

https://dspace.org/

Systematics of quark/gluon tagging

Gras, P.; Höche, S.; Kar, D.; Larkoski, A.; Lönnblad, L. ...

2017-07
Springer Verlag
http://hdl.handle.net/10539/23387
Gras, P. et al. 2017. Systematics of quark/gluon tagging. JOURNAL OF HIGH ENERGY PHYSICS 2017(7), Article number 91.

Downloaded from a DSpace repository. https://dspace.org/

JHEP

Published for SISSA by Springer

Received: May 2, 2017
Accepted: July 5, 2017
Published: July 18, 2017

Systematics of quark/gluon tagging

Philippe Gras,[a] **Stefan Höche,**[b] **Deepak Kar,**[c] **Andrew Larkoski,**[d] **Leif Lönnblad,**[e] **Simon Plätzer,**[f,g] **Andrzej Siódmok,**[h,i] **Peter Skands,**[j] **Gregory Soyez**[k,1] **and Jesse Thaler**[l,1]

[a] *IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France*
[b] *SLAC National Accelerator Laboratory, Menlo Park, CA 94025, U.S.A.*
[c] *School of Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa*
[d] *Physics Department, Reed College, Portland, OR 97202, U.S.A.*
[e] *Department of Astronomy and Theoretical Physics, Lund University, Lund, Sweden*
[f] *Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, U.K.*
[g] *Particle Physics Group, School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, U.K.*
[h] *CERN, TH Department, CH-1211 Geneva, Switzerland*
[i] *Institute of Nuclear Physics, Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Kraków, Poland*
[j] *School of Physics and Astronomy, Monash University, Clayton, VIC-3800, Australia*
[k] *IPhT, CEA Saclay, CNRS UMR 3681, F-91191 Gif-sur-Yvette, France*
[l] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

E-mail: philippe.gras@cern.ch, shoeche@slac.stanford.edu, deepak.kar@cern.ch, larkoski@reed.edu, leif.lonnblad@thep.lu.se, simon.platzer@durham.ac.uk, Andrzej.Siodmok@ifj.edu.pl, peter.skands@monash.edu, gregory.soyez@ipht.fr, jthaler@mit.edu

[1] Coordinators for an initial version of this study appearing in ref. [1].

Open Access, © The Authors.
Article funded by SCOAP[3].

https://doi.org/10.1007/JHEP07(2017)091

JHEP07(2017)091

Abstract: By measuring the substructure of a jet, one can assign it a "quark" or "gluon" tag. In the eikonal (double-logarithmic) limit, quark/gluon discrimination is determined solely by the color factor of the initiating parton (C_F versus C_A). In this paper, we confront the challenges faced when going beyond this leading-order understanding, using both parton-shower generators and first-principles calculations to assess the impact of higher-order perturbative and nonperturbative physics. Working in the idealized context of electron-positron collisions, where one can define a proxy for quark and gluon jets based on the Lorentz structure of the production vertex, we find a fascinating interplay between perturbative shower effects and nonperturbative hadronization effects. Turning to proton-proton collisions, we highlight a core set of measurements that would constrain current uncertainties in quark/gluon tagging and improve the overall modeling of jets at the Large Hadron Collider.

Keywords: Jets, QCD Phenomenology

ArXiv ePrint: 1704.03878

JHEP07(2017)091

Contents

1 Overview **1**

2 What is a quark/gluon jet? **5**

3 Quantifying tagging performance **7**
- 3.1 Generalized angularities 7
- 3.2 Classifier separation 10

4 Analytic quark/gluon predictions **12**
- 4.1 Casimir scaling at LL 12
- 4.2 NLL resummation 14
- 4.3 Nonperturbative shape function 14

5 Idealized quark/gluon discrimination **18**
- 5.1 Baseline analysis 18
- 5.2 Parameter dependence 21
- 5.3 Impact of generator settings 24

6 Quark/gluon tagging at the LHC **31**
- 6.1 Defining enriched samples 31
- 6.2 Baseline analysis 33
- 6.3 Parameter dependence 36

7 Summary and recommendations **38**

JHEP07(2017)091

1 Overview

Jets are robust tools for studying short-distance collisions involving quarks and gluons. With a suitable jet definition, one can connect jet measurements made on clusters of hadrons to perturbative calculations made on clusters of partons. More ambitiously, one can try to tag jets with a suitably-defined flavor label, thereby enhancing the fraction of, say, quark-tagged jets over gluon-tagged jets. This is relevant for searches for physics beyond the standard model, where signals of interest are often dominated by quarks while the corresponding backgrounds are dominated by gluons. A wide variety of quark/gluon discriminants have been proposed [2–19], and there is a growing catalog of quark/gluon studies at the Large Hadron Collider (LHC) [20–25].

In order to achieve robust quark/gluon tagging, though, one needs theoretical and experimental control over quark/gluon radiation patterns. At the level of eikonal partons, a hard quark radiates soft gluons proportional to its $C_F = 4/3$ color factor while a hard

gluon radiates soft gluons proportional to $C_A = 3$, and quark/gluon tagging performance is simply a function of C_A/C_F. As we will see, quark/gluon discrimination performance is highly sensitive to perturbative effects beyond the eikonal limit, such as $g \to q\overline{q}$ splittings and color coherence, as well as to nonperturbative effects such as color reconnection and hadronization. While these effects are modeled (to differing degrees) in parton-shower generators, they are relatively unconstrained by existing collider measurements, especially in the gluon channel.

The goal of this paper is to highlight these uncertainties in quark/gluon tagging, using both parton-shower generators and first-principles calculations. We start in the idealized context of electron-positron collisions, where one can study final-state quark/gluon radiation patterns in the absence of initial-state complications. Here, we find modest differences in the predicted distributions for quark/gluon discriminants, which then translate to large differences in the predicted quark/gluon separation power. Motivated by these uncertainties, we propose a set of LHC measurements that should help improve the modeling of jets in general and quark/gluon tagging in particular. A summary and outline of this paper follows.

A common misconception about quark/gluon tagging is that it is an intrinsically ill-defined problem. Of course, quark and gluon partons carry color while jets are composed of color-singlet hadrons, so the labels "quark" and "gluon" are fundamentally ambiguous. But this is philosophically no different from the fact that a "jet" is fundamentally ambiguous and one must therefore always specify a concrete jet finding procedure. As discussed in section 2, one can indeed create a well-defined quark/gluon tagging procedure based on unambiguous hadron-level measurements. In this way, even if what one means by "quark" or "gluon" is based on a naive or ambiguous concept (like Born-level cross sections or eikonal limits), quark/gluon discrimination is still a well-defined technique for enhancing desired signals over unwanted backgrounds.

In order to quantify quark/gluon discrimination power, there is a wide range of possible quark/gluon discriminants and a similarly large range of performance metrics, both discussed in section 3. As a concrete set of discriminants, we consider the generalized angularities λ_β^κ [14] (see also [26–29]),

$$\lambda_\beta^\kappa = \sum_{i \in \text{jet}} z_i^\kappa \theta_i^\beta, \tag{1.1}$$

with the notation to be explained in section 3.1. We consider five different (κ, β) working points, which roughly map onto five variables in common use in the literature:

$(0,0)$	$(2,0)$	$(1,0.5)$	$(1,1)$	$(1,2)$
multiplicity	p_T^D	LHA	width	mass

(1.2)

Here, multiplicity is the hadron multiplicity within the jet, p_T^D was defined in refs. [11, 12], LHA refers to the "Les Houches Angularity" (named after the workshop venue where this study was initiated [1]), width is closely related to jet broadening [30–32], and mass is closely related to jet thrust [33]. To quantify discrimination performance, we focus on

JHEP07(2017)091

classifier separation (a default output of the TMVA package [34]):

$$\Delta = \frac{1}{2}\int \mathrm{d}\lambda \, \frac{\big(p_q(\lambda) - p_g(\lambda)\big)^2}{p_q(\lambda) + p_g(\lambda)}, \tag{1.3}$$

where p_q (p_g) is the probability distribution for λ in a generated quark jet (gluon jet) sample. This and other potential performance metrics are discussed in section 3.2.

To gain a baseline analytic understanding, we use resummed calculations in section 4 to provide a first-order approximation for quark/gluon radiation patterns. For $\kappa = 1$, the generalized angularities are infrared and collinear (IRC) safe, and therefore calculable in (resummed) perturbation theory. At leading-logarithmic (LL) accuracy, the IRC-safe angularities satisfy a property called Casimir scaling, and the resulting classifier separation Δ is a universal function of C_A/C_F, independent of the value of β. At present, the distributions for generalized angularities are known to next-to-leading-logarithmic (NLL) accuracy [13, 14]. Here, we include the resummation of the leading-color nonglobal logarithms [35], though we neglect the resummation of pure jet-radius logarithms [36], and soft single-logarithmic corrections proportional to powers of the jet radius. These NLL calculations are effectively at parton-level, so to obtain hadron-level distributions, we estimate the impact of nonperturbative effects using shape functions [37, 38].

To gain a more realistic understanding with a full hadronization model, we use parton-shower generators in section 5 to predict quark/gluon separation power. In an idealized setup with e^+e^- collisions, we can use the following processes as proxies for quark and gluon jets:

$$\text{“quark jets”}: \quad e^+e^- \to (\gamma/Z)^* \to u\bar{u}, \tag{1.4}$$

$$\text{“gluon jets”}: \quad e^+e^- \to h^* \to gg, \tag{1.5}$$

where h is the Higgs boson. These processes are physically distinguishable by the quantum numbers of the associated color-singlet production operator, giving a way to define truth-level quarks and gluons labels without reference to the final state.[1] We compare seven different parton-shower generators both before hadronization (“parton level”) and after hadronization (“hadron level”):

- Pythia 8.215 [39],
- Herwig++ 2.7.1 [40, 41],[2]
- Sherpa 2.2.1 [43],
- Vincia 2.001 [44],

[1] Of course, the quantum numbers of the color singlet operator are not measurable event by event. The idea here is to have a fundamental definition of “quark” and “gluon” that does not reference QCD partons directly.

[2] We use the default angular-ordered shower for these studies. We also performed preliminary tests with Herwig 7.0.3 [42].

JHEP07(2017)091

- Deductor 1.0.2 [45] (with hadronization performed by Pythia 8.212),[3]
- Ariadne 5.0.β [46],[4]
- Dire 1.0.0 [48] (with cluster hadronization performed by Sherpa 2.1.1).

To test other generators, the analysis code used for this study is available as a Rivet routine [49], which can be downloaded from https://github.com/gsoyez/lh2015-qg.

As we will see, the differences between these generators arise from physics at the interface between perturbative showering and nonperturbative fragmentation. One might think that the largest differences between generators would appear for IRC-unsafe observables like multiplicity and p_T^D, where nonperturbative hadronization plays an important role. Surprisingly, comparably-sized differences are also seen for the IRC-safe angularities, indicating that these generators have different behavior even at the level of the perturbative final-state shower. In section 5.2, we study these differences as a function of the collision energy Q, the jet radius R, and the strong coupling constant α_s, showing that the generators have somewhat different discrimination trends. In section 5.3, we compare the default parton shower configurations to physically-motivated changes, showing that modest changes to the shower/hadronization parameters can give rather large differences in quark/gluon separation power.

At the end of the day, most of the disagreement between generators is due to gluon radiation patterns. This is not so surprising, since most of these generators have been tuned to reproduce distributions from e^+e^- colliders, and quark (but less so gluon) radiation patterns are highly constrained by event shape measurements at LEP [50–53]. In section 6, we suggest a possible analysis strategy at the LHC to specifically constrain gluon radiation patterns. At a hadron collider, the distinction between quark jets and gluon jets is rather subtle, since radiation patterns depend on color connections between the measured final-state jets and the unmeasured initial-state partons. That said, we find that one can already learn a lot from hadron-level measurements, without trying to isolate "pure" quark or gluon samples. In particular, we advocate measuring the generalized angularities on quark/gluon enriched samples:

$$\text{"quark enriched"}: \quad pp \to Z + \text{jet}, \tag{1.6}$$

$$\text{"gluon enriched"}: \quad pp \to \text{dijets}, \tag{1.7}$$

where "enriched" means that the Born-level process contributing to these channels is dominated by the corresponding jet flavor. By making judicious kinematic cuts, we could further flavor-enrich these samples [54], though we will not pursue that in this paper for simplicity.

We present our final recommendations and conclusions in section 7. The main take home message from this study is that, contrary to the standard lore, the e^+e^- measurements currently used for tuning are insufficient to constrain uncertainties in the final state shower. There are alternative e^+e^- measurements, however, that can play an important

[3] Note that this Deductor plus Pythia combination has not yet been tuned to data.

[4] This version of Ariadne is not yet public, but available from the author on request. For e^+e^- collisions, the physics is the same as in Ariadne 4 [47].

JHEP07(2017)091

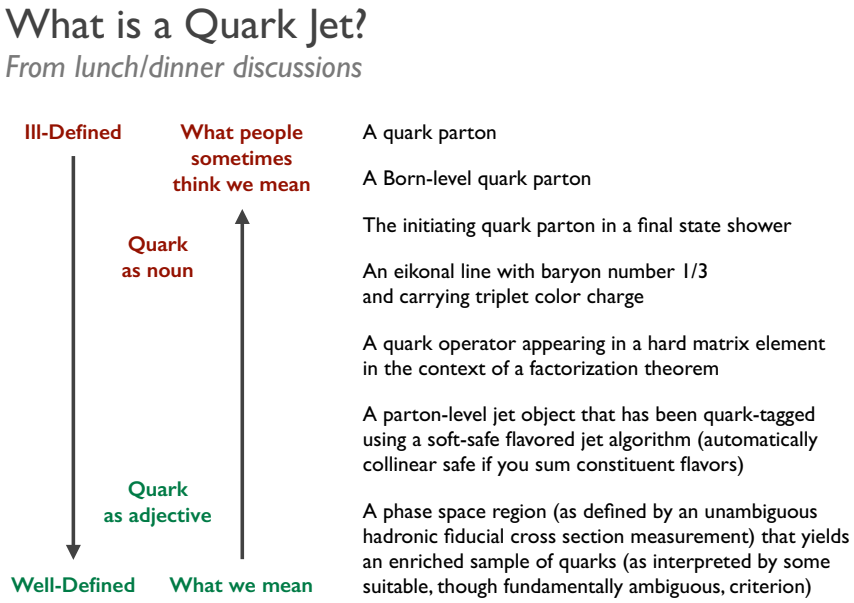


Figure 1. Original slide from the June 10, 2015 summary report of the quark/gluon Les Houches subgroup [1].

role in constraining gluon radiation patterns. Ultimately, gluon-enriched measurements at the LHC will be crucial to achieve robust quark/gluon discrimination.

2 What is a quark/gluon jet?

As part of the 2015 Les Houches workshop on "Physics at TeV Colliders" [1], an attempt was made to define exactly what is meant by a "quark jet" or "gluon jet" (see figure 1). Here are some suggested options for defining a quark jet, in (approximate) order from most ill-defined to most well-defined. Related statement can be made for gluon jets.

A quark jet is...

- *A quark parton.* This definition (incorrectly) assumes that there is a one-to-one map between a jet and its initiating parton. Because it neglects the important role of additional radiation in determining the structure of a jet, we immediately dismiss this definition.

- *A Born-level quark parton.* This definition at least acknowledges the importance of radiative corrections to jet production, but it leaves open the question of how exactly to define the underlying Born-level process from an observed final state. (For one answer valid at the parton level, see flavored jet algorithms below.)

- *An initiating quark parton in a final state parton shower.* We suspect that this is the definition most LHC experimentalists have in mind. This definition assumes that the parton-shower history is meaningful, though, which may not be the case beyond the

JHEP07(2017)091

strongly-ordered or LL approximations. Because the parton shower is semi-classical, this definition neglects the impact of genuinely quantum radiative corrections as well as nonperturbative hadronization.

- *A maximum-p_T quark parton within a jet in a final state parton shower.* This definition uses the hardest parton within the active jet area encountered at any stage of the shower evolution, including the initial hard scattering process. This "max-p_T" prescription is a variant on the initiating parton prescription above (see further discussion in ref. [55]). It differs from the initiating parton by a calculable amount in a LL shower [36] and is based on the same (naive) assumption that the parton-shower history is meaningful.

- *An eikonal line with baryon number 1/3 and carrying triplet color charge.* This is another semi-classical definition that attempts to use a well-defined limit of QCD to define quarks in terms of light-like Wilson lines. Philosophically, this is similar to the parton-shower picture, with a similar concern about how to extrapolate this definition away from the strict eikonal limit.

- *A parton-level jet object that has been quark-tagged using an IRC-safe flavored jet algorithm.* This is the strategy adopted in ref. [56]. While this definition neglects the impact of hadronization, it does allow for the calculation of quark jet cross sections at all perturbative orders, including quantum corrections.

The unifying theme in the above definitions is that they try to identify a quark as an object unto itself, without reference to the specific final state of interest. However, it is well-known that a "quark" in one process may not look like a "quark" in other process, due to color correlations with the rest of the event, especially the initial state in pp collisions. The next definition attempts to deal with the process dependence in defining quarks.

- *A quark operator appearing in a hard matrix element in the context of a factorization theorem.*This is similar to the attitude taken in ref. [54]. In the context of a well-defined cross section measurement, one can (sometimes) go to a limit of phase space where the hard production of short-distance quarks and gluons factorizes from the subsequent long-distance fragmentation. This yields a nice (gauge-covariant) operator definition of a quark jet, which can be made precise for observables based on jet grooming [57, 58]. That said, even if a factorization theorem does exist for the measurement of interest, this definition is potentially ambiguous beyond leading power.

The definition we adopt for this study is inspired by the idea that one should think about quark/gluon tagging in the context of a specific measurement, regardless of whether the observable in question has a rigorous factorization theorem.

- *A phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks (as interpreted by some suitable, though fundamentally ambiguous, criterion).* Here, the goal is to *tag* a phase space region as being quark-like, rather than try to determine a truth definition of

JHEP07(2017)091

a quark. This definition has the advantage of being explicitly tied to hadronic final states and to the discriminant variables of interest. *The main challenge with this definition is how to determine the criterion that corresponds to successful quark enrichment.* For that, we have to rely to some degree on the other less well-defined notions of what a quark jet is.

To better understand this last definition, consider a quark/gluon discriminant λ. Since λ can be measured on any jet, one can unambiguously determine the cross section $\mathrm{d}\sigma/\mathrm{d}\lambda$ for any jet sample of interest. But measuring λ does not directly lead to the probability that the jet is a quark jet, nor to the probability distribution $p_q(\lambda)$ for λ within a quark jet sample. Rather, the process of measuring λ must be followed by a separate process of interpreting how the value of λ should be used as part of an analysis.

For example, the user could choose that small λ jets should be tagged as "quark-like" while large λ jets should be tagged as "gluon-like". Alternatively, the user might combine λ with other discriminant variables as part of a more sophisticated classification scheme. The key point is that one first measures hadron-level discriminant variables on a final state of interest, and only later does one interpret exactly what those discriminants accomplish (which could be different depending on the physics goals of a specific analysis). Typically, one might use a Born-level or eikonal analysis to define which regions of phase space should be associated with "quarks" or "gluons", but even if these phase space regions are based on naive or ambiguous logic, λ itself is a well-defined discriminant variable.

In section 5, we will consider the generalized angularities λ_β^κ as our discriminant variables and we will assess the degree to which the measured values of λ_β^κ agree with a quark/gluon interpretation based on Born-level production modes. This is clearly an idealization, though one that makes some sense in the context of e^+e^- collisions, since truth-level "quark" and "gluon" labels can be defined by the Lorentz structure of the production vertex. In section 6, we will recommend that the LHC experiments perform measurements of λ_β^κ in well-defined hadron-level final states, without necessarily attempting to determine separate $p_q(\lambda_\beta^\kappa)$ and $p_g(\lambda_\beta^\kappa)$ distributions. Eventually, one would want to use these hadron-level measurements to infer something about parton-level quark/gluon radiation patterns. Even without that interpretation step, though, direct measurements of $\mathrm{d}\sigma/\mathrm{d}\lambda_\beta^\kappa$ would provide valuable information for parton-shower tuning. This in turn would help λ_β^κ become a more robust and powerful discriminant in searches for new physics beyond the standard model.

3 Quantifying tagging performance

3.1 Generalized angularities

A wide variety of quark/gluon discriminants have been proposed (see ref. [9] for an extensive catalog), but here we limit ourselves to a two-parameter family of generalized angularities [14], shown in figure 2. These are defined as (repeating eq. (1.1) for convenience)

$$\lambda_\beta^\kappa = \sum_{i \in \mathrm{jet}} z_i^\kappa \theta_i^\beta, \tag{3.1}$$

JHEP07(2017)091

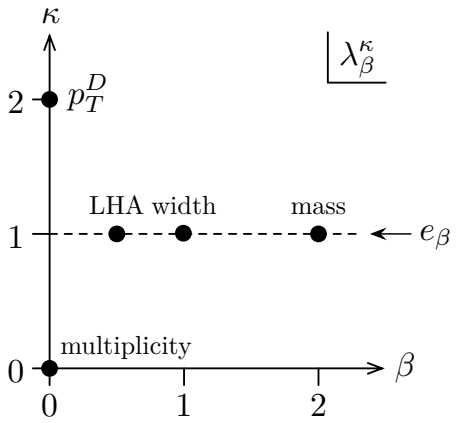


Figure 2. Two-parameter family of generalized angularities, adapted from ref. [14]. The dots correspond to the five benchmark angularities used in this study, with "LHA" referring to the Les Houches Angularity. The horizontal line at $\kappa = 1$ corresponds to the IRC-safe angularities, $e_\beta = \lambda^1_\beta$.

where i runs over the jet constituents, $z_i \in [0, 1]$ is a momentum fraction, and $\theta_i \in [0, 1]$ is a (normalized) angle to the jet axis. The parameters $\kappa \geq 0$ and $\beta \geq 0$ determine the momentum and angle weighting, respectively. For $\kappa = 1$, the generalized angularities are IRC safe and hence calculable in perturbation theory [29] (see also [28, 59–62]), and we will sometimes use the shorthand

$$e_\beta \equiv \lambda^1_\beta. \tag{3.2}$$

For general $\kappa \neq 1$, there are quasi-perturbative techniques based on generalized fragmentation functions [14] (see also [10, 63–65]). In our parton-shower studies, we determine λ^κ_β using all constituents of a jet, though one could also consider using charged-particle-only angularities to improve robustness to pileup (at the expense of losing some particle-level information).

For our e^+e^- study, we cluster jets with FastJet 3.2.1 [66, 67] using the ee-variant of the anti-k_t algorithm [68], with $|\vec{p}|$-ordered winner-take-all recombination [29, 69, 70] to determine the jet axis $\hat{n}$. Unlike standard E-scheme recombination [71], the winner-take-all scheme yields a jet axis $\hat{n}$ that does not necessarily align with the jet three-momentum $\vec{p}$; this turns out to be a desirable feature for avoiding soft recoil effects [13, 29, 30, 72, 73]. We define

$$z_i \equiv \frac{E_i}{E_J}, \qquad \theta_i \equiv \frac{\Omega_{i\hat{n}}}{R}, \tag{3.3}$$

where E_J is the jet energy, E_i is the particle energy, $\Omega_{i\hat{n}}$ is the opening angle to the jet axis, and R is the jet radius (taken to be $R = 0.6$ by default, unless explicitly mentioned otherwise).

For our pp study, we use the standard pp version of anti-k_t with E-scheme recombination, defining

$$z_i \equiv \frac{p_{Ti}}{\sum_{j\in\text{jet}} p_{Tj}}, \qquad \theta_i \equiv \frac{R_{i\hat{n}}}{R}, \tag{3.4}$$

where p_{Ti} is the particle transverse momentum and $R_{i\hat{n}}$ is the rapidity-azimuth distance to the jet axis. To define a recoil-free axis, we recluster the jet using the Cambridge/Aachen

JHEP07(2017)091

(C/A) algorithm [74, 75] with p_T-ordered winner-take-all recombination. Note that with this choice of recombination scheme, the p_T of the recoil-free axis is effectively the scalar sum $\sum_{j\in\text{jet}} p_{Tj}$ used in eq. (3.4). In addition to directly measuring the angularities, we also want to test the impact of jet grooming (see e.g. [76–79]). As one grooming example, we use the modified mass drop tagger (mMDT) with $\mu = 1$ [76, 80] (equivalently, soft drop declustering with $\beta = 0$ [81]). This grooming procedure starts from the C/A-reclustered jet, which yields an angular-ordered clustering tree. This tree is then declustered, removing the softer branch until

$$\frac{\min[p_{T1}, p_{T2}]}{p_{T1} + p_{T2}} > z_{\text{cut}}, \tag{3.5}$$

where 1 and 2 label the two branches of a splitting. By applying the mMDT procedure, we can test how quark/gluon discrimination performance and robustness is affected by removing soft radiation from a jet. For concreteness, we always set $z_{\text{cut}} = 0.1$ to match the studies in refs. [81–84].

By adjusting κ and β in the angularities, one can probe different aspects of the jet fragmentation. We consider five benchmark values for (κ, β) indicated by the black dots in figure 2:

$$\begin{aligned}
(0,0) &= \text{hadron multiplicity},\\
(2,0) &\Rightarrow p_T^D \text{[11, 12] (specifically } \lambda_0^2 = (p_T^D)^2),\\
(1,0.5) &= \text{Les Houches Angularity (LHA)},\\
(1,1) &= \text{width or broadening [30–32]},\\
(1,2) &\Rightarrow \text{mass or thrust [33] (specifically } \lambda_2^1 \simeq m_{\text{jet}}^2/E_{\text{jet}}^2).
\end{aligned} \tag{3.6}$$

Except for the LHA, these angularities (or their close cousins) have already been used in quark/gluon discrimination studies. The LHA has been included to have an IRC safe angularity that weights energies more heavily than angles, similar in spirit to the $\beta = 0.2$ value advocated in ref. [13] for energy correlation functions. Most of the results in this paper are shown in terms of the LHA; results for the other four benchmark values are available in the source files as supplementary material of this paper, where each figure in this paper corresponds to multipage file.

For the IRC-safe case of $\kappa = 1$, there is an alternative version of the angularities based on energy correlation functions [13] (see also [73, 85, 86]),

$$\text{ecf}_\beta = \sum_{i<j\in\text{jet}} z_i z_j \theta_{ij}^\beta \simeq \lambda_\beta^1, \tag{3.7}$$

where equality holds in the extreme eikonal limit.[5] For the e^+e^- case, the pairwise angle θ_{ij} is typically normalized to the jet radius as $\theta_{ij} \equiv \Omega_{ij}/R$. To avoid a proliferation of curves, we will not show any results for ecf_β. We will also neglect quark/gluon discriminants that take into account azimuthal asymmetries within the jet, though observables like the covariance tensor [9] and 2-subjettiness [87–89] can improve quark/gluon discrimination.

[5] This equality also relies on using a recoil-free axis choice $\hat{n}$ to define θ_i. Amusingly, $\lim_{\beta\to 0} \text{ecf}_\beta = (1 - \lambda_0^2)/2$ (i.e. $\kappa = 2$, $\beta = 0$), such that the $\beta \to 0$ limit of the IRC-safe energy correlation functions corresponds to the IRC-unsafe p_T^D.

JHEP07(2017)091

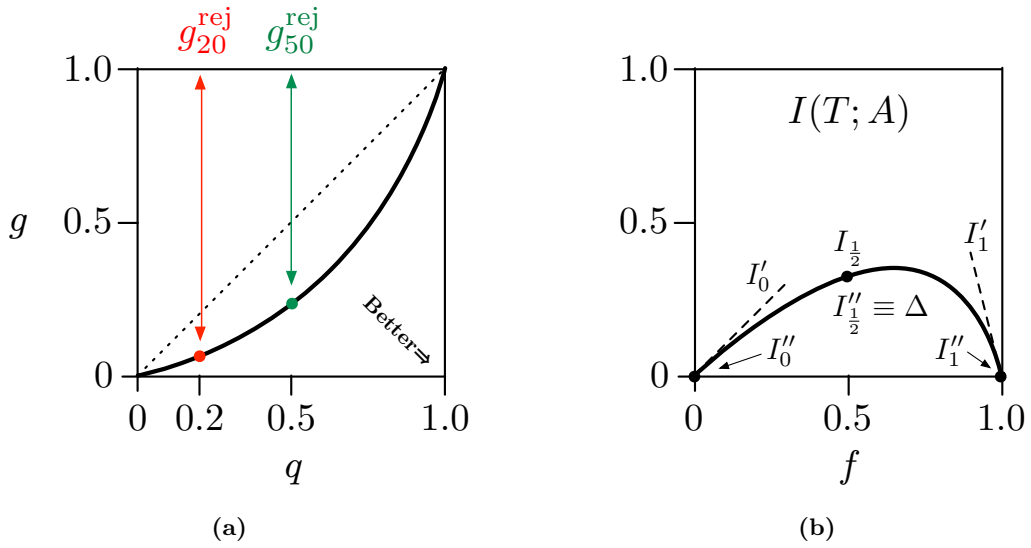


Figure 3. Alternative metrics for discrimination power. (a) A ROC curve, showing the gluon rejection rate at fixed 20% and 50% quark efficiency. While not shown in this paper, some of this ROC information is available in the source files as supplementary material. (b) Mutual information $I(T;A)$ as a function of the quark fraction f, showing the relationship to classifier separation $\Delta \equiv I''_{\frac{1}{2}}$ and other information-theoretic quantities. The mutual information at $f = 1/2$ (i.e. $I_{\frac{1}{2}}$) is also available in the source files as supplementary material of this paper.

See ref. [16] for a related study of quark/gluon systematics for shower deconstruction [90–92] and energy correlation functions [13].

3.2 Classifier separation

Since we will be testing many parton-shower variants, we need a way to quantify quark/gluon separation power in a robust way that can easily be summarized by a single number. For that purpose we use classifier separation (repeating and reorganizing eq. (1.3) for convenience),

$$\Delta = \frac{1}{2}\int \mathrm{d}\lambda \, \frac{\big(p_q(\lambda) - p_g(\lambda)\big)^2}{p_q(\lambda) + p_g(\lambda)} = 1 - 2\int \mathrm{d}\lambda \, \frac{p_q(\lambda)\, p_g(\lambda)}{p_q(\lambda) + p_g(\lambda)}, \tag{3.8}$$

where p_q (p_g) is the probability distribution for the quark jet (gluon jet) sample as a function of the classifier λ. Here, $\Delta = 0$ corresponds to no discrimination power and $\Delta = 1$ corresponds to perfect discrimination power.

A more common way to talk about discrimination power is in terms of receiver operating characteristic (ROC) curves, shown in figure 3a. At a point (q,g) on the ROC curve, where $q, g \in [0, 1]$, one can define a selection that yields q efficiency for quarks and g mistag rate for gluons, or equivalently, a $(1 - g)$ efficiency for gluons for a $(1 - q)$ mistag rate for quarks. There are various ways to turn the ROC curve into a single number, and in the source files attached as supplementary material to this paper, every figure for Δ is part of

JHEP07(2017)091

a multipage file that also has results for

$$\{g_{20}^{\text{rej}}, g_{50}^{\text{rej}}\}: \quad \text{Gluon rejection rate at } \{20\%, 50\%\} \text{ quark efficiency;} \tag{3.9}$$

$$\{q_{20}^{\text{rej}}, q_{50}^{\text{rej}}\}: \quad \text{Quark rejection rate at } \{20\%, 50\%\} \text{ gluon efficiency;} \tag{3.10}$$

$$s^{\text{rej}}: \quad \text{Symmetric rejection rate at } s^{\text{rej}} \text{ efficiency.} \tag{3.11}$$

Since we are more interested in understanding the relative performance between parton showers rather than the absolute performance, we will not show full ROC curves in this paper, though they can be easily derived from the p_q and p_g distributions. If one observable has an everywhere better ROC curve than another (i.e. it is Pareto optimal), then it will also have a larger Δ value. The converse is not true, however, since depending on the desired working point, a "bad" discriminant as measured by Δ might still be "good" by another metric. In that sense, Δ contains less information than the full ROC curve.

An alternative way to quantify discrimination power is through mutual information, which counts the number of "bits" of information gained from measuring a discriminant variable (see ref. [14]). Given a sample with quark fraction $f \in [0, 1]$ and gluon fraction $(1 - f)$, the mutual information with the truth (a.k.a. the truth overlap) is

$$I(T; \Lambda) = \int \mathrm{d}\lambda \left(f\, p_q(\lambda) \log_2 \frac{p_q(\lambda)}{p_{\text{tot}}(\lambda)} + (1 - f)\, p_g(\lambda) \log_2 \frac{p_g(\lambda)}{p_{\text{tot}}(\lambda)} \right), \tag{3.12}$$

where $T = \{q, g\}$ is the set of truth labels, $\Lambda = \{\lambda\}$ is the (continuous) set of discriminant values, and

$$p_{\text{tot}}(\lambda) = f\, p_q(\lambda) + (1 - f)\, p_g(\lambda). \tag{3.13}$$

The choice $f = \frac{1}{2}$ was used in ref. [14] and is also available in the source files as supplementary material of this paper,

$$I(T; A)\big|_{f=\frac{1}{2}} \equiv I_{\frac{1}{2}}, \tag{3.14}$$

though other f choices are plausible.

Though we will not use mutual information in this study, it is amusing to note that the second derivative of $I(T; \Lambda)$ with respect to f is related to classifier separation as

$$-\frac{\log 2}{4} \frac{\partial^2 I(T; \Lambda)}{\partial f^2}\bigg|_{f=\frac{1}{2}} \equiv I''_{\frac{1}{2}} = \Delta. \tag{3.15}$$

More broadly, the dependence of $I(T; A)$ on f can be related to other concepts in statistics, as visualized in figure 3b. At $f = 0$ and $f = 1$, the mutual information itself is zero, but the derivatives are:

$$\frac{\partial I}{\partial f}\bigg|_{f=0} \equiv I'_0 = \int \mathrm{d}\lambda\, p_q(\lambda) \log_2 \frac{p_q(\lambda)}{p_g(\lambda)}, \quad -\log 2\, \frac{\partial^2 I}{\partial f^2}\bigg|_{f=0} \equiv I''_0 = \int \mathrm{d}\lambda\, \frac{p_q(\lambda)^2}{p_g(\lambda)}, \tag{3.16}$$

$$-\frac{\partial I}{\partial f}\bigg|_{f=1} \equiv I'_1 = \int \mathrm{d}\lambda\, p_g(\lambda) \log_2 \frac{p_g(\lambda)}{p_q(\lambda)}, \quad -\log 2\, \frac{\partial^2 I}{\partial f^2}\bigg|_{f=1} \equiv I''_1 = \int \mathrm{d}\lambda\, \frac{p_g(\lambda)^2}{p_q(\lambda)}. \tag{3.17}$$

JHEP07(2017)091

The first derivative is sometimes called relative entropy and the second derivative is sometimes called discrimination significance. Unlike classifier separation, these later metrics do not treat quark and gluon distributions symmetrically.

One advantage of Δ over $I(T;\Lambda)$ is that the integrand in eq. (3.8) is easier to interpret, since it tracks the fractional difference between the signal and background at a given value of λ.[6] Specifically, by plotting

$$\frac{\mathrm{d}\Delta}{\mathrm{d}\lambda} = \frac{1}{2}\frac{\big(p_q(\lambda) - p_g(\lambda)\big)^2}{p_q(\lambda) + p_g(\lambda)}, \tag{3.18}$$

one can easily identify which regions of phase space contribute the most to quark/gluon discrimination. One can then ask whether or not the regions exhibiting the most separation power are under sufficient theoretical control, including both the size of perturbative uncertainties and the impact of nonperturbative corrections.

4 Analytic quark/gluon predictions

For the IRC-safe angularities with $\kappa = 1$ (namely $e_{0.5}$, e_1, and e_2 from eq. (3.6)), we can use analytic calculations to get a baseline expectation for the degree of quark/gluon separation. At LL accuracy, a jet effectively consists of a single soft gluon emission from a hard quark/gluon, with a suitable Sudakov form factor coming from vetoing additional radiation. At this order, the strong coupling constant is fixed and only the leading splitting function is used. In particular, the IRC-safe angularities at LL order satisfy a property called Casimir scaling (reviewed below), such that the discrimination power is independent of β.

At NLL order, the jet is described by multiple gluon emissions from a hard quark/gluon, including the effects of α_s running and subleading terms in the splitting function, but neglecting matrix element corrections and energy-momentum conservation. For the IRC-safe angularities, the NLL-accurate distributions were calculated in refs. [13, 14] (see also [26, 28, 29, 73, 93]). For the generalized angularities, generalized fragmentation functions [10, 14, 63–65] were used to extend the NLL calculation beyond the IRC-safe regime, though we will not use that technique in the present paper. To date, the impact of soft nonperturbative physics has not been included in the distributions for the IRC-safe angularities, but we do so below.

4.1 Casimir scaling at LL

As shown in refs. [13, 14], the IRC-safe angularities satisfy Casimir scaling at LL accuracy, which implies that the quark/gluon discrimination performance only depends on the color factor ratio C_A/C_F. To see this, let us first introduce the notation for the cumulative distribution $\Sigma(\lambda)$, which is defined by

$$\Sigma(\lambda) = \int_0^\lambda \mathrm{d}\lambda'\, p(\lambda'), \qquad p(\lambda) = \frac{\mathrm{d}\Sigma}{\mathrm{d}\lambda}. \tag{4.1}$$

[6] Another advantage of Δ over $I(T;\Lambda)$ arises when trying to assign statistical uncertainties to finite Monte Carlo samples. Since Δ is defined as a simple integral, one can use standard error propagation to assign uncertainties to Δ. By contrast, because of the logarithms in the $I(T;\Lambda)$ integrand, one has to be careful about a potential binning bias [14].

JHEP07(2017)091

If an observable satisfies Casimir scaling, then the quark and gluon cumulative distributions can be written as

$$\Sigma_q(\lambda) = e^{-C_F\, r(\lambda)}, \qquad \Sigma_g(\lambda) = e^{-C_A\, r(\lambda)}, \tag{4.2}$$

where $r(\lambda)$ is a monotonically decreasing function of λ. Here, the only difference between the quark and gluon distributions is in the color factors $C_F = 4/3$ versus $C_A = 3$.

At LL accuracy, the distributions for the IRC-safe angularities take precisely this form [13, 14],[7]

$$\Sigma_i^{\mathrm{LL}}(e_\beta) = \exp\left[-\frac{\alpha_s C_i}{\pi\beta}\log^2 e_\beta\right], \tag{4.3}$$

where i labels the jet flavor. This LL result can be understood from the fact that quarks and gluons have the same leading splitting function up to an overall multiplicative color factor

$$P_i(z) \simeq \frac{2C_i}{z}, \tag{4.4}$$

and therefore the Sudakov form factor (which is what appears in eq. (4.3)) differs only by the color factor in the exponent.

Observables that satisfy Casimir scaling have universal ROC curves [13] and universal truth overlaps [14], which are independent of the precise functional form of $r(\lambda)$ and only depends on the ratio C_A/C_F. We can derive the same universality for classifier separation. Using

$$p_i(\lambda) = -C_i\, r'(\lambda)\, e^{-C_i\, r(\lambda)} \tag{4.5}$$

and the change of variables $u \equiv e^{-C_F\, r(\lambda)}$, we have

$$\Delta = -\frac{1}{2}\int \mathrm{d}\lambda\, r'(\lambda)\frac{\left(C_F\, e^{-C_F\, r(\lambda)} - C_A\, e^{-C_A\, r(\lambda)}\right)^2}{C_F\, e^{-C_F\, r(\lambda)} + C_A\, e^{-C_A\, r(\lambda)}}, \tag{4.6}$$

$$= \frac{1}{2}\int_0^1 \frac{\mathrm{d}u}{u}\frac{\left(u - (C_A/C_F)u^{C_A/C_F}\right)^2}{u + (C_A/C_F)u^{C_A/C_F}}, \tag{4.7}$$

$$= 2\left({}_2F_1\left[1, \frac{C_F}{C_A - C_F}; \frac{C_A}{C_A - C_F}; -\frac{C_A}{C_F}\right] - \frac{1}{2}\right), \tag{4.8}$$

where ${}_2F_1[a,b;c;z]$ is the hypergeometric function.[8] For the case of QCD with $C_A/C_F = 9/4$,

$$\Delta_{\mathrm{QCD}} \simeq 0.1286. \tag{4.9}$$

We mark this benchmark value with an arrow on the subsequent plots for reference.[9]

Going beyond LL accuracy, Casimir scaling is typically violated, and Δ depends on the precise observable in question. Thus, all of the differences we see in our subsequent studies are effects that are truly higher-order or nonperturbative.

[7] Strictly speaking, eq. (4.3) is only valid in the fixed-coupling approximation. Running-coupling corrections already arise at LL accuracy, replacing $\alpha_s \log^2 e_\beta$ by an all-orders series $g_1(\alpha_s \log e_\beta)\log e_\beta$. This does not affect the property of Casimir scaling.

[8] An alternative way to derive this result is to take $I(T;A)$ from ref. [14] and use eq. (3.15) to extract Δ.

[9] In large N_c QCD where $C_A/C_F \to 2$, $\Delta \to \ln 3 - 1 \simeq 0.0986$, so quark/gluon separation is expected to be more challenging as $N_c \to \infty$.

JHEP07(2017)091

4.2 NLL resummation

Going to NLL accuracy is straightforward for global logarithms, using the formalism of ref. [73]. Nonglobal logarithms have not been included in previous angularity calculations, but we include them here using their numerical extraction in the large N_c limit [35]. We always assume that the jet radius R is order 1 so we can ignore $\log R$ resummation [36].

The cumulative distribution for an IRC safe angularity e_β takes the form [35, 73]

$$\Sigma(e_\beta) = \frac{e^{-\gamma_E R'(e_\beta)}}{\Gamma\left(1 + R'(e_\beta)\right)} e^{-R(e_\beta)} e^{-f_{\text{NGL}}(e_\beta)}, \tag{4.10}$$

where $R(e_\beta)$ is known as the radiator function, $f_{\text{NGL}}(e_\beta)$ encodes nonglobal logarithms, γ_E is the Euler-Mascheroni constant, Γ is the gamma function, and primes indicate logarithmic derivatives:

$$R'(e_\beta) = -\frac{\partial}{\partial \log e_\beta} R(e_\beta). \tag{4.11}$$

The $e^{-R(e_\beta)}$ factor in eq. (4.10) is just the Sudakov form factor, which exhibits Casimir scaling at LL accuracy, and the prefactor containing $R'(e_\beta)$ captures the effect of multiple emissions on the e_β distribution. The $e^{-f_{\text{NGL}}(e_\beta)}$ factor comes from eq. (18) of ref. [35], where it is called $\mathcal{S}$. Note that f_{NGL} is proportional to $C_i\, C_A$, so it effectively preserves Casimir scaling; for this reason, the inclusion of nonglobal logarithms is not expected to have a large impact on quark/gluon separation power.

For the IRC-safe angularities, the radiator function is [13, 14]

$$R(e_\beta) = C_i \int_0^1 \frac{\mathrm{d}\theta}{\theta} \int_0^1 \mathrm{d}z\, p_i(z) \frac{\alpha_s(k_t)}{\pi} \Theta\left(z\theta^\beta - e_\beta\right), \tag{4.12}$$

and the strong coupling is evaluated with two-loop running at the k_t emission scale

$$k_t = z\,\theta\,R\,E_J. \tag{4.13}$$

The reduced splitting functions (i.e. splitting functions summed over all allowed $1 \to 2$ processes) are

$$p_q(z) = \frac{1 + (1-z)^2}{z}, \qquad p_g(z) = 2\frac{1-z}{z} + z(1-z) + \frac{T_R\, n_f}{C_A}(z^2 + (1-z)^2), \tag{4.14}$$

where $T_R = 1/2$ and we take the number of light quark flavors to be $n_f = 5$. Following ref. [14], both $R(e_\beta)$ and $R'(e_\beta)$ are truncated to only keep terms that are formally of NLL accuracy.

4.3 Nonperturbative shape function

The quark/gluon studies in refs. [13, 14] used solely the distributions as calculated in eq. (4.10) (with $f_{\text{NGL}} = 0$). As we will see, nonperturbative hadronization has a big effect in our parton-shower studies, so we would like to include the corresponding effect in our analytic results.

JHEP07(2017)091

The IRC-safe angularities are additive observables, meaning that at leading power in the small e_β limit, one can decompose them into separate contributions from perturbative and nonperturbative modes:

$$e_\beta \simeq e_\beta^{(\mathrm{pert})} + e_\beta^{(\mathrm{NP})}. \tag{4.15}$$

One can then convolve the perturbative distribution for $\hat{e}_\beta \equiv e_\beta^{(\mathrm{pert})}$ with a nonperturbative shape function F that describes the distribution of $\epsilon \equiv e_\beta^{(\mathrm{NP})}$ [37, 38] (see also [94–98]),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}e_\beta} = \int \mathrm{d}\hat{e}_\beta \, \mathrm{d}\epsilon \, \hat{\sigma}(\hat{e}_\beta) \, F(\epsilon) \, \delta\left(e_\beta - \hat{e}_\beta - \epsilon\right), \tag{4.16}$$

where $\hat{\sigma} \equiv \mathrm{d}\hat{\sigma}/\mathrm{d}\hat{e}_\beta$ refers to the perturbative result. The shape function prescription gives sensible results in the small e_β limit, but it breaks down at large values of e_β, since the convolution in eq. (4.16) can yield e_β values that extend beyond the physical range. To address this, we need to smoothly turn off the nonperturbative shift as $\hat{e}_\beta$ approaches the physical endpoint $e_\beta^{\max}$. There is no unique way to do this, but we find sensible results using

$$\frac{\mathrm{d}\sigma}{\mathrm{d}e_\beta} = \int \mathrm{d}\hat{e}_\beta \, \mathrm{d}\epsilon \, \hat{\sigma}(\hat{e}_\beta) \, F(\epsilon) \, \delta\left(e_\beta - e_\beta^{\max} \frac{\hat{e}_\beta + \epsilon}{e_\beta^{\max} + \epsilon}\right), \tag{4.17}$$

which ensures that the cross section normalization is not modified even when the impact of the shape function is suppressed. For simplicity, we take $e_\beta^{\max} = 1$ for all of our distributions, though in practice the perturbative distributions do not extend out that far.

The shape function F has to be extracted from data, but we can use simple parametrizations that account for some aspects of its known behavior:

$$F(\epsilon) = \frac{\pi\epsilon}{2\epsilon_0^2} \exp\left[-\frac{\pi}{4}\frac{\epsilon^2}{\epsilon_0^2}\right], \qquad F_{\mathrm{alt}}(\epsilon) = \frac{4\epsilon}{\epsilon_0^2} \exp\left[-\frac{2\epsilon}{\epsilon_0}\right]. \tag{4.18}$$

Both of these functions go linearly to zero at $\epsilon \to 0$, fall exponentially as $\epsilon \to \infty$, are normalized by $\int \mathrm{d}\epsilon \, F(\epsilon) = 1$, and have an expectation value

$$\langle \epsilon \rangle = \int \mathrm{d}\epsilon \, \epsilon \, F(\epsilon) = \epsilon_0. \tag{4.19}$$

The parameter ϵ_0 can therefore be interpreted as the average shift of the perturbative distribution from nonperturbative effects. The second form in eq. (4.18) was used in ref. [99], but we take the first form as our default since it has a less pronounced high-side tail.

Following ref. [59] (see also [73, 94]), one can estimate ϵ_0 as a function of β. Nonperturbative modes have $k_t \simeq \Lambda_{\mathrm{QCD}}$, so eq. (4.13) implies the relationship

$$z\,\theta \simeq \frac{\Lambda_{\mathrm{QCD}}}{R\,E_J}. \tag{4.20}$$

Appealing to local parton-hadron duality [100], we can estimate the average contribution to $z\,\theta^\beta$ from nonperturbative soft gluon emissions as

$$\begin{aligned} \epsilon_0 &= \frac{\Omega_0}{R\,E_J} \int_0^1 \frac{\mathrm{d}z}{z} \int_0^1 \frac{\mathrm{d}\theta}{\theta} \, z\,\theta^\beta \, \delta\left(z\theta - \frac{\Xi_0}{R\,E_J}\right) \\ &= \frac{1}{\beta - 1} \frac{\Omega_0}{R\,E_J} \left(1 - \left(\frac{\Xi_0}{R\,E_J}\right)^{\beta-1}\right), \end{aligned} \tag{4.21}$$

JHEP07(2017)091

where we are using the soft (and collinear) gluon matrix element to determine the phase space integration, and Ω_0 and Ξ_0 are nonperturbative parameters that are both of order Λ_{QCD}. This estimate of ϵ_0 can also be understood by considering two types of nonperturbative modes:

$$\text{NP Soft}: \quad z \simeq \frac{\Lambda_{\text{QCD}}}{R\,E_J}, \quad \theta \simeq 1, \tag{4.22}$$

$$\text{NP Collinear}: \quad z \simeq 1, \qquad \theta \simeq \frac{\Lambda_{\text{QCD}}}{R\,E_J}. \tag{4.23}$$

These contribute to the angularities as

$$e_\beta^{(\text{NP})} \simeq \underbrace{\frac{\Lambda_{\text{QCD}}}{R\,E_J}}_{\text{soft}} + \underbrace{\left(\frac{\Lambda_{\text{QCD}}}{R\,E_J}\right)^\beta}_{\text{collinear}}. \tag{4.24}$$

We see that soft modes dominate for $\beta > 1$, collinear modes dominate for $\beta < 1$, and soft and collinear modes are equally important for $\beta = 1$. This behavior is indeed encoded in eq. (4.21), which smoothly interpolates between these regimes, yielding a logarithmic structure of $\frac{\Omega_0}{R\,E_J} \log \frac{R\,E_J}{\Xi_0}$ for $\beta = 1$ exactly.

Comparing quarks and gluons, we expect the overall size of the nonperturbative shift Ω_0 should scale proportional to the Casimir factors, as in eq. (4.4). The scaling of Ξ_0 is less clear, since it controls nonperturbative collinear radiation, which is less well studied than nonperturbative soft radiation. For our baseline distributions, we assume that Ξ_0 also obeys Casimir scaling:

$$\frac{\Omega_0^g}{\Omega_0^q} = \frac{\Xi_0^g}{\Xi_0^q} = \frac{C_A}{C_F}. \tag{4.25}$$

By tying Ω_0 and Ξ_0 together, this has the effect of reducing the phase space for nonperturbative emissions from gluons, which is particularly important for $\beta < 1$. Ideally, one would want a more rigorous justification for the assumptions in eq. (4.25) (as well as the convolution structure in eq. (4.17)), though that is beyond the scope of the present work.

To test whether eq. (4.21) is a plausible estimate for nonperturbative corrections, we take the parton-shower generators studied in the next section and study how the average value of e_β shifts as hadronization is turned off and on, and use that to estimate ϵ_0. We emphasize that a hadronization model used with a parton shower is not the same as a shape function in an analytic calculation, so one has to be careful drawing conclusions about the size of ϵ_0 from a study like this. In particular, effects that are captured by $F(\epsilon)$ in an analytic calculation could either be part of the perturbative showering or nonperturbative hadronization in a generator. That said, we expect that the scaling of the e_β shift as a function of E_J, R, and β should be roughly the same.

In figure 4, we show the size of the e_β shift as a function of $R\,E_J$ for the three benchmark β values, where the band indicates the minimum and maximum shifts seen among Pythia, Herwig, Sherpa, Vincia, Deductor, Ariadne, and Dire.[10] We then compare to the expected shift from eq. (4.21) with

$$\Omega_0^i = C_i \times 0.23\,\text{GeV}, \qquad \Xi_0^i = C_i \times 0.37\,\text{GeV}, \tag{4.26}$$

[10] We verified that the overall conclusions do not change when considering the separate scaling of R and E_J.

JHEP07(2017)091

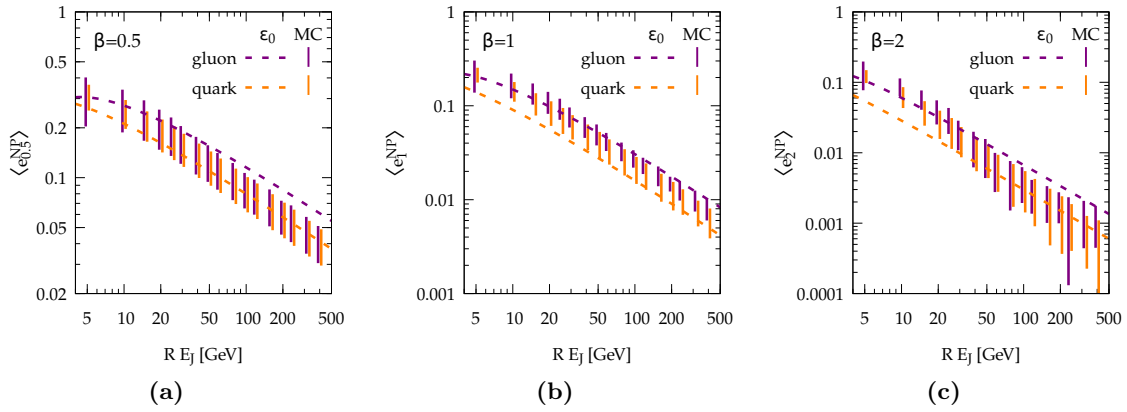


Figure 4. Average nonperturbative shifts to the IRC-safe angularities as a function of $R\,E_J$ for (a) $\beta = 0.5$, (b) $\beta = 1$, and (c) $\beta = 2$. The vertical bands correspond to the range of shifts seen by turning hadronization off and on in the different parton-shower generators. The dashed line corresponds to the best fit to Ω_0 and Ξ_0, assuming the functional form for ϵ_0 in eq. (4.21) and the assumption of Casimir scaling in eq. (4.25). While the $R\,E_J$ trend agrees, the hadronization corrections (as implemented in the parton showers) do not appear to exhibit Casimir scaling.

where these values are obtained by doing a (logarithmic) fit to all of the parton-shower shift values. While the parton shower trends with $R\,E_J$ and β roughly agree with eq. (4.21), there is no evidence for the Casimir scaling hypothesis in eq. (4.25).[11] This is most likely because parton showers already achieve some degree of Casimir scaling through multiple (perturbative) soft gluon emissions from the shower. Despite this caveat, we use the extracted values from eq. (4.26) for our baseline distributions below.

We also consider two alternative scaling behaviors for Ω_0 and Ξ_0. The first alternative is motivated by the observation that, as far as the perturbative soft gluon matrix element is concerned, the Casimir factor affects the rate of soft gluon emissions but not the associated kinematics. Thus, one might expect the overall Ω_0 factor in eq. (4.21) to respect Casimir scaling, but not the Ξ_0 factor inside the delta function. We therefore test a variant with

$$\text{No } C_i \text{ in } \Xi_0: \qquad \Omega_0^i = C_i \times 0.22\,\text{GeV}, \qquad \Xi_0^i = 0.70\,\text{GeV}, \tag{4.27}$$

where again these values are estimated by fitting to the parton shower e_β shifts. As shown in figure 15a below, eq. (4.27) leads to a dramatic increase in the predicted quark/gluon separation power for $\beta < 1$. The second alternative is motivated by the absence of any evidence of Casimir scaling in the parton showers from figure 4. We therefore try taking both nonperturbative parameters to be independent of C_i, with

$$\text{No } C_i \text{ in } \epsilon_0: \qquad \Omega_0^i = 0.44\,\text{GeV}, \qquad \Xi_0^i = 0.70\,\text{GeV}, \tag{4.28}$$

which leads to a corresponding decrease in separation power, since the nonperturbative shape function now has the same behavior for quarks and gluons.

[11] This conclusion is not simply an artifact of the fitting procedure, as none of the individual generators show evidence for Casimir scaling in the nonperturbative shift either.

JHEP07(2017)091

5 Idealized quark/gluon discrimination

We now turn to parton-shower studies of quark/gluon discrimination, starting with the idealized case of e^+e^- collisions. While far less complicated than quark/gluon tagging in the LHC environment, this e^+e^- case study demonstrates the importance of final-state evolution for quark/gluon discrimination, independent from initial-state complications arising in pp collisions. A Rivet routine [49] for this analysis can be downloaded from https://github.com/gsoyez/lh2015-qg under `MC_LHQG_EE.cc`.

To define the truth-level jet flavor, we use a simple definition: a quark jet is a jet produced by a parton-shower event generator in $e^+e^- \to (\gamma/Z)^* \to u\bar{u}$ hard scattering, while a gluon jet is a jet produced in $e^+e^- \to h^* \to gg$. Of course, an $e^+e^- \to u\bar{u}$ event can become a $e^+e^- \to u\bar{u}g$ event after one step of shower evolution, just as $e^+e^- \to gg$ can become $e^+e^- \to gu\bar{u}$. This illustrates the inescapable ambiguity in defining jet flavor.[12] To partially mitigate the effect of wide-angle emissions, we restrict our analysis to jets that satisfy

$$\frac{E_{\text{jet}}}{Q/2} > 0.8, \tag{5.1}$$

where Q is the center-of-mass collision energy, allowing for up to two jets studied per event. Note that this condition acts as a restriction on out-of-jet radiation, which already suppresses to some extent non-global effects [35].[13] There is also the ambiguity of which parton shower to use, so we investigate quark/gluon radiation patterns in several event generators: Pythia 8.215 [39], Herwig 2.7.1 [40, 41], Sherpa 2.2.1 [43], Vincia 2.001 [44], Deductor 1.0.2 [45] (with hadronization performed by Pythia), Ariadne 5.0.β [46], and Dire 1.0.0 [48] (with cluster hadronization performed by Sherpa).

5.1 Baseline analysis

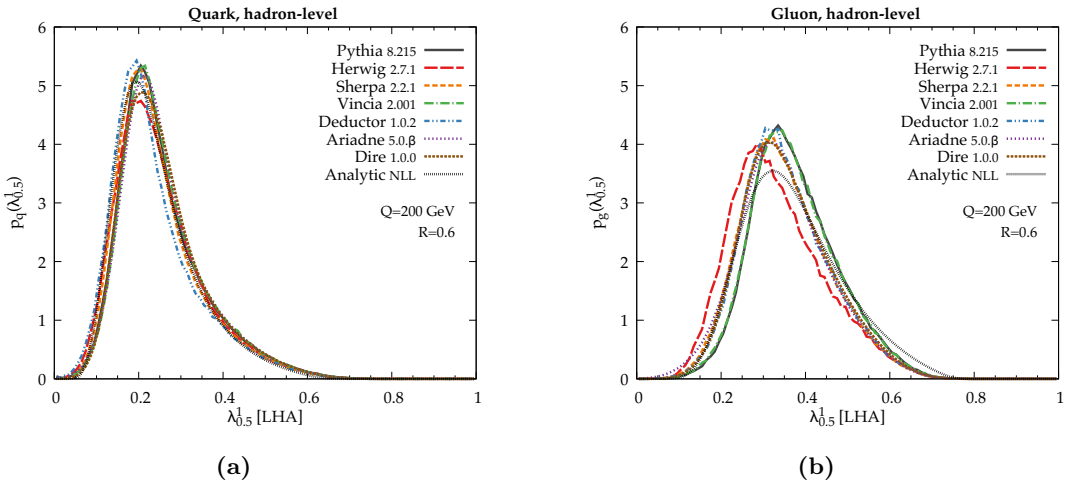
In figure 5, we show hadron-level distributions of the LHA (i.e. $e_{0.5} = \lambda^1_{0.5}$) in the quark sample ($p_q$) and gluon sample ($p_g$), comparing the baseline settings of seven different parton-shower generators with a center-of-mass collision energy of $Q = 200\,\text{GeV}$ and jet radius $R = 0.6$. In the quark sample in figure 5a, there is relatively little variation between the generators, which is not surprising since most of these programs have been tuned to match LEP data (though LEP never measured the LHA itself). Turning to the gluon sample in figure 5b, we see somewhat larger variations between the generators; this is expected since there is no data to directly constrain $e^+e^- \to gg$ (though there are indirect tests from LEP; see section 7). It is satisfying that for both the quark and gluon samples, the analytic NLL results from section 4 peak at roughly the same locations as the parton showers. In the source files attached as supplementary material to this paper, one can see comparable levels of agreement for the two other IRC-safe angularities (e_1 and e_2).

In figure 5c, we plot the integrand of classifier separation, $\mathrm{d}\Delta/\mathrm{d}\lambda$ from eq. (3.18). This shows where in the LHA phase space the actual discrimination power lies, with large values

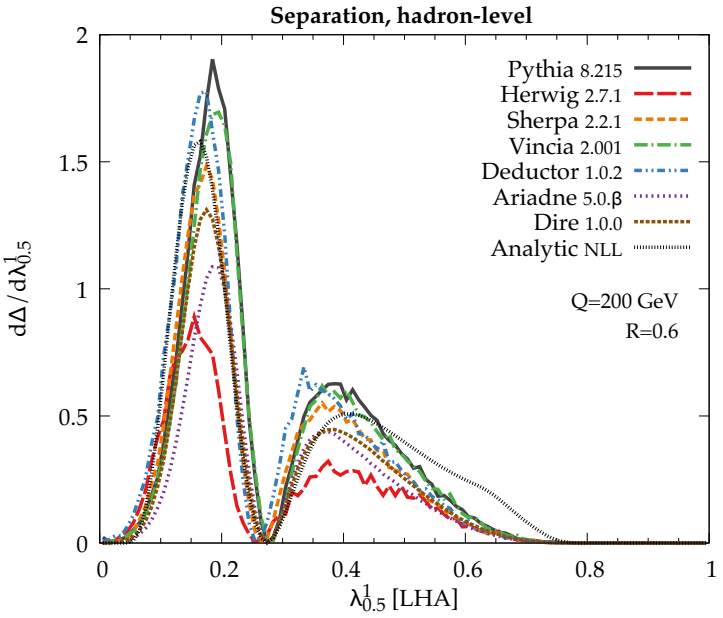
[12] In an e^+e^- context, our definition at least respects the Lorentz structure of the production vertex, so in that sense it is a fundamental definition that does not reference (ambiguous) quark or gluon partons directly.

[13] Note that we have not included the effect of eq. (5.1) in our analytic calculation, which in principle affects the functional form of f_{NGL} for non-global logarithms.

JHEP07(2017)091



(a) (b)



(c)

Figure 5. Hadron-level distributions of the LHA for (a) the $e^+e^- \to u\bar{u}$ ("quark jet") sample, (b) the $e^+e^- \to gg$ ("gluon jet") sample, and (c) the classifier separation integrand in eq. (3.18). Seven parton-shower generators — Pythia 8.215, Herwig 2.7.1, Sherpa 2.2.1, Vincia 2.001, Deductor 1.0.2, Ariadne 5.0.β, and Dire 1.0.0 — are run at their baseline settings with center-of-mass energy $Q = 200\,\text{GeV}$ and jet radius $R = 0.6$. We also show the analytic NLL results from section 4.

of the integrand corresponding to places where the quark and gluon distributions are most dissimilar. Now we see considerable differences between the generators, reproducing the well-known fact that Pythia is more optimistic about quark/gluon separation compared to Herwig [20]. The predicted discrimination power from the other five generators and the NLL calculation are intermediate between these extremes.

JHEP07(2017)091

One might expect that the differences between generators are due simply to their having different hadronization models. It seems, however, that the differences already appear at the parton level prior to hadronization. We should say at the outset that it is nearly impossible to do a true apples-to-apples comparison of parton-level results, since these generators are interfaced to different hadronization models, and only the hadron-level comparison is physically meaningful. In particular, the crossover between the perturbative and nonperturbative regions is ambiguous and each of these showers has a different effective shower cutoff scale, resulting in different amounts of radiation being generated in the showering versus hadronization steps.[14] Similarly, for the parton-level NLL results, small values of the angularities are artificially suppressed by the $\alpha_s \to \infty$ Landau pole, which enhances the Sudakov exponent.[15]

With that caveat in mind, we show parton-level results in figure 6. One immediately notices that three of the generators — Herwig, Sherpa, and Deductor — yield a large population of events where the perturbative shower generates no emissions, even in the gluon sample. This gives $\lambda^1_{0.5} = 0$ such that non-zero values of the LHA are generated only by the hadronization model. By contrast, Pythia and Vincia give overall larger values of the LHA from the perturbative shower alone, with Ariadne and Dire yielding intermediate results. As mentioned above, some of this difference can be explained simply by the different shower cutoff scales used in each generator, but it probably also reflects a difference in how semi-perturbative gluon splittings are treated. Since figure 5a shows that all generators give similar distributions for quark jets after hadronization, we conclude that understanding quark/gluon discrimination is a challenge at the interface between perturbative showering and nonperturbative hadronization.

To summarize the overall discrimination power, we integrate eq. (3.18) to obtain the value of classifier separation Δ for the LHA. This is shown in figure 7, which also includes the four other benchmark angularities from eq. (3.6). There is a rather large spread in predicted discrimination power between the generators, especially at hadron level in figure 7a. While such differences might be expected for IRC-unsafe angularities (multiplicity and p_T^D) which depend on nonperturbative modeling, these differences persist even for the IRC-safe angularities at parton level (see figure 7b).[16] This suggests a more fundamental difference between the generators that is already present in the perturbative shower.

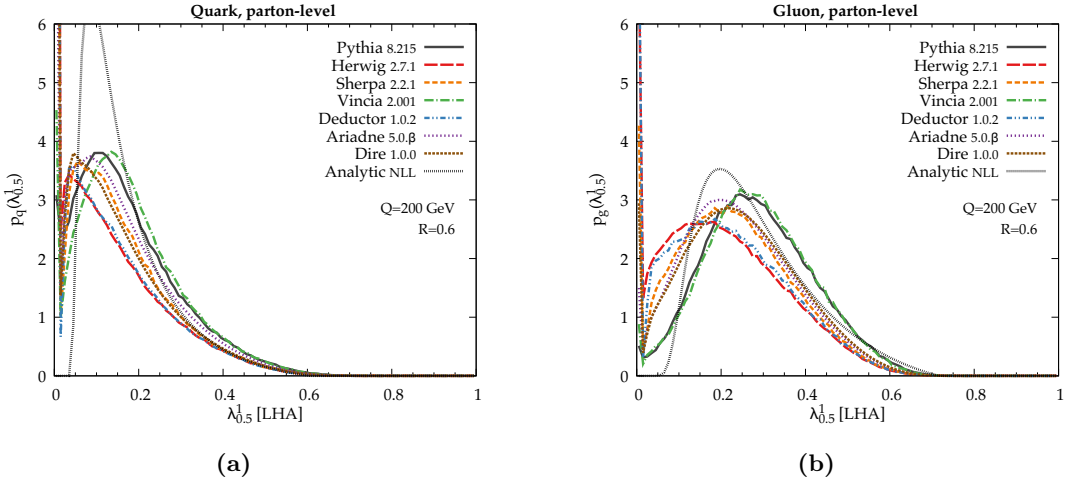
For the IRC-safe angularities with $\kappa = 1$, there is a generic trend seen by all of the hadron-level generators that discrimination power decreases as β increases. This trend agrees with the study performed in ref. [13] and our NLL calculation here, but disagrees with the ATLAS study in ref. [20], which found flat (or even increasing) discrimination power with increasing β. Understanding this β trend will therefore be crucial for understanding quark/gluon radiation patterns.

[14] In general, generators based on string hadronization tend to use a lower shower cutoff scale ($\sim 0.5\,\text{GeV}$) compared to those based on cluster hadronization ($\sim 1\,\text{GeV}$).

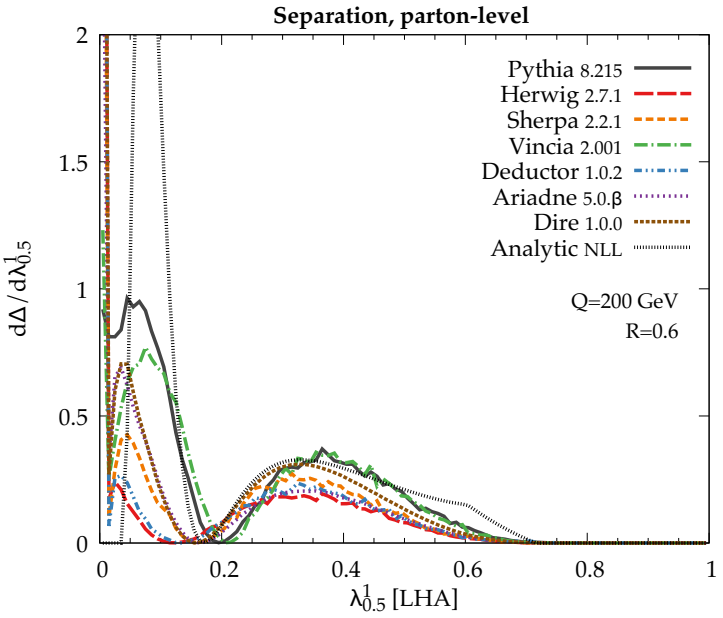
[15] An alternative approach would be to freeze α_s at scales below Λ_{QCD} or extend it into the nonperturbative region as suggested in refs. [101–103]. Either way, this region of phase space is dominated by the nonperturbative shape function, which is absent from the "parton-level" distributions.

[16] It is interesting that four of the generators — Herwig, Sherpa, Deductor, and Ariadne — have a comparatively narrow spread in predicted discrimination power at parton level, though this spread increases dramatically at hadron level.

JHEP07(2017)091



(a) (b)



(c)

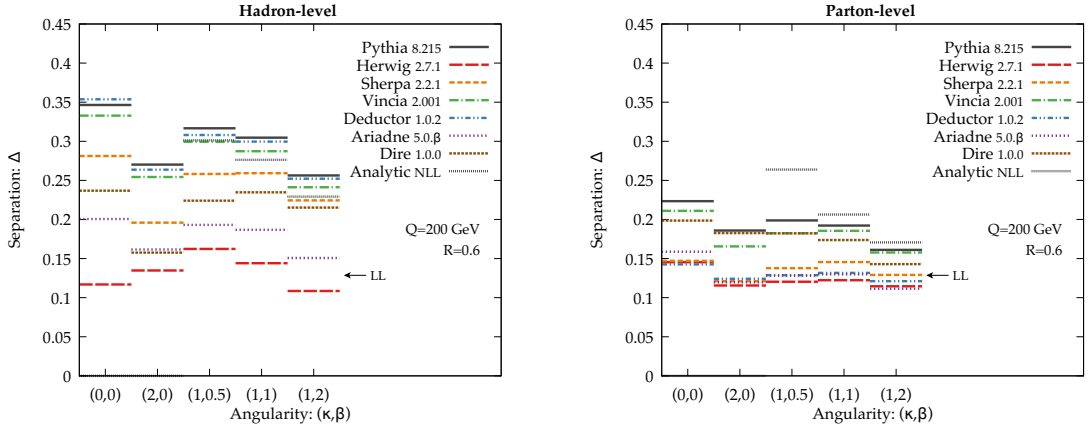
Figure 6. Same as figure 5, but at the parton level. Note that HERWIG, SHERPA, and DEDUCTOR all have cross section spikes at $\lambda^1_{0.5} = 0$ that extend above the plotted range.

5.2 Parameter dependence

Given the large absolute differences in discrimination power seen above, we next want to check if the parton-shower generators exhibit similar or dissimilar trends as parameters are varied. We perform three parameter sweeps, using the boldface values below as defaults:

$$\begin{aligned} \text{Collision Energy}: &\quad Q = \{50, 100, \mathbf{200}, 400, 800\}\,\text{GeV}, \\ \text{Jet Radius}: &\quad R = \{0.2, 0.4, \mathbf{0.6}, 0.8, 1.0\}, \\ \text{Strong Coupling}: &\quad \alpha_s/\alpha_{s0} = \{0.8, 0.9, \mathbf{1.0}, 1.1, 1.2\}, \end{aligned} \tag{5.2}$$

JHEP07(2017)091



(a) **(b)**

Figure 7. Classifier separation Δ for the five benchmark angularities in eq. (3.6), determined from the various generators at (a) hadron level and (b) parton level. The first two columns correspond to IRC-unsafe distributions (multiplicity and p_T^D), while the last three columns are the IRC-safe angularities. The LHA (i.e. $\kappa = 1$, $\beta = 1/2$) is shown in the middle column. Results in terms of ROC values appear in the source files as supplementary material of this paper, for this and subsequent plots. The label "LL" indicates the value from eq. (4.9) predicted by Casimir scaling.

where α_{s0} is the default value of the strong coupling, which is different between the generators (and sometimes different between different aspects of the same generator).

The resulting values of Δ for the LHA are shown in figure 8, at both the hadron level and parton level. There are number of surprising features in these plots. Perhaps the most obvious (and seen already in figure 7) is that even for the IRC-safe angularities, the effect of hadronization is rather large, both on the absolute scale of discrimination and the trends. The main exception to this is Herwig, which does not exhibit as much of a shift from hadronization, though an effect is still present.

The next surprising feature is that the parton-level trends for sweeping α_s do not necessarily correspond to those for sweeping Q and R. According to the perturbative NLL logic in ref. [13], quark/gluon discrimination should depend on α_s evaluated at the scale $QR/2$, with larger values of $\alpha_s(QR/2)$ leading to improved discrimination power. This is indeed seen in the parton-level curves obtained from the analytic NLL calculation in section 4, and parton-level Pythia, Herwig, Vincia, Ariadne, and Dire also show improved performance with larger α_s. However, larger values of Q and R correspond to smaller values of α_s, so the NLL logic would predict that increasing Q or R should lead to worse discrimination power. Instead, at parton-level, all of the generators show the opposite Q and R trend from the analytic NLL result.

One reason to expect quark/gluon discrimination to improve at higher energies is that the phase space available for shower evolution increases as Q increases. The scale μ of the shower splitting is $\mu_0^2 < \mu^2 < Q^2$, where $\mu_0 = \mathcal{O}(\text{GeV})$ is the shower cutoff scale. With more range for shower evolution at higher Q, there is a greater possibility to see that a quark jet is different from a gluon jet. Similarly, larger values of R allow

JHEP07(2017)091

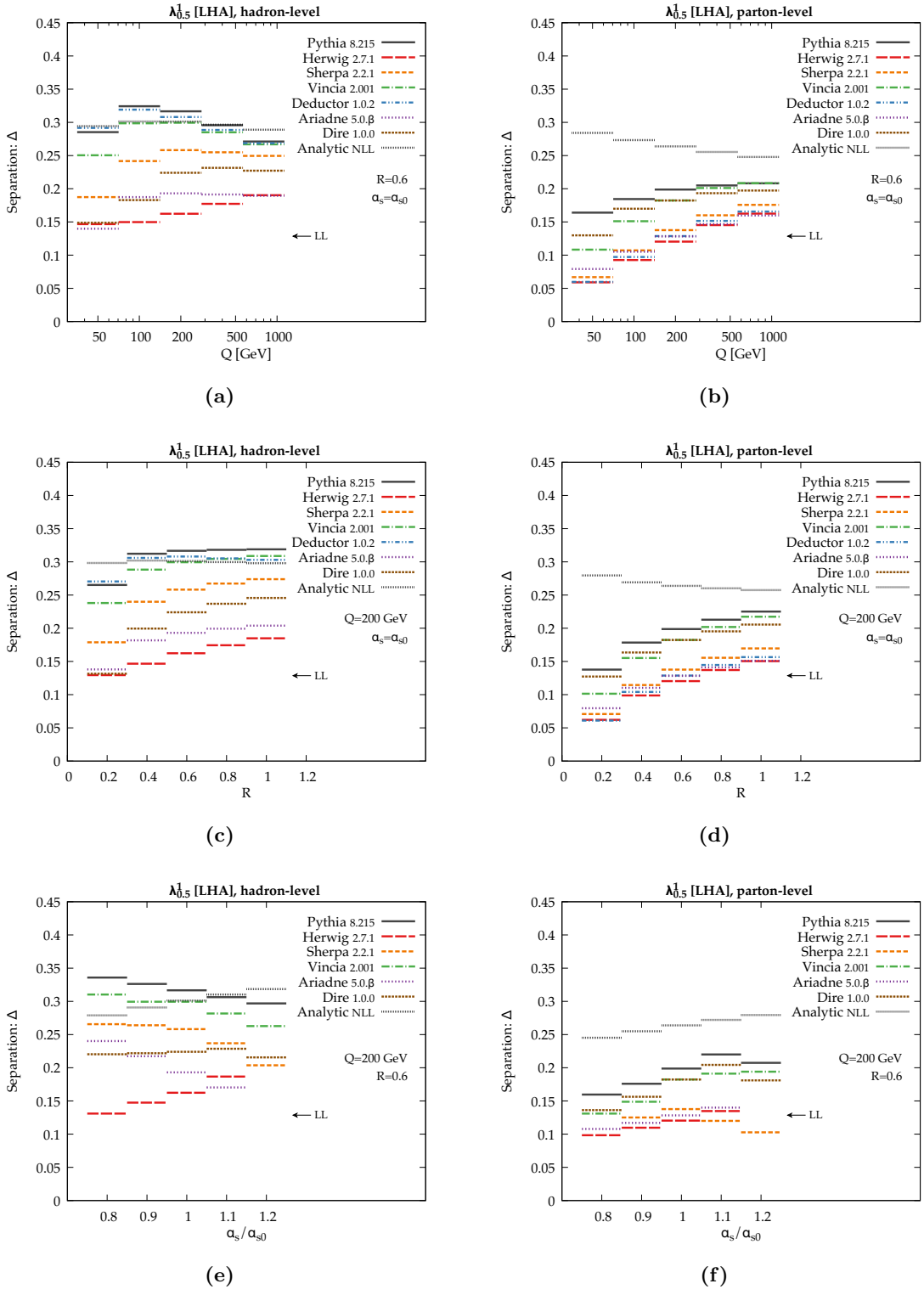


Figure 8. Classifier separation Δ for the LHA, sweeping the collision energy Q (top row), jet radius R (middle row), and coupling constant α_s/α_{s0} (bottom row). Results are shown at hadron level (left column) and parton level (right column).

JHEP07(2017)091

for more emissions within a jet, and from scaling symmetry, one expects that parton-level discrimination power should depend on the combination QR.[17] By contrast, the NLL logic says that quark/gluon discrimination should be dominated by the leading emission(s) in a jet, and since α_s is smaller at higher values of QR, those leading emissions are more similar between quarks and gluons. Given these two different but equally plausible logics, both of which undoubtably play some role in the complete story, this motivates experimental tests of quark/gluon separation as a function of Q and R.

For many of the generators, going from parton-level to hadron-level reverses or flattens the Q and α_s trends, though the R trends are more stable. For the NLL results, including the shape function from section 4.3 leads to an overall improvement in discrimination power and a slight flattening of the Q and R trends, though the difference between parton-level and hadron-level is not nearly as dramatic as for the parton showers. This is further evidence that the boundary between perturbative and nonperturbative physics is ambiguous, and hadron-level comparisons are the most meaningful.

5.3 Impact of generator settings

Formally, parton-shower generators are only accurate to modified leading-logarithmic (MLL) accuracy, though they include physically important effects like energy/momentum conservation and matrix element corrections that go beyond MLL. We can assess the impact of these higher-order effects by changing the baseline parameter settings in each parton-shower generator. We will also explore similar kinds of changes for the NLL analytic calculation.

Because each generator is different, we cannot always make the same changes for each generator. Similarly, the spread in discrimination power shown below should *not* be seen as representing the intrinsic uncertainties in the shower, since many of these changes we explore are not physically plausible. The goal of these plots is to demonstrate possible areas where small parameter changes could have a large impact on quark/gluon discrimination. Ultimately, collider data and higher-order calculations will be essential for understanding the origin of quark/gluon differences. In all cases, we show both hadron-level and parton-level results, even if a setting is only expected to have an impact at the hadron level.

Our Pythia baseline is based on the Monash 2013 tune, with parameters described in ref. [105]. In figure 9, we consider the following Pythia variations:

- Pythia: no $g \to q\bar{q}$. While the dominant gluon splitting in the parton shower is $g \to gg$, Pythia — and every other shower in this study — also generates the subleading $g \to q\bar{q}$ splittings by default. This variation turns off $g \to q\bar{q}$, which makes gluon jets look more gluon-like, thereby increasing the separation power.

- Pythia: no ME. The first emission in Pythia is improved by applying a matrix element correction [106], but this variation turns those corrections off, showing the impact of non-singular terms. No matrix element correction is available for $h^* \to gg$,

[17] At small values of R, one has to worry about the flavor purity of a jet sample, since scale evolution can change the leading parton flavor [36, 104]. Similarly, the restriction in eq. (5.1) can impose a non-trivial bias on the jet flavor at small R.

JHEP07(2017)091

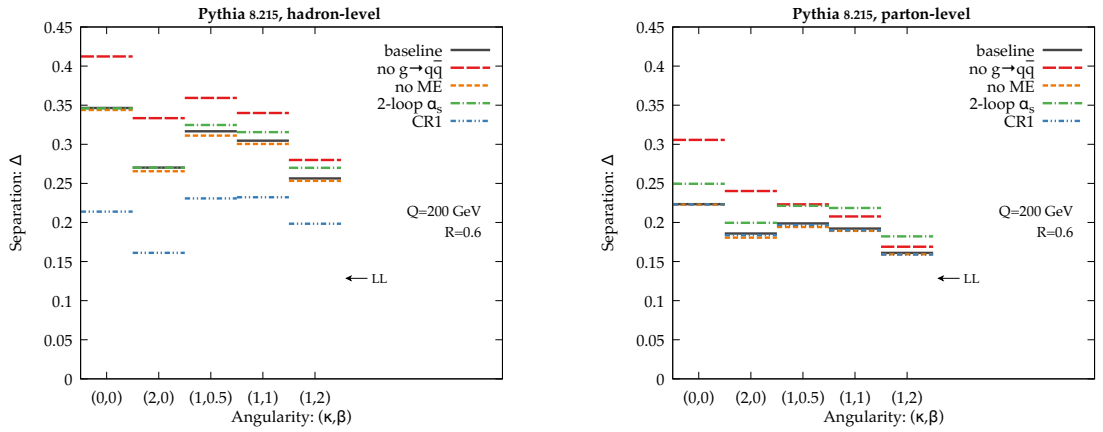


Figure 9. Settings variations for PYTHIA 8.215. Shown are (a) hadron-level and (b) parton-level results for the classifier separation Δ derived from the five benchmark angularities.

though, so the true impact of these corrections might be larger than the relatively small effect seen for this variation.

- PYTHIA: 2-LOOP α_s. The default PYTHIA setting is to use 1-loop running for α_s. This variation turns on 2-loop running for α_s, which has a small (beneficial) effect at parton level which is washed out at hadron level.

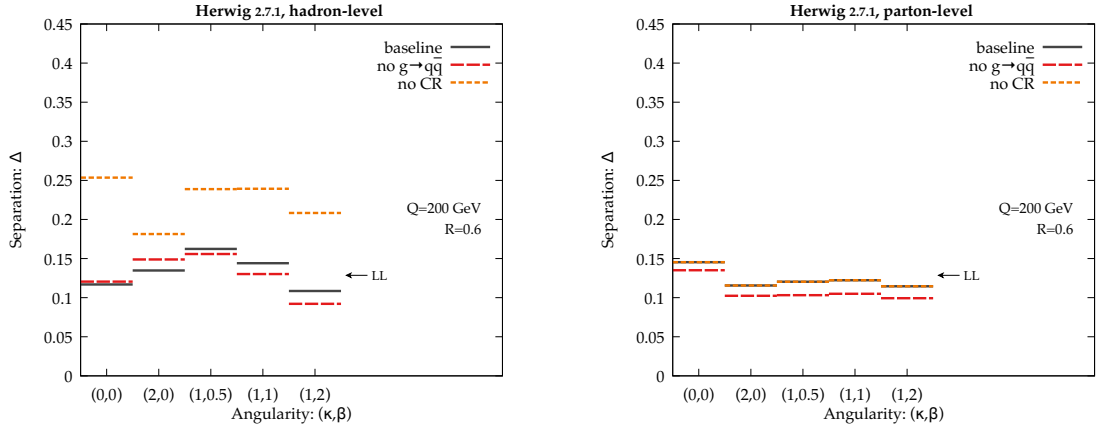
- PYTHIA: CR1. Often, one thinks of color reconnection as being primarily important for hadron colliders, but even at a lepton collider, color reconnection will change the Lund strings used for hadronization. Compared to the baseline, this variation uses an alternative "SU(3)"-based color reconnection model [107] (i.e. `ColourReconnection:mode = 1`). No attempts were made to retune any of the other hadronization parameters (as would normally be mandated in a tuning context), so this change simply illustrates the effect of switching on this reconnection model with default parameters, leaving all other parameters unchanged. At parton level, this variation has no effect as expected. At hadron level, this variation considerably degrades quark/gluon separation compared to the baseline.

The most surprising PYTHIA effect is the large potential impact of the color reconnection model, which is also important for the HERWIG generator described next.

Our HERWIG baseline uses version 2.7.1, with improved modeling of underlying event [108] and the most recent UE-EE-5-MRST tune [109], which is able to describe the double-parton scattering cross section [110] and underlying event data from $\sqrt{s} = 300\,\text{GeV}$ to $\sqrt{s} = 7\,\text{TeV}$. In figure 10, we consider the following HERWIG variations:

- HERWIG: NO $g \to q\bar{q}$. Turning off $g \to q\bar{q}$ splittings in HERWIG has the reverse behavior as seen in PYTHIA, leading to slightly worse discrimination power, though the effect is modest.

JHEP07(2017)091



(a) **(b)**

Figure 10. Same as figure 9, but for HERWIG 2.7.1.

- HERWIG: NO CR. The variation turns off color reconnections in HERWIG. This has no effect at parton level, as expected. At hadron level, this variation for HERWIG gives a rather dramatic improvement in quark/gluon discrimination power. We think this arises since color reconnection in HERWIG allows any color-anticolor pair to reconnect, even if they arose from an initially color octet configuration. By turning off color reconnection, the gluons look more octet-like, explaining the improvement seen.

The importance of color reconnections in HERWIG is a big surprise from this study, motivating future detailed studies into which color reconnection models are most realistic when compared to data. In the future, we also plan to test the default angular-ordered HERWIG shower against an alternative dipole shower [111].

Our SHERPA baseline uses matrix element corrections for the first two emissions ($N_{\text{jet}} = 2$) with CKKW-style matching [112]. In figure 11, we consider the following SHERPA variations:

- SHERPA: NO $g \to q\bar{q}$. Turning off $g \to q\bar{q}$ splittings in SHERPA has a negligible effect at parton level, but it leads to a large jump in discrimination power at hadron level, again due to an interplay between the perturbative shower and nonperturbative hadronization.

- SHERPA: $N_{\text{jet}} = 1$. This variation only performs CKKW matching for the first emission, leading to negligible changes in the discrimination performance.

- SHERPA: $N_{\text{jet}} = 0$. Turning off all matrix element corrections in SHERPA slightly decreases the predicted quark/gluon discrimination power, in agreement with the behavior of PYTHIA.

Within SHERPA, matrix element corrections appear to have a very small effect at parton level. The large changes seen at hadron level from turning off $g \to q\bar{q}$ splittings motivates further investigations into the shower/hadronization interface.

JHEP07(2017)091

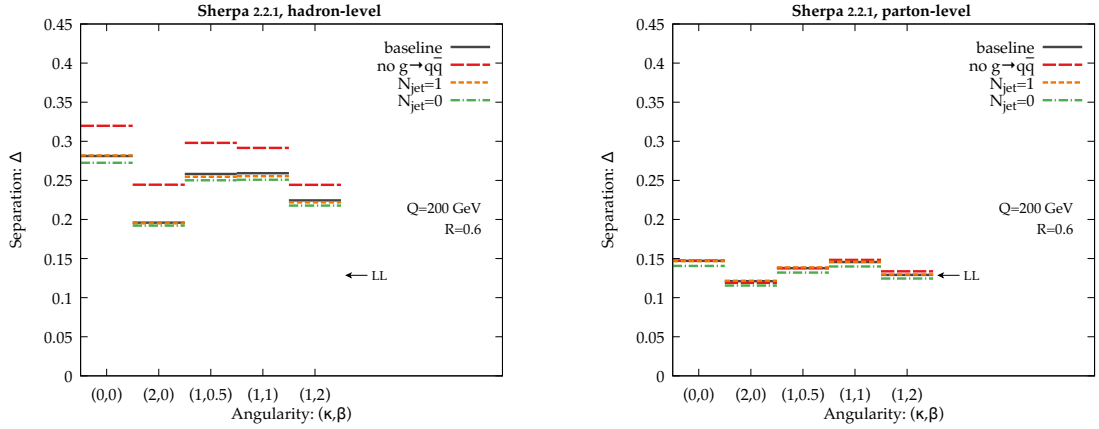


Figure 11. Same as figure 9, but for SHERPA 2.2.1.

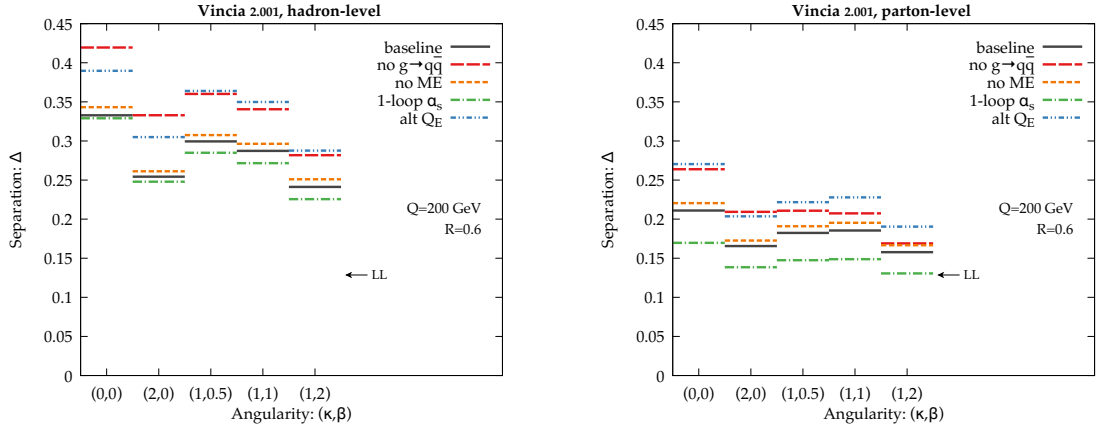


Figure 12. Same as figure 9, but for VINCIA 2.001.

Our VINCIA baseline uses the default setup for version 2.001 [44], which includes "smooth ordering" and LO matrix-element corrections [113] up to $\mathcal{O}(\alpha_s^3)$ for both $e^+e^- \to q\bar{q}$ and $e^+e^- \to gg$. The coupling α_s is evaluated with 2-loop running defined by $\alpha_s(M_Z) = 0.118$ (reinterpreted according to the CMW scheme [114]) with $\mu_R = 0.6p_\perp$ as the renormalization scale for gluon emissions and $\mu_R = 0.5m_{q\bar{q}}$ for $g \to q\bar{q}$ branchings. In figure 12, we consider the following VINCIA variations:

- VINCIA: NO $g \to q\bar{q}$. This variation turns off $g \to q\bar{q}$, leading to the expected increase in separation power as seen in PYTHIA.
- VINCIA: NO ME. By default, each $2 \to 3$ antenna in VINCIA has an associated matrix element correction factor. Since the antennae are already rather close to the true matrix elements, turning off these matrix elements has a modest effect on quark/gluon discrimination power.

JHEP07(2017)091

- VINCIA: 1-LOOP α_s. This variation switches from 2-loop to 1-loop α_s running, yielding a parton-level difference which goes in the same direction as the equivalent PYTHIA variation (note the baseline in PYTHIA is 1-loop running) and a modest hadron-level difference, again in agreement with the observation for PYTHIA.

- VINCIA: ALT Q_E. By default, VINCIA uses a transverse-momentum scale (the same as in ARIADNE) as the evolution variable for gluon emissions. This variation instead uses a virtuality-like quantity. This changes the Sudakov factors to slightly enhance wide-angle emissions over collinear ones (see e.g. [115]). The resulting increase in separation power is mainly due to increased activity in the $H \to gg$ shower.

Since VINCIA and PYTHIA share the same hadronisation model and both have dipole-style showers, it is not surprising that they exhibit similar behaviors as parameters are changed. The biggest surprise is the significant change observed when using an alternative shower evolution variable ("ALT Q_E"), which persists at hadron level. Although this variation is theoretically disfavored (the default $p_\perp$ evolution variable has been shown to reproduce the logarithmic structure of the $q\bar{q} \to qg\bar{q}$ antenna function to second order in α_s [116]), formal control of the ambiguity would depend on one-loop corrections. It would therefore be interesting to determine the extent to which multi-leg NLO merging techniques (such as UNLOPS [117]) would reduce it, and/or whether second-order corrections to the shower kernels are required (for which only a proof of concept currently exists [118]).

Our DEDUCTOR baseline uses leading color plus (LC+) showering, which includes some subleading color structures. We find that switching from LC+ to LC showering at parton level has a negligible impact on quark/gluon discrimination power. When DEDUCTOR interfaces with the default tune of PYTHIA 8.212 for hadronization, only leading color is used in the showering, such that partons with their LC color information can be directly passed to the Lund string model. No DEDUCTOR variations are shown here, though it would be interesting to study the effect of $g \to q\bar{q}$ splitting in future work.

Our ARIADNE baseline is based on a beta release of version 5. In figure 13, we consider the following ARIADNE variation:

- ARIADNE: NO $g \to q\bar{q}$. This variation turns off $g \to q\bar{q}$, leading to modest improvement in separation power, similar in magnitude to HERWIG though in the opposite direction.

- ARIADNE: NO SWING. Swing refers to color reconnections performed during the perturbative cascade, where dipoles in the same color state are allowed to reconnect in a way which prefers low-mass dipoles [46, 119]. Turning off swing has an effect already at parton level, which is amplified at hadron level, leading to improved quark/gluon separation.

Like for PYTHIA and HERWIG, color reconnections play a surprisingly important role in ARIADNE.

Our DIRE baseline is based on the initial release, interfaced with SHERPA for cluster hadronization. In figure 14, we consider the following DIRE variations:

JHEP07(2017)091

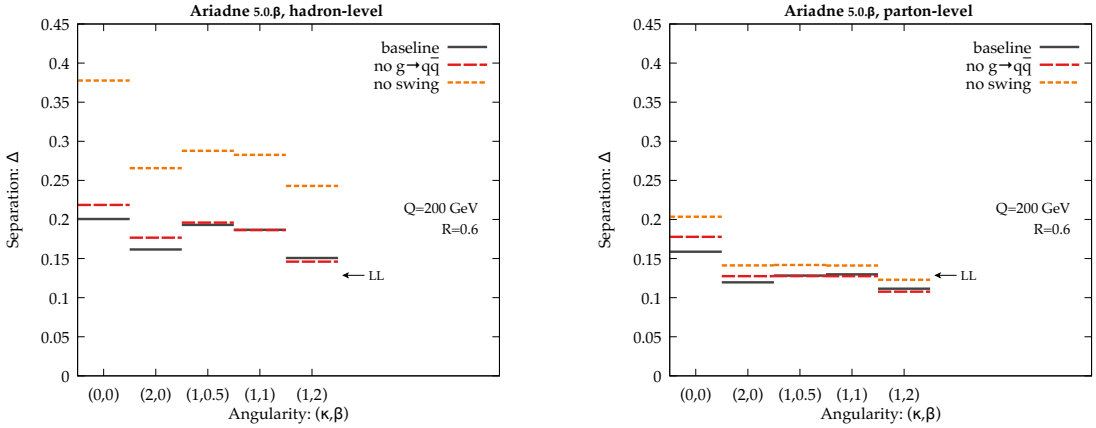


Figure 13. Same as figure 9, but for ARIADNE 5.0.β.

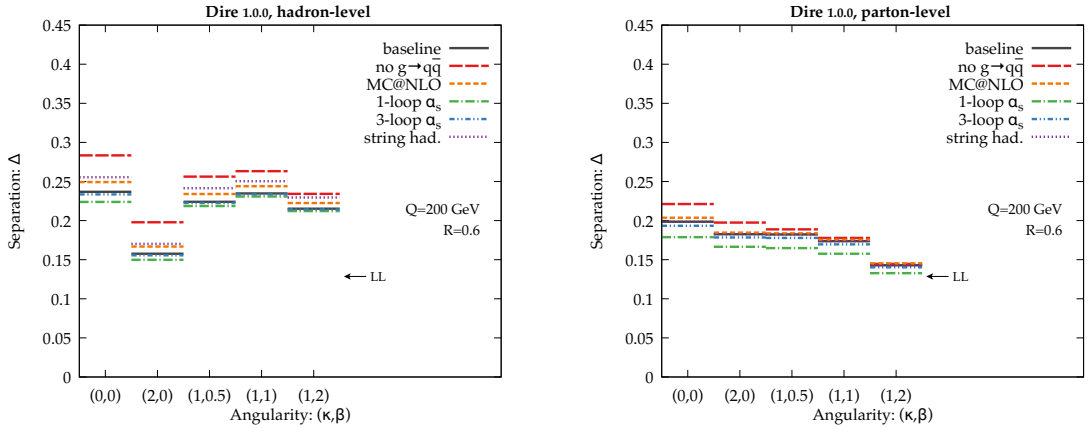
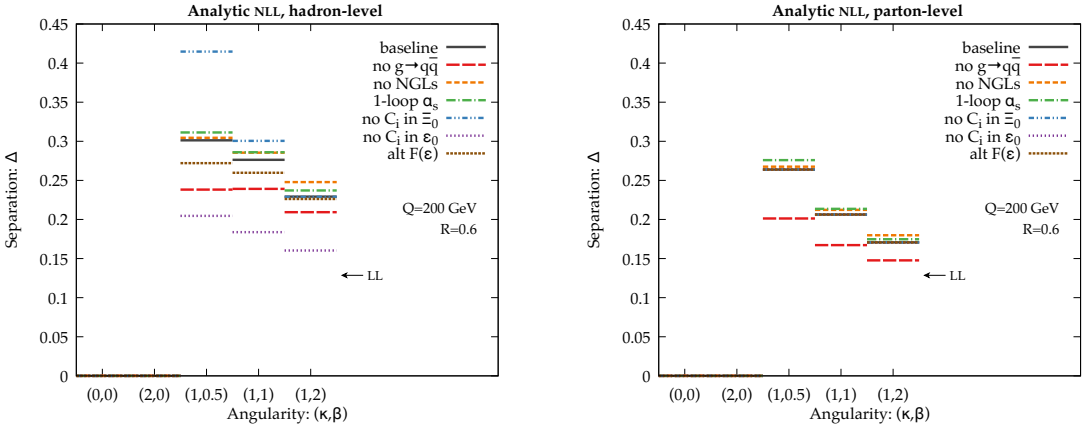


Figure 14. Same as figure 9, but for DIRE 1.0.0.

- DIRE: NO $g \to q\bar{q}$. This variation turns off $g \to q\bar{q}$, yielding an improvement in separation power at both the parton level and hadron level, intermediate between ARIADNE and VINCIA.
- DIRE: MC@NLO. This variation uses MC@NLO [120] as implemented in SHERPA to provide a matrix element correction. The discrimination power slightly improves at both parton and hadron level, though not that much, since the DIRE shower already is very close to capturing the matrix element for the first emission.
- DIRE: 1-LOOP α_s. The default within DIRE is to perform 2-loop α_s running. This variation uses just 1-loop running, with a slight degradation of discrimination power.
- DIRE: 3-LOOP α_s. Using 3-loop running also degrades performance, but by a very small amount.

JHEP07(2017)091



(a) **(b)**

Figure 15. Same as figure 9, but for the analytic NLL calculation from section 4.

- DIRE: STRING HAD. This variation uses PYTHIA for Lund string fragmentation, which only has an effect at hadron level. This leads to a modest improvement in discrimination power, suggesting that long-range color connections can play an important role in quark/gluon discrimination. Note that the shower cutoff scale is the same for cluster and string fragmentation in DIRE.

Of the generators we tested, DIRE is the only one that interfaces with two different hadronization routines, motivating further studies into the differences between cluster and string fragmentation.

Finally in figure 15, we consider the analytic NLL calculation from section 4. Here, we can only study the IRC-safe angularities with $\kappa = 1$.

- ANALYTIC NLL: NO $g \to q\bar{q}$. To turn off gluon splitting to quarks, we set $n_f = 0$ in eq. (4.14), without adjusting the running of α_s. This effectively decreases the number of emissions from gluons, making them look more quark-like. The resulting decrease in discrimination power is the opposite of the behavior seen in the parton-shower generators (except HERWIG), suggesting that at higher perturbative orders, the effect of $g \to q\bar{q}$ will go beyond just changing the reduced splitting functions.

- ANALYTIC NLL: NO NGLS. Here, we set $f_{\text{NGL}} = 0$ in eq. (4.10). Since nonglobal logarithms obey Casimir scaling in the $N_c \to \infty$ limit, this is expected to have a mild impact on quark/gluon separation power, which is indeed the case.

- ANALYTIC NLL: 1-LOOP α_s. The analytic NLL calculation uses 2-loop α_s running by default. This option uses only 1-loop running, which has a relatively small (beneficial) impact.

- ANALYTIC NLL: NO C_i IN Ξ_0. The default choice for the average nonperturbative shift ϵ_0 assumes Casimir scaling as in eq. (4.26). This option uses instead the shift

JHEP07(2017)091

in eq. (4.27), which only has Casimir scaling for Ω_0 and not for Ξ_0. This makes a dramatic impact for $\beta < 1$ at hadron-level, since Ξ_0 dominantly controls the impact of nonperturbative collinear emissions. Specifically, the default ϵ_0 scales like C_i^β for $\beta \ll 1$ whereas this option has linear scaling with C_i, leading to increased discrimination power.

- Analytic NLL: no C_i in ϵ_0. This option uses the nonperturbative shift in eq. (4.28), which is the same for quarks and gluons. As expected, this reduces the difference between quark and gluon jets at hadron-level, leading to a large reduction in discrimination power.

- Analytic NLL: alt $F(\epsilon)$. Here, we change the functional form of the shape function in eq. (4.18) from $F(\epsilon)$ to $F_{\text{alt}}(\epsilon)$, keeping the same value of ϵ_0. Since F_{alt} has a larger high-side tail, there is more overlap of the quark and gluon distributions, reducing somewhat the discrimination power.

The key lesson from this analytic study is that the form of the nonperturbative shape function has a large effect on quark/gluon discrimination power, especially the assumed dependence of ϵ_0 on the Casimir factor. So while higher-order perturbative calculations of quark/gluon radiation patterns are essential, quantitative control over nonperturbative physics will be required to make robust statements about the predicted discrimination power.

6 Quark/gluon tagging at the LHC

It is clear from our e^+e^- study that quark/gluon radiation patterns face considerable theoretical uncertainties, as seen from the differing behaviors of parton-shower generators and from the importance of the shape function in the analytic calculation. This is true even accounting only for final-state physics effects, so additional initial-state complications can only increase the uncertainties faced in pp collisions at the LHC. Beyond just the application to quark/gluon tagging, this is an important challenge for any analysis that uses jets. For example, a proper experimental determination of jet energy scale corrections requires robust parton-shower tools that correctly model effects like out-of-cone radiation.

Eventually, one would like to perform improved analytic calculations to address these radiation pattern uncertainties. In the near term, though, measurements from the LHC will be essential for improving the parton-shower modeling of jets. In this section, we perform an example LHC analysis that highlights the kind of information one can gain about quark/gluon radiation patterns, despite the additional complications faced by hadronic collisions.

6.1 Defining enriched samples

As discussed in section 2, there is no way to isolate pure samples of quark or gluon jets at the LHC, but one can isolate quark/gluon-enriched samples, as defined by the flavor label of the jet in the corresponding Born-level partonic process. As shown in figure 16, the

JHEP07(2017)091

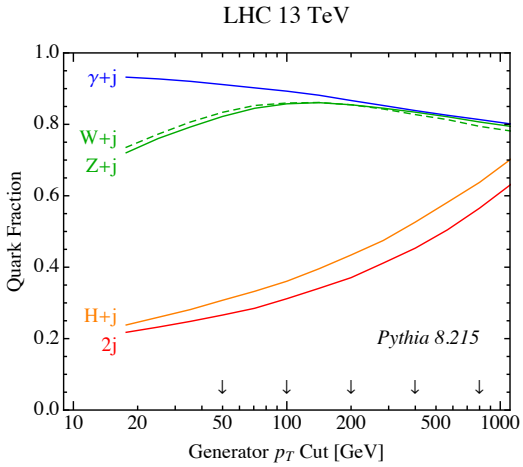


Figure 16. Quark fraction of jets at parton level, as defined by the Born-level parton flavor. The arrows indicate the $p_T^{\min}$ values used for the study in section 6.3.

Born-level jet in $W/Z/\gamma+$jet is more than 70% quark enriched over the entire jet p_T range of interest. For jets softer than around 200 GeV, the Born-level jet in dijets or $H+$jet is more than 60% gluon enriched, with that fraction decreasing as the jet p_T increases. More sophisticated enrichment procedures are described in ref. [54].

In principle, one could try to "diagonalize" some combination of vector boson plus jet and dijet samples in order to define separate quark or gluon samples (see e.g. [20]). In the spirit of section 2, though, we think it is more beneficial for the LHC experiments to perform process-specific measurements without trying to directly determine their quark and gluon composition. Instead of quark/gluon separation, here we ask the more well-defined question of whether one can tell "the jet in Z plus jets" (quark-enriched) apart from "the jet in dijets" (gluon-enriched).[18] In a similar spirit, one could test for differences within a single jet sample, such as comparing central jets versus forward jets in dijet production. This process-based strategy can help sidestep the known process dependance of defining quarks and gluons at the LHC, where color correlations have an important impact on observed jet radiation patterns.

For this study, we study proton-proton collisions at the 13 TeV LHC. We consider four different hadron-level generators — Pythia 8.215 [39], Herwig 2.7.1 [40, 41], Sherpa 2.2.1 [43], and Vincia 2.001 [44] — using $Z \to \mu^+\mu^-$ plus jets as our quark-enriched sample and dijets as our gluon-enriched sample. All of these generators are used with their default settings, including underlying event modeling and hadronization. We set $R = 0.4$ as the default jet radius, with jets defined by the anti-k_t algorithm, in keeping with current jet studies at the LHC, exploring other values in section 6.3. Hadrons with rapidity $|y| < 2.5$ are used for jet clustering, and the resulting jets are restricted to have $|y_{\text{jet}}| < 1.5$. We apply a minimum p_T cut with default value $p_T^{\min} = 100$ GeV, similar in spirit to the $Q/2$ value used in the e^+e^- study, though the precise meaning of $p_T^{\min}$ differs between the two samples.

[18] See, however, ref. [121] for a machine-learning approach to handle mixed quark/gluon samples.

JHEP07(2017)091

The specific analysis routines used for this pp study are available from https://github.com/gsoyez/lh2015-qg. For the Z plus jets analysis (with RIVET routine `MC_LHQG_Zjet.cc`), the selection criteria for the reconstructed Z boson and jet are:

$$pp \to Z + j \text{ (“quark-enriched”)}: \quad p_T^Z > p_T^{\min}, \quad \frac{p_T^{\text{jet}}}{p_T^Z} > 0.8, \qquad |y_{\text{jet}} - y_Z| < 1.0. \tag{6.1}$$

In addition, we apply a $p_T > 5\,\text{GeV}$ cut on each muon. For the dijet analysis (with RIVET routine `MC_LHQG_dijet.cc`), our selection is based on the two hardest jets (labeled 1 and 2), both of which are used for analysis if they satisfy:

$$pp \to 2j \text{ (“gluon-enriched”)}: \quad \frac{p_{T,1}+p_{T,2}}{2} > p_T^{\min}, \quad \frac{p_{T,2}}{p_{T,1}} > 0.8, \quad |y_1 - y_2| < 1.0. \tag{6.2}$$

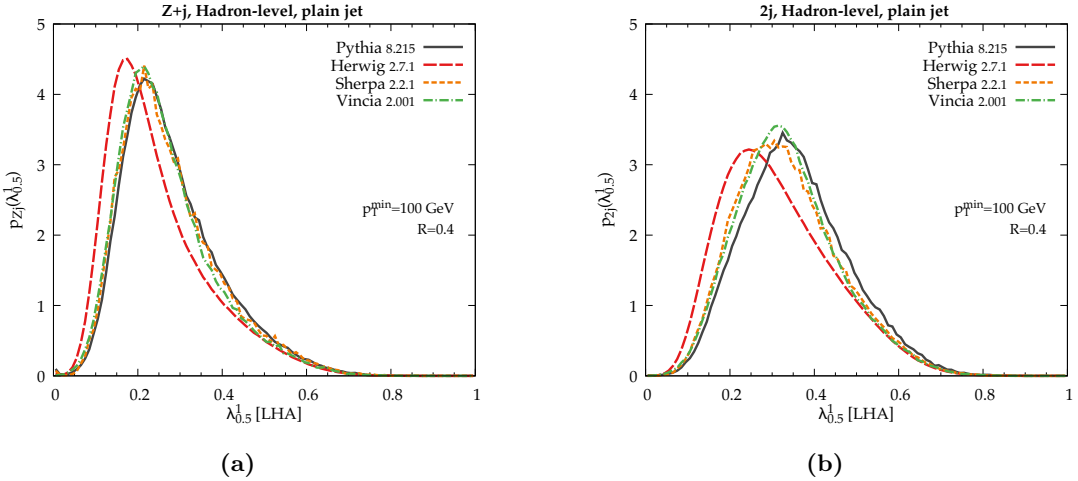
We study the same five benchmark angularities from eq. (3.6), but we also test the impact of soft radiation removal using mMDT grooming ($\mu = 1$ and $z_{\text{cut}} = 0.1$), with the grooming condition given in eq. (3.5). Prior to both the computation of λ_β^κ and the application of the mMDT procedure, the jet constituents are reclustered with the C/A algorithm, using the winner-take-all recombination scheme.

6.2 Baseline analysis

In figure 17, we show LHA distributions for the quark-enriched and gluon-enriched samples, using the default values $p_T^{\min} = 100\,\text{GeV}$ and $R = 0.4$. For the quark-enriched Z plus jets sample, all of the generators except HERWIG yield similar distributions, as expected from the e^+e^- study where the various generators broadly agreed on quark radiation patterns. For the gluon-enriched dijet sample, the difference between generators grows noticeably, yielding disagreements that are even larger than the e^+e^- study. This difference is apparent also in the $\mathrm{d}\Delta/\mathrm{d}\lambda$ distribution, where PYTHIA and VINCIA predict substantially larger separation power than HERWIG, again in agreement with the e^+e^- study. Results from SHERPA appear to be intermediate between these extremes, though the integrated Δ value turns out to be similar to HERWIG (see figure 19a). Already from these raw distributions, we see that LHC jet shape measurements would help constrain parton-shower uncertainties, especially for gluon-enriched jets.

We next turn to the impact of jet grooming. Often jet grooming is described as a strategy to mitigate jet contamination from pileup, underlying event, and initial-state radiation [76–79]. Even at the level of final-state radiation, though, grooming modifies the observed jet radiation patterns in ways that are interesting from the quark/gluon discrimination perspective [80, 81]. The impact of grooming is shown for the LHA after mMDT in figure 18. In general, grooming pushes jet shapes to smaller values, since the effect of grooming is to remove soft peripheral radiation from a jet. If the parton showers differed primarily in their treatment of wide-angle soft radiation, then one would expect grooming to bring the distributions into closer agreement. Instead, we see that the generator differences persist even after grooming, suggesting that the parton showers differ already in their treatment of collinear radiation, despite using the same underlying collinear splitting kernels. This motivates LHC measurements of groomed jet shapes to better understand the description of collinear physics.

JHEP07(2017)091



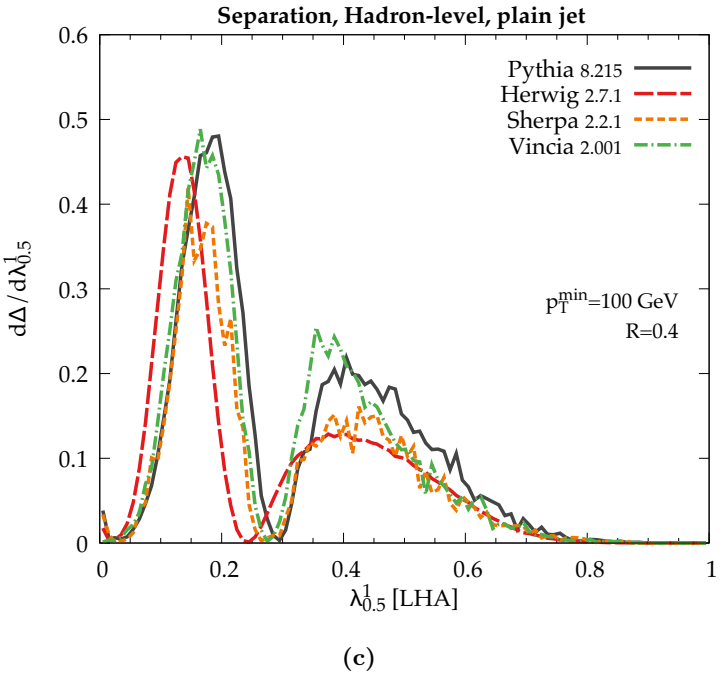
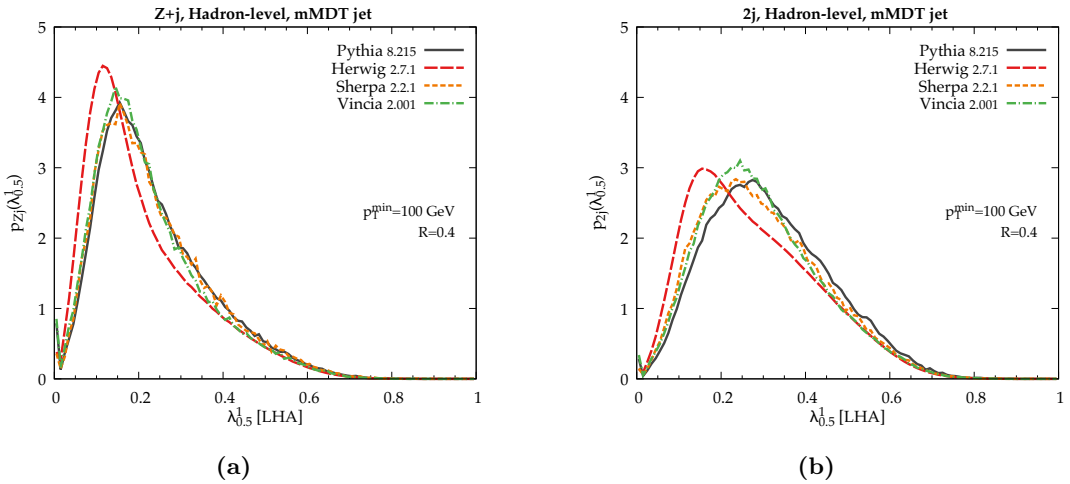


Figure 17. Distributions of the LHA at the LHC for (a) the $pp \to Z + j$ ("quark-enriched") sample, (b) the $pp \to 2j$ ("gluon-enriched") sample, and (c) the classifier separation integrand in eq. (3.18). Four parton-shower generators — Pythia 8.215, Herwig 2.7.1, Sherpa 2.2.1, and Vincia 2.001 — are run at their baseline settings with $p_T^{\min} = 100\,\text{GeV}$ and jet radius $R = 0.4$. Note that the plotted range is different from figure 5.

In figure 19, we plot the classifier separation Δ for all five benchmark angularities, with and without jet grooming. Compared to the e^+e^- study, the overall degree of discrimination power is reduced, as expected because the $Z + j$ and dijet processes do not yield pure quark/gluon samples. The spread between the generators is fairly large, with the expected trend that Pythia is more optimistic about quark/gluon separation than Her-

JHEP07(2017)091



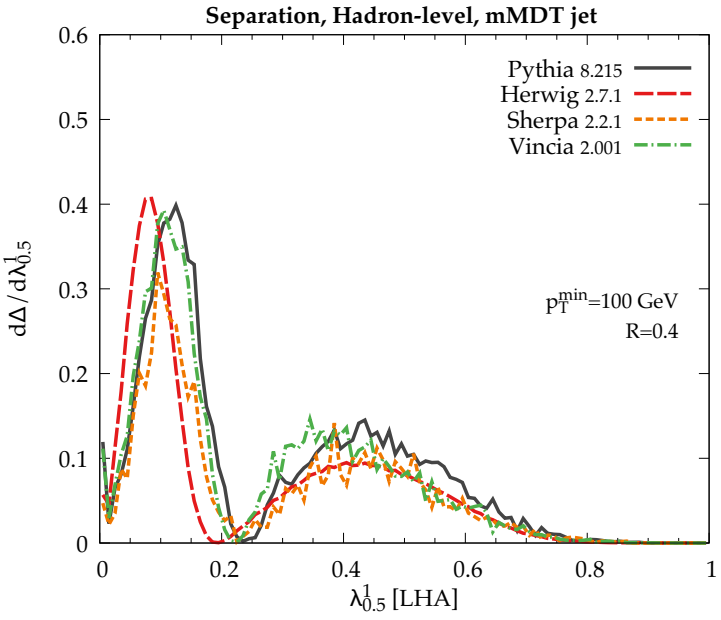
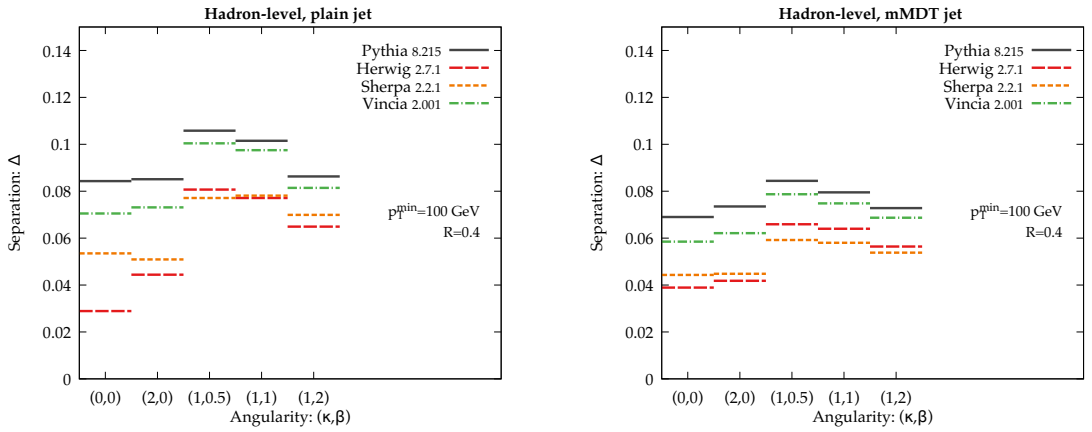


Figure 18. Same as figure 17 but after mMDT jet grooming with $\mu = 1$ and $z_{\text{cut}} = 0.1$.

wig. We see that Vincia has somewhat smaller predicted separation power than Pythia. Though their raw distributions differ, the discrimination power in Sherpa is comparable to Herwig, with the ordering roughly flipped for unsafe versus safe observables.

One surprising outcome of this study is the relatively modest impact of grooming on discrimination power. From the calculations in refs. [80, 81], one generically expects quark/gluon discrimination power to degrade after jet grooming, since the soft radiation that is being removed carries information about the color structure of the jet. This predicted degradation, however, is only seen modestly here, possibly because soft correlations with the initial state already blurred the distributions in the ungroomed case. One advan-

JHEP07(2017)091



(a) **(b)**

Figure 19. Classifier separation Δ for the five benchmark angularities in our LHC study, determined from the various generators (a) using all jet constituents and (b) after mMDT grooming. Note that the plotted range is different from figure 7, reflecting the decreased discrimination power in the realistic pp case compared to the idealized e^+e^- study with pure quark/gluon samples.

tage of working with groomed samples is that jet grooming reduces the process dependence in quark/gluon radiation patterns [57, 58]. In this way, groomed angularities should yield a more robust theoretical definition for quark and gluon jets, with only a small performance penalty.

6.3 Parameter dependence

To test parameter dependence, we now consider five different minimum p_T values and five different jet radii, with the boldface values corresponding to the defaults used above:[19]

$$\begin{aligned} \text{Minimum } p_T: &\quad p_T^{\min} = \{50, \mathbf{100}, 200, 400, 800\}\,\text{GeV}, \\ \text{Jet Radius}: &\quad R = \{0.2, \mathbf{0.4}, 0.6, 0.8, 1.0\}. \end{aligned} \tag{6.3}$$

These p_T values are effectively twice those used for the e^+e^- study in eq. (5.2) (where $E_{\text{jet}} \simeq Q/2$), and one should keep in mind from figure 16 that the degree of quark/gluon enrichment changes as a function of $p_T^{\min}$.

The results of sweeping $p_T^{\min}$ and R are shown in figure 20. In general, increasing $p_T^{\min}$ leads to a degradation of separation power, though this is due in large part to the change in sample composition shown in figure 16. In the context of these mixed samples, it is difficult to disentangle the impact of reduced quark/gluon enrichment at high p_T with actual trends in discrimination power (cf. figure 8a from the e^+e^- study). That said, the trends are sufficiently different between generators that the relative differences cannot be ascribed to sample composition alone.

[19] As a technical note, in order to test all values of $p_T^{\min}$ in a single Monte Carlo run, we generate p_T-weighted events in each generator.

JHEP07(2017)091

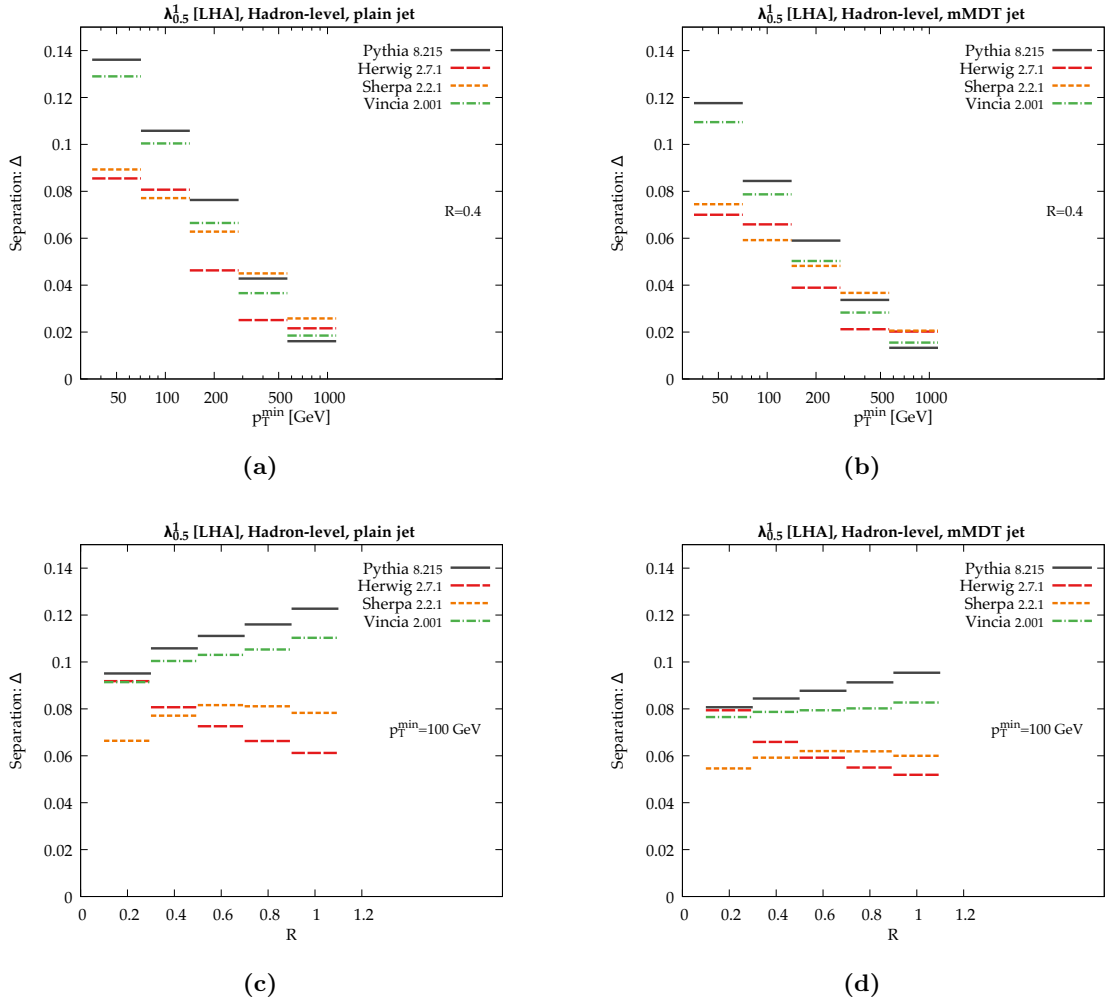


Figure 20. Classifier separation Δ for the LHA in our LHC study, sweeping $p_T^{\min}$ (top row) and jet radius R (bottom row). Results are shown with all jet constituents (left column) and after mMDT grooming (right column).

With respect to changing the jet radius R, the discrimination trends are noticeably different between the generators. In Pythia and Vincia, the discrimination power rises with increasing jet radius, whereas in Herwig, the discrimination power degrades with larger R; Sherpa has relatively little R dependence. The trends are rather similar before and after jet grooming, again pointing to differences between the generators in collinear physics and not just soft physics. We conclude that varying $p_T^{\min}$ and R provides important information about quark/gluon radiation patterns that cannot captured by focusing on a single kinematic regime. We therefore encourage LHC measurements of jet shapes at multiple energy scales with multiple jet radii.

JHEP07(2017)091

7 Summary and recommendations

By measuring the substructure of jets, one can gain valuable information about the relative quark/gluon composition of a jet sample. The challenge we have identified in this study is that the precise radiation patterns of quark and gluon jets are poorly understood, in the sense that parton-shower generators give rather different predictions for absolute quark/gluon discrimination power as well as relative trends as a function of the jet kinematics. From our analytic NLL studies including nonglobal logarithms and shape functions, we see that both perturbative and nonperturbative physics play an important role in determining jet shape distributions. That said, analytic calculations are not yet at the level of accuracy where they could directly guide the tuning of event generators. Therefore, LHC measurements are the best near-term strategy to constrain quark/gluon radiation patterns and enable quark/gluon discrimination to become a robust experimental tool.

Our five benchmark angularities probe both the perturbative and nonperturbative structure of jets, so we think they would be a good starting point for a more comprehensive quark/gluon jet shape analysis at the LHC. In this spirit, we are encouraged by the track multiplicity study of ref. [25], though for parton-shower tuning is it is important to have measurements not only of jet shape averages but also of the full jet shape probability distributions. In terms of specific measurements that should be highest priority for ATLAS and CMS, our study has not revealed a silver bullet. Rather, all of the observables studied in this paper show similar levels of disagreement between generators, so a systematic LHC study of even one observable is likely to offer crucial new information.

What does seem to be essential is to make LHC measurements at multiple jet p_T scales with multiple jet radii R in multiple different quark/gluon-enriched samples. Unfolded distributions would be the most useful for constraining parton-shower uncertainties, but even detector-level measurements compared to detector-simulated parton showers could help spot troubling trends. For the IRC-safe angularities in particular, studying the β dependence would help separate information about collinear and soft radiation patterns, especially given the fact that the β trends seen in the parton-shower generators here disagree with those seen in ref. [20]. In addition, measurements of both groomed and ungroomed jet shapes could help disentangle collinear versus soft effects.

If possible, it would be interesting to study the LHA ($\beta = 1/2$) on archival LEP data, since this angularity probes the core of jets in a new way, distinct from broadening-like ($\beta = 1$) or thrust-like ($\beta = 2$) observables. Among the IRC-safe angularities studied here, the LHA has the best predicted discrimination power, making it (and other $0 < \beta < 1$ angularities) a well-motivated target for future lepton collider measurements. Similarly, it would be worthwhile to improve our analytic understanding of the LHA. From figure 5c, we see that the LHA has discrimination power both at small values of $\lambda^1_{0.5}$ (where nonperturbative corrections play an important role) as well as at larger values of $\lambda^1_{0.5}$ (where fixed-order corrections are important). Therefore, one must go beyond an NLL understanding to accurately describe the quark/gluon performance of the LHA.

The key lesson to parton-shower authors is that, contrary to some standard lore, existing LEP measurements used for tuning *do not* constrain all of the relevant aspects of the

JHEP07(2017)091

final state parton shower. While we have extensive information about quark-jet radiation patterns from LEP event shapes, gluon-jet radiation patterns are largely unconstrained. This has important implications for parton-shower tuning strategies, since LHC data can and should be used to adjust final-state shower parameters. For example, the ATLAS A14 tune of Pythia has a 10% lower value of α_s in the final-state shower compared to the Monash tune, which yields better agreement with charged-particle multiplicity distributions [25]. However, A14 has not been tested on LEP event shapes, suggesting that a global tuning strategy is needed. In addition, it is worth mentioning that similar quark/gluon studies have been carried out in deep inelastic electron-proton scattering [122], which offer an intermediate step between pp and e^+e^- collisions, and this ep data could also be valuable for parton-shower tuning.

Interestingly, there are LEP measurements that do constrain gluon radiation patterns, as recently summarized by K. Hamacher in ref. [123]. Unfortunately, these are not currently implemented in Rivet [49] and, to our knowledge, are not used in any present-day parton-shower tuning strategy. The cleanest LEP studies focused on $Z \to b\bar{b}g$ events, i.e. 3-jet events with two heavy-flavor tagged jets [124, 125]. By applying appropriate event-selection cuts, these studies identified "symmetric events" where the gluon was relatively isolated [126–129]; this strategy was extended to more general 3-jet topologies using Lorentz-invariant $p_\perp$ scales [130]. In the rare case that the two tagged jets appeared in the same hemisphere of an event, the opposite hemisphere could be used to define an inclusive sample of gluon-like jets [131–134]. With a relatively pure gluon-jet sample, one could then study various aspects of gluon-jet fragmentation, including hadron multiplicity, single-hadron energy fractions, and y-splitting scales. It should be noted, however, that in at least some of the above analyses, the corrections to hadron level made use of Monte Carlo truth information to correct not only for photon initial-state radiation and detector effects but also for impurities in the gluon-jet selection. We therefore encourage efforts to determine the extent to which Rivet implementations of these measurements are practicable, and to begin that process if the corrections are deemed to be sufficiently model independent. This would enable a broader suite of LEP measurements to be included in the next round of parton-shower tunes and in global comparisons such as MCplots[20] [135].

In a similar spirit, a future high-luminosity lepton collider would allow for measurements of the above processes with a high precision (see e.g. [123]). At sufficiently high collision energies, one can also measure other interesting processes, such as associated Higgs production with the Higgs boson decaying to bottom quarks or gluons. Such measurements would provide an invaluable source of data in the context of quark/gluon discrimination and, more generically, for parton-shower tuning.

Based on this study, we have identified three aspects of the final-state parton shower that deserve closer scrutiny.

- *Gluon splitting to quarks.* Some of the largest differences between generators came from turning on and off the $g \to q\bar{q}$ splitting process. While Pythia, Sherpa, Vincia, Ariadne and Dire suggest that (unphysically) turning off $g \to q\bar{q}$ would im-

[20] http://mcplots.cern.ch.

JHEP07(2017)091

prove quark/gluon separation, Herwig (and the analytic calculation from section 4) suggests the opposite conclusion. Beyond quark/gluon discrimination, it would be helpful to identify other contexts where $g \to q\bar{q}$ might play an important role (see e.g. [136]).

- *Color reconnection in the final state.* Color reconnection is often thought of as an issue mainly at hadron colliders, but we have seen that it can have an impact in e^+e^- collisions as well. This is particularly the case with the default color reconnection model in Herwig, since it allows the reconnection of color/anticolor lines even if they originally came from an octet configuration. We also saw large changes from "Pythia: CR1" and "Ariadne: no swing", suggesting that one should revisit color reconnection physics when tuning generators to LEP data.

- *Reconsidering α_s defaults*: in the context of parton-shower tuning, the value of α_s used internally within a code need not match the world average value, since higher-order effects not captured by the shower can often be mimicked by adjusting α_s. That said, one has to be careful whether a value of α_s tuned for one process is really appropriate for another. For example, Pythia uses a relatively large value of α_s in its final-state shower, which allows it to match LEP event shape data. The same value of α_s, though, probably also leads to too much radiation within gluon jets.

Finally, we want to emphasize that despite the uncertainties currently present in parton-shower generators [137–140], parton showers in particular (and QCD resummation techniques more generally) will be essential for understanding quark/gluon discrimination. Fixed-order QCD calculations cannot reliably probe the very soft and very collinear structure of jets, which is precisely where valuable information about quark/gluon radiation patterns reside. Given the ubiquity and value of parton-shower generators, improving the understanding of quark/gluon discrimination will assist every jet study at the LHC.

Acknowledgments

We thank Marat Freytsis and Davison Soper for contributing to the original Les Houches study [1] that formed the basis for this work. We particularly thank Davison Soper for contributing plots based on Deductor. We thank L'École de Physique des Houches and the organizers of the 2015 Les Houches workshop on "Physics at TeV Colliders" for a stimulating environment while this work was being initiated. We benefitted from helpful discussions and feedback from Samuel Bein, Andy Buckley, Jon Butterworth, Mario Campanelli, Peter Loch, Ben Nachman, Zoltan Nagy, Chris Pollard, Stefan Prestel, Salvatore Rappoccio, Gavin Salam, Alexander Schmidt, Frank Tackmann, and Wouter Waalewijn. The work of SH is supported by the U.S. Department of Energy (DOE) under contract DE-AC02-76SF00515. DK acknowledges support from the Science Faculty Research Council, University of Witwatersrand. The work of LL, SP, and AS is supported in part by the MCnetITN FP7 Marie Curie Initial Training Network, contract PITN-GA-2012-315877. LL is also supported by the Swedish Research Council (contracts 621-2012-2283 and 621-2013-4287).

JHEP07(2017)091

SP acknowledges support from a FP7 Marie Curie Intra European Fellowship under Grant Agreement PIEF-GA-2013-628739. AS acknowledges support from COST Action CA15213 THOR and the National Science Centre, Poland Grant No. 2016/23/D/ST2/02605. PS is the recipient of an Australian Research Council Future Fellowship, FT130100744. The work of GS is supported in part by the Paris-Saclay IDEX under the IDEOPTIMALJE grant, by the French Agence Nationale de la Recherche, under grant ANR-15-CE31-0016, and by the ERC Advanced Grant Higgs@LHC (No. 321133). The work of JT is supported by the DOE under grant contract numbers DE-SC-00012567 and DE-SC-00015476.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

JHEP07(2017)091

References

[1] J.R. Andersen et al., *Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*, arXiv:1605.04692 [INSPIRE].

[2] H.P. Nilles and K.H. Streng, *Quark-Gluon Separation in Three Jet Events*, *Phys. Rev.* **D 23** (1981) 1944.

[3] L.M. Jones, *Tests for Determining the Parton Ancestor of a Hadron Jet*, *Phys. Rev.* **D 39** (1989) 2550.

[4] Z. Fodor, *How to See the Differences Between Quark and Gluon Jets*, *Phys. Rev.* **D 41** (1990) 1726.

[5] L. Jones, *Towards a systematic jet classification*, *Phys. Rev.* **D 42** (1990) 811 [INSPIRE].

[6] L. Lönnblad, C. Peterson and T. Rognvaldsson, *Using neural networks to identify jets*, *Nucl. Phys.* **B 349** (1991) 675 [INSPIRE].

[7] J. Pumplin, *How to tell quark jets from gluon jets*, *Phys. Rev.* **D 44** (1991) 2025.

[8] J. Gallicchio and M.D. Schwartz, *Quark and Gluon Tagging at the LHC*, *Phys. Rev. Lett.* **107** (2011) 172001 [arXiv:1106.3076] [INSPIRE].

[9] J. Gallicchio and M.D. Schwartz, *Quark and Gluon Jet Substructure*, *JHEP* **04** (2013) 090 [arXiv:1211.7038] [INSPIRE].

[10] D. Krohn, M.D. Schwartz, T. Lin and W.J. Waalewijn, *Jet Charge at the LHC*, *Phys. Rev. Lett.* **110** (2013) 212001 [arXiv:1209.2421] [INSPIRE].

[11] F. Pandolfi and D. Del Re, *Search for the Standard Model Higgs Boson in the $H \to ZZ \to llqq$ Decay Channel at CMS*. Ph.D. Thesis, Zurich, ETH (2012).

[12] CMS collaboration, *Search for a Higgs boson in the decay channel $H \to ZZ^*$ to $q\bar{q}\ell^-\ell^+$ in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **04** (2012) 036 [arXiv:1202.1416] [INSPIRE].

[13] A.J. Larkoski, G.P. Salam and J. Thaler, *Energy Correlation Functions for Jet Substructure*, *JHEP* **06** (2013) 108 [arXiv:1305.0007] [INSPIRE].

[14] A.J. Larkoski, J. Thaler and W.J. Waalewijn, *Gaining (Mutual) Information about Quark/Gluon Discrimination*, *JHEP* **11** (2014) 129 [arXiv:1408.3122] [INSPIRE].

[15] B. Bhattacherjee, S. Mukhopadhyay, M.M. Nojiri, Y. Sakaki and B.R. Webber, *Associated jet and subjet rates in light-quark and gluon jet discrimination*, *JHEP* **04** (2015) 131 [arXiv:1501.04794] [INSPIRE].

[16] D. Ferreira de Lima, P. Petrov, D. Soper and M. Spannowsky, *quark-gluon tagging with Shower Deconstruction: Unearthing dark matter and Higgs couplings*, *Phys. Rev.* **D 95** (2017) 034001 [arXiv:1607.06031] [INSPIRE].

[17] B. Bhattacherjee, S. Mukhopadhyay, M.M. Nojiri, Y. Sakaki and B.R. Webber, *quark-gluon discrimination in the search for gluino pair production at the LHC*, *JHEP* **01** (2017) 044 [arXiv:1609.08781] [INSPIRE].

[18] P.T. Komiske, E.M. Metodiev and M.D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, *JHEP* **01** (2017) 110 [arXiv:1612.01551] [INSPIRE].

[19] J. Davighi and P. Harris, *Fractal based observables to probe jet substructure of quarks and gluons*, arXiv:1703.00914 [INSPIRE].

[20] ATLAS collaboration, *Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C 74** (2014) 3023 [arXiv:1405.6583] [INSPIRE].

[21] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C 75** (2015) 17 [arXiv:1406.0076] [INSPIRE].

[22] CMS collaboration, *Measurement of electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C 75** (2015) 66 [arXiv:1410.3153] [INSPIRE].

[23] ATLAS collaboration, *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **12** (2015) 055 [arXiv:1506.00962] [INSPIRE].

[24] CMS collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$*, *Phys. Rev.* **D 92** (2015) 032008 [arXiv:1506.01010] [INSPIRE].

[25] ATLAS collaboration, *Measurement of the charged-particle multiplicity inside jets from $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, *Eur. Phys. J.* **C 76** (2016) 322 [arXiv:1602.00988] [INSPIRE].

[26] C.F. Berger, T. Kucs and G.F. Sterman, *Event shape/energy flow correlations*, *Phys. Rev.* **D 68** (2003) 014012 [hep-ph/0303051] [INSPIRE].

[27] L.G. Almeida, S.J. Lee, G. Perez, G.F. Sterman, I. Sung and J. Virzi, *Substructure of high-p_T Jets at the LHC*, *Phys. Rev.* **D 79** (2009) 074017 [arXiv:0807.0234] [INSPIRE].

[28] S.D. Ellis, C.K. Vermilion, J.R. Walsh, A. Hornig and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, *JHEP* **11** (2010) 101 [arXiv:1001.0014] [INSPIRE].

[29] A.J. Larkoski, D. Neill and J. Thaler, *Jet Shapes with the Broadening Axis*, *JHEP* **04** (2014) 017 [arXiv:1401.2158] [INSPIRE].

[30] S. Catani, G. Turnock and B.R. Webber, *Jet broadening measures in e^+e^- annihilation*, *Phys. Lett.* **B 295** (1992) 269 [INSPIRE].

JHEP07(2017)091

[31] P.E.L. Rakow and B.R. Webber, *Transverse Momentum Moments of Hadron Distributions in QCD Jets*, *Nucl. Phys.* **B 191** (1981) 63 [INSPIRE].

[32] R.K. Ellis and B.R. Webber, *QCD Jet Broadening in Hadron Hadron Collisions*, *Conf. Proc.* **C 860623** (1986) 74 [INSPIRE].

[33] E. Farhi, *A QCD Test for Jets*, *Phys. Rev. Lett.* **39** (1977) 1587 [INSPIRE].

[34] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS(ACAT)040 [physics/0703039] [INSPIRE].

[35] M. Dasgupta and G.P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett.* **B 512** (2001) 323 [hep-ph/0104277] [INSPIRE].

[36] M. Dasgupta, F. Dreyer, G.P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039 [arXiv:1411.5182] [INSPIRE].

[37] G.P. Korchemsky and G.F. Sterman, *Power corrections to event shapes and factorization*, *Nucl. Phys.* **B 555** (1999) 335 [hep-ph/9902341] [INSPIRE].

[38] G.P. Korchemsky and S. Tafat, *On power corrections to the event shape distributions in QCD*, *JHEP* **10** (2000) 010 [hep-ph/0007005] [INSPIRE].

[39] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [INSPIRE].

[40] M. Bahr et al., *HERWIG++ Physics and Manual*, *Eur. Phys. J.* **C 58** (2008) 639 [arXiv:0803.0883] [INSPIRE].

[41] J. Bellm et al., *HERWIG++ 2.7 Release Note*, arXiv:1310.6877 [INSPIRE].

[42] J. Bellm et al., *HERWIG 7.0/HERWIG++ 3.0 release note*, *Eur. Phys. J.* **C 76** (2016) 196 [arXiv:1512.01178] [INSPIRE].

[43] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622] [INSPIRE].

[44] N. Fischer, S. Prestel, M. Ritzmann and P. Skands, *Vincia for Hadron Colliders*, *Eur. Phys. J.* **C 76** (2016) 589 [arXiv:1605.06142] [INSPIRE].

[45] Z. Nagy and D.E. Soper, *A parton shower based on factorization of the quantum density matrix*, *JHEP* **06** (2014) 097 [arXiv:1401.6364] [INSPIRE].

[46] C. Flensburg, G. Gustafson and L. Lönnblad, *Inclusive and Exclusive Observables from Dipoles in High Energy Collisions*, *JHEP* **08** (2011) 103 [arXiv:1103.4321] [INSPIRE].

[47] L. Lönnblad, *ARIADNE version 4: A program for simulation of QCD cascades implementing the color dipole model*, *Comput. Phys. Commun.* **71** (1992) 15 [INSPIRE].

[48] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, *Eur. Phys. J.* **C 75** (2015) 461 [arXiv:1506.05057] [INSPIRE].

[49] A. Buckley et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803 [arXiv:1003.0694] [INSPIRE].

[50] ALEPH collaboration, A. Heister et al., *Studies of QCD at e^+e^- centre-of-mass energies between 91-GeV and 209-GeV*, *Eur. Phys. J.* **C 35** (2004) 457 [INSPIRE].

[51] DELPHI collaboration, J. Abdallah et al., *A Study of the energy evolution of event shape distributions and their means with the DELPHI detector at LEP*, *Eur. Phys. J.* **C 29** (2003) 285 [hep-ex/0307048] [INSPIRE].

JHEP07(2017)091

[52] L3 collaboration, P. Achard et al., *Studies of hadronic event structure in e^+e^- annihilation from 30-GeV to 209-GeV with the L3 detector*, *Phys. Rept.* **399** (2004) 71 [hep-ex/0406049] [INSPIRE].

[53] OPAL collaboration, G. Abbiendi et al., *Measurement of event shape distributions and moments in $e^+e^- \to$ hadrons at 91–209 GeV and a determination of α_s*, *Eur. Phys. J.* **C 40** (2005) 287 [hep-ex/0503051] [INSPIRE].

[54] J. Gallicchio and M.D. Schwartz, *Pure Samples of Quark and Gluon Jets at the LHC*, *JHEP* **10** (2011) 103 [arXiv:1104.1175] [INSPIRE].

[55] A. Buckley and C. Pollard, *QCD-aware partonic jet clustering for truth-jet flavour labelling*, *Eur. Phys. J.* **C 76** (2016) 71 [arXiv:1507.00508] [INSPIRE].

[56] A. Banfi, G.P. Salam and G. Zanderighi, *Infrared safe definition of jet flavor*, *Eur. Phys. J.* **C 47** (2006) 113 [hep-ph/0601139] [INSPIRE].

[57] C. Frye, A.J. Larkoski, M.D. Schwartz and K. Yan, *Precision physics with pile-up insensitive observables*, arXiv:1603.06375 [INSPIRE].

[58] C. Frye, A.J. Larkoski, M.D. Schwartz and K. Yan, *Factorization for groomed jet substructure beyond the next-to-leading logarithm*, *JHEP* **07** (2016) 064 [arXiv:1603.09338] [INSPIRE].

[59] A.J. Larkoski and J. Thaler, *Unsafe but Calculable: Ratios of Angularities in Perturbative QCD*, *JHEP* **09** (2013) 137 [arXiv:1307.1699] [INSPIRE].

[60] A.J. Larkoski, I. Moult and D. Neill, *Toward Multi-Differential Cross sections: Measuring Two Angularities on a Single Jet*, *JHEP* **09** (2014) 046 [arXiv:1401.4458] [INSPIRE].

[61] M. Procura, W.J. Waalewijn and L. Zeune, *Resummation of Double-Differential Cross sections and Fully-Unintegrated Parton Distribution Functions*, *JHEP* **02** (2015) 117 [arXiv:1410.6483] [INSPIRE].

[62] A. Hornig, Y. Makris and T. Mehen, *Jet Shapes in Dijet Events at the LHC in SCET*, *JHEP* **04** (2016) 097 [arXiv:1601.01319] [INSPIRE].

[63] W.J. Waalewijn, *Calculating the Charge of a Jet*, *Phys. Rev.* **D 86** (2012) 094030 [arXiv:1209.3019] [INSPIRE].

[64] H.-M. Chang, M. Procura, J. Thaler and W.J. Waalewijn, *Calculating Track-Based Observables for the LHC*, *Phys. Rev. Lett.* **111** (2013) 102002 [arXiv:1303.6637] [INSPIRE].

[65] H.-M. Chang, M. Procura, J. Thaler and W.J. Waalewijn, *Calculating Track Thrust with Track Functions*, *Phys. Rev.* **D 88** (2013) 034030 [arXiv:1306.6630] [INSPIRE].

[66] M. Cacciari and G.P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B 641** (2006) 57 [hep-ph/0512210] [INSPIRE].

[67] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C 72** (2012) 1896 [arXiv:1111.6097] [INSPIRE].

[68] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189] [INSPIRE].

[69] D. Bertolini, T. Chan and J. Thaler, *Jet Observables Without Jet Algorithms*, *JHEP* **04** (2014) 013 [arXiv:1310.7584] [INSPIRE].

[70] G. Salam, *E_t^∞ Scheme*, unpublished.

JHEP07(2017)091

[71] G.C. Blazey et al., *Run II jet physics*, hep-ex/0005012 [INSPIRE].

[72] Y.L. Dokshitzer, A. Lucenti, G. Marchesini and G.P. Salam, *On the QCD analysis of jet broadening*, *JHEP* **01** (1998) 011 [hep-ph/9801324] [INSPIRE].

[73] A. Banfi, G.P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073 [hep-ph/0407286] [INSPIRE].

[74] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323] [INSPIRE].

[75] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, hep-ph/9907280 [INSPIRE].

[76] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [arXiv:0802.2470] [INSPIRE].

[77] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, *Phys. Rev.* **D 80** (2009) 051501 [arXiv:0903.5081] [INSPIRE].

[78] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, *Phys. Rev.* **D 81** (2010) 094023 [arXiv:0912.0033] [INSPIRE].

[79] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [arXiv:0912.1342] [INSPIRE].

[80] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *JHEP* **09** (2013) 029 [arXiv:1307.0007] [INSPIRE].

[81] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657] [INSPIRE].

[82] A.J. Larkoski and J. Thaler, *Aspects of jets at* 100 *TeV*, *Phys. Rev.* **D 90** (2014) 034010 [arXiv:1406.7011] [INSPIRE].

[83] A.J. Larkoski, S. Marzani and J. Thaler, *Sudakov Safety in Perturbative QCD*, *Phys. Rev.* **D 91** (2015) 111501 [arXiv:1502.01719] [INSPIRE].

[84] CMS collaboration, *Splitting function in pp and PbPb collisions at 5.02 TeV*, CMS-PAS-HIN-16-006.

[85] M. Jankowiak and A.J. Larkoski, *Jet Substructure Without Trees*, *JHEP* **06** (2011) 057 [arXiv:1104.1646] [INSPIRE].

[86] I. Moult, L. Necib and J. Thaler, *New Angles on Energy Correlation Functions*, *JHEP* **12** (2016) 153 [arXiv:1609.07483] [INSPIRE].

[87] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [arXiv:1011.2268] [INSPIRE].

[88] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, *JHEP* **02** (2012) 093 [arXiv:1108.2701] [INSPIRE].

[89] G.P. Salam, L. Schunk and G. Soyez, *Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging*, *JHEP* **03** (2017) 022 [arXiv:1612.03917] [INSPIRE].

[90] D.E. Soper and M. Spannowsky, *Finding physics signals with shower deconstruction*, *Phys. Rev.* **D 84** (2011) 074002 [arXiv:1102.3480] [INSPIRE].

JHEP07(2017)091

[91] D.E. Soper and M. Spannowsky, *Finding top quarks with shower deconstruction*, *Phys. Rev.* **D 87** (2013) 054012 [arXiv:1211.3140] [INSPIRE].

[92] D.E. Soper and M. Spannowsky, *Finding physics signals with event deconstruction*, *Phys. Rev.* **D 89** (2014) 094005 [arXiv:1402.1189] [INSPIRE].

[93] A. Hornig, C. Lee and G. Ovanesyan, *Effective Predictions of Event Shapes: Factorized, Resummed and Gapped Angularity Distributions*, *JHEP* **05** (2009) 122 [arXiv:0901.3780] [INSPIRE].

[94] A.V. Manohar and M.B. Wise, *Power suppressed corrections to hadronic event shapes*, *Phys. Lett.* **B 344** (1995) 407 [hep-ph/9406392] [INSPIRE].

[95] Y.L. Dokshitzer and B.R. Webber, *Calculation of power corrections to hadronic event shapes*, *Phys. Lett.* **B 352** (1995) 451 [hep-ph/9504219] [INSPIRE].

[96] G.P. Salam and D. Wicke, *Hadron masses and power corrections to event shapes*, *JHEP* **05** (2001) 061 [hep-ph/0102343] [INSPIRE].

[97] C. Lee and G.F. Sterman, *Momentum Flow Correlations from Event Shapes: Factorized Soft Gluons and Soft-Collinear Effective Theory*, *Phys. Rev.* **D 75** (2007) 014022 [hep-ph/0611061] [INSPIRE].

[98] V. Mateu, I.W. Stewart and J. Thaler, *Power Corrections to Event Shapes with Mass-Dependent Operators*, *Phys. Rev.* **D 87** (2013) 014025 [arXiv:1209.3781] [INSPIRE].

[99] I.W. Stewart, F.J. Tackmann and W.J. Waalewijn, *Dissecting Soft Radiation with Factorization*, *Phys. Rev. Lett.* **114** (2015) 092001 [arXiv:1405.6722] [INSPIRE].

[100] Y.I. Azimov, Y.L. Dokshitzer, V.A. Khoze and S.I. Troyan, *Similarity of Parton and Hadron Spectra in QCD Jets*, *Z. Phys.* **C 27** (1985) 65 [INSPIRE].

[101] Y.L. Dokshitzer, G. Marchesini and B.R. Webber, *Dispersive approach to power behaved contributions in QCD hard processes*, *Nucl. Phys.* **B 469** (1996) 93 [hep-ph/9512336] [INSPIRE].

[102] A. Guffanti and G.E. Smye, *Nonperturbative effects in the W and Z transverse momentum distribution*, *JHEP* **10** (2000) 025 [hep-ph/0007190] [INSPIRE].

[103] S. Gieseke, M.H. Seymour and A. Siodmok, *A model of non-perturbative gluon emission in an initial state parton shower*, *JHEP* **06** (2008) 001 [arXiv:0712.1199] [INSPIRE].

[104] M. Dasgupta, F.A. Dreyer, G.P. Salam and G. Soyez, *Inclusive jet spectrum for small-radius jets*, *JHEP* **06** (2016) 057 [arXiv:1602.01110] [INSPIRE].

[105] P. Skands, S. Carrazza and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, *Eur. Phys. J.* **C 74** (2014) 3024 [arXiv:1404.5630] [INSPIRE].

[106] G. Miu and T. Sjöstrand, *W production in an improved parton shower approach*, *Phys. Lett.* **B 449** (1999) 313 [hep-ph/9812455] [INSPIRE].

[107] J.R. Christiansen and P.Z. Skands, *String Formation Beyond Leading Colour*, *JHEP* **08** (2015) 003 [arXiv:1505.01681] [INSPIRE].

[108] S. Gieseke, C. Rohr and A. Siodmok, *Colour reconnections in HERWIG++*, *Eur. Phys. J.* **C 72** (2012) 2225 [arXiv:1206.0041] [INSPIRE].

[109] M.H. Seymour and A. Siodmok, *Constraining MPI models using $\sigma_{\rm eff}$ and recent Tevatron and LHC Underlying Event data*, *JHEP* **10** (2013) 113 [arXiv:1307.5015] [INSPIRE].

JHEP07(2017)091

[110] M. Bahr, M. Myska, M.H. Seymour and A. Siodmok, *Extracting $\sigma_{\text{effective}}$ from the CDF γ+3jets measurement*, *JHEP* **03** (2013) 129 [arXiv:1302.4325] [INSPIRE].

[111] S. Platzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in HERWIG++*, *Eur. Phys. J.* **C 72** (2012) 2187 [arXiv:1109.6256] [INSPIRE].

[112] S. Catani, F. Krauss, R. Kuhn and B.R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231] [INSPIRE].

[113] W.T. Giele, D.A. Kosower and P.Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev.* **D 84** (2011) 054003 [arXiv:1102.2126] [INSPIRE].

[114] S. Catani, B.R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, *Nucl. Phys.* **B 349** (1991) 635 [INSPIRE].

[115] N. Fischer, S. Gieseke, S. Plätzer and P. Skands, *Revisiting radiation patterns in e^+e^- collisions*, *Eur. Phys. J.* **C 74** (2014) 2831 [arXiv:1402.3186] [INSPIRE].

[116] L. Hartgring, E. Laenen and P. Skands, *Antenna Showers with One-Loop Matrix Elements*, *JHEP* **10** (2013) 127 [arXiv:1303.4974] [INSPIRE].

[117] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, *JHEP* **03** (2013) 166 [arXiv:1211.7278] [INSPIRE].

[118] H.T. Li and P. Skands, *A framework for second-order parton showers*, *Phys. Lett.* **B 771** (2017) 59 [arXiv:1611.00013] [INSPIRE].

[119] C. Bierlich, G. Gustafson, L. Lönnblad and A. Tarasov, *Effects of Overlapping Strings in pp Collisions*, *JHEP* **03** (2015) 148 [arXiv:1412.6259] [INSPIRE].

[120] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244] [INSPIRE].

[121] L.M. Dery, B. Nachman, F. Rubbo and A. Schwartzman, *Weakly Supervised Classification in High Energy Physics*, *JHEP* **05** (2017) 145 [arXiv:1702.00414] [INSPIRE].

[122] ZEUS collaboration, S. Chekanov et al., *Substructure dependence of jet cross sections at HERA and determination of α_s*, *Nucl. Phys.* **B 700** (2004) 3 [hep-ex/0405065] [INSPIRE].

[123] D. d'Enterria et al., *Parton Radiation and Fragmentation from LHC to FCC-ee*, arXiv:1702.01329.

[124] OPAL collaboration, G. Alexander et al., *A Direct observation of quark-gluon jet differences at LEP*, *Phys. Lett.* **B 265** (1991) 462 [INSPIRE].

[125] OPAL collaboration, P.D. Acton et al., *A Study of differences between quark and gluon jets using vertex tagging of quark jets*, *Z. Phys.* **C 58** (1993) 387 [INSPIRE].

[126] ALEPH collaboration, D. Buskulic et al., *Quark and gluon jet properties in symmetric three jet events*, *Phys. Lett.* **B 384** (1996) 353 [INSPIRE].

[127] DELPHI collaboration, P. Abreu et al., *Energy dependence of the differences between the quark and gluon jet fragmentation*, *Z. Phys.* **C 70** (1996) 179 [INSPIRE].

[128] OPAL collaboration, R. Akers et al., *A model independent measurement of quark and gluon jet properties and differences*, *Z. Phys.* **C 68** (1995) 179 [INSPIRE].

[129] DELPHI collaboration, P. Abreu et al., *Investigation of the splitting of quark and gluon jets*, *Eur. Phys. J.* **C 4** (1998) 1 [INSPIRE].

JHEP07(2017)091

[130] DELPHI collaboration, P. Abreu et al., *Measurement of the gluon fragmentation function and a comparison of the scaling violation in gluon and quark jets*, *Eur. Phys. J.* **C 13** (2000) 573 [INSPIRE].

[131] OPAL collaboration, G. Alexander et al., *Test of QCD analytic predictions for the multiplicity ratio between gluon and quark jets*, *Phys. Lett.* **B 388** (1996) 659 [INSPIRE].

[132] OPAL collaboration, K. Ackerstaff et al., *Multiplicity distributions of gluon and quark jets and tests of QCD analytic predictions*, *Eur. Phys. J.* **C 1** (1998) 479 [hep-ex/9708029] [INSPIRE].

[133] OPAL collaboration, G. Abbiendi et al., *Experimental properties of gluon and quark jets from a point source*, *Eur. Phys. J.* **C 11** (1999) 217 [hep-ex/9903027] [INSPIRE].

[134] OPAL collaboration, G. Abbiendi et al., *Tests of models of color reconnection and a search for glueballs using gluon jets with a rapidity gap*, *Eur. Phys. J.* **C 35** (2004) 293 [hep-ex/0306021] [INSPIRE].

[135] A. Karneyeu, L. Mijovic, S. Prestel and P.Z. Skands, *MCPLOTS: a particle physics resource based on volunteer computing*, *Eur. Phys. J.* **C 74** (2014) 2714 [arXiv:1306.3436] [INSPIRE].

[136] P. Ilten, N.L. Rodd, J. Thaler and M. Williams, *Disentangling Heavy Flavor at Colliders*, arXiv:1702.02947 [INSPIRE].

[137] J. Bellm, G. Nail, S. Plätzer, P. Schichtel and A. Siódmok, *Parton Shower Uncertainties with HERWIG 7: Benchmarks at Leading Order*, *Eur. Phys. J.* **C 76** (2016) 665 [arXiv:1605.01338] [INSPIRE].

[138] J. Bellm, S. Plätzer, P. Richardson, A. Siódmok and S. Webster, *Reweighting Parton Showers*, *Phys. Rev.* **D 94** (2016) 034028 [arXiv:1605.08256] [INSPIRE].

[139] S. Mrenna and P. Skands, *Automated Parton-Shower Variations in PYTHIA 8*, *Phys. Rev.* **D 94** (2016) 074005 [arXiv:1605.08352] [INSPIRE].

[140] E. Bothmann, M. Schönherr and S. Schumann, *Reweighting QCD matrix-element and parton-shower calculations*, *Eur. Phys. J.* **C 76** (2016) 590 [arXiv:1606.08753] [INSPIRE].

JHEP07(2017)091