PROCEEDINGS OF SCIENCE

Background Modelling in the ATLAS $H \to \gamma\gamma$ Differential Cross Section Analysis

N. Gillwald[a,*] **on behalf of the ATLAS Collaboration**

[a] *Deutsches Elektronen-Synchrotron DESY,*
Notkestrasse 85, Hamburg, Germany

E-mail: nils.gillwald@desy.de

Measurements of fiducial integrated and differential cross sections of the Higgs boson decaying into a pair of photons are presented with an emphasis on discussing the background modelling strategy. The measurements use the full ATLAS dataset corresponding to 139 fb^{-1} of proton-proton collision data recorded during the 2015 - 2018 Run-2 of the Large Hadron Collider. Several diphoton and jet distributions related to the event kinematics as well as the multiplicity of the jets produced in association with the Higgs boson are measured and show good agreement with the Standard Model prediction. All presented results are unfolded to particle level. The integrated fiducial production cross section is measured to be 65.2 ± 7.1 fb, which is in good agreement with the Standard Model prediction of 63.6 ± 3.3 fb.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

*Speaker

PoS(EPS-HEP2021)572

ATL-PHYS-PROC-2021-069
07 October 2021

1. Introduction

During Run-2 of the Large Hadron Collider (LHC) [1] data-taking program running from 2015 to 2018, the ATLAS experiment [2] recorded 139 fb^{-1} of high-quality proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV. With this larger data set and updated analysis techniques, the precision of the measurements has been improved significantly compared to previous measurements. Differential cross section measurements in the $H \to \gamma\gamma$ decay channel [3] based on this data are summarized in the following.

2. Analysis Strategy

A selection of variables sensitive to various Higgs boson properties are measured inclusively in the Higgs boson production modes: The transverse momentum (p_T) distribution of the $\gamma\gamma$ system $p_T^{\gamma\gamma}$, the rapidity of the $\gamma\gamma$ system $y_{\gamma\gamma}$, and the azimuthal angle between the two highest p_T jets $\Delta\Phi_{jj}$, among others. They are sensitive to properties of perturbative QCD, couplings to b and c quarks and potential unknown heavy particles, the parton distribution functions of the colliding protons, and to Higgs boson spin and CP properties.

The analysis requires events with two isolated photons with a p_T of at least 35% and 25% of the invariant $\gamma\gamma$ mass in a pseudorapidity[1] range of $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. Further, their invariant mass must fulfill $m_{\gamma\gamma} \in [105, 160]$ GeV. Hadronic jets are reconstructed using an $R = 0.4$ anti-k$_t$ algorithm, must have a $p_T > 30$ GeV and an absolute rapidity $|y| < 4.4$. The definition of the fiducial measurement region closely matches this selection, in order to not introduce model dependencies due to acceptance effects.

A simultaneous profile likelihood fit to the $\gamma\gamma$ invariant mass $m_{\gamma\gamma}$ in all bins of the variable of interest extracts the signal yields, which is illustrated in Figure 1. The fit correlates the Higgs boson mass and the systematic uncertainties between all bins of a variable. To determine the signal model shape, a double-sided Crystal Ball (DSCB) function is fit to Monte Carlo (MC) simulated $H \to \gamma\gamma$ templates. The parametrization of the background is determined using MC-based templates. Its parameters are left to freely float in the final fit to data. More details on the background modelling strategy are given in Section 3. The extracted reconstruction-level signal yields are unfolded to particle level cross sections using bin-by-bin correction factors. The unfolding corrects for detector efficiency and resolution effects, allowing comparisons to theoretical predictions and results from other experiments.

3. Background Modelling Strategy

The parametrization of the background used in each bin of the profile likelihood fit is determined by a series of signal + background fits to background-only MC templates. These are built from the main backgrounds to this analysis: irreducible continuous $\gamma\gamma$ as well as reducible continuous γj and

[1] ATLAS uses a right-handed coordinate system, with the nominal interaction point at the origin. The y-axis points to the earth's surface, the x-axis points to the center of the LHC ring, and the z-axis points along the beam pipe. In the transverse detector plane, cylindrical coordinates (r, ϕ) are used, where ϕ is defined as the azimuthal angle around the beam pipe. The pseudorapidity is defined as a function of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

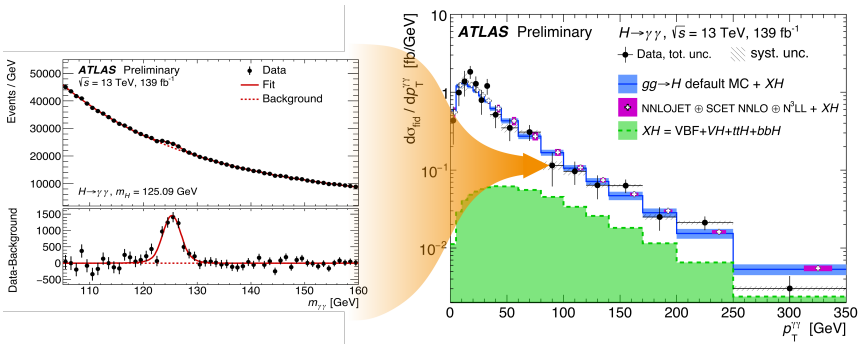
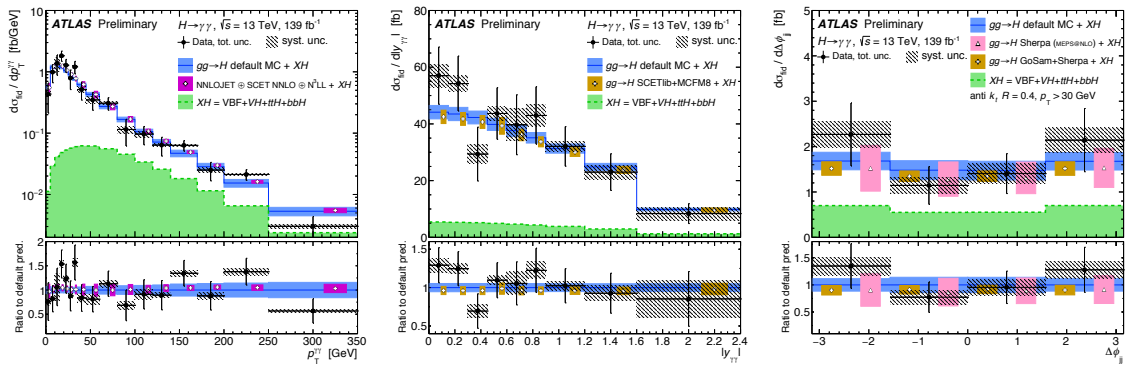


Figure 1: Schematic illustration of the differential cross section measurement for the $p_T^{\gamma\gamma}$ distribution [3]. For each data point in the plot on the right, the corresponding $m_{\gamma\gamma}$ distribution is fit in a simultaneous profile likelihood fit. Note that the plot on the left is only for illustration, it shows the $m_{\gamma\gamma}$ distribution integrated over $p_T^{\gamma\gamma}$.

jj events, in which jets were misidentified as photons. The relative contributions of these different backgrounds are estimated using a double two-dimensional sideband method, separately inverting the selection criteria for each photon. The integrated fiducial purities are $(75 \pm 4)\%$, $(22 \pm 3)\%$, and $(3 \pm 1)\%$, respectively. The template is then built by combining the different distributions according to their purities. For this, the shape of the $\gamma\gamma$ distribution is taken directly from MC simulation. The γj shape is obtained by reweighting the $\gamma\gamma$ shape using a fit to data control regions. The jj contribution has no impact on the final shape of the template and is neglected.

The total background has a smoothly falling distribution, of which the true form is unknown. An inappropriate choice of functional form to describe its shape could lead to a bias on the extracted signal yield. This bias is called *spurious signal*, and the function choice is optimized to limit the influence of this potential bias on the measurement. Possible functions which are considered as candidates for the background model are exponential functions of a first, second, or third order polynomial, third, fourth, or fifth order Bernstein polynomials, and a power-law function. In order to estimate the spurious signal for every function in every bin, signal + background fits are conducted to the background-only template with each candidate function. The fits are conducted by increasing the position of the peak of the DSCB in steps of 0.5 GeV from 121 to 129 GeV. The maximal signal yield of all fits is then used as the spurious signal of the probed model. The scan is done in order to circumvent a potential underestimation of the bias at the nominal Higgs boson mass, in cases where e.g. the bias is < 0 below 125 GeV, and > 0 above 125 GeV. Functions which have a spurious signal $\pm x$ times its statistical uncertainty of less than 10% of the expected Standard Model (SM) signal yield or less than 20% of the expected statistical uncertainty on the signal yield are considered as candidate functions for the background model. Here, x is a factor varied between 0, 1, and 2 and is only relaxed if no function passes with a lower factor. Additionally, functions are required to have a χ^2 probability of $> 1\%$ to describe the data sidebands, and an F-Test is conducted on the data sidebands to check if a higher-order function is needed to describe the background shape. If several functions pass these criteria, the model with the lowest number of degrees of freedom is chosen. The spurious signal is then used as a systematic uncertainty on the background model choice in the final fit to data.



(a) **(b)** **(c)**

Figure 2: Measured differential cross sections as a function of (a) $p_T^{\gamma\gamma}$, (b) $|y_{\gamma\gamma}|$, and (c) $\Delta\phi_{jj}$ [3]. The measurements are compared to selected SM theoretical predictions.

Since the $H \to \gamma\gamma$ analysis has a very small S/B ratio, the model choice and the spurious signal are very sensitive to even small fluctuations in the background templates. Such fluctuations could lead to an overestimation of the spurious signal. This would lead to a suboptimal function choice and overestimated systematic uncertainties. To circumvent this, the background templates are smoothed using Gaussian Process Regression (GPR) smoothing before the function selection procedure. This has been verified not to introduce any additional bias on the spurious signal estimation. While generating more events for the background templates would have a similar effect, it would require generating $O(10^9)$ additional events, so the GPR method is preferred.

4. Results and Conclusion

A selection of the measured differential cross sections compared to different SM predictions is presented in Figure 2. The measured integrated fiducial cross section times the $H \to \gamma\gamma$ branching ratio of $\sigma_{\text{fid}} = 65.2 \pm 4.5$ (stat.) $\pm$ 5.6 (syst.) $\pm$ 0.3(theo.) fb is in agreement with the SM prediction of 63.6 ± 3.3 fb. The integrated fiducial cross section is limited by the systematic uncertainties, and the differential cross sections are limited by the statistical uncertainty. All measured cross sections show a good agreement with the SM predictions.

Measurements with data collected by the ATLAS experiment during Run-2 of the LHC provide an improved precision view on Higgs boson properties. The $H \to \gamma\gamma$ channel is an excellent channel to conduct such measurements. To exploit its full potential, a reliable background model is essential.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 678215).

References

[1] L. Evans and P. Bryant, *JINST, 3:S08001*, 2008

[2] ATLAS Collaboration, *JINST 3 S08003*, 2008

[3] ATLAS Collaboration, *ATLAS-CONF-2019-029*, https://cds.cern.ch/record/2682800, 2019