

Search for Resonant Leptoquark Production Coupling to an Electron and a b-Quark using Run 2 pp Collision Data of the ATLAS Experiment

von

Christopher Maximilian Engel

Masterarbeit in Physik
vorgelegt dem Fachbereich Physik, Mathematik und Informatik (FB 08)
der Johannes Gutenberg-Universität Mainz
am 11. Oktober 2024

1. Gutachter: Prof. Dr. Stefan Tapprogge
2. Gutachter: Prof. Dr. Achim Denig

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel (dazu zählen auch KI-basierte Anwendungen oder Werkzeuge) benutzt habe. Sämtliche wörtlichen oder sinngemäßen Übernahmen und Zitate sind kenntlich gemacht und nachgewiesen. Ich versichere, dass ich keine Hilfsmittel verwendet habe, deren Nutzung die Prüferin oder der Prüfer explizit ausgeschlossen hat.

Im Anhang A.3 („Used KI-Tools“) habe ich die verwendeten KI-Tools dokumentiert. Mit Abgabe der vorliegenden Leistung übernehme ich die Verantwortung für das eingereichte Gesamtprodukt. Ich verantworte damit auch jegliche KI-generierten Inhalte, die ich in meine Arbeit übernommen habe. Die Richtigkeit übernommener (KI-generierter) Aussagen und Inhalte habe ich nach bestem Wissen und Gewissen geprüft.

Ich habe die Arbeit nicht zum Erwerb eines anderen Leistungsnachweises in gleicher oder ähnlicher Form eingereicht.

Mir ist bekannt, dass ein Verstoß gegen die genannten Punkte prüfungsrechtliche Konsequenzen hat und insbesondere dazu führen kann, dass die Studien- und Prüfungsleistung als mit „nicht bestanden“ bewertet wird. Die Einschreibung kann für bis zu zwei Jahre widerrufen werden, wenn Studierende zweimal oder häufiger bei Prüfungsleistungen täuschen (§ 69 Abs. 4 und 5 HochSchG).

Mainz, den 11.10.2024

Christopher Maximilian Engel
ETAP
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität D-55099 Mainz
chengel@students.uni-mainz.de

Abstract

This master thesis sets an improved expected exclusion limit on a scalar $\tilde{S}_1$ Leptoquark model solely coupling to an electron and a b -quark as part of an ongoing research by ATLAS. The exclusion limit is gained by searching for a resonant Leptoquark production using full Run 2 data of the ATLAS experiment which corresponds to a proton-proton collision with a center-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 140 fb^{-1} . Additionally, the simulation of the signal model uses a lepton parton distribution function and has a next-to-leading-order signal precision. This next-to-leading-order precision leads to the production of final states with one and two leptons motivating a search in both final states. After optimizing the exclusion significance in both cases separately, they are recombined in a statistical analysis. The results of this thesis leads to the conclusion that the ongoing resonant search is capable of excluding new regions. These new expected exclusion limits range from Leptoquarks with masses of 1.8 TeV and a coupling λ of 1.2 to Leptoquarks with masses of 2.6 TeV and a coupling λ of 3.5 at a 95% confidence limit.

Zusammenfassung

Diese Masterarbeit setzt eine verbesserte erwartete Ausschlussgrenze auf ein skalares $\tilde{S}_1$ Leptoquark Modell, welches nur an ein Elektron und ein b -quark koppelt, als Teil einer laufenden Forschung von ATLAS. Diese Ausschlussgrenze wird erreicht, indem nach der resonanten Produktion eines Leptoquarks mit den gesamten Run 2 Daten des ATLAS Experiments, welche einer Proton-Proton Kollision mit einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV und einer integrierten Luminosität von 140 fb^{-1} entsprechen, gesucht wird. Zusätzlich benutzt die Simulation des Signalmodells eine Lepton Partondichtefunktion und hat eine "next-to-leading order" Signalpräzision. Diese "next-to-leading order" Präzision führt zu der Produktion von Endzuständen mit einem oder zwei Leptonen, was zu einer Suche in beiden Endzuständen motiviert. Nachdem die Ausschlusssignifikanzen in beiden Fällen optimiert wurden, werden sie wieder in der statistischen Analyse kombiniert. Die Ergebnisse dieser Masterarbeit führen zu der Schlussfolgerung, dass die laufende Suche nach resonant produzierten Leptoquarks in der Lage ist, neue Regionen auszuschließen. Diese neuen erwarteten Ausschlussgrenzen reichen von Leptoquarks mit Massen von 1.8 TeV und einer Kopplung λ von 1.2 bis zu Leptoquarks mit Massen von 2.6 TeV und einer Kopplung λ von 3.5 mit einer 95%igen Konfidenzgrenze.

Contents

Abstract	iii
Zusammenfassung	iv
1. Introduction	1
2. Theoretical Background	3
2.1. Standard Model of Physics	3
2.1.1. Elementary Particles	3
2.1.2. Mathematical Formulation	6
2.2. Proton-Proton Collisions	11
2.2.1. Conceptual Quantities	12
2.2.2. Feynman Formalism	13
2.2.3. Parton Distribution Functions	15
2.2.4. Proton-Proton Interaction	18
2.3. Going Beyond the Standard Model	20
2.3.1. Limits of the Standard Model	20
2.3.2. Leptoquarks	22
3. ATLAS Experiment	27
3.1. Large Hadron Collider	27
3.2. Coordinate System and Common Kinematic Variables	29
3.3. Inner Detector	31
3.4. Calorimeter System	33
3.5. Muon Spectrometer	36
3.6. Trigger and Data Acquisition	37
3.7. Object Reconstruction and Identification	38
3.7.1. Electrons	38
3.7.2. Muons	39
3.7.3. Jets	40
3.7.4. b-Tagging	41
3.7.5. Missing Transverse Energy	42
3.7.6. Overlap Removal	42
3.8. Monte Carlo Method	43
3.9. Data Formats	45
4. Data and Monte Carlo Samples	46
4.1. Collision Data	46

Contents

4.2. Simulated Signal	48
4.3. Simulated Background	50
4.4. Fake Electron Estimation	54
4.5. Analysis Software	55
5. Event Selection	57
5.1. Object Selection	57
5.1.1. Electron Selection	57
5.1.2. Muon Selection	58
5.1.3. Jet Selection	58
5.2. Preselection	59
5.3. Kinematic Distributions	61
5.3.1. Leading Lepton	62
5.3.2. Leading Jet	65
5.3.3. Lepton-Jet System	70
5.4. Cuts	75
5.4.1. Number of Leptons	75
5.4.2. Lepton Flavour	76
5.4.3. Number of b-jets	78
5.4.4. Significance of Missing Transverse Energy	78
5.4.5. Mass of Lepton-Lepton system	81
5.4.6. Ratio of Transverse Momentum of Leading Lepton and Mass of Lepton-Jet System	82
5.4.7. Distance Between Leading Lepton and Jet in Collider Coordinates	82
5.5. Region Definitions	86
5.6. Cut Efficiencies	94
5.7. Dijet Sample Performance	95
6. Statistical Analysis	103
6.1. Statistical Framework	103
6.2. Systematic Uncertainty	107
7. Results	109
7.1. Fit Results	109
7.2. Expected Exclusion Limit	129
8. Conclusion and Outlook	136
A. Appendix	138
A.1. Tables and Figures	138
A.1.1. Subleading Lepton Kinematics	138
A.2. Used KI-Tools	140
B. Bibliography	141

1. Introduction

Physics is a very old field of science originating from the ancient urge to understand the world around oneself and the motivation to accomplish previously impossible tasks using analytical thinking with the second one building the bridge to inventions and engineering. This dates back more than two thousand years. A prime example at around 250 BC would be Archimedes which, among many other things, introduced the fundamentals of statics and the Archimedes' Principle describing the displacement of liquids and the accompanying force letting objects flow.

The understanding of physical processes is ever expanding and evolving with every new theory being based on the whole knowledge of previous physicists and conclusions while generalizing the observations. The probably most famous example for this is the fundamental understanding of classical mechanics and gravity from Isaac Newton in the late 17th century which describes the motion of objects moving at speeds commonly found at scales accessible by humans at this time with incredible precision. Classical mechanics was the leading theory to describe motion and gravity for over 200 years and is still commonly used to date due to its simplicity. But in 1905 and 1915, Albert Einstein introduced special relativity [1] and general relativity [2] respectively, proving that Newtons gravity concept is an effective theory only applicable at low energies. When going to higher energies, becoming relevant at scales some orders of magnitude smaller than the speed of light or masses at the scale of planetary masses, relativistic mechanics have to be applied to describe the system correctly. This is not contradicting Newtons observations as relativistic mechanics asymptotically lead into classical mechanics when going to lower energies. Without the discoveries of Newton, special and general relativity could not have been developed being closely related and intertwined.

The same most likely holds true for the Standard Model of particle physics being a great theory explaining the fundamental particles known today with unprecedented precision, even predicting the Higgs boson which was only discovered in 2012 [3, 4]. But the Standard Model does not explain everything with many questions left open. Some of these open questions are the asymmetry between matter and anti-matter in the universe, the absence of a dark matter candidate and a possible unification of the fundamental forces. This motivates to the creation of extended physical frameworks which would be called Grand Unified Theory in the case of the unification of the strong force, the weak force and the electromagnetic force. There are many different Grand Unified Theories proposed but all of them take the Standard Model as baseline and lead asymptotically into the Standard Model when going to lower energies.

To discover the boundaries of physics, particle accelerators have been developed and build for almost 100 years. The largest and highest energetic accelerator to date is

1. Introduction

the Large Hadron Collider [5] at CERN near Geneva, Switzerland where experiments such as ATLAS [6] and CMS [7] are located. It is a proton-proton collider with a center-of-mass energy of $\sqrt{s} = 13$ TeV during Run 2 which is the data used in this thesis. These scales of energies allows to uncover the limits of the Standard Model by e.g. searching for new particles. One of these particles is the Leptoquark which is a hypothetical boson carrying both a lepton and baryon number. It is predicted by many Standard Model extensions, especially Grand Unified Theories, directly connecting leptons and quarks. A discovery of such a particle would pave the way for Grand Unified Theories.

This thesis is part of an active ATLAS group searching for the resonant production of a Leptoquark using a scalar $\tilde{S}_1$ Leptoquark model with an electric charge of $Q = -4/3$. The search includes Leptoquarks coupling to different charged leptons and quarks and this thesis uses Leptoquarks only coupling to an electron and a b -quark. For this, a next-to-leading-order (NLO) signal precision is used which includes a lepton parton distribution function necessary for the resonant production. A next-to-leading-order signal precision for the resonant production of Leptoquarks results in two leptons in the final state around 30% of the time which motivates a search in final states with 1 electron and 2 electrons with each having an additional b -jet originating from the b -quark.

To cover the theoretical knowledge needed for this thesis chapter 2 gives a rough overview of the Standard Model both from a more general and a more mathematical perspective, explains proton-proton collisions in more detail and talks about physics beyond the Standard Model including properties and current constraints on Leptoquarks. Chapter 3 provides information on the ATLAS experiment, its detector components, how the measured data is used to reconstruct them into physical objects and how to simulate events creating an elaborated prediction of the measured data. Then, specifics on the collision data and the simulated data and how they are created is elaborated on in chapter 4. After having produced simulated backgrounds and signals, the following chapters describe the main focus of this thesis. It starts with chapter 5 describing the selected objects and some basic preselection cuts to reduce the amount of data having to be processed. After that, the kinematic distributions are investigated and understood to explore the behaviour of the final state. Then, cuts are identified and optimized which maximize background rejection while also keeping a good signal efficiency. Additionally, signal regions, control regions and validation regions are defined each serving different purposes. Signal regions maximize the sensitivity to find a Leptoquark, control regions are used to correct for mismodeling of Monte Carlo Simulations by fitting the simulation to data and validation region ensure that the corrections gained from the control regions can be applied to signal regions. After having an optimal sensitivity to find deviations from the Standard Model, chapter 6 talks about the inner workings of the statistical analysis and the systematic uncertainties included in this thesis. Afterwards, the results of the statistical analysis are provided in chapter 7 including tests of the results and expected exclusion limits on Leptoquark models. Finally, conclusions, suggestions for the next steps and ideas to improve the results of this thesis are provided in chapter 8.

2. Theoretical Background

This chapter will tackle the theoretical basics necessary to understand the motivation and physical conclusions made in this thesis. First, the Standard Model of particle physics will be presented by introducing the elementary particles and explaining the basic mathematical structure of quantum field theory. Using the Standard Model the phenomenology of proton-proton collisions are explained including the calculation of Feynman diagrams. After that, open questions within the Standard Model are raised and possible theories that go beyond the Standard Model are discussed. Finally, the theory behind the Leptoquarks is given including the current exclusion limits.

2.1. Standard Model of Physics

The Standard Model (SM) unifies a lot of physical observations by providing a mathematical description of three of the four fundamental forces with gravity being not included. First the fundamental particles including the mediators of the forces are described before going into further detail on the mathematical formulations describing these particles. This section is inspired by [8] which is in turn based on the textbooks [9, 10, 11] which were also used independently. Throughout this entire thesis natural units are used which sets the reduced Planck constant [12], defined as $\hbar = \frac{h}{2\pi}$, and the speed of light c equal to one: $\hbar = c = 1$. This leads to quantities such as masses, momenta and energies sharing the same units of energy which are in particle physics usually expressed in terms of "electron Volt" eV, where 1 eV is the energy necessary for an electron to move through a difference in electric potential of 1 Volt and e is the absolute electric charge of an electron. The original units can be restored by multiplying with the correct amount of $\hbar$ and c , e.g. masses can be obtained by dividing by c^2 .

2.1.1. Elementary Particles

The Standard Model of Physics is based on elementary point-like particles which can be grouped into two categories: bosons and fermions. Bosons have an integer spin, whereas fermions have a half-integer spin ($1/2, 3/2, \dots$ etc.) and follow the Bose-Einstein [13] and the Fermi-Dirac statistics [14] respectively. This results in different macroscopic behaviours. Most notably the Pauli exclusion principle [15] which holds for fermions but not for bosons and states that two fermions cannot have an identical quantum state. There are a total of 12 fundamental fermions which all have a spin of $1/2$ and five fundamental bosons. Fermions are the constituents of matter and can interact via one of the three fundamental forces: the electromagnetic

2. Theoretical Background

force, the weak force and the strong force. These forces are mediated with spin-1 gauge bosons, photons γ , the W^+ and W^- boson, the Z boson and gluons. The last fundamental particle is the most recently discovered spin-0 Higgs boson [3, 4], named after Peter Higgs.

The fermions can be further split into two groups: the leptons and the quarks. Leptons can be either charged or neutral and can therefore interact electromagnetically and weakly or only weakly. The three charged leptons have an electric charge of $Q = -1e$ and are the electron, the muon and the tau with each belonging to its own generation and being heavier than the previous one. Their charge neutral partners are the neutrinos with each charged lepton having its own associated neutrino, the electron neutrino ν_e , the muon neutrino ν_μ , and the tau neutrino ν_τ . Neutrinos are considered massless within the Standard Model although observed neutrino oscillations [16, 17, 18] require massive neutrinos to exist. Yet no neutrino mass has been measured so far and only upper limits were possible to be set. Additionally, the ordering of the neutrino masses is not established and only the absolute value of the differences between them has been measured so far.

Each lepton has its own antiparticle which has the same properties, including the same mass, of the corresponding lepton but is charge conjugated. This means in the case of the leptons having the opposite electric charge. Antiparticles are denoted by a bar over the particle symbol and are usually named after the particle with the prefix "anti-", e.g. the antiparticle of a muon μ is an antimuon $\bar{\mu}$. The only exception is the antiparticle of the electron e , the positron $\bar{e}$. A summary of the information for the leptons is given in table 2.1.

Table 2.1.: Overview of lepton properties. Values marked with an asterisk are considered this way in the SM but experiments only give upper or lower limits on these values. All values are taken out of the PDG (particle data group) [19].

Feature	Electron	Electron neutrino	Muon	Muon neutrino	Tau	Tau neutrino
Generation	1		2		3	
Symbol	e	ν_e	μ	ν_μ	τ	ν_τ
Mass [MeV]	0.511	0 *	105.66	0 *	$1.78 \cdot 10^3$	0 *
Mean lifetime [s]	stable *	stable *	$2.20 \cdot 10^{-6}$	stable *	$2.90 \cdot 10^{-13}$	stable *
Electric charge [e]	-1	0	-1	0	-1	0

The second group of fermions are the quarks which are charged particles with an electric charge as well as a color charge which lets them interact electromagnetically, weakly or strongly. Up-type quarks have an electric charge of $Q = 2/3e$ and down-type quarks have a charge of $Q = -1/3e$. The possible color charges are either red, blue or green with antiquarks carrying the respective anti-colors anti-red, anti-blue or anti-green. Each type of quarks has three generations giving a total of 6 quarks: up, down, charm, strange, top, bottom ordered by generation with the

2. Theoretical Background

up-type quark being the first one named per generation. The mass of the particles is the only property changing between generations. Due to confinement [20, 21, 22] color charges cannot exist freely which forces quarks to create hadrons, color neutral bound states. If two quarks create a bound state it is called a meson such as a pion and if three quarks create a bound state it is called a baryon. The most common and well known baryons are the proton and the neutron. Table 2.2 gives a summary overview of the quarks and its properties.

Table 2.2.: Overview of quark properties. All values are taken out of the PDG [19].

Feature	Up	Down	Charm	Strange	Top	Bottom
Generation	1		2		3	
Symbol	u	d	c	s	t	b
Mass [MeV]	2.16	4.70	$1.27 \cdot 10^3$	93.5	$172.6 \cdot 10^3$	$4.18 \cdot 10^3$
Electric charge [e]	$2/3$	$-1/3$	$2/3$	$-1/3$	$2/3$	$-1/3$

The fundamental bosons consist of the force mediator gauge bosons with spin-1 and the Higgs boson with spin-0. The boson mediating the electromagnetic interaction is the massless photon γ . It couples to electrically charged particles including charged leptons, quarks and the charged W bosons. Due to it being massless the electromagnetic field has an infinite range yet its field strength decreases quadratically.

The weak force is mediated with the three gauge bosons W^+ , W^- , and Z with the first two carrying electric charge. Unlike the photon these bosons are massive which gives the weak interaction a finite range. The corresponding charge for the coupling is the weak isospin I_3 . Due to only left-handed (and right-handed anti-) particles having a non zero weak isospin the weak force couples only to these particles. A prominent example of weak interaction is the beta decay.

Gluons g are massless, couple to color charged particles and mediate the strong force. They carry color charges themselves always having a color and an anti-color. There are three colors in total which lead to eight possible combinations of color charges a gluon can have. The combinations can be described using the Gell-Mann matrices [23]. Gluons having color charges allow for gluon self-coupling unlike photons which are not electrically charged. This self-coupling limits the range of the strong force making the interaction range not infinite although gluons are massless. The strong force explains the stability of a nucleus despite same charged protons being contained. A summary of the gauge bosons is given in table 2.3.

The Higgs boson H is a result of the spontaneous symmetry breaking which lead to the Higgs mechanism, already postulated in 1964 [24, 25, 26]. For more information on spontaneous symmetry breaking and the Higgs mechanism see section 2.1.2. The Higgs boson has a mass of 125.2 GeV [19] and is the excitation of the Higgs field. By coupling to the Higgs field, particles gain their mass.

2. Theoretical Background

Table 2.3.: Overview of gauge boson properties. All values are taken out of the PDG [19].

Feature	Photon	$W^\pm$ boson	Z boson	Gluon
Symbol	γ	$W^\pm$	Z	g
Mass [GeV]	0	80.37	91.19	0
Force	electromagnetic	weak	weak	strong
Range [m]	∞	10^{-18}	10^{-18}	10^{-15}

2.1.2. Mathematical Formulation

The theory of the Standard Model is a non-abelian gauge theory that combines field theory and quantum mechanics including, but not being limited to, the Schrödinger equation [27, 28], the Dirac equation [29, 30] and the uncertainty principle [31]. Additionally, special relativity [1] is also included in the Standard Model. The theory itself is a type of quantum field theory (QFT) and is defined by the local gauge symmetry group

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (2.1)$$

$SU(n)$ is the special unitary group and $U(n)$ is the unitary group with n degrees. $SU(3)_C$ is the group defining the space of color charge which is described by quantum chromodynamics (QCD) and $SU(2)_L \times U(1)_Y$ is the symmetry group described by electroweak theory (EW). The group $SU(2)_L$ defines the space of left-handed particles and $U(1)_Y$ defines the space of hypercharge. Due to spontaneous symmetry breaking of the EW symmetry the group $U(1)_{EM}$ is created being described by quantum electrodynamics (QED).

In order to describe fields one needs to define the Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ which depends on a field ϕ and its derivative $\partial_\mu \phi = \frac{\partial \phi}{\partial x^\mu}$ in all four space-time dimensions with $x^\mu = (t, x, y, z)$ being the 4-vector of the position. From now on the Lagrangian density will be referred to as Lagrangian and the Einstein summation convention is used which means that an index is summed over if it appears twice so e.g.:

$$\partial_\mu p^\mu = \sum_\mu \partial_\mu p^\mu = \partial_0 p^0 + \partial_1 p^1 + \partial_2 p^2 + \partial_3 p^3. \quad (2.2)$$

Additionally, superscript and subscript indices state whether vectors are covariant or contravariant. Throughout this thesis the Minkowski metric is used which replaces the contraction matrix with a 4x4 matrix having zero entries apart from the diagonal $(-1, 1, 1, 1)$. The contraction matrix used in equation 2.2 was the identity matrix. Using the Minkowski metric this now results in:

$$\partial_\mu p^\mu = -\partial_0 p^0 + \partial_1 p^1 + \partial_2 p^2 + \partial_3 p^3. \quad (2.3)$$

2. Theoretical Background

The Lagrangian is defined to be Lorentz invariant and the action S is defined as:

$$S = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x, \quad (2.4)$$

where d^4x denotes the four-dimensional integration. The principle of least action [32] requires the Lagrangian to be resistant to small variations of S and therefore needs to fulfill the equation $\delta S = 0$ in order for the Lagrangian to describe equations of motion. Requiring this results in the Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0. \quad (2.5)$$

According to Noether's Theorem [33] for every continuous symmetry of the action with Euler-Lagrange equations being fulfilled there is an associated conserved current.

Quantum electrodynamics (QED) is based on the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \quad (2.6)$$

which describes a spin-1/2 fermion field with mass m using the spinor field ψ and γ^μ are the Dirac matrices [29]. The Dirac equation follows from the equations of motion which means the Dirac Lagrangian can be determined to be

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad (2.7)$$

with the Dirac adjoint $\bar{\psi} = \psi^\dagger \gamma^0$ and $\psi^\dagger$ being the Hermitian adjoint. The Dirac Lagrangian is invariant under global phase transformations

$$\psi \rightarrow e^{i\theta} \psi, \text{ with } \theta \in \mathbb{R} \quad (2.8)$$

because the Dirac adjoint gets factor $e^{-i\theta}$ which cancels when combining $\bar{\psi}\psi$. This Lagrangian is however not invariant under local phase transformations

$$\psi(x) \rightarrow e^{i\theta(x)} \psi(x), \text{ with } \theta \in \mathbb{R} \quad (2.9)$$

giving the Lagrangian an additional term

$$\mathcal{L} \rightarrow \mathcal{L} - (\partial_\mu \theta) \bar{\psi} \gamma^\mu \psi. \quad (2.10)$$

This term needs to be canceled which is done by introducing the vector boson field A_μ that transforms under local gauge transformations as

$$A_\mu \rightarrow A_\mu + \frac{1}{e} \partial_\mu \theta, \quad (2.11)$$

where e is the elementary charge. This results in the new Lagrangian

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \quad (2.12)$$

2. Theoretical Background

with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ being the electromagnetic tensor originally used in the Maxwell equations and introducing the covariant derivative $D_\mu \equiv \partial_\mu + ieA_\mu$. Expanding the covariant derivative gives the expression

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - (e\bar{\psi}\gamma^\mu\psi)A_\mu, \quad (2.13)$$

where every term describes different parts of QED. The first term describes the fermion field, the second term describes the vector boson field and the last term describes the interaction of both. Note that the second term is required to have no mass term for the Lagrangian to be invariant and the third term containing two spinor and one vector field suggesting that two fermions always couple to one boson in QED.

Quantum chromodynamics (QCD) is based on the $SU(3)$ group and the derivation is similar to the one described for QED but it starts with a Lagrangian from a set of fermions with three color charges

$$\mathcal{L} = \bar{\psi}_j(i\gamma^\mu \partial_\mu - m)\psi_j, \text{ with } j \in \{r, g, b\}. \quad (2.14)$$

This finally leads to the Lagrangian

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}_i(i\gamma^\mu D_{\mu;ij} - m\delta_{ij})\psi_j - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}, \quad (2.15)$$

where the covariant derivative is defined as $D_{\mu;ij} = \partial_\mu \delta_{ij} - igT_{a;ij}A_\mu^a$ and the gluon field strength tensor is defined as $G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_{b;\mu}A_{c;\nu}$. The Kronecker-Delta δ_{ij} is defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \quad (2.16)$$

the indices are $a, b, c \in \{1, \dots, 8\}$ describing the different gluon fields and $i, j \in \{r, g, b\}$ describing the color charges. The structure constants of the $SU(3)$ are given by f^{abc} , g is the coupling strength and T^a describes the generators of the underlying symmetry group, here being given by the eight Gell-Mann matrices λ^a with $T^a = \frac{1}{2}\lambda^a$.

Electroweak theory (EW) is the unification of the electromagnetic and weak force and was postulated by Glashow [34] and independently confirmed by Weinberg [35] and Salam [36]. Being defined by the groups $SU(2) \times U(1)_Y$ EW has four vector bosons, three from $SU(2)$ and one from $U(1)_Y$, denoted as W_μ^1 , W_μ^2 , W_μ^3 and B_μ respectively. As briefly mentioned before, $SU(2)$ creates the left-handed space which is the space of particles with antiparallel spin and momentum vectors whereas right-handed particles have parallel spin and momentum vectors. All fundamental fermions can be left-handed and right-handed apart from neutrinos which are always left-handed due to them being massless and therefore moving at the speed of light.

2. Theoretical Background

W_μ^i with $i \in \{1, 2, 3\}$ couple to the weak isospin I_3 and B_μ couples to the hypercharge Y . The final EW Lagrangian is

$$\mathcal{L}_{\text{EW}} = \bar{\psi}_k^R i\gamma^\mu D_\mu \psi_k^R + \bar{\psi}_j^L i\gamma^\mu D_\mu \psi_j^L - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (2.17)$$

with the first two terms describing the fermion fields and the last two terms the gauge boson fields. ψ_k^R are the right-handed fermion singlets with k being the types of particles (no neutrinos) and ψ_j^L are the left-handed fermion doublets with j being all quark and lepton generations. The field strength tensors of the gauge bosons are defined as:

$$W_a^{\mu\nu} = \partial^\mu W_a^\nu - \partial^\nu W_a^\mu + g\epsilon^{abc} W_\mu^b W_\nu^c, \quad (2.18)$$

$$B^{\mu\nu} = \partial^\mu W^\nu - \partial^\nu W^\mu. \quad (2.19)$$

The covariant derivative is $D_\mu = \partial_\mu - ig\tau_a W_\mu^a + ig'\frac{Y}{2}B_\mu$ with g and g' being the $SU(2)$ and $U(1)$ coupling constants respectively. The generators τ_a are the generators of $SU(2)$ being described by the Pauli-Matrices $\tau_a = \frac{\sigma_a}{2}$. All four introduced vector bosons here are massless due to them having no mass terms in order for the Lagrangian to stay invariant under local gauge transformations and do not correspond to the physical vector bosons measured.

In order for the vector bosons to receive mass, spontaneous symmetry breaking and the Higgs mechanism needs to be introduced. For this a complex scalar doublet field is introduced:

$$\phi(x) = \begin{pmatrix} \phi^\dagger(x) \\ \phi^0(x) \end{pmatrix}, \quad (2.20)$$

which is usually referred to as the Higgs field. The corresponding Lagrangian has the form

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi), \quad (2.21)$$

with the covariant derivative being the same before $D_\mu = \partial_\mu - ig\tau_a W_\mu^a - ig'\frac{Y}{2}B_\mu$. The Higgs potential $V(\phi^\dagger \phi)$ is

$$V(\phi^\dagger \phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \quad (2.22)$$

where $\mu^2, \lambda > 0$. The Higgs Lagrangian is added to the previously defined $\mathcal{L}_{EW}$ to create the new Lagrangian. The vacuum expectation value can be extracted out of this potential to be $v = \frac{\mu}{\sqrt{\lambda}} e^{i\theta}$ with $\theta \in \mathbb{R}$. This describes a circle in the complex plane with infinitely many solutions. Choosing $\theta = 0$ keeps the results in the real plane and additionally gives the absolute value of $v = \frac{\mu}{\sqrt{\lambda}} \approx 246$ GeV. Expanding the Higgs field around the vacuum expectation value v yields:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \quad (2.23)$$

2. Theoretical Background

where H is the real scalar Higgs field. Using this new expression of the Higgs field for the kinetic term of the Lagrangian gives

$$(D_\mu \phi)^\dagger (D^\mu \phi) = v^2 \frac{g^2}{8} \left[W_\mu^1 W_1^\mu + W_\mu^2 W_2^\mu + W_\mu^3 W_3^\mu + \left(\frac{g'}{g} B_\mu - W_\mu^3 \right) \left(\frac{g'}{g} B^\mu - W_3^\mu \right) \right] - \frac{1}{2} \lambda v^2 h^2 + \text{additional terms}, \quad (2.24)$$

only showing the resulting mass terms for three massive vector bosons and the Higgs boson. To retrieve the massive gauge bosons $W^\pm$ the linear combination of W^1 and W^2 is used:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2). \quad (2.25)$$

The Z boson and the photon originate from the rotation of W^3 and B with the usual rotation matrix as:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & -\sin(\theta_W) \\ \sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) W_\mu^3 - \sin(\theta_W) B_\mu \\ \sin(\theta_W) W_\mu^3 + \cos(\theta_W) B_\mu \end{pmatrix}, \quad (2.26)$$

where θ_W is the weak mixing angle. There are many common relations of the weak mixing angle θ_W , the coupling strengths g and g' and boson masses m_W and m_Z , some of them are:

$$\cos(\theta_W) = \frac{m_W}{m_Z}, \quad (2.27)$$

$$m_W = \frac{1}{2} v g = \frac{v}{2} \sqrt{g^2 + g'^2} \cos(\theta_W), \quad (2.28)$$

$$e = g \sin(\theta_W) = g' \cos(\theta_W). \quad (2.29)$$

Equation 2.27 predicts Z to be heavier than W due to $0 < \cos(\theta_W) < 1$. The values 0 and 1 are not possible because then no mixing would occur and one would return to the massless bosons B and W^3 . The last term in equation 2.24 gives an expression for the Higgs mass being $m_H = \sqrt{\lambda} v^2$ which can be written as $m_H = \mu$ when using the previously introduced definition of v .

Now that the bosons have gained their mass one might notice that the fermions do not have a mass term in $\mathcal{L}_{EW}$ either, again because the mass terms would break $SU(2)$ symmetry. To fix this the so-called Yukawa coupling is introduced which describes the interaction of spinor fields with scalar fields, here the fermion fields and the Higgs field. The contributions of this Lagrangian to the mass term is given by:

$$\mathcal{L}_{\text{Yukawa}} = Y_{ij}^d \bar{Q}^i H d_R^j - Y_{ij}^u \bar{Q}^i \tilde{H} u_R^j - Y_{ij}^e \bar{L}^i H e_R^j + \text{h.c.}, \quad (2.30)$$

with h.c. denoting the Hermitian conjugates of the shown terms. The Yukawa-matrices Y^f contain the information of the fermion mass but are not diagonal so the mass eigenstates are not directly readable. Q^i and L^i represent the Quark and Lepton left-handed doublets respectively where $i \in \{1, 2, 3\}$ gives the generation. d_R^j ,

2. Theoretical Background

u_R^j and e_R^j resemble the down-type quarks, up-type quarks and the charged leptons of generation j . There are no right-handed neutrinos in the SM so only three terms are introduced. H is the Higgs doublet and $\tilde{H} = i\sigma_2 H^*$ where σ_2 is the second Pauli-matrix. After symmetry breaking and changing the basis into the mass basis the fermion masses can be extracted:

$$\mathcal{L}_{\text{Yukawa}} = -m_j^d \bar{d}_L^j d_R^j - m_j^u \bar{u}_L^j u_R^j - m_j^e \bar{e}_L^j d_R^j, \quad (2.31)$$

where m_j^d , m_j^u and m_j^e are the masses of the down-type quarks, the up-type quarks and the charged leptons respectively. The basis of the left-handed quarks is shifted by the unitary matrices U_u and U_d and the basis of the right-handed quarks is shifted by the unitary matrices K_u and K_d with up-type and down-type quarks respectively. These matrices fulfill

$$Y_u = U_u M_u K_u^\dagger, \quad (2.32)$$

where M_u is the diagonal mass matrix of the up-type quarks. This works analogously for down-type quarks. Adding these terms turns U_u , U_d , K_u and K_d into identity matrices leaving only M_u and M_d . The previously introduced m_j^d and m_j^u can be expressed as $\frac{v}{\sqrt{2}}M_d$ and $\frac{v}{\sqrt{2}}M_u$ respectively. The approach for the charged lepton mass works similarly. The change of basis also modifies the kinetic terms keeping dependencies of the $W^\pm$ coupling to quarks with a unitary matrix $V = U_u^\dagger U_d$ which is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix and can be described with four free parameters. Three angles θ_{12} , θ_{23} and θ_{13} and one phase factor δ . The current values without uncertainties of the CKM matrix are [19]:

$$\begin{aligned} |V_{\text{CKM}}| &= \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.0037 \\ 0.225 & 0.973 & 0.042 \\ 0.009 & 0.041 & 0.9991 \end{pmatrix} \\ &= \begin{pmatrix} |c_{12}c_{13}| & |s_{12}c_{13}| & |s_{13}e^{-i\delta}| \\ |-s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta}| & |c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta}| & |s_{23}c_{13}| \\ |s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta}| & |-c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta}| & |c_{23}c_{13}| \end{pmatrix}, \end{aligned} \quad (2.33)$$

suggesting that W bosons strongly prefer to couple to same generation quarks. c_{ij} and s_{ij} represent $\cos(\theta_{ij})$ and $\sin(\theta_{ij})$. The parameters have been measured to be (written without uncertainties) [19]: $\sin(\theta_{12}) = 0.225$, $\sin(\theta_{23}) = 0.042$, $\sin(\theta_{13}) = 0.0037$ and $\delta = 1.15$.

2.2. Proton-Proton Collisions

One way to test the Standard Model thoroughly and search for new physics are collider experiments. There are many different possible collider setups including one moving particle with the other one being at rest or two particles moving towards to each other. These objects could be elementary particles like electrons, could be whole atoms like lead or hadrons like protons. In this thesis the experiment is using two

2. Theoretical Background

protons moving towards to each other. Additionally, hadrons are composite particles containing partons and at the energies reached at LHC the partons interact with each other rather than the whole proton.

This section will introduce quantities relevant for proton-proton collisions and give an introduction to the Feynman formalism including Feynman diagrams. Then, the parton model and the parton distribution functions (PDFs) are explained, followed by a phenomenological description of the scattering process.

2.2.1. Conceptual Quantities

In scattering experiments one is interested in the probability of a final state $\langle f|$ starting from an initial state $\langle i|$. These states are using the bra-ket notation and are taken at $t = \pm\infty$ assuming no other interactions occurring. Introducing the scattering matrix $\mathcal{S}$, also called the $\mathcal{S}$ matrix, the relation

$$\mathcal{S}_{fi} = \langle f|\mathcal{S}|i\rangle \quad (2.34)$$

can be created to describe the transition amplitude $\mathcal{S}_{fi}$ of the initial state going into the final state after interacting as described by the scattering matrix. Noticing the scattering matrix is essentially the time-evolution operator from $-\infty$ to $+\infty$ it can be expressed by [37]

$$\mathcal{S} = \mathcal{T} \exp \left(-i \int_{-\infty}^{+\infty} d^4x \mathcal{H}(x) \right), \quad (2.35)$$

with $\mathcal{T}$ being the time-ordered product and $\mathcal{H}(x)$ being the Hamiltonian density. Expanding yields the Dyson series:

$$\mathcal{S} = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots d^4x_n \mathcal{T}(\mathcal{H}_1(t_1) \dots \mathcal{H}_n(t_n)), \quad (2.36)$$

with the Hamiltonian densities describing the interactions being time-ordered. Instead of the $\mathcal{S}$ matrix the matrix element $\mathcal{M}$ is usually used which is related to the $\mathcal{S}$ matrix by

$$\langle f|\mathcal{S} - \mathbb{1}|i\rangle = (2\pi)^4 \delta^4 \left(\sum p_i - \sum p_f \right) \mathcal{M}, \quad (2.37)$$

with the identity matrix $\mathbb{1}$ and the delta distribution requiring 4-momentum conservation of initial to final momenta. The identity matrix describes a theory where all initial particles are identical to the final particles implying no interactions. For more information on how to calculate $\mathcal{M}$ see section 2.2.2.

The measurable quantity is not the matrix element but rather the cross section σ which gives a measure of probability for a specific process to occur when two particles collide. The cross section is given in units of area, most commonly in barns which is expressed in SI units as $1 \text{ b} = 10^{-28} \text{ m}^2$. The classical interpretation for that would be the area two colliding objects see from one another which would give a probability how likely they are to hit. Due to forces being mediated over distances the cross section does not correspond to the sizes of the particles in quantum mechanics. There are

2. Theoretical Background

two different types of cross section, the total cross section and the differential cross section. The total cross section σ is as described above, the measure of probability for the process to occur, the differential cross section $\frac{d\sigma}{d\Omega}$ does additionally take the solid angle into account with $d\Omega = \sin\theta d\theta d\phi$. This gives the cross section depending on the angle. The total cross section can be gained by integrating the differential cross section over the whole solid angle.

$$\sigma = \int_0^{2\pi} \int_0^\pi \frac{d\sigma}{d\Omega} \sin\theta d\theta d\phi, \quad (2.38)$$

which gives a factor of 4π if $\frac{d\sigma}{d\Omega}$ is independent of θ and ϕ .

The most general formula possible for the differential cross section for two particles going to n final state particles ($p_1 + p_2 \rightarrow p'_1 + p'_2 + \dots + p'_n$) is [37]:

$$\frac{d\sigma}{d\Omega} = (2\pi)^4 \frac{N_1 N_2 \prod_k \frac{1}{g_k!}}{4\sqrt{(p_1 p_2)^2 - m_1^2 m_2^2}} \cdot \prod_{i=1}^n \frac{N'_i d^3 p'_i}{2E'_i (2\pi)^3} \cdot \delta^4 \left(p_1 + p_2 - \sum_{i=1}^n p'_i \right) |\mathcal{M}|^2. \quad (2.39)$$

If the polarizations are not measured the initial spins have to be averaged and the final spins have to be summed over. N_i represents the normalization factors for the different particles involved. For spin-0 bosons and photons this is 1, for spin-1/2 this is $2m$, where m is the mass of this particle. g_k counts the multiplicity of particles of species k in case the final states contain indistinguishable particles, the delta distribution again ensures 4-momentum conservation.

Combining the total cross section with the instantaneous luminosity $\mathcal{L}$ gives a quantity of the number of events detected for a specific dN within a time frame dt :

$$\mathcal{L}\sigma = \frac{dN}{dt}. \quad (2.40)$$

The instantaneous luminosity is independent of the specific process and the final state particles but rather depends on the beam properties (see section 3.1). The instantaneous luminosity can be used to get the integrated luminosity L by integrating over the time.

Lastly, the center-of-mass (COM) energy $\sqrt{s}$ is the energy of the collision in the center-of-mass frame. It is usually denoted in the terms of the Mandelstam variable $s = (p_1 + p_2)^2$.

2.2.2. Feynman Formalism

The calculation of the matrix element $\mathcal{M}$ is highly non-trivial. Richard Feynman introduced the Feynman diagrams [38] to visualize the underlying mathematical processes. Each line and each meeting point of the lines, called vertex, represent a mathematical term. The typical drawing convention is to use straight solid lines for fermions, dashed lines for scalar bosons, wavy solid lines for vector bosons with the exception being gluons which are drawn with curly lines. Fermion lines have arrows attached which point into the direction of movement in time. Anti-fermions can be considered

2. Theoretical Background

to be fermions moving backwards in time and therefore have arrows pointing back in time. The arrows of fermions can by construction not point at the same vertex. Throughout this whole thesis the time axis points to the right in all Feynman diagrams. Figure 2.1 shows the drawing conventions in case of a quark-antiquark pair producing a gluon and two bosons turning into a Higgs boson (a process called vector boson fusion).

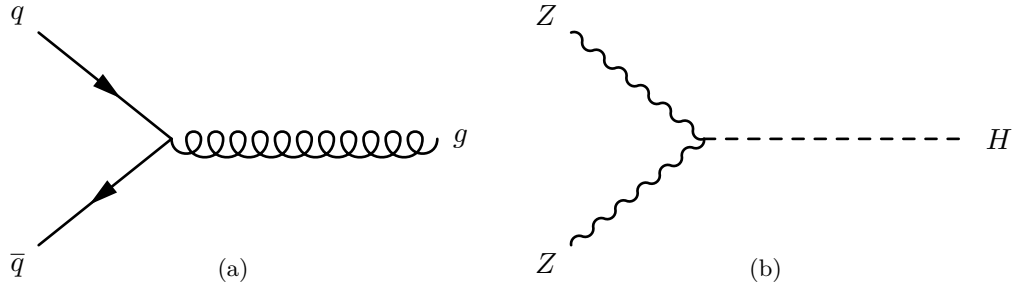


Figure 2.1.: Exemplary Feynman diagrams of a quark-antiquark pair producing a gluon and a vector boson fusion.

Drawing the desired process as a Feynman Diagram and then using the Feynman formalism to turn it into a mathematical expression is a way to calculate the matrix element. As an exemplary calculation the Feynman diagram in figure 2.2 is considered which shows the annihilation of an electron-positron (positively charged electron) pair and a following pair production of the same.

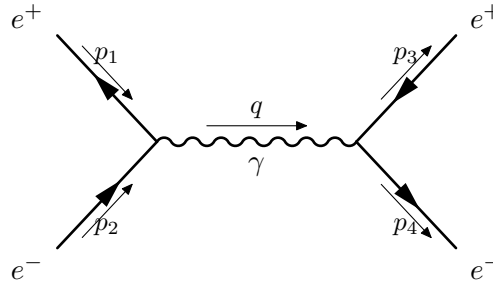


Figure 2.2.: Exemplary Feynman diagram of an electron-positron annihilation and the following pair production.

The process is called s-channel because it uses the Mandelstam variable s during the calculation. Considering that an incoming fermion line is $u(p_i)$, an outgoing fermion line is $\bar{u}(p_i)$, the incoming anti-fermion line is $\bar{v}(p_i)$, the outgoing anti-fermion line is $v(p_i)$, a vertex of a QED process (two charged particles coupling to a photon) is $ie\gamma^\mu$, and the term for an internal photon line, also called photon propagator, is $\frac{-i\eta_{\mu\nu}}{q^2}$ with $\eta_{\mu\nu}$ being the Minkowski metric and q being the momentum of the propagator. Additionally, when creating the matrix element, the fermions lines always have to be

2. Theoretical Background

followed backwards due to e.g. $\bar{u}(p_4)\gamma^\mu v(p_3) \neq v(p_3)\gamma^\mu \bar{u}(p_4)$ not being equal. Taking the product of all mathematical terms while using $s = (p_1 + p_2)^2$ yields:

$$\begin{aligned} i\mathcal{M} &= [\bar{u}(p_4)ie\gamma^\mu v(p_3)] \frac{-i\eta_{\mu\nu}}{q^2} [\bar{v}(p_1)ie\gamma^\nu u(p_2)] \\ &= -\frac{e^2}{s} [\bar{u}(p_4)v(p_3)] [\bar{v}(p_1)u(p_2)]. \end{aligned} \quad (2.41)$$

This is the result of one Feynman diagram but when allowing for more vertices there are an infinite number of Feynman diagrams possible with the same incoming and outgoing particles which all need to be considered and summed when calculating the total matrix element squared $|\mathcal{M}|^2$. Luckily, these higher orders are usually suppressed because of the added factor e^2 every time the number of vertices increases by two. Higher order corrections can stem from vertex or propagator corrections. The Feynman diagram shown in figure 2.2 is called a leading-order (LO) diagram, and every higher order has a "next-to" added, so next-to-leading-order (NLO), next-to-next-to-leading order (NNLO) etc. Exemplary NLO diagrams for the LO diagram of figure 2.2 is given in figure 2.3.

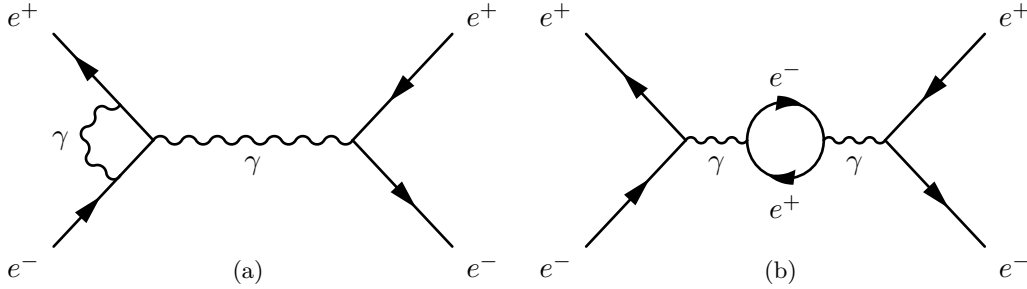


Figure 2.3.: Exemplary NLO Feynman diagrams of an electron-positron annihilation and the following pair production. The left diagram shows a vertex correction and the right diagram shows a propagator correction.

Feynman diagrams containing loops can cause the matrix element to diverge. To still be able to get reasonable results instead of infinities a renormalization scheme is introduced which cancel these divergencies with counterterms.

2.2.3. Parton Distribution Functions

Due to the high energies reached in particle accelerators, instead of the protons interacting with each other as a whole, it is the constituents of the protons, the partons, which interact with each other. These partons are described by the parton model [39]. The proton consists of three valence quarks, two up quarks and one down quark. These protons are held together with the strong interaction via gluon exchange. These gluons also radiate new gluons or can create new quarks with pair production which are called sea quarks. These sea quarks can also recombine back into gluons creating a

2. Theoretical Background

pair production and annihilation cycle. The probability of a parton with a given fraction of the total proton momentum is given by the parton distribution function (PDF) $f_{a;A}(x_a, Q^2)$ [40]. The momentum fraction x_a is the fraction of a parton a of the total proton momentum $p_a = x_a \cdot p_A$ of a proton A and $Q^2 = -q^2$ is the momentum scale given by the energy of the exchange particle, also called the propagator in the last section. The momentum scale changes the distributions of the PDFs as shown in figure 2.4 where at low energies the valence quarks have a higher overall contribution compared to higher energies where the sea quarks and gluons have a significantly higher likelihood to carry a large momentum fraction. The sea quarks and gluons follow a falling distribution whereas the up and down quarks have a peak due to being valence quarks.

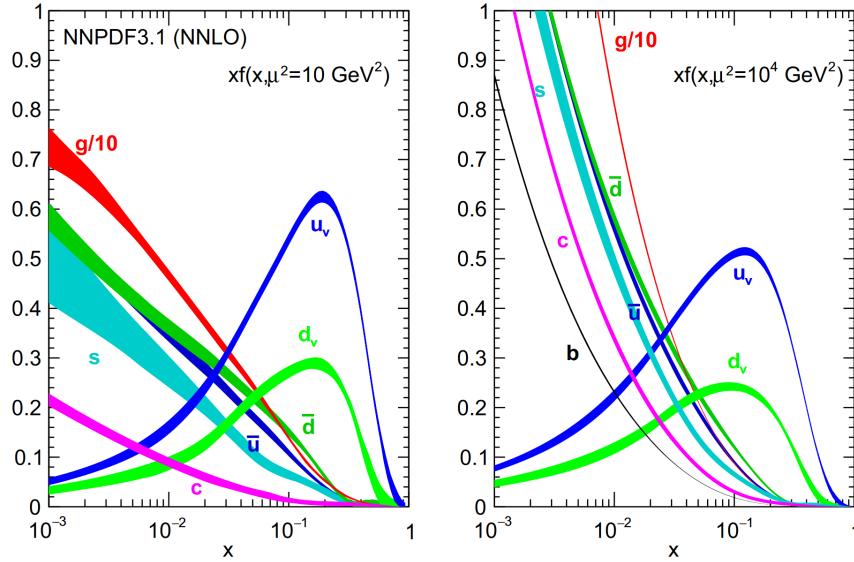


Figure 2.4.: Two PDFs at different momentum scales using the NNPDF3.1 proton PDF. The quarks denoted with a lower v are the valence quarks of the proton and the left figure has a momentum scale of 10 GeV^2 and the right figure is at 10^4 GeV^2 [41].

Knowing the partons interact with each other, the contributing momentum to the Feynman diagrams is the momentum of the specific parton p_a instead of the proton momentum p_A . Assuming protons A and B, to get the cross section of a process $AB \rightarrow X$ one needs to integrate over all possible momentum combinations being able to create the desired final state $ab \rightarrow X$ all having their own cross section $\tilde{\sigma}$. Using the PDF of parton a of Proton A $f_{a;A}$ and the PDF of parton b of Proton B $f_{b;B}$ this results in formula

$$\sigma_{AB \rightarrow X} = \int \int f_{a;A}(x_a, Q^2) f_{b;B}(x_b, Q^2) \tilde{\sigma}_{ab \rightarrow X} dx_a dx_b, \quad (2.42)$$

where $\tilde{\sigma}_{ab \rightarrow X}$ is calculated as in equation 2.39.

The exact momentum fraction dependency of the PDFs cannot be determined

2. Theoretical Background

theoretically due to perturbation theory of QCD breaking down at energy scales required to determine the PDFs. Instead, a global fit has to be applied using data from either deep inelastic scattering experiments like ZEUS [42] or hadron colliders like ATLAS. Using experimental data, the momentum fraction dependency can be determined but the momentum scale Q^2 dependency has to be extrapolated using the DGLAP formalism [43, 44]. For this a reference scale Q_0^2 is chosen, e.g. $Q_0^2 = 1 \text{ GeV}^2$ as done in "NNPDF3.0NLO" [45] which is one of the PDFs used in this thesis. At this scale the PDF can be parameterized using the results of the measurements. Then, the DGLAP formalism, named after Dokshitzer, Gribov, Lipatov, Altarelli and Parisi, is used to extrapolate these to the desired momentum scale which is. Additionally, data which was not used for the determination of the PDF at the reference scale is used to reduce the uncertainties on the PDF further by extrapolating to the corresponding momentum scale and cross checking the results.

The uncertainties of the PDFs are based on the uncertainties of the fitting data. These uncertainties could correlate in general which requires the use of the Hessian method [46] for a proper estimation of the uncertainty. Also, uncertainties of input parameters such as the strong coupling constant or the masses of quarks increase the uncertainties of PDFs. These uncertainties appear as systematic uncertainties in the analysis and need to be taken into account when calculating total uncertainties.

In addition to gluons, quarks can also radiate photons due to their charged nature. This gives rise to a photon PDF which is again used to create a lepton PDF [47]. Leptons are also constituents of the proton due to pair production by photons. Figure 2.5 shows the luminosity of leptons and photons interacting with different particles at a pp-collision at a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$.

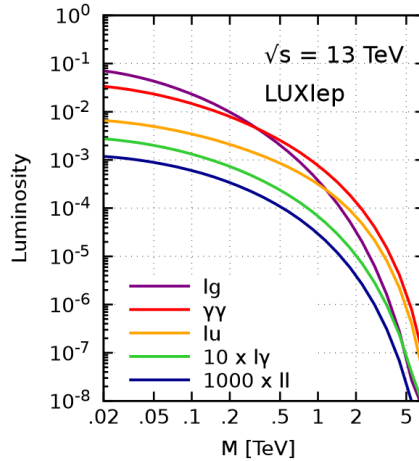


Figure 2.5.: Luminosities of lepton-gluon (violet), photon-photon (red), lepton-up (orange), lepton-photon (green) and lepton-lepton (blue) interactions plotted against the invariant mass of the parton-parton system at a center-of-mass energy of 13TeV. The used PDF is LUXlep [47].

2. Theoretical Background

2.2.4. Proton-Proton Interaction

Electron-electron scattering is comparably simple to model and simulate due to it being fundamental particles with a very well known energy. In proton-proton collisions this is not the case. The interacting partons only carry a fraction of the proton momenta without having discrete values and instead having whole distributions described by PDFs. Additionally, the interacting parton itself is unknown. It could be any of the particles having a PDF in the proton. Figure 2.6 visualizes this, where both protons contribute with an unknown parton with a momentum of $x_i p_i$, $i \in \{1, 2\}$ to the hard scattering process and produce the proton remnants X_1 and X_2 .

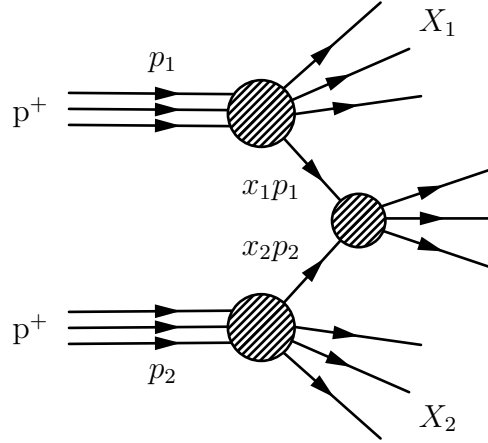


Figure 2.6.: Illustration of the hard scattering process in a proton-proton collision leading to the proton remnants X_1 and X_2 and the hard scattering process.

After the hard scattering process a plethora of additional interactions occur. These will be explained step by step and resemble the order Monte Carlo generators process these interactions when simulating proton-proton collisions (as will be described in ref 3.8). Apart from hard interactions also soft processes take place. They are commonly called parton showers and describe soft emissions, which can happen before or after the hard process mimicking initial or final state radiation. Soft emissions are collinear gluons (or rarely photons) having either low momentum or being emitted under small angles which are calculated using the DGLAP evolution [43, 44]. After having simulated a parton shower there are a lot of coloured particles present. Due to confinement these particles are not allowed to exist freely and need to recombine into hadrons, which is called hadronization. These hadrons are mostly unstable and decay at some point which also needs to be considered. The probability of a particle to decay depends on the mean lifetime τ and follows an exponential decay

$$N(t) = N_0 \exp\left(-\frac{t}{\tau}\right). \quad (2.43)$$

2. Theoretical Background

Here, $N(t)$ is the number of particles after a time t , N_0 is the number of starting particles. If one is interested in the chance for one particle to decay $1 - N(t)$ yields the chance for a particle to have decayed after a time t after setting $N_0 = 1$. If a particle has multiple decay channels the mean lifetime is given by

$$\tau_{\text{total}} = \frac{\prod_i \tau_i}{\sum_i \tau_i}, \quad (2.44)$$

with τ_i representing the mean lifetime of each process independently. In addition to the already mentioned processes the protons remnants or parts of them can also interact in some way. This is called the underlying process and again leads to new parton showers, hadronization and decays. An exemplary proton-proton collision is shown in figure 2.7.

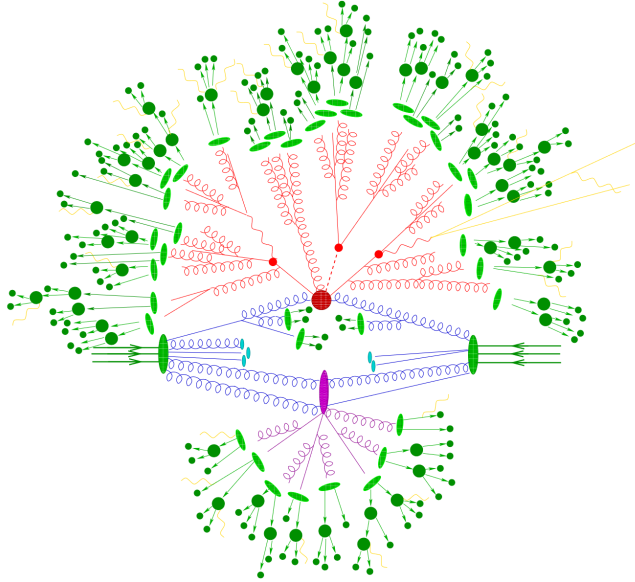


Figure 2.7.: Illustration of a proton-proton. The large green discs are the protons, the blue lines depict the original proton components and the larger red circle represents the hard scattering process. The red small vertices and lines corresponds to the parton showers, the light green discs are the hadronization and the connecting dark green circles and lines are the hadron decays. Lastly, the purple object symbolize the underlying events and the yellow lines are soft photon radiations[48].

This approach only considers the interaction of one proton with another proton. In reality not just two protons interact but in the proton beam each proton has a probability to interact. These secondary interactions are called pile-up and contribute a large part of the background in events. There are two main types of pile-up: in-time pile-up and out-of-time pile-up [49].

In-time pile-up describes pile-up created from additional collisions in the same bunch

2. Theoretical Background

crossing strongly depending on the number of interactions per bunch crossing. This is again dependent on the instantaneous luminosity.

The second type of pile-up, out-of-time pile-up, describes additional interactions not originating from the bunch crossing where the primary vertex is placed. This mainly happens due to detector components having a larger response time than the bunch crossing gap of 25 ns.

Other pile-up contributions include random hits in the muon system, a proton bunch scraping a collimator device which produces muons parallel to the beam-line or collisions of protons on residual gas inside the beam-pipe.

2.3. Going Beyond the Standard Model

Despite the excellent accuracy of the Standard Model describing particle physics experiments there are still questions the Standard Model cannot answer. This section will go into further details of the limits of the SM and then describes Leptoquarks, a hypothetical particle which could handle some of the limitations of the SM. Any physics trying to expand the Standard Model is called physics beyond the Standard Model (BSM).

2.3.1. Limits of the Standard Model

The most notable shortcoming of the SM is the complete absence of one fundamental force: gravity. In particular, general relativity (GR) [2] and the SM were not created with a possible unification in mind and need to be treated separately. The inclusion of gravity into the SM would be a first step towards a theory of everything (TOE) which would require a theory of quantum gravity.

Also, neutrino oscillations [16, 17] have been measured which requires neutrino masses. Neutrino masses are not predicted by the Standard Model and have been measured to be more than 5 orders of magnitude smaller than the electron mass ($\nu_e < 0.8$ eV [19]). This discrepancy can also not be explained by the SM.

Observation at cosmological scales such as galaxy rotations lead to the conclusion that dark matter has to make up a large portion of the total matter in the universe. Dark matter has to not interact electromagnetically and has to have gravitational effects by having a mass. The Standard Model does not provide a suitable dark matter candidate.

Additionally, cosmological observations also predict the existence of dark energy in order to explain the accelerating expansion of the universe. Current approximations set the dark energy around 70% of the total matter in the universe, dark matter at 25% and baryonic matter (matter contained in the SM) at around 5%. These values are only rough estimates and dedicated matter composition papers have large uncertainties on these values [50].

Other unexplained features include the matter-antimatter asymmetry in the observable universe. The Standard Model does not prefer matter over anti-matter and predicts an approximate equality of them. The observed universe has an

2. Theoretical Background

abundance of matter compared to anti-matter. Baryogenesis could be an explanation for this but this has three conditions according to Sakharov [51]: Baryon number violation, charge-parity (CP) violation and interactions out of the thermal equilibrium. CP violation has been observed but the observed value is several orders of magnitude too small to explain the large discrepancy between matter and anti-matter.

In addition to the observations there are still some oddities of the Standard Model which cannot be explained currently. The first one being the 19 free parameters of the Standard Model which do not have a theoretical prediction and need to be determined experimentally. These 19 parameters are: the 9 fermion masses, the three CKM angles and the CKM phase factor, the three gauge couplings of $SU(3)$, $SU(2)$ and $U(1)$, the QCD vacuum angle, the vacuum expectation value and the Higgs mass. These 19 parameters describe one possible combination in the parameter space where other parameter combinations could be defined as long as they absorb all 19 degrees of freedom. The second one being the hierarchy problem [52] referring to the large discrepancies in order of magnitudes for some observed values. One example being that the weak force is more than 20 orders of magnitude stronger than the gravitational force. One postulated solution for that would be the introduction of extra dimensions [53]. The third one being the apparent symmetry between leptons and quarks with both having three generations with two particles each without having an underlying symmetry requiring this.

The Standard Model predicts lepton flavour universality. This means that at energy scales way above the lepton masses the branching ratios to all leptons generations should be identical. Multiple measurements measured notable deviations from the expected ratio of 1 showing a deviation of more than 2σ [54, 55]. This is not enough to be counted as a discovery due to not being above the discovery significance of 5σ but could hint to BSM physics. Further studies are necessary.

Approaches to unify the strong force with the electromagnetic force and the weak force similar to the electroweak unification have been made. These are called Grand Unified Theories (GUT) [56] and need to introduce additional symmetries. These theories are based on symmetry groups such as $SU(5)$ or $SO(10)$ where quarks and leptons share multiplets strongly correlating them. In these models there would be a boson connecting them, the Leptoquark.

There are many postulated solutions to the shortcomings of the Standard Model. One of them being supersymmetry [57] giving every particle a supersymmetric partner. This partner has a changed spin by $1/2$ giving bosons a supersymmetric spin- $1/2$ fermion partner and fermions a supersymmetric spin-1 boson partner. Due to supersymmetry breaking these partners would be heavier than the SM particles known.

2. Theoretical Background

2.3.2. Leptoquarks

Leptoquarks (LQ) are hypothetical particles carrying a lepton number and a baryon number and therefore coupling to a lepton and a quark. They also have to carry color charge due to coupling to one quark. LQs appear in many BSM physics including GUTs with a $SU(5)$ symmetry [56], a $SU(15)$ symmetry [58] or E_6 superstring models [59, 60] which have an additional $U(1)$ symmetry. They naturally arise in these theories being the boson connecting the lepton and quark families which could explain the similarities between the lepton and quark families. Due to the possibility to couple differently to different generations, LQ would be able to explain the possible violation of lepton flavour universality if new measurements on lepton flavour violation manage to surpass the 5σ threshold.

Apart from the fact that they have a lepton number and baryon number very little is constrained in terms of quantum numbers and properties and depends on the theory which predicts the LQ [61]. They could be scalar bosons as well as vector bosons, can have absolute electrical charges of $1/3$, $2/3$, $4/3$ or $5/3$. Depending on the charge the lepton and baryon number also change depending whether the LQ couples to the particles or anti-particles. A LQ always has a anti-particle variant replacing the contributing fermions with their respective anti-particle. The investigated LQ in this thesis has an electric charge of $Q = -4/3$, a spin of 0, is an $SU(2)$ singlet and is usually referred to as an $\tilde{S}_1$ LQ. A comprehensive list of all possible LQs directly coupling to SM fermions is given in table 2.4, where $SU(3)_C$, $SU(2)_W$ and $U(1)_Y$ indicate the QCD representation, the weak isospin and the weak hypercharge respectively which are the correspondent charges. Additionally, the fermion number $F = L + 3B$ is introduced which is either 0 when one of the fermions is an anti-particle and -2 if this is not the case. For more information on the current constraints of every LQ model, see [62].

Table 2.4.: Different possible Leptoquarks with their respective quantum numbers [19].

Spin	F	$SU(3)_C$	$SU(2)_W$	$U(1)_Y$	Allowed coupling
0	-2	$\bar{3}$	1	$1/3$	$\bar{q}_L^c \ell_L$ or $\bar{u}_R^c e_R$
0	-2	$\bar{3}$	1	$4/3$	$\bar{d}_R^c e_R$
0	-2	$\bar{3}$	3	$1/3$	$\bar{q}_L^c \ell_L$
1	-2	$\bar{3}$	2	$5/6$	$\bar{q}_L^c \gamma^\mu e_R$ or $\bar{d}_R^c \gamma^\mu \ell_L$
1	-2	$\bar{3}$	2	$-1/6$	$\bar{u}_R^c \gamma^\mu \ell_L$
0	0	3	2	$7/6$	$\bar{q}_L e_R$ or $\bar{u}_R \ell_L$
0	0	3	2	$1/6$	$\bar{d}_R \ell_L$
1	0	3	1	$2/3$	$\bar{q}_L \gamma^\mu \ell_L$ or $\bar{d}_R \gamma^\mu e_R$
1	0	3	1	$5/3$	$\bar{u}_R \gamma^\mu e_R$
1	0	3	3	$2/3$	$\bar{q}_L \gamma^\mu \ell_L$

2. Theoretical Background

In addition to setting the specific type of LQ searched for, the Yukawa couplings can be set to simplify searches. This analysis focuses on a LQ exclusively coupling to an electron and a b -quark (eb) by setting all other couplings to zero. For results of other channels in a resonant LQ search see the internal note of the ongoing search of ATLAS [63].

There are multiple ways known on how to search for LQs in hadron colliders namely the pair-production (PP) [64, 65], single-production (SP) [66, 67] and limits from the Drell-Yan process (DY) [68, 69]. This is additionally accompanied by constraints set by the weak charge Q_W in channels containing first generation quarks due to LQs modifying Q_W if they were to exist [70]. These constraints do not apply to the investigated electron and b -quark (eb) channel because Q_W measurements were only done for the weak charge of protons [71], cesium [72] and thallium [72] which are all containing only valence quarks of the first generation .

Each search has its own advantages being able to exclude different regions. Pair-production is mostly independent of the coupling to fermions by coupling to gluons but is strongly constrained in the mass region due to having to produce two real LQs allowing to exclude low mass regions. The Drell-Yan search is mostly independent on the LQ mass by creating a virtual LQ but is heavily constrained by the coupling being able to exclude high coupling regions. The single-production is both dependent on the mass and the coupling being able to possibly constrain the intermediate regions. A leading order Feynman diagram for pair-production, single-production and the DY-limit is given in figure 2.8 and the current constraints in the eb channel by the different searches are given in figure 2.9. The single-production search is not able to constrain the exclusion region further in the eb channel for Run 1 and partial Run 2 results.

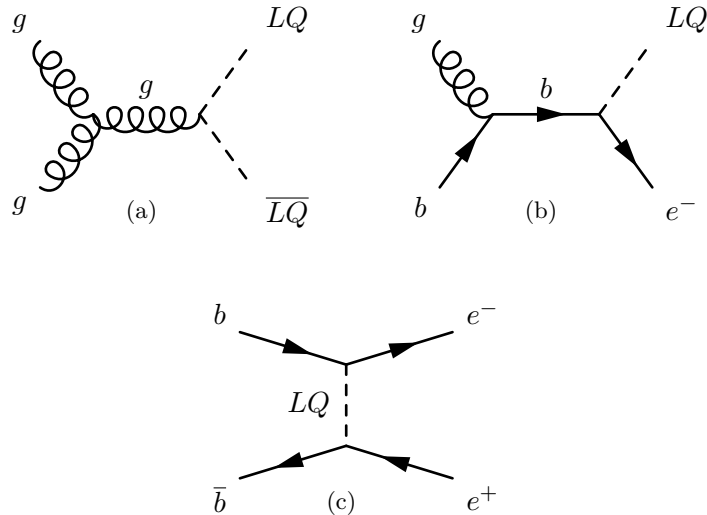


Figure 2.8.: Exemplary LO Feynman diagrams of LQ pair-production (a), single-production (b) and Drell-Yan production (c).

2. Theoretical Background

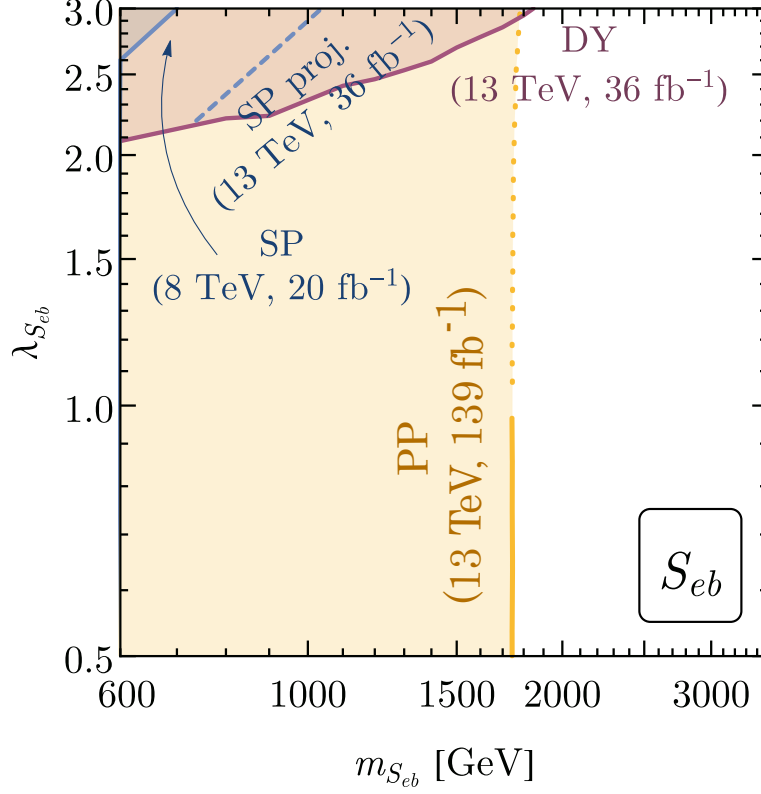


Figure 2.9.: 95% CL exclusion limits of previous searches using different approaches. This is based on [70] but the pair-production limit has been replaced with a more recent and more constraining result given by [73] which in turn extracted the limit from the search given by [64] using the whole Run 2 dataset. The dashed blue line is the expected gain from a single-production search using 36 fb^{-1} instead of 20 fb^{-1} and a higher center-of-mass energy.

Due to the recent addition of an electron PDF for the proton a new method of detecting LQs has emerged: resonant LQ production [74, 75]. This allows to probe the intermediate mass and coupling regions similar to the single-production search but estimations expect a better exclusion limit for the resonant production. Figure 2.10 shows the LO resonant production of a LQ with an electron directly originating out of the proton. The achieved Feynman diagram precision is NLO and uses the diagrams given in figure 2.11. Many NLO diagrams give an additional lepton in the final state giving rise for a resonant LQ search with one final state lepton and two final state leptons.

2. Theoretical Background

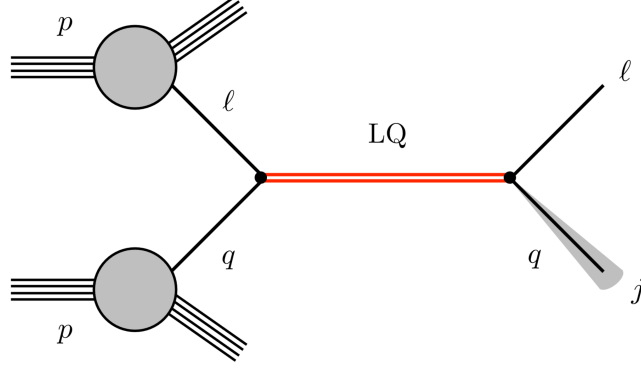


Figure 2.10.: LO Feynman diagram of a resonant LQ production directly showing the that the quark and lepton are coming out of the protons. The final state quark creates a particle shower and is identified as a jet [74].

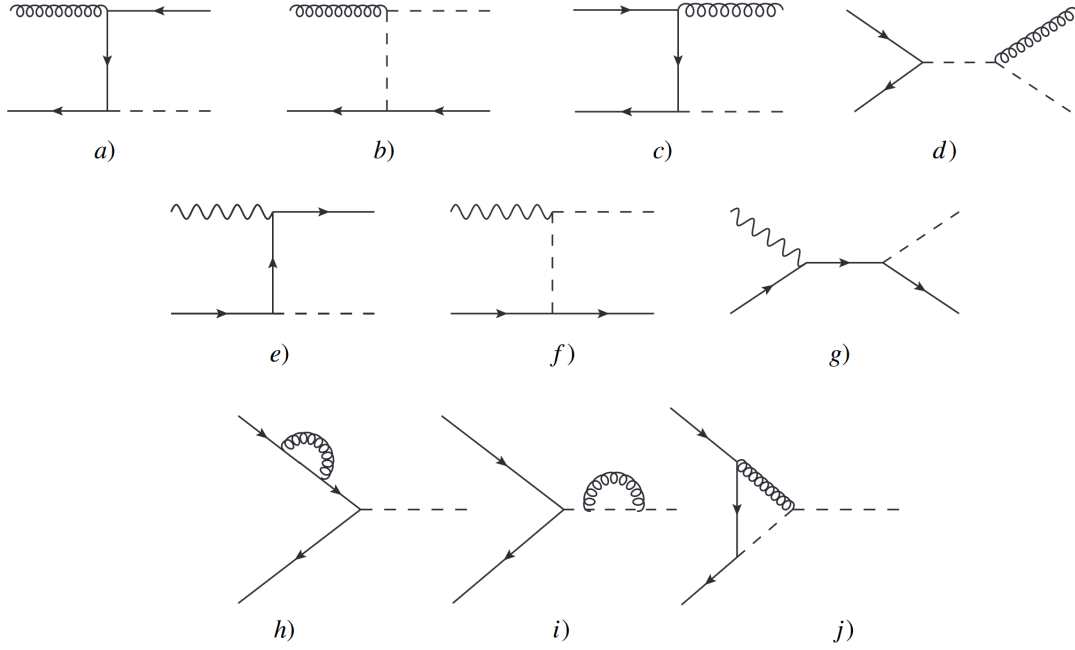


Figure 2.11.: NLO Feynman diagrams being considered for the resonant LQ production. Diagrams a) and b) are $g + e \rightarrow b + LQ$ processes, c) and d) are $b + e \rightarrow g + LQ$ processes, e), f) and g) are $\gamma + b \rightarrow e + LQ$ processes and h), i) and j) are the virtual corrections considered [76].

2. Theoretical Background

The predicted exclusion limit gained from a resonant search is given in figure 2.12. The given limit is for the electron-down-quark channel, also showing the constraint given by Q_W and is just to visualize that there is an expected gain to be seen. The exact numbers and limits are not transferable to the eb channel and are not the most recent results of the electron-down-quark channel either.

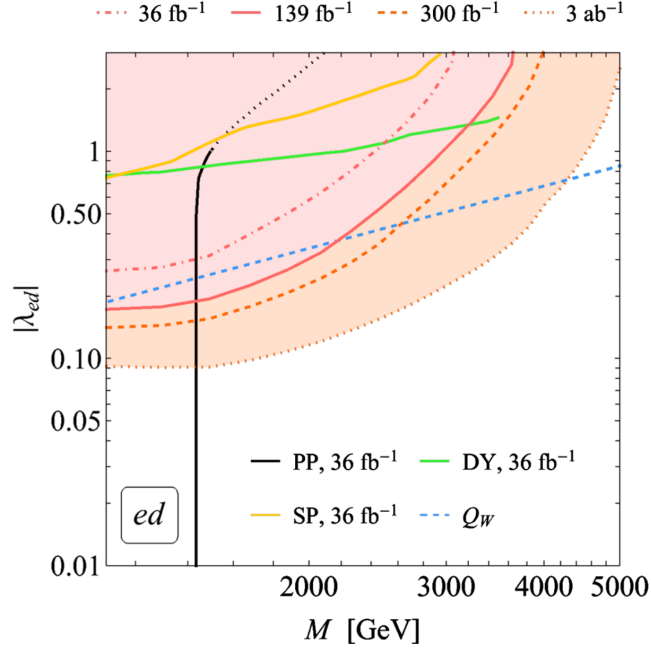


Figure 2.12.: 95% CL exclusion limits of previous searches compared to the expected exclusion limits gained in a resonant LQ production in the eb channel. The weak charge Q_W limit is included and predictions for a resonant LQ search with four different integrated luminosities are given. [74].

3. ATLAS Experiment

The following chapter provides information on the ATLAS (A Toroidal LHC ApparatuS) experiment [6, 77, 78] and the Large Hadron Collider (LHC) [5], which the ATLAS experiment is part of. The common coordinate system at ATLAS and common kinematic variables are given and defined. The discussed components of the ATLAS experiment describe the experimental setup during the full Run 2 data-taking period from 2015 to 2018. These components include the tracking detector, the electromagnetic and the hadronic calorimeters, the forward calorimeter and the muon spectrometer. Additionally, the trigger and data acquisition, object reconstruction and identification, the inner working of Monte Carlo simulations, and the common data formats at the ATLAS experiment are elaborated on.

3.1. Large Hadron Collider

The Large Hadron Collider (LHC) is a hadron-hadron collider located at the European Organization for Nuclear Research (CERN, derived from Conseil européen pour la recherche nucléaire) accelerator complex near Geneva, Switzerland. It has a circumference of 26.7 km and had a center-of-mass energy of $\sqrt{s} = 13$ TeV and reached an integrated luminosity of 140 fb^{-1} during Run 2, which is the data basis for this thesis. It is the worlds largest and most energetic particle accelerator at the time of this publication. This allows for the possibility to search for physics beyond the standard model including, but not being exclusive to, the Higgs sector, Supersymmetry, Quantum Black Holes and Leptoquarks. The LHC consists of two rings due to using same charged particles and it uses the tunnel of its predecessor, the Large Electron Positron Collider (LEP). To reach this unmatched center-of-mass energy several pre-accelerators are mandatory. The protons originate from ionizing hydrogen and are accelerated up to 50 MeV in the linear accelerator LINAC2. They then enter a series of circular accelerators, namely the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) which gives them an energy of 450 GeV per proton. At this point they are entering the LHC and are accelerated up to 6.5 TeV. The accelerator complex of CERN used for proton-proton collisions at the LHC is shown in figure 3.1 including all four interaction points of the LHC.

The experiments located at the four interaction points of the LHC are ATLAS, CMS [7], ALICE [80] and LHCb [81]. ATLAS and CMS are multipurpose detectors, whereas ALICE is specialized for ion collisions when the LHC runs in ion beam mode using typically lead ions. LHCb uses the same proton beams as ATLAS and CMS but focuses on high precision measurements of processes involving bottom quarks. The proton beams consist of 2808 bunches with a spacing of 25 ns each having roughly

3. ATLAS Experiment

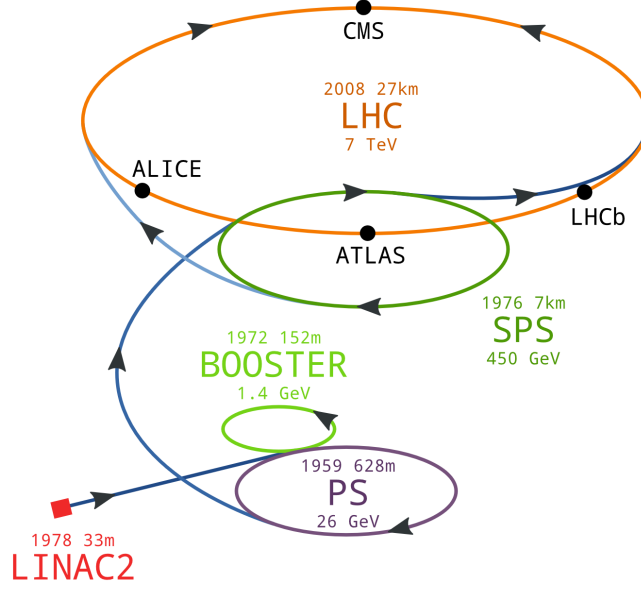


Figure 3.1.: Schematic of the accelerator complex at CERN with all used accelerators and beam lines used to accelerate the protons used in the LHC. Information includes the construction date, the circumference of circular accelerators, length of linear accelerators and maximum beam energy [79].

$1.5 \cdot 10^{11}$ protons and have a revolution frequency of 11.2 kHz [82]. The beam dimensions transverse to the beam direction are $16.6 \mu\text{m}$ in both x and y direction. For the definition of the axis see section 3.2. The integrated luminosity provided by the LHC is only dependent on the accelerator properties given by [83]:

$$\mathcal{L} = \frac{I_{b1} I_{b2} f_{rev} n_b}{2\pi \sqrt{(\sigma_{1x}^2 + \sigma_{2x}^2)(\sigma_{1y}^2 + \sigma_{2y}^2)}}, \quad (3.1)$$

with the two beam intensities I_{b1} and I_{b2} , the revolution frequency f_{rev} , the number of bunches n_b , and the beam dimensions σ_x and σ_y for both beams transverse to the beam direction. The highest instantaneous luminosity measured during Run 2 was $\mathcal{L} = 21 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. The instantaneous luminosity decreases over time due to loss of protons. To reduce this effect several bunch fillings are done but the starting instantaneous luminosity deviates every time. This deviation of the starting instantaneous luminosity after each bunch fill is shown in figure 3.2 for 2018.

3. ATLAS Experiment

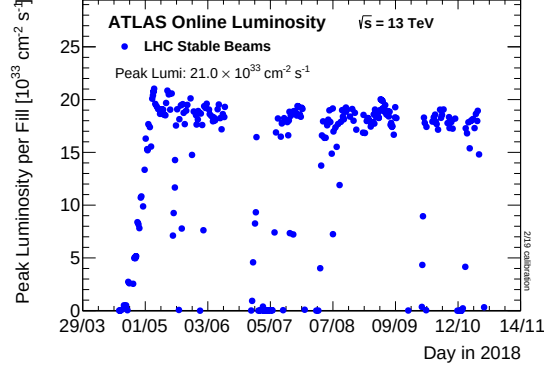


Figure 3.2.: Peak instantaneous luminosity per bunch fill during 2018 [84].

3.2. Coordinate System and Common Kinematic Variables

The ATLAS detector is constructed around the beam axis at the interaction point. Its goal is to measure, reconstruct and identify as many different particles as possible. It is symmetric with respect to the interaction point and has a rotational symmetry around the beam axis. It covers most of the solid angle in order to minimize the missed particles. A schematic inside view of the detector is shown in figure 3.3 by hiding a side of the detector. The Pixel detector, semiconductor tracker (SCT) and transition radiation tracker (TRT) create the tracking detector, also called the inner detector (ID), of the ATLAS experiment. The other components shown are the muon chambers being part of the muon spectrometer (MS), the different magnets and the sampling calorimeters.

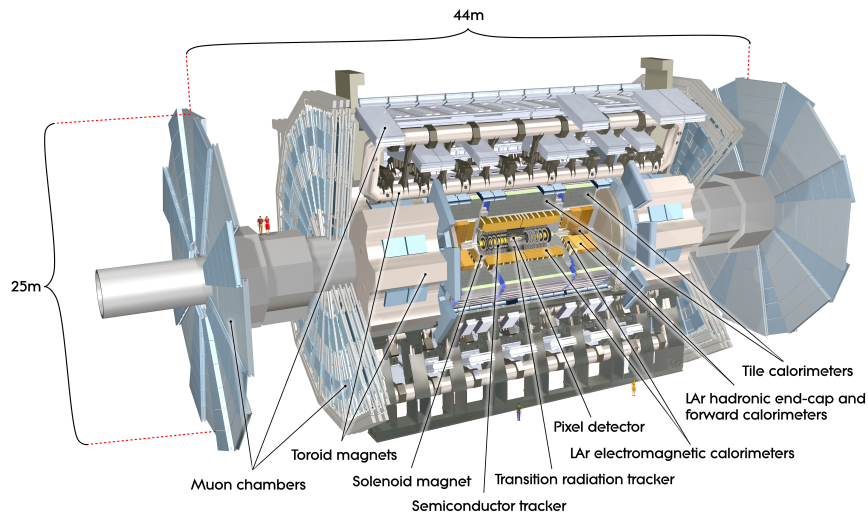


Figure 3.3.: Schematic of the ATLAS detector showing the inner components [85].

3. ATLAS Experiment

Before the different components of the ATLAS experiment are described in further detail, the common coordinate system used in the experiment and the relevant kinematic variables are elaborated.

A right-handed Cartesian coordinate system is used with the origin placed at the interaction point. The z -axis is along the beam axis so the x - y plane is transverse to the beam line. The x -axis points toward the center of the LHC and the y -axis points upwards. The azimuthal angle ϕ has values from $-\pi$ to π , is positioned in the x - y plane and is defined with respect to the x -axis. Lastly, the polar angle θ is defined with respect to z , goes from 0 to π and gives the angle of a particle's flight path to the beam line.

Because θ is not Lorentz-invariant under boosts along the beam-axis the rapidity y is used instead. y is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (3.2)$$

for a given particle with energy E and a momentum along the z -axis p_z . Additionally, one can define the pseudorapidity η as: axis the rapidity y is used instead. y is defined as:

$$\eta = \frac{1}{2} \ln \left(\frac{p + p_z}{p - p_z} \right), \quad (3.3)$$

with p being the total momenta. For massless particles the pseudorapidity is equal to the rapidity. This also holds approximately for highly relativistic particles. Both quantities are still not Lorentz-invariant under boosts along the beam-axis but the differences are, so $\eta_1 - \eta_2$ is invariant. Pseudorapidity can also be expressed in terms of the polar angle θ with the relation

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.4)$$

From this one can see that η is 0 when in the transverse plane and $+\infty$ or $-\infty$ when along the positive or negative z -axis respectively.

Because the detector cannot measure particles going into the beam line, energies and momenta are usually expressed with respect to the transverse plane:

$$p_T = \sqrt{p_x^2 + p_y^2} = |p| \sin(\theta), \quad (3.5)$$

$$E_T = \sqrt{p_T^2 + m^2}, \quad (3.6)$$

which reduces to $E_T = p_T$ for massless particles. Distances of particle directions ΔR are expressed as differences in η and ϕ with equation

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \quad (3.7)$$

Other used variables include z_0 and d_0 being the distance of the reconstructed vertex from the primary vertex in z -axis and the transverse plane respectively. Also the

3. ATLAS Experiment

missing transverse energy $\vec{E}_T^{\text{miss}}$ can be defined with the negative sum of the transverse energy $\vec{E}_T^i$ of every particle as:

$$\vec{E}_T^{\text{miss}} = - \sum_i \vec{E}_T^i. \quad (3.8)$$

For more information on the determination of $\vec{E}_T^{\text{miss}}$ see section 3.7.5. Analogously, the missing transverse momentum can be defined as the negative sum of the transverse momenta of all particles. This analysis additionally uses the significance of the missing transverse momentum $\mathcal{S}(p_T^{\text{miss}})$ defined in [86] because it yields a better discrimination between the signal and background and follows the equation

$$\mathcal{S}(p_T^{\text{miss}}) = \frac{p_T^{\text{miss}}}{\sqrt{\sigma_L^2 (1 - \rho_{LT}^2)}}, \quad (3.9)$$

where p_T^{miss} is the missing transverse momentum, σ_L is the resolution longitudinal to p_T^{miss} and ρ_{LT} is the correlation between longitudinal and transverse resolutions with respect to p_T^{miss} . This equation takes effects like noise into account and expresses how likely it is that a measured value of p_T^{miss} is only due to noise fluctuations and not from a weakly interacting particle such as a neutrino. Another variable that can be defined is the transverse mass of two objects m_T which is defined as:

$$m_T = \sqrt{2p_T^A p_T^B (1 - \cos(\Delta\phi))}, \quad (3.10)$$

with p_T^A and p_T^B being the momenta of the particles A and B respectively and $\Delta\phi$ being the difference in azimuthal angle. The transverse mass used in this thesis defines the transverse mass of the leading lepton and the missing transverse momentum $m_T^{\ell 1, \text{miss}}$ which is just referred to as the transverse mass m_T for the rest of the thesis.

3.3. Inner Detector

The inner detectors [87, 88, 89] purpose is to provide tracking information of charged particles. This is done by using three detector components namely the pixel detector [90, 91], the semiconductor tracker (SCT) [92] and the transition radiation tracker (TRT) [93]. The whole system is surrounded by the central solenoid [94] which provides a magnetic field along the beam axis of 2 T in the inner detector to be able to determine the transverse momentum of charged particles by studying the curvature. The detector has a coverage of up to $|\eta| = 2.5$ and achieves a momentum resolution of $\frac{\Delta p_T}{p_T} = 0.05\% \cdot p_T[\text{GeV}] + 1\%$. Figure 3.4 shows the inner detector of the sub-systems split into the barrel and end-cap region. The solenoid is not shown.

The pixel detector is the innermost part of the detector and in the central region it has four layers [96] of silicon pixel modules placed cylindrically around the beam line and a size of $50 \times 400 \mu\text{m}^2$. In both end-cap regions the modules are mounted on three disks being placed perpendicular to the beam. The two main goals of the pixel

3. ATLAS Experiment

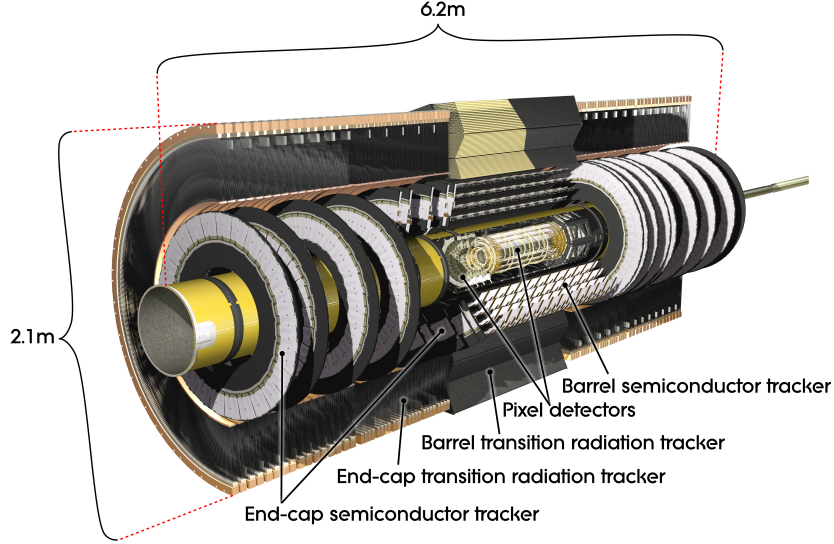


Figure 3.4.: Schematic of the ATLAS inner detector showing the inner components [95].

detector are high resolution particle tracks and reconstruction of multiple interaction vertices especially the primary vertex. It has a total number of around 80 million pixels providing a resolution of 115 μm in beam direction and 10 μm in the transverse plane.

The SCT is build around the pixel detector which uses silicon microstrips mounted on a total of eight layers in the central region where two are used at the same time to determine one space point which yields a total of four additional space points. Each end-cap region has the silicon strips placed on nine discs where each layer is twisted by 40 mrad compared to the previous one. The total resolution of the SCT is 580 μm in beam direction and 10 μm in the transverse plane and has a total of 6.3 million readout channels.

The outermost component of the inner detector is the TRT which functions as a drift chamber. It is made of roughly 300,000 straw tubes filled with 70% xenon but also 27% CO_2 and 3% O_2 . In the end-cap region they are placed perpendicular to the beam line. These tubes are surrounded by transition radiation material which emits transition radiation when a charged particle passes through it. The intensity is proportional to the Lorentz factor $\gamma = \frac{E}{m}$ with E and m being the respective energy and mass of the passing particle. Charged particles ionize the gases in the tubes when passing them by creating free electrons. These electrons then drift to wires which have a potential in the order of kV applied. In addition to more tracking information with the drift chambers, which have a transverse resolution of 130 μm , the intensity of the radiation is used to specify the type of particle by determining the mass after an energy determination in the calorimeter system thus e.g. differentiating electrons

3. ATLAS Experiment

from hadrons. The TRT has a total of approximately 350 thousand readout channels.

3.4. Calorimeter System

This section will go into further detail on the different calorimeters located at the ATLAS experiment around the ID. The purpose of these calorimeters is to measure the energies of the particles produced in the proton-proton collision. This is done for all particles apart from specific exclusions like muons and neutrinos. The calorimeters are optimized for electromagnetic showers and hadronic showers separately by using electromagnetic calorimeters (ECAL) and hadronic calorimeters (HCAL).

An electromagnetic shower [97] is induced by electrons, photons and positrons. There are two processes creating the rather narrow shower shape: bremsstrahlung and pair production of an electron-positron pair. For pair production the photon needs to have energy of at least twice the electron mass $E_\gamma > 2m_e$. These processes are repeated until the photon does not have enough energy or the electron/positron energy is low enough that the energy loss due bremsstrahlung is overshadowed by effects like ionization. Due to this being fixed thresholds, the final number of charged particles strictly correlate to the number of total energy the particle had.

Hadronic showers [98] on the other hand have a plenitude of effects which makes them larger with larger variances in the shape of the shower. The dominating effects are the inelastic scattering of light hadrons on the detector material, production of hadrons and the de-excitation of previously excited nuclei.

Figure 3.5 shows the calorimeter system of ATLAS split into the different sub-systems. All calorimeters at ATLAS are sampling calorimeters meaning that they have passive and active materials layered. The passive material initializes and drives the showering process and the active material is used to be able to determine the number of charged particles. The used passive material is dependent on the sub-system but the active material used is mostly liquid argon (LAr) with the exception being the tile barrel and the tile extended barrel using scintillating tiles as active material. Due to the more compact shape of electromagnetic showers compared to hadronic showers the ECALs are positioned before the HCALs (closer to the interaction point).

ECALs [100] have an accordion like structure and can be separated into the EM barrel and the electromagnetic end-cap (EMEC) with the EM barrel covering $|\eta| < 1.475$ and the EMEC covering $1.375 < |\eta| < 3.2$. Both the EM barrel as well as the EMEC use lead as the passive material. Although they have an overlap region the sensitivity in the so called barrel-endcap crack region from $1.37 < |\eta| < 1.52$ is lower due reduced thickness of the detector material. This crack is necessary for the cables and services into the ID. The thickness of the ECALs is expressed in radiation lengths X_0 which is the distance where on average the energy of an electron is reduced to $1/e$ of its original energy. The thickness of the EM barrel region is more than $22 X_0$ and the EMEC reaches more than $24 X_0$. The relative energy resolution of the ECALs can be

3. ATLAS Experiment

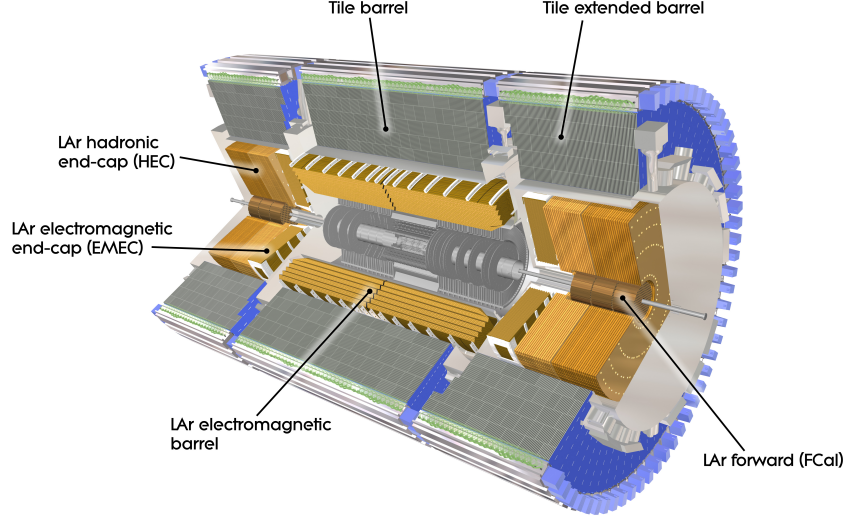


Figure 3.5.: Schematic of the ATLAS calorimeter system showing the inner components [99].

expressed with the equation as defined in [101]

$$\frac{\Delta E}{E} = \frac{0.1}{\sqrt{E}} + \frac{b}{E} + 0.7\%, \quad (3.11)$$

where E is given in GeV and the first term is the sampling term, the second being the noise term and the last being the constant term. The noise term can be expressed as roughly $350 \cdot \cosh(\eta)$ MeV. All terms are η dependent and the given values hold for low $|\eta|$. The granularity of the EM barrel differs between the three layers. The first layer, the one closest to the interaction point, has a granularity of 0.003×0.1 in $\eta \times \phi$, the second layer has 0.025×0.025 and the third layer also has 0.025×0.025 . The EMEC behaves identical in the second and third layer and the first layer is very similar but changes to 0.006×0.1 in $\eta \times \phi$ in the range $2.0 < |\eta| < 2.5$. In the range $2.5 < |\eta| < 3.2$ it has a granularity of 0.1×0.1 .

HCALs can be separated into the tile barrel [102], the tile extended barrel [102] and the hadronic end-cap (HEC). They cover $|\eta| < 1$, $0.8 < |\eta| < 1.7$ and $1.5 < |\eta| < 3.2$ respectively. The passive material for the tile calorimeters is iron and HEC uses copper for its passive material. For HCALs the thickness is expressed in interaction lengths λ which is defined as the mean distance a hadron travels until it scatters inelastically with the material. The detector provides at least 10λ in a pseudorapidity region of $|\eta| < 4.9$. This is visualized in figure 3.6 and ensures that as much energy as possible is stopped and not lost for energy and E_T^{miss} calculations. Additionally, this prevents unwanted particles entering the muon spectrometer. The relative energy resolution

3. ATLAS Experiment

for energies E in GeV of the HCALs is given by [103] as being roughly

$$\frac{\Delta E}{E} = \frac{0.5}{\sqrt{E}} + 3\%. \quad (3.12)$$

The granularity of the tile barrel and the tile extended barrel is 0.1×0.1 in $\eta \times \phi$. The HEC has the same granularity below $|\eta| = 2.5$ and a granularity of 0.2×0.2 beyond that.

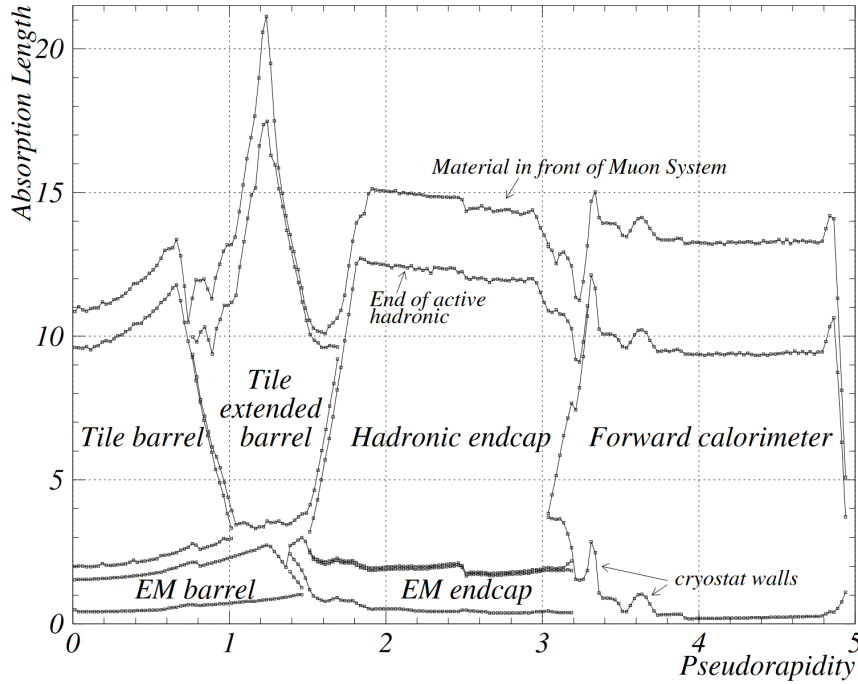


Figure 3.6.: Total absorption length in the ATLAS calorimeter system given as a function of η [104].

The forward calorimeter (FCal) [105] is designed to cover the high $|\eta|$ region not charted by the other calorimeters. It has a range of $3.1 < |\eta| < 4.9$ and functions as ECAL and HCAL simultaneously by having three layers which are specialized. The first is usually called the electromagnetic layer and the other two are called the hadronic layers because the respective showers are mostly contained within these layers. To be able to achieve this the passive material for the first layer is copper and for the second and third layer it is tungsten. Due to the limited size of the calorimeter the relative energy resolution is worse compared to the other calorimeters and is given by

$$\frac{\Delta E}{E} = \frac{1}{\sqrt{E}} + 10\%. \quad (3.13)$$

Although the reconstructed objects in this thesis only consider η ranges up to 2.5 (see section 5.1), the calorimeter is still relevant for the determination of E_T^{miss} as well as $\mathcal{S}(p_T^{\text{miss}})$.

3.5. Muon Spectrometer

Muons require a dedicated detector to be reconstructed. This is because they only deposit energies of about 3 GeV in the calorimeters due to their high mass. For this, the muon spectrometer (MS) [106] is used which is positioned around the calorimeters and comprises the outermost part of the ATLAS detector. It covers up to $|\eta| < 2.7$ and uses, similar to the ID, magnetic fields to bend the tracks and reconstruct momenta precisely. The toroids provide a bending power of 1.5 to 5.5 Tm in the barrel region ($|\eta| < 1.4$) and 1.0 to 7.5 Tm in the end-caps ($1.6 < |\eta| < 2.7$) [107, 108, 109]. The transition region ($1.4 < |\eta| < 1.6$) has a bending power of 2.5 to 6 Tm which is achieved by having a total of 8 toroids placed throughout the MS. The MS is composed of four detector sub-systems which fulfil two different purposes. All sub-systems and toroids of the MS are shown in figure 3.7.

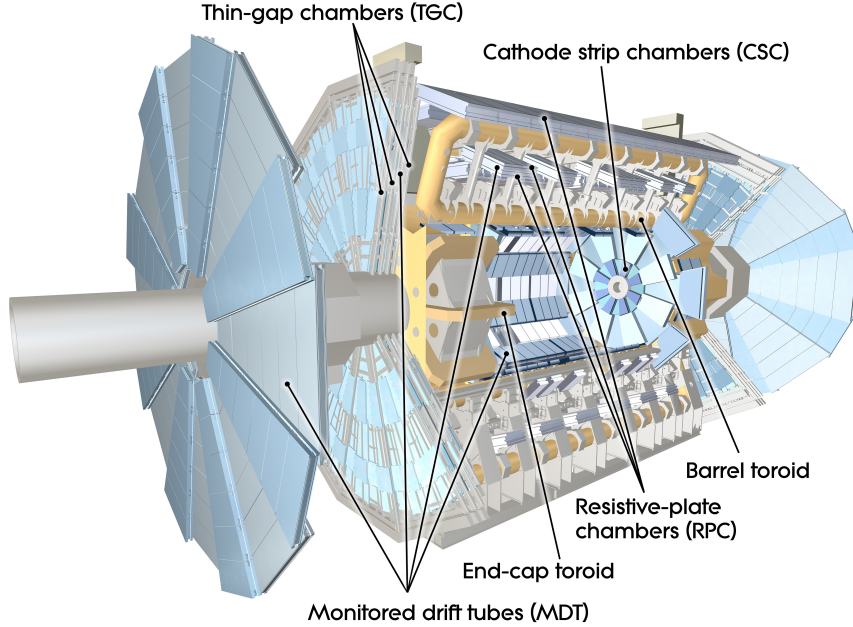


Figure 3.7.: Schematic of the ATLAS muon spectrometer showing the inner components [110].

The first task is the ability to trigger on muons, which the resistive-plate chambers and the thin-gap chambers are optimized for. For this, they need a time resolution of less than the bunch spacing of 25 ns and need to be able to identify a muon candidate well within the trigger decision time of 2.5 μ s (see section 3.6). The spacial resolution they provide is in the order of mm.

The second task the MS needs to handle is the high precision measurement of muon momenta which is accomplished with the monitor drift tube and the cathode strip chambers. They have a spacial resolution in the order of μ m. The relative momentum resolution of a muon with an energy of 1 TeV is 10%. Lastly, there is a gap in the

3. ATLAS Experiment

precision chambers of the MS at around $\eta = 0$ for cables and services into the inner part of the detector.

3.6. Trigger and Data Acquisition

The LHC provides a bunch crossing every 25 ns which corresponds to a collision rate of 40 MHz. With each event taking roughly 1.5MB of disk space this would lead to 60 TB/s of data being stored if all events were to be saved. This is not possible and trigger systems were developed to reduce the number of saved events by selecting events which may have a physical interest. This is done in two main steps: a hardware based level 1 trigger (L1) [111] and a software based high level trigger (HLT) [112]. Together with other systems they build the Trigger and Data Acquisition (TDAQ) [113] system.

The L1 trigger has a decision time of 2.5 μ s and the data is stored temporarily within this time frame and afterwards either deleted or further processed depending on the trigger decision. It receives reduced cell granularity of the calorimeters and basic track information of the muon tracker. With this information a relevant signature is searched with a trigger menu. Each trigger has a specific requirement, e.g. an electron with a transverse energy of more than 20 GeV. Possible relevant signatures are: jets, electrons, photons, taus and muons having to surpass certain thresholds. Muon signatures only use the muon spectrometer information to trigger, whereas the other signature candidates only use the calorimeter information. The region in which this signature was measured is marked as a Region Of Interest. If one of the triggers in the trigger menu accept the event, it is saved. The L1 trigger reduces the event rate to 100 kHz and the remaining events are then passed to the HLT.

The HLT has access to the full calorimeter and tracking information within the region of interest. This allows for more precise but still fast identification algorithms to be used. Again, the signatures have to surpass certain thresholds. If one of the triggers in the trigger menu accept the event the event is saved on disk. The event rate is reduced to 1.5 kHz and the data is reconstructed which reduces the file size to roughly 1MB. This leads to a total 1.5GB/s being saved to disk storage.

Some triggers in both the HLT and L1 have a prescaling which lets it reject a portion of the suitable events. This is usually done for frequently firing triggers with e.g. a very low energy threshold to still have data in this region without exploding the total trigger rate.

Due to being limited to fast reconstruction algorithms, having coarser cell granularity or having only a part of the detector information the energy values can deviate from the values gained from offline reconstruction. This means that an offline reconstructed electron with an energy of 18 GeV could have passed a L1 trigger which requires more than 20 GeV for an electron or vice versa where it wasn't saved although it would have had enough in the offline reconstruction. For this reason one usually tries to require triggers in an analysis to be in the plateau region where these effects less likely [114].

3.7. Object Reconstruction and Identification

In order to evaluate the measurements of the detector, objects need to be reconstructed. The detector itself measures cell energies without knowing which particle(s) deposited these energies. By applying track and energy reconstructions these energies can be assigned to a physical object which passed the detector. The reconstruction and identification steps done for each relevant particle in this thesis are discussed in this section.

3.7.1. Electrons

The electron reconstruction within $|\eta| < 2.5$ can be broken down into two major steps. The deposition of energy in the electromagnetic calorimeter and the measurement of a matching track in the inner detector. The reconstruction method has to be changed above $|\eta| > 2.5$ because there is no more tracking information available (see section 3.3) but this will not be discussed in further detail.

A so called topo-cluster algorithm [115, 116] is used to build proto-clusters in the electromagnetic and hadronic calorimeter. Using a likelihood approach for the deposited energies, the shower is either marked as an electromagnetic cluster or a hadronic cluster depending on the fraction of energy deposited in each calorimeter. In case it is marked as an electromagnetic cluster a matching of the tracking information is required [117]. For that, the ATLAS Global χ^2 Track Fitter [118] is used to reconstruct tracks within the inner detector by reading out the pixel and SCT detectors. These tracks are then rather loosely matched to the electromagnetic cluster considering possible energy losses including bremsstrahlung by applying the Gaussian Sum Filter [119]. If multiple tracks met the requirements the best fitting track is determined using an algorithm which includes the information on ΔR of the center of the cluster and the track.

The 4-momentum is then defined by taking the information from both the track and the electromagnetic cluster. The energy is directly taken from the cluster and the η and ϕ values are calculated using the track information.

In case an electron candidate is found a likelihood approach [120] is applied to determine the chance of the candidate being a real electron rather than background event such as a hadronic jet or a converted photon. There are a total of 19 discriminating parameters used which include, but are not exclusive to: 1) shower width, 2) amount of energy in the hadronic layers and 3) track starting point. The first two help in distinguishing from hadronic jets whereas the third one serves for the rejection of converted photons. Depending on the likelihood result the candidate is put into one of three identification categories: Loose, Medium or Tight. Each of them has increasingly higher likelihood thresholds to be accepted. This reduces both the number of real electrons accepted and the number of fake electrons but has bigger impact on the number of fake electrons. The specific efficiency depends on the transverse energy but is set to be 93%, 88% and 80% over the whole E_T

3. ATLAS Experiment

region for Loose, Medium and Tight respectively. For example the Tight operating point changes from an efficiency of 58% at $E_T = 4.5$ GeV to an efficiency of 88% at $E_T = 100$ GeV. In all cases over the whole E_T the background rejection is above 95%. More information on the identification can be found in the internal ATLAS documentation of reference [121].

Finally, also isolation criteria can be applied to reduce background contributions. There are two different types to cover. The first one is the option of having calorimeter isolation that limits the maximum allowed energy in a cone around the electron energy deposit. The second one is a tracking isolation limiting the sum of transverse energy around a cone of the electron track. The isolation criteria can be applied simultaneously or mutually exclusive depending on the use case. The efficiency of the isolation working points depends on the E_T but is usually located somewhere above 90% for transverse energies above 40 GeV. For more details on the different kinds of isolation criteria to apply, see the internal ATLAS documentation [122].

3.7.2. Muons

The muon reconstruction [123, 124, 125] uses the information of the inner detector and the muon spectrometer in order to identify a muon. There are several different methods to determine a muon candidate increasing for example the acceptance at especially high or low η values (extrapolated muons and calorimeter-tagged muons respectively). Due to the fact that this analysis only considers combined muons this type is the only one discussed in further detail here.

Combined muons candidates are determined by independently reconstructing tracks in the inner detector and the muon spectrometer. The inner detector reconstruction is similar to the reconstruction of the electron track without requiring a matched shower or considering bremsstrahlung whereas the muon spectrometer uses a different χ^2 fit from the hits in the middle layers and then testing the hits in the inner and outer layers. To obtain these tracks and to get a candidate either the outside-in strategy extrapolates the muon spectrometer track into the inner detector or the inside-out strategy extrapolates the inner detector track into the muon spectrometer. In both cases the extrapolated track is then matched to the track from this detector part.

The Muon candidates are then identified as one of five identification categories using a likelihood approach of it being a real muon and not a background event such as muon originating from a kaon or pion decay. Some of the identification parameters are the difference of the p_T of the inner detector and muon spectrometers tracks over the p_T of the combined track or the χ^2 value of the combined fit. The five categories are: Loose, Medium, Tight, Low- p_T and High- p_T . The Loose selection has the highest reconstruction efficiency, Medium has the lowest uncertainties with reasonable purity and efficiency and Tight maximizes purity. The Low- p_T and High- p_T are, as the name suggests, optimized for low p_T muons (> 3 GeV) and high p_T muons (> 100

3. ATLAS Experiment

GeV) respectively while keeping uncertainties as low as possible.

In particular starting from $p_T > 3$ GeV the Loose working point has an efficiency between 90%-98% increasing for higher energies. The Medium has an efficiency between 70%-97% and Tight has an efficiency between 36%-93% (no muons below $p_T < 4$ GeV allowed in the Tight working point). The plateaus are reached at different transverse momenta, roughly 5 GeV for Loose and Medium and 20 GeV for Tight. The background rejection of the tight requirements are between 3 to 10 times the background rejection of the Loose working point. The background rejection of the Loose working point varies between 1.2% and 0.12% decreasing with higher transverse momenta. The Low- p_T working point provides similar signal efficiencies than the Loose working point, especially at low transverse momenta, while providing a better background rejection. The High- p_T working point has the worst efficiency of all working points, with around 80% starting from $p_T > 20$ GeV but has the best momentum resolution.

The isolation criteria applicable in order to reduce possible backgrounds are very similar to the electron isolation criteria but differ in the specific values which are optimal. It is possible to use tracking isolation, calorimeter isolation or both. The specific efficiencies depend on the used method but mostly reach a plateau close to 100% between 30-80 GeV and mostly start at a efficiency of 50-70% at 4 GeV. Further information on the isolation cuts for muons is found in the internal ATLAS note [122].

3.7.3. Jets

Jet signatures are the most prominent objects reconstructed in hadron colliders. They do not consist of any single particle but rather are made of many final state particles being collimated. They originate, among other things, from gluons and quarks hadronizing due to confinement. The reconstruction [126, 127], similar to the reconstruction of electrons, starts with the creation of topo-clusters [115] in the electromagnetic and hadronic calorimeters. These topo-clusters are created around cells with an energy to noise ratio above 4σ and neighboring cells are added as long as the energy to noise ratio stays above 2σ . By also using tracker information of the charged particles the particle flow algorithm [128] corrects energy and momentum measurements and decreases the respective uncertainties by e.g. having less background from pile-up.

These topo-clusters are then combined with using a clustering algorithm based on distances. Common algorithms include the k_t [129], the anti- k_t [130] and the Cambridge/Aachen [131] algorithms. The one used at ATLAS is the anti- k_t algorithm which clusters around high p_T bins. The equation shown is for the anti- k_t algorithm but the named algorithms essentially only differ in the sign of the exponent of the transverse momenta in the following equations:

$$d_{ij} = \min \left(k_{ti}^{-2}, k_{tj}^{2p} \right) \frac{\Delta_{ij}^2}{R^2}, \quad (3.14)$$

3. ATLAS Experiment

$$d_{iB} = k_{ti}^{-2}, \quad (3.15)$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the geometrical distance between the two objects using the azimuth angle $\phi_{i,j}$ and rapidity $y_{i,j}$ of the objects i and j , $k_{ti,tj}$ are the transverse momenta of the objects i, j and R is the radius parameter, usually set to be 0.4. Also, d_{ij} and d_{iB} are the distance parameters between the two objects and between object i and the beam. If $d_{ij} < d_{iB}$ the clusters are combined otherwise i is defined as a jet and removed from further distance comparisons. Figure 3.8 visualizes the results of the anti- k_t algorithm using an exemplary set of energy depositions consisting of a parton-level event and many soft events (see section 2.2).

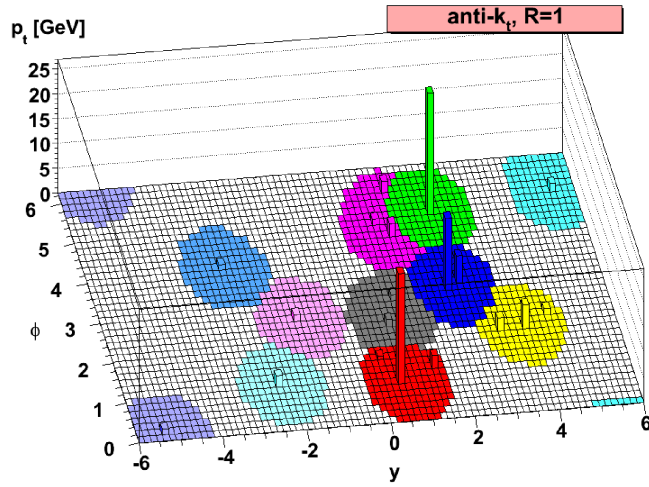


Figure 3.8.: Results of the anti- k_t algorithm with one parton-level event and 10^4 soft events with an R of 1. Each color corresponds to a different jet object shown in the ϕ - y plane where the height of the pillars give the p_T of the bin [130].

One way to additionally reduce background contributions is the use of a Jet Vertex Tagger (JVT) [132, 133] which reconstructs the production vertex and returns a quantity how likely it is that the tested jet originates from the primary vertex. This mostly serves for rejection of pile-up vertices and detector noise.

3.7.4. b-Tagging

The jet reconstruction does not yield any information on the type of jet. Due to differences in physical processes the topology of jets differ depending on the creating process so they can be tagged. A b -quark mostly decays into a W boson and a lighter quark which gives a different substructure than just a light quark. Additionally, due to the lifetime of the b -quark it is possible to measure a displaced vertex. Some jets types are hard to distinguish, e.g. an u and d quark create almost identical looking jet structures. So the final tagging efficiencies also vary on the jet type and a b -jet has the most unique structure which makes it the best performing at jet-tagging.

3. ATLAS Experiment

The b -tagging is done in two major steps. First, low-level algorithms are used to reconstruct the characteristic properties of b -jets. These algorithms consider, but are not limited to, the impact parameters IP2D and IP3D [134], secondary vertex finding algorithms (e.g. SV1 [135]) or multi-vertex finding algorithms such as JetFitter [136]. Second, the outputs of these algorithms are used as inputs into boosted decision trees and (deep) neural networks such as MV2 and DL1 [137]. Both calculate likelihood ratios and return the probability whether the tested jet is a b -jet. A exemplary setting of 77% means that this number of b -jets are successfully tagged.

3.7.5. Missing Transverse Energy

The missing transverse energy [127] is an important quantity in describing particles the detector cannot detect such as neutrinos or other non-interacting particles. The missing transverse energy is used instead of the total missing energy because the detector cannot measure particles leaving through the beam line which would lead to large uncertainties. The missing transverse energy on the other hand should theoretically (almost, due to that fact that the beams have a crossing angle and are not perfectly head on [138, 139],) sum up to zero due to 4-momentum conservation if the energies of all present particles were known. In reality this is not true due to noise fluctuations, nonetheless it is the only way to estimate the contribution of weakly interacting particles. The missing transverse energy E_T^{miss} is defined via E_x^{miss} and E_y^{miss} as:

$$E_{x,y}^{\text{miss}} = \sum E_{x,y}^e + \sum E_{x,y}^\gamma + \sum E_{x,y}^\tau + \sum E_{x,y}^{\text{jets}} + \sum E_{x,y}^\mu + \sum E_{x,y}^{\text{soft}}, \quad (3.16)$$

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \quad (3.17)$$

where the sums contain all the identified particles that fall into the specific category. The only exception being $\sum E_{x,y}^{\text{soft}}$ which is the sum of all identified tracks which were not assigned to any particle coming from soft events. Instead of just using the energy cells themselves, the energies of the reconstructed objects are summed up to be able to use object based calibrations and reduce systematic errors [140, 141].

3.7.6. Overlap Removal

All of the reconstructions mentioned above are done independently from each other which can lead to overlaps. To avoid energy cells being assigned to multiple particles or tracks being assigned multiple times an overlap removal algorithm is applied. This algorithm uses the checks in table 3.1 to determine which object the energy cell should be assigned to. For more details on the overlap removal the ATLAS documentation [142] and the internal ATLAS documentation [143] are available.

3. ATLAS Experiment

Table 3.1.: Overlap removal criteria applied to the defined objects in this analysis. If the criteria is fulfilled the object in "Reject" is removed when being compared to "Against" [142, 143].

Reject	Against	Criteria
Muon	Electron	is Calo-Muon and shared ID track
Electron	Muon	shared ID track
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < 0.4$
Jet	Muon	$\text{NumTrack} < 3$ and (ghost-associated or $\Delta R < 0.2$)
Muon	Jet	$\Delta R < 0.4$

3.8. Monte Carlo Method

In order to predict the results the Standard Model or BSM physics would yield when measuring data with an experiment so called Monte Carlo (MC) simulations are used. Monte Carlo simulations, named after the Monte Carlo Casino in Monte Carlo, Monaco, simulates a physics process first on "truth" level and then determines the detector responses to such a process. More on that in the next paragraphs. In particle physics due to quantization results can not be predicted with certainty and can only be expressed in terms of percentages and likelihoods. In this sense predicting the outcome of a particle physics experiment is somewhat like gambling. To still be able to draw conclusions the results are expressed in significances, so how far they deviate from the expectation value (more on this in chapter 6). To be able to do this one needs to thoroughly study the expected results both for the case where the Standard Model holds and for the case new physics are present (like a Leptoquark in this thesis). MC is doing this by using random number generators (RNG) to mimic the uncertainty of nature. This means that doing the same process twice results in different looking events due to all things being probability based.

The first level in MC production is "truth" level which means that the interaction takes place without interacting with anything else, so e.g. no detector material interactions. For this a plethora of MC generators are used [144] all having their own advantages and disadvantages. A great majority of them share a common order of execution with the first one being the calculation of the hard process by calculating all possible Feynman diagrams up to a certain order. Possible orders are e.g. leading-order (LO), next-to-leading-order (NLO) or next-to-next-to-leading-order (NNLO). Starting from NLO additional initial state radiations (ISR) and final state radiations (FSR) are allowed.

After the hard process, the steps discussed in detail in section 2.2.4 are processed being: parton showers, hadronization, decay of unstable particles and simulation of underlying events. Parton showers simulate soft gluon and photon emissions. To ensure the separation from hard process ISR and FSR from soft emissions certain recombination methods are employed to avoid double counting events, namely match-

3. ATLAS Experiment

ing and merging [48]. The simulation of the hadronization process is done in different ways depending on the used MC generator. The most prominent phenomenological approaches are the string model [145, 146] and the cluster model [147, 148]. Then, the unstable particles which are created from the hadronization process can decay with a certain probability depending on the mean lifetime and the underlying events are simulated.

After having simulated the "truth" level the detector response needs to be estimated to be able to compare the simulation to real data. The detector simulation of the ATLAS detector is done using GEANT4 [149]. This program is able to describe the interaction of particle with matter and has very precise information on the detector materials and dimensions. It is also able to simulate the exact response including smearing effect of the different components, e.g. the energy cells, and how the components digitize this information. This GEANT4 information is combined in a digitization process with pile-up and noise contributions. This has the advantage of not having to simulate the whole event again if the pile-up conditions are changed. From this point on the simulated event can be handled identically to a real event and is able to be saved in the same data formats. Figure 3.9 visualizes the different data formats of simulated vs. real events and the point they converge. For more information on the data formats for real events see the next section.

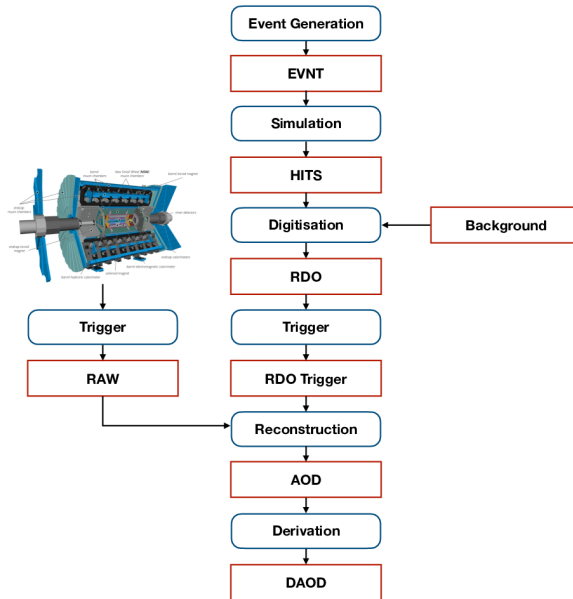


Figure 3.9.: Illustration of the data formats for simulated and real events. EVNT is the truth level information, HITS is a detector simulation format and Background is the format for noise and pile-up. RDO contains the exact same information as the whole detector transmits to the triggers, and RDO trigger corresponds to the RAW format.[150].

3. ATLAS Experiment

The number of produced MC events does not correspond to the number of expected events, it just decreases the statistical uncertainty of the simulated process and is only limited by processing power. After producing the desired amount of events, weights need to be applied to represent the SM process. The purpose of the event generator weight depends on the specific generator used and e.g. accounts for double counting or shape corrections and always needs to be applied. Then the events are normalized, scaled to the cross section of the specific process simulated and then scaled to the integrated luminosity of the collider experiment, here 140 fb^{-1} . Additionally, scale factors need to be applied which accommodate for inefficiencies. This includes efficiency corrections like b -tagging efficiencies, JVT efficiencies, reconstruction efficiencies, isolation efficiencies or pile-up corrections.

3.9. Data Formats

There are a plethora of different data formats used at ATLAS [151] for real events. The initial data format is RAW which consists of the full cell information of the whole ATLAS detector. These information are used to apply reconstruction algorithms (see section 3.7) and calibrations and are then saved as Event Summary Data (ESD) which is solely done on the LHC Computing Grid [152, 153]. This file is then re-calibrated and formatted to an Analysis Data Object (xAOD) [154]. The size of these files is still in the Petabyte range which emphasizes preselection for the analysis.

This is done in three main steps: "skimming", "thinning" and "slimming". Skimming is the removal of whole events by requiring for example the event to have more than two reconstructed electrons. Thinning removes selected objects from every event for example removing all muons because muons would not be needed for a search. Slimming removes variables from all events and objects which would for example be the ϕ value for every object. (Cutting away ϕ information is not advised at all because even if it is not used it is an important control parameter.)

The resulting data format is called derived xAOD (DAOD) which have a file size in the order of Terabyte. Finally, the DAODs can be stored in a n-tuple format which can be read and modified by the data analysis framework ROOT [155].

4. Data and Monte Carlo Samples

This chapter contains all the information concerning the measured and simulated datasets. This includes the collision data measured in Run 2 (section 4.1) and the Monte Carlo Simulation of the signal and the background samples (sections 4.2 and 4.3).

4.1. Collision Data

This search uses the Run 2 pp collision data of the LHC, which has an integrated luminosity of 140 fb^{-1} and a center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$. The data taking period stretches from 2015-2018. All used events are part of the Good Runs List which describes the part of data collected which matches to requirements to be used in an analysis for example all required detector parts working properly [82]. The resulting luminosity is shown in figure 4.1 where the value of the integrated luminosity is stated as 139 fb^{-1} in contrary to the used $140.1 \pm 1.2 \text{ fb}^{-1}$ which is the number recommended in [82]. This deviation most likely arises from a slight change in recommended luminosity since the creation of this plot. Nevertheless, it visualizes still very accurately the increase in integrated luminosity throughout Run 2 and the efficiency of delivered to usable luminosity. The internal ATLAS documentation for the Good Runs List can be found in reference [156].

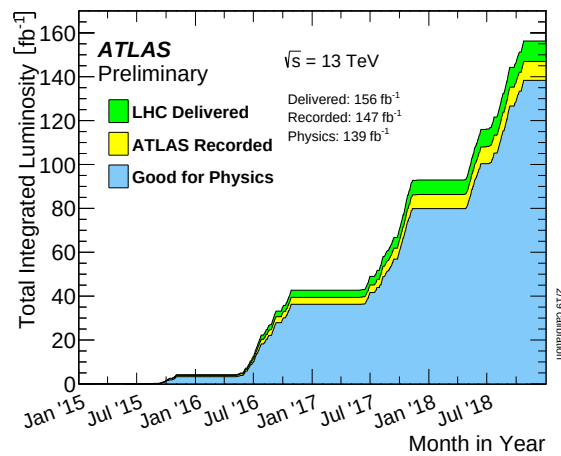


Figure 4.1.: Visualization of the integrated luminosity delivered by the LHC, recorded by ATLAS and used for physics analysis. Note that 139 fb^{-1} is an old estimation and this analysis uses 140 fb^{-1} [84].

4. Data and Monte Carlo Samples

Additionally, all events are selected using different single lepton HLTs depending on the lowest available unrescaled thresholds for each data taking period. These triggers are listed in table 4.1 and work accumulatively. The most important parts of the trigger name are the first two options where the minimum energy and the minimum likelihood ratio is defined. A trigger with "e24_lhmedium" for example requires an electron with $E_T > 24$ GeV (e24) and a medium likelihood ratio (lhmedium). These likelihood ratios are using neural networks to determine how likely it is that the observed particle is an electron. The settings Tight, Medium and Loose only define an acceptance threshold. Lastly, the other options give additional information on the HLT, e.g. "L1EM22VH" requires the level 1 trigger with this name to fire which in turn fires when an electron of more than 22 GeV is measured. The triggers applied follow the internal ATLAS recommendations at [157].

Table 4.1.: Summary of single lepton trigger list which needs to be triggered to be used in the analysis. Only 2015 has its own trigger list whereas the other years use the same triggers. Only one trigger needs to fire for the event to be accepted.

Period	Electron Triggers
2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose
2016	HLT_e26_lhtight_nod0_ivarloose
2017	HLT_e60_lhmedium_L1EM22VHI
2018	HLT_e140_lhloose_L1EM22VHI

During the data taking period the number of collisions per bunch crossing have changed greatly and increased mainly due to increasing instantaneous luminosity. This results in larger backgrounds in the form of larger pile-up which needs to be accounted for. The interaction distributions for each year separately and accumulated are shown in figure 4.2 where the expected number of interactions of the whole Run 2 ($\langle\mu\rangle = 33.7$) more than doubles the average number of interactions in 2015 ($\langle\mu\rangle = 13.4$).

4. Data and Monte Carlo Samples

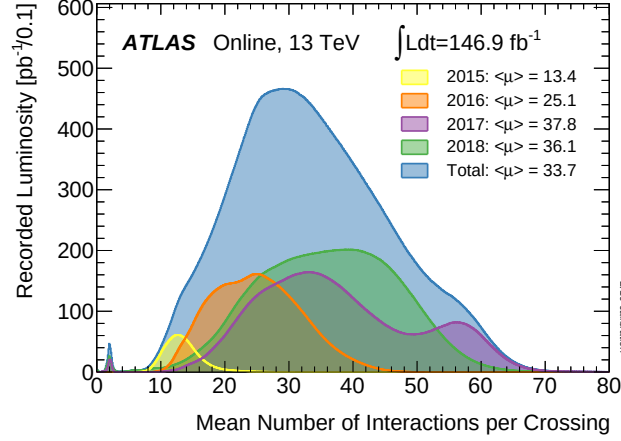


Figure 4.2.: Visualization of the interactions per bunch crossing for the Run 2 data taking period both accumulated and split in each year including the expected number of interactions [84].

4.2. Simulated Signal

The simulated signal is a scalar Leptoquark $\tilde{S}_1$ coupling only to an electron and a b -quark resulting in a branching fraction of 1. There is a total set of 24 different parameter combinations produced having masses from 1 to 3.5 TeV in steps of 0.5 TeV and couplings from 0.5 to 3.5 in steps of 1. The corresponding dataset identifiers (DSIDs) are 602449-602472. All cross sections are listed in table 4.2 with all other relevant information for each produced set of parameters. The highest cross section of $8.74 \cdot 10^{-2}$ fb has the sample with a mass of 1 TeV and a coupling of 3.5 and the lowest cross section is $5.18 \cdot 10^{-7}$ fb given by the 3.5 TeV, 0.5 coupling sample. The generators used to produce the signal are POWHEG [158, 159, 160] and HERWIG 7 [161] using a state-of-the-art resonant Leptoquark routine [76]. This differs from the phenomenological resonant Leptoquark search which used MADGRAPH [162] and PYTHIA 8 [163, 164] for the simulation [74]. The used lepton PDF has NLO precision and is called "LUXlep-NNPDF 31_nlo_as_0118_luxqed" (LUXlep) [47]. It is evolved out of the photon PDF due to the fact that the electrons produced in the proton originate from photons. Due to having NLO precision it is possible to have a second lepton in the final state (see section 2.3.2) which allows for searches in the 1 lepton and 2 lepton region.

4. Data and Monte Carlo Samples

Table 4.2.: Signal sample list used in this analysis including the mass, coupling, number of generated events surviving preselection cuts (see section 5.2), cross sections extracted from POWHEG, used generator and DSID values.

Mass [TeV]	Coupling	Preselection events	Cross section [pb]	Generator	DSID
1	0.5	55781	$1.42 \cdot 10^{-3}$	PowhegHerwig7	602449
	1.5	51171	$1.35 \cdot 10^{-2}$		602450
	2.5	42976	$4.09 \cdot 10^{-2}$		602451
	3.5	35076	$8.74 \cdot 10^{-2}$		602452
1.5	0.5	64898	$1.69 \cdot 10^{-4}$		602453
	1.5	59138	$1.71 \cdot 10^{-3}$		602454
	2.5	50361	$5.76 \cdot 10^{-3}$		602455
	3.5	41538	$1.36 \cdot 10^{-2}$		602456
2	0.5	66545	$3.01 \cdot 10^{-5}$		602457
	1.5	57844	$3.38 \cdot 10^{-4}$		602458
	2.5	46763	$1.30 \cdot 10^{-3}$		602459
	3.5	38819	$3.42 \cdot 10^{-3}$		602460
2.5	0.5	66184	$6.66 \cdot 10^{-6}$		602461
	1.5	53816	$8.92 \cdot 10^{-5}$		602462
	2.5	42430	$4.02 \cdot 10^{-4}$		602463
	3.5	34550	$1.18 \cdot 10^{-3}$		602464
3	0.5	64915	$1.75 \cdot 10^{-6}$		602465
	1.5	47604	$3.00 \cdot 10^{-5}$		602466
	2.5	36178	$1.58 \cdot 10^{-4}$		602467
	3.5	30752	$5.04 \cdot 10^{-4}$		602468
3.5	0.5	61899	$5.18 \cdot 10^{-7}$		602469
	1.5	39763	$1.21 \cdot 10^{-5}$		602470
	2.5	31370	$7.33 \cdot 10^{-5}$		602471
	3.5	27327	$2.46 \cdot 10^{-4}$		602472

4.3. Simulated Background

In order to mimic a signal event at least one highly energetic lepton and jet need to be identified. In order to do this the Feynman diagram needs a particle which creates a jet such as a gluon or a quark and a lepton (or a jet creating particles faking a lepton which is strongly suppressed).

The therefore considered background processes for this analysis are: W +jets, Z +jets, single top, top anti-top ($t\bar{t}$), dijet (Dijet when referring the sample), diboson semi- as well as fully leptonic, lepton-quark scattering and top anti-top with a vector boson ($t\bar{t}V$). The dominating background processes are W +jets, Z +jets, dijets and single top and $t\bar{t}$ which will be combined and referred to as "Top". The remaining processes will be referred to as "Others" in all plots to reduce cluttering because their combined contribution is still lower than a single dominating background.

The W +jets process is the dominating background when considering a final state with one lepton. The W boson has to decay leptonically into a charged lepton and its corresponding neutrino in order to produce the necessary lepton in order to fake a signal event. The corresponding Feynman diagrams are shown in figure 4.3.

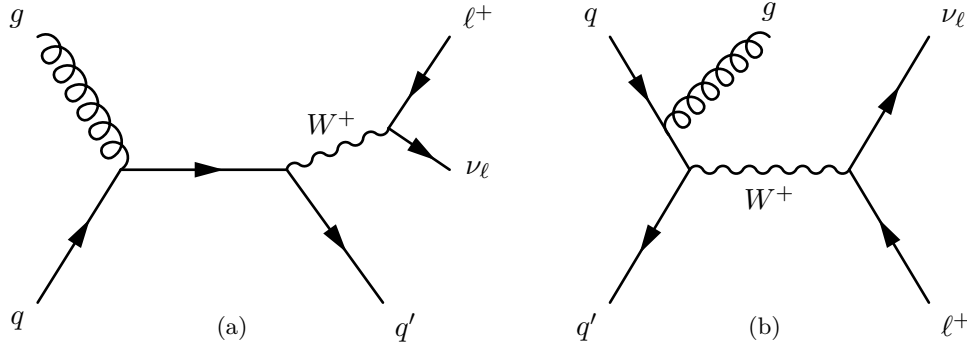


Figure 4.3.: Selection of contributing W +jets diagrams. W^- diagrams are also possible if the incoming particles are charge conjugated.

Z +jets is the dominating background when considering a final state with two leptons. A selection of contributing leading order Feynman diagrams is given in figure 4.4. In all of these the Z boson needs to decay leptonically in order to produce the required amount of leptons.

4. Data and Monte Carlo Samples

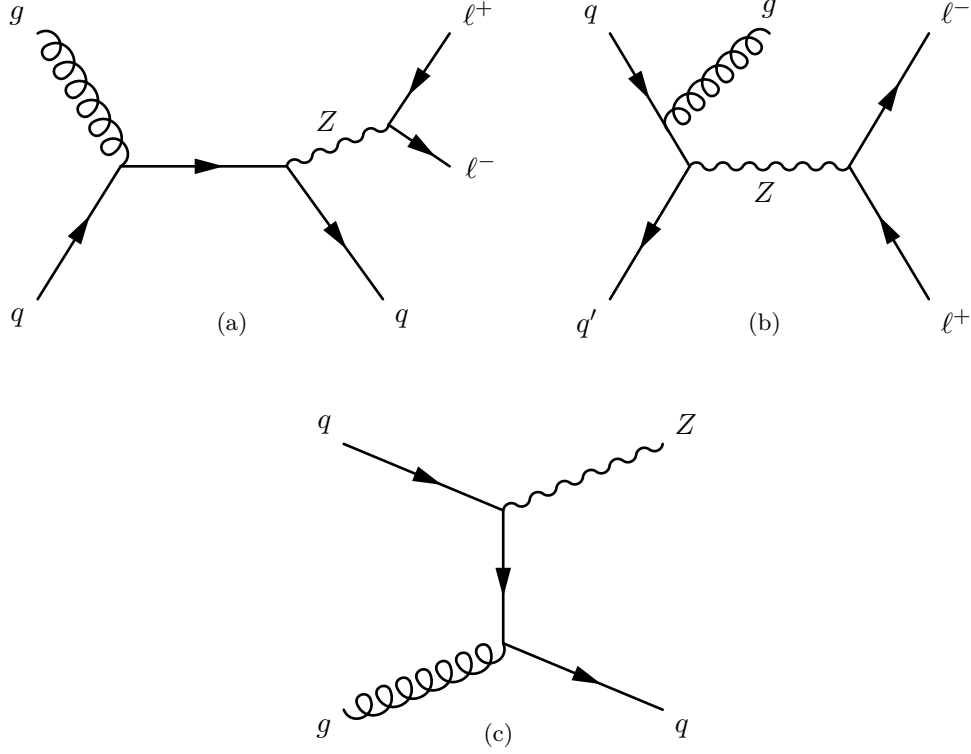


Figure 4.4.: Selection of contributing Z + jets diagrams.

The Top samples dominate the background at higher b -jet multiplicities. This is because a top quark almost exclusively decays into a b -quark and a W boson. The corresponding Feynman diagrams for a leptonic and hadronic W decay are shown in figure 4.5. Some of the contributing diagrams are shown in figure 4.6 where the single top s-channel and t-channel diagrams, the single top with an additional W boson and the $t\bar{t}$ diagrams are drawn.

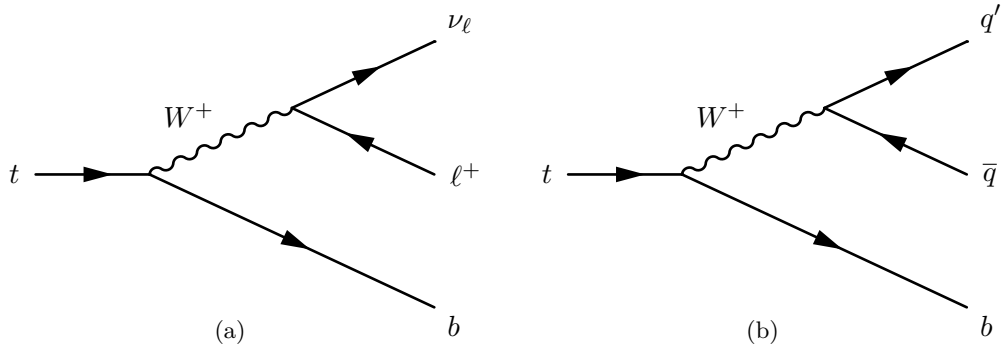


Figure 4.5.: Most common decay modes of the top quark with the W boson either decaying leptonically (a) or hadronically (b).

4. Data and Monte Carlo Samples

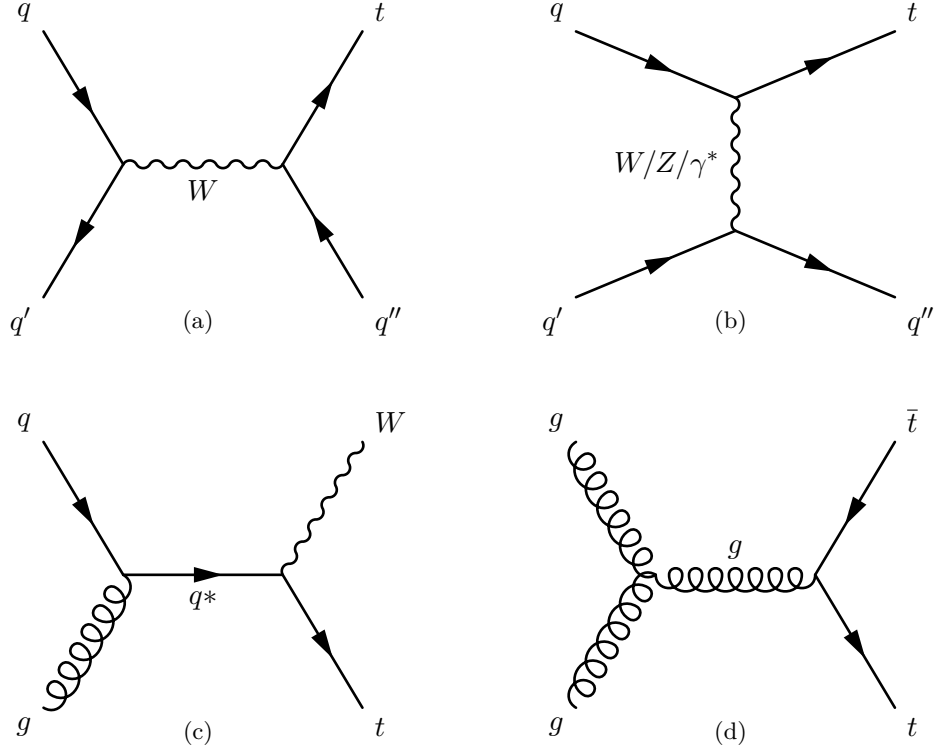


Figure 4.6.: Selection of contributing "Top" diagrams. This includes s-channel (a) and t-channel (b) productions of a single top quark which forbid q'' to be a top quark to not have two top quarks in the final state. Additionally, processes with a W boson in addition to a top quark are considered (c) and a $t\bar{t}$ diagram (d) is shown.

Dijet samples have the highest contribution to one lepton final states when not applying any isolation criteria on the lepton (see section 3.7). This is because of the very high cross section of this process due to the sheer amount of possible leading order diagrams with relatively high cross sections. One of the jets needs to be misidentified as a lepton from either a coincidental lepton flying in the same direction or a highly energetic lepton being created in the showering process. A small selection of possible Feynman diagrams is shown in figure 4.7.

The remaining processes include diboson decays where both decay leptonically (fully-leptonic) and cases where one decays hadronically and one leptonically (semi-leptonic). It also includes $t\bar{t}$ events with an additional vector boson ($t\bar{t}V$) and processes where a lepton scatters on a quark. One contributing Feynman diagram for each of these backgrounds is drawn in figure 4.8.

4. Data and Monte Carlo Samples

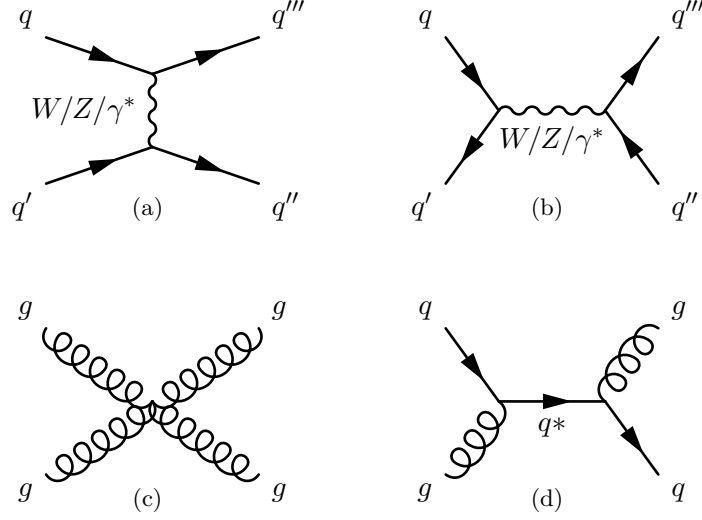


Figure 4.7.: Small selection of contributing dijet diagrams. One of the resulting jets is misidentified as a lepton. Depending on the exchange boson and channel in (a) and (b) the quarks either don't change flavour, differ in flavour or are their own anti particles.

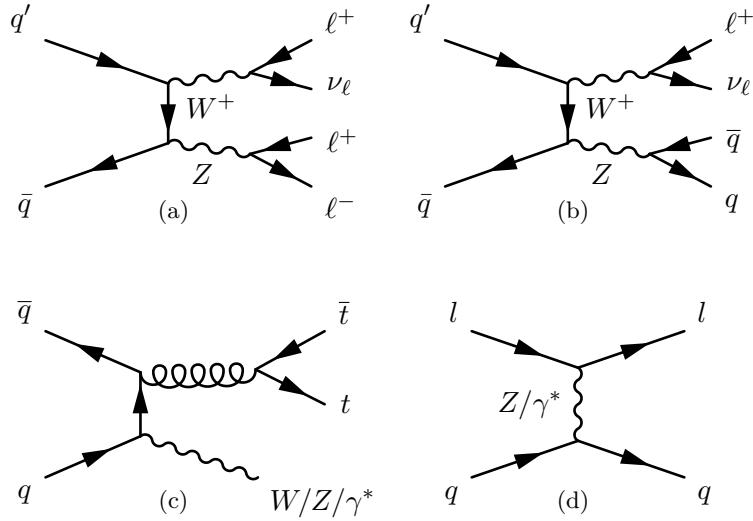


Figure 4.8.: One contribution diagram for each background part of the "Others" background. This includes diboson processes where both decay leptonically (a), dibosons with a semileptonic decay (b), $t\bar{t}$ in combination with a vector boson (c) and a lepton-quark scattering process (d).

4. Data and Monte Carlo Samples

All backgrounds follow the internal instructions of the Physics Modeling Group (PMG) recommendations [165] if possible. Table 4.3 summarizes all background samples naming the processes and including their respective number of generated events after applying preselection cuts (see section 5.2), cross sections, generators and DSIDs. The used generators for the backgrounds include SHERPA 2.2.11, 2.2.12, 2.2.14 and 2.2.2 [166, 167], PYTHIA 8.230 [163, 164], POWHEG BOX v2 [158, 159, 160], MADGRAPH 3.3.4 and MADGRAPH5_AMC@NLO 2.3.3 [162] and HERWIG 7.2.3 [161]. The used PDFs are "NNPDF2.3LO " [168] when using PYTHIA and "NNPDF3.0NLO " [45] when using other generators. The only exception being LUXlep for the lepton-quark scattering using the same PDF as the signal samples. Further information on the samples can be found in the internal note of the current resonant Leptoquark search of ATLAS [63].

Table 4.3.: Background sample list used in this analysis including the specific processes, number of generated events surviving preselection cuts (see section 5.2), cross sections, used generators and DSID values. EW and QCD for the diboson sample represent electroweak and quantum chromodynamic processes respectively.

Background	Process	Preselection events	Cross section [pb]	Generator	DSID
$W + \text{jets}$	$W \rightarrow e\nu$	23846711	$2.18 \cdot 10^4$	SHERPA 2.2.11	700338-700340
	$W \rightarrow \mu\nu$		$2.18 \cdot 10^4$		700341-700343
	$W \rightarrow \tau\nu$		$2.18 \cdot 10^4$		700344-700349
$Z + \text{jets}$	$Z \rightarrow ee$	33013657	$2.22 \cdot 10^3$	SHERPA 2.2.11	700320-700322
	$Z \rightarrow \mu\mu$		$2.22 \cdot 10^3$		700323-700325
	$Z \rightarrow \tau\tau$		$2.24 \cdot 10^3$	SHERPA 2.2.14	700792-700794
Top	s-channel	4094	$1.03 \cdot 10^1$	POWHEG BOX v2 + PYTHIA 8.230	410644, 410645
	t-channel	72066	$2.17 \cdot 10^2$		410658, 410659
	tW	71591	$7.17 \cdot 10^1$		601352, 601355
	$t\bar{t}$	545300	$7.30 \cdot 10^2$		410480, 410481
Dijet		41047	$1.91 \cdot 10^9$	PYTHIA 8.230	364701-364712
Diboson	Fully-leptonic EW	646226	$2.99 \cdot 10^{-1}$	SHERPA 2.2.12	700587-700590
	Fully-leptonic QCD		$2.17 \cdot 10^1$		700600-700605
	Semi-leptonic QCD	1869631	$5.09 \cdot 10^{-1}$	SHERPA 2.2.2	345705, 345706, 345718, 345723
			$1.63 \cdot 10^0$		364302-364305
			$7.50 \cdot 10^1$		SHERPA 2.2.11
$t\bar{t}V$		206951	$1.34 \cdot 10^0$	MADGRAPH5_AMC@NLO 2.3.3 + PYTHIA 9.230	410155-410157, 410218-410220
l - q scattering	$e + q$	144418	$1.44 \cdot 10^0$	MADGRAPH 3.3.4 + HERWIG 7.2.3	525888-525894
	$\mu + q$		$1.32 \cdot 10^0$		525895-525901

4.4. Fake Electron Estimation

Precisely simulating the possibility of the detector misidentifying a jet as an electron can be quite challenging using MC. To control the MC behaviour, a data-driven background estimation is used which uses measured collision data to estimate the background contribution of events containing fake electrons. This data-driven approach is only used for the case of one final state electron due to dijet being a diminishing background with two final state electrons. The specific approach used is called "asymptotic matrix method" [169] and needs two selection criteria for the electron, in

4. Data and Monte Carlo Samples

this analysis they are called "baseline" and "signal" electrons (see section 5.1). Signal electrons are a subset of baseline electrons fulfilling tighter requirements. Baseline electrons which are not part of the signal electrons are called "loose" electrons. The fraction of real electrons that are baseline electrons being also signal electrons is called real efficiency ϵ_r and the fraction of fake electrons that are baseline electrons being also signal electrons is called fake efficiency ϵ_f . They are therefore defined as:

$$N_r^s = \epsilon_r N_r^b, \quad (4.1)$$

$$N_f^s = \epsilon_f N_f^b, \quad (4.2)$$

where N_r^s and N_f^s are the numbers of real and fake electrons in the set of signal electrons and N_r^b and N_f^b are the numbers of real and fake electrons in the set of baseline electrons. The real efficiency can be extracted from MC simulations because MC simulations already use the real efficiency to correct differences in MC to data. The fake efficiency needs to be taken from data samples containing different events than the ones being estimated. Using both efficiencies, a matrix correlating the number of signal N^s and loose electrons N^l with the number of real baseline electrons N_r^b and fake baseline electrons N_f^b can be created. The matrix is

$$\begin{pmatrix} N^s \\ N^l \end{pmatrix} = \begin{pmatrix} \epsilon_r & \epsilon_f \\ 1 - \epsilon_r & 1 - \epsilon_f \end{pmatrix} \begin{pmatrix} N_r^b \\ N_f^b \end{pmatrix}, \quad (4.3)$$

which can be used to get the expression

$$N_f^b = \frac{1}{\epsilon_r - \epsilon_f} \left[(\epsilon_r - 1) N^s + \epsilon_r N^l \right]. \quad (4.4)$$

Using this in 4.2, the number of fake electrons in the set of signal electrons can be estimated. The efficiencies depend on several factors such as the electron momentum which leads to every event needing to be weighted differently. The applied weights are:

$$w_i = \begin{cases} \frac{\epsilon_{fi}}{\epsilon_{ri} - \epsilon_{fi}} (\epsilon_{ri} - 1), & \text{if signal electron} \\ \frac{\epsilon_{fi}}{\epsilon_{ri} - \epsilon_{fi}} (\epsilon_{ri}), & \text{if baseline electron} \end{cases}, \quad (4.5)$$

with ϵ_{ri} and ϵ_{fi} being the real and fake efficiency of the corresponding electron i . Due to ϵ_{ri} always being smaller than one, the weight of signal electrons is always negative. The total sum of weights is exactly the number of fake electrons being in the signal electron set:

$$N_f^s = \sum w_i. \quad (4.6)$$

4.5. Analysis Software

Throughout the analysis several different programs come into play. This section will go over all analysis steps of this analysis, name the used programs and claiming whether

4. Data and Monte Carlo Samples

the steps were part of this thesis or provided externally.

The production process of the MC sample up to the Ntuple level was not part of this thesis but are rather provided by the LQ resonant search group of the ATLAS collaboration using, among other programs, the internally available LJNtupleMaker program [170]. The used analysis release is AnalysisBase, version 25.2.2 and the derivation tag is p6026.

First looks at kinematics and root files were done using SWAN [171], a Jupyter notebook implementation with access to the CERN cloud which allows for both C++ (using g++ 13.1.0 as compiler) and Python 3 [172] code which were both used and tested. The extensive inspection of all kinematic distributions, testing of different cuts and optimization of these is done using a fully self written program from scratch in C++ in a modular design based on classes. The goal of this program is maximum flexibility in changing parameters for cutting and plotting decisions. It uses the ROOT library v.6.32.02 [155] to implement the input and output of ROOT files via the RDataFrame class. This program was used on LXPLUS [173], an interactive interface to use the computing clusters provided by CERN. Additionally, it was also modified to work remotely by sending jobs with HTCondor [174] to the computing clusters for calculations which took up to three days.

During the creation of this thesis this program was replaced by a internally available program named LJHistoMaker [175], being programmed in both Python and C++, to increase the visual quality of the plots and ensure similar looking plots compared to the ones created by the ATLAS resonant LQ group. This program was modified by adding some functionalities and modifying others. It was originally used on LXPLUS/SWAN but was later moved to "ATLAS Analysis Facility at UChicago" (UChicago) [176] because the added systematic uncertainties increased the files sizes of the MC samples greatly (roughly a factor of 40) which would not have fit in the available space on AFS and EOS, the file storages of LXPLUS and SWAN. Here again a mixture of interactive coding and notebook usage are used. This altered version of LJHistoMaker was used to create a majority of the plots in this thesis.

The statistical analysis is based on LJStatAnalysis, also internally available under [177]. This was originally created for tutorials, was already somewhat adjusted to fit an e-d channel resonant LQ search and then heavily changed to create the statistical analysis seen in this thesis. It was also originally used on LXPLUS/SWAN and later moved to UChicago. This program uses cabinetry [178, 179, 180] which is a wrapper of numerous libraries including additional functionalities of fits and visualizations. The hypothesis test itself uses pyhf [181, 182].

Note that the used internal links in this section are linking to programs used in a currently ongoing search and might provide utilities and files which were not available during the creation of this thesis.

5. Event Selection

This chapter goes into further detail on the selection of events which are considered in the statistical analysis (see chapter 6). To do this, first the different requirements on the reconstructed objects are stated in section 5.1, then the preselection cuts are shown in section 5.2. After that, to be able to motivate the decisions of the cuts, the kinematic distributions of the leading lepton and jet, the second most energetic lepton and the lepton-jet system with preselection are shown in section 5.3. In this context the different distributions are explained including the use of pseudorapidity for the leading jet. The following section 5.4 will treat the optimized cuts by showing N-1 plots. The region definitions and explaining the chosen Control Regions, Validation Regions and Signal Regions (CRs, VRs, SRs respectively) is done in section 5.5. The programs used are discussed in previous section 4.5.

5.1. Object Selection

This section will go into further detail on the object selection by stating the attributes an object has to have to be considered in the analysis. Additionally, the working points applied to the objects are mentioned. This list is split into the different objects reconstructed being the different lepton types and jet types.

5.1.1. Electron Selection

For electrons to be considered in this analysis they need to fulfil the baseline requirements in table 5.1. These requirements can be split into categories: kinematic requirements, vertex requirements and working point requirement. The kinematic requirements apply cuts on p_T and η , where the vertex requirements make sure that the track to vertex association fulfils criteria for $z_0 \sin \theta$ and $d_0/\sigma(d_0)$. Finally, the working point requirement chooses different identification and isolation working points. If the tighter signal criteria is fulfilled the electron is considered to be a baseline as well as a signal electron. Further information on the electron selection can be found on the internal ATLAS recommendations page in reference [183].

5. Event Selection

Table 5.1.: Selection criteria for baseline and signal Electrons.

Feature	Baseline Criterion	Signal Criterion
p_T	$> 10 \text{ GeV}$	$> 20 \text{ GeV}$
$ \eta $	< 2.47	< 2.47
$ z_0 \sin \theta $	$< 0.5 \text{ mm}$	$< 0.5 \text{ mm}$
$ d_0/\sigma(d_0) $	-	< 5
Identification working point	LooseBL	Tight
Isolation working point	-	HighPtCaloOnly

5.1.2. Muon Selection

Muons follow similar selection criteria as electrons. The specific requirements necessary for the muons to be considered baseline or signal muons are summarized in table 5.2. The kinematic requirements are mostly identical compared to the electron requirements apart from its slight increase in η coverage. The vertex requirements in transversal direction are tighter for signal muons. The defined working points are different and apply different algorithms for identification and isolation. For more information on the muon selection the internal recommendations page [184] can be looked up.

Table 5.2.: Selection criteria for baseline and signal Muons.

Feature	Baseline Criterion	Signal Criterion
p_T	$> 10 \text{ GeV}$	$> 20 \text{ GeV}$
$ \eta $	< 2.5	< 2.5
$ z_0 \sin \theta $	$< 0.5 \text{ mm}$	$< 0.5 \text{ mm}$
$ d_0/\sigma(d_0) $	-	< 3
Identification working point	High-pT	High-pT
Isolation working point	-	PflowTight_VarRad

5.1.3. Jet Selection

The jet selection is listed in table 5.3 and deviates quite a lot from the lepton selections. The first thing is that there is no baseline object defined and all considered jets are signal jets. Secondly, the working points and track to vertex association requirements are switched for other types of requirements. The kinematic requirements on p_T and η are identical to the requirements for the leptons. If the b -tagging requirement is additionally fulfilled the identified jet is considered a b -jet. These criteria are based on the ones applied for the official resonant LQ search currently done by ATLAS as the same samples are used. More information on that can be found in the internal note of the ongoing LQ search [63].

5. Event Selection

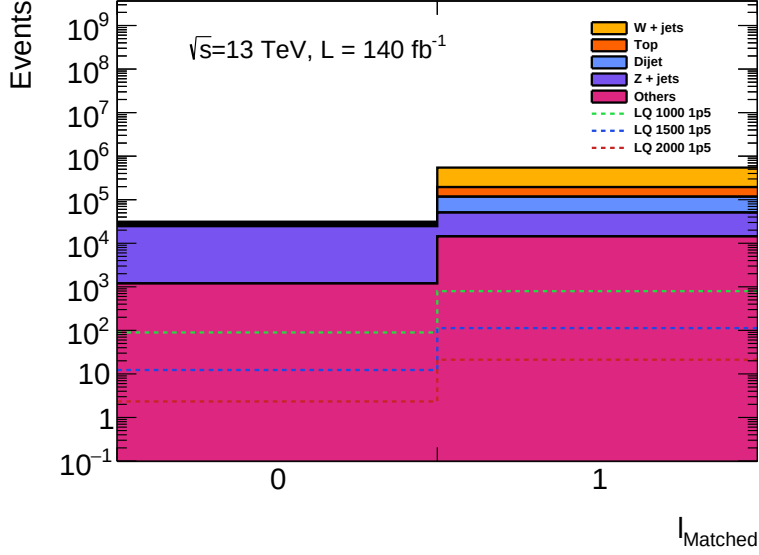
Table 5.3.: Selection criteria for Jets. For b -jets the b -tagging requirement additionally applies.

Feature	Signal Criterion
p_T	$> 20 \text{ GeV}$
$ \eta $	< 2.5
Clustering Algorithm	Anti- k_t
Radius R	0.4
Pileup mitigation	JVT FixedEffPt
b -tagging	DL1dv01 FixedCutBEff 77%

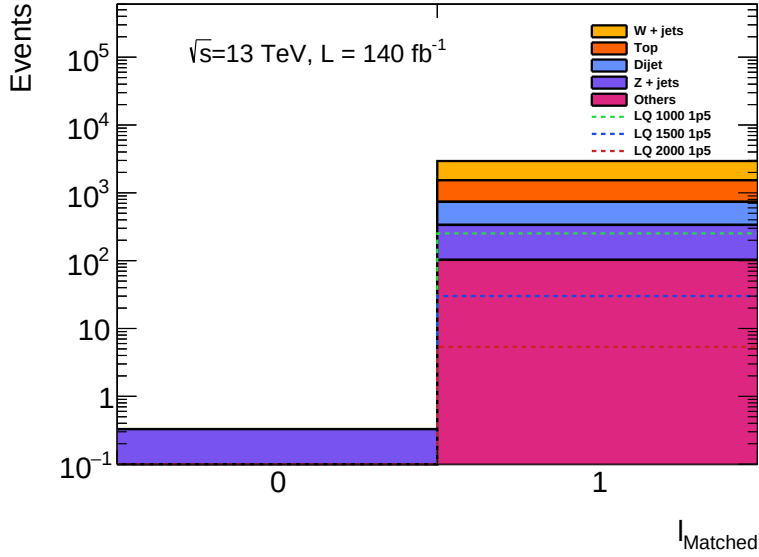
5.2. Preselection

The preselection cuts include cuts to ensure a optimal working point and cuts reducing file sizes by cutting phase space which is not needed for this search. The event is required to have one or two same flavoured baseline and the same amount of signal leptons (this allows for potential low energy leptons below the baseline threshold), $m_{\ell j} > 700$ because regions below that will not contribute to the statistical analysis, the leading lepton and jet both are required to have a $p_T > 130 \text{ GeV}$ and the event must pass the single-lepton triggers mentioned in table 4.1. Additionally, the offline reconstruction is matched to the trigger signature by requiring $+1 \text{ GeV}$ on the offline (reconstructed) p_T . This ensures that the triggers are used in their respective plateau regions by following the trigger signature group recommendation [185]. In the dataset used in this analysis the 2 lepton (2L) events were not correctly filled and the Boolean responsible for this matching is always set to 0. This should have a negligible impact due to the p_T requirement already applied on the leading lepton which is above most trigger thresholds. Additionally, the constraints on the lepton kinematics which will be elaborated on in section 5.4 will reduce the impact further. Figure 5.1 visualizes the matching Boolean for the 1 lepton (1L) region with only the precuts and all cuts introduced in section 5.4. With only precuts applied matching removes roughly 5% of the events and with all cuts applied matching removes less than one event which corresponds to the removal at the order of 0.01% of total events. The behaviour of the matching is expected to be the same for the 2L case.

5. Event Selection



(a) Only preselection applied.



(b) All cuts applied defined in section 5.4.

Figure 5.1.: The plots show the Boolean determining whether the lepton is matched or not for the 1 lepton region. Both plots contain all backgrounds and some selected signal models with a coupling of 1.5 and do not have a selection on the matching done. The top plot shows the preselection and bottom plot contains all cuts.

5. Event Selection

Applying only this preselection already removes most of the dijet events, which without them would be the dominating background, by having tighter requirements on the lepton kinematics. Figure 5.2 shows the number of signal leptons with preselection before applying matching and a cut on signal leptons.

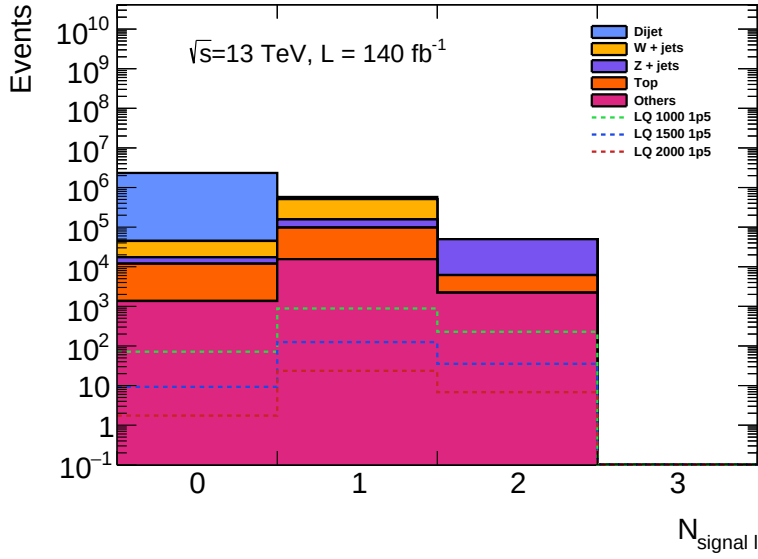


Figure 5.2.: Visualization of the number of signal leptons after preselection before applying restrictions to number of signal leptons and matching. The legend is sorted for the background which means that the dijet background contains the most amount of events.

Table 5.4 gives a short summary of the applied cuts which are referred to as precuts throughout this thesis.

Table 5.4.: Applied cuts which are considered precuts and always applied except if explicitly stated otherwise. The leptons have the same flavour.

Feature	$N_{\text{signal } \ell}$	$m_{\ell j}$ [GeV]	$p_T^{\ell 1}$ [GeV]	$p_T^{j 1}$ [GeV]	Pass HLT	Trigger Matching
Cut value	1 or 2	> 700	> 130	> 130	Yes	Yes

5.3. Kinematic Distributions

To have a general understanding of the kinematics involved and to get an idea for the possible cuts to apply the kinematics have to be inspected. The important objects to look at here are the leading lepton, the leading jet and the lepton-jet system. These Kinematics are shown with all backgrounds modeled and a set of signal models

5. Event Selection

which shows the mass as well as the coupling dependence. There is a total of five LQ models, three of them have masses with 1, 1.5 and 2 TeV and a coupling of 1.5. The other two also have a mass of 1.5 TeV and couplings of 0.5 and 2.5.

5.3.1. Leading Lepton

Starting with the leading lepton figures 5.3 to 5.6 show the energy, p_T , ϕ and η distributions. There are no requirements on the lepton flavour at preselection level so these plots include muons for the backgrounds. The energy distributions for the background follow a falling distribution. The argumentation for the general falling nature of all energy, p_T and mass distributions is the falling differential cross section at higher values which occurs from a shrinking phase space. The energy distributions for the signals rises until it reaches half of the LQ mass and then also follows a continuously falling distribution. The peak at half the LQ mass can be explained by that fact that it is a two-body decay. This means in the rest frame of the LQ the decay products are back-to-back and gain half of the momentum each to hold 4-momentum conservation.

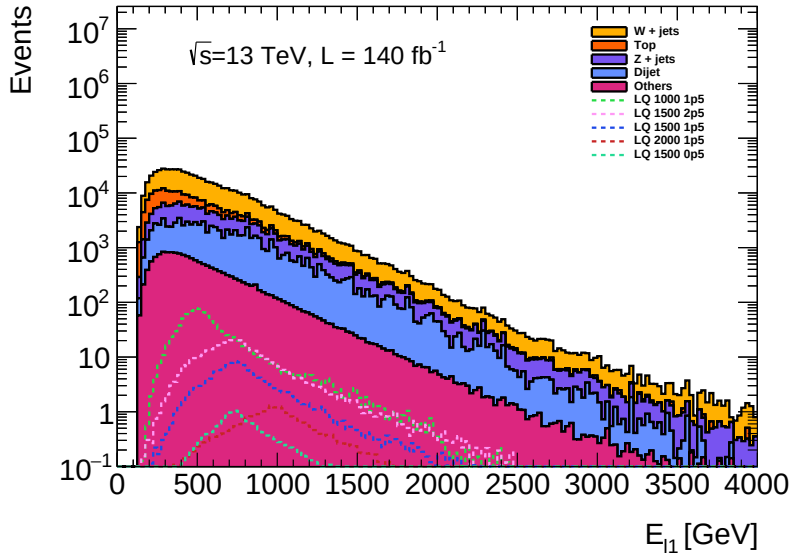


Figure 5.3.: Energy distribution of the leading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

The p_T distribution visualized in figure 5.4 shows a falling distribution for the background. The signal follows a Jacobian distribution with the Jacobian peak again

5. Event Selection

at half the LQ mass. The shape of the Jacobian can be explained using this short calculation:

$$|p| = \frac{M_{LQ}}{2} \quad (5.1)$$

$$p_T = |p| \cdot \sin(\theta) = \frac{M_{LQ}}{2} \cdot \sin(\theta). \quad (5.2)$$

These two relations are already known. The first one comes from 4-momentum conservation in the rest-frame and the second one was already discussed in section 3.2. To get the differential cross section in terms of p_T one defines:

$$\frac{d\sigma}{dp_T} = \frac{d\sigma}{d\cos\theta} \cdot \frac{d\cos\theta}{dp_T} \quad (5.3)$$

To be able to calculate $\frac{d\cos\theta}{dp_T}$, first $\cos\theta$ needs to be expressed in terms of p_T . This is done by using the Pythagorean identity $\sin^2(\theta) + \cos^2(\theta) = 1$. Solving 5.2 for $\sin(\theta)$ and squaring it gives

$$\sin^2(\theta) = \left(\frac{2p_T}{M_{LQ}} \right)^2. \quad (5.4)$$

By adding $\cos^2(\theta)$ and solving for $\cos(\theta)$ expresses $\cos(\theta)$ in terms of p_T

$$\cos(\theta) = \pm \sqrt{1 - \left(\frac{2p_T}{M_{LQ}} \right)^2} = \pm \frac{2}{M_{LQ}} \cdot \sqrt{\left(\frac{M_{LQ}}{2} \right)^2 - p_T^2}. \quad (5.5)$$

Now $\frac{d\cos\theta}{dp_T}$ can be calculated which is:

$$\frac{d\cos\theta}{dp_T} = \mp \frac{2p_T}{M_{LQ}} \cdot \frac{1}{\sqrt{\left(\frac{M_{LQ}}{2} \right)^2 - p_T^2}}. \quad (5.6)$$

If one now combines equation 5.3 and 5.6 the result is:

$$\frac{d\sigma}{dp_T} = \frac{d\sigma}{d\cos\theta} \cdot \frac{2p_T}{M_{LQ}} \cdot \frac{1}{\sqrt{\left(\frac{M_{LQ}}{2} \right)^2 - p_T^2}}. \quad (5.7)$$

The positive sign is chosen to describe a positive cross section yielding a positive amount of events. Equation 5.7 clearly shows that the cross section maximizes for $p_T = \frac{M_{LQ}}{2}$ and then has a steep drop off.

The ϕ and η distribution of the leading lepton (figures 5.5 and 5.6) behave as expected. The process is independent of ϕ apart from the barrel-endcap crack in the electromagnetic calorimeter (see section 3.4) visible in the phi distribution. The background is almost independent of η in the observed range and the signal has a small decline towards higher absolute values due to the fact that a massive LQ is less likely to have a highly longitudinally boosted lepton. Neither the mass, nor the rapidity is interesting to look at for the lepton due to it being highly relativistic.

5. Event Selection

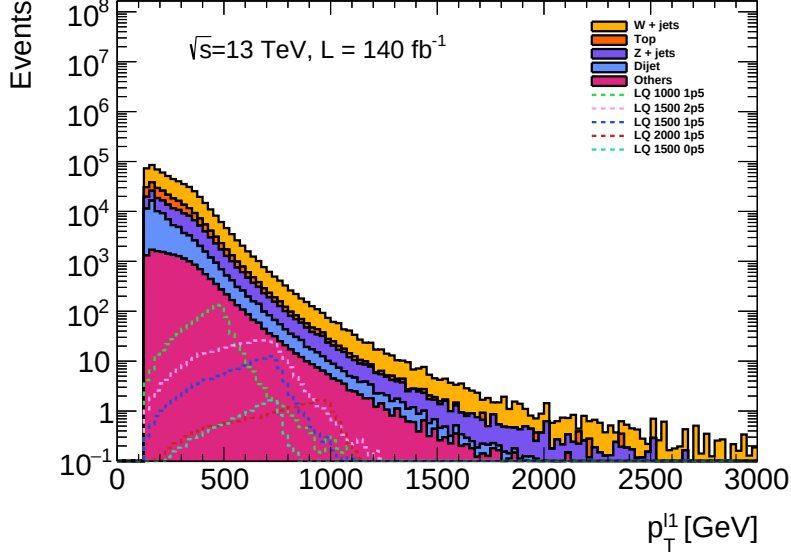


Figure 5.4.: p_T distribution of the leading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

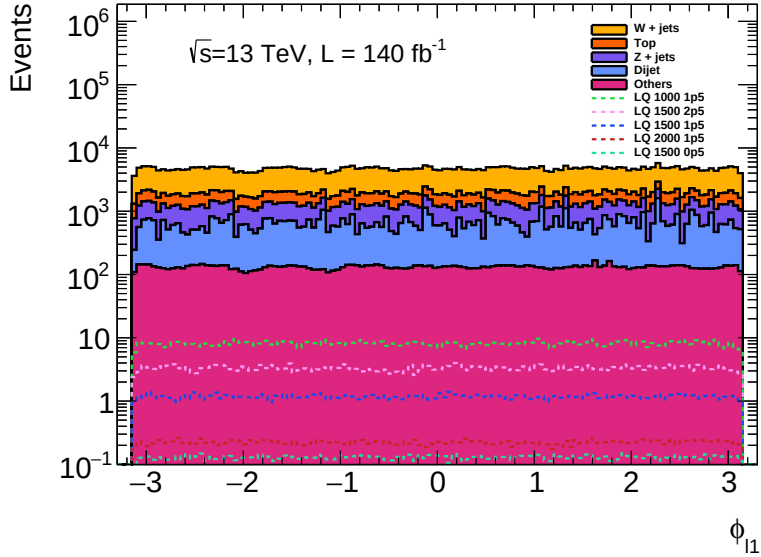


Figure 5.5.: ϕ distribution of the leading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

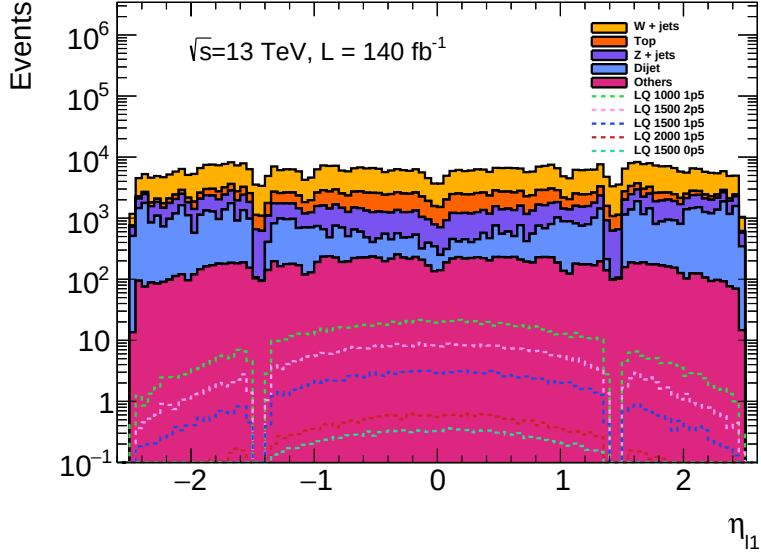


Figure 5.6.: η distribution of the leading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5.3.2. Leading Jet

The jet distributions of the energy, p_T, ϕ and η are shown in figures 5.7 to 5.10 and are very similar to the lepton distributions apart from some smearing effects in the energies due to the more difficult object reconstruction and the disappearing dip at the barrel-endcap crack position in the electromagnetic calorimeter. This can be explained by realizing that the jet deposits most of its energy in the hadronic calorimeter and that the spatial dimensions of the jet are larger. The dijet peaks in the ϕ and η distributions are due to single large weight events. These events are not of any concern because they will be removed by applying the cuts introduced in section 5.4 so they will not contribute in any way to the statistical analysis.

5. Event Selection

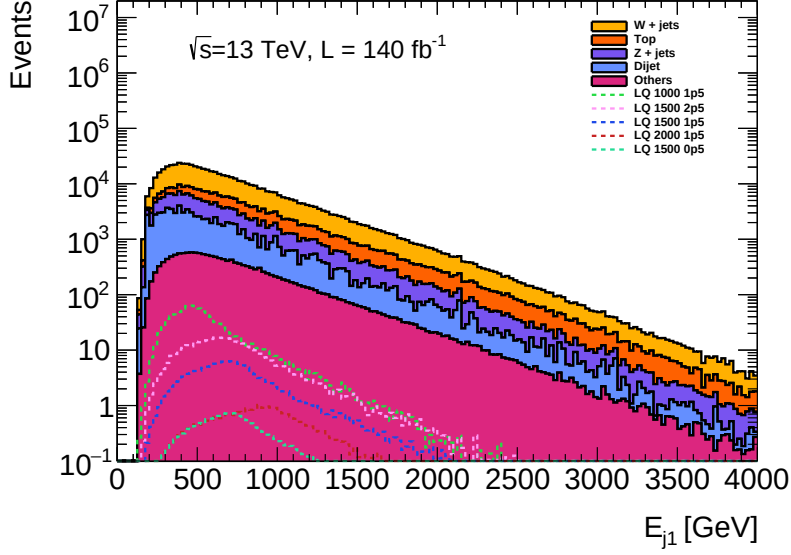


Figure 5.7.: Energy distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

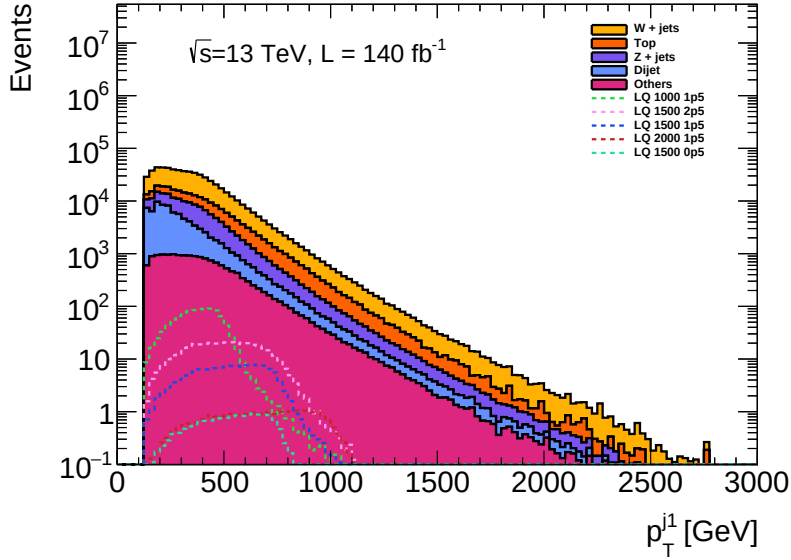


Figure 5.8.: p_T distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

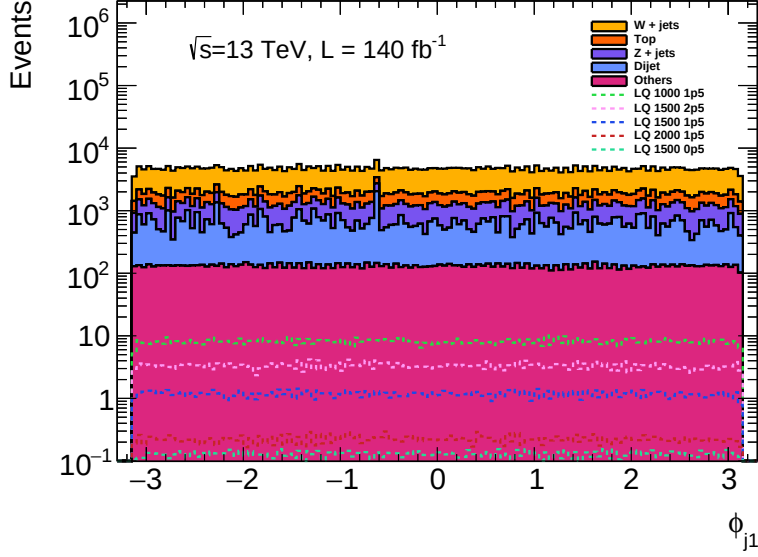


Figure 5.9.: ϕ distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

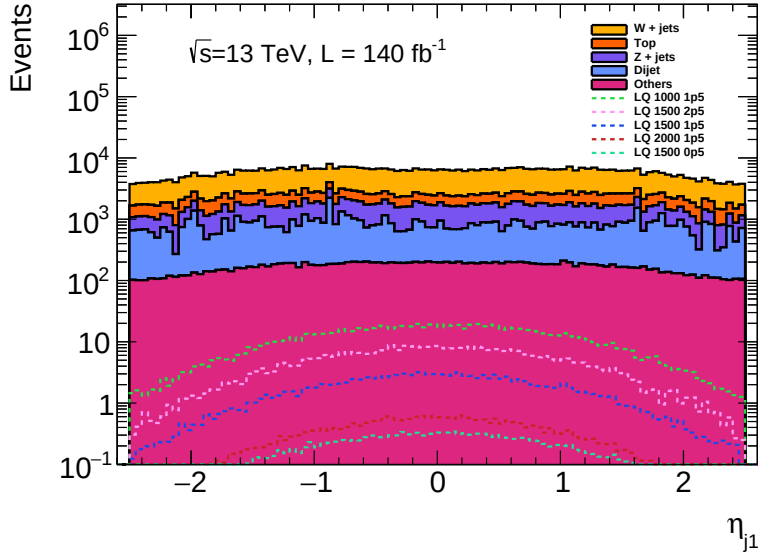


Figure 5.10.: η distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

For the leading jet it is worth looking at the mass distribution and the difference between the rapidity y and the pseudorapidity η . The mass distribution is given in figure 5.11 and there are two important properties to mention. The first being the relatively high mass going up to the hundreds of GeV range, which seems large enough to possibly have an impact on the rapidity. At closer inspection there are comparably few events left at high masses. For example the number of events at 100 GeV already is at 10% of the maximum value and 0.1% at 200 GeV. Nonetheless the rapidity will be checked for completeness. The second interesting property are the peaks of the backgrounds at the W mass ($m_W \approx 80.37$ GeV [19]) combined with a Z mass peak ($m_Z \approx 91.19$ GeV [19]) and a peak at the top mass ($m_Z \approx 172.6$ GeV [19]). These peaks originate from the hadronic decay of a top quark or a boson which lets them propagate their mass into the jet. This does not lead to a peak in the W +jets and Z +jets background because there the bosons have to decay leptonically in contrary to the diboson background (which is combined in "Others"). The top samples show a Top and W peak because a top can decay into a W , which again could decay hadronically.

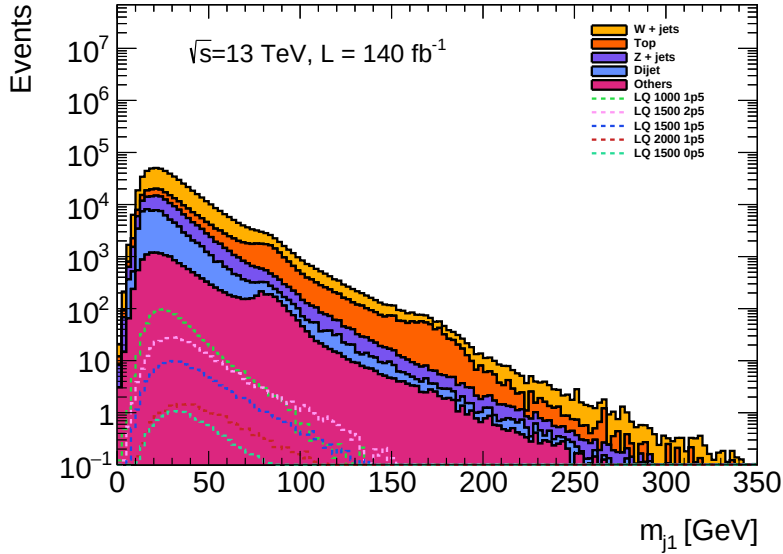


Figure 5.11.: Mass distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

When now looking at the rapidity in figure 5.12 it looks indistinguishable from the pseudorapidity distribution without overlaying it. To still visualize the discrepancies the absolute difference of η and y $\Delta(\eta y) = |\eta - y|$ is shown in figure 5.13. There, one can see that the difference does not exceed 0.035 with more than one event. So it is very appropriate to keep using the pseudorapidity instead of the rapidity especially if one considers that the upcoming cuts in the following section will reduce the number of jets with lower transverse momentum. Again, the single bin peaks in the dijet sample originate from single events with large weights which are removed as a side effect of the cuts described in section 5.4.

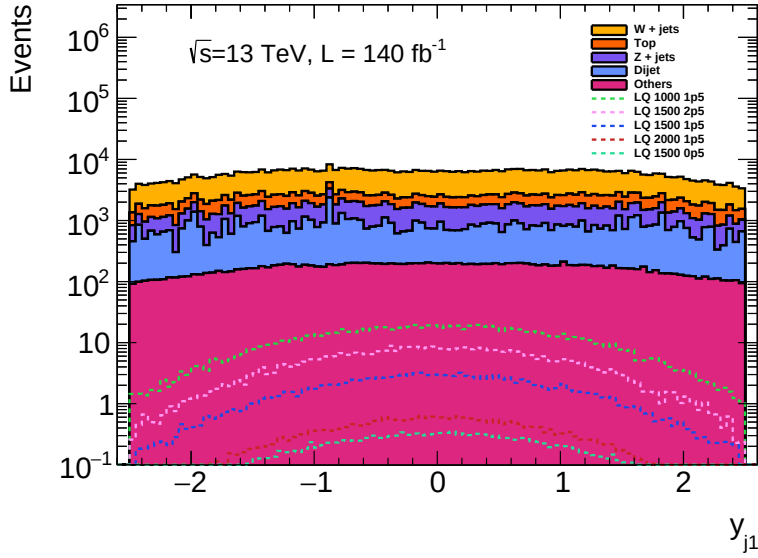


Figure 5.12.: Rapidity distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

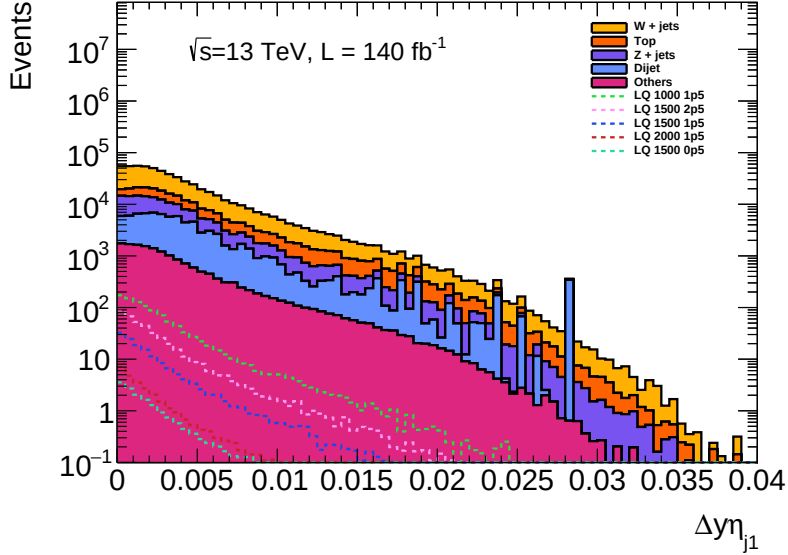


Figure 5.13.: $\Delta(\eta_j)$ distribution of the leading jet. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5.3.3. Lepton-Jet System

Now that the leading lepton and the leading jet have been looked at the lepton-jet system can be investigated considering only leading particles. Here there are more kinematic variables possible to investigate. In addition to all the ones tackled before there are also the spatial differences of the producing particles, namely $\Delta\phi$, $\Delta\eta$ and ΔR (see section 3.2 for definition).

Starting with the energy distribution in figure 5.14 one can see that the background again follows a falling distribution and the signal samples peak at their corresponding LQ masses and then also follow a falling distribution.

The p_T distribution in figure 5.15 has a similar shape for signal and background becoming ever less likely with higher p_T .

The ϕ distribution in figure 5.16 for the lepton-jet system shows, as expected, no dependence on ϕ . The η distribution, seen in figure 5.17, on the other hand completely differs from the other η distributions that were shown before. This is because the LQ (or rather the invariant mass of the lepton-jet system) is very heavy and can not be considered relativistic. Therefore the pseudorapidity is no relevant quantity and should not be used in the lepton-jet system.

5. Event Selection

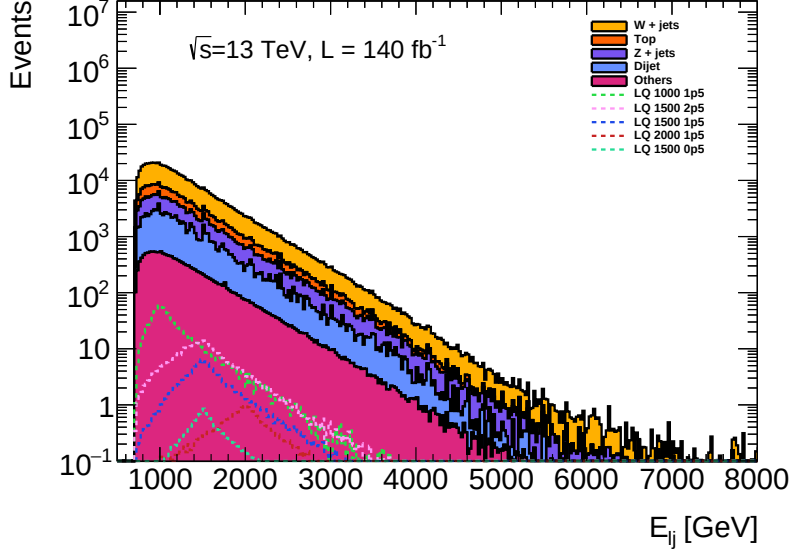


Figure 5.14.: Energy distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

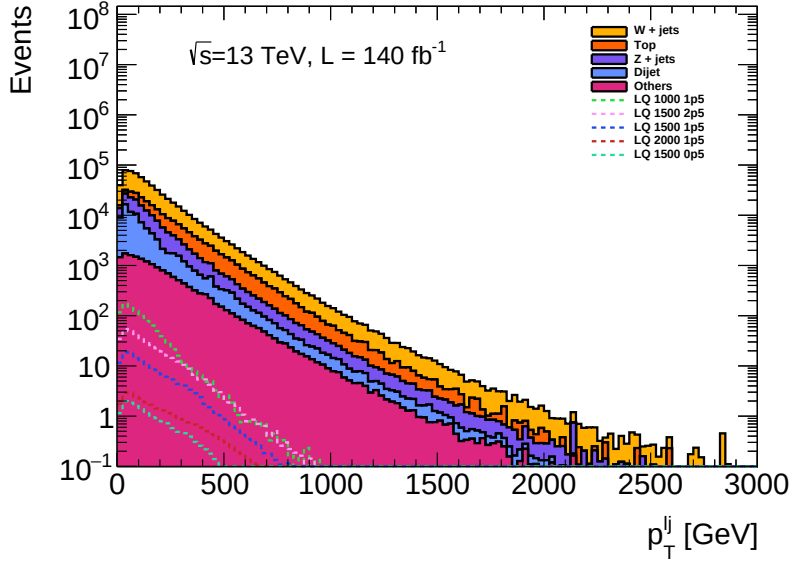


Figure 5.15.: p_T distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

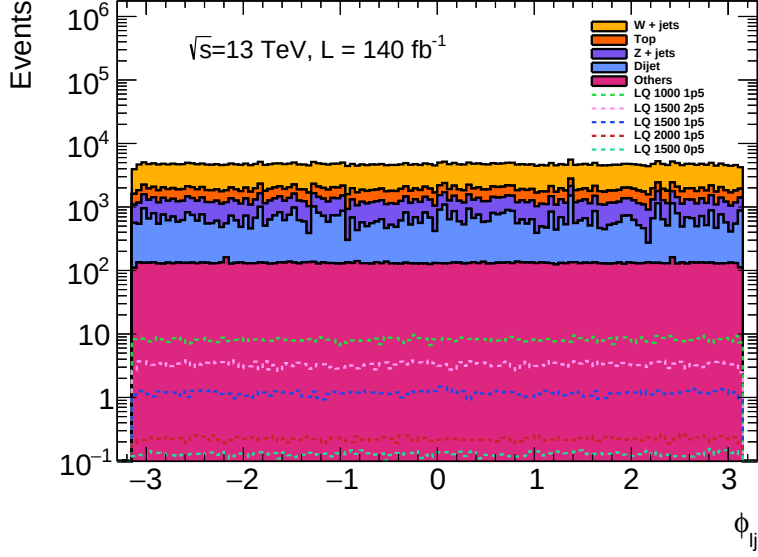


Figure 5.16.: ϕ distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

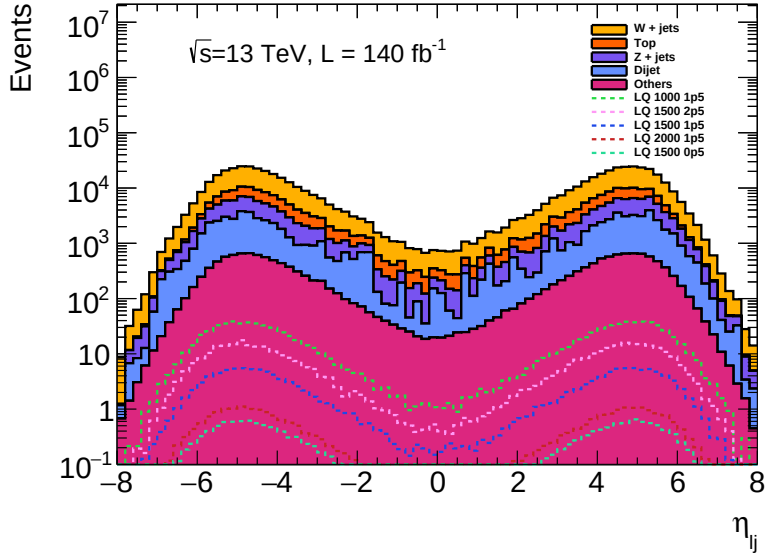


Figure 5.17.: η distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

The mass of the lepton-jet system is shown in figure 5.18. There it becomes really clear that the LQ masses are very high because in most cases, they are far greater than the p_T of the lepton-jet system in each event. The background follows, once again, a falling distribution. The signal on the other hand has a unique shape having multiple constituents contributing. In general a resonance would be fully described by a relativistic Breit-Wigner distribution using the resonant mass and the decay width

$$\Gamma = \frac{\lambda^2}{16\pi} \cdot m_{LQ}. \quad (5.8)$$

The Breit-Wigner distribution is symmetric around the mass, which is not the case in this study. There are two main effects creating this change of shape. The first one being final state radiation (FSR) reducing the energy of the measured particle due to parts of the energy being not considered. The second effect is the probability of the two partons producing the LQ having the required energy. This is called the parton-parton luminosity function and is essentially a folding of the two PDFs used. The simplified conclusion to take out of this is that for higher energies there are less available momentum fraction combinations possible to reach the required energy, which makes them less likely to occur. One more thing to take from this plot is that this is the first plot where the coupling significantly changes the shape instead of just increasing the number of events. The reason for that is that the coupling increases the decay width which combined with the FSR and parton-parton luminosity effect almost removes the resonant signature for high coupling (e.g LQ with 1.5 TeV mass and 2.5 coupling in figure 5.18). This has a strong impact on the analysis because it strongly impacted the signal region definitions outlined in section 5.5 by optimizing the region size depending on the expected coupling sensitivity.

The last property to look at for the lepton-jet system is the real rapidity to replace the pseudorapidity which is shown in figure 5.19. Here, the U-shape is not visible in the η distribution. It still differs from all previous η/y distributions due to the fact that the heavy system does not allow for very high rapidity because this would require a boost in beam direction which is kinematically strongly suppressed.

The distributions of the subleading lepton were also investigated, but they do not yield any relevant information for this thesis. The energy and p_T follow falling distributions for all samples and the angular distributions are similar to the leading lepton distributions. The figures supporting these statements can be found in appendix A.1.1.

5. Event Selection

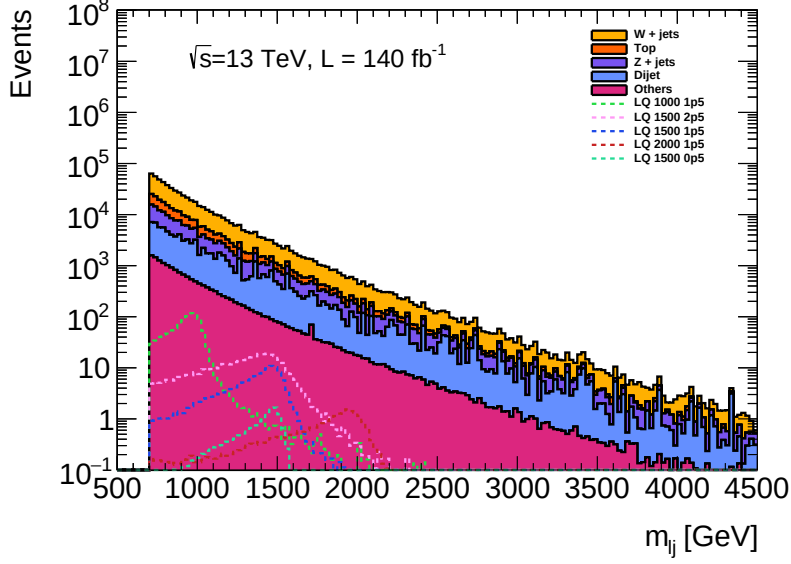


Figure 5.18.: Mass distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

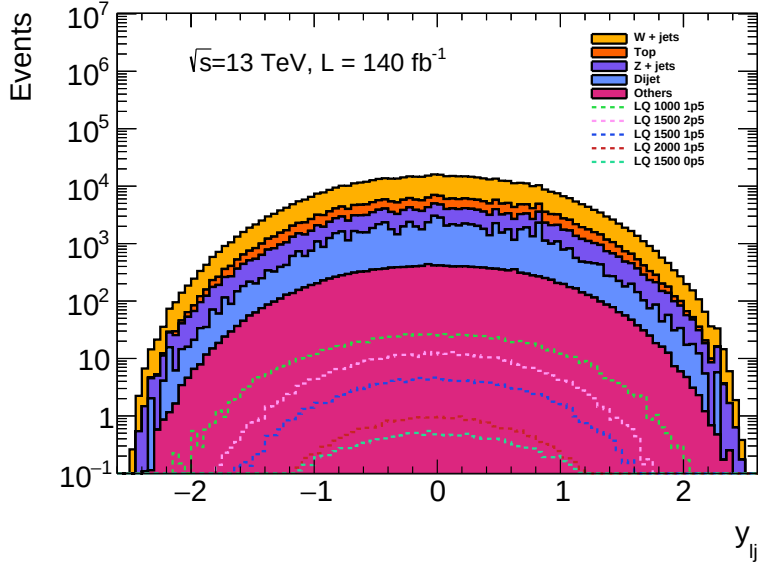


Figure 5.19.: Rapidity distribution of the lepton-jet system. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

5.4. Cuts

Now that all relevant kinematic distributions have been looked at and have been understood it is possible to apply sophisticated cuts to maximize significances. For this not only the kinematic properties but also the physical processes of background events have been considered (see section 4.3), e.g. that the W +jets background always produces a neutrino.

All following plots in this section are so called "N-1" plots which means all cuts are applied apart from the one explicitly talked about. They also have a cut of $m_{\ell j} > 900$ to only look at the signal regions defined in the next section. Additionally, the significances are plotted as well. They describe the significance of each bin separately and do not correspond to the final significances calculated during the statistical analysis but are just to be taken as a guideline to be able to roughly estimate the best cut. The used formula is:

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)} \right] \right)}, \quad (5.9)$$

which is defined in the ATLAS internal reference [186] and holds for $n \geq b$ and is the case because the signal model increases the number of expected events. n is the total number of events including the signal and b is the number of background events. σ is the uncertainty on the number of background events which is estimated to be $\sigma = \sqrt{(0.2b)^2 + \sqrt{b}^2}$ with $(0.2b)^2$ being a flat systematic uncertainty and $\sqrt{b}^2$ being the statistical uncertainty.

As discussed in the previous section, the shape of the kinematics only slightly change when changing the LQ models mass or coupling apart for varying peaks in the energy and mass distribution of the lepton-jet system as well as the energy and p_T distributions of the leading lepton and jet. Using this observation, to simplify the optimization of the cuts it is done for a 2 TeV 1.5 coupling LQ sample. The cuts are optimized by getting the best cut for each variable after preselection. The best cut is chosen to be the one which yields the highest significance in the 2 TeV signal region (see section 5.5). Then the cuts are iterated by applying N-1 plots and again extracting the best cut. This is repeated until no further changes in cuts appear. Lastly it is confirmed that these cuts also work very well for other LQ models with varying masses and couplings. The following plots use five different quark models, all with a coupling of 1.5 and with masses ranging from 1 TeV to 3 TeV.

5.4.1. Number of Leptons

Before looking at specific cuts the analysis is split into the one lepton (1L) and two leptons (2L) region with exactly 1 signal lepton and exactly 2 signal leptons respectively. This allows for independent optimization of the regions which are then recombined for the exclusion limit. As seen in figure 5.20, for the 1L region the W +jets in combination with the Top background dominate and in the 2L region the Z +jets

5. Event Selection

in combination with again the Top background dominate. This separate optimization allows for changing the cuts which mainly remove $W + \text{jets}/Z + \text{jets}$ in one of the regions to maximize statistics.

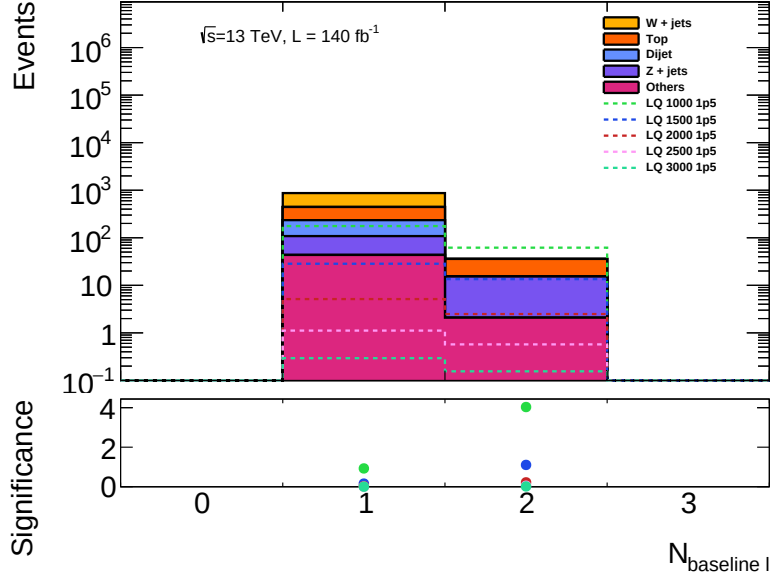
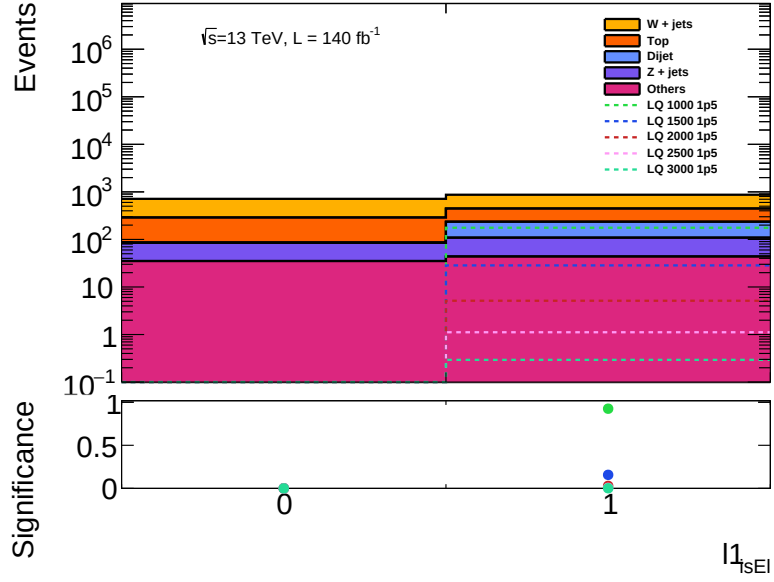


Figure 5.20.: Number of signal leptons after applying $N-1$ cuts. Both the background and signal events are sorted by the total number of events separately.

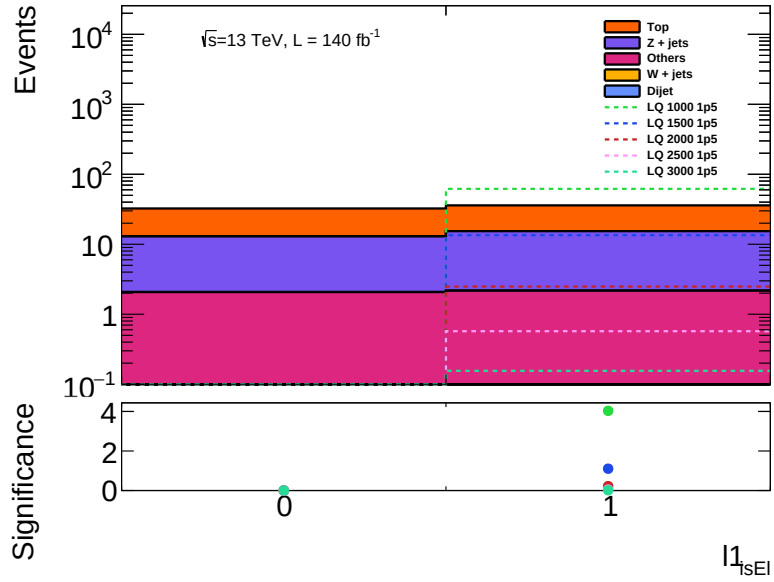
5.4.2. Lepton Flavour

One of the most obvious cuts is requiring that the leptons involved are always electrons and not muons. The signal model allows only for coupling to electrons and not to muons, the background sample did allow for muon interactions which makes this a lossless cut. This is visualized in figure 5.21 for the 1L and 2L region separately.

5. Event Selection



(a) 1L distribution.



(b) 2L distribution.

Figure 5.21.: Boolean whether the first (and second by definition) lepton is an electron after applying N-1 cuts for the 1L and 2L region. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

5.4.3. Number of b -jets

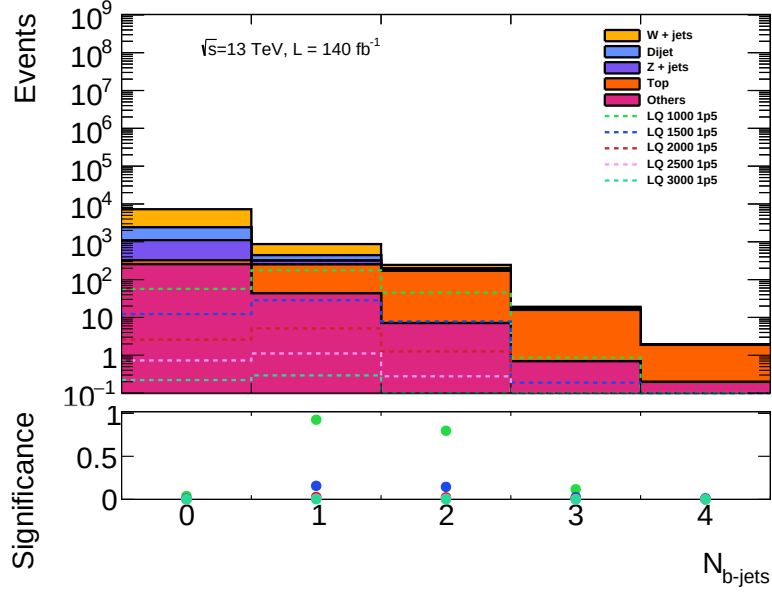
The next cut is a cut on the number of b -jets which serves two main purposes. On one hand, having b -tagging reduces all background not containing any b -quarks. This mainly reduces backgrounds which can produce any type of quark without favouring b -quarks but does not reduce the backgrounds containing top quarks because they almost exclusively decay via a b -quark. On the other hand, requiring a maximum amount of b -jets suppresses the Top background which dominates the high b -jet regions. The number of b -jets for each event is shown in figure 5.22 for the 1L and 2L region separately.

It was tested whether allowing for one or two b -jets would increase significance but it turned out that allowing only one b -jet is the best option.

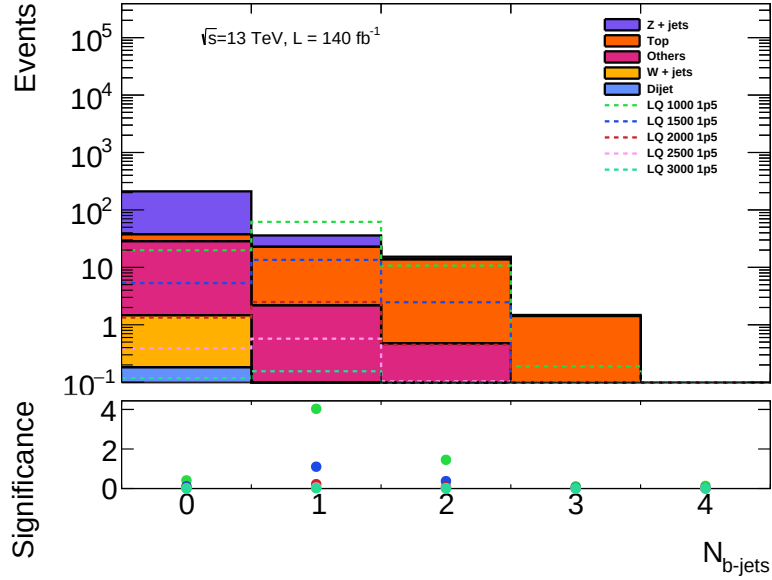
5.4.4. Significance of Missing Transverse Energy

Using a cut on $\mathcal{S}(p_T^{\text{miss}})$ arises from the idea that there are no neutrinos originating from the primary vertex of the signal model but every W decay produces neutrinos. This greatly reduces the W +jets as well as the Top background because top quarks decay into a W which produces a neutrino when decaying leptonically. The $\mathcal{S}(p_T^{\text{miss}})$ distribution is shown in figure 5.23 for the 1L and 2L region where the optimal cut differs. The chosen cut for the 1L region is $\mathcal{S}(p_T^{\text{miss}}) < 2.5$ and for the 2L region it is chosen to be $\mathcal{S}(p_T^{\text{miss}}) < 4$ to maintain enough statistics in the 2L region.

5. Event Selection



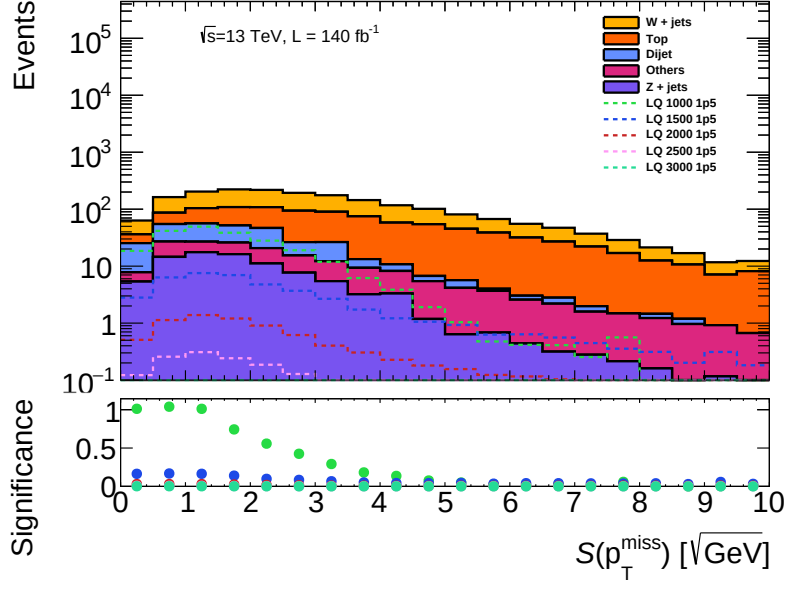
(a) 1L distribution.



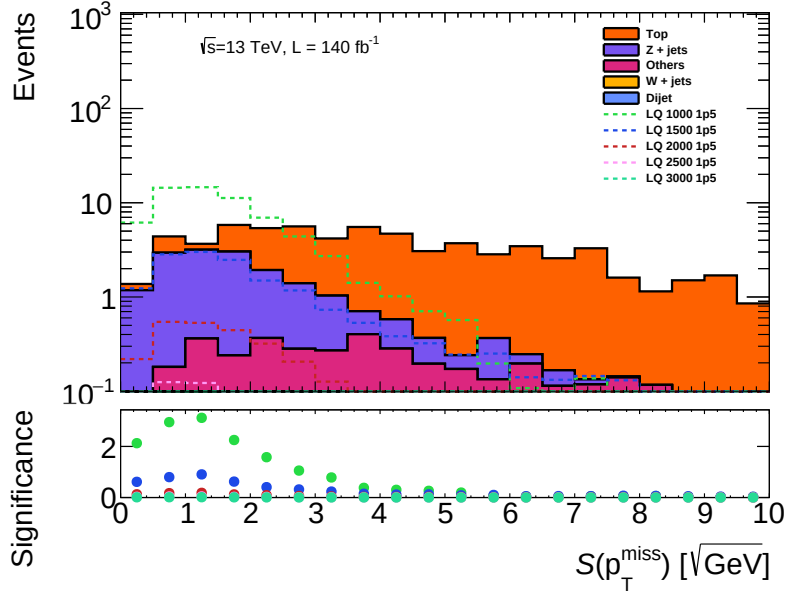
(b) 2L distribution.

Figure 5.22.: Number of b -jets after applying N-1 cuts for the 1L and 2L region. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection



(a) 1L distribution.



(b) 2L distribution.

Figure 5.23.: $\mathcal{S}(p_T^{\text{miss}})$ distribution after applying N-1 cuts for the 1L and 2L region. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

5.4.5. Mass of Lepton-Lepton system

If an event has two leptons it is possible to calculate the invariant mass of the lepton-lepton system. This is done in figure 5.24 for the 2L region because the 1L does not have the second lepton required to produce this plot. Due to the fact that the Z boson decays into two same flavoured leptons when decaying leptonically, the Z -peak is clearly visible in this distribution for the Z + jets background but also all the other backgrounds that can contain a Z boson, namely the $t\bar{t}V$ and Diboson backgrounds, contained in the Others background. These backgrounds are the backgrounds reduced when applying this cut. A possible cut would be a window cut around the Z mass but it was actually found that it is optimal to just apply a lower cut of $m_{\ell\ell} > 300$ GeV.

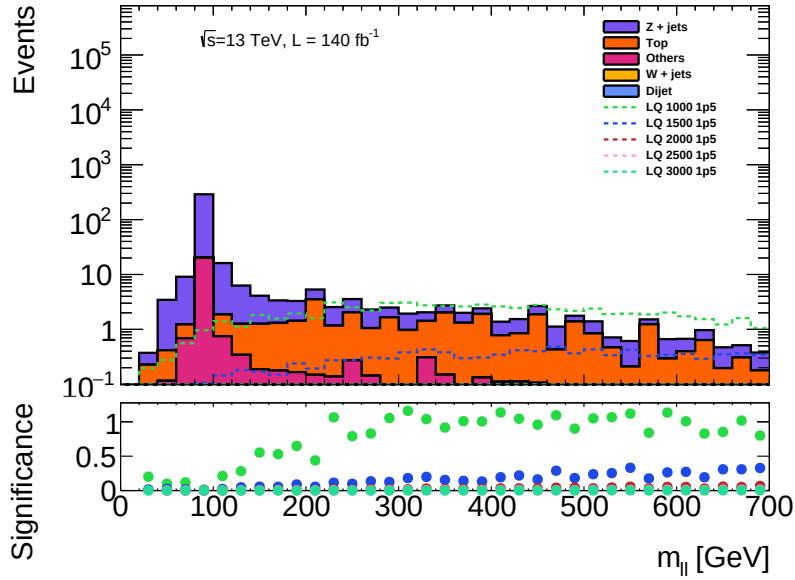


Figure 5.24.: $m_{\ell\ell}$ distribution after applying N-1 cuts for the 2L region. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

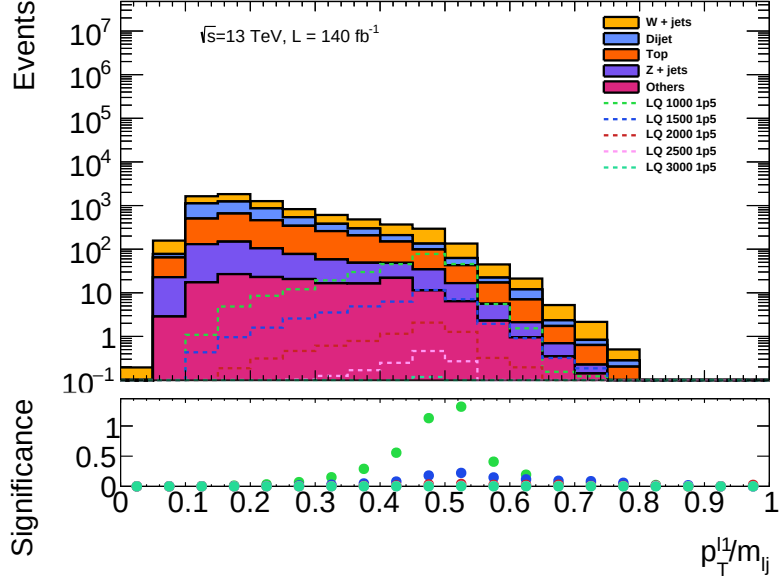
5.4.6. Ratio of Transverse Momentum of Leading Lepton and Mass of Lepton-Jet System

The usage of the ratio $p_T^{\ell 1}/m_{\ell j}$ instead of a simple cut on $p_T^{\ell 1}$ allows for a dynamic cut depending on the invariant mass in the lepton-jet system. This ensures that also for high masses the leading lepton possesses a large fraction of the total energy of the lepton-jet system. As a reminder, in the rest frame of the LQ the electron has half of the energy of the LQ mass. Applying possible FSR, smearing and boosting to lab frame effects results in the distributions shown in figure 5.25. This cut does not reduce any background in particular but rather all backgrounds. The final cuts are $p_T^{\ell 1}/m_{\ell j} > 0.4$ in the 1L region and $p_T^{\ell 1}/m_{\ell j} > 0.3$ in the 2L region to, once again, maintain high enough statistics to maximize the significance gained. It was also tried to apply upper cuts or a similar cut on the leading jet but it did not yield any improvements so they were dropped.

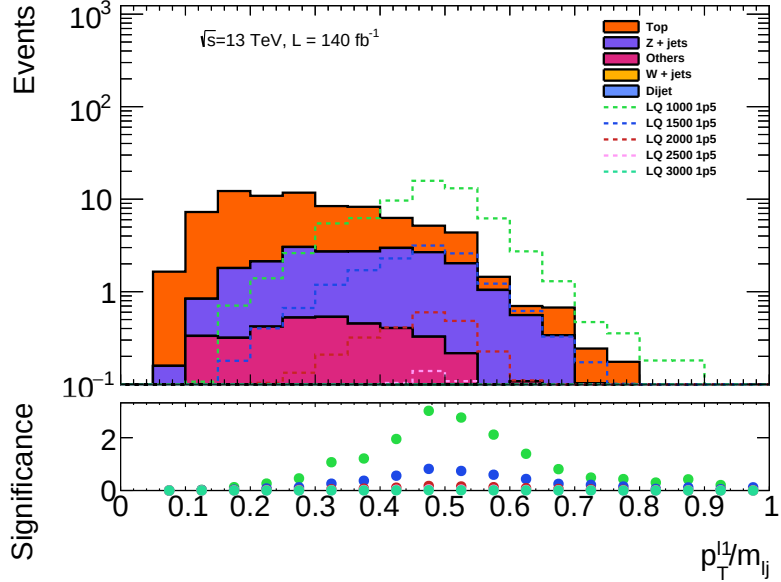
5.4.7. Distance Between Leading Lepton and Jet in Collider Coordinates

A cut on $\Delta R(\ell j)$ was not used in this analysis. It was tested thoroughly and it was applied throughout the most part of the work for this thesis but it turned out that removing it and tweaking the other cuts slightly yields better results. The reason for initially introducing this cut is that the LQ is comparably heavy and is therefore only weakly boosted. In the rest frame the LQ decays back-to-back which gives $\Delta R(\ell j) = 2\pi$, because the difference in ϕ is 2π and the difference in eta is zero. By weakly boosting it most of the events are expected to be in a window around 2π . The reason why this did not improve the exclusion limit is because it uses the same physical reasoning as the cut on $p_T^{\ell 1}/m_{\ell j}$ which makes these cuts very strongly correlated. Figures 5.26 and 5.27 visualize this by showing $\Delta R(\ell j)$ with and without the cut on $p_T^{\ell 1}/m_{\ell j}$ for the 1L region and the 2L region. The shape of the background changes greatly and becomes very similar to the signal shape after applying the cut.

5. Event Selection



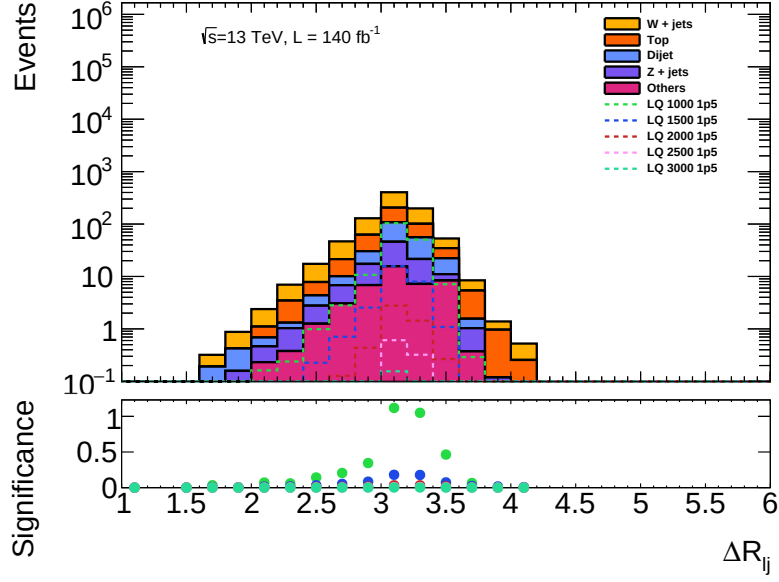
(a) 1L distribution.



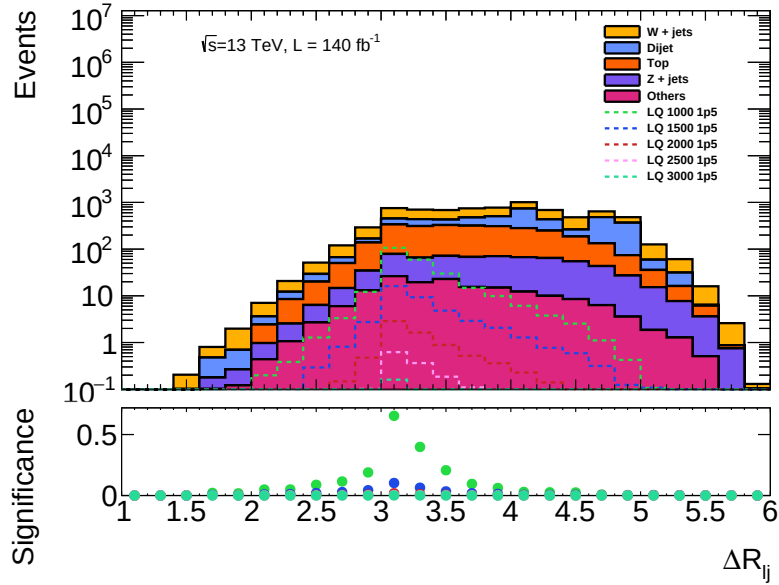
(b) 2L distribution.

Figure 5.25.: p_T^{l1}/m_{lj} distribution after applying N-1 cuts for the 1L and 2L region. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection



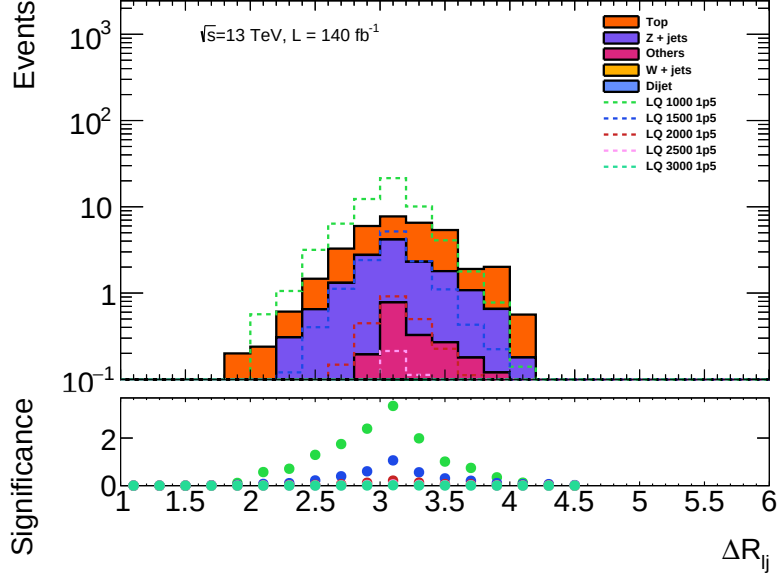
(a) 1L distribution with a cut on $p_T^{\ell 1}/m_{\ell j}$ applied.



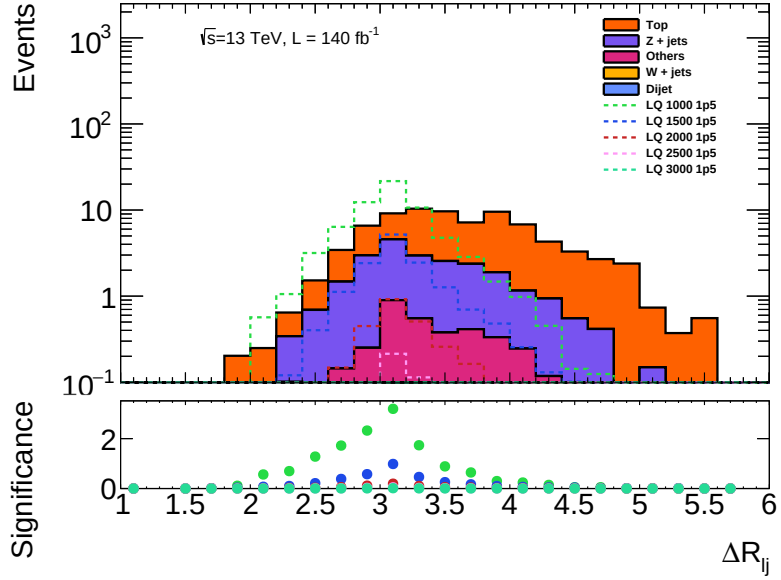
(b) 1L distribution without a cut on $p_T^{\ell 1}/m_{\ell j}$ applied.

Figure 5.26.: $\Delta R(\ell j)$ distribution after applying N-1 cuts for the 1L region with and without a cut on $p_T^{\ell 1}/m_{\ell j}$. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection



(a) 2L distribution with a cut on $p_T^{\ell 1}/m_{\ell j}$ applied.



(b) 2L distribution without a cut on $p_T^{\ell 1}/m_{\ell j}$ applied.

Figure 5.27.: $\Delta R(\ell j)$ distribution after applying N-1 cuts for the 2L region with and without a cut on $p_T^{\ell 1}/m_{\ell j}$. Both the background and signal events are sorted by the total number of events separately.

5. Event Selection

5.5. Region Definitions

The above mentioned cuts are maximizing the sensitivity to the existence of a LQ with the simulated masses. These cuts will be applied to the different signal regions (SR). There are multiple SRs defined to be sensitive to a larger range of masses as they are dominant in very different $m_{\ell j}$ ranges which was already shown in 5.18. The $m_{\ell j}$ regions need to increase in size for higher masses due to the decay width being related to the mass linearly and the coupling quadratically as shown in equation 5.8. It was also made sure that the bin size always stays larger than the energy resolution. Additionally, for higher masses the exclusion sensitivity lowers and the regions were therefore optimized for higher couplings. The cuts for the SRs can be found in tables 5.5 and 5.6 for the 1L and 2L region. In addition to a SR for each mass hypothesis that was produced there are also intermediate regions, namely the 1.2 TeV, 1.7 TeV, 2.2 TeV and 2.7 TeV regions. The purpose for these is to cover the whole $m_{\ell j}$ region but they are not optimized for such LQ masses. If there would be any excess in events measured within these regions they would have to be readjusted.

Table 5.5.: 1 lepton signal regions (SR-1L) definitions.

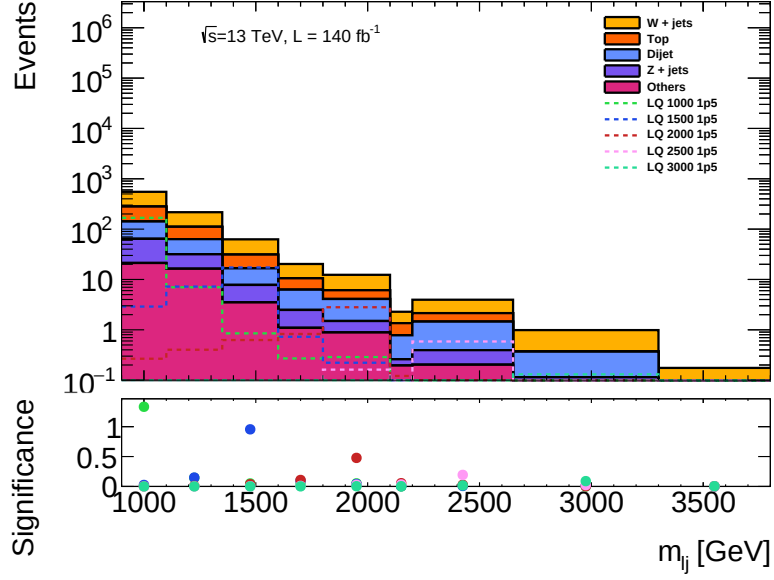
Feature	SR-1L-1TeV	SR-1L-1p2TeV	SR-1L-1p5TeV	SR-1L-1p7TeV	SR-1L-2TeV	SR-1L-2p2TeV	SR-1L-2p5TeV	SR-1L-3TeV	SR-1L-Overflow
$m_{\ell j}$ [GeV]	[900, 1100]	[1100, 1350]	[1350, 1600]	[1600, 1800]	[1800, 2100]	[2100, 2200]	[2200, 2650]	[2650, 3300]	> 3300
$N_{\text{baseline } \ell}$					= 1				
$N_{\text{signal } \ell}$					= 1				
$N_{b\text{-jets}}$					= 1				
$\mathcal{S}(p_{\text{T}}^{\text{miss}})$ [GeV $^{\frac{1}{2}}$]					< 2.5				
$p_{\text{T}}^{\ell 1}/m_{\ell j}$					> 0.4				

Table 5.6.: 2 lepton signal regions (SR-2L) definitions.

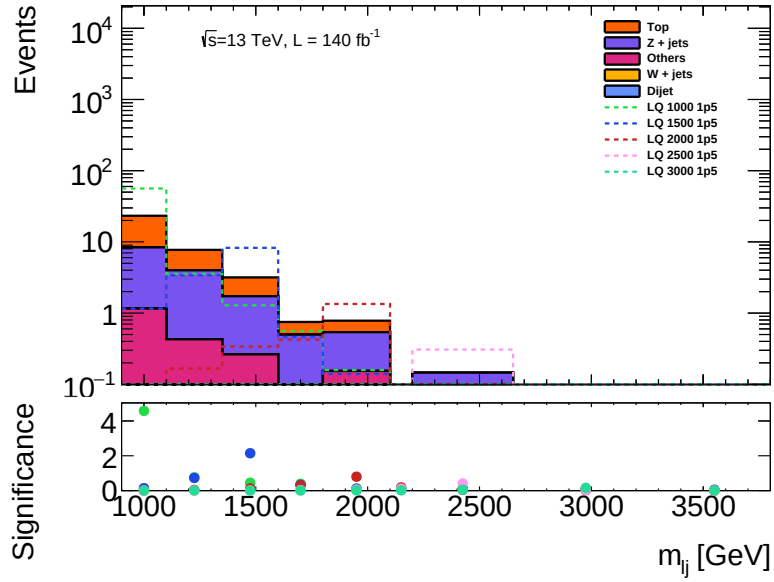
Feature	SR-2L-1TeV	SR-2L-1p2TeV	SR-2L-1p5TeV	SR-2L-1p7TeV	SR-2L-2TeV	SR-2L-2p2TeV	SR-2L-2p5TeV	SR-2L-3TeV	SR-2L-Overflow
$m_{\ell j}$ [GeV]	[900, 1100]	[1100, 1350]	[1350, 1600]	[1600, 1800]	[1800, 2100]	[2100, 2200]	[2200, 2650]	[2650, 3300]	> 3300
$N_{\text{baseline } \ell}$					= 2				
$N_{\text{signal } \ell}$					= 2				
$N_{b\text{-jets}}$					= 1				
$m_{\ell\ell}$ [GeV]					> 300				
$\mathcal{S}(p_{\text{T}}^{\text{miss}})$ [GeV $^{\frac{1}{2}}$]					< 4				
$p_{\text{T}}^{\ell 1}/m_{\ell j}$					> 0.3				

Figure 5.28 visualizes the defined SRs by showing the $m_{\ell j}$ distribution with the new binning. There it becomes clear that the corresponding mass hypothesis has the highest significance in its designated bin which was intended. Additionally, the 2L region has a way higher significance due to the background becoming roughly one order of magnitude smaller when the signal samples only decreases by 50%. Nonetheless the 2L region provides way less statistics which will affect the statistical analysis negatively.

5. Event Selection



(a) 1L SR.



(b) 2L SR.

Figure 5.28.: m_{lj} pre-fit distribution in the 1L and 2L SR showing the signal to background significance.

5. Event Selection

Other types of regions to define are the control regions (CR) and the validation regions (VR). Both of them are orthogonal to the SR and to each other. These regions serve the purpose of correcting the MC simulation in respect to some sort of mismodeling by introducing normalization factors. To get these normalization factors the number of events in the measured data is compared to the number of simulated events in the CR and then set to be equal by having normalization factors for the major backgrounds in the SR. These major backgrounds can be identified in the pie chart of figure 5.29 which shows the composition of the whole 1L SR and the whole 2L SR. The major backgrounds are: W +jets, Z +jets, and Top. Dijet could maybe also be considered a major background in the 1L region but for dijet it is difficult to find a CR which has a good purity, enough dijet events and few other events. Also, it should be sufficient to absorb this normalization by the W +jets and Top normalization factor.

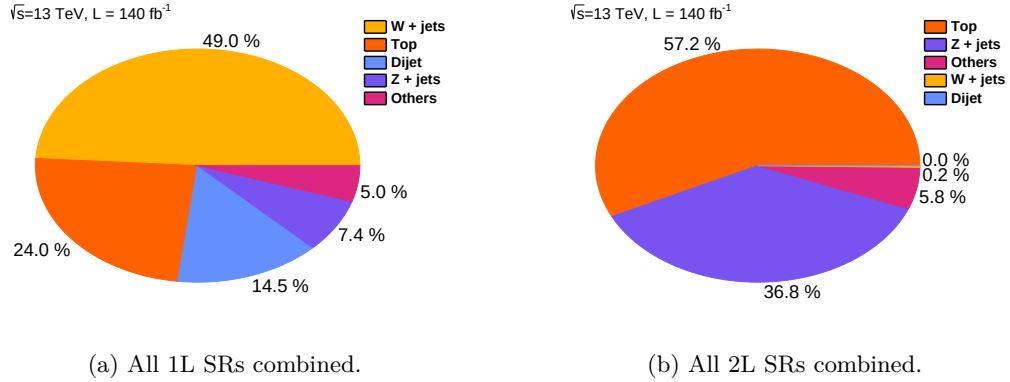


Figure 5.29.: Pie charts of the background composition in the 1L and 2L SR. The mass dependent SRs are combined here by setting the cut $900 < m_{\ell j} < 3900$ GeV. All cuts from section 5.4 are applied but no normalization factors are used.

The VRs are then used to confirm that the normalization factors are applicable for the SRs. To be able to make this statement the VRs need to be as close as possible to the SRs while still staying orthogonal. To ensure this orthogonality, one of the cuts is inverted and, in case of the CRs, the $m_{\ell j}$ region is shifted. The cuts for the CRs and VRs are shown in tables 5.7 to 5.9 for the W +jets regions WCR and WVR, the Z +jets regions ZCR and ZVR and the Top regions TCR and TVR respectively.

5. Event Selection

Table 5.7.: W control and validation regions (WCR/WVR) definitions.

Feature	WCR	WVR-1TeV	WVR-1p2TeV	WVR-1p5TeV	WVR-1p7TeV	WVR-2TeV	WVR-Overflow
$m_{\ell j}$ [GeV]	[700, 900]	[900, 1100]	[1100, 1350]	[1350, 1600]	[1600, 1800]	[1800, 2100]	> 2100
$N_{\text{baseline } \ell}$				= 1			
$N_{\text{signal } \ell}$				= 1			
$N_{b\text{-jets}}$				= 1			
$\mathcal{S}(p_{\text{T}}^{\text{miss}})$ [GeV $^{\frac{1}{2}}$]				> 2.5			
$p_{\text{T}}^{\ell 1}/m_{\ell j}$				> 0.4			

Table 5.8.: Z control and validation regions (ZCR/ZVR) definitions.

Feature	ZCR	ZVR-1TeV	ZVR-1p2TeV	ZVR-1p5TeV	ZVR-1p7TeV	ZVR-2TeV	ZVR-Overflow
$m_{\ell j}$ [GeV]	[700, 900]	[900, 1100]	[1100, 1350]	[1350, 1600]	[1600, 1800]	[1800, 2100]	> 2100
$N_{\text{baseline } \ell}$				= 2			
$N_{\text{signal } \ell}$				= 2			
$N_{b\text{-jets}}$				= 1			
$m_{\ell\ell}$ [GeV]				< 300			
$\mathcal{S}(p_{\text{T}}^{\text{miss}})$ [GeV $^{\frac{1}{2}}$]				< 4			
$p_{\text{T}}^{\ell 1}/m_{\ell j}$				> 0.3			

Table 5.9.: T control and validation regions (TCR/TVR) definitions.

Feature	TCR	TVR-1TeV	TVR-1p2TeV	TVR-1p5TeV	TVR-1p7TeV	TVR-2TeV	TVR-Overflow
$m_{\ell j}$ [GeV]	[700, 900]	[900, 1100]	[1100, 1350]	[1350, 1600]	[1600, 1800]	[1800, 2100]	> 2100
$N_{\text{baseline } \ell}$				= 1 = 2			
$N_{\text{signal } \ell}$				= 1 = 2			
$N_{b\text{-jets}}$				> 1 > 1			
$\mathcal{S}(p_{\text{T}}^{\text{miss}})$ [GeV $^{\frac{1}{2}}$]				< 2.5 < 4			
$p_{\text{T}}^{\ell 1}/m_{\ell j}$				> 0.4 > 0.3			

5. Event Selection

The W regions invert $\mathcal{S}(p_T^{\text{miss}})$, the Z regions invert $m_{\ell\ell}$ and the Top regions require more than one b -jet. Additionally, the $m_{\ell\ell}$ cut was dropped for the Top region and the 1L and 2L regions are considered together because top strongly contributes to both and separating them would decrease the statistics too much. This was tested and it was additionally found out that the normalization factor stayed very similar when comparing splitting and combining of the Top regions so combining them to reduce the uncertainty seems reasonable.

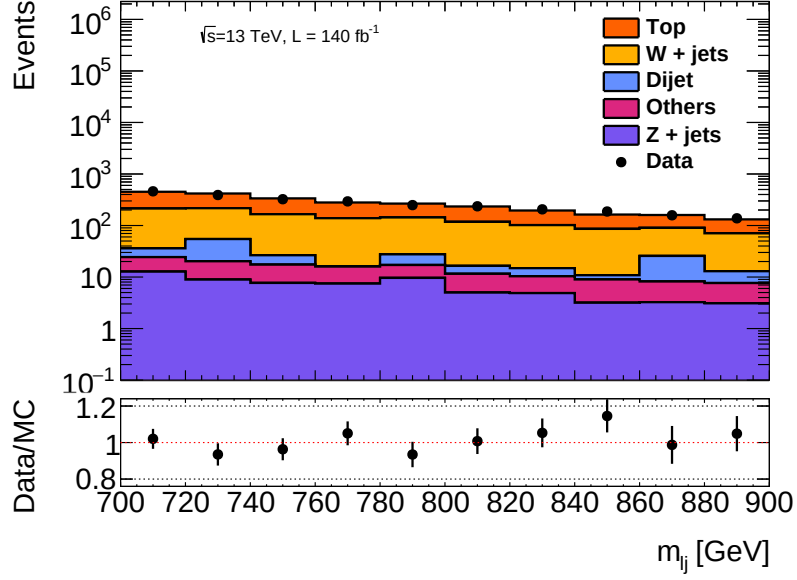
The CRs have a $m_{\ell j}$ range of 700 - 900 GeV and the VRs have the same $m_{\ell j}$ regions as their corresponding SRs. The exception being the high mass VRs starting from the 2.2 TeV region which are combined to one large region to ensure at least one expected background event in these regions. CRs going from 500 - 900 GeV were tested to get even more statistics but the number of events is already large enough (see figures 5.30 to 5.31) in the CRs and the slight reduction in uncertainty on the normalization parameter is not worth the fact that the median of the CR moves further away from the SR.

Possible CRs and even VRs with a b -veto for the W and Z regions were tested and discarded because the topology of b -jets deviates from the topology of light jets (see section 3.7.4). This was visible during the testing phase of the CRs and VRs in a larger offset of data to MC in the VRs after applying normalization factors.

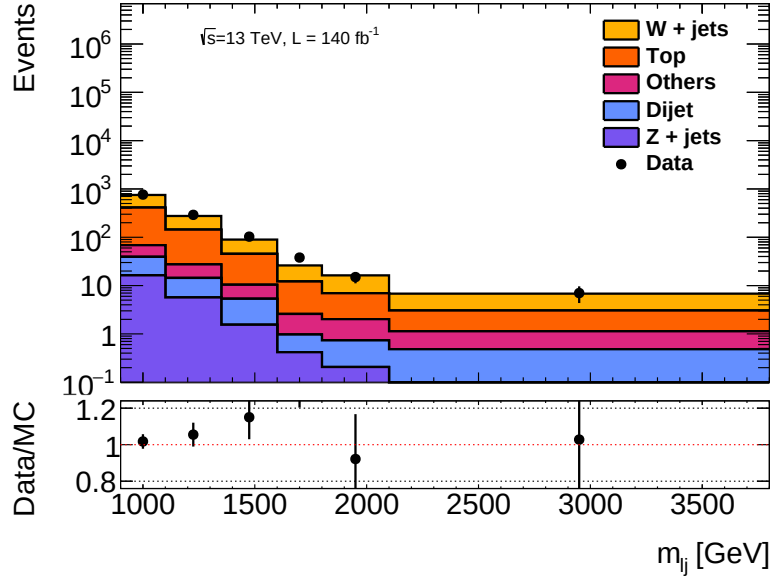
To check that the CRs are chosen wisely figures 5.30, 5.31 and 5.32 show the different CRs binned in steps of 20 GeV. The data/MC ratio fit for both the W and Top region is always within 1σ apart from the 840-860 GeV bin in both cases. The Z region shows a clear offset of the data/MC ratio of roughly 20% within an uncertainty of 1σ which the normalization factor should be able correct. The purity for the Z region is good, for the Top region a lot of impurities are already existent, but tightening the cuts by e.g. requiring more than two b -jet reduces the statistics too much which was tested thoroughly. The worst purity has the W region where there are actually more Top events than W events. Increasing the purity here could mainly be done by applying a b -veto but as mentioned before, the normalization factor would not be very well applicable for the SR then.

Figures 5.30, 5.31 and 5.32 also show all VRs and visualize that even for the last bin the analysis does not go below 1 expected background event. Additionally one sees that the Data/MC agreement shows a larger disagreement than the CR but is still mostly within 1σ . Section 7.1 shows the same CR and VR plots after applying the normalization factors.

5. Event Selection



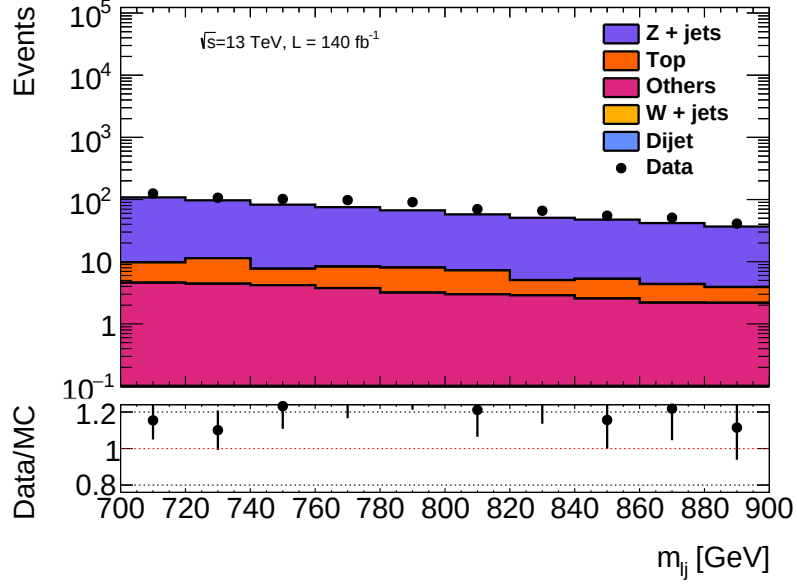
(a) WCR.



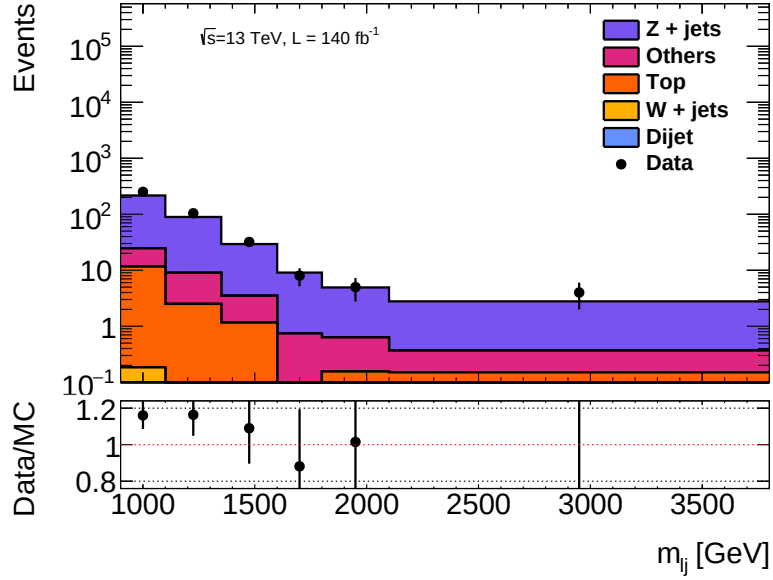
(b) WVR.

Figure 5.30.: m_{lj} pre-fit distributions in the WCR, and WVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

5. Event Selection



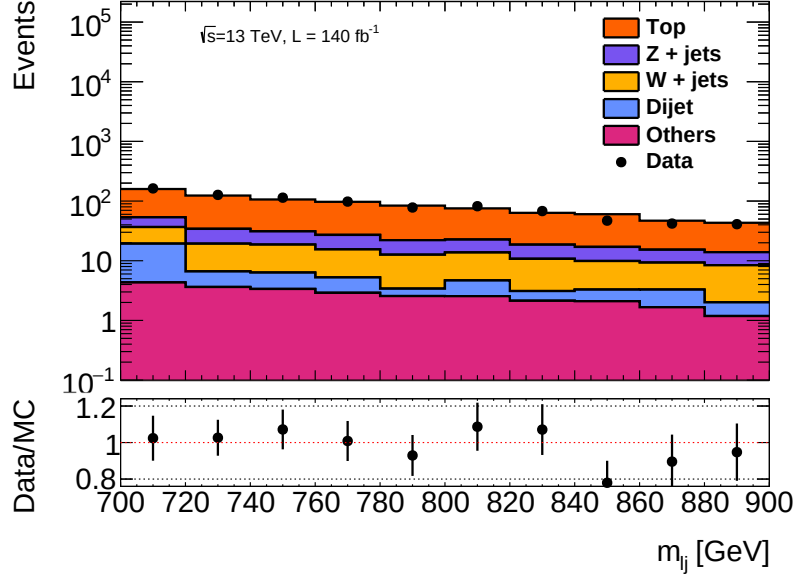
(a) ZCR.



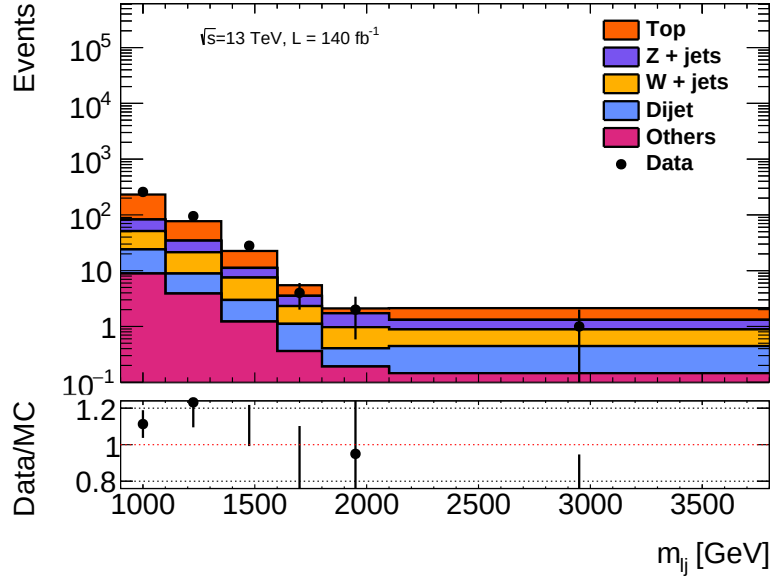
(b) ZVR.

Figure 5.31.: m_{lj} pre-fit distributions in the ZCR, and ZVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

5. Event Selection



(a) TCR.



(b) TVR.

Figure 5.32.: m_{lj} pre-fit distributions in the TCR, and TVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

5. Event Selection

5.6. Cut Efficiencies

The cuts discussed in section 5.4 try to keep as much signal efficiency as possible while suppressing most of the background in the SR. To show the efficiencies of the applied cuts, the cut flow tables of the WCR and the SR-1L-2TeV are given in tables 5.10 and 5.11 respectively. The used signal is a LQ with a mass of 2 TeV and a coupling of 1.5. The WCR includes data whereas the SR does not include data points, due to not having access to unblinded data in the SR. The signal contamination of the WCR is very low, the Top contribution is quite high but is unavoidable by having $N_{b\text{-jets}} = 1$ to ensure the validity in the SR. The cut flow table for SR-1L-2TeV shows great improvements with every applied cut. Especially the cut on the SR itself provides a great efficiency discrepancy of the Signal to the background of more than a factor of 30. The only exception being the restriction on the number of leptons providing a worse efficiency for the Signal than for the $Z + \text{jets}$, Top and Others background. This is not concerning due to having a dedicated optimization for two leptons.

Table 5.10.: Cut flow table of the WCR. The expected number of events predicted by MC are rounded to the next integer value. The signal sample is a 2 TeV LQ with a coupling of 1.5.

Sample	Quantity	Preselection	Region	$N_{\text{signal } \ell}$	ℓ flavour	$N_{b\text{-jets}}$	$\mathcal{S}(p_{\text{T}}^{\text{miss}})$	$p_{\text{T}}^{\ell 1}/m_{\ell j}$
$W + \text{jets}$	Number of Events	351680	202161	202147	112824	8700	5159	1103
	efficiency [%]	100.00	57.48	99.99	55.81	7.71	59.29	21.38
	cumulative efficiency [%]	100.00	57.48	57.48	32.08	2.47	1.47	0.31
$Z + \text{jets}$	Number of Events	79792	45146	20214	10754	940	270	66
	efficiency [%]	100.00	56.58	44.77	53.20	8.74	28.71	24.37
	cumulative efficiency [%]	100.00	56.58	25.33	13.48	1.18	0.34	0.08
Top	Number of Events	81963	50172	47831	28061	12806	9412	1293
	efficiency [%]	100.00	61.21	95.33	58.67	45.64	73.50	13.74
	cumulative efficiency [%]	100.00	61.21	58.36	34.24	15.62	11.48	1.58
Dijet	Number of Events	66159	32808	32805	31595	2545	648	100
	efficiency [%]	100.00	49.59	99.99	96.31	8.06	25.45	15.42
	cumulative efficiency [%]	100.00	49.59	49.59	47.76	3.85	0.98	0.15
Others	Number of Events	16570	8910	7744	4392	531	336	76
	efficiency [%]	100.00	53.77	86.91	56.72	12.09	63.24	22.63
	cumulative efficiency [%]	100.00	53.77	46.73	26.51	3.20	2.03	0.46
Signal	Number of Events	28	1	1	1	1	0	0
	efficiency [%]	100.00	4.13	78.47	100.00	60.94	36.24	69.20
	cumulative efficiency [%]	100.00	4.13	3.24	3.24	1.98	0.72	0.50
Data	Number of Events	586118	342151	310930	190970	26255	15829	2644
	efficiency [%]	100.00	58.38	90.88	61.42	13.75	60.29	16.70
	cumulative efficiency [%]	100.00	58.38	53.05	32.58	4.48	2.70	0.45

5. Event Selection

Table 5.11.: Cut flow table of the SR-1L-2TeV. The expected number of events predicted by MC are rounded to the next integer value. The signal sample is a 2 TeV LQ with a coupling of 1.5.

Sample	Quantity	Preselection	Region	$N_{\text{signal } \ell}$	ℓ flavour	$N_{b\text{-jets}}$	$\mathcal{S}(p_{\text{T}}^{\text{miss}})$	$p_{\text{T}}^{\ell}/m_{\ell j}$
W + jets	Number of Events	351680	4412	4411	2376	203	82	6
	efficiency [%]	100.00	1.25	99.98	53.86	8.54	40.21	7.67
	cumulative efficiency [%]	100.00	1.25	1.25	0.68	0.06	0.02	0.00
Z + jets	Number of Events	79792	1087	541	281	25	18	1
	efficiency [%]	100.00	1.36	49.71	52.03	8.72	73.99	3.35
	cumulative efficiency [%]	100.00	1.36	0.68	0.35	0.03	0.02	0.00
Top	Number of Events	81963	631	596	340	160	39	2
	efficiency [%]	100.00	0.77	94.42	57.02	47.13	24.60	5.13
	cumulative efficiency [%]	100.00	0.77	0.73	0.41	0.20	0.05	0.00
Dijet	Number of Events	66159	1060	1060	1054	183	105	3
	efficiency [%]	100.00	1.60	100.00	99.41	17.40	57.54	2.46
	cumulative efficiency [%]	100.00	1.60	1.60	1.59	0.28	0.16	0.00
Others	Number of Events	16570	251	215	117	13	5	1
	efficiency [%]	100.00	1.51	85.91	54.47	11.15	36.98	18.30
	cumulative efficiency [%]	100.00	1.51	1.30	0.71	0.08	0.03	0.01
Signal	Number of Events	28	13	10	10	6	4	3
	efficiency [%]	100.00	46.01	75.48	100.00	56.68	76.57	66.46
	cumulative efficiency [%]	100.00	46.01	34.73	34.73	19.68	15.07	10.02

5.7. Dijet Sample Performance

This section will compare the results of the dijet sample with the results gained using a data-driven fake estimation instead which will be referred to as "Fake" in the plots. For this, the additional cut of $m_T > 20$ GeV is introduced to stay orthogonal to the region used for the fake electron estimation. The m_T distribution in the combined 1L SRs is shown in figure 5.33 by using $m_{\ell j} > 900$. The region with $m_T < 20$ provides high statistics while trying to cut as less signal as possible.

5. Event Selection

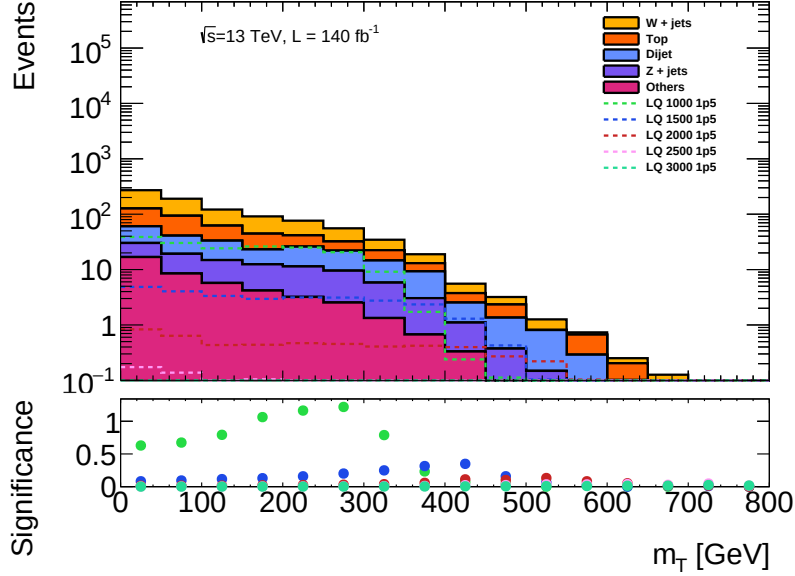


Figure 5.33.: m_T distributions in the SR after applying all cuts. Both the background and signal events are sorted by the total number of events separately.

All effected CRs and SRs are analysed being the WCR, the TCR and the 1L SR combining all SRs with one lepton. The regions with two lepton final states are not tested due to the dijet background having a negligible amount of events for two leptons in the final state. The TCR uses the cut on m_T only for cases with 1 lepton, the two lepton selection remains unchanged. The data-driven background contains older samples using AnalysisBase, version 24.2.24 instead of AnalysisBase, version 25.2.11 and the derivation tag p5855 instead of p6026 due to not having a data-driven sample available for the newest derivation. These differences in versions are only marginal and are just mentioned for completeness. Other changes include changes in the muon calibration, which should also have almost no impact due to muons being rejected, the introduction of an electron efficiency scale factor and the update of the b -tagging scale factor recommendations. The last two might create slight differences and have to be taken into account when looking at the distributions. All future figures in this section use the older production version for all samples to create consistencies between the them. The 1L SR did yield a negative number of events for the fake lepton estimation. This is not forbidden using the asymptotic matrix method (see section 4.4) but does not make much sense to use and is therefore not advised. Nonetheless, the two samples will be compared in the CRs.

First of all, the Dijet sample is directly compared to the fake estimation in the WCR in figure 5.34 for the $m_{\ell j}$ distribution. There, the Dijet sample is most of the times higher or close to the fake estimation while mostly being within one order of magnitude with the exception of one bin being within two orders of magnitude. The

5. Event Selection

uncertainty on the fake estimation is not drawn.

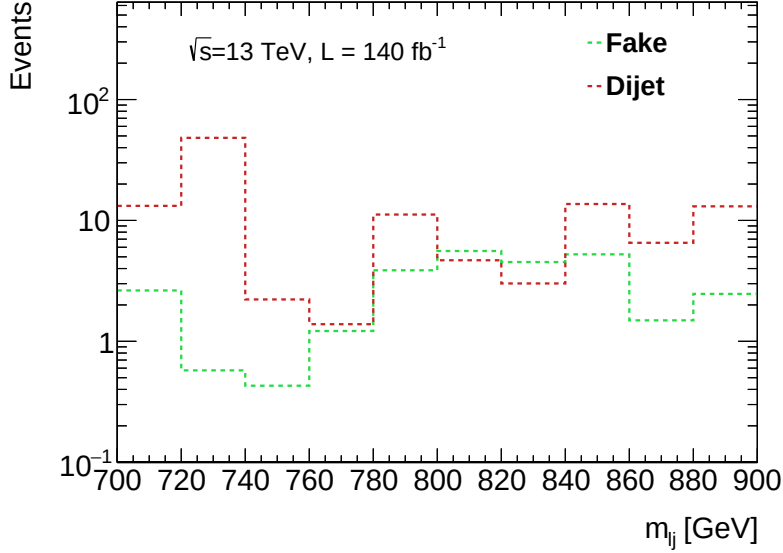
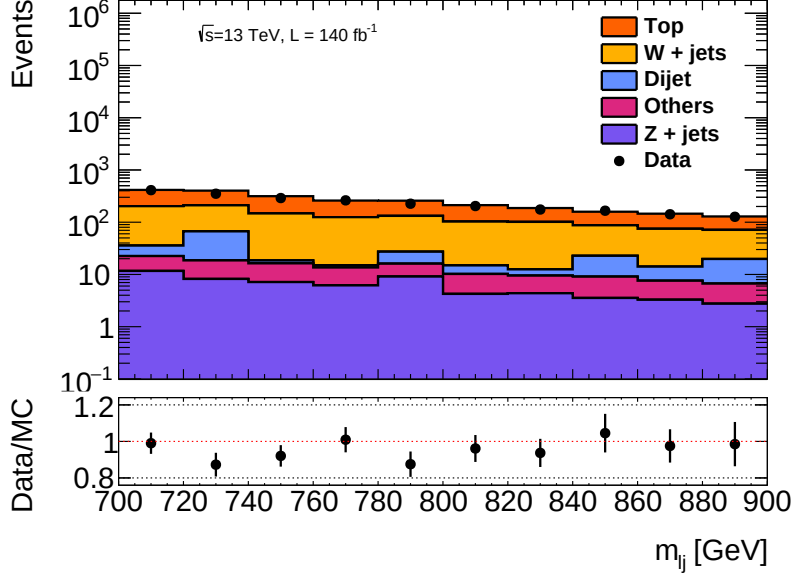


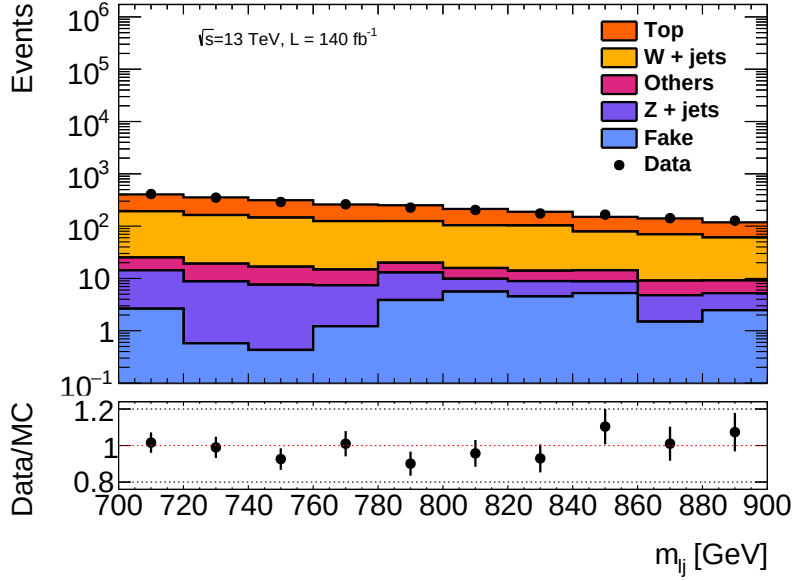
Figure 5.34.: m_{lj} distributions in the WCR after applying all cuts comparing the event yield of the Dijet sample with the electron fake estimation. The used samples do not use the newest version are included for comparing the sample not for the absolute yields.

To put these differences in perspective both samples are combined with the full background estimation and compared to the collision data yield. For that, the m_{lj} , the leading lepton p_T and the leading jet p_T distribution for the WCR are given in figures 5.35, 5.36 and 5.37 respectively. Most notable is the change from being the third largest background to being the smallest. For the m_{lj} distribution the fake estimation reduces the overestimation in the second bin but instead introduces a slight underestimation in the last three bins. The other bins change minimally and the general shape remains the same. For the leading lepton p_T distribution is the Dijet showing a good distribution. The last data point in the 600-700 GeV does not contain any Dijet events and is still overestimated showing the low impact of the Dijet sample. The fake electron estimation fit the data slightly better at low momenta but the bins starting with 500 GeV are greatly overestimated by the data-driven method. The leading jet p_T distributions behave similar to the leading lepton p_T distributions. The low momentum region is slightly better described by the fake electron estimation. The high momentum region seems to be described very similarly with both samples.

5. Event Selection



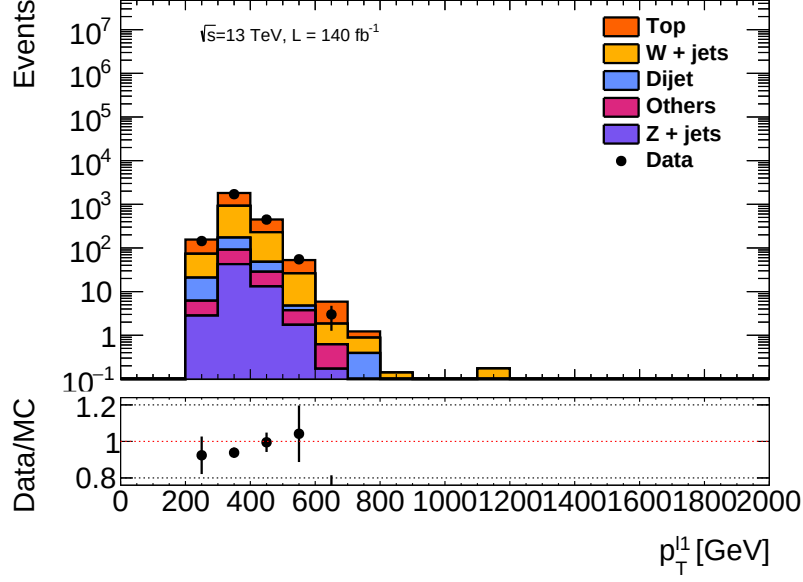
(a) Dijet.



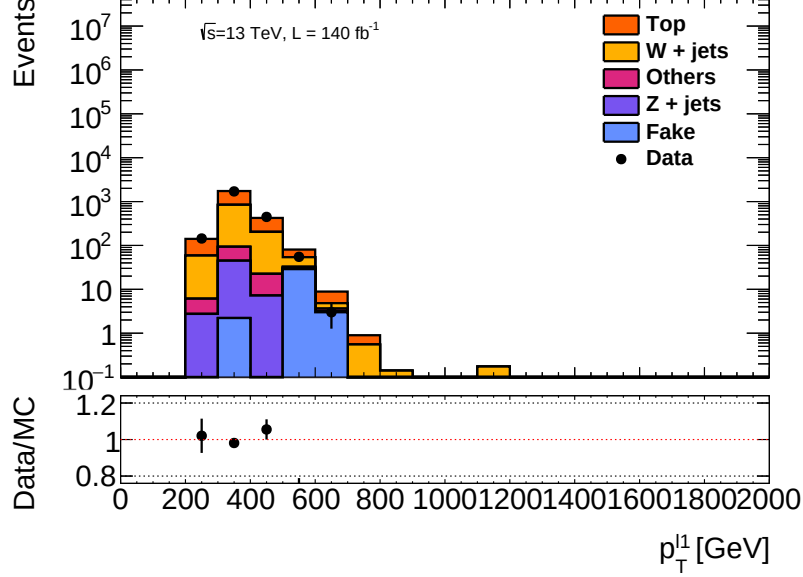
(b) Electron Fake Estimation.

Figure 5.35.: $m_{\ell j}$ distributions in the WCR after applying all cuts showing the difference in the total background estimation when replacing the Dijet sample with the electron fake estimation. The used samples do not use the newest version are included for comparing the sample not for the absolute yields.

5. Event Selection



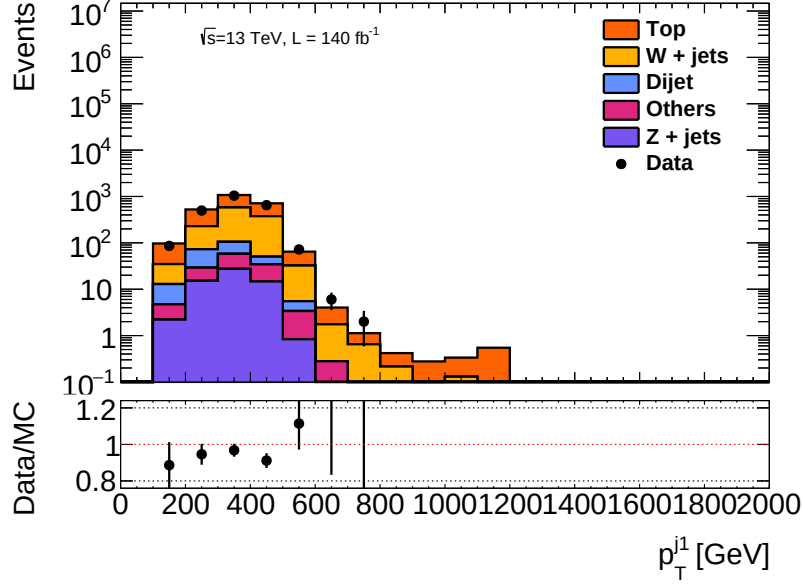
(a) Dijet.



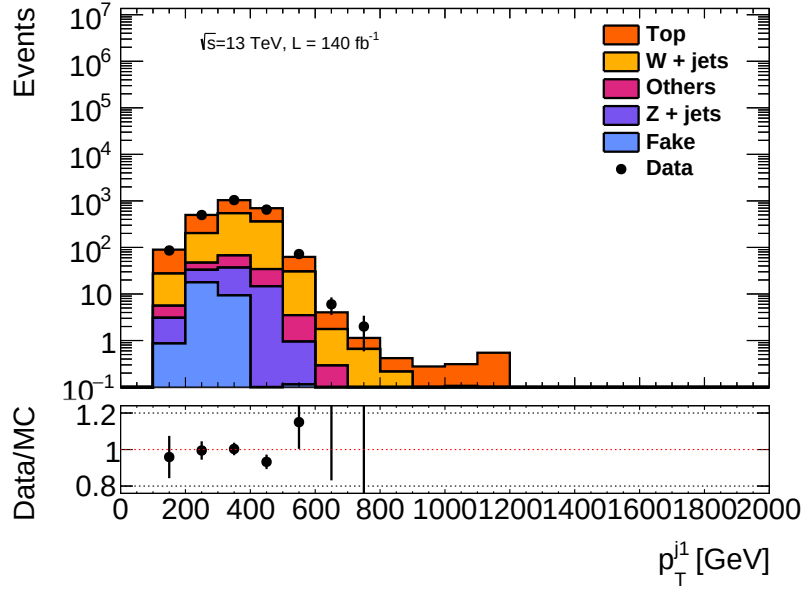
(b) Electron Fake Estimation.

Figure 5.36.: p_T distributions of the leading lepton in the WCR after applying all cuts showing the difference in the total background estimation when replacing the Dijet sample with the electron fake estimation. The used samples do not use the newest version are included for comparing the sample not for the absolute yields.

5. Event Selection



(a) Dijet.



(b) Electron Fake Estimation.

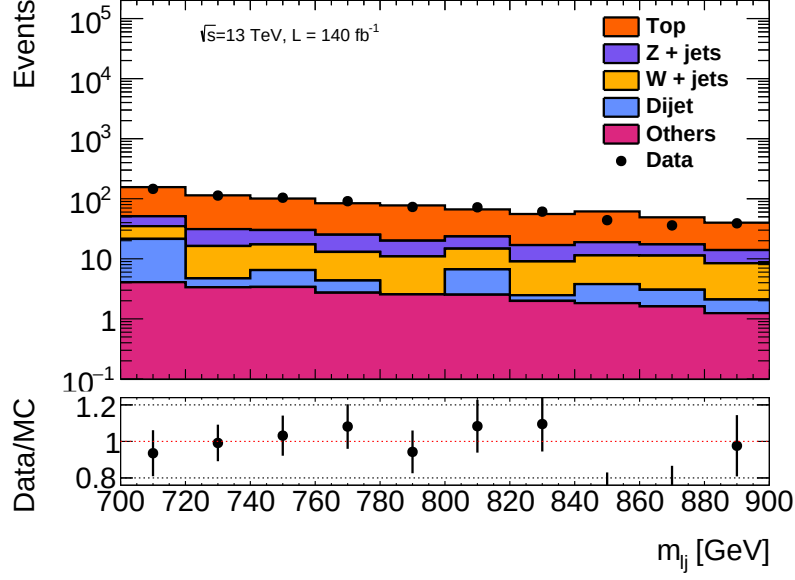
Figure 5.37.: p_T distributions of the leading jet in the WCR after applying all cuts showing the difference in the total background estimation when replacing the Dijet sample with the electron fake estimation. The used samples do not use the newest version are included for comparing the sample not for the absolute yields.

5. Event Selection

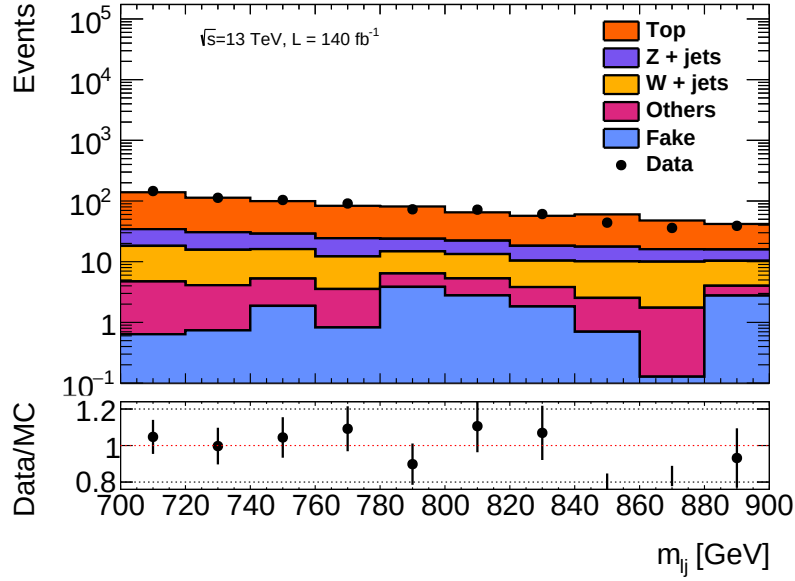
Lastly the TCR is tested. Due to Dijet having a small contribution in the TCR any changes are minimal and only the $m_{\ell j}$ distribution comparison is given in figure 5.38. At first glance, the Data/MC ratio might look identical and they are in fact very close together. Only the first bin has a stronger deviation, all other changes in the ratio are within 0.05 and well within the uncertainties. This distribution clearly shows the marginal difference the two different samples produce.

Due to these comparably small differences induced from the fake electron estimation, the negative yields in the SR and the fact that the newest sample wasn't available, the rest of this thesis keeps using the dijet sample. Any flat mismodeled increase or decrease in events of the Dijet sample compared to the fake electron sample is absorbed by the normalization factors.

5. Event Selection



(a) Dijet.



(b) Electron Fake Estimation.

Figure 5.38.: m_{lj} distributions of the leading lepton in the TCR after applying all cuts showing the difference in the total background estimation when replacing the Dijet sample with the electron fake estimation. The used samples do not use the newest version are included for comparing the sample not for the absolute yields.

6. Statistical Analysis

This chapter gives the necessary background information needed to understand and recreate the results given in chapter 7. For this the general approach in a statistical analysis is explained and also the systematic uncertainties used for this thesis are listed.

6.1. Statistical Framework

This section is heavily inspired by [187] and [188]. When additional references are used they are directly mentioned.

To be able to do an estimate on the sensitivity of a measurement one needs to know which signal models can be excluded if only the standard model were to be measured. This is called the exclusion significance and differs from the discovery significance which is used to confirm a discovery of a particle where usually a significance of 5σ is required. For the exclusion significance a confidence level of 95% is required which equals to a p-value of 0.05 and can be transformed into a significance of $Z = 1.64$ with the relation:

$$Z = \Phi^{-1}(1 - p), \quad (6.1)$$

where Φ^{-1} is the inverse of the cumulative distribution of a standard Gaussian. This is using the definition of a one-sided confidence limit because the signal model is expected to always increase the number of expected events and never decrease it. This relation is also visualized in figure 6.1.

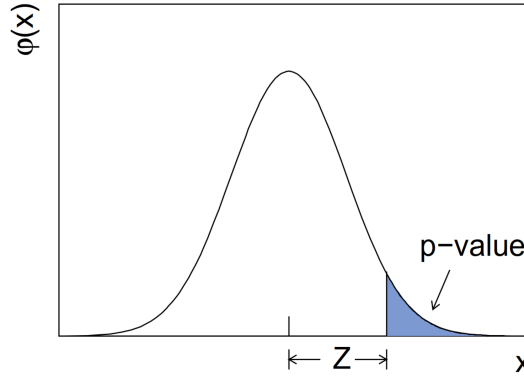


Figure 6.1.: Relation of the p-value and the significance Z for a standard Gaussian.[187]

6. Statistical Analysis

The expected number of events for $n = [n_1, ..n_N]$ with N bins can be expressed as

$$E[n_i] = \mu s_i + b_i, \quad (6.2)$$

where μ is the signal strength parameter of the signal model which means that $\mu = 0$ corresponds to the background only hypothesis and $\mu = 1$ corresponds to the signal+background hypothesis. s_i and b_i are the number of signal and background events in bin i which are gained by calculating

$$s_i = s_{tot} \int_{\text{bin } i} f_s(x, \theta_s) dx \text{ and} \quad (6.3)$$

$$b_i = b_{tot} \int_{\text{bin } i} f_b(x, \theta_b) dx. \quad (6.4)$$

Here s_{tot} and b_{tot} are the total number of signal/background events and $f_s(x, \theta_s)$ and $f_b(x, \theta_b)$ are their respective probability density functions (pdfs) which depend on the measured quantity x and the shape parameters θ_s and θ_b . Both shape parameters as well as b_{tot} can be considered nuisance parameters and are denoted as θ in future equations. These nuisance parameters can be constrained by constructing control regions where no signal event is expected. The expectation values of the control regions with M bins and $m = [n_1, ..n_M]$ are expressed as

$$E[m_i] = u_i(\theta). \quad (6.5)$$

The probabilities P of each bin to measure ϵ events when γ events are expected follow a Poisson distribution results in equation:

$$P(\epsilon, \gamma) = \frac{\gamma^\epsilon}{\epsilon!} e^{-\gamma}. \quad (6.6)$$

To get the likelihood function one needs to take the product of all Poisson probabilities:

$$L(\mu, \theta) = \prod_{j=1}^N \frac{(\mu s_j + b_j)^{n_j}}{n_j!} \cdot e^{-(\mu s_j + b_j)} \cdot \prod_{k=1}^M \frac{(u_k)^{m_k}}{m_k!} \cdot e^{-(u_k)} \quad (6.7)$$

To have a measure how well a value of μ performs the profile likelihood ratio is introduced as:

$$\lambda(\mu) = \frac{L(\mu, \theta_{\text{opti}})}{L(\mu_{\text{best}}, \theta_{\text{best}})}. \quad (6.8)$$

$L(\mu, \theta_{\text{opti}})$ is the likelihood function which is maximized for a fixed μ . $L(\mu_{\text{best}}, \theta_{\text{best}})$ on the other hand is the maximized likelihood function for any μ . The maximization is done by varying the values of θ .

This leads to λ always being $0 \leq \lambda \leq 1$ which allows for the use of the test statistic

$$q_\mu = \begin{cases} -2\ln(\lambda(\mu)) & \mu_{\text{best}} \leq \mu \\ 0 & \mu_{\text{best}} \geq \mu \end{cases} \quad (6.9)$$

6. Statistical Analysis

where q_μ has values from 0 to infinity with higher values having larger disagreement with the measured values. The $\mu_{\text{best}} \geq \mu$ case is to cover the possibility to have more measured values than the current μ would expect. This region is not considered for the rejection because one might falsely reject this value of μ otherwise.

One can now define the p-value as

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\text{inf}} f(q_\mu|\mu) dq_\mu, \quad (6.10)$$

where $f(q_\mu|\mu)$ is again the pdf of q_μ while using μ as the hypothesis. $q_{\mu,\text{obs}}$ is the observed value of the test statistic. Figure 6.2 visualizes this relation. Note that t_μ in this figure corresponds to q_μ .

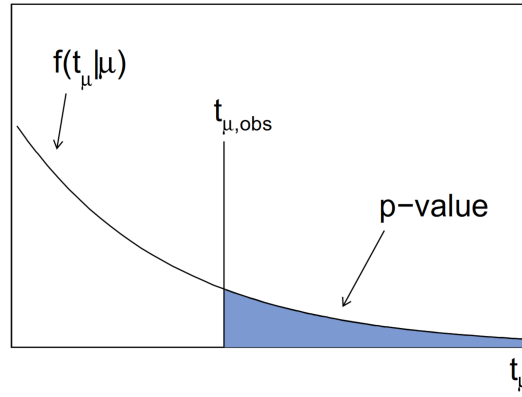


Figure 6.2.: Relation of the p-value and the pdf $f(t_\mu|\mu)$ when observing the test statistic $t_{\mu,\text{obs}}$. [187]

Using these relations the toy model [189] approach can be used to receive the shape by creating a large amount of pseudo-datasets and calculating the upper limit of μ for all of them. With this one can integrate the values and get the significances by counting the number of events having a lower value than the one tested. For example, a significance of 1σ is reached when 68% of events have a lower value than the one tested. This method of estimating the pdf takes a lot of computational time.

As an alternative the asymptotic method [187] was developed to reduce the computational time needed to get an exclusion limit. The asymptotic method uses the Wald approximation introduced in [190] with which q_μ can be simplified to

$$q_\mu = \begin{cases} \frac{(\mu - \tilde{\mu})^2}{\sigma^2} & \mu_{\text{best}} \leq \mu, \\ 0 & \mu_{\text{best}} \geq \mu, \end{cases} \quad (6.11)$$

where $\tilde{\mu}$ follows a Gaussian around the mean value μ' with a standard deviation of σ . From this, the pdf is found to be

$$f(q_\mu|\mu') = \Phi\left(\frac{\mu' - \mu}{\sigma}\right) \delta(q_\mu) + \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q_\mu}} \exp\left[-\frac{1}{2} \left(\sqrt{q_\mu} - \frac{\mu - \mu'}{\sigma}\right)^2\right] \quad (6.12)$$

6. Statistical Analysis

The cumulative distribution is then given by

$$F(q_\mu|\mu') = \Phi\left(\sqrt{q_\mu} - \frac{\mu - \mu'}{\sigma}\right), \quad (6.13)$$

which reduces to

$$F(q_\mu|\mu) = \Phi(\sqrt{q_\mu}), \quad (6.14)$$

when $\mu' = \mu$ holds. Combining equation 6.10 and 6.14 one can see that the p-value is given with:

$$p_\mu = 1 - F(q_\mu|\mu) = 1 - \Phi(\sqrt{q_\mu}). \quad (6.15)$$

Combining this now with equation 6.1 gives an expression for the significance

$$Z_\mu = \Phi^{-1}(1 - p_\mu) = \sqrt{q_\mu}. \quad (6.16)$$

Using these last two equations the p-value and the significance can be calculated without having to simulate countless of pseudo-events. The asymptotic method has the drawback of being an approximation and it breaks down at low number of expected background events. This was tested by [187] and depending on the accuracy targeted, the minimum number of expected background events is advised to be between 2-20. The final confidence limit used is CL_s which is required to be < 0.05 to reject a hypothesis, was introduced in [191] and is defined as

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1 - p_b}. \quad (6.17)$$

This does not only take the signal+background hypothesis into account but also the background only hypothesis. Such an approach is rather conservative due to having generally higher confidence limits but it ensures that signals which produce a pdf close to the background only hypothesis are not rejected too easily. Signal hypotheses which are close to the background hypothesis suffer from impurities by the background distribution and may be rejected falsely otherwise. If the hypotheses are strongly separated the denominator goes to 1 and CL_s reduces to CL_{s+b} . Figure 6.3 visualizes both p-values with respect to the observed test statistic, in this case being the mean value of the background only pdf.

The used program for the statistical analysis is mentioned in section 4.5 in further detail.

6. Statistical Analysis

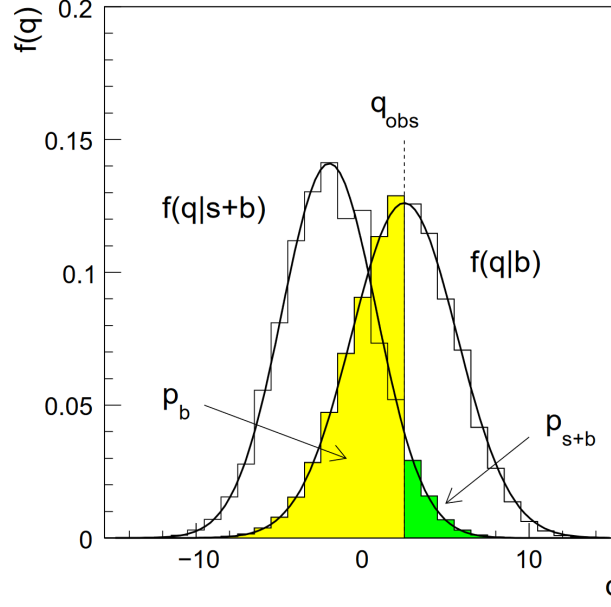


Figure 6.3.: Illustration of the p-values p_{s+b} and p_b when observing the test statistic q_{obs} . [187]

After having a p-value, the exclusion limit needs to be created. This is done by testing multiple signal hypothesis in a specified phase space and receiving a p-value for all of them. This thesis uses a grid-like structure which is used to interpolate the p-values. By interpolating using the four closest values (for the 2-D grid) it is possible to draw a 95% exclusion limit line where the p-values are equal to 0.05.

In addition, one usually includes so called σ -bands which represent the statistical fluctuation of data by a standard deviation σ . This does not correspond to the change of the exclusion significance but rather is used as a visual indication where the observed limit with a probability given in terms of σ will be. Using figure 6.3, this means that the width of $f(q|b)$ is used for σ instead of the width of $f(q|s+b)$.

6.2. Systematic Uncertainty

Systematic uncertainties are very important to quantify and estimate the uncertainty on the results presented in the next chapter. For this, many different systematic uncertainties went into the statistical analysis. Nonetheless, not all systematics are included for example the uncertainty on the electron isolation energy, the systematics going into the muon identification or any theoretical uncertainties just to name a few. To account for potential underestimation of the uncertainty a 20% flat dummy uncertainty has been introduced in each SR. This means that every yield in the signal region has an added 20% uncertainty. This should be more than enough to cover the missing systematic uncertainty and might even overestimate the uncertainty. The

6. Statistical Analysis

added systematic uncertainties are: electron reconstruction efficiency with an uncertainty of roughly 0.25%, electron identification efficiency with an uncertainty of around 1.5% , electron trigger efficiency with roughly 0.3% uncertainty, JVT efficiency with an uncertainty varying between 5-20% and several b -tagging efficiency uncertainties contributing to the total uncertainty with roughly 2-5% depending on the specific process. These uncertainties change for every event and the provided values are averages of these uncertainties.

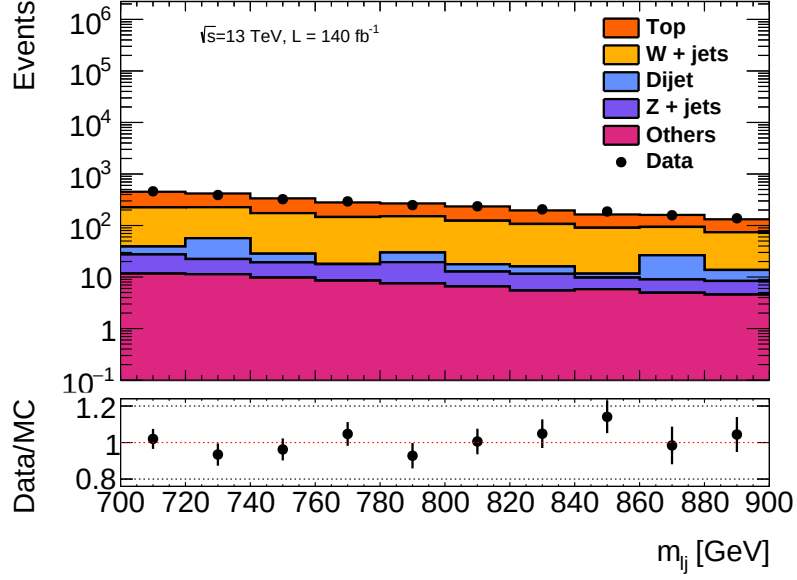
7. Results

This chapter contains the final results of this thesis including post-fit plots of CRs, VRs and SR composition. The expected exclusion limit using the asymptotic method is presented for the 1L and 2L channel as well as the combined results including a yield table and the comparison with a toy model approach. Finally fit parameters, and fit control plots such as a correlation matrix and a pulls plot are given.

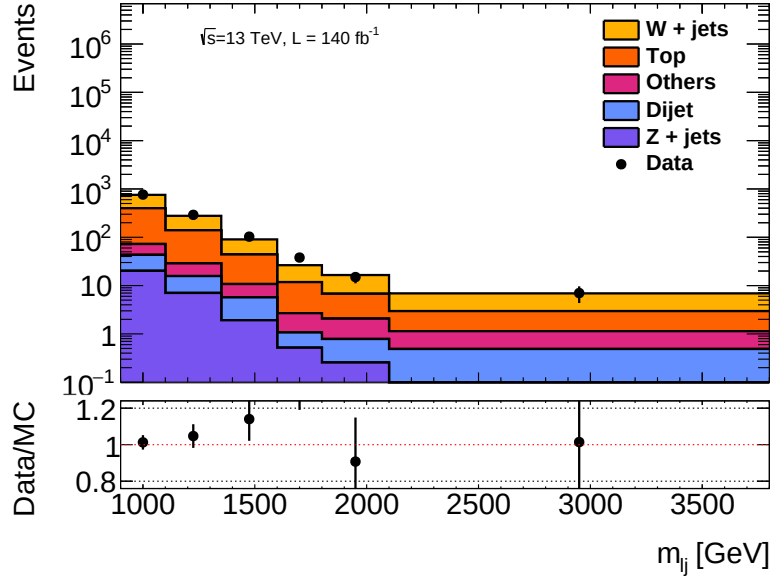
7.1. Fit Results

The normalization factors are determined using a maximum likelihood fit tweaking the normalization factors so that MC reflects the data yields in the CRs. They are: " $\mu_{\text{Top_el}}$ " = 0.9530 ± 0.2478 , " $\mu_{\text{Wjets_el}}$ " = 1.0468 ± 0.1495 , " $\mu_{\text{Zjets_el}}$ " = 1.2393 ± 0.1815 . The CRs and VRs are shown post-fit in figures 7.1, 7.2 and 7.3 using these normalization factors comparing them to the pre-fit plots in section 5.5. They show slight adjustments for the WCR and TCR and the whole offset in the ZCR is corrected now. All CRs are still having one bin clearly outside of the 1σ range which can be expected statistically for ten bins each. No clear decline or increase of the Data/MC ratio throughout the CR is noticeable. If there was a structure to be guessed it would be a sine but this is rather far fetched and the offset is well within the statistical fluctuation. The VR plots change as well but not as notably as the ZCR. They are generally closer to the data now compared to the pre-fit plots.

7. Results



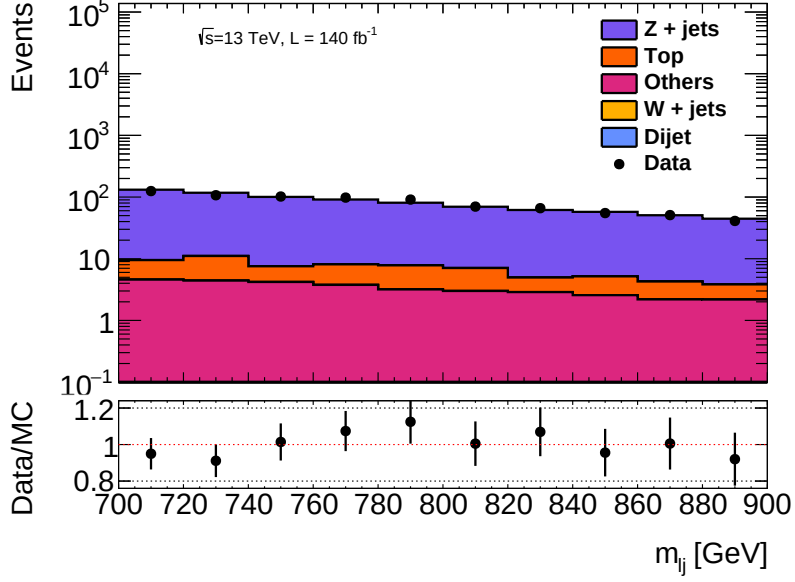
(a) WCR.



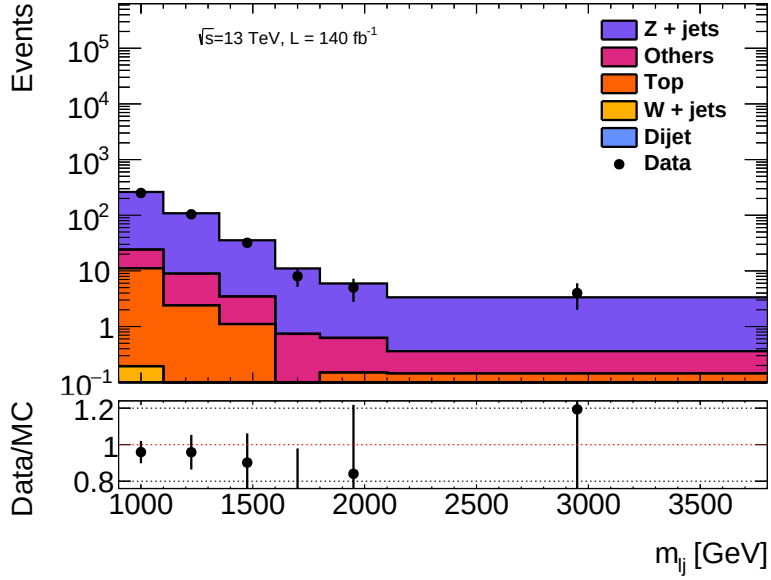
(b) WVR.

Figure 7.1.: m_{lj} post-fit distributions in the WCR, and WVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

7. Results



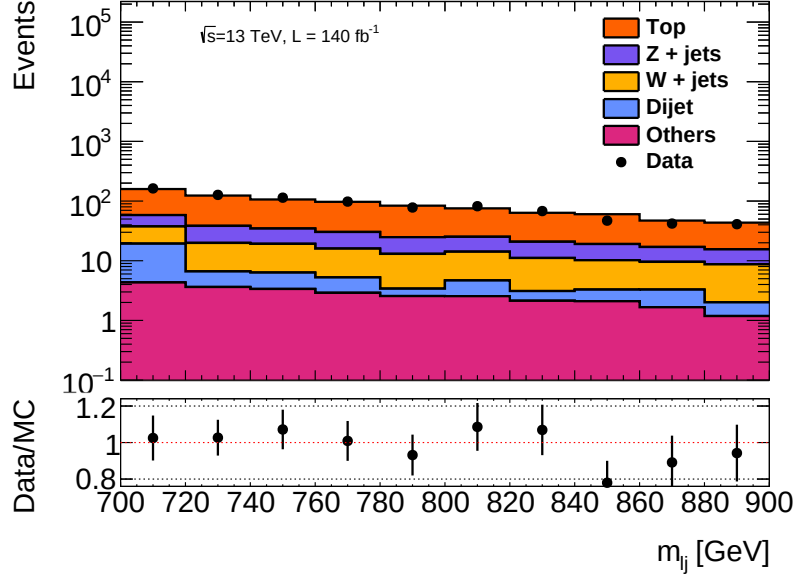
(a) ZCR.



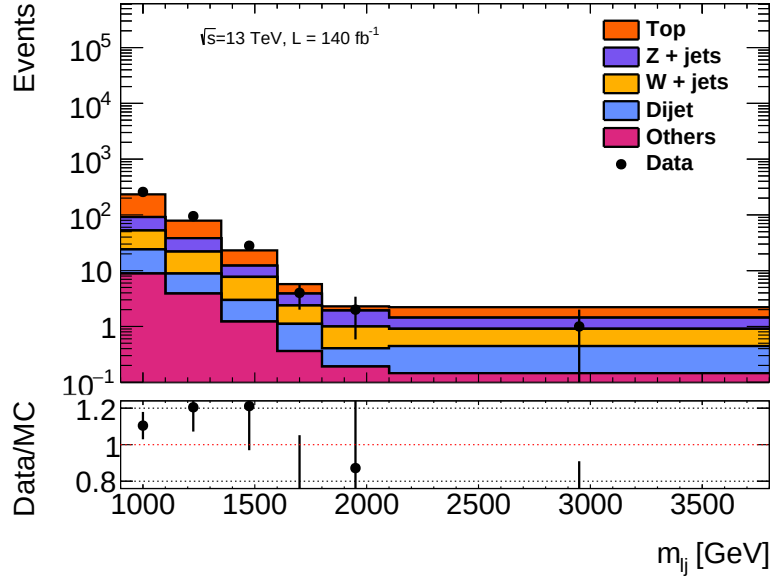
(b) ZVR.

Figure 7.2.: m_{lj} post-fit distributions in the ZCR, and ZVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

7. Results



(a) TCR.



(b) TVR.

Figure 7.3.: m_{lj} post-fit distributions in the TCR, and TVR showing the data to MC agreement. Note that the binning in the CR is only shown for illustrative purposes, the CRs are not binned in the fit.

7. Results

Figure 7.4 shows the post-fit CR and VR significances. The CR significances are 0 by definition because the MC was fit to describe these perfectly. Most of the VRs are well within 1σ and only a few are within 2σ , which is again expected by statistical fluctuation, especially if considering that these values do not apply any uncertainty on the MC values. For the Top VRs, data is slightly underestimated for low masses and slightly overestimated for 1.7 TeV upwards. The WVRs seem to have a slope to underestimate data more and more up to 1.7 TeV. At higher masses the MC estimation fits very well. For the ZVRs the MC seems to very slightly overestimate the data but stays well within 1σ .

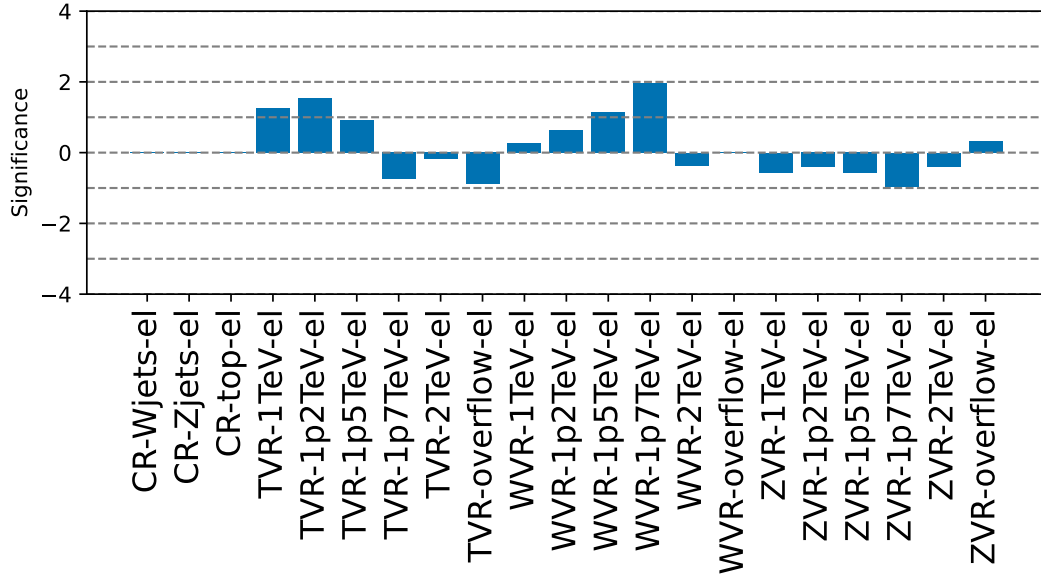


Figure 7.4.: Significances of the CRs and VRs. These significances only consider the uncertainty on the data and neglect the uncertainty of MC simulation.

Figure 7.5 shows the 1L and 2L SR post-fit compositions. Comparing this with the pre-fit compositions (which were shown in figure 5.29) the contribution of the $Z + \text{jets}$ background has increased greatly in the 2L region with a change of roughly 6% while the 1L region stays similar only changing up to 2% for one specific background. The ordering of background contributions in each channel remains unchanged.

7. Results

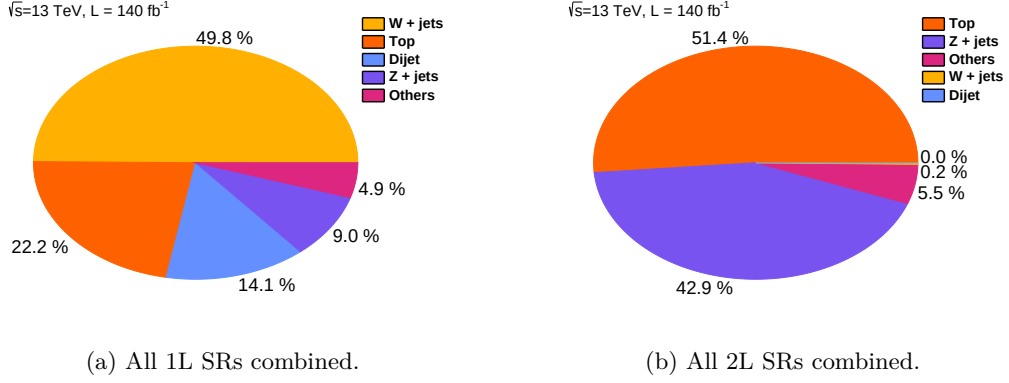
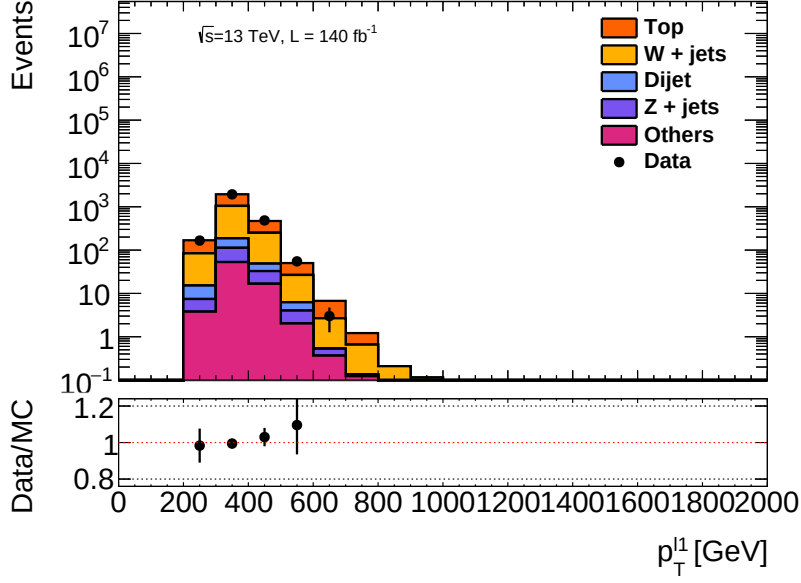


Figure 7.5.: Post-fit pie charts of the background composition in the 1L and 2L SR. The mass dependent SRs are combined here by setting the cut $900 < m_{\ell j} < 3900$ GeV. All cuts from section 5.4 are applied with the normalization factors being used.

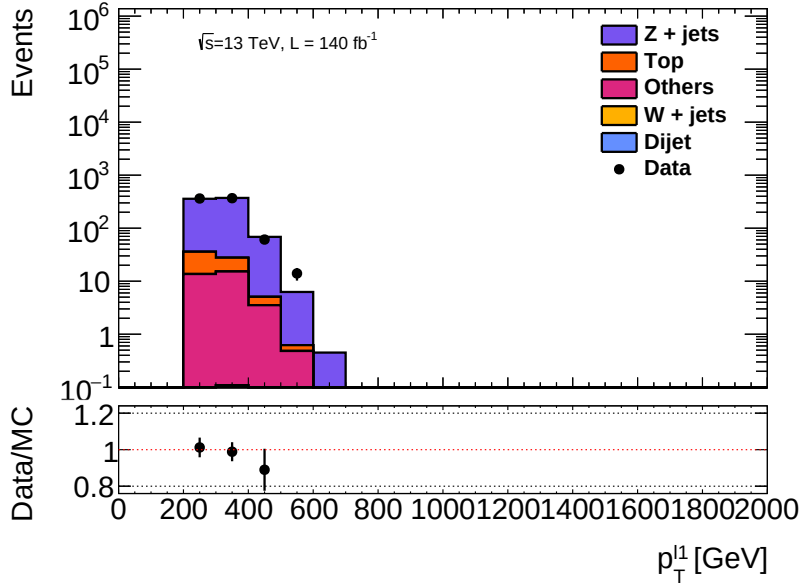
The MC to data agreement needs to be confirmed to ensure the modeling of the data is sufficient and no major oversights were made. For this, ratio plots of relevant distributions are investigated. This includes selected kinematic distributions as well as distributions of the parameters which were cut on. The WCR and the ZCR are used to not be sensitive for any signal. The WCR is used to show the behaviour with one lepton and the ZCR is used for two lepton case. The shown errors only consider statistical uncertainty of the data yields and do not consider any uncertainties of the MC yields.

First the leading lepton p_T , ϕ and η distributions are shown in figures 7.6, 7.7 and 7.8 respectively for both the WCR and ZCR region. The p_T distributions provide good agreement of MC to data at low momenta. The agreement falls with higher momenta but stays within 1σ apart from the last data point which deviates more than 2σ in both regions. Above 700 GeV for the WCR and 600 GeV for the ZCR no data points are available. This is viable due to the MC expectation being at or below 1 event in both regions. The ϕ distributions do not have any unexpected slopes and only have fluctuations around the expected values. These fluctuations go up to roughly 2.5σ but mostly stay within 1σ . The η distributions show good agreement at low absolute η but seems to fall off at higher absolute η values. This is due to decreasing statistics but could also hint to a mismodeling of the leptons in the EMEC. All in all, the general shape of the distribution is followed and these deviations are still within 2σ .

7. Results



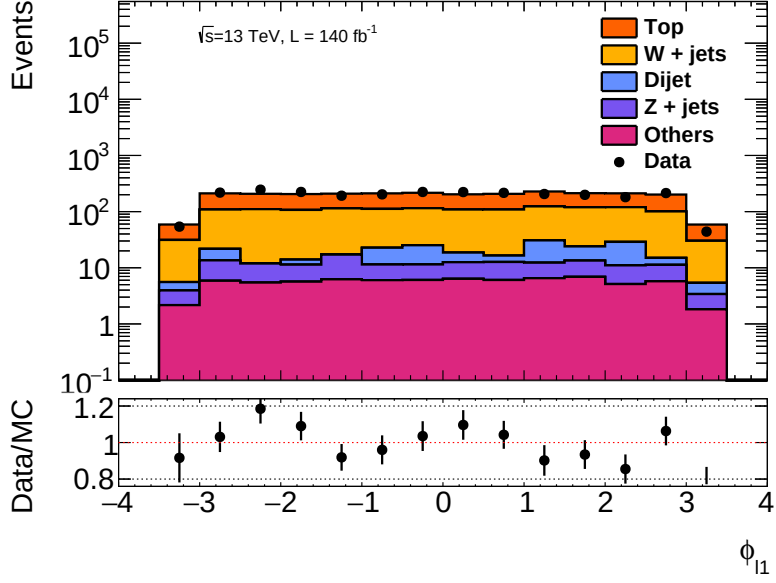
(a) WCR.



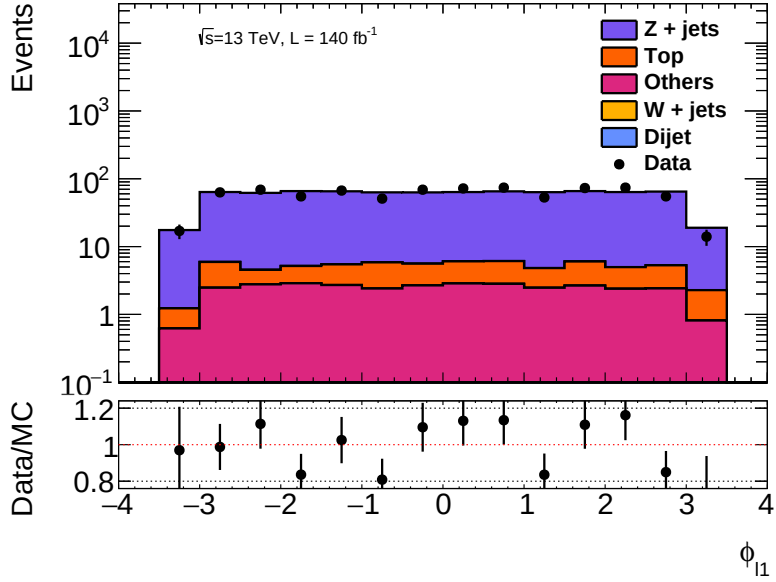
(b) ZCR.

Figure 7.6.: Post-fit p_T distributions of the leading lepton in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



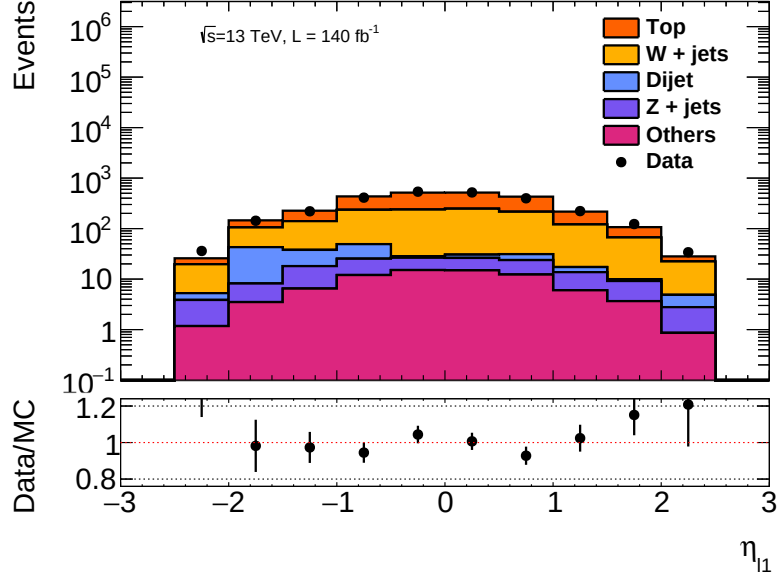
(a) WCR.



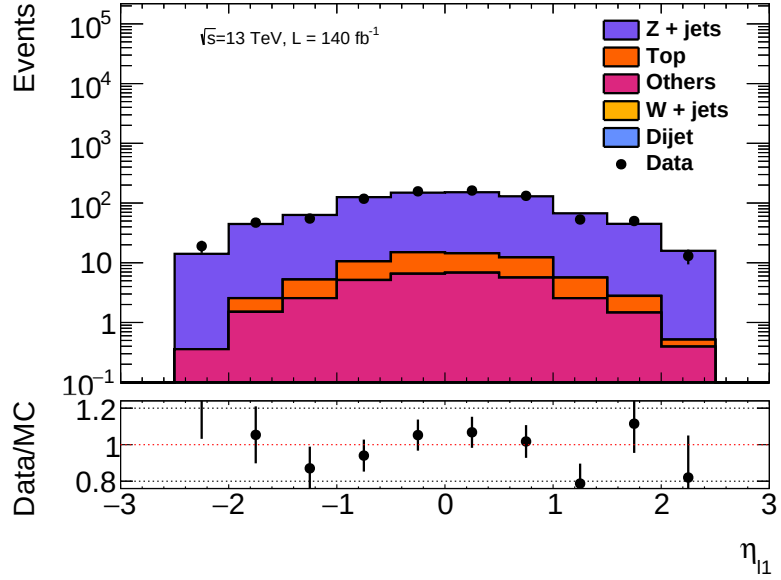
(b) ZCR.

Figure 7.7.: Post-fit ϕ distributions of the leading lepton in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



(a) WCR.



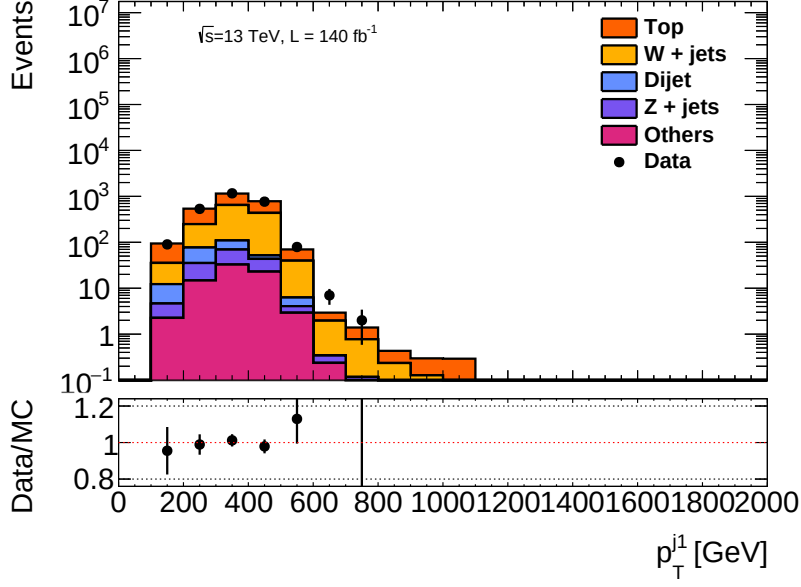
(b) ZCR.

Figure 7.8.: Post-fit η distributions of the leading lepton in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

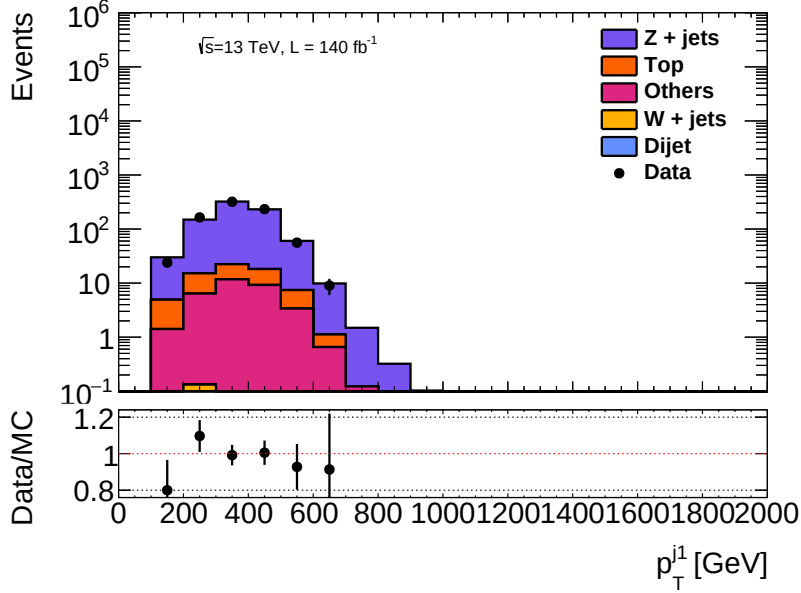
7. Results

Next up are the leading jet distributions of p_T , ϕ and η shown in figures 7.9, 7.10 and 7.11. Starting with the p_T distributions, the agreement for both regions are very good apart from one bin in the WCR where the measured data exceeds the expected events by more than 20%. Most other regions are very close and have a deviation of less than 10%. The ϕ distributions do not have an unexpected slope. The ZCR is almost completely within 1σ apart from the first bin. The WCR mostly agrees but three bins exceed 2σ . Both of them are located near $|\phi| = 2$. These differences could not be resolved and would need further investigation if more precision in this specific region were to be necessary but when using the whole region the agreement is sufficient. The η distribution of the WCR shows a really strong agreement of MC to collision data. The ZCR also has a strong agreement in most regions but shows a larger disagreement in two neighboring bins where one is overestimated and one is underestimated. The overall shape of the distribution seems to fit the data and both offsets are within 2σ .

7. Results



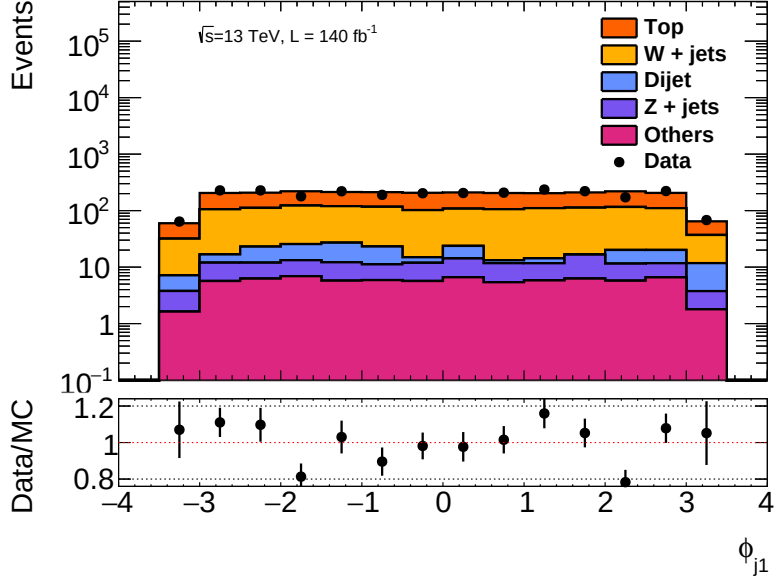
(a) WCR.



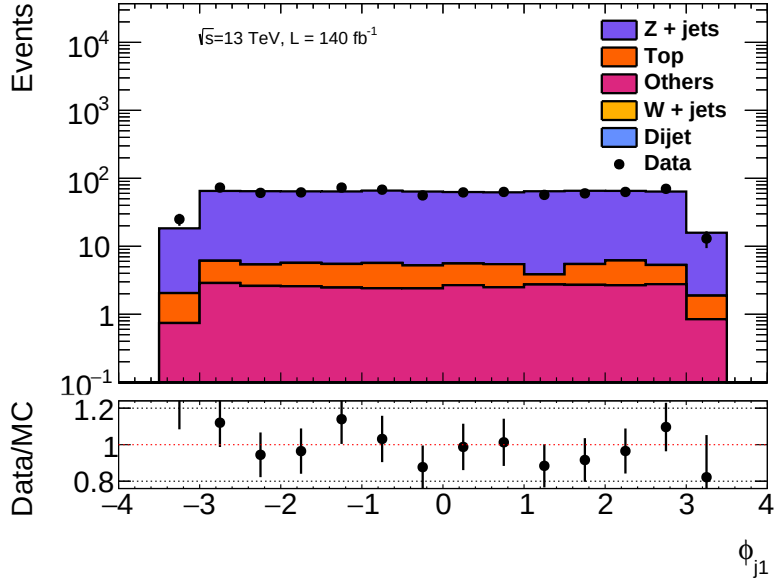
(b) ZCR.

Figure 7.9.: Post-fit p_T distributions of the leading jet in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



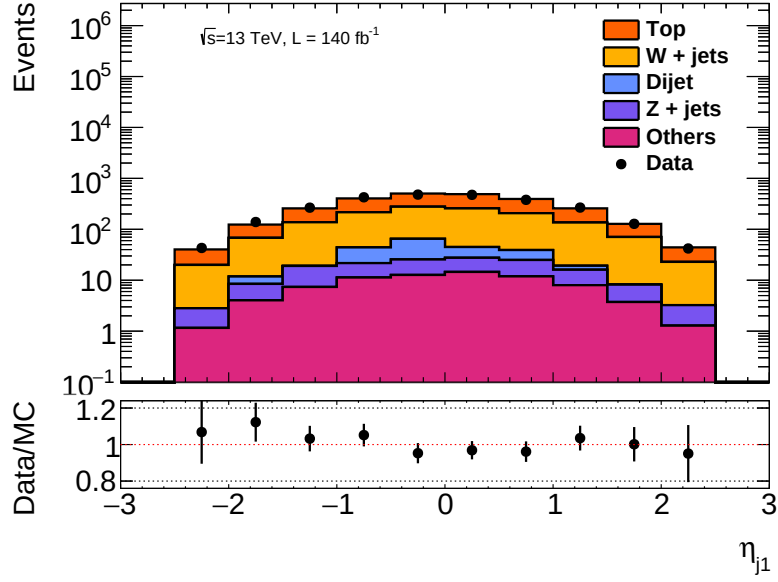
(a) WCR.



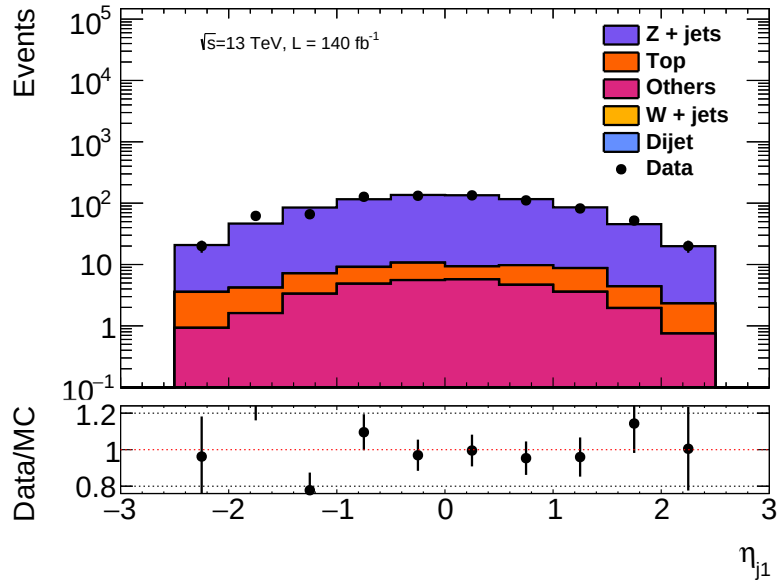
(b) ZCR.

Figure 7.10.: Post-fit ϕ distributions of the leading jet in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



(a) WCR.



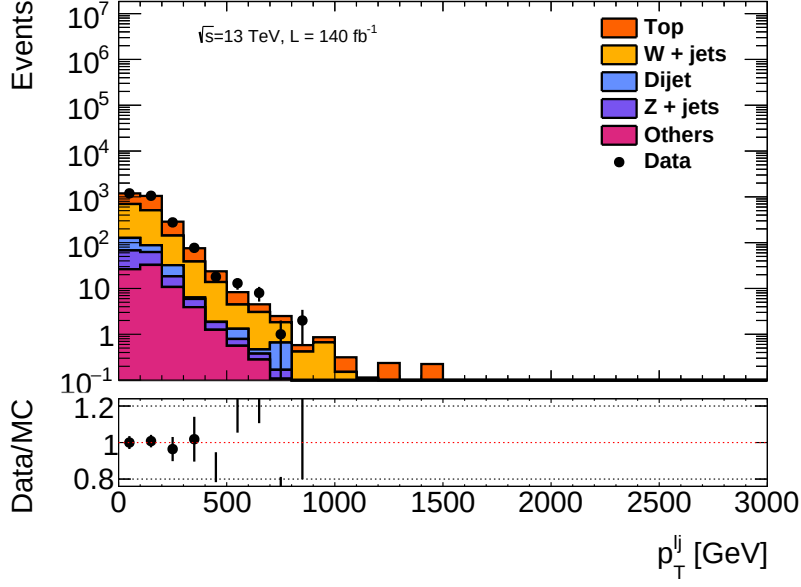
(b) ZCR.

Figure 7.11.: Post-fit η distributions of the leading jet in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

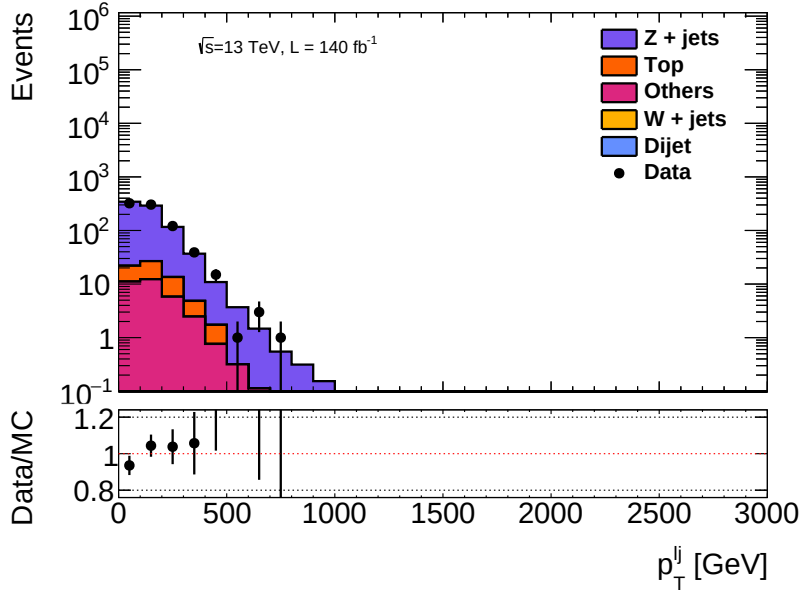
7. Results

After having looked at the leading lepton and the leading jet, the distributions of lepton-jet system can be checked. Shown in figures 7.12, 7.13 and 7.14 are the p_T , ϕ and rapidity y distributions respectively. The y distribution replaces η for heavy particles and heavy invariant mass systems, which is necessary as shown in section 5.3.3. The p_T distribution fits well at low momenta up to 400 GeV in both regions. After that the statistics drop too much to make a definite conclusion, but most bins are within 1σ and the rest are within 2σ . The ϕ distributions for the WCR and ZCR have clear agreement with data modulo possible fluctuations. The shape of the y distributions are also in great agreement with the data, the ZCR has all data points within 1σ and WCR is mostly within 1σ , the rest is within 2σ .

7. Results



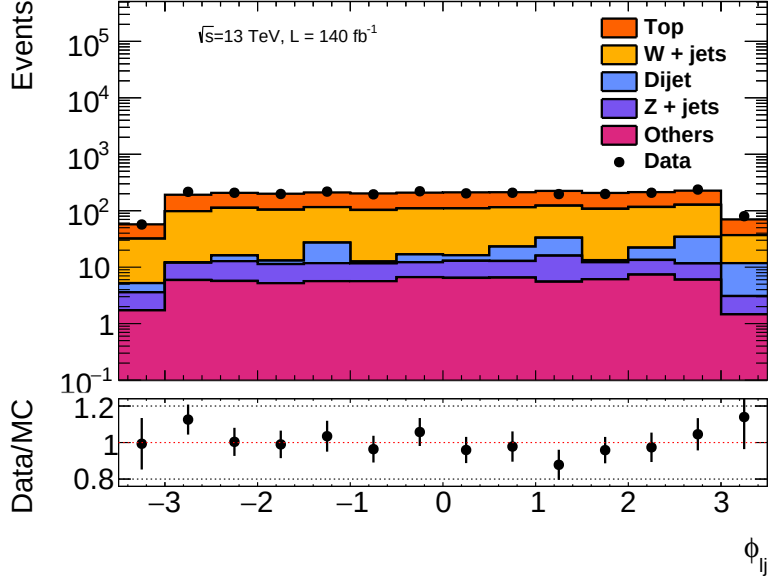
(a) WCR.



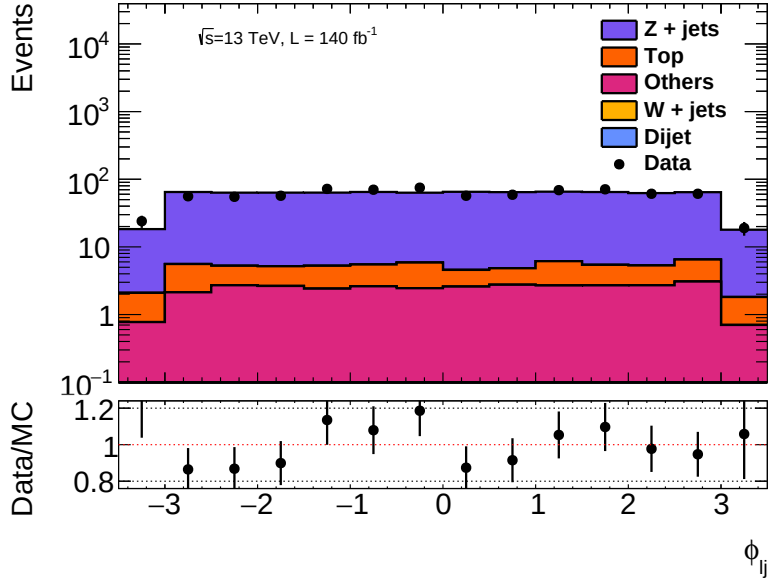
(b) ZCR.

Figure 7.12.: Post-fit p_T distributions of the lepton-jet system in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



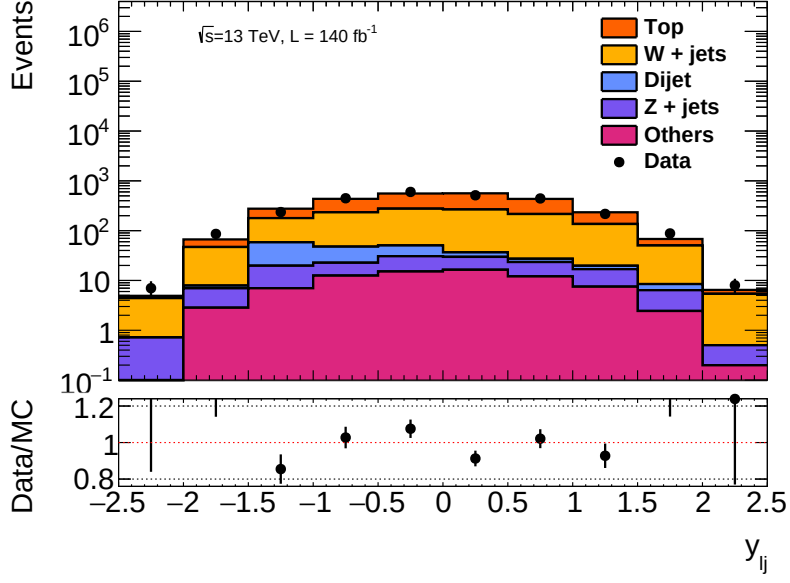
(a) WCR.



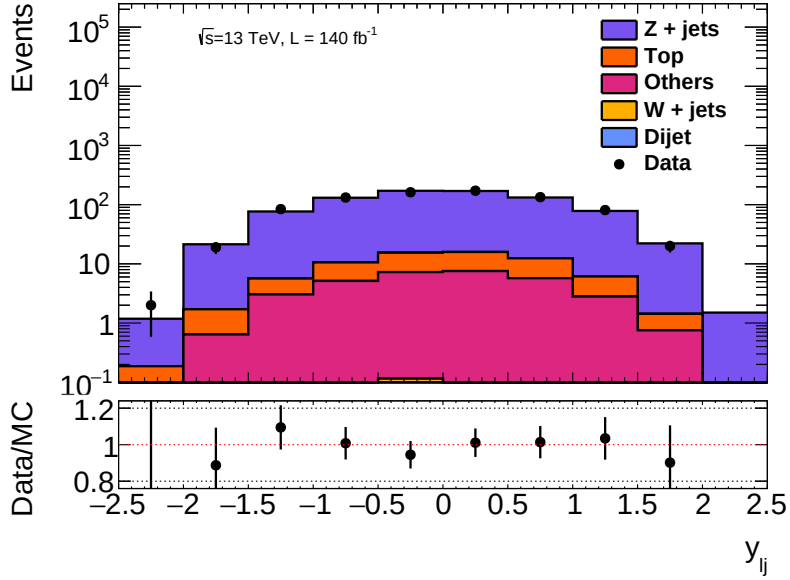
(b) ZCR.

Figure 7.13.: Post-fit ϕ distributions of the lepton-jet system in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



(a) WCR.



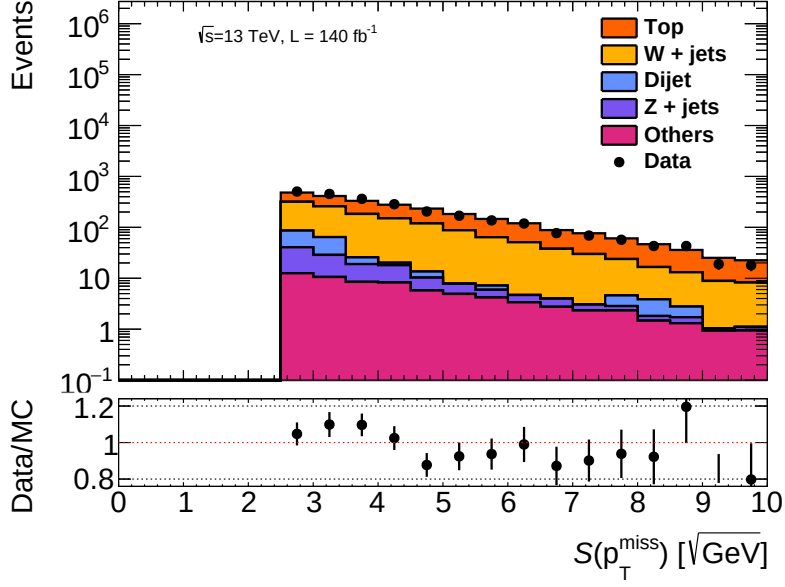
(b) ZCR.

Figure 7.14.: Post-fit y distributions of the lepton-jet system in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

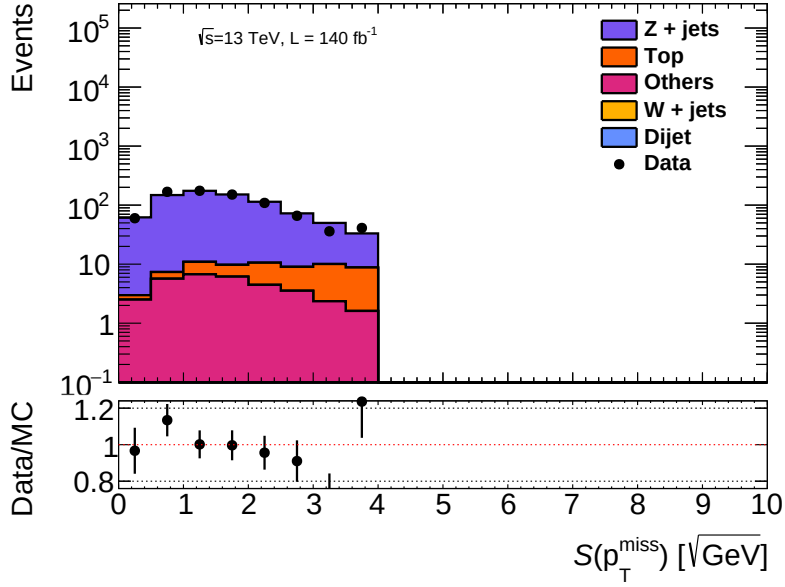
7. Results

Lastly, the cut parameters are inspected, namely $\mathcal{S}(p_T^{\text{miss}})$ in figure 7.15, $p_T^{\ell 1}/m_{\ell j}$ in figure 7.16 and $m_{\ell\ell}$ in figure 7.17. The $\mathcal{S}(p_T^{\text{miss}})$ distributions are quite different due to being the inverted cut in the WCR. The general shape agrees for both regions. The WCR stays within 2σ throughout, often being within or close to 1σ and the ZCR is often within 1σ and always within 2σ . There might be a slight slope visible in the data/MC ratio of the ZCR but it could very well be statistical fluctuations. The distribution of $p_T^{\ell 1}/m_{\ell j}$ also agree in the general shape of the distributions but when going to a high ratio the MC seems to increasingly disagree with collision data. This is still within 2σ but hints to a possible mismodeling of the tail especially in the ZCR. Again, due to limited statistics this could just be fluctuation since the region up to 0.6 for the WCR and 0.5 for the ZCR agrees very well with the data. When looking at the $m_{\ell\ell}$ distribution the Z -peak is clearly visible in MC and data with both having a ratio of close to 1 at the Z -peak. The region below and above the Z mass is dominated by low statistics but seems to agree within the errors with the data. There is no $m_{\ell\ell}$ distribution for the WCR shown because this region requires only one lepton making it impossible to create $m_{\ell\ell}$.

7. Results



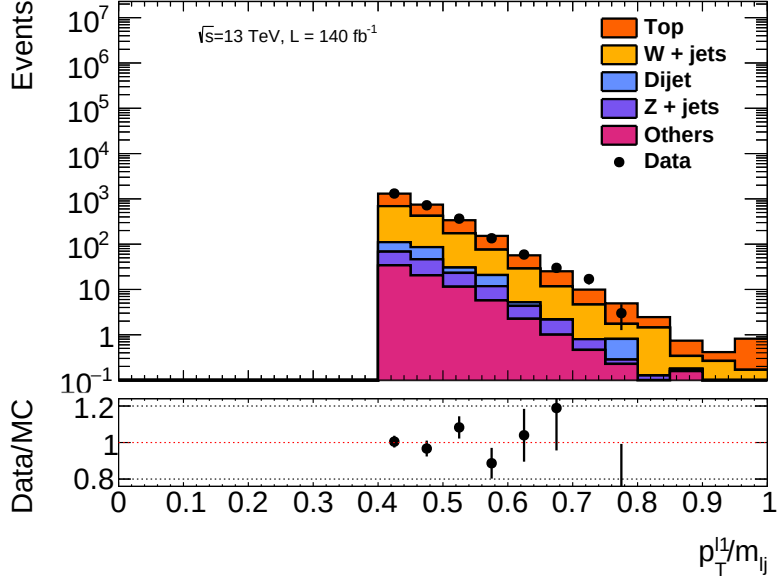
(a) WCR.



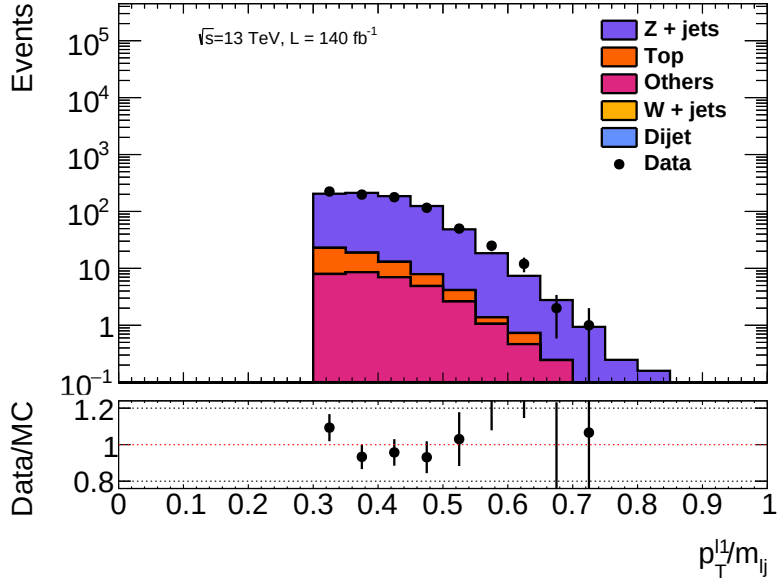
(b) ZCR.

Figure 7.15.: Post-fit $S(p_T^{\text{miss}})$ distributions in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results



(a) WCR.



(b) ZCR.

Figure 7.16.: Post-fit p_T^1/m_{lj} distributions in the WCR and ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7. Results

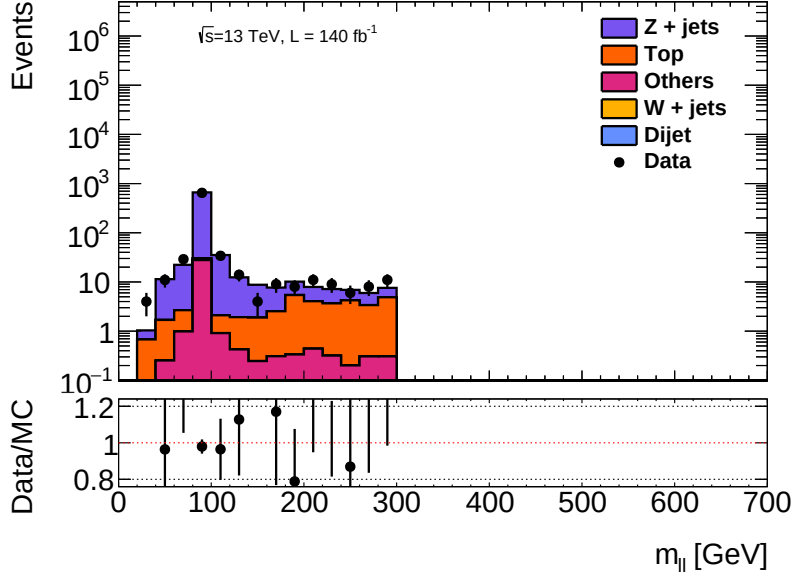


Figure 7.17.: Post-fit $m_{\ell\ell}$ distribution in the ZCR. All final cuts and the normalization factors are applied. The background events are sorted by the total number of events in the region.

7.2. Expected Exclusion Limit

Before looking at the combined expected exclusion limit which gives the final results of this thesis the separated expected exclusion limits are shown. Figure 7.18 shows both exclusion limits within the same plot to have a side-by-side comparison. λ_{31} represents the coupling to a third generation quark and a first generation lepton and the figure also contains current limits on the exclusion limit for LQs coupling to an electron and a b -quark. These limits are taken from [64] for the pair-production limit and from [70] for the Drell-Yan limit where the Drell-Yan limit is extracted from the corresponding figure (see figure 2.9). Both regions give roughly the same exclusion limit with both capable of covering a new phase space in the 1.8 to 2.2 TeV region with LQs of at least a coupling of 1.8 if not considering the σ -bands. As it turns out the 2L region does not provide a better exclusion limit although previous plots (e.g. figure 5.28) suggested that the significance is greatly increased for the 2L region. The reason for that is the smaller statistics which take a huge part in calculating the exclusion limit during the statistical analysis. The multiple crossings of the exclusion limits have no physical reason because the decision, on which of these two exclusion limits perform better, should not be dependent on the LQ mass. The kinematics shown in section 5.3 do not suggest any LQ mass dependent anomalies regarding the different lepton channels. The only reason can be that these differences are simply statistical fluctuations which is also hinted at by the fact that the two exclusion limits are very

7. Results

well within the 1σ bands of each other.

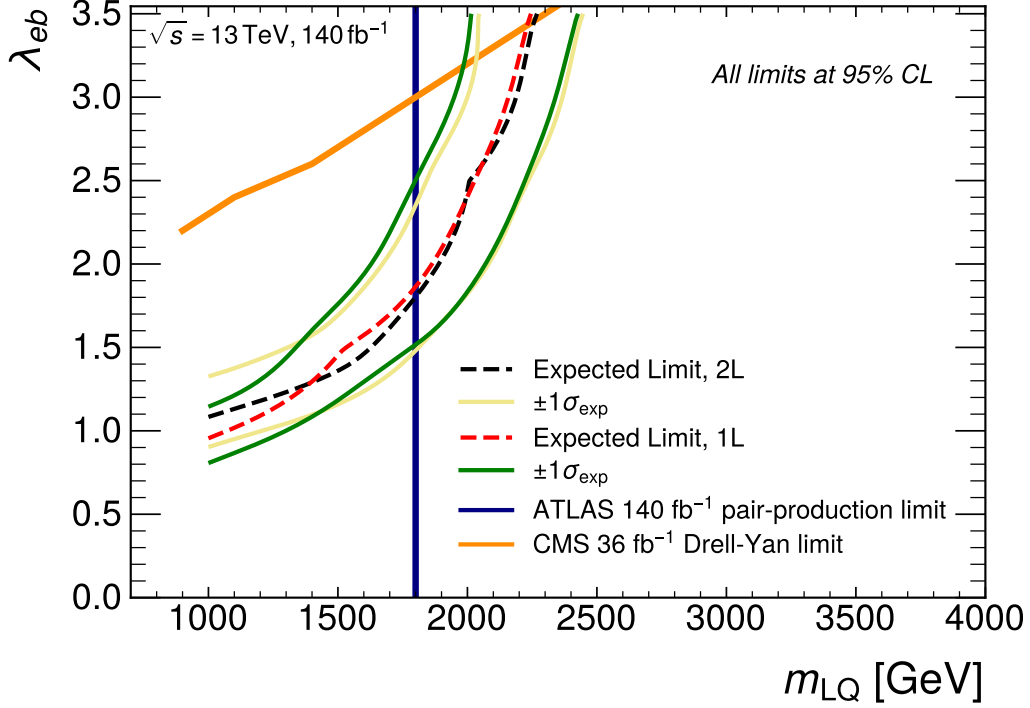


Figure 7.18.: Expected exclusion limit for the 1L and 2L region when used separately including 1σ bands.

Coming to the combined exclusion limit, the final values for each signal hypothesis are found in table 7.1. The LQ model with 1 TeV mass and a coupling of 3.5 and the LQ model with a mass of 3.5 TeV and a coupling of 0.5 were both dropped from the final analysis. The reason to do this is different for both.

The 1 TeV, 3.5 coupling model was dropped because it produced NaNs when applying the asymptotic method which could not be fully resolved. The most likely reason for this occurrence are wrongly filled systematics. This is not too worrying because this model is way above the final exclusion limit and does not change the result of the fit at all. The second one was dropped because the cross section of a LQ (coupling to e and b with a mass of 3.5 TeV and a coupling of 0.5) is just too small to produce any significant amount of events during the whole Run 2 of the ATLAS experiment. This results in the asymptotic method yielding a p-value of 1 because of the falling floating number precision around 1 [192]. This results in a significance of - infinity which also does not work in this plot. Dropping this is also not problematic because this model is way below the exclusion limit and there are plenty of other models closer to the exclusion limit shaping the curve. This means that removing this model does not change the final result as well.

7. Results

Table 7.1.: Exclusion Limit table for the combined exclusion limit using the asymptotic method. The $x_{\pm 1\sigma}$ parameters are the p-values and the significances of the $\pm 1\sigma$ -bands.

Mass [GeV]	Coupling	p-value	p-value $_{+1\sigma}$	p-value $_{-1\sigma}$	Significance	Significance $_{+1\sigma}$	Significance $_{-1\sigma}$
1000	0.5	0.2971	0.5741	0.1295	0.5328	-0.1867	1.1289
1000	1.5	$4.631 \cdot 10^{-9}$	$6.979 \cdot 10^{-7}$	$2.172 \cdot 10^{-11}$	5.744	4.826	6.592
1000	2.5	$2.469 \cdot 10^{-22}$	$1.648 \cdot 10^{-18}$	$2.574 \cdot 10^{-26}$	9.650	8.701	10.55
1000	3.5	NaN	NaN	NaN	NaN	NaN	NaN
1500	0.5	0.4494	0.7087	0.2491	0.1270	-0.5496	0.6773
1500	1.5	$1.653 \cdot 10^{-5}$	$5.600 \cdot 10^{-4}$	$3.509 \cdot 10^{-7}$	4.151	3.258	4.961
1500	2.5	$1.240 \cdot 10^{-13}$	$8.512 \cdot 10^{-11}$	$1.266 \cdot 10^{-16}$	7.320	6.386	8.194
1500	3.5	$2.068 \cdot 10^{-23}$	$1.766 \cdot 10^{-19}$	$1.684 \cdot 10^{-27}$	9.901	8.951	10.80
2000	0.5	0.7562	0.8971	0.5989	-0.6940	-1.265	-0.2504
2000	1.5	0.05697	0.2177	0.01163	1.581	0.7801	2.269
2000	2.5	$2.704 \cdot 10^{-4}$	$4.896 \cdot 10^{-3}$	$1.087 \cdot 10^{-5}$	3.460	2.583	4.246
2000	3.5	$3.306 \cdot 10^{-7}$	$2.402 \cdot 10^{-5}$	$3.238 \cdot 10^{-9}$	4.972	4.065	5.804
2500	0.5	0.9037	0.96313	0.8266	-1.303	-1.788	-0.9407
2500	1.5	0.4241	0.6888	0.2268	0.1915	-0.4924	0.7495
2500	2.5	0.1160	0.3373	0.03190	1.195	0.4199	1.854
2500	3.5	0.02615	0.1313	$3.983 \cdot 10^{-3}$	1.941	1.120	2.653
3000	0.5	0.9699	0.9889	0.9435	-1.879	-2.288	-1.585
3000	1.5	0.7870	0.9119	0.6429	-0.7961	-1.353	-0.3663
3000	2.5	0.5401	0.7734	0.3366	-0.1006	-0.7502	0.4218
3000	3.5	0.3113	0.5884	0.1392	0.4921	-0.2234	1.0839
3500	0.5	1.0	1.0	1.0	$-\infty$	$-\infty$	$-\infty$
3500	1.5	0.9234	0.9710	0.86037	-1.428	-1.896	-1.082
3500	2.5	0.8293	0.9313	0.7063	-0.9512	-1.486	-0.5426
3500	3.5	0.6871	0.8615	0.5064	-0.4876	-1.087	-0.01592

7. Results

The expected exclusion limit plot is shown in figure 7.19. The exclusion limit combines somewhat the results of the 1L and 2L by considering more SRs in the statistical analysis. This results in better exclusion limits than the separated ones and also a tighter σ -band. The sensitive mass range for the expected exclusion limit is from 1.8 to 2.6 TeV with minimal coupling of 1.2, again not considering the σ -band. The covered phase space increased drastically by using the 1L and 2L regions together.

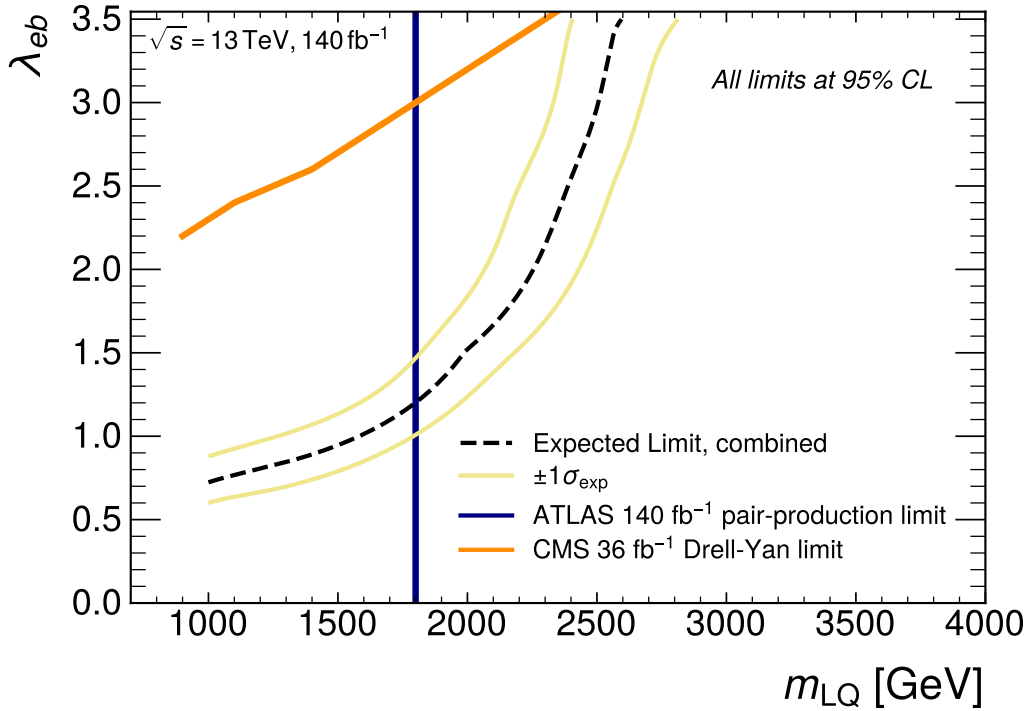


Figure 7.19.: Expected exclusion limit after combining the 1L and 2L region results including 1σ bands.

In order to verify the results of the asymptotic method, a toy model cross check has been employed. The results for each signal hypothesis are shown in table 7.2 when applying the toy model. This created many results which are not applicable, like significances of $\pm\infty$, p-values of 0 or 1 and NaNs. The toy model approach had a couple of difficulties which could not be resolved within the time frame of this thesis. One of the reasons for the values of $0/\infty$ could be the p-values being set to 0 if they become smaller than 10^{-4} .

Nonetheless, these results can be used to confirm that the asymptotic method mostly agrees with the toy model results in the most sensitive region for models of 2.5 TeV and 3 TeV and the low coupling 2 TeV samples. Apart from the 1 TeV, 0.5 coupling model, the p-values agree when within $10^{-4} < p < 0.97$. Out of this range the asymptotic method starts to deviate from the toy model results which could originate from

7. Results

errors of the toy model code as well as the breakdown of the asymptotic method approach. Most likely it is a combination of both. The values of the σ -bands also seem to deviate more than the central values. The values vary both up and down seeming to apply an uncertainty but not an obvious bias on the results.

Table 7.2.: Exclusion Limit table for the combined exclusion limit using the toy model approach. The $x_{\pm 1\sigma}$ parameters are the p-values and the significances of the $\pm 1\sigma$ -bands.

Mass [GeV]	Coupling	p-value	p-value $_{+1\sigma}$	p-value $_{-1\sigma}$	Significance	Significance $_{+1\sigma}$	Significance $_{-1\sigma}$
1000	0.5	0.1139	0.3901	0.02524	1.206	-0.2791	1.956
1000	1.5	0.0	0.0	0.0	∞	∞	∞
1000	2.5	0.0	0.0	0.0	∞	∞	∞
1000	3.5	NaN	NaN	NaN	NaN	NaN	NaN
1500	0.5	0.4286	1.0	0.2674	0.1798	$-\infty$	0.6206
1500	1.5	0.0	$5.685 \cdot 10^{-4}$	0.0	∞	3.254	∞
1500	2.5	0.0	0.0	0.0	∞	∞	∞
1500	3.5	0.0	0.0	0.0	∞	∞	∞
2000	0.5	0.7710	1.0	0.6120	-0.7421	$-\infty$	-0.2846
2000	1.5	0.05944	0.1988	0.02657	1.560	-0.8459	1.934
2000	2.5	NaN	NaN	NaN	NaN	NaN	NaN
2000	3.5	0.0	0.0	0.0	∞	∞	∞
2500	0.5	0.9283	1.	0.8379	-1.464	$-\infty$	-0.9858
2500	1.5	0.4208	1.0	0.3433	0.1998	$-\infty$	0.4035
2500	2.5	0.1012	0.2845	0.05689	1.275	0.5697	1.581
2500	3.5	0.02639	0.09582	$3.041 \cdot 10^{-3}$	1.937	1.306	2.743
3000	0.5	0.9741	1.0	0.9347	-1.945	$-\infty$	-1.512
3000	1.5	0.8267	1.0	0.7695	-0.9412	$-\infty$	-0.7372
3000	2.5	0.5344	1.0	0.4282	-0.08639	$-\infty$	0.1808
3000	3.5	0.3052	0.5525	0.1722	0.5095	-0.1321	0.9454
3500	0.5	1.0	1.001	0.9511	$-\infty$	NaN	-1.656
3500	1.5	0.9618	1.0	0.9397	-1.7722	$-\infty$	-1.552
3500	2.5	0.8614	1.0	0.7209	-1.086	$-\infty$	-0.5857
3500	3.5	0.6849	1.0	0.5855	-0.4814	$-\infty$	-0.2160

The results of the fitting procedure also need to be verified to give meaningful and sensible outputs to ensure that the expected exclusion limits can be assumed to be applicable. For this, the fit parameters including the normalization factors are given in a pulls and correlation matrix. The results of the normalization factors are shown in figure 7.20 in combination with the systematic uncertainties which are all set to 0.0000 ± 0.9934 . The systematic uncertainties starting with "EL" correspond to uncertainties related to the electron, "FT" are uncertainties related to b -tagging and "JET" handles uncertainties on the jet reconstruction, namely the JVT efficiency. Figure 7.21 shows the correlation matrix of this fit. The most important things here are that the systematic errors are uncorrelated and that they are in general not or weakly correlated to the normalization factors. The only exception being the JVT efficiency systematic which is expected to correlate to the normalization factors. The correlation of the normalization factors to one another is expected because they directly impact the number of events in the CRs. These plots give rise to the conclusion that the fitting procedure was successful and did not include any unwanted correlations or

7. Results

unexpected large errors.

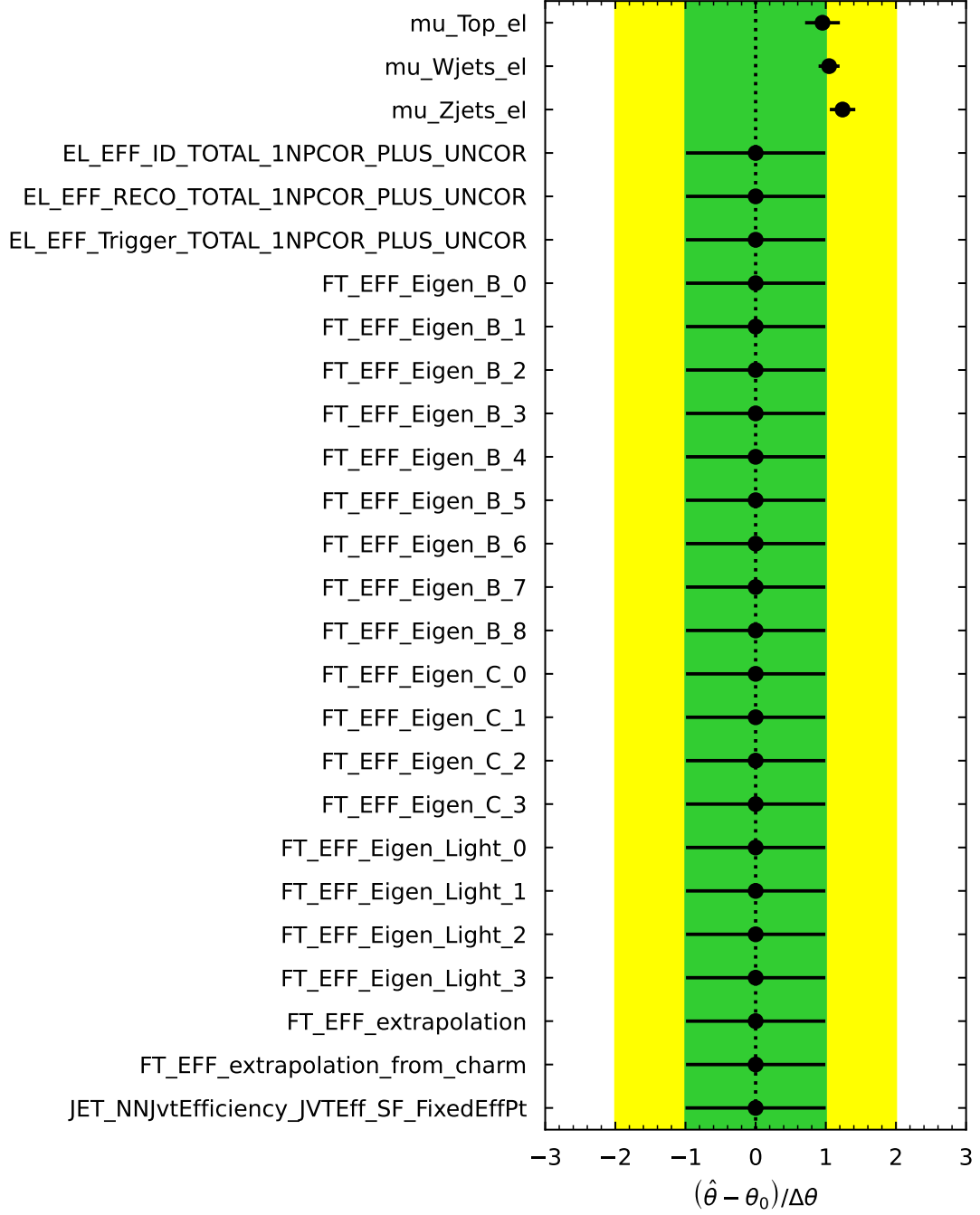


Figure 7.20.: Pulls plot of the fit which gives the fit value with its uncertainty for the normalization parameters and all implemented systematics.

7. Results

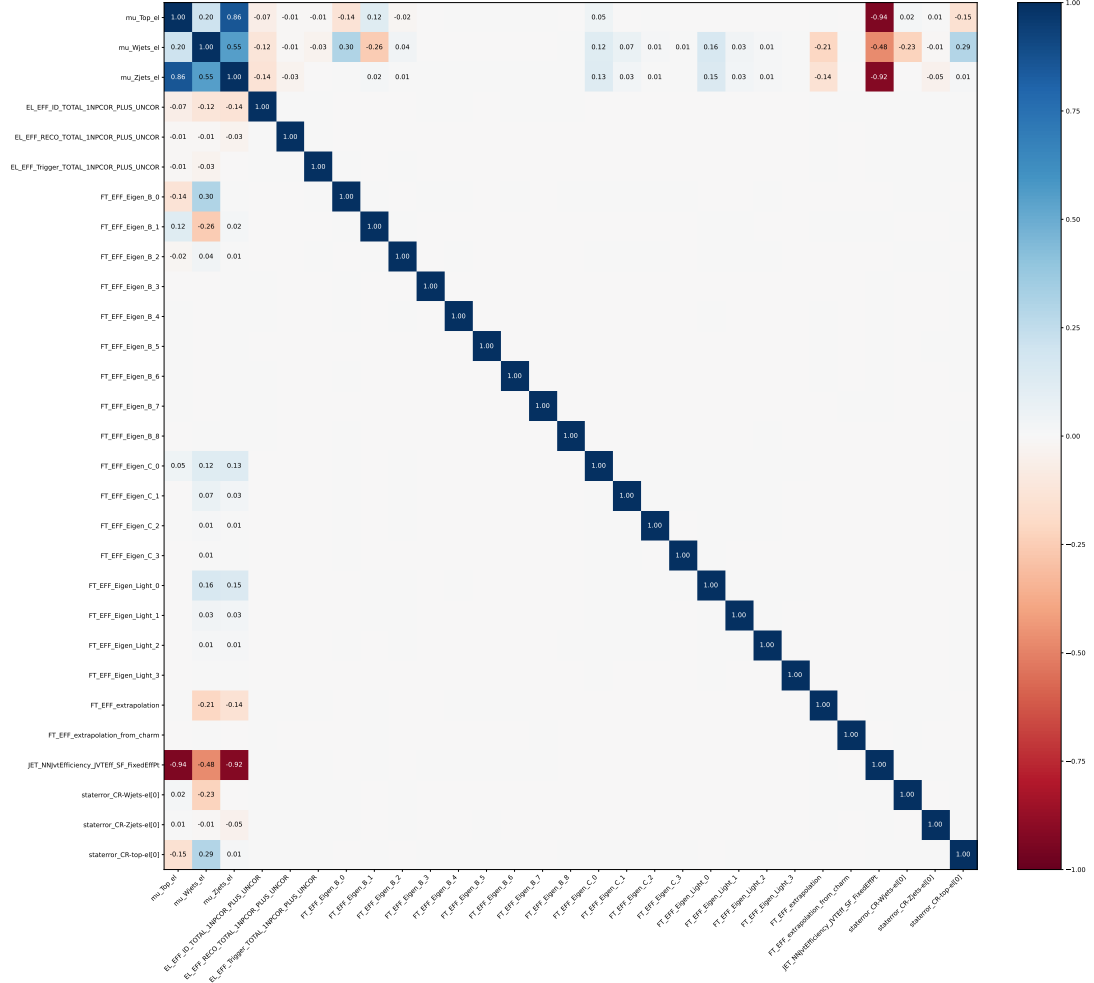


Figure 7.21.: Correlation plot of the fit. The ordering of the parameters (top to bottom, left to right) are the three normalization parameters of Top, W, Z, the systematic errors as ordered in figure 7.20 and the three statistical uncertainties for W, Z and Top.

8. Conclusion and Outlook

This thesis tackled the question whether the resonant Leptoquark production is capable of giving new exclusion limits in the electron + b -quark channel using the whole Run 2 data of the ATLAS experiment having a center-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 140 fb^{-1} . The used Leptoquark model is a scalar $\tilde{S}_1$ Leptoquark only coupling to an electron and a b -quark and the resonant search uses state-of-the-art NLO calculations including a dedicated lepton PDF.

To gain the best possible exclusion limits a cut flow is introduced trying to maximize background rejection while keeping the signal efficiency as high as possible. The dominant background processes are simulated using Monte Carlo simulations to estimate the expected data yield. For this, a control region is defined free from any signal contamination, the MC samples are scaled to fit the data yields in this region and are then tested in a validation region. The resonant structure is then scanned for in the invariant mass distribution of the lepton-jet system. Additionally, the background contribution containing dijets has been tested to estimate using a data-driven background estimation but it turned out to give similar yields as the MC simulation of these backgrounds. A statistical analysis was employed to create an expected exclusion limit using the asymptotic method.

The results clearly show a possible improvement in the exclusion limits compared to previous searches in the pair-production channel, single-production channel or Drell-Yan searches being sensitive to previously unreachable masses and couplings. Due to the fact that the resonant Leptoquark production is both mass and coupling dependent, no clear mass or coupling threshold is set but rather an excluded area is defined in the mass-coupling plane. This newly covered area ranges from 1.8 TeV and a coupling λ of 1.2 to a mass of 2.6 TeV with a coupling λ of 3.5 at a 95% confidence limit.

Future steps would be adding the missing systematic uncertainties and removing the applied flat dummy systematic in the signal regions that were added to account for the missing systematics. Also, cross checking the results with a fully working toy model would be very sensible because the statistics in some regions are low enough that the asymptotic method applied might break down and overestimate the exclusion limit. A limited toy model comparison was done but some limitations occurred which could not be resolved within the time frame of this thesis. To expand on this thesis one could add Run 3 data which should also yield better exclusion limits due to the higher combined statistics, but also the higher center-of-mass energy resulting in increased cross sections for a potential LQ production. Another intuitive and logical next step to do is getting an observed exclusion limit in addition to

8. *Conclusion and Outlook*

an expected exclusion limit by unblinding the data, which was unfortunately not yet allowed since an internal approval of the analysis was pending at the time of writing this thesis. Nonetheless, the results of this thesis showed that the resonant Leptoquark production is a promising approach to create a better exclusion limit and the insights gained in this thesis will be integrated in a publication by ATLAS being released in the near future.

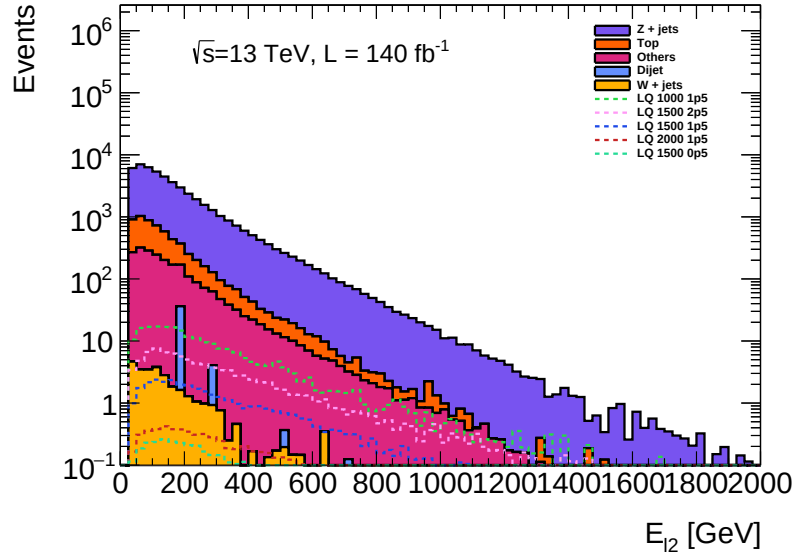
A. Appendix

A.1. Tables and Figures

A.1.1. Subleading Lepton Kinematics

The following plots will show the kinematics of the subleading lepton mentioned in section 5.3.3.

Figure A.1.: Energy distribution of the subleading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.



A. Appendix

Figure A.2.: p_T distribution of the subleading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.

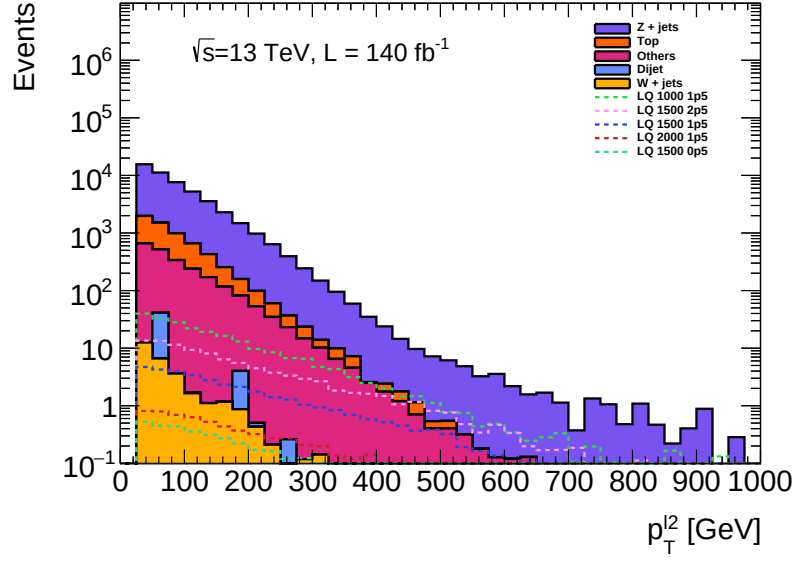
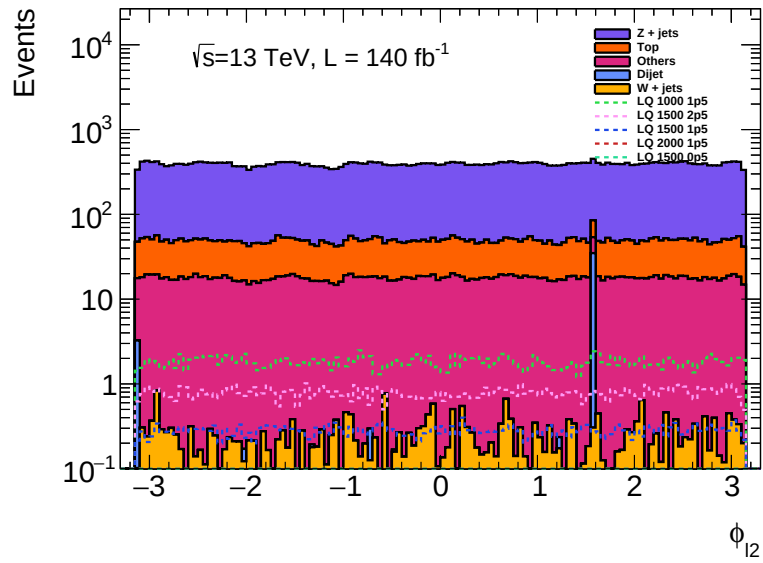
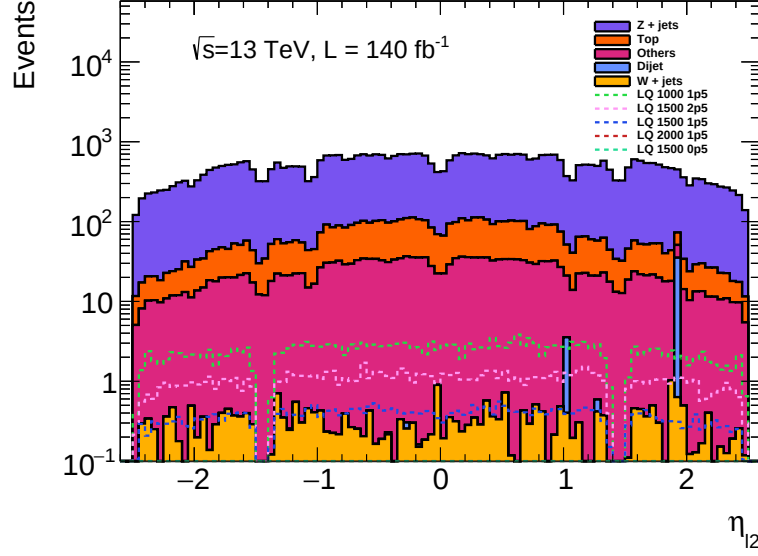


Figure A.3.: ϕ distribution of the subleading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.



A. Appendix

Figure A.4.: η distribution of the subleading lepton. Only preselections are applied. Both the background and signal events are sorted by the total number of events separately.



A.2. Used KI-Tools

There have been no KI-Tools used throughout the creation of this thesis.

B. Bibliography

- [1] Albert Einstein.
“Zur Elektrodynamik bewegter Körper”.
In: *Annalen Phys.* 322.10 (1905), pp. 891–921.
DOI: 10.1002/andp.19053221004.
- [2] Albert Einstein.
“Zur Allgemeinen Relativitätstheorie”.
In: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1915 (1915).
[Addendum: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1915, 799–801 (1915)], pp. 778–786.
- [3] G. Aad et al.
“Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”.
In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29.
ISSN: 0370-2693.
DOI: 10.1016/j.physletb.2012.08.020.
- [4] S. Chatrchyan et al.
“Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”.
In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61.
ISSN: 0370-2693.
URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [5] Lyndon Evans and Philip Bryant.
“LHC Machine”.
In: *JINST* 3 (Aug. 2008), S08001.
DOI: 10.1088/1748-0221/3/08/S08001.
- [6] G. Aad et al.
“The ATLAS Experiment at the CERN Large Hadron Collider”.
In: *JINST* 3 (Aug. 2008), S08003.
DOI: 10.1088/1748-0221/3/08/S08003.
- [7] S. Chatrchyan et al.
“The CMS Experiment at the CERN LHC”.
In: *JINST* 3 (Aug. 2008), S08004.
DOI: 10.1088/1748-0221/3/08/S08004.
- [8] Frederic Fischer.

B. Bibliography

- “Search for long-lived axion-like particles from flavour-changing neutral current exotic Top quark decays with the ATLAS experiment”.
MA thesis. Mainz U., July 2024.
- [9] David J Griffiths.
Introduction to elementary particles; 2nd rev. version.
Physics textbook.
New York, NY: Wiley, 2008.
URL: <https://cds.cern.ch/record/111880>.
- [10] Matthew D. Schwartz.
Quantum Field Theory and the Standard Model.
Cambridge University Press, Mar. 2014.
ISBN: 978-1-107-03473-0.
- [11] Mark Thomson.
Modern particle physics.
New York: Cambridge University Press, 2013.
ISBN: 978-1-107-03426-6.
DOI: 10.1017/CB09781139525367.
- [12] Max Planck.
“Über irreversible Strahlungsvorgänge”.
In:
Von Kirchhoff bis Planck: Theorie der Wärmestrahlung in historisch-kritischer Darstellung.
Ed. by Hans-Georg Schöpf.
Wiesbaden: Vieweg+Teubner Verlag, 1978,
Pp. 166–171.
ISBN: 978-3-663-13885-3.
DOI: 10.1007/978-3-663-13885-3_13.
- [13] S. N. Bose.
“Plancks Gesetz und Lichtquantenhypothese”.
In: *Zeitschrift für Physik* 26.1 (Dec. 1924), pp. 178–181.
ISSN: 0044-3328.
DOI: 10.1007/BF01327326.
- [14] Enrico Fermi.
On the Quantization of the Monoatomic Ideal Gas.
Translation by Alberto Zannoni.
Dec. 1999.
DOI: 10.48550/arXiv.cond-mat/9912229.
- [15] W. Pauli.
“Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren”.
In: *Zeitschrift für Physik* 31.1 (Feb. 1925), pp. 765–783.
ISSN: 0044-3328.

B. Bibliography

DOI: 10.1007/BF02980631.

- [16] B. T. Cleveland et al.
“Measurement of the solar electron neutrino flux with the Homestake chlorine detector”.
In: *Astrophys. J.* 496 (1998), pp. 505–526.
DOI: 10.1086/305343.
- [17] Q. R. Ahmad et al.
“Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory”.
In: *Physical Review Letters* 89.1 (June 2002).
ISSN: 1079-7114.
DOI: 10.1103/physrevlett.89.011301.
- [18] G. Bellini et al.
“Neutrino Oscillations”.
In: *Advances in High Energy Physics* 2014 (2014), pp. 1–28.
ISSN: 1687-7365.
DOI: 10.1155/2014/191960.
- [19] S. Navas et al.
“Review of particle physics”.
In: *Phys. Rev. D* 110.3 (2024), p. 030001.
DOI: 10.1103/PhysRevD.110.030001.
- [20] Gerard 't Hooft.
“A Two-Dimensional Model for Mesons”.
In: *Nucl. Phys. B* 75 (1974), pp. 461–470.
DOI: 10.1016/0550-3213(74)90088-1.
- [21] David J. Gross and Frank Wilczek.
“Ultraviolet Behavior of Nonabelian Gauge Theories”.
In: *Phys. Rev. Lett.* 30 (1973). Ed. by J. C. Taylor, pp. 1343–1346.
DOI: 10.1103/PhysRevLett.30.1343.
- [22] Kenneth G. Wilson.
“Confinement of Quarks”.
In: *Phys. Rev. D* 10 (1974). Ed. by J. C. Taylor, pp. 2445–2459.
DOI: 10.1103/PhysRevD.10.2445.
- [23] Murray Gell-Mann.
“The Eightfold Way: A Theory of strong interaction symmetry”.
In: (Mar. 1961).
DOI: 10.2172/4008239.
- [24] Peter W. Higgs.
“Broken symmetries, massless particles and gauge fields”.
In: *Phys. Lett.* 12 (1964), pp. 132–133.
DOI: 10.1016/0031-9163(64)91136-9.

B. Bibliography

- [25] F. Englert and R. Brout.
“Broken Symmetry and the Mass of Gauge Vector Mesons”.
In: *Phys. Rev. Lett.* 13 (Aug. 1964). Ed. by J. C. Taylor, pp. 321–323.
DOI: 10.1103/PhysRevLett.13.321.
- [26] Peter W. Higgs.
“Spontaneous Symmetry Breakdown without Massless Bosons”.
In: *Phys. Rev.* 145 (1966), pp. 1156–1163.
DOI: 10.1103/PhysRev.145.1156.
- [27] E. Schrödinger.
“Quantisierung als Eigenwertproblem”.
In: *Annalen Phys.* 384.4 (1926), pp. 361–376.
DOI: 10.1002/andp.19263840404.
- [28] E. Schrödinger.
“An Undulatory Theory of the Mechanics of Atoms and Molecules”.
In: *Phys. Rev.* 28.6 (1926), p. 1049.
DOI: 10.1103/PhysRev.28.1049.
- [29] Paul A. M. Dirac.
“The quantum theory of the electron”.
In: *Proc. Roy. Soc. Lond. A* 117 (1928), pp. 610–624.
DOI: 10.1098/rspa.1928.0023.
- [30] Paul A. M. Dirac.
“The quantum theory of the electron. Part II”.
In: *Proc. Roy. Soc. Lond. A* 118 (1928), pp. 351–161.
DOI: 10.1098/rspa.1928.0056.
- [31] W. Heisenberg.
“Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik”.
In: *Z. Phys.* 43 (1927), pp. 172–198.
DOI: 10.1007/BF01397280.
- [32] Richard Phillips Feynman.
“The principle of least action in quantum mechanics”.
PhD thesis. Princeton U., 1942.
DOI: 10.1142/9789812567635_0001.
- [33] E. Noether.
“Invariante Variationsprobleme”. ger.
In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257.
URL: <http://eudml.org/doc/59024>.
- [34] S. L. Glashow.
“Partial Symmetries of Weak Interactions”.
In: *Nucl. Phys.* 22 (1961), pp. 579–588.

B. Bibliography

- DOI: 10.1016/0029-5582(61)90469-2.
- [35] Steven Weinberg.
“A Model of Leptons”.
In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
DOI: 10.1103/PhysRevLett.19.1264.
- [36] Abdus Salam.
“Weak and Electromagnetic Interactions”.
In: *Conf. Proc. C* 680519 (1968), pp. 367–377.
DOI: 10.1142/9789812795915_0034.
- [37] Walter Greiner and Joachim Reinhardt.
Field Quantization.
Springer, 1996.
ISBN: 978-3-642-61485-9.
DOI: 10.1007/978-3-642-61485-9.
- [38] R. P. Feynman.
“The Theory of positrons”.
In: *Phys. Rev.* 76 (6 Sept. 1949), pp. 749–759.
DOI: 10.1103/PhysRev.76.749.
- [39] Tung-Mow Yan and Sidney D. Drell.
“The Parton Model and its Applications”.
In: *Int. J. Mod. Phys. A* 29.30 (2014), p. 1430071.
DOI: 10.1142/S0217751X14300713.
arXiv: 1409.0051 [hep-ph].
- [40] R. P. Feynman.
“The behavior of hadron collisions at extreme energies”.
In: *Conf. Proc. C* 690905 (1969), pp. 237–258.
- [41] Richard D. Ball et al.
“Parton distributions from high-precision collider data”.
In: *Eur. Phys. J. C* 77.10 (2017), p. 663.
DOI: 10.1140/epjc/s10052-017-5199-5.
arXiv: 1706.00428 [hep-ph].
- [42] S. Chekanov et al.
“Measurement of the neutral current cross-section and $F(2)$ structure function for deep inelastic $e + p$ scattering at HERA”.
In: *Eur. Phys. J. C* 21 (2001), pp. 443–471.
DOI: 10.1007/s100520100749.
arXiv: hep-ex/0105090.
- [43] Guido Altarelli and G. Parisi.
“Asymptotic Freedom in Parton Language”.
In: *Nucl. Phys. B* 126 (1977), pp. 298–318.
DOI: 10.1016/0550-3213(77)90384-4.

B. Bibliography

- [44] Yuri L. Dokshitzer.
“Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.”
In: *Sov. Phys. JETP* 46 (1977), pp. 641–653.
- [45] Richard D. Ball et al.
“Parton distributions for the LHC Run II”.
In: *JHEP* 2015.4 (Apr. 2015), p. 040.
ISSN: 1029-8479.
DOI: 10.1007/JHEP04(2015)040.
arXiv: 1410.8849 [hep-ph].
- [46] J. Pumplin et al.
“Uncertainties of predictions from parton distribution functions. 2. The Hessian method”.
In: *Phys. Rev. D* 65 (2001), p. 014013.
DOI: 10.1103/PhysRevD.65.014013.
arXiv: hep-ph/0101032.
- [47] Luca Buonocore et al.
“Leptons in the proton”.
In: *JHEP* 08.08 (2020), p. 019.
DOI: 10.1007/JHEP08(2020)019.
arXiv: 2005.06477 [hep-ph].
- [48] Stefan Höche.
“Introduction to Parton-Shower Event Generators”.
In:
Journeys Through the Precision Frontier: Amplitudes for Colliders.
World Scientific, Nov. 2015.
Chap. 5, pp. 235–295.
DOI: 10.1142/9789814678766_0005.
arXiv: 1411.4085 [hep-ph].
- [49] Zachary Marshall.
“Simulation of Pile-up in the ATLAS Experiment”.
In: *J. Phys. Conf. Ser.* 513 (2014). Ed. by D. L. Groep and D. Bonacorsi,
p. 022024.
DOI: 10.1088/1742-6596/513/2/022024.
- [50] B. C. Paul, S. Ghose, and P. Thakur.
“Emergent Universe from a composition of matter, exotic matter and dark energy”.
In: *Monthly Notices of the Royal Astronomical Society* 413.1 (Apr. 2011),
pp. 686–690.
ISSN: 0035-8711.
DOI: 10.1111/j.1365-2966.2010.18177.x.

B. Bibliography

- eprint: <https://academic.oup.com/mnras/article-pdf/413/1/686/18593530/mnras0413-0686.pdf>.
- [51] A. D. Sakharov.
“Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe”.
In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35.
DOI: 10.1070/PU1991v034n05ABEH002497.
- [52] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali.
“The Hierarchy problem and new dimensions at a millimeter”.
In: *Phys. Lett. B* 429.3 (1998), pp. 263–272.
DOI: 10.1016/S0370-2693(98)00466-3.
arXiv: hep-ph/9803315.
- [53] Lisa Randall and Raman Sundrum.
“Large Mass Hierarchy from a Small Extra Dimension”.
In: *Phys. Rev. Lett.* 83 (17 Oct. 1999), pp. 3370–3373.
DOI: 10.1103/PhysRevLett.83.3370.
arXiv: hep-ph/9905221.
- [54] R. Aaij et al.
“Test of Lepton Universality Using $B^+ \rightarrow K^+ \ell^+ \ell^-$ Decays”.
In: *Phys. Rev. Lett.* 113 (15 Oct. 2014), p. 151601.
DOI: 10.1103/PhysRevLett.113.151601.
- [55] R. Aaij et al.
“Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays”.
In: *Journal of High Energy Physics* 2017.8 (Aug. 2017), p. 55.
ISSN: 1029-8479.
DOI: 10.1007/JHEP08(2017)055.
- [56] H. Georgi and S. L. Glashow.
“Unity of All Elementary-Particle Forces”.
In: *Phys. Rev. Lett.* 32 (8 Feb. 1974), pp. 438–441.
DOI: 10.1103/PhysRevLett.32.438.
- [57] Stephen P. Martin.
“A Supersymmetry primer”.
In: *Adv. Ser. Direct. High Energy Phys.* 18 (1998). Ed. by Gordon L. Kane, pp. 1–98.
DOI: 10.1142/9789812839657_0001.
arXiv: hep-ph/9709356.
- [58] Paul H. Frampton and Bum-Hoon Lee.
“SU(15) grand unification”.
In: *Phys. Rev. Lett.* 64 (6 Feb. 1990), pp. 619–621.
DOI: 10.1103/PhysRevLett.64.619.
- [59] Edward Witten.

B. Bibliography

- “Symmetry breaking patterns in superstring models”.
In: *Nucl. Phys. B* 258 (1985), pp. 75–100.
ISSN: 0550-3213.
DOI: 10.1016/0550-3213(85)90603-0.
- [60] Michael Dine et al.
“Superstring Model Building”.
In: *Nucl. Phys. B* 259.4 (1985), pp. 549–571.
ISSN: 0550-3213.
DOI: 10.1016/0550-3213(85)90001-X.
- [61] I. Doršner et al.
“Physics of leptoquarks in precision experiments and at particle colliders”.
In: *Physics Reports* 641 (June 2016), pp. 1–68.
ISSN: 0370-1573.
DOI: 10.1016/j.physrep.2016.06.001.
- [62] Georges Aad et al.
Leptoquark summary plot for scalar or vector models.
Tech. rep.
Geneva: CERN, 2024.
URL: <https://cds.cern.ch/record/2903898>.
- [63] Michael Holzbock et al.
Search for resonant leptoquark production in lepton-jet invariant mass spectra in Run 2 and Run 3 pp collision data of the ATLAS experiment.
Tech. rep.
Internal documentation, restricted access.
Geneva: CERN, 2024.
URL: <https://cds.cern.ch/record/2903750>.
- [64] Georges Aad et al.
“Search for pairs of scalar leptoquarks decaying into quarks and electrons or muons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector”.
In: *Journal of High Energy Physics* 2020.10 (Oct. 2020), p. 112.
ISSN: 1029-8479.
DOI: 10.1007/jhep10(2020)112.
- [65] Anton Wolf.
“Search for pair-produced first and second generation scalar leptoquarks in pp collisions at $\sqrt{s} = 13$ TeV & work with the liquid argon purity monitoring system in the ATLAS experiment”.
Presented 05 Mar 2021.
PhD thesis. Mainz U., May 2020.
URL: <https://cds.cern.ch/record/2754027>.
- [66] Vardan Khachatryan et al.
“Search for single production of scalar leptoquarks in proton-proton collisions at $\sqrt{s} = 8$ TeV”.

B. Bibliography

- In: *Phys. Rev. D* 93.3 (Feb. 2016). [Erratum: *Phys.Rev.D* 95, 039906 (2017)], p. 032005.
DOI: 10.1103/PhysRevD.93.032005.
- [67] Holger Arnold Herr.
“Search for New Heavy Charged Bosons and Singly Produced Leptoquarks with the ATLAS Experiment”.
Presented 04 Feb 2021.
PhD thesis. Mainz U., Oct. 2020.
URL: <https://cds.cern.ch/record/2752952>.
- [68] Albert M Sirunyan et al.
“Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”.
In: *Journal of High Energy Physics* 2018.201 (June 2018), p. 120.
ISSN: 1029-8479.
DOI: 10.1007/JHEP06(2018)120.
- [69] Ulrich Haisch, Luc Schnell, and Stefan Schulte.
“On Drell-Yan production of scalar leptoquarks coupling to heavy-quark flavours”.
In: *Journal of High Energy Physics* 2022.11 (Nov. 2022), p. 106.
ISSN: 1029-8479.
DOI: 10.1007/JHEP11(2022)106.
- [70] Martin Schmaltz and Yi-Ming Zhong.
“The leptoquark Hunter’s guide: large coupling”.
In: *Journal of High Energy Physics* 2019.01 (Jan. 2019).
ISSN: 1029-8479.
DOI: 10.1007/jhep01(2019)132.
- [71] D. Androić et al.
“Precision measurement of the weak charge of the proton”.
In: *Nature* 557.7704 (May 2018), pp. 207–211.
ISSN: 1476-4687.
DOI: 10.1038/s41586-018-0096-0.
- [72] C. Patrignani et al.
“Review of Particle Physics”.
In: *Chin. Phys. C* 40.10 (2016), p. 100001.
DOI: 10.1088/1674-1137/40/10/100001.
- [73] Georges Aad et al.
“Combination of searches for pair-produced leptoquarks at $\sqrt{s} = 13$ TeV with the ATLAS detector”.
In: *Physics Letters B* 854 (June 2024), p. 138736.
ISSN: 0370-2693.
DOI: 10.1016/j.physletb.2024.138736.

B. Bibliography

- [74] Luca Buonocore et al.
“Lepton-Quark Collisions at the Large Hadron Collider”.
In: *Phys. Rev. Lett.* 125 (23 Dec. 2020), p. 231804.
ISSN: 1079-7114.
DOI: 10.1103/PhysRevLett.125.231804.
- [75] Admir Greljo and Nudžeim Selimović.
“Lepton-quark fusion at Hadron colliders, precisely”.
In: *Journal of High Energy Physics* 2021.3 (Mar. 2021), p. 279.
ISSN: 1029-8479.
DOI: 10.1007/JHEP03(2021)279.
- [76] Luca Buonocore et al.
“Resonant leptoquark at NLO with POWHEG”.
In: *Journal of High Energy Physics* 2022.11 (Nov. 2022), p. 129.
ISSN: 1029-8479.
DOI: 10.1007/JHEP11(2022)129.
URL: [https://doi.org/10.1007/JHEP11\(2022\)129](https://doi.org/10.1007/JHEP11(2022)129).
- [77] A. Airapetian et al.
ATLAS detector and physics performance: Technical Design Report, 1.
Technical design report. ATLAS.
Geneva: CERN, 1999.
URL: <https://cds.cern.ch/record/391176>.
- [78] A. Airapetian et al.
ATLAS detector and physics performance: Technical Design Report, 2.
Technical design report. ATLAS.
Geneva: CERN, 1999.
URL: <https://cds.cern.ch/record/391177>.
- [79] Oliver Chris Endner.
“Search for New Physics Using Jets in Proton-Proton Collisions”.
Presented 19 May 2016.
PhD thesis. Mainz U., Dec. 2015.
URL: <https://cds.cern.ch/record/2154631>.
- [80] K. Aamodt et al.
“The ALICE experiment at the CERN LHC”.
In: *JINST* 3 (Aug. 2008), S08002.
DOI: 10.1088/1748-0221/3/08/S08002.
- [81] A. Augusto Alves Jr. et al.
“The LHCb Detector at the LHC”.
In: *JINST* 3 (Aug. 2008), S08005.
DOI: 10.1088/1748-0221/3/08/S08005.
- [82] Georges Aad et al.

B. Bibliography

- “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”.
- In: *Eur. Phys. J. C* 83.10 (Oct. 2023), p. 982.
- ISSN: 1434-6052.
- DOI: 10.1140/epjc/s10052-023-11747-w.
- [83] G Papotti.
“Luminosity analysis”.
- In: *Chamonix 2011 Workshop on LHC Performance* (2011). Specific DOI did not work and DOI to conference is given, pp. 266–269.
- DOI: 10.5170/CERN-2011-005.
- URL: <https://cds.cern.ch/record/1353869>.
- [84] Eric Torrence.
Public ATLAS Luminosity Results for Run-2 of the LHC.
Version r86.
- URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [85] Joao Pequeno.
“Computer generated image of the whole ATLAS detector”.
- 2008.
- URL: <https://cds.cern.ch/record/1095924>.
- [86] Georges Aad et al.
The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions.
Tech. rep.
Geneva: CERN, Feb. 2024.
arXiv: 2402.05858 [hep-ex].
URL: <https://cds.cern.ch/record/2888751>.
- [87] S Haywood et al.
ATLAS inner detector: Technical Design Report, 1.
Technical design report. ATLAS.
Geneva: CERN, 1997.
URL: <https://cds.cern.ch/record/331063>.
- [88] S Haywood et al.
ATLAS inner detector: Technical Design Report, 2.
Technical design report. ATLAS.
Geneva: CERN, 1997.
URL: <https://cds.cern.ch/record/331064>.
- [89] T Cornelissen et al.
Concepts, Design and Implementation of the ATLAS New Tracking (NEWT).
Tech. rep.
Geneva: CERN, 2007.
URL: <https://cds.cern.ch/record/1020106>.

B. Bibliography

- [90] Norbert Wermes and G Hallewel.
ATLAS pixel detector: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, 1998.
URL: <https://cds.cern.ch/record/381263>.
- [91] G. Aad et al.
“ATLAS pixel detector electronics and sensors”.
In: *JINST* 3 (July 2008), P07007.
DOI: 10.1088/1748-0221/3/07/P07007.
URL: <https://cds.cern.ch/record/1119279>.
- [92] Georges Aad et al.
“Operation and performance of the ATLAS semiconductor tracker”.
In: *JINST* 9 (Aug. 2014), P08009.
ISSN: 1748-0221.
DOI: 10.1088/1748-0221/9/08/P08009.
arXiv: 1404.7473 [hep-ex].
- [93] Morad Aaboud et al.
“Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties”.
In: *JINST* 12.05 (May 2017), P05002.
ISSN: 1748-0221.
DOI: 10.1088/1748-0221/12/05/P05002.
arXiv: 1702.06473 [hep-ex].
- [94] ATLAS collaboration.
ATLAS central solenoid: Technical Design Report.
Technical design report. ATLAS.
Electronic version not available.
Geneva: CERN, Apr. 1997.
DOI: 10.17181/CERN.ZZVJ.2JYE.
- [95] Joao Pequenaio.
“Computer generated image of the ATLAS inner detector”.
2008.
URL: <https://cds.cern.ch/record/1095926>.
- [96] M Capeans et al.
ATLAS Insertable B-Layer Technical Design Report.
Tech. rep.
CERN, Sept. 2010.
URL: <https://cds.cern.ch/record/1291633>.
- [97] E. Barberio et al.
“Fast simulation of electromagnetic showers in the ATLAS calorimeter: Frozen showers”.
In: *J. Phys.: Conf. Ser.* 160 (2009), p. 012082.

B. Bibliography

- DOI: 10.1088/1742-6596/160/1/012082.
URL: <http://cds.cern.ch/record/1269391>.
- [98] P. Amaral et al.
“Hadronic Shower Development in Iron-Scintillator Tile Calorimetry”.
In: *Nucl. Instrum. Methods Phys. Res., A* 443.1 (2000), pp. 51–70.
ISSN: 0168-9002.
DOI: 10.1016/S0168-9002(99)01020-7.
URL: <http://cds.cern.ch/record/386137>.
- [99] Joao Pequeno.
“Computer Generated image of the ATLAS calorimeter”.
2008.
URL: <https://cds.cern.ch/record/1095927>.
- [100] ATLAS collaboration.
ATLAS liquid-argon calorimeter: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, Dec. 1996.
DOI: 10.17181/CERN.FWRW.F00Q.
- [101] Georges Aad et al.
“Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”.
In: *The European Physical Journal C* 74.10 (Oct. 2014), p. 3071.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-014-3071-4.
- [102] ATLAS collaboration.
ATLAS tile calorimeter: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, Dec. 1996.
DOI: 10.17181/CERN.JRBJ.7028.
- [103] Georges Aad et al.
Operation and performance of the ATLAS tile calorimeter in LHC Run 2.
Tech. rep.
Geneva: CERN, 2024.
arXiv: 2401.16034.
URL: <http://cds.cern.ch/record/2887689>.
- [104] A. Airapetian et al.
ATLAS calorimeter performance: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, Dec. 1996.
URL: <https://cds.cern.ch/record/331059>.
- [105] A. Artamonov et al.
“The ATLAS Forward Calorimeter”.

B. Bibliography

- In: *JINST* 3 (2008), P02010.
DOI: 10.1088/1748-0221/3/02/P02010.
URL: <http://cds.cern.ch/record/1094547>.
- [106] ATLAS collaboration.
ATLAS muon spectrometer: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, June 1997.
URL: <http://cds.cern.ch/record/331068>.
- [107] ATLAS collaboration.
ATLAS magnet system: Technical Design Report, 1.
Technical design report. ATLAS.
Geneva: CERN, Apr. 1997.
DOI: 10.17181/CERN.905C.VDTM.
- [108] J P Badiou et al.
ATLAS barrel toroid: Technical Design Report.
Technical design report. ATLAS.
Electronic version not available.
Geneva: CERN, Apr. 1997.
DOI: 10.17181/CERN.RF2A.CP5T.
- [109] ATLAS collaboration.
ATLAS end-cap toroids: Technical Design Report.
Technical design report. ATLAS.
Electronic version not available.
Geneva: CERN, Apr. 1997.
DOI: 10.17181/CERN.P03D.WQLV.
- [110] Joao Pequenaio.
“Computer generated image of the ATLAS Muons subsystem”.
2008.
URL: <https://cds.cern.ch/record/1095929>.
- [111] ATLAS collaboration.
ATLAS level-1 trigger: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, June 1998.
URL: <http://cds.cern.ch/record/381429>.
- [112] Peter Jenni et al.
ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.
Technical design report. ATLAS.
Geneva: CERN, July 2003.
URL: <http://cds.cern.ch/record/616089>.
- [113] William Panduro Vazquez.

B. Bibliography

- “The ATLAS Data Acquisition system in LHC Run 2”.
In: *J. Phys. Conf. Ser.* 898.3 (2017). Ed. by Richard Mount and Craig Tull, p. 032017.
DOI: 10.1088/1742-6596/898/3/032017.
- [114] Morad Aaboud et al.
“Performance of the ATLAS trigger system in 2015”.
In: *The European Physical Journal C* 77.5 (May 2017), p. 317.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-017-4852-3.
- [115] Georges Aad et al.
“Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”.
In: *The European Physical Journal C* 77.7 (July 2017), p. 490.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-017-5004-5.
- [116] W Lampl et al.
Calorimeter Clustering Algorithms: Description and Performance.
Tech. rep.
Geneva: CERN, May 2008.
URL: <https://cds.cern.ch/record/1099735>.
- [117] Morad Aaboud et al.
“Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”.
In: *The European Physical Journal C* 79.8 (Aug. 2019), p. 639.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-019-7140-6.
URL: <https://doi.org/10.1140/epjc/s10052-019-7140-6>.
- [118] Thijs G. Cornelissen et al.
“The global χ^2 track fitter in ATLAS”.
In: *J. Phys. Conf. Ser.* 119 (2008). Ed. by Randall Sobie, Reda Tafirout, and Jana Thomson, p. 032013.
DOI: 10.1088/1742-6596/119/3/032013.
URL: <http://cds.cern.ch/record/1176901>.
- [119] ATLAS collaboration.
Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung.
Tech. rep.
Geneva: CERN, May 2012.
URL: <https://cds.cern.ch/record/1449796>.
- [120] Georges Aad et al.
“Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data”.

B. Bibliography

- In: *JINST* 14.12 (Dec. 2019), P12006.
DOI: 10.1088/1748-0221/14/12/P12006.
- [121] Philip Sommer et al.
Electron and Photon Selection and Identification for Run2 and Run3.
Version r186. Internal documentation, restricted access.
URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EGammaIdentificationRun2>.
- [122] Yohei Yamaguchi.
Isolation and Fake Forum: Summary of Rel. 22 studies.
Version r12. Internal documentation, restricted access.
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationFakeForumRel22>.
- [123] Georges Aad et al.
“Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”.
In: *Eur. Phys. J. C* 81.7 (July 2021), p. 578.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-021-09233-2.
- [124] Georges Aad et al.
“Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}=13$ TeV”.
In: *Eur. Phys. J. C* 76.5 (May 2016), p. 292.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-016-4120-y.
- [125] Georges Aad et al.
“Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data”.
In: *Eur. Phys. J. C* 74.11 (Nov. 2014), p. 3130.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-014-3130-x.
- [126] Georges Aad et al.
“Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”.
In: *Eur. Phys. J. C* 75.1 (Jan. 2015), p. 17.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-014-3190-y.
- [127] A. Hrynevich.
“ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2”.
In: *JINST* 12.06 (June 2017). Ed. by Lev Shekhtman, p. C06038.
DOI: 10.1088/1748-0221/12/06/C06038.

B. Bibliography

- [128] Morad Aaboud et al.
“Jet reconstruction and performance using particle flow with the ATLAS Detector”.
In: *Eur. Phys. J. C* 77.7 (July 2017), p. 466.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-017-5031-2.
- [129] S. Catani et al.
“Longitudinally-invariant $k_{\perp}$ -clustering algorithms for hadron-hadron collisions”.
In: *Nucl. Phys. B* 406.1 (1993), pp. 187–224.
ISSN: 0550-3213.
DOI: 10.1016/0550-3213(93)90166-M.
URL: <http://cds.cern.ch/record/246812>.
- [130] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez.
“The anti- k_t jet clustering algorithm”.
In: *JHEP* 04.04 (Apr. 2008), p. 063.
ISSN: 1029-8479.
DOI: 10.1088/1126-6708/2008/04/063.
arXiv: 0802.1189 [hep-ph].
- [131] CMS collaboration.
A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging.
Tech. rep.
Geneva: CERN, July 2009.
URL: <http://cds.cern.ch/record/1194489>.
- [132] ATLAS collaboration.
Tagging and suppression of pileup jets with the ATLAS detector.
Tech. rep.
Geneva: CERN, May 2014.
URL: <https://cds.cern.ch/record/1700870>.
- [133] Georges Aad et al.
“Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”.
In: *Eur. Phys. J. C* 76.11 (Oct. 2016), p. 581.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-016-4395-z.
- [134] ATLAS collaboration.
Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run.
Tech. rep.
Geneva: CERN, July 2017.
URL: <https://cds.cern.ch/record/2273281>.
- [135] ATLAS collaboration.

B. Bibliography

- Secondary vertex finding for jet flavour identification with the ATLAS detector.*
Tech. rep.
Geneva: CERN, June 2017.
URL: <https://cds.cern.ch/record/2270366>.
- [136] ATLAS collaboration.
Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC.
Tech. rep.
Geneva: CERN, 2018.
URL: <https://cds.cern.ch/record/2645405>.
- [137] Georges Aad et al.
“ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV”.
In: *Eur. Phys. J. C* 79.11 (Nov. 2019), p. 970.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-019-7450-8.
arXiv: 1907.05120 [hep-ex].
- [138] Howard S. Matis et al.
“The BRAN luminosity detectors for the LHC”.
In: *Nucl. Instrum. Meth. A* 848 (2017), pp. 114–126.
ISSN: 0168-9002.
DOI: 10.1016/j.nima.2016.12.019.
arXiv: 1612.01238 [physics.acc-ph].
- [139] Nikolaos Karastathis et al.
“Crossing Angle Anti-Leveling at the LHC in 2017”.
In: *J. Phys. Conf. Ser.* 1067.2 (Sept. 2018), p. 022004.
DOI: 10.1088/1742-6596/1067/2/022004.
- [140] Morad Aaboud et al.
“Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV”.
In: *Eur. Phys. J. C* 78.11 (Nov. 2018), p. 903.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-018-6288-9.
arXiv: 1802.08168 [hep-ex].
- [141] G. Aad et al.
“Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector”.
In: *Eur. Phys. J. C* 77.4 (Apr. 2017), p. 241.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-017-4780-2.
arXiv: 1609.09324 [hep-ex].
- [142] Steve Farrell.

B. Bibliography

- PhysicsAnalysis/AnalysisCommon/AssociationUtils* - Overlap removal tools.
Version 456b2a7a.
URL: <https://gitlab.cern.ch/atlas/athena/tree/21.2/PhysicsAnalysis/AnalysisCommon/AssociationUtils>.
- [143] D Adams et al.
Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014.
Tech. rep.
Internal documentation, restricted access.
Geneva: CERN, 2014.
URL: <https://cds.cern.ch/record/1743654>.
- [144] Michael H. Seymour and Marilyn Marx.
“Monte Carlo Event Generators”.
In:
LHC Phenomenology.
Ed. by Einan Gardi, Nigel Glover, and Aidan Robson.
Cham: Springer International Publishing, 2015,
Pp. 287–319.
ISBN: 978-3-319-05362-2.
DOI: 10.1007/978-3-319-05362-2_8.
arXiv: 1304.6677 [hep-ph].
- [145] X. Artru and G. Mennessier.
“String model and multiproduction”.
In: *Nucl. Phys. B* 70.1 (Feb. 1974), pp. 93–115.
ISSN: 0550-3213.
DOI: 10.1016/0550-3213(74)90360-5.
- [146] Bo Andersson et al.
“Parton Fragmentation and String Dynamics”.
In: *Phys. Rept.* 97.2 (July 1983), pp. 31–145.
ISSN: 0370-1573.
DOI: 10.1016/0370-1573(83)90080-7.
- [147] Geoffrey C. Fox and Stephen Wolfram.
“A Model for Parton Showers in QCD”.
In: *Nucl. Phys. B* 168.2 (June 1980), pp. 285–295.
ISSN: 0550-3213.
DOI: 10.1016/0550-3213(80)90111-X.
- [148] D. Amati and G. Veneziano.
“Preconfinement as a Property of Perturbative QCD”.
In: *Phys. Lett. B* 83.1 (Apr. 1979), pp. 87–92.
ISSN: 0370-2693.
DOI: 10.1016/0370-2693(79)90896-7.
- [149] S. Agostinelli et al.

B. Bibliography

- “GEANT4—a simulation toolkit”.
In: *Nucl. Instrum. Meth. A* 506.3 (July 2003), pp. 250–303.
ISSN: 0168-9002.
DOI: 10.1016/S0168-9002(03)01368-8.
URL: <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [150] Georges Aad et al.
Software and computing for Run 3 of the ATLAS experiment at the LHC.
Tech. rep.
Geneva: CERN, Apr. 2024.
arXiv: 2404.06335 [hep-ex].
URL: <https://cds.cern.ch/record/2895022>.
- [151] G. Duckeck et al., eds.
ATLAS Computing: technical design report.
Technical design report. ATLAS.
Geneva: CERN, June 2005.
URL: <https://cds.cern.ch/record/837738>.
- [152] K. Bos et al.
LHC computing Grid: Technical Design Report.
Technical design report. LCG.
Geneva: CERN, June 2005.
URL: <https://cds.cern.ch/record/840543>.
- [153] R. W. L. Jones.
“ATLAS computing and the GRID”.
In: *Nucl. Instrum. Meth. A* 502.2 (2003). Ed. by V. A. Ilyin, V. V. Korenkov,
and D. Perret-Gallix, pp. 372–375.
ISSN: 0168-9002.
DOI: 10.1016/S0168-9002(03)00446-7.
- [154] James Catmore et al.
“A new petabyte-scale data derivation framework for ATLAS”.
In: *J. Phys. Conf. Ser.* 664.7 (2015), p. 072007.
DOI: 10.1088/1742-6596/664/7/072007.
- [155] R. Brun and F. Rademakers.
“ROOT — An object oriented data analysis framework”.
In: *Nucl. Instrum. Meth. A* 389.1 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86.
ISSN: 0168-9002.
DOI: 10.1016/S0168-9002(97)00048-X.
- [156] Anyes Taffard, Emma Kuwertz, and Pamela Ferrari.
GoodRunListsForAnalysisRun2.
Version r179. Internal documentation, restricted access.

B. Bibliography

- URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/GoodRunListsForAnalysisRun2>.
- [157] Batool Safarzadeh Samani.
E/gamma trigger recommendations for Run 2 data.
Version r26. Internal documentation, restricted access.
URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/TrigEgammaRecommendations2015>.
- [158] Stefano Frixione, Paolo Nason, and Carlo Oleari.
“Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”.
In: *JHEP* 2007 (Nov. 2007), p. 070.
ISSN: 1029-8479.
DOI: 10.1088/1126-6708/2007/11/070.
arXiv: 0709.2092 [hep-ph].
- [159] Simone Alioli et al.
“A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”.
In: *JHEP* 06 (June 2010), p. 043.
ISSN: 1029-8479.
DOI: 10.1007/JHEP06(2010)043.
arXiv: 1002.2581 [hep-ph].
- [160] Paolo Nason.
“A New method for combining NLO QCD with shower Monte Carlo algorithms”.
In: *JHEP* 2004.11 (Dec. 2004), p. 040.
ISSN: 1029-8479.
DOI: 10.1088/1126-6708/2004/11/040.
arXiv: hep-ph/0409146.
- [161] Johannes Bellm et al.
“Herwig 7.0/Herwig++ 3.0 release note”.
In: *Eur. Phys. J. C* 76.4 (Apr. 2016), p. 196.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-016-4018-8.
arXiv: 1512.01178 [hep-ph].
- [162] J. Alwall et al.
“The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”.
In: *JHEP* 2014.7 (July 2014), p. 079.
ISSN: 1029-8479.
DOI: 10.1007/JHEP07(2014)079.
arXiv: 1405.0301 [hep-ph].
- [163] Torbjörn Sjöstrand.

B. Bibliography

- “The PYTHIA Event Generator: Past, Present and Future”.
In: *Comput. Phys. Commun.* 246 (Jan. 2020), p. 106910.
ISSN: 0010-4655.
DOI: 10.1016/j.cpc.2019.106910.
arXiv: 1907.09874 [hep-ph].
- [164] Torbjörn Sjöstrand et al.
“An introduction to PYTHIA 8.2”.
In: *Comput. Phys. Commun.* 191 (June 2015), pp. 159–177.
ISSN: 0010-4655.
DOI: 10.1016/j.cpc.2015.01.024.
arXiv: 1410.3012 [hep-ph].
- [165] Simone Amoroso et al.
PMG references document.
Tech. rep.
Internal documentation, restricted access.
Geneva: CERN, 2019.
URL: <https://cds.cern.ch/record/2678867>.
- [166] Peter E. Freeman, Stephen Doe, and Aneta Siemiginowska.
“Sherpa: a mission-independent data analysis application”.
In: *Astronomical Data Analysis*.
Ed. by Jean-Luc Starck and Fionn D. Murtagh.
Vol. 4477.
International Society for Optics and Photonics. SPIE, Nov. 2001,
Pp. 76–87.
DOI: 10.1117/12.447161.
- [167] Enrico Bothmann et al.
“Event Generation with Sherpa 2.2”.
In: *SciPost Phys.* 7.3 (Sept. 2019), p. 034.
ISSN: 2542-4653.
DOI: 10.21468/SciPostPhys.7.3.034.
arXiv: 1905.09127 [hep-ph].
- [168] Richard D. Ball et al.
“Parton distributions with LHC data”.
In: *Nucl. Phys. B* 867.2 (Feb. 2013), pp. 244–289.
ISSN: 0550-3213.
DOI: 10.1016/j.nuclphysb.2012.10.003.
arXiv: 1207.1303 [hep-ph].
- [169] G. Aad et al.
“Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC”.
In: *JINST* 18.11 (2023), T11004.
DOI: 10.1088/1748-0221/18/11/T11004.

B. Bibliography

- arXiv: 2211.16178 [hep-ex].
- [170] Michael Holzbock et al.
LJNtupleMaker.
Version 7129ac12. Internal documentation, restricted access.
URL: <https://gitlab.cern.ch/atlas-phys/exot/jdm/ana-exot-2021-30/LJNtupleMaker>.
- [171] Danilo Piparo et al.
“SWAN: A service for interactive analysis in the cloud”.
In: *Future Generation Computer Systems* 78 (2018), pp. 1071–1078.
ISSN: 0167-739X.
DOI: <https://doi.org/10.1016/j.future.2016.11.035>.
- [172] Guido Van Rossum and Fred L. Drake.
Python 3 Reference Manual.
Scotts Valley, CA: CreateSpace, 2009.
ISBN: 1441412697.
- [173] CERN.
The LXPLUS Service.
Version EL9. Documentation of LXPLUS.
URL: <https://lxplusdoc.web.cern.ch/>.
- [174] CERN.
CERN Batch Service User Guide.
Documentation of HTCondor.
URL: <https://batchdocs.web.cern.ch/index.html>.
- [175] Michael Holzbock et al.
LJHistoMaker.
Version 1c767ffc. Internal documentation, restricted access.
URL: <https://gitlab.cern.ch/atlas-phys/exot/jdm/ana-exot-2021-30/LJHistoMaker>.
- [176] CERN.
ATLAS Analysis Facility at UChicago.
About page of UChicago.
URL: <https://af.uchicago.edu/about>.
- [177] Michael Holzbock et al.
LJStatAnalysis.
Version 408948a9. Internal documentation, restricted access.
URL: <https://gitlab.cern.ch/atlas-phys/exot/jdm/ana-exot-2021-30/LJStatAnalysis>.
- [178] Kyle Cranmer and Alexander Held.
“Building and steering binned template fits with cabinetry”.
In: *EPJ Web Conf.* 251 (Aug. 2021), p. 03067.
DOI: 10.1051/epjconf/202125103067.

B. Bibliography

- [179] Alexander Held et al.
Documentation of Cabinetry.
URL: <https://cabinetry.readthedocs.io/en/latest/>.
- [180] Alexander Held et al.
Repository of Cabinetry.
URL: <https://github.com/scikit-hep/cabinetry>.
- [181] Lukas Heinrich, Matthew Feickert, and Giordon Stark.
pyhf: v0.7.6.
Version 0.7.6. <https://github.com/scikit-hep/pyhf/releases/tag/v0.7.6>.
DOI: 10.5281/zenodo.1169739.
- [182] Lukas Heinrich et al.
“pyhf: pure-Python implementation of HistFactory statistical models”.
In: *Journal of Open Source Software* 6.58 (2021), p. 2823.
DOI: 10.21105/joss.02823.
- [183] Sandrine Laplace.
EGamma recommendations for Run2 release 21.
Version r56. Internal documentation, restricted access.
URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EGammaRecommendationsR21>.
- [184] Stefano Zambito.
MCPAnalysisGuidelinesMC16.
Version r68. Internal documentation, restricted access.
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC16>.
- [185] Georges Aad et al.
“Performance of electron and photon triggers in ATLAS during LHC Run 2”.
In: *Eur. Phys. J. C* 80.1 (Jan. 2020), p. 47.
ISSN: 1434-6052.
DOI: 10.1140/epjc/s10052-019-7500-2.
arXiv: 1909.00761 [hep-ex].
- [186] William Buttinger and Michel Lefebvre.
Formulae for Estimating Significance.
Tech. rep.
Internal documentation, restricted access.
Geneva: CERN, 2018.
URL: <https://cds.cern.ch/record/2643488>.
- [187] Glen Cowan et al.
“Asymptotic formulae for likelihood-based tests of new physics”.
In: *Eur. Phys. J. C* 71.2 (Feb. 2011). [Erratum: *Eur.Phys.J.C* 73, 2501 (2013)],
p. 1554.
ISSN: 1434-6052.

B. Bibliography

- DOI: 10.1140/epjc/s10052-011-1554-0.
arXiv: 1007.1727 [physics.data-an].
- [188] G. Aad et al.
Expected performance of the ATLAS experiment: detector, trigger and physics.
Geneva: CERN, Jan. 2009.
URL: <https://cds.cern.ch/record/1125884>.
- [189] ATLAS collaboration, CMS collaboration, and The LHC Higgs Combination Group.
Procedure for the LHC Higgs boson search combination in Summer 2011.
Tech. rep.
Geneva: CERN, Aug. 2011.
URL: <https://cds.cern.ch/record/1379837>.
- [190] Abraham Wald.
“Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large”.
In: *Transactions of the American Mathematical Society* 54.3 (Nov. 1943), pp. 426–482.
ISSN: 00029947, 10886850.
eprint: <https://www.ams.org/journals/tran/1943-054-03/S0002-9947-1943-0012401-3/S0002-9947-1943-0012401-3.pdf>.
URL: <http://www.jstor.org/stable/1990256>.
- [191] Alexander L. Read.
“Presentation of search results: The CL_s technique”.
In: *J. Phys. G* 28.10 (Sept. 2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704.
DOI: 10.1088/0954-3899/28/10/313.
- [192] Eric Goubault.
“Static Analyses of the Precision of Floating-Point Operations”.
In: *Static Analysis*.
Ed. by Patrick Cousot.
Berlin, Heidelberg: Springer Berlin Heidelberg, 2001,
Pp. 234–259.