IL NUOVO CIMENTO **38 C** (2015) 56
DOI 10.1393/ncc/i2015-15056-7

Colloquia: IFAE 2014

Multivariate analysis in the search for the Higgs boson produced in association with $W \to \ell\nu(\ell = e, \mu)$ in the decay channel $WW^{(*)} \to \ell\nu\ell\nu(\ell = e, \mu)$

Daniele Puddu
INFN, Sezione di Roma Tre and Università degli Studi Roma Tre - Rome, Italy

received 7 January 2015

Summary. — The Higgs boson production in the WH and ZH associated modes, collectively considered as VH associated production, provides important information on the Higgs boson couplings to gauge bosons. In particular, the WH associated production mode with the Higgs boson decaying into two W bosons, gives a direct measurement of the Higgs to W boson coupling constant. A Multivariate Analysis (MVA) based on Boosted Decision Trees (BDT) has been used to enhance the sensitivity and reject the main $WZ/W\gamma^*$ and $ZZ^{(*)}$ backgrounds in the data collected at 8 TeV centre-of-mass energy with the ATLAS detector at LHC.

PACS 12.15.-y – Electroweak interactions.
PACS 12.60.Fr – Extensions of electroweak Higgs sector.

1. – Description of the channel

The signature studied in this analysis consists in three leptons with total charge ±1, eventually in presence of missing transverse momentum. The main backgrounds with three real isolated leptons are due to the diboson production $WZ/W\gamma^*$, as well as the $ZZ^{(*)}$ production with an undetected lepton. These backgrounds cannot be reduced by the application of tight lepton identification criteria and are characterised by the presence of at least one pair of Same-Flavour Opposite-Sign (SFOS) leptons.

At a significantly lower rate, but comparable to the signal, triboson production, in particular $WWW^{(*)}$, represents an irreducible background, while the associated production of $t\bar{t}$ pairs with vector bosons can be reduced through a top-veto based on the identification of no b-jets in the final state.

Final states with fewer than three prompt leptons and/or without real missing transverse momentum may contribute to the background due to instrumental effects. Fake leptons include both jets which have been misidentified as leptons and real non-isolated leptons from light flavour, beauty and charm decays. Background processes with two prompt leptons, such as WW, $Z+$jets, $t\bar{t}$ and Wt production, must be accompanied by a fake lepton to enter the selection, these can be reduced through isolation requirements on the three leptons. Final states with only one prompt lepton, such as W boson production or single top quarks produced through the s-channel or t-channel, would require two fake leptons and are strongly suppressed by isolation requirements.

The leptons in the event are classified by identifying ℓ_0 as the lepton with unique charge, ℓ_1 as the lepton closer in $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ to ℓ_0, and ℓ_2 as the remaining one.

2. – The analysis

A MVA analysis has been introduced in order to improve the significance with respect to a Cut-Based analysis [1] in this channel.

The training of the BDT has been performed using the WH ($m_H = 125\,\mathrm{Gev}$) as signal and the $WZ/W\gamma^*$ and ZZ^* as backgrounds.

The choice of the input discriminating variables has been optimised targeting the best separation between the signal and the background. The lepton related variables and two kind of Missing Energy Transverse have been chosen. These are the following:

$$p_{\mathrm{T}}^{\ell_0}, p_{\mathrm{T}}^{\ell_1}, p_{\mathrm{T}}^{\ell_2}, |\Sigma \mathbf{p}_{\mathrm{T}}^{\mathrm{lep}}|, m_{\ell_0\ell_1}, m_{\ell_0\ell_2}, \Delta R_{\ell_0\ell_1}, E_{\mathrm{T,STVF}}^{\mathrm{miss}}, E_{\mathrm{T,Track}}^{\mathrm{miss}}.$$

An interesting region where BDT could learn the differences between signal and the main backgrounds has been chosen: at least a SFOS lepton pair, at most one jet, no b-tagged jets and $E_{\mathrm{T,STVF}}^{\mathrm{miss}} > 15\,\mathrm{Gev}$ have been required. All the background samples are then used in the final classification. A cut-based selection has been performed to reject the other backgrounds:

Cut 1: 3 isolated leptons ($p_{\mathrm{T}} > 15\,\mathrm{Gev}$), trigger, total charge ± 1;

Cut 2: "jet-veto", at most one jet with $p_{\mathrm{T}} > 25\,\mathrm{Gev}$;

Cut 3: "top-veto", no b-tagged jets with $p_{\mathrm{T}} > 20\,\mathrm{Gev}$;

Cut 4: $E_{\mathrm{T,STVF}}^{\mathrm{miss}} > 30\,\mathrm{Gev}$ and $E_{\mathrm{T,TrackClj}}^{\mathrm{miss}} > 20\,\mathrm{Gev}$ ($E_{\mathrm{T,TrackClj}}^{\mathrm{miss}}$ has been preferred to $E_{\mathrm{T,Track}}^{\mathrm{miss}}$ used in the training);

Cut 5: $|m_{\ell\ell} - m_Z| > 25\,\mathrm{Gev}$, $m_{\ell\ell}^{min} > 12\,\mathrm{Gev}$, $m_{\ell\ell}^{max} < 200\,\mathrm{Gev}$;

Cut 6: $\Delta R_{\ell_0\ell_1} < 2.0$.

After the above selections, "sub-signal regions" have been defined using the BDT Output with the bounds $(-1.0, 0.0, 0.4, 0.6, 0.8, 0.9, 1.0)$. Each of these regions has been included into the global fit (sect. **3**).

3. – Results

Control Regions, defined in a phase space disjoint but close to the signal phase space, are used to normalise the prediction of some of the backgrounds to the yield estimated using data. These are: $WZ/W\gamma^*$, the dominant background; ZZ^*, the second background; Top; $Z + \mathrm{jets}$; $Z\gamma$. All the information will be extracted from the data with a global fit, including the Normalization Factors and the observed upper limit on the VH Cross Section. The fit procedure takes into account all the detector systematics, the correlations and statistical errors.

REFERENCES

[1] ATLAS Collaboration, *"Search for associated production of the Higgs boson in the $WH \rightarrow WWW^{(*)} \rightarrow \ell\nu\ell\nu\ell\nu$ and $ZH \rightarrow ZWW^{(*)} \rightarrow \ell\ell\ell\nu\ell\nu$ channels with the ATLAS detector at the LHC"*. ATLAS-CONF-2013-075, 2014. http://cds.cern.ch/record/1562933.