

Boosted top quarks in the peak region with N³LL resummationBrad Bachu^{1,2}, André H. Hoang^{3,4}, Vicent Mateu^{5,6}, Aditya Pathak^{1,3,7} and Iain W. Stewart¹¹*Center for Theoretical Physics, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA*²*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*³*University of Vienna, Faculty of Physics, Boltzmannngasse 5, A-1090 Wien, Austria*⁴*Erwin Schrödinger International Institute for Mathematical Physics, University of Vienna,
Boltzmannngasse 9, A-1090 Wien, Austria*⁵*Departamento de Física Fundamental e IUFFyM, Universidad de Salamanca,
E-37008 Salamanca, Spain*⁶*Instituto de Física Teórica UAM-CSIC, E-28049 Madrid, Spain*⁷*University of Manchester, School of Physics and Astronomy, Manchester M13 9PL, United Kingdom*

(Received 20 January 2021; accepted 1 June 2021; published 30 July 2021)

We present results for the 2-jettiness differential distribution for boosted top quark pairs produced in e^+e^- collisions in the peak region accounting for QCD large-logarithm resummation at next-to-next-to-next-to-leading logarithmic order and fixed-order corrections to matrix elements at next-to-next-to-leading order calculated in the framework of soft-collinear effective theory and boosted heavy quark effective theory. Electroweak and finite-width effects are included at leading order. We study the perturbative convergence of the cross section in the pole and MSR mass schemes, with and without soft gap subtractions. We find that there is a partial cancellation between the pole mass and soft function renormalons. When renormalon subtractions concerning the top mass and the soft function are implemented, the perturbative uncertainties are, however, systematically smaller and an improvement in the stability of the peak position is observed. We find that the top MSR mass may be determined with perturbative uncertainties well below 100 MeV from the peak position of the 2-jettiness distribution. This result has important applications for Monte Carlo top quark mass calibrations.

DOI: [10.1103/PhysRevD.104.014026](https://doi.org/10.1103/PhysRevD.104.014026)**I. INTRODUCTION**

The top quark mass m_t is one of the most important parameters of the Standard Model (SM). In conjunction with the Higgs boson mass, it is an essential input for studies of the stability of the SM electroweak vacuum [1–6], and it plays an important role in precision electroweak fits [7]. The most precise determinations of the top mass to-date come from so-called “direct measurements,” that are based on the kinematic reconstruction of the final-state top quark decay products and the comparison of the resulting kinematic distributions with parton-shower Monte Carlo (MC) simulations. The current world average for direct measurements reads $m_t^{\text{MC}} = 172.76 \pm 0.30$ GeV [8] and projections for the HL-LHC indicate that uncertainties as small as 200 MeV can be reached for individual measurements [9].

The interpretation of these measurements is, however, (as reviewed below) impacted by an additional ambiguity from the lack of understanding of the field-theoretic meaning of the top mass parameter encoded in the MC event generators [9–12]. This ambiguity is not yet precisely quantified and should be considered at the GeV level, i.e., it is comparable to the uncertainties quoted by the experimental analyses [13]. Carrying out first-principle theoretical predictions of kinematic distributions that exhibit high sensitivity to the top mass is a challenging program in light of the disparate energy scales that enter the top production and the measurements on the decay products. These lead to large logarithms of ratios of these scales that require resummation. Furthermore, due to nonperturbative corrections it is necessary to take into account the hadronic nature of the final state. While these effects can be accounted for via MC event generators, it comes at the cost of the limited perturbative [14,15] and conceptual precision of the MC description of perturbative and nonperturbative effects. This is the origin of the interpretation problem of the top mass parameter employed in the MC generators (see Ref. [13] for further discussion).

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

In order to achieve a precise top quark mass determination we therefore need an observable that has the required kinematic sensitivity and is theoretically tractable, such that (a) it can be reliably calculated in perturbation theory within a specific short-distance top quark mass scheme, and (b) nonperturbative effects can be consistently quantified from a field theory perspective. Such calculations can be carried out in the framework of effective field theories (EFTs) that are systematically improvable in their power-counting expansion, as well as in perturbation theory where resummation of large logarithms now reaches next-to-next-to-next-to-leading logarithmic (N³LL) accuracy for a number of applications [16–19]. Moreover, the framework of EFTs offers ways to rigorously describe and quantify nonperturbative effects due to hadronization [20–23].

Therefore, for such an observable, EFT-based calculations may naturally incorporate all the distinct features of parton-shower MC simulations while retaining a systematic connection to field theory and thus theoretical control. Hence, EFTs offer promising prospects for precision collider physics and for developing diagnostic tools for improving MC simulations [23,24]. Such a framework has recently been applied to the broad program of top quark mass measurements, including a factorized description of a hadron-level differential top jet mass spectrum for jets initiated by boosted top quarks in the peak region at a future e^+e^- collider [25,26], boosted top jets with soft drop grooming at the LHC [23], a calibration of the MC top quark mass parameter [27] (based on the work of Refs. [25,26]), and parton-level studies of the correlation of the MC top quark mass parameter with the parton shower evolution cutoff [15]. The sensitivity of event-shape variable definitions with respect to quark mass effects has also been recently studied in Refs. [28,29] in the context of fixed-order and resummed perturbation theory.

In this work we continue this effort by extending the perturbative calculations of boosted tops in the peak region at an e^+e^- collider from the N²LL resummation of logarithms and $\mathcal{O}(\alpha_s)$ fixed-order matrix elements used in Ref. [27], to N³LL resummation with next-to-next-to-leading-order (NNLO) [$\mathcal{O}(\alpha_s^2)$] fixed-order matrix elements.

A. Status of top mass measurements

Recent direct top quark mass measurements have yielded the results $m_t^{\text{MC}} = 172.26 \pm 0.61$ GeV (CMS) [30] and $m_t^{\text{MC}} = 172.69 \pm 0.48$ GeV (ATLAS) [31] at the LHC, and $m_t^{\text{MC}} = 174.34 \pm 0.64$ GeV [32] at the Tevatron. The top mass superscript MC signifies that the direct measurements extract the top quark mass parameter coded in the MC generators used for the analyses. The lack of understanding of a precise field-theoretic definition of m_t^{MC} (and hence its relation to short-distance masses defined in the context of quantum field theory, which is the preferred input parameter for high-precision theoretical predictions) results in an

additional conceptual uncertainty in how m_t^{MC} should be related to the pole mass m_t^{pole} or a short-distance mass such as the $\overline{\text{MS}}$ mass $\overline{m}_t^{(6)}(\mu)$ or the MSR mass $m_t^{\text{MSR}}(R)$ [33–35]. This ambiguity is not included in the individual quoted experimental uncertainties but should be considered to be (at least) comparable [13]. Analyses that can shed light on a more complete and quantitative understanding of these issues, as well as more first-principle aspects of MC event generators, are underway [15,27,36]. Other recent related studies include analyses of the theoretical limitations concerning the modeling of the dynamics in the top quark production and decay [37], finite lifetime [38], hadronization effects [39], and observable infrared sensitivity [40].

The so-called top quark pole-mass measurements based on the total cross section [41], for which precise theoretical predictions expressed in terms of the pole mass renormalization scheme have been employed, have generally yielded results with larger uncertainties than the direct measurements: $m_t^{\text{pole}} = 172.9_{-2.6}^{+2.5}$ GeV (ATLAS) [42] and $m_t^{\text{pole}} = 172.7_{-2.7}^{+2.4}$ GeV (CMS) [43]. The relatively large errors result from the uncertainty in the normalization of the inclusive cross section (both in the measurement and the theoretical prediction) and its relatively weak dependence on m_t . Employing differential cross sections such as leptonic distributions, and using matched NLO + PS (parton shower) MC simulations results in an enhanced top mass sensitivity [44]. Such differential measurements have been included in the world average $m_t^{\text{pole}} = 172.4 \pm 0.7$ GeV [8] for the pole mass measurements which is in good agreement with the corresponding world average of direct measurements mentioned above. Recent precision measurements, also using matched NLO + PS MC simulations for the theoretical predictions, accounting for $t\bar{t}$ + jet final states [45] ($m_t^{\text{pole}} = 171.1_{-1.0}^{+1.2}$ GeV [46]) and the reconstructed top-antitop invariant mass distribution ($m_t^{\text{pole}} = 170.5 \pm 0.8$ GeV from a simultaneous α_s fit [47]) have comparable uncertainties but have, however, resulted in significantly lower m_t values, posing some tension that may partly arise from missing theoretical input in predictions for the corresponding differential cross sections [13,48,49].

It is well known that the pole mass additionally suffers from a conceptual $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon ambiguity which in dedicated analyses was estimated to amount to 110 MeV in Ref. [50] and to 250 MeV in Ref. [35].¹ This ambiguity does not represent an uncertainty due to perturbative truncation or limited information, but signifies the principle conceptual imprecision in assigning a definite value to m_t^{pole} . The ambiguity is not related to any physical effect,

¹The quoted numbers for the pole mass renormalon ambiguity arise in the context of finite bottom and charm masses. When charm and bottom quarks are treated as massless quarks, the infrared sensitivity decreases and pole mass ambiguity is smaller.

but inherent to the unphysical nature of the pole mass renormalization condition. It can therefore be avoided by expressing cross sections obtained in perturbation theory in terms of a short-distance mass scheme at an appropriate renormalization scale, which can also improve the overall convergence of the perturbative series at the first few orders. As far as the spread of the above-mentioned recent pole mass measurements is concerned, the pole mass renormalon ambiguity likely plays no role, because it is smaller than the quoted uncertainties of these measurements. For the direct top quark mass measurements, the pole mass renormalon problem has been discussed intensely in the context of the frequently used approach of identifying m_t^{MC} and m_t^{pole} . However, in Ref. [15] it was shown analytically for the simple case of the 2-jettiness distribution in $e^+e^- \rightarrow t\bar{t} + X$, that the quark mass parameter associated to a NLL-precise parton shower is in general not the pole mass, but a low-scale short-distance mass that depends on the value of the shower cut and may differ from the pole mass by an amount larger than the pole mass ambiguity.² So the problem of how to properly interpret the MC top quark mass m_t^{MC} in terms of a well-defined and ambiguity-free field theory mass is not related to the pole mass renormalon ambiguity, but to the limited theoretical precision of the state-of-the-art MC event generators and to ignorance concerning MC systematics.

Motivated by the interpretation issues of the direct top mass measurements and the still large uncertainties in the pole mass measurements from inclusive and differential cross sections, a number of alternative methods to measure m_t have been proposed, which are based on differential cross sections with respect to variables constructed from top decay products exhibiting strong kinematic top mass sensitivity. The observables these analyses are based on include the M_{T2} variable and variants of it [51,52], the lepton b -jet invariant mass [53], the shape of b -jet and B -meson energy distributions [54], and the J/ψ and lepton invariant masses [55,56]. Conceptually, these observables are governed by parton shower dynamics as well as various nonperturbative effects, in a way analogous to the direct reconstruction method (albeit with differing systematics). This is because they are also based on the idea of assessing the kinematics of decaying (colored) top quark particles through simulations obtained from parton-shower MCs. Their reliance on these MCs further makes their potential extension to higher logarithmic precision nontrivial, as the theoretical precision of parton showers is quite observable dependent. Additionally, the nonperturbative effects of hadronization are accounted for through a multiparameter

²In Ref. [15] it was shown that the quark mass parameter associated to a NLL-precise parton shower based on the coherent branching algorithm with shower cut Q_0 is a Q_0 -dependent short-distance mass $m_t^{\text{CB}}(Q_0)$ that differs from the pole mass by the amount $m_t^{\text{CB}}(Q_0) - m_t^{\text{pole}} = -(2/3)\alpha_s(Q_0)Q_0$.

MC hadronization model. Here, since a systematic way to quantify the intrinsic MC hadronization uncertainty does not yet exist, uncertainties are typically being estimated by comparing predictions based on different models.

B. Top mass determination using effective field theories

In Refs. [25,26] a framework of EFTs was developed to describe boosted top quarks in the peak region at a future e^+e^- collider. Using soft-collinear effective theory (SCET) [57–61] a factorization theorem for the double-differential hemisphere-mass cross section in $e^+e^- \rightarrow t\bar{t} + X$ was derived in the boosted top quark limit, with center-of-mass energy $E_{\text{cm}} = Q \gg m_t$. The two invariant masses M_t and $M_{\bar{t}}$ are defined using all particles in each of the two hemispheres that are determined by the event’s thrust axis described below. The peak refers to the region where the $M_t - M_{\bar{t}}$ double-differential distribution exhibits the hemisphere-mass top and antitop resonances. One key feature of this factorization formula is that the most important hadronization effects are parametrized via a convolution with a top-mass- and Q -independent nonperturbative shape function, which is field-theoretically defined from a vacuum matrix element of Wilson lines. The universality of the factorization formula states that, at least in principle, data for massless dijet events (obtained even at past experiments such as LEP) could be used to fix the nonperturbative shape function and to make the analysis independent of estimates of nonperturbative corrections obtained from MC event generators. The other key feature is that the dependence on the renormalization scheme of the top quark mass is fully controlled so that one can make and test predictions in any scheme to the extent that higher-order perturbative corrections are incorporated. In the tail of the distribution an operator expansion can be applied such that the nonperturbative corrections are dominated by a single parameter.

The dijet hemisphere-mass cross section exhibits a clear peak at the top and antitop resonances which are directly sensitive to the value of the top quark mass. However, the resonance location in M_t and $M_{\bar{t}}$ is not directly at the top quark mass due to radiative effects related to ultra-collinear and large-angle soft radiation, as well as hadronization and finite-width effects. The ultra-collinear radiation refers to radiation that is soft in the (anti)top quark rest frame, but becomes collinear due to the (anti)top quark boost. In the factorization theorem this ultra-collinear radiation is described by boosted versions of heavy quark effective theory (HQET) [62–66], called boosted HQET or simply bHQET.³ The typical scales of the ultra-collinear radiation in the peak region range between the top width Γ_t and ~ 10 GeV, which also quantifies the typical “off-shellness”

³Here the letter “b” in bHQET stands for the fact that the two HQET theories for the top and antitop quarks are boosted in opposite directions.

$\sim(q^2 - m_t^2)/m_t$ of the decaying top quarks. The top mass dependence of the location and shape of the observable peak, along with the top quark scheme dependence, is specified by ultra-collinear radiation effects which can be calculated perturbatively. Furthermore, the leading-order electroweak effects come from the top width which is fully encoded in Breit-Wigner-modified top quark propagators. The large-angle soft radiation is only sensitive to the collinear top quark color flow (which is fully taken over by the top decay products in the boosted limit) and describes soft momentum exchange between the two hemispheres. It is governed by scales below the top quark width, and its nonperturbative contributions constitute the shape function. Due to the boost of the top quarks, the effect of large-angle soft radiation on the peak locations is enhanced by a factor Q/m_t .

The above-mentioned framework was employed in Ref. [27], where a calibration of the m_t^{MC} parameter in Pythia 8.205 was carried out using the $\text{N}^2\text{LL} + \mathcal{O}(\alpha_s)$ prediction for the 2-jettiness event shape in $e^+e^- \rightarrow t\bar{t} + X$ collisions defined as

$$\tau_2 \equiv 1 - \max_{\hat{n}_t} \frac{\sum_i |\hat{n}_t \cdot \vec{p}_i|}{Q}, \quad (1)$$

where the sum runs over all produced particles i in the event and \hat{n}_t is the thrust axis that maximizes the sum in Eq. (1). Since a sum over all final-state momenta is involved in Eq. (1), we restrict ourselves to hadronically decaying top and antitop jets, which is accomplished by simply including the corresponding branching fraction in the Born cross section σ_0 in the factorization formula discussed below. In the limit $\tau_2 - 2m_t^2/Q^2 \ll 1$, i.e., when focusing on the peak region, the event shape corresponds (up to power corrections) to the sum of the hemisphere-invariant masses, such that

$$\tau_2 = \frac{M_t^2 + M_{\bar{t}}^2}{Q^2} + \mathcal{O}(\tau_2^2), \quad (2)$$

and thus has the same kinematic sensitivity to the top mass as the double-differential cross section considered in Refs. [25,26]. Being a global event shape, the τ_2 differential cross section is furthermore free from nonglobal logarithms [67,68]. The results of the calibration carried out in Ref. [27] showed that m_t^{MC} in Pythia 8.205 cannot be simply identified with the pole mass, but it is numerically close to the MSR mass $m_t^{\text{MSR}}(1 \text{ GeV})$. Compatible numerical results were obtained in the analysis of jets from top quarks produced in pp collisions with soft drop grooming in Ref. [36] based on theoretical calculations at NLL order.

In this work we improve the calculation of the 2-jettiness event shape cross section in $e^+e^- \rightarrow t\bar{t} + X$ by including NNLO fixed-order matrix elements with N^3LL resummation [referred to as $\text{N}^3\text{LL} + \mathcal{O}(\alpha_s^2)$]. By now, all the

ingredients that enter the factorization formula in Eq. (3) below are individually known to the accuracy needed to achieve this precision for the 2-jettiness cross section, in particular the 2-loop heavy quark jet function [69], full 2-loop thrust soft function [70,71], and the N^3LL result for the Wilson coefficient for the matching at the top mass scale [72]. We also include the recently calculated analytic 4-loop result for the cusp anomalous dimension [73], which despite having a tiny numerical impact, is an important formal ingredient to obtain N^3LL accuracy. We consistently combine these ingredients to obtain the $\text{N}^3\text{LL} + \mathcal{O}(\alpha_s^2)$ boosted top cross section fully analytically, and study the convergence of the resummed perturbation theory. We exclusively consider the e^+e^- 2-jettiness cross section in the boosted and bHQET limits where power corrections in m_t/Q and in the top quarks' off-shellness $\sim(q^2 - m_t^2)/m_t$ are neglected. These corrections, even though they are formally power-suppressed, can be non-negligible for phenomenological analyses in the peak region. With the inclusion of these power corrections, the results obtained in this article will serve to improve analyses such as the top quark mass calibration carried out in Ref. [27], which we leave to future work.

The outline of the paper is as follows. We first present the factorization formula in Sec. II, describing its key features and elements. We describe in Sec. III the implementation of the cross section in terms of short-distance mass schemes and a renormalon-free soft function in Sec. IV. The setup of observable-dependent renormalization scales is discussed in Sec. VI. Finally, in Sec. VII we combine all of the pieces to calculate the resummed cross section and study its perturbative convergence when the renormalons in the pole mass and the soft function are either subtracted or left unsubtracted. Using an analysis of the 2-jettiness peak location we draw conclusions on the perturbative uncertainties of a determination of the MSR mass and the pole mass. In the Appendices we review and state additional details for the various ingredients that are needed in this analysis. We conclude in Sec. VIII.

II. FACTORIZATION THEOREM IN THE PEAK REGION

In Ref. [25] two factorized expressions for the differential cross section were derived that are valid in the peak and tail regions of the double hemisphere-mass, or equivalently 2-jettiness, distribution. In the tail region the fluctuations in the top mass can be large, such that $M_{t,\bar{t}}^2 - m_t^2 \sim m_t^2$, and a factorization formula based on SCET with massive particles can be derived. On the other hand, in the peak region, the off-shellness of the decaying top is constrained such that $M_{t,\bar{t}}^2 - m_t^2 \ll m_t^2$, necessitating an additional factorization and a resummation accounting for the top width as an additional relevant scale that is carried out in the bHQET framework. As already mentioned, in this work we focus on

the factorization in the peak region accounting exclusively for the bHQET contributions, leaving off-shellness power corrections described in the SCET factorization (that become essential in the tail region) and the inclusion of m_t/Q SCET power corrections to future work. The factorization formula in the peak region is given by

$$\begin{aligned} \frac{1}{\sigma_0} \frac{d\sigma}{d\tau_2} &= m_t Q^2 H_{\text{evol}}^{(5,6)}(Q, m_t, Q, \mu; \mu_H, \mu_m) \\ &\times \int d\ell d\hat{s} U_B^{(5)}(\hat{s}_\tau - Q\ell - \hat{s}, \mu, \mu_B) J_{B,\tau_2}^{(5)}(\hat{s}, \Gamma, \delta m, \mu_B) \\ &\times \int d\ell' dk U_S^{(5)}(\ell - \ell', \mu, \mu_S) \hat{S}_{\tau_2}^{(5)}(\ell' - k, \bar{\delta}, \mu_S) \\ &\times F(k - 2\Delta), \end{aligned} \quad (3)$$

where we have the boost parameter Q given by

$$Q \equiv \frac{Q}{m_t} \quad (4)$$

and have defined the off-shellness variable \hat{s}_τ as⁴

$$\hat{s}_\tau \equiv \frac{Q^2 \tau_2 - 2m_t^2}{m_t}. \quad (5)$$

Equation (3) involves various perturbative ingredients, including an evolved matching function $H_{\text{evol}}^{(5,6)}$, jet and soft functions $J_{B,\tau_2}^{(5)}$ and $\hat{S}_{\tau_2}^{(5)}$, and evolution kernels $U_B^{(5)}$ and $U_S^{(5)}$. It also includes a nonperturbative shape function F , whose independence from other parameters is a prediction of the factorization theorem. These ingredients will be discussed in detail in the subsections below. Equation (3) applies in the peak region where $\hat{s}_\tau \sim \Gamma \ll m_t$, with $\Gamma \gtrsim 2\Gamma_t$ being the effective width of the distribution broadened by radiation, hadronization, and finite-width effects. In this region, where the τ_2 distribution exhibits a resonance, the 2-jettiness variable is (up to power corrections) directly related to the sum of the squared hemisphere masses defined with respect to the thrust axis, as given in Eq. (2). It is therefore convenient to define the inclusive jet mass variable M_J ,

$$M_J^2 \equiv \frac{1}{2} Q^2 \tau_2, \quad (6)$$

which inherits some of the features of a reconstructed top invariant mass, albeit being based on a hemisphere top jet. The M_J distribution peaks close to m_t , but is in addition affected (with respect to peak position as well as the width of the observed peak resonance) by large-angle soft radiation exchanged between the two hemispheres. The widening of the peak due to top-decay width and soft QCD effects,

however, does not affect the kinematic sensitivity of the observable, and normalizing the M_J distribution enables the uncertainties in the M_J peak location to be taken as a direct measure for the uncertainties in the associated top mass determination.

The factorization formula separates perturbative contributions from the hard local interactions involving the scales Q and m_t (encoded in the hard factor $H_{\text{evol}}^{(5,6)}$), dynamical effects associated to large-angle soft radiation (accounted for in the soft function $S_{\tau_2}^{(5)}$), and dynamical effects due to ultra-collinear radiation (contained in the jet function $J_{B,\tau_2}^{(5)}$). The jet function $J_{B,\tau_2}^{(5)}$ incorporates the leading-order effects due to the top quark width Γ and also carries, because of its peaked structure, the main top quark mass sensitivity of the τ_2 distribution. This allows testing at high precision the impact of using either the pole mass scheme m_t^{pole} or a suitable short-distance mass m_t^{sd} . This is indicated by the argument δm_t , where⁵

$$\delta m \equiv m_t^{\text{pole}} - m_t^{\text{sd}} \quad (7)$$

is the perturbative series for the difference between the pole and the adopted short-distance masses. The choice $\delta m = 0$ implies the use of the pole mass scheme. We note that there is also top quark mass (scheme) dependence in the hard factor $H_{\text{evol}}^{(5,6)}$, indicated by the argument m_t , which, however, only affects the normalization of the τ_2 distribution and is very subdominant compared to the main sensitivity to the top quark mass. The different character of the top mass dependence in the hard and bHQET jet functions is discussed in more detail below.

For the τ_2 distribution in the peak region for boosted top quarks, we have the hierarchy $Q \gg m_t \gg \{\hat{s}_\tau, \Gamma\}$. This leads to large logarithms, which are resummed via renormalization group (RG) equations for the corresponding perturbative matrix elements. This resummation is implemented through the evolution factors $U_i^{(n)}$ [$i = H, v$ (or m), S, B], which RG evolve each of the functions appearing in the factorization theorem from their natural scales μ_i to a common final scale μ . The quantities μ_i ($i = H, m, S, B$) are renormalization scales indicating the natural physical values of the quantum fluctuations encoded in the respective factors. These scales are varied in the final results in order to assess the theoretical uncertainties due to missing higher-order contributions. The choice of μ is arbitrary and the factorized prediction is (strictly) invariant under changes of this μ . In contrast the dependence on the initial scales μ_i only cancels out order by order in resummed perturbation theory. The scale μ is typically set equal to one

⁴The limit $\hat{s}_\tau \rightarrow 0$ corresponds to the tree-level kinematics for $e^+e^- \rightarrow t\bar{t}$, where $\tau_2 = 2m_t^2/Q^2$.

⁵In the context of HQET, the mass scheme correction δm is called the residual mass term.

of the renormalization scales μ_i ($i = H, m, S, J$) such that one of the renormalization evolution factors disappears.

The superscripts “(5)” and “(5,6)” indicate the number of active dynamical flavors relevant for the momentum scales of the respective quantum effects, where “(5)” and “(6)” indicate scales below and above the top mass, respectively.

Due to the simple inclusive character of the 2-jettiness (or the M_J) distribution, the leading-order nonperturbative effects arise from the low-energy dynamics of the large-angle soft radiation and are encoded in the (hadronization) shape function F , which is convolved with the perturbative soft function $S_{r_2}^{(5)}$. Physically, the shape function incorporates the leading effects of hadronization and controls the amount of nonperturbative radiation being exchanged between the two hemispheres. Even though the details of the shape function form must be determined from experimental data, the way how the shape function appears in the factorized cross section represents a very strong theoretical constraint on hadronization.

In the following we briefly review the physical aspects of all functions appearing in the factorization theorem one by one. For a detailed discussion on how the bHQET factorization theorem of Eq. (3) connects to the corresponding formula in the tail region, and possible alternative setups to organize the renormalization group evolution, we refer to Ref. [74].

A. The hard function

The factorized cross section involves a two-step matching from QCD to SCET, and then from SCET to bHQET, at the scales $\mu_H \sim Q$ and $\mu_m \sim m_t$, respectively. Therefore, the difference between 6 or 5 flavors is related to the top quark being a dynamical degree of freedom or not. The resulting hard matching coefficients in Eq. (3), together with their RG evolution kernels, are collectively written as

$$\begin{aligned} H_{\text{evol}}^{(5,6)}(Q, m_t, q, \mu; \mu_H, \mu_m) \\ \equiv H_Q^{(6)}(Q, \mu_H) \\ \times U_{H_Q}^{(6)}(Q, \mu_H, \mu_m) H_m^{(6)}(m_t, q, \mu_m) U_v^{(5)}(q, \mu_m, \mu). \end{aligned} \quad (8)$$

The term U_{H_Q} evolves the SCET hard function H_Q [75–80] from $\mu_H \sim Q$ to $\mu_m \sim m_t$ and resums large logarithms of Q/m_t in the cross section. U_v is responsible for the evolution of the bHQET current between $\mu_m \sim m_t$ and μ , which we assume is smaller than μ_m , and only depends on the top quark boost factor q defined in Eq. (4). The hard matching at the top quark mass scale μ_m , which encodes off-shell top quark quantum fluctuations that arise in the heavy quark limit, is given by H_m [26,72]. Since the matching is performed at the top quark mass, one can express H_m in terms of α_s with either 5 or 6 active flavors. In the numerical analysis below we choose six. The effect

of this freedom in the scheme choice is, however, tiny and numerically irrelevant.

We now discuss the parameter $q = Q/m_t$ [see Eq. (4)] appearing in the mass mode matching factor $H_m^{(6)}$ and the bHQET current evolution kernel $U_v^{(5)}$. In the peak region, for the bHQET factorization treatment of the top quark dynamics, the momentum of the nearly on-shell top quarks is parametrized as $p_{t,\bar{t}}^\mu = m_t v_{t,\bar{t}}^\mu + k^\mu$, where $v_{t,\bar{t}}^\mu$ (with $v_{t,\bar{t}}^2 = 1$) is the 4-velocity of the energetic (anti)top quarks and k^μ is a small residual momentum accounting for the fluctuations caused by the low-energy radiation [in the (anti)top quark rest frame], such that $k^\mu \ll m_t$ and one can expand the dynamical effects to leading power in $k^\mu/m_t \sim \Gamma_t/m_t$ [in the (anti)top rest frame]. The reference 4-velocities $v_{t,\bar{t}}^\mu$ are defined by

$$v_t^\mu = (q^{-1}, q, \vec{0}_\perp), \quad v_{\bar{t}}^\mu = (q, q^{-1}, \vec{0}_\perp), \quad (9)$$

using light-cone coordinates defined relative to the thrust axis $n^\mu = (1, \vec{n}_t)$ (which we take along the top direction), such that $p^\mu = (n \cdot p, \vec{n} \cdot p, \vec{p}_\perp)$, where $\vec{n}^\mu = (1, -\vec{n}_t)$ is an auxiliary vector satisfying $n \cdot \vec{n} = 2$ and $n^2 = \vec{n}^2 = 0$. The parameter q appearing in the definition of the reference velocities is related to the top quark boost and defined in Eq. (4), such that the on-shell (anti)top 4-velocity would approach $v_{t,\bar{t}}^\mu$ in the boosted limit $Q \gg m_t$ in the absence of any radiation. The mass mode matching factor $H_m^{(6)}$ and the bHQET current evolution kernel $U_v^{(5)}$ depend on the reference velocities v_t^μ and $v_{\bar{t}}^\mu$ and thus on q through the scalar products $v_{t,\bar{t}} \cdot k$ appearing in the bHQET Feynman diagrams.

The choice of $v_{t,\bar{t}}^\mu$ is ambiguous with respect to higher-order power corrections of $\mathcal{O}(k^\mu/m_t)$ [$\sim \mathcal{O}(\Gamma_t/m_t)$ in the (anti)top rest frame] and entails a symmetry of bHQET factorization with respect to changes of $v_{t,\bar{t}}^\mu$ and thus of q , which is one aspect of a more general class of symmetry transformations called reparametrization invariance [81] that connects different orders in the bHQET $1/m_t$ expansion.

As a consequence, there is a power-suppressed freedom in the choice of the boost parameter q , which in the factorization theorem of Eq. (3) appears in the function $H_{\text{evol}}^{(5,6)}$ and the momentum argument of the jet-function evolution factor $U_B^{(5)}$. The m_t parameter appearing in q is, however, not associated to any particular top mass renormalization scheme, but for the final evaluation it just has to be numerically chosen close to the invariant mass of the top quark in the resonance region. It is therefore not mandatory to reexpand the boost parameter q when expressing the top quark mass in a short-distance renormalization scheme. The other consequence is, that variation $q \rightarrow q + \delta q$ with $\delta q \sim Q\Gamma_t/m_t^2$ in the peak region account for uncertainties due to the truncation of power corrections in the context of the leading-power bHQET factorization.

We conclude by noting that actually setting the parameter m_t in q equal to the pole mass artificially reintroduces the pole mass renormalon (albeit with formally power-suppressed numerical effects) which cannot be cured by a reexpansion in terms of a short-distance mass, because there are no associated factorially divergent perturbative corrections in other parts of the factorization theorem. It is therefore not advisable to identify the parameter m_t in q as the pole mass.

The top mass parameter m_t that appears in the argument of $H_m^{(6)}$ in Eq. (8) is identical to the mass used in the threshold decoupling relation of the strong coupling when transitioning between 6 and 5 flavors. This m_t can be expressed in either the pole or $\overline{\text{MS}}$ scheme (with the same mass scheme employed in the $\alpha_s^{(6)} \leftrightarrow \alpha_s^{(5)}$ decoupling relation). Since H_m only involves logarithms of the ratio μ_m/m_t (in addition to logs of q) it is also insensitive to the $\mathcal{O}(\Gamma_t)$ fluctuations at the top threshold that are suppressed by $\mathcal{O}(\Gamma_t/m_t)$. The difference between the choice of pole or $\overline{\text{MS}}$ masses is accordingly found to be also numerically subleading.

B. The ultra-collinear sector

Upon integrating out the off-shell (small) components of the top quark field in bHQET, the leading-order dynamics of the remaining ultra-collinear fluctuations along the top and antitop quark directions is captured by the 2-jettiness bHQET jet function, defined as

$$J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \Gamma_t, \delta m, \mu_B) = \int d\hat{s} J_B^{(5)}(\hat{s}, \Gamma_t, \delta m, \mu_B) J_B^{(5)}(\hat{s}_\tau - \hat{s}, \Gamma_t, \delta m, \mu_B). \quad (10)$$

Here ‘‘ultra’’ distinguishes the collinear modes in the peak region $\hat{s}_\tau \gtrsim 2\Gamma_t$ [which are soft in the (anti)top quark rest frame] from higher virtuality collinear modes appearing in the tail region $\hat{s}_\tau \sim m_t$, where bHQET off-shellness power corrections become large and SCET provides the adequate description for collinear radiation. Here $J_B(\hat{s}_t, \delta m_t, \Gamma_t, \mu_B)$ is the familiar bHQET jet function that appears in di-hemisphere-mass and jet mass distributions [15,26,69]. At leading order, the finite top width effects of the bHQET jet function $J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \Gamma_t, \delta m, \mu_B)$ can be expressed as a convolution of its stable-top version $J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \Gamma_t = 0, \delta m, \mu_B)$ with an inclusive Breit-Wigner function [26],

$$J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \Gamma_t, \delta m, \mu_B) = \int \frac{d\hat{s}'}{\pi} \frac{2\Gamma_t}{(2\Gamma_t)^2 + (\hat{s}_\tau - \hat{s}')^2} \times J_{B,\tau_2}^{(5)}(\hat{s}', \Gamma_t = 0, \delta m, \mu_B), \quad (11)$$

where the factor $2\Gamma_t$ arises from accounting for the widths from both top and antitop quarks. This implies that the measurement on $t\bar{t}$ final states is fully inclusive in the decay

products as well as any radiation from them. The consistency of this inclusive treatment is ensured by considering the boosted limit $Q \gg m_t$ where the decay products from the top and antitop quarks are collimated back to back along the direction of the thrust axis in distinct hemispheres. Both the boosted top decay products as well as gluon and light-quark radiation encoded in J_B have typical angles $\sim 2m_t/Q$ relative to these axes.⁶ The variable \hat{s}_t appearing in the (single) hemisphere jet functions $J_B^{(5)}$ in Eq. (10) is equivalent to $\hat{s}_t = 2v_t \cdot k$, where k is the total residual momentum of the collimated system (after removing the contribution from the top quark mass). This captures the invariant mass of the decaying top quark system together with its ultra-collinear radiation, up to terms of $\mathcal{O}(\hat{s}_t^2/m_t)$ [25]. In the peak region, the ultra-collinear fluctuations have a virtuality of $\mathcal{O}(\hat{s}_{t,\bar{t}} \gtrsim \Gamma_t)$, which leads to $\hat{s}_\tau = \hat{s}_t + \hat{s}_{\bar{t}} \gtrsim 2\Gamma_t$. The natural choice for the renormalization scale of the 2-jettiness bHQET jet function is $\mu_B \sim \hat{s}_\tau \sim (Q^2\tau_2 - 2m_t^2)/m_t$.

For $\Gamma_t = 0$ the 2-jettiness bHQET jet function $J_{B,\tau_2}^{(5)}$ has support only for non-negative \hat{s}_τ and equals $\delta(\hat{s}_\tau)/m_t$ at tree level. It is this threshold behavior which causes the strong top mass sensitivity of the τ_2 distribution in the peak region. We emphasize, however, that while the partonic threshold is at $\hat{s}_\tau = 0$, the observable peak position exhibited by the entire factorization theorem of Eq. (3) is determined coherently from the effects of the ultra-collinear and large-angle soft radiation together with the Breit-Wigner smearing. The well-known pole mass renormalon problem arises from higher-order perturbative corrections in $J_{B,\tau_2}^{(5)}$ in the pole mass scheme, and is encoded in the size of the coefficients of plus distributions in \hat{s}_τ . The pole mass renormalon can be remedied by using, instead of the pole mass m_t^{pole} , a suitable short-distance mass scheme m_t^{sd} in the definition of \hat{s}_τ in Eq. (5). Because the bHQET jet functions are defined strictly at leading order in the $1/m_t$ expansion, switching the top mass renormalization scheme requires that one also accounts for the mass scheme correction

$$\delta m(R) = m_t^{\text{pole}} - m_t^{\text{sd}}(R) = \sum_{i=0} \left[\frac{\alpha_s^{(5)}(\mu_B)}{4\pi} \right]^i \delta m_i(R), \quad (12)$$

strictly to leading order in the $1/m_t$ expansion. This leads to the generic form

$$J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \Gamma_t = 0, \delta m(R), \mu_B) \equiv \frac{1}{m_t^2} \sum_{i=0} \left[\frac{\alpha_s^{(5)}(\mu_B)}{4\pi} \right]^i \sum_{j=-1}^{2i+1} \frac{B_{ij}}{\mu_B} \mathcal{L}_j \left(\frac{\hat{s}_\tau - 4\delta m(R)}{\mu_B} \right), \quad (13)$$

⁶Other global event shapes such as the C -parameter are more sensitive to the kinematical distribution of the top decay products, such that the Breit-Wigner approximation is inaccurate even in the boosted limit [82].

where the B_{ij} are constant coefficients. This yields the 2-jettiness bHQET jet function for stable top quarks in an arbitrary (short-distance) mass scheme m_t^{sd} , where it is *still strictly mandatory* to expand the dependence on δm consistently in powers of the strong coupling $\alpha_s^{(5)}(\mu_B)$ such that the pole mass renormalon cancels order by order. Setting $\delta m = 0$, one recovers the corresponding result in the pole mass scheme, which was calculated up to $\mathcal{O}(\alpha_s^2)$ in Refs. [26,69]. Here, $\mathcal{L}_k(x) = [\Theta(x) \frac{\log^k(x)}{x}]_+$ is the standard plus function distribution with a vanishing integral over the range $x \in [0, 1]$ for $k \geq 0$ and $\mathcal{L}_{-1}(x) = \delta(x)$.

From the expression in Eq. (13) we can also clearly see that the bHQET power counting requires that the perturbative series for the mass scheme correction $\delta m(R)$ obeys the scaling $\delta m \sim \hat{s}_\tau \ll m_t$. This shows from a power-counting point of view why low-scale short-distance masses have to be employed and the $\overline{\text{MS}}$ mass, which has $\delta m \sim m_t$, is forbidden. Such low-scale short-distance masses always involve an infrared subtraction scale R , which is necessary to eliminate the large $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon corrections appearing in the pole mass scheme [33]. To avoid upsetting the bHQET jet function power counting it is important that the series coefficients defining the low-scale short-distance mass in Eq. (12) have the property $\delta m_i(R) \propto R$, where the infrared scale R is parametrically close to the typical ultra-collinear scale, i.e., $R \sim \mu_B \sim \hat{s}_\tau$. In our analysis we employ the MSR mass $m_t^{\text{MSR}}(R)$ [33–35,83] that satisfies this requirement, as explained in more detail in Sec. III.

We finally note that the overall factor $(1/m_t)^2$ appearing in the generic expression for the 2-jettiness bHQET jet function $J_{B,\tau_2}^{(5)}$ in Eq. (13) arises from Eq. (10) since each J_B has a $1/m_t$ factor [26]. The renormalization scheme dependence of this factor is power suppressed in the leading-order bHQET expansion and therefore in principle not specified at the level of the factorization theorem of Eq. (3).

C. The soft sector and nonperturbative effects

The thrust partonic soft function $\hat{S}_{\tau_2}^{(5)}$ accounts for the effects of large-angle soft radiation with respect to the thrust axis [16,17,70,71,84,85]. In the context of inclusive observables, such as the global event shape 2-jettiness, it accounts for the so-called ultrasoft modes, which have scale fluctuations of the order of $\mu_S \sim \hat{s}_\tau/Q$, parametrically smaller than the typical scale μ_B for the ultra-collinear modes. Thus, μ_S represents the smallest perturbative scale relevant for the τ_2 distribution.

For global event shapes such as τ_2 or jet-based observables without jet grooming, the leading nonperturbative effects from hadronization can be described via convolution with a shape function $F(k - 2\Delta)$,

$$S_{\tau_2}(\ell, \mu_S) = \int_0^\ell dk \hat{S}_{\tau_2}^{(5)}(\ell - k, \bar{\delta} = 0, \mu_S) F(k - 2\Delta), \quad (14)$$

where $F(k)$ has support for $k \geq 0$, peaks at $k \sim \Lambda_{\text{QCD}}$ [21,86], and is by definition normalized to unity, $\int_0^\infty dk F(k) = 1$. Here, Δ is a model parameter which accounts for the minimum hadronic energy deposit in each hemisphere (hence the factor of 2), referred to as the “gap.” In the tail region where $\ell \sim \mu_S \gg \Lambda_{\text{QCD}}$, one can expand for large ℓ and the most important nonperturbative effect is encoded in the first moment $\overline{\Omega}_1$ of the shape function,

$$\begin{aligned} S_{\tau_2}(\ell \gg \Lambda_{\text{QCD}}, \mu_S) &= \hat{S}_{\tau_2}^{(5)}(\ell, \bar{\delta} = 0, \mu_S) \\ &\quad - 2\overline{\Omega}_1 \hat{S}_{\tau_2}^{(5)'}(\ell, \bar{\delta} = 0, \mu_S) + \dots, \\ 2\overline{\Omega}_1 &\equiv 2\Delta + \int_0^\infty dk k F(k), \end{aligned} \quad (15)$$

where $\overline{\Omega}_1$ can also be expressed as a vacuum matrix element of soft Wilson lines [20,22], and the ellipses represent higher-order terms in the expansion. The concrete form of the shape function we use for the generic numerical examination carried out in Sec. VII is provided in Appendix G.

As we move further into the peak region, μ_S decreases and eventually approaches the nonperturbative scale Λ_{QCD} . Thus, in the peak region $\mu_S \sim 1 \text{ GeV} > \Lambda_{\text{QCD}}$ and the effects of the shape function have to be accounted for exactly in terms of the convolution of Eq. (14). Even though this in principle implies that an infinite amount of information could be required to fix the analytic form of the shape function, the normalization and the requirement that all moments of F exist, together with the fact that factorization strictly demands convolution, allow us to reliably constrain the form of the shape function in terms of a few parameters by means of an expansion in optimally designed basis functions [87]; see, e.g., Refs. [27,88]. In practice, determining the first moment of the shape function fixes the bulk of the information encoded in it for the whole τ_2 spectrum.

The factorization into partonic soft and nonperturbative shape functions displayed in Eq. (14) depends on the regularization and renormalization schemes that are employed for the computation of the partonic soft function. The argument “ $\bar{\delta} = 0$ ” shown in Eq. (14) stands for the standard $\overline{\text{MS}}$ scheme. As was shown in Ref. [21], these prescriptions entail that the partonic soft function has an $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon which affects the partonic threshold at $\ell = 0$ for the perturbative ultrasoft radiation [21,89] in a way very similar to how the $\mathcal{O}(\Lambda_{\text{QCD}})$ pole mass renormalon affects the threshold \hat{s}_τ for the partonic ultra-collinear radiation. Both features are physically disentangled by the fact that the factorization theorem in Eq. (3) predicts that the ultrasoft effects on the hemisphere jet masses (and thus τ_2) are enhanced by a factor q compared to the ultra-collinear effects. This also entails that for a top quark mass determination, different c.m. energies Q and simultaneous

fits including parameters of the shape function must be considered to lift the degeneracy between m_t and hadronization effects. This is in close analogy to the α_s determinations from e^+e^- event-shape data carried out in Refs. [19,90,91], where nonperturbative effects were not fixed from hadronization corrections in MC event generators, but from a simultaneous fit using data obtained for different c.m. energies Q .

The difference between the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons affecting the partonic soft and bHQET jet functions is that the former cancels inside Eq. (15) with the nonperturbative matrix element $\overline{\Omega}_1$, while the latter is only an artifact of the pole mass and nonexistent when employing a short-distance mass scheme. One can remove the partonic soft function renormalon in an analogous way by reexpressing the first moment $\overline{\Omega}_1$ in a new scheme that includes a perturbative subtraction:

$$\begin{aligned} 2\Omega_1(R_s) &\equiv \int_0^\infty dk k F(k) + 2\overline{\Delta}(R_s), \\ \overline{\Delta}(R_s) &\equiv \Delta - \bar{\delta}(R_s), \end{aligned} \quad (16)$$

where

$$\bar{\delta}(R_s) = R_s \sum_{i=0} \left[\frac{\alpha_s^{(5)}(\mu_S)}{4\pi} \right]^i \bar{\delta}_i \quad (17)$$

is a perturbative series constructed such that it has exactly the same renormalon as $\overline{\Omega}_1$ and the $\overline{\text{MS}}$ renormalized partonic soft function $\hat{S}_{\tau_2}^{(5)}(\ell, \bar{\delta} = 0, \mu_S)$. Equation (16) implies that

$$\begin{aligned} \Omega_1(R_s) &= \overline{\Omega}_1 - \bar{\delta}(R_s) \\ &= \overline{\Omega}_1 - \Delta + \overline{\Delta}(R_s). \end{aligned} \quad (18)$$

This entails the introduction of the scale $R_s \gtrsim \Lambda_{\text{QCD}}$ (in analogy to the scale R for the MSR mass discussed below), which effectively represents an infrared cut for the partonic soft function which is then free from the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon ambiguity. This results in $\overline{\Delta}(R_s)$, and hence $\Omega_1(R_s)$, being renormalon free, and that instead of the $\overline{\text{MS}}$ renormalized partonic soft function $\hat{S}_{\tau_2}^{(5)}(\ell, \bar{\delta} = 0, \mu_S)$ one employs the ‘‘gap subtracted’’ partonic soft function [21]:

$$\begin{aligned} \hat{S}_{\tau_2}(\ell, \bar{\delta}(R_s), \mu_S) &\equiv \hat{S}_{\tau_2}(\ell - 2\bar{\delta}(R_s), \mu_S) \\ &= \sum_{i=0} \left[\frac{\alpha_s^{(5)}(\mu_S)}{4\pi} \right]^i \sum_{j=-1}^{2i+1} \frac{S_{ij}}{\mu_S} \mathcal{L}_j \left(\frac{\ell - 2\bar{\delta}(R_s)}{\mu_S} \right). \end{aligned} \quad (19)$$

In analogy to the partonic 2-jettiness bHQET jet function, it is *strictly mandatory* to expand the dependence on $\bar{\delta}$

consistently in powers of the strong coupling $\alpha_s^{(5)}(\mu_S)$ such that the soft function $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon consistently cancels order by order. Furthermore, to avoid upsetting the soft function power counting and to avoid the appearance of large logarithms in the subtraction it is mandatory that R_s is parametrically close to the typical soft scale, i.e., $R_s \sim \mu_S \sim \hat{s}_\tau/q$. Setting $\bar{\delta} = 0$, one recovers the $\overline{\text{MS}}$ renormalized partonic soft function. Hence, the hadron-level soft function becomes

$$S_{\tau_2}(\ell, \mu) = \int_0^\ell dk \hat{S}_{\tau_2}(\ell - k, \bar{\delta}(R_s), \mu) F(k - 2\overline{\Delta}(R_s)). \quad (20)$$

We discuss the precise definition of the scheme that defines $\overline{\Delta}(R_s)$ in Sec. IV.

At this point we note that there is a partial cancellation between the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons in the partonic soft and bHQET jet functions as the corresponding ambiguities and the associated diverging behavior of the perturbative series are equally severe but have opposite sign [15]. The extent of the cancellation is Q dependent through the factor q that enters the factorization convolution between the partonic soft and bHQET jet functions in Eq. (3). As a consequence, the impact of the soft function renormalon in the 2-jettiness distribution increases with Q , while the impact of the jet function renormalon does not. This means that the overall effect of both $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons may be hidden for a certain range of Q values when the pole mass is used in the jet function and no gap subtraction is carried out for the partonic soft function. As we show in Sec. VII this indeed happens within the range of Q values relevant for top mass determinations. However, since simultaneous fits for different c.m. energies Q are mandatory to independently determine the top mass and the parameters of the shape function without degeneracy, the impact of the individual $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons cannot be avoided, so that the perturbative uncertainties in analyses accounting for both types of renormalon subtractions are systematically smaller than those when renormalon subtractions are not implemented.

III. TOP MASS SCHEMES

In our analysis we employ two renormalization schemes for the top quark mass: the pole mass m_t^{pole} and the scale-dependent MSR mass $m_t^{\text{MSR},(5)}(R)$ [33–35]. The pole mass scheme has—from a technical point of view—the simplest implementation because we can set the residual mass term δm appearing in Eqs. (3) and (13) to zero and all entries of m_t discussed before to m_t^{pole} . However, employing the pole mass scheme entails an $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon ambiguity, which visibly destabilizes the order-by-order behavior of the cross section in the peak region, as we show explicitly in Sec. VII. This can be systematically avoided by using a suitable short-distance mass scheme. The MSR mass

$m_t^{\text{MSR},(5)}(R)$ is a short-distance scheme that is derived from the $\overline{\text{MS}}$ mass and represents an extension of the $\overline{\text{MS}}$ mass concept: while the $\overline{\text{MS}}$ mass $\overline{m}_t^{(6)}(\mu)$ is suitable for scales $\mu \geq m_t$, the MSR mass is appropriate for scales $R \leq m_t$.

For the $\overline{\text{MS}}$ mass we will use the notation $\overline{m}_t^{(6)} \equiv \overline{m}_t^{(6)}(\mu = \overline{m}_t^{(6)})$ below. In the approximation that all quark flavors lighter than the top are massless (which we adopt in our analysis), the defining series for $m_t^{\text{pole}} - m_t^{\text{MSR},(5)}(R)$ is obtained⁷ from the corresponding series for $m_t^{\text{pole}} - \overline{m}_t^{(6)}$ by removing all corrections arising from the self-energy diagrams with top quark loops and by setting $\overline{m}_t^{(6)} \rightarrow R$ as well as $\alpha_s^{(6)} \rightarrow \alpha_s^{(5)}$:

$$\delta m(R) = m_t^{\text{pole}} - m_t^{\text{MSR},(5)}(R) = R \sum_{i=1} \left[\frac{\alpha_s^{(5)}(R)}{4\pi} \right]^i a_i. \quad (21)$$

We give the details of the MSR scheme definition and the numerical values of the coefficients in Appendix D. This means that the MSR mass is a scheme derived from the $\overline{\text{MS}}$ mass, but where the virtual off-shell fluctuations in the on-shell self-energy from scales beyond R (which includes virtual top quark effects) are integrated out. The MSR mass is therefore designed for top-mass-dependent observables sensitive to soft QCD dynamics.

The scale R can be interpreted as the resolution scale below which virtual self-energy and real ultra-collinear radiation is treated as unresolved, so that only self-energy contributions above R are absorbed into the mass. This means that $m_t^{\text{MSR},(5)}(R)$ for $R < m_t$ and $\overline{m}_t^{(6)}(R)$ for $R > m_t$ contain self-energy contributions coming only from scales above R . This interpretation entails that in the limit $R \rightarrow 0$, where all virtual self-energy and real ultra-collinear radiation is treated as resolved and all virtual self-energy contributions are absorbed in the mass, we approach the pole mass, which is precisely expressed in Eq. (21). The renormalon ambiguity of the pole mass can thus be seen to be associated with the problem that the limit $R \rightarrow 0$ involves crossing the Landau pole of the strong coupling which *a priori* cannot be carried out in an unambiguous way.

Here we use the interpretation in Ref. [34], where the top MSR mass $m_t^{\text{MSR},(5)}(R)$ is regarded as the 5-flavor extension of the 6-flavor $\overline{\text{MS}}$ mass $\overline{m}_t^{(6)}(R)$ for scales R below the top quark mass and where both mass schemes are matched at the scale $R = \overline{m}_t^{(6)}$. The matching relation is given in Eq. (D3). For scales $R < \overline{m}_t$ the MSR mass evolves with the R -evolution equation

$$\frac{d}{d \log(R)} m_t^{\text{MSR},(5)}(R) = -R \sum_{n=0} \gamma_n^R \left[\frac{\alpha_s^{(5)}(R)}{4\pi} \right]^{n+1}, \quad (22)$$

where γ_n^R are obtained from the coefficients a_i in Eq. (21) using the procedure outlined in Appendix F.

As explained in Sec. II B, the consistent use of a short-distance mass in Eq. (3) entails that the MSR mass scale R satisfies the parametric relation $R \sim \mu_B \sim \hat{s}_\tau$, which means that R depends on τ_2 and $m_t^{\text{MSR},(5)}$ adopts the status of a dynamical scale-dependent “mass coupling” in complete analogy to the well-known concept of the scale- and flavor-number-dependent strong coupling α_s . This dynamical treatment of the top quark MSR scheme resums important large logarithms via R -evolution as shown in Eq. (22). The reader should note that the RHS of the R -evolution equation is linear in R , which differs from the common logarithmic renormalization group equations. This linear evolution is an essential aspect of properly treating the physical mass effects that govern the resonance/close-to-mass-shell dynamics of heavy colored particles.

In our numerical analysis we use $\overline{m}_t^{(6)} = \overline{m}_t^{(6)}(\overline{m}_t^{(6)})$ as the standard reference mass value, which we quote as our main input and from which we then calculate the MSR or $\overline{\text{MS}}$ masses at the respective scales needed within the factorization formula in Eq. (3). For the flavor-number-dependent strong coupling $\alpha_s^{(5,6)}(\mu)$ we always use matching at 4 loops [92–94] and running at 4 loops [95,96], where the flavor matching is carried out at $\mu = \overline{m}_t^{(6)}$.

We finally note that the analytic properties of the pole mass $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon in terms of knowledge on the large-order behavior of the perturbation series obtained in the pole mass scheme are by now very well understood; see, e.g., Refs. [34,50] for the case where all quarks except for the top are assumed massless and Ref. [35] where finite bottom and charm masses are included systematically.⁸ One of the most interesting observations in this context is that the large-order asymptotic behavior (for some unknown reason) universally sets in at $\mathcal{O}(\alpha_s)$ and is already well saturated at $\mathcal{O}(\alpha_s^2)$, which is one of the reasons why the pole mass renormalon has received significant attention in the literature. Another useful (but also confusing) consequence of this fact is that it is very easy to devise different types of low-scale short-distance masses from either physical [25,69,97–101] or conceptual considerations [33–35,102]. All are—as long as the resolution scale R is assigned appropriate values—similarly effective in minimizing mass-related QCD corrections and stabilizing the perturbation series already at $\mathcal{O}(\alpha_s^{1,2})$. The MSR mass $m_t^{\text{MSR},(5)}(R)$ provides a unifying concept to connect all

⁷Here we employ the scheme that was called “natural MSR” mass in Ref. [34].

⁸This knowledge implies a more precise understanding of the size of the renormalon ambiguity, but not that the ambiguity itself is eliminated.

low-scale short-distance masses with the $\overline{\text{MS}}$ mass via its renormalization group equation given in Eq. (22).

IV. SOFT GAP SUBTRACTION

The $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon of the soft function $\hat{S}_{\tau_2}^{(5)}$ has large-order properties very similar to the $\mathcal{O}(\Lambda_{\text{QCD}})$ pole mass renormalon in the bHQET jet function. It is known to differ from the latter just due to a different normalization [21] and the sign difference already mentioned at the end of Sec. II. However, compared to the pole mass renormalon, the soft function $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon is harder to pinpoint quantitatively at low orders due to the soft function anomalous dimension and hadron mass effects [22,103]. Therefore, the soft function renormalon does not seem to exhibit the same universality as the pole mass renormalon (even though it is phenomenologically equally relevant). Apart from e^+e^- event shapes [104], it is still largely unknown whether or in which way the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon in $\hat{S}_{\tau_2}^{(5)}$ appears in a universal manner in soft functions relevant for collider observables which are linearly sensitive to large-angle soft radiation.

It has so far been the practice to define the renormalon subtraction series $\bar{\delta}(R_s)$ in Eq. (17) directly from the soft function; see Refs. [21,84] for such approaches. In our analysis we employ a modified version of the prescription suggested in Ref. [84]. It is based on the Fourier transform of the renormalon-subtracted partonic soft function

$$\begin{aligned} \tilde{S}_{\tau_2}^{(5)}(y, \bar{\delta}(R_s), \mu_S) &= \int d\ell e^{-iy\ell} \hat{S}_{\tau_2}^{(5)}(\ell, \bar{\delta}(R_s), \mu_S) \\ &= \tilde{S}_{\tau_2}^{(5)}(y, \bar{\delta} = 0, \mu_S) e^{-2i\bar{\delta}(R_s)y}, \end{aligned} \quad (23)$$

where the gap subtraction series $\bar{\delta}(R_s)$ is factored into the exponential factor shown in the second line. It is therefore possible to define an expression for $\bar{\delta}(R_s)$ which cancels the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon of the soft function by the condition

$$\begin{aligned} \bar{\delta}(R_s) &\equiv \frac{R_s}{2} \log \left[\tilde{S}_{\tau_2}^{(5)} \left(\frac{1}{iR_s}, \bar{\delta} = 0, R_s \right) \right] \\ &= \frac{R_s}{2} \sum_{i=1} \left[\frac{\alpha_s^{(5)}(R_s)}{4\pi} \right]^i \sum_{j=0}^{i+1} s_{ij} \gamma_E^j, \end{aligned} \quad (24)$$

where s_{ij} are coefficients of the fixed-order series expansion of $\log[\tilde{S}_{\tau_2}^{(5)}(y, \mu)]$ shown below explicitly in Eq. (28). When the gap subtraction series $\bar{\delta}(R_s)$ is used in the factorization theorem, it is crucial that the renormalization scale of the strong coupling $\alpha_s(R_s)$ is reexpressed in terms of $\alpha_s(\mu_S)$, the coupling used in the series for the soft function, to ensure a systematic order-by-order cancellation of the renormalon. This is detailed in Appendix E.

This definition of the gap subtraction can be contrasted with the one used in Ref. [84] where the subtraction series

was instead related to a derivative of the soft function logarithm:

$$\bar{\delta}^{\text{Ref.}[84]}(\mu_S, R_s) \equiv \frac{R_s e^{\gamma_E} d \log \tilde{S}_{\tau_2}^{(5)}(y, \mu_S)}{2 d \log(iy)} \Big|_{iy e^{\gamma_E} = \frac{1}{R_s}}. \quad (25)$$

In this definition, the scale of the strong coupling is μ_S by construction, and the gap subtraction inherits a nontrivial anomalous dimension in μ_S from the soft function. In Appendix E we describe a set of generic gap subtraction schemes that include Eqs. (24) and (25) as special cases.

While both definitions in Eqs. (24) and (25) are perfectly viable subtraction schemes (i.e., equally effective at asymptotic large orders), the series $\bar{\delta}^{\text{Ref.}[84]}$ in Eq. (25) is numerically zero at $\mathcal{O}(\alpha_s)$ for $\mu_S = R_s$ because the one-loop noncusp anomalous dimension vanishes, $\gamma_0^{S_r} = 0$. This necessitates choosing R_s strictly below μ_S in the peak region to reduce the size of the $\mathcal{O}(\alpha_s)$ correction. This can, however, be problematic when considering $\mathcal{O}(\alpha_s^2)$ corrections because in the peak region the soft scale $\mu_S \sim \hat{s}_\tau/q$ is already parametrically smaller than the top quark width such that setting $R_s > \mu_S$ can lead to instabilities. On the other hand, for the gap subtraction definition in Eqs. (24), we have $s_{10} = 15.053$ at one loop which, in addition to being nonzero, is also numerically sizable allowing for the implementation of an effective gap subtraction for the more natural setting $R_s = \mu_S$. We will therefore adopt the gap subtraction scheme defined in Eq. (24) in our analysis.

The R -evolution of the gap parameter corresponding to the scheme defined by $\bar{\delta}$ in Eq. (24) is given by

$$\bar{\Delta}(R_1) - \bar{\Delta}(R_0) = - \sum_{n=0}^{\infty} \bar{\gamma}_n^R \int_{R_0}^{R_1} dR \left[\frac{\alpha_s^{(5)}(R)}{4\pi} \right]^{n+1}, \quad (26)$$

where the anomalous dimension coefficients $\bar{\gamma}_n^R$ are derived from the fixed-order coefficients s_{ij} in Eq. (24) following the steps laid out in Appendix F.

In our numerical analysis we take the first moment value for $\Omega_1^{\text{Ref.}[84]}$ (2 GeV) obtained in Ref. [91] as the input and determine from it $\Omega_1(2 \text{ GeV})$ in the gap scheme defined in Eq. (24) as well as $\bar{\Omega}_1$ for the case of no gap subtraction. For the no-gap case we use $\Delta = 0.1 \text{ GeV}$ as the input value of the gap parameter in Eq. (15) which fixes the analytic form of the shape function $F(k)$. From this we calculate $\bar{\Delta}(R_s = 2 \text{ GeV})$ in the scheme of Eq. (24) using Eq. (16). This input then unambiguously fixes the value of $\bar{\Delta}(R_s)$ at any scale R_s using the R -evolution equation (26), thus determining the form of the gap-subtracted shape function $F(k - 2\bar{\Delta}(R_s))$ in Eq. (20). For the R -evolution in Eq. (26) we employ 2-loop precision [70,71,84]. For $F(k)$ we adopt the parametrization given in Ref. [87]; see Appendix G for explicit expressions. We note that this approach implies that the form of the shape function $F(k - 2\bar{\Delta}(R_s))$ entering the

factorization theorem shown in Eq. (3) depends dynamically on the value of the physical 2-jettiness variable τ_2 . In the peak region this dependence is, however, quite weak because the scale μ_S saturates; see Sec. VI.

V. N³LL RESUMMED CROSS SECTION

For the numerical evaluation of the factorization formula in Eq. (3), we find it convenient to work in the position (Fourier) space where the convolutions involving the bHQET jet and soft functions and their respective renormalization evolution factors become simple products. (The results in this section could equally well be expressed using the Laplace transform.)

A. Stable-top cross section without renormalon subtractions

We first discuss the stable top quark cross section for the case of having neither soft function gap nor mass subtractions. The Fourier transforms of the stable top 2-jettiness bHQET jet and soft functions are defined via

$$\begin{aligned} J_{B,\tau_2}^{(5)}(\hat{s}_\tau, \mu) &= \int \frac{dx}{2\pi} e^{ix\hat{s}_\tau} \tilde{J}_{B,\tau_2}^{(5)}(x, \mu), \\ \hat{S}_{\tau_2}^{(5)}(\ell^+, \mu) &= \int \frac{dy}{2\pi} e^{iy\ell^+} \tilde{S}_{\tau_2}^{(5)}(y, \mu), \end{aligned} \quad (27)$$

where for brevity we have dropped the zero arguments $\delta m = 0$, $\bar{\delta} = 0$, and $\Gamma_i = 0$. Thus, the position space jet and soft functions have the form

$$\begin{aligned} m_t^2 \tilde{J}_{B,\tau_2}^{(5)}(x, \mu_B) &= \exp \left\{ \sum_{i=1} \left[\frac{\alpha_s^{(5)}(\mu_B)}{4\pi} \right]^i \sum_{j=0}^{i+1} b_{ij} \log^j(i e^{\gamma_E} x \mu_B) \right\}, \\ \tilde{S}_{\tau_2}^{(5)}(y, \mu_S) &= \exp \left\{ \sum_{i=1} \left[\frac{\alpha_s^{(5)}(\mu_S)}{4\pi} \right]^i \sum_{j=0}^{i+1} s_{ij} \log^j(i e^{\gamma_E} y \mu_S) \right\}, \end{aligned} \quad (28)$$

such that Eq. (3) can be written as

$$\begin{aligned} \frac{d\hat{\sigma}^{\Gamma_i \rightarrow 0}}{d\tau_2}(\hat{s}_\tau) &= \sigma_0 m_t Q^2 \int \frac{dx}{2\pi} e^{ix\hat{s}_\tau} \\ &\times H_{\text{evol}}^{(5,6)}(Q, m_t, Q, \mu; \mu_H, \mu_m) e^{K_B^{(5)}(\mu, \mu_B) + K_S^{(5)}(\mu, \mu_B)} \\ &\times (i e^{\gamma_E} x \mu_B)^{\omega_B^{(5)}(\mu, \mu_B)} (i e^{\gamma_E} Q x \mu_S)^{\omega_S^{(5)}(\mu, \mu_S)} \\ &\times \tilde{J}_{B,\tau_2}^{(5)}(x, \mu_B) \tilde{S}_{\tau_2}^{(5)}(Qx, \mu_S), \end{aligned} \quad (29)$$

where due the convolution in Eq. (3) the soft function in position space is evaluated at $y = Qx$ and the evolution kernels $K_i^{(n_f)}$ and $\omega_i^{(n_f)}$ are defined in Appendix A. The RG-evolved hard factor in Eq. (29) is given by

$$\begin{aligned} H_{\text{evol}}^{(5,6)}(Q, m_t, Q, \mu; \mu_H, \mu_m) &\equiv H_Q^{(6)}(Q, \mu_H) H_m^{(6)}(m_t, Q, \mu_m) e^{K_{HQ}^{(6)}(\mu_m, \mu_H) + K_v^{(5)}(\mu_m, \mu)} \\ &\times \left(\frac{\mu_H}{Q} \right)^{\omega_{HQ}^{(6)}(\mu_m, \mu_H)} Q^{-\omega_v^{(5)}(\mu_m, \mu)}. \end{aligned} \quad (30)$$

Here μ is the common final renormalization scale of all of the RG evolution factors. Taking the inverse Fourier transform back to distribution space, we find

$$\frac{d\hat{\sigma}^{\Gamma_i \rightarrow 0}}{d\tau_2}(\hat{s}_\tau) = \frac{d\hat{\sigma}^{(0)}(\hat{s}_\tau, \partial_\Omega)}{d\tau_2} \frac{e^{\gamma_E \Omega}}{\Gamma(-\Omega)} \Big|_{\Omega = \tilde{\omega}^{(5)}(\mu_S, \mu_B)}, \quad (31)$$

where for later convenience we have defined the following function of the derivative operator ∂_Ω :

$$\begin{aligned} \frac{d\hat{\sigma}^{(0)}}{d\tau_2}(\hat{s}_\tau, \partial_\Omega) &\equiv \sigma_0 \frac{m_t Q^2}{\hat{s}_\tau} H_{\text{evol}}^{(5,6)}(Q, m_t, Q, \mu; \mu_H, \mu_m) \\ &\times e^{K_B^{(5)}(\mu, \mu_B) + K_S^{(5)}(\mu, \mu_B)} \left(\frac{\mu_B}{\hat{s}_\tau} \right)^{\omega_B^{(5)}(\mu, \mu_B)} \left(\frac{Q\mu_S}{\hat{s}_\tau} \right)^{\omega_S^{(5)}(\mu, \mu_S)} \\ &\times \tilde{J}_{B,\tau_2}^{(5)} \left[\partial_\Omega + \log \left(\frac{\mu_B}{\hat{s}_\tau} \right) \right] \tilde{S}_{\tau_2}^{(5)} \left[\partial_\Omega + \log \left(\frac{Q\mu_S}{\hat{s}_\tau} \right) \right]. \end{aligned} \quad (32)$$

It acts on the function of Ω shown in Eq. (31), and the outcome is evaluated at the following μ -independent evolution kernel between the bHQET jet and soft scales:

$$\tilde{\omega}^{(5)}(\mu_S, \mu_B) \equiv \omega_S^{(5)}(\mu, \mu_S) + \omega_B^{(5)}(\mu, \mu_B). \quad (33)$$

We note that since the jet scale is always above the soft scale, one has $\tilde{\omega}^{(5)} < 0$. In Eq. (32) the arguments of the position-space jet and soft functions are understood to replace the corresponding logarithms shown in Eq. (28). For the sake of brevity we have suppressed the arguments μ_B and μ_S that appear in the running coupling, as shown in Eq. (28). The meaning of the superscript “(0)” on $\hat{\sigma}$ will be clarified below in Eq. (39). We also note that the dependence of the result in Eq. (32) on \hat{s}_τ is defined in terms of rational power plus distributions which have support for $\hat{s}_\tau \geq 0$ in the case of stable top quarks [26]. To account for the fixed-order corrections contained in the product of the functions $H_Q^{(6)}$, $H_m^{(6)}$, $J_{B,\tau_2}^{(5)}$, and $\tilde{S}_{\tau_2}^{(5)}$ at N^kLO we expand their product strictly to $\mathcal{O}(\alpha_s^k)$. The relevant formulas for the evolution kernels and their Fourier transforms are presented in Appendix A. We collect the numerical results for the anomalous dimensions in Appendix B and the fixed-order expressions for all of the factorization functions $H_Q^{(6)}$, $H_m^{(6)}$, $J_{B,\tau_2}^{(5)}$, and $\tilde{S}_{\tau_2}^{(5)}$ up to NNLO in Appendix C.

TABLE I. Loop corrections required for specified orders. In the last two columns γ_R and δ refer to either soft or MSR-mass subtractions.

	Cusp	Noncusp	Matching	$\beta[\alpha_s]$	γ_R	δ
LL	1	...	Tree	1
NLL	2	1	Tree	2	1	...
N ² LL	3	2	1	3	2	1
N ³ LL	4	3	2	4	3	2
NLL'	2	1	1	2	1	1
N ² LL'	3	2	2	3	2	2

Finally, since we carry out resummation at the level of the differential cross section, when implementing the cross section at $N^k\text{LL} + \mathcal{O}(\alpha_s^{k-1})$ accuracy for $k \geq 1$ (referred to as “unprimed” orders), we explicitly incorporate the $\mathcal{O}(\alpha_s^k)$ plus-function boundary condition [105] in order to correctly sum up logarithms that are counted as $N^k\text{LL}$ in the exponent of the cumulative distribution. This amounts to including the *single logarithmic* terms appearing at $\mathcal{O}(\alpha_s^k)$ in the jet and soft functions. For the “primed” orders $N^k\text{LL}'$, or equivalently, $N^k\text{LL} + \mathcal{O}(\alpha_s^k)$ accuracy, this is not necessary as the $\mathcal{O}(\alpha_s^k)$ fixed-order matching already includes this single logarithmic term. The loop order of the theoretical ingredients for the primed and unprimed orders are summarized in Table I. We refer to Ref. [91] for further details on primed and unprimed orders.

B. Renormalon subtractions

We now describe how the renormalon subtractions with respect to the top quark mass and soft function gap are to be included starting from the unsubtracted stable top cross section in Eq. (31). First we recall that the δm dependence in the bHQET jet function $J_{B,\tau_2}^{(5)}$ results from reexpressing the pole mass m_t^{pole} contained in \hat{s}_τ in terms of the MSR mass. From Eq. (5) we have

$$\begin{aligned} \hat{s}_\tau &= \frac{Q^2\tau_2 - 2[m_t^{\text{MSR}}(R) + \delta m(R)]^2}{m_t^{\text{MSR}}(R) + \delta m(R)} \\ &= \hat{s}_\tau^{\text{MSR}}(R) - 4\delta m(R) + \mathcal{O}\left(\frac{\alpha_s\Gamma_t}{m_t}\right), \end{aligned} \quad (34)$$

where

$$\hat{s}_\tau^{\text{MSR}}(R) \equiv \frac{Q^2\tau_2 - 2[m_t^{\text{MSR}}(R)]^2}{m_t^{\text{MSR}}(R)}. \quad (35)$$

The two terms shown in the second line of Eq. (34) represent those to be accounted for in Eq. (3) since power-suppressed contributions in the peak region must be systematically dropped for consistency. For the term $\delta m(R)$, which contains the pole mass renormalon ambiguity, this is particularly important to achieve the order-by-order cancellation of the

pole mass renormalon. Thus, the stable-top bHQET jet function using Eq. (13) in position space can be expressed up to NNLO as

$$\begin{aligned} \tilde{J}_{B,\tau_2}^{(5)}(x, \delta m, \mu_B) &= \int d\hat{s}' e^{-ix\hat{s}'} J_{B,\tau_2}^{(5)}(\hat{s}' - 4\delta m, \mu_B) \\ &= \left[1 - (ix)4\delta m + (ix)^2 \frac{(4\delta m)^2}{2!} \right] J_{B,\tau_2}^{(5)}(x, \mu_B), \end{aligned} \quad (36)$$

where we have dropped the zero argument $\Gamma_t = 0$ in the jet function and the argument R in the mass subtraction $\delta m(R)$ for simplicity. We have kept only terms at most quadratic in $\delta m(R)$ so that the pole mass renormalon can be consistently canceled to $\mathcal{O}(\alpha_s^2)$. Likewise, the soft gap subtraction can be incorporated using Eq. (23) such that

$$\begin{aligned} \tilde{S}_{\tau_2}^{(5)}(qx, \bar{\delta}, \mu_S) &= e^{-2iq\bar{\delta}x} \tilde{S}_{\tau_2}^{(5)}(qx, \mu_S) \\ &= \left[1 - (ix)2q\bar{\delta} + (ix)^2 \frac{(2q\bar{\delta})^2}{2!} \right] \tilde{S}_{\tau_2}^{(5)}(qx, \mu_S), \end{aligned} \quad (37)$$

where we dropped the argument R_s in the gap subtraction $\bar{\delta}(R_s)$. Including the subtraction terms in Eqs. (36) and (37) and *strictly* expanding to $\mathcal{O}(\alpha_s^2)$, we arrive at the following expression for the renormalon-subtracted cross section for stable top quarks:

$$\begin{aligned} \frac{d\hat{\sigma}^{\Gamma_t=0}}{d\tau_2}(\hat{s}_\tau, \delta m(R), \bar{\delta}(R_s)) &= \sum_{n=0}^2 \frac{d\hat{\sigma}^{(n)}}{d\tau_2}(\hat{s}_\tau^{\text{MSR}}(R), R_s, \partial_\Omega) \frac{e^{\gamma_E\Omega}}{\Gamma(-\Omega)} \Big|_{\Omega=\bar{\omega}^{(5)}(\mu_S, \mu_B)+n}. \end{aligned} \quad (38)$$

Here we have

$$\begin{aligned} \frac{d\hat{\sigma}^{(n)}}{d\tau_2}(\hat{s}_\tau^{\text{MSR}}(R), R_s, \partial_\Omega) &= \frac{(-1)^n}{n!} \left[\frac{\delta_{\text{tot}}(R, R_s)}{e^{\gamma_E\hat{s}_\tau^{\text{MSR}}(R)}} \right]^n \frac{d\hat{\sigma}^{(0)}(\hat{s}_\tau^{\text{MSR}}(R), \partial_\Omega)}{d\tau_2}, \end{aligned} \quad (39)$$

where the total subtraction series δ_{tot} has the form

$$\delta_{\text{tot}}(R, R_s) \equiv 4\delta_m(R) + 2q\bar{\delta}(R_s), \quad (40)$$

and $d\hat{\sigma}^{(0)}/d\tau_2$ is given in Eq. (32). It is essential to consistently drop terms of $\mathcal{O}(\alpha_s^3)$ and higher (concerning the fixed-order corrections in the hard, jet, and soft functions as well as the renormalon subtractions) in the product in Eq. (39). Note that the derivatives in the sum over n in Eq. (38) are evaluated at $\Omega = \bar{\omega} + n$. For simplicity we will

display the functions $\hat{s}_\tau^{\text{MSR}}$, δ_m , $\bar{\delta}$, and δ_{tot} without their arguments R and R_s below.

C. Including the top width

From Eq. (11) we see that the cross section for unstable top quarks involves an additional Breit-Wigner convolution such that⁹

$$\begin{aligned} \frac{d\hat{\sigma}}{d\tau_2}(\hat{s}_\tau^{\text{MSR}}, \delta m, \bar{\delta}) &= \int_0^\infty \frac{d\hat{s}'}{\pi} \frac{2\Gamma_t}{(2\Gamma_t)^2 + (\hat{s}_\tau^{\text{MSR}} - \hat{s}')^2} \\ &\times \frac{d\hat{\sigma}^{\Gamma_t=0}}{d\tau_2}(\hat{s}', \delta m, \bar{\delta}), \end{aligned} \quad (41)$$

where as a result of the stable top cross section, the integration is bounded below by $\hat{s}' = 0$. However, $\hat{s}_\tau^{\text{MSR}}$ (or \hat{s}_τ in the pole mass scheme) for the unstable top 2-jettiness cross section can be negative as well. To incorporate this convolution in Eq. (38) we first note that the cross section is proportional to $1/(\hat{s}')^{\tilde{\omega}^{(5)}+1}$, as can be seen from Eq. (32) using Eq. (33). This factor can be brought to the right of the ∂_Ω derivatives using the identity

$$\begin{aligned} \frac{1}{(\hat{s}')^{\Omega+1}} \left[\partial_\Omega + \log\left(\frac{\mu_B}{\hat{s}'}\right) \right]^n \\ = \frac{1}{(\sqrt{\hat{s}_\tau^2 + (2\Gamma_t)^2})^{\Omega+1}} \\ \times \left[\partial_\Omega + \log\left(\frac{\mu_B}{\sqrt{\hat{s}_\tau^2 + (2\Gamma_t)^2}}\right) \right]^n \left(\frac{\sqrt{\hat{s}_\tau^2 + (2\Gamma_t)^2}}{\hat{s}'} \right)^{\Omega+1}, \end{aligned} \quad (42)$$

where we remind the reader that \hat{s}' is the integration variable in Eq. (41) whereas \hat{s}_τ (or $\hat{s}_\tau^{\text{MSR}}$ in the MSR mass scheme) is related to the τ_2 measurement as defined in Eq. (5). The analogous relation also holds for the ∂_Ω derivatives associated with the soft function. Hence, we need to evaluate the convolution

$$\begin{aligned} \mathcal{I}(\Omega, \hat{s}_\tau, 2\Gamma_t) &\equiv \frac{1}{\Gamma(-\Omega)} \int_0^\infty \frac{d\hat{s}'}{\pi} \frac{2\Gamma_t}{(2\Gamma_t)^2 + (\hat{s}_\tau - \hat{s}')^2} \\ &\times \left(\frac{\sqrt{\hat{s}_\tau^2 + (2\Gamma_t)^2}}{\hat{s}'} \right)^{\Omega+1} \\ &= \frac{\phi(\frac{\hat{s}_\tau}{2\Gamma_t})\Gamma(2+\Omega)}{\Gamma(1+(1+\Omega)\phi)\Gamma(1-(1+\Omega)\phi)}, \end{aligned} \quad (43)$$

which has a smooth $\Omega \rightarrow 0$ limit, and where we have defined

⁹The analytical results shown in this section were first derived in Ref. [27].

$$\phi(x) \equiv \frac{1}{2} + \frac{1}{\pi} \arctan(x). \quad (44)$$

In the limit $\Gamma_t \rightarrow 0$ one smoothly recovers the stable-top results. For $\Gamma_t = 0$ we find that $\mathcal{I} = 0$ when $\hat{s}_\tau < 0$. Using these expressions in Eqs. (38) and (41), we now arrive at the final expression for the unstable-top cross section with renormalon subtractions:

$$\begin{aligned} \frac{d\hat{\sigma}}{d\tau_2}(\hat{s}_\tau^{\text{MSR}}, \delta m, \bar{\delta}) &= \sum_{n=0}^2 \frac{d\hat{\sigma}^{(n)}}{d\tau_2} \left(\sqrt{(\hat{s}_\tau^{\text{MSR}})^2 + (2\Gamma_t)^2}, \partial_\Omega \right) \\ &\times e^{\gamma_E \Omega} \mathcal{I}(\Omega, \hat{s}_\tau, 2\Gamma_t) \Big|_{\Omega=\tilde{\omega}^{(5)}(\mu_s, \mu_B)+n}. \end{aligned} \quad (45)$$

Finally, the hadron-level cross section is obtained from the partonic cross section via convolution with the nonperturbative shape function:

$$\frac{d\sigma}{d\tau_2}(\tau_2) = \int dk \frac{d\hat{\sigma}}{d\tau_2}(\hat{s}_\tau^{\text{MSR}}(R) - \rho k) F(k - 2\bar{\Delta}(R_s)). \quad (46)$$

VI. PROFILE FUNCTIONS

To properly sum large logarithms we use τ_2 -dependent renormalization scales $\mu_i(\tau_2)$, $R(\tau_2)$, and $R_s(\tau_2)$, called profile functions [87,91]. They have canonical scaling in the resummation regions, and freeze at a perturbative scale in the resonance region to avoid the breakdown of perturbation theory for anomalous dimensions. In the far tail region, they become equal to the hard scale to reproduce the fixed-order perturbative expansion with a common scale μ . They are expressed in terms of seven parameters which can be varied to estimate perturbative uncertainties. Following Refs. [27,106], we employ a natural generalization of the profile functions devised for massless event shapes in Ref. [19], to which they reduce in the massless limit.

The strategy to estimate perturbative uncertainties involves varying all of the profile functions up and down by at most a factor of 2 and 1/2, respectively, including a shape-dependent variation in the jet scale, as well as varying the value at which the soft scale freezes in the nonperturbative region. We show bands for the latter two variations in Fig. 1, and indicate the factor of 2 variations by arrows. We scan over these profile variations by generating a sample of 500 profiles where all their parameters are varied simultaneously with random choices within their allowed ranges (see Ref. [91] for details on this general approach). The concrete form of the profile functions and how their parameters are varied are given below. The total uncertainty is determined by the envelope of the resulting cross sections.

For the hard and mass matching scales we use τ_2 -independent functions $\mu_H = e_H Q$ and $\mu_m = \sqrt{e_H} m_t$, respectively, which depend on the center-of-mass energy and top quark mass, as well as on a free parameter e_H ,

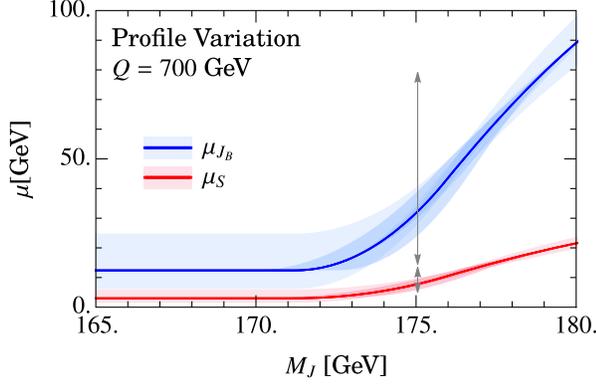


FIG. 1. Examples of variation of the jet and soft scales to estimate perturbative uncertainty. The e_{J_0} parameter is related to a “trumpet” variation in the jet scale that turns off in the fixed-order region. For the soft scale we vary μ_0 , the perturbative value that the soft scale freezes to in the nonperturbative region.

which has the default value 1 and in our random scan is varied in the interval $[0.5, 2]$. The variations of the two matching scales are correlated to keep the correct hierarchies. For the top quark mass parameter m_t , which appears in the matching scale μ_m as well as in other variables entering the profile functions discussed below, we adopt the numerical value $m_t = m^{\text{MSR}}(R = 5 \text{ GeV})$.

Our τ_2 -dependent profile functions are implemented through the soft scale $\mu_S(\tau_2)$, which in the peak region is parametrized with the following piecewise function:

$$\mu_S = \begin{cases} e_{\text{np}}\mu_0, & \tau_2^{\text{min}} \leq \tau_2 \leq t_0, \\ \zeta[e_{\text{np}}\mu_0, \mu_S(\tau_2 > t_1), t_0, t_1, \tau_2], & t_0 < \tau_2 < t_1, \\ \left(1 + \frac{n_s e_{\text{slope}}}{n_s + \hat{m}_t - \tau_2^{\text{min}}}\right) r_s \mu_H(\tau_2 - \tau_2^{\text{min}}), & t_1 \leq \tau_2 < t_2, \end{cases} \quad (47)$$

where $\tau_2^{\text{min}} = 2\hat{m}_t^2$ with $\hat{m}_t \equiv m_t/Q$. We refer to the three corresponding intervals as the “nonperturbative,” “bHQET,” and “SCET-resummation” regions. The function $\zeta[f_a(\tau_2), f_b(\tau_2), t_a, t_b, \tau_2]$ smoothly connects any two linear functions $f_{a,b}(\tau_2)$ that end/begin at the points $t_{a,b}$. This is achieved by means of two quadratic polynomials of τ_2 smoothly joined at $\tau_2 = (t_a + t_b)/2$, where the explicit formula can be found in Ref. [19]. The parameter n_s has the default value 0.5 and is varied in the range $|n_s - 0.5| \leq 0.025$. Its effect in the peak region is relatively mild. The default slope in the SCET-resummation region is set by $r_s = 2$ and using the default value 0 for the variable e_{slope} . Slope variations are implemented by varying e_{slope} in the interval $[1/1.13 - 1, 1.13 - 1]$. Note that the rescaling factor to the left of r_s approaches 1 in the massless limit. In a similar way, the parameter affecting the flat nonperturbative region is e_{np} . Its default value is 1 and it is varied in the interval $[1/\sqrt{2}, \sqrt{2}]$ with $\mu_0 = 3 \text{ GeV}$. The values of

the transition points $t_{0,1}(m_t, Q)$ depend on the mass and center-of-mass energy and have the form

$$\begin{aligned} t_0 &= \frac{2}{(Q/1 \text{ GeV})} + \frac{d_0}{(Q/1 \text{ GeV})^{0.5}} + \tau_2^{\text{min}}, \\ t_1 &= \frac{2.25}{(Q/1 \text{ GeV})^{0.75}} + \frac{d_1}{(Q/1 \text{ GeV})^{0.5}} + \tau_2^{\text{min}}, \\ t_2 &= n_2 + \hat{m}_t, \end{aligned} \quad (48)$$

with $|d_{0,1}| \leq 0.05$ and $|n_2 - 0.25| \leq 0.025$. Their default values are $d_{0,1} = 0.05$ and $n_2 = 0.25$. For the energies and masses considered in this article one has $t_0 < t_1 < t_2$. For the jet scale profile function $\mu_J(\tau_2)$ we first define $\tilde{\mu}_J(\tau_2) = \sqrt{e_H} \mu_S(\tau_2)/\hat{m}_t$ and $t_s = n_s + \hat{m}_t$, and then use the piecewise function

$$\mu_J = \begin{cases} [1 + \tilde{e}_J(t_0 - t_s)^2] \tilde{\mu}_J(t_0), & \tau_2^{\text{min}} \leq \tau_2 \leq t_0, \\ \zeta[\mu_J(t_0), \mu_J(\tau_2 > t_1), t_0, t_1, \tau_2], & t_0 < \tau_2 < t_1, \\ [1 + \tilde{e}_J(\tau_2 - t_s)^2] \tilde{\mu}_J(\tau_2), & t_1 \leq \tau_2 < t_2. \end{cases} \quad (49)$$

Here the jet-function parameter \tilde{e}_J is defined with a rescaling factor

$$\tilde{e}_J = e_J \left[\frac{n_s - (t_0 - \tau_2^{\text{min}})}{t_s - t_0} \right]^2, \quad (50)$$

with variations $|e_J| \leq 1.5$ used for assessing uncertainties and the default value $e_J = 0$. For $m_t = 0$ we recover $e_J = \tilde{e}_J$. We set the mass and soft-function renormalon subtraction scales to their respective canonical values: $R_s(\tau_2) = \mu_S(\tau_2)$ and $R(\tau_2) = \mu_J(\tau_2)$.

VII. NUMERICAL ANALYSIS

In this section we present a numerical analysis of the bHQET $N^3\text{LL} + \mathcal{O}(\alpha_s^2)$ 2-jettiness peak region cross section based on the factorization formula of Eq. (3). We remind the reader that this factorization theorem is based on the bHQET limit and does not account for subleading terms related to higher powers of \hat{s}_τ (kinematic power corrections) and m_t/Q (mass power corrections). As was shown in Refs. [27,82,106], these corrections are formally power suppressed in the peak region, but still important numerically for a realistic phenomenological analysis concerning the top mass dependence of the 2-jettiness peak-region line shape. In the following we therefore carry out a generic numerical analysis pointing out important features of the bHQET 2-jettiness peak region cross section at this order related to the convergence of the perturbative series, as well as the impact of the jet and soft function renormalons and the improvement related to their subtractions. A phenomenological analysis aiming for a systematic study of

other sources of theoretical uncertainties is postponed to future work.

For our analysis we have two independent codes to implement all cross sections: one in FORTRAN [107] and one in c++ using SCETlib [108]. Numerical integrations are carried out using either QUADPACK [109] or the GSL library [110].

Before getting to results, some comments on our parametric input are in order. For the subsequent discussion we use $\alpha_s^{(5)}(m_Z) = 0.118$ with $m_Z = 91.1876$ GeV as input for the strong coupling. As input for the top quark mass we take the standard $\overline{\text{MS}}$ mass $\overline{m}_t^{(6)} \equiv \overline{m}_t^{(6)}(\overline{m}_t^{(6)}) = 160$ GeV. The conversion (matching) to the MSR top mass at the scale $R = \overline{m}_t^{(6)}$ is based on the formulas derived in Ref. [34] employed at $\mathcal{O}(\alpha_s^2)$ and given in Eqs. (D4) and (D5). The evolution of the MSR top mass to the renormalon subtraction profile scale R is obtained via Eq. (22) using the $\mathcal{O}(\alpha_s^3)$ R -anomalous dimensions, such that

$$m_t^{\text{MSR}}(R = 2 \text{ GeV}) = 169.537 \text{ GeV}, \quad (51)$$

for the MSR top mass at 2 GeV, which one can interpret as the (renormalon-free) kinematic mass that governs the top-mass dependence of the peak position. The same top mass value is used in the boost parameter ϱ defined in Eq. (4) when the MSR mass scheme is used. For comparison we will also discuss results for the cross section in the pole mass scheme and without gap subtraction. At this point one has to recall that the pole mass is, due to its renormalon ambiguity, an order-dependent concept, where the size of the fixed-order corrections in its relation to a short-distance mass at a given order depends on the renormalization scale of the short-distance mass. Thus, to achieve comparable theoretical predictions employing the MSR and the pole mass schemes (i.e., with peak positions that are compatible), it is essential to apply fixed-order conversion from the MSR to the pole mass at the renormalization scale of the MSR mass that is employed in the peak region of the distribution (which is the part of the distribution that carries the highest top quark mass sensitivity). Furthermore, the order of conversion has to match the fixed-order input used for the theoretical calculation [13,35]. In the peak region of the distribution the appropriate renormalization scale is just the bHQET jet function scale μ_J , around 10 GeV; see Sec. VI. Thus, the proper way to determine the pole mass is to use the fixed-order conversion from $m_t^{\text{MSR}}(10 \text{ GeV})$, as was pointed out in Ref. [27]. Therefore, to determine the pole mass for the following analysis we first determine $m_t^{\text{MSR}}(10 \text{ GeV})$ using Eq. (51) and R evolution, and then apply fixed-order conversion to the pole scheme at the scale $R = 10$ GeV. As we are employing $\mathcal{O}(\alpha_s^2)$ fixed-order corrections to the bHQET jet function at the highest N³LL order of our analysis, this fixed-order conversion

must be carried out at two loops. Using this procedure, we find

$$m_t^{\text{pole}} = 169.718 \text{ GeV}. \quad (52)$$

This input value for the top quark pole mass ensures that at N³LL the peak position in the pole mass scheme is compatible with that obtained in the MSR mass scheme.

Last, in order to fix the form of the nonperturbative model function $F(k)$ [see Eq. (14) and Appendix G], we have to specify values for the first moment Ω_1 . In analogy to the top quark mass, there are also different schemes for Ω_1 related to the definition of the gap subtraction parameter. Thus, the corresponding values for Ω_1 also have to be fixed with some care. Here we aim to adopt values for Ω_1 consistent with the e^+e^- thrust analysis of Ref. [91], using the fact that the same universal soft function that is given in Eq. (14) also appears in the thrust factorization theorem for massless quark production. The value of Ω_1 determined in Ref. [91], $\Omega_1^{\text{Ref.}[84]}(R_s = 2 \text{ GeV}) = 0.323 \pm 0.045$ GeV, was based on the gap subtraction scheme suggested in Ref. [84] [explained in detail in and below Eq. (25)]. For our analysis it needs to be converted to the gap scheme of Eq. (24) adopted in this work, referred to as $\Omega_1(R_s)$, as well as to the unsubtracted $\overline{\text{MS}}$ gap scheme $\overline{\Omega}_1$ (which still contains the soft function renormalon). For these conversions we must use the 2-loop fixed-order formulas in Eqs. (E5) and (G4), respectively, since at the highest N³LL order of our analysis we employ the soft function at $\mathcal{O}(\alpha_s^2)$. This gives

$$\begin{aligned} \Omega_1(R_s = 2 \text{ GeV}) &= 0.739 \text{ GeV}, \\ \overline{\Omega}_1 &= 0.276 \text{ GeV}, \end{aligned} \quad (53)$$

where we convert at the scale $R_s = 2$ GeV, the typical value of the soft function profile in the peak region. Fixing $\Delta = 0.1$ GeV for the unsubtracted gap [see Eq. (15)] and using Eq. (16), we obtain $\overline{\Delta}(R_s = 2 \text{ GeV}) = 0.563$ GeV for the gap term entering Eq. (20). With this choice of parameters we can use the same shape function parameters as employed in Ref. [91], which corresponds to taking $\lambda = 0.349$ GeV and $c_2 = 0.05$ in Eq. (G1). We then use Eq. (26) with $\mathcal{O}(\alpha_s^2)$ running to determine $\overline{\Delta}(R_s)$ at the τ_2 -dependent R_s values required by the gap subtraction profile function.

In Fig. 2 the 2-jettiness differential cross section in the MSR mass scheme with gap subtraction is shown for $Q = 700$ GeV as a function of the inclusive jet mass variable M_J [see Eq. (6)] at all primed and unprimed orders up to N³LL for the default set of profile functions (see Sec. VI). Here the primed orders include all contributions of the unprimed with the addition of the fixed-order matrix elements at one higher order in α_s . Our results at NNLL' and N³LL are new and have not been analyzed before in the literature.

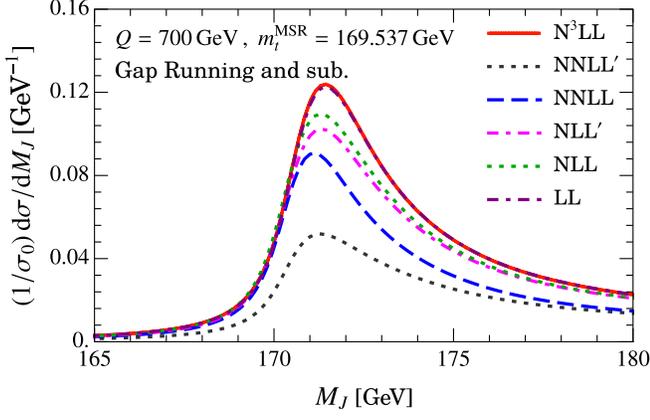


FIG. 2. Perturbative convergence of the 2-jettiness cross section, with all curves normalized with the Born cross section σ_0 . The peak location shows excellent convergence, while the normalization corrections are significant and exhibit slower convergence.

In Fig. 2 all curves are normalized to the Born-level massless total cross section σ_0 . The behavior of the curves at the different orders therefore reflects the effects of the perturbative corrections to the shape as well as to the normalization of the cross section. We see that, apart from sizable normalization corrections which happen to be positive for all subsequent orders, the convergence with respect to the peak location and the shape is excellent. From the difference in the curves obtained from primed and unprimed orders we can also see that the effects on the normalization from higher-order corrections in the renormalization group equations (and thus the resummation of large logarithmic terms) are smaller than those in the fixed-order matrix elements.

In order to analyze the effects of the higher-order corrections to the distribution shape and its order-dependent perturbative uncertainty, it is useful to normalize the curves from the different orders to a common M_J interval. In Fig. 3 the 2-jettiness differential cross sections at $Q = 700$ GeV (upper panels) and $Q = 2000$ GeV (lower panels) are shown in the MSR mass scheme with gap subtractions using default profile functions. The results are normalized to the M_J interval displayed in the respective panels at NLL (green dotted line), NNLL (blue dash-dotted line), and N³LL (red solid line). Primed orders are not displayed to avoid cluttering. We also display uncertainty bands with the corresponding colors at each of these three orders. These bands are derived by determining the upper and lower values of the distributions (for each M_J value) obtained by considering 500 profile functions generated randomly within the profile function parameter ranges given in Sec. VI. To generate the bands, each cross section from a given profile is normalized to the displayed M_J range. The central curves exhibit excellent perturbative convergence for the shape. The width of each band illustrates the size of the perturbative uncertainty, which

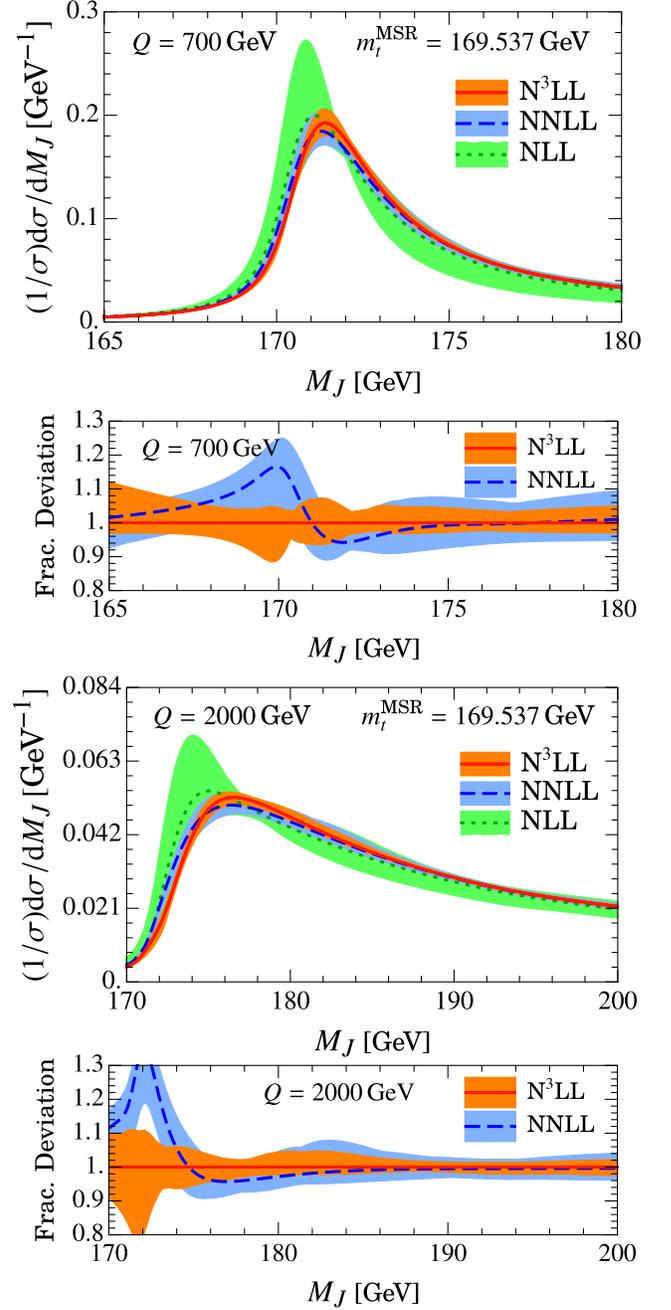


FIG. 3. Perturbative convergence and uncertainty bands for self-normalized cross sections at $Q = 700$ GeV (upper panels) and $Q = 2000$ GeV (lower panels) in the MSR mass scheme and with gap subtractions. All curves are normalized over the displayed ranges. The two smaller panels show the same results at the two highest orders, but as a fractional deviation from the central N³LL result.

nically decreases with increasing order. For better visibility the error bands and lines are displayed once more in the lower parts of each plot showing the fractional deviation from the central N³LL curve. At $Q = 700$ GeV for $M_J \geq 171$ GeV the relative uncertainty in the peak region is $\pm(4-10)\%$ at NNLL and $\pm(3-7)\%$ at N³LL. In contrast,

at $Q = 2000$ GeV for $M_J \geq 175$ GeV the relative uncertainty in the peak region is $\pm(3-8)\%$ at NNLL and $\pm(1-5)\%$ at N³LL.

In Fig. 4 we show for comparison the analogous results for cross sections in which the pole scheme for the top quark mass is employed without gap subtractions. We will return and discuss this figure in more detail below.

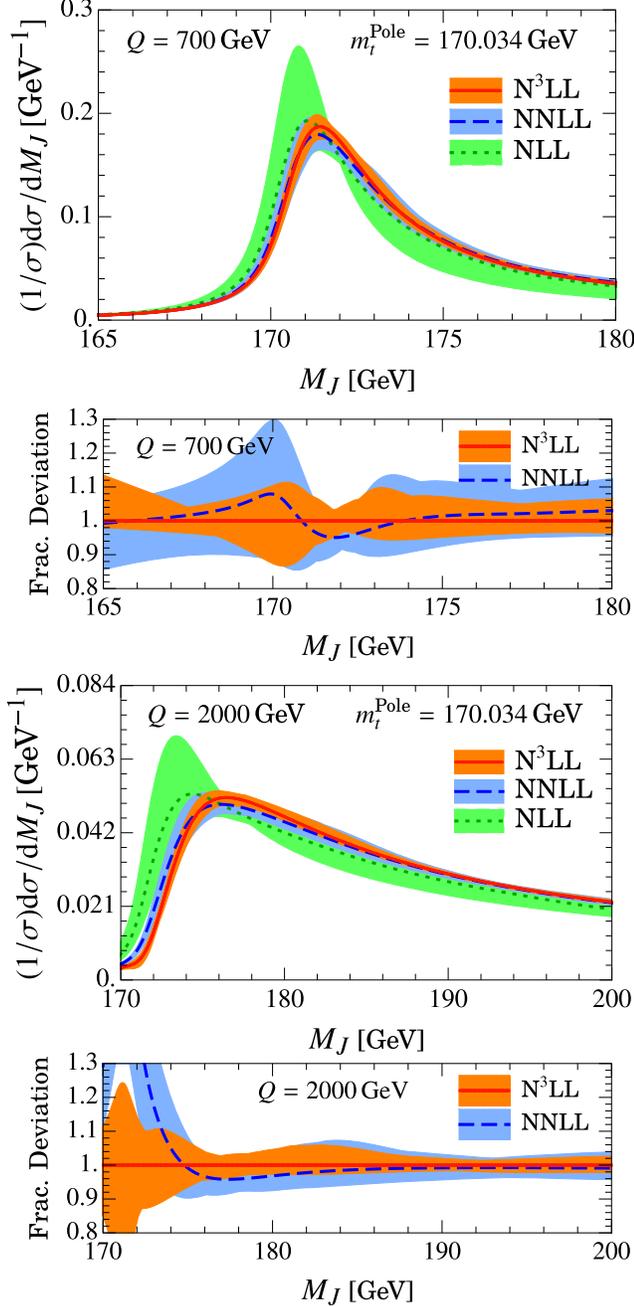


FIG. 4. Perturbative convergence and uncertainty bands for self-normalized cross sections at $Q = 700$ GeV (upper panels) and $Q = 2000$ GeV (lower panels) in the pole mass scheme and without gap subtractions. The two smaller panels show the same results at the two highest orders, but as a fractional deviation from the central N³LL result.

It is instructive to first examine the importance and interplay of the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons contained in the perturbative fixed-order series of the bHQET jet and partonic soft functions. To illustrate the impact of the pole mass renormalon in the bHQET jet function, we display in the upper panel of Fig. 5 the 2-jettiness cross section in the

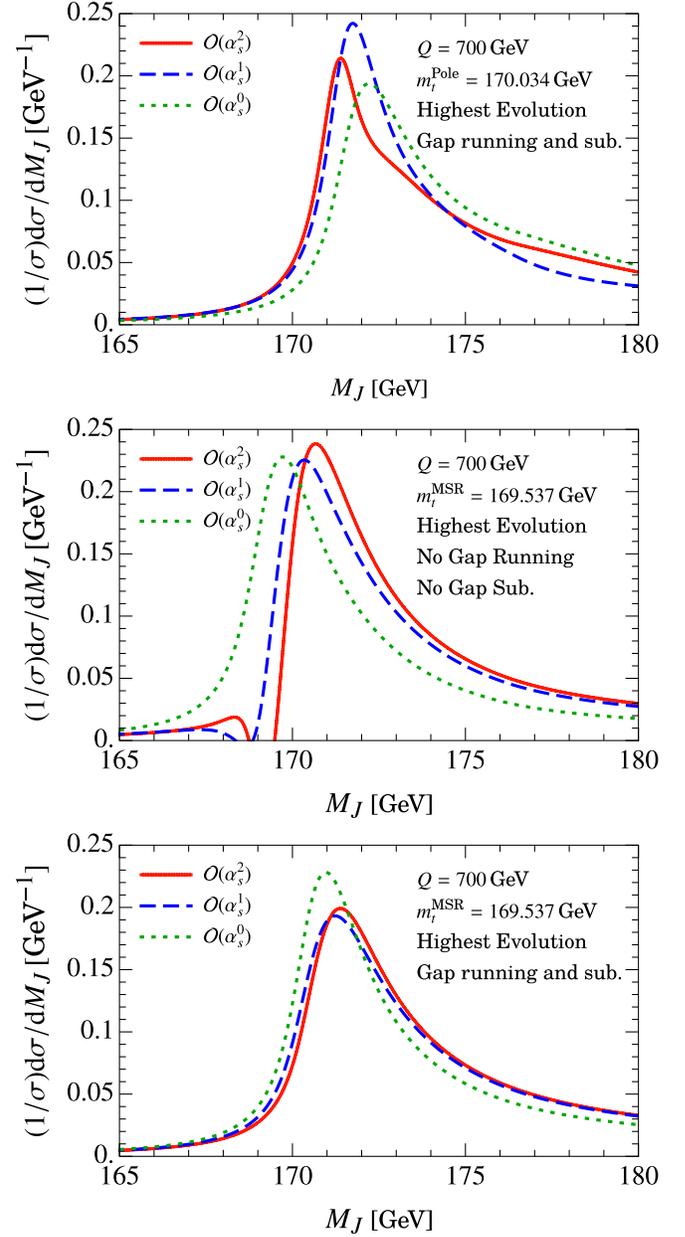


FIG. 5. Analysis of the pole mass and soft function renormalon's impact on the convergence of perturbation theory at $\mathcal{O}(\alpha_s^0)$, $\mathcal{O}(\alpha_s)$, and $\mathcal{O}(\alpha_s^2)$. The top panel uses the jet function in the pole mass scheme, but includes gap subtractions for the soft function. The middle panel does not include gap subtractions, but utilizes the jet function in the MSR mass scheme. The bottom panel employs the MSR mass scheme and gap subtractions, and thus removes the leading renormalons in both the jet and soft functions.

pole mass scheme for $Q = 700$ GeV using the default profile functions, consistently expanding all fixed-order matrix elements entering the factorization theorem (i.e., the bHQET jet function, the soft function and the hard function) to $\mathcal{O}(\alpha_s^0)$ (dotted green), $\mathcal{O}(\alpha_s)$ (dashed blue), and $\mathcal{O}(\alpha_s^2)$ (solid red) and consistently using gap subtractions. Thus, renormalon subtractions associated with the pole mass renormalon are not included, while the soft function renormalon is still removed systematically. Since the renormalization group evolution, which predominantly affects the normalization, does not contain any renormalon effects, we adopt the highest-order N³LL anomalous dimensions for all renormalization-group resummation factors, so that the behavior of the three curves is focused on the pole mass renormalon in the unsubtracted jet function. The curves clearly exhibit the well-known pole mass renormalon problem which causes the peak position to systematically shift towards smaller jet masses with increasing order. At the level of the 2-loop bHQET jet function itself, this behavior was discussed in Ref. [69]. In a fit to data this behavior would correspond to a pole mass value that systematically increases with the perturbative order. This is the known behavior of the perturbative series for the pole mass in terms of a short-distance mass [13,35,50]. Furthermore, the curves show some instabilities in its shape, in particular in the form of the distribution at and above the peak.

In order to illustrate the impact of the soft function renormalon we display in the middle panel of Fig. 5 the 2-jettiness cross section *without gap subtractions*, again for $Q = 700$ GeV using the default profile functions and consistently expanding all fixed-order matrix elements entering the factorization theorem, but this time using the MSR top quark mass scheme (and the standard \overline{MS} mass in the hard function). Here subtractions associated with the soft function renormalon are not included, while the pole mass renormalon is removed systematically. As in the upper panel, we adopt the highest-order N³LL anomalous dimensions for all renormalization-group resummation factors, so that the three curves focus on the behavior due to the soft function renormalon. We see that the soft function renormalon causes the peak position to systematically shift towards larger jet masses with increasing order. In a fit to data this behavior would correspond to an MSR mass value that systematically decreases with the perturbative order. Furthermore, the curves at $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ show considerable shape instabilities in the region below the peak, where the cross section can even become negative. We note that the impact of the soft function renormalon increases with the c.m. energy Q . This dependence on Q arises from the boost factor $\varrho = Q/m_t$ appearing in the convolution integral shown in Eq. (3), which is also manifest in Eq. (40).

Finally, in the lower panel of Fig. 5 we display the corresponding three curves once again, but systematically

accounting for the subtractions associated to both the pole mass and soft function renormalons, by using the MSR top quark mass scheme and gap subtractions, respectively. We now observe very good convergence of the peak position and, furthermore, no instabilities in the shape of the distribution are visible.

The upper and middle panels of Fig. 5 also nicely illustrate the presence of a partial cancellation of the jet and soft function renormalon effects since they have opposite signs. In the combined order-by-order cross sections both corrections thus partially cancel when the pole mass scheme is employed and no gap subtraction is applied for the soft function. Even though this partial cancellation arises between two independent physical effects and should therefore be considered as accidental from a principle point of view, it does undeniably take place in the physical regions of c.m. energies where high-precision extractions of the top mass can be carried out. One may therefore ask the question whether this cancellation may in principle allow for a pole mass determination where the impact of the pole mass renormalon could be tamed or even avoided altogether. At this point we would like to remind the reader that for a top mass determination from data (or MC pseudodata) simultaneous fits of the peak region 2-jettiness distribution for several Q values are needed to disentangle the dependence on the top quark mass and the shape function parameters. So there is a strong degeneracy concerning the dependence on the top quark mass and the shape function parameters, and in particular its first moment Ω_1 . Given that there are strong cancellations between corrections affecting both of these dependences, it can be expected that they degrade the overall precision of such an analysis. Furthermore, since the amount of mutual cancellation between the pole mass and soft function renormalons is Q dependent, it is expected that such fits for theoretical predictions without any renormalon subtractions will exhibit larger theoretical uncertainties compared to those where the pole mass and soft function renormalons are separately and independently subtracted. Such an extensive analysis is, however, beyond the scope of this work.

Furthermore, it should also be pointed out that the shape function appearing in Eq. (3) is universal and appears in the same form also in the factorization theorem for the e^+e^- thrust distribution below the top pair threshold, where precise information on its parameters can be extracted from available e^+e^- data [91] at significantly smaller values of Q . The thrust distribution for massless quark production is sensitive to the same soft function renormalon, but does not have any top mass dependence. Thus, if information on the renormalon-free shape function parameters obtained from e^+e^- thrust data is systematically accounted for, it is unavoidable that renormalon effects must be properly handled for top quark mass determinations from the 2-jettiness distribution.

Finally, it is instructive to also have a closer look at the 2-jettiness cross section without any renormalon subtraction. In Fig. 4 the 2-jettiness differential cross section at $Q = 700$ GeV (upper panels) and $Q = 2000$ GeV (lower panels) in the pole mass scheme and without gap subtractions are shown for the default profile functions. The results are normalized to the M_J interval displayed in the respective panels at NLL (green dotted line), NNLL (blue dash-dotted line), and N^3 LL (red solid line). We also display uncertainty bands with the corresponding colors at each of the three orders. These bands are again derived by determining the upper and lower values of the distributions (for each M_J value) obtained by considering 500 profile functions generated randomly within the profile function parameter ranges. Apart from the fact that neither the pole mass nor the soft function renormalons are subtracted, the setup used for all curves and uncertainty bands in Fig. 4 is precisely the same as the one used for Fig. 3. We see that the perturbative behavior concerning the convergence and the perturbative uncertainties is also good even without any renormalon subtraction. This underlines the partial cancellation of the jet and soft function renormalons. However, a closer inspection shows that the perturbative uncertainty bands are narrower when the subtraction of all renormalons is taken care of systematically. This is visible in the fractional deviation plots, where the N^3 LL renormalon-subtracted predictions in Fig. 3 exhibit an average uncertainty of $\pm 3.8\%$ at $Q = 700$ GeV for $M_J \geq 171$ GeV compared to $\pm 5.5\%$ for the predictions without any renormalon subtraction in Fig. 4. In contrast, for $Q = 2000$ GeV Fig. 3 has an average uncertainty of $\pm 2.4\%$ for $M_J \geq 175$ GeV compared to $\pm 2.9\%$ for the predictions without any renormalon subtraction in Fig. 4.

An interesting aspect of our definition for the jet mass variable M_J [defined using 2-jettiness in Eq. (6)] is that it is normalized in a way such that it can be seen as a direct measure for the top quark mass. Therefore, the behavior of the peak position for the M_J distribution allows us to draw conclusions on the size of the perturbative uncertainties of a top mass determination from the peak position. To avoid outliers we discard the two highest and two lowest points in the scan so as to better represent the bulk of the points. In Fig. 6 we show the peak positions of the curves at $Q = 700$ GeV and $Q = 2000$ GeV for the default profile functions and their perturbative uncertainties, estimated from the 500 random profile functions. Results are shown in the MSR mass scheme with gap subtraction (MSR, red) and in the pole mass scheme without gap subtractions (pole, blue) at NLL, NNLL, and N^3 LL. The central values, shown by dots, correspond to the default profile scales, so the perturbative uncertainties are asymmetric. The corresponding numbers are also given in Table II. The results show that the perturbative uncertainty is systematically smaller when the top quark mass and soft function

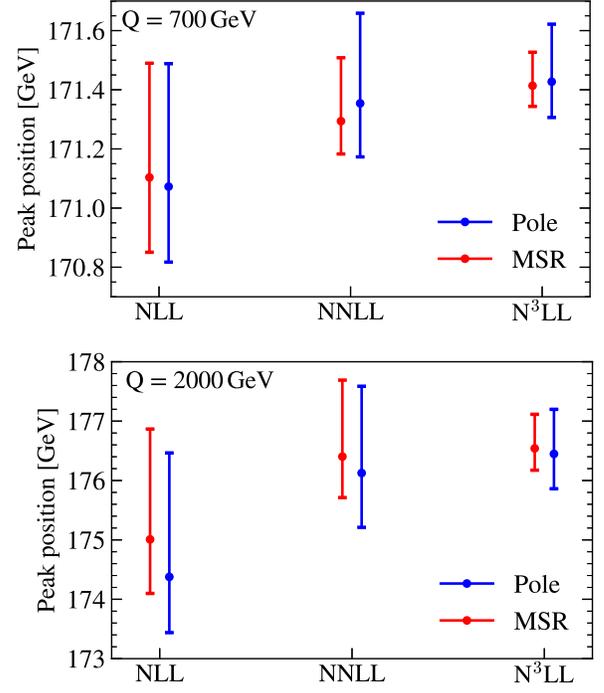


FIG. 6. Peak positions at $Q = 700$ GeV (upper panel) and $Q = 2000$ GeV (lower panel) for cross sections in the MSR (red) and pole schemes (blue) at NLL, NNLL, and N^3 LL accuracy. The error bars are obtained from a flat random scan over 500 parameters and the central value corresponds to the default profile.

renormalons are subtracted. For $Q = 700$ GeV, where we have the highest top quark mass sensitivity, using renormalon subtractions leads to an uncertainty in the peak location of around ± 85 MeV at N^3 LL order. Without renormalon subtractions the uncertainty at this order increases to around ± 150 MeV, which is almost a factor of 2 larger. For $Q = 2000$ GeV the uncertainties are larger for the analysis without renormalon subtraction as well (around ± 450 MeV with renormalon subtraction compared to around ± 650 MeV without renormalon subtractions, both at N^3 LL order). Here the difference is less pronounced because the overall top mass quark sensitivity decreases for larger Q values and the overall uncertainties increase. Analogous behavior is also visible at lower orders. Our results indicate that the MSR mass may be extracted with

TABLE II. Peak positions at different perturbative orders using the MSR and pole mass schemes, as shown in Fig. 6.

Mass scheme	Q [GeV]	Peak Positions [GeV]		
		NLL	NNLL	N^3 LL
MSR	700	$171.104^{+0.386}_{-0.253}$	$171.294^{+0.214}_{-0.111}$	$171.414^{+0.113}_{-0.070}$
	2000	$175.008^{+1.858}_{-0.910}$	$176.403^{+1.287}_{-0.690}$	$176.541^{+0.574}_{-0.367}$
Pole	700	$171.073^{+0.416}_{-0.255}$	$171.354^{+0.305}_{-0.181}$	$171.427^{+0.195}_{-0.121}$
	2000	$174.377^{+2.087}_{-0.938}$	$176.126^{+1.461}_{-0.915}$	$176.448^{+0.750}_{-0.587}$

an uncertainty of well below 100 MeV, while the pole mass uncertainty is at the level of 150 MeV. Interestingly, this is about the size of the top quark pole mass renormalon ambiguity of 166 MeV that was estimated recently in Ref. [35] for the case of massless charm and bottom quarks (which is the approximation we use in our analysis). Note that in an earlier analysis in Ref. [50] the top quark pole mass renormalon ambiguity was estimated as the smaller value of 67 MeV (also for massless charm and bottom quarks).

The results we have obtained in this simple analysis of the peak positions—taken by themselves—do not contradict the view that the top quark pole mass can be extracted from the 2-jettiness cross section with perturbative uncertainties below the pole mass renormalon ambiguity, but they also show that at least at N³LL order the precision is not (yet) sufficient to achieve that goal and that higher-order corrections beyond this order would be mandatory to get there. On the other hand, the results also support the view that, even though the 2-jettiness cross section exhibits a cancellation between the pole mass and soft function renormalons, the pole mass can still not be extracted with a precision below its renormalon ambiguity. In any case, using renormalon subtractions, and in particular the MSR mass scheme, will yield substantially higher precision and smaller perturbative uncertainties.

At this point we would like to again mention that in mass determinations from data (or MC pseudodata) simultaneous fits of the peak region 2-jettiness distribution for several Q values are needed to disentangle the dependence on the top quark mass and the shape function parameters, and that the whole distribution in the peak region (rather than just the peak position) would enter such fits. As mentioned before, however, this kind of study requires that also off-shell and m_t/Q power-suppressed contributions are included, as their effects can be non-negligible depending on how the cross section is normalized.

The dominant such QCD corrections to the factorization theorem in the bHQET region come from two sources: mass power corrections appearing as higher-order terms in Eq. (2), and corrections to the perturbative singular structures. (Additional nonsingular kinematic power corrections are very small at one loop and hence irrelevant.) The former are universal at any order in α_s and shift the distribution to the right by $\mathcal{O}(m_t^4/Q^4)$ but are trivial to incorporate. The latter are known analytically to $\mathcal{O}(\alpha_s)$ in QCD [28]: at tree level one gets a modification of the coefficient of the delta function, while at $\mathcal{O}(\alpha_s)$ the plus distribution coefficient is also affected. One can include these mass corrections by a suitable modification of the hard and jet functions (see, e.g., Ref. [29] for more details). These mass power corrections decrease the cross sections in Fig. 2 by 5% (beyond NLL) everywhere except for the region to the left of the peak where the effects are smaller. However, if the cross section is normalized the effect of

these power corrections drops below a percent, becoming negligible in all relevant regions. It is reasonable to believe that these power corrections will be of similar form and size once $\mathcal{O}(\alpha_s^2)$ are added. A complete analysis that accounts for these effects will require the inclusion of the $\mathcal{O}(\alpha_s^2)$ correction to the primary massive quark SCET jet function that was computed recently in Ref. [74] and shall be carried out in future work.

VIII. CONCLUSIONS

In this article we have presented results for the 2-jettiness differential distribution for boosted tops produced in e^+e^- collisions in the peak region, accounting for the resummation of large QCD logarithms at N³LL order and fixed-order corrections to the hard, soft, and jet function matrix elements at next-to-next-to-leading order [$\mathcal{O}(\alpha_s^2)$], calculated in the framework of soft-collinear effective theory and boosted heavy quark effective theory. We have systematically removed the $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalons contained in the soft and jet functions by using a gap subtraction as well as the MSR mass, and have provided a numerical analysis indicating that the perturbative uncertainties of a determination of the top quark MSR mass from the N³LL + $\mathcal{O}(\alpha_s^2)$ prediction at a c.m. energy of $Q = 700$ GeV are well below the level of 100 MeV. For future reference all theoretical formulas are given explicitly in several appendices.

An interesting aspect of the 2-jettiness distribution is that the soft- and jet-function renormalons partially cancel each other for center-of-mass energies above 700 GeV where the boosted top quark approximation is valid and precise top mass determinations can be carried out. This cancellation arises because the soft-function and pole-mass renormalons enter with different signs. While these two renormalons represent two physically independent infrared sensitivities, the cancellation allows for rather stable and convergent predictions in the pole mass scheme, if at the same time the soft function renormalon is also left unsubtracted. However, the resulting perturbative uncertainties are still systematically larger compared to the predictions where both renormalons are independently removed.

The analysis done here based on boosted heavy quark effective theory neglects subleading collinear off-shell corrections, which have been determined recently at $\mathcal{O}(\alpha_s^2)$ in Ref. [74] and shall be accounted for in future work.

ACKNOWLEDGMENTS

This work was supported in part by FWF Austrian Science Fund under the Project No. P28535-N27, the Spanish MINECO Ramón y Cajal program (RYC-2014-16022), the Office of Nuclear Physics of the U.S. Department of Energy under the Grant No. DE-SCD011090, the MECD Grants No. FPA2016-78645-P and No. PID2019-105439 GB-C22, the IFT Centro de

Excelencia Severo Ochoa Program under Grant SEV-2012-0249, the EU STRONG-2020 project under the program H2020-INFRAIA-2018-1, Grant Agreement No. 824093 and the COST Action CA16201 PARTICLEFACE. I. S. was also supported by the Simons Foundation through the Investigator Grant No. 327942. We acknowledge partial support by the FWF Austrian Science Fund under the Doctoral Program ‘‘Particles and Interactions’’ No. W1252-N27. A. P. is a member of the Lancaster-Manchester-Sheffield Consortium for Fundamental Physics, which is supported by the UK Science and Technology Facilities Council (STFC) under Grant No. ST/T001038/1. V. M. thanks the University of Vienna and MIT for hospitality while parts of this work were completed. B. B. is partially supported by the Government of the Republic of Trinidad and Tobago. B. B. also thanks the University of Vienna and DESY for hospitality while parts of this work were completed. Figures are made using the MATHEMATICA package PLHOT [111].

APPENDIX A: FORMULAS

The factorization theorem presented in Eq. (3) is expressed in momentum space where the jet and soft functions, along with the evolution factors, are distributions involving series in plus and delta functions. We find it convenient to combine the ingredients in position space where the convolutions become simple products. Below our notation and definitions with variable mass dimension follow Ref. [26]. For earlier work on the associated resummation formulas, see Refs. [112–114]. For a function $\mathcal{F}(q, \mu)$ that depends on a momentum-space variable q with mass dimensions $j_{\mathcal{F}}$ and the renormalization scale μ , the Fourier transform in position space is defined as

$$\tilde{\mathcal{F}}(x, \mu) = \int_{-\infty}^{\infty} dq e^{-iqx} \mathcal{F}(q, \mu), \quad (\text{A1})$$

where x has mass dimensions $-j_{\mathcal{F}}$. The position-space anomalous dimension permits writing the corresponding RG relation as a local equality,

$$\mu \frac{d}{d\mu} \tilde{\mathcal{F}}(x, \mu) = \tilde{\gamma}_{\mathcal{F}}(x, \mu) \tilde{\mathcal{F}}(x, \mu), \quad (\text{A2})$$

such that RG-evolved position-space soft and jet functions can be expressed as a regular product:

$$\begin{aligned} \eta(\Gamma, \mu, \mu_0) = & -\frac{\Gamma_0}{2\beta_0} \left\{ \log r + \frac{\alpha_0}{4\pi} \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) (r-1) + \frac{1}{2} \left(\frac{\alpha_0}{4\pi} \right)^2 \left(\frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} + \frac{\Gamma_2}{\Gamma_0} - \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} \right) (r^2-1) \right. \\ & \left. + \frac{1}{3} \left(\frac{\alpha_0}{4\pi} \right)^3 \left[\frac{\Gamma_3}{\Gamma_0} - \frac{\beta_3}{\beta_0} + \frac{\Gamma_1}{\Gamma_0} \left(\frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} \right) - \frac{\beta_1}{\beta_0} \left(\frac{\beta_1^2}{\beta_0^2} - 2\frac{\beta_2}{\beta_0} + \frac{\Gamma_2}{\Gamma_0} \right) \right] (r^3-1) \right\}, \end{aligned} \quad (\text{A7})$$

and

$$\begin{aligned} \mathcal{F}(q, \mu) &= \int dq' U_{\mathcal{F}}(q-q', \mu, \mu_0) \mathcal{F}(q', \mu_0) \\ &= \int \frac{dx}{2\pi} e^{iqx} \tilde{U}_{\mathcal{F}}(x, \mu, \mu_0) \tilde{\mathcal{F}}(x, \mu_0). \end{aligned} \quad (\text{A3})$$

The factorization function H_Q is not a distribution, but a simple function of the center-of-mass energy Q , with mass dimension $j_H = 1$ in our convention. One can treat factorization, and position-space jet and soft functions on the same footing simply by using $\mathcal{Q} = \{Q, 1/(ie^{\gamma_E} x)\}$, $F = \{H_Q, \tilde{\mathcal{F}}\}$, and $\gamma_F = \{\gamma_H, \tilde{\gamma}_{\mathcal{F}}\}$ [we describe the evolution of the bHQET current in Eq. (B3)]. In this way, we express the RG evolution and evolution from μ_0 to μ as

$$\begin{aligned} \mu \frac{d}{d\mu} F(\mathcal{Q}, \mu) &= \left[\Gamma_F[\alpha_s] \log \left(\frac{\mu}{Q} \right) + \gamma_F[\alpha_s] \right] F(\mu, \mu_0), \\ F(\mathcal{Q}, \mu) &= U_F(\mu, \mu_0; \mathcal{Q}) F(\mu_0, \mathcal{Q}) \\ &\equiv e^{K_F(\mu, \mu_0)} \left(\frac{\mu_0^{j_F}}{Q} \right)^{\omega_F(\mu, \mathcal{Q})} F(\mathcal{Q}, \mu_0), \end{aligned} \quad (\text{A4})$$

where $\Gamma_F[\alpha_s]$ resums double logarithms and is proportional to the universal cusp anomalous dimension $\Gamma^{\text{cusp}}[\alpha_s]$, and $\gamma_F[\alpha_s]$ is the noncusp anomalous dimension. The evolution kernels K_F and ω_F are defined as [$\alpha_0 \equiv \alpha_s(\mu_0)$, $\alpha_\mu \equiv \alpha_s(\mu)$]

$$\begin{aligned} K_F(\mu, \mu_0) &= K_F^\Gamma(\mu, \mu_0) + K_F^\gamma(\mu, \mu_0), \\ K_F^\Gamma(\mu, \mu_0) &= j_F \int_{\alpha_0}^{\alpha_\mu} \frac{d\alpha}{\beta(\alpha)} \Gamma_F[\alpha] \int_{\alpha_0}^{\alpha} \frac{d\alpha'}{\beta(\alpha')}, \\ K_F^\gamma(\mu, \mu_0) &= \int_{\alpha_0}^{\alpha_\mu} \frac{d\alpha}{\beta(\alpha)} \gamma_F[\alpha], \\ \omega_F(\mu, \mu_0) &= \int_{\alpha_0}^{\alpha_\mu} \frac{d\alpha}{\beta(\alpha)} \Gamma_F[\alpha]. \end{aligned} \quad (\text{A5})$$

The results of the evolution kernels at N³LL are given by

$$\begin{aligned} K_F^\Gamma(\mu, \mu_0) &= j_F K(\Gamma_F, \mu, \mu_0), \\ K_F^\gamma(\mu, \mu_0) &= \eta(\gamma_F, \mu, \mu_0), \\ \omega_F(\mu, \mu_0) &= \eta(\Gamma_F, \mu, \mu_0), \end{aligned} \quad (\text{A6})$$

where

$$\begin{aligned}
K(\Gamma, \mu, \mu_0) = & \frac{\Gamma_0}{4\beta_0^2} \left\{ \frac{4\pi}{r\alpha_0} (r \log r + 1 - r) + \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) (r - 1 - \log r) - \frac{\beta_1}{2\beta_0} \log^2 r \right. \\
& + \frac{\alpha_0}{4\pi} \left[\left(\frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} - \frac{\beta_1^2}{\beta_0^2} \right) (r - 1 - r \log r) - B_2 \log r + \left(\frac{\Gamma_2}{\Gamma_0} - \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} + B_2 \right) \frac{r^2 - 1}{2} + \left(\frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} - \frac{\Gamma_2}{\Gamma_0} \right) (r - 1) \right] \\
& + \left(\frac{\alpha_0}{4\pi} \right)^2 \left[\left(B_2 \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) + \frac{B_3}{2} \right) \frac{r^2 - 1}{2} + \left(\frac{\Gamma_3}{\Gamma_0} - \frac{\Gamma_2\beta_1}{\Gamma_0\beta_0} + \frac{B_2\Gamma_1}{\Gamma_0} + B_3 \right) \left(\frac{r^3 - 1}{3} - \frac{r^2 - 1}{2} \right) \right. \\
& \left. - \frac{\beta_1}{2\beta_0} \left(\frac{\Gamma_2}{\Gamma_0} - \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} + B_2 \right) \left(r^2 \log r - \frac{r^2 - 1}{2} \right) - \frac{B_3}{2} \log r - B_2 \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) (r - 1) \right] \left. \right\}, \quad (\text{A8})
\end{aligned}$$

where $r = \alpha_\mu/\alpha_0$ depends on the 4-loop running coupling, and the B_i coefficients take the following values: $B_2 = \beta_1^2/\beta_0^2 - \beta_2/\beta_0$ and $B_3 = 2\beta_1\beta_2/\beta_0^2 - \beta_1^3/\beta_0^3 - \beta_3/\beta_0$. The series expansions of the QCD beta function and the cusp and noncusp pieces of a generic SCET anomalous dimension are written as

$$\begin{aligned}
\beta[\alpha_s] &= -2\alpha_s \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi} \right)^{n+1}, \\
\Gamma[\alpha_s] &= \sum_{n=0}^{\infty} \Gamma_n \left(\frac{\alpha_s}{4\pi} \right)^{n+1}, \quad (\text{A9})
\end{aligned}$$

where here the Γ 's either stand for Γ_F , Γ^{cusp} , or γ_F . The numerical expressions for the universal cusp anomalous dimension coefficients for 5 and 6 flavors are given by [73,115,116]

$$\begin{aligned}
\{\Gamma_i^{\text{cusp}}\}_{0 \leq i \leq 3}^{(5)} &= \{4, 27.633, 179.406, 141.254\}, \\
\{\Gamma_i^{\text{cusp}}\}_{0 \leq i \leq 3}^{(6)} &= \{4, 23.188, 35.497, -2581.527\}, \quad (\text{A10})
\end{aligned}$$

where the 4-loop cusp anomalous dimension is obtained from recent work in Refs. [73,117]. Note that in our convention we do not include a factor of C_F in the definition of $\Gamma^{\text{cusp}}[\alpha_s]$, but we do include this factor for the $\Gamma_F[\alpha_s]$'s for various functions given below in Appendix B. Finally, we quote the numerical results for the QCD beta function up to four loops [95,96,118,119]:

$$\begin{aligned}
\{\beta_i^{(5)}\}_{0 \leq i \leq 3} &= \{23/3, 116/3, 180.907, 4826.16\}, \\
\{\beta_i^{(6)}\}_{0 \leq i \leq 3} &= \{7, 26, -32.5, 2472.28\}. \quad (\text{A11})
\end{aligned}$$

APPENDIX B: ANOMALOUS DIMENSIONS

The RG equations for the various functions appearing in Eq. (3) are as follows:

$$\begin{aligned}
\mu \frac{d}{d\mu} \log[H_Q^{(6)}(Q, \mu)] &= \Gamma_{H_Q}[\alpha_s] \log\left(\frac{\mu}{Q}\right) + \gamma_{H_Q}[\alpha_s], \\
\mu \frac{d}{d\mu} \log[\mathcal{J}_v^{(5)}(q, \mu)] &= \Gamma_v[\alpha_s] \log\left(\frac{1}{Q}\right) + \gamma_v[\alpha_s] \\
&\equiv \gamma_v(q), \\
\mu \frac{d}{d\mu} \log[\tilde{J}_{B,\tau_2}^{(5)}(x, \mu)] &= \Gamma_{J_B^c}[\alpha_s] \log(i e^{\gamma_E} x \mu) + \gamma_{J_B^c}[\alpha_s] \\
&\equiv \tilde{\gamma}_{J_B^c}(x, \mu), \\
\mu \frac{d}{d\mu} \log[\tilde{S}_{\tau_2}^{(5)}(y, \mu)] &= \Gamma_{S_\tau}[\alpha_s] \log(i e^{\gamma_E} y \mu) + \gamma_{S_\tau}[\alpha_s] \\
&\equiv \tilde{\gamma}_{S_\tau}(x, \mu), \quad (\text{B1})
\end{aligned}$$

where \tilde{J}_{B,τ_2} and \tilde{S}_{τ_2} are the Fourier transforms of the thrust, unsubtracted, stable bHQET jet and soft functions defined in Eqs. (11) and (23). Note that we have $\gamma_{H_Q} = 2\gamma_H$, $\gamma_{J_B^c} = 2\gamma_B$, and $\gamma_{S_\tau} = 2\gamma_S$, with $\gamma_{H,S}$ given as in Ref. [91] and γ_B defined in Ref. [26]. In Eq. (3), instead of running H_m we RG evolve the squared matrix element of the bHQET current, \mathcal{J}_v , defined as

$$\mathcal{J}_v^{(5)} = |\langle \mathcal{J}_{\text{bHQET}}^{(5)} \rangle|^2. \quad (\text{B2})$$

The RG evolution of $\mathcal{J}_v^{(5)}$, using the results in Eq. (B1) for the evolution factor U_v shown in Eq. (8), reads

$$U_v(Q, \mu_m, \mu) = e^{K_v^{(5)}(\mu_m, \mu)} Q^{-\omega_v^{(5)}(\mu_m, \mu)}, \quad (\text{B3})$$

where, keeping in line with our notation, we emphasize that μ_m is the final scale up to which $\mathcal{J}_v^{(5)}$ is RG evolved. The bHQET current $\mathcal{J}_{\text{bHQET}}^{(5)}$ [26,72] is given by

$$\mathcal{J}_{\text{bHQET}} = \bar{h}_{v_i} W_n Y_n^\dagger \Gamma_n^\mu Y_{\bar{n}} W_{\bar{n}}^\dagger h_{v_j}, \quad (\text{B4})$$

where h_{v_i} are the heavy-quark fields describing top and antitop quarks; $W_{n,\bar{n}}$ are the Wilson lines formed from ultra-collinear gluons, such that in position space $W_n^\dagger(x) = \text{P exp}[ig \int_0^\infty ds \bar{n} \cdot A_n(\bar{n}s + x)]$; $Y_{n,\bar{n}}$ are similarly defined Wilson lines with ultrasoft gluons, and $\Gamma_v^\mu = \gamma^\mu$ and

$\Gamma_a^\mu = \gamma^\mu \gamma_5$. The RG consistency in the bHQET sector implies the following constraint:

$$\gamma_v(\varrho) = \tilde{\gamma}_{J_B^c}(x, \mu) + \tilde{\gamma}_{S_\tau}(qx, \mu), \quad (\text{B5})$$

with $\gamma_v(\varrho)$, $\tilde{\gamma}_{J_B^c}(x, \mu)$, and $\tilde{\gamma}_{S_\tau}(qx, \mu)$ defined above in Eq. (B1). This leads to a cancellation of the $\log(\mu)$ dependence in the anomalous dimension of the bHQET current and implies $\Gamma_v[\alpha_s] = \Gamma_{J_B^c}[\alpha_s] = -\Gamma_{S_\tau}[\alpha_s]$, $\gamma_v[\alpha_s] = \gamma_{S_\tau}[\alpha_s] + \gamma_{J_B^c}[\alpha_s]$. Since the ‘‘cusp’’ piece in Eq. (B5) is μ independent, we find that the kernel K_v^Γ [see Eq. (A5)] does not appear in the RG evolution of the bHQET matrix element in Eq. (30), but K_v^γ and ω_v do. With the convention in Eq. (B1) the cusp and the noncusp pieces have the following values [26,69,72,76,78,114,116,120–124]:

$$\begin{aligned} \Gamma_{J_B^c}[\alpha_s] &= \Gamma_v[\alpha_s] = -\Gamma_{S_\tau}[\alpha_s] = 4C_F \Gamma^{\text{cusp}(5)}[\alpha_s], \\ \Gamma_{H_Q}[\alpha_s] &= -4C_F \Gamma^{\text{cusp}(6)}[\alpha_s], \\ \{\mathcal{Y}_i^{H_Q}\}_{0 \leq i \leq 2} &= \{-16, 32.669, -21.044\}, \\ \{\mathcal{Y}_i^v\}_{0 \leq i \leq 2} &= \{32/3, 86.619, 477.753\}, \\ \{\mathcal{Y}_i^{J_B^c}\}_{0 \leq i \leq 2} &= \{32/3, -65.803, -840.284\}, \\ \{\mathcal{Y}_i^{S_\tau}\}_{0 \leq i \leq 2} &= \{0, 152.422, 1318.037\}. \end{aligned} \quad (\text{B6})$$

Using the formulas for the evolution factors in Eq. (A4), the RG evolution equations in Eq. (B1), and the anomalous dimensions in Eq. (B6), one arrives at the resummed expressions in Eqs. (29) and (30). From Eq. (B1) we see that all dynamical momentum variables have dimensions of energy so that $j_F = 1$ for all the evolution functions that we consider here.

APPENDIX C: FIXED-ORDER RESULTS

We now state the results for the 2-loop matrix elements and factorization functions that appear in our analysis. We show the results for the logarithm of the functions since it simplifies the structure of the results.

1. 2-loop results

The SCET hard matching function $H_Q^{(6)}$ with 6-flavor coupling is given by [75–79,125]

$$\begin{aligned} \log[H_Q^{(6)}(Q, \mu)] &= \frac{\alpha_s^{(6)}(\mu)}{4\pi} [9.372 - 16L_Q - 10.667L_Q^2] \\ &+ \left[\frac{\alpha_s^{(6)}(\mu)}{4\pi} \right]^2 [305.454 - 163.879L_Q \\ &- 173.835L_Q^2 - 49.778L_Q^3], \\ L_Q &= \log\left(\frac{\mu}{Q}\right). \end{aligned} \quad (\text{C1})$$

For the hard matching function at the top mass scale $H_m^{(6)}$ we state the result with m_t expressed in either the pole or $\overline{\text{MS}}$ scheme [26,72]:

$$\begin{aligned} \log[H_m^{\text{pole}(6)}(m_t^{\text{pole}}, \varrho, \mu)] &= \frac{\alpha_s^{(6)}(\mu)}{4\pi} [15.053 - 2.667L_m + 2.667L_m^2] \\ &+ \left[\frac{\alpha_s^{(6)}(\mu)}{4\pi} \right]^2 [152.578 - 45.728L_m + 26.5699L_m^2 \\ &- 6.222L_m^3 + \log(\varrho)(-5.531 - 5.926L_m - 1.778L_m^2)], \end{aligned} \quad (\text{C2})$$

$$\begin{aligned} \log[H_m^{\overline{\text{MS}}(6)}(\bar{m}_t, \varrho, \mu)] &= \log[H_m^{\text{pole}(6)}(\bar{m}_t, \varrho, \mu)] \\ &+ \left[\frac{\alpha_s^{(6)}(\mu)}{4\pi} \right]^2 (56.889L_{\bar{m}} - 28.444), \end{aligned} \quad (\text{C3})$$

where $L_m = \log[(m_t^{\text{pole}})^2/\mu^2]$, $L_{\bar{m}} = \log(\bar{m}_t^2/\mu^2)$, and ϱ has been defined in Eq. (4).

The 2-loop fixed-order result for the unsubtracted, stable-top bHQET jet function in position space is given by [26,69]

$$\begin{aligned} \log[m_t^2 \tilde{J}_{B,\tau_2}^{(5)}(x, \mu)] &= \frac{\alpha_s^{(5)}(\mu)}{4\pi} [15.053 + 10.667\tilde{L}_B + 10.667\tilde{L}_B^2] \\ &+ \left[\frac{\alpha_s^{(5)}(\mu)}{4\pi} \right]^2 [310.954 + 165.012\tilde{L}_B + 155.465\tilde{L}_B^2 \\ &+ 54.518\tilde{L}_B^3], \quad \tilde{L}_B = \log(i e^{\gamma_E} x \mu). \end{aligned} \quad (\text{C4})$$

Likewise, the unsubtracted soft function in position space reads [26,70,71]

$$\begin{aligned} \log[\tilde{S}_{\tau_2}^{(5)}(y, \mu)] &= \frac{\alpha_s^{(5)}(\mu)}{4\pi} [-13.159 - 10.667\tilde{L}_S^2] \\ &+ \left[\frac{\alpha_s^{(5)}(\mu)}{4\pi} \right]^2 [-81.361 - 49.357\tilde{L}_S \\ &- 73.687\tilde{L}_S^2 - 54.518\tilde{L}_S^3], \\ \tilde{L}_S &= \log(i e^{\gamma_E} y \mu). \end{aligned} \quad (\text{C5})$$

2. Generating fixed-order terms

We now describe a helpful algorithm that allows one to generate the fixed-order expansion of the position-space matrix elements and factorization functions discussed here from the nonlogarithmic coefficients and their anomalous dimensions in Eq. (B6). This applies to the matching

coefficient $H_Q^{(6)}$, and the bHQET jet and the soft matrix elements, and can be used to reproduce the results stated above in Eqs. (C1), (C4), and (C5). Note that the algorithm described below must be generalized in an obvious way for $H_m^{(6)}$ to obtain the results in Eqs. (C2) and (C3), since its running results from the anomalous dimensions of the bHQET current and the SCET matching coefficient (with different number of dynamical flavors in the running coupling) and includes an additional rapidity logarithm.

Consider a position-space matrix element or a factorization function $F(\mu, Q)$ having the generic form

$$F(\mu, Q) = A \exp \left\{ \sum_{m=1} a_{mn}^F \left[\frac{\alpha_s(\mu)}{4\pi} \right]^m \sum_{n=0}^{m+1} \log^n \left(\frac{\mu^{j_F}}{Q} \right) \right\}, \quad (\text{C6})$$

where we have $A = 1/m_1^2$ for the thrust position-space bHQET jet function and $A = 1$ in the other cases; see Eq. (28). The constant terms a_{m0}^F serve as independent data, whereas other coefficients can then be determined by anomalous dimensions and the beta function. Thus, the a_{m0}^F serve as boundary condition data for the RG differential equations in Eq. (A4). The logarithmic terms a_{mn}^F for $n \geq 1$ then can be expressed as

$$a_{mn}^F = \frac{1}{j_F^n} (a_{mn}^F[\beta] + a_{mn}^F[\beta, \gamma_F] + a_{mn}^F[\beta, \Gamma_F]), \quad (\text{C7})$$

where the three terms result from the running of the coupling, and the noncusp and cusp pieces of the anomalous dimension of the given function. The coefficients in Eq. (C7) can be obtained via the following recursion relations:

$$\begin{aligned} a_{mn}^F[\beta] &= \frac{2}{n} \sum_{i=n}^{m-1} i a_{i(n-1)}^F[\beta] \beta_{m-i-1}^{(n_f)}, & 1 \leq n \leq m-1, \\ a_{mn}^F[\beta, \gamma_F] &= \frac{2}{n} \sum_{i=n-1}^{m-1} i a_{i(n-1)}^F[\beta, \gamma_F] \beta_{m-i-1}^{(n_f)}, & 2 \leq n \leq m, \\ a_{mn}^F[\beta, \Gamma_F] &= \frac{2}{n} \sum_{i=n-2}^{m-1} i a_{i(n-1)}^F[\beta, \Gamma_F] \beta_{m-i-1}^{(n_f)}, & 3 \leq n \leq m+1, \end{aligned} \quad (\text{C8})$$

with $m > 1$ in order to have a sensible upper limit. The starting values of the three series (with $m \geq 1$) are given by

$$\begin{aligned} a_{m0}^F[\beta] &= a_{m0}^F, & a_{m1}^F[\beta, \gamma_F] &= \gamma_{m-1}^F, \\ a_{m2}^F[\beta, \Gamma_F] &= \frac{j_F}{2} \Gamma_{m-1}^F, \end{aligned} \quad (\text{C9})$$

with $a_{m0}^F[\beta, \gamma_F] = a_{m0}^F[\beta, \Gamma_F] = a_{m1}^F[\beta, \Gamma_F] = 0$. Here, the integer j_F corresponds to the dimension of the

momentum-space variable q as it appears naturally in logarithms,¹⁰ as shown in Eq. (A4). The dependence on j_F is factorized as in Eq. (C7) and only enters the $a_{mn}^F[\beta, \Gamma_F]$ coefficients through the boundary condition in Eq. (C9). For the factorization functions in Eq. (A4) this is simply set to 1. The constant terms of the SCET matrix elements up to NNLO are as follows:

$$\begin{aligned} \{a_{i,0}^{H_Q}\}_{1 \leq i \leq 2} &= \{9.372, 305.454\}, \\ \{a_{i,0}^{J_{B,\tau_2}}\}_{1 \leq i \leq 2} &= \{15.053, 310.954\}, \\ \{a_{i,0}^{S_{\tau_2}}\}_{1 \leq i \leq 2} &= \{-13.159, -81.361\}. \end{aligned} \quad (\text{C10})$$

APPENDIX D: MSR MASS

The defining series for the standard $\overline{\text{MS}}$ mass $\overline{m}_t^{(6)} \equiv \overline{m}_t^{(6)}(\overline{m}_t^{(6)})$ reads

$$\delta \overline{m} \equiv m_t^{\text{pole}} - \overline{m}_t^{(6)} = \overline{m}_t^{(6)} \sum_{i=1} \left[\frac{\alpha_s^{(6)}(\overline{m}_t^{(6)})}{4\pi} \right]^i a_i^{(6)}(5, 1), \quad (\text{D1})$$

where the notation $a_i^{(n_f)}(n_\ell, n_h)$ refers to the coefficient for n_ℓ massless and n_h heavy flavors with the running coupling expressed in the n_f -flavor scheme (note that in general $n_f \neq n_\ell + n_h$). The defining relation for the MSR mass¹¹ is given by

$$\begin{aligned} \delta m(R) &= m_t^{\text{pole}} - m_t^{\text{MSR},(5)}(R) \\ &= R \sum_{i=1} \left[\frac{\alpha_s^{(5)}(R)}{4\pi} \right]^i a_i^{(5)}(5, 0), \end{aligned} \quad (\text{D2})$$

where $n_h = 0$ signifies that the virtual self-energy corrections coming from top quark virtual loops have been integrated out. Using these results, we can write down the matching relation between the $\overline{\text{MS}}$ and MSR masses at the scale $\overline{m}_t^{(6)}$:

$$m_t^{\text{MSR},(5)}(\overline{m}_t^{(6)}) - \overline{m}_t^{(6)} = \overline{m}_t^{(6)} \sum_{i=1} \left[\frac{\alpha_s^{(5)}(\overline{m}_t^{(6)})}{4\pi} \right]^i \Delta a_i^{(5)}, \quad (\text{D3})$$

with

$$\Delta a_i^{(5)} = a_i^{(5)}(5, 1) - a_i^{(5)}(5, 0). \quad (\text{D4})$$

¹⁰Even though from Eq. (B1) one can see that $j_F = 1$ for all the functions we consider in our analysis, for sake of generality we have left it explicit in the formulas above.

¹¹We remind the reader that we adopt the ‘‘natural’’ MSR scheme as defined in Ref. [34] and refer to it as just the MSR mass.

The coefficients have the following numerical values [126–131]:

$$\begin{aligned} \{a_i^{(5)}(5, 0)\}_{1 \leq i \leq 3} &= \{5.333, 130.128, 4582.535\}, \\ \{a_i^{(5)}(5, 1)\}_{1 \leq i \leq 3} &= \{5.333, 131.785, 4699.703\}. \end{aligned} \quad (\text{D5})$$

APPENDIX E: SOFT GAP SUBTRACTION SCHEMES

We can generically start the construction of the gap subtraction series $\bar{\delta}$ that shall cancel the soft function renormalon in Eq. (23) by considering the following general condition:

$$\frac{d^n}{d \log^n(iy)} \log[\tilde{S}_{\tau_2}^{(5)}(y, \mu_\delta) e^{-2iy\bar{\delta}(\mu_\delta, R_s, n, \xi)}]_{y=\frac{\xi}{iR_s}} = 0, \quad (\text{E1})$$

where $n \geq 0$ and $\xi \sim \mathcal{O}(1)$ is an auxiliary parameter. This condition specifies a physical renormalon-free ‘‘momentum-subtraction-like’’ scheme which defines the soft function by imposing a condition on it at a point in position space. Here $\tilde{S}_{\tau_2}^{(5)}(y, \mu_\delta)$ is the $\overline{\text{MS}}$ soft function and μ_δ is a reference renormalization scale that can be chosen independently of R_s . Solving Eq. (E1) for $\bar{\delta}$ gives

$$\bar{\delta}(\mu_\delta, R_s; n, \xi) \equiv \frac{R_s}{2\xi} \frac{d^n}{d \log^n(iy)} \log[\tilde{S}_{\tau_2}^{(5)}(y, \mu_\delta)]_{iy=\frac{\xi}{R_s}}. \quad (\text{E2})$$

This defines a range of gap subtraction schemes for different choices of n and ξ . The renormalon in the soft function is not influenced by the terms depending on the cusp or noncusp anomalous dimensions, and hence appears only in the terms $a_{ij}^{S_{\tau_2}}[\beta]$ using the notation of Eq. (C7). For a given n , these terms enter the series for $\bar{\delta}$ at $\mathcal{O}(\alpha_s^{n+1})$. The choices $n = 0, 1$ yield the two subtraction schemes given in Eqs. (24) and (25). We do not consider schemes with $n \geq 2$ where the $a_{ij}[\beta]$ enter at $\mathcal{O}(\alpha_s^3)$ and beyond.

The scheme used in Refs. [84,91] corresponds to the choice $\xi = e^{\gamma_E}$ and $\mu_\delta = \mu_s$, yielding Eq. (25). Here we instead employ the $n = 0$ scheme in Eq. (24), setting $\xi = 1$ and $\mu_\delta = R_s$ instead, which makes the gap parameter $\bar{\Delta}$ independent of the renormalization scale μ_s of the soft function. This yields

$$\begin{aligned} \bar{\delta}(R_s) &\equiv \bar{\delta}(R_s, R_s; 0, 1) = R_s \sum_{i=1} \left[\frac{\alpha_s^{(5)}(R_s)}{4\pi} \right]^i d_{i0}, \\ \text{where } d_{i0} &= \frac{1}{2} \sum_{j=0}^{i+1} s_{ij} \gamma_E^j, \end{aligned} \quad (\text{E3})$$

where the s_{ij} 's are simply the constant terms and coefficients of powers of logarithms $\tilde{L}_S^j = \log^j(i e^{\gamma_E} y \mu)$ in the

fixed-order expansion of $\log[S_{\tau_2}^{(5)}(y, \mu)]$ at $\mathcal{O}(\alpha_s^i)$ in Eq. (C5) [see Eq. (28)]. We remind the reader that when $\bar{\delta}(R_s)$ is used in the factorization theorem, it is crucial that $\bar{\delta}$ is treated as a series expansion in $\alpha_s^{(5)}(\mu_s)$, the same coupling used in the series for the soft function. Therefore, the final expression used for our analysis is

$$\begin{aligned} \bar{\delta}(R_s) &= R_s \sum_{i=1} \left[\frac{\alpha_s^{(5)}(\mu_s)}{4\pi} \right]^i d_{ij} \log^j \left(\frac{\mu_s}{R_s} \right), \\ \text{with } d_{ij} &= \frac{2}{j} \sum_{k=j}^{i-1} k d_{k(j-1)} \beta_{i-k-1}^{(5)}. \end{aligned} \quad (\text{E4})$$

Finally, we can relate the leading power correction $\Omega_1(R_s)$ between the two subtraction schemes:

$$\begin{aligned} \Omega_1(R_s) - \Omega_1^{\text{Refs.}[84,91]}(R_s, R_s) &= R_s [\bar{\delta}^{\text{Refs.}[84,91]}(R_s, R_s) - \bar{\delta}(R_s)] \\ &= R_s \left\{ 8.357 \frac{\alpha_s^{(5)}(R_s)}{4\pi} + 28.489 \left[\frac{\alpha_s^{(5)}(R_s)}{4\pi} \right]^2 \right\}. \end{aligned} \quad (\text{E5})$$

APPENDIX F: R EVOLUTION

We consider a generic perturbative series depending linearly and logarithmically on the scale R that has the following form¹²:

$$f(R) = R \sum_{i=1}^{\infty} \left[\frac{\alpha_s(R)}{4\pi} \right]^i f_i. \quad (\text{F1})$$

The evolution equation of $f(R)$ with respect to R is then given by

$$\frac{df(R)}{d \ln R} = R \frac{df(R)}{dR} = R \sum_{n=0}^{\infty} \gamma_n^{f,R} \left[\frac{\alpha_s(R)}{4\pi} \right]^{n+1}, \quad (\text{F2})$$

with the R anomalous dimension coefficients being

$$\gamma_n^{f,R} = f_{n+1} - 2 \sum_{j=0}^{n-1} (n-j) \beta_j f_{n-j}, \quad (n \geq 1), \quad (\text{F3})$$

where $\gamma_0^{f,R} = f_1$. The crucial aspect of using the RG equation in Eqs. (F2) is that the R -evolution anomalous dimension in Eq. (F3) is $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon free if the series in Eq. (F1) contains an $\mathcal{O}(\Lambda_{\text{QCD}})$ renormalon [33,34]. This is a fundamental ingredient in the construction of renormalon subtractions in distributions where the subtraction scale depends on the value of kinematic

¹²In this Appendix it is assumed that α_s runs with n_f active flavors.

quantities [21,132]. The solution of the R -evolution equation is straightforward and given by

$$f(R_1) - f(R_0) = \sum_{n=0}^{\infty} \gamma_n^{f,R} \int_{R_0}^{R_1} dR \left[\frac{\alpha_s(R)}{4\pi} \right]^{n+1}. \quad (\text{F4})$$

Using Eq. (F3), we can derive the anomalous dimensions for the R -evolution of the MSR mass and the gap subtractions by using the following values for f_i :

$$\begin{aligned} \delta m(R), \text{ Eq. (D2): } f_i &= a_i^{(5)}(5, 0), \\ \bar{\delta}(R_s), \text{ Eq. (E3): } f_i &= d_{i0}. \end{aligned} \quad (\text{F5})$$

Note that the R -evolution equations for the MSR mass m_i^{MSR} and the gap parameter $\bar{\Delta}(R_s)$ in Eqs. (22) and (26) are defined with an additional minus sign of the evolution equations of the corresponding subtraction terms which are referred to in Eq. (F4). For the R -anomalous dimensions we get

$$\begin{aligned} \{\gamma_i^{\text{MSR}}\}_{0 \leq i \leq 3} &= \{5.333, 48.350, 179.501\}, \\ \{\gamma_i^{\Delta}\}_{0 \leq i \leq 2} &= \{-8.357, 55.693\}. \end{aligned} \quad (\text{F6})$$

APPENDIX G: NONPERTURBATIVE MODEL FUNCTION

We use the nonperturbative model function $F(k)$ that has the form

$$F(k) = S_{\tau}^{\text{mod}}(k, \lambda, \{c_i\}) \equiv \frac{1}{\lambda} \left[\sum_{n=0}^N c_n f_n \left(\frac{k}{\lambda} \right) \right]^2, \quad (\text{G1})$$

where the basis functions are [87]

$$\begin{aligned} f_n(z) &= 8 \sqrt{\frac{2z^3(2n+1)}{3}} e^{-2z} P_n(g(z)), \\ g(z) &= \frac{2}{3} [3 - e^{-4z}(3 + 12z + 24z^2 + 32z^3)] - 1, \end{aligned} \quad (\text{G2})$$

and P_n are Legendre polynomials. For $\sum_i c_i^2 = 1$ the norm of $S_{\tau}^{\text{mod}}(k)$ is unity, i.e., $\Omega_0 = 1$. The choice of basis in Eqs. (G1) and (G2) depends on specifying one dimensionful parameter λ which is characteristic for the width of the soft function. Following Ref. [91], in our analysis we set $N = 2$ and $c_1 = 0$, such that the first moment using Eq. (15) is given by

$$\bar{\Omega}_1 = \Delta + \frac{\lambda}{2} [c_0^2 + 0.201c_0c_2 + 1.100c_2^2], \quad (\text{G3})$$

where the normalization condition $c_0^2 + c_2^2 = 1$ can be used to express $c_0 > 0$ in terms of c_1 . For our numerical analyses we take $c_2 = 0.05$, $\Delta = 0.1$ GeV, and the value of $\bar{\Omega}_1$ then fixes λ . The $\Omega_1(R_s)$ including gap running can be evaluated using Eqs. (18) and (E3). For a reference scale of $R_s = 2$ GeV, we find

$$\begin{aligned} \Omega_1(2 \text{ GeV}) &= \bar{\Omega}_1 - (2 \text{ GeV}) \sum_{i=1}^2 \left[\frac{\alpha_s^{(5)}(2 \text{ GeV})}{4\pi} \right]^i \delta_{i0} \\ &= \bar{\Omega}_1 + 0.463 \text{ GeV}. \end{aligned} \quad (\text{G4})$$

-
- [1] N. Cabibbo, L. Maiani, G. Parisi, and R. Petronzio, *Nucl. Phys.* **B158**, 295 (1979).
 - [2] S. Alekhin, A. Djouadi, and S. Moch, *Phys. Lett. B* **716**, 214 (2012).
 - [3] D. Buttazzo, G. Degrossi, P. P. Giardino, G. F. Giudice, F. Sala, A. Salvio, and A. Strumia, *J. High Energy Phys.* **12** (2013) 089.
 - [4] A. Andreassen, W. Frost, and M. D. Schwartz, *Phys. Rev. Lett.* **113**, 241801 (2014).
 - [5] V. Branchina and E. Messina, *Phys. Rev. Lett.* **111**, 241801 (2013).
 - [6] V. Branchina, E. Messina, and A. Platania, *J. High Energy Phys.* **09** (2014) 182.
 - [7] M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer (Gfitter Group), *Eur. Phys. J. C* **74**, 3046 (2014).
 - [8] P. Zyla *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
 - [9] P. Azzi *et al.*, *CERN Yellow Rep. Monogr.* **7**, 1 (2019).
 - [10] A. H. Hoang and I. W. Stewart, *Nucl. Phys. B, Proc. Suppl.* **185**, 220 (2008).
 - [11] A. H. Hoang, in *7th International Workshop on Top Quark Physics (TOP2014) Cannes, France, 2014* (2014) arXiv: 1412.3649.
 - [12] G. Corcella, *Front. Phys.* **7**, 54 (2019).
 - [13] A. H. Hoang, *Ann. Rev. Nucl. Part. Sci.* **70**, 225 (2020).
 - [14] M. Dasgupta, F. A. Dreyer, K. Hamilton, P. F. Monni, and G. P. Salam, *J. High Energy Phys.* **09** (2018) 033.
 - [15] A. H. Hoang, S. Plätzer, and D. Samitz, *J. High Energy Phys.* **10** (2018) 200.
 - [16] T. Becher and M. D. Schwartz, *J. High Energy Phys.* **07** (2008) 034.

- [17] Y.-T. Chien and M. D. Schwartz, *J. High Energy Phys.* **08** (2010) 058.
- [18] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Phys. Rev. D* **86**, 094002 (2012).
- [19] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, *Phys. Rev. D* **91**, 094017 (2015).
- [20] C. Lee and G. Sterman, eConf **C0601121**, A001 (2006).
- [21] A. H. Hoang and I. W. Stewart, *Phys. Lett. B* **660**, 483 (2008).
- [22] V. Mateu, I. W. Stewart, and J. Thaler, *Phys. Rev. D* **87**, 014025 (2013).
- [23] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, *J. High Energy Phys.* **12** (2019) 002.
- [24] J. Mo, F. J. Tackmann, and W. J. Waalewijn, *Eur. Phys. J. C* **77**, 770 (2017).
- [25] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, *Phys. Rev. D* **77**, 074010 (2008).
- [26] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, *Phys. Rev. D* **77**, 114003 (2008).
- [27] M. Butenschoen, B. Dehnadi, A. H. Hoang, V. Mateu, M. Preisser, and I. W. Stewart, *Phys. Rev. Lett.* **117**, 232001 (2016).
- [28] C. Lepenik and V. Mateu, *J. High Energy Phys.* **03** (2020) 024.
- [29] A. Bris, V. Mateu, and M. Preisser, *J. High Energy Phys.* **09** (2020) 132.
- [30] A. M. Sirunyan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **79**, 313 (2019).
- [31] M. Aaboud *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **79**, 290 (2019).
- [32] T. E. W. Group (CDF and D0 Collaboration), arXiv: 1407.2682.
- [33] A. H. Hoang, A. Jain, I. Scimemi, and I. W. Stewart, *Phys. Rev. Lett.* **101**, 151602 (2008).
- [34] A. H. Hoang, A. Jain, C. Lepenik, V. Mateu, M. Preisser, I. Scimemi, and I. W. Stewart, *J. High Energy Phys.* **04** (2018) 003.
- [35] A. H. Hoang, C. Lepenik, and M. Preisser, *J. High Energy Phys.* **09** (2017) 099.
- [36] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, *Phys. Rev. D* **100**, 074021 (2019).
- [37] S. Ferrario Ravasio, T. Ježo, P. Nason, and C. Oleari, *Eur. Phys. J. C* **78**, 458 (2018); **79**, A859 (2019).
- [38] G. Heinrich, A. Maier, R. Nisius, J. Schlenk, M. Schulze, L. Scyboz, and J. Winter, *J. High Energy Phys.* **07** (2018) 129.
- [39] G. Corcella, R. Franceschini, and D. Kim, *Nucl. Phys.* **B929**, 485 (2018).
- [40] P. Nason, in *11th International Workshop on Top Quark Physics (TOP2018) Bad Neuenahr, Germany, 2018* (2019) [arXiv:1901.04737].
- [41] M. Czakon, P. Fiedler, and A. Mitov, *Phys. Rev. Lett.* **110**, 252004 (2013).
- [42] G. Aad *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **74**, 3109 (2014); **76**, A642 (2016).
- [43] A. M. Sirunyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **09** (2017) 051.
- [44] M. Aaboud *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **77**, 804 (2017).
- [45] S. Alioli, P. Fernandez, J. Fuster, A. Irlles, S.-O. Moch, P. Uwer, and M. Vos, *Eur. Phys. J. C* **73**, 2438 (2013).
- [46] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **11** (2019) 150.
- [47] A. M. Sirunyan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **80**, 658 (2020).
- [48] W.-L. Ju, G. Wang, X. Wang, X. Xu, Y. Xu, and L. L. Yang, *J. High Energy Phys.* **06** (2020) 158.
- [49] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, and J. Mazzitelli, *J. High Energy Phys.* **08** (2020) 027.
- [50] M. Beneke, P. Marquard, P. Nason, and M. Steinhauser, *Phys. Lett. B* **775**, 63 (2017).
- [51] C. G. Lester and D. J. Summers, *Phys. Lett. B* **463**, 99 (1999).
- [52] S. Chatrchyan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **73**, 2494 (2013).
- [53] CMS Collaboration, Determination of the top-quark mass from the $m(\text{lb})$ distribution in dileptonic $t\bar{t}$ events at $\sqrt{s} = 8 \text{ TeV}$, Report No. CMS-PAS-TOP-14-014, 2014.
- [54] K. Agashe, R. Franceschini, D. Kim, and M. Schulze, *Eur. Phys. J. C* **76**, 636 (2016).
- [55] V. Khachatryan *et al.* (CMS Collaboration), *J. High Energy Phys.* **12** (2016) 123.
- [56] V. Khachatryan *et al.* (CMS Collaboration), *Phys. Rev. D* **93**, 092006 (2016).
- [57] C. W. Bauer, S. Fleming, and M. E. Luke, *Phys. Rev. D* **63**, 014006 (2000).
- [58] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *Phys. Rev. D* **63**, 114020 (2001).
- [59] C. W. Bauer, D. Pirjol, and I. W. Stewart, *Phys. Rev. D* **65**, 054022 (2002).
- [60] C. W. Bauer and I. W. Stewart, *Phys. Lett. B* **516**, 134 (2001).
- [61] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, *Phys. Rev. D* **66**, 014017 (2002).
- [62] E. Eichten and B. R. Hill, *Phys. Lett. B* **234**, 511 (1990).
- [63] N. Isgur and M. B. Wise, *Phys. Lett. B* **232**, 113 (1989).
- [64] N. Isgur and M. B. Wise, *Phys. Lett. B* **237**, 527 (1990).
- [65] B. Grinstein, *Nucl. Phys.* **B339**, 253 (1990).
- [66] H. Georgi, *Phys. Lett. B* **240**, 447 (1990).
- [67] M. Dasgupta and G. P. Salam, *Phys. Lett. B* **512**, 323 (2001).
- [68] M. Dasgupta and G. P. Salam, *J. High Energy Phys.* **08** (2002) 032.
- [69] A. Jain, I. Scimemi, and I. W. Stewart, *Phys. Rev. D* **77**, 094008 (2008).
- [70] P. F. Monni, T. Gehrmann, and G. Luisoni, *J. High Energy Phys.* **08** (2011) 010.
- [71] R. Kelley, M. D. Schwartz, R. M. Schabinger, and H. X. Zhu, *Phys. Rev. D* **84**, 045022 (2011).
- [72] A. H. Hoang, A. Pathak, P. Pietrulewicz, and I. W. Stewart, *J. High Energy Phys.* **12** (2015) 059.
- [73] J. M. Henn, G. P. Korchemsky, and B. Mistlberger, *J. High Energy Phys.* **04** (2020) 018.
- [74] A. H. Hoang, C. Lepenik, and M. Stahlhofen, *J. High Energy Phys.* **08** (2019) 112.
- [75] T. Matsuura and W. L. van Neerven, *Z. Phys. C* **38**, 623 (1988).
- [76] T. Matsuura, S. C. van der Marck, and W. L. van Neerven, *Nucl. Phys.* **B319**, 570 (1989).

- [77] T. Gehrmann, T. Huber, and D. Maitre, *Phys. Lett. B* **622**, 295 (2005).
- [78] S. Moch, J. A. M. Vermaseren, and A. Vogt, *J. High Energy Phys.* **08** (2005) 049.
- [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] M. E. Luke and A. V. Manohar, *Phys. Lett. B* **286**, 348 (1992).
- [82] M. Preisser, Jet shapes with massive quarks for e^+e^- annihilation, Ph.D. thesis, Vienna University, 2018.
- [83] V. Mateu and P. G. Ortega, *J. High Energy Phys.* **01** (2018) 122.
- [84] A. H. Hoang and S. Kluth, [arXiv:0806.3852](https://arxiv.org/abs/0806.3852).
- [85] A. Hornig, C. Lee, I. W. Stewart, J. R. Walsh, and S. Zuberi, *J. High Energy Phys.* **08** (2011) 054.
- [86] G. P. Korchemsky and S. Tafat, *J. High Energy Phys.* **10** (2000) 010.
- [87] Z. Ligeti, I. W. Stewart, and F. J. Tackmann, *Phys. Rev. D* **78**, 114014 (2008).
- [88] F. U. Bernlochner, H. Lacker, Z. Ligeti, I. W. Stewart, F. J. Tackmann, and K. Tackmann (SIMBA Collaboration), [arXiv:2007.04320](https://arxiv.org/abs/2007.04320).
- [89] E. Gardi, *J. High Energy Phys.* **04** (2000) 030.
- [90] R. A. Davison and B. R. Webber, *Eur. Phys. J. C* **59**, 13 (2009).
- [91] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Phys. Rev. D* **83**, 074021 (2011).
- [92] K. Chetyrkin, B. A. Kniehl, and M. Steinhauser, *Nucl. Phys.* **B510**, 61 (1998).
- [93] K. Chetyrkin, J. H. Kuhn, and C. Sturm, *Nucl. Phys.* **B744**, 121 (2006).
- [94] Y. Schroder and M. Steinhauser, *J. High Energy Phys.* **01** (2006) 051.
- [95] S. A. Larin and J. A. M. Vermaseren, *Phys. Lett. B* **303**, 334 (1993).
- [96] M. Czakon, *Nucl. Phys.* **B710**, 485 (2005).
- [97] A. Czarnecki, K. Melnikov, and N. Uraltsev, *Phys. Rev. Lett.* **80**, 3189 (1998).
- [98] M. Beneke, *Phys. Lett. B* **434**, 115 (1998).
- [99] A. H. Hoang, Z. Ligeti, and A. V. Manohar, *Phys. Rev. Lett.* **82**, 277 (1999).
- [100] A. H. Hoang, Z. Ligeti, and A. V. Manohar, *Phys. Rev. D* **59**, 074017 (1999).
- [101] A. Hoang, *Phys. Rev. D* **61**, 034005 (1999).
- [102] A. Pineda, *J. High Energy Phys.* **06** (2001) 022.
- [103] G. P. Salam and D. Wicke, *J. High Energy Phys.* **05** (2001) 061.
- [104] C. Lee and G. Sterman, *Phys. Rev. D* **75**, 014022 (2007).
- [105] L. G. Almeida, S. D. Ellis, C. Lee, G. Sterman, I. Sung, and J. R. Walsh, *J. High Energy Phys.* **04** (2014) 174.
- [106] B. Dehnadi, Heavy quark mass determinations with sum rules and jets, Ph.D. thesis, Vienna University, 2016.
- [107] *GFortran, Gnu compiler collection (gcc), Version 5.0.0* (Copyright (C) 2014 Free Software Foundation, Inc., 2014).
- [108] M. A. Ebert, J. K. L. Michel, F. J. Tackmann *et al.*, DESY-17-099 (2018), webpage: <http://scetlib.desy.de>.
- [109] R. Piessens, E. d. Kapenga, C. Uberhuber, and D. Kahaner, *QUADPACK: A Subroutine Package for Automatic Integration* (Springer-Verlag, Berlin, 1983).
- [110] G. P. Contributors, GSL—GNU scientific library—GNU project—free software foundation (FSF),” é (2010).
- [111] A. Pathak (to be made public) (2020).
- [112] G. P. Korchemsky and G. Marchesini, *Phys. Lett. B* **313**, 433 (1993).
- [113] C. Balzereit, T. Mannel, and W. Kilian, *Phys. Rev. D* **58**, 114029 (1998).
- [114] M. Neubert, *Eur. Phys. J. C* **40**, 165 (2005).
- [115] G. P. Korchemsky and A. V. Radyushkin, *Nucl. Phys.* **B283**, 342 (1987).
- [116] S. Moch, J. A. M. Vermaseren, and A. Vogt, *Nucl. Phys.* **B688**, 101 (2004).
- [117] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, *Phys. Lett. B* **782**, 627 (2018).
- [118] O. V. Tarasov, A. A. Vladimirov, and A. Y. Zharkov, *Phys. Lett.* **93B**, 429 (1980).
- [119] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, *Phys. Lett. B* **400**, 379 (1997).
- [120] W. L. van Neerven, *Nucl. Phys.* **B268**, 453 (1986).
- [121] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, *Nucl. Phys.* **B407**, 3 (1993).
- [122] A. Vogt, *Phys. Lett. B* **497**, 228 (2001).
- [123] A. Idilbi, X.-d. Ji, and F. Yuan, *Nucl. Phys.* **B753**, 42 (2006).
- [124] T. Becher, M. Neubert, and B. D. Pecjak, *J. High Energy Phys.* **01** (2007) 076.
- [125] R. N. Lee, A. V. Smirnov, and V. A. Smirnov, *J. High Energy Phys.* **04** (2010) 020.
- [126] R. Tarrach, *Nucl. Phys.* **B183**, 384 (1981).
- [127] N. Gray, D. J. Broadhurst, W. Grafe, and K. Schilcher, *Z. Phys. C* **48**, 673 (1990).
- [128] K. G. Chetyrkin and M. Steinhauser, *Phys. Rev. Lett.* **83**, 4001 (1999).
- [129] K. G. Chetyrkin and M. Steinhauser, *Nucl. Phys.* **B573**, 617 (2000).
- [130] K. Melnikov and T. v. Ritbergen, *Phys. Lett. B* **482**, 99 (2000).
- [131] P. Marquard, L. Mihaila, J. H. Piclum, and M. Steinhauser, *Nucl. Phys.* **B773**, 1 (2007).
- [132] A. H. Hoang, A. Jain, I. Scimemi, and I. W. Stewart, *Phys. Rev. D* **82**, 011501 (2010).