hep-ph 1206.4312

arXiv:1206.4312v3 [hep-ph] 24 Sep 2012

DESY 12-104

Resummation Properties of Jet Vetoes at the LHC

Frank J. Tackmann,[1] Jonathan R. Walsh,[2] and Saba Zuberi[2]

[1] *Theory Group, Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany*

[2] *Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720*

(Dated: June 19, 2012)

Jet vetoes play an important role at the LHC in the search for the Higgs and ultimately in precise measurements of its properties. Many Higgs analyses divide the cross section into exclusive jet bins to maximize the sensitivity in different production and decay channels. For a given jet category, the veto on additional jets introduces sensitivity to soft and collinear emissions, which causes logarithms in the perturbative expansion that need to be resummed to obtain precise predictions. We study the higher-order resummation properties of several conceptually distinct kinematic variables that can be used to veto jets in hadronic collisions. We consider two inclusive variables, the scalar sum over p_T and beam thrust, and two corresponding exclusive variables based on jet algorithms, namely the largest p_T and largest beam thrust of a jet. The inclusive variables can, in principle, be resummed to higher orders. We show that for the jet-based variables, there are dual effects due to clustering in the jet algorithm for both large and small jet radius R that make a complete resummation at or beyond next-to-leading logarithmic order (NLL) challenging. For $R \sim 1$, the clustering of soft and collinear emissions gives $\mathcal{O}(1)$ contributions starting at next-to-next-to-leading logarithmic order that are not reproduced by an all-orders soft-collinear factorization formula and therefore are not automatically resummed by it. For $R \ll 1$, clustering induces logarithms of R that contribute at NLL in the exponent of the cross section, which cannot be resummed with currently available methods. We explicitly compute the leading jet clustering effects at $\mathcal{O}(\alpha_s^2)$ and comment on their numerical size.

I. INTRODUCTION

The LHC is at the brink of discovering (or ruling out) a Standard-Model-like Higgs boson [1, 2]. Many Higgs and new-physics analyses divide the data into categories based on the number of hard jets in the final state. This "jet-binning" significantly enhances the experimental sensitivity to various combinations of Higgs production and decay channels. The experimental distinction between different channels is of course an important requirement to determine the Higgs properties. Any coupling measurement requires precise and reliable theory predictions of the jet cross sections used in the experimental analysis.

By vetoing jets, one restricts the phase space for additional emissions. This makes the cross section sensitive to soft and collinear radiation, which induces Sudakov double logarithms of the jet-veto variable at each order in perturbation theory. At small enough cuts the logarithmic corrections become large and dominate the perturbative series, which degrades the reliability of fixed-order perturbation theory. To obtain precise predictions and robust uncertainty estimates it is important to understand the structure of large logarithms and ideally resum them to all orders in perturbation theory. The form of the jet-veto logarithms and the precise structure of the logarithmic series depends on the details of how the jet veto is imposed.

In this paper we consider four kinematic variables that are conceptually and theoretically distinct, and study the prospects for resumming the respective large logarithms when these (types of) variables are used to veto central jets. For definiteness we will discuss the process $gg \to H+0$ jets in the following, but our analysis applies to the production of any color-singlet in conjunction with a veto on central jets. It also carries over to generic processes where one requires N signal jets in the final state and places a veto on additional jets, i.e. when measuring an exclusive N-jet cross section.

In principle, one can think of various kinematic variables that can be employed to enforce a veto on hard emissions. The four variables we will discuss are summarized in Table I and are described next. They are classified as either inclusive or exclusive (jet-based) variables and according to their sensitivity to either the virtuality or p_T of emissions.

The inclusive variables we consider are beam thrust [3] and the scalar sum of p_T,

$$\mathcal{T}_B = \sum_m |\vec{p}_{Tm}|\, e^{-|y_m - Y|}\,, \tag{1}$$

$$E_T = \sum_m |\vec{p}_{Tm}|\,, \tag{2}$$

where $\vec{p}_{Tm}$ and y_m are the transverse momenta and rapidities with respect to the beam axis of all particles in the final state, but excluding the Higgs decay products, and Y is the rapidity of the Higgs boson.[1] For these inclusive variables one sums over all hadrons m in the final state, which at the perturbative level corresponds to constraining the sum over all emissions. As a result such

[1] For cases like $H \to WW$ where Y cannot be measured directly, one can also consider the analog of $\mathcal{T}_B$ defined in the hadronic center-of-mass frame by setting $Y = 0$ in Eq. (1).

inclusive variables are theoretically the cleanest and best understood. The complete resummation of logarithms at small $\mathcal{T}_B$ and E_T has been carried out to next-to-next-to-leading logarithmic order (NNLL) [4, 5] and next-to-leading logarithmic order (NLL) [6], respectively, and the extension to higher logarithmic orders poses no conceptual difficulties. The inclusive nature of these variables makes it challenging to measure them in a hadron-collider environment. These experimental issues can be mitigated by summing over jets rather than hadrons, as is done in current event-shape measurements at the LHC [7, 8], or by only summing over charged tracks, which can then be corrected for in the measurement.

The difference between $\mathcal{T}_B$ and E_T is in the different rapidity weighting of emissions. In the rest frame of the Higgs, where $Y = 0$, $\mathcal{T}_B$ is equivalent to the sum over the small light-cone component of momenta, $\mathcal{T}_B = \sum_k (E_k - |p_k^z|) \sim t/m_H$, where t is the spacelike virtuality of the colliding hard partons after initial-state radiation. Hence, by measuring beam thrust $\mathcal{T}_B$ one is sensitive to the virtuality scale of emissions. In contrast, by measuring E_T one is sensitive to the p_T scale of emissions, and one might call it "beam broadening". This difference in sensitivity to virtuality vs. p_T causes the logarithmic series at small $\mathcal{T}_B$ or E_T to have very different structures. In the parton shower, this difference is analogous to the different Sudakov form factors for virtuality and p_T ordered showers.

The exclusive jet-based variables we consider are the largest beam-thrust or p_T of a jet, given by

$$\mathcal{T}_j = \max_{m \in j(R)} |\vec{p}_{Tm}|\, e^{-|y_m - Y|}\,, \tag{3}$$

$$p_{Tj} = \max_{m \in j(R)} |\vec{p}_{Tm}|\,, \tag{4}$$

where $\vec{p}_{Tj}$ and y_j are the jet's transverse momentum and rapidity. They have the same sensitivity to the virtuality or p_T scale of emissions as their inclusive counterparts $\mathcal{T}_B$ and E_T in Eqs. (1) and (2). The exclusive variables are based on identifying jets $j(R)$ of size R and considering the largest contribution from a jet. Perturbatively, this corresponds to constraining the maximum of all emissions at a typical "resolution scale" R, which is in contrast to constraining the sum of emissions as in the inclusive variables. We focus on jets using the k_{T} class of algorithms [9–13]. The jet-based variables are more straightforward to use experimentally, and a p_T veto on jets using the anti-k_{T} algorithm is the common choice in experiments. On the other hand, the variable's inherent dependence on the jet algorithm and the resolution scale R make them considerably less tractable theoretically. The resummation for p_{Tj} was considered recently in Refs. [14, 15].

We implement the veto on central jets by putting a constraint

$$\mathcal{T}_B \le \mathcal{T}^{\mathrm{cut}}\,, \qquad E_T \le p_T^{\mathrm{cut}}\,,$$
$$\mathcal{T}_j \le \mathcal{T}^{\mathrm{cut}}\,, \qquad p_{Tj} \le p_T^{\mathrm{cut}}\,, \tag{5}$$

	sensitive to	
	virtuality	p_T
inclusive	$\mathcal{T}_B$	E_T
jet-based (exclusive)	$\mathcal{T}_j$	p_{Tj}
constraint	$\le \mathcal{T}^{\mathrm{cut}}$	$\le p_T^{\mathrm{cut}}$

TABLE I: Classification of different jet-veto variables.

and define the small parameter λ by

$$\lambda^2 = \mathcal{T}^{\mathrm{cut}}/m_H \qquad \text{or} \qquad \lambda = p_T^{\mathrm{cut}}/m_H\,. \tag{6}$$

The cross section with a $\mathcal{T}^{\mathrm{cut}}$ or p_T^{cut} veto contains Sudakov logarithms $\alpha_s^n \ln^m \lambda$ with $m \le 2n$. In the region of small λ the logarithms dominate the cross section and we want to resum them. At the same time, any nonlogarithmic contributions that depend on $\mathcal{T}^{\mathrm{cut}}$ or p_T^{cut} are suppressed by relative powers of λ, and can be added to the resummed result.

For the jet-based variables, one can distinguish two cases of how the jet size R is counted relative to λ, we can consider either $R \sim \lambda$ or $R \gg \lambda$.[2] For this distinction it is irrelevant whether one considers $R \sim 1$ or $R \ll 1$. We stress that which of these two formal cases is in the end more appropriate in practice, for given numerical values of R and $\mathcal{T}^{\mathrm{cut}}/m_H$ or p_T^{cut}/m_H, is a separate question that needs to be studied numerically and will be addressed in Sec. V.

As we will show in this paper, there are competing effects in either limit arising from the dependence on the jet algorithm, which spoil the complete logarithmic resummation. Jet algorithm dependent effects first arise at $\mathcal{O}(\alpha_s^2)$ in the form

$$\alpha_s^2\, f_{\mathrm{alg}}^{(2)}(R)\, \ln \lambda\,. \tag{7}$$

The function $f_{\mathrm{alg}}^{(2)}(R)$ depends on the algorithm and the jet-veto variable, and in general contains terms of $\mathcal{O}(R^2)$, constant terms, and $\ln R$ terms. These were calculated in Ref. [14] for the p_{Tj} veto.

For $R \gg \lambda$ the jet algorithm mixes soft and collinear contributions in the measurement, giving rise to a term in Eq. (7) of the form

$$\alpha_s^2 R^2 \ln \lambda\,, \tag{8}$$

at leading order in R. When keeping the full R dependence, these soft-collinear mixing terms inhibit an all-order soft-collinear factorization in the measurement at

[2] This counting is a natural proxy for the two general cases of R either scaling as a positive power of λ or R not scaling with λ. The difference amounts to whether we formally count $\mathcal{O}(R)$ terms as power corrections in λ or not and logarithms of R as $\ln \lambda$ or not.

leading power in λ. In Ref. [15], a factorization formula for the p_{Tj} veto was presented in the limit $\lambda \ll R \sim 1$. Their derivation does not account for the effect of these mixing terms to all orders, effectively assuming that they are power-suppressed, and hence breaks down for $R \sim 1$.

For $R \sim \lambda$, the mixing terms inhibiting the soft-collinear factorization of the measurement can be regarded as power corrections in λ, so factorization goes through. However, in this limit clustering effects that change the boundary of the jet at each order in α_s introduce corrections that depend on $\ln R$. These are referred to as clustering logarithms and were first pointed out in Ref. [16] and were studied in Refs. [17–21]. The clustering logarithms in $f^{(2)}_{\text{alg}}(R)$ give a contribution of the form

$$\alpha_s^2 \, \ln R \, \ln \lambda \,. \tag{9}$$

They are distinct from those previously studied since they are associated with collinear rather than soft divergences within each jet. The $\mathcal{O}(\alpha_s^2)$ term in Eq. (7) is the first in an all-orders series of terms of the form

$$\alpha_s^n \, f^{(n)}_{\text{alg}}(R) \, \ln \lambda \,, \qquad f^{(n)}_{\text{alg}}(R) \supset \ln^k R \,, \;\; k \leq n-1 \,. \tag{10}$$

For $R \sim \lambda$ we have to count $\ln R \sim \ln \lambda$, so these clustering logarithms give a new contribution at NLL at each higher order in α_s. Hence, they spoil the complete logarithmic resummation at NLL and beyond.

At a formal level, this means one is stuck between a rock and a hard place. We would like to consider R as small to justify not resumming soft-collinear mixing terms, by treating them as unresummed power corrections. But at the same time we would like to treat R as large, to avoid having to count $\ln R$ terms as large logarithms.

A possible way forward would be to resum those logarithms whose all-order series is known while simultaneously quantifying the effect of the terms that cannot be resummed. If these effects can be appropriately folded into uncertainty estimates, then a reliable theoretical prediction can be obtained. The formal power counting is less important in this case, and effectively the two formal limits for R are unified by reflecting the important terms from each case in the perturbative uncertainties.

The remainder of this paper is organized as follows. In Sec. II we present an overview of the necessary steps to achieve soft-collinear factorization and resummation, focusing on the properties of the measurement. We present the factorization formulas for the inclusive observables in Sec. II A and for the exclusive observables in Sec. II B. In Sec. III we discuss the soft-collinear mixing in detail and show that it gives an $\mathcal{O}(R^2)$ contribution to the rate. We demonstrate this by power counting as well as explicit calculation of the mixing terms at $\mathcal{O}(\alpha_s^2)$, the details of which are given in Appendix B. In Sec. IV we discuss the clustering logarithms in detail and calculate their contribution at $\mathcal{O}(\alpha_s^2)$, with the details of the calculation given in Appendix C. We use non-Abelian exponentiation and the collinear sensitivity of the clustering logarithms to show that their contribution to the cross section exponentiates and for $R \sim \lambda$ contributes at NLL in the exponent. In Sec. V we summarize our findings and give an outlook based on a numerical analysis of the size of the $\mathcal{O}(\alpha_s^2)$ mixing and clustering terms for different values of R.

II. FACTORIZATION FORMULAS

The large Sudakov logarithms in the jet-vetoed cross sections arise as a remnant of the cancellation of soft and collinear IR divergences between virtual and real contributions. Their resummation is thus intimately tied to the universal structure of QCD amplitudes in the soft and collinear limit. Furthermore, resummation relies on the fact that soft gluon emissions from energetic particles are eikonal and that the total soft eikonal matrix element factorizes from the remaining amplitude.

A convenient framework to study the logarithmic structure of the cross section is provided by soft-collinear effective theory (SCET) [22–27], which makes the soft-collinear limit of QCD manifest at a Lagrangian and operator level using a systematic power expansion in the small parameter λ. The resummation of Sudakov logarithms is then achieved by standard effective-theory methods through a systematic scale separation and renormalization group evolution between the scales. For a detailed discussion of this procedure in the context of SCET we refer the reader to the literature. In the following, we give a schematic overview of the basic steps, concentrating on the features that are important for our further analysis.

After matching full QCD onto SCET, the cross section in SCET for $gg \to H$ with no hard jets in the final states has the schematic form (for details see e.g. Refs. [3, 5])

$$\mathrm{d}\sigma_{gg\to H} \sim |C_{ggH}(\mu)|^2 \big\langle p_a p_b \big| \mathcal{O}_{ggH}(\mu)^\dagger \, \widehat{\mathcal{M}} \, \mathcal{O}_{ggH}(\mu) \big| p_a p_b \big\rangle \,. \tag{11}$$

The incoming (anti)protons have momenta

$$p^\mu_{a,b} = E_{\text{cm}} \frac{n^\mu_{a,b}}{2} \quad \text{with} \quad n^\mu_a = (1, \hat{z}) \,, \quad n^\mu_b = (1, -\hat{z}) \,. \tag{12}$$

The Wilson coefficient $C_{ggH}(\mu)$ arises from matching onto the $\mathcal{O}_{ggH}(\mu)$ operator in SCET and encodes the hard-scattering contributions, which live at the hard scale $\mu_H \sim m_H$. The measurement operator $\widehat{\mathcal{M}}$ implements the phase-space cuts and measurements on the final state. The QCD dynamics in the soft and collinear limits, at leading order in the expansion parameter λ, are encoded in the SCET operator matrix element. At this point, the renormalization group evolution (RGE) of the hard Wilson coefficient can be used to sum logarithms of the form $\ln(\mu_{cs}/m_H) \sim \ln \lambda$ that arise as the ratio of some low soft-collinear scale $\mu_{sc} \sim \lambda m_H$ (set by the measurement in the matrix element) and the hard interaction scale $\sim m_H$.

The matrix element in Eq. (11) contains further logarithms of λ due to the different scaling of soft and collinear contributions. These logarithms are part of the full logarithmic structure of the cross section, and their resummation requires the separation of soft and collinear contributions. In SCET, this separation proceeds in two steps. First, the decoupling of soft emissions happens via a field redefinition of the collinear quark and gluon fields [25], after which $\mathcal{O}_{ggH}$ takes the form

$$\mathcal{O}_{ggH} = H\,\mathcal{O}_a\,\mathcal{O}_s\,\mathcal{O}_b = H\,\mathcal{B}^{\mu}_{n_a\perp} T\big[\mathcal{Y}^{\dagger}_{n_a}\mathcal{Y}_{n_b}\big]\mathcal{B}_{n_b\perp\mu}\,. \quad (13)$$

Here, H is the Higgs field, $\mathcal{O}_a$ and $\mathcal{O}_b$ are collinear gluon fields ($\mathcal{B}_{n\perp}$) in the forward (n_a) and backward (n_b) beam directions, and $\mathcal{O}_s$ is an operator of soft lightlike Wilson lines ($\mathcal{Y}_n$) along the n_a and n_b directions. After the soft-collinear decoupling, the SCET Lagrangian contains no interactions between soft and collinear fields at leading order in λ.

Second, the measurement $\widehat{\mathcal{M}}$ must be separated into collinear and soft components that act independently on the soft and collinear final states. Schematically,

$$\widehat{\mathcal{M}} = \widehat{\mathcal{M}}_a \times \widehat{\mathcal{M}}_b \times \widehat{\mathcal{M}}_s + \delta\widehat{\mathcal{M}}\,. \quad (14)$$

The operators $\widehat{\mathcal{M}}_a$, $\widehat{\mathcal{M}}_b$, and $\widehat{\mathcal{M}}_s$ are obtained by restricting $\widehat{\mathcal{M}}$ to act only on n_a-collinear, n_b-collinear, and soft fields respectively. The total contribution from the remainder $\delta\widehat{\mathcal{M}}$ must be power suppressed in λ (such that to all orders it can contribute at most terms that scale like $\lambda\ln^n\lambda$).

With these two ingredients, the matrix element in Eq. (11) factorizes into independent soft and collinear matrix elements,

$$\begin{aligned} B_a(\mu) &\sim \big\langle p_a\big|\mathcal{O}_a^{\dagger}\,\widehat{\mathcal{M}}_a\,\mathcal{O}_a\big|p_a\big\rangle(\mu)\,,\\ B_b(\mu) &\sim \big\langle p_b\big|\mathcal{O}_b^{\dagger}\,\widehat{\mathcal{M}}_b\,\mathcal{O}_b\big|p_b\big\rangle(\mu)\,,\\ S(\mu) &\sim \big\langle 0\big|\mathcal{O}_s^{\dagger}\,\widehat{\mathcal{M}}_s\,\mathcal{O}_s\big|0\big\rangle(\mu)\,. \end{aligned} \quad (15)$$

Here S is a soft function and B is a beam function [3, 28, 29], which describes the collinear initial-state radiation from the parton entering the hard interaction, and can be calculated perturbatively by matching onto the standard parton distribution functions [see Eq. (B3)]. It follows that the cross section factorizes as well,

$$\mathrm{d}\sigma_{gg\to H} \sim |C_{ggH}(\mu)|^2\,\big[B_a(\mu)\times B_b(\mu)\times S(\mu)\big]\,. \quad (16)$$

The collinear and soft matrix elements are renormalized objects and depend on a renormalization (or separation) scale μ. The RGE running between the natural collinear and soft scales then resums logarithms of the form $\ln(\mu_c/\mu_s)\sim\ln\lambda$ that are present in the soft-collinear matrix element itself.

From this discussion, we can see that resumming the logarithmic series in $\ln\lambda$ requires an explicit separation of the degrees of freedom whose associated scale depends on λ. In the effective-field theory context the resummation is then achieved by RGE methods. The complete all-order resummation requires this separation to hold to all orders in perturbation theory. This is guaranteed for the hard-scattering factorization in Eq. (11), which essentially amounts to expanding QCD in the soft-collinear limit. It also holds for the soft-collinear operator decoupling, which is independent of the measurement. However, for the measurement function it means that the separation in Eq. (14) has to hold for any number of soft and collinear particles in the final state. As we will see below, this requirement is satisfied by the inclusive observables, but for the jet-based observables it provides a nontrivial constraint and only holds for $R\sim\lambda$.

In the following we discuss the soft-collinear factorization properties of our four observables and the resulting factorization formulas for the cross section. We consider the inclusive observables in Sec. II A and the jet-based ones in Sec. II B. In Appendix A we discuss the RGE constraints on the factorization formulas for these observables.

A. Inclusive Observables

The full measurement operator $\widehat{\mathcal{M}}$ entering Eq. (11) directly follows from the definitions of the observables in Eqs. (1) and (2). For example, for beam thrust its action on a given final state with a set of momenta $\{p_m\}$ is

$$\mathcal{M}(\mathcal{T}_B) = \delta\Big(\mathcal{T}_B - \sum_m |\vec{p}_{Tm}|\,e^{-|y_m-Y|}\Big)\,, \quad (17)$$

and analogously for E_T. Since the inclusive observables simply sum over particles, we can write them as a sum over separate contributions from n_a-collinear, n_b-collinear, and soft particles,

$$\begin{aligned} \mathcal{T}_B &= \mathcal{T}_{Ba} + \mathcal{T}_{Bb} + \mathcal{T}_{Bs}\,,\\ E_T &= E_{Ta} + E_{Tb} + E_{Ts}\,. \end{aligned} \quad (18)$$

This means there is no contribution $\delta\mathcal{M}$ at leading power which mixes the different sectors, and we can factorize the measurement function in terms of a convolution,

$$\begin{aligned} \mathcal{M}(k) = \int \mathrm{d}k_a\,\mathrm{d}k_b\,\mathrm{d}k_s\,\delta(k-k_a-k_b-k_s)\\ \times\,\mathcal{M}_a(k_a)\,\mathcal{M}_b(k_b)\,\mathcal{M}_s(k_s)\,, \end{aligned} \quad (19)$$

where k here stands for either $\mathcal{T}_B$ or E_T. In the case of $\mathcal{T}_B$ we have

$$\begin{aligned} \mathcal{M}_a(\mathcal{T}_{Ba}) &= \delta\Big(\mathcal{T}_{Ba} - \sum_{m\in n_a-\mathrm{coll}} |\vec{p}_{Tm}|\,e^{Y-y_m}\Big)\,,\\ \mathcal{M}_a(\mathcal{T}_{Bb}) &= \delta\Big(\mathcal{T}_{Bb} - \sum_{m\in n_b-\mathrm{coll}} |\vec{p}_{Tm}|\,e^{y_m-Y}\Big)\,,\\ \mathcal{M}_s(\mathcal{T}_{Bs}) &= \delta\Big(\mathcal{T}_{Bs} - \sum_{m\in\mathrm{soft}} |\vec{p}_{Tm}|\,e^{-|y_m-Y|}\Big)\,. \end{aligned} \quad (20)$$

Note that for the collinear sectors, the contribution from the respective opposite hemisphere does not contribute, because it is power-suppressed. Similarly, for E_T we have

$$\begin{aligned}
\mathcal{M}_a(E_{Ta}) &= \delta\Big(E_{Ta} - \sum_{m\in n_a-\text{coll}} |\vec{p}_{Tm}|\Big)\,,\\
\mathcal{M}_a(E_{Tb}) &= \delta\Big(E_{Tb} - \sum_{m\in n_b-\text{coll}} |\vec{p}_{Tm}|\Big)\,,\\
\mathcal{M}_s(E_{Ts}) &= \delta\Big(E_{Ts} - \sum_{m\in \text{soft}} |\vec{p}_{Tm}|\Big)\,. \qquad (21)
\end{aligned}$$

Using n_a and n_b we define light-cone coordinates,

$$p^\mu = (n_a\cdot p, n_b\cdot p, p_\perp)\,, \quad p^\mu = n_a\cdot p\,\frac{n_b^\mu}{2} + n_b\cdot p\,\frac{n_a^\mu}{2} + p_\perp^\mu\,. \qquad (22)$$

In terms of these, n_a-collinear particles have momentum scaling $p_a \sim m_H(\lambda^2, 1, \lambda)$, while n_b-collinear particles have momentum scaling $p_b \sim m_H(1, \lambda^2, \lambda)$. The soft and collinear contributions to $\mathcal{T}_B$ can be written as

$$\begin{aligned}
\mathcal{T}_{Ba} &= e^Y n_a\cdot P_a\,, \qquad \mathcal{T}_{Bb} = e^{-Y} n_b\cdot P_b\,,\\
\mathcal{T}_{Bs} &= \sum_{m\in\text{soft}} \min\{e^Y n_a\cdot p_m,\, e^{-Y} n_b\cdot p_m\}\,, \qquad (23)
\end{aligned}$$

where $P_{a,b}$ is the total momentum of all $n_{a,b}$-collinear final-state particles. Since $\mathcal{T}_B$ measures the small light-cone components, soft particles that contribute to the measurement of $\mathcal{T}_B$ have ultrasoft (us) momentum scaling $p_{us} \sim m_H(\lambda^2, \lambda^2, \lambda^2)$. The appropriate version of SCET for this case is called SCET_I. Since $p_{a,b}^2 \sim \lambda^2 m_H^2 \gg p_{us}^2 \sim \lambda^4 m_H^2$, soft and collinear degrees of freedom in SCET_I are separated in virtuality and the RGE running in SCET_I describes evolution in invariant mass.

The factorization formula for the beam thrust distribution is [3, 5]

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{T}_B} &= \sigma_0\, H_{gg}(m_H,\mu) \int \mathrm{d}Y \int \mathrm{d}k_a\,\mathrm{d}k_b\\
&\quad\times B_g(m_H k_a, x_a, \mu)\, B_g(m_H k_b, x_b, \mu)\\
&\quad\times S_B^{gg}(\mathcal{T}_B - k_a - k_b, \mu)\,, \qquad (24)
\end{aligned}$$

where

$$x_a = \frac{m_H}{E_\text{cm}}\, e^Y\,, \qquad x_b = \frac{m_H}{E_\text{cm}}\, e^{-Y}\,, \qquad (25)$$

and

$$\sigma_0 = \frac{\sqrt{2} G_F\, m_H^2}{576\pi E_\text{cm}^2}\,. \qquad (26)$$

The convolution between beam and soft functions in Eq. (24) is a direct consequence of the convolution of the measurement in Eq. (19).

In contrast to $\mathcal{T}_B$, E_T measures the perpendicular component of momentum, which means soft particles contributing to the measurement of E_T have soft momentum scaling $p_s \sim m_H(\lambda, \lambda, \lambda)$. The appropriate version of SCET for this case is called SCET_II. In SCET_II, soft and collinear degrees of freedom have the same virtuality scaling, $p_{a,b}^2 \sim p_s^2 \sim \lambda^2 m_H^2$, but are still separated in rapidity. As a result, the RGE running in SCET_II describes both evolution in invariant mass as well as rapidity [30, 31].

The factorization formula for the E_T distribution is [6]

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}E_T} &= \sigma_0\, H_{gg}(m_H,\mu) \int \mathrm{d}Y \int \mathrm{d}k_a\,\mathrm{d}k_b\\
&\quad\times B_g(m_H, k_a, x_a, \mu, \nu)\, B_g(m_H, k_b, x_b, \mu, \nu)\\
&\quad\times S_T^{gg}\Big(E_T - k_a - k_b, \mu, \nu\Big)\,. \qquad (27)
\end{aligned}$$

The soft and beam functions are different from those in Eq. (24) because they contain a different measurement. Beam functions with p_T-dependence were studied in SCET in Refs. [15, 30–36]. Since the beam and soft functions in Eq. (27) are renormalized and RG evolved in rapidity, they depend on another scale ν from this running.

Higher-order resummation at small $\mathcal{T}_B$ and E_T can be carried out systematically using Eqs. (24) and (27). Extending the $\mathcal{T}_B$ resummation to N^3LL requires determining the two-loop beam and soft functions, while resuming E_T to NNLL requires determining the two-loop beam and soft non-cusp anomalous dimensions.

B. Exclusive Observables

Vetoes on individual jets rather than an inclusive veto on the final state requires a more careful understanding of the role of the jet algorithm. When putting a cut on $\mathcal{T}_j$ or p_{Tj} in Eqs. (3) and (4), the full measurement operator acting on the complete final state in Eq. (11) is

$$\begin{aligned}
\mathcal{M}^\text{jet}(\mathcal{T}^\text{cut}) &= \prod_{m\in j(R)} \theta\left(|\vec{p}_{Tm}| e^{-|y_m - Y|} < \mathcal{T}^\text{cut}\right)\,,\\
\mathcal{M}^\text{jet}(p_T^\text{cut}) &= \prod_{m\in j(R)} \theta\left(|\vec{p}_{Tm}| < p_T^\text{cut}\right)\,, \qquad (28)
\end{aligned}$$

where we consider jets $j(R)$ defined by the k_T class of algorithms. Note that integrating $\mathcal{T}_j$ or p_{Tj} up to a cut turns the maximum condition in their definitions in Eqs. (3) and (4) into the simple product of θ functions in Eq. (28). Hence, for the exclusive variables it is more convenient to consider the integrated cross section with a cut rather than the differential spectrum as in the inclusive variables.

We now want to separate the measurement function into components that act independently on soft and collinear final states. That is, we want to express the full jet veto in Eq. (28) as

$$\begin{aligned}
\mathcal{M}^\text{jet}(k^\text{cut}) &= \mathcal{M}_a^\text{jet}(k^\text{cut})\, \mathcal{M}_b^\text{jet}(k^\text{cut})\, \mathcal{M}_s^\text{jet}(k^\text{cut})\\
&\quad + \delta\mathcal{M}^\text{jet}(k^\text{cut})\,, \qquad (29)
\end{aligned}$$

where here and in the remainder of this section, $k^{\rm cut}$ stands for either $\mathcal{T}^{\rm cut}$ or $p_T^{\rm cut}$. The functions $\mathcal{M}_i^{\rm jet}(k^{\rm cut})$ for $i=a,b,s$ are defined by the full measurement applied to n_a-collinear, n_b-collinear, and soft particles, respectively, which also defines the remainder $\delta\mathcal{M}^{\rm jet}(k^{\rm cut})$. For this separation to be meaningful, the remainder should be power-suppressed, which requires that the full measurement $\mathcal{M}^{\rm jet}$ does not mix constraints on collinear and soft particles. This is a nontrivial condition, since the veto on any individual jet is not allowed to mix constraints between sectors. If a jet has a collinear component p_c and a soft component p_s, then the veto condition, e.g. for $p_T^{\rm cut}$, is

$$|\vec{p}_{Tc}+\vec{p}_{Ts}|<p_T^{\rm cut}\,. \tag{30}$$

This prevents $\mathcal{M}^{\rm jet}$ from factorizing into separate soft and collinear components and gives a contribution to $\delta\mathcal{M}^{\rm jet}$. Therefore, to preserve factorization this scenario should only happen with a power-suppressed rate, such that at leading power each jet contains either only soft or only collinear final states. In that case, we can perform the veto separately on jets in each sector, i.e., we can write the product over all jets in Eq. (28) as products over soft and collinear jets as in Eq. (29).

We shall show that for $R\sim\lambda$, $\delta\mathcal{M}^{\rm jet}$ indeed gives a power-suppressed contribution to the rate. In this limit, the following factorization formula holds for the $\mathcal{T}_j$ veto:

$$\begin{aligned}\sigma(\mathcal{T}^{\rm cut}) = \sigma_0 H_{gg}(m_H,\mu)\int dY\, B_g^{\rm jet}\left(m_H\mathcal{T}^{\rm cut},x_a,\mu\right)\\ \times B_g^{\rm jet}\left(m_H\mathcal{T}^{\rm cut},x_b,\mu\right) S_{gg}^{\rm jet}(\mathcal{T}^{\rm cut},\mu)\,,\end{aligned} \tag{31}$$

where x_a, x_b and σ_0 are defined in Eqs. (25) and (26). In the same limit, an analogous factorization formula holds for the p_{Tj} veto:

$$\begin{aligned}\sigma(p_T^{\rm cut}) = \sigma_0 H_{gg}(m_H,\mu)\int dY\, B_g^{\rm jet}\left(m_H,p_T^{\rm cut},x_a,\mu,\nu\right)\\ \times B_g^{\rm jet}\left(m_H,p_T^{\rm cut},x_b,\mu,\nu\right) S_{gg}^{\rm jet}(p_T^{\rm cut},\mu,\nu)\,.\end{aligned} \tag{32}$$

An equivalent form of this factorization formula was derived in Ref. [15], and the NLL resummation for $p_T^{\rm cut}$ was performed in Refs. [14, 15].

In Ref. [14] jet-algorithm dependent effects were calculated at fixed $\mathcal{O}(\alpha_s^2)$. These results were interpreted in Ref. [15] in terms of a two-loop anomalous dimension and used to extend the resummation based on Eq. (32) to NNLL, working in the limit $\lambda\ll R\sim 1$ to avoid $\ln R$ clustering logarithms. However, we will show explicitly in the following that for $R\sim 1$ Eqs. (31) and (32) do not reproduce the all-order structure of QCD beyond NLL. We will also see that for $R\sim\lambda$, where the factorization formula holds, only some parts of the $\mathcal{O}(\alpha_s^2)$ contributions from Ref. [14] are correctly interpreted in terms of anomalous dimensions and used in conjunction with Eq. (32).

Since $\mathcal{M}^{\rm jet}$ in Eq. (29) is a simple product, and not a convolution as for the inclusive variables in Eq. (19), the factorized cross sections in Eqs. (31) and (32) now contain a product of beam and soft functions rather than a convolution. Also, each function explicitly depends on the jet algorithm used, in addition to the jet-veto variable itself, and includes R-dependent clustering effects. Note that the $\mathcal{O}(\alpha_s)$ results do not yet depend on the effects of the jet algorithm. Nevertheless, the resummed cross sections are different for the inclusive and exclusive observables starting at NLL because of the different structures of their factorization theorems. This reflects the fact that constraining the sum of emissions provides a very different phase-space constraint than constraining each individual emission for more than one emission.

To understand the role of the jet algorithm in vetoing on individual jets and how it impacts the logarithmic series, it is useful to express the measurement function in the form

$$\begin{aligned}\mathcal{M}^{\rm jet} = &\left(\mathcal{M}_a+\Delta\mathcal{M}_a^{\rm jet}\right)\left(\mathcal{M}_b+\Delta\mathcal{M}_b^{\rm jet}\right)\left(\mathcal{M}_s+\Delta\mathcal{M}_s^{\rm jet}\right)\\ &+\delta\mathcal{M}^{\rm jet}\,,\end{aligned} \tag{33}$$

where $\Delta\mathcal{M}_i^{\rm jet}$ is defined to contain the jet-algorithm dependence within each of the collinear and soft sectors,

$$\mathcal{M}_i^{\rm jet}(k^{\rm cut}) = \mathcal{M}_i(k^{\rm cut}) + \Delta\mathcal{M}_i^{\rm jet}(k^{\rm cut})\,, \tag{34}$$

for $i=a,b,s$. The definition of $\Delta\mathcal{M}_i^{\rm jet}$ is subtle, since it depends on what we define the corrections due to clustering *relative* to; namely it depends on the precise choice of $\mathcal{M}_i$, which is independent of R and the jet algorithm. To study the effect of clustering we choose $\mathcal{M}_i$ to be the inclusive $\mathcal{T}_B$ or E_T measurement,

$$\mathcal{M}_i(k^{\rm cut}) = \theta\Bigg(\sum_{m\in i^{th}-{\rm sector}} k_m < k^{\rm cut}\Bigg)\,, \tag{35}$$

where k_m is the $\mathcal{T}_B$ or E_T contribution from each particle.

At $\mathcal{O}(\alpha_s)$ there are no jet-algorithm effects, since there is only a single, either soft or collinear, final-state particle. This means that $\delta\mathcal{M}^{(1)}$ and $\Delta\mathcal{M}_i^{(1)}$ are zero, and $\mathcal{M}^{{\rm jet}\,(1)}$ reduces to a sum over terms with one nontrivial constraint in each sector,

$$\mathcal{M}^{{\rm jet}\,(1)}(k^{\rm cut}) = \sum_{i=a,b,s}\theta(k_i<k^{\rm cut})\,. \tag{36}$$

This gives the same $\mathcal{O}(\alpha_s)$ contribution as the integral over the $k=\mathcal{T}$ or E_T distribution with $k<k^{\rm cut}$.

Starting at $\mathcal{O}(\alpha_s^2)$ the role of the jet algorithm must be understood. When $R\gg\lambda$, soft-collinear mixing effects are important. They give a correction to the cross section of the form

$$\begin{aligned}\delta\sigma_{SC}^{\rm jet} \sim\, &|C_{ggH}(\mu)|^2\\ &\times\left\langle p_a p_b\right|\mathcal{O}_{ggH}(\mu)^\dagger\delta\widehat{\mathcal{M}}^{\rm jet}\,\mathcal{O}_{ggH}(\mu)\left|p_a p_b\right\rangle\,,\end{aligned} \tag{37}$$

FIG. 1: (a) The contribution to soft-collinear mixing at α_s^2 from the independent emission of a collinear and soft gluon clustered into a single jet, relevant when $R \gg \lambda$. (b) Clustering corrections in the beam and soft functions from correlated emissions, relevant when $R \sim \lambda$.

where the measurement operator $\delta\widehat{\mathcal{M}}^{\rm jet}$ is defined by Eq. (29) as the difference between the full measurement and its restriction to the soft and collinear sectors. At $\mathcal{O}(\alpha_s^2)$, the soft-collinear mixing contribution arises from the independent emission of a collinear and soft particle as shown in Fig. 1(a). It is given by the measurement function

$$\delta\mathcal{M}^{{\rm jet}\,(2)}(k^{\rm cut}) = \delta\mathcal{M}_{as}^{{\rm jet}\,(2)}(k^{\rm cut}) + \delta\mathcal{M}_{bs}^{{\rm jet}\,(2)}(k^{\rm cut})\,, \quad (38)$$

where the two terms correspond to a soft particle clustering with either a n_a-collinear or n_b-collinear particle. For the $k_{\rm T}$-class algorithms, which cluster particles with $\Delta R < R$,

$$\begin{aligned}\delta\mathcal{M}_{as}^{{\rm jet}\,(2)}(k^{\rm cut}) = \theta(\Delta R_{as} < R)\,\big[&\theta(k_{\rm jet} < k^{\rm cut}) \\ &- \theta(k_a < k^{\rm cut})\theta(k_s < k^{\rm cut})\big]\,, \end{aligned} \quad (39)$$

where $k_{\rm jet}$ is the value of the jet observable after combining the soft and collinear particles. Since $\delta\mathcal{M}^{\rm jet}$ is defined as the correction to the factorized measurement, the second term subtracts the corresponding $\mathcal{O}(\alpha_s^2)$ contribution from $\mathcal{M}_a^{{\rm jet}(1)}\mathcal{M}_s^{{\rm jet}(1)}$.

In Sec. III, we calculate the (leading) contribution from Eq. (39) to the cross section explicitly, and we will see that it has the form $\sim \alpha_s^2 R^2 \ln(m_H/k^{\rm cut})$ for both $\mathcal{T}_j$ and p_{Tj} vetoes. There are also factorized clustering corrections of the same form in $\Delta\mathcal{M}_i^{\rm jet}$ (for $i = a, b, s$) from two independent emissions within each sector, and which are part of the two-loop soft and beam functions. The total QCD contribution from independent emissions thus arises as the sum of the factorized contributions and the soft-collinear mixing contribution. This shows explicitly that the factorized cross sections in Eqs. (31) and (32) do not reproduce the full NNLL structure of QCD for general $R \sim 1$, since they do not contain the soft-collinear mixing contributions.

When $R \sim \lambda$, clustering corrections from independent emissions, including the soft-collinear mixing term $\delta\mathcal{M}^{\rm jet}$, can be regarded as power corrections. In this limit a different type of clustering corrections in $\Delta\mathcal{M}_i^{\rm jet}$ become important, arising from correlated emissions within the beam and soft functions, as shown in Fig. 1(b). At $\mathcal{O}(\alpha_s^2)$ they are described by the measurement function

$$\begin{aligned}\Delta\mathcal{M}_i^{{\rm jet}\,(2)}(k^{\rm cut}) &= \Big\{\theta(\Delta R_{12} < R)\theta(k_{\rm jet} < k^{\rm cut}) \\ &\quad + \theta(\Delta R_{12} > R)\theta(k_1 < k^{\rm cut})\theta(k_2 < k^{\rm cut})\Big\} \\ &\quad - \theta(k_1 + k_2 < k^{\rm cut})\,, \end{aligned} \quad (40)$$

where particles 1 and 2 both belong to sector i. In Sec. IV we calculate the contributions from Eq. (40), and we will see that they have the form $\alpha_s^2 \ln R \ln(m_H/k^{\rm cut})$. We also discuss the higher-order structure of clustering logarithms and argue that at $\mathcal{O}(\alpha_s^n)$ terms of the form $\alpha_s^n \ln^{n-1} R \ln(m_H/k^{\rm cut})$ contribute in the exponent of the cross section. For $R \sim \lambda$ these terms are NLL. Since the coefficients of the clustering logarithms at each order contain a genuinely new algorithm-dependent contribution, which will generically be unrelated to lower orders, they cannot be resummed with present methods [21].

1. *Rapidity Cutoffs on Jets*

In an experiment, the physical limitations of the detector impose a rapidity cutoff on measured jets. The parton luminosities naturally suppress forward, high-p_T jets, but in experimental analyses even moderate rapidity cutoffs (e.g., a cutoff of $y_{\rm cut} = 2.5$) are used. Monte Carlo studies of the dependence on the cutoff with a p_T veto on jets find a negligible effect for $y_{\rm cut} \sim 4.5$ but an $\mathcal{O}(10\%)$ effect for $y_{\rm cut} \sim 2.5$ in the range of typical $p_T^{\rm cut}$ values [5, 14].

The rapidity cutoff $y_{\rm cut}$ regulates the rapidity divergences present in the soft and beam functions for the p_{Tj} veto. In the bare functions, this effectively converts rapidity divergences into factors of $y_{\rm cut}$. At $\mathcal{O}(\alpha_s)$, the RG structure of the hard, beam, and soft functions implies that these divergences have no effect on the fixed-order logarithms. The first order that the effect of the rapidity cutoff on the resummation can be observed is at $\mathcal{O}(\alpha_s^2)$, and, in principle, the cutoff could affect the $\mathcal{O}(\alpha_s^2 \ln^2\lambda)$ and $\mathcal{O}(\alpha_s^2 \ln\lambda)$ terms. Although we do not consider these effects here, it would be interesting to study their impact on the logarithmic structure and resummation in more detail.

III. SOFT-COLLINEAR MIXING

The jet algorithm used to define $\mathcal{T}_j$ and p_{Tj} gives rise to jets that can contain both soft[3] and collinear particles, and the collinear beam radiation can give rise to multiple

[3] Our discussion in this section is mostly insensitive to whether the soft radiation is described by ultrasoft modes (a $\mathcal{T}_j$ veto, SCET$_{\rm I}$) or soft modes (a p_{Tj} veto, SCET$_{\rm II}$), but we will point out when subtleties arise.

jets. Placing a veto on such jets mixes the phase space constraints in each sector through conditions of the form

$$\mathcal{T}_c + \mathcal{T}_s \leq \mathcal{T}^{\rm cut}\,, \qquad |\vec{p}_{Tc} + \vec{p}_{Ts}| \leq p_T^{\rm cut}\,, \qquad (41)$$

and this mixing contributes to $\delta\mathcal{M}^{\rm jet}$ in Eq. (29). In this section we show that these soft-collinear mixing contributions scale like R^2, and so R must scale as λ for these effects to be power suppressed.

Clustering jet algorithms build jets by merging particles in the final state according to a distance metric ρ_{ij} between particles and a metric ρ_i for single particles. In each clustering step, the minimum metric determines the action of the algorithm. If the minimum is a pairwise metric then that pair is merged into a new particle, and if the minimum is a single particle then that particle is promoted to a jet. For the anti-$k_{\rm T}$ algorithm [13], these metrics are

$$\rho_{ij} = \min\bigl(p_{Ti}^{-1}, p_{Tj}^{-1}\bigr)\frac{\Delta R_{ij}}{R}\,, \qquad \rho_i = p_{Ti}^{-1}\,. \qquad (42)$$

The algorithm will cluster particles with separation $\Delta R_{ij} < R$ together.[4]

Phase-space constraints from jet algorithms have been studied in SCET in Refs. [37–40]. By applying the canonical scaling of soft and collinear particles to the anti-$k_{\rm T}$ metrics in Eq. (42) we can determine the *typical* behavior of the algorithm. For example, for the case of a veto on the jet p_T, the metrics scale as

$$\rho_{cc} \sim \rho_{ss} \sim \frac{\lambda^{-1}}{R}\,, \quad \rho_{sc} \sim \frac{\lambda^{-1}}{R}\ln\frac{1}{\lambda}\,, \quad \rho_c \sim \rho_s \sim \lambda^{-1}\,. \qquad (43)$$

This naive power counting suggests that clustering of soft and collinear particles will typically not occur, since soft particles are at central rapidities and collinear particles are at large rapidities and as a result $\rho_{sc} \gg \{\rho_{cc}, \rho_{ss}, \rho_c, \rho_s\}$ (for $R \sim 1$ or smaller). The ordering $\rho_{sc} > \rho_{ss}, \rho_{cc}$ implies that collinear and soft particles will separately cluster among themselves before clustering with each other. The ordering $\rho_{sc} > \rho_c, \rho_s$ implies that (groups of) collinear and soft particles will be promoted to jets before any soft-collinear clustering between them can take place. A similar argument applies to measuring $\mathcal{T}_j$.

Power counting along these lines, counting $\lambda \ll R \ll \ln(1/\lambda)$, was used in Ref. [15] to argue that, at leading power, soft and collinear particles do not occupy the same jet. However, the canonical scaling in Eq. (43) is not sufficient to demonstrate that soft collinear clustering gives only a power-suppressed contribution to the cross section [41]. In particular, there is a collinear enhancement in soft emissions along the beam directions (i.e., the direction of the soft Wilson lines). This is straightforward to see from the matrix element for eikonal emission,

$$\mathcal{A}_{\rm eikonal} \sim \frac{{\rm d}p_T}{p_T}\,{\rm d}y\,{\rm d}\phi\,. \qquad (44)$$

The matrix element is flat in rapidity, meaning soft particles will populate jets at a forward rapidities at an equal rate to jets at central rapidities. For a jet j in a given direction, the rate for a soft particle to be radiated within the cone of the jet scales as

$$\int_{-\infty}^{\infty}{\rm d}y_s \int_0^{\pi}\frac{{\rm d}\phi_s}{\pi}\,\theta(\Delta R_{js} < R) \sim R^2\,. \qquad (45)$$

If $R \sim \lambda$, this rate is power suppressed and can be neglected. We will see this scaling confirmed in the explicit calculations of soft-collinear mixing below.

The geometry of jets at a hadron collider changes significantly in the forward region. For central rapidities, the clustering condition $\Delta R_{ij} < R$ is approximated by $\theta_{ij} < R$, where θ is the physical angle between i and j. At forward rapidities, the jet subtends a smaller angle R_θ with respect to the beam axis. The size of the jets in ϕ remains unchanged, so jets become elongated at forward rapidities. A comparison between jets in $y-\phi$ space and $\theta-\phi$ space is shown in Fig. 2. If a jet is at rapidity y_j with a corresponding angle to the beam θ_j, then

$$R_\theta \approx 2\theta_j \sinh R\,. \qquad (46)$$

At the characteristic angle $\theta_c \sim \lambda$ of collinear emissions, the size of the jet scales approximately as λR. This means that collinear initial-state radiation from the incoming hard partons can create multiple jets of size R in the final state, since the size of the jet is smaller than (potentially parametrically smaller than) the total size of the collinear beam sector, which is an angle of order λ around the beam direction. This is important because if each collinear sector created a single jet (around the beam), a veto on jets containing both soft and collinear particles would require $\delta\mathcal{M}^{\rm jet}$ to only depend on the total

FIG. 2: The shape of jets in two spaces: $y-\phi$ (left), the geometry where jets are found, and $\theta-\phi$ (right), which is the physical geometry. As jets become forward, their size in θ shrinks. In the right figure, the dashed blue line shows the approximate shape of jets if ΔR_{ij} is replaced with the physical angle θ_{ij} between particles.

[4] In some cases, a clustering sequence can pull two particles farther than R away from each other. Such configurations tend to be uncommon.

collinear momentum and the measurement could therefore be expressed in a factorizable way. However this is not the case here. From this analysis, we see that jets can contain both soft and collinear particles. Vetoing on such jets requires a constraint on the combined soft and collinear contribution to the veto variable, as in Eq. (41). Therefore, soft-collinear factorization will be inhibited by these jets for $R \sim 1$.

We can calculate the correction to the cross section from soft-collinear mixing, $\delta\sigma^{\rm jet}_{SC}$ in Eq. (37), for both jet-based veto observables at $\mathcal{O}(\alpha_s^2)$. The mixing correction is given by the measurement function $\delta\mathcal{M}^{\rm jet\,(2)}$ in Eqs. (38) and (39) inserted into matrix elements of $\mathcal{O}_{ggH}$ with one soft and one collinear emission. Because of the soft-collinear decoupling this is simply the product of the $\mathcal{O}(\alpha_s)$ matrix elements for one soft and one collinear emission.

The form of the soft-collinear mixing terms are the same as the $\mathcal{O}(\alpha_s^2)$ corrections when two independent emissions within either the soft or collinear sectors cluster (which is contained in $\Delta\mathcal{M}_i^{\rm jet}$). Therefore, the total correction from clustering of two independent emissions to the cross section is given by

$$\Delta\sigma^{\rm indep\,(2)} = \Delta\sigma^{\rm indep\,(2)}_{SS} + \Delta\sigma^{\rm indep\,(2)}_{CC} + \delta\sigma^{\rm jet\,(2)}_{SC}\,, \quad (47)$$

where SS, CC, and SC denote the independent emission contributions from the soft sector, collinear sector, and from soft-collinear mixing, respectively. In addition to the soft-collinear mixing term, we also calculate the correction from the soft sector, $\Delta\sigma^{\rm indep}_{SS}$.[5] The calculations are given in Appendix B, where we work to leading power in R, $\mathcal{O}(R^2)$, and drop corrections of $\mathcal{O}(R^4)$.

The $\mathcal{T}_j$ veto is easier to calculate, as dimensional regularization fully handles all divergences. The soft-collinear mixing, $\delta\sigma^{\rm jet}_{SC}$, is the sum of contributions from each beam, $\delta\sigma^{\rm jet}_{SC_a} + \delta\sigma^{\rm jet}_{SC_b}$ and is given in Eq. (B27). This includes a nontrivial zero-bin subtraction in the collinear sector. The two contributions $\delta\sigma^{\rm jet}_{SC_a}$ and $\delta\sigma^{\rm jet}_{SC_b}$ are identical, and their sum is

$$\delta\sigma^{\rm jet\,(2)}_{SC}(\mathcal{T}^{\rm cut}) = -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{3}R^2\ln\frac{m_H}{\mathcal{T}^{\rm cut}}\,. \quad (48)$$

Here $\sigma_{\rm LO}$ is the leading-order cross section, which is given in Eq. (B6). We can see explicitly that the soft-collinear mixing contains a single logarithm and scales as R^2, as expected from the naive scaling calculation in Eq. (45). The soft independent emission contribution, see Eq. (B12), is

$$\Delta\sigma^{\rm indep\,(2)}_{SS}(\mathcal{T}^{\rm cut}) = -\sigma_{\rm LO}\frac{1}{\epsilon}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\bigg(\frac{\mu^2}{\mathcal{T}^{\rm cut\,2}}\bigg)^{2\epsilon}\frac{\pi^2}{12}R^2\,. \quad (49)$$

[5] Determining the $\mathcal{O}(\alpha_s^2)$ beam function matrix elements is significantly more involved, and is not required to demonstrate the presence of soft-collinear mixing. It is therefore beyond the scope of this work.

The remaining contribution at this order is from the independent emission of two collinear particles, $\Delta\sigma^{\rm indep}_{CC}$, whose form can be constrained by two facts. First, since the effective theory reproduces the IR structure of QCD, the collinear contribution must cancel the $1/\epsilon$ dependence in $\Delta\sigma^{\rm indep}_{SS}$. Second, the scale dependence of the collinear matrix element can be determined by power counting. It follows that the $\Delta\sigma^{\rm indep}_{CC}$ contribution must have the form

$$\Delta\sigma^{\rm indep\,(2)}_{CC}(\mathcal{T}^{\rm cut}) = \sigma_{\rm LO}\frac{1}{\epsilon}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\bigg(\frac{\mu^2}{m_H\,\mathcal{T}^{\rm cut}}\bigg)^{2\epsilon}\frac{\pi^2}{12}R^2\,. \quad (50)$$

Combining all three contributions, the leading R-dependent clustering effect from independent emissions with a cut on $\mathcal{T}_j$ is given by

$$\sigma^{\rm indep\,(2)}(\mathcal{T}^{\rm cut}) = -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{2}R^2\ln\frac{m_H}{\mathcal{T}^{\rm cut}}\,. \quad (51)$$

For the p_{Tj} veto, the total $\mathcal{O}(\alpha_s^2)$ clustering effect from independent emissions in QCD, which includes the soft-collinear contributions, was calculated in Ref. [14] and found to be

$$\Delta\sigma^{\rm indep\,(2)}(p_T^{\rm cut}) = -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{3}R^2\,\ln\frac{m_H}{p_T^{\rm cut}}\,. \quad (52)$$

This result will serve as a partial cross check on our results. To calculate the soft-collinear mixing contribution for the p_{Tj} veto, we have to regulate rapidity divergences. We use the analytic regulator [42] in this case, for which ν plays the same role as μ in dimensional regularization, and α the role of ϵ. The soft-collinear mixing contributions are different for $\delta\sigma^{\rm jet}_{SC_a}$ and $\delta\sigma^{\rm jet}_{SC_b}$, because of the asymmetry in the regulator between the two collinear sectors. We find

$$\begin{aligned}\delta\sigma^{\rm jet\,(2)}_{SC}(p_T^{\rm cut}) &= \delta\sigma^{\rm jet\,(2)}_{SC_a} + \delta\sigma^{\rm jet\,(2)}_{SC_b}\\ &= -\sigma_{\rm LO}\frac{1}{\alpha}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{6}R^2\\ &\quad\times\bigg[\bigg(\frac{\nu\,m_H}{(p_T^{\rm cut})^2}\bigg)^{2\alpha} - \bigg(\frac{\nu}{m_H}\bigg)^{2\alpha}\bigg]\\ &= -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{2\pi^2}{3}R^2\,\ln\frac{m_H}{p_T^{\rm cut}}\,. \quad (53)\end{aligned}$$

As for the $\mathcal{T}_j$ veto, the soft-collinear mixing for the p_{Tj} veto contains a single logarithm and scales as R^2. The soft independent emission contribution for p_{Tj} is scaleless and thus vanishes

$$\Delta\sigma^{\rm indep\,(2)}_{SS}(p_T^{\rm cut}) = 0\,. \quad (54)$$

Hence, to reproduce the full independent emission result in Eq. (52), the collinear contribution must be

$$\Delta\sigma^{\rm indep\,(2)}_{CC}(p_T^{\rm cut}) = \sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{3}R^2\,\ln\frac{m_H}{p_T^{\rm cut}}\,. \quad (55)$$

The soft-collinear mixing contributions in Eqs. (48) and (53) are not isolated to $\mathcal{O}(\alpha_s^2)$. At higher orders, additional emissions can generate a tower of Sudakov double logarithms on top of these results and also generate higher-order mixing effects. Accurate uncertainty estimates require a better understanding of the soft-collinear mixing terms at all orders.

IV. CLUSTERING LOGARITHMS

We now turn to the form of clustering logarithms in the cross section. When $R \sim \lambda$, the cross section $\sigma(k^{\rm cut})$ satisfies the factorization formulas in Eqs. (31) and (32). Clustering of final state soft and collinear particles can give rise to logarithms of R that become important when $R \ll 1$. These logarithms arise as the remnant of a collinear divergence between particles in the jet, and are associated with connected webs (c-webs) in the matrix element [43, 44].

To see this, consider a two-parton final state with a p_{Tj} veto. The measurement function can be written as a constraint on each particle's p_T plus a correction factor for when they are clustered into a jet,

$$\mathcal{M}^{\text{jet }(2)} = \mathcal{M}^{(2)}_{\rm veto} + \Delta\mathcal{M}^{\text{jet }(2)}_{\rm veto}\,, \tag{56}$$

where

$$\begin{aligned}
\mathcal{M}^{(2)}_{\rm veto} &= \theta(p_{T1} < p_T^{\rm cut})\,\theta(p_{T2} < p_T^{\rm cut})\,,\\
\Delta\mathcal{M}^{\text{jet }(2)}_{\rm veto} &= \theta(\Delta R_{12} < R)\Big[\theta(|\vec{p}_{T1}+\vec{p}_{T2}| < p_T^{\rm cut})\\
&\quad - \theta(p_{T1} < p_T^{\rm cut})\,\theta(p_{T2} < p_T^{\rm cut})\Big]\,.
\end{aligned} \tag{57}$$

Each individual measurement, $\mathcal{M}^{(2)}_{\rm veto}$ and $\Delta\mathcal{M}^{\text{jet }(2)}_{\rm veto}$, is separately IR unsafe if the matrix element has a collinear singularity between the two particles. Since the total measurement, $\mathcal{M}^{\text{jet }(2)}$, is IR safe, the divergence cancels. The remnant is logarithmic sensitivity to R, schematically

$$\begin{aligned}
&\mathcal{M}^{(2)}_{\rm veto} \sim \frac{1}{\epsilon}\,, \qquad \Delta\mathcal{M}^{\text{jet }(2)}_{\rm veto} \sim -\frac{1}{\epsilon}R^{-2\epsilon}\,,\\
&\mathcal{M}^{(2)}_{\rm veto} + \Delta\mathcal{M}^{\text{jet }(2)}_{\rm veto} \sim \ln R\,.
\end{aligned} \tag{58}$$

This structure persists for more particles in the final state. In general, if there are n final-state particles, there are at most $n-1$ collinear singularities between them, each of which leads to a factor of $\ln R$ from clustering effects. We will show that the general form of the leading clustering logarithm at $\mathcal{O}(\alpha_s^n)$ is

$$\sigma_{\rm LO}\left(\frac{\alpha_s}{\pi}\right)^n C^{(n)}_{p_T}(\ln R)\,\ln\frac{m_H}{p_T^{\rm cut}}\,, \tag{59}$$

where $C^{(n)}_{p_T}(\ln R)$ contains at most $n-1$ logarithms of R and the same form holds for $\mathcal{T}^{\rm cut}$.

A. $\mathcal{O}(\alpha_s^2)$ Clustering Logarithms

Although the division in Eq. (56) helps determine the order of the clustering logarithms, it is not well suited to define the clustering correction. Instead, we use the division in Sec. II B, defining the clustering logarithms relative to the inclusive measurement.

At $\mathcal{O}(\alpha_s^2)$, the clustering logarithms for the p_T veto have been calculated in Ref. [14]. We perform the soft function clustering calculation for both p_{Tj} and $\mathcal{T}_j$ in Appendix C, since it is instructive to see the RG structure explicitly. The bare soft function corrections for $\mathcal{T}_j$ are given in Eq. (C18) and for p_{Tj} in Eq. (C17), where for the latter we use the rapidity regulator [30, 31]. The bare corrections are UV-divergent and give a contribution to the two-loop soft anomalous dimensions of

$$\begin{aligned}
\Delta\gamma_S^{(2)}(\mathcal{T}^{\rm cut},\mu) &= \left(\frac{\alpha_s}{\pi}\right)^2 C^{(2)}_{\mathcal{T}}(\ln R)\,,\\
\Delta\gamma_S^{\nu(2)}(p_T^{\rm cut},\nu) &= \left(\frac{\alpha_s}{\pi}\right)^2 C^{(2)}_{p_T}(\ln R)\,,
\end{aligned} \tag{60}$$

where the coefficients are given by

$$\begin{aligned}
C^{(2)}_{\mathcal{T}}(\ln R) &= C_A^2\Big(\frac{131-12\pi^2-132\ln 2}{18}\ln R - 0.936\Big)\\
&\quad + C_A T_R n_f\Big(-\frac{23-24\ln 2}{9}\ln R + 0.748\Big)\,,\\
C^{(2)}_{p_T}(\ln R) &= C_A^2\Big(\frac{131-12\pi^2-132\ln 2}{18}\ln R - 1.12\Big)\\
&\quad + C_A T_R n_f\Big(-\frac{23-24\ln 2}{9}\ln R + 0.764\Big)\,.
\end{aligned} \tag{61}$$

For the p_T veto, the contribution to the cross section stemming from this coefficient agrees with the results in Ref. [14] for the $\ln R$ term. The constant terms depend the observable the clustering effect is defined relative to. We use the inclusive E_T measurement, as in Eq. (40), for this purpose, whereas Ref. [14] uses the total p_T, which for two particles is $p_T = |\vec{p}_{T1}+\vec{p}_{T2}|$. Either choice is possible and we have checked that we reproduce the constant terms in Ref. [14] when alternatively using their definition of the clustering correction.

The soft anomalous dimension from these clustering effects must be canceled by the anomalous dimensions of the beam functions, since the total $\mathcal{O}_{ggH}$ matrix element has only UV divergences that match the hard function, and which are unrelated to clustering. It is worthwhile to note that since no collinear singularities exist between soft and collinear particles, the cancellation of the divergences in the soft function from clustering must come entirely from the beam functions. It cannot come from a soft-collinear mixing term, which is power suppressed when $R \sim \lambda$. The cancellation of the beam and soft anomalous dimensions from clustering leaves a fixed-order contribution that has a logarithm of the ratio of the beam and soft scales. Note that the scale accompanying μ or ν in the fixed-order beam and soft functions is fixed

by power counting. For the p_T veto, the beam ν scale is m_H and the soft ν scale is $p_T^{\rm cut}$, while for the $\mathcal{T}$ veto, the beam scale is $\sqrt{m_H \mathcal{T}^{\rm cut}}$ and the soft scale is $\mathcal{T}^{\rm cut}$. Hence, the clustering effect at $\mathcal{O}(\alpha_s^2)$ in the cross section is

$$\begin{aligned}\Delta\sigma^{(2)}(\mathcal{T}^{\rm cut}) &= \sigma_{\rm LO}\left(\frac{\alpha_s}{\pi}\right)^2 \ln\frac{m_H}{\mathcal{T}^{\rm cut}}\,\frac{1}{2}C_{\mathcal{T}}^{(2)}(\ln R)\,,\\ \Delta\sigma^{(2)}(p_T^{\rm cut}) &= \sigma_{\rm LO}\left(\frac{\alpha_s}{\pi}\right)^2 \ln\frac{m_H}{p_T^{\rm cut}}\,C_{p_T}^{(2)}(\ln R)\,. \end{aligned} \tag{62}$$

This form matches that in Eq. (59). Because these terms are connected with the anomalous dimension, it is possible to resum the logarithm of $\ln(m_H/k^{\rm cut})$ they contain. Since the anomalous dimension contributions from clustering involve only the beam and soft functions, the total evolution factor from RG evolution for each veto observable is

$$\begin{aligned}U_\Delta^{(2)}(\mu_S,\mu_J) &= \exp\left[\int_{\mu_S}^{\mu_J}\frac{{\rm d}\mu}{\mu}\left(\frac{\alpha_s(\mu)}{\pi}\right)^2 C_{\mathcal{T}}^{(2)}(\ln R)\right],\\ U_\Delta^{(2)}(\nu_S,\nu_J) &= \exp\left[\int_{\nu_S}^{\nu_J}\frac{{\rm d}\nu}{\nu}\left(\frac{\alpha_s(\mu)}{\pi}\right)^2 C_{p_T}^{(2)}(\ln R)\right].\end{aligned} \tag{63}$$

Although the logarithms of R are exponentiated here, they are not connected with any scale in the effective theory. Below, we will show that the logarithms of R indeed appear in the exponent, meaning the above procedure is correct. However, the non-cusp anomalous dimension at $\mathcal{O}(\alpha_s^n)$ contains a term $\sim \ln^{n-1}R$, so for small R the perturbative series for the anomalous dimensions contains large unresummed logarithms. Equivalently, the perturbative series in the exponent is not resummed and contains terms $\alpha_s^n \ln^{n-1} R \ln(m_H/k^{\rm cut})$. For $R \sim \lambda$ these are unresummed NLL corrections. Hence, the exponentiation of the $\mathcal{O}(\alpha_s^2)$ clustering logarithms is doing nothing to tame the NLL clustering logarithms at higher orders. The complete resummation of the clustering logarithms would require to calculate the entire tower of coefficients simultaneously.

B. Higher-Order Structure of Clustering Logarithms

The higher-order clustering logarithms are simplest to examine in the soft function, and RG constraints can be used to relate them to the beam functions and determine the overall effect on the cross section. The soft function at $\mathcal{O}(\alpha_s^n)$ can be written as

$$S^{(n)}(p_T^{\rm cut}) = \int {\rm d}\Phi_n\, \mathcal{A}_n(\Phi_n)\, \mathcal{M}_n(\Phi_n, p_T^{\rm cut})\,, \tag{64}$$

where Φ_n and $\mathcal{A}_n$ are the n-particle phase space and matrix element. Non-Abelian exponentiation [43–46] implies that the eikonal matrix elements exponentiate, and can be factorized into c-web contributions, where a c-web is a diagram connecting eikonal lines that cannot be separated into lower order c-webs by cutting each eikonal line once. This means that the amplitude can be written as

$$\mathcal{A}_n(\Phi_n) = \sum_W N_W \left[\prod_{w\in W}\mathcal{A}_w(\Phi_{n_w})\right], \tag{65}$$

where W is a set of c-webs that partition the n-particle state, and w is an individual c-web with n_w particles. N_W is a combinatoric factor from exponentiation of the matrix elements and $\mathcal{A}_w$ is the matrix element for the c-web w.

This factorization is useful because collinear singularities between final-state partons only exist when they are in the same c-web. Therefore if two partons in different c-webs are clustered into the same jet then the rate is suppressed by the area of the jet in rapidity and azimuthal angle, which is $\mathcal{O}(R^2)$. At leading power, this implies that each jet's constituents are in the same c-web, and for a given set of c-webs W the measurement function can be factorized over this set:

$$\mathcal{M}_n(\Phi_n, p_T^{\rm cut}) = \prod_{w\in W}\mathcal{M}_{n_w}(\Phi_{n_w})\,. \tag{66}$$

Note that $\mathcal{M}_{n_w}$ is the same measurement function as $\mathcal{M}_n$ but over Φ_{n_w} instead of Φ_n. Using the exponentiation of matrix elements,

$$S(p_T^{\rm cut}) = \exp\left[\sum_{\text{c-webs } w} S_w(p_T^{\rm cut})\right], \tag{67}$$

where

$$S_w = \int {\rm d}\Phi_{n_w}\, \mathcal{A}(\Phi_{n_w})\, \mathcal{M}_{n_w}(\Phi_{n_w}, p_T^{\rm cut})\,. \tag{68}$$

As argued in Appendix A, RG invariance implies that the clustering logarithms can enter into at most the non-cusp anomalous dimension for the beam and soft functions at each order. For the p_T jet veto, the clustering logarithms will be a part of the ν non-cusp anomalous dimension.

Just as they did at $\mathcal{O}(\alpha_s^2)$, collinear divergences between particles in the c-web will produce logarithms of R as a finite artifact of the divergence. Since there are at most $n_w - 1$ collinear divergences in a c-web of n_w particles, this implies there are at most $n_w - 1$ logarithms of R. Therefore, the general form of the clustering logarithms in the soft function S_w for the c-web is, in the $\mathcal{T}$ veto case,

$$\Delta S_w(\mathcal{T}^{\rm cut}) = \left(\frac{\alpha_s}{\pi}\right)^{n_w} C_{\mathcal{T}}^{(n_w)}(\ln R)\, \ln\frac{\mu}{\mathcal{T}^{\rm cut}}\,, \tag{69}$$

and in the p_T veto case,

$$\Delta S_w(p_T^{\rm cut}) = \left(\frac{\alpha_s}{\pi}\right)^{n_w} C_{p_T}^{(n_w)}(\ln R)\, \ln\frac{\nu}{p_T^{\rm cut}}\,. \tag{70}$$

FIG. 3: Numerical size of various $\mathcal{O}(\alpha_s^2)$ terms for a cut on p_{Tj} (left panel) and on $\mathcal{T}_j$ (right panel) for $m_H = 125\,\text{GeV}$ at a 8 TeV LHC. The terms are scaled by the leading order cross section, σ_{LO}. We show two values of the jet radius, $R = 0.4$ (solid curves) and $R = 1.0$ (dashed curves) for the k_{T} class of jet algorithms. Shown are the $\mathcal{O}(\alpha_s^2)$ contribution to the total inclusive cross section with no veto, $\sigma^{(2)}(\infty)$, the full $\mathcal{O}(\alpha_s^2)$ contribution with a jet veto, $\sigma^{(2)}(k^{\text{cut}})$ with $k^{\text{cut}} = p_T^{\text{cut}}$ or $\mathcal{T}^{\text{cut}}$, the $\mathcal{O}(\alpha_s^2)$ clustering logarithm corrections, $\Delta\sigma^{(2)}(\ln R)$, and the soft-collinear mixing contribution, $\delta\sigma_{SC}^{(2)}(R^2)$. At small R, the clustering logarithms are sizable while the soft-collinear mixing terms are small. At large R the situation is reversed.

Both $C_{\mathcal{T}}^{(n_w)}$ and $C_{p_T}^{(n_w)}$ contain at most $n_w - 1$ powers of $\ln R$. These give rise to non-cusp anomalous dimension contributions. RG evolution of the soft function and beam function clustering logarithms from the c-web w gives rise to an all-orders contribution to σ of the form

$$U_\Delta^{(n_w)}(\mu_S, \mu_J) = \exp\left[\int_{\mu_S}^{\mu_J} \frac{\mathrm{d}\mu}{\mu} \left(\frac{\alpha_s(\mu)}{\pi}\right)^{n_w} C_{\mathcal{T}}^{(n_w)}(\ln R)\right],$$
$$U_\Delta^{(n_w)}(\nu_S, \nu_J) = \exp\left[\int_{\nu_S}^{\nu_J} \frac{\mathrm{d}\nu}{\nu} \left(\frac{\alpha_s(\mu)}{\pi}\right)^{n_w} C_{p_T}^{(n_w)}(\ln R)\right]. \tag{71}$$

Since $C^{(n_w)}$ contains a $\ln^{n_w-1} R$ term, when $R \sim \lambda$ this series is NLL. These terms are not directly obtained by resummation of lower order terms, underscoring the breakdown of resummation unless the coefficients of the clustering logarithms can all be calculated at once.

V. CONCLUSIONS

Jet vetoes are an important part of many experimental analyses. To obtain precise theoretical predictions, the resummation of large logarithms of the jet-veto scale is important, but they must account for effects of the jet algorithm. The algorithm generates two effects that can inhibit resummation, namely soft-collinear mixing and clustering logarithms. From a formal perspective, there are two scaling limits to consider, $R \sim \lambda$ and $R \gg \lambda$. Soft-collinear mixing arises at large values of the jet radius $R \gg \lambda$ from correlations between soft and collinear radiation into the same jet. This mixing affects the resummation of logarithms of the veto scale at NNLL and beyond and must be accounted for in order to prove an all-orders factorization theorem. When $R \sim \lambda$, soft-collinear mixing effects are power suppressed, but clustering logarithms are important, which are a remnant of collinear singularities between particles in the jet that cancel in the total cross section. In the small R limit, new clustering logarithms at NLL arise at each order and cannot currently be resummed. Therefore, there is a tradeoff between these effects depending on the value of R chosen, and it is important to understand the relative sizes of these effects and incorporate them into estimates of theory uncertainties.

In Sec. II, we contrasted the factorization picture for inclusive observables $\mathcal{T}_B$ and E_T with the exclusive jet-based observables $\mathcal{T}_j$ and p_{Tj}. The factorization and resummation for the inclusive observables are well understood. In this paper we have explored the effects of soft-collinear mixing and clustering on the factorization and resummation for the exclusive observables. In Secs. III and IV, we calculated the soft-collinear mixing and clustering logarithms at $\mathcal{O}(\alpha_s^2)$.

The resummation for a p_{Tj} veto was recently considered in Refs. [14, 15]. In Ref. [14], the resummation was performed to NLL, but the clustering logarithms were only included at fixed $\mathcal{O}(\alpha_s^2)$. Our calculation of the leading clustering logarithms agree with the analytic results given in Ref. [14]. Reference [15] works in the limit $\lambda \ll R \sim 1$. They argue that in this limit the factorization formula in Eq. (32) holds up to $\mathcal{O}(\lambda)$ power corrections, and based on this extend the resummation to NNLL. While clustering logarithms do not pose a problem for $R \sim 1$, we disagree that the factorization formula in Eq. (32) holds in this limit. Since soft-collinear mixing contributions are in fact not power suppressed for $R \sim 1$, the factorization formula does not reproduce the full NNLL structure of the cross section at leading power.

Therefore, the impact of these terms on the perturbative uncertainties found in Ref. [15] should be examined.

We can investigate the numerical size of the jet-algorithm effects using Higgs production through gluon fusion as an example. The cross section for a veto $k^{\rm cut}$ on $\mathcal{T}_j$ or p_{Tj} can be expanded in fixed order as

$$\sigma(k^{\rm cut}) = \sigma_{\rm LO} + \sigma^{(1)}(k^{\rm cut}) + \sigma^{(2)}(k^{\rm cut}) + \mathcal{O}(\alpha_s^3)\,, \quad (72)$$

where $\sigma^{(n)}$ is $\mathcal{O}(\alpha_s^n)$ relative to $\sigma_{\rm LO}$. The $\mathcal{O}(\alpha_s)$ terms, $\sigma^{(1)}$, do not depend on the jet algorithm, and hence are the same for $\mathcal{T}_j$, p_{Tj} and $\mathcal{T}_B$, E_T, respectively. The $\mathcal{O}(\alpha_s^2)$ terms, $\sigma^{(2)}$, include the effect of clustering logarithms, $\Delta\sigma^{(2)}(\ln R)$, and soft-collinear mixing, $\delta\sigma^{(2)}(R^2)$,

$$\sigma^{(2)}(k^{\rm cut}) \supset \Delta\sigma^{(2)}(\ln R)\,,\, \delta\sigma^{(2)}(R^2)\,. \quad (73)$$

In Fig. 3, we plot these terms in units of $\sigma_{\rm LO}$ as a function of $p_T^{\rm cut}$ (left panel) and $\mathcal{T}^{\rm cut}$ (right panel), for $R = 0.4$ (solid curves) and $R = 1.0$ (dashed curves). For comparison, we also show the full $\mathcal{O}(\alpha_s^2)$ correction to the vetoed cross section, $\sigma^{(2)}(k^{\rm cut})$, as well as the $\mathcal{O}(\alpha_s^2)$ correction to the total cross section without a veto, $\sigma^{(2)}(k^{\rm cut} = \infty)$ [47–52].

The tradeoff between soft-collinear mixing and clustering logarithms is clear in Fig. 3. For $R = 0.4$, the clustering logarithms dominate over the mixing terms. For $R = 1.0$, the mixing terms dominate over the clustering logarithms. In each case, the size of the numerically more important term is an appreciable fraction of the $\mathcal{O}(\alpha_s^2)$ correction. The difference between $\sigma^{(2)}(k^{\rm cut})$ and $\sigma^{(2)}(\infty)$ primarily comes from the logarithmic terms, while the absolute size of $\sigma^{(2)}(k^{\rm cut})$ relative to $\sigma^{(2)}(\infty)$ is indicative of the substantial cancellation between the jet-veto logarithms and the large NNLO K-factor for the total cross section [53]. Given that depending on R the clustering or soft-collinear mixing terms are a sizable fraction of the logarithms $\ln(m_H/k^{\rm cut})$ that one is trying to resum, it is crucial to understand their size at higher orders and correctly fold them into resummed uncertainty estimates. This has not been done in previous studies of the Higgs cross section where resummation for the p_T veto is performed.

The standard jet radii in Higgs analyses are $R = 0.4$ for ATLAS and $R = 0.5$ for CMS. As seen in Fig. 3, with these values, at $\mathcal{O}(\alpha_s^2)$ the clustering logarithms are numerically important and the mixing terms can be regarded as power corrections. This suggests that the phenomenologically relevant limit is $R \sim \lambda$, where soft-collinear factorization can be applied but clustering logarithms should be regarded as NLL. To proceed one should study the impact of the clustering logarithms on the resummed perturbative series in order to properly take them into account in a resummed uncertainty estimate. For this purpose, a calculation of the $\mathcal{O}(\alpha_s^3)$ clustering logarithms would provide very useful information.

Note added in proofs: After this work was completed, Ref. [54] appeared, where a NNLL resummation formula for the cross section with a p_T jet veto is presented for $R \sim 1$ that accounts for the contributions due to soft-collinear mixing. This formula is equivalent to modifying Eq. (32) to include these terms by hand, and is consistent with our analysis that Eq. (32) by itself is insufficient to perform the resummation beyond NLL when $R \sim 1$. (Ref. [54] does not discuss the case $R \sim \lambda$.) The structure of the $\mathcal{O}(\alpha_s^2)$ soft-collinear mixing terms suggests that for the p_T case one may be able to absorb them into the beam and soft functions in Eq. (32) by performing additional collinear zero-bin subtractions in the soft sector. It is an open question and beyond the scope of this work if such an approach can be extended to all orders in α_s such that Eq. (32) becomes a valid all-order factorization theorem.

Acknowledgments

We thank Andrea Banfi, Thomas Becher, and Matthias Neubert for discussions. We thank Iain Stewart for comments on the manuscript. J.W. and S.Z. thank the DESY Theory Group for hospitality while portions of this work were completed. This work was supported in part by the DFG Emmy-Noether Grant No. TA 867/1-1, by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under the Contract No. DE-AC02-05CH11231. J.W. was supported by the LHC Theory Initiative, under the National Science Foundation Grant No. PHY-0705682.

Appendix A: Renormalization Group Constraints

The anomalous dimensions of the beam and soft functions for the four observables that we study are constrained by RG invariance. We give a brief summary of the properties of the anomalous dimensions and the general form for each observable. For the exclusive observables, we work in terms of the veto variable $\mathcal{T}^{\rm cut}$ or $p_T^{\rm cut}$, and these forms are useful when discussing clustering logarithms in Sec. IV.

The hard function is universal for all the observables used, and the hard anomalous dimension is

$$\gamma_H^{gg}(m_H,\mu) = 2\Gamma^g_{\rm cusp}[\alpha_s(\mu)]\ln\frac{m_H^2}{\mu^2} + 2\gamma_H^g[\alpha_s(\mu)]\,. \quad (A1)$$

Consistency of the factorization theorem requires that the hard, beam, and soft anomalous dimensions cancel. For the four observables, this consistency is slightly different.

For beam thrust, the consistency relation is

$$\gamma_H^{gg}(m_H,\mu)\delta(\mathcal{T}_B) + 2m_H\gamma_B^g(m_H\mathcal{T}_B,\mu) + \gamma_S^{gg}(\mathcal{T}_B,\mu) = 0\,. \quad (A2)$$

Analogously, for the $\mathcal{T}_j < \mathcal{T}^{\rm cut}$ veto on jets, the consistency relation is

$$\gamma_H^{gg}(m_H,\mu) + 2\gamma_B^g(m_H\mathcal{T}^{\rm cut},\mu) + \gamma_S^{gg}(\mathcal{T}^{\rm cut},\mu) = 0\,. \quad (A3)$$

Both of these factorization theorems are described by $\mathrm{SCET_I}$. The scale dependence in the beam and soft functions is fixed by power counting; the beam function scale is $\mu_B^2 = m_H\mathcal{T}$ and the soft function scale is $\mu_S = \mathcal{T}$, where $\mathcal{T} = \mathcal{T}_B$ or $\mathcal{T}^{\mathrm{cut}}$. This constrains the beam and soft anomalous dimensions, so that the coefficients of the cusp anomalous dimension are fixed relative to that of the hard function. For beam thrust, these anomalous dimensions are

$$\begin{aligned}\gamma_B^g(t,\mu) &= -2\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\,\frac{1}{\mu^2}\mathcal{L}_0\Big(\frac{t}{\mu^2}\Big) + \gamma_B^g[\alpha_s(\mu)]\,\delta(t)\,,\\ \gamma_S^{gg}(k,\mu) &= 4\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\,\frac{1}{\mu}\mathcal{L}_0\Big(\frac{k}{\mu}\Big) + \gamma_S^{gg}[\alpha_s(\mu)]\,\delta(k)\,,\end{aligned} \tag{A4}$$

where $\mathcal{L}_0(x) = [\theta(x)/x]_+$ denotes the usual plus distribution. For a $\mathcal{T}_j$ veto on jets, the anomalous dimensions are

$$\begin{aligned}\gamma_B^g(m_H\mathcal{T}^{\mathrm{cut}},\mu) &= -2\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\ln\frac{m_H\mathcal{T}^{\mathrm{cut}}}{\mu^2} \\ &\quad + \gamma^g_{B\,\mathrm{jet}}[\alpha_s(\mu)]\,,\\ \gamma_S^{gg}(\mathcal{T}^{\mathrm{cut}},\mu) &= 4\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\ln\frac{\mathcal{T}^{\mathrm{cut}}}{\mu} \\ &\quad + \gamma^{gg}_{S\,\mathrm{jet}}[\alpha_s(\mu)]\,.\end{aligned} \tag{A5}$$

The non-cusp anomalous dimensions for these two observables agree at $\mathcal{O}(\alpha_s)$, but due to the different structure of their factorization theorems (beam thrust contains a convolution while the $\mathcal{T}_j$ veto does not), the resummed distributions only coincide at leading logarithmic order.

For the p_T-based observables, which are described by $\mathrm{SCET_{II}}$, rapidity divergences exist and must be separately regulated. The rapidity renormalization group formalizes the regularization of rapidity divergences by introducing separate anomalous dimensions for the beam (or jet) and soft functions [30, 31]. This rapidity renormalization group works much like the traditional RGE, with a scale ν that is the analog to the usual μ scale.

Power counting and renormalization group invariance provide strong constraints on the anomalous dimensions for p_T-based observables. Rapidity divergences are regulated by factors inserted into soft and collinear Wilson lines; in each sector the regulator and its scaling are

$$\begin{aligned}\text{beam:}\quad & \nu^\eta|\bar n\cdot P_g|^{-\eta} \sim \nu^\eta\big(\lambda^0\big)^{-\eta}\,,\\ \text{soft:}\quad & \nu^\eta|P_{3g}|^{-\eta} \sim \nu^\eta\big(\lambda\big)^{-\eta}\,,\end{aligned} \tag{A6}$$

where P^μ is the momentum operator, and the group momentum (P_g) of connected webs of gluons is regulated. In each beam function the large component of momentum is regulated. This power counting implies that in the rapidity RG the beam function scale is $\nu_B = m_H$ and the soft function scale $\nu_S = p_T$, where $p_T = E_T$ or p_T^{cut}. In addition, in the standard RG in μ, the beam and soft functions live at the low scale, $\mu_B = \mu_S = p_T$.

For the p_T-based observables, there are two consistency relations: one in μ-space and one in ν-space. For the inclusive E_T observable, the consistency relations are

$$\begin{aligned}0 &= \gamma_H^{gg}(m_H,\mu)\delta(E_T) + 2\gamma_B^{\mu,g}(m_H,E_T,\mu,\nu) \\ &\quad + \gamma_S^{\mu,gg}(E_T,\mu,\nu)\,,\\ 0 &= 2\gamma_B^{\nu,g}(E_T,\mu,\nu) + \gamma_S^{\nu,gg}(E_T,\mu,\nu)\,.\end{aligned} \tag{A7}$$

And for a p_{Tj} veto on jets, the consistency relations are

$$\begin{aligned}0 &= \gamma_H^{gg}(m_H,\mu) + 2\gamma_B^{\mu,g}(m_H,p_T^{\mathrm{cut}},\mu,\nu) \\ &\quad + \gamma_S^{\mu,gg}(p_T^{\mathrm{cut}},\mu,\nu)\,,\\ 0 &= 2\gamma_B^{\nu,g}(p_T^{\mathrm{cut}},\mu,\nu) + \gamma_S^{\nu,gg}(p_T^{\mathrm{cut}},\mu,\nu)\,.\end{aligned} \tag{A8}$$

The general forms of the beam and soft anomalous dimensions can be constrained through these relations as well as the exactness condition,

$$\frac{\mathrm{d}}{\mathrm{d}\ln\mu}\gamma_F^\nu(\mu,\nu) = \frac{\mathrm{d}}{\mathrm{d}\ln\nu}\gamma_F^\mu(\mu,\nu)\,, \tag{A9}$$

for $F = B, S$. We ignore running coupling effects in this exactness relation that will contribute to the ν anomalous dimensions starting at $\mathcal{O}(\alpha_s^2)$. For E_T, the μ anomalous dimensions have the form

$$\begin{aligned}&\gamma_B^{\mu,g}(m_H,E_T,\mu,\nu) \\ &\quad= \bigg\{2\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\ln\frac{\nu}{m_H} + \gamma_B^{\mu,g}[\alpha_s(\mu)]\bigg\}\delta(E_T) \\ &\qquad + \gamma^\mu[\alpha_s(\mu)]\,\frac{1}{\mu}\mathcal{L}_0\Big(\frac{E_T}{\mu}\Big)\,,\\ &\gamma_S^{\mu,g}(E_T,\mu,\nu) \\ &\quad= \bigg\{4\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\ln\frac{\mu}{\nu} + \gamma_S^{\mu,g}[\alpha_s(\mu)]\bigg\}\delta(E_T) \\ &\qquad - 2\gamma^\mu[\alpha_s(\mu)]\,\frac{1}{\mu}\mathcal{L}_0\Big(\frac{E_T}{\mu}\Big)\,,\end{aligned} \tag{A10}$$

and the ν anomalous dimensions have the form

$$\begin{aligned}\gamma_B^{\nu,g}(E_T,\mu,\nu) &= -2\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\,\frac{1}{\mu}\mathcal{L}_0\Big(\frac{E_T}{\mu}\Big) \\ &\quad + \gamma^{\nu,g}[\alpha_s(\mu)]\,\delta(E_T)\,,\\ \gamma_S^{\nu,g}(E_T,\mu,\nu) &= 4\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\,\frac{1}{\mu}\mathcal{L}_0\Big(\frac{E_T}{\mu}\Big) \\ &\quad - 2\gamma^{\nu,g}[\alpha_s(\mu)]\,\delta(E_T)\,.\end{aligned} \tag{A11}$$

The cusp anomalous dimension dependence in the hard function fixes every part of the beam and soft anomalous dimensions except for the non-cusp terms and a cusp term $\gamma^\mu[\alpha_s(\mu)]$ in the beam and soft μ anomalous dimensions that cancels between them. Because the beam and soft functions have the same μ scale, as long as they are RG evolved in μ in the same way this γ^μ term will cancel. At fixed-order, this cancellation is guaranteed and has no effect on the fixed-order logarithms. It is absent in the anomalous dimensions at $\mathcal{O}(\alpha_s)$, and may be

absent at all orders. Additionally, there is a constraint on the non-cusp anomalous dimensions from Eq. (A7), $2\gamma_H^g[\alpha_s] + 2\gamma_B^{\mu,g}[\alpha_s] + \gamma_S^{\mu,gg}[\alpha_s] = 0$. If the beam and soft functions are RG evolved in μ in the same way, then it is irrelevant how the μ non-cusp anomalous dimensions are divided between them, since only the sum matters.

For the p_{Tj} veto the μ anomalous dimensions have the form

$$\begin{aligned}\gamma_B^{\mu,g}(m_H, p_T^{\text{cut}}, \mu, \nu) &= 2\Gamma^g_{\text{cusp}}[\alpha_s(\mu)] \ln\frac{\nu}{m_H} \\ &\quad + \gamma^{\mu,g}_{B\,\text{jet}}[\alpha_s(\mu)] + \gamma^{\mu}[\alpha_s(\mu)] \ln\frac{p_T^{\text{cut}}}{\mu}\,, \\ \gamma_S^{\mu,g}(p_T^{\text{cut}}, \mu, \nu) &= 4\Gamma^g_{\text{cusp}}[\alpha_s(\mu)] \ln\frac{\mu}{\nu} + \gamma^{\mu,gg}_{S\,\text{jet}}[\alpha_s(\mu)] \\ &\quad - 2\gamma^{\mu}[\alpha_s(\mu)] \ln\frac{p_T^{\text{cut}}}{\mu}\,, \end{aligned} \tag{A12}$$

and the ν anomalous dimensions are

$$\begin{aligned}\gamma_B^{\nu,g}(p_T^{\text{cut}}, \mu, \nu) &= -2\Gamma^g_{\text{cusp}}[\alpha_s(\mu)] \ln\frac{p_T^{\text{cut}}}{\mu} + \gamma^{\nu,g}_{\text{jet}}[\alpha_s(\mu)]\,, \\ \gamma_S^{\nu,g}(p_T^{\text{cut}}, \mu, \nu) &= 4\Gamma^g_{\text{cusp}}[\alpha_s(\mu)] \ln\frac{p_T^{\text{cut}}}{\mu} - 2\gamma^{\nu,g}_{\text{jet}}[\alpha_s(\mu)]\,. \end{aligned} \tag{A13}$$

The same considerations as for the E_T observable apply here. All the cusp parts are fixed by RG invariance with the hard function, and the only nontrivial unconstrained terms are the non-cusp ν anomalous dimensions of the beam and soft functions.

If clustering effects are associated with divergences in the beam and soft functions, then they can only impact the non-cusp beam and soft anomalous dimensions. This implies that clustering corrections have at most a single logarithm of the veto variable at each order. However, as we saw in Sec. IV, they can produce additional logarithms of the jet radius R, because these are not associated with any scale and are therefore not constrained by the RGE.

Appendix B: Soft-Collinear Mixing at $\mathcal{O}(\alpha_s^2)$

In this appendix we present the $\mathcal{O}(\alpha_s^2)$ calculations of the soft-collinear mixing contribution to the cross section, denoted as $\delta\sigma_{SC}^{\text{jet}\,(2)} = \delta\sigma_{SC_a}^{\text{jet}\,(2)} + \delta\sigma_{SC_b}^{\text{jet}\,(2)}$. As explained in Sec. III, this contribution arises from a correlation in the measurement between single independent emissions from the soft and collinear sectors, and formally breaks factorization if it is not power suppressed. These mixing terms have the same form as the independent emission clustering terms at $\mathcal{O}(\alpha_s^2)$ in the soft and collinear sectors, and the total $\mathcal{O}(\alpha_s^2)$ clustering correction to the independent emission contribution to the cross section is given by [see Eq. (47)]

$$\Delta\sigma^{\text{indep}\,(2)} = \Delta\sigma_{SS}^{\text{indep}\,(2)} + \Delta\sigma_{CC}^{\text{indep}\,(2)} + \delta\sigma_{SC}^{\text{jet}\,(2)}\,. \tag{B1}$$

The crucial result is that the soft-collinear mixing terms are NNLL terms that scale as R^2, requiring R to scale as λ for them to be formally power suppressed. We also calculate the $\mathcal{O}(\alpha_s^2)$ clustering contribution from independent emissions in the soft function, $\Delta\sigma_{SS}^{\text{indep}\,(2)}$.

The bare soft functions at $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ for independent emission in $\overline{\text{MS}}$ (without a measurement function) are

$$\begin{aligned} S^{\text{b}\,(1)} &= \frac{\alpha_s C_A}{\pi} \frac{e^{\gamma_E\epsilon}\mu^{2\epsilon}}{\Gamma(1-\epsilon)} \int_0^\infty \mathrm{d}k^+\mathrm{d}k^-\,(k^+k^-)^{-1-\epsilon}\,, \\ S^{\text{b}\,(2)} &= \frac{1}{2}\big[S^{\text{b}\,(1)}\big]^2\,, \end{aligned} \tag{B2}$$

where the label b stands for bare. We have implemented the on-shell conditions and left only the $+,-$ light-cone components unintegrated.

The gluon beam function B_g can be written as the convolution between a perturbative function $\mathcal{I}(t,z,\mu)$ and the parton distribution function $f_g(x,\mu)$ [5, 28],

$$B_g(t,x,\mu) = \sum_{j=\{g,q,\bar{q}\}} \int_x^1 \frac{\mathrm{d}z}{z}\,\mathcal{I}_{gj}(t,z,\mu)\,f_j\Big(\frac{x}{z},\mu\Big)\,, \tag{B3}$$

where $t = m_H k^+$ is the spacelike virtuality of the gluon entering in the hard interaction. The tree level, $B_g^{(0)}$, bare $\mathcal{O}(\alpha_s)$ naive, $\widetilde{B}_g^{\text{b}\,(1)}$, and zero-bin, $B_{g,0}^{\text{b}\,(1)}$, beam functions (without a measurement function) are [5]

$$\begin{aligned} B_g^{(0)} &= \delta_{gj}\frac{1}{m_H}\delta(k^+)\delta(1-z)\,, \\ \widetilde{B}_g^{\text{b}\,(1)} &= \frac{\alpha_s C_A}{2\pi}\theta(z)\hat{P}_{gg}(z)\frac{1}{m_H}\frac{e^{\gamma_E\epsilon}\mu^{2\epsilon}}{\Gamma(1-\epsilon)}\int_0^\infty \mathrm{d}q^+\mathrm{d}q^- \\ &\quad \times (q^+)^{-1-\epsilon}(q^-)^{-\epsilon}\,\delta\Big(q^- - m_H\frac{1-z}{z}\Big)\,, \\ B_{g,0}^{\text{b}\,(1)} &= \frac{\alpha_s C_A}{\pi}\frac{1}{m_H}\delta(1-z)\frac{e^{\gamma_E\epsilon}\mu^{2\epsilon}}{\Gamma(1-\epsilon)} \\ &\quad \times \int_0^\infty \mathrm{d}q^+\mathrm{d}q^-\,(q^+)^{-1-\epsilon}(q^-)^{-1-\epsilon}\,, \\ B_g^{\text{b}\,(1)} &= \widetilde{B}_g^{\text{b}\,(1)} - B_{g,0}^{\text{b}\,(1)}\,, \end{aligned} \tag{B4}$$

where

$$\hat{P}_{gg}(z) = 2\Big[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\Big]\,. \tag{B5}$$

We have dropped the quark contributions to the beam function as they will not produce logarithms in the mixing terms. In $\widetilde{B}_g^{\text{b}\,(1)}$, there is a singularity as $z\to 1$ in the unregularized splitting function $\hat{P}_{gg}$ that is regulated by q^-. When considering the full measurement (both the factorized and soft-collinear mixing terms), there is an IR divergence in $\mathcal{I}$ that is canceled by the parton distribution functions; however, this cancellation occurs in the factorized measurement function and is not present in the mixing term. The soft-collinear mixing correction has a single divergence and $B_g^{\text{b}\,(1)}$ is proportional to $\delta(1-z)$, meaning there is no nontrivial convolution with

the parton distribution functions. Therefore we will just compute the coefficient of the leading order cross section, $\sigma_{\rm LO}$, which is given by

$$\sigma_{\rm LO} = \sigma_0\, H_{gg}^{(0)}(m_H,\mu) \int {\rm d}Y\, f_g(x_a,\mu)\, f_g(x_b,\mu)\,, \quad \text{(B6)}$$

with $x_{a,b} = (m_H/E_{\rm cm})e^{\pm Y}$.

1. Clustering Effects for Independent Emissions in the Soft Function

We perform the soft function calculations first, as the techniques and results will be used later in the soft-collinear mixing terms. The independent emission terms are of the form

$$\Delta\sigma_{SS}^{\rm indep\,(2)}(k^{\rm cut}) = S^{\rm b\,(2)} \int_0^\pi \frac{{\rm d}\Delta\phi}{\pi}\, \Delta\mathcal{M}_s^{\rm jet\,(2)}\,. \quad \text{(B7)}$$

The integral over $\Delta\phi$ does not need to be regulated since there is no collinear divergence between the final-state particles. The measurement function for a $\mathcal{T}_j$ veto is

$$\Delta\mathcal{M}_s^{\rm jet\,(2)}(\mathcal{T}^{\rm cut}) = 2\theta(\Delta R_{12} < R)\Big[\theta(k_1^+ + k_2^+ < \mathcal{T}^{\rm cut}) - \theta(k_1^+ < \mathcal{T}^{\rm cut})\theta(k_2^+ < \mathcal{T}^{\rm cut})\Big]\,. \quad \text{(B8)}$$

We have multiplied by 2 to account for the case where the two gluons are both in the other hemisphere (where $\mathcal{T}$ is equal to the minus component of momenta), and dropped the $\mathcal{O}(R^4)$ region of phase space where the two gluons are in different hemispheres but still cluster. For a p_{Tj} veto, the measurement function is

$$\Delta\mathcal{M}_s^{\rm jet\,(2)}(p_T^{\rm cut}) = 2\theta(\Delta R_{12} < R)\Big[\theta(p_{T1} + p_{T2} < p_T^{\rm cut}) - \theta(p_{T1} < p_T^{\rm cut})\theta(p_{T2} < p_T^{\rm cut})\Big]\,. \quad \text{(B9)}$$

In this case we drop the $\mathcal{O}(R^4)$ correction from using the vector sum over transverse momenta in the combined constraint.

To evaluate the matrix element, we use the variables

$$\begin{aligned} \Delta y &= \frac{1}{2}\ln\frac{k_1^- k_2^+}{k_1^+ k_2^-}\,, & y_t &= \frac{1}{4}\ln\frac{k_1^- k_2^-}{k_1^+ k_2^+}\,, \\ \mathcal{T}^{\rm cut}\,:\,\mathcal{T}_1 &= k_1^+\,, & \mathcal{T}_2 &= k_2^+\,, \\ p_T^{\rm cut}\,:\,p_{T1} &= \sqrt{k_1^+ k_1^-}\,, & p_{T2} &= \sqrt{k_2^+ k_2^-}\,. \end{aligned} \quad \text{(B10)}$$

For the $\mathcal{T}_j$ veto, the soft-function contribution is

$$\begin{aligned} &\Delta\sigma_{SS}^{\rm indep\,(2)}(\mathcal{T}^{\rm cut}) \quad \text{(B11)} \\ &= \frac{\sigma_{\rm LO}}{2}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{(e^{\gamma_E}\mu^2)^{2\epsilon}}{\Gamma(1-\epsilon)^2}\, 8 \int_0^\infty {\rm d}y_t\, e^{-4\epsilon y_t} \int_{-\infty}^\infty {\rm d}\Delta y \\ &\quad\times \int_0^\pi \frac{{\rm d}\Delta\phi}{\pi}\,\theta(\Delta R_{12} < R) \int_0^\infty {\rm d}\mathcal{T}_1\, {\rm d}\mathcal{T}_2\, (\mathcal{T}_1\,\mathcal{T}_2)^{-1-2\epsilon} \\ &\quad\times \Big[\theta(\mathcal{T}_1 + \mathcal{T}_2 < \mathcal{T}^{\rm cut}) - \theta(\mathcal{T}_1 < \mathcal{T}^{\rm cut})\theta(\mathcal{T}_2 < \mathcal{T}^{\rm cut})\Big]\,. \end{aligned}$$

Only the integral over y_t produces a divergence, and the integrals are straightforward to evaluate. The divergent term is

$$\Delta\sigma_{SS}^{\rm indep\,(2)}(\mathcal{T}^{\rm cut}) = -\sigma_{\rm LO}\frac{1}{\epsilon}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\Big(\frac{\mu}{\mathcal{T}^{\rm cut}}\Big)^{4\epsilon}\frac{\pi^2}{12}R^2\,. \quad \text{(B12)}$$

For p_{Tj}, we need to regulate the rapidity divergences, for which we use the analytic regulator [42]. In this case, the regulator factor is

$$\nu^{2\alpha}(k_1^+ k_2^+)^{-\alpha} = \nu^{2\alpha}(p_{T1}p_{T2})^{-\alpha} e^{4\alpha y_t}\,. \quad \text{(B13)}$$

The amplitude and measurement function are independent of y_t. Hence, the soft function is proportional to

$$\begin{aligned} \int_{-\infty}^\infty {\rm d}y_t\, e^{4\alpha y_t} &= \int_{-\infty}^x {\rm d}y_t\, e^{4\alpha y_t} + \int_x^\infty {\rm d}y_t\, e^{4\alpha y_t} \\ &= \frac{1}{4\alpha}\big[e^{4\alpha x} - e^{4\alpha x}\big] = 0\,. \end{aligned} \quad \text{(B14)}$$

We made the reason why this integral is zero explicit by breaking the range of integration into two parts: $(-\infty, x)$ and (x,∞). In the lower range $\alpha > 0$ regulates the integral and in the upper range $\alpha < 0$ regulates the integral, and the two terms cancel. (This is precisely equivalent to the case of a scaleless integral in pure dimensional regularization with α here playing the role of ϵ.) Thus,

$$\Delta\sigma_{SS}^{\rm indep\,(2)}(p_T^{\rm cut}) = 0\,. \quad \text{(B15)}$$

2. Soft-Collinear Mixing Terms

The mixing term between soft and n_a-collinear emissions at $\mathcal{O}(\alpha_s^2)$ is given by

$$(SC_a) = B_g^{\rm b\,(1)} S^{\rm b\,(1)} \int_0^\pi \frac{{\rm d}\Delta\phi}{\pi}\, \delta\mathcal{M}_{as}^{\rm jet\,(2)}\,, \quad \text{(B16)}$$

and similarly for (SC_b). We have included the integral over the relative azimuthal angle $\Delta\phi$ between the soft and collinear particles. Since we are concerned with the leading divergences, we do not need to regulate ϕ. The $\mathcal{O}(\alpha_s^2)$ measurement corrections for $\mathcal{T}_j$ and p_{Tj}, $\delta\mathcal{M}^{\rm jet\,(2)}(\mathcal{T}^{\rm cut})$ and $\delta\mathcal{M}^{\rm jet\,(2)}(p_T^{\rm cut})$, are

$$\begin{aligned} \delta\mathcal{M}^{\rm jet\,(2)}(\mathcal{T}^{\rm cut}) &= \theta(\Delta R_{sc} < R)\Big[\theta(\mathcal{T}_c + \mathcal{T}_s < \mathcal{T}^{\rm cut}) - \theta(\mathcal{T}_c < \mathcal{T}^{\rm cut})\,\theta(\mathcal{T}_s < \mathcal{T}^{\rm cut})\Big]\,, \\ \delta\mathcal{M}^{\rm jet\,(2)}(p_T^{\rm cut}) &= \theta(\Delta R_{sc} < R)\Big[\theta(p_{Tc} + p_{Ts} < p_T^{\rm cut}) - \theta(p_{Tc} < p_T^{\rm cut})\,\theta(p_{Ts} < p_T^{\rm cut})\Big]\,. \end{aligned} \quad \text{(B17)}$$

For the $\mathcal{T}_j$ veto, the two soft-collinear mixing terms, (SC_a) and (SC_b), are equal. They are made up of the naive and zero-bin beam function contributions, so that

$$(SC) = (S\widetilde{C}) - (SC^{(0)})\,. \tag{B18}$$

The zero-bin of the beam function in each collinear sector is equal to the soft function contribution, meaning

$$(SC^{(0)}) = 2(SS)\,. \tag{B19}$$

Changing variables from k^- to Δy,

$$k^- = k^+ e^{2(y_c - \Delta y)}\,, \quad e^{-2y_c} = \frac{q^+ z}{m_H(1-z)}\,. \tag{B20}$$

the total correction from the naive beam function term in the mixing is

$$\begin{aligned}
&(S\widetilde{C})(\mathcal{T}^{\rm cut}) \\
&\quad = 2\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{(e^{\gamma_E}\mu^2)^{2\epsilon}}{\Gamma(1-\epsilon)^2}\,\theta(z)\hat{P}_{gg}(z)\Big(\frac{z}{1-z}\Big)^{2\epsilon}\frac{1}{m_H^{1+2\epsilon}} \\
&\qquad \times \int_{-\infty}^{\infty}\!\mathrm{d}\Delta y\, e^{2\epsilon\Delta y}\int_0^{\pi}\frac{\mathrm{d}\Delta\phi}{\pi}\theta(\Delta R < R) \\
&\qquad \times \int_0^{\infty}\!\mathrm{d}q^+\mathrm{d}k^+\,(q^+k^+)^{-1-2\epsilon} \\
&\qquad \times \Big[\theta(k^+ + q^+ < \mathcal{T}^{\rm cut}) - \theta(q^+ < \mathcal{T}^{\rm cut})\theta(k^+ < \mathcal{T}^{\rm cut})\Big].
\end{aligned} \tag{B21}$$

The remaining integrals are finite, so we can set $\epsilon = 0$ in the integrand after pulling out an overall scale dependence of $(\mathcal{T}^{\rm cut})^{-2\epsilon}$. The integrals over Δy and $\Delta\phi$ give

$$\int_{-\infty}^{\infty}\!\mathrm{d}\Delta y\int_0^{\pi}\frac{\mathrm{d}\Delta\phi}{\pi}\,\theta(\Delta R < R) = \frac{1}{2}R^2\,, \tag{B22}$$

and the integrals over q^+ and k^+ give

$$\begin{aligned}
\int_0^{\infty}\!\mathrm{d}q^+\mathrm{d}k^+\,\frac{1}{q^+k^+}\Big[&\theta(k^+ + q^+ < \mathcal{T}^{\rm cut}) \\
&- \theta(q^+ < \mathcal{T}^{\rm cut})\,\theta(k^+ < \mathcal{T}^{\rm cut})\Big] = -\frac{\pi^2}{6}\,.
\end{aligned} \tag{B23}$$

Thus,

$$\begin{aligned}
(S\widetilde{C})(\mathcal{T}^{\rm cut}) &= -\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\theta(z)}{m_H}\hat{P}_{gg}(z)\Big(\frac{z}{1-z}\Big)^{2\epsilon} \\
&\quad\times\left(\frac{\mu^2}{m_H\mathcal{T}^{\rm cut}}\right)^{2\epsilon}\frac{\pi^2}{6}R^2.
\end{aligned} \tag{B24}$$

Expanding z as a distribution about $z = 1$ and keeping only the divergent term,

$$\theta(z)\hat{P}_{gg}(z)\Big(\frac{z}{1-z}\Big)^{2\epsilon} = -\frac{1}{\epsilon}\delta(1-z) + \mathcal{O}(\epsilon^0)\,, \tag{B25}$$

the soft-collinear mixing term is

$$(S\widetilde{C})(\mathcal{T}^{\rm cut}) = \frac{1}{\epsilon}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\delta(1-z)}{m_H}\left(\frac{\mu^2}{m_H\mathcal{T}^{\rm cut}}\right)^{2\epsilon}\frac{\pi^2}{6}R^2\,. \tag{B26}$$

Subtracting the zero-bin contribution, which is twice Eq. (B12), the correction to the cross section becomes

$$\begin{aligned}
\delta\sigma_{SC}^{\rm jet\,(2)}(\mathcal{T}^{\rm cut}) &= \sigma_{\rm LO}\frac{1}{\epsilon}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{6}R^2 \\
&\quad\times\left[\left(\frac{\mu^2}{m_H\mathcal{T}^{\rm cut}}\right)^{2\epsilon} - \left(\frac{\mu^2}{(\mathcal{T}^{\rm cut})^2}\right)^{2\epsilon}\right] \\
&= -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{\pi^2}{3}R^2\,\ln\frac{m_H}{\mathcal{T}^{\rm cut}}\,.
\end{aligned} \tag{B27}$$

For the p_{Tj} veto, the soft independent emission, and therefore the zero-bin contribution, are zero. Hence, the soft-collinear mixing terms are just given by the naive contribution. To evaluate the matrix element, we change variables from q^+, k^+ to p_{Tc}, p_{Ts}, where

$$q^+ = \frac{p_{Tc}^2 z}{m_H(1-z)}\,, \quad k^+ = p_{Ts}\,e^{\Delta y}\,\frac{m_H(1-z)}{p_{Tc}z}\,. \tag{B28}$$

We also use the same change of variables from k^- to Δy as in the $\mathcal{T}_j$ veto. There are again rapidity divergences not regulated by ϵ for which we use the analytic regulator. The regulator factor for (SC_a) is

$$\nu^{2\alpha}(q^+k^+)^{-\alpha} = \nu^{2\alpha}(p_{Tc}p_{Ts})^{-\alpha}\left(\frac{p_{Tc}z}{m_H(1-z)}\right)^{-2\alpha}. \tag{B29}$$

The integrals over Δy and $\Delta\phi$ can be performed as before, and we find

$$\begin{aligned}
(SC_a)(p_T^{\rm cut}) &= \Big(\frac{\alpha_s C_A}{\pi}\Big)^2\frac{(e^{\gamma_E}\mu^2)^{2\epsilon}}{\Gamma(1-\epsilon)^2}\frac{1}{m_H}\,\theta(z)\hat{P}_{gg}(z)\Big(\frac{1-z}{z}\Big)^{2\alpha}(\nu m_H)^{2\alpha}\,R^2 \\
&\quad\times\int_0^{\infty}\!\mathrm{d}p_{Tc}\,\mathrm{d}p_{Ts}\,p_{Tc}^{-1-2\epsilon-3\alpha}p_{Ts}^{-1-2\epsilon-\alpha}\Big[\theta(p_{Tc}+p_{Ts} < p_T^{\rm cut}) - \theta(p_{Tc} < p_T^{\rm cut})\,\theta(p_{Ts} < p_T^{\rm cut})\Big]\,.
\end{aligned} \tag{B30}$$

We can set $\epsilon = 0$ everywhere, as there is only a rapidity divergence as $z \to 1$. Performing the final integrals, we

find

$$(SC_a)(p_T^{\rm cut}) = -\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{\theta(z)}{m_H}\,\hat{P}_{gg}(z)\Big(\frac{1-z}{z}\Big)^{2\alpha} \times \left(\frac{\nu m_H}{(p_T^{\rm cut})^2}\right)^{2\alpha} \frac{\pi^2}{6} R^2\,. \tag{B31}$$

Expanding in z to extract the divergence, we get

$$(SC_a)(p_T^{\rm cut}) = -\frac{1}{\alpha}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{\delta(1-z)}{m_H}\left(\frac{\nu m_H}{(p_T^{\rm cut})^2}\right)^{2\alpha} \frac{\pi^2}{6} R^2\,. \tag{B32}$$

This gives the following correction to the cross section

$$\delta\sigma_{SC_a}^{\rm jet\,(2)}(p_T^{\rm cut}) = -\sigma_{\rm LO}\frac{1}{\alpha}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \left(\frac{\nu m_H}{(p_T^{\rm cut})^2}\right)^{2\alpha} \frac{\pi^2}{6} R^2\,. \tag{B33}$$

The regulator for the mixing between soft and n_b-collinear sectors gives a difference scale dependence. In this case the label component of collinear momentum is regulated, which effectively amounts to regulating the minus momentum component in the above calculation. This changes the scale dependence to $(\nu/m_H)^{2\alpha}$ with an overall minus sign:

$$\left(\frac{\nu m_H}{(p_T^{\rm cut})^2}\right)^{2\alpha} \to -(\nu/m_H)^{2\alpha}\,. \tag{B34}$$

Thus the entire soft-collinear mixing contribution to the cross section, $\delta\sigma_{SC_a}^{\rm jet} + \delta\sigma_{SC_b}^{\rm jet}$, is

$$\begin{aligned}\delta\sigma_{SC}^{\rm jet\,(2)}(p_T^{\rm cut}) &= -\sigma_{\rm LO}\frac{1}{\alpha}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{\pi^2}{6} R^2 \\ &\quad\times \left[\left(\frac{\nu m_H}{(p_T^{\rm cut})^2}\right)^{2\alpha} - \left(\frac{\nu}{m_H}\right)^{2\alpha}\right] \\ &= -\sigma_{\rm LO}\Big(\frac{\alpha_s C_A}{\pi}\Big)^2 \frac{2\pi^2}{3} R^2\,\ln\frac{m_H}{p_T^{\rm cut}}\,. \end{aligned} \tag{B35}$$

Appendix C: Clustering Logarithms in the Soft Function at $\mathcal{O}(\alpha_s^2)$

Clustering logarithms first occur at $\mathcal{O}(\alpha_s^2)$, and they are easiest to calculate in the soft function. RG invariance can be used to extract the beam function contributions, as explained in Sec. IV.

The definition of what is a clustering effect is subtle because one must define what the effect is *relative* to. One ostensibly natural option is to define it relative to the cross section if no clustering takes place. However, as explained in Sec. IV, this measurement is infrared unsafe due to collinear singularities between partons. A more sensible "primary" measurement are the inclusive $\mathcal{T}_B$ or E_T measurements. These measurements are IR safe, and when a set of particles becomes collinear their contribution to the primary measurement is the same as the contribution to the observable that is vetoed (the p_T or $\mathcal{T}$ of the jet). Here, we only keep the divergent term with the appropriate $p_T^{\rm cut}$ and $\mathcal{T}^{\rm cut}$ scale dependence, as this is the piece connected to the soft and beam functions. The finite terms do not take part in the resummation and can be captured as usual by matching the resummed result to the full fixed-order result at NNLO.

The $\mathcal{O}(\alpha_s^2)$ soft measurement function for the clustering correction relative to the inclusive measurements is [see Eq. (40)]

$$\begin{aligned}&\Delta\mathcal{M}_s^{\rm jet\,(2)}(k^{\rm cut}) \\ &\quad= \Big\{\theta(\Delta R_{12} < R)\theta(k_{\rm jet} < k^{\rm cut}) \\ &\qquad + \theta(\Delta R_{12} > R)\theta(k_1 < k^{\rm cut})\theta(k_2 < k^{\rm cut})\Big\} \\ &\qquad - \theta(k_1 + k_2 < k^{\rm cut})\,, \end{aligned} \tag{C1}$$

where $k = \mathcal{T}$ or p_T. In this measurement function, $k_{\rm jet}$ is the observable for the clustered pair. For $k = \mathcal{T}$, $\mathcal{T}_{\rm jet} = \mathcal{T}_1 + \mathcal{T}_2$ except when the jet spans the boundary at the Higgs rapidity. For $k = p_T$, the scalar sum of transverse momenta is not the same as the magnitude of the vector sum. However, when $R \sim \lambda$, these differences become power-suppressed for both the $\mathcal{T}$ and p_T observables. Since this is the limit we are working in, and we are neglecting power-suppressed terms in this limit, the measurement functions for the clustering effect simplify to

$$\begin{aligned}\Delta\mathcal{M}^{\rm jet\,(2)}(k^{\rm cut}) = \theta(\Delta R_{12} > R)\Big[&\theta(k_1 < k^{\rm cut})\theta(k_2 < k^{\rm cut}) \\ &- \theta(k_1 + k_2 < k^{\rm cut})\Big]\,. \end{aligned} \tag{C2}$$

The phase space constraint on ΔR_{12} suggests a convenient set of coordinates for the calculation. In terms of the rapidity y_i, azimuthal angle ϕ_i, and observable k_i (where $k = \mathcal{T}$ or p_T), the coordinates we use are

$$\begin{aligned} y_t &= \frac{1}{2}(y_1 + y_2)\,, & \phi_t &= \frac{1}{2}(\phi_1 + \phi_2)\,, \\ \Delta y &= y_1 - y_2\,, & \Delta\phi &= \phi_1 - \phi_2\,, \\ k_t &= k_1 + k_2\,, & z &= \frac{k_1}{k_t}\,. \end{aligned} \tag{C3}$$

The full $\mathcal{O}(\alpha_s^2)$ soft function matrix elements can be found in Ref. [55]. Using $k = p_T$, in terms of these coordinates the non-Abelian soft matrix elements are

$$\mathcal{A}_A = 4g^4 C_A^2 \, \frac{1}{p_{Tt}^4 \, z^2(1-z)^2} \, \frac{1}{\cosh\Delta y - \cos\Delta\phi} \, \frac{1}{z^2+(1-z)^2+2z(1-z)\cosh\Delta y}$$
$$\times \left\{ (z^2+(1-z)^2)\cos\Delta\phi - z(1-z)(1-\cos\Delta\phi\cosh\Delta y) \right.$$
$$\left. + (1-\epsilon)\frac{z^2(1-z)^2\sinh^2\Delta y}{(\cosh\Delta y-\cos\Delta\phi)(z^2+(1-z)^2+2z(1-z)\cosh\Delta y)} \right\},$$
$$\mathcal{A}_f = 4g^4 C_A T_R n_f \, \frac{1}{p_{Tt}^4 \, z^2(1-z)^2} \, \frac{1}{(\cosh\Delta y - \cos\Delta\phi)} \, \frac{1}{(z^2+(1-z)^2+2z(1-z)\cosh\Delta y)}$$
$$\times \left\{ z(1-z) - \frac{2z^2(1-z)^2\sinh^2\Delta y}{(\cosh\Delta y-\cos\Delta\phi)(z^2+(1-z)^2+2z(1-z)\cosh\Delta y)} \right\}. \tag{C4}$$

In the small ΔR limit, the matrix elements simplify to

$$\mathcal{A}_A^R = 4g^4 C_A^2 \, \frac{1}{p_{Tt}^4 z^2(1-z)^2} \, \frac{2}{\Delta R^2} \left[z^2+(1-z)^2 + \frac{2z^2(1-z)^2\Delta y^2}{\Delta R^2} \right],$$
$$\mathcal{A}_f^R = 4g^4 C_A T_R n_f \, \frac{1}{p_{Tt}^4 z^2(1-z)^2} \, \frac{2}{\Delta R^2} \left[z(1-z) - \frac{4z^2(1-z)^2\Delta y^2}{\Delta R^2} \right]. \tag{C5}$$

We start with the calculation for p_{Tj}. The measurement function can be written as

$$\Delta\mathcal{M}_s^{\text{jet}}(p_T^{\text{cut}}) = \theta(\Delta R > R)\,\theta\left[p_T^{\text{cut}} < p_{Tt} < p_T^{\text{cut}} \frac{1}{\max(z,1-z)} \right]. \tag{C6}$$

The matrix element and measurement function are independent of y_t, meaning the integral over y_t is unregulated. To regulate this rapidity divergence we use the rapidity regulator, which regulates the z-component of the group momentum for the c-web through the factor

$$\nu^\eta |2P_{3g}|^{-\eta} = \nu^\eta \, p_{Tt}^{-\eta} \left| 2z\sinh\Big(y_t + \frac{1}{2}\Delta y\Big) + 2(1-z)\sinh\Big(y_t - \frac{1}{2}\Delta y\Big) \right|^{-\eta}. \tag{C7}$$

Integrating over y_t then gives a single $1/\eta$ divergence and this is the only divergent part of the calculation,

$$\nu^\eta p_{Tt}^{-\eta} \int_{-\infty}^{\infty} \mathrm{d}y_t \left| 2z\sinh\Big(y_t + \frac{1}{2}\Delta y\Big) + 2(1-z)\sinh\Big(y_t - \frac{1}{2}\Delta y\Big) \right|^{-\eta} = \frac{2}{\eta}\nu^\eta p_{Tt}^{-\eta} + \mathcal{O}(\eta^0). \tag{C8}$$

The matrix element scales simply with p_{Tt} since it is the only dimensionful variable, and so p_{Tt} can be easily integrated against the measurement function including the regulator factor. This integral is

$$\int_0^\infty \mathrm{d}p_{Tt} \, p_{Tt}^{-1-4\epsilon-\eta} \, \theta\left[p_T^{\text{cut}} < p_{Tt} < p_T^{\text{cut}} \frac{1}{\max(z,1-z)} \right] = -(p_T^{\text{cut}})^{-4\epsilon-\eta} \ln[\max(z,1-z)]. \tag{C9}$$

Carrying through the integrals over the on-shell conditions, the soft function correction for the p_T veto is

$$\Delta S^{\text{b}\,(2)}(p_T^{\text{cut}}) = -\frac{8}{\eta} \, \frac{1}{(4\pi)^4} \left(\frac{\nu}{p_T^{\text{cut}}} \right)^\eta \int_0^1 \mathrm{d}z \int_{-\infty}^{\infty} \mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi} \, \frac{\ln[\max(z,1-z)]}{z(1-z)} \, \theta(\Delta R > R) \Big[p_{Tt}^4 z^2(1-z)^2 \mathcal{A}(z,\Delta y,\Delta\phi) \Big]. \tag{C10}$$

The remaining integrals are finite and have the form

$$a\ln R + b + \mathcal{O}(R). \tag{C11}$$

We determine a analytically and extract b numerically. To determine the coefficient of $\ln R$, we rewrite the matrix element as the difference

$$\mathcal{A} = (\mathcal{A} - \mathcal{A}^R) + \mathcal{A}^R, \tag{C12}$$

where $\mathcal{A}^R$ is the matrix element expanded in the small R limit, given in Eq. (C5). The difference $\mathcal{A}-\mathcal{A}^R$ is finite as $R\to 0$, meaning that up to $\mathcal{O}(R^2)$ corrections,

$$\int \mathrm{d}\Delta y \mathrm{d}\Delta\phi\, \theta(\Delta R > R)\,(\mathcal{A}-\mathcal{A}^R) = \int \mathrm{d}\Delta y\, \mathrm{d}\Delta\phi\,(\mathcal{A}-\mathcal{A}^R) + \mathcal{O}(R^2)\,. \tag{C13}$$

Integrating the matrix element $\mathcal{A}^R$ is simple, using the relations

$$\begin{aligned}
&\int_{-\infty}^{\infty}\mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi}\,\frac{1}{\Delta R^2}\,\theta(\Delta R > R) = -\ln R + \ln 2\pi + \mathcal{O}(R)\,,\\
&\int_{-\infty}^{\infty} d\Delta y \int_0^\pi \frac{d\Delta\phi}{\pi}\,\frac{2\Delta y^2}{\Delta R^4}\,\theta(\Delta R > R) = -\ln R + \ln 2\pi + \frac{1}{2} + \mathcal{O}(R)\,.
\end{aligned} \tag{C14}$$

Carrying through the integrals, we obtain

$$\begin{aligned}
&\int_0^1 \mathrm{d}z \int_{-\infty}^{\infty}\mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi}\,\frac{\ln[\max(z,1-z)]}{z(1-z)}\,\Big[p_{Tt}^4 z^2(1-z)^2\mathcal{A}_A^R\Big]\theta(\Delta R > R)\\
&\quad = 4g^4 C_A^2\Big\{-\frac{1}{36}(131-12\pi^2-132\ln 2)\ln R + \frac{1}{72}\big[-13+12\ln 2+2\ln(2\pi)(131-12\pi^2-132\ln 2)\big]\Big\}\,,\\
&\int_0^1 \mathrm{d}z \int_{-\infty}^{\infty}\mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi}\,\frac{\ln[\max(z,1-z)]}{z(1-z)}\Big[p_{Tt}^4 z^2(1-z)^2\mathcal{A}_f^R\Big]\theta(\Delta R > R)\\
&\quad = 4g^4 C_A T_R n_f\Big\{\frac{1}{18}(23-24\ln 2)\ln R + \frac{1}{36}\big[13-12\ln 2-2\ln(2\pi)(23-24\ln 2)\big]\Big\}\,.
\end{aligned} \tag{C15}$$

Performing the full integrals numerically, we get

$$\begin{aligned}
&\int_0^1 \mathrm{d}z \int_{-\infty}^{\infty}\mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi}\,\frac{\ln[\max(z,1-z)]}{z(1-z)}\Big[p_{Tt}^4 z^2(1-z)^2(\mathcal{A}_A-\mathcal{A}_A^R)\Big] = 4g^4 C_A^2\,(4.66)\,,\\
&\int_0^1 \mathrm{d}z \int_{-\infty}^{\infty}\mathrm{d}\Delta y \int_0^\pi \frac{\mathrm{d}\Delta\phi}{\pi}\,\frac{\ln[\max(z,1-z)]}{z(1-z)}\Big[p_{Tt}^4 z^2(1-z)^2(\mathcal{A}_f-\mathcal{A}_f^R)\Big] = 4g^4 C_A T_R n_f\,(0.138)\,.
\end{aligned} \tag{C16}$$

Thus, the final result for the leading divergent corrections due to clustering in the bare soft function is

$$\begin{aligned}
\Delta S^{\mathrm{b}\,(2)}(p_T^{\mathrm{cut}}) = \frac{1}{\eta}\Big(\frac{\alpha_s}{\pi}\Big)^2\Big(\frac{\nu}{p_T^{\mathrm{cut}}}\Big)^\eta\Big\{&C_A^2\Big[\frac{1}{18}(131-12\pi^2-132\ln 2)\ln R - 1.12\Big]\\
&+ C_A T_R n_f\Big[-\frac{1}{9}(23-24\ln 2)\ln R + 0.764\Big]\Big\}\,.
\end{aligned} \tag{C17}$$

Note that the constant terms, -1.12 and $+0.764$, depend on the choice of inclusive observable that the clustering effect is defined relative to. Hence, they are different from those found in Ref. [14], since here we use $E_T = |\vec{p}_{T1}|+|\vec{p}_{T2}|$ while Ref. [14] uses the vector sum $|\vec{p}_{T1}+\vec{p}_{T2}|$. When using the latter in our calculation we reproduce the results in Ref. [14].

The steps to calculate the clustering logarithms for $\mathcal{T}_j$ proceed analogously. The calculation is slightly more tedious due to the fact that a particle's contribution to the observable changes depending on what hemisphere it is in. Therefore, there are two regions of rapidities $y_{1,2}$ of the two particles to consider:

i) $y_1, y_2 > 0$ and $y_1, y_2 < 0$: Both particles are in the same hemisphere. These configurations will contribute to the divergent terms.

ii) $y_1 > 0, y_2 < 0$ and $y_1 < 0, y_2 > 0$: The particles are in opposite hemispheres. This region does not contribute to the divergent terms. The region of phase space where particles can cluster scales as $\mathcal{O}(R^4)$ and can be neglected.

The variables in Eq. (C3) are also useful for this calculation. Carrying through the calculation, we find the divergent terms contributing to the bare soft function

$$\begin{aligned}
\Delta S^{\mathrm{b}\,(2)}(\mathcal{T}^{\mathrm{cut}}) = \frac{1}{4\epsilon}\Big(\frac{\alpha_s}{\pi}\Big)^2\Big(\frac{\mu}{\mathcal{T}^{\mathrm{cut}}}\Big)^{4\epsilon}\Big\{&C_A^2\Big[\frac{1}{18}\big(131-12\pi^2-132\ln 2\big)\ln R - 0.937\Big]\\
&+ C_A T_R n_f\Big[-\frac{1}{9}\big(23-24\ln 2\big)\ln R + 0.747\Big]\Big\}\,.
\end{aligned} \tag{C18}$$

Note that the divergent $\ln R$ terms are the same as for p_{Tj}, while the constant terms differ.

[1] ATLAS Collaboration, G. Aad *et al.*, Phys. Lett. B **710**, 49 (2012), [arXiv:1202.1408].
[2] CMS Collaboration, S. Chatrchyan *et al.*, Phys. Lett. B **710**, 26 (2012), [arXiv:1202.1488].
[3] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. D **81**, 094035 (2010), [arXiv:0910.0467].
[4] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. Lett. **106**, 032001 (2011), [arXiv:1005.4060].
[5] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **04**, 092 (2011), [arXiv:1012.4480].
[6] A. Papaefstathiou, J. M. Smillie, and B. R. Webber, JHEP **04**, 084 (2010), [arXiv:1002.4375].
[7] CMS Collaboration, V. Khachatryan *et al.*, Phys. Lett. B **699**, 48 (2011), [arXiv:1102.0068].
[8] ATLAS Collaboration, G. Aad *et al.*, arXiv:1206.2135.
[9] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. B **269**, 432 (1991).
[10] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
[11] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993), [hep-ph/9305266].
[12] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997), [hep-ph/9707323].
[13] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), [arXiv:0802.1189].
[14] A. Banfi, G. P. Salam, and G. Zanderighi, JHEP **06**, 159 (2012), [arXiv:1203.5773].
[15] T. Becher and M. Neubert, JHEP **1207**, 108 (2012), [arXiv:1205.3806].
[16] A. Banfi and M. Dasgupta, Phys. Lett. B **628**, 49 (2005), [hep-ph/0508159].
[17] Y. Delenda, R. Appleby, M. Dasgupta, and A. Banfi, JHEP **12**, 044 (2006), [hep-ph/0610242].
[18] K. Khelifa-Kerfa, JHEP **02**, 072 (2012), [arXiv:1111.2016].
[19] A. Hornig, C. Lee, J. R. Walsh, and S. Zuberi, JHEP **01**, 149 (2012), [arXiv:1110.0004].
[20] R. Kelley, J. R. Walsh, and S. Zuberi, arXiv:1202.2361.
[21] R. Kelley, J. R. Walsh, and S. Zuberi, arXiv:1203.2923.
[22] C. W. Bauer, S. Fleming, and M. E. Luke, Phys. Rev. D **63**, 014006 (2000), [hep-ph/0005275].
[23] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, Phys. Rev. D **63**, 114020 (2001), [hep-ph/0011336].
[24] C. W. Bauer and I. W. Stewart, Phys. Lett. B **516**, 134 (2001), [hep-ph/0107001].
[25] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002), [hep-ph/0109045].
[26] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. D **66**, 014017 (2002), [hep-ph/0202088].
[27] M. Beneke, A. Chapovsky, M. Diehl, and T. Feldmann, Nucl. Phys. **B643**, 431 (2002), [hep-ph/0206152].
[28] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **09**, 005 (2010), [arXiv:1002.2213].
[29] S. Fleming, A. K. Leibovich, and T. Mehen, Phys. Rev. D **74**, 114004 (2006), [hep-ph/0607121].
[30] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, Phys. Rev. Lett. **108**, 151601 (2012), [arXiv:1104.0881].
[31] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, JHEP **05**, 084 (2012), [arXiv:1202.0814].
[32] S. Mantry and F. Petriello, Phys. Rev. D **81**, 093007 (2010), [arXiv:0911.4135].
[33] S. Mantry and F. Petriello, Phys. Rev. D **83**, 053007 (2011), [arXiv:1007.3773].
[34] T. Becher and M. Neubert, Eur. Phys. J. **C71**, 1665 (2011), [arXiv:1007.4005].
[35] A. Jain, M. Procura, and W. J. Waalewijn, JHEP **04**, 132 (2012), [arXiv:1110.0839].
[36] M. Garcia-Echevarria, A. Idilbi, and I. Scimemi, JHEP **07**, 002 (2012), [arXiv:1111.4996].
[37] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, JHEP **11**, 101 (2010), [arXiv:1001.0014].
[38] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion, and J. R. Walsh, Phys. Lett. B **689**, 82 (2010), [arXiv:0912.0262].
[39] W. M.-Y. Cheung, M. Luke, and S. Zuberi, Phys. Rev. D **80**, 114021 (2009), [arXiv:0910.2479].
[40] T. T. Jouttenus, Phys. Rev. D **81**, 094017 (2010), [arXiv:0912.5509].
[41] J. R. Walsh and S. Zuberi, arXiv:1110.5333.
[42] T. Becher and G. Bell, Phys. Lett. B **713**, 41 (2012), [arXiv:1112.3907].
[43] J. G. M. Gatheral, Phys. Lett. B **133**, 90 (1983).
[44] J. Frenkel and J. C. Taylor, Nucl. Phys. **B246**, 231 (1984).
[45] E. Gardi, E. Laenen, G. Stavenga, and C. D. White, JHEP **11**, 155 (2010), [arXiv:1008.0098].
[46] A. Mitov, G. Sterman, and I. Sung, Phys.Rev. D **82**, 096010 (2010), [arXiv:1008.0099].
[47] C. Anastasiou, K. Melnikov, and F. Petriello, Phys. Rev. Lett. **93**, 262002 (2004), [hep-ph/0409088].
[48] C. Anastasiou, K. Melnikov, and F. Petriello, Nucl. Phys. **B724**, 197 (2005), [hep-ph/0501130].
[49] S. Catani and M. Grazzini, Phys. Rev. Lett. **98**, 222002 (2007), [hep-ph/0703012].
[50] M. Grazzini, JHEP **02**, 043 (2008), [arXiv:0801.3232].
[51] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999), [hep-ph/9905386].
[52] J. M. Campbell, R. K. Ellis, and C. Williams, Phys. Rev. D **81**, 074023 (2010), [arXiv:1001.4495].
[53] I. W. Stewart and F. J. Tackmann, Phys. Rev. D **85**, 034011 (2012), [arXiv:1107.2117].
[54] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, arXiv:1206.4998.
[55] A. Hornig, C. Lee, I. W. Stewart, J. R. Walsh, and S. Zuberi, JHEP **08**, 054 (2011), [arXiv:1105.4628].