Study of the discrepancy between $t\bar{t}$+heavy flavours background samples produced using Matrix Element and Parton Shower calculations

Migle Stankaityte

Supervisors: Dr. Judith Katzy and Dr. John Stakely Keller

Deutsches Elektronen-Synchrotron DESY, Hamburg

School of Physics and Astronomy
University of Birmingham
Birmingham, B15 2TT

September 7, 2016

Abstract

A discrepancy was found in tt+heavy flavour background samples produced for investigation of ttH production. An excess of events was found in the region of low (< 0.35) c and b hadron p_T to jet p_T fraction for samples generated using parton shower (PS) calculations for $c\bar{c}$ and $b\bar{b}$ pairs, whereas it was not seen in samples where the heavy flavour pair production is done using matrix element (ME) calculations. The excess was more prominent in the $t\bar{t}c\bar{c}$ sample and for jets containing only one c or b hadron thus the research focused on these cases. It was found that the excess is not due to a kinematic sculpting effect or due to a hadronisation problem in the PS calculations. It was also seen that c quarks of non-gluonic origin were contributing to the excess in the Powheg+Pythia sample. Lastly, observations were made supporting the hypothesis that low energy c hadrons originate from quarks produced in soft gluon splitting.

Contents

1 Introduction

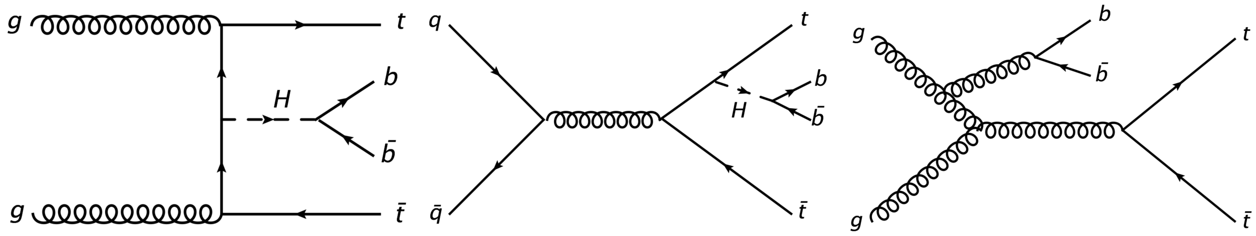
Since the discovery of the Higgs Boson at the Large Hadron Colider (LHC) in July 2012 it has been found in various decay channels. However the research in many areas associated with the Higgs boson is still ongoing. One of the most interesting of these areas is the production of the Higgs boson in association with a pair of top quarks ($t\bar{t}H$) and subsequent $H \to b\bar{b}$ decay as study of these channels would allow us to directly measure the Yukawa coupling of the top quark (y_t)[1].

In order to better differentiate the signal from the background, accurate Monte Carlo (MC) simulations of the background are needed. In recent ATLAS investigations background simulations were produced using next-to-leading order (NLO) matrix element (ME) calculations for the $t\bar{t}$ quark pair production and parton shower (PS) calculations for the $b\bar{b}$ or $c\bar{c}$ quark pair production[2]. Due to certain limitations in the PS calculations it is considered that using ME calculations for the b or c quark production might be more appropriate[2]. However, differences between the samples produced using ME and PS methods have been observed in the region of low hadron-jet transverse momentum (p_T) fraction and need to be understood. This project investigates the possible reasons for these discrepancies.

2 Theory and Background

2.1 $t\bar{t}c\bar{c}$ and $t\bar{t}b\bar{b}$ Background

The research done during this project aims to aid the search of the Higgs boson produced in association with a pair of top quarks and subsequently decaying into two bottom quarks. Two examples of the described process can be seen in figures 1a and 1b. The $t\bar{t}H$ production mode allows direct measurement of the top quark Yukawa coupling (y_t) as opposed to other production and decay processes like gluon fusion production (ggH) or diphoton decay ($H \rightarrow \gamma\gamma$) modes, where the coupling can only be measured through loop effects[3].



(a) (b) (c)

Figure 1: Examples of (a) and (b) $t\bar{t}H$ production with subsequent $H \rightarrow b\bar{b}$ decay, and (c) top quark production in association with b-jets[1].

The background is mainly dominated by the production of a top quark pair in association with heavy flavour quarks[1]. An example can be seen in figure 1c. The two background signals investigated in this report are $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$. The $t\bar{t}b\bar{b}$ background has the same final state as the interaction of interest, whereas the $t\bar{t}c\bar{c}$ background is significant as c-jets can be misidentified as b-jets.

2.2 Monte Carlo Event Generators

In order to better distinguish the $t\bar{t}H \rightarrow t\bar{t}b\bar{b}$ signal from the $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ background Monte Carlo (MC) simulations of these backgrounds must be performed.

A MC generator uses pseudo-random numbers in order to simulate various particle physics phenomena. The simulation starts with a high energy hard scattering process[4][5]. In this case it is a quark anti-quark pair produced by a pair of gluons (figure 1a) or a light quark anti-quark pair (figure 1b) from the colliding protons. As the QCD coupling (α_s) is weak at high energies[6] the partons can be assumed to be free particles and perturbative ME calculations can be used[7]. Depending on the level of accuracy desired, perturbation theory is typically applied to leading order (LO) or next-to-leading order (NLO). After the hard scattering process the gluons can radiate further producing parton showers. These parton showers represent higher-order corrections to the hard scattering process and require non perturbative calculations, which are called parton shower (PS) calculations. In addition there are remnants of the two initial beam hadrons and soft initial and final state radiation present[8]. These processes are called the *underlying event* for which a separate calculation process is used, which is not discussed in detail in this report. An illustration of a MC simulation of a proton-proton collision can be seen in figure 2.

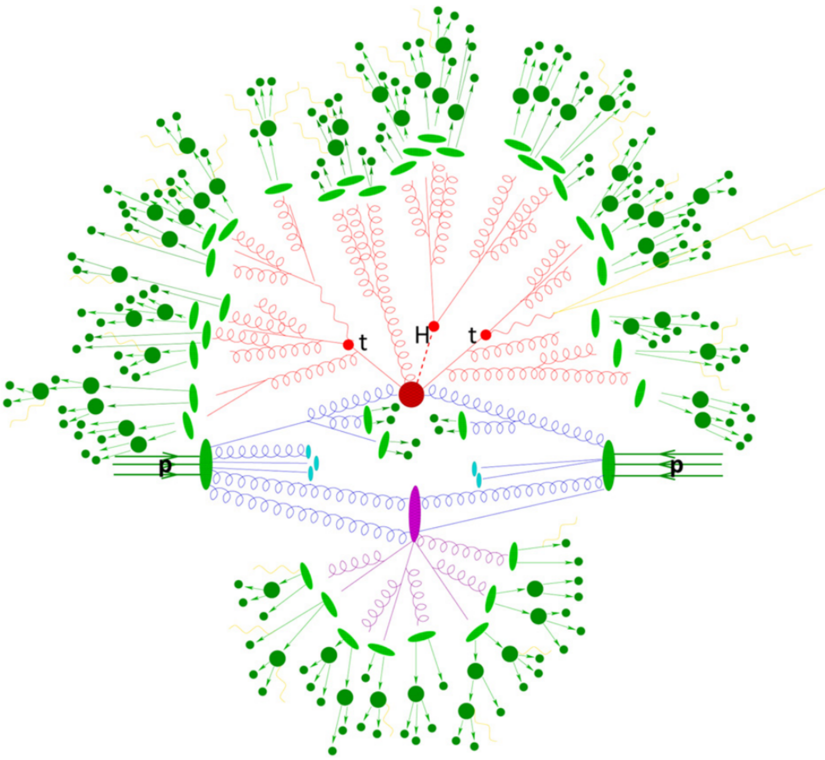


Figure 2: Illustration of a MC simulation of a proton-proton collision. ME calculations are used for the hard scattering process (red circles). Non-perturbative calculations are applied for parton showers (red), the hadronisation of final state partons (light green), and decaying hadrons (dark green). The underlying event includes the secondary interaction (purple) and the proton remnants shown (blue)[8]. *Source: ©with kind permissions from Springer Science + Business Media.*

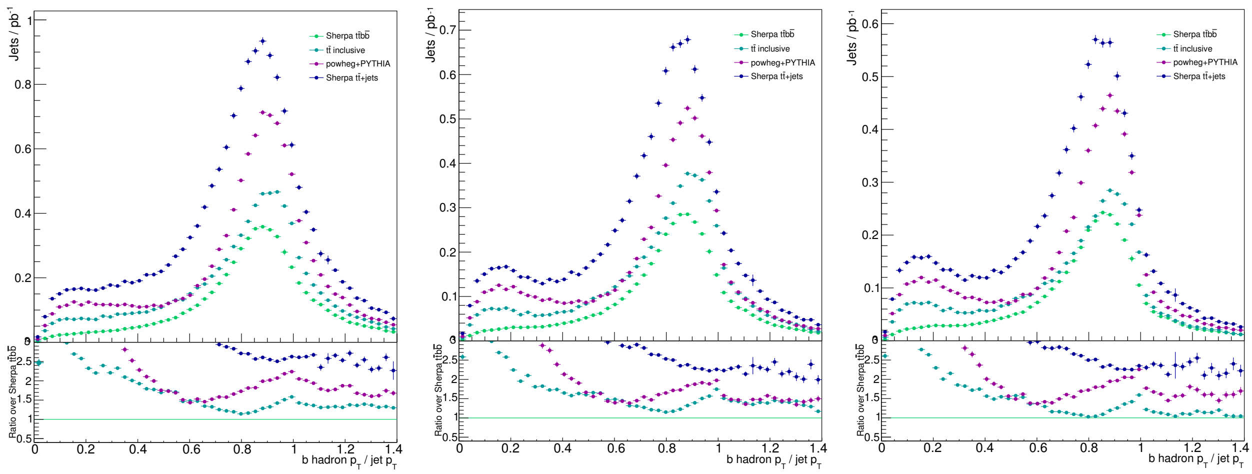
In the case of $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ background two methods of simulation are considered. In both cases the top quark pair production is calculated using ME calculations. Then the $b\bar{b}$ or $c\bar{c}$ production can either be simulated using PS or ME calculations. Due to the nature of the calculation methods the former approach would better describe softer close by quarks, whereas the latter should be better for quarks with higher transverse momentum and angular separation[2].

In this project the $t\bar{t}c\bar{c}$ events generated using solely ME calculations were produced with MadGraph5_aMC@NLO which combines the MC event generator with next-to-leading order (NLO) calculations[9]. The renormalisation and factorisation scale was set to $H_T/4$ where H_T corresponds to the scalar sum of the jet momentum in the event[2]. These events are referred to as the $t\bar{t}c\bar{c}$ sample. The $t\bar{t}b\bar{b}$ events generated with solely ME calculations were generated using the Sherpa v2.1.1 generator with the MEPS@NLO setup[2] and are referred to as the Sherpa $t\bar{t}b\bar{b}$ sample. Event samples generated using the ME only for $t\bar{t}$ generation and the PS for $b\bar{b}$ and $c\bar{c}$ generation were produced using three set-ups: Herwig++ and MadGraph5_aMC@NLO - referred to as the $t\bar{t}$ inclusive sample; Sherpa and MEPS@NLO - referred to as the Sherpa $t\bar{t}$+jets sample; and using the PYTHIA event generator with the POWHEG BOX framework for implementing NLO calculations - referred to as the Powheg+Pythia sample. The detailed generator configurations used for event production can be found in the *Studies of tt+cc production with* MadGraph5_aMC@NLO *and* Herwig++ *for the ATLAS experiment*[2] ATLAS note.

2.3 Discrepancy of Samples using ME and PS for Heavy Flavour Simulation

As mentioned in section 2.2 two calculation methods are considered for the $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ background simulation - using ME or PS to simulate $b\bar{b}$ and $c\bar{c}$ production. When considering substituting one for the other it is important to compare the two methods and understand what is possibly lacking in one compared to the other.

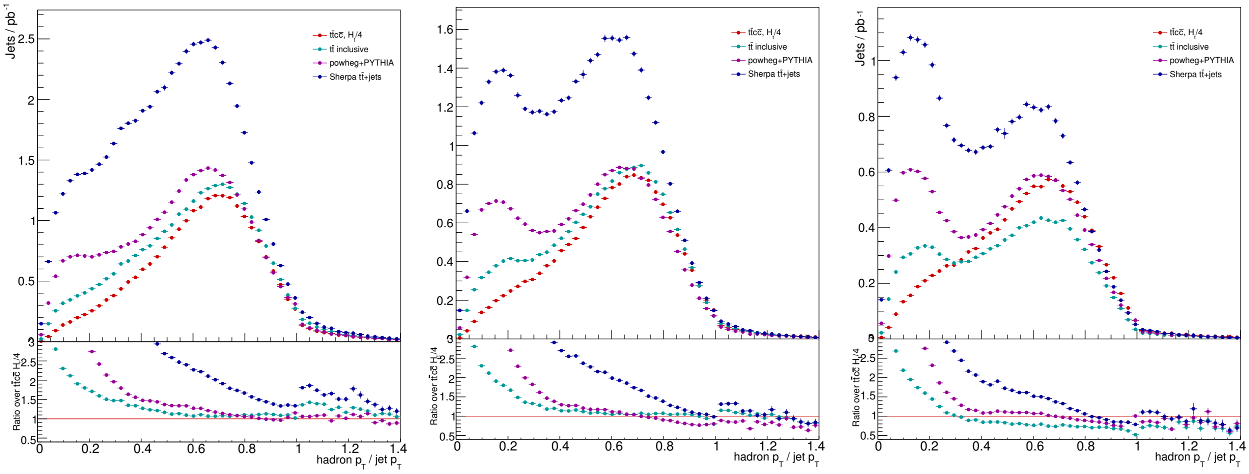
In this project a discrepancy found in a previous analysis is investigated (refer to *Studies of tt+cc production with* MadGraph5_aMC@NLO *and* Herwig++ *for the ATLAS experiment*[2] ATLAS note). The samples produced using PS calculations for the heavy flavour production have been found to have an excess of events in the region of low (< 0.35) hadron p_T to overall jet p_T ratio when compared to the the samples generated using ME. This appears to be the case for both the $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ backgrounds as seen in figures 3 and 4.



(a) (b) (c)

Figure 3: The bottom hadron p_T to jet p_T fraction with (a) 15GeV jet p_T cut-off for all jets, (b) 25GeV jet p_T cut-off for all jets and (c) 25GeV jet p_T cut-off for jets with 1 bottom hadron in the jet.

In both cases the excess is more prominent in the samples with a larger jet p_T cut-off as can be seen by comparing figure 3a with figure 3b, and figure 4a with figure 4b. This implies that in samples using PS calculations there are more events of high energy jets with low energy corresponding hadrons. Furthermore, the problem is a lot more prominent in jets with only one associated c or b hadron respectively as seen when comparing figure 3b with 3c, and 4b with 4c. It is also noticeable that the excess is comparably larger in the Sherpa $t\bar{t}$+jets and the Powheg+Pythia samples than the $t\bar{t}$ inclusive sample.



(a) (b) (c)

Figure 4: The charm hadron p_T to jet p_T fraction with (a) 15GeV jet p_T cut-off for all jets, (b) 25GeV jet p_T cut-off for all jets and (c) 25GeV jet p_T cut-off for jets with 1 charm hadron in the jet.

Three hypotheses were made in accordance with the observed behaviour. Firstly, there is a possibility that the excess is not due to any physical process and is due to the cuts made on the samples. As the excess is only seen in the samples with PS calculations it is also possible that there is a problem in the hadronisation process and high p_T quarks are producing low p_T hadrons initiating the jets. Another possibility is that the hadrons causing the bump are produced by soft gluon splitting into two b or c quarks producing hadrons which do not initiate the jet, but are close to a jet which is consequently misidentified as a b or a c jet. As the quarks produced by a gluon should in theory be largely separated one quark is either lost or is associated with a different jet which would explain why the excess is so prominent in jets associated with one b or c hadron.

As the discrepancy is apparent in both $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ background samples and is more prominent $t\bar{t}c\bar{c}$ samples this project focuses the investigations on $t\bar{t}c\bar{c}$ samples.

3 Experimental Procedure and Results

The project began with replication of the results of the previous investigation in which the discrepancy was found. This was done in order to make sure that a correct starting point was established and further investigations were in accordance with previous results. The results were successfully replicated, however this report does not go in to further detail on the matter.

3.1 Event Selection and n-tuple Generation

In order to test the hypothesis made, investigations of the origin of the c hadrons and subsequently the c quarks are needed. Therefore appropriate event information must be selected and stored.

The initial n-tuples provided for previous result replication included the p_T, pseudorapidity η and polar angle ϕ of the top quarks, jets, hadrons in the jets (and their PDG ID) and final state electrons and muons. The new n-tuples were expanded such that the information on charm quarks in the event and their mothering particles was stored as well. However the electron and muon information in the starting n-tuples was purely needed to impose fiducial cuts in the analysis. As these cuts were not applied in this investigation the electron and muon information was not stored.

Firstly all stored jets had a pseudo rapidity cut of $|\eta| < 2.5$ as these are typical cuts made for $t\bar{t}$ production[2].

Certain checks had to be implemented during the selection process so as to avoid double counting of particles. Top quarks tend to radiate a photon or a gluon before decay. Some simulation software stores a particle that has radiated as a new particle which is mothered by the old one. In order to make an appropriate comparison of the different generators only the final top quark (after the radiation) was stored. Similarly charm quarks can radiate before the decay, however as only the origin of these quarks is of interest only the initial charm quark information (before radiation) was stored. Furthermore, charm quarks can be produced in the decays of bottom quarks. As the jet initiating quarks are of interest, these c quarks were removed as well. Likewise c hadrons mothered by other c or b hadrons were removed.

Another check applied during the n-tuple creation process was calculating the number of c quarks and comparing it to the number of c hadrons in the event. The number of quarks in the event (N_c+$N_{\bar{c}}$) should match the total number of c hadrons ($N_{\text{all } c \text{ hadrons}}$) plus an extra count for every charmonium ($N_{\text{charmonium}}$) in the event such that

$$N_c + N_{\bar{c}} = N_{\text{all } c \text{ hadrons}} + N_{\text{charmonium}}. \tag{1}$$

If this condition was met, it is implied that the correct number of quarks was stored. The implication of said condition uncovered a problem in Powheg+Pythia event generation. It was found that if charm quarks mothered by W bosons radiate, the mother stored in the event log was not a charm quark but a W boson. This needed to be accounted for and thus only the information on the c quark with the highest p_T was stored.

As samples produced with Sherpa have very different event data generation and storage to other generators implying the counting checks described in this section proved complicated and it was decided to not use Sherpa samples for the remainder of the investigation.

Lastly, as this analysis investigates the additional heavy flavour quark jets certain selection criteria had to be applied at the analysis level. Top quarks almost always decay to Wb, thus the jets initiated by b quarks coming from top decays were removed. Jets initiated by c quarks mothered by W bosons from top decay were removed as well.

3.2 Experimental Method

Testing of all hypotheses required matching quarks to the hadrons of jets. Most of the investigation was done with jets that are matched to a single c hadron as the discrepancy is most pronounced for these type of jets. A c quark is matched to a c hadron if

$$\Delta R = \Delta\eta \oplus \Delta\phi < 0.4 \tag{2}$$

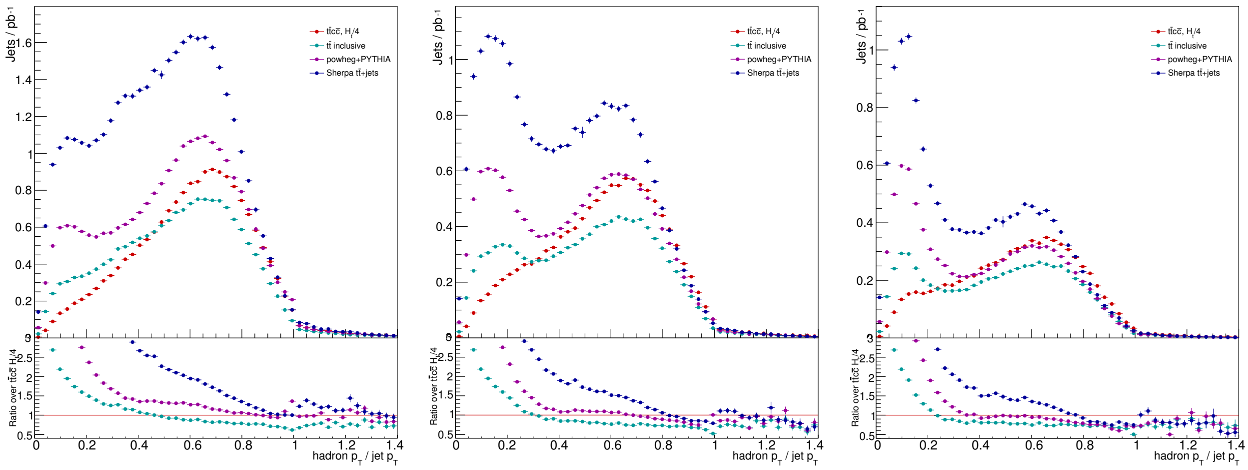
where ΔR is a quantity describing angular separation between two objects, $\Delta\eta$ is the difference in pseudorapidities and $\Delta\phi$ is the difference in azimuthal angle of the two objects, which in this case are the quark and the hadron.

As the discrepancy is most significant for jets associated to one c hadron, the investigation focuses on these type of jets. Also, the results are divided into two c hadron p_T fraction regions: low (p_T fraction < 0.35) which contains the excess, and high (p_T fraction > 0.35) where all samples appear to be overall consistent.

3.3 Results

3.3.1 Hypothesis: Kinematic Sculpting Effect

The first hypothesis made suggests that the excess is a kinematic sculpting effect, meaning that the excess is due to the cuts made on the sample. This can be tested by observing the excess with different minimum jet p_T cuts on jets with a single hadron as displayed in figure 5.



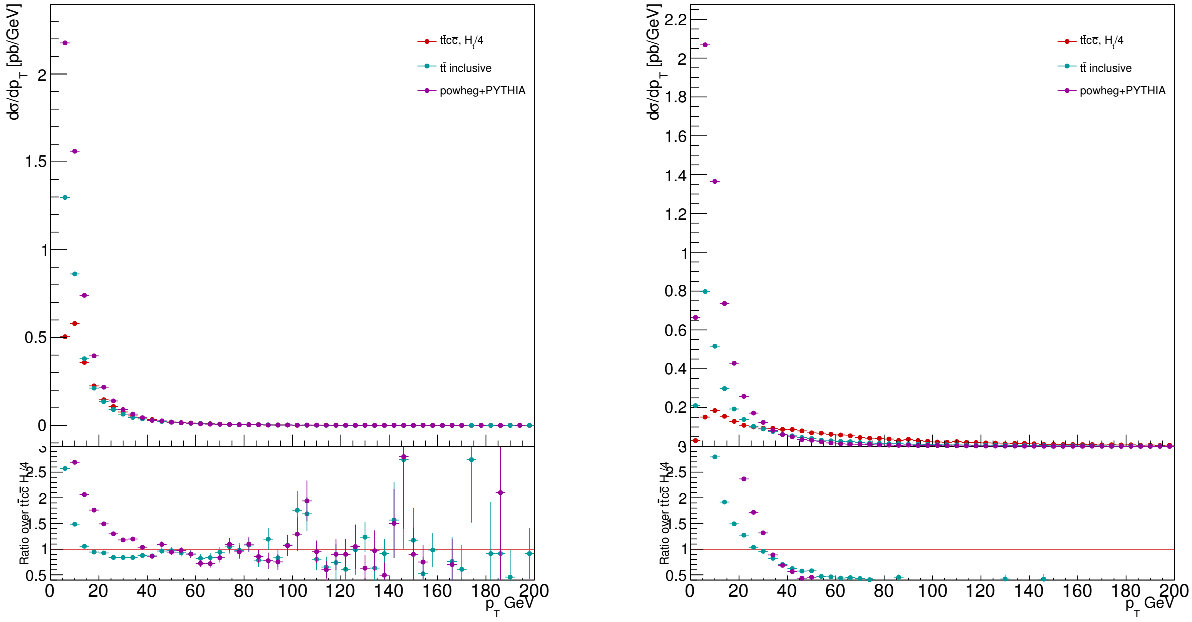
(a) (b) (c)

Figure 5: The charm hadron p_T to jet p_T fraction with (a) 15GeV (b) 25GeV and (c) 40GeV jet p_T cut-off for jets with one c hadron.

It can be seen in figure 5 that the number of events in the excess is not affected by any of the made cuts and stays the same in the samples with 15GeV, 25GeV and 40GeV jet p_T cut-offs. This means that the hypothesis that the discrepancy is due to cuts made on the samples is incorrect and can be rejected.

3.3.2 Hypothesis: Hadronisation Problem

One of the hypothesis made suggests that as the excess is only seen in the samples where additional jets are simulated using PS calculations there might be a problem in the hadronisation process such that high energy quarks initiating the jets are producing low energy c hadrons.



(a) (b)

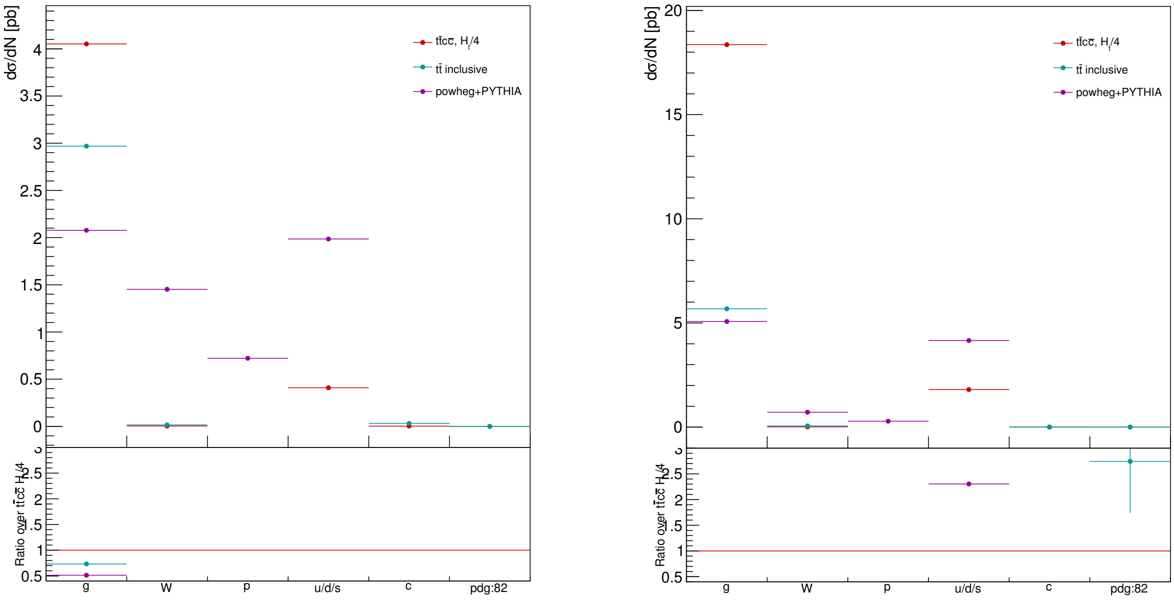
Figure 6: (a) c hadron and (b) c quark p_T in the low hadron p_T fraction region (< 0.35) for jets associated with one hadron with a 25GeV jet p_T cut-off.

Figure 6 shows that both the c hadrons (figure 6a) associated to jets with one c hadron and the charm quarks matched to these hadrons (figure 6b) in the low c hadron p_T fraction region have comparably low p_T. Thus it is clear that there is no problem in the hadronisation process and the low p_T c hadrons are produced by low p_T charm quarks. This observation also agrees with the second hypothesis stating that the excess is due to soft gluons from other jets splitting into low p_T charm quarks which in turn produce c hadrons with low p_T.

3.3.3 Quark Origin Investigation

To get a better understanding of the possible underlying processes related to the c hadrons causing the excess the mother PDG IDs of the charm quarks associated to c hadrons of jets with one c hadron were observed. The findings are shown in figure 7 and the precise percentages of mothers can be found in table 1 in appendix A.1.

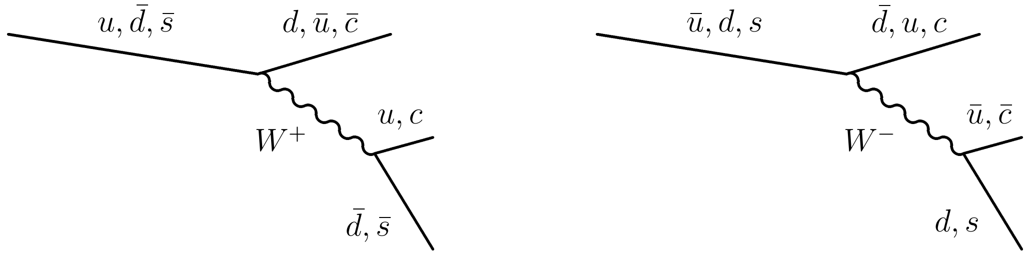
Both in $t\bar{t}c\bar{c}$ and $t\bar{t}$ inclusive samples a large majority of the quarks originate from gluons, whereas in the Powheg+Pythia sample only a third of the charm quarks in the excess region (low p_T fraction) and half of the charm quarks in the non-excess region (high p_T fraction) are mothered by gluons, with the rest coming from different origins such as W bosons, protons and u, d and s quarks.



(a) (b)

Figure 7: The PDG IDs of mothers of c quarks associated to the hadron of jets with one c hadron in the (a) low (< 0.35) and (b) high (> 0.35) c hadron p_T fraction region with 25GeV jet p_T cut-off.

Theoretically the charm quarks coming from W bosons and $u/d/s$ quarks can be produced through weak processes depicted in figure 8. Manual checks were performed at n-tuple generation level. It was found that the mothers of W bosons mothering c quarks (excluding W bosons coming from top quarks) appear to always be either u, d, s quarks or gluons. The u, d, s quarks as mothers agree with the theoretical expectation. Finding gluons to be W boson mothers is likely due to some peculiarity in the event record storage and possibly requires further research not concluded in this project.



(a) (b)

Figure 8: Feynman diagrams of the weak processes assumed to be the origin of c quarks mothered by W bosons and $u/d/s$ quarks.

In order to check if charm quarks coming from $u/d/s$ quarks were produced in the weak process depicted in figure 8 it was required to check if the $u/d/s$ quarks mothering the charm quarks had any daughters that were W bosons. Due to the storage space saving methods implemented in the event record storage the information on some of the $u/d/s$ quark daughters was lost. The recovered daughter information suggests that some of the $u/d/s$ quark daughters are indeed W bosons, however it cannot be stated with complete certainty if this is the case for all of the quarks. Overall, the observations at n-tuple level appear to agree with the theoretical assumption of c quarks produced through weak processes.

The contribution to the excess of c hadrons produced by c quarks from different origins was observed. This was done by looking at the c hadron p_T fraction of c hadrons associated with quarks exclusively from each origin. The findings can be seen in figure 9.

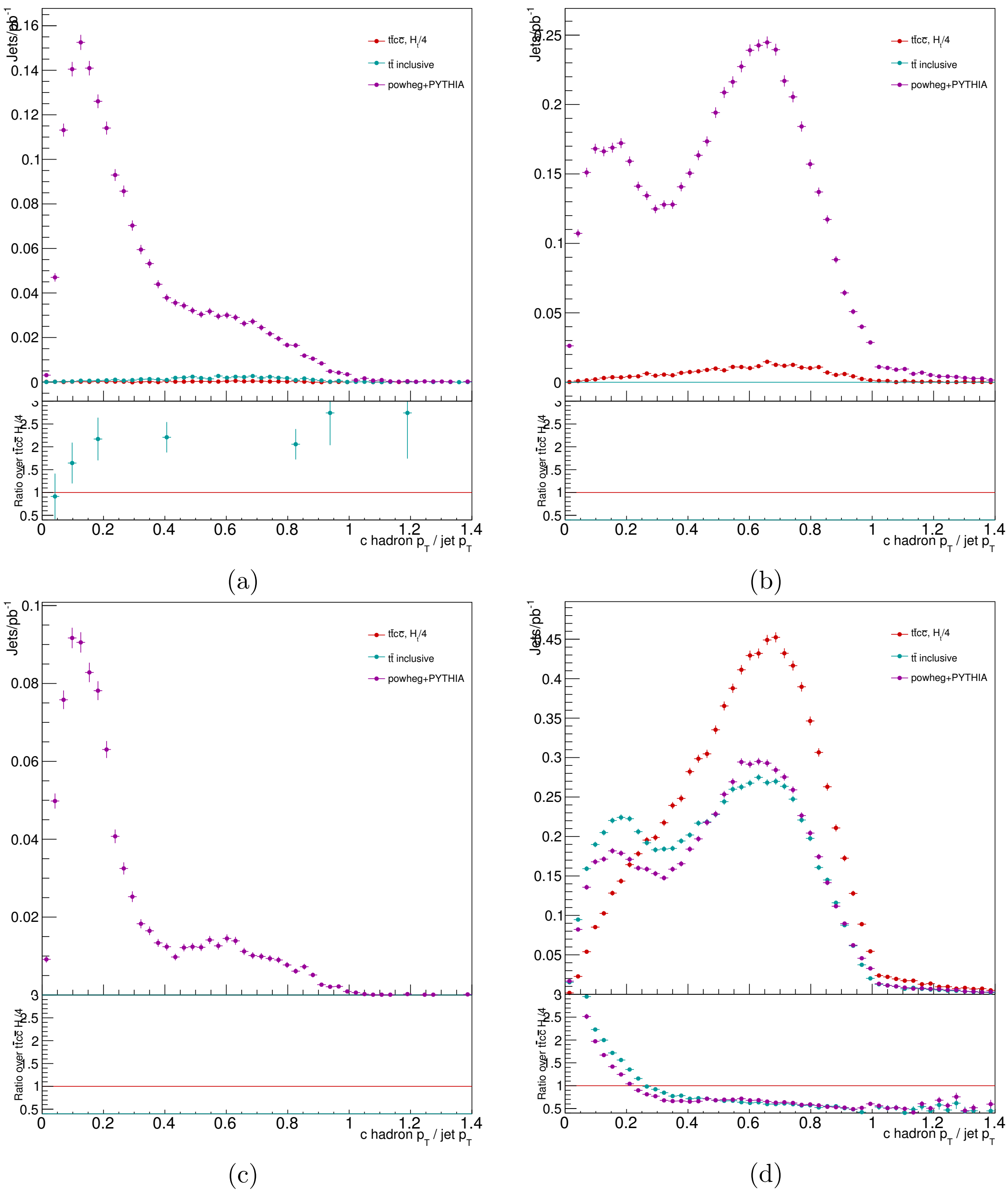


Figure 9: c hadron p_T fraction for 25GeV jet p_T cut-off of jets with one c hadron, where all matched c quarks are only mothered by (a) W bosons, (b) $u/d/s$ quarks, (c) protons and (d) gluons.

In figures 9a, 9b and 9c it is apparent that quarks from non-gluon origins in the Powheg+Pythia sample contribute to the excess. Figure 9d shows that without c hadrons from quarks of non-gluon origins the Powheg+Pythia excess becomes similar to that of $t\bar{t}$ inclusive sample, whereas without this selection it was a lot more prominent than in the $t\bar{t}$ inclusive sample (refer to section 2.3). Thus this explains why Powheg+Pythia samples had a more significant excess. In addition, as the weak processes would not create b quarks this also possibly explains why the excess is less prominent in the $t\bar{t}b\bar{b}$ samples than $t\bar{t}c\bar{c}$ samples, however further research should be done to verify this.

It is reasonable that c quarks originating from weak processes will have a lower p_T, as the quarks are not coming from the initial hard scattering process. However the p_T of all of the protons was stored in the event record as 0, thus it is not possible to easily determine the origin of the protons or determine the reason as to why it is contributing to the excess. This was not investigated in this report and requires further research.

3.3.4 Hypothesis: Soft Gluon Splitting

As seen in figure 9d the discrepancy is still apparent when only c hadrons associated with c quarks mothered by gluons are selected. This gives reason to think that the discrepancy is related to gluons. Thus the hypothesis that the excess is due to soft gluons splitting into quarks which in turn form low p_T c hadrons was investigated.

Firstly, as the excess is the most apparent in jets with one c hadron it was important to test that there is no generator fault which results in a gluon producing a single charm quark. It can be seen in figure 10 that all events where a gluon is mothering a c quark associated to the c hadron of a jet with a single c hadron have at least two charm quarks. This indeed suggests that there is no generator fault.

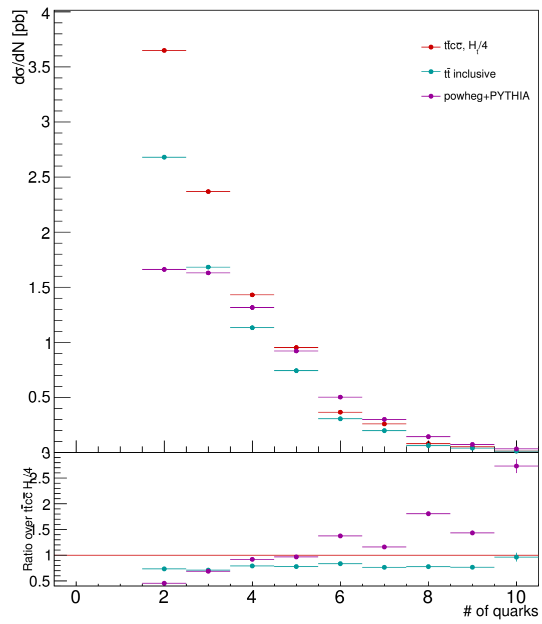
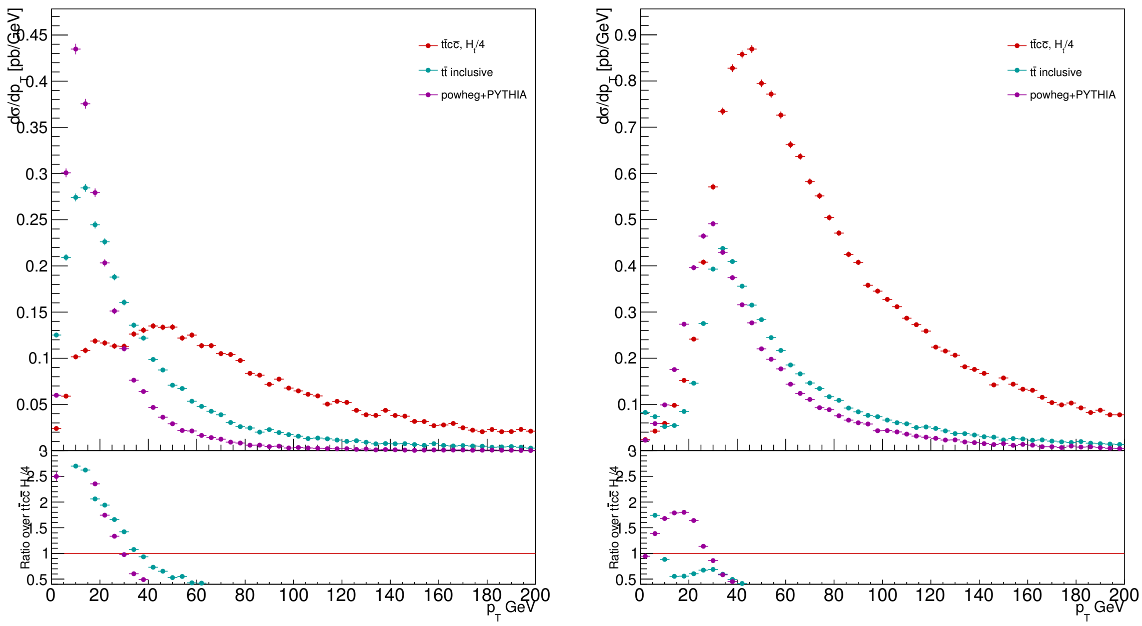


Figure 10: Number of charm quarks in an event where a charm quark coming from a gluon is associated with a c hadron from a jet with a single c hadron.

Next the p_T of mothering gluons is observed in both the high and low c hadron p_T fraction regions and is shown in figure 11. As every generator stored the gluon p_T in the event record differently, the gluon p_T seen in figure 11 was found by calculating the p_T of the $c\bar{c}$ system mothered by the gluon in question.



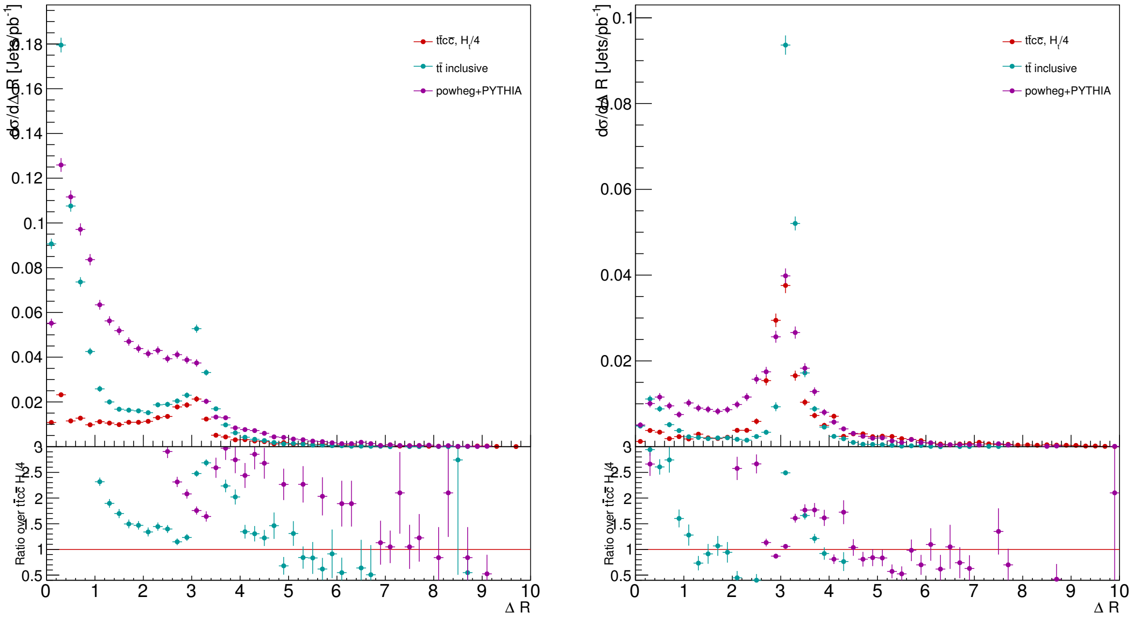
(a) (b)

Figure 11: p_T of gluon ($c\bar{c}$ system) mothering a charm quark associated with the c hadron of a jet with a single c hadron in the (a) low (< 0.35) and (b) high (> 0.35) c hadron p_T fraction regions.

Looking at figure 11b it can be said that all three investigated samples have similar gluon p_T distributions in the high c hadron p_T fraction region. In contrast, figure 11a shows that the gluons mothering c quarks associated with c hadrons of jets with a single c hadron in the Powheg+Pythia and $t\bar{t}$ inclusive samples are significantly softer than those in the $t\bar{t}c\bar{c}$ sample in the excess region. This agrees with the soft gluon splitting hypothesis.

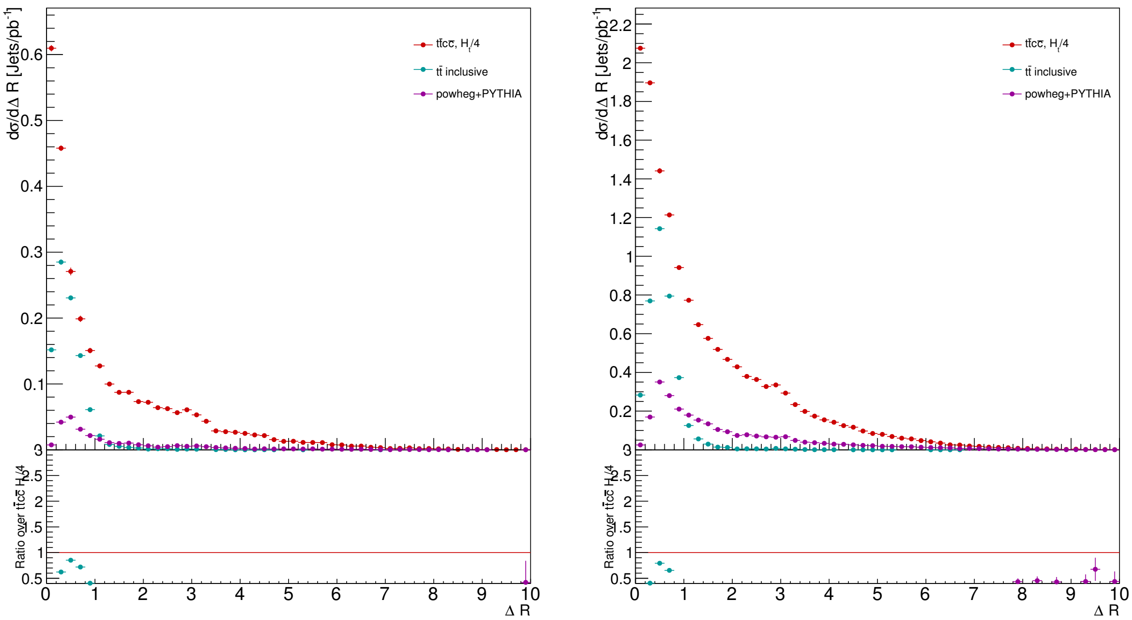
The hypothesis in question was tested further by looking at the value of ΔR, which depicts the separation between the two charm quarks coming from a gluon where one quark (or both) is associated to the c hadron of a jet with a single c hadron. The findings were divided into four situations by varying hadron p_T fraction regions (low < 0.35 and high > 0.35) and gluon p_T regions (low $p_T < 15$GeV and high $p_T > 40$GeV) and can be seen in figures 12 and 13.

As expected the separation of quarks from high p_T gluons (figure 13) peaks at low values and is consistent between all samples in both low and high c hadron p_T fraction regions. In the case of quark mothered by soft gluons (figure 12) a ΔR peak approximately around π is visible in both low and high c hadron p_T fraction regions, meaning that there are soft gluons apparent in the samples producing largely separated quarks. This is indeed consistent with the made hypothesis.



(a) (b)

Figure 12: ΔR between the two charm quarks mothered by low $p_T < 15$GeV gluons where one quark (or both) is associated to the c hadron of a jet with a single c hadron in the (a) low (< 0.35) and (b) high (> 0.35) c hadron p_T fraction regions.



(a) (b)

Figure 13: ΔR between the two charm quarks mothered by high $p_T > 40$GeV gluons where one quark (or both) is associated to the c hadron of a jet with a single c hadron in the (a) low (< 0.35) and (b) high (> 0.35) c hadron p_T fraction regions.

In the low c hadron p_T region the separation of charm quarks mothered by soft gluons have a clear peak at small separation values as seen in figure 12a. This can possibly be explained by kinematics. The two quarks are produced back-to-back in the gluon frame-of-reference. As we select one quark to have a small separation with a c hadron in a jet the direction of one of the quarks will be close to the direction of the jet and the other one will be roughly opposite. The relevant diagram can be seen in figure 14a. The peak at small separation can be explained if the gluon has just enough energy to boost the opposing gluon in the direction of the jet as seen in the diagram in figure 14b.

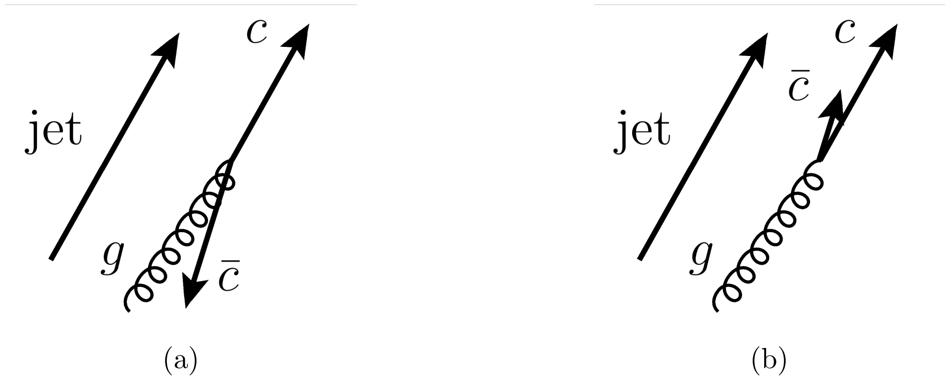


Figure 14: A pair of charm quarks produced by a soft gluon in the frame-of-reference of (a) the gluon and (b) the detector.

Overall all observations made in this section are consistent with the hypothesis that the excess of high p_T jets with low p_T c hadrons is at least partly due to soft gluons splitting into a pair of c quarks which in turn produce c hadrons which match to jets leading to identification of jets as c jets. Further investigation into this hypothesis is needed. Theoretically this process can happen with a soft gluon splitting into two b quarks causing an excess in the $t\bar{t}b\bar{b}$ samples as well, however further investigation not done in this project should be conducted.

4 Conclusion

When conducting research involving Monte Carlo simulations it is essential to compare different methods of simulation in order to select the method most fitting for the required task. When comparing simulation methods for $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ background a discrepancy was uncovered between samples generated using ME calculations for $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ production and samples produced using ME calculations for $t\bar{t}$ and PS calculations for $b\bar{b}$ or $c\bar{c}$ production. The latter samples were found to have an excess in the region of low (< 0.35) b or c hadron p_T to jet p_T fraction. This project focused on the investigation of this discrepancy. As the excess was more prominent in $t\bar{t}c\bar{c}$ samples and for jets with one c or b hadron the investigation focused on these cases.

Firstly previous results were replicated and new n-tuples were produced containing the required information for the investigation. As the excess was not affected by minimum jet p_T cuts the hypothesis that the discrepancy is due to kinematic sculpting effects was discarded. Comparing the p_T of c hadrons of jets with a single hadron and the c quarks associated with them it was found that both the c hadrons and c quarks have low p_T and thus the excess was seen to not be caused by a hadronisation problem in the PS calculations. It was also seen that the reason as to why the excess was more prominent in the Powheg+Pythia sample was due to quarks from non-gluonic origins such as weak decays. As weak decays would not produce b quarks in such quantity this also would explain why the excess is less prominent in the $t\bar{t}b\bar{b}$ samples, however this requires further investigation. The contribution of c quarks mothered by protons to the excess will also need further observation. Lastly, a hypothesis suggesting that the excess is caused by c hadrons coming from c quarks originating from soft gluon splitting was investigated. It was seen that in the excess region the gluons mothering c quarks associated with c hadrons of jets with one c hadron from samples using PS calculations for heavy flavour simulation are significantly softer than those using ME calculations. More so it was seen that the c quarks produced by gluons of high p_T (> 40GeV) have small separation (ΔR), whereas soft gluons were seen to produce largely separated c quarks. Thus the observations made agree with the soft gluon splitting hypothesis, however further investigation into the topic is needed.

Overall the progress made during this project will be useful for further research and observations made will have to be taken into account in further analysis.

5 Acknowledgements

I would like to thank my supervisors Judith Katzy and John Stakely Keller for providing me with this interesting project and the help and support that they offered during it. I also would like to thank the members of the ATLAS group at DESY for the warm welcome I received during my time at DESY. Lastly, I offer my thanks to the organisers of the DESY summer student program for making this invaluable experience possible.

Appendix A Experimental Procedure and Results

A.1 Quark Origin Investigation

		gluon	W boson	proton	$u/d/s(\bar{u}/\bar{d}/\bar{s})$
$t\bar{t}c\bar{c}$	Low	93.85%	0.11%	0.00%	5.99%
	High	94.36%	0.06%	0.00%	5.57%
$t\bar{t}$ inclusive	Low	98.60%	0.50%	0.00%	0.00%
	High	98.99%	0.89%	0.00%	0.00%
Powheg+Pythia	Low	33.31%	23.28%	11.57%	31.84%
	High	49.59%	6.98%	2.79%	40.64%

Table 1: Precentage of mother species of charm quarks associated with c hadrons of jets with one c hadron in the low (< 0.35) and high (> 0.35) c hadron p_T fraction regions.

References

[1] Georges Aad et al. “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Eur. Phys. J.* C75.7 (2015), p. 349. DOI: 10.1140/epjc/s10052-015-3543-1. arXiv: 1503.05066 [hep-ex].

[2] *Studies of tt+cc production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment.* Tech. rep. ATL-PHYS-PUB-2016-011. Geneva: CERN, May 2016.

[3] Georges Aad et al. "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 05 (2016), p. 160. DOI: 10.1007/JHEP05(2016)160. arXiv: 1604.03812 [hep-ex].

[4] B. Webber. "Parton shower Monte Carlo event generators". In: *Scholarpedia* 6.12 (2011). revision #128236, p. 10662.

[5] Stephen Mrenna. *Talk: Monte Carlo & Event Generators*. Computing Division, Generators and Detector Simulation Group, Fermilab. 2009.

[6] D.H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 2000. ISBN: 9780521621960.

[7] K. A. Olive et al. "Review of Particle Physics: 40. Monte Carlo Event Generators". In: *Chin. Phys.* C38 (2014). DOI: 10.1088/1674-1137/38/9/090001.

[8] Judith M. Katzy. "QCD Monte-Carlo model tunes for the LHC". In: *Prog. Part. Nucl. Phys.* 73 (2013), pp. 141–187. DOI: 10.1016/j.ppnp.2013.08.002.

[9] S. Frixione et al. *The MC@NLO Package*. (Accessed on 30/08/2016). URL: http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO/.