PROCEEDINGS OF SCIENCE

Higgs $\rightarrow$ bb

Elisa GUIDO
INFN Genova
E-mail: `elisa.guido@cern.ch`

Andrea RIZZI
Università degli Studi di Pisa and INFN Pisa
E-mail: `andrea.rizzi@cern.ch`

We present the current results obtained by ATLAS and CMS experiments at the LHC in the search for the Higgs to b-quarks decay mode, in the case of a production of the Higgs associated both to a vector boson and to a t-quarks pair. We also enlighten the more critical items expected for the extension of the analyses to the higher energy and luminosity conditions foreseen for the future phases of the LHC data taking.

VI Italian workshop on p-p physics at the LHC
May 8-10, 2013
Acquario di Genova, Ponte Spinola, Area Porto Antico, Genova, Italy

1. Current results

At a mass of ~ 125 GeV, the standard model (SM) Higgs boson (H) is expected to decay predominantly into a bottom-antibottom quark pair. The observation of the Higgs to b-quarks decay (Hbb) is thus crucial in order to provide constraints on the nature of the newly discovered boson [1, 2]. The Hbb is currently analyzed by both the ATLAS and the CMS experiments in two different production channels, in association either with a weak vector boson (V) with V=Z,W$^{\pm}$ (VH), or with a t-quarks pair (ttH). At present, the data collected at the center-of-mass energy $\sqrt{s}$=8 TeV are only partially analyzed, and the complete results are expected soon.

1.1 Search for a Higgs boson produced in association with a vector boson and decaying to bottom quarks (VH)

For the current most up to date results [3, 4], both the experiments analyze a sample consisting of ~ 5 fb^{-1} of data at $\sqrt{s}$=7 TeV, and $\sim 12-13$ fb^{-1} of data at $\sqrt{s}$=8 TeV. The decay channels analyzed are $ZH \rightarrow \nu\bar{\nu}b\bar{b}$, $WH \rightarrow \ell\bar{\nu}b\bar{b}$ and $ZH \rightarrow \ell^{+}\ell^{-}b\bar{b}$, characterized by final state topologies with zero, one or two leptons, besides the two b-jets from the H decay. Different triggers and event selection are applied for the three possible final states. In order to maximize the sensitivity, events are further divided in different categories, depending on the transverse momentum of the vector boson $p_T(V)$. The main backgrounds for this search are $Z+b$-jets, $W+b$-jets and top events; they are modeled with simulated samples and their shape and yield are then checked on data *control regions*, properly defined in order to be enriched in one of the background topologies, while depleted in signal events. The CMS experiment applies a multivariate analysis (MVA), while the ATLAS experiment performs a cut-based analysis. In order to extract the signal strength for a given mass hypothesis, both the experiments perform a fit from simulation to data of the MVA distribution in the CMS analysis, and of the di-b-jet invariant mass in the ATLAS analysis, respectively. The systematic uncertainties on the signal strength, mainly related to the b-tagging efficiency and to the limited amount of the simulated events used to mimic the backgrounds, are treated as nuisance parameters in the fit. Upper limits (ULs) at the 95% of confidence level (CL) are set on the production cross section normalized to the SM expectation. The observed and expected ULs for a Higgs mass of 125 GeV are 2.45 and 1.15 times the SM prediction for the CMS analysis [3], and 1.8 and 1.9 for the ATLAS analysis [4] , respectively. This search is therefore quite near to the SM sensitivity, with an excess of signal events corresponding to a local significance of $\sim 2\sigma$ for the CMS analysis.

1.2 Search for a Higgs boson produced in association with a t-quarks pair and decaying to bottom quarks (ttH)

The data analyzed by the two experiments in the current most up to date results are different: the CMS experiment uses ~ 5 fb^{-1} of data at $\sqrt{s}$=7 TeV, and ~ 5 fb^{-1} of data at $\sqrt{s}$=8 TeV [5], while the ATLAS experiment analyzes only ~ 5 fb^{-1} of data at $\sqrt{s}$=7 TeV [6]. This search considers final states where, besides the Hbb decay, the two top-quarks decay into $W+b$ and the W bosons decay either both leptonically, or one to $\ell^{+}\nu, \ell^{-}\overline{\nu}$ and the other one to $q\bar{q}$. The events are thus characterized by one or two high-p_T leptons, large missing transverse energy, and high jet and b-jet multiplicity. Jet multiplicities are used in order to subdivide the events into different categories and variable signal sensitivity. The dominant background in the regions with the highest signal sensitivity is due to $t\bar{t}$+jets events, and it is mainly irreducible. All the background contributions are simulated and checked by using *control regions* on data. In order to discriminate between signal and background, the CMS analysis is based on an MVA approach, while the ATLAS analysis on a cut-based one. The signal strength for a given mass hypothesis is extracted from a fit from simulation to data, performed simultaneously to all the categories, of the MVA output distribution in the CMS analysis, while in the ATLAS analysis of the di-b-jet invariant mass distribution, after kinematically reconstructing the $t\bar{t}$ system, in the two highest signal sensitivity categories, and to the scalar sum of p_T(jet) for all the others. The systematic uncertainties on the signal strength are treated as nuisance parameters in the fit. No statistically significant signal is found over the background expectations, and 95% of CL ULs are set on the production cross section normalized to the SM expectation. The observed and expected ULs for a Higgs mass of 125 GeV are 5.8 and 5.2 times the SM prediction for the CMS analysis [5], and 13.1 and 10.5 for the ATLAS analysis [6], respectively.

2. Critical items for higher energy and luminosity

2.1 Theoretical predictions

The measurement of the Hbb yield, in all its different production channels, is still quite far from the precision measurements regime. The current theoretical predictions have instead already reached this precision for the VH and ttH production mechanisms, including NLO or NNLO calculations not only for the inclusive cross section but also for some of the most important differntial variables [7, 8] [9, 10] [11] [12] [13] [14, 15, 16] [17, 18] [19, 20]. For this reason the theoretical uncertainties on the signal cross sections are not the limiting factor for the Hbb studies, as long as the calculations performed at $7-8$ TeV will be performed also for the LHC high energy expected in 2015 of ~ 13 TeV.

On the other hand the experimental measurement is often limited by the large uncertainties on the background predictions. In the various ATLAS and CMS Hbb analyses, normalization corrections or differential corrections are applied to the SM processes that are predicted from simulation. Normalization corrections are applied as simple scale factors (SF) to the background yields, while differential corrections are typically implemented with event weighting.

This kind of corrections often are affected by large systematic uncertainties that, in prospects, will be a limiting factor in the analysis sensitivity. Both the SF and the weight functions are typically computed in *control regions* of the phase space that are expected to be signal depleted. The extrapolation of the measured SF or weight function to the signal enriched phase space regions is done assuming that the same correction be valid in both regions. This procedure is affected by sizable uncertainties, as there is no direct way to check whether the simulation is mispredicting the background component in the same way in the two regions. In some cases there is not even the possibility of defining signal free control regions and pure simulation predictions are used with huge systematic uncertainties.

The event weighting techniques are typically used for p_T spectrum correction of $W/Z+jet$ or $t\bar{t}+jets$ backgrounds, while significant SF are observed for example in the case of $W/W+1\,b-jet$ process. In the case of the $t\bar{t}b\bar{b}$ irreducible ttH background there is no clean control region to compute a correction in ATLAS and CMS analyses, and an uncertainty of 50% is ascribed to the

simulation prediction of these background.
Some of the largest deviations observed between simulation prediction and measurements on data are for processes dominated by, or involving in large part, the gluon to $b\bar{b}$ splitting process. At the parton level, gluon splitting (GSP) is typically defined as the $b\bar{b}$ production where one of the two b is soft or where the two b are collinear. This phase space is difficult to calculate in perturbative QCD and it is handled in LHC simulations by the so called parton shower generators. Typically only one of the two b-quarks is detected in the experiment as a b-jet, either because the second is too soft or because the hadronization of the two b-quarks largely overlaps producing a single jet. Deviations up to a factor of two for this processes are measured by both ATLAS and CMS [21, 22, 23].
The large mispredictions in GSP can be a limiting factor for example in the ttH analysis, where the number of jets and number of b-tag counting is used to define the different regions and a global fit of the yields is used, as explained in one of the previous sections. Indeed, for example, the ratio of the “4jets - 3 tag” to the “6 jets - 4 tags” of the $tt+b(b)$ background yields can be mispredicted by the simulation, because of underestimated GSP, and the global fit could then be wrongly constrained.

2.2 Limitations from systematic uncertainties

The current ATLAS and CMS analyses for ttH and VH productions, with Hbb decay, are heavily relying on simulation to predict the SM processes such as di-boson production, $t\bar{t}$, $W/Z+jets$, single top, etc. In order to have a precise prediction, a huge number of simulated event was needed for 2012 analyses (as an example, the CMS VH analysis alone processed about half a billion data events and used a similar amount of Monte Carlo generated simulation events). Despite the large number of MC events available, the systematic uncertainties arising from statistical uncertainties in the simulated samples are among the largest ones. Those uncertainties are going to become a limiting factor when, with higher luminosity, the statistical uncertainty of the data will shrink. There are a few possibility that the experiments will need to investigate in the next years to overcome this problem. One option is to further increase the statistical power of the simulation, either with more events or with better phase space slicing, but this will probably require more computing power, in particular when considering that higher energy and higher pile-up events take longer to be fully simulated. It should also be noted that not all generators allow this kind of slicing. Another possibility would come from a completely data-driven background estimate, similar to what is done in the Higgs to diphoton searches [24, 25]. In this case the main problem come from the much worse mass resolution, compared to the diphoton analysis, possibly leading to a large systematic uncertainty depending on the functional shape chosen for the background fit. This technique is usable with the large data statistics that will be available in the mid-low boost regions (*i.e.* ∼50-100 GeV) of the VH analysis.

2.3 Signal and background cross section scaling

An important factor to take into account is the scaling of the production cross section with the LHC $\sqrt{s}$. In particular, the $t\bar{t}$ background increases with $\sqrt{s}$ more than the VH signal cross section, while it increases less than the ttH one. The higher energy regime will then change the S/B of the different Hbb modes, possibly changing also the relative importance, in term of sensitivity, of the various channels. For example, the $WH \rightarrow l\nu bb$ and the $ZH \rightarrow \nu\nu bb$ have a large $t\bar{t}$ background contribution, while the $ZH \rightarrow llbb$ analysis is less contaminated by this background. Currently the Zll mode, due to lower cross section, is less sensitive than the VH channels, but the reduced S/B in the other two modes can change the picture. Similarly the ttH mode, penalized by the very low cross section at $7-8$ TeV, is expected to have a performance boost at $\sim$ 13 TeV.

2.4 Event selection and reconstruction

The higher number of collision per beam crossing (pile-up) foreseen for the future LHC runs will be another challenge for the Hbb search, as for most data analyses at the LHC. The higher pile-up, but also the higher center-of-mass energy, is expected to increase the rates of the triggers currently used for the Higgs to b-quarks search. Some of those triggers could become non sustainable and the raise of their thresholds could be needed. We expect the most critical triggers to be those of $ZH \rightarrow \nu\nu bb$ where the missing energy signature of the Z is exploited.

In term of offline reconstruction we can identify as critical items the b-tagging and jet energy resolution in the new environment, the removal of pile-up jets, the missing energy resolution and the lepton isolation used to keep multijet backgrounds under control.

An additional topic that should be reconsidered for the higher energy running is the one related to jet substructures [26]. The jet substructure techniques are currently being studied in the two experiments but so far no striking improvement using such techniques was reported by ATLAS or CMS. The substructures method was considered to be potentially helpful producing a better mass resolution and being able to distinguish the two b-jets at very high Higgs boost. Different substructures algorithms have been tested in new physics searches involving jets with energy of several hundreds of GeV, but the currently explored Hbb phase space was not yet entering such high p_T regime. With higher $\sqrt{s}$ and higher integrated luminosity we expect to start populating the region above $\sim$ 400 GeV of Higgs boost where jet-merging could become an issue for a simple di-jet analysis. A key role could then be played by the substructure technique and is being studied by the two collaborations.

3. Conclusions

We presented the current state of CMS and ATLAS analyses searching for Higgs decay to b-quarks. While the final results on 2012 data taking are expected in the next months, we already started looking ahead for the next LHC running. Although the current precision of the measurement of the Hbb production is largely dominated by the statistical uncertainties, the major source of systematic limitations have been presented. Critical items, on which the experiments and the theorists should focus they efforts towards precision measurements, have been reviewed.

References

[1] [ATLAS Collaboration]. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,. Phys. Lett. B **716** (2012) 1 [arXiv:1207.7214 [hep-ex]].

[2] [CMS Collaboration]. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys. Lett. B **716** (2012) 30 [arXiv:1207.7235 [hep-ex]].

[3] [CMS Collaboration] Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks for HCP 2012. CMS-PAS-HIG-12-044.

[4] [ATLAS Collaboration] Coupling properties of the new Higgs-like boson observed with the ATLAS detector at the LHC. ATLAS-CONF-2012-127.

[5] [CMS Collaboration] Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC. arXiv:1303.0763 [hep-ex].

[6] [ATLAS Collaboration] Search for the Standard Model Higgs boson produced in association with top quarks in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. ATLAS-CONF-2012-135.

[7] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck. EW corrections to Higgs strahlung at the Tevatron and the LHC with HAWK. 2011.

[8] G. Ferrera, M. Grazzini, and F. Tramontano. Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO. *Phys.Rev.Lett.*, 107:152003, 2011.

[9] S.L. Glashow, Dimitri V. Nanopoulos, and A. Yildiz. Associated Production of Higgs Bosons and Z Particles. *Phys.Rev.*, D18:1724–1727, 1978.

[10] Z. Kunszt, Z. Trocsanyi, and W. James Stirling. Clear signal of intermediate mass Higgs boson production at LHC and SSC. *Phys.Lett.*, B271:247–255, 1991.

[11] T. Han and S. Willenbrock. QCD correction to the $pp \to WH$ and ZH total cross-sections. *Phys. Lett.*, B273:167–172, 1991.

[12] O. Brein, A. Djouadi, and R. Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579:149–156, 2004.

[13] M. L. Ciccolini, S. Dittmaier, and M. Krämer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev.*, D68:073003, 2003.

[14] R. Raitio and W. W. Wada. Higgs boson production at large transverse momentum in QCD. *Phys.Rev.*, D19:941, 1979.

[15] J. N. Ng and P. Zakarauskas. A QCD parton calculation of conjoined production of Higgs bosons and heavy flavors in p anti-p collision. *Phys.Rev.*, D29:876, 1984.

[16] Z. Kunszt. Associated Production of Heavy Higgs Boson with Top Quarks. *Nucl.Phys.*, B247:339, 1984.

[17] W. Beenakker et al. Higgs radiation off top quarks at the Tevatron and the LHC. *Phys. Rev. Lett.*, 87:201805, 2001.

[18] W. Beenakker et al. NLO QCD corrections to ttbar H production in hadron collisions. *Nucl. Phys.*, B653:151–203, 2003.

[19] S. Dawson, L. H. Orr, L. Reina, and D. Wackeroth. Associated top quark Higgs boson production at the LHC. *Phys. Rev.*, D67:071503, 2003.

[20] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth. Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections. *Phys. Rev.*, D68:034022, 2003.

[21] [ATLAS Collaboration] Measurement of the cross-section for W boson production in association with b-jets in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector. arXiv:1302.2929 [hep-ex].

[22] [CMS Collaboration] Measurement of $B\overline{B}$ Angular Correlations based on Secondary Vertex Reconstruction at $\sqrt{s} = 7$ TeV. JHEP **1103** (2011) 136. [arXiv:1102.3194 [hep-ex]].

[23] [CMS Collaboration] Public Analysis Summary EWK-11-015, Measurement of B anti-B angular correlations in the process pp->Zbb (paper in preparation), 2012.

[24] [ATLAS Collaboration]. Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data. ATLAS-CONF-2013-012.

[25] [CMS Collaboration]. Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel. CMS-PAS-HIG-13-001.

[26] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam Jet substructure as a new Higgs search channel at the LHC. arXiv:0810.0409 [hep-ph].