The Higgs as a portal to the hidden sector via an analysis of $H \to Z_d Z_d \to 4l$ using the ATLAS detector

D A Unwuchola[1,2]
[1] University of Johannesburg, Johannesburg, South Africa
[2] On behalf of the ATLAS Collaboration

E-mail: `doomnull@gmail.com`

Abstract. The Standard Model (SM) has well known deficiencies, and there is clearly need for new physics beyond the SM. The particles manifesting the new physics would interact at most weakly with the SM particles, and hence they are termed dark. The Higgs boson is potentially a favourable route for the production of the dark particles. There are a large class of theories where couplings or mixings at the Higgs level leads to exotic Higgs decays, which nonetheless do not significantly disturb the known physics below the Higgs level. This is therefore a significant potential discovery opportunity. We present the motivation and progress made in the studies which have been carried out as part of designing the search for the exotic decay of the SM Higgs which proceeds via a dark force back to SM four leptons, $H \to Z_d Z_d \to 4l$ from the LHC run 1 data using the ATLAS detector.

1. Introduction

Beyond Standard Model (BSM) phenomenology presents the Higgs boson as a sensitive discovery channel for particles that couple weakly to Standard Model (SM) particles. The potential for a new sector to exist without disturbing the existing good agreement between the SM and all previous data including LHC data arises because the new sector primarily couples to the SM at the Higgs level. This reflects the universality of the Higgs interaction to many theoretical incarnations of ”new Physics”. The Higgs therefore provides a connection to new physics not charged under SM forces. In addition, the branching ratio of the Higgs in the SM is narrow and the branching ratio of the Higgs to exotic particles can be arranged within the details of the models to be relatively unconstrained by the existing data-theory agreements. This accommodates a significant fraction of its total decay width being available for exotic decays. In a very extensive article, Curtin et al [1] argue that this is therefore an extremely significant window to BSM Physics at the LHC. This motivates the term ”Higgs portal” for this type of access to the discovery of possible new dark particles. Essentially it will manifest via an exotic Higgs decay and can be directly searched for. The phenomenolgy is very rich, and we have chosen a final state signature whereby the discovered Higgs is actually the lightest partner of a Higgs multiplet that arises from the mixing between the BSM dark sector and the SM Higgs. This new Higgs decays via a dark force particle in analogy to the SM, a new dark and light neutral boson, Z_d, which in turn decays back to SM leptons. The coupling of the Z_d to SM leptons is set at the maximum value accommodated by the current agreement between theory and experiment. It is further possible, that even in this case, the phenomenology can be arranged such that the branching ratio of the $Z_d \to 4l$ is essentially 100%. Our discovery signature is therefore $H \to Z_d Z_d \to 4l$. In the rest of the paper, we present an overview of the analysis procedure. The results which include data are will be presented in a future paper.

2. Analysis

The analysis for the exotic Higgs decay parallels that of the four lepton Higgs discovery channel, $H \rightarrow ZZ^* \rightarrow 4l$, since they share the same primary and final states. The major difference arises from the fact that in the intermediate state, we have an unknown Z_d mass, and so we replace the cut for the on-shell Z by a requirement that both the Z_d bosons are on shell so that their masses are equal. There are of course some other differences, some of which are mentioned below.

2.1. Pre-selections, Cuts

The event and object selections, kinematic requirements and trigger matching are the same as in the four lepton Higgs discovery analysis using the ATLAS detector. The isolation requirements for each lepton within the detector, the separation between a lepton pair (ΔR) and the significance for the reconstruction of the interaction point (IP) are also the same. The reader is referred to the Higgs discovery paper [2] and references therein for details. These similarities are shown in green in the table reproduced as figure 1.

A subtlety arises in the consideration of which association to use in the reconstruction of the bosonic parents for the lepton pairs in the final state quadruplet. When all four leptons have the same flavour (but of course there are two opposite sign pairs), then there is an ambiguity in the association of the two di-lepton pairs. This is resolved in the standard analysis by the requirement that one of the dilepton pairs reconstructs to give the correct mass of the SM Z. In this exotic decay analysis, the pairs have the association chosen to minimise the mass difference in the two reconstructed Z_d bosons, which are both required to be on-shell.

2.2. Definition of the Signal Region

The cuts defining the Signal Region (SR), are also different between the exotic analysis and the standard analysis. One can note in the table the following cuts - selection of quadruplets consistent with the discovered Higgs as a parent, $100 < m4l < 150$ GeV, the requirement for the di-lepton pairs to reconstruct to an on-shell Z_d, $m4l < 63$ GeV and the requirement that each reconstructed Z_d has the same mass, $\Delta m < 15$ GeV. The cuts have boundaries which are the results of studies to optimise them, and which recognise systematic effects in the ATLAS detector. The Δm is in fact still being studied, so the value given here is nominal (see section 2.4).

The consideration of various processes which may masquerade as signal, has led to the introduction of vetoes arising from Zs, J/Ψs and Υs. Cuts are applied in these cases on the di-lepton events at a late stage in the analysis. As the association procedure of dilepton pairs mentioned above may be ambiguous for these backgrounds, the veto is applied if any of the two di-lepton associations falls foul of it.

These differences in the SR definition are shown in red in the same table. The analysis code can be switched to be consistent with either the standard or the exotic analysis. The cut-flow (measurements of the cut efficiency at each stage) using the analysis code configured for the standard analysis was compared to the cut-flow as established for the initial discovery and current study of the discovered Higgs, and the excellent agreement here is a partial benchmark for validating the analysis code to be deployed in the Z_d analysis.

2.3. Monte Carlo study of the signal and background samples

The process $H \rightarrow ZZ^* \rightarrow 4l$ of the standard analysis, which is the signal for the SM Higgs discovery, is actually a background for the Z_d analysis. In addition electroweak vector boson pair production from $q\bar{q}$ annihilation and loop-induced gluon fusion, $ZZ \rightarrow 4l$ is a further irreducible background. In these cases there is a sizeable correlated production of objects that can manifest in the Z^*Z^* region. These processes are well modelled by Monte Carlo (MC). Other processes considered for the contribution to background include the production and decay of $t\bar{t}$,

Table 1. A summary of the cuts applied in four lepton Higgs discovery channel $H \rightarrow ZZ^* \rightarrow 4l$ and the current analysis $H \rightarrow Z_d Z_d \rightarrow 4l$. The table is explained in the text.

Cuts	$H \rightarrow ZZ \rightarrow 4l$	$H \rightarrow Z_d Z_d \rightarrow 4l$
4 l, kinematics, trigger matching	SAME	SAME
Primary pair mass (m_{12})	$50 < m_{12} < 106$	No cut!
Secondary pair mass (m_{34})	$X < m_{34} < 115 (X = f(m_{4l}))$	No cut!
1 unique quadruplet	m_{12} closest to m_z	$\lvert\Delta m\rvert = \lvert m_{12} - m_{34}\rvert$ minimal
$\Delta R > 0.1[0.2]$ for SS [OS] pairs	SAME	SAME
track and call isolation	SAME	SAME
IP significance	SAME	SAME
Signal Region		$\lvert\Delta M\rvert < 10$ GeV $100 < m_{4l} < 150$ GeV $m_{12}, m_{34} < 63$ GeV Z veto

$Zb\bar{b}$ and Z+jets and which produce leptons or jets faking leptons. The J/Ψ and Υ have also been considered as they can be correlated with a Z through the associated quarkonium production mechanism from double parton scattering. These backgrounds have been also modelled in MC, however the statistics available are less than optimal. The Parton Distribution Functions (PDF) sets on which primary scatterings are based, the event generators for the processes mentioned and the description of the full ATLAS simulation based on the GEANT4 framework are mostly described in reference [2].

The signal has been modelled using a particular incarnation of a Higgs Portal model in the $H \rightarrow Z_d Z_d \rightarrow 4l$ exotic decay signature, the Wells Model [3, 4]. This is done at the generator level in order to represent the initial lepton quadruplet kinematic distributions and angular correlations. The actual signal rate has been scaled to the SM rate for $H \rightarrow ZZ^* \rightarrow 4l$ simply for convenient comparison. The use of a model at generator level allows a study of backgrounds and cut optimisation. The final analysis proceeds in some sense rather more model independently, as it is a limit setting procedure based on testing for an excess over the well studied backgrounds in the SR. The left and right parts of figure 1 and the left part of figure 2 indicate the contribution of the signal and the various background distributions viewed as a function of the reconstructed mass for the full lepton quadruplet, the two di-lepton pairs and the mass difference between the two di-lepton pairs. The signal is shown for 2 masses at the limits of our SR for the Z_d. The rest of the description of these plots can be found in the captions. Only the case for the $4e$ final state is shown. There are of course subtleties associated with the cut efficiency and MC distribution validation and therefore the treatment of each different final state, even if in general, there are similarities. These distributions are frozen after the IP level cut, just before the SR cuts are applied. In the di-lepton case of figure 1 at the left, in the case of the ZZ background, a peak is visible at the SM Z boson mass, and a tail at lower mass comes from events containing an off-shell Z boson.

In the case of the lepton quadruplet of figure 1 at the right, the peak at 125 GeV corresponds to $H \rightarrow Z_d Z_d \rightarrow 4l$ events and also to $H \rightarrow ZZ^* \rightarrow 4l$ events. The $ZZ^* \rightarrow 4l$ background is mostly confined to $m4l > 180$ GeV (two on-shell Z bosons). At the region of the SM Z mass we see a peak from the s-channel single on-shell Z production where one of the decay leptons emits an off-shell Z. It is clear that this process still reconstructs to a single Z mass in the $m4l$ spectrum. In the case of the mass difference between the two di-lepton pairs of figure 1 at the left, we see as expected a signal peak in the region of low mass difference, and also a peak in the ZZ background.

In the case of the right section of figure 2, a 2D view of the MC signal and background is provided, which further illuminates the SR cuts to be applied. Further description of the figure is provided in the caption.

The SR cuts mentioned above clean the background in the SR region very effetively, so that it is dominated by the signal. Evidence for this is not shown here, due to space limitations.

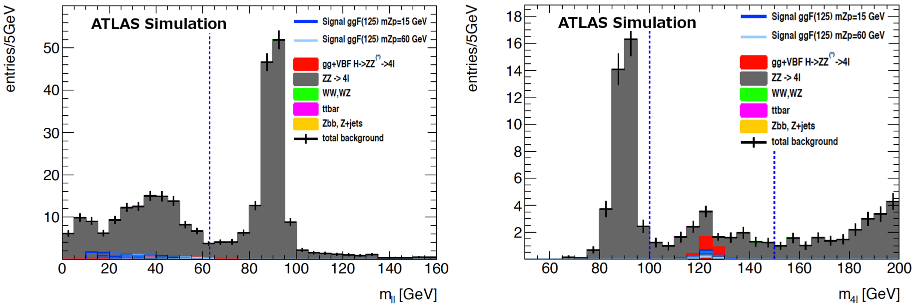

Figure 1. Left : Dilepton invariant mass for all signal and background samples, in the $4e$ final state. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement symbolized by the blue dashed line at 63 GeV. Right : Four-lepton invariant mass for all signal and background samples, in the $4e$ final state. The blue dashed lines show the $100 < m4l < 150$ GeV cut applied on this variable to define the Signal Region. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement.

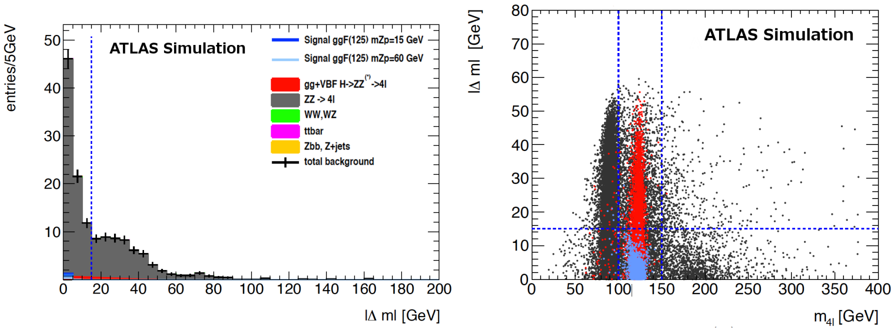

Figure 2. Left : Absolute mass difference for all signal and background samples, in the $4e$ final state. The blue dashed line shows the cut applied on this variable to define the Signal Region. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement. Right : Mass difference Δm as a function of the four-lepton invariant mass $m4l$ for background $ZZ^* \to 4l$ in dark gray, $H \to ZZ^* \to 4l$ in red) and signal $H \to Z_d Z_d \to 4l$ with $mZ_d = 50$ GeV, in blue) events, in the $4e$ final states. The blue dashed lines show the limits of the loose Signal Region : $\Delta m < 15$ GeV and $100 < m4l < 150$ GeV.

2.4. Background estimation and Systematics

The $H \to ZZ^* \to 4l$ and $ZZ \to 4l$ processes are estimated by MC simulation and subtracted. Other di-boson process such as WW and WZ which have survived the cuts are also well modelled and can be removed based on MC simulation. The remaining $t\bar{t}$, $Zb\bar{b}$ and Z+jets are estimated with a data driven method, known as the ABCD method. The ABCD method is applied to extrapolate the remaining backgrounds from a more general Control Region (CR) into the SR. In the ABCD method, two uncorrelated variables (observables) are selected. Passing or failing a cut on these variables defines a SR and CR for each of them. In the 2D product space of these two variables, one will then have a single SR region quadrant and three other CR region quadrants. The backgrounds to be modelled by MC are first removed, and the remaining events (data or MC depending on whether its analysis or a study) will be distributed in these four quadrants. The quadrant representing the SR in both variables can then be estimated by considering that the ratios of events in either horizontally or vertically neighbouring quadrants must be equal (assumption of non-correlation in the two variables). In practice, the non-correlation can be estimated from MC, and this can be applied as a correction. This analysis has now converged on the selection of the two uncorrelated variables as the isolation requirements on the each dilepton pair (the two leptons of the pair pass or fail the cut combining the track-based isolation, calorimeter-based isolation and impact parameter significance requirements). The method is applied to the SR as defined for all cuts. A final procedure to optimise the Δm cut, based on achieving the best possible exclusion limit, is still in process.

The limit setting procedure for discovery or exclusion is based on the best possible knowledge of the statistical and systematic uncertainties in all quantities involved in the quantitative understanding of the data. In the case of the systematic uncertainties, theoretical systematics contribute in the MC evaluation process. These are related to the EW and QCD models themselves as well as to the PDFs which lead ultimately to uncertainties on the cross sections. Then there are the detector systematics relevant to the electron and muon identification efficiency and uncertainty on the signal yields. There is the luminosity uncertainty which affects the scaling of the MC yields and the evaluation of the signal yields. There is also the data driven systematics, as applicable to the ABCD method. This consists in the estimates of the non-correlation and to the uncertainties due MC statistics in the method development. This process is almost completed for this analysis.

Table 2. Acceptance times efficiency for signal and background in the Signal Region (defined with $|\Delta m| < 15$ GeV), for the $4e$ final state. The expected number of signal events is normalized to the SM $H \to ZZ^* \to 4l$ rate in the $4e$ channel.

Process	Events processed	Events in the SR	Acc. × eff.	Expected events ($20.7 fb^{-1}$)
Signal ($m_{Z_d} = 20$ GeV)	29908	1674.5	5.6%	6.16
Signal ($m_{Z_d} = 50$ GeV)	29939	2054.8	6.86%	7.56
(gg+VBF) $H \to ZZ^* \to 4l$	119.11	0.42	0.36%	0.42
$ZZ \to 4l$	$1.81 \cdot 10^5$	0.12	$6.44 \cdot 10^{-5}$%	0.12
WW, WZ	60170	0	0%	0
$t\bar{t}$	$5.50 \cdot 10^5$	0.037	$6.65 \cdot 10^{-6}$%	0.037
Zbb, Z+jets	$7.14 \cdot 10^7$	0	0%	0
(Z+) low mass	55577	0	0%	0

2.5. Derivation of a limit for the case of exclusion or discovery.

Table 2 presents the signal and background events expected for the $4e$ channel in the SR where the expected number of signal events is normalized to the SM $H \to ZZ^* \to 4l$ rate. The data can also be processed by the analysis code and the number of data events in each SR bin can be obtained. It will then be possible to compare the expected background and the data, considering the uncertainties on each.

The signal strength μ, is defined as the ratio of the $H \to Z_d Z_d \to 4l$ rate relative to the SM $H \to ZZ^* \to 4l$ rate:

$$\mu = \frac{\sigma \times BR(H \to Z_d Z_d \to 4l)}{[\sigma \times BR(H \to ZZ^* \to 4l)]_{\mathrm{SM}}} \tag{1}$$

A quantitative and robust interpretation of this data as limits for the case of exclusion will be computed from a maximum likelihood fit to the numbers of events in the various signal regions following the CL_s modified frequentist formalism [5, 6] with the profile likelihood test statistic [7].

As an indication of the type of result that can be expected, the result of Curtin based on the analysis in reference 3 is presented in figure 3.

The case for discovery is closely related and not discussed further here.

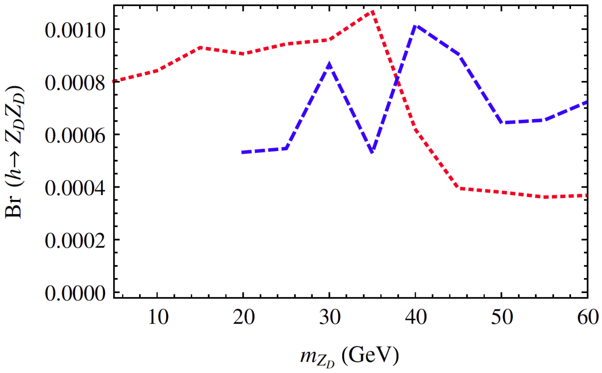

Figure 3. A 95% CL exclusion contour for the exotic Higgs decay $H \to Z_d Z_d$ based on the assumption of $BR(Z_d \to 2l) = 30\%$ taken from Curtin [1] figure 30. In the case of this anlysis, the limit will be presented on the excluded signal strength μ, defined as the ratio of the $H \to Z_d Z_d \to 4l$ rate relative to the SM $H \to ZZ^* \to 4l$ rate. This diagram is very important to be referenced as it gives a good picture of what our result would look like.

3. Conclusion

In conclusion, we have presented the motivation for a study of the exotic Higgs decay $H \to Z_d Z_d \to 4l$. This is based on using the discovered Higgs as a portal to new physics BSM. The analysis is well progressed, and has been discussed to the stage of limit setting, in the case of either exclusion or discovery.

Acknowledgments

The authors acknowledge the support of the National Research Foundation (NRF) and Department of Science and Technology, both of South Africa, who have supported the research though the SA-CERN progamme. The NRF contributed to bursaries and the NRF and the Universities of Johannesburg and Cape Town contributed to post doctoral fellowships.

References

[1] Curtin D, Essig R, Gori S, Jaiswal P, Katz A, Liu T, Liu Z, McKeen D, Shelton J, Strassler M, Surujon Z, Tweedie B and Zhong Y M 2014 *Phys. Rev. D* **90**(7) 075004 URL `http://link.aps.org/doi/10.1103/PhysRevD.90.075004`

[2] The ATLAS Collaboration 2012 *Physics Letters B* **716** 1 – 29 ISSN 0370-2693 URL `http://www.sciencedirect.com/science/article/pii/S037026931200857X`

[3] Kane G and Pierce A (eds) 2008 *Perspective on LHC Physics* (World Scientific) ISBN 978-981-277-975-5

[4] Gopalakrishna S, Jung S and Wells J D 2008 *Physics Review D* **78** 055002

[5] Read A L 2002 Presentation of search results: The cl(s) technique Tech. Rep. J.Phys.G G28 (2002) 2693 – 2704

[6] Cousins R D and Highland V L 1992 *Nucl. Instrum. Meth.* **A 320** 331–335

[7] Cowan G, Cranmer K, Gross E and Vitells O 2011 *Eur. Phys. J.* **C 71** 1554 (*Preprint* `1007.1727`)