Draft version 0.3

ATLAS NOTE

January 9, 2017

Search for Higgs pair production with decays to WW(jjjj) and $\gamma\gamma$ in 36.5 fb^{-1} proton-proton collision data at 13 TeV in the ATLAS detector

ATLAS Collaboration[1], Yu Zhang[1], Jason Veatch[3.], Kai Henßen[3.], Stan Lai[3.]

[1.] *Institute of High Energy Physics, Chinses Academy of Siences, Beijing, China*
[2.] *University of Texas at Dallas*
[3.] *II. Physikalisches Institut, Georg-August-Universität, Göttingen, Germany*

Abstract

A search is performed for resonant and non-resonant Higgs pair production with one Higgs boson decaying to full hadronic WW^* and the other to $\gamma\gamma$ using proton-proton collision data corresponding to an integrated luminosity of 36.5 fb^{-1} at a 13 TeV centre-of-mass energy recorded with the ATLAS detector. No deviation from the Standard Model prediction is observed. The observed (expected) upper limit at 95% confidence level on the cross section for $gg \rightarrow hh$ is XXX pb (XXX pb) for the non-resonant Higgs pair production. For resonant Higgs pair production, the observed (expected) upper limits at 95% confidence level on cross section times the branching ratio of $X \rightarrow hh$ range from XXX pb (XXX pb) to XXX pb (XXX pb) as a function of the resonant mass from 260 GeV to 3 TeV assuming that the narrow-width approximation holds.

Contents

1 Statements

1.1 version 0.0

22.1 fb^{-1} of data and 260 GeV signal sample are used for first selection optimization.
To do list:

- optimize the jet combination. Selecting leading 4 or 3 jets is not the optimal strategy.
- check signal of other mass point and investigate more kinematic variables for further selection optimization.
- signal and background modeling.
- systematics study.

2 Introduction

2.1 Theoretical motivation

2.1.1 Non-resonant hh production

2.1.2 Resonant hh production

3 Data and Monte Carlo samples

3.1 Data samples

The data samples used in this analysis correspond to the data recorded by ATLAS in the whole 2015 (3.2 fb^{-1}) and 33.3 fb^{-1} of 2016, which sums up to an inegrated luminosity of 36.5 fb^{-1}. The whole dataset is recorded with all subsystems of ATLAS operational [1].

3.2 Monte Carlo samples

SM single Higgs bacgkrounds and signals are estimated with MC samples that are documented in this section, while the continuum photon background of the SM processes with multiphotons and multijets is estimated in sideband [2] with the data-driven method as described in Section 7.3.

The simulation under MC15c configuration is used in the analysis. The samples are generated with the consideration of multiple interactions per bunch crossing by introducing pileup noise at the stage of digitization. MC15c configuration incorporates the pileup condition that is an average of the actual pileup condition in 2015 data and an estimation for 2016 data.

3.2.1 MC samples for signals

Signal samples are generated with MadGraph5_aMC@NLO [1]. For both non-resonant and resonant productions, the event generation is performed using a next-to-leading-order SM Higgs pair model developped by the Cosmology, Particle Physics and Phenomenology (CP3) theory group [2]. Events are generated with a Higgs Effective Field Theory (HEFT) using aMC@NLO method [3] and are reweighted to take into account top quark mass dependence. The top mass can become an important effect [4], particularly for the non-resonant case. The shower is implemented by Herwig++ [5] with UEEE5 underlying-event tune [6], and the PDF set CTEQ6L1 [7] is used. The heavy scalar, H, is assumed to have a narrow width. Technically its decay width is set to 10 MeV in the event generation for the following masses: 260 GeV, 300 GeV, 400 GeV, 500 GeV, 750 GeV, 1 TeV, 1.5 TeV, 2 TeV, 2.5 TeV, and 3 TeV. The card used in MadGraph5 for signal event generations is attached The generator level filter *ParentChildFilter* implements the selection of these decay products. Details on the signal samples are listed in Table 1. All signal samples are produced with the ATLAS fast simulation framework (AF2).

3.2.2 MC samples for SM single Higgs backgrounds

Simulated samples for SM single Higgs background are produced to investigate the components of this background in $m_{\gamma\gamma}$ and to estimate their contributions. The SM single Higgs background considered here is assumed to be produced via five production modes: gluon-gluon fusion (ggh), vector boson fusion, (VBF), Higgsstrahlung (Wh and Zh) and Higgs associated production with a pair of top quarks ($t\bar{t}h$), where h is the light (SM-like) 125 GeV Higgs boson. These samples are simulated using the full ATLAS simulation and reconstruction chain. The mass of the SM Higgs boson is set to 125 GeV. More details on generator, parton shower and simulation tags are listed in Table 2.

The cross sections at $\sqrt{s}$ = 13 TeV corresponding to each production mode are listed in Table 3. In the analysis, these cross sections will be multiplied by the $h \to \gamma\gamma$ branching ratio of 0.00228, since all simulated samples are produced with SM Higgs decaying into photon pairs.

[1] Good Run Lists are data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml for 2015 data and data16_13TeV.periodAllYear_DetStatus-v82-pro20-12_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml for 2016 data

[2] The sideband is defined as $m_{\gamma\gamma} \in [105, 160]$ GeV excluding the Higgs mass window as defined in Section 5.1.

DSID	Processes	Generators, tunes and PDFs	Tags
342621	non-resonant	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e4419\_a766\_a821\_r7676\_p2691$
343756	$X \to hh$, 260 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343758	$X \to hh$, 300 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343761	$X \to hh$, 400 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343763	$X \to hh$, 500 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343818	$X \to hh$, 750 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343819	$X \to hh$, 1000 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343820	$X \to hh$, 1500 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343821	$X \to hh$, 2000 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343822	$X \to hh$, 2500 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$
343823	$X \to hh$, 3000 GeV	MadGraph + Herwigpp UEEE5 CTEQ6L1	$e5153\_a766\_a821\_r7676\_p2691$

Table 1: Simulated signal samples

DSID	Processes	Generators, tunes and PDFs	Tags
341000	ggh	Powheg+Pythia8 AZNLO CTEQ6L1	$e3806\_s2608\_r7772\_r7676\_p2669$
341001	VBF	Powheg+Pythia8 AZNLO CTEQ6L1	$e3806\_s2608\_r7772\_r7676\_p2669$
341067	Wh	Pythia8 A14 NNPDF2.3LO	$e3796\_s2608\_s2183\_r7772\_r7676\_p2669$
341068	Zh	Pythia8 A14 NNPDF2.3LO	$e3796\_s2608\_s2183\_r7772\_r7676\_p2669$
341069	$t\bar{t}h$	Pythia8 A14 NNPDF2.3LO	$e3796\_s2608\_s2183\_r7772\_r7676\_p2669$

Table 2: Simulated SM single Higgs background samples.

production	cross sections
ggh	48.52 pb
VBF	3.779 pb
Wh	1.369 pb
Zh	0.8824 pb
$t\bar{t}h$	0.5065 pb
$gg \to hh$	33.41 fb

Table 3: Cross sections for SM single Higgs processes at $\sqrt{s} = 13$ TeV with $m_h = 125.09$ GeV and the SM Higgs pair productions, $gg \to hh$.

4 Object definition

The object definition is similar to what is used by the HGam group. The analysis framework of $hh \to \gamma\gamma WW^*$ is based on the HGamAnalysisFramework that is centrally developed by HGam group. The tag of the framework is HGamAnalysisFramework-00-02-55-11 which is used to produce official MxAOD samples of version h013a.

4.1 Photons

- The E_{T} of leading (sub-leading) photon is required to be larger than 25 GeV.
- The $|\eta|$ of photon is considered up to 2.37, vetoing the crack region $1.37 < |\eta| < 1.52$.
- Tight photons are required as is the default in HGam group. The photon identification algorithm is based on the lateral and longitudinal energy profiles of the shower measured in the electromagnetic calorimeter.
- The isolation working point FixedCutLoose is used. It is one of the recommended points from the isolation forum. Photons are required to pass both calorimeter-based and track-based isolation requirements.
- Photons are passed through the e/γ ambiguity tool, as is the default in the HGam group. The ambiguity tool is developed to discriminate photons and electrons that can otherwise have overlapping selections. In particular, converted photons from electrons in the silicon can lead to large e/γ fake rate. The ambiguity tool makes requirements on the number of silicon hits and the conversion rates to keep this rate under control without significant loss of signal efficiency.

4.2 Jets

- The anti-k_t algorithm [8] with the size parameter of $R = 0.4$ is used to reconstruct jets from topological clusters in the calorimeters that are calibrated to the EM scale.
- Jets undergo an energy calibration
- Jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$.
- Jets from pileup are rejected by applying a JVT (Jet Vertex Tagger) cut. The jet is rejected if JVT< 0.59 for $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$.
- Events with a jet passing the LooseBad cut are rejected. The LooseBad jet quality requirement is designed to reject fake jets caused by detector readout problems and non-collision backgrounds.

4.3 Electrons

Electrons are reconstructed from energy clusters in the EM calorimeter matched with tracks reconstructed in the inner detector.

- E_{T} is required to be larger than 10 GeV.
- $|\eta|$ is required to be less than 2.47 vetoing the transition region with $1.37 < |\eta| < 1.52$.
- The $|d_0|$ significance ($d_0/\sigma(d_0)$) with respect to the primary vertex in the event is required to be less than 5.

- The $|z_0|$ with respect to the primary vertex in the event is required to be less than 0.5mm.
- Identification: Medium quality electrons are used.
- Isolation: Loose electrons are used.

4.4 Muons

Muons are reconstructed from tracks in the inner detector and the muon spectrometer.

- p_T is required to be larger than 10 GeV.
- $|\eta|$ is required to be less than 2.7.
- The $|d_0|$ significance with respect to the primary vertex in the event is required to be less than 3.
- The $|z_0|$ with respect to the primary vertex in the event is required to be less than 0.5mm.
- Identification: Medium quality muons are used.
- Isolation: GradientLoose is used.

4.5 Overlap removal

Since the collections of objects are reconstructed using different algorithms in parallel (i.e. there no check to prevent a single cluster or track from being included in the reconstruction of two different object) it is necessary to implement a set of rules to remove objects nearby each other to avoid double counting. The rules are implemented sequentially as defined below:

- The two leading photons are always kept.
- Electrons with $\Delta R(e, \gamma) < 0.4$ are removed.
- Jets with $\Delta R(jet, \gamma) < 0.4$ are removed.
- Jets with $\Delta R(jet, e) < 0.2$ are removed.
- Muons with $\Delta R(\mu, \gamma) < 0.4$ or $\Delta R(\mu, jet) < 0.4$ are removed
- Electrons with $\Delta R(e, jet) < 0.4$ are removed.

5 Event selection

The event selection procedure identifies two photons and then applies requirements on the multiplicities of jets in order to increase the signal purity and background rejection for events with multi-jets. This analysis selects events with a boosted topology as well as events with a resolved topology. The event selection for the analysis starts with the full di-photon selection from the $h \to \gamma\gamma$ analysis in RUN II to select two high p_{T} isolated photons.

5.1 Common selection

- **Trigger**: Events are required to pass at least one of the following diphoton triggers, using a logical OR: `HLT_g35_loose_g25_loose` or `HLT_g35_medium_g25_medium` or `HLT_2g50_loose` or `HLT_2g20_tight`.
- **Good Run List and Detector Quality**: Events must belong to the luminosity blocks specified in the Good Run Lists:
  - data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml for 2015 data
  - data16_13TeV.periodAllYear_DetStatus-v82-pro20-12_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml for 2016 data

  These GRLs reject events with data integrity errors in the calorimeters and incomplete events where some detector information is missing are rejected, as well as events which are corrupted due to power supply trips in the tile calorimeter.
- **Primary Vertex**: The primary vertex is selected using the neural network algorithm from HGam group. The photons' four momenta, JVT and track isolation are corrected with respect to this origin, and the mass of the diphoton system is accordingly recalculated.
- **2 loose photons**: At least two loose photons with $E_{\mathrm{T}} > 25$ GeV and within the detector acceptance are selected.
- The other cuts on photons involving **Identification (tight ID), Isolation, Rel.Pt cuts**. The relative p_{T} cut requires the p_{T} of leading (sub-leading) photon to be larger than 0.35(0.25) of diphoton invariant mass. The diphoton invariant mass is required to be within the range $m_{\gamma\gamma} \in [105, 160]$ GeV.
- **Higgs mass window**: $|m_{\gamma\gamma} - m_h| < 2\sigma_{m_{\gamma\gamma}}$ is also required where $m_h = 125.09$ GeV is the measured SM Higgs boson mass and $\sigma_{m_{\gamma\gamma}} = 1.7$ GeV is the experimental diphoton mass resolution.
- **Lepton veto**: Events are required to contain exactly zero electrons or muons.
- ***b*-veto**: In order to suppress backgrounds with top quarks and ensure orthogonality to other hh searches ($bb\gamma\gamma$, $bbbb$, $bb\tau\tau$, etc.), the event is rejected if there are any b-tagged jets. The b-tagger is MV2c10 with a b-tagging efficiency of 70%.

The efficiencies of common event selection are listed in Table 4. These efficiencies are derived for signals from simulated samples. After the selection of the two photons, the signal efficiencies range from 38.0% to 43.0%, while after the additional selection on the jets, the leptons and the tight mass window on the di-photon, the signal efficiencies range from 5.65% to 10.7%, for a resonant mass from 260 and 500 GeV.

	SM	Resonant hh			
	Higgs pair	260 GeV	300 GeV	400 GeV	500 GeV
All Events	100.0%	100.0%	100.0%	100.0%	100.0%
Duplicate	100.0%	100.0%	100.0%	100.0%	100.0%
GRL	100.0%	100.0%	100.0%	100.0%	100.0%
Pass Trigger	73.7%	68.5%	69.6%	71.9%	74.6%
Detector Quality	73.7%	68.5%	69.6%	71.9%	74.6%
has PV	73.7%	68.5%	69.6%	71.9%	74.6%
2 loose photons	59.3%	56.9%	56.5%	57.6%	59.7%
Trig Match	59.0%	56.6%	56.3%	57.3%	59.9%
Tight ID	49.8%	46.8%	46.2%	48.1%	50.8%
Isolation	45.2%	40.2%	40.2%	43.4%	46.5%
Rel.Pt cuts	41.7%	37.5%	36.4%	39.4%	43.0%
$105 < m_{\gamma\gamma} < 160$ GeV	41.6%	37.4%	36.3%	39.2%	42.8%

Table 4: Efficiencies for the common event selection criteria

5.2 Boosted selection

- **Large-R jet multiplicity**: The boosted selection is sensitive to events in which both W-bosons are sufficiently boosted that the decay products of each are fully contained in large-R jets. Events are required to contain ≥ 2 large-R jets.

- **On-shell W-boson identification**: ≥ 1 large-R jet is required to have a mass consistent with $m_W = 80.3$ GeV.

- **Identification of 2-prong decays**: Substructure variables will be used to identify the large-R jet(s) containing the decay products of one or both of the W-bosons. This will be updated as the selection becomes finalized.

5.3 Resolved selection

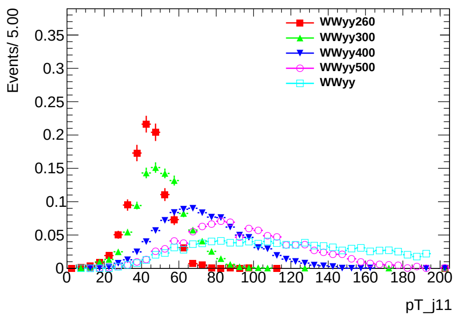
- **Orthogonality with the boosted selection**: Events that fail the large-R jet multiplicity requirement in the boosted selection are considered for the resolved selection.

- **Jet multiplicity**: Considering the jet p_{T} at truth level, the two categories are defined by exact 3 jets or at least 4 jets to enlarge signal efficiency.

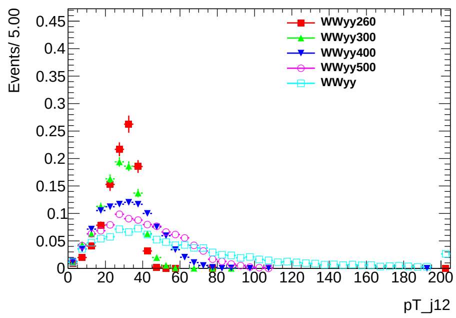
6 Selection optimization

The character of signal events is estimated by MC. The estimation of SM Higgs background is from MC. Data sideband $|m_{\gamma\gamma} - 125.09| > 3.4GeV$ is used to model the continuum background. $\sigma(pp \rightarrow hh)$ is supposed to be 1 pb. The events are split to exact 3 jet category and at least 4 jet category because the signal jets from W boson are very soft and the pT threshold of 25 GeV could kill many signal jets. Figure 1 shows the pT of signal jets at truth level. The general selection optimization is as following. First, define a mass window of diphoton + 3 or 4 jets containing 85 % signal events for resonant search. Second, a scan on $pT_{\gamma\gamma}$ is performed to improve the significance since the SM Higgs will be boosted in di-Higgs event. The significance is calculated by Eq 1. The yield of BSM signal and SM Higgs is estimated from MC. The continuum background yield is estimated from an exponential fit on data sideband.

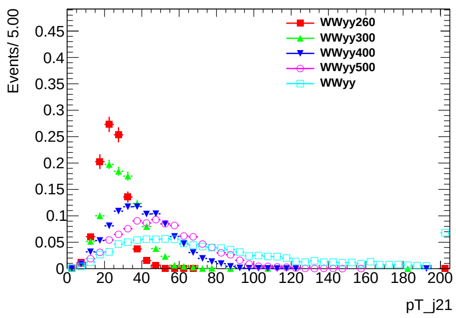
$$\sigma_{signal} = \sqrt{2 \times \left\{ (N_{signal} + N_{SM\ Higgs} + N_{continuum}) \times ln(1 + \frac{N_{signal}}{N_{SM\ Higgs} + N_{continuum}}) - N_{signal} \right\}} \quad (1)$$



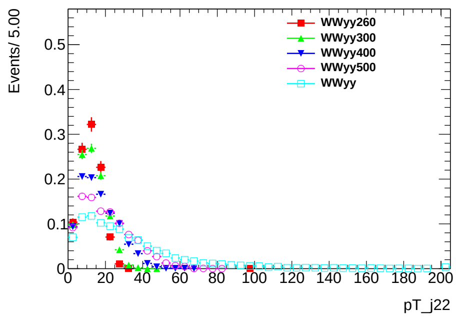
(a) leading jet of on-shell W



(b) subleading jet of on-shell W



(c) leading jet of off-shell W

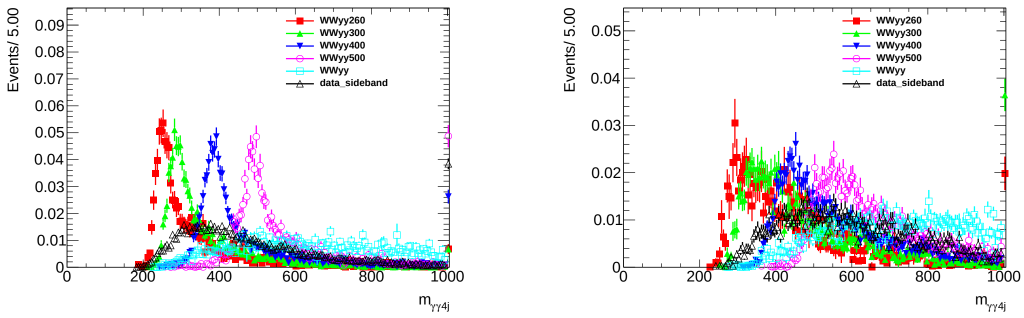


(d) leading jet of off-shell W

Figure 1: pT of four signal jets at truth level

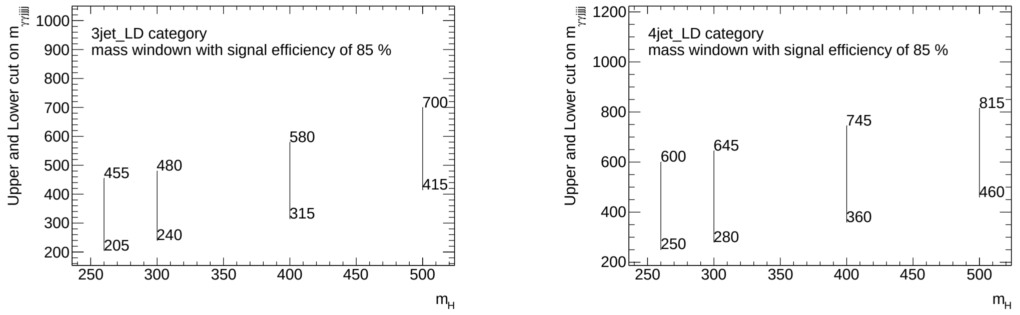
6.1 Strategy 1

The most simple selection is considered in this section. The leading 3 or 4 central jets are selected as WW^* candidate. Additional mass constraint of capital Higgs and $p_{\gamma\gamma}$ cut are introduced. Figure 2 shows the invariant mass of diphoton plus 3 or 4 jets and Figure 3 defines the mass window containing 85 % of signal yields. The distribution of $pT_{\gamma\gamma}$ is shown in Figure 4. A scan on $pT_{\gamma\gamma}$ and fit on data sideband are performed . Figure 5 6 7 8 9 show the result of optimization. Table 5 summarize the result.



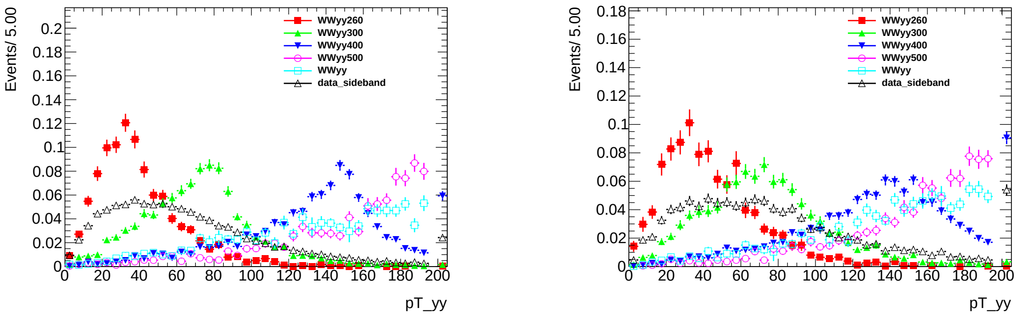
(a) 3 jet category (b) 4 jet category

Figure 2: invariant mass of $\gamma\gamma + 3\ or\ 4\, jets$



(a) 3 jet category (b) 4 jet category

Figure 3: mass window of $m_{\gamma\gamma4(3)j}$ containing 85 % of signal events

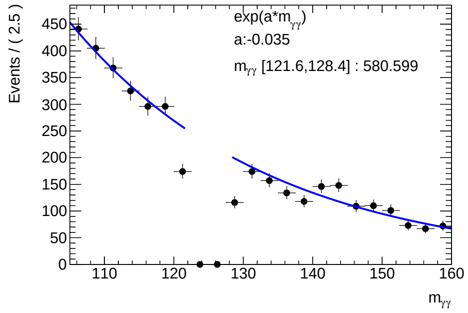


(a) 3 jet category (b) 4 jet category

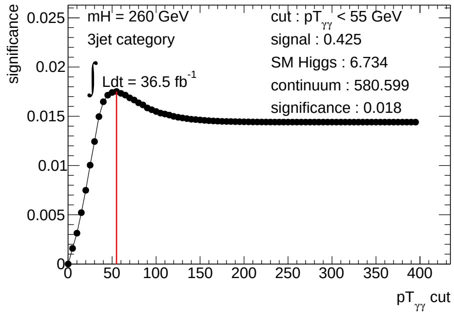
Figure 4: $pT_{\gamma\gamma}$ distribution in 3 jet and 4 jet category

	non-res		260 GeV		300 GeV		400 GeV		500 GeV	
mass window [GeV]	-	-	[205, 455]	[250, 600]	[240, 480]	[280, 645]	[315, 580]	[360, 745]	[415, 700]	[460, 815]
$pT_{\gamma\gamma}[GeV]$ cut	> 155	> 155	< 55	< 70	> 40	> 25	> 120	> 120	> 175	> 180
signal yield	0.48	0.64	0.43	0.24	0.50	0.37	0.48	0.48	0.51	0.55
SM Higgs	13.12	7.72	6.73	2.02	18.18	6.87	7.88	4.05	4.05	1.98
continuum	85.82	64.28	580	180	779	289	120	51	31	13
significance	0.048	0.076	0.018	0.018	0.018	0.021	0.042	0.064	0.086	0.142
combined significance	0.090		0.025		0.028		0.077		0.166	

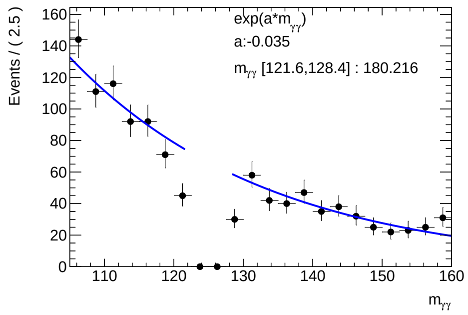
Table 5: summary of selection, yield and significance



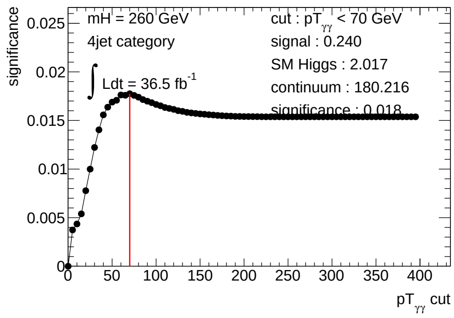
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

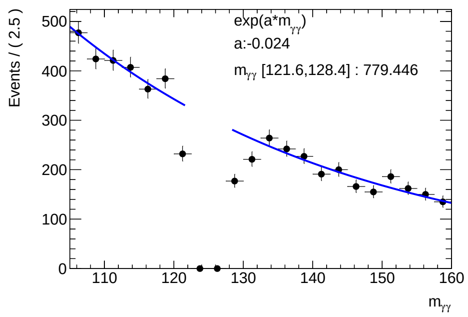


(c) cut scan of 3 jet category

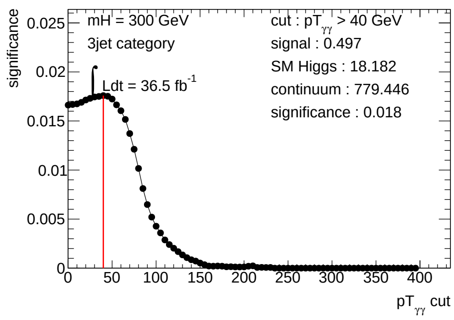


(d) cut scan of 4 jet category

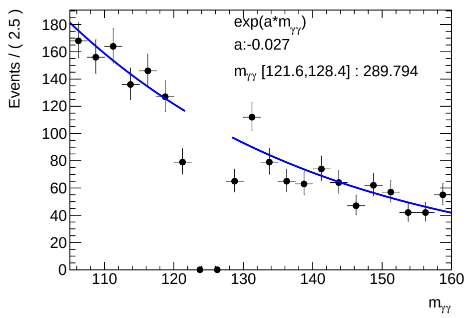
Figure 5: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 260 GeV.



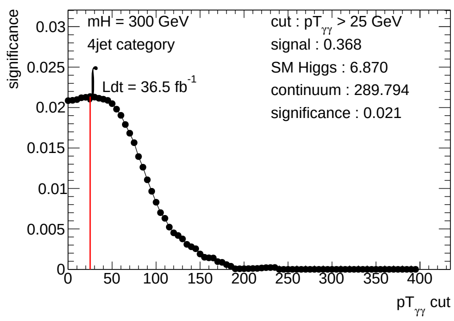
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

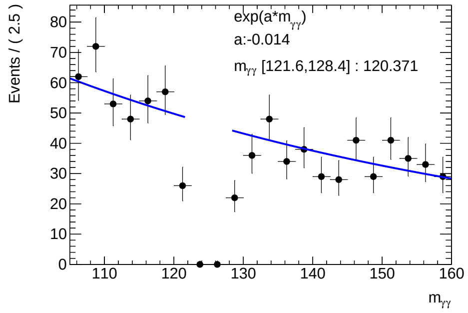


(c) cut scan of 3 jet category

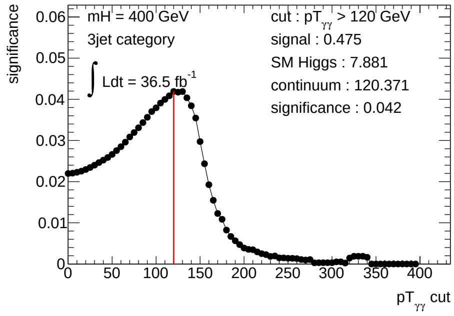


(d) cut scan of 4 jet category

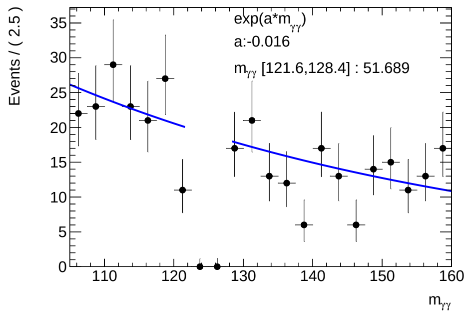
Figure 6: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 300 GeV.



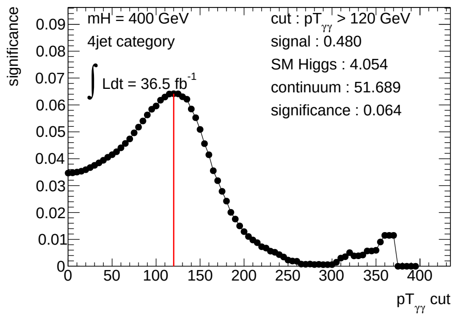
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

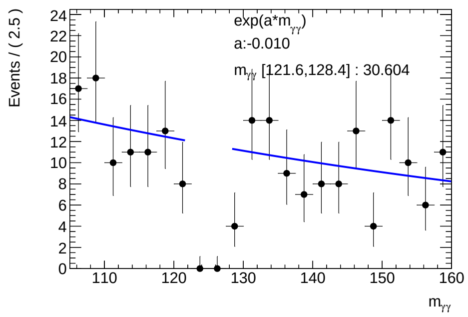


(c) cut scan of 3 jet category

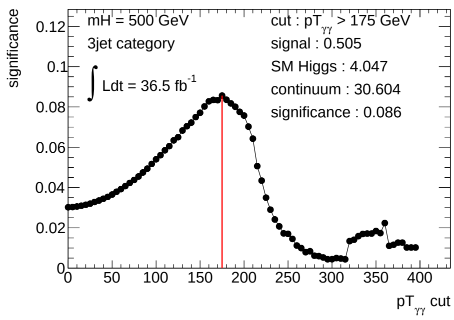


(d) cut scan of 4 jet category

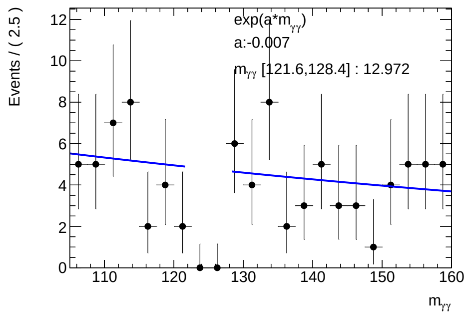
Figure 7: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 400 GeV.



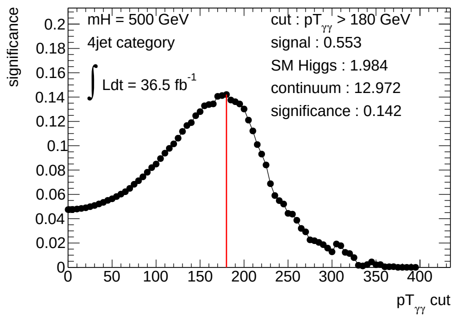
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

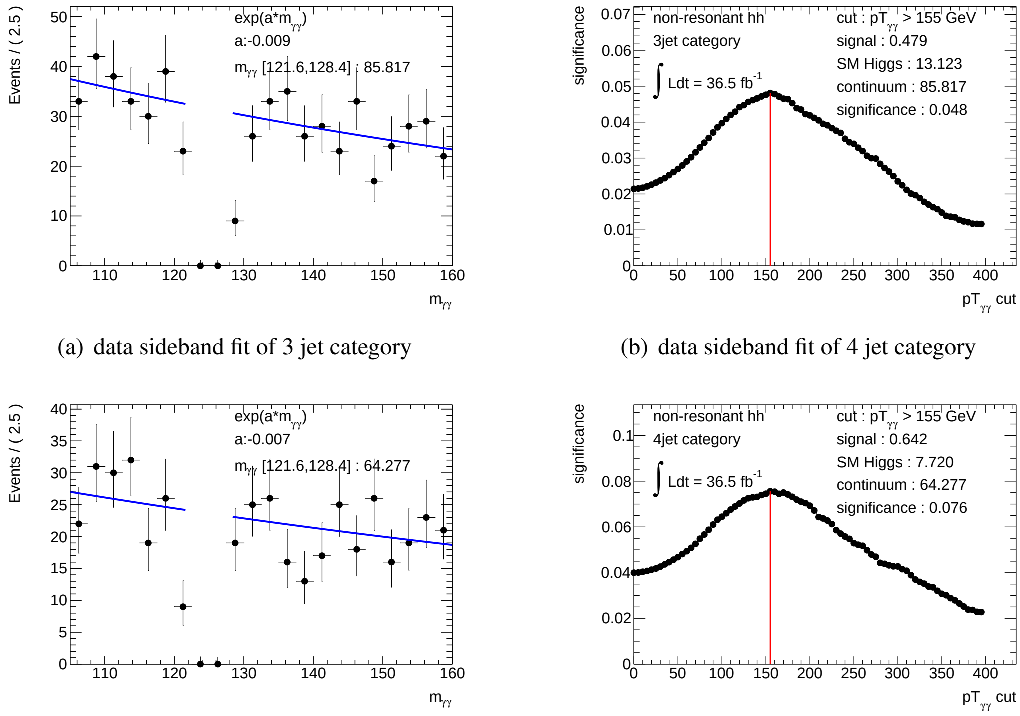


(c) cut scan of 3 jet category



(d) cut scan of 4 jet category

Figure 8: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 500 GeV.



(a) data sideband fit of 3 jet category

(b) data sideband fit of 4 jet category

(c) cut scan of 3 jet category

(d) cut scan of 4 jet category

Figure 9: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for non-resonant.

6.2 Strategy 2

6.2.1 Jet combination

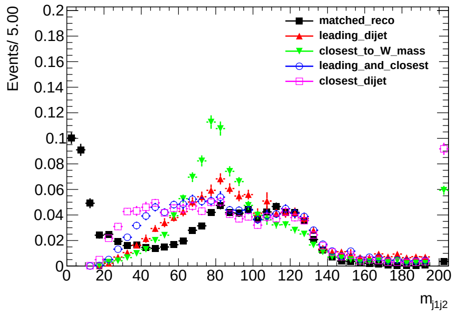
Some strategies are considered to reconstruct one on-shell W boson. The details are listed in Table 6. Figure 10 shows the invariant mass distribution in different signal samples. The method of dijet closest to W mass has best mass resolution and it is used to selection the on-shell W. Another dijet system which invariant mass is close 40 GeV is taken as an off-shell W candidate. Figure 11 shows the 2D distribution of invariant mass of selected on-shell and off-shell W boson. From the plot, a rough W mass window cut, $|m_{on-shell\ W} - 80GeV| < 20GeV$ and $m_{off-shell\ W} < 80GeV$, is determined. After a further investigation, it is found that this constraint is not helpful to the sinificance, so this constraint is not introduced finally. More details about jet combination is discussed in Appendix **??**.

method to select on-shell W	description	match efficiency of on-shell W candidate
leading 4 jets	select the 4 leading jets. take leading two jets as on-shell W and subleading two jets as off-shell W.	1
leading and closest	select a leading jet and another closest as on-shell W boson	1
dijet closest to W mass	select dijet which invariant mass is closest to W mass	1
closest dijet	select the closest dijet	1

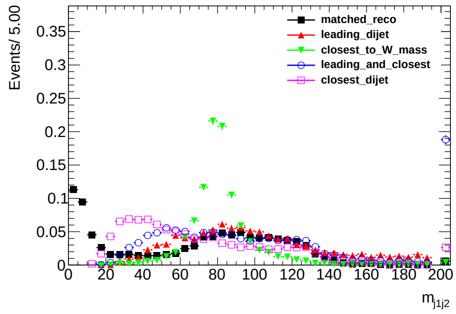
Table 6: method description of W boson reconstruction

6.2.2 Mass window of diphoton plus jet system

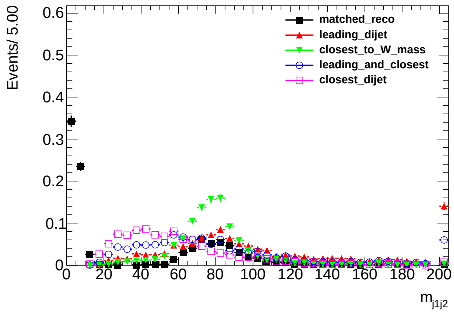
In the resonant search, mass spectrum of diphoton plus 3 or 4 jets can indicate the mass of capital Higgs, so a mass constrain is performed in resonant analysis. The mass distribution of $\gamma\gamma$+3(4) jets is shown in Figure 12 after jet selection and W mass constrain. A mass window which contains 85 % signal events is defined in Figure 13.



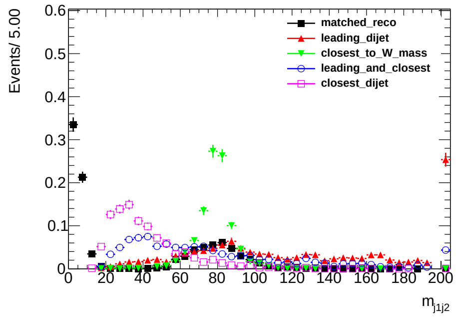
(a) 3 jet category of non-resonnat



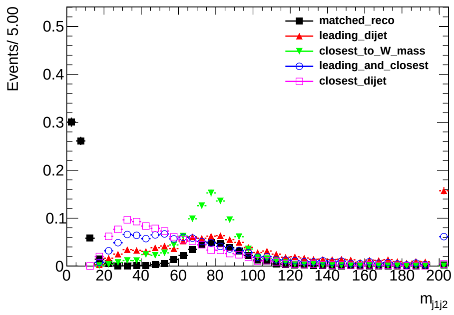
(b) 4 jet category of non-resonant



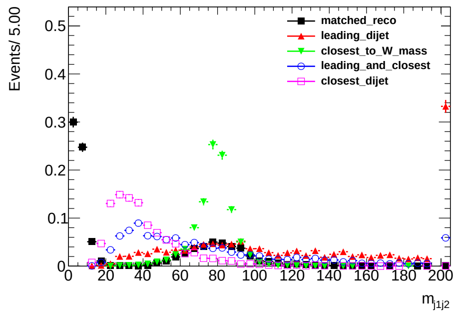
(c) 3 jet category of $m_H = 260$ GeV



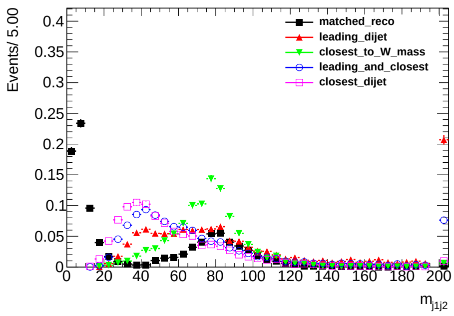
(d) 4 jet category of $m_H = 260$ GeV



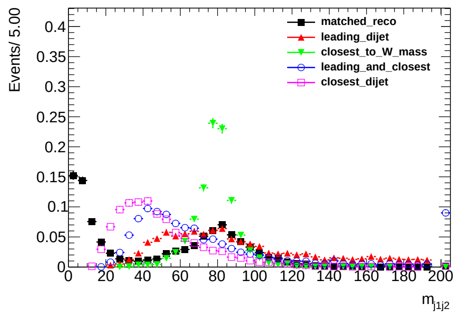
(e) 3 jet category of $m_H = 300$ GeV



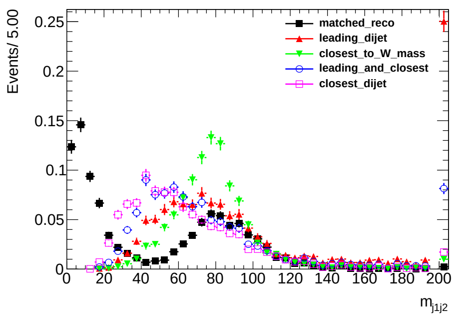
(f) 4 jet category of $m_H = 300$ GeV



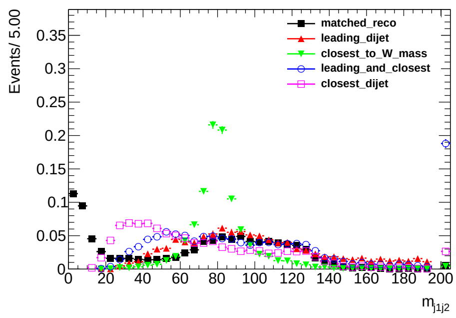
(g) 3 jet category of $m_H = 400$ GeV



(h) 4 jet category of $m_H = 400$ GeV

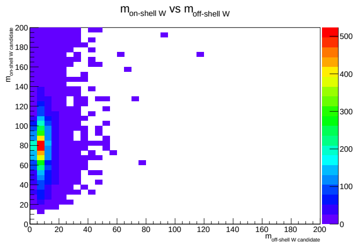


(i) 3 jet category of $m_H = 500$ GeV

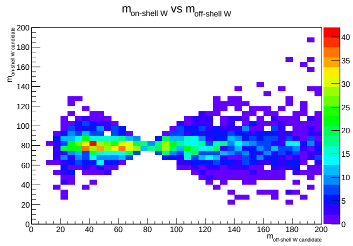


(j) 4 jet category of $m_H = 500$ GeV

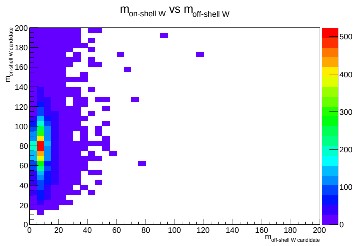
Figure 10: invariant mass distribution of on-shell W candidate in different signal samples



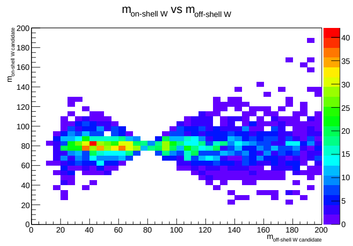
(a) 3 jet category of data sideband



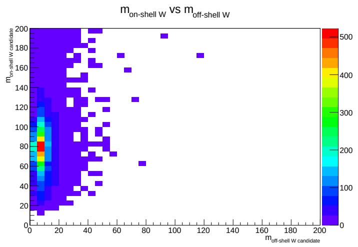
(b) 4 jet category of data sideband



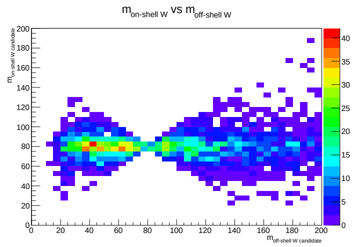
(c) 3 jet category of non-resonant



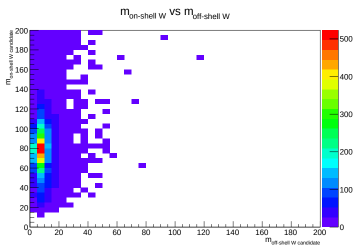
(d) 4 jet category of non-resonant



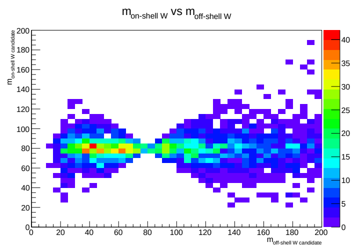
(e) 3 jet category of m_H = 260 GeV



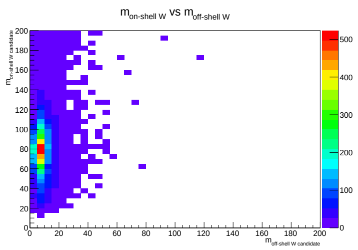
(f) 4 jet category of m_H = 260 GeV



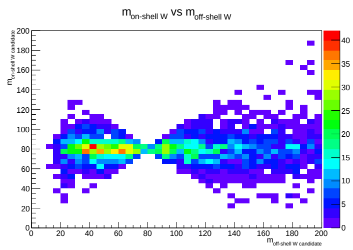
(g) 3 jet category of m_H = 300 GeV



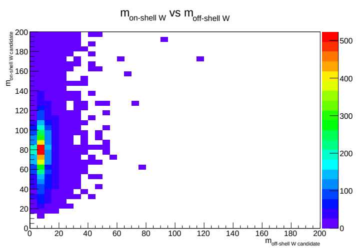
(h) 4 jet category of m_H = 300 GeV



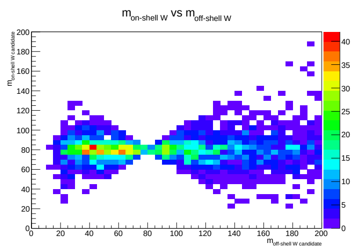
(i) 3 jet category of m_H = 400 GeV



(j) 4 jet category of m_H = 400 GeV

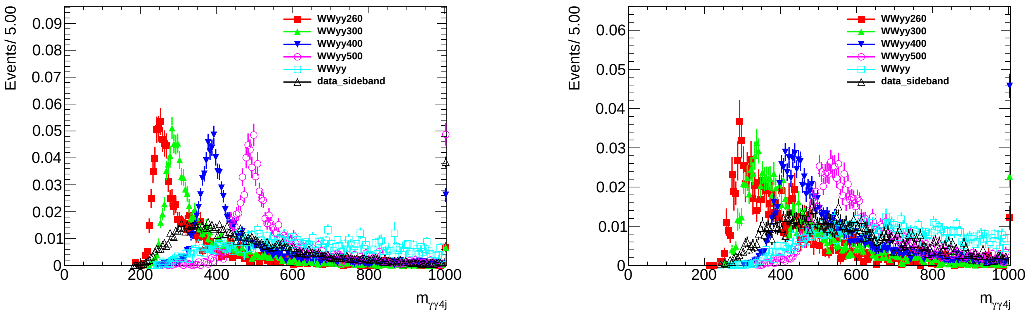


(k) 3 jet category of m_H = 500 GeV



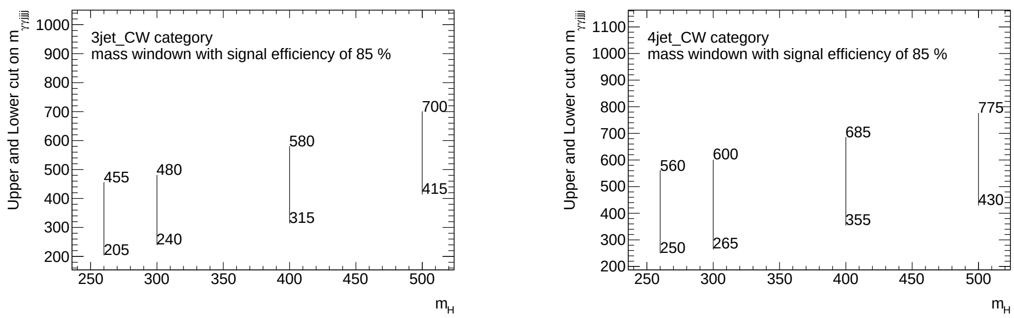
(l) 4 jet category of m_H = 500 GeV

Figure 11: 2D distribution of selected on-shell W mass vs off-shell W mass



(a) leading jet of on-shell W (b) subleading jet of on-shell W

Figure 12: pT of four signal jets at truth level



(a) mass window of 3 jet category (b) mass window of 4 jet category

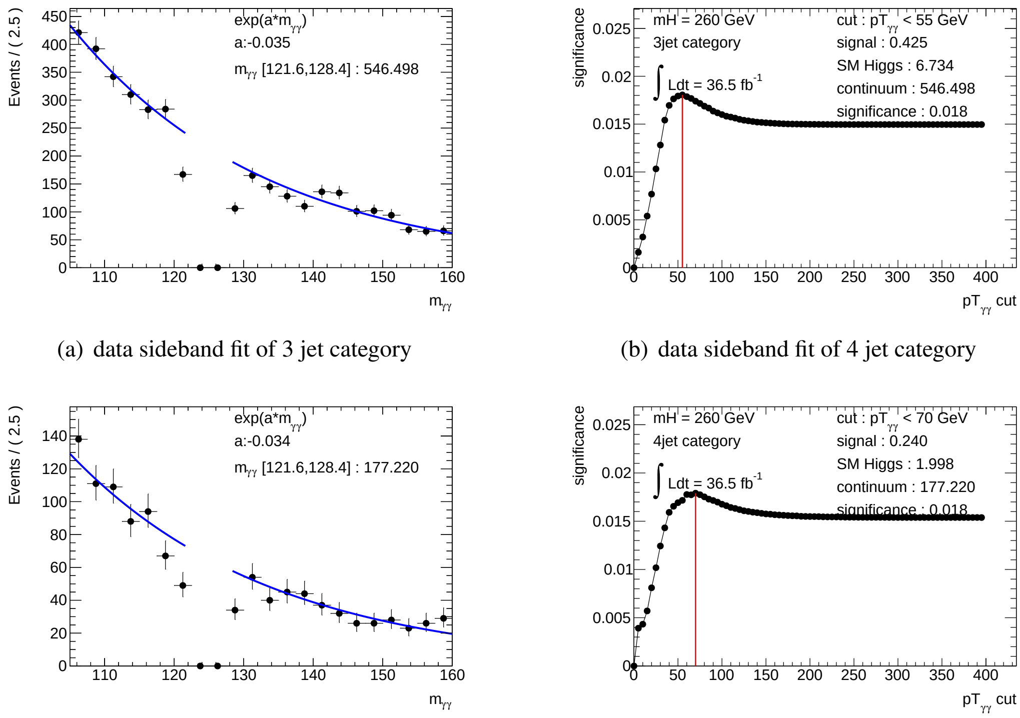
Figure 13: mass window containing 85 % signal events

	non-res		260 GeV		300 GeV		400 GeV		500 GeV	
mass window [GeV]	-	-	[205, 455]	[250, 560]	[240, 480]	[265, 600]	[315, 580]	[355, 685]	[415, 700]	[430, 775]
$pT_{\gamma\gamma}[GeV]$ cut	> 155	> 155	< 55	< 70	> 40	> 25	> 130	> 120	> 175	> 170
signal yield	0.48	0.64	0.43	0.24	0.50	0.37	0.48	0.48	0.51	0.55
SM Higgs	13.12	7.72	6.73	2.00	18.18	7.00	7.88	6.46	4.10	2.70
continuum	85.82	64.28	546	177	754	293	81	53	28	19
significance	0.048	0.076	0.018	0.018	0.018	0.022	0.044	0.065	0.089	0.139
combined significance	0.090		0.025		0.028		0.077		0.166	

Table 7: summary of selection, yield and significance

6.2.3 $pT_{\gamma\gamma}$ cut

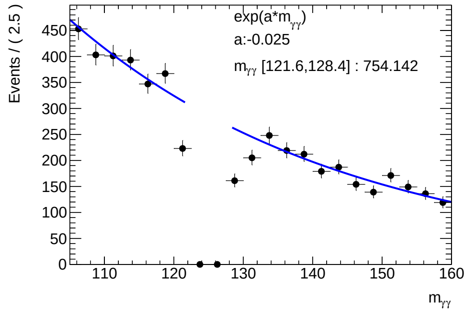
For both resonant and non-resonant search, $pT_{\gamma\gamma}$ has good separation power. Its distribution is shown in Figure 4. The two SM Higgs are more boosted with higher mass of capital Higgs. A scan on $pT_{\gamma\gamma}$ is performed to determine the best cut value for resonant and non-resonant search. The background yield in tight mass window is extracted from an exponential fit on data sideband. The expected signifcance can be calculated with expected signal yield, SM Higgs yield and continuum background yield. Figure 14 15 16 17 18 discuss the detailed scan procedure. All the expected numbers are listed in Table 7. Significance is calculated by Eq 1.



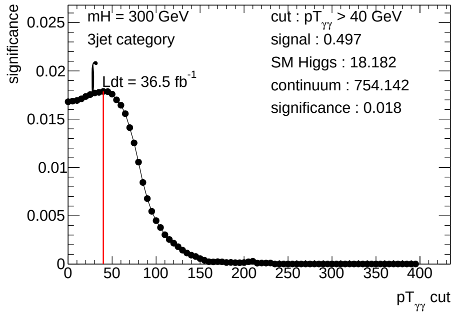
(a) data sideband fit of 3 jet category (b) data sideband fit of 4 jet category

(c) cut scan of 3 jet category (d) cut scan of 4 jet category

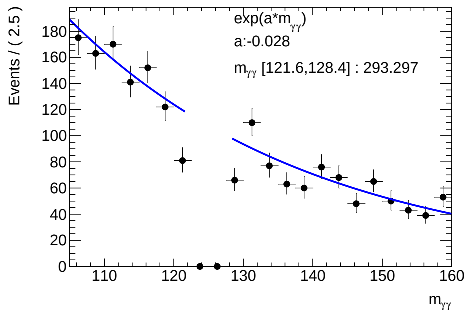
Figure 14: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 260 GeV.



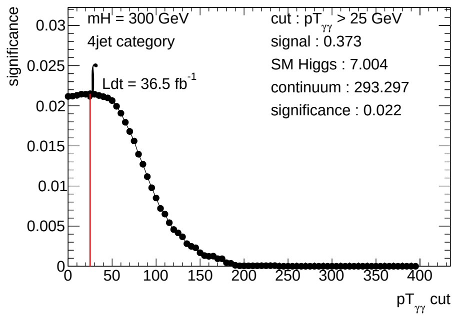
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

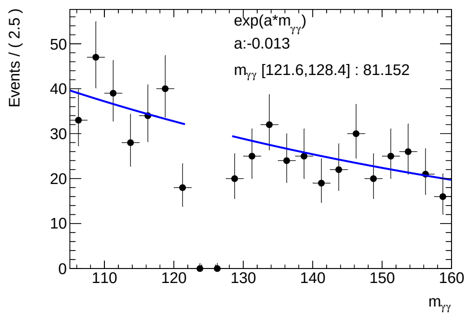


(c) cut scan of 3 jet category

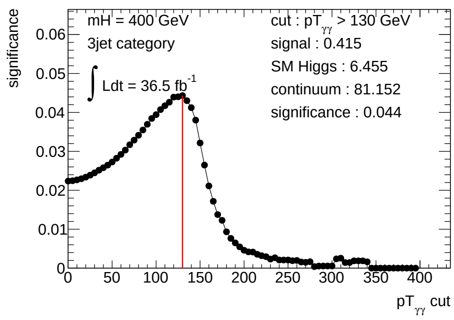


(d) cut scan of 4 jet category

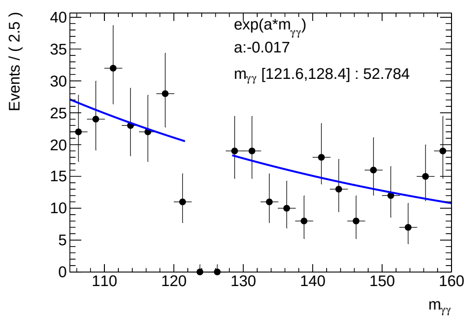
Figure 15: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 300 GeV.



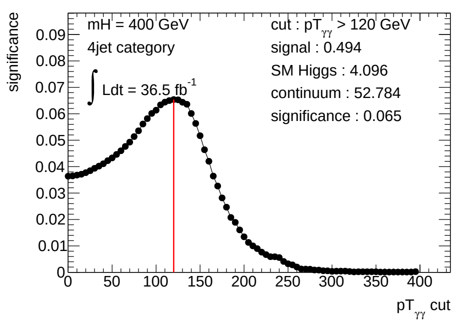
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

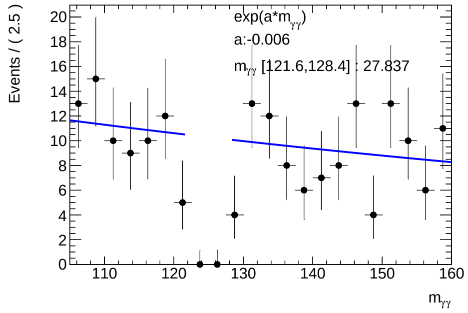


(c) cut scan of 3 jet category

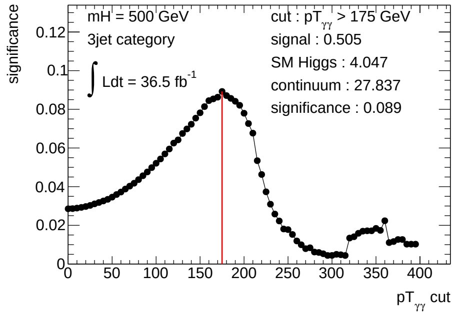


(d) cut scan of 4 jet category

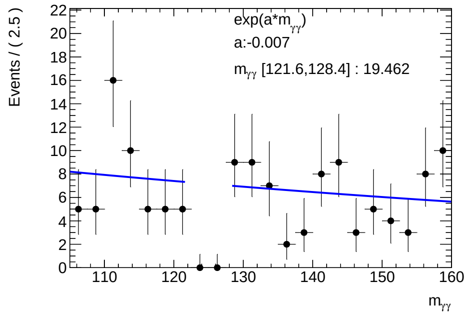
Figure 16: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 400 GeV.



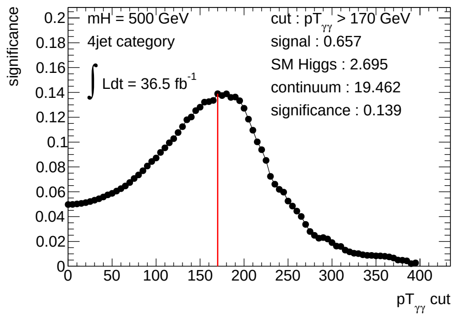
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category

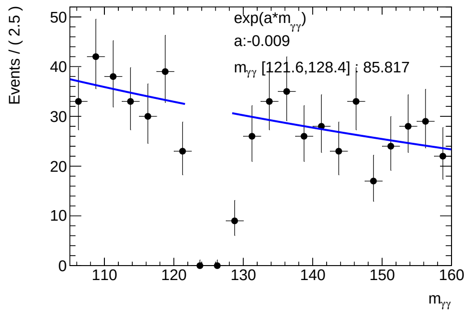


(c) cut scan of 3 jet category

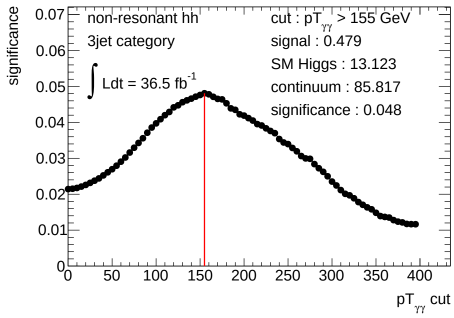


(d) cut scan of 4 jet category

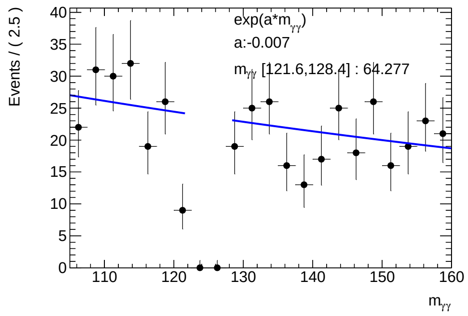
Figure 17: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for m_H = 500 GeV.



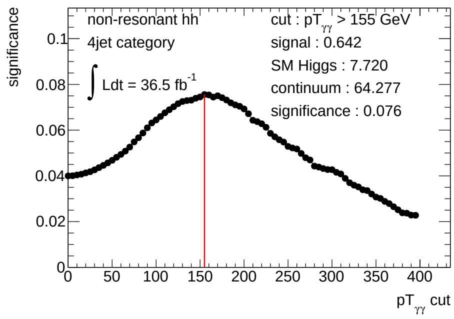
(a) data sideband fit of 3 jet category



(b) data sideband fit of 4 jet category



(c) cut scan of 3 jet category



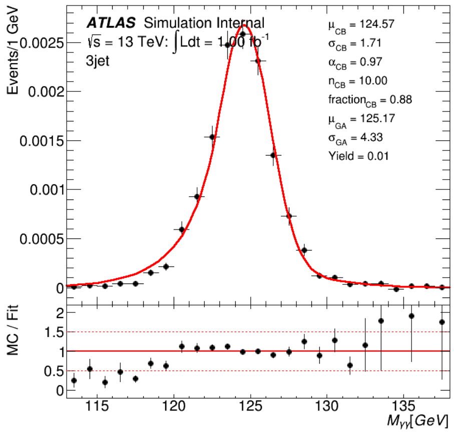
(d) cut scan of 4 jet category

Figure 18: scan the $pT_{\gamma\gamma}$ cut and plot the data sideband fit for non-resonant.

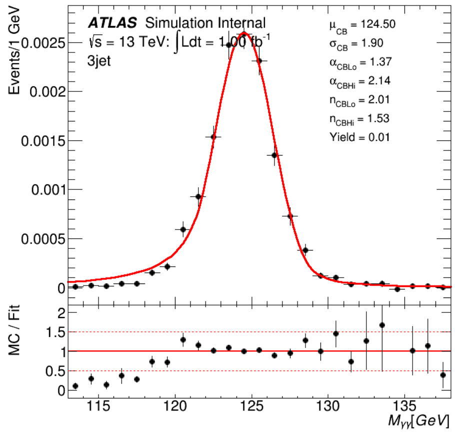
7 Signal and background estimations

7.1 Signal modeling

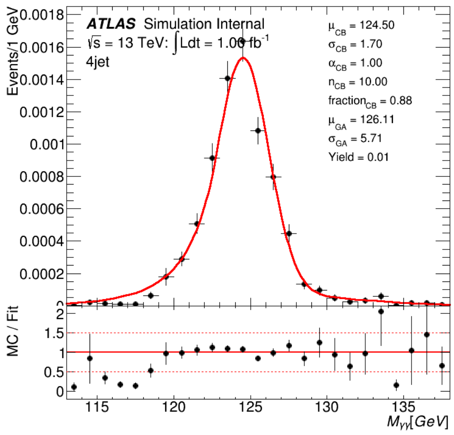
Similar to $h \to \gamma\gamma$ analysis, the signal shape of this analysis can be modelled by Double-Sided Crystal-ball(DSCB) function of Crystal-ball plus Gaussian function(CBGA). Figure 19 show the fit on $m_{\gamma\gamma}$.



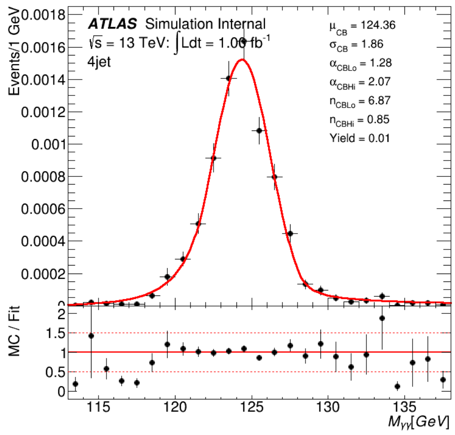
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category



(c) CBGA fit in 4 jet category

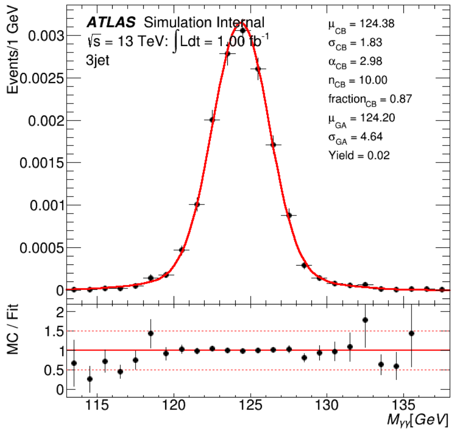


(d) DSCB fit in 4 jet category

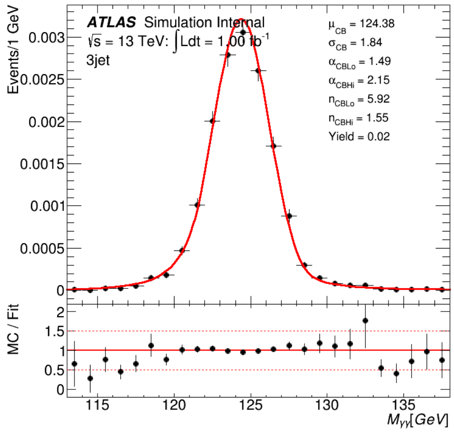
Figure 19: CBGA and DSCB fit in m_H = 260 GeV

7.2 Simulation of Higgs background processes

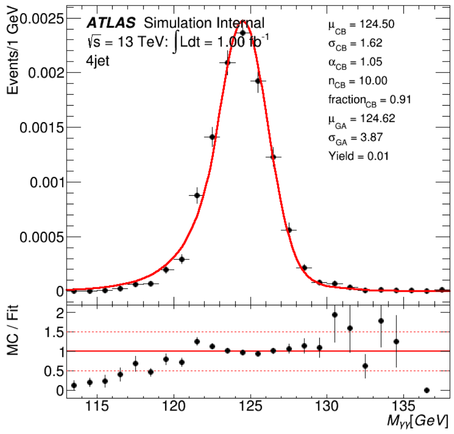
Standard Model production of a single Higgs boson with the $h \to \gamma\gamma$ decay mode are estimated using Monte Carlo simulation. Other decay modes are not considered because they do not contribute to the Higgs-mass peak in the $m_{\gamma\gamma}$ spectrum. The processes considered are gluon-gluon fusion, vector boson fusion, Higgsstrahlung, and Higgs production :in association with $t\bar{t}$. In gluon-gluon fusion events, all jets are the result of ISR. In vector boson fusion events, ISR and FSR are responsible for the extra jets in addition to the forward jets from the scattered quarks in vector boson fusion and the hadronic decay products of the W- or Z-boson in Higgsstrahlung. In the case of $t\bar{t}$+h events, a sufficient number of



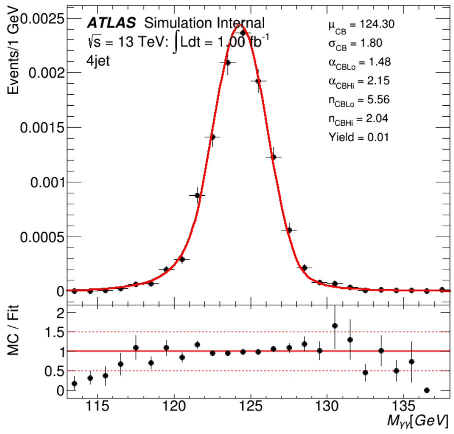
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

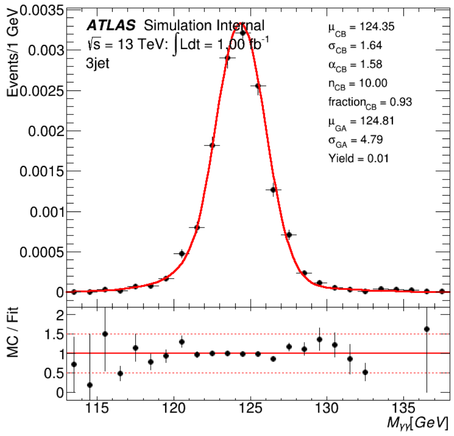


(c) CBGA fit in 4 jet category

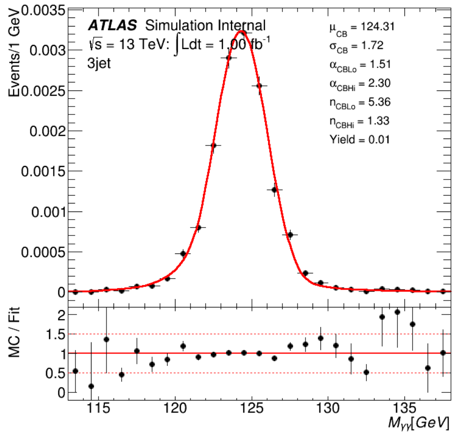


(d) DSCB fit in 4 jet category

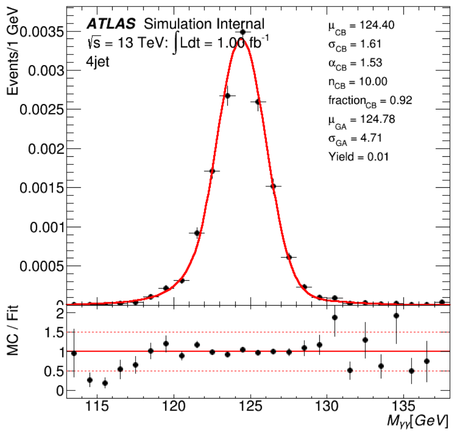
Figure 20: CBGA and DSCB fit in m_H = 300 GeV



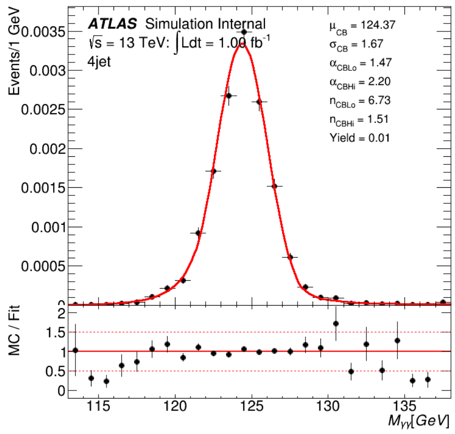
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

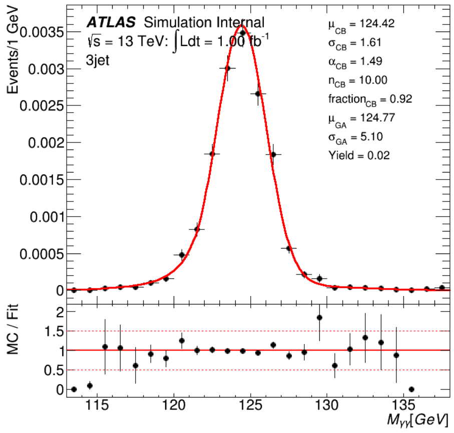


(c) CBGA fit in 4 jet category

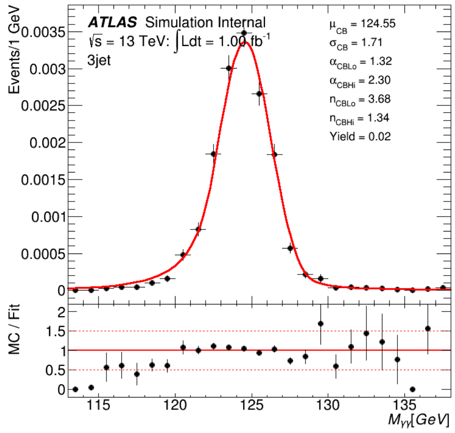


(d) DSCB fit in 4 jet category

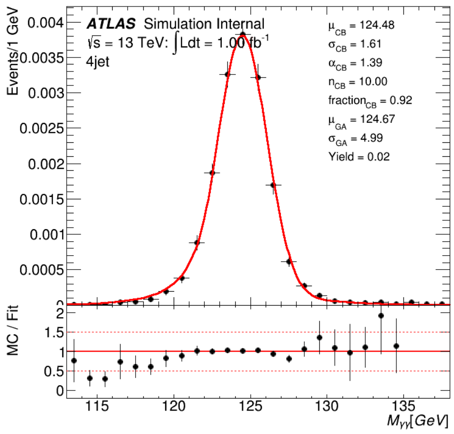
Figure 21: CBGA and DSCB fit in m_H = 400 GeV



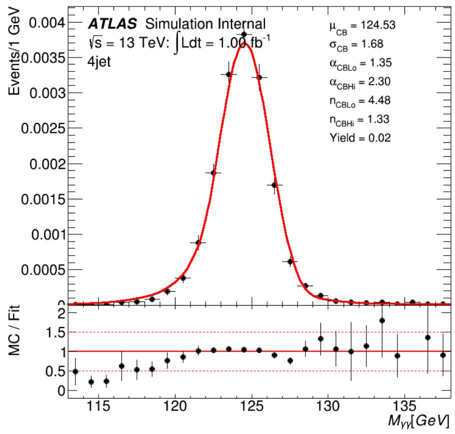
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

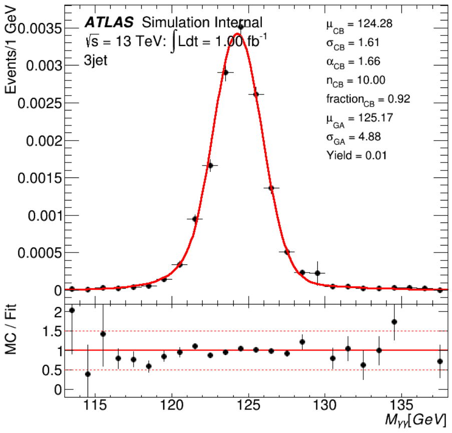


(c) CBGA fit in 4 jet category

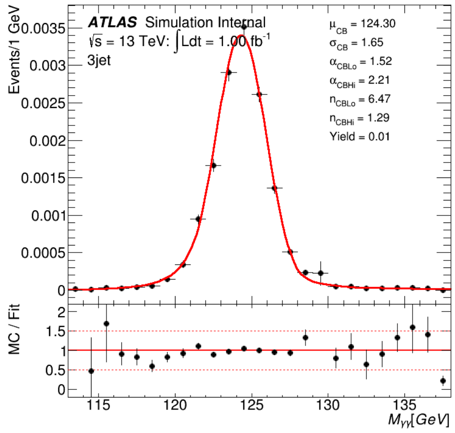


(d) DSCB fit in 4 jet category

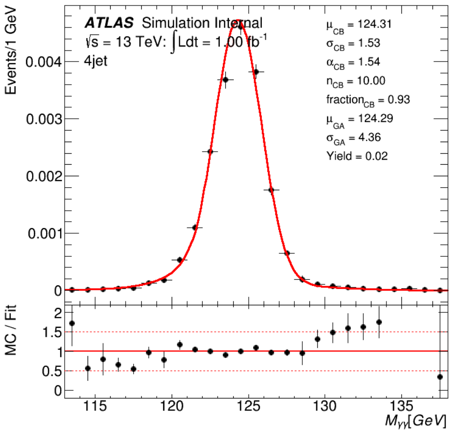
Figure 22: CBGA and DSCB fit in m_H = 500 GeV



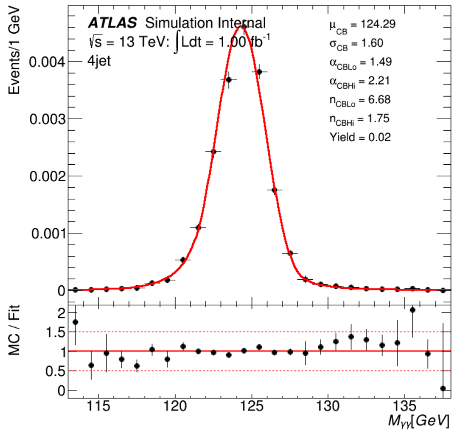
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category



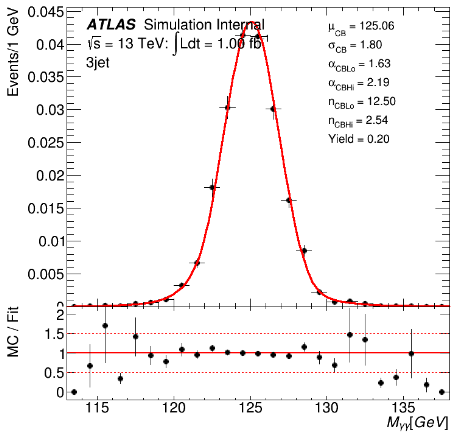
(c) CBGA fit in 4 jet category



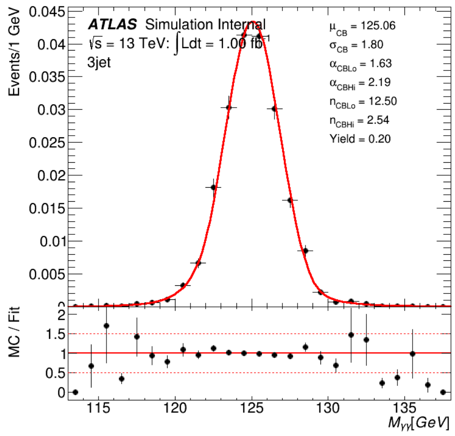
(d) DSCB fit in 4 jet category

Figure 23: CBGA and DSCB fit in non-resonant search

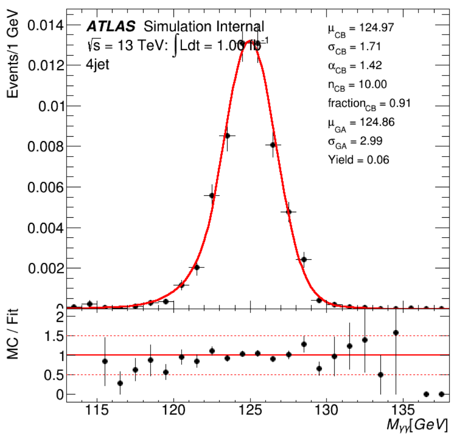
jets are produced from the decay of the two top quarks. The $m_{\gamma\gamma}$ shape SM Higgs is also modeled by Double-Sided Crystal-ball or Crystal-ball plus Gaussian function.



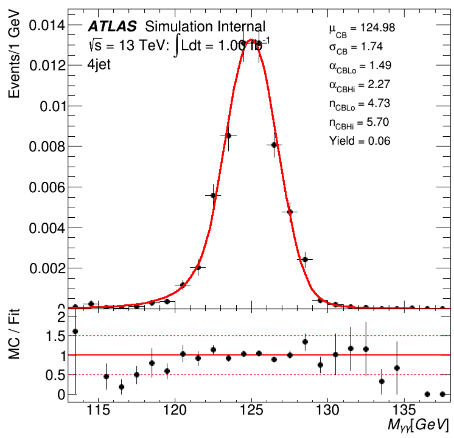
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category



(c) CBGA fit in 4 jet category



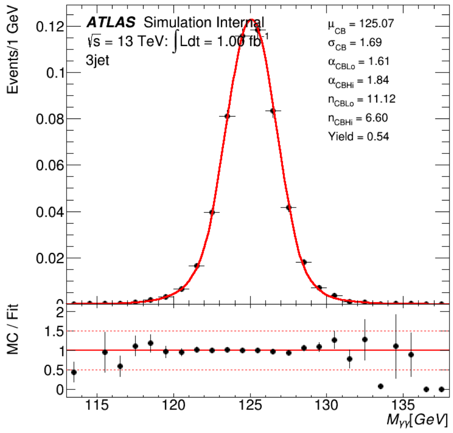
(d) DSCB fit in 4 jet category

Figure 24: CBGA and DSCB fit of SM Higgs background in $m_H = 260$ GeV

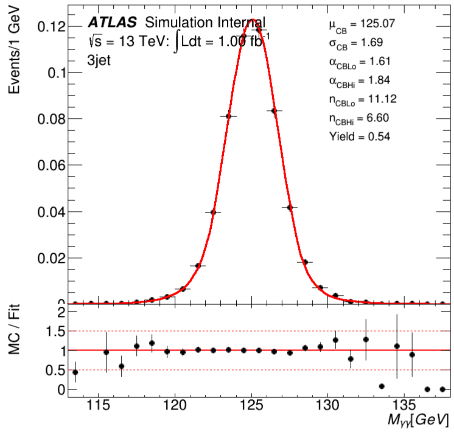
7.3 Estimating continuum background processes

The continuum background consist of $\gamma\gamma$, $\gamma - jet$ and jet-jet events. The method of spurious signal is used to choose the optimal function to describe the continuum background shape. The principle is to perform S+B fit to large statistic background-only MC sample. The fitted yield is called spurious signal, N_{sp}. The N_{sp} must pass some requirements. It must be smaller than 10 % of the expected signal yield and 20 % of the fitted signal uncertainty. If all the candidate functions pass the criteria, the function with least degree of freedom is chosen. One sample of 100M fast simulation diphoton plus up to 3 jets is produced in HGam group and it is used in spurious signal analysis. The candidate function could be exponential, 2nd-exponential, bernstein polynominal. Figure 33 shows the spurious signal test result.

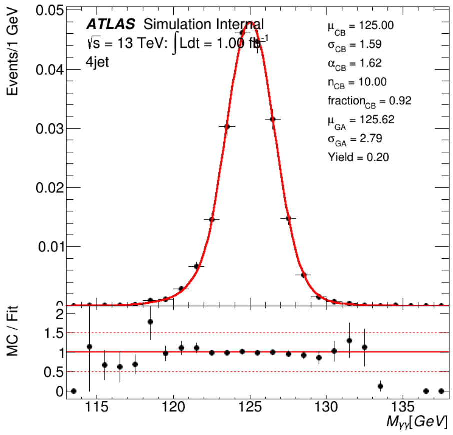
hbtp!



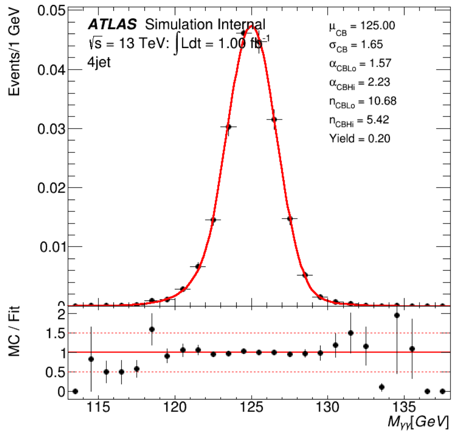
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

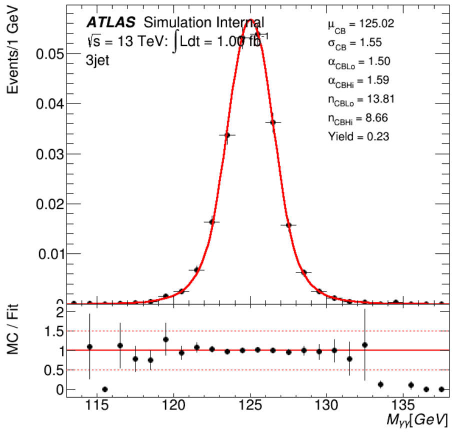


(c) CBGA fit in 4 jet category

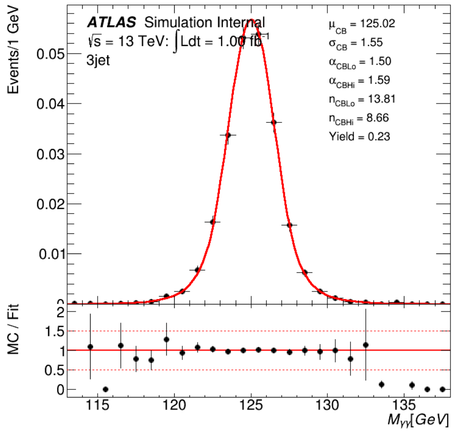


(d) DSCB fit in 4 jet category

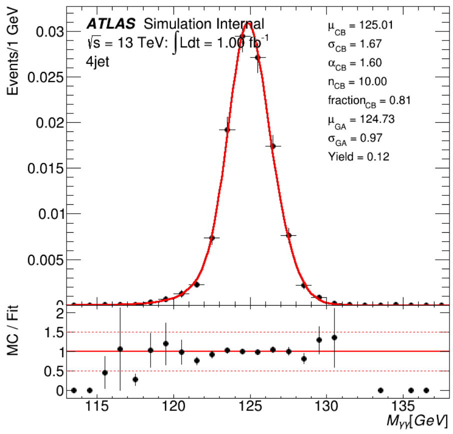
Figure 25: CBGA and DSCB fit of SM Higgs background in m_H = 300 GeV



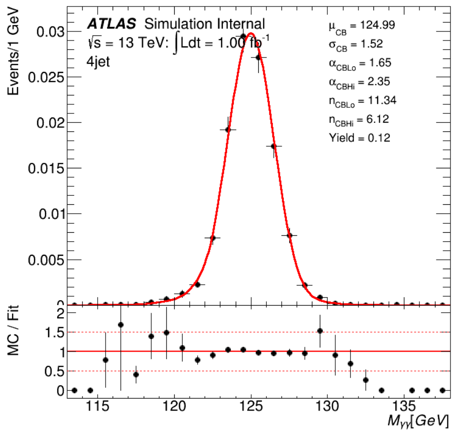
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

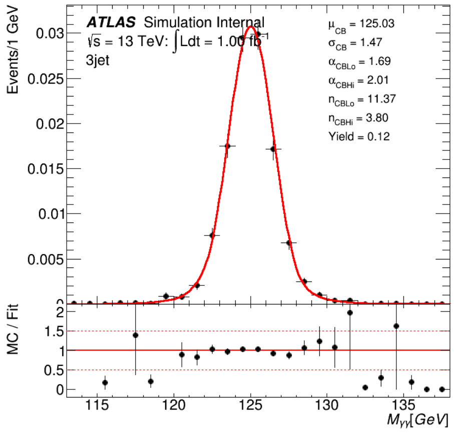


(c) CBGA fit in 4 jet category

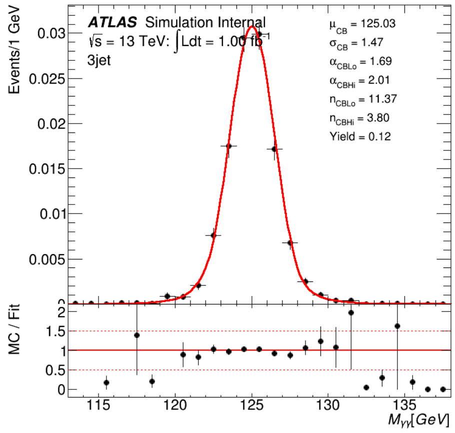


(d) DSCB fit in 4 jet category

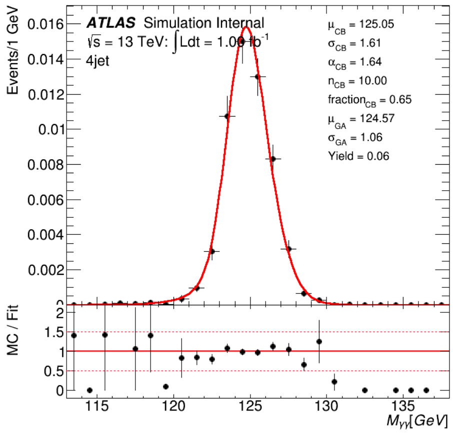
Figure 26: CBGA and DSCB fit of SM Higgs background in m_H = 400 GeV



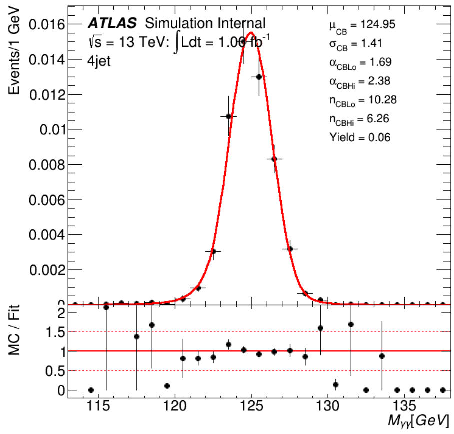
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

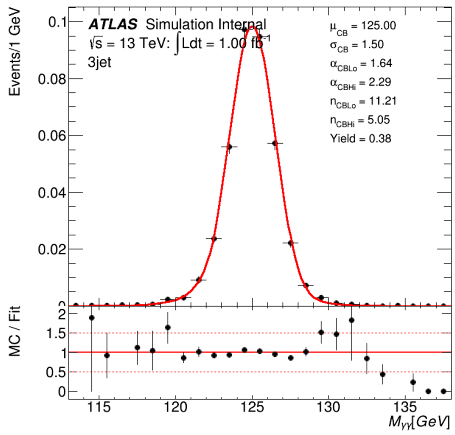


(c) CBGA fit in 4 jet category

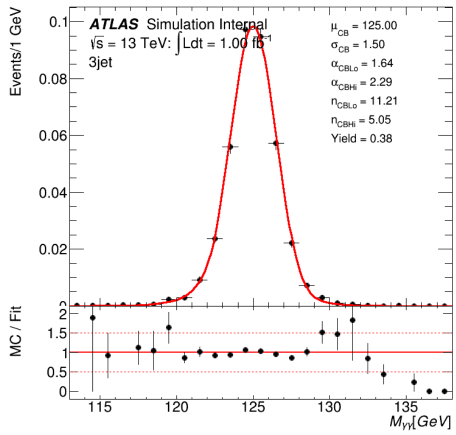


(d) DSCB fit in 4 jet category

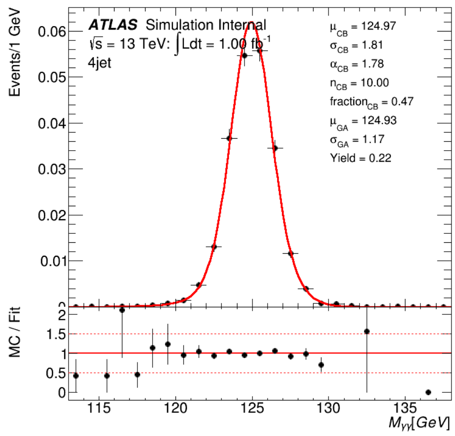
Figure 27: CBGA and DSCB fit of SM Higgs background in m_H = 500 GeV



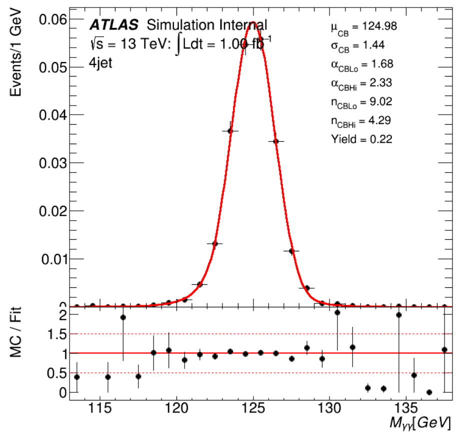
(a) CBGA fit in 3 jet category



(b) DSCB fit in 3 jet category

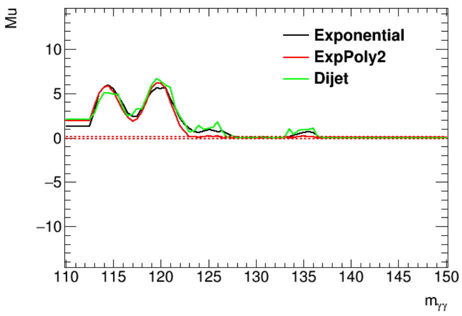


(c) CBGA fit in 4 jet category

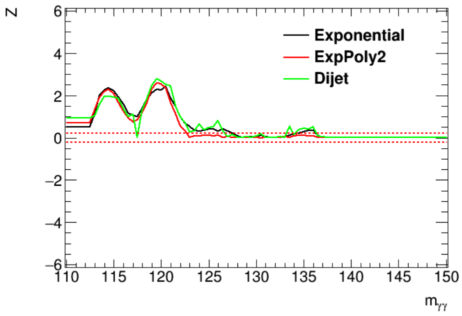


(d) DSCB fit in 4 jet category

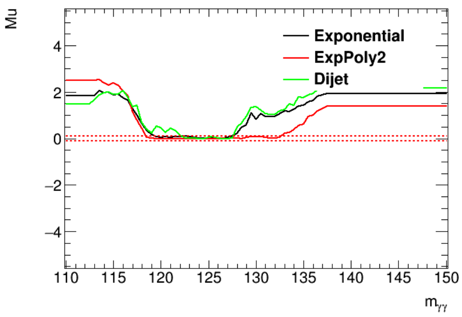
Figure 28: CBGA and DSCB fit of SM Higgs background in non-resonant search



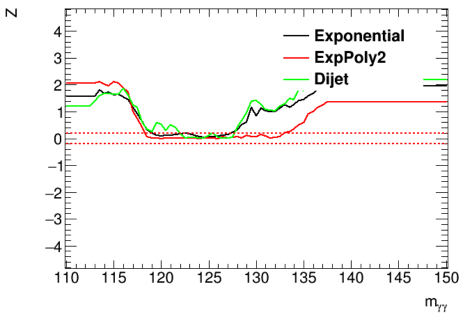
(a) S/S_{ref} of 3jet category for non-resonant



(b) $S/\Delta S$ of 3jet category for non-resonant

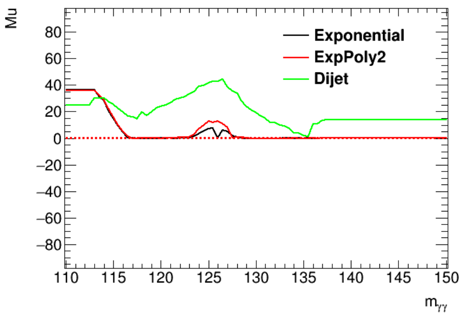


(c) S/S_{ref} of 3jet category for non-resonant

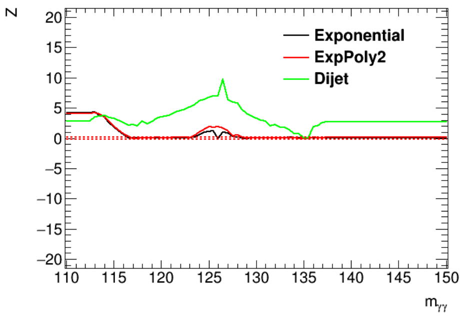


(d) $S/\Delta S$ of 3jet category for non-resonant

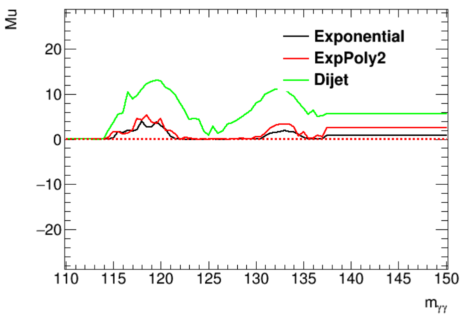
Figure 29:



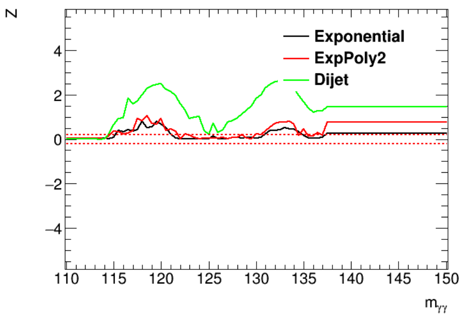
(a) S/S_{ref} of 3jet category for $m_H = 260GeV$



(b) $S/\Delta S$ of 3jet category for $m_H = 260GeV$

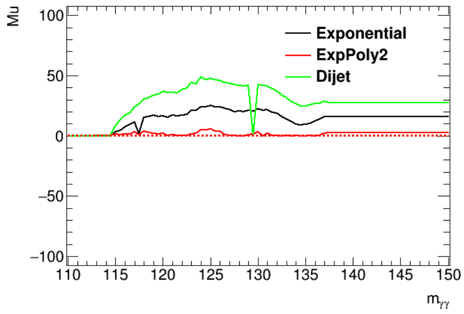


(c) S/S_{ref} of 3jet category for $m_H = 260GeV$

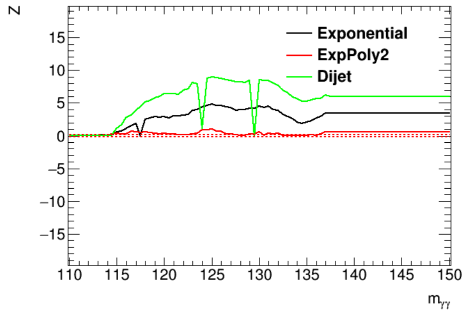


(d) $S/\Delta S$ of 3jet category for $m_H = 260GeV$

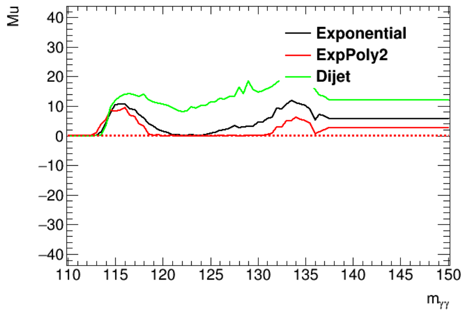
Figure 30:



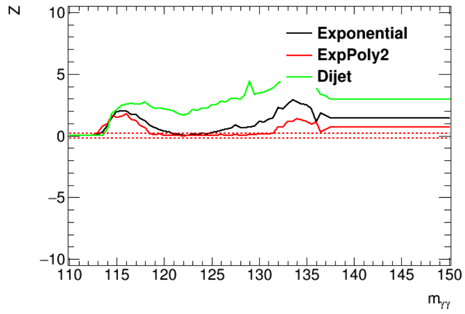
(a) S/S_{ref} of 3jet category for $m_H = 300GeV$



(b) $S/\Delta S$ of 3jet category for $m_H = 300GeV$

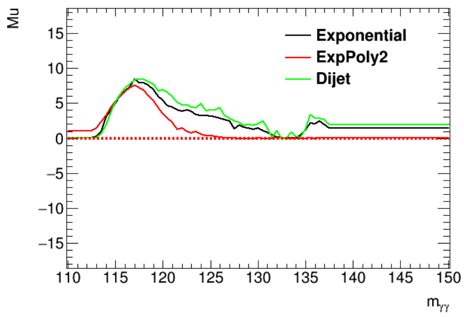


(c) S/S_{ref} of 3jet category for $m_H = 300GeV$

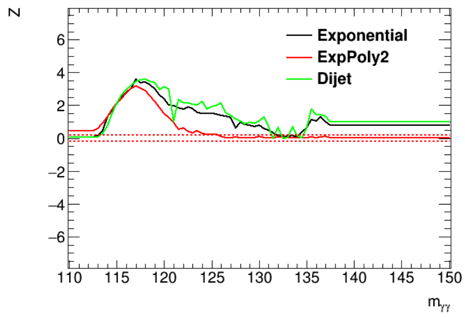


(d) $S/\Delta S$ of 3jet category for $m_H = 300GeV$

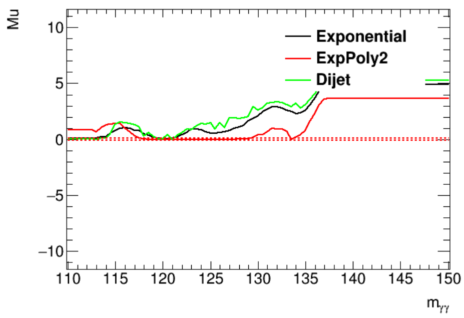
Figure 31:



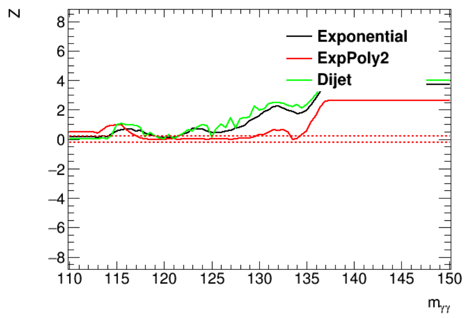
(a) S/S_{ref} of 3jet category for $m_H = 400GeV$



(b) $S/\Delta S$ of 3jet category for $m_H = 400GeV$

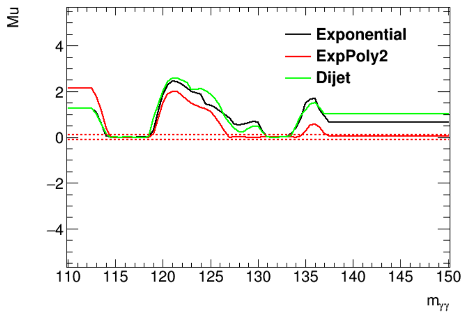


(c) S/S_{ref} of 3jet category for $m_H = 400GeV$

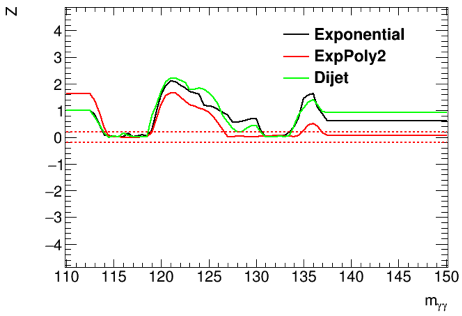


(d) $S/\Delta S$ of 3jet category for $m_H = 400GeV$

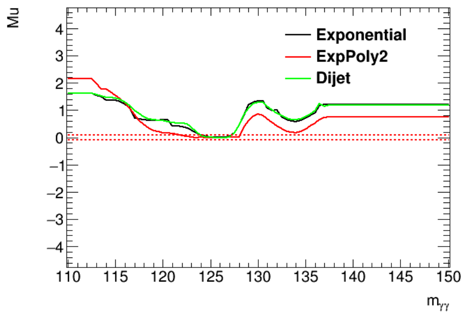
Figure 32:



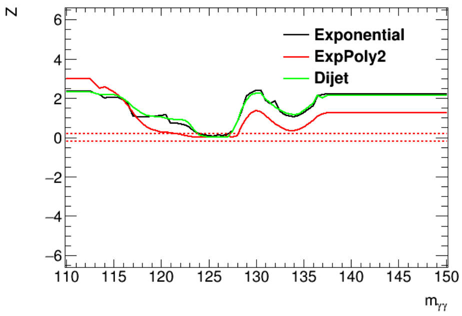
(a) S/S_{ref} of 3jet category for $m_H = 500GeV$



(b) $S/\Delta S$ of 3jet category for $m_H = 500GeV$



(c) S/S_{ref} of 3jet category for $m_H = 500GeV$



(d) $S/\Delta S$ of 3jet category for $m_H = 500GeV$

Figure 33: $S/\Delta S$ and S/S_{ref} in each category

8 Systematic uncertainties

8.1 Luminosity uncertainty

8.2 Theory uncertainties

8.2.1 Cross-section

8.2.2 PDF

8.3 Object uncertainties

8.3.1 Leptons

8.3.2 Photons

8.3.3 Jets

8.3.4 *b*-tagging

8.4 Sideband fit uncertainties

9 Statistical interpretation

10 Unblinded result

11 Summary

References

[1] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[2] Web pages. F. Maltoni, Higgs pair production. https://cp3.irmp.ucl.ac.be/projects/madgraph/wiki/HiggsPairProduction, Dec 2013.

[3] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[4] J. Grigo, J. Hoff, and M. Steinhauser, *Higgs boson pair production: top quark mass effects at NLO and NNLO*, Nucl. Phys. **B900** (2015) 412–430, arXiv:1508.00909 [hep-ph].

[5] J. Bellm et al., *Herwig++ 2.7 Release Note*, arXiv:1310.6877 [hep-ph].

[6] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[7] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C. P. Yuan, *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[8] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) no. 04, 063.

Appendices

to be finished