Search for a resonance in the diphoton plus b-jet final states in the ttH and bbH production

Esra Mohammed Shrif[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, PO Box 722, Somerset West 7129, South Africa

E-mail: shrif.esra@gmail.com

Abstract. We propose a simple analysis to search for resonance in the diphoton and at least one b-tagged jet final states with ttH and bbH models. This region has never been checked and needs to be checked now in run II and in the future run III data. We search for the new resonance in the range above 160 GeV. The analysis uses Run II proton-proton (pp) collision data with an integrated luminosity of 140 fb^{-1} recorded at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with the ATLAS detector. In this work, we discuss event selections and signal optimization. In addition, we compare data to MC simulation in the control region. A 2D scan of the variables, number of central jets N_{cjet} and $\Delta\phi_{\gamma\gamma}$ is made to optimize the search.

1. Introduction

The diphoton channel provides a clean experimental signature with an invariant mass resolution that can be reconstructed with high precision [1, 2]. The Higgs to diphoton decay $h \to \gamma\gamma$ channel was one of the most important channels that led to the discovery of the Higgs boson [3, 4]. Recently, the unconfirmed resonance with mass around 750 GeV also observed in the diphoton channel [5, 6, 7]. An excess in the diphoton spectrum at high-energy, could be interpreted as the decay of a hypothetical particle with spin-0 or spin-2.

The searches for new high-mass resonances decaying into two photons was performed, using CERN Large Hadron Collider (LHC) pp collision data recorded by the ATLAS and CMS detectors since run I [8, 9]. This proceeding presents a search for a resonance in the diphoton plus b-tagged jet final states. In this channel, the QCD background is highly suppressed. These final states can be the result from top or bottom quark associated productions. We search for the diphoton resonance in high mass region, through the $gg/q\bar{q} \to t\bar{t}H$, $b\bar{b}H$ process, examples of tree-level Feynman diagrams are given in Figure 1. Where, H is Higgs like spin zero particle with a mass range [160 GeV, 1 TeV] decaying to diphoton final state in this study. To understand these final states, for instance, the top quark, t, decays to a W^+ boson, and a quark $q = b, s$ or d. Here, W^+ boson decay into a lepton and a neutrino, $t \to W^+q \to \ell^+\nu q$. The case for anti-top quark decay is $\bar{t} \to W^-q \to \ell^-\nu\bar{q}$, where, $\bar{q} = \bar{d}, \bar{b}$ or $\bar{d}$. These quarks will turn into jets as they cannot exist freely. The plan is to search for any resonance of diphoton and setup limits on the production cross-section times branching ratio ($\sigma\times$BR($H \to \gamma\gamma$)) to the ttH and bbH models with $m_H > 160$ GeV.

 ISBN: 978-0-620-88875-2

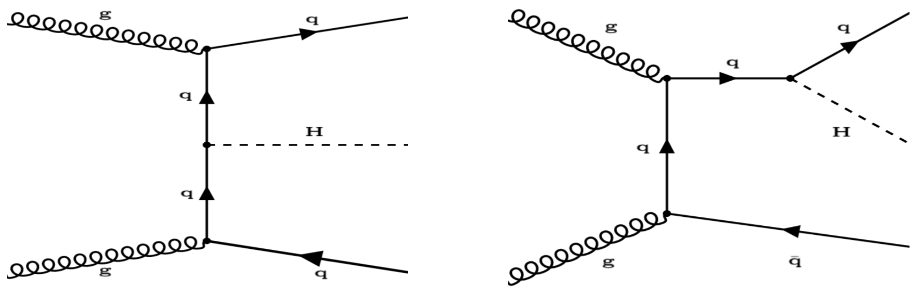


Figure 1. Feynman diagrams show two different production processes of the heavy boson H with a pair of top or bottom quarks, q refers to either top or bottom quark.

2. Samples and Event Selection

Monte Carlo (MC) simulated events are used for the estimation of signal and background processes. The MADGRAPH aMC@NLO [10] event generator interfaced to PYTHIA8 [11] is used to produce the ttH signal samples, while the bbH samples are simulated using PYTHIA8. The dominant background samples that are used here are the SM diphoton production and γ+jet; contributions also come from $bb\gamma\gamma$, $V\gamma$ (W/Z), $V\gamma\gamma$, $tt\gamma\gamma$ and $tt\gamma$ (one fake photon from electron). The γ+jet and diphoton samples are normalised to 20% and 80% respectively of the data at the inclusive case. The dataset used in this analysis encompasses all data that was recorded with the ATLAS detector, between 2015 and 2018, in pp collisions with an integrated luminosity of $140\,\text{fb}^{-1}$ at a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

After an event passes the quality selection (after selecting two well identified photons), it has to pass additional kinematic and geometrical requirements to be considered for the analysis. The event are selected with at least two photons, the transverse momentum p_T of the leading and sub-leading photon should be greater than 40 GeV, and 30 GeV, respectively. The invariant mass of the diphoton system $m_{\gamma\gamma} > 130$ GeV. It is required to have at least one b-tagged jet. To avoid the SM Higgs peak events are then categorised into three regions:

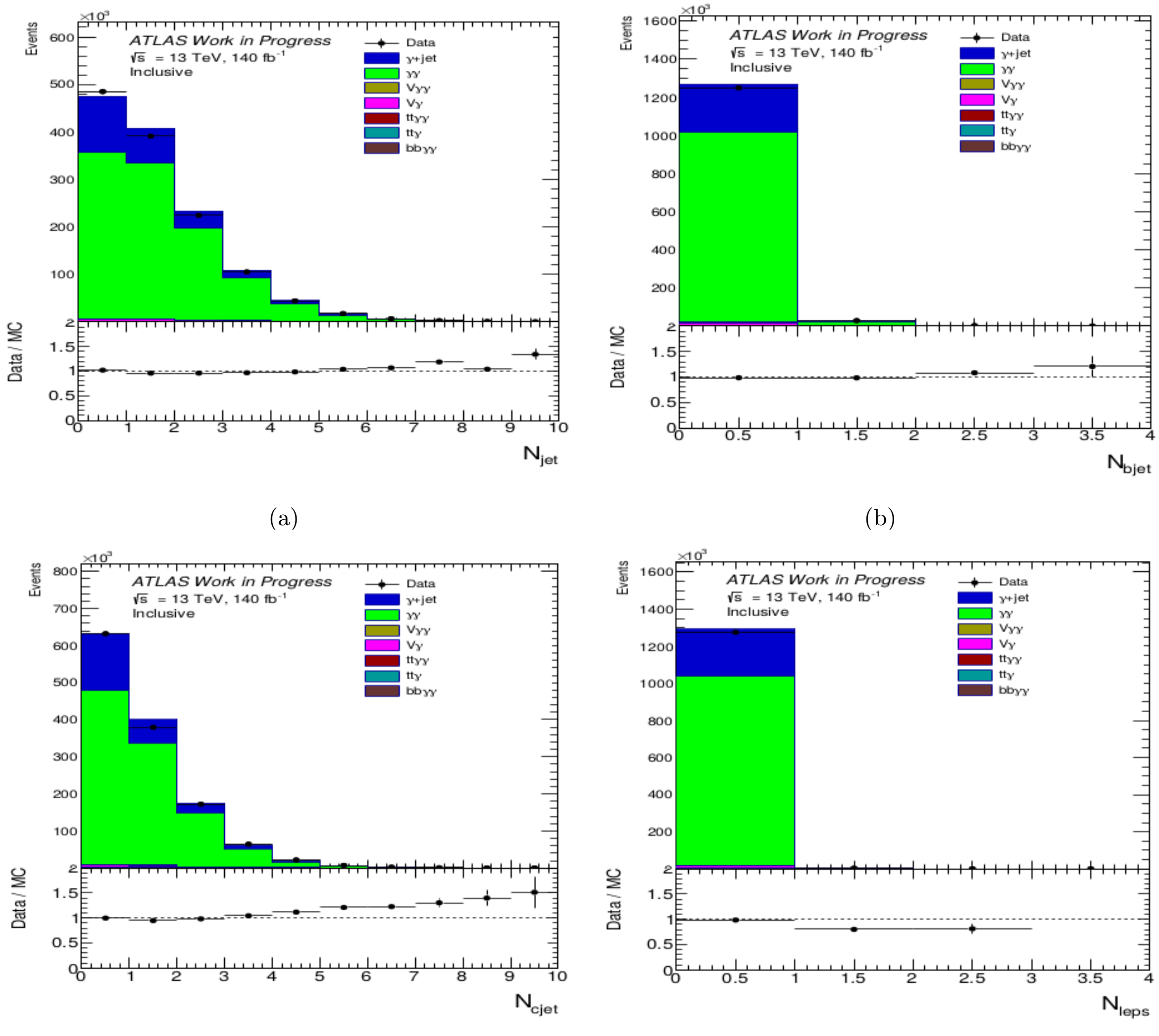
- At least one lepton and at least one b-tagged jet.
- Zero lepton and exactly one b-tagged jet.
- Zero lepton and at least two b-tagged jet.

3. Analysis

After the application of the requirements described above over the dataset and the MC (background and signal) samples, we tabulated the number of the events that passed the cuts in Table 1. It's important to note that the signal events are not normalised to anything; they are just entries. The background is normalised with its corresponding cross-section, total weight, efficiency and the data luminosity. The normalizations of the $\gamma\gamma$ and γ+jet contributions are respectively fixed to 80% and 20% of the data yield in the inclusive case. Figures 2 and 3 display the kinematic distributions corresponding to different variables for the inclusive case (just after preselections). The agreement between data and total background is reasonable in both the cores and tails.

Cuts	Data	Total Bkg	ttH180	ttH250	ttH500	ttH750	bbH180	bbH250	bbH500	bbH750
Inclusive	1278499	1301087	68150	78406	91070	98070	20091	23860	28326	29894
$N_{bjet} \geq 1$	30119	30393	52563	59397	67764	69720	7875	10772	16332	18961
$N_{bjet} = 1$ & $N_{leps} = 0$	27658	28045	22437	24468	30630	31478	7246	9954	14892	17034
$N_{bjet} \geq 2$ & $N_{leps} = 0$	2100	1912	14731	16805	17180	18316	623	795	1423	1891
$N_{bjet} \geq 1$ & $N_{leps} \geq 1$	361	436	15395	18122	19953	19924	6	23	16	35

Table 1. The number of events that are survived after the application of the event selections described in section 2 for data and MC corresponding to various regions.



(a) (b) (c) (d)

Figure 2. Kinematic distributions comparing the data to background at the inclusive case (after selecting two photons with $p_T^{1\gamma} > 40$ GeV and $p_T^{2\gamma} > 30$ GeV and $m_{\gamma\gamma} > 130$ GeV). The multiplicity of (a) jets and (b) b-tagged jets, (c) Number of central jets and (d) Number of leptons.

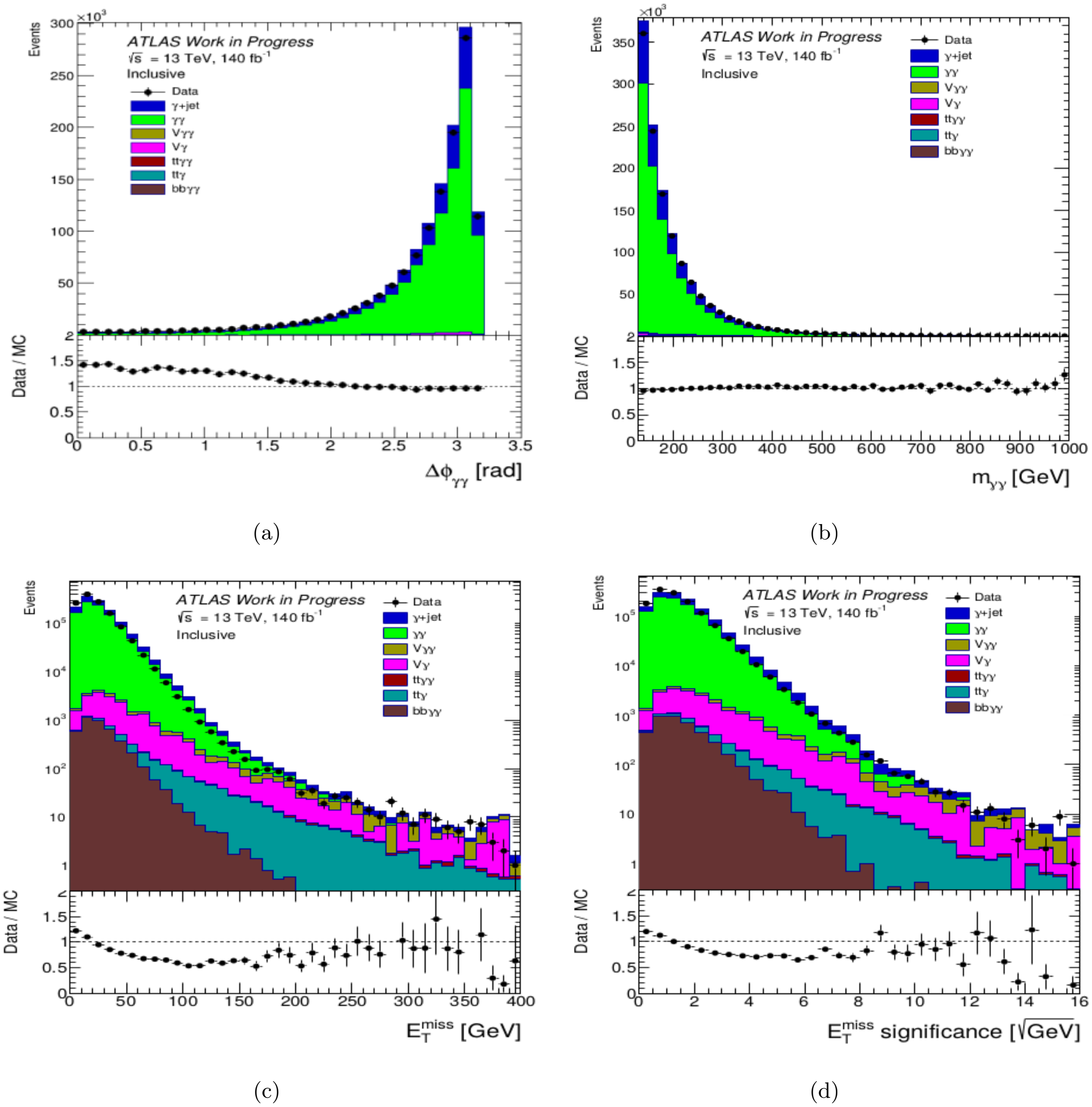


Figure 3. Kinematic distributions comparing the data to background at the inclusive case (after selecting two photons with $p_T^{1\gamma} > 40$ GeV and $p_T^{2\gamma} > 30$ GeV and $m_{\gamma\gamma} > 130$ GeV). (a) the azimuthal angle between the two photons $\Delta\phi_{\gamma\gamma}$, (b) the invariant mass of the diphoton system, (c) the missing transverse energy $E_T^{\rm miss}$ and (d) $E_T^{\rm miss}$ significance ($E_T^{\rm miss}/\sqrt{\sum E_T^{\rm miss}}$).

4. Optimization

We use the ttH sample with $m_H = 180\,\text{GeV}$ as the signal sample for optimization. The reason for choosing the ttH sample is because it has events with zero lepton and one or two b-tagged jet, and events with one lepton and one b-tagged jet. It also has more signatures that are different from the background samples. The bbH is almost the same as the background (nothing to optimize). Therefore, we optimize the ttH, and at the same time, we examine the bbH and $Z(bb)H$. Since the ttH with m_H in the range 105 to 160 GeV is examined in the nominal

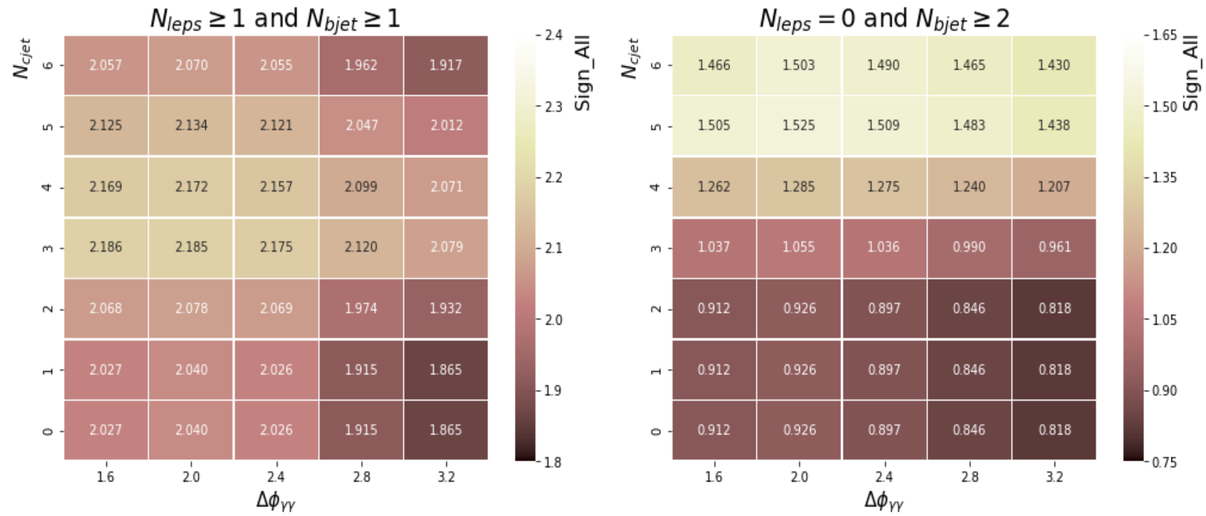


Figure 4. Optimization results for region ($N_{leps} \geq 1$ and $\mathrm{N_{bjet}} \geq 1$) and ($N_{leps} = 0$ and $\mathrm{N_{bjet}} \geq 2$). Sign_All indicate the quadrature sum of significance for all four bins in each region.

Region	$N_{leps} \geq 1$ and $N_{bjet} \geq 1$	$N_{leps} = 0$ and $N_{bjet} \geq 2$	$N_{leps} = 0$ and $N_{bjet} = 1$
NC-Significance	0.87	0.37	0.15
C-Significance	2.19	1.53	1.01
N_{cjet}	3	5	5
$\Delta\phi_{\gamma\gamma}$	1.6	2	2

Table 2. NC and C-significance with the corresponding N_{cjet} and $\Delta\phi_{\gamma\gamma}$ cuts for each region.

Higgs analysis, so that in this optimization, we use the lowest available ttH signal sample which corresponds to the $m_H = 180$ GeV. All background samples are mixed and normalised to 140 fb^{-1} for the optimization.

We optimize the variables N_{cjet} and $\Delta\phi_{\gamma\gamma}$ to select the cut that gives the maximum significance. These variables are chosen because they give better significance compared to the other variables that we tested. In each region (Section 2), we further split events into four bins by making a 2D scan on $\Delta\phi_{\gamma\gamma}$ and N_{cjet}. The significance is calculated using the formula $S/\sqrt{S+B}$, and this is to avoid having infinite significance in the case where B is zero. Where S and B are the signal and background events, respectively. Here, the signal is normalised to 50 events in the inclusive case, while the background is calculated in a small range of the diphoton invariant mass spectrum ($180 \pm 10\,$GeV). Then we use the quadrature sum of significance for all four bins in each region. Figure 4, shows the result of the optimization for two of our regions. While in Table 2, we show the Non-Categorized "NC" and the maximum categorized "C" significance that we got for all regions together with the corresponding optimal $\Delta\phi_{\gamma\gamma}$ and N_{cjet} cuts for each region. There is a big improvement in the significance after the selection cuts.

5. Summary and Future Work

A simple analysis is proposed to search for a new heavy resonance decaying into two photons in association with at least one b-tagged jet. Kinematic distributions with data to background comparisons have been shown. Optimization on truth sample (preselection) has also been presented, to mimic the analysis procedure and choose the maximum significance cut as the final event selection.

In the future, and each category, a signal plus background fit will be performed to extract the signal from the data. The double-sided crystal ball method will be used for the signal parameterization according to the diphoton mass. In contrast, an analytical function will be used to describe the background of the whole spectrum. We will test the ttH and bbH model and give the limits on the production cross-section times the branching ratio.

Acknowledgment

E. Shrif would like to acknowledge the financial support of the Schlumberger Foundation, Faculty for the Future.

References

[1] Aaboud M *et al.* [ATLAS Collaboration], Phys. Lett. B 775, 105 (2017) [arXiv:1707.04147].
[2] Molinaro E and Vignaroli N, Mod. Phys. Lett. A 32, no. 27, 1730024 (2017) [arXiv:1707.00926].
[3] Aad G *et al.* [ATLAS Collaboration], Phys. Lett. B 716, 1 (2012) [arXiv:1207.7214].
[4] Chatrchyan S *et al.* [CMS Collaboration], Phys. Lett. B 716, 30 (2012) [arXiv:1207.7235].
[5] The ATLAS collaboration, ATLAS-CONF-2015-081.
[6] CMS Collaboration,CMS-PAS-EXO-15-004.
[7] The ATLAS collaboration, ATLAS-CONF-2016-059.
[8] Lenzi B *et al.* [ATLAS Collaboration] PoS ICHEP 2016, 436 (2016) [arXiv:1611.03308].
[9] Khachatryan V *et al.* [CMS Collaboration], Phys. Lett. B 767, 147 (2017) [arXiv:1609.02507].
[10] Alwall J *et al.*, JHEP 1407, 079 (2014) [arXiv:1405.0301].
[11] Sjostrand T, Mrenna S and Skands P Z, Comput. Phys. Commun. 178, 852 (2008) [arXiv:0710.3820].