EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

CERN-EP-2033-001
April 14, 2033

Search for the decay of an X boson into photons with the ATLAS detector

Guillermin Jules[1], Lecomte Samuel[2]

Abstract

This paper presents the results of the observation of data produced by proton-proton collisions at the HL-LHC at center-of-mass energy of 14 TeV. The analysis was focused on events where decay to two photons was observed. The optimal selection was obtained from three simulated samples. The observed (expected) mass of the boson should be between 500 and 1000 GeV, according to the theory *(X, 2033)* based on the standard model. Observation of the $\gamma\gamma$ decay channel showed no evidence of an excess of events compared to the background with an upper limit on signal strength of 0.3736 μ.

[1] Université Clermont Auvergne (ATLAS Collaboration), France
[2] Université Clermont Auvergne (ATLAS Collaboration), France

Contents

1. Introduction

The purpose of this experiment is to detect a potential boson called "X-Boson" which would largely solve the problems related to dark matter. The existence of this boson was theoretically demonstrated by the theoretical physicist Marianne X *(X, 2033)*, it would have a mass between 500 and 1000 GeV and would therefore be observable at the HL-LHC, especially from the ATLAS experiment. In particular, we will observe its decay channel into two photons in order to determine whether or not this boson exists.

2. The ATLAS Detector

The ATLAS detector *(Atlas Collaboration, 2008)* covers almost the entire solid angle about the proton–proton interaction point. It consists of an inner tracking detector, electromagnetic and hadronic calorimeters, and a muon spectrometer.

Charged-particle tracks and interaction vertices are reconstructed using information from the inner detector (ID). The ID consists of a silicon pixel detector, of a silicon microstrip detector, and of a transition radiation tracker (TRT). The ID is immersed in a 2 T axial magnetic field provided by a thin superconducting solenoid. The silicon detectors provide precision tracking over the pseudorapidity interval $|\eta| < 2.5$, while the TRT offers additional tracking and substantial discrimination between electrons and charged hadrons for $|\eta| < 2.0$.

The solenoid is surrounded by electromagnetic (EM) and hadronic calorimeters allowing energy measurements of photons, electrons and hadronic jets and discrimination between the different particle types. The EM calorimeter is a lead/liquid-argon (LAr) sampling calorimeter. It consists of a barrel section, covering the pseudorapidity region $|\eta| < 1.475$, and of two endcap sections, covering $1.375 < |\eta| < 3.2$. The EM calorimeter is divided in three layers,

longitudinally in depth, for $|\eta| < 2.5$, and in two layers for $2.5 < |\eta| < 3.2$. In the regions $|\eta| < 1.4$ and $1.5 < |\eta| < 2.4$, the first layer has a fine η segmentation to discriminate isolated photons from neutral hadrons decaying to pairs of close-by photons. It also allows, together with the information from the second layer, where most of the energy is collected, a measurement of the shower direction without assumptions on the photon production point. In the range of $|\eta| < 1.8$ a presampler layer allows corrections to be made for energy losses upstream of the calorimeter. The hadronic calorimeter reconstructs hadronic showers using steel absorbers and scintillator, or either copper or copper–tungsten absorbers immersed in a LAr active medium.

A muon spectrometer surrounds the calorimeter. It comprises separate trigger and precision tracking chambers in the magnetic field provided by three large air-core toroids.

3. Dataset

3.1. Datasets presentation

The first part of the work is based on the analysis of signals simulating the decay of a boson into two photons using Monte Carlo event generators *(Sjöstrand & all, 2015)* with the following parameters. The mass of the X-boson is located between 500 and 1000 GeV, the cross section is set to $\sigma = 5.4$ pb for proton-proton collisions at $\sqrt{s} = 14$ TeV and a filter efficiency of 0.17. The background signal is generated with SHERPA 5.11 event generator *(Bothmann & all, 2019)* considering a cross section $\sigma = 1286.5$ pb and a filter efficiency of 0.081. As we can see, we have two types of simulated data: Firstly, a set of three signals with masses within the assumed interval of the X-boson. Secondly, a simulated background that allows us, by comparing it to the first three signals, to make our cuts while reducing the background noise as much as possible.

We then proceeded to the analysis of the observed data collected by the ATLAS detector during Run III of the HL-LHC between 2031 and 2033. During this period, the integrated luminosity was $1298.1\ \text{fb}^{-1}$.

3.2. Events selection

We have several samples of 10,000 events for different potential masses of the X-boson. We start here with the one called signal_{1000} which corresponds to a potential mass of 1000 GeV. For each event we observe several variables for each of the two photons:

- Transverse momentum: momentum of the photon, in an ideal experiment, this value should be zero.
- Jet: shower of hadron, quarks decays into shower
- Phi: azimuthal angle
- Eta $= -ln(tan\frac{\theta}{2})$, pseudorapidity
- Isolation: particle (lepton) isolated or in the middle of a jet
- Missing energy: Missing energy from collision of proton beams

From the study of the three samples for different potential masses, we can calculate the invariant mass of the two-photon system. We expect to obtain an invariant mass similar to the mass of the desired boson. The formula is based on transverse momenta (p_T), pseudorapidity (η) and azimuthal angle (ϕ):

$$m_{inv} = \sqrt{2p_{T_1}p_{T_2}\,cosh(\eta_1 - \eta_2) - cos(\phi_1 - \phi_2)} \tag{1}$$

By applying (1) to our datasets, we get a result in accordance with our expectations. Indeed, we observe a peak of events in the interval [400; 1000] GeV where the X-boson must be located. In addition, the simulated background decreases well as expected, we have to refine our data selection to be able to exploit them. We also normalize the samples to have a realistic view of the probability of collision between 2 photons. We therefore assign weight to each sample. To calculate this normalization, we use:

$$N_{exp} = \mathcal{L} \times \sigma \times \epsilon \times \frac{N_{sel}}{N_{gen}} \tag{2}$$

Where N_{exp} is the expected number of events in collisions, $\mathcal{L}$ is the integrated luminosity of the collected dataset, σ is the cross-section for the process under consideration, ϵ is the filter-efficiency (if we restrict the event generation for example to a particular decay type of the particle. Here, $X \rightarrow \gamma\gamma$), N_{sel} is the number of selected events in our analysis, and N_{gen} is the number of events generated, before any other selections. In our case we take $N_{sel} = 1$ and N_{gen} is the total number of events in the sample. We also take into account the uncertainties given by: $w \times \sqrt{N}$ where N is the number of entries in that bin and w is the weight. We represent the statistical and systematic uncertainties in the histograms Figure 1:

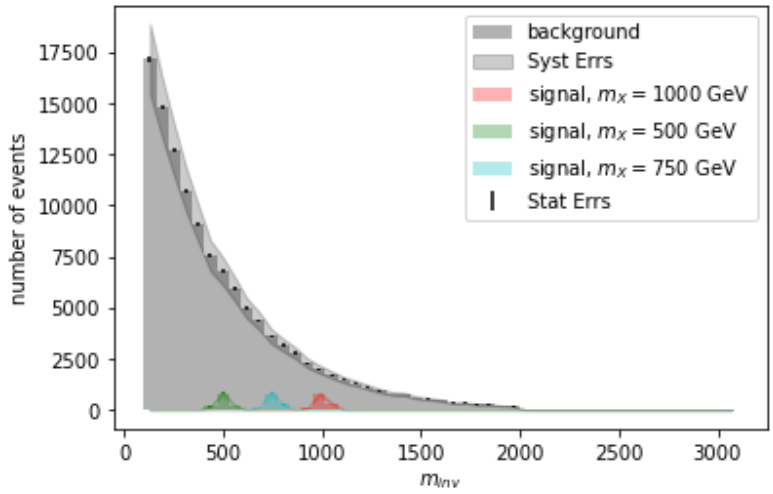


Figure 1: Invariant mass plot with uncertainties and normalization for simulated samples.

We then establish the search region for the X-boson, the theory predicts a mass in the interval [400;1100] GeV. We observe that the data of the simulated samples are located in this interval. We were able to count the number of events of each signal in this region and make plots of each variable in the same window.

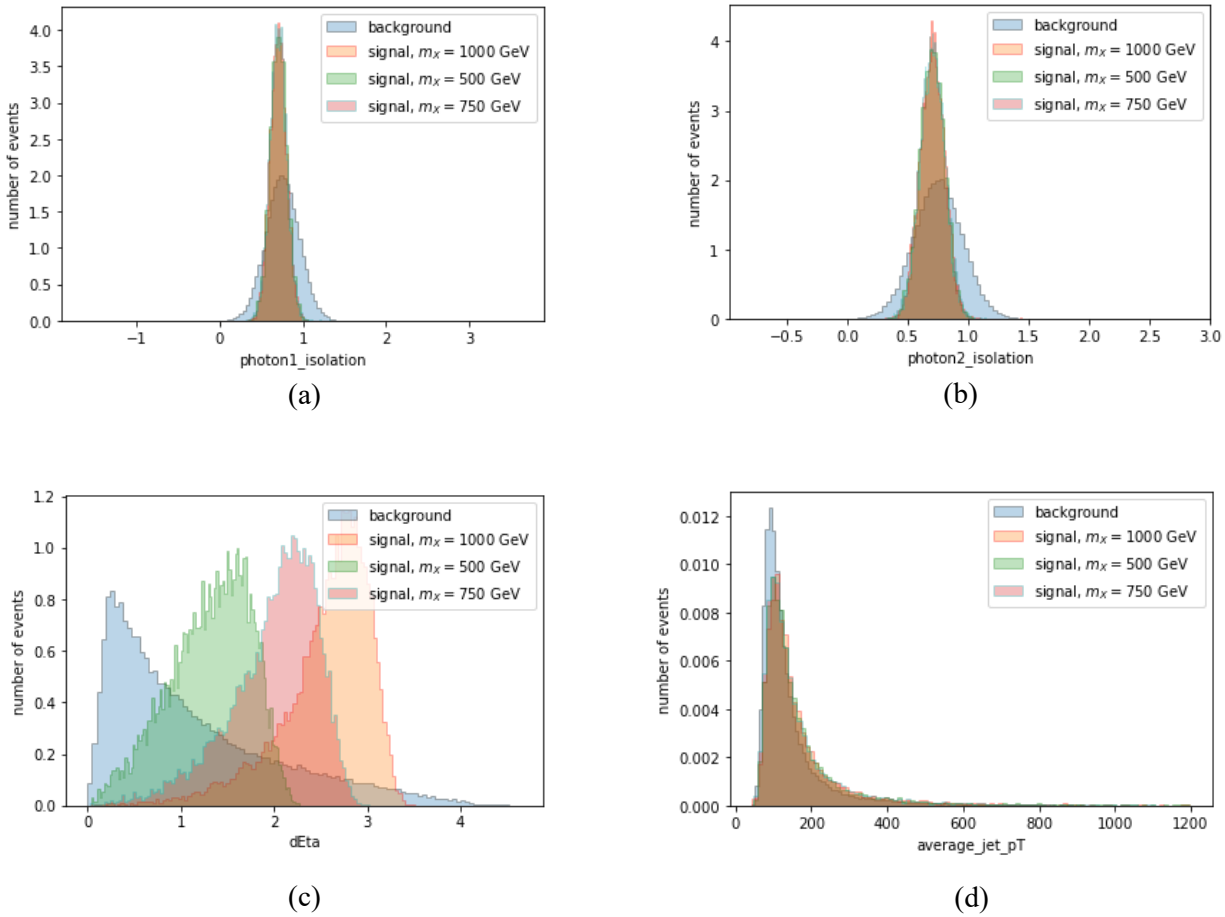


Figure 2: Plot of simulated samples' variables, only 4 of them are shown here. The green, red, and orange signals represent the simulated data and in blue we have the simulated background.

These four graphs Figure 2 show a signal that stands out from the background, it is primarily on these variables that we must act to clarify our selection. We now apply several cuts to the different variables mentioned above to optimize the Search Region and therefore reduce the background. It is important to know the impact that each cut has on the samples, which is why we introduce the sensitivity such as:

$$Z = \frac{N_s}{\sqrt{N_b}} \tag{3}$$

Where N_s is the weighted number of signal events passing the cut, and N_b is the weighted number of background events. If the new cut increases the sensitivity, then it is useful for the analysis. Before starting the cuts, we get the following values reported in Table 1:

Cut	Background	$Signal_{1000}$	$Signal_{500}$	$Signal_{750}$	$Z\text{-}signal_{1000}$	$Z\text{-}signal_{500}$	$Z\text{-}signal_{750}$	Z-average
None	135270.46	1191.66	1191.66	1191.66	3.24	3.24	3.24	3.24
Invariant mass window	46445.11	1190.58	1191.30	1191.66	5.53	5.53	5.53	5.53

Table 1: Preliminary results obtained after setting the search region to the invariant mass interval [400;1100] GeV.

We encounter our four simulated samples again together with the measurement of the sensitivity (Z) for each of them and immediately notice that the simple fact of focusing our analysis on the invariant mass interval of the X-boson already greatly improves the Z value. The choice of cuts is important because it is its quality that will allow us to get a sensitive analysis. For some variables, it is possible that we choose to cut off a portion of the signal in order to remove a larger portion of the background signal. The cuts were made on all variables, except those included in the calculation of invariant mass in order to not disturb our previous results. By optimizing them as much as possible, we obtain the following values.

Cut	Background	$Signal_{1000}$	$Signal_{500}$	$Signal_{750}$	$Z\text{-}signal_{1000}$	$Z\text{-}signal_{500}$	$Z\text{-}signal_{750}$	Z-average
None	135270.46	1191.66	1191.66	1191.66	3.24	3.24	3.24	3.24
Invariant mass window	46445.11	1190.58	1191.30	1191.66	5.53	5.53	5.53	5.53
photon1_isolation > 0.49	26165.36	1063.79	1065.94	1062.96	6.58	6.59	6.57	6.58
photon1_isolation < 0.87	26165.36	1063.79	1065.94	1062.96	6.58	6.59	6.57	6.58
photon1_isolation < 0.87	26165.36	1063.79	1065.94	1062.96	6.58	6.59	6.57	6.58
photon2_isolation < 0.88	17349.79	1015.05	1015.53	1007.54	7.71	7.71	7.65	7.69
n_jets < 3.94	11171.99	872.65	865.14	867.29	8.26	8.19	8.21	8.22
average_jet_pT > 43.36	11171.99	872.65	865.14	867.29	8.26	8.19	8.21	8.22
lead_jets_dPhi > -0.77	10012.72	869.67	862.04	864.19	8.69	8.61	8.64	8.65
lead_jets_dPhi < 6.29	8082.41	802.46	795.55	791.02	8.93	8.85	8.80	8.86
n_leptons < 1.41	3285.72	583.32	586.65	582.72	10.18	10.23	10.17	10.19

Table 2: Additive calculation of the sensitivity of each cut for each simulated sample.

Table 2 show that after all the cuts, the sensitivity is around 10, which represents more than three times greater than that which we had without the cuts. In addition, the number of events in the background has been reduced by about a factor of 40 while that in the signal samples by a factor of 2, the optimization of the Search Region is therefore successful.
From these cuts, we can again plot a histogram of invariant mass (m_{inv}) for all simulated samples, taking into account normalization and uncertainties.

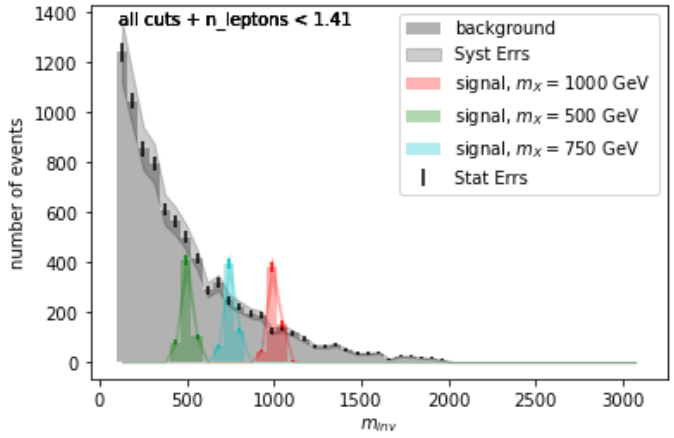


Figure 3: Histogram of invariant mass with all cuts taken into account.

Comparing Figure 1 and Figure 3, we immediately notice that thanks to the cuts made, the three signals representing the masses of the boson $m_X = 500; 750$ and 1000 GeV stand out more from the background. The number of events of this signal in the interval [400;1100] GeV decreased from 46445 to 3286 while retaining a significant proportion of the events of the other signals.

3.3. Signal and background modeling

Finally, we compute the amount of background in the signal region in order to see if we have an excess of events that might reveal the presence of the X-boson. To accomplish this step, we pick up the previous data again but with the window on the invariant mass reversed. We parameterize the background histogram with a descending exponential function and then integrate between 400 and 1100 GeV to determine the number of events included in the signal.

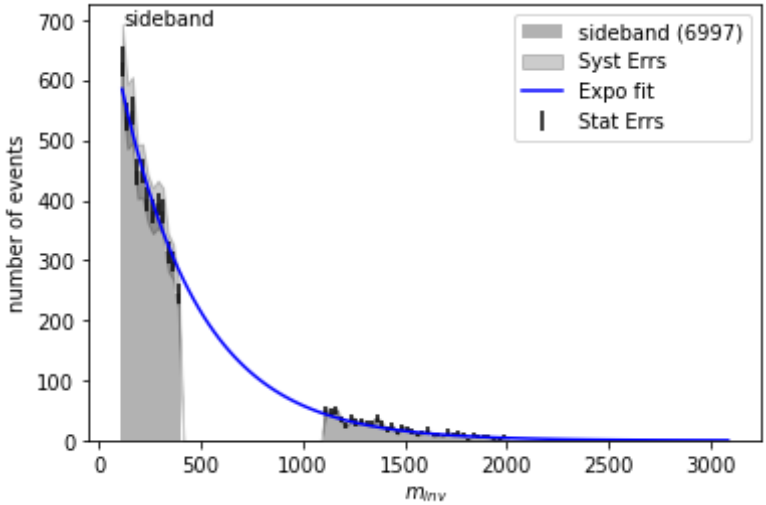


Figure 4:Data adjustment using a descending exponential parameter function (783.578951; 0.002598).

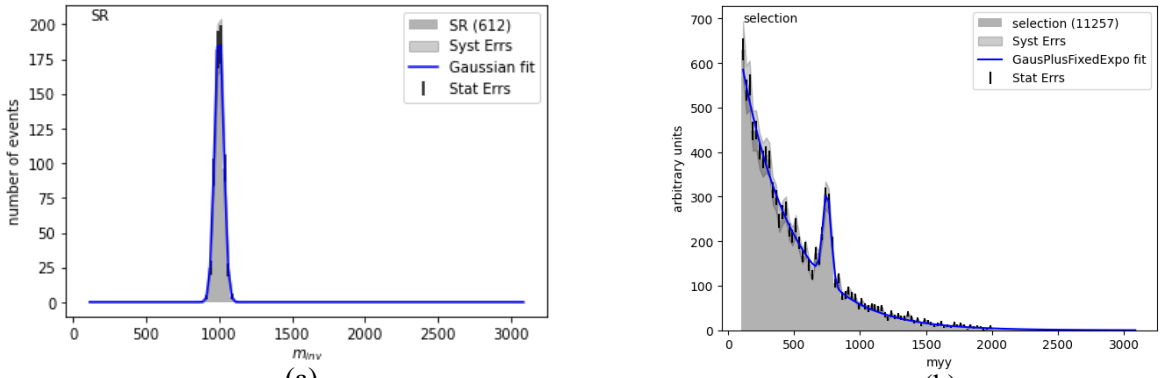
From this work, we estimate that there are (3578.9982 ± 109.4550) background events in the region [400;1100] GeV. The uncertainty about this value comes from our fit, we calculate the parameters of the function passing best through all the points but due to its finite precision, the result varies from one iteration to another. As a result, we calculate the significance of an excess, which represents the level of confidence we have between an expected value and an observed value. This value is determined for two different data choices: first the background alone and then by adding $signal_{1000}$. We got the following results:

```
SM only: obs = 3645.54  exp = 3579.00 +/- 109.46   s=0.6 sigma
SM+X1000 only: obs = 4258.17  exp = 3579.00 +/- 109.46   s=6.2 sigma
```

Figure 5: Calculation of significance (s), for background (SM) and background + nominal signal (SM+1000).

The hypothesis we take into account is that there is no X-boson in these simulated datasets. The results Figure 5 show that, a significance close to 1σ probably corresponds to a statistical fluctuation, while a result close to 5σ undoubtedly corresponds to the discovery of this boson here. In order to quantify it, we calculate its mass and cross-section from an adjustment. Since the events of the nominal signal are distributed in Gaussian, we adjust the data from this same function by taking the same parameters of the exponential as for Figure 4.

We determine Figure 6 *(a)* that the mass of the X-boson is (1000.19 ± 1.24) GeV, with a cross-section $\sigma = (5.39 \pm 0.32)$ pb. Finally, by applying this adjustment to the total signal, we notice Figure 6 *(b)* a perturbation on the mass of the particle, caused by the addition of background events.



(a) (b)

Figure 6: Final results of the analysis, (a) represents the Gaussian of the signal centered on a mass of 1000 GeV and (b) represents the final signal with the addition of background.

So far, we have worked with a set of simulated data. It is part of the unblinding procedure which consists of performing most of the work on simulated data in order to verify the proper functioning of our numerical tools and, above all, not to forget to process some of the data. Our analysis method is therefore ready to be confronted with the observed data, with the certainty of generating a minimum of statistical fluctuations.

4. Systematic uncertainties

During our study, we were tackled with two types of uncertainty. First, the statistical uncertainty is caused by the limited size of our samples. We can limit it by repeating the same calculation many times in order to obtain a more reliable mean value than the first approximation. We have calculated this uncertainty from the following formula $w \times \sqrt{N}$ with N the number of events in each bin and w the fixed weight. These results can be found in Figure 1.
Second, systematic uncertainties are errors that recur with each measurement or calculation, caused by, for example, bad calibration or incorrect choice of parameterization. These systematic errors are present and identical throughout the dataset, but by identifying them, it is possible to correct them with the application of a corrective factor.

5. Results for the $\gamma\gamma$ decay window on observed data

We now turn to the analysis of the observed data. The data_4 file contains 134,926 events. This dataset is analyzed using the method we developed in Part 3 with simulated samples. Firstly, we must add the invariant mass calculation and the uncertainties in order to be able to make the first figures.

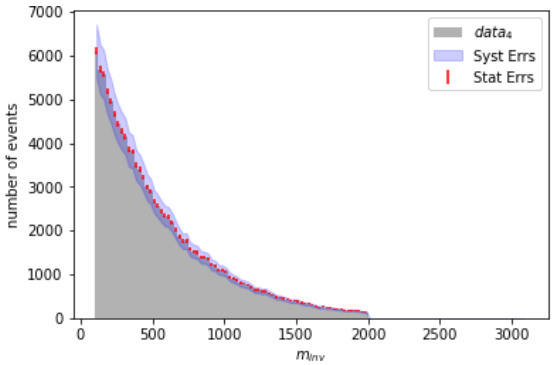


Figure 7:Invariant mass graph for data_4, statistical and systematic uncertainties are also shown.

By analyzing Figure 7, we note that there is no visible sign, for the moment, of the presence of the desired boson. The next step is to add the window on the invariant mass and apply all the cuts that we defined in Part 3.2.

Once the boson's search area is precisely delimited, we do the opposite by selecting events outside the [400;1100] GeV window of the invariant mass. As before, we make an adjustment according to an exponential law. This allows us to estimate the number of background events in the estimated mass range of the X-boson.

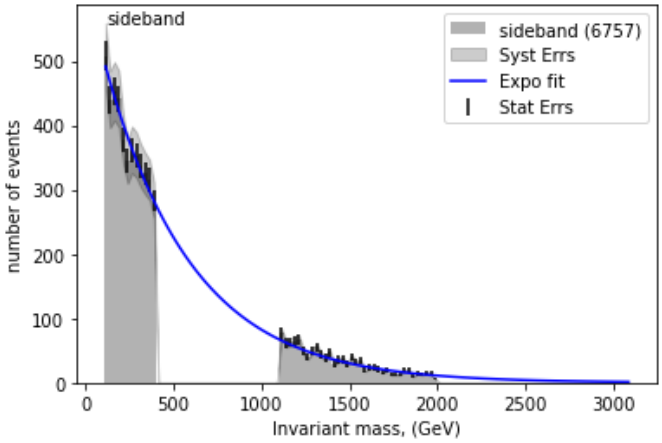


Figure 8:Adjustment using an exponential function with parameters (616.886765; 0.002015).

Based on the fit shown in Figure 8, we estimate that there are (4129.0362 ± 117.9092) background events in the interval [400;1100] GeV. This data allows us to obtain a significance $s = 0.5\sigma$. It seems that we are dealing with only a small statistical fluctuation that does not allow any conjecture as to the existence of the X-boson. In addition, we were able to calculate a limit on the cross-section. From here, we can produce a “Brazil Plot” Figure 9 to account for this limit. This representation replaces the cross-section with a signal strength index (μ) which is a percentage. Thus, taking the nominal cross-section for the simulated data which was equal to 5.4 pb corresponds to μ=1. In our case, we determined that we could exclude signals of strength above 0.3736 μ, corresponding to an effective cross section of 2,01744 pb.

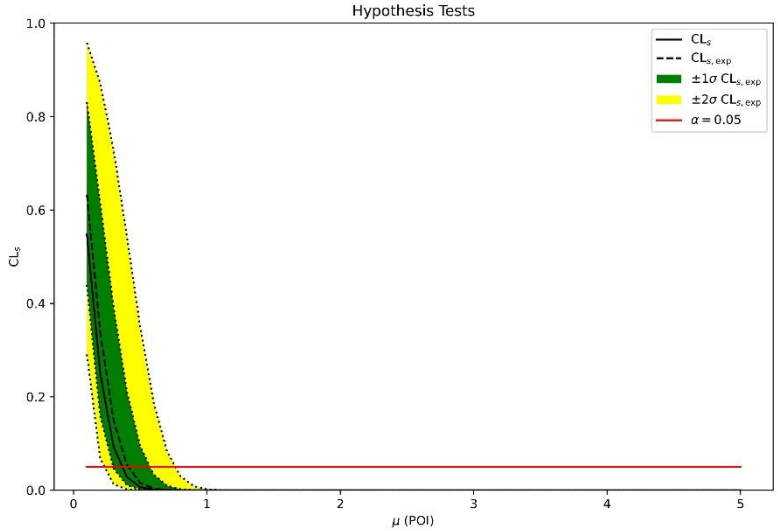


Figure 9: "Brazil plot" of our data allowing us to obtain an upper limit of 0.37μ.

6. Conclusion

Measurements of the mass and cross-section of the X-boson through its diphoton decay channel were made from proton-proton collision data collected by the ATLAS experiment at the HL-LHC. The data were taken at the center of mass at an energy $\sqrt{s} = 14$ TeV, corresponding to an integrated luminosity of $1298.1\ \text{fb}^{-1}$. Measurements showed no evidence of the X-boson in the [400;1100] GeV invariant mass window.

The uncertainties reported are mainly statistical uncertainties related to the numerical methods used. The confidence level is measured at $s = 0.5\sigma$, which shows that we are only experiencing statistical fluctuations. Finally, we estimated that we can exclude signals of strength above $\mu = 0.3736$, or $\sigma = 2{,}01744$ pb.

7. References

Atlas Collaboration. (2008). The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*. doi:10.1088/1748-0221/3/08/s08003

Bothmann, E., & all. (2019). Event generation with Sherpa. *SciPost physics*.

Sjöstrand, T., & all. (2015). An introduction to PYTHIA 8.2. *Computer Physics Communications*. doi:10.1016/j.cpc.2015.01.024

X, M. (2033). *Extension of the standard model: theory on the existence of an X boson.*

8. Appendix