Jet Sudakov Safety in pp collisions

Shivesh Mandalia

Supervisors: Phillip Harris and Andre David

October 31, 2015

Abstract

In this paper the Sudakov safety, which is an extension of infrared and collinear safety, of the soft drop parameter z_g has been tested for pp collisions using the `Pythia` MC generator. The Sudakov safe nature of z_g allows one to compute its distributions from a fixed scale. The universality of z_g for jets soft dropped with $\beta = 0$ has also been shown to be true for quark and gluon jets, which means that this distribution is independent of the renormalisation scale. The effect that hadronization has on this scale invariance inside `Pythia` has also been tested. MC data after detector simulation (`QCD MC 13 TeV`) and real data (`QCD Data`) have been subject to similar analysis and it is found that the universality of z_g is held. The application of this parameter is then motivated for the study of the quark-gluon plasma in heavy ion collisions.

1 QCD Basics

Quantum Chromodynamics (QCD) is the theory describing the interactions between quarks and gluons. The six spin-$\frac{1}{2}$ quarks each carry electric and colour charges and colour is exchanged by eight massless gluons. These gluons are themselves charged, which leads to behaviour such as self interactions. Quarks and gluons are not stable particles by themselves and must be bound together to form colour singlets called hadrons. Another distinguishing feature of QCD is its dependence on the energy scale at which the particles are probed. Unlike Quantum Electrodynamics (QED), the running couping constant of QCD **decreases** as the energy scale is increased implying more radiation at lower energy scales. This phenomenon is known as 'asymptotic freedom'. A consequence of asymptotic freedom is that perturbation theory is applicable only at higher energies i.e. shorter distance scales. This means that theoretical calculations of observables become increasing difficult to calculate at lower energies. A consequence of this is that only parts of QCD can be expressed in terms of the fundamental parameters of QCD, whereas other parameters must be expressed by models or functions that are constrained by data. The process of separating the non-perturbative calculations from the perturbative calculations is known as factorization.

1.1 Factorization

Factorization is one of the main tools for calculating observables in QCD. Factorizing allows us to make theoretical predictions of observables such as the cross-section of events by separating long distance behaviour from short distance behaviour. The short distance behaviour can be calculated perturbatively, however the magnitude of the coupling at long distances makes perturbative calculations difficult and instead this behaviour is described by functions that are constrained by experimental data at a given energy scale. Once these have been obtained at a particular energy scale, the function can be evolved to different energy scales using a renormalisation group equation (RGE) known as the DGLAP[1] equation.

Protons are comprised of three valence quarks along with a sea of quarks and gluons, constantly being emitted and destroyed inside the proton. These fluctuations involve relatively small momentum transfers as they are confined to the size of the proton. Hard probes such as high energy photons interacting via. DIS interact over much shorter distances and thus shorter timescales (due to relativistic time dilation). As a result, factorization theorem allows us to separate the cross-section for the non-perturbative parton density function (PDF), which parameterizes the distribution of partons in the proton, with the perturbatively calculable partonic scattering cross-section.

Factorization theorem also has utility after the initial hard scatter. The fragmentation of the quarks and gluons here is factorized into a fragmentation function (FF), which represents the probability of a given parton to fragment into a particular hadron - carrying a certain fraction of the partons energy.

Both the PDF and the FF are functions of the splitting function and so they have similar characteristics. The splitting function is the distribution of x, the fractional difference in energy between each final state daughter of an initial parton and the initial parton (Equation 1).

$$x = \frac{E_{\text{final state daughter}}}{E_{\text{initial parton}}} \tag{1}$$

Figure 1 shows the splitting function for pp collisions. A discrimination has been made for quark initiated jets and gluon initiated jets and from this we can see that the fraction of energy given to the produced particle for the gluon jets is lower when compared to quark jets. Expanding on this, we can draw the conclusion that gluon jets overall have more splittings (vertexes) until the produced particles are final state. This can be explained by examining the modes in which quarks and gluons fragment into. Quarks themselves can only produce a quark and a gluon, however gluons have a probability to either produce a quark/antiquark, a quark and a gluon or two gluons. Gluons have more fragmentation channels so this is the reason we find that gluon jets have more splittings.

2 Jet Physics

In nature, quarks and gluons cannot exist singularly as the strong force forbids coloured particles from being isolated. Colour confinement only allows the existence of particles which have a net zero colour charge (hadrons). In high energy pp DIS events, the struck parton forms hadrons as it becomes displaced from the nucleus. The

[1] DGLAP: Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

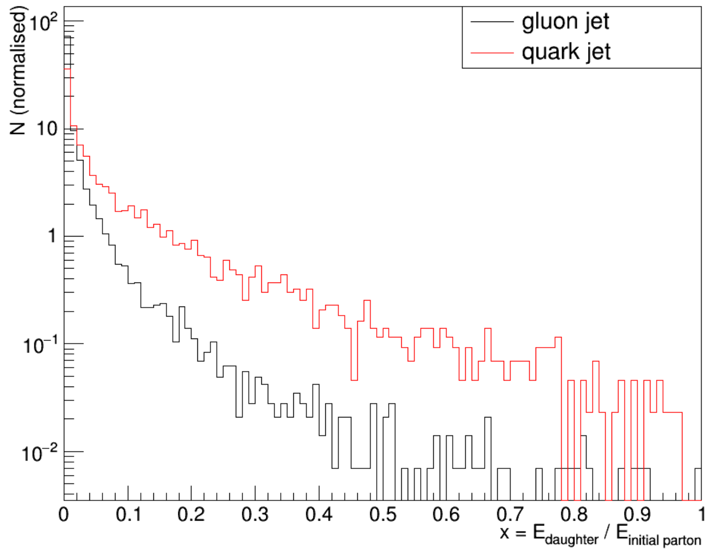

Figure 1: The splitting function for a 13 TeV pp collision. Highest energy quark/gluon with $p_T > 20$ GeV was chosen as the initial parton. Plotted here is the distribution of x, the fractional energy difference between the final state daughters of the initial parton and the initial parton (Equation 1). Histogram has been normalised so that the integral of both distributions are equal.

act of this production of hadrons is called hadronization. The particles created by the hadronization process can fragment themselves until they reach a stable state and as a result a tight cone of hadrons, referred to as a jet, is observed.

2.1 Infrared and Collinear Divergences

Not only does the coupling constant need to be small in order to avoid difficult theoretical calculations in pQCD but the observable must satisfy a further requirement called **Infrared and Collinear Safety** (IRC). An observable is IRC safe if it has:

1. Safety against soft radiation

   Extra emission of soft (low energy) particles should not change the value of the observable

2. Safety against collinear radiation

   Splitting any particle into two comoving particles with the same total four momentum should not change the value of the observable

If these conditions are met, then the long distance non-perturbative corrections will be suppressed and so the observables of interest will lead to finite results. This also means that IRC safe observables guarantee infrared corrections at large distances are small.

2.2 Jet Clustering

Jet clustering is the process by which one takes a collection of final state particles and applies a jet clustering algorithm in order to group particles coming from the same source and to form a tree-like structure of the resulting jet. Since we measure IRC safe observables, for reasons discussed in the previous section, only IRC safe jet clustering algorithms can be compared with theory. This means that the clustering algorithm must be insensitive to arbitrary soft emissions and collinear splittings.

An illustration of the behaviour of an IRC unsafe algorithm is shown in Figure 2. The presence of soft radiation between two jets may cause a merging of the jets that would not occur in the absence of the soft radiation as illustrated by the two leftmost figures. For illustration on the right, the dashed configuration fails to produce a jet because its energy is split among several detector towers whereas the solid configuration produces a jet because its energy is more narrowly distributed.

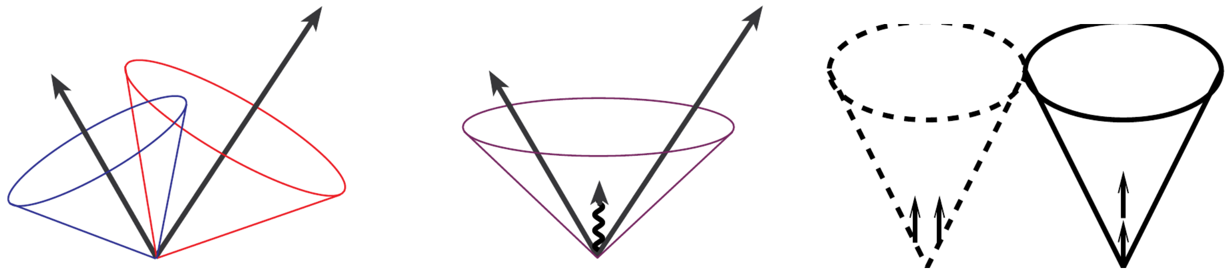

Figure 2: An illustration of soft and collinear sensitivity in jet reconstruction.[1]

The standard examples of IRC safe jet clustering algorithms are the k_T algorithm [2], anti-k_T algorithm [3] and the Cambridge/Aachen (CA) algorithm [4]. In each of these algorithms, a distance metric d_{ij} is calculated between each final state particles and is minimised in order to select which two particles to combine until the smallest distance is from particle i to the beam, d_{iB}, at which point the structure is labelled as a jet. The distance metrics are defined as

$$d_{ij} = \min\left(p_{Ti}^{2\gamma}, p_{Tj}^{2\gamma}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{2}$$

$$d_{iB} = p_{Ti}^{2\gamma} \tag{3}$$

where $\Delta R_{ij}^2 = (\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2$, with $\Delta\eta_{ij}$ being the separation in pseudorapidity, $\Delta\phi_{ij}^2$ the separation in azimuth and R being the jet radius. The parameter γ is dependent on the algorithm chosen, so for the anti-k_T algorithm $\gamma = -1$, the CA algorithm $\gamma = 0$ and the k_T algorithm $\gamma = 1$. For this paper, we will focus on the CA algorithm, which has the advantage that the tree structure closely resembles the structure of QCD bremsstrahlung, which will become useful in the next section when we analyse and apply cuts to the vertexes in this tree.

2.3 Jet Grooming

Once jet clustering has molded the final state particles into a jet with a tree-like structure, jet grooming can be performed. At each vertex in the tree a grooming condition is tested to determine whether or not to modify the vertex. By doing this, wide-angle soft radiation is removed from the jet in order to minimise effects of contamination from initial state radiation (ISR), underlying event (UE) and multiple hadron scattering (pileup). Examples of jet grooming algorithms are Pruning, Trimming, Filtering and the one we will turn our attention to - Soft Drop declustering [5].

The soft drop grooming condition is given by:

$$\frac{\min\left(p_{T1}, p_{T2}\right)}{p_{T1} + p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta} \tag{4}$$

where z_{cut} and β are free parameters, R_0 is the jet radius and R_{ij} is the angular separation. This condition is tested at every vertex in the jet tree. If it evaluates as being true, then the algorithm simply moves onto the next vertex. If it does not satisfy this condition then the lower p_T particle/jet is selected and then discarded.

Figure 3 shows the effect of applying the soft drop groom on the mass of the jet. This plot was created by first selecting the highest energy quark/gluon with a $p_T > 200$ GeV. The final state daughters of this chosen parton were found and they were clustered using the CA algorithm with $R_0 = 0.4$ and a p_T cut of 200 GeV. The jet with the minimum R_{ij} separation between it and the initial parton was found and then this jet was soft dropped using $\beta = 0$ and $z_{cut} = 0.1$. The jet mass minus the effective mass of the initial parton was then added to the histogram as the y-axis value with the p_T of the jet being the x-axis value. Looking at the comparison between the original and the soft dropped jet masses, we can see that the soft drop grooming decreases the mass of the jet, making it closer to the value of the initial parton mass.

Note that in this paper, all plots are made using 13 TeV pp collisions using `PYTHIA` v8.186. The jet clustering and grooming is done using `FastJet` v3.1.0 and `Recursive Tools contrib` 1.017.

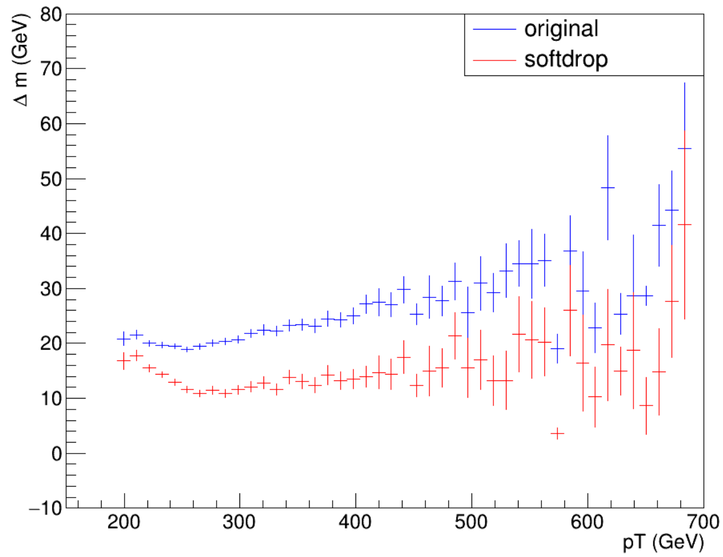

Figure 3: The jet mass with minimum R_{ij} separation to the initial parton minus the effective mass of the initial parton is plotted as a function of the jet p_T. Distribution before (blue) and after the soft drop groom (red) are shown.

3 Sudakov Safety

So far IRC safety has been the guiding principle for determining which observables are calculable in pQCD. However, there exists non-IRC safe observables which can be calculated by using non-perturbative objects such as PDFs, FFs (see Section 1.1) to absorb the singularities. Such observables are called Sudakov safe [6]. In that sense, Sudakov safety is an extension of IRC safety, however this does not mean that all IRC unsafe observables are Sudakov safe. A more formal definition is given as follows. Let observable u be IRC unsafe and choose observable s such that it regulates all singularities of u so that the probability of measuring u given s, $p(u|s)$, is finite at all perturbative orders. To calculate $p(u)$ we can integrate over s:

$$p(u) = \int \mathrm{d}s\, p(s)\, p(u|s) \tag{5}$$

More generally, we can choose a vector of IRC safe observables to regulate u,

$$p(u) = \int \mathrm{d}^n\mathbf{s}\, p(\mathbf{s})\, p(u|\mathbf{s}) \tag{6}$$

An example of a Sudakov safe observable that we shall focus on is the soft drop parameter z_g (see Section 2.3), which shall be regulated using the groomed jet radius r_g.

$$r_g = \frac{R_{12}}{R_0} \qquad\qquad z_g = \frac{\min\left(p_{T1}, p_{T2}\right)}{p_{T1} + p_{T2}} \tag{7}$$

$$p(z_g) = \frac{1}{\sigma}\frac{d\sigma}{dz_g} = \int d\,r_g\, p(r_g)\, p(z_g|r_g) \tag{8}$$

What can be found is that for different modes of the soft drop groom related to the value of β, the safety of the soft drop parameter varies. Table 1 summarises this behaviour. For $\beta < 0$, z_g is IRC safe (and also Sudakov safe), so here z_g has a well defined expression in α_s. When $\beta > 0$, z_g is Sudakov safe but is no longer IRC safe which means the z_g distribution no longer has an analytic dependence on α_s. The most interesting case is when $\beta = 0$. Here, the distribution of z_g is independent of α_s at the first order. There are only collinear divergences present at this particular β value, and these can be absorbed into a unique analytically calculable FF, which has the crucial property that the collinear divergences approximately cancel order-by-order in α_s to all orders. This FF

	Safety	Divergences	Expansion
$\beta < 0$	IRC	None	α_s^n
$\beta = 0$	IRC via.FF	Collinear Only	α_s^{n-1}
$\beta > 0$	Sudakov	Collinear & Soft-Coll.	$\alpha_s^{n/2}$

Table 1: As β is adjusted, $p(z_g)$ interpolates between IRC-safe and two Sudakov-safe behaviours, related to divergences in z_g. Here, $n \geq 1$ ranges over positive integers. [6]

can also be evolved to different renormalisation scales using a RGE (see Section 1.1), however the renormalised FF asymptotes to the original FF as the renormalisation scale is increased and so at this range the renormalised FF becomes independent of the renormalisation scale.

The approximate independence of both higher order effects and renormalisation scale at $\beta = 0$ makes z_g an interesting parameter to study further.

3.1 $P(z_g)$ plots

The properties of z_g have so far been tested in Ref [6] using the `Herwig` MC generator at 13 TeV for quark jets. $P(z_g)$ was plotted for jet p_T values > 50 GeV to > 2 TeV as shown in Figure 4. Note that hadronization effects were turned off and only quark jets were considered in the `Herwig` simulation. As expected, the distribution of z_g is approximately independent of the jet p_T and thus this demonstrates its scale independence.

The procedure used to calculate z_g is laid out in Figure 5. Figure 4 was firstly reproduced using `Pythia` instead of `Heriwg`. Following from this, the requirement of selecting a quark jet was changed to instead choose gluon jet and then for both quark and gluon jets, the hadronization effects in `Pythia` were switched on. These plots are shown in Figures 6 and 7, and ratios of quark/gluon and hadronization on/off are also shown in Figures 8 and 9. It can be seen that for the case where hadronization effects were switched off, the quark/gluon initiated jets have slight differences however they are both consistent with Figure 4. If on the other hand hadronization effects are present, the quark/gluon distributions are no longer completely consistent with Figure 4, or even each other, however the basic shape of $P(z_g)$ is still present. The effect of turning on hadronization breaks the p_T

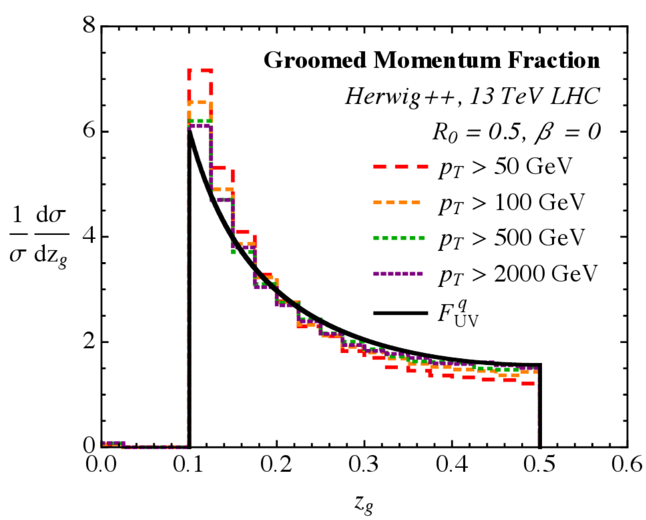

Figure 4: Distributions of z_g for $\beta = 0$ and $z_{\text{cut}} = 0.1$ at the 13 TeV LHC, as simulated by `Herwig++` 2.6.3 with hadronization effects turned off. The p_T of the jets ranges from 50 GeV to 2 TeV, and the asymptotic distribution for quark jets, F^q_{UV} is solid black. [6]

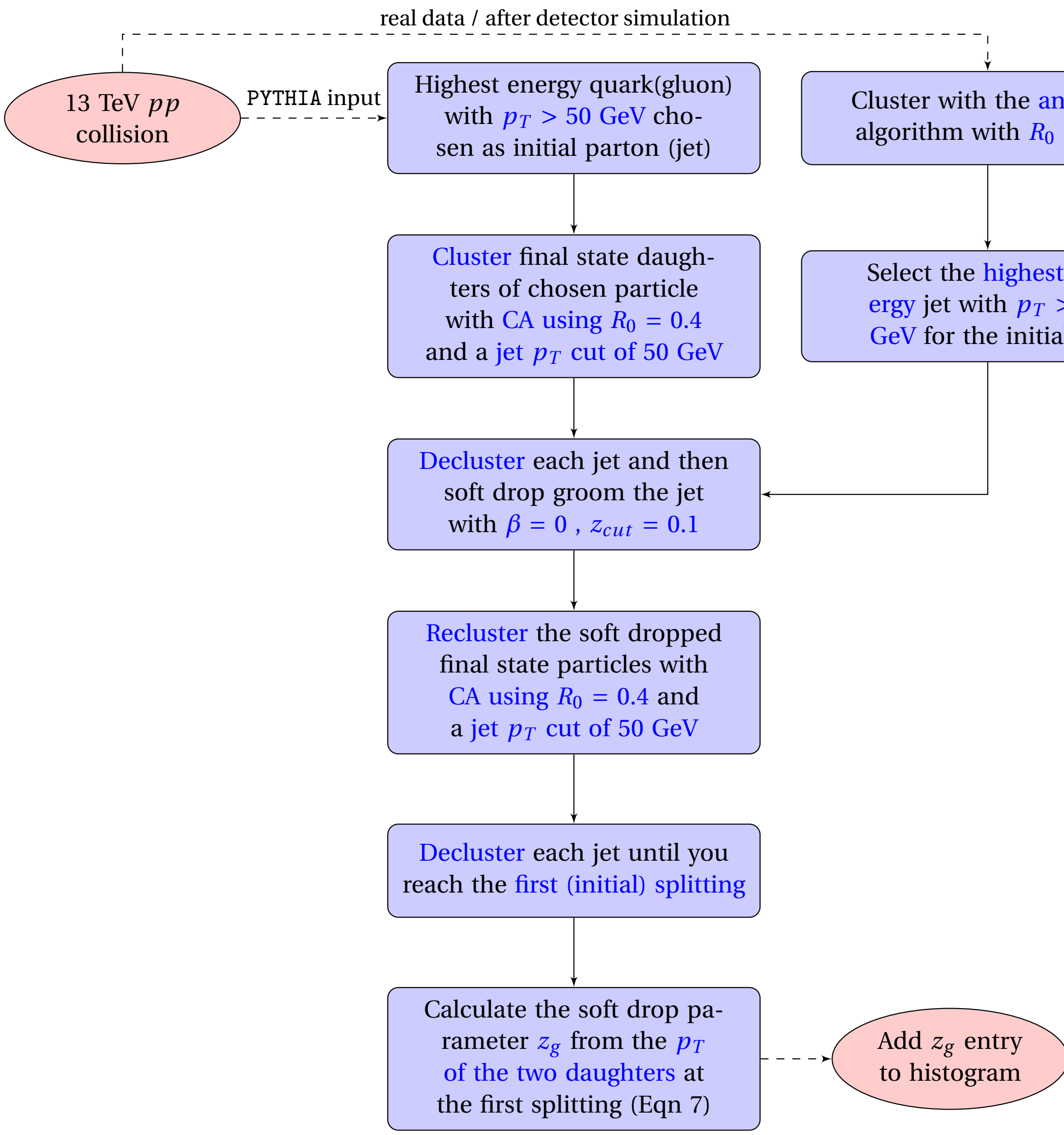

Figure 5: Pseudocode for obtaining the soft drop parameter value z_g for a pp collision.

(a) Quark initiated jet (b) Gluon initiated jet

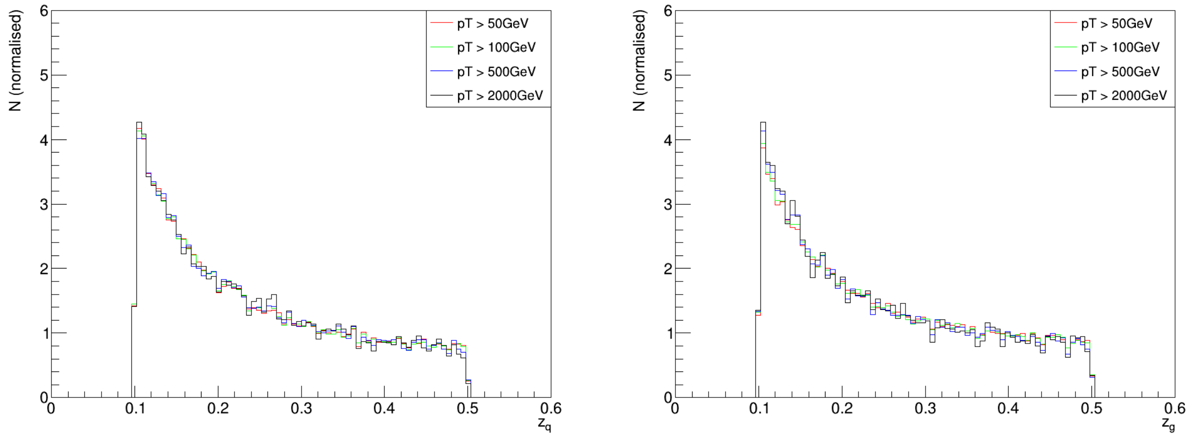

Figure 6: $P(z_g)$ distribution for quark and gluon initiated jets, with hadronization effects turned off in `Pythia`

(a) Quark initiated jet (b) Gluon initiated jet

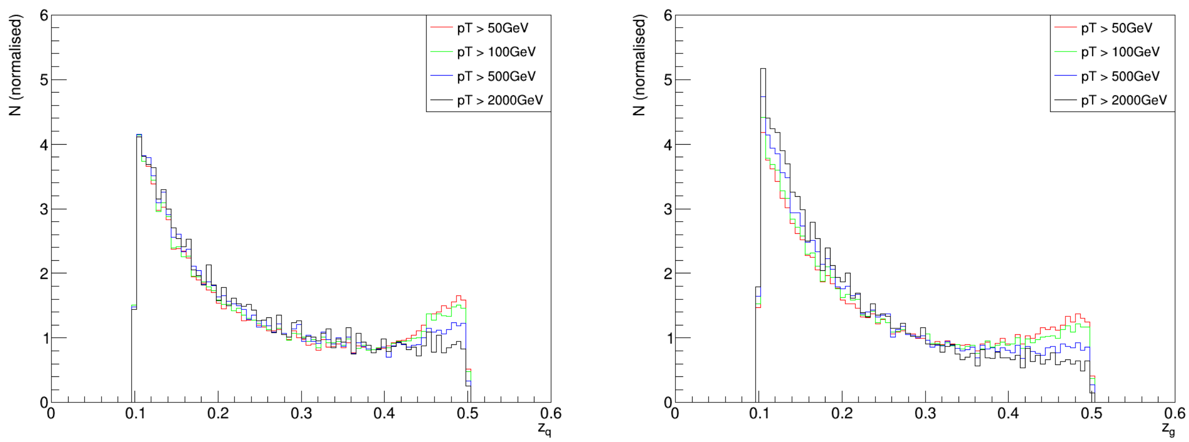

Figure 7: $P(z_g)$ distribution for quark and gluon initiated jets, with hadronization effects turned on in `Pythia`

(a) No hadronization effects (b) With hadronization effects

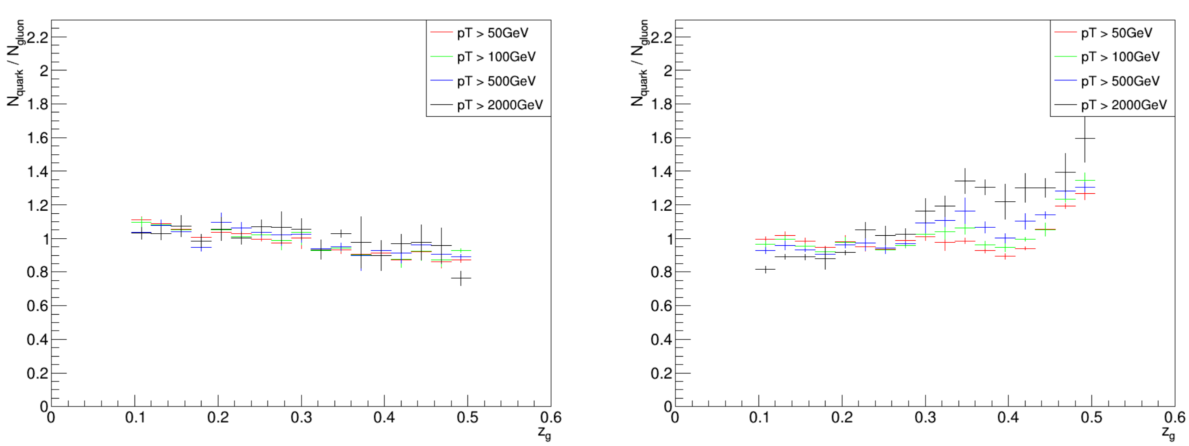

Figure 8: Ratio of $P(z_g)$ for quark and gluon jets, with and without hadronization effects turned on in `Pythia`

(a) Quark initiated jet (b) Gluon initiated jet

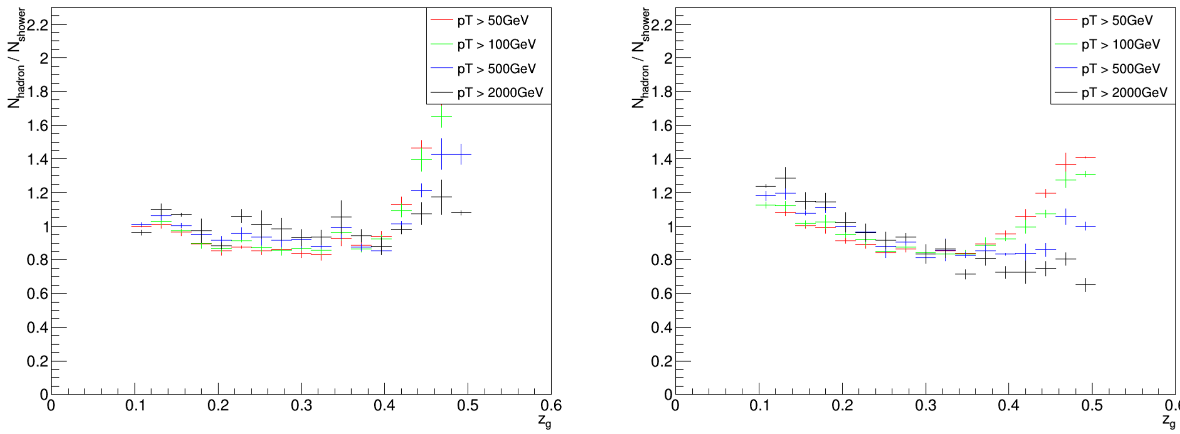

Figure 9: Ratio of $P(z_g)$ with hadronization on/off in `Pythia`, for quark and gluon jets.

(a) No hadronization effects (b) With hadronization effects

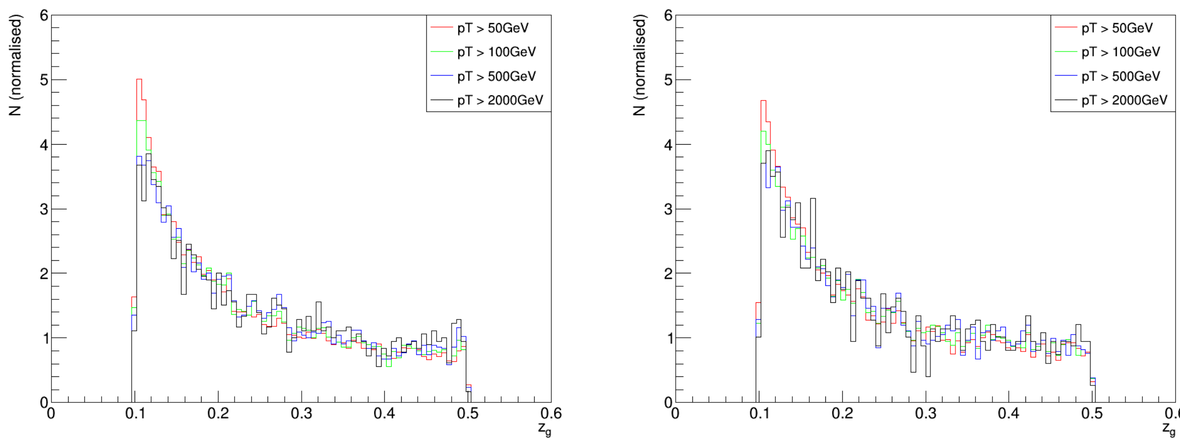

Figure 10: $P(z_g)$ distribution by first using the anti-k_T algorithm to select the initial jet (Figure 5), with hadronization effects turned on in `Pythia`

(a) `QCD MC 13 TeV` (b) `QCD DATA`

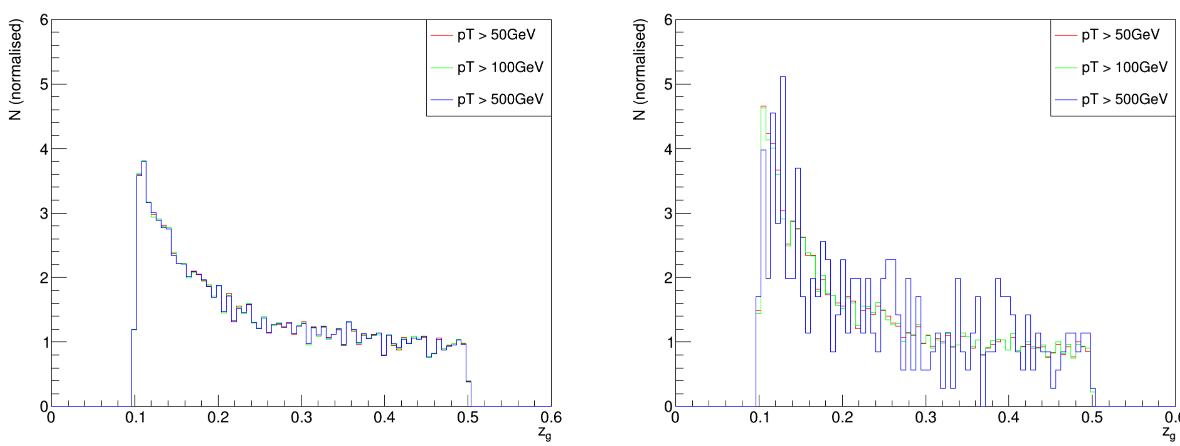

Figure 11: $P(z_g)$ distribution for MC data after detector simulation (`QCD MC 13 TeV`) and for real data (`QCD DATA`)

independence of $P(z_g)$ at high z_g values which can be seen in Figure 7 by the fact that there are relativity more high z_g events at lower p_T's. Beyond being caused by allowing hadronization effects, the underlying reason for these differences is not yet well understood.

Now we move onto using MC data which has been put through a detector simulation. Only the final state particles are detected in the experiment so we can no longer directly select the quark/gluon jet. Instead, the entire collection of final state particles are clustered using the anti-kT algorithm and the highest energy jet with $p_T > 50$ GeV is selected. From there, the same analysis is applied as in the previous case. The full procedure is again illustrated in Figure 5. Before moving onto the analysis using data after detector simulation, to be able to compare before and after detector effects, this procedure was used with final state particles after the `Pythia` simulation. Figure 10 shows the result with and without hadronization effects.

Figure 11a shows the $P(z_g)$ distribution using `QCD 13 TeV MC` data, which is a `Pythia` event sample after detector simulation. Note the $p_T > 2000$ GeV plot was removed as there were no jets that satisfied this criteria. The independence of $P(z_g)$ to the p_T and thus renormalisation scale is again apparent, so now we can move on to using real data. Figure 11b shows $P(z_g)$ for `QCD Data`. The statistics at high p_T is again relatively low, so there are wild fluctuations at high p_T however the general trend is similar to the MC level plots.

4 Heavy Ion Collisions

One interesting application of the soft drop parameter is in parameterizing new states of matter. In heavy ion experiments, such as Pb-Pb collisions, the properties of the "soup" of quarks and gluons (quark-gluon plasma) that forms at extremely high temperatures and densities is not well understood [7]. The independence of $P(z_g)$ to the renormalisation scale for pp collisions means that, whether or not it is independent of the scale, the soft drop parameter will have utility in describing the quark-gluon plasma. If the soft drop parameter does not have a similar distribution to quark/gluon jets then it can be used as a parameter to quantify the quark-gluon plasma. On the other hand, if it does have a similar distribution to quark/gluon jets, then it can be used as a foundation on which to build better theoretical models due to its scale invariance and its similarity with quark/gluon jets.

Using a MC simulation of Pb-Pb collisions at 2760 GeV, the same steps laid out in Figure 5 were applied to the data sample. Figure 12 shows the result of this analysis. At first glance, the plots look similar to that of quark/gluons, however the peak at low z_g is much more pronounced for Pb-Pb collisions. Also, the number of jets with high (> 500 GeV) p_T is much lower when comparing to quark/gluons.

Further analysis would comprise of testing the effect of centrality on $P(z_g)$, which is a measure of the degree to which two nuclei overlap [8]. A low centrality would mean a greater contribution from the quark-gluon plasma as the collision is more direct (head-on). This would allow one to analyse the increasing effect of the quark-gluon plasma by gradually decreasing the centrality value. Also, one could test the effect of choosing to plot the z_g value not the initial vertex, but at the second or third etc. vertex to see if the scale invariance is still present.

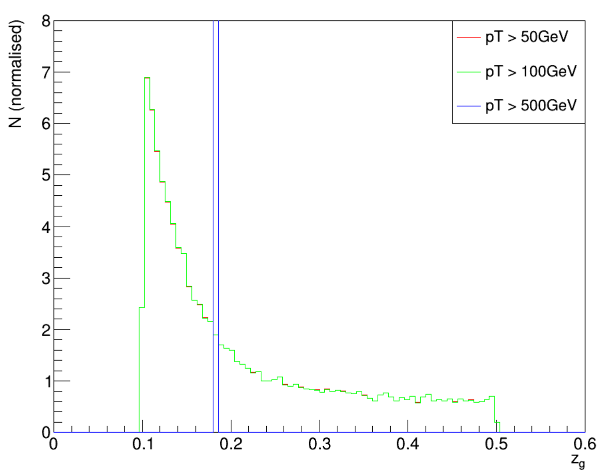

Figure 12: $P(z_g)$ distribution for Pb-Pb MC collisions at 2760 GeV

References

[1] Gerald C. Blazey et al. “Run II jet physics”. *QCD and weak boson physics in Run II. Proceedings, Batavia, USA, March 4-6, June 3-4, November 4-6, 1999*. 2000, pp. 47–77. arXiv: hep-ex/0005012 [hep-ex]. URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-00-092.

[2] S. Catani et al. “New clustering algorithm for multi - jet cross-sections in e+ e- annihilation”. *Phys. Lett.* B269 (1991), pp. 432–438. DOI: 10.1016/0370-2693(91)90196-W.

[3] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The Anti-k(t) jet clustering algorithm”. *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[4] Yuri L. Dokshitzer et al. “Better jet clustering algorithms”. *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[5] Andrew J. Larkoski et al. “Soft Drop”. *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].

[6] Andrew J. Larkoski, Simone Marzani, and Jesse Thaler. “Sudakov Safety in Perturbative QCD”. *Phys. Rev.* D91.11 (2015), p. 111501. DOI: 10.1103/PhysRevD.91.111501. arXiv: 1502.01719 [hep-ph].

[7] Xiaofeng Luo et al. “Volume fluctuation and auto-correlation effects in the moment analysis of net-proton multiplicity distributions in heavy-ion collisions”. *Journal of Physics G: Nuclear and Particle Physics* 40.10 (2013), p. 105104. URL: http://stacks.iop.org/0954-3899/40/i=10/a=105104.

[8] Betty Abelev et al. “Centrality determination of Pb-Pb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV with ALICE”. *Phys. Rev.* C88.4 (2013), p. 044909. DOI: 10.1103/PhysRevC.88.044909. arXiv: 1301.4361 [nucl-ex].