# $V/H+$ jet production with the Cambridge/Aachen algorithm

K. Khelifa-Kerfa*

*Department of Physics and Materials Sciences, College of Arts and Sciences, Qatar University, P.O. Box 2713, Doha, Qatar; Department of Physics, Faculty of Science and Technology, Relizane University, Relizane 48000, Algeria; Laboratory of Physics of Experimental Techniques and Applications, University of Medea, Medea 26000, Algeria; Laboratory of Thin Layers and Advanced Technologies, Relizane University, Relizane 48000, Algeria and Laboratory of Mathematics and Applications, University of Chlef, Chlef 02000, Algeria*

(Received 18 July 2026; accepted 6 October 2026; published 26 October 2026)

We present fixed-order perturbative calculations up to four-loop order for a generic nonglobal QCD observable in hadron-hadron collisions. Specifically, we study the invariant-mass distribution of the highest-$p_t$ jet produced in association with a vector boson or a Higgs boson, where jets are defined using the Cambridge/Aachen sequential recombination algorithm. This work is part of a series of papers [K. Khelifa-Kerfa and Y. Delenda, J. High Energy Phys. 03 (2015) 094; K. Khelifa-Kerfa, Eur. Phys. J. C **85**, 133 (2025); K. Khelifa-Kerfa, Phys. Rev. D **111**, 094018 (2025); K. Khelifa-Kerfa and M. Benghanem, Eur. Phys. J. C **85**, 845 (2025); K. Khelifa-Kerfa, J. High Energy Phys. 10 (2024) 079; K. Khelifa-Kerfa and M. Benghanem, J. High Energy Phys. 02 (2026) 009.]examining the impact of various jet algorithms on the perturbative structure of nonglobal observables across different collision environments. We find that, at fixed order, the Cambridge/Aachen algorithm typically suppresses the large nonglobal logarithms more effectively than the anti-$k_t$ and $k_t$ algorithms; a notable exception occurs at four-loop order for the purely gluonic channel $gg \to gH$, for which the combined clustering and nonglobal logarithm contribution in Cambridge/Aachen clustering exceeds that of $k_t$ over a wide range of jet radii. At all orders, the Cambridge/Aachen algorithm performs comparably to the $k_t$ algorithm. Furthermore, comparisons of all-orders resummed form factors reveal that finite-$N_c$ corrections remain at the percent level, consistent with previous findings.

DOI: 10.1103/gzgr-gf82

## I. INTRODUCTION

Precision calculations of jet observables at hadron colliders, particularly at the Large Hadron Collider (LHC) at CERN, demand an increasingly refined understanding of QCD radiation patterns. Among the most phenomenologically important of these observables is the invariant mass of a jet produced in association with an electroweak boson or a Higgs boson—a configuration that serves simultaneously as a stringent test of the Standard Model (SM) and as a sensitive probe of new physics through its jet substructure [1,2]. The invariant jet-mass distribution is especially valuable because it encodes information about color flow, underlying-event activity, hadronization, and the interplay of multiple logarithmic hierarchies—all of which are directly relevant to precision measurements at the LHC.

In the limit of small normalized jet mass, $m_j^2/p_t^2 \ll 1$, fixed-order perturbation theory becomes inadequate owing to the proliferation of logarithmically enhanced terms. These arise from multiple sources. The first category, commonly termed global logarithms, comprises contributions from soft-collinear and hard-collinear radiation off all incoming and outgoing hard partons; upon resummation these exponentiate into the well-known Sudakov form factor. The second, phenomenologically more challenging, category comprises *nonglobal logarithms* (NGLs) [3,4], which arise from soft-gluon correlations connecting spatially separated phase-space regions—specifically, secondary emissions from inside the jet that are induced by harder radiation deposited outside. NGLs are a manifestation of the genuinely non-Abelian color dynamics of QCD and resist straightforward exponentiation. Despite sustained theoretical efforts [5–12], their complete resummation remains an open problem, particularly beyond the leading-color approximation.

*Contact author: kkerfa@qu.edu.qa

*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.*

A third class of large logarithms, known as clustering logarithms (CLs), arises whenever jets are defined by sequential recombination algorithms other than the anti-$k_t$ one [13]. First recognized in the context of the $k_t$ algorithm by Banfi and Dasgupta [14], and investigated further in Refs. [15,16], these contributions stem from the algorithm-induced recombination of Abelian, independent primary gluon emissions. Such recombination distorts the real-virtual cancellations that ordinarily occur for primary emissions, thereby introducing a new tower of single logarithms at each perturbative order. Crucially, whereas clustering suppresses the magnitude of the standard NGLs—by preventing certain gluon configurations from contributing to the jet mass—it simultaneously generates CLs [17–19], so that the total logarithmic content of the observable cannot be reduced simply by a judicious choice of algorithm.

The Cambridge/Aachen (C/A) algorithm [20,21] occupies a distinguished position among sequential recombination schemes. Unlike the $k_t$ algorithm [22,23]—whose distance measure weights angular separations by the smaller of the two particles' transverse momenta, causing the softest emissions to merge first—the C/A algorithm employs a purely geometric distance measure that is independent of any energy scale. This angular-ordered clustering sequence closely mirrors the structure of QCD coherence [24] and is widely exploited in jet substructure studies, particularly in the context of jet grooming [2]. It is also the structural foundation of modern jet-grooming techniques: since the C/A clustering history is purely angular-ordered, it can be traversed backward from wide angle to narrow angle, providing the well-defined declustering sequence on which groomers such as the modified mass-drop tagger and soft drop operate [2,25,26]. Furthermore, among the three members of the generalized sequential recombination family—anti-$k_t$ [13], $k_t$, and C/A—the C/A algorithm has been shown to yield the smallest nonglobal logarithmic contributions [17,18,27], making it theoretically attractive for analyses where logarithmic accuracy is paramount.

Analytic understanding of NGLs and CLs in hadronic processes has advanced significantly in recent years. Fixed-order calculations of NGLs for the jet-mass distribution in $Z+$ jet and dijet production at two loops were presented in [28], while the companion paper [29] extended these to include both NGLs and CLs in the $k_t$ and C/A algorithms, also at two loops, for Higgs- and vector-boson associated jet production. multiloop analyses were subsequently carried out for $e^+e^-$ annihilation processes [5,7,19,30–32], revealing that the fixed-order CL series exponentiates and that the perturbative expansion converges rapidly. For hadronic collisions, a systematic study of NGLs up to four loops with full color dependence and complete jet-radius dependence in the anti-$k_t$ algorithm was reported in [33] for $V/H+$ jet processes, while the companion work [34] extended these calculations to the $k_t$ algorithm, providing the first four-loop results encompassing both CLs and NGLs for the jet mass in a hadronic environment.

The present paper continues this program by extending the work of [34] to the Cambridge/Aachen algorithm. We compute the complete fixed-order distribution of the normalized invariant mass squared of the leading jet in $V/H+$ jet production, incorporating contributions from both CLs and NGLs through four-loop order in perturbation theory. Our calculations are carried out in the eikonal approximation with strong energy ordering among the final-state soft gluons, ensuring single-logarithmic accuracy. By retaining the full dependence on color factors and the jet-radius parameter $R$ for all relevant partonic channels, we provide semianalytical expressions that are directly applicable to phenomenological studies.

Several notable features distinguish the C/A case from both the anti-$k_t$ and $k_t$ results. First, as is also the case at two loops, the C/A algorithm yields NGL and CL coefficients that are identical to those of $k_t$ clustering. Second, and in contrast to the $k_t$ situation studied in [34], the purely geometric nature of the C/A distance measure introduces distinct higher-loop clustering structures starting at three-loop order, which we derive and analyze in detail. The resulting NGL and CL contributions are smaller in magnitude than in the $k_t$ case, and smaller still compared to the anti-$k_t$ case. Third, we find that the NGL and CL coefficients exhibit the universal "boundary" or "edge" effect—approaching finite nonzero constants as $R \to 0$—consistent with observations in $e^+e^-$ calculations [18,19,30] and in earlier hadronic results [29,33,34]. Fourth, the all-orders analysis of C/A clustering shows that it produces nonglobal effects that are almost identical in size to the $k_t$ case, particularly for values of the jet-mass observable where the exponentiation of the fixed-order results is reliable. This observation is further confirmed by a detailed study of the specific $Z+$ jet process. These features make the C/A and $k_t$ algorithms preferable to anti-$k_t$ for mitigating the impact of nonglobal logarithms. Fifth, the finite-$N_c$ corrections are shown to remain at the percent level for phenomenologically relevant values of the observable. This is in excellent agreement with earlier results reported for both lepton and hadron processes (see, for instance, [8,33,35]).

It is worth noting that, unlike the cases of the anti-$k_t$ and $k_t$ algorithms, the all-orders numerical resummation of nonglobal observables defined with the C/A algorithm is severely limited by the availability of dedicated codes. The only available code for C/A clustering is that of Ref. [36], which, however, is restricted to observables with purely soft singularities, such as the energy-flow (or gap-between-jets) observable. The invariant jet-mass observable exhibits both soft and collinear singularities and therefore falls outside the scope of that code. The Dasgupta-Salam Monte Carlo (MC) code [3], upon which the code of Ref. [36] is based, covers observables with both soft and collinear singularities but is restricted to the anti-$k_t$ and $k_t$ algorithms, with no

implementation for C/A. Consequently, no all-orders numerical results are available against which to benchmark our analytical C/A calculations; we therefore compare instead to the $k_t$ results produced by the MC code of [3].

Throughout the calculation we treat super-leading logarithms (SLLs) [12,37,38] as decoupled from the standard NGL and CL contributions at the loop orders considered here; any potential interleaving between these two classes first enters at five-loop order for $2 \to 1$ processes [12,33,38].

The paper is organized as follows. Section II introduces the kinematic framework, the observable definition, the C/A jet algorithm, and the structure of the resummed jet-mass distribution. Section III presents the three- and four-loop calculations of CLs and NGLs for all partonic channels. Section IV compares the fixed-order analytical predictions against all-orders numerical resummation from the Monte Carlo code of [3]. Section V summarizes our conclusions and discusses prospects for further work.

## II. DEFINITIONS

### A. Kinematics

We study the production of a single hard jet in association with a Higgs boson $H$ or a vector boson $V \in \{Z, W^{\pm}, \gamma\}$ at a proton-proton collider. The relevant partonic Born channels are identical to those discussed in [33,34]: for vector-boson-associated production there are two topologically distinct processes,

$$(\delta_1)\colon\ q\bar{q} \to gV/H, \qquad (\delta_2)\colon\ qg \to qV/H, \tag{1a}$$

while Higgs production acquires an additional all-gluon channel,

$$(\delta_3)\colon\ gg \to gH. \tag{1b}$$

The channel $\bar{q}g \to \bar{q}V/H$ is color-equivalent to $(\delta_2)$ and need not be tracked separately. From the perspective of our soft-gluon calculations, all three channels involve exactly three hard colored legs together with a color-singlet boson; they differ only through their Born cross sections and the corresponding color Casimir scalars. The Higgs-boson channel $gg \to gH$ proceeds via an effective vertex encoding the heavy top-quark loop.

We denote the four-momenta of the incoming partons by $p_a$ and $p_b$, the outgoing hard parton by $p_j$, and those of the subsequently emitted soft gluons by $k_i$ $(i = 1, \ldots, m)$. In the center-of-mass frame these are parametrized as

$$p_a = x_a \frac{\sqrt{s}}{2}(1,0,0,1), \qquad p_b = x_b \frac{\sqrt{s}}{2}(1,0,0,-1), \tag{2a}$$

and

$$p_j = p_t(\cosh y, \cos\varphi, \sin\varphi, \sinh y), \tag{2b}$$

$$k_i = k_{ti}(\cosh\eta_i, \cos\phi_i, \sin\phi_i, \sinh\eta_i), \tag{2c}$$

where $\sqrt{s}$ is the partonic center-of-mass energy; $x_a$ and $x_b$ are the momentum fractions carried by the colliding partons, described by standard parton distribution functions currently available through next-to-next-to-next-to-leading-order ($\mathrm{N^3LO}$) accuracy [39]; $p_t$, $y$, and $\varphi$ are the transverse momentum, rapidity, and azimuthal angle of the outgoing hard parton with respect to the beam axis; and $k_{ti}$, $\eta_i$, and $\phi_i$ are the corresponding kinematic quantities for the $i$th soft gluon. All hard partons are treated as massless; extensions incorporating heavy-quark mass effects are reserved for future work. Recoil corrections enter at next-to-next-to-leading logarithmic accuracy and are accordingly neglected throughout [40].

To facilitate the angular integrations, it is convenient to introduce polar variables $(r_i, \theta_i)$ centered on the hard-jet axis:

$$\eta_i - y = R r_i \cos\theta_i, \qquad \phi_i - \varphi = R r_i \sin\theta_i, \tag{3}$$

with $r_i > 0$ and $0 < \theta_i < 2\pi$, so that $\mathrm{d}\eta_i \mathrm{d}\phi_i = R^2 r_i \mathrm{d}r_i \mathrm{d}\theta_i$.

The one-loop dipole antenna function, encoding the universal pattern of soft-gluon radiation from a color dipole $(\alpha, \beta)$, is given by

$$w^i_{\alpha\beta} = \frac{k_{ti}^2}{2} \frac{(p_\alpha \cdot p_\beta)}{(p_\alpha \cdot k_i)(k_i \cdot p_\beta)}. \tag{4}$$

The corresponding color factors for the three Born dipoles $\Delta_\delta = \{(aj), (bj), (ab)\}$ in channel $\delta$ read [41,42]

$$\mathcal{C}_{q\bar{q}} = \mathcal{C}_{qq} = 2\mathrm{C_F} - \mathrm{C_A}, \qquad \mathcal{C}_{qg} = \mathcal{C}_{gg} = \mathrm{C_A}, \tag{5}$$

where the color Casimir scalars are $\mathrm{C_F} = (N_c^2 - 1)/(2N_c)$ and $\mathrm{C_A} = N_c$, with $N_c = 3$ in QCD.

### B. Observable and jet algorithm

The observable under study is the normalized invariant mass squared of the leading hard jet,

$$\varrho = \frac{m_j^2}{p_t^2} = \frac{1}{p_t^2}\left(p_j + \sum\nolimits_{i \in j} k_i\right)^2 = \sum_{i \in j} \varrho_i + \mathcal{O}\left(\frac{k_t^2}{p_t^2}\right), \tag{6}$$

where the sum runs over all soft gluons recombined into the jet by the clustering procedure. The individual contribution of gluon $k_i$ to $\varrho$ is

$$\varrho_i = \frac{2(p_j \cdot k_i)}{p_t^2} = 2\xi_i[\cosh(\eta_i - y) - \cos(\phi_i - \varphi)], \tag{7}$$

where $\xi_i \equiv k_{ti}/p_t$ and terms of order $k_t^2/p_t^2$ have been dropped in the eikonal limit. Expanding in powers of the jet

radius $R$ via the parametrization (3) yields

$$\varrho_i = \xi_i \left[ R^2 r_i^2 + \frac{R^4 r_i^4}{12} \cos(2\theta_i) + \cdots \right], \tag{8}$$

and retaining only the leading piece—which contributes at single-logarithmic accuracy—gives

$$\varrho_i \simeq R^2 \xi_i r_i^2. \tag{9}$$

The Cambridge/Aachen (C/A) jet algorithm [20,21] is a member of the inclusive generalized sequential recombination family [43], parametrized by the interparticle and beam distances

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}, \qquad d_{iB} = 1, \tag{10}$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$. The algorithm proceeds iteratively: at each step one identifies the smallest distance among all $\{d_{ij}, d_{iB}\}$; if it corresponds to a pair $(i, j)$, the two particles are merged into a pseudojet via the $E$-scheme recombination, in which the four-momenta of the two particles are added directly; if it corresponds to a beam distance $d_{iB}$, particle $i$ is promoted to a final-state jet and removed from the list. This cycle continues until no particles remain.

The defining characteristic of the C/A algorithm—distinguishing it from both the $k_t$ [22,23] and anti-$k_t$ [13] schemes—is that its distance measure is entirely geometric, carrying no dependence on the particles' transverse momenta. Consequently, the clustering sequence is governed purely by angular proximity: the pair with the smallest angular separation is always merged first, irrespective of the relative softness of the two constituents.

By contrast, the $k_t$ algorithm weights angular separations by $\min(p_{ti}^2, p_{tj}^2)$, ensuring that the softest particles cluster earliest. In the eikonal approximation with strong energy ordering $\xi_1 \gg \xi_2 \gg \cdots \gg \xi_m$, the pairwise condition for two soft gluons to merge before either is individually declared a jet takes the same geometric form for both the $k_t$ and C/A algorithms:

$$\Delta R_{ij}^2 < R^2 \Leftrightarrow r_i^2 + r_j^2 - 2 r_i r_j \cos(\theta_i - \theta_j) < 1. \tag{11}$$

The two algorithms therefore produce identical results at two-loop order, where only a single pair of soft gluons needs to be considered. The genuinely distinct behavior of the C/A scheme emerges at three-loop order and beyond, where the ordering of clustering steps in a multiparticle configuration can differ significantly. Specifically, in the $k_t$ algorithm, the softest emission always clusters first with its nearest neighbor—the energy-ordered clustering priority—while in C/A the angularly closest pair clusters first regardless of the energy hierarchy. This distinction generates different irreducible clustering structures at three-loop order and beyond, which are the primary focus of the present work.

### C. Observable distribution

At next-to-leading logarithmic (NLL) accuracy, the differential jet-mass cross section in a partonic channel $\delta$ takes the factorized form [28,29,33]

$$\frac{\mathrm{d}\Sigma_\delta(\rho)}{\mathrm{d}\mathcal{B}_\delta} = \int_0^\rho \frac{\mathrm{d}^2\sigma_\delta}{\mathrm{d}\mathcal{B}_\delta \, \mathrm{d}\varrho} \mathrm{d}\varrho, \tag{12}$$

where $\rho$ is the jet-mass veto, $\mathcal{B}_\delta$ encodes the Born phase space including parton distribution functions and the hard-scattering matrix element squared, and $\mathrm{d}\sigma_\delta$ is the fully differential partonic cross section for channel $\delta$. Throughout this paper we distinguish between the integration variable $\varrho$ and the jet-mass veto $\rho$, with $\varrho \le \rho$. The integrated distribution, summed over all contributing channels, is then

$$\Sigma(\rho) = \sum_\delta \int \mathrm{d}\mathcal{B}_\delta \frac{\mathrm{d}\Sigma_\delta(\rho)}{\mathrm{d}\mathcal{B}_\delta} \Xi_\mathcal{B}, \tag{13}$$

where $\Xi_\mathcal{B}$ represents a set of Born-level kinematic cuts. In the region of small jet mass, $\rho \ll 1$, the cross section is dominated by logarithmically enhanced contributions and may be written as

$$\frac{\mathrm{d}\Sigma_\delta(\rho)}{\mathrm{d}\mathcal{B}_\delta} = \frac{\mathrm{d}\sigma_{0,\delta}}{\mathrm{d}\mathcal{B}_\delta} f_{\mathcal{B},\delta}(\rho)[1 + \mathcal{O}(\alpha_s)], \tag{14}$$

where $\mathrm{d}\sigma_{0,\delta}/\mathrm{d}\mathcal{B}_\delta$ is the Born-level partonic differential cross section and the function $f_{\mathcal{B},\delta}(\rho)$ resums the hierarchy of large logarithms.

For the C/A algorithm, this resummed factor decomposes as [27,29,33,34,40]

$$f_{\mathcal{B},\delta}(\rho) = f_{\mathcal{B},\delta}^{\mathrm{global}}(\rho) \mathcal{S}_\delta(\rho) \mathcal{C}_\delta(\rho), \tag{15}$$

where $f_{\mathcal{B},\delta}^{\mathrm{global}}(\rho)$ is the algorithm-independent Sudakov form factor arising from the exponentiation of a single soft-gluon emission—accounting for soft-collinear, hard-collinear, and soft-wide-angle primary radiations from all incoming and outgoing hard partons. This global factor has been computed at NLL accuracy for the same class of observables in [28,29] and will not be rederived here; we refer the reader to those works for explicit expressions.

The two remaining factors in (15) are algorithm-dependent and constitute the main objects of study in this paper:

(i) $\mathcal{S}_\delta(\rho)$ encodes the resummation of leading non-global logarithms arising from correlated secondary gluon emissions, modulated by the C/A clustering constraints on the available phase space.

(ii) $\mathcal{C}_\delta(\rho)$ captures the resummation of clustering logarithms generated by Abelian, independent primary emissions that are reshuffled into or out of the jet by the C/A recombination procedure.

No closed-form analytical all-orders expressions are available for either $\mathcal{S}_\delta$ or $\mathcal{C}_\delta$ with C/A clustering; the present paper provides the first systematic fixed-order expansion of both quantities through four-loop order in the hadronic environment. We expand

$$f_{\mathcal{B},\delta}(\rho) = 1 + f^{(1)}_{\mathcal{B},\delta}(\rho) + f^{(2)}_{\mathcal{B},\delta}(\rho) + f^{(3)}_{\mathcal{B},\delta}(\rho) + f^{(4)}_{\mathcal{B},\delta}(\rho) + \cdots, \tag{16}$$

where $f^{(n)}_{\mathcal{B},\delta}(\rho)$ denotes the $n$-loop contribution of order $\alpha_s^n$ and receives separate CL and NGL pieces,

$$f^{(n)}_{\mathcal{B},\delta}(\rho) \supset \mathcal{C}_{n,\delta}(\rho) + \mathcal{S}_{n,\delta}(\rho), \qquad n \geq 2. \tag{17}$$

The $m$-loop contribution to the jet-mass fraction is expressed via the measurement operator $\hat{\mathcal{U}}_m$, first introduced in [7] and developed further in [5,32], as

$$f^{(m)}_{\mathcal{B},\delta}(\rho) = \sum_X \int_{\xi_1>\xi_2>\cdots>\xi_m} \left(\prod_{i=1}^m \mathrm{d}\Phi_i\right) \hat{\mathcal{U}}_m \mathcal{W}^X_{1\cdots m,\delta} \times \Xi^{(\mathrm{C/A})}_m(k_1,\ldots,k_m), \tag{18}$$

where $\mathcal{W}^X_{1\cdots m,\delta}$ is the eikonal amplitude squared for $m$ strongly energy-ordered soft gluons in configuration $X$ for channel $\delta$ [41,42], the phase-space element for gluon $i$ is

$$\mathrm{d}\Phi_i = \bar{\alpha}_s \frac{\mathrm{d}\xi_i}{\xi_i}\mathrm{d}\eta_i \frac{\mathrm{d}\phi_i}{2\pi} = \bar{\alpha}_s \frac{\mathrm{d}\xi_i}{\xi_i} R^2 r_i \mathrm{d}r_i \frac{\mathrm{d}\theta_i}{2\pi}, \tag{19}$$

with $\bar{\alpha}_s \equiv \alpha_s/\pi$, and $\Xi^{(\mathrm{C/A})}_m(k_1,\ldots,k_m)$ is the C/A clustering function encoding the phase-space constraints imposed by the algorithm on the $m$-gluon configuration. The measurement operator factorizes at each loop order as $\hat{\mathcal{U}}_m = \prod_{i=1}^m \hat{u}_i$, with

$$\hat{u}_i = 1 - \Theta^{\mathrm{R}}_i \Theta^{\rho}_i \Theta^{\mathrm{in}}_i, \tag{20}$$

where $\Theta^{\mathrm{R}}_i = 1$ for a real emission and zero otherwise, $\Theta^{\mathrm{in}}_i = \Theta(1-r_i^2)$ enforces that gluon $i$ lies inside the jet after clustering, and $\Theta^{\rho}_i = \Theta(\varrho_i - \rho)$ restricts to configurations in which gluon $i$'s contribution to the jet mass exceeds the veto $\rho$.

Strong ordering of the gluon transverse momenta, $\xi_1 \gg \xi_2 \gg \cdots \gg \xi_m$, greatly simplifies the clustering sequence: the softer of any two particles always satisfies the C/A interparticle distance condition before the beam distance, so the C/A clustering function $\Xi^{(\mathrm{C/A})}_m$ reduces to a product of step functions in the polar variables $(r_i, \theta_i)$. The explicit forms of these functions at three- and four-loop order, together with their interplay with the eikonal amplitudes squared, are derived in the following section.

Explicit details of the Born cross sections, parton distribution functions, and the global Sudakov factor are presented in [28,29] and will not be reproduced here. In the following, we concentrate entirely on the perturbative computation of $\mathcal{C}_\delta(\rho)$ and $\mathcal{S}_\delta(\rho)$ up to four-loop order.

## III. FIXED-ORDER CALCULATIONS

It is straightforward to verify from the definition of the sequential recombination algorithms—as presented, for instance, in Ref. [43]—that all jet algorithms perform identically at one-loop order. Equation (18) gives, up to NLL accuracy [29,33,34],

$$f^{(1)}_{\mathcal{B},\delta}(\rho) = -\bar{\alpha}_s \left[ (\mathcal{C}_{aj} + \mathcal{C}_{bj}) \frac{L^2}{4} + \left( \mathcal{C}_{ab} \frac{R^2}{2} + (\mathcal{C}_{aj} + \mathcal{C}_{bj}) h(R) \right) L \right], \tag{21}$$

where $L \equiv \ln(R^2/\rho)$, $h(R) = R^2/8 + R^4/576 + \mathcal{O}(R^6)$, and the color Casimir scalars are defined in Eq. (5). The exponentiation of this one-loop result, after accounting for the running coupling, produces the resummed Sudakov form factor $f^{\mathrm{global}}_{\mathcal{B},\delta}(\rho)$ discussed in the paragraph following Eq. (15); explicit expressions can be found in Refs. [28,29].

At two-loop order, the $k_t$ and C/A algorithms perform identically, but differently from the anti-$k_t$ algorithm. For completeness, we recall here the two-loop results for the $k_t$ case from our previous work [34]. The invariant jet-mass distribution at this order takes the form

$$f^{(2)}_{\mathcal{B},\delta}(\rho) = \frac{1}{2!}\left[f^{(1)}_{\mathcal{B},\delta}(\rho)\right]^2 + \mathcal{C}_{2,\delta}(\rho) + \mathcal{S}_{2,\delta}(\rho), \tag{22}$$

where the last two terms are the two-loop CL and NGL contributions, respectively. They take the form

$$\mathcal{C}_{2,\delta}(\rho) = \frac{1}{2!}\bar{\alpha}_s^2 L^2 \mathcal{F}_{2,\delta}(R), \qquad \mathcal{S}_{2,\delta}(\rho) = -\frac{1}{2!}\bar{\alpha}_s^2 L^2 \mathcal{G}_{2,\delta}(R), \tag{23}$$

where the $R$-dependent two-loop coefficients $\mathcal{F}_{2,\delta}(R)$ and $\mathcal{G}_{2,\delta}(R)$ are given as power series in $R$ in Eqs. (3.23) and (3.28) of [34] for all three partonic channels. The first distinction between the $k_t$ and C/A algorithms arises at three-loop order, which is the subject of the next section.

### A. Three-loop order

The three-loop eikonal amplitude, summed over all possible gluon configurations $X$ for a specific partonic channel $\delta$, can be deduced from that of $e^+e^-$ annihilation

processes, presented in detail in our previous paper [27], following the procedure outlined in Ref. [41]. It reads

$$\sum_X \hat{\mathcal{U}}_3 \mathcal{W}^{X,\mathrm{C/A}}_{123,\delta} = \sum_X \hat{\mathcal{U}}_3 \mathcal{W}^{X,\mathrm{k_t}}_{123,\delta} - \left(\prod_{i=1}^{3} \Theta_i^{\rho}\right) \Theta_1^{\mathrm{out}} \Theta_3^{\mathrm{in}} \Omega_{12}\Omega_{23}\bar{\Omega}_{13}\Delta_{123} \times \left[\mathcal{W}^{\mathrm{VVR}}_{123,\delta} + \mathcal{W}^{\mathrm{RVR}}_{123,\delta} + \mathcal{W}^{\mathrm{RRR}}_{123,\delta}\right], \tag{24}$$

where the first term on the right-hand side (rhs) is the result for the $k_t$ jet algorithm, reported in Eq. (3.29) of [34], $\Omega_{ij} = \Theta(d_{jB} - d_{ij})$, and $\Delta_{ijk} = \Theta(d_{jk} - d_{ij})$ with $i < j$. When the constraint $\Omega_{ij} = 1$, the pair of gluons $(i, j)$ is clustered first into a pseudojet with a four-momentum that lies along that of the harder of the two particles. If, instead, $\Omega_{ij} = 0$—and hence $\bar{\Omega}_{ij} \equiv 1 - \Omega_{ij} = 1$—the pair of gluons $(i, j)$ is not clustered together and each remains at its initial location prior to the application of the C/A algorithm. In both cases, the phase space is nonetheless modified by the algorithm relative to its unclustered form.

When $\Delta_{ijk} = 1$, the pair of particles $(i, j)$ is clustered first, having the smallest distance. The opposite case, $\Delta_{ijk} = 0$, corresponds to the pair $(j, k)$ being clustered first. If, for example, both $\Delta_{123}$ and $\Delta_{231}$ are equal to unity, there are two competing small distances, $d_{12}$ and $d_{23}$, that must be manually ordered for the analytical calculation to cover the full available phase space: (a) $d_{12} < d_{23}$: gluons $k_1$ and $k_2$ are clustered first, and by strong energy ordering the resulting pseudojet lies along the harder gluon $k_1$; gluon $k_3$ is then left unclustered; (b) $d_{23} < d_{12}$: gluons $k_2$ and $k_3$ are clustered first into a pseudojet along the harder gluon $k_2$, which is in turn clustered with $k_1$; gluon $k_1$ is then left unclustered. This simple example illustrates how the ordering of the distance measures can alter the final gluon configuration, and hence each gluon's contribution to the observable distribution. Such an ordering requirement is absent for both the anti-$k_t$ and $k_t$ algorithms, since their distance measures are automatically ordered by the transverse momenta that already appear in their definitions.

It is worth noting that, in Eq. (24), the location of gluon $k_2$ is not explicitly shown and therefore does not play a decisive role in the C/A clustering correction. This can be explained straightforwardly as follows: from the product of the constraints $\Omega_{12}$ and $\Omega_{23}$ it is evident that gluon $k_2$ will always be clustered either with the hardest gluon $k_1$ or the softest gluon $k_3$. In the first case, $d_{12} < d_{23}$, $k_2$ is dragged by $k_1$ regardless of its initial position (in or out of the jet), and the resulting pseudojet is essentially $k_1$ by strong energy ordering; one is then left with only gluons $k_1$ and $k_3$ to determine whether the configuration contributes to the jet mass. In the second case, $d_{12} > d_{23}$, $k_2$ drags $k_3$ into a pseudojet that lies essentially along $k_2$, which is itself then dragged by the harder gluon $k_1$, yielding a final pseudojet along $k_1$. In both cases, $k_2$ plays the role of transporting $k_3$ toward or away from $k_1$ depending on their relative angular separation—a condition that is independent of whether $k_2$ lies inside or outside the jet.

In our recent paper on C/A clustering in $e^+e^-$ [27], we discussed in full detail how the C/A clustering proceeds at three-loop order. We listed all possible scenarios and orderings that must be taken into account and combined them to arrive at the analog of Eq. (24) for the $e^+e^-$ case. The calculations for $V/H$ + jet follow exactly the same procedure for all three channels, the only difference being in the corresponding Born color factors.

Substituting the explicit formulas for the eikonal amplitudes squared from Ref. [41], one obtains, for the sum in the square bracket in Eq. (24),

$$\mathcal{W}^{\mathrm{VVR}}_{123,\delta} + \mathcal{W}^{\mathrm{RVR}}_{123,\delta} + \mathcal{W}^{\mathrm{RRR}}_{123,\delta} = \prod_{\sigma=1}^{3} \left( \sum_{(i_\sigma j_\sigma) \in \Delta_\delta} \mathcal{C}_{i_\sigma j_\sigma} w^{\sigma}_{i_\sigma j_\sigma} \right) + \sum_{\substack{k,\ell,m=1 \\ k\neq\ell\neq m}}^{3} \left( \mathrm{C_A} \sum_{(ij)\in\Delta_\delta} \mathcal{C}_{ij} \mathcal{A}^{k\ell}_{ij} \right) \left( \sum_{(i'j')\in\Delta_\delta} \mathcal{C}_{i'j'} w^{m}_{i'j'} \right) + \mathrm{C_A^2} \sum_{(ij)\in\Delta_\delta} \mathcal{C}_{ij} (\mathcal{A}^{12}_{ij}\bar{\mathcal{A}}^{13}_{ij} + \mathcal{B}^{123}_{ij}) + \mathcal{Q}_\delta \sum_{(ijk)\in\pi_\delta} (\mathcal{G}^{k1}_{ij}(2,3) + 2 \leftrightarrow 3), \tag{25}$$

where the one-loop antenna function is given in Eq. (4), the set of all dipoles in channel $\delta$ is $\Delta_\delta = \{(ab), (aj), (bj)\}$, the set of triplets in channel $\delta$ is $\pi_\delta = \{(abj), (ajb), (bja)\}$, the color Casimir scalars are defined in Eq. (5), and the quadrupole color factor, $\mathcal{Q}_\delta$, takes the channel-dependent values [41]

$$\mathcal{Q}_{\delta_1} = \mathcal{Q}_{\delta_2} = \mathrm{C_A^2}(\mathrm{C_A} - 2\mathrm{C_F}), \qquad \mathcal{Q}_{\delta_3} = 6\mathrm{C_A}. \tag{26}$$

The higher-order antenna functions are defined as follows: [41,42]

$$\mathcal{A}^{ij}_{\alpha\beta} = w^i_{\alpha\beta}(w^j_{\alpha i} + w^j_{i\beta} - w^j_{\alpha\beta}), \tag{27a}$$

$$\mathcal{B}^{ijk}_{\alpha\beta} = w^i_{\alpha\beta}(\mathcal{A}^{jk}_{\alpha i} + \mathcal{A}^{jk}_{i\beta} - \mathcal{A}^{jk}_{\alpha\beta}), \tag{27b}$$

$$\mathcal{G}^{ij}_{\alpha\beta}(m,n) = w^j_{\alpha\beta} T^{ij}_{\alpha\beta}(n) U^{ij}_{\alpha\beta}(m), \tag{27c}$$

with the cross-channel functions

$$T^{ij}_{\alpha\beta}(n) = w^n_{\alpha\beta} + w^n_{ij} - w^n_{\alpha i} - w^n_{\beta j}, \tag{27d}$$

$$U^{ij}_{\alpha\beta}(n) = w^n_{\alpha\beta} + w^n_{ij} - w^n_{\alpha j} - w^n_{\beta i}. \tag{27e}$$

The first term in Eq. (25) is the source of CLs, since it arises from independent primary emissions whose real-virtual cancellation is disturbed by the clustering, whereas the remaining three terms give rise to NGLs. Upon integration, one can show that the perturbative distribution at this loop order takes the form—analogous to the $k_t$ case [34]—

$$f^{(3),\mathrm{C/A}}_{\mathcal{B},\delta}(\rho) = \frac{1}{3!}\left[f^{(1)}_{\mathcal{B},\delta}\right]^3 + f^{(1)}_{\mathcal{B},\delta} \times [\mathcal{C}_{2,\delta} + \mathcal{S}_{2,\delta}] + \mathcal{C}^{\mathrm{C/A}}_{3,\delta} + \mathcal{S}^{\mathrm{C/A}}_{3,\delta}. \tag{28}$$

The first two terms on the rhs of Eq. (28) have already been computed and are identical to the $k_t$ case. The last two terms, $\mathcal{C}^{\mathrm{C/A}}_{3,\delta}$ and $\mathcal{S}^{\mathrm{C/A}}_{3,\delta}$, are the new contributions arising from the application of C/A clustering to the invariant jet-mass distribution at three-loop order. We discuss them in detail below.

### 1. CLs

The primary-emission piece in Eq. (25) gives rise to clustering logarithms and may be cast in the usual factorized form [5,18,28,29,44]

$$\mathcal{C}^{\mathrm{C/A}}_{3,\delta}(\rho) = -\frac{1}{3!}\bar{\alpha}_s^3 L^3 \mathcal{F}^{\mathrm{C/A}}_{3,\delta}(R), \tag{29a}$$

where the full C/A coefficient for CLs at three-loop order decomposes as

$$\mathcal{F}^{\mathrm{C/A}}_{3,\delta} = \mathcal{F}^{k_t}_{3,\delta} + \tilde{\mathcal{F}}^{\mathrm{C/A}}_{3,\delta}, \tag{29b}$$

where $\mathcal{F}^{k_t}_{3,\delta}$ is the contribution already present in $k_t$ clustering. Recall that the C/A algorithm strictly encompasses $k_t$: every configuration accounted for by $k_t$ clustering is automatically included in C/A clustering, with C/A generating additional contributions not captured by $k_t$. This has been discussed extensively in Ref. [27]. The pure C/A correction at this order, $\tilde{\mathcal{F}}^{\mathrm{C/A}}_{3,\delta}$, reads

$$\tilde{\mathcal{F}}^{\mathrm{C/A}}_{3,\delta}(R) = R^6 \int_{1_{\mathrm{out}}} \left[\int_{2_{\mathrm{in}}} + \int_{2_{\mathrm{out}}}\right] \int_{3_{\mathrm{in}}} \Omega_{12}\Omega_{23}\bar{\Omega}_{13}\Delta_{123} \times \left[\prod_{\sigma=1}^{3} \sum_{(i_\sigma j_\sigma)\in\Delta_\delta} \mathcal{C}_{i_\sigma j_\sigma} w^{\sigma}_{i_\sigma j_\sigma}\right], \tag{29c}$$

where the following integral shorthand notation is adopted [34]:

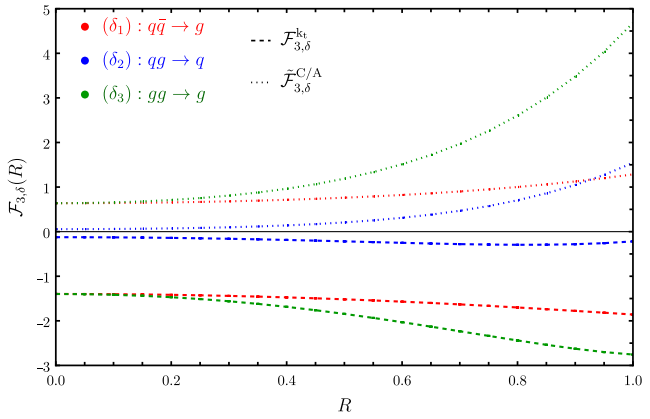



FIG. 1. The pure C/A corrections to the CL coefficients at three-loop order as a function of the jet radius $R$ (dotted curves) for all three Born channels. The dashed curves correspond to the CL coefficients in the $k_t$ clustering case.

$$\int_{i_{\mathrm{in}}} \equiv \int_0^1 r_i \mathrm{d}r_i \int_0^{2\pi} \frac{\mathrm{d}\theta_i}{2\pi}, \qquad \int_{i_{\mathrm{out}}} \equiv \int_1^{\pi/|R\sin\theta_i|} r_i \mathrm{d}r_i \int_0^{2\pi} \frac{\mathrm{d}\theta_i}{2\pi}, \tag{30}$$

where the upper limit on the second integral has been discussed in our previous works (see, for instance, the paragraph following Eq. (3.26) of Ref. [34]). To account for integration over all positions of gluon $k_2$, we exploit the completeness relation $\int_{2_{\mathrm{in}}} + \int_{2_{\mathrm{out}}} = 1$. Both integrals are evaluated numerically using the multidimensional CUBA library [45] and summed to yield the results displayed in Fig. 1. The numerical values of $\mathcal{F}^{k_t}_{3,\delta}$ and $\mathcal{F}^{\mathrm{C/A}}_{3,\delta}$ at some representative values of $R$ are shown in Table I.

TABLE I. Numerical values of the three-loop CL and NGL coefficients, $\mathcal{F}_{3,\delta}$ and $\mathcal{G}_{3,\delta}$, for the $k_t$ and C/A algorithms at representative jet radii $R$ for all three Born channels.

| $R$ | Channels | $F^{k_t}_{3,\delta}$ | $F^{C/A}_{3,\delta}$ | $G^{k_t}_{3,\delta}$ | $G^{C/A}_{3,\delta}$ |
|---|---|---|---|---|---|
| 0.4 | $(\delta_1)$ | −1.47 | −0.76 | 23.81 | 5.06 |
| | $(\delta_2)$ | −0.19 | −0.05 | 13.16 | 4.94 |
| | $(\delta_3)$ | −1.69 | −0.72 | 28.86 | 7.04 |
| 0.6 | $(\delta_1)$ | −1.57 | −0.75 | 24.66 | 5.02 |
| | $(\delta_2)$ | −0.25 | −0.06 | 8.08 | −3.05 |
| | $(\delta_3)$ | −2.03 | −0.52 | 23.63 | −2.97 |
| 0.7 | $(\delta_1)$ | −1.63 | −0.73 | 25.32 | 5.09 |
| | $(\delta_2)$ | −0.28 | −0.19 | 3.40 | −9.69 |
| | $(\delta_3)$ | −2.23 | −0.27 | 19.25 | −10.67 |
| 0.8 | $(\delta_1)$ | −1.70 | −0.70 | 26.10 | 5.20 |
| | $(\delta_2)$ | −0.28 | −0.29 | 0.41 | −17.82 |
| | $(\delta_3)$ | −2.44 | 0.16 | 14.23 | −19.61 |
| 1.0 | $(\delta_1)$ | −1.85 | −0.57 | 27.73 | 5.13 |
| | $(\delta_2)$ | −0.22 | 1.32 | −15.49 | −36.99 |
| | $(\delta_3)$ | −2.76 | 1.92 | 4.95 | −38.96 |

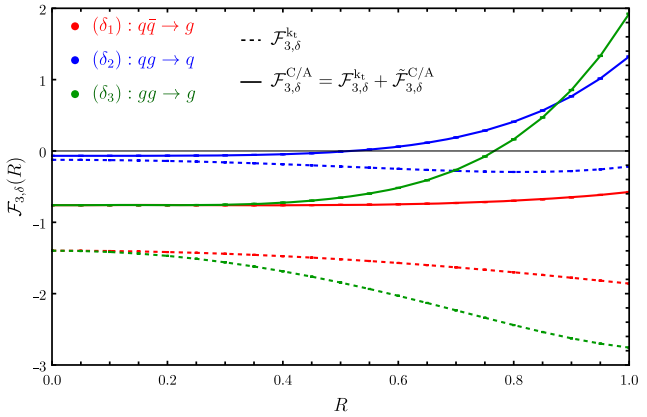



FIG. 2. Comparison of the full C/A (solid) and $k_t$ (dashed) CL coefficients at three-loop order as a function of $R$ for all three Born channels.

Figure 1 compares $\tilde{\mathcal{F}}^{C/A}_{3,\delta}$ with $\mathcal{F}^{k_t}_{3,\delta}$ for all three Born channels as a function of the jet radius $R$. The pure C/A corrections are generally of opposite sign to the $k_t$ coefficients and smaller in magnitude; they thus serve to reduce the impact of the $k_t$ contributions.

In Fig. 2 we plot the full C/A CL coefficient alongside the $k_t$ one. As expected, combining the $k_t$ contribution with the pure C/A correction $\tilde{\mathcal{F}}^{C/A}_{3,\delta}$ yields a full C/A CL coefficient that is smaller in magnitude than its $k_t$ counterpart. This is an advantageous feature of the C/A algorithm, since it is phenomenologically desirable to minimize or entirely eliminate the contribution of CLs to the observable distribution. Several important observations follow from Fig. 2:

(i) Unlike the CL coefficients for $k_t$ clustering, which are negative for all values of $R$, those of C/A change sign across the range of $R$. This is especially true for channels $(\delta_2)$ and $(\delta_3)$. It follows from Eq. (29a) that the CL contribution to the jet-mass distribution in $k_t$ clustering is positive across the whole range of $R$ and for all Born channels. A similar behavior is seen for channel $(\delta_1)$ in C/A clustering. By contrast, channels $(\delta_2)$ and $(\delta_3)$ do not maintain a constant sign over all values of $R$. Specifically, channel $(\delta_2)$ has a vanishing CL contribution at $R \sim 0.5$, while channel $(\delta_3)$ has a vanishing CL contribution at $R \sim 0.78$. These particular values of $R$ could in principle be exploited to eliminate the CL contribution entirely. However, one must also verify the magnitude of the NGLs at such values, since the ideal scenario would be for both CLs and NGLs to be simultaneously suppressed or to mutually cancel.

(ii) For small values of $R$, the CL coefficients for channels $(\delta_1)$ and $(\delta_3)$ are identical in size, owing to the equality of their color factors—both channels possess an outgoing hard gluon in the final state. They do, however, become distinct at larger values of $R$, with the divergence beginning around $R \sim 0.35$ for C/A and $R \sim 0.15$ for $k_t$.

(iii) The edge (boundary) effect is evident for both clustering algorithms. Specifically, as the jet radius shrinks to zero, the CL coefficients do not tend to zero but instead approach a finite constant. This behavior has been explained in detail in our previous papers (see, for instance, Refs. [18,33,44]).

(iv) In the limit $R \to 0$, one recovers the $e^+e^-$ findings of Ref. [27]:

$$\lim_{R\to 0} \mathcal{F}^{C/A}_{3,\delta_1}(R) = \lim_{R\to 0} \mathcal{F}^{C/A}_{3,\delta_3}(R) = -0.76 \simeq -0.028 C_A^3,$$
$$\lim_{R\to 0} \mathcal{F}^{C/A}_{3,\delta_2}(R) = -0.068 \simeq -0.028 C_F^3. \quad (31)$$

(v) The $R$-dependence of the CL coefficients in C/A clustering differs among the three channels, with channel $(\delta_3)$ exhibiting the strongest $R$-dependence and channel $(\delta_1)$ the weakest. Moreover, the $R$-dependence of the CL coefficients is more pronounced in C/A clustering than in $k_t$ clustering. This may be attributed to the fact that the C/A algorithm admits more gluon configurations that contribute to the jet-mass observable than $k_t$; furthermore, such configurations depend on $R$ through the angular distance measure.

(vi) Quantitatively, the CL coefficients in C/A clustering are reduced by approximately 45% relative to the $k_t$ case for all three channels at $R = 0$. For channel $(\delta_1)$, this reduction grows steadily, reaching approximately 70% at $R = 1$. For channels $(\delta_2)$ and $(\delta_3)$, the reduction reaches 100% at the values of $R$ where the C/A CL coefficients vanish. For channel $(\delta_2)$, the C/A CL coefficient subsequently changes sign and grows to exceed the magnitude of its $k_t$ counterpart for $R \gtrsim 0.8$. For channel $(\delta_3)$, the reduction diminishes to approximately 30% at $R \sim 1$, with the C/A coefficient carrying the opposite sign to that of $k_t$.

In the following section, we discuss the NGL contribution to the jet-mass distribution at three-loop order.

### 2. NGLs

The computation of NGL coefficients follows an analogous procedure to that of CLs. In Eq. (25) there are three terms contributing to NGLs at this loop order. The first term, proportional to $\mathcal{A}^{k\ell}_{ij} \times w^m_{i'j'}$, arises from interference between the one- and two-loop perturbative distributions. The second, proportional to $\mathcal{A}^{12}_{ij}\bar{\mathcal{A}}^{13}_{ij}$ and $\mathcal{B}^{123}_{ij}$, corresponds to two-loop dipole-dipole interference and a genuine three-loop dipole contribution. The last term, proportional to $\mathcal{G}^{k\ell}_{ij}$, corresponds to the quadrupole (ghost) contribution (see Ref. [41] for full details). The total NGL contribution from all three contributions can be shown to take the usual form

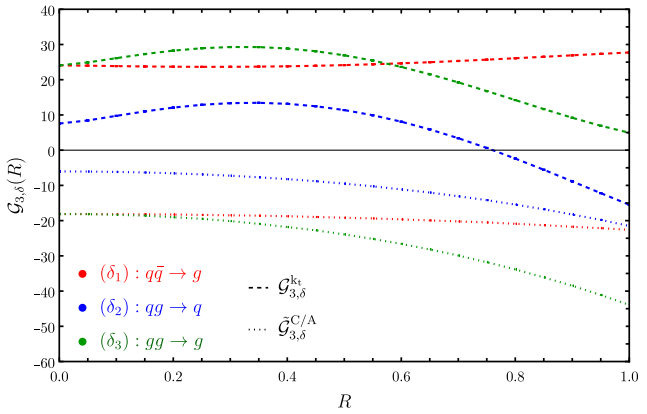



FIG. 3. The pure C/A corrections to the NGL coefficients at three-loop order as a function of the jet radius $R$ (dotted curves) for all three Born channels. The dashed curves correspond to the NGL coefficients in the $k_t$ clustering case.

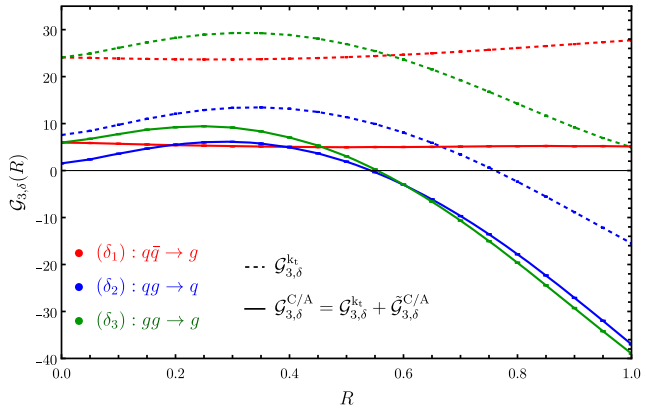



FIG. 4. Comparison of the full C/A (solid) and $k_t$ (dashed) NGL coefficients at three-loop order as a function of $R$ for all three Born channels.

$$\mathcal{S}^{C/A}_{3,\delta}(\rho) = +\frac{1}{3!}\bar{\alpha}_s^3 L^3 \mathcal{G}^{C/A}_{3,\delta}(R), \tag{32}$$

where the NGL coefficient is given by

$$\mathcal{G}^{C/A}_{3,\delta}(R) = \mathcal{G}^{k_t}_{3,\delta}(R) + \tilde{\mathcal{G}}^{C/A}_{3,\delta}(R), \tag{33}$$

with the first term corresponding to the $k_t$ contribution, already computed in Ref. [34], and the second term being the pure C/A correction at this loop order. Recall that each of these two terms contains all three aforementioned contributions, namely interference, dipole, and quadrupole. These integrals may be evaluated numerically using, for instance, the CUBA library. The results are plotted in Fig. 3. Analogously to the CL case, the pure C/A corrections to the NGL distribution are of opposite sign to the $k_t$ ones and of comparably smaller magnitude; they thus serve to reduce the NGL contribution relative to both $k_t$ and anti-$k_t$. The latter usually produces the largest NGL contribution (see, for instance, Refs. [33,34]). Note that, unlike the $k_t$ case, the C/A corrections maintain a consistent sign across the whole range of $R$ for all three channels.

In Fig. 4 we plot the full C/A NGL coefficients for all three Born channels as a function of the jet radius alongside their $k_t$ counterparts. Several observations follow from Fig. 4:

(i) As anticipated, the NGL coefficients for C/A clustering are generally smaller in magnitude than those for $k_t$. However, this pattern is not consistent over the whole range of $R$. There are intervals of $R$ in which the C/A coefficients exceed those of $k_t$; this occurs after the coefficients change sign.

(ii) The boundary effect persists across all channels and all jet algorithms.

(iii) The NGL coefficients in C/A clustering for channels $(\delta_2)$ and $(\delta_3)$ vanish at $R \sim 0.55$. This is to be compared with only one channel for $k_t$ clustering, namely channel $(\delta_2)$, which vanishes at $R \sim 0.78$.

(iv) The reduction in the magnitude of the NGL coefficients in C/A clustering relative to $k_t$ is approximately 75%–79% for all three Born channels at $R = 0$. For channel $(\delta_1)$, this reduction continues steadily, reaching approximately 81% at $R = 1$. Note that for this channel there is no change of sign. For channels $(\delta_2)$ and $(\delta_3)$, the C/A NGL coefficients grow to exceed those of $k_t$ beyond the value of $R$ at which they vanish (i.e., beyond $R \sim 0.55$).

(v) In the limit of small $R$, one obtains from Fig. 4:

$$\lim_{R\to 0} \mathcal{G}^{C/A}_{3,\delta_1} = \lim_{R\to 0} \mathcal{G}^{C/A}_{3,\delta_3} = 5.95,$$
$$\lim_{R\to 0} \mathcal{G}^{C/A}_{3,\delta_2} = 1.54. \tag{34}$$

These values differ considerably from those reported for the $e^+e^-$ dijet-mass observable [27], where $C_F^2 C_A(0.43) + C_F C_A^2(0.37) \simeq 6.73$ and $C_A^3(0.43 + 0.37) \simeq 21.6$. The latter were also found not to match those calculated for the hemisphere-mass observable [46]. The small-$R$ limit of the $k_t$ NGL coefficients does coincide with those computed for the $e^+e^-$ dijet-mass observable, as illustrated in Ref. [34]. They read (see also Fig. 4)

$$\lim_{R\to 0} \mathcal{G}^{k_t}_{3,\delta_1} = \lim_{R\to 0} \mathcal{G}^{k_t}_{3,\delta_3} = 7.5,$$
$$\lim_{R\to 0} \mathcal{G}^{k_t}_{3,\delta_2} = 24.07. \tag{35}$$

Note that the $e^+e^-$ results in C/A clustering are close to the $k_t$ ones at small $R$ because the pure C/A corrections are quite small in that case. By contrast, the pure C/A corrections in the $V/H$ + jet processes studied in this paper are significant, reaching well

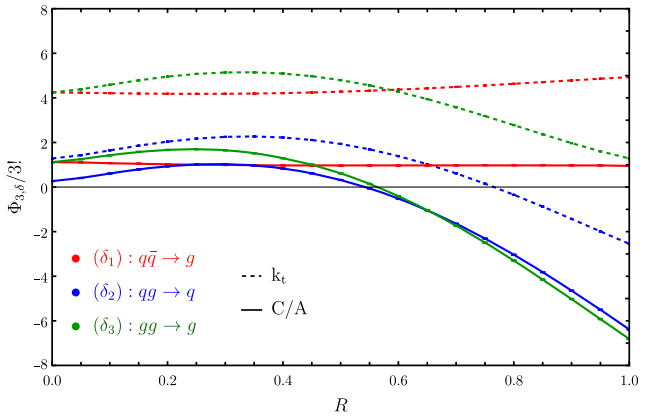



FIG. 5. Comparison of the sum of CL and NGL coefficients for C/A (solid) and $k_t$ (dashed) clustering at three-loop order as a function of $R$ for all three Born channels.

beyond 70% for small $R$, as discussed above (see Fig. 3). This may be attributed to the fact that, according to Eq. (25), all eikonal amplitudes squared are summed together as they all share the same phase-space constraint—unlike the $k_t$ case, in which the eikonal amplitudes squared are multiplied by different phase-space constraints, causing a partial cancellation among them (see Eqs. (3.48) of Ref. [34]).

To assess the combined impact of both types of nonglobal logarithms—CLs and NGLs—on the distribution of the invariant-mass observable at this loop order, we plot, in Fig. 5, their sum

$$\mathcal{C}_{3,\delta}(\rho) + \mathcal{S}_{3,\delta}(\rho) = \frac{1}{3!}\bar{\alpha}_s^3 L^3 \Phi_{3,\delta}(R), \tag{36}$$

for both $k_t$ and C/A clusterings, where $\Phi_{3,\delta} \equiv -\mathcal{F}_{3,\delta} + \mathcal{G}_{3,\delta}$. We observe that for $R \lesssim 0.7$ the sum of both coefficients is of order unity or smaller for all three channels in the C/A clustering case. In fact, for channels $(\delta_2)$ and $(\delta_3)$ the combined nonglobal effect vanishes for $R \sim 0.55$–$0.58$. A similar behavior is seen for $k_t$, but only for channel $(\delta_2)$ in the range $R \in [0, 1]$. This confirms the conclusion of Ref. [27] for $e^+e^-$ processes that the C/A jet algorithm is preferred to $k_t$—and over anti-$k_t$—in mitigating the impact of large nonglobal logarithms over a wide range of $R$, and in eliminating them entirely for specific values of $R$. This conclusion holds, however, only at three-loop order. One must verify it at higher loop orders as well as at all orders. The former task is carried out in the next section, in which we address the impact of the said logarithms at four loops.

### B. Four-loop order

The four-loop calculation of the invariant jet-mass distribution for the leading high-$p_t$ jet follows an identical strategy to the three-loop case presented in the previous section. The details are based on the results of Refs. [27,33,34,41]. The full structure of the C/A integrand is identical to that computed for the $e^+e^-$ case in [27], the only difference being in the eikonal amplitudes squared, which are instead taken from [41]. The main steps leading to the final results are as follows:

(i) One enumerates all initial in-jet and out-of-jet gluon configurations at four loops. These include configurations—prior to applying the C/A algorithm—in which only gluon $k_4$ is inside the jet and all others are outside; gluons $k_3$ and $k_4$ are inside and the remaining two are outside; and so on. For each such configuration one considers all possible pairwise clusterings among the four gluons, characterized by the constraints $\Delta_{12j}$, $\Delta_{13j}$, $\Delta_{14j}$, $\Delta_{23j}$, $\Delta_{24j}$, and $\Delta_{34j}$. Each clustering constraint of the form $\Delta_{\alpha\beta j}$, where $j$ denotes the leading jet, takes the value 1 or 0. If it equals 1, gluons $k_\alpha$ and $k_\beta$ are clustered first, i.e., they have the smallest pairwise distance measure. If it equals 0, the smallest distance is that between the softer gluon $k_\beta$ and the jet, and the two gluons are not clustered together. There is a total of $2^6 = 64$ such configurations to be considered. Full details are given in Ref. [27].

(ii) For configurations in which more than one $\Delta_{\alpha\beta j}$ equals 1, all possible orderings of the smallest distances must be considered. For example, if $\Delta_{12j} = \Delta_{13j} = 1$, one must consider both $d_{12} > d_{13}$ and $d_{12} < d_{13}$. There is a total of 1957 such orderings across the 64 configurations.

(iii) For each of the above gluon configurations there are 16 possible real/virtual assignments, corresponding to each of the four gluons being either real (R) or virtual (V). Of these, only 8 need to be computed explicitly; the remaining 8 follow by symmetry [5,33]. The 8 independent configurations are: RRRR, RRVR, RVRR, VRRR, RVVR, VRVR, VVRV, and VVVR. Here, for instance, RRRR denotes all four gluons real, and RVVR denotes gluons $k_1$ and $k_4$ real with $k_2$ and $k_3$ virtual. The explicit formulas for the corresponding eikonal amplitudes squared are given in [41].

(iv) The C/A algorithm is applied to each configuration to determine the final in-jet or out-of-jet status of each of the four gluons. Only gluons that end up inside the measured jet contribute to its mass. NGLs and CLs arise from configurations in which the real-emission contribution to the jet mass is not fully canceled by the corresponding virtual corrections.

(v) Accounting for all of the above configurations is a tedious and complex task that can, nonetheless, be fully automated. To this end, we have developed a PYTHON code to perform the bookkeeping, together with a *Mathematica* script to organize the output expressions. Since the C/A clustering contains

the $k_t$ clustering as a subset, the PYTHON code is written to omit contributions already accounted for in the $k_t$ case, so that the final output consists of the pure new C/A corrections. The $k_t$ contributions have already been computed in our previous work [34].

(vi) The resulting pure C/A expressions constitute the integrand to be evaluated numerically. Of all the phase-space regions enumerated above, only four produce nonvanishing real-virtual miscancellations and hence contribute to the jet-mass distribution, namely: $1_{\rm out}2_{\rm out}3_{\rm out}4_{\rm in}$, $1_{\rm out}2_{\rm out}3_{\rm in}4_{\rm in}$, $1_{\rm out}2_{\rm in}3_{\rm out}4_{\rm in}$, and $1_{\rm out}2_{\rm in}3_{\rm in}4_{\rm in}$.

Following the above analysis, one can show that the invariant jet-mass distribution at four-loop order takes an analogous form to those at two- and three-loop order, Eqs. (22) and (28) (see also Refs. [27,34]):

$$f_{\mathcal{B},\delta}^{(4),\mathrm{C/A}}(\rho) = \frac{1}{4!}\left[f_{\mathcal{B},\delta}^{(1)}\right]^4 + f_{\mathcal{B},\delta}^{(1)} \times \left[\mathcal{C}_{3,\delta}^{\mathrm{C/A}} + \mathcal{S}_{3,\delta}^{\mathrm{C/A}}\right] + \frac{1}{2!}\left[f_{\mathcal{B},\delta}^{(1)}\right]^2 \times [\mathcal{C}_{2,\delta} + \mathcal{S}_{2,\delta}] + \frac{1}{2!}[\mathcal{C}_{2,\delta}]^2 + \frac{1}{2!}[\mathcal{S}_{2,\delta}]^2 + \mathcal{C}_{2,\delta} \times \mathcal{S}_{2,\delta} + \mathcal{C}_{4,\delta}^{\mathrm{C/A}} + \mathcal{S}_{4,\delta}^{\mathrm{C/A}}. \tag{37}$$

The last two terms are the new four-loop CL and NGL contributions. They admit the usual factorized form

$$\mathcal{C}_{4,\delta}^{\mathrm{C/A}}(\rho) = +\frac{1}{4!}\bar{\alpha}_s^4 L^4 \mathcal{F}_{4,\delta}^{\mathrm{C/A}}(R), \tag{38a}$$

$$\mathcal{S}_{4,\delta}^{\mathrm{C/A}}(\rho) = -\frac{1}{4!}\bar{\alpha}_s^4 L^4 \mathcal{G}_{4,\delta}^{\mathrm{C/A}}(R). \tag{38b}$$

Although Eq. (38) define $\mathcal{F}_{4,\delta}^{\mathrm{C/A}}$ and $\mathcal{G}_{4,\delta}^{\mathrm{C/A}}$ as the formal coefficients of the four-loop CL and NGL contributions separately, the complex interplay among color factors, kinematics, and jet-algorithm constraints at this order prevents their independent numerical determination for C/A clustering. Unlike the situation at two- and three-loop order—where the primary-emission piece of the integrand (the source of CLs) and the correlated-emission piece (the source of NGLs) can be isolated term by term—the C/A clustering constraints at four loops generate phase-space constraints that couple both sources inseparably. Only their combination

$$\mathcal{C}_{4,\delta}^{\mathrm{C/A}}(\rho) + \mathcal{S}_{4,\delta}^{\mathrm{C/A}}(\rho) = \frac{1}{4!}\bar{\alpha}_s^4 L^4 \underbrace{(\mathcal{F}_{4,\delta}^{\mathrm{C/A}}(R) - \mathcal{G}_{4,\delta}^{\mathrm{C/A}}(R))}_{\equiv \Phi_{4,\delta}^{\mathrm{C/A}}(R)} \tag{39}$$

is numerically accessible, where we have introduced the shorthand $\Phi_{4,\delta}^{A}(R) \equiv \mathcal{F}_{4,\delta}^{A}(R) - \mathcal{G}_{4,\delta}^{A}(R)$ for $A \in \{\mathrm{anti}-k_t, k_t, \mathrm{C/A}\}$. Numerical values of $\Phi_{4,\delta}^{A}/4!$ for both the $k_t$ and C/A algorithms, together with the pure C/A correction $\Delta\Phi_{4,\delta}/4! \equiv (\Phi_{4,\delta}^{\mathrm{C/A}} - \Phi_{4,\delta}^{\mathrm{k_t}})/4!$, are given in Table II at representative values of $R$. All values are obtained by numerical integration using the CUBA library. A negative value of $\Phi_{4,\delta}^{A}/4!$ indicates that NGLs dominate CLs ($\mathcal{G}_{4,\delta}^{A} > \mathcal{F}_{4,\delta}^{A}$); the single positive C/A entry for channel $(\delta_3)$ at $R = 1.0$ signals a local reversal of this hierarchy. A positive $\Delta\Phi_{4,\delta}/4!$ means C/A reduces the nonglobal coefficient relative to $k_t$ and a negative value means the opposite.

TABLE II. Numerical values of the combined four-loop CL and NGL coefficient $\Phi_{4,\delta}^{A}/4!$, Eqs. (38) and (39), for the $k_t$ and C/A algorithms, and the pure C/A correction $\Delta\Phi_{4,\delta}/4!$, at representative jet radii $R$ for all three Born channels.

| **R** | Channel | $\boldsymbol{\Phi_{4,\delta}^{k_t}/4!}$ | $\boldsymbol{\Phi_{4,\delta}^{C/A}/4!}$ | $\boldsymbol{\Delta\Phi_{4,\delta}/4!}$ |
|---|---|---|---|---|
| 0.4 | $(\delta_1)$ | −2.37 | −1.82 | 0.55 |
| | $(\delta_2)$ | −1.85 | −1.78 | 0.07 |
| | $(\delta_3)$ | −4.78 | −6.50 | −1.72 |
| 0.6 | $(\delta_1)$ | −2.45 | −1.93 | 0.52 |
| | $(\delta_2)$ | −2.49 | −2.42 | 0.07 |
| | $(\delta_3)$ | −3.96 | −5.70 | −1.74 |
| 0.7 | $(\delta_1)$ | −2.09 | −1.46 | 0.63 |
| | $(\delta_2)$ | −3.46 | −3.14 | 0.32 |
| | $(\delta_3)$ | −4.32 | −5.75 | −1.43 |
| 0.8 | $(\delta_1)$ | −2.22 | −1.63 | 0.59 |
| | $(\delta_2)$ | −4.26 | −3.46 | 0.80 |
| | $(\delta_3)$ | −4.07 | −4.62 | −0.55 |
| 1.0 | $(\delta_1)$ | −1.84 | −1.07 | 0.77 |
| | $(\delta_2)$ | −4.74 | −2.30 | 2.44 |
| | $(\delta_3)$ | −0.98 | 1.22 | 2.20 |

Figure 6 shows the combined quantity $\Phi_{4,\delta}/4!$ for the pure C/A corrections alongside the $k_t$ results at four-loop order, for all three channels as a function of the jet radius $R$. The combined effect of CLs and NGLs for the pure C/A correction is relatively small and of opposite sign to the $k_t$ result for channels $(\delta_1)$ and $(\delta_2)$, confirming that C/A consistently reduces the nonglobal logarithmic contribution relative to $k_t$ for these two channels. The corresponding result for the pure gluon-fusion channel $(\delta_3)$ is both relatively large (reaching $\Delta\Phi_{4,\delta_3}/4! \approx -1.74$ at $R = 0.6$; see Table II) and of the same sign as the $k_t$ result up to $R \sim 0.85$ where it vanishes and flips sign.

Figure 7 depicts the combined CL and NGL effect at four-loop order, $\Phi_{4,\delta}$, for all three channels as a function of $R$, comparing all three sequential recombination algorithms: anti-$k_t$, $k_t$, and C/A. The following observations are worth noting:

(i) Channels $(\delta_2)$ and $(\delta_3)$ follow the pattern seen at three-loop order: both exhibit a reduced nonglobal logarithmic impact compared to both anti-$k_t$ and $k_t$, though no optimal jet radius exists at which this

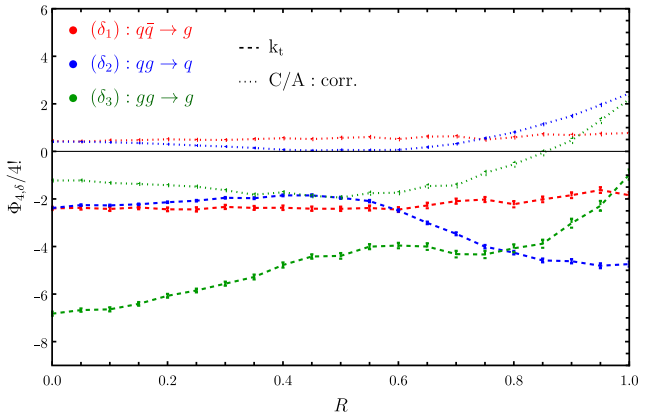



FIG. 6. The combined CL and NGL coefficient for the pure C/A corrections (dotted) and the full $k_t$ clustering (dashed) at four-loop order, shown as a function of $R$ for all three Born channels.

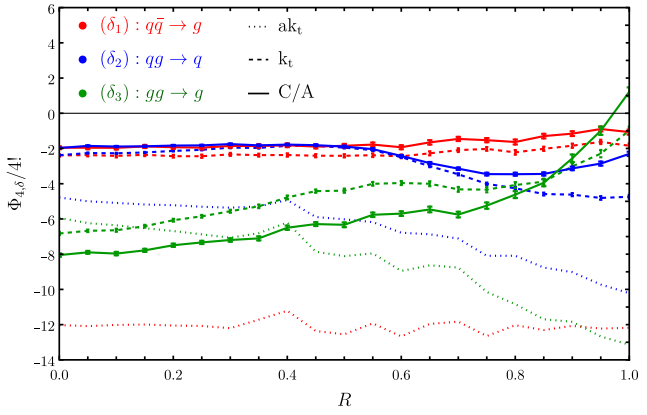



FIG. 7. The combined CL and NGL coefficient for the full C/A (solid), full $k_t$ (dashed), and anti-$k_t$ (dotted) clusterings at four-loop order, shown as a function of $R$ for all three Born channels.

impact vanishes entirely. Channel $(\delta_2)$ for C/A is nonetheless comparable to $k_t$ over a wide range of $R$. The reduction in the combined CL and NGL coefficient reaches 84% and 59% for channels $(\delta_1)$ and $(\delta_2)$, respectively, compared to the anti-$k_t$ case, while the corresponding reductions for $k_t$ are approximately 80% and 50%. This confirms the relatively small corrections induced by C/A clustering compared to $k_t$. At the phenomenologically relevant jet radius $R = 0.7$, the reductions are 88% and 56% for C/A and 82% and 51% for $k_t$ for these two channels, further confirming the superiority of the C/A algorithm in suppressing nonglobal effects over a wide range of jet radii.

(ii) Channel $(\delta_3)$, on the other hand, deviates from the three-loop pattern. For $R \lesssim 0.4$, the combined nonglobal effect for C/A exceeds that for both anti-$k_t$ and $k_t$ by at most 35% (at $R = 0$). For $R \gtrsim 0.4$, the nonglobal effect for C/A is continuously reduced relative to anti-$k_t$, reaching a reduction of approximately 100% at $R \sim 0.97$, corresponding to the complete elimination of this effect; it then changes sign beyond this jet radius. Compared to $k_t$ clustering, the nonglobal effect for C/A is larger for $R \lesssim 0.9$. Hence, for channel $(\delta_3)$, the $k_t$ algorithm is preferred to C/A in suppressing the impact of nonglobal logarithms.

(iii) In the small-$R$ limit, channels $(\delta_1)$ and $(\delta_2)$ overlap at values of approximately $-1.96$ and $-2.38$ for C/A and $k_t$ clusterings, respectively. This differs from the two- and three-loop cases, in which channels $(\delta_1)$ and $(\delta_3)$ overlap in the small-$R$ limit. This indicates that the gluon-fusion channel behaves in an unexpected manner starting at four-loop order. Indeed, it has been observed in previous works (see, e.g., [5,42]) that several patterns established at lower loop orders break down at four loops, both at the level of eikonal amplitudes and, consequently, at the cross-section level. It is also worth noting that the values obtained in the limit $R \to 0$ differ from those found in the $e^+e^-$ dijet case [27].

An appealing feature of the three-loop results is that the combined CL and NGL coefficient vanishes at specific values of the jet radius for channels $(\delta_2)$ and $(\delta_3)$ in C/A clustering ($R \approx 0.55$–$0.58$, Fig. 5), raising the prospect of an "optimal" radius at which nonglobal logarithms are eliminated. However, several issues temper this observation. First, even at three-loop order, channel $(\delta_1)$ admits no such zero; thus, no single value of $R$ simultaneously suppresses nonglobal effects across all three Born channels. The phenomenologically relevant quantity is the Born-cross-section-weighted combination

$$\propto \sum_{\delta} \int \mathrm{d}\mathcal{B}_{\delta} (\mathcal{F}^{\mathrm{C/A}}_{3,\delta} - \mathcal{G}^{\mathrm{C/A}}_{3,\delta}), \tag{40}$$

which is not computed here and is unlikely to vanish at the same $R$ as the individual channels. Second, and more fundamentally, the vanishing of the three-loop coefficient at $R \approx 0.55$ is an order-by-order cancellation between CLs and NGLs, not a signal of a physical suppression mechanism: the four-loop contributions to the same channels are nonzero at that radius (Fig. 7 and Table II), confirming that the cancellation does not persist beyond three-loop order. Establishing whether a physically meaningful optimal radius exists at all orders would require either fixed-order calculations beyond four loops—a task that is currently infeasible given the rapidly growing complexity of the eikonal amplitudes, the intricate geometry and combinatorics of multigluon configurations, and the sheer number of phase-space orderings that must be enumerated—or a dedicated all-orders numerical resummation for C/A clustering. The latter path is not free of obstacles: the only available all-orders C/A code [36] is restricted to purely

soft observables (such as the energy-flow observable), since the BMS equation [6] on which it rests is formulated under the assumption of purely soft singularities. The invariant jet-mass observable exhibits both soft and collinear singularities, which places it outside the scope of both the BMS framework and the aforementioned numerical code. The Dasgupta-Salam code seems, currently, the only promising short-term avenue and might be modified to cover the C/A clustering. Such a task will be pursued in future publications. Another, long-term option is to extend the BMS equation to cases where the final-state soft partons can be clustered with one of the sequential recombination jet algorithms.

To quantify the phenomenological impact of C/A clustering, it is instructive to perform an all-orders estimate of the effect of the computed nonglobal logarithms on the resummed invariant jet-mass distribution. This is the subject of the next section.

## IV. ESTIMATION OF ALL-ORDERS EFFECTS

In order to estimate the contribution of the missing higher-order terms in the CL and NGL series to the invariant jet-mass distribution, and to investigate the impact of the nonglobal part on the overall distribution, we follow the standard procedure (see, for instance, [17,18,27,33,34,44,47]). Specifically, based on the fixed-order perturbative calculations up to four-loop order presented above, we assume that both CLs and NGLs exponentiate, and that their resummed form factors appearing in Eq. (15) take, at NLL accuracy, the form

$$\mathcal{C}_\delta^{\mathrm{C/A}}(t) = \exp\left[\sum_{n=2}^{4} \frac{(-1)^n}{n!} \mathcal{F}_{n,\delta}^{\mathrm{C/A}}(R)(2t)^n\right], \tag{41a}$$

$$\mathcal{S}_\delta^{\mathrm{C/A}}(t) = \exp\left[-\sum_{n=2}^{4} \frac{(-1)^n}{n!} \mathcal{G}_{n,\delta}^{\mathrm{C/A}}(R)(2t)^n\right], \tag{41b}$$

where $\mathcal{F}_{n,\delta}^{\mathrm{C/A}}$ and $\mathcal{G}_{n,\delta}^{\mathrm{C/A}}$ are the CL and NGL coefficients, respectively, computed in the previous section. The evolution variable $t$ is related to the invariant jet-mass observable $\rho$ at one-loop order—sufficient for NLL accuracy—by

$$t = -\frac{1}{4\pi\beta_0} \ln\left[1 - 2\alpha_s(p_t)\beta_0 L\right], \tag{42}$$

with $L = \ln(R^2/\rho)$ and the leading-order QCD beta-function coefficient $\beta_0 = (11\mathrm{C_A} - 2n_f)/(12\pi)$, where $n_f$ is the number of active quark flavors. These fixed-order-based resummed form factors are then compared to the output of specialized numerical Monte Carlo codes, such as the one developed by Dasgupta and Salam [3]. The latter are parametrized as follows:

$$\mathcal{C}_\delta^{\mathrm{MC}}(t) = \exp\left[\sum_{(ij)\in\Delta_\delta} \mathcal{C}_{ij}^2 \mathcal{F}_{2,\delta}^{(ij)}(R) f_{ij}(t)\right], \tag{43a}$$

$$\mathcal{S}_\delta^{\mathrm{MC}}(t) = \exp\left[-\mathrm{C_A}\sum_{(ij)\in\Delta_\delta} \mathcal{C}_{ij} \mathcal{G}_{2,\delta}^{(ij)}(R) \tilde{f}_{ij}(t)\right], \tag{43b}$$

where $\mathcal{F}_{2,\delta}^{(ij)}$ and $\mathcal{G}_{2,\delta}^{(ij)}$ are the two-loop CL and NGL coefficients, respectively, computed in our previous paper [34] (recall that the C/A coefficients are identical to the $k_t$ ones at this loop order). The Born color factors $\mathcal{C}_{ij}$ are defined in Eq. (5). The fitting functionals $f_{ij}$ and $\tilde{f}_{ij}$ are given by

$$f_{ij}(t) = \frac{1 + (\lambda_{ij}t)^2}{1 + (\sigma_{ij}t)^{\gamma_{ij}}} t^2. \tag{44}$$

Both $f_{ij}$ and $\tilde{f}_{ij}$ share the same functional form but differ in the values of the fitting parameters $\lambda_{ij}$, $\sigma_{ij}$, and $\gamma_{ij}$. The values of these parameters for the $k_t$ case at $R = 0.7$ are given in Eqs. (4.3) and (4.6) of Ref. [34] and are also summarized in Table III.

As noted previously, the MC code of [3] does not implement the C/A jet algorithm. Moreover, the numerical code of [36], which does implement all three sequential recombination algorithms including C/A, is restricted to observables with purely soft singularities and does not cover observables with simultaneous soft and collinear singularities, such as the invariant jet-mass studied here. Both codes furthermore operate in the large-$N_c$ limit, whereas our analytical calculations retain the full finite-$N_c$ color structure; no all-orders numerical tool covering C/A clustering at finite $N_c$ for this class of observables is currently available. The only existing all-orders numerical result at finite $N_c$ concerns the energy-flow observable in Higgs decays for the anti-$k_t$ algorithm [35]. We therefore compare our analytical calculations to the output of the MC code of [3] for the anti-$k_t$ and $k_t$ algorithms, which serves to estimate the impact of C/A clustering at all orders relative to those two algorithms. A similar analysis was performed in the $e^+e^-$ dijet process in our previous work [27].

TABLE III. Fitting parameters $\lambda_{ij}$, $\sigma_{ij}$, and $\gamma_{ij}$ appearing in Eq. (44) for $R = 0.7$, corresponding to $\mathcal{C}_\delta^{\mathrm{MC}}(t)$ and $\mathcal{S}_\delta^{\mathrm{MC}}(t)$ (or, equivalently, to $f_{ij}$ and $\tilde{f}_{ij}$).

| Form factor | Channel | $\lambda_{ij}/\mathrm{C_A}$ | $\sigma_{ij}/\mathrm{C_A}$ | $\gamma_{ij}$ |
|---|---|---|---|---|
| $\mathcal{C}_\delta^{\mathrm{MC}}(t)$ | $aj, bj$ | 19.71 | 8.43 | 2.18 |
| | $ab$ | 3.34 | 2.14 | 2.18 |
| $\mathcal{S}_\delta^{\mathrm{MC}}(t)$ | $aj, bj$ | 2.29 | 10.00 | 0.88 |
| | $ab$ | 1.12 | 4.00 | 0.68 |

Because at four-loop order the CL and NGL coefficients could not be individually disentangled for C/A clustering, only their difference, $\mathcal{F}^{\mathrm{C/A}}_{4,\delta}(R) - \mathcal{G}^{\mathrm{C/A}}_{4,\delta}(R)$, has been determined numerically. This combination alone suffices to construct the physically relevant product $\mathcal{C}_\delta(t) \times \mathcal{S}_\delta(t)$, since multiplying the two exponential form factors of Eq. (41) yields a single exponential whose argument is a sum, over loop order, of terms proportional to $\mathcal{F}_{n,\delta}(R) - \mathcal{G}_{n,\delta}(R)$:

$$f^{\mathrm{ng}}_\delta(t) \equiv \mathcal{C}_\delta(t) \times \mathcal{S}_\delta(t) = \exp\left[\sum_{n=2}^{3} \frac{(-1)^n}{n!}\Big(\mathcal{F}_{n,\delta}(R) - \mathcal{G}_{n,\delta}(R)\Big)(2t)^n + \frac{1}{4!}\Big(\mathcal{F}_{4,\delta}(R) - \mathcal{G}_{4,\delta}(R)\Big)(2t)^4\right]. \tag{45}$$

At $n = 2, 3$, the individual coefficients $\mathcal{F}_{n,\delta}(R)$ and $\mathcal{G}_{n,\delta}(R)$ are known separately for all three jet algorithms, as detailed in the discussion of the two- and three-loop CL and NGL contributions above. At $n = 4$, this remains the case for the anti-$k_t$ and $k_t$ algorithms; for C/A clustering, however, we substitute directly the numerically determined combination $\Phi^{\mathrm{C/A}}_{4,\delta}(R) = \mathcal{F}^{\mathrm{C/A}}_{4,\delta}(R) - \mathcal{G}^{\mathrm{C/A}}_{4,\delta}(R)$ into Eq. (45), without evaluating $\mathcal{C}^{\mathrm{C/A}}_\delta(t)$ and $\mathcal{S}^{\mathrm{C/A}}_\delta(t)$ individually at this order.

We plot, in Fig. 8, the product $f^{\mathrm{ng}}_\delta(t)$ of Eq. (45), together with $\mathcal{C}^{\mathrm{MC}}_\delta(t) \times \mathcal{S}^{\mathrm{MC}}_\delta(t)$, for all three jet algorithms and all Born channels. Recall that for the anti-$k_t$ algorithm there is no CL contribution, so $\mathcal{C}^{\mathrm{ak_t}}_\delta(t) = 1$. The labels "2-loop," "3-loop," and "4-loop" correspond to truncating the sum in Eq. (45) at $n = 2$, 3, and 4, respectively.

Recalling that $f^{\mathrm{ng}}_\delta$ multiplies the Sudakov form factor in Eq. (15), Fig. 8 clearly shows, as anticipated, that the anti-$k_t$ algorithm produces the most severe suppression of the total distribution—that is, the largest reduction in the overall size of the invariant jet-mass distribution across all three channels. The $k_t$ algorithm produces a considerably smaller reduction over a wide range of the variable $t$, making it a more favorable choice as it more effectively mitigates the net impact of nonglobal logarithms and brings the total distribution closer to the pure Sudakov result (see Ref. [34] for further details).

The exponentiation of the fixed-order results is reliable for small values of $t$, specifically $t \lesssim 0.15$, corresponding to $\rho \gtrsim 0.005$. This is confirmed by comparing the $k_t$ fixed-order exponentiation against the $k_t$ MC curve, for which both are available. In this range, the C/A effect is comparable to the $k_t$ one. For larger values of $t$, where the fixed-order exponentiation is no longer reliable, the C/A effect appears more aggressive than $k_t$, shifting the curves toward the anti-$k_t$ result. This apparent contradiction with the three- and four-loop fixed-order findings—which show C/A outperforming $k_t$ in suppressing nonglobal logarithms—further confirms that Fig. 8 should only be trusted in the range $t \lesssim 0.15$–$0.2$, which is the phenomenologically relevant region. As noted in previous works (see, for instance, [5,17,28]), the two-loop results provide the best approximation to the all-orders MC curves over a wide range of $t$. Given the absence of an all-orders numerical estimate for C/A clustering, our analysis shows that the C/A and $k_t$ algorithms perform almost equally in reducing the impact of nonglobal logarithms at all orders. Nonetheless, the fixed-order calculations clearly indicate a preference for C/A over $k_t$ in this regard.

Figure 9 quantifies the effect of finite-$N_c$ corrections on the exponential NGL form factor constructed from the fixed-order coefficients, for channels $(\delta_1)$ and $(\delta_2)$ at $R = 0.7$. Recall that the third all-gluon Born channel $(\delta_3)$ receives no finite-$N_c$ corrections. For the phenomenologically relevant range $t \leq 0.15$, the finite-$N_c$ corrections

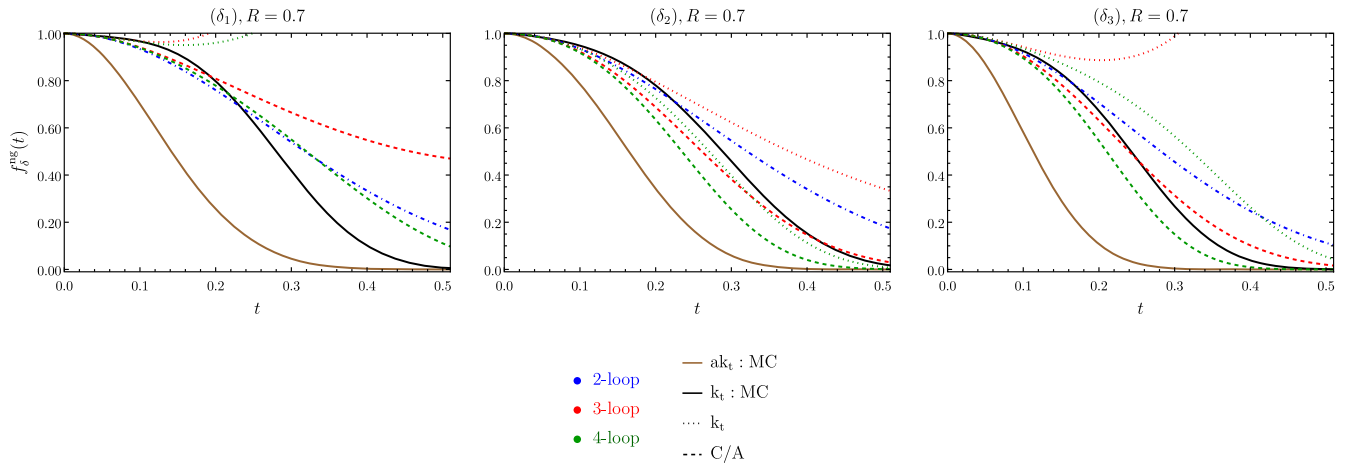



FIG. 8. Comparisons of the fixed-order exponentiated results at two-, three-, and four-loop order for the $k_t$ (dotted) and C/A (dashed) algorithms against the output of the MC code of [3] for the anti-$k_t$ (brown solid) and $k_t$ (black solid) algorithms. Results are shown for jet radius $R = 0.7$ and for all three Born channels.

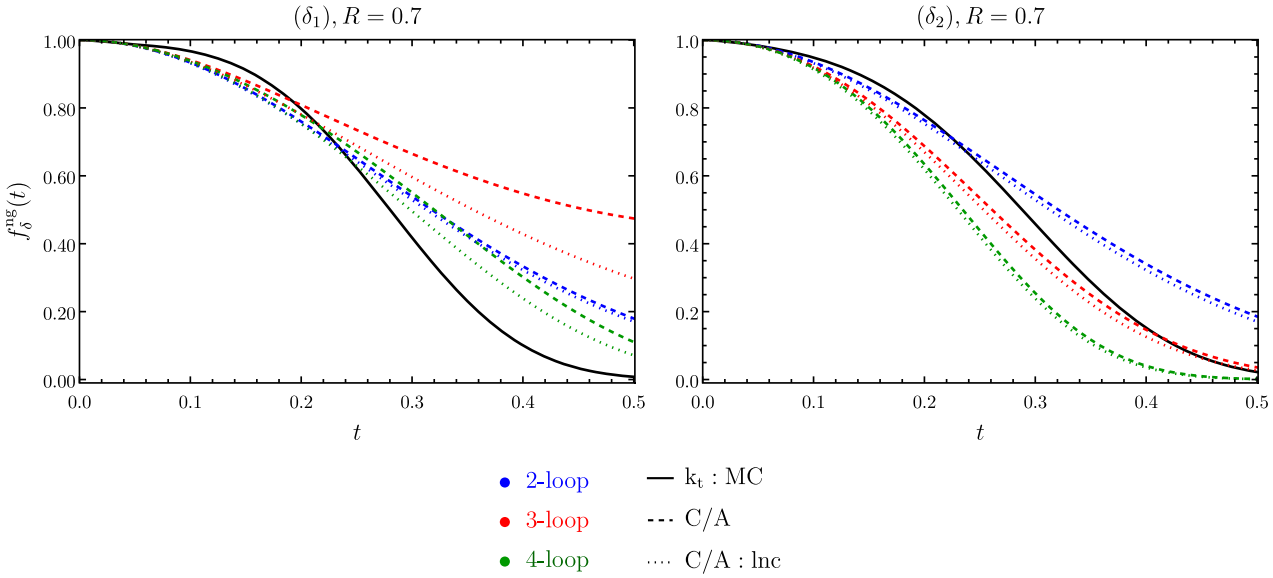



FIG. 9. Comparisons of the fixed-order exponentiated results at two-, three-, and four-loop order for the C/A jet algorithm at finite- (dashed) and large-$N_c$ (dotted). Results are shown for jet radius $R = 0.7$ and for the two distinct Born channels.

remain below 2% for both channels, in agreement with previous findings [8,33,35].

To assess the phenomenological impact of nonglobal logarithms in the C/A clustering more concretely, we show in Fig. 10 the normalized invariant jet-mass differential cross section of the leading-$p_t$ ($p_t > 200$ GeV) jet, summed over the two relevant Born channels, $(\delta_1)$ and $(\delta_2)$, for the specific hadronic process $Z$ + jet at $\sqrt{s} =$ 7 TeV and $R = 0.6$.

As anticipated from Fig. 8, the impact of C/A clustering is of the same magnitude as that of $k_t$ clustering. The position of the peak is shifted slightly to the right by approximately the same amount at all three loop orders, with the reduction in peak height decreasing from approximately 6.43% at two-loop order to 5.67% at four-loop order. These values are nearly identical to those found for the $k_t$ case in Ref. [34]. Compared to the anti-$k_t$ clustering, which induces a reduction of approximately 15.5% in the peak height, the C/A and $k_t$ algorithms perform comparably and both significantly better at reducing the overall impact of nonglobal logarithms. Figure 11 depicts the ratio of the $k_t$ four-loop and C/A three- and four-loop exponential approximations to the two-loop result, further confirming that up to four loops the various truncated approximations exhibit a comparable behavior over a wide range of the invariant jet mass.

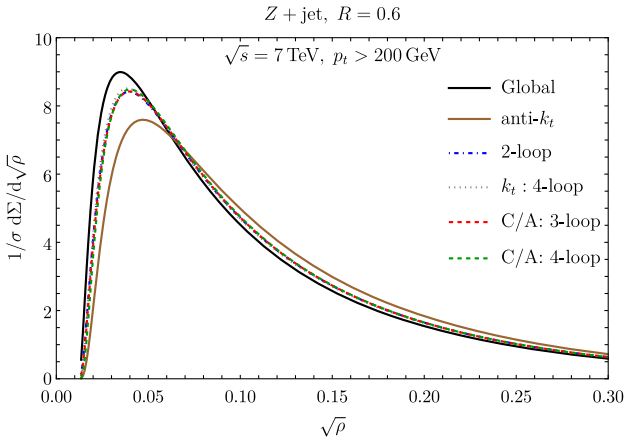



FIG. 10. Various approximations to the normalized differential resummed distribution in the (square-root of the) invariant jet-mass squared for the leading-$p_t$ jet in $Z$ + jet production at $\sqrt{s} = 7$ TeV and $R = 0.6$. The $k_t$ curves are shown for reference. The two-loop exponential is identical for both $k_t$ and C/A; the three- and four-loop results differ between the two algorithms.

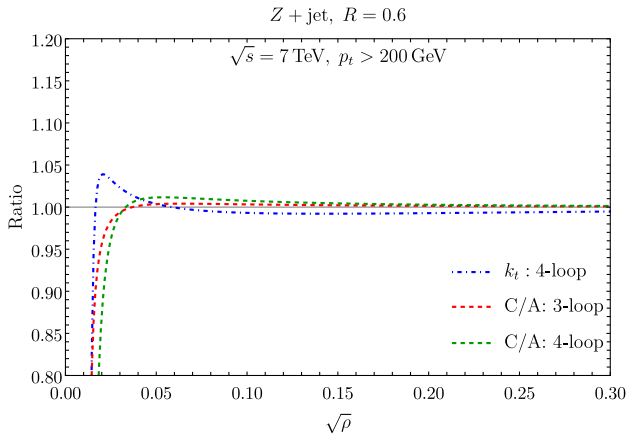



FIG. 11. Ratio of the $k_t$ four-loop and C/A three- and four-loop approximations to the two-loop result.

## V. CONCLUSION

The present paper is an extension of our previous series of works on the structure and impact of nonglobal logarithms appearing in the distribution of generic nonglobal jet observables, where the final-state jets are defined using the three well-known sequential recombination algorithms: anti-$k_t$, $k_t$, and C/A. Those calculations covered jet production in both leptonic and hadronic environments and included the full jet-radius dependence as well as finite-$N_c$ corrections. They were carried out at fixed perturbative order up to four loops. The observable distribution clearly exhibits an exponentiation pattern, which is exploited to construct an all-orders estimate in terms of the coefficients computed at each loop order. The resulting resummed distributions are then validated against the output of numerical MC codes wherever available.

In the present work we have computed the normalized invariant jet-mass distribution for the leading-$p_t$ jet in Higgs- and vector-boson production in association with a jet—a process that has been widely studied by both the phenomenology and experimental communities. The final-state jets are clustered using the Cambridge/Aachen algorithm for various jet radii. Fixed-order calculations are performed through four-loop order in perturbation theory. Applying C/A clustering in this hadronic setting is substantially more complex than the $k_t$ case, owing to the fact that the C/A distance measure is independent of the transverse momenta of the final-state partons; consequently, all possible orderings of the pairwise distances must be taken into account. This task has been fully automated using scripts written in PYTHON together with *Mathematica*.

At two-loop order, the C/A clustering is equivalent to $k_t$, and the results are therefore identical to those reported in our previous work [34]. At three- and four-loop order, the C/A algorithm produces contributions that differ from those of $k_t$, and these are presented here for the first time in the literature. In general, the effect of nonglobal logarithms on the total observable distribution is smaller for C/A than for $k_t$, a feature that holds for the CL and NGL contributions individually as well as for their combined effect. A notable exception occurs for the purely gluonic channel $gg \to gH$ at four-loop order, where the combined CL and NGL contribution in C/A clustering exceeds that of $k_t$ for jet radii $R \lesssim 0.9$, thus making $k_t$ the preferred algorithm for this specific channel over that range. The four-loop calculation is the most demanding, and as a result a separation of the CL and NGL contributions has not been possible at this order. Nevertheless, it is their combined effect that ultimately enters the observable distribution.

The all-orders analysis reveals that the resummed form factor constructed from the fixed-order calculations can be reliably trusted only for $\rho \gtrsim 0.005$ (corresponding to $m_j^2 \gtrsim 0.005 p_t^2$). Furthermore, the overall impact of nonglobal logarithms for C/A clustering has been shown to be of the same order as that for $k_t$ clustering. Both algorithms perform significantly better than anti-$k_t$ in suppressing the impact of these logarithms on the final observable distribution over a wide range of jet radii.

The present work completes the fixed-order study of the effect of the three sequential recombination algorithms on nonglobal observables through four-loop order, for the specific processes involving two and three hard colored legs, namely $e^+e^-$ annihilation and $V/H +$ jet production. Future work should address higher loop orders and/or tackle the all-orders resummation directly, since it appears difficult to reliably infer the latter from fixed-order calculations alone. Other hadronic processes that warrant careful investigation include extending the multiloop analysis to dijet production, which involves a considerably more complex color and kinematic structure.

## DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.


[1] R. Kogler *et al.*, Rev. Mod. Phys. **91**, 045003 (2019).
[2] A. J. Larkoski, I. Moult, and B. Nachman, Phys. Rep. **841**, 1 (2020).
[3] M. Dasgupta and G. P. Salam, Phys. Lett. B **512**, 323 (2001).
[4] M. Dasgupta and G. P. Salam, J. High Energy Phys. 03 (2002) 017.
[5] K. Khelifa-Kerfa and Y. Delenda, J. High Energy Phys. 03 (2015) 094.
[6] A. Banfi, G. Marchesini, and G. Smye, J. High Energy Phys. 08 (2002) 006.
[7] M. D. Schwartz and H. X. Zhu, Phys. Rev. D **90**, 065004 (2014).
[8] Y. Hatta and T. Ueda, Nucl. Phys. **B874**, 808 (2013).
[9] A. Banfi, F. A. Dreyer, and P. F. Monni, J. High Energy Phys. 10 (2021) 006.
[10] A. Banfi, F. A. Dreyer, and P. F. Monni, J. High Energy Phys. 03 (2022) 135.
[11] T. Becher, N. Schalch, and X. Xu, Phys. Rev. Lett. **132**, 081602 (2024).
[12] T. Becher, M. Neubert, D. Y. Shao, and M. Stillger, J. High Energy Phys. 12 (2023) 116.

[13] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (2008) 063.
[14] A. Banfi and M. Dasgupta, Phys. Lett. B **628**, 49 (2005).
[15] R. B. Appleby and M. H. Seymour, J. High Energy Phys. 12 (2002) 063.
[16] Y. Delenda, R. Appleby, M. Dasgupta, and A. Banfi, J. High Energy Phys. 12 (2006) 044.
[17] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, and S. Marzani, J. High Energy Phys. 08 (2010) 064.
[18] K. Khelifa-Kerfa, J. High Energy Phys. 02 (2012) 072.
[19] Y. Delenda and K. Khelifa-Kerfa, J. High Energy Phys. 09 (2012) 109.
[20] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (1997) 001.
[21] M. Wobisch and T. Wengler, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)* (1998), pp. 270–279.
[22] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
[23] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993).
[24] G. P. Salam, Eur. Phys. J. C **67**, 637 (2010).
[25] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, J. High Energy Phys. 09 (2013) 029.
[26] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, J. High Energy Phys. 05 (2014) 146.
[27] K. Khelifa-Kerfa, SciPost Phys. **20**, 094 (2026).
[28] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, J. High Energy Phys. 10 (2012) 126.
[29] N. Ziani, K. Khelifa-Kerfa, and Y. Delenda, Eur. Phys. J. C **81**, 570 (2021).
[30] K. Khelifa-Kerfa, Eur. Phys. J. C **85**, 133 (2025).
[31] K. Khelifa-Kerfa, Phys. Rev. D **111**, 094018 (2025).
[32] K. Khelifa-Kerfa, Phys. Rev. D **111**, 014027 (2025).
[33] K. Khelifa-Kerfa, J. High Energy Phys. 10 (2024) 079.
[34] K. Khelifa-Kerfa and M. Benghanem, J. High Energy Phys. 02 (2026) 009.
[35] Y. Hatta and T. Ueda, Nucl. Phys. **B962**, 115273 (2021).
[36] T. Becher and J. Haag, J. High Energy Phys. 01 (2024) 155.
[37] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, J. High Energy Phys. 08 (2006) 059.
[38] T. Becher, M. Neubert, and D. Y. Shao, Phys. Rev. Lett. **127**, 212002 (2021).
[39] A. Barontini, N. Laurenti, and J. Rojo, *Proc. Sci.*, DIS2024 (2025) 039.
[40] A. Banfi, G. P. Salam, and G. Zanderighi, J. High Energy Phys. 03 (2005) 073.
[41] K. Khelifa-Kerfa and Y. Delenda, Phys. Lett. B **809**, 135768 (2020).
[42] Y. Delenda and K. Khelifa-Kerfa, Phys. Rev. D **93**, 054027 (2016).
[43] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012).
[44] K. Khelifa-Kerfa, QCD resummation for high-$p_T$ jet shapes at hadron colliders, Ph.D. thesis, Manchester University, 2012.
[45] T. Hahn, Comput. Phys. Commun. **168**, 78 (2005).
[46] K. Khelifa-Kerfa and M. Benghanem, Eur. Phys. J. C **85**, 845 (2025).
[47] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, Nucl. Phys. **B407**, 3 (1993).