Study of Higgs pair production with $H \to b\bar{b}$ and $H \to WW \to qq\ell\nu$ for an upgraded CMS detector at the High Luminosity LHC

A. Hinzmann, B. Kilminster, C. Lange & I. Neutelings

University of Zurich

December 2015

Abstract

A study of the Higgs boson pair production where one Higgs boson decays into $b\bar{b}$ quarks and one into WW bosons in the semi-leptonic final state with a $t\bar{t}$ background is presented. The study uses simulated pp collisions at $\sqrt{s} = 14$ TeV in an upgraded CMS detector at the High Luminosity LHC assuming an integrated luminosity $L = 3000$ fb^{-1}. Kinematic variables are examined for a multivariate analysis with a Boosted Decision Tree.

1 Introduction

[under construction]

The announcement of the discovery of the Higgs boson on the 4th of July in 2012 marks the most important milestone of particle physics in the last ten years. Almost fifty years after its first proposal [1, 2], the discovery of the Higgs boson confirms the symmetry breaking of the $SU_L(2) \times U(1)_Y$ in the Standard Model by which vector bosons gain mass without spoiling gauge invariance.

This study takes a look at a new channel of Higgs pair production, which provides the most direct measurement of the Higgs boson self-coupling λ. This self-coupling might give valueable information of physics beyond the Standard Model.

Due to large backgrounds, the Higgs boson in particular was hard to find, even with the most powerful particle collider to date. Finding two Higgs boson from Higgs pair production $H \to HH$ will also prove challenging.

2 Samples

The signal and background processes are simulated with Monte Carlo samples. These only contain $bbWW \to bbqq\ell\nu$ at generator level, where events with a W-boson decaying into a tau lepton are excluded. Both generation, parton shower and hadronization are done in Pythia6. The samples were finally reconstructed with Delphes for the CMS Phase II technical proposal. Since the jets list in Delphes contains the leptons, jets within a cone of $\Delta R = 0.2$ from a lepton and an energy difference $|p_T^j - p_T^\ell|/p_T^\ell = 0.4$ are removed from the list.

Table 1: Cross sections at NNLO and $\sqrt{s} = 14$ TeV [4][5], branching ratios $\mathcal{B}$ (excluding events with a $W \to \tau\nu$) [7][8][9] and number of Monte Carlo events per process in the samples.

process	$\sigma\mathcal{B}$ [fb]	branching ratio $\mathcal{B}$	number of MC events
HH	**40**		
$HH \to bbWW \to bbqq\ell\nu$	2.88	0.072	166 483
$HH \to bbWW \to bb\ell\nu\ell\nu$	0.44	0.011	22 812
$t\bar{t}$	**984 500**		
$t\bar{t} \to bbWW \to bbqq\ell\nu$	282 552	0.287	164 661
$t\bar{t} \to bbWW \to bb\ell\nu\ell\nu$	44 303	0.045	22 546

3 Event preselection & clean-up

We select from the samples events with at least two b-jets with $p_T > 30$ GeV and $|\eta| < 2.5$, at least four jets with $p_T > 20$ GeV and $|\eta| < 2.5$, exactly one lepton with $p_T > 20$ GeV and $|\eta| < 2.5$ and missing transverse energy $\not{E}_T > 20$ GeV.

Further clean-up cuts, 60 GeV $< M_{\mathrm{bb}} <$ 160 GeV and $\Delta R_{\mathrm{bb}} < 3$ GeV, remove a significant amount of background with out affecting the signal too much. Figure 1 shows the two variables before these cuts.

In case of more than two b-jets, the b-jet pair closest in ΔR_{bb} is used for M_{bb} and other b-tagged jets are then regarded as light jets. Figure 2 shows the jet and b-jet multiplicity after the clean-up cuts.

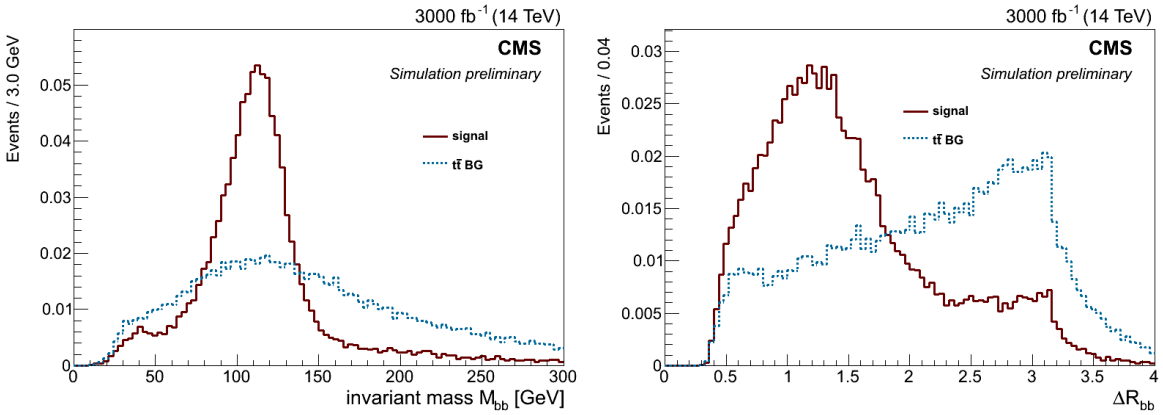

Figure 1: M_{bb} and ΔR_{bb} before clean-up.

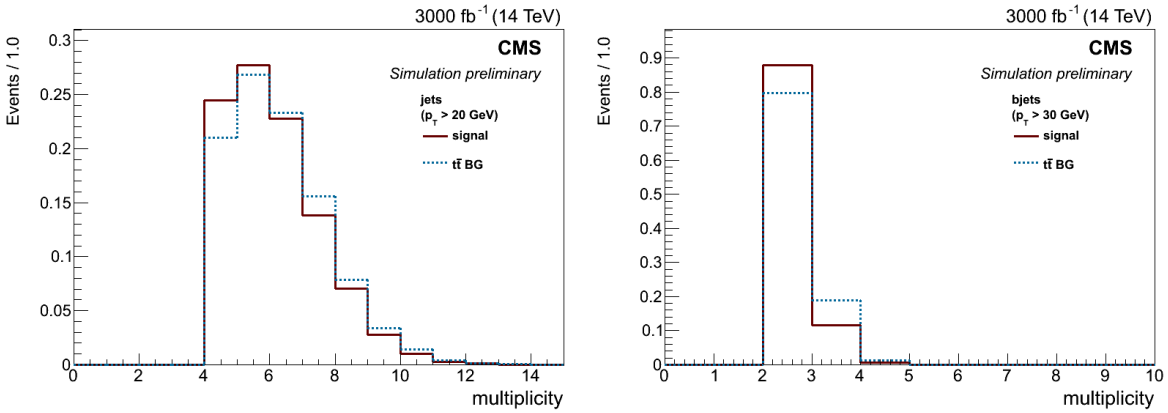

Figure 2: Multiplicities of $p_T > 20$ GeV jets and $p_T > 30$ GeV.

Table 2: Significance $P = S/(1+\sqrt{B})$ and yields $S := N(\mathrm{HH})$ and $B := N(\mathrm{t\bar{t}})$ with NNLO cross sections at $\sqrt{s} = 14$ TeV and with integrated luminosity $L = 3000$ fb^{-1}.

Selection level	P	S	B
Initial bbWW $\to$ bb$qq\ell\nu$ sample	0.297	8640	847 654 500
Selection	0.109	1496	189 235 942
Clean-up	0.130	1153	78 762 511

4 Multivariate analysis

The TMVA's boosted decision tree (BDT) is used for the multivariate analysis. The following are input variables for the BDT: p_T^{bb} of the two b-tagged jets, p_T^{jj} of the two leading "light" jets, p_T^{ℓ} of the leading lepton, $\not{E}_T$, p_T^{bb}, $p_T^{\mathrm{b}_2\ell}$, $p_T^{j_1\ell}$, $\Delta R_{j_1\ell}$, $\Delta R_{j_2\ell}$, $\Delta R_{\mathrm{b}_1\ell}$, $\Delta R_{\mathrm{b}_2\ell}$, ΔR_{bb}, ΔR_{jj}, $\Delta R_{jj,l}$, $\Delta R_{jj,\mathrm{b}_1}$,

$\Delta\phi_{j_1\ell,\mathrm{bb}}$, M_{bb}, M_{jjl}, M_{jj,b_1}, $M_{jj,\mathrm{b}}$, $M_{\mathrm{b}_2\ell\nu}$. $M_{\mathrm{b}_2\mathrm{l}}$ and $M_T^{\ell\nu}$. Here j_1 denotes the light jet closest to the lepton, and j_2 the second closest, while b_1 denotes the b-tagged jet farthest to the lepton and b_2 the second farthest. To exploit the top mass, two invariant masses reconstruct a leptonic and hadronic top as follows: the two leading jets and closest b-jet second closest to the lepton (i.e. b_1 in case of only two b-tagged jets) form M_{jj,b_1} and the lepton, reconstructed neutrino and b-jet closest to the lepton make $M_{\mathrm{b}_2\ell\nu}$. The neutrino here is reconstructed assuming its transverse momentum p_T^ν is given by the missing transverse energy and its longitudinal component p_z^ν is (the real part of) the solution of $M_{\mathrm{W}}^2 = (p_\ell + p_\nu)^2$. The transverse mass $M_T^{\ell\nu}$ is defined as

$$M_T^{\ell\nu} = \sqrt{2p_T^\ell \not{E}_T (1 - \cos\Delta\phi_{\ell,\not{p}_T})}. \tag{1}$$

All variables are shown Figs. 4-10.

5 Results

The final BDT output and background rejection versus signal efficiency of the test sample is shown in Fig. 11. A cut is made at 0.44, yielding a significance of $P = 0.37$, 27 signal events and 5153 background events at an integrated lumininosity $L = 3000$ fb^{-1}.

6 Conclusions

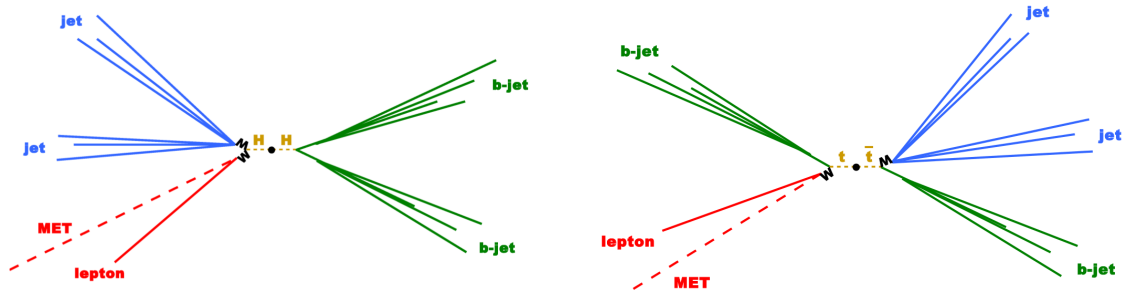

Figure 3: Sketch of a boosted Higgs boson pair and a boosted $\mathrm{t\bar{t}}$ pair.

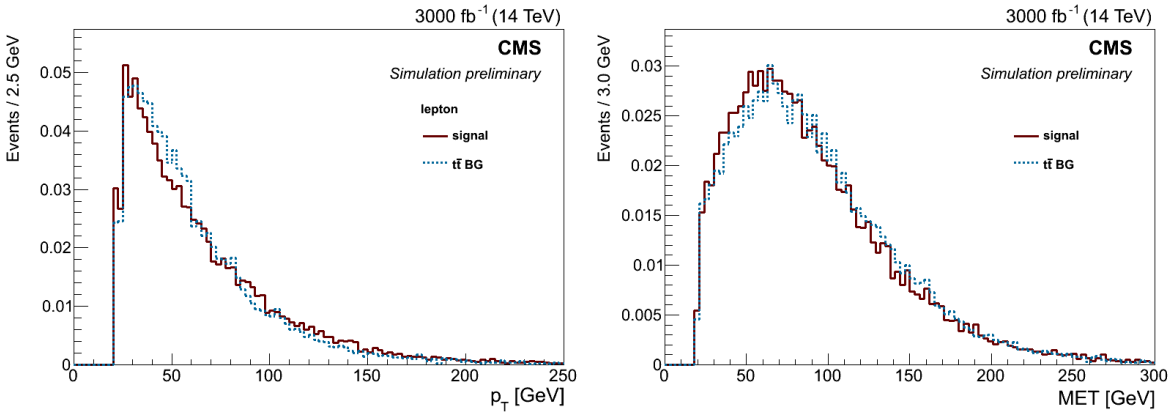

Figure 4: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: transverse momentum p_T of the lepton and missing transverse energy $\not{E}_T$.

References

[1] F. Englert & R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys. Rev. Lett. **13**, 321 (August, 1964) doi:10.1103/PhysRevLett.13.321

[2] P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett. **13**, 508 (October, 1964) doi:10.1103/PhysRevLett.13.321

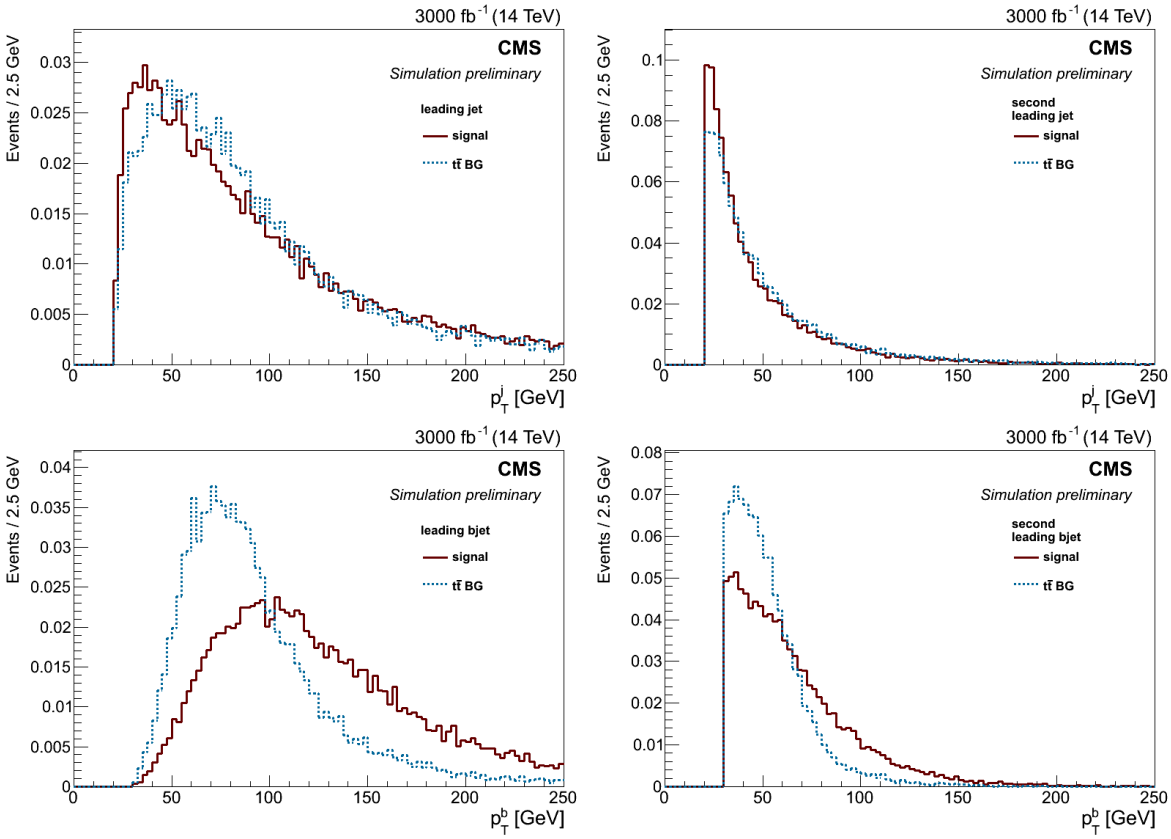

Figure 5: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: transverse momentum p_T for the two leading jets and two leading b-jets.

[3] C. Delaere *et al.*, *Study of HH production with $H \to bb$, $H \to WW \to \ell\nu\ell\nu$ for an upgraded CMS detector at the HL-LHC*, CMS draft analysis note 2014/141.

[4] D. de Florian & J. Mazzitelli, *Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*. Phys. Rev. Lett. **111** (Nov, 2013) 201801, doi:10.1103/PhysRevLett.111.201801, arXiv:1309.6594.

[5] *NNLO+NNLL top-quark-pair cross sections - ATLAS-CMS recommended predictions for top-quark-pair cross sections using the Top++v2.0 program (M. Czakon, A. Mitov, 2013)*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO#Top_quark_pair_cross_sections_at.

[6] R. Frederix *et al.*, *Higgs pair production at the LHC with NLO and parton-shower effects*, Phys. Rev. Lett. **B723** 142 (May, 2014), doi:10.1016/j.physletb.2014.03.026, arXiv:1401.7340.

[7] *Higgs cross sections for European Strategy studies in 2012*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/HiggsEuropeanStrategy2012#SM_Higgs_decay_branching_ratio_M.

[8] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ in Top-Quark-Pair Decays Using Dilepton Events and the Full CDF Run II Data Set*, Phys. Rev. Lett. **112**, 221801 (June, 2014), doi:10.1103/PhysRevLett.112.221801, arXiv:1404.3392.

[9] J. Beringer *et al.* (Particle Data Group), PR **D86**, 010001 (2012) and 2013 partial update for the 2014 edition (http://pdg.lbl.gov/2013/listings/rpp2013-list-w-boson.pdf).

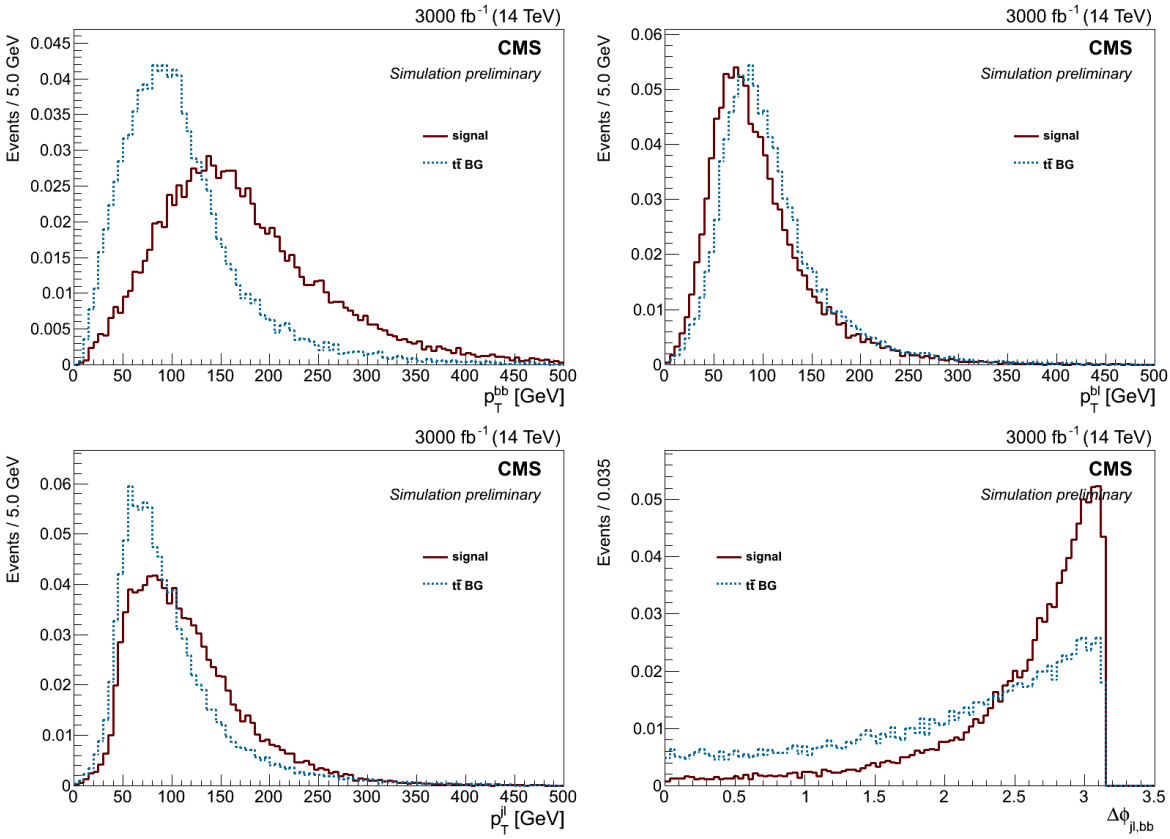

Figure 6: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: p_T^{bb}, p_T^{jj}, $p_T^{j_1\ell}$ and $\Delta\phi_{j_1\ell,\mathrm{bb}}$.

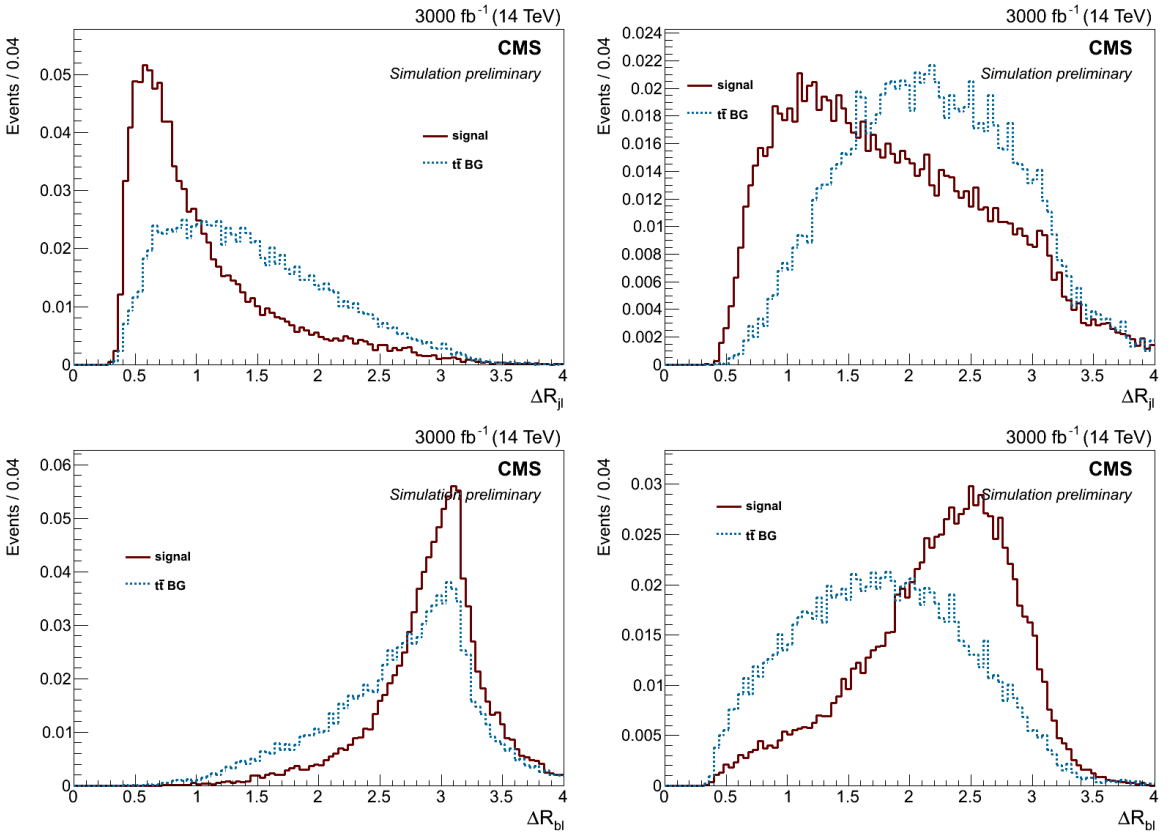

Figure 7: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: $\Delta R_{j_1\ell}$, $\Delta R_{j_2\ell}$, $\Delta R_{b_1\ell}$ and $\Delta R_{b_2\ell}$.

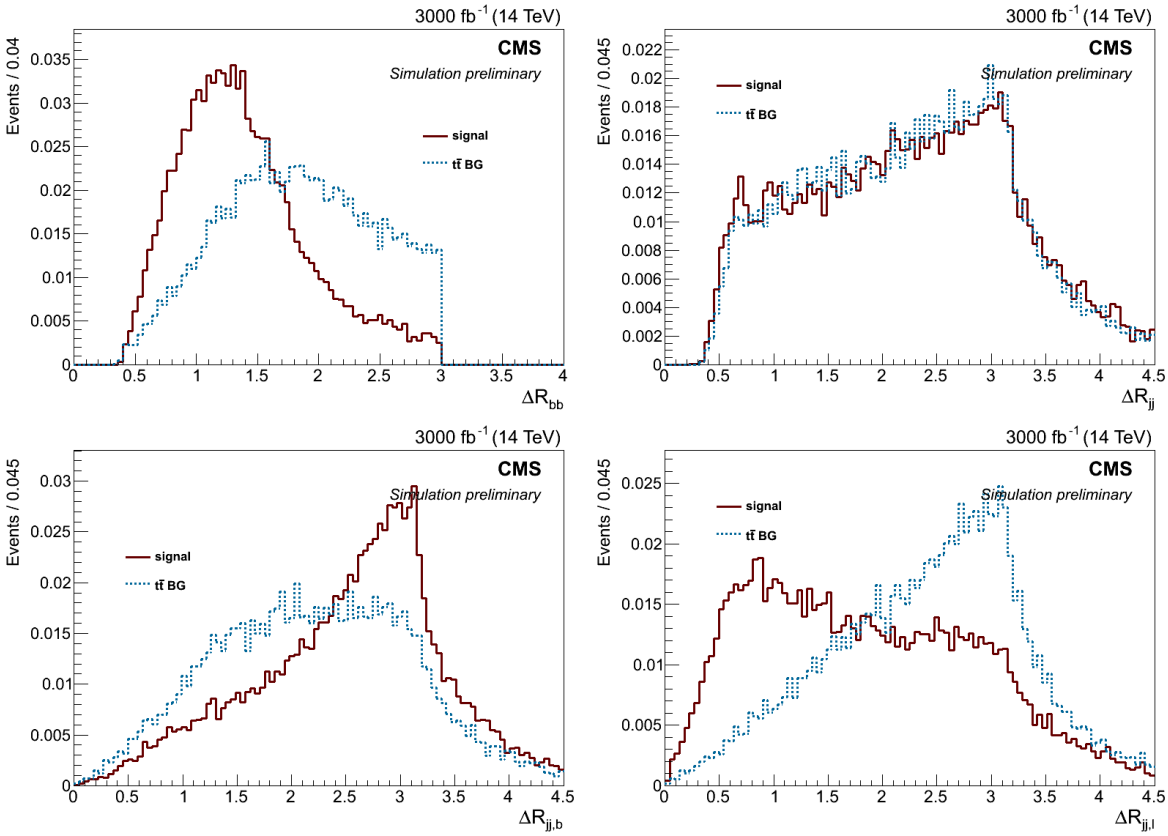

Figure 8: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: ΔR_{bb}, ΔR_{jj}, $\Delta R_{jj,b_1}$ and $\Delta R_{jj,\ell}$.

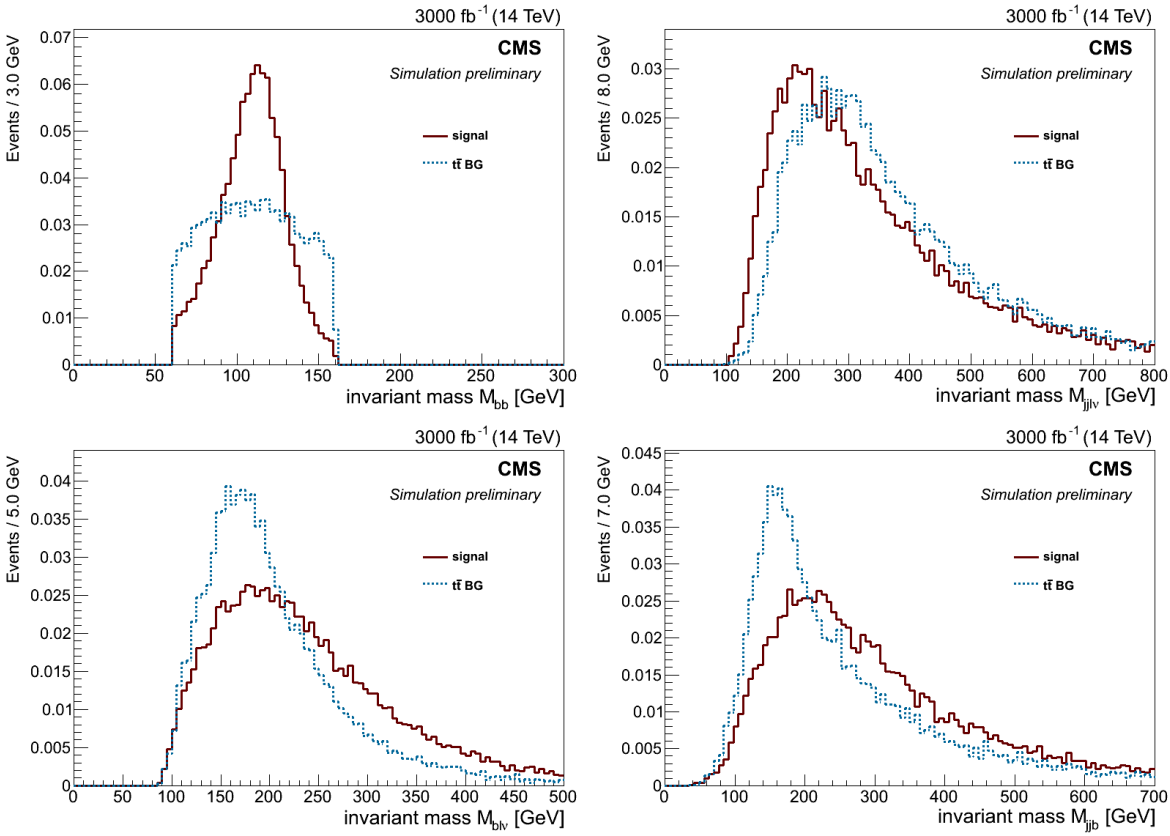

Figure 9: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: Higgs mass reconstructions M_{bb} and $M_{jj\ell\nu}$ and top mass reconstructions $M_{jj\mathrm{b}_1}$ and $M_{\mathrm{b}_2\ell\nu}$.

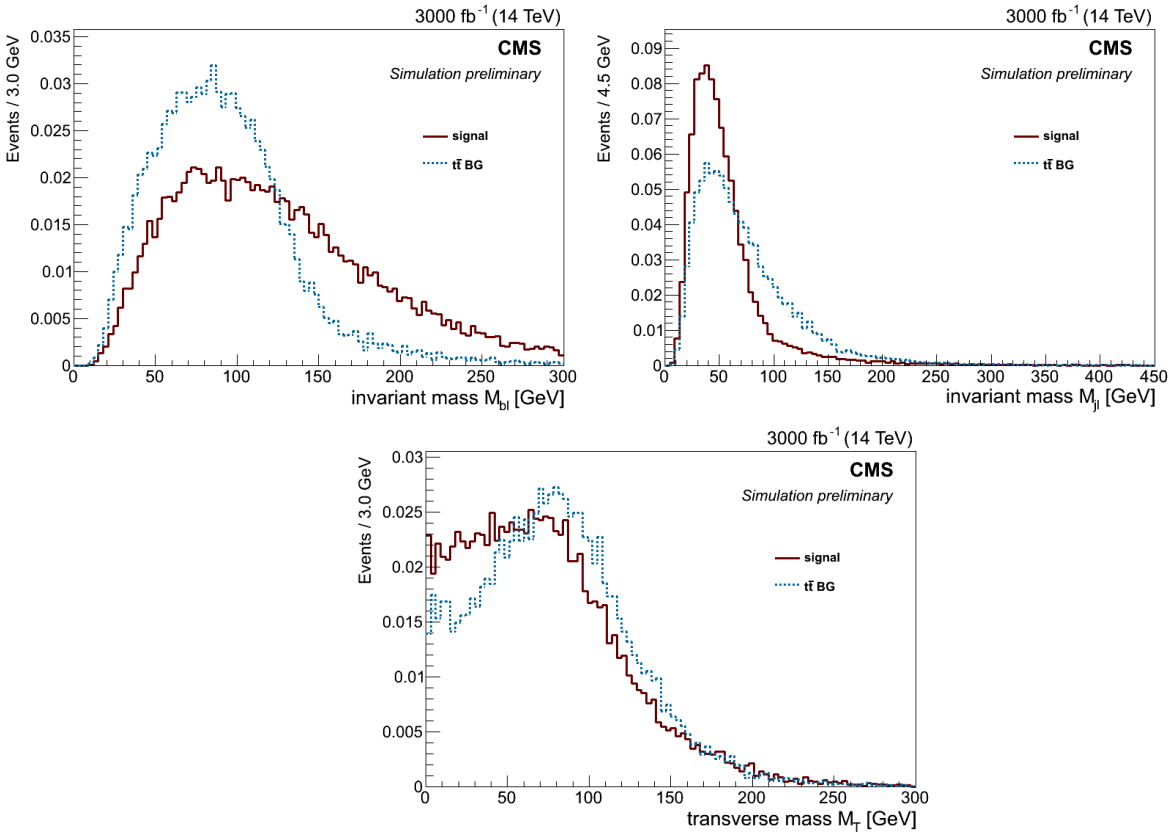

Figure 10: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: $M_{b_2 l}$ and $M_T^{\ell\nu}$ (see Eq. (1)).

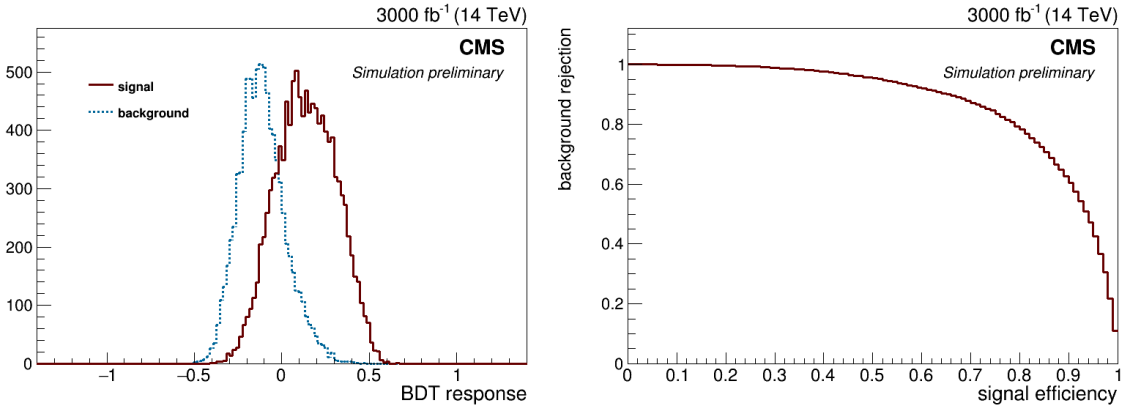

Figure 11: Final BDT output and background rejection versus signal efficiency.