High mass VBF categorization for narrow-width resonance searches in the $H \to ZZ \to 4\ell$ channel with the ATLAS detector

Mzwandile Thabede[1], **Theodota Lagouri**[1], **Xifeng Ruan**[1] **and Bruce Mellado**[1,2]

[1] School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa

[2] iThemba LABS, National Research Foundation, P.O. Box 722, Somerset West, 7129, South Africa

E-mail: mzwandile.thabede@cern.ch

Abstract. The search for a heavy resonance (H) decaying to four leptons in the $H \to ZZ \to 4\ell$ ($\ell = \mu$ or e) channel can be conducted in at least two main production modes. This analysis channel is also one of the main decay channels for the Higgs boson which was discovered at the Large Hadron Collider (LHC) in 2012. In this proceeding, we categorize four lepton events using a cut-based approach in events with two or more jets (N-jets ≥ 2). This study focuses on the signal optimization for the $H \to ZZ \to 4\ell$ channel and the separation of vector boson fusion (VBF) events from the events produced via the gluon-gluon Fusion (ggF) production mechanism. The study is conducted using the ATLAS full Run 2 Monte Carlo with the total luminosity of 140 fb^{-1}. By defining a leading jet to be a jet with the largest transverse momentum, the optimal selection cuts on the invariant mass (m_{jj}) and the pseudo-rapidity difference between two leading jets ($\Delta\eta_{jj}$) can be obtained using signal selection efficiency and background rejection 2D maps. The new cuts obtained are then compared with the current optimal VBF selection to check if there is any improvement in the sensitivity. It was found that the standard cut-based VBF selection is still good enough to use for event categorization in the four-lepton channel with high selection efficiency and background rejection in the 0.2 TeV $\leq \mathrm{m}_H \leq$ 1 TeV mass range.

1. Introduction

After the discovery of the Standard Model Higgs boson in 2012 by ATLAS and CMS collaborations [1, 2], much of the work done has been to perform the measurements of its physical properties and to search for physics beyond the Standard Model. In hadron colliders, heavy Higgs bosons are produced mainly via the gluon fusion and vector boson fusion production mechanisms. This makes it possible to search separately for resonances in each production mode. To this end, it is necessary to distinguish events from each production mode by defining a set of selection requirements (criteria) to be applied on all events that enter the signal region. Given that the ggF production mode is the most dominant one at the LHC in terms of the Higgs boson production cross-section [3], one may ask if the sensitivity of the Higgs boson searches to the VBF production mode can be improved using Run 2 Monte Carlo from the ATLAS detector. This study attempts to answer this question by adopting a cut-based categorization method.

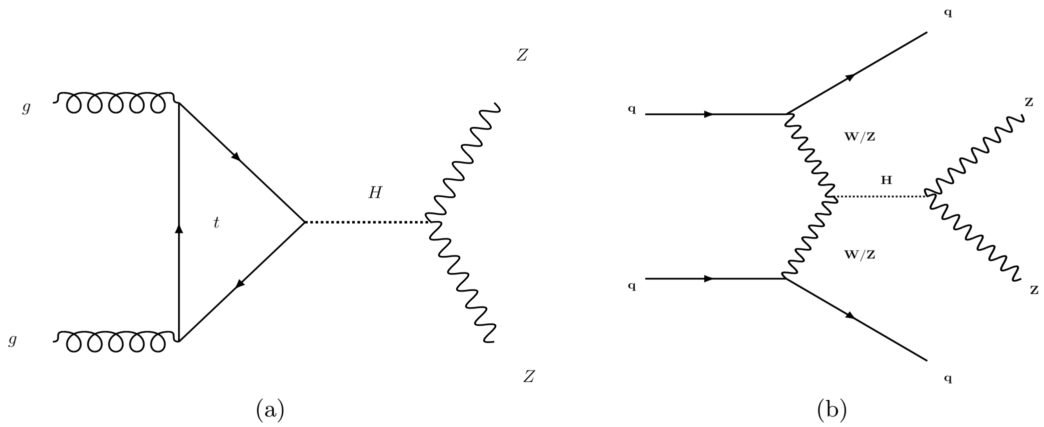


Figure 1: The leading production mechanisms for the Higgs boson in hadron colliders are (a) the gluon fusion and (b) the vector boson fusion production modes.

The selection for the signal events is based on the characteristic 2-jet signature of the VBF production mode ($pp \to H(\to ZZ)jj$), which is an important feature that is used to distinguish between ggF and VBF categories. Details of this selection and optimization study is discussed in this study where we assume the production of a narrow-width resonance (NWA). Searches for heavy resonances in the $H \to ZZ \to 4\ell$ decay channel can be performed separately in the ggF and in the VBF production modes provided that a criteria for selecting clean signal (VBF) events has been clearly defined. To classify events as either VBF or ggF, a method based on the signal selection efficiency and background rejection is adopted. This is achieved by choosing events with two or more jets with p_T greater than 30 GeV. Signal events will also satisfy the optimized cut on the di-jet invariant mass and the separation in η[1] between the two leading jets ($\eta_{j1} - \eta_{j2}$). The categorization method employed in this study provides the optimal selection cuts that can be used for all samples in the mass range considered. The Feynman diagrams for the ggF and VBF production modes are shown in figure 1.

2. Event Selection

This section provides a summary of the event selection [4, 5] as applied to all events in the four lepton channel. All events passing the standard four-lepton event selection are used for categorization. Only events with exactly four electron and/or muon combinations (one quadruplet) in the final state are selected. Events are classified into three different channels according to the flavor of leptons in the final state ($4e$, 4μ and $2\mu 2e$). For high mass studies, both Z-bosons are on-shell so it is only the measurement resolution that decides which way the pairing of leptons occurs. 2μ2e and 2e2μ quadruplets are selected by choosing di-lepton pairs with the invariant mass closest to the mass of the Z-boson (≈ 91.2 GeV) [6] so instead of using four channels in the analysis, 2μ2e and 2e2μ are combined into one channel called 2μ2e.

2.1. Final State Object Selection

To reconstruct the final state objects for the four lepton channel, electrons are required to have $p_T \geq 7$ GeV and with $|\eta| \leq 2.47$. Requirements for muon candidates is $p_T \geq 5$ GeV and with $|\eta| \leq 2.7$. Not more than one calorimeter-tagged, segment-tagged or stand-alone muon candidate in the pseudo-rapidity range $2.5 \leq |\eta| \leq 2.7$ is allowed per quadruplet. A jet is a composite object

[1] The pseudo-rapidity is defined in terms of the polar angle (θ) as $\eta = -\ln\tan(\theta/2)$.

contained in a narrow cone. The identification and reconstruction of an object as a jet begins on the topological clusters of cells in the calorimeter, used as input for the jet-finding algorithms. Jet objects are reconstructed using the Anti-k$_T$ algorithm [7]. This is a clustering algorithm with a radius parameter $R = 0.4$ and is implemented in the FastJet package [8] which uses the particle flow algorithm [9] as input. Jet objects in the four-lepton channel are required to satisfy $\mathrm{p}_T \geq 30$ GeV and $|\eta| \leq 4.5$. Pile-up jets can be suppressed by applying a cut on the jet vertex fraction (JVF) [10, 11] which requires that all jet objects must originate from the primary vertex.

2.2. Lepton Pairs and Quadruplet Formation

The Higgs boson candidate is obtained by selecting one quadruplet (two same flavour opposite sign lepton pairs) with $p_T \geq 20, 15, 10$ GeV for the first three leading leptons. The separation requirement of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \geq 0.10$ (0.20) for same- (different-) flavour leptons are required and all quadruplets containing a pair of leptons with invariant mass smaller than 5 GeV are vetoed. This requirement on the di-lepton invariant mass helps to reduce the contamination from J/ψ mesons in the analysis. The leading lepton pair is chosen by selecting a pair with invariant mass (m_{12}) closest to the mass of the Z-boson and is required to be between 50 GeV and 106 GeV. The invariant mass of the third and fourth leptons in the quadruplet (m_{34}) is required to be in the range ($m_{th} \leq m_{34} \leq 115$ GeV) where m_{th} is the threshold mass and ranges between 12 GeV and 50 GeV, depending on the invariant mass of the quadruplet and it remains 50 GeV for high mass studies ($m_{4l} \geq 190$ GeV). The threshold (m_{th}) increases linearly with m_{4l} from the lower value of 12 GeV to the highest value of 50 GeV in the range 140 GeV $\leq m_{4l} \leq$ 190 GeV.

2.3. Impact Parameter and Isolation Requirements

Leptons are required to be isolated using track-based and calorimeter-based isolation discriminants. The track-based isolation discriminant, defined as the sum of the transverse momenta of all tracks inside a cone of size $\Delta\mathrm{R} = 0.3$ around the muon and $\Delta\mathrm{R} = 0.2$ around an electron excluding the lepton track, divided by the lepton p_T, must not be larger than 0.15. The calorimeter-based isolation discriminant similarly adds up the cluster E_T values used to reconstruct jets within a cone of width $\Delta\mathrm{R} = 0.2$ around the barycentre of the candidate lepton. The E_T sum divided by the p_T of the lepton is required to be less than 0.3 for muons and 0.2 for electrons. The requirement on the transverse impact parameter significance $|d_0/\sigma d_0| \leq 3$ (5) for muons (electrons) must also be satisfied [12]. Track-based isolation requirements, calorimeter-based isolation requirements, primary vertex constraint as well as the impact parameter requirements on lepton candidates help with the suppression of the sub-dominant $t\bar{t}$ and $Z + jets$ background contributions [13]. The requirement for tracks to originate from the primary vertex (defined to be a vertex with the largest p_T sum) is used to suppress contribution from pile-up [14] events.

2.4. Overlap Removal

To remove overlap between leptons of different flavors, electrons that share the same Inner Detector track with a muon are removed, unless the muon in question is a segment-tagged muon or a calorimeter-tagged muon with no track in the muon spectrometer. Jet objects that appear within $\Delta\mathrm{R} = 0.2$ of an electron or within $\Delta\mathrm{R} = 0.1$ of a muon are removed.

3. Optimization of VBF Signal

In this section, we discuss the optimization of the m_{jj} and $\Delta\eta_{jj}$ selection cuts for the four-lepton channel. The method employed in this study is based on the signal selection efficiency ($\epsilon(S)$)

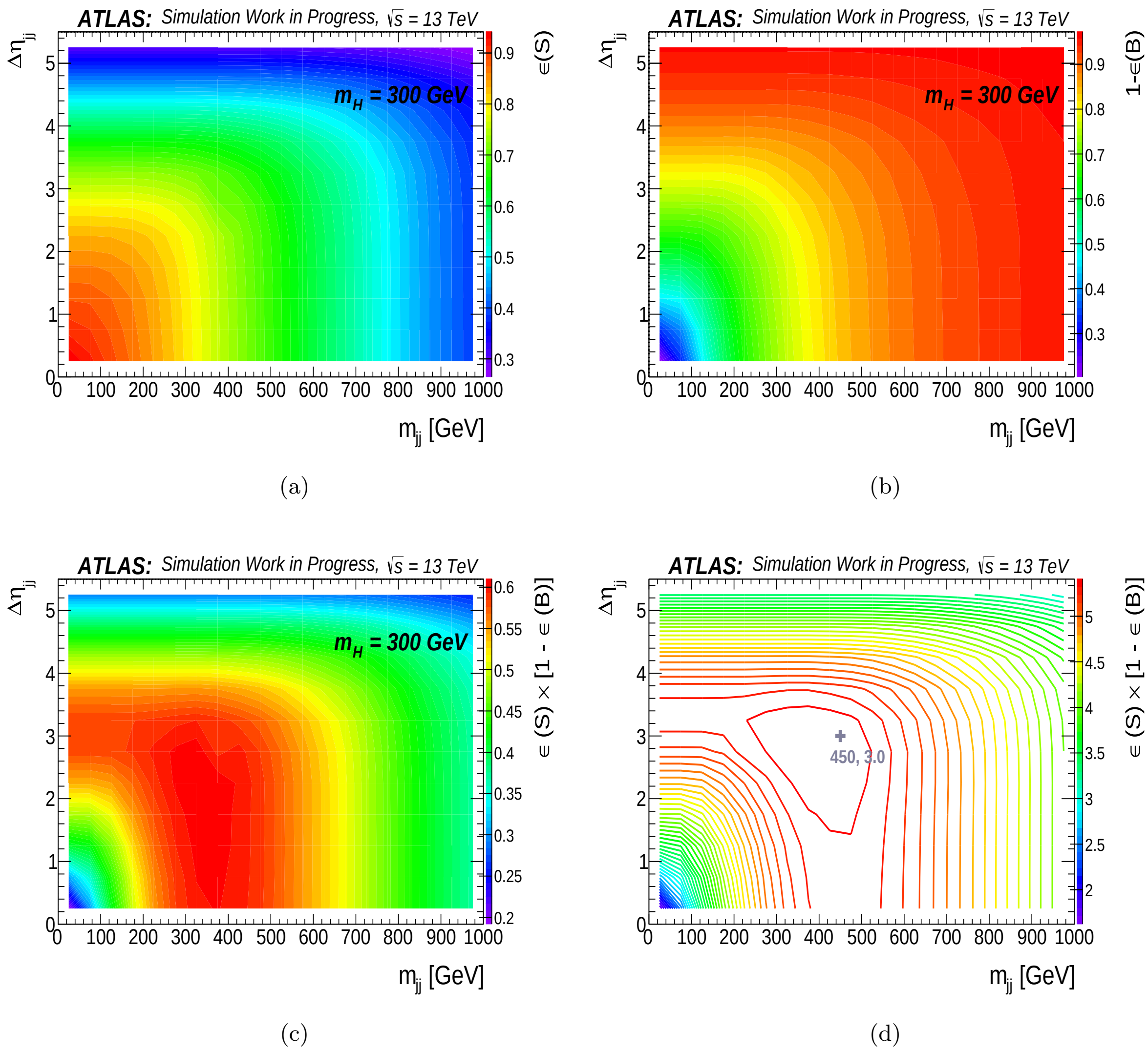
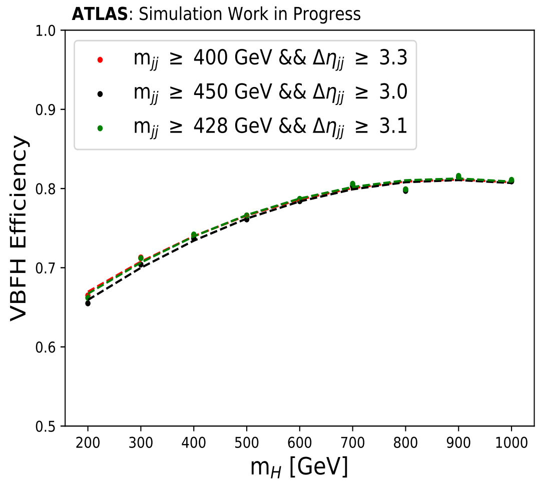
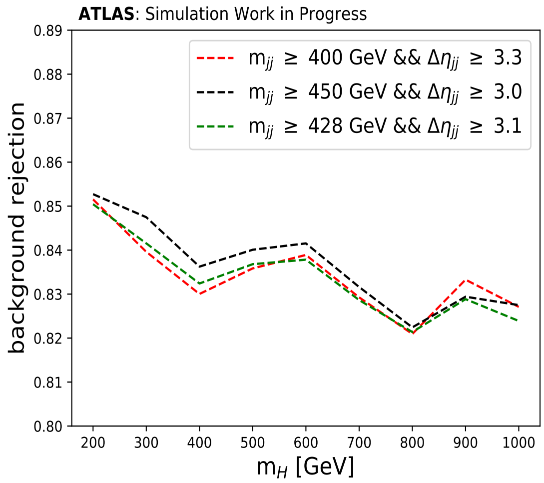


Figure 2: The (a) signal selection efficiency (ϵ(S)), (b) background rejection (1-ϵ(B)), (c) the product ϵ(S)×[1-ϵ(B)] for m_H = 300 GeV mass point and (d) ϵ(S)×[1-ϵ(B)] for all mass points from m_H = 200 GeV to m_H = 1 TeV.

and background rejection (1-$\epsilon(B)$) 2D maps. The signal selection efficiency is calculated as the ratio of the number events after the VBF selection cuts to the number of events before the VBF selection cuts were applied. One efficiency map is displayed in figure 2 for the m_H = 300 GeV mass points. To define background (ggF) rejection in the VBF signal region, a fraction of non-VBF events that are removed during the selection is calculated and it is the measure of the ability of each applied selection cut to reject ggF events in the VBF signal region. Background rejection and selection efficiency are calculated in 0.2 TeV to 1 TeV mass range. Various combinations of m_{jj} (0 GeV to 1000 GeV) and $\Delta\eta_{jj}$ (0 to 5.5) cuts are investigated in figure 2b and the optimal selection for each mass point is chosen by finding the cuts corresponding to the maximum value of the product ϵ(S)×[1-ϵ(B)]. This method provides selection cuts with good ability to reject non-VBF events and still maintain high signal selection efficiency in the given mass range. Choosing a very tight cut may result in excellent background rejection but very low signal efficiency. The opposite is also true, choosing a loose cut may result in high signal selection efficiency but poor background rejection. This trade-off can be seen clearly in figure 3 as a small increase in



(a)



(b)

Figure 3: The (a) Selection efficiency and (b) background rejection as a function of mass of the resonance for the optimal selection in the VBF category.

background rejection and a small decrease in the signal efficiency for the new cuts. The overall selection that can be used for all mass points in the mass range considered in this study can be obtained by combining all ϵ(S)$\times$[1-ϵ(B)] 2D maps. The maximum value of the product ϵ(S)$\times$[1-ϵ(B)] obtained in figure 2d reveals the new optimal selection for all nine mass points that were used. The currently used standard VBF selection requires events with two or more jets with p_T greater than 30 GeV, that are well separated in η and satisfying $m_{jj} \geq 400$ GeV and $\Delta\eta_{jj} \geq 3.3$. Comparison of the new cuts with the standard VBF selection is discussed in section 4.

4. Comparison

To test the overall performance of the VBF selection obtained in section 3, the efficiency and background rejection were calculated for each optimal selection and drawn as a function of the resonance mass. The final optimal selection is decided by comparing the new cuts with the current standard cut-based selection in the VBF category. In figure 3, a third set of cuts ($m_{jj} \geq 428$ GeV and $\Delta\eta_{jj} \geq 3.1$) is included and this was obtained by taking the average of the cut values over all mass points that are used in this study. The black line in figure 3 is the new selection obtained from section 3 by summing all ϵ(S)$\times$[1-ϵ(B)] maps, the green line is a new selection cut obtained by taking the average over all mass points and the red line shows the currently used standard cut which is used here as a reference cut-based selection. This comparison is done to look for improvements in the selection efficiency and background rejection for the chosen optimal selection with respect to the standard cuts. It can be seen in figure 3 that the current selection is still good enough to use as cut-based VBF selection with high signal selection efficiency across all mass points.

5. Summary

The categorization of events into VBF and ggF categories was performed using selection efficiency and background rejection 2D maps for ggFH and VBFH signal samples in the 0.2 TeV $\leq m_H \leq$ 1 TeV mass range. The optimization is done using Monte Carlo samples and by choosing events with two or more jets and requiring that all events passing the optimal selection be categorized into the VBF-enriched category. The remaining events are classified in any one of the three ggF categories depending on the lepton-flavour composition of the final state. It was found that the currently used selection is still good enough for use in the cut-based categorization of VBF and ggF events with full Run 2 Monte Carlo corresponding to the total integrated luminosity of 140 fb^{-1}. It is recommended therefore to keep the current cuts ($\mathrm{m}_{jj} \geq$ 400 GeV and $\Delta\eta_{jj} \geq 3.3$) as optimal cut-based selection to use for narrow resonance searches in the $H \to ZZ \to 4\ell$ channel.

References

[1] Aad G *et al.* (ATLAS) 2012 *Phys. Lett.* **B716** 1–29 (*Preprint* 1207.7214)
[2] Chatrchyan S *et al.* (CMS) 2012 *Phys. Lett.* **B716** 30–61 (*Preprint* 1207.7235)
[3] Hussein M Y 2017 Higgs boson production at the lhc (*Preprint* 1703.03952)
[4] Aad G *et al.* (ATLAS) 2015 *Phys. Rev.* **D91** 012006 (*Preprint* 1408.5191)
[5] Aad G, Abajyan T and et-al 2013 *Physics Letters B* **723** 15 – 32 ISSN 0370-2693
[6] Achard P, Adriani O and et-al 2004 *Physics Letters B* **585** 42 – 52 ISSN 0370-2693
[7] Cacciari M, Salam G P and Soyez G 2008 *JHEP* **04** 063 (*Preprint* 0802.1189)
[8] Cacciari M, Salam G P and Soyez G 2012 *Eur. Phys. J.* **C72** 1896 (*Preprint* 1111.6097)
[9] Aaboud M *et al.* (ATLAS) 2017 *Eur. Phys. J.* **C77** 466 (*Preprint* 1703.10485)
[10] Aaboud M *et al.* (ATLAS) 2017 *Eur. Phys. J.* **C77** 580 [Erratum: Eur. Phys. J.C77, no.10,712(2017)] (*Preprint* 1705.02211)
[11] Tomiwa K 2017 *Journal of Physics: Conference Series* **802** 012012
[12] Aaboud Morad and others (ATLAS) 2017 *JHEP* **10** 132 (*Preprint* 1708.02810)
[13] Sirunyan, Albert M and et-al (CMS) 2017 *JHEP* **11** 047 (*Preprint* 1706.09936)
[14] Marshall Z 2014 *Journal of Physics: Conference Series* **513** 022024