Two- and three-loop data for the groomed jet mass

Adam Kardos[*]

University of Debrecen, 4010 Debrecen, PO Box 105, Hungary

Andrew J. Larkoski[†]

Physics Department, Reed College, Portland, Oregon 97202, USA

Zoltán Trócsányi[‡]

Institute for Theoretical Physics, ELTE Eötvös Loránd University,
Pázmány Péter 1/A, H-1117 Budapest, Hungary
and MTA-DE Particle Physics Research Group, University of Debrecen,
4010 Debrecen, PO Box 105, Hungary

We discuss the status of resummation of large logarithmic contributions to groomed event shapes of hadronic final states in electron-positron annihilation. We identify the missing ingredients needed for next-to-next-to-next-to-leading logarithmic (NNNLL) resummation of the modified mass mass drop tagger (mMDT) groomed jet mass in e^+e^- collisions: the low-scale collinear-soft constants at two-loop accuracy, $c_{S_c}^{(2)}$, and the three-loop noncusp anomalous dimension of the global soft function, $\gamma_S^{(2)}$. We present a method for extracting those constants using fixed-order codes: the EVENT2 program to obtain the color coefficients of $c_{S_c}^{(2)}$, and MCCSM for extracting $\gamma_S^{(2)}$. We present all necessary formulae for resummation of the mMDT groomed heavy jet mass distribution at NNNLL accuracy.

I. INTRODUCTION

At the Large Hadron Collider (LHC), quantum chromodynamics (QCD) has been studied at an unprecedented precision in both experiment and theory. Major advances have especially recently been made in the field of jet substructure in which the internal dynamics of jets are exploited to learn information about the short-distance physics; see Refs. [1,2] for recent reviews. Because jets are identified as restricted regions of phase space, particles that land in a jet can come from numerous sources from outside of the jet that contaminate a simple theoretical picture of the jet's dynamics. To mitigate this contamination radiation, which may come from initial-state radiation, secondary parton scatterings, or secondary proton scatterings (pile-up), analysis techniques broadly referred to as jet grooming have been introduced. The jet groomers

systematically identify emissions in the jet that likely came from such contamination and remove them from consideration. The best theoretically understood jet groomers are the modified mass mass drop tagger (mMDT) [3,4] and soft drop [5] algorithms, due to their unique feature of elimination of nonglobal logarithms (NGLs) [6], leading correlations between in-jet and out-of-jet scales.

The elimination of NGLs of mMDT and soft drop groomers means that calculations of observables can be meaningfully extended beyond leading-logarithmic accuracy. References [7,8] derived an all-orders factorization theorem for observables like the jet mass measured on mMDT/soft drop groomed jets, in the limit that the jet mass is small with respect to the jet energy. This factorization theorem enabled calculation to next-to-next-to-leading logarithmic (NNLL) accuracy for jet observables, resulting in predictions with controlled theoretical uncertainties over a wide, and experimentally-relevant, region of phase space. Many other calculations for mMDT/soft drop groomed jet mass (and related observables) have followed [9–21], and both ATLAS and CMS have measured these groomed jet masses in collisions [22–24] and compared directly to theoretical predictions. Further calculations have been performed of groomed observables that are sensitive to additional structure in jets [25–28], and related grooming

procedures have been introduced [29] that exploit the nice theoretical properties of these groomers.

With the high theoretical precision achievable with mMDT/soft drop and the continual experimental measurements, a realistic goal of this program is a first extraction of underlying physical quantities from jet substructure. First studies have demonstrated that distributions of observables on these groomed jets are very sensitive to the value of α_s [30]. Further, grooming promises to provide an unambiguous theoretical definition of the top quark mass, and a universal robustness to its value [31]. To reach the necessary theoretical precision for these measurements requires not only high logarithmic accuracy, but also high fixed-order accuracy and well-understood nonperturbative corrections. Next-to-next-to-leading fixed order (NNLO) predictions for mMDT and soft drop groomed events have been produced using the MCCSM (Monte Carlo for the CoLoRFulNNLO Subtraction Method) code [32,33] for jet production in e^+e^- collisions, and a detailed analysis of nonperturbative corrections within the context of the factorization theorem for groomed jets has been presented [34]. Further, advances in NNLO multijet production in hadron collisions like at the LHC have been steadily made over the past few years [35–44].

For precision extraction of the strong coupling, for example, logarithmic accuracy should extend to next-to-next-to-next-to-leading logarithmic order (NNNLL), the first order at which simple additive matching to NNLO fixed-order is possible. NNNLL is the order to which thrust [45,46] and C-parameter [47,48] are resummed for precision extraction of α_s from LEP data. Such high precision is possible because improving logarithmic accuracy requires calculation of universal anomalous dimensions and constants for the functions that appear in the factorization theorem for these observables. In the case of mMDT/soft drop groomed jet mass, the factorization theorem for the process $e^+e^- \to$ hadrons reads

$$\frac{1}{\sigma_0}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\tau_\mathrm{L}\mathrm{d}\tau_\mathrm{R}} = H(Q^2)S(z_\mathrm{cut})[J(\tau_\mathrm{L}) \otimes S_c(\tau_\mathrm{L}, z_\mathrm{cut})]$$
$$\times [J(\tau_\mathrm{R}) \otimes S_c(\tau_\mathrm{R}, z_\mathrm{cut})], \quad (1)$$

where σ_0 is the leading-order $e^+e^- \to q\bar{q}$ cross section, z_cut is a parameter of the mMDT/soft drop groomer, and τ_L, τ_R are the left- and right-hemisphere masses, respectively. Extending this factorization theorem beyond NNLL requires the two-loop constants for each of the functions on the right, as well as their three-loop anomalous dimensions. The hard $H(Q^2)$ and jet functions $J(\tau)$ are known completely through three loops, and the global soft function $S(z_\mathrm{cut})$ is known completely through two loops [49–51]. Consistency of the factorization theorem requires the sum of anomalous dimensions of the functions to vanish so that the cross section is renormalization-group invariant. Therefore, for NNNLL resummation, we only need to

calculate the two-loop constants of the collinear-soft function $S_c(\tau, z_\mathrm{cut})$ and the three-loop global soft function anomalous dimension.

In this paper, we use numerical fixed-order event generators to extract these necessary ingredients. We will restrict our analysis to jet masses groomed with mMDT for simplicity and compactness. In this case, we use EVENT2 [52] to find the two-loop constants $c_{S_c}^{\mathrm{mMDT}}$ of the collinear-soft function (in Laplace conjugate space) to be

$$c_{S_c}^{\mathrm{mMDT}} = \left(\frac{\alpha_s}{4\pi}\right)^2 [C_F^2(22 \pm 4) + C_F C_A(41 \pm 1)$$
$$+ C_F T_R n_f(14.4 \pm 0.1)], \quad (2)$$

separated into distinct color channels where $C_F = 4/3$, $C_A = 3$, and $T_R = 1/2$ in QCD, and n_f is the number of active quark flavors. Using MCCSM, we find the three-loop anomalous dimension of the global soft function γ_S^{mMDT} to be

$$\gamma_S^{\mathrm{mMDT}} = \left(\frac{\alpha_s}{4\pi}\right)^3 [-11600 \pm 2000] \quad (n_f = 5), \quad (3)$$

where we fix the number of active quarks to $n_f = 5$. These results enable resummation to NNNLL accuracy for jet substructure observables for the first time, which we present in a companion paper [53].

The outline of this paper is as follows. In Sec. II, we review the mMDT grooming algorithm and its all-orders factorization theorem. In Sec. III, we discuss our procedure for extracting the two-loop constants of the collinear-soft function. In Sec. IV, we extract the three-loop anomalous dimension of the global soft function using numerical results from MCCSM. We conclude in Sec. V. Several Appendixes provide the analytical and numerical details of the results provided in the main text.

II. mMDT GROOMING AND FACTORIZATION THEOREM

In this paper, we restrict our analysis to hadronic final states in electron-positron annihilation where high fixed-order results are available and can be used to extract currently-unknown data for resummation. As such, we will only discuss the grooming algorithm and observables for e^+e^- collisions, and we point readers interested in other applications to the cited literature. In this section, we will review the definition of the grooming algorithm, as well as the all-orders factorization theorem that we exploit.

The mMDT groomer [3], or soft drop with angular exponent $\beta = 0$ [5], proceeds as follows:
  (1) Divide the final state of an $e^+e^- \to$ hadrons event into two hemispheres in any infrared and collinear safe way. This could include separation with thrust

[54,55], or in a recoil-free manner using broadening axes [56,57].

(2) In each hemisphere, run the Cambridge/Aachen jet algorithm [58,59] to produce an angular-ordered pairwise clustering history of particles. The clustering metric appropriate for e^+e^- collisions is

$$d_{ij} = 1 - \cos\theta_{ij}, \qquad (4)$$

where θ_{ij} is the angle between particles i and j in the same hemisphere.

(3) Starting with the left hemisphere and at widest angle, step through the Cambridge/Aachen particle branching tree. At each branching in the tree, test if

$$\frac{\min[E_i, E_j]}{E_i + E_j} > z_{\text{cut}}, \qquad (5)$$

where i and j are the daughter particles at that branching and z_{cut} is some fixed numerical value where $0 \leq z_{\text{cut}} < 0.5$. If this is true, stop and return all particles that remain in the left hemisphere. If it is false, remove the lower energy branch, and continue to the next branching at smaller angle. Repeat for the right hemisphere.

(4) Once the groomer has terminated, any observable can be measured on the particles that remain in the two hemispheres.

We will refer to this procedure as the mMDT groomer or simply mMDT throughout this paper.

From this procedure, Ref. [8] proved an all-orders factorization theorem for the cross section differential in the groomed hemisphere masses τ_L and τ_R, where

$$\tau_i = \frac{m_i^2}{E_i^2}, \qquad (6)$$

for mass m_i and energy E_i of hemisphere i. For $\tau_L, \tau_R \ll z_{\text{cut}} \ll 1$, the cross section factorizes into:

$$\frac{1}{\sigma_0} \frac{d^2\sigma}{d\tau_L d\tau_R} = H(Q^2) S(z_{\text{cut}}) [J(\tau_L) \otimes S_c(\tau_L, z_{\text{cut}})]$$
$$\times [J(\tau_R) \otimes S_c(\tau_R, z_{\text{cut}})], \qquad (7)$$

where σ_0 is the leading-order cross section for $e^+e^- \to q\bar{q}$, $H(Q^2)$ is the hard function for quark–antiquark production in e^+e^- collisions, $S(z_{\text{cut}})$ is the global soft function for mMDT grooming, $J(\tau_i)$ is the quark jet function for hemisphere mass τ_i, $S_c(\tau_i, z_{\text{cut}})$ is the collinear-soft function for hemisphere mass τ_i with mMDT grooming, and we suppress the dependence on the renormalization scale μ in all functions. The symbol $\otimes$ denotes convolution over the hemisphere mass τ_i. In most of the expressions that follow, for simplicity we will set $\sigma_0 = 1$ GeV^{-2}. While we

will calculate data relevant for the functions in this factorization theorem in this paper, we will do so through the single-differential cross section of the groomed heavy hemisphere mass, defined as

$$\frac{1}{\sigma_0} \frac{d\sigma_g}{d\rho} \equiv \int d\tau_L d\tau_R \frac{1}{\sigma_0} \frac{d^2\sigma}{d\tau_L d\tau_R} [\Theta(\tau_L - \tau_R)\delta(\rho - \tau_L)$$
$$+ \Theta(\tau_R - \tau_L)\delta(\rho - \tau_R)]. \qquad (8)$$

The subscript g on the cross section denotes that this is groomed. Additionally, we note that this definition of the heavy hemisphere mass differs from the standard definition in the ungroomed case when the heavy hemisphere mass is defined as:

$$\rho = \frac{\max[m_L^2, m_R^2]}{Q^2}, \qquad (9)$$

with Q being the center-of-mass energy. When hemispheres are groomed, the grooming eliminates their dominant correlations, and so it is more natural to define the groomed mass with respect to the hemisphere energy, and not the center-of-mass energy.

The factorization theorem is simplest to analyze in Laplace space, where we Laplace transform in both τ_L and τ_R to eliminate the convolutions. In Laplace space, the cross section becomes a simple product:

$$\frac{\sigma(\nu_L, \nu_R)}{\sigma_0} = H(Q^2) S(z_{\text{cut}}) \tilde{J}(\nu_L) \tilde{S}_c(\nu_L, z_{\text{cut}}) \tilde{J}(\nu_R) \tilde{S}_c(\nu_R, z_{\text{cut}}),$$
$$(10)$$

where ν_L (ν_R) is the Laplace conjugate of τ_L (τ_R). In this product form, each function in the factorization theorem satisfies a simple renormalization group equation (RGE),

$$\mu \frac{\partial \tilde{F}}{\partial \mu} = \left(d_F \Gamma_{\text{cusp}} \log \frac{\mu^2}{\mu_F^2} + \gamma_F \right) \tilde{F}, \quad (\tilde{F} = H, S, \tilde{J}, \tilde{S}_c)$$
$$(11)$$

where d_F is a constant, μ_F is the canonical scale, and γ_F is the noncusp anomalous dimension particular to $\tilde{F}$. Γ_{cusp} is the cusp anomalous dimension for back-to-back light-like Wilson lines in the fundamental representation of color SU(3). Large logarithms of hemisphere masses can be resummed to all orders in α_s through this renormalization group equation. It can be solved exactly, and we present its explicit solution through α_s^3 in Appendix B.

The order to which logarithms can be resummed through this renormalization group equation depends on the accuracy to which its components are calculated. For the canonical definition of logarithmic accuracy [60], Table I shows the order in α_s to which the components of the renomalization group equation are needed. In this table,

TABLE I. $\alpha_{\rm s}$-order of ingredients needed for resummation to the logarithmic accuracy given. $\Gamma_{\rm cusp}$ is the cusp anomalous dimension, γ_F is the noncusp anomalous dimension for function $\tilde{F}$, β is the QCD β-function, and c_F are the low-scale constants for function $\tilde{F}$. The final column shows the relative order to which the resummed cross section can be additively matched to fixed-order.

	$\Gamma_{\rm cusp}$	γ_F	β	c_F	Matching
LL	$\alpha_{\rm s}$		$\alpha_{\rm s}$		
NLL	$\alpha_{\rm s}^2$	$\alpha_{\rm s}$	$\alpha_{\rm s}^2$		$\alpha_{\rm s}$
NNLL	$\alpha_{\rm s}^3$	$\alpha_{\rm s}^2$	$\alpha_{\rm s}^3$	$\alpha_{\rm s}$	$\alpha_{\rm s}^2$
NNNLL	$\alpha_{\rm s}^4$	$\alpha_{\rm s}^3$	$\alpha_{\rm s}^4$	$\alpha_{\rm s}^2$	$\alpha_{\rm s}^3$

"LL" denotes leading logarithmic accuracy, "NLL" is next-to-leading logarithmic accuracy, etc. The factorization theorem successfully resums all logarithms of both ρ and $z_{\rm cut}$ simultaneously, to leading power in the limit where $\rho \ll z_{\rm cut} \ll 1$. Thus "NLL," for example, refers to the resummation of all terms up through those of the form $\alpha_s^n \log^n \rho$, $\alpha_s^n \log^n z_{\rm cut}$, and $\alpha_s^n \log^n \frac{\rho}{z_{\rm cut}}$ in the exponent of the cross section cumulative in the mass ρ. Reference [8] resummed the mMDT groomed mass distribution to NNLL accuracy, requiring the new calculation of one-loop constants and two-loop noncusp anomalous dimensions. The goal of this paper is to extend the order of the known components to accomplish NNNLL resummation.

The cusp anomalous dimension is now known to four-loop order [61–74], and the QCD β-function is as well [75]. The hard and jet functions are known through three-loop order [45,75–81], as they are relevant for resummation of a broad class of observables. The global mMDT soft function is now completely known to two-loops [8,49–51], while only the two-loop noncusp anomalous dimension of the collinear-soft function is known. Thus, to achieve NNNLL accurate resummation of mMDT groomed mass, we need the two-loop collinear-soft function constants, and the three-loop noncusp anomalous dimensions of the soft and collinear-soft functions. Actually, because the cross section is renormalization-group invariant, the sum of noncusp anomalous dimensions of the functions in the factorization theorem must vanish:

$$0 = \gamma_H + \gamma_S + 2\gamma_J + 2\gamma_{S_c}. \qquad (12)$$

Hence, only the two-loop constants of the collinear-soft function and the three-loop noncusp anomalous dimension of the global soft function are needed to accomplish resummation at NNNLL accuracy. In the following sections, we numerically extract these values from fixed-order codes. All results needed for solving the renomalization group equations to $\mathcal{O}(\alpha_s^3)$ are provided in the Appendixes.

Before moving to the numerical extraction, it is useful to comment on the region of phase space in which mMDT

grooming acts nontrivially. Assuming that the grooming parameter is small, $z_{\rm cut} \ll 1$ formally, then only low-energy particles in each hemisphere are groomed away. To leading power in a soft particle's energy E_s, the mMDT constraint is

$$E_s > E_H z_{\rm cut} = \frac{Q}{2} z_{\rm cut}, \qquad (13)$$

where $E_H = Q/2$ is the hemisphere energy, which is half of the center-of-mass collision energy in the soft limit. For such a soft emission that just passes mMDT, its contribution to the hemisphere mass is

$$\rho = \frac{m_H^2}{E_H^2} = \frac{2E_H E_s(1 - \cos\theta_s)}{E_H^2} \leq 2\frac{E_s}{E_H}, \qquad (14)$$

where the inequality corresponds to a soft emission right at the hemisphere boundary. Taking the upper bound as the parametric scaling of the soft particle's energy, we have

$$E_s \sim \frac{\rho E_H}{2}. \qquad (15)$$

Then, using the mMDT constraint, the value of the hemisphere mass for such an emission is

$$\rho \gtrsim 2z_{\rm cut}, \qquad (16)$$

and mMDT strongly affects the cross section for values of ρ smaller than this.

III. EXTRACTION OF TWO-LOOP CONSTANTS

In this section, we will present a method for numerical extraction of the two-loop constant terms of the collinear-soft function $c_{S_c}^{(2)}$ for the factorization theorem formulated in Laplace space, Eq. (10). Our procedure for doing so will be to relate the cross section differential in the groomed hemisphere mass ρ to the total cross section. We can express the leading-power (LP) differential cross section for $\rho \to 0$ as a sum of terms with support exclusively at $\rho = 0$ and terms with support away from 0:

$$\frac{d\sigma_{\rm g,LP}}{d\rho} = D_{\delta,\rm g}\delta(\rho) + \frac{d\sigma_{\rm g}^{\rm sing}}{d\rho}. \qquad (17)$$

The terms that have support away from $\rho = 0$ are defined to integrate to 0 on $\rho \in [0, 1]$. The total cross section can then be written as:

$$\sigma_{\rm tot} = D_{\delta,\rm g} + \int_0^1 d\rho \left(\frac{d\sigma_{\rm g}}{d\rho} - \frac{d\sigma_{\rm g}^{\rm sing}}{d\rho} \right), \qquad (18)$$

where $d\sigma_{\rm g}/d\rho$ is the full differential cross section of the groomed hemisphere mass for all $\rho > 0$.

The desired unknown two-loop constants are contained in the δ-function coefficient, $D_{\delta,\text{g}}$. One approach to determining them is to numerically evaluate the integral that remains using the known total cross section for $e^+e^- \to$ hadrons and a fixed-order code. Performing this particular numerical integral is somewhat complicated by the fact that grooming introduces multiple regions and so the integrand exhibits cusps and has support all the way out to the maximal value of ρ. We will significantly rewrite this expression into a more directly useful form so that the numerical integrals we must perform are restricted to the region where grooming turns on, $\rho \sim z_{\text{cut}}$. First, we separate the integral into two parts,

$$\int_0^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right) = \int_0^{2z_{\text{cut}}} \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right)$$
$$+ \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \frac{d\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho}. \tag{19}$$

Below $2z_{\text{cut}}$ mMDT grooming does, while above $2z_{\text{cut}}$ it does not, dominantly enforce constraints on the radiation. Then we add and subtract the singular differential cross section for ungroomed heavy hemisphere mass to the second integral on the right of Eq. (19):

$$\int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} = \int_{2z_{\text{cut}}}^4 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho}\Theta(1-\rho) - \frac{\mathrm{d}\sigma}{\mathrm{d}\rho} \right) + \int_{2z_{\text{cut}}}^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma}{\mathrm{d}\rho}. \tag{20}$$

For $\rho \sim 1 \gg z_{\text{cut}}$, mMDT grooming has no effect, so the first integral in this expression is dominated by the region near $\rho \simeq z_{\text{cut}}$, and we can drop the step function Θ. Integrals on the right extend to 4, rather than 1 as simple consequence of the factor of 4 difference in definition of the groomed versus ungroomed heavy hemisphere mass, as discussed in Sec. II. For the integral exclusively of the ungroomed heavy hemisphere mass cross section, we can write it as

$$\int_{2z_{\text{cut}}}^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma}{\mathrm{d}\rho} = \int_0^4 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} \right) + \mathcal{O}(z_{\text{cut}})$$
$$+ \int_{2z_{\text{cut}}}^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho}. \tag{21}$$

Assuming $z_{\text{cut}} \ll 1$, the lower bound of the first integral can be set to 0, neglecting power corrections that vanish as $z_{\text{cut}} \to 0$.

After implementing all these identities, the total cross section can be written as

$$\sigma_{\text{tot}} = D_{\delta,\text{g}} + \int_0^{2z_{\text{cut}}} \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right)$$
$$+ \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma}{\mathrm{d}\rho} \right) + \int_0^4 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} \right)$$
$$+ \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right)$$
$$+ \int_1^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} + \mathcal{O}(z_{\text{cut}}). \tag{22}$$

Further progress can be made by replacing the total cross section σ_{tot} with its expression as an integral over the ungroomed hemisphere mass cross section. Writing the ungroomed hemisphere mass cross section at LP as

$$\frac{\mathrm{d}\sigma_{\text{LP}}}{\mathrm{d}\rho} = D_\delta \delta(\rho) + \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho}, \tag{23}$$

the total cross section can also be expressed as

$$\sigma_{\text{tot}} = D_\delta + \int_0^4 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} \right). \tag{24}$$

Using Eqs. (22) and (24), we find the relationship

$$D_\delta = D_{\delta,\text{g}} + \int_0^{2z_{\text{cut}}} \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right) + \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma}{\mathrm{d}\rho} \right)$$
$$+ \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing}}}{\mathrm{d}\rho} \right) + \int_1^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma^{\text{sing}}}{\mathrm{d}\rho}. \tag{25}$$

Through two-loop order, the only unknowns in this equation are the collinear-soft function constants $c_{S_c}^{(2)}$ within the $D_{\delta,\text{g}}$ term and the first two integrals. However, those integrals can be numerically evaluated with a fixed-order code. We will first validate this relationship at one-loop, and then use it at two-loops to determine $c_{S_c}^{(2)}$.

A. Test at one-loop

At one-loop order, Eq. (25) simplifies significantly. For $\rho > 2z_{\text{cut}}$, the groomed and ungroomed differential cross sections are identical, and so the third term in Eq. (25) is identically zero. Hence, the one-loop relationship between the groomed and ungroomed cross sections can be expressed as

$$D_\delta^{(1)} - D_{\delta,\text{g}}^{(1)} - \int_{2z_{\text{cut}}}^1 \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma^{\text{sing},(1)}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing},(1)}}{\mathrm{d}\rho} \right)$$
$$- \int_1^4 \mathrm{d}\rho \frac{\mathrm{d}\sigma^{\text{sing},(1)}}{\mathrm{d}\rho} = \int_0^{2z_{\text{cut}}} \mathrm{d}\rho \left(\frac{\mathrm{d}\sigma_{\text{g}}^{(1)}}{\mathrm{d}\rho} - \frac{\mathrm{d}\sigma_{\text{g}}^{\text{sing},(1)}}{\mathrm{d}\rho} \right). \tag{26}$$

This equation has been written so that everything on the left is known, and the right side can be evaluated numerically. Using the formulas from Appendix C, the left side evaluates to

$$D_\delta^{(1)} - D_{\delta,g}^{(1)} - \int_{2z_{cut}}^1 d\rho \left(\frac{d\sigma^{sing,(1)}}{d\rho} - \frac{d\sigma_g^{sing,(1)}}{d\rho} \right)$$

$$- \int_1^4 d\rho \frac{d\sigma^{sing,(1)}}{d\rho} = \frac{\alpha_s C_F}{2\pi} \left(\frac{\pi^2}{3} - 2\log^2 2 \right)$$

$$\simeq 2.32896 \frac{\alpha_s C_F}{2\pi}. \tag{27}$$

To evaluate the right side of Eq. (26), we generate about 10^{13} $e^+e^- \to q\bar{q}g$ events in EVENT2 [52]. The events are groomed with mMDT with a range of z_{cut} values from 10^{-5} to 10^{-2}, in powers of $\sqrt{10} \simeq 3.16$. On the groomed events, we then measure the heavy hemisphere mass. For each value of z_{cut}, we calculate the integrand of the integral on the right side of Eq. (26) as a table of values ranging from

$\rho \in [e^{-20}, 2z_{cut}]$, in steps of powers of $e^{0.1}$. To compute the integral, we then fit a smooth interpolating function to the table and numerically integrate. The result of this procedure is shown in Fig. 1. Here, we plot the result of the integral divided by the coupling factor:

$$\text{Integral} = \frac{2\pi}{\alpha_s C_F} \int_0^{2z_{cut}} d\rho \left(\frac{d\sigma_g^{(1)}}{d\rho} - \frac{d\sigma_g^{sing,(1)}}{d\rho} \right). \tag{28}$$

As z_{cut} decreases, the integral values are seen to converge to the exact expected value computed in Eq. (27).

B. Fit at two-loops

Having validated the procedure at one-loop, we move on to using it at two-loops to determine the constant terms of the collinear-soft function, $c_{S_c}^{(2)}$. Unlike at one-loop, all terms in Eq. (25) are generically nonzero, so we have the equality:

$$D_\delta^{(2)} - D_{\delta,g}^{(2)} - \int_{2z_{cut}}^1 d\rho \left(\frac{d\sigma^{sing,(2)}}{d\rho} - \frac{d\sigma_g^{sing,(2)}}{d\rho} \right) - \int_1^4 d\rho \frac{d\sigma^{sing,(2)}}{d\rho}$$

$$= \int_0^{2z_{cut}} d\rho \left(\frac{d\sigma_g^{(2)}}{d\rho} - \frac{d\sigma_g^{sing,(2)}}{d\rho} \right) + \int_{2z_{cut}}^1 d\rho \left(\frac{d\sigma_g^{(2)}}{d\rho} - \frac{d\sigma^{(2)}}{d\rho} \right). \tag{29}$$

The right side of the equality is unknown and must be evaluated numerically. The left side is completely known, except for the two-loop constant of the collinear-soft function, $c_{S_c}^{(2)}$. Using the results of Appendix C and the singular ungroomed heavy hemisphere mass distribution calculated in Refs. [82–84], the left side can be expressed as:

$$D_\delta^{(2)} - D_{\delta,g}^{(2)} - \int_{2z_{cut}}^1 d\rho \left(\frac{d\sigma^{sing,(2)}}{d\rho} - \frac{d\sigma_g^{sing,(2)}}{d\rho} \right) - \int_1^4 d\rho \frac{d\sigma^{sing,(2)}}{d\rho}$$

$$= \left(\frac{\alpha_s}{2\pi} \right)^2 C_F \Bigg[C_F \left(-2 \left(\frac{\pi^2}{3} - 2\log^2 2 \right) \log^2 z_{cut} + \left(-\frac{9}{4} + 6\log^2 2 - \pi^2 - 8\log^3 2 \right. \right.$$

$$+ \frac{8}{3}\pi^2 \log 2 + 8\zeta_3 \right) \log z_{cut} + \frac{9}{4}\log 2 + 2\log^2 2 - \frac{\pi^2}{3} - 6\log^3 2 + 2\pi^2 \log 2 + 6\zeta_3$$

$$+ 2\log^4 2 - \frac{4}{3}\pi^2 \log^2 2 + \frac{61}{360}\pi^4 - 8\zeta_3 \log 2 + \frac{\log 2}{4}\gamma_{S,C_F}^{(1)} - \frac{c_{S,C_F}^{(2)}}{4} - \frac{c_{S_c,C_F}^{(2)}}{2} \Bigg)$$

$$+ C_A \left(-\frac{11}{3} \left(\frac{\pi^2}{3} - 2\log^2 2 \right) \log z_{cut} - \frac{508}{81} - \frac{202}{27}\log 2 - \frac{67}{9}\log^2 2 - \frac{67}{24}\pi^2 - \frac{11}{3}\log^3 2 \right.$$

$$+ \frac{11}{36}\pi^2 \log 2 + \frac{88}{9}\zeta_3 + \frac{\pi^2}{3}\log^2 2 + \frac{11}{30}\pi^4 + 4\zeta(\bar{3},\bar{1}) + \frac{\log 2}{4}\gamma_{S,C_A}^{(1)} - \frac{c_{S,C_A}^{(2)}}{4} - \frac{c_{S_c,C_A}^{(2)}}{2} \Bigg)$$

$$+ T_R n_f \left(\frac{4}{3} \left(\frac{\pi^2}{3} - 2\log^2 2 \right) \log z_{cut} - \frac{34}{81} + \frac{56}{27}\log 2 + \frac{20}{9}\log^2 2 + \frac{19}{18}\pi^2 + \frac{4}{3}\log^3 2 \right.$$

$$- \frac{\pi^2}{9}\log 2 - \frac{32}{9}\zeta_3 + \frac{\log 2}{4}\gamma_{S,n_f}^{(1)} - \frac{c_{S,n_f}^{(2)}}{4} - \frac{c_{S_c,n_f}^{(2)}}{2} \Bigg) \Bigg]. \tag{30}$$
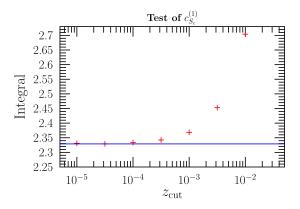
FIG. 1.   Plot of the numerical integral of Eq. (28) as a function of $z_{\rm cut}$ as evaluated in EVENT2. The exact value for $z_{\rm cut} \to 0$ is plotted as the solid line.

In this expression, $\zeta(\bar{3},\bar{1})$ is a multiple zeta value, where

$$\zeta(\bar{3},\bar{1}) \simeq -0.117876. \qquad (31)$$

The coefficients of the two-loop anomalous dimension of the soft function $\gamma_S^{(1)}$ were calculated in Refs. [8,49,50] and are

$$\gamma_S^{(1)} = 2C_F\Big(C_F\gamma_{S,C_F}^{(1)} + C_A\gamma_{S,C_A}^{(1)} + T_Rn_f\gamma_{S,n_f}^{(1)}\Big)$$
$$\simeq 2C_F(17.0055C_F - 20.4487C_A - 10.9322T_Rn_f). \qquad (32)$$

The two-loop constants of the soft function $c_S^{(2)}$ were recently calculated by the SOFTSERVE collaboration [49–51]:

$$c_S^{(2)} = C_F^2 c_{S,C_F}^{(2)} + C_FC_A c_{S,C_A}^{(2)} + C_FT_Rn_f c_{S,n_f}^{(2)}$$
$$\simeq 74.523C_F^2 - 103.70C_FC_A + 29.048C_FT_Rn_f. \qquad (33)$$

The only unknowns in this expression are therefore the two-loop constants of the collinear-soft function. As with the soft function's anomalous dimension and constants, we break up the collinear-soft constant into distinct color channel coefficients:

$$c_{S_c}^{(2)} = C_F^2 c_{S_c,C_F}^{(2)} + C_FC_A c_{S_c,C_A}^{(2)} + C_FT_Rn_f c_{S_c,n_f}^{(2)}. \qquad (34)$$

To evaluate the right side of Eq. (29), we perform essentially the same procedure as at one-loop. We generate about 10^{13} $e^+e^- \to$ hadrons events at $\mathcal{O}(\alpha_s^2)$ in EVENT2 and mMDT groom the final states with values of $z_{\rm cut}$ ranging from 10^{-5} to 10^{-2} in factors of $\sqrt{10}$. On these groomed events, we measure the heavy hemisphere mass ρ, and bin the resulting cross section over the range of $\log\rho \in [-20,0]$, in steps of 0.1. For the second integral on the right side of Eq. (29), we also measure the ungroomed heavy hemisphere mass over the same range.

We then construct a smooth interpolating function for both integrands in Eq. (29) which we then numerically integrate. Because of the finite cutoffs inherent in EVENT2, we cannot practically extend the first integral all the way down to 0. Instead, we integrate over the reduced range of

$$\rho \in \left[\frac{2z_{\rm cut}}{400}, 2z_{\rm cut}\right]. \qquad (35)$$

We have additionally verified that increasing or decreasing the lower bound by a factor of 2 only affects the results within quoted uncertainties. With the integrals thus calculated, we then solve Eq. (29) for each color channel's collinear-soft function two-loop constant coefficient.

The results of this procedure are plotted in Fig. 2. Here, we show the result of the two-loop constant extraction at each value of $z_{\rm cut}$ considered. We observe clear convergence as $z_{\rm cut} \to 0$ for the C_A and T_Rn_f channel coefficients, with the extracted value denoted by the dotted line, and the shaded band is our claimed uncertainty. For the C_F channel, the convergence is less clear, with the extracted value of the two-loop constant appearing to diverge as $z_{\rm cut} \to 0$. In other contexts, it has been noted that the extraction of anomalous dimensions and constants in the C_F channel with EVENT2 is problematic [8,82], unless a significantly larger number of events are generated. Thus, we take as our extracted value of the C_F channel constant the lowest value determined through this procedure. Further, if the approach to the true value were exponential in $\log z_{\rm cut}$, as would be expected, then the difference between the extracted constants at different $z_{\rm cut}$ values that differ by a factor should uniformly decrease. Thus, we take our uncertainty to be twice the difference between the extracted value at $z_{\rm cut} = 0.001$ and $z_{\rm cut} = 0.00316$.[1]

Finally, we obtain for the coefficients in the color decomposition of the two-loop constant of the collinear-soft function as

$$c_{S_c}^{(2)} = C_F^2(22\pm4) + C_FC_A(41\pm1) + C_FT_Rn_f(14.4\pm0.1). \qquad (36)$$

IV. EXTRACTION OF THREE-LOOP ANOMALOUS DIMENSIONS

We now turn to extraction of the three-loop noncusp anomalous dimension of the mMDT global soft function, $\gamma_S^{(2)}$. With the results quoted in the Appendixes and the

---

[1]We have also attempted to use MCCSM to determine the two-loop collinear-soft constant in the C_F channel, given the results in the C_A and n_fT_R channels from EVENT2. However, we were never able to achieve the numerical precision of EVENT2 with comparable computational run times and as deep in the infrared, so could not perform a reliable extraction with MCCSM.
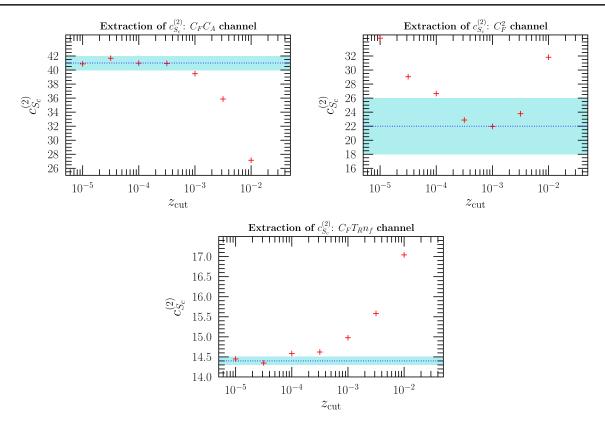
FIG. 2. Plots of the values extracted from EVENT2 of the two-loop constants separated into distinct color channels of the collinear-soft function, as a function of $z_{\rm cut}$. Our claimed extracted value and its uncertainty is illustrated by the dotted line and shaded band, respectively.

extracted value of the two-loop constant of the collinear-soft function, we are able to solve the renormalization group equation through $\mathcal{O}(\alpha_s^3)$ order, and construct the differential cross section by inverse Laplace transformation. For $\rho > 0$, the only unknown piece of the cross section is $\gamma_S^{(2)}$ and so its value can be extracted through comparison of the analytic prediction in the regime where $\rho \ll z_{\rm cut} \ll 1$ to a next-to-next-to-leading fixed-order $e^+e^- \to$ hadrons event generator. For some fixed values of $z_{\rm cut}$, we can then fit the analytic cross section to match the predictions of the NNLO code for the value of the anomalous dimension and extrapolate to $z_{\rm cut} \to 0$ to determine the $\gamma_S^{(2)}$.

The first step of this procedure is to solve the renormalization group equations to $\mathcal{O}(\alpha_s^3)$ to identify the cross section predicted by the factorization theorem. The $\mathcal{O}(\alpha_s^3)$ coefficient of the expansion is given in Eq. (C4) where we substituted explicitly the color factors of QCD ($C_F = 4/3$, $C_A = 3$, $T_R = 1/2$), and set the number of active quarks to $n_f = 5$. We have also just used the central values of the collinear-soft constants, with no accounting for uncertainties. The uncertainties that we will find for $\gamma_S^{(2)}$ will more than account for uncertainties in the collinear-soft function constants, so we can safely ignore them. In this logarithmic form of the cross section, the anomalous dimension $\gamma_S^{(2)}$

appears as a constant offset of the cross section. This suggests a procedure for its numerical extraction.

As the mMDT grooming removes double logarithms in ρ to all orders, the singular term at $\mathcal{O}(\alpha_s^3)$, $D_{C,g}$, is only quadratic in $\log \rho$. Thus, our task of determining $\gamma_S^{(2)}$ is reduced to finding the constant term of a parabola. A parabola exhibits greatest sensitivity to the constant where its slope is smallest; that is, near its extremum, in our case a minimum. Therefore, we would like to fit for $\gamma_S^{(2)}$ near the minimum of the logarithmic parabolic cross section. However, the factorization theorem is only valid for $\rho \ll z_{\rm cut} \ll 1$, so the fit is only sensible if the minimum is compatible with this singular limit. These two competing requirements, fitting near the minimum and existing in the factorization regime, will greatly restrict the range of $z_{\rm cut}$ that we can consider.

Taking the derivative of the function $D_{C,g}$ given in Eq. (C4) to determine the point $\rho_{\rm min}$ at which the cross section is minimized, we can then compare $\rho_{\rm min}$ to $z_{\rm cut}$. We plot the result of this minimization and comparison to $z_{\rm cut}$ in Fig. 3 where the solid curve is the ratio $z_{\rm cut}/\rho_{\rm min}$, as a function of $z_{\rm cut}$. The fixed ratios of 1 and 10 are also shown to guide the eye. We find that only for $z_{\rm cut} \gtrsim 0.01$ is $\rho_{\rm min}$ smaller than $z_{\rm cut}$, and only for $z_{\rm cut} \gtrsim 0.03$ is there an order of magnitude between them. As our factorization theorem

Ratio of $z_{\rm cut}$ to the position of the minimum of $D_C(z_{\rm cut}, \ln(\rho))$

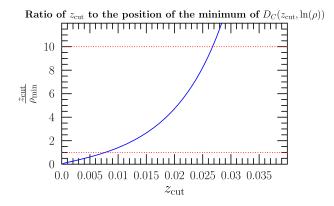FIG. 3.   Plot of the ratio of $z_{\rm cut}$ to the location of the minimum of the factorization theorem cross section at $\mathcal{O}(\alpha_s^3)$ logarithmic in ρ, $\rho_{\rm min}$. Lines corresponding to ratios of 1 and 10 have been added to guide the eye.

requires a parametric separation between ρ and $z_{\rm cut}$, we will only consider values of $z_{\rm cut}$ larger than 0.03 in our fit for the anomalous dimension. Further, demanding that $z_{\rm cut}$ is relatively small, we will restrict to an upper bound on $z_{\rm cut}$ of 0.1. While a limited range in $z_{\rm cut}$, and relatively large values, we will see that it is sufficient for an extraction of the anomalous dimension.

For numerical generation of the cross section for the process $e^+ e^- \to$ hadrons at $\mathcal{O}(\alpha_s^3)$ we use the MCCSM code [32]. MCCSM has previously been used to calculate distributions of standard event shapes [85–87] as well as mMDT and soft drop groomed event shapes [33] in $e^+ e^-$ collisions. To validate the implementation of the mMDT grooming we compared the resulting distributions at LO and NLO accuracy from the independent implementations in MCCSM and EVENT2. We found perfect agreement between the predictions of the two codes.
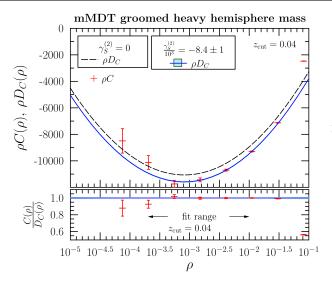
For extraction of the three-loop soft function anomalous dimension, we computed the distribution of the mMDT groomed heavy hemisphere mass for four values of $z_{\rm cut}$ that satisfies the criteria established earlier: $z_{\rm cut} = 0.04$, 0.06, 0.08, 0.1. Any Monte Carlo integrator must contain a lower cut $y_{\rm min}$ on the phase space that any scaled two-particle invariant of the event must exceed. Double-precision arithmetic limits the lowest value of $y_{\rm min}$ to the range $10^{-8 \pm 1}$. As for the extraction of $\gamma_S^{(2)}$ we need stable predictions for $\rho \ll z_{\rm cut}$ we have to investigate carefully the lower limit on ρ where our predictions are still reliable at a given $y_{\rm min}$. We found that we can use $y_{\rm min} = 10^{-8}$ for computing the phase-space integrals of the real-virtual contributions and $y_{\rm min} = 10^{-7}$ for the double-real contributions to obtain realiable distributions for $\rho \geq 5 \times 10^{-5}$ [88]. We used these values of the technical cut to generate 250 batches of events for the double-virtual and real-virtual contributions and 10 thousand batches for the most challenging double-real integrals, with 100 million events in each batch. The events were binned logarithmically, with

10 equal bins in the range $-10 < \log(\rho) < 0$, i.e., we computed the coefficient of the distribution $\rho \frac{d\sigma_g}{d\rho}$ directly. First we performed a flat average of the results of the batches of events. The resulting distributions were fairly stable and smooth except for the first two bins with smallest values of ρ due to occasional misbinning of the squared matrix element and subtraction terms. We filtered the results of the individual batches of events by selecting the ten batches with largest uncertainty in the second bin. Removing the batches from the average one-by-one, we found compatible results for the distributions, but gradually decreasing uncertainty in the first two bins. We produced our final distributions without removing any of the events, but employing a weighted (by the inverse of the uncertainties) average of the individual results of batches. We have verified that the resulting distributions are consistent with the unfiltered and unweighted events, though with significantly smaller uncertainties in each histogram bin. The resulting distributions are provided in Appendix D for each value of $z_{\rm cut}$ that was generated, and we also show it for $z_{\rm cut} = 0.04$ in Fig. 4 (left).

With these NNLO events generated, we can then compare MCCSM to our analytic expression in Eq. (C4). We first plot the two cross sections over one another to demonstrate sensitivity to the soft function anomalous dimension. At left in Fig. 4, we plot the MCCSM cross section for the mMDT groomed heavy hemisphere mass with $z_{\rm cut} = 0.04$, and provide two curves of the corresponding analytic prediction in which the soft function anomalous dimension takes one of two values: $\gamma_S^{(2)} = 0$ and $\gamma_S^{(2)} = -8400$, with the shaded band representing an uncertainty in its value of ± 1000. Clear sensitivity to the value of the anomalous dimension is observed, and the MCCSM results prefer a negative value of $\gamma_S^{(2)}$. This is illustrated more directly in the right panel of Fig. 4. Here, we have set the analytic prediction equal to the MCCSM data and solved for $\gamma_S^{(2)}$ at each value of ρ. Below about $\rho \simeq 0.01$, the value of the anomalous dimension at each data point is consistent within uncertainties of one another, demonstrating independence of ρ, as $\gamma_S^{(2)}$ must.

For each value of $z_{\rm cut}$ that we generated, we determined the anomalous dimension through the following procedure. Taking the plot on the right of Fig. 4 (and the corresponding plots for other values of $z_{\rm cut}$), we fit for the anomalous dimension from data points in the range $\rho \in [0.0002, 0.01]$.[2] This range ensures that the largest value of ρ is still significantly smaller than $z_{\rm cut}$. The contribution of a data point to the anomalous dimension and its uncertainty are weighted by the quoted variance of

---

[2]We have verified that the extracted central value of $\gamma_S^{(2)}$ is consistent within uncertainties using a fit range of $\rho \in [0.000075, 0.004]$, though with larger uncertainties.
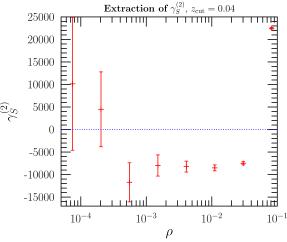
FIG. 4. Left: Plot comparing the distribution of mMDT groomed heavy hemisphere mass at $\mathcal{O}(\alpha_s^3)$ from MCCSM to the analytic prediction with $\gamma_S^{(2)} = 0$ and $\gamma_S^{(2)} = -8400 \pm 1000$. Overall coupling dependence has been stripped off. Right: Plot showing the extracted value of $\gamma_S^{(2)}$ from MCCSM at each ρ point. For both plots, $z_{\text{cut}} = 0.04$.
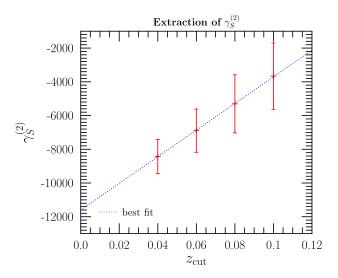

FIG. 5. Plot of $\gamma_S^{(2)}$, with uncertainties, determined at each value of z_{cut}. The best-fit line is dotted and its intercept at $z_{\text{cut}} = 0$ is the quoted extracted value of $\gamma_S^{(2)}$.

that data point. So, the points at larger ρ with smaller uncertainty dominate the fit. The result of this fitting procedure is shown in Fig. 5. As z_{cut} decreases, the value of the extracted anomalous dimension also decreases, and the central values lie almost perfectly on a line. Extrapolating the points in the $z_{\text{cut}} \to 0$ limit produces the desired three-loop anomalous dimension. Performing a weighted fit of the points in Fig. 5 to a line, we find the intercept, that is, $\gamma_S^{(2)}$, to be:

$$\gamma_S^{(2)} = -11600 \pm 2000 \quad (n_f = 5). \tag{37}$$

We expect this claimed uncertainty to be a conservative overestimate due to the extremely linear behavior of the central values.[3]

V. CONCLUSIONS AND OUTLOOK

Using numerical predictions at NLO and NNLO accuracies, we were able to extract two-loop constants and three-loop anomalous dimensions for the mMDT groomed mass factorization theorem in $e^+ e^- \to$ hadrons events. As a result, we are able to resum the groomed mass distribution to NNNLL accuracy, extending the predictions of Ref. [8] to one higher logarithmic order. We present NNNLL + NNLO resummed and matched predictions for the heavy hemisphere mMDT groomed jet mass in a companion paper [53].

The results presented here also enable higher-precision calculations for groomed mass distributions for jets produced in hadron collisions. The cross section for groomed mass ρ of jets at a hadron collider can be expressed as

$$\frac{d\sigma}{d\rho} = \sum_{i \in q,g} \mathcal{N}_i(p_T, y, z_{\text{cut}}, R) J_i(\rho) \otimes S_{c,i}(\rho, z_{\text{cut}}). \tag{38}$$

At a hadron collider, both quark and gluon jets can be produced, so we have to sum over both possibilities. In the factorization limit $\rho \ll z_{\text{cut}} \ll 1$, all radiation that contributes to ρ must be collinear, so the only dependence on the groomed jet mass lives in the universal jet and collinear-soft

---

[3]There are additional uncertainties in MCCSM due to a numerical fit of part of the NNLO matrix element. We find that the uncertainty stated here in the value of the anomalous dimension is representative of uncertainties that come from this fit in MCCSM.

functions, appropriate for quarks or gluons. The normalization factor $\mathcal{N}_i(p_T, y, z_{\text{cut}}, R)$ encodes the relative fraction of jets in the sample that are quark or gluon flavor, and can in principle depend on all other scales in the event, except for ρ: the jet's transverse momentum p_T, rapidity y, the value of z_{cut}, and the jet radius R. The results presented here enable NNNLL resummation of all ρ dependence at a hadron collider.

Extension of the mMDT groomed jet mass factorization theorem to three-loop order is one necessary step for complete perturbative control over the groomed jet mass distribution. However, with multiple scales in a groomed jet, there are multiple regimes in which one may need to resum large logarithms. No NGLs in the groomed mass ρ are present in the regime where $\rho \ll z_{\text{cut}}$, by the collinearity of emissions that remain after grooming. Yet, the NGLs are still there, they have just been pushed into the regime in which $\rho \simeq z_{\text{cut}}$. Typical values of z_{cut} of about 0.1 may produce numerically relevant NGLs that should be resummed to control them in the region of the cross section where grooming begins to dominate. Conversely, our factorization theorem requires that $z_{\text{cut}} \ll 1$ to ensure that only soft emissions are groomed away. However, the typical $z_{\text{cut}} \simeq 0.1$ may not be such a good approximation to $z_{\text{cut}} \ll 1$, and finite z_{cut} contributions should be included. This was studied in the original mMDT paper [3], but a factorization theorem for this regime does not yet exist.

Additionally, our results in this paper are restricted to the mMDT groomer, or soft drop with $\beta = 0$. To extract the two-loop collinear-soft constants with $\beta > 0$, essentially the same procedure could be used as proposed in this paper, with appropriate modifications for the phase space boundaries when $\beta \neq 0$. An analysis for general soft drop angular exponent β is much more challenging at three-loops. Our extraction of the three-loop anomalous dimension for mMDT grooming was simplified because mMDT removes double logarithms of the mass, but for $\beta > 0$, double logarithms generically exist. What was a problem of finding the constant term of a parabola here turns into finding the constant term of a fifth-degree polynomial for general β. A significant amount of data are known for general soft drop grooming at two-loop accuracy [49–51], but complete resummation to NNNLL accuracy will be challenging. Soft drop grooming with nonzero β can provide a handle on quark versus gluon fractions in a jet sample at a hadron collider [30] as well as provide just the right amount of grooming for boosted top quarks [31]. Three-loop data for soft drop grooming is therefore highly desired. We hope that the work presented in this paper inspires further precision predictions for jet physics.

ACKNOWLEDGMENTS

APPENDIX A: RESUMMATION INGREDIENTS

1. β-function and cusp anomalous dimension

The QCD β-function is defined to be

$$\beta(\alpha_s) = \mu \frac{\partial \alpha_s}{\partial \mu} = -2\alpha_s \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}. \quad (A1)$$

For expansion of the cross section to three-loop order, we need the β-function to two-loop order [89–91]. The first two coefficients are

$$\beta_0 = \frac{11}{3} C_A - \frac{4}{3} T_R n_f,$$
$$\beta_1 = \frac{34}{3} C_A^2 - 4 T_R n_f \left(C_F + \frac{5}{3} C_A\right). \quad (A2)$$

Correspondingly, we need the cusp anomalous dimension

$$\Gamma_{\text{cusp}} = \sum_{n=0}^{\infty} \Gamma_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1} \quad (A3)$$

to three-loop order. The first three coefficients of the cusp anomalous dimension are [92–95]:

$$\Gamma_0 = 4,$$
$$\Gamma_1 = 4C_A \left(\frac{67}{9} - \frac{\pi^2}{3}\right) - \frac{80}{9} T_R n_f,$$
$$\Gamma_2 = 4C_A^2 \left(\frac{245}{6} - \frac{134\pi^2}{27} + \frac{11\pi^4}{45} + \frac{22}{3}\zeta_3\right)$$
$$+ 32C_A T_R n_f \left(-\frac{209}{108} + \frac{5\pi^2}{27} - \frac{7}{3}\zeta_3\right)$$
$$+ 4C_F T_R n_f \left(16\zeta_3 - \frac{55}{3}\right) - \frac{64}{27} T_R^2 n_f^2. \quad (A4)$$

2. Noncusp anomalous dimensions and constants

The noncusp anomalous dimensions γ and constants c can be expanded in similar series in α_s as the β-function and cusp anomalous dimensions. We have

$$\gamma = \sum_{n=0}^{\infty} \gamma^{(n)} \left(\frac{\alpha_s}{4\pi}\right)^{n+1}, \qquad c = \sum_{n=1}^{\infty} c^{(n)} \left(\frac{\alpha_s}{4\pi}\right)^n. \quad (A5)$$

To $\mathcal{O}(\alpha_s^3)$ accuracy in terms with support away from 0, we need the noncusp anomalous dimension to α_s^3 order and the constants to α_s^2 order.

a. Hard function

The first three noncusp anomalous dimensions of the hard function (quark form factor) are [76,78,96,97]:

$$\gamma_H^{(0)} = -12C_F,$$

$$\gamma_H^{(1)} = (-6 + 8\pi^2 - 96\zeta_3)C_F^2 + \left(-\frac{1922}{27} - \frac{22}{3}\pi^2 + 104\zeta_3\right)C_F C_A + \left(\frac{520}{27} + \frac{8}{3}\pi^2\right)C_F n_f T_R,$$

$$\gamma_H^{(2)} = 2C_F\left[C_F^2\left(-29 - 6\pi^2 - \frac{16\pi^4}{5} - 136\zeta_3 + \frac{32\pi^2}{3}\zeta_3 + 480\zeta_5\right)\right.$$

$$+ C_F C_A\left(-\frac{151}{2} + \frac{410\pi^2}{9} + \frac{494\pi^4}{135} - \frac{1688}{3}\zeta_3 - \frac{16\pi^2}{3}\zeta_3 - 240\zeta_5\right)$$

$$+ C_A^2\left(-\frac{139345}{1458} - \frac{7163\pi^2}{243} - \frac{83\pi^4}{45} + \frac{7052}{9}\zeta_3 - \frac{88\pi^2}{9}\zeta_3 - 272\zeta_5\right)$$

$$+ C_F T_F n_f\left(\frac{5906}{27} - \frac{52\pi^2}{9} - \frac{56\pi^4}{27} + \frac{1024}{9}\zeta_3\right)$$

$$\left.+ C_A T_F n_f\left(-\frac{34636}{729} + \frac{5188\pi^2}{243} + \frac{44\pi^4}{45} - \frac{3856}{27}\zeta_3\right) + T_F^2 n_f^2\left(\frac{19336}{729} - \frac{80\pi^2}{27} - \frac{64}{27}\zeta_3\right)\right]. \tag{A6}$$

The first two constants of the hard function are [76,77]:

$$c_H^{(1)} = C_F\left(-16 + \frac{7\pi^2}{3}\right),$$

$$c_H^{(2)} = C_F^2\left(\frac{511}{4} - \frac{83}{3}\pi^2 + \frac{67}{30}\pi^4 - 60\zeta_3\right) + C_F C_A\left(-\frac{51157}{324} + \frac{1061}{54}\pi^2 - \frac{8}{45}\pi^4 + \frac{626}{9}\zeta_3\right)$$

$$+ C_F T_R n_f\left(\frac{4085}{81} - \frac{182}{27}\pi^2 + \frac{8}{9}\zeta_3\right). \tag{A7}$$

b. Jet function

The first three noncusp anomalous dimensions of the quark jet function are [75,76,78]:

$$\gamma_J^{(0)} = 6C_F,$$

$$\gamma_J^{(1)} = C_F^2(3 - 4\pi^2 + 48\zeta_3) + C_F C_A\left(\frac{1769}{27} + \frac{22}{9}\pi^2 - 80\zeta_3\right) + C_F T_R n_f\left(-\frac{484}{27} - \frac{8}{9}\pi^2\right),$$

$$\gamma_J^{(2)} = -2C_F\left[C_F^2\left(-\frac{29}{2} - 3\pi^2 - \frac{8\pi^4}{5} - 68\zeta_3 + \frac{16\pi^2}{3}\zeta_3 + 240\zeta_5\right)\right.$$

$$+ C_F C_A\left(-\frac{151}{4} + \frac{205\pi^2}{9} + \frac{247\pi^4}{135} - \frac{844}{3}\zeta_3 - \frac{8\pi^2}{3}\zeta_3 - 120\zeta_5\right)$$

$$+ C_A^2\left(-\frac{412907}{2916} - \frac{419\pi^2}{243} - \frac{19\pi^4}{10} + \frac{5500}{9}\zeta_3 - \frac{88\pi^2}{9}\zeta_3 - 232\zeta_5\right)$$

$$+ C_F T_F n_f\left(\frac{4664}{27} - \frac{32\pi^2}{9} - \frac{164\pi^4}{135} + \frac{208}{9}\zeta_3\right)$$

$$\left.+ C_A T_F n_f\left(-\frac{5476}{729} + \frac{1180\pi^2}{243} + \frac{46\pi^4}{45} - \frac{2656}{27}\zeta_3\right) + T_F^2 n_f^2\left(\frac{13828}{729} - \frac{80\pi^2}{81} - \frac{256}{27}\zeta_3\right)\right]. \tag{A8}$$

The first two constants of the jet function are [45]:

$$c_J^{(1)} = C_F\left(7 - \frac{2\pi^2}{3}\right),$$

$$c_J^{(2)} = C_F^2\left(\frac{205}{8} - \frac{97\pi^2}{12} + \frac{61\pi^4}{90} - 6\zeta_3\right)$$

$$+ C_F C_A\left(\frac{53129}{648} - \frac{155\pi^2}{36} - \frac{37\pi^4}{180} - 18\zeta_3\right)$$

$$+ C_F T_R n_f\left(-\frac{4057}{162} + \frac{13\pi^2}{9}\right). \tag{A9}$$

c. mMDT global soft function

The first two coefficients of the noncusp anomalous dimensions of the mMDT global soft function were calculated previously [8,50] and the third coefficient is new in this work:

$$\gamma_S^{(0)} = 0,$$

$$\gamma_S^{(1)} = 2C_F[17.0055 C_F - 20.4487 C_A - 10.9322 T_R n_f],$$

$$\gamma_S^{(2)} = -11600 \pm 2000 \quad (n_f = 5). \tag{A10}$$

In $\gamma_S^{(1)}$, the coefficient of the C_F^2 term is known analytically:

$$\frac{16\pi}{3}\text{Cl}_2\left(\frac{\pi}{3}\right) \simeq 17.0055, \tag{A11}$$

where $\text{Cl}_2(x)$ is the Clausen function:

$$\text{Cl}_2\left(\frac{\pi}{3}\right) = 1.01494160\ldots. \tag{A12}$$

The $C_F C_A$ and $C_F T_R n_f$ coefficients in $\gamma_S^{(1)}$ are not known analytically, so we present their decimal expansion to 6 significant figures. The value of $\gamma_S^{(2)}$ is evaluated with five active quarks, $n_f = 5$.

The first constant of the mMDT soft function was calculated in [8] and the second constant was calculated by the SOFTSERVE collaboration [49–51]:

$$c_S^{(1)} = -\pi^2 C_F,$$

$$c_S^{(2)} = 74.523 C_F^2 - 103.70 C_F C_A + 29.048 C_F T_R n_f. \tag{A13}$$

Uncertainties in the values of these numerical expressions occur in the last digit, which we do not quote because they are vastly smaller than other relevant uncertainties.

d. mMDT collinear-soft function

By renormalization group invariance, the anomalous dimension of the collinear-soft function can be found to all-orders through its relationship with the other anomalous dimensions in the factorization theorem:

$$\gamma_{S_c} = -\frac{\gamma_H + \gamma_S}{2} - \gamma_J. \tag{A14}$$

As such, we will not present the fixed-order expansion of the collinear-soft function's anomalous dimension here. The first coefficient of the collinear-soft function's constant terms was calculated in Ref. [8] and the second coefficient is new to this work:

$$c_{S_c}^{(1)} = 0,$$

$$c_{S_c}^{(2)} = C_F^2(22 \pm 4) + C_F C_A(41 \pm 1)$$

$$+ C_F T_R n_f(14.4 \pm 0.1). \tag{A15}$$

APPENDIX B: FIXED-ORDER EXPANSION

The hemisphere mMDT groomed jet mass factorization theorem in $e^+e^- \to$ hadrons collisions is

$$\frac{1}{\sigma_0}\frac{d^2\sigma}{d\tau_L d\tau_R} = H(Q^2)S(z_{\text{cut}})[J(\tau_L) \otimes S_c(\tau_L, z_{\text{cut}})]$$

$$\times [J(\tau_R) \otimes S_c(\tau_R, z_{\text{cut}})], \tag{B1}$$

where $H(Q^2)$ is the hard function for quark–antiquark production in e^+e^- collisions, $S(z_{\text{cut}})$ is the global soft function for mMDT grooming, $J(\tau_i)$ is the quark jet function for hemisphere mass τ_i, and $S_c(\tau_i, z_{\text{cut}})$ is the collinear-soft function for hemisphere mass τ_i with mMDT grooming. The symbol $\otimes$ denotes convolution over the hemisphere mass τ_i. The convolution can be transformed into a simple product by Laplace transforming the jet and collinear-soft functions appropriately. The Laplace transform of a function F in the factorization theorem is

$$\tilde{F}(\nu) = \int_0^\infty d\tau e^{-\tau\nu} F(\tau), \tag{B2}$$

and the Laplace-transformed factorization theorem is

$$\frac{\sigma(\nu_L, \nu_R)}{\sigma_0} = H(Q^2)S(z_{\text{cut}})\tilde{J}(\nu_L)\tilde{S}_c(\nu_L, z_{\text{cut}})\tilde{J}(\nu_R)\tilde{S}_c(\nu_R, z_{\text{cut}}). \tag{B3}$$

All functions in this factorization theorem satisfy a simple renormalization group equation:

$$\mu\frac{\partial\tilde{F}}{\partial\mu}=\left(d_F\Gamma_{\text{cusp}}\log\frac{\mu^2}{\mu_F^2}+\gamma_F\right)\tilde{F}, \quad \tilde{F}=H,S,\tilde{J},\tilde{S}_c \tag{B4}$$

where μ is the renormalization scale. The expression in parentheses is $\tilde{F}$'s anomalous dimension where Γ_{cusp} is the cusp anomalous dimension, d_F is a coefficient particular to function F, μ_F is the canonical scale of F, and γ_F is the noncusp anomalous dimension of $\tilde{F}$.

The anomalous dimension equation can be solved iteratively in powers of α_s. Through $\mathcal{O}(\alpha_s^3)$, the solution is

$$\begin{aligned}
\tilde{F} = 1 &+ \frac{\alpha_s}{4\pi}\left[\frac{d_F\Gamma_0}{4}\log^2\frac{\mu^2}{\mu_F^2}+\frac{\gamma_F^{(0)}}{2}\log\frac{\mu^2}{\mu_F^2}+c_F^{(1)}\right] + \left(\frac{\alpha_s}{4\pi}\right)^2\left[\frac{d_F^2\Gamma_0^2}{32}\log^4\frac{\mu^2}{\mu_F^2}+\frac{d_F\Gamma_0}{4}\left(\frac{\beta_0}{3}+\frac{\gamma_F^{(0)}}{2}\right)\log^3\frac{\mu^2}{\mu_F^2}\right.\\
&+\left(\frac{\beta_0\gamma_F^{(0)}}{4}+\frac{(\gamma_F^{(0)})^2}{8}+\frac{d_F\Gamma_0 c_F^{(1)}}{4}+\frac{d_F\Gamma_1}{4}\right)\log^2\frac{\mu^2}{\mu_F^2}+\left(\beta_0 c_F^{(1)}+\frac{\gamma_F^{(0)}c_F^{(1)}}{2}+\frac{\gamma_F^{(1)}}{2}\right)\log\frac{\mu^2}{\mu_F^2}+c_F^{(2)}\right]\\
&+\left(\frac{\alpha_s}{4\pi}\right)^3\left[\frac{d_F^3\Gamma_0^3}{384}\log^6\frac{\mu^2}{\mu_F^2}+\frac{d_F^2\Gamma_0^2}{16}\left(\frac{\beta_0}{3}+\frac{\gamma_F^{(0)}}{4}\right)\log^5\frac{\mu^2}{\mu_F^2}+d_F\Gamma_0\left(\frac{\beta_0^2}{24}+\frac{5}{48}\beta_0\gamma_F^{(0)}+\frac{(\gamma_F^{(0)})^2}{32}+\frac{d_F\Gamma_0 c_F^{(1)}}{32}+\frac{d_F\Gamma_1}{16}\right)\log^4\frac{\mu^2}{\mu_F^2}\right.\\
&+\left(\frac{\beta_0^2\gamma_F^{(0)}}{6}+\frac{\beta_0(\gamma_F^{(0)})^2}{8}+\frac{(\gamma_F^{(0)})^3}{48}+\frac{d_F\Gamma_0\beta_0 c_F^{(1)}}{3}+\frac{d_F\Gamma_0\beta_1}{12}+\frac{d_F\Gamma_0\gamma_F^{(0)}c_F^{(1)}}{8}+\frac{d_F\Gamma_0\gamma_F^{(1)}}{8}\right.\\
&\left.+\frac{d_F\Gamma_1\beta_0}{6}+\frac{d_F\Gamma_1\gamma_F^{(0)}}{8}\right)\log^3\frac{\mu^2}{\mu_F^2}+\left(\beta_0^2 c_F^{(1)}+\frac{3}{4}\beta_0\gamma_F^{(0)}c_F^{(1)}+\frac{\beta_1\gamma_F^{(0)}}{4}+\frac{(\gamma_F^{(0)})^2 c_F^{(1)}}{8}\right.\\
&\left.+\frac{d_F\Gamma_0 c_F^{(2)}}{4}+\frac{\beta_0\gamma_F^{(1)}}{2}+\frac{\gamma_F^{(0)}\gamma_F^{(1)}}{4}+\frac{d_F\Gamma_1 c_F^{(1)}}{4}+\frac{d_F\Gamma_2}{4}\right)\log^2\frac{\mu^2}{\mu_F^2}\\
&\left.+\left(2\beta_0 c_F^{(2)}+\beta_1 c_F^{(1)}+\frac{\gamma_F^{(0)}c_F^{(2)}}{2}+\frac{\gamma_F^{(1)}c_F^{(1)}}{2}+\frac{\gamma_F^{(2)}}{2}\right)\log\frac{\mu^2}{\mu_F^2}+c_F^{(3)}\right]+\mathcal{O}(\alpha_s^4).
\end{aligned} \tag{B5}$$

For the functions in the factorization theorem, the d_F coefficients are

$$\begin{aligned}
d_H &= -2C_F, & d_J &= 2C_F,\\
d_S &= 2C_F, & d_{S_c} &= -2C_F.
\end{aligned} \tag{B6}$$

The canonical scales μ_F^2 of the Laplace-space functions are

$$\begin{aligned}
\mu_H^2 &= Q^2, & \mu_J^2 &= \frac{Q^2}{4\nu},\\
\mu_S^2 &= z_{\text{cut}}^2 Q^2, & \mu_{S_c}^2 &= \frac{z_{\text{cut}}Q^2}{4\nu}.
\end{aligned} \tag{B7}$$

Then, to find the cross section in real space through $\mathcal{O}(\alpha_s^3)$, we multiply the Laplace-space functions together, expand in α_s, and then inverse Laplace transform. Because mMDT grooming removes double logarithms, the inverse Laplace transformation is relatively easy, and one only needs three transformations:

$$\begin{aligned}
\mathcal{L}^{-1}(\log\nu) &= -\left(\frac{1}{\tau}\right)_+,\\
\mathcal{L}^{-1}(\log^2\nu) &= 2\left(\frac{\log\tau}{\tau}\right)_+ - \frac{\pi^2}{6}\delta(\tau),\\
\mathcal{L}^{-1}(\log^3\nu) &= -3\left(\frac{\log^2\tau}{\tau}\right)_+ + \frac{\pi^2}{2}\left(\frac{1}{\tau}\right)_+ - 2\zeta_3\delta(\tau).
\end{aligned} \tag{B8}$$

The $+$-functions are defined to integrate to 0 on $\tau\in[0,1]$.

APPENDIX C: mMDT GROOMED HEAVY HEMISPHERE MASS RESULTS

Through $\mathcal{O}(\alpha_s^3)$, the singular cross section for the heavy hemisphere mMDT groomed mass can be written as

$$\frac{d\sigma_{\mathrm{g,LP}}}{d\rho} = \delta(\rho)D_{\delta,\mathrm{g}} + \frac{\alpha_s}{2\pi}(D_{A,\mathrm{g}}(\rho))_+ + \left(\frac{\alpha_s}{2\pi}\right)^2 (D_{B,\mathrm{g}}(\rho))_+ + \left(\frac{\alpha_s}{2\pi}\right)^3 (D_{C,\mathrm{g}}(\rho))_+. \tag{C1}$$

The coefficient functions are

$$
\begin{aligned}
D_{\delta,\mathrm{g}} &= 1 + \frac{\alpha_s}{4\pi}\left[C_F(-2 + 12\log 2 + 16\log 2\log z_{\mathrm{cut}} + 4\log^2 z_{\mathrm{cut}})\right] \\
&+ \left(\frac{\alpha_s}{4\pi}\right)^2\Bigg[C_F^2\bigg(4 + \frac{4}{3}\pi^2 - \frac{7}{90}\pi^4 - 72\zeta_3 - 18\log 2 + 72\log^2 2 + 96\zeta_3\log 2 + c^{(2)}_{S,C_F} + 2c^{(2)}_{S_c,C_F} - 4\pi^2\log(4z_{\mathrm{cut}}) \\
&- \gamma^{(1)}_{S,C_F}\log(4z_{\mathrm{cut}}) - 8\log z_{\mathrm{cut}}\log(16z_{\mathrm{cut}}) - \frac{8}{3}\pi^2\log^2 z_{\mathrm{cut}} + 192\log^2 2\log z_{\mathrm{cut}} + 48\log 2\log^2 z_{\mathrm{cut}} + 128\log^2 2\log^2 z_{\mathrm{cut}} \\
&+ 64\log 2\log^3 z_{\mathrm{cut}} + 8\log^4 z_{\mathrm{cut}}\bigg) + C_F C_A\bigg(\frac{493}{81} + \frac{497}{54}\pi^2 - \frac{53}{90}\pi^4 + \frac{4694}{27}\log 2 - \frac{22}{9}\pi^2\log 2 + 44\log^2 2 + \frac{302}{9}\zeta_3 \\
&- 104\zeta_3\log 2 + c^{(2)}_{S,C_A} + 2c^{(2)}_{S_c,C_A} - \gamma^{(1)}_{S,C_A}\log(4z_{\mathrm{cut}}) + \frac{808}{27}\log z_{\mathrm{cut}} + \frac{1072}{9}\log 2\log z_{\mathrm{cut}} + \frac{11}{3}\pi^2\log z_{\mathrm{cut}} \\
&- 28\zeta_3\log z_{\mathrm{cut}} - \frac{4}{3}\pi^2\log z_{\mathrm{cut}}\log(16z_{\mathrm{cut}}) + \frac{268}{9}\log^2 z_{\mathrm{cut}} + \frac{176}{3}\log^2 2\log z_{\mathrm{cut}} - \frac{44}{3}\log^2 z_{\mathrm{cut}}\log(4z_{\mathrm{cut}})\bigg) \\
&+ C_F n_f T_R\bigg(\frac{28}{81} - \frac{86}{27}\pi^2 - \frac{1528}{27}\log 2 + \frac{8}{9}\pi^2\log 2 - 16\log^2 2 + \frac{8}{9}\zeta_3 + c^{(2)}_{S,n_f} + 2c^{(2)}_{S_c,n_f} - \gamma^{(1)}_{S,n_f}\log(4z_{\mathrm{cut}}) \\
&- \frac{224}{27}\log z_{\mathrm{cut}} - \frac{4}{3}\pi^2\log z_{\mathrm{cut}} - \frac{64}{3}\log^2 2\log z_{\mathrm{cut}} - \frac{80}{9}\log z_{\mathrm{cut}}\log(16z_{\mathrm{cut}}) + \frac{16}{3}\log^2 z_{\mathrm{cut}}\log(4z_{\mathrm{cut}})\bigg)\Bigg]. \tag{C2}
\end{aligned}
$$

In this expression, we have left the two-loop soft and collinear-soft function constants implicit, as well as the two-loop noncusp anomalous dimension of the soft function. In the subscripts of these terms, we have identified the corresponding color channel, which is just the coefficient of that term in the Eqs. (A10), (A13), and (A15).

The terms with support away from $\rho = 0$ are

$$
\begin{aligned}
D_{A,\mathrm{g}}(\rho) &= \frac{1}{\rho}[C_F(-3 - 4\log z_{\mathrm{cut}})], \\
D_{B,\mathrm{g}}(\rho) &= \frac{1}{\rho}\Bigg[C_F^2\bigg((3 + 4\log z_{\mathrm{cut}})^2\log\rho - 18\log 2 + 4\log z_{\mathrm{cut}} - 48\log 2\log z_{\mathrm{cut}} \\
&- 6\log^2 z_{\mathrm{cut}} - 32\log 2\log^2 z_{\mathrm{cut}} - 8\log^3 z_{\mathrm{cut}} + \frac{9}{4} + \pi^2 - 12\zeta_3 + \frac{\gamma^{(1)}_{S,C_F}}{4}\bigg) \\
&+ C_F C_A\bigg(\frac{11}{6}(3 + 4\log z_{\mathrm{cut}})\log\rho - 11\log 2 + \left(-\frac{134}{9} + \frac{2}{3}\pi^2\right)\log z_{\mathrm{cut}} \\
&+ \frac{11}{3}\log\frac{z_{\mathrm{cut}}}{16}\log z_{\mathrm{cut}} - \frac{2347}{108} + \frac{11}{36}\pi^2 + 13\zeta_3 + \frac{\gamma^{(1)}_{S,C_A}}{4}\bigg) \\
&+ C_F n_f T_R\bigg(-\frac{2}{3}(3 + 4\log z_{\mathrm{cut}})\log\rho + 4\log 2 + \frac{40}{9}\log z_{\mathrm{cut}} + \frac{16}{3}\log 2\log z_{\mathrm{cut}} \\
&- \frac{4}{3}\log^2 z_{\mathrm{cut}} + \frac{191}{27} - \frac{\pi^2}{9} + \frac{\gamma^{(1)}_{S,n_f}}{4}\bigg)\Bigg]. \tag{C3}
\end{aligned}
$$

The $D_{C,\mathrm{g}}$ coefficient is extremely unwieldy, so we only quote here the approximate numerical function:

TABLE II. Table of values with uncertainties of the α_s^3 factor C in the form of the differential cross section of Eq. (D1) for the mMDT groomed heavy hemisphere mass and the quoted value of z_{cut}. These values correspond to five active quarks, $n_f = 5$.

$\log \rho$	$z_{\text{cut}} = 0.04$	$z_{\text{cut}} = 0.06$	$z_{\text{cut}} = 0.08$	$z_{\text{cut}} = 0.1$
-9.5	-8495 ± 922	-9361 ± 831	-7872 ± 804	-6543 ± 762
-8.5	-10127 ± 519	-9529 ± 489	-7835 ± 456	-5973 ± 444
-7.5	-11742 ± 271	-10185 ± 253	-8245 ± 235	-6553 ± 218
-6.5	-11435 ± 147	-9183 ± 134	-7256 ± 120	-5653 ± 112
-5.5	-10701 ± 76	-8040 ± 70	-6089 ± 60	-4560 ± 56
-4.5	-9293 ± 41	-6615 ± 35	-4817 ± 32	-3506 ± 29
-3.5	-7127 ± 21	-4892 ± 18	-3429 ± 16	-2390 ± 15
-2.5	-2475 ± 12	-2247 ± 10	-1642 ± 9	-1175 ± 8
-1.5	2803 ± 5	2334 ± 4	1781 ± 4	1442 ± 4
-0.5	1608.0 ± 1.2	1593.8 ± 1.1	1574.6 ± 1.1	1548.7 ± 1.1

$$
\begin{aligned}
D_{C,\text{g}} \simeq \frac{1}{\rho} \Bigg[& (-75.85\log^3 z_{\text{cut}} - 334.22\log^2 z_{\text{cut}} - 451.70\log z_{\text{cut}} - 182.78)\log^2 \rho \\
& + (75.85\log^4 z_{\text{cut}} + 269.56\log^3 z_{\text{cut}} + 1008.64\log^2 z_{\text{cut}} + 1762.95\log z_{\text{cut}} + 877.52)\log \rho \\
& - 18.96\log^5 z_{\text{cut}} - 37.59\log^4 z_{\text{cut}} - 230.06\log^3 z_{\text{cut}} - 724.49\log^2 z_{\text{cut}} - 1641.62\log z_{\text{cut}} - 1944.97 + \frac{\gamma_S^{(2)}}{16} \Bigg]. \quad \text{(C4)}
\end{aligned}
$$

Here we substituted explicitly the color factors of QCD ($C_F = 4/3$, $C_A = 3$, $T_R = 1/2$), and set the number of active quarks to $n_f = 5$. We have also just used the central values of the collinear-soft constants, with no accounting for their uncertainties.

APPENDIX D: MCCSM DATA AT $\mathcal{O}(\alpha_s^3)$

Generating the data from MCCSM for the extraction of the three-loop noncusp anomalous dimension of the soft function $\gamma_S^{(2)}$ took about a century of CPU time. For faster evaluation later, we provide the mMDT groomed heavy hemisphere distributions here, with $z_{\text{cut}} = 0.04, 0.06, 0.08,$ 0.1. For the differential cross section of the mMDT groomed heavy hemisphere mass, MCCSM calculates the ρ-dependent factors A, B, and C at leading, next-to-leading, and next-to-next-to-leading fixed order, respectively. These can then be combined to evaluate the cross section through $\mathcal{O}(\alpha_s^3)$ as:

$$
\frac{\rho}{\sigma_0} \frac{d\sigma_{\text{FO}}}{d\rho} = \frac{\alpha_s}{2\pi} A + \left(\frac{\alpha_s}{2\pi}\right)^2 \left[B + A\beta_0 \log\frac{\mu}{Q} \right] + \left(\frac{\alpha_s}{2\pi}\right)^3 \left[C + 2B\beta_0 \log\frac{\mu}{Q} + A\left(\frac{\beta_1}{2}\log\frac{\mu}{Q} + \beta_0^2\log^2\frac{\mu}{Q}\right) \right]. \quad \text{(D1)}
$$

Here, $\alpha_s = \alpha_s(\mu)$, the strong coupling evaluated at the renormalization scale μ, β_0 and β_1 are the QCD β-function coefficients, and Q is the center-of-mass collision energy. When $\mu = Q$, the terms proportional to the β-function vanish, and the expression simplifies. To extract $\gamma_S^{(2)}$, we need the distribution of C as a function of ρ. The value of C for $\log \rho \in [-9.5, -0.5]$ in steps of 1 and all z_{cut} values we consider are listed in Table II. The entries in this table were determined using the filtering and weighting procedure described in Sec. IV.

[1] A. J. Larkoski, I. Moult, and B. Nachman, Phys. Rep. **841**, 1 (2020).

[2] R. Kogler et al., Rev. Mod. Phys. **91**, 045003 (2019).

[3] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, J. High Energy Phys. 09 (2013) 029.

[4] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, Eur. Phys. J. C **73**, 2623 (2013).

[5] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, J. High Energy Phys. 05 (2014) 146.

[6] M. Dasgupta and G. P. Salam, Phys. Lett. B **512**, 323 (2001).

[7] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, arXiv:1603.06375.

[8] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, J. High Energy Phys. 07 (2016) 064.

[9] S. Marzani, L. Schunk, and G. Soyez, J. High Energy Phys. 07 (2017) 132.

[10] S. Marzani, L. Schunk, and G. Soyez, Eur. Phys. J. C **78**, 96 (2018).

[11] Y. Makris, D. Neill, and V. Vaidya, J. High Energy Phys. 07 (2018) 167.

[12] Z.-B. Kang, K. Lee, X. Liu, and F. Ringer, J. High Energy Phys. 10 (2018) 137.

[13] J. Chay and C. Kim, J. Korean Phys. Soc. **74**, 439 (2019).

[14] J. Baron, S. Marzani, and V. Theeuwes, J. High Energy Phys. 08 (2018) 105; 05 (2019) 056(E).

[15] Y. Makris and V. Vaidya, J. High Energy Phys. 10 (2018) 019.

[16] Z.-B. Kang, K. Lee, X. Liu, and F. Ringer, Phys. Lett. B **793**, 41 (2019).

[17] C. Lee, P. Shrivastava, and V. Vaidya, J. High Energy Phys. 09 (2019) 045.

[18] S. Marzani, D. Reichelt, S. Schumann, G. Soyez, and V. Theeuwes, J. High Energy Phys. 11 (2019) 179.

[19] D. Gutierrez-Reyes, Y. Makris, V. Vaidya, I. Scimemi, and L. Zoppi, J. High Energy Phys. 08 (2019) 161.

[20] P. Cal, D. Neill, F. Ringer, and W. J. Waalewijn, J. High Energy Phys. 04 (2020) 211.

[21] Z.-B. Kang, K. Lee, X. Liu, D. Neill, and F. Ringer, J. High Energy Phys. 02 (2020) 054.

[22] M. Aaboud et al. (ATLAS Collaboration), Phys. Rev. Lett. **121**, 092001 (2018).

[23] A. M. Sirunyan et al. (CMS Collaboration), J. High Energy Phys. 11 (2018) 113.

[24] G. Aad et al. (ATLAS Collaboration), Phys. Rev. D **101**, 052007 (2020).

[25] G. P. Salam, L. Schunk, and G. Soyez, J. High Energy Phys. 03 (2017) 022.

[26] A. J. Larkoski, I. Moult, and D. Neill, arXiv:1708.06760.

[27] A. J. Larkoski, I. Moult, and D. Neill, J. High Energy Phys. 02 (2018) 144.

[28] D. Napoletano and G. Soyez, J. High Energy Phys. 12 (2018) 031.

[29] Y.-T. Chien and I. W. Stewart, J. High Energy Phys. 06 (2020) 064.

[30] J. Bendavid et al., in Proceedings of the Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report (2018) [arXiv:1803.07977].

[31] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, Phys. Rev. D **100**, 074021 (2019).

[32] A. Kardos, G. Somogyi, and Z. Trócsányi, Proc. Sci. LL2016 (**2016**) 021.

[33] A. Kardos, G. Somogyi, and Z. Trócsányi, Phys. Lett. B **786**, 313 (2018).

[34] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, J. High Energy Phys. 12 (2019) 002.

[35] S. Badger, C. Brnnum-Hansen, H. B. Hartanto, and T. Peraro, Phys. Rev. Lett. **120**, 092001 (2018).

[36] S. Abreu, F. Febres Cordero, H. Ita, B. Page, and M. Zeng, Phys. Rev. D **97**, 116014 (2018).

[37] H. A. Chawdhry, M. A. Lim, and A. Mitov, Phys. Rev. D **99**, 076011 (2019).

[38] S. Abreu, F. Febres Cordero, H. Ita, B. Page, and V. Sotnikov, J. High Energy Phys. 11 (2018) 116.

[39] S. Badger, C. Brnnum-Hansen, H. B. Hartanto, and T. Peraro, J. High Energy Phys. 01 (2019) 186.

[40] S. Abreu, J. Dormans, F. Febres Cordero, H. Ita, and B. Page, Phys. Rev. Lett. **122**, 082002 (2019).

[41] D. Chicherin, T. Gehrmann, J. M. Henn, P. Wasser, Y. Zhang, and S. Zoia, Phys. Rev. Lett. **123**, 041603 (2019).

[42] S. Abreu, J. Dormans, F. Febres Cordero, H. Ita, B. Page, and V. Sotnikov, J. High Energy Phys. 05 (2019) 084.

[43] S. Badger, D. Chicherin, T. Gehrmann, G. Heinrich, J. M. Henn, T. Peraro, P. Wasser, Y. Zhang, and S. Zoia, Phys. Rev. Lett. **123**, 071601 (2019).

[44] H. B. Hartanto, S. Badger, C. Brnnum-Hansen, and T. Peraro, J. High Energy Phys. 09 (2019) 119.

[45] T. Becher and M. D. Schwartz, J. High Energy Phys. 07 (2008) 034.

[46] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, Phys. Rev. D **83**, 074021 (2011).

[47] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, Phys. Rev. D **91**, 094017 (2015).

[48] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, Phys. Rev. D **91**, 094018 (2015).

[49] G. Bell, R. Rahn, and J. Talbert, J. High Energy Phys. 07 (2019) 101.

[50] G. Bell, R. Rahn, and J. Talbert, Nucl. Phys. **B936**, 520 (2018).

[51] G. Bell, R. Rahn, and J. Talbert, arXiv:2004.08396.

[52] S. Catani and M. H. Seymour, Nucl. Phys. **B485**, 291 (1997); **510**, 503(E) (1998).

[53] A. Kardos, A. J. Larkoski, and Z. Trócsányi, arXiv:2002.00942.

[54] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).

[55] E. Farhi, Phys. Rev. Lett. **39**, 1587 (1977).

[56] H. Georgi and M. Machacek, Phys. Rev. Lett. **39**, 1237 (1977).

[57] A. J. Larkoski, D. Neill, and J. Thaler, J. High Energy Phys. 04 (2014) 017.

[58] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (1997) 001.

[59] M. Wobisch and T. Wengler, Hadronization corrections to jet cross-sections in deep inelastic scattering, in Monte Carlo Generators for HERA Physics. Proceedings, Workshop, Hamburg, Germany, 1998–1999 (DESY, Hamburg, Germany, 1998), pp. 270–279.

[60] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, Nucl. Phys. **B407**, 3 (1993).

[61] J. A. Gracey, Phys. Lett. B **322**, 141 (1994).

[62] M. Beneke and V. M. Braun, Nucl. Phys. **B454**, 253 (1995).

[63] J. Davies, A. Vogt, B. Ruijl, T. Ueda, and J. A. M. Vermaseren, Nucl. Phys. **B915**, 335 (2017).

[64] J. Henn, A. V. Smirnov, V. A. Smirnov, M. Steinhauser, and R. N. Lee, J. High Energy Phys. 03 (2017) 139.

[65] J. M. Henn, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, J. High Energy Phys. 05 (2016) 066.

[66] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, J. High Energy Phys. 10 (2017) 041.

[67] A. Grozin, J. High Energy Phys. 06 (2018) 073; 01 (2019) 134(A).

[68] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, Phys. Lett. B **782**, 627 (2018).

[69] R. Brser, A. Grozin, J. M. Henn, and M. Stahlhofen, J. High Energy Phys. 05 (2019) 186.

[70] J. M. Henn, T. Peraro, M. Stahlhofen, and P. Wasser, Phys. Rev. Lett. **122**, 201602 (2019).

[71] A. von Manteuffel and R. M. Schabinger, Phys. Rev. D **99**, 094014 (2019).

[72] R. N. Lee, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, J. High Energy Phys. 02 (2019) 172.

[73] J. M. Henn, G. P. Korchemsky, and B. Mistlberger, J. High Energy Phys. 04 (2020) 018.

[74] A. von Manteuffel, E. Panzer, and R. M. Schabinger, Phys. Rev. Lett. **124**, 162001 (2020).

[75] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, Phys. Lett. B **400**, 379 (1997).

[76] S. Moch, J. A. M. Vermaseren, and A. Vogt, J. High Energy Phys. 08 (2005) 049.

[77] T. Gehrmann, T. Huber, and D. Maitre, Phys. Lett. B **622**, 295 (2005).

[78] T. Becher, M. Neubert, and B. D. Pecjak, J. High Energy Phys. 01 (2007) 076.

[79] P. A. Baikov, K. G. Chetyrkin, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Phys. Rev. Lett. **102**, 212002 (2009).

[80] R. N. Lee, A. V. Smirnov, and V. A. Smirnov, J. High Energy Phys. 04 (2010) 020.

[81] T. Gehrmann, E. W. N. Glover, T. Huber, N. Ikizlerli, and C. Studerus, J. High Energy Phys. 06 (2010) 094.

[82] Y.-T. Chien and M. D. Schwartz, J. High Energy Phys. 08 (2010) 058.

[83] R. Kelley, M. D. Schwartz, R. M. Schabinger, and H. X. Zhu, Phys. Rev. D **84**, 045022 (2011).

[84] P. F. Monni, T. Gehrmann, and G. Luisoni, J. High Energy Phys. 08 (2011) 010.

[85] V. Del Duca, C. Duhr, A. Kardos, G. Somogyi, and Z. Trócsányi, Phys. Rev. Lett. **117**, 152004 (2016).

[86] V. Del Duca, C. Duhr, A. Kardos, G. Somogyi, Z. Szőr, Z. Trócsányi, and Z. Tulipánt, Phys. Rev. D **94**, 074019 (2016).

[87] Z. Tulipánt, A. Kardos, and G. Somogyi, Eur. Phys. J. C **77**, 749 (2017).

[88] A. Kardos, A. Larkoski, and Z. Trócsányi, Acta Phys. Pol. B **50**, 1891 (2019).

[89] W. E. Caswell, Phys. Rev. Lett. **33**, 244 (1974).

[90] D. R. T. Jones, Nucl. Phys. **B75**, 531 (1974).

[91] E. Egorian and O. V. Tarasov, Teor. Mat. Fiz. **41**, 26 (1979) [Theor. Math. Phys. **41**, 863 (1979)].

[92] G. P. Korchemsky and A. V. Radyushkin, Nucl. Phys. **B283**, 342 (1987).

[93] A. Vogt, Phys. Lett. B **497**, 228 (2001).

[94] C. F. Berger, Phys. Rev. D **66**, 116002 (2002).

[95] S. Moch, J. A. M. Vermaseren, and A. Vogt, Phys. Lett. B **625**, 245 (2005).

[96] W. L. van Neerven, Nucl. Phys. **B268**, 453 (1986).

[97] T. Matsuura, S. C. van der Marck, and W. L. van Neerven, Nucl. Phys. **B319**, 570 (1989).