1. SYSTEMATICS OF QUARK/GLUON TAGGING[1,2,3]

By measuring the substructure of a jet, one can assign it a "quark" or "gluon" tag. In the eikonal limit, quark/gluon discrimination is determined solely by the color factor of the initiating parton (C_F versus C_A). In this section, we confront the challenges faced when going beyond this leading-order understanding, using parton shower generators 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.

1.1 Overview

Jets are a robust tool 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 [1, 2, 3, 4, 5, 6, 7, 8], and there is a growing catalog of quark/gluon studies at the Large Hadron Collider (LHC) [9, 10, 11, 12, 13, 14].

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 quark radiates proportional to its $C_F = 4/3$ color factor while a gluon radiates proportional to $C_A = 3$. In this section, we demonstrate that quark/gluon discrimination performance is highly sensitive to subleading perturbative effects beyond the eikonal limit, such as $g \to q\bar{q}$ 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 section is to highlight these uncertainties, which then suggests a set of future LHC measurements that will improve the modeling of jets in general and quark/gluon tagging in particular.

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 Sec. 1.2, one can indeed create a well-defined quark/gluon tagging procedure based 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.

There are a wide range of possible quark/gluon discriminants and a similarly large range of ways to quantify discrimination power. As a concrete set of discriminants, we consider the generalized angularities λ_β^κ [7] (see also [15, 16, 17, 18]),

$$\lambda_\beta^\kappa = \sum_{i \in \text{jet}} z_i^\kappa \theta_i^\beta, \tag{1}$$

[1] Section coordinators: Gregory Soyez and Jesse Thaler

[2] Contributing authors: Marat Freytsis, Philippe Gras, Deepak Kar, Leif Lönnblad, Simon Plätzer, Andrzej Siodmok, Peter Skands, and Davison Soper

[3] Additional input: Samuel Bein, Andy Buckley, Jon Butterworth, Mario Campanelli, Andrew Larkoski, Peter Loch, Ben Nachman, Zoltan Nagy, Chris Pollard, Salvatore Rappoccio, Gavin Salam, Alexander Schmidt, Frank Tackmann, and Wouter Waalewijn

with the notation to be explained in Sec. 1.3. We consider five different (κ, β) working points, which roughly map onto five variables in common use in the literature:

$$
\begin{array}{ccccc}
(0,0) & (2,0) & (1,0.5) & (1,1) & (1,2) \\
\text{multiplicity} & p_T^D & \text{LHA} & \text{width} & \text{mass}
\end{array}
\tag{2}
$$

Here, multiplicity is the hadron multiplicity within the jet, p_T^D was defined in Refs. [4, 5], LHA refers to the "Les Houches Angularity" (named after the venue of this workshop), width is closely related to jet broadening [19, 20, 21], and mass is closely related to jet thrust [22]. To quantify discrimination performance, we focus on classifier separation (a default output of TMVA [23]):

$$
\Delta = \frac{1}{2} \int d\lambda \, \frac{\left(p_q(\lambda) - p_g(\lambda)\right)^2}{p_q(\lambda) + p_g(\lambda)},
\tag{3}
$$

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 Sec. 1.4.

We begin our parton shower generator predictions for quark/gluon discrimination in Sec. 1.5, using an idealized setup with e^+e^- collisions. Here, we can use the following processes as proxies for quark and gluon jets:

$$
\text{"quark jets"}: \quad e^+e^- \to (\gamma/Z)^* \to u\bar{u},
\tag{4}
$$

$$
\text{"gluon jets"}: \quad e^+e^- \to h^* \to gg,
\tag{5}
$$

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.[4] We compare six different parton shower generators both before hadronization ("parton level") and after hadronization ("hadron level"):

- PYTHIA 8.205 [24, 25],
- HERWIG++ 2.7.1 [26, 27],
- SHERPA 2.1.1 [28],
- VINCIA 1.201 [29],
- DEDUCTOR 1.0.2 [30] (with hadronization performed by PYTHIA),[5]
- ARIADNE 5.0.β [31].[6]

In the future, we plan to make HERWIG 7 [33] and DIRE [34] predictions about quark/gluon discrimination as well as investigate predictions from analytic resummation [6, 7].

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 infrared-and-collinear (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 Sec. 1.6, we study these differences as a function of collision energy Q, jet radius R, and strong coupling constant α_s, showing that the generators have somewhat different discrimination trends. In Sec. 1.7, 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.

---

[4] 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.

[5] Note that this DEDUCTOR plus PYTHIA combination has not yet been tuned to data.

[6] 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 [32].

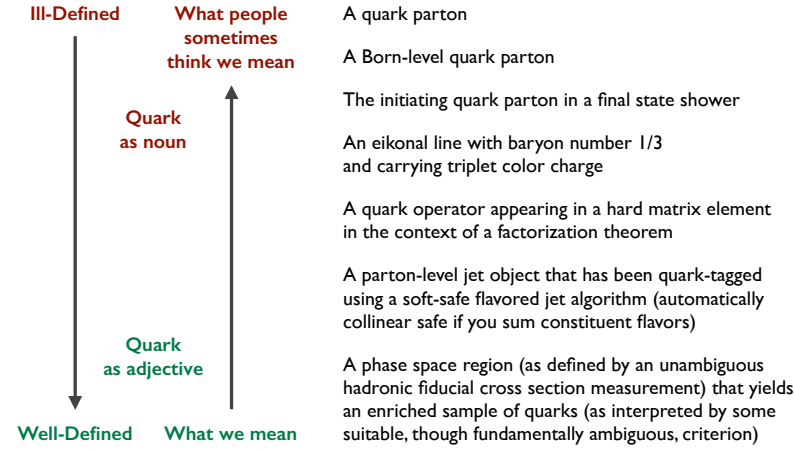## What is a Quark Jet?
From lunch/dinner discussions

Ill-Defined · What people sometimes think we mean · A quark parton

A Born-level quark parton

The initiating quark parton in a final state shower

Quark as noun · An eikonal line with baryon number 1/3 and carrying triplet color charge

A quark operator appearing in a hard matrix element in the context of a factorization theorem

A parton-level jet object that has been quark-tagged using a soft-safe flavored jet algorithm (automatically collinear safe if you sum constituent flavors)

Quark as adjective · 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)

Well-Defined · What we mean

Fig. 1: Original slide from the June 10, 2015 summary report of the quark/gluon Les Houches subgroup.

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 [35, 36, 37, 38]. In Sec. 1.8, 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 suspect that much can be learned from hadron-level measurements, even without isolating "pure" quark or gluon samples.

We present our final recommendations and conclusions in Sec. 1.9. The main take home message from this study is that, contrary to the standard lore, existing measurements at e^+e^- colliders are insufficient to constrain uncertainties in the final state shower. Therefore, gluon-enriched measurements at the LHC will be crucial to achieve robust quark/gluon discrimination.

1.2 What is a Quark/Gluon Jet?

As part of the 2015 Les Houches workshop, an attempt was made to define exactly what is meant by a "quark jet" or "gluon jet" (see Fig. 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 strongly-ordered or leading-logarithmic

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 "max-p_T" prescription is a variant on the initiating parton prescription above (see further discussion in [39]). It differs from the initiating parton by a calculable amount in a leading logarithmic shower [40] 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 [41]. 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 [42]. 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. 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, but it tries to avoid relying on the presence of a 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 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 will 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 $d\sigma/d\lambda$ 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 this physics goals of a specific analysis). Typically, one might use a Born-level or eikonal analysis to define which regions of
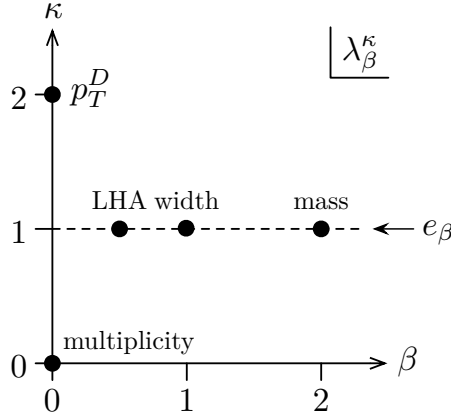
Fig. 2: Two-parameter family of generalized angularities, adapted from [7]. The dots correspond to the five benchmark angularities used in this study. The horizontal line at $\kappa = 1$ corresponds to the IRC safe angularities, $e_\beta = \lambda_\beta^1$.

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 Sec. 1.5, 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 the truth-level "quark" and "gluon" labels are defined by the Lorentz structure of the production vertex. In Sec. 1.9, 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 $d\sigma/d\lambda_\beta^\kappa$ 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.

1.3 Generalized Angularities

A wide variety of quark/gluon discriminants have been proposed (see [2] for an extensive catalog), but here we limit ourselves to a two-parameter family of generalized angularities [7], shown in Fig. 2. These are defined as (repeating Eq. (1) for convenience)

$$\lambda_\beta^\kappa = \sum_{i \in \text{jet}} z_i^\kappa \theta_i^\beta, \tag{6}$$

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 [18] (see also [17, 43, 44, 45, 46]). For general $\kappa \neq 1$, there are quasi-perturbative techniques based on generalized fragmentation functions [7] (see also [3, 47, 48, 49]). In our parton shower studies, we determine λ_β^κ 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 using the ee-variant of the anti-k_t algorithm [50], with $|\vec{p}|$-ordered winner-take-all recombination [18, 51, 52] to determine the jet axis $\hat{n}$. Unlike standard E-scheme recombination [53], 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

[6, 18, 19, 54, 55]. We define

$$z_i \equiv \frac{E_i}{E_{\text{jet}}}, \qquad \theta_i \equiv \frac{\Omega_{i\hat{n}}}{R}, \tag{7}$$

where 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). To translate our ee study to an eventual pp study (left to future work), one would use the standard pp version of anti-k_t with p_T-ordered winner-take-all recombination, defining

$$z_i \equiv \frac{p_{Ti}}{\sum_{j\in\text{jet}} p_{Tj}}, \qquad \theta_i \equiv \frac{R_{i\hat{n}}}{R}, \tag{8}$$

where p_{Ti} is the particle transverse momentum and $R_{i\hat{n}}$ is the rapidity-azimuth distance to the jet axis.

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

$$\begin{aligned}
(0,0) &= \text{hadron multiplicity}, \\
(2,0) &\Rightarrow p_T^D \text{ [4, 5] (specifically } \lambda_0^2 = (p_T^D)^2), \\
(1,0.5) &= \text{Les Houches Angularity (LHA)}, \\
(1,1) &= \text{width or broadening [19, 20, 21]}, \\
(1,2) &\Rightarrow \text{mass or thrust [22] (specifically } \lambda_2^1 \simeq m_{\text{jet}}^2 / E_{\text{jet}}^2).
\end{aligned} \tag{9}$$

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. [6].

For the IRC safe case of $\kappa = 1$, there is an alternative version of the angularities based on energy correlation functions [6] (see also [55, 56]),

$$\text{ecf}_\beta = \sum_{i<j\in\text{jet}} z_i z_j \theta_{ij}^\beta \simeq \lambda_\beta^1, \tag{10}$$

where equality holds in the extreme eikonal limit.[7] 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 [2] and 2-subjettiness [57, 58] can improve quark/gluon discrimination.

1.4 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 Eq. (3) for convenience),

$$\Delta = \frac{1}{2} \int d\lambda \, \frac{\left(p_q(\lambda) - p_g(\lambda)\right)^2}{p_q(\lambda) + p_g(\lambda)}, \tag{11}$$

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. At a point (q,g) on the ROC curve, where $q, g \in [0, 1]$, one can define a selection

---

[7]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.

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. To turn the ROC curve into a single number, it is common to report the gluon rejection rate at, say, 50% quark efficiency. Since we are more interested in understanding the relative performance between parton showers rather than the absolute performance, we will not show ROC curves here, 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 [7]). 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

$$I(T; \Lambda) = \int d\lambda \left(f\, p_q(\lambda) \log_2 \frac{p_q(\lambda)}{p_{\text{tot}}(\lambda)} + (1 - f)\, p_g(\lambda) \log_2 \frac{p_g(\lambda)}{p_{\text{tot}}(\lambda)} \right), \tag{12}$$

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

$$p_{\text{tot}}(\lambda) = f\, p_q(\lambda) + (1 - f)\, p_g(\lambda). \tag{13}$$

The choice $f = 1/2$ was used in Ref. [7], 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

$$-\frac{\log 2}{4} \frac{\partial^2 I(T; \Lambda)}{\partial f^2}\bigg|_{f=\frac{1}{2}} = \Delta. \tag{14}$$

One advantage of Δ over $I(T; \Lambda)$ is that the integrand in Eq. (11) is easier to interpret, since it tracks the fractional difference between the signal and background at a given value of λ.[8] Specifically, by plotting

$$\frac{d\Delta}{d\lambda} = \frac{1}{2} \frac{\left(p_q(\lambda) - p_g(\lambda)\right)^2}{p_q(\lambda) + p_g(\lambda)}, \tag{15}$$

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.

1.5 Idealized Quark/Gluon Discrimination

Our parton shower studies are based on the idealized case of discriminating quark and gluon jets in $e^+ e^-$ collisions. As we will see, this example demonstrates the importance of final state evolution for quark/gluon discrimination, separate from initial state complications arising in pp collisions.

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

---

[8]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 [7].

inescapable ambiguity in defining jet flavor.[9] To partially mitigate the effect of wide-angle emissions, we restrict our analysis to jets that satisfy

$$\frac{E_{\text{jet}}}{Q/2} > 0.8, \tag{16}$$

with up to two jets studied per event. There is also the ambiguity of which parton shower to use, and we investigate this ambiguity by looking at results from several event generators: PYTHIA 8.205 [24, 25], HERWIG++ 2.7.1 [26, 27], SHERPA 2.1.1 [28], VINCIA 1.201 [29], DEDUCTOR 1.0.2 [30] (with hadronization performed by PYTHIA), and ARIADNE 5.0.β [31].

In Fig. 3, we show hadron-level distributions of the LHA (i.e. $\lambda_{0.5}^1$) in the quark sample (p_q) and gluon sample (p_g), comparing the baseline settings of six different parton shower generators with a center-of-mass collision energy of $Q = 200$ GeV and jet radius $R = 0.6$. In the quark sample in Fig. 3a, 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 Fig. 3b, we see somewhat larger variations between the generators; this is expected since there is no data to directly constrain $e^+ e^- \rightarrow gg$. In Fig. 3c, we plot the integrand of classifier separation, $d\Delta/d\lambda$ from Eq. (15). This shows where in the LHA phase space the actual discrimination power lies, with large values 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 [9]. The predicted discrimination power from the other four generators are intermediate between these extremes.

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.

With that caveat in mind, we show parton-level results in Fig. 4. 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. This gives $\lambda_{0.5}^1 = 0$ 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. 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 Fig. 3a 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. (15) to obtain the value of classifier separation Δ for the LHA. This is shown in Fig. 5, which also includes the four other benchmark angularities from Eq. (9). There is a rather large spread in predicted discrimination power between the generators, especially at hadron level in Fig. 5a. 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 Fig. 5b).[10] This suggests a more fundamental

---

[9]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.

[10]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.
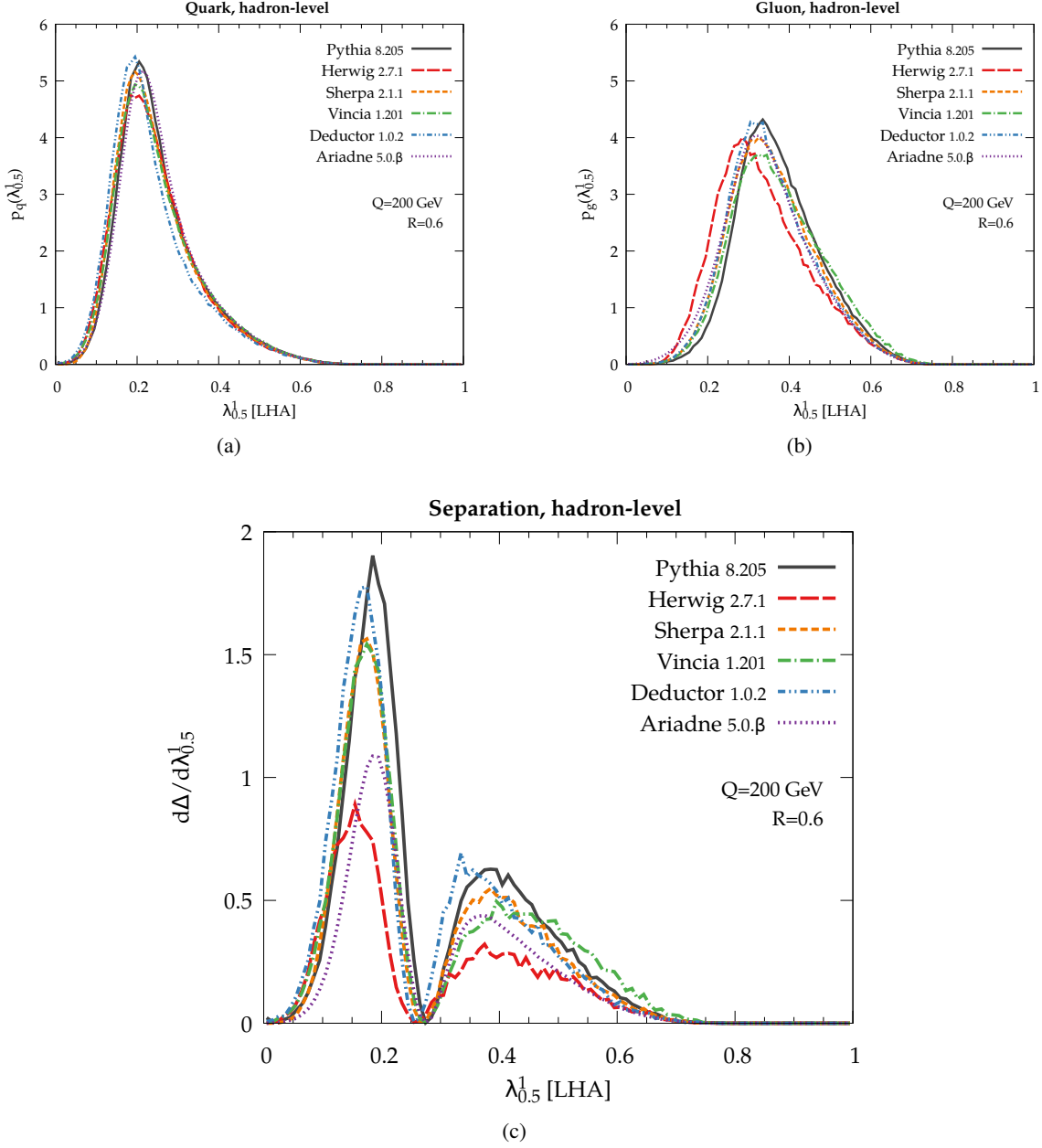
Fig. 3: Hadron-level distributions of the LHA for (a) the $e^+e^- \to u\bar{u}$ ("quark jet") sample, (b) the $e^+e^- \to gg$ ("gluon jet") sample, and (c) the classifier separation integrand in Eq. (15). Six parton shower generators—PYTHIA 8.205, HERWIG++ 2.7.1, SHERPA 2.1.1, VINCIA 1.201, DEDUCTOR 1.0.2, and ARIADNE 5.0.β—are run at their baseline settings with center-of-mass energy $Q = 200$ GeV and jet radius $R = 0.6$.
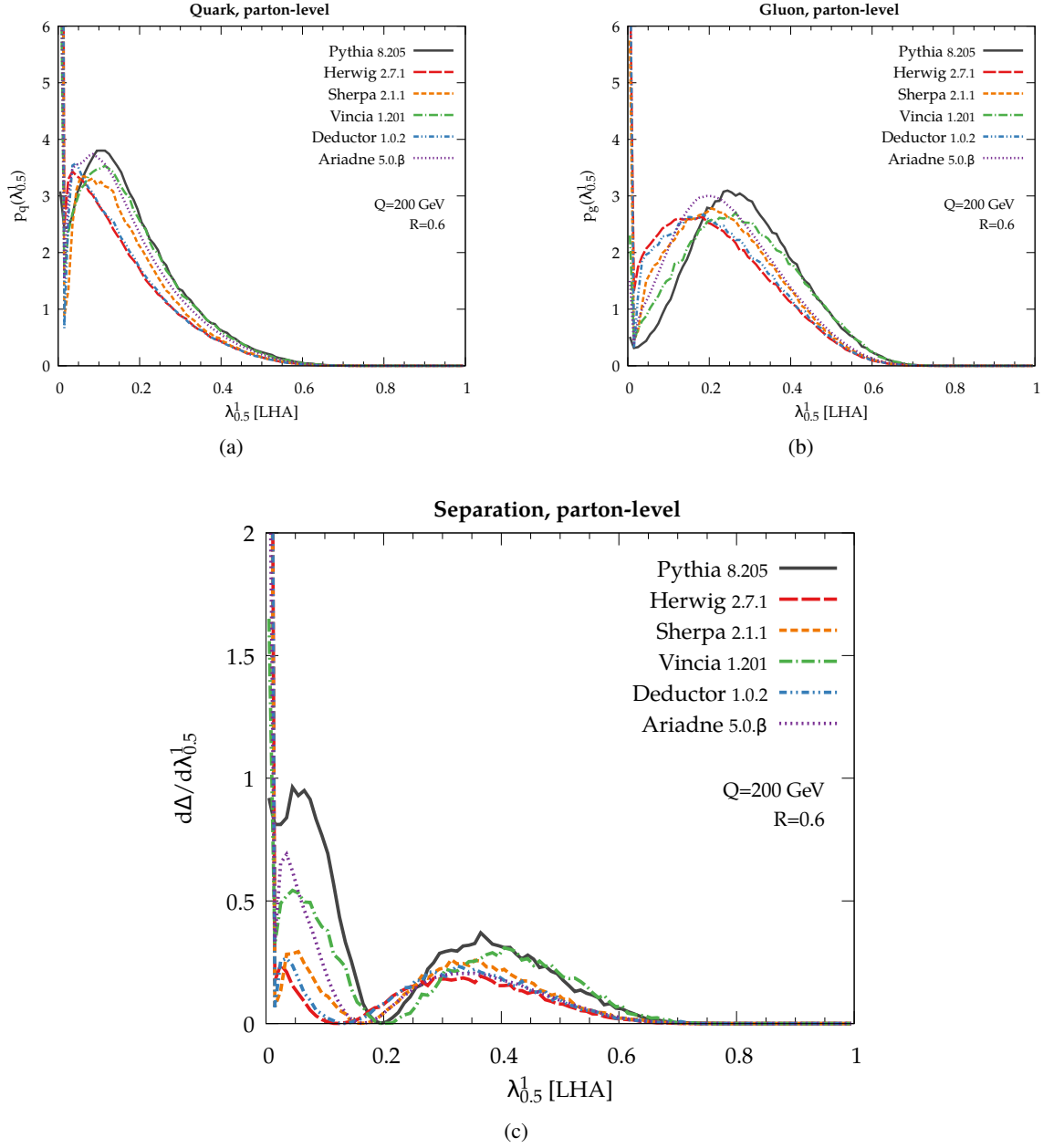
Quark, parton-level

Pythia 8.205
Herwig 2.7.1
Sherpa 2.1.1
Vincia 1.201
Deductor 1.0.2
Ariadne 5.0.β

Q=200 GeV
R=0.6

$p_q(\lambda_{0.5})$

$\lambda_{0.5}^1$ [LHA]

(a)

Gluon, parton-level

Pythia 8.205
Herwig 2.7.1
Sherpa 2.1.1
Vincia 1.201
Deductor 1.0.2
Ariadne 5.0.β

Q=200 GeV
R=0.6

$p_g(\lambda_{0.5})$

$\lambda_{0.5}^1$ [LHA]

(b)

Separation, parton-level

Pythia 8.205
Herwig 2.7.1
Sherpa 2.1.1
Vincia 1.201
Deductor 1.0.2
Ariadne 5.0.β

Q=200 GeV
R=0.6

$d\Delta/d\lambda_{0.5}^1$

$\lambda_{0.5}^1$ [LHA]

(c)

Fig. 4: Same as Fig. 3, but at the parton level. Note that HERWIG, SHERPA, and DEDUCTOR all have cross section spikes at $\lambda_{0.5}^1 = 0$ that extend above the plotted range.
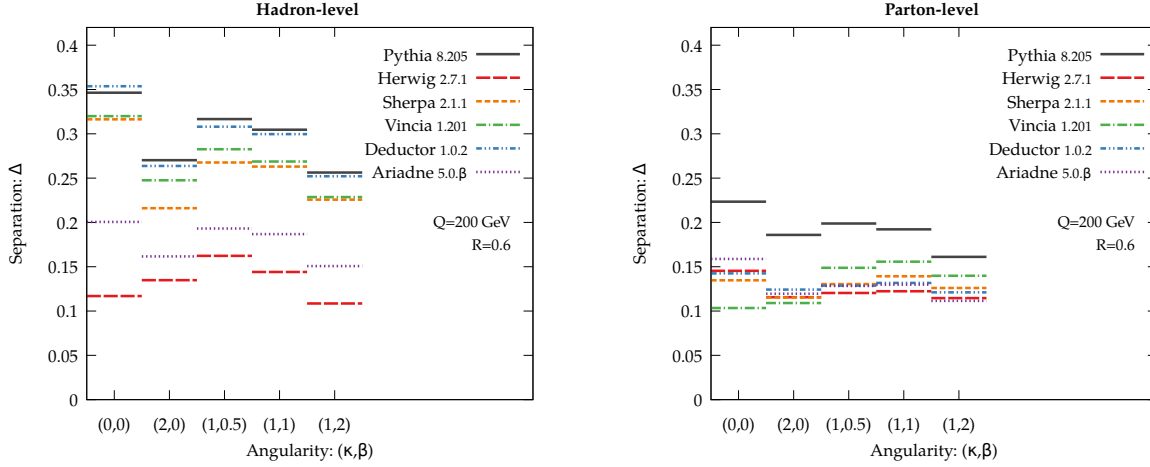
Fig. 5: Classifier separation Δ for the five benchmark angularities in Eq. (9), determined from the various generators at (a) hadron level and (b) parton level. The first two columns correspond to IRC unsafe distributions (multiplicity and p_T^D), while the last three columns are the IRC safe angularities. The LHA (i.e. $\kappa = 1$, $\beta = 1/2$) is shown in the middle column.

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. [6], but disagrees with the ATLAS study in Ref. [9], 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.

1.6 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:

$$\text{Collision Energy}: Q = \{50, 100, \mathbf{200}, 400, 800\} \text{ GeV},$$
$$\text{Jet Radius}: R = \{0.2, 0.4, \mathbf{0.6}, 0.8, 1.0\}, \tag{17}$$
$$\text{Strong Coupling}: \alpha_s/\alpha_{s0} = \{0.8, 0.9, \mathbf{1.0}, 1.1, 1.2\},$$

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 Fig. 6, 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 Fig. 5) 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 an effect 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 next-to-leading-logarithmic (NLL) logic in Ref. [6], 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. Indeed, PYTHIA, HERWIG, and ARIADNE do 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, all of the generators show the opposite trend.
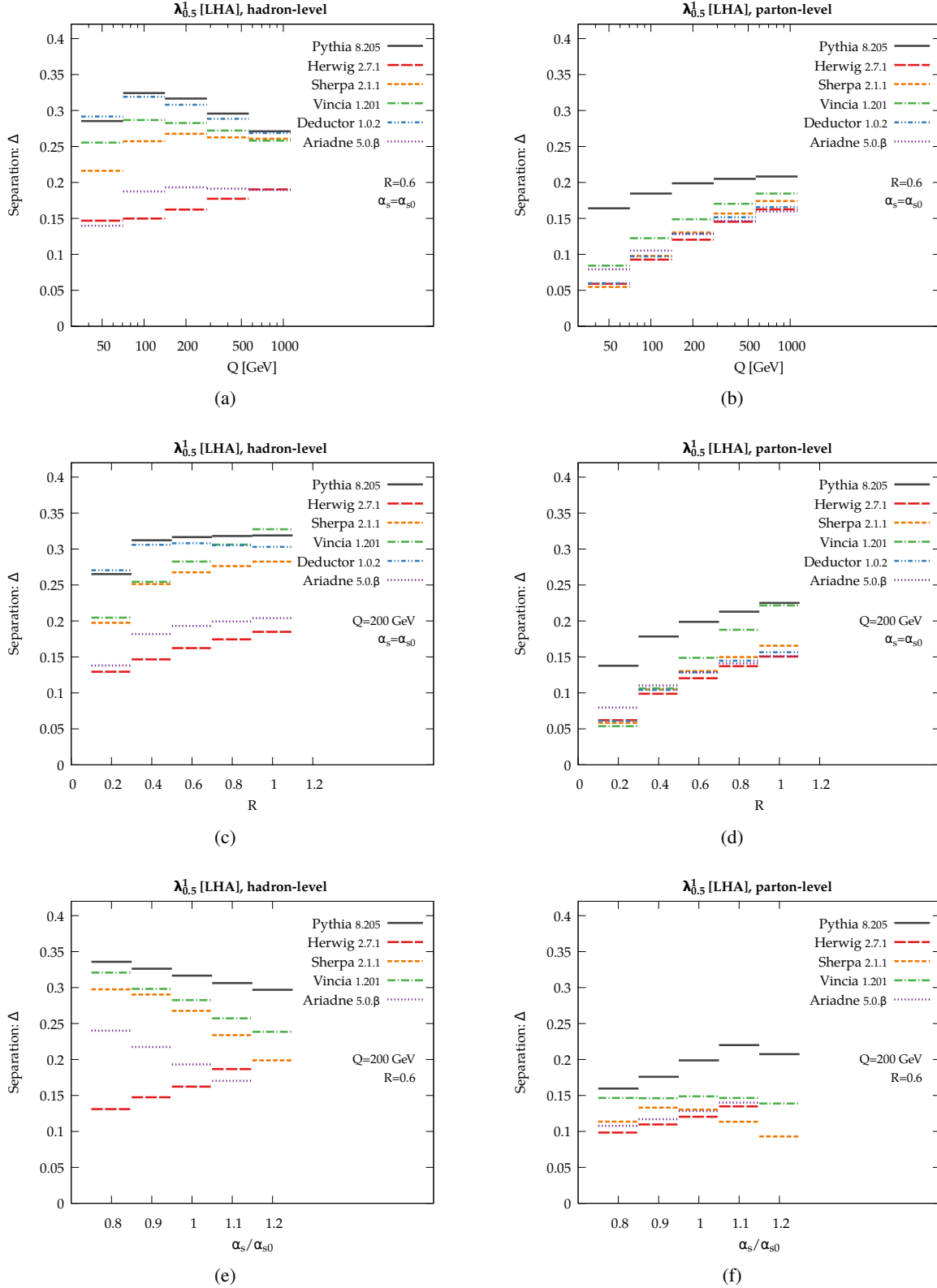
Fig. 6: Classifier separation Δ for the LHA, sweeping the collision energy Q (top row), jet radius R (middle row), and coupling constant α_s/α_{s0} (bottom row). Results are shown at hadron level (left column) and parton level (right column).

One reason to expect quark/gluon discrimination to improve as higher energies is that that 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 for more emissions within a jet, and from scaling symmetry, one expects that parton-level discrimination power should depend on the combination QR.[11] 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 probably 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. The study in Ref. [6] did not include nonperturbative hadronization corrections, so we do not yet have a theoretical expectation for the impact of hadronization. In future work, we plan to follow Ref. [43] and include nonperturbative hadronization corrections via shape functions [60, 61, 62, 63, 64, 65, 66]. With confusions already at parton level, though, further perturbative calculations beyond NLL accuracy are also needed. For example, at NLL accuracy, one does account for the fact that a jet can contain multiple perturbative emissions, but those emissions are treated as if they themselves do not radiate. By contrast, parton showers allow every emission to reradiate, which might be driving the Q and R discrimination trends.

1.7 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.

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 motivated. 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. [67]. In Fig. 7, 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 [68], but this variation turns those corrections off, showing the impact of non-singular terms. This has relatively little effect in our study, however, since there is no matrix element correction available for $h^* \to gg$.

- 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

---

[11] 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 [40, 59].
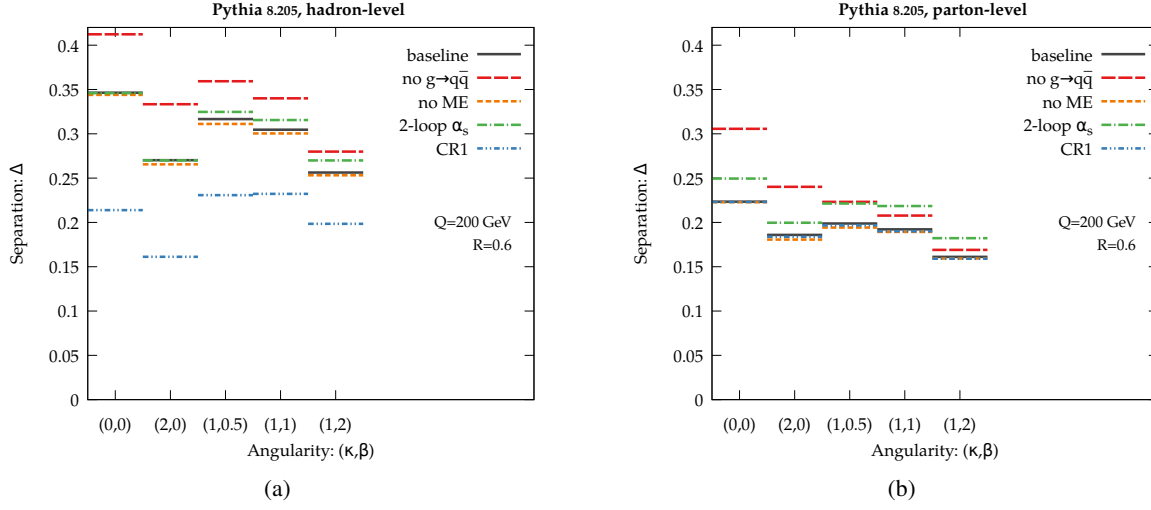
Fig. 7: Settings variations for PYTHIA 8.205. Shown are (a) hadron-level and (b) parton-level results for the classifier separation Δ derived from the five benchmark angularities.
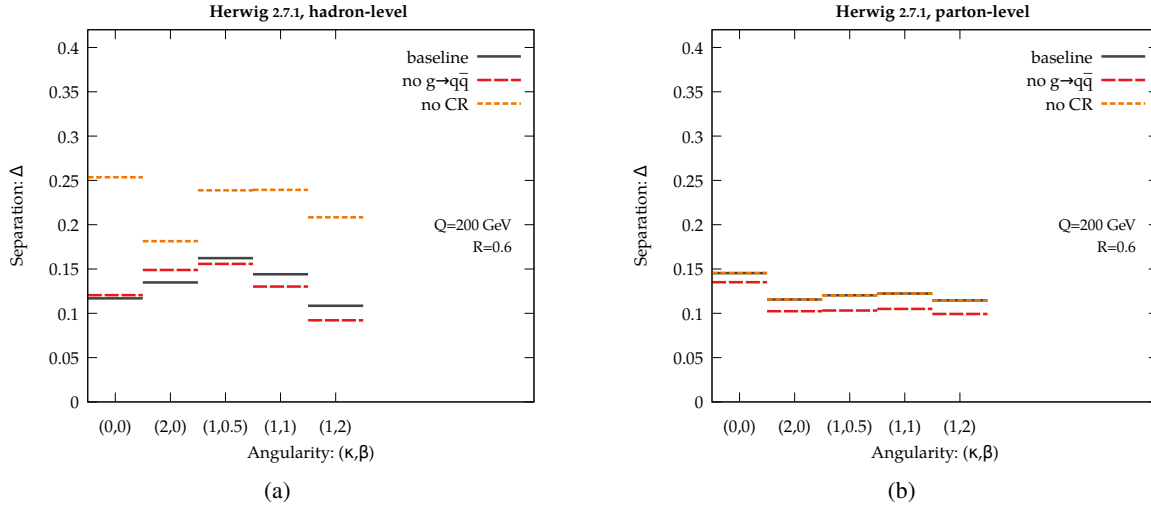

Fig. 8: Same as Fig. 7, but for HERWIG 2.7.1.

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 [69] (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 decreases 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 [70] and the most recent UE-EE-5-MRST tune [71], which is able to describe the double-parton scattering cross section [72] and underlying event data from $\sqrt{s} = 300$ GeV to $\sqrt{s} = 7$ TeV. In Fig. 8, we consider the following HERWIG variations:

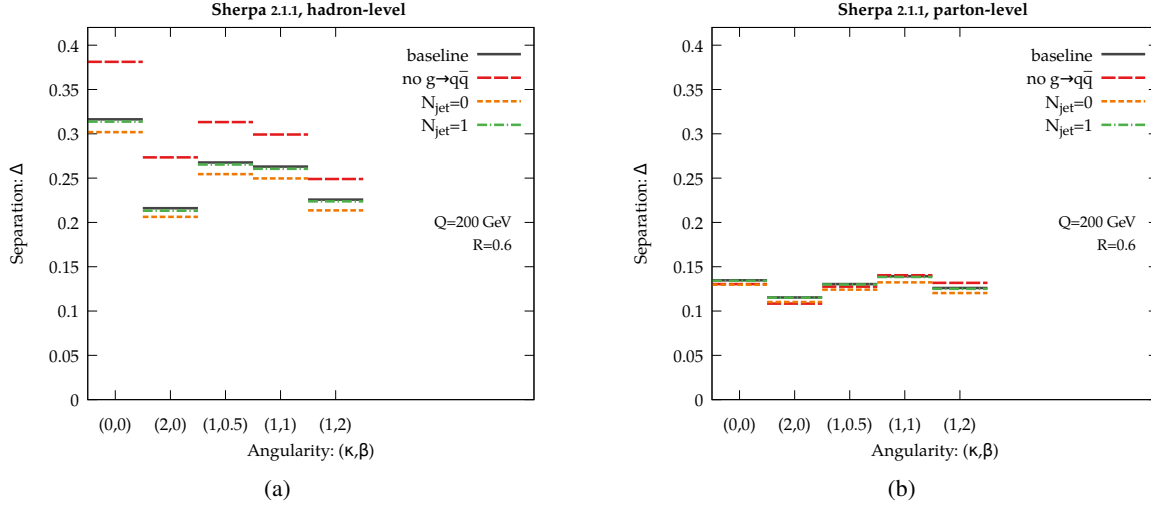- HERWIG: NO $g \to q\bar{q}$. Turning off $g \to q\bar{q}$ splittings in HERWIG has the reverse behavior as seen

Fig. 9: Same as Fig. 7, but for SHERPA 2.1.1.

in PYTHIA, leading to slightly worse discrimination power, though the effect is modest.

- 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 [73].

Our SHERPA baseline uses matrix element corrections for the first two emissions ($N_{\text{jet}} = 2$) with CKKW-style matching [74]. In Fig. 9, 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.

Our VINCIA baseline is based on the "jeppsson5" tune. While VINCIA has NLO matrix elements for $e^+e^- \to q\bar{q}$, it does not have them for $e^+e^- \to gg$, so we will use LO VINCIA throughout. In Fig. 10, 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 antenna are already rather close to the true matrix elements, turning off these matrix elements has a modest effect on quark/gluon discrimination power.
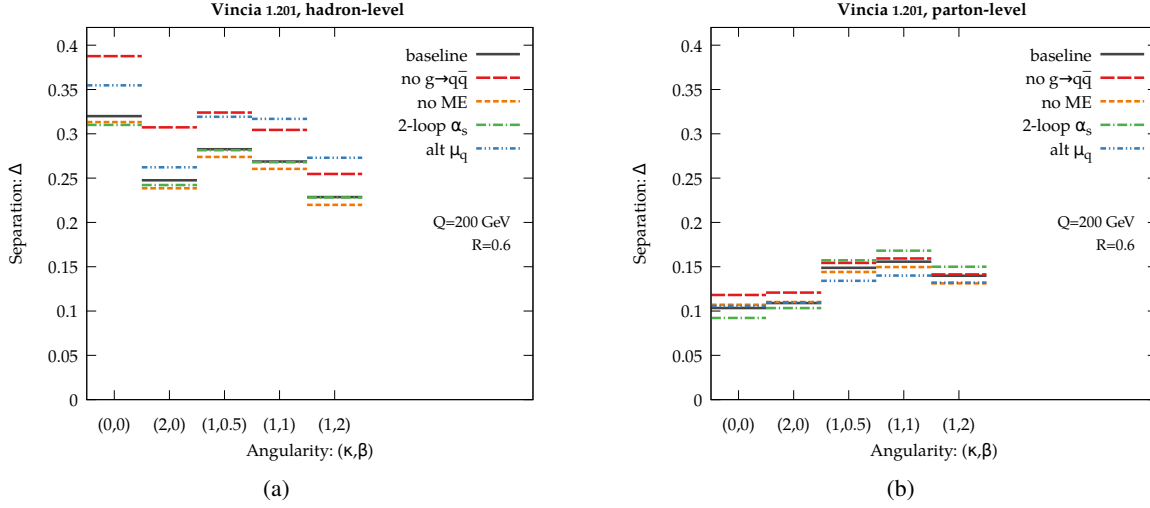
Fig. 10: Same as Fig. 7, but for VINCIA 1.201.

- VINCIA: 2-LOOP α_s. Like for PYTHIA, this variation switches from 1-loop to 2-loop α_s running, yielding a modest parton-level difference and almost no hadron-level difference.
- VINCIA: ALT μ_q. By default, the α_s scale used in VINCIA is determined by the transverse momentum of the corresponding antenna. In this variation, the α_s scale is set to half the invariant mass of the mother antenna. This slightly decreases the discrimination power at hadron level, but increases the discrimination power at parton level, again showing the complicated interplay of perturbative and nonperturbative effects.

Since VINCIA and PYTHIA share the same underlying engine, it is not surprising that they exhibit similar behaviors as parameters are changed. The biggest surprise is the way that changing the α_s scale for the antenna can lead to different trends at parton and hadron level.

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 PYTHIA 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, but we anticipate studying the effect of $g \rightarrow q\bar{q}$ splitting in future work.

Finally, our ARIADNE baseline is based on a beta release of version 5. In Fig. 11, we consider the following ARIADNE variation:

- ARIADNE: NO $g \rightarrow q\bar{q}$. This variation turns off $g \rightarrow q\bar{q}$, leading to modest change in separation power, similar in magnitude to HERWIG.
- 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 [31, 75]. 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.

1.8 Looking Towards 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. 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. Beyond just the application to quark/gluon tagging, this is an important
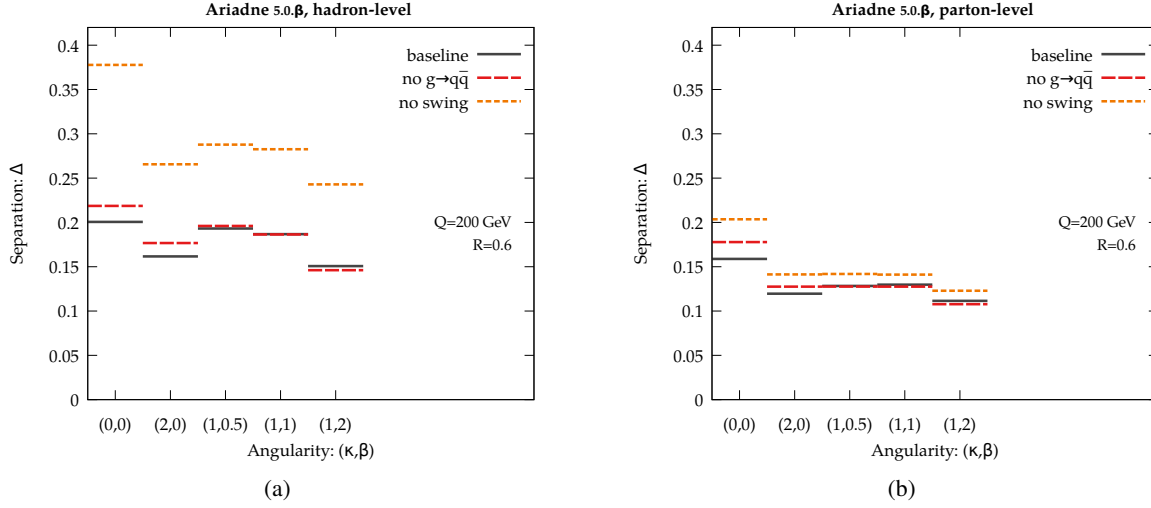
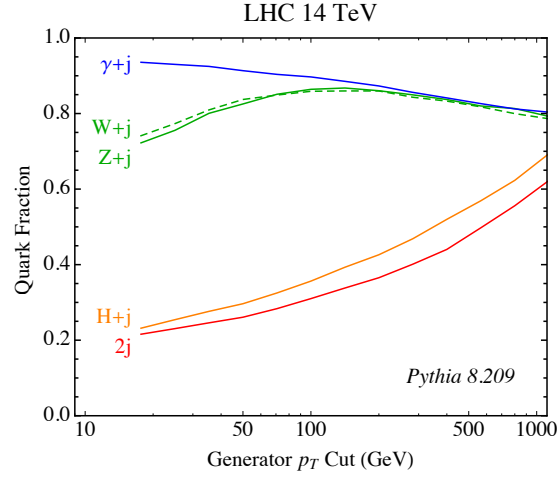Fig. 11: Same as Fig. 7, but for ARIADNE 5.0.β.

Fig. 12: Quark fraction of jets at parton level, as defined by the Born-level parton flavor.

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.

At the LHC, there is no way to isolate pure samples of quark or gluon jets, 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 Fig. 12, the 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. [42].

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. [9]). In the spirit of Sec. 1.2, 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. For example, instead of quark/gluon separation, one can ask the more well-defined question about whether one can tell "the jet in Z plus jets" (quark-enriched) apart from "the jet in dijets" (gluon-enriched). Similarly, one could test for differences

within a jet sample, such as comparing central jets versus forward jets in dijet production. This process-based strategy is also helpful for sidestepping the known process dependance of defining quarks and gluons at the LHC, where color correlations to the rest of the event impede a universal definition of quark and gluon jets.

Already, one would learn a lot from unfolded measurements at the LHC of the generalized angularities in a variety of quark/gluon enriched samples. Even detector-level measurement comparing a wide range of generators (and generator modifications) would help in understanding how to improve quark/gluon modeling. The five benchmark angularities in Eq. (9) probe both the perturbative and nonperturbative structure of jets and are therefore a good starting point for a more comprehensive quark/gluon jet shape analysis. In this spirit, we are encouraged by the track multiplicity study of Ref. [14], 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.

Finally, we think it would be useful to perform LHC jet shape measurements after jet grooming (see e.g. [76, 77, 78, 79]). Often jet grooming is described as a pileup mitigation strategy, but even independent of pileup, grooming modifies the observed jet radiation patterns in ways that are interesting from the quark/gluon perspective. Using techniques from Refs. [80, 81], one can calculate the quark/gluon discrimination power for angularities after jet grooming. Those calculations generically predict that groomed jet shapes should have reduced quark/gluon discrimination power compared to ungroomed jet shapes. That said, because techniques like soft drop [81] remove soft radiation from jets, they tend to reduce the process dependence in quark/gluon radiation patterns, and may therefore yield a more robust theoretical definition for quark and gluon jets.

1.9   Summary and Recommendations

By measuring the substructure of jets, one can gain valuable information about the 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 is 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. At the moment, analytic calculations are not at a level of accuracy where they can offer a useful guide. 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.

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 generalized angularities studied here showed similar levels of disagreement between generators, so a systematic LHC study of even one observable is likely to offer crucial new information. What is essential is to make measurements at multiple jet p_T scales with multiple jet radii R in multiple different jet samples. Unfolded distributions would be most useful for parton shower tuning, but even detector-level measurements compared to detector-simulated parton showers would 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. [9].

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, LHA has the best predicted discrimination power, making it (and other $0 < \beta < 1$ angularities) a well-motivated target for future lepton collider measurements.

The key lesson to parton shower authors is that, contrary to some standard lore, existing LEP measurements *do not* constrain all of the relevant aspects of the 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 charge particle multiplicity distributions [14]. 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 [82], which offer an intermediate step between pp and e^+e^- collisions, and this ep data could also be valuable 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, and ARIADNE suggest that (unphysically) turning off $g \to q\bar{q}$ would improve quark/gluon separation, HERWIG (and the analytic calculation in Ref. [6]) 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.

- *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 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, 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.

References

[1] J. Gallicchio and M. D. Schwartz, *Phys. Rev. Lett.* **107** (2011) 172001, [`1106.3076`].

[2] J. Gallicchio and M. D. Schwartz, *JHEP* **04** (2013) 090, [`1211.7038`].

[3] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, *Phys. Rev. Lett.* **110** (2013), no. 21 212001, [`1209.2421`].

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

[5] S. Chatrchyan *et. al.*,, **CMS** Collaboration *JHEP* **04** (2012) 036, [`1202.1416`].

[6] A. J. Larkoski, G. P. Salam, and J. Thaler, *JHEP* **06** (2013) 108, [`1305.0007`].

[7] A. J. Larkoski, J. Thaler, and W. J. Waalewijn, *JHEP* **11** (2014) 129, [`1408.3122`].

[8] B. Bhattacherjee, S. Mukhopadhyay, M. M. Nojiri, Y. Sakaki, and B. R. Webber, *JHEP* **04** (2015) 131, [`1501.04794`].

[9] G. Aad *et. al.*,, **ATLAS** Collaboration *Eur. Phys. J.* **C74** (2014), no. 8 3023, [`1405.6583`].

[10] G. Aad *et. al.*,, **ATLAS** Collaboration *Eur. Phys. J.* **C75** (2015) 17, [`1406.0076`].

[11] V. Khachatryan *et. al.*,, **CMS** Collaboration *Eur. Phys. J.* **C75** (2015), no. 2 66, [`1410.3153`].

[12] G. Aad *et. al.*,, **ATLAS** Collaboration *JHEP* **12** (2015) 055, [`1506.00962`].

[13] V. Khachatryan *et. al.*,, **CMS** Collaboration *Phys. Rev.* **D92** (2015), no. 3 032008, [`1506.01010`].

[14] G. Aad *et. al.*,, **ATLAS** Collaboration `1602.00988`.

[15] C. F. Berger, T. Kucs, and G. F. Sterman, *Phys. Rev.* **D68** (2003) 014012, [`hep-ph/0303051`].

[16] L. G. Almeida, S. J. Lee, G. Perez, G. F. Sterman, I. Sung, and J. Virzi, *Phys. Rev.* **D79** (2009) 074017, [`0807.0234`].

[17] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, *JHEP* **11** (2010) 101, [`1001.0014`].

[18] A. J. Larkoski, D. Neill, and J. Thaler, *JHEP* **04** (2014) 017, [`1401.2158`].

[19] S. Catani, G. Turnock, and B. R. Webber, *Phys. Lett.* **B295** (1992) 269–276.

[20] P. E. L. Rakow and B. R. Webber, *Nucl. Phys.* **B191** (1981) 63.

[21] R. K. Ellis and B. R. Webber, *Conf. Proc.* **C860623** (1986) 74.

[22] E. Farhi, *Phys. Rev. Lett.* **39** (1977) 1587–1588.

[23] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla, *ArXiv Physics e-prints* (Mar., 2007) [`physics/0703039`].

[24] T. Sjostrand, S. Mrenna, and P. Z. Skands, *JHEP* **05** (2006) 026, [`hep-ph/0603175`].

[25] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *Comput. Phys. Commun.* **191** (2015) 159–177, [`1410.3012`].

[26] M. Bahr *et. al.*, *Eur. Phys. J.* **C58** (2008) 639–707, [`0803.0883`].

[27] J. Bellm *et. al.*, 1310.6877.

[28] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *JHEP* **02** (2009) 007, [0811.4622].

[29] W. T. Giele, L. Hartgring, D. A. Kosower, E. Laenen, A. J. Larkoski, J. J. Lopez-Villarejo, M. Ritzmann, and P. Skands, *PoS* **DIS2013** (2013) 165, [1307.1060].

[30] Z. Nagy and D. E. Soper, *JHEP* **06** (2014) 097, [1401.6364].

[31] C. Flensburg, G. Gustafson, and L. Lonnblad, *JHEP* **08** (2011) 103, [1103.4321].

[32] L. Lonnblad, *Comput. Phys. Commun.* **71** (1992) 15–31.

[33] J. Bellm *et. al.*, 1512.01178.

[34] S. Höche and S. Prestel, *Eur. Phys. J.* **C75** (2015), no. 9 461, [1506.05057].

[35] A. Heister *et. al.*,, **ALEPH** Collaboration *Eur. Phys. J.* **C35** (2004) 457–486.

[36] J. Abdallah *et. al.*,, **DELPHI** Collaboration *Eur. Phys. J.* **C29** (2003) 285–312, [hep-ex/0307048].

[37] P. Achard *et. al.*,, **L3** Collaboration *Phys. Rept.* **399** (2004) 71–174, [hep-ex/0406049].

[38] G. Abbiendi *et. al.*,, **OPAL** Collaboration *Eur. Phys. J.* **C40** (2005) 287–316, [hep-ex/0503051].

[39] A. Buckley and C. Pollard, 1507.00508.

[40] M. Dasgupta, F. Dreyer, G. P. Salam, and G. Soyez, *JHEP* **04** (2015) 039, [1411.5182].

[41] A. Banfi, G. P. Salam, and G. Zanderighi, *Eur. Phys. J.* **C47** (2006) 113–124, [hep-ph/0601139].

[42] J. Gallicchio and M. D. Schwartz, *JHEP* **10** (2011) 103, [1104.1175].

[43] A. J. Larkoski and J. Thaler, *JHEP* **09** (2013) 137, [1307.1699].

[44] A. J. Larkoski, I. Moult, and D. Neill, *JHEP* **09** (2014) 046, [1401.4458].

[45] M. Procura, W. J. Waalewijn, and L. Zeune, *JHEP* **02** (2015) 117, [1410.6483].

[46] A. Hornig, Y. Makris, and T. Mehen, 1601.01319.

[47] W. J. Waalewijn, *Phys. Rev.* **D86** (2012) 094030, [1209.3019].

[48] H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, *Phys. Rev. Lett.* **111** (2013) 102002, [1303.6637].

[49] H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, *Phys. Rev.* **D88** (2013) 034030, [1306.6630].

[50] M. Cacciari, G. P. Salam, and G. Soyez, *JHEP* **04** (2008) 063, [0802.1189].

[51] D. Bertolini, T. Chan, and J. Thaler, *JHEP* **04** (2014) 013, [1310.7584].

[52] G. Salam, *Unpublished*.

[53] G. C. Blazey *et. al.*, in *QCD and weak boson physics in Run II. Proceedings, Batavia, USA, March 4-6, June 3-4, November 4-6, 1999*, pp. 47–77, 2000. `hep-ex/0005012`.

[54] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. P. Salam, *JHEP* **01** (1998) 011, [`hep-ph/9801324`].

[55] A. Banfi, G. P. Salam, and G. Zanderighi, *JHEP* **03** (2005) 073, [`hep-ph/0407286`].

[56] M. Jankowiak and A. J. Larkoski, *JHEP* **06** (2011) 057, [`1104.1646`].

[57] J. Thaler and K. Van Tilburg, *JHEP* **03** (2011) 015, [`1011.2268`].

[58] J. Thaler and K. Van Tilburg, *JHEP* **02** (2012) 093, [`1108.2701`].

[59] M. Dasgupta, F. A. Dreyer, G. P. Salam, and G. Soyez, `1602.01110`.

[60] A. V. Manohar and M. B. Wise, *Phys. Lett.* **B344** (1995) 407–412, [`hep-ph/9406392`].

[61] Y. L. Dokshitzer and B. R. Webber, *Phys. Lett.* **B352** (1995) 451–455, [`hep-ph/9504219`].

[62] G. P. Korchemsky and G. F. Sterman, *Nucl. Phys.* **B555** (1999) 335–351, [`hep-ph/9902341`].

[63] G. P. Korchemsky and S. Tafat, *JHEP* **10** (2000) 010, [`hep-ph/0007005`].

[64] G. P. Salam and D. Wicke, *JHEP* **05** (2001) 061, [`hep-ph/0102343`].

[65] C. Lee and G. F. Sterman, *Phys. Rev.* **D75** (2007) 014022, [`hep-ph/0611061`].

[66] V. Mateu, I. W. Stewart, and J. Thaler, *Phys. Rev.* **D87** (2013), no. 1 014025, [`1209.3781`].

[67] P. Skands, S. Carrazza, and J. Rojo, *Eur. Phys. J.* **C74** (2014), no. 8 3024, [`1404.5630`].

[68] G. Miu and T. Sjostrand, *Phys. Lett.* **B449** (1999) 313–320, [`hep-ph/9812455`].

[69] J. R. Christiansen and P. Z. Skands, *JHEP* **08** (2015) 003, [`1505.01681`].

[70] S. Gieseke, C. Rohr, and A. Siodmok, *Eur. Phys. J.* **C72** (2012) 2225, [`1206.0041`].

[71] M. H. Seymour and A. Siodmok, *JHEP* **10** (2013) 113, [`1307.5015`].

[72] M. Bahr, M. Myska, M. H. Seymour, and A. Siodmok, *JHEP* **03** (2013) 129, [`1302.4325`].

[73] S. Platzer and S. Gieseke, *Eur. Phys. J.* **C72** (2012) 2187, [`1109.6256`].

[74] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *JHEP* **11** (2001) 063, [`hep-ph/0109231`].

[75] C. Bierlich, G. Gustafson, L. Lönnblad, and A. Tarasov, *JHEP* **03** (2015) 148, [`1412.6259`].

[76] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Phys. Rev. Lett.* **100** (2008) 242001, [`0802.2470`].

[77] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Phys. Rev.* **D80** (2009) 051501, [`0903.5081`].

[78] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Phys. Rev.* **D81** (2010) 094023, [`0912.0033`].

[79] D. Krohn, J. Thaler, and L.-T. Wang, *JHEP* **02** (2010) 084, [`0912.1342`].

[80] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, *JHEP* **09** (2013) 029, [`1307.0007`].

[81] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, *JHEP* **05** (2014) 146, [`1402.2657`].

[82] S. Chekanov *et. al.*,, **ZEUS** Collaboration *Nucl. Phys.* **B700** (2004) 3–50, [`hep-ex/0405065`].