CERN Non-Member States Summer Student Program 2017

Optimization of event selections for tH-production

project report by:

Zhenisbek Tagay

Supervisor: Richard Polifka (ATLAS)

11-Aug-2017

Abstract

This summer student project studies the perspectives for the search of tH-production in di-photonic decay channel of Higgs-boson. Monte-Carlo simulation data of 3000 fb^{-1} integrated luminosity corresponding to proton-proton collisions of center-of-mass energy 13 TeV at ATLAS detector are used in this analysis.

The process being searched for, unlike other Higgs processes (e.g. ttH->γγ), is sensitive to the sign of Yukawa coupling between top-quark and Higgs-boson. However, tH-production has a very small cross section in SM, which complicates the searches. This works aims to find suitable selection criteria, which in its turn will suppress the background and provide enough statistics of tH signal for further searches.

Introduction

One of the properties of Higgs boson which hasn't been measured yet is the Yukawa coupling to top-quark. Since the coupling of Higgs-boson is proportional to the mass of fermions, coupling to the top-quark is very large. There are two possible processes in which Yukawa coupling can be measured directly: top-antitop-quark production with emitted Higgs-boson ($t\bar{t}H$), and single-top quark production with Higgs-boson (tH). The former one is not sensitive to the sign of Yukawa coupling, hence is not considered as a signal in this work.

There are 3 possible tH-production models. In the first model, b-quark and W-boson produce single-top quark and Higgs-boson is emitted (tHqb). In the second model, gluon and b-quark produce W-boson, single top-quark, and Higgs-boson is emitted (WtH). In the third model, which has significantly less cross section than first two models and, hence, can be neglected, quark and anti-quark interact to produce single top-quark, b-quark and Higgs-boson is emitted.

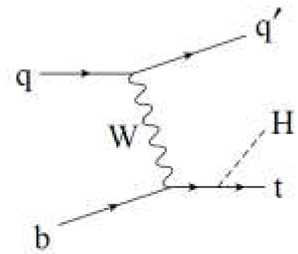

Figure 1. tHqb

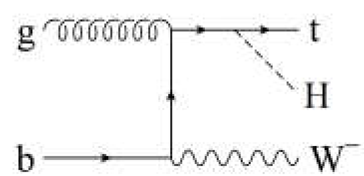

Figure 2. WtH

The deviation of Yukawa coupling from its Standard Model value will lead to dramatic increase in cross-section of tH-production, which the main motivation of this study.

Both leptonic and hadronic decay-channels of coming out W-boson are analyzed in this work.

The LHC

Most fundamental particles have very small lifetime and, therefore, can't be observed in nature. The alternative way to study this kind of particles is to produce them by accelerating and colliding other stable particles. The Large Hadron Collider (hereinafter the LHC) [1] is the world's largest and most powerful particle accelerator. The LHC is the circular accelerator with 27 km circumference. It has two beam lines, in which particles travel in opposite directions, and collide in 4 places, where beam lines cross. The rotation of protons is held by a strong magnetic field produced by superconducting magnets. These magnets operate at a temperature of 1.9 K, therefore liquid helium is used in the cooling system of an accelerator.

Since June 2015 (Run 2 period) [2] the LHC conducts proton-proton collisions with a centre-of-mass energy of 13 TeV. The maximum instantaneous luminosity is $1.7\times10^{34}\ cm^{-2}\ s^{-1}$. The Run 2 aims ti collect 100 fb^{-1} of data in three years. The Run 3 (2020-2022) will be operating at centre-of-mass energy of 14 TeV with maximum instantaneous luminosity of $2\times10^{34}\ cm^{-2}\ s^{-1}$ in order to collect 300 fb^{-1} of data by the end of 2022. Finally, LHC will be upgrarded to High Luminosity LHC with peak luminonsity of $5\times10^{34}\ cm^{-2}\ s^{-1}$ to collect 3000 fb^{-1} of data by 2035.

The ATLAS detector.

The largest detector at LHC is ATLAS (**A** **T**oroidal **L**HC **A**pparatu**s**)[3]. It is 44 m long, 25 m in diameter and located at one of the 4 interaction points. Its main purpose to identify and track particles coming from proton-proton collisions. The detector itself is made up of many layers of cylindrically arranged instruments, aimed to detect and reconstruct different types of particles.The schematic of the detector is shown in Figure 1.

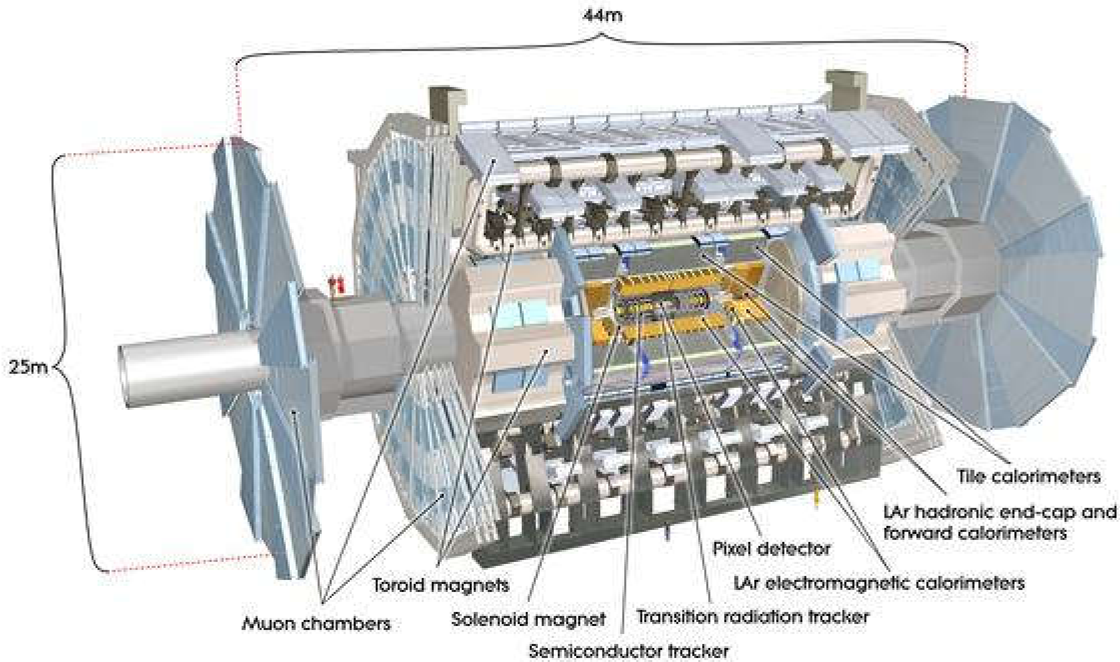

Figure 3. The ATLAS detector

The four major components of ATLAS detector are the Inner Detector, the Calorimeter, the Muon Spectrometer and Magnet System.

The Inner Detector is the first part of ATLAS detector to see particles [4]. It measures the direction, momentum, and charge of electrically-charged particles. The inner detector is consists of 3 parts: Pixel Detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT). Pixel detector, the closest part of the interaction point, is made up of 80 million pixels (channels) with a total area of 1.7 m^2. SCT, the next layer, consists of 6 million readout strips (one per 80μm). TRT is made up of drift tubes filled with xenon and carbon dioxide. The gas is ionized when charged particles go through it. The magnetic field inside the Inner Detector is 2 T.

The Calorimeter is used to measure the energy particles lose in it [5]. Almost all particles, except muons and neutrinos, interact with a calorimeter and deposit whole energy in it by developing a particle shower. Calorimeters usually consist of “passive” high-density materials (e.g. lead) and “active” media, such as liquid argon.

Electromagnetic calorimeters measure the energy of photons and electrons, whereas hadronic calorimeters are used to sample energies of hadrons and reconstruct jets.

The muon spectrometer

Muons are particles which pass through the Inner detector and Calorimeter undetected [6]. The muon spectrometer, made up of 4,000 muon chambers detects and measures the momenta of muons. Large toroidal superconducting magnets are used to deflect incoming muons, and thereby the momenta of particles can be measured.

Analysis Strategy

Monte-Carlo simulation data of 3000 fb^{-1} corresponding to a proton-proton collision at center-of-mass energy 13 TeV are used in this analysis. The analysis strategy and object definitions are based on a master thesis on searches for tH-production [7].

The parameter which is sensitive to Higgs-boson production is the diphoton invariant mass: it should match the well-known mass of Higgs-boson. Diphoton invariant mass should range between 105 and 160 GeV. The leading photon is required to have $E_T > 0.35\, m_{\gamma\gamma}$, and the subleading one $E_T > 0.25\, m_{\gamma\gamma}$. Otherwise, the event is rejected.

There are 2 types of background processes for tH-production: Higgs background and continuum background. Higgs background includes processes where Higgs-boson decaying into 2 photons is involved. The Higgs background with the highest cross section is $t\bar{t}H$. The Higgs background must be suppresses by applying relevant selection criteria. The main goal of this work is to separate signal from the Higgs background with much more cross section.

The continuum background includes processing, involving photons which meet the selection criteria but don't originate from Higgs-boson decay. This type of background is extrapolated separately from Higgs processes. Unfortunately, continuum background was not included in this work, since it is still on

Before event selection, so called so-called Upgrade Performance Functions are applied to particles. These functions simulate the behavior of particles in Upgrade detector (starting from Run 4).

Two event selections based on decay channels for W-boson are developed: leptonic and hadronic. In the leptonic channel, only events with 1 electron or muon are considered for analysis. Tau-leptons are not considered due to difficulties with reconstruction.

Signal has a relatively small expected number of events with a comparison to the background. In order to quantify the performance of our analysis selection, expected significance is calculated. It is defined as $Z_0 = {}^{S}\!/_{\sqrt{B}}$, where S stands for number of expected tH signal events and B for background. It is expected that events from the continuous background will play a dominant role in the background estimation, however due to lack of simulation, they are absent in this analysis. Therefore, only the separation of Higgs signal and background is studied.

In order to exploit different shapes of discriminating variables of signal and background, categorisation is often implemented. In this analysis, it is based on pseudorapidity (η) of the jets present in the event. In this way, the combined significance of individual categories yields larger sensitivity than if the whole analysis would be considered integrated over η. The combined significance is then calculated as square root of quadratic sum of individual significances

Object definitions

Photons are reconstructed as energy deposits in the electromagnetic calorimeter in pseudorapidity $|\eta|<2.37$. The space between barrel and endcap calorimeters (i.e. $1.3<|\eta|<1.56$) is excluded. We require having at least 2 photons with invariant mass in the range between 105 GeV and 160 GeV since they must originate from Higgs-decay. Finally, the leading photon needs to have $E_T>0.35m_{\gamma\gamma}$, and subleading photon $E_T>0.25m_{\gamma\gamma}$.

Electrons are reconstructed from energy deposits on electromagnetic calorimeter in the region $|\eta|<2.47$. Also, electrons need to have one track in Inner Detector associated with them. Transverse energy threshold for selection is $E_T>15$ GeV. Finally, electrons must be isolated. If there is a photon in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}=0.4$, the electron is rejected to avoid double-counting.

Muons are reconstructed by tracks both in the Muon Spectrometer and the Inner Detector in the region $|\eta|<2.7$. They are required to have transverse momentum greater than 10 GeV. Last, muons should be isolated. If there is a photon in a cone with $\Delta R=0.4$, it is rejected, for the same reason as electrons.

Jets are reconstructed from energy deposits in the calorimeter. Jets are formed from clusters of cells in the calorimeter. In this work jets are required to have at least 30 GeV transverse momentum and be in the region $|\eta|<4.5$. Similar to other particles, electrons should be isolated. If there is a photon in a cone with $\Delta R=0.4$ or an electron in cone with $\Delta R=0.2$, the jet is rejected.

Jets containing b-jets differ from other jets by their long lifetime. Therefore, a secondary vertex can be associated with them. In this study, 70% working point is used in the b-tagging algorithm. The requirements for b-jets are the same as for usual jets, except for isolation criteria: if there is a photon in a cone with $\Delta R=0.4$, the photon is rejected, b-jet is accepted. The reason is because b-jet can produce a photon.

All aforementioned particles are required to pass corresponding efficiency tests, given by Upgrade Performance Functions.

Analysis

Diphoton invariant mass

As it was mentioned before, in tH-production 2 leading photons must originate from Higgs-decay. To test this, the invariant mass of leading and subleading photons was calculated for each selected event in the signal Monte Carlo containing Higgs boson of mass $m_H = 125\ GeV$ decaying to a pair of two photons. The histogram below illustrates the distribution of invariant mass.

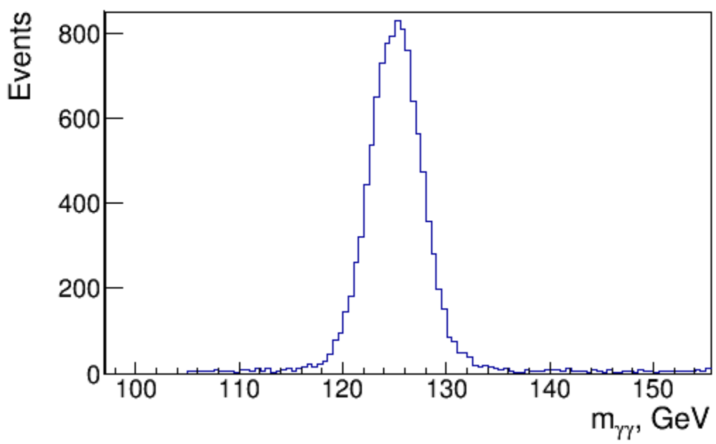

Figure 4. Diphoton Invariant Mass

Leptonic selection

In this decay channel, W-boson decays into a lepton and its neutrino. Before calculating significance, one needs to see which selection have more statistics for signal than one for the background. The distributions for the number of jets are presented in the figure below.

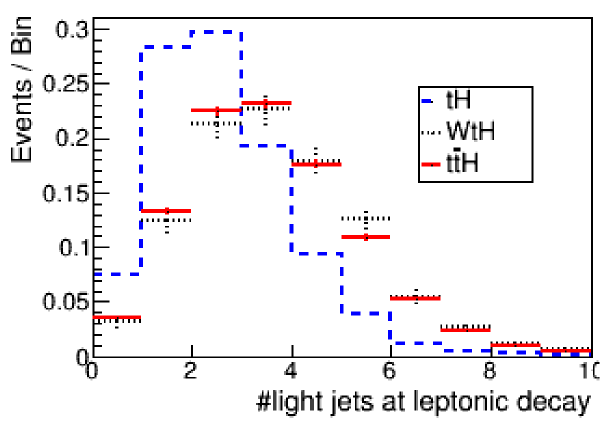

Figure 6. Distribution for the number of light jets

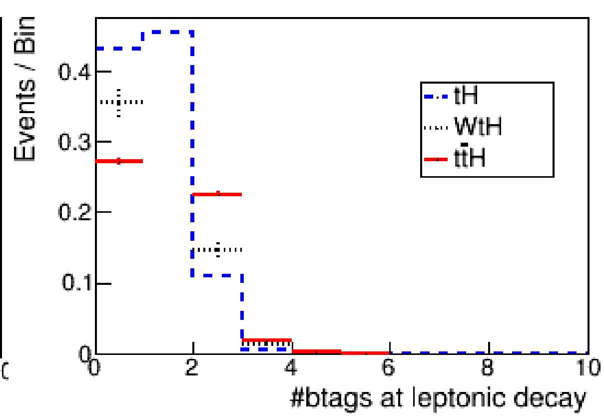

Figure 5. Distribution for the number of b-tags

From the distributions, it is clear that one should consider events with at most 2 light jets since in that case signal has much better statistics than the background. Regarding b-tags, the background has better statistics in any case. However, we can set our selection criteria to exactly 1 b-tag, where the difference between the background and signal statistics is minimal. Significances for mentioned categories are presented below.

Leptonic selection						
	1b[1];<=2 lj[2]		1b; <=2 lj ; 0 frw. j[3]		>=1b ;<=2 lj; >=1 frw.j	
tH	28.09	±0.97	16.55	±0.75	11.55	±0.62
WtH	12.76	±0.63	10.88	±0.58	1.88	±0.24
ttH	228.03	±3.18	191.04	±2.91	36.99	±1.28
S	40.85	±1.16	27.43	±0.95	13.43	±0.66
B	228.03	±3.18	191.04	±2.91	36.99	±1.28
Significance	2.71	±0.08	1.98	±0.07	2.21	±0.12
			Combined significance: **2.96**			

Table 1. Leptonic selection

The leptonic category is split into 2 sub-categories: with at least forward jet, and without it. From the table, we see that one with forwarding jet has a bit more contribution to the combined significance.

Hadronic selection

Now we look at jet distributions for hadronic decay channel, presented below.

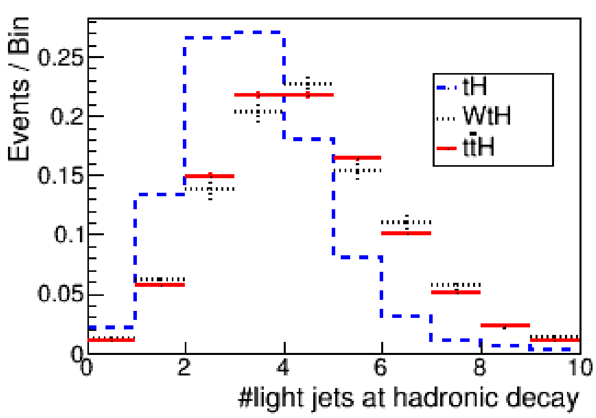

Figure 8. Distribution for the number of light jets

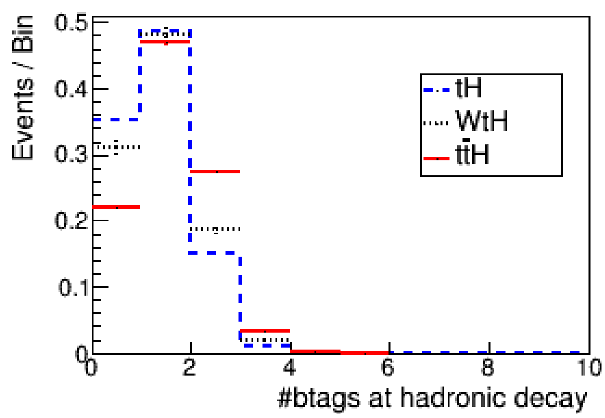

Figure 9. Distribution for the number of b-tags

Based on distributions above, one would select events with at most 3 light jets, because in that selection we have more signal than the background. A number of b-tags will remain 1, as it was in the leptonic channel. The significance for these selections is given in the next page.

---

[1] b-tags
[2] light jets
[3]forward jets (b-tags + light jets)

Hadronic						
	1b, <=3 lj		1b, <=3 lj, >=1 frw lj		1b, <=3 lj, 0 frw. lj	
tH	115.80	±1.97	45.58	±1.23	70.22	±1.53
WtH	26.84	±0.91	4.71	±0.37	22.14	±0.00
ttH	516.60	±4.78	90.53	±2.00	426.06	±0.82
S	142.64	±2.17	50.29	±1.28	92.36	±4.34
B	516.60	±4.78	90.53	±2.00	426.06	±0.82
Significance	6.28	±0.10	5.29	±0.15	4.47	±0.21
			Combined significance: **6.93**			

Table 2. Hadronic decay, light jet splitting

Here we see that this selection has the biggest significance among all categories. Splitting according to the number of forward light jets shows that the events with at least one light jet in the forward region have bigger significance, and, therefore, contribute more to the combined significance than the ones without the forward light jet.

Hadronic						
	1b, <=3 lj		1b cent., <=3 lj		1b frw.,<=3 lj	
tH	115.8	±1.97	106.94	±1.89	8.86	±0.55
WtH	26.84	±0.91	26.13	±0.89	0.71	±0.15
ttH	516.6	±4.78	500.03	±4.71	16.57	±0.86
S	142.64	±2.17	133.07	±2.09	9.57	±0.57
B	516.6	±4.78	500.03	±4.71	16.57	±0.86
Significance	6.28	±0.10	5.95	±0.10	2.35	±0.15
			Combined significance: **6.40**			

Table 3. Hadronic selection, b-tag splitting

Splitting the only b-tag in this category into central and forward implies that the hadronic selections with one central jet will contribute more to the combined significance.

Since the combined significance, in this case, is less than in the previous splitting of hadronic decay, the previous one will be used further.

The combination of leptonic and hadronic categories yields:

	S	B	significance
Leptonic	40.85	228.03	2.96
Hadronic	142.64	516.60	6.93
Combined			7.53

Table 4. Combination of categories

Summary and Outlook

This work was performed as a CERN summer student project. Its main goal was to develop event selection criteria in order to increase the significance of very rare signal over its background. This work was motivated by existing master thesis on the related topic [7]. In comparison to that work, we already have applied Upgrade Smearing Functions to particles, added pileup jets and used new tracking and b-tagging efficiencies. Moreover, b-tags are considered both in central and forward regions.

Unfortunately the continuum background was not taken into account due to technical difficulties, therefore the absolute value of significance should not be considered final for this analysis. However due to the fact that the missing background is expected to contribute similarly to all categories, the relative importance (and the combination over inclusive analysis) of individual categories can be inferred from this study

This work was performed in context of the ATLAS High Granularity Timing Detector phase 2 upgrade which focuses on improving the pile up rejection in the forward region of the upgraded ATLAS detector at the HL-LHC. From results above we can already conclude, that due to the relative sensitivity of this analysis to the forward categories, an improved reconstruction of the events in the forward region can improve the expected sensitivity of the tH production at the HL-LHC.

The work is in progress and the final result is expected to be part of the internal ATLAS High Granularity Timing Detector Initial Design Report.

References

[1] *The Large Hadron Collider*. https://home.cern/topics/large-hadron-collider

[2] *LHC ten year schedule*. http://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-schedule-update.pdf

[3] *ATLAS Experiment. Detector & Technology.* https://atlas.cern/discover/detector

[4] *ATLAS Experiment. The Inner Detector.* https://atlas.cern/discover/detector/inner-detector

[5] *ATLAS Experiment. The Calorimeter.* https://atlas.cern/discover/detector/calorimeter

[6] *ATLAS Experiment. The Muon Spectrometer.* https://atlas.cern/discover/detector/muon-spectrometer

[7] *Studies on the prospects for the search for tH-production with* $H \rightarrow \gamma\gamma$ *at* $\sqrt{s}$ *= 13 TeV with the ATLAS detector.* Isabel Nitsche, master thesis