Background estimation for multilepton and b-jets analysis at ATLAS at the LHC

Jeremiah Monnakgotla[1] **, Lebohang Mokoena**[1]**, Lehumo Mashishi**[1]**, Yesenia Hernández Jiménez**[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, PO Box 722, Somerset West 7129, South Africa

E-mail: `jeremiah.kgomotso.monnakgotla@cern.ch`

Abstract. Background estimation is one of the most important aspects for all the analysis at the Large Hadron Collider. This proceeding presents the background estimation for the $A \to ZH$ search with the heavy scalar H decaying into a pair of Higgs, or Higgs-like Scalar bosons. The final state for this search is formed by three leptons, at least two b-tagged jets and low jet multiplicity. The $t\bar{t}Z$ process is the dominant background contribution for this analysis. The second important background is the WZ for which a control region is defined to check the modeling and the normalization of this process.

1. Introduction

This work performs a background estimation for the Madala hypothesis [3] in the multi-lepton and b-jets production at the LHC. This hypothesis considers a production of a CP-odd scalar A which decays to H and the Z bosons leaving a final state of three leptons and at least two b-tagged jets and low jet multiplicity [4]. A wide range of analysis have been performed by ATLAS and CMS collaborations observe discrepancies between data and Monte Carlo (MC) which can be explained by the Madala hypothesis. Firstly in CMS, a supersymmetry search with the production of $t\bar{t}Z$ decay with three lepton and b-tagged jets was used as one of the control regions (CRs) for background normalization [1]. Secondly, the measurement of the $t\bar{t}Z$ cross section observe a discrepancy at low jet multiplicity [2]. Furthermore, the observed discrepancies between data and MC has opened a window for searches Beyond Standard Model (BSM).It is clear that the LHC data is not well understood. With that in view this proceeding describes the background estimation for the $A \to ZH$ search. In Addition, the technique of control region (CR) is implemented by inverting one of requirements in the signal region (SR).

2. Signal Model

In the model depicted in Figure 1, a new heavy scalar boson decays into the heavy scalar (H) which subsequently decays into a pair of Higgs, or Higgs-like scalar (S) bosons leaving a final state of three leptons and at least two b-jets.Therefore the S and the h decay to WW bosons (decay semi-leptonically) and two b-tagged jets. The Z boson decays to a pair of leptons and h is the standard model (SM) Higgs boson with the mass of 125 GeV. The masses of the H and

S are assumed to be 250 GeV and 145 GeV respectively as discussed in ref [[3, 4]]. The mass of CP odd-Higgs A is expected to be greater than sum of the masses of H and Z boson.

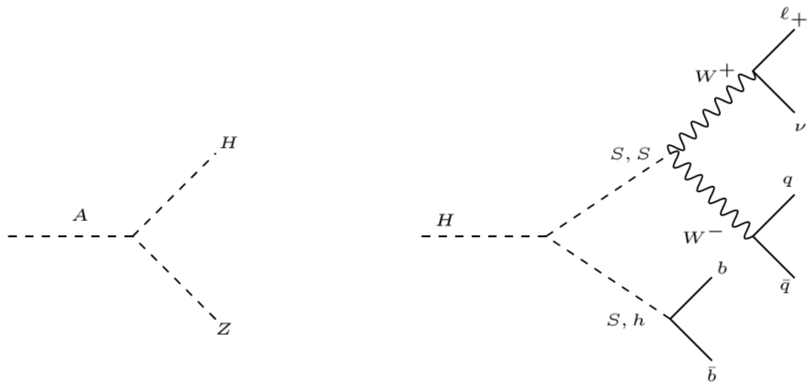

Figure 1. The Feynman diagram of the $A \to ZH$ production (left) and of Higgs and Higgs-like scalar bosons.

3. Signal Region Definitions

In this analysis two different SRs are defined to be $3l - Z{-}2b2$j and $3l - Z{-}2b3$j. The final state is characterized by events containing exactly three leptons with the same flavour lepton pair closest to the Z mass, at least at least two b-tagged jets and two or three jets. it can be clearly seen that the signal sensitivity for $A \to ZH$ is high for events having exactly two or three jets [4]. Events should contain at least one pair of leptons with same-flavor and opposite sign (SFOS). In addition, the invariant mass of the SFOS pair ($m_{l_{\pm}l_{\mp}}$) must be compatible with the Z boson mass (m_Z) in a window of 20 GeV. The leading lepton is required to have a transverse momentum (P_{T}) above 27 GeV while for other two leptons this threshold is lowered to 20 GeV. The three leptons have a total charge of ± 1.The exact signal definitions and event selection are summarized in Table 1

Table 1. Summary of the event selection in the signal region.

Variable	signal region $3l - Z{-}2b2$j	signal region $3l - Z{-}2b3$j
Number of Leptons	3	
Leading Lepton P_T	> 27 GeV	
Second and Third Lepton P_T	> 20 GeV	
Lepton Total Charge	± 1	
One SFOS Lepton Pair	required	
Z-like SFOS Pair	$\lvert m_{\ell_{\pm}\ell_{\mp}} - m_Z \rvert < 10$ GeV	
$N_{b-tagged\ jets}$	≥ 2	≥ 2
N_{jets}	$= 2$	$= 3$

4. Analysis Strategy

The background estimation is evaluated by defining CRs to be orthogonal to the SRs. The SRs are set to be characterized by events with low jet multiplicity, therefore the CRs will be expected to contain events with high jet multiplicity. The jet multiplicity distribution in Figure 2 with events containing at least two b-tagged jets, clearly illustrates the SR and the CR.

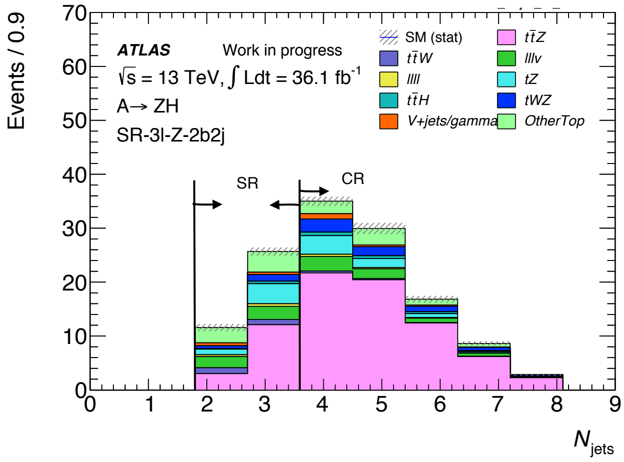

Figure 2. Jet multiplicity distribution after the selection of at least two b-tagged jets.

5. Background Estimation

In the three leptons final state the SM background arises from $t\bar{t}Z$, WZ and other top ($t\bar{t}$) processes. These SM backgrounds are predicted using MC simulation which are commonly known as the irreducible background. A strategy is developed to estimate these background processes which involve the construction of CRs which depends on the requirement of the SR. Henceforth four CRs are constructed to evaluate the dominant background processes and to observe events which are topologically similar to the signal events. The first three CRs required SFOS pair whose mass is closest to the Z boson mass.

5.1. Control region for the $t\bar{t}Z$

The $t\bar{t}Z$ background is particularly important in the three leptons analysis, and it is estimated by the use of two CRs which are labeled as: $3l - Z{-}1b4$j and $3l - Z{-}2b4$j. The former requires at least four jets with exactly one b-tagged jet. The later requires at least four jets with at least two b-tagged jets. With these requirement the $t\bar{t}Z$ process contribute 70 % and 40 % with respect to each CR. The distributions of the invariant mass of the lepton pair are shown in Figure 3 and Figure 4.

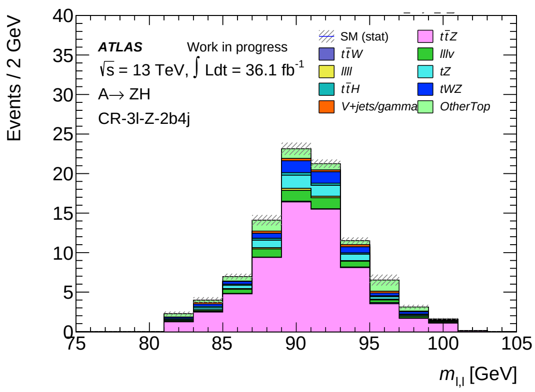

Figure 3. Invariant mass of Z-like lepton pair after the selection of at least four jets with at least two b-tagged jets.

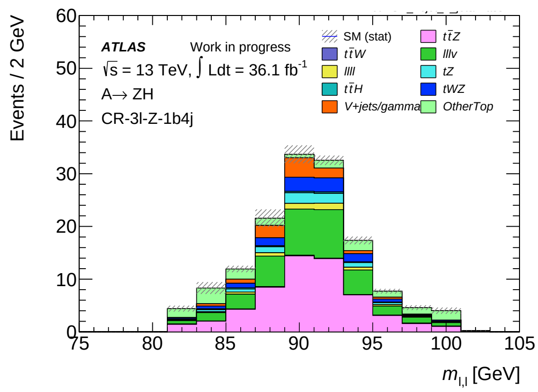

Figure 4. Invariant mass of Z-like lepton pair after the selection of at least four jets with exactly 1 b-tagged jet.

5.2. Control Region for the WZ

The third CR which is sensitive to the WZ process is labeled as: $3l - WZ$. This CR select events with at least three jets and zero b-tagged jets. In addition, missing transverse momentum is required to be greater than 40 GeV. With these requirement the WZ contribution is roughly 80.1% of the total background. A very small percentage arise from the V+jets/γ and ZZ processes as shown in Figure 5 and Figure 6.

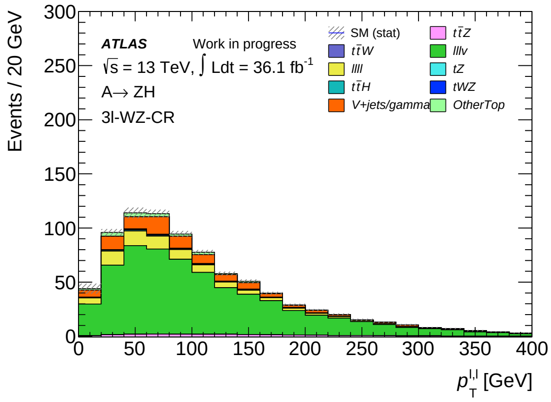

Figure 5. P_T of Z-like lepton pair after selecting at least three jet with zero b-tagged jet.

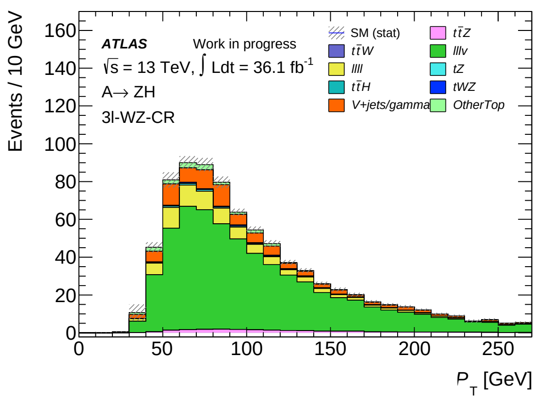

Figure 6. P_T leading lepton after selecting at least three jet with zero b-tagged jet.

5.3. Control Region for the Z mass veto

The $3l - noZ - 2b4$j region select events with exactly three leptons, at least four jets, at least two b-tagged jets and veto events containing $Z-$like lepton pair. In Figure 7 the leading lepton P_T illustrates that the other top process dominates, the contribution of $t\bar{t}Z$ is highly reduced since the events associated with Z boson decaying leptonically are rejected. With these requirements the other top ($t\bar{t}$) contribute roughly 50.2% of the total background, where other top ($t\bar{t}$) is the combination of $t\bar{t}$, 3*top* and 4*top* SM processes but dominated by $t\bar{t}$ process. The $t\bar{t}Z$ background contribute approximately 24.8% of the total background. Table 2 show the expected numbers of events in the CRs.

Table 2. The expected event yields in the CRs for an integrated luminosity of 36.1 fb^{-1}. Uncertainties are statistical only.

	$3l - Z{-}1b4$j	$3l - Z{-}2b4$j	$3l - Z{-}0b3$j	$3l - noZ{-}2b4$j
$t\bar{t}Z$	57.32 ± 0.57	64.13 ± 0.61	20.83 ± 0.33	13.44 ± 0.29
$t\bar{t}W$	0.80 ± 0.08	0.75 ± 0.09	0.47 ± 0.06	4.60 ± 0.21
WZ	36.99 ± 1.25	6.01 ± 0.44	616.42 ± 5.28	0.82 ± 0.15
ZZ	4.57 ± 0.21	0.91 ± 0.10	81.84 ± 0.95	0.35 ± 0.09
tZ	7.66 ± 0.20	6.09 ± 0.18	6.03 ± 0.17	0.54 ± 0.05
$t\bar{t}H$	1.63 ± 0.10	1.80 ± 0.10	0.51 ± 0.05	5.46 ± 0.19
tWZ	10.91 ± 0.51	5.92 ± 0.40	6.83 ± 0.39	0.76 ± 0.13
V+jets/γ	10.32 ± 2.18	1.67 ± 0.33	94.55 ± 8.39	0.53 ± 0.14
$OtherTop$	16.11 ± 1.66	7.22 ± 1.24	19.14 ± 1.39	25.92 ± 1.85
Total Background	146.30 ± 3.13	94.51 ± 1.56	846.62 ± 10.06	52.41 ± 1.91

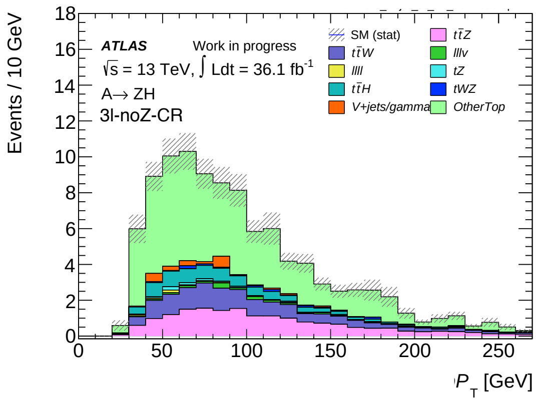

Figure 7. Leading lepton P_T after the selection of at least four jets with at least two b-tagged jets.

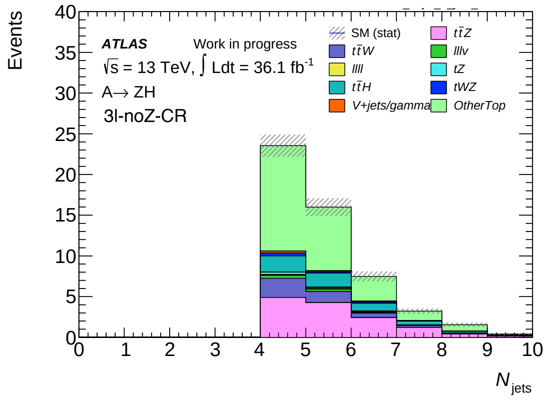

Figure 8. Jet multiplicity distribution after the selection of at least four jets with at least two b-tagged jets.

6. Conclusion

The proceeding presents the background estimation for the $A \to ZH$ search, which is performed by constructing four CRs with their requirements orthogonal to the SRs. The main background processes are the $t\bar{t}Z$, WZ and "other top" ($t\bar{t}$). The analysis will be performed with the full Run 2 data set. The next step is to check the performance of the MC for different SM processes and extrapolate their contributions to the SR.

7. References

[1] Sirunyan AM 2018 *Phys. Rev.* **D97** 9 Phys. Rev. D 97, 032009 (2018) [hep-ex]

[2] Sirunyan AM 2018 *JHEP* **08** 9 JHEP 08 (2018) 011 [hep-ex]

[3] von Buddenbrock S Chakrabarty N Cornell AS et al 2015 5 arXiv:1506.00612 [hep-ph]

[4] von Buddenbrock S Chakrabarty N Cornell AS et al 2016 *Eur. Phys. J. C* **76** 16 Eur. Phys. J. C (2016) 76: 580