Quark vs. Gluon Jets

Gregory Soyez[1]

[1] IPhT, CEA Saclay

Abstract: The ability to discriminate between quark and gluon jets has many applications in collider physics. In this contribution we briefly report on a work, initiated in the context of the 2015 Les Houches "Physics at TeV Colliders" workshop, which compares the quark/gluon tagging performance predicted by different Monte-Carlo generators. We discuss measurements at the LHC and at a FCC-ee that would further constrain quark/gluon tagging and Monte Carlo generators.

Introduction

Designing an method to effectively separate quark- and gluon-initiated jets is a longstanding open question. (see e.g. [1] for a series of possible candidates). It is usually done via jet substructure observables like jet shapes which exploit differences in the radiation pattern of quarks and gluons. In general, we are interested in developing quark/gluon discrimination tools that go as far as possible beyond the naive C_A v. C_F Casimir scaling and are able to do so with limited and controlled theoretical uncertainties.

A key question in that respect is how well current (parton-shower) Monte-Carlo generators agree on their respective predictions for the quark-gluon discriminating power. In this contribution, we report on a study presented in Ref [2] where we show based on an idealised case that the results obtained for the quark/gluon discriminating power differ sizeably between Monte Carlo generators.

For more details, we refer directly to Section IV.5 of Ref. [2] and references therein. Most of the results presented below are taken from an extended version in preparation, Ref [3].

Are quark and gluon jets well-defined?

Since quarks and gluons can branch into one another, are ill-defined concepts beyond the lowest order of the perturbative series, and are not directly observed in the final state of the collisions, the concept of a quark and a gluon jet might itself seem ill-defined at first sight.

Rather than trying to determine a truth definition of a quark or a gluon, our approach is to consider a more practical approach, tied to the hadronic final state. We therefore define *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).* We note that one still needs to determine the criterion that corresponds to a successful quark enrichment and for that, we have to rely to some degree on a less well-defined notion of what a quark jet is.

In a way, we can see this as using "quark" and "gluon" as adjectives and not as nouns.

Comparisons between different generators in an idealised study

We have systematically tested the performance of quark/gluon tagging predicted by different Monte Carlo generators in an idealised setup. We have considered $e^+e^- \to Z \to u\bar{u}$ as a source of quark jets and $\mu^+\mu^- \to H \to gg$ as a source of gluon jets.

As a discriminating variable, we have studied generalised angularities [4] for which (for $\kappa = 1$) there also exists analytic results at the NLL accuracy:

$$\lambda_\alpha^\kappa = \frac{1}{E_{\text{jet}}^\kappa R^\alpha} \sum_{i \in \text{jet}} E_i^\kappa \Delta R_i^\alpha. \tag{1}$$

The full study includes several working points in the (κ, β) parameter space but we focus here on the IRC-safe "Les Houches Angularity" (LHA) $\lambda_{0.5}^1$.

To quantify the discriminating power, we use the following quantity

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

built from the probability distributions p_q and p_g for the quark and gluon samples as a function of the LHA. In a way, Δ can be seen as a measure of the significance of the difference between the quark and gluon probabilities. It also has the advantage that the integrand can be plotted as a function of λ to see where the discriminating power gets its larger contributions.

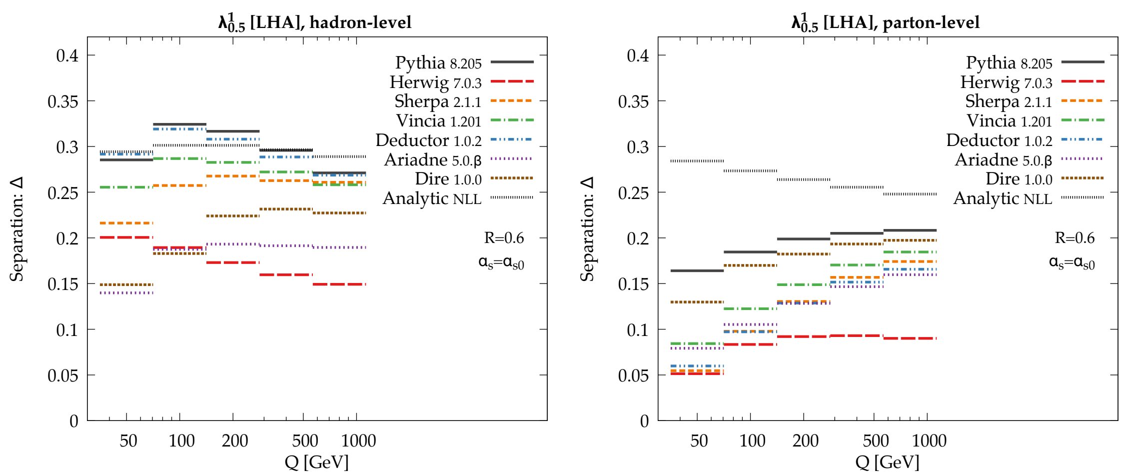

Figure 1: Dependence of the quark/gluon discriminating power (Δ) as a function of the energy (Q) of the collision for different Monte Carlo generators. Left: hadron level; right: parton level.

The quark and gluon processes have been simulated using a series of Monte-Carlo generators. Currently: Pythia 8.205 [5], Herwig 7.0.3 [6], Sherpa 2.1.1 [7], Vincia 1.201 [8], Deductor 1.0.2 [9], Ariadne 5.0.β [10] and Dire 1.0.0 [11]. We also include analytic results at NLL with a modelling of hadronisation effects. The probability distributions $p_q(\lambda)$ and $p_g(\lambda)$, as well as the discriminant Δ can be measured for different angularities, varying fundamental parameters like the energy Q of the collision, the jet radius R and the value used for the strong coupling constant at the Z mass.

In order to pinpoint what ingredients in the generators drive the discriminating power, we have also varied a few chosen internal knobs in each of the generators. See Refs. [2] and [3] for details.

The code used for the analysis is developed in the Rivet 2.4 [12] framework with jet clustering and manipulation done using FastJet 3.1 [13]. It is publicly available in the `lh2015-qg` repository on `GitHub`.

Figure 1 shows an example of our findings: the dependence of Δ on the centre-of-mass energy of the collision. We observe rather large differences between the generators under consideration, both at parton and at hadron level, with Pythia predicting a large discriminating power and Herwig a much smaller one. We also see that non-perturbative effects have a large impact on Δ. Differences are however already substantial at the perturbative level, i.e. in the parton shower. These differences can mostly be traced back to $p_g(\lambda)$ which is currently poorly constrained while $p_q(\lambda)$ is reasonably well constrained e.g. from LEP data. Large differences are also seen for other angularities (both IRC safe and unsafe) and quality measures and call for a better understanding and better constraints on both the perturbative shower and the non-perturbative corrections.

Possible measurements at the LHC

It is natural to wonder if one can perform dedicated measurements at the LHC to help constrain the large differences observed above. A simple option is to measure (generalised) angularity distributions* and the corresponding separation variables for dijet (gluon enriched) and Z+jet (quark enriched) events. Note that we want to report results directly for each processes without making any model-dependent effort to recover “quark” and “gluon” results. In particular, the separation Δ should be computed directly between the Z+jet and dijet distributions.

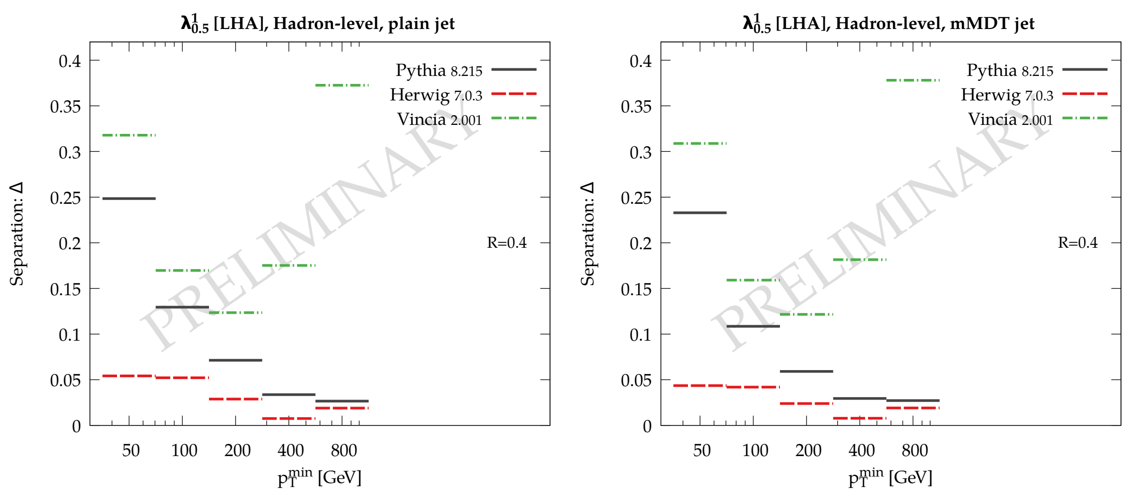

Figure 2: Separation Δ between the LHA measured on jets from Z+jet and dijet events as a function of the cut $p_T^{\min}$ on the jet transverse momentum. Left: angularities are computed using all the constituents of the jet; right: an mMDT procedure is applied before computing the angularities.

As for the ee study-case presented above, there are several parameters that can be varied to further probe the kinematic dependence: the scale dependence can be probed by varying the cut on the

*now defined using the p_t of the jet constituents instead of their energy jet p_t and the angular dependence can be studied either by varying the jet radius, or by measuring the (generalised) angularities on a jet groomed with the (modified) mass-drop procedure (mMDT).

This study — which can also be found on the `lh2015-qg` `GitHub` repository — is still in progress but preliminary results are presented in Figure 2, where we show the separation Δ as a function of the cut $p_T^{\min}$ on the jet transverse momentum. The same patterns as in the previous case are observed with Pythia and Vincia predicting much larger separations than Herwig. these differences remain after applying a mMDT procedure, suggesting that the differences are already present in the description of small-angle physics. Measuring these quantities at the LHC would definitely help to further constrain the Monte Carlo generators.

Possible measurements at an FCC-ee

With a lower hadronic activity, the environment of e^+e^- collisions is far more conducive to precision measurements. Some additional information about quark/gluon discrimination and new constraints on parton-shower generators could possibly already be available from a re-analysis of LEP data with the tools described above, but a new circular collider at a higher energy and with higher statistics would definitely bring in valuable information in many respects.

Since LEP data is already extensively used in Monte Carlo tuning and provide mostly a quark-enriched sample of jets, one observes much smaller differences between generators for the distributions obtained from quark jets than for gluon jets. One should therefore target to build a clean gluon-enriched sample.

A first process of interest is be to look at 3-jet events with 2 b-tagged jets, where the third jet would provide a clean gluon enriched sample. This would largely benefit from the high luminosity and energy coverage expected at a FCC-ee.

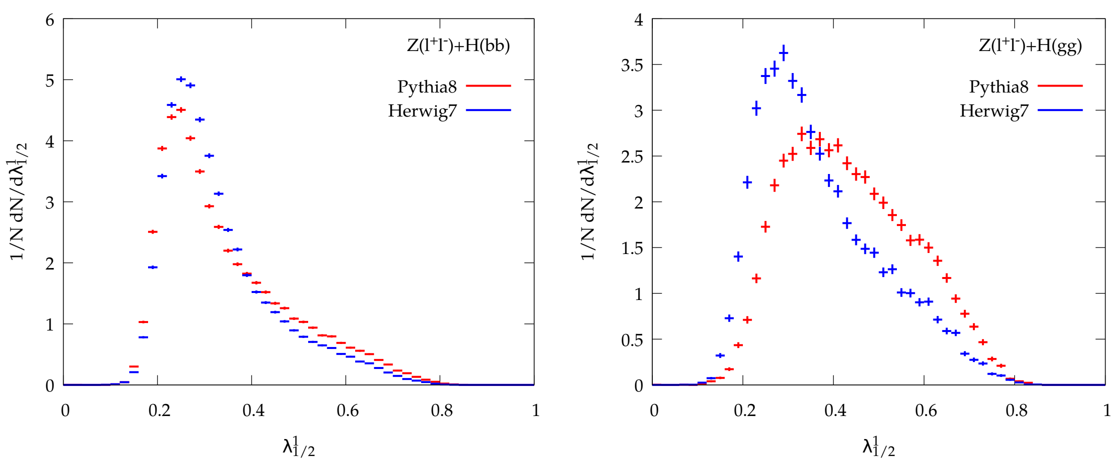

Figure 3: LHA distributions observed from Pythia8 and Herwig7 simulations for $e^+e^- \to Z(\ell^+\ell^-)H(b\bar{b})$ (left) and $e^+e^- \to Z(\ell^+\ell^-)H(gg)$ (right) events. We assume perfect b tagging.

Another process of interest is associated Higgs production at $\sqrt{s} = 240$ GeV, where the Higgs can either decay to a $b\bar{b}$ pair or to gluons. Figure 3 shows the distributions obtained for the LHA after selecting a pair of leptons within 20 GeV of the Z mass and requiring that the two jets are within

15 GeV of the Higgs mass. We clearly see a decent agreement between Pythia and Herwig for the $b\bar{b}$ sample, with much larger differences in the gluon-enriched sample.

The plot includes expected (ideal) statistical uncertainties for an integrated luminosity of 2.5 ab^{-1}, corresponding to 2.1 million HZ events, including about 80000 events with the Z decaying to a e^+e^- or $\mu^+\mu^-$ pair and the Higgs boson decaying to a $b\bar{b}$ pair, and about 12000 events where the Higgs boson decays to a gluon pair instead.

We clearly see how such a measurement, possible at a FCC-ee, would bring crucial information for the development of Monte Carlo event generators. It would also help developing better quark/gluon taggers, a tool of broad application in collider physics.

References

[1] J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. **107** (2011) 172001 [arXiv:1106.3076 [hep-ph]]; JHEP **1304** (2013) 090 [arXiv:1211.7038 [hep-ph]].

[2] J. R. Andersen *et al.*, [arXiv:1605.04692 [hep-ph]].

[3] P. Gras, S. Hoeche, D. Kar, A. Larkoski, L. Lönnblad, S. Plätzer, S. Prestel, A. Siódmok, P. Skands, G. Soyez and J. Thaler, in preparation.

[4] A. J. Larkoski, J. Thaler and W. J. Waalewijn, JHEP **1411** (2014) 129 [arXiv:1408.3122 [hep-ph]].

[5] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605** (2006) 026 [hep-ph/0603175]; T. Sjstrand *et al.*, Comput. Phys. Commun. **191** (2015) 159 [arXiv:1410.3012 [hep-ph]].

[6] M. Bahr *et al.*, Eur. Phys. J. C **58** (2008) 639 doi:10.1140/epjc/s10052-008-0798-9 [arXiv:0803.0883 [hep-ph]]; J. Bellm *et al.*, [arXiv:1310.6877 [hep-ph]]; J. Bellm *et al.*, Eur. Phys. J. C **76** (2016) no.4, 196 [arXiv:1512.01178 [hep-ph]].

[7] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, JHEP **0902** (2009) 007 [arXiv:0811.4622 [hep-ph]].

[8] W. T. Giele, L. Hartgring, D. A. Kosower, E. Laenen, A. J. Larkoski, J. J. Lopez-Villarejo, M. Ritzmann and P. Skands, PoS DIS **2013** (2013) 165 [arXiv:1307.1060 [hep-ph]].

[9] Z. Nagy and D. E. Soper, JHEP **1406** (2014) 097 [arXiv:1401.6364 [hep-ph]].

[10] C. Flensburg, G. Gustafson and L. Lonnblad, JHEP **1108** (2011) 103 [arXiv:1103.4321 [hep-ph]].

[11] S. Hche and S. Prestel, Eur. Phys. J. C **75** (2015) no.9, 461 [arXiv:1506.05057 [hep-ph]].

[12] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz and F. Siegert, Comput. Phys. Commun. **184** (2013) 2803 [arXiv:1003.0694 [hep-ph]].

[13] M. Cacciari and G. P. Salam, Phys. Lett. B **641** (2006) 57 [hep-ph/0512210]; M. Cacciari, G. P. Salam and G. Soyez, Eur. Phys. J. C **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].