analysis for search of VBF Higgs to $\gamma\gamma$

November 9, 2015

Contents

1 Introduction

Introduction

2 Data and MC samples

Data and MC samples

2.1 Data Samples

Data Samples

2.2 MC Samples

MC samples

2.2.1 Signal Process

Signal Process

2.2.2 Background Process

Background Process

3 Event selection and reconstruction

Event selection and reconstruction

4 Discriminating variables

The VBF process has an unique signature,with two forward jets and little QCD radiation in the central region from hard interaction

The main non-resonant background in VBF H $\rightarrow \gamma\gamma$ is from Standard Model QCD process:$\gamma\gamma$ pairs in association with at least two jets,single γ events with at least three jets of which one jet fakes a photon,and multi-jet events where two jets fake two photons.Another important background to VBF Higgs $\gamma\gamma$ is resonant ggF Higgs $\gamma\gamma$ in association with at least two jets.Here Sherpa samples of $\gamma\gamma$ plus jets are used to simulate $\gamma\gamma$ plus jets background.γ+jet and jet-jet background are estimated by reverse identification(RevID) and reverse isolation(RevIso) samples from data,which are required to have one or two photons fail photon identification or isolation.

Variables	Definition
m_{jj}	Invariant mass of dijet
$\Delta\eta_{jj}$	Pseudo-rapidity separation of dijet
$\Delta\Phi_{\gamma\gamma,jj}$	Azimuthal angle between diphoton and dijet system
p_{Tt}	Diphoton p_T projected perpendicular to the diphoton thrust axis
$\Delta R^{min}_{\gamma,j}$	Minimum ΔR between either leadingsubleading photon and leadingsubleading jet
$\eta^{Zeppenfeld}$	$\mid\eta_{\gamma\gamma} - 0.5 * (\eta_{j1} + \eta_{j2})\mid$

Table 1: variables and definition

	$\gamma 1$	$\gamma 2$	$\gamma\gamma$
j1	$\gamma 1, j1$	$\gamma 2, j1$	$\gamma\gamma, j1$
j2	$\gamma 1, j2$	$\gamma 2, j2$	$\gamma\gamma, j2$
jj	$\gamma 1, jj$	$\gamma 2, jj$	$\gamma\gamma, jj$

Table 2: relevant system

4.1 Selecting the list of variables for the initial training

Compared with background events,di-jet of VBF signal is in forward region with hight transverse momentum and large rapidity separation.The decay products of Higgs lie in the central-rapidity region.Some discriminating variables are given from this unique angle correlation.Following variables are use in Run1 analysis in table **??** and the distribution is showed in figure **??** after VBF preselection,which requires at least two jets and $\Delta\eta_{jj} > 2, \eta^{Zepp} < 5$.

4.2 Optimizing the combination of input variables

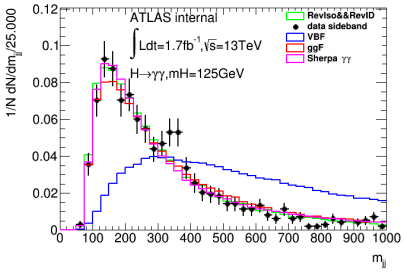
some other variables are introduced to discriminate VBF signal from ggF background and non-resonant background.The relevant system is listed in table **??**.Some variables of each system is concerned ,such as $p_T, scalarsumofp_T, scalarsumofmomentum, invariantm$ $rapidity, \Delta p_T, \Delta\Phi, \Delta R, p_T/P, E, E/M$.

5 Cut-based analysis

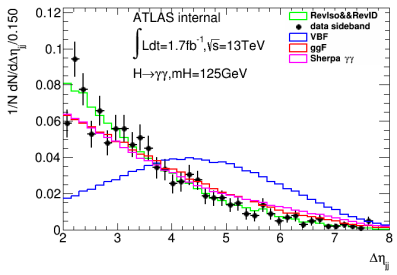
Cut-based analysis Cut-based analysis is based on TMVA package and a rectangular cut classification method,which means doing simple cut on each discriminating variablee.The combination of input variables is the same as Run1 analysis ,$m_{jj}, \Delta\eta_{jj}, \Delta\Phi_{\gamma\gamma,jj}, \Delta R^{min}_{\gamma,j}, \eta_{Zepp}$ for cut-based tight category and $m_{jj}, \Delta\eta_{jj}, \Delta\Phi_{\gamma\gamma,jj}$ for cut-based loose category.For training,VBF sample is used as signal sample and Sherpa $\gamma\gamma$,RevID and RevIso are used as background samples.Background decomposition is updated for run2.The fraction of $\gamma\gamma$ background is 80% and the fraction of γj and jj is 20%.

The cut-based training is performed by

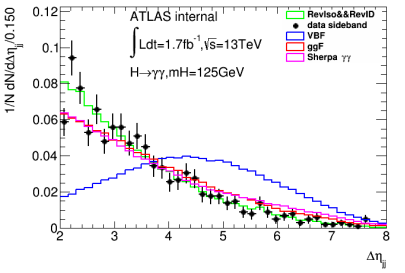
'BookMethod(TMVA::Types::kCuts, "Cuts",



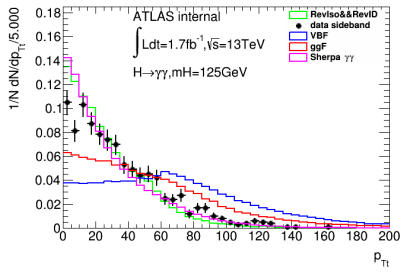
(a) m_{jj}



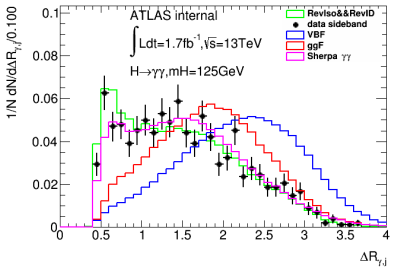
(b) $\Delta\eta_{jj}$



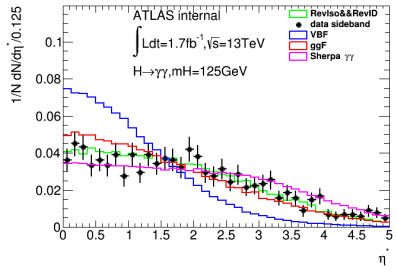
(c) $\Delta\Phi_{\gamma\gamma\,jj}$



(d) $p_{Tt\gamma\gamma}$



(e) $\Delta R^{min}_{\gamma,j}$



(f) η^{Zepp}

Figure 1: distribution of discriminating variables in different samples

"!H:!V:FitMethod=MC:EffSel:SampleSize=200000:VarProp=FSmart"
)'
The cut-based response is obtained by 'reader-¿Evaluate("cuts",signal efficiency)'.With the response of different signal efficiency ,the expected significance can be calculated with equation (**??**).

$$\sigma_{VBF} = \sqrt{2 \times ((N_{VBF} + N_{ggF} + N_{background}) \times ln(1 + \frac{N_{VBF}}{N_{ggF} + N_{background}}) - N_{VBF})} \quad (1)$$

where the VBF signal number N_{VBF} and ggF background number N_{ggF} is obtained from MC,and the non-resonant background number is obtained from a fit with data sideband in $m_{\gamma\gamma}$ spectrum. The work point is determined where the significance is largest.The plot of significance versus signal efficiency is showed in figure **??**.After VBF preselection,at least two jets and $\Delta\eta_{jj} > 2 and \eta^{Zepp} < 5$,the events are required to pass to make up cut-based category. The cut-based category requires and the remaining events make up cut-based loose category.

6 BDT analysis

BDT analysis

7 Validation of the BDT

Validation of the BDT

8 Systematic uncertainty

Systematic uncertainty

9 Statistical results

Statistical results

10 Results and conclusions

Results and conclusions