

Dipartimento di Fisica e Astronomia “Augusto Righi”
Corso di Laurea in Fisica

Optimization of c-jet selection in Z+jets production at the ATLAS experiment

Relatore:
Prof.ssa Laura Fabbri

Presentata da:
Alberto Tanzi

Anno Accademico 2024/2025

Abstract

La fisica delle particelle si occupa dei costituenti fondamentali della materia e delle loro proprietà, definendo anche le interazioni fondamentali che governano la materia. Il Large Hadron Collider (LHC), installato presso il CERN di Ginevra, è l'acceleratore di particelle più potente al momento disponibile, in grado di far collidere protoni e ioni pesanti al fine di produrre enormi quantità di dati per la ricerca. Questi dati sono resi disponibili attraverso diversi rivelatori di particelle posti lungo la circonferenza di LHC, uno di questi, il più grande, è ATLAS. ATLAS è un rivelatore per la scoperta di nuova fisica, attraverso l'osservazione di possibili nuove particelle o di incongruenze dei modelli teorici al momento più accreditati. Sebbene fino ad oggi gran parte dei risultati degli esperimenti nel campo della fisica delle particelle siano in accordo con le predizioni del Modello Standard della fisica delle particelle, sono ancora numerose le domande a cui la teoria non è in grado di dare risposta, ragione per cui nei prossimi anni esperimenti come ATLAS potranno darci risultati sorprendenti. L'esperimento ATLAS ambisce a fornire nuove verifiche del Modello Standard a scale energetiche mai esplorate prima, a cercare fenomeni che sono al di là degli attuali quadri teorici, oltre che a studiare il bosone di Higgs, della cui scoperta ATLAS è stato uno dei responsabili. Questa tesi si occupa di misure della produzione di bosoni Z (che decadono successivamente in leptoni) in associazione con jet generati da c-quark. L'analisi utilizza previsioni fornite da generatori di eventi Monte Carlo il cui ruolo è stato di simulare l'esito di collisioni protone-protone all'energia del centro di massa di $\sqrt{s} = 13.6$ TeV, con luminosità integrata di 29 fb^{-1} . Il processo che regola il fenomeno d'interesse è governato dall'interazione forte che ha luogo tra i costituenti dei protoni, ovvero quark e gluoni, dei quali i primi sono responsabili della formazione dei jet attraverso l'adronizzazione. Per il decadimento leptonic del bosone Z sono presi in considerazione il canale elettronico e il canale muonico, ovvero i rispettivi casi in cui il bosone Z decade in elettroni e muoni. Questo tipo di studio è fondamentale per verificare il modello della Cromodinamica Quantistica in regime perturbativo, nonché per descrivere come la quantità di moto è distribuita tra le componenti interne dei protoni. Inoltre, poiché il processo studiato è uno dei fondi dominanti per il decadimento del bosone di Higgs, tale analisi è fondamentale per approfondire la conoscenza delle costanti di accoppiamento di quest'ultimo. Nello specifico questa tesi si occupa di studiare il rapporto tra previsioni di segnale e misure di fondo dell'energia trasversa mancante (MET) attraverso grandezze derivanti da suddette quantità, ed è volta a valutare come ottimizzare il taglio di variabile sulla MET. Nel capitolo 1 si presenta una breve descrizione del Modello Standard della fisica delle particelle, con particolare attenzione a jet e bosone Z, insieme allo stato dell'arte delle misure di decadimenti coinvolgenti tale bosone. Il capitolo 2 si focalizza sul complesso di LHC e sul rivelatore ATLAS, concludendosi con una sezione dedicata al come avviene la ricostruzione del segnale lasciato dalle particelle di interesse e dai jet. Infine, nel capitolo 3 sono riportati i dati e i metodi di analisi applicati, mostrando come non sia necessario operare alcun taglio di variabile sull'energia trasversa mancante al fine di ottimizzare il rapporto tra segnale e fondo. Seguono poi le conclusioni in cui viene brevemente esposto l'esito dell'analisi svolta.

Abstract

Particle physics is the field of physics that deals with the fundamental constituents of matter and with their properties, defining the fundamental interactions between the components of matter as well. The Large Hadron Collider (LHC), located at CERN in Geneva, is a particle accelerator capable of colliding protons and heavy ions with the aim of producing large amount of data for research. Said data is gathered through different particle detectors placed along LHC's circumference, ATLAS being one of said detectors and acting as a fundamental component in the validation of theoretical models and in pushing the boundaries of scientific research. Although most of the results of particle physics experiments performed until now have shown agreement with the Standard Model of particle physics' predictions, many are the questions that the current theoretical frameworks cannot provide an answer for, which is the reason as to why experiments such as ATLAS still play a fundamental role in research. The ATLAS experiment aims at exploring new energy scales to challenge the Standard Model, and at searching for new phenomena that are not accounted for in the current theories, other than at studying the Higgs boson, which was discovered through ATLAS and the CMS detector. This thesis focuses on measurements of the production of leptonically decaying Z bosons in association with jets originating from c-quarks. The analysis was performed using predictions provided by Monte Carlo event generators, through which it was possible to simulate the outcome of measurements of proton-proton collisions at center of mass energy of $\sqrt{s} = 13.6$ TeV, with an integrated luminosity of 29 fb^{-1} . The process behind the decay in question is dominated by the strong interaction that takes place between the protons' constituents, namely quarks and gluons, of which the former are also responsible for the formation of jets, through the phenomenon of hadronization. As for the lepton decay, electron and muon channels are taken into account. This type of study is fundamental to probe perturbative quantum chromodynamics, other than to describe how momentum is distributed between the components of the protons. Furthermore, since the process under study is one the dominant backgrounds for the Higgs boson decay, this analysis is key to deepen the knowledge of said boson's coupling constants. This thesis focuses on the study of the ratio between signal measurements and background measurements involving missing transverse energy (MET), and it is aimed at evaluating how to optimize the variable cut on MET. In Chapter 1 a brief description of the Standard Model is presented, with a focus on jets and the physics of the Z boson, as well as the state of the art of Z boson measurements. Chapter 2 focuses on the LHC complex and on the ATLAS detector, ending with a section explaining how signal reconstruction is performed over jets, electrons and muons. Finally, Chapter 3 reports the elaborated plots and the analysis technique, pointing out how it is not necessary to operate a cut on missing transverse energy in order to optimize the signal-to-background ratio. In the Conclusion the outcome of the analysis is briefly illustrated.

Contents

1	Theoretical Overview	1
1.1	The Standard Model of Particle Physics	1
1.1.1	Running couplings	4
1.1.2	The Higgs boson and mechanism	5
1.2	Proton-Proton Interactions at LHC	6
1.2.1	Monte Carlo Simulations	8
1.3	The Z Boson	8
1.3.1	Discovery and Properties	9
1.3.2	Z Boson in association with jets of heavy quarks	10
1.3.3	State of the art of Z+jets measurements	13
2	The ATLAS Apparatus at LHC	16
2.1	The Large Hadron Collider	16
2.1.1	The detectors at LHC	19
2.1.2	Luminosity	19
2.2	The ATLAS Detector	21
2.2.1	The ATLAS reference frame	22
2.2.2	The inner detector	24
2.2.3	System of magnets	26
2.2.4	The calorimetry system	27
2.2.5	Muon Spectrometer	29
2.2.6	Trigger and DCS	30
2.3	Object Reconstruction and Identification	30
2.3.1	Electrons	31
2.3.2	Muons	32
2.3.3	Jets	33
2.3.4	Flavor Tagging	35
2.3.5	Missing transverse energy (MET)	36

3	Data Analysis	38
3.1	Event Generators	38
3.2	Signal-to-Background optimization	40
3.2.1	ROC Curves	44
	Conclusion	51

Chapter 1

Theoretical Overview

The Standard Model of Particle Physics (SM) is currently the theory that best describes the nature of fundamental particles and forces. Throughout history, several experiments have attempted to identify flaws or limitations within this theory. Despite there being phenomena that the SM cannot fully explain yet, it remains the most tested and validated theory to date. In this chapter, the SM is discussed in detail in Section 1.1. In Section 1.2 the peculiar theoretical aspects of proton-proton (pp) collisions are addressed, and Section 1.3 delves into the current understanding of the production of the Z boson in association with b-jets in pp collisions.

1.1 The Standard Model of Particle Physics

The SM is a gauge quantum field theory that describes three out of the four fundamental interactions currently known: electromagnetic, weak, and strong. A way of including gravity in the framework of the SM has yet to be found, however, due to the weakness of gravity at the energetic scales of particle physics, the gravitational interaction is negligible at such scales (Tab.1.1).

Strong	Electromagnetic	Weak	Gravitational
1	10^{-2}	10^{-7}	10^{-39}

Table 1.1: Order of magnitude of the relative strenghts of the fundamental interactions[1].

Many high precision experiments confirmed the validity of the SM framework and its effectiveness in accurately describing all observed particles. The last milestone of the SM was the discovery of the Higgs boson at LHC in 2012, a particle which was predicted by the theory many years before. The SM describes and organizes the fundamental particles (Fig.1.1), dividing them into fermions, which obey to

Fermi-Dirac statistics and have a semi-integer spin, and bosons, which obey to Bose-Einstein statistics and have either integer or zero spin. The model introduces the existence of matter and antimatter as well, with the latter being composed by particles whose quantum numbers have all opposite sign with respect to the particles that make up matter. The confirmation of the existence of antimatter came in 1932 with the discovery of positrons, which are electrons' antiparticles. However there are also phenomena that cannot be explained by this theory, such as the unification of all forces at very high energy including gravity, the nature of Dark Matter, and the asymmetry between particles and antiparticles. Due to these limitations, various theories beyond the SM [2] have been proposed, in which the SM maintains the role of theory valid at low energies. However these new theories have yet to be confirmed.

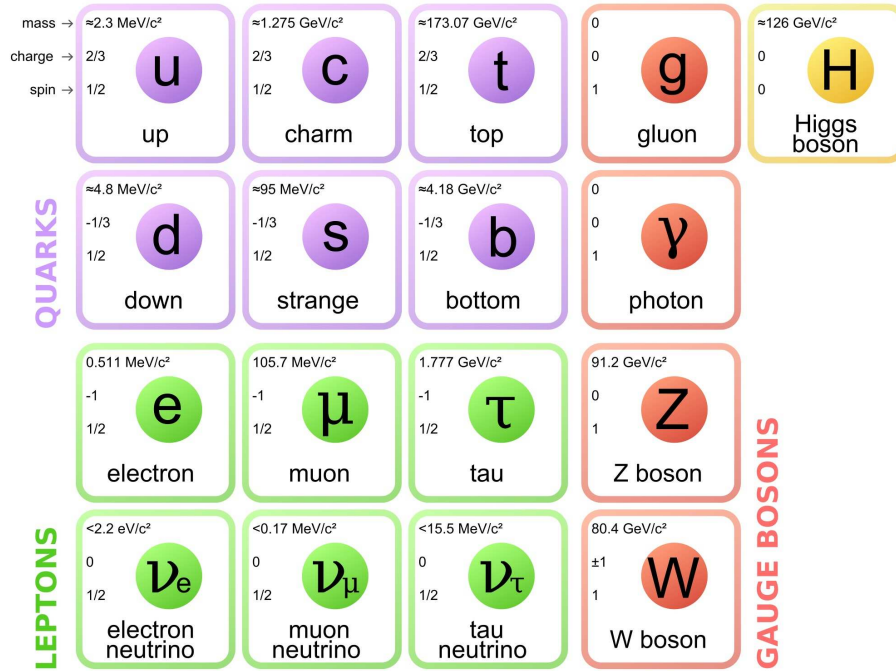


Figure 1.1: Particles predicted by the SM. In the green are leptons, which, along with quarks in the blue, are particles of matter. In the red and yellow are the bosons, which mediate interactions, the red boxes represent vector bosons and the yellow ones represent the scalar Higgs boson. The mass, charge and spin of the particles are on the top left of each box, in descending order.

Fermions

Fermions are particles with half-integer spin, such as leptons and quarks. There are three families of leptons, each differing in mass, as illustrated in Fig.1.1. The first family is composed of the electron (e) and the electron neutrino (ν_e), the second family consists of the muon (μ) and the muon neutrino (ν_μ), the third and last family consists of the tau (τ) and the tau neutrino (ν_τ). Neutrinos were initially predicted

to be massless by the SM, but experimental evidences, such as neutrino oscillations, indicate that their mass, although small, is not equal to zero [3]. The mass of other leptons increases moving from the first up to the third family. Electrons, muons and taus have the same electric charge e , while neutrinos of all families are chargeless. Leptons can be observed as free particles. Quarks are fundamental particles with fractional electric charges of either $-\frac{1}{3}e$ or $\frac{2}{3}e$. They are divided into three families, each containing two quarks, as shown in Fig.1.1: up (u) and down (d), charm (c) and strange (s), top (t) and bottom (b). Regarding quarks, as for leptons, those in the first generation are lighter, while those in the last generation are heavier. In addition to electric charge, quarks also carry color charge, which can be in one of three states: red, blue, or green. Due to the phenomenon of color confinement in strong interactions, quarks cannot be observed as free particles but only in bound states known as hadrons. Hadrons must therefore be color-neutral and can be baryons or mesons. Baryons consist of three quarks bound together, while mesons are composed of a quark-antiquark pair. Protons and neutrons are examples of baryons and are made up of the quark combinations (uud) and (udd), respectively. Each fermion has an associated antiparticle, known as an antifermion, which has the same mass as its corresponding fermion but opposite charge. For instance, the positron (e^+) is the antiparticle of the electron (e^-), is positively charged and has a leptonic number of -1 [4].

Bosons

Bosons are particles with integer spin that act as mediators of the fundamental forces. The bosons predicted by the SM are: the photon, 8 gluons, the W^+ and W^- bosons, the Z boson, and the Higgs boson. The photon is the mediator of the electromagnetic interaction, which affects electrically charged particles. The gluon is the mediator of the strong interaction, binding quarks together within hadrons. Gluons carry color charges and interact by exchanging this charge with quarks, thereby participating in strong interactions directly. The W and Z bosons mediate the weak interaction, which is responsible for processes like beta decay in atomic nuclei. The W bosons exist with both positive and negative electric charge (W^+ and W^- , respectively), while the Z boson is electrically neutral. The W and Z bosons are massive, in contrast to photons and gluons, which are massless. The Higgs boson is a fundamental particle whose discovery confirmed the Higgs mechanism, through which elementary particles acquire their mass [4].

All the interactions in the SM are modelled, in analogy with classical electromagnetism, as excitations of a field diffused through space, that in the case of the standard model is quantized. Each field's quanta are the aforementioned bosons, the particles through the exchange of which interactions take place.

The SM forms a group of symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, and interactions are described requiring gauge invariance (both local and global) under said groups. $SU(3)_C$ accounts for the color symmetry of quarks and describes strong interactions. The required invariance under this group introduces eight different fields, related to the eight gluons described by Quantum Chromodynamics (QCD). $SU(2)_L$ and $U(1)_Y$ account for the weak isospin and hypercharge symmetry and describe weak and electromagnetic interactions[5].

1.1.1 Running couplings

A fundamental parameter that determines the strength of the particle interactions is the coupling constant. The couplings, both electromagnetic and strong, are not constant values but rather they vary with energy [6]. In QED, the behaviour of the coupling depends on the energy of the probe. The QED coupling running is expressed as:

$$\alpha_{em}(Q^2) = \frac{\alpha_{em}(\mu^2)}{1 - \frac{\alpha_{em}(\mu^2)}{3\pi} \ln(\frac{Q^2}{\mu^2})} \quad (1.1)$$

where Q^2 is the energy of the probe, μ^2 is the reference scale, often taken as $\mu^2 = m_e^2$, where m_e is the mass of the electron. The equation shows how the intensity of the interaction depends on the energy of the probe: as the energy increases, the coupling strength increases and vice versa. The larger the coupling, the more intense will be the interaction. Indeed, as Q^2 increases, the denominator decreases and $\alpha(Q^2)$ increases. The couplings, both electromagnetic and strong, are also expressed as "e" and " g_s ", respectively. In natural units, the electromagnetic coupling is also expressed as $\alpha = \frac{e^2}{4\pi}$, therefore e and α are proportional.

The QCD coupling running is expressed as:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi} (11n_C - 2n_F) \ln(\frac{Q^2}{\mu^2})} \quad (1.2)$$

where n_C is the number of colors, n_F is the number of flavors and $\alpha_s = \frac{g_s^2}{4\pi}$. The different sign in the denominator leads to the opposite behavior. In this case in-

creasing the energy of the probe, i.e. probing at short distances, causes the particles involved to behave as free particles: this phenomenon is known as *asymptotic freedom*. At lower energies and large distances, the coupling becomes stronger, leading to the phenomenon of *confinement*, which explains why colored particles cannot be observed directly [4]. In Fig.1.2 are illustrated the behaviors of the QED and QCD couplings as a function of the energy of the probe.

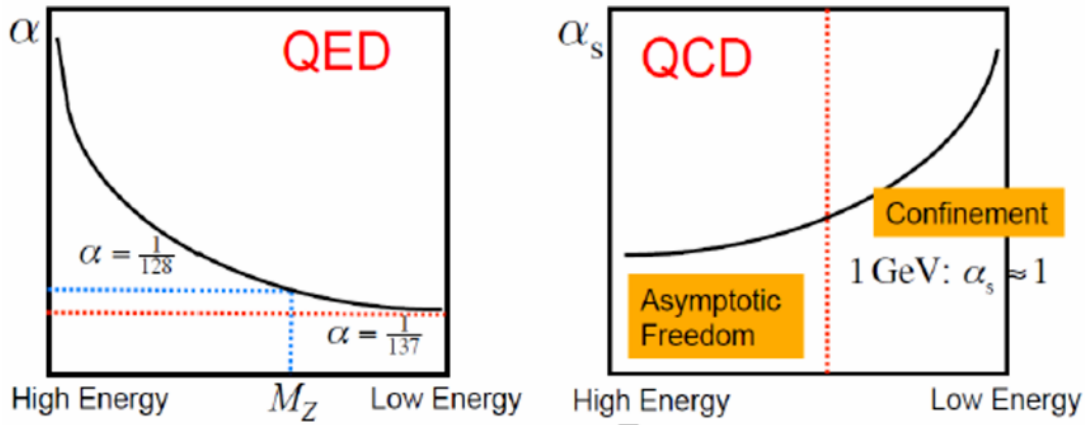


Figure 1.2: Running of the coupling constants as functions of Q^2 , the QED coupling is represented on the left, while the QCD coupling parameter's behavior is represented on the right[4].

1.1.2 The Higgs boson and mechanism

In the 1970's, Glashow, Weinberg, Salam, t'Hooft et al., came up with a theory aimed at unifying the electromagnetic and weak interaction. That theory is now a fundamental part of the SM and is commonly referred to as *electroweak theory*. The main problem that had to be dealt with during the developement of the theory was caused by the fact that on one hand weak interactions demanded for massive quanta (due to their short range) to be carried by, while on the other hand, electromagnetic quantum theories required massless mediators (due to their infinite range). The solution to that problem was brought forth by a group of researchers that included P. Higgs, through the introduction of the Higgs mechanism [7], which predicted the existence of an addicitional boson, the Higgs boson, which brings along a condensed field, the Higgs field, which can interact with the weak interaction's mediators, granting them the property of mass. Hypothesized for the first time in 1964, the Higgs boson was finally observed in 2012 at CERN by the ATLAS and CMS experiments held at the LHC [5]. The potential of the Higgs field is expressed by the following equation:

$$V(\phi) = \frac{1}{2}\mu^2(\phi\phi^*)^2 + \frac{1}{4}\lambda(\phi\phi^*)^4 \quad (1.3)$$

Where ϕ is the Higgs field, λ an adimensional parameter that expresses the intensity of the interaction, and μ can be interpreted as a mass term. As shown in Fig.1.3, the potential for $\mu^2 < 0$ presents two minimum values for $\phi = \pm\sqrt{-\frac{\mu^2}{\lambda}} \equiv \nu$, said values are the vacuum expectation values. According to this potential, the Higgs boson itself presents a mass [5]:

$$m_H^2 = -2\mu^2 = 2\lambda\nu^2 \quad (1.4)$$

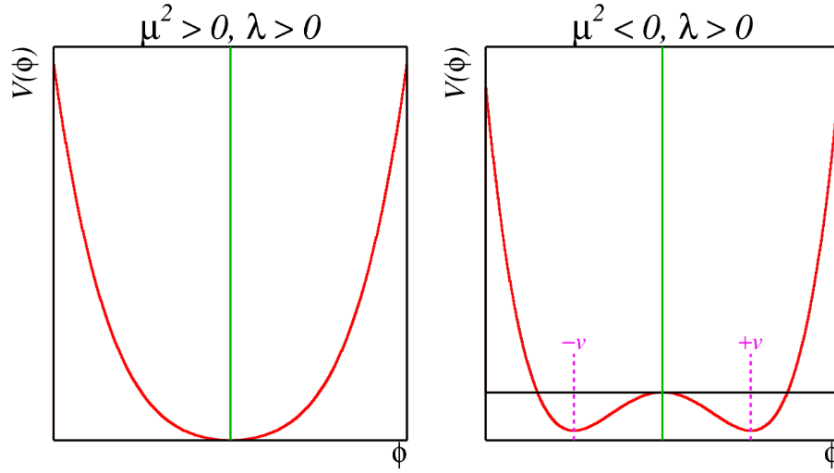


Figure 1.3: Potential $V(\phi)$ of the Higgs field ϕ for $\mu^2 > 0$ on the left and $\mu^2 < 0$ on the right. In the graph on the right it is apparent that the minimum values of $V(\phi)$, corresponding to the vacuum expectation values, are where $\phi \neq 0$.

1.2 Proton-Proton Interactions at LHC

Proton-proton (pp) collisions are crucial experiments in high-energy physics in which "bunches" of protons are accelerated to extremely high velocities and made to collide. When two protons collide, their constituents interact in a variety of ways: the scattering can be elastic, when the particles in the final state are the same as those in the initial state, or inelastic. The interactions analyzed in this thesis involve inelastic scattering. Most of the inelastic interactions are "soft", characterized by low momentum transfer and non-perturbative effects, while hard interactions, described by perturbative Quantum Chromodynamics (pQCD), involve high momentum transfer. Therefore, in the study of pp collisions at high energy, both perturbative and non-perturbative QCD processes are involved. The process through which the inner

constituents of the protons (quarks and gluons) begin to interact directly is referred to as deep-inelastic scattering. In these interactions, pQCD allows for the calculation of cross-sections at fixed order using Parton Distribution Functions (PDFs) and squared matrix elements (MEs). While the interaction takes place, soft and collinear partons can be emitted from the initial or final states through the phenomenon of parton showering (PS). Parton showers lead the system from the hard scattering scale to the hadronization scale by stimulating successive emissions resulting in a cascade of lower-energy partons, which can be modeled through a series of splittings. Due to confinement, quarks and gluons cannot exist in a state that is not bound with another quark in order to make a composed neutral-colored particle, which is why quarks and gluons in the final state recombine to form colorless hadrons through a process called hadronization.

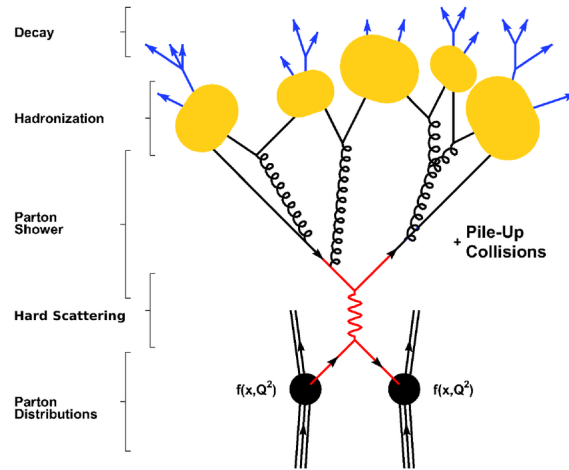


Figure 1.4: Schematization of a pp collision, where $f(x, Q^2)$ indicate the parton distribution functions.

The hadrons produced then tend to move in collimated directions, reflecting the momentum and the direction of the partons that the hadronization process started from, said collimated groups of hadrons are referred to as "jets". In the study of pp collisions, additional complexities can arise due to pileup events, namely, soft collisions resulting from other parton-parton interactions occurring simultaneously. Fig.1.4 shows a schematic representation of a collision between two protons. The two objects at the bottom represent the protons, the internal structure of which is described by parton distribution functions (PDFs). During the collision, two partons interact in a hard scattering process. The remaining partons within the protons contribute to what is known as the *underlying event*. The elementary particles produced in the hard scattering process constitute the parton level, and, through the process of fragmentation, they transition from the parton level to the particle level, where quarks and gluons are combined into hadrons [4].

1.2.1 Monte Carlo Simulations

In the context of this work, hadronization actually refers to the specific model used by a Monte Carlo generator to describe the transition from partonic to final state. The two most recent models are based on an isotropic and longitudinal phase space: the “string” and the “cluster” model, depicted in Fig.1.5. The string model is based on the hypothesis that quarks correspond to the final points of a string and gluons are deviations from it. Along the string, partons are ordered according to their color. This model provides a predictive picture of the primary hadron movement in space-time. This structure can be applied to multi-hadron configurations. A possible disadvantage is the presence of many parameters linked to quark flavours, which must be tuned with data. The cluster model is based on the confinement features of the parton showering, ending with the reconstruction of parton clusters with low mass values; it starts with a non perturbative splitting of gluons in $q\bar{q}$ pairs, which then originate clusters. Although the cluster model is less precise compared to the string model, it provides a complete description of events at high energy when combined with PS, by using a limited number of parameters [8].

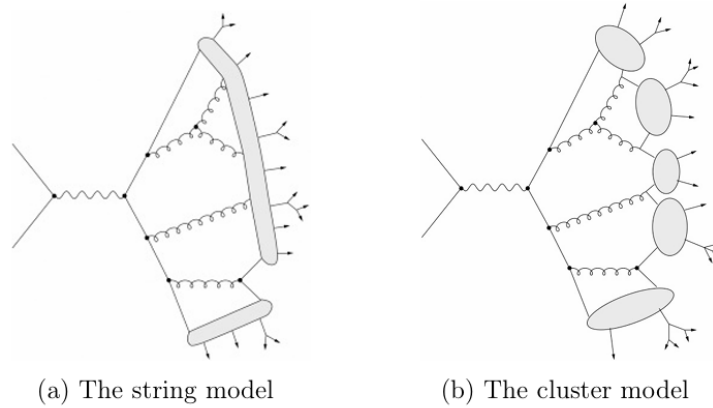


Figure 1.5: (a) The string hadronization model directly transforms partonic systems in hadrons. (b) The cluster model uses clusters as intermediate steps, with a typical mass of a few GeV.

1.3 The Z Boson

Since its discovery in 1983 by the UA1 experiment [9], the Z boson has been and still is an object of interest in the high energy physics field. In this Section, the Z boson’s fundamental properties are described, focusing on the reason why, after more than 30 years from its discovery, it is still an important ingredient in the particle physics research.

1.3.1 Discovery and Properties

The Z boson was discovered in 1983 by the UA1 collaboration at the $S\bar{p}pS$ collider at CERN [9]. UA1 the primary goal was the search for the massive bosons involved in the electroweak interaction. After the discovery of the W bosons in events with single isolated electrons and missing energy, the UA2 detector observed eight events interpreted in terms of the reaction:

$$\bar{p}p \rightarrow Z + \text{anything} \quad \text{with} \quad Z \rightarrow e^+e^- \quad (1.5)$$

in a data sample corresponding to a total integrated luminosity of 131 nb^{-1} . From those events, the UA2 collaboration measured the mass of the Z boson as:

$$M_Z = (91.9 \pm 1.3 \pm 1.4) \text{ GeV} \quad (1.6)$$

where the first error accounts for the analysis strategy and the second one is the systematic error.

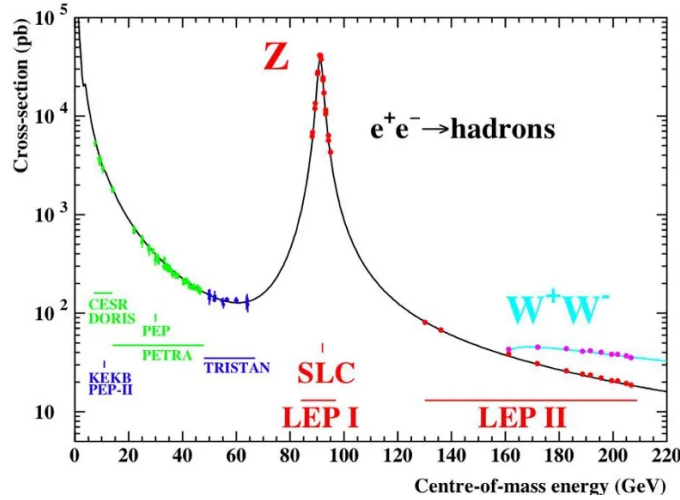


Figure 1.6: Measurements of the e^+e^- cross section, the red data account for the Z boson. It's worth noting that the Z boson Cross-Section follows the Breit-Wigner distribution, the peak corresponding to the mass of the particle [10].

The properties of the Z boson were then studied in great detail at LEP, an electron-antielectron collider at CERN until the year 2000, when the LHC took the place of LEP (Fig.1.6). During the first phase (LEP1) the centre-of-mass energy was kept in an interval of $\pm 3 \text{ GeV}$ around m_Z ; this allowed experiments to collect a total amount of 18 million Z decays from 1989 to 1995 and to perform precision measurements [10, 11]. The Z boson mass (m_Z), the total width (Γ_Z) and the partial widths into fermions (leptons) ($\Gamma_{f\bar{f}}$), have been determined from an analysis of the production cross sections of fermionic final states coming from e^+e^- colli-

sions. The shape of the cross section variation around the Z peak is parametrised by a Breit-Wigner function, with an energy-dependent total width (Fig.1.6). The determination of the aforementioned parameters was performed through an analytic expression of this cross section in terms of the parameters themselves and by fitting the calculated cross section to the measured one, varying the parameters involved.

$$\sigma(s) = \sigma_{f\bar{f}}^0 \frac{s\Gamma_Z^2}{(s - m_Z^2)^2 + \frac{s^2}{m_Z^2}\Gamma_Z^2} \quad (1.7)$$

Where $\sigma_{f\bar{f}}^0$ represents the cross section for the process $e^+e^- \rightarrow f\bar{f}$ at $\sqrt{s} = m_Z$. If the final state involves the fermionic e^+e^- couple, the above cross section must be integrated to take into account small contributions from photons exchange and γ -Z interference. The world average value of the Z boson mass and total width are [12]:

$$m_Z = (91.1876 \pm 0.0021)\text{GeV} \quad \Gamma_Z = (2.4952 \pm 0.0023)\text{GeV} \quad (1.8)$$

The cross section can be expressed as a function of the partial widths of the Z possible decays $\Gamma_{f\bar{f}}$:

$$\sigma_{f\bar{f}}^0 = \frac{12\pi}{m_Z^2} \frac{\Gamma_{e^+e^-}^2 \Gamma_{f\bar{f}}}{\Gamma_Z^2} \quad (1.9)$$

In Tab.1.2 the Z partial widths and the respective branching ratios are reported for each decay channel [12].

Decay Channel	Partial width ($\Gamma_{f\bar{f}}$) (MeV)	Branching ratio ($\Gamma_{f\bar{f}}/\Gamma_Z$) (%)
e^+e^-	89.91 ± 0.12	3.363 ± 0.004
$\mu^+\mu^-$	83.99 ± 0.18	3.366 ± 0.007
$\tau^+\tau^-$	84.08 ± 0.22	3.370 ± 0.008
invisible	499.0 ± 1.5	20.00 ± 0.06
hadrons	1744 ± 2.0	69.91 ± 0.06

Table 1.2: Z bosons partial widths and branching ratios for each decay mode.

1.3.2 Z Boson in association with jets of heavy quarks

At LEP there was no way to study the Z boson in association with jets, since the initial state was made up of only electrons and positrons. At the LHC, where the initial state is made up by gluons and quarks (the internal components of protons),

it is possible to study in great detail processes such as the production of a Z boson in association with jets by exploiting the large centre-of-mass energy of pp collisions. The production of a Z boson in association with b- or c-jets is fundamental in studies that involve the Higgs boson coupling, making up most of the background for decays such as the one depicted in Fig.1.10. This thesis mainly focuses on the analysis of the Z boson in association with c-jets.

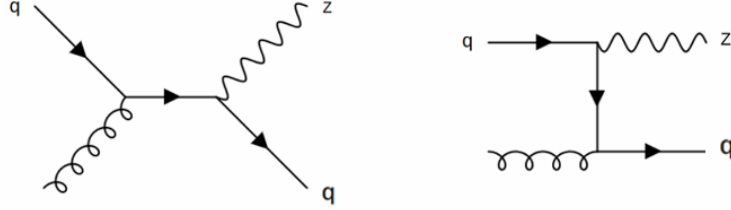


Figure 1.7: Feynman diagrams for the production of a Z boson in association with one jet of quarks.

Since protons have an internal structure, as discussed previously in Section 1.1, partons (quarks and gluons) are the actual protagonists of pp interactions. Quarks have different flavours, each with different mass, that allow for a distinction into light and heavy quarks. Heavy quarks, namely charm (c), bottom (b) and top (t) quarks, happen to have the strong coupling constant in the perturbative region at their mass scale.

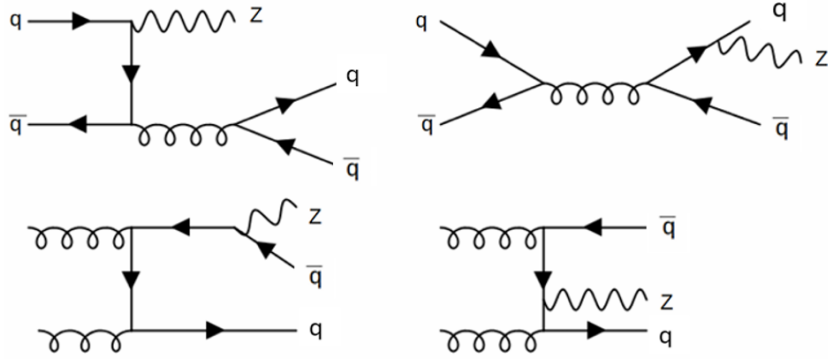


Figure 1.8: Feynman diagrams for the production of a Z boson in association with two jets of quarks.

The signature measured in this thesis consists of two leptons with opposite charge, same flavour and isolated, originating from the decay of the Z boson, and two c-jets resulting from the hadronization of c-quarks. This kind of study provides an

important test of pQCD, applicable in the regime of high energy and short distance interactions. Moreover, they are crucial in the understanding of the contribution that heavy quarks add to parton distribution functions (PDFs). The production of Z+c-jets is a significant background for several analyses at LHC, including the study of the Higgs boson that is produced in association with a Z boson (ZH), as shown in Fig.1.9. In this process, the Z boson decays into leptons and the Higgs boson decays into charm quark pairs ($c\bar{c}$) or bottom quark pairs ($b\bar{b}$).

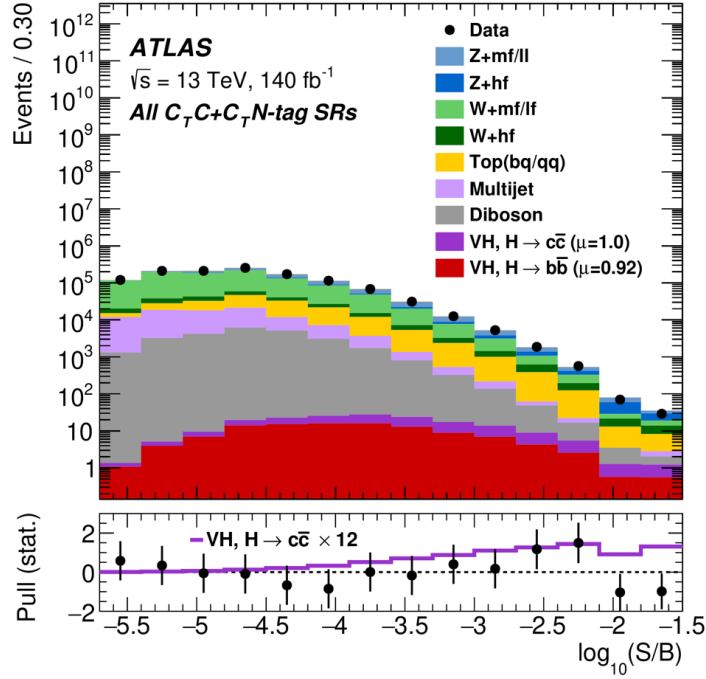


Figure 1.9: The event yields as a function of $\log_{10} (S/B)$ for data, background and a Higgs boson signal for the nominal VH fit. The Higgs boson decays into charm quark pairs ($c\bar{c}$) and bottom quark pairs ($b\bar{b}$). For more details see [13].

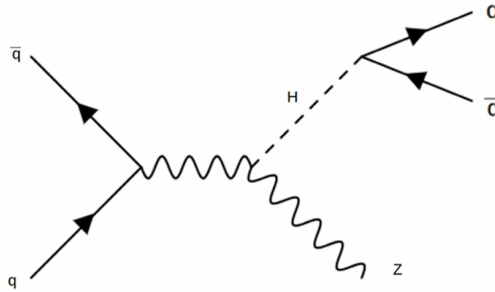


Figure 1.10: Feynman diagram for the production of the Higgs boson decaying in two jets in association with a Z boson.

The Higgs boson decay into a pair of c-quarks is one of the decay branches

predicted by the SM for a Higgs boson with a mass of 125 GeV, corresponding to the mass of the Higgs boson discovered at LHC in 2012, although it has not been measured yet. The decay of the Higgs boson into $c\bar{c}$ pairs is, in fact, one of the SM processes not yet observed. In contrast to $b\bar{b}$ pairs, which have a branching ratio of $\sim 58\%$, the branching ratio of $c\bar{c}$ pairs is estimated to be of the order of $\sim 2.9\%$. In addition to that, due to the large multijet background, this decay can only be studied in scenarios where the signal-to-background ratio is more favorable, namely in associated production with a Vector boson (V). The downside to that is a significant reduction in the production rate. In this context, the production of $W/Z + H(c\bar{c})$, where two c-jets are present in the final state along with the leptonic decay of the W ($W \rightarrow \ell\nu$) or the Z ($Z \rightarrow \ell\ell$), represents one of the leading topologies for studying Higgs physics at the LHC and the Higgs coupling to c-quarks. Since this final state has never been observed, improving the precision of the measurements requires a better understanding of the $W/Z+b$ -jets and $W/Z+c$ -jets background, which currently represents one of the main limiting factors [14].

1.3.3 State of the art of Z+jets measurements

The inclusive and differential cross-sections for the production of Z bosons in association with b-jets have been studied in proton-antiproton ($p\bar{p}$) collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV by the CDF and D0 experiments [15, 16] and in pp collisions at $\sqrt{s} = 7$ TeV at the LHC by the ATLAS and CMS experiments [17, 18]. The CMS experiment then extended these measurements to $\sqrt{s} = 8$ TeV [19]. Both experiments performed studies at $\sqrt{s} = 13$ TeV using the full Run-2 dataset [20, 21]. The most recent ATLAS result at $\sqrt{s} = 13$ TeV full Run-2 dataset measured both the inclusive and differential cross-sections for Z boson production (decaying into electrons or muons) in association with at least one or at least two b-jets, or at least one c-jet.

But while differential measurements of $Z+b$ -jets in a fiducial phase space in pp collisions have been performed at LHC at $\sqrt{s} = 7, 8$, and 13 TeV. The topology under study is constituted by a Z boson in association with a large-radius jet in the boosted regime, coming from a leptonic decay. The reason why a large radius jet is employed is because these measurements take place in the boosted regime, which arises when jets have a high p_T , causing them to be so close together that they cannot be distinguished from one another experimentally, as shown in Fig.1.11. The large-radius jet is then decomposed into smaller-radius sub-jets. The sub-jets are used as proxies for b-quarks originating from the jet. The measurements of the decay products involving at least one b-jet allow testing for b-quark PDF models, and

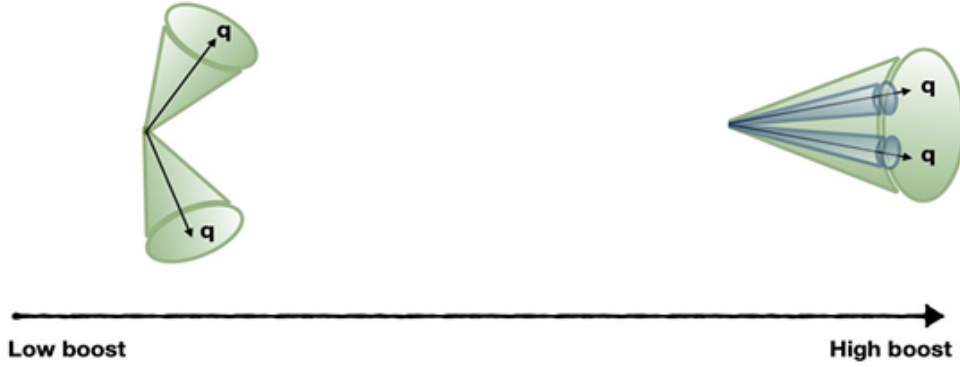


Figure 1.11: Representation of how the resolution of jets change from resolved regime, at low boost, on the left, to boosted regime, at high boost, on the right.

can be used to constrain them, while measurements in association with at least two b-jets allow testing for the gluon splitting ($g \rightarrow b\bar{b}$) mechanism. Said measurements will provide a significant extension in the high-energy regime to the $Z + \geq 1$ b-jets measurements in the standard resolved analysis [22].

On the topic of the methodology of how said measurements are performed, jets of particles are reconstructed at the ATLAS experiment using the hadronic calorimeter. To achieve high angular resolution, inner-detector tracks have been used in place of calorimeter information to construct the sub-jets. Identification of the track-jets as likely (or unlikely) to have been initiated by b- or c-quarks is provided by multivariate jet-tagging algorithms. They exploit the long lifetime in the case of b-hadrons produced by quark hadronization, relying on the impact parameter and on the reconstructed secondary vertex information of the tracks associated with the jets [23]. Systematic uncertainties of the jet selection efficiency in the boosted regime are of the order of 10 – 20%, but are expected to be reduced. To correct the experimental result for the detector effects, such as resolutions and efficiencies, deep-learning neural networks are deployed. Compared to the standard approaches these methods do not require histogramming the reconstructed quantities and arbitrarily high-dimensional data can be used, which allows for new observables to be measured even after the measurement process has come to an end. The boosted $Z + \text{jets}$ analysis will provide the first information on the modeling of observables that serve to discriminate between QCD and Higgs production and could even hint at new physics processes. The main goal of this thesis is to analyze signal of $Z \geq 1$ c-jet, $Z \geq 2$ c-jet, and inclusive $Z + \text{c-jets}$ signal, since they take on an important role in the study of the H boson decay into jets. To do so the first step will be to optimize

the c-jet selection through the quantities of efficiency, defined as $\frac{S_c}{S_{tot}}$, where S_c is the c-jet signal that survived a given variable cut, and S_{tot} is the total signal, and background rejection, defined as $1 - \frac{B_c}{B_{tot}}$, where B_c is the background that survived the cut of the process where a given amount of c-jets were involved, and B_{tot} is the total background. Specifically, in this work, the variable cuts were operated over the missing transverse energy (MET), and the analysis was aimed at optimizing the cut.

Chapter 2

The ATLAS Apparatus at LHC

Founded in 1954, the Conseil Européen pour la Recherche Nucléaire (CERN) is the world largest research center for particle physics. Based in Geneva, CERN was one of the first European joint projects, counting 22 member nations that cooperate to probe the fundamental structure of the universe nowadays. The complex and large scientific instrumentation involves accelerators to boost beams of particles up to very high energy, and detectors to collect and observe the results of the collisions. The Large Hadron Collider (LHC) is the most recent accelerator complex at CERN. Built in a 27 km long circular tunnel and situated from 50 m to 175 m underground, the main LHC detectors are: ATLAS, CMS, LHCb and ALICE. The analysis presented in this thesis has been performed using predictions of events taking place inside the ATLAS detector, the features of which are the subject of this chapter. In Section 2.1 an overview of the LHC is presented, Section 2.2 delves into the components of the ATLAS apparatus, finally Section 2.3 expands on the theme of object reconstruction, with a focus on jets.

2.1 The Large Hadron Collider

LHC was built to address the challenges of the post-LEP (1989-2000) and post-Tevatron (1987-2011) era, exploring the TeV energy scale. As previously stated, the LHC consists of a 27 km long ring where beams of particles travel in opposite directions and are eventually made to collide. The beams are kept apart in two ultra-high vacuum chambers at a pressure of 10^{-10} torr, and collide in four possible points, corresponding to the detectors' locations. The proton-proton (pp) collision mode is the primary operational setup, although LHC can hold heavy ion collisions as well, specifically of lead nuclei [24].

LHC completed two data-taking campaigns: Run1 (up to 2012) where the proton collision energy was $\sqrt{s} = 7$ and 8 TeV, Run 2 (from 2015 to 2018) where the energy

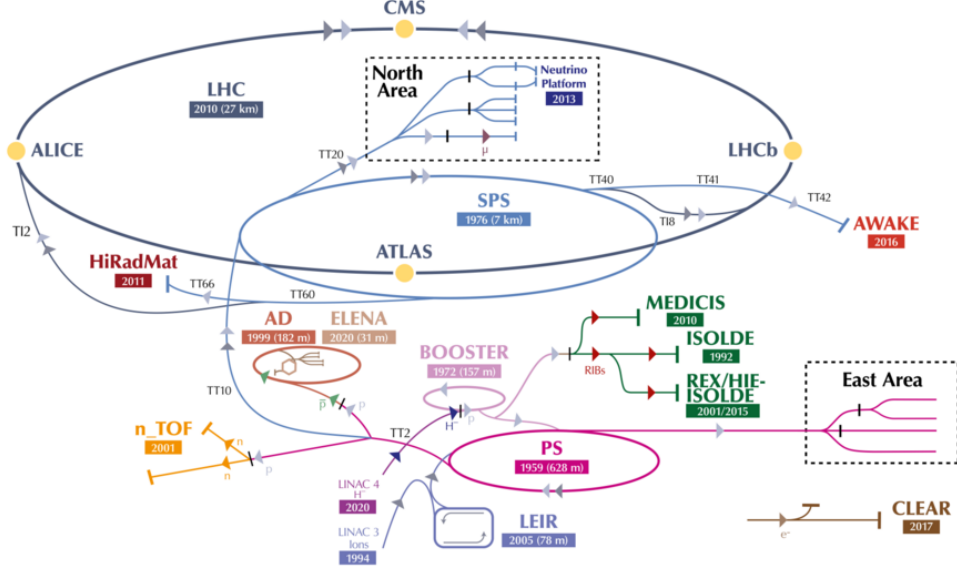


Figure 2.1: Schematic layout of the CERN acceleration complex and experiments.

was $\sqrt{s} = 13$ TeV. Since 2022, the Run 3 campaign is on-going at $\sqrt{s} = 13.6$ TeV.

In the accelerator, particles are accelerated in a vacuum in order to reduce the chance of collisions with air molecules. The protons' trajectory is controlled through dipole magnets, while quadrupole magnets keep the beam focused. The accelerating cavities are electromagnetic resonators that accelerate the beam of particles. Superconducting electromagnets are deployed in order to create a strong enough magnetic field, allowing for tighter turns. The dipoles become superconducting below a temperature of 10 K, therefore the LHC has to operate at 1.9 K to reach a magnetic field of almost 8.3 T. Such low temperatures are managed with a cooling system that uses superfluid helium. The LHC is composed of eight arcs and eight insertions. The arcs contain the dipole bending magnets, while the insertions are long straight sections, each one with two transition regions. The sector is the region between two insertion points and at each sector there are different working units that take care of magnet installation, hardware, powering.

The acceleration process takes advantage of the whole CERN complex, shown in Fig.2.1, the LHC is where the final step takes place. Protons obtained by ionizing hydrogen atoms are first accelerated up to 50 MeV by the linear accelerator LINAC, then the energy of the particles is increased through circular systems. Beams are injected in the following order into the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), reaching 1.4 GeV, 25 GeV and 450 GeV, respectively, before being led into the LHC. The PS and the SPS provide protons in bunches, which were separated by 25 ns in Run 2; a maximum of 2808 bunches, containing about 10^{11} protons each, can be injected into the

	Design	2012	2015	2016
Beam energy (TeV)	7	4	6.5	6.5
Dipole magnetic field (T)	8.33	~ 6.3	~ 8.0	~ 8.0
Number of protons per bunch	$1.15 \cdot 10^{11}$	$1.7 \cdot 10^{11}$	$1.2 \cdot 10^{11}$	$1.18 \cdot 10^{11}$
Maximum number of proton bunches	2808	1374	2244	2076
Number of bunch places	3064	3064	3064	3064
Bunch separations (ns)	25	50	50-25	25
Peak luminosity ($\text{cm}^{-2}\text{s}^{-1}$)	10^{34}	$7.7 \cdot 10^{33}$	$5.0 \cdot 10^{33}$	$1.4 \cdot 10^{34}$
Integrated recorded luminosity (fb^{-1})	80-120	21.3	3.9	36.5
Average pile-up interactions	19	37	15	23

Table 2.1: An overview of design goal specifications of LHC, with performance related parameters during LHC operations in 2012, 2015 and 2016.

LHC ring. Relevant LHC parameters are reported in Tab.2.1 [25].

LHC accelerates two beams in opposite directions before colliding them in four interaction points distributed along its ring, each hosting a detector. ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are the two multipurpose experiments of LHC, while LHCb (Large Hadron Collider beauty) and ALICE (A Large Ion Collider Experiment) are focused respectively on studies concerning flavor physics, and heavy ion collisions.

Fig.2.2 reports the time schedule of LHC.

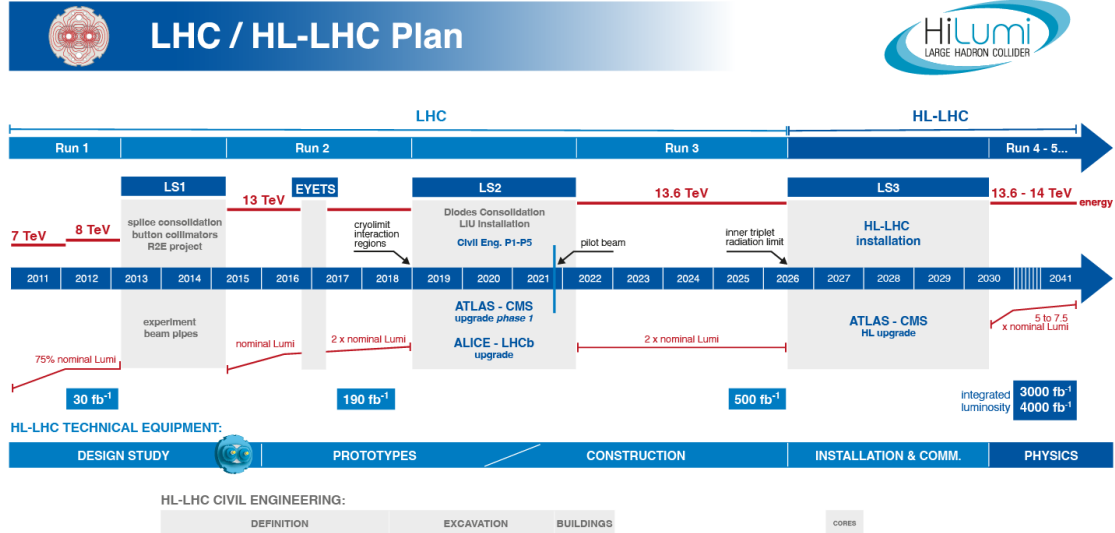


Figure 2.2: Time schedule of LHC, involving active periods and technical shut down, from Run 1 to last upgrade to High Luminosity LHC [26].

2.1.1 The detectors at LHC

The main detectors at LHC are ATLAS, CMS, ALICE and LHCb, located at the interaction points. There are also three other additional detectors: TOTEM, LHCf, and MoEDAL, detecting forward emitted particles and particles along the beam, nearby collision points. A few of their main features are: the high speed response, the good spatial segmentation, other than the ability to select only interesting events.

ATLAS and CMS (A Toroidal LHC ApparatuS and Compact Muons Solenoid) are the biggest detectors at LHC. Their purpose is to study the Higgs boson properties and to investigate new physics beyond the SM. Although the technologies involved are different, the two detectors perform similarly.

LHCb is focused on CP violation and the matter-antimatter asymmetry by studying the decay of states containing b and c quarks.

ALICE (A Large Ion Collider Experiment) is designed to detect the products of heavy ions collisions. It focuses on strongly interacting matter at extreme energy densities in order to study the quark-gluon plasma.

TOTEM (TOTal cross section Elastic scattering and diffraction dissociation Measurement at the LHC) is used to monitor LHC luminosity, and make precise elastic and diffractive measurements of the total pp cross-sections in the forward region.

LHCf uses forward particles to simulate high-energy cosmic rays. It's placed at 140 m from the interaction point of ATLAS.

MoEDAL started in 2015, it is placed in the same interaction point of LHCb and it is used to search for magnetic monopoles.

2.1.2 Luminosity

To assess the performance of the accelerator, a useful quantity is the instantaneous luminosity L , defined as the number of interactions per unit of time and area:

$$L = \frac{1}{\sigma} \frac{dN}{dt} \quad (2.1)$$

where $\frac{dN}{dt}$ is the event rate of a given physics process and σ is the cross section. The instantaneous luminosity is expressed in units of $\text{cm}^{-2}\text{s}^{-1}$. Indeed, the

integrated luminosity is the number of interactions per unit of area:

$$\mathcal{L} = \int L dt \quad (2.2)$$

In the case of a collider, the instantaneous luminosity is measured as:

$$L = f_r n_b \frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \quad (2.3)$$

where f_r is the beam revolution frequency, n_b is the number of bunches in each beam, N_1 and N_2 are the number of protons in each beam, σ_x and σ_y are the gaussian transverse profiles of the beams. The formula works if the collision is head-on, otherwise, multiplication by a corrective factor S that depends on the crossing angle is necessary.

Luminosity measurements for the ATLAS detector are handled by forward detectors such as LUCID (Luminosity measurement Using Cherenkov Detectors), which is the official ATLAS luminosity monitor since the beginning of Run 2. In Fig.2.3 it is shown how the integrated luminosity changes around a year. During Run 1 LHC delivered to ATLAS 28 fb^{-1} of data, of which 25 fb^{-1} were usable for physics analyses. During Run 2 the delivered data added up to 156 fb^{-1} , of which 140 fb^{-1} were usable for physics analyses. The expected luminosity that LHC will deliver by the end of Run 3 is about 300 fb^{-1} .

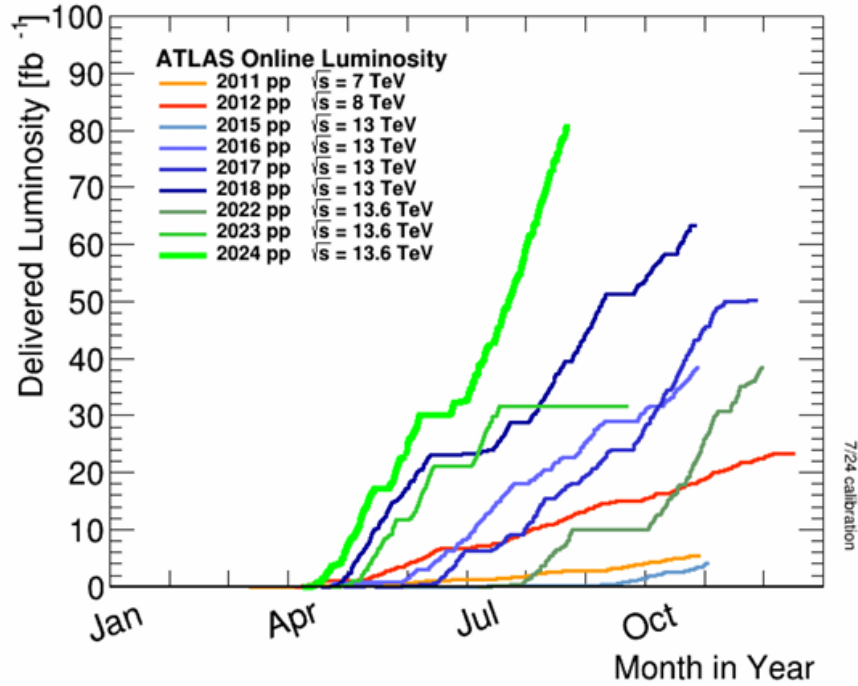


Figure 2.3: Integrated luminosity as a function of data acquisition period [4].

The formula that relates luminosity and beam parameters is as follows:

$$L = \frac{I_1 I_2 f \gamma}{4\pi \epsilon_n \beta^* e^2} F \quad (2.4)$$

where I_1 and I_2 are the currents of the colliding beams, e is the charge of the electron, γ is a factor that increases with the beam energy, ϵ_n is the normalized emittance, β^* is the beta function at the interaction point, F is the geometric reduction factor. Luminosity is also influenced by both pile-up and underlying events. Pile-up $\langle\mu\rangle$ is the mean number of pp interactions per bunch crossing. Since ideally the events collected for physical analysis should originate from a single interaction, a high pile-up poses a challenge for detectors, especially taking into account the fact that with increasing luminosity, pile-up increases as well [4].

In Fig.2.4, the mean number of interactions per bunch crossing is shown for each ATLAS Run.

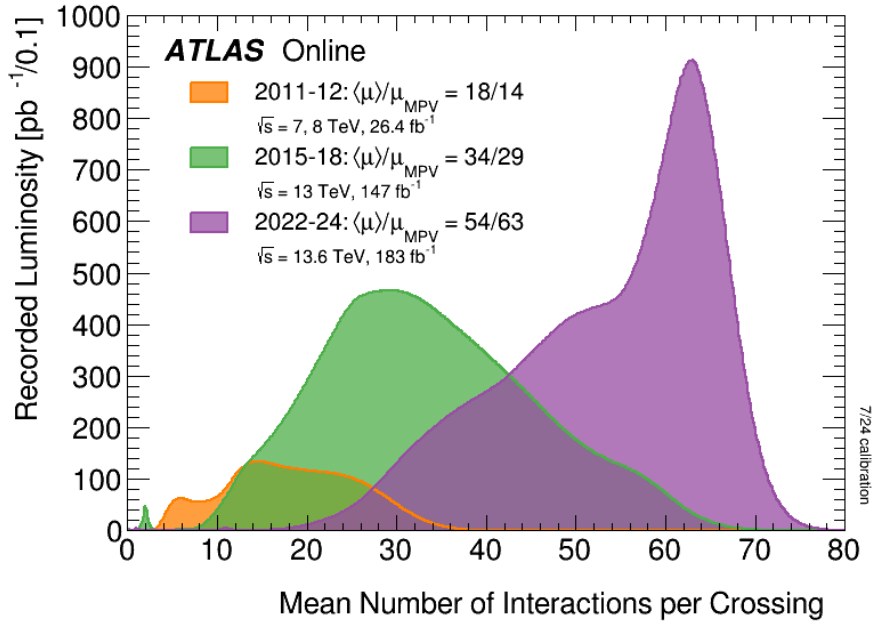


Figure 2.4: Integrated Luminosity as a function of the mean number of interactions per crossing, during Run 1 (in the orange), Run 2 (in the green), and Run 3 (in the purple).

2.2 The ATLAS Detector

ATLAS has a total length of 42 m, a diameter of 22 m and a weight of 7000 t, making it the largest detector at the LHC [27]. It is cylindrically shaped, symmetric with respect to the interaction point (IP), with a solid angle coverage of almost 4π ,

and concentric layers of sub-detectors placed around the pipe where the particle beams travel. The detector is organized in a central barrel and two end-caps, its layout is depicted in Fig.2.5. Since the detector's purpose is to reconstruct the particles generated in the collisions and their respective energy, it should provide as much solid angle coverage as possible, around the IP, while also being fast enough to cover a relevant portion of the events occurring with a high rate. The previously mentioned sub-detectors lining ATLAS play an important role in the identification of the particles resulting from the interactions between bunches and will be individually described in the following subsections.

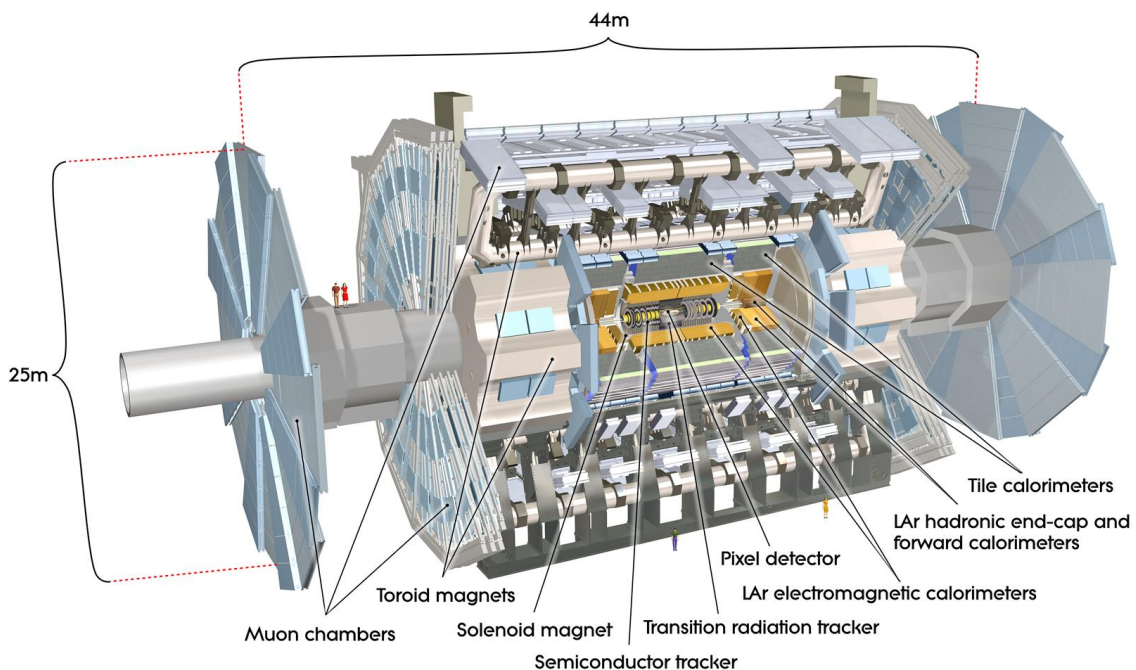


Figure 2.5: The ATLAS detector layout.

2.2.1 The ATLAS reference frame

The ATLAS coordinate system, as shown in Fig.2.6, is right handed: the x-axis points to the center of the LHC ring, the z-axis follows the tunnel while the y-axis points upwards [8]. The origin of the coordinate system is the geometric center of the detector. Polar coordinates are used in the (R, ϕ) plane as well. The angle ϕ is the one in the transverse plane between the x and y axes, and its expression is

$$\phi = \arctan \left(\frac{p_y}{p_x} \right) \quad (2.5)$$

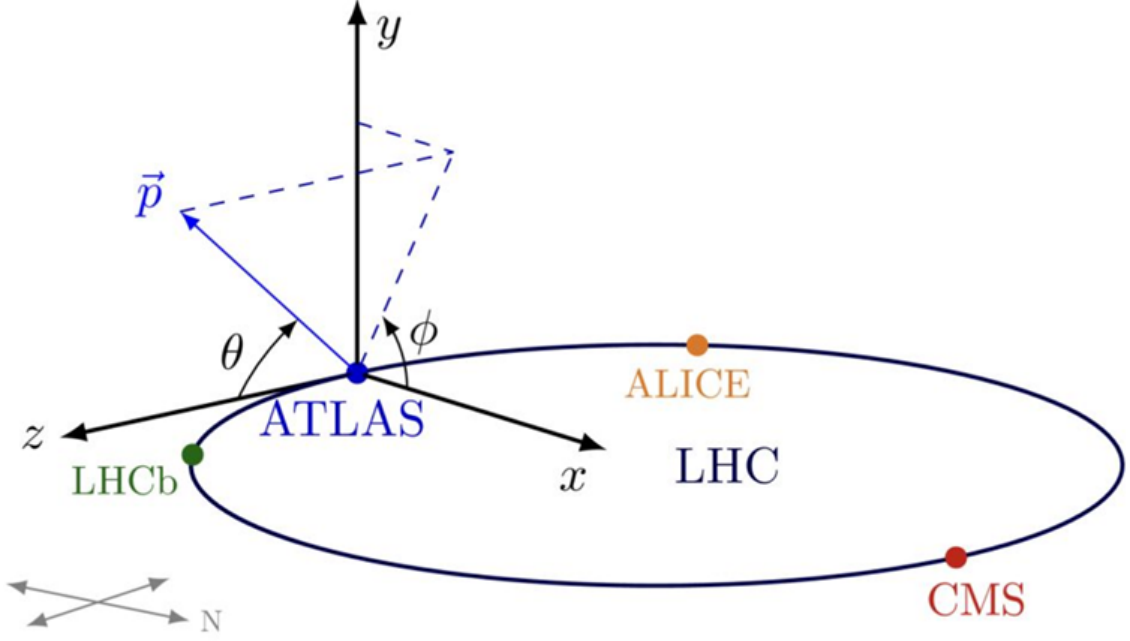


Figure 2.6: ATLAS frame of reference.

where p_y and p_x are the momenta along the x and y axes. The angle θ is the angle between the z-axis and the transverse plane. Another fundamental quantity is the rapidity y , defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (2.6)$$

In the relativistic limit, where mass is negligible, a quantity called pseudorapidity η , whose differences are invariant for boosts along the z-axis, is employed. a visual representation of η is reported in Fig.2.7 and it is defined as:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.7)$$

The transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$, the transverse energy E_T and the missing transverse energy E_T^{miss} are defined in the x-y plane. ΔR is the distance in the $\eta - \phi$ space, and it is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.8)$$

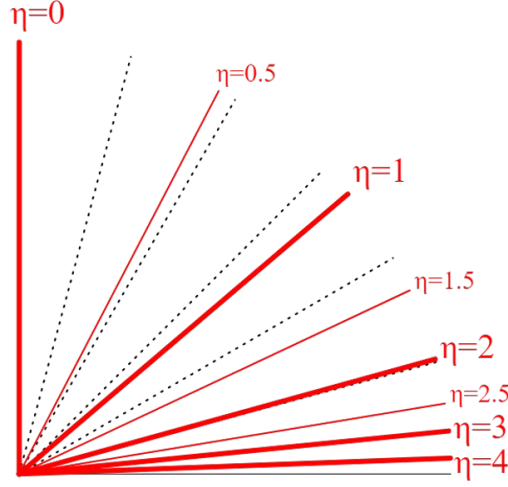


Figure 2.7: Pseudorapidity visualization. As the pseudorapidity increases, the particle track gets closer to parallel to the beam direction.

The following parameters are employed to parametrize the track:

- the polar angles ϕ_0 and θ_0 ;
- the transverse and longitudinal impact parameters z_0 and d_0 ;
- the ratio between the charge and the momentum of the particle.

2.2.2 The inner detector

Particle tracking is handled by the Inner Detector (ID), shown in Fig.2.8 the closest to the IP. Composed of concentric layers of detecting material, the inner detector has a total radius of 1.2 m and length of 6.2 m [27]. In terms of pseudorapidity that would correspond to a region of $|\eta| < 2.5$, beyond that region the detector can no longer rely on the ID, and the system of calorimeters and the muon spectrometer come into play. The ID mainly reconstructs momenta and charged tracks. A central solenoid generating a magnetic field of 2 T allows the ID to measure charge and momentum from the way particles are bent by interacting with the magnetic field. Extrapolating tracks to their origin point in the beam-pipe allows for vertices reconstruction, the primary vertex corresponding to the point in space where the pp collision took place, and the secondary vertex to the position of heavy particle decay. The ID is also divided into sub-detectors itself, these are, from the inner layer out: the Insertable B-Layer, the Pixel and the Silicon microstrip of the Semi Conductor Trackers, used along with the Transition Radiation Tracker.

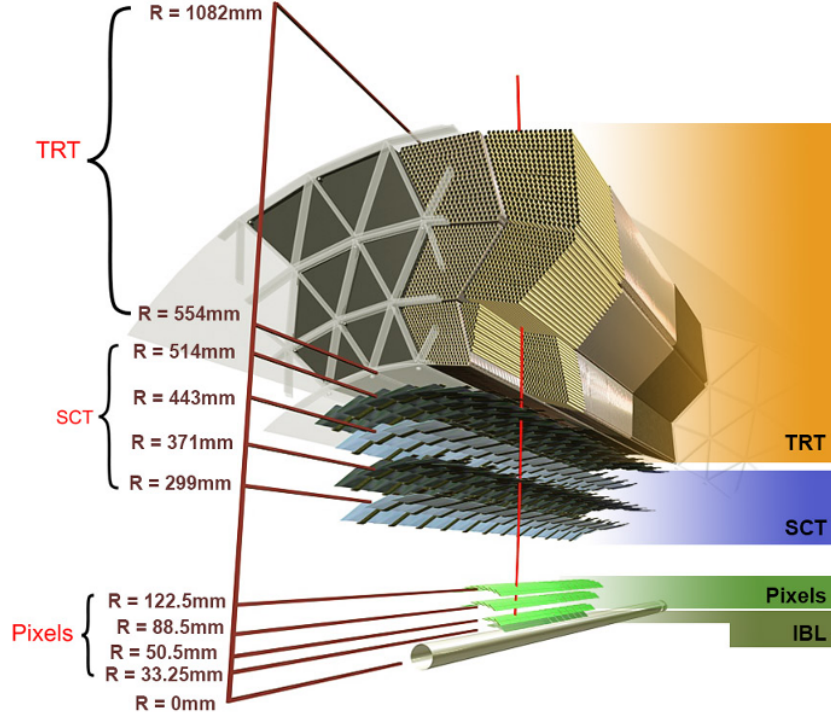


Figure 2.8: An illustration of the ATLAS Inner Detector, including the new insertable B-layer (IBL) added between Run 1 and Run 2. The red line indicates the trajectory of a hypothetical particle with $p_T = 10$ GeV and $\eta = 0.3$.

Insertable B-Layer and Pixel Detector

The first barrel layer is the Insertable B-Layer (IBL), at 3.3 cm from the beam axis, made of pixels with a size of $50 \times 250 \mu\text{m}^2$ and a $200 \mu\text{m}$ thickness, in the high $|z|$ region the thickness is of $230 \mu\text{m}$. The IBL was added during the LS2 (see Fig.2.2) to improve measurement quality over the impact parameter, vertex reconstruction and over the b-tagging performance [28, 29].

At 5.05 cm distance from the beam axis, the B-Layer is found, along with Layer 1 at 8.85 cm and Layer 2 at 12.255 cm, these distances and the layers' dimensions were chosen in order to maximize the probability that a particle crossing one layer will end up crossing the other two layers as well.

Because of its close distance to the barrel, where particle density peaks, the Pixel Detector has a thin granularity, and its pixels are made by the same pixels as the IBL, differing only in shape, being $50 \times 400 \mu\text{m}^2$ in size and $250 \mu\text{m}$ in thickness. The end caps are made up by the same pixels as the layers following the IBL. In terms of pseudorapidity, the whole pixel detector covers a range of $|\eta| < 2.5$ [5, 8].

Semi-Conductor Tracker

The Semi-Conductor Tracker (SCT), is a strip detector composed by a central barrel and two end caps, that employs silicon microstrips composed by p implants on n substrates. It can only measure a single coordinate on the sensor plane but in each barrel layer a pair of sensors with strips that aren't parallel is placed, allowing to measure both the transverse plane and the z -axis coordinates. Divided into four cylinders of radii that go from 299 mm to 515 mm, the barrel covers a range of pseudrapidity $|\eta| < 1.1 - 1.4$ for the length of 1492 mm. The remaining range coverage going up to $|\eta| = 2.5$ is provided by 9 end caps disks of 56 cm radius [5].

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) is the outermost section of the ID, it is a system of tubular trackers that are 4 mm in diameter, made of kapton and carbon fibers, filled with a mixture of Xenon, Argon, CO_2 , O_2 , and a gold-plated wire of Tungsten. The TRT contains 52544 of said tubes, with length amounting to 1.5 m, parallel to the beam axis. The tubes cover a pseudrapidity range of $|\eta| < 1$, with a radius that goes from 0.5 m to 1.1 m. The remaining range $1 < |\eta| < 2$ and $0.8 < |z| < 2.7$ m is managed by the end caps. The passage of ionizing particles through the tubes induces a signal on the anodes, which is one way the TRT can identify particles. Other than through ionization processes, the TRT, being composed of layers of materials with different refraction indices, makes it possible to identify light and heavy particles from the radiation they emit in the transition from one material to another depending on their velocity [5].

2.2.3 System of magnets

The magnet system [30], shown in Fig.2.9, is a core component of the ATLAS detector, not only does it give ATLAS its original shape, but it also provides a strong magnetic field, fundamental to bend particles' trajectories enough to allow accurate measurements and track reconstruction. The ATLAS detector can be broken down to two main systems of magnets: a Solenoid and a Barrel Toroid (BT), along with two End-Cap Toroids (ECT). A schematic representation of the ATLAS magnet system is shown in Fig.2.9.

- The **solenoid** is 5.3 m long and with a diameter of 2.5 m, its axis is parallel to the beam-pipe. It is placed around the ID providing the 2 T magnetic field which allows for charge and momentum measurements. It was designed to have as low material thickness as possible in front of the calorimetry system [8].

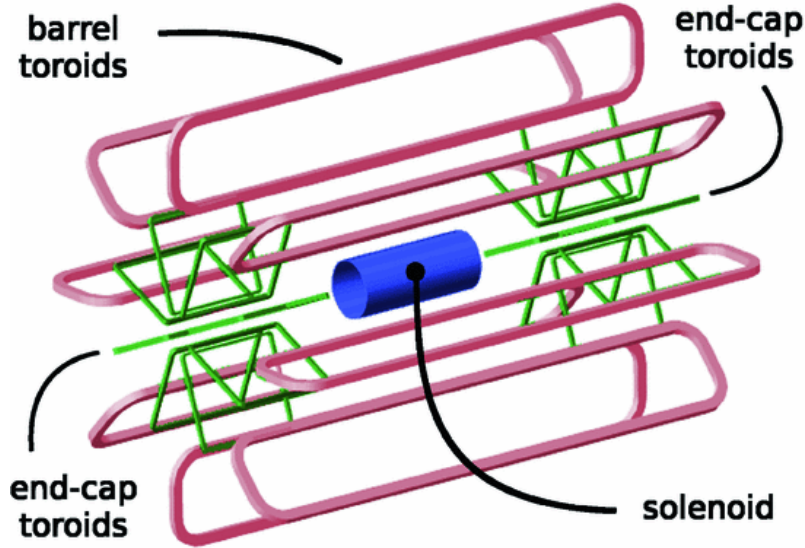


Figure 2.9: The ATLAS magnets system.

- The **toroid** system provides a 4 T field, mainly perpendicular to muons' trajectories. It is composed by 8 BTs of 25 m in length (with the inner core measuring 9.4 m and the outer diameter 20.1 m), and two ECTs that are 5 m long (with the inner core measuring 1.64 m and the outer diameter 10.7 m). The BTs bend particles in the pseudorapidity range of $|\eta| < 1$, the ECTs provide for a range of $1.4 < |\eta| < 2.7$, while a combination of both BTs and ECTs manages the transition region ranging from $\eta = 1$ to $\eta = 1.4$ [8].

2.2.4 The calorimetry system

The layers following the ID inside ATLAS are the calorimeters (shown in Fig.2.10), providing coverage up to $|\eta| < 4.9$ [27]. The calorimeters deployed in the ATLAS detector are made up of two main types of material: the absorber and the active material. The interaction of particles with the absorber triggers a cascade of secondary particles (particle shower), while the active material generates measurable signals, allowing energy measurements via reconstructing the primary particle's energy through the sum of the energies deposited inside the calorimeters by the secondary products. The central section of the system has also a high enough granularity to reconstruct the primary vertex and the interacting particle's direction. Since photons, electrons and hadrons interact differently with matter, different types of calorimeters are employed in order to measure different particles' energy. Hermeticity is a fundamental feature for a calorimeter to achieve high energy resolution, focusing on jets, the large radius of the toroidal magnet muon system is key for the calorimeters to contain jets.

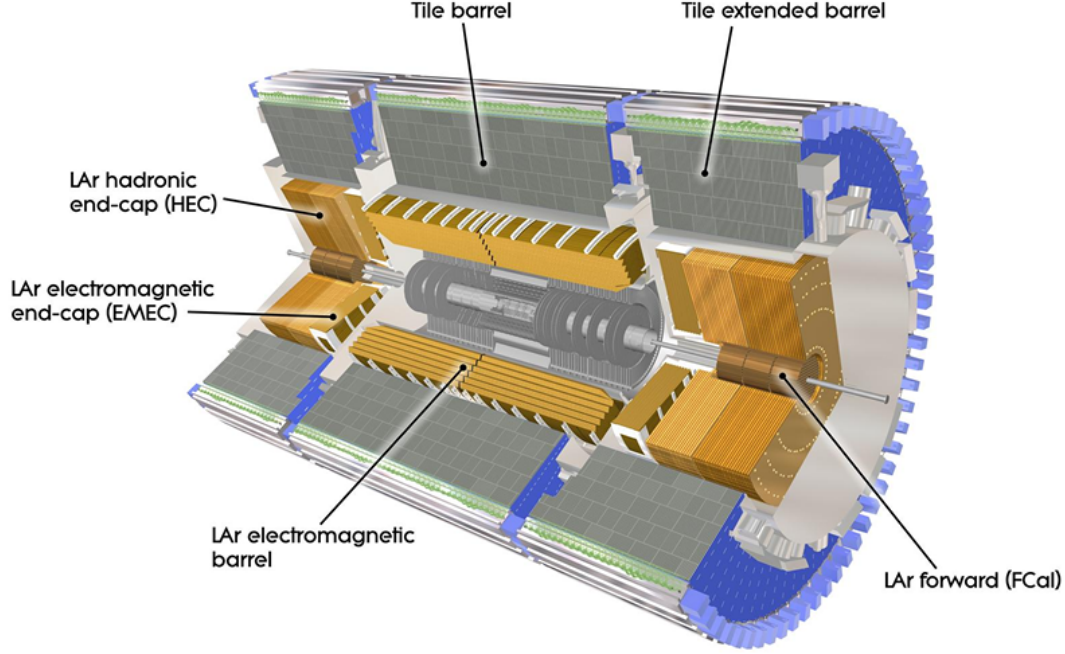


Figure 2.10: Overall view of the ATLAS calorimetry system

Electromagnetic Calorimeter

Also known as Liquid Argon calorimeter (LAr), the Electromagnetic Calorimeter (ECAL) is made up of four sampling calorimeters, using liquid argon as the active material and lead as the passive material. The ECAL is shaped as an accordion, achieving complete azimuthal coverage. The barrel (EMB) covers a range of $|\eta| < 1.475$ while the end-caps (EMEC) account for the range $1.375 < |\eta| < 3.2$. The ECAL is divided into three layers [27]. The first layer is made up of strips and it allows to separate charged pions from neutral pions, in front of the first strip, in the range $|\eta| < 1.8$, a pre-sampler provides additional information about the loss of energy that photons and electrons went through before reaching the calorimeter. The middle and outer layers consist of towers, the outer layer is designed to detect particle showers caused by photons and electrons with an energy of $E > 50$ GeV. The segmented design is key to achieve high precision spatial measurements and helps tracing back photons coming from a primary vertex.

Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is made up of two sampling calorimeters of different composition, the barrel shaped Tile Calorimeter (TileCal) and two end-caps (HEC). The TileCal uses scintillators as active material and iron as the absorber. It is composed by a long barrel (LB) in the central region accounting for $|\eta| < 1$ and two extended barrels (EB) covering the interval $0.8 < |\eta| < 1.7$ separated by a

gap filled with scintillators to recover energy losses occurring between the two barrel regions. The HEC accounts for the range $1.5 < |\eta| < 3.2$, it is composed by two wheels of 2 m outer radius, using copper as the passive material and liquid argon as the active material.

Forward Calorimeter

The Forward Calorimeter (FCal) covers a pseudorapidity region of $3.1 < |\eta| < 4.9$ close to the beam-pipe. It is divided into three modules, providing measurements for both electromagnetic and hadronic particles. It has a cylindrical shape and deploys liquid argon as the active material and copper in the first module as the passive material for electromagnetic measurements, while tungsten is used in the second and third module as the passive material for hadronic interactions. Being close to the beam-pipe, the FCal is subject to high levels of radiation, for that reason, it employs an intrinsic radiation hard technology. Despite having a lower energy resolution FCal plays an important role in measurement concerning the missing transverse energy (MET) and the reconstruction of very forwarded jets [5].

2.2.5 Muon Spectrometer

A specific detector is used to measure the momentum of muons and identify them: the muon spectrometer (MS). Muons travel in the detector farther than other charged particles generated in the collisions, since they hardly interact with materials. Muons emit bremsstrahlung radiation way less than electrons and have a long lifetime, which is why the muon spectrometer is placed in the outermost part of ATLAS within the magnetic field, which is necessary to perform momentum measurements via trajectory bending.

The spectrometer is divided into three parts: a barrel, made up of three more layers covering a range of $\eta < 1.05$, and two endcaps, wheels perpendicular to the beam direction covering a range of $1.05 < |\eta| < 2.7$.

The MS is also composed of several sub-detectors, such as Monitored Drift Tubes (MDTs) which provide precise measurements of the muons' momenta, and Cathode Strip Chambers (CSCs), accounting for the high pseudorapidity region. Knowing the law that binds momentum, curvature radius, and magnetic field, $p = 0.3BR$, it is possible to obtain momenta measurements since B is fixed and R is provided by the MDTs and the CSCs. CSCs are multi-wire proportional chambers that contain cathode planes divided in strips perpendicularly to the beam direction, making it possible to measure on both the bending and the transverse planes. Finally, CSCs' time resolution complements the MDTs. The MDTs provide a momentum

resolution for 1 TeV muons of $\sigma_{p_T}/p_T \sim 10\%$, dropping to 2-3% for muons with lower momentum. For low-momentum muons, measurements of the MS must be complemented with measurements from the ID.

Another type of sub-detectors deployed in the MS are fast response chambers managing on-line triggering. Namely, the MS makes use of Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). The latter type of sub-detectors can identify whether an event has occurred from the bunch crossing with an accuracy of 99%. The RPC sub-system is made up of planes arranged in three cylindrical layers around the z -axis, and in order to measure ϕ and pseudorapidity, each plane is made up of two additional layers. TGCs are complementary to MTDs and provide measurements of both the bending plane coordinate and ϕ .

2.2.6 Trigger and DCS

In ATLAS events happen with a frequency of 40 MHz, which has to be reduced by the Trigger and Data Acquisition (TDAQ) system in order to reach an amount of data that can be stored and processed. The TDAQ selects in a short amount of time which of the occurring events are worth analyzing, reducing the event rate from 40 MHz to 1 kHz. The system during Run 1 consisted of three trigger levels of increased selectivity: Level 1 (L1), Level 2 (L2) and the Event Filter. During Run 2 the triggers underwent several improvements, which brought L2 and the Event Filter to be merged into one single filter: the High Level Trigger (HLT) [31].

L1 takes on the first selection, it is hardware based and, with a decision time of 25 μs , it reduces the event rate to 100 kHz, identifying high p_T muons, electrons, photons, jets, taus decaying into hadrons, and missing energy. L1 takes information from the MS and from the calorimeters to locate Regions of Interest (RoIs), which represent the position of the object that sent off the trigger in the $\eta - \phi$ plane.

The HLT, which is software based, receives the reduced data from L1 and, with a 200 ms decision time, it reduces the event rate to approximately 1 kHz, finalizing the data selection process for off-line analysis.

The trigger system is monitored by the Detector Control System (DCS), active during all periods of data collection. The DCS monitors every aspect of ATLAS, making sure that everything works fine.

2.3 Object Reconstruction and Identification

Exploiting the technologies described in the previous sections, ATLAS can perform particle and particle track reconstruction, by combining the signals deposited in

different parts of the detector. A depiction of particle tracks reconstructed from an event at ATLAS is shown in Fig.2.11. The main particles of interest in this thesis are leptons, specifically electrons and muons, and jets made up by c-quarks.

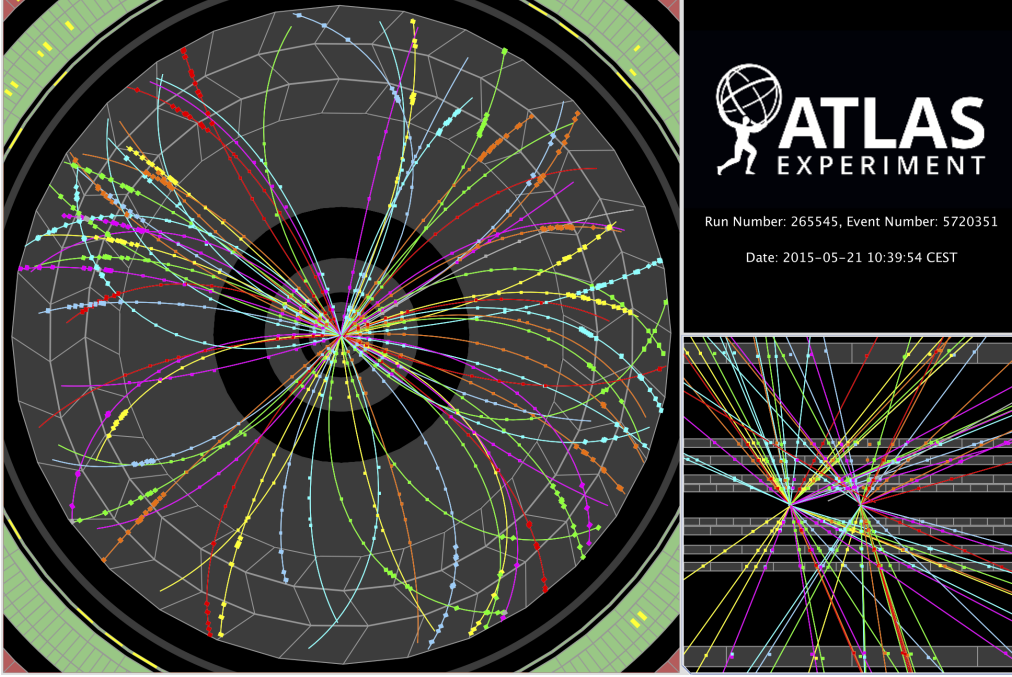


Figure 2.11: Display of a pp collision event recorded by ATLAS on May 21st 2015, at 13 TeV collision energy. Tracks reconstructed from hits in the inner tracking detector are shown to originate from two interaction points, indicating a pile-up event [32].

2.3.1 Electrons

Electrons emit photons via bremsstrahlung, which at high energy cause the formation of electron-positron pairs, generating an electromagnetic shower. Electrons are charged particles, which means they leave a signal in the ID's pixel detector and in the SCT detector. The reconstruction is performed through the formation of hits clusters, that are then converted via pattern recognition algorithm to points in the 3-dimensional space. Further filters are applied for tracks that do not meet the initial criteria, providing precise reconstructions [4].

The EMCal is employed as well, since electrons deposit energy in that calorimeter, too. Said energy is obtained through the sum of the energy deposited in each layer of the calorimeter and deploying algorithms that form clusters by identifying localized deposits of energy, and to determine which tracks are worth keeping.

Eventually the ID tracks are matched with the calorimeter's clusters. The distance in η and ϕ between tracks, the cluster's barycenter in the second layer of the

EMCal, the hits number in the pixel detector, are some of the criteria through which the primary track is selected.

Electrons themselves are identified through a likelihood method (LH). The LH is able to separate signal (prompt electrons) and background (electron from electromagnetic showers, jets, etc.). The LH takes both the signal (L_S) and the background (L_B) likelihoods, the product of probability density functions and other parameters (x_i):

$$L_{S(B)}(x) = \prod_{i=1}^n P_{S(B),i}(x_i) \quad (2.9)$$

where $P_{S,i}(x_i)$ is the value of the signal probability distribution function (pdf) for the i -th at x_i , and $P_{B,i}(x_i)$ is corresponding value for the background PDF.

Then the discriminant d_L is calculated as:

$$d_L = \frac{L_S}{L_S + L_B} \quad (2.10)$$

peaking at one for signals electrons, and at zero for background.

The minimum value of the LH discriminant defines three operating points (OPs) for physics analyses: LooseLH, MediumLH, and TightLH. In this analysis the electrons (as well as the muons) must have transverse momentum of $p_T > 27$ GeV and $|\eta| < 2.5$.

2.3.2 Muons

Muons leave tracks in the ID and in the MS, in the two detectors, reconstruction is performed independently. The process begins with the search for hit patterns in each muon chamber [33].

The MDTs employ a Hough transform to identify aligned hits, which are then fit with straight lines. In the CSCs the search for hits segments is performed in the $\eta - \phi$ plane.

The track are reconstructed by combining segments from different layers, starting from the central layers, where usually more hits are available, to the outer and inner layers. To form a track the required minimum is of two matching segments.

Muons are further categorized depending on the nature of the reconstruction (i.e. the algorithms employed and the sub-detectors involved):

- Combined (CB): independently tracked by the ID and the MS, and combined through a global fit.
- Segment-Tagged (ST): if the candidate track in the ID matches at least one

MS segment, especially useful for muons with a low transverse momentum p_T or for areas which the MS cannot cover or has low coverage over.

- Calorimeter-Tagged (CT): if the candidate track in the ID matches an energy deposit in the calorimetry system, with a minimum ionizing particle.
- Extrapolated (ME): reconstructed with the only use of the MS, extrapolated to the interaction point, especially useful for regions in the pseudorapidity range of $2.5 < |\eta| < 2.7$, where the ID cannot provide any coverage.

Similarly to electrons, muons are identified through four OPs: Loose, Medium, Tight and High- p_T . For the Loose OP, CT muons and ST muons are considered only in the pseudorapidity region of $|\eta| < 0.1$, CB and ME muons are not constrained. The Loose OP maximizes the reconstruction efficiency. The Medium criteria takes into account only ME and CB tracks, and is aimed at minimizing systematic uncertainties. The Tight OP takes only CB muons into account, as long as they are reported to have hits in at least two stations of the MS and satisfy the Medium OP. The Tight OP, while being less efficient than other criteria, maximizes the purity of the detected muons. Finally the High- p_T criterion selects only muons that satisfy the Medium OP, and maximizes the momentum resolution for tracks with $p_T > 100$ GeV. In this analysis the muons (as well as the electrons) must have transverse momentum of $p_T > 27$ GeV and $|\eta| < 2.5$.

2.3.3 Jets

Hadronic collisions at ATLAS produce large numbers of quarks and gluons, which hadronize immediately due to confinement, producing a collimated particle shower, the net momentum of which is equal to the initiating particle's. The particle shower in question is addressed as "jet", and happens extremely often in collisions taking place at LHC. Jet tracks are the ensemble of ID tracks left by the particles that made up the jet. However, when it comes to track reconstruction, jet is more often the term referring to the calorimeter object, reconstructed from cone-shaped energy deposits. Traces of jets could also be found, other than in the ID and in the calorimeters, in the MS, whenever a muon is emitted as a component of the jet, or if the HCAL ends up not absorbing some of the energy of the jet.

The jet reconstruction starts with a topological clustering algorithm, which forms clusters from a high energy cell [34]. In order to be eligible for clusterization, the cells are required to have an energy significance of at least 4σ above the noise level, defined as the quadratic sum of electronic signal and pile-up signal. Nearby cells which have an energy significance that is 2σ above the noise level, are added to build

topo-clusters via iteration. Topo-clusters are used to group together neighboring cells with relevant deposits of energy in the $\eta - \phi$ plane, to reconstruct jets. After the topo-cluster reconstruction, clusters are divided into energy categories through a local maximum energy criterion.

In jet reconstruction the main strategy adopted is aimed at determining whether the separation between two energy clusters is large enough to originate from two separate particle showers, rather than belonging to the same jet. The anti- k_t algorithm [35] recombines clusters with the following procedure:

1. for each cluster i , calculate the distance d_{ij} from all other clusters j

$$d_{ij} = \min(p_{T,i}^{2k}, p_{T,j}^{2k}) \frac{\Delta R_{ij}^2}{R^2} \quad (2.11)$$

where ΔR_{ij}^2 is the angular distance between cluster i and cluster j in the $\eta - \phi$ plane, R is the cone radius of the jet, finally k is an anti- k_t algorithm parameter fixed to the value of -1 .

2. for every cluster i , evaluate the distance from the beam:

$$d_{iB} = p_{T,i}^{2k} \quad (2.12)$$

3. find the minimum distance between d_{ij} and d_{iB}
4. eventually, if the minimum value is d_{ij} , combine i and j into one pseudo-jet and repeat from the top. In case d_{iB} is not the minimum value, consider i as a final state and ignore it in the next iterations.

After reconstructing the jet, another crucial point is to get a correct energy measurement. That requires several corrections due to different effects such as the lower response of the HCAL with respect to the ECAL, the energy deposited outside the calorimeter and outside the jet, pile-up effects and noise thresholds [36, 37, 38]:

- a correction to the jet origin, which, without affecting the energy of the jet, points the jet direction towards the primary vertex.
- a pile-up correction, obtained through in situ measurements, that depends on the number of primary vertices, the jet pseudorapidity and the spacing of the bunch.
- energy and η calibration, that account for energy loss in inert material or for energy deposited outside of the jet cone.

- a residual in situ calibration is applied only to jets in collision data and it is determined from collision events by comparing the energy of the reconstructed jet to a reference object (usually a generic particle or another jets).

To distinguish Z+c-jets from pile-up, ATLAS deploys a track-based tagging approach: the Jet Vertex Tagger (JVT). The JVT is a discriminant which makes use of variables obtained from the sum of the transverse momentum of jet tracks from the ID and the amount of primary vertices. The algorithm employed for the reconstruction is a multivariate (MVA) technique named k-Nearest Neighborhood, and outputs the probability for the subject jet to have originated from a hard-scatter vertex. The probability for a jet to be of signal-type, is calculated as the ratio between the amount of hard-scatter jets and the amount of hard-scatter+pile-up jets that are found in the local neighborhood of the subject jet. Z($\rightarrow \mu\mu$)+jets events are employed to assess the efficiency for JVT in data, exploiting a tag-and-probe procedure [39], and calibrations are provided in bin of p_T of the Z boson.

2.3.4 Flavor Tagging

Jets that originate from the hadronization of b-quarks play a fundamental role in several analyses within the ATLAS program, including Standard Model precision measurements, and Higgs studies. b-hadrons have many unique properties that can be used to recognize them. According to the current theory of the Standard Model, b-hadrons decay is suppressed through the CKM matrix, resulting in a longer life-span, hence a longer traveled distance in the inner detector before decay occurs, compared to other quarks. In the laboratory frame that means that a b-hadron will get to travel a few millimeters before undergoing decay. b-quarks are also the heaviest quark that can hadronize, resulting in a heavy hadronic signature. And lastly, a b-hadron produces a higher charged track multiplicity on average than lighter hadrons [8].

MV2 is the algorithm developed for b-jet tagging during RUN 2 [40]. All the algorithms that provide the variables for MV2 take advantage of the long b-quark lifetime, said algorithms are: a likelihood-based combination of the transverse and longitudinal impact parameter significances; the presence of a secondary vertex; the reconstruction of the b-quark decay chain by looking for a shared direction which connects the primary vertex to both the bottom and the charm decay vertices.

The b-jet tagging efficiency is measured for jets in the range $|\eta| < 2.5$ and $p_T > 20$ GeV for many OPs. Four OPs are defined corresponding to 60%, 70%, 77% and 85% b-jet tagging efficiencies.

The calibration of b-tagging algorithms calls for a corrective factor, which is used on simulated data in order to match the b-tagging efficiency measured. Said factor is derived for both correctly identified b-jets and light jets or c-jets that were mistakenly b-tagged. Studies involving the efficiency of said algorithms are performed on top pair samples, because they are enriched in b-jets, benefiting from the high production rate at LHC. There are different methods for b-tagging calibration [41]:

1. the **tag and counting** method fits the multiplicity of b-tagged jets in $t\bar{t}$ events. Most of the events caused by the decay of the t-quark are expected to contain exactly two real jets, according to the SM. Although the number of reconstructed b-jets may amount to two, since b-jets can be lost outside the detector coverage, or there could be additional jets from gluon-gluon splitting, or from c-jets and light-jets being mistakenly tagged as b-jets.
2. the **kinematic selection** method measures the b-tagging rate of the leading jets. It relies on the knowledge of the flavor composition of the $t\bar{t}$ signal and background samples, to then extract the b-jet tagging efficiency through measurements of the fraction of b-tagged jets in the data.
3. the **kinematic fit** method uses a fit of the $t\bar{t}$ event topology to extract a purified sample of b-jets, over which the b-tagging efficiency gets evaluated. Based on a χ^2 fit, it uses the masses of the two t-quarks and W bosons as constraints.
4. the **combinatorial likelihood** method improves the precision of the kinematic selection, by taking advantage of the kinematic correlations between jets involved in the event. It relies on an a priori knowledge of the flavor composition of the $t\bar{t}$ samples.

c-taggers are a retrained version of the b-tagging algorithm. They are optimized to maximize c-jet efficiency for a given b- or light-jet rejection, and they are calibrated using control channels enriched in c-jets such as W+c-jets or D* mesons.

2.3.5 Missing transverse energy (MET)

The ATLAS geometry and coverage allows to reconstruct most of the products of the collisions. The conservation of momentum in the plane transverse to the beam axis (x-y) implies that the sum of the transverse momentum vectors of the collision products should amount to zero. E_T^{miss} is a quantity defined as the negative vectorial

sum of energy in the transverse plane of all objects reconstructed in the event [42], and it is employed to measure the momentum balance of the event, that quantity is known as Missing Transverse Energy (MET). Because neutrinos do not interact with any of the detectors employed in ATLAS, according to the SM, a high E_T^{miss} event suggests the presence of undetected high- p_T neutrinos. However E_T^{miss} can also come from other sources, such as beam halo muons, cosmic muons and possible physics beyond the SM.

The E_T^{miss} components on the x-y plane are computed as:

$$\begin{cases} E_x^{miss} = - \sum_{i=1}^{N_{cell}} E_i \sin \theta_i \cos \phi_i \\ E_y^{miss} = - \sum_{i=1}^{N_{cell}} E_i \sin \theta_i \sin \phi_i \end{cases} \quad (2.13)$$

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (2.14)$$

where E_i , θ_i and ϕ_i are, respectively, the calorimeter cell energy, the polar angle and the azimuthal angle. The sum excludes cells flagged as noisy by the ATLAS database.

Chapter 3

Data Analysis

Optimization of the signal-to-background ratio for processes involving the production of a Z boson and c-jets is the main goal of this thesis. To perform said task the prediction has been provided by the MC generators Sherpa, MadGraph with the FxFx algorithm, Pythia and Powheg. Said data have then been analyzed and plotted through the use of ROOT, and the results are presented in the following sections, with a focus on Missing Transverse Energy.

3.1 Event Generators

Due to the impossibility to perform exact calculations in non-perturbative QCD, to fragmentation through different materials, and to the geometry of the detectors, there is no access to any sort of analytical approach to the calculation of cinematic distributions for the events taking place inside the collider. However, the detector's responses can be simulated through MC generators, offering a valid aid in data validation, refining event selection criteria, and in the evaluation of systematic uncertainties. MC generators provide a tool to simulate a high number of random events originating from a chosen process, and, therefore, to recreate cinematic distributions. Statistical fluctuations leading to uncertainty can easily be reduced by incrementing the number of simulated events.

The generators deployed to simulate signal for $Z \rightarrow ll$ are MadGraph+FxFx combined with Pythia8, and Sherpa 2.2.14, while Powheg combined with Pythia8 was used to simulate $t\bar{t}$. A few examples of MC signal and background predictions for different variables are shown in Fig.3.1.

MadGraph is employed to generate Matrix Elements (ME) for specific processes, in this case it was used for MEs up to three partons at NLO. Pythia8, which is a general-purpose generator that models everything from the hard scattering to the hadronization phase, was combined with MadGraph to separately generate the

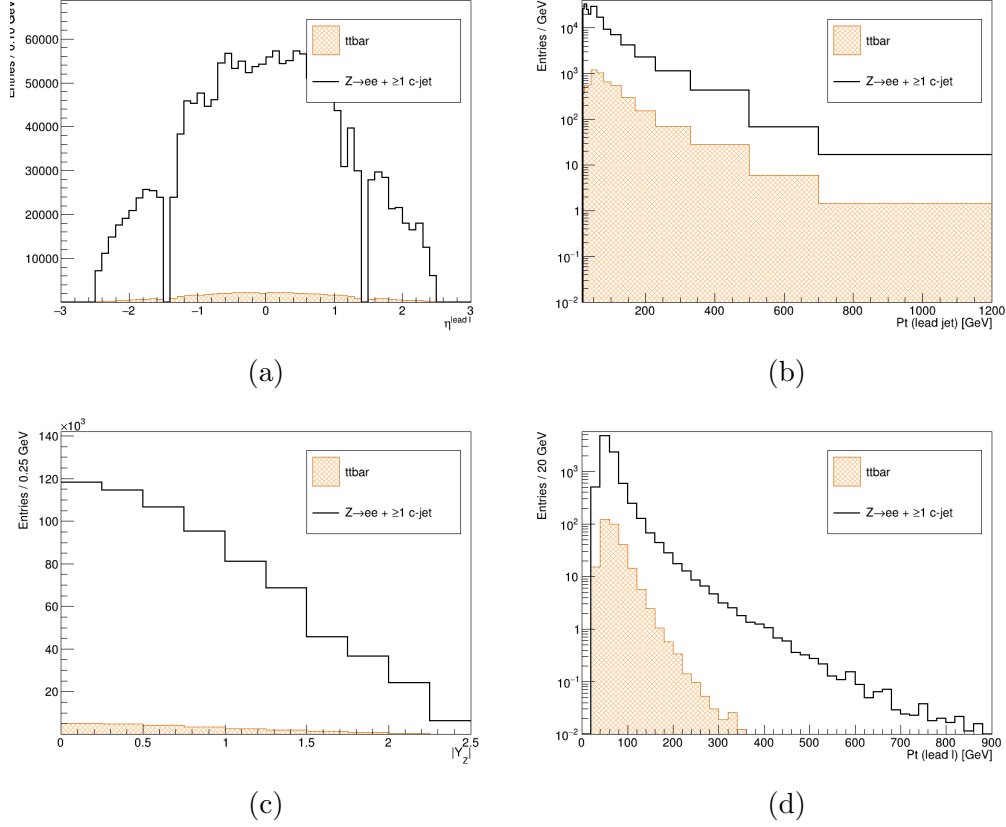


Figure 3.1: Examples of Signal entries, represented as a black line, and Background entries, in the orange. In the (a) histogram predictions for the pseudorapidity η of the leading lepton is shown, in (b) the transverse momentum of the leading jet is shown, in (c) the rapidity of the Z boson is shown, finally in (d) the transverse momentum of the leading lepton is shown.

parton showering and the hadronization leading to jet formation, while FxFx was the algorithm used to merge the different jet multiplicities. Sherpa 2.2.14 was used to generate both matrix elements and the processes of parton showering and hadronization. Sherpa is a general-purpose generator that uses the cluster model for hadronization, and it separates itself from Pythia by integrating ME calculations with parton showers, standing out for its precision. All the predictions generated come with a theoretical uncertainty, mainly the one coming from the QCD scale and the PDF uncertainties. Processes involving $t\bar{t}$ were generated through Powheg, used for Next-to-Leading Order (NLO) calculations in parton shower simulations, combined with Pythia8 [4]. The predictions were simulated at a center-of-mass energy of 13.6 TeV, with a luminosity of 29 fb^{-1} .

Process	Generator	Order of pQCD in ME
$Z \rightarrow \ell\ell$	MG+FxFx+Py8	0-3p NLO
	Sherpa 2.2.14	0-2p NLO, 3-5 LO
$t\bar{t}$	Powheg+Py8	NLO

Table 3.1: List of the processes under study, their generators, and the number of partons at the leading order (LO) of pQCD in the matrix element (ME) calculation.

3.2 Signal-to-Background optimization

As previously discussed in Chapter 2, high values of E_T^{miss} can be caused by the presence of neutrinos, and, since neutrinos are not involved in the final state of the Z boson production processes that this analysis focuses on, a cut on E_T^{miss} can be required to reduce the background.

One of the main sources of background is the leptonic decay of $t\bar{t}$, which sees the W bosons involved decaying into leptons. Said source of background can be reduced by operating the invariant mass cut, which in this analysis has been performed for lepton masses of $76 < m_{\ell\ell} < 106$ GeV as shown in Fig.3.2 for the electron channel, and in Fig.3.3 for the muon channel.

The main contribution of this work on c-jet selection is to verify whether or not a cut on MET could improve the selection.

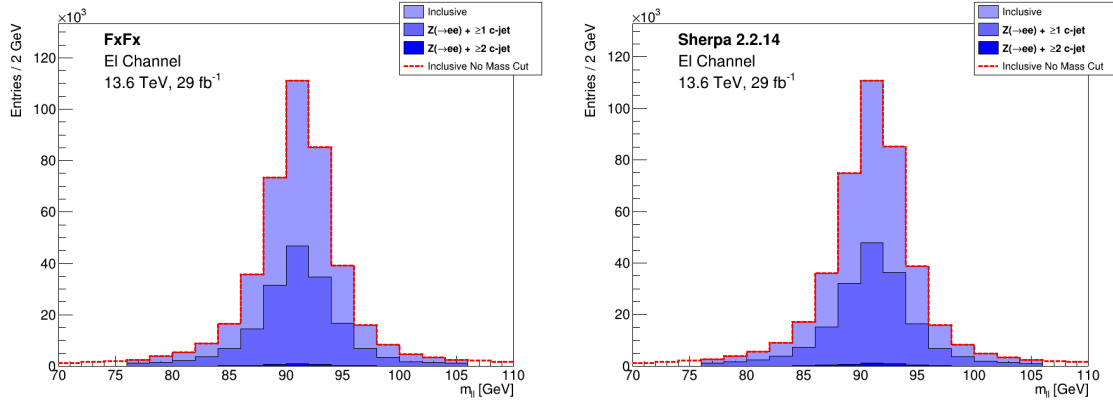


Figure 3.2: Monte Carlo distributions of signal events for the leptons' masses $m_{\ell\ell}$ for the electron decay channel. The red dashed line highlights the inclusive data prior to the cut, showing the distribution beyond the cut's edges. The signal of the plot on the left was generated via MG+FxFx, the one on the right via Sherpa 2.2.14

Despite the mass cut, background can still be relevant at high jet transverse momentum p_T . To further reduce the $t\bar{t}$ contribution a E_T^{miss} cut can be performed. However, when operating cuts parts of both the background and the signal get lost, which is why analysis is needed to assess whether or not the cut is advantageous. For this purpose there are two quantities that should be taken on account: Signal

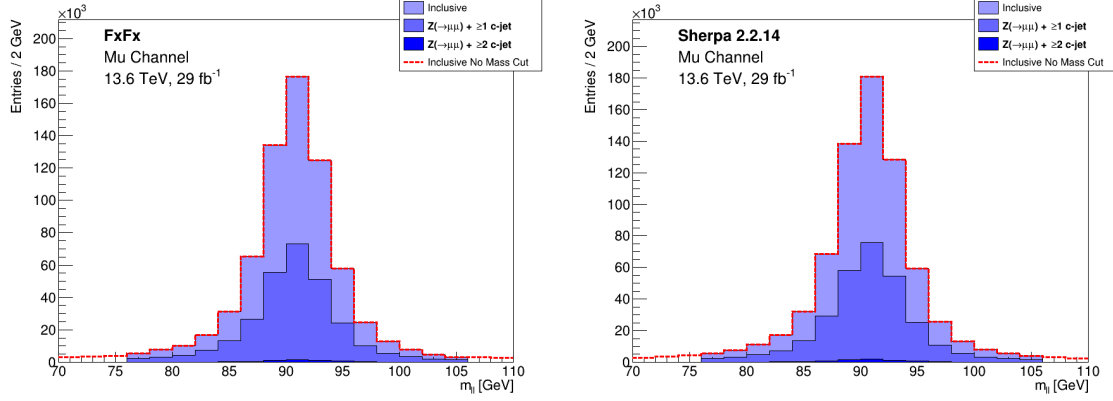


Figure 3.3: Monte Carlo distributions of signal events for the leptons' masses m_{11} for the muon decay channel. The red dashed line highlights the inclusive data prior to the cut, showing the distribution beyond the cut's edges. The signal of the plot on the left was generated via MG+FxFx, the one on the right via Sherpa 2.2.14.

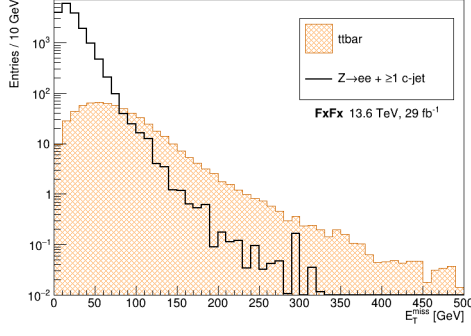
Efficiency (ϵ_S) and Background Rejection ($1 - \epsilon_B$), which are respectively defined as:

$$\epsilon_S = \frac{S_{cut}}{S_{tot}} \quad (3.1)$$

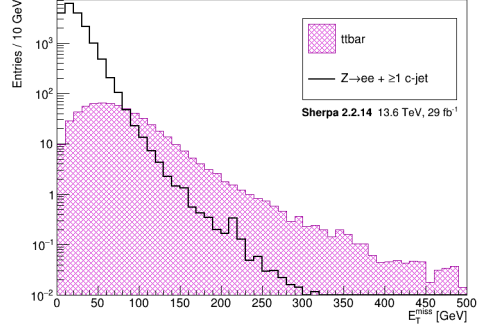
where S_{cut} is the number of signal events below the cut and S_{tot} is the total number of signal events,

$$1 - \epsilon_B = 1 - \frac{B_{cut}}{B_{tot}} \quad (3.2)$$

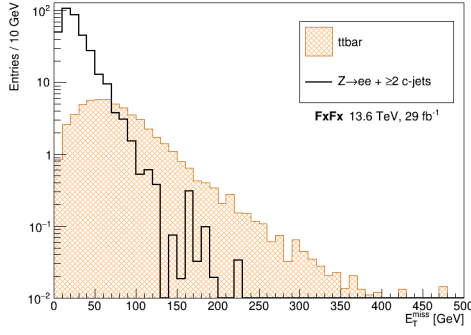
where ϵ_B is the efficiency of the cut on the $t\bar{t}$ sample, B_{cut} is the number of background events passing the cut and B_{tot} is the total amount of background events. The choice of the cut on the E_T^{miss} should be aimed at maximizing both the Signal Efficiency and the Background Rejection. In Fig.3.4 and Fig.3.5 the signal plot is represented as a black line, and it is superimposed over the background, which is plotted in the orange for predictions generated by MG+FxFx and Pythia8, in the purple for predictions generated by Sherpa.



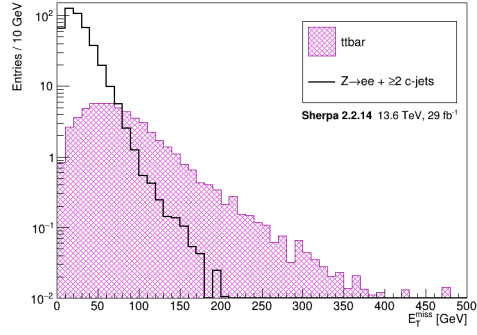
(a)



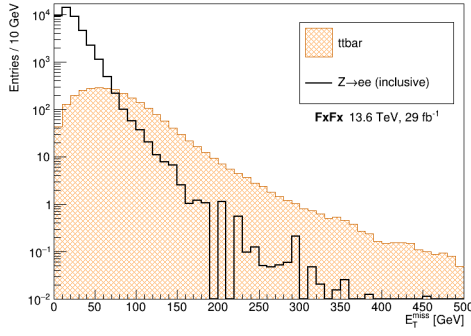
(b)



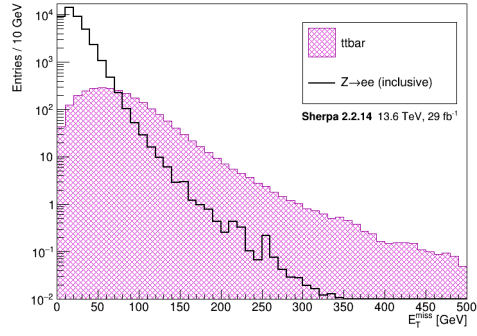
(c)



(d)

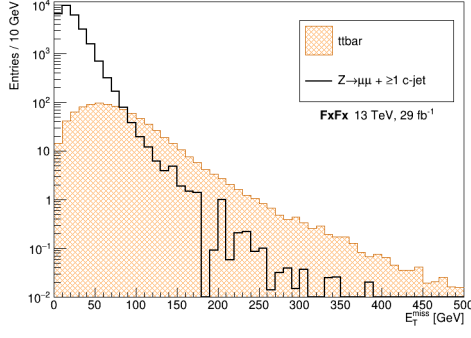


(e)

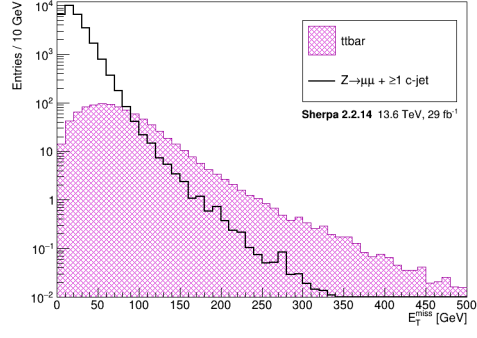


(f)

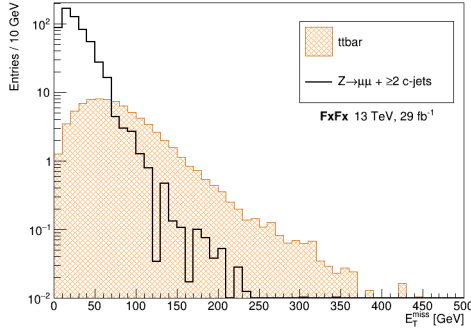
Figure 3.4: Monte Carlo distributions of $t\bar{t}$ and signal events for the electron channel, on the left column distributions generated by MadGraph FxFx are shown, while on the right columns distributions from Sherpa are shown. The top row compares signal classified as containing one or ore c-jets, the middle row shows signal classified as containing two or more c-jets, the bottom row shows inclusive signal.



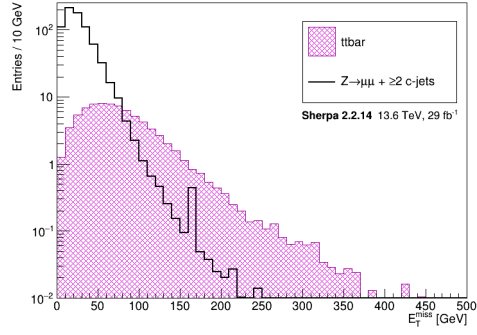
(a)



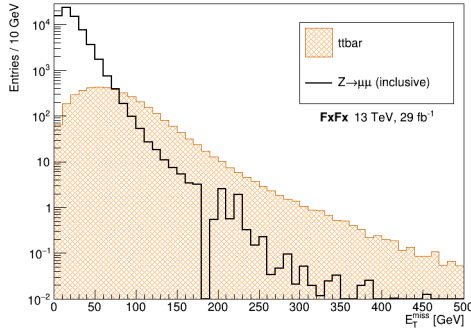
(b)



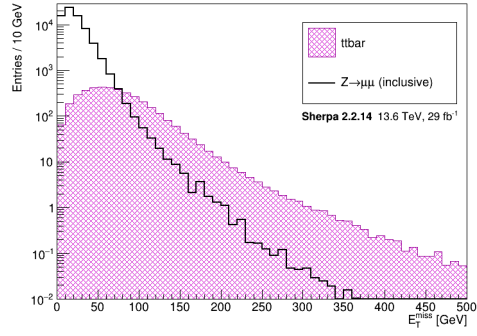
(c)



(d)



(e)



(f)

Figure 3.5: Monte Carlo distributions of $t\bar{t}$ and signal events for the muon channel, on the left column distributions generated by MadGraph FxFx are shown, while on the right columns distributions from Sherpa are shown. The top row compares signal classified as containing one or more c-jets, the middle row shows signal classified as containing two or more c-jets, the bottom row shows inclusive signal.

3.2.1 ROC Curves

By plotting the Background Rejection as a function of Signal Efficiency for every possible cut, an empirical ROC (Receiver Operating Characteristic) curve is built. In the following plots the main performed cuts are highlighted. These are at the energies of 40, 50, 60, 70 and 80 GeV. The ROC curves for both channels and for the different MC generators are shown in Fig.3.6 and Fig.3.8.

In order to find the optimal value for the E_T^{miss} cut, there is a quantity called significance, which is defined as $S/\sqrt{S+B}$, where S and B are respectively the number of the signal and background events that pass the cut. The value of E_T^{miss} that maximizes significance is taken as the optimal value for the cut. The specificity plots are shown for each decay channel and MC generator in Fig.3.7 and Fig.3.9. A comparison between the specificity function before and after the mass cut is shown in Fig.3.10. In Tab.3.2 the number of signal events and background events are shown with their respective uncertainties, which were calculated for the total entries as for independent measurements:

$$\sigma_{tot}^2 = \sum_{i=1}^N \sigma_i^2 \quad (3.3)$$

and then propagated for the Background Rejection and the Signal Efficiency using:

$$\sigma_{tot}^2 = \frac{\partial f^2}{\partial S} \sigma_S^2 + \frac{\partial f^2}{\partial B} \sigma_B^2 \quad (3.4)$$

where f is the generic expression for which the uncertainty may be evaluated, S is the signal and B is the background, with σ_S and σ_B being their respective uncertainties. In Tab.3.3 and in Tab.3.4 efficiencies and background rejection are shown in percentages with their respective uncertainty.

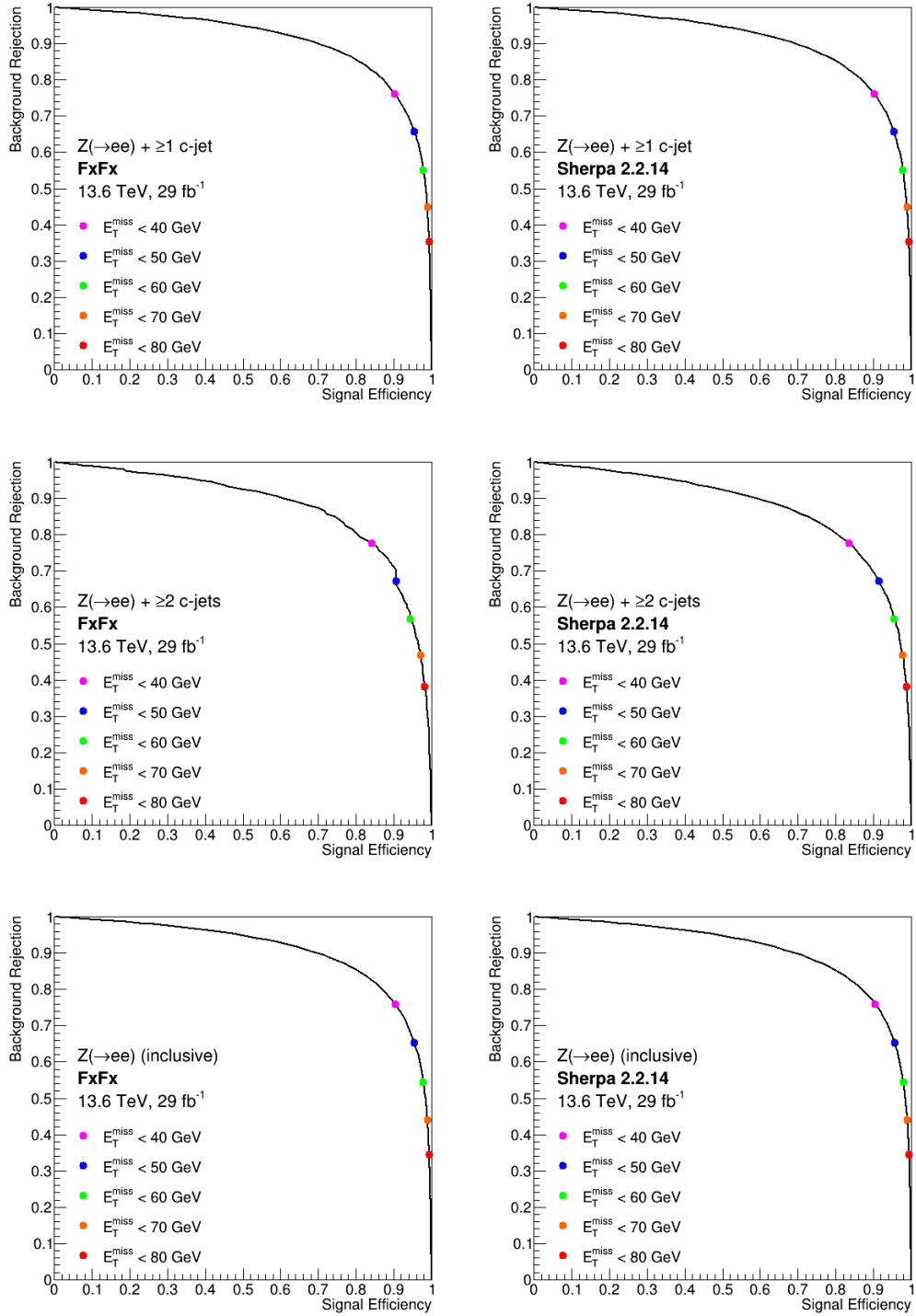
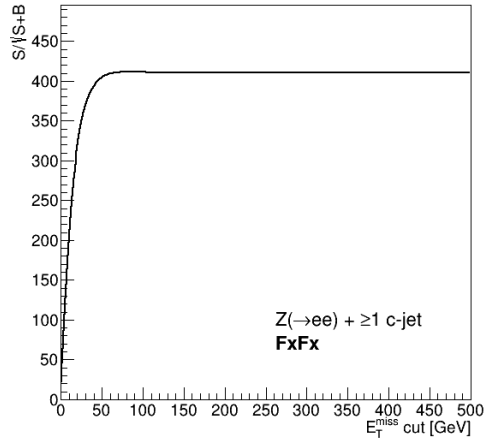
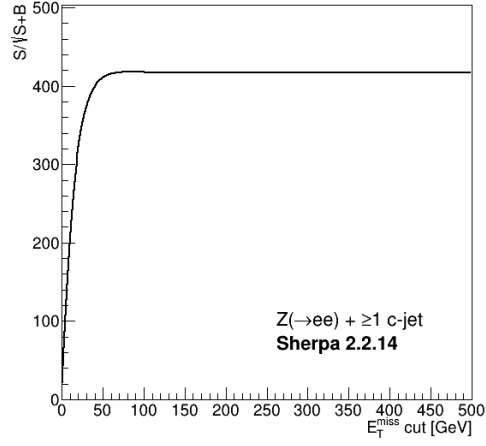


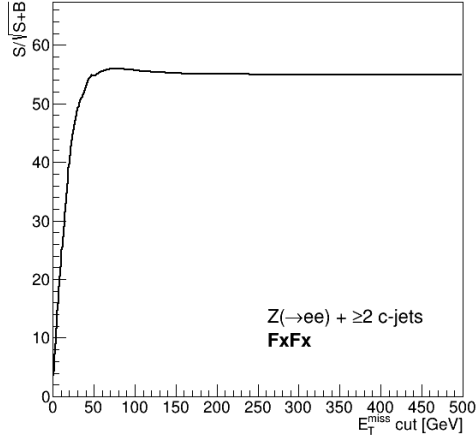
Figure 3.6: ROC empirical curves for the electron channel built by integrating the histogram content after every possible cut on each bin.



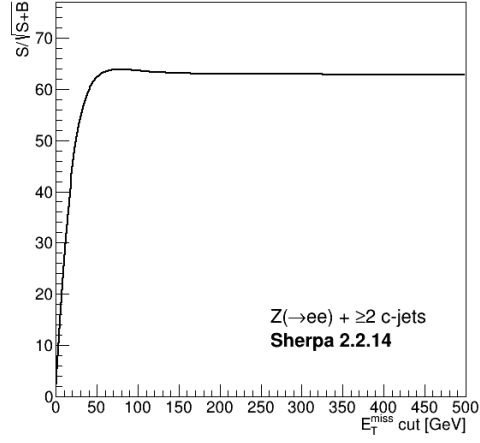
(a)



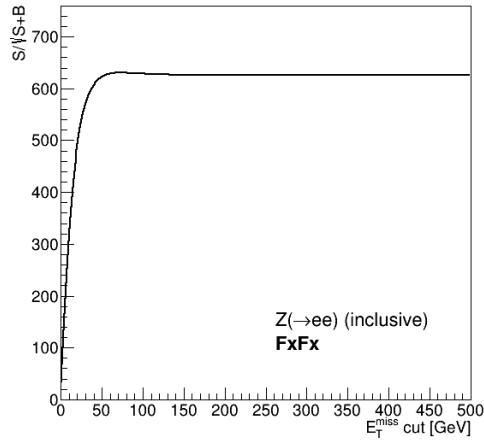
(b)



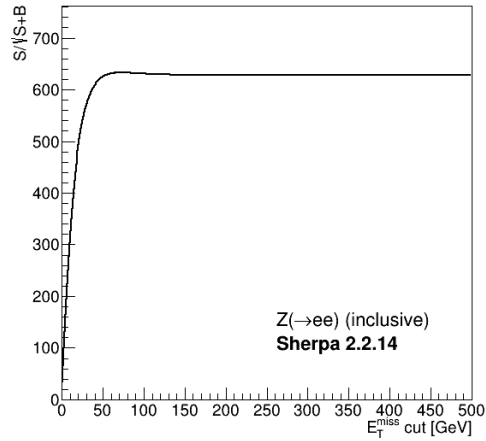
(c)



(d)



(e)



(f)

Figure 3.7: Specificity as a function of the missing transverse energy for the electron channel. As it is shown in the four plots, generated by FxFx on the left, and by Sherpa on the right, there is no maximum point, which means that the optimal cut would be to infinity.

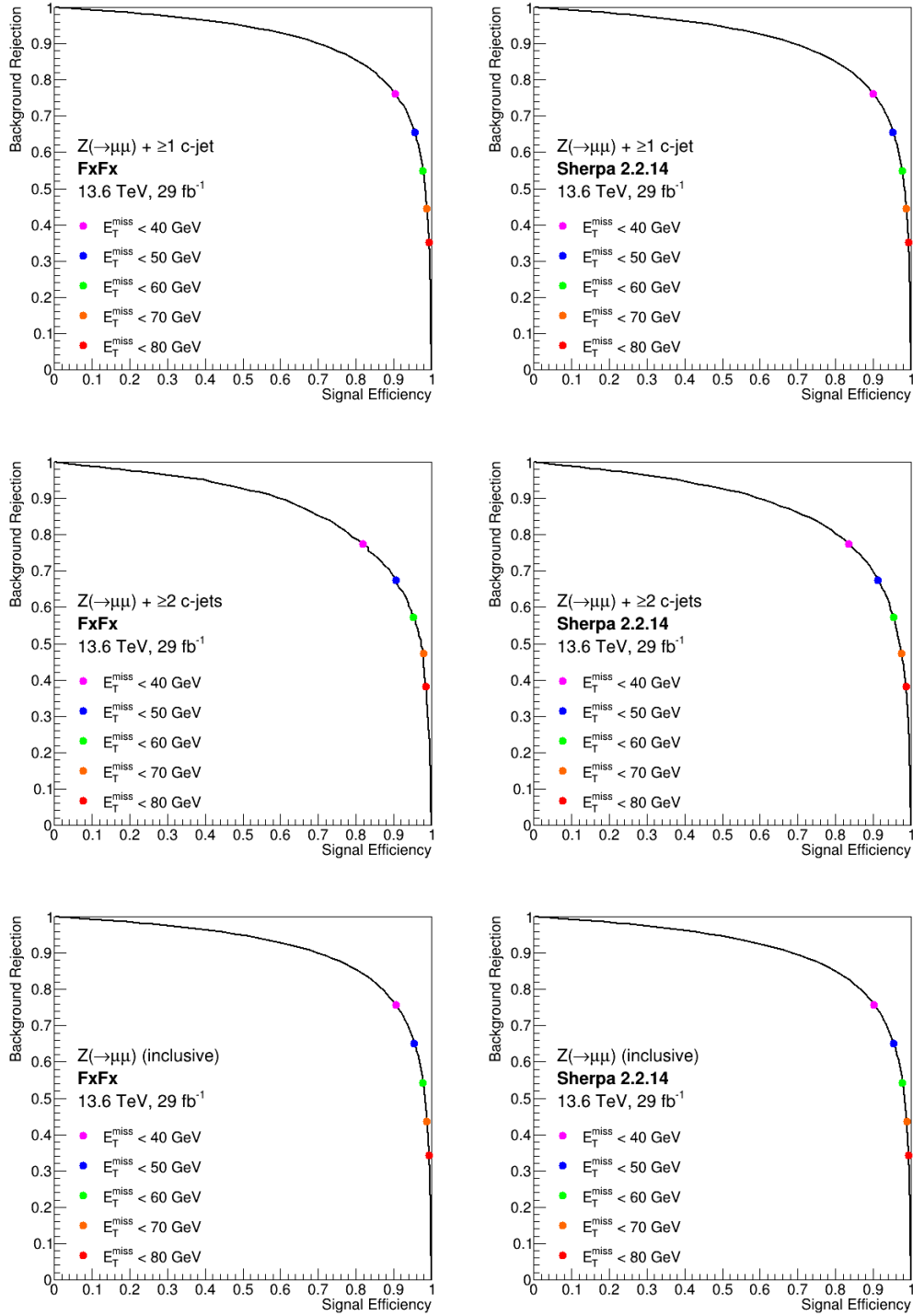


Figure 3.8: ROC empirical curves for the muon channel built by integrating the histogram content after every possible cut on each bin.

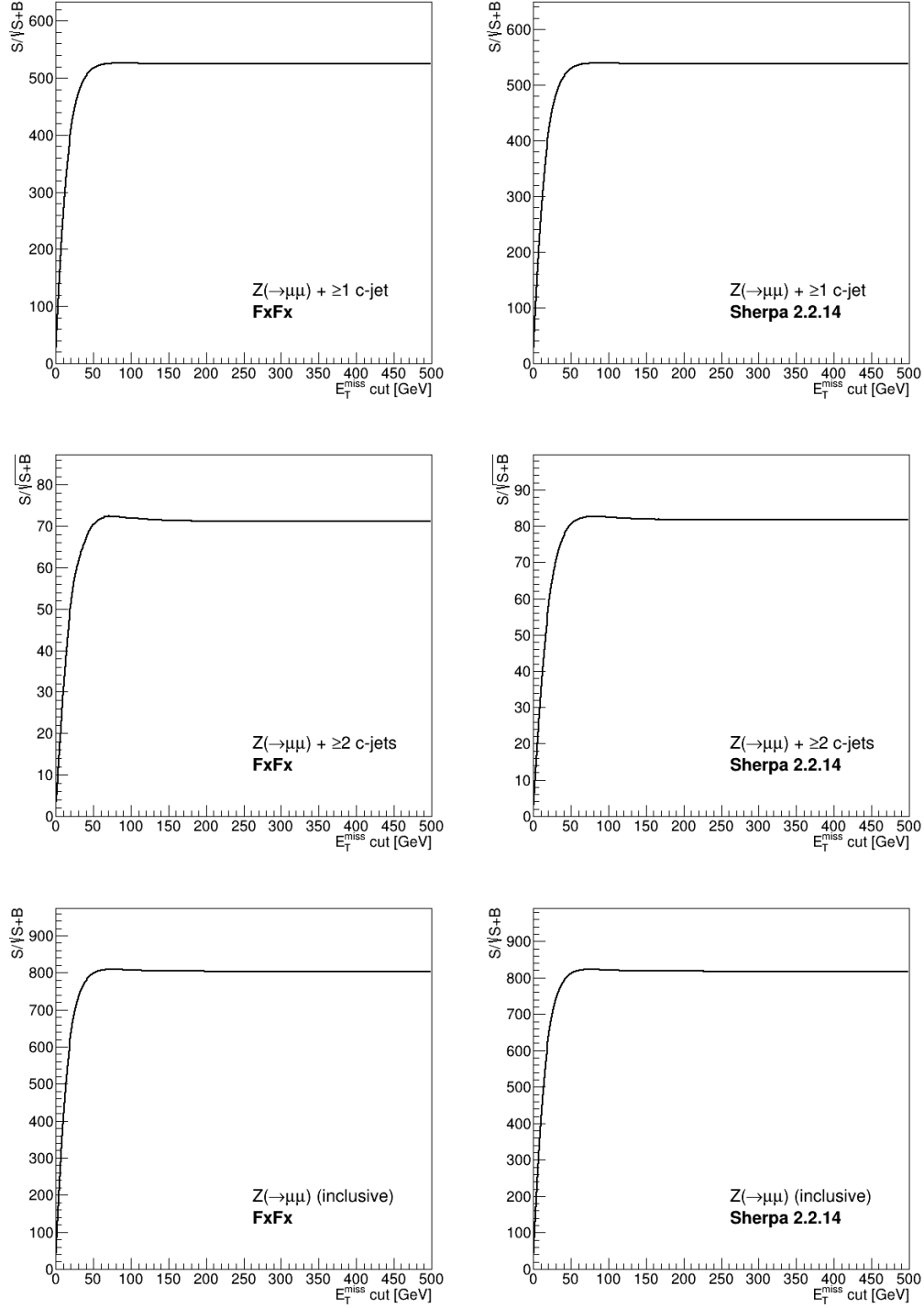


Figure 3.9: Specificity as a function of the missing transverse energy for the muon channel. As it is shown in the four plots, generated by FxFx on the left, and by Sherpa on the right, there is no maximum point, which means that the optimal cut would be to infinity.

	MG+FxFx	Sherpa	Powheg+Py8
Electron channel			
$Z+ \geq 1$ c-jet	$(174 \pm 2) \cdot 10^3$	$(179.4 \pm 0.5) \cdot 10^3$	—
$Z+ \geq 2$ c-jet	$(35 \pm 2) \cdot 10^2$	$(44.5 \pm 0.5) \cdot 10^2$	—
$Z(inclusive)$	$(416 \pm 3) \cdot 10^3$	$(419.1 \pm 0.8) \cdot 10^3$	—
$t\bar{t} \geq 1$ c-jet	—	—	$(608 \pm 1) \cdot 10$
$t\bar{t} \geq 2$ c-jet	—	—	558 ± 5
$t\bar{t}(inclusive)$	—	—	$(267.6 \pm 0.3) \cdot 10^2$
Muon channel			
$Z+ \geq 1$ c-jet	$(284 \pm 2) \cdot 10^3$	$(297.6 \pm 0.6) \cdot 10^3$	—
$Z+ \geq 2$ c-jet	$(58 \pm 2) \cdot 10^2$	$(73.9 \pm 0.7) \cdot 10^2$	—
$Z(inclusive)$	$(683 \pm 3) \cdot 10^3$	$(706 \pm 1) \cdot 10^3$	—
$t\bar{t} \geq 1$ c-jet	—	—	$(882 \pm 2) \cdot 10$
$t\bar{t} \geq 2$ c-jet	—	—	790 ± 5
$t\bar{t}(inclusive)$	—	—	$(391.9 \pm 0.4) \cdot 10^2$

Table 3.2: Entries of signal and background for the different processes for the electron and the muon channels.

Cuts (GeV)		40	50	60	70	80
Electron channel ($Z+\geq 1$ c-jets)						
MG+FxFx	ϵ_S (%)	90.20 ± 1.33	95.45 ± 1.37	97.85 ± 1.39	98.92 ± 1.40	99.41 ± 1.40
	$1 - \epsilon_B$ (%)	76.24 ± 0.13	65.75 ± 0.17	55.05 ± 0.20	44.81 ± 0.23	35.42 ± 0.25
Sherpa	ϵ_S (%)	90.24 ± 0.36	95.42 ± 0.37	97.86 ± 0.38	98.93 ± 0.38	99.46 ± 0.38
	$1 - \epsilon_B$ (%)	76.24 ± 0.13	65.75 ± 0.17	55.05 ± 0.20	44.81 ± 0.23	35.42 ± 0.25
Electron channel ($Z+\geq 2$ c-jets)						
MG+FxFx	ϵ_S (%)	84.24 ± 5.27	90.71 ± 5.58	94.41 ± 5.73	97.10 ± 5.84	98.23 ± 5.88
	$1 - \epsilon_B$ (%)	77.59 ± 0.43	67.22 ± 0.54	56.93 ± 0.64	46.79 ± 0.74	38.06 ± 0.82
Sherpa	ϵ_S (%)	83.55 ± 1.43	91.48 ± 1.53	95.55 ± 1.57	97.73 ± 1.60	98.82 ± 1.61
	$1 - \epsilon_B$ (%)	77.59 ± 0.43	67.22 ± 0.54	56.93 ± 0.64	46.79 ± 0.74	38.06 ± 0.82
Electron channel (inclusive)						
MG+FxFx	ϵ_S (%)	90.42 ± 0.80	95.42 ± 0.83	97.85 ± 0.84	98.95 ± 0.85	99.42 ± 0.85
	$1 - \epsilon_B$ (%)	75.92 ± 0.06	65.31 ± 0.08	54.42 ± 0.10	43.94 ± 0.11	34.38 ± 0.12
Sherpa	ϵ_S (%)	90.49 ± 0.25	95.58 ± 0.26	97.95 ± 0.26	98.99 ± 0.26	99.49 ± 0.27
	$1 - \epsilon_B$ (%)	75.92 ± 0.06	65.31 ± 0.08	54.42 ± 0.10	43.94 ± 0.11	34.38 ± 0.12

Table 3.3: Signal efficiency (ϵ_S) and background rejection ($1 - \epsilon_B$) percentages with associated uncertainties for different processes in the electron channel. It is worth noticing how for the selected cuts, efficiency never falls below 80%.

Cuts (GeV)		40	50	60	70	80
Muon channel ($Z+\geq 1$ c-jets)						
MG+FxFx	ϵ_S (%)	90.51 ± 0.82	95.60 ± 0.85	97.86 ± 0.86	98.89 ± 0.87	99.45 ± 0.87
	$1 - \epsilon_B$ (%)	76.10 ± 0.11	65.59 ± 0.13	54.79 ± 0.16	44.39 ± 0.18	35.05 ± 0.20
Sherpa	ϵ_S (%)	90.11 ± 0.28	95.31 ± 0.29	97.74 ± 0.29	98.88 ± 0.30	99.42 ± 0.30
	$1 - \epsilon_B$ (%)	76.10 ± 0.11	65.59 ± 0.13	54.79 ± 0.16	44.39 ± 0.18	35.05 ± 0.20
Muon channel ($Z+\geq 2$ c-jets)						
MG+FxFx	ϵ_S (%)	81.97 ± 4.03	90.60 ± 4.30	95.29 ± 4.44	97.89 ± 4.51	98.60 ± 4.5
	$1 - \epsilon_B$ (%)	77.51 ± 0.34	67.48 ± 0.43	57.29 ± 0.51	47.33 ± 0.58	38.11 ± 0.65
Sherpa	ϵ_S (%)	83.47 ± 1.17	91.36 ± 1.24	95.41 ± 1.27	97.58 ± 1.30	98.73 ± 1.31
	$1 - \epsilon_B$ (%)	77.51 ± 0.34	67.48 ± 0.43	57.29 ± 0.51	47.33 ± 0.58	38.11 ± 0.65
Muon channel (inclusive)						
MG+FxFx	ϵ_S (%)	90.59 ± 0.55	95.54 ± 0.57	97.85 ± 0.58	98.88 ± 0.58	99.43 ± 0.58
	$1 - \epsilon_B$ (%)	75.74 ± 0.05	65.10 ± 0.06	54.14 ± 0.08	43.63 ± 0.09	34.20 ± 0.10
Sherpa	ϵ_S (%)	90.32 ± 0.19	95.46 ± 0.20	97.84 ± 0.20	98.92 ± 0.20	99.44 ± 0.20
	$1 - \epsilon_B$ (%)	75.74 ± 0.05	65.10 ± 0.06	54.14 ± 0.08	43.63 ± 0.09	34.20 ± 0.10

Table 3.4: Signal efficiency (ϵ_S) and background rejection ($1 - \epsilon_B$) percentages with associated uncertainties for different processes in the muon channel. It is worth noticing how for the selected cuts, efficiency never falls below 80%.

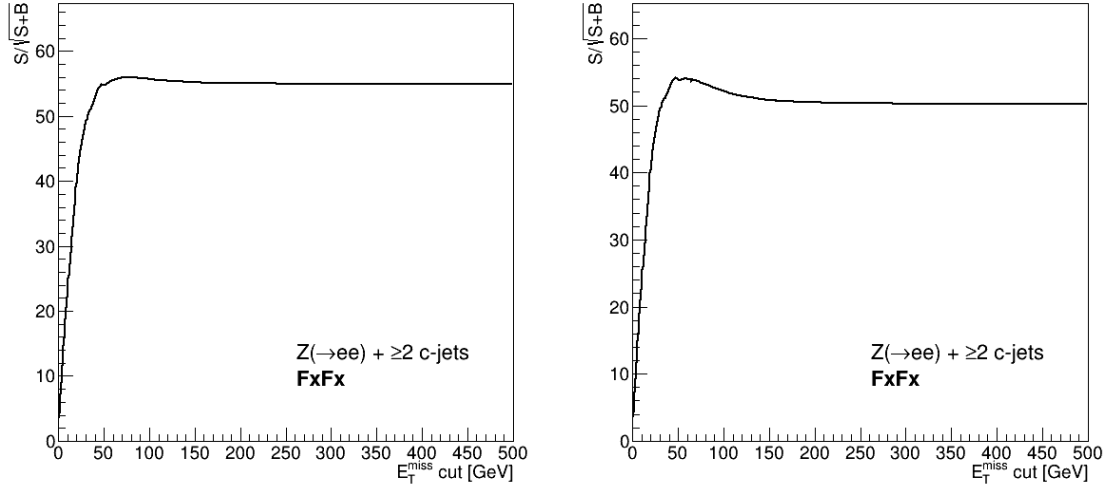


Figure 3.10: Comparison between significance function obtained from data which underwent an invariant mass cut, on the left, with the same function from data that was not cut. It is clear that in the case of the data that did not undergo cutting, there is a maximum point, around 50 GeV, signifying that in this case there exists an optimal cut value.

Conclusion

The goal of this thesis is to compare the efficiency of different cuts over the E_T^{miss} and to test whether out of all the possible cuts there were any to stand out at optimizing both the signal efficiency ϵ_S and the Background Rejection ($1 - \epsilon_B$). The analysis was carried out through the ROOT data analysis framework, using data generated by Monte Carlo generators Sherpa, MadGraph+FxPx combined with Py8, and Powheg+Py8 which simulated proton-proton collision processes at $\sqrt{s} = 13.6$ TeV, for the production of the following types of signal events: $Z(\rightarrow ee)+\geq 1$ c-jet, $Z(\rightarrow ee)+\geq 2$ c-jet, $Z(\rightarrow \mu\mu)+\geq 1$ c-jet, $Z(\rightarrow \mu\mu)+\geq 2$ c-jet, and inclusive signal for both decay channels. The signal was overlayed with the background for a first qualitative comparison, then, by looping over every possible MET cut, starting from no cut, to cutting everything on the plot, the ROC curves were built for each process. The selected cuts were at 40, 50, 60, 70, and 80 GeV, and for all of the selected cuts $\epsilon_S > 80\%$ while also $1 - \epsilon_B > 34\%$. However, following an analysis of the specificity ($S/\sqrt{S+B}$) as a function of the MET, it turned out that the optimal cut value would be to infinity, since the function soon reaches a plateau, suggesting there being no need for a cut on MET in this context. Finally, a look at plots built from predictions which did not undergo any invariant mass cut prior to analysis, shows that the mass cut does, in fact, matter in lowering the ratio between signal and $t\bar{t}$, since without it, a cut on E_T^{miss} would be suitable, while with it the cut on MET is unnecessary.

Bibliography

- [1] Zannoni Raul. *Studio delle prestazioni di algoritmi per la selezione di jet provenienti da quark top ad elevato impulso trasverso*. BSc thesis. 2016.
- [2] J.D. Lykken. “Beyond the Standard Model”. In: *CERN Yellow Report* (2010). arXiv:1005.1676. URL: <https://doi.org/10.48550/arXiv.1005.1676>.
- [3] G. Bellini et al. “Neutrino Oscillations”. In: *Advances in High Energy Physics* (2013). arXiv:1310.7858v1. URL: <https://doi.org/10.48550/arXiv.1310.7858>.
- [4] Santoro Giuseppina. “Measurements of the production of a Z boson in association with two b-jets in proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS experiment”. MA thesis. University of Calabria, 2024.
- [5] Dradi Michele. *Search for Supersymmetry in the ATLAS experiment using soft displaced tracks and a Deep Neural Network*. BSc thesis. 2021.
- [6] F. Halzen and A.D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley and Sons, 1984.
- [7] *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. Physics Letters B 716 (2012) 30-61.
- [8] Vittori Camilla. “Inclusive Z boson production and in association with b-jets in proton-proton collisions at 13 TeV with the ATLAS experiment”. PhD thesis. Alma Mater Studiorum, University of Bologna, 2018.
- [9] UA2 Collaboration (P. Bagnaia et al.) *Evidence for $Z \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider*. Physics Letters 129B (1983) 130-140.
- [10] URL: <https://cerncourier.com/a/the-w-and-z-at-lep/>.
- [11] K.A. Olive et al. *Particle Data Group*. Chin. Phys. C 38 (2014) 090001 and 2015 update. URL: <http://pdg.lbl.gov>.
- [12] C. Patrignani et al. *Particle Data Group*. Chin. Phys. C, 40, 100001 (2016).

- [13] ATLAS Collaboration, Aad G., Aakvaag E., et al. *Measurements of WH and ZH production with Higgs boson decays into bottom quarks and direct constraints on the charm Yukawa coupling in 13 TeV pp collisions with the ATLAS detector*. J. High Energ. Phys. 2025, 75 (2025).
- [14] ATLAS Collaboration. *Measurements of WH and ZH production in the $Z \rightarrow b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector*. CERN-EP-2020-087, 08 March 2021. arXiv:2007.02873v2.
- [15] CDF Collaboration. *Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$* . Phys. Rev. D 79 (2009) 052008. arXiv: 0812.4458[hep-ex].
- [16] D0 Collaboration. *Measurement of the $p\bar{p} \rightarrow W+b+X$ production cross section at $\sqrt{s} = 1.96 \text{ TeV}$* . Phys. Lett. B 718 (2013) 1314. arXiv: 1210.0627[hep-ex].
- [17] ATLAS Collaboration. *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*. JHEP 10 (2014) 141. arXiv: 1407.3643[hep-ex].
- [18] CMS Collaboration. *Measurement of the $Z/\gamma^*+b\text{jet}$ cross section in pp collisions at $\sqrt{s} = 7 \text{ TeV}$* . JHEP 06 (2012) 126. arXiv:1204.1643[hep-ex].
- [19] A. Tumasyan et al. (CMS Collaboration). V. Khachatryan A. M. Sirunyan. *Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8 \text{ TeV}$* . Eur. Phys. J. C 77, 11 (2017) pp.751.
- [20] A. Tumasyan et al. (CMS Collaboration). *Measurement of the production cross section for Z+b-jet in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* . Phys. Rev. D 105 (2022) 092014.
- [21] ATLAS Collaboration. *Measurements of the production cross section for a Z boson in association with b- or c-jets in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*. CERN-EP-2024-081. arXiv:2403.15093v1.
- [22] ATLAS Collaboration. *Measurements of the production cross section for a Z boson in association with b-jets*. JHEP 07 (2020) 044.
- [23] ATLAS Collaboration. *ATLAS b-jet identification performance*. Eur. Phys. J. C 79 (2019) 970. URL: <https://www.hepdata.net>.
- [24] CERN. *LHC Machine*. 2008 JINST 3 S08003, ISSN:1748-0221.
- [25] Lyndon Evans and Philip Bryant. *LHC Machine*. JINST, 3 S08001, 2008.
- [26] URL: <https://project-hl-lhc-industry.web.cern.ch/content/project-schedule>.

- [27] ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. 2008 JINST 3 S08003.
- [28] ATLAS Collaboration. *Insertable B-Layer Technical Design Report*. ATLAS-TDR-019.
- [29] F. Hugging. *The ATLAS Pixel Insertable B-Layer (IBL)*. arXiv:1012.2742v1.
- [30] ATLAS Collaboration. *Magnet system and magnetic field*. 2008 JINST 3 S08003.
- [31] ATLAS Collaboration. *The ATLAS Data Acquisition and High Level Trigger System*. JINST 11 (2016) P06008.
- [32] URL: <https://atlas.cern/updates/news/atlas-records-first-test-collisions-13-tev>.
- [33] ATLAS Collaboration. *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $s=13$ TeV*. CERN-EP-2016-033. arXiv:0802.1189.
- [34] ATLAS Collaboration. *Jet kinematic distributions in proton-proton collisions at $\sqrt{s} = 900$ GeV with the ATLAS detector*, ATLAS-CONF-2010-001.
- [35] G. P. Salam and G. Soyez. *The anti- k_t jet clustering algorithm*. JHEP 0804:063, 2008.
- [36] ATLAS Collaboration. *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*. ATLAS-CONF-2011-028.
- [37] ATLAS Collaboration. *Collaboration, Determination of the ATLAS jet energy measurement uncertainty using photon-jet events in proton-proton collisions at $\sqrt{s}=7$ TeV*. ATLAS-CONF-2011-03.
- [38] ATLAS Collaboration. *Probing the measurement of jet energies with the ATLAS detector using Z+jet events from proton-proton collisions at $\sqrt{s}=7$ TeV*. ATLAS-CONF-2011-053.
- [39] ATLAS Collaboration. *Tagging and suppression of pile-up jets with the ATLAS detector*. ATLAS-CONF-2014-018.
- [40] ATLAS Collaboration. *Expected performance of the ATLAS b-tagging algorithms in Run-2*. ATL-PHYS-PUB-2015-022.
- [41] ATLAS Collaboration. *Performance of b-Jet Identification in the ATLAS Experiment*. arXiv:1512.01094v2.
- [42] ATLAS Collaboration. *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*. ATL-COM-PHYS-2015-1408.