PROCEEDINGS OF SCIENCE

Higgs boson search and properties measurement in the $H \to \gamma\gamma$ decay channel

Paolo Meridiani
INFN Sezione di Roma
E-mail: paolo.meridiani@roma1.infn.it

Ruggero Turra
Università degli Studi di Milano & INFN
E-mail: ruggero.turra@mi.infn.it

The observation of decays into photon pairs played a leading role in the discovery of the new Higgs-like boson with mass around 125 GeV announced by the ATLAS and CMS collaborations at the LHC. The main characteristics of the data analyses are presented and compared between the two experiments. The evolution of the analyses and future prospects are discussed.

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

PoS(LHCPP2013)002

The Higgs mechanism is one of the best-motivated explanation of the electroweak symmetry breaking. In the standard model (SM), this mechanism in its simplest form predicts the existence of only one elementary scalar neutral particle, the Higgs boson.

The decay into two photons ($H \to \gamma\gamma$) has always been considered as a "discovery channel" for a low-mass Higgs boson at the Large Hadron Collider (LHC). The SM $H \to \gamma\gamma$ decay is predicted to have a very small branching ratio (2.2×10^{-3} for $m_H = 120$ GeV). The search of the Higgs boson in this channel is characterised by a small signal-over-background ratio, a very simple signature (two isolated photons with relatively high momentum transverse to the beam axis) and a sensitivity which is mostly driven by the experimental resolution of the invariant mass of the diphoton system. The main backgrounds are represented by an irreducible component coming from the production of two isolated prompt photons, and a reducible one from QCD processes, where at least one jet is misreconstructed as an isolated photon.

The analysis strategy adopted by the ATLAS and CMS collaborations is very similar, a narrow peak is searched in the diphoton mass spectrum. In addition, diphoton events are categorized according to the expected invariant mass resolution and signal likelihood to improve the statistical sensitivity of the search and to add the capability to tag specific Higgs boson production modes (e.g. vector boson fusion Higgs production (VBF) or production associated with a vector boson (VH)).

CMS presents two analyses [1], one cut-based (based on simple rectangular cuts) and a second one based on a multivariate (MVA) discriminator, considered as the main one given the better expected sensitivity. A cut-based analysis is presented by ATLAS [2], and in addition a first study of spin [3]. The whole 7 TeV and 8 TeV datasets have been used in the latest updates of these analyses: $4.8 + 20.7$ fb^{-1} for ATLAS [2] and $5.1 + 19.6$ fb^{-1} for CMS [1].

The performance of the two electromagnetic calorimeters (ECAL) is somewhat different, following the different choices taken during the design stage of the two experiments. CMS achieves a better photon energy resolution thanks to its homogeneous ECAL [4],[5],[6] and a smaller amount of material in front of ECAL (in ATLAS the calorimeters are placed outside the coil of the inner solenoid). On the contrary, the ATLAS calorimeter [7] can provide a direction measurement thanks to its longitudinal segmentation and a better discrimination between γ and π^0 thanks to the fine segmentation of the first layer in the pseudorapidity η direction. The ECAL performance in terms of energy resolution and response stability versus time have been checked with electrons from $Z \to ee$ and $W \to e\nu$ [6], [8], [9] by both experiments.

The sensivity of the analysis is directly proportional to the achieved diphoton mass resolution. Apart for the photon energy resolution which is driven by the ECAL performance, it is fundamental to measure precisely the photon direction to keep the contribution to the diphoton mass resolution from the angle mesurement small. It is then required to know precisely the longitudinal position of the Higgs production vertex (along the beam-line). The identification of the primary vertex becomes increasingly difficult in presence of a large number of interaction vertices for beam crossing (pileup). For this purpose the the tracks recoiling against the diphoton system can be used; in addition, ATLAS can take advantage of the measurement of the electromagnetic shower direction and of the direction of the conversion tracks in case of a converted photon. The resolution on the longitudinal position of the primary vertex is 15 mm from the calorimetric pointing and 6 mm when the conversion direction is used. CMS can do something similar only using the converted photons. The final choice of the vertex position is obtained from a multivariate algorithm which combines all the information available. The overall invariant mass resolution of the diphoton system, quantified by the FWHM/2.35, is 1.77 GeV for ATLAS and 1.64 GeV for CMS for an Higgs boson with mass $m_H = 125$ GeV, reaching a resolution of about 1% when both photons impinge in the central part of the detector.

Some differences are present for the kinematic cut applied at the photon selection stage: ATLAS uses a fixed cut for the two photons $p_T > 30, 40$ GeV while CMS uses a cut scaling with the invariant mass : $p_T > m_{\gamma\gamma}/3$, $p_T > m_{\gamma\gamma}/4$. Photon identification requirements are based on shower shape in the calorimeter (both lateral and longitudinal shape) and isolation requirements enforced both in the inner detector and in the calorimetric compartment. ATLAS uses a multivariate photon identification algorithm for the 7 TeV dataset and a cut-based approach in the 8 TeV dataset, CMS uses both approaches for the cut-based and multivariate analysis respectively. The performance of the photon identification algorithms have been compared between data and simulations using control samples from $Z \to ee$, where electrons are reconstruced as if they were photons , and $Z \to ll\gamma$ decays, showing an overall agreement at the percent level [10],[11],[12]. After the diphoton selection, the purity of events with two prompt isolated photons in the invariant mass range between 100 and 180 GeV is 75^{+3}_{-4}% for ATLAS and around 70% for CMS .

Both experiments divide the diphoton events in exclusive categories to tag a specific Higgs boson production mode (VBF or VH) requiring additional jets, large missing transverse energy or the presence of additional leptons; the remaining events are divided in inclusive categories with different signal to background ratio and diphoton mass resolution. ATLAS analysis has 9 inclusive categories based on the presence of a converted photon, the position in the calorimeter and transverse momentum of the diphoton system. CMS has 4 inclusive categories for the main analysis, defined using the output of a multivariate discriminator (BDT). The multivariate discriminator takes as input several variables, as the photon identification score of the photons, the expected diphoton mass resolution, the kinematic properties of the two photons and the probability to have correctly identified the right vertex. The output of the BDT is validated by comparing the distribution of the output variable in data and MC simulation with $Z \to ee$ events where electrons are reconstructed as photons. The two analyses reach similar $S/\sqrt{B}$ ratio and expected p-value (defined as the probablity to have a value of the test statistics as extreme as the observed one for the background only hypothesis) in the inclusive sample and in the different categories.

In the analyses of both experiments, the background spectrum is estimated from data, fitting the observed diphoton mass distribution with a parametric function in each individual category, albeit the method to choose of the parametrization is different. ATLAS employs high statistics background MC samples to define possible "truth models" for the background. Several functions are tested against these background samples; the one which minimises the bias in estimating the background in the signal region is chosen. The estimated bias is also used to assign a systematic uncertainty on the background estimation. CMS instead derives the truth models from the data, identifying different classes of functions able to give a good fit of the data. Pseudo-data are generated from these truth models and fitted with different test functions. Functions with increasing number of degrees of freedom are exploited: the chosen parametrization is the one able to give an average bias in the estimated number of background events in the signal region smaller then 20% of the statistical uncertainty. This allows the systematic uncertainty on the background estimation to be neglected, effectively including it as a statistical uncertainty.

While the expected significances are quite similar between the two experiments, the observed ones are quite different: ATLAS quotes 7.4 (4.1 expected), CMS 3.2 (4.2 expected) using the MVA analysis and 3.9 (3.5 expected) for the cut-based. This translates into different best-fit signal strengths for the new boson $\hat{\mu} = \frac{\sigma}{\sigma_{SM}}$: $1.65 \pm 0.24(\text{stat})^{+0.25}_{-0.18}(\text{sys})$ for ATLAS and $0.78^{+0.28}_{-0.26}$ for CMS using the MVA analysis and $1.11^{+0.32}_{-0.30}$ using the cut-based. The measured mass is $126.8 \pm 0.2(stat) \pm 0.7(syst)$ GeV for ATLAS and $125.4 \pm 0.5(stat) \pm 0.6(syst)$ GeV for CMS. The uncertainty on the mass measurement is dominated by the systematic uncertainties, coming mostly from the precision of the knowledge of the photon energy calibration.

ATLAS also quotes the signal strength for specific production modes: $\mu_{ggH+t\bar{t}H} = 1.6^{+0.3}_{-0.3}(\text{stat})^{+0.3}_{-0.2}(\text{syst}), \mu_{VBF} = 1.7^{+0.8}_{-0.8}(\text{stat})^{+0.5}_{-0.4}(\text{syst}), \mu_{VH} = 1.8^{+1.5}_{-1.3}(\text{stat})^{+0.3}_{-0.3}(\text{syst})$.

A review of the evolution of the ATLAS and CMS analyses has also been presented. The expected significance (defined with the gaussian one-sided convention for the p-value) follows the expected $\sqrt{L}$ behaviour, where L is the integrated luminosity corresponding to the dataset analyzed. For the cut-based analyses the sensitivity grows a little faster, reflecting the improvements in the analyses. Similarly, the uncertainty on $\hat{\mu}$ improves as $1/\sqrt{L}$, showing that the nature of this error is still mostly statistical, and is anticipated to improve when more data become available.

At variance with this, the uncertainty on the mass is already limited by the systematic uncertainty related to the calibration of the photon energy which is derived from electrons coming from the Z boson decay: data/MC energy calibration factors are computed for different categories of electrons to set the Z boson mass peak at the same position predicted by the MC simulation. The error of this procedure comes from the extrapolation from electrons to photons and from the Z boson to the $H(125)$ boson mass scale. These extrapolations rely on MC simulations, in particular the Geant4 [13] simulation of the showers and the implementation of the material descriptions. Reduction of the extrapolation errors is then connected to the improvements of the simulation and of the description of the material in front of the detectors which both experiments are pursuing. The use of photons from the $Z \to \mu\mu\gamma$ decay is currently used as a cross-check in both experiments, somehow limited by the even larger scale extrapolation and by the small amount of data.

Among the various consistency check of the analyses, the compatibility of the two CMS analyses (cut-based and MVA) and of the observed diphoton mass resolution in ATLAS have been discussed. In the fit to the diphoton mass for signal extraction, the estimated energy resolution is pulled at about 2σ away from its nominal value. Dedicated studies revealed no indication that the systematic uncertainty on the resolution is underestimated; the large pull can then be due to a statistical effect arising from background fluctuations.

The compatibility of the two CMS analyses (cut based and MVA) was also studied. The two analyses share about 50% of the selected events. To estimate their compatibility the correlation among the two analyses has to be estimated. This was performed estimating the variance of the difference between the $\hat{\mu}$ among the two analyses using the “jackknife delete-d resampling” [14]. Considering the whole dataset the two analyses are estimated to be compatible within 1.5σ.

References

[1] CMS Collaboration. Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel. Technical Report CMS-PAS-HIG-13-001, CERN, Geneva, Mar 2013.

[2] 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. Technical Report ATLAS-CONF-2013-012, CERN, Geneva, Mar 2013.

[3] ATLAS Collaboration. Study of the spin of the Higgs-like boson in the two photon decay channel using 20.7 fb^{-1} of pp collisions collected at $\sqrt{s} = 8$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2013-029, CERN, Geneva, Mar 2013.

[4] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[5] CMS Collaboration. CMS ECAL Performance on 2011 data. Technical Report CMS-DP-2012-002, CERN, Geneva, Feb 2012.

[6] CMS Collaboration. 2012 ECAL detector performance plots. Technical Report CMS-DP-2013-007, CERN, Geneva, Mar 2013.

[7] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[8] ATLAS Collaboration. Electron energy response stability with pile-up in 2012 data up to HCP-dataset. Technical Report ATL-COM-PHYS-2012-1668, CERN, Geneva, Nov 2012.

[9] ATLAS Collaboration. Electron energy response stability with time in 2012 data up to HCP-dataset. Technical Report ATL-COM-PHYS-2012-1593, CERN, Geneva, Nov 2012.

[10] CMS Collaboration. Photon ID performance with 19.6 fb^{-1} of data collected at $\sqrt{s} = 8$ TeV with the CMS detector. Technical Report CMS-DP-2013-010, CERN, Geneva, Apr 2013.

[11] ATLAS Collaboration. Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{1} of pp collision data collected in 2011. Technical Report ATLAS-CONF-2012-123, CERN, Geneva, Aug 2012.

[12] ATLAS Collaboration. Photon Efficiency Measurement with Radiative Photons in Z events ($Z \to ll\gamma$). Technical Report ATL-COM-PHYS-2013-244, CERN, Geneva, Mar 2013.

[13] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois, M. Asai, G. Barrand, R. Capra, S. Chauvie, R. Chytracek, G. A P Cirrone, G. Cooperman, G. Cosmo, G. Cuttone, G. G. Daquino, M. Donszelmann, M. Dressel, G. Folger, F. Foppiano, J. Generowicz, V. Grichine, S. Guatelli, P. Gumplinger, A. Heikkinen, I. Hrivnacova, A. Howard, S. Incerti, V. Ivanchenko, T. Johnson, F. Jones, T. Koi, R. Kokoulin, M. Kossov, H. Kurashige, V. Lara, S. Larsson, F. Lei, O. Link, F. Longo, M. Maire, A. Mantero, B. Mascialino, I. McLaren, P.M. Lorenzo, K. Minamimoto, K. Murakami, P. Nieminen, L. Pandola, S. Parlati, L. Peralta, J. Perl, A. Pfeiffer, M.G. Pia, A. Ribon, P. Rodrigues, G. Russo, S. Sadilov, G. Santin, T. Sasaki, D. Smith, N. Starkov, S. Tanaka, E. Tcherniaev, B. Tome, A. Trindade, P. Truscott, L. Urban, M. Verderi, A. Walkden, J. P. Wellisch, D.C. Williams, D. Wright, and H. Yoshida. Geant4 developments and applications. *Nuclear Science, IEEE Transactions on*, 53(1):270–278, 2006.

[14] M. H. Quenouille. Approximate tests of correlation in time-series 3. *Mathematical Proceedings of the Cambridge Philosophical Society*, 45:483–484, 6 1949.