Search for $t\bar{t}$ resonances with ATLAS

by

Afaldy Hayeeteh
afaldy.hayeeteh@gmail.com

and

Md Rounak Jahan Raj
rounakjahanraj@gmail.com

Submitted to the Department of Physics

in fulfillment of Advanced Laboratory course for the degrees

of International Master on Advanced methods in Particle Physics

at

Technische Universität Dortmund

May 2025

Contents

Chapter 1

Introduction and Search Strategy

The Standard Model of particle physics (SM) is a well-established theory that describes the fundamental particles and how they interact. Despite its many successes, the SM has some important limitations. For example, it does not explain dark matter, does not include gravity, and has several unexplained parameters such as why particles have the masses they do.

The main goal of this experiment is to search for signs of new physics beyond the Standard Model (BSM). These BSM theories try to solve the SM's problems by predicting new particles and forces. A popular example is a hypothetical heavy particle called the Z' boson, which is similar to the SM Z boson but heavier. The top quark is the heaviest elementary particle, with a mass of about 173 GeV [2]. It was discovered at the Tevatron collider in the United States. Since 2010, top quark studies have been carried out mainly at the Large Hadron Collider (LHC) by looking at proton-proton (pp) collisions that produce $t\bar{t}$ pairs.

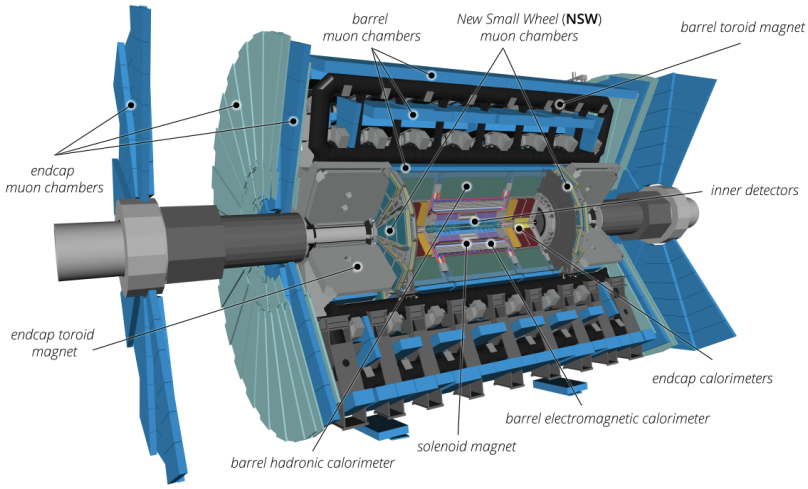


Figure 1-1: The ATLAS detector showing the main sub-systems [1].

One of the main experiments studying $t\bar{t}$ pairs events is the ATLAS detector at the LHC

1-1. ATLAS is a detector designed to detect almost any kind of new physics. It consists of several subsystems that work together to track particles and measure their energies and momenta. Although ATLAS does not detect top quarks directly, it observes the particles that result from top quark decays.

The detector can be divided into four main parts: the inner detector, the calorimeters, the muon spectrometer, and the magnet system. The inner detector is the closest to the collision point and tracks charged particles. It includes a silicon pixel detector which has very high resolution useful for tagging b-jets, silicon strips, and a transition radiation tracker (TRT) that helps separate electrons and positrons by detecting transition radiation. Surrounding the inner detector are the calorimeters, which absorb particles and measure their energy.

The muon spectrometer forms the outermost layer. Since muons can pass through the other layers without losing much energy. The inner detector is surrounded by a solenoid magnet that bends charged particle tracks allowing their momentum to be calculated. One limitation of detector including ATLAS is that it cannot directly detect neutrinos, as they interact only weakly and usually pass through undetected.

In this experiment, $t\bar{t}$ pairs are mainly produced through gluon-gluon fusion. We focus on the lepton+jets decay channel, where one top quark decays into a lepton (electron or muon) and a neutrino, and the other decays into jets, as illustrated in Figure 1-2.

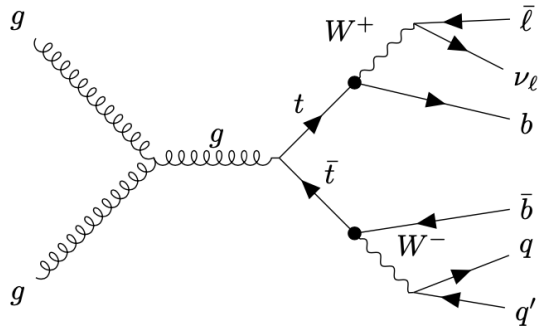


Figure 1-2: Feynman diagram showing $t\bar{t}$ production through gluon-gluon fusion and decay in the lepton+jets channel.

This channel is chosen because it provides a good balance between background and signal visibility. To account for the presence of neutrinos, missing transverse momentum (E_T^{miss}) is used as a key observable. To identify jets originating from b-quarks, b-tagging algorithms are applied. To perform the analysis, events are selected using criteria that enhance the signal over the background. These include requirements on the number of jets, a high-p_T lepton, and missing transverse energy (E_T^{miss}) to account for neutrinos. Backgrounds are estimated using Monte Carlo (MC) simulations and validated with real data. A final discriminant is used in a statistical test to search for deviations from the Standard Model prediction. A signal in this channel would suggest new physics.

Chapter 2

Dataset Description

Table 2.1 summarizes the datasets and simulated samples used in this analysis. It includes the recorded proton-proton collision data from 2016 as well as various Monte Carlo simulated backgrounds and signal samples.

Table 2.1: pp collisions recorded in 2016 at $\sqrt{s} = 13$ TeV

Sample	**Description**
Data	Events selected with a single-lepton trigger: electrons with $E_T > 25$ GeV or muons with $p_T > 25$ GeV.
SM $t\bar{t}$ background	Simulated at next-to-leading order (NLO) with parton shower.
Other backgrounds	Single-top events using best available theoretical predictions.
	W+jets and Z+jets events
Z' signal	
Mass points: 500–3000 GeV	Narrow width assumption ($\Gamma/m \approx 1\%$).
	SM-like coupling to $t\bar{t}$ with full detector simulation.

Chapter 3

Data Analysis

This section presents an analysis of data collected by ATLAS from pp collisions at $\sqrt{s} =$ 13 TeV in 2016, corresponding to datasets in Table 2.1. The analysis focuses solely on the lepton+jets channel: one charged lepton, one neutrino, and four or more jets, two of which are b-tagged.

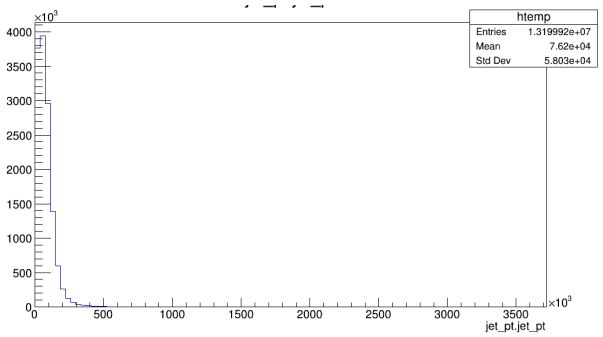
3.1 Event selection

To select events corresponding to the lepton+jets channel, the dataset was filtered according to the criteria listed in Table 3.1. These criteria aim to suppress Standard Model backgrounds to increase the chance of observing the $t\bar{t}$ signal in the lepton+jets channel.

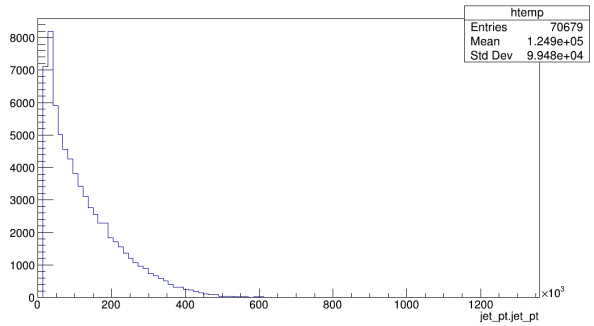
Table 3.1: Event Selection Criteria for the Lepton+Jets Channel

Criteria	**Requirement**
Number of leptons	Exactly 1 (single-lepton channel)
Lepton p_T	> 50 GeV
Lepton $\|\eta\|$	< 2.5 (detector acceptance)
Jet p_T	At least one jet with $p_T > 80$ GeV
Number of jets	3 – 5 jets (targeting 4-jet $t\bar{t}$ events)
b-tagging	At least one jet with MV2c10 > 0.83

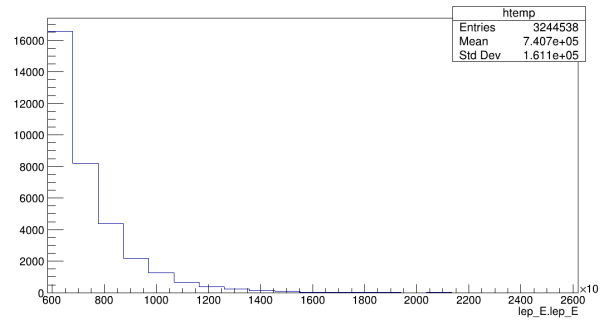
After applying event selection, Figure 3-1 presents examples of kinematic quantities illustrating low-level observables for the SM $t\bar{t}$ background and a Z' signal with $m_{Z'} = 1000$ GeV. Figures 3-1 (a) and (b) show the jet p_T distributions for the background of SM $t\bar{t}$ and the signal Z', respectively. In both cases, the jet p_T values are mostly below 600 GeV. Similarly, the lepton energy distributions lep_E shown in Figures 3-1 (c) and (d) are mostly below 1000 GeV. It is clear that the values in both plots overlap which means the cut based on this observable will not work well as it suppresses both the Standard Model background and the signal. This also occurs with other low-level observables such as missing transverse energy.



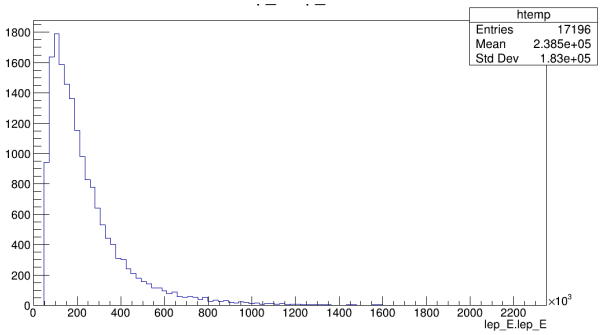
(a) Jet p_T distribution for the SM $t\bar{t}$ background



(b) Jet p_T distribution for the Z' signal



(c) Lepton energy distribution lep_E for the SM $t\bar{t}$ background



(d) Lepton energy distribution lep_E for the Z' signal

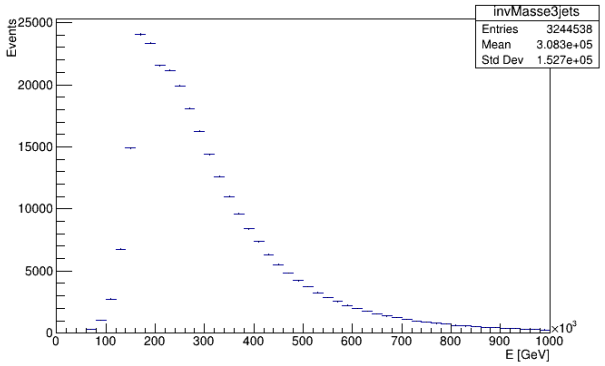
Figure 3-1: Distributions of kinematic variables after event selection. (a) and (b) show jet transverse momentum (p_T) for the SM $t\bar{t}$ background and Z' signal, respectively. (c) and (d) show lepton energy lep_E distributions for the SM $t\bar{t}$ background and Z' signal, respectively

The discriminating power of the low-level observables is insufficient for further analysis. Therefore, high-level observables are introduced to improve the separation between signal and background. High-level observables use combined information from low-level observables to better tell the signal apart from the background. This analysis identifies several suitable high-level variables, selected from the following:

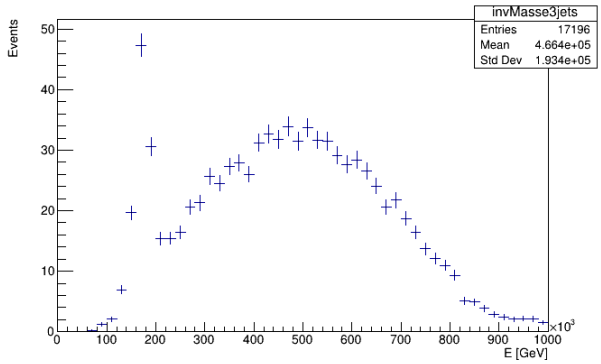
- E_{T}^{miss}, the missing transverse momentum magnitude;
- Azimuthal angle difference between E_{T}^{miss} and the lepton;
- Invariant mass of the three highest-p_T jets;
- Invariant mass of the four highest-p_T jets, lepton, and neutrino.
- Pseudorapidity of the system of four highest-p_T jets, lepton, and neutrino.

Through trial and error, the invariant masses of the three and four highest-p_T jets, lepton, and neutrino were selected as extra event selection criteria. Figure 3-2 shows a comparison between the SM $t\bar{t}$ background and the Z' signal with $m = 1000$ GeV. The three highest-p_T jets in panels (a) and (b) exhibit strong discriminative power: for the SM $t\bar{t}$, the distribution

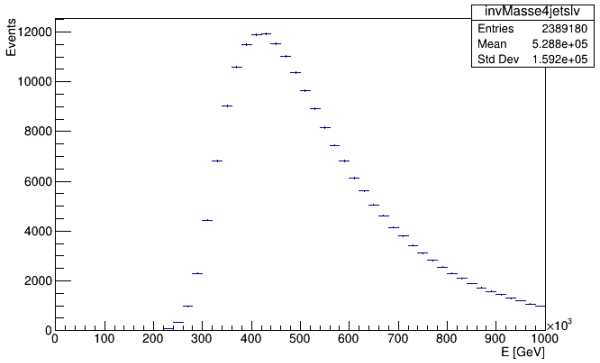
peaks near 100 GeV, while the Z' signal shows a Gaussian-like peak around 500 GeV. Similarly, the four highest-p_T jets in panels (c) and (d) display a clear skew in opposite directions between the two samples.



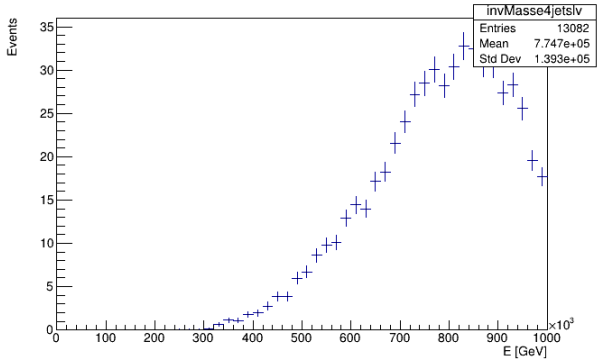
(a) Jet p_T distribution for the three highest-p_T jets, lepton, and neutrino in the SM $t\bar{t}$ background



(b) Jet p_T distribution for the three highest-p_T jets in the Z' signal



(c) Jet p_T distribution for the 4th highest-p_T, lepton, and neutrino jet in the SM $t\bar{t}$ background



(d) Jet p_T distribution for the 4th highest-p_T jet in the Z' signal

Figure 3-2: Comparison between the SM $t\bar{t}$ background and the Z' signal with $m = 1000$ GeV. Panels (a) and (b) show the jet p_T distributions for the three highest-p_T jets, where the SM $t\bar{t}$ background peaks near 100 GeV while the Z' signal peaks around 500 GeV. Panels (c) and (d) show the jet p_T distributions of the 4th highest-p_T jet, lepton, and neutrino, exhibiting opposite skewness between the SM background and the signal.

Until this point, the event selection using both low- and high-level observables filters the data expected to be used for next analysis.

3.2 Agreement of Simulation and Data

The next step in the analysis is to ensure data–MC agreement so that any sign of new physics can be studied under the background-only hypothesis. First, considering that the expected number of events N_{exp} represents the number of events detected for a given process.

$$N_{\text{exp}} = \mathcal{L} \cdot \sigma \cdot \varepsilon \cdot A \tag{3.1}$$

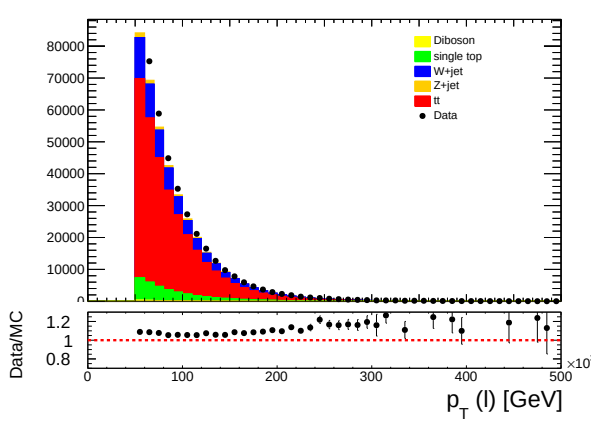
Where $\mathcal{L}$ is the integrated luminosity, σ is the production cross-section of the process, ε is the selection efficiency, and A is the detector acceptance.

To achieve the background-only hypothesis, all simulated events must be normalized to the integrated luminosity of the dataset. Each MC event is assigned a weight w which adjusts the contribution of that event so that the total sum of weighted events matches N_{exp}:

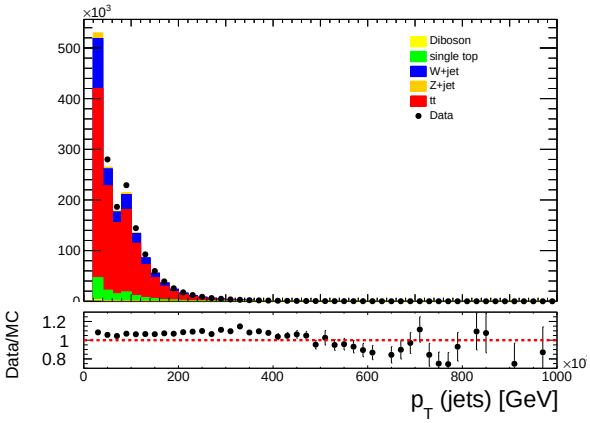
$$w = \frac{w_{\text{MC}} \cdot \mathcal{L} \cdot \sigma}{\sum w_{\text{MC}}} \tag{3.2}$$

Here, w_{MC} is the generator weight of the event and $\sum w_{\text{MC}}$ is the sum of all generator weights before any selection is applied. This weighting ensures the MC sample is normalized to the expected number of events observed in the data. So far Monte Carlo event weights w scale individual events so their sum matches the expected total number of events N_{exp} ensuring the simulation correctly represents the interested process.

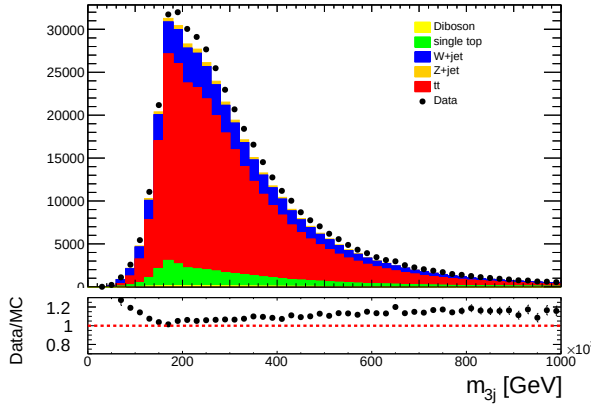
Figure 3-3 shows example distributions in a stacked plot comparing the agreement between simulation and data. It is evident that there are regions with significant uncertainties where the data do not align well with the simulation, suggesting that improvements in the underlying theory or simulation techniques may be necessary. Such effects are clearly observed in the p_T distributions of the lepton (a) and jets (b). However for high-level observables such as the highest three- (c) and four-p_T (d) jets both the simulation and data show good agreement. Hence, this level of agreement is sufficient to perform a statistical analysis for the search for a $t\bar{t}$ resonance.



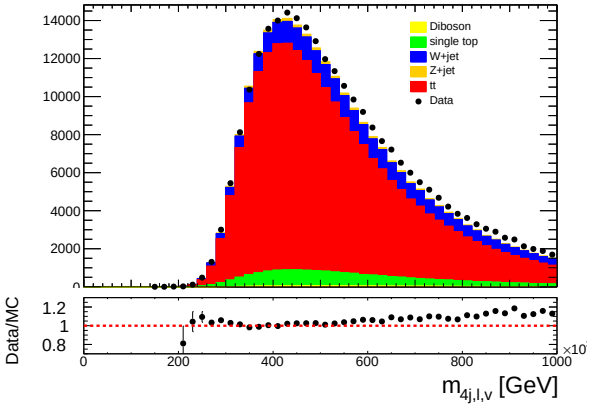
(a) Lepton p_T distribution



(b) Jet p_T distribution



(c) Invariant mass of the three highest-p_T jets



(d) Invariant mass of the four highest-p_T jets

Figure 3-3: Example distributions shown in a stacked plot comparing simulation and data. (a) and (b) reveal some disagreement in the lepton and jet p_T distributions, In contrast, (c) and (d), which represent higher-level observables, demonstrate good agreement between data and simulation.

3.3 Statistical analysis

The final step is to to evaluate how well the data agree with the Standard Model background. To quantify this agreement, calculating χ^2 tests if data agrees with the background-only hypothesis. A large χ^2 value signals possible new physics and helps set exclusion limits on models. The chi-square χ^2 is defined as

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{\sigma_i^2} \tag{3.3}$$

where O_i is the observed number of events in bin i, E_i is the expected number of events from the SM background in the same bin, and σ_i is the total uncertainty associated with the expected value. However, the χ^2 alone does not indicate whether the difference is from

by random chance. The background-only hypothesis is accepted or rejected based on the p-value defined as:

$$p = P\left(\chi^2_{\text{dof}} \geq \chi^2_{\text{data}}\right)$$

where the degrees of freedom (dof) is defined as

$$\text{dof} = \text{number of bins} - \text{number of fitted parameters} - 1.$$

Finally, the data can be used for a statistical analysis. Table 3.2 shows the results before and after including systematic uncertainties. A flat 14% uncertainty was added from a 4% uncertainty in the luminosity and about 10% from other sources like b-tagging and lepton identification.

Statistic	**Before Uncertainty**	**After Uncertainty**
χ^2	705.232	18.9422
p-value	1.71×10^{-116}	0.99998

Table 3.2: Comparison of χ^2 and p-value before and after including uncertainty.

Based on the data before including uncertainties, the χ^2 value is 705.232 with a p-value of 1.71×10^{-116}, which is far below the 5σ threshold of 5.7×10^{-7}. This indicates a discovery because it matches an extremely significant deviation from the background-only hypothesis. However, after including uncertainties, the p-value increases to 0.99998 showing no significant evidence of the Z' signal.

3.4 Theoretical Predictions and Exclusion Limits

Building on this result, we now turn to the theoretical interpretation and constraints derived from the data. Figure 3-4 displays the experimental limits on the production of the Z'-boson as a function of its mass.

The black curve shows the observed 95% confidence level limit and represents the maximum signal allowed by the data for each mass. The blue curve shows the predicted signal from a Z'-boson at each mass. If the blue curve is above the black curve, the predicted signal is larger than what the experiment allows. This means those masses are excluded at the 95% confidence level. No excess was seen up to about 2 TeV, so Z'-bosons below this mass are excluded. For higher masses, the predicted signal is too small to test with the current data, and those masses are still possible. This comparison shows which masses are allowed or ruled out by the experiment.

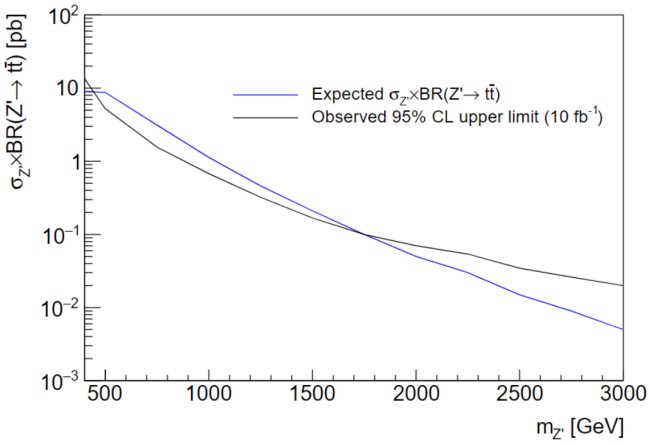


Figure 3-4: Theoretical limits and 95% confidence level exclusion limits on the production cross section for various Z'-boson masses. The x-axis shows the possible mass of the Z'-boson, while the y-axis shows the production cross section multiplied by the branching ratio for decay into top quarks

Chapter 4

Conclusion

A search was conducted for a heavy Z'-boson that decays into pairs of top quarks $t\bar{t}$ using data collected by the ATLAS detector at $\sqrt{s} = 13$ TeV. Distributions were compared between the data and predictions from the Standard Model. At first, a large difference was observed between the data and predictions, indicated by a very low p-value 1.7×10^{-116} before considering uncertainties. However, after systematic uncertainties included, the agreement improved significantly, with the p-value increasing to 0.99998. This means no strong evidence for the presence of a Z'-boson was found.

The high observable invariant mass of the $t\bar{t}$ pairs was used because it helps classify new particle signals from known background processes. Z' masses between about 410 GeV and 1800 GeV were excluded at a 95% confidence level. For masses above 2 TeV, the current data is not sensitive enough to draw conclusions.

Overall, limits on the existence of the Z'-boson have been placed through this analysis, and We, as master's students, have gained experience and understanding that will help us in future searches for new physics beyond the Standard Model.

Literatures

[1] G. Aad et al. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation (JINST)*, 3:S08003, 2008.

[2] R. L. Workman et al. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2022:083C01, 2022.