Study of the sensitivity of jet parameter distributions to the strong coupling constant α_s in $Z^0 \rightarrow e^+e^-$ events at ATLAS.

Guillaume Kirsch
Magdalen College, Oxford

Thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

University of Oxford, Trinity Term 2009

Study of the sensitivity of jet parameter distributions to the strong coupling constant α_s in $Z^0 \rightarrow e^+e^-$ events at ATLAS.

Guillaume Kirsch,
Magdalen College, Oxford

Thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy.

University of Oxford, Trinity Term 2009

Abstract

The ATLAS experiment is one of the four large detectors located at the Large Hadron Collider at CERN near Geneva, Switzerland. It is expected to start collecting data from proton-proton collisions up to 14 TeV centre-of-mass energy. ATLAS is a general-purpose detector aimed at gaining a better understanding of current theories of particle physics as well as the potential discovery of new physics beyond the Standard Model.

Due to the large hadronic content in each interaction, leptons are expected to be the primary signature to look for in many processes, hence the importance of reconstructing them efficiently. This thesis presents an algorithm to measure the efficiency of the electron reconstruction software and the expected performance from data. The results show our capacity to evaluate the efficiency at different integrated luminosities over a large kinematic range under different background conditions.

We also validate the development of an ATLAS specific simulation framework aimed at fast event processing. In order to achieve this, this thesis compares outputs of the different simulation frameworks, both a full Geant4 simulation and the ATLAS specific fast simulation ATLFASTII. The study compares the distributions in a standard $Z^0 \rightarrow e^+e^-$ sample and assesses the consistency of different background processes with the event selection cuts.

Finally, we use locally produced fast simulation samples to assess the sensitivity to modifications of the strong coupling constant on the number of $Z^0 \rightarrow e^+e^-$ events associated with additional jets. This sensitivity is compared to the expected experimental uncertainties, such as the choice of jet algorithms, jet energy scales or theoretical uncertainties such as the parton distribution function uncertainty.

This thesis was typeset with LaTeX.

Published at the University of Oxford, Oxford, United Kingdom.

À mes grands-mères,
À ma famille

To my grand-mothers,
To my family

“The fundamental principle of science, the definition almost, is this: the sole test of the validity of any idea is experiment.”

Richard P. Feynman

Acknowledgements

The last three years have been both intense and stimulating. Many people and organisation are to be thanked for providing teaching, stability and entertainment. A nominal list of friends would be too long, so where I thank them collectively, I shall hope that they understand how important their contribution has been and I shall make sure my thanks reach them individually.

To start with, I want to convey my thanks where it is due, to the researchers and staff of the Particle Physics sub-department at the University of Oxford. This thesis would not have been possible without each and every single one of them. Here, I would like to thank in particular my fellow students in the ATLAS group, whose help, ideas and patience I have abused during all this time. In particular, I would like to thank Ellie Dobson, Kristin Lohwasser, and Maria Fiascaris. These thanks also extend to the students from the other groups, whose outsider point-of-view has been a refreshing source of ideas on a regular basis. Here I would like to record the contributions of Erik Devetak, Gregory Moss, John Dale, and Matthew Warden for the heated discussions , be it around a meal, a cup of tea or a pint. The academics are of course not to be forgotten. I would like to thank my supervisor, Dr. Jeff Tseng, as well as Dr Tony Weidberg, who despite already having students of his own, has spared me enough of his time to guide me through the pitfalls of my research. Finally, apart from the general thanks I would like to thank the people with who I have shared an office during the last three years and I do hope I have not annoyed them unduly. In particular, I would like to thank Ian Preston for proof-reading this thesis. My thanks further extend to the many people I have collaborated with in the ATLAS collaboration at large, be they fellow students in other institutions, post-docs or group conveners.

It is also fair to thank the many interesting students I have met in Oxford, outside of the Physics department. Whether in Magdalen College or in the Oxford University Athletics Club, the persons I have met have made my stay in Oxford a memorable one and have have fulfilled my expectations of meeting an interesting, intelligent and cosmopolitan society by attending one of the leading English speaking universities. British, American, Australian, New Zealander, European or other, they have all helped me widen my view of the world and my address book now reads like an atlas, pun intended.

I would also like to thank the people I met during my stay at CERN , on both sides of the border. The people working at CERN on the one hand, for the stimulating insights into the physics of high energies, the non physicists of the Ain-Est Athletisme club on the other hand for their light heartedness after a hard day of research.

My thanks also extends to the sponsors of my doctorate. First of all The MathWorks for believing in my capacities and for their kind welcome. They made my stays in Cambridge memorable and without their financial support, my study at Oxford would probably have been in jeopardy. The same has to be said about the Ecole Polytechnique, without which

I would probably never have formed this project and Magdalen College which provided so much more than just money, by allowing me to study among friends in such beautiful surroundings.

Last, but by no means least I would like to thank my family. Despite not always understanding what I was doing, or why I was doing it, they have supported me without fault for afar. I may not have gone back to France very often due to the highly time-consuming nature of research, but I have never lost contact with any of them. Because English is not their cup of tea for most of them, the next section will cater for them more specifically.

Remerciements

Les trois dernières années passées à Oxford et à Genève ont été riches en enseignements et en rebondissements. J'ai eu la chance d'y rencontrer des personnes passionnantes et de participer à des clubs très amicaux. La liste des personnes à remercier serait trop longue à établir nominativement, aussi j'espère que les personnes concernées se reconnaîtront et je prendrai soin d'apporter mes remerciements en personne où ils sont dus.

Il est naturel de débuter cette liste par les chercheurs et membres du personnels du département de physique des particules. Cette thèse n'aurait jamais vu le jour sans eux et je les remercie du fond du coeur pour leur patience, leurs conseils et leur soutien. En particulier, je tiens à souligner la contribution à la réussite de cette thèse des autres étudiants du département, qu'ils travaillent sur ATLAS ou non. Leur apport en terme de remarques scientifiques, amitié et détente a énormément contribué à soutenir ma motivation et bonne humeur face aux difficultés. Je tiens également à remercier les membres du groupe ATLAS pour leurs idées, mon directeur de thèse, mais surtout je tiens à faire une mention particulière également pour le Dr. Tony Weidberg et pour le temps qu'il m'a consacré malgré son emploi du temps chargé. Bien sûr, mes remerciements ne s'arrłtent pas à la porte de l'université d'Oxford et s'étendent à tous les membres d'ATLAS avec lesquels j'ai collaboré.

Il n'est que juste d'inclure dans ces remerciements les fascinants étudiants que j'ai rencontrés à Oxford en dehors du département de Physique. Que ce soit au Magdalen College ou au club d'athlétisme de l'université d'Oxford, j'ai rencontré des gens intelligents, intéressants et cosmopolites qui ont rempli toutes les espérances fondées en partant pour Oxford. Qu'ils soient anglais, américains, australiens, néo-zélandais, européens ou autre, ils ont élargi mon opinion du monde et mon carnet d'adresse se lit maintenant comme un atlas.

Je tiens également à remercier les personnes qui ont contribué à faire de mon séjour au CERN un moment agréable. Bien qu'ils ne liront probablement jamais ces quelques lignes, je tiens à remercier les membres de l'Ain-Est Athlétisme pour leur camaraderie qui était très rafraîchissante après une longue journée de travail.

Je tiens également à remercier les sponsors de cette aventure. L'école Polytechnique tout d'abord, sans laquelle je n'aurais sans doute jamais eu cette idée folle de continuer l'école. Le Magdalen College ensuite, qui m'a fourni, non seulement une bourse, mais également un cadre idyllique et m'a permis de rencontrer des gens passionnants. Enfin, The Mathworks, pour avoir cru en moi et avoir fait de mes étés à Cambridge des interludes intéressants et stimulants.

Enfin, cette liste ne serait pas complète sans un remerciement appuyé à ma famille. Malgré la distance qui nous sépare, et la difficulté de mon sujet d'étude, ils m'ont toujours soutenu sans conditions ce qui a rendu mes choix d'autant plus faciles. Je ne suis sans doute pas allé les voir assez souvent à cause du peu de temps libre que me laissait mes recherches,

mais j’espère que ce court effort de traduction leur fera comprendre à quel point il était important pour moi qu’ils soient toujours présents à mon esprit.

Preface

A few weeks before ATLAS and the LHC start collecting their first data, we have reached a crucial time to prepare for fast and smooth studies of the current hot physics topics. It is necessary to check we have all the tools ready as far as possible for a timely study of the detector performances.

This thesis describes two efforts to prepare for early ATLAS physics analyses as well as a preliminary study of feasibility for one physics study.

After a brief introduction presenting some theories of particle physics and the main features of the components of the ATLAS experiment (Chapter 1), a first part is devoted to the measurement of the electron reconstruction efficiency in Chapters 2 to 5:

- Chapter 2 describes in more detail the hardware and software configurations used for electron reconstruction at ATLAS, as well as the selection procedure to get as clean a sample of electrons as possible.

- Chapter 3 describes the mathematical framework which explains how to extract the efficiency under study using the parameters we collected in Chapter 2.

- Chapter 4 presents the backgrounds considered to our $Z^0 \to e^+e^-$signal process. It presents a method to statistically separate the background events from the signal ones and how to extract a systematic error associated with this procedure. It also presents a special treatment for QCD-based backgrounds because of the lack of available statistics.

- Chapter 5 finally summarises the results we obtained in two ATLAS organised exercises using this method. It demonstrates our capacity to perform the measurement under different background assumptions.

In a second study, we validate, in Chapter 6, the electron reconstruction in the ATLAS fast simulation framework. We verify the accuracy of the electron distributions and the difference of a benchmark selection on a selection of process between normal full simulation and the fast simulation framework ATLFASTII.

Finally, in Chapter 7, we use the results we presented in the previous chapters to study the sensitivity of distribution ratios to the value of the strong coupling constant $\alpha_s(M_{Z^0})$.

Chapter 8 summarizes the results presented throughout this thesis.

Appendices A, B describe the datasets used in Chapters 5 and 6 respectively to get the results presented in these chapters.

Appendix C finally presents the configuration files used to generate the datasets used for the analysis in Chapter 7.

Contents

List of Figures

List of Tables

Glossary

Athena Name of the ATLAS software framework

ATLAS A Toroidal LHC ApparatuS, general purpose high-energy physics detector built as one of the four LHC experiments

ATLFASTII ATLAS fast(er) reconstruction framework, based on a full simulation of the inner detector and a fast simulation of the calorimeters

CERN European Organisation for Nuclear Research, located near Geneva, Switzerland

CSC Computing Software Challenge, name of one the of ATLAS computing exercises

egamma Name given to an electromagnetic object candidate, either electron or photon

ElectronLoose Standard set of identification cuts, characterized by a high efficiency, but low rejection against jets.

ElectronMedium Standard set of identification cuts, characterized by a medium efficiency and medium rejection against jets. These cuts are a superset of ElectronLoose.

ElectronTight Standard set of identification cuts, characterized by a medium efficiency and very high rejection against jets. These cuts are a superset of ElectronMedium.

$\mathbf{E_T}$ Transverse energy of a particle in the *x*–*y* plane

η Pseudorapidity, defined as $\eta = -\ln\tan(\theta/2)$ with θ the polar angle

$\mathbf{E_T^{miss}}$ Transverse missing energy of an event in the *x*–*y* plane

FDR Full Dress Rehearsal, name of one the of ATLAS computing exercises

IP Interaction Point. Point of collision between two protons resulting in a hard interaction

IsEM Is ElectroMagnetic, software bitmask recording which identification cuts have been passed by an egamma object candidate

JES Jet Energy Scale. Estimated error between the measurement of a Jet energy and its true energy

LEP Large Electron Positron Collider, a 27 km long electron–positron collider stopped in 2001. The LHC is using the same tunnel that used to host LEP.

LHC Large Hadron Collider, a 27 km long proton–proton collider with a design center of mass energy of $\sqrt{s} = 14$ TeV and luminosity of $\mathscr{L} = 10^{34}\ cm^{-2}s^{-1}$

MSSM Minimum Supersymmetric Standard Model

parton single particle (quark or gluon), constituent of a hadron, used by the generators to compute hard-processes. Either in- or out-going.

$\mathbf{p_T}$ Transverse momentum of a particle in the x–y plane

QCD Quantum ChromoDynamic, theory of strong interactions based on the colour exchange between quarks through the mediation of gluons

RoI Region of Interest. Detector region of interest for potential trigger object reconstruction

SCT Silicon Strip detector, one of the ATLAS inner detector sub–systems

SM Standard Model of Particle Physics

SUSY Supersymmetry

TeVatron a 6.3 km long proton–antiproton collider with a center of mass energy of $\sqrt{s} = 1.96$ TeV and luminosity of $\mathscr{L} = 3 \cdot 10^{32}\ cm^{-2}s^{-1}$

track Reconstructed path of a charged particle as it progresses through the detector

TRT Transition Radiation Tracker, a straw–based drift chamber. One of the ATLAS inner detector sub–systems

Chapter 1

Introduction

This chapter presents an overall introduction to the most current theories in high-energy physics. It describes a brief history of the discoveries over the last decades and an introduction to possible extensions beyond the currently accepted results. It also describes the main features of the LHC and the ATLAS detector that will be referred to later in this thesis.

Our understanding of matter and energy is described in terms of the interactions between elementary particles. Throughout the centuries, the goal of physicists has been to reduce the number of laws governing the behaviour of all know forms of matter to a minimal set of fundamental equations and theories. The aim is to identify a single theory that unifies all theories into a theory of everything from which all other known laws are derived, either directly or from approximations.

1.1 Theories in Particle Physics

Over the last 112 years since the discovery of the electron by J.J Thomson in 1897 [1], the particle physics community has developed a relatively universal and well-accepted description of what are considered elementary particles. How to extend the model beyond what is already well established by past experiments is still under debate. It is hoped by a large part of the community that the results we will collect from the Large Hadron Collider will help us strengthen our current theories and pave the way to new discoveries.

1.1.1 The Standard Model of Particle Physics

The Standard Model (SM) [2, 3] is currently the most widely accepted model describing the fundamental particles and their interactions. It is a theory describing three of the four known fundamental interactions. It was initiated in 1963 by Sheldon Glashow's discovery of a way to combine the electromagnetic and weak interactions [4], and took its modern form in 1967 after Weinberg and Salam incorporated the Higgs mechanism into Glashow's Electroweak theory [2]. The theory of strong interactions only acquired its modern form around 1973-1974 and was completely integrated to the Standard Model after the discovery of the gluon at the Deutsches Elektronen Synchrotron in 1979 [5, 6].

Over the last three decades, almost all experimental tests of the forces described in the Standard Model have agreed with its predictions. The ability of the SM to successfully predict physical phenomena was first observed with the discovery of neutral current weak interactions in 1973 [7]. The observation of the intermediate weak bosons, $W^{\pm}$ and Z^0, at the European Organization for Nuclear Research (CERN) in 1983 further increased the confidence in that model as these particle exhibited the expected properties [8]. The electroweak parameters and quantitative predictions from Quantum Chromodynamics (QCD) were subsequently tested with high precision at the Large Electron Positron collider (LEP) [9, 10], the Stanford Linear Collider (SLC) [11], the TeVatron [12] and HERA [13]. Accurate measurements at the Z^0 resonance energies in electron-positron collisions showed no evidence for deviations from the SM predictions. The $W^{\pm}$ boson and top quark mass measurements performed in electron-positron collisions at LEP II (for the W [14]) and in proton-antiproton collisions at the TeVatron [15] are also consistent with the SM expectations.

The theory is presented as a spontaneously broken gauge quantum field theory. The most fundamental principles of nature are explained through particles and their interactions where the interactions themselves are also considered as the exchange of virtual particles. Particles or groups of particles are represented as quantum fields, i.e. n-dimensional complex fields, noted by an operator $\psi_i(x)$, where the norm of the field is the probability of finding that particle at the coordinate x. Since the norm is the only measurable quantity, the phase of the field is left as a free parameter. To comply with special relativity, it is necessary to be able to assign the phase θ arbitrarily at any point in the phase-space without affecting the physical state of a particle. Assuming this, all measurements in an Abelian theory should be invariant under the transformation

$$\psi(x) \rightarrow e^{i\theta(x)}\psi(x).$$

This is the basis of local gauge invariance in the SM and any possible extension. In the SM, electroweak interactions result from a local $SU(2)_L \times U(1)_Y$ symmetry which is spontaneously broken via the Higgs mechanism [16] and leads to the existence of four electroweak bosons, γ, $W^{\pm}$ and Z^0, the last three being massive. Strong interactions are mediated by eight gluons which arise from a local $SU(3)_C$ symmetry.

Matter fields representing fermions are organized into families, with left-handed[1] and right-handed[2] particles transforming respectively as weak isodoublets and weak isosinglets. Only the left-handed component of particles and the right-handed component of their anti-particles participate in charged weak interactions in the SM, which is known as parity violation. Masses of particles are introduced by their coupling to a scalar field known as the Higgs field. However, despite intensive searches at LEP and TeVatron, the Higgs particle has not yet been discovered. It remains as the only missing particle in the SM.

Despite the very good agreement of SM predictions with measurements in the last decades, there is some evidence that the minimal version with 19 free parameters used at the moment, does not describe nature sufficiently. There is, for example, evidence that neutrinos have non-zero masses, which the minimal SM does not allow. However, the SM can be extended to be consistent with these results by adding ten more free parameters which are not predicted by any theory and thus, need to be determined experimentally. Similarly,

[1]The projection of the spin onto the direction of movement is anti-parallel to the momentum in the limit of massless particles

[2]The projection of the spin onto the direction of movement is parallel to the momentum in the limit of massless particles

Standard Model Particles			
Family	Interaction eigenstates	Mass eigenstates	Spin
Leptons	$e^-_{L/R}, \mu^-_{L/R}, \tau^-_{L/R}$	e^-, μ^-, τ^-	1/2
	$\nu_{eL}, \nu_{\mu_L}, \nu_{\tau L}$	ν_e, ν_μ, ν_τ	1/2
Quarks	$u'_{L/R}, c'_{L/R}, t'_{L/R}$	u, c, t	1/2
	$d'_{L/R}, s'_{L/R}, b'_{L/R}$	d, s, b	1/2
Bosons	B^0, W^0	γ, Z^0	1
	H^0	H^0	0
	$W^\pm$	$W^\pm$	1
	g	g	1

Table 1.1: Particle spectrum of the SM. Labels “L” and “R” denote left- and right-handed electroweak eigenstates.

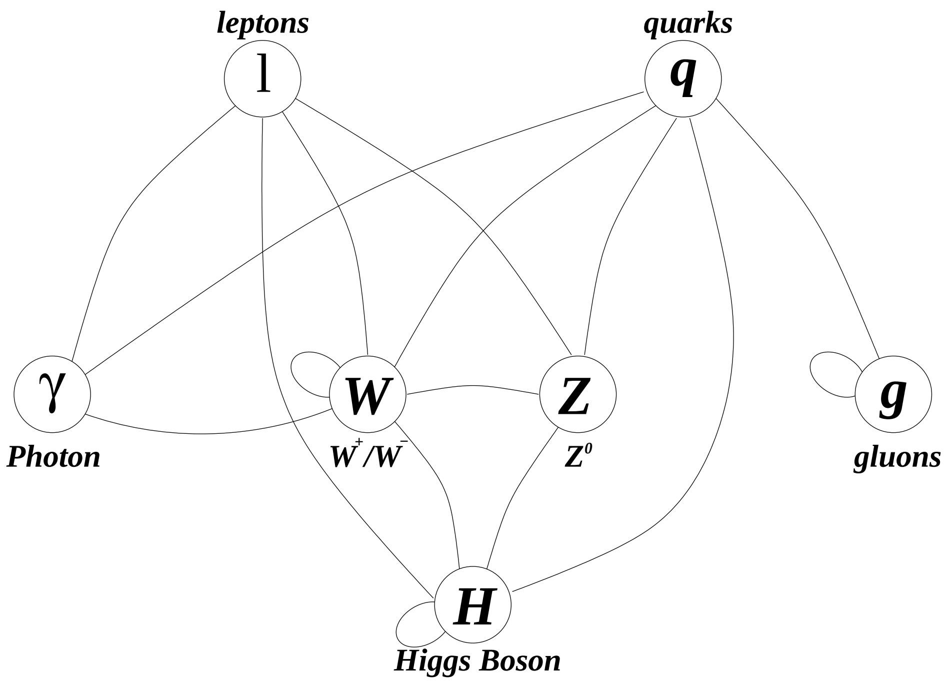


Figure 1.1: Summary of the interactions between the particles described by the Standard Model.

measurements of the anomalous magnetic moment of the muon at Brookhaven [17], and the measurement of the cold matter density in our universe by WMAP [18], give a hint that physics even beyond the extended SM should exist. Astrophysics observations make it clear for example that either our understanding of gravity based on Einstein's theory of General Relativity is wrong, or that particles forming cold dark matter, which have so far escaped our detection, must exist. Similarly, the Axion, a pseudo-scalar particle supposed to explain the non-existence of CP-violation in the strong interaction, has not be discovered yet. It could theoretically be an ingredient of such cold dark matter, but may not explain the measured dark matter density sufficiently [19] and is not technically part of the SM. Other SM shortcomings include the matter-antimatter asymmetry observed in the universe which is not explained at the moment either.

There also are theoretical motivations for physics beyond the SM. One such example is the hierarchy problem concerning the quadratically divergent fermion loop corrections to the Higgs boson mass. New physics is required to happen at the TeV scale to constrain the Higgs mass in the area of a few hundred GeV and thus make the SM consistent with recent W and top mass measurements. Additionally for a reason of symmetry, theorists are aiming for the unification of the gauge couplings. Repeatedly, history has shown that a better understanding of nature leads to a unification of theories.

1.1.2 Neutral weak boson production

This thesis revolves around the production of the neutral weak boson Z^0. At the LHC, this happens primaraly by a process call Drell-Yan process where a quark-antiquark annihilates into a Z^0 boson. The study of the Z boson is very important at the LHC as it is produced with a large cross-section relative to other processes. We expect around 1400 pb^{-1} at 14 TeV for the $Z^0 \rightarrow e^+e^-$ channel alone. It also has a very clean signature which make it a useful tool for calibrations and performance studies.

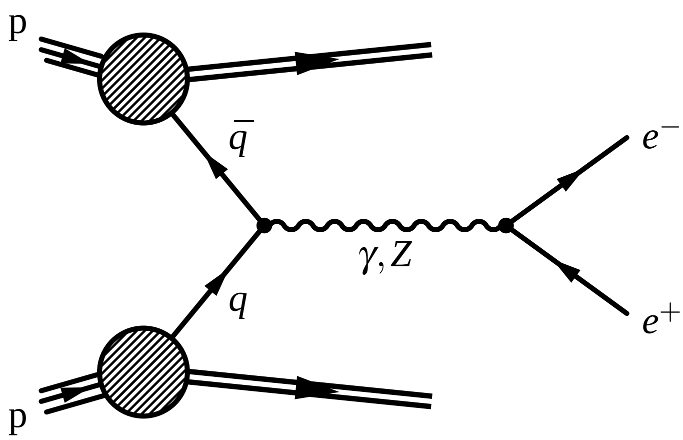


Figure 1.2: Feynman diagram of a simple $Z^0 \rightarrow e^+e^-$ event.

The main features of the $Z^0 \rightarrow e^+e^-$ process are :

- large cross-section, higher than most of its backgrounds and will provide us with large statistics;
- no intrinsic missing transverse energy;
- two high p_T leptons;
- isolated lepton trigger signature (isolated lepton);

1.1.3 Supersymmetry

Many theoretical aspects justify the addition of an extra type of symmetry beyond the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetries of the SM fields. Since the discovery of the Pauli Principle, particles have been divided into two groups, fermions with half-integer spins of the form $\frac{2n+1}{2}$ and bosons with integer spins. The additional symmetry requires new particles to be added to form multiplets, so that a supersymmetric operation on the representation of a particle belonging to one group transforms it into a member of the other group. Thus the spin is modified but all other properties (mass, couplings) remain the same. Hence, the electron with spin $\frac{1}{2}$ and a mass of 0.511 MeV would have a bosonic superpartner,

Minimum Supersymmetric Standard Model Particles					
Interaction	Mass	Spin	Interaction	Mass	Spin
$e^-_{L/R}, \mu^-_{L/R}, \tau^-_{L/R}$	e^-, μ^-, τ^-	1/2	$\tilde{e}^-_{L/R}, \tilde{\mu}^-_{L/R}, \tilde{\tau}^-_{L/R}$	$\tilde{e}^-_{1/2}, \tilde{\mu}^-_{1/2}, \tilde{\tau}^-_{1/2}$	0
$\nu_{eL}, \nu_{\mu_L}, \nu_{\tau L}$	ν_e, ν_μ, ν_τ	1/2	$\tilde{\nu}_{eL}, \tilde{\nu}_{\mu L}, \tilde{\nu}_{\tau L}$	$\tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau$	0
$u'_{L/R}, c'_{L/R}, t'_{L/R}$	u, c, t	1/2	$\tilde{u}'_{L/R}, \tilde{c}'_{L/R}, \tilde{t}'_{L/R}$	$\tilde{u}_{1/2}, \tilde{c}_{1/2}, \tilde{t}_{1/2}$	0
$d'_{L/R}, s'_{L/R}, b'_{L/R}$	d, s, b	1/2	$\tilde{d}'_{L/R}, \tilde{s}'_{L/R}, \tilde{b}'_{L/R}$	$\tilde{d}_{1/2}, \tilde{s}_{1/2}, \tilde{b}_{1/2}$	0
B^0, W^0	γ, Z^0	1	$\tilde{B}^0, \tilde{W}^0$	$\tilde{\chi}^0_1, \tilde{\chi}^0_2, \tilde{\chi}^0_3, \tilde{\chi}^0_4,$	1/2
H^0_u, H^0_d	$h, H, A^0, A^\pm$	0	$\tilde{H}^0_u, \tilde{H}^0_d$		1/2
$W^\pm$	$W^\pm$	1	$\tilde{W}^\pm$	$\tilde{\chi}^\pm_1, \tilde{\chi}^\pm_2$	1/2
$H^\pm$	$H^\pm$	1	$\tilde{H}^\pm$		1/2
g	g	1	$\tilde{g}$	$\tilde{g}$	1/2

Table 1.2: Particle spectrum of the MSSM. Labels "L" and "R" denote left- and right-handed electroweak eigenstates., The extended SM has an additional Higgs doublet. Mass eigenstates $x_{1/2}$ are defined as $x_{1/2} = x_{L/R} \cdot \cos\theta \pm x_{R/L} \cdot \sin\theta$. Since θ is close to zero for electrons, muons, up, down, charm and strange quarks, their mass eigenstates are mostly denoted with "L/R" instead of "1/2" in literature.

in most theories with spin 0, and the same mass. However, since no supersymmetric partner has been discovered to date, it is certain that, if Supersymmetry exists at a high energy scale, it is broken at the scales accessible for experiments at the moment.

The simplest possible extension of the SM with a minimal particle content is called the Minimal Supersymmetric Standard Model (MSSM) [20]. The MSSM presents an additional Higgs doublet among its "standard" particles, describes more than twice as many particles as the SM and solves some of the open theoretical questions about the SM.

One of the main motivations for supersymmetry comes from the quadratically divergent contributions to the squared Higgs mass. The quantum mechanical interactions of the Higgs boson including the fermion loops at one-loop-level cause an unphysical quadratic divergence of the Higgs mass and unless there is an effective cancellation, the natural size of the Higgs mass is the highest scale possible. Supersymmetry reduces the size of the quantum corrections by having automatic cancellations between fermionic and bosonic Higgs interactions. If supersymmetry is restored at the weak scale, then the Higgs mass is related to supersymmetry breaking, which can be induced from small non-perturbative effects, explaining the vastly different scales between the weak interactions and gravitational interactions.

In many supersymmetric Standard Models, there is a stable heavy particle (such as the lightest neutralino) which could serve as a WIMP (weakly interacting massive particle) dark matter candidate. The existence of a supersymmetric dark matter candidate is closely tied to the stability of the lightest supersymmetric particle.

One other theoretical incentive for the existence of supersymmetry existing at the weak scale is gauge coupling unification. The evolution of the value of the three gauge coupling constants of the Standard Model with the energy scale is sensitive to the particle content of the theory. These coupling constants do not quite meet together at a common energy scale if we run them using the Standard Model. With the addition of supersymmetry, such a match is possible, which would help reduce the number of free parameters in the model. The MSSM provides the possibility of gauge coupling unification at a scale of about 10^{16} GeV [20].

However, the MSSM also leaves open some important questions. It does not give any hint as to the source of CP-violation nor is it able to make predictions for any particle mass. Furthermore, the MSSM still is a theory without quantum gravity.

In order to confirm the predictions of the MSSM, it will first be necessary to find some of the new particles given in Table 1.2, and measure their mass and interactions. In the MSSM, each superpartner has the same couplings as its SM partner, hence, in a second step, the interactions, couplings and the spin of the new particles would have to be determined.

1.2 Experimentally probing Nature at the highest energies

To study the most fundamental components of matter and their interactions, it is necessary to study them at always decreasing length scales. That is equivalent to studying them at always increasing energies. During the past century, technologies have been developed to accelerate and collide charged particles at energies which unveil nature down to a fraction of a femtometre. Such energies allow us to reproduce the effects of a temperature of 10^{15}

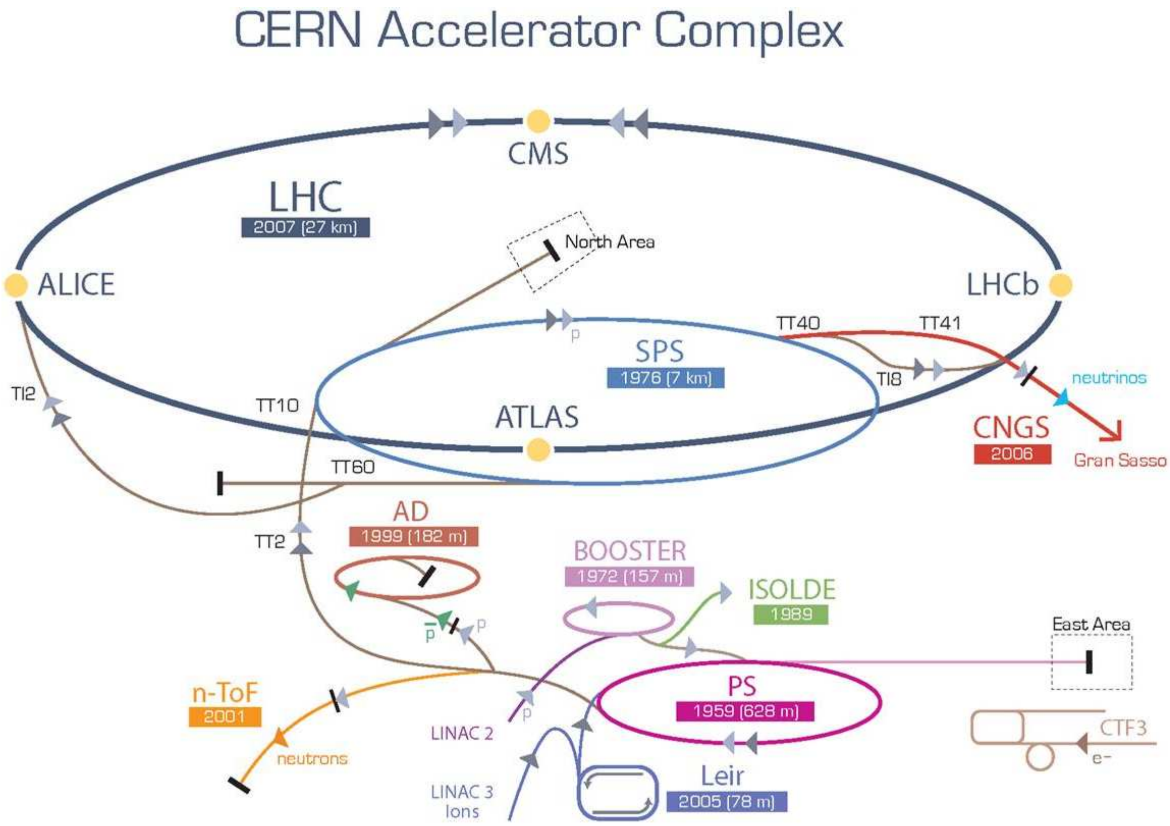


Figure 1.3: Diagram of the LHC accelerator among the CERN accelerator complex

Kelvins (in lead ion collisions), which existed a fraction of a nano-second after the Big-Bang. Experiments of that scale are very expensive, and therefore, only possible thanks to large international funding and collaboration. In addition to the contribution to fundamental knowledge about particles we get from their results, these experiments have a far-reaching impact on technical developments which is then contributed back to society.

1.2.1 The Large Hadron Collider

The construction of the Large Hadron Collider (LHC) [26] was approved by the CERN's Council in December 1994. The LHC has been installed in a 27 kilometre circumference tunnel which is about 100 metres underground and was used before that by the Large Electron Positron collider (LEP) until 2001. When it reaches it design performance, the LHC will collide protons at a center-of-mass energy $\sqrt{s}$ of up to 14 TeV, the highest ever reached in the history of human-made particle collisions. The design luminosity $\mathscr{L}$ is 10^{34} $cm^{-2}s^{-1}$. Furthermore, lead-ion collisions are planned with a total $\sqrt{s}$ of 1148 TeV and a luminosity of 10^{27} $cm^{-2}s^{-1}$.

The experimental program of the LHC will include the dedicated heavy-ion detector ALICE (A Large Ion Collider Experiment), the specialised B-physics spectrometer LHCb and the general purpose detectors CMS (Compact Muon Solenoid) and ATLAS (Section 1.2.2).

The accelerator consists of two interleaved synchrotron rings, whose main elements are 1232 superconducting NbTi dipole and 392 quadrupole magnets operating in superfluid helium at a temperature of 1.9 K. The dipole magnets guide the particles along the ring while the quadrupole magnets keep the particle bunches focussed. The nominal field for LHC magnets to handle 7 TeV beams is 8.34 Tesla, and the goal is to achieve 9 Tesla. Two 450 GeV beams produced in the Super Proton Synchrotron (SPS) will be injected into the LHC, and then accelerated by superconducting cavities, 8 cavities per ring, operating at 2 MV and 400 MHz.

It is certain that the LHC is going to provide an exciting insight into the most fundamental laws of nature. There are several possible scenarios of what might, or might not, be discovered in this new range of energy. The Higgs particle, which was predicted over 40 years ago, could finally be unveiled. On the other hand, if there is no Higgs, our current understanding of physics might be shaken fundamentally. It is clear that, if the SM describes physics at the investigated energy scale, the Higgs could be discovered already with an integrated luminosity of 30 fb^{-1} or less depending on its mass [21]. Its non-existence would cast very serious doubts on the consistency of the SM. Furthermore, there are numerous theories which predict various different particles within the energy range of the LHC. All these theories can be tested and some of them, like most supersymmetric models, provide very clear signatures which could be discovered at the LHC. Unfortunately, most theories can adjust their parameters such that they cannot be excluded completely. For example, in some areas of the MSSM, in regions of moderate $\tan(\beta)^3$ and large values of $A_0{}^4$ [22], the lightest Higgs boson might be very difficult to detect and all other particles would be invisible within the investigated range of energy [21].

History has shown that each step into a new area of physics challenges existing theories and opens our eyes towards better ones. Furthermore, it is clear that the discoveries at the LHC will have a strong impact on the future of particle physics, especially on the question of whether a linear collider should be build or not. Also, the design of such a collider depends on the energy range of the particles which may be discovered. In any case, the LHC will provide important experimental data for theories which have been developed and worked on in the past decades.

For the discovery of new physics, a precise knowledge about the experimental conditions is extremely important. Besides the understanding of various other uncertainties, such as luminosity, alignment and calibration, the knowledge about uncertainties coming from parton distribution functions in the proton (PDF) and higher order corrections in the perturbative description of the SM is crucial.

1.2.2 A Toroidal LHC ApparatuS (ATLAS)

ATLAS is a general-purpose experiment to study proton-proton collisions at the LHC. It will be run by a collaboration of 2100 physicists working in more than 165 universities and laboratories from 37 countries. The detector design was optimised to maximise the possible range for physics analyses, from precision measurement of SM parameters like W and top masses to the searches for the Higgs boson or other symmetry-breaking mechanisms, supersymmetric particles, and other indications of the existence of physics beyond the SM (leptoquarks, microscopic black-holes, ...).

[3] relative contribution that the two Higgs doublets contribute to the electroweak symmetry breaking
[4] trilinear scalar couplings to A_0 at the GUT scale

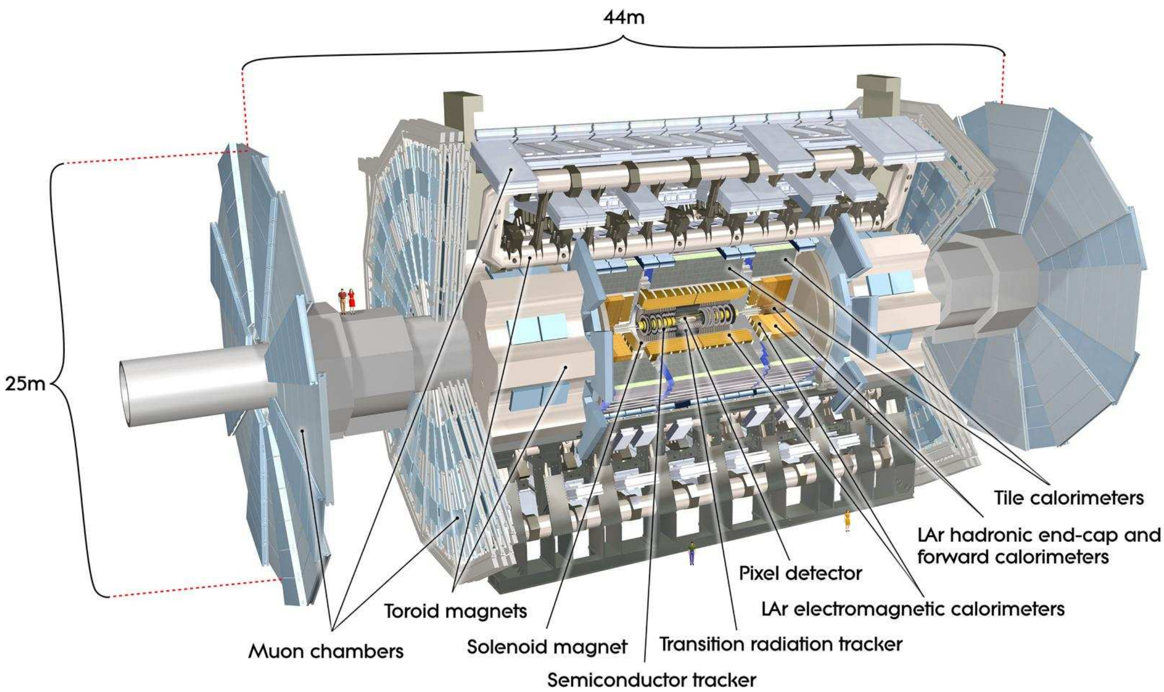


Figure 1.4: A computer-generated view of the ATLAS detector. The dimensions of the detector are ~ 25 m in height and ~ 44 m in length. The weight of about 7000 tonnes is close to that of the Eiffel Tower (7300 tonnes).

The ATLAS detector, represented on Figure 1.4, is built as a large cylinder 25 meters in diameter and 44 meter long centred around the beam pipe. Special emphasis is given to a region about 60 cm long in the centre of the detector where the two beams will collide. The nominal interaction point (IP) is defined as the origin of the coordinate system, while the beam direction defines the z-axis, and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The positive z-axis was chosen to form a right-handed coordinate-system. Hence the A-side of the detector, facing the Geneva lake, is defined as that with positive z and the C-side, facing the Jura, is that with negative z. The radius r of an element is defined as the distance between that element and the beam pipe in the x-y plane. The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$ (in the case of massive objects such as Z^0 bosons, the rapidity $y = 1/2\ln[(E+p_z)/(E-p_z)]$ is used). The transverse momentum p_T, the transverse energy E_T, and the missing transverse energy E_T^{miss} are defined in the x-y plane unless stated otherwise. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The ATLAS detector has been built forward-backward symmetric with respect to the nominal IP. The magnet configuration comprises a thin superconducting solenoid surrounding the inner-detector cavity, and three large superconducting toroids (one barrel and two end-caps) arranged with an eight-fold azimuthal symmetry around the calorimeters. This fundamental choice has driven the design of the rest of the detector.

The inner detector (Section 1.2.2.1, Figure 1.6) is immersed in a 2 T solenoidal field. Pattern recognition, momentum and vertex measurements, and electron identification are achieved with a combination of discrete, high-resolution semiconductor pixel and strip detectors in the inner part of the tracking volume, and straw-tube tracking detectors with the capability to generate and detect transition radiation in its outer part.

The calorimeters (Section 1.2.2.2, Figure 1.12) provide the energy reconstruction in AT-

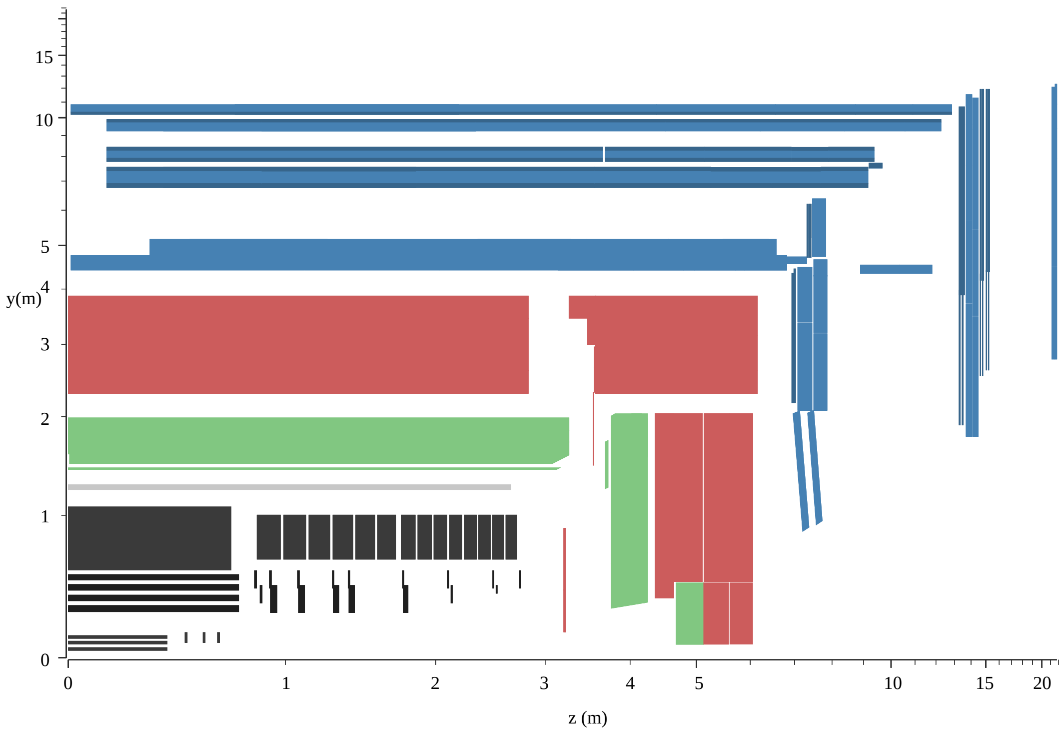


Figure 1.5: View of the ATLAS detector in the y-z plane. The detector is symmetric on the other side of the y axis and present a cylindrical symmetry around the z axis. The inner detector is represented in black, the solenoid in grey, the calorimeters in green and red, and the muon spectrometer in blue.

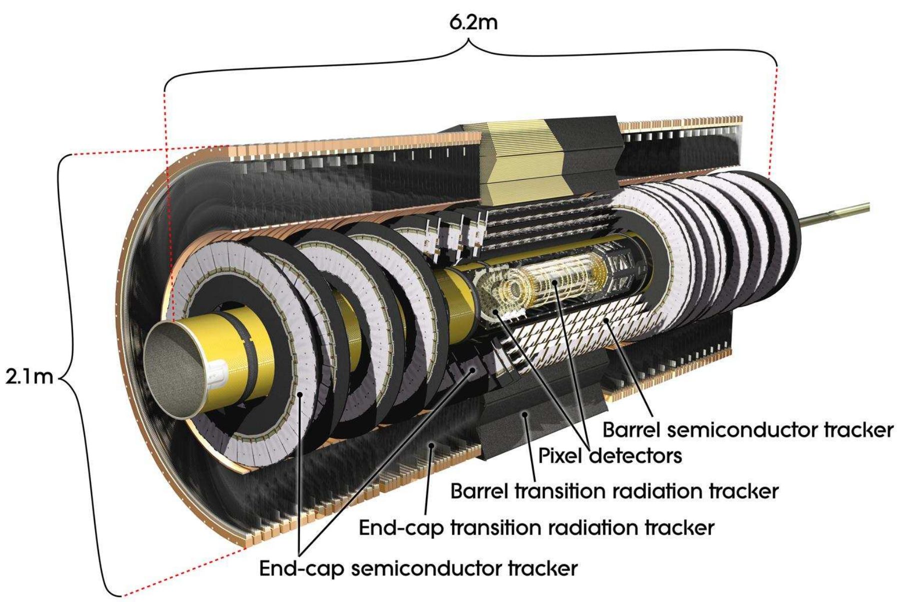


Figure 1.6: Perspective cut-away view of the inner detector. The different sub-structures are labelled. The outer radius of the inner detector is 103 cm, for a total length of 6.2 m

LAS. High granularity liquid-argon (LAr) electromagnetic sampling calorimeters, with excellent performance in terms of energy and position resolution, cover the pseudorapidity range $|\eta| < 3.2$. The hadronic calorimetry in the range $|\eta| < 1.7$ is provided by a scintillator-tile calorimeter, which is separated into a large barrel and two smaller extended barrel cylinders, one on either side of the central barrel. In the end-caps ($|\eta| > 1.5$), LAr technology is also used for the hadronic calorimeters, matching the outer $|\eta|$ limits of end-cap electromagnetic calorimeters. The LAr forward calorimeters provide both electromagnetic and hadronic energy measurements, and extend the pseudorapidity coverage to $|\eta| < 4.9$.

The calorimeter is surrounded by the muon spectrometer (section 1.2.2.3). The air-core toroid system, with a long barrel and two inserted end-cap magnets, generates strong bending power in a large volume within a light and open structure. Multiple-scattering effects are thereby minimised, and excellent muon momentum resolution is achieved with three layers of high precision tracking chambers.

1.2.2.1 Inner detector

The inner detector is the innermost part of the ATLAS detector. It consists of three subdetectors: from the beam pipe out, a silicon pixel detector (PIXEL), a silicon strip detector (SCT), and a transition radiation tracker detector (TRT).

The inner detector was designed to provide tracking in a pseudo-rapidity range up to $|\eta| < 2.5$. However, for budget reason, the last big wheels of the TRT end-cap were not built, so that the TRT coverage only extends to a pseudo-rapidity of $|\eta| < 2.0$.

The PIXEL (Figure 1.8) is designed for high-precision particle tracking close the interaction point. Its precision allows finding relatively long-lived particles such as b-mesons and τ-leptons. The detector is constituted of 1744 modules which are silicon wafers of $16.4 \times 60.8\ mm^2$, each covered by 47268 pixels. Out of these pixels, 41984 have a size of $50 \times 400\ \mu m^2$ each, while the remaining 5284 each have a size of $50 \times 600\ \mu m^2$ to cover

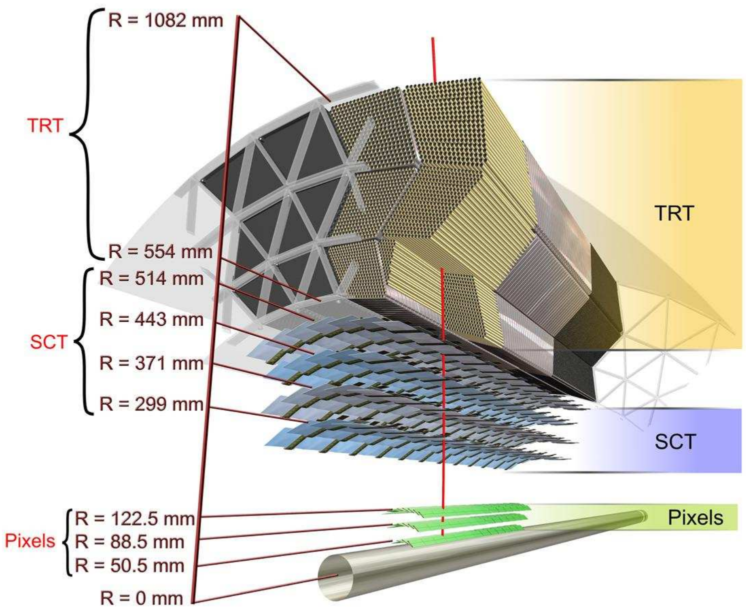


Figure 1.7: Drawing showing the sensors and structural elements traversed by a charged track in the barrel inner detector. The track traverses successively the beryllium beam-pipe, the three cylindrical silicon-pixel layers, the four cylindrical double layers of barrel silicon microstrip sensors (SCT), and approximately 36 axial straws contained in the barrel transition radiation tracker modules within their support structure. [21]

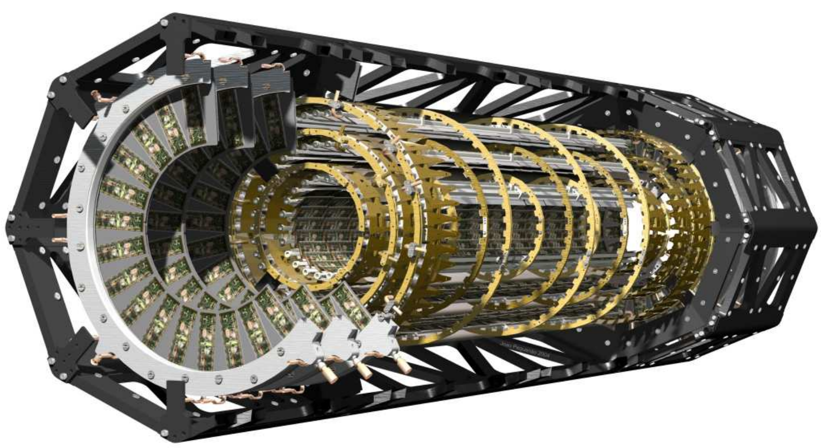


Figure 1.8: Computer-generated 3D view of the PIXEL layout, with three barrel layers, three disks on each side and its support structure. The full length is 1600 mm for an outer radius of 200 mm. [21]

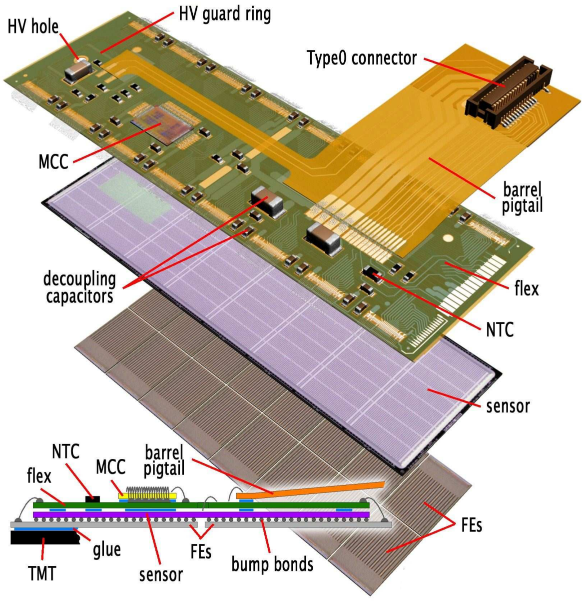


Figure 1.9: Schematic view of a barrel PIXEL module (top) illustrating the major pixel hybrid and sensor elements, including the MCC (module-control chip), the front-end (FE) chips, the NTC thermistors, the high-voltage (HV) elements and the Type0 signal connector. Also shown (bottom) is a plan view showing the bump-bonding of the silicon pixel sensors to the polyimide electronics substrate. [21]

the gap between their read-out chips. Figure 1.9 shows a schematic view of a barrel pixel module. The PIXEL consists of three cylindrical barrel layers at radii 50.5 mm, 88.5 mm and 122.5 mm, respectively (see Figure 1.7). The innermost layer is also later referred to as the b-layer, as it is the most critical for the resolution needed to reconstruct secondary vertices, which a critical discrimination factor to separate b-, c- and light quark jets. The 1456 barrel modules are glued to carbon staves which are inclined by an azimuthal angle of 20 degree around their long axis. This angle called the Lorentz angle, ensures that the deposited charges drift straight to the anode in the solenoid magnetic field. There also are three end-cap disks on each side to increase coverage in pseudorapidity while maintaining the number of hits for each track. Each disk is made of 8 sectors having 6 modules each. The heat of the modules is evacuated through the flex hybrid board to aluminium cooling tubes fixed to the carbon support structure.

The SCT (Figure 1.10) provides precision track measurements in the intermediate radial

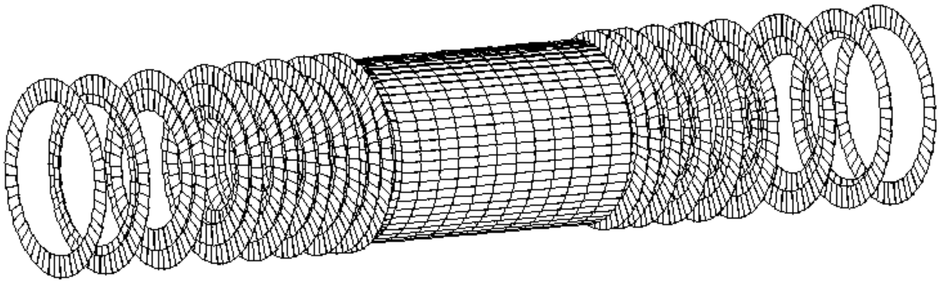


Figure 1.10: Schematic view of the SCT layout, with barrel and 18 disks. The full length is 5440 mm for an outer radius of 514 mm. [21]

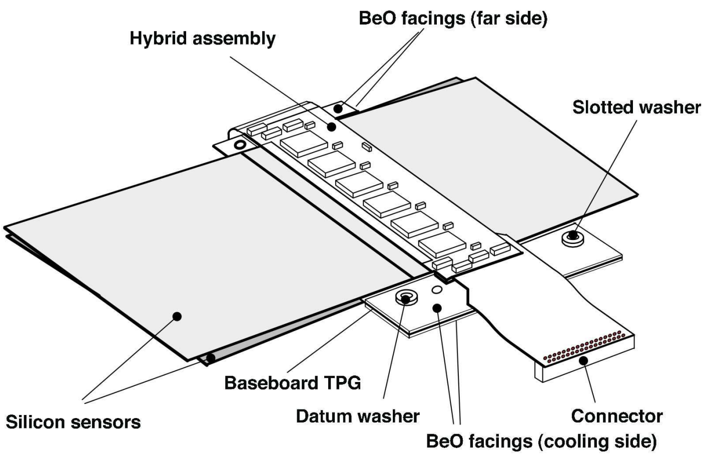


Figure 1.11: Schematic view of a barrel SCT module. The readout electronics is located in the centre of the module. The electronics board is also connected to the cooling system to dissipate the excess heat from the silicon. [21]

region while limiting cost. It contributes to the measurement of momentum, impact parameter and vertex position. The 2112 SCT barrel modules are mounted on four carbon-fibre cylinders with radii of 299 mm, 271 mm, 443 mm, 514 mm and a length of 1498 mm symmetric around the nominal IP. Another 1976 modules are mounted on 18 SCT end-cap disks. Barrel modules consist of four wafers of 768 microstrips mounted in two planes which are glued together back-to-back at a 40 mrad stereo angle as presented in figure 1.11. The pitch in the barrel is 80 μm, and in the end-cap it varies with radius from 55 to 95 μm because the modules are wedge-shaped to accommodate the wheel structure. The two planes with a small stereo angle allow the determination of the hit position along the strips. The design is a compromise between reducing cost and number of readout channels on the one hand and improving the spatial resolution on the other hand. Just like the PIXEL barrel modules, the SCT barrel modules are mounted on their support structure with a small Lorentz angle around the beam direction as seen on Figure 1.7. The spatial resolution is 17 μm in the direction perpendicular to the strips and 580 μm in the direction parallel to the strips [23]. The SCT allows a coverage of the pseudorapidity range up to $|\eta| < 2.5$.

The final part of the ATLAS Inner Detector is a combined straw-based drift chamber tracker and a transition radiation detector as seen on the outside of figures 1.6 and 1.7. The TRT barrel is made of 52544 axial straws (aligned to the beam pipe) which are 4 mm in diameter and about 150 cm in length. They are located between 56 cm and 107 cm away from the beam pipe. The TRT end-cap detectors contain a total of 246760 radial straws (perpendicular of the beam pipe) at radii between 64 cm and 103 cm. The TRT provides on average, 36 two-dimensional measurement points per track with a resolution[5] of less than 0.170 mm [24]. The higher signal collected when particles emit transition radiation facilitates distinguishing electrons from pions. The technology used consists of Kapton straws interleaved with polypropylene foils or fibres acting as radiators, and a gas mixture made of Xenon (70%) –for good X-ray absorption– with the addition of CO_2 (27%) and O_2 (3%) to increase the electron drift velocity and for photon-quenching.

[5] the resolution is dependent on the gas mixture

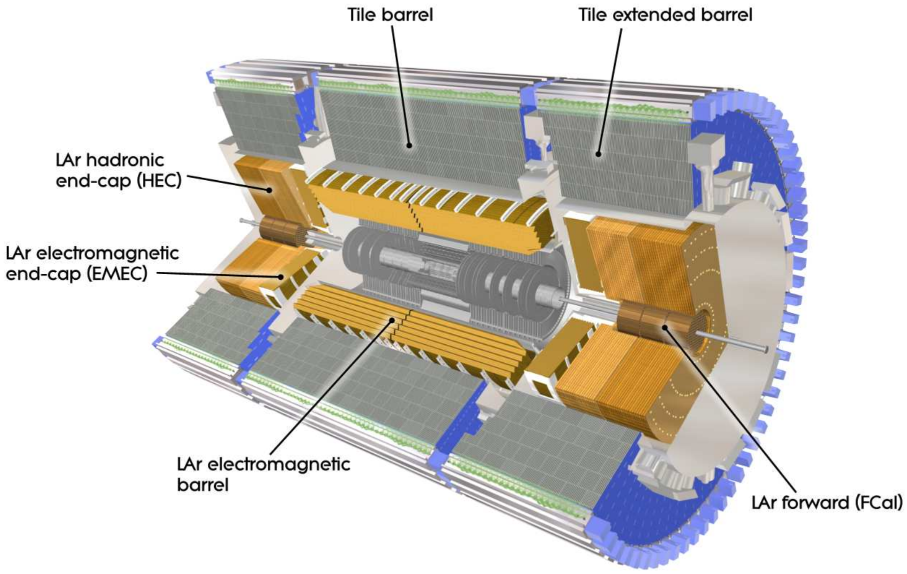


Figure 1.12: Cut-away view of the ATLAS calorimeter system.

1.2.2.2 Calorimeters

The ATLAS calorimeters surround the central tracking detectors and are separated from these by the solenoidal magnet as seen on Figure 1.5 and 1.12. Their aim is to measure the energy of particles originating from the proton-proton interactions at the IP. They consist of an electromagnetic (EM) calorimeter covering the pseudorapidity region $|\eta| < 3.2$, a hadronic barrel calorimeter covering $|\eta| < 1.7$, hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering the region $3.1 < |\eta| < 4.9$ as seen on figure 1.13. Particles are stopped in the material of the calorimeters in order to determine their energies. The main fraction of the energy gets absorbed in layers of high-density material. The EM calorimeter which is responsible for the absorption of electromagnetically interacting particles like electrons, positrons and photons, is based on a lead and stainless steel structure. Its high granularity makes it very precise for the measurement of both the amount of energy absorbed and the location of the deposited energy. Angles of particle trajectories can be measured within about $50 - 60\ mrad/\sqrt{E\ (GeV)}$ [21].

The hadronic calorimeter uses steel plates in the central and tungsten in the forward

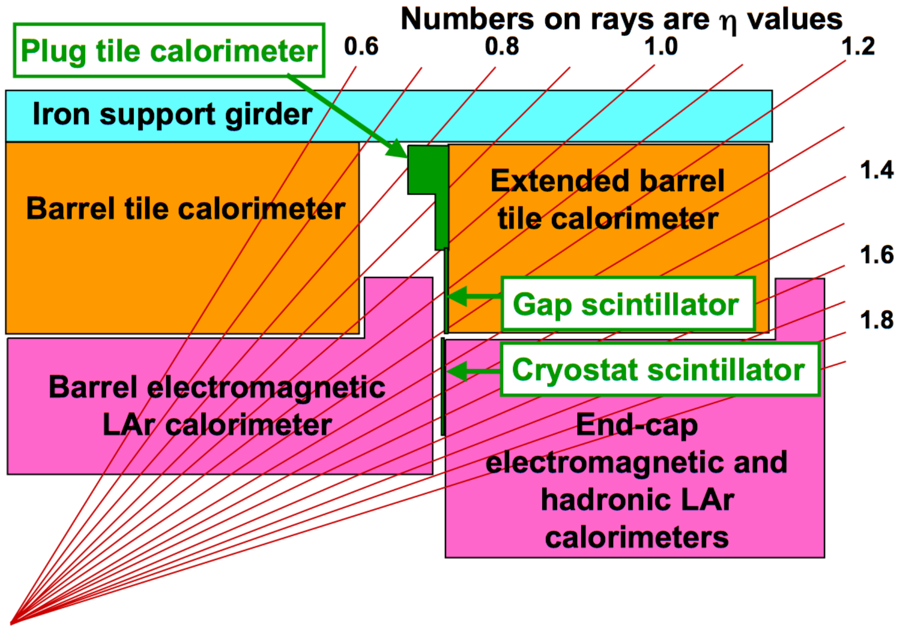


Figure 1.13: Schematic of the transition region between the barrel and end-cap cryostats, where additional scintillator elements are installed to provide corrections for energy lost in inactive material (not shown), such as the liquid-argon cryostats and the inner-detector services.

calorimeter as absorption material to measure the energy of strongly interacting particles. It localises particles with a precision of about 100 $mrad/\sqrt{E}$ GeV [24]. Charge particles, produced in the absorption process of the primary particles, ionize the liquid argon which is interleaved between the absorber plates. Due to an applied electric field, electrons released drift, inducing a current on the electrode structure. The total current is proportional to the energy of the incoming particle. The described type of calorimeter is called a sampling calorimeter because only a small sample of the deposited energy is actually measured in the active medium. The correct calibration of the calorimeters is therefore crucial for their physics performance. Initial calibration is extrapolated from extensive test beam data collected on the different module types. Further improvements involve a complex procedure which is partly based on using particles originating from known decays of SM particles.

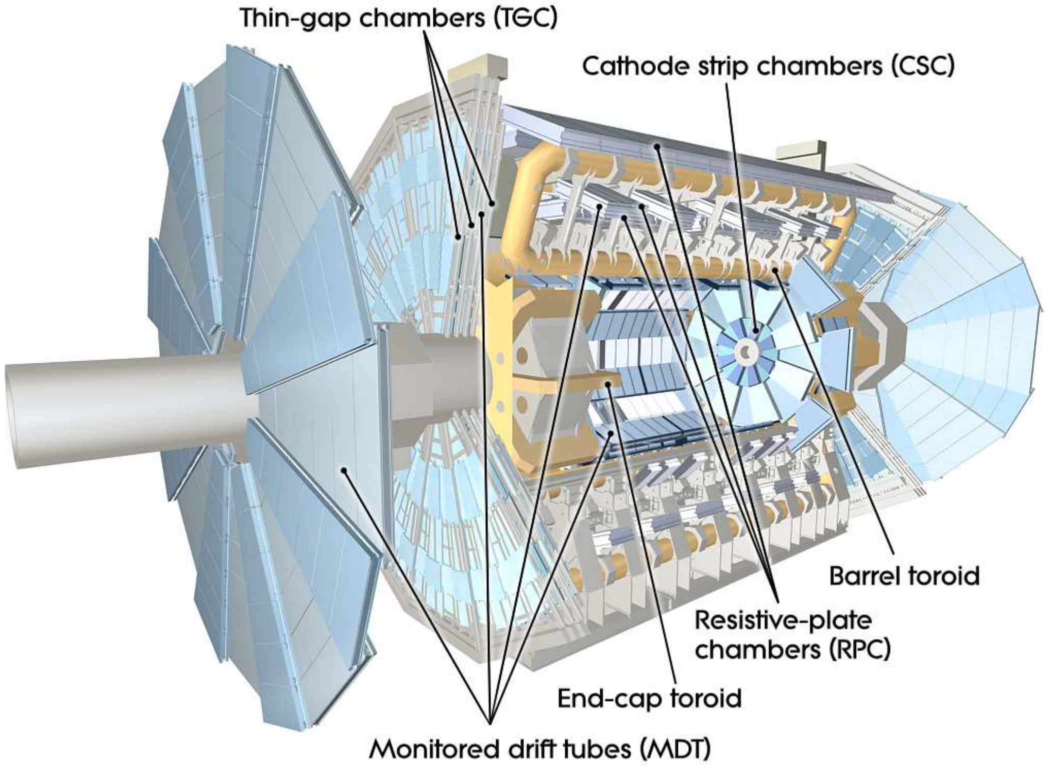


Figure 1.14: Cut-away view of the ATLAS muon system.

1.2.2.3 Muon spectrometer

A schematic layout of the muon spectrometer is shown in Figure 1.14. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. This magnet configuration provides a field which is mostly orthogonal to the muon trajectories, while minimising the degradation of resolution due to multiple scattering. The anticipated high level of particle flux has had a major impact on the choice and design of the spectrometer instrumentation, affecting performance parameters such as rate capability, granularity, ageing properties, and radiation hardness.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis; in the transition and end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers. Figure 1.14 shows the arrangement of the different detectors. Over most of the η-range, a precision measurement of the track coordinates in the principal bending direction of the magnetic field is provided by Monitored Drift Tubes (MDT's). The mechanical isolation in the drift tubes of each sense wire from its neighbours guarantees a robust and reliable operation. At large pseudorapidities, Cathode Strip Chambers (CSC's, which are multiwire proportional chambers with cathodes segmented into strips) with higher granularity are used in the innermost plane over $2 < |\eta| < 2.7$, to withstand the demanding rate and background conditions. The stringent requirements on the relative alignment of the muon chamber layers are met by the combination of precision mechanical-assembly techniques and optical alignment systems both within and between muon chambers. The trigger system covers the pseudorapidity range $|\eta| < 2.4$. Resistive Plate Chambers (RPC's) are used in the barrel and Thin Gap Chambers (TGC's) in the end-cap regions. The trigger chambers for the muon spectrometer serve a threefold purpose: to provide bunch-crossing identification, to provide well-defined p_T thresholds for the trigger, and to measure the muon coordinate in the direction orthogonal to that determined by the precision tracking chambers.

1.2.2.4 Trigger system

The Trigger and Data Acquisition (collectively TDAQ) systems, the timing- and trigger-control logic, and the Detector Control System (DCS) are partitioned into sub-systems, typically associated with sub-detectors, which have the same logical components and building blocks. The trigger system has three distinct levels: L1, L2, and the event filter (EF) .Each trigger level refines the decisions made at the previous level and, where necessary, applies additional selection criteria. L1 is a fully electronic dedicated hardware trigger, whereas L2 and EF are software based trigger levels. The data acquisition system receives and buffers

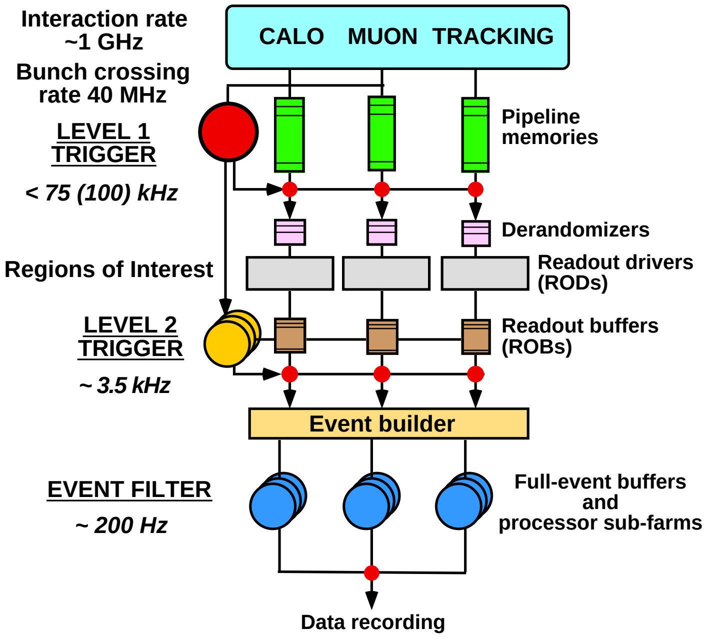


Figure 1.15: Diagram of the trigger levels. L1 bases its decision on crude calorimeter and muon system data. L2 uses the full granularity of the detector in the ROI's pointed out by L1. EF uses the full detector information. Timing requirements for each level are also described.

the event data from the detector-specific readout electronics, at the L1 trigger accept rate, over 1600 point-to-point readout links. The first level uses a limited amount of the total detector information to make a decision in less than 2.5 μs, reducing the rate to about 75 kHz [21]. The two higher levels access more detector information for a final rate of up to 200 Hz with an event size of approximately 1.3 Mbyte.

The L1 trigger searches for high transverse-momentum muons, electrons, photons, jets, and τ-leptons decaying into hadrons, as well as large missing and total transverse energy. Its selection is based on information from a subset of detectors. High transverse-momentum muons are identified using trigger chambers in the barrel and end-cap regions of the spectrometer. Calorimeter selections are based on reduced-granularity information from all the calorimeters. Results from the L1 muon and calorimeter triggers are processed by the central trigger processor, which implements a trigger "menu" made up of combinations of trigger selections. Pre-scaling of trigger menu items is also available, allowing optimal use of the bandwidth as luminosity and background conditions change. Events passing the L1 trigger selection are transferred to the next stages of the detector-specific electronics and subsequently to the data acquisition via point-to-point links. In each event, the L1 trigger also defines one or more Regions-of-Interest (RoI's), i.e. the geographical coordinates in η and ϕ, of those regions within the detector where its selection process has identified interesting features. The RoI data include information on the type of feature identified and the criteria passed, e.g. a threshold. This information is subsequently used by the high-level trigger.

The L2 selection is seeded by the RoI information provided by the L1 trigger over a dedicated data path. L2 selections use, at full granularity and precision, all the available detector data within the RoI's (approximately 2% of the total event data). The L2 menus are designed to reduce the trigger rate to approximately 3.5 kHz, with an event processing time of about 40 ms, averaged over all events. The final stage of the event selection is carried out by the event filter, which reduces the event rate to roughly 200 Hz. Its selections are implemented using offline analysis procedures within an average event processing time of the order of four seconds.

The combined detector and trigger infrastructure are called the online part of the ATLAS analysis process as it operated in real–time as the data arrives and everything that is rejected by the trigger is lost forever and cannot be recovered.

1.2.2.5 ATLAS Software framework

Athena is the ATLAS general purpose offline software framework. It is build around a core project, Gaudi, that is common with the LHCb experiment. It provides facilities for various needs such as event generation, simulation, reconstruction and detector description. It is structured in packages which contain services and commonly used algorithms. Each package may have explicitly defined dependencies on other packages. The handling of input and output data as well as configuration of run specific parameters is done through a so-called "jobOption"-file, which is a Python script describing which algorithms to run and which values are to be used for configuration. The Athena framework is under continuous development by ATLAS physicists and new versions are released approximately every month. As data-taking approaches, it is converging towards a stabilisation of the data persistency formats. The analysis algorithms on the other end are evolving quickly and are expected to be regularly maintained to follow modifications in other packages and changes in the knowledge of the detector performance.

The offline software contains a generally complete description of the ATLAS detector as built. This description is used by the reconstruction and is subject to changes as more information about, for example, inner detector module alignment or calorimeter calibrations, becomes available. The reconstruction of a collision is performed by a collection of various tools. For the inner detector reconstruction, hit clusters are first built, then space-points[6] are constructed, then finally, a track finder is run to reconstruct charge particle trajectories. Similarly, energy clusters are reconstructed in the calorimeters by adding energy depositions in adjacent cells and are added to well lined tracks to build particle candidates. If no matched tracks are found, neutral object candidates are defined (photon, π^0, neutron,…)

[6] in the SCT, at least one hit on each side of a module is needed to construct one space-point

Chapter 2

Measuring the electron reconstruction efficiency at ATLAS

This chapter describes the electron reconstruction at ATLAS. It presents the software algorithm under study and the method to measure its efficiency from data with minimum reliance on Monte-Carlo. It presents the definition of the Tag and Probe method used for this measurement. The actual measurement and the method used to deal with the background are presented in the following chapters.

At a hadron machine like the LHC, lepton identification is a very important tool as the presence of leptons provides a relatively clean environment and is used in many signatures of both existing and new physics. Lepton identification can be split in two different aspects: on the one side, the probability that a lepton is indeed identified as a lepton, which is called lepton efficiency, and on the other side, the probability for a different particle or jet to be mistaken for a lepton, which is called the fake rate.

This work concentrates mainly on an algorithm to measure the electron efficiency, i.e., the probability of an electron to be identified as an electron by the combination of hardware and software. The electron efficiency is a very important parameter that enters in the measurement of many cross-sections and so it is important to obtain an accurate measurement of it even with early data.

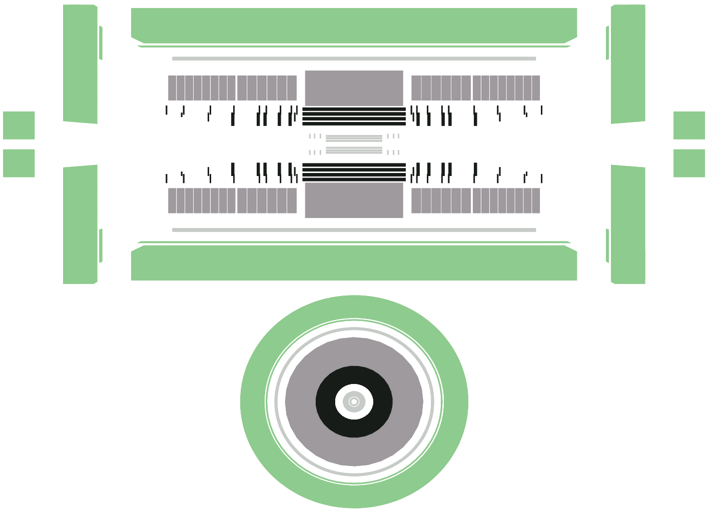


Figure 2.1: Views of the main detectors involved in electron reconstruction. From the centre out, PIXEL, SCT, TRT and Liquid-Argon Calorimeter. The solenoid is also represented.

2.1 Electron reconstruction algorithms

This section describes the main features of the electron reconstruction algorithms at ATLAS whose efficiency we try to assess.

2.1.1 Detector hardware involved

The electron reconstruction at ATLAS mainly involves two detector components:

- the inner detector tracking is used to determine the trajectory of the charged particles (i.e. track), and is used to separate electron and photons (see Section 1.2.2.1);
- the calorimeters are used to measure the energy of the particles. For electrons, the liquid argon electro-magnetic calorimeter is the most important sice most electrons don't reach the hadronic calorimeter (see Section 1.2.2.2).

The principle of electron reconstruction is that an electron passing through the detector should leave a track in the inner detector before depositing its energy in the Liquid-Argon electro-magnetic calorimeter. The software algorithm is built with this principle in mind.

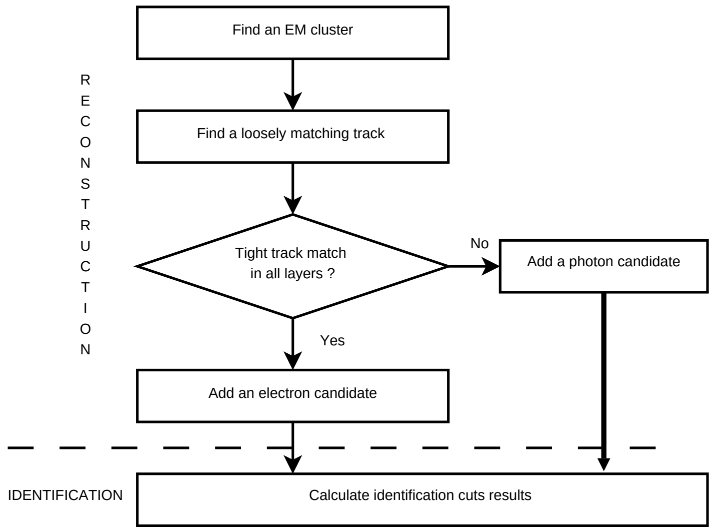


Figure 2.2: Diagram of the egamma electron algorithm. See text for details in section 2.1.2

The reconstruction software is split between two algorithms which are aimed at electrons originating from different processes.

The so-called "soft-electron" algorithm is geared toward low energy electron reconstruction, namely electrons originating from low mass objects like J/ψ[1] or even Υ[2] or electrons from decay of b- of c-jets.

2.1.2 High-p_T software algorithm

The main electron algorithm is named "egamma" and is aimed at high p_T electrons originating from the hard process. It is the main algorithm for electroweak physics studies.

The egamma algorithm is organized in two parts. The first part reconstructs the clusters in the calorimeter to create electron/photon candidates, the second part tries to associate a track to differentiate electron from photon candidates. A third separate part defines standard cuts to increase the probability to get the correct particle identification.

[1]The mass of the $J/\psi(1S)$ is 3096.916 MeV [3]
[2]The mass of the $\Upsilon(4S)$ is 10579.4 MeV [3]

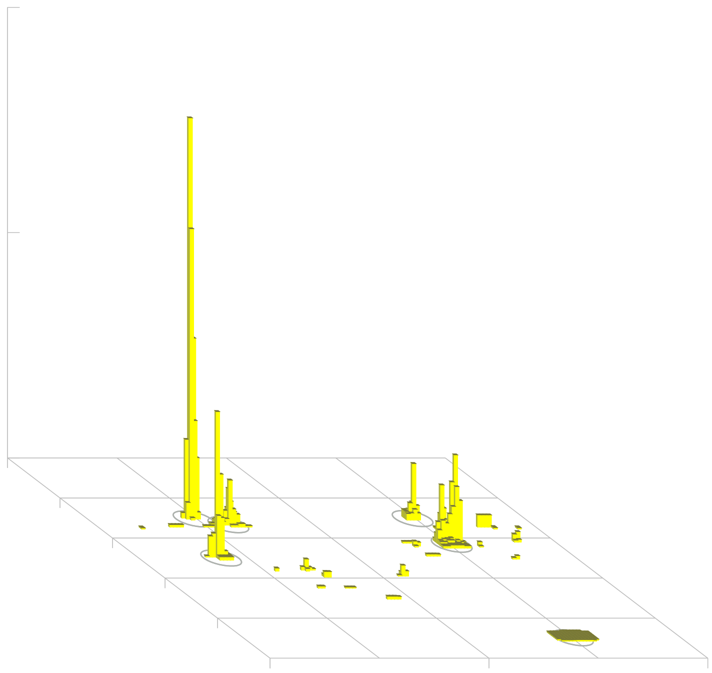


Figure 2.3: η-ϕ view of the energy deposition in the electro-magnetic calorimeter. The energy is summed by projective tower.

Layer	$\Delta\eta$	$\Delta\phi$	ΔR
Layer 1	0.0031	0.0980	4.3 X_0
Layer 2	0.025	0.0245	16 X_0
Layer 3	0.5	0.0245	2 X_0

Table 2.1: Granularity of the electro-magnetic calorimeter

2.1.2.1 Clustering algorithm

The default clustering algorithm is a Sliding Window Clustering algorithm. The name derives from the fact that the algorithm looks at narrow windows in $\eta \times \phi$ to find the clusters.

The clusters are reconstructed using a three step process :

1. the tower building;

2. the pre-cluster (seed) finding;

3. the cluster filling.

Parameter	**Value**
Tower building	
Calorimeter type	Electro-Magnetic (LAr)
η range	$[-2.5, 2.5]$
$\eta \times \phi$ divisions ($\Delta\eta \times \Delta\phi$)	200×256 ($.025 \times .025$)
Pre-cluster finding	
Size of window ($\eta \times \phi$ towers)	5×5
E_T^{thres} (GeV)	3
Duplicate separation ($\Delta\eta_{dupl} \times \Delta\phi_{dupl}$)	2×2
Cluster filling	
cluster size ($\eta \times \phi$)	3×5, 3×7 or 5×5,

Table 2.2: Parameters used in the clustering of the egamma reconstruction algorithm

The tower building phase consists in dividing the $\eta - \phi$ plane in a large number of squares (≈ 50000) and adding the energy of all the cells in each square, giving the representation of the energy deposition in Figure 2.3. For electro-magnetic clusters, only cells in the electro-magnetic calorimeter are considered. The electro-magnetic calorimeter segmentation can be seen on Figure 2.5 as well as its three layer structure. The granularity of the different parts of the electro-magnetic calorimeter are summarized in Table 2.1.

The pre-cluster finding consists of making "windows", i.e. groups, of the previously built towers (typically 5×5 towers in $\eta \times \phi$). A pre-cluster is formed from the accumulated energy of such a window is above a threshold E_T^{thres}. The size of the window and the threshold are optimised to obtain the best efficiency of the pre-cluster search and to limit the rate of fake pre-clusters due to noise. Duplicated and/or overlapping clusters are removed by imposing a minimum separation between the cluster centres, otherwise only the most energetic cluster is kept. Clusters are circled in grey on Figure 2.3.

Finally, the cluster filling consists of including in the cluster all the cells in a rectangle centred around the seed position. Different sizes can be used and the calibration is applied depending of the object type (electrons or photons). The identification criteria (see Section 2.1.2.3) are applied to each cluster size and the one that passes the most criteria is kept.

Full details on the clustering algorithm as well as the values of the parameters are de-

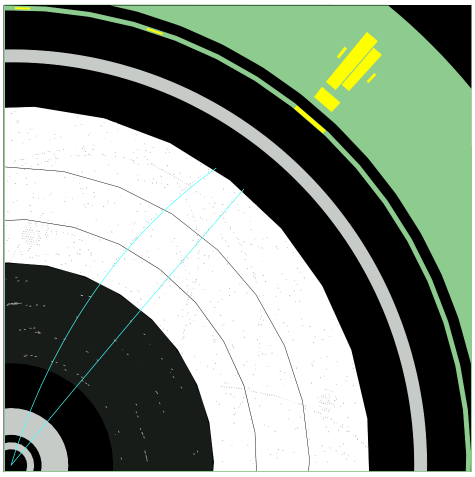


Figure 2.4: The extrapolation of the track to the calorimeter must match with the cluster centre.

scribed in the ATLAS Detector Paper [21]. The parameters used in version 14 of the ATLAS software can be found in Table 2.2.

2.1.2.2 Track matching

After an electro-magnetic cluster has been identified, one needs to determine whether it is a photon or an electron candidate. This is dependent on whether a track can be matched to this cluster.

This is done in two steps. First a crude matching is done. The $\eta - \phi$ direction parameters of the track at the interaction vertex are loosely matched to the cluster barycentre. Then, for the tracks that fulfil that requirement, the $\eta - \phi$ position is extrapolated to each layer of the calorimeter and a close match is required.

If more than one track passes the extrapolation requirement as in Figure 2.4, the software only keeps the one such that the ratio of energy to momentum (E/p) is closest to 1. The energy is extracted from the energy of the cluster, whereas the momentum is calculated from the track curvature. Again, particular values for the cuts can be found in the ATLAS Detector Paper [21].

Parameter	Value
crude matching η	< 0.05
crude matching ϕ	< 0.1
tight matching η	< 0.05
tight matching ϕ	$[-0.1, 0.05]$
E/p matching	< 10

Table 2.3: Parameters used in tracking part of the egamma reconstruction algorithm

Clusters for which a reasonable track-match exists are considered electron candidates, while the others are kept as photon candidates. This concludes the reconstruction part of the algorithm.

Table 2.3 summarizes the parameters used in the egamma track-matching algorithm.

2.1.2.3 Identification

Immediately after a candidate has been discovered, identification tests are performed on electron candidate to separate real electrons from fakes according to different criteria. These are not part of the reconstruction process as such and so their performance is not the object of this measurement. But given that they are used by the measurement to reject the background, we describe them here as well.

The following calorimeter-based discriminating variables are used:

Hadronic leakage is the ratio of the transverse energy deposited in the hadronic calorimeter to the total transverse energy of the electron;

Second layer electrons deposit most of their energy in the second sampling layer of the calorimeter as defined on Figure 2.5. In this layer two variables are used:

- **Shower shape** most of the energy should be deposited close to the centre of the shower. The shape is defined by the ratio of the energy deposited in the core of the shower to the energy deposited in a wider circle;
- **Shower width** electrons have narrower showers than hadronic jets;

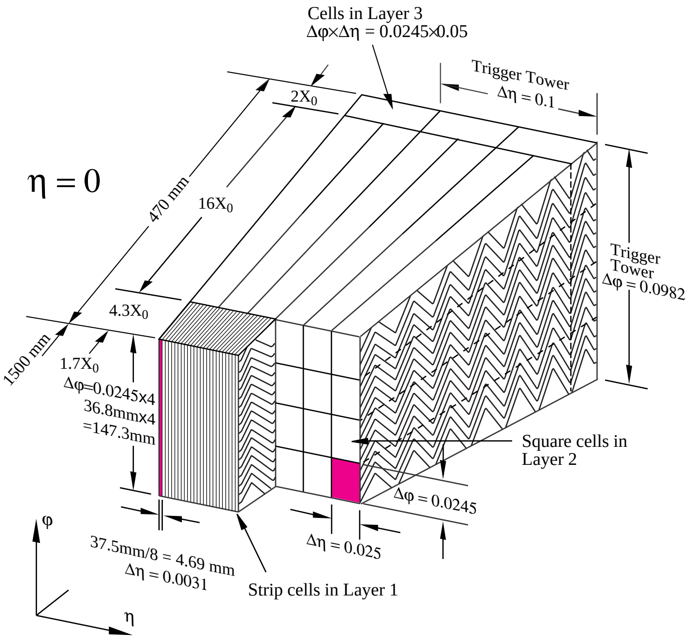


Figure 2.5: Sketch of a EM calorimeter barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ. The granularity in eta and phi of the cells of each of the three layers and of the trigger towers is also shown.

First Layer with its very fine granularity in rapidity as pictured on Figure 2.5, it can be used to detect substructures within a shower. Pion jets (π^0/η) are found to have often two maxima as they mostly decay to two photons. The test searches for a second maximum close to the hottest cell in this layer.

Isolation we can ensure the object is isolated by requiring a hollow cone around the object core to not exceed a certain energy relative to the amount of energy deposited in its core.

All these cuts are dependent on the pseudo-rapidity to account for the detector geometry. The values at central rapidity for a 20 GeV electron are recorded in Table 2.4.

In addition to the calorimeter based identification, a certain number of track based criteria are also defined :

Track-quality require a certain number of hits to remove track fitter fakes:

- at least nine precision hits (PIXEL+SCT);

Cut Name	**Value**
Hadronic Leakage	$E_T^{had1}/E_T^{e} < 0.02$
Shower Shape	$E_{3\times7}^{EM2}/E_{7\times7}^{EM2} > 0.77$
Shower Width	$width^{EM2} < 0.014$
First Layer 1	$E_{2^{nd}max}^{EM1}/E_{min\ between\ max}^{EM1} < 0.15$
First Layer 2	$width_{total}^{EM1} < 4$
First Layer 3	$(E_{5strips}^{EM1} - E_{3strips}^{EM1})/(E_{3strips}^{EM1}) < 0.45$
First Layer 4	$width_{3strips}^{EM1} < 0.80$
Isolation	$E_T^{cone20}/E_T < 0.5$

Table 2.4: Calorimeter-based identification cuts for a 20 Gev Electron at $\eta = 0$

- at least two hits in the pixel layers, one of the hits being in the b-layer;
- a small transverse impact parameter $|d_0| < 1$ mm.

spacial matching tighter cuts on the extrapolated $\Delta\eta$ and $\Delta\phi$ matching between the track and the cluster;

energy matching tighter cuts on the E/p matching between the track and the cluster;

transition radiation an electron should have a high number of high threshold hits in the TRT from the higher transition radiation while traversing the radiating material.

All these cuts are again dependent on the pseudo-rapidity to better match the detector geometry. The values at central rapidity for a 20 GeV electron are recorded in Table 2.5.

We can choose to apply any combination of these cuts to reduce the contamination of fake jets at the cost of reducing the efficiency of the identification.

There are however some predefined sets of selection defined by the egamma working group that are designed to be good compromise between efficiency and purity [25]. Rejections and efficiencies are given in Table 2.6.

Cut Name	**Value**
PIXEL+SCT	$N_{hits} > 9$
PIXEL	$N_{hits}^{PIXEL} > 2$
B-Layer	$N_{hits}^{b-layer}! = 0$
Impact Parameter	$\lvert d_0 \rvert < 1mm$
Matching $\Delta\eta$	$\lvert\Delta\eta\rvert < 0.005$
Matching $\Delta\phi$	$\lvert\Delta\phi\rvert < 0.02$
Energy Matching	$0.8 < E/p < 2.5$
TRT Hits	$N_{hits}^{TRT} \geq 15$
Transition Radiation Hits	$N^{TR}/N_{hits}^{TRT} > 0.08$

Table 2.5: Track-based identification cuts for a 20 GeV E_T Electron at $\eta = 0$

IsEM ElectronLoose This set of cuts performs a simple electron identification based only on limited information from the calorimeters. Cuts are applied on the hadronic leakage and on shower-shape variables, derived from only the middle layer of the EM calorimeter (lateral shower shape and lateral shower width). This set of cuts provides excellent identification efficiency, but low background rejection.

IsEM ElectronMedium This set of cuts improves the background rejection by adding cuts on the strips in the first layer of the EM calorimeter and on the tracking variables.

IsEM ElectronTight This set of cuts makes use of all the particle-identification tools currently available for electrons. In addition to the cuts used in the medium set, cuts are applied on the number of vertexing-layer hits (to reject electrons from conversions), on the number of hits in the TRT, on the ratio of high-threshold hits to the number of hits in the TRT (to reject the dominant background from charged hadrons).

<table>
<tr><th rowspan="3">Cuts</th><th colspan="2">$E_T > 17$ GeV</th><th colspan="2">$E_T > 8$ GeV</th></tr>
<tr><th>Efficiency (%)</th><th>Jet rejection</th><th>Efficiency (%)</th><th>Jet rejection</th></tr>
<tr><th>electrons from $Z^0 \rightarrow e^+e^-$</th><th></th><th>Single electrons ($E_T = 10$ GeV)</th><th></th></tr>
<tr><td>Loose</td><td>87.96 ± 0.07</td><td>567 ± 1</td><td>75.8 ± 0.1</td><td>513 ± 2</td></tr>
<tr><td>Medium</td><td>77.29 ± 0.06</td><td>2184 ± 13</td><td>64.8 ± 0.1</td><td>1288 ± 10</td></tr>
<tr><td>Tight</td><td>64.22 ± 0.07</td><td>$(9.8 \pm 0.4) \cdot 10^4$</td><td>48.5 ± 0.1</td><td>$(5.8 \pm 0.3) \cdot 10^4$</td></tr>
</table>

Table 2.6: Expected efficiencies for isolated electrons and corresponding jet background rejections for the three standard levels of cuts used for electron identification. The results are shown for the simulated filtered di-jet and minimum-bias samples, corresponding respectively to E_T-thresholds of 17 GeV (left) and 8 GeV (right). Rejection is calculated as the number of rejected jet for every accepted one in a pure light jet sample. The quoted errors are statistical [25].

2.1.2.4 Other refinements

Apart from the main features described previously, additional refinements are gradually being introduced into the ATLAS software.

Conversion recovery This algorithm tries to recover converted primary photons by looking at nearly parallel tracks. It can decide that an electron is in fact originating from a converted photon and as such move the object to the photon candidate collection. Because of the high number of tracks in the ATLAS tracker, fakes are a concern and some primary electrons may falsely be labelled as converted photons.

Bremsstrahlung recovery These algorithms try to recover hard bremsstrahlung of the electron track. They try to correct the curvature of the track to get a better fit and extract a better measurement of the overall electron momentum.

Backtracking This algorithm tries to reconstruct a track by starting from a calorimeter hot cell and going toward the detector centre. The primary focus is to reconstruct tracks from late conversions in the tracker. Again, because of the high occupancy in the tracker, fake tracks may be reconstructed leading to reconstruction problems.

2.1.3 Low p_T electron software algorithm

This algorithm is track based, as opposed to the egamma algorithm. First the algorithm searches for high quality tracks. Then theses tracks are extrapolated to the calorimeter where a cluster is built. The same identification variables as for the egamma algorithm are then available.

The high track quality is assessed by the following requirements:

- at least two hits in the pixel detector, one of which being in the b-layer;
- a minimal number of precision hits (pixels+SCT);
- small impact parameter of the track in the transverse plane $|d_0| < 1$ mm;
- at least one high threshold hit in the TRT detector along the track, corresponding to at least one instance of transition radiation; (if $|\eta| < 2$)
- at least 20 hits in the TRT detector along the track. (if $|\eta| < 2$)

After a good track has been selected, a fixed size cluster is constructed around the extrapolation of the track in the second sampling of the calorimeter. To reduce fake jets, further requirements are imposed on the cluster reconstructed after the extrapolation:

- $E/p > 0.7$ where p is extracted from the track and E is the cluster energy;
- $E1/E > 0.03$ (Energy in the first EM layer over full cluster energy);
- $E3/E < 0.5$ (Energy in the third EM layer over full cluster energy).

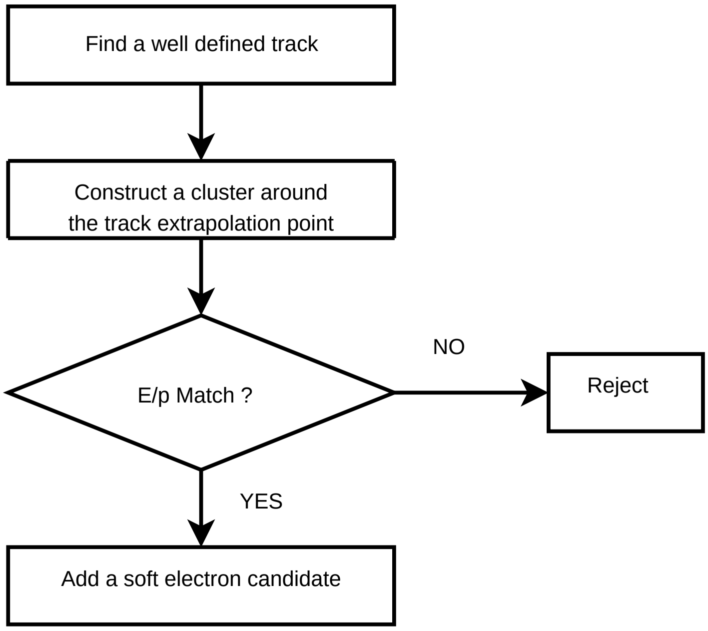


Figure 2.6: Diagram of the soft-electron algorithm

Cut Name	**Value**
PIXEL	$N_{hits}^{PIXEL} \geq 2$
B-Layer	$N_{hits}^{b-layer}! = 0$
PIXEL+SCT	$N_{hits} \geq 7$
Impact Parameter	$\|d_0\| < 1$ mm
TRT Hits	$N_{hits}^{TRT} \geq 20$
Transition Radiation Hits	$N^{TR}/N_{hits}^{TRT} > 0.05$
Minimum p_T	$p_T > 2$ GeV
Energy Matcht	$E/p > 0.7$
First EM Layer	$E1/E > 0.03$
Third EM Layer	$E3/E < 0.5$

Table 2.7: Requirement on track and cluster for a soft electron reconstruction

2.2 Definition of the measurement

At an experiment like ATLAS, electron identification is the result of many hardware components and software algorithms. In the ATLAS software definitions, electrons are selected at different levels :

- on-line, there are three different definitions matching the three trigger levels at ATLAS.
- off-line, the reconstruction starts by finding cluster/track combinations that are likely electron candidates
- subsequently, additional tests are performed on the candidates in an identification step.

The way the trigger system tries to identify electrons is described in the “Electron and Photon” Chapter of the ATLAS Technical Report [25]. This thesis does not try to determine the performance of the trigger system.

This works focuses on an algorithm to try to measure the offline efficiency for a real electron to be reconstructed as an electron candidate by the software. Its purpose is to measure the efficiency of the first step of the off-line reconstruction process described above.

The efficiency of the identification step is studied separately in the “Electron and Photon” Chapter of the ATLAS Technical Report [25].

The aim of this work is the measurement of the efficiency of the first step of the reconstruction which we will refer to simply as “reconstruction” as opposed to “identification”, which will refer to the second step of the software reconstruction process.

This is confusing because many papers, in particular cross-section measurements, refer to the electron efficiency as the efficiency of the combined reconstruction and identification steps. But since many analyses uses different identification cut sets, this leads to a set of different “electron efficiencies”.

The ATLAS performance group point of view has been that studying the reconstruction and the identification separately is more efficient. It allows different analyses to use different identification cuts and understand their differences. Moreover, the two steps may have slightly different systematics which can then be studied separately.

What this thesis will refer to as the “electron efficiency” is henceforth really the electron *reconstruction* efficiency, while we will refer to the efficiency of the identification process as the “identification efficiency”.

The same technique can be used to measure the identification efficiency relative to the reconstruction efficiency as well as the trigger efficiency at each trigger level. Such studies are presented in [25].

2.3 Measurement framework

Measuring an efficiency is a question of defining a very clean sample of objects and counting how many of these pass the requirements.

In our case, that means trying to find “real” high-p_T electrons coming from the hard process, and measuring how many of these are reconstructed by the software. Defining “real” electrons in data is difficult, as unlike Monte-Carlo production, we have no prior knowledge of what to expect.

In order to obtain this knowledge from data, we apply a “tag and probe” method to get a good electron sample.

2.3.1 Tag and probe principle

The principle of a tag and probe method is to select an event type which has several distinguishable signatures, one of them being the object of the investigation.

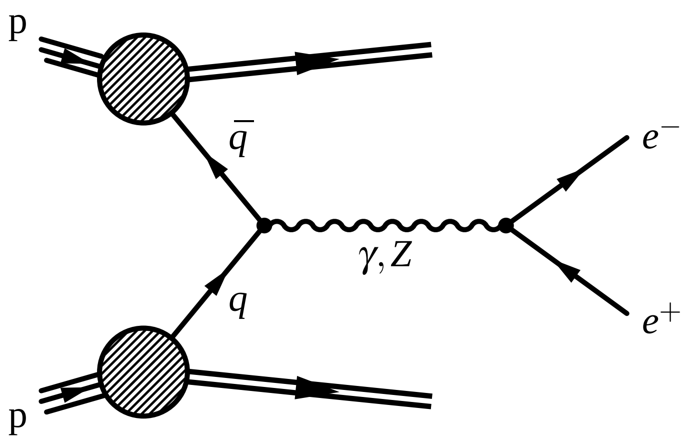


Figure 2.7: Feynman diagram of a simple $Z^0 \rightarrow e^+e^-$ event.

In our case, we elected to select the $Z^0 \rightarrow e^+e^-$ process for several reasons:

- it has a relatively large cross-section[3], higher than most of its backgrounds and will provide us with large statistics;
- no missing transverse energy;
- two high p_T leptons;
- one isolated trigger signature (isolated lepton);

We select a clean sample by using all the signatures of the event except the one under investigation, so that we can then study the remaining one separately.

The selection consists of finding events which triggered on one high energy lepton. We can then add a requirement for very little missing transverse energy (to account for small detector mismeasurements) and at least one good electron. Here, "good electron" means one reconstructed electron which passes all the identification cuts described earlier in Section 2.1.2.3.

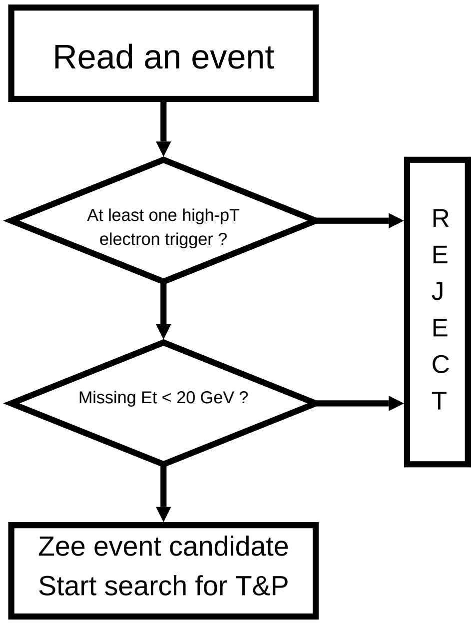


Figure 2.8: Diagram of the event selection

2.3.2 Global event selection

Given that we are looking for electron decays of the Z^0 boson, we decided to put a requirement on at least one high-p_T electron in the event.

In order to achieve that, we scan through the e-gamma stream of the ATLAS output and keep only events that pass a determined trigger chain at all three levels.

In release 13, for the CSC[4] exercise that led to the publication of the ATLAS Technical Report [25], we chose the chain corresponding to one isolated electron with a minimum p_T of 15 GeV. That choice was made because that chain was the one with the lowest threshold without being prescaled.

In release 14, for the FDR[5] exercise, for the same reasons, the chosen chain had a minimum p_T of 20 GeV.

What is more, since a $Z^0 \to e^+e^-$ event has intrinsically no missing transverse energy,

[3] we expect $\approx 1400 pb$.
[4] Computing System Commissioning
[5] Full Dress Rehearsal

we fix a maximum threshold on the measured missing E_T in the event. This introduces a minimum bias in the selection while rejecting backgrounds with neutrinos. The choice of the threshold will mostly depend on the resolution achieved in the measurement of the transverse energy.

For the CSC and FDR exercises, that maximum was fixed at 20 GeV. That threshold was chosen to keep a large signal fraction while rejecting most of the $W \to e\nu$ + fake electron background. It also limits contamination from some other processes like electronic decays of $t\bar{t}$ or $Z^0 \to \tau\tau$ events.

All these thresholds will of course need to be re-evaluated with data, once the trigger menu is finalised.

2.3.3 Object Selection

Once an event has passed the global event selection described above, we study the different reconstructed objects individually to find good candidates and reject fakes.

We first try to find our Tag, i.e. a well defined electron, before looking for a potential probe electron.

2.3.3.1 Tag selection

As described in subsection 2.3.1, we use the signature from a $Z^0 \to e^+e^-$ event to improve the purity of our sample.

We use as tight a selection as we can on the first lepton from the decay. We achieve that by imposing the following requirements on one single reconstructed electron:

1. it was reconstructed by the e-gamma algorithm;
2. it is within the tracking volume ($|\eta| < 2.4$);

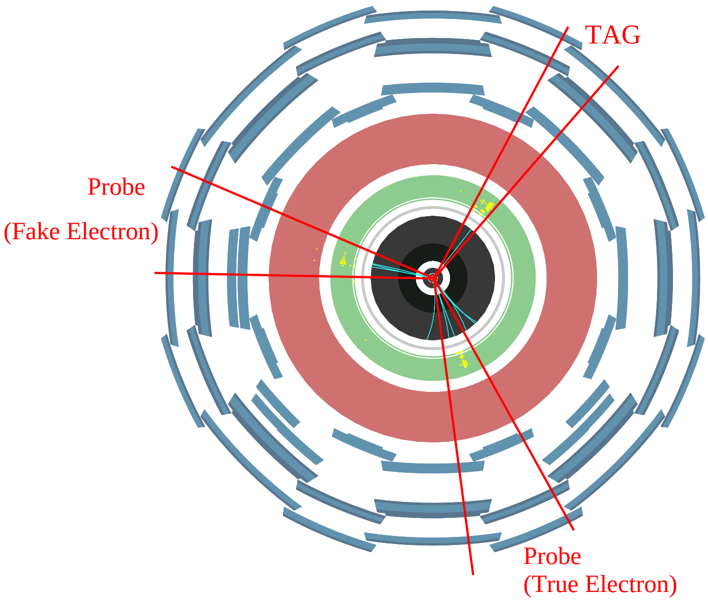


Figure 2.9: Selection of Tag and Probe object in the ATLAS detector. Event with more than one Probe candidate for one Tag.

3. it is outside the calorimeter cracks to avoid poor energy measurements ($1.37 < |\eta| < 1.52$);

4. its p_T is higher than the chosen trigger threshold;

5. it matches in ΔR, at each trigger level of our chosen chain, with an object that passed the chain;

6. it passes the tightest identification requirements we have (IsEM ElectronTight).

We consider all electrons passing every requirement mentioned above as being a very good electron candidate. We then decide to use them as one of the legs of a Z^0 decay and so "Tag" the event.

If there are more than one such electron in a event, we consider each one in turn as a new "Tag" electron. We will explain in chapter 3 how to deal with double counting issues.

After we have identified a Tag, we probe through the rest of the event for a candidate for the second leg of the Z^0 decay.

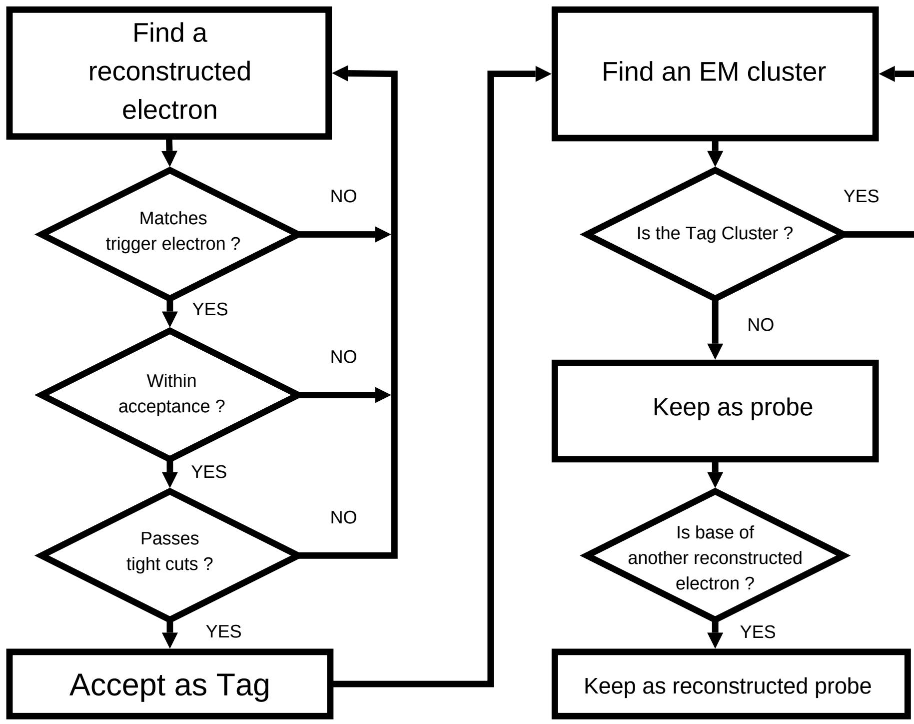


Figure 2.10: Tag and Probe selection algorithm

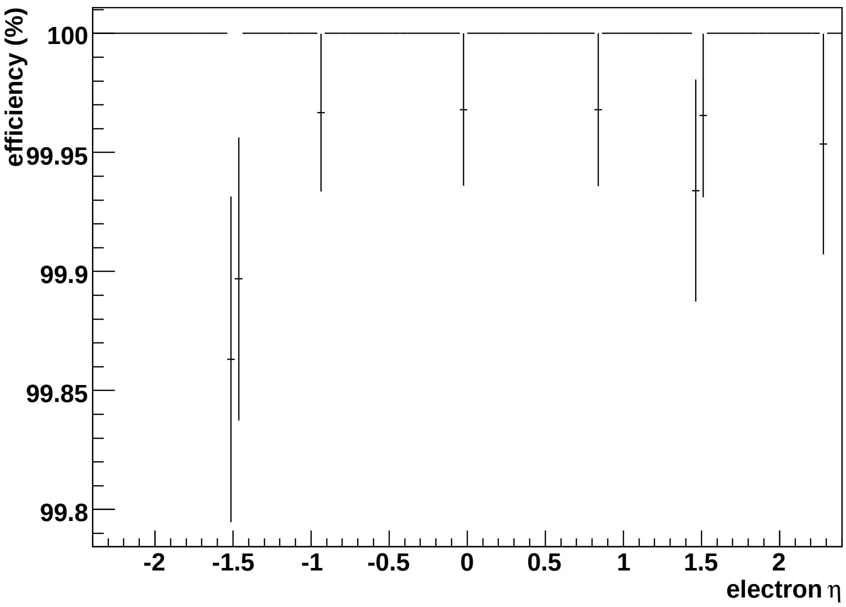
2.3.3.2 Probe selection

Ideally, we would like to measure the "absolute" electron efficiency. For that, we would need to know precisely where the other true electron is and whether it was reconstructed.

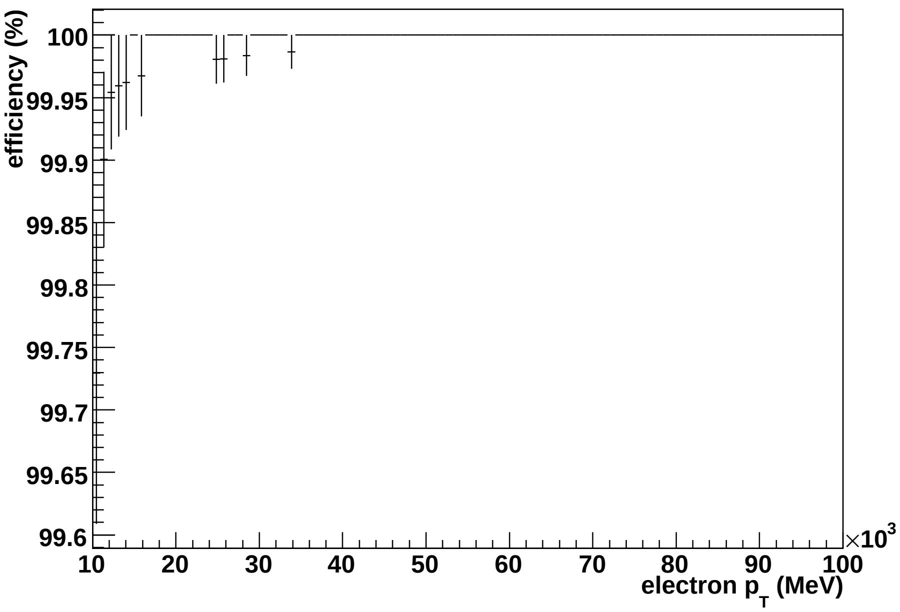
However in data this knowledge is not available and so we have to use the closest thing we can find.

In the ATLAS case, we decided to use any electro-magnetic cluster as an electron candidate. The reason is that the cluster finding algorithm has a much lower E_T threshold than the energy of the electrons coming from a Z^0 decay. As such it is expected to have a very good efficiency for high p_T electrons. A study made from Monte-Carlo ATLAS simulation, indicated that the efficiency is much better than 99.9% over most of the η range as illustrated on figure 2.11.

We have investigated further criteria that can optionally be applied, such as a minimum separation in ϕ between the Tag and the Probe. Unfortunately, this criterion was found to

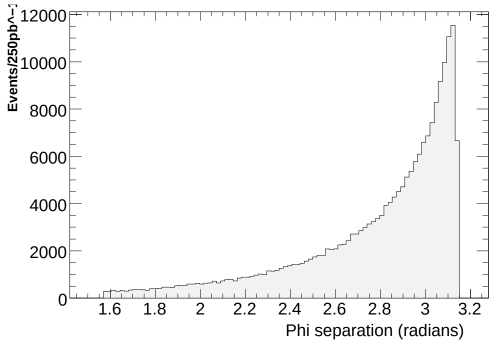


(a) cluster efficiency as a function of η ($p_T > 10$ GeV)

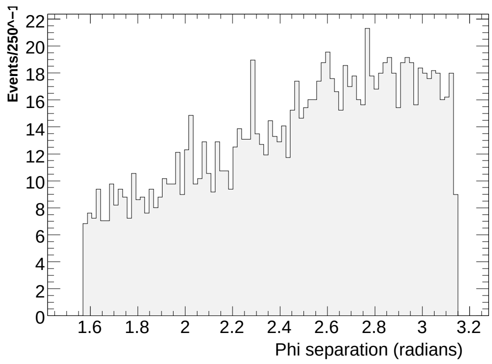


(b) cluster efficiency as a function of p_T ($|\eta| < 2.47$)

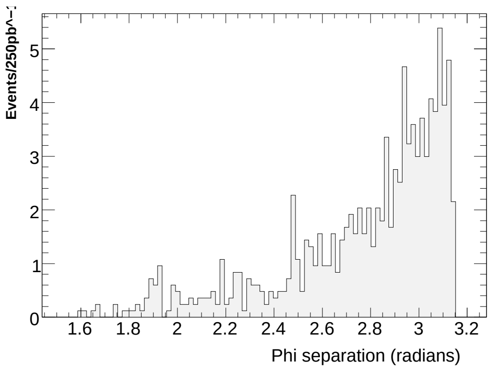
Figure 2.11: Efficiency of the clustering algorithm for high-p_T electrons.



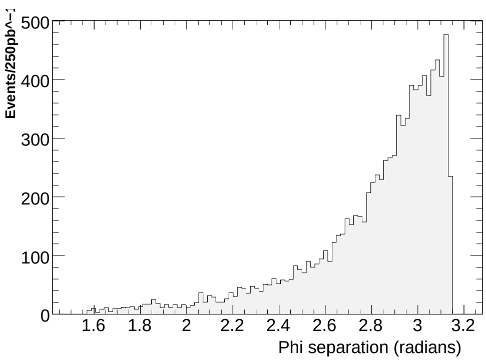
(a) $pp \to Z^0 \to e^+e^-$



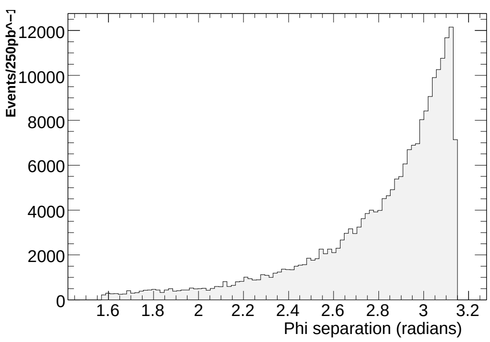
(b) $pp \to Z^0 \to \tau^+\tau^-$



(c) $pp \to t\overline{t}$



(d) $pp \to W + jets$



(e) Multijets (estimated)

Figure 2.12: distance in ϕ between the Tag and the Probe for the signal and some background samples.

criterion	Tag	Probe	reconstructed Probe
EM cluster	☑	☑	☑
Minimum E_T	☑	☑	☑
$\|\eta\|$ cut	☑	☑	☑
Track matching	☑	☐	☑
Tight identification	☑	☐	☐
Passed trigger	☑	☐	☐

Table 2.8: Informations available and criteria passes by each Tag or Probe object

have a low discriminating power against the most common backgrounds (see Figure 2.12). What is more, it may be problematic in case of highly boosted Z^0 events, as the angle can be reduced by the boost.

Once we have selected a Tag and a probe, we can reconstruct the Z^0 they supposedly originated from. We will later use the invariant mass line shape to statistically separate the signal from the background.

2.3.4 Probe matching criterion

Once we have identified a suitable probe candidate, we should have a criterion to check whether that candidate was indeed reconstructed as an electron by the software. That is achieved by checking whether there is a reconstructed e-gamma electron whose cluster is our probe cluster.

Other versions of this test could rely on $\Delta\eta$-$\Delta\phi$ matching, but this relies on a matching parameter that will be hard to estimate from data, and it does not account for duplicate cluster finding problems. It could associate more than one cluster with one reconstructed electron.

Table 2.8 summarises the different cuts applied to our different objects. It reminds us that the set of Tags is a subset of both probes and reconstructed probes.

2.4 Summary

This chapter presented the ATLAS hardware and the software framework for electron reconstruction. It also described the aim of the electron reconstruction efficiency measurement.

The next chapter will describe the mathematical framework in which the measurement is performed as a test of what we can achieve with the ATLAS detector.

Chapter 3

Calculating the electron efficiency

> This chapter provides a proof of the mathematical framework we use to extract the efficiency from data. It assumes we have at our disposal Tag and Probe pairs as described in the previous chapter. The method to remove the background to get to the true number of Tag and Probe pairs is described in the next chapter.

Once we have identified Tag and Probe pairs, we need to find the correct way to extract the efficiency from the numbers we collected.

This chapter describes the derivation and calculation used to extract the efficiency. We calculate the efficiency in two cases:

- the overall efficiency, i.e. integrated over all other variables;
- differential efficiencies, i.e. as a function of a given variable.

3.1 Extracting the overall efficiency

This section will describe the proper way to extract the efficiency from the tag and probe selection. To make things clearer, we will first describe the way to get the efficiency in an inclusive case, before we expand the method to the differential case which is more useful in practice.

3.1.1 Definition of the parameters

After background subtraction (see chapter 4), once we have isolated $Z^0 \to e^+e^-$ events, we can extract from our selection:

- the total number of tag and preselected probe pairs: N_1;
- the number of tag and probe pairs in the subset where the probe is reconstructed as an electron : N_2.

The measurement of the efficiency ε we are looking for, is then a function of two other parameters:

- T : the efficiency for a reconstructed electron to pass the tighter Tag cuts (trigger and Identification cut requirement).
- N : the total number of Z^0 events where both our electrons were in acceptance (η, p_T, missing $E_T$requirements);

In all the following computations, the following assumptions are made:

Charge symmetry we assume that the efficiency to reconstruct an electron is the same as the efficiency to reconstruct a positron. This will be checked, with data, by measuring separately the efficiency for pairs with electron tags and for pairs with positron tags. If the efficiencies agree, as they do in the Monte-Carlo simulation, then the two samples will be merged to improve the statistics;

Tags are subset of probes which means that we can factorize the efficiency to reconstruct a tag as the efficiency to reconstruct a probe times a “tagging efficiency”, which is the efficiency to pass the trigger and the tight identification cuts. See Table 2.8 for the differences between tag and probe conditions.

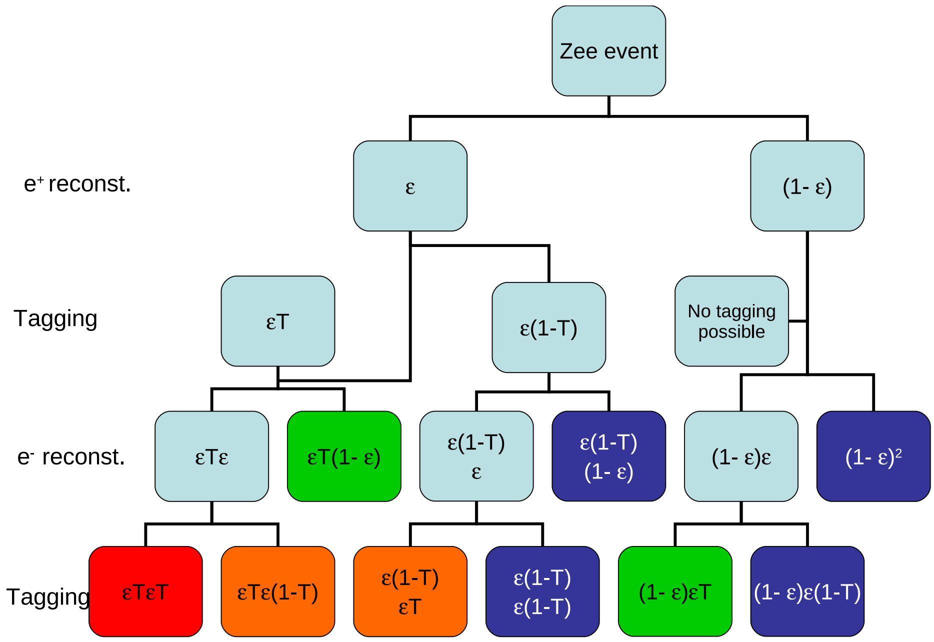


Figure 3.1: Probability for an event to pass different selections. Colours refer to the cases enumerated in Section 3.1.2.

3.1.2 Derivation of the inclusive efficiency

We can then distinguish four kinds of events in our signal sample from their signature in our detector : see Figure 3.1:

1. The events we are blind to, because they have no reconstructed electron or none that pass the Tag selection (dark blue);

2. The events where only a Tag is found, but the probe is not reconstructed (green);

3. The events where only one of the electrons pass the Tag selection and where the probe is matched to an offline electron (orange);

4. The events where both of the electrons can pass the Tag selection and, in both cases, the probe is matched (red).

The last category of event is double counted as we use each electron as a Tag in turn as we explained in section 2.3.3.1.

Hence, the total number of pairs we will see will be

$$N_1 = N_{green} + N_{orange} + 2 \cdot N_{red}, \tag{3.1}$$

whereas the number of tag and reconstructed probe pairs seen is

$$N_2 = N_{orange} + 2 \cdot N_{red}. \tag{3.2}$$

From the Figure 3.1, we extract that from a total number N of $Z^0 \to e^+e^-$ events seen at ATLAS, we get:

$$\begin{aligned} N_{green} &= 2 \cdot N \cdot T\varepsilon(1-\varepsilon); \\ N_{orange} &= 2 \cdot N \cdot (T\varepsilon \cdot \varepsilon(1-T)); \\ 2N_{red} &= 2 \cdot N \cdot (T\varepsilon)^2; \end{aligned}$$

Hence, we end up with: $N_1 = 2NT\varepsilon$ and $N_2 = 2NT\varepsilon^2$, so that simple algebra gives us the calculation for the inclusive electron efficiency ε:

$$\frac{N_2}{N_1} = \frac{2NT\varepsilon^2}{2NT\varepsilon}$$

$$\Rightarrow \varepsilon = \frac{N_2}{N_1} \tag{3.3}$$

The errors on the two numbers N_1 and N_2 are correlated seeing as N_2 is a subsample of N_1. Errors can be computed from N_2 and $N_3 = N_1 - N_2$ (sample of non reconstructed electrons). However, toy Monte-Carlo tests shows that the errors on the ratio is overly conservative when calculated from N_2 and N_3. A better approximation is obtained from a binomial error compuation using N_1 and N_2.

Unfortunately, this number is not very useful in practice. Indeed, it is very dependant on the distribution of the electrons originating from the Z^0 decay. If the electron efficiency is very dependent on the electron p_T, as we suppose it is, based on previous experiments, a sample with a different distribution would modify this value. Even the simple bias introduced by the p_T threshold of the trigger can lead to different results.

As a consequence, it is clear that an inclusive efficiency is of limited use to the collaboration, and we must try to measure the efficiency differentially, as a function of, e.g., the pseudo-rapidity or the momentum.

3.2 Deriving a differential efficiency

Ideally, we would like to get a functional form for the efficiency. Unfortunately, neither the method nor the statistics we will collect in the first years enable us to do so. As such, we will describe a method to extract a binned efficiency, with the knowledge that as the bin size reduces, we will be less and less dependent on the distribution of our sample. The previous section can be described as a particular case of this one with only a single bin.

We call this parameter-dependent efficiency a “differential efficiency” by analogy with the “differential cross-section” terminology. Yet, while in the cross-section case, it is purely a differential form, the “differential efficiency” must be convoluted by the weighted number of events in each bin to get the complete efficiency:

$$\varepsilon_{inclusive} = \frac{1}{N} \int \frac{\partial \varepsilon}{\partial \eta} \delta(\eta) d\eta,$$

where the measure $\delta(\eta)$ is the number of events where the probe has this value of η.

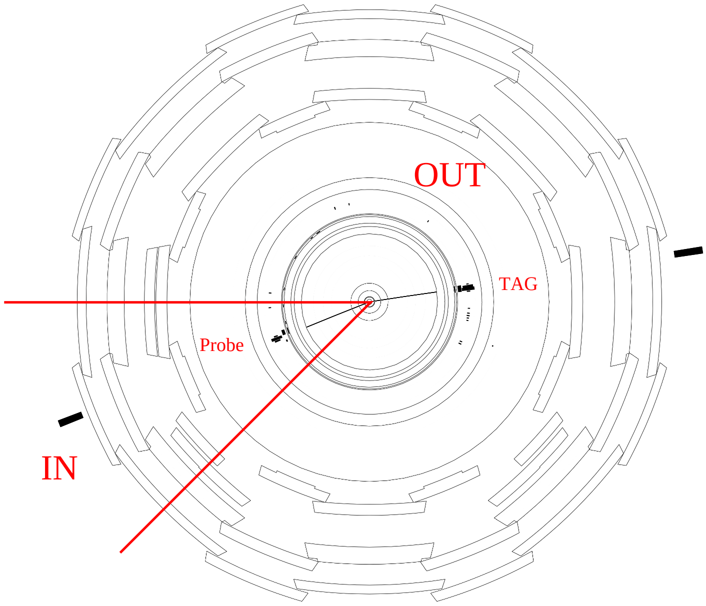


Figure 3.2: Studying the efficiency in a region in ϕ of the detector

3.2.0.1 Definition of the parameters

Once we have isolated Z^0 events, we select one bin of the distribution to study. We can then extract from our selection:

- the total number of Tag&Probe pairs where the probe is in this bin: N_1;
- the number of Tag&Probe pairs where the probe is in this bin and where the probe is reconstructed as an electron : N_2.

This case needs to be broken in two separate cases.

- The case where both the Tag and the probe fall into the same bin (*in* case);
- The case where the Tag and the probe fall into different bins (*out* case), as in Figure 3.2.

The first case is similar to the one treated in the previous section. The second case is only slightly different. We must first introduce new parameters: two parameters describing the "inside" region and three describing the "outside" region to measure ε, the electron efficiency of the bin under investigation ;

- T^{in} : the efficiency for a reconstructed electron to pass the tighter Tag cuts in the current bin;
- N^{in} the number of events where both electrons are in the same bin;
- ε^{out} : the inclusive electron efficiency for this sample outside of the bin under investigation ;
- T^{out} : the efficiency for a reconstructed electron to pass the tighter Tag cuts outside of the reconstruction region.
- N^{out} the number of events where one electron is in the right bin and one is outside;

3.2.1 Computing the efficiency

We can then reconstruct the two numbers N_1 and N_2 as defined in section 3.1 from these parameters.

N_1 is constructed from two contributions.

1. The one where the Tag is outside: N_1^{out};
2. the one where the both are in the same bin: N_1^{in}

N_1^{in} is computed in the same way as for the inclusive case, as it is exactly the same situation. We can just take equation 3.1 and replace the variables accordingly.

$$\begin{aligned}
N_1^{out} &= N^{out} \cdot \varepsilon^{out} T^{out} \\
N_1^{in} &= N^{in} \cdot \left(2\varepsilon(1-\varepsilon) T^{in} + 2\varepsilon^2 T^{in}(1-T^{in}) + 2\varepsilon^2 T^{in^2} \right) \\
&= 2N^{in} \varepsilon T^{in} \\
N_1 &= N_1^{in} + N_1^{out} \\
&= N^{out} \varepsilon^{out} T^{out} + 2N^{in} \varepsilon T^{in}
\end{aligned}$$

Similarly, N_2 is constructed from two contributions as well.

1. The one where the Tag is outside: N_2^{out};

2. the one where the both are in the same bin: N_2^{in}.

Again, N_2^{in} yields a similar form to equation 3.2.

$$\begin{aligned}
N_2^{out} &= N^{out} \cdot \varepsilon^{out} T^{out} \cdot \varepsilon \\
N_2^{in} &= N^{in} \cdot \left(2\varepsilon^2 T^{in}(1-T^{in}) + 2\varepsilon^2 T^{in^2} \right) \\
&= 2N^{in} \varepsilon^2 T^{in} \\
N_2 &= N_2^{in} + N_2^{out} \\
&= \left(N^{out} \varepsilon^{out} T^{out} + 2N^{in} \varepsilon T^{in} \right) \varepsilon
\end{aligned}$$

From these formulae, it is obvious that the efficiency we are looking for can again be extracted by the simple formula:

$$\frac{N_2}{N_1} = \frac{\left(N^{out} \varepsilon^{out} T^{out} + 2N^{in} \varepsilon T^{in} \right) \varepsilon}{N^{out} \varepsilon^{out} T^{out} + 2N^{in} \varepsilon T^{in}}$$

$$\Rightarrow \varepsilon = \frac{N_2}{N_1} \tag{3.4}$$

This solution is exactly the same as the one from the inclusive case (see equation 3.3). This is because the inclusive case is just a special case, where the *in* region is the whole detector over the whole kinematic range.

3.2.2 Usage

As we can see from equation 3.4, the result does not depend on the efficiencies outside of the current bin ε^{out}. As such it is independent, to the first order, from the distribution of the Tag electrons. The only dependence left is due to the correlations between the two electrons.

3.3 Summary

This chapter presented the mathematical framework used to extract the electron efficiency. We first presented the simpler but less interesting case of an inclusive efficiency in the $Z^0 \rightarrow e^+e^-$ channel, then the much more useful binned efficiency computation. The only limitation to the binning is given by the statistics and our capacity to subtract the background.

The next chapter will describe the method used to extract the number of signal pairs from the total number of Tag and Probe pairs in each bin.

Chapter 4

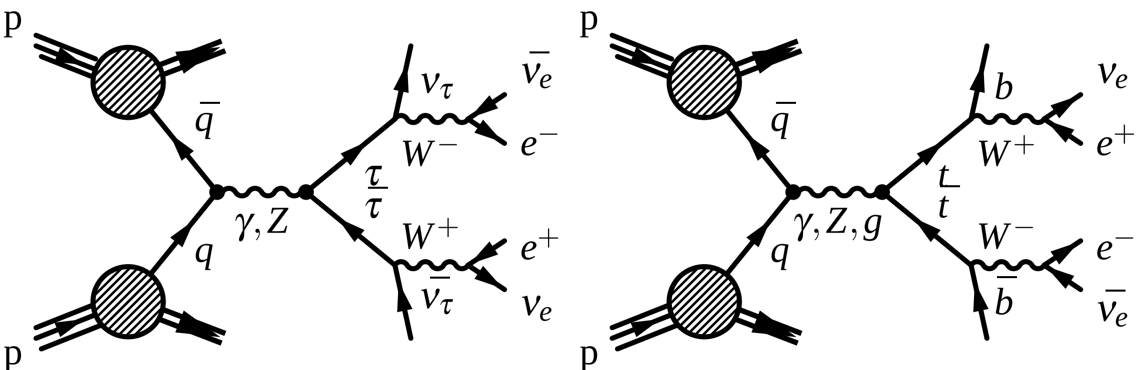
Background subtraction to $Z^0 \rightarrow e^+e^-$

This chapter describes the steps taken to separate backgrounds from signal in order to extract the number of real Tag and Probe pairs in our sample. We discuss the different backgrounds and their contribution, as well as a method to assess the systematics associated with this method.

Even after the steps we described in chapter 2, we are still left with some remaining background contamination. Some background indeed have the exact same signature as the $Z^0 \rightarrow e^+e^-$process. This chapters describes the processes we have identified that will account for sizeable background to our measurements, the steps we have taken to subtract them, and finally, how we have prepared for the background estimation using Monte-Carlo samples.

4.1 List of $Z^0 \rightarrow e^+e^-$ backgrounds

We have identified a few backgrounds that can fake a $Z^0 \rightarrow e^+e^-$ signal . Some are *real* in the sense that they really have the same true signature as our signal, i.e., two high-p_T electrons with low missing transverse energy. Some are faking one or more electron, in most cases a jet is wrongly identified as an electron by our software.



(a) $pp \to Z^0 \to \tau^+\tau^-$ (b) $pp \to t\bar{t}$

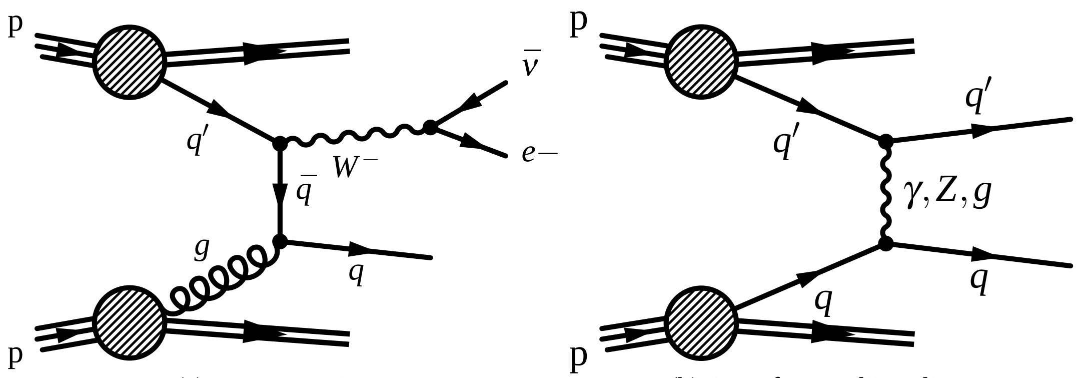
Figure 4.1: Feynman diagrams of the main backgrounds with two real electrons.

We will list below what we identified as the main backgrounds to our measurement. The list is not exhaustive, but it is believed that the cross-sections of the omitted processes are too low to significantly affect the precision of the measurement in the first few years. The contributions from the included backgrounds already present very varied numbers of events. The expected contributions from the main backgrounds are plotted in figure 4.3. They are extracted by running the selection described in Chapter 2 on samples containing only the quoted process. A description of the datasets used for these is given in Appendix A.

A small background also comes from $Z^0 \to e^+e^-$ event, where one of the true reconstructed electrons is associated with a cluster from a non-electron object. This background is called combinatorial and has a small efect on the reconstructed Z^0 mass line shape in the $Z^0 \to e^+e^-$ sample.

4.1.1 *Real* backgrounds

These backgrounds are the backgrounds with a large enough cross-section, which produce two high-p_T electrons. They are backgrounds in the sense that they distort the Z^0 mass resonance peak and would need to be accounted for in a cross section measurement. For the efficiency measurement, given that they do produce two real electrons, they can be accepted as signal.



(a) $pp \to W + jets$ (b) One of many dijets diagrams

Figure 4.2: Example Feynman diagrams of the main backgrounds from fakes.

4.1.1.1 $pp \to Z^0 \to \tau^+\tau^-$

The $Z^0 \to \tau^+\tau^- \to e^+\nu_e\overline{\nu_\tau}e^-\overline{\nu_e}\nu_\tau$ can under some circumstances have a low missing transverse energy. This leaves just two isolated high-p_T electrons. The invariant mass line shape contribution from this background is a peak which is much wider than the $Z^0 \to e^+e^-$ resonance and at a lower energy so that the contamination to the $Z^0 \to e^+e^-$ peak is small.

4.1.1.2 $pp \to t\bar{t}$

The $t\bar{t} \to (e^+\nu_e b)\left(e^-\overline{\nu_e}\bar{b}\right)$ may also appear to have a low missing transverse energy. Once the detector is well understood, we may be able to use the b-tagging to reject these events. This will not be desirable at the beginning of the detector operation and would also be detrimental to Z^0+jets studies. The di-electron mass spectrum peaks right under the Z^0 peak, but it is wider and the small cross section reduces its contribution to a fraction of our total number of events. We expect a good representation of that background in our Monte-Carlo simulations.

4.1.2 Reducible backgrounds from fake electrons

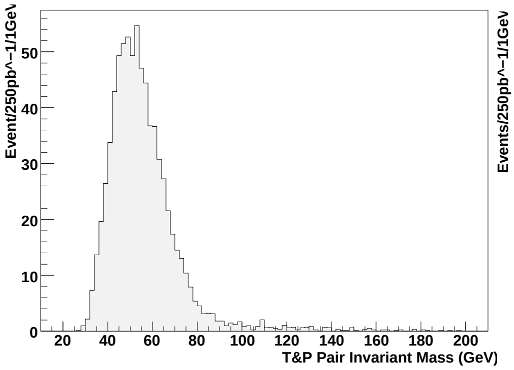
These backgrounds are backgrounds which do not have the right signature. They are missing one, or two, of the electrons, but because another object in the event fakes an electron, they are picked up by our selection. Despite the rejection power we expect from our event selection, these processes have high enough cross-sections that they still represent a sizeable contribution. The normalisation of these contributions is difficult to determine from Monte-Carlo simulations. We will need to determine the rejection power of our requirements from data before we can determine their contribution accurately. Due to their large cross-sections, it is difficult to simulate enough events to represent a sizeable luminosity. We will describe a way to compensate for the shortage of simulated events in Section 4.4.

4.1.2.1 $pp \rightarrow W + jets$

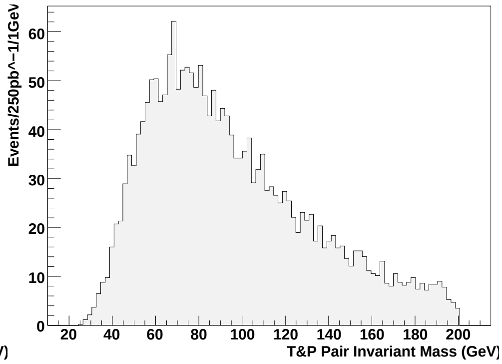
$W \rightarrow e\nu$ signal is normally rejected by our missing E_T requirements, but if the neutrino is emitted in a direction parallel to the beam pipe, the apparent missing transverse energy will be low. The Tag and trigger requirements can then be fulfilled by the lepton from the W, and it is possible for an additional jet to fake an electron. From Monte-Carlo simulation we expect the contribution to have similar features as the $t\bar{t}$ one, but a much higher contribution due to the difference in cross-section.

4.1.2.2 Multi-jet final states

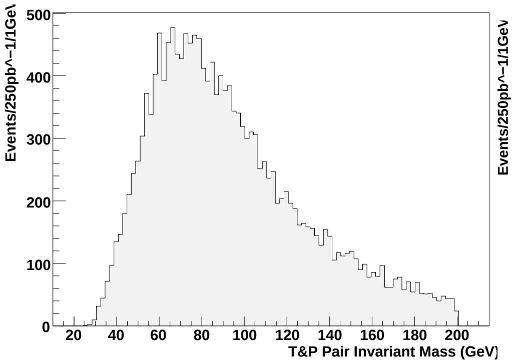
These processes do not have any real high-p_T electrons. But as their cross-sections will be so much higher than any other single process, it is expected that the contribution of fake leptons will still be our largest background, in particular the single fake electron associated with a cluster contribution. This will obviously depend on the rejection power of the selection criteria we apply.



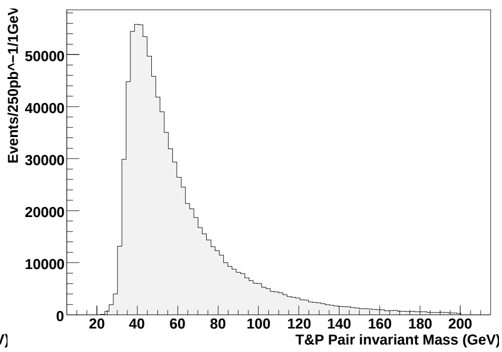
(a) $pp \rightarrow Z^0 \rightarrow \tau^+\tau^-$



(b) $pp \rightarrow t\bar{t}$



(c) $pp \rightarrow W + jets$



(d) estimated dijets shape and normalisation

Figure 4.3: Background contributions from our main backgrounds for a luminosity of $250pb^{-1}$. The spectra are extracted from datasets containing only the mentioned process.

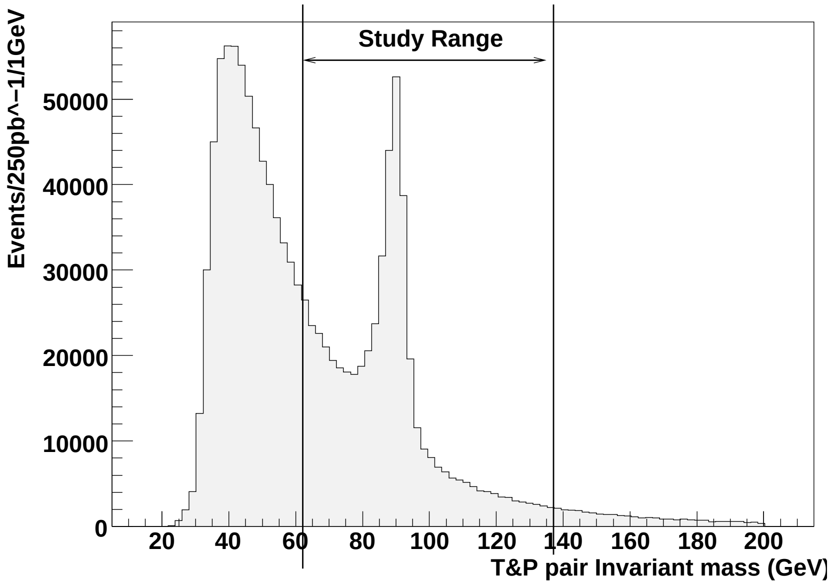


Figure 4.4: Invariant mass of Tag and Probe pairs as determined in a Monte-Carlo exercise.

4.2 Background removal algorithms

Now that we have identified which backgrounds will contribute to our measurement, we must describe the steps we plan to use to remove their contributions from our measurement. This section will describe two possible techniques.

4.2.1 Data preparation

After we have collected di-electron pair candidates, we make use of the Z^0 mass line shape to separate the background from the signal.

Suppose we want to study the number of Z^0 events with certain characteristics, e.g. one of the electrons in a certain η range, or events with at least two jets. We extract the subset of pairs were that condition was fulfilled, and we plot the invariant mass of the pair. This gives us an invariant mass line shape as in Figure 4.4.

Since the turn on of the curve is determined by the threshold effect in the trigger and the software, we place a lower cut to remove the lower tail that is not easily modelled. Similarly, we place a high cut to remove the high mass region were the statistics is expected to be poor. We then end up with a relatively large range which contains both background and signal. We have different options to separate the two contributions.

4.2.2 Background contribution determination

We have identified two different techniques to determine the background contribution to our sample. Both make the assumption that the background distribution is falling exponentially in our range. The function modelling the background will obviously need to be checked with data as it becomes available.

4.2.2.1 Z^0 mass line shape fit

The first technique consists in using a fit describing the peak and the background to determine the contribution from each one.

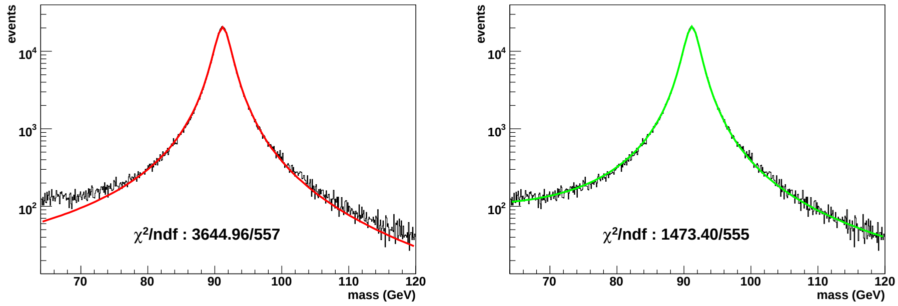
We use the Minuit fitter in ROOT for all the technical aspect of the fitting process.

The background contribution is modelled by an exponential $e^{-a \cdot m}$. We have checked that this does describe the shape of our Monte-Carlo background correctly. We will use the first data to try and determine how good this distribution matches the line shape on a large distribution range and modify the model accordingly.

The signal is modelled based on three contributions:

- the truth Z^0 mass distribution from the theory;
- a correction for radiative loses on the electrons and detector resolution.

The true Z^0 mass distribution is simply modelled by the theoretical relativistic Breit-Wigner distribution as seen on Figure 4.5a.



(a) Original Breit-Wigner function (b) Breit-Wigner function plus γ^* exchange $BW + \exp^{-\alpha x}$

Figure 4.5: Possible true Z^0 mass line shapes.

$$BW(m, m_{Z^0}, \Gamma_{Z^0}) = \frac{\Gamma_{Z^0}^2 m_{Z^0}^2}{\left[(m^2 - m_{Z^0}^2)^2 + \Gamma_{Z^0}^2 m_{Z^0}^2\right]}$$

We see from Figure 4.5b that the side bands are better described by adding an exponential function describing the photon echange term. However, we elected not to account for this term as it is smeared by the resolution function and then it represents a smaller effect than the background contribution, even than from combinatorial mis-association in the Z^0 sample.

Another possible contribution to the theoretical line shape is a correction due to the parton distribution functions. Because low-x collisions are more frequent than energetic ones, the shape is distorted by a factor which is an inversely proportional power function of the boson mass ($\times m_{Z^0}^{-\beta}$). But again the effect is smeared by the detector resolution and there would be a strong correlation between the parameters of the resolution function and β if this contribution were added to the fit parameters.

The correction for radiation losses and energy resolution is modelled by a Crystal Ball distribution [26]. This function was developed by the Crystal Ball collaboration to model energy losses. It consists of a Gaussian core with a power law low tail while keeping a continuous first derivative. The low tail models final states radiations and bremsstrahlung losses.

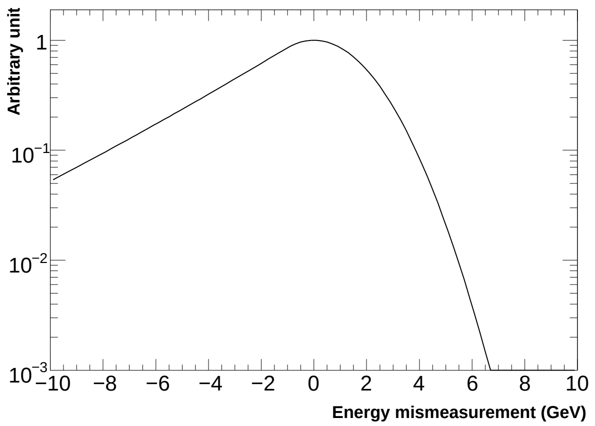


Figure 4.6: Crystal-Ball function with arbitrary parameters.

The parameters $\bar{x}$ and σ describe the behaviour of the energy resolution. $\bar{x}$ describes the offset on the energy scale, while σ describes the resolution. α and n model the radiation loses.

$$CB(x,\alpha,n,\bar{x},\sigma) = \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), & for \quad \frac{x-\bar{x}}{\sigma} > -\alpha \\ A\cdot\left(B-\frac{(x-\bar{x})}{\sigma}\right)^{-n}, & for \quad \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}$$

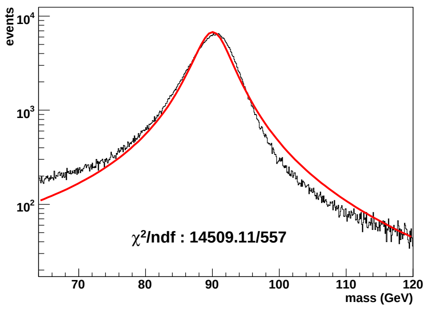
where

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{\alpha^2}{2}\right) \\ B &= \frac{n}{|\alpha|} - |\alpha| \end{aligned}$$

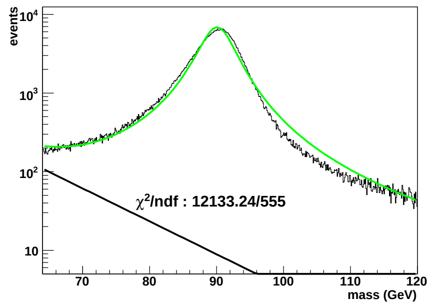
The final form is obtained by convoluting the Breit-Wigner distribution in mass with the Crystal Ball distribution, to account for the detector effects. Figures 4.7a and 4.7b show that the simple Breit-Wigner distribution is not good enough for describing the peak asymmetry arising from energy losses in the detector.

The convolution describes the peak much better (Figure 4.7c) but underestimates the side bands from the combinatorial background and photon exchange term. The addition of an exponential background corrects this discrepancy as seen from Figure 4.7d.

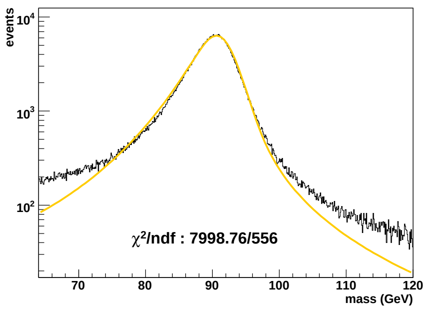
The value of $\bar{x}$ in the Crystal Ball definition is fixed at zero, as the convolution implies that the values of $\bar{x}$ and m_{Z^0} are correlated. By fixing $\bar{x}$, we ensure that the fitted value of m_{Z^0}



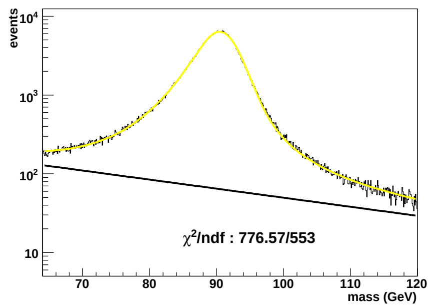
(a) Simple Breit-Wigner fit



(b) Breit-Wigner fit plus background



(c) convolution fit



(d) convolution fit plus background

Figure 4.7: Possible reconstructed Z^0 mass line shapes.

is shifted by the energy scale offset. We let the mass m_{Z^0} run free to account for this wrong energy scale calibration. The width Γ_{Z^0} is fixed to the LEP measured value as we use the Crystal Ball σ to account for energy resolution.

$$\begin{aligned} S(m, m_{Z^0}, \beta, \alpha, n, \bar{x}, \sigma) &= [BW(y, m_{Z^0}, \Gamma_{Z^0}) * CB(y, \alpha, n, 0, \sigma)](m) \\ &= \int BW(y, m_{Z^0}, \Gamma_{Z^0}) \cdot CB(m-y, \alpha, n, 0, \sigma) dy \end{aligned}$$

We can then extract the number of signal events in a narrow window around the peak by two different means:

- integrate the signal function;
- integrate the background function and subtract it from the total number of events.

In both cases, we restrict ourselves to a narrower window than the one over which the fit is performed to avoid any edge effects and to stay in a region where the signal to background ratio is slightly better. This consists of performing a fit on a large window 64 GeV–120 GeV, but only integrating the signal in a region of 80 GeV–100 GeV.

4.2.2.2 Side band extrapolation

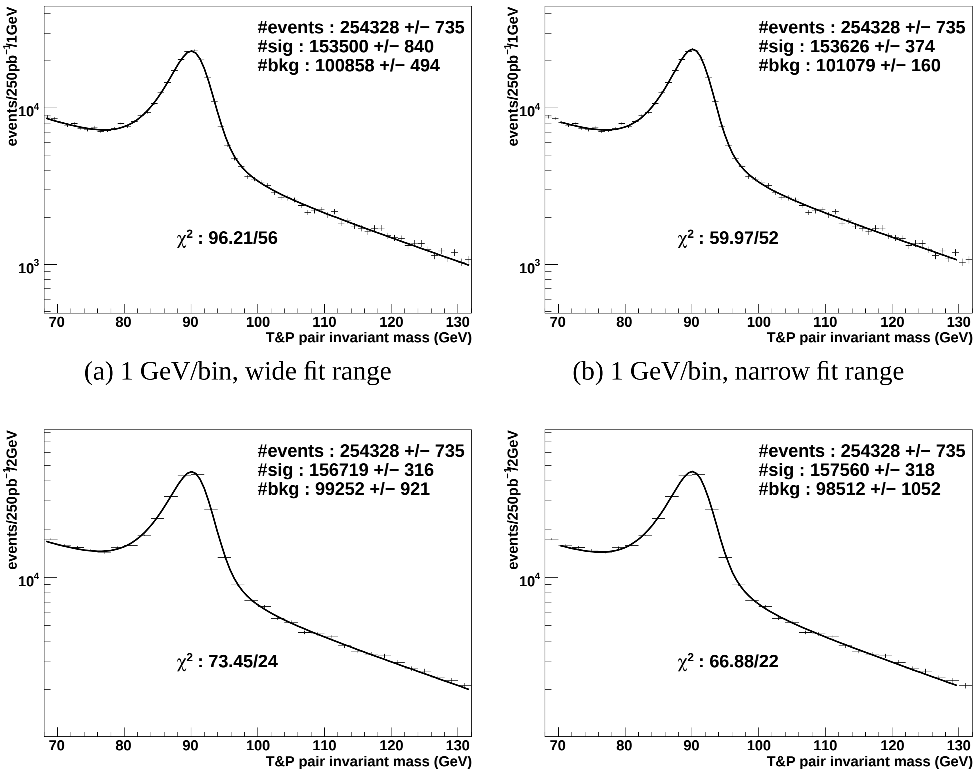
In cases where the turn on threshold of the mass distribution is low enough, and we have sizeable side bands far from the peak where the signal contribution is negligible on both sides, we can extrapolate the background under the peak assuming a background distribution shape. An example of such a configuration is shown on Figure 4.4.

This method will be difficult to use at the start of ATLAS operation, as there is no measure of the quality of the extrapolation to warn us if our assumptions are wrong, unlike the χ^2 for the fit. If the background is well described, however, it has the advantage to require much less statistics to give accurate results compared to the fit. Unfortunately, as the trigger threshold increases, the turn-on curve gets closer to the Z^0 peak and the low mass side band may become unusable.

4.3 Background removal systematics estimation

The process we just described in section 4.2 depends on some parameters that are not determined from data but are defined manually in the code:

- the large range over which we study the mass distribution;
- the smaller range we consider to be our peak region;
- the bin width when we plot our distribution for the fit method;
- the distribution functional form.



(a) 1 GeV/bin, wide fit range

(b) 1 GeV/bin, narrow fit range

(c) 2 GeV/bin, wide fit range

(d) 2 GeV/bin, narrow fit range

Figure 4.8: Fits using different parameters. The quoted numbers are the number of pairs in a range of 80–100 GeV around the Z^0 mass.

While the functional form is not something easily modified, since it relies on our understanding of the measurement, the other parameters are really arbitrary and only roughly optimized for speed and statistical reasons.

To measure the effect of such parameters, we perform the same measurement multiple times by varying slightly the conditions:

binning effect we double the number of bins;

fit effect we reduce the fit range by one bin on each side;

integration effect we reduce the peak range by one bin on each side.

Ideally, we should of course get the same answer every time, except when we reduce the peak range. In practice, the answers are different, in some cases by quite a large margin.

The Figure 4.8 show one such spectrum fitted in four different configurations. The quoted numbers are the number of pairs in a range of $80-100$ GeV around the Z^0 mass.

We decided to quote as the result for number of pairs, the weighted average of all these measurements as some fits fail due to poor statistics, and, as a systematic error, the standard deviation between these. We keep as a statistical error the error from the nominal set of parameters, as it is the one with the largest statistics, therefore the one whose fit is the best constrained.

In conclusion, we have completely described a method to extract the number of di-electron pairs produced through a $Z^0 \to e^+e^-$process from a background contaminated sample. This includes a measure of the systematic error linked with the separation from the background and the signal.

4.4 Evaluation of the background from Monte-Carlo simulations

The plots we have shown so far in figures 4.3 or 4.8 show the expected contribution of the background in Monte-Carlo simulations of the ATLAS detector.

Most of these were extracted by fully simulating the outcome of the given process as it passes through the detector. We simulated more than $250pb^{-1}$ worth of $Z^0 \to e^+e^-$ events. We then took events from our main backgrounds scaled to the same luminosity. In cases where we did not produce enough events to match the required integrated luminosity, we attributed weights to the events we collected to correct for the luminosity ratio.

However, this is a highly time consuming process. For a process like dijet production, it was computationally impossible to simulate enough events to match the simulated luminosity for the $Z^0 \to e^+e^-$ sample. Even with more than six million simulated events, the total dijets sample only represent around $0.02pb^{-1}$. Even at the lower luminosity expected at start-up,

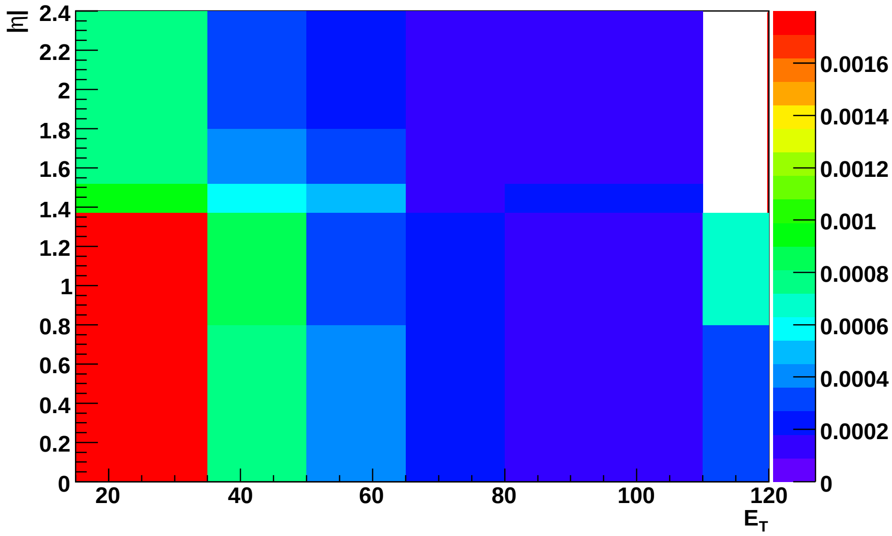


Figure 4.9: Probability for a jet to fake a electron passing all identification cuts (IsEM ElectronTight).

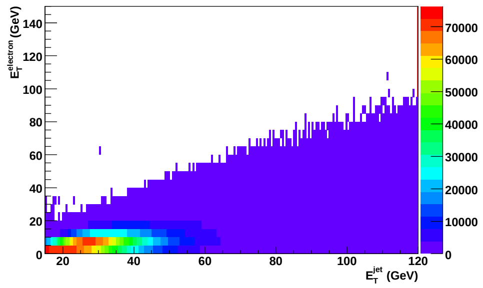
this represents but a few minutes of running.

The main problem is that, if we attribute a weight to the events passing our whole selection, the weight is very large (≈ 50000) and we end up with a spectrum with very sharp peaks which is un-physical and is quite clearly an artefact of our simulation shortcomings.

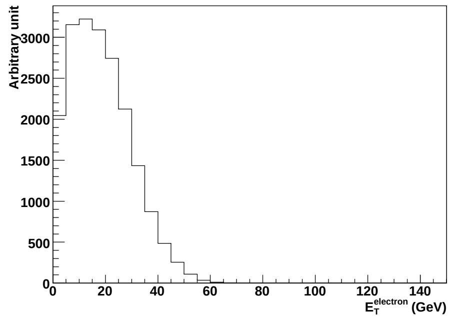
We therefore needed to devise a technique to get a smooth estimation of the background shape and normalisation. We elected to use a method based on fake rate estimation.

4.4.1 Electron fake rate estimation

The main reason we get sharp peaks in our estimation of the background from the dijets process, is that our selection has a large rejection factor against dijets. The background contribution mainly comes from the fact that the dijets cross-section is so large. One way to get a smooth background estimate is that instead of having a weight of 1 for events passing our selection and a weight of 0 for the others, we can attribute to an event a weight corresponding to the *probability* that an event of this type passes our selection.



(a) fake electron E_T vs. E_T^{jet}



(b) fake electron E_T distribution for a jet of $E_T = 80$ GeV

Figure 4.10: Energy scale corrections for fake electrons.

Say we have an event with two jets with properties (E_T^1, η^1) and (E_T^2, η^2). We can then compute the probability for each of the jets to fake an electron with our different requirements. The weight of an event is computed by the probability of one of the jets to fake the first electron multiplied by the probability for the second jet to fake the other. As an example, Figure 4.9 shows the probability for a jet to fake an electron passing all the identification cuts. This is plotted as a function of E_T and η, as these two variables are the ones which are expected to have the most influence on the fake rate.

The difference with the number quoted in table 2.6 comes from the use of a different sample. Table 2.6 was computed from a broad simulation of jets in the ATLAS detector, whereas figure 4.9 was computed from a sample to which a designated filter was applied. The aim of the filter is to select generator level jets which present an increased probability to fake leptons, hence the higher fake rate in this sample. The rate is only moderately higher because the filter only needs one jet to pass the selection whereas the fake rate is calculated with all jets in the sample.

4.4.2 Energy scale corrections estimation

We still need to correct the reconstructed energy, as the electro-magnetic energy scale will be different from the jet one. To do just that, we plot a two dimensional histogram of the jet

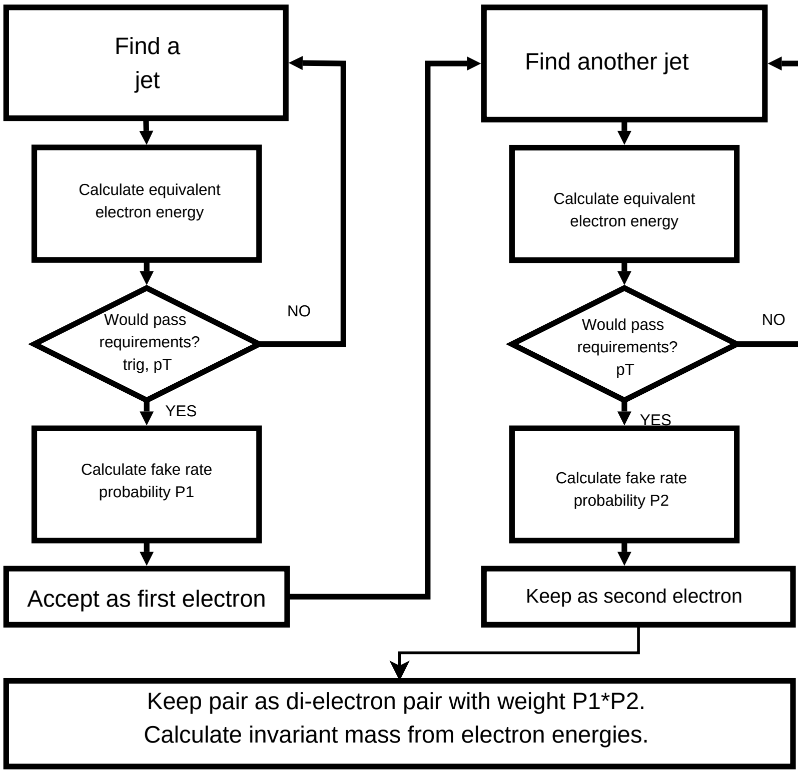


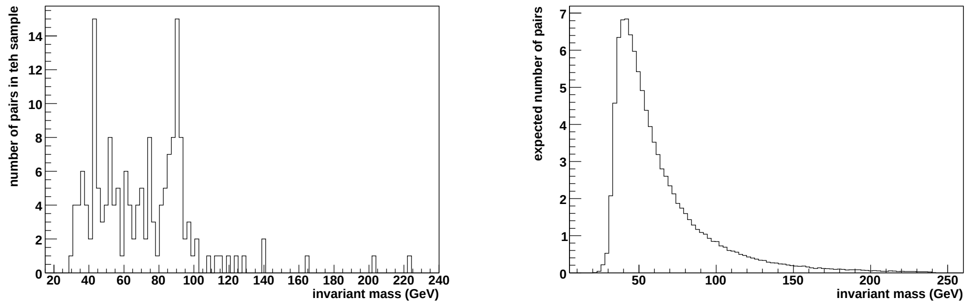
Figure 4.11: Algorithm to get the background from a large dijets sample.

E_T against the fake electron E_T for all jets that were falsely reconstructed as an electron as in Figure 4.10a. Then, when we get any jet with a transverse energy of E_T^{jet}, we select a slice of that histogram along E_T^{jet}, which gives us the distribution of the electron energy for such a jet as in Figure 4.10b. We then pick a random number from that distribution. This allows us to get the right distribution for the electron energies.

4.4.3 Background normalisation and shape estimation

The algorithm used to extract the background shape from a large background sample is described in figure 4.11.

For each event in the sample, we loop over the jets in the event. For each jet, we create a fake electron using the energy correction described in section 4.4.2. We then check whether that electron would pass the kinematic acceptance cuts for the first electron in our normal analysis. If it does, we keep that electron and calculate the probability for that jet to pass the



(a) original selection on dijets sample (b) smooth background estimate from fake rate

Figure 4.12: Dijets background contribution.

other requirements on the electron (trigger levels, identification flags) using the appropriate fake rate as described in 4.4.1.

We then go over all the other jets in the same event trying to find another potential electron. For each, we create again a fake electron. Just like for the first one, we check this fake electron would pass the selection for the second electron in our analysis (cluster, p_T, η, etc.). Again we calculate the probability for that jet to fake our second electron selection.

Now armed with both electrons, we save the pair as a di-electron candidate with a weight equal to the probability that both jets were faking the right kind of electron. This method provides us with a very large number of pairs, but each with a very low weight, hence providing us with a smooth spectrum with a low normalisation.

To get the normalisation right, we add the requirement that our smooth spectrum must have the same integral as the ragged one we get from the normal analysis on the same dataset so that we end up with a smooth spectrum representing the integrated luminosity of our original sample. This indicated that the fake rate method get the normalisation wrong by 3%. We can then just multiply the weight of the whole sample to match the luminosity of our signal sample without risking isolated sharp peaks as we can see in Figure 4.12.

The peak at the Z^0 mass in Figure 4.12a, is due to the presence of a few ($\sim$ 10 in 4 million) real Z^0 production during the scattering process. This effect is overwhelmingly dominated by the fake jets which is why is does not appear in our estimate. What is more, such events contain a real Z^0 boson decaying to electron and should be counted as signal, not background.

Chapter 5

Measuring the electron reconstruction efficiency during computing exercises

This chapter describes the attempt we made at measuring the electron reconstruction at ATLAS as described in Section 2.2 during two separate computing exercises in 2007-2008. It presents the measured efficiency comparing it with the Monte-Carlo truth when available. The efficiency is presented differentially when the statistics permit it.

During the run-up to the LHC startup, the ATLAS community has been preparing for data taking by organizing computing exercises. The idea is to use Monte-Carlo simulated data to train and assess the different algorithms needed to reconstruct events and prepare the necessary tools for physics analyses. A first series of exercises was organized, before 2006, under the name Data Challenge whose aim was to assess the scalability of the ATLAS computing architecture to the volume of data expected from the experiment.

Then, between 2006 and 2008, the Computing Software Challenge (CSC) was meant to test all aspects of the ATLAS software from triggers to complete physics analyses. The results were then compiled by the collaboration and were published as *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics* [25].

Finally, in 2008, another exercise was set-up to test the data pipeline and reconstruction infrastructure. The so-called Full Dress Rehearsal (FDR) simulated a few hours of ATLAS operation at low luminosity.

5.1 Electron reconstruction efficiency measurement during the CSC exercise

The CSC exercise[1] was geared toward preparing for physics analyses. As such, data was prepared to simulate most expected processes that we should be able to study at ATLAS, from well-known SM processes like $W^+ \to e^+ \nu_e$ to different beyond the SM processes like SUSY. These samples are produced separately to be able to study precisely the contribution of each on a given analysis. They obviously need to be combined according to their relative cross-section and sizes.

5.1.1 Samples and software versions considered

As we explained in chapter 4, one signal sample and four backgrounds are considered. Our signal is obviously $Z^0 \to e^+e^-$ for the reasons we explained in section 2.3.1, while the four backgrounds are :

- $pp \to Z^0 \to \tau^+\tau^-$;
- $pp \to t\bar{t}$;
- $pp \to W + jets$;
- $pp \to jets$

The list of datasets used and the software versions for the different simulation steps is given in Appendix A.

[1]While the original samples used in the CSC exercise and published in [25] were generated with a centre of mass energy of 14 TeV and were simulated and reconstructed with previous versions of the ATLAS software, the results presented in this section have been updated to use data samples generated with a centre of mass energy of 10 TeV and simulated and reconstructed with more recent (Release 14) software.

Process	number of events	cross-section (pb)	filter efficiency
$pp \to Z^0 \to e^+e^-$	909909	1143.96	0.86
$pp \to Z^0 \to \tau^+\tau^-$	97907	1128.37	1.0
$pp \to t\bar{t}$	1027675	373.59	0.55
$pp \to W \to e\nu$	964526	11764.6	0.88
$pp \to jet\ +\ jet$	1003804	$1.461 \cdot 10^9$	0.0706

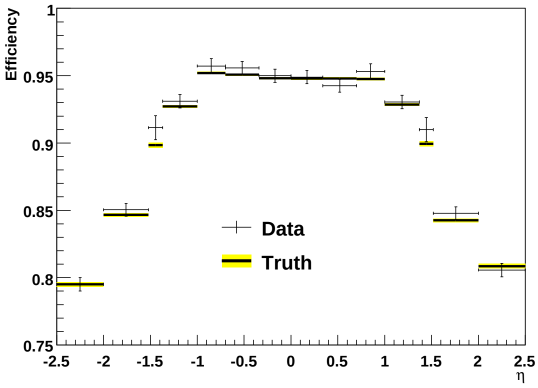
Table 5.1: Cross-sections and sizes of the samples used for the CSC like analysis.

We processed the event samples listed in Table 5.1 and using the appropriate weights for each event to get an equivalent sample of integrated luminosity 924 pb^{-1}.

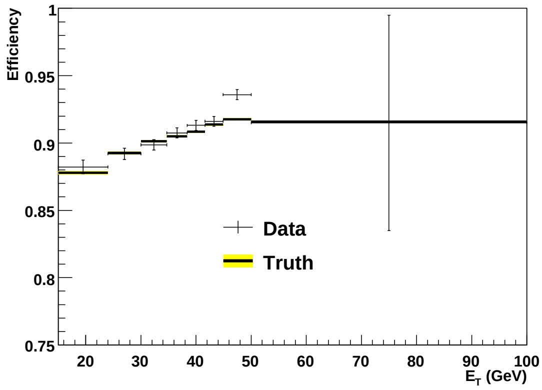
All the results presented in this chapter have been obtained using Athena version 14.2.25.0 and ROOT version 5.18.

The trigger menu (list of triggers) used in these samples is the one expected to be used by ATLAS once the LHC reaches a luminosity of 10^{32} $cm^{-2}s^{-1}$. In this menu, the lowest non-prescaled trigger chain is calibrated for an isolated 15 GeV electron. It consists of one 13 GeV region of interest (L1_EM13I) and two high-level trigger objects (L2_e15i_medium and EF_e15i_medium). This dictates that our acceptances will be for electrons above 15 GeV and with $|\eta| < 2.5$.

All results in this chapter are compared to "Truth" information. We selected events for which both truth electrons passed the same acceptance cuts as the reconstructed ones to account for biases due to the selection as explained in Chapter 2. Then we studied which fraction was effectively reconstructed as offline electrons, thus yielding the "true electron efficiency". The criteria for electron resconstruction is that a reconstructed electron can be found within $\Delta R < 0.2$ of the true electron in the $\eta - -\phi$ plane. This is plotted in black with yellow statistical errors throughout this chapter.



(a) Electron efficiency as a function of η



(b) Electron efficiency as a function of E_T

Figure 5.1: Electron efficiencies for a signal only sample. Errors include systematics.

5.1.2 Signal only tests

This first section compares the result we get from our algorithm if we only consider the signal sample. It can be considered an ideal case where our selection (2.3) is perfect and we only analysed $Z^0 \rightarrow e^+e^-$events. There are nevertheless a few fake e^+e^- pairs remaining due mainly to wrong combinations between a real tight electron and a cluster coming from another object, e.g. a jet.

Hence we still perform the background removal explained in section 4.2.2.1. The combinatorial background is small compared to the one originating from other processes but should not be discarded. For the overall sample, we estimated the number of background events to 12121 ± 849 out of 483831 ± 696 events whose electron-cluster pair has an invariant mass in the range 82–100 GeV (2.5%) and 10500 ± 774 out of the 439835 ± 663 events whose cluster is matched to an electron (2.38%).

This background contribution in the simulation is small, but the addition of pile-up once the LHC reaches the design luminosity will increase this contribution because of our very loose requirements on the probe electron.

Table 5.2 and Figure 5.1a show the measured efficiency compared to the Monte-Carlo truth as a function of η. Table 5.2 also shows the difference between the two as the ra-

η	Measured			Monte Carlo		difference
	mean	stat.	syst.	mean	stat.	(σ)
-2.50–-2.00	79.51	0.22	0.44	79.50	0.17	0.01
-2.00–-1.52	85.04	0.18	0.44	84.66	0.14	0.63
-1.52–-1.37	91.14	0.26	0.86	89.85	0.20	1.18
-1.37–-1.00	93.11	0.13	0.49	92.71	0.10	0.64
-1.00–-0.70	95.72	0.11	0.53	95.20	0.09	0.83
-0.70–-0.34	95.57	0.10	0.47	95.07	0.08	0.90
-0.34–0.00	95.00	0.11	0.47	94.82	0.08	0.31
0.00–0.34	94.89	0.11	0.47	94.79	0.08	0.18
0.34–0.70	94.26	0.11	0.46	94.80	0.08	- 0.96
0.70–1.00	95.32	0.12	0.56	94.76	0.09	0.84
1.00–1.37	93.05	0.13	0.49	92.86	0.10	0.31
1.37–1.52	90.99	0.26	0.86	89.93	0.19	0.98
1.52–2.00	84.78	0.18	0.44	84.26	0.14	0.84
2.00–2.50	80.57	0.21	0.45	80.85	0.17	-0.43

Table 5.2: Electron efficiency as a function of η for a signal only sample

tio $\frac{\varepsilon_{data}-\varepsilon_{mc}}{error_{data}}$. The quoted systematic uncertainty is associated with the background subtraction and the different choices we made for the binning of distribution.

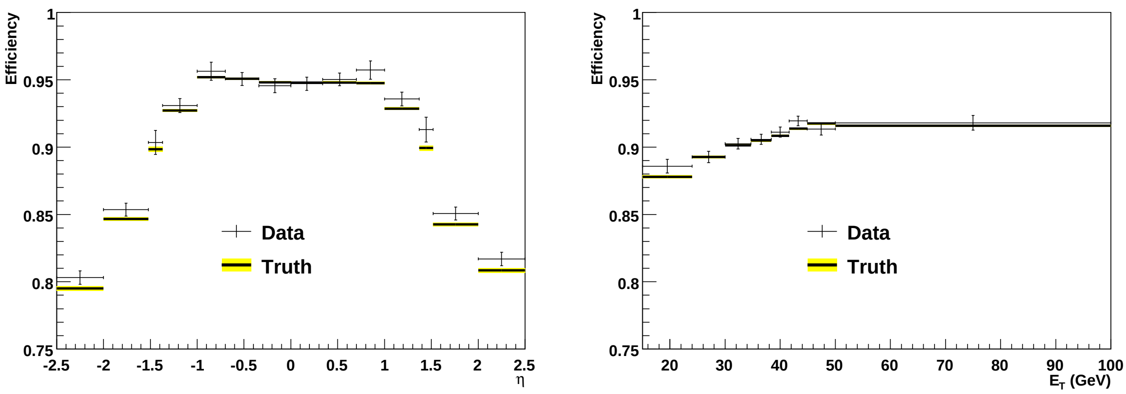
The fact that the measured uncertainty is consistently within one sigma from the MC truth indicates that the errors on the measured and MC values are correlated (mostly the statistical component) but proves our capacity to conduct the measurement with accuracy.

What is more, the asymmetry observed in η is an artefact of the geometry used in the CSC exercise. Additional dead material was introduced in the simulation for half the detector to study its impact on the performance. In practise, after having ascertained that the detector is symmetric using this method, we will fold the detector to study $|\eta|$, hence doubling our statistics.

The results agree over the whole range, proving our capacity to measure the efficiency in a clean sample with minimum bias. The binning was chosen to match the geometry of the detector while maintaining high enough statistics in each bin. The crack region in the calorimeter ($1.37 < \eta < 1.52$) is treated separately because the resolution is expected to be much worse and so it disturbs the line shape of the Z^0 resonance.

E_T (GeV)	Measured			Monte Carlo		difference
	mean	stat.	syst.	mean	stat.	(σ)
15.00–24.00	88.22	0.17	0.48	87.80	0.12	0.68
24.00–30.00	89.20	0.14	0.40	89.27	0.10	-0.14
30.00–34.70	89.86	0.12	0.37	90.14	0.09	-0.58
34.70–38.40	90.75	0.11	0.36	90.49	0.09	0.56
38.40–41.60	91.31	0.11	0.36	90.85	0.09	1.00
41.60–44.90	91.61	0.10	0.35	91.38	0.09	0.52
44.90–50.00	93.59	0.09	0.36	91.75	0.08	3.99
50.00–100.00	91.50	0.11	8.00	91.58	0.08	-0.01

Table 5.3: Electron efficiency as a function of E_T (GeV) for a signal only sample



(a) Electron efficiency as a function of η

(b) Electron efficiency as a function of E_T

Figure 5.2: Electron efficiencies for a sample without dijets. Errors include systematics.

Considering the distribution as a function of E_T (see Figure 5.1b and Table 5.3), the agreement is again very good, with an overestimation at large values of E_T ($41\ GeV < E_T < 50\ GeV$).

5.1.3 Signal and non-QCD backgrounds

As we explained in section 4.4, QCD represents a major challenge for the simulation of ATLAS data due to its large cross-section at the LHC. Since it undergoes a particular treatment, we will consider it separately. As Table 5.1 shows, we analysed enough events from our other backgrounds so that we can just reweigh our events without creating artefacts from outstanding events through reweighting (artificial spikes in distributions).

η	Measured			Monte Carlo		difference
	mean	stat.	syst.	mean	stat.	(σ)
-2.50–-2.00	80.31	0.21	0.46	79.50	0.17	1.20
-2.00–-1.52	85.36	0.18	0.44	84.66	0.14	1.14
-1.52–-1.37	90.35	0.27	0.85	89.85	0.20	0.46
-1.37–-1.00	93.09	0.13	0.50	92.71	0.10	0.61
-1.00–-0.70	95.63	0.11	0.66	95.20	0.09	0.57
-0.70–-0.34	95.06	0.11	0.47	95.07	0.08	-0.01
-0.34–0.00	94.56	0.11	0.50	94.82	0.08	-0.43
0.00–0.34	94.71	0.11	0.47	94.79	0.08	-0.14
0.34–0.70	95.03	0.11	0.47	94.80	0.08	0.42
0.70–1.00	95.72	0.11	0.66	94.76	0.09	1.26
1.00–1.37	93.58	0.13	0.49	92.86	0.10	1.18
1.37–1.52	91.31	0.26	0.89	89.93	0.19	1.23
1.52–2.00	85.07	0.18	0.44	84.26	0.14	1.32
2.00–2.50	81.70	0.21	0.46	80.85	0.17	1.27

Table 5.4: Electron efficiency as a function of η for a sample without dijets.

In the case where we collected a sample containing our signal along with our three other backgrounds, the contribution of background increased to 22841 ± 957 out of 494473 ± 754 simple electron-cluster pairs (4.61%) and to 19404 ± 845 out of 449027 ± 716 matched pairs (4.32%). This doubles our background contribution but has limited impact on our capacity to measure the efficiency.

As we can see from Figure 5.2a, the addition of these backgrounds has little impact on our mean estimate. We observe a systematic overestimation in the end-cap. Comparing the results from Table 5.4 with Table 5.2, we can notice a small increase in the systematic errors coming from a higher uncertainty on the background estimation.

Figure 5.2b and Table 5.5 again show an excellent agreement except at very large E_T. This can be explained by looking at the fit. At large E_T the background events do not follow the exponential background description as we can see on Figure 5.3.

This leads to a poor background estimation in that region and, as a consequence, biases the result. This bias is anticipated to distort the spectrum as our selection of a Tag with a minimum E_T of 15 GeV and a probe of minimum 50 GeV the phase-space for low-mass

E_T (GeV)	Measured			Monte Carlo		difference
	mean	stat.	syst.	mean	stat.	(σ)
15.00–24.00	88.59	0.17	0.47	87.80	0.12	1.27
24.00–30.00	89.27	0.14	0.40	89.27	0.10	0.00
30.00–34.70	90.25	0.12	0.37	90.14	0.09	0.23
34.70–38.40	90.58	0.11	0.36	90.49	0.09	0.19
38.40–41.60	91.12	0.11	0.36	90.85	0.09	0.58
41.60–44.90	91.94	0.10	0.35	91.38	0.09	1.27
44.90–50.00	91.35	0.11	0.43	91.75	0.08	-0.77
50.00–100.00	91.81	0.11	0.54	91.58	0.08	0.35

Table 5.5: Electron efficiency as a function of E_T (GeV) for a sample without dijets.

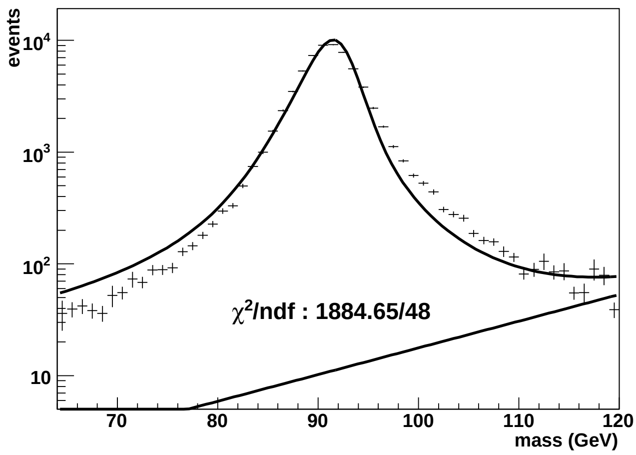


Figure 5.3: Invariant mass spectrum of tag and probe pair which have $44.90\,GeV < E_T^{probe} < 50\,GeV$ Sample includes signal and non QCD backgrounds. The straight line represents the background contribution evaluated from the fit.

Sample	All pairs	Match pairs
Signal only	471710 ± 487	429337 ± 399
Signal and non-QCD backgrounds	471632 ± 589	429623 ± 449
All backgrounds	469527 ± 833	422794 ± 1342

Table 5.6: Estimated number of signal events for different sample mixing

pairs is much reduced. Only boosted pairs where the tag and the probe are bunched together will fill the left side of the spectrum.

High-energy fakes are less likely, as they mostly punch through to the hadronic calorimeter, but they still distort the spectrum. We didn't take any action at this point to do a special fit for these bins as the effects of QCD fake still needs to be added and are expected to dominate.

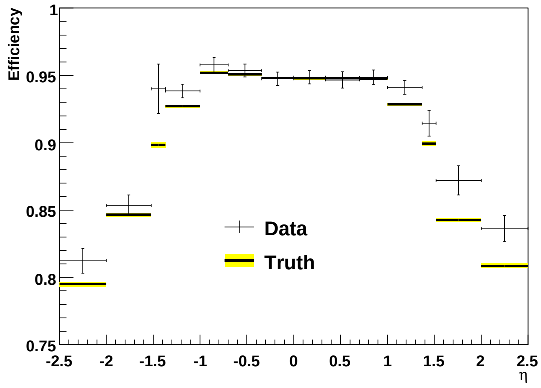
5.1.4 Complete sample with estimated QCD background

As we explained in section 4.4, the contribution from QCD has to be estimated in our simulation data. It is by far the largest contribution to our background, in particular for low energy probes, since most fakes will come from low energy jets.

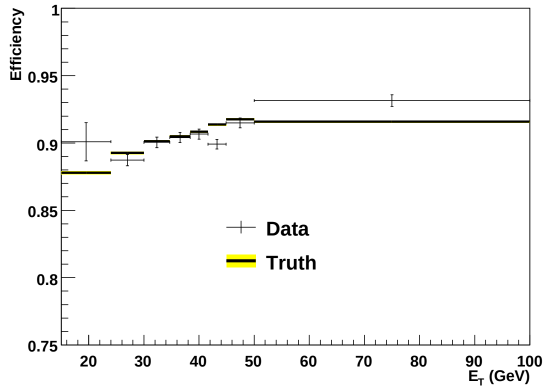
After we added in the QCD, the contribution of background increased to 226818 ± 2484 out of 696345 ± 2123 simple electron-cluster pairs (33.04%) and to 207016 ± 2512 out of 629870 ± 2123 matched pairs (32.86%). This huge increase in background is the main worry at the LHC as the QCD contribution at these energies suffers from very large uncertainties.

In this exercise, as we can see from Table 5.6, despite the fact that the signal to noise ratio dropped from 40:1 to 2:1 with the introduction of different background, we managed to keep a consistent estimate of the *real* number of pairs overall (1.5% variation).

The main effect of the QCD background addition is a large increase in the systematic error for all distributions. Large background effects happen at large pseudo-rapidity and low E_T, where most of the fakes are concentrated. Apart from the η and E_T distribution, Table 5.9 shows the effect of correlations between the two variables. The *diff* column indicates



(a) Electron efficiency as a function of η



(b) Electron efficiency as a function of E_T

Figure 5.4: Electron efficiencies for a realistic sample including an estimated QCD background. Errors include systematics.

the difference between the measured efficiency and the truth in terms of the error on the measured efficiency.

Just like for the non QCD case, the degradation of performance at low and high E_T comes from distortions to the line shape as illustrated on Figure 5.5.

At low E_T (Figure 5.5a), the signal to noise ratio becomes an issue. In this case in particular, where our QCD background estimation method implies a large uncertainty for each bin, the fit quality cannot be assessed effectively. Considering that the absolute level of the background in data is still relatively unknown due to the uncertainty on the cross-section, one may have to raise the threshold at which to select its electrons.

At high E_T (Figure 5.5b), the background distribution looks like a flat pedestal. Once again, it is poorly described by our falling exponential and a different function will be used in this case if data proves to follow this behaviour. A test performed on the spectrum shown on Figure 5.6 shows that changing the background from an exponential to a quadratic function ($f(m) = a \cdot m^2 + b \cdot m + c$) improves the χ^2 for the fits (85.92/48 down from 94.03/48). The uncertainty is still large, in particular at low and high invariant mass. Once first data is available, systematic studies will be carried on to find the function giving the best description of the line shape for each regime.

η	Measured			Monte Carlo		difference
	mean	stat.	syst.	mean	stat.	(σ)
-2.50–-2.00	81.24	0.21	0.89	79.50	0.17	1.60
-2.00–-1.52	85.36	0.18	0.75	84.66	0.14	0.77
-1.52–-1.37	94.01	0.21	1.84	89.85	0.20	2.04
-1.37–-1.00	93.84	0.12	0.49	92.71	0.10	1.86
-1.00–-0.70	95.78	0.11	0.54	95.20	0.09	0.91
-0.70–-0.34	95.37	0.10	0.46	95.07	0.08	0.54
-0.34–0.00	94.75	0.11	0.48	94.82	0.08	-0.11
0.00–0.34	94.87	0.11	0.48	94.79	0.08	-0.14
0.34–0.70	94.67	0.11	0.60	94.80	0.08	-0.19
0.70–1.00	94.86	0.12	0.53	94.76	0.09	0.15
1.00–1.37	94.11	0.12	0.50	92.86	0.10	2.02
1.37–1.52	91.45	0.28	0.92	89.93	0.19	1.32
1.52–2.00	87.20	0.17	1.07	84.26	0.14	2.41
2.00–2.50	83.62	0.21	0.94	80.85	0.17	2.44

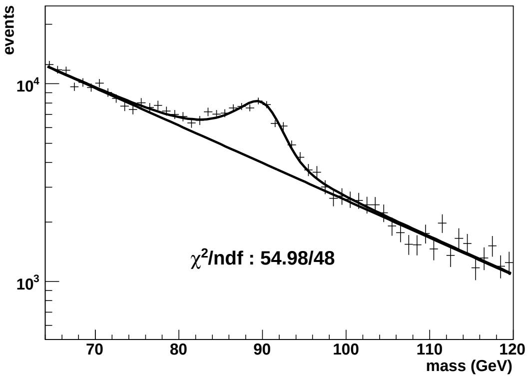
Table 5.7: Electron efficiency as a function of η for a realistic sample including an estimated QCD background.

E_T (GeV)	Measured			Monte Carlo		difference	χ^2/NDF
	mean	stat.	syst.	mean	stat.	(σ)	
15.00–24.00	90.10	0.16	1.41	87.80	0.12	1.50	54.98/48
24.00–30.00	88.73	0.14	0.40	89.27	0.10	-1.03	57.38/48
30.00–34.70	90.05	0.12	0.37	90.14	0.09	-0.19	55.59/48
34.70–38.40	90.41	0.12	0.36	90.49	0.09	-0.18	50.82/48
38.40–41.60	90.66	0.11	0.35	90.85	0.09	-0.41	97.87/48
41.60–44.90	89.91	0.11	0.34	91.38	0.09	-3.29	90.64/48
44.90–50.00	91.49	0.11	0.35	91.75	0.08	-0.58	307.66/48
50.00–100.00	93.15	0.10	0.42	91.58	0.08	3.04	94.03/48

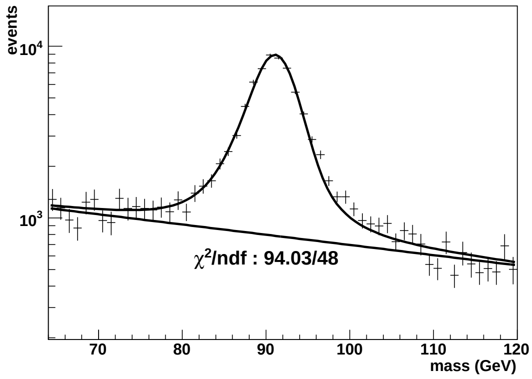
Table 5.8: Electron efficiency as a function of E_T (GeV) for a realistic sample including an estimated QCD background.

$\|\eta\|$ \ E_T (GeV)	15.00–32.40				32.40–42.43				42.43–75.00			
	mean	stat.	syst.	diff.(σ)	mean	stat.	syst.	diff.(σ)	mean	stat.	syst.	diff.(σ)
0.00 – 0.70	95.40	0.11	0.52	2.15	94.24	0.09	0.37	1.21	95.42	0.08	0.44	0.06
0.70 – 1.37	92.31	0.15	0.64	0.33	92.69	0.11	0.39	1.98	93.04	0.11	0.41	3.57
1.37 – 1.52	89.88	0.37	1.79	0.61	91.35	0.29	0.95	1.18	91.54	0.32	1.06	0.15
1.52 – 2.00	81.71	0.26	1.48	0.78	84.06	0.21	0.50	0.11	86.63	0.20	0.59	0.81
2.00 – 2.50	83.02	0.24	1.50	1.86	80.58	0.26	0.54	0.54	79.60	0.28	0.57	1.28

Table 5.9: Electron efficiency as a function of E_T (GeV) and $|\eta|$ for a realistic sample including an estimated QCD background.



(a) 15 $GeV < E_T^{probe} < 24$ GeV



(b) 50 $GeV < E_T^{probe} < 100$ GeV

Figure 5.5: Invariant mass spectrum of tag and probe pairs. Sample includes all background samples. The straight line represents the background contribution evaluated from the fit.

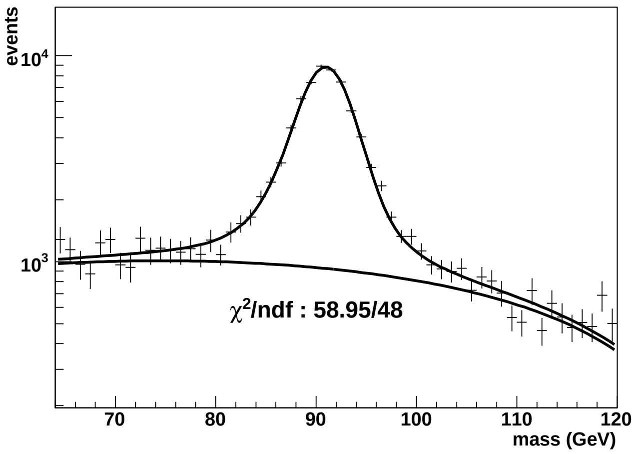


Figure 5.6: Invariant mass spectrum of tag and probe pair which have $E_T^{probe} > 50$ GeV. Sample includes all background samples. Background fit is described as a parabola. The parabola represents the background contribution evaluated from the fit.

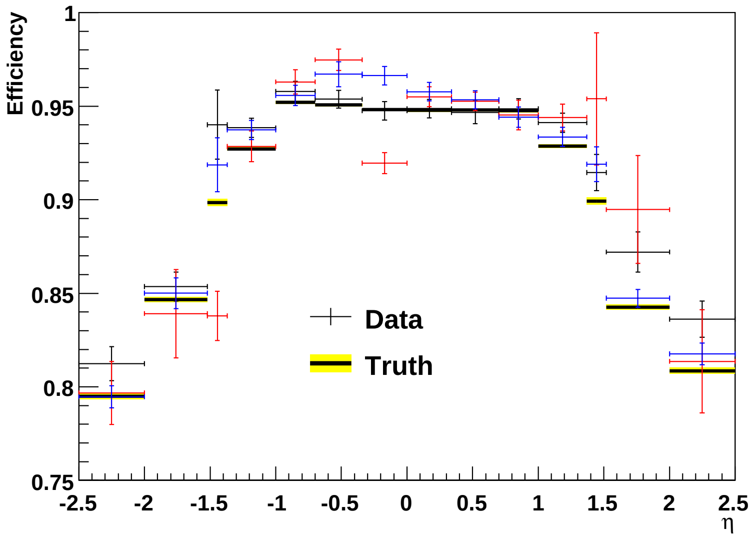


Figure 5.7: Electron efficiency as a function of η for a realistic sample. Black is the sample with nominal background, blue is the half background and red is the sample with background and a half.

5.1.5 Different background normalisations

As we explained in Chapter 4, our knowledge of the QCD background is very limited. The expected cross-section of the dijet process suffers an uncertainty of $\pm 50\%$. As such it is important for us to check that our methods yields consistent results in the different situations. We performed the same analysis as in subsection 5.1.4, but by using weights of 0.5, 1 and 1.5 for QCD background events.

As expected, we can see from Figures 5.7 and 5.8 that the signal to noise ratio has an inverse effect on the uncertainty of our measurement. For a luminosity of 1 fb^{-1}, the implied uncertainty raises from 1% to about 3%. The overall capacity to measurement the efficiency is not fundamentally affected.

5.1.6 Other results on this type of analysis

As described earlier, the same exercise has been carried out for the ATLAS collaboration for [25]. Results can bee seen in the Chapter *Electrons and Photons*,pages 88–91 and in the

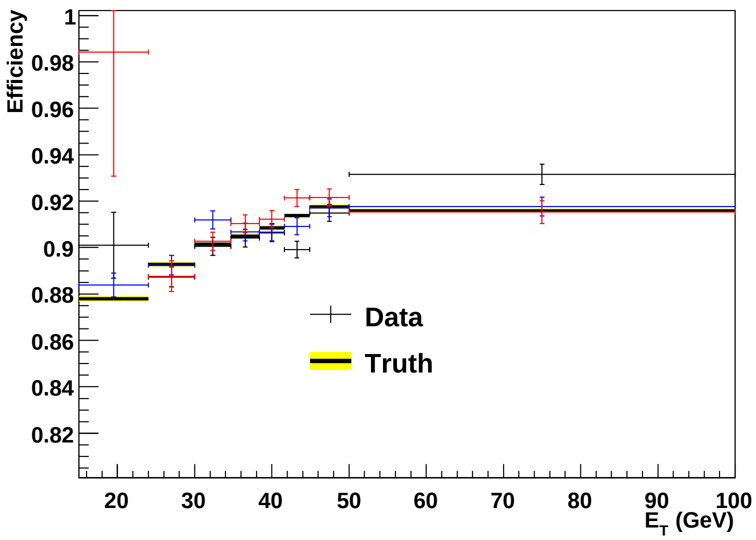


Figure 5.8: Electron efficiency as a function of p_T for a realistic sample. Black is the sample with nominal background, blue is the half background and red is the sample with background and a half.

Chapter *Standard Model*, pages 761–763. Both results focused on a lower luminosity (100 and 50 pb^{-1} resp.).

5.2 Electron reconstruction efficiency measurement for a FDR-style exercise

The FDR (Full Dress Rehearsal) was the last computing exercise organized to test the ATLAS software before the planned start-up of the LHC in September 2008. It was organized as a blind exercise for the different ATLAS analyses.

5.2.1 Samples and software versions considered

Fully simulated events from the CSC datasets were mixed to simulate the expected output of the detector. The simulated hits were inserted in the ATLAS data pipeline to test the trigger and reconstruction frameworks.

Run number	dataset	luminosity (nb^{-1})	events
52280	fdr08_run2.0052280.physics_Egamma.merge.AOD.o3_f47_m26	348.00	90588
52283	fdr08_run2.0052280.physics_Egamma.merge.AOD.o3_f47_m26	349.20	41324

Table 5.10: Runs included in FDR-like analysis

The FDR-2 data was a series of runs played over several days. Each day's series of runs is supposed to correspond approximately to an LHC "fill". The two runs used for this analysis correspond to misaligned data reconstructed with the geometry ATLAS-CSC-02-00-00, both at $10^{31}\ cm^{-2}s^{-1}$ average luminosity. This geometry is the one extracted by the ATLAS alignment group during the CSC exercise and represents a realistic estimate of our knowledge of the *real* detector geometry after a few months of data taking.

Finally, the software release used to reconstruct these events was AtlasTier0-14.2.20.6.

The runs considered for this study are listed in Table 5.10. The difference in the number of events between the two samples comes from the fact that run number 52280 also included a filtered dijets sample to increase the fake electron probability. Altogether, the QCD contribution is largely underestimated, as the simulated statistics available at ATLAS in insufficient to match the luminosity used for other samples.

The trigger menu used was different from the one used in the CSC algorithm. As a result, we used a different trigger chain to select $Z^0 \rightarrow e^+e^-$candidate events. Among the available non-prescale trigger chains, the isolated 20 GeV electron chain is the closest to the lower threshold of 15 GeV we want to study. This chain consists of a Level 1 18 GeV region of interest (L1_EM18) and two high-level trigger objects (L2_e20i and EF_e20i).

5.2.2 Efficiency determination in very early data

The limited luminosity of the sample (697.2 nb^{-1}) combined with the lack of QCD contribution limit our ability to estimate the background contribution in this sample. The background is much lower than what we anticipated in the CSC exercise, with the contribution of back-

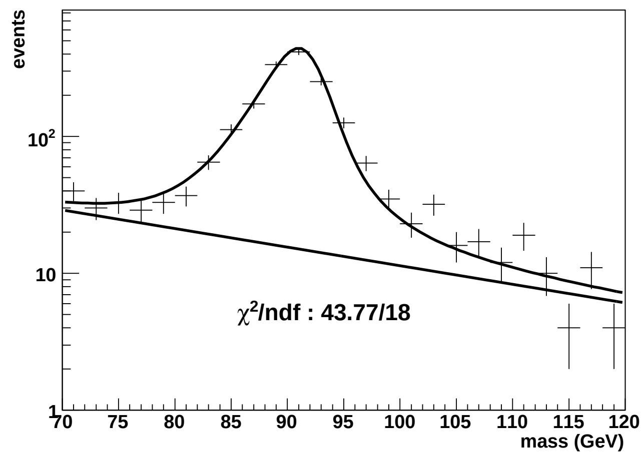
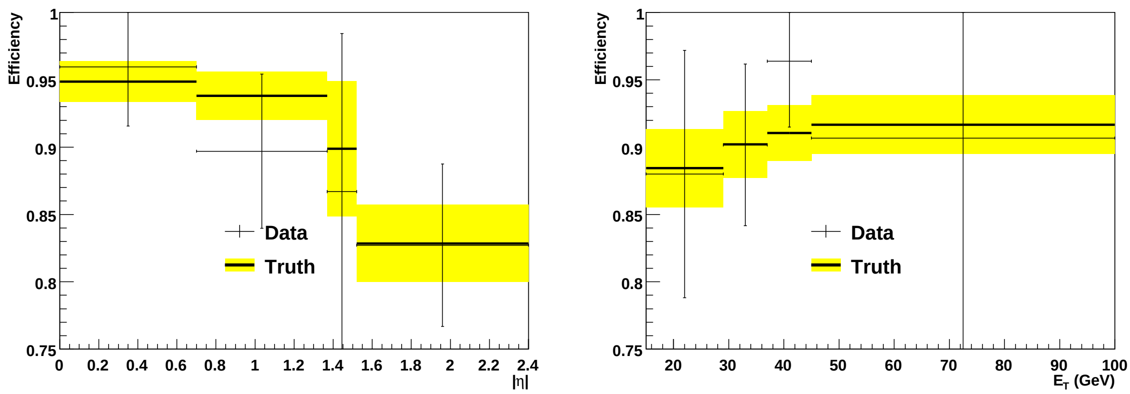


Figure 5.9: Invariant mass spectrum of tag and probe pair in the FDR sample.



(a) Electron efficiency as a function of $|\eta|$ (b) Electron efficiency as a function of E_T

Figure 5.10: Electron efficiencies in FDR sample. Errors include systematics.

ground being limited to 128 ± 52 out of 1577 ± 37 simple electron-cluster pairs (8.11%) and to 78 ± 52 out of 1392 ± 40 matched pairs (5.6%). This background estimation is compatible with the background estimate we got from the CSC estimate in the non-QCD sample (Section 5.1.3). The increase can be explained by the inclusion of some other dilepton samples and the small QCD sample in run 52280.

The Figure 5.10 shows the measured efficiency in the FDR sample. Given that the FDR sample doesn't include any Monte Carlo truth information, the measured efficiency is overlaid with the MC information extracted from the CSC dataset for comparison. The statistical uncertainty on the MC truth corresponds to the same luminosity as the FDR sample.

$\|\eta\|$	Measured mean	stat.	syst.	Monte Carlo mean	difference (σ)
0.00–0.70	95.96	0.84	4.31	94.87	0.25
0.70–1.37	89.70	1.17	5.60	93.81	0.71
1.37–1.52	86.70	3.92	11.06	89.89	0.27
1.52–2.40	82.72	2.13	5.65	82.85	0.02

Table 5.11: Electron Efficiency as a function of $|\eta|$ in FDR sample

E_T (GeV)	Measured mean	stat.	syst.	Monte Carlo mean	difference (σ)
15.00–29.00	88.01	2.44	8.86	88.45	0.05
29.00–37.00	90.17	1.88	5.71	90.21	0.01
37.00–45.00	96.38	0.87	4.81	91.05	1.08
45.00–100.00	90.68	1.24	67.78	91.66	0.01

Table 5.12: Electron Efficiency as a function of E_T (GeV) in FDR sample

Despite the challenging luminosity, these results desmonstrate the fact that we can get insight as to the behaviour of the reconstruction to the level of 5% after just a couple of running hours at this luminosity in absence of QCD background. We should be able to measure the efficiency to within 5% in a handfull of bins.

5.3 ATLAS collaborative effort

These results have been presented to the ATLAS collaboration and have appeared in talks at conferences and proceedings.

The methods and algorithms have been incorporated in a global collaborative effort to measure the detector performance from data. Called InSituPerformance, this effort regroups efficiencies and resolutions measurement for the different types of object ATLAS will try to reconstruct. In particular, the event selection and the background subtraction method have been converted to an Athena algorithm and tool respectively and are now part of the official ATLAS release. The background subtraction tool in particular is available to use by any analyses in PhysicsAnalysis/AnalysisCommon/InsituPerformance/EGammaPerformance.

The InSituPerformance framework is aimed at providing an automated system to measure ATLAS performances with data to be then used by other analyses like cross-section calculations.

The current effort points to the fact that we sould be able to measure the efficiency to within 5% in a few hours and to within 1% to 3%in a few of months depending on the background conditions at the LHC.

Chapter 6

Validation of the ATLAS fast simulation reconstruction – ATLFASTII

This chapter describes the work done with the fast simulation validation group to verify the usability of the fast simulation ATLFASTII to simulate $Z^0 \rightarrow e^+e^-$ events and most backgrounds. It describes the fast simulation software used for the reconstruction and compares the fast and full simulation results.

Due to the complexity of the ATLAS detector, in particular its high granularity, simulating the accurate propagation of particles is a time-consuming process. At ATLAS, we evaluate the average time to simulate the propagation of all the particles of an event from the centre of the detector to the muon system to about fifteen minutes. This makes this solution impractical for large scale productions or tests involving different sets of parameters. Other solutions have been proposed that sacrifice accuracy for speed. The usability of such solutions to study physics processes must be checked, i.e. validated, against the output of the full detector simulation

6.1 The ATLAS fast simulation reconstruction framework – ATLFASTII

The normal ATLAS reconstruction framework, referred to as full simulation, is based on a high precision description of the detector geometry and the material budget in the detector. All known particle-matter interaction effects are modelled using the Geant4 software program [27, 28]. This high precision has a price in speed that makes its usage difficult for comparison between the data and the Monte-Carlo simulation for processes with a large cross-section.

Different options have been proposed to improve the speed of the simulation. They are all based on the same principle.

6.1.1 Principle of fast simulation of particle reconstruction

The most costly part of the full simulation is the computation of the evolution of each individual particle coming out of a hard interaction. The precise tracking of each particle, of its interactions, and of the interactions of each particle produced during any decay require significant computing time. The simulation of particles passing through the calorimeters in particular is very costly because of the high material density in the detector.

Fast simulation strategies are based on the principle of average detector responses. For each particle type, many single particles are fully simulated. We then collect the distribution of the main reconstructed physical parameters that we consider are important for physics studies (transverse energy and momentum, shower shape variables, . . .). When we then want to reconstruct another particle of that type in fast simulation, we take the corresponding distributions for particles with the same basic truth parameters (pseudo-rapidity, momentum), and draw random values from these distributions. This is similar to the principle we presented to simulate a large number of fake electrons with a corrected energy scale from a dijet

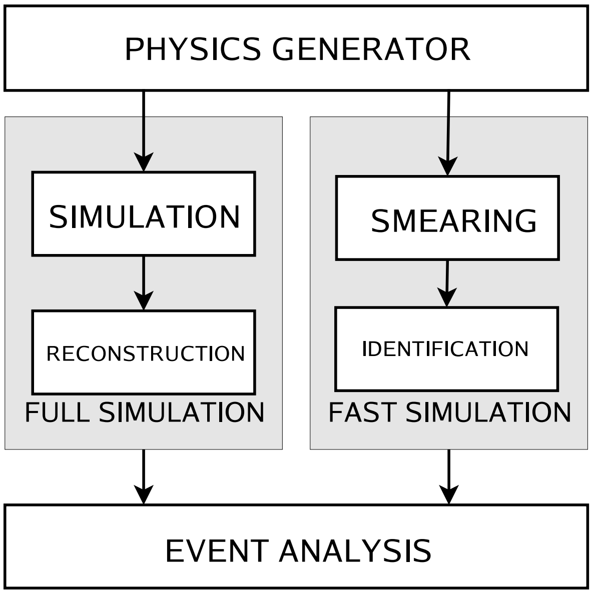


Figure 6.1: Difference between the full and fast reconstruction.

sample in section 4.4.2. This process is called smearing and can produce a reconstructed object straight from the parton-level generation stage.

For boolean or quantified information (identification flags,. . .), we use the full simulation to study the efficiency of a particular flag and use a random number generator to make sure the result follows the right distribution. Again, particles are binned for optimal statistics.

This approach, despite being much faster, presents a lot of limitations. Due to the limited statistics available to create the initial distributions, one has to make concessions as to the precision of the distributions. More importantly, it is very difficult to take into consideration the correlations between the different observables, such as, for example, the interactions between different energy depositions in busy event topologies.

6.1.2 ATLFASTII fast simulation

To improve the overall accuracy of the fast reconstruction, ATLFASTII, the current fast simulation strategy in ATLAS, combines features of the fast and full simulations. It uses a full simulation of the inner detector (PIXEL, SCT and TRT) to get a good representation of the particle tracking and a fast simulation of the calorimeters (FastCaloSim). This limits the time-saving aspect of the fast simulation to a factor of 3 but improves accuracy compared with older versions which used a crude fast(er) tracking algorithm. This work investigates the state of ATLFASTII in Athena 14.2.25.8.

Newer algorithms are being tested as well to use a more modern fast tracking system (ATLFASTII-F) that would save more time. Another extension is to produce detector hits instead of reconstructed objects: this project is called ATLFASTII-D (for digitization). This allows to run trigger code out-of-the-box as well. This is intended to be the default ATLFASTII simulation in Athena release 15 and later.

6.1.3 Validation of individual object reconstruction

The first step of the validation is to check that the smearing process does produce objects that reproduce all the main features of the full simulation. This is generally done using single particle samples to be able to check that each part of the detector geometry is reproduced effectively. This work was carried out by the ATLFASTII validation group and will be presented in a public note added to [25] later in 2009.

6.1.4 Validation through benchmark analyses

The final validation is carried out by comparing the output of full and fast simulations on complete analyses. Some physics processes are chosen (Higgs, W, $t\bar{t}$) to be reconstructed by both full and fast simulation. We then compare the efficiency of the different selection cuts to verify they comply between the two approaches. These analyses are also run on the different backgrounds to check the usability of ATLFASTII for full data analysis.

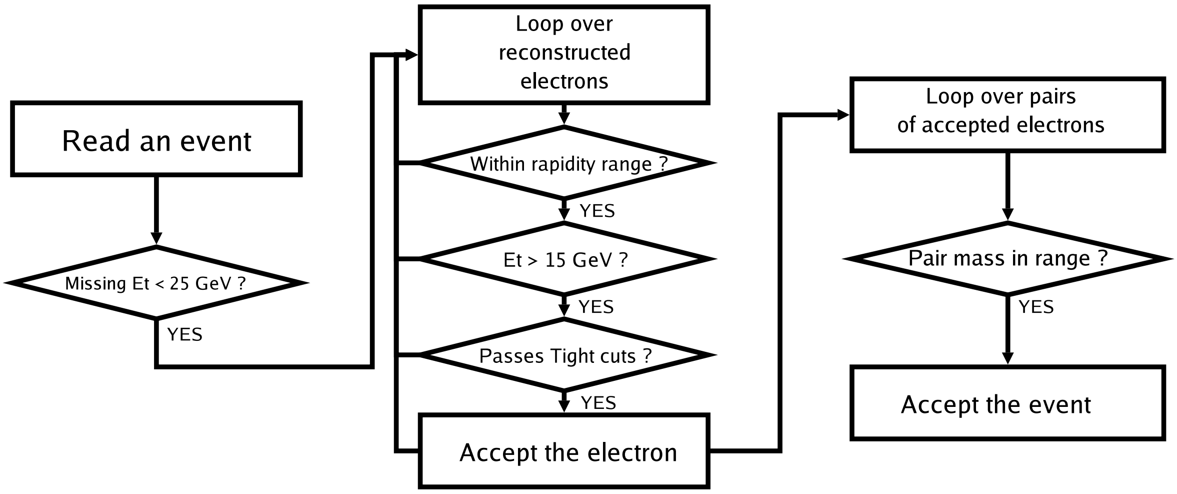


Figure 6.2: Event selection algorithm.

6.1.5 Post-simulation corrections

Since the constants used by the reconstruction are frozen with each release, as the understanding of the detector improves, the output of the simulation for a given release differs from our best understanding at any given time. In order to avoid re-simulating every event, every time a constant changes, a post-simulation correction scheme has been devised. It takes the finished output from the fast simulation and slightly modifies some of the observables (identification flags, momentum, ...) to match our best understanding of the detector performances. These corrections can be applied to be saved to new datasets for later use, or can be performed on the fly to match different running conditions.

The reason we need to apply such corrections is that a fixed version of ATLFAST is released with built-in values of the smearing distributions and identification efficiencies. As new full simulation versions are released or as more data gets collected, the ATLFAST output drifts from our up-to-date knowledge of the detector. The corrections allow us to reduce the difference without rerunning the whole fast simulation framework.

6.2 Benchmark $Z^0 \rightarrow e^+e^-$ analysis validation

The aim of the validation exercise is to design a simple selection for $Z^0 \rightarrow e^+e^-$ events and run it on the signal dataset and its main background to verify whether the fast simulation complies with the full simulation. This would justify the use of fast simulation for feasability studies that do not require detailed detector level cuts.

6.2.1 Signal selection

The selection is very similar to the one described in section 2.3. The selection has to be simplified slightly as not all features of full simulation are available in the fast simulation in version 14.

6.2.1.1 Event selection

In particular, as no trigger information is available, the selection skips that step entirely. We retain the missing transverse selection. The maximum cut-off is increased to 25 GeV. It will still be an important discrimination factor against the backgrounds, but since the object selection has a higher discrimination factor than the one used for the electron efficiency, we can afford to increase our acceptance slightly, without letting any more background in our selection.

6.2.1.2 Objects selection

The object selection consists of requiring two reconstructed electrons with tight identification selection cuts (see ElectronTight in Section 2.1.2.3). Because of the inner detector geometry, we still restrict electrons to an absolute pseudo-rapidity below 2.5. Tight identification was chosen because it is the largest sample of electron with the same identification cut. Medium would include electrons which are medium but not tight. This is a fraction of the signal

($\sim$ 10%) for which we expect the fast simulation to be less reliable because the sample from which to extract the constant is much smaller.

Despite the lack of trigger threshold, we still retained the minimum transverse energy threshold at 15 GeV, as it is what we expect to use when data first become available.

As we are interested in pairs around the Z^0 mass peak, we reject pairs with a transverse mass below 50 GeV and those over 120 GeV.

6.2.1.3 Datasets representation

The validation compares three different setups of the ATLAS reconstruction software. The full simulation dataset is plotted in black throughout this chapter. The standard fast simulation results appear in red while the fast simulation with post-simulation corrections is plotted in blue.

The full list of datasets is given in Appendix B.

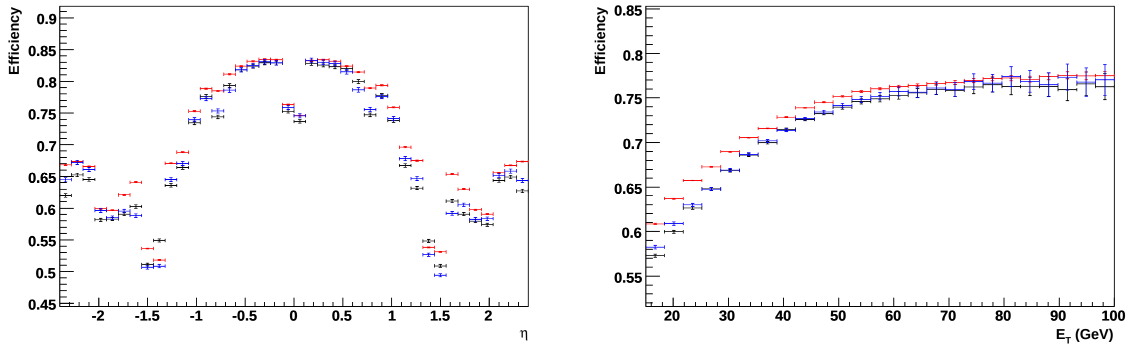
6.2.2 Electron distributions

The first step is checking that the electrons follow the same distributions in full and fast simulations.

6.2.2.1 Efficiencies

We started by checking that the electron efficiency is the same throughout all samples. Any deviation would bias the final number of events passing any selection.

For that we compared the three $Z^0 \rightarrow e^+e^-$ datasets: we measured in each the probability for a generator-level electron to be reconstructed as an electron with tight cuts.



(a) Tight electron efficiency as a function of η (b) Tight electron efficiency as a function of E_T

Figure 6.3: Comparison of tight electron efficiencies. Distribution colours are defined in Section 6.2.1.3.

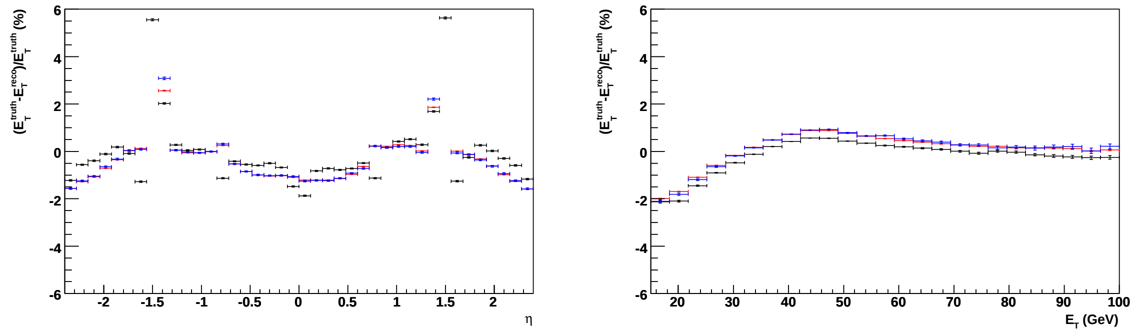
As we can see from Figure 6.3, the efficiency is higher in the fast simulated sample than in the fully reconstructed one. The deviation is clearly visible at low E_T where the difference reaches 4%. In pseudo-rapidity, we see the main difference in the crack region ($1.37 < |\eta| < 1.52$). There also is a large difference at $|\eta| \sim 0.8$ where the full simulation efficiency drops while the fast simulation efficiency remains level with respect to the neighbouring bins.

The application of efficiency correction on this sample shows clearly the important effect they can have. The behaviour of the full simulation is much better reproduced over the whole range of pseudo-rapidity and energy. The corrections are less efficient in the calorimeter transition regions ($|\eta| \sim 0.8$ or $|\eta| \sim 1.5$). This is due to the limited statistics that prevent us modelling very localized effects more efficiently.

6.2.2.2 Energy scale

We then tried to assess whether the energy scale is the same in all samples. In order to achieve this, we looked at the relative difference in transverse energy between the generator-level electrons and their reconstructed counter parts.

Figure 6.4 shows that the energy scale is very consistent between the fast and the full simulations. The match shows an overestimation as a function of E_T of the order of 0.5%.



(a) Tight electron energy scale as a function of η (b) Tight electron energy scale as a function of E_T

Figure 6.4: Comparison of Tight electron energy scales. Distribution colours are defined in Section 6.2.1.3.

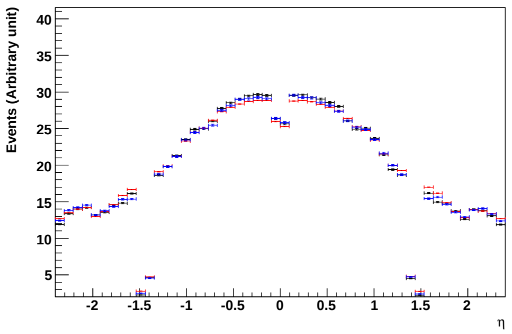
Figure 6.4a show that while the energy is well estimated in the barrel region, the crack region ($1.37 < |\eta| < 1.52$) and the end-caps ($|\eta| > 1.52$) present sizeable deviations. These effects come from the fact that the energy scale in these regions present highly non-gaussian distributions which make the fast simulation less reliable.

Overall the fast simulation smooths out the effect of the detector geometry due to the averaging effect of the smearing. The used version of the post-simulation corrections does not attempt to correct the energy scale, hence the lack of difference between the two samples.

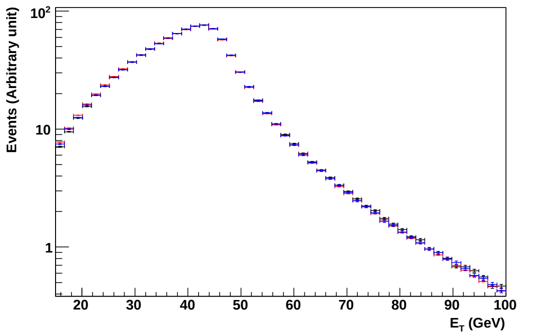
The overall effect of the complicated structure in pseudo-rapidity leads to a small bias over the whole energy range as we can clearly see from Figure 6.4b. This means that on average, the electrons from the fast simulation are slightly less energetic than the ones from full simulation by about 0.5%.

6.2.2.3 Electron parameter distributions

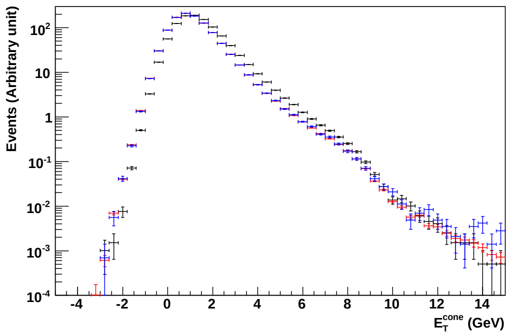
After having observed the small differences in efficiency and energy scale, we can carry out the selections described earlier to keep only tight electrons coming from Z^0 decays. Because each datasets contain a different number of events, we overlay the normalized distributions, to hide the differences in luminosity. Any difference in the efficiency of the event selection



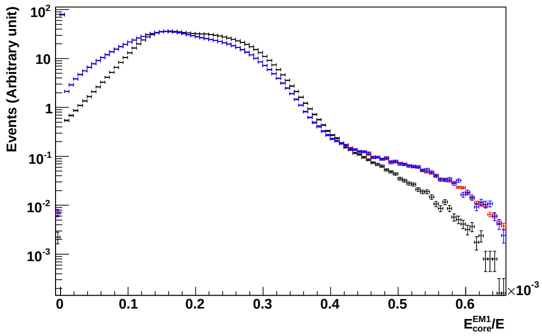
(a) Electron distribution as a function of η



(b) Electron distribution as a function of E_T



(c) Distribution of E_T in a cone with half-opening angle 0.2, with exclusion of a window of size 7×5 in electromagnetic calorimeter



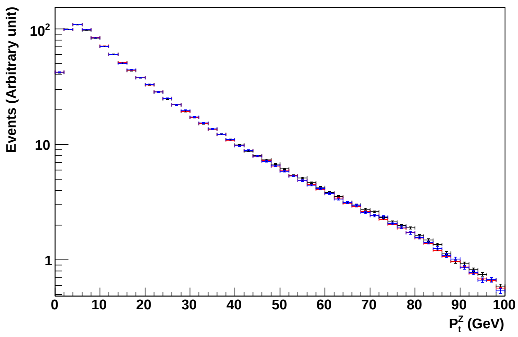
(d) Distribution of the fraction of the energy reconstructed in the first compartment of the electromagnetic calorimeter (see text for details)

Figure 6.5: Comparison of electron distributions. Distribution colours are defined in Section 6.2.1.3.

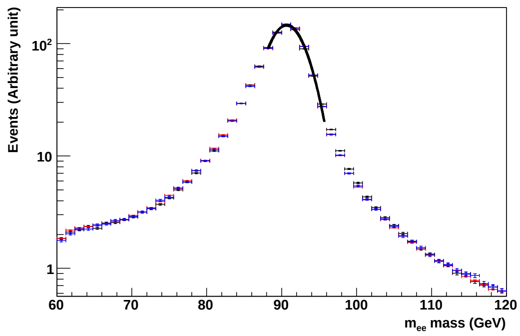
will be detailed later in Section 6.2.4.

Figures 6.5a and 6.5b show that the agreement is between the full and fast simulation is generally good (excluding crack region). The fast simulation only exhibits a small underestimation in the central barrel region.

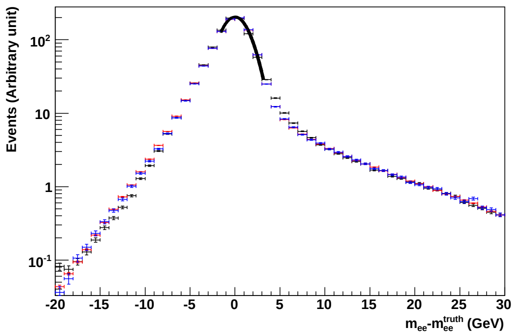
Unfortunately, the fast simulation is less well tuned to simulate precise detector effects like the parameters of the electro-magnetic shower. Figure 6.5c shows that the transverse energy in a hollow cone around an electron is under estimated by the fast simulation. This is a difficult variable to simulate because it is heavily affected by final state radiation, bremsstrahlung and clustering effects. Most shower-shape related variables are not correctly modelled. Figure 6.5d shows another distribution which varies significantly from the full simulation. The distribution is the distribution of the fraction of the energy deposited in the



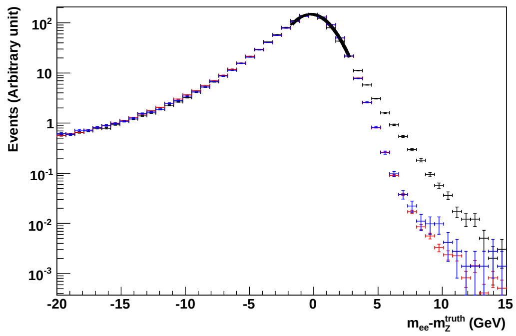
(a) Transverse momentum of the reconstructed Z^0 boson



(b) Invariant mass of the reconstructed electron pairs



(c) Difference between the invariant mass of the reconstructed electron pairs and the invariant mass of the true electron pairs



(d) Difference between the invariant mass of the reconstructed electron pairs and the mass of the true Z^0 boson

Figure 6.6: Distributions of parameters of the reconstructed electron pairs. Distribution colours are defined in Section 6.2.1.3.

electro-magnetic calorimeter in a window of five strips centred around the maximum energy strip in the first longitudinal layer to the total energy of the shower. The reason that the shapes differ so wildly is that the parametrisation was extracted using an outdated version of the full simulation and that the distribution had not been updated since.

These two distributions show that the fast simulation, in the tested version, is ill-suited for analyses that require this level of detail.

6.2.3 Electron pair distributions

It is important that a simulation maintains the correlation between electrons coming from a Z^0 decay. Figure 6.6a shows that the transverse momentum of the reconstructed Z^0 boson is

dataset	Mean (GeV) mean ±error	Width (GeV) mean± error
full sim	90.618±.005	2.567± .005
fast sim	90.696±.002	2.510± .002
fast sim with corrections	90.700±.005	2.510± .006

Table 6.1: Parameters of the gaussian fit of the Z^0 invariant mass peak in Figure 6.6b

dataset	Mean (GeV) mean ±error	Width (GeV) mean± error
full sim	-0.066±.003	1.593± .003
fast sim	0.046 ±.001	1.600± .001
fast sim with corrections	0.048 ±.003	1.600± .004

Table 6.2: Parameters of the gaussian fit of the distribution of the difference between the invariant mass of the electron pair and the invariant mass of the true electron pair in Figure 6.6c

conserved between the samples. As we can see, the fast simulation reproduces the spectrum of the full simulation.

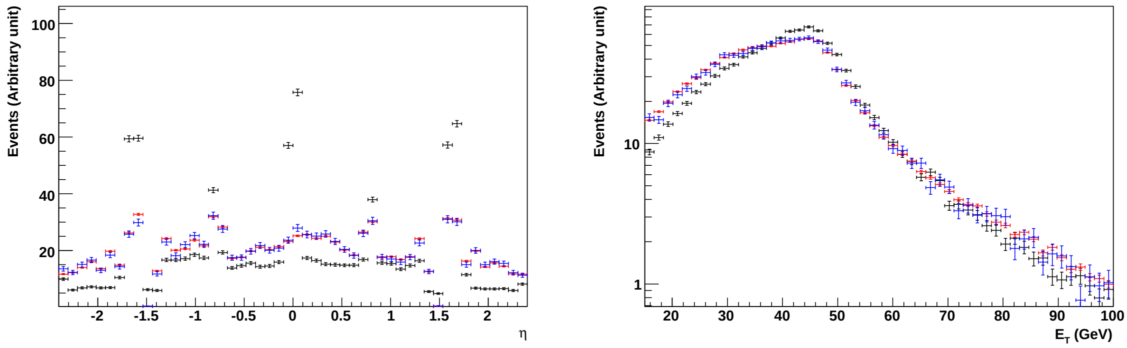
Similarly, the invariant mass distributions plotted on Figure 6.6b show a very good agreement. To check the agreement of the peaks, as well as the check of the tails, a gaussian fit is shown. A gaussian is fitted iteratively in a range $[-\sigma, 2\cdot\sigma]$ around the distribution mean until it stabilizes. The parameters of the fits are displayed in Table 6.1. The fast simulation agrees again well with the full simulation.

Figure 6.6c shows the difference between the invariant mass of the reconstructed pair and the one of the true electron pair, i.e. the true electrons after final state radiations might have happened. Again the distributions match and the gaussian fit parameters displayed in Table 6.2 are in agreement. The small difference in the mean can be explained by the small difference in energy scale we already noticed, but overall the effect is negligible when compared with the width of the peak.

Figure 6.6d shows the difference between the invariant mass of the reconstructed pair and the mass of the true Z^0 boson in the event. Despite the fact that the fit parameters show that the peak is largely unaffected by the fast simulation, we clearly see a lack of fast simulation

dataset	Mean (GeV) mean ±error	Width (GeV) mean± error
full sim	-0.303±.004	1.527± .003
fast sim	-0.173±.001	1.506± .001
fast sim with corrections	-0.172±.004	1.507± .003

Table 6.3: Parameters of the gaussian fit of the distribution of the difference between the invariant mass of the electron pair and the invariant mass of the true Z^0 boson in Figure 6.6d



(a) Electron distribution as a function of η

(b) Electron distribution as a function of E_T

Figure 6.7: Comparison of tight electron distribution for events where $m^{reco}_{e^+e^-} - m^{truth}_Z > 3$ GeV. Distribution colours are defined in Section 6.2.1.3.

events in the high tail of the distribution.

Figure 6.7 shows the distribution of the electron parameters for precisely the electrons in events in the high tail ($m^{reco}_{e^+e^-} - m^{truth}_Z > 3$ GeV). We can clearly identify the detector regions which lead to the lack of events. They correspond to transition detector regions where the calorimeter segments are connected (see Figure 1.13). The full simulation predicts that the energy of the pair is much more likely to be overestimated if one of the electrons fall in the regions $|\eta| \sim 0$, $|\eta| \sim 0.8$ or $|\eta| \sim 1.5$. This results from a large non-linearity of the reconstructed energy in these regions. While the fast simulation does present some non-linear behaviour in these regions, it does not correctly reproduce the effects happening in these regions.

The difference between the distributions in Figure 6.6d is hard to see in Figure 6.6b. The 1% or below effect is spread out over many more bins which make a difference in shape invisible to the fit or the naked eye.

The E_T distribution in Figure 6.7b also shows that the fast simulation overestimates the contribution of low E_T electrons compared to the full simulation. Coupled with the small difference in energy scale between the simulations (Figure 6.4b), this leads to the lack of events where the reconstructed electron pair mass is higher than the true mass.

The post-simulation corrections have negligible influence on the distribution shapes. However the difference in efficiencies does have a noticeable impact on the total number of pairs collected as we will see in the Section 6.2.4.

6.2.4 Selection cut flow tables

One of the most important aspects of fast simulation is that it reproduces as precisely as possible the selection applied on the full analysis. One way to check this is to compare the cut flow: we observe which fraction of the signal survives each of our cuts.

The different cuts are as follows:

Missing E_T The event passes the selection for a maximum missing transverse energy of 25 GeV;

Acceptance The number of pairs of electrons in these events where both electron pass the pseudo-rapidity and minimum momentum cuts. There may be more than one pair per event in the case of fakes;

Identification The number of such pairs where both electrons pass the identification criteria (ElectronTight);

Charge The number of such pairs where both electrons have opposite charge;

Mass The number of such pairs with an invariant mass in the window $[60\text{ GeV}, 120\text{ GeV}]$.

cut	G4		AFII		Corrected AFII	
	events	cut eff.	events	cut eff.	events	cut eff.
Full sample	983806±447		983806±143		983806±527	
Missing E_T	965093±443	98%	954329±141	97%	954324±519	97%
Acceptance	425240±294	44%	425154±94	45%	424904±346	45%
Identification	223638±213	53%	231457±69	54%	222231±250	52%
Charge	222551±213	100%	230305±69	100%	221087±250	99%
Mass	218478±211	98%	226031±68	98%	216963±247	98%

Table 6.4: Cut flow table for the $Z^0 \rightarrow e^+e^-$signal. Number of events is scaled to 1 fb^{-1} and statistical errors are MC statistics scaled to the same luminosity.

We then compare the effect of each cut on the signal and three backgrounds. The list of datasets used is listed in Appendix B. We do not compare the output of the dijet sample because its large cross section will require a special treatment for each analysis which will be too analysis specific. This section will present the cut flow tables for the following samples:

- $Z^0 \rightarrow e^+e^-$;
- $Z^0 \rightarrow \tau^+\tau^-$;
- $t\bar{t}$;
- $W + jets$.

The efficiency of the different cuts for the signal $Z^0 \rightarrow e^+e^-$ sample, as seen in Table 6.4, gives us many indications of the strength and weaknesses of ATLFASTII.

The number of events left after the Missing E_T cut is understimated by the fast simulation by approximately 1%. Both samples present the same distribution in pseudo-rapidity and transverse momentum as we've already seen on Figure 6.5, hence the efficiency of the Acceptance cut is the same through all samples. But we have also noted that the probability to pass the tight identification cuts is slightly higher for the fast simulation which leads to a $\sim$ 2% increase in the selection efficiency for the Identification cut.

The identical values of efficiencies of the Charge cut indicates that both simulations have the same small charge misidentification rate.

cut	G4		AFII		Corrected AFII	
	events	cut eff.	events	cut eff.	events	cut eff.
Full sample	1128370±1458		1128370±2523		1128370±3402	
Missing E_T	903350±1305	80%	905839±2261	80%	905804±3048	80%
Acceptance	31370±243	3%	33699±436	4%	33646±587	4%
Identification	1002±43	3%	1207±83	4%	1159±109	3%
Charge	982±43	98%	1151±81	95%	1118±107	96%
Mass	266±22	27%	310±42	27%	339±59	30%

Table 6.5: Cut flow table for the $Z^0 \rightarrow \tau^+\tau^-$ background. Number of events is scaled to 1 fb^{-1} and statistical errors are MC statistics scaled to the same luminosity.

Finally, the fact that the energy scales were mostly identical between the samples leads to the fact that the reconstructed Z^0 boson masses follow the same distribution in both simulations, hence the identical efficiency for both Mass cuts.

Overall, normalized to the integrated luminosity of the full simulation sample, this leads to a 3.46% overestimation in the number of events for the fast simulation.

Once the post-simulation corrections have been applied, the discrepancy of the identification cut is reduced. This directly leads to a difference in the normalized number of events of barely 0.9%, as none of the other cuts are overly affected.

The different backgrounds present different behaviours. $Z^0 \rightarrow \tau^+\tau^-$, as we can see from Table 6.5, presents good consistency in the efficiency of the Missing E_T and Acceptance cuts.

We also find that the increased efficiency in the corrected fast simulation leads to a small increase of the efficiency of the Identification cut again. This is particularly important because the difference in efficiency is highest in the low E_T region where most of the electrons from taus are. The Charge and Mass cut efficiencies are statistically compatible between full and fast simulations.

Overall, the normalized fast simulation sample presents 16.5% more events than the full simulation but this number has to be considered relative to the 8.3% statistical uncertainty on the full simulation number.

cut	G4		AFII		Corrected AFII	
	events	cut eff.	events	cut eff.	events	cut eff.
Full sample	205475±147		205475±92		205475±118	
Missing E_T	25401±52	12%	26689±33	13%	26675±43	13%
Acceptance	33527±59	132%	40041±41	150%	40068±52	150%
Identification	208±5	1%	228±3	1%	223±4	1%
Charge	194±5	93%	211±3	92%	205±4	92%
Mass	84±3	43%	93±2	44%	89±2	43%

Table 6.6: Cut flow table for the $t\bar{t}$ background. Number of events is scaled to 1 fb^{-1} and statistical errors are MC statistics scaled to the same luminosity.

For this sample, the post-simulation corrections increases the difference. The better efficiency cuts out fractionally more background even though the efficiency is still overestimated in the low E_T region. However, the efficiency corrections, varying with energy, increases the invariant mass of the remaining pairs, leading to more of them passing the Mass cut.

The $t\bar{t}$ background presents the same properties as the $Z^0 \rightarrow \tau^+\tau^-$ one as we can see from Table 6.6. The Missing E_T cut is consistent throughout all samples.

The Acceptance cut presents an efficiency of more than 100% because the extra jets increase the number of fake electrons, and therefore, the number of pairs we can reconstruct. If we have three electrons (including one fake), we can indeed reconstruct three possible pairs. The identification cuts are nevertheless identical as the rejection factor of the tight identification cuts rejects nearly all fake electrons.

Again the Charge and Mass cuts do not present any disparity between the full and the fast simulation.

Overall, the fast simulation over estimated this background by 10.7% but also the full simulation presents a statistical uncertainty of 4%.

The effect for this sample of post-simulation corrections is opposed to the one for $Z^0 \rightarrow \tau^+\tau^-$. The efficiency for high energy electrons implies an invariant mass reconstruction closer to the one from the full simulation. It means more pairs are failing the Mass cut because their invariant mass is too high. The remaining small difference in the normalized

cut	G4		AFII		Corrected AFII	
	events	cut eff.	events	cut eff.	events	cut eff.
Full sample	10352848±22191		10352848±59772		10352848±59772	
Missing E_T	2516223±10940	24%	2609263±30007	25%	2609263±30007	25%
Acceptance	56459±1639	2%	72815±5013	3%	72815±5013	3%
Identification	143±82	0%	345±345	0%	0±345	0%
Charge	95±67	67%	0±345	0%	0±345	0%
Mass	48±48	50%	0±345	0%	0±345	0%

Table 6.7: Cut flow table for the $W + jets$ background. Number of events is scaled to 1 fb^{-1} and statistical errors are MC statistics scaled to the same luminosity.

sample is mostly due to the increased number of fakes which leads to a higher number of pairs passing the Acceptance cut, which biases the rest of our selection.

The $W + jets$ background sample (Table 6.7) shows that both simulations predict virtually no background from this sample. Virtually every event fails our Identification cut as there is only one true electron in each event. The remaining survivors are completely dealt with by the opposite charge requirement.

There are no obvious differences introduced by the post-simulation corrections.

Overall, even in such a simplified analysis with selection cuts tuned to be comparable with fast simulation, the agreement is limited but sufficient for feasability studies.

6.3 Summary

The fast improvement of the fast simulation has made it very usefull for feasibility study that do not require precise information of energy deposition in the detector. The current reconstruction allows a selection of event to within 5% at a cost a third of the one of the full simulation. Improvements being developped will futher reduce the cost by a factor of 3 to 5 which will make very large scale simulation much more manageable.

Chapter 7

Sensitivity to the strong coupling constant α_s at the Large Hadron Collider

This chapter studies the possible influence of the strong coupling constant on different distributions. It aims at analysing the impact of different values of the parameter $\alpha_s(M_{Z^0})$ on distributions measured from the production of Z^0 bosons with associated jets in the $Z^0 \to e^+e^-$ channel for an integrated luminosity of 1 fb^{-1}.

Quantum Chromodynamics (QCD) will play a major role in LHC analyses. It is a major source of uncertainty on the rate of most hadronic events (hard multi-jets interactions) and on the underlying event structure. We hope that the use of LHC data will enable us to further constrain its parameters, thus improving our theoretical predictions and our understanding of hadronization. Many QCD effects need to be understood to improve our models, from parton distributions functions in the proton structure to parton showering in hard interactions.

Most of these effects are governed by a limited number of parameters. One of these parameters appears in all QCD equations: the strong coupling constant α_s. This chapter tries to study if the production of $Z^0 \to e^+e^-$ events associated with jets at the LHC has the sensitivity necessary to be competitive with the existing measurements.

7.1 The strong coupling constant

7.1.1 Theoretical framework

The effective QCD coupling α_s has a strong dependence on the renormalization scale. This dependence is the reason that ensure that quarks are bound at small scales whilst being asymptotically free at high energy. The Particle Data Group [3] describes this dependence as being controlled by the β-function:

$$\begin{aligned}
\mu\frac{\partial\alpha_s}{\partial\mu} &= 2\beta(\alpha_s) = -\frac{\beta_0}{2\pi}{\alpha_s}^2 - \frac{\beta_1}{4\pi^2}{\alpha_s}^3 - \frac{\beta_2}{64\pi^3}{\alpha_s}^4 - \ldots, \\
\beta_0 &= 11 - \frac{2}{3}n_f, \\
\beta_1 &= 51 - \frac{19}{3}n_f, \\
\beta_2 &= 2857 - \frac{5033}{9}n_f + \frac{325}{27}n_f^2;
\end{aligned}$$

where n_f is the number of quarks with mass less than the energy scale μ. Although this expression governs the running of the coupling, it does not determine its normalisation. That must be determined from experiment. This constant of integration is the one fundamental constant of QCD. For convenience and to allow for the comparison of the value with other experiments, it has become standard to choose M_{Z^0} as a fixed-reference scale. The value at other values of the scale μ can be obtained from $\log(\mu^2/M_{Z^0}^2) = \int_{\alpha_s(M_{Z^0})}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)}$.

Measurements of $\alpha_s(M_{Z^0})$ have been performed in deep-inelastic scattering experiments through global fits [29] as well as in high-energy hadron collisions [12] and at perturbative QCD level in e^+e^- collision at LEP [30]. The values of $\alpha_s(M_{Z^0})$ going into world average computations are shown in Figure 7.1. The current world average is estimated at $\alpha_s(M_{Z^0}) = 0.1176 \pm 0.002$.

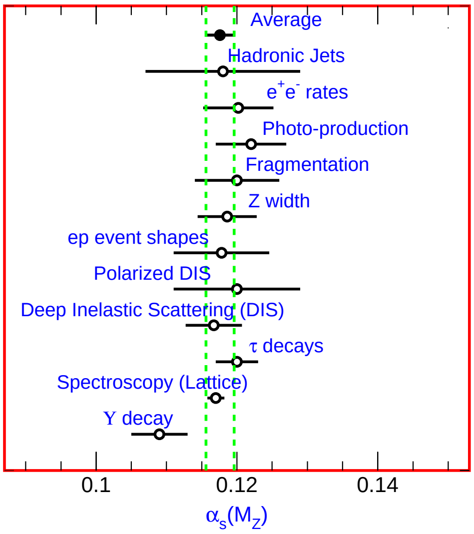


Figure 7.1: Summary of the value of $\alpha_s(M_{Z^0})$ from various processes. The values shown indicate the process and the measured value of α_s extrapolated to $\alpha_s(M_{Z^0})$. The error shown is the total error including theoretical uncertainties. The average which comes from these measurements is also shown. [3]

7.1.2 Motivations

As we already stated, the effective QCD coupling α_s plays a major role at the LHC. While it is correlated with other parameters in parton distribution functions and parton shower implementations, it is easy to understand its contribution to hard-interaction rates.

The simplest Feynman diagram for Z^0 production at the LHC is given by Figure 7.2. It contains no strong interaction vertex. However, at tree level, Z^0 production associated with one hard jet displays one qqg or $gq\bar{q}$ initiated hard initial state radiation as seen on Figure 7.3.

Each strong interaction vertex carries an α_s-proportional factor in rate, such that we expect the ratio $R_0 = \frac{Z^0+1jet}{Z^0+0jets}$ to be a function of the value of $\alpha_s(M_{Z^0})$. The QCD vertex factor is given from Feynman rules by $-ig\gamma^\mu T^a_{ij}$, where a and μ are the indices carried by the gluon propagator, i and j the ones from the quarks, and g is the coupling such that $\alpha_s = g^2/4\pi$ [31]. In the squared matrix element, we always obtain even powers of g, hence the α_s dependence of the cross-sections of the considered processes.

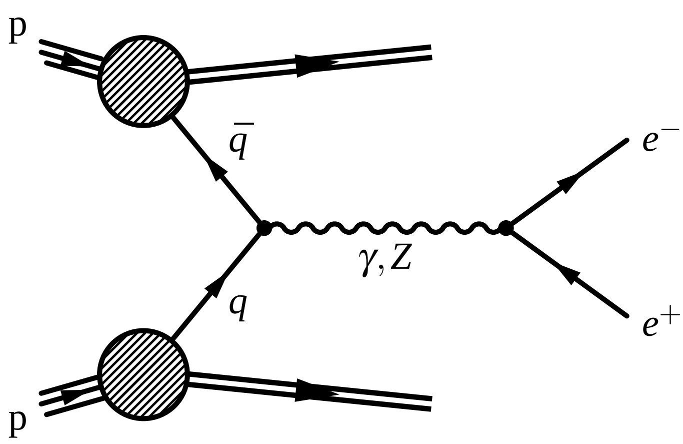


Figure 7.2: Feynman diagram of the simplest $Z^0 \to e^+e^-$ event.

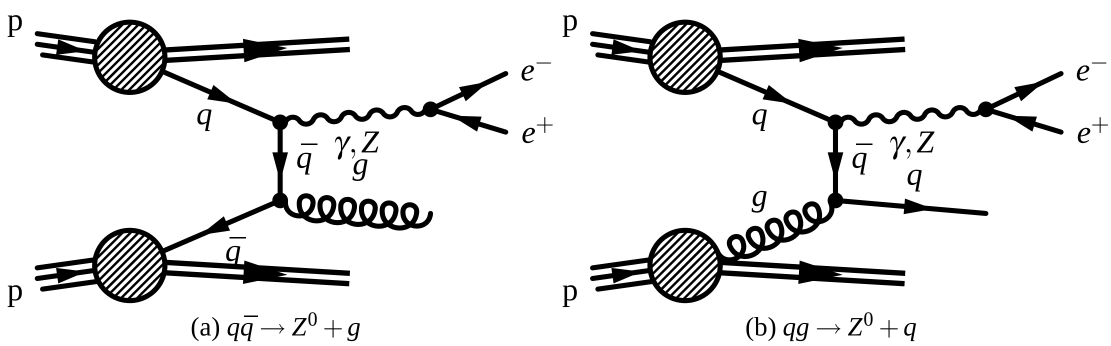


Figure 7.3: Feynman diagrams of two tree-level Z^0 +1 jet events.

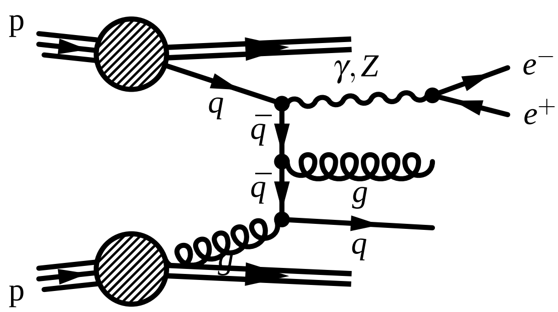


Figure 7.4: Feynman diagram of a tree-level Z^0 +2 jets event.

Each subsequent quark or gluon radiation vertex carries the same factor, so that higher order tree level diagrams can also be used to extract the strong coupling constant such as the one drawn on Figure 7.4.

For this reason, measuring not only the ratio $R_0 = \frac{Z^0+1jet}{Z^0+0jets}$, but the ratio between any two distributions for different numbers of associated jets may contribute to the study of the influence of α_s.

The use of $Z^0 \rightarrow e^+e^-$ events is justified by the high-cross section and the clean signatures we already described in Section 2.3.1. Background subtraction is also easier than in most other channels, thanks to the procedure we described in Chapter 4.

7.1.3 Data preparation

As we already explained, the constant $\alpha_s(M_{Z^0})$, or simply α_s, is fundamental in all physics processes involving QCD. Hence, to simulate events with a particular value of α_s, that value must be consistently used by every part of the simulation chain.

By default in release 14, the ATLAS simulation framework uses a value of $\alpha_s(M_{Z^0}) = 0.118$. The default parton distribution used is CTEQ6L1 (from LHAPDF-5.6.0 [32]) which is a Leading Order (LO) fit using a LO expression for the running of α_s. Pythia, and any other generator is configured to use the same value for hard interactions and parton shower algorithms.

Because we wanted to simulate events with different values of α_s, we had to simulate events using a different software configuration.

7.1.3.1 Choice of Parton Distribution Functions

Parton Distribution Function (PDF) fits are generally performed for fixed values of α_s. As we already stated, CTEQ6L1 uses the value 0.118 [33]. However, the CTEQ collaboration also releases PDFs using different values of α_s. The PDF set CTEQ6AB is a set of 20 PDFs using 10 different values of α_s that span the range $0.110 < \alpha_s(M_{Z^0}) < 0.128$ with 0.002 increments [34]. These PDFs are Next-to-Leading Order (NLO) in α_s but use two slightly different methods for the running of the constant.

Set A, CTEQ6A the original NLO definition is given by

$$\alpha_s(\mu) = c_1[1 - c_2 \ln(L)/L]/L,$$

where $L = \ln(\mu^2/\Lambda^2)$, $c_1 = 4\pi/\beta_0$, and $c2 = 16\pi^4/\beta_1$. The parameter Λ depends on n_f, and hence takes on different values Λ_{n_f} when μ crosses each quark threshold. This is the definition used in all others CTEQ global analyses.

Set B, CTEQ6B this definition solves exactly the partial differential equation

$$\mu \frac{\partial \alpha_s}{\partial \mu} = -\frac{\beta_0}{2\pi} {\alpha_s}^2 - \frac{\beta_1}{4\pi^2} {\alpha_s}^3.$$

The other theoretical assumptions and functional parametrizations are identical to other CTEQ6 series, so that PDFs CTEQ6A118 and CTEQ6B118 are nearly identical to PDF CTEQ6M (LO fit, NLO α_s), the best fit PDF from the CTEQ collaboration.

The range of $\alpha_s(M_{Z^0})$ covered by the CTEQ6AB series is much wider than the currently accepted 1σ error range on the world average. It explores this extended range because the lowest and highest values (0.110 and 0.128) represent outlying values that have been obtained by individual experiments.

We decided to use the CTEQ6A series because it allows for an easier comparison with other CTEQ PDF sets, in particular the PDF error set of the CTEQ6M series (see 7.4.5).

7.1.3.2 Choice of event generator

The ATLAS software can be interfaced with a number of different event generators, from leading-order generators (PYTHIA [35], ALPGEN [36]) to Next-to-Leading Order ones (MC@NLO [37]). The different generators have different limitations that make the simulation of our events difficult:

- PYTHIA uses the Leading Log approximation, which is known to underestimate the hard-jet emission process. This is a major issue for Z^0+jets events generation;

- ALPGEN, MC@NLO and others have a limited set of PDFs hard-coded, which mean we cannot use the CTEQ6A series. What is more, they do not offer any mechanism to set the value of α_s consistently.

For these reasons, we elected to use the generator SHERPA 1.1.2 [38]. SHERPA uses a tree-level matrix element generator for the calculation of hard-scattering processes. It allows us to choose any PDF set in the LHAPDF package and the value of α_s can be set consistently for both hard-scattering and parton shower processing.

The configuration files used to simulate events for Z+up-to-3 jets with $\alpha_s = 0.110$ at 14 TeV can be found in Appendix C. The events were produced in 2008 and in winter 2009, when it was expected that the 10 TeV phase of collisions would last but a few months to tune the machine. Hence it was believed that we would only collect a few hundred inverse picobarns at 10 TeV before moving on to the design energy. Therefore the centre-of-mass energy was set to 14 TeV. It has been announced since that the LHC would start at 7 TeV centre-of-mass energy before increasing to 10 TeV for an extended period of time, collecting more than the target 1 fb^{-1}.

Unfortunately, this means that all the background samples in 2009 have been centrally generated at 10 TeV. Since we do not attempt a precise measurement of the level of background in our sample, but only aim to estimate the uncertainty which the background sub-

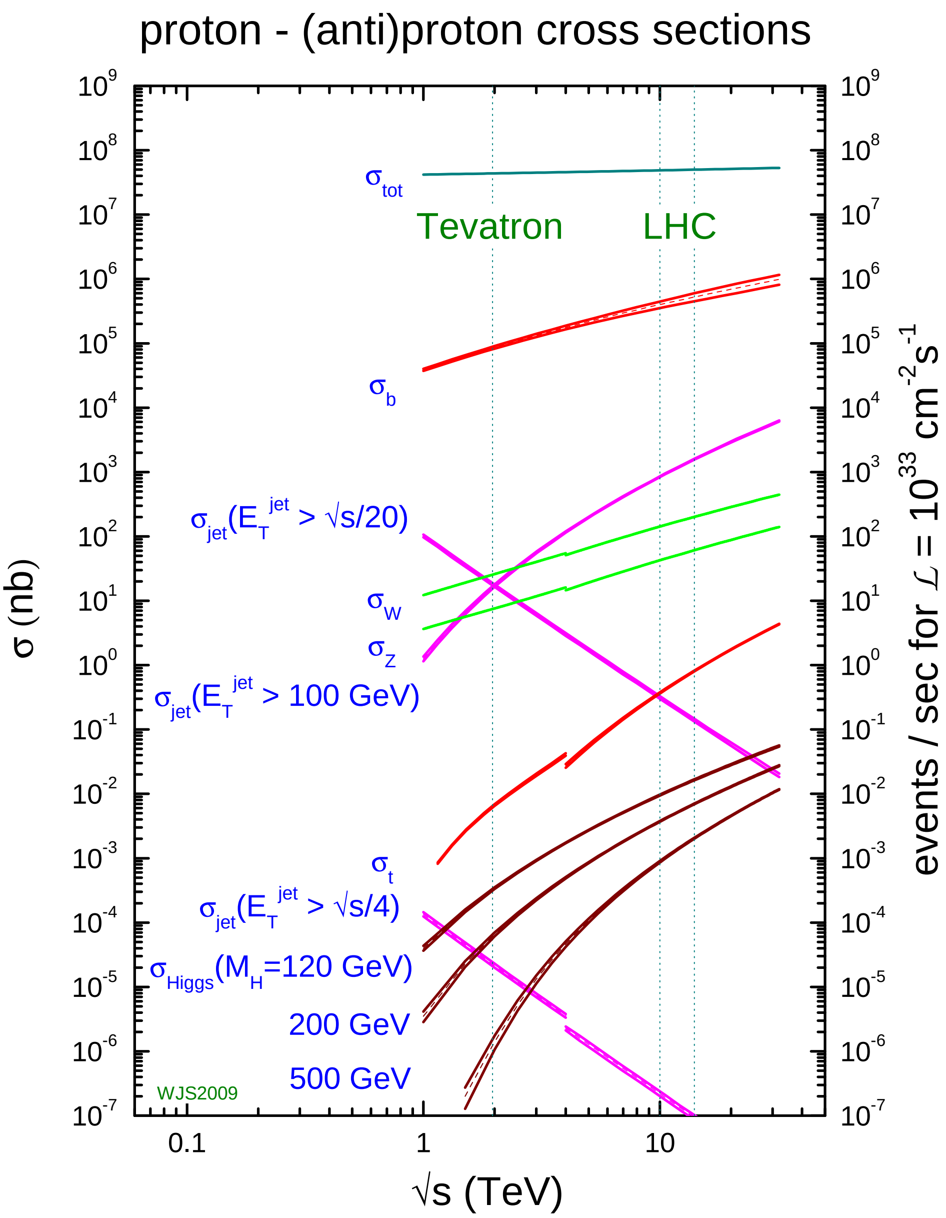


Figure 7.5: Cross-section of various SM processes as a function of the centre-of-mass energy [39]

$\alpha_s(M_{Z^0})$	number of events	cross-section (pb)	luminosity (fb^{-1})
0.110	773041	684.79	1.12
0.118	786926	799.39	.98
0.128	790921	911.96	.86

Table 7.1: Cross-sections and sizes of the privately produced samples of $Z^0 \to e^+e^- + jets$.

traction method introduces on our measurement, this is treated as one further source of uncertainty on the level of background. Figure 7.5 shows how the cross-sections of various processes are affected by this change in available centre-of-mass energy.

We decided to only simulate events with up to 3 jets for the following reasons :

- the computation time needed to prepare the matrix elements for all diagrams for Z+4 jets events was too high;
- the cross section for Z+3 jets is, according to SHERPA for $\alpha_s = 0.118$, 32.44 pb out of the inclusive cross section for Z+up-to-3 jets of 799.39 pb (4%). The Z+4 jets cross section would be of the order of 10 pb or less, which would leave poor statistics for the background subtraction;
- we do not expect our background representation to be very accurate for a larger number of jets before Monte-Carlo tuning using real ATLAS data.

7.1.3.3 Choice of detector simulation

Given the high time consumption of the full simulation, and the private nature of the event production, we selected to simulate the events using ATLFASTII as described in Chapter 6. The configuration was the same as the one used for the official validation samples in Athena 14.2.25.8.

7.1.3.4 Simulated data

We simulated events with three different values of α_s(0.110, 0.118, 0.128). Table 7.1 shows the number of events and cross sections of the three simulated datasets. The cross sections are the ones given by SHERPA and include the filter efficiencies. The filter applied on the generation is described in the configuration file Selector.dat (see Appendix C). The filters are :

- $|\eta^{e^+/e^-}| < 3$,
- $p_T^{e^+/e^-} > 8$ GeV,
- $m_{Z^0} > 50$ GeV,
- $E_T^{jet} > 20$ GeV, for each jet.

As we already mentioned, the values of α_s span a range which is wider than the error on the world average. The reason behind this choice was to study the full range of values obtained by different experiments. We aim at comparing the result of the simulations with uncertainties from PDFs and Jet Energy Scale (JES). Precision measurements of α_s will require higher integrated luminosities as well as Monte-Carlo tuning to ATLAS data.

7.2 Event Selection

The selection of $Z^0 \rightarrow e^+e^-$ events is carried out following similar selection cuts to those outlined in Section 2.3

7.2.1 Global event selection

As we already pointed out in Chapter 6, no trigger information is available in ATLFASTII reconstruction in release 14. The global event selection is therefore only based on a maximum missing transverse energy cut $E_T^{miss} < 25GeV$.

7.2.2 Electron selection

To maximize the background rejection, we only retain as valid electrons, the electrons passing the ElectronTight set of identification cuts in the pseudorapidity range $|\eta| < 2.5$ and with a minimum transverse energy of 15 GeV, which would correspond to our trigger requirement at low luminosities (10^{32} cm^{-2}s^{-1}). Only events which contain exactly two such electrons are kept in the analysis.

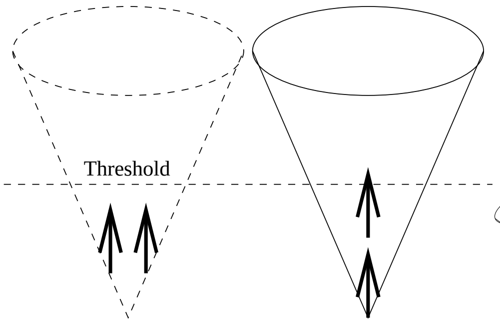
7.2.3 Jet selection

Jet counting is an intrinsic part of a Z + jets study. Hence the selection of jets is an important step and some care has to be taken.

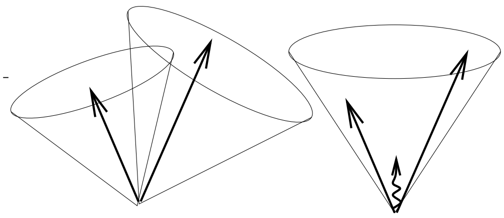
The first consideration is geometric. Because of the coverage of the hadronic calorimeter in ATLAS (see Section 1.2.2.2), we restrict ourselves away from the forward calorimeters which are expected to be harder to calibrate effectively. Hence we only study jets in the pseudorapidity range $|\eta| < 3$.

7.2.3.1 Jet algorithms

ATLAS, like most recent high-energy physics experiment, employs different schemes for jet reconstruction. The most common ones have been in use for years at Tevatron [40] are called Cone and Kt algorithms.



(a) Collinear safety issue



(b) Infrared safety issue

Figure 7.6: Illustration of collinear (left) and infrared safety (right) issues. See text for details

- Cone algorithms form jets by associating together particles whose energy deposition lie within a circle of a specific radius R in the $\eta \times \phi$ plane. ATLAS uses two different values for the parameter R of 0.4 and 0.7. The algorithm seeds from high-energy calorimeter energy depositions and iterates until stable energy weighted cones are obtained.

- Kt algorithms [41] are recursively aggregative algorithms. At each step, they calculate the distance between two pre-clusters i and j using the formula $d_{ij} = \min(p^2_{T,i}, p^2_{T,j})\frac{\Delta R^2_{ij}}{D^2}$ where D is a parameter of the jet algorithm. If this distance is smaller than both $p^2_{T,i}$ and $p^2_{T,j}$ the two clusters are merged. The algorithm produces a list of jets separated by at least $\Delta R > D$. ATLAS uses again two different values for D of 0.4 and 0.6.

In addition, ATLAS also implemented a new algorithm: Seedless Infrared Safe Cone jet algorithm (SISCone), which is a variation of cone algorithms supposed to provide more infrared and collinear safety [42]. Collinear safety implies the algorithm should not be sensitive to threshold effects if, instead of one high-energy particle, the jet is composed of multiple collinear softer ones, as seen on Figure 7.6a. Similarly, infrared safety implies the identified jets should not depend on the presence of soft radiation as seen on Figure 7.6b. Typically, seeded algorithms are also collinear safe if the seed energy threshold is much lower than the scale at which jets are studied.

7.2.3.2 Jet calibrations

For reconstructed data, full simulation and ATLFASTII provide two calibration schemes for jet algorithms. The first is based on calorimeter towers and uses a global calibration method derived from the one used at the H1 experiment at DESY [43]. The other one is based on a Local Calibration (LC) of calorimeter cells. Both methods are presented in Section 10.5.3 of [21].

7.2.3.3 Choice of jet collection

After all reconstruction and calibration choices have been applied, we end up with seven choices of jets collections in ATLFASTII reconstructed events:

Cone 0.4 H1 Jets reconstructed with a Cone algorithm of size 0.4 and calibrated using the global H1 method;

Cone 0.4 LC Jets reconstructed with a Cone algorithm of size 0.4 and calibrated using the local calibration method;

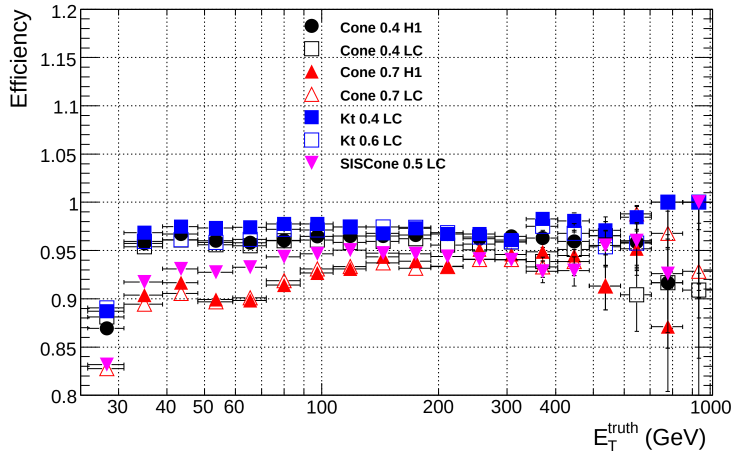
Cone 0.7 H1 Jets reconstructed with a Cone algorithm of size 0.7 and calibrated using the global H1 method;

Cone 0.7 LC Jets reconstructed with a Cone algorithm of size 0.7 and calibrated using the local calibration method;

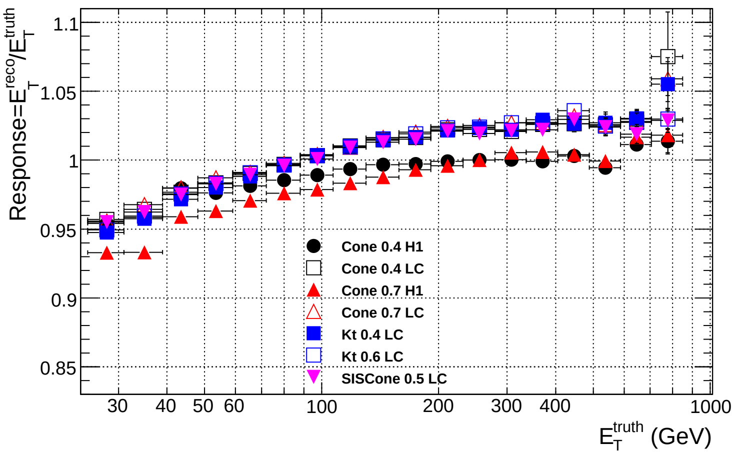
Kt 0.4 LC Jets reconstructed with a Kt algorithm of parameter 0.4 and calibrated using the local calibration method;

Kt 0.6 LC Jets reconstructed with a Kt algorithm of parameter 0.6 and calibrated using the local calibration method;

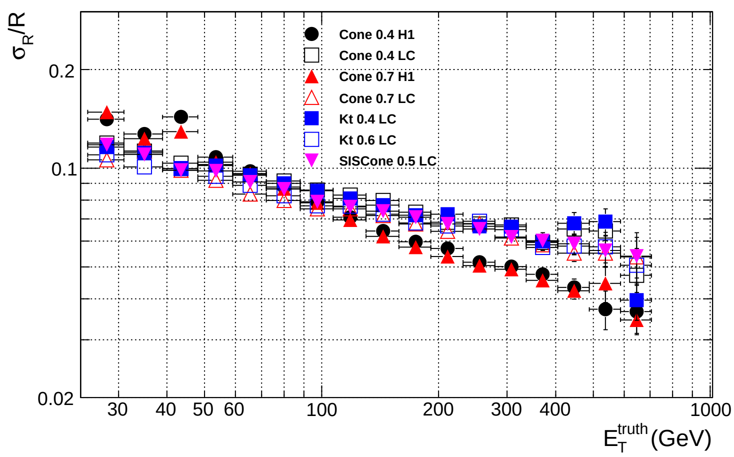
SISCone 0.5 LC Jets reconstructed with a SISCone algorithm of size 0.5 and calibrated using the local calibration method;



(a) Jet algorithms efficiencies



(b) Jet algorithms linearities



(c) Jet algorithms resolutions

Figure 7.7: Performance of the different ATLAS jet algorithms as a function of E_T^{jet}

As we aim at performing a jet counting experiment, the most important parameters from a point of view of a jet algorithm performance are the efficiency and the linearity. The following discussion is based on the plots from Figure 7.7, obtained by applying the Jet-Performance package directly to our private sample with $\alpha_s(M_{Z^0}) = 0.118$. For methods to measure these parameters on ATLAS data, refer to [25].

The efficiency of a jet algorithm (Figure 7.7a) is defined as the probability of a jet reconstructed from truth particles to be reconstructed in the reconstructed data, i.e. does the offline software find one Cone4 jet for each truth Cone4 jet. The jet algorithms are compared with their truth equivalent as they should exhibit the same properties. The matching is done by comparing the distance between the two jets with the jet algorithm parameter P : $\Delta R < P/2$. No energy matching is performed so that the same matching can be used to study efficiencies, linearities and resolutions. Figure 7.7a shows that Cone7 jets have the lowest efficiencies with SISCone jets. On the other end of the scale, Kt 0.4 LC clearly has the highest efficiency with $\sim 97\%$ in the range $40\ GeV < E_T < 200\ GeV$

The linearity (Figure 7.7b) is defined as the ratio of the reconstructed jet energy over the truth jet energy. It measures the average shift in energy introduced by the calorimeter and the reconstruction software. No algorithm is exactly linear. The global H1 calibration does a better job at high energy than the local calibration scheme in these simulated data. However, we expect most of our jets to have low energy so that local calibration shows better performances in this high interest region as well as 0.4 cone H1. Furthermore, the energy sensitivity is very important close to the jet cut-off threshold, so that again, linearity at low energies is more important.

Finally, the resolution (Figure 7.7c) is the standard deviation of the reconstructed energy for a set of jets of a given energy. Figure 7.7c displays resolutions after corrections of the reconstructed jet energy by the linearity extracted from Figure 7.7b. Here again, the global H1 calibration outperforms the local calibration method at high energies. Most other algorithms perform similarly over the whole range.

Global	E_T^{miss}	$<$	25 GeV
Electrons	$\lvert\eta^e\rvert$	$<$	2.5
	E_T^e	$>$	15 GeV
	Electron ID Cuts	$==$	Tight
	$N_{electrons}$	$==$	2
Jets	Jet collection	$==$	Kt 0.4 LC
	$\lvert\eta^{jet}\rvert$	$<$	3
	E_T^{jet}	$>$	40 GeV
Overlap	ΔR_{e-jet}	$>$	0.4

Table 7.2: Summary of selection used for α_s study

Since most of the algorithms perform very closely as far as energy reconstruction is performed, we decided to use the Kt 0.4 LC jets as they present a clear edge in term of efficiency for this sample. For similar reasons, we decided to study jets only above 40 GeV as the efficiency drops rapidly at lower energies.

Furthermore, the probability of reconstructing jets from pile-up interactions (interactions between two protons other than the ones producing the Z^0 boson) decreases rapidly with the jet transverse energy. A high threshold of 40 GeV implies that virtually all our jets will have been produced by the hard interaction.

Our jet selection then consists of all jets in the Kt 0.4 LC collection with $|\eta| < 3$, $E_T >$ 40 GeV and which do not overlap with any of our selected electrons. The selection cuts for our selection are summarized in Table 7.2

7.3 Event counting

For easier comparison between reconstructed data and truth level information, it is conventional to unfold the reconstructed data to hadron-level, i.e. the state of the event generation after parton showering and fragmentation took place, where we can also reconstruct truth jets. The typical formula used is :

$$N^{truth}_{n_{jets}=n} = \frac{N^{data}_{n_{jets}=n} - BG}{A \cdot \varepsilon_{electron}{}^2 \cdot \varepsilon_{trigger}},$$

where :

$N^{data}_{n_{jets}=n}$ is the number of reconstructed events in data, with n truth jets, which pass all our selection cuts;

A is the acceptance, i.e. the probability that a truth n-jets event present the right topological configuration to be reconstructed and selected.

$\varepsilon_{electron}$ is the probability of a truth electron to be reconstructed as an ElectronTight offline electron. The probability is squared because of the presence of two electrons;

$\varepsilon_{trigger}$ is the probability of an event with two tight electrons to pass the trigger;

Because the efficiency for electron reconstruction is not uniform as a function of either η or E_T, we weight each event independently to increase our precision. It is very difficult to measure efficiency for a given sample whereas we expect the efficiency to reconstruct individual object to be process-independent, so that we can use the information extracted using the method presented in Chapters 2 to 5.

As we presented in Chapter 5, we can measure the reconstruction efficiency as a 2-D distribution. As the same can be done for the efficiency of an offline electron to pass the ElectronTight cuts, we can use a 2-D map similar to the one presented on Figure 7.8 to extract $\varepsilon_{electron}$ for each electron individually.

Trigger information is not available in ATLFASTII, which is equivalent to $\varepsilon_{trigger} = 1$;

In this preliminary study, we select our truth objects using the same selection as our offline ones. As such, the acceptance correction is $A = 1$. This ensures that distributions such as $p_T^{Z^0}$ are not distorted by our electron requirements in terms of pseudo-rapidity and energy.

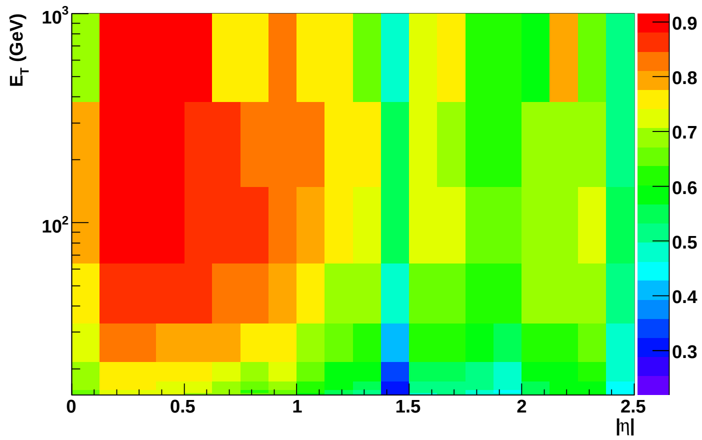


Figure 7.8: 2-D efficiency for tight electron reconstruction

It makes the comparison between parton-level measurement and reconstructed objects easier but it will be different for comparison with CMS results and combined studies. Truth level calculations are here compared with reconstructed calculations using the same acceptance for electrons.

The number $N^{data}_{n_{jets}=n}$ is harder to estimate. Because of the inefficiencies in the jet reconstruction algorithm, events are shifted between bins of n_{jets}. Events with one truth jet are reconstructed as events with no jets with a probability $1-\varepsilon_{jet}$, where ε_{jet} is the probability of a truth jet to be reconstructed as an offline jet. So that, if we label N^R_i the number of events where i jets were reconstructed (R for Reconstructed) and N^{data}_i the number of reconstructed events (with two reconstructed electrons and low missing transverse energy) which contain i truth jets, then N^{data}_i can be obtained by solving the system based on the upper triangular matrix:

$$\begin{bmatrix} N^R_0 \\ N^R_1 \\ N^R_2 \\ N^R_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1-\varepsilon_{jet} & (1-\varepsilon_{jet})^2 & (1-\varepsilon_{jet})^3 & \dots \\ 0 & \varepsilon_{jet} & \binom{2}{1}\varepsilon_{jet}(1-\varepsilon_{jet}) & \binom{3}{1}\varepsilon_{jet}(1-\varepsilon_{jet})_2 & \dots \\ 0 & 0 & (\varepsilon_{jet})^2 & \binom{3}{2}\varepsilon^2_{jet}(1-\varepsilon_{jet}) & \dots \\ 0 & 0 & 0 & (\varepsilon_{jet})^3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} N^{data}_0 \\ N^{data}_1 \\ N^{data}_2 \\ N^{data}_3 \\ \vdots \end{bmatrix}$$

Using the work by the Jet/Etmiss group in ATLAS, we should be able to extract the efficiency similar to the one presented on Figure 7.7a. Because this correction is applied on the integrated set and not event by event, it is important that the efficiency be as uniform as possible. The efficiency presents an important turn-on curve below our 40 GeV threshold, which reinforces the necessity for a high jet energy threshold. This matrix assumes that the jet efficiency is independent of the jet multiplicity in the event. This can be assumed to be the case as jets are local objects, unless the jets start to overlap. Because we study only events with few jets compared with the available $\eta - \phi$ space in the calorimeters, we assumed that this assumption holds.

Each jet energy will also be corrected for non-linearity using results of the type shown on Figure 7.7b as underestimating the jet energy would mean reconstructed jets would not pass the energy threshold while the truth jets would.

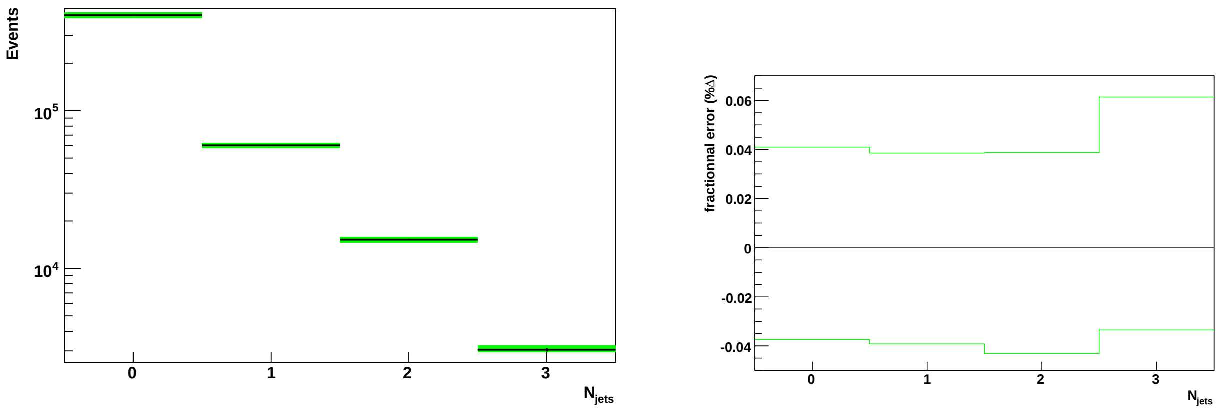
7.4 Sources of uncertainty

7.4.1 Background subtraction

We described in Chapter 4 a method to subtract background events from a $Z^0 \to e^+e^-$ event sample. As we described in Chapter 4, this procedure yields a systematic uncertainty linked with the choice of initial parameters. We use standard error propagation techniques to evaluate the resulting error on the measured distribution.

7.4.2 Electron efficiency uncertainty

The second source of uncertainty is the one arising from our measurement of the electron efficiency from data. Because we use this efficiency when we unfold the number of events to hadron level, this leads to an experimental uncertainty.



(a) Absolute uncertainty (b) fractional uncertainty

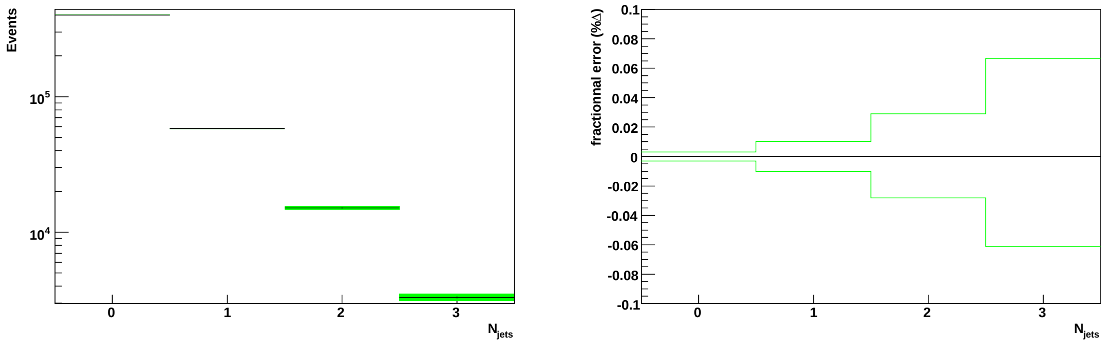
Figure 7.9: Uncertainty due to the electron efficiency on the measured number of jets

We estimate this error by varying the estimated electron efficiency. From the result of Chapter 5, we expect to be able to measure the electron efficiency to within 2% or below. Therefore, we measure the distributions with an uncertainty of ±2%, which is slightly worse than the precision the ATLAS collaboration hopes to achieve with 1 fb^{-1}. One such distribution can be observed on Figure 7.9.

7.4.3 Jet efficiency uncertainty

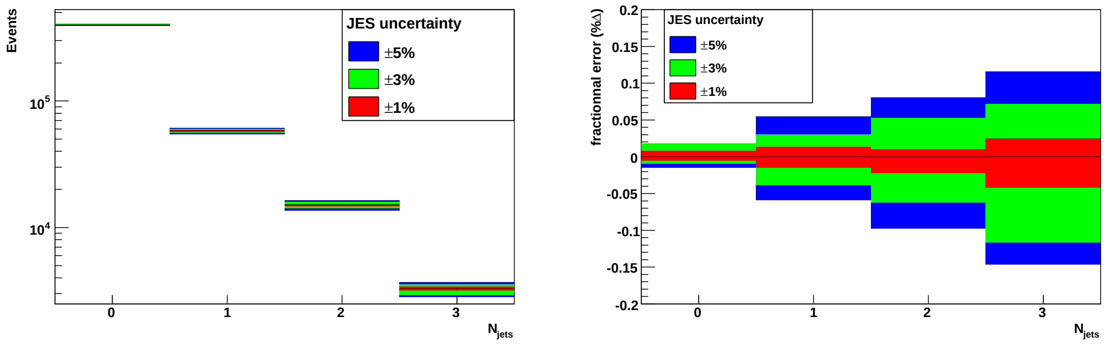
Just like electrons, jets are a source of uncertainty. This uncertainty arises from the limited accuracy of our measurement of the jet efficiencies from data. Despite the large statistics, it will be difficult to select a clean jet sample. Since we use the efficiency when we unfold the number of events to hadron level, this also leads to an experimental uncertainty.

We estimate this error by varying the estimated jet efficiency. It is difficult to estimate which precision the efficiency determination will reach. Different methods have been proposed including methods using jet balancing methods against the recoil of electron-magnetic objects like photons or Z^0 bosons. The expected performance of these methods are however very uncertain. Therefore, we decided to assume an uncertainty comparable to the one arising from the electron efficiency of ±2%. The resulting uncertainties can be observed in



(a) Absolute uncertainty (b) fractional uncertainty

Figure 7.10: Uncertainty due to the jet efficiency on the measured number of jets



(a) Absolute uncertainty (b) fractional uncertainty

Figure 7.11: Jet energy scale uncertainty on the measured number of jets

Figure 7.10.

7.4.4 Jet energy scale uncertainty

The second source of uncertainty is linked to the fact that we only accept jets above a certain threshold. Because of that, we might underestimate the number of jets in the event if we underestimate the energy or overestimate it in the other case. The jet energy scale error result from remaining uncertainties after our the estimation of the linearity of the jet algorithm calibration.

We estimate this error by varying the estimation of the jet energies. For this study, we measure the distributions with uncertainties of ±1%, ±3% and ±5%, the latter being the precision the ATLAS collaboration hopes to achieve with 1 fb^{-1}. A JES of ±5% has been quickly achieved at previous experiments, including those at the TeVatron, so that it is a reasonable expectation that the ATLAS calibration experts may achieve a similar precision.

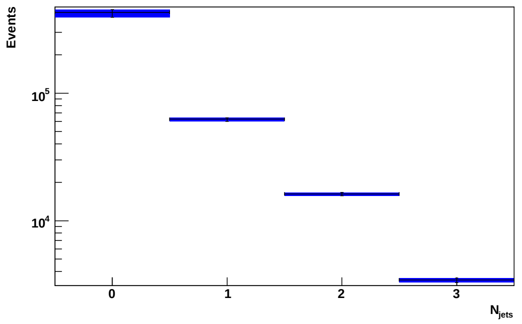
All the sources of uncertainty listed are linked with our limited knowledge of the occurring processes. However, our understanding of the underlying theory is also limited. Apart from uncertainties due to missing higher order corrections, the main source of uncertainty is linked with our knowledge of the Parton Distribution Functions.

7.4.5 PDF uncertainty

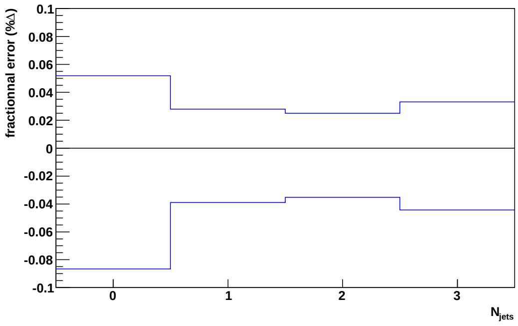
We already mentioned the impact that the choice of PDF can have on the event generation process. PDF are extracted from global fits to data from many experiments. The fits, apart from a fixed value of α_s, contain, for the CTEQ series, 22 free parameters, also called eigenvectors. To study the effect of 1σ errors on the estimation of these parameters, the CTEQ collaboration releases what they call 44 error PDFs [44], which are PDFs where one of these eigenvectors is modified up or down so that the χ^2 of the fit varies by 10 [44].

To study the influence of the eigenvector-based errors, we use a technique called PDF reweighting [45]. For an event with parton 1 carrying a momentum fraction x_1, parton 2 carrying a momentum fraction x_2 and undergoing a hard-process of scale Q^2, we can calculate the probabilities of such an event happening with two PDFs A and B, respectively p_A and p_B. For a whole dataset simulated with PDF A, we can obtain distributions similar to those obtained from a dataset simulated with PDF B by reweighting each event by the corresponding fraction p_B/p_A.

The PDF uncertainty is then calculated by adding in quadrature, the errors computed as the difference between the values obtained from the nominal PDF ($V_{nominal}$) and the values



(a) Absolute uncertainty from PDF set CTEQ6M



(b) Fractional uncertainty from PDF set CTEQ6M

Figure 7.12: Error estimation from PDF uncertainty on the measured number of jets

obtained from each of the eigenvector variations ($V_{-1\sigma}$ and $V_{+1\sigma}$):

$$\begin{aligned} error^{+} &= \max(0, V_{-1\sigma} - V_{nominal}, V_{+1\sigma} - V_{nominal}), \\ error^{-} &= \max(0, V_{nominal} - V_{-1\sigma}, V_{nominal} - V_{+1\sigma}). \end{aligned}$$

This gives us “positive” and “negative” errors linked with the PDF uncertainty as seen on Figure 7.12a.

Because the 44 error PDFs are provided by the CTEQ collaboration as variation of the CTEQ6M best fit PDF, we calculate the errors relative to that central value and then apply them to the value measured with the CTEQ6A118 PDF.

7.5 α_s sensitivity

This section describes the measurement of some distributions and studies the corresponding sensitivity to α_s.

	R_0	R_1	R_2
Measured from reconstructed data ($\times 100$)			
value	14.56	25.96	21.86
statistical error	0.07	0.25	0.44
background systematics	0.09	0.35	0.64
electron efficiency $\pm 2\%$	$^{+0.00}_{-0.03}$	$^{+0.03}_{-0.46}$	$^{+0.49}_{-0.00}$
jet efficiency $\pm 2\%$	$^{+0.20}_{-0.19}$	$^{+0.48}_{-0.47}$	$^{+0.80}_{-0.74}$
JES $\pm 1\%$	$^{+0.28}_{-0.32}$	$^{+0.00}_{-0.20}$	$^{+0.32}_{-0.44}$
JES $\pm 3\%$	$^{+0.59}_{-0.81}$	$^{+0.57}_{-0.65}$	$^{+0.39}_{-1.26}$
JES $\pm 5\%$	$^{+1.03}_{-1.10}$	$^{+0.63}_{-1.08}$	$^{+0.72}_{-1.17}$
Measured from parton-level ($\times 100$)			
value	14.53	25.97	21.14
statistical error	0.06	0.23	0.40
PDF uncertainty	$^{+1.43}_{-1.03}$	$^{+0.80}_{-0.79}$	$^{+0.68}_{-0.68}$

Table 7.3: Measurement of the ratio $R_N = \frac{Z^0 \to e^+e^- + (N+1)\ jets}{Z^0 \to e^+e^- + N\ jets}$. The errors associated with different uncertainties are quoted. All values have been scaled by 100 for readability.

R_N	$\alpha_s = 0.118$		$\alpha_s = 0.110$			$\alpha_s = 0.128$		
	mean ×100	stat.	mean ×100	stat.	diff (σ)	mean ×100	stat.	diff (σ)
R_0	14.53	0.06	14.23	0.06	3.35	16.87	0.07	25.10
R_1	25.97	0.23	24.96	0.23	3.11	27.49	0.22	4.74
R_2	21.14	0.40	19.02	0.39	3.80	20.74	0.36	0.75

Table 7.4: Truth value of R_N for samples with different values of α_s. The means and errors have been scaled by a factor of 100 for readability. Diff indicates the separation between a sample and the 118 reference sample in terms of statistical error.

7.5.1 Ratio $R_N = \frac{Z^0 \to e^+e^- + (N+1)\ jets}{Z^0 \to e^+e^- + N\ jets}$

As we discussed in Section 7.1, the primary quantity to observe a sensitivity to α_s is the ratio $R_N = \frac{Z^0 \to e^+e^- + (N+1)\ jets}{Z^0 \to e^+e^- + N\ jets}$, in particular, $R_0 = \frac{Z^0 \to e^+e^- + 1\ jet}{Z^0 \to e^+e^- + 0\ jets}$, with its high statistics. Table 7.3 compares the measured ratio of the ratios R_0, R_1 and R_2 with their generator-level values and the uncertainty from the different sources listed in Section 7.4. For readability, the values in Table 7.3 have been multiplied by a factor of 100, i.e. the measured value of R_0 is 0.1456 with a statistical error of 0.0007.

We notice that statistical errors, background subtraction systematic errors and electron efficiency are the smallest sources of uncertainty, while, uncertainties on jet parameters are the largest source of uncertainty on the measured value of the ratio.

PDF uncertainties on the theoretical value are however the principal source of uncertainty and are comparable with the largest uncertainty on the reconstructed measured value. This is worrisome, as any α_s sensitivity necessitates to be able to distinguish values at parton-level from dataset simulated with different α_s values.

However, while Table 7.3 shows our ability to measure the first few ratios R_N using reasonably calibrated ATLAS data, Table 7.4 shows that the statistical power to separate the samples with different values of α_s is insufficient.

The separation in R_0 between the samples $\alpha_s = 0.110$ and $\alpha_s = 0.118$ for R_0 is only 0.003, while the theoretical error on the value form $\alpha_s = 0.118$ from PDF uncertainties alone is ${}^{+0.0143}_{-0.0103}$.

Most of the PDF uncertainty comes from the fact that most of the electroweak interactions at the LHC are produced at a regime where the PDFs are poorly constrained (low-x gluon collisions). It is possible that the PDF uncertainty will be rapidly reduced using, in particular, W production [46] at the LHC.

If the PDF uncertainties are constrained enough that the different samples become distinguishable ($|R_N^{\alpha_s=0.118} - R_N^{\alpha_s=0.110}| \gtrsim 3\sigma$), then the limiting constraint becomes the jet energy scale. Bringing the jet energy scale to within 1% would enable us to separate the true value (0.118 in this study) from the extreme values 0.110 and 0.128. Further sensitivity would require more precise knowledge of the jet algorithm performances. While it is unclear how far we can improve the jet efficiency in the LHC environment, its measurement is difficult using only data and mostly relies on Monte-Carlo tuning.

Because the emission of initial radiation hard-jets will generally result in a transverse boost of the exchanged quark, it is possible that looking at transverse parameter distributions will yield a better sensitivity.

7.5.2 Ratio R_N as a function of $p_T^{Z^0}$

Table 7.5 shows the measurement of R_0 in four bins of $p_T^{Z^0}$. This distribution is probably the most efficient to study R_0 as most events with no jets will have very low $p_T^{Z^0}$ while events with more than one jet will see the Z^0 recoil against the hadronic components of the event.

We study this distribution in the range $[0-60]$ GeV. Figure 7.13 shows that the statistics in the bin $[45-60]$ GeV are still sufficient for $N_{jets} == 0$ to study the the ratio. These high p_T events are produced in association with many low-p_T jets which fail our threshold.

As we can see easily from Table 7.5 and Figure 7.14, the PDF uncertainty on R_0 is not a flat function of $p_T^{Z^0}$. The uncertainty is much lower for low values of $p_T^{Z^0}$ because most of the

$p_T^{Z^0}$ (GeV)	0 – 15	15 – 30	30 – 45	45 – 60
Measured from reconstructed data ($\times$100)				
value	0.52	4.73	28.84	112.00
statistical error	0.01	0.07	0.33	1.45
background systematics	0.02	0.10	0.47	2.06
electron efficiency $\pm$2%	$^{+0.01}_{-0.00}$	$^{+0.00}_{-0.04}$	$^{+0.43}_{-0.00}$	$^{+0.00}_{-0.86}$
jet efficiency $\pm$2%	$^{+0.00}_{-0.00}$	$^{+0.02}_{-0.02}$	$^{+0.53}_{-0.51}$	$^{+4.49}_{-4.15}$
JES $\pm$1%	$^{+0.01}_{-0.01}$	$^{+0.08}_{-0.30}$	$^{+0.93}_{-1.48}$	$^{+3.71}_{-6.54}$
JES $\pm$3%	$^{+0.03}_{-0.05}$	$^{+0.37}_{-0.66}$	$^{+3.64}_{-3.96}$	$^{+9.39}_{-14.93}$
JES $\pm$5%	$^{+0.06}_{-0.07}$	$^{+0.61}_{-0.95}$	$^{+5.65}_{-5.64}$	$^{+18.21}_{-22.19}$
Measured from parton-level Monte-Carlo ($\times$100)				
value	0.52	4.24	25.65	106.23
statistical error	0.01	0.07	0.29	1.30
PDF uncertainty	$^{+0.05}_{-0.05}$	$^{+0.23}_{-0.22}$	$^{+0.43}_{-0.38}$	$^{+1.92}_{-1.98}$

Table 7.5: Measurement of the ratio R_0 as a function of $p_T^{Z^0}$. The errors associated with different uncertainties are quoted. All values have been scaled by 100 for readability.

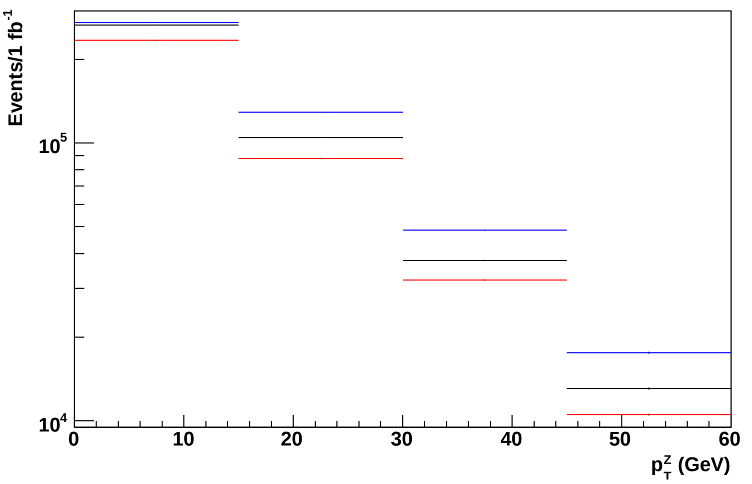


Figure 7.13: Number of events in the sample where $N_{jets} == 0$ as a function of $p_T^{Z^0}$ for samples with different values of α_s. Black is nominal $\alpha_s = 0.118$, blue is $\alpha_s = 0.110$ and red is $\alpha_s = 0.128$.

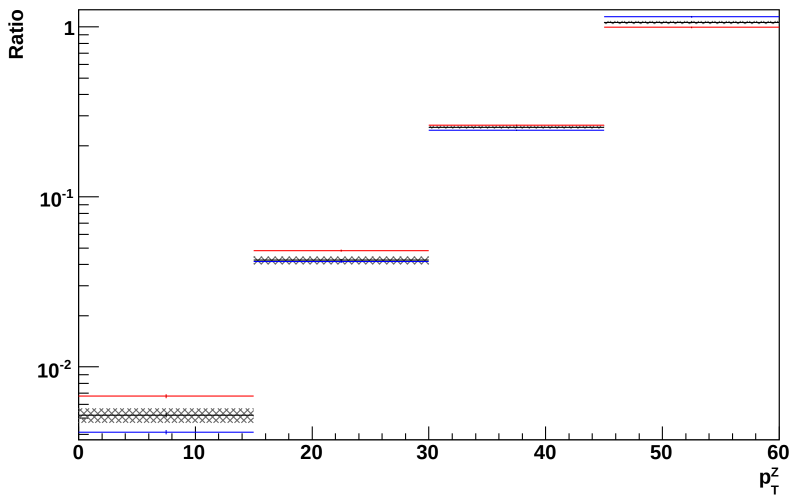


Figure 7.14: Truth value of R_0 as a function of $p_T^{Z^0}$ for samples with different values of α_s. Black is nominal $\alpha_s = 0.118$, blue is $\alpha_s = 0.110$ and red is $\alpha_s = 0.128$. The shaded areas represent the PDF uncertainty on the $\alpha_s = 0.118$ value.

$p_T^{Z^0}$ (GeV)	$\alpha_s = 0.118$		$\alpha_s = 0.110$			$\alpha_s = 0.128$		
	mean $\times 100$	stat.	mean $\times 100$	stat.	diff (σ)	mean $\times 100$	stat.	diff (σ)
0.00–15.00	0.52	0.01	0.41	0.01	5.63	0.67	0.02	6.91
15.00–30.00	4.24	0.07	4.16	0.07	0.83	4.82	0.07	6.25
30.00–45.00	25.65	0.29	24.65	0.29	2.41	26.42	0.28	1.91
45.00–60.00	106.23	1.30	114.72	1.44	4.37	99.43	1.14	3.92

Table 7.6: Truth values of R_0 as a function of $p_T^{Z^0}$ for samples with different values of α_s

events for $Z^0 \rightarrow e^+e^- + 0$ jets are produced by quark-antiquark interactions for which the uncertainty is lower. A high value of $p_T^{Z^0}$ betrays a hard jet emission, which is more sensitive to gluon contributions as seen in the Feynman diagrams on Figure 7.3 and 7.4.

Overall, the uncertainty in this first bin is much smaller than the error on the integrated R_0 seen in Table 7.3 and a JES uncertainty of $\pm 5\%$ only induces an uncertainty which is double what we expect from our background subtraction.

The same effect is observed when comparing the truth values obtained using the different values of α_s in Figure 7.14 and Table 7.6. At low $p_T^{Z^0}$ the separation between $R_0^{0.110}$ and $R_0^{0.118}$ is of the order of 2σ in terms of theoretical uncertainty (including PDF uncertainty). In comparison, the uncertainty on the measured value from reconstructed data is of the order

$E_T^{Leading jet}$ (GeV)	40–60	60–80	80–100	100–120
Measured from reconstructed data ($\times$100)				
value	9.22	27.06	40.47	55.00
statistical error	0.18	0.50	0.96	1.62
background systematics	0.26	0.74	1.38	2.29
electron efficiency $\pm$2%	$^{+0.03}_{-0.08}$	$^{+0.00}_{-0.38}$	$^{+0.00}_{-0.40}$	$^{+0.98}_{-0.00}$
jet efficiency $\pm$2%	$^{+0.19}_{-0.19}$	$^{+0.75}_{-0.71}$	$^{+0.85}_{-0.82}$	$^{+1.36}_{-1.30}$
JES $\pm$1%	$^{+0.32}_{-0.43}$	$^{+0.36}_{-0.75}$	$^{+0.00}_{-1.25}$	$^{+1.31}_{-1.17}$
JES $\pm$3%	$^{+1.49}_{-1.25}$	$^{+1.26}_{-2.38}$	$^{+1.18}_{-3.10}$	$^{+1.93}_{-4.37}$
JES $\pm$5%	$^{+2.51}_{-1.71}$	$^{+2.30}_{-3.50}$	$^{+3.23}_{-3.94}$	$^{+3.09}_{-4.86}$
Measured from parton-level Monte-Carlo ($\times$100)				
value	9.00	27.19	43.28	57.43
statistical error	0.17	0.49	0.97	1.64
PDF uncertainty	$^{+0.28}_{-0.29}$	$^{+0.80}_{-0.78}$	$^{+1.27}_{-1.25}$	$^{+1.83}_{-1.73}$

Table 7.7: Measurement of the ratio R_1 as a function of $E_T^{Leading jet}$. The errors associated with different uncertainties are quoted. All values have been scaled by 100 for readability.

of 0.0005, which would be the limit for a 95% exclusion for the extreme values of α_s.

Other bins of the same distribution could also yield some separation if the PDF uncertainty were to be reduced significantly. For the most part, the efforts needed to include these bins in a measurement would be the same as the one needed for the inclusive R_N ratios.

7.5.3 Ratio R_N as a function of $E_T^{Leading jet}$

The distribution of R_N as a function of $E_T^{Leading jet}$ is another distribution where we expect α_s to induce a change in shape. Unfortunately, this means discarding the $Z^0 \rightarrow e^+e^- + 0$ jets subsample, which is the largest part of the $Z^0 \rightarrow e^+e^-$sample which means the statistics are limited. This implies that the statistical error on the measurement is larger and becomes

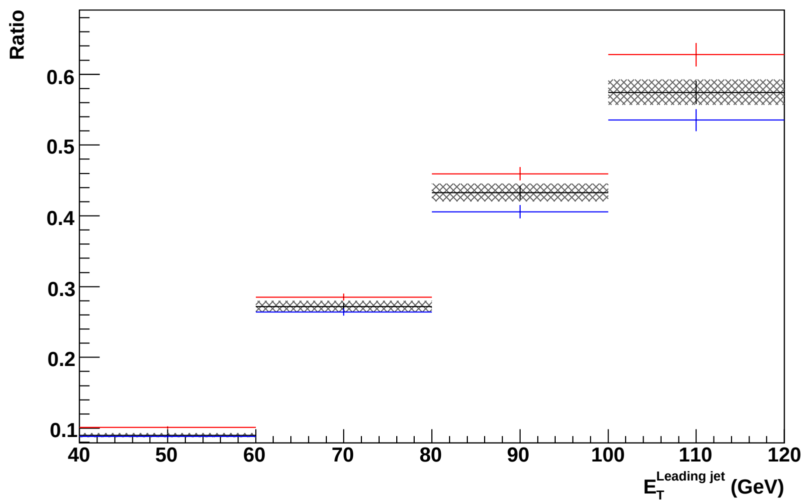


Figure 7.15: Truth value of R_1 as a function of $E_T^{Leading jet}$ for samples with different values of α_s. Black is nominal $\alpha_s = 0.118$, blue is $\alpha_s = 0.110$ and red is $\alpha_s = 0.128$. The shaded areas represent the PDF uncertainty on the $\alpha_s = 0.118$ value.

$E_T^{Leading jet}$ (GeV)	$\alpha_s = 0.118$		$\alpha_s = 0.110$			$\alpha_s = 0.128$		
	mean $\times 100$	stat.	mean $\times 100$	stat.	diff (σ)	mean $\times 100$	stat.	diff (σ)
40.00–60.00	9.00	0.17	8.80	0.17	0.79	10.08	0.17	4.42
60.00–80.00	27.19	0.49	26.42	0.50	1.11	28.52	0.48	1.94
80.00–100.00	43.28	0.97	40.57	0.94	2.01	45.93	0.95	1.96
100.00–120.00	57.43	1.64	53.53	1.57	1.72	62.78	1.66	2.30

Table 7.8: Truth value of R_0 as a function of $E_T^{Leading jet}$ for samples with different values of α_s

comparable with the PDF uncertainty over the whole range as seen in Table 7.7.

A different value of α_s modifies the probability of the emission of hard jets. Hence the sensitivity in the leading jet E_T distribution is mostly observed at higher values of $E_T^{Leading jet}$ as we can see on Figure 7.15 and Table 7.8.

Again, the cumulated measurement uncertainty using the best JES estimate is comparable with the theoretical uncertainty including the PDF uncertainty and with the separation of the measured ratio for different values of α_s.

7.6 Summary

We have shown that we can measure the value of the ratio $R_N = \frac{Z^0 \to e^+ e^- + (N+1)\ jets}{Z^0 \to e^+ e^- + N\ jets}$ with an accuracy about equal to the current theoretical uncertainty. Unfortunately, the current theory does not provide enough separating power to be able to use this information to constrain the value of α_s.

Improvements to the global PDF fits from LHC data will rapidly reduce the theoretical uncertainties as the gluon structure of the proton becomes better constrained at low x. It will then be a question of how well we manage to calibrate the detector with early data. In particular, understanding the jet algorithms performances will be crucial to get an accurate estimate of the ratio in this channel.

Chapter 8

Conclusions

Just as the LHC is gearing up to take the lead in high-energy physics research, the ATLAS collaboration is preparing to meet the challenges of this new frontier. The search for new physics predicted by various theories will require precise knowledge of the detector hardware condition, software environment and a good understanding of the existing theories in this new regime. This thesis is part of this collaborative effort.

To start with, a method has been developed to measure the electron reconstruction efficiency in $Z^0 \to e^+e^-$ events, using both initial data as it becomes available and data collected at higher integrated luminosities for higher precision and/or granularity. The initial measurement will be used to identify possible under-performing detector regions and, if found, understand whether the problems are hardware or software related. More precise measurements will be available later for the benefit of all other physics groups to use in their analyses, such as the many cross-section measurements ATLAS will perform.

The algorithm has been shown to perform in different background conditions as the hadronic background at high energy is still unknown. The background subtraction method presented is adaptable to different running conditions of the LHC and does not rely on Monte-Carlo predictions.

The precision of the measurement depends on the background conditions and on the statistics available. It is however expected that the electron efficiency will be measured with an accuracy of $\pm 2\%$ with a data sample of 1 fb^{-1} for electrons with an intermediate transverse energy, 20 GeV $< E_T <$ 100 GeV.

Because of the precision of the ATLAS simulation software, simulation of events is a highly time consuming process. New physics studies and private “toy” production will require a faster way to simulate new physics conditions than the standard Geant4-based reconstruction model. A faster lighter reconstruction framework was investigated and compared with the original framework, showing sufficient accuracy for first feasability studies of possible physics analyses. Simulations of different processes were studied and limitied consistency was achieved for a benchmark analysis relying mostly on electron identification and kinematics. Further improvement are, however needed if the fast simulation and reconstruction framework is to be used for analyses with more complex detector-based selection cuts.

Finally, the influence of α_s on different distributions in Z + jets events was studied. Despite limited sensitivity to relatively large variations of the parameter, accurate measurements have been performed. Improvements to the parton distribution function with LHC data should increase our sensitivity which will then mainly be limited by our knowledge of the jet energy scale and jet reconstruction afficiency.

Appendix A

Datasets used in the CSC exercise

The datasets names are :

$pp \to Z^0 \to e^+e^-$ mc08.106050.PythiaZee_1Lepton.recon.AOD.e347_s462_r604/ ;

$pp \to Z^0 \to \tau^+\tau^-$ mc08.106052.PythiaZtautau.recon.AOD.e347_s462_r604/;

$pp \to t\bar{t}$ mc08.105200.T1_McAtNlo_Jimmy.recon.AOD.e357_s462_r635/;

$pp \to W \to e\nu$ mc08.106020.PythiaWenu_1Lepton.recon.AOD.e352_s462_r604/;

$pp \to jet + jet$ mc08.105802.JF17_pythia_jet_filter.recon.AOD.e347_s462_r604/

The final part of each dataset name indicates which releases were used for the different stages of the centralized ATLAS production. They can be interpreted as follows :

e347 the events were generated using version 14.2.0.2;

e352 the events were generated using version 14.2.20.1;

e357 the events were generated using version 14.2.21.1;

s462 the events were fully simulated (Geant4) using version 14.2.10.1;

r604 the events were reconstructed using version 14.2.25.2;

r635 the events were reconstructed using version 14.2.25.8;

Appendix B

Datasets used in the validation of ATLFASTII

The datasets used for both types of reconstruction always come from the same generation datasets. We compared datasets for the signal and the four backgrounds we described in section 4.

B.1 $pp \to Z^0 \to e^+e^-$ datasets

simulation step	dataset name
Generation	mc08.106050.PythiaZee_1Lepton.evgen.EVNT.e347/
Full simulation	mc08.106050.PythiaZee_1Lepton.merge.AOD.e347_s462_r635_t53/
Fast simluation	mc08.106050.PythiaZee_1Lepton.merge.AOD.e347_a84_t53/
Corrected fast simulation	user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.106050. PythiaZee_1Lepton.merge.AOD.e347_a84_t53_tid060104 user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.106050. PythiaZee_1Lepton.merge.AOD.e347_a84_t53_tid060105 user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.106050. PythiaZee_1Lepton.merge.AOD.e347_a84_t53_tid060106

Table B.1: Datasets involved in ATLFASTII validation for $pp \to Z^0 \to e^+e^-$

The dataset involved are listed in Table B.1.

All datasets were generated using Athena version 14.2.0.2.

The full simulation was run using version 14.2.10.1 and the reconstruction run using 14.2.25.8. Finally the output files were merged using version 14.2.25.6

The fast simulation was performed using version 14.2.25.8 and the files were subsequently merged using version 14.2.25.6.

The fast simulation with post-simulation corrections were applied in Release 14.2.25 using the packages MonteCarloReactTools-00-02-04-12 and MonteCarloReactUtils-00-00-12.

B.2 $pp \to Z^0 \to \tau^+\tau^-$ datasets

simulation step	dataset name
Generation	mc08.106052.PythiaZtautau.evgen.EVNT.e347/
Full simulation	mc08.106052.PythiaZtautau.merge.AOD.e347_s462_r635_t53/
Fast simluation	mc08.106052.PythiaZtautau.merge.AOD.e347_a84_t53/
Corrected Fast simulation	user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.106052. PythiaZtautau.merge.AOD.e347_a84_t53_tid056545

Table B.2: Datasets involved in ATLFASTII validation for $pp \to Z^0 \to \tau^+\tau^-$

The dataset involved are listed in Table B.2.

All datasets were generated using Athena version 14.2.0.2.

The full simulation was run using version 14.2.10.1 and the reconstruction run using 14.2.25.8. Finally the output files were merged using version 14.2.25.6

The fast simulation dataset was performed out using version 14.2.25.8 and the files were subsequently merged using version 14.2.25.6.

The fast simulation with post-simulation corrections were applied in Release 14.2.25 using the packages MonteCarloReactTools-00-02-04-12 and MonteCarloReactUtils-00-00-12.

B.3 $pp \to t\bar{t}$ datasets

simulation step	dataset name
Generation	mc08.105200.T1_McAtNlo_Jimmy.evgen.EVNT.e357/
Full simulation	mc08.105200.T1_McAtNlo_Jimmy.merge.AOD.e357_s462_r635_t53/
Fast simluation	mc08.105200.T1_McAtNlo_Jimmy.merge.AOD.e357_a84_t53/
Corrected Fast simulation	user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.105200. T1_McAtNlo_Jimmy.merge.AOD.e357_a84_t53_tid059047 user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.105200. T1_McAtNlo_Jimmy.merge.AOD.e357_a84_t53_tid059048 user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.105200. T1_McAtNlo_Jimmy.merge.AOD.e357_a84_t53_tid059049

Table B.3: Datasets involved in ATLFASTII validation for $pp \to t\bar{t}$

The dataset involved are listed in Table B.3.

All datasets were generated using Athena version 14.2.21.1.

The full simulation was run using version 14.2.10.1 and the reconstruction run using 14.2.25.8. Finally the output files were merged using version 14.2.25.6

The fast simulation was performed using version 14.2.25.8 and the files were subsequently merged using version 14.2.25.6.

The fast simulation with post-simulation corrections were applied in Release 14.2.25 using the packages MonteCarloReactTools-00-02-04-12 and MonteCarloReactUtils-00-00-12.

B.4 $pp \to W \to e\nu$ datasets

simulation step	dataset name
Generation	mc08.106020.PythiaWenu_1Lepton.evgen.EVNT.e352/
Full simulation	mc08.106020.PythiaWenu_1Lepton.recon.AOD.e352_s462_r635/
Fast simluation	mc08.106020.PythiaWenu_1Lepton.recon.AOD.e352_a84/
Corrected Fast simulation	user09.MarkHohlfeld.ganga.MonteCarloReact.mc08.106020. PythiaWenu_1Lepton.merge.AOD.e352_a84_t53_tid071425

Table B.4: Datasets involved in ATLFASTII validation for $pp \to W \to e\nu$

The dataset involved are listed in Table B.4.

All datasets were generated using Athena version 14.2.20.1 and reconstructed from dataset

The full simulation dataset used was run using version 14.2.10.1 and the reconstruction run using 14.2.25.8.

The fast simulation dataset used was carried out using version 14.2.25.8.

The fast simulation with post-simulation corrections were applied in Release 14.2.25 using the packages MonteCarloReactTools-00-02-04-12 and MonteCarloReactUtils-00-00-12.

Appendix C

SHERPA configuration files

C.1 Analysis.dat

```
BEGIN_ANALYSIS {
    LEVEL  Showers
    PATH_PIECE  _Analysis
    Statistics  FinalState
} END_ANALYSIS
```

C.2 Model.dat

```
!-----------------------------------------------------------------------
!-- Model parameters --------------------------------------------------
!-----------------------------------------------------------------------
MODEL                   = SM        ! Model
!
! SM parameters
!
EW_SCHEME               = 0         ! which parameters define
   the ew sector.
ALPHAS(MZ)              = 0.110     ! strong coupling at scale
   M_Z
ALPHAS(default)         = 0.0800    ! strong coupling
ORDER_ALPHAS            = 1         ! NLO
1/ALPHAQED(0)           = 137.036   ! inverse of alpha QED in
   the Thomson limit
1/ALPHAQED(default)     = 132.51    ! inverse of alpha QED
```

```
SIN2THETAW             = 0.2222    ! Weinberg angle at scale
   M_Z
VEV                    = 246.      ! Higgs vev
LAMBDA                 = 0.47591   ! SM Higgs self coupling
CKMORDER               = 0         ! order of expansion of CKM
   matrix in Cabibbo angle
CABIBBO                = 0.22      ! Cabibbo angle (Wolfenstein
    parametrization)
A                      = 0.85      ! A (Wolfenstein
   parametrization)
RHO                    = 0.50      ! rho (Wolfenstein
   parametrization)
ETA                    = 0.50      ! eta (Wolfenstein
   parametrization)
!
! SUSY parameters
!
SLHA_INPUT             = LesHouches_SPS1A.dat
!
! ADD parameters
!
N_ED                   = 2          ! Number of extra
   dimensions in ADD model
G_NEWTON               = 6.707e-39 ! Newton gravity constant
M_S                    = 2.5e3      ! string scale for ADD
   model, meaning depends on KK_CONVENTION
M_CUT                  = 2.5e3      ! cut-off scale for the c.m
   . energy
KK_CONVENTION          = 1          ! 0=const mass
                                    ! 1=simplified sum(HLZ)
                                       2=exact sum(HLZ)
                                    ! 3=Hewett +1    4=Hewett
                                       -1
                                    ! 5=GRW (Lambda_T(GRW)=M_S
                                       for virtual Graviton
                                       exchange
                                    !        or M_D=M_S for
                                       real Gravtion production
                                       )
                                    ! only KK_CONVENTION's 1,2
                                       or 5 are applicable for
                                       real Graviton production
                                       !

!
! Particles Tau Decays
```

```
!
STABLE[15]=0
!
! Particles top
!
Mass[6]=172.5
!
! Particles W
!
MASS[24]=80.403
WIDTH[24]=2.141
!
! Particles Z
!
MASS[23]=91.1876
WIDTH[23]=2.4952
```

C.3 Processes.dat

```
%-----------------------------------------------------------------
%-- Processes to calculate ---------------------------------------
%-----------------------------------------------------------------
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% jet jet -> electron positron + n-jets
%
Process : 93 93 -> 11 -11 93{3}
Print_Graphs
Order electroweak : 2
End process
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% jet jet -> mu+ mu- + n-jets
%
% Process : 93 93 -> 13 -13 93{3}
% Print_Graphs
% Order electroweak : 2
% End process
% %
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% %
% % jet jet -> electron positron + n-jets
% %
% Process : 93 93 -> 15 -15 93{3}
```

```
% Print_Graphs
% Order electroweak : 2
% End process
% %
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

C.4 Run.dat

```
(beam){
  !------------------------------------------------------------------------
  !-- Beam parameters -----------------------------------------------------
  !------------------------------------------------------------------------
  BEAM_1            = 2212                  ! possible beam
     particles:  P+, P-, e+, e-
  BEAM_ENERGY_1     = 7000.                 ! in GeV
  BEAM_POL_1        = 0.                    ! Polarization
     degree -1 ... 1
  BEAM_SPECTRUM_1 = Monochromatic           ! Monochromatic,
     Laser_Backscattering,Simple_Compton
  K_PERP_MEAN_1 = 0.2
  K_PERP_SIGMA_1          = 0.8
  !
  BEAM_2            = 2212                  ! possible beam
     particles:  P+, P-, e+, e-
  BEAM_ENERGY_2     = 7000.                 ! in GeV
  BEAM_POL_2        = 0.                    ! Polarization
     degree -1 ... 1
  BEAM_SPECTRUM_2 = Monochromatic           ! Monochromatic,
     Laser_Backscattering,Simple_Compton
  K_PERP_MEAN_2 = 0.2
  K_PERP_SIGMA_2          = 0.8
  !
  BEAM_SMIN         = 1.e-10                ! Minimal fraction
     of nominal s after beam spectra
  BEAM_SMAX         = 1.0                   ! Maximal fraction
     of nominal s after beam spectra
  !
  E_LASER_1            = 1.17e-9            ! Laser energy in
    GeV
  P_LASER_1            = 0.                 ! Laser polarization
     +-1
  E_LASER_2            = 1.17e-9            ! Laser energy in
    GeV
  P_LASER_2            = 0.                 ! Laser polarization
     +-1
```

```
  LASER_MODE           = 0                    ! 0 = all, 1,2,3 =
     individual components
  LASER_ANGLES         = Off                  ! On/Off
  LASER_NONLINEARITY   = On                   ! On/Off
}(beam)
(isr){
  !------------------------------------------------------------
  !-- ISR parameters ------------------------------------------
  !------------------------------------------------------------
  BUNCH_1          = 2212       ! possible beam particles: P+,
     P-, e+, e-
  ISR_1            = On         ! On/Off
  !
  BUNCH_2          = 2212       ! possible beam particles: P+,
     P-, e+, e-
  ISR_2            = On         ! On/Off
  !
  ISR_SMIN         = 1.e-10     ! Minimal fraction of nominal s
     for parton after ISR
  ISR_SMAX         = 1.0        ! Maximal fraction of nominal s
     for parton after ISR
  !
  ISR_E_ORDER      = 1          ! Perturbative order of
     electron structure function
  ISR_E_SCHEME     = 2          ! Beta-scheme : 0,1,2 , default
     = 2
  !
  PDF_SET          = cteq6AB.LHgrid    ! the CTEQ6 PDFs cteq6m
     ,cteq6d,cteq6l,cteq6l1
                                ! can be accessed when
                                   PDF_GRID_PATH = CTEQ6Grid
                                !
                                ! to use LHAPDF list one of the
                                   PDF sets,
                                ! , e.g. Alekhin_100.LHpdf, that
                                   comes with your
                                ! LHAPDF installation, the full
                                   list can be found in
                                  ! 'lhapdf-config --pdfsets-path
                                     ', in addition set
                                ! PDF_GRID_PATH = PDFsets
                                !
                                  ! or use the interface to
                                     MRST99 through
                                ! PDF_GRID_PATH = MRST99Grid
                                  !
```

```
  PDF_SET_VERSION = 0
  !
  PDF_GRID_PATH    = PDFsets ! CTEQ6Grid,PDFsets,MRST99Grid
}(isr)
(me){
  !------------------------------------------------------------
  !-- ME generators -------------------------------------------
  !------------------------------------------------------------
  ME_SIGNAL_GENERATOR      = Amegic       ! Internal or Amegic
  EVENT_GENERATION_MODE = Unweighted
  !
  COUPLING_SCHEME          = Running_alpha_S ! Fixed   (default
     value) or
                                          ! Running (s for ME
                                             generators, pt for
                                             2->2)
  YUKAWA_MASSES            = Fixed        ! Fixed   (polemass) or
                                          ! Running (higgs mass)
  YUKAWA_MASSES_FACTOR     = 1.           ! Additional prefactor
     for yukawas
  !
  SCALE_SCHEME             = CKKW         ! see 'http://www.
     sherpa-mc.de/scales.html'
  KFACTOR_SCHEME           = 1            ! 1=as(pt2)/as(ecms)^
     nstrong,default=1
  SUDAKOV_WEIGHT           = 1            ! apply sudakov weight
     on single events
  SCALE_FACTOR             = 1.           ! factor the scale is
     multiplied with
}(me)
(integration){
  !------------------------------------------------------------
  !-- Phase space setup ---------------------------------------
  !------------------------------------------------------------
  ERROR =  1.e-2   ! Error by calculating matrix-elements
  !
  FINISH_OPTIMIZATION = On ! Integrate until optimization is
     completed
  !
  INTEGRATOR = 6   ! Phasespace : Rambo=0, Sarge=1,
                   !              Rambo + Sarge=2,
                 !              Rambo + Multichannel=3,
                 !              Multichannel:
                 !              =4 (Original channel
                    generator)
```

```
                  !              =5 (Original+extra channels
                     for competing peaks
                  !                  like in ->Higgs->ZZ->q qb
                      q qb)
                  !              =6 (New channel generator)
                  ! Choice 4-6 only affects channel generation
                     !
  VEGAS = On
}(integration)
(mi){
  !===================================================!
  !          Underlying Event Setup file             !
  !===================================================!
  !
  ! general parameters
  !
  MI_HANDLER     = Amisic                ! Amisic / None
  !
  ! hard underlying event parameters
  !
  CREATE_GRID 93 93 -> 93 93             ! processes to
     generate
  PS_ERROR          = 1.0e-2             ! error for
     integration
  REGULATE_XS       = 0                  ! regulate cross
     section
  XS_REGULATION     = 2.45               ! regulation
     parameter
  SCALE_MIN         = 2.45               ! minimum scale
  RESCALE_EXPONENT    = 0.16             ! rescaling exponent
  REFERENCE_SCALE           = 1800.0            ! reference
     energy scale
  PROFILE_FUNCTION    = Gaussian                ! Gaussian /
     Double_Gaussian
  PROFILE_PARAMETERS  = 1.0 0.5 0.5      ! size (must be 1),
     coresize ,
                                         ! matter fraction
}(mi)
(shower){
  !-------------------------------------------------------------
  !-- Parton showers --------------------------------------------
  !-------------------------------------------------------------
  SHOWER_GENERATOR  = Apacic    !
  !
  FSR_SHOWER        = 1         !  1=On,0=Off
  ISR_SHOWER        = 1         !  1=On,0=Off
```

```
  !
  IS_PT2MIN         = 4.        !  IS Shower cutoff
  FS_PT2MIN         = 1.        !  FS Shower cutoff
}(shower)
(fragmentation){
  !----------------------------------------------------------------------
  !-- Fragmentation parameters ------------------------------------------
  !----------------------------------------------------------------------
  FRAGMENTATION = Ahadic        ! Off, Lund (Pythia string
     fragmentation) or Ahadic
  DECAYMODEL    = Hadrons       ! Lund or Hadrons

  !----------------------------------------------------------------------
  !-- Photons parameters ------------------------------------------------
  !----------------------------------------------------------------------
  YFS_MODE        = 2          ! settings for Photons-module:
                               ! 0        - off/no corrections
                               ! 1        - soft only
                               ! 2        - soft + order(alpha
                                  ) hard correction (default)
  YFS_USE_ME      = 1          ! use exact MEs for hard
     correction if possible
  YFS_IR_CUTOFF   = 1E-3       ! IR cut-off

  MAX_PROPER_LIFETIME = 10
}(fragmentation)
```

C.5 Selector.dat

```
!----------------------------------------------------------------------
!-- Parton level selectors ---------------------------------------------
!----------------------------------------------------------------------
JetFinder   sqr(20/E_CMS) 1.
Mass 11 -11 50 10000
PT 11 8. 1000000.
PT -11 8. 1000000.
Rapidity 11 -3.0 3.0
Rapidity -11 -3.0 3.0
```

Bibliography

[1] Per F. Dahl. “Flash of the Cathode Rays: A History of J.J. Thomson’s Electron.” CRC Press (1997).

[2] S. Weinberg. “A model of leptons”. *Phys. Rev. Lett.*, **19** (1967) 1264–1266.

[3] C. Amsler & al. “Review of Particle Physics”. *Journal of Physics G*, **33** (2006) 1+. `http://pdg.lbl.gov`

[4] S.L. GLASHOW. “Partial – Symmetries of Weak Interactions”. *Selected Papers on Gauge Theory of Weak and Electromagnetic Interactions*, pages 171–180.

[5] DP Barber, U. Becker, H. Benda, A. Boehm, JG Branson, J. Bron, D. Buikman, J. Burger, CC Chang, HS Chen et al. “Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA”. *Physical Review Letters*, **43**, 12 (1979) 830–833.

[6] R. Brandelik, W. Braunschweig, K. Gather, V. Kadansky, K. Lübelsmeyer, P. Mättig, H.U. Martyn, G. Peise, J. Rimkus, HG Sander et al. “Evidence for planar events in e+ e- annihilation at high energies”. *Physics Letters B*, **86**, 2 (1979) 243–249.

[7] FJ Hasert, S. Kabe, W. Krenz, J. Von Krogh, D. Lanske, J. Morfin, K. Schultze, H. Weerts, GH Bertrand-Coremans, J. Sacton et al. “Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment”. *Physics Letters B*, **46**, 1 (1973) 138–140.

[8] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups. “Precision Electroweak Measurements on the Z Resonance”. *Phys. Rept.*, **427** (2006) 257.

[9] PD Acton, G. Alexander, J. Allison, PP Allport, KJ Anderson, S. Arcelli, A. Astbury, D. Axen, G. Azuelos, GA Bahan et al. “A determination ofα s () at LEP using resummed QCD calculations”. *Zeitschrift f*
”ur Physik C Particles and Fields, **59**, 1 (1993) 1–19.

[10] S. Bethke & JE Pilcher. "Tests of perturbative QCD at LEP". *Annual Review of Nuclear and Particle Science*, **42**, 1 (1992) 251–289.

[11] S. Bethke. "Experimental Results on QCD and Jets at LEP and SLC". PITHA-94-29, Talk given at Tennessee International Symposium on Radiative Corrections: Status and Outlook, Gatlinburg, TN, 27 June-1 July 1994.

[12] T. Affolder, H. Akimoto, A. Akopian, MG Albrow, P. Amaral, D. Amidei, K. Anikeev, J. Antos, G. Apollinari, T. Arisawa et al. "Measurement of the Strong Coupling Constant from Inclusive Jet Production at the Tevatron $\bar{p}p$ Collider". *Physical review letters*, **88**, 4 (2002) 42001.

[13] R.D. Ball & S. Forte. "Inclusive measurement of the strong coupling at HERA". *Arxiv preprint hep-ph/9607289*.

[14] M. Acciarri, O. Adriani, M. AguilarBenitez, S. Ahlen, J. Alcaraz, G. Alemanni, J. Allaby, A. Aloisio, G. Alverson, MG Alviggi et al. "Measurements of mass, width and gauge couplings of the W boson at LEP".

[15] F. Abe, MG Albrow, D. Amidei, J. Antos, C. Anway-Wiese, G. Apollinari, H. Areti, M. Atac, P. Auchincloss, F. Azfar et al. "Measurement of the W boson mass". *Physical Review D*, **52**, 9 (1995) 4784–4827.

[16] PW Higgs. "Broken symmetries, massless particles and gauge fields, Phys". *Rev. Lett*, **12** (1964) 132.

[17] GW Bennett, B. Bousquet, HN Brown, G. Bunce, RM Carey, P. Cushman, GT Danby, PT Debevec, M. Deile, H. Deng et al. "Measurement of the Negative Muon Anomalous Magnetic Moment to 0.7 ppm". *Physical Review Letters*, **92**, 16 (2004) 161802.

[18] DN Spergel, L. Verde, HV Peiris, E. Komatsu, MR Nolta, CL Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut et al. "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) 1 Observations: Determination of Cosmological Parameters". *The Astrophysical Journal Supplement Series*, **148**, 1 (2003) 175–194.

[19] S.J. Asztalos, E. Daw, H. Peng, LJ Rosenberg, DB Yu, C. Hagmann, D. Kinion, W. Stoefft, K. van Bibber & J. LaVeigne. "EXPERIMENTAL CONSTRAINTS ON THE AXION DARK MATTER HALO DENSITY". *ASTROPHYSICAL JOURNAL*, **571**, 2 (2002) 27–30.

[20] I.J.R. Aitchison. "Supersymmetry and the MSSM: An Elementary Introduction". *Arxiv preprint hep-ph/0505105*.

[21] G. Aad & al. "ATLAS Collaboration". *The ATLAS Experiment at the CERN Large Hadron Collider, JINST*, **3**.

[22] S.P. Martin. “A Supersymmetry Primer”. *Arxiv preprint hep-ph/9709356*.

[23] A. Abdesselam, T. Akimoto, PP Allport, J. Alonso, B. Anderson, L. Andricek, F. Anghinolfi, RJ Apsimon, G. Barbier, AJ Barr et al. “The barrel modules of the ATLAS semiconductor tracker”. *Nuclear Inst. and Methods in Physics Research, A*, **568**, 2 (2006) 642–671.

[24] A.T.R. Detector. “Technical Design Report”. *CERN/LHCC*, **21** (2001) 2001.

[25] A. Collab. “Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics”. Technical report, CERN-OPEN-2008-020, Geneva, 2008, to appear.

[26] JE Gaiser. “Charmonium spectroscopy from radiative decays of the J/psi and psi”. Ph.D. thesis, Stanford Univ., Stanford, CA (1983).

[27] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand et al. “GEANT4 – a simulation toolkit”. *Nuclear Inst. and Methods in Physics Research, A*, **506**, 3 (2003) 250–303.

[28] J. Allison, K. Amako, J. Apostolakis, H. Araujo, PA Dubois, M. Asai, G. Barrand, R. Capra, S. Chauvie, R. Chytracek et al. “Geant4 developments and applications”. *IEEE Transactions on Nuclear Science*, **53**, 1 Part 2 (2006) 270–278.

[29] G. Altarelli, R.D. Ball, S. Forte & G. Ridolfi. “Determination of the Bjorken sum and strong coupling from polarized structure functions”. *Nuclear Physics-Section B*, **496**, 1 (1997) 337–357.

[30] R.W.L. Jones, M. Ford, G.P. Salam, H. Stenzel & D. Wicke. “High Energy Physics-Phenomenology Title: Theoretical uncertainties on α_s from event-shape variables in e^+e^- annihilations”. *Journal reference: JHEP*, **312** (2003) 007.

[31] J. Campbell. “Higher Order Corrections to Multijet Production in Electron-Positron Annihilation”. Ph.D. thesis, University of Durham, Durham, UK (1998).

[32] M. R. Whalley, D. Bourilkov & R. C. Group. “The Les Houches Accord PDFs (LHAPDF) and Lhaglue”.

[33] J. Pumplin et al. “New generation of parton distributions with uncertainties from global QCD analysis”. *JHEP*, **07** (2002) 012.

[34] J. Pumplin, A. Belyaev, J. Huston, D. Stump & W.K. Tung. “Parton distributions and the strong coupling: CTEQ6AB PDFs”. *JHEP*, **2** (2006) 032.

[35] T. Sjostrand, S. Mrenna & P. Skands. “Pythia 6.4–Physics and Manual,[hep-ph/0603175]”. Technical report, FERMILAB-PUB-06-052-CD-T],[LU TP 06-13], JHEP 05, 026, doi: 10.1088/1126-6708/2006/05/026 (2006).

[36] M.L. Mangano, F. Piccinini, A.D. Polosa, M. Moretti & R. Pittau. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". *Journal of High Energy Physics*, **2003**, 7 (2003) 1–1.

[37] S. Frixione & B.R. Webber. "Matching NLO QCD computations and parton shower simulations". *Journal of High Energy Physics*, **2002**, 6 (2002) 2–2.

[38] T. Gleisberg et al. "Event generation with SHERPA 1.1". *JHEP*, **02** (2009) 007.

[39] W.J. Stirling. "Progress in Parton Distribution Functions and Implications for LHC". In "Presented at the XXXVIII International Symposium on Multiparticle Dynamics (ISMD 2008)", volume 15, page 20 (2008).

[40] G.C. Blazey, J.R. Dittmann, S.D. Ellis, V.D. Elvira, K. Frame, S. Grinstein, R. Hirosky, R. Piegaia, H. Schellman, R. Snihur et al. "Run II jet physics". *Arxiv preprint hep-ex/0005012*.

[41] Stephen D. Ellis & Davison E. Soper. "Successive combination jet algorithm for hadron collisions". *Phys. Rev.*, **D48** (1993) 3160–3166.

[42] GP Salam & G. Soyez. "A practical seedless infrared-safe cone jet algorithm, JHEP 05 (2007) 086". *arXiv*, **704**.

[43] I. Abt, T. Ahmed, S. Aid, V. Andreev, B. Andrieu, R.D. Appuhn, C. Arnault, M. Arpagaus, A. Babaev, E. Banas et al. "The tracking calorimeter and muon detectors of the H1 experiment at Hera". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **386** (1997) 348–396.

[44] Daniel Stump et al. "Inclusive jet production, parton distributions, and the search for new physics". *JHEP*, **10** (2003) 046.

[45] F. Heinemann. "Robust Track Based Alignment of the ATLAS Silicon Detectors and Assessing Parton Distribution Uncertainties in Drell-Yan Processes". Ph.D. thesis, Oxford Univ. Oxford (2007).

[46] A. Tricoli, AM Cooper-Sarkar & C. Gwenlan. "Uncertainties on W and Z production at the LHC". *Arxiv preprint hep-ex/0509002*.