Report on comparison between analyses for semi- and dileptonic channels of HH production at HL-LHC

Izaak Neutelings (University of Zurich)

December 10, 2015

This report compares results of selection level cuts of the analysis of HH $\to$ bbWW $\to$ bb$\ell\nu\ell\nu$ with the analysis note by C. Delaere *et al.* [1] and a new HH $\to$ bbWW $\to$ bb$qq\ell\nu$ analysis.

1 Samples

The Monte Carlo samples of the analysis note were generated by MadGraph_aMC@NLO, and the parton shower and hadronization was done in Pythia6. Our samples were fully produced in Pythia6. All samples were finally reconstructed with Delphes for the CMS Phase II technical proposal. For our analysis, the samples only contain bbWW $\to$ bb$\ell\nu\ell\nu$ or bbWW $\to$ bb$qq\ell\nu$ at generator level, where taus coming from a W-boson are excluded.

2 Event selections & clean-up

To reproduce the results by C. Delaere *et al.*, the event selections and clean-up as described in their analysis note are applied to our samples. Similar selections are made in the semileptonic case, however note that there are no cuts equivalent to $\Delta R_{\ell\ell}$, $M_{\ell\ell}$ or $\Delta\phi_{\mathrm{bb},\ell\ell}$ at generator or clean-up level for the semileptonic case, leading to bigger differences between the two cases. All cuts are compared in Table 1.

3 Cross section & branching ratios

For our analysis, we have been using the cross sections at $\sqrt{s} = 14$ TeV given in Table 2. The analysis note lists next leading order σ_{LO} with k-factor k_{NNLO} of the samples, all listed in Table 3.

The branching ratios are found using

$$\begin{aligned}
\mathcal{B}(\mathrm{HH} \to \mathrm{bbWW}) &= 2\mathcal{B}(\mathrm{H} \to \mathrm{bb})\mathcal{B}(\mathrm{H} \to \mathrm{WW}) \simeq 0.248 \\
\mathcal{B}(\mathrm{t\bar{t}} \to \mathrm{bWbW}) &= \mathcal{B}(\mathrm{t} \to \mathrm{bW})^2 \simeq 0.997 \\
\mathcal{B}(\mathrm{WW} \to \ell\nu\ell\nu) &= \mathcal{B}(\mathrm{W} \to \ell\nu)^2 \simeq 0.046 \\
\mathcal{B}(\mathrm{WW} \to qq\ell\nu) &= 2\mathcal{B}(\mathrm{W} \to qq)\mathcal{B}(\mathrm{W} \to \ell\nu) \simeq 0.288
\end{aligned}$$

with numbers from [5], [6] and [7]:

$$\begin{aligned}
\mathcal{B}(\mathrm{HH} \to \mathrm{bbWW} \to \mathrm{bb}\ell\nu\ell\nu) &\simeq 0.011 \\
\mathcal{B}(\mathrm{HH} \to \mathrm{bbWW} \to \mathrm{bb}qq\ell\nu) &\simeq 0.072 \\
\mathcal{B}(\mathrm{t\bar{t}} \to \mathrm{bbWW} \to \mathrm{bb}qq\ell\nu\ell\nu) &\simeq 0.045 \\
\mathcal{B}(\mathrm{t\bar{t}} \to \mathrm{bbWW} \to \mathrm{bb}qq\ell\nu) &\simeq 0.287
\end{aligned}$$

4 Significance & yield

Using the yield $N = \sigma\mathcal{B}L$ with an integrated luminosity $L = 3000$ fb^{-1}, the Punzi significance is calculated as:

$$P = \frac{S}{1+\sqrt{B}} \tag{1}$$

Table 1: Event selectiona and clean-up: comparison between the dileptonic and semileptonic final state.

dileptonic final state	semileptonic final state
Generator level filter on background	
b-quarks: $p_T > 15$ GeV	b-quarks: $p_T > 15$ GeV
leptons: $p_T > 15$ GeV, $\|\eta\| < 2.5$	lepton: $p_T > 15$ GeV, $\|\eta\| < 2.5$
$\Delta R_{\ell\ell} < 2.5$	
Selection	
two b-tagged jets: $p_T > 30$ GeV, $\|\eta\| < 2.5$	min. two b-tagged jets: $p_T > 30$ GeV, $\|\eta\| < 2.5$
	min. four jets: $p_T > 20$ GeV, $\|\eta\| < 2.5$
two oppositely charged leptons with:	one lepton with:
muons: $p_T > 20$ GeV, $\|\eta\| < 2.5$	muon: $p_T > 20$ GeV, $\|\eta\| < 2.5$
electrons: $p_T > 25$ GeV, $\|\eta\| < 2.5$	electron: $p_T > 25$ GeV, $\|\eta\| < 2.5$
$\not{E}_T > 20$ GeV	$\not{E}_T > 20$ GeV
Clean-up	
60 GeV $< M_{\mathrm{bb}} <$ 160 GeV	60 GeV $< M_{\mathrm{bb}} <$ 160 GeV
$\Delta R_{\mathrm{bb}} < 3.1$ GeV	$\Delta R_{\mathrm{bb}} < 3$ GeV
$M_{ll} < 85$ GeV	
$\Delta R_{\ell\ell} < 2$	
$\Delta\phi_{\mathrm{bb},\ell\ell} < 1.7$	

Table 2: Cross sections at NNLO and $\sqrt{s} = 14$ TeV [2][3], branching ratios $\mathcal{B}$ (excluding W $\to \tau\bar{\tau}$) [5][6][7] and number of Monte Carlo events per process in our analysis sample.

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

Table 3: Cross sections for their Monte Carlo samples listed in the analysis note by C. Deleare *et al.* [1].

process	σ_{LO} [fb]	k_{NNLO}	number of MC events
HH $\to$ bbWW $\to$ bb$\ell\nu\ell\nu$	0.163	2.3	1.1M
$\mathrm{t\bar{t}}$ full leptonic	9030	1.85	4.8M

Table 4: Number of MC events in our sample at each level of selection.

	dileptonic final state		semileptonic final state	
selection level	signal	background	signal	background
bbWW $\to$ bb$\ell\nu\ell\nu$, bb$q q \ell\nu$	22812	22546	166483	164661
Generator level filter on background	22812	8339	166483	137880
Selection	2571	1636	28821	36760
Clean-up	2066	280	22225	15300

Table 5: Comparison of the significance P and yields $S := N(\mathrm{HH})$ and $B := N(\mathrm{t\bar{t}})$ between the semileptonic and dileptonic final state for our results using the cross sections at NNLO from Table 2 and the results by C. Deleare *et al.* at $\sqrt{s} = 14$ TeV and with integrated luminosity $L = 3000\ \mathrm{fb}^{-1}$. The factor $k_{\mathcal{B}} = \sqrt{\mathcal{B}(\mathrm{WW} \to \ell\nu\ell\nu)/\mathcal{B}(\mathrm{WW} \to \mathrm{qq}\ell\nu)} \simeq 0.397$ allows for comparison.

	dileptonic final state			semileptonic final state			
	P	S	B	P	$k_{\mathcal{B}}P$	S	B
Initial bbWW → bb$qq\ell\nu$, bb$\ell\nu\ell\nu$ sample							
Assuming NNLO	0.115	1320	132 907 500	0.297	0.118	8640	847 654 500
Generator filter on background							
Assuming NNLO	0.188	1320	49 157 972	0.324	0.129	8640	709 789 218
Selection							
Assuming NNLO	0.048	149	9 644 135	0.109	0.043	1496	189 235 942
C. Delaere *et al.*	0.043	113	6 759 579				
Clean-up							
Assuming NNLO	0.093	120	1 650 586	0.130	0.052	1153	78 762 511
C. Delaere *et al.*	0.075	90	1 437 144				
Neural network							
C. Delaere *et al.*	0.60	37	3875				

with yields $S := N(\mathrm{HH})$ and $B := N(\mathrm{t\bar{t}})$. To obtain the significance at after each round of cuts, we multiply the yields with the respective selection efficiency. The total number of Monte Carlo events per sample in our analysis is listed in Table 2.

5 Results

Results are summarized in Table 5. To compare the dileptonic to the semileptonic case, the significance of the latter is also scaled by $k_{\mathcal{B}} = \sqrt{\mathcal{B}(\mathrm{WW} \to \ell\nu\ell\nu)/\mathcal{B}(\mathrm{WW} \to \mathrm{qq}\ell\nu)} \simeq 0.397$.

References

[1] 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.

[2] 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`.

[3] *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`.

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

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

[6] 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`.

[7] 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`).

Appendix A: Preliminary results on multivariate analysis of $\mathrm{HH} \to \mathrm{bbWW} \to \mathrm{bb}q q \ell \nu$

The TMVA's boosted decision tree (BDT) is used for the multivariate analysis on $\mathrm{HH} \to \mathrm{bbWW} \to \mathrm{bb}qq\ell\nu$ with background $\mathrm{t\bar{t}} \to \mathrm{bbWW} \to \mathrm{bb}qq\ell\nu$. Samples, selection and clean-up are described in sections 1 to 2. 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}, $M_{jj,\mathrm{b}}$, $M_{\mathrm{b}_2\ell\nu}$. $M_{\mathrm{b}_2\ell}$ 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. 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. 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 (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

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

All variables are shown Figs. 2-8.

The final BDT output and background rejection versus signal efficiency of the test sample is shown in Fig. 9. A cut is made at 0.44, yielding a significance of $P = 0.37$ (see Eq. (1)), 27 signal events and 5153 background events at an integrated lumininosity $L = 3000\ \mathrm{fb}^{-1}$.

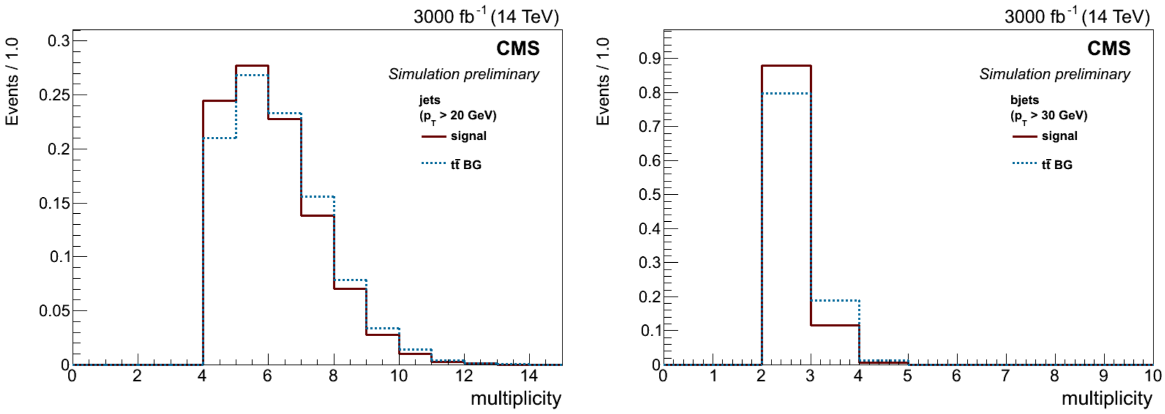

Figure 1: Multiplicities of $p_T > 20$ GeV jets and $p_T > 30$ GeV.

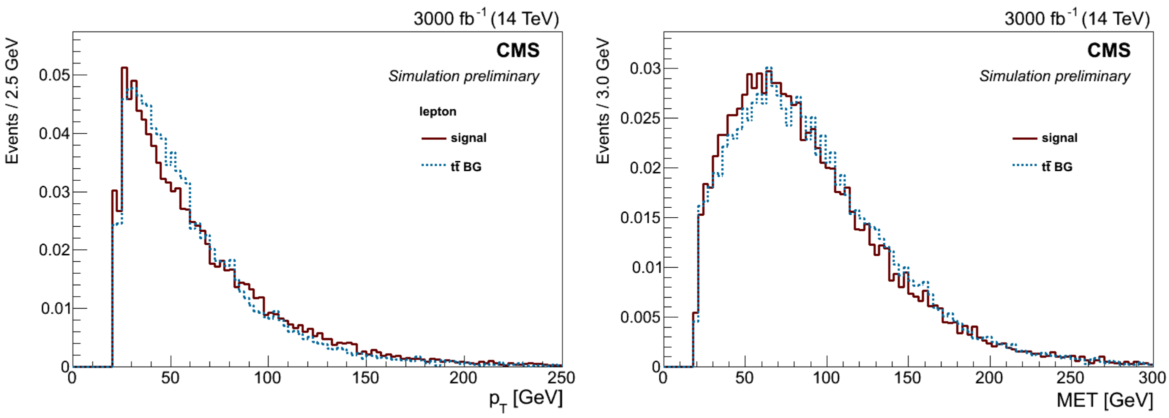

Figure 2: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: transverse momentum p_T of the lepton and missing transverse energy $\not{E}_T$.

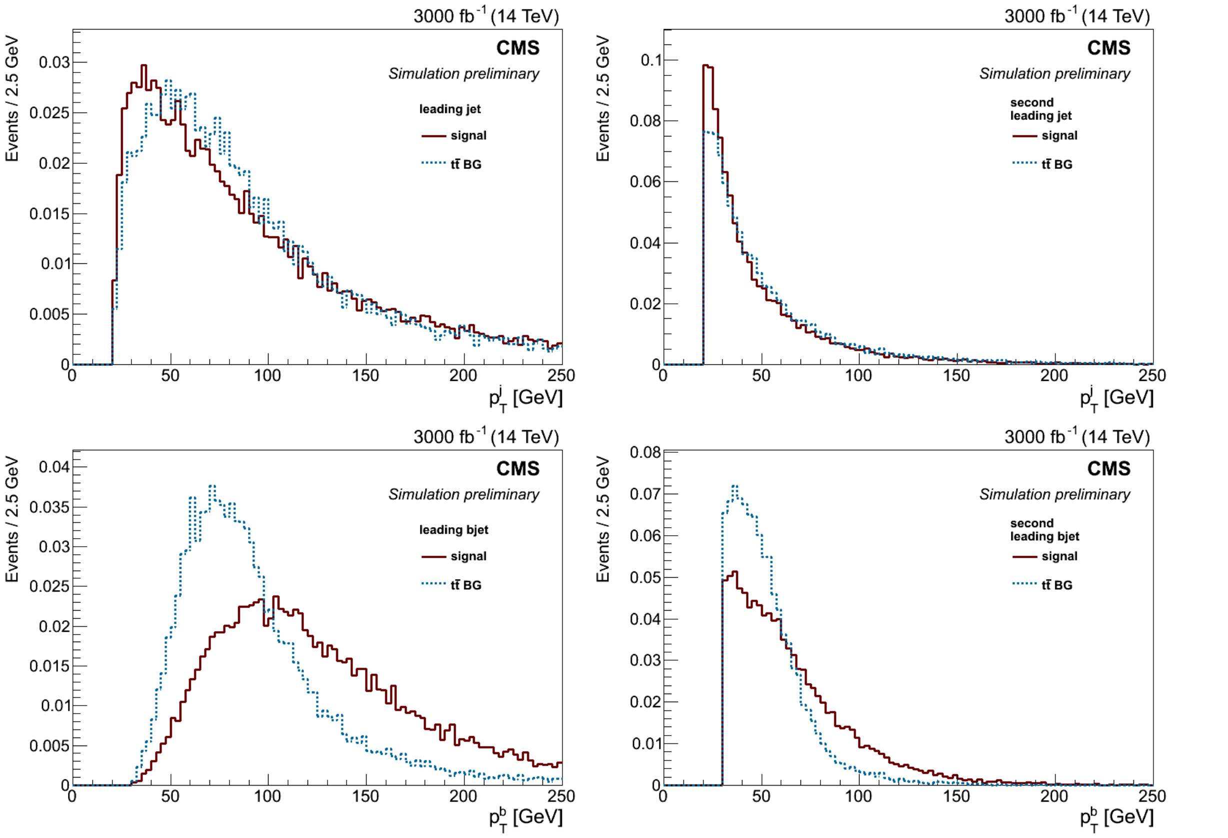

Figure 3: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: transverse momentum p_T for the two leading jets and two leading b-jets.

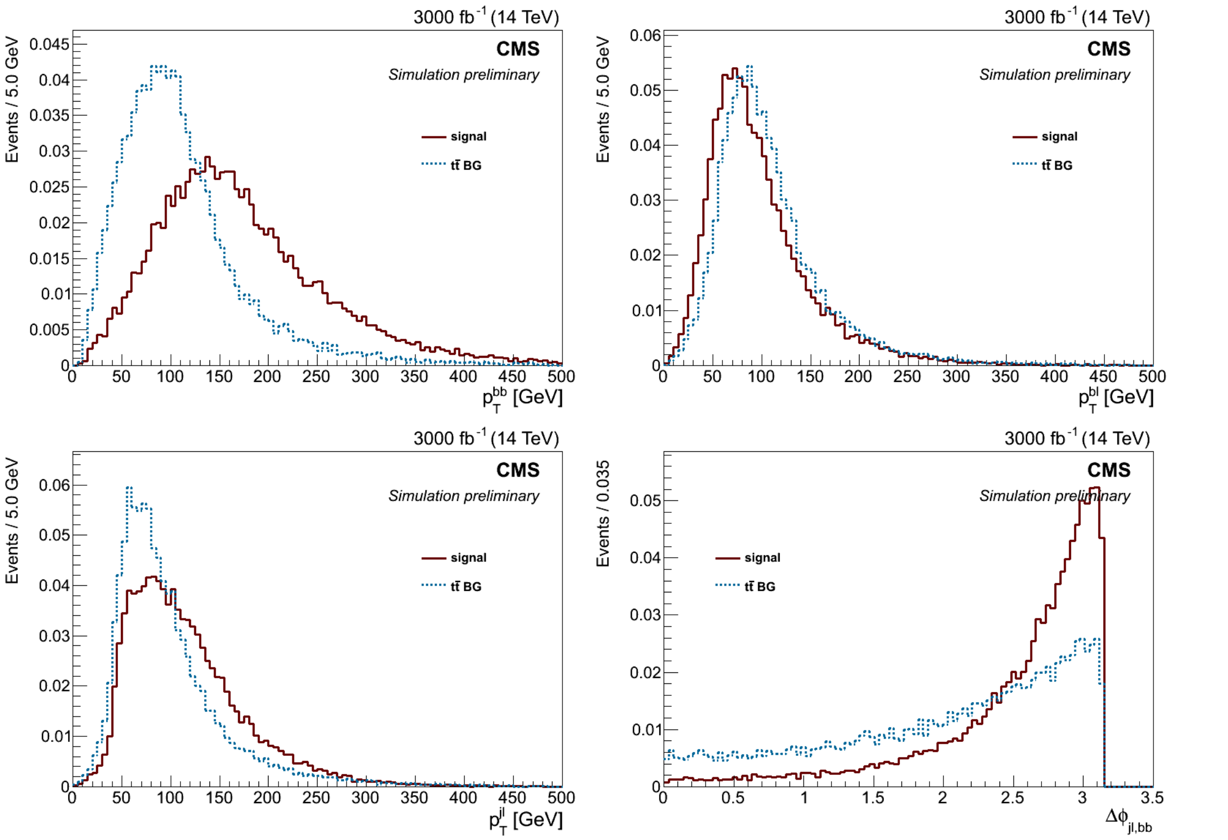

Figure 4: Variables distribution of HH (red) and $t\bar{t}$ (blue) for the neural network: $p_T^{\rm bb}$, p_T^{jj}, $p_T^{j_1\ell}$ and $\Delta\phi_{j_1\ell,\rm bb}$.

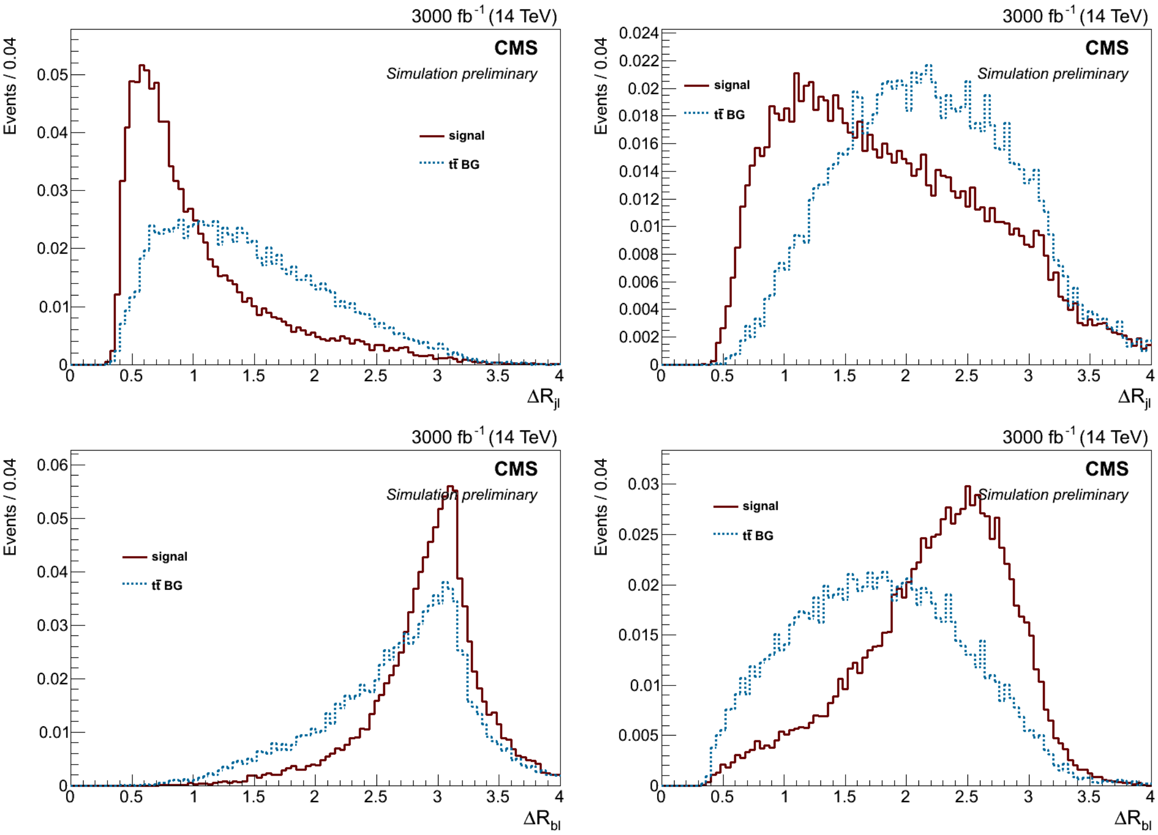

Figure 5: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: $\Delta R_{j_1\ell}$, $\Delta R_{j_2\ell}$, $\Delta R_{b_1\ell}$ and $\Delta R_{b_2\ell}$.

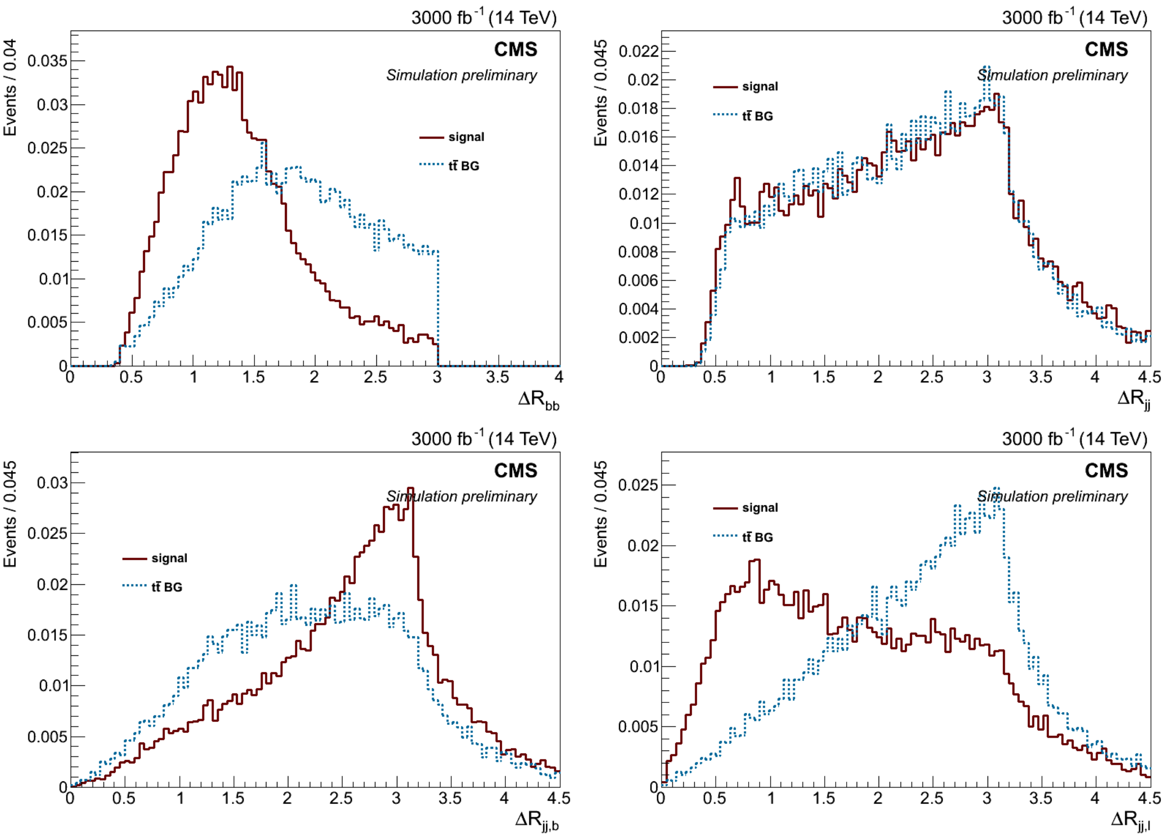

Figure 6: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: ΔR_{bb}, ΔR_{jj}, $\Delta R_{jj,b_1}$ and $\Delta R_{jj,\ell}$.

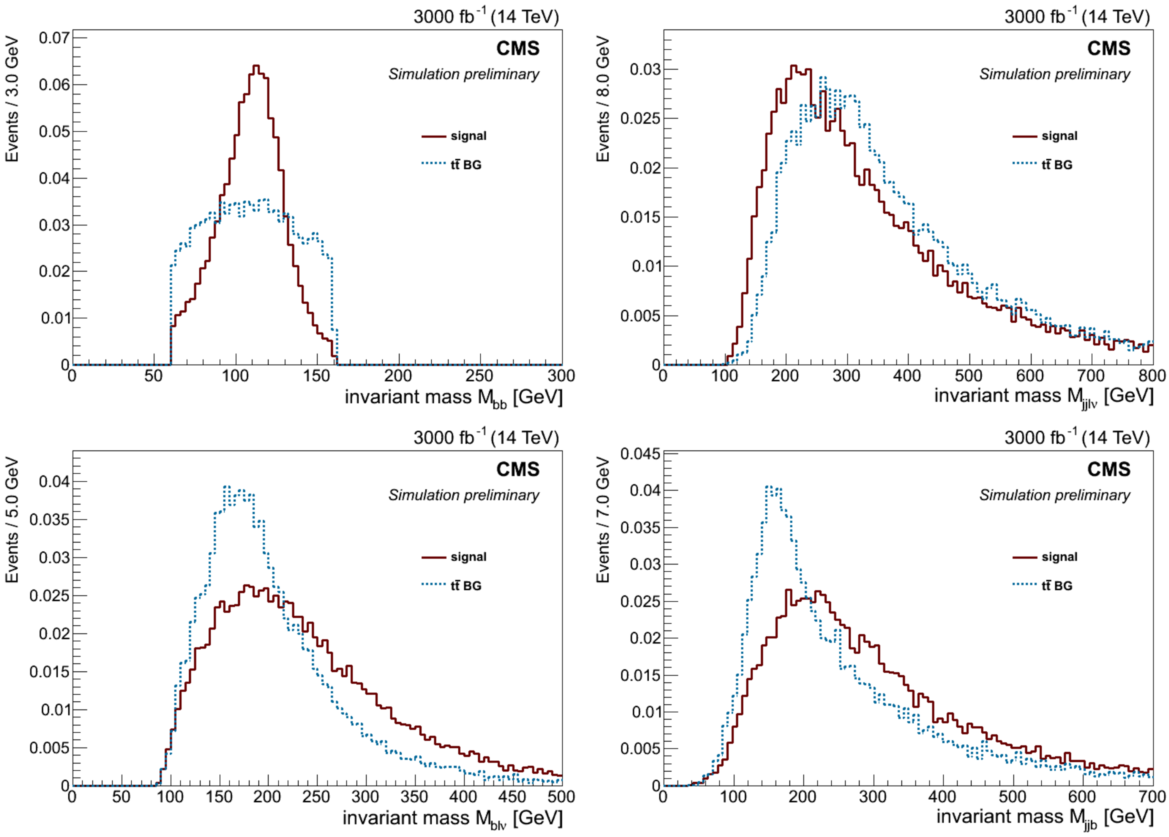

Figure 7: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: Higgs mass reconstructions M_{bb} and $M_{jj\ell\nu}$ and top mass reconstructions $M_{jj\mathrm{b}_1}$ and $M_{\mathrm{b}_2\ell\nu}$.

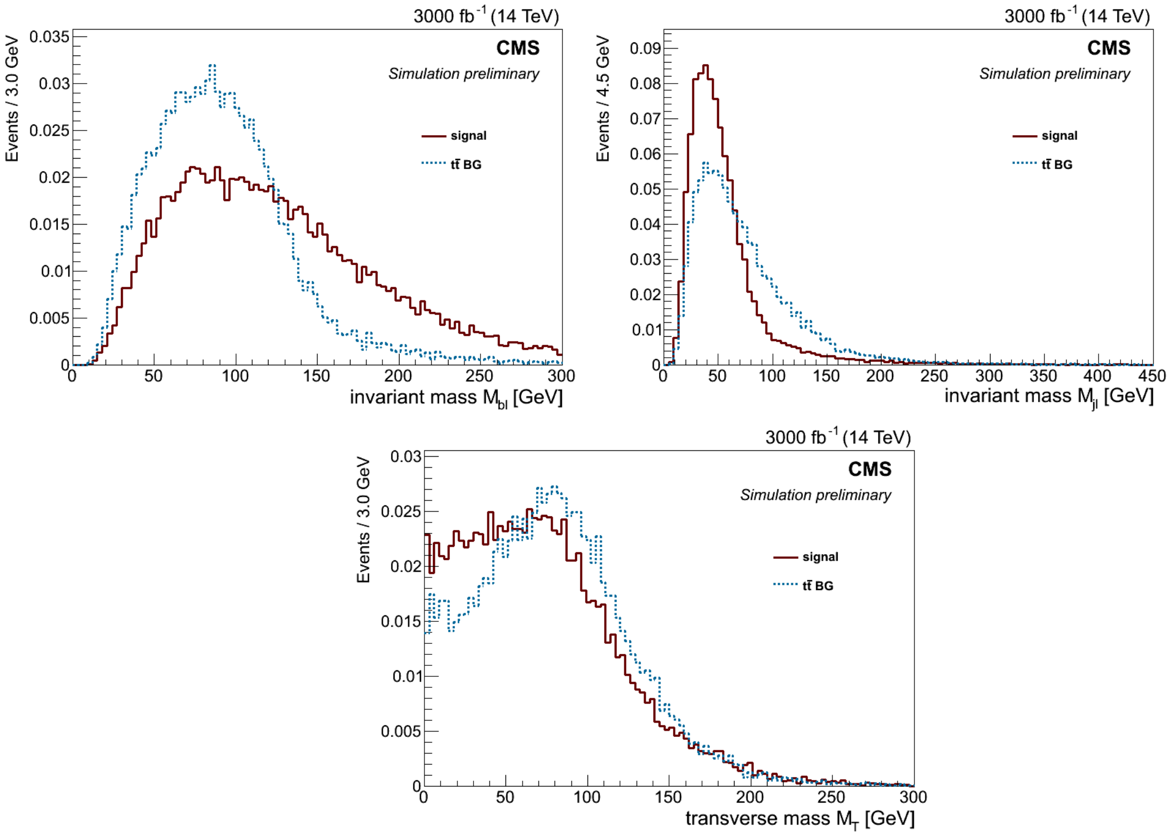

Figure 8: Variables distribution of HH (red) and $\mathrm{t\bar{t}}$ (blue) for the neural network: $M_{\mathrm{b}_2\ell}$ and $M_T^{\ell\nu}$ (see Eq. (2)).

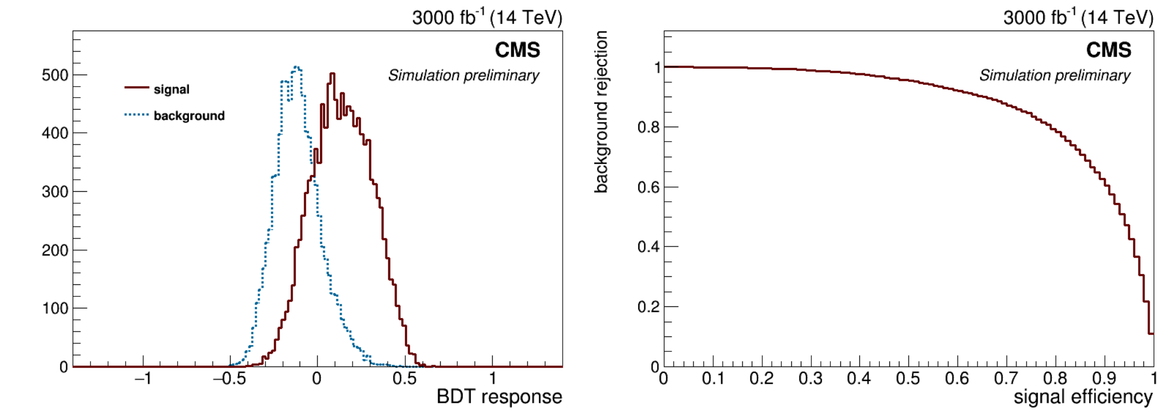

Figure 9: Final BDT output and background rejection versus signal efficiency.