Optimization of charm tagging algorithms and feasibility study of VH(cc) channel

CERN Summer Student Project Report

Joakim Sjunnebo

Supervisors:
Heather Gray and Timothy Paul Scanlon

September 12, 2014

Abstract

This report is a summary of the work I have done as a summer student at CERN. I have worked on two different projects, thus, the report is divided into two parts, each describing one of the projects. The first part describes a study aimed at optimizing charm tagging algorithms. The second part is a summary of a feasibility study of the VH(cc) channel. The two projects are tied together as results from the first project are used in the second project.

Chapter 1

Optimization of charm tagging algorithms

1.1 Introduction

The identification of jets originating from b-quarks is an important part of the LHC physics program. Charm tagging has been used in some physics analyses, but it is still associated with low efficiencies and the full potential has likely not yet been exploited. The possibility to apply charm tagging would be of great interest in the LHC physics program, as it would enable the analysis of the $H \to c\bar{c}$ channel to be performed, as well as open up possibilities to analyze various beyond the Standard Model signals, such as supersymmetry [1, 2, 3].

There are a range of different b-tagging algorithms in use at ATLAS, as well as a handful of algorithms designed to tag c-jets. So far, the performance of the c-tagging algorithms has not been fully studied, therefore further performance gains are still possible.

This note presents a study of the properties and c-tagging performance of three flavour tagging algorithms; MV2, JetFitterCharm and GAIA. MV2 is a b-tagging algorithm, but is in this analysis utilized in such a way that it performs c-tagging instead. JetFitterCharm is a c-tagger which has been used for some time in ATLAS physics analyses, whereas GAIA is a relatively new c-tagging algorithm. The properties of the algorithm outputs have been studied and, based on those properties, a handful of potential classifiers have been proposed. The c-tagging performance of these classifiers has been determined in terms of the c-jet efficiency, b-jet rejection and light-flavour jet rejection, and a few operating points for use in future physics analyses have been proposed.

1.2 Simulated data samples

The studies presented in this report were performed using a simulated $t\bar{t}$ sample from the official MC12 Monte Carlo production. The sample was generated using Powheg+Pythia with the Perugia 2011C tune. Jets were required to have transverse momentum $p_{\mathrm{T}} > 20$ GeV and pseudorapidity $|\eta| < 2.5$. Furthermore, the Jet Vertex Fraction, defined as the summed transverse momentum of the tracks associated to a jet consistent with originating from the selected primary vertex divided by the summed transverse momentum of all tracks associated to a jet, was required to be larger than 0.5. The total number of events used in the analysis was 10 million and the number of jets passing the selection cuts was approximately 4 million.

1.3 Jet flavour tagging algorithms

A schematic diagram of the current ATLAS flavour tagging framework is shown in Fig. 1.1. Tagging begins by applying the four algorithms at the top. Each of these algorithms use a different approach to tag the jet flavour, the two most common approaches involve calculating the impact parameters and reconstructing the secondary vertex [4]. After these algorithms have been applied, all the information in a jet has been distilled down to several dozen parameters. These are then used as inputs to various likelihood lookup tables and neural networks before reaching the final tagging algorithms MV1, MV1c, MV2 and JetFitterCharm. MV1 is a widely used b-tagging algorithm in ATLAS physics analyses. It is implemented as an artificial neural network trained with b-jets as signal and light-flavour jets as background and provides a very large light-flavour jet rejection for a given b-jet efficiency [5]. MV1c is similar to MV1 but is trained with c-jets as background instead of light-flavour jets, resulting in a higher c-jet rejection for a given b-jet efficiency, at the cost of a slightly reduced light-flavour jet rejection. MV2 and JetFitterCharm are both used extensively in this study and will be described in more detail in the next two sections.

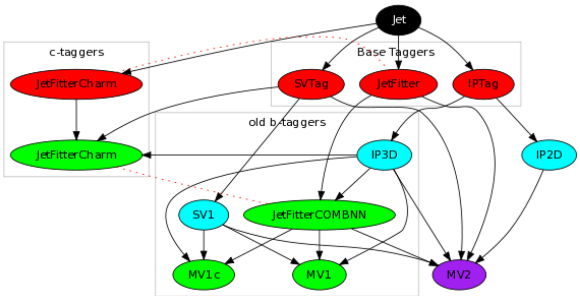


Figure 1.1: The ATLAS flavour tagging information flow. Red nodes are tracking or vertexing algorithms, cyan likelihood lookup tables, green neural networks and magenta boosted decision trees.

1.3.1 MV2

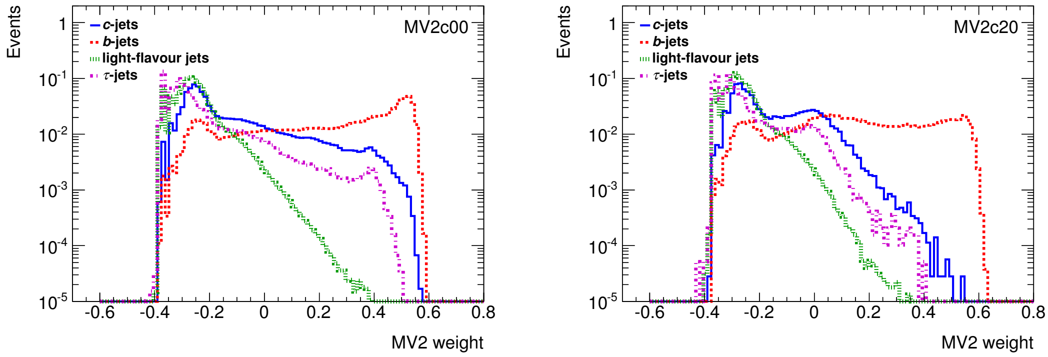
MV2 is an improved version of MV1 and MV1c and is implemented as a boosted decision tree with 25 input variables. There are several flavours of MV2 which differ by the composition of the samples used in training. For example, MV2c00 is trained with a background composed of 100% light-flavour jets and MV2c20 is trained with 20% c-jets and 80% light-flavour jets. Hence, MV2 offers several different choices depending on the desired performance in terms of light-flavour jet rejection and c-jet rejection. MV2 calculates and outputs a tag weight for each jet. The distribution of this weight for each of the different jet flavours, including jets initiated by τ-leptons, is shown in Fig. 1.2(a) for MV2c00 and Fig. 1.2(b) for MV2c20. Depending on how the cut on the tag weight is applied, MV2 can be used both as a b-tagger and a c-tagger. A cut which selects all jets with tag weights above a certain value will predominantly select b-jets. Conversely, a cut which selects all jets with tag weights between two certain values will, if these two values are selected carefully, select c-jets and reject the other flavours.

Since MV2c20 is trained against both light-flavour jets and c-jets, it has a larger separation between b-jets and c-jets than MV2c00. By comparing the two distributions it is evident that MV2c00 has a higher concentration of c-jets in the center than MV2c20. Therefore, MV2c00 has been used in this analysis in order to study the c-tagging performance of MV2.

1.3.2 JetFitterCharm

JetFitter is an algorithm which exploits the topological structure of weak b- and c-hadron decays inside the jet. It searches for a common line on which the primary vertex and the b- and c-vertices lie, as well as their position on this line, calculating an approximated flight path for the b-hadron. A variation of JetFitter, JetFitterCharm, is an artificial neural network trained to select c-jets. It maps the input variables to three posterior output variables; P_b, P_c and P_u, roughly corresponding to the probability that a jet is b-quark, c-quark or light quark or gluon initiated. The sum of the output probabilities is equal to one, which means that only two of the probabilities contain unique information.

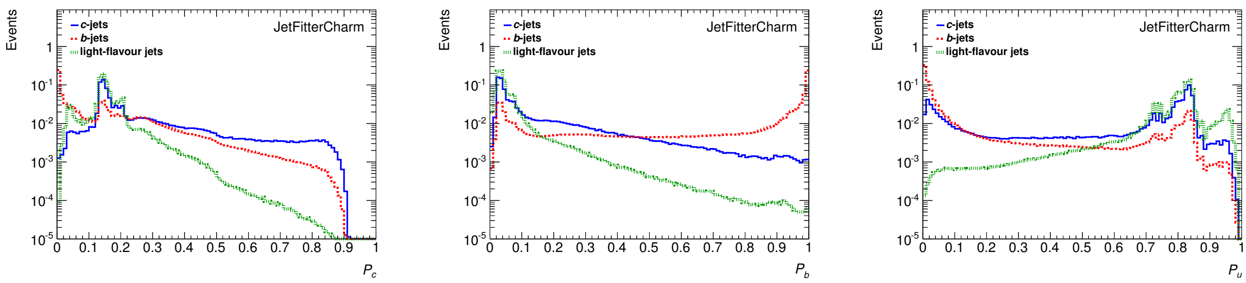
Figure 1.3 shows the distribution of each of the three output probabilities for b-jets, c-jets and light-flavour jets. From these distributions, the separation of jet flavours and discriminating power



(a) MV2c00 (b) MV2c20

Figure 1.2: MV2 tag weight distributions for c-jets, b-jets, light-flavour jets and τ-jets. Figure 1.2(a) shows the MV2c00 distribution and Figure 1.2(b) shows MV2c20.

of each output probability can be seen. P_b separates b-jets from the other flavours very well since the majority of the b-jets are concentrated in the right part of the distribution, where P_b has a peak, whereas the other two flavours are most abundant in the opposite part. P_c lacks a sharp peak at high values, but still the c-jets are concentrated to the right part of the distribution and the separation is decent. P_u also lacks a peak at the highest values and the separation of light-flavour jets from the other two flavours are quite poor. This means that a lot of c-jets have to be rejected in order to reject light-flavour jets, and that there is a trade-off between c-jet efficiency and light-flavour jet rejection.

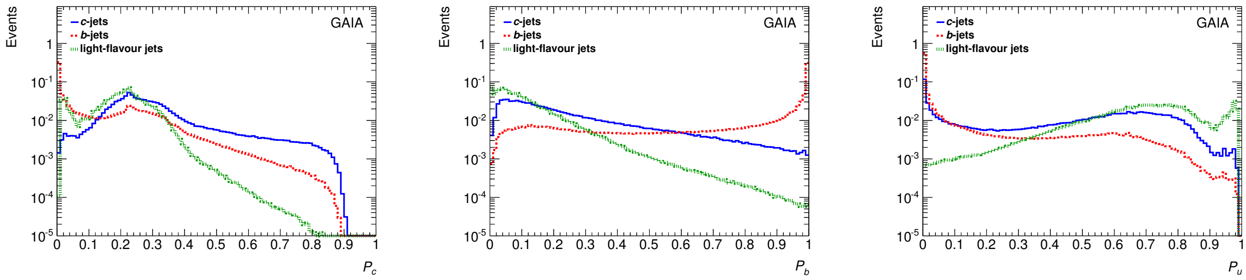


(a) P_c (b) P_b (c) P_u

Figure 1.3: Distribution of jet flavours for the three JetFitterCharm outputs P_c, P_b and P_u.

1.3.3 GAIA

GAIA is a general-purpose jet flavour identification algorithm implemented as a deep learning neural network [6]. It was designed to integrate the output from existing flavour tagging algorithms and reduce the number of intermediate taggers without losing performance. Like JetFitterCharm, GAIA has three output probabilities; P_b, P_c and P_u. Figure 1.4 shows the b-jet, c-jet and light-flavour jet distributions for each of the output probabilities. Both P_c and P_b show similar jet flavour separations compared to JetFitterCharm, whereas P_u seems to separate the light-flavour jets slightly better for GAIA than JetFittterCharm.



(a) P_c (b) P_b (c) P_u

Figure 1.4: Distribution of jet flavours for the three GAIA outputs P_c, P_b and P_u.

1.4 Cuts and method

When constructing a classifier, the goal is to get a good separation between the c-jets and the rest of the jets. A large overlap between the jets means that it is difficult to place a cut which gives a high c-efficiency and also gives a high b-jet rejection and light-flavour jet rejection. Since c-jets are similar to both b-jets and light-flavour jets, there is a trade-off between having good b-jet rejection and light-flavour jet rejection. This means that a classifier which is able to separate c-jets and b-jets well is not able to separate c-jets and light-flavour jets very well. Because of this trade-off, there is no universal optimal choice in selecting a classifier and cut, it depends on the intended application of the flavour tagging. When used in a specific physics analysis, the signal and background composition determines the optimal efficiencies for that analysis. For example, if the background is dominated by b-jets and the signal by c-jets, it is more important to have a high c-tagging efficiency and b-rejection than a high light-flavour jet rejection.

The approach taken in this analysis has been to study the distribution of jet flavours for different combinations of the algorithm outputs to find classifiers which have large separations between the jet flavours. The performance of each classifier was determined by varying the position of the cuts over the entire distribution and calculating the c-jet efficiency, b-jet rejection and light-flavour jet rejection for each position.

Another method consisted of training a boosted decision tree to make a classifier. Different combinations of input variables were used in the training, including the output probabilities P_b, P_c and P_u, as well as p_T and η. Several configurations and options were tested, including the number of trees in the forest, the boosting type and parameter, the maximum depth of the trees and the minimum node size, as well as the flavour composition of the training sample. However, no significant gain in performance was found when using the BDT. Therefore, the details of this study are omitted from this report.

1.4.1 MV2

The MV2 weight distribution shows that the majority of the b-jets have large tag weights. A cut which rejects all jets with a tag weight above a certain value was applied in order to remove those b-jets. Similarly, the majority of the light-flavour jets have large negative tag weights and a cut which rejects all jets with tag weights below a certain value was applied to remove light-flavour jets. This leaves all jets in the center of the distribution, which contains a considerable amount of the c-jets.

1.4.2 JetFitterCharm and GAIA

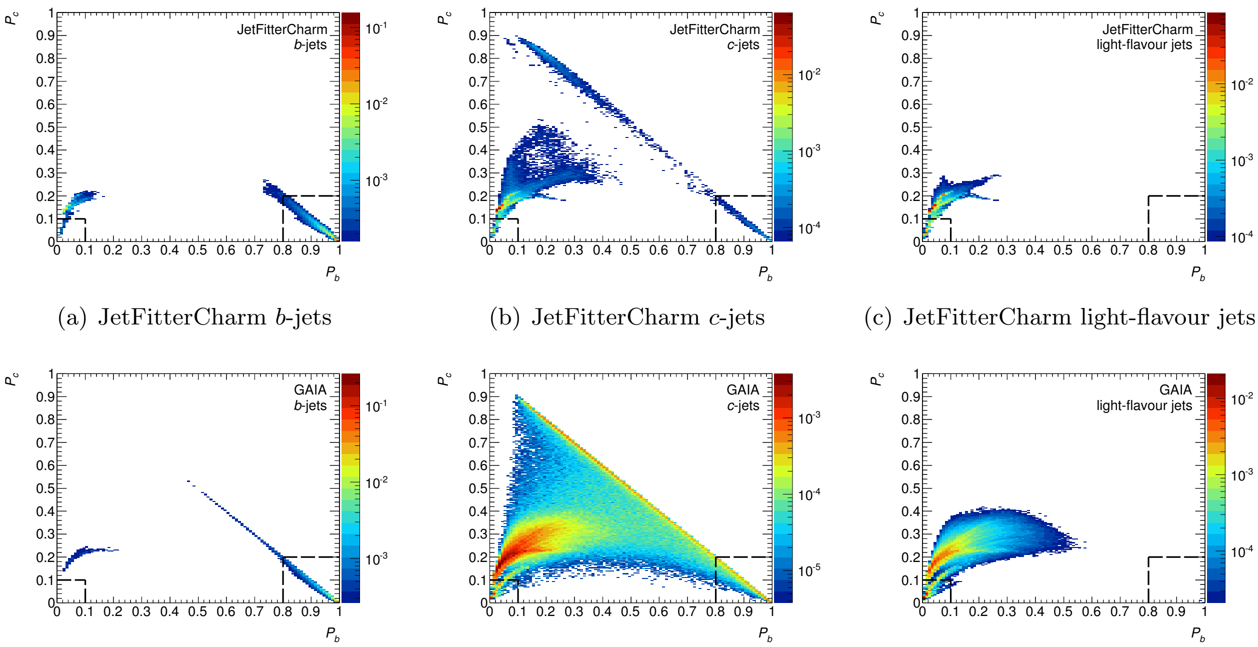
Since both JetFitterCharm and GAIA have the same output probabilities, the same type of cuts were applied to them, which allowed for a direct comparison of the performance of the two algorithms. Two different cuts were studied and analyzed.

$P_b - P_c$ plane cut

Figure 1.5 shows the the distribution of b-jets, c-jets and light-flavour jets in the $P_b - P_c$ plane for JetFitterCharm and GAIA. The vast majority of the b-jets are located in the lower-right region of the plane. Since the concentration of c-jets and light-flavour jets is low in this region, it can be cut away in order to reject b-jets. A similar cut was applied to the lower-left region of the plane in order to reject light-flavour jets. The cut was implemented as two squares, one with a corner in the lower-left corner of the plane and the other with a corner in the lower-right corner of the plane. A jet was accepted if

$$(P_c > c_1 \,||\, P_b > c_1) \;\&\&\; (P_c > c_2 \,||\, P_b < 1 - c_2) \tag{1.1}$$

was true, where c_1 and c_2 are the two cut values and correspond to the size of the squares.



(a) JetFitterCharm b-jets (b) JetFitterCharm c-jets (c) JetFitterCharm light-flavour jets

(d) GAIA b-jets (e) GAIA c-jets (f) GAIA light-flavour jets

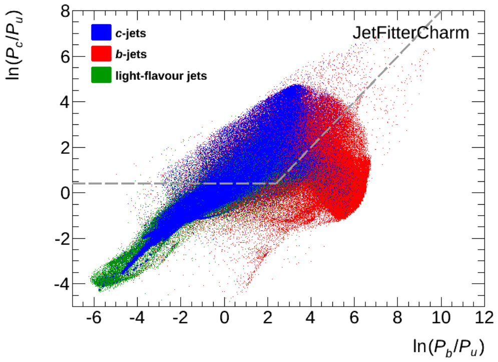
Figure 1.5: Distribution of jets in the $P_b - P_c$ plane for b-jets, c-jets and light-flavour jets for JetFitterCharm and GAIA. The black dashed lines indicate the cuts. The cut defined by the left square is placed to reject light-flavour jets and the right square to reject b-jets.

$\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut

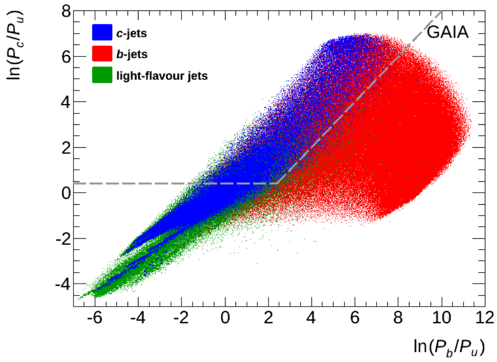
The distribution of jet flavours in the $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane is shown in Fig. 1.6(a) for JetFitterCharm and Fig. 1.6(b) for GAIA. There is a large accumulation of b-jets in the right part of the plane, a region which is largely empty of c-jets and light-flavour jets. This region was rejected by placing a sloped cut in the plane, roughly tangential to the c-jet–b-jet boundary. The light-flavour jets are concentrated to the lower part of the plane, so a horizontal cut was applied to reject these. More specifically, a jet was selected if

$$\ln(P_c/P_u) > \ln(P_b/P_u) + c_1 \;\&\&\; \ln(P_c/P_u) > c_2 \tag{1.2}$$

was true.



(a) JetFitterCharm

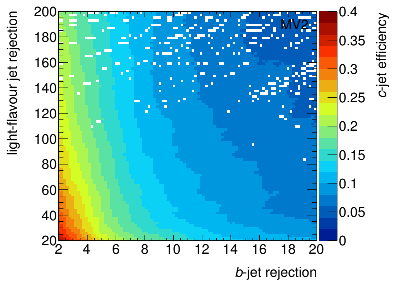


(b) GAIA

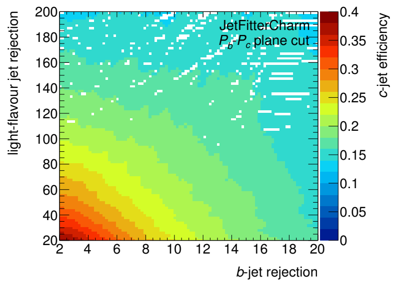
Figure 1.6: Distribution of jets in the $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane for 1.6(b) JetFitterCharm and 1.6(b) GAIA. The gray dashed lines indicate the cuts. The sloped cut is placed to reject b-jets and the horizontal cut to reject light-flavour jets.

1.5 Results

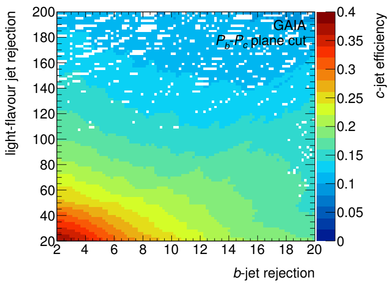
The two cut values for the two-dimensional cuts described in Section 1.4 were varied independently of each other and the resulting efficiencies and rejection rates were calculated for each pair of cut values. The results are shown in Fig. 1.7 where the c-jet efficiency is color coded and shown as a function of the b-jet rejection and light-flavour jet rejection.



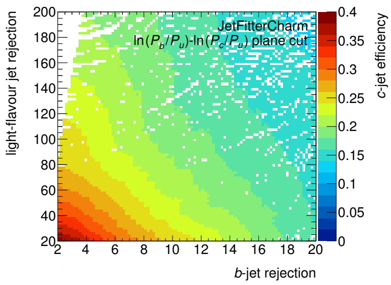
(a) MV2



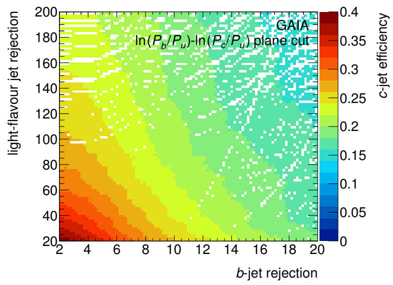
(b) JetFitterCharm $P_b - P_c$ plane cut



(c) GAIA $P_b - P_c$ plane cut



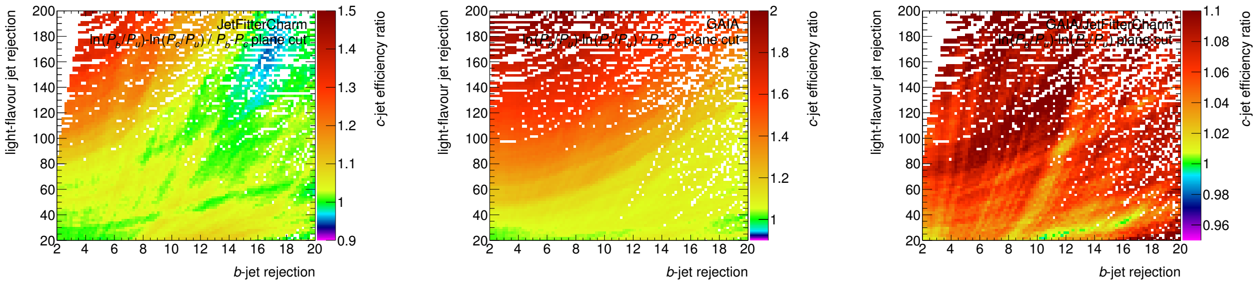
(d) JetFitterCharm $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut



(e) GAIA $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut

Figure 1.7: c-tagging performance plot showing the c-jet efficiency in color coding as a function of b-jet rejection and light-flavour jet rejection. Figure 1.7(a) shows the MV2 performance, 1.7(b) JetFitterCharm $P_b - P_c$ plane cut, 1.7(c) GAIA $P_b - P_c$ plane cut, 1.7(d) JetFitterCharm $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut and 1.7(e) GAIA $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut.

To compare the performance of the different cuts and tagging algorithms, ratios of the performance plots for the two-dimensional cuts for GAIA and JetFitterCharm are shown in Fig. 1.8. Here the color coding represents the ratio of c-jet efficiencies and is shown as a function of the b-jet rejection and light-flavour jet rejection. The $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane cut performs better than the $P_b - P_c$ plane cut in most regions for both JetfitterCharm and GAIA. The ratio of the two $\ln(P_b/P_u) - \ln(P_c/P_u)$ performance plots shows that the GAIA version of this classifier generally performs slightly better than JetFitterCharm, with up to about 10% higher c-jet efficiencies.



(a) JetFitterCharm comparison

(b) GAIA comparison

(c) JetFitterCharm and GAIA comparison

Figure 1.8: Ratios of GAIA and JetFitterCharm performance plots showing the c-jet efficiency ratio color coded as a function of c-rejection and light-flavour jet rejection. Figure 1.8(a) shows the ratio of the JetFitterCharm $\ln(P_b/P_u) - \ln(P_c/P_u)$ and $P_b - P_c$ performance plots, 1.8(b) shows the ratio of the GAIA $\ln(P_b/P_u) - \ln(P_c/P_u)$ and $P_b - P_c$ performance plots and 1.8(c) shows the ratio of the GAIA and JetFitterCharm $\ln(P_b/P_u) - \ln(P_c/P_u)$ performance plots.

Based on the performance study and the results above, a few operating points have been proposed for use in physics analyses. Since the best performance was achieved by GAIA using the $\ln(P_b/P_u) - \ln(P_c/P_u)$ classifier, it was used to calculate all the operating point efficiencies. The proposed operating points are summarized in Table 1.1, where the achieved efficiencies as well as the cut values are given for each point. The bold points are suggested for calibration. Three points with relatively high c-jet efficiency, 45%, 40% and 35%, have been proposed for analyses where a high c-jet efficiency is required. Two more c-jet efficiencies have been proposed, 30% and 20%, each with three different combinations of the rejection factors; one with a high b-jet rejection, one with a high light-flavour jet rejection and one with a compromise between the rejection factors.

Table 1.1: Efficiencies for proposed operating points. All efficiencies are calculated using the $\ln(P_b/P_u) - \ln(P_c/P_u)$ classifier for GAIA. c_1 and c_2 are the two cut values defined in Eq. 1.2. RF stands for Rejection Factor which is the reciprocal of the efficiency. The bold points are suggested for calibration.

c-jet eff. (%)	b-jet RF	l-jet RF	c_1	c_2
45	1.7	11.1	−5.90	−0.31
40	2.3	17.0	−3.70	−0.12
35	**2.7**	**32.3**	**−3.40**	**0.20**
30	**3.9**	**51.0**	**−2.30**	**0.46**
30	2.7	57.5	−4.00	0.60
30	**7.4**	**17.5**	**−0.60**	**−0.20**
20	8.1	213.7	−1.58	1.46
20	6.4	303.0	−2.40	1.80
20	**11.1**	**101.0**	**−0.70**	**0.75**

1.6 Conclusions

The c-tagging performance of three jet flavour tagging algorithms has been studied. Based on the properties of the algorithm outputs, different classifiers have been proposed, and by cutting on those classifiers, the c-jet efficiency, b-jet rejection and light-flavour jet rejection has been determined. MV2, being a b-tagging algorithm, performed less optimally than the c-tagging algorithms JetFitterCharm and GAIA. Of the proposed classifiers, the cut in the $\ln(P_b/P_u) - \ln(P_c/P_u)$ plane provided the highest c-jet efficiency for both JetFitterCharm and GAIA. Compared to each other, GAIA performed slightly better than JetFitterCharm for this classifier, with up to about 10% higher c-jet efficiencies for equivalent b-jet rejection and light-flavour jet rejection.

The proposed operating points and efficiencies can be compared to the c-tagging operating point used in some of the current ATLAS physics analyses [2, 3]. This operating point have a c-jet efficiency of 19%, a b-jet rejection of 8.0 and a light-flavour jet rejection of 200. One of the operating points proposed in this note has a similar performance, with a c-jet efficiency of 20%, b-jet rejection of 8.1 and a light-flavour jet rejection of 213.7. A higher b-jet efficiency can be achieved at the cost of a lower light-flavour jet rejection, for example with the proposed operating point with a c-jet efficiency of 20%, b-jet rejection of 11.1 and a light-flavour jet rejection of 101. A higher c-jet efficiency can also be achieved, but at the cost of a lower b-jet rejection and light-flavour jet rejection, for example with the proposed operating point with a c-jet efficiency of 30%, a b-jet rejection of 3.9 and a light-flavour jet rejection of 51.

Beside these operating points, others have been proposed as well, reflecting the different requirements posed on the tagging efficiencies in different physics analyses. Since the performance of the classifiers developed in this study show promising results and that the c-tagging performance could be improved from current levels, it makes sense to calibrate some or all of the suggested operating points, so that they may be used in future physics analyses.

Chapter 2

VH(cc) feasibility study

2.1 Introduction

The discovery of the Higgs boson was the main achievement during Run-1 of the LHC. This discovery was made using the decays of the Higgs to bosons [7, 8]. Therefore, the observation and measurement of the coupling of the Higgs boson to fermions is of great interest. Thus far evidence has been found of the Higgs coupling to third-generation fermions: τ leptons and b quarks [9, 10], however, as yet, no measurements have had sensitivity to second-generation fermions. The decay of the Higgs to charm quarks has a branching ratio of 3%, but this is the largest branching ratio of the Higgs to second-generation fermions. The small branching ratio and the low tagging efficiencies still associated with charm tagging make the measurement of the decay to a pair of charm quarks very difficult.

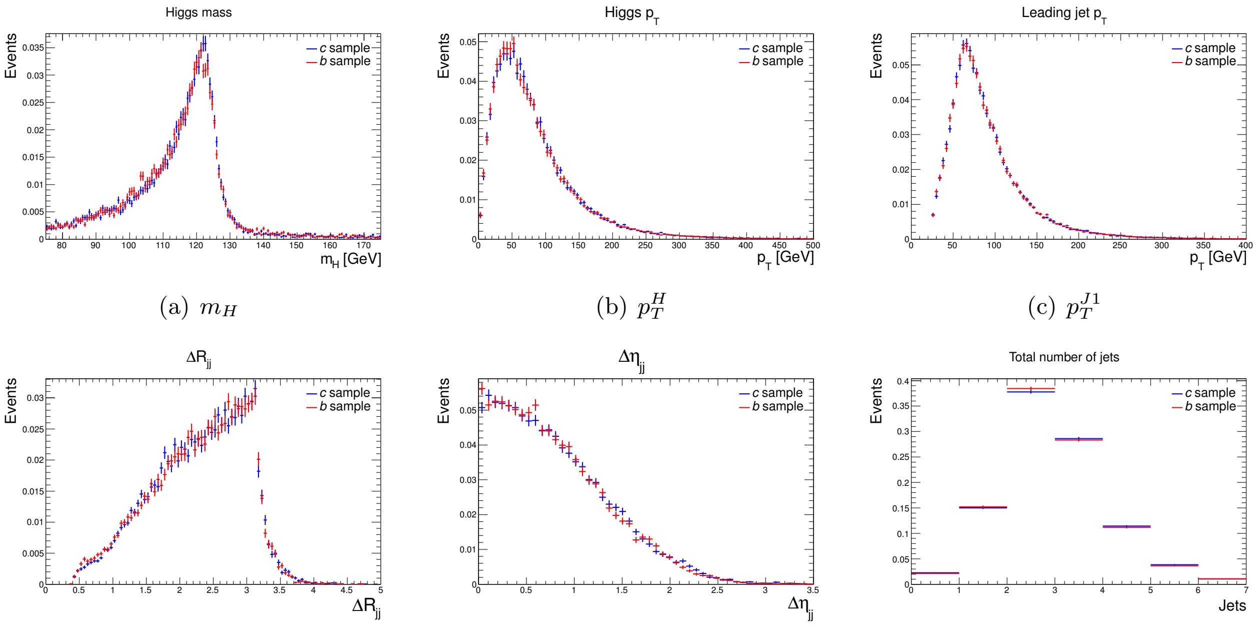
This report presents a summary of a feasibility study for the observation of the Higgs boson decaying into a pair of charm quarks, with the Higgs boson produced in association with a vector boson. The study is divided into two parts. In the first part, yields from the existing VH(bb) analysis are altered to make an estimation of the expected yields for the VH(cc) channel. In the second part, a truth-level analysis has been performed to make the same estimation.

2.2 Estimation using VH(bb) yields

The 2-tag 0-lepton, 1-lepton and 2-lepton prefit yields from [11] were used as a starting point. These yields were then unfolded to get the pre-tag yields, which was done by scaling the prefit yields with the jet flavour tagging efficiencies for the respective channels. To do this, an assumption had to be made about the flavour composition of the samples. For example, the VH(bb) yields were divided by the square of the b-tagging efficiency. A study of the kinematics of both the VH(bb) and VH(cc) events were performed. For this, a signal sample was generated using Powheg+Pythia8 with the AU2CT10 tune. The study was done by comparing the distributions of important variables such as the dijet mass, transverse momentum, ΔR and the number of jets in each event for both channels. The results, shown in Fig. 2.1, show that the kinematics are very similar for the two channels. Therefore, the VH(cc) yield was introduced as the VH(bb) yield multiplied by the ratio of branching ratios. By introducing typical c-tagging efficiencies and multiplying the yields by these efficiencies, an estimation of the expected yields for the VH(cc) channel was obtained.

The analysis crucially depends on the c-tagging performance. Therefore, the study discussed in Chapter 1 was performed to determine optimal tagging points to use in this analysis. Based on that study, an operating point which provided a c-jet efficiency of 30%, a b-jet rejection of 3.9 and a light-flavour jet rejection of 51 was determined. With these efficiencies, the significance was estimated to be 0.01 for all three channels with an integrated luminosity of 20 fb^{-1}. If the integrated luminosity was increased to 3000 fb^{-1}, the significances were estimated to be 0.15 for the 2-lepton, 0.16 for the 1-lepton and 0.19 for the 0-lepton channel.

The expected significance was also calculated as a function of the tagging efficiencies and the integrated luminosity. The results, shown in Fig. 2.2, show that the expected significance is very low unless the c-tagging performance and integrated luminosity is very high. The 0-lepton, 1-lepton



(a) m_H (b) p_T^H (c) p_T^{J1}

(d) $\Delta R(j,j)$ (e) $\Delta\eta(j,j)$ (f) Number of jets

Figure 2.1: Comparison of kinematics between VH(bb) and VH(cc) events.

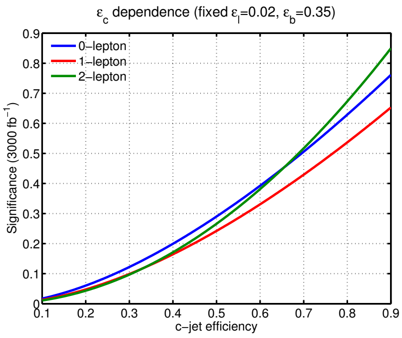
and 2-lepton channels have very different dependence on the tagging efficiencies, which is a result of the different background composition of these channels. The significance is seen to increase with c-jet efficiency for all three channels. The b-jet efficiency only has a small effect on the 0-lepton and 1-lepton channels, and a very large effect on the 2-lepton channel. Conversely, the light-flavour jet efficiency only has a large effect on the 0-lepton and 1-lepton channels, and almost no effect on the 2-lepton channel. The significance increased with integrated luminosity for all three channels. For an observation to be made of this channel, the integrated luminosity and c-tagging performance has to be very high.

2.3 Truth-level analysis

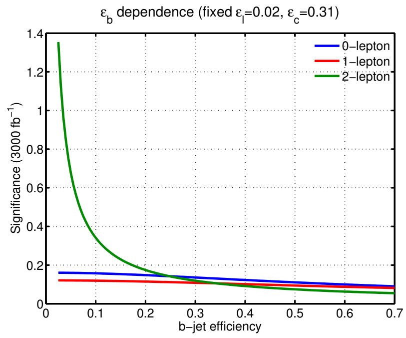
A truth-level study was performed to study the expected signal and background composition for the VH(cc) channel. Only $Z \to \mu\mu$ events were considered in this analysis, as they provide the clearest signal. The analysis was performed by accessing the truth information of the Monte Carlo events and only selecting events with two c-jets and a Z boson decaying to two muons. This was done on different samples, both signal and background, and from this, the expected yields were obtained. Figure 2.3 shows some of the results from this analysis, with distributions of the Higgs boson mass, the leading jet p_T, ΔR between the leading jets, and the number of jets, shown with the expected signal and background composition. The ratio of the VH(cc) to VH(bb) yield was 0.05, which was expected since the ratio of the branching ratios for the two decays modes is 0.05. The significance was calculated from the truth-level analysis yields as well and was found to be approximately 0.07.

2.4 Conclusions

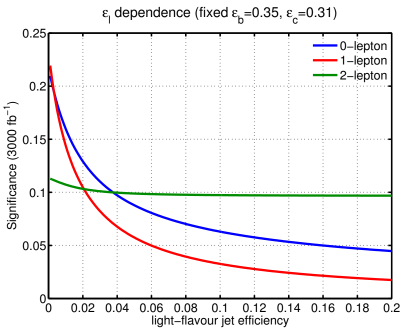
A feasibility study of the VH(cc) channel has been carried out using two different approaches. The estimated significances in the two approaches was found to be approximately 0.01 and 0.07, for an integrated luminosity of 20 fb^{-1}. With one of the optimized c-tagging operating points from the other part of this project and an integrated luminosity of 3000 fb^{-1}, the significance was estimated to be only 0.1. This shows that for the significances to reach a level high enough to make an observation, the c-tagging performance and integrated luminosity has to be very high.



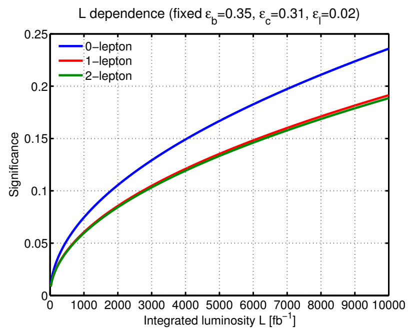
(a) c-jet efficiency



(b) b-jet efficiency

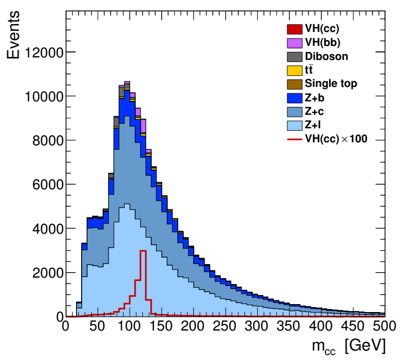


(c) light-flavour jet efficiency

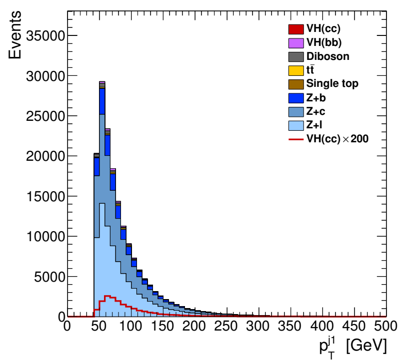


(d) Integrated luminosity

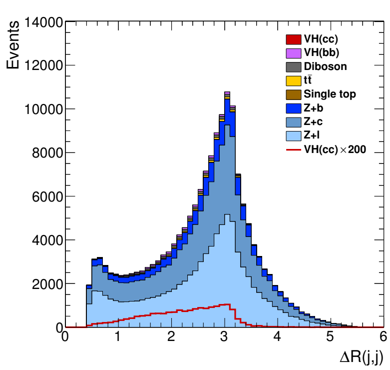
Figure 2.2: Expected VH(cc) significance as a function of tagging efficiencies and integrated luminosity.



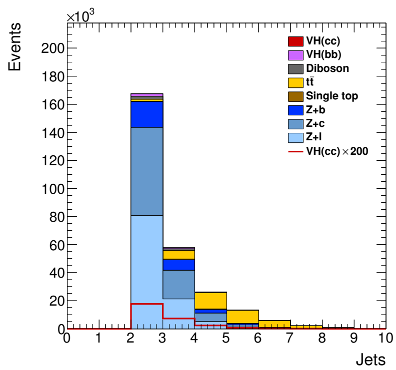
(a) m_{cc}



(b) Leading jet p_T



(c) $\Delta R(j, j)$



(d) Number of jets

Figure 2.3: Distribution of Higgs boson mass, leading jet p_T, ΔR between leading jets and total number of jets, showing the signal and background composition.

Bibliography

[1] C. Delaunay, T. Golling, G. Perez, and Y. Soreq, *Charming the Higgs*, Phys.Rev. **D89** (2014) 033014, arXiv:1310.7029 [hep-ph].

[2] ATLAS Collaboration, *Search for Scalar Top Quarks in final states with two charm jets and missing transverse momentum at $\sqrt{s} = 8$ TeV (supporting INT note)*, ATLAS-COM-PHYS-2013-1642. https://cds.cern.ch/record/1636856/files/ATL-COM-PHYS-2013-1642.pdf.

[3] ATLAS Collaboration, *Search for Scalar Charm Production with the ATLAS Detector in pp Collisions at $\sqrt{s} = 8$ TeV – supporting documentation*, ATLAS-COM-PHYS-2014-391. https://cds.cern.ch/record/1698012/files/ATL-COM-PHYS-2014-391.pdf.

[4] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102. https://cds.cern.ch/record/1369219/files/ATLAS-CONF-2011-102.pdf.

[5] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, ATLAS-COM-PHYS-2012-824. https://cds.cern.ch/record/1455862/files/ATL-COM-PHYS-2012-824.pdf.

[6] Y. Bengio, A. Courville, and P. Vincent, *Unsupervised Feature Learning and Deep Learning: A Review and New Perspectives*, CoRR **1206.5538** (2012). http://arxiv.org/abs/1206.5538.

[7] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[8] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[9] ATLAS Collaboration, *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2013-108, CERN, Geneva, Nov, 2013. https://cds.cern.ch/record/1632191/files/ATLAS-CONF-2013-108.pdf.

[10] ATLAS Collaboration, *Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2013-079, CERN, Geneva, Jul, 2013. https://cds.cern.ch/record/1563235/files/ATLAS-CONF-2013-079.pdf.

[11] ATLAS Collaboration, *Search for the production of a Standard Model Higgs Boson produced in association with a Vector Boson and decaying to a pair of b-quarks*, Tech. Rep. ATL-COM-2014-051, CERN, Geneva, Jul, 2014.