$Z + \gamma$ Differential Cross Section Measurements and the Digital Timing Calibration of the Level-1 Calorimeter Trigger Cluster Processor System in ATLAS

Joseph Neil Lilley

Thesis submitted for the degree of
Doctor of Philosophy

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

November 2010

Abstract

This thesis investigates the reconstruction of $Z(\rightarrow ee)\gamma$ events with the ATLAS detector at the LHC. The capabilities of the detector are explored for the initial run scenario with a proton-proton centre of mass collision energy of $\sqrt{s} = 7\,\text{TeV}$, and an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$. Monte Carlo simulations are used to predict the expected precision of a differential cross-section measurement for initial state radiation $Z + \gamma$ events, both with respect to the transverse momentum of the photon, $p_T(\gamma)$, and the three body $ee\gamma$ invariant mass. A bin-by-bin correction is used to account for the signal selection efficiency and purity, and to correct the measured (simulated) distribution back to the theoretical prediction. The main backgrounds are found to be from the final state radiation $Z + \gamma$ process, and from jets faking photons in $Z \rightarrow ee$ events. The possible QCD multijet background is studied using a fake-rate method, and found to be negligible for the particle identification cuts used in the analysis. The main systematic uncertainties on the differential cross-section measurements are explored with Monte Carlo simulations, and found to be of a similar scale to the statistical errors for the chosen distribution binning.

The three body $ee\gamma$ invariant mass distribution was then used as the basis of an exclusion study on new particles decaying to the $Z(ee)\gamma$ final state. Under the assumption that the measured data agrees with the Standard Model prediction, exclusion limits were placed at 95% confidence level on the cross-section times branching ratio for a new scalar (modelled by SM Higgs process), or vector (based on a low-scale technicolor process) particle hypothesis, for particles in the mass range 200 to 900 GeV. Limits of the order $\mathcal{O}(0.01)$ - $\mathcal{O}(0.1)\,\text{pb}$ on the cross section times branching ratios are predicted, which would improve on the equivalent limits previously calculated by the DØ experiment at the Tevatron collider, albeit in a different $\sqrt{s}$ region, where cross-sections will generally be higher for new massive particles.

In addition to the $Z\gamma$ measurements, a digital timing calibration procedure was developed for the Cluster Processor (CP) subsystem of the level-1 calorimeter trigger. This work was essential to providing a repeatable and robust mechanism for timing in the digital processing in the CP system, a necessary ingredient for a robust and reliable trigger system; a pre-requisite of any physics analysis. This calibration procedure is described here.

Author's Contribution

ATLAS's membership comprises approximately 3000 physicists, from more than 157 universities and laboratories, spanning some 38 countries worldwide; this is a testament to its triumph as a collaborative enterprise. It is only through this collaborative aspect that the work presented in this thesis has been possible. Unless otherwise stated, or referenced, the work in this thesis is my own.

With respect to the work carried out on the level-1 calorimeter trigger Cluster Processor calibration, as pointed out in chapter 3, the software which I created fits into a large pre-existing framework. The calibration scans described in chapter 3 are ostensibly my own work, but would not have been possible without the expert support and guidance from the L1Calo team.

The analysis of $Z\gamma$ events relied on the use of simulated data which in the large part was produced by the collaboration, and distributed over the LHC computing grid; and for the new particle study where I simulated the events myself, I am thankful to the Monte Carlo generator, and ATLAS simulation authors. The analysis code used to analyse the data samples, and produce the results presented in chapter 5, was authored by myself, but was however dependent on the use of the ATLAS simulation framework, and the ROOT framework for presenting the results.

Acknowledgements

It is with pleasure that I take the opportunity to acknowledge the people and efforts which have made this work possible. Firstly I would like to thank Dave Charlton, for his expert guidance, and patience, throughout the duration of my research, and more acutely in the closing stages. In getting to grips with L1Calo in the first instance, Gilles Mahout was indispensable, and later with my time at CERN; Murrough, Steve, Bruce, and the rest of L1Calo were a pleasure to work with. I would also like to thank the Science and Technology Facilities Council for funding my research, and allowing me to spend a fruitful period living and working at CERN.

My fellow research students and members of the particle physics group, in particular the flow of cohabitants through west 316 in Birmingham, and R1 at CERN, you have provided innumerable responses to my least intelligent questions and remarks, for which some mention is deserved (if not something stronger). Specifically; Chris, Mark, Dave, and Tim have been my coding heroes.

I would like to collectively thank all of my friends outside of physics, many of whom still do not believe that I am in fact a physicist, I hope this is evidence enough. You have all had to put up with a lot of complaining from me about physics, computers, particles, and other such things which do not really concern you, but through which you afford me the support and encouragement that has kept me going. I would particularly like to thank Gabriel, for support above and beyond the call of duty.

Last but by no means least, I would like to thank Colette and George (a.k.a mum and dad), a constant stream of encouragement, if only so that one day I might finish, and get a real job. Finally, I would like to thank my brother Kevin, who can not be here to see me finish, but who was, and will continue to be, my inspiration.

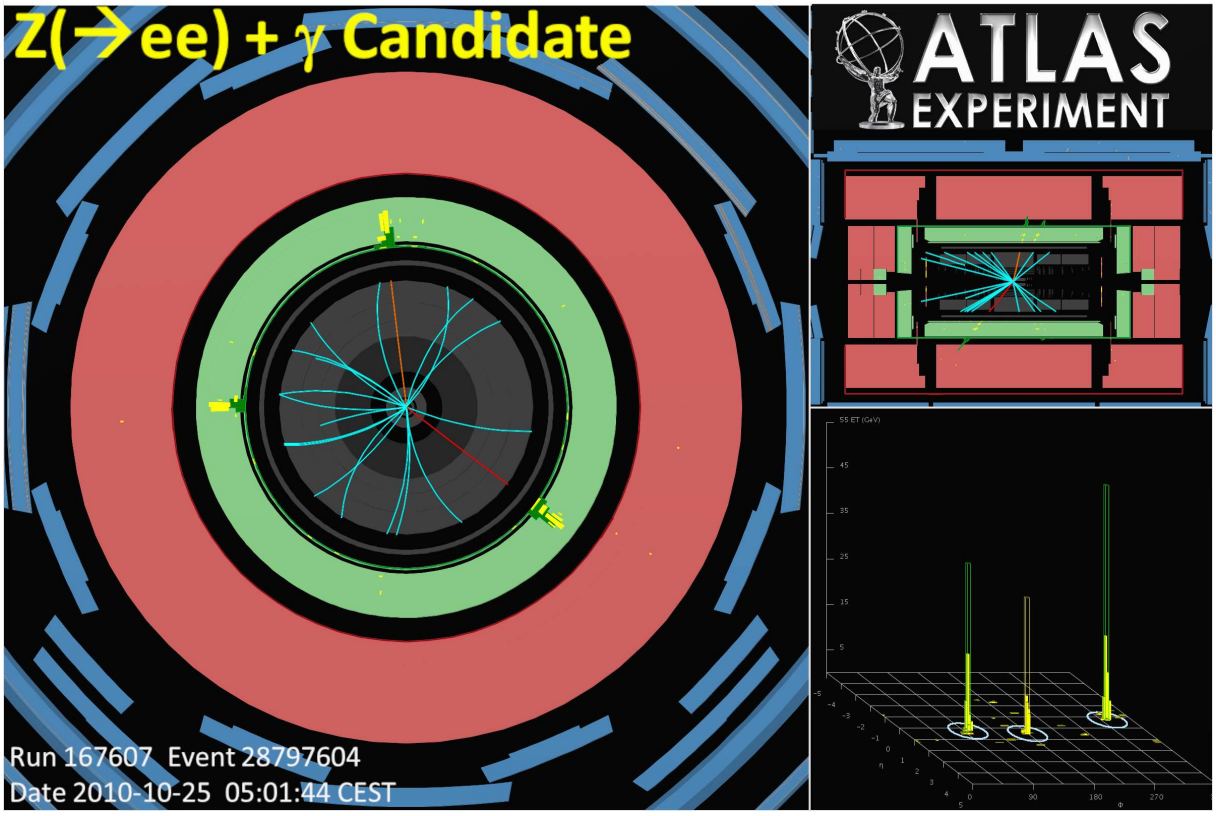


Figure 1: $Z(\rightarrow ee)\gamma$ candidate event display picture from ATLAS 2010 7 TeV physics run.

Contents

List of Figures

List of Tables

Chapter 1

Theory

This chapter begins with an introduction to the Standard Model (SM), the accepted framework used to describe the particle content and interactions of matter and forces (excluding gravity), at the most fundamental level. In order to place the measurements presented in this thesis into context, a summarised description of various aspects of SM theory are presented, including; the SM as a gauge theory, electroweak physics, how predictions of particle interactions are related to the underlying theory (Lagrangian), how calculations are performed by the experimentalist (Monte Carlo generators), and some properties of $Z\gamma$ events in the context of ATLAS. Natural units, $\hbar = c = 1$, are adopted throughout.

1.1 Introducing the Standard Model

The fundamental particles of the Standard Model fall into two categories; the fermions (intrinsic spin 1/2, obeying Fermi-Dirac statistics) which are the quarks and leptons which make up matter, and the bosons (integer intrinsic spin, obeying Bose-Einstein statistics) which mediate the electromagnetic, weak nuclear, and strong forces. In the Standard Model these particles are fundamental and thus have no sub-structure. The fermions are

divided into groups depending on the forces with which they interact; quarks which are subject to the strong force, and leptons which are not. The fermions appear to be arranged naturally into three generations, of which only the first is present in ordinary matter; e.g. protons are essentially made of u and d quarks (see table 1.1). The three generations are copies of each other differing only in mass; the reason for which is an important open question in particle physics. The Standard Model is often discussed in sectors relating to the different forces; the electromagnetic (EM) sector governed by the quantum field theory of quantum electrodynamics (QED); the weak sector which is described in a unified form with electromagnetism by electroweak theory (EW); and the strongly interacting sector which is described by quantum chromodynamics (QCD).

Particle	Generation			Electric charge	Interaction
	1^{st}	2^{nd}	3^{rd}		
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	$+\frac{2}{3}$ $-\frac{1}{3}$	EM, Weak, Strong
Leptons	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}$	0 -1	Weak EM, Weak

Table 1.1: Fermions in the Standard Model.

Table 1.2 lists the vector (spin-1) bosons in the Standard Model, note the absence of a description of gravity here. Despite this being a desirable feature of a complete theory, it does not affect the usefulness of the Standard Model at describing particle interactions, since at the energy scales of LHC interactions the strength of gravity is negligible. On a basic level, the influence the forces have on the fermions is controlled by the exchange of, and coupling to, the various bosons. The boson describing the exchange force for electromagnetism is the familiar massless photon γ; this couples to any fermion carrying

electric charge. The weak force differs somewhat in that it is mediated by massive bosons, the charged $W^{\pm}$ and neutral Z, which have masses $M_W = 80.4\,\text{GeV}$, and $M_Z = 91.2\,\text{GeV}$ [1]. These couple to all fermions with a similar coupling strength to that for photons, however the rate of weak processes are greatly suppressed relative to electromagnetic processes because the calculations involving the exchange of weak bosons pick up a term inversely proportional to powers of the mass of the weak boson. The mechanism by which the weak bosons gain mass is the Higgs mechanism, which is presented in section 1.4.1.

For the exchange of the strong force there are 8 massless gluons, g; these couple to any particle carrying "colour charge", which in addition to the quarks includes the gluons themselves. The characteristics of this self-coupling aspect give rise to the differing variation of the strength of the strong force with distance, compared to the electromagnetic and weak forces (see section 1.5.1).

Force	Boson(s)		Electric charge
EM	photon	γ	0
Weak	Z boson	Z^0	0
	W boson	W^+	$+1$
		W^-	-1
Strong	gluon (8 types)	g	0

Table 1.2: Vector bosons in the Standard Model.

1.2 Standard Model as a Gauge Field Theory

The modern way to describe nature at the particle level is to define the Lagrangian[1], $\mathcal{L}$. The Lagrangian contains terms corresponding to the quantum fields describing the

[1]Following common practise, what is referred to here as simply the Lagrangian, $\mathcal{L}$, is in fact the Lagrangian *density* , where the actual Lagrangian $L = \int \mathcal{L} d^3x$.

fermions and bosons, and dictates the dynamic evolution of the system. From the Lagrangian one can in principle (but not in practise) use quantum theory to calculate the behaviour of all systems, or composite objects (nuclei, atoms, e.t.c.) built from the particles in its specification. It is analogous to the Lagrangian in classical field theory, where by solving the Euler-Lagrange equation (equation 1.1), one derives the equations of motion of the system.

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) = 0 \tag{1.1}$$

The Standard Model Lagrangian is invariant under certain symmetry transformations (transformations which leave the equations of motion invariant). In addition to Lorentz invariance, the other type of symmetry of the Lagrangian is a local *gauge* symmetry; the power of this concept is introduced using the example of QED in section 1.2.1. There is a powerful relationship between physical symmetries and conserved quantities, this is embodied in Noether's theorem which states that for each physical symmetry of a system there is a corresponding conserved current (physical quantity) [2]. For example in QED, the conserved current due to the gauge symmetry is electric charge, which means that this quantum number must be conserved at any particle interaction vertex.

1.2.1 QED as a Gauge Theory

Quantum electrodynamics (QED) can be thought of as a prototype for the Standard Model, and is the simplest example of the power of gauge symmetries. The Lagrangian describing a freely propagating fermion of mass m is given by equation 1.2.

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \tag{1.2}$$

Here ψ are spinors representing the fermion fields, with the adjoint spinor $\bar{\psi} = \psi^\dagger \gamma^0$, and γ^μ are the Dirac gamma matrices[2] (codifying the spin information). A local gauge transformation is $\psi \longrightarrow \psi' = e^{iq\Lambda(x)}\psi(x)$, where $\Lambda(x)$ is a real valued function of posi-

tion. This transformation belongs to the unitary Abelian group of symmetry transformations $U(1)$; the term Abelian means that the group transformations commute (e.g. $e^{iq\Lambda_1(x)}e^{iq\Lambda_2(x)} = e^{iq\Lambda_2(x)}e^{iq\Lambda_1(x)}$).

The introduction of this gauge transformation breaks the invariance of the original Lagrangian, introducing an extra term (since $ie^{-iq\Lambda(x)}\bar{\psi}\gamma^\mu\partial_\mu(e^{iq\Lambda(x)}\psi) \neq i\bar{\psi}\gamma^\mu\partial_\mu\psi$). A new field A^μ, may be introduced to restore the invariance of the Lagrangian, by considering the *interaction* of the fermion field with this gauge field A^μ, via a term $-e\bar{\psi}\gamma^\mu A_\mu\psi$ (where e is the charge of the electron). The derivative ∂_μ is redefined as the covariant derivative $\mathcal{D}_\mu$ thus:

$$\mathcal{D}_\mu = \partial_\mu + ieA_\mu(x) \tag{1.3}$$

which must transform under gauge transformation in the same way as the field itself, such that:

$$\mathcal{D}_\mu\psi \longrightarrow \mathcal{D}'_\mu\psi' = e^{ie\Lambda(x)}\mathcal{D}_\mu\psi \tag{1.4}$$

with,

$$A_\mu \longrightarrow A'_\mu = \partial_\mu - \frac{1}{e}\partial_\mu\Lambda(x). \tag{1.5}$$

The Lagrangian (implicitly including the interaction term with the gauge field) can now be written in the covariant form:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi, \tag{1.6}$$

so by imposing a local $U(1)$ gauge invariance on the free Lagrangian, the interaction term associated with the gauge field A^μ is recovered, this is identified as the physical photon field. The full QED Lagrangian must include a kinetic[2] term, to describe the propogation of the photon field [3]. This is achieved by defining the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which is invariant under local gauge transformation. If this factor enters like $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, it is both Lorentz invariant (indices are summed over) and the Lagrange equations of motion exactly match the relativistic form of Maxwell's equations

[2]The Lagrangian $L = T - V$, where T is the kinetic, and V the potential term.

[4]. The full QED Lagrangian can be written thus:

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma^{\mu}\mathcal{D}_{\mu}\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.7}$$

The symmetry principle also tells us that the photon must be massless, since any simple mass term in the Lagrangian (quadratic in A_{μ}) would break the local gauge invariance.

1.3 From Lagrangian to Measurement

To unlock the predictive power of the Standard Model one must be able, starting with the Lagrangian, to predict rates and distributions of particle interactions which can be measured by experiment; experiment is the arbiter when it comes to the validation of a theory. The accepted way to visualise fundamental particle interactions is via the Feynman diagram, as in figure 1.1. Not only does this act as a useful visual aid, it is moreover a calculational tool, where the lines and vertices have corresponding mathematical terms (Feynman rules) which are used to construct the probability amplitude calculations (matrix elements) for the process which the diagram represents.

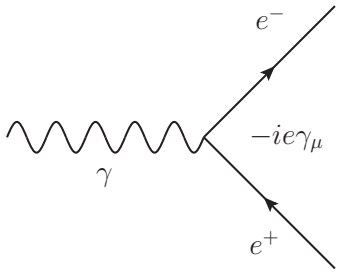


Figure 1.1: Feynman diagram for a photon coulping to e^+e^- pair, Feynman rule vertex factor is shown, $-ie\gamma_{\mu}$.

The calculation of probability amplitudes (and following this, cross-sections) which employ Feynman's rules are grounded in perturbation theory. The perturbation series is an expansion in powers of the coupling parameter (in QED this is e), so in order to be valid, the value of the coupling must be sufficiently small for the perturbation series

to rapidly converge. This consideration has important ramifications for LHC physics, where due to the strength of the strong coupling for quarks bound in hadrons (where $\alpha_s(Q^2) > 1$), non-perturbative treatments must also be employed (see section 1.5.2).

1.4 Electroweak Physics

In the 1960's Glashow[5], Weinberg[6] and Salam[7] unified the electromagnetic and weak interactions into the *electroweak* gauge theory, which exhibits symmetry under transformation of the local gauge group $SU(2)_L \times U(1)_Y$. The $SU(2)_L$ transformation is a rotation in *weak isospin* space, where left-handed fermions are arranged into weak isospin doublets, and the right-handed fermions as singlets under weak hypercharge Y (these are shown in table 1.3). The weak hypercharge is twice the difference between the electrical charge Q and weak isospin (z-projection) T_Z, i.e. $Q = T_Z + \frac{Y}{2}$. Note that the right-handed fermions do not possess weak isospin, and accordingly do not interact with the charged weak current ($W^{\pm}$); consequently right-handed neutrinos are not observable in nature. Table 1.3 shows the electroweak singlets and doublets.

Left-handed doublets								
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$	$\begin{pmatrix} u \\ d' \end{pmatrix}$	$\begin{pmatrix} c \\ s' \end{pmatrix}$	$\begin{pmatrix} t \\ b' \end{pmatrix}$			
Right-handed singlets								
$(e)_R$	$(\mu)_R$	$(\tau)_R$	$(u)_R$	$(d)_R$	$(c)_R$	$(s)_R$	$(t)_R$	$(b)_R$

Table 1.3: Electro-weak doublets and singlets.

A full exploration of the electroweak Lagrangian is beyond the scope of this thesis (see for example [8]), but some important consequences are summarised here. In order to satisfy the invariance of the Lagrangian under a local $SU(2)_L \times U(1)_Y$ gauge symmetry,

a triplet of scalar fields W_i^μ, $i = 1, 2, 3$ coupling to weak isospin, and a singlet coupling to weak hypercharge B^μ are required. These fields are related to the physical $W^\pm, Z$ and A_μ (EM) fields accordingly:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \pm i W_\mu^2 \right), \tag{1.8}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \tag{1.9}$$

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W. \tag{1.10}$$

Here, θ_W is the weak mixing angle describing the mixing of the fields, and relating the strength of the weak coupling g, to the electromagnetic coupling e, $e = g \sin\theta_W$. As a consequence of the non-abelian nature of the $SU(2)_L$ transformations, interaction terms between the W_i fields enter in order to conserve invariance. These result in both triple gauge coupling and quartic gauge coupling vertices, which represent interaction terms between the bosons. The part of the electroweak Lagrangian describing the TGC's is as follows[8]:

$$\begin{aligned}\mathcal{L}_{TGC} = &- ie \cot\theta_W \{ (\partial^\mu W^\nu - \partial^\nu W^\mu) W_\mu^\dagger Z_\nu - (\partial^\mu W^{\nu\dagger} - \partial^\nu W^{\mu\dagger}) W_\mu Z_\nu + W_\mu W_\nu^\dagger (\partial^\mu Z^\nu - \partial^\nu Z^\mu) \} \\ &- ie \{ (\partial^\mu W^\nu - \partial^\nu W^\mu) W_\mu^\dagger A_\nu - (\partial^\mu W^{\nu\dagger} - \partial^\nu W^{\mu\dagger}) W_\mu A_\nu + W_\mu W_\nu^\dagger (\partial^\mu A^\nu - \partial^\nu A^\mu) \}.\end{aligned} \tag{1.11}$$

From this form of the Lagrangian, it follows that the only couplings are between WWZ, and $WW\gamma$ fields; there is no tree level coupling in the SM involving only Z bosons and photons.

1.4.1 The Higgs Mechanism

The introduction of an explicit mass term for the bosons in the electroweak Lagrangian would violate the $SU(2)_L \times U(1)_Y$ local gauge invariance. This problem is solved in an elegant manner by the Higgs mechanism. This involves the introduction of new spin-0

scalar fields called the Higgs fields, a complex weak isospin doublet. One component of these fields has a non-zero vacuum expectation value, meaning that the overall Lagrangian is invariant but the physical ground state is not. A choice of minimum in the potential is chosen such that one component of the Higgs doublet is non-zero. The electroweak Lagrangian is redefined in terms of this non-zero Higgs ground state; this is referred to as the *spontaneous symmetry breaking*. New terms are generated which correspond to mass terms for the W and Z bosons, leaving a massless field corresponding to the photon, and predicting a new massive neutral gauge boson, the Higgs boson. This prescription leaves the underlying gauge symmetry which generated the electroweak bosons intact. The theory does not directly predict a mass for the Higgs boson, but Standard Model constraints (to protect the unitarity of vector boson scattering amplitudes[9]) suggests a mass below about 1 TeV; as yet the Higgs boson has not been detected by experiment, but constraints on the possible mass have been placed at previous experiments[10]. The discovery of the Higgs boson is one of the primary motivations for ATLAS and the LHC.

1.5 Modelling Proton-Proton Interactions

At the LHC, protons collide together at an unprecedented energy, allowing previously inaccessible regions of the Standard Model, and beyond, to be explored. QCD is at the heart of these collisions, responsible for describing the underlying structure of quarks and gluons within the protons, and controlling the dynamics of the hard scattering of partons at the point of collision. The experimental data recorded by ATLAS must be comparable to the underlying theory; in order to understand this procedure, some important aspects of QCD must be considered.

1.5.1 Aspects of QCD

QCD is a gauge field theory, which is invariant under $SU(3)$ group transformations in colour charge space. The colour charge is labelled in units of Red, Blue and Green, with their associated anti-colour partners. Free objects must be colour neutral (a colour singlet state); these are the baryons made from a colour neutral triplet of quarks (e.g. RGB, or $\bar{R}\bar{G}\bar{B}$), and the mesons formed by colour-anti-colour combinations (e.g. $R\bar{R}$,$G\bar{G}$,$B\bar{B}$). As previously mentioned, there are 8 gluons which are themselves coloured (carrying 2 units of colour charge), giving rise to both cubic and quartic interaction vertices between the gluons.

The phenomenological result of these self-coupling terms is manifest in the nature of the running of the strong coupling α_S, as a function of Q^2 (the 4-momentum transfer of a particle exchange). The result is an anti-screening effect[3], where the strong force increases with increasing separation. As a result, free quarks are never seen; as quarks move apart it becomes energetically favourable to create new quark anti-quark pairs from the vacuum. These combine with, and follow the original quark to form a *jet* of hadrons, in a process called *hadronisation*. It is the hadronic jets which are detectable in ATLAS, not the original partons. Considering the running of the coupling in the other direction, is the concept of *asymptotic freedom*; which expresses the fact that only at high Q^2 (short distances) can the partons be considered free (unbound).

1.5.2 Parton Distribution Functions

When two protons collide, the underlying scattering is that of two partons (quarks or gluons), which carry some unknown reduced fraction of the proton's momentum. A full calculation of a such an interaction would require knowledge of the individual parton's initial state wave functions in a region where perturbation theory can not be applied (large α_S). To address this problem, inspiration comes from the parton model of the

[3]The opposite of the charge screening seen when probing the charge of electrons.

proton, which was originally developed in the late 1960's to explain observations from deep inelastic scattering experiments [11]. Here structure functions describe the internal structure of the proton; these are dependent on the momentum transfer of the probe, Q^2, and x, the fraction of the proton's momentum carried by the struck quark.

Today, Parton Distribution Functions (PDFs) describe the probability densities for finding a parton of a given species, f, carrying a longitudinal momentum fraction (relative to the proton's direction), x, for a momentum transfer, Q^2. These distributions are extracted from fits to an ensemble of data from hard scattering experiments, see for example [12]. The full process calculation is then separated into a part representing the hard scatter (where perturbation theory can be applied), which is convolved with parton distribution functions describing the initial state; and fragmentation functions, describing the fragmentation of quarks into hadrons (if predicting hadron distributions).

1.5.3 Monte Carlo Generators

Calculations based in quantum mechanics are inherently probabilistic; this is the reason that experiments collect large statistical sets of data. Calculations exploit the probabilistic nature of the theory by using random numbers to generate a large ensemble of unique predictions (events), which taken collectively reproduce the expected statistical distributions of events. In practise, calculations are performed by computer programs generically know as Monte Carlo event generators [13]. Many different generators are available, incorporating a diverse range of theoretical models.

The main generator used in this study (Pythia[14] for the signal prediction) simulates all aspects of the physical process; both perturbative (hard scatter matrix element calculation) and non-perturbative (PDF, fragmentation, hadronisation), as shown schematically in figure 1.2. The output of the generator for each simulated event is a list of 4-vectors representing the final state particles. The distribution of final state particles from a large ensemble of events represents the theoretical prediction (governed by the underlying La-

grangian of the Standard Model), which is compared to the experimentally derived data distributions.

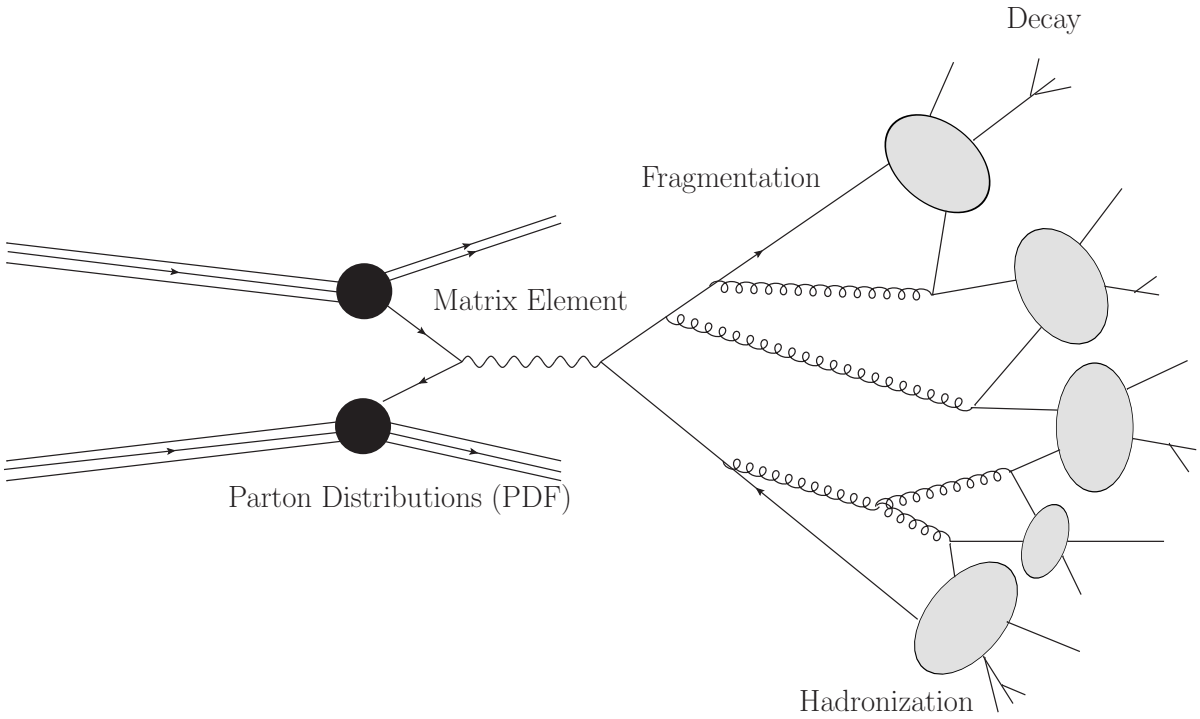
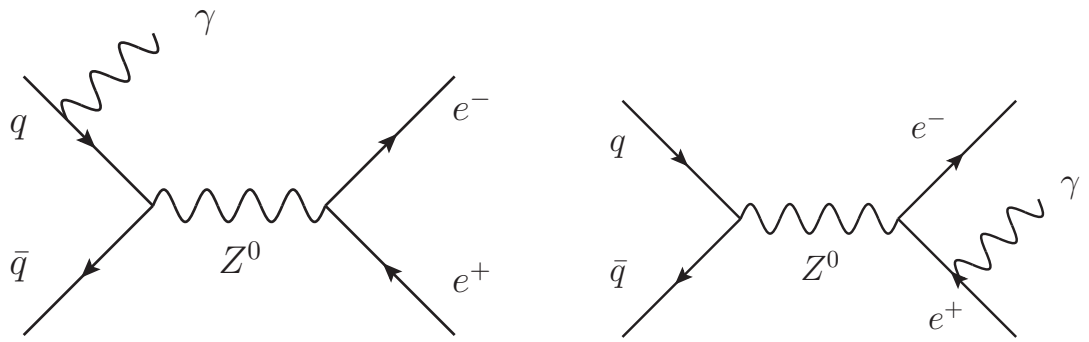


Figure 1.2: Schematic representation of the processes simulated in a Monte Carlo generator such as Pythia.

1.6 $Z\gamma$ Production at the LHC

As mentioned in section 1.4, there is no vertex coupling of the form $VZ\gamma$, where $V = Z, \gamma$. As a result, measuring the $Z\gamma$ cross-section is a direct probe of the non-abelian structure of the electroweak Lagrangian. According to the Standard Model, $Z\gamma$ (where $Z \to e^+e^-$) states are produced at lowest order via the diagrams represented in figure 1.3. These are referred to as Initial State Radiation (ISR), where the photon is emitted from an initial quark leg; and Final State Radiation (FSR), where the photon is radiated by one of the leptons from the Z decay.

A full amplitude for the $e^+e^-\gamma$ final state includes the interference term between the ISR and FSR diagrams. This interference is not simulated in Pythia, but a parton level comparison between Pythia (no interference term) and `Sherpa` [15] (interference term



(a) ISR (b) FSR

Figure 1.3: (a) LO Feynman diagram for ISR signal process. (b) LO Feynman diagram for FSR process.

included) showed this to be a negligible effect given the constraints on the phase space applied in the final signal event selection (the event selection is described in section 5.4). The Feynman diagrams in figures 1.3(a) and 1.3(b) represent the lowest, or leading order (LO) processes. For the purposes of this analysis, the ISR diagram is considered to be the signal process; this is because the kinematical properties of potential new physics events (see section 1.6.2) will be more ISR-like than FSR-like.

1.6.1 Higher Order Corrections to $Z\gamma$ Processes

At the LHC the effect of higher order QCD corrections in the $Z\gamma$ channel are predicted to be significant, and increasing with the transverse momentum of the photon[16]. At the time of this study, the only suitable Monte Carlo signal data sample was produced at LO only in Pythia. Pythia provides a reasonable approximation to the effect of higher order corrections with a non-perturbative fragmentation approach[14]. Given that the focus of this study is on the capabilities of the ATLAS detector, and that the final cross-sections will be normalised by fits to real data, this approach is satisfactory. If one is to claim a discovery of a deviation from a Standard Model prediction, a higher order calculation should be used, such as that provided by the `Sherpa`, `AYLEN`[16], or `BHO`[17] generators. These generators differ in a number of ways from Pythia, for example they calculate the interference term between ISR and FSR diagrams, as well as allowing for the specification

of anomalous triple gauge couplings (discussed in section 1.6.2).

1.6.2 A Possible Window on New Physics

As previously mentioned, the cross-section of the $Z\gamma$ process is precisely predicted by the Standard Model[18]. However the Standard Model in its current form is most likely not the final theory, but an effective theory working at the scale probed with current experiments. There are at present a number of extensions to the Standard Model which predict a contribution to the $Z\gamma$ final state, including new particles entering through higher order effects (SUSY, charged Higgs), low-scale technicolor theory (LSTC)[19], gravitons[20], and the SM Higgs.

The effect of some generic (unknown) new physics can be studied by introducing an anomalous triple gauge coupling term into an effective Lagrangian; such a coupling term would produce a deviation in the cross-section from the expected Standard Model value[18]. Limits have been placed on the strengths of these couplings by previous experiments [21], and as yet no deviations from the Standard Model have been observed. With sufficient integrated luminosity, ATLAS is expected to further constrain the limits on these couplings[22]. Given that the focus of this study is on what is achievable in the early running period of ATLAS[4], this approach is replaced in favour of a simpler benchmark search for a new particle decay to the $ee\gamma$ final state. The capability of ATLAS in placing exclusion limits on such generic new physics, by its (scalar or vector) particle decay to the $Z\gamma$ di-boson state is presented in section 5.8.

1.6.3 Previous $Z\gamma$ Measurements

The Z boson was discovered in 1983, at the UA1[23] and UA2[24] experiments at the Super Proton Synchrotron (SPS) at CERN. At the same time the first radiative photons

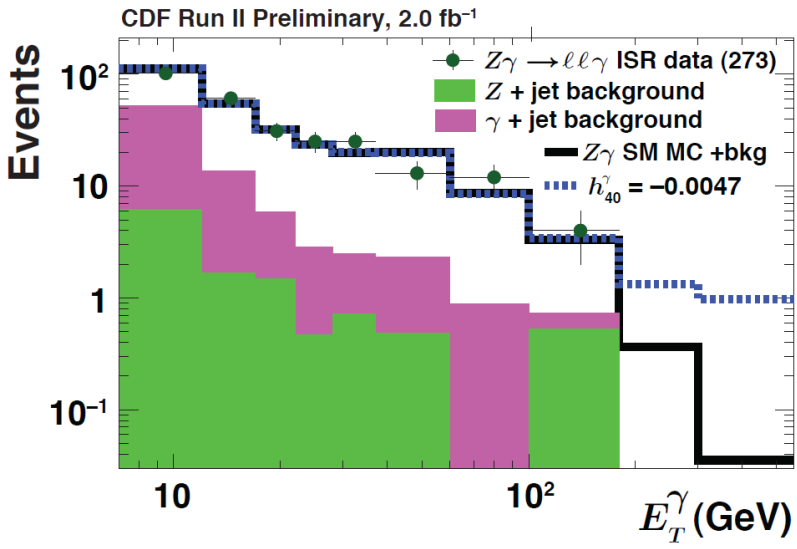
[4] An integrated luminosity of $1\,\text{fb}^{-1}$ at a proton-proton centre of mass energy $\sqrt{s} = 7\,\text{TeV}$.

in Z events were identified.

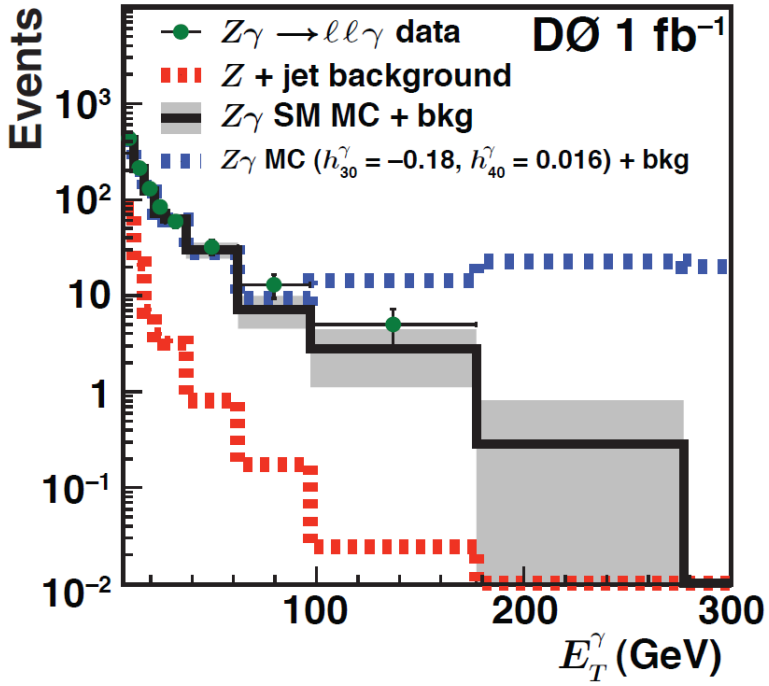
Later, at the large electron positron (LEP) collider, an e^+e^- collider operating at up to $\sqrt{s} = 209$ GeV, between 1989 and 2000 at CERN, radiative Z events were routinely recorded; both in Z decays during 1989-1995, when $\sqrt{s} \approx m_Z$, and at LEP-2 during 1996-2000, when $\sqrt{s}$ was between $161 - 209\,\text{GeV}$. In the later period (LEP2), these events were used to calculate the centre of mass collision energy[25], and also to place exclusion limits on the size of anomalous triple gauge couplings between Z bosons and photons[26][27].

The most recent $Z\gamma$ investigations have been taking place at the DØ and CDF experiments at the Tevatron proton-anti-proton collider, with $\sqrt{s} = 1.96\,\text{TeV}$ centre of mass collision energy. The results from these studies are most directly comparable to that expected with ATLAS. The combined results from these experiments, for an integrated luminosity of between $1 - 2\,\text{fb}^{-1}$, can be found for example in [28]. Figure 1.4 shows the measured photon transverse energy distributions, from both CDF (figure 1.4(a)), and DØ (figure 1.4(b)); where a handful of events with $E_T(\gamma) > 100\,\text{GeV}$ have been observed.

Limits on the cross-section times branching ratios of possible new vector and scalar particles decaying to the $Z\gamma$ final state have been placed at the DØ experiment [29]; these results are shown in figure 1.5, and act as a point of reference for the corresponding results simulated with ATLAS in section 5.8.4

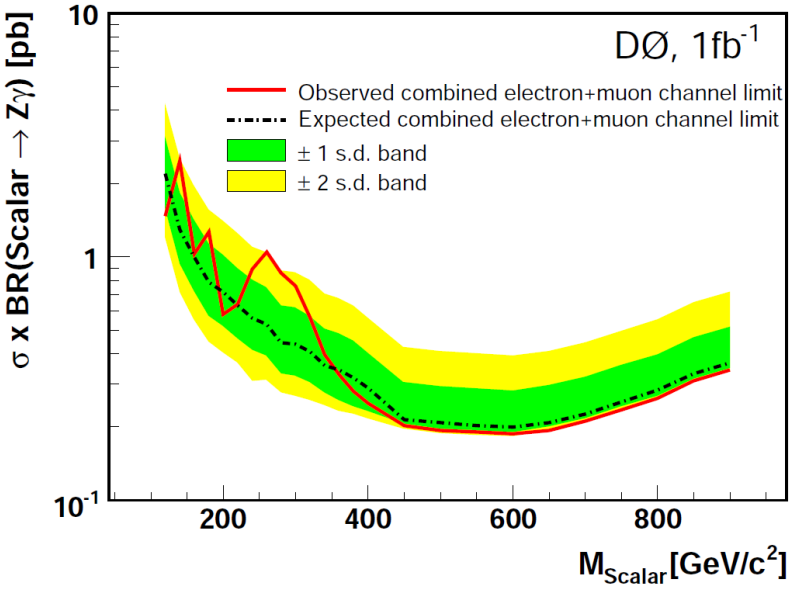


(a)

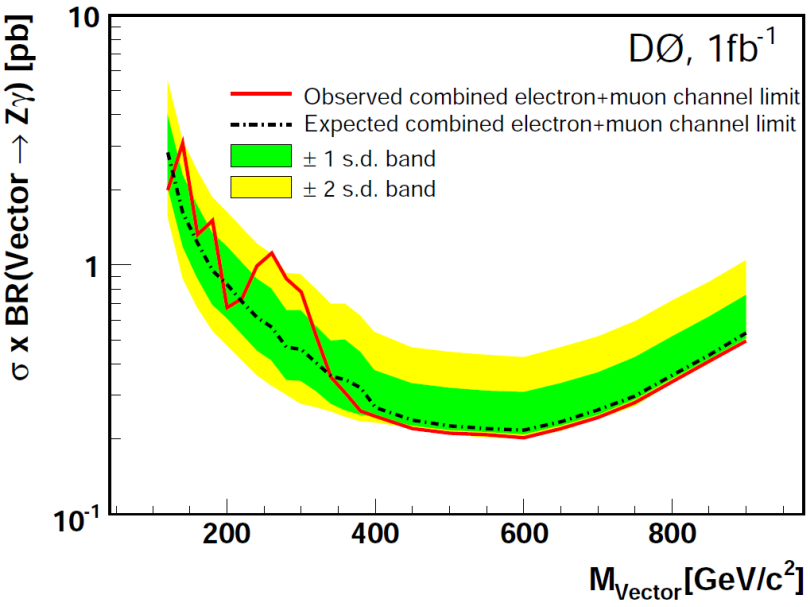


(b)

Figure 1.4: (a) Photon E_T spectra in $Z\gamma$ production for the CDF analysis. (b) Photon E_T spectra in $Z\gamma$ production for the DØ analysis. Both plots taken from [28]. The dashed blue line represents the scenario of a possible anomalous $Z\gamma$ coupling term.



(a)



(b)

Figure 1.5: (a) DØ observed and expected 95% C.L. limits for a scalar particle decaying into $Z\gamma$ as a function of the particle mass. The bands represent the 1 s.d. (dark) and 2 s.d. (light) uncertainties on the expected limit.(b) Equivalent plot for a vector particle. Plots taken from [29].

Chapter 2

The ATLAS Detector

ATLAS is one of four particle detectors situated around the 27 km circumference Large Hadron Collider at CERN, near Geneva (Switzerland). ATLAS has a varied and ambitious physics program, this is reflected in the "general purpose" design of the detector. ATLAS must be capable of discovering whichever previously unseen physics is manifest at the TeV energy scale; whether it be a Standard Model-like Higgs boson decay (this is used as a benchmark for the detector subsystems, given the many decay modes, and dependence on detector resolution), cascade decays of super-symmetric particles, heavy resonances, or more exotic theories involving extra-dimensions, black-holes, and gravitons. The common element of all these signatures is their eventual decay into Standard Model particles which interact with the detector matter in a well understood way (except for those particles which interact very weakly with detector materials, such as neutrinos, or as yet undiscovered particles like the Lightest Stable Particle (LSP) of R-parity conserving super-symmetry; these will leave a detectable missing energy signature). In the following chapter, a summary of the LHC accelerator and an overview of the ATLAS detector and its various subsystems are presented.

2.1 The Large Hadron Collider

The Large Hadron Collider is the world's largest, and highest energy, particle accelerator, and is designed to accelerate bunches of protons up to a centre of mass collision energy of 14 TeV. The LHC accelerator is housed in the 27 km long tunnel originally built for its predecessor accelerator, LEP, and is situated about 100 m underground. The use of this pre-existing tunnel places physical constraints on the achievable collision energy, given the presently available accelerator technologies. The LHC uses RF cavities to accelerate the proton beams, and a system of super-conducting dipole and quadrupole magnets to bend and focus the beams of protons, which are brought to collide at four interaction points around the ring. At these points sit the four detectors of the LHC; ATLAS[30], CMS[31], LHCb[32], and ALICE[33]. Figure 2.1 shows the layout of accelerators at the CERN complex.

The protons start their journey as hydrogen atoms; extra electrons are injected to form negative hydrogen ions H^-, these are then accelerated by a high potential electrode, and the electrons are stripped off as the ions pass through a carbon foil. The resulting protons enter a series of accelerators, which sequentially increase in energy. The protons pass from the LINAC2 (50 MeV) to the Proton Synchrotron Booster (PSB - 1.4 GeV), and are next injected to the Proton Synchrotron (PS - 26 GeV) where the LHC bunching structure is applied with 40 MHz and 80 MHz RF cavities, resulting in a bunch spacing of 25 ns. These bunches then enter the Super Proton Synchrotron where they are accelerated up to 450 GeV, before being injected into the LHC ring where the counter-rotating bunches are accelerated by the system of superconducting dipole magnets to the design energy of 7 TeV per beam. The beams of protons are focussed and steered into collision at the interaction points by a system of superconducting quadrupole magnets, which when operating with the designed beam parameters results in a luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

In addition to protons, the LHC has been designed to collide heavy (lead, gold) ions for a portion of the run schedule. These collisions are the focus of the ALICE experiment,

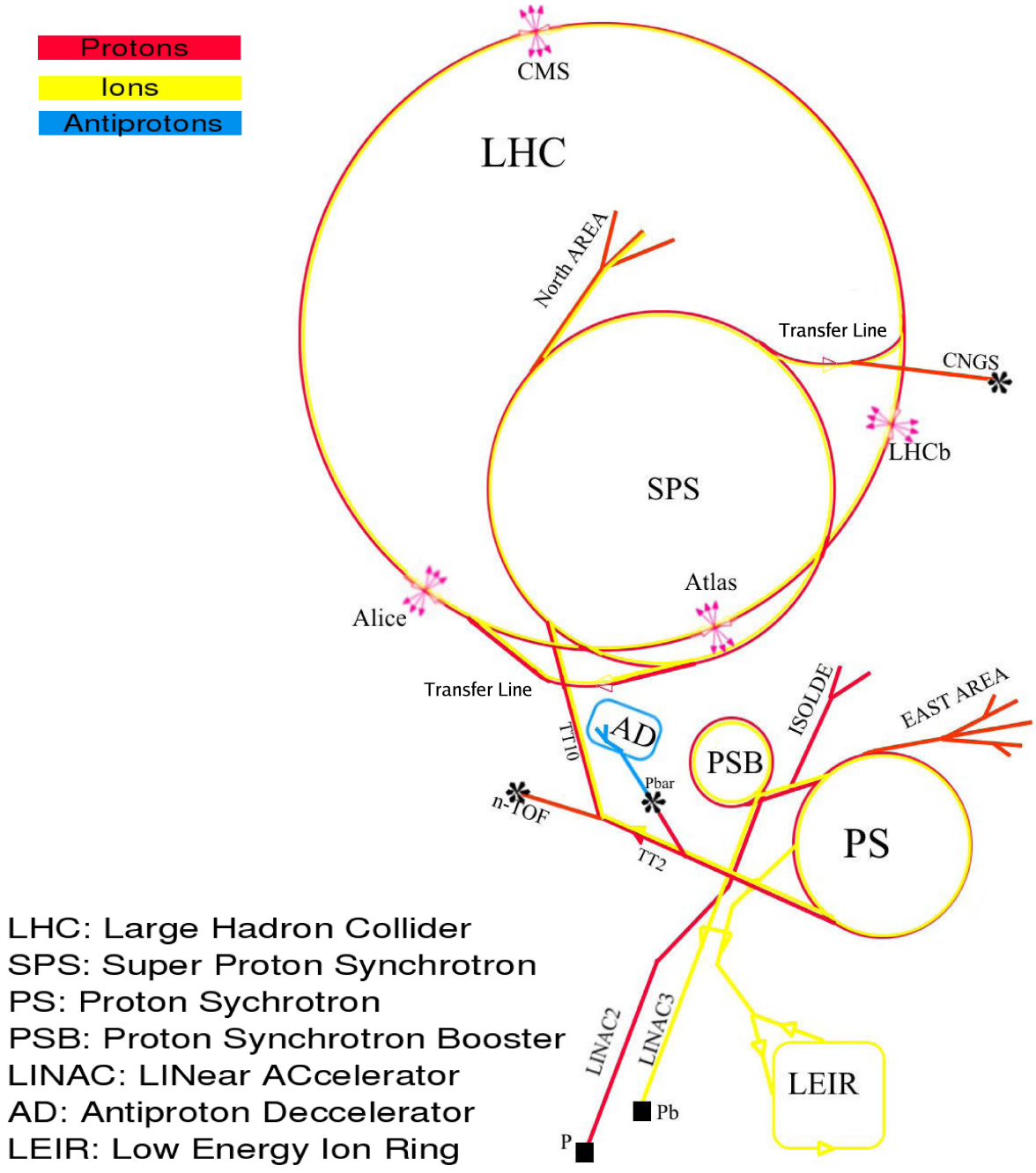


Figure 2.1: Layout of accelerators at CERN site, the red path represent protons, the yellow path - heavy ions(adapted from[34])

whose physics program exploits the high energy density of the ion-ion collisions, focussing on the study of quark-gluon plasmas, and high multiplicity physics. ATLAS and CMS also plan a heavy-ion physics program.

2.1.1 Operational Experience

The LHC saw the first injection of proton beams in September 2008, but progress was halted after only 9 days due to a technical failure. The hardware fault was fixed in the ensuing months, and the first collisions of beams at 900 GeV were recorded by the LHC detectors in November 2009. Following the 2009/2010 winter shut-down period, a successful period of beam commissioning came to fruition with collisions at a record 7 TeV centre of mass energy in March 2010. Taking into consideration safety concerns, this energy is adopted as the operational energy for proton collisions, until the planned 2012 shut-down when modifications to the accelerator can be made.

Since March 2010, the LHC has been running at 3.5 TeV per beam, with efforts focussed on a gradual increase in luminosity. During 2010, the total integrated luminosity recorded by the ATLAS detector was $\mathcal{L} \approx 45\,\mathrm{pb}^{-1}$, with an instantaneous luminosity of $\approx 1 \times 10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. It is expected that by the end of the first running period (2012), $\approx 1\,\mathrm{fb}^{-1}$ integrated luminosity will be collected at $\sqrt{s} = 7\,\mathrm{TeV}$, with ATLAS. It is for this reason that the $1\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 7\,\mathrm{TeV}$ scenario is explored in the analysis of $Z\gamma$ events presented later in the thesis.

2.1.2 Pile-up

Increasing the luminosity is an important task for the LHC, since the only way to collect more events of rare processes (with low cross-sections), is to collect more statistics (higher luminosity). This is achievable by increasing the number of bunches, as well as the number of protons in each bunch; however this has the side-effect of creating more collisions per

bunch-crossing. At the design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, there are expected to be an average of 23 collisions per bunch crossing. This effect is known as pile-up, and must be taken into account in any physics analysis. However, with the instantaneous luminosity achievable in the early running period, pile-up will not be a serious problem (events triggered at a luminosity of $10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ are expected to contain 1 collision on average).

2.2 The ATLAS Detector

ATLAS is the largest of the four LHC detectors, measuring 25 m in diameter and 44 m in length, weighing about 7,000 tonnes, and sitting about 100 m underground. Its design is built upon the knowledge earned from previous particle detectors, sharing the common format of layers of different detector subsystems each designed for measuring different types of particles, capturing as many of the outgoing particles as is feasible, in order to achieve the fullest picture of what is going on in every collision event. The technical specifications of ATLAS are optimised to meet the requirements of the physics program, which, due to the interaction conditions at the LHC, are:

- Fast, radiation-hard detector elements.
- High granularity for track-finding, and to reduce confusion of signals from nearby tracks.
- Large acceptance in pseudo-rapidity η, and full azimuthal coverage ϕ.
- Good charged particle momentum resolution and reconstruction efficiency.
- Excellent electromagnetic calorimetry with good electron/photon identification, complementary hadronic coverage for jet and missing transverse energy (E_T^{miss}) measurements.
- Standalone muon identification, with unambiguous charge identification, and excellent momentum resolution, over a wide range of momenta.

- An efficient, flexible, and robust triggering system.

Figure 2.2 shows the layout of the main ATLAS detector. Details of the various sub-systems of ATLAS are described in the following sections.

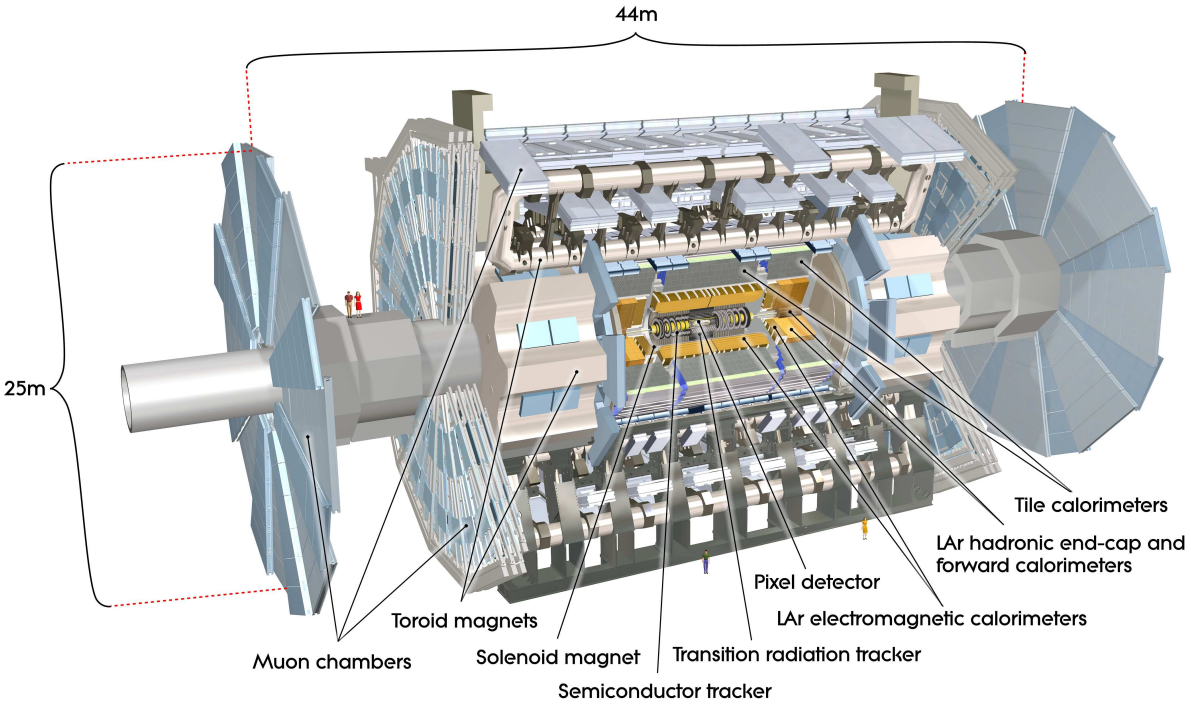


Figure 2.2: The ATLAS detector[30].

2.2.1 ATLAS Co-ordinate system

The co-ordinate system used in ATLAS is defined by the beam direction, which lies along the z-axis. The x-y plane lies transverse to the beam direction, with the positive x-axis pointing from the interaction point towards the centre of the LHC ring, and the positive y-axis pointing upwards. The nominal interaction point is at the origin of the co-ordinate system (0,0,0). The polar angle θ is the angle of deviation from the beam axis, and the azimuthal angle ϕ is the angle about the beam axis. A useful quantity for LHC physics is the pseudo-rapidity η, which is defined as $\eta = -\ln\tan(\theta/2)$. Other useful quantities such as transverse[1] momentum p_T, and transverse energy E_T, are measured in the x-y plane.

[1]transverse quantities, calculated orthogonally to the beam-pipe axis z, are useful in proton-proton collisions since the momentum in the transverse direction is 0 before a collision, and so momentum

The separation distance ΔR is commonly used, and is defined in pseudo-rapidity-azimuth space as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

2.3 Inner Detector Tracking

The inner detector (ID) is the closest detector system to the interaction point, covering the pseudo-rapidity range up to $|\eta| = 2.5$. It is comprised of three independent, but complementary sub-detectors employing different technologies based on the proximity to the beam (radiation hardness), required granularity (occupancy) to achieve sufficient momentum resolution measurements, and cost. The inner detector sits inside a 2 T solenoid magnet, which supplies the magnetic field required for charged particle momentum calculations. Figure 2.3 depicts the layout of the inner detector sub-systems, and figure 2.4 shows the radial distribution of ID layers. An important consideration in the design of the ID is the amount of dead material; this needs to be kept to a minimum to keep the number of conversions and unnecessary interactions to a minimum, whilst at the same time balancing the rich set of features which are required.

2.3.1 Pixel Detector

The inner-most detector is the pixel detector, which is comprised of three pixel layers arranged in concentric cylinders about the beam axis, with flat disks perpendicular to the beam-axis at either end. Each track typically crosses three pixel layers. There are a total of 80.4 million channels [30], with each pixel measuring a minimum of $50 \times 400\mu\text{m}^2$ in R-$\phi \times z$ (barrel). This results in a typical accuracy in the barrel region of 10μm in R-ϕ, and 115μm in z; and in the end-cap disks 10μm in R-ϕ, and 115μm in R. The pixel detector is invaluable to the identification of secondary vertices and the measurement of decay lengths. Due to the proximity to the beam (see figure 2.4), and the associated

conservation can be applied to the outgoing particles.

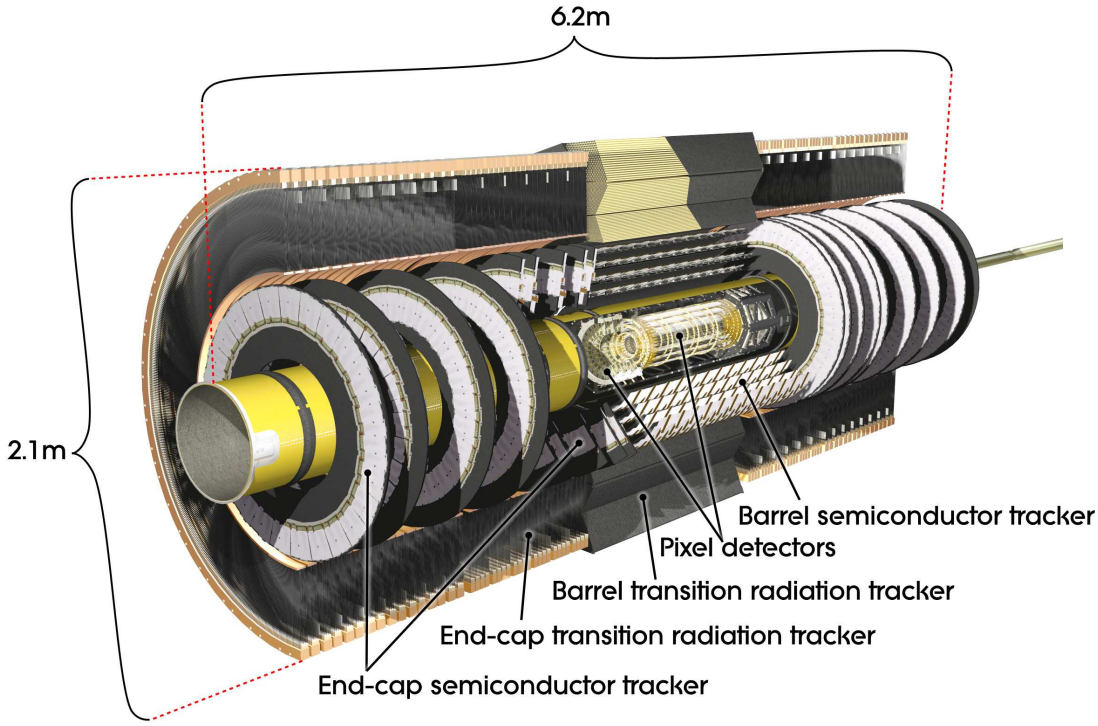


Figure 2.3: Layout of the Inner Detector system[30].

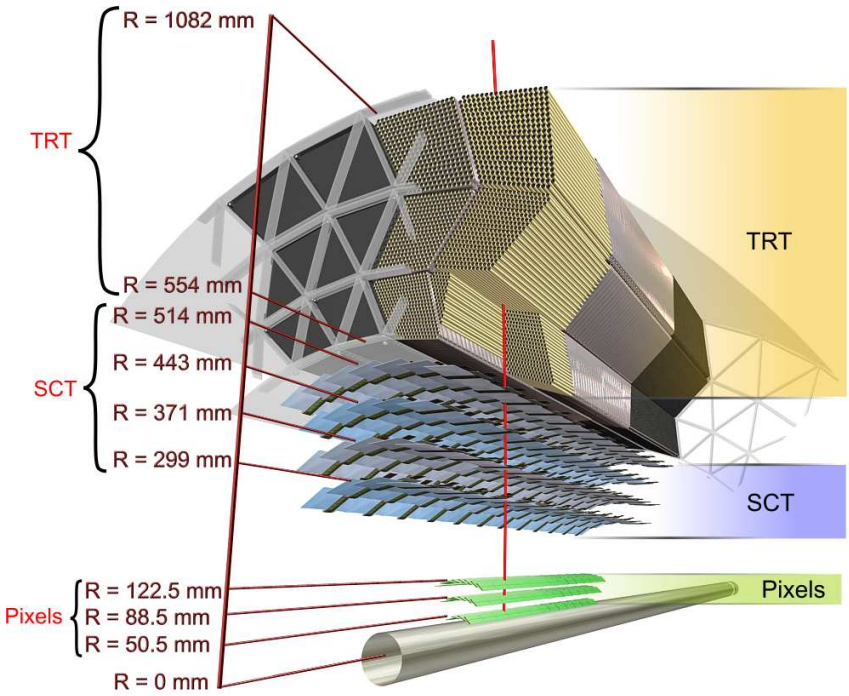


Figure 2.4: Radial distribution of inner detector barrel layers[30].

accumulation of radiation damage, it is planned that the pixel detector inner layer will eventually be replaced. The inner-most layer of the pixel detector is often referred to as the b-layer.

2.3.2 Semi-Conductor Tracker

The next sub-detector out from the pixel detector is the semiconductor tracker (SCT); it consists of approximately 6.3 million read-out channels. In the barrel region, each particle track crosses four SCT layers (four space points), each layer consisting of two wafers of silicon micro-strip detector glued back-to-back with a stereo angle of 40 mrad; each wafer contains more than 750 strips, each separated by 80μm. The angle between the wafers in the SCT modules allows for a z measurement in addition to R-ϕ. In the end-caps the SCT strips run radially outward, again with a 40 mrad stereo angle, allowing for measurement in R in addition to R-ϕ. The achievable accuracy of measurements in the barrel are 17μm (R-ϕ), and 580μm (z); and in the end-caps 17μm (R-ϕ), and 580μm (R).

2.3.3 Transition Radiation Tracker

The last sub-detector comprising the ID is the transition radiation tracker (TRT). Due to the larger radius (volume), lower track density, and in order to save on cost, the TRT uses a cheaper technology than the pixel and SCT detectors. It consists of a large number of 4 mm diameter straw tubes with a central 30 μm gold-plated tungsten wire, the straws are 144 cm in length (37 cm in the end-caps). The straws run parallel to the beam in the barrel region and radially outwards in the end-caps, covering a pseudo-rapidity range up to $|\eta| = 2.0$. The barrel TRT provides an extra 36 space points to the track fitting (per track), and is capable of a measurement only in R-ϕ with an accuracy of 130 μm per straw. Particle identification is enhanced by the use of Xenon gas in the tubes, which facilitates the observation of transition radiation photons; the probability to radiate is proportional to the Lorentz factor, γ, of the particle, which is particularly useful for the

identification of electrons with their low rest mass.

The combination of high precision tracking in the pixel and SCT, with the extended track fitting range, and 36 space points provided by the TRT, result in high precision R-ϕ and z measurements, and robust pattern recognition leading to a good track momentum measurement.

2.4 Calorimetry

In order to satisfy the requirements of the ambitious physics program, the calorimetry in ATLAS must be capable of high precision measurements of electron and photon energies, provide a flexible and robust jet reconstruction, and a good missing transverse energy measurement. The calorimetry is split into electromagnetic (ECAL) and hadronic (HCAL) sub-detectors, each employing different technologies based on η range and required granularity. The calorimeters work on the basic principle that energetic particles will interact with the active material of the detectors and radiate particles in a cascade, creating a shower of particles (depending on the energy of the particle and material of detector); the calorimetry profiles the shape/distribution of these showers. The direction of the shower provides a measure on the direction of flight of the particle; the width and shape of showers help in the discrimination of particle type (electron/photon showers are more compact than hadronic showers). A cut-away illustration of the calorimeters can be seen in figure 2.5.

2.4.1 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a sampling calorimeter which employs liquid argon (LAr) as the active material, inter-spliced with layers of a lead (Pb) absorber. It is split into a barrel region ($|\eta| < 1.475$), and end-cap regions ($1.375 < |\eta| < 3.2$). The layers have an accordion-like structure (this can be seen in figure 2.6), which provides a smooth

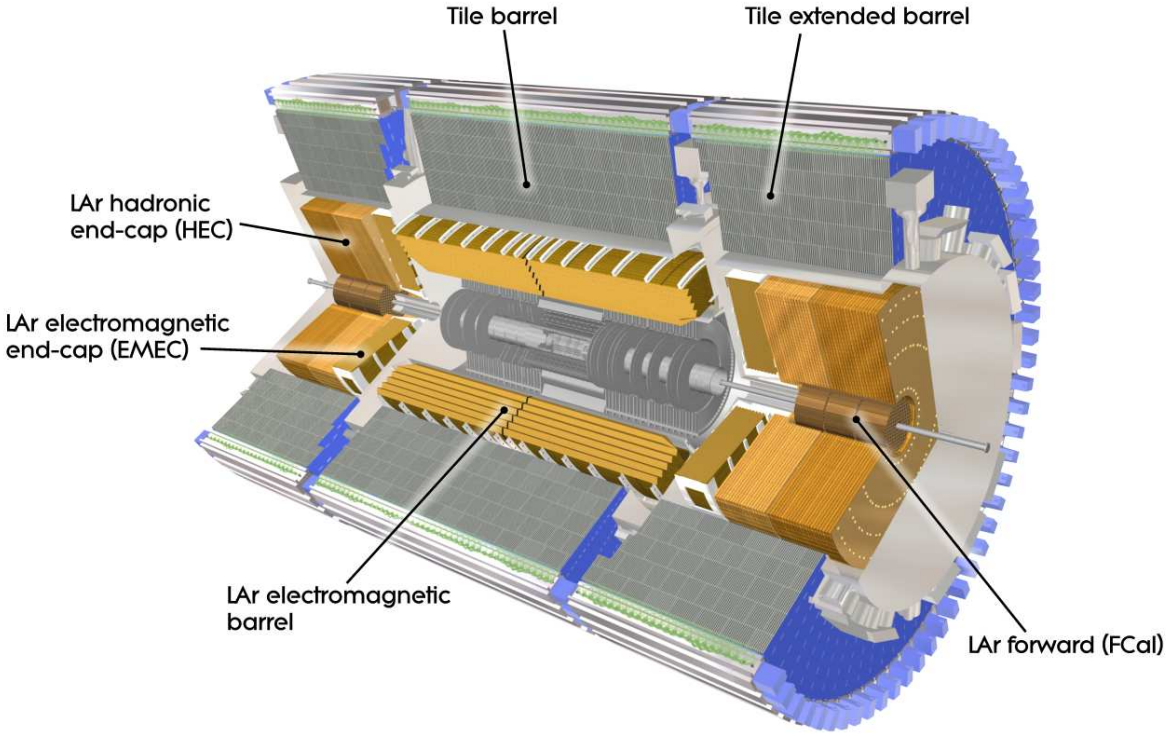


Figure 2.5: Cut-away sketch of ATLAS Calorimetry[30].

crack-free ϕ coverage. In the barrel region there are three layers to the detector; the first layer has fine-granularity sampling strips (up to $|\eta| < 2.5$) which aid in the discrimination of $\pi^0 \rightarrow \gamma\gamma$ decays from isolated photons; the second layer has a coarser granularity and provides the majority of the absorption depth, this layer is useful for the study of width and isolation of showers; the final layer has a coarser granularity than the second layer (since it is further out), and extends the depth to $22X_0$ radiation lengths ($36X_0$ in the end-caps). A sketch of a LAr barrel module depicting the cell granularity in the three layers can be seen in figure 2.6.

In the pseudo-rapidity range $3.1 < |\eta| < 4.9$ sits the forward calorimeter (FCal). This is made up of three layers, the first of which is a LAr electromagnetic calorimeter; here the lead absorber is substituted with copper, which is more suited to the higher radiation in this region.

In the region $|\eta| < 1.8$ there is an extra layer of LAr before the first layer of ECAL, this acts as a pre-sampler; it is used to correct for energy losses in the inner detector, solenoid

and cryostat. To keep the amount of dead material before the detector to a minimum the LAr calorimeter and the solenoid share a common vacuum vessel, removing the need for two extra walls, as would be the case if separate vacua were used.

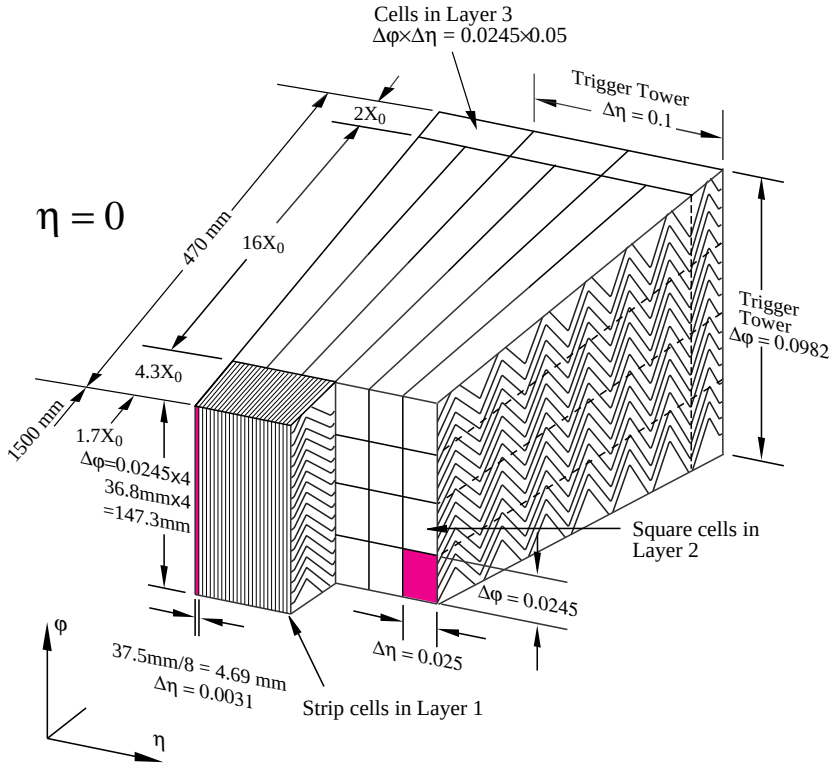


Figure 2.6: A depiction of a LAr barrel module, showing the cell granularity in the three layers, and the accordion-like structure [30].

2.4.2 Hadronic Calorimeter

The hadronic calorimeter (HCAL) is responsible for measuring the showers created by hadronic jets which are normally initiated in the EM calorimeter, but deposit the majority of their energy in the HCAL. The HCAL is split into three different regions, which employ different technologies due to the radiation yields at varying pseudo-rapidities. In the region $|\eta| < 1.7$ is the scintillating tile calorimeter (TileCal); this is divided into central ($|\eta| < 1.0$) and extended barrel ($0.8 < |\eta| < 1.7$) parts, and uses a steel absorber with scintillating plastic tiles as the active material. Scintillation light is read out by wavelength-shifting fibres into photo-multiplier tubes, these are grouped such that the read-out is projective from the interaction point in η. The TileCal extends radially from 2.28 m out to 4.25 m,

a depth of about 10 interaction lengths λ; this is sufficient to absorb all but the highest energy hadronic jets, keeping punch-through to the muon system at a minimum.

The region $1.5 < |\eta| < 3.2$ is covered by the LAr hadronic end-cap calorimeter (HEC); this is a LAr-copper sampling calorimeter which shares the same cryostat as the LAr EM end-cap to keep the material budget at a minimum. It consists of two independent wheels in each end-cap, comprised of 32 wedge-shaped modules each. The modules each have 2 layers, resulting in 4 sampling layers in each end-cap region. In the forward region $3.1 < |\eta| < 4.9$ sits the LAr forward calorimeter (FCal) (as previously mentioned), the outer two layers are for hadronic calorimetry. They employ tungsten absorber in place of the copper used in the first layer, which is more appropriate for measuring hadrons in this pseudo-rapidity region.

2.5 Muon Detector System

The outermost layer of subdetectors comprise the muon system; it is designed to provide a momentum measurement from the fitting of a track to a number of position measurements recorded as the muons travel out through the detector, following a curved trajectory supplied by the up to 4 T magnetic field of the toroid system. The super-conducting air-core toroid system is made of 8 coils in the barrel region, and a further eight coils in each end-cap, and is designed to provide a magnetic field orthogonal to the direction of flight of the muons. The muon system consists of four separate sub-detectors, namely the Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). The MDT and CSC are responsible for accurate measurements of position, whilst the RPC and TGC are used for triggering purposes, and have excellent timing resolution. The layout of the muon sub-detectors can be seen in figure 2.7.

The MDT is comprised of $30\,mm$ diameter tubes ranging in length between $1\,m$ and

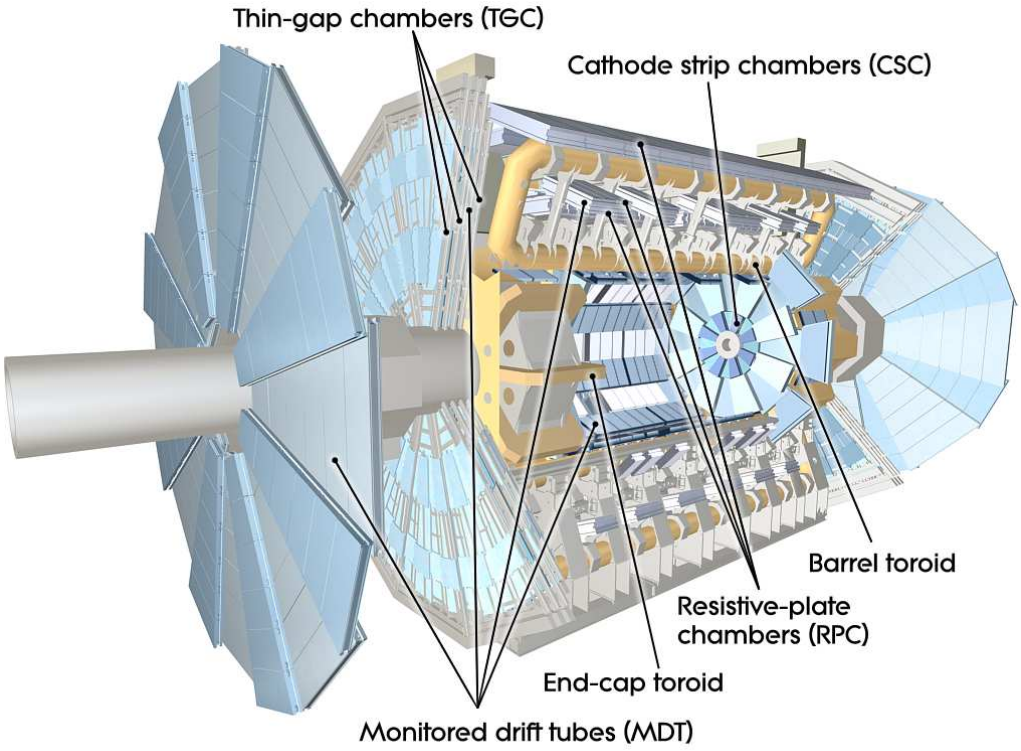


Figure 2.7: Cut-away view of muon detector system and toroidal magnets[30].

$6\,m$ arranged into modules made from two layers of tubes separated by a space; these modules are spread over three to eight layers (cylindrical layers in the barrel region, and flat layers perpendicular to the beam in the end-caps), covering a pseudo-rapidity range up to $|\eta| < 2.7$. The MDT system provides on average 20 position measurements per muon track. In the region $2.0 < |\eta| < 2.7$ the muon rate is high, it is here that the high granularity CSC sits. The CSC is a system of proportional chambers with cathode planes segmented into orthogonal strips, allowing for a simultaneous measurement of ϕ and R, crucial in this high track density region. The other muon detectors are the RPC, providing a measurement of ϕ and z in the pseudo-rapidity range $|\eta| < 1.05$; and the TGC which covers $1.05 < |\eta| < 2.4$, and measures ϕ and R. Despite the coarse granularity of the RPC and TGC, they are invaluable in that their timing resolution is sufficient to enable bunch crossing identification, and trigger the readout of the high resolution MDT and CSC detectors (via the Level-1 trigger system).

2.6 Forward Detectors and Luminosity Measurement

In addition to the main detector elements presented above, there are three further detector sub-systems situated at increasing distances from the interaction point. The LUCID detector is the first; situated at ± 17 m from the interaction point, surrounding the beam-pipe at the edge of the cavern. It uses Cherenkov imaging to provide online luminosity measurements. The next system is the Zero-Degree Calorimeter (ZDC); this is located at ± 140 m along the beam pipe from the interaction point, and is designed to measure forward neutrons from heavy-ion collisions. The final sub-detector is ALFA; situated at a distance of ± 240 m from the interaction point, it uses Roman pot detectors, and provides an absolute luminosity measurement.

2.7 Trigger system

A trigger is a system which preferentially selects the most interesting physics events out of a sea of background events, which due to interaction conditions at the LHC is dominated by soft QCD processes. The proton-proton interaction rate at the LHC, at design luminosity, is approximately 1 GHz (40 MHz bunch-crossing frequency, times an average ~ 25 collisions per bunch), but the event storage rate is limited to around 400 Hz (for events averaging 1.5 MB in size), which is limited by the distribution capabilities of the external computer resources. The trigger is responsible for reducing this rate, whilst at the same time selecting events which contain useful physics. These are usually events with a large momentum transfer between interacting partons, resulting in outgoing particles with significant momentum transverse to the beam pipe (high p_T). The functionality of the trigger system is motivated by the requirements of the physics goals; muons, electrons, taus, photons, jets, missing energy (neutrinos or new physics signatures), and minimum-bias events must all be capable of passing the trigger.

The ATLAS trigger system[30] is divided into 3 levels (See Figure 2.8), level-1 is a

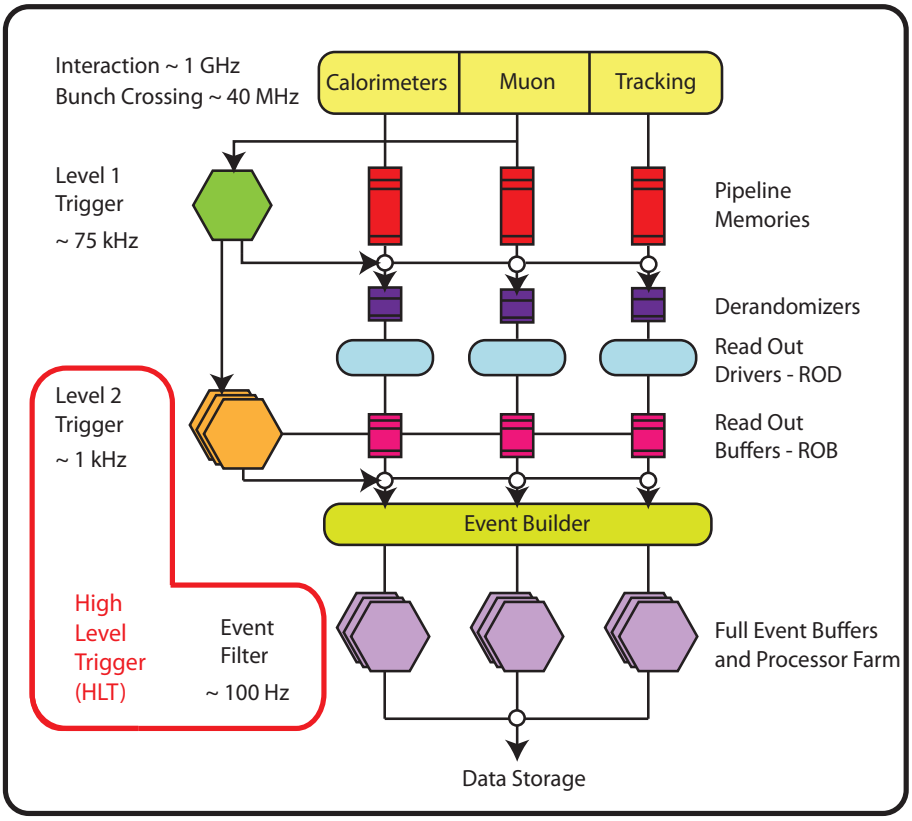


Figure 2.8: Schematic of ATLAS Trigger system.

hardware based system which reduces the output event rate to a maximum of 75 kHz in a specified latency of 2.5 μs. The latency is constrained by the front-end electronics of various sub-detectors, and the pipeline lengths of the data acquisition/readout (DAQ) system. In order to make the decision as quickly as possible, level-1 uses relatively coarse information from the detector. Information from the level-1 detector subsystems (calorimeters, muon trigger chambers, e.t.c) is examined for each event; energy thresholds are applied and passing objects counted. This information is combined across the sub-detectors (object types), and compared against the trigger menu (the selection of items to be recorded) in the Central Trigger Processor (CTP). If an event passes the CTP's requirements, a level-1 Accept (L1A) decision is made, and the Region of Interest builder (RoIB) sends a Region of Interest (RoI), containing geographical information about the trigger item(s) to level-2.

Level-2 is a system of software based selection algorithms running on processor farms; the algorithms are seeded by an RoI received after the level-1 processing. Level-2 collects

the full readout of data from the detector elements relevant to the RoI. Level-2 refines the decision made at level-1, reducing the rate to about 1 kHz in a latency of about 10 microseconds. Events accepted by level-2 are sent to the *Event Builder*, which assembles the event ready for further examination by level-3.

Level-3 is known as the Event Filter (EF), and collectively with level-2 forms the High Level Trigger (HLT). The event filter selects events using a sequence of software based algorithms. The EF has access to the full event data, since it is called after the event building step. The EF has an output rate of about 200 Hz, and selection and classification takes a few seconds. Events passing the EF are written to disk, and distributed worldwide via the LHC Computing Grid (LCG)[35].

Chapter 3

Level-1 Calorimeter Trigger Timing Calibration

This chapter relates to the work carried out in collaboration with the Level-1 Calorimeter trigger (L1Calo) group, which formed a substantial part of my PhD research. The majority of interesting physics processes being studied at the LHC will contain electrons, photons, jets, tau's, or missing transverse energy signatures, for which L1Calo is responsible for the initial identification. It is important that L1Calo is a reliable, stable, and well calibrated system; if L1Calo fails, most of the interesting events will not be recorded, and the extensive physics program of ATLAS can not be accomplished. The focus of the work carried out with L1Calo was in the development and testing of a digital timing calibration procedure for the Cluster Processor (CP) system. The work involved the design, implementation, and testing of a software based calibration routine. The task included the development of software for the control of the Clock Alignment Module (CAM); software to analyse and determine the most appropriate digital clock settings on each Cluster Processor Module (CPM) board, and store these results in the online conditions database (COOL). The software which I developed is incorporated into the substantial, pre-existing L1Calo software framework. The calibration procedure was initially developed with the use of the Birmingham test-rig (a local, scaled down version of the full L1Calo system

in ATLAS); and later tested, and successfully deployed in the full L1Calo system within ATLAS. In the following sections, the CP system and digital timing requirements are explained, along with the functionality of software which has been developed, and examples of results from a typical calibration run.

3.1 Level-1 Calorimeter Trigger System

The L1Calo system provides the first level trigger decision for electrons/photons, taus/hadrons, jets, sum-E_T and E_T^{miss}, based on information from the calorimetry; this task is constrained by the detector readout hardware (the limit of the detector front-end pipeline memory buffers), and the 40 MHz bunch-crossing frequency, to be completed within 2.5μs. Given the complexity of the task, and the stringent latency constraints, a system of purpose built hardware processors were designed and built to achieve the goal. For each event, the multiplicities of identified trigger objects passing predefined E_T thresholds are sent to the Central Trigger Processor (CTP), where information about different object types is combined and compared to the trigger menu (this menu specifies which objects are to be selected for different energy thresholds), and an overall Level 1 Accept (L1A) or reject decision is made. When there is an L1A pass decision, geographical information about the relevant trigger items is sent from L1Calo to the Level-2 trigger in the form of a Region of Interest (RoI), where it is used as a seed for the HLT algorithms. At the same time, buffered information from all L1Calo sub-systems is read out by the data acquisition system (DAQ) to facilitate a full monitoring of the performance of the system, and provide important diagnostic information should any unforeseen faults or instabilities arise.

The L1Calo hardware is split over three main sub-systems, the first being the Pre-Processor (PPr) system , which converts the analogue calorimeter signals into a calibrated, digital signal for use by the level-1 trigger algorithms. The other main sub-systems are the Cluster Processor (CP) system, where electron and photon candidates are identified;

and the Jet/Energy-sum Processor (JEP) system, where jets, missing energy, and energy sums are identified. Within the CP and JEP systems, Common Merger Modules (CMMs) combine the crate, and system level results (the multiplicities of objects passing thresholds); and a system of Read Out Drivers (RODs) communicate level-1 data to external systems. An overview of the L1Calo system can be seen in figure 3.1.

3.1.1 Pre-Processor Subsystem

The Pre-Processor System (PPr)[36] is where the analogue signals from the calorimetry are digitised, associated to the correct bunch crossing (not a trivial task considering the analogue calorimeter pulses have a long time-base which spans several bunch crossings), and assigned a calibrated E_T value. There are 7200 trigger towers (each with a granularity of $\sim 0.1 \times 0.1$ in $\Delta\eta \times \Delta\phi$), these are received via receiver stations (where the hadronic calorimeter signal gains are calibrated to convert energy to E_T), into eight PPr crates, each of which contains 16 Pre-Processor modules (PPMs). On-board the PPMs, ten-bit Flash Analogue-to-Digital Converters (FADCs) digitise the signals with a sampling frequency of 40.08 MHz (the LHC bunch crossing frequency). Next, bunch crossing identification is performed (BCID), and a look-up table is used to provide a calibrated E_T value (with pedestal subtraction and noise threshold applied). Finally, bunch-crossing multiplexing (BCMux) is performed on the towers sent to the Cluster Processor, where the data from two channels is combined into one channel[1], halving the number of required connections to the CP system (this is useful, given the tight space/size restrictions on the L1Calo system). Towers are combined into 0.2×0.2 ($\Delta\eta \times \Delta\phi$) jet elements for the Jet Energy Processor (JEP), and the CP and JEP signals are serialised and transmitted in parallel via 480 Mbit/s Low Voltage Differential Signal (LVDS) cables to the CP and JEP subsystems, which are situated in different crates.

[1]This relies on the fact that the BCID logic blanks the consecutive tower in a channel after a BC has been identified; two consecutive bunch crossings of multiplexed data can be used to identify the correct tower and BCID of the multiplexed data with only 1 BCMux bit.

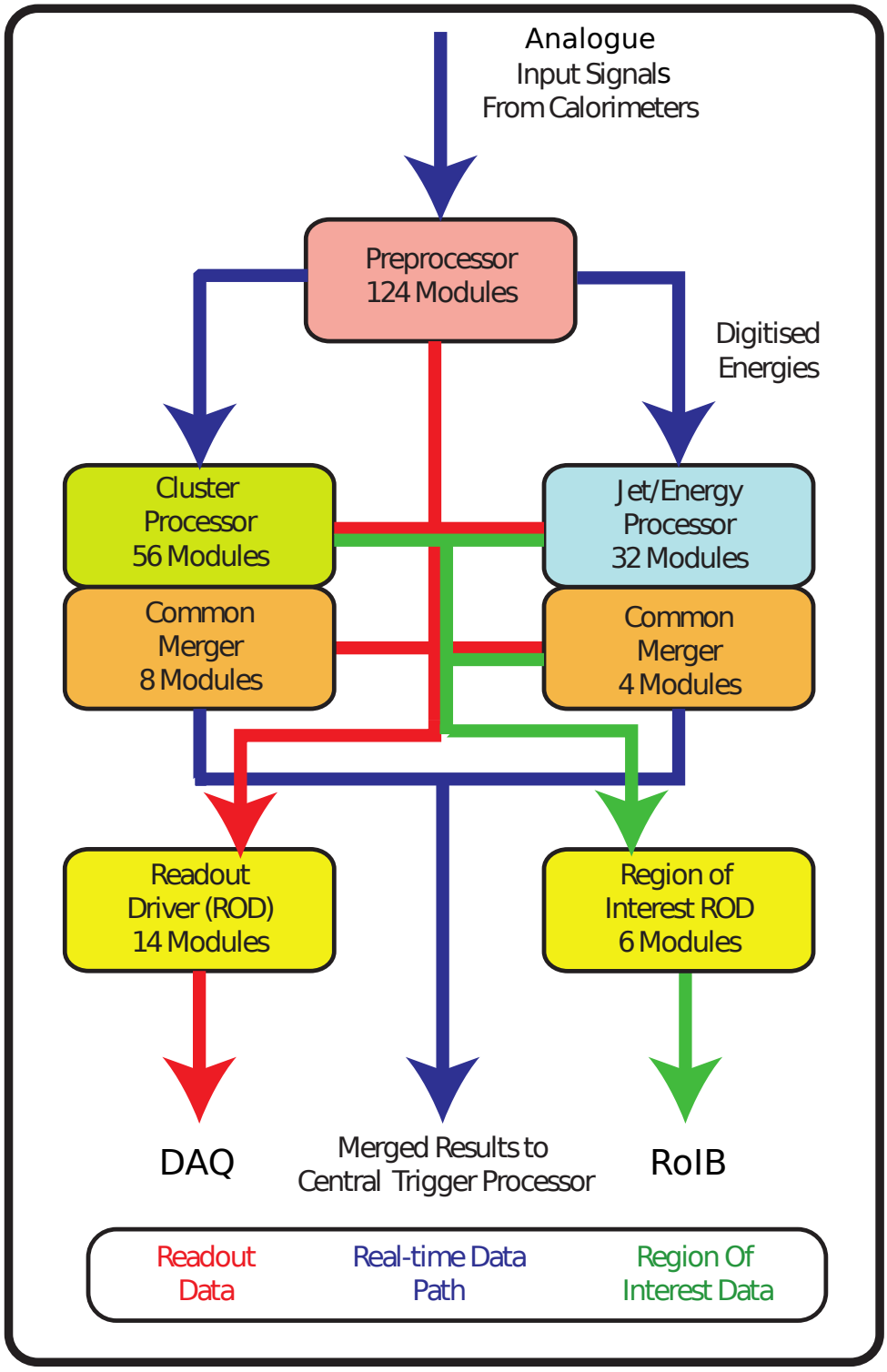


Figure 3.1: Summary of L1Calo subsystem layout.

3.1.2 Cluster and Jet Processor subsystems

The CP and JEP sub-systems perform similar tasks, and as a result share common architectural and hardware features. Both the CP electron/photon and τ algorithms, and the JEP jet algorithm, employ overlapping "sliding window" searches to locate features. The CP system is split over four 9U VME crates, each crate containing 14 Cluster Processor Modules (CPMs), and dealing with 90° in azimuth, and a pseudorapidity range extending to $|\eta| = 2.5$, as shown in figure 3.2.

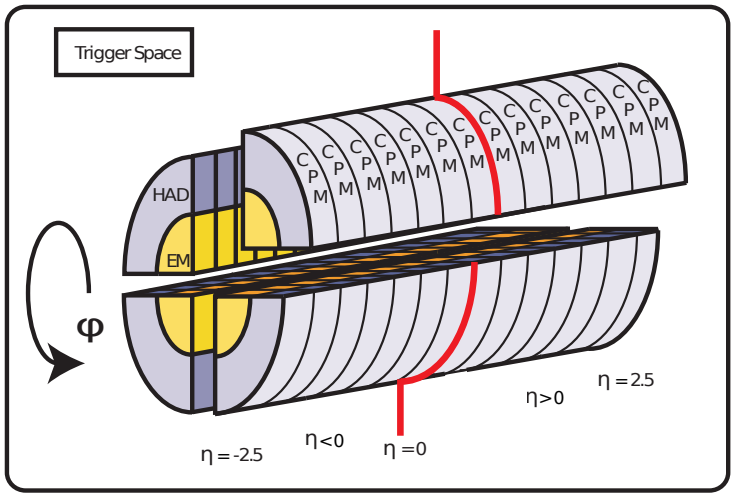


Figure 3.2: Geographic distribution of trigger towers to CP system. Each azimuthal quadrant corresponds to one crate of electronics.

The JEP system is split over only two crates, each crate dealing with two opposing quadrants in azimuth, with each crate containing 16 Jet Energy Modules (JEMs); the jet trigger extends to a pseudorapidity range of $|\eta| = 3.2$, and sum-E_T and E_T^{miss} triggers to $|\eta| = 4.9$. This increased coverage is because the JEP receives input from the Forward Calorimeters (FCAL), in contrast to the CP system which does not. Both CP and JEP systems need to share overlapping data with adjacent modules in order to execute the search algorithms; this is achieved in η via the Fan-in-Fan-out (FIFO) of trigger towers over a high speed custom designed point-to-point back-plane [37], and in ϕ by duplicate towers being sent to azimuthally adjacent crates from the PPr system. The results from the algorithms are summed and merged at the crate level in the Common Merger Modules (CMMs) in both the CP and JEP crates. The system wide results are also merged in

dedicated CMMs, prior to being sent to the CTP, where the overall Level-1 decision is made.

3.2 Cluster Processor System

As previously mentioned, the CP system is split over four crates; figure 3.3 shows the module layout in a single crate. Each crate contains a Single Board Computer (SBC, slot 1) which controls the software based VME register access, and hosts run-time software for controlling the crate modules; CMMs occupy slots 3 and 20, where crate level results are merged; slots 5 to 18 house the 14 CPMs, and a Timing Control Module (TCM) sits in slot 20. The TCM receives and converts the optical clock (the 40.08 MHz LHC clock) signal from the Timing, Trigger, and Control (TTC) system, which is then distributed via the back-plane to the TTCdec (see section 3.2.2) cards on each CPM. The final module in the CP crate is the Clock Alignment Module (CAM) sitting in slot 2, this is used in the timing calibration procedures presented later on, and is described in section 3.3.

3.2.1 CP Algorithms

The Cluster Processor electron/photon algorithm identifies (2×2) clusters of electromagnetic trigger towers where at least one of the four (1×2) or (2×1) component sums exceeds one of 16 predefined thresholds. A set of programmable fixed-value isolation-veto thresholds are set for the 12-tower ring surrounding the (2×2) EM core, the corresponding (2×2) HAD core, and its surrounding ring also. The τ/hadron algorithm differs in that the (2×1) or (1×2) EM tower sums are added to the corresponding (2×2) HAD core behind, before threshold comparison. Accordingly, in the τ/hadron case, isolation thresholds are available only for the 12-tower rings surrounding the (2×2) EM and HAD cores. Figure 3.4 shows a representation of the CP algorithm "sliding windows".

The algorithms[38] run over all possible (4×4) windows, so the windows overlap and

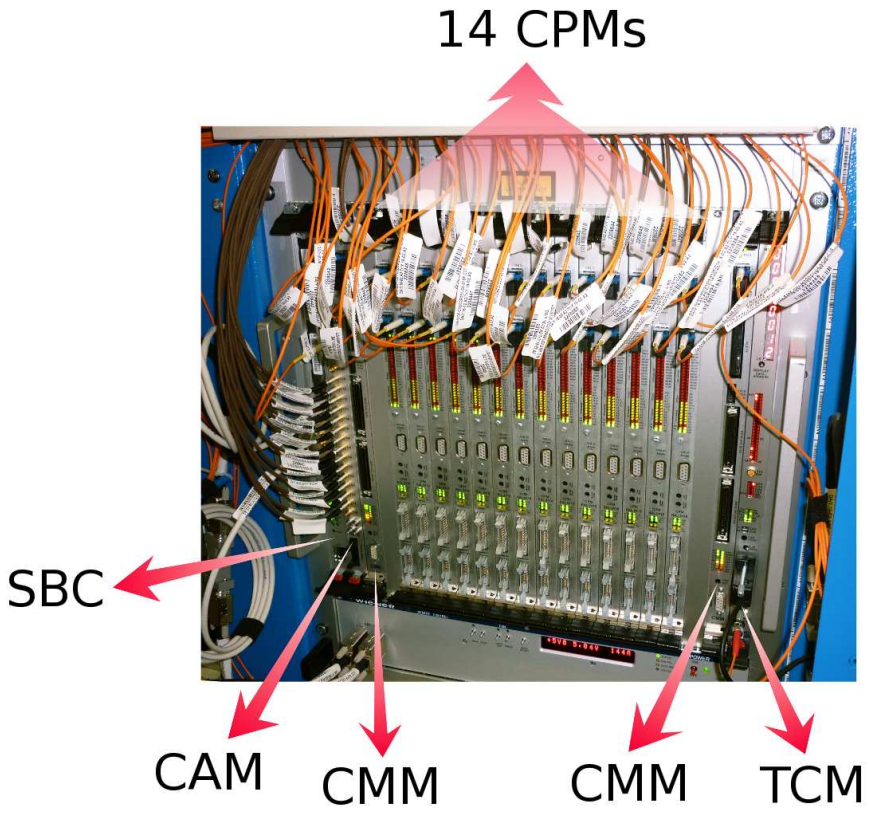
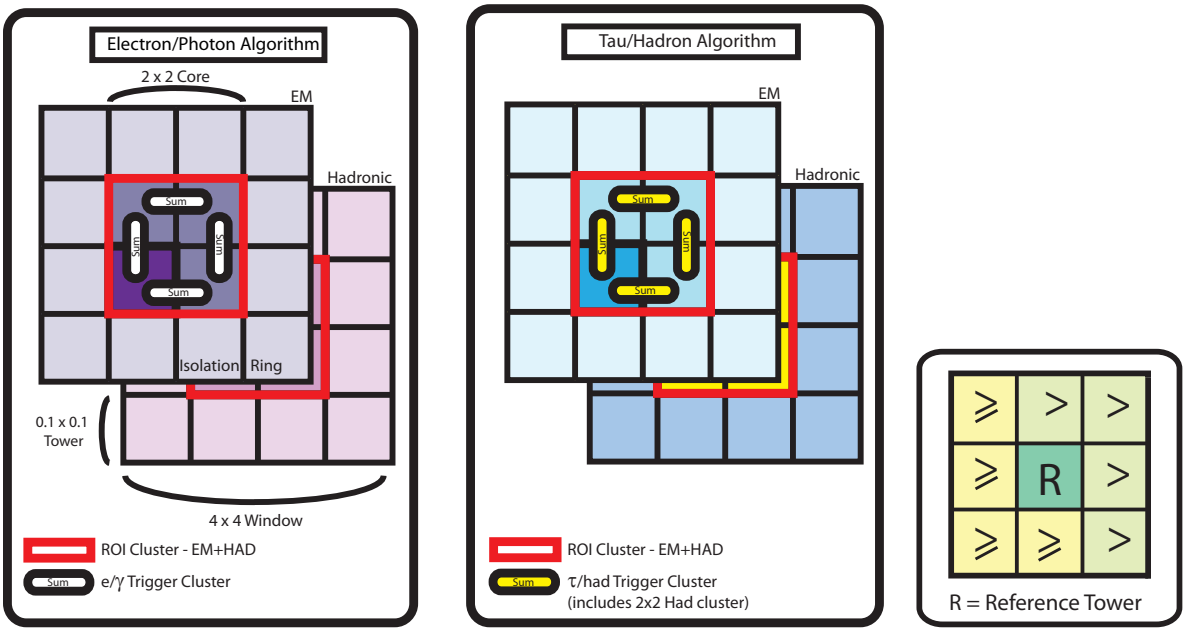


Figure 3.3: CP crate module layout: Slot 1 = SBC; Slot 2 = CAM; Slots 3,20 = CMM; Slots 5-18 = CPM; Slot 21 = TCM; Slots 4,19 = Empty.



(a) (b)

Figure 3.4: (a)CP $(e/\gamma)/(\tau/had)$ algorithm sliding windows. The darkest cell in the 4×4 core is the reference tower . (b) Comparison of Trigger Towers for peak-finding decision.

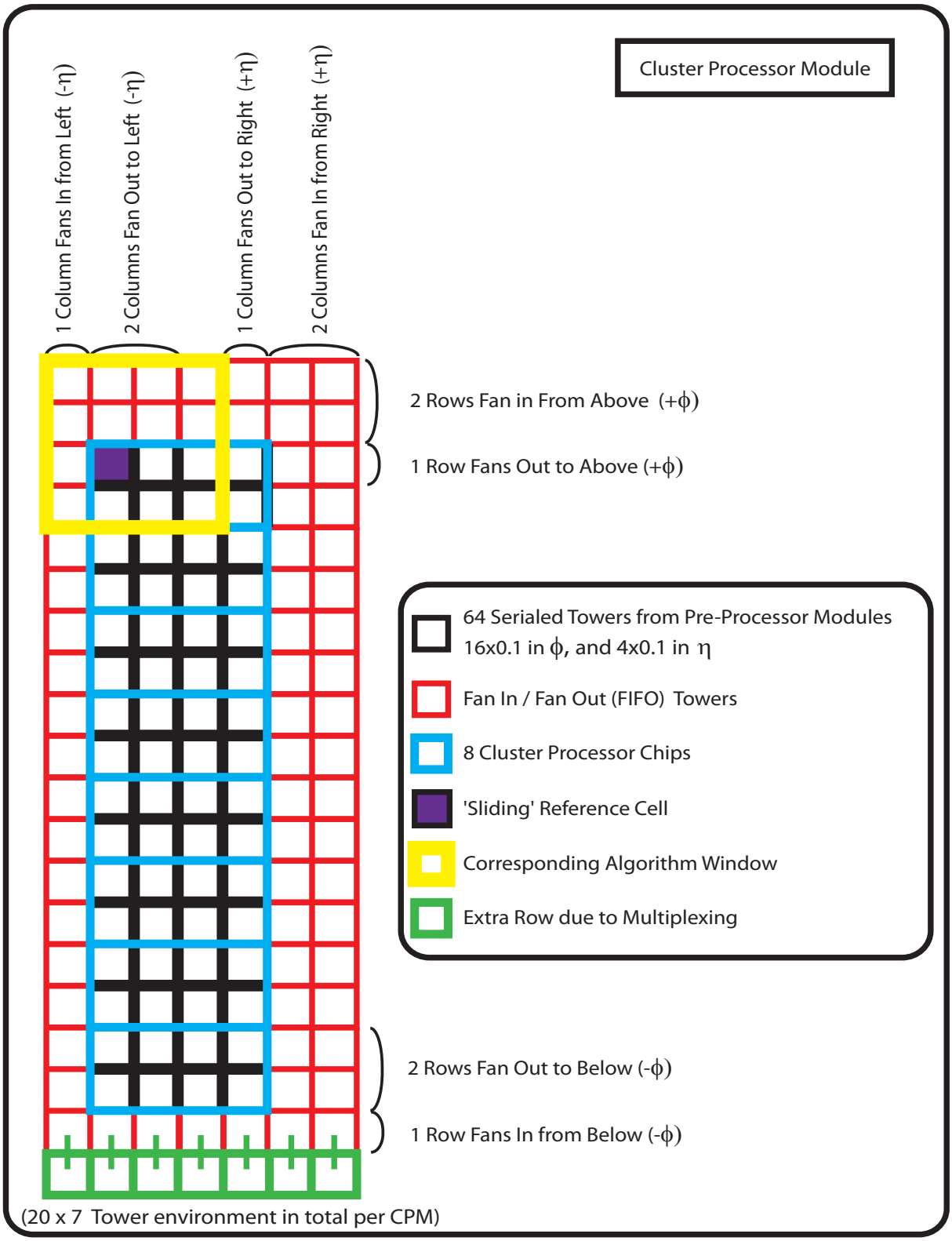


Figure 3.5: Trigger tower allocations for a single CPM, including FIFO towers.

"slide" in steps of one trigger tower element. The overlapping implies a cluster may satisfy the algorithm in two or more neighbouring (4×4) windows. This double counting is avoided by imposing that the sum of the (2×2) central towers (EM + HAD) be a local maximum, with respect to the eight nearest overlapping neighbours. The comparison is made greater than in four of the neighbours and greater than, or equal to in the other four (see Fig 3.4(b)) to resolve problems when comparing digital sums with identical values.

In order to cover the entire $\eta - \phi$ range of trigger space, each CPM contains 8 FPGA processor chips (CP chips), with each chip processing 8 overlapping (4×4) windows. Figure 3.5 shows the trigger tower layout including FIFO, being processed by one CPM. In order to keep the board-to-board transmission low, a large boundary region is used to allow the peak-finding to be performed within each card. The multiplicities of identified clusters from each CP chip are transmitted via the crate back-plane to the CMM for merging of crate/system level results.

3.2.2 CPM Module Layout and Timing Controls

Figure 3.6 shows the layout of components on a CPM. Each CPM receives and de-serialises 80 LVDS cable inputs from the PPr system, via connectors through the back-plane. At the serialiser FPGA chips, data is buffered in pipeline memories, for read out via the data acquisition system should an L1A be broadcast (for monitoring and diagnostic/testing purposes). The serialiser FPGAs clock in the data stream at a frequency of 40.08MHz; the data is strobed with a clock signal broadcast from a daughter module called the TTC decoder (TTCdec) card. The de-serialised data stream is re-timed to 160MHz for clocking in at the CP chip (where there are pipeline memories, for readout by data acquisition, for monitoring/testing).

The timing chip on the TTCdec card decodes the LHC machine clock broadcast from the TTC system, via the TCM, across the crate back-plane. The TTCdec transmits the clock signal via two high resolution phase shifters, with a resolution of 104 ps (this is

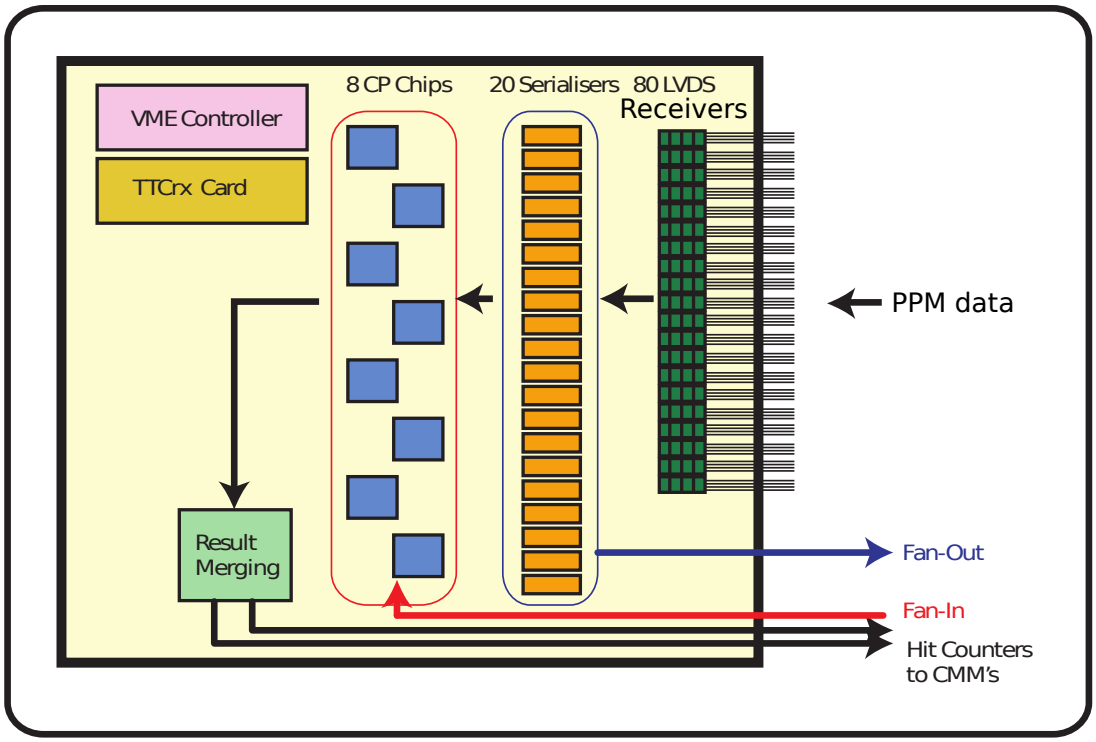


Figure 3.6: Cluster Processor Module board layout.

called a TTC *step*). The two clocks are referred to as `Clock40Des1` (40 MHz Serialiser clock) and `Clock40Des2` (160 MHz CP Chip clock); the names originate from the fact that these clock phases are intended to "deskew" out of time channels. In addition to the `Clock40Des1` and `Clock40Des2` phases, further delays and phase shifts can be applied to the clocks at the serialiser/CP chips. At each serialiser chip, an optional delay of 1 TTC step can be applied to the `Clock40Des1` clock; and at each CP chip there is the option of applying a phase delay of $\frac{1}{4}$, $\frac{1}{2}$, or $\frac{3}{4}$ of a TTC step to each of the 12 fan-in towers only (these towers, coming from adjacent modules, may require a delay to bring them in phase). Figure 3.7 clarifies which towers are Fan-in, or on-board towers, for a single CP chip.

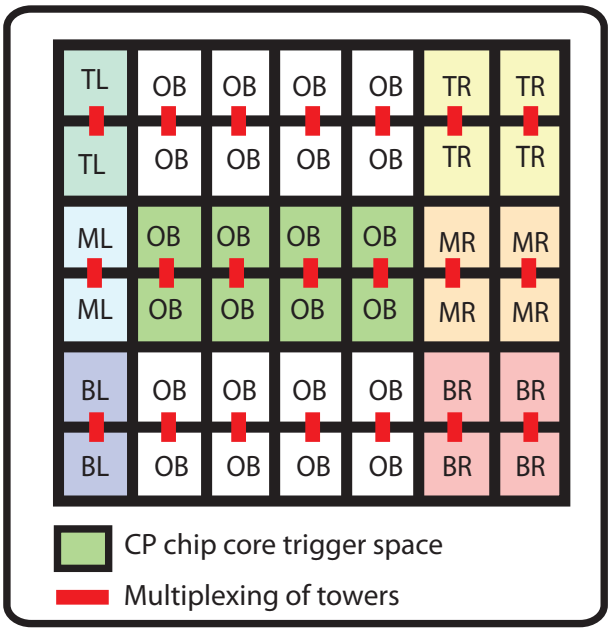


Figure 3.7: The 12 programmable CP-chip clock phases can only be applied to the Fan-In channels. These are: bottom left (BL), middle left (ML), top left (TL), bottom right (BR), middle right (MR), and top right (TR) trigger towers (6 EM + 6 HAD). The on-board (OB) towers are the towers recieved directly into the module via LVDS (not fanned-in across the backplane from adjacent modules). Each multiplexed pair of towers corresponds to one channel on the CP-chip input (see figure 3.5 to see how this relates to the module level).

3.2.3 CPM Digital Timing Calibration

The CPM system is designed as a parallel pipelined[2] synchronous digital system, that is to say that the data path is strobed relative to the LHC clock, via the on-board `Clock40Des1` (serialiser) and `Clock40Des2` (CP chip) clocks. In order that the data is correctly processed by the module, the on-board clock signal(s) must be aligned with the incoming data, as in figure 3.8.

If the clock and the data stream are not properly aligned, i.e. the incoming data is strobed in the transition between a digital 0 or 1, then the data may be misinterpreted. To monitor the data quality, the incoming byte-stream contains a parity error bit for each word. The parity error bit is used to identify when data has been badly strobed. When the phases of the on-board clock and incoming data-stream are appropriately aligned there is

[2]The data from each bunch-crossing follows the previous data along the "pipe", the pipeline memory buffers are sufficient to store all bunch crossings within the latency of the system, so that no data is lost before an L1A is received.

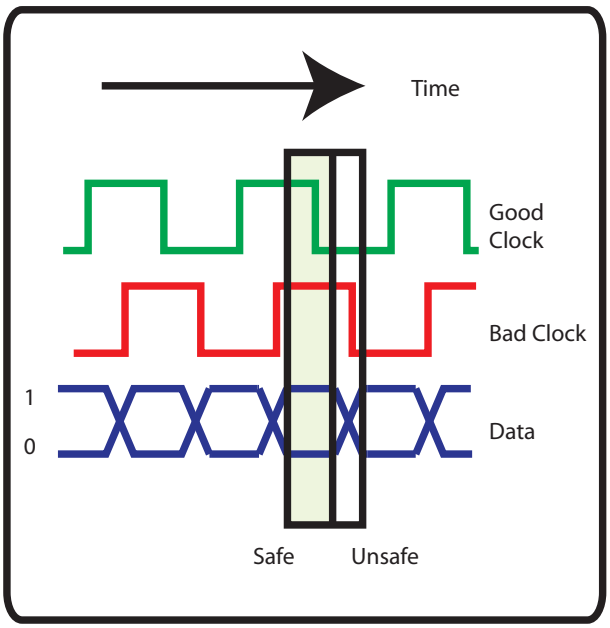


Figure 3.8: Digital clock lock in of data stream, for the case of data being strobed on the clock falling edge.

a parity-error free (safe) window. The goal of the calibration procedure is two fold; at the serialisers, the `Clock40Des1` should be derived for each module which ensures clock synchronisation across all modules, is within the timing safety margin, and is commensurate with keeping the latency to an acceptable minimum. The `Clock40Des2` timing at the CP chips is more difficult, since an average parity error free window here is as narrow as 2 ns; here the optimal clock phases for each configurable channel must be derived, such that the timing margins in each CP chip are as wide as possible, with a `Clock40Des2` value that is safe for all chips on a single board.

The calibration procedures developed for the calibration of the CP digital timing are software based, and rely on the trigger hardware being configured in a special mode of running. During the calibration procedures it is necessary for the digital clocks to be adjusted in an incremental fashion over their permissible ranges; it is following this technique that the term *scan* was adopted. The details of the various scans which have been developed to calibrate the digital timing of the CP system are described in detail in sections 3.5 and 3.6.

3.3 Clock Alignment Module

Setting equal `Clock40Des1` values on all CPM's in a crate does not ensure synchronicity in this clock across the modules (this is also true in the JEP crate), for reasons such as the effect of temperature variation on TTCdec components [39], and differences in path length for the clock propagation across the back-plane. In order to accurately measure the differences in clock phases between modules, the Clock Alignment Module (CAM) was designed. Figure 3.9 summarises the CAM.

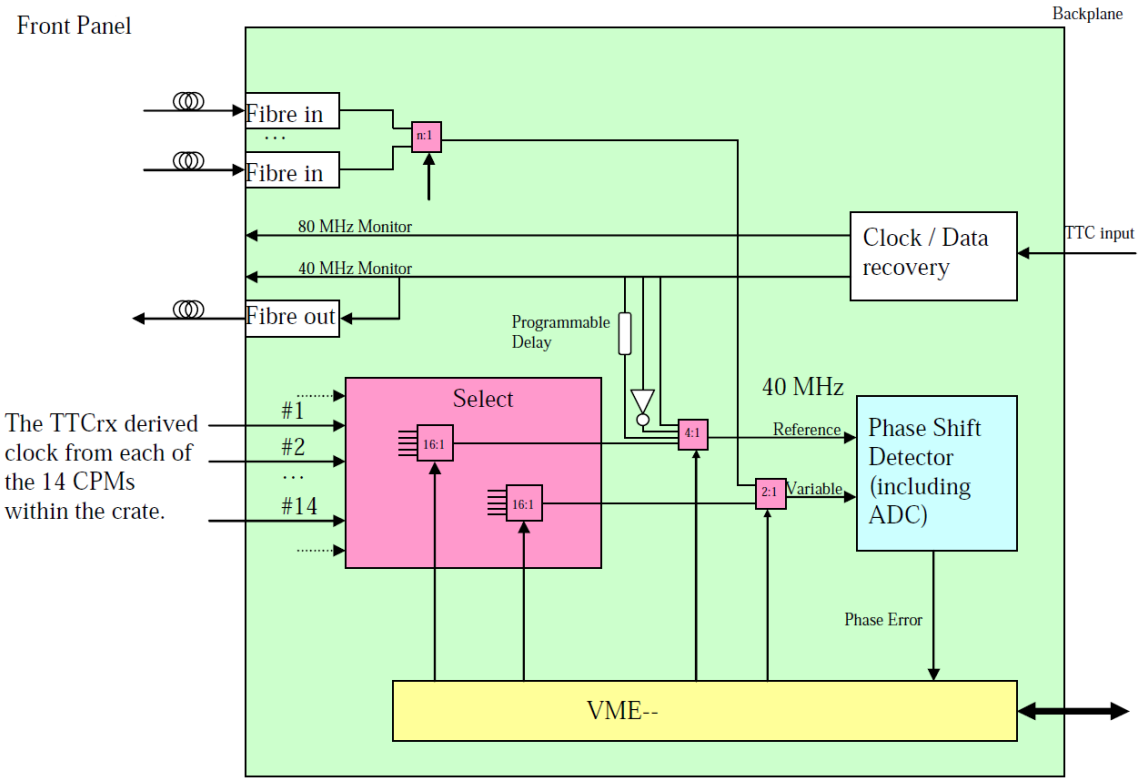


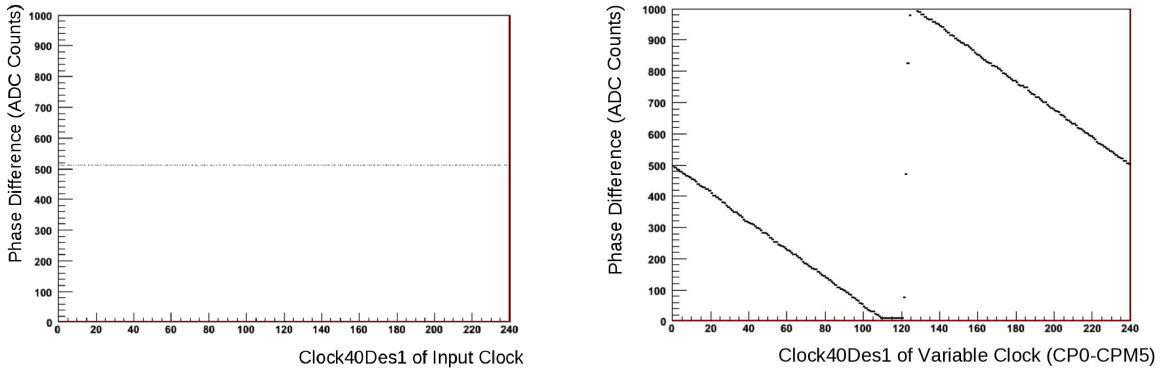
Figure 3.9: Clock Alignment Module layout.

The CAM receives inputs of the `Clock40Des1` (or optionally `Clock40Des2`) signal from the 14 CPM's in a crate via cables into the front panel, as well as deriving an internal copy of the 40.08 MHz TTC clock received from the TCM via the back-plane. On-board the CAM is the logic to select (via VME messages) a *reference*, and a *variable* clock (from any of the input clocks), and compare/readout the relative phase difference. The 10-bit phase shift-detecting ADC chip has a resolution of 25 ps ($\frac{1}{4}$ of a TTC step). As well as its use in the `Clock40Des1` calibration procedure, the CAM provides real-time monitoring

of the phase differences in TTC clocks between modules during the normal physics run; every few seconds the phase differences are calculated between modules, and any shifts greater than a predefined threshold are flagged up as an error to the run controller. This allows a fast identification of drifting clocks/TTC hardware faults during run-time.

3.3.1 CAM Test Software

Before the newly produced CAM could be fully integrated into the CP system, both the hardware components and software used to control the module need to be fully tested and proven both error-free and stable in regular use. A software based procedure was created to measure the pedestal and gain of the CAM phase comparison ADC chip. This first checks the pedestal value, by measuring the phase difference in ADC counts when the same clock is selected for both reference and variable input. In order to check the uniformity of the pedestal, the input clocks are cycled over their entire phase (240 TTC steps), see figure 3.10(a) . Next, the check of the gain is performed by comparing the phase difference (in ADC counts) between the on-board (CAM) TTC clock and the input clocks from each CPM in turn, with the `Clock40Des1` of each input CPM input signal cycled through its whole phase, see figure 3.10(b). The feature in figure 3.10(b) at `Clock40Des1`= 120 occurs when the relative phase difference is π, i.e. completely out of phase.



(a) CAM Pedestal (b) CAM Gain

Figure 3.10: (a) An example CAM pedestal result, showing the value of the phase difference (in ADC counts) for two "in-phase" clocks. (b) The ADC response to the phase difference between clocks, when one clock is cycled throught its whole phase (2π), and the reference clock kept constant.

Results of the type shown in figure 3.10 confirm the expected behaviour of the module under test, and also confirm that the software access to the VME registers on the CAM works, and that the integration of the module into the L1Calo on-line run software environment is a success. The pedestal value calculated from the pedestal scan is used in the calculation of the relative phase difference between modules, in normal units (i.e. the conversion from ADC counts to TTC steps). Accordingly, at the end of a successful scan the pedestal value is stored (after a successful validation procedure) in an appropriate `COOL` folder (see section 3.4.1) for retrieval during the normal CAM operational mode.

3.4 L1Calo Software

The L1Calo software is responsible for controlling the operation of the trigger within the ATLAS run environment[40]. The software serves many distinct and important roles; from the ad-hoc VME access of module registers and spy memories (there are about 300 VME modules with unique register and memory maps); configuration and control of calibration and monitoring tasks; and the day-to day stable running as part of an integrated ATLAS physics run. The majority of the L1Calo software framework is coded in C++, with some Java libraries in addition. The framework is split into approximately 75 individual software packages, which are grouped into about 8 categories with different internal and external dependencies, an overview of which can be seen in figure 3.11. The majority of the code written for the calibration routines described in the following sections fits into the "Calibration Procedures and Analysis" folder of figure 3.11.

3.4.1 COOL Database

There are about 50 configurable parameters for each trigger channel (7200 towers at the PPMs) in the system; these could for example have been derived in a special calibration run, or calculated from an analysis of data. These parameters must be stored and retrieved

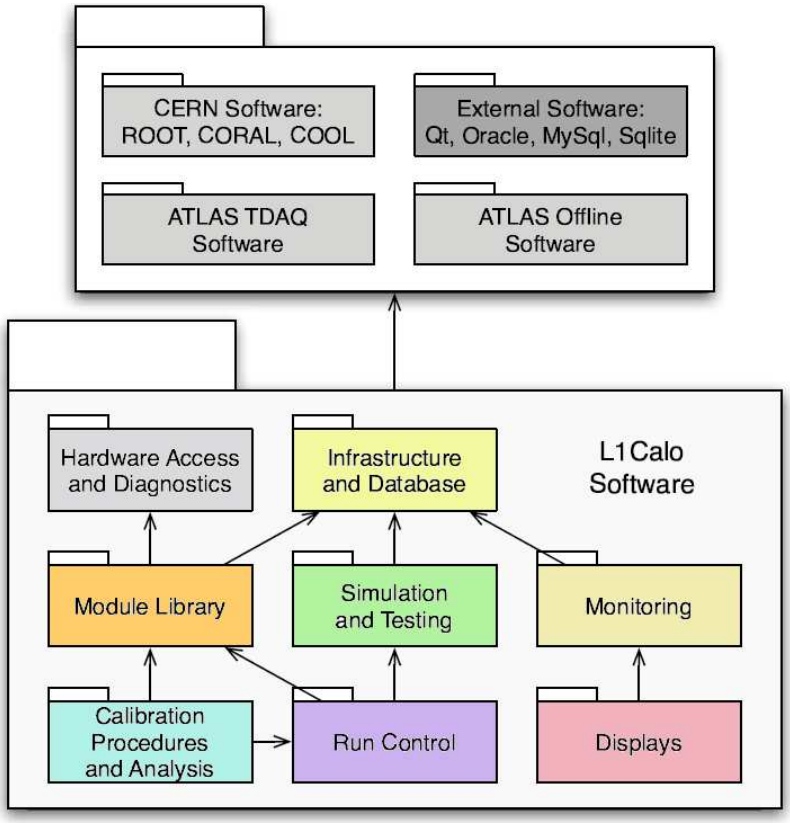


Figure 3.11: L1Calo software architecture summary[40].

by various software tasks at various times, both internally within L1Calo, and also by external ATLAS software. To facilitate this functionality, these configurable parameters are stored in the L1Calo `COOL` relational database [41]. The `COOL` database is organised into *folders*; these folders relate to different tasks and the associated parameters, such as module settings, run control parameters, and calibration scan results. Within each folder, each module/channel has a unique `COOL` channel ID; and an associated set of parameters which are stored along with a corresponding time-stamp known as the Interval of Validity (IoV). The IoV makes it possible to monitor the histories of any configurable parameter, since every time a parameter is updated, the IoV is adjusted accordingly.

3.4.2 Timing Calibration Run Software

The software calibration scans which have been developed to derive the timing constants in the CP system are based on a special software partition/mode of running, separate

from the everyday operational environment, which require only a subset of the hardware system to be available (PPr and CP crates for the `Clock40Des1` timing, and CP system only for `Clock40Des2`). A full bit-wise simulation of the data-path is used by the scans, which can be loaded/read-out at various points in the data-path through spy-memories (for the CP system, these are the spy memories at the serialiser and CP chips). The simulation is configurable, such that it can simulate various test vector data patterns which are useful for diagnostic tests of the data-path/logic.

The scans are set-up and executed from the normal run control software, which is used in the everyday running of the trigger; some parameters must be specified when setting up the run, such as which test vector pattern to use with the simulation. The user must select which partition to use, and can additionally modify the partition to exclude hardware components when necessary. Next, the partition is *configured*; this is where the settings controlling the behaviour of the scan are read from the `COOL` database, broadcast to modules, and the byte-stream simulation of the data-path computed. The program controlling all of the calibration scans is called the `CpmKicker` (this sits in the "Calibration Procedures and Analysis" part of the software framework, as shown in figure 3.11).

Once the run-type has been successfully configured, the *run* command is issued from the run controller software, and the scan commences. During a scan the `CpmKicker` can issue messages over VME, for example to sequentially increment the clock phases being calibrated over some predefined range. At the same time parity error bits are collected from the relevant spy memories, and `ROOT`[42] histograms constructed. The histograms are both sent to the On-line Histogram Presenter (OHP) for on-the-fly monitoring of the scan progress, and also saved for archival/analysis at the end of the scan. Next, the `CpmKicker` instantiates various algorithms to derive calibration settings from the parity error information. On the completion of a successful scan, useful results are written to a scan results folder (specific to the run-type) of the `COOL` database. A separate software script is used to validate the results from any scan; successfully updated results are written to the calibrated module settings folder, ready for loading onto the module in the configure

stage of a normal physics run.

3.4.3 Timing Window Algorithms

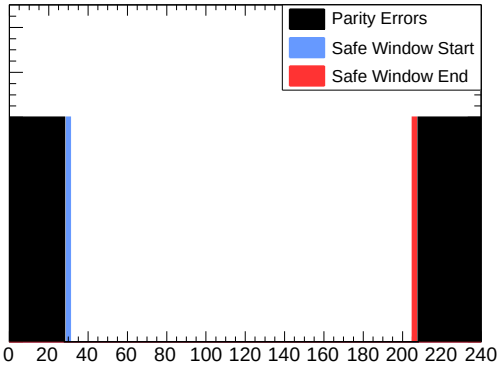
Once the parity error histograms have been collected from a timing scan (this is described in sections 3.5 and 3.6), algorithms are used to identify the openings, closings, widths and centres of the parity error free regions in each channel (Serialiser or CP chip channel). The algorithms must be able to robustly identify several types of parity error distribution; some examples of these are shown in figure 3.12. After an iterative cycle of development and testing, the algorithms correctly identify, and calculate the safe timing window parameters of all of the expected types of parity error distribution, or in the case of an erroneous channel, identify it as such.

3.5 Serialiser Timing Calibration

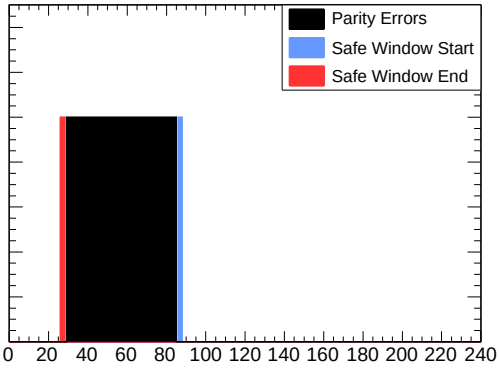
The first step in the CP system timing calibration procedure is the alignment of `Clock40Des1` across all CPMs. This is a two-fold process; firstly, the clocks are all aligned using the CAMs; next, a `Clock40Des1` timing scan is performed, and the value of the clocks which have been lined up in the first step are checked to ensure that they fall within the parity error free (safe) region. There is some circularity in the process, since an appropriate clock value must be used in the first step in order to provide a successful result from the second step; if necessary the two steps can be repeated in an iterative way to converge on safe results across the system.

3.5.1 Lining up the Clocks

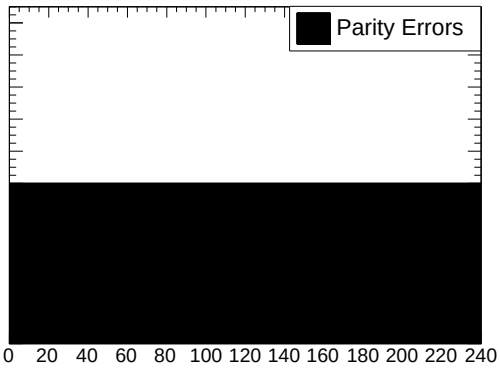
For simplicity, it was decided that the CAMs would be used in the mode where the `Clock40Des1` of each CPM in a crate, in turn, is compared to that of CPM1 (The CPM



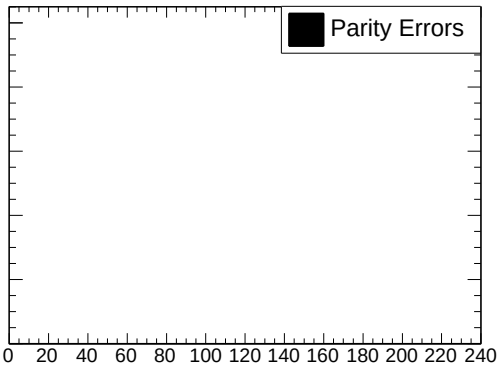
(a) Parity error free window in middle of phase range.



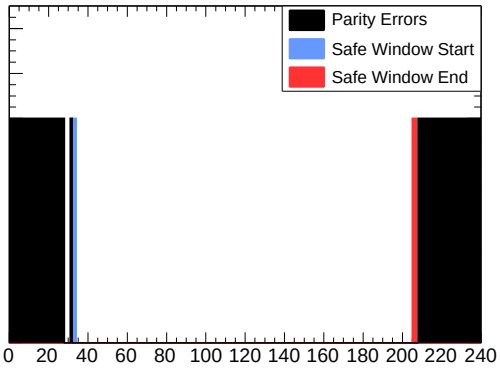
(b) Parity error region in middle of phase range.



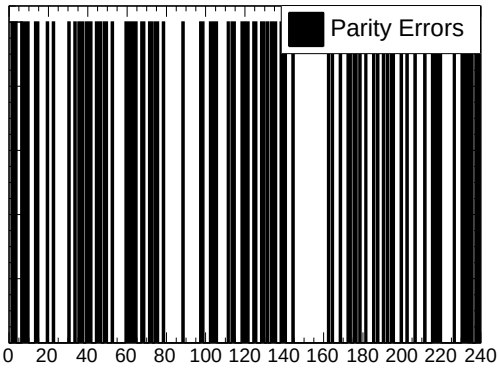
(c) Parity errors everywhere - fault



(d) No parity errors - no data



(e) Clock jitter at edge of Parity error region.



(f) Intermittent parity errors - hardware fault

Figure 3.12: Example parity error distributions.

occupying slot 5, see figure 3.3) in that crate. Accordingly, before a CAM-based scan can be performed to align the `Clock40Des1` phases within the individual crates, the `Clock40Des1` of CPM1s between the crates (the four crates are referred to as CP0, CP1, CP2, and CP3) must be aligned, such that on the completion of a successful scan, the timing is inline across the whole CP system. This task is not currently incorporated into the run control based software calibration routine, and has to be done manually using direct access to the VME timing registers on the CPMs (using the stand alone L1Calo VME access software application) in combination with an oscilloscope. Crate CP3 is chosen as the absolute point of reference; its `Clock40Des1` signal is routed via a cable from the CPM front panel connector to an oscilloscope, where it is compared in turn to the signals from CPM1 in CP0, CP1, and CP2. The `Clock40Des1` phases of CP0-CPM1, CP1-CPM1, and CP2-CPM1 are manually adjusted until the clock signals are as close as possible (aligned using the rising clock edge) on the scope. The calibrated `Clock40Des1` phases for CPM1 in each crate is then written to the (validated) module parameters folder in the `COOL` database, ready for loading onto the module at run-time, and for the CAM based scan that follows.

The next stage is to align the `Clock40Des1` phases across all CPMs in each crate. This is done using a calibration run type, the `cam-scan`. The `cam-scan` cycles through the CPMs, setting each one in turn as the *variable* clock source on the CAM, whilst keeping the *reference* clock set to CPM1. The phase offset in ADC counts is read out for each comparison, and converted to a difference in multiples of TTC step (with a precision of $\frac{1}{4}$ TTC step, using the CAM ADC pedestal value from `COOL`). Should the `Clock40Des1` from any module be greater than half a TTC step out of phase with the reference clock, a new value is derived which should bring the clock more in line with the reference. The newly derived `Clock40Des1` phases are automatically written to the appropriate calibration scan results folder in the `COOL` database. A standalone software script is then used to copy the results to the *validated* module settings folder in `COOL`, if the results are deemed appropriate (there was no error during the scan). It is the validated values which are loaded onto a module in the *configuration* stage of a run. If the `Clock40Des1` values are

found to be grossly out of time following the initial `cam-scan`, the scan can be repeated and the results validated in an iterative way until the results have converged (in practice this is not usually more than 2 scans).

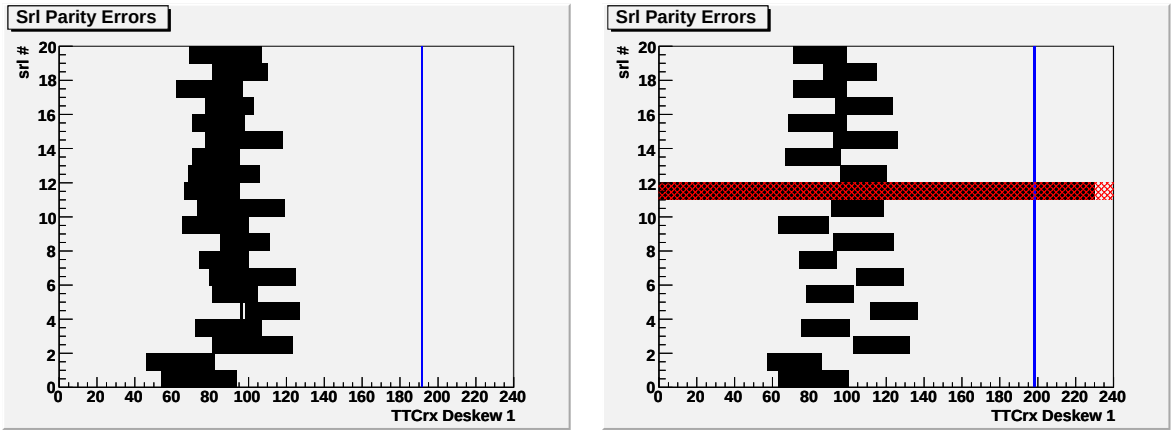
3.5.2 Clock40Des1 Parity Error Scan

Once the `Clock40Des1` phases have been calibrated/aligned across the whole CP system, it must be confirmed that the calibrated value of `Clock40Des1` for each module falls safely within the parity error free region of each serialiser's timing window; this is done using the `cpm-serialiser-scan`. The `cpm-serialiser-scan` builds parity error timing windows for all serialisers in the CP system. It does this by sequentially incrementing the `Clock40Des1` phase in each module through its whole phase (240 TTC steps), at each step collecting the parity error bits from the spy memories at the serialisers. The data used is `PpmPlayback` data; this is simulated data which is loaded into memory at the PPMs, and follows the normal data-path route, over the LVDS cables, to the CPM serialisers. This allows any timing delays due to differences in connectivity between PPMs and CPMs to be identified.

The parity error information is summarised in the form of histograms, and an algorithm finds the safe timing region (see section 3.4.3). The scan requires that the `Clock40Des1` must be more than a predefined minimum number of TTC steps clear of the edges of the parity error window. This reference value is called `minWindow`, and is configured in the `COOL` database (the current value used is 10 TTC steps). At the end of the scan, any errors are highlighted (e.g. if the `Clock40Des1` falls within an error region, or is less than `minWindow` TTC steps from the edge of a parity error region); if this happens, the whole timing process must be repeated, starting with a new value of `Clock40Des1` in CP3-CPM1. At the end of the scan, the parity error histograms (one for each module) are overlaid with the calibrated value of `Clock40Des1`, and any erroneous channels are highlighted.

3.5.3 Results

Figure 3.13 shows some typical results from a `cpm-serialiser-scan` at the module level; the error channel highlighted in figure 3.13(b) was the only error channel (1 out of 1120) identified during the first full scan of the installed system, and was subsequently found to be a hardware fault (which was later fixed). Figure 3.14 shows an example of calibrated `Clock40Des1` values derived from a system wide scan.



(a) CP0-CPM2 (b) CP1-CPM7

Figure 3.13: (a) Example result from cpm-serialiser-scan. (b) Example of an erroneous channel result.

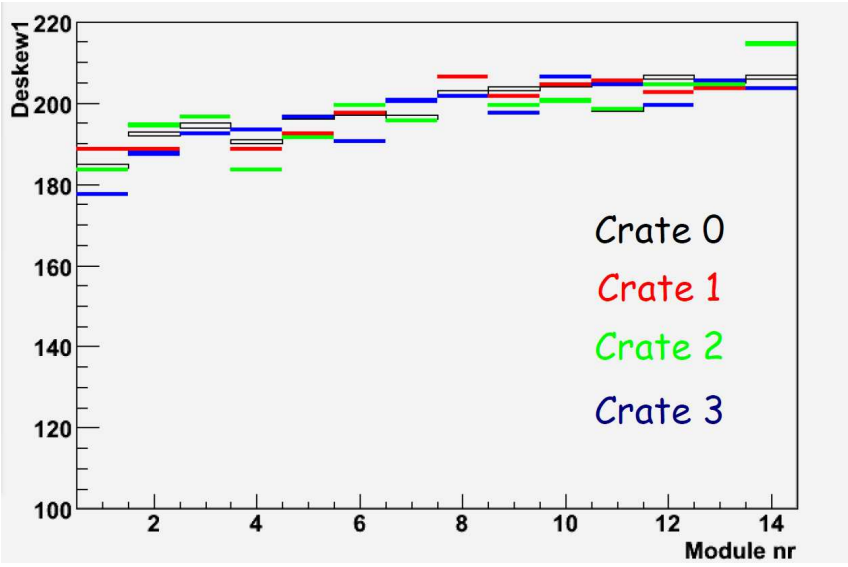


Figure 3.14: Summary of Clock40Des1 (deskew1) values derived across the CP system from a real scan.

3.6 CP Chip Timing Calibration

The CP chip timing calibration is composed of two tasks. The first task is to optimise the phases of the fan-in towers in each chip (12 phase offsets per chip, see figure 3.7), such that the safe timing window is as wide as possible for each chip. Once the optimal phases have been found, a single `Clock40Des2` phase is derived for each CPM, from the average of the centre points of each chip's safe timing window.

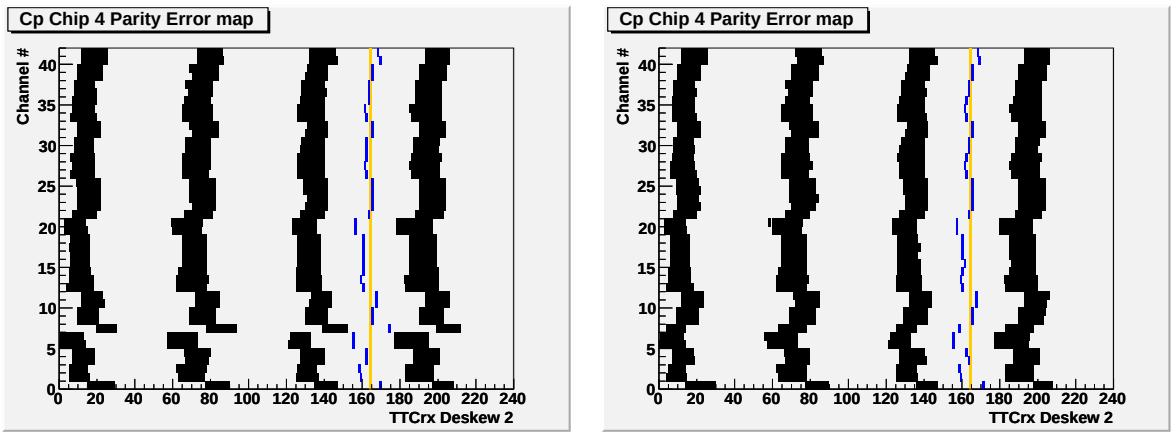
3.6.1 cpm-algorithm-scan

A special calibration run type called the `cpm-algorithm-scan` was developed to facilitate the timing calibration of the `Clock40Des2` clock, and timing phases for the CP chips. The scan uses either `PpmPlayback` or `CpmPlayback` data; the latter is simulated data loaded into the memories at the CPM's serialisers, then following the normal data-path through the CP chip logic. The scan increments the `Clock40Des2` phase through its entire range in 240 TTC steps; at each step of the scan, parity error bits are read out for the 42 trigger tower channels (see figure 3.7) in each of the 8 CP chips on a module. Once the scan has collected all the parity errors, algorithms derive the timing window parameters for each channel (window start, end, centre, and width), as well as identifying any erroneous or masked out channels.

Channels which are known to have problems are identifiable by a channel error-code, which is stored alongside the channel's calibration parameters in the `COOL` database. The scan both masks known error channels from the analysis, and writes new error codes to the `COOL` database upon the identification of new errors (in addition to flagging up the errors to the user at the end of the run).

The average safe window centre positions of the channels corresponding to the on-board towers (the 30 channels which can not have a phase shift applied) is calculated. Next, the difference between the safe window centres of the 12 adjustable channels, and

the average of the non-adjustable channels is calculated. In channels where the absolute difference is greater than 10 TTC steps, a new value of the programmable phase delay is suggested, in order to shift the parity error window more in-line with the average of the on-board (non adjustable) windows. Figure 3.15 shows an example of the parity error distribution in a single chip, taken with both original phase settings in 3.15(a), and updated phases suggested by the `cpm-algorithm-scan` in 3.15(b); note that the timing for chip channel 8 has been shifted as a result of the scan, creating a wider safe window overall.



(a) First scan (b) Second scan

Figure 3.15: (a) Example parity error histogram from cpm-algorithm-scan for a single CP chip (first scan, before new phase suggested). The orange line is the average of the centre points of the "safe windows" of the on-board (non-adjustable) towers; the blue lines are the centre positions of "safe windows" for each channel. (b) Modified result after the new phases suggested by the cpm-algorithm-scan have been applied.

After the phase calculation step, there are two possible outcomes; either new phases have been suggested for some channels, or all the phases are found to be optimal. If new phases have been suggested, these are written to the appropriate results folder in `COOL`. At this stage the scan is over, and the results stored in `COOL` must be validated and copied to the calibrated module settings folder. The scan must then be repeated; this process should be repeated until there is convergence on the phases (optimal phases), which in practice takes one or two scans at most. In the case that the result does not converge (if for example a change in phase iterates on the edge of being 10 TTC steps away from the on-board average), the next step can be executed as a separate run type (i.e. forced to

calculate the average, even if new phases are suggested by the first step).

Once the scan has found that all phases are optimal, it goes on to derive the safest value of `Clock40Des2` for each module. This is calculated from the average of the centre points of all channels in a module ($42 \times 8 = 336$ channels). The derived value of `Clock40Des2` is then written to the appropriate `COOL` results folder, where it must next be validated for use in the calibrated system. Figure 3.16 shows a typical `cpm-algorithm-scan` module level result. If there was an erroneous channel identified in the scan, or an average value of `Clock40Des2` is derived which is not in the safe window for all channels, then the histogram is overlaid with a red box to indicate such, as in figure 3.13(b).

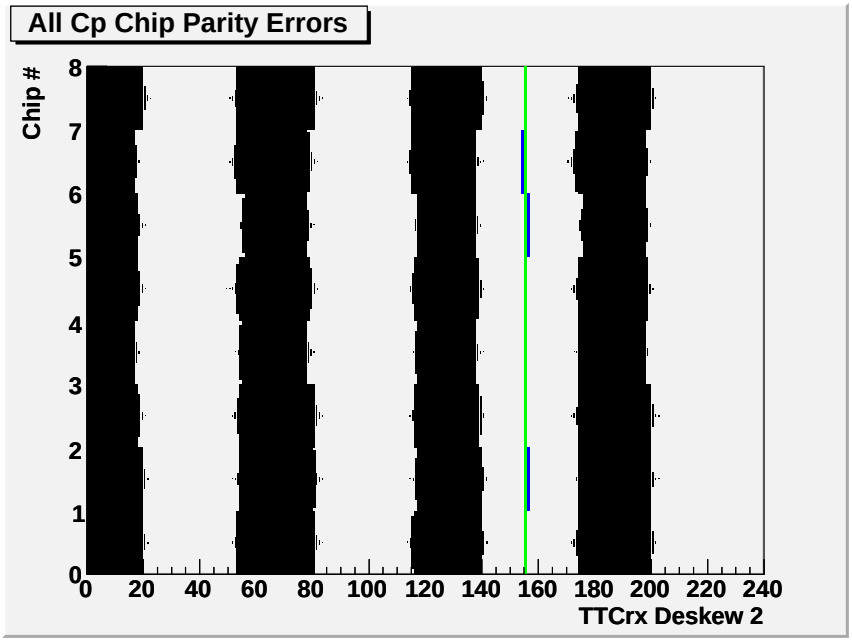


Figure 3.16: Example of Module level parity error histogram from cpm-algorithm-scan. The blue lines are the individual chip "safe window" centre points, and the green line is the average - the Clock40Des2 phase which is the final result.

3.7 Conclusions

The timing calibration methods described in the preceding sections were developed, tested, debugged, and finally deployed in the final system in the ATLAS counting room over a period of about two years. During this time, the CAM module was developed and inte-

grated into the calibration procedure. In addition, the timing scans identified a number of hardware faults which were subsequently fixed. The calibration routines have proven to be an invaluable resource, without which re-calibrating the digital timing of the CP system would be unfeasible in a practical, repeatable way. The software is used on a regular basis to provide checks on the timing calibration, and is ready to be used in the event that the system must be re-timed from scratch.

The CP chip timing calibration, which is where the timing margins are most critical, has turned out to be very successful. Following the `cpm-algorithm-scan` phase calibration, the average safe timing window at the CP chip has gone from approximately 2.0 ns wide to 3.0 ns wide; this is a significant improvement, leading to a safer, and more stable trigger.

Despite the many successes of the work carried out here, there is space for possible modifications/improvements to the calibration procedures. The most obvious of these is the option for a software controlled solution to the inter-crate timing for the `Clock40Des1` phase. This is presently done manually, but could in principle be automated. It should be noted however that this manual calibration of four channels is a small inconvenience compared to the total number of channels that would need to be manually calibrated, in the absence of the automated scans which are presented here.

Chapter 4

Particle Reconstruction and Identification

This chapter describes the standard reconstruction and identification of particles in $Z(ee)\gamma$ events in ATLAS. Signal events are identified by electrons and photons, with the main background coming from jets in $Z + jet$ events passing the photon reconstruction algorithm and selection cuts. The electron and photon candidates are formed from clusters of energy deposited in the calorimeters, in combination with tracking information from the Inner Detector. Electrons are distinguished from photons by the matching of a track to a calorimeter cluster (unconverted photons leave no track), but an ambiguity arises when photons convert to electron-positron pairs in the tracking region, since both objects possess a track matched to a cluster. In the following sections the standard cut based selection of electrons and photons employed in ATLAS, as well as the removal of overlapping particles and ambiguities (when a particle is reconstructed as electron/photon/jet candidate simultaneously) that arise in $Z\gamma$ events is described.

4.1 ATLAS Event Format

Any event which is selected by the trigger system is read out in `RAW` format, this contains all data read out from the detector, that is unique to that event. This data is then reconstructed by `ATHENA`, the ATLAS reconstruction software, where it is calibrated and summarised into Event Summary Data (`ESD`) format. This format still contains the majority of detector information, including all cells and hits, and is consequently too large to be widely distributed/replicated on the LHC Grid. The `ESD` file is therefore processed to create an Analysis Object Data (`AOD`) file; this is much smaller in size and contains lists of reconstructed physics objects, such as electrons, photons, jets, muons, tracks, and calorimeter clusters. The `AOD` contains the level of information appropriate for most physics analyses (any task which does not require access to the full set of cell, or track hits).

Simulated data from Monte Carlo event generators (see section 1.5.3) consists of lists of 4-vectors representing the final state particles of the simulated process; these must be converted into a special format (`EVGEN`) in order to be read in to the `ATHENA` framework. The response of the detector hardware to the generated particles is simulated by the `GEANT4`[43] software, the results of which are stored in a `HITS` file; a digitisation simulation is typically run together with the event reconstruction to produce the `ESD` format. The final `AOD` file is equivalent to that produced from real data, except that the simulated `AOD` contains Monte Carlo truth information, in addition.

4.2 Electron Reconstruction

There are two separate approaches to the reconstruction of electrons in ATLAS, these algorithms are referred to as `Electron` and `Softe`. The first (`Electron`) is designed for isolated, high-p_T electrons; it is seeded by calorimeter energy clusters followed by a search for a matching track in the Inner Detector. The second (`Softe`) approach is designed

for soft electrons, such as those from b decays; it is seeded by the tracks, and searches for matching calorimeter clusters. Since the $Z\gamma$ signal contains high-p_T isolated electrons (from the Z boson decay), it was required that all of the electron candidates considered in the analysis be authored by the `Electron` algorithm (either solely, or in combination with the `Softe` algorithm). In many cases, both `Electron` and `Softe` algorithms will identify the same real electron. If these two electron candidates are sufficiently close, only one electron candidate enters the electron container. The choice of algorithm used for the $Z(ee)\gamma$ analysis, along with the issue of overlap between electron candidates authored by different algorithms is dealt with in section 4.2.2.

4.2.1 Author *Electron* Algorithm

The first step in the electron reconstruction is the creation of seed clusters out of calorimeter energy deposits; energy deposits are formed into 0.025×0.025 ($\Delta\eta \times \Delta\phi$) cells. A *sliding window* algorithm [44] creates fixed size rectangular clusters of cells, located to contain the local energy maximum. For electrons, a rectangle of 3×7 cells ($\Delta\eta \times \Delta\phi$) is optimal in the second sampling layer of the barrel ECAL (see section 2.4.1); the elongation in ϕ is due to the bending of electrons in the solenoid field and Bremsstrahlung radiation of soft photons in this direction. In the end-caps, clusters are formed from 5×5 cells, since the η spread of showers increases with $|\eta|$. Next, the cluster-to-track matching is performed.

The inner detector track candidate is extrapolated to the second sampling layer in the ECAL. During the extrapolation care is taken to account for Bremsstrahlung losses[45]; if the impact position of the extrapolated track matches the cluster position to within $|\Delta\eta| < 0.05$ and $|\Delta\phi| < 0.1$, it is considered a good match. If a track does not have silicon hits, the η requirement is overlooked; this is because the resolution of the TRT η measurement is limited, so only a loose constraint on η is applied (e.g. a barrel segment track matching a cluster in the same barrel segment). It is possible for several tracks to be

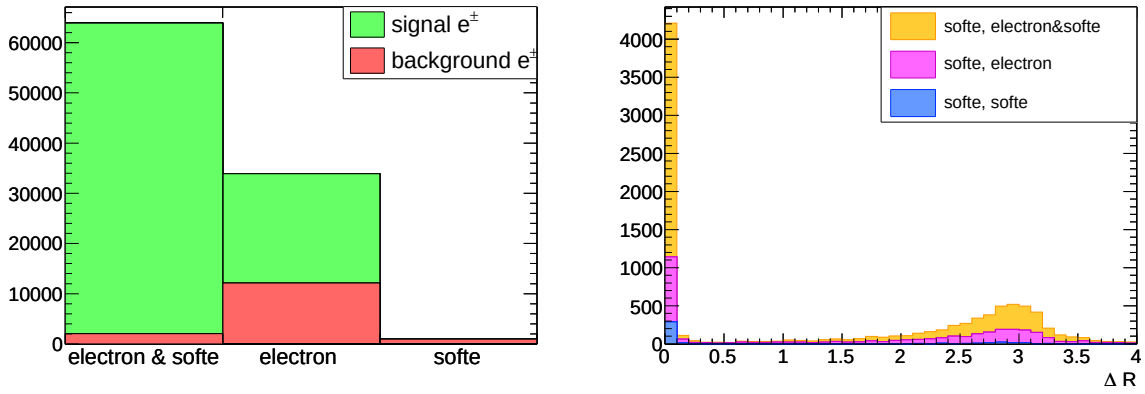
matched to the same cluster, in these cases the quality of the match can be determined by looking at the separation, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, between track and cluster; the closest match being chosen. Finally, the electron four-momentum is computed using the additional tracking information from the best-matched track. The η and ϕ directions attributed to the electron object are taken from the track parameters when silicon hits are available; candidates with no silicon hits use the cluster's η-pointing for the η direction.

4.2.2 Electron Overlap Removal

Before an event can be considered in the $Z(ee)\gamma$ analysis, it is important that any ambiguities/duplications created by the reconstruction algorithms are resolved. As mentioned above, a real electron may give rise to several overlapping reconstructed candidates (with different tracks/authors), and as such appear more than once in the electron `AOD` container. It is non-physical to keep duplicate particles in the event record, so only the most appropriate particle is kept for further analysis, and the others (authored by the least appropriate algorithm) are discarded. A simple procedure is to choose a preference of author; selecting the electrons reconstructed by the highest preference author first, and then removing any reconstructed electrons authored by a lower preference algorithm within a predefined separation cone.

As mentioned previously, electrons must have been reconstructed by the `Electron` algorithm to be considered in the analysis. Many of the good electron candidates are identified by both electron reconstruction algorithms simultaneously. Figure 4.1 shows the breakdown of electron author for signal and background electrons in Pythia $Z \to ee$ events.

A preference is chosen such that electrons within the fiducial acceptance ($p_T > 20\,\text{GeV}$, $|\eta| < 2.4$) authored by both `Electron` and `Softe` algorithms are chosen first, followed by those authored solely by the `Electron` algorithm. Accordingly, any electrons authored solely by the `Electron` author, falling within $\Delta R < 0.1$ of a `Electron` and `Softe` electron,



(a) (b)

Figure 4.1: (a) The electron reconstruction algorithm author details for signal (truth matched to a Z boson) and background (no match to a Z boson) electrons in the Pythia $Z + \gamma$ sample. (b) The separation cone ΔR between all unique pairs of electrons in Pythia $Z + \gamma$ events, of whom one of the electrons is authored solely by the "Softe" algorithm (the author combinations are described in the legend).

are removed from the study (a ΔR cone of 0.1 was found sufficient to accept electrons from Z's up to a high transverse momentum of at least 600 GeV). Following on from figure 4.1, electrons which are only authored by the `Softe` algorithm are not considered in the analysis; since they are predominantly ($\approx 90\%$) background, represent a small fraction of the total number of electrons, and as shown in figure 4.1(b), are found to overlap with those electrons authored by `Electron` and `Softe`, and `Electron` authors.

4.3 Photon Reconstruction

There are two different algorithms employed in the reconstruction of photon candidates in ATLAS; these are tailored for both *unconverted* and *converted* photons. A conversion is when a photon interacts with the material of the detector (tracker) producing an e^+e^- pair. Ideally, unconverted photons can be identified from electromagnetic energy clusters which have no matching track in the Inner Detector. The clusters are identified in a similar way to that described previously for electrons, but with a fixed 5×5 cell cluster size ($\Delta\eta\times\Delta\phi$). A selection of quantities related to the cluster are stored ready for optimisation of the selection later on. It should be noted that about 10% of unconverted photons are

mistakenly reconstructed as electrons; this is due to low-p_T tracks (typically < 2GeV) being wrongly matched to unconverted photon clusters. These photons are recovered by a dedicated procedure after the reconstruction of electrons has been completed (any electron candidates with TRT only hits on the track, and with track $p_T < 2\,\text{GeV}$, with $E/p > 10$; where E is cluster energy, and p the track momentum, is considered an unconverted photon).

The converted photons are identified by the matching of calorimeter clusters to *conversion vertex candidates* reconstructed by the Inner Detector software. There is an inherent overlap between the converted photon candidates and the reconstructed electron candidates, the treatment of which has evolved along with the development of ATLAS software. The procedure for the recovery of converted photons presented here is the version employed with the software analysing early data from ATLAS (release 15).

4.3.1 Converted Photon Recovery

The conversion vertex candidates identified by the Inner Detector software are one of two types; those which have two inner detector tracks matched to a vertex, and those with only one track. In the case of two matched tracks, the vertex is reconstructed using a constrained fit with the two sets of track parameters, with the assumption of a massless photon, allowing the four-momentum of the original converted photon to be calculated. The vertex candidates with only one track are most likely due to soft (one of the electrons having $p_T < 0.5\,\text{GeV}$), or co-linear (symmetric high p_T photons where the tracks are too close together to resolve as separate) electrons emerging from the conversion. In these cases no constrained vertex fit can be performed, and the location of the vertex is taken as the first space-point of the track; consequently for these candidates the original photon four-momentum can not be computed.

The next step in the recovery of converted photons is the matching of conversion vertex candidates to calorimeter energy clusters. For conversion vertex candidates with two

tracks, where the two track momenta are similar (within a factor of 4), both tracks are individually extrapolated to the second sampling layer of the ECAL; if the extrapolated points are both within 0.05×0.1 ($\Delta\eta \times \Delta\phi$) of the centre of the cluster, the conversion vertex candidate is considered a good match. For double-track conversion vertex candidates where one of the tracks has a much smaller momenta than the other (more than a factor of 4), the four-momentum of the original photon (calculated from the constrained fit of the track parameters described above) is used to extrapolate the photon candidates trajectory to the ECAL, where the match can be tested. For single-track conversion vertex candidates, the track measurement is extrapolated to the second sampling layer of the ECAL.

At this stage all candidates with a track (single or double track conversion vertex candidate) matched to a cluster are considered to be *electron* objects; this can include more than one conversion vertex candidate for the same cluster (electron object), if more than one conversion vertex candidate has been matched to it (this is likely in electromagnetic showers where several conversion vertices exist in a small $\Delta\eta - \Delta\phi$ region). These multiple matches are sorted into an order of preference based on quality, such that double-track candidates have precedence over single track candidates; if there are several double-track candidates, the one at the smallest radius of separation is considered the best.

The next step in the procedure recovers the converted photon objects from the collection of electrons. For electrons with a conversion vertex candidate (one, or several), the track best matched to the cluster (closest in $\Delta\eta \times \Delta\phi$ from the cluster centre) is compared to the track(s) belonging to the best conversion vertex candidate for that cluster. If the two tracks coincide, then the electron is considered as a converted photon. If the track does not match any of those belonging to the conversion vertex candidate, but the track p_T is smaller than that of the candidate converted photon, it is also considered as a converted photon; otherwise it is just considered as an electron. Finally, for electrons without assigned conversion vertex candidates, if the track is made only of TRT hits and has a momentum $p_T > 2\,\text{GeV}$, with $E/p < 10$, it is considered a converted photon

(even though there is no conversion vertex); here the first track-point is assigned as a single-track conversion vertex.

The recovered converted photon objects are then rebuilt using a 3×7 ($\Delta\eta \times \Delta\phi$) cell window, and a dedicated converted photon calibration is applied (since converted photons appear differently from unconverted photons in the calorimeter, e.g. lateral shower width); various parameters belonging to the object are stored for later use in the cut based selection procedure. The reconstructed tracks are re-fitted under the electron assumption to correct for Bremsstrahlung losses, and the constrained vertex fit is repeated; this results in a more accurate calculation of the original photon four-momentum. For converted photons with TRT only tracks, the photon four-momentum is calculated using the η-pointing of the calorimeter cluster, as in the case of unconverted photons.

The final collection of photons then contains all of the converted and unconverted photons in one place. They are distinguished by the assignment of an author; `Photon` for unconverted, and `RConv` for converted photons. Figure 4.2 depicts the radial distributions of simulated reconstructed converted photon conversion vertex positions in the Pythia signal $Z + \gamma$ sample, coloured by detector subsystem; figure 4.3 shows the spatial coordinates of conversion vertices in the pixel detector.

4.3.2 Conversion Overlaps with Electrons in $Z\gamma$ Events

At the end of the converted photon recovery procedure, the reconstructed electron objects from which recoveries were made are still present in the collection of particles in the `AOD`, these overlap with each converted photon recovered. It is non-physical to keep these electrons in the analysis, therefore the first analysis step requires that electrons closer than $\Delta R < 0.1$ to a converted photon (this photon is required to pass the fiducial acceptance requirements for signal photons; $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$) are removed from the analysis. This procedure removes the dependence on the order in which electrons and photons are later analysed; without this removal procedure, and if electrons are selected

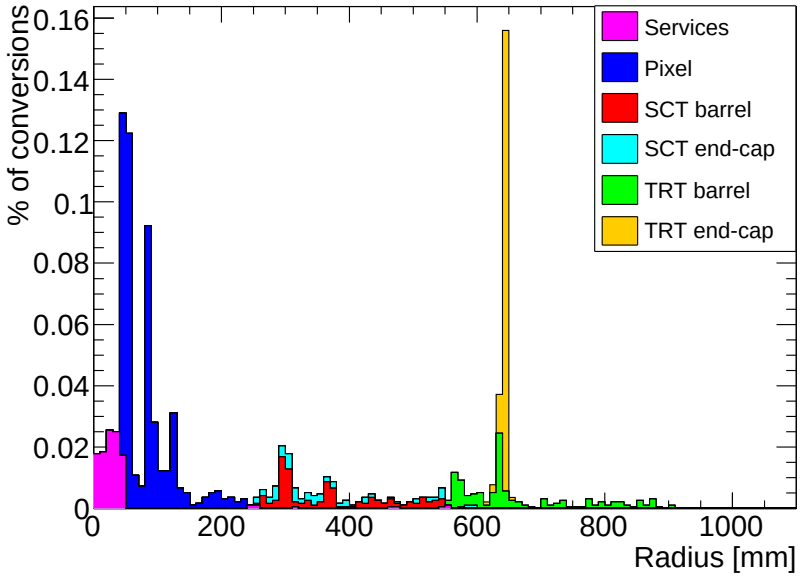


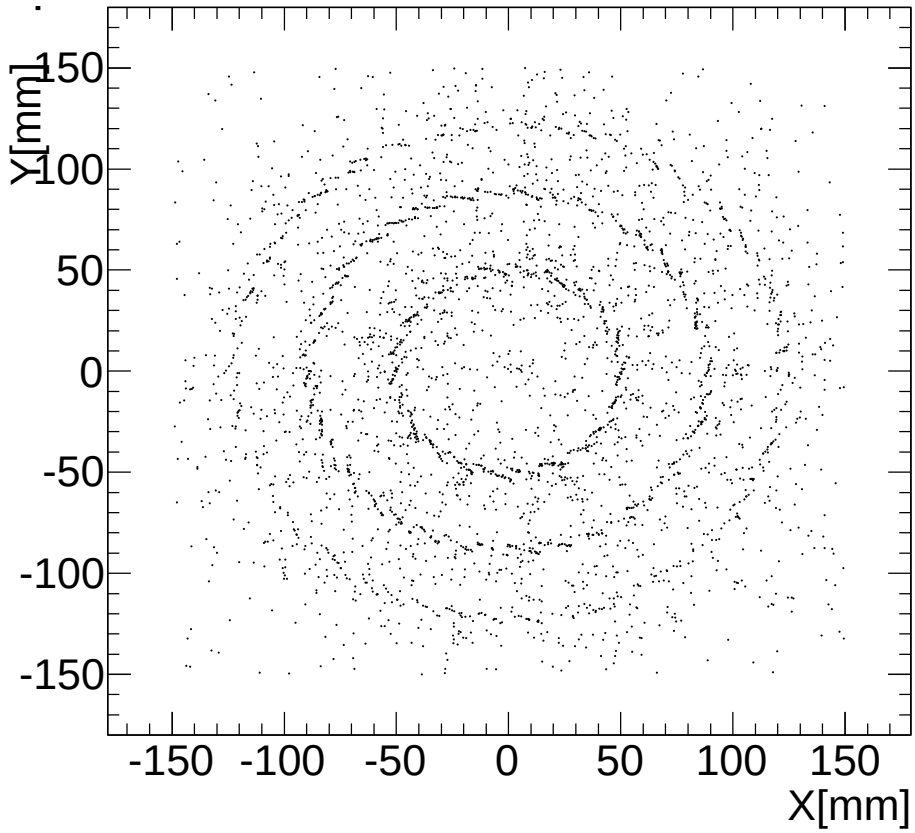
Figure 4.2: Radius of conversion vertex for converted signal photons, identified by Inner Detector subsystem, normalised to unity.

before photons in the analysis, it is possible that the photon efficiency would suffer, because reconstructed electrons and photons are required by the analysis to be separated by $\Delta R(e,\gamma) > 0.4$ (see section 5.4.5 for details).

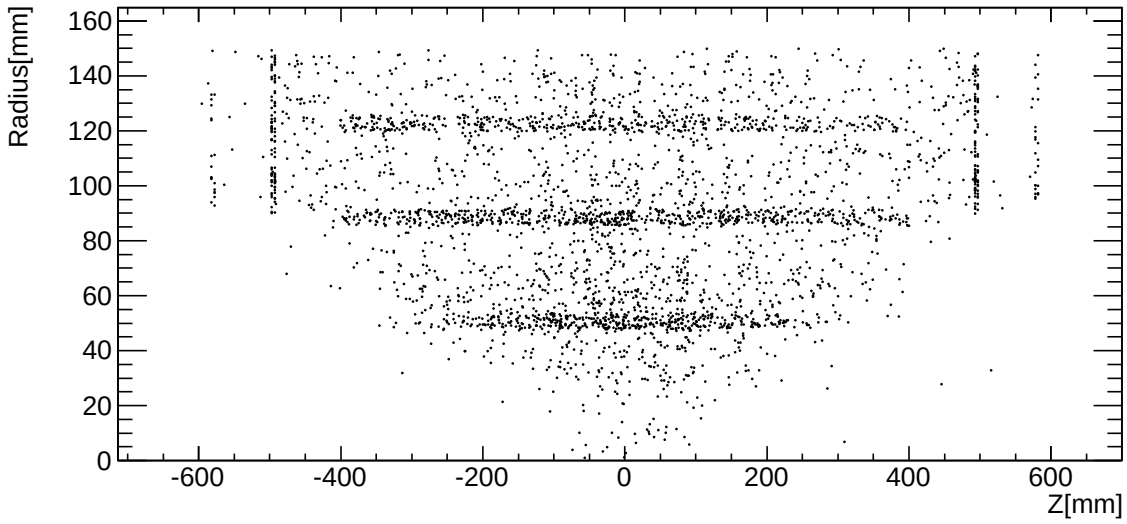
4.4 Cut based selection of Electrons and Photons

The largest background to isolated electrons and photons in ATLAS is in the form of QCD jets that pass the electron/photon reconstruction algorithms; these are referred to as *fake* electrons/photons. In order to reduce the QCD background, the application of a set of rectangular[1] selection cuts based on calorimetry and tracking information which discriminate between signal and background has been developed [46]. These selection cuts are collectively referred to in the jargon as `isEm`, and for electrons have been optimised to provide `Loose`, `Medium`, and `Tight` reference levels of selection (with, or without isolation criteria); each level applying a stricter subset of cuts than the previous, to further reduce

[1]These cuts are rectangular (a straight-forward binary cut on each variable) for any given particle; although the values of these cuts are derived using a multivariate technique.



(a)



(b)

Figure 4.3: (a) Conversion vertex positions in the X-Y plane (pixel detector only).(b) Conversion vertex positions in the Z-R plane (pixel detector only); the material structure of the pixel layers can clearly be seen.

the background (at the cost of a loss in signal reconstruction efficiency). For photons, only `Loose` and `Tight` reference selections are defined.

4.4.1 Electron Identification Cuts

The `Loose` electron selection is based on shower shape variables from the second sampling layer of the ECAL, along with hadronic leakage information. At the `Medium` level, in addition to the `Loose` cuts, information from the strip layer (first layer of ECAL) is included, as well as track quality and track-cluster matching information. The `Tight` selection goes further, building on the `Medium` cuts, utilising the identification capabilities of the TRT, as well as adding in the pixel b-layer, and E/p requirements. Table 4.1 shows the variables which are used in the electron `isEm` selection for the `Loose`, `Medium`, and `Tight` levels. The values of the cuts on the variables are optimised using a multivariate analysis of signal to background (maximising the signal efficiency for a given level of jet rejection); this is performed in a number of distinct cluster E_T, and η regions [46].

4.4.2 Electron Reconstruction Efficiency

Previous studies have indicated that for the identification of Z bosons from their decay to electrons, the `isEm` selection `MediumIso` provides good background rejection against QCD fakes[46]. The definition of `MediumIso` is the same as for the `Medium` definition presented in table 4.1, with the addition of calorimeter isolation requirements (the details of the isolation criteria can be found in [47]). The choice of `MediumIso` has in addition to the referenced study been independantly predicted to provide a sufficient rejection of QCD fakes for the $Z\gamma$ analysis (see section 5.3.6).

As a cross-check with other studies the reconstruction efficiency for signal electrons has been calculated for `Loose`, `MediumIso`, and `Tight` electron selections, for electrons from the $Z \to ee$ decay in the $Z\gamma$ signal sample (see section 5.2.1 for a description of the signal

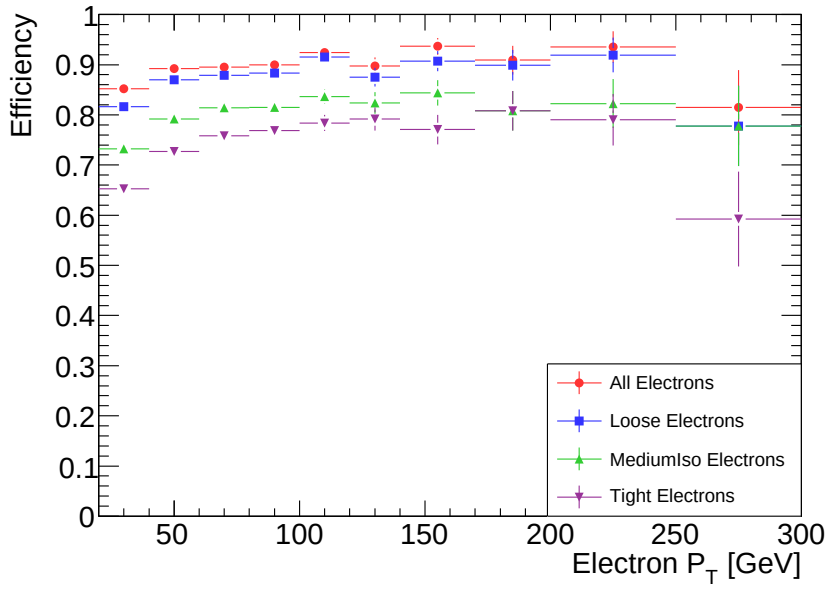
Type	Description	Formula
	Loose	
Acceptance	$\|\eta\| < 2.47$	
ECAL 2^{nd} layer	Ratio of energy in 3×7 to 7×7 ($\eta \times \phi$) cells in 2^{nd} ECAL layer.	$R_\eta = \frac{E_{3\times7}}{E_{7\times7}}$
	Lateral width of shower using all cells in a 5×5 window in 2^{nd} ECAL layer.	$w_2 = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$
HCAL Leakage	Ratio of E_T in 1^{st} layer of HCAL, to that of the whole cluster.	$R_{had_1} = \frac{E_T^{had_1}}{E_T}$
	Ratio of E_T in entire HCAL, to that of the whole cluster.	$R_{had} = \frac{E_T^{had}}{E_T}$
	Medium	
ECAL 1^{st} layer	Total shower width using approx 20×2 strip cells ($0.0625\,\Delta\eta \times 0.2\,\Delta\phi$), where i_{max} is strip containing local energy maximum.	$w_{stot} = \sqrt{\frac{\sum E_i (i - i_{max})^2}{\sum E_i}}$
	Ratio of energy difference between 1^{st} and 2^{nd} largest strip energy deposit, over the sum of these energies.	$E_{ratio} = \frac{E_{1st_{\max}} - E_{2nd_{\max}}}{E_{1st_{\max}} + E_{2nd_{\max}}}$
Track quality	Number of pixel hits > 0.	
	Combined number of pixel + SCT hits > 6.	
	Transverse impact parameter < 5mm.	
Track matching	$\Delta\eta < 0.01$ between cluster and track.	
	Tight	
b-layer	Number of B-layer hits > 0.	
Track matching	$\Delta\phi < 0.02$ between cluster and track.	
	$\Delta\eta < 0.005$ between cluster and track.	
	Ratio of cluster energy to track momentum.	E/p
	Transverse impact parameter < 1mm.	
TRT	Number of TRT hits.	
	Ratio of high threshold (pulse height) to normal TRT hits.	

Table 4.1: Variables/cuts used in the isEm selection of Loose, Medium, and Tight electrons.

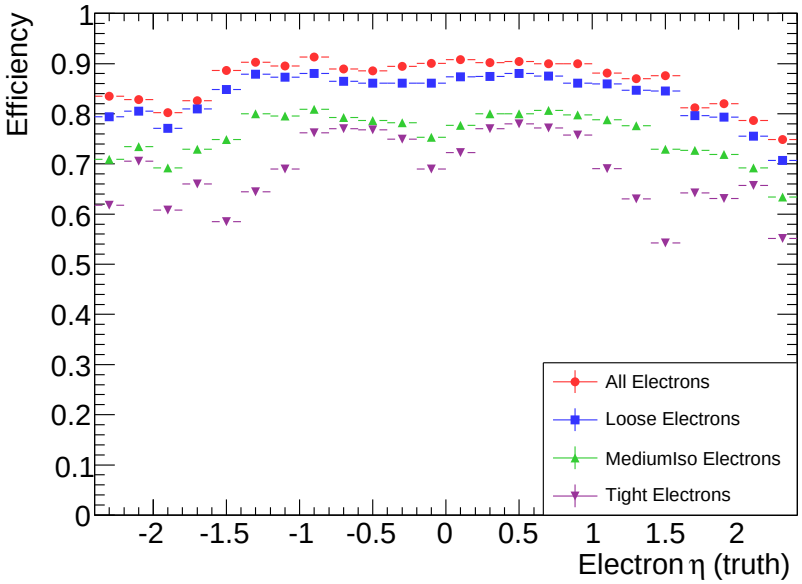
Monte Carlo data sample). The electron reconstruction efficiency is defined as the ratio of signal electrons which have passed selection cuts and been correctly reconstructed, over the total number of signal electrons generated (MC truth) in each bin of the histogram. In order to identify the origin of reconstructed electron candidates, a truth match is performed as described in section 5.2.3, and the Monte Carlo decay tree of the matched truth particle is checked to decide whether or not it is an electron from the $Z \rightarrow ee$ decay. The resulting distributions are shown in figure 4.4, where both the reconstructed and truth electrons are required to pass the fiducial acceptance criteria $p_T > 20\,\text{GeV}$, and $|\eta| < 2.4$; it should be noted that reconstructed electrons overlapping with converted photons have been discarded as per the procedure described in section 4.3.2. The average efficiencies found here ($\sim 85\%$ for `Loose`, $\sim 75\%$ for `MediumIso`, and $\sim 65\%$ for `Tight`) are in agreement with those reported elsewhere [48].

4.4.3 Photon isEM Cuts

The photon `isEm` selection consists of two reference sets of cuts, namely `Loose` and `Tight`. The `Loose` selection shares a common set of variable cuts to that for electrons (`Loose`), and is primarily used for triggering purposes, providing an excellent efficiency for a reasonable fake rate[46]. The `Tight` photon requirement is optimised to provide good rejection against the background from π^0's decaying to photons; it specifically uses an enhanced set of strip layer cuts, which due to the high granularity in this layer of the ECAL offers good γ-π^0 separation. This enforces a stricter pseudo-rapidity acceptance requirement, since the strip-layer only covers $|\eta| < 2.37$ with the crack region ($1.37 < |\eta| < 1.52$) excluded. The selection cuts are optimised separately for converted and unconverted photons, using a multivariate analysis to balance signal reconstruction efficiency against background rejection. Table 4.2 shows the variables which are used in the photon `isEm` selection procedure.



(a)



(b)

Figure 4.4: (a) Electron reconstruction efficiency for default (all) reconstructed electrons, and Loose, MediumIso and Tight electron selections; for electrons from $Z \rightarrow ee$ decay in $Z\gamma$ events, binned in p_T of truth electron. (b) binned in η of truth electron.

Type	Description	Formula
	Loose	
HCAL leakage	Ratio of energy in 1^{st} HCAL layer to cluster energy, R_{had_1}.	(see table 4.1)
	Ratio of energy in entire HCAL to cluster energy, R_{had}.	(see table 4.1)
ECAL 2^{nd} layer	Ratio in η of cell energies in 3×7 to 7×7 window, R_η.	(see table 4.1)
	Lateral shower width, w_2.	(see table 4.1)
	Tight	
Acceptance	$\lvert\eta\rvert < 2.37$ and exclusion of crack region $1.37 < \lvert\eta\rvert < 1.52$.	
ECAL 2^{nd} layer	Ratio of energy in 3×3 to 3×7 ($\eta \times \phi$) cell windows.	$R_\phi = \frac{E_{3\times3}}{E_{3\times7}}$
ECAL 1^{st} layer	Total lateral shower width, w_{stot}.	(see table 4.1)
	Shower width for three strips around maximum	$w_{s3} = \sqrt{\frac{\sum E_i (i-i_{max})^2}{\sum E_i}}$
	Fraction of energy within 7 strips, but outside central 3.	$F_{side} = \frac{E(\pm3)-E(\pm1)}{E(\pm1)}$
	Energy difference between strip of lowest energy between leading two strips, and second largest strip.	$\Delta E = E_{2^{nd}\max} - E_{\min}$
	Difference between two largest energy strips, divided by the sum of these energies E_{ratio}.	(see table 4.1)

Table 4.2: Variables/cuts used in the isEm selection of Loose and Tight photons.

4.4.4 Photon Reconstruction Efficiency

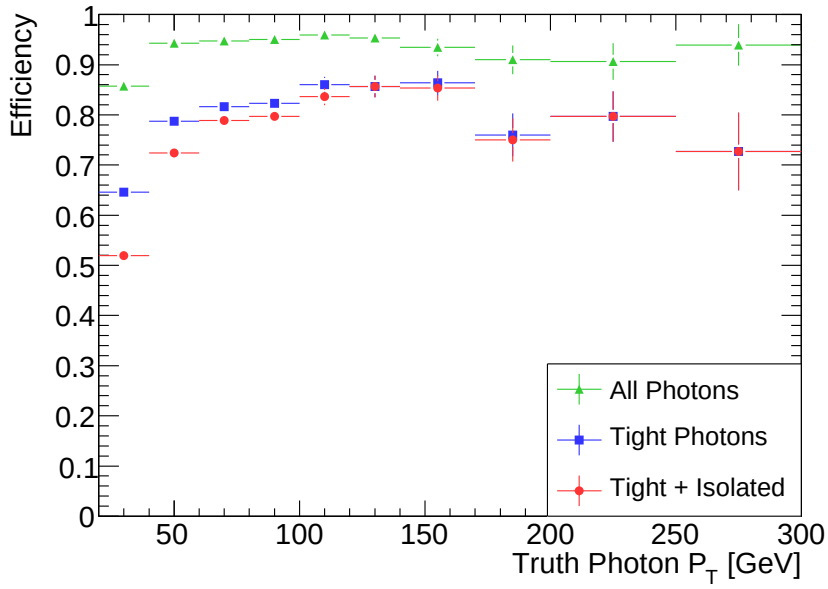
In the analysis of photons in $Z\gamma$ events, the photon `isEm` selection `Tight` is used; the motivation for this choice is the rejection power against the QCD background (jets in Z events being reconstructed as photons), as presented in section 5.4.5. The reconstruction efficiency for signal photons passing the `isEm Tight` selection in $Z\gamma$ events, without any constraints on electrons in the event, has been calculated to provide a point of comparison to other photon studies[48].

The efficiency is defined in each bin as the ratio of well reconstructed photons passing the selection criteria, to the total number of photons generated in that bin (MC truth photons). In order to assess whether or not a reconstructed photon is a signal photon a proximity truth matching has been performed ($\Delta R < 0.1$), and the Monte Carlo decay tree has been tested to ensure the photon is a signal photon (and not a background photon e.g. from $\pi^0 \rightarrow \gamma\gamma$ decay in hadronic jets). Figure 4.5 shows the reconstruction efficiency for all reconstructed photons (without any `isEm` requirement), `Tight` photons, and `Tight` isolated photons. The isolation criteria [2] is applied to further suppress the QCD background, see section 5.4.5 for details. A marked drop in the reconstruction efficiency can be seen in figure 4.5(b), for `Tight` and isolated `Tight` photons, in the region about $|\eta| \approx 0.5$; this is expected, as the region $1.37 < |\eta| < 1.52$ is exluded as part of the `Tight` cut.

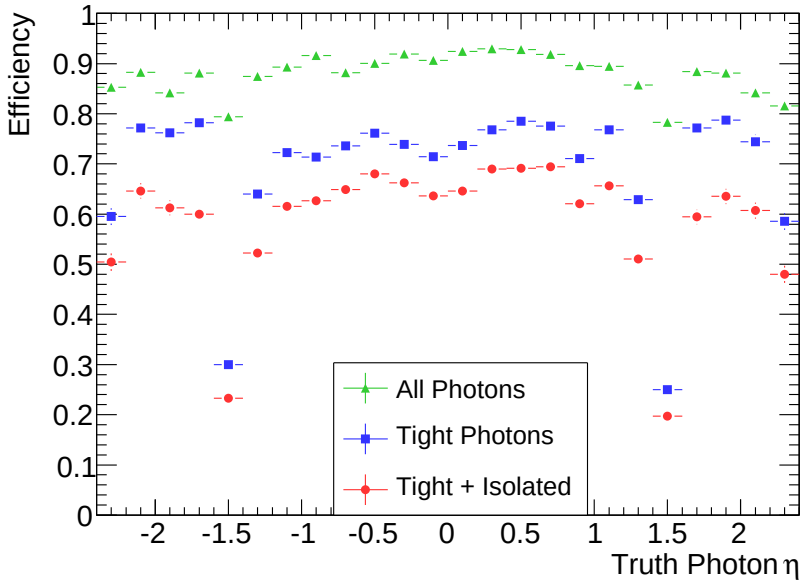
4.5 Jet Reconstruction

In order to parameterise the QCD background contribution to the signal yield, a method of jet reconstruction is chosen. The study of hadronic jets is an important aspect of most physics analysis in ATLAS; the QCD multi jet production cross-section is very large (see section 5.3), and jets from fragmentation are produced in abundance. The goal of a jet

[2] $\frac{E_{cone,0.4}-E_{cluster}}{E_{cluster}} < 0.1$, Energy sum in a cone of opening 0.4 about the cluster centre, minus the cluster energy, divided by the cluster transverse energy, must be less than 0.1



(a)



(b)

Figure 4.5: (a) Photon reconstruction efficiency for all reconstructed photons, Tight, and Isolated Tight photon selections; for ISR photons in $Z\gamma$ events, binned in p_T of truth electron. (b) binned in η of truth photon.

reconstruction algorithm is to capture the energy deposited by the cascade of hadrons and interpret the four-momentum of the jet; this is algorithm-dependent, since there is no absolute definition of a jet possible. The basic principle of a jet reconstruction algorithm is to cluster together energy deposits in the HCAL until it is decided that the jet is well contained, followed by an energy calibration to account for detector effects.

Since the analysis is not dependent on the jet reconstruction, the choice of jet algorithm is not important here. For the purposes of this study, the chosen algorithm is called `Cone4H1TopoJet` in the jargon, which is a fixed cone algorithm of cone radius 0.4; the details of which can be found in the reference [48].

Chapter 5

Measurements of $Z\gamma$ Production

This chapter describes the selection of $Z + \gamma$ events where the $Z \to e^+e^-$, and the subsequent measurement of the photon differential cross-section, relative to the transverse momentum (p_T) of the photon, or three body $ee\gamma$ invariant mass. The goal is to see how well one could measure this spectrum using ATLAS, with $\mathcal{L} = 1\,\text{fb}^{-1}$ integrated luminosity at $\sqrt{s} = 7\,\text{TeV}$. In addition the $ee\gamma$ invariant mass distribution is used to estimate 95% Confidence Level exclusion limits on the cross-section of a generic new particle (scalar or vector) decay to $Z\gamma$.

5.1 Analysis Method

The photon differential cross-section, binned in the photon transverse momentum, $p_T(\gamma)$, or in the three-body $ee\gamma$ invariant mass, $M_{ee\gamma}$, is calculated from data using the following equation:

$$\left(\frac{d\sigma_{(pp\to Z+\gamma\to e^+e^-\gamma)}}{dX}\right)_i = \frac{N_i}{\mathcal{L}\,\varepsilon_i\,(\Delta X_i)} \times \mathcal{P}_i \tag{5.1}$$

Here X is the variable being measured (either $p_T(\gamma)$ or $M_{ee\gamma}$), N_i is the total number of events recorded in the i-th X bin, passing all the event selection criteria; $\mathcal{L}$ is the

integrated luminosity, ε_i is the signal reconstruction efficiency in the i-th bin (including effects of reconstruction, selection cuts, and trigger), ΔX_i is the width of the i-th bin, and $\mathcal{P}_i$ is the data purity[1] calculated after all of the selection criteria have been applied. The use of bin-by-bin correction factors is discussed in section 5.4.7, where bin migrations are found to be small enough to validate the use of this method, in place of a more complex unfolding strategy.

Monte Carlo simulations are used throughout this analysis in order to understand and prepare for important effects in the analysis, such that a robust analysis is ready to be applied to real data, once sufficient data is collected. To simulate the presence of real data, N_i in equation 5.1 is derived from a pseudo-data sample; a Monte Carlo sample containing a combination of the signal and background samples, with the correct fractions of events admixed according to the individual cross-sections of the constituent samples, and the integrated luminosity chosen for the analysis. The efficiencies and purities used to correct the measured distribution back to the underlying cross-section (the Monte Carlo generator *truth level* cross-section) are all calculated from a statistically independent subset of the samples used to derive N_i. The calculation of the correction factors is described in the following sections. The motivation for the choice of binning employed in the measurements is presented in section 5.4.7.

In order to compare between different theoretical models or predictions for the cross-section measurements, the measured differential cross-section may have to be further corrected, to account for any differences in fiducial acceptance applied in the theoretical model. The differential cross-sections derived in this study are subject to the fiducial constraints of $p_T(e) > 20\,\text{GeV}$, $|\eta(e)| < 2.4$, $p_T(\gamma) > 20\,\text{GeV}$, $|\eta(\gamma)| < 2.4$, and $\Delta R(e,\gamma) > 0.4$. If the theoretical distribution is calculated for a more inclusive fiducial acceptance than that measured here, a further correction factor must be derived.

[1]The use of a purity factor is only valid for a simulation study; with real data, a method using $N - B$, where B is the expected number of background events, would be adopted.

5.2 Monte Carlo Data Samples

An integrated luminosity of $1\,\mathrm{fb}^{-1}$ is the amount of data that ATLAS can be expected to collect in its first run period ($2010 - 2011$). The analysis uses data samples from a large-scale Monte Carlo production, simulated with a 7 TeV proton-proton centre of mass collision energy. Unless stated otherwise, the samples have been produced by the ATLAS production team. These samples are produced in accordance with the `MC09` tune[49]; where the internal settings and simulation options of the Monte Carlo generators have been optimised with inputs from recent experiments, a key ingredient being minimum-bias and underlying event data from the Tevatron experiment.

5.2.1 Main Signal and Background Samples

Signal - ISR

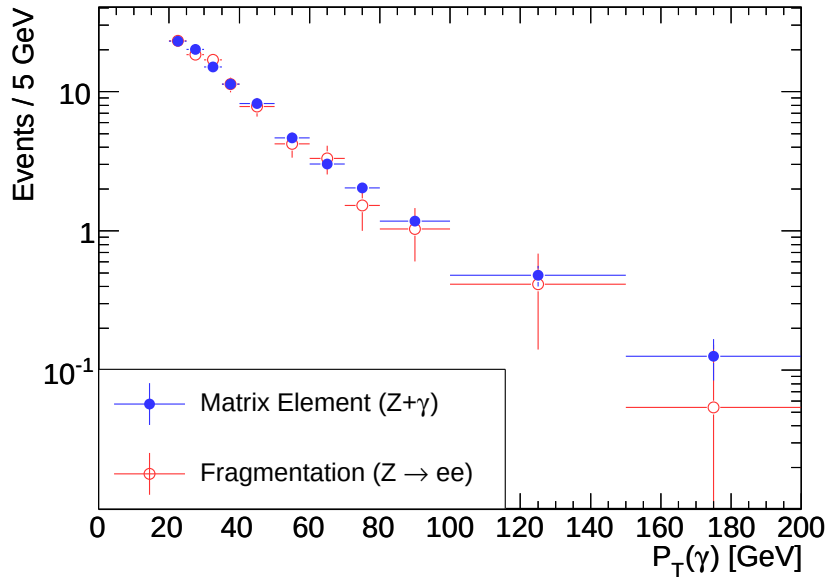
The signal (ISR) Monte Carlo sample is produced with the Pythia event generator[14], using a $Z + \gamma$ LO matrix element (figure 1.3(a)), where the Z is forced to decay to electrons. Pythia's internal routines for simulating FSR are switched off, in favour of the dedicated sub-program `PHOTOS`[50]. In addition to the standard Pythia parameters (`MC09` tune), a minimum transverse momentum is set on the hard process of 10 GeV. A generator level photon filter is applied which accepts events likely to contain at least one photon with $p_T > 15$GeV, and $|\eta| < 2.7$, in combination with a multi-lepton filter which accepts events containing at least two leptons with $p_T > 15$GeV, and $|\eta| < 2.7$. About 25% of the Pythia generated events pass both filters and are subsequently passed through the full ATLAS detector simulation chain; this filtering improves the use of available computer resources. Each event in this sample will contain a hard photon originating from the matrix element, referred to as Initial State Radiation (ISR), and a Z boson. Additional photons may be radiated from the leptons of the Z decay (FSR). In events containing more than one photon within the acceptance, the highest p_T photon is considered as the

signal photon.

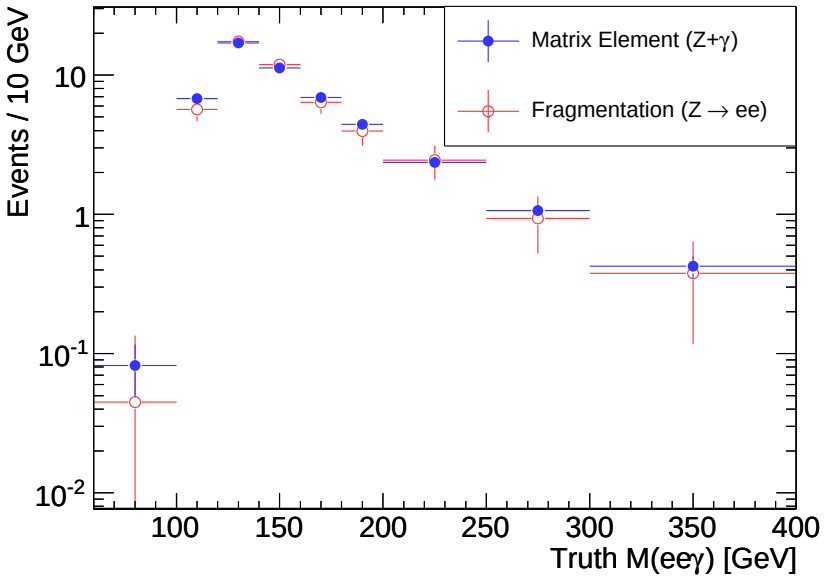
Background - FSR

The main source of background to the ISR signal comes from the FSR process shown in figure 1.3(b), which is simulated by the Pythia inclusive $Z \rightarrow ee$ process, where no photon is specified in the matrix element. There is an invariant mass constraint of $M_Z > 60\,\text{GeV}$ imposed in the generator, and the Z is forced to decay to electrons. These events contain at least two electrons, with some events containing photons coming from many possible sources. These include FSR photons (these photons are simulated with the `PHOTOS` sub-program), photons from hadron decays (e.g. $\pi^0 \rightarrow \gamma\gamma$), fake photons (parts of jets mistakenly reconstructed as photons), and also ISR photons which are simulated by Pythia's fragmentation model.

The ISR photons in the $Z \rightarrow ee$ inclusive sample may be considered as good signal photons; figure 5.1 shows a comparison between analysed distributions of ISR photons coming from the signal ($Z + \gamma$, photon defined in matrix element) and background ($Z \rightarrow ee$, photon originates from fragmentation routine) samples. It can be seen that the Pythia fragmentation model provides a consistent description to the full matrix element calculation in this kinematic range, however the statistics in the signal ($Z + \gamma$) sample are considerably better. For this reason the $Z + \gamma$ sample is hereafter used, and referred to, as the signal sample. The inclusive $Z \rightarrow ee$ sample is used as the main source of background photons (FSR). It is assumed safe to use these two samples in conjunction, as long as any double counting of ISR photons between the two samples is avoided. Accordingly, all ISR photons originating from Pythia's fragmentation model in the inclusive $Z \rightarrow ee$ are hereafter removed from the study. It should be noted that the reason the photons in the inclusive sample can be separated unambiguously into FSR and ISR is that Pythia does not calculate the interference term between the two diagrams. Table 5.1 shows the number of events, cross-sections, and corresponding integrated luminosities of the signal and main background samples.



(a)



(b)

Figure 5.1: (a) Comparison of photon p_T distributions for matrix element ISR photons in the signal ($Z + \gamma$) sample, and fragmentation ISR photons from the background ($Z \rightarrow ee$) sample, for events passing the signal selection criteria (see section 5.4), normalised to an integrated luminosity of $\mathcal{L} = 1fb^{-1}$. (b) Similar comparison for $ee\gamma$ invariant mass distribution.

Data-set	Process	Generator	Cross Section (pb)	Events	$\int \mathcal{L} dt$ (fb^{-1})
105168	$Z(ee) + \gamma$	Pythia	0.69	49,896	72.3
106046	$Z \rightarrow e^+e^-$	Pythia	855.8	4,758,621	5.6

Table 5.1: Data-set number and cross-sections (corrected for filter efficiency where appropriate) for signal and main background samples, along with the number of available events and equivalent integrated luminosity.

5.2.2 Additional Background Processes

Any other processes from which two electrons and a photon are reconstructed within the specified fiducial acceptance, and which pass the selection cuts, may contaminate the signal distribution beyond the FSR contribution discussed above (entering through $\mathcal{P}_i$ in equation 5.1). Naturally, the choice of selection cuts is motivated to preferentially accept the signal events, and reject the background events. There are a number of further physics channels which could potentially contribute to the background; these include events containing two real electrons and a real photon, such as di-boson (ZZ, WZ, WW), and $t\bar{t}$ events containing radiative photons. It is also possible that events contain one, two, or even three fake electrons/photons (when particles comprising jets are reconstructed as electrons/photons) at the same time, in processes such as $Z \rightarrow ee$ with a fake photon (this is simulated by the inclusive $Z \rightarrow ee$ sample previously mentioned), $\gamma + jet$ and QCD multi-jets. Details concerning the Monte Carlo samples used to simulate these potential additional background processes, along with their associated cross-sections, can be seen in table 5.2.

After running the full analysis selection chain (as detailed in section 5.4) on the samples mentioned in table 5.2, the background contribution from the $t\bar{t}$, ZZ, WW, and WZ processes was found to be insignificant. No events from the $W \rightarrow e\nu$ sample passed the signal selection cuts. Despite the poor statistics of an equivalent 0.1 fb^{-1} of data; 1 event here would enter as ≈ 10 in the final sample, which compared to the signal yield, $\mathcal{O}(100)$, is not severe. A simple cut on the missing transverse energy, E_T^{miss}, should further reduce

Data-set	Process	Generator	Cross Section (pb)	Events	$\int \mathcal{L} dt$ (fb^{-1})
105200	$t\bar{t}$ (t1)	MC@NLO[51]	207.7	199,828	0.96
105986	ZZ	Herwig[52]	1.51	249,725	165.4
105985	WW	Herwig	15.56	50,000	3.21
105987	WZ	Herwig	4.87	50,000	10.27
106020	$W \rightarrow e\nu$	Pythia	10352.3	997,296	0.10
108087	$\gamma + jet$	Pythia	104264	4,994,464	0.048

Table 5.2: Data-set number and cross-sections (corrected for filter efficiency where appropriate) for additional background samples, along with the number of available events and equivalent integrated luminosity.

this background if found to be needed. As a point of reference, table 5.3 shows the number of these additional background events which survive the signal selection cuts; these cuts are explained in detail later, in section 5.4.

Cut	$t\bar{t}$	ZZ	WZ	WW	$W \rightarrow e\nu$
$\mathcal{A}$	4099	23	72	41	1793
$\mathcal{B}$	2615	18	20	5	1122
$\mathcal{C}$	3	6	4	0	0
$\mathcal{D}$	2	5	3	0	0
$\mathcal{E}$	2	4	2	0	0

Table 5.3: Table showing the number of expected events passing the signal selection cuts, for the additional background processes, with an integrated luminosity of $\mathcal{L} = 1\,fb^{-1}$. The cuts are briefly: ($\mathcal{A}$) Event contains two e's and one γ; ($\mathcal{B}$) Separation $\Delta R(e,\gamma) < 2.4$, transverse momentum $p_T(e,\gamma) > 20\,GeV$, and pseudorapidity $|\eta(e,\gamma)| < 2.4$; ($\mathcal{C}$) isEm identification selections: electron = mediumIso, photon = tight; ($\mathcal{D}$) photon isolation criteria; ($\mathcal{E}$) Invariant mass cut, $M_{ee\gamma} > (2M_Z - M_{ee})$. The details of these cuts can be found in section 5.4.

The $\gamma + jet$ sample size is below that required to achieve a statistically significant result; however following on from conclusions made from the study of the QCD multi-jet background in section 5.3, it is also a background which should be reduced to an insignificant level after the imposition of the signal selection cuts. The QCD multi-jet background is a special case, where due to the enormous cross-section compared to the signal process, an alternative method must be employed to provide a useful background

estimate with sufficient precision; this is presented in section 5.3.

5.2.3 Monte Carlo Truth Information

At various points in the analysis of the Monte Carlo data it has been necessary to identify which Monte Carlo generator particles (called "truth" particles in the ATLAS jargon) are responsible for energy deposits/reconstructed particles after the detector simulation. Because there is no unambiguous record of the link between truth and reconstructed particles (due to the nature of the detector simulation), a *truth matching* technique, which relies on a proximity search between truth and reconstructed particles, is employed. For electrons and photons, a study of the spatial separation, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, indicates that a separation cone of $\Delta R < 0.1$ is sufficient to perform a reliable matching. If there is a match within $\Delta R < 0.1$, the reconstructed particle is identified as originating from the truth particle.

Following this truth matching, the Monte Carlo truth information can be navigated in order to identify the origin of any reconstructed particles. From inspection of the *parents* of a truth matched particle, it is possible to locate the production vertex where it originated. It should be noted that these vertices are not like Feynman vertices, they are purely a book-keeping record which is specific to each individual monte-carlo generator. Following this method, reconstructed photons in the analysis are identified as ISR, FSR, from hadron decays, or fake (no match to a truth photon).

5.3 QCD Background

Any hadronic jet falling within the fiducial acceptance criteria of the signal particles (electrons and photons must satisfy $|\eta| < 2.4$, and $p_T > 20\,\text{GeV}$) may mistakenly be reconstructed as an electron or photon (using the algorithms described in sections 4.2 and 4.3). Previous studies have indicated that about one in every $10^3 - 10^4$ (p_T and

identification selection dependent) hadronic jets in ATLAS could fake an electron or photon candidate [48]. The probability that three jets in one event are reconstructed as two electrons and one photon, could naively be assumed to be the order of 1 in $10^9 - 10^{12}$ (ignoring any correlation factors when considering the likelihood that multiple jets fake e/γ's in a single event[2]). This probability is very small, but to ascertain whether it is of any significance, one must consider the QCD di-jet cross-sections, and that the number of expected ISR signal events passing all selection criteria for an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$ is of $\mathcal{O}(100)$.

In order to estimate the expected number of $pp \to 3j+X$ ($|\eta_j| < 2.4$, and $p_T > 20\,\text{GeV}$) events at $\sqrt{s} = 7\,\text{TeV}$, a simple study was conducted on a large QCD Monte Carlo sample (JF17 Di-jet sample, see section 5.3.2 for details). Figure 5.2 shows the predicted yield of events containing at least three hadronic (`Cone4H1TopoJets`, see section 4.5) jets, for an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$, which is approximately 10^{10} events. Combining the small probability of an event containing two fake electrons and a fake photon, with the large number of expected events, it is clear that further investigation of this potential background contribution is necessary. In practice, the number of available simulated QCD multi-jet events is too low (see section 5.3.2) to facilitate a straightforward study of this potential background source; to address this issue, a "fake rate" method was developed.

5.3.1 Jet Fake Rate method

The basic principle of the "fake rate method" is to derive a p_T dependent estimate of the rate at which QCD jets are reconstructed as fake electrons and photons. These "fake rates" are derived from the filtered Pythia JF17 di-jet sample. An independent, unfiltered sample (samples J0 - J4 in section 5.3.2) is then used to construct "fake" signal events, where all eligible jets ($|\eta| < 2.4$, and $p_T > 20\,\text{GeV}$) are assumed to fake an electron/photon object with a weight proportional to the previously derived fake rate. In

[2]The effect of correlations in events containing multiple fake e/γ's has been explored in section 5.3.6, where it is predicted to be insignificant

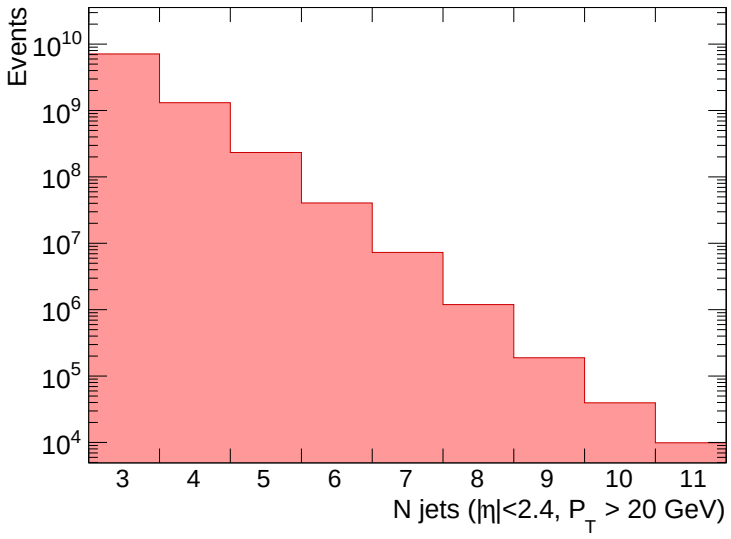


Figure 5.2: The expected number of events containing N jets ($|\eta| < 2.4$, $p_T > 20GeV$), derived from the JF17 di-jet Monte Carlo sample, normalised to an integrated luminosity of $\mathcal{L} = 1fb^{-1}$.

events containing more than three eligible jets, all independent three-wise combinations are constructed, and all combinations of electron-electron-photon are considered. The large number of events assures a statistically smooth distribution, whilst the weighting from the fake rate approximates the normalisation.

5.3.2 QCD Monte Carlo samples

Binned Pythia Di-jet Samples

The ATLAS Monte Carlo production of QCD di-jet events (using the Pythia event generator) is divided into binned samples based on the transverse momentum imparted to the outgoing jets from the primary interaction (Pythia `CKIN 3` parameter [14]). The sample binning and cross-sections are shown in table 5.4.

Inspection of table 5.4 reveals that the number of simulated events in bins J0 - J4 is significantly too few to provide predictions with sufficient statistical weight, for $1\,\mathrm{fb}^{-1}$ integrated luminosity, without using the fake rate method. The J5 and J6 samples are produced in a sufficient quantity to derive statistically significant results by running the

Data-set	Pythia `CKIN 3` range [GeV]	σ (pb)	N Events	Scale Factor to $1\,\mathrm{fb}^{-1}$
J0	$8-17$	9.9×10^{9}	$1,399,184$	7.0×10^{6}
J1	$17-35$	6.8×10^{8}	$1,395,383$	4.9×10^{5}
J2	$35-70$	4.1×10^{7}	$1,398,078$	2.9×10^{4}
J3	$70-140$	2.2×10^{6}	$1,397,430$	1.6×10^{3}
J4	$140-280$	8.8×10^{4}	$1,397,401$	62.8
J5	$280-560$	2.3×10^{3}	$1,389,612$	1.7
J6	$560-1120$	33.6	$1,355,653$	0.025

Table 5.4: Pythia QCD di-jet samples, cross-sections, numbers of simulated events, and scale factors to $1fb^{-1}$ integrated luminosity.

standard analysis routines on them (where no contribution to the background was found), and as such the fake rate study is confined to samples J0 to J4.

Pythia JF17 Filtered Di-jet Sample

The JF17 sample is a QCD di-jet sample which has been filtered at the event generation stage (before full simulation), in order to preferentially accept events which may contain a photon/electron like object (cluster of momentum). The filter algorithm considers all generated particles in an event, combining them into 0.6×0.6 ($\eta\times\phi$) cells. Events where the p_T in any 2×2 cell window is greater than 17 GeV are kept. The use of the filter enhances the statistical weight of the study, since only events passing the filter (approximately 8% of all generated events) are fully simulated, improving the effectiveness of the computational power available. The sample contains $9,987,342$ events, which combined with a production cross-section of $\sigma = 1.1472\times10^{9}\,\mathrm{pb}$, and filter efficiency of $\varepsilon = 0.0852$, corresponds to an integrated luminosity of $\mathcal{L} = 0.1\,\mu\mathrm{b}^{-1}$.

In addition to di-jet events, the JF17 sample also contains Z, W, $t\bar{t}$, and $\gamma + jet$ processes. These events are vetoed from the study, since they contain prompt electrons and photons, which are already considered elsewhere in the analysis.

Pythia JF17 Filter Efficiency

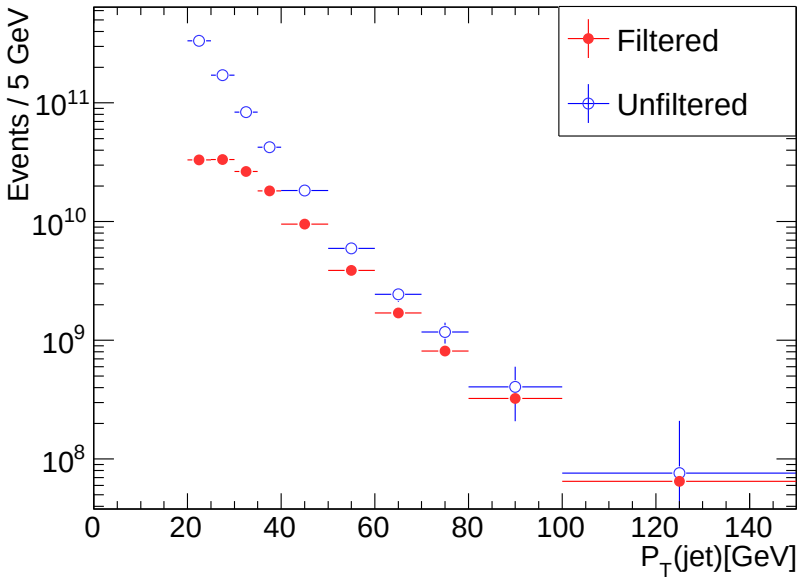
The efficiency of the JF17 filter is not flat with respect to the transverse momentum of the jet in question. Since the QCD background fake-rate method depends on the calculation of p_T dependent fake-rates, it is important that the efficiency of the JF17 filter is known as a function of the jet's transverse momentum, so that the correct p_T-dependent normalisation is applied when calculating fake rate distributions with the JF17 sample. In order to study the efficiency of the JF17 Filter, a sample of approximately $50,000$ Pythia di-jet events were simulated[3], using the same generation parameters as the JF17 sample, but without the filter included. Figure 5.3(a) shows the jet rate (`Cone4H1TopoJet`, $p_T > 20\,\text{GeV}$, and $|\eta| < 2.4$) in the filtered and unfiltered samples, binned in the p_T of the jet, and normalised to an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$. The filter efficiency is defined simply as the ratio of the filtered and unfiltered distributions ($\varepsilon_i = N_i(Filtered)/N_i(Unfiltered)$), and is shown in figure 5.3(b). Figure 5.3(b) tells us that events containing high-p_T jets are more likely to pass the JF17 filter, as expected from the definition of the filter.

5.3.3 Single Electron Fake rate

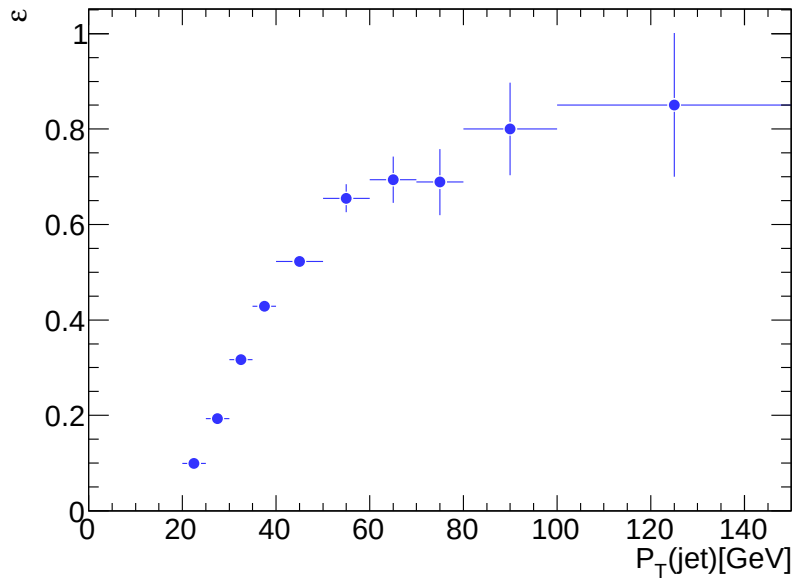
A fake electron in this study is regarded as any reconstructed electron ($p_T > 20$ GeV, $|\eta| < 2.4$ and `MediumIso` ID) which either does not have a truth electron match (within a proximity match $\Delta R < 0.1$), or is collinear with a hadronic jet (in this sense an electron from a b-jet is considered a fake). Each reconstructed electron object is matched to the `Cone4H1TopoJets` jet of closest separation, this is referred to as the matched jet.

Figure 5.4 shows the response in p_T of fake `MediumIso` electrons with respect to the matched jet. High p_T jets are most often reconstructed as electrons with a much lower p_T. Accordingly, the fake rate is calculated in jet p_T bins (for convenience, the binning is chosen to match that of the signal measurement).

[3]Simulated with a parametrised fast simulation of the detector (ATLFASTII).



(a)



(b)

Figure 5.3: Figure (a) shows the jet p_T distributions for filtered (JF17) and unfiltered Pythia di-jet events (Jet $p_T > 20\ GeV$, $|\eta| < 2.4$), normalised to an integrated luminosity of $\mathcal{L} = 1\, fb^{-1}$. Figure (b) shows the resulting effective filter selection efficiency for a jet, calculated from the ratio of the distributions in figure (a).

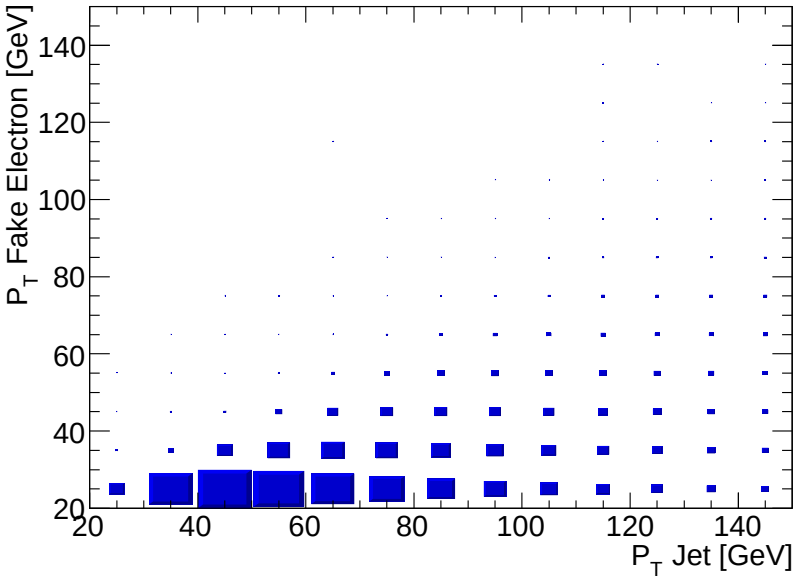


Figure 5.4: p_T of fake "MediumIso" electrons, versus p_T of truth matched jet; the size of the box represents the occupation density.

The electron fake rate is defined as the number of fake electrons divided by the total number of jets in a given bin of jet p_T, $\varepsilon_i = N_i(\textit{fake electrons})/N_i(\textit{jets})$, where the number of jets is corrected according to the JF17 relative filter efficiency derived from figure 5.3. The fake rates for `MediumIso` electrons, binned in the p_T of the matched jet, can be seen in figure 5.5. These distributions are consistent with the spread of p_T between fake electron and jet (figure 5.4), and the expected overall fake rates [48].

5.3.4 Single Photon Fake Rate

A fake photon in this study is regarded as any reconstructed photon ($p_T > 20\,\text{GeV}$, $|\eta| < 2.4$ and `Tight` ID) which either does not have a truth photon match (within a $\Delta R < 0.1$), or is collinear with a hadronic jet (in this sense photons from a π^0 decay are considered as fakes). Each reconstructed photon object is matched to the `Cone4H1TopoJets` jet of closest separation; this is referred to as the matched jet.

The photon fake rate is defined as the number of fake photons divided by the total number of jets in a given bin of jet p_T. The fake rates for `Tight` isolated photons (see

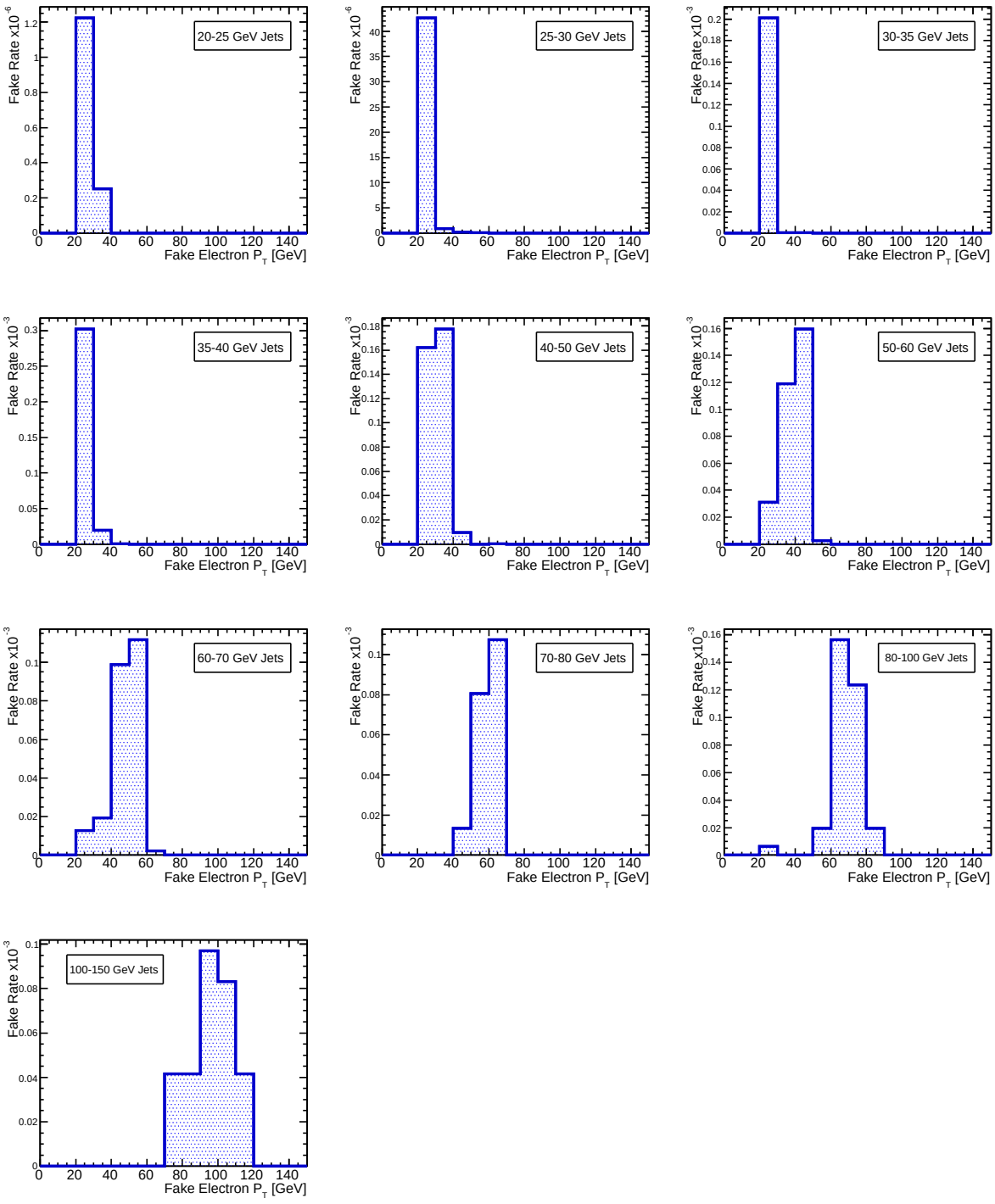


Figure 5.5: "MediumIso" electron fake rates in $20-25$*,* $25-30$*,* $30-35$*,* $35-40$*,* $40-50$*,* $50-60$*,* $60-70$*,* $70-80$*,* $80-100$*, and* $100-150$ *GeV bins in jet* p_T*, as a function of fake electron* p_T*.*

section 5.4.5 for details of the isolation requirement), binned in the p_T of the matched jet, can be seen in figure 5.6.

5.3.5 Fake $Z(\rightarrow ee)\gamma$ Event Rate

The fake event rate is approximated by analysing the binned J0 to J4 di-jet samples and selecting events containing at least three jets ($p_T > 20\,\text{GeV}$, $|\eta| < 2.4$). All three-wise combinations of jets in selected events are constructed, and to each unique combination, all three permutations of one photon, and two electrons, are made. The jet which has been chosen to fake a photon is randomly assigned a p_T based on the corresponding histogram in figure 5.6, whereby a fake photon p_T bin is chosen by generating a random number (normalised to the area of the histogram); the p_T bin containing the point at which the integral of the distribution equates the random number is chosen. Within this randomly selected bin, a second random number is generated to select a random value of p_T. The direction (η, ϕ) of this fake photon is chosen to match that of the jet being faked. The fake electrons are created in the same manner, using the electron fake-rates shown in figure 5.5.

Accordingly, Z boson and $Z+\gamma$ composite objects are created from the combinations of fake electron and photon, and the weights of the fake electrons and photons, corresponding to their derived p_T-dependent fake rates, are multiplied together to approximate the rate at which all three objects are faked together. This is a simple assumption, ignoring any effects due to correlations in events containing two fake objects (e.g. as may be the case with electrons from b-jets). A study of Z boson (di-electron) fakes, where correlations are accounted for (section 5.3.6), indicates that such correlations are not large enough to have a significant effect. Only events where the "fake Z" mass is greater than 60 GeV are kept, in line with the same mass selection in the standard analysis. The final estimates of the QCD background contribution from the binned di-jet samples (Pythia) J0-J4, to the signal photon p_T and $M(ee\gamma)$ distributions, are shown in figure 5.7.

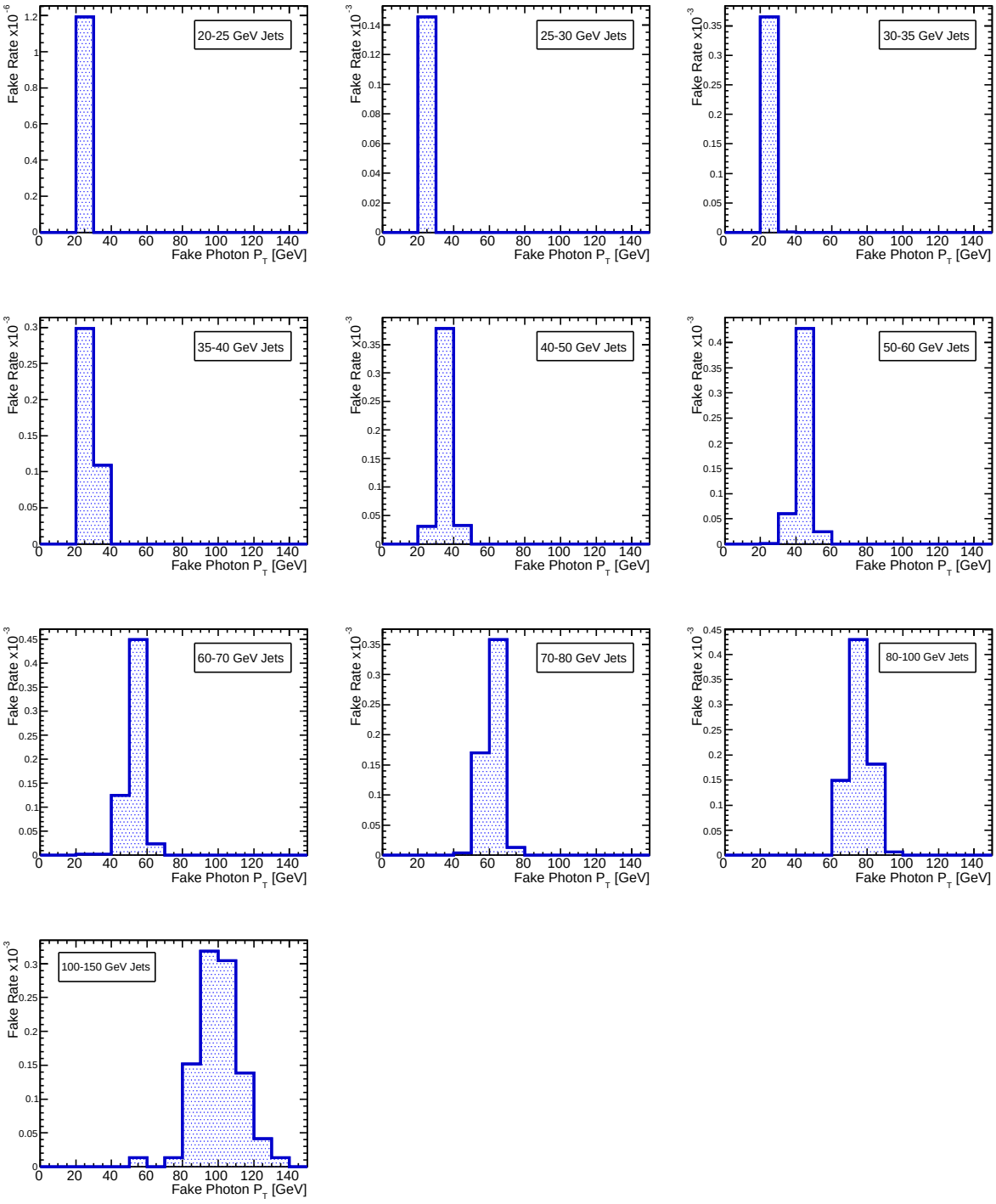
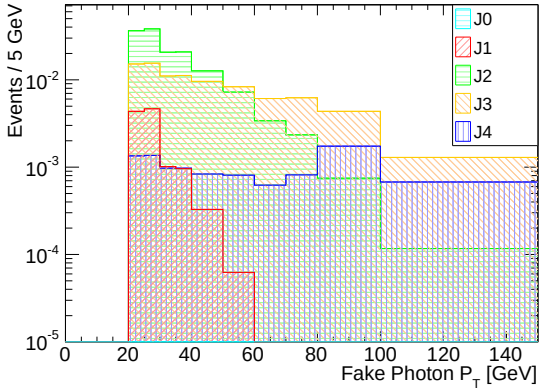
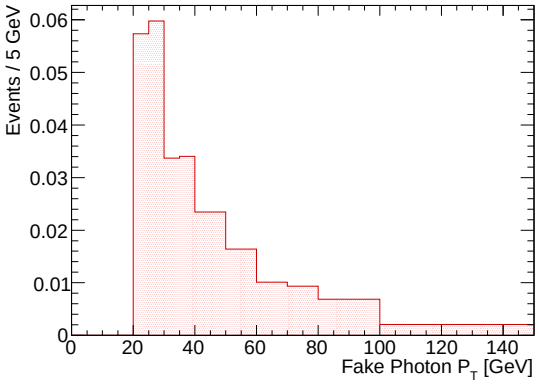


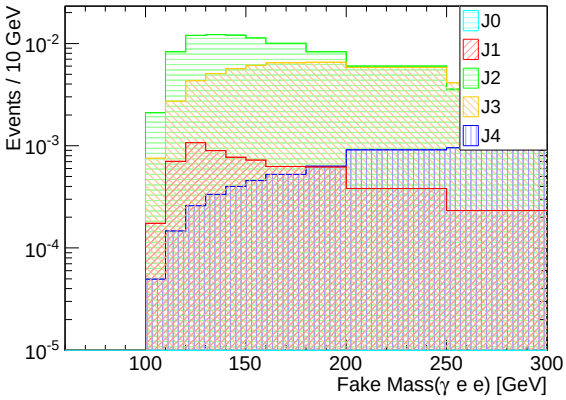
Figure 5.6: "Tight" isolated photon fake rates in $20-25$, $25-30$, $30-35$, $35-40$, $40-50$, $50-60$, $60-70$, $70-80$, $80-100$, *and* $100-150$ *GeV bins in jet* p_T, *as a function of fake photon* p_T.



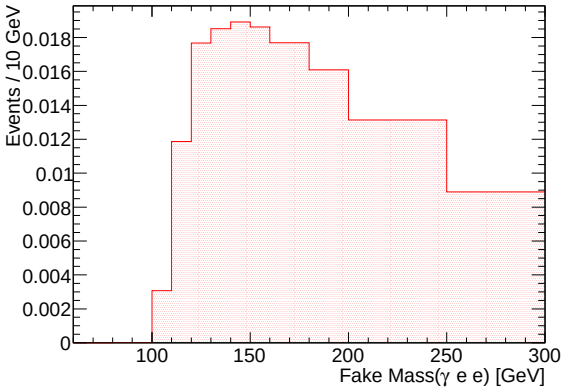
(a) Binned contributions.



(b) Sum of binned contributions.



(c) Binned contributions.



(d) Sum of binned contributions.

Figure 5.7: (a) Estimated photon p_T background contribution from binned di-jet samples J0, J1, J2, J3, and J4, using approximate fake rate method, in bins of fake photon p_T. (b) Sum of binned J0-J4 contributions. (c) J0 - J4 binned $M(ee\gamma)$ contribution. (d) J0 - J4 summed $M(ee\gamma)$ contribution. Distributions normalised to an integrated luminosity of $\mathcal{L} = 1\, fb^{-1}$.

5.3.6 Validation of Approximate Fake Rate Method

The validity of the fake-rate method can be tested by considering a process with a less sophisticated reconstruction than the full $Z+\gamma$ signal, requiring in the first instance only one single fake particle per event. There is a sufficient number of simulated events containing a single fake particle to facilitate a direct comparison between the proposed fake-rate method, and the standard analysis routine, with an adequate statistical significance. The first test comparison is made between `Tight` photons predicted by the fake-rate method, where every jet ($|\eta| < 2.4$, and $p_T > 20\,\text{GeV}$) in samples J0 - J4 is assigned a fake photon according to the method described in section 5.3.5, and of `Tight` photons identified by the standard analysis routine in the same samples. Figure 5.8 shows the comparison of photon transverse momentum, $p_T(\gamma)$, between the two methods, confirming that the fake-rate technique successfully reproduces the single photon spectrum shape, with the normalisation correct to within about 20%.

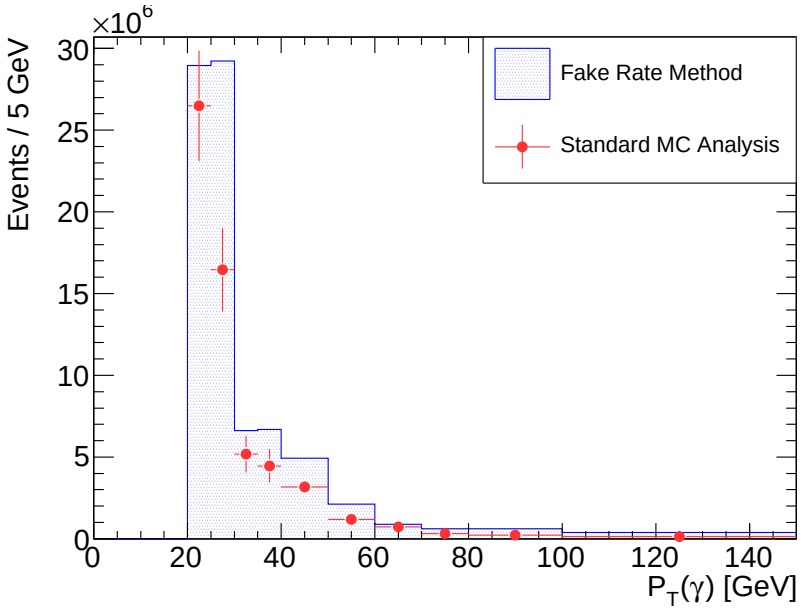


Figure 5.8: Comparison of the p_T spectrum of "Tight", isolated photons derived using the fake-rate method, and using the standard analysis routine, for binned J0 - J4 di-jet samples, normalised to $\mathcal{L} = 1\,fb^{-1}$ integrated luminosity. The structure in the statistical errors originates from the fact that the distributions are the sum of the binned di-jet distributions, each having a different normalisation factor.

A second comparison is performed between the `Loose` di-electron fake-rate and stan-

dard `Loose` di-electron reconstruction. This gives an indication of the possible effects of ignoring correlations in events containing two fake electrons. The di-electron fake-rate is calculated via the construction of all unique pairings of single fake-electrons in an event, with each di-electron pair possessing a weight equal to the product of the fake-rates of the individual electrons comprising the pair. A comparison of the invariant mass spectrum of such di-electron pairs produced using this fake-rate method, along with that of di-electron pairs reconstructed by the standard analysis routine, is shown in figure 5.9.

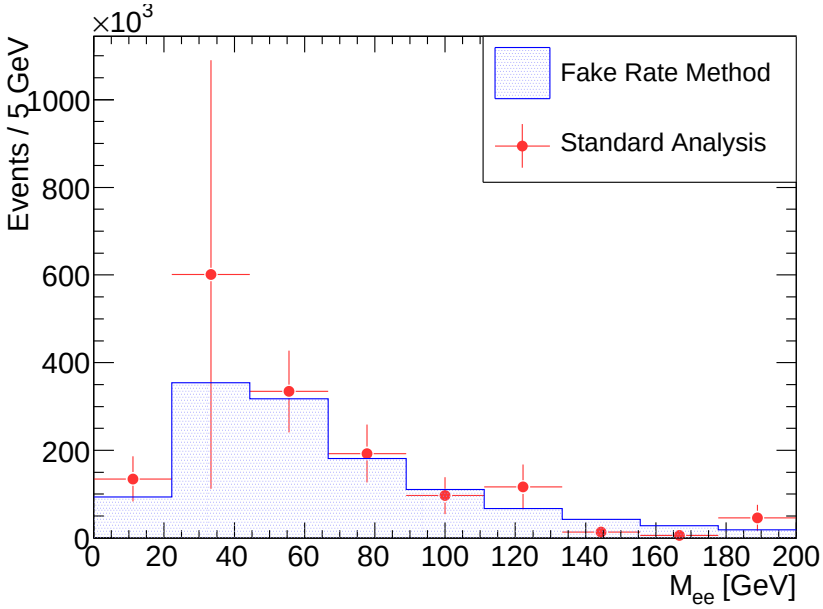


Figure 5.9: Comparison of the invariant mass of "Loose" di-electron pairs derived using the fake-rate method, and using the standard analysis routine, for binned J0 - J4 di-jet samples, normalised to $\mathcal{L} = 1fb^{-1}$ integrated luminosity. The structure in the statistical errors originates from the fact that the distributions are the sum of the binned di-jet distributions, each having a different normalisation factor.

Figure 5.9 indicates that both the shape, and normalisation of the invariant mass distribution, is reproduced well by the fake-rate method, albeit with a limited statistical weight.

QCD Background Study Conclusions

A fake rate method was applied to study the possible contribution to the $ee\gamma$ final state from QCD multi-jet events, where the jets are mistakenly reconstructed as electrons and

photons. Figures 5.7(b) and 5.7(d) show the estimated fake background contributions, where much less than one event is expected to enter either distribution. This when compared to $\mathcal{O}(100)$ signal events (figures 5.19(a) and 5.19(b)) is negligible. Checks were made of the method by comparing estimates derived using the fake-rate method for single photons (`Tight`), and Z bosons (`Loose` di-electron pairs), to the distributions derived from the straightforward analysis routine. An agreement was found between the methods to within 20%, and as such the final estimates of the QCD background should have a similar uncertainty coming from the fake-rate method. It should be noted that this uncertainty is in the fake-rate method, and not in the fake electron/photon rate; this uncertainty is also too small to give this background any importance (see section 5.7.4 for comments on the jet fake rate uncertainty). It is following this study that the QCD background is no longer considered as an important background to the analysis, and is hereafter ignored.

5.4 Event Selection

The following section motivates and describes the event selection critera, and kinematical cuts, which have been chosen to optimise the $Z(\rightarrow ee)\gamma$ ISR differential cross section study. This includes the trigger, choice of reference selection (`isEm`) for electrons and photons, fiducial acceptance criteria (p_T, η, $\Delta R(e,\gamma)$), photon isolation criteria, and a cut in the $ee - ee\gamma$ invariant mass plane; as well as the motivation for the bin widths for the final differential cross section distributions.

5.4.1 Trigger Requirement

The first stage in selecting a signal event is to choose a trigger condition, since only events passing the trigger requirement will be analysed when ATLAS collects real physics data. In order to keep the maximum number of signal events, the most inclusive electron (`EF_e20_loose`) and di-electron (`EF_2e5_medium`) triggers available in the current menu

of trigger items are chosen. Since the analysis requires two electrons with $p_T > 20\,\text{GeV}$ within $|\eta| < 2.4$, these trigger items applied in combination are sufficient to select all of the good signal events. The efficiency of the trigger at selecting good signal events was estimated using simulated data, after all other acceptance criteria had been applied, and was found to be 100%, as expected.

5.4.2 Particle Overlap Removal

The first step in the analysis of events passing the trigger is the removal of overlapping objects created by the particle reconstruction algorithms. The full details of this overlap removal can be found in sections 4.2.2 and 4.3.2; the procedure can be summarised as follows:

- Photons (unconverted and converted) passing the fiducial acceptance criteria ($p_T > 20\,\text{GeV}$ within $|\eta| < 2.4$) are selected.
- Electrons($p_T > 20\,\text{GeV}$ within $|\eta| < 2.4$) authored by `Electron` and `Softe` algorithms simultaneously, and separated by $\Delta R > 0.4$ from pre-selected photons are selected.
- Electrons ($p_T > 20\,\text{GeV}$ within $|\eta| < 2.4$) authored solely by `Electron` algorithm, and separated by $\Delta R > 0.1$ from any other pre-selected electron (and $\Delta R > 0.4$ form pre-selected photons) are selected.

5.4.3 Di-electron Acceptance

The identification of signal events begins with the selection of di-electron pairs, where both electrons must satisfy the acceptance requirements $|\eta| < 2.4$, $p_T > 20\,\text{GeV}$, and in addition pass the standard electron identification `MediumIso` reconstruction criteria. The choice of `MediumIso` ID is motivated by the rejection power with respect to hadronic jets;

the electron fake rate drops by a factor of at least 6 for `MediumIso` electrons, with respect to `loose` electrons [46].

5.4.4 Z Boson Selection

Z boson candidates are constructed from any identified di-electron pairs passing the electron acceptance criteria, with an invariant mass constraint of $M_{ee} > 60\,\text{GeV}$. In a small number of signal events (2.5%), where there are more than two good reconstructed electrons, more than one candidate Z boson may be identified; in such events it must be decided how one can preferentially choose the correct pair. In simulated $Z\gamma$ events containing more than one Z boson candidate, characteristics of the *good* (both electrons truth matched to a Z) and *bad* (unmatched) electron combinations were compared. Following this inspection, the di-electron pair of opposite charge (if such a pairing exists), closest to the true Z boson mass is regarded as the correct pair. This is effective at choosing the correct pair most often, but at the cost of the introduction of a small (but measurable) bias towards the Z mass[4]. If no pair of zero charge sum exists, the event is rejected (background events are more likely to contain like-sign pairs); using this procedure 0.7% of good ISR events are rejected.

Figure 5.10 shows the reconstructed Z mass from the $Z \to ee$ sample, along with the expected QCD contribution calculated from the fake rate method described in section 5.3.1 (adapted to `MediumIso` di-electron pairs).

5.4.5 Photon Selection

In addition to the Z boson candidate, a selected event must contain a reconstructed photon object within $|\eta| < 2.4$, $p_T > 20\,\text{GeV}$, separated from the electrons by $\Delta R(\gamma, e) > 0.4$, and pass the standard photon reconstruction classification `Tight`. The photon is also required

[4]This bias is not accounted for in the analysis, since the contribution is $< 1\%$.

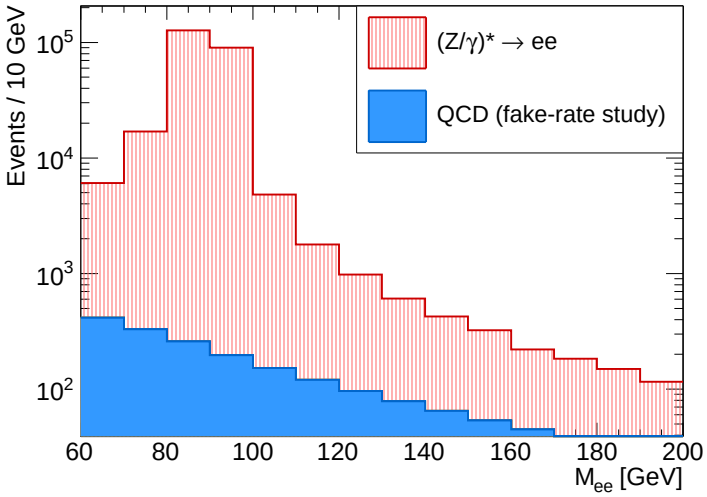


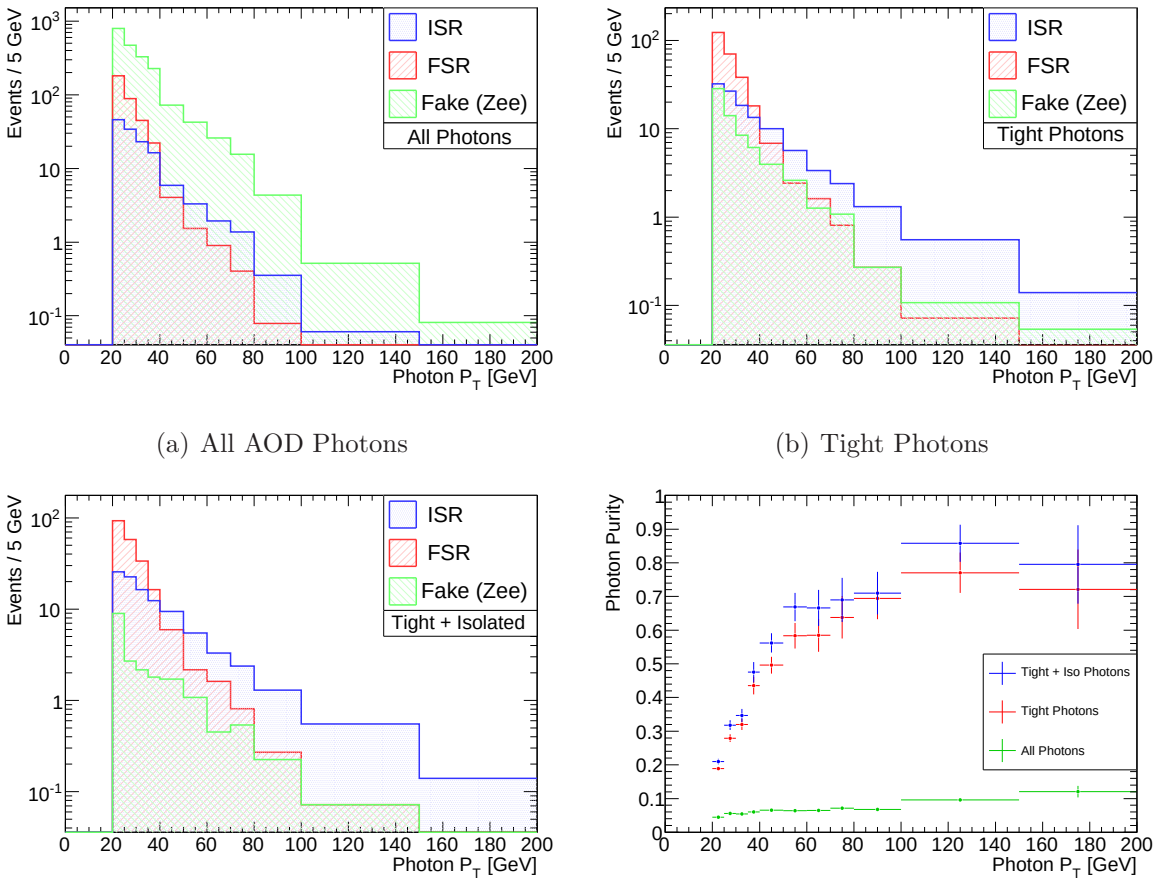
Figure 5.10: Reconstructed M_{ee} distribution for all selected Z boson candidates in the $Z \rightarrow ee$ sample, and the expected QCD contribution (calculated using the fake-rate method), for an integrated luminosity of $\mathcal{L} = 1fb^{-1}$.

to pass an isolation criteria, where the sum of energies in a cone of 0.4 solid angle centred on the photon direction, $E_T(cone\,0.4)$, minus the energy of the photon candidate, divided by the energy of the photon candidate, must be less than 0.1:

$$\frac{E_T(cone\,0.4) - E_T(\gamma)}{E_T(\gamma)} < 0.1. \tag{5.2}$$

The choice of a `Tight` *Isolated* photon is motivated by the rejection power against hadronic jet fakes. Figure 5.11 shows a comparison between the signal and background contributions to the $p_T(\gamma)$ distribution for `Tight` isolated photons, `Tight` without isolation, and all AOD photons (default, without and isEm ID requirements), in events passing the Z boson cuts detailed in section 5.4.4.

As expected, figures 5.11(a) - 5.11(d) show that the `tight` ID criteria do a good job at rejecting fake photons, with a reduction of 97%, whilst only reducing the signal (ISR) yield by 20%. The isolation criterion further suppresses the background contribution from fake photons by 68%, with again a relatively small associated drop in signal yield of 10%. Table 5.5 summarises the expected yield of photons (in events passing the Z selection cuts only) for `tight` isolated, `tight`, and all AOD photons, with an integrated luminosity



(a) All AOD Photons

(b) Tight Photons

(c) Tight isolated Photons

(d) Photon Purity

Figure 5.11: (a) Photon p_T distribution in events passing Z selection, for all default AOD photons (no isEm ID requirement), for signal (ISR) and backgrounds (FSR, Fakes [from $Z \rightarrow ee$ sample, including photons from hadrons such as $\pi^0 \rightarrow \gamma\gamma$]), normalised to an integrated luminosity of $\mathcal{L} = 1\,fb^{-1}$. (b) Distribution for tight photons. (c) Distribution for tight, isolated photons. (d) Purity of photon selection $\left(\mathcal{P}_i = \frac{N_i(Good\ Reco)}{N_i(All\ Reco)}\right)$ for the different photon selections. Other backgrounds are neglected.

of $\mathcal{L} = 1\,\text{fb}^{-1}$.

Photon Origin	Number of expected events		
	All Photons	`Tight` Photons	`Tight` Isolated Photons
ISR	184	147	131
FSR	366	274	223
Fake	2585	78	25

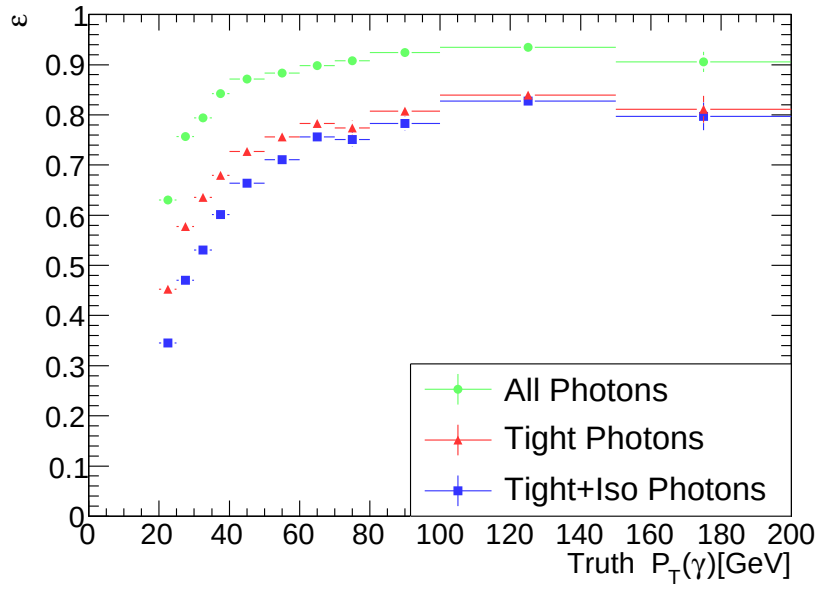
Table 5.5: Number of expected events which are selected by the various photon ID criteria in events passing Z selection cuts (from ISR signal sample, and FSR background sample), for an integrated luminosity $\mathcal{L} = 1\,fb^{-1}$. Other backgrounds neglected.

The efficiency with which signal photons are reconstructed (without any electron requirements or Z boson cuts) is defined in equation 5.3:

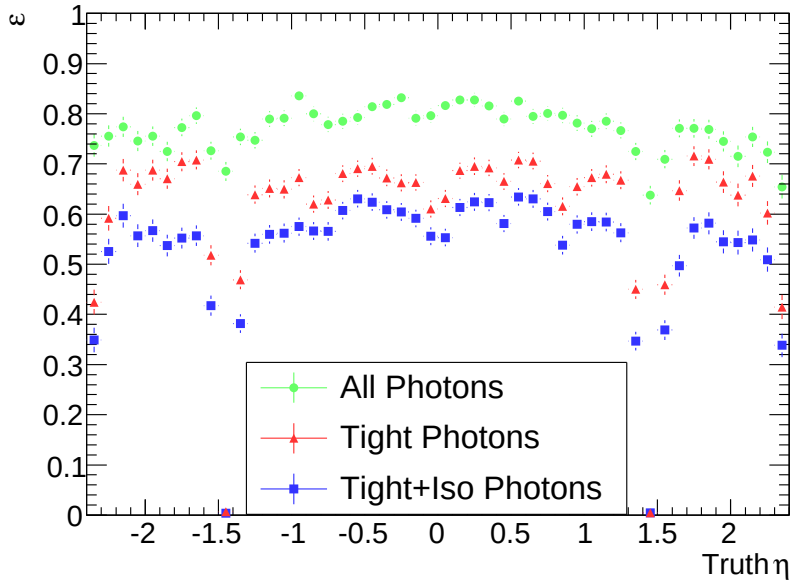
$$\varepsilon_i = \frac{N_i(\text{Reco \& Truth})}{N_i(\text{Truth})}, \tag{5.3}$$

here N_i(Reco & Truth) is the number of reconstructed truth matched photons in the $i-$th bin, and N_i(Truth) is the number of truth photons in the $i-$th bin, within the same fiducial acceptance. Figure 5.12 shows the signal photon efficiency, ε_i, in bins of photon transverse momentum and pseudo-rapidity η.

The reconstruction efficiencies shown in figures 5.12(a) and 5.12(b) are consistent with previous studies [46]. The drop in efficiency in the pseudo-rapidity distribution around $|\eta| \approx 1.5$ is due to the geometry of the calorimeter (where the barrel and end-cap meet; see section 2.4), this region, $1.37 < |\eta| < 1.52$ is excluded from the `Tight` photon identification criteria.



(a)



(b)

Figure 5.12: (a) Reconstruction efficiency, ε, of all ISR photons in the signal ($Z + \gamma$) sample, for all AOD photons, Tight photons, and Tight isolated photons, in bins of $p_T(\gamma)$. (b) ε, in bins of photon pseudo-rapidity, η.

5.4.6 $Z\gamma$ Invariant Mass Cut

After the photon selection criteria have been applied, the remaining background to the ISR signal comes mostly from FSR photons in the inclusive $Z \rightarrow ee$ channel (see figure 5.11). A simple and robust way to select the ISR events, and reject the FSR events, is to consider the invariant mass of the ee and $ee\gamma$ systems. In ISR events, the ee invariant mass will tend to cluster around the true Z boson mass, and the $ee\gamma$ invariant mass will tend to be greater than the Z boson mass; conversely in FSR events, the $ee\gamma$ mass will tend to cluster around the true Z boson mass, with the ee mass tending to be less than the Z boson mass. Figure 5.13 demonstrates the simulated invariant mass distributions of the ee and $ee\gamma$ systems.

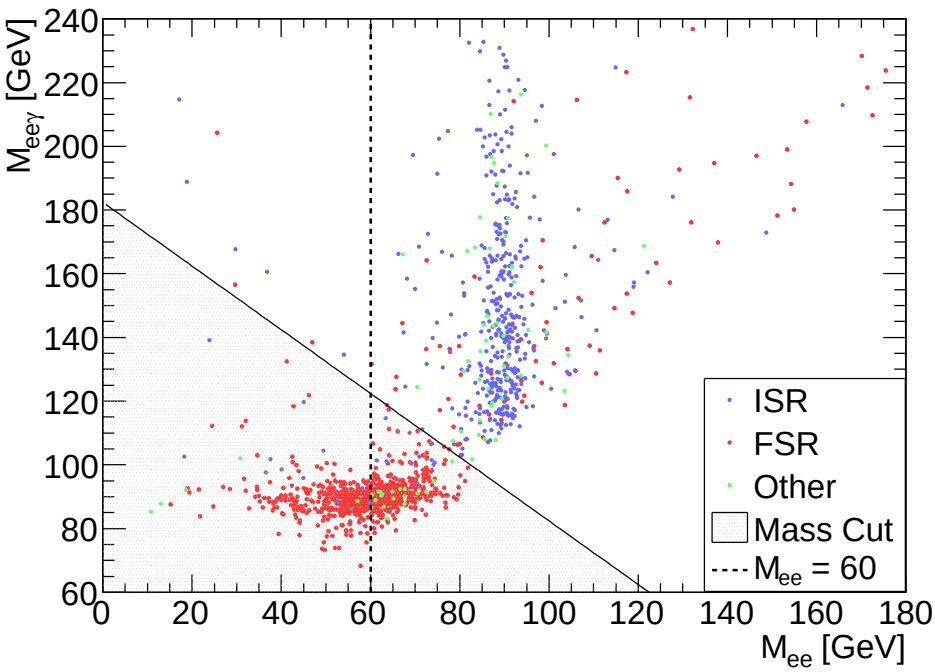


Figure 5.13: M_{ee} versus $M_{ee\gamma}$ for ISR, FSR and Fake photons, with selection cut $M_{ee\gamma} > (2M_Z - M_{ee})$ overlaid, for an expected integrated luminosity of $\mathcal{L} = 5fb^{-1}$.

It can be seen from figure 5.13 that a simple cut in the M_{ee}-$M_{ee\gamma}$ plane can effectively distinguish between the majority of ISR and FSR events. The cut chosen is $M_{ee\gamma} > (2M_Z - M_{ee})$. There are a number of alternative, and more sophisticated techniques available to distinguish between ISR and FSR events; including Neural Networks, Boosted

Decision Trees, and log likelihood approaches [48]. These techniques rely on a multivariate combination of experimental observables, many of which may not be fully understood until the experiment has reached a level of operational maturity. Since this thesis is concerned with the short to medium term capabilities of ATLAS, and the performance of the binary mass cut is already good, the most robust and well understood method has been adopted.

5.4.7 Distribution Bin Width

The choice of histogram binning for the distributions being measured is both important, and non-trivial. Smaller bin widths will tend to result in a better representation of the exact shape of the distribution (in the absence of statistical errors). However, if the bin width is too narrow, the statistical error could grow too large to provide a precise result, and bin migrations due to resolution effects will become large. Since a bin-by-bin correction method is being used in the analysis, it is important that migrations are kept small. A useful way of studying the effect of bin migration is to consider whether a photon's reconstructed p_T or $M_{ee\gamma}$ falls in the same bin as that in which it was generated. We can consider this as the stability, $\mathcal{S}_i$, in the i-th bin, where:

$$\mathcal{S}_i = \frac{N(\text{Gen \& Reco})_i}{N(\text{Gen})_i}. \tag{5.4}$$

Here $N(\text{Gen \& Reco})_i$ is the number of photons generated and reconstructed in the same bin (p_T or $M_{ee\gamma}$), and $N(\text{Gen})_i$ is the number generated in the same bin, but reconstructed in any bin. If all of the photons were reconstructed in the same bin as that in which they were generated, the stability would equal unity. If the bin width is too narrow, resolution effects will cause the migration of reconstructed photons to adjacent bins. The stability should be high (it will not reach unity, because there are always some migrations at the edges of bins), and flat across the desired range.

Figure 5.14 shows the stability variation for a number of bin widths for the photon p_T distribution; by inspection, the smallest acceptable bins in each p_T region are chosen, resulting in a variable binning corresponding to $[20-25]$, $[25-30]$, $[30-35]$, $[35-40]$, $[40-50]$, $[50-60]$, $[60-70]$, $[70-80]$, $[80-100]$, $[100-150]$, and $[150-200+]$ GeV bins. This choice is a compromise between keeping the bin widths as anarrow as possible, whilst maintaining a high, and reasonably flat stability.

Figure 5.15 shows the stability variation for a number of bin widths for the $ee\gamma$ invariant mass distribution; following on from the stability calculation, the smallest acceptable bins in each region are chosen, resulting in a variable binning for the $M_{ee\gamma}$ distribution corresponding to $[100-120]$, $[100-120]$, $[120-140]$, $[140-160]$, $[160-180]$, $[180-200]$, $[200-250]$, $[250-300]$, and $[300-400+]$ GeV bins. For both the p_T and $M_{ee\gamma}$ distributions the highest bin includes any outliers; this is important in order to catch any potential anomalous, or unexpected result from real data, which may be an indication of new physics.

5.4.8 Uncorrected $p_T(\gamma)$ and $M_{ee\gamma}$ Distributions

The $p_T(\gamma)$ and $M_{ee\gamma}$ distributions created from the events passing all of the selection criteria are displayed in figure 5.16. The breakdown of the different contributions in the Monte Carlo samples is shown, normalised to an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$, along with the yield from the statistically independent pseudo-data sample, representing N_i in equation 5.1, shown with statistical errors only. By comparison of figure 5.16 to figure 5.11(c), the rejection power of the $ee\gamma$ invariant mass cut with respect to the FSR background can be clearly seen. The agreement between the pseudo-data sample and summed Monte Carlo distributions indicates that the splitting of the one large Monte Carlo sample into two independent samples is self consistent (the normalisations match).

The number of expected events expected to pass all selection criteria, for ISR, FSR, and fake photons in $Z \to ee$ events, with an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$, are

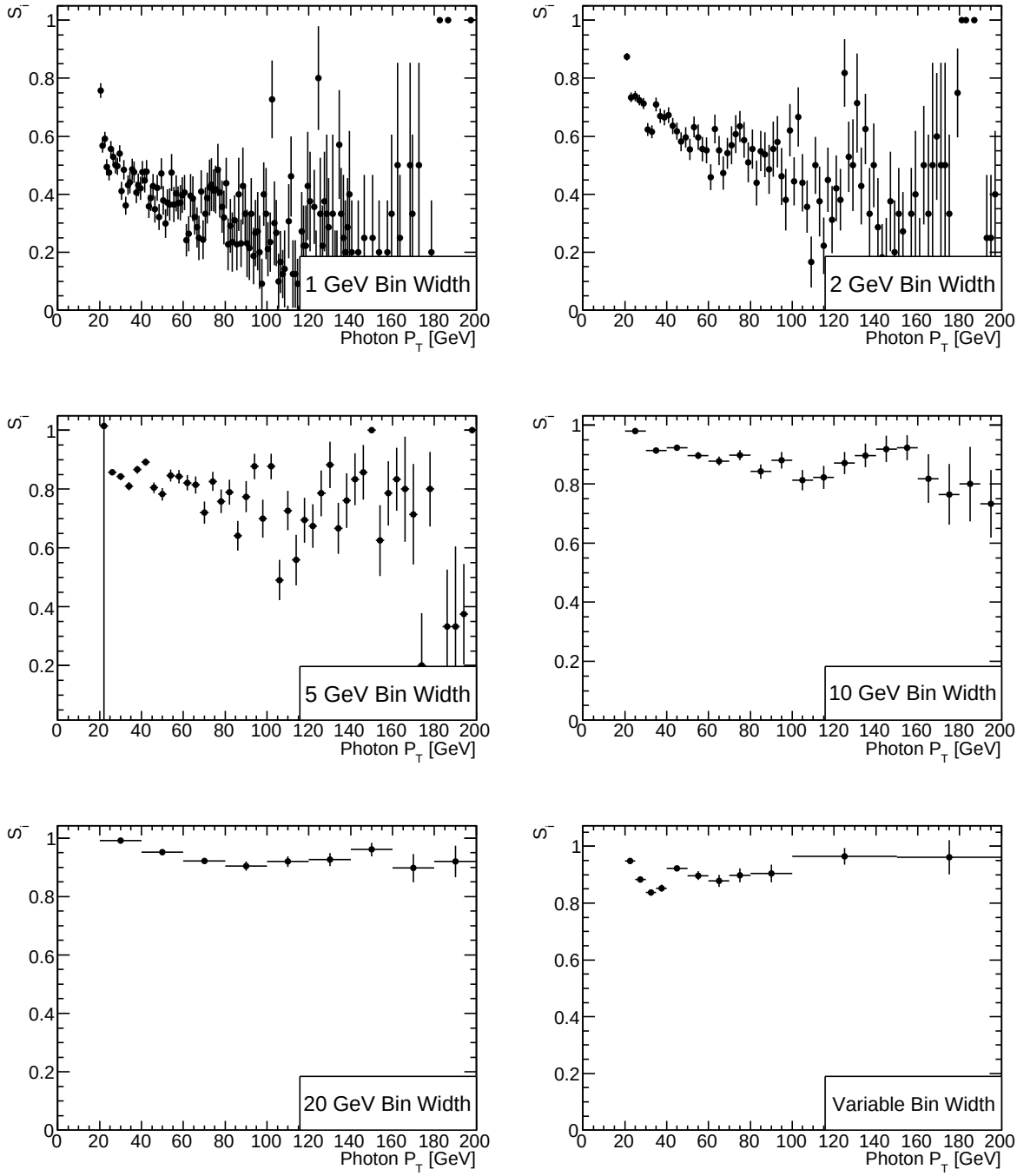


Figure 5.14: Photon p_T bin width stability for various bin widths, including the final choice of variable bin width: $[20-25]$, $[25-30]$, $[30-35]$, $[35-40]$, $[40-50]$, $[50-60]$, $[60-70]$, $[70-80]$, $[80-100]$, $[100-150]$, *and* $[150-200+]$ *GeV bins.*

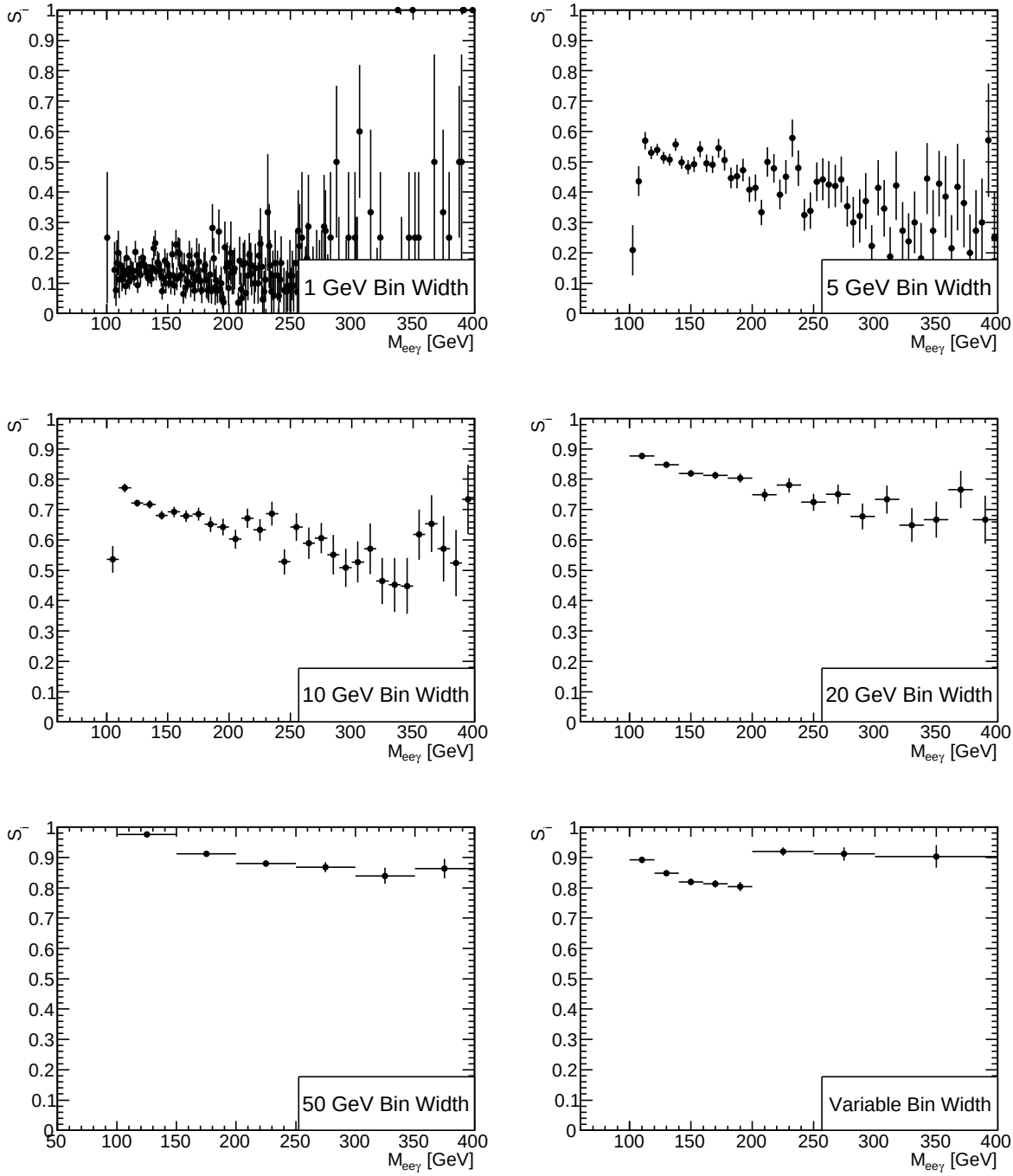
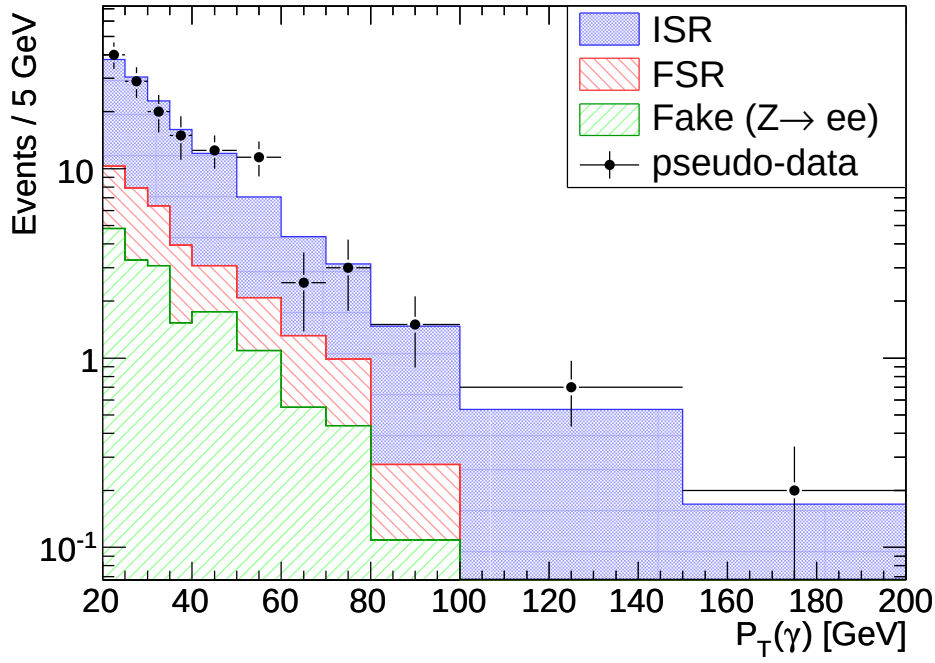
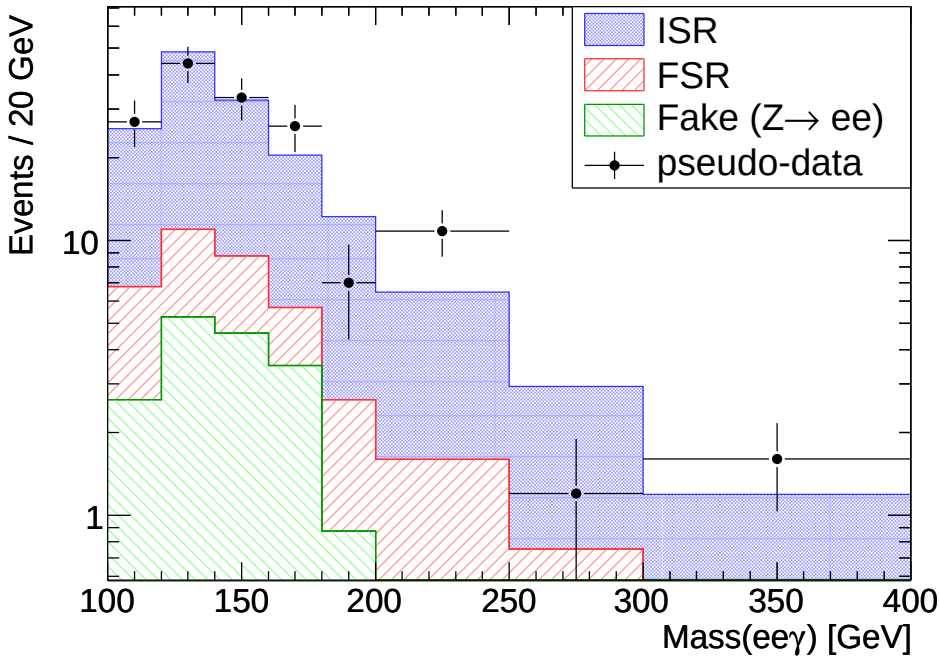


Figure 5.15: $M_{ee\gamma}$ bin width stability for various bin widths, including the final variable bins choice of $[100-120]$, $[100-120]$, $[120-140]$, $[140-160]$, $[160-180]$, $[180-200]$, $[200-250]$, $[250-300]$, *and* $[300-400+]$ *GeV bins.*



(a)



(b)

Figure 5.16: (a) $p_T(\gamma)$ distribution of events passing all selection cuts, with breakdown of photon type (MC contributions are added together/stacked), normalised to an integrated luminosity of $\mathcal{L} = 1fb^{-1}$, with comparison to pseudo-data sample. (b) $M_{ee\gamma}$ distribution.

shown in table 5.6

Photon Origin	Number of expected events
ISR	126
FSR	24
Fake	22
Total	172

Table 5.6: Number of expected events passing all selection criteria, for ISR, FSR and fake photons in $Z \to ee$ events, for an integrated luminosity $\mathcal{L} = 1\,fb^{-1}$. Other backgrounds neglected.

5.5 Signal Efficiencies and Purities

This section describes the derivation of the correction factors in equation 5.1 which are used to correct the measured distribution back to the predicted differential cross-section. The following distributions are created using the full amount of simulated data available (removing the reserved $1\,\text{fb}^{-1}$, which is for the statistically independent pseudo-data sample). Because the signal is defined as the ISR process only, the $Z+\gamma$ sample alone is used to derive the efficiency factors; for the purity factors, the background contamination is included, and the $Z \to ee$ sample is used accordingly, in addition to the $Z+\gamma$ sample.

5.5.1 Signal Selection Efficiency

The signal reconstruction efficiency, ε_i, is a measure of how successfully the analysis selects good signal events. That is, how many of the signal events in each bin (p_T, or $M_{ee\gamma}$) pass the trigger, contain two reconstructed electrons ($|\eta| < 2.4$, $p_T > 20\,\text{GeV}$, `MediumIso` ID, $M_{ee} > 60\,\text{GeV}$) and a photon ($|\eta| < 2.4$, $p_T > 20\,\text{GeV}$, `Tight` ID, and isolation criteria), with a separation cone between electron and photon of at least 0.4, whilst satisfying the invariant mass constraint $M_{ee\gamma} > (182.4\,\text{GeV} - M_{ee})$. The reconstruction efficiency is

defined as follows:

$$\varepsilon_i = \frac{N^{ISR}(\text{Reco \& Truth})_i}{N^{ISR}(\text{Truth})_i}. \tag{5.5}$$

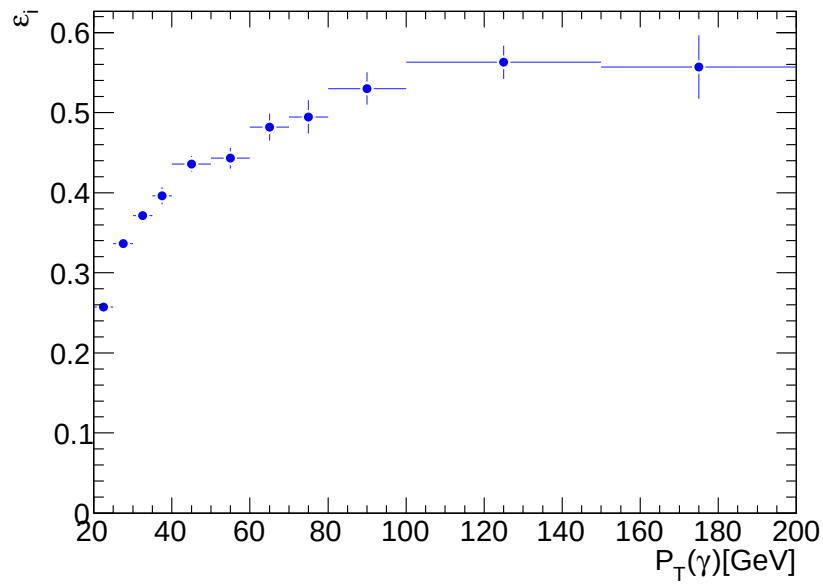
Here $N^{ISR}(\text{Reco \& Truth})_i$ is the number of ISR photons in the i-th bin ($p_T(\gamma)$ or $M_{ee\gamma}$) which have been reconstructed, and successfully matched to the Monte Carlo truth record, confirming they are indeed ISR in origin (note that the truth photons may not have been generated in the same bin as that in which they are reconstructed, as demonstrated previously); and $N^{ISR}(\text{Truth})_i$ is the number of photons expected purely by inspection of the Monte Carlo truth particles, in the same bin i. The ε_i distribution is calculated from the $Z+\gamma$ signal sample, and carries the underlying assumption that the event contains two truth electrons ($|\eta| < 2.4$, $p_T > 20\,\text{GeV}$), and one truth photon ($|\eta| < 2.4$, $p_T > 20\,\text{GeV}$), with $\Delta R(e, \gamma) > 0.4$. The derived efficiency distributions for the photon p_T, and $ee\gamma$ invariant mass measurements, can be seen in figure 5.17.

5.5.2 Signal Selection Purity

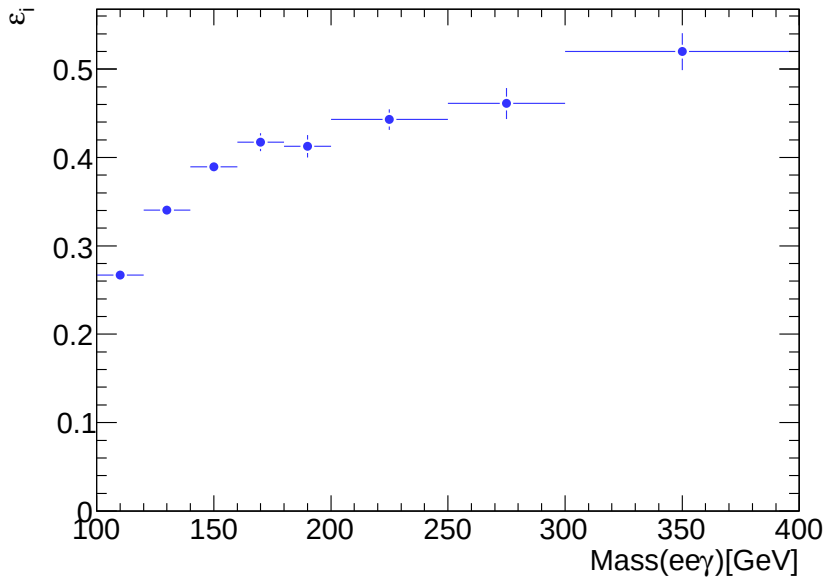
In order to quantify the background contamination of the measured distribution N_i, a purity factor $\mathcal{P}_i$ must be calculated. The signal event purity is defined as:

$$\mathcal{P}_i = \frac{N^{ISR}(\text{Reco \& Truth})_i}{N^{All}(\text{Reco})_i}, \tag{5.6}$$

here $N^{ISR}(\text{Reco \& Truth})_i$ is the same as in equation 5.5; and $N^{All}(\text{Reco})_i$ is the total number of photons reconstructed in the i-th p_T or $M_{ee\gamma}$ bin, inclusive of contributions from the background sources. Figure 5.18 shows the derived purity distributions for the p_T and $M_{ee\gamma}$ measurements. In practice only the FSR and fake photons in $Z \to ee$ events are considered; all other backgrounds were found to be negligible (see section 5.2.2), and are ignored in the purity calculation.

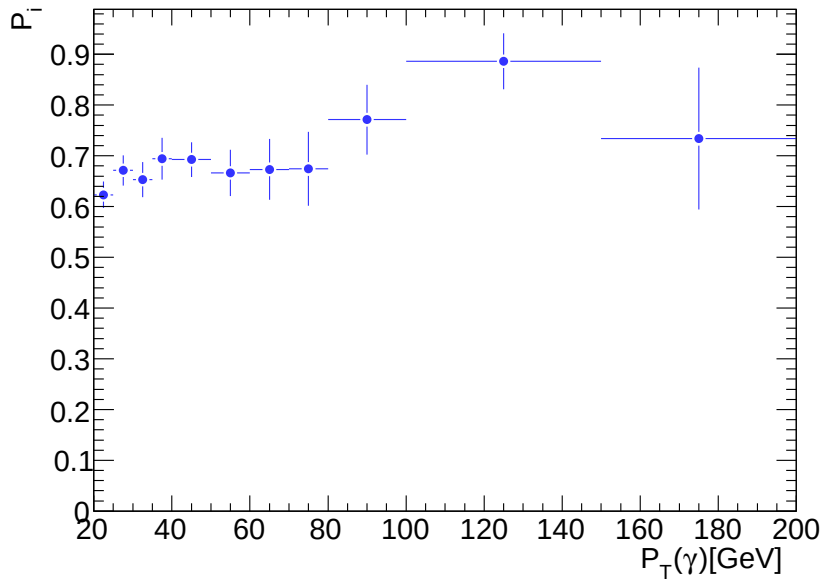


(a)

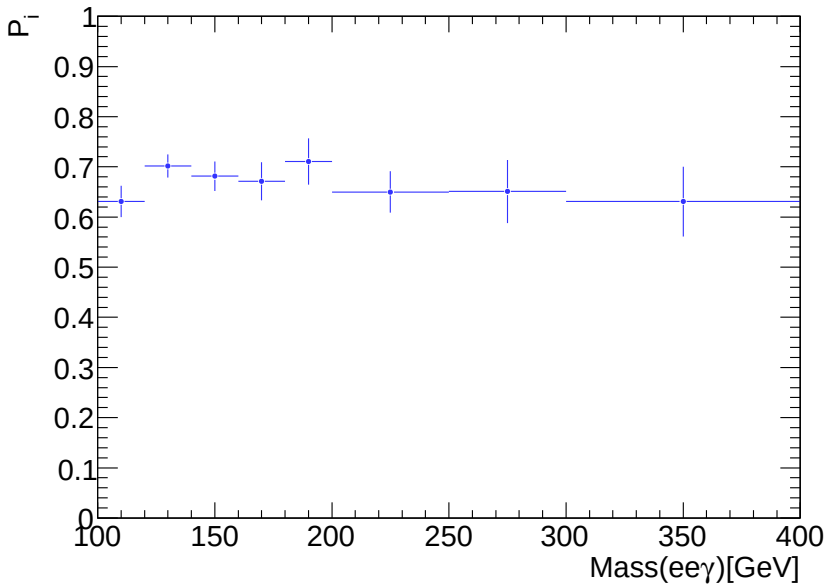


(b)

Figure 5.17: (a) Signal event reconstruction efficiency ε_i, in bins of $p_T(\gamma)$. (b) Signal event reconstruction efficiency ε_i, in bins of $M_{ee\gamma}$. Errors are statistical only.



(a)



(b)

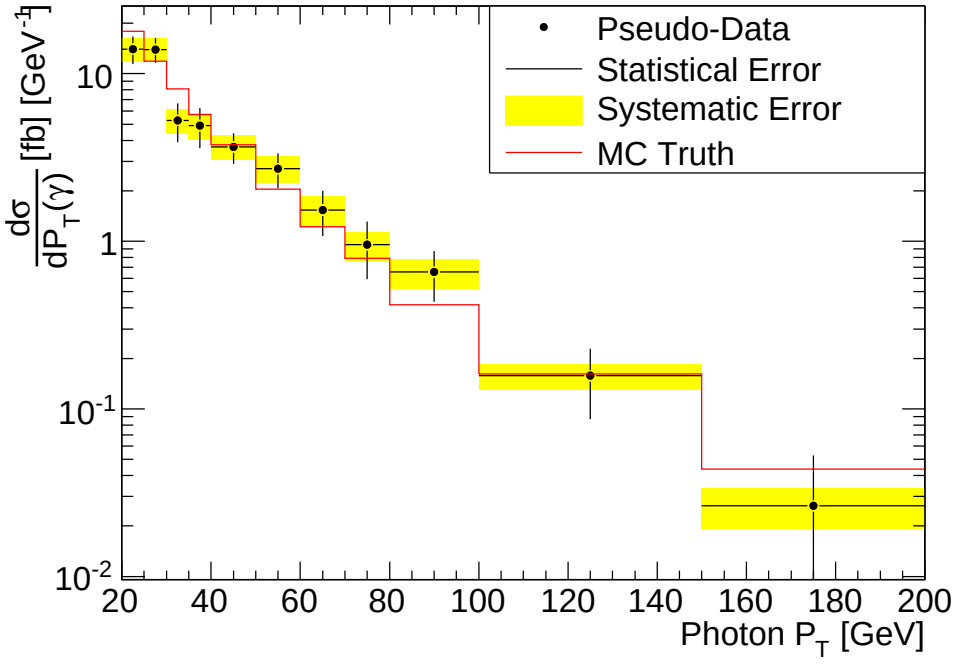
Figure 5.18: (a) Reconstruction purity $\mathcal{P}_i$, in bins of $p_T(\gamma)$. (b) Reconstruction purity $\mathcal{P}_i$, in bins of $M_{ee\gamma}$. Errors are statistical only.

5.6 $p_T(\gamma)$ and $M_{ee\gamma}$ Differential Cross-section Distributions

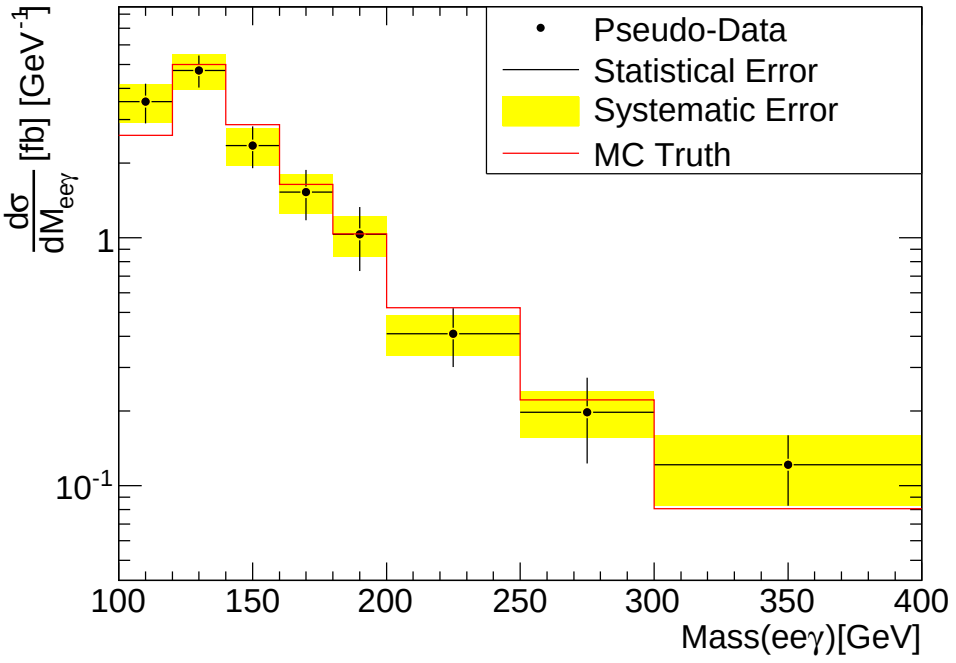
The final stage of the analysis simulation is to apply the correction factors detailed in equation 5.1, bin-by-bin, to the uncorrected measured (pseudo-data) distributions presented in figure 5.16. In order to verify the self-consistency of the analysis chain, the corrected pseudo-data sample is compared to an independent Monte Carlo generator level distribution (no detector simulation), created using the same Pythia matrix element (this is the theoretical prediction); this acts as a closure test on the analysis. To enter the generator level distribution, generated events are required to contain electrons and photons within the fiducial acceptance criteria: $|\eta| < 2.4$, $p_T > 20\,\text{GeV}$, $\Delta R(e,\gamma) > 0.4$, and $M_{ee} > 60\,\text{GeV}$. These derived generator level distributions are then corrected for luminosity and bin width, to be comparable to the corrected pseudo-data distributions. Figure 5.19 shows the pseudo-data/MC comparison for the photon p_T, and $M_{ee\gamma}$ invariant mass, differential cross-section measurements. The systematic errors shown in these plots are discussed in the following sections.

5.7 Systematic Uncertainties

The measured distributions are subject to uncertainty from a number of systematic effects, of varying importance. The systematic effects which are expected to have the largest effect on the measurement uncertainty (EM Energy Scale, p_T Resolution, particle ID efficiency, Jet Fake uncertainty, Monte-Carlo statistics, Luminosity uncertainty, and theoretical uncertainty) are investigated in this section.



(a)



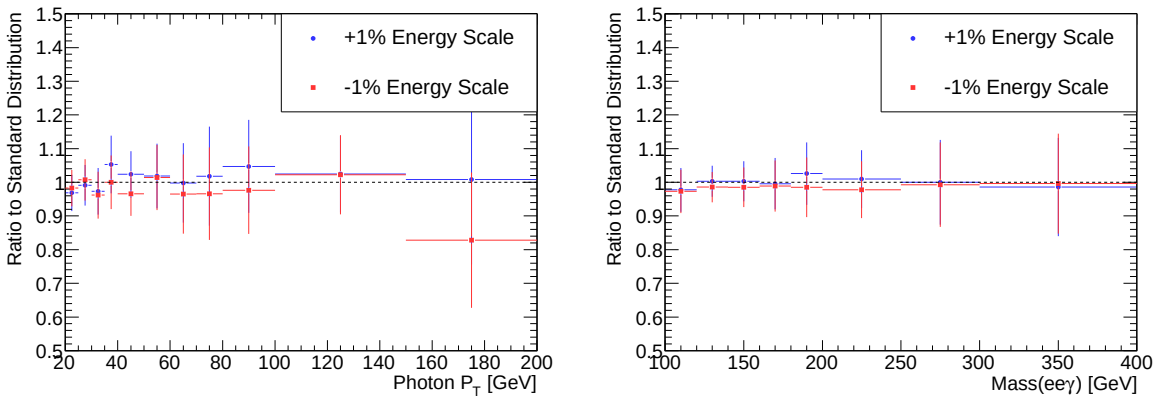
(b)

Figure 5.19: (a) Photon p_T differential cross-section from pseudo-data, versus MC truth, for an integrated luminosity of $\mathcal{L} = 1fb^{-1}$. (b) $ee\gamma$ invariant mass distribution.

5.7.1 EM Energy Scale Systematic

It is expected that the EM energy scale will be known to a precision of 1% or better by the time that 1 fb^{-1} of data has been collected[48]. This will be done using the $Z \rightarrow ee$ process, where the Z mass is known very precisely from previous experiments[53], correcting the energy scale to reproduce the precisely known quantity. In order to estimate the effect of such an uncertainty in the energy scale, the Monte Carlo data was re-analysed, shifting the energy of every electron and photon by $\pm 1\%$, before any fiducial acceptances or cuts were applied. The full analysis chain was then applied to the energy shifted particles, resulting in new predictions for the differential cross-section distributions (the efficiency and purity correction factors used were those calculated in sections 5.5.1 and 5.5.2 using the un-shifted electrons and photons), which were then compared to the default (no energy shift) case.

Figure 5.20 shows the ratio of the shifted distributions to the original distributions. The large statistical errors are due to the small number of background photons entering the measurements from the large $Z \rightarrow ee$ background sample; the statistical fluctuations are considered as a systematic uncertainty separately in section 5.7.5. The total error from this source is below about $\approx 5\%$ for the $p_T(\gamma)$ spectrum, and below about $\approx 3\%$ for the $M_{ee\gamma}$ spectrum.



(a) (b)

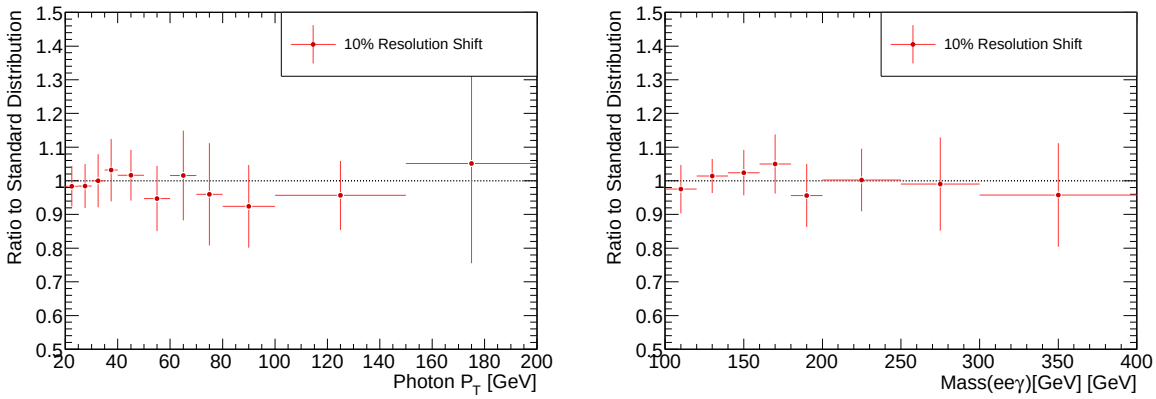
Figure 5.20: (a) Systematic effect of a $\pm 1\%$ uncertainty in EM energy scale, binned in $p_T(\gamma)$. (b) $M_{ee\gamma}$ distribution.

5.7.2 p_T Resolution Systematic

The p_T resolution of electrons and photons is estimated to be known with an uncertainty of 10% by the time $1\,\text{fb}^{-1}$ integrated luminosity has been recorded (this can again be calculated using previous Z boson measurements)[48]. In order to estimate the effect this uncertainty has on the measured distributions, an analysis was conducted where for every truth matched electron and photon, the reconstructed p_T was recalculated thus:

$$p'_T = p_T(Truth) + (1.1 \times [p_T(Reco) - p_T(Truth)]) \tag{5.7}$$

Following this, new efficiency and purity correction factors were derived, and the analysis re-run using the original p_T values, but with the resolution shifted correction factors. Figure 5.21 shows the ratio of the resolution adjusted differential cross-section distributions, to the standard case. As mentioned previously, the statistical fluctuations are large; these are dealt with as a separate uncertainty is section 5.7.5. The uncertainty associated with a 10% p_T resolution shift on the measured $p_T(\gamma)$ and $M_{ee\gamma}$ distributions is about $\approx 5\%$ across the p_T and $M_{ee\gamma}$ ranges.



(a) (b)

Figure 5.21: (a) Systematic effect of a 10% uncertainty in p_T resolution for electrons and photons, binned in $p_T(\gamma)$. (b) $M_{ee\gamma}$ distribution.

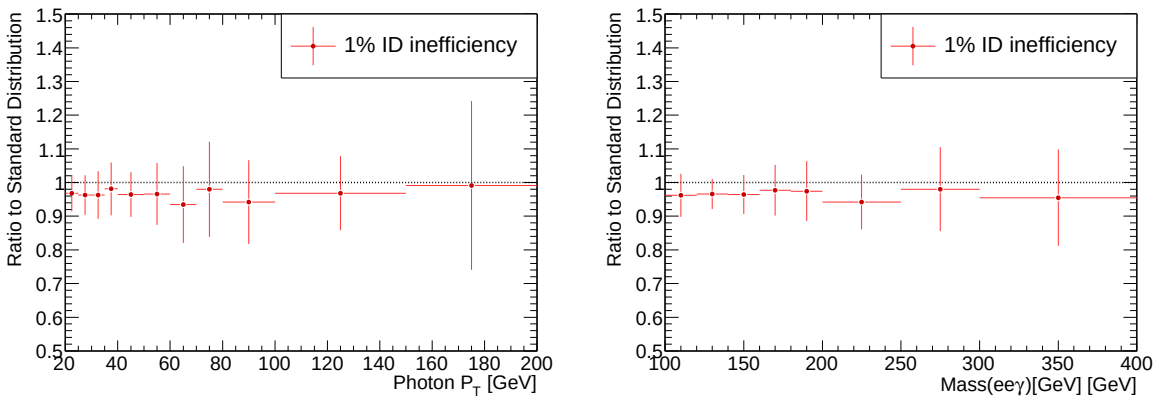
5.7.3 Particle ID Efficiency Systematic

An error on the efficiency of 1% or better is expected for the particle ID of electrons and photons by the time that $\mathcal{L} = 1\,\text{fb}^{-1}$ integrated luminosity has been recorded[48]. For electrons, this can be measured from data using a tag-and-probe method[54]. The method can be applied to a sample containing two objects available for selection (e.g. the two electrons in $Z \to ee$ events), one electron is selected and acts as a *tag*. The tag electron helps to select a pure sample; the other electron in the event is the *probe*, this is not required to pass identification cuts. The reconstruction efficiency can be calculated by considering the number of times that the probe electron is correctly identified by the reconstruction software. For photons, the reconstruction efficiency is calculated based on the similarity between electron and photon signatures in the detector(s). The efficiency here must be calculated using Monte Carlo techniques, since an equivalent channel to $Z \to ee$ does not exist to enable the application of a tag-and-probe method for photons.

The uncertainty on the differential cross-section measurements caused by this inefficiency is calculated by rejecting a random 1% of all identified electrons and photons, and re-running the full analysis chain (using the efficiency and purity factors previously calculated with no inefficiency applied). Figure 5.22 shows the ratio of differential cross-section measurements for the case of a 1% inefficiency in particle ID, with respect to the standard distributions. Again, the statistical fluctuations are large; as previously mentioned these are considered independantly in section 5.7.5. The uncertainty on the measured $p_T(\gamma)$ and $M_{ee\gamma}$ distributions, due to a 1% inefficiency in electron/photon ID, can be seen to be consistent with the 3% that is expected (multiplying the 1% inefficiency, by the number of reconstructed particles).

5.7.4 Jet Fake Uncertainty Systematic

The electron and photon fake rates are expected to have an uncertainty of 20% with $1\,\text{fb}^{-1}$ integrated luminosity [48]. This can be measured by looking in the side-bands



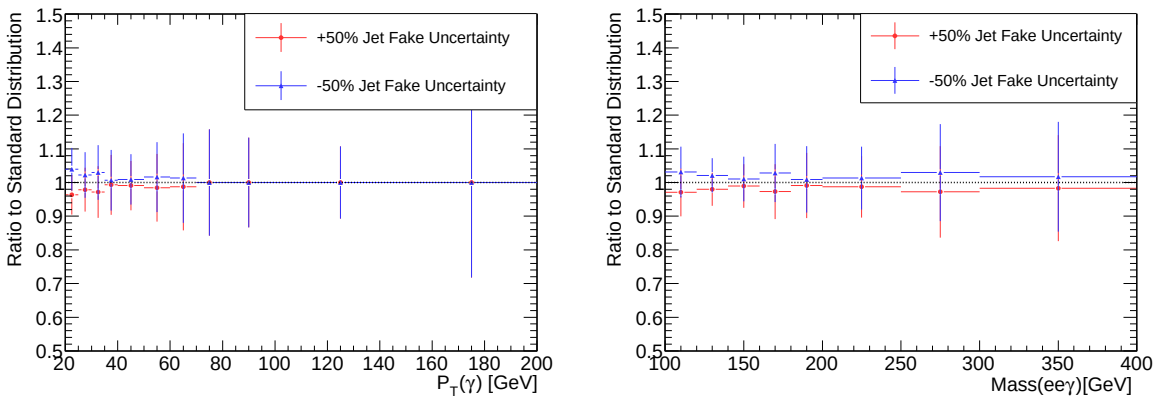
(a) (b)

Figure 5.22: (a) Systematic effect of a 1% inefficiency in particle ID for electrons and photons, binned in $p_T(\gamma)$. (b) $M_{ee\gamma}$ distribution.

of some experimental variable distribution, and comparing Monte Carlo predictions of signal and background (jet) contributions, to the measured data. Assuming a pessimistic 50% uncertainty on the electron and photon fake rates, the combined uncertainty on the QCD multi-jet background (see section 5.3.1) would still be insignificant, because this background is so low relative to the signal distribution.

Since fake photons in Z events form a sizeable background to the differential cross-section measurements, the effect of a 50% uncertainty in the photon fake rate is investigated. The distributions from fake-photons in Z events passing the signal selection are scaled by $\pm 50\%$, and new purity correction factors are derived using the scaled distributions. The standard analysis is then re-run using the adjusted purity factors; figure 5.23 shows the ratio of the adjusted differential cross-section measurements to the standard (un-scaled) distributions.

Despite statistical limitations (see section 5.7.5), the total uncertainty on the differential $p_T(\gamma)$ and $M_{ee\gamma}$ cross section measurements due to a 50% uncertainty in the photon fake rate from jets is below about $\approx 5\%$ in total, across the range of $p_T(\gamma)$ and $M_{ee\gamma}$ values.



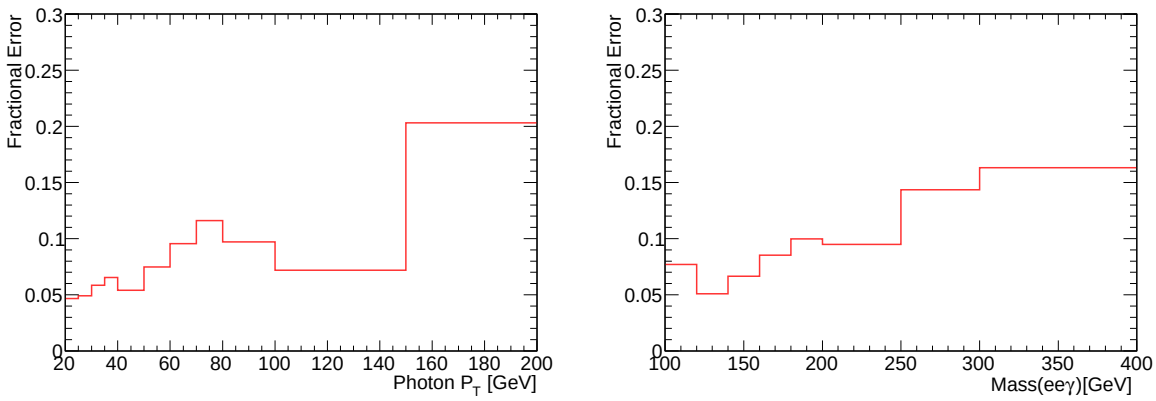
(a) (b)

Figure 5.23: (a) Systematic effect of a ±50% uncertainty in the photon fake-rate, binned in $p_T(\gamma)$. (b) $M_{ee\gamma}$ distribution.

5.7.5 Monte Carlo Statistics Systematic

An uncertainty in the differential cross-section measurements (and the systematic uncertainties derived in the preceeding sections) is introduced by the limited statistics available in the Monte Carlo samples used to derive the efficiency and purity correction factors. The size of this uncertainty is here estimated from the addition in quadrature of the statistical errors (bin-by-bin) in the efficiency and purity factor distributions. The statistical uncertainties on the efficiency and purity distributions are treated individually as binomial. Figure 5.24 shows the relative uncertainty on the differential cross-section measurements introduced by the limited Monte-Carlo statistics available.

The total error introduced to the differential $p_T(\gamma)$ and $M_{ee\gamma}$ cross section measurements from the limited Monte Carlo statistics is between 5 and 10% for most bins, growing to between 15 and 20% in the bin with the lowest number of expected events. It should be noted that this systematic effect is reducible; with the simulation of more events, the uncertainty due to the limited statistics will fall.



(a) (b)

Figure 5.24: (a) Systematic effect of Monte-Carlo statistics in $p_T(\gamma)$ distribution. (b) $M_{ee\gamma}$ distribution.

5.7.6 Luminosity Uncertainty

At the time of writing [5], the luminosity measurement in ATLAS with 7 TeV collisions has been calibrated using beam separation (Van Der Meer) scans with a measured uncertainty of 11%[55]. This method does not rely on a priori knowledge of physics cross-sections, unlike other methods, and has proven useful in the early running period where the theoretical uncertainties on the cross-sections used by other methods are too large to provide competitive alternative measurements. Accordingly, an 11% uncertainty on the luminosity is adopted, which is a pessimistic scenario given that with $\mathcal{L} = 1\,\text{fb}^{-1}$ integrated luminosity this uncertainty will have reduced.

5.7.7 Theoretical Uncertainty

The choice of Monte Carlo generator sets the precise values of event reconstruction efficiency, and purity, in each bin of the measured distributions; as well as predicting the exact nature of bin migrations. Different generators may predict different acceptances, and thus a different final (efficiency and purity corrected) distribution. As mentioned in section 1.6, there are a number of available Monte Carlo generators for simulating the $Z\gamma$

[5] by November 2010 an integrated luminosity of $\mathcal{L} \approx 45\,\text{pb}^{-1}$ has been recorded by ATLAS.

process, and for each generator a number of configurable parameters that must be set.

The generators differ in a number of ways, including whether they are LO, NLO, or NLL[6]; whether they include interference between ISR and FSR diagrams or not; how they deal with non-perturbative processes (fragmentation, hadronisation); how the matrix element is matched to the parton shower, e.t.c. Also important at the LHC is the choice of Parton Density Function (see section 1.5.2), since there are a number of available sets, tuned to different experimental results, and at different orders in the coupling constant α_s (these are designed to match the order at which the hard process is calculated). A comprehensive analysis of the various elements that make up the theoretical uncertainty is beyond the scope of this study, but a study comparing the results from an alternative Monte Carlo generator is presented here.

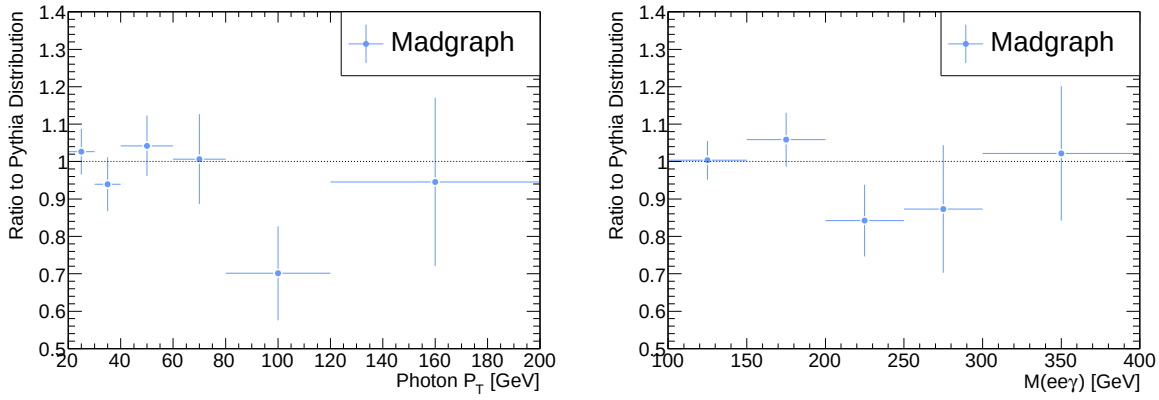
Madgraph Generator Comparison

An investigation into the theoretical uncertainty has been conducted by comparing the differential cross-section distributions as predicted by Pythia, to those predicted by the generator Madgraph[56]. One difference between the two samples is the use of different PDF sets; the Pythia ISR sample uses the MRST[57] set, and Madgraph uses the CTEQ6L1 set[12]. In addition, Madgraph computes the process $pp \rightarrow ee\gamma$, including the ISR diagram, FSR diagram and the interference term; whereas Pythia only computes the ISR diagram.

For the Madgraph sample, reconstructed ISR photons were identified by comparison of a proximity matched truth photon ($\Delta R < 0.1$) to the Monte Carlo truth record. Signal reconstruction efficiency and purity distributions, in accordance with the procedure presented for the Pythia sample (sections 5.5.1 and 5.5.2) were derived. The same background sample (Pythia $Z \rightarrow ee$) was used in the derivation of the purity distribution for both the Madgraph and Pythia signal (ISR) samples. The differential cross-section

[6]NLL = Next-to-Leading-Log - the higher order logarithmic correction terms only; essentially the virtual corrections (no emissions).

distributions calculated using the Madgraph sample were normalised to the Pythia distribution, since it is the difference in shape which is of importance here. Due to the limited statistics available for the Madgraph sample, the binning chosen for each distribution is wider than for the original Pythia prediction. Figures 5.25(a) and 5.25(b) show the ratio of the Madgraph differential cross-section distributions to the Pythia reference.



(a) (b)

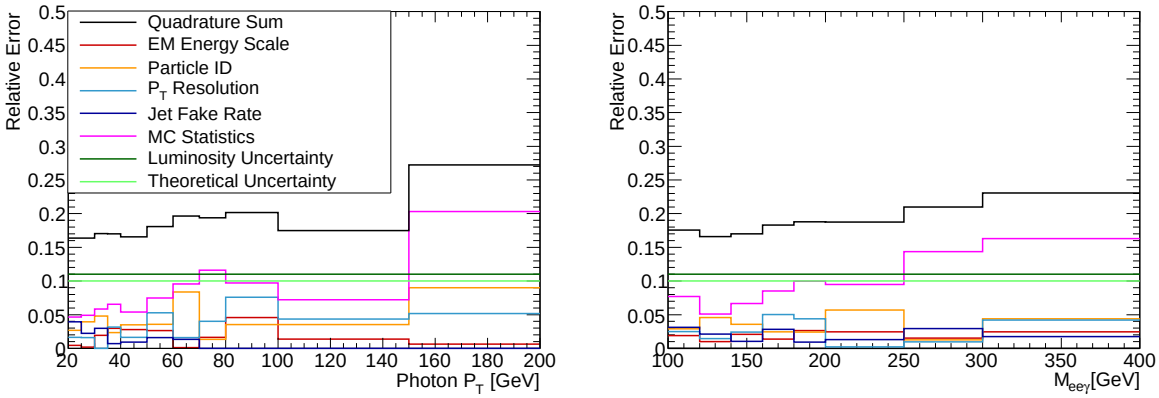
Figure 5.25: (a) The ratio of Madgraph to Pythia ISR differential cross-section predictions, in $p_T(\gamma)$. (b) Ratio of $M_{ee\gamma}$ distributions, statistical errors only.

Figure 5.25 shows that there is no marked systematic difference in shape of the differential cross-section predictions in either $p_T(\gamma)$ or $M_{ee\gamma}$. Although the study is limited by the statistics available, an uncertainty of 10% is adopted to represent the combined theoretical uncertainty; this is comparable to the uncertainty suggested by a previous study[22]. Clearly, with more statistics available from different generator samples, as well as a study on similar effects in the background samples, a more rigorous constraint on the theoretical uncertainty can be imposed.

5.7.8 Combined Systematic Uncertainty

The different contributions of systematic uncertainty presented in the preceeding sections may be added together in quadrature to provide an overall estimate of the level of systematic uncertainty in each measurement. This procedure is correct if the different contributions are uncorrelated, and is at worst an overestimate, if it is found that there

is some correlation between the different contributions. Figure 5.26 shows the relative error of each of the systematic effects which have been investigated, along with the combined uncertainty calculated from the addition in quadrature of the individual effects. From inspection of figure 5.26, the overall systematic uncertainty could most readily be reduced by an improvement in the available Monte Carlo statistics, by generating more signal events. The combined systematic uncertainties are shown relative to the $p_T(\gamma)$ and $M_{ee\gamma}$ distributions in figure 5.19; where it can be seen that the combined systematic uncertainties are of a similar order to, or smaller than, the statistical uncertainties in each bin.



(a) (b)

Figure 5.26: (a) Breakdown of systematic uncertainties, and quadratic total sum for $p_T(\gamma)$ distribution. (b) $M_{ee\gamma}$ distribution.

5.8 New Particle Decays to $Z\gamma$

As mentioned in section 1.6.2, the Standard Model in its current form is not the final (complete) theory of particle physics. The $pp \to Z(ee)\gamma$ measurements presented in the previous section can form the basis of a search for some new physics process, since it is possible that some new particle may be produced at LHC energies and decay into the same final state ($ee\gamma$), thus contributing to the measured distribution, and causing a deviation from the Standard Model prediction. The most sensitive distribution for the observation of such a decay will be the $M_{ee\gamma}$ invariant mass distribution.

The predicted cross-sections times branching ratios of specific models of a production and decay (such as the SM Higgs scalar[58], or vector techniparticle (ω_T, ρ_T) [59]) are too low for these processes to be observed, with an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$ in ATLAS. What can be done however is to consider more general cases where the cross-section times branching ratio is not known, but these models are used to predict the kinematic properties of production and decay. Exclusion limits on the cross-sections of these processes can then be calculated from the measured distribution. Specifically, if the data are consistent with the Standard Model predictions, exclusion confidence limits can be placed. In the following, the previously held notation is adapted, and the term "signal" is defined as a putative scalar or vector decay contribution, with the term "background" denoting the combined Standard Model contributions from ISR, FSR, and fake $Z\gamma$ events, passing the previously defined selection cuts.

5.8.1 Monte Carlo Samples

The scalar particle hypothesis is simulated by the standard model Higgs boson process using the Pythia Monte Carlo event generator, where only decays of the Higgs particle to $Z(ee) + \gamma$ are kept. This sample is not ideal, since the variation of the particles width is peculiar to this model, but is adopted in the absence of a more apropriate sample at the time of writing. For the vector particle process, the low-scale technicolor (LSTC) straw-man model[60] implemented in Pythia was used to produce the techni-vectors ω_T, ρ_T, and their decay into the $Z(ee)+\gamma$ final state. The Pythia generator parameters relating to the masses and widths of the LSTC particles can be found in appendix A. Figure 5.27 shows the corresponding Feynman diagrams for the decay of these scalar and vector particles to the $ee\gamma$ final state.

A fast simulation of the ATLAS detector was used to simulate the signal processes. A total of $5,000$ events were simulated at each of the mass points 200, 300, 400, 500, 600, 700, 800 and 900 GeV, for both the scalar and vector processes. The background

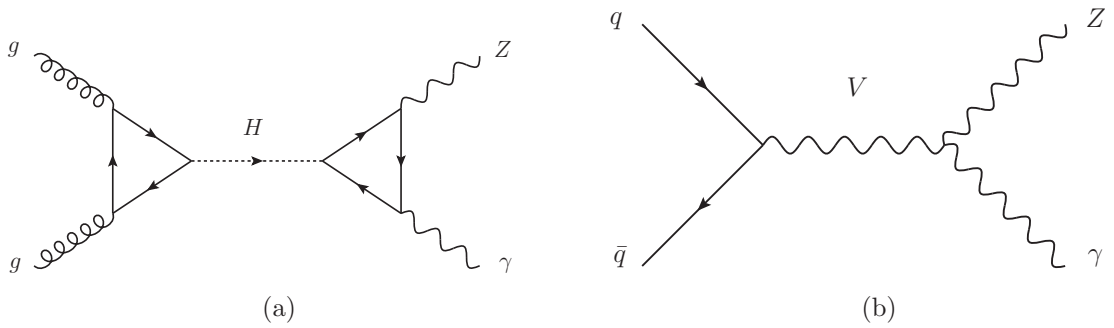


Figure 5.27: (a) Example Feynman diagram of production and decay of SM Higgs scalar to $Z\gamma$ final state. (b) Example Feynman diagram of some generic vector particle, V, decaying to $Z\gamma$ final state.

to these processes are the Standard Model $ee\gamma$ (ISR, FSR, Fake) events already studied; accordingly, the same samples described in section 5.2.1 are used here. Any interference term between the Signal (new particle) and background (SM $Z\gamma$) processes is neglected.

5.8.2 Event Selection and Reconstruction Efficiency

Events are required to pass the same selection criteria as for the Standard Model $Z\gamma$ ISR signal selection described in section 5.4, that is: two electrons (`MediumIso`, $p_T > 20\,\text{GeV}$, $|\eta| < 2.4$, $M_{ee} > 60\,\text{GeV}$), and a photon (`Tight` with isolation, $p_T > 20\,\text{GeV}$, $|\eta| < 2.4$), separated by $\Delta R(e, \gamma) > 0.4$, and the invariant mass cut $M_{ee\gamma} > (2M_Z - M_{ee})$. The binning regime of the $M_{ee\gamma}$ distribution was chosen according to the predicted resolution for the signal processes. Figure 5.28 shows the reconstructed $M_{ee\gamma}$ shift for the Higgs decay channel, with contributions from all generated mass points. The chosen bin widths are 20 GeV in the region [100 - 500 GeV], and 50 GeV in the region [500 - 1000 GeV].

Figure 5.29 shows an example reconstructed $M_{ee\gamma}$ distribution for a 700 GeV Higgs scalar, and vector particle, where it can be seen that the Pythia Higgs model predicts a wide distribution compared to the narrow width of the vector particle. This larger width at high masses is a feature of the specific SM Higgs model, as implemented in Pythia, and should be considered when comparing the final results.

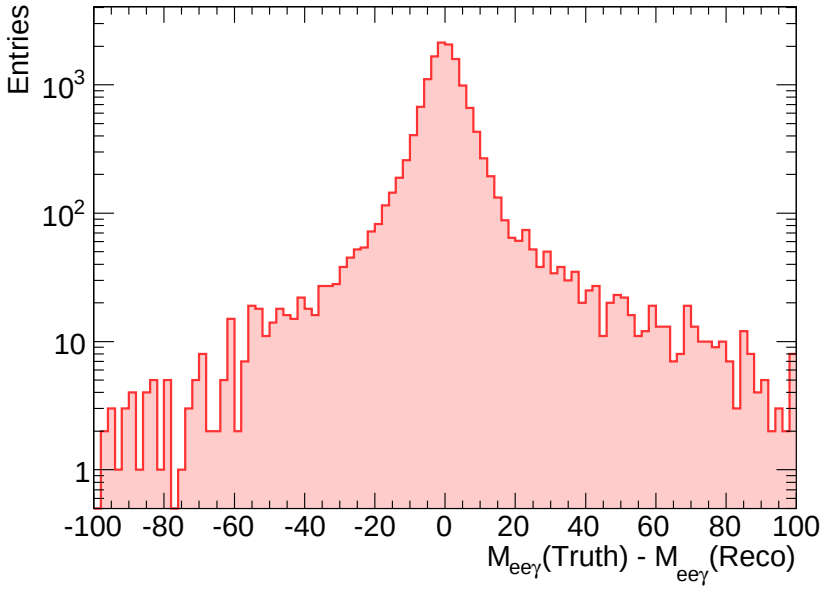


Figure 5.28: Reconstructed $M_{ee\gamma}$ shift relative to generated mass, for SM Higgs decays to $Z\gamma$, for all generated mass points.

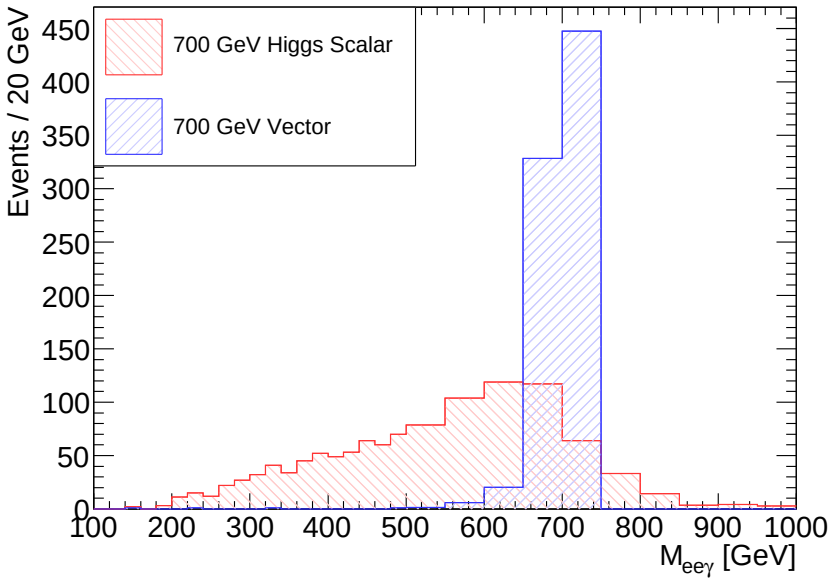


Figure 5.29: Reconstructed $M_{ee\gamma}$ for a 700 GeV scalar Higgs, and vector particle decays, with 5,000 events.

Reconstruction Efficiency

The global signal reconstruction efficiency is defined as the total number of well reconstructed signal events passing all selection cuts, divided by the total number of generated events (truth) expected in the same fiducial acceptance. This has been calculated for each of the vector and scalar mass points, and can be seen in figure 5.30; where it is observed that the efficiency for the vector hypothesis is $\approx 5\%$ higher on average (this similarity is despite the larger predicted width of the Higgs scalar at high masses in the Pythia model).

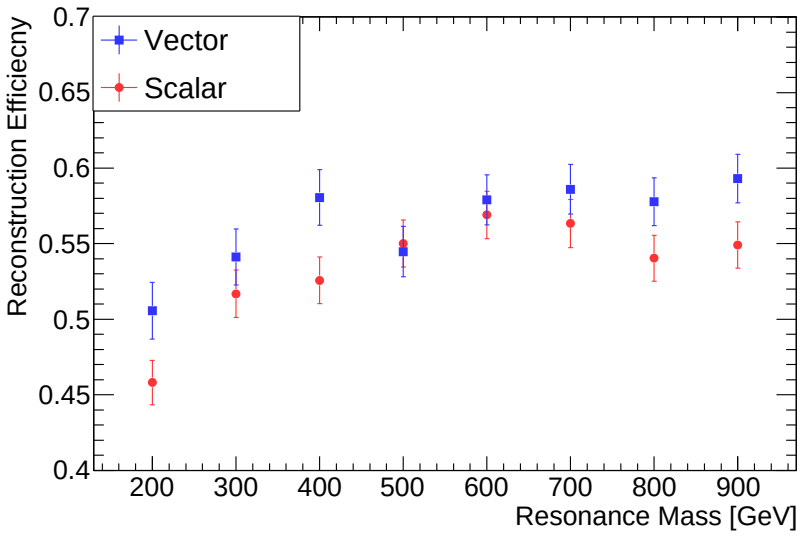


Figure 5.30: Global signal reconstruction efficiencies as a function of particle mass, for vector and Higgs-like scalar particles.

5.8.3 Exclusion Limit Setting Software

In order to calculate the exclusion limits for an observation of a new (Higgs-like scalar, or vector) particle decay to $Z\gamma$, given the background only hypothesis of the standard model $Z\gamma$ distribution presented in section 5.6, a modified frequentist statistical method is employed [61]. The signal-like outcomes are discriminated from background-like outcomes using a Poisson likelihood ratio test statistic X:

$$X = \prod_{n}^{i=1} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{d_i}}{d_i!} \Big/ \frac{e^{-b_i} b_i^{d_i}}{d_i!}, \tag{5.8}$$

where n is the number of bins i, s_i is the estimated signal in bin i, b_i is the estimated background in bin i, and d_i is the number of observed candidates (data) in bin i. It is the Log-Likelihood Ratio (LLR) which is then calculated, which is defined as:

$$LLR = -2\ln X = 2\sum_{i=1}^{n} s_i - 2\sum_{i=1}^{n} d_i \ln\left(1 + \frac{s_i}{b_i}\right). \tag{5.9}$$

The confidence level (CL) for excluding the possibility of a new particle signal on top of the background (the $s+b$ hypothesis), is defined as the probability that the test statistic X, is equal to or less than that observed in data X_{obs}, given the hypothesised levels of signal and background:

$$CL_{s+b} = P_{s+b}(X \leq X_{obs}), \tag{5.10}$$

where P_{s+b} is the sum of Poisson probabilities:

$$P_{s+b}(X \leq X_{obs}) = \sum_{X(\{d_i'\}) \leq X(\{d_i\})} \prod_{i=1}^{n} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{d_i'}}{d_i'!}. \tag{5.11}$$

Here $X(\{d_i\})$ is the test statistic computed for the observed set of candidates in each channel $\{d_i\}$, and the sum runs over all possible outcomes $\{d_i'\}$ which have test statistics less than or equal to the observed one. The modified frequentist CL is then given by the ratio of the signal plus background hypothesis, to the background only hypothesis, $CL_s = CL_{s+b}/CL_b$. The new physics hypothesis is excluded at 95% if $CL_s = 0.05$, i.e. $CL = 1 - CL_s$.

In practice, the computation of the confidence level is performed with the `mclimit` [62] software, within the `ROOT`[42] framework. The program runs many thousands of pseudo-experiments in each bin (each bin is treated as a separate counting experiment), in order to calculate the sum in equation 5.11. Since there are no measured data available, the pseudo-data sample[7] used in figure 5.19(b) is adopted as d_i. In order to calculate the exclusion limits on cross-section times branching ratio, the program scales the s_i (new

[7]This sample contains ISR, FSR and fake events mixed in proportion to their cross-sections to represent an expected yield of Standard Model events (i.e. background only hypothesis) for an integrated luminosity $\mathcal{L} = 1\,\text{fb}^{-1}$.

particle) distributions freely, repeating the CL_s calculation until a scale factor is found which just excludes the new particle $(s+b)$ hypothesis at 95% CL; this scale factor gives the limit.

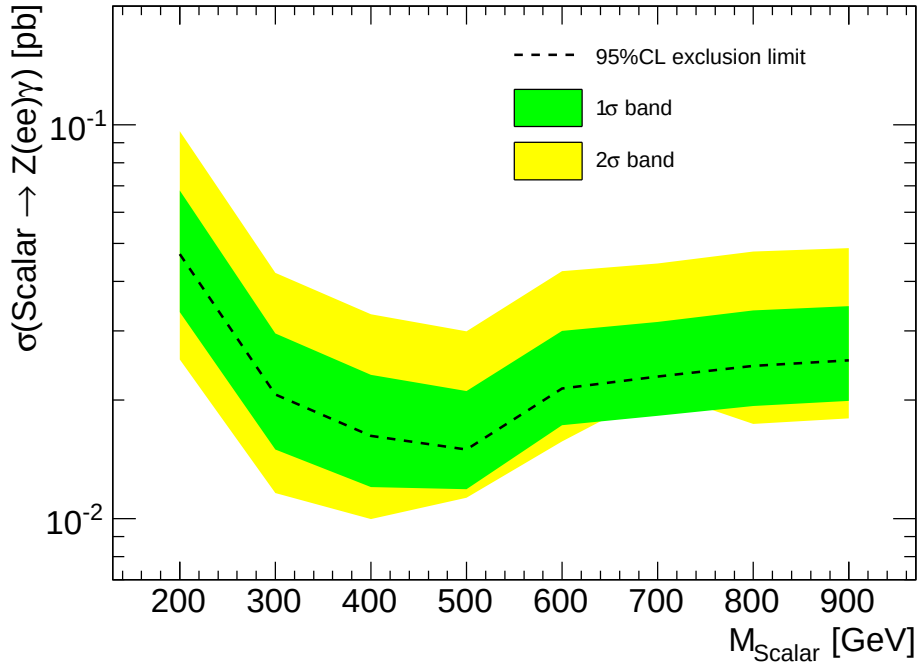
Systematic Uncertainties

Systematic uncertainties are incorporated into the calculation in the form of gaussian-distributed nuisance parameters. The program categorises these uncertainties into those which are a multiplicative scale factor, and those which change the shape of the distribution. The program also allows for the specification of correlation of uncertainties between the signal and background channels. For the purposes of this study, the following uncertainties are adopted.

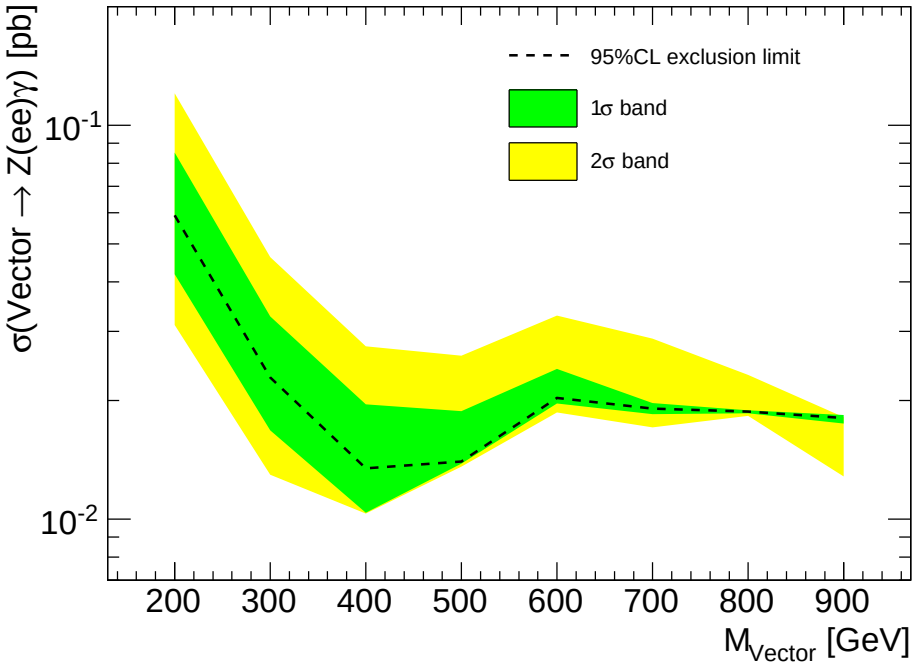
- A scale error of 11% from luminosity uncertainty, for both signal and background distributions, with 100% correlation.
- A scale error of 11% from combined systematic effects (theoretical uncertainty, EM energy scale, p_T resolution, jet fake uncertainty, particle ID inefficiency) for the background sample only.
- A scale error of 10% for combined systematic effects (theoretical uncertainty, ID inefficiency) on the signal sample only.
- Shape errors for the signal sample representing the effects of a 1% EM energy scale uncertainty; these are dealt with by supplying histogram templates of the shifted signal distributions accounting for the uncertainty (see section 5.7.1).
- Shape error for the signal sample representing the effects of a 10% resolution uncertainty (see section 5.7.2).

5.8.4 Results

The exclusion limits at 95% CL calculated for the vector and scalar particle hypotheses, along with the 1σ, and 2σ uncertainty bands, are shown in figure 5.31. At high masses (above 700 GeV) in the vector case, the uncertainty bands collapse; this is because there are no data points (background only hypothesis) in the signal region. The same effect is not seen in the scalar hypothesis, because the specific Higgs model predicts large widths over this mass region, and the signal distribution does cover an area where there are (background) data points. The results of this simulation study compare favourably to previous results derived with the DØ experiment at the Tevatron, [29], albeit in a different $\sqrt{s}$ region, and with different fiducial acceptance criteria. This study shows an improvement of $\mathcal{O}(10)$ - $\mathcal{O}(100)$ with respect to the corresponding cross-section times branching ratio limits achieved at DØ.



(a) Higgs-like Scalar particle



(b) Vector particle.

Figure 5.31: (a) Predicted 95% CL exclusion limit on cross-section times branching ratio for a Higgs-like scalar particle decay to $Z(ee)\gamma$ final state as implemented in Pythia, for $\mathcal{L} = 1\,fb^{-1}$ integrated luminosity. (b) Same for the vector particle hypothesis.

Chapter 6

Conclusion

The ATLAS trigger system has been reliably selecting interesting physics events thoughout the data-taking period of ATLAS[63]. The first part of the work reported in this thesis was concerned with the digital timing calibration of the Cluster Processor subsystem of the level-1 calorimeter trigger. A procedure was developed, debugged, and implemented which allowed for the calibration of the digital timing parameters of the Cluster Processor system, an essential part of providing a robust and stable level-1 trigger operation. The calibration procedures are now deployed, and successfully in use, in the running trigger system.

The next focus of work was in the measurement of $Z(\rightarrow ee) + \gamma$ final states with ATLAS. Due to the operational calendar of the LHC, a Monte Carlo simulation based study of the potential of ATLAS to measure this process was undertaken. The study focused on what would be achievable in a period of early running, with an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$, at a proton-proton centre of mass collision energy of $\sqrt{s} = 7\,\text{TeV}$. The ISR process was adopted as the signal process, and measurements of the differential cross-section binned in both the transverse momentum of the photon, $p_T(\gamma)$, and three-body $ee\gamma$ invariant mass, $M_{ee\gamma}$, were derived using a bin-by-bin correction method. This involved a study of the different background contributions, and a survey of the various

systematic uncertainties which will affect the accuracy of the final measurements. It was found that the systematic uncertainties are expected to be of a similar order to the statistical uncertainties with this amount of collected data; the combined systematic uncertainty could most readily be reduced by simulating more Monte Carlo events (both signal and background).

Finally, the $M_{ee\gamma}$ measurement was used as the basis for a study of hypothetical new particles decaying to the $ee\gamma$ final state. The ability of ATLAS to reconstruct a SM Higgs-like scalar, and vector particle decay to the $Z(ee)\gamma$ state was explored, for particles in the mass range $200 - 900\,\text{GeV}$. Following this, exclusion limits were placed on the cross-section times branching ratio of such new processes, as a function of their mass, assuming that the (simulated) data collected fits the Standard Model hypothesis. The results of the study indicate that ATLAS should vastly improve on the sensitivity to new physics in this channel, even with $1\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$, over that achieved previously at the Tevatron collider[29].

Appendix A

LSTC Parameters in Pythia

Table A.1 summarises the generation parameters specific to the LSTC model, that were set in Pythia for the production of LSTC techniparticles used for the study in section 5.8.

	Kinematic Range [GeV]		Mass [GeV]							
Mass Point [GeV]	`ckin 1`	`ckin 2`	π_T^0	$\pi_T^\pm$	π_T'	a_T^0	$a_T^\pm$	ρ_T^0	$\rho_T^\pm$	ω_T
200	190	210	150	150	300	220	220	200	200	200
300	290	310	200	200	400	330	330	300	300	300
400	390	410	250	250	500	440	440	400	400	400
500	490	510	300	300	600	550	550	500	500	500
600	590	610	350	350	700	660	660	600	600	600
700	690	710	400	400	800	770	770	700	700	700
800	790	810	450	450	900	880	880	800	800	800
900	890	910	500	500	1000	990	990	900	900	900

Table A.1: Parameters set in Pythia for the LSTC model.

Bibliography

[1] Particle Data Group, K. Nakamura, Review of particle physics, J. Phys. **G37**, 075021 (2010).

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, 1995).

[3] M. Kramer and Soler., Large hadron collider phenomenology, proceedings of the fifty seventh Scottish Universities Summer School in Physics, St. Andrews, 17 August to 29 August 2003, 57th Scottish Universties Summer School in Physics: LHC Phenomenology (SUSSP 2003) 17-29 Aug 2003, St. Andrews, Scotland, United Kingdom.

[4] M. Kaku, *Quantum field theory: A Modern introduction* (Oxford Univ. Pr, 1993).

[5] S. L. Glashow, Partial Symmetries of Weak Interactions, Nucl. Phys. **22**, 579 (1961).

[6] S. Weinberg, A Model of Leptons, Phys. Rev. Lett. **19**, 1264 (1967).

[7] A. Salam and J. C. Ward, Electromagnetic and weak interactions, Phys. Lett. **13**, 168 (1964).

[8] A. Pich, The Standard model of electroweak interactions, (2005), hep-ph/0502010.

[9] B. W. Lee, C. Quigg, and H. B. Thacker, Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass, Phys. Rev. **D16**, 1519 (1977).

[10] CDF and D0, B. Penning, Combined upper limit for SM Higgs at the Tevatron, PoS **EPS-HEP2009**, 231 (2009).

[11] J. D. Bjorken and E. A. Paschos, Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon, Phys. Rev. **185**, 1975 (1969).

[12] J. Pumplin *et al.*, New generation of parton distributions with uncertainties from global QCD analysis, JHEP **07**, 012 (2002), hep-ph/0201195.

[13] M. A. Dobbs *et al.*, Les Houches guidebook to Monte Carlo generators for hadron collider physics, (2004), hep-ph/0403045.

[14] T. Sjöstrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP **05**, 026 (2006), hep-ph/0603175.

[15] T. Gleisberg *et al.*, Event generation with SHERPA 1.1, JHEP **02**, 007 (2009), 0811.4622.

[16] D. De Florian and A. Signer, W gamma and Z gamma production at hadron colliders, Eur. Phys. J. **C16**, 105 (2000), hep-ph/0002138.

[17] U. Baur, T. Han, and J. Ohnemus, QCD corrections and anomalous couplings in $Z\gamma$ production at hadron colliders, Phys. Rev. **D57**, 2823 (1998), hep-ph/9710416.

[18] U. Baur and E. L. Berger, Probing the weak-boson sector in $Z\gamma$ production at hadron colliders, Phys. Rev. D **47**, 4889 (1993).

[19] G. Brooijmans *et al.*, New Physics at the LHC. A Les Houches Report: Physics at TeV Colliders 2009 - New Physics Working Group, (2010), 1005.1229.

[20] J. P. Skittrall, Production of a Z boson and a photon via a Randall- Sundrum-type graviton at the Large Hadron Collider, Eur. Phys. J. **C60**, 291 (2009), 0809.4383.

[21] A. Oh, Limits on anomalous neutral current triple gauge boson couplings, PoS **hep2001**, 123 (2001).

[22] S. Hassani, CERN Report No. ATL-PHYS-2003-023, 2002 (unpublished).

[23] UA1, G. Arnison *et al.*, Experimental observation of lepton pairs of invariant mass around 95-GeV/c**2 at the CERN SPS collider, Phys. Lett. **B126**, 398 (1983).

[24] UA2, P. Bagnaia *et al.*, Evidence for Z0 → e+ e- at the CERN anti-p p collider, Phys. Lett. **B129**, 130 (1983).

[25] OPAL, G. Abbiendi *et al.*, Determination of the LEP beam energy using radiative fermion-pair events, Phys. Lett. **B604**, 31 (2004), hep-ex/0408130.

[26] ALEPH, J. Alcaraz *et al.*, A Combination of preliminary electroweak measurements and constraints on the standard model, (2006), hep-ex/0612034.

[27] L3, P. Achard *et al.*, Study of the $e^+e^- \to Z\gamma$ process at LEP and limits on triple neutral-gauge-boson couplings, Phys. Lett. **B597**, 119 (2004), hep-ex/0407012.

[28] CDF and D0, A. L. Lyon, Testing the Standard Model with $W\gamma$ and $Z\gamma$ at the Tevatron, (2008), 0810.3766.

[29] A. V. Ferapontov, Search for particles decaying into $Z\gamma$ at D0, (2008), 0810.1751.

[30] ATLAS, G. Aad *et al.*, The ATLAS Experiment at the CERN Large Hadron Collider, JINST **3**, S08003 (2008).

[31] CMS, R. Adolphi *et al.*, The CMS experiment at the CERN LHC, JINST **3**, S08004 (2008).

[32] LHCb, A. A. Alves *et al.*, The LHCb Detector at the LHC, JINST **3**, S08005 (2008).

[33] ALICE, K. Aamodt *et al.*, The ALICE experiment at the CERN LHC, JINST **3**, S08002 (2008).

[34] http://cdsweb.cern.ch/collection/Photos .

[35] J. Elmsheuser *et al.*, Distributed analysis using GANGA on the EGEE/LCG infrastructure, J. Phys.: Conf. Ser. **119**, 072014 (2008).

[36] P. Hanke, ATLAS Level-1 Calorimeter Trigger Prototype Processor Backplane (version 1.1), (2001), http://hepwww.rl.ac.uk/Atlas-L1/Modules/PPr/PPMod_Wrup.pdf.

[37] S. Silverstein, The Pre-Processor Module (PPM) for the ATLAS Level-1 Calorimeter Trigger, (2009), http://hepwww.rl.ac.uk/Atlas-L1/Modules/Backplane/BackplaneFDR1_1.pdf.

[38] E. F. Eisenhandler, CERN Report No. CERN-ATL-DAQ-2004-011, 2004 (unpublished).

[39] J. R. A. Booth, *Sensitivity to anomalous triple gauge couplings in the W-gamma channel and a Calibration study on the level-1 calorimeter trigger of the ATLAS detector*, PhD thesis, Birmingham U., Birmingham, 2008.

[40] M. Landon *et al.*, Control, test and monitoring software framework for the ATLAS level-1 calorimeter trigger, ATL-DAQ-PROC-2008-004.

[41] F. Viegas, R. Hawkings, and G. Dimitrov, Relational databases for conditions data and event selection in ATLAS, J. Phys.: Conf. Ser. **119**, 042032 (2008).

[42] R. Brun and F. Rademakers, ROOT: An object oriented data analysis framework, Nucl. Instrum. Meth. **A389**, 81 (1997).

[43] GEANT4, S. Agostinelli *et al.*, GEANT4: A simulation toolkit, Nucl. Instrum. Meth. **A506**, 250 (2003).

[44] W. Lampl *et al.*, CERN Report No. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, 2008 (unpublished).

[45] V. Kartvelishvili, Electron bremsstrahlung recovery in ATLAS, Nuclear Physics B - Proceedings Supplements **172**, 208 (2007), Proceedings of the 10th Topical Seminar on Innovative Particle and Radiation Detectors, Proceedings of the 10th Topical Seminar on Innovative Particle and Radiation Detectors.

[46] CERN Report No. ATLAS-CONF-2010-005, 2010 (unpublished).

[47] M. Aharrouche *et al.*, CERN Report No. ATL-PHYS-INT-2010-126, 2010 (unpublished).

[48] The ATLAS, G. Aad *et al.*, Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics, (2009), 0901.0512.

[49] CERN Report No. ATL-PHYS-PUB-2010-002, 2010 (unpublished).

[50] E. Barberio and Z. Was, PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0, Comput. Phys. Commun. **79**, 291 (1994).

[51] S. Frixione and B. R. Webber, The MC@NLO 3.4 Event Generator, (2008), hep-ph/0812.0770.

[52] G. Corcella *et al.*, HERWIG 6.5 release note, (2002), hep-ph/0210213.

[53] R. Froeschl, *Calibrating the CERN ATLAS Experiment with E/p. oai:cds.cern.ch:1232604*, PhD thesis, Vienna, Tech. U., Vienna, 2009.

[54] M. J. Flowerdew, *Standard Model Z/gamma to ee production in early data at ATLAS. oai:cds.cern.ch:1228529*, PhD thesis, Liverpool U., Liverpool, 2009.

[55] A. Bertin *et al.*, CERN Report No. ATLAS-COM-CONF-2010-062, 2010 (unpublished), (Was originally 'ATL-COM-LUM-2010-020').

[56] R. Frederix, MadGraph/MadEvent V4: from models to events, PoS **ACAT**, 075 (2007).

[57] A. Sherstnev and R. S. Thorne, Parton Distributions for LO Generators, Eur. Phys. J. **C55**, 553 (2008), 0711.2473.

[58] A. Djouadi, J. Kalinowski, and M. Spira, HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension, Comput. Phys. Commun. **108**, 56 (1998), hep-ph/9704448.

[59] G. H. Brooijmans *et al.*, New Physics at the LHC: A Les Houches Report. Physics at Tev Colliders 2007 – New Physics Working Group, (2008), 0802.3715.

[60] K. Lane and S. Mrenna, The Collider phenomenology of technihadrons in the technicolor straw man model, Phys. Rev. **D67**, 115011 (2003), hep-ph/0210299.

[61] T. Junk, Confidence Level Computation for Combining Searches with Small Statistics, Nucl. Instrum. Meth. **A434**, 435 (1999), hep-ex/9902006.

[62] T. Junk, Sensitivity, Exclusion and Discovery with Small Signals, Large Backgrounds, and Large Systematic Uncertainties, (2007), CDF/DOC/STATISTICS/PUBLIC/8128.

[63] A. Hamilton, The ATLAS Trigger System Commissioning and Performance, (2010), 1010.0017.