B(\omega,\vec{p}_T,\mu,\nu)}{\mathrm{d}\nu} = -\frac{1}{2}\int\mathrm{d}^2\vec{k}_T\,\gamma_\nu(\vec{k}_T,\mu)\,B(\omega,\vec{p}_T-\vec{k}_T,\mu,\nu) \tag{5.27}$$

are very similar to the soft function RGEs, and hence both iterative solution eqs. (5.10) and (5.11) as well as the all-order solutions eqs. (5.13) and (5.21) can be applied to the beam function upon proper replacement of anomalous dimensions and boundary terms.

The main difference to the soft function is that the canonical rapidity scale is $\nu_B = \omega$ rather than $\nu_B = p_T|_+$, and hence does not require distributional scale setting. The solution of the μ-RGE at the canonical ν-scale is

$$B(\omega,\vec{p}_T,\mu,\nu_B=\omega) = \delta(\vec{p}_T)\,\mathcal{B}(\omega,\mu) \tag{5.28}$$
$$+\left[\frac{1}{2\pi p_T}\frac{\mathrm{d}}{\mathrm{d}p_T}\mathcal{B}(\omega,p_T)\exp\left\{\int_{p_T}^\mu\frac{\mathrm{d}\mu'}{\mu'}\,\gamma_B(\omega,\mu',\nu_B=\omega)\right\}\right]_+^\mu.$$

The boundary term is defined as

$$\theta(p_T)\mathcal{B}(\omega,p_T) = \int_{|\vec{k}_T|\leq p_T}\mathrm{d}^2\vec{k}_T\,B(\vec{k}_T,\mu=p_T,\nu_B=\omega) \tag{5.29}$$

and is expanded as

$$\mathcal{B}(\omega,\mu) = \sum_{n=0}^\infty B_n(\omega,\mu)\left[\frac{\alpha_s(\mu)}{4\pi}\right]^n. \tag{5.30}$$

It is more complicated than for the soft function, because the beam functions are further matched onto PDFs by an operator product expansion (see ref. [62] for more details),

$$B_a(\omega,\vec{p}_T,\mu,\nu) = \sum_i\int\frac{\mathrm{d}z}{z}\,\mathcal{I}_{ai}(z,\vec{p}_T,\mu,\nu)\,f_i\Big(\frac{\omega/E_{\mathrm{cm}}}{z},\mu\Big). \tag{5.31}$$

Here B_a is the beam function for flavor a and f_i is the PDF for flavor i. (We ignore for the moment that gluon beam functions furthermore have a helicity structure.) While applying eq. (5.29) to eq. (5.31) eliminates all distributions in $\vec{p}_T$ in the $\mathcal{I}_{ai}$ matching kernels, the boundary term still involves a convolution in z with the PDF.

The rapidity evolution kernel is obtained from eq. (5.21) by replacing all $\nu_0 = k_i|_+$ by $\nu_0 = \nu_B \sim \omega$. Since this scale is independent of p_T, the ν-integrals can be pulled out as

$$B(\omega,\vec{p}_T,\mu,\nu) = \int\mathrm{d}^2\vec{k}_T\,V_B(\vec{k}_T,\nu,\nu_B)\,B(\omega,\vec{p}_T-\vec{k}_T,\mu,\nu_B)\,,$$

$$V_B(\vec{k}_T,\nu,\nu_B) = \delta(\vec{k}_T) + \sum_{n=0}^\infty\frac{1}{n!}\left(-\frac{1}{2}\ln\frac{\nu}{\nu_B}\right)^n(\gamma_\nu\otimes^n)(\vec{k}_T,\mu)\,, \tag{5.32}$$

where the $-1/2$ arises from $\gamma_{\nu,B} = -\gamma_\nu/2$. Hence, for the beam function we recover the simple exponential in convolution space with the fixed-order boundary condition $B(\omega,\vec{p}_T,\mu,\nu_B)$. This allows for a simple scale variation by choosing ν_B to be a suitable profile.
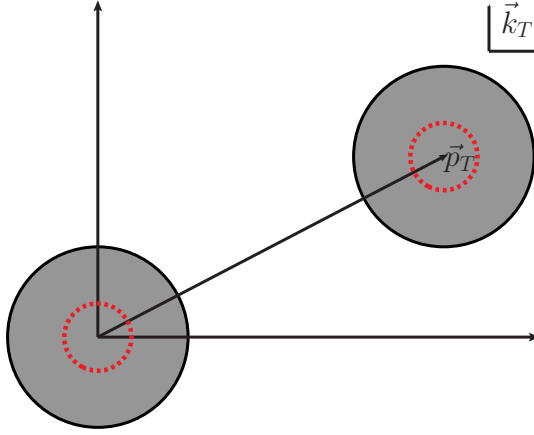
Figure 2. Illustration of division of the integration range of $\vec{k}_T$ in discs around the two singularities at $\vec{0}$ and $\vec{p}_T$. The grey discs represent the disc radius $p_T \gg \Lambda \gg \Lambda_{\mathrm{QCD}}$. The red, dashed line illustrates the location of the Landau poles at $|\vec{k}_T| = \Lambda_{\mathrm{QCD}}$ and $|\vec{p}_T - \vec{k}_T| = \Lambda_{\mathrm{QCD}}$.

5.3 Perturbativity of convolutions

The solution eq. (5.21) of the rapidity RGE contains multiple convolutions of the form $(\gamma_\nu \otimes S)(\vec{p}_T)$ or $(\gamma_\nu \otimes B)(\vec{p}_T)$. These convolutions naturally probe $\alpha_s(|\vec{k}_T|)$ in the nonperturbative regime $|\vec{k}_T| \lesssim \Lambda_{\mathrm{QCD}}$. Fortunately, nonperturbative effects turn out to be suppressed by $\mathcal{O}(\Lambda_{\mathrm{QCD}}^2/p_T^2)$, which ultimately means that the q_T spectrum is also perturbative up to corrections $\mathcal{O}(\Lambda_{\mathrm{QCD}}^2/q_T^2)$, as one would naively expect.

To show this, we consider the convolution $(f_1 \otimes f_2)(\vec{p}_T, \mu)$ of two generic distributions

$$\left[f_1(|\vec{k}_T|) \right]_+^\mu, \qquad \left[f_2(|\vec{k}_T|) \right]_+^\mu, \tag{5.33}$$

which only depend on the magnitude $|\vec{k}_T|$, which is the typical case we are interested in. In general, the distributions can depend on further parameters which we suppress, as they do not affect the following calculation. We also do not include $\delta(\vec{k}_T)$ terms, as they trivially factor out of the convolution $f_1 \otimes f_2$.

Assuming perturbative $|\vec{p}_T|, \mu \gg \Lambda_{\mathrm{QCD}}$, we introduce an auxiliary parameter Λ fulfilling $\Lambda_{\mathrm{QCD}} \ll \Lambda \ll p_T$ to split the integration range of the convolution

$$(f_1 \otimes f_2)(\vec{p}_T, \mu) = \int \mathrm{d}^2\vec{k}_T \left[f_1(|\vec{p}_T - \vec{k}_T|) \right]_+^\mu \left[f_2(|\vec{k}_T|) \right]_+^\mu \tag{5.34}$$

into three pieces (see fig. 2), two discs around $\vec{0}$ and $\vec{p}_T$ of radius Λ, and the rest of the plane.

The convolution integral then splits into three pieces,

$$(f_1 \otimes f_2)(\vec{p}_T, \mu) = \int_{B_\Lambda(\vec{0})} \mathrm{d}^2\vec{k}_T \, f_1(|\vec{p}_T - \vec{k}_T|) \left[f_2(|\vec{k}_T|) \right]_+^\mu + \int_{B_\Lambda(\vec{p}_T)} \mathrm{d}^2\vec{k}_T \left[f_1(|\vec{p}_T - \vec{k}_T|) \right]_+^\mu f_2(|\vec{k}_T|)$$

$$+ \int_{\mathbb{R}^2 \setminus (B_\Lambda(\vec{0}) \cup B_\Lambda(\vec{p}_T))} \mathrm{d}^2\vec{k}_T \, f_1(|\vec{p}_T - \vec{k}_T|) f_2(|\vec{k}_T|) \,, \tag{5.35}$$

where $B_\Lambda(\vec{p}_T)$ denotes the ball of radius Λ around the origin $\vec{p}_T$. In the first integral, the range is restricted to be close the origin, and hence the plus prescription of f_1 can be dropped, and vice versa for the second integral. The third integral is separated from both poles by $\Lambda \gg \Lambda_{QCD}$ such that both distributions can be dropped. Furthermore, it also avoids both Landau poles at $|\vec{k}_T| = \Lambda_{QCD}$ and $|\vec{p}_T - \vec{k}_T| = \Lambda_{QCD}$, and hence is completely perturbative. This isolates all the nonperturbative contributions into the first two integrals.

We now focus on the first integral. Since by construction $|\vec{k}_T| \leq \Lambda \ll |\vec{p}_T|$, we can approximate $|\vec{p}_T - \vec{k}_T|^2 = |\vec{p}_T|^2 \big(1 + \mathcal{O}(\vec{k}_T^2/\vec{p}_T^2)\big)$. Assuming that the functions only depend on the squared magnitudes, we find

$$
\begin{aligned}
\int_{B_\Lambda(\vec{0})} \mathrm{d}^2\vec{k}_T \, f_1|(\vec{p}_T - \vec{k}_T|)\Big[f_2(|\vec{k}_T|)\Big]_+^\mu &= \int_{B_\Lambda(\vec{0})} \mathrm{d}^2\vec{k}_T \, f_1(|\vec{p}_T|)\Big[1 + \mathcal{O}\Big(\frac{\vec{k}_T^2}{\vec{p}_T^2}\Big)\Big]\Big[f_2(|\vec{k}_T|)\Big]_+^\mu \\
&= f_1(|\vec{p}_T|)\Big[1 + \mathcal{O}\Big(\frac{\Lambda^2}{\vec{p}_T^2}\Big)\Big]\int_{|\vec{k}_T|\leq\Lambda} \mathrm{d}^2\vec{k}_T \Big[f_2(|\vec{k}_T|)\Big]_+^\mu \\
&= f_1(|\vec{p}_T|)\Big[1 + \mathcal{O}\Big(\frac{\Lambda^2}{\vec{p}_T^2}\Big)\Big]2\pi\int_\mu^\Lambda \mathrm{d}k_T \, k_T f_2(k_T)\,.
\end{aligned}
$$
(5.36)

Here we used that the plus distribution by definition removes all contributions for $|\vec{k}_T| \leq \mu$ and therefore its integral is only sensitive to the upper integration limit Λ. Hence, we find that the leading nonperturbative effect is suppressed as $\mathcal{O}(\Lambda^2/\vec{p}_T^2)$.

The same approximation can be applied to the second integral in eq. (5.35), such that eq. (5.35) becomes

$$
\begin{aligned}
(f_1 \otimes f_2)(\vec{p}_T, \mu) = \Big[1 + \mathcal{O}\Big(\frac{\Lambda^2}{p_T^2}\Big)\Big]2\pi\int_\mu^\Lambda \mathrm{d}k_T \, k_T \Big[f_1(p_T)f_2(k_T) + f_1(k_T)f_2(p_T)\Big] + \\
+ \int_{\mathbb{R}^2\setminus(B_\Lambda(\vec{0})\cup B_\Lambda(\vec{p}_T))} \mathrm{d}^2\vec{k}_T \, f_1(\vec{p}_T - \vec{k}_T)f_2(\vec{k}_T)\,.
\end{aligned}
$$
(5.37)

In conclusion, we find that convolutions $(f_1 \otimes f_2)(\vec{p}_T, \mu)$ are well behaved with nonperturbative contributions suppressed as $\mathcal{O}(\Lambda^2/\vec{p}_T^2)$, as long as both $|\vec{p}_T|$ and μ themselves are perturbative, $|\vec{p}_T|, \mu \gg \Lambda_{QCD}$.

Illustration: The above result can be illustrated by an explicit example, which also shows a different strategy to evaluate such convolutions. We consider the convolution $(f \otimes f)(\vec{p}_T, \mu)$ of

$$
f(\vec{p}_T, \mu) = \Big[\frac{\alpha_s(|\vec{p}_T|)}{\vec{p}_T^2}\Big]_+^\mu\,,
$$
(5.38)

which is the prototype of nonperturbative convolutions appearing in the q_T spectrum. Even without an explicit calculation, its generic form is known from its μ-dependence. From

$$
\mu\frac{\mathrm{d}f(\vec{p}_T, \mu)}{\mathrm{d}\mu} = -2\pi\alpha_s(\mu)\delta(\vec{p}_T)\,,
$$
(5.39)

it follows that

$$
\mu\frac{\mathrm{d}(f \otimes f)(\vec{p}_T, \mu)}{\mathrm{d}\mu} = -4\pi\alpha_s(\mu)\Big[\frac{\alpha_s(|\vec{p}_T|)}{\vec{p}_T^2}\Big]_+^\mu\,.
$$
(5.40)

Integrating this according to the prescription eq. (2.79), we find

$$(f \otimes f)(\vec{p}_T, \mu) = (f \otimes f)(\vec{p}_T, \mu_0 = p_T|_+) - 4\pi \int_{\mu_0 = p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \alpha_s(\mu') \left[\frac{\alpha_s(|\vec{p}_T|)}{\vec{p}_T^2} \right]_+^{\mu'}$$

$$= (f \otimes f)(\vec{p}_T, \mu_0 = p_T|_+) - \left[\frac{4\pi\alpha_s(|\vec{p}_T|)}{\vec{p}_T^2} \int_{p_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \alpha_s(\mu') \right]_+^{\mu}. \quad (5.41)$$

The plus distribution is perturbative as long as both $|\vec{p}_T|, \mu \gg \Lambda_{\mathrm{QCD}}$, which is precisely the regime we have considered in the above general derivation. The structure of the boundary term is not known, but its most generic structure to be μ_0-independent is

$$(f \otimes f)(\vec{p}_T, \mu_0 = p_T|_+) = \left[\frac{-\Lambda_{\mathrm{QCD}}/|\vec{p}_T| \cdot F'(\Lambda_{\mathrm{QCD}}/|\vec{p}_T|)}{2\pi\vec{p}_T^2} \right]_+^{\xi} + F(\Lambda_{\mathrm{QCD}}/\xi)\delta(\vec{p}_T), \quad (5.42)$$

where the dependence on the arbitrary parameter ξ cancels between the two terms. Since $(f \otimes f)$ scales as $1/\vec{p}_T^2$, F must be a scalar function. It can hence only depend on $\Lambda_{\mathrm{QCD}}/|\vec{p}_T|$, as no other scales can be combined into a scaleless number. To be any physically reasonable function, F should vanish (or at most become constant) for $\Lambda_{\mathrm{QCD}} \to 0$. Furthermore we expect it to only depend on the magnitude $\vec{p}_T^2$, and hence the boundary term should scale as $\mathcal{O}(\Lambda_{\mathrm{QCD}}^2/\vec{p}_T^2)$, just as expected from the general calculation.

6 The resummed transverse-momentum spectrum

We now discuss in more detail the final result for the resummed transverse momentum spectrum eq. (3.1),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \sigma_0\, H(Q, \mu) \int \mathrm{d}^2\vec{k}_a\, \mathrm{d}^2\vec{k}_b\, \mathrm{d}^2\vec{k}_s\, \delta(\vec{q}_T - \vec{k}_a - \vec{k}_b - \vec{k}_s)$$

$$\times B_a(\omega_a, \vec{k}_a, \mu, \nu)\, B_b(\omega_b, \vec{k}_b, \mu, \nu)\, S(\vec{k}_s, \mu, \nu), \quad (6.1)$$

where $\omega_{a,b} = Q^{\pm Y}$. All distributional logarithms in the q_T spectrum are fully resummed in momentum space by using the RG-evolved hard, beam, and soft functions in eq. (6.1).

To simplify the complicated structure of the ν-evolution, we use eq. (5.32) to evolve the beam functions $B_{a,b}$ from their natural rapidity scales $\nu_{a,b} \sim \omega_{a,b}$ to ν. Since both kernels are simple exponentials in convolution space, they can be combined into a single evolution kernel, which due to $2\gamma_{\nu,B} = -\gamma_{\nu,S} = -\gamma_\nu$ precisely yields the kernel used to shift the soft function, eq. (5.22),

$$(B_a \otimes B_b)(\vec{p}_T, \mu, \nu) = \int \mathrm{d}^2\vec{k}_T\, \mathrm{d}^2\vec{k}_a\, \mathrm{d}^2\vec{k}_b\, \delta(\vec{p}_T - \vec{k}_T - \vec{k}_a - \vec{k}_b)$$

$$\times V_S(\vec{k}_T, \mu, \nu_B, \nu)\, B_a(\omega_a, \vec{k}_a, \mu, \nu_a)\, B_b(\omega_b, \vec{k}_b, \mu, \nu_b), \quad (6.2)$$

where $\nu_B = \sqrt{\nu_1\nu_2}$. Since beam functions inside the final convolution are evaluated at their natural scales $\nu_{a,b} \sim \omega_{a,b}$, they are free of large rapidity logarithms.

In addition to soft and beam evolution, we also need the RG evolution of the hard function, which is straightforwardly obtained from eq. (3.6),

$$H(Q, \mu) = H(Q, \mu_H) \exp\left[\int_{\mu_H}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu')\right],$$ (6.3)

where $\mu_H \sim Q$ ensures that $H(Q, \mu_H)$ is free of large logarithms.

To assemble the cross section eq. (6.1), we set the arbitrary scale μ to the total transverse momentum q_T. As usual, this requires distributional scale setting, $\mu = q_T|_+ \equiv \mu_T|_+$.[6] We keep the symbolic notation μ_T rather than q_T to make the origin of all factors of q_T and μ_T in the final formula clear. We then find

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \sigma_0\, H(Q, \mu_H) \frac{1}{2\pi q_T} \frac{\mathrm{d}}{\mathrm{d}q_T} \int_{|\vec{p}_T| \leq q_T} \mathrm{d}^2\vec{p}_T$$

$$\times \exp\left[\int_{\mu_H}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu')\right] \int \mathrm{d}^2\vec{k}_a\, \mathrm{d}^2\vec{k}_b\, \mathrm{d}^2\vec{k}_s\, \delta(\vec{p}_T - \vec{k}_a - \vec{k}_b - \vec{k}_s) \int \mathrm{d}^2\vec{k}'_s$$

$$\times \left[\delta(\vec{k}_s - \vec{k}'_s) + \sum_{n=1}^{\infty} \prod_{i=1}^{n} \int_{k_{i-1}|_+}^{\nu_{i-1}} \frac{\mathrm{d}\nu_i}{\nu_i} \int \mathrm{d}^2\vec{k}_i\, \gamma_\nu(\vec{k}_{i-1} - \vec{k}_i, \mu_T)\, \delta\left(\vec{k}_s - \vec{k}'_s - \sum_i \vec{k}_i\right)\right]$$

$$\times B_a(\omega_a, \vec{k}_a, \mu_T, \nu_a)\, B_b(\omega_b, \vec{k}_b, \mu_T, \nu_b)\, S(\vec{k}'_s, \mu_T, k'_s|_+).$$ (6.4)

The first line contains the μ_T-independent pieces and the cumulant integral from the distributional scale setting $\mu = \mu_T|_+$. The second line makes explicit that the μ-evolution resums logarithms $\ln(\mu_T/\mu_H) \sim \ln(q_T/Q)$. The third line contains the ν-evolution kernel, eq. (5.21), to evolve the soft function to $\nu_0 \equiv \nu_B = \sqrt{\nu_a \nu_b}$. The last line contains beam and soft functions evaluated at their natural ν-scales, and hence all rapidity logarithms in eq. (6.4) are fully resummed.

Note that eq. (6.4) could be the starting point for a numerical evaluation. Although the infinite number of convolutions cannot be calculated in closed form, one could for example evaluate the result iteratively and truncate the sum in eq. (6.4) once a desired numerical accuracy is reached.

One of the original motivations to carry out the momentum-space resummation (see refs. [14, 21, 82]) was to avoid the intrinsic nonperturbative sensitivity at large $q_T \gg \Lambda_{\mathrm{QCD}}$ arising in the Fourier-space resummation (see sec. 7.1.1 for more details). From eq. (6.4) it is clear that the analogous sensitivity is still present in the momentum-space resummation, namely the cross section is intrinsically sensitive to the nonperturbative contributions from the rapidity anomalous dimension, since the convolutions probe $|\vec{k}'_i| \lesssim \Lambda_{\mathrm{QCD}}$. Fortunately, in either case these effects turn out to be suppressed by $\Lambda_{\mathrm{QCD}}^2/q_T^2$, as shown in sec. 5.3.

The last line of eq. (6.4) contains the μ-evolved beam and soft functions given by eqs. (5.13) and (5.28). To investigate its structure in more detail, we omit the pure δ-terms in eqs. (5.13) and (5.28), which trivially factor out of the convolution, and for simplicity

---

[6]Alternatively, one could choose $\mu = \mu_H$, but then the final structure is less intuitive.

write everything into a single plus distribution. This gives

$$B_a(\omega_a, \vec{k}_a, \mu_T, \nu_a) \, B_b(\omega_b, \vec{k}_b, \mu_T, \nu_b) \, S(\vec{k}_s, \mu_T, k_s|_+) \tag{6.5}$$
$$\supset \left[\frac{1}{2\pi k_a} \frac{\mathrm{d}}{\mathrm{d}k_a} \frac{1}{2\pi k_b} \frac{\mathrm{d}}{\mathrm{d}k_b} \frac{1}{2\pi k_s} \frac{\mathrm{d}}{\mathrm{d}k_s} \mathcal{B}(\omega_a, k_a) \, \mathcal{B}(\omega_b, k_b) \, \mathcal{S}[\alpha_s(k_s)] \right.$$
$$\times \left. \exp\left\{ \int_{k_a}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, \gamma_B(\omega_a, \mu', \omega_a) + \int_{k_b}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, \gamma_B(\omega_b, \mu', \omega_b) + \int_{k_s}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, \gamma_S(\mu', k_s) \right\} \right]_+^{\mu_T} .$$

The first line of the plus distribution contains the pure fixed-order boundary terms, which are free of any logarithmic distributions. The exponential explicitly contains μ-evolutions from the convolution momenta to the scale μ_T. These are required to properly satisfy the exact path independence of the μ- and ν-evolution. This can be checked explicitly by verifying the μ_T-independence of eq. (6.4). For practical purposes, it might be sufficient to use the fixed-order expansion for eq. (6.5), but since this would deviate from the strict resummation order, it would have to be verified numerically.

6.1 Illustration at LL

To illustrate eq. (6.4), we evaluate it to leading-logarithmic (LL) order, which is strictly defined by keeping the LO boundary terms for hard, beam, and soft function, keeping Γ_0 and β_0 in the anomalous dimensions, and dropping all noncusp anomalous dimensions. In this limit, the μ-evolved beam function boundary term becomes

$$B(\omega, \vec{p}_T, \mu_T, \nu = \omega) = \delta(\vec{p}_T) f(\omega, \mu_T) + \left[\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} f(\omega, p_T) \right]_+^{\mu_T} \approx \delta(\vec{p}_T) f(\omega, \mu_T) , \quad (6.6)$$

where the μ-evolution drops out because $\gamma_B(\mu, \nu = \omega)$ vanishes at LL. We can also drop the PDF evolution, since it is a subleading logarithmic effect (i.e. the PDF anomalous dimensions are counted as noncusp contributions). The μ-evolved soft function boundary term at LL is

$$S(\vec{k}_s, \mu_T, k_s|_+) = \delta(\vec{k}_s) + \left[\frac{1}{2\pi k_s} \frac{\mathrm{d}}{\mathrm{d}k_s} \exp\left\{ \int_{k_s}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \ln \frac{\mu'}{k_s} \right\} \right]_+^{\mu_T} . \tag{6.7}$$

With strict canonical scales, $\nu_B = Q$ and $\mu_T = q_T$, eq. (6.4) is then given by

$$\frac{\mathrm{d}\sigma^{\mathrm{LL}}}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \sigma_0 \frac{1}{2\pi q_T} \frac{\mathrm{d}}{\mathrm{d}q_T} f_a(\omega_a, \mu_T) \, f_b(\omega_b, \mu_T) \int_{|\vec{p}_T| \leq q_T} \mathrm{d}^2 \vec{p}_T \, \exp\left[\int_{\mu_H}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, \gamma_H(Q, \mu') \right] \int \mathrm{d}^2 \vec{k}_s$$

$$\times \left[\delta(\vec{p}_T - \vec{k}_s) + \sum_{n=1}^{\infty} \prod_{i=1}^{n} \int_{k_{i-1}|_+}^{\nu_{i-1}} \frac{\mathrm{d}\nu_i}{\nu_i} \int \mathrm{d}^2 \vec{k}_i \, \gamma_\nu(\vec{k}_{i-1} - \vec{k}_i, \mu_T) \, \delta\left(\vec{p}_T - \vec{k}_s - \sum_i \vec{k}_i \right) \right]$$

$$\times \left(\delta(\vec{k}_s) + \left[\frac{1}{2\pi k_s} \frac{\mathrm{d}}{\mathrm{d}k_s} \exp\left\{ \int_{k_s}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \, 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \ln \frac{\mu'}{k_s} \right\} \right]_+^{\mu_T} \right) . \tag{6.8}$$

In this form, it is somewhat reminiscent of the form suggested by Dokshitzer, Dyakonov, and Troyan for the LL cross section in refs. [83, 84],

$$\frac{\mathrm{d}\sigma^{(\mathrm{DDT})}}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \frac{\sigma_0}{\pi} \frac{\mathrm{d}}{\mathrm{d}q_T^2} f_a(\omega_a, q_T) \, f_b(\omega_b, q_T) \, e^{\mathbb{S}(Q, q_T)} . \tag{6.9}$$

However, comparing this to eq. (6.8), $e^{\mathbb{S}}$ would be given by the full cumulant integral of eq. (6.8), which has an exponential structure but is by no means a simple exponential. Only if one were to neglect all convolutions, i.e. keep only the first $\delta(\vec{p}_T - \vec{k}_s)$ term in the rapidity evolution in the second line of eq. (6.8), one would find the expected Sudakov factor $\mathbb{S} = \int_Q^{q_T} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu')$. In a full LL resummation, the rapidity evolution forbids such a simple relation, and as a consequence the simple DDT form eq. (6.9) cannot hold for the q_T spectrum. (At higher orders, the effect of PDF running neglected in eq. (6.6) would similarly spoil this.) In fact, the DDT formula was never derived as LL solution of a factorization theorem, but from a summation of ladder diagrams relevant at LL [83].

An interesting feature of eq. (6.8) is that after carrying out all convolutions, the result can always be written in the form $\delta(\vec{p}_T) + \sum c_n \mathcal{L}_n(\vec{p}_T, \mu_T) + \sum d_m \mathcal{L}_m(\vec{p}_T, \nu_B)$. With strict canonical scale setting, $\mu_T = q_T|_+$, the first sum vanishes due to $\mathcal{L}_n(\vec{q}_T, q_T|_+) = 0$. In contrast, the second sum is converted into $\mathcal{L}_m(\vec{p}_T, \nu_B) \to \mathcal{L}_m(\vec{q}_T, \nu_B)$. Since $\nu_B \sim Q$, this precisely yields the rapidity logarithms $\ln(Q/q_T)$, which however do not have a simple exponential structure any more. For noncanonical scales, the $\mathcal{L}_n(\vec{p}_T, \mu_T)$ would not exactly vanish, but yield small corrections that probe the all-order logarithmic structure.

As an illustrative example, we carry out the first few convolutions, but ignore α_s-running for simplicity. This is analogous to the simple study in sec. 5.1.3. We find

$$
\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \sigma_0 \frac{1}{2\pi q_T} \frac{\mathrm{d}}{\mathrm{d}q_T} \theta(q_T) f_a(\omega_a, q_T) f_b(\omega_b, q_T) \exp\left[-\frac{\Gamma_{\mathrm{cusp}}}{2} \ln^2 \frac{Q^2}{q_T^2} \right]
$$
$$
\times \left[1 - 2\Gamma_{\mathrm{cusp}}^2 \zeta_3 \ln \frac{Q^2}{q_T^2} + \Gamma_{\mathrm{cusp}}^3 \left(\frac{2\zeta_3}{3} \ln^3 \frac{Q^2}{q_T^2} + 6\zeta_5 \ln \frac{Q^2}{q_T^2} \right) \right.
$$
$$
\left. + \Gamma_{\mathrm{cusp}}^4 \left(-4\zeta_5 \ln^3 \frac{Q^2}{q_T^2} + 10\zeta_3^2 \ln^2 \frac{Q^2}{q_T^2} - 30\zeta_7 \ln \frac{Q^2}{q_T^2} \right) + \mathcal{O}(\Gamma_{\mathrm{cusp}}^5) \right]. \quad (6.10)
$$

We have not evaluated the derivative, as in this form the terms in square bracket are exactly the result of the rapidity evolution. It clearly induces apparent subleading terms in the cross section, which however are part of the strict LL order that is based on the expansion of the anomalous dimensions.

For comparison, the Fourier-resummed spectrum (see sec. 5.1.3 for more details) with its canonical scale choices $\nu_S = \mu = b_0/b_T$, neglecting again α_s-running, is given by

$$
\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}Y \mathrm{d}\vec{q}_T} = \sigma_0 f_a(\omega_a, \mu) f_b(\omega_b, \mu) \int \frac{\mathrm{d}^2 \vec{b}_T}{(2\pi)^2} e^{i\vec{b}_T \cdot \vec{q}_T} \exp\left[-\frac{\Gamma_{\mathrm{cusp}}}{2} \ln^2 \frac{Q^2 b_T^2}{b_0^2} \right]. \quad (6.11)
$$

As before, we neglect running in the PDFs for simplicity to evaluate them at the random scale μ. The Sudakov double logarithm is clearly a well-behaved function that drops fast for both $\vec{b}_T \to 0$ and $\vec{b}_T \to \infty$, and hence the inverse Fourier integral must be fine as well. To compare this expression to eq. (6.10), we split $\ln(Qb_T/b_0) = \ln(Q/\mu) + \ln(\mu b_T/b_0)$, with μ arbitrary. Expanding the integrand in Γ_{cusp}, we only need to Fourier-transform powers

of logarithms $\ln^n(\mu b_T/b_0)$, which is well known (see appendix C.2). Setting $\mu = q_T|_+$ gives

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0 \frac{1}{2\pi q_T} \frac{\mathrm{d}}{\mathrm{d}q_T} \theta(q_T) f_a(\omega_a, q_T)\, f_b(\omega_b, q_T) \exp\left[-\frac{\Gamma_{\mathrm{cusp}}}{2} \ln^2 \frac{Q^2}{q_T^2}\right]$$
$$\times \left[1 - 2\Gamma_{\mathrm{cusp}}^2 \zeta_3 \ln \frac{Q^2}{q_T^2} + \Gamma_{\mathrm{cusp}}^3 \left(\frac{2\zeta_3}{3}\ln^3 \frac{Q^2}{q_T^2} + 6\zeta_5 \ln \frac{Q^2}{q_T^2} - \frac{10}{3}\zeta_3^2\right)\right.$$
$$\left. + \Gamma_{\mathrm{cusp}}^4 \left(-4\zeta_5 \ln^3 \frac{Q^2}{q_T^2} + 10\zeta_3^2 \ln^2 \frac{Q^2}{q_T^2} - 30\zeta_7 \ln \frac{Q^2}{q_T^2} + 28\zeta_3\zeta_5\right) + \mathcal{O}(\Gamma_{\mathrm{cusp}}^5)\right]. \tag{6.12}$$

The result agrees with eq. (6.10) up to the constant terms $-10/3\,\zeta_3^2\Gamma_{\mathrm{cusp}}^3$ and $28\zeta_3\zeta_5\Gamma_{\mathrm{cusp}}^4$. These nonlogarithmic terms are examples of different boundary terms induced by using the Fourier-space boundary conditions. The important observation is that the logarithmic structure exactly matches in both attempts.

Lastly, we also compare this to the result from using the naive rapidity evolution of the soft function. In this case, eq. (6.1) is evaluated at $\mu = \mu_T \sim q_T$, and the soft function is taken from eqs. (3.19) and (3.20),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0\, H(Q, \mu_H) \exp\left[\int_{\mu_H}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu')\right] \int \mathrm{d}^2\vec{k}_a\, \mathrm{d}^2\vec{k}_b\, \mathrm{d}^2\vec{k}_s\, \delta(\vec{q}_T - \vec{k}_a - \vec{k}_b - \vec{k}_s)$$
$$\times B_a(\omega_a, \vec{k}_a, \mu_T, \nu_a)\, B_b(\omega_b, \vec{k}_b, \mu_T, \nu_b)$$
$$\times \int \mathrm{d}^2\vec{k}_T V_S(\vec{k}_s - \vec{k}_T, \mu_T, \nu_B, \nu_S) S(\vec{k}_s, \mu_T, \nu_S). \tag{6.13}$$

Assuming the LO boundary conditions for the hard, beam, and soft functions, we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0\, f_a(\omega_a, \mu_T)\, f_b(\omega_b, \mu_T) \exp\left[\int_{\mu_H}^{\mu_T} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu')\right]$$
$$\times \left[\delta(\vec{q}_T) + \sum_{n=1}^{\infty} \frac{1}{n!} \ln^n \frac{\nu_B}{\nu_S} (\gamma_\nu \otimes^n)(\vec{q}_T, \mu_T)\right]. \tag{6.14}$$

Setting $\mu_T = \nu_S = q_T|_+$, we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0 \frac{1}{2\pi q_T} \frac{\mathrm{d}}{\mathrm{d}q_T} \theta(q_T) f_a(\omega_a, q_T)\, f_b(\omega_b, q_T) \exp\left[-\frac{\Gamma_{\mathrm{cusp}}}{2} \ln^2 \frac{Q^2}{q_T^2}\right]$$
$$\times \left[1 + \frac{2}{3}\Gamma_{\mathrm{cusp}}^3 \ln^3 \frac{Q^2}{q_T^2} \zeta_3 + \mathcal{O}(\Gamma_{\mathrm{cusp}}^5)\right]. \tag{6.15}$$

The obtained highest logarithmic term $\Gamma_{\mathrm{cusp}}^3 \ln^3(Q^2/q_T^2)$ matches the one in eqs. (6.10) and (6.12), but all other terms are missing. The second line of eq. (6.15) is nothing but the familiar factor $e^{-2\gamma_E\omega_s}\Gamma(1-\omega_s)/\Gamma(1+\omega_s)$ from eq. (3.27), which diverges at $\omega_s = 2\Gamma_{\mathrm{cusp}}\ln(Q/q_T) = 1$. The fact that this term also appears in the correctly resummed results in eqs. (6.10) and (6.12) is quite surprising as one might expect it to get modified in order to alleviate the spurious divergence. On the other hand, since the naive rapidity evolution kernel eq. (3.27) does correctly *shift* the rapidity logarithms, it must in fact also appear in eqs. (6.10) and (6.12), and so it seems that both the $\vec{p}_T$-space and $\vec{b}_T$-space resummation still contain this divergence. However, we know that eq. (6.11) is well

behaved, and hence all the additional logarithmic terms in eqs. (6.10) and (6.12) that are not present in eq. (6.15) must conspire to cancel the divergence in the spectrum. Hence, we find the peculiar feature that apparent-NNLL and higher terms cancel the divergence caused by the apparent-NLL terms in the strict LL spectrum. A detailed numerical study of this effect would be an interesting extension of this work.

This simple exercise clearly shows that counting logarithms $\ln(Q/q_T)$ in the q_T spectrum is an intrinsically ill-defined notion. Instead, the resummation order should be defined strictly and unambiguously through the perturbative order of anomalous dimensions and pure boundary terms entering all RGEs. This is done in both our momentum-space and the Fourier-space resummation, which both lead to well-behaved results.

7 Comparison to the literature

In this section, we discuss the relation of our results to the Fourier-space resummation in the original CSS formulation and various existing implementations (sec. 7.1), as well as other approaches for a direct momentum-space resummation (secs. 7.2 and 7.3).

7.1 Resummation in Fourier space

7.1.1 Comparison to CSS formalism

The equivalence of the Fourier-space resummation in the original CSS formulation and performing the RG evolution in Fourier space in the context of SCET has been discussed several times before [16, 85] (see e.g. refs. [30, 86–88] for one-dimensional cases like thrust or threshold resummation). It is not unexpected that the two formulations are equivalent, since both are based on the same underlying factorization in the soft-collinear approximation.

Interestingly, in their first paper [38] in the context of back-to-back jets in e^+e^--collisions, Collins and Soper formulated an evolution equation in transverse momentum space that precisely corresponds to the rapidity RGE in our framework; compare eq. (6.4) in ref. [38] with eq. (3.14), where the rapidity scale ν essentially corresponds to the scale ζ in the CSS formulation. However, due to the much simpler formulation in Fourier space, as first noted by Parisi and Petronzio [74], they carried out their analysis in Fourier space. Studying in detail the properties of the inverse transform, with particular focus on the effect of the (nonperturbative) large b-region [39], then paved the road to resumming the q_T-spectrum in hadronic collisions [40].

The CSS formula of ref. [40] can be easily related to the SCET formalism by Fourier transforming the beam and soft functions and their RGEs given in sec. 3.1. For example, for the soft function we have

$$\mu\frac{\mathrm{d}\tilde{S}(\vec{b}_T,\mu,\nu)}{\mathrm{d}\mu} = \gamma_S(\mu,\nu)\tilde{S}(\vec{b}_T,\mu,\nu),$$

$$\nu\frac{\mathrm{d}\tilde{S}(\vec{b}_T,\mu,\nu)}{\mathrm{d}\nu} = \tilde{\gamma}_\nu(\vec{b}_T,\mu)\tilde{S}(\vec{b}_T,\mu,\nu), \qquad \mu\frac{\mathrm{d}\tilde{\gamma}_\nu(\vec{b}_T,\mu)}{\mathrm{d}\mu} = -4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)], \qquad (7.1)$$

and most importantly the rapidity RGE becomes multiplicative. Since logarithms in impact parameter space always depend on $\mu b_T/b_0$ or $\nu b_T/b_0$, all logarithms are minimized with boundary scale choice $\mu_0 = \nu_0 = b_0/b_T$. With this choice, the above RGEs are solved by

$$\tilde{S}(\vec{b}_T, \mu, \nu) = \exp\left[\int_{b_0/b_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_S(\mu', \nu)\right] \exp\left[\tilde{\gamma}_\nu(\vec{b}_T, b_0/b_T) \ln \frac{\nu b_T}{b_0}\right] \tilde{S}(\vec{b}_T, b_0/b_T, b_0/b_T),$$

$$\tilde{\gamma}_\nu(\vec{b}_T, \mu) = -\int_{b_0/b_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} 4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] + \tilde{\gamma}_\nu[\alpha_s(b_0/b_T)]. \tag{7.2}$$

Note that by first solving the ν-RGE at the low scale and then the μ-RGE, $\tilde{\gamma}_\nu$ enters evaluated at its natural scale $\mu_0 = b_0/b_T$ for which it reduces to the pure fixed-order boundary $\tilde{\gamma}_\nu[\alpha_s(b_0/b_T)]$. In this way, the resummation of $\tilde{\gamma}_\nu$ becomes trivial in the Fourier space evolution.

The beam functions are similarly resummed with boundary scales $\mu_0 = b_0/b_T, \nu_0 = \omega_i$,

$$\tilde{B}_a(\omega, \vec{b}_T, \mu, \nu) = \exp\left[\int_{b_0/b_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \gamma_B(\mu', \nu; \omega)\right] \exp\left[-\frac{1}{2}\tilde{\gamma}_\nu(\vec{b}_T, b_0/b_T) \ln \frac{\nu}{\omega}\right]$$

$$\times \tilde{B}_a(\omega, \vec{b}_T, \mu = b_0/b_T, \nu = \omega), \tag{7.3}$$

and can furthermore be matched onto PDFs using an operator product expansion,

$$\tilde{B}_a(\omega, \vec{b}_T, \mu, \nu) = \sum_i \int \frac{\mathrm{d}z}{z} \tilde{\mathcal{I}}_{ai}(z, \vec{b}_T, \mu, \nu) f_i\left(\frac{\omega/E_{\mathrm{cm}}}{z}, \mu\right), \tag{7.4}$$

where E_{cm} is hadronic center-of-mass energy, $\tilde{\mathcal{I}}_{ai}$ are perturbative matching coefficients, and f_i are standard PDFs. Altogether, the resummed cross section eq. (3.1) takes the simple form

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0 \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2} e^{i\vec{b}_T \cdot \vec{q}_T} \sum_{i,j} \int \frac{\mathrm{d}z_a \mathrm{d}z_b}{z_a z_b} f_i\left(\frac{\omega_a/E_{\mathrm{cm}}}{z_a}, \frac{b_0}{b_T}\right) f_j\left(\frac{\omega_b/E_{\mathrm{cm}}}{z_b}, \frac{b_0}{b_T}\right)$$

$$\times \exp\left[-\int_{b_0/b_T}^{Q} \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu') + \ln \frac{Qb_T}{b_0} \tilde{\gamma}_\nu[\alpha_s(b_0/b_T)]\right]$$

$$\times H(Q, \mu = Q) \tilde{\mathcal{I}}_{ai}(z_a, b_T) \tilde{\mathcal{I}}_{bj}(z_b, b_T) \tilde{S}(b_T). \tag{7.5}$$

The functions $\tilde{\mathcal{I}}_{ai}(z, b_T) \equiv \tilde{\mathcal{I}}_{ai}(z, b_T, \mu_0 = b_0/b_T, \nu_0 = \omega)$ and $\tilde{S}(b_T) \equiv \tilde{S}(\vec{b}_T, \mu_0 = \nu_0 = b_0/b_T)$ are all free of logarithms. We have explicitly split the Sudakov exponential into a piece originating from the μ-RGEs and one from the rapidity RGE. This result should be compared to the CSS formula [40],

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q\mathrm{d}Y\mathrm{d}\vec{q}_T} = \sigma_0 \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2} e^{i\vec{b}_T \cdot \vec{q}_T} \sum_{i,j} \int \frac{\mathrm{d}z_a \mathrm{d}z_b}{z_a z_b} f_i\left(\frac{\omega_a/E_{\mathrm{cm}}}{z_a}, \frac{b_0}{b_T}\right) f_j\left(\frac{\omega_b/E_{\mathrm{cm}}}{z_b}, \frac{b_0}{b_T}\right) \tag{7.6}$$

$$\times C_{ai}(z_a, b_T) C_{bj}(z_b, b_T) \exp\left[-\int_{(b_0/b_T)^2}^{Q^2} \frac{\mathrm{d}\bar{\mu}^2}{\bar{\mu}^2} \left(\ln \frac{Q^2}{\bar{\mu}^2} A[\alpha_s(\bar{\mu})] + 2B[\alpha_s(\bar{\mu})]\right)\right].$$

Comparing eqs. (7.5) and (7.6), we see that the functions C_{aj} correspond to the product of hard and soft function with the beam function matching kernels $\tilde{\mathcal{I}}_{ai}$, showing that they

can be further split into a process-dependent hard virtual and a process-independent soft-collinear piece, as was noticed also in refs. [46, 50]. The evolution kernels are related by

$$A(\alpha_s) = \Gamma_{\text{cusp}}(\alpha_s) + \beta(\alpha_s)\frac{\mathrm{d}\tilde{\gamma}_\nu(\alpha_s)}{\mathrm{d}\alpha_s}\,, \quad 2B(\alpha_s) = \gamma_H(\alpha_s) - \tilde{\gamma}_\nu(\alpha_s)\,. \tag{7.7}$$

The important observation that $A(\alpha_s)$ receives a contribution from the rapidity anomalous dimension in addition to the cusp anomalous dimension was first made in ref. [16].

To compare to our canonical resummation in momentum space, first note that setting $\nu_0 = b_0/b_T$ inside the Fourier integral is the analog of choosing the scale $\nu_0 = k_T|_+$ inside the convolutions building up the ν-evolution kernel. Similarly, choosing $\mu_0 = b_0/b_T$ for both beam and soft function inside the Fourier integral corresponds to choosing $\mu_0 = p_T|_+$ when solving the μ-RGE for $S(\vec{p}_T)$ and $B(\vec{p}_T)$, which then accordingly enters the convolutions $B \otimes B \otimes S$. In practice, the Fourier-space resummation is of course more easily implemented because the measurement constraint $\vec{q}_T = \sum_i \vec{k}_i$ is translated into a common impact parameter $\vec{b}_T$ for all convolved functions, and hence the iteratively defined convolution exponential in eq. (5.21) turns into a standard exponential function in Fourier space. As discussed on general grounds in sec. 2.5 and in detail for γ_ν in sec. 4.3, the formal difference between the different spaces arises from the fact that one uses different boundary conditions which induce different subleading terms to all orders in α_s. Hence with a robust estimate of theoretical uncertainties at any given order, one would expect that both techniques yield results compatible within their uncertainties, but a direct comparison of both results would provide another interesting way to assess theory uncertainties, in particular of nonperturbative effects.

7.1.2 Practical implementations

In the following we briefly comment on several implementations (without claiming to be exhaustive), which perform the resummation fully or partially in b space. As reference to compare to we take either the canonical b-space or the canonical momentum-space evolution we have derived, as both techniques reproduce *all* logarithms in the fixed-order reexpansion in their respective space. We only focus on the effects of deviating from these canonical scale choices and the strict resummation order. A detailed discussion regarding the phenomenologically important aspects of matching to the full fixed-order and assessing theory uncertainties can be found e.g. in ref. [85].

The original CSS formula is the basis of refs. [42–45] as well as refs. [46–55]. For the latter, the canonical Fourier-space logarithms are replaced by

$$\ln(b_T^2\mu^2/b_0^2) \to \ln(1 + b_T^2\mu^2/b_0^2)\,. \tag{7.8}$$

The benefit of shifting the argument of the logarithm is that it suppresses the region $p_T \sim 1/b_T \gg Q$. This should effectively suppress the contributions from energetic emissions, and it would be interesting to study its effect on the small q_T region compared to the strict canonical resummation. Furthermore it ensures that integrating over the $\vec{q}_T$-spectrum restores the inclusive cross section. The Landau pole intrinsic to the calculation is treated

using the minimal prescription [89, 90] by deforming the integration contour around the Landau pole, which is valid as long as q_T is sufficiently perturbative.

The first SCET-based calculation[7] was carried out in refs. [16–18] in the so-called collinear anomaly framework using an analytic regulator [64–66]. The resummation of rapidity logarithms there corresponds to performing the ν-evolution in Fourier space with fixed canonical scale choice $\nu_0 = b_0/b_T$. However, the low scale μ_0 is set to q_T, and the rapidity anomalous dimension γ_ν (corresponding to the anomaly coefficient F in their framework) is not resummed but expanded to a certain fixed order (referred to as ϵ-expansion). This looses some of the strict resummation accuracy, and in light of our discussion, one might expect that energetic emissions are incorrectly treated, which could affect the region of very small q_T. The naive divergence in the cross section (see sec. 3.4) is further avoided by choosing $\mu_0 = q_T + q_*$. The offset $q_* \approx 2$ GeV for Drell-Yan and $q_* \approx 8$ GeV for Higgs production is chosen large enough to explicitly avoid the divergence. However, it does so by essentially turning off the resummation below $q_T \lesssim q_*$, which then also drops the sensitivity to nonperturbative effects in the rapidity evolution kernel.

In refs. [56–58], a factorization theorem has been derived using the δ-regulator [56, 67], and has been applied e.g. in refs. [19, 20]. Here, the rapidity evolution is also performed in Fourier space, while the μ-RGEs are solved in momentum space, but in contrast to the above the rapidity anomalous dimension (in their notation the D function) is fully resummed. The low scale μ_0 is chosen as $\mu_0 = q_T + Q_0$, where $Q_0 \approx 2$ GeV for both Drell-Yan and Higgs production serves as a cutoff for the nonperturbative region. While at very small $q_T \lesssim Q_0$ deviations from the canonical resummation are expected, these can be effectively absorbed into the nonperturbative contributions that become relevant in this region. Indeed ref. [19] takes great care to assess nonperturbative effects in the Drell-Yan spectrum.

Lastly, the factorization theorem of refs. [15, 33] has been applied to Higgs production in ref. [85]. They employ a canonical resummation fully in b-space, similar to the CSS approach, but instead of shifting the arguments of the b-space logarithms, the resummation is turned off with profile scales in b-space whose form is based on the final value of q_T. The nonperturbative large b_T-region is avoided by explicitly cutting off the Fourier integration at $b_T \leq 2$ GeV^{-1}, while verifying that changing the cut in $1.5 - 3$ GeV^{-1} only produces a negligible variation. Because of the generic suppression of nonperturbative effects by $\Lambda_{\rm QCD}^2/q_T^2$, this can be expected to hold as long as q_T is sufficiently perturbative.

7.2 Early approaches for direct q_T-space resummation

There have been several attempts in the past to carry out the resummation directly in momentum space. They typically attempt to explicitly count logarithms $\ln(Q/q_T)$ in the q_T spectrum. As we discussed before, this is dangerous as it can easily lead one to discard apparent subleading contributions that are seemingly unimportant but are actually relevant. The original DDT approach [83, 84] introduced the LL cross section in the form of

---

[7]Earlier attempts of $\vec{q}_T$-resummation in SCET [91–93] missed the effects of rapidity divergences.

eq. (6.9),

$$\frac{\mathrm{d}\sigma^{(\mathrm{DDT})}}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} = \frac{\sigma_0}{\pi}\frac{\mathrm{d}}{\mathrm{d}q_T^2}f_a(\omega_a,q_T)\,f_b(\omega_b,q_T)\,e^{\mathbb{S}(Q,q_T)}\,. \tag{7.9}$$

Its connection to our full LL solution was already discussed in sec. 6.

Ref. [14] tried to extend eq. (7.9) to NLL, counting logarithms $L = \ln(Q/q_T)$ in the Sudakov exponent $\mathbb{S}$. In this counting, ref. [14] finds the cross section

$$\frac{\mathrm{d}\sigma^{(\mathrm{DDT})}}{\mathrm{d}Q^2\mathrm{d}Y\mathrm{d}\vec{q}_T} \propto Q^2\frac{\mathrm{d}}{\mathrm{d}q_T^2}\left[e^{\mathbb{S}}\frac{\Gamma(1+h/2)}{\Gamma(1-h/2)}\right], \qquad h = 4\Gamma_{\mathrm{cusp}}[\alpha_s(q_T)]\ln\frac{q_T}{Q}\,, \tag{7.10}$$

which diverges for $h = -2$. This can be directly related to our results, where the log-counting in the exponent corresponds to using the naive solution of the rapidity RGE in eq. (3.20) and keeping the rapidity anomalous dimension γ_ν at a fixed order rather than fully resuming it. As we saw in sec. 3.4, the result for the rapidity evolution kernel in this approximation [see eq. (3.27)] contains exactly the same spurious divergence at $\omega_s \equiv -h/2 = 1$ as eq. (7.10). As we have argued, this is caused by the incorrect treatment of energetic emissions, which become increasingly important for small q_T. This agrees with ref. [14], where it is remarked that energy conservation constraints are not correctly implemented in the resummation formula. This is remedied by basing the logarithmic order counting on the anomalous dimensions, and hence is not a flaw of the factorization theorem itself.

The second DDT-based approach, ref. [82], counts logarithms $L = \ln(Q/q_T)$ directly in the cross section, i.e. only counting $\alpha_s L^2 \sim 1$. Interestingly, this actually happens to lead to a numerically well-defined prediction of the cross section. To see this, consider again the naive rapidity evolution kernel eq. (3.20),

$$V(\vec{p}_T,\mu,\nu,\nu_0) = \delta(\vec{p}_T) + \sum_{n=1}^{\infty}\frac{1}{n!}\ln^n\frac{\nu}{\nu_0}(\gamma_\nu\otimes^n)(\vec{p}_T,\mu)\,. \tag{7.11}$$

Counting logarithms in the cross section is equivalent to truncating the n-fold convolution $\gamma_\nu\otimes^n$ at the desired accuracy. For example, at LL_σ where $\gamma_\nu(\vec{p}_T,\mu) = 2\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)]\mathcal{L}_0(\vec{p}_T,\mu)$, one would only keep the first term of the sum

$$(\gamma_\nu\otimes^n)(\vec{k}_T,\mu) = (2\Gamma_{\mathrm{cusp}})^n n\left[\mathcal{L}_{n-1}(\vec{k}_T,\mu) + 4\zeta_3\binom{n-1}{n-4}\mathcal{L}_{n-4}(\vec{k}_T,\mu) + \cdots\right], \tag{7.12}$$

but treat the first subleading term $\mathcal{L}_{n-4}$ as a $\mathrm{N^3LL}_\sigma$ correction. The evolution kernel would then be given by

$$V(\vec{p}_T,\mu,\nu,\nu_0) = \delta(\vec{p}_T) + \omega_s\mathcal{L}^{\omega_s}(\vec{p}_T,\mu)\,, \tag{7.13}$$

where $\omega_s = 2\Gamma_{\mathrm{cusp}}[\alpha_s(\mu)]\ln(\nu/\nu_0)$. Comparing to eqs. (3.27) and (7.10), where all terms in eq. (7.12) are kept, only keeping the first term in eq. (7.12) completely removes the divergence in the kernel. However, this cross-section counting is of course only applicable in an intermediate q_T range and only includes a small subset of logarithms compared to the full resummation.

A different approach was adopted in refs. [21–23], which attempt to obtain an explicit momentum-space expression of the inverse Fourier transform of the b-space result. The resummed b-space result $\mathrm{d}\tilde{\sigma}/\mathrm{d}b_T^2$ is then expanded in terms of logarithms $L_b = \ln(b_T^2\mu^2/b_0^2)$, whose inverse Fourier transform is known for arbitrary powers L_b^n, see appendix C.2. This allows one to construct a series in momentum space that approximates the Fourier transform to an in principle arbitrary precision. However, the resummation itself is effectively still performed canonically in b space.

7.3 Coherent branching formalism

Very recently, a different approach using the coherent branching formalism of refs. [25, 26] has been proposed in ref. [24], which is not based on a factorization theorem or solving the q_T evolution equations.

While a detailed comparison to the NNLL result given in ref. [24] would be very interesting, we leave it for future work and in the following concentrate on comparing to their NLL result, which is already instructive. The starting point of their derivation is the cumulative cross section obtained by summing over any number of independent emissions

$$\Sigma(q_T) = \int_0^{q_T} \mathrm{d}k_T \, \frac{\mathrm{d}\sigma(k_T)}{\mathrm{d}k_T} = \sigma_0 \int_0^\infty \langle \mathrm{d}k_1 \rangle \, R'(k_1)\, e^{-R(\epsilon k_1)}$$
$$\times \sum_{n=0}^\infty \frac{1}{n!} \prod_{i=2}^{n+1} \int_{\epsilon k_1}^{k_1} \langle \mathrm{d}k_i \rangle \, R'(k_i)\, \theta\left(|\vec{q}_T| - \left| \sum_j \vec{k}_j \right| \right), \quad (7.14)$$

where the phase-space measure is $\langle \mathrm{d}k_T \rangle = \frac{\mathrm{d}k_T}{k_T}\frac{\mathrm{d}\phi}{2\pi}$. The differential spectrum follows to be

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\vec{q}_T} = \sigma_0 \int \mathrm{d}^2\vec{k}_1 \, \frac{R'(k_1)}{2\pi k_1^2} e^{-R(\epsilon k_1)} \sum_{n=0}^\infty \frac{1}{n!} \prod_{i=2}^{n+1} \int_{\epsilon k_1 < |\vec{k}_i| < k_1} \mathrm{d}^2\vec{k}_i \, \frac{R'(k_i)}{2\pi k_i^2}\, \delta\left(\vec{q}_T - \sum_j \vec{k}_j \right), \quad (7.15)$$

where we use a notation resembling our convolution notation. Here, the hardest emission $\vec{k}_1$ has been singled out. The parameter $\epsilon \ll 1$ reflects that emissions with $k_i < \epsilon k_1$ are unresolved. Correspondingly, the exponential $e^{-R(\epsilon k_1)}$ encodes the Sudakov suppression of having no emission between scales ϵk_1 and Q. The radiator R is given at NLL by

$$R(k_T) = \int_{k_T}^Q \frac{\mathrm{d}\mu'}{\mu'} \left(4\Gamma_{\mathrm{cusp}}[\alpha_s(\mu')] \ln \frac{Q}{\mu'} - \frac{\alpha_s(\mu')}{\pi} \beta_0 \right) = \int_{k_T}^Q \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q, \mu'), \quad (7.16)$$

where we converted from the Catani-Marchesini-Webber scheme [94] used in ref. [24] to the $\overline{\mathrm{MS}}$-scheme. Its derivative R' evaluates to[8]

$$R'(k_T) = -k_T \frac{\mathrm{d}R(k_T)}{\mathrm{d}k_T} = 4\Gamma_{\mathrm{cusp}}[\alpha_s(k_T)] \ln \frac{Q}{k_T}. \quad (7.17)$$

---

[8]The results in eqs. (7.14) and (7.15) are technically only NLL accurate up to the fact that the PDFs in σ_0 are evaluated at fixed μ_F rather than k_1, which however is irrelevant for the present discussion. We have dropped the constant $\beta_0\alpha_s/\pi$ here, which would be canceled if the PDFs were evaluated at k_1.

The structure of eq. (7.15) is closely related to our results, as it essentially contains an infinite number of convolutions, with the exception that the hardest emission is explicitly singled out. Furthermore the unresolved regions $0 < k_i < \epsilon k_1$ that are cut out of the k_i integrals ($i \geq 2$) are already captured in the Sudakov exponent. Letting $\epsilon \to 0$ is then equivalent to the cancellation of IR singularities encoded in the plus distributions in our framework. Each factor R' corresponds to a ν-anomalous dimension times its rapidity logarithm,

$$\frac{R'(k_i)}{2\pi k_i^2} = \frac{4\Gamma_{\rm cusp}[\alpha_s(k_i)]\ln(Q/k_i)}{2\pi k_i^2} = \int_{k_i}^Q \frac{{\rm d}\nu'}{\nu'}\frac{4\Gamma_{\rm cusp}[\alpha_s(k_i)]}{2\pi k_i^2} = \int_{k_i}^Q \frac{{\rm d}\nu'}{\nu'}\,\gamma_\nu(\vec{k}_i, \mu)\,, \quad (7.18)$$

where we assumed $k_i > 0$ to drop the plus prescription in γ_ν. In summary, eq. (7.15) agrees very well with our momentum-space resummed cross section that follows from the factorization theorem, up to the different treatment of the cancellation of IR divergences.

Differences arise when the resummed spectrum is further expanded to obtain NLL accuracy, defined by counting logarithms in the cumulative cross section. Expanding the individual emission momenta around $k_i \sim q_T$ and counting logarithms $\ln(Q/q_T)$, they also reproduce the spurious divergence in the cross section. This is precisely equivalent to using the naive rapidity resummation, which does not treat energetic emissions correctly, as discussed in secs. 3.2 and 3.4.

To circumvent this, ref. [24] instead expands all radiators appearing in eq. (7.14) around the hardest emission k_1. At NLL, the necessary expansions are

$$R(\epsilon k_1) = R(k_1) + R'(k_1)\ln\frac{1}{\epsilon} + \cdots\,, \quad R'(k_i) = R'(k_1) + \cdots\,. \quad (7.19)$$

A simplification is that all radiators $R(k_i)$ and $R'(k_i)$ are now evaluated at $\alpha_s(k_1)$ rather than $\alpha_s(k_i)$, thereby removing the nonperturbative effects that would otherwise be present in the rapidity evolution. This procedure hence fundamentally resums logarithms of $\ln(Q/k_1)$ to NLL, where the resummation accuracy is defined by explicitly counting logarithms in the cumulative cross section (using exponent counting). Ref. [24] then argues that the formal accuracy in terms of counting logarithms $\ln(Q/q_T)$ will be the same, and only differ by subleading terms from the naive result.

Compared to our exact solution, this procedure effectively corresponds to approximating the rapidity logarithms $\ln(\nu/k_i)$ in the exact rapidity evolution kernel by logarithms $\ln(\nu/k_1)$, which allows to pull them out of the convolutions. Since this avoids the spurious singularity, one might expect that this approximation is safer than the naive one of taking $\ln(\nu/k_i) \sim \ln(\nu/q_T)$. On the other hand, these differences are all of apparent subleading nature, and it would be interesting to study in more detail to what extent this approach reproduces the subleading terms in the q_T spectrum that are included in the strict LL and NLL evolution in either momentum or Fourier space.

8 Conclusion

We have investigated solving differential equations for arbitrary distributions. These arise naturally in differential spectra of observables resolving additional soft or collinear QCD

radiation, where the cancellation of IR singularities is encoded through plus distributions, and the all-order logarithmic structure is fully encoded in (renormalization group) evolution equations. Solving these equations allows the resummation of the logarithmic distributions to all orders, provided that the boundary term in the solution is free of logarithms, such that it can be reliably calculated in fixed-order perturbation theory.

We have introduced a technique for distributional scale setting $\mu = k|_+$ that allows one to treat logarithmic distributions like ordinary logarithms. In particular, it eliminates any logarithms contained in the boundary condition of plus distributions such as $[\mu/k]_+^\mu$. It can be straightforwardly applied to solve distributional differential equations for both one-dimensional and two-dimensional distributions, where it ensures that the appearing boundary term is free of any logarithmic distributions. This allows one to perform the RG evolution and resummation directly in distribution (momentum) space. It thus enables the implementation of profile scales, the transition and matching to the full fixed-order distribution, and the estimation of perturbative uncertainties through scale variations directly in distribution space.

The technique has been applied to obtain the resummation of the transverse momentum (q_T) spectrum for the first time by solving the associated evolution equations in momentum space. We showed that a well-known spurious singularity in the spectrum arises from wrong scale setting, i.e. wrong boundary terms in the RG evolution, causing an incorrect treatment of energetic emissions, and which is cured by a proper distributional scale setting. This yields a well-defined resummation of the q_T spectrum, whose resummation accuracy is strictly defined by the perturbative expansion of the associated anomalous dimensions (and boundary terms) without requiring to count explicit powers of logarithms. We indeed find that trying to specify the logarithmic accuracy by explicitly counting logarithms $\ln(Q/q_T)$ in the spectrum can be ill defined, and partly is the reason for the spurious divergences encountered in previous attempts to perform resummation in momentum space.

Previous attempts at a momentum-space resummation were partially motivated by trying to avoid nonperturbative effects in the q_T spectrum, which are unavoidable in the Fourier space resummation due to integrating $\alpha_s(1/b_T)$ over its Landau pole. We find that analogous nonperturbative effects also appear in the rapidity evolution kernel in the strict momentum-space resummation, because real emissions with momentum $\vec{k}_i$ naturally scale with $\alpha_s(k_i)$, which is necessary to suppress energetic emissions. We discussed how the nonperturbative contributions to the rapidity anomalous dimension can be isolated in momentum space, which closely reproduces the common treatment in Fourier space. We also showed that nonperturbative effects arising from integrating over the $\vec{k}_i \to 0$ region inside convolutions are generically suppressed as $\Lambda_{\mathrm{QCD}}^2/q_T^2$, such that they do not spoil the predictivity of the resummation for perturbative q_T.

The correct momentum-space rapidity evolution involves an intricate iterative convolution structure. Its numerical implementation is nontrivial, which we plan to address in future work. The rapidity evolution has so far been performed in Fourier b_T space, where logarithms $\ln(Qb_T)$ rather than $\ln(Q/q_T)$ are resummed. While both approaches are formally equivalent, our general analysis shows that the boundary conditions employed in the

evolution intrinsically differ to all orders. In the future, it would be interesting to compare them numerically, as they probe different subleading terms to all orders, and this could also provide new insight into nonperturbative effects.

Acknowledgments

A   Notation and conventions

Most of this paper is devoted to transverse momentum dependent distributions, but sec. 2 also discusses one-dimensional distributions. To distinguish the two cases, the momentum-space arguments are typically called k for the one-dimensional case and $\vec{p}_T, \vec{k}_T, \vec{q}_T$ for the two-dimensional case. The subscript $_T$ is always added to make the distinction clear. For magnitudes of vectors, we drop the explicit vector sign and simply write $|\vec{p}_T|^2 = p_T^2$ etc.

A.1   Fourier transformations

By default we work in distribution space. The conjugate functions in Fourier space are always denoted with a tilde. The Fourier conjugate variable in the one-dimensional case is typically called y, while for the two-dimensional case it is called $\vec{b}_T$. Our conventions for the Fourier transformation for the one-dimensional case are

$$f(k) = \int \frac{\mathrm{d}y}{2\pi}\, e^{+ixy} \tilde{f}(y)\,, \tag{A.1}$$

$$\tilde{f}(y) = \int \mathrm{d}k\, e^{-iky} f(k)\,, \tag{A.2}$$

and for the two-dimensional case

$$f(\vec{p}_T) = \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2}\, e^{+i\vec{p}_T\cdot\vec{b}_T}\, \tilde{f}(\vec{b}_T)\,, \tag{A.3}$$

$$\tilde{f}(\vec{b}_T) = \int \mathrm{d}^2\vec{p}_T\, e^{-i\vec{p}_T\cdot\vec{b}_T} f(\vec{p}_T)\,. \tag{A.4}$$

For azimuthally symmetric functions, $f(\vec{p}_T) \equiv f(p_T)$, the latter simplify to

$$f(p_T) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}b_T\, b_T J_0(b_T p_T)\, \tilde{f}(b_T)\,, \tag{A.5}$$

$$\tilde{f}(b_T) = 2\pi \int_0^\infty \mathrm{d}p_T\, p_T J_0(b_T p_T) f(p_T)\,, \tag{A.6}$$

where $J_0(x)$ is the 0th-order Bessel function.

A.2 Convolutions

One-dimensional convolutions are defined as

$$(f \otimes g)(k, \ldots) \equiv \int \mathrm{d}k' \, f(k - k', \ldots) \, g(k', \ldots), \tag{A.7}$$

where the dots stand for possible additional arguments of the functions. Multiple convolutions are abbreviated as

$$(f \otimes^n)(k) \equiv \int \mathrm{d}k_1 \ldots \mathrm{d}k_n \, f(k_1) \ldots f(k_n) \, \delta(k - k_1 - \cdots - k_n). \tag{A.8}$$

Two-dimensional convolutions are defined as

$$
\begin{aligned}
(f \otimes g)(\vec{p}_T, \ldots) &= \int \mathrm{d}^2\vec{k}_1 \mathrm{d}^2\vec{k}_2 \, f(\vec{k}_1, \ldots) \, g(\vec{k}_2, \ldots) \, \delta(\vec{p}_T - \vec{k}_1 - \vec{k}_2) \\
&= \int \mathrm{d}^2\vec{k}_T \, f(\vec{p}_T - \vec{k}_T, \ldots) \, g(\vec{k}_T, \ldots),
\end{aligned}
\tag{A.9}
$$

where the dots stand again for possible additional arguments of the functions. Multiple convolutions are abbreviated as

$$(f \otimes^n)(\vec{p}_T) \equiv \int \mathrm{d}^2\vec{k}_1 \cdots \mathrm{d}^2\vec{k}_n \, f(\vec{k}_1) \ldots f(\vec{k}_n) \, \delta(\vec{p}_T - \vec{k}_1 - \cdots - \vec{k}_n). \tag{A.10}$$

A.3 Fixed-order perturbative expansions

We make frequent use of fixed-order expansions in α_s. The expansion coefficients of beta function and cusp anomalous dimension are defined as

$$\frac{\mathrm{d}\alpha_s}{\mathrm{d}\ln\mu} = \beta(\alpha_s) = -2\alpha_s \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}, \qquad \Gamma_{\mathrm{cusp}}(\alpha_s) = \sum_{n=0}^{\infty} \Gamma_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}. \tag{A.11}$$

Similarly, the constant noncusp pieces of all anomalous dimensions are expanded as

$$\gamma(\alpha_s) = \sum_{n=0}^{\infty} \gamma_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}. \tag{A.12}$$

The soft function, and analogously other functions, are expanded as

$$S(\vec{p}_T, \mu, \nu) = \sum_{n=0}^{\infty} S^{(n)}(\vec{p}_T, \mu, \nu) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^n. \tag{A.13}$$

B One-dimensional plus distributions

For completeness we collect and extend the definitions and formulas for one-dimensional plus distributions from refs. [27, 95].

B.1 Definition

For a function $g(x)$ that has support for $x \geq 0$ and diverges less than $1/x^2$ for $x \to 0$, the defining properties of its plus distributions with boundary condition $x_0 > 0$ are

$$\left[\theta(x)g(x)\right]_+^{x_0} = \theta(x)\,g(x) \qquad \text{for } x \neq 0 , \tag{B.1}$$

$$\int^{x_0} \mathrm{d}x' \left[\theta(x')g(x')\right]_+^{x_0} = 0 . \tag{B.2}$$

The lower limit of integration is kept implicit and formally has to include the singularity at $x = 0$. An explicit definition is given by [27]

$$\begin{aligned}
\left[g(x)\right]_+^{x_0} &\equiv \left[\theta(x)g(x)\right]_+^{x_0} \\
&= \lim_{\epsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}x}\left[\theta(x-\epsilon)\int_{x_0}^{x} \mathrm{d}x'\, g(x')\right] \\
&= \lim_{\epsilon \to 0}\left[\theta(x-\epsilon)g(x) + \delta(x-\epsilon)\int_{x_0}^{x} \mathrm{d}x'\, g(x')\right] .
\end{aligned} \tag{B.3}$$

Equivalently, we have

$$\begin{aligned}
\lim_{\epsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}x}\left[\theta(x-\epsilon)\,G(x)\right] &= \lim_{\epsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}x}\left[\theta(x-\epsilon)\big(G(x) - G(x_0) + G(x_0)\big)\right] \\
&= \left[\theta(x)\frac{\mathrm{d}G(x)}{\mathrm{d}x}\right]_+^{x_0} + \delta(x)\,G(x_0) .
\end{aligned} \tag{B.4}$$

From the above definitions, it follows that the boundary condition can be shifted using

$$\left[\theta(x)g(x)\right]_+^{x_0} = \left[\theta(x)g(x)\right]_+^{x_1} + \delta(x)\int_{x_0}^{x_1} \mathrm{d}x'\, g(x') . \tag{B.5}$$

The derivative with respect to the boundary value is thus given by

$$\frac{\mathrm{d}}{\mathrm{d}x_0}\left[\theta(x)g(x)\right]_+^{x_0} = -g(x_0)\,\delta(x) . \tag{B.6}$$

More generally, if g itself depends on x_0, we have

$$\frac{\mathrm{d}}{\mathrm{d}x_0}\left[\theta(x)g(x,x_0)\right]_+^{x_0} = \left[\frac{\mathrm{d}}{\mathrm{d}x_0}g(x,x_0)\right]_+^{x_0} - g(x_0,x_0)\,\delta(x) . \tag{B.7}$$

Following ref. [27], we denote the standard plus distributions with boundary condition $x_0 = 1$ as

$$\mathcal{L}_n(x) = \left[\frac{\theta(x)\ln^n x}{x}\right]_+^1 , \qquad \mathcal{L}^a(x) = \left[\frac{\theta(x)}{x^{1-a}}\right]_+^1 . \tag{B.8}$$

In addition, for dimensionful arguments, we define

$$\mathcal{L}_n(k,\mu) \equiv \frac{1}{\mu}\mathcal{L}_n\!\left(\frac{k}{\mu}\right) \equiv \left[\frac{\theta(k)}{k}\ln^n \frac{k}{\mu}\right]_+^\mu , \qquad \mathcal{L}^a(k,\mu) \equiv \frac{1}{\mu}\mathcal{L}^a\!\left(\frac{k}{\mu}\right) \equiv \left[\frac{\theta(k)}{k}\left(\frac{k}{\mu}\right)^a\right]_+^\mu \tag{B.9}$$

They are related by

$$\mathcal{L}_n(x) = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\mathcal{L}^a(x)\bigg|_{a=0} , \qquad \mathcal{L}_n(k,\mu) = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\mathcal{L}^a(k,\mu)\bigg|_{a=0} . \tag{B.10}$$

$\mathcal{L}_n(k,\mu)$	FT$[\mathcal{L}_n(k,\mu)]$
$\mathcal{L}_0(k,\mu)$	$-L_y$
$\mathcal{L}_1(k,\mu)$	$+\frac{1}{2}L_y^2 + \frac{\pi^2}{12}$
$\mathcal{L}_2(k,\mu)$	$-\frac{1}{3}L_y^3 - \frac{\pi^2}{6}L_y - \frac{2}{3}\zeta_3$
$\mathcal{L}_3(k,\mu)$	$+\frac{1}{4}L_y^4 + \frac{\pi^2}{4}L_y^2 + 2\zeta_3 L_y + \frac{3}{80}\pi^4$
$\mathcal{L}_4(k,\mu)$	$-\frac{1}{5}L_y^5 - \frac{\pi^2}{3}L_y^3 - 4\zeta_3 L_y^2 - \frac{3}{20}\pi^4 L_y - \left(\frac{2}{3}\pi^2\zeta_3 + \frac{24}{5}\zeta_5\right)$
$\mathcal{L}_5(k,\mu)$	$+\frac{1}{6}L_y^6 + \frac{5}{12}\pi^2 L_y^4 + \frac{20}{3}\zeta_3 L_y^3 + \frac{3}{8}\pi^4 L_y^2 + \left(\frac{10}{3}\pi^2 + 24\zeta_5\right)L_y + \left(\frac{61}{1008}\pi^6 + \frac{20}{3}\zeta_3^2\right)$

Table 1. Fourier transform of $\mathcal{L}_n(k,\mu)$ for $n \leq 5$. Results are expressed in terms of $L_y = \ln(iy\mu e^{\gamma_E})$. See eq. (B.13).

L_y^n	FT$^{-1}[L_y^n]$
1	$\delta(k)$
L_y	$-\mathcal{L}_0(k,\mu)$
L_y^2	$2\mathcal{L}_1(k,\mu) - \frac{\pi^2}{6}\delta(k)$
L_y^3	$-3\mathcal{L}_2(k,\mu) + \frac{\pi^2}{2}\mathcal{L}_0(k,\mu) - 2\zeta_3\delta(k)$
L_y^4	$4\mathcal{L}_3(k,\mu) - 2\pi^2\mathcal{L}_1(k,\mu) + 8\zeta_3\mathcal{L}_0(k,\mu) + \frac{\pi^4}{60}\delta(k)$
L_y^5	$-5\mathcal{L}_4(k,\mu) + 5\pi^2\mathcal{L}_2(k,\mu) - 40\zeta_3\mathcal{L}_1(k,\mu) - \frac{\pi^4}{12}\mathcal{L}_0(k,\mu) + \left(\frac{10}{3}\pi^2\zeta_3 - 24\zeta_5\right)\delta(k)$
L_y^6	$6\mathcal{L}_5(k,\mu) - 10\pi^2\mathcal{L}_3(k,\mu) + 120\zeta_3\mathcal{L}_2(k,\mu) + \frac{\pi^4}{2}\mathcal{L}_1(k,\mu) + (144\zeta_5 - 20\pi^2\zeta_3)\mathcal{L}_0(k,\mu)$ $+\left(40\zeta_3^2 - \frac{5}{168}\pi^6\right)\delta(k)$

Table 2. Fourier transform of $L_y^n = \ln^n(iy\mu e^{\gamma_E})$ for $n \leq 6$. See eq. (B.14).

B.2 Fourier transformation

The Fourier transformation of a plus function $\mathcal{L}^a(k,\mu)$ with respect to $y = y - i0$ is given by

$$\int \mathrm{d}k\, e^{-iky}\mathcal{L}^a(k,\mu) = \frac{1}{a}\left[(i\mu y)^{-a}\Gamma(1+a) - 1\right] = \frac{1}{a}\left[e^{-aL_y}R_1(a) - 1\right], \tag{B.11}$$

which holds for $a > -1$ through analytic continuation. In the second step we introduced the abbreviations

$$L_y = \ln(iy\mu e^{\gamma_E}), \qquad R_1(a) = e^{\gamma_E a}\Gamma(1+a). \tag{B.12}$$

$R_1^{(n)}$	Exact value	Numerical value
$R_1^{(0)}$	1	1
$R_1^{(1)}$	0	0
$R_1^{(2)}$	$\frac{\pi^2}{6}$	$\approx 1.64 = 0.822 \times 2!$
$R_1^{(3)}$	$-2\zeta_3$	$\approx -2.40 = -0.401 \times 3!$
$R_1^{(4)}$	$\frac{3\pi^2}{20}$	$\approx 14.6 = 0.609 \times 4!$
$R_1^{(5)}$	$-\frac{10\pi^2}{3}\zeta_3 - 24\zeta_5$	$\approx -64.4 = -0.537 \times 5!$
$R_1^{(6)}$	$\frac{61\pi^6}{168} + 40\zeta_3^2$	$\approx 406.9 = 0.565 \times 6!$
$R_1^{(7)}$	$-\frac{21\pi^4}{2}\zeta_3 - 84\pi^2\zeta_5 - 720\zeta_7$	$\approx -2815 = -0.559 \times 7!$

Table 3. The first values of $R_1^{(n)}$, defined in eq. (B.15).

$\tilde{R}_1^{(n)}$	Exact value	Numerical value
$\tilde{R}_1^{(0)}$	1	1
$\tilde{R}_1^{(1)}$	0	0
$\tilde{R}_1^{(2)}$	$-\frac{\pi^2}{6}$	≈ -1.64
$\tilde{R}_1^{(3)}$	$-2\zeta_3$	≈ -2.40
$\tilde{R}_1^{(4)}$	$\frac{\pi^4}{60}$	≈ 1.62
$\tilde{R}_1^{(5)}$	$\frac{10\pi^2}{3}\zeta_3 - 24\zeta_5$	≈ 14.7
$\tilde{R}_1^{(6)}$	$-\frac{5\pi^6}{168} + 40\zeta_3^2$	≈ 29.2
$\tilde{R}_1^{(7)}$	$-\frac{7\pi^4}{6}\zeta_3 + 84\pi^2\zeta_5 - 720\zeta_7$	≈ -2.96

Table 4. The first values of $\tilde{R}_1^{(n)}$, defined in eq. (B.16).

An explicit form of the Fourier transform for plus functions $\mathcal{L}_n(k,\mu)$ follows from eq. (B.10),

$$
\int \mathrm{d}k \, e^{-iky} \mathcal{L}_n(k,\mu) = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\bigg|_{a=0} \frac{1}{a}\left[e^{-aL_y}R_1(a) - 1\right]
$$
$$
= \frac{1}{n+1}\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}L_y^k \, R_1^{(n+1-k)}, \qquad (\text{B.13})
$$

and the inverse is given by [95]

$$\int \frac{\mathrm{d}y}{2\pi} e^{iky} L_y^n = \frac{\mathrm{d}^n}{\mathrm{d}a^n} \left\{ \frac{-a}{R_1(-a)} \mathcal{L}^{-a}(k) + \frac{\delta(k)}{R_1(-a)} \right\}_{a=0}$$

$$= \sum_{k=0}^{n-1} (-1)^{k+1} n \binom{n-1}{k} \tilde{R}_1^{(n-k-1)} \mathcal{L}_k(k,\mu) + \tilde{R}_1^{(n)} \delta(k) , \tag{B.14}$$

where the quantities $R_1^{(n)}, \tilde{R}_1^{(n)}$ occurring in the coefficients are defined as

$$R_1^{(n)} = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\Big|_{a=0} e^{\gamma_E a} \Gamma(1+a) , \tag{B.15}$$

$$\tilde{R}_1^{(n)} = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\Big|_{a=0} \frac{e^{\gamma_E a}}{\Gamma(1-a)} . \tag{B.16}$$

They fulfill the sum rule

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k R_1^{(n-k)} \tilde{R}_1^{(k)} = 0 . \tag{B.17}$$

The values can be conveniently obtained from the expansions

$$\sum_{n=0}^{\infty} \frac{R_1^{(n)} x^n}{n!} = \exp\left[+ \sum_{n=1}^{\infty} \frac{\zeta(n+1)}{n+1} (-x)^{n+1} \right] , \tag{B.18}$$

$$\sum_{n=0}^{\infty} \frac{\tilde{R}_1^{(n)} x^n}{n!} = \exp\left[- \sum_{n=1}^{\infty} \frac{\zeta(n+1)}{n+1} x^{n+1} \right] . \tag{B.19}$$

Numerically we find the asymptotic behavior

$$R_1^{(n)} \approx 0.561 \times (-1)^n n! \qquad \text{for } n \gg 1 . \tag{B.20}$$

The factorial behavior of $R_1^{(n)}$ reflects that the Taylor series of $R_1(a) = e^{\gamma_E a} \Gamma(1+a)$ only converges for $|a| < 1$. For $R_1^{(n)}$ there is no such simple asymptotic behavior, but due to the infinite radius of convergence of $\tilde{R}_1(a) = e^{\gamma_E a}/\Gamma(1-a)$, it is expected to grow much less severely. This is confirmed by the first values shown in 4.

For illustration, table 1 shows the Fourier transforms of $\mathcal{L}_n(k,\mu)$ for $n \leq 5$, table 2 the Fourier transforms of L_y^n for $n \leq 6$. The first few $R_1^{(n)}$ and $\tilde{R}_1^{(n)}$ are given in tables 3 and 4.

B.3 Convolutions

The convolution of two plus distributions is given by

$$\int \mathrm{d}x' \, \mathcal{L}^a(x-x') \, \mathcal{L}^b(x') = \left[\mathcal{L}^{a+b}(x) + \frac{\delta(x)}{a+b} \right] V_1(a,b)$$

$$+ \left(\frac{1}{a} + \frac{1}{b} \right) \mathcal{L}^{a+b}(x) - \frac{1}{b} \mathcal{L}^a(x) - \frac{1}{a} \mathcal{L}^b(x) , \tag{B.21}$$

where $V_1(a,b)$ is defined by

$$V_1(a,b) = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)} - \frac{1}{a} - \frac{1}{b} , \tag{B.22}$$

which satisfies $V_1(0,0) = 0$. Taking derivatives with respect to a and b we can get the corresponding formulas for convolutions of the form $\mathcal{L}_n \otimes \mathcal{L}^a$ and $\mathcal{L}_m \otimes \mathcal{L}_n$. The explicit results can be found in Appendix B of ref. [27].

An important special case are multiple convolutions of $\mathcal{L}_n$, which is given by

$$(\mathcal{L}_0 \otimes^n)(k, \mu) = (-1)^n \sum_{i=0}^{n-1} (-1)^{i+1} n \binom{n-1}{i} \tilde{R}_1^{(n-i-1)} \mathcal{L}_i(k, \mu) + (-1)^n \tilde{R}_1^{(n)} \delta(k) \,. \quad \text{(B.23)}$$

C Two-dimensional plus distributions

C.1 Definition

Plus distributions in momentum space are uniquely defined by the two conditions

$$\left[g(\vec{p}_T)\right]_+^\mu = g(\vec{p}_T) \qquad \text{for } |\vec{p}_T| > 0 \qquad \text{(C.1)}$$

$$\int_{|\vec{p}_T| \leq \mu} \mathrm{d}^2 \vec{p}_T \left[g(\vec{p}_T)\right]_+^\mu = 0 \,, \qquad \text{(C.2)}$$

where $g(\vec{p}_T)$ diverges at most as $1/p_T^2$ for $p_T \to 0$. The boundary condition can be shifted using

$$\left[g(\vec{p}_T)\right]_+^\mu = \left[g(\vec{p}_T)\right]_+^\xi \mp \delta(\vec{p}_T) \int_{\xi < |\vec{k}_T| < \mu} \mathrm{d}^2 \vec{k}_T \, g(\vec{k}_T) \,, \qquad \text{(C.3)}$$

where the $-$ sign holds for $\xi < \mu$ and the $+$ sign for $\xi > \mu$. For azimuthally symmetric functions, $g(\vec{p}_T) \equiv g(|\vec{p}_T|)$, this simplifies to

$$\left[g(|\vec{p}_T|)\right]_+^\mu = \left[g(|\vec{p}_T|)\right]_+^\xi - 2\pi \delta(\vec{p}_T) \int_\xi^\mu \mathrm{d}k_T \, k_T \, g(k_T) \,. \qquad \text{(C.4)}$$

It is important to keep in mind that it is nevertheless defined as a two-dimensional distribution, even though the distribution effectively contains a scalar function.

It follows that the derivative with respect to the boundary condition is given by

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \left[g(|\vec{p}_T|)\right]_+^\mu = -2\pi \mu^2 g(\mu) \, \delta(\vec{p}_T) \,. \qquad \text{(C.5)}$$

More generally, if g itself depends on μ, then

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \left[g(|\vec{p}_T|, \mu)\right]_+^\mu = -2\pi \mu^2 g(\mu, \mu) \, \delta(\vec{p}_T) + \left[\mu \frac{\mathrm{d}g(|\vec{p}_T|, \mu)}{\mathrm{d}\mu}\right]_+^\mu \,. \qquad \text{(C.6)}$$

For a scalar input function $f(p_T)$, the azimuthally symmetric two-dimensional plus distribution can be defined by the derivative

$$\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \theta(p_T) = \delta(\vec{p}_T) \,, \qquad \int_{|\vec{k}_T| \leq p_T} \mathrm{d}^2 \vec{k}_T \, \delta(\vec{k}_T) = \theta(p_T) \,,$$

$$\frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \left[\theta(p_T) f(p_T)\right] = \left[\frac{1}{2\pi p_T} \frac{\mathrm{d}f(p_T)}{\mathrm{d}p_T}\right]_+^\xi + \delta(\vec{p}_T) f(\xi) \,. \qquad \text{(C.7)}$$

Note that the appearance of the θ-function is crucial to render the derivative well defined in form of a plus distribution.

We define the frequently occurring logarithmic plus distributions as

$$\mathcal{L}^a(\vec{p}_T, \mu) \equiv \frac{1}{\pi\mu^2}\left[\left(\frac{\vec{p}_T^2}{\mu^2}\right)^{a-1}\right]_+^\mu \equiv \frac{1}{\pi\mu^2}\mathcal{L}^a\left(\frac{\vec{p}_T^2}{\mu^2}\right), \tag{C.8}$$

$$\mathcal{L}_n(\vec{p}_T, \mu) \equiv \frac{1}{\pi\mu^2}\left[\frac{\mu^2}{\vec{p}_T^2}\ln^n\frac{\vec{p}_T^2}{\mu^2}\right]_+^\mu \equiv \frac{1}{\pi\mu^2}\mathcal{L}_n\left(\frac{\vec{p}_T^2}{\mu^2}\right), \tag{C.9}$$

where the $\mathcal{L}^a(x)$ and $\mathcal{L}_n(x)$ on the right-hand side are the standard ones for dimensionless arguments with boundary condition $x_0 = 1$ as defined in eq. (B.8). They are related by

$$\mathcal{L}_n(\vec{p}_T, \mu) = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\mathcal{L}^a(\vec{p}_T, \mu)\bigg|_{a=0}. \tag{C.10}$$

Their cumulant and inverse cumulants are

$$\int_{|\vec{k}_T|\leq p_T}\mathrm{d}^2\vec{k}_T\,\mathcal{L}_n(\vec{k}_T, \mu) = \frac{\theta(p_T)}{n+1}\ln^{n+1}\frac{p_T^2}{\mu^2},$$

$$\frac{1}{2\pi p_T}\frac{\mathrm{d}}{\mathrm{d}p_T}\left[\theta(p_T)\ln^n\frac{p_T^2}{\mu^2}\right] = n\,\mathcal{L}_{n-1}(\vec{p}_T, \mu) \qquad \text{for } n \geq 1. \tag{C.11}$$

Finally note that our definition of $\mathcal{L}_n(\vec{p}_T, \mu)$ is related to that in ref. [15] by

$$\mathcal{L}_n(\vec{p}_T, \mu) = 2(-1)^n\mathcal{L}_n^0(\mu, \vec{p}_T; \mu), \tag{C.12}$$

where $\mathcal{L}_n^0(\mu, \vec{p}_T; \mu)$ is defined in eq. (F.1) in ref. [15].

C.2   Fourier transformation

The Fourier transform of $\mathcal{L}^a(\vec{p}_T, \mu)$ is given by

$$\int\mathrm{d}^2\vec{p}_T\,e^{-i\vec{p}_T\cdot\vec{b}_T}\mathcal{L}^a(\vec{p}_T, \mu) = \frac{1}{a}\left[\left(\frac{b_T^2\mu^2}{4}\right)^{-a}\frac{\Gamma(1+a)}{\Gamma(1-a)} - 1\right]. \tag{C.13}$$

It is convenient to express the Fourier transform of $\mathcal{L}_n(\vec{p}_T, \mu)$ as polynomials of

$$L_b \equiv \ln\frac{b_T^2\mu^2}{b_0^2}, \qquad b_0 = 2e^{-\gamma_E} \approx 1.12291\ldots. \tag{C.14}$$

The Fourier transform of $\mathcal{L}_n(\vec{p}_T, \mu)$ then follows from eq. (C.13) using eq. (C.10),

$$\int\mathrm{d}^2\vec{p}_T\,e^{-i\vec{p}_T\cdot\vec{b}_T}\mathcal{L}_n(\vec{p}_T, \mu) = \frac{\mathrm{d}^n}{\mathrm{d}a^n}\frac{1}{a}\left[\left(\frac{b_T^2\mu^2}{4}\right)^{-a}\frac{\Gamma(1+a)}{\Gamma(1-a)} - 1\right]\bigg|_{a=0}$$

$$= \frac{1}{n+1}\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}R_2^{(n+1-k)}L_b^k. \tag{C.15}$$

The inverse transformation is given by

$$\int \frac{\mathrm{d}^2 \vec{b}_T}{(2\pi)^2} e^{i\vec{p}_T \cdot \vec{b}_T} L_b^n = \frac{\mathrm{d}^n}{\mathrm{d}a^n} \left[-aR_2(a)\mathcal{L}^{-a}(\vec{p}_T, \mu) + R_2(a)\delta(\vec{p}_T) \right]\Big|_{a=0}$$

$$= \sum_{k=0}^{n-1} (-1)^{k+1} n \binom{n-1}{k} R_2^{(n-k-1)} \mathcal{L}_k(\vec{p}_T, \mu) + R_2^{(n)}\delta(\vec{p}_T). \qquad (C.16)$$

The constants $R_2^{(n)}$ occurring in the coefficients are defined as

$$R_2^{(n)} = \frac{\mathrm{d}^n}{\mathrm{d}a^n} e^{2\gamma_E a} \frac{\Gamma(1+a)}{\Gamma(1-a)}\Big|_{a=0}. \qquad (C.17)$$

They fulfill the useful property

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} R_2^{(k)} R_2^{(n-k)} = 0. \qquad (C.18)$$

By integrating eq. (C.16) over $|\vec{p}_T| \leq \mu$, it also follows that

$$R_2^{(n)} = \int_0^\infty \mathrm{d}x \, J_1(x) \ln^n \frac{x^2}{b_0^2}. \qquad (C.19)$$

They can also be obtained from the relation [21]

$$\sum_{n=0}^{\infty} \frac{R_2^{(n)}}{n!} x^n = \exp\left[-2 \sum_{n=1}^{\infty} \frac{\zeta(2n+1)}{2n+1} x^{2n+1} \right]. \qquad (C.20)$$

Approximating the zeta-function with its asymptotic value $\zeta(n) \approx 1$, it is easy to see from this equation that $R_2^{(n)}$ scales as $n!$ for large n. Numerically we find

$$R_2^{(n)} \approx 0.315 \times (-1)^n n! \qquad \text{for } n \gg 1. \qquad (C.21)$$

For illustration, we list in table 5 the Fourier transforms of $\mathcal{L}_n(\vec{p}_T, \mu)$ for $n \leq 5$, and in table 6 the inverse Fourier transforms of L_b^n for $n \leq 6$. The first few $R_2^{(n)}$ are given in table 7.

C.3 Convolutions

Convolutions of plus functions can be conveniently calculated by transforming to impact parameter space with eqs. (C.13) and (C.15), where convolutions become simple products, and then Fourier transforming back using eq. (C.16).

The convolution of two $\mathcal{L}^a$ is given by [15]

$$(\mathcal{L}^a \otimes \mathcal{L}^b)(\vec{p}_T, \mu) = V_2(a, b)\mathcal{L}^{a+b}(\vec{p}_T, \mu) - \frac{\mathcal{L}^a(\vec{p}_T, \mu)}{b} - \frac{\mathcal{L}^b(\vec{p}_T, \mu)}{a} + \left(\frac{V_2(a, b)}{a+b} - \frac{1}{ab} \right)\delta(\vec{p}_T),$$

$$(C.22)$$

$\mathcal{L}_n(\vec{p}_T, \mu)$	$\mathrm{FT}[\mathcal{L}_n(\vec{p}_T, \mu)]$
$\mathcal{L}_0(\vec{p}_T, \mu)$	$-L_b$
$\mathcal{L}_1(\vec{p}_T, \mu)$	$+\frac{1}{2}L_b^2$
$\mathcal{L}_2(\vec{p}_T, \mu)$	$-\frac{1}{3}L_b^3 - \frac{4}{3}\zeta_3$
$\mathcal{L}_3(\vec{p}_T, \mu)$	$+\frac{1}{4}L_b^4 + 4\zeta_3 L_b$
$\mathcal{L}_4(\vec{p}_T, \mu)$	$-\frac{1}{5}L_b^5 - 8\zeta_3 L_b^2 - \frac{48}{5}\zeta_5$
$\mathcal{L}_5(\vec{p}_T, \mu)$	$+\frac{1}{6}L_b^6 + \frac{40}{3}\zeta_3 L_b^3 + 48\zeta_5 L_b + \frac{80}{3}\zeta_3^2$

Table 5. Fourier transform of $\mathcal{L}_n(\vec{p}_T, \mu)$ to $\vec{b}_T$-space for $n \leq 5$. Results are expressed in terms of $L_b = \ln(b_T^2 \mu^2 e^{2\gamma_E}/4)$. See eq. (C.15).

L_b^n	$\mathrm{FT}^{-1}[L_b^n]$
1	$\delta(\vec{p}_T)$
L_b	$-\mathcal{L}_0(\vec{p}_T, \mu)$
L_b^2	$+2\mathcal{L}_1(\vec{p}_T, \mu)$
L_b^3	$-3\mathcal{L}_2(\vec{p}_T, \mu) - 4\zeta_3 \delta(\vec{p}_T)$
L_b^4	$+4\mathcal{L}_3(\vec{p}_T, \mu) + 16\zeta_3 \mathcal{L}_0(\vec{p}_T, \mu)$
L_b^5	$-5\mathcal{L}_4(\vec{p}_T, \mu) - 80\zeta_3 \mathcal{L}_1(\vec{p}_T, \mu) - 48\zeta_5 \delta(\vec{p}_T)$
L_b^6	$+6\mathcal{L}_5(\vec{p}_T, \mu) + 240\zeta_3 \mathcal{L}_2(\vec{p}_T, \mu) + 288\zeta_5 \mathcal{L}_0(\vec{p}_T, \mu) + 160\zeta_3^2 \delta(\vec{p}_T)$

Table 6. Fourier transform of $L_b^n = \ln^n(b_T^2 \mu^2 e^{2\gamma_E}/4)$ to $\vec{p}_T$-space for $n \leq 6$. See eq. (C.16).

where

$$V_2(a, b) = \frac{\Gamma(1 - a - b)}{\Gamma(1 - a)\Gamma(1 - b)} \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}. \tag{C.23}$$

Convolutions of type $\mathcal{L}_n \otimes \mathcal{L}^a$ and $\mathcal{L}_n \otimes \mathcal{L}_m$ can be obtained by applying eq. (C.10) to both sides of the above relation by carefully taking the derivative with respect to a and b. An important special case are multiple convolutions of $\mathcal{L}_0(\vec{p}_T, \mu)$, which are found to be

$$(\mathcal{L}_0^{\otimes n})(\vec{p}_T, \mu) = \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2} e^{i\vec{p}_T \cdot \vec{b}_T}(-L_b)^n \tag{C.24}$$

$$= (-1)^n \sum_{k=0}^{n-1} (-1)^{k+1} n \binom{n-1}{k} R_2^{(n-k-1)} \mathcal{L}_k(\vec{p}_T, \mu) + (-1)^n R_2^{(n)} \delta(\vec{p}_T).$$

$R_2^{(n)}$	Exact value	Numerical value
$R_2^{(0)}$	1	1
$R_2^{(1)}$	0	0
$R_2^{(2)}$	0	0
$R_2^{(3)}$	$-4\zeta_3$	$\approx -4.81 = -0.801 \times 3!$
$R_2^{(4)}$	0	0
$R_2^{(5)}$	$-48\zeta_5$	$\approx -49.8 = -0.415 \times 5!$
$R_2^{(6)}$	$160\zeta_3^2$	$\approx 231.2 = 0.321 \times 6!$
$R_2^{(7)}$	$-1440\zeta_7$	$\approx -1452 = -0.288 \times 7!$

Table 7. The first values of $R_2^{(n)}$, defined in eq. (C.17).

C.4 Integral relations

Here we collect a few important integrals with distributional scale setting according to eq. (2.79),

$$\int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} f(\vec{p}_T, \mu') \equiv \frac{1}{2\pi p_T} \frac{\mathrm{d}}{\mathrm{d}p_T} \int_{k_T \leq p_T} \mathrm{d}^2\vec{k}_T \int_{p_T}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} f(\vec{k}_T, \mu') \,, \tag{C.25}$$

which are useful to resum the soft function iteratively using eqs. (5.10) and (5.11)

$$\int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \delta(\vec{p}_T) \ln^n \frac{\mu'^2}{\mu^2} = -\frac{1}{2} \mathcal{L}_n(\vec{p}_T, \mu) \,, \tag{C.26}$$

$$\int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \mathcal{L}_m(\vec{p}_T, \mu') \ln^n \frac{\mu'^2}{\mu^2} = -\frac{1}{2} \frac{m!\,n!}{(n+m+1)!} \mathcal{L}_{n+m+1}(\vec{p}_T, \mu) \,, \tag{C.27}$$

$$\int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \delta(\vec{p}_T) \ln^n \frac{\mu'}{\nu = p_T|_+} = (-1/2)^{n+1} \mathcal{L}_n(\vec{p}_T, \mu) \,, \tag{C.28}$$

$$\int_{p_T|_+}^{\mu} \frac{\mathrm{d}\mu'}{\mu'} \mathcal{L}_m(\vec{p}_T, \mu') \ln^n \frac{\mu'}{\nu = p_T|_+} = \frac{(-1/2)^{n+1}}{m+1} \mathcal{L}_{n+m+1}(\vec{p}_T, \mu) \,. \tag{C.29}$$

Setting $\nu = p_T|_+$ inside the last two integrals only makes sense within the distributional prescription.

Furthermore, the following two-dimensional generalization of eq. (2.24) is also useful,

$$\delta(\vec{p}_T) \ln^{n+1} \frac{\mu_0^2}{\mu^2} \bigg|_{\mu_0 = p_T|_+} = (n+1)\mathcal{L}_n(\vec{p}_T, \mu) \,,$$

$$(m+1)\mathcal{L}_m(\vec{p}_T, \mu) \ln^n \frac{\mu_0^2}{\mu^2} \bigg|_{\mu_0 = p_T|_+} = (n+m+1)\mathcal{L}_{n+m}(\vec{p}_T, \mu) \,,$$

$$\mathcal{L}_m(\vec{p}_T, \mu_0) \ln^n \frac{\mu_0^2}{\mu^2} \bigg|_{\mu_0 = p_T|_+} = 0 \,, \tag{C.30}$$

which holds for $n, m \geq 0$.

D Fixed-order expansion of the soft function

For completeness we give the fixed order expansion of the soft function $S(\vec{p}_T, \mu, \nu)$ to three loops, derived using distributional scale setting in sec. 5, and show the relation to the soft function predicted through scale setting in Fourier space. Similar expressions for the rapidity anomalous dimension γ_ν are given in secs. 4.2 and 4.3. To simplify the formulas, we make use of the known one-loop result

$$\gamma_{\nu\,0} = 0\,, \qquad \gamma_{S\,0} = 0\,. \tag{D.1}$$

The soft function is expanded as

$$S(\vec{p}_T, \mu, \nu) = \sum_{n=0}^{\infty} S^{(n)}(\vec{p}_T, \mu, \nu) \left[\frac{\alpha_s(\mu)}{4\pi}\right]^n\,. \tag{D.2}$$

The coefficients through $\mathcal{O}(\alpha_s^3)$ are expressed in terms of $L_\nu \equiv \ln(\mu/\nu)$ and $\mathcal{L}_n \equiv \mathcal{L}_n(\vec{p}_T, \mu)$,

$$S^{(0)} = \delta(\vec{p}_T)\,, \tag{D.3}$$

$$S^{(1)} = -\Gamma_0\,\mathcal{L}_1 - 2\Gamma_0\,L_\nu\,\mathcal{L}_0 + S_1\delta(\vec{p}_T)\,, \tag{D.4}$$

$$\begin{aligned}
S^{(2)} =& \frac{\Gamma_0^2}{2}\,\mathcal{L}_3 + \mathcal{L}_2(3\Gamma_0^2\,L_\nu + \beta_0\Gamma_0) + \mathcal{L}_1(4\Gamma_0^2\,L_\nu^2 + 2\beta_0\Gamma_0\,L_\nu - \Gamma_1 - \Gamma_0 S_1) \\
&+ \mathcal{L}_0\Big[L_\nu(-2\Gamma_1 - 2\Gamma_0 S_1) - \frac{1}{2}(\gamma_{S\,1} + \gamma_{\nu\,1}) - \beta_0 S_1 + 2\Gamma_0^2\zeta_3\Big] \\
&+ \delta(\vec{p}_T)\big[L_\nu(-\gamma_{\nu\,1} + 4\Gamma_0^2\zeta_3) + S_2\big]\,, \tag{D.5}
\end{aligned}$$

$$\begin{aligned}
S^{(3)} =& -\frac{\Gamma_0^3}{8}\,\mathcal{L}_5 + \mathcal{L}_4\Big(-\frac{5}{4}\Gamma_0^3 L_\nu - \frac{5}{6}\beta_0\Gamma_0^2\Big) \\
&+ \mathcal{L}_3\Big(-4\Gamma_0^3\,L_\nu^2 - \frac{14}{3}\beta_0\Gamma_0^2\,L_\nu - \beta_0^2\Gamma_0 + \Gamma_0\Gamma_1 + \frac{1}{2}\Gamma_0^2 S_1\Big) \\
&+ \mathcal{L}_2\Big[-4\Gamma_0^3 L_\nu^3 - 6\beta_0\Gamma_0^2 L_\nu^2 + L_\nu\big(-2\beta_0^2\Gamma_0 + 6\Gamma_0\Gamma_1 + 3\Gamma_0^2 S_1\big) \\
&\qquad + \beta_1\Gamma_0 + 2\beta_0\Gamma_1 + \frac{3}{4}\Gamma_0(\gamma_{S\,1} + \gamma_{\nu\,1}) + \frac{5}{2}\beta_0\Gamma_0 S_1 - 5\Gamma_0^3\zeta_3\Big] \\
&+ \mathcal{L}_1\Big[L_\nu^2\big(8\Gamma_0\Gamma_1 + 4\Gamma_0^2 S_1\big) \\
&\qquad + L_\nu\big(2\beta_1\Gamma_0 + 4\beta_0\Gamma_1 + 2\Gamma_0\gamma_{S\,1} + 3\Gamma_0\gamma_{\nu\,1} + 6\beta_0\Gamma_0 S_1 - 20\Gamma_0^3\zeta_3\big) \\
&\qquad - \Gamma_2 + \beta_0\gamma_{S\,1} + 2\beta_0\gamma_{\nu\,1} + 2\beta_0^2 S_1 - \Gamma_1 S_1 - \Gamma_0 S_2 - 12\beta_0\Gamma_0^2\zeta_3\Big] \\
&+ \mathcal{L}_0\Big[L_\nu^2(2\Gamma_0\gamma_{\nu\,1} - 16\Gamma_0^3\zeta_3) + L_\nu(-2\Gamma_2 + 2\beta_0\gamma_{\nu\,1} - 2\Gamma_1 S_1 - 2\Gamma_0 S_2 - 16\beta_0\Gamma_0^2\zeta_3) \\
&\qquad - \frac{1}{2}(\gamma_{S\,2} + \gamma_{\nu\,2}) - \frac{1}{2}(2\beta_1 + \gamma_{S\,1} + \gamma_{\nu\,1})S_1 - 2\beta_0 S_2 + 4\Gamma_0\Gamma_1\zeta_3 + 2\Gamma_0^2 S_1\zeta_3 - 6\Gamma_0^3\zeta_5\Big] \\
&+ \delta(\vec{p}_T)\Big[-\frac{16}{3}\Gamma_0^3\zeta_3 L_\nu^3 - 8\beta_0\Gamma_0^2\zeta_3 L_\nu^2 \\
&\qquad + L_\nu\big(-\gamma_{\nu\,2} - \gamma_{\nu\,1}S_1 + 8\Gamma_0\Gamma_1\zeta_3 + 4\Gamma_0^2 S_1\zeta_3 - 12\Gamma_0^3\zeta_5\big) + S_3\Big]\,. \tag{D.6}
\end{aligned}$$

The results through $\mathcal{O}(\alpha_s^2)$ agree with the explicit calculation using the η-regulator in ref. [62]. The S_n denote the constants as derived using distributional setting, see sec. 5. They are related to $\tilde{S}_n$, the correspond constants with scale setting in Fourier space, by

$$S_1 = \tilde{S}_1 \,, \tag{D.7}$$

$$S_2 = \tilde{S}_2 + \frac{4}{3}\beta_0\Gamma_0\zeta_3 \,, \tag{D.8}$$

$$S_3 = \tilde{S}_3 + \frac{4}{3}\beta_1\Gamma_0\zeta_3 + \frac{8}{3}\beta_0\Gamma_1\zeta_3 + \Gamma_0(\gamma_{S\,1} + \gamma_{\nu\,1})\zeta_3 + \frac{10}{3}\beta_0\Gamma_0 S_1\zeta_3 - \frac{10}{3}\Gamma_0^3\zeta_3^2 - 8\beta_0\Gamma_0^2\zeta_5 \,. \tag{D.9}$$

References

[1] ATLAS collaboration, G. Aad et al., *Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS*, *JHEP* **09** (2014) 112, [arXiv:1407.4222].

[2] ATLAS collaboration, G. Aad et al., *Fiducial and differential cross sections of Higgs boson production measured in the four-lepton decay channel in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector*, *Phys. Lett.* **B738** (2014) 234–253, [arXiv:1408.3226].

[3] ATLAS collaboration, G. Aad et al., *Measurements of the Total and Differential Higgs Boson Production Cross Sections Combining the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ Decay Channels at $\sqrt{s}=8TeV$ with the ATLAS Detector*, *Phys. Rev. Lett.* **115** (2015) 091801, [arXiv:1504.05833].

[4] ATLAS collaboration, G. Aad et al., *Measurement of fiducial differential cross sections of gluon-fusion production of Higgs bosons decaying to $WW^* \to e\nu\mu\nu$ with the ATLAS detector at $\sqrt{s} = 8$ TeV*, *JHEP* **08** (2016) 104, [arXiv:1604.02997].

[5] ATLAS collaboration, T. A. collaboration, *Measurement of fiducial, differential and production cross sections in the $H \to \gamma\gamma$ decay channel with 13.3 fb^{-1} of 13 TeV proton-proton collision data with the ATLAS detector*, .

[6] CMS collaboration, V. Khachatryan et al., *Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **04** (2016) 005, [arXiv:1512.08377].

[7] CMS collaboration, V. Khachatryan et al., *Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C76** (2016) 13, [arXiv:1508.07819].

[8] CMS collaboration, V. Khachatryan et al., *Measurement of the transverse momentum spectrum of the Higgs boson produced in pp collisions at $\sqrt{s} = 8$ TeV using $H \to WW$ decays*, *Submitted to: JHEP* (2016) , [arXiv:1606.01522].

[9] ATLAS collaboration, G. Aad et al., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **09** (2014) 145, [arXiv:1406.3660].

[10] ATLAS collaboration, G. Aad et al., *Measurement of the transverse momentum and ϕ_η^* distributions of DrellYan lepton pairs in protonproton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) 291, [arXiv:1512.02192].

[11] CMS collaboration, S. Chatrchyan et al., *Measurement of the Rapidity and Transverse Momentum Distributions of Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev.* **D85** (2012) 032002, [`arXiv:1110.4973`].

[12] CMS collaboration, V. Khachatryan et al., *Measurement of the Z boson differential cross section in transverse momentum and rapidity in protonproton collisions at 8 TeV*, *Phys. Lett.* **B749** (2015) 187–209, [`arXiv:1504.03511`].

[13] CMS collaboration, V. Khachatryan et al., *Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *Submitted to: JHEP* (2016) , [`arXiv:1606.05864`].

[14] S. Frixione, P. Nason and G. Ridolfi, *Problems in the resummation of soft gluon effects in the transverse momentum distributions of massive vector bosons in hadronic collisions*, *Nucl. Phys.* **B542** (1999) 311–328, [`hep-ph/9809367`].

[15] J.-Y. Chiu, A. Jain, D. Neill and I. Z. Rothstein, *A Formalism for the Systematic Treatment of Rapidity Logarithms in Quantum Field Theory*, *JHEP* **05** (2012) 084, [`arXiv:1202.0814`].

[16] T. Becher and M. Neubert, *Drell-Yan Production at Small q_T, Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J.* **C71** (2011) 1665, [`arXiv:1007.4005`].

[17] T. Becher, M. Neubert and D. Wilhelm, *Electroweak Gauge-Boson Production at Small q_T: Infrared Safety from the Collinear Anomaly*, *JHEP* **02** (2012) 124, [`arXiv:1109.6027`].

[18] T. Becher, M. Neubert and D. Wilhelm, *Higgs-Boson Production at Small Transverse Momentum*, *JHEP* **05** (2013) 110, [`arXiv:1212.2621`].

[19] U. D'Alesio, M. G. Echevarria, S. Melis and I. Scimemi, *Non-perturbative QCD effects in q_T spectra of Drell-Yan and Z-boson production*, *JHEP* **11** (2014) 098, [`arXiv:1407.3311`].

[20] M. G. Echevarria, T. Kasemets, P. J. Mulders and C. Pisano, *QCD evolution of (un)polarized gluon TMDPDFs and the Higgs q_T-distribution*, *JHEP* **07** (2015) 158, [`arXiv:1502.05354`].

[21] A. Kulesza and W. J. Stirling, *Sudakov logarithm resummation in transverse momentum space for electroweak boson production at hadron colliders*, *Nucl. Phys.* **B555** (1999) 279–305, [`hep-ph/9902234`].

[22] A. Kulesza and W. J. Stirling, *On the resummation of subleading logarithms in the transverse momentum distribution of vector bosons produced at hadron colliders*, *JHEP* **01** (2000) 016, [`hep-ph/9909271`].

[23] A. Kulesza and W. J. Stirling, *Soft gluon resummation in transverse momentum space for electroweak boson production at hadron colliders*, *Eur. Phys. J.* **C20** (2001) 349–356, [`hep-ph/0103089`].

[24] P. F. Monni, E. Re and P. Torrielli, *Higgs Transverse-Momentum Resummation in Direct Space*, *Phys. Rev. Lett.* **116** (2016) 242001, [`arXiv:1604.02191`].

[25] A. Banfi, G. P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073, [`hep-ph/0407286`].

[26] A. Banfi, H. McAslan, P. F. Monni and G. Zanderighi, *A general method for the resummation of event-shape distributions in e^+e annihilation*, *JHEP* **05** (2015) 102, [`arXiv:1412.2126`].

[27] Z. Ligeti, I. W. Stewart and F. J. Tackmann, *Treating the b quark distribution function with reliable uncertainties*, *Phys. Rev.* **D78** (2008) 114014, [`arXiv:0807.1926`].

[28] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu and I. W. Stewart, *Thrust at N^3LL with Power Corrections and a Precision Global Fit for $\alpha_s(m_Z)$*, *Phys. Rev.* **D83** (2011) 074021, [arXiv:1006.3080].

[29] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *Higgs Production with a Central Jet Veto at NNLL+NNLO*, *JHEP* **04** (2011) 092, [arXiv:1012.4480].

[30] L. G. Almeida, S. D. Ellis, C. Lee, G. Sterman, I. Sung and J. R. Walsh, *Comparing and counting logs in direct and effective methods of QCD resummation*, *JHEP* **04** (2014) 174, [arXiv:1401.4460].

[31] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, *Phys. Rev.* **D85** (2012) 034011, [arXiv:1107.2117].

[32] I. W. Stewart, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Jet p_T resummation in Higgs production at $NNLL' + NNLO$*, *Phys. Rev.* **D89** (2014) 054001, [arXiv:1307.1808].

[33] J.-y. Chiu, A. Jain, D. Neill and I. Z. Rothstein, *The Rapidity Renormalization Group*, *Phys. Rev. Lett.* **108** (2012) 151601, [arXiv:1104.0881].

[34] C. W. Bauer, S. Fleming and M. E. Luke, *Summing Sudakov logarithms in $B \to X_s\gamma$ in effective field theory*, *Phys. Rev.* **D63** (2000) 014006, [hep-ph/0005275].

[35] C. W. Bauer, S. Fleming, D. Pirjol and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, *Phys. Rev.* **D63** (2001) 114020, [hep-ph/0011336].

[36] C. W. Bauer and I. W. Stewart, *Invariant operators in collinear effective theory*, *Phys. Lett.* **B516** (2001) 134–142, [hep-ph/0107001].

[37] C. W. Bauer, D. Pirjol and I. W. Stewart, *Soft collinear factorization in effective field theory*, *Phys. Rev.* **D65** (2002) 054022, [hep-ph/0109045].

[38] J. C. Collins and D. E. Soper, *Back-To-Back Jets in QCD*, *Nucl. Phys.* **B193** (1981) 381. [Erratum: Nucl. Phys.B213,545(1983)].

[39] J. C. Collins and D. E. Soper, *Back-To-Back Jets: Fourier Transform from B to K-Transverse*, *Nucl. Phys.* **B197** (1982) 446–476.

[40] J. C. Collins, D. E. Soper and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, *Nucl. Phys.* **B250** (1985) 199–224.

[41] J. Collins, *Foundations of perturbative QCD*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge Univ. Press, New York, NY, 2011.

[42] C. Balazs and C. P. Yuan, *Soft gluon effects on lepton pairs at hadron colliders*, *Phys. Rev.* **D56** (1997) 5558–5583, [hep-ph/9704258].

[43] C. Balazs and C. P. Yuan, *Higgs boson production at the LHC with soft gluon effects*, *Phys. Lett.* **B478** (2000) 192–198, [hep-ph/0001103].

[44] C. Balazs, E. L. Berger, P. M. Nadolsky and C. P. Yuan, *All-orders resummation for diphoton production at hadron colliders*, *Phys. Lett.* **B637** (2006) 235–240, [hep-ph/0603037].

[45] J. Wang, C. S. Li, Z. Li, C. P. Yuan and H. T. Li, *Improved Resummation Prediction on Higgs Production at Hadron Colliders*, *Phys. Rev.* **D86** (2012) 094026, [arXiv:1205.4311].

[46] S. Catani, D. de Florian and M. Grazzini, *Universality of nonleading logarithmic contributions in transverse momentum distributions*, *Nucl. Phys.* **B596** (2001) 299–312, [hep-ph/0008184].

[47] D. de Florian and M. Grazzini, *Next-to-next-to-leading logarithmic corrections at small transverse momentum in hadronic collisions*, *Phys. Rev. Lett.* **85** (2000) 4678–4681, [hep-ph/0008152].

[48] D. de Florian and M. Grazzini, *The Structure of large logarithmic corrections at small transverse momentum in hadronic collisions*, *Nucl. Phys.* **B616** (2001) 247–285, [hep-ph/0108273].

[49] S. Catani and M. Grazzini, *QCD transverse-momentum resummation in gluon fusion processes*, *Nucl. Phys.* **B845** (2011) 297–323, [arXiv:1011.3918].

[50] S. Catani, L. Cieri, D. de Florian, G. Ferrera and M. Grazzini, *Universality of transverse-momentum resummation and hard factors at the NNLO*, *Nucl. Phys.* **B881** (2014) 414–443, [arXiv:1311.1654].

[51] G. Bozzi, S. Catani, G. Ferrera, D. de Florian and M. Grazzini, *Production of Drell-Yan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy*, *Phys. Lett.* **B696** (2011) 207–213, [arXiv:1007.2351].

[52] S. Catani, D. de Florian, G. Ferrera and M. Grazzini, *Vector boson production at hadron colliders: transverse-momentum resummation and leptonic decay*, *JHEP* **12** (2015) 047, [arXiv:1507.06937].

[53] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Higgs boson production at the LHC: Transverse-momentum resummation and rapidity dependence*, *Nucl. Phys.* **B791** (2008) 1–19, [arXiv:0705.3887].

[54] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to 2\gamma\gamma$, $H \to WW \to l\nu l\nu$ and $H \to ZZ \to 4l$ decay modes*, *JHEP* **06** (2012) 132, [arXiv:1203.6321].

[55] L. Cieri, F. Coradeschi and D. de Florian, *Diphoton production at hadron colliders: transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy*, *JHEP* **06** (2015) 185, [arXiv:1505.03162].

[56] M. G. Echevarria, A. Idilbi and I. Scimemi, *Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone*, *JHEP* **07** (2012) 002, [arXiv:1111.4996].

[57] M. G. Echevarra, A. Idilbi and I. Scimemi, *Soft and Collinear Factorization and Transverse Momentum Dependent Parton Distribution Functions*, *Phys. Lett.* **B726** (2013) 795–801, [arXiv:1211.1947].

[58] M. G. Echevarria, A. Idilbi and I. Scimemi, *Unified treatment of the QCD evolution of all (un-)polarized transverse momentum dependent functions: Collins function as a study case*, *Phys. Rev.* **D90** (2014) 014003, [arXiv:1402.0869].

[59] T. C. Rogers, *An Overview of Transverse Momentum Dependent Factorization and Evolution*, *Eur. Phys. J.* **A52** (2016) 153, [arXiv:1509.04766].

[60] I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *Factorization at the LHC: From PDFs to Initial State Jets*, *Phys. Rev. D* **81** (2010) 094035, [arXiv:0910.0467].

[61] I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *The Quark Beam Function at NNLL*, *JHEP* **09** (2010) 005, [arXiv:1002.2213].

[62] T. Lübbert, J. Oredsson and M. Stahlhofen, *Rapidity renormalized TMD soft and beam functions at two loops*, *JHEP* **03** (2016) 168, [`arXiv:1602.01829`].

[63] X.-d. Ji, J.-p. Ma and F. Yuan, *QCD factorization for semi-inclusive deep-inelastic scattering at low transverse momentum*, *Phys. Rev.* **D71** (2005) 034005, [`hep-ph/0404183`].

[64] M. Beneke and T. Feldmann, *Factorization of heavy to light form-factors in soft collinear effective theory*, *Nucl. Phys.* **B685** (2004) 249–296, [`hep-ph/0311335`].

[65] J.-y. Chiu, F. Golf, R. Kelley and A. V. Manohar, *Electroweak Sudakov corrections using effective field theory*, *Phys. Rev. Lett.* **100** (2008) 021802, [`arXiv:0709.2377`].

[66] T. Becher and G. Bell, *Analytic Regularization in Soft-Collinear Effective Theory*, *Phys. Lett.* **B713** (2012) 41–46, [`arXiv:1112.3907`].

[67] J.-y. Chiu, A. Fuhrer, A. H. Hoang, R. Kelley and A. V. Manohar, *Soft-Collinear Factorization and Zero-Bin Subtractions*, *Phys. Rev.* **D79** (2009) 053007, [`arXiv:0901.1332`].

[68] Y. Li, D. Neill and H. X. Zhu, *An Exponential Regulator for Rapidity Divergences*, *Submitted to: Phys. Rev. D* (2016) , [`arXiv:1604.00392`].

[69] S. Catani and M. Grazzini, *Higgs Boson Production at Hadron Colliders: Hard-Collinear Coefficients at the NNLO*, *Eur. Phys. J.* **C72** (2012) 2013, [`arXiv:1106.4652`]. [Erratum: Eur. Phys. J.C72,2132(2012)].

[70] S. Catani, L. Cieri, D. de Florian, G. Ferrera and M. Grazzini, *Vector boson production at hadron colliders: hard-collinear coefficients at the NNLO*, *Eur. Phys. J.* **C72** (2012) 2195, [`arXiv:1209.0158`].

[71] T. Gehrmann, T. Luebbert and L. L. Yang, *Calculation of the transverse parton distribution functions at next-to-next-to-leading order*, *JHEP* **06** (2014) 155, [`arXiv:1403.6451`].

[72] M. G. Echevarria, I. Scimemi and A. Vladimirov, *Universal transverse momentum dependent soft function at NNLO*, *Phys. Rev.* **D93** (2016) 054004, [`arXiv:1511.05590`].

[73] M. G. Echevarria, I. Scimemi and A. Vladimirov, *Unpolarized Transverse Momentum Dependent Parton Distribution and Fragmentation Functions at next-to-next-to-leading order*, *JHEP* **09** (2016) 004, [`arXiv:1604.07869`].

[74] G. Parisi and R. Petronzio, *Small Transverse Momentum Distributions in Hard Processes*, *Nucl. Phys.* **B154** (1979) 427–440.

[75] J. R. Gaunt, *Glauber Gluons and Multiple Parton Interactions*, *JHEP* **07** (2014) 110, [`arXiv:1405.2080`].

[76] A. H. Hoang and I. W. Stewart, *Designing gapped soft functions for jet production*, *Phys. Lett.* **B660** (2008) 483–493, [`arXiv:0709.3519`].

[77] I. Scimemi and A. Vladimirov, *Power corrections and renormalons in Transverse Momentum Distributions*, [`arXiv:1609.06047`].

[78] G. P. Korchemsky and G. F. Sterman, *Nonperturbative corrections in resummed cross-sections*, *Nucl. Phys.* **B437** (1995) 415–432, [`hep-ph/9411211`].

[79] M. G. Echevarria, A. Idilbi, A. Schfer and I. Scimemi, *Model-Independent Evolution of Transverse Momentum Dependent Distribution Functions (TMDs) at NNLL*, *Eur. Phys. J.* **C73** (2013) 2636, [`arXiv:1208.1281`].

[80] J. Collins and T. Rogers, *Understanding the large-distance behavior of transverse-momentum-dependent parton densities and the Collins-Soper evolution kernel*, *Phys. Rev.* **D91** (2015) 074020, [`arXiv:1412.3820`].

[81] T. Becher and G. Bell, *Enhanced nonperturbative effects through the collinear anomaly*, *Phys. Rev. Lett.* **112** (2014) 182002, [`arXiv:1312.5327`].

[82] R. K. Ellis and S. Veseli, *W and Z transverse momentum distributions: Resummation in q_T space*, *Nucl. Phys.* **B511** (1998) 649–669, [`hep-ph/9706526`].

[83] Y. L. Dokshitzer, D. Diakonov and S. I. Troian, *Hard Processes in Quantum Chromodynamics*, *Phys. Rept.* **58** (1980) 269–395.

[84] Y. L. Dokshitzer, D. Diakonov and S. I. Troian, *On the Transverse Momentum Distribution of Massive Lepton Pairs*, *Phys. Lett.* **B79** (1978) 269–272.

[85] D. Neill, I. Z. Rothstein and V. Vaidya, *The Higgs Transverse Momentum Distribution at NNLL and its Theoretical Errors*, *JHEP* **12** (2015) 097, [`arXiv:1503.00005`].

[86] T. Becher, M. Neubert and B. D. Pecjak, *Factorization and Momentum-Space Resummation in Deep-Inelastic Scattering*, *JHEP* **01** (2007) 076, [`hep-ph/0607228`].

[87] G. Sterman and M. Zeng, *Quantifying Comparisons of Threshold Resummations*, *JHEP* **05** (2014) 132, [`arXiv:1312.5397`].

[88] M. Bonvini, S. Forte, G. Ridolfi and L. Rottoli, *Resummation prescriptions and ambiguities in SCET vs. direct QCD: Higgs production as a case study*, *JHEP* **01** (2015) 046, [`arXiv:1409.0864`].

[89] E. Laenen, G. F. Sterman and W. Vogelsang, *Higher order QCD corrections in prompt photon production*, *Phys. Rev. Lett.* **84** (2000) 4296–4299, [`hep-ph/0002078`].

[90] A. Kulesza, G. F. Sterman and W. Vogelsang, *Joint resummation in electroweak boson production*, *Phys. Rev.* **D66** (2002) 014011, [`hep-ph/0202251`].

[91] Y. Gao, C. S. Li and J. J. Liu, *Transverse momentum resummation for Higgs production in soft-collinear effective theory*, *Phys. Rev.* **D72** (2005) 114020, [`hep-ph/0501229`].

[92] A. Idilbi, X.-d. Ji and F. Yuan, *Transverse momentum distribution through soft-gluon resummation in effective field theory*, *Phys. Lett.* **B625** (2005) 253–263, [`hep-ph/0507196`].

[93] S. Mantry and F. Petriello, *Factorization and Resummation of Higgs Boson Differential Distributions in Soft-Collinear Effective Theory*, *Phys. Rev.* **D81** (2010) 093007, [`arXiv:0911.4135`].

[94] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys.* **B349** (1991) 635–654.

[95] A. Jain, I. Scimemi and I. W. Stewart, *Two-loop Jet-Function and Jet-Mass for Top Quarks*, *Phys. Rev.* **D77** (2008) 094008, [`arXiv:0801.0743`].