Analysing events with the Z boson in the ATLAS experiment

Ksenija Kovalenka (10485506)
School of Physics and Astronomy
The University of Manchester
(Dated: May 15, 2022)

The cross-section value for the Z boson decay to two leptons $Z \rightarrow l^+l^-$ was obtained using ATLAS Open Data of pp collisions at $\sqrt{s} = 13\,TeV$. The data set corresponds to an integrated luminosity of 10.064 fb^{-1}. The result for the total cross-section of the leptonic branch was obtained to be $\sigma_Z^{ll} = 1.973 \pm 0.001(stat.) \pm 0.003(syst.) \pm 0.034(lumi.)\ nb$ by combining results for electron and muon channels. Data driven background estimation is used to reduce a systematic uncertainty. The calculated cross-section is consistent with the measurement obtained by the ATLAS collaboration [1].

I. INTRODUCTION

The Standard Model of particle physics is aiming to describe fundamental building blocks of matter and their interactions with a relatively small number of particles [2]. As with any other physics theory, its predictions have to be tested experimentally. The Large Hadron Collider (LHC) is the largest particle physics experiment built for this purpose [3]. This experiment was aiming to measure the cross-section of the Z boson decay to 2 leptons using data from the ATLAS experiment at the LHC collected in 2016 [4].

II. DATA COLLECTION

The ATLAS Open Data [4] used in this experiment contained events recorded at the ATLAS detector system and Monte Carlo data sets for a number of processes. The data set is corresponding to the integrated luminosity of 10.064 fb^{-1} with an uncertainty of 1.7% [5]. Following variables provided in the data sets were used in the analysis: invariant mass, transverse momentum, pseudorapidity and isolation variables (more details in Section IV A).

The ATLAS detector has a similar onion ring structure as a typical general-purpose detector. A layered system in the inner detector measures a position of a particle at several distances from the collision point to reconstruct the tracks of passing particles. There are Hadron and EM calorimeters in the next layer stopping the particles to measure the energy deposit they leave. The superconducting magnet system is providing solenoid and toroidal fields to bend the paths of the charged particles which helps to measure their charge and momentum. Muons are measured separately in the muon spectrometer because they are weakly interacting with the other parts of the detector [6].

Monte Carlo (MC) is a program designed to simulate the detector response to the specific processes happening in the proton-proton collisions. All probabilities of different events and the detector response to them are simulated using the best current theoretical knowledge. MC generated data sets for a range of processes were used in this experiment to decide on selection criteria for the real data set. Selection criteria were required to isolate Z events from the background. MC was used to estimate the background count in the final cross-section calculation as well (see Section III) [1].

III. THEORY

The target process for this experiment was Z boson decay to two leptons $q\bar{q} \rightarrow Z \rightarrow l^+l^-$. The Feynman diagram for the process is shown in Fig. 1.

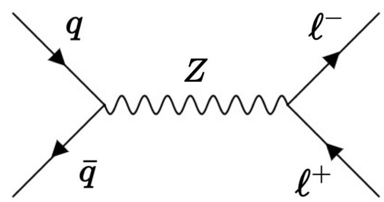

FIG. 1. Feynman diagram of Z boson production and decay.

The cross-section is describing the probability of a specific reaction to occur. It is expressed in the units of the area due to the origin of first calculations produced for firing particles at the target material. For the processes occurring at a particle accelerator, the cross-section σ is given by

$$\sigma(Z \rightarrow l^+l^-) = \frac{N_s - N_{bg}}{\epsilon \int Ldt}, \quad (1)$$

where N_s is the number of selected event from the real data set, N_{bg} is the estimate of the number of background events in the selected region, ϵ is the efficiency and $\int Ldt$ is the integrated luminosity corresponding to the total amount of events in the data set. The efficiency is estimated as a fraction of the number of events generated by MC for the specific process and the total number of MC events. Lepton universality suggests that the cross-sections in the electon and muon channels should be the same [7].

IV. ANALYTICAL TECHNIQUE

A. Event Selection Criteria

The main signature of the $Z \to l^+l^-$ event is the presence of two isolated leptons of the same type and opposite charge, because it is a neutral current weak interaction [2]. Basic selection cuts applied to the data set included these requirements.

The Z boson is produced at a relatively low boost in the beam direction due to its high mass and so the decay products will have a low boost as well, which means low longitudinal and high transverse momentum. The leptons must therefore have high values of transverse momentum p_t. The invariant mass distribution peak of two leptons must peak around the mass of the Z boson $m_Z \approx 91\,GeV$, which can be seen if figure 2. Further cuts were applied by comparing the real data and the MC event count for the distributions of the different variables, after the basic selection cuts were applied. We were aiming to cut out the regions where MC was poorly simulating the data and which were also justified by the expected properties of the leptonic pair discussed above. Some of the regions were investigated further by examining the invariant mass distributions of the isolated background and checking if it followed what is expected from signal events. If the invariant mass distribution looked differently, the events could be removed. The stacked plots of MC and real data in the electron channel for the invariant mass distribution for the selected signal region and for one of the background regions are shown in figure 2.

The range of the final selection cuts applied is summarised in Table I.

TABLE I. Final Selection Cuts

electrons	muons
$68GeV \leq m_{inv}^{ll}$	$70GeV \leq m_{inv}^{ll} \leq 120GeV$
$35GeV \leq p_t(1^{st}\ lepton)$	$32GeV \leq p_t(1^{st}\ lepton)$
$28GeV \leq p_t(2^{nd}\ lepton)$	$28GeV \leq p_t(2^{nd}\ lepton) \leq 97GeV$

The systematic uncertainty originated from the poor simulation of the signal by the MC reflected in the efficiency and errors in the background estimation. The variation in the selection criteria was used to estimate the systematic uncertainty in the final result by looking at the cross-section variation when the selection cuts for the specific variable were tightened or relaxed. In addition to familiar kinematic variables, isolation variables such as ptcone and etcone were used. These measurements are describing the sum of the transverse momentum of tracks and the energy deposit in the hadronic calorimeters respectively in the cone around the track of the particle of interest. We expect signal events to have small values for the isolation variables [4]. Combined ptcone of two leptons upper cut at $4-8GeV$ was introduced along with varying m_{inv}^{ll} and p_t cuts at this stage.

Other variables and cuts such as upper cuts on p_t around $80-100GeV$ were explored but not used in the final analysis because they showed to have a negligible effect on the cross-section calculation.

The statistical uncertainty was estimated by looking at the dominant error in the Equation 1, which was identified to be the selected events. Assuming Poisson statistics, the error was estimated to be the square root of the total counts.

B. Background Estimation

The background count was estimated using MC generate data sets for the following events: $Z \to \tau^+\tau^-$, $W^\pm \to l^\pm \bar{\nu}_l$ and $t\bar{t}$ decay to leptons. In addition, the data-driven so-called ABCD method was used to estimate the background which didn't have a MC generated data set. The main contribution to such background could be $b\bar{b}$ decay to leptons and the poor simulation of the W events by the MC. Four regions A, B, C and D correspond to the specific change in the basic selection cuts of charge and lepton type while keeping the rest of the selection criteria fixed. The schematic diagram of four regions used in data-driven background estimation is shown in Figure 1.

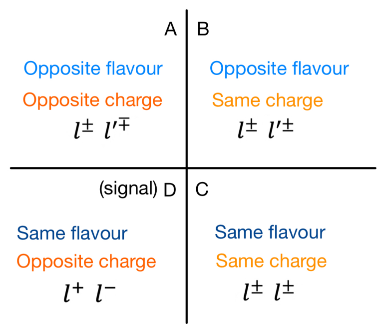

FIG. 1. ABCD method regions with corresponding selection criteria for the events.

Region D contained the signal while the rest were designed to only contain background. MC data for the background regions discussed above was removed from all of the regions to ensure that only unsimulated counts are left. The main assumption was that the same proportion of the background is in the region C compared to B as well as in the region A compared to D, allowing to estimate the amount of background in region D as $N_{bg}(D) = \frac{N_{bg}(A) \times N_{bg}(C)}{N_{bg}(B)}$ [8]. MC predicted the background count of around 20000 from about 3.5 million selected events, while ABCD estimate of around 10000 counts was added to it.

V. RESULTS

The variation of cross-section for both muons and electrons is shown in Figure 3, where each point corresponds to a variation in one of the selection cuts. It was decided to quote the final value as the average between the least

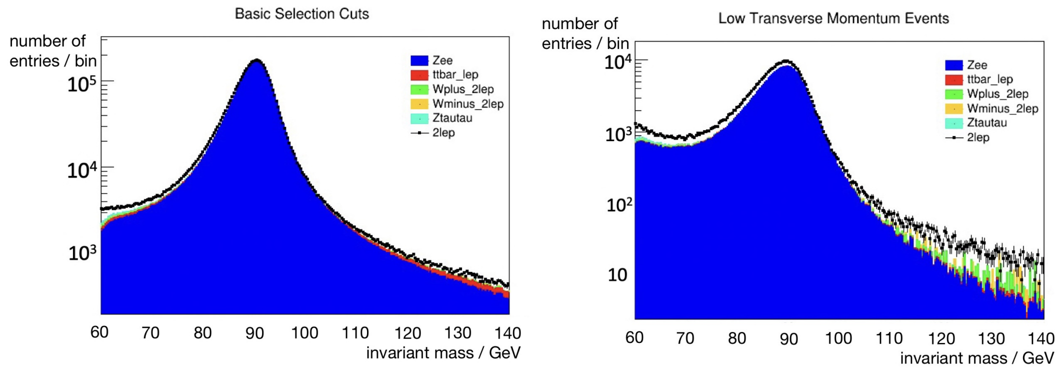

FIG. 2. MC and real data stacked plots for the signal region with the basic selection cuts applied on the left and the background region: $35GeV \geq p_t(1^{st}\ lepton)$ and $28GeV \geq p_t(2^{nd}\ lepton)$ on the right. 2lep in the legend is denoting the real data set.

and greatest value with statistical uncertainty as difference between the average and the edge values.

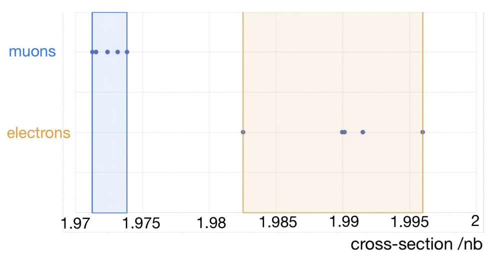

FIG. 3. Cross-section variation as different selection cuts are applied for electron and muon channel.

The final cross-section values obtained are $\sigma_Z^{ee} = 1.989 \pm 0.001(stat.) \pm 0.013(syst.) \pm 0.034(lumi.)\ nb$ for electrons and $\sigma_Z^{\mu\mu} = 1.973 \pm 0.001(stat.) \pm 0.003(syst.) \pm 0.034(lumi.)\ nb$ for muons. The values calculated for muons and electrons are equal within the uncertainty which is consistent with lepton universality mentioned in Section III. The total leptonic branch cross-section was obtained assuming lepton universality, therefore treating cross-section values for electrons and muons as independent measurements and combining results for them as a weighted mean. The final result is $\sigma_Z^{ll} = 1.973 \pm 0.001(stat.) \pm 0.003(syst.) \pm 0.034(lumi.)\ nb$ for the combined cross-section. It can be seen that the muon cross-section completely dominates in the total cross-section calculation as its systematic uncertainty is the order of magnitude smaller than one for electrons.

The systematic uncertainty on the values might be underestimated due to lack of number and variation in the selection cuts for different variables as a result of limited computation time available. Nevertheless, all cross-section values obtained are consistent with the ATLAS collaboration analysis for the same center of mass energy calculation, which is $\sigma_Z^{ll} = 1.981 \pm 0.007(stat.) \pm 0.038(syst.) \pm 0.042(lumi.)\ nb$ [1].

VI. CONCLUSION

The total Z boson decay cross-section for the leptonic branch was obtained to be $\sigma_Z^{ll} = 1.973 \pm 0.001(stat.) \pm 0.003(syst.) \pm 0.034(lumi.)\ nb$ using ATLAS Open Data for the pp collisions at the LHC. An effect of a wide range of selection cuts for different variables on the cross-section was investigated; a group of them were used to estimate the systematic uncertainty. Background counts were carefully eliminated using MC simulation and data-driven background estimation method. Final value is consistent with the cross-section stated by the ATLAS collaboration [1], but the uncertainties are likely to be underestimated and need further investigation.

[1] ATLAS Collaboration, Measurement of w± and z-boson production cross sections in pp collisions at s=13 tev with the atlas detector, Physics Letters B **759**, 601 (2016).

[2] B. R. Martin and G. Shaw, *Particle Physics* (John Wiley and Sons, 2008).

[3] CERN, The large hadron collider, `https://home.cern/science/accelerators/large-hadron-collider`, accessed: 2022-05-12.

[4] *Review of the 13 TeV ATLAS Open Data release*, Tech. Rep. (CERN, Geneva, 2020).

[5] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Tech. Rep. (CERN, Geneva, 2019).

[6] ATLAS Collaboration, The ATLAS experiment at the CERN large hadron collider, Journal of Instrumentation **3** (08).

[7] *Validation of Monte Carlo event generators in the ATLAS Collaboration for LHC Run 2*, Tech. Rep. (CERN, Geneva, 2016).

[8] S. Choi, Improved extrapolation methods of data-driven background estimation in high-energy physics (2019).