Monitoring fast discoveries at the LHC: Implementing jet searches in the TADA framework at the ATLAS detector

Bachelor's Project, Spring Semester

Author: Vilhelm Ekberg

June 2, 2016

Dr. Caterina Doglioni

Prof. Torsten Åkesson

Particle Physics

LUND UNIVERSITY

Abstract

In this project, part of the 2015 ATLAS data from $\sqrt{s} = 13$ TeV proton-proton collisions was studied with the TAG Data (TADA) framework. The project also serves as a general introduction to the data analysis tools used in high-energy physics experiments. However, since the TADA data had not yet been reviewed by the ATLAS collaboration, it was omitted in this thesis. It was revealed that some of the distributions were affected by trigger biases, which were removed by implementing a kinematic cut on the events. Furthermore, modifications of some TADA scripts were performed in order to successfully run the analysis. These could be beneficial to other Lund University students.

Acknowledgements

First of all, I wish to express my sincerest gratitude towards my supervisor Dr. Doglioni. You have provided invaluable support, which I could never have completed the project without. Whenever I stumbled across challenges in the programming, you would always point me in the right direction. Gabriel and Markus also deserve recognition for helping me with the TADA framework. Furthermore, I want to acknowledge my administrator Florido. You have been very supportive in setting up the data analysis tools, and helping me out with technical problem. Juno, Prim, and Sebastian, who have been working in the same room as I, also deserve my appreciation for their uplifting spirits and positive energy. Lastly, I want to thank my parents for encouraging me in my project, and for keeping our home a tranquil place for studying and working.

List of Acronyms

- **ADD** Arkani-Hamed Dimopoulos Dvali
- **DM** Dark Matter
- **EM** Electromagnetic
- **FCAL** Forward Calorimeter
- **HEC** Hadronic End-Cap
- **HLT** High Level Trigger
- **ID** Inner Detector
- **L1** Level-1
- **LHC** Large Hadron Collider
- **LO** Leading Order
- **MC** Monte Carlo
- **MS** Muon Spectrometer
- **NLO** Next-to-Leading Order
- **QBH** Quantum Black Hole
- **QCD** Quantum Chromodynamics
- **ROI** Region of Interest
- **SM** Standard Model
- **TADA** TAG Data
- **WIMP** Weakly Interacting Massive Particle

Contents

1 Introduction

1.1 Aim of the project

The study is part of the ATLAS experimental effort to probe for new physics beyond the *Standard Model* (*SM*). Today, this model explains much of the known physics at the elementary level. Nevertheless, the search is motivated by both theoretical shortcomings and empirical observations unaccounted for in the model. The ATLAS experiment is based on high-energy proton-proton collisions in the *Large Hadron Collider* (*LHC*) (fig. 1), a synchrotron with a circumference of 27 km [1]. It has recently been upgraded to provide collisions with centre-of-mass energies, $\sqrt{s}$, of 13 TeV [2]. Following the collisions, collimated streams of particles called *jets* are formed [2]. The kinematic properties of the jets will be analysed by the ATLAS, a multi-purpose particle detector [3]. These results are then displayed in different distributions, which may be compared with the expected influence of exotic physics [3, 4].

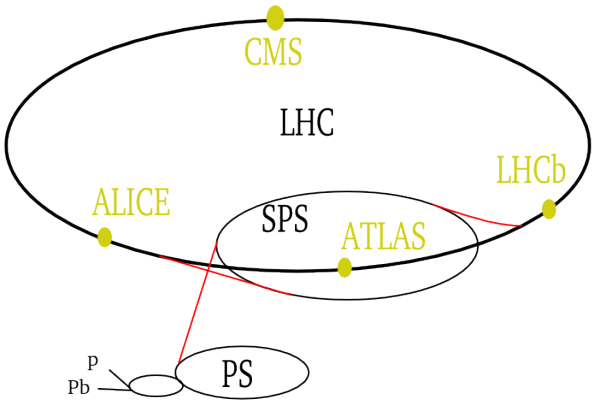

Figure 1: The LHC accelerator system along with the detectors situated around it. Figure taken from ref. [5].

In this project, the *TAG Data* (*TADA*) *Physics Monitoring* framework allows for *online* analysis of the ATLAS data [6]. This means the results can be analysed directly upon being recorded [6]. Furthermore, TADA only incorporates a limited amount of data, which allows the distributions to be produced shortly within data taking [6]. Thus, hints of new physics signals or potential problems in the distributions may be discovered quickly [6].

The specific aim of this project is to improve some TADA distributions from the 2015 data at $\sqrt{s} = 13$ TeV. This will be done by determining a selection cut to be applied on the events. However, since the TADA data has not yet been reviewed by the ATLAS collaboration, it will not be displayed in this thesis.

The remainder of section 1 is devoted to an overview of the SM and branches of new physics expected at the LHC. In section 2, the ATLAS detector and the theoretical framework surrounding the jets are described. The results are then discussed in section 3, with an outlook in section 4 summarising the main findings. Chapter 5 is reserved for appendix material. Lastly, natural units are used throughout in this thesis.

1.2 The Standard Model

The fermionic *quarks* and *leptons*, which have neither structure nor excited states, represent the fundamental particles known to date [1]. They interact mutually through four fundamental interactions: *the gravitational force*, *the electromagnetic* (*EM*) *force*, the *weak force*, and the *strong force* [1]. The interactions are mediated between the particles through *gauge bosons* [1]. Another component of the model is the *Higgs mechanism*, through which the quarks and leptons obtain their masses [1].

The SM is a result of several years of probing for the fundamental constituents of the Universe. Although the atomic theory was proposed already in Ancient Greece [7], it was not until 1911 that the nucleus was uncovered experimentally by Rutherford [8]. In 1897, Thomson discovered the electron - the first of six known leptons - through cathode ray experiments [9]. Additional discoveries of leptons, including that of the *muon* in 1936, would continue until the turn of the millennium [1]. Furthermore, between 1967 and 1973, inelastic lepton-proton scatterings provided the first evidence for smaller components within the nucleus - the quarks [10]. The Higgs mechanism was proposed in 1964, but only in 2012 was the existence of the *Higgs boson*, an important particle in this mechanism, verified experimentally at the LHC [11].

1.3 Quantum Chromodynamics

The theory of the strong interaction, which is mediated by the massless *gluon*, is called *Quantum Chromodynamics* (*QCD*) [1]. Only the quarks and gluons may interact strongly, and six *flavours* categorised into three *generations* have been discovered so far (fig. 2) [1]. The non-fundamental *hadrons*, which the proton is an example of, are in turn formed from the quarks [1].

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

Figure 2: The six known quarks, where: $u = up$, $c = charm$, $t = top$, $d = down$, $s = strange$ and $b = bottom$.

In QCD, the quantum numbers *color hypercharge*, Y^C, and *color isospin*, I_3^C, are assigned to the quarks and gluons [1]. Furthermore, *color confinement* prohibits them from existing individually, and the colour charge allows for self-interactions of the gluons [1]. The latter gives rise to *asymptotic freedom*, which suppresses the strength of the strong force at short distances of propagation [4]. On the other hand, the colour confinement causes the force to become strong over longer distances of propagation [4].

The process in which jets are formed is known as *hadroniztaion* [4]. The basics of this process can be understood by considering a quark and an antiquark interacting through the strong force. As their mutual separation increases, the colour confinement causes their mutual potential energy to increase [4]. When the separation is large enough, the creation of two new quark-antiquark pairs will be energetically favourable [4, 12]. A progression of these events gives rise to hadronization [4].

During high-energy proton-proton collisions, both quarks and gluons are involved in the interactions [13]. Thus, the impacts will result in them scattering and hadronizing into the collimated streams of hadrons called jets [4, 14]. The momentum of a jet is the sum of the individual momenta of the hadrons, and equals that of the quarks or gluons from which it originated [1, 15].

1.4 New physics

One empirical finding which the SM fails to explain is the pronounced difference in strength between the gravitational force and the other three fundamental forces [1]. A possible model of new physics, which addresses this issue, is the *ADD-model* [16]. In this configuration, gravity may act over 4+n dimensions, while the remaining interactions are restricted to a four-dimensional membrane [16]. A central aspect of the ADD-model is the relation between the (4+n)-dimensional *Planck Mass*, at which the strengths of all four forces become unified, and the n additional dimensions [4, 16]. The hierarchy problem stems from the four-dimensional Planck Mass being in the order of 10^{19} GeV [16, 17]. However, the existence of additional dimensions would imply a considerably lower (4+n)-dimensional Planck Mass [16,17]. Thus, the production of *Quantum Black Holes* (*QBHs*) at the LHC would be possible [4, 17].

If only the known matter in the SM is considered, observations in the 1930s regarding the observed rotational velocities of galaxies cannot be explained [18]. This motivated the introduction of *Dark Matter* (*DM*) - a new type of matter beyond the SM [16, 18]. It has been estimated that DM constitutes roughly 85 % of all matter [1], with its identity possibly attributable to *Weakly Interacting Massive Particles* (*WIMPs*) [16]. These were presumably non-relativistic in the early universe, so-called *cold dark matter* [16]. WIMPs represent additional exotic system expected to be detected for energies achievable at the LHC [16]. One more subdivision of new physics is the *four-fermion contact interaction*, which has a characteristic energy scale Λ [3, 19]. This suggests the existence of smaller components within the quarks, which are fundamental particles according to the SM [3,19]. The influence of the interaction is expected to increase for energies close to the energy scale Λ [3, 19], which ATLAS experiments could potentially help set limits on.

2 Experimental aspects

2.1 Jet kinematics

In the discussion of jet kinematics, the Cartesian x-,y- and z-axes constitute a right-handed coordinate system originating from the centre of the ATLAS detector [15]. The x-axis points towards the centre of the LHC ring, the y-axis points perpendicularly upwards relative to the horizontal line, and the z-axis runs along the tangent to the LHC ring [4]. It is also convenient to introduce the *polar angle*, θ, and *azimuthal angle*, ϕ, which are measured upwards with respect to the positive z- and x-axes, respectively [4]. (Fig. 1,3)

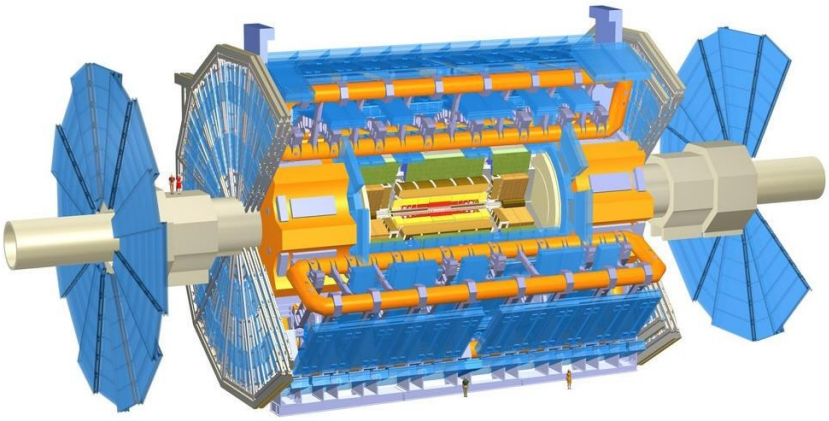

Figure 3: Schematic overview of the ATLAS detector. Figure taken from ref. [20].

Furthermore, the *rapidity* is given by

$$y = \frac{1}{2}\ln(\frac{E + p_z}{E - p_z}) \tag{2.1}$$

where E is the energy of the jet and p_z is the z-axis component of the jet momentum [4]. Additionally, y^* is defined as

$$y^* = \frac{1}{2}(y_1 - y_2), \tag{2.2}$$

with y_1 and y_2 being the rapidities of the two jets with the highest p_T, where p_T is the y-axis component of the jet momentum [3]. In this thesis, the jet with the highest transverse momentum is referred to as the *leading* jet, and the jet with the second-to-highest transverse momentum is called the *subleading* jet [21]. They will also collectively be referred to as a *dijet* [15].

Furthermore, the *pseudodrapidity* is defined by

$$\eta = \frac{1}{2}\ln(\frac{|p| + p_z}{|p| - p_z}) = -\ln[\tan(\frac{\theta}{2})] \tag{2.3}$$

where $|p|$ is the absolute value of the jet momentum [4]. For the pseudorapidities $\eta_{1,2}$ of the leading and subleading jets one defines [4]

$$\eta^* = \frac{1}{2}(\eta_1 - \eta_2). \tag{2.4}$$

It can be seen that the approximation $y \approx \eta$ is valid for jets with high absolute momenta relative to their masses [4].

Another variable to be introduced is

$$\chi = e^{|y_1 - y_2|} = e^{2|y^*|}, \tag{2.5}$$

where it is evident that $\chi \geq 1$ [4]. Lastly, the *rapidity boost* of the two leading jets relative to the collision centre of the protons is defined by [22]

$$y_B = \frac{1}{2}(y_1 + y_2). \tag{2.6}$$

The *dijet invariant mass*, m_{jj}, is equivalent to the energy of particles produced at rest in the collisions, which then decay into dijets [1, 15]. In the *dijet invariant mass distribution*, to be introduced later, it is of interest to look for "bumps" [15]. These are caused by increased number of events for certain m_{jj}, and could indicate the creation of new particles, also referred to as *resonances* [1,15]. The resonances have widths in m_{jj} due to their finite lifetimes and the Heisenberg uncertainty principle [1].

For the following derivation, a dijet formed from a QCD scattering event at leading order (LO), with two particles in the initial and final states, is considered [4, 15]. An expression relating m_{jj} to the transverse momenta p_{T1} and p_{T2} of the leading and subleading jets, respectively, is given by [4]

$$m_{jj} = \sqrt{p_{T1}p_{T2}}\sqrt{e^{2|y^*|} + e^{-2|y^*|} - 2\cos(\Delta\phi)}. \tag{2.7}$$

Here, $\Delta\phi$ is the absolute azimuthal angle between the jets [4]. In order to simplify the calculations, we consider the situation in which the momenta of the colliding particles cancel. Thus, the jets will be emitted "back-to-back" in ϕ and have equal p_T. As a result, the above expression becomes [4, 15]:

$$m_{jj} = \sqrt{p_T * p_T}\sqrt{e^{2|y^*|} + e^{-2|y^*|} + 2} = p_T(e^{|y^*|} + e^{-|y^*|}). \tag{2.8}$$

There are also next-to-leading order (NLO) processes with 3 particles in the final state, the contributions of which are neglected in the expression above [4, 15].

2.2 The ATLAS detector

The cylindrical geometry of the ATLAS provides an almost complete coverage of all angles around the collision centre (fig. 3). The innermost component is the *inner detector* (*ID*), which measures the momenta of charged particles [19]. It covers the region $|\eta| < 2.5$, with the surrounding solenoid providing a 2 T magnetic field [23]. The next layer is the *EM calorimeter* extending over $|\eta| < 3.2$ [4]. This is used to measure the energies deposited by EM showers [1]. Outside of the EM calorimeter, the *hadronic calorimeter* measures the energies deposited by hadronic showers for $|\eta| < 3.2$ [4, 19]. This consists of a *barrel region* covering $|\eta| < 1.7$ and two *hadronic end-cap calorimeters* (*HECs*) extending over $1.5 < |\eta| < 3.2$ [23]. The outermost component is the *muon spectrometer* (*MS*), which measures the momenta of muons for $|\eta| < 2.7$ [24]. Lastly, the *forward calorimeters* (*FCALs*) in the region $3.1 < |\eta| < 4.9$ provide calorimetric measurements of both EM and hadronic showers [23].

A jet finding algorithm is used to reconstruct the jets from the formed particles in the fragmentation [19, 25]. In the ATLAS detector, three-dimensional *topoclusters* are formed around calorimeter cells registering energy deposits with good signal-to-noise ratios [19, 25]. The noise partly stems from electronic fluctuations in the calorimeters and from *pile-up*, where events from several proton-proton collisions are detected simultaneously by the calorimeters [26]. These topoclusters are used as inputs to the *anti-k_T algorithm* [19]:

$$d_{i,j} = min(p_{Ti}^{-2}, p_{Tj}^{-2})\frac{\Delta y^2 + \Delta\phi^2}{R^2},$$

$$d_{i,B} = p_{Ti}^{-2}.$$

Here, $d_{i,j}$ is the distance between particles i and j, with their transverse momenta given by $p_{Ti,Tj}$ [19]. The recombination parameter R represents a geometrical section $\sqrt{y^2+\phi^2}$ in the detector [11]. The differences in rapidity and azimuthal angle between particles i and j are given by Δy and $\Delta\phi$, respectively [19]. Merging of i and j occurs if $d_{i,j} < d_{i,B}$, where $d_{i,B}$ is the distance between particle i and the proton beam [19].

ATLAS uses a system of triggers to discard unwanted events [27]. The *level-1* (*L1*) *triggers* are designed to search for *Regions of Interest* (*ROIs*), where jets might be found [27]. Only those regions will be forwarded to the *high level triggers* (*HLTs*) [27]. One example of a HLT is the HLT_j360, which only selects events where the p_T of at least one jet is above the trigger threshold at 360 GeV [4]. This threshold is approximately in accordance with the scale of *offline* jets, i.e. reconstructed jets that are not analysed at the moment of data taking.

The *trigger efficiency* represents the ratio of the number of events selected by a trigger to the number of entries selected by a *reference trigger*. It represents the probability of an event being selected by the trigger under investigation, given that a jet with a certain p_T is created in the event [4]. Therefore, the trigger efficiency may be evaluated as a function of the jet p_T [4]. The reference trigger must be fully efficient at the energy of interest, and it must have a lower threshold than the trigger under investigation.

2.3 Dijet distributions and kinematic cuts

One way to evaluate the results is trough the earlier mentioned dijet invariant mass distribution. Here, the number of registered events are displayed as a function of m_{jj} [1, 14, 28]. A QCD model for the background m_{jj} distribution, which is described in ref. [14, 28], is then fitted to the data points. New resonances are expected to cause significant discrepancies between the data and the fit for narrow intervals in m_{jj}, resulting in the formation of "bumps" [15,28]. The BUMPHUNTER algorithm returns the *p-value* of the deviation between the data and the model least likely stemming from statistical background fluctuations [29]. The lower the p-value, the higher the possibility of the deviation being attributable to a new resonance [29, 30].

In the χ *distribution*, the number of entries are plotted against χ in bins of m_{jj} [31]. Here, new physics is expected to to cause the number of entries to increase as χ decreases [31]. SM physics, on the other hand, is expected to result in a distribution that is more or less constant as a function of χ [31]. It is also anticipated that exotics physics will be prominent for high m_{jj} bins, while QCD processes will dominate for lower m_{jj} bins [31].

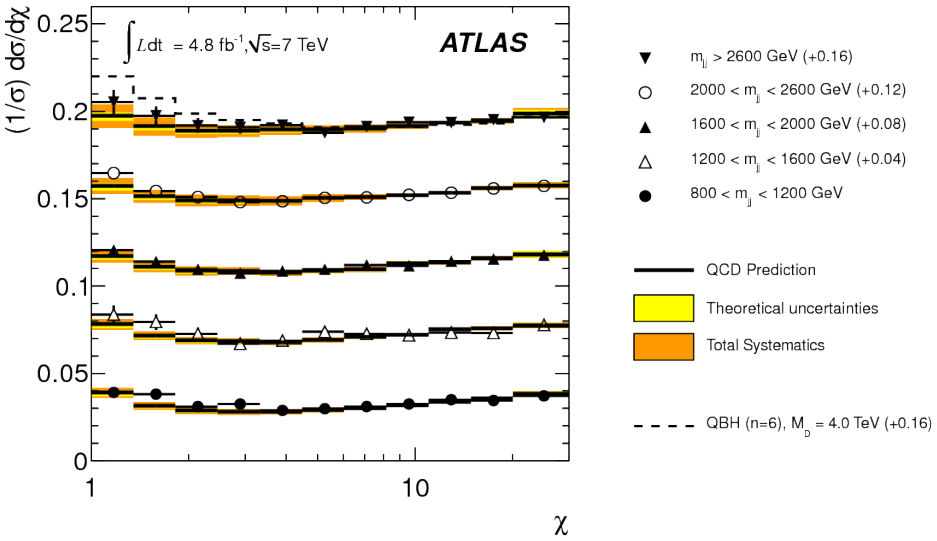

Figure 4: An example of a χ distribution at $\sqrt{s} = 13$ TeV, with the expected influence of QBHs included. Figure taken from ref. [31].

Furthermore, the F_χ is defined as the ratio of the number of events with *central* dijets to the number of events fulfilling the $|y^*|$ cut to be explained later [3, 31]:

$$F_\chi = \frac{N(|y^*| < 0.6)}{N(|y^*| < 1.7)}. \tag{2.9}$$

Here, events with $|y^*| < 0.6$ are expected to show signs of new physics, whereas QCD physics become more important as $|y^*|$ approaches 1.7 [3]. In the F_χ distribution, this variable is evaluated as a function of m_{jj} [3, 31]. A quark contact interaction is expected to cause an increase in F_χ as m_{jj} approaches the characteristic energy scale Λ of the interaction [3,31]. By comparison, F_χ should be more or less constant with m_{jj} if no new physics is involved [3, 31].

The *partonic cross-section* represents the cross-section of the scattering events between the quarks and gluons in the proton-proton collisions [4, 15]. This quantity is related to the number of registered events by the detector [4, 15], and is expected to be altered by new physics [15]. The sensitivity of the search to the partonic cross-section, and therefore the influence of exotic physics, can be optimized by limiting the range of y_B [15, 19].

Furthermore, a kinematic cut $p_{T,min}$ is enforced to restrict the p_T range of the leading jet. In each trigger, this cut must be set where the trigger efficiency is at least 99.5 %, in order to avoid trigger biases [4]. Additionally, the cut $|y^*|_{max}$ is applied with respect to the maximum $|y^*|$ for which the ATLAS can detect jets and to the $|y_B|$ cut [4]. By inserting $p_{T,min}$ and $|y^*|_{max}$ into expression (2.8), the minimum m_{jj} for which the kinematic cuts do not introduce a bias is determined [4, 15]:

$$m_{jj,min} = p_{T,min}(e^{|y^*|_{max}} + e^{-|y^*|_{max}}). \tag{2.10}$$

As earlier mentioned, the leading and subleading jets were assumed to have equal p_T in this expression, and only LO QCD scattering events were considered [4, 15]. If NLO processes are also considered, expression (2.10) must be modified by a factor of $\sqrt{2}$ [4].

3 Results

3.1 Fitting the trigger efficiency curve with a Python fit routine

The trigger efficiency had already been measured as a function of the jet p_T for the HLT_j360, which was used in producing the 2015 dataset at $\sqrt{s} = 13$ TeV. It was of interest to determine the p_T for which the trigger efficiency reached 99.5 %. A modified version, $a * \mathrm{erf}((p_T - b) * GeV^{-1}) + c$, of the Gauss error function, $\mathrm{erf}(x) = (2/\sqrt{\pi}) * \int_0^x e^{-t^2} \mathrm{dt}$, with a,b and c representing fit parameters, was used to fit these data points. For comparison, an empirically motivated formula, $(50\ \%) * \mathrm{erf}((p_T - a)/\sqrt{2 * b^2}) * (1 - c) + (50\ \%) * (1 - c)$, was also utilised as a fit function. The regression was conducted in the Python 2.7.6 language with the native fit routine *scipy.optimize.curve_fit()*.

The Python routine was designed to minimise the χ^2 of the fit, defined as $\Sigma_i (y_i - f_i)^2/\sigma_i^2$. This definition of χ^2 would also be used when comparing the accuracies of the fits. Here, y_i represented the data points, f_i the efficiency values of the fit functions at the data points, and σ_i corresponded to the weights. The errors of the data efficiency values were used as weights for the fitting. Data points with errors rounded down to zero were assumed to have unspecified errors. Those data points were thus treated as unweighted, meaning the corresponding weights were set to 1.

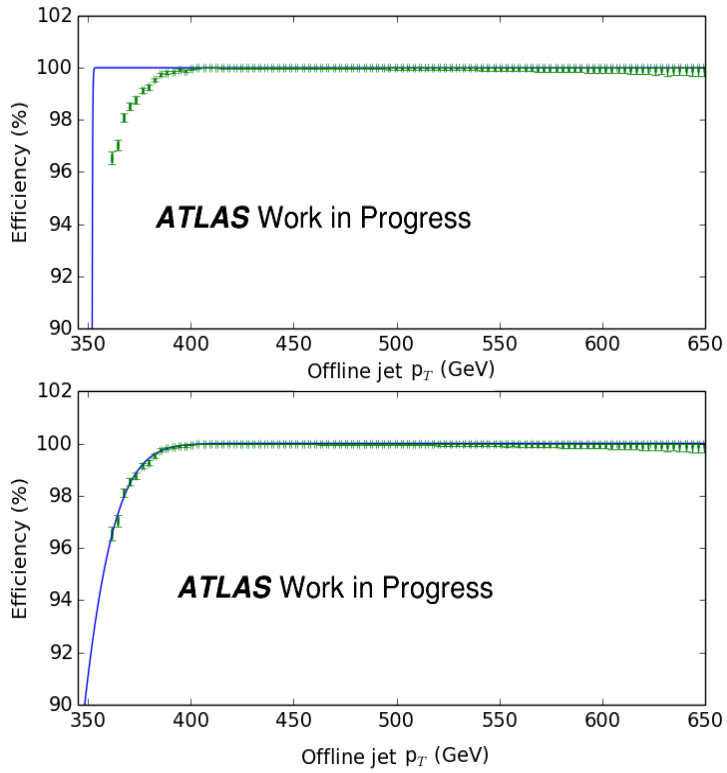

Figure 5: The data points and the solid lines representing the fit functions are shown, with the upper plot displaying the modified error function and the lower plot displaying the empirical function. Data points where the offline jet p_T was below the trigger threshold at 360 GeV were rejected.

The χ^2 of the modified error function was computed to 225.39 for 94 degrees of freedom, and for the empirical function it was calculated to 3.09 (fig. 5). From these results, it is concluded that the fit with the modified error function was very inaccurate, while the empirical function agreed well with the data. This was also evident from a rough comparison of the two fits by eye (fig. 5).

As can be seen, the modified error function fit closely resembled a leap function (fig. 5), which a trigger efficiency curve should theoretically behave as [4]. Therefore, this regression highlighted the disagreement between theory and the trend displayed by the data points before the plateau region (fig. 5).

The p_T values of the *full-efficiency points*, where the efficiency values of the fit functions first exceeded 99.5 %, were also determined. These were calculated analytically by incorporating the returned fit parameters. Consequently, it was calculated to 352.76 GeV for the modified error function. As for the empirical function, it was calculated to 380.89 GeV.

3.2 Fitting the trigger efficiency curve with a native ROOT fit routine

While still using Python 2.7.6, the ROOT fit routine $TH1 :: Fit()$ was also used for fitting the same two functions in the previous section. In this technique, the data points first had to be converted into histograms for the fitting. The comparison of the two techniques was motivated by $TH1 :: Fit()$ being more widely used in ATLAS projects.

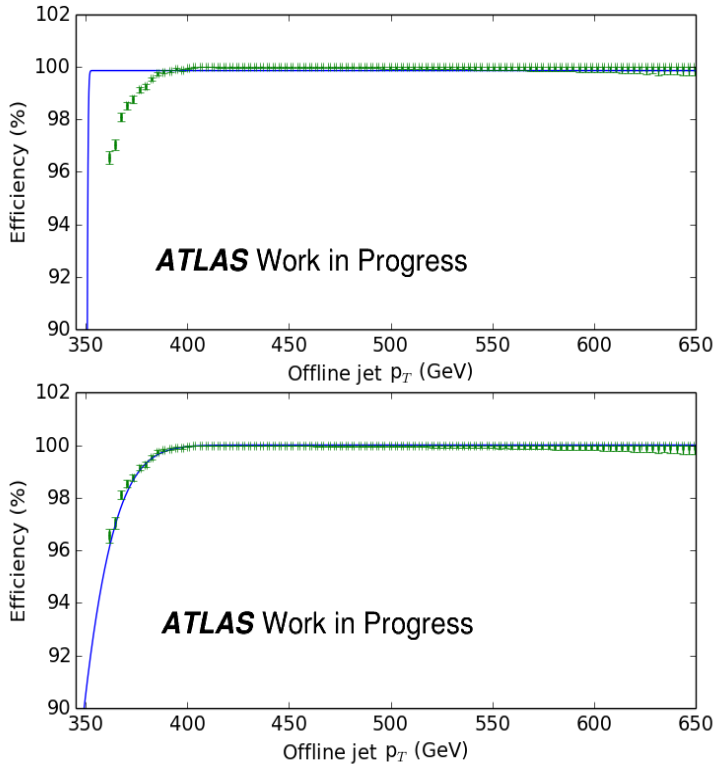

Figure 6: As in fig. 5, the upper plot displays the modified error function whereas the empirical function is shown in the lower graph. Again, data points for which the offline jet p_T was below the trigger threshold at 360 GeV were excluded. The number of bins in the histograms, excluding under- and overflow bins, equalled the number of data points.

Here, the χ^2 for the modified error function was computed to 1003.40, and for the empirical function it was calculated to 10.00. The full-efficiency points were computed to 351.91 GeV and 382.74 GeV for the modified error function and empirical function, respectively. From a study of the curves (fig. 6), as well as χ^2, the inadequate agreement between the data and the modified error function is clear. By contrast, the empirical function provided a significantly more accurate regression.

The conversion of the data points into TH1 histograms affected the binning of the offline jet p_T during the fits with the ROOT technique, compared with the Python routine. Furthermore, bins with errors rounded down to zero were excluded from the fits with the former technique, while all data points were used in the fits with the latter technique. The two routines also had slightly different built-in definitions of the χ^2 to be minimised in the fits (although the Python definition was used when evaluating the accuracies of the regressions). These aspects all contributed to the significant disagreements in the χ^2 values between the two routines.

Based on the evaluation of χ^2, the usage of the empirical function resulted in more accurate fits than the modified error function with both techniques. Following the same line of reasoning, the native Python function *scipy.optimize.curve_fit()* provided a better regression with the empirical fit function. To conclude, the most accurate value for the full-efficiency point was thus 380.89 GeV. Because of errors in the returned fit parameters, there were uncertainties in the full-efficiency points. However, estimations of those have been omitted in this thesis.

3.3 Applying p_T cuts on a simulated m_{jj} distribution

By using the Monte Carlo (MC) program PYTHIA 8.1 [12,32], two m_{jj} distributions from proton-proton collisions at $\sqrt{s} = 13$ TeV were simulated. In one of the distributions, a cut on the leading jet p_T was applied. The purpose was to limit the p_T range of the leading jet to where the trigger efficiency of the HLT_j360 was at least 99.5 %. This would have been done in a real analysis to remove the trigger bias. Since the full-efficiency point had been determined to 380.89 GeV in the previous subsections, the cut was set to 410 GeV to ensure a larger consistency with other ATLAS analyses. A p_T cut at 50 GeV was enforced on the subleading jet, in order to exclude events where the two leading jets were too unbalanced in the y- and z-axes. Additional kinematic cuts used by Nele Boelaert on the 2010 proton-proton collision data at $\sqrt{s} = 7$ TeV were also introduced in both distributions [4]. Consequently, the restrictions $|y^*| < 1.7$, $|y_B| < 0.75$ and $|\eta_{1,2}| < 2.8$ were utilized. The reasons for the cuts on $|y^*|$ and $|y_B|$ in that study were explained in subsection 2.3. By requiring $|\eta_{1,2}| < 2.8$, events in detector regions where the jet energy scale was not well understood were rejected [4].

The ratios between the number of events in the two distributions, as well as the corresponding error bars, were obtained with the ROOT function *TGraphAsymmErrors::Divide()*. In this method, the distribution with the implemented p_T cuts was correctly treated as a subset of the other distribution [33]. Furthermore, a confidence interval of 68.3 % was used, and the Beta parameters a and b for the prior were both set to 1 [33].

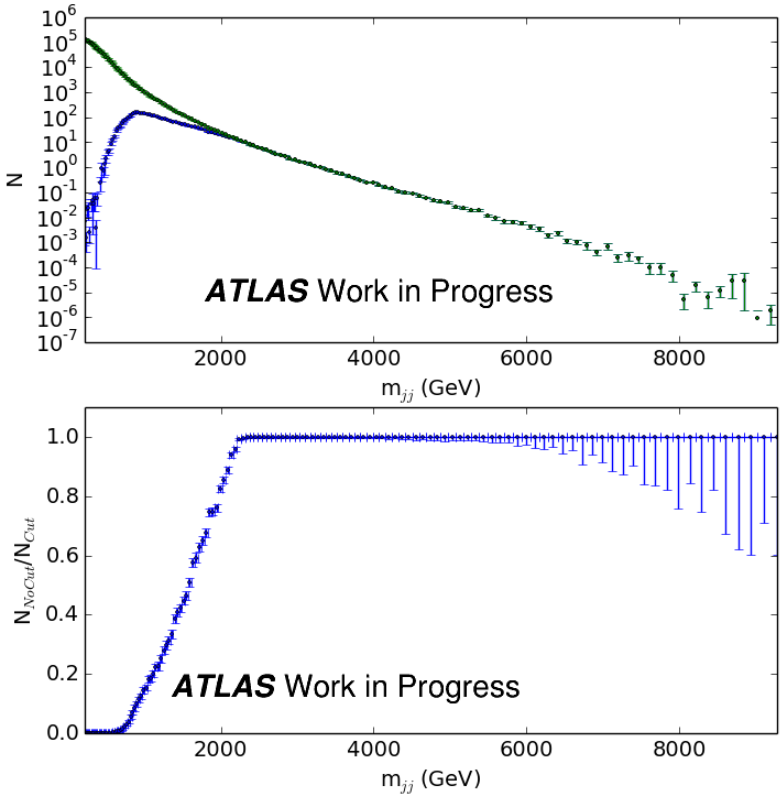

Figure 7: The two invariant mass distributions, with and without the enforced p_T cuts, have been plotted in the upper graph. Up to the point where the two distributions align, the topmost data points correspond to the distribution without the implemented p_T cuts. In the lower graph, the ratios N_{NoCut}/N_{Cut} between the number of events in the two distributions are displayed.

The m_{jj} value where the two distributions were equal within a precision of 0.16 % was determined to 2347 GeV. This was done by scanning the lower graph in fig. 7 from left to right with a for-loop, in order to find the first data point where the efficiency was 1. An analytical calculation of the m_{jj} value was also done for comparison, by inserting the kinematic cuts on $|y^*|$ and on the leading jet p_T into expression (2.10):

$$m_{jj,min} = (410 \text{ GeV}) * (e^{1.7} + e^{-1.7}) = 2320 \text{ GeV} \approx 2300 \text{ GeV}.$$

PYTHIA utilizes LO QCD scattering events with two particles in the initial and final states [12,32]. The program also includes showers of quarks and gluons in both the initial and final states, which realistically allows for events with more than two jets [12,32]. In the analytical calculation, the restriction to LO scatterings only meant that solely events with two jets were regarded [4,15]. Despite this, a good agreement between the results from the two methods was still obtained. On the other hand, that the two leading MC jets were required to be fairly well balanced in the y- and z-axes supports why the results should be somewhat similar. This is because the two leading jets were assumed to have equal p_T in the analytical calculation [4,15].

However, QCD scattering processes of higher orders were not included in the two methods [4,12,15,32]. Furthermore, the two leading jets might not be well balanced in the y- and z-axes in reality. These aspects added some uncertainties to the determined m_{jj} values. It was thus concluded that, for real data, the bias from the kinematic cuts

disappeared for m_{jj} above 2.5 TeV. In the next subsection, this would be used as a kinematic cut in the TADA F_χ distributions.

3.4 Implementing the m_{jj} cut in TADA

The 2015 results from proton-proton collisions at $\sqrt{s} = 13$ TeV at the LHC were studied with the TADA framework [6]. This procedure made use of several scripts already written by ATLAS collaborators, which incorporated both Python 2.7.4 and C++ 4.8.4 [6]. Some of the scripts had to be modified in order to successfully install and run the TADA scripts. Those changes may be found in Appendix A.

As mentioned in the introduction, the data has been omitted, although the selection cuts were applied to both the data and the MC samples representing the SM background. The samples were all generated by PYTHIA 8.1, and the TADA framework also simulated effects of the ATLAS detector [6]. Only background contributions from events with 3-6 jets were included, as the other contributions were negligibly small.

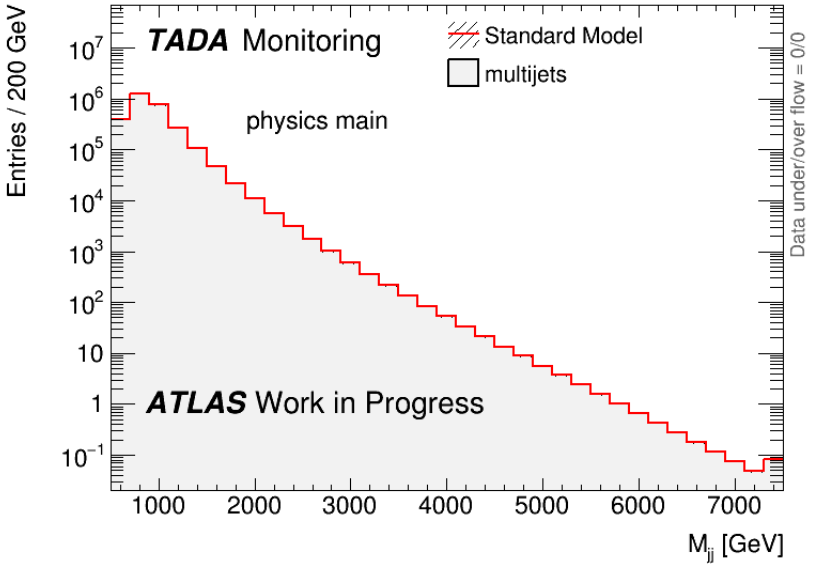

Figure 8: The invariant mass distribution, where the events were required to satisfy $p_{T1,T2} > 40$ GeV, $|\eta^*| < 0.6$ and $|\eta_{1,2}| < 1.7$. The cut $|\eta^*| < 0.6$ defined the central dijets as in subsection 2.3, although $|\eta^*|$ was used here instead of $|y^*|$. Furthermore, the cuts $|\eta_{1,2}| < 1.7$ meant that only events within the barrel region of the hadronic calorimeter were included [23].

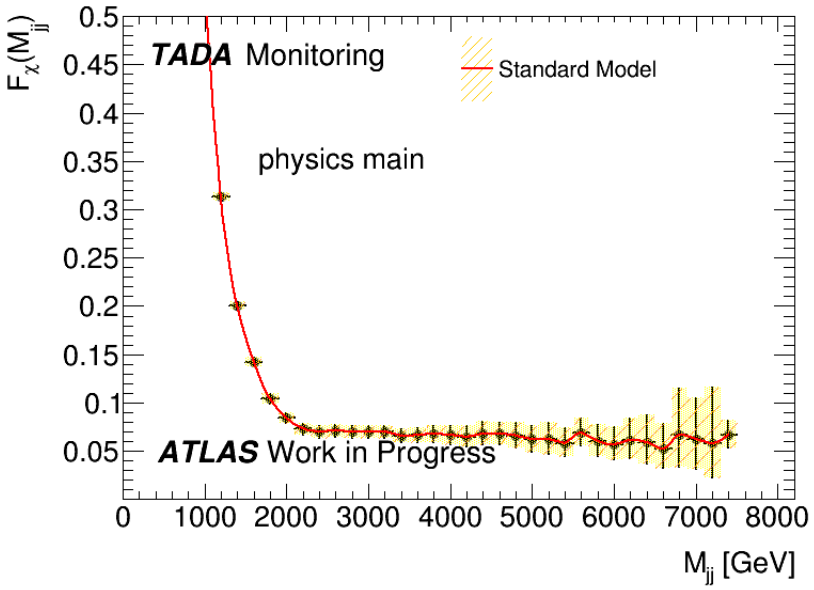

Figure 9: The F_χ distribution is shown, where F_χ was defined as $N(|\eta^*| < 0.6)/N(|\eta^*| < 1.7)$. Consequently, this definition of F_χ differed from the one in the introduction, since $|\eta^*|$ was used rather than $|y^*|$. Here, the selections $p_{T1,T2} > 40$ GeV and $|\eta^*| < 0.6$ were implemented on the events.

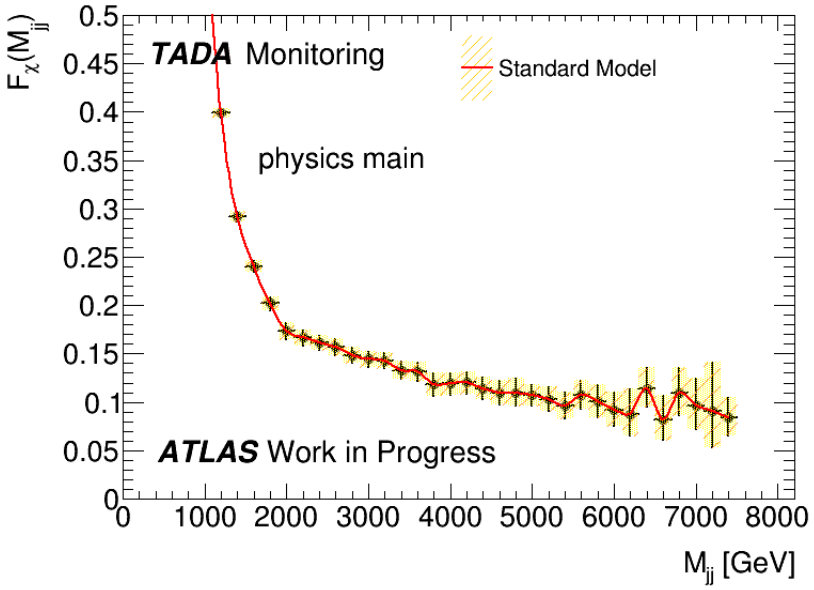

Figure 10: The F_χ distribution, where the events had to fulfil $p_{T1,T2} > 40$ GeV, $|\eta^*| < 0.6$ and $|\eta^*| < 1.7$, is displayed.

Since events with $|\eta^*| < 0.6$ were a subset of events with $|\eta^*| < 1.7$, the *TGraphAsymmErrors::Divide()* function was used to obtain the F_χ values and associated error bars. As in subsection 3.3, a confidence interval of 68.3 % was used, and the Beta parameters a and b for the prior were both set to 1. However, both of the F_χ distributions increased sharply for lower dijet invariant masses, which should not have occurred for SM physics [3, 31] (fig. 9,10).

In subsections 3.1 and 3.2, the trigger efficiency curve was fitted for the HLT_j360. A p_T cut of 410 GeV, where the trigger efficiency in the most accurate fit exceeded 99.5 %, was then applied on the leading jet (fig. 5). In subsection 3.3, it was concluded that this cut would introduce a bias for entries with $|y^*| < 1.7$ in the interval $m_{jj} < 2.5$ TeV. Even though the procedure in subsection 3.3 incorporated y^* rather than η^*, this dijet invariant mass range agreed well with the interval of the escalating regions (fig. 9,10). Therefore, the increases stemmed from the trigger not being close to fully efficient for events with $|\eta^*| < 1.7$ in this m_{jj} interval. In order to remove the trigger biases, only events also fulfilling $m_{jj} > 2.5$ TeV should be selected in the F_χ distributions (fig. 9, 10). Lastly, the increases in the F_χ distributions were sharp because the trigger efficiency dropped rapidly for jet p_T below the plateau region (fig. 5,9,10).

For events with $|y^*| < 0.6$, a 410 GeV p_T cut on the leading jet would only introduce a bias for events with $m_{jj} < 1$ TeV, as revealed by a calculation with expression (2.10). This dijet invariant mass range was found to the far left in the m_{jj} distribution (fig. 8). Despite the difference between y and η, the calculation hints at the obsoleteness in restricting the p_T range of the leading jet in fig. 8.

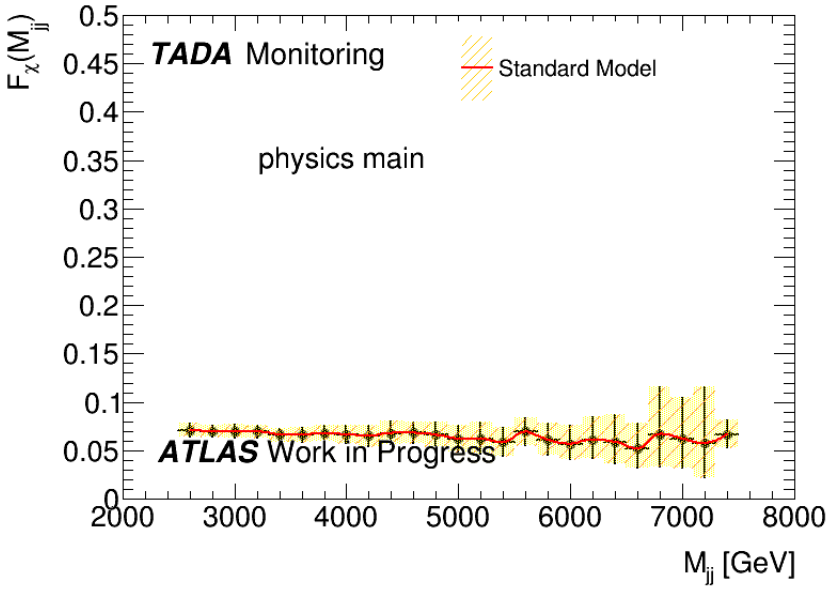

Figure 11: This F_χ distribution has been obtained with the selections $p_{T1,T2} > 40$ GeV and $|\eta^*| < 0.6$. However, only events which also fulfilled $m_{jj} > 2.5$ TeV were included.

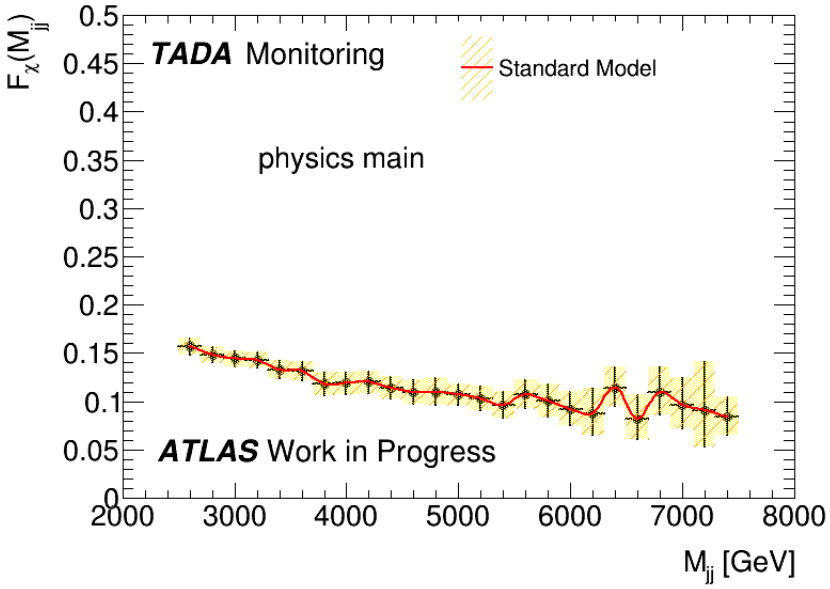

Figure 12: The F_χ distribution, which was obtained with the selections $p_{T1,T2} > 40$ GeV, $|\eta^*| < 0.6$ and $|\eta^*| < 1.7$, is displayed. Here, the distribution is also restricted to events where $m_{jj} > 2.5$ TeV.

As can be seen, the entire escalating region in fig. 11 has been removed, whereas there was still a small increase for lower dijet invariant masses in fig. 12. The latter was due to the additional selection cuts $|\eta_{1,2}| < 1.7$, which excluded certain events in the region $|\eta^*| < 1.7$, but no events within $|\eta^*| < 0.6$. As the m_{jj} decreases, the background dijet production rate increases [14, 28]. Therefore, the rejection of entries in the denominator of the F_χ increased in significance as the m_{jj} decreased in fig. 12.

4 Outlook

In this project, the TADA framework was used to study the 2015 data from proton-proton collisions at $\sqrt{s} = 13$ TeV. The TADA ran on a limited dataset, which allowed fragments of potential new physics signals to be seen shortly within data taking [6]. However, the F_χ distributions unexpectedly increased significantly for smaller dijet invariant masses. Here, the F_χ was defined as $N(|\eta^*| < 0.6)/N(|\eta^*| < 1.7)$.

Firstly, the trigger efficiency curve for the one-jet trigger HLT_j360, used in obtaining the 2015 data, was studied. Then, a p_T cut of 410 GeV, well into the plateau region of the curve, was enforced on the leading jet. Through an MC simulations in PYTHIA 8.1 [12, 32], two m_{jj} distributions were simulated, both with the restriction $|y^*| < 1.7$, but only one of them with the p_T cut. By comparing the two distributions, as well as utilizing the cuts on $|y^*|$ and p_T in an analytical calculation, it was determined that the two distributions became equal for $m_{jj} > 2.5$ TeV. The conclusion was that the trigger efficiency was insufficient for m_{jj} below 2.5 TeV, which introduced trigger biases in the F_χ distributions. By only selecting entries with $m_{jj} > 2.5$ TeV, large portions of the escalating regions were removed.

In upcoming versions of the TADA software, the restriction to events with $m_{jj} > 2.5$ TeV should be considered in order to avoid trigger biases in the F_χ distributions. Furthermore, the changes made to the TADA scripts could potentially be beneficial to

other LU students. They have therefore been proposed for the inclusion in an updated version of the README.

In this project, I have seen how data from the cylindrical ATLAS particle detector can be displayed in different distributions. These distributions incorporate the kinematic properties of jets, which are used to probe for hints of new physics. I have also learned about some programming languages used for producing the distributions. Additionally, it has been seen how technical limitations of the ATLAS detector, among others, motivate the application of kinematic cuts. The effects of these on some distributions have been studied in this project.

5 Appendix

A Modifications of TADA scripts

All the scripts mentioned here were found under the folder *PhysicsAnalysis/FastPhysMonitoring/FastPhysTagMon/trunk* at **https://svnweb.cern.ch/trac/atlasoff/browser**. The first change, in line 8 of *setup_athena.sh*, concerned the setup process of the TADA framework:

```
export RUCIO_ACCOUNT=???
```

Here, *???* should be replaced by the CERN account name of the user in question.

The batch queue for the jobs was also redefined as *1nh*. Thus, the changes were as follows in lines 17-21 of the script *_setup_.sh*:

```
# export TADAQUEUE='atlastzprod'
# export SHORTTADAQUEUE='atlastzprod'
# define default batch queue for TADA jobs
export TADAQUEUE='1nh'
export SHORTTADAQUEUE='1nh'
```

Accordingly, in lines 7-8 of *trunk/scripts/seqSubmit.py* the following alterations were made:

```
queue='1nh'
shortqueue='1nh'
```

The scripts used to produce the plots also had to be modified. Firstly, in lines 41-47 of *scripts/SFrameWrapper.py*, the function *fix()* was introduced. Consequently, illegal syntaxes to the Python command *xrdcp* were avoided, meaning the MC samples could successfully be imported. The changes were as follows:

```
def fix(arg):
    arg=arg.replace("|","")
    return(arg)

# copy input files
for f in TheSample.Files:
    f=fix(f)
```

Furthermore, in lines 259-266 of the same script the function *escape()* had to be created:

```
def escape(str):
    str = str.replace("'","&apos;")
    str = str.replace("<","&lt;")
    str = str.replace(">","&gt;")
    return str

for key in sel.keys() :
    f.write( "<Item Name=\"" + key + sid + "\" Value='" +
             escape(unicode(sel[key])) + "'/>\n" )
```

The purpose was to replace special characters in XML-files critical for the production of the raw histograms.

Finally, the following changes were made in lines 64-71 of the script *scripts/mergeSamplesBatch.py*:

```
#from ROOT import *
#turn warnings back on
#gErrorIgnoreLevel=kInfo


import ROOT
gROOT.SetBatch(True)
gROOT.Reset()
```

This was necessary in order to prevent a "segmentation fault" during the post-processing.

References

[1] Martin BR, Shaw G. Particle Physics-Third Edition. Chichester, United Kingdom: John Wiley & Sons Ltd; 2008.

[2] ATLAS Collaboration. Search for New Phenomena in Dijet Mass and Angular Distributions from pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. Phys Lett B [Internet]. 2016 [updated 2016 Mar 3; cited 2016 Feb 8];754:302-322. Available from: http://arxiv.org/abs/1512.01530 DOI: 10.1016/j.physletb.2016.01.032

[3] ATLAS Collaboration. Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS Detector. New J Phys [Internet]. 2011 [cited 2016 Feb 5];13(5):053044. Available from: http://arxiv.org/abs/1103.3864 DOI: 10.1088/1367-2630/13/5/053044

[4] Boelaert N. Dijet angular distributions in proton-proton collisions: At $\sqrt{s} = 7$ TeV and $\sqrt{s} = 14$ TeV. Lund, Sweden: Lund University; 2010.

[5] Horvath A. The LHC experiments and the preaccelerators [print]. 2006. 1 print: color, 744 x 524 pix. Link to license: http://creativecommons.org /licenses/by-sa/4.0/legalcode

[6] Köneke K. ATLAS Fast Physics Monitoring. In: Lepton Photon 2011; 2011 Aug 22-27; Mumbai, India. Available from: https://elsing.web.cern.ch/elsing/material/ATL-COM-DAPR-2011-011.pdf.

[7] Narayan RH. The Conception of Atom in Greek and Indian Physics. In: Proceedings of the 17th International Conference on Vedanta; 2007 Sep 19-21; Oxford, United Kingdom. Available from: http://arxiv.org/abs/0712.0764.

[8] Kragh H. Rutherford, Radioactivity, and the Atomic Nucleus. 2012 Feb; Madrid, Spain. Available from: http://arxiv.org/abs/1202.0954.

[9] Falconer I. Corpuscles, Electrons and Cathode Rays: JJ Thomson and the 'Discovery of the Electron'. Br J Hist Sci [Internet]. 1987 [cited 2016 Mar 14];20(3):241-76. Available from: http://www.jstor.org/stable/pdf/4026358.pdf?acceptTC=true DOI: http://dx.doi.org/10.1017/S0007087400023955

[10] Riordan M. The Discovery of Quarks. Sci [Internet]. 1992 [cited 2016 Mar 14];256(5061):1287-93. Available from: http://inspirehep.net/record333489/files/slac-pub-5724.pdf DOI: 10.1126/science.256.5061.1287

[11] Plehn T. Lectures on LHC Physics. Berlin, Germany: Springer International Publishing; 2014.

[12] Mrenna S, Sjöstrand T, Skands PZ. PYTHIA 6.4 Physics and Manual. J High Energy Phys [Internet]. 2006 [updated 2006 May 12; cited 2016 Mar 2];5:026. Available from: http://arxiv.org/abs/hep-ph/0603175 DOI: 10.1088/1126-6708/2006/05/026

[13] Sjöstrand T, Skands PZ. Multiple Interactions and the Structure of Beam Remnants. J High Energy Phys [Internet]. 2004 [updated 2004 May 19; cited 2016 Mar 20];8(3):053. Available from: http://arxiv.org/pdf/hep-ph/0402078 DOI: http://dx.doi.org/10.1088/1126-6708/2004/03/053

[14] ATLAS Collaboration. Search for New Physics in the Dijet Mass Distribution using 1 fb^{-1} of pp Collision Data at $\sqrt{s} = 7$ TeV collected by the ATLAS Detector. Phys Lett B [Internet]. 2011 [updated 2011 Nov 22; cited 2016 Feb 12];708(1):37-54. Available from: http://arxiv.org/abs/1108.6311 DOI: 10.1016/j.physletb.2012.01.035

[15] Bryngemark L. Search for new Phenomena in Dijet Angular Distributions at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV. Lund, Sweden: Lund University; 2016.

[16] ATLAS Collaboration. Search for dark matter candidates and large extra dimensions in events with a jet and missing transverse momentum with the ATLAS detector. J High Energy Phys [Internet]. 2013 [updated 2013 Apr 13; cited 2016 Feb 7];4:1-51. Available from: http://arxiv.org/abs/1210.4491 DOI: 10.1007/JHEP04(2013)075

[17] ATLAS Collaboration. Search for Quantum Black Hole Production in High-Invariant-Mass Lepton+Jet Final States Using *pp* Collisions at $\sqrt{s} =$ 8 TeV and the ATLAS Detector. Phys Rev Lett [Internet]. 2014 [updated 2014 Mar 6; cited 2016 Feb 12];112(9):091804. Available from: http://arxiv.org/abs/1311.2006 DOI: 10.1103/PhysRevLett.112.091804

[18] Kane G, Pierce A. Perspectives on LHC Physics. Hackensack, N.J.: World Scientific Publishing Co.; 2008.

[19] DeViveiros PO. Search for Contact Interactions using Dijet Angular Distributions with the ATLAS Detector at the CERN Large Hadron Collider. Toronto, Canada: University of Toronto; 2011.

[20] [ATLAS experiment] [print]. Argonne National Laboratory; 2012. 1 print: color, 995 x 536 pix. Link to license: http://creativecommons.org/licenses /by-sa/4.0/legalcode

[21] ATLAS Collaboration. Measurement of dijet cross sections in *pp* collisions at 7 TeV centre-of-mass energy using the ATLAS detector. J High Energy Phys [Internet]. 2014 [updated 2014 May 28; cited 2016 Feb 7];5(59):1-67. Available from: http://arxiv.org/abs/1312.3524 DOI: 10.1007/JHEP05(2014)059

[22] CMS Collaboration. Measurement of Dijet Angular Distributions and Search for Quark Compositeness in *pp* Collisions at $\sqrt{s} = 7$ TeV. Phys Rev Lett [Internet]. 2011 [cited 2016 Mar 4];106(20):201804. Available from: http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.106.201804 DOI: http://dx.doi.org/10.1103/PhysRevLett.106.201804

[23] ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Eur Phys J C Part Fields [Internet]. 2015 [updated 2015 Jan 28; cited 2016 Feb 7];75(1):1-101. Available from: http://arxiv.org/abs/1406.0076 DOI: 10.1140/epjc/s10052-014-3190-y

[24] ATLAS Collaboration. Search for contact interactions and large extra dimensions in the dilepton channel using proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. Eur Phys J C [Internet]. 2014 [updated 2014 Dec 16; cited 2016 Feb 7];74(12):3134. Available from: http://arxiv.org/abs/1407.2410 DOI: 10.1140/epjc/s10052-014-3134-6

[25] ATLAS Collaboration. Search for New Phenomena in Dijet Angular Distributions in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV Measured with the ATLAS Detector. Phys Rev Lett [Internet]. 2015 [updated 2015 Jun 18; cited 2016 Feb 7];114(22):221802. Available from: http://arxiv.org/abs/1504.00357 DOI: 10.1103/PhysRevLett.114.221802

[26] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. Eur Phys J C Part Fields [Internet]. 2016 [cited 2016 Mar 14]. Available from: http://arxiv.org/abs/1603.02934

[27] Abbott B, Blair R, Crone G, Green B, Love J, Proudfoot J, Rifki O, Vazquez WP, Vandelli W, Zhang J. The evolution of the region of interest builder for the ATLAS experiment at CERN. J Instrum [Internet]. 2016 [cited 2016 Mar 21];11(2):02080. Available from: http://iopscience.iop.org/article/10.1088/1748-0221/11/02/C02080 DOI: http://dx.doi.org/10.1088/1748-0221/11/02/C02080

[28] ATLAS Collaboration. Search for new phenomena in the dijet mass distribution using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector. Phys Rev D Part Fields [Internet]. 2015 [cited 2016 Mar 24];91(5):052007. Available from: http://journals.aps.org/prd/abstract/10.1103/PhysRevD.91.052007 DOI: http://dx.doi.org/10.1103/PhysRevD.91.052007

[29] Choudalakis G. Dijet Searches for New Physics in ATLAS. In: Proceedings of the DPF-2011 Conference; 2011 Aug 8-13; Providence, R.I.. Available from: http://arxiv.org/abs/1109.2144.

[30] Choudalakis G. On hypothesis testing, trials factor, hypertests and the BumpHunter. In: PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding; 2011 Jan 17-20; Geneva, Switzerland. Available from: http://arxiv.org/abs/1101.0390.

[31] ATLAS Collaboration. ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV. J High Energy Phys [Internet]. 2013 [updated 2012 Nov 9; cited 2016 Mar 21];1:1-46. Available from: http://arxiv.org/abs/1210.1718 DOI: 10.1007/JHEP01(2013)029

[32] Mrenna S, Sjöstrand T, Skands PZ. A brief introduction to PYTHIA 8.1. Comput Phys Comm [Internet]. 2015 [cited 2016 Mar 21];178(11):852-867. Available from: http://arxiv.org/abs/0710.3820 DOI: 10.1016/j.cpc.2008.01.036

[33] Casadei D. Estimating the selection efficiency. J Instrum [Internet]. 2012 Aug [updated 2012 Jul 25; cited 2016 Feb 29];7(8):08021. Available from: http://arxiv.org/abs/0908.0130 DOI: 10.1088/1748-0221/7/08/P08021