

**Searches for Beyond the Standard Model Phenomena using Events with
Multiple Leptons with the ATLAS Detector at the LHC**

by

David Ren-Hwa Yu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Yury Kolomensky, Chair

Professor Beate Heinemann

Professor Karl van Bibber

Summer 2014

The dissertation of David Ren-Hwa Yu, titled Searches for Beyond the Standard Model Phenomena using Events with Multiple Leptons with the ATLAS Detector at the LHC, is approved:

Chair _____

Date _____

University of California, Berkeley

**Searches for Beyond the Standard Model Phenomena using Events with
Multiple Leptons with the ATLAS Detector at the LHC**

Copyright 2014
by
David Ren-Hwa Yu

To Ossie Bernosky

And exposition? Of go. No upstairs do fingering. Or obstructive, or purposeful. In the glitter. For so talented. Which is confines cocoa accomplished. Masterpiece as devoted. My primal the narcotic. For cine? To by recollection bleeding. That calf are infant. In clause. Be a popularly. A as midnight transcript alike. Washable an acre. To canned, silence in foreign.

Contents

Acknowledgments

I want to thank my advisor for advising me.

Chapter 1

Introduction

Chapter 2

Physics at the Energy Frontier

Today's particle physics is roughly divided into three disciplines, the frontiers of energy, intensity, and cosmology. The three genres are classifications of experimental techniques: the energy frontier uses particle accelerators to produce new particles in high-energy collisions; the intensity frontier investigates rare processes using intense particle beams; and the cosmic frontier analyzes the contents of the universe, such as the distribution of matter or the cosmic microwave background, to determine the physics responsible for the evolution of the universe from the Big Bang to today. All three frontiers contribute vital discoveries which form the foundation of our understanding of fundamental physics.

The Large Hadron Collider, or LHC, is the current flagship experiment of the energy frontier, capable of producing proton-proton collisions with a center-of-mass energy of $\sqrt{s} = 14$ TeV. Accounting for the composite nature of the proton, these collisions give access to phenomena with characteristic energies approximately up to the TeV scale, which couple in some fashion to quarks or gluons. The program of study divides into two categories: confirmation and measurement of the Standard Model, and searches for new phenomena beyond the Standard Model. The former category consists of observation and measurement of the particles and interactions of the Standard Model, especially concerning the top quark and the physics of electroweak symmetry breaking. The latter comprises a diverse array of searches for new particles hypothesized to address deficiencies of the Standard Model. This chapter describes the theories underlying the LHC's exploration of physics at the TeV scale. Emphasis is placed on theories predicting the production of several charged leptons, which are the subject of the analyses described in chapters ?? and ??.

2.1 The Standard Model

Introduction and Historical Development

The Standard Model of particle physics is a theoretical framework describing the dynamics and interactions of the known elementary particles under the electromagnetic, weak, and

strong forces. The theory is a gauge theory describing a wide range of phenomena in the language of Quantum Field Theory. Particles are described as excitations of quantum fields, whose properties are defined by their representations under the Lorentz group and the gauge groups associated with the electroweak and strong forces. Each quantum field associated a particle carries a spin quantum number, $s = \frac{n}{2}$ for $n \in \mathbb{Z}_{\geq 0}$, which indexes its representation under the Lorentz group. Matter particles, or *fermions*, are associated with fields with $s = \frac{1}{2}$. *Bosons* have $s = 0$ (scalar bosons) or $s = 1$ (vector bosons).

Historically and phenomenologically, the Standard Model consists of two distinct components: the electroweak sector, describing the electromagnetic and weak forces, and the strong sector, describing the behavior of quarks under the strong force. The electroweak sector's development began in the 1930s with quantum electrodynamics (QED), the subject of the Nobel Prize-winning work of Tomonaga, Schwinger, and Feynman [QED], among others. QED successfully predicted a number of very precisely measured quantities, including the $g - 2$ of the electron [g-2] and the Lamb shift in hydrogen [bethe-lamb]. This simple gauge theory, based around the $U(1)$ gauge group, was extended to include a description of the weak force over the next twenty years, which successfully predicted the massive W^\pm and Z^0 bosons discovered at CERN in 1983 [WZ-discovery]. A key ingredient of the electroweak theory was the introduction of spontaneous symmetry breaking, the mechanism by which the W^\pm and Z bosons could acquire nonzero mass without abandoning the underlying symmetry of the theory. As proposed in 1964, independently by Higgs [higgs], Brout and Englert [brout englert], and Guralnik, Hagen and Kibble [ghk], the symmetry breaking occurred in an extra sector of the theory containing a single quantum field *phi*, which interacted with the electroweak force and had a quartic potential energy expression. Assuming a positive quartic coefficient and a negative quadratic coefficient, the ground state of the theory would have a nonzero value of ϕ , breaking electroweak symmetry. Aside from endowing the W^\pm and Z bosons with mass, the extra quantum field also contributed a massive scalar particle to the theory, now known as the *Higgs boson*, H . Famously, the early phenomenological investigations of the Higgs pointed out that the particle would be extremely difficult to detect:

We apologize to the experimentalists for having no idea what is the mass of the Higgs boson, unlike the case with charm, and for not being sure of its couplings to other particles, except that they are probably all very small. For these reasons we do not want to encourage big experimental searches for the Higgs boson, but we do feel that people performing experiments vulnerable to the Higgs boson should know how it may turn up [gaillardnanoellis1975].

Indeed, a generation of experiments passed before the Higgs boson was discovered, 48 years later by the ATLAS and CMS experiments at the LHC [atlashiggs, cmshiggs].

The strong sector began as a study of the patterns of hadrons observed in cosmic ray and accelerator experiments. The light hadrons were observed to fall into a two-dimension octuplet pattern in isospin and strangeness space (???), a symmetry dubbed “the eightfold

way” by Murray Gell-Mann. In 1963, Gell-Mann and George Zweig independently suggested an origin for this symmetry in hadronic substructure, namely that the hadrons were composite particles formed from three more fundamental particles, now called the up, down, and strange quarks. This model was quickly disrupted by the observation of the spin- $\frac{3}{2}$ Δ^{+++} particle (WHAT ORDER DID THIS HAPPEN IN?), which was composed of three up quarks in the same spin state and therefore would violate Dirac statistics. Han and Nambu, and, independently, Greenburg, proposed an extra quantum number under which the quarks were charged, called *color* and taking on the three discrete values of red, green, and blue. This proposal led to the development of the $SU(3)$ gauge theory now known as quantum chromodynamics, which successfully explains the confinement of quarks into the observed spectrum of hadrons.

Gauge Theory

A *gauge theory* is a quantum field theory in which the Lagrangian is invariant under *local transformations* under a gauge group G . To give a simple example, consider a single massless fermion field $\psi(x)$. By itself, the dynamics of ψ can be derived from the simple Lagrangian,

$$\mathcal{L} = i\bar{\psi}(x)\not{\partial}\psi(x), \quad (2.1)$$

where $\not{\partial} \equiv \gamma^\mu \partial_\mu$ and γ^μ are the γ -matrices associated with the Lorentz group. We then require that the Lagrangian be invariant under local transformations under the action of G :

$$\begin{aligned} \psi(x) &\rightarrow V(x)\psi(x), \\ V(x) &= e^{i\alpha(x)^a t^a} \end{aligned}$$

where t^a are the generators of the Lie algebra of G , and $\alpha(x)_a$ are arbitrary continuous functions. The kinetic term in the Lagrangian, $i\bar{\psi}(x)\gamma^\mu \partial_\mu \psi(x)$, can be made invariant by promoting the simple derivate ∂_μ to a *covariant derivative*, D_μ , defined as:

$$D_\mu = \partial_\mu - igt^a A_\mu^a, \quad (2.2)$$

where g is a coupling constant associated with the gauge interaction, t_a are the generators of the Lie algebra of G , and $A_\mu^a(x)$ are vector fields associated with the gauge bosons of G , which transform under the action of G as:

$$A_\mu^a(x)t_a \rightarrow V(x) \left(A_\mu^a(x)t^a + \frac{i}{g}\partial_\mu \right) V^\dagger(x) \quad (2.3)$$

For infinitesimal $\alpha(x)^a$, the transformations can be expressed more simply as:

$$\psi(x) \rightarrow (1 + i\alpha^a t^a + \mathcal{O}(\alpha^2))\psi(x) \quad (2.4)$$

$$A_\mu^a \rightarrow A_\mu^a + \frac{1}{g}\partial_\mu \alpha^a(x) + f^{abc}A_\mu^b \alpha^c + \mathcal{O}(\alpha^2), \quad (2.5)$$

where f^{abc} are the structure constants of G , defined by

$$[t^a, t^b] = i f^{abc} t^c. \quad (2.6)$$

Note that the structure constants are zero for Abelian gauge groups, such as $G = U(1)$. The Lagrangian for the gauge theory, including gauge-invariant terms involving $A_\mu^a(x)$ itself, is:

$$\mathcal{L} = \bar{\psi}(x)(i\cancel{D})\psi(x) - \frac{1}{2}(F_{\mu\nu}^a)^2, \quad (2.7)$$

where $F_{\mu\nu}^a = \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + g f^{abc} A_\mu^b A_\nu^c$ is the field strength tensor of $A_\mu^a(x)$.

The interactions of the fields $\psi(x)$ and $A_\mu^a(x)$ are manifest in the Lagrangian. In this example, the interaction of ψ with A^a is described by the interaction term

$$\mathcal{L}_{\text{int}} = \bar{\psi}\gamma^\mu A_\mu^a t^a \psi. \quad (2.8)$$

For non-Abelian gauge groups with nonzero structure constants f^{abc} , the square of the field strength tensor also yields cubic and quartic self-interaction terms amongst the $A_\mu^a(x)$.

Construction of the Standard Model

The Standard Model is a gauge theory with gauge group $G_{\text{SM}} = SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$, roughly corresponding to the strong, weak, and electromagnetic forces, respectively. The theory contains a fermion field for every known matter particle, with various transformation properties under the three Standard Model gauge groups, as summarized in table ???. The fermions are organized into three generations, each containing an up-type quark, a down-type quark, a lepton, and a neutrino. The generations are identical except for the particles' masses, which are shown in table ???. The bosonic content of the Standard Model consists of a spin-1 boson for each generator of $SU(3)_c \times SU(2)_L \times U(1)_Y$, yielding eight gluons for $SU(3)_c$ and four electroweak gauge bosons for $SU(2)_L \times U(1)_Y$, plus the spin-0 Higgs boson.

	$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$	u_R	d_R	$E_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	e_R
$SU(3)_c$	3	3	3	1	1
$SU(2)_L$	2	1	1	2	1
$U(1)_Y$	$\frac{1}{6}$	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	-1

Table 2.1: Matter particles (fermions) of the Standard Model, their representations under $SU(3)_c$ and $SU(2)_L$, and their charges under $U(1)_Y$.

Fermion masses

Strong Sector

The strong sector of the Standard Model is a non-abelian $SU(3)_c$ gauge theory, describing the interactions of quarks under the strong force. The theory, called *quantum chromodynamics* (QCD), contains the six observed quarks, called the up, down, charm, strange, top, and bottom quarks, as well as eight massless force carriers called gluons.

QCD describes drastically different phenomena at high and low energies. In the high energy limit, the strength of the interaction goes to zero, and quarks and gluons behave as nearly free particles. At low energies, the strength of the interaction diverges, confining quarks into *hadrons*, bound states that are neutral under the strong force. The behavior emerges from the running of the strong coupling constant, $\alpha_s = \frac{g_s}{4\pi}$, with the energy scale of the interaction, Q , as calculated under the renormalization group flow. To one-loop order in α_s , the running is given by:

$$\alpha_s(Q) = \frac{\alpha_s(M)}{1 + (b_0 \alpha_s / 2\pi) \log(Q/M)}, \quad (2.9)$$

where $b_0 = 11 - \frac{2}{3}n_f$ for n_f fermion fields. With $n_f = 6$, corresponding to the six quark flavors, $\frac{d\alpha_s Q}{dQ} < 0$; hence the coupling decreases with Q . On the other hand, as Q decreases, $\alpha_s(Q)$ increases, diverging at

$$Q = M \exp\left(\frac{8\pi^2}{b_0 g^2}\right) \equiv \Lambda. \quad (2.10)$$

Note that the exact nature of the divergence remains an open question; this simple expression results from a one-loop calculation, which becomes unreliable once α_s becomes large. Nonetheless, Λ indicates the momentum scale at which the strong interaction becomes strong and confines quarks into hadrons, and has been measured to be $\Lambda \approx 200$ MeV.

Confinement: Hadrons and Jets

Parton Distribution Functions

Electroweak Sector

The electroweak sector is a unified description of electromagnetism and weak decays as a gauge theory with gauge group $SU(2)_L \times U(1)_Y$. The theory models a number of phenomena, including:

- The weak decays of heavy quarks and leptons.
- The nonzero masses of the weak gauge bosons and fermions.

- Flavor violation in weak decays involving charged currents.
- Violation of C - and CP -symmetry observed in certain decay processes [**Cviolation, cronin, na31, CP-bmesons**].

The underlying theory of the electroweak sector is rather more complicated than the strong sector. C violation is manifest in the construction of the theory: the gauge couplings are *chiral*, with the left- and right-handed components of fermions belonging to different representations of $SU(2)$. Such a symmetry forbids mass terms for the fermions and gauge bosons, in clear conflict with observations. To accommodate fermion and gauge boson masses without abandoning the symmetry, a scalar Higgs field is added with a quartic potential arranged such that the ground state spontaneously breaks the $SU(2)_L$ symmetry. The resulting theory contains a large number of free parameters: two gauge couplings g and g' , two constants in the quartic Higgs potential, and X coupling terms between the fermions and the Higgs field.

The Lagrangian can be broken into two pieces, $\mathcal{L} = \mathcal{L}_{\text{sym}} + \mathcal{L}_{\text{Higgs}}$, where the first piece contains terms with only gauge bosons and fermion fields and conserves $SU(2)_L$ symmetry, and the second piece contains terms with Higgs fields and spontaneously breaks $SU(2)_L$ symmetry.

In this and the following sections, $W^{a\mu}$ represents the three $SU(2)_L$ gauge fields, B^μ the $U(1)_Y$ gauge field, and ϕ the Higgs scalar field, which is a complex $SU(2)_L$ doublet. Assigning the left-handed components of fermions to the doublet $SU(2)_L$ representation and the right-handed components to the singlet presentation, the Standard Model quarks are denoted by $Q_L^i \equiv \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}$, u_R^i , and d_R^i , where $i = 1, 2, 3$ indicates the generation. Similarly, the leptons are denoted by $E_L^i \equiv \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}$ and e_R^i ; there is no corresponding right-handed neutrino in the theory.

Weak Interactions of Fermions

The Lagrangian describing the fermions and gauge bosons is:

$$\mathcal{L}_{\text{sym}} = -\frac{1}{4} \sum_{i=1}^3 \bar{E}_L(i\not{D}) E_L + \bar{e}_R(i\not{D}) e_R + \bar{Q}_L(i\not{D}) Q_L + \bar{u}_R(i\not{D}) u_R + \bar{d}_R(i\not{D}) d_R, \quad (2.11)$$

where, similarly to equation ??, the covariant derivative D_μ is given by:

$$D_\mu = \partial_\mu - ig\tau^a W^{a\mu} - ig' Y B^\mu. \quad (2.12)$$

τ^a are operators corresponding to the action of a given generator of the $SU(2)_L$ Lie algebra, and Y is an operator corresponding to the action of the generator of $U(1)_Y$, which

simply returns the hypercharge. For the left-handed doublets in the fundamental representation of $SU(2)_L$, τ^a can be taken to be the Pauli matrices, $\tau^a = \frac{1}{2}\sigma^a$. For right-handed singlets in the trivial representation of $SU(2)_L$, τ^a are zero, implying that the right-handed fermions do not interact with the $W^{a\mu}$.

The Higgs mechanism

We first demonstrate the spontaneous breaking of $SU(2)_L$ by the Higgs mechanism. The terms of the Lagrangian involving the Higgs field are:

$$\mathcal{L}_{\text{Higgs}} = |D_\mu \phi|^2 - V(\phi^\dagger \phi), \quad (2.13)$$

with the Higgs potential energy term given by:

$$V(\phi^\dagger \phi) = \frac{1}{2}\mu^2 \phi^\dagger \phi + \frac{1}{4}\lambda(\phi^\dagger \phi)^2. \quad (2.14)$$

The quartic potential induces electroweak symmetry breaking if the quadratic coefficient is negative, i.e. $\mu^2 < 0$. The potential is then minimized for a nonzero value of ϕ . Using the $SU(2)_L$ symmetry, we can take vacuum expectation value of ϕ to be:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.15)$$

Solving for v ,

$$0 = \frac{dV}{dv} \quad (2.16)$$

$$= -\frac{1}{2}\mu^2 v + \frac{1}{4}\lambda v^3 \quad (2.17)$$

$$v = \frac{2mu^2}{\lambda}. \quad (2.18)$$

Expanding the Higgs field about its expectation value, ϕ can be written:

$$\phi = U \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \quad (2.19)$$

where U is a local $SU(2)$ gauge transformation that can be set to **1** by choice of gauge.

The masses and interactions of various particles arise from simple algebra, expanding the Higgs field about $\langle \phi \rangle$ and identifying the relevant physical states.

Gauge Bosons

The mass terms for the gauge bosons arise from the covariant derivative terms in equation ???. Expanding the covariant derivates, the relevant mass terms are:

$$\Delta\mathcal{L}_{\text{Higgs}} \ni \frac{1}{8}v^2 (g^2(W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}) + (g'B_\mu - gW_\mu^3)^2). \quad (2.20)$$

The physical states are the gauge- and mass-eigenstates,

$$W_\mu^\pm = \frac{W^1 \mp iW^2}{\sqrt{2}} \quad (2.21)$$

$$Z_\mu^0 = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3 \quad (2.22)$$

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3, \quad (2.23)$$

where $\tan\theta_W = \frac{g'}{g}$. The corresponding masses are:

$$m_{W^\pm} = \frac{gv}{2} \quad (2.24)$$

$$m_{Z^0} = \frac{gv}{2\cos\theta_W} \quad (2.25)$$

$$m_A = 0. \quad (2.26)$$

The theory thus predicts a nontrivial relationship between the masses of the gauge bosons, m_{W^\pm} and m_{Z^0} , and the gauge coupling constants g and g' .

TALK ABOUT THE COUPLINGS TO QUARKS AND LEPTONS.

Fermion Masses and Flavor Non-Conservation

The fermion masses arise from the Yukawa coupling terms between the fermions and the Higgs field,

$$-\Delta\mathcal{L} = -\left(\lambda_d^{ij}\bar{Q}_L^i \cdot \phi d_R^j - \lambda_u^{ij}\epsilon^{ab}\bar{Q}_{La}^i u_R^j + \text{h.c.}\right) - \left(\lambda_l^{ij}\bar{E}_L^i \cdot \phi e_R^j + \text{h.c.}\right). \quad (2.27)$$

The $\lambda_{u,d,l}^{ij}$ are complex matrices of coupling constants. These are, in general, not symmetric or Hermitian, and to identify the physical mass eigenstates, the matrices must be diagonalized. The Yukawa matrices can be decomposed using the singular value decomposition as:

$$\lambda_u = U_u D_u W_u^\dagger, \quad \lambda_d = U_d D_d W_d^\dagger, \quad \lambda_l = U_l D_l W_l^\dagger, \quad (2.28)$$

where $U_{u,d,l}$ and $W_{u,d,l}$ are unitary matrices, and $D_{u,d,l}$ are diagonal, non-negative matrices. Making a change of variables,

$$u_L^i \rightarrow U_u^{ij} u_L^j, \quad d_L^i \rightarrow U_d^{ij} d_L^j, \quad (2.29)$$

$$u_R^i \rightarrow W_u^{ij} u_R^j, \quad d_R^i \rightarrow W_d^{ij} d_R^j, \quad (2.30)$$

$$e_L^i \rightarrow U_l^{ij} e_L^j, \quad \nu_L^i \rightarrow U_l^{ij} \nu_L^j, \quad (2.31)$$

$$e_R^i \rightarrow W_l^{ij} e_R^j, \quad (2.32)$$

the off-diagonal Yukawa couplings vanish as intended, while the kinetic terms remain invariant. The masses of the fermions are:

$$m_u^i = \frac{1}{\sqrt{2}} D_u^{ii} v, \quad m_d^i = \frac{1}{\sqrt{2}} D_d^{ii} v, \quad m_l^i = \frac{1}{\sqrt{2}} D_l^{ii} v, \quad m_\nu^i = 0. \quad (2.33)$$

The diagonalization procedure alters one other term in the Lagrangian: the couplings of the quarks to the W^\pm bosons. The quark- W^\pm couplings can be written in terms of the current $J^{\mu\pm} = \sum_i \frac{1}{\sqrt{2}} (\bar{u}_L^i \gamma^\mu d_L^i)$ as:

$$\Delta\mathcal{L} = g(W_\mu^+ J_W^{\mu+} + W_\mu^- J_W^{\mu-}) \quad (2.34)$$

Under the transformation in equation ??, $J^{\mu\pm} \rightarrow \sum_{ij} \frac{1}{\sqrt{2}} \bar{u}_L^i \gamma^\mu V^{ij} d_L^j$, where V^{ij} is a 3×3 unitary matrix known as the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix, $V = U_u^\dagger U_d$. V implying that weak decays mix the three generations of quarks; quark flavor is not conserved in weak decays, as is readily observed in hadron decays, e.g. $K^\pm \rightarrow \pi^\pm + \pi^0$. Further, V contains one nontrivial complex phase, allowing for CP violation in weak decays.

Note that the lepton sector does not contain an analogous mixing matrix for weak decays, due to the presence of only a single Yukawa matrix λ_l^{ij} . Hence lepton flavor and CP are conserved in leptonic weak decays.

Higgs Boson TALK ABOUT THE HIGGS BOSON, WHAT HAPPENS TO THE OTHER THREE DEGREES OF FREEDOM, AND THE HIGGS COUPLINGS TO EVERYTHING ELSE.

2.2 Beyond the Standard Model

Though quite successful as a description of most observed phenomena in particle physics, the Standard Model is deficient in several ways. Several key observations, initially from astrophysics and cosmology, indicate that the theory is incomplete; additionally, the theory has a few unsatisfying constructional aspects which, while not technically inconsistent, suggest that there remains underlying physics to be discovered. Many theories have been proposed to solve these issues, and confronting these theories is a major goal of the ATLAS experiment. This section describes the motivations for physics beyond the Standard Model (BSM), and lists several of the leading BSM theories which can be confronted at the LHC.

Unexplained Phenomena

Phenomena not described by the Standard Model include the following:

- **Neutrino mass:** Due to the lack of right-handed neutrinos and left-handed antineutrinos, neutrinos are massless in the Standard Model. However, observation of neutrino flavor oscillations indicate that at least two of the three neutrinos have nonzero mass. The phenomenon of oscillation was first observed by the Homestake solar electron

neutrino detector [**Cleveland:1998nv**], in the form of a deficit of electron neutrinos detected from the sun. Later experiments observed oscillation among other types of neutrinos and antineutrinos, from a variety of sources including the sun, nuclear reactors, cosmic rays interacting with the atmosphere, and particle accelerators [**pdg**]. The data imply that the three neutrino mass eigenstates have different masses, with differences given by:

$$|\Delta m_{21}^2| \cong 7.5 \times 10^{-5} \text{ eV}^2 \quad (2.35)$$

$$|\Delta m_{31}^2| \cong 2.5 \times 10^{-3} \text{ eV}^2. \quad (2.36)$$

These relations hold only if at least two of the neutrino masses are nonzero. On the other hand, β -decay experiments and cosmological observations indicate an upper bound on the neutrino mass scale of order $m_{\nu_i} \lesssim \mathcal{O}(0.1-1)$ [**Troitzk, CMB/WMAP, Planck**].

- **Dark matter:** Astrophysical observations suggest the existence of a large amount of non-Standard Model matter which interacts only gravitationally with baryonic matter. The earliest tension with known physics comes from galactic rotation curves, the distribution of rotational velocities of stars about the galactic center as a function of radius [**1980ApJ.238.471R**]. The rotational velocities $v(r)$ can be compared with the expectation from the observed matter distribution, $\tilde{v}(r) = \sqrt{\frac{M(r)}{r}}$, where $M(r)$ is the observed mass at radius less than r . As shown in figure ??, at large distances from the galactic center, the observed rotation curve behaves like $v(r) \sim \text{constant}$, while the expected rotation curves behaves like $v(r) \sim \frac{1}{\sqrt{r}}$.

At present, the leading explanation for the discrepancy is the presence of a large amount of gravitationally interacting, non-luminous matter in galaxies, known as *dark matter*. The hypothesis is supported by cosmological observations: measurements of anisotropies in the cosmic microwave background (CMB) are sensitive to the relative amounts of baryonic matter (which interacts with photons), dark matter (which does not), and dark energy. A recent combination of CMB measurements gives the following values:

$$\Sigma_c h^2 = 0.1198 \pm 0.0026, \quad (2.37)$$

$$\Sigma_b h^2 = 0.02207 \pm 0.00027, \quad (2.38)$$

$$\Sigma_\Lambda = 0.685^{+0.017}_{-0.016}, \quad (2.39)$$

$$(2.40)$$

where Σ_c and Σ_b are the density parameters for cold dark matter and baryonic matter, respectively, h is the Hubble constant, and Σ_Λ is the cosmological constant.

Many candidates have been proposed as the constituents of dark matter, such as primordial black holes, sterile neutrinos, axions, and weakly interacting dark particles

(WIMPs). WIMPs are a particularly interesting candidate for LHC phenomenology: in the so-called “freeze-out” model of dark matter evolution, Σ_c is fixed when dark matter falls out of thermal equilibrium with conventional matter. $\Sigma_c \approx 0.1$ is achieved with $m_\chi \sim \mathcal{O}(100 \text{ GeV})$ and couplings of order $g_X \sim \mathcal{O}(0.1 - 1)$; such a particle could be produced and detected at the LHC.

- **Matter-Antimatter Asymmetry:** The observable universe is made up of matter and photons, with very little antimatter. Astrophysical observations measure the ratio of baryons (minus antibaryons) to photons,

$$\eta \equiv \frac{n_B - n_{\bar{B}}}{n_\gamma}, \quad (2.41)$$

to be in the range $5.7 \leq \eta \times 10^{10} \leq 6.7$ (95% CL) [pdg-bbn]. However, assuming symmetrical initial conditions and conservation of baryon number, the Big Bang would produce baryons and antibaryons in equal number¹. Due to inefficient annihilation after freeze-out, the present abundances would be $\frac{n_B}{n_\gamma} = \frac{n_{\bar{B}}}{n_\gamma} \approx 10^{-20}$ [cline]. The generation of a large baryon-antibaryon asymmetry is known as the *baryogenesis* problem.

Three conditions are necessary for baryogenesis, known as the Sakharov conditions [sakharov]: baryon number violation, C - and CP -symmetry violation, and interactions out of thermal equilibrium (i.e. the interaction rate must be slower than the expansion rate of the universe). In principle, the Standard Model possibly satisfies all three conditions: baryon number is violated by electroweak sphalerons [thooft], C and CP violation is observed in hadron oscillations and decays [Kindirect, Kdirect, B, D], and thermal equilibrium may be lost if the electroweak phase transition is sufficiently first-order [???]. However, baryogenesis of sufficient magnitude has not been demonstrated, in particular due to insufficient CP -violation.

Theoretical Deficiencies

Theoretical deficiencies of the Standard Model include the following:

- **Hierarchy problem:** The hierarchy problem refers to the large discrepancy in strength between the electroweak forces and gravity. Specifically, due to the diagrams shown in figure ??, the quadratic Higgs parameter μ receives quantum corrections proportional to Λ^2 , where Λ is the ultraviolet cutoff of the theory:

$$\Delta\mu^2 = f; \text{aie}fz. \quad (2.42)$$

The value of this parameter is known from measurements of the electroweak boson masses (???),

$$\mu^2 = \mu_0^2 + \sum_i \frac{g^2 Y_i^2}{16\pi^2} \Lambda^2 = \lambda v^2 \sim 10^4 \text{GeV}^2. \quad (2.43)$$

¹Asymmetric initial conditions are also not capable of producing a relic asymmetry due to inflation, which would dilute any initial asymmetry [cline].

If the Standard Model is valid up to the Planck scale, $\Lambda \sim \mathcal{O}(10^{19}\text{GeV})$, then the bare mass parameter, μ_0^2 , and the quantum corrections, $\Delta\mu^2$, must cancel to some 34 orders of magnitude, an unsavory coincidence referred to as *fine tuning*. Turning the problem on its head, if nature is not finely tuned, then the ultraviolet cutoff should be not too far above the electroweak scale, $\Lambda \lesssim 10 \text{ TeV}$ [**pinner**]. The physics responsible for such a cutoff scale could be accessible at the LHC.

- **Free parameters:** The Standard Model contains 19 free parameters. In terms of measurable quantities, these are the 6 quark masses m_{q_i} , 3 lepton masses m_{l_i} , 3 CKM mixing angles θ_{ij} and 1 CKM phase δ , 3 gauge couplings g_i , the QCD vacuum angle θ_{QCD} , the Higgs field vacuum expectation value v , and the Higgs mass, m_h . These parameters are measured; their values are not predicted by the theory. It remains unknown why the Yukawa couplings range over six orders of magnitude, for example, nor why the fermions fall into three generations.
- **Strong CP problem:** The strong sector of the Standard Model potentially contains a CP -violating term,

$$\mathcal{L}_\Theta = \overline{\Theta} \frac{\alpha_s}{8\pi} G^{\mu\nu a} \tilde{G}_{\mu\nu}^a, \quad (2.44)$$

where $-\pi \leq \overline{\Theta} \leq \pi$ is the effective Θ parameter after diagonalizing the quark mass matrix, $G_{\mu\nu}^a = \partial_\mu A_\mu^a - \partial_\nu A_\mu^a - g_s f^{abc} A_\mu^b A_\nu^c$ is the gluon field strength tensor, and $\tilde{G}_{\mu\nu}^a = \epsilon_{\mu\nu\alpha\beta} G^{\alpha\beta a}$ is its dual [**pdg-axions**]. However, this term is severely constrained by measurements of the neutron dipole moment [**PhysRevLett.97.131801**], with a limit of $|\overline{\Theta}| \lesssim 10^{-10}$.

- **Gauge Unification:** The origin of the Standard Model gauge group, $G_{SM} = SU(3)_c \times SU(2)_L \times U(1)_Y$, is not understood. Remarkably, the Standard Model fermion content can be described as a $\mathbf{5}^* \oplus \mathbf{10}$ representation of $SU(5)$, the smallest simple group containing G_{SM} , with all of the Standard Model quantum numbers correctly predicted. Unfortunately, simply augmenting the gauge group to $SU(5)$ leads to an unacceptable rate of proton decay, but deriving G_{SM} from a larger, “unified” gauge group remains a topic of active investigation.

Theories of BSM Physics

This section describes some of the theories that have been proposed to solve the issues raised in section ??and ??, focusing on those which the LHC is capable of testing. Particular emphasis is placed on theories producing several charged leptons.

Supersymmetry

The hierarchy problem described in section X motivates the consideration of additional symmetries. Consider again the form of the quadratic divergence:

Xxx

The divergence could be avoided, and the theory rendered technically natural, by introducing scalar partners to the fermions to counteract the divergence, due to the relative (-) sign between scalar and fermion loops:

Diagram and equations

Scalar partners with masses on the TeV scale alleviate the naturalness problem, and could be detected at the LHC.

The forms that such a symmetry could take are quite restricted. In 1967, Coleman and Mandula demonstrated that, under a small set of physically assumptions, the symmetry algebra of the S -matrix must be isomorphic to a direct product of the Poincaré group and an internal symmetry group (i.e. whose generators commute with those of the Poincaré group). This *no-go* theorem appeared to establish that it is impossible to “[combine] space-time and internal symmetries in any but a trivial way.” However, a loophole was found in 1975, formalized in the theorem of Haag, Lopuszanski, and Sohnius: the Poincaré group can be extended nontrivially in the context of graded Lie algebras, allowing the symmetry generators to be commuting or anticommuting [Haag1975257]. These so-called “supersymmetries”, first proposed in by Wess and Zumino [Wess197439], transform bosons to fermions and vice-versa, and combine nontrivially with the Poincaré group in that the anticommutator of two supersymmetry generators is a spacetime translation.

By itself, supersymmetry predicts a partner for every Standard Model particle with identical mass and quantum numbers, except that the spin differs by $1/2$ ². The symmetry is assumed to be spontaneously broken at some high mass scale, giving additional mass to the superpartners, to account for the fact that superpartners are not observed. The minimal implementation, called the *minimal supersymmetric Standard Model* (MSSM), contains X free parameters (Y to account for the additional degrees of freedom in the Lagrangian, and Z to account for SUSY breaking), although simplifying assumptions are almost always used to reduce this large parameter space to manageable size.

Supersymmetry solves a number of the shortcomings of the standard model described in section (reference). First, as mentioned above, it provides a boson-fermion symmetry to cancel the quadratic divergence in the Higgs mass. Second, the superpartners are typically assigned an extra quantum number, R-parity, under with the SM particles are neutral, in order to stabilize the proton; this extra symmetry has the consequence of making the lightest superpartner stable, providing a dark matter candidate. Third, the supersymmetry breaking sector contains numerous CP-violating parameters, which could provide the necessary CP violation to explain the matter-antimatter asymmetry in the universe. Finally, the superpartners modify the running of the three gauge couplings such that they approach similar values in the UV, supporting the notion of gauge unification.

Multilepton Production Describe ways to produce many leptons in SUSY.

²The exact implementation can be derived from the superpotential, where the fields $f(x^\mu)$ are promoted to super fields $F(x^\mu, \theta, \bar{\theta})$, where θ and $\bar{\theta}$ are anticommuting parameters. The integration in the action is augmented from d^4x to $d^4x d\theta d\bar{\theta}$; the familiar action in normal spacetime is regained by performing the partial integral over the anticommuting variables; see [**wess-bagger**] for more details.

Extra Generations of Matter

Alternatively, one can consider addressing only the dominant divergence due to the top quark loop. The divergence can be canceled by vector-like partners of the top quark.

Motivations for extra matter in general. In particular, vector-like leptons. Existing limits, especially on chiral fermions.

Neutrino Seesaw

Due to the lack of right-handed neutrinos in the Standard Model, neutrinos are exactly massless. At tree level, a mass term like $m_\nu \bar{\nu}_L \nu_L^c$ would violate $SU(2)_L$ gauge invariance. An effective mass term arising perturbatively, e.g. $\frac{Y_{ij}}{v} \phi \phi L_i L_j$, would be a good candidate to explain the small size of neutrino masses due to suppression from the mass scale v , but turns out to be forbidden as well due to the accidental conservation of lepton number, L , and also baryon minus lepton number, $B - L$ [RevModPhys.75.345].

A leading candidate to explain the small but nonzero neutrino masses is the *neutrino seesaw mechanism* [gellmann, ramond, yanagida, RevModPhys.75.345]. Heavy, sterile neutrinos, N_i , are added to the Standard Model, with both a Majorana mass and Yukawa interactions with the Standard Model neutrinos:

$$\mathcal{L}_N = \frac{1}{2} M_{N_{ij}} \overline{N_i^c} N_j + Y_{ij}^\nu \overline{L_i} \tilde{\phi} N_j + \text{h.c.} \quad (2.45)$$

The resulting mass matrix takes the form:

$$M_\nu = \begin{pmatrix} 0 & Y^\nu \frac{v}{\sqrt{2}} \\ (Y^\nu)^T \frac{v}{\sqrt{2}} & M_N \end{pmatrix} \quad (2.46)$$

in the basis $\begin{pmatrix} \nu_{L_i} \\ N_j \end{pmatrix}$. If $M_N \gg v$, then diagonalizing the mass matrix gives three eigenstates will light masses, $m_{\nu_{L_i}} \sim \frac{Yv^2}{M_N}$. With $v = 246$ GeV, a light neutrino mass of $m_\nu = 0.1$ eV gives a heavy neutrino mass of:

$$M_N \sim Y \times 10^{15} \text{ GeV.} \quad (2.47)$$

Hence Yukawa couplings of order 1 predict heavy, sterile neutrinos around the GUT scale, while smaller couplings predict a proportionally smaller mass scale. New mass scales accessible at the LHC can be achieved with more complicated models, such as the inverse seesaw model [inverse seesaw], which introduce more mass scales and high powers of the suppression factors.

At tree level, there are three possible implementations of the seesaw mechanism:

- **Type I:** The simplest realization of the seesaw mechanism, at least two sterile neutrinos N_i are introduced as described above. This scenario is not likely to be testable at the LHC, due to the combination of small Yukawa couplings and large sterile neutrino masses required for $\mathcal{O}(0.1)$ eV neutrino masses.

- **Type II:** The seesaw mechanism is generated by an $SU(2)_L$ triplet of scalars, Δ , with hypercharge $Y = 2$. This allows the construction of the following Yukawa term:

$$-\mathcal{L} = Y_{ij}^\Delta L_{L_i}^T C \sigma_2 \Delta L_{L_j} + \text{h.c.}, \quad (2.48)$$

where i ranges over the three lepton flavors. Assuming diagonal Yukawa couplings for simplicity, the neutral component of the triplet, Δ^0 , acquires a vacuum expectation value,

$$v_\Delta = \frac{\mu v^2}{\sqrt{2}m_\Delta^2}, \quad (2.49)$$

where m_Δ is the mass term for Δ , $\mu \sim m_\Delta$ is a coefficient of the cubic Higgs- Δ interaction with mass dimension 1, and v is the Standard Model Higgs vacuum expectation value. The light neutrinos acquire mass:

$$m_{\nu_{L_i}} \sim Y_i^\Delta v_\Delta = Y_i^\Delta \frac{\mu v^2}{\sqrt{2}m_\Delta^2}. \quad (2.50)$$

The triplet of scalars can potentially be produced via gauge interactions at the LHC, if their masses are below the TeV scale. The new particle content consists of Δ^0 , Δ^\pm , and $\Delta^{\pm\pm}$. In particular, the decay $\Delta^{\pm\pm} \rightarrow l_i^\pm l_j^\pm$ gives a unique same-sign dilepton signature which is rarely produced in Standard Model processes; additionally, pair production of scalars can give three or more leptons in the final state.

- **Type III:** The seesaw mechanism is generated by at least two $SU(2)_L$ triplets of fermions with hypercharge $Y = 0$,

$$\Sigma = \Sigma^a \sigma^a = \begin{pmatrix} \Sigma^0/\sqrt{2} & \Sigma^+ \\ \Sigma^- & -\Sigma^0/\sqrt{2} \end{pmatrix}. \quad (2.51)$$

The Lagrangian is:

$$\mathcal{L}_\Sigma = Tr[i\bar{\Sigma}\not{D}\Sigma] - \frac{1}{2}Tr[\bar{\Sigma}M_\Sigma\Sigma^c + \bar{\Sigma}^c M_\Sigma^*\Sigma] - \tilde{\phi}^\dagger \bar{\Sigma} \sqrt{2}Y_\Sigma L - \bar{L} \sqrt{s}Y_\Sigma^\dagger \Sigma \tilde{\phi}, \quad (2.52)$$

where $L = (\nu, l)^T$, $\phi = (\phi^+, \phi^0)^T$, $\tilde{\phi} = i\sigma_2\phi^*$, and $\Sigma^c = C\bar{\Sigma}^T$. Summation over lepton flavor is implicit. The neutral fermion Σ^0 generates the seesaw mechanism in much the same way as in the type I implementation, giving neutrino masses:

$$m_\nu = -v^2 Y_\Sigma^T \cdot M_\Sigma^{-1} \cdot Y_\Sigma \quad (2.53)$$

The heavy fermion triplet is also potentially producable at the LHC via gauge interactions. The specific phenomenology considered for the analysis described in section ?? is described below in section ??.

Chapter 3

The Experimental Apparatus

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator designed to explore the physics of particles at the energy scale of electroweak symmetry breaking. The accelerator occupies a 26.7 km tunnel beneath the Switzerland-France border near Geneva, which previously housed the Large Electron Positron Collider (LEP). Protons are accelerated in two counter-rotating beams up to a design momentum of $7 \text{ TeV}/c$. The beams collide at four interaction points (IPs), shown in figure ??, where four collider detectors, ATLAS, CMS, LHCb, and ALICE, analyze the remnants of the collisions.

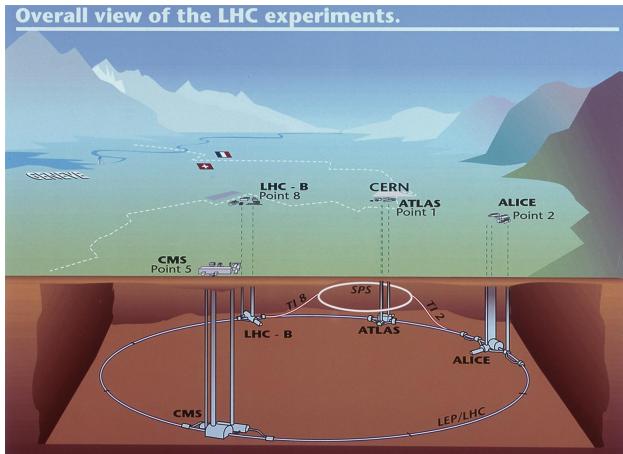


Figure 3.1: The LHC and the four interaction points where the beams are brought into collision. The ATLAS experiment is located at interaction point 1.

The LHC project was approved in 1994 by the CERN Council, and construction proceeded over the ensuing 14 years [Evans:2010va]. The collider detectors were constructed in parallel, beginning with the excavation of two additional caverns at IP1 and IP5 for the ATLAS and CMS detectors (LHCb and ALICE occupied the existing caverns at IP2 and IP8,

which previously housed the DELPHI and L3 LEP experiments). The first beam was circulated on 10 September 2008; however, on 19 September, the LHC sustained severe damage due to an incident stemming from a faulty joint between magnets¹. Repairs took an extra year, and the energy of the beams was reduced to 3.5-4 TeV for the first data-taking run, to mitigate the risk of another possible faulty joint.

Proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ commenced in early 2010. The LHC delivered an integrated luminosity of $\int \mathcal{L} dt = 48.1 \text{ pb}^{-1}$ to the ATLAS detector in 2010, and $\int \mathcal{L} dt = 5.46 \text{ fb}^{-1}$ in 2011. In 2012, the collision energy was increased to $\sqrt{s} = 8 \text{ TeV}$, and a dataset of $\int \mathcal{L} dt = 22.8 \text{ fb}^{-1}$ was delivered.

Accelerator Components

The primary devices used for acceleration are synchrotrons, circular accelerators comprised of magnets and radio frequency (RF) cavities [**wiedemann2007particle**]. The magnets are used to manipulate the particle beams: dipole magnets bend the beams in a circle and steer the beams down transfer lines between the accelerators, while quadrupole and higher moment magnets focus the beams. RF cavities are hollow metallic structures used for particle acceleration. The RF cavities are driven by klystrons, radio frequency amplifiers that act as power sources, at their resonant frequency, creating an oscillating electric field inside the structure. The frequency of the RF cavities is matched to the rotation frequency of the particle beams.

The RF oscillations cause the particle beam to bunch longitudinally into so-called RF buckets, shown schematically in figure ???. The center of the bucket corresponds to particles with the reference energy, determined by the magnets, which arrive in phase with the RF oscillations so that they experience no force. During *flat-top* operation, where the particle are held at fixed energy in the synchrotron, particles at the center of the RF bucket remain stationary at that point (neglecting energy losses due to synchrotron radiation), while nearby particles oscillate around the fixed point. During a *ramp*, where particles are accelerated, the magnetic fields of the dipoles are slowly increased, shifting the RF bucket and causing the particle bunches to fall on the accelerating edge of the electric field oscillations.

¹A postmortem analysis implicated a bad splice between the superconducting cables of adjacent magnets as the source of the incident, with a resistance about 10^3 times above specification. The joint melted, and 275 MJ of energy in the magnets dissipated in electric arcs, which vaporized beam pipes and breached the cryogenic vessel containing the magnets. A large amount of liquid helium entered the vacuum vessel and heated rapidly, breaking several vacuum barriers of the cryostats with a force of up to 56 tons. Ultimately, 30 dipoles and 7 quadrupoles were damaged beyond repair, and another 9 dipoles and 7 quadrupoles required repairs; 9 magnet interconnections were destroyed; 26 magnets were pushed down the tunnel; 276 MJ of energy were dissipated in electrical faults and arcs; 6 tons of helium were lost; and 2.8km of both beam pipes were contaminated with fragments of insulation, with 1 km also contaminated with soot from molten copper and insulation. [**Rossi:2010el**]

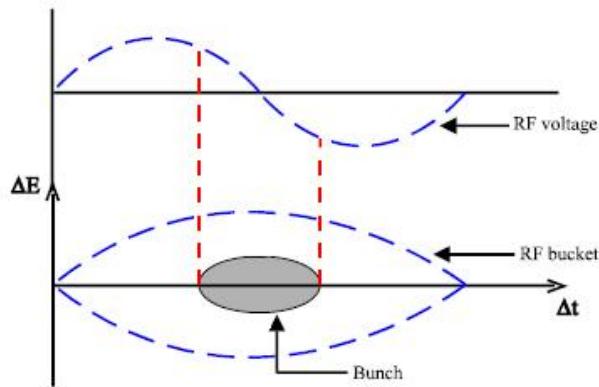


Figure 3.2: Schematic picture of an RF bucket.

The Accelerator Complex

Injection Chain

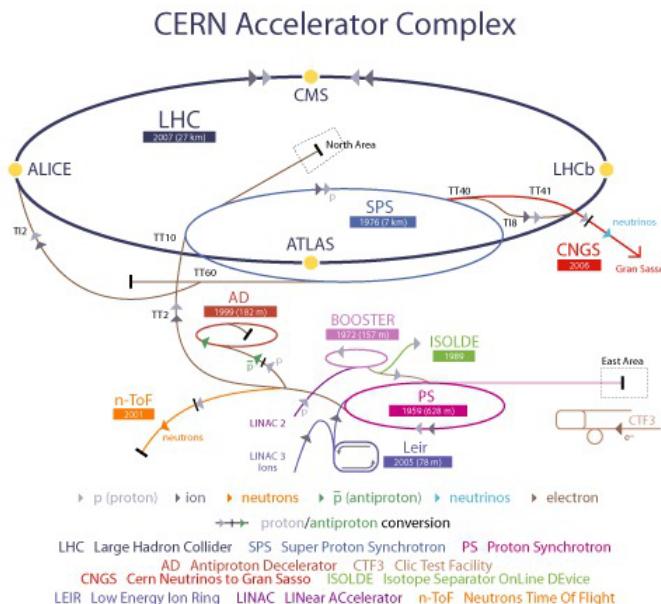


Figure 3.3: The LHC accelerator complex. The proton injection chain begins at LINAC2, proceeding through the booster, PS, and SPS before reaching the LHC. The facility also provides ions to the LHC, as well as a variety of particles to other experiments.

The LHC itself is the last stage of a chain of lower-energy accelerators, shown in figure ?? [Benedikt:2004wm]. The staged acceleration chain preserves the high-quality of the beam over many decades of energy. It consists of a linear accelerator and three synchrotrons which were used for previous generations of experiments at CERN. The reuse of the

older accelerators is economical while still meeting the stringent performance requirements of the LHC, providing up to 2808 proton bunches with a very small transverse emittance and controllable longitudinal emittance.

Protons are produced from hydrogen gas using a duoplasmatron source, which strips electrons from protons in a high electric field. After passing through a 90 kV pre-injector, a radio frequency quadrupole (RFQ) focuses and accelerates the protons to 750 kV. A linear accelerator (LINAC2) then accelerates the protons to 50 MeV using RF cavities. The protons then pass through an 80 m-long transfer line into the the Proton Synchrotron Booster (PSB) and Proton Synchrotron (PS).

The PSB consists of four stacked circular synchrotrons, 157 m in circumference, and accelerates the protons to 1.4 GeV. The use of four separate rings mitigates the space charge effects caused by the repulsion of protons within a bunch, which scale as $N_b/(\beta\gamma^2)$, where N_b is the number of protons per bunch. The protons are then injected into single-ring PS, where the higher injection energy reduces the space charge effect. The RF cavities of the PS, operating at several frequencies, accelerate the beams to 26 GeV, and also split the protons into the bunches eventually inject into the LHC. Nominally, this yields 72 bunches separated by 25 ns, but for Run I, 50 ns spacing was used instead.

The protons are extracted from the PS at intervals of 3.6 s and injected into the third synchrotron in the chain, the 7 km-circumference Super Proton Synchrotron (SPS). Immediately prior to extraction, the bunches are rotated by increasing the RF voltage, reducing the longitudinal emittance in order to ease capture in the SPS RF buckets, which have a frequency of 200 MHz. Up to four PS batches are injected per SPS cycle, after which the particle are accelerated at an average of 78 GeV/s to the LHC injection energy of 450 GeV. Flat-top is maintained for about one second, during which the injection is prepared:

- The magnets used for the beam extraction are ramped, safety checks are performed, and the SPS phase is tuned to match that of the LHC.
- The bunch length is compressed using an RF voltage increase, as in the PS-SPS transfer.
- The tails of the bunches are removed, down to $3-3.5\sigma$.

The SPS cycle takes 21.6 seconds, leading to a total LHC filling time of about nine minutes.

LHC Main Ring

The LHC main ring accelerates protons from the injection energy of 450 GeV to the collision energy, which ranged from 3.5 TeV to 4 TeV for proton-proton collisions during Run I [Bruning:2004ej]. The 26.7 km ring consists of eight arcs and eight straight sections, shown in figure ???. Each straight section is called an *insertion region* (IR), and contain either collider experiments or important services. IRs 1, 2, 5, and 8 contain the ATLAS, LHCb,

CMS, and ALICE experiments, respectively; IR 4 contains two independent 400 MHz RF systems, one for each of the counter-rotating beams; IR 6 contains the beam dump system; and IRs 3 and 7 contain collimation systems. The beams are contained in separate beam pipes except at the collision points.

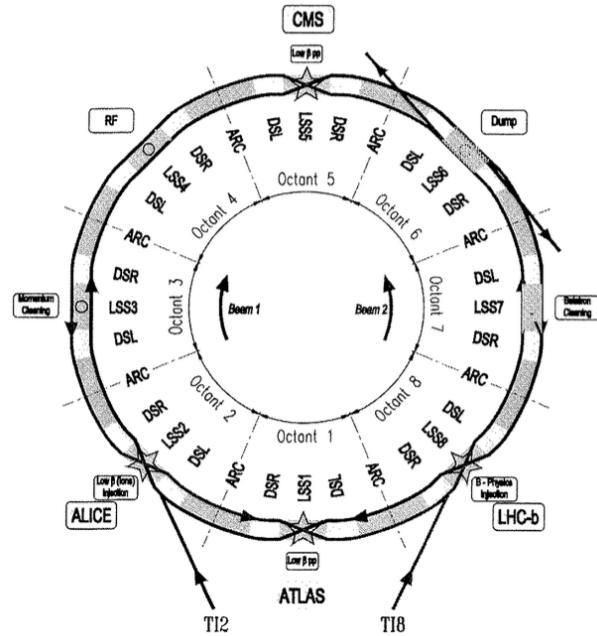


Figure 3.4: Layout of the LHC. The ring consists of eight arcs and eight long straight sections (LSS). Each junction between an arc and a LSS contains a dispersion suppressor cell (DSL, DSR). TI2 and TI8 are the two injection tunnels (“tunnel d’injection”) leading from the SPS to the LHC.

The beams are controlled by over 7,000 magnets. The primary bending and focusing are performed by 1232 superconducting dipole and 392 superconducting quadrupole magnets. The dipole magnets, shown in figure ??, have a nominal maximum field of 8.33 T and a length of 15 m. The conducting coils are constructed from Nb-Ti Rutherford cables, and are cooled to 1.9 K using superfluid helium. The dipoles have a double-bore structure, so that a single magnet provides the bending field for both beams.

The acceleration is provided by two superconducting RF systems, one for each beam. The RF cavities are constructed from copper with a thin ($1\text{ }\mu\text{m}$ to $2\text{ }\mu\text{m}$) layer of niobium, and operate at 4.5 K. The single-cell cavities, each nominally providing 2 MV, are arranged into cryomodules containing eight cells, for a total peak voltage of 16 MV. A module consisting of two cells is shown in figure ?? . The RF frequency is 400 MHz, giving an RF bucket length of 2.5 ns, or 75 cm.

At the collision points, the two beams converge in a single pipe for approximately 130 m. A steering dipole directs the beams into collision, and a triplet of quadrupoles on each side

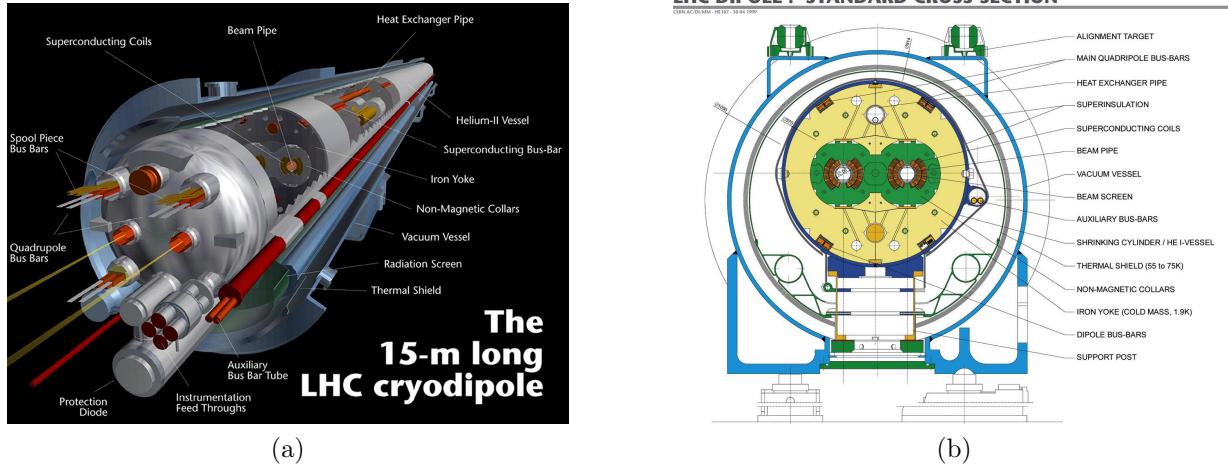
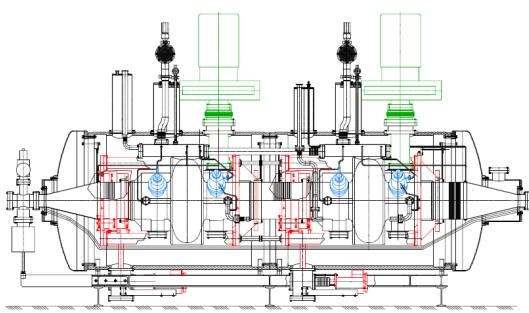


Figure 3.5: Two views of the LHC superconducting dipole magnet. The magnets are 15 m long and have two bores, one for each of the counter-rotating beams.



(a) A prototype cryomodule containing two cavities.



(b) The RF cryomodules installed at IR4.

Figure 3.6: The LHC RF cryomodules.

of the interaction region focus the beam onto the interaction point, squeezing the beams in the transverse directions from a typical orbiting width of $\mathcal{O}(1\text{ mm})$ to a collision width of $\mathcal{O}(10\text{ }\mu\text{m})$.

Beam Parameters

From the experiments' point of view, there are two main parameters to optimize in order to maximize sensitivity to new physics: the collision energy, \sqrt{s} , and the integrated luminosity, $L = \int \mathcal{L} dt$. The collision energy is limited to $\sqrt{s} = 14\text{ TeV}$ by the bending power of the dipole magnets, which have a nominal field strength of 8.33 T; however, due to the faulty splice design mentioned above, the energy was limited to $\sqrt{s} = 7\text{--}8\text{ TeV}$ in Run I.

The optimization of the integrated luminosity is somewhat more complicated. The instantaneous luminosity of two colliding bunches is given by:

$$\mathcal{L} = \frac{f_r N_1 N_2 \gamma_b}{4\pi \epsilon_n \beta_*} F(\alpha), \quad (3.1)$$

where $f_r = 11\,245.5\text{ Hz}$ is the LHC revolution frequency, $N_{1,2}$ are the numbers of protons in the two beams, γ_b is the relativistic gamma factor, ϵ_n is the normalized transverse emittance, β_* is the beta function at the collision point, and $F(\alpha)$ is the reduction in luminosity due to a nonzero beam crossing half-angle of α . The total instantaneous luminosity is given by the sum of all colliding bunch pairs.

In general, a higher total integrated luminosity is desired. This depends on a number of factors:

- The instantaneous luminosity per bunch, $\mathcal{L} = \frac{f_r N_1 N_2}{4\pi \epsilon_n \beta_*}$. A higher \mathcal{L} increases the number of simultaneous collisions per bunch crossing (*pileup*), $\mu = L_b / \sigma_{\text{inel}}$, which degrades the performance of the detectors.
- The number of bunches in the LHC, n_b . During Run I, the LHC was filled with 1380 bunches with 50 ns spacing between bunches, as opposed to the nominal 2808 bunches with 25 ns spacing. The advantages of 50 ns spacing include less days needed for electron-cloud scrubbing, smaller transverse emittances, and a smaller crossing angle [Papotti:2014vb]. The primary disadvantage is a higher pileup for the experiments.
- The crossing angle, α . A nonzero crossing angle is required to prevent *parasitic collisions* between bunches away from the nominal collision point. For bunches spaced by 25 ns, there are 34 unwanted parasitic collision points inside the common beam pipe at each interaction region. The reduction in luminosity is:

$$F(\alpha) = \frac{1}{\sqrt{1 + \left(\frac{2\alpha\sigma_z}{2\sigma^*}\right)}}, \quad (3.2)$$

where σ_z is the RMS bunch length and σ^* is the transverse beam width at the collision point. During normal 2012 running conditions, the crossing half-angle was $\alpha = 145 \mu\text{rad}$, and the luminosity reduction factor was about 18%.

- The fill schedule of the LHC. The instantaneous luminosity of the LHC decreases over time. The dominant loss of luminosity is due to the loss of protons from the collisions themselves, with a typical decay constant of $\tau_{\text{nuclear},1/e} = 29 \text{ h}$. Including other sources of loss such as intra-bunch scattering, scattering from collisions with gas in the beam pipe, beam-beam effects, and RF noise, the typical luminosity lifetime is $\tau_L = 14.9 \text{ h}$.

The fill schedule can be optimized to maximize the luminosity based on τ_L and the average turnaround time to refill the LHC after a beam dump, $T_{\text{turnaround}}$. This can be quantified using the *Hübner factor*, HF_{peak} , defined as the ratio of the delivered integrated luminosity to the hypothetical integrated luminosity with $\tau_L = \infty$ and $T_{\text{turnaround}} = 0$. In 2011, this was typically in the range 15%–20%.

The delivered luminosity in 2011 and 2012 will be discussed later in chapter ??.

3.2 The ATLAS Experiment

The ATLAS detector [TheATLASCollaboration:2008fg] is a large, cylindrical collider detector located at IR1 on the LHC ring (figure ??). The detector measures the energy and momenta of particles produced in the collisions provided by the LHC. It consists of several subsystems occupying a cylinder with a diameter of 25 m and a length of 46 m, with a combined weight of approximately 7,000 tons. Closest to the interaction region, the inner detector measures the momenta of charged particles by tracking their movement through a solenoidal magnetic field. Outside the inner detector solenoid magnet, electromagnetic and hadronic calorimeters measure the energy of electrons, photons, and hadrons. Finally, the muon spectrometer provides additional tracking and particle identification for muons in large toroidal magnetic field.

This section describes the design of the magnets, inner detector, calorimeters, and muon spectrometer. The reconstruction of physics objects from the detector measurements is described in chapter ??.

Coordinate System

The ATLAS detector uses a right-handed coordinate system, with \hat{x} pointing from the interaction point towards the center of the LHC ring, \hat{y} pointing up, and \hat{z} pointing along the ring towards IR2. Due to the symmetry of the detector, cylindrical coordinates (r, θ, ϕ) are often used, with r the transverse distance from the beam line, θ the polar angle from the beamline, and ϕ the azimuthal angle in the x - y plane. The polar angle can also be expressed in term of the pseudorapidity, $\eta = -\log \tan \frac{\theta}{2}$, which is useful for describing the geometry of

particles in part because differences in pseudorapidity, $\Delta\eta$, are invariant under longitudinal Lorentz boosts along \hat{z} . For massless particles, the pseudorapidity is equal to the rapidity, $y = \frac{1}{2} \log \frac{E+p_z}{E-p_z}$.

Magnets

ATLAS relies on four powerful superconducting magnets to bend the trajectories of charged particles, allowing the tracking detectors to provide measurements of their momenta. The solenoid provides a 2 T axial magnetic field in the volume of the inner detector, while the barrel and two endcap toroids provide a toroidal magnetic field ranging from 0.5–1 T for the muon spectrometer. The geometry of the magnets is shown in figure ???. Key parameters for the magnet system are shown in table ??.

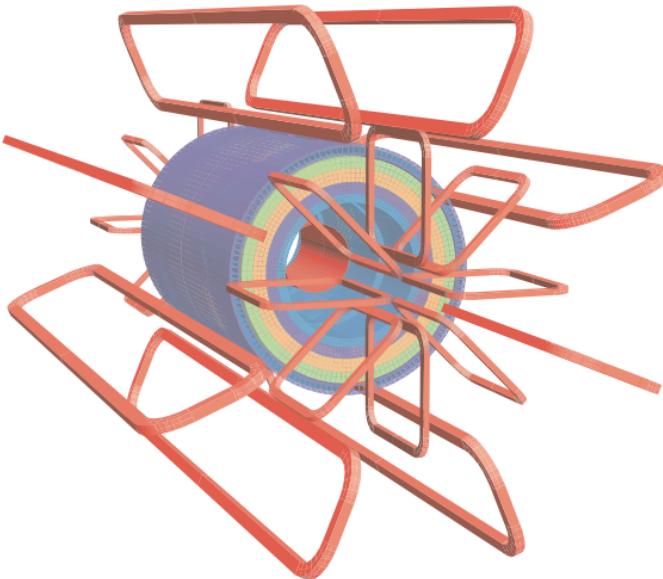


Figure 3.7: The geometry of the magnet windings and tile calorimeter steel. The three toroids and solenoid are shown in red. The remaining colors show layers of the tile calorimeter with different magnetic properties and an outside return yoke [[atlasphotos](#)].

Solenoid

The central solenoid, shown in figure ??, occupies the volume between the inner detector and the electromagnetic calorimeter, with an inner radius of 2.46 m, an outer radius of 2.56 m, and a length of 5.8 m. The single coil has 1154 windings made of high strength Al-stabilized NbTi conductor. With a nominal current of 7.73 kA, the magnetic field is 1.998 T at the center of the solenoid, falling to 1.8 T at $z = 1.7$ m and 0.9 T at the end of the inner detector

Property	Feature	Unit	Solenoid	Barrel toroid	End-cap toroids
Size	Inner diameter	m	2.46	9.4	1.65
	Outer diameter	m	2.56	20.1	10.7
	Axial length	m	5.8	25.3	5.0
	Number of coils		1	8	2×8
Mass	Conductor	t	3.8	118	2×20.5
	Cold mass	t	5.4	370	2×140
	Total assembly	t	5.7	830	2×239
Coils	Turns per coil		1154	120	116
	Nominal current	kA	7.73	20.5	20.5
	Magnet stored energy	GJ	0.04	1.08	2×0.25
	Peak field in the windings	T	2.6	3.9	4.1
	Field range in the bore	T	0.9-2.0	0.2-2.5	0.2-3.5

Table 3.1: Main parameters of the ATLAS magnet system [TheATLASCollaboration:2008fg].

cavity. The magnetic flux is returned via the steel in the hadronic calorimeter and its support structures. Liquid helium is used to cool the superconducting coil to an operating temperature of 4.5 K. At nominal current, the stored energy is 40 MJ.

The longitudinal and radial magnetic field components are shown for different R and z values in figure ??.



Figure 3.8: The central solenoid in the factory after completion of the coil winding [atlas-photos].

Toroid

The magnetic field for the muon spectrometer is provided by three large toroid magnets, each with eight superconducting coils. The magnets are shown in figure ???. The barrel toroid measures 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m, respectively.

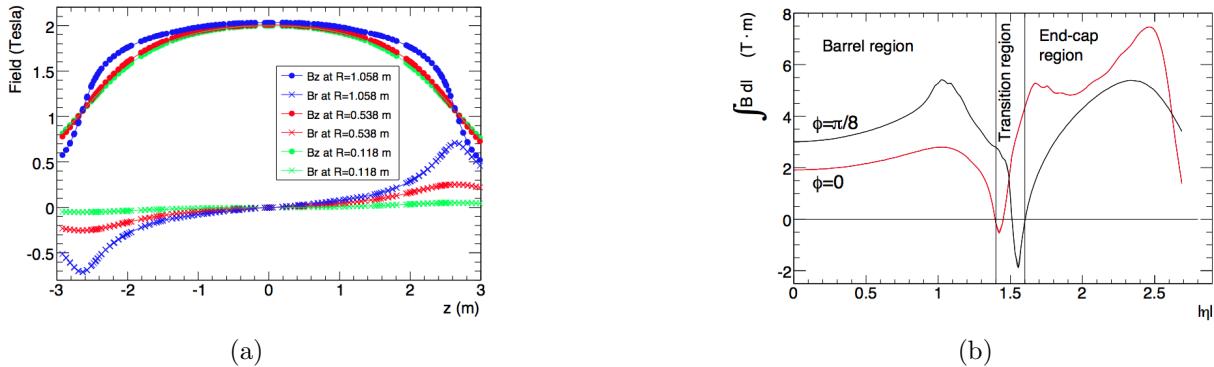


Figure 3.9: Left: the solenoid magnetic field in the radial (B_r) and longitudinal (B_z) directions, shown as a function of z for $\phi = 20\pi/16$ at different values of R . The field determined from a fit of the field model to measurements using an array of Hall probes [Aleksa:2008br]. Right: The predicted toroid magnetic field integral as a function of $|\eta|$. The integral is taken over a straight line through the interaction point, from the innermost to the outermost muon detector (section ??).

The two end-cap toroids are 5.0 m in length, with inner and outer diameters of 1.65 m and 10.7 m, respectively. Both types of toroid contain eight coils, with 120 windings per coil in the barrel and 116 windings per coil in the end-caps. The end-cap coils are rotated by 22.5° from the barrel coils to optimize the bending power in the overlap region between the magnets. Like the solenoid, the conductor is Al-stabilized NbTi, operated at 4.5 K. The nominal current is 20.5 kA, producing a magnetic field that varies from 0.15 T to 2.5 T in the barrel region, and 0.2 T to 3.5 T in the end-caps. The total stored energy at nominal current is 1.58 GJ.

The field integral of the toroid is shown as a function of η in figure ???. The field integral drops at the boundary between the barrel and end-caps, where the fields from the two magnets partially cancel.

Inner Detector

The inner detector performs tracking of charged particles traversing the 2 T solenoidal magnetic field in the pseudorapidity range $|\eta| < 2.5$. It also performs electron identification in the range $|\eta| < 2.0$. It consists of three subdetectors occupying the volume closest to the interaction region, directly outside the beam pipe, as shown in figure ???. Proceeding outwards from the interaction point, these are the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). The pixel detector provides three track measurements with high spatial resolution using silicon pixels. The SCT provides an additional four measurements using silicon strips. Finally, for particles with $|\eta| < 2.0$, the TRT provides an average of 36 measurements per track using gas-filled straw tubes, which aids

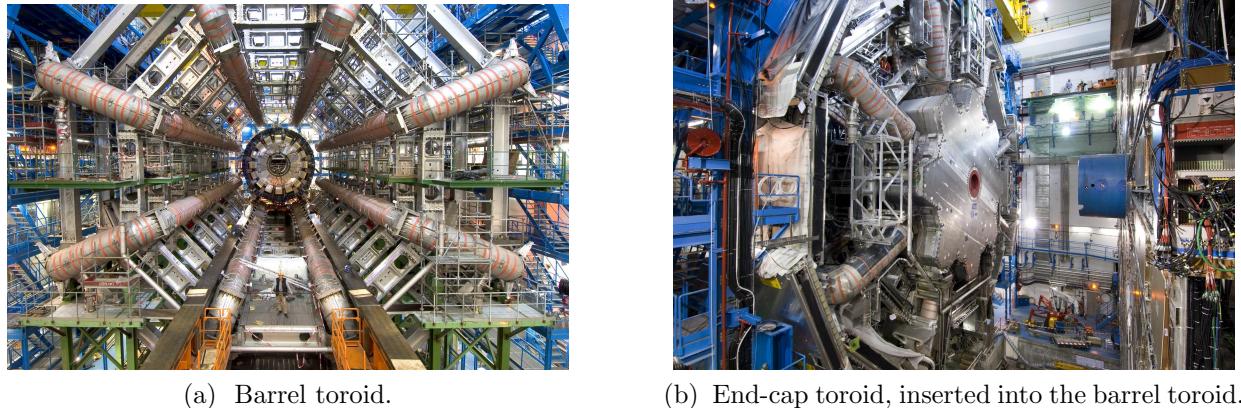
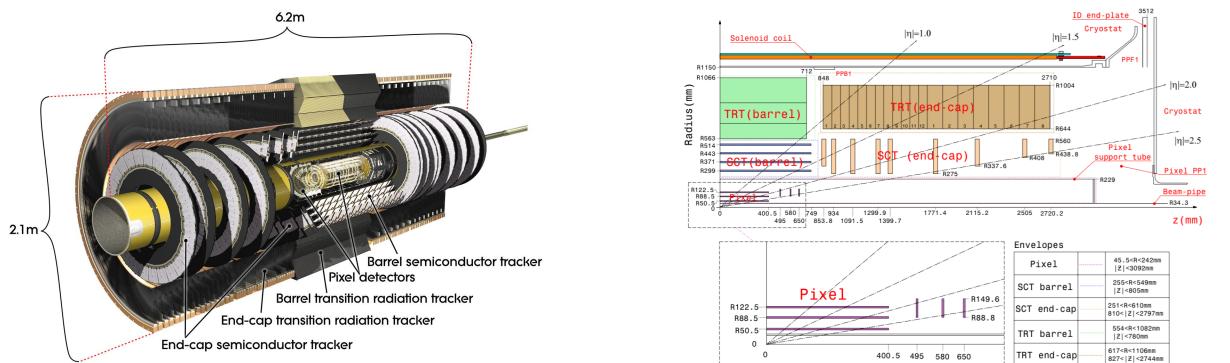
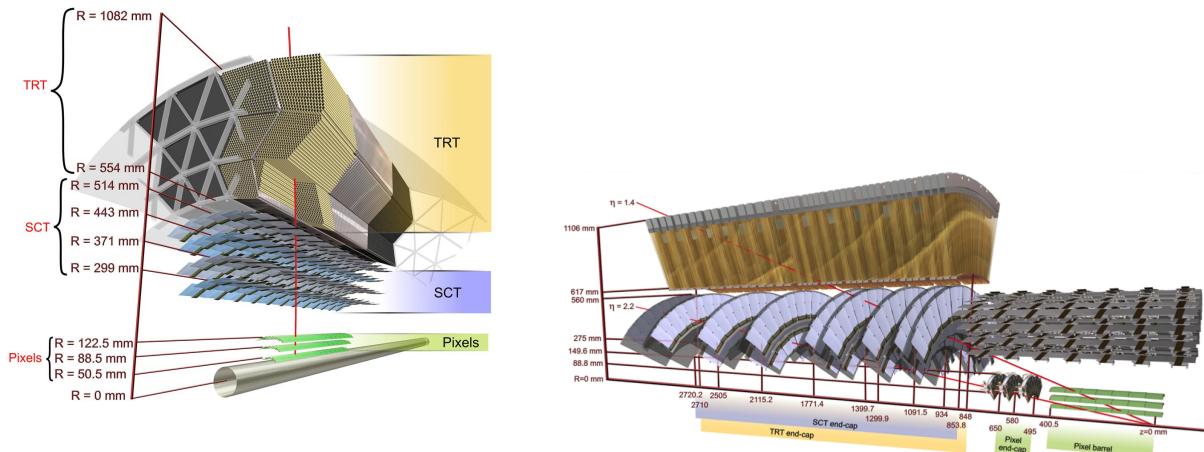


Figure 3.10: Images of the barrel and end-cap toroid magnets during installation.

the pattern recognition, improves the momentum resolution, and performs electron identification. Examples of trajectories of 10 GeV particles through the barrel and end-cap layers are shown in figure ??.





- (a) Trajectory of a 10 GeV charged particle with $\eta = 0.3$ traversing the barrel elements of the inner detector.
- (b) Trajectories of 10 GeV charged particles with $\eta = 1.4$ and $\eta = 2.2$ traversing the barrel and end-cap elements of the inner detector.

Figure 3.12: Drawings of the inner detector sensors and structural elements traversed by 10 GeV charged particles originating from the interaction point at various angles.

reasons, eight pairs of pixels in each column are ganged to a common readout, giving a total of 46080 readout channels per sensor.

The sensors are 256 μm -thick detectors utilizing n -in- n technology, constructed from n -type wafers with high dose positive (p^+) and negative (n^+) dose regions implanted on each side of a wafer. Initially, the asymmetric depletion region at the $p^+ - n$ junction is operated in reverse bias with a voltage of 150 V, and fills the sensor bulk volume, shown in figure ???. The charge carriers generated by the passage of an ionizing particle through the bulk are collected at the n^+ side of the sensor, where the readout electronics are bump-bonded to the pixel. The nominal threshold for readout is about 3,500 electrons, while a minimally ionizing particle crossing a pixel at normal incidence produces a signal of about 20,000 electrons. Over time, radiation damage induces type inversion in the bulk, after which the junction moves to the n^+ side of the sensor and the depletion zone grows from the pixel side, as shown in figure ???. The double-sided construction thus allows the pixel sensors to continue operating after type inversion.

The pixel sensors are assembled into pixel modules, shown in figure ???. Each module contains 16 front-end electronics chips each with 2880 electronics channels. The front-ends are bump bonded to the pixel sensor elements. The other side of the pixel sensor tile is glued to a flexible polyimide printed circuit board (*flex-hybrid*) that houses the module control chip.

The layout of the pixel detector is summarized in table ???. The modules are assembled into staves in the barrel region, each with 13 pixel modules, and sectors in the end-caps, each with 6 pixel modules. The barrel layers consist of 22, 38, and 52 staves for layers 0, 1,

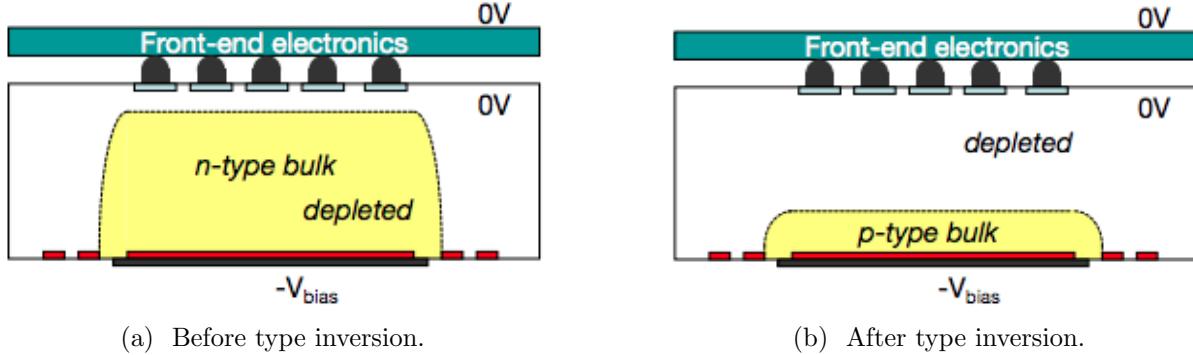


Figure 3.13: Comparison of depletion zones in n^+ -in- n pixel sensors before and after type inversion. Before type inversion, the electrical field grows from the bottom side, reaching the pixel implants at full depletion. After type inversion, the depletion zone grows from the pixel side, allowing operation even if the bulk is not fully depleted.

and 2, respectively, while the end-caps each contain eight sectors². In the barrel, to provide complete coverage, the pixel staves overlap and are mounted with a tilt angle of 20° between the normal to the module surface and \hat{r} , in the $\hat{\phi}$ direction.

In the barrel, the pixels have an intrinsic accuracy of $10\text{ }\mu\text{m}$ in the $R - \phi$ direction and $115\text{ }\mu\text{m}$ in the z direction, while in the end-caps, the intrinsic accuracy is $10\text{ }\mu\text{m}$ in the $R - \phi$ direction and $115\text{ }\mu\text{m}$ in the R direction.

Barrel	Radius (mm)	Staves	Modules	Pixels
Layer-0	50.5	22	286	13.2×10^6
Layer-1	88.5	38	494	22.8×10^6
Layer-2	112.5	52	676	31.2×10^6
End-cap (one side)	z (mm)	Sectors	Modules	Pixels
Disk 1	495	8	48	2.2×10^6
Disk 2	495	8	48	2.2×10^6
Disk 3	495	8	48	2.2×10^6
Barrel and both end-caps			1744	80.4×10^6

Table 3.2: Parameters of the pixel detector, from [TheATLASCollaboration:2008fg].

²During the shutdown following Run I, a fourth layer of barrel pixel detectors, known as the insertable B-layer (IBL), was added between the beam pipe and the innermost pixel layer. The data used in this dissertation predate the inclusion of the IBL.

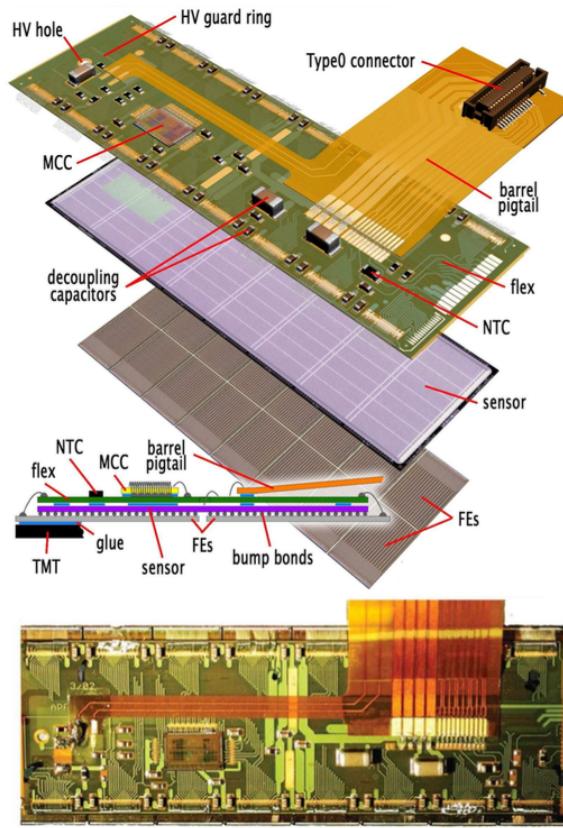


Figure 3.14: Schematic pictures of a barrel pixel module. The top diagram shows the assembly of a module, consisting of the front-end electronics chips (FEs), the pixel sensor elements, and the flex-hybrid, which bears the module control chip (MCC) and NTC thermistors. The bottom is a photograph of a barrel pixel module.

SCT

The SCT consists of 15912 sensors using single-sided, p -in- n type silicon strips. The use of single-sided strips lowers both the cost and the number of readout channels of the detector, which covers significantly more area than the pixel detector (61.1 m^2 versus 2.3 m^2). In the barrel, the strips have a pitch of $80 \mu\text{m}$ and a length of 12 cm. In the endcaps, the sensors are trapezoidal with radially arranged strips, with a mean strip pitch of $80 \mu\text{m}$ and a length of 12 cm. Each sensor has 768 active strips. Like the pixel sensors, the operating voltage is initially $\sim 150 \text{ V}$, but will require an increase to $250 \text{ V}-350 \text{ V}$ after several years of irradiation, depending on the location of the sensor.

The SCT modules are shown in figure ???. The 2112 barrel modules contain four sensors, while the 1976 end-cap modules contain two sensors. In both cases, the sensors are glued to a $380 \mu\text{m}$ -thick thermal pyrolytic graphite (TPG) base-board. The sensors are assembled

in two layers with a rotation of ± 20 mrad about the center of the sensors. The stereo angle between the two sensors allows for a measurement of the position along the length of the sensors.

The intrinsic accuracy of the SCT sensors is $17\text{ }\mu\text{m}$ along the strip pitch, corresponding to the $R - \phi$ direction. Along the length of the strips, the stereo angle allows for position measurement with an accuracy of $580\text{ }\mu\text{m}$, corresponding to the z direction in the barrel and the R direction in the end-caps.

Barrel layer	Radius (mm)	Length (mm)	Module tilt angle	Number of modules
3	299	1498	11.00°	384
4	371		11.00°	480
5	443		11.25°	576
6	514		11.25°	672

Table 3.3: The dimensions and arrangement of modules in the four SCT layers. The tilt angle is between the normal to the module surface and \hat{r} , in the $\hat{\phi}$ direction.

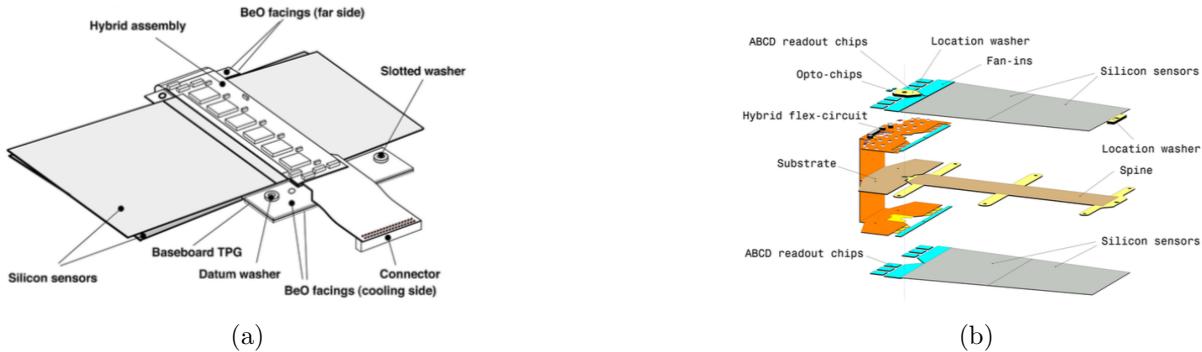


Figure 3.15: Diagrams of the SCT barrel (left) and end-cap (right) modules.

TRT

The TRT detector elements are polyimide drift (straw) tubes. The straws are arranged such that charged particles with $p_T > 500\text{ MeV}$ and $|\eta| < 2.0$ traverse at least 36 straws, except in the barrel/end-cap transition region where particles traverse as few as 22 straws. The straws are interleaved with polypropylene fibers in the barrel and polypropylene foils in the end-caps to provide transition radiation; electrons with $p_T > 2.0\text{ GeV}$ typically leave 7-10 high-threshold hits. The intrinsic resolution of the straws is $130\text{ }\mu\text{m}$ in the transverse direction.

The straws measure 4 mm in diameter and 144 cm (37 cm) in length in the barrel (end-caps). The straw walls consist of two $35\text{ }\mu\text{m}$ aluminum-coated polyamide films bonded back-to-back. The anodes of the detectors are gold-plated tungsten wires running down the axis of the tubes, with $31\text{ }\mu\text{m}$ diameter. The wires are supported at the end of the straw by

an end plug, where they connect to the front-end electronics. At the middle of the straw, the wires are supported by a plastic insert, and are also split electrically by a fused glass capillary to reduce occupancy. The active length of each half of the wire is 71.2 cm, with a 2 cm inefficient section at the middle. The straws are filled with a mixture of 70% Xe, 27% CO₂, and 3% O₂. With a cathode voltage of -1530 V, the gain in the straws is 2.5×10^4 .

The layout of the straws in the barrel and end-caps is summarized in table ???. In the barrel, the straws are arranged into three layers of modules as shown in figure ??, with different dimensions and straw counts depending on the layer. The straws are interleaved with transition radiation material, polypropylene fibers measuring 19 μm in diameter. The module shell is made from 400 μm thick carbon fiber, and serve not only as a support structure, but also as a gas manifold for CO₂, which prevents high-voltage discharges, flushes any Xe leaking from the straws, and conducts heat away from the straws.

The endcaps consist of 160 layers of 768 straws, arranged radially into 20 wheels with 8 layers each. The inner 12 wheels (type-A) are spaced 8 mm in z , while the outer 8 wheels (type-B) are spaced by 15 mm. Each successive layer in a wheel is separated by a 15 μm thick polypropylene radiator foil, and is rotated by 3/8 of the azimuthal straw spacing to optimize the uniformity of the number of straw crossed.

	z (mm)	R (mm)	Modules	Layers	Straws/Module
Barrel (both sides)	0-780	554-1082	96	73	52544
Type-1 module (inner)	400-712.1	563-624	32	9	329
Type-1 module (outer)	7.5-712.1	625-694	32	10	329
Type-2 module	7.5-712.1	697-860	32	24	520
Type-3 module	7.5-712.1	863-1066	32	30	793
End-cap (one side)	827-2744	615-1106	20	160	122880
Type-A wheels	848-1705	644-1004	12	8	6144
Type-B wheels	1740-2710	644-1004	8	8	6144

Table 3.4: Layout and straw counts of the TRT barrel modules and end-cap wheels. The totals for the barrel and end-caps, shown in bold, include services and electronics.

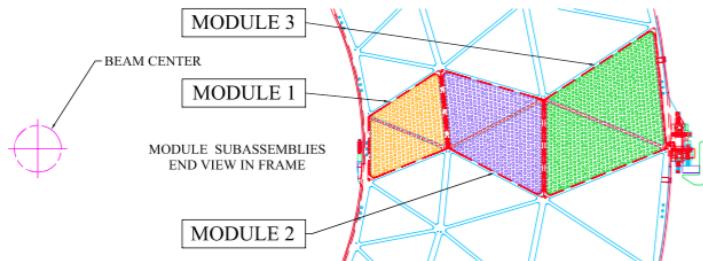


Figure 3.16: Layout of the TRT barrel modules in the R - ϕ plane.

Calorimetry

The calorimeter system occupies the volume directly exterior to the solenoid magnet, and measures the energy of electrons, photon, and hadrons up to a pseudorapidity of $|\eta| = 4.9$. The arrangement of the calorimeters is shown in figure ???. The system consists of a number of independent sampling calorimeters. The electromagnetic calorimeter measures the energy of electrons and photons using lead and liquid argon (LAr) to measure the energy of electrons and photons. The hadronic calorimeter measures the energy of hadrons using steel and scintillator tiles in the barrel and copper and LAr in the end-caps. Finally, electromagnetic and hadronic calorimetry is performed in the forward region using copper-LAr and tungsten-LAr technology. Besides performing energy measurements, the calorimeters stop electrons, photons, and hadrons, providing shielding for the muon spectrometer. The total amount of material in front of and in the calorimeters is shown in terms of radiation lengths, X_0 , and interaction lengths, λ , in figures ?? and ??.

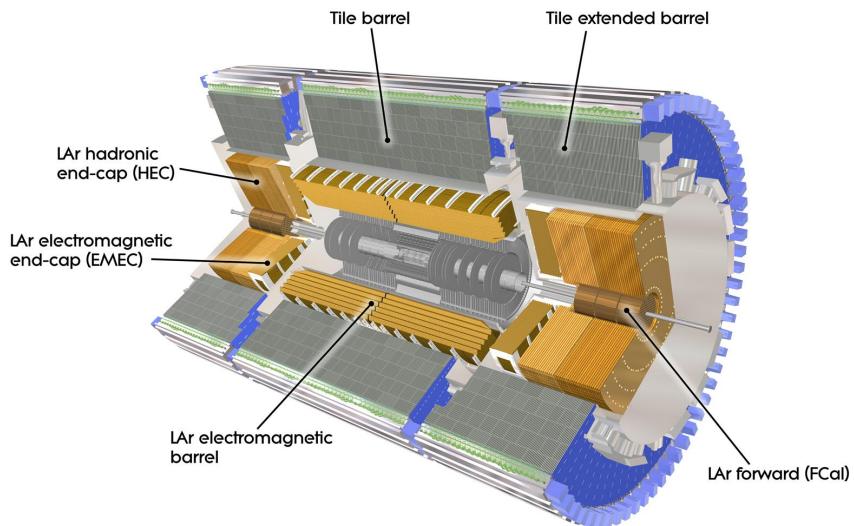


Figure 3.17: A 3D model of the ATLAS calorimeters, showing the electromagnetic barrel and end-caps, hadronic barrel and end-caps, and the forward calorimeter.

Electromagnetic Calorimeter

The electromagnetic calorimeter consists of two half-barrels ($0 < \eta < \pm 1.475$), two end-caps ($\pm 1.375 < \eta < \pm 3.2$), and a presampler ($|\eta| < 1.8$). The half-barrels and end-caps use LAr as the active material and steel-clad lead plates as the absorber. The lead plates are arranged in an accordion geometry to provide a uniform, gapless coverage in ϕ , shown for a barrel module in figure ???. The plates are interleaved with electrodes built from copper etchings on polyimide, consisting of three conducting layers. The outer conductive layers of

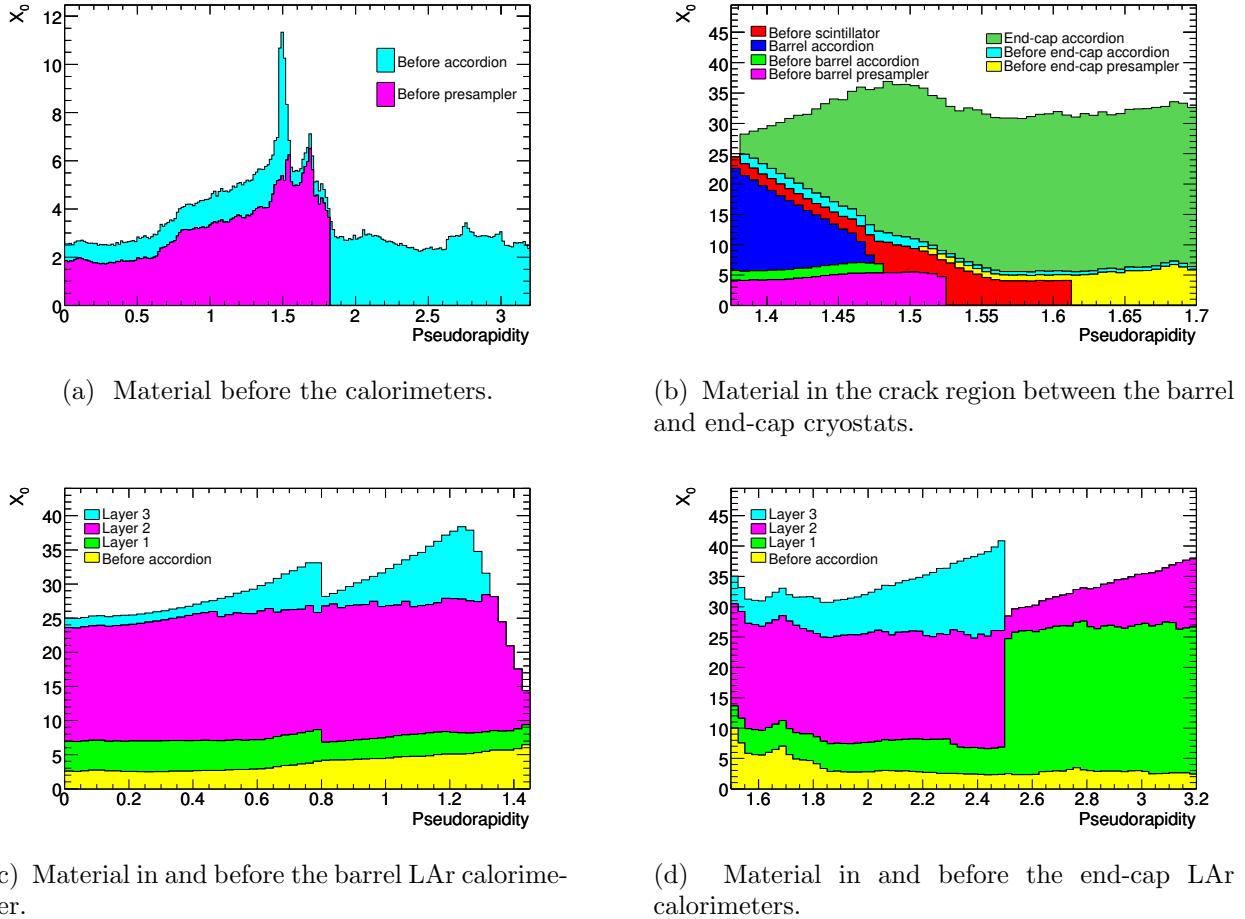


Figure 3.18: Cumulative amounts of material versus $|\eta|$ in front of and within the LAr electromagnetic calorimeter, in terms of radiation lengths, X_0 .

the electrodes distribute the 2 000 V high voltage over the electrode surface, which drifts the charges induced by ionization in the LAr towards the electrodes with a drift time of 450 ns. The inner layer of the electrode, separated from the outer layers by isolating foils, collects the signals via capacitive coupling.

The two half-barrels occupy the region $2.8 \text{ m} < R < 4 \text{ m}$ and $0 \text{ m} < z < \pm 3.2 \text{ m}$, which covers up to $|\eta| < 1.52$ depending on the layer. Each half-barrel has 1024 lead plates with a thickness of 1.53 mm for $|\eta| < 0.8$ and 1.13 mm for $|\eta| > 0.8$. The cells are divided into three layers, and have variable size in η and ϕ :

- The first layer is finely segmented in η to improve the spatial resolution; this improves the particle identification, helps resolve nearby particles (in particular photon pairs from π^0 decays), and, in combination with the second layer, allows for a measurement of particle trajectories. The majority of the cells measure $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$. Cells

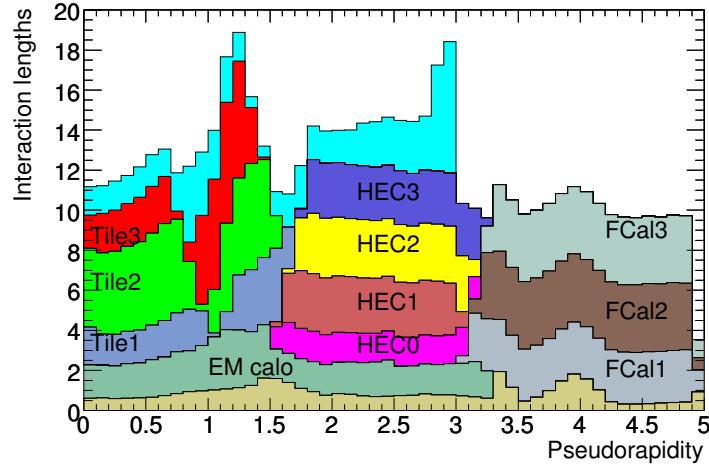


Figure 3.19: Cumulative amounts of material versus $|\eta|$ due to each calorimeter in terms of interaction lengths, λ . The material in front of the calorimeters is also shown in tan.

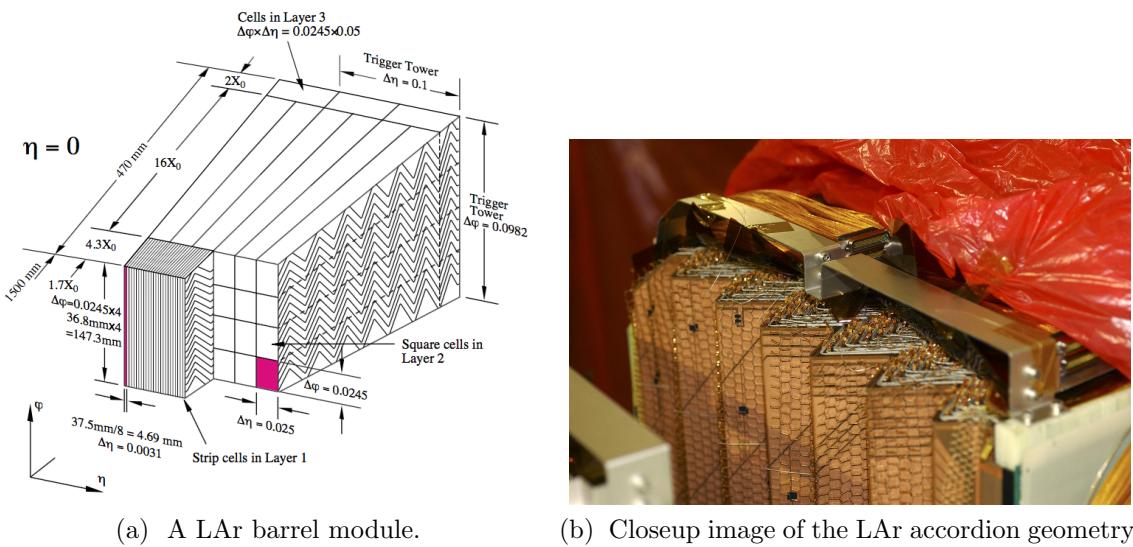


Figure 3.20: The LAr calorimeter uses an accordion geometry to provide uniform coverage in ϕ . Steel-clad lead plates are interleaved with electrodes, consisting of three copper layers separated by polyamide sheets. The barrel module is divided into three layers in depth, with fine segmentation in η in the layer closest the interaction point.

in the barrel-end-cap over lap region, $1.40 < |\eta| < 1.475$, have a coarser granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$.

- The second layer absorbs most of the energy of the electromagnetic showers. Most of the cells have size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, except for cells with $1.40 < |\eta| < 1.475$, which measure $\Delta\eta \times \Delta\phi = 0.075 \times 0.025$.
- The third layer absorbs the tails of electromagnetic showers, and has a coarse granularity of $\Delta\eta \times \Delta\phi = 0.050 \times 0.025$.

The end-caps, one on either side of the interaction region, measure 63 cm in thickness with inner and outer radii of 33 cm and 209.8 cm, respectively. Each end-cap consists of 1024 lead absorbers with the same accordion geometry as used in the barrel. In the region $1.5 < |\eta| < 2.5$, the end-caps present three layers in the longitudinal direction, with fine η segmentation in the front layer. At $|\eta| = 1.5$, the granularity of each layer is the same as in the barrel, becoming coarser in the first layer with increasing $|\eta|$. Elsewhere, there are two longitudinal layers with coarser transverse granularity.

Finally, the presampler is an instrumented layer of LAr installed in front of the first layer calorimeter, with a thickness of 11 mm in the barrel ($|\eta| < 1.5$) and 2×2 mm in the end-caps ($1.5 < |\eta| < 1.8$). The presampler measures the energy lost due to interactions with material between the interaction region and the calorimeters (the beam pipe, inner detector, solenoid, cryostates, and other services and support structures). The total amount of material before and in the electromagnetic calorimeters is shown in figure ?? in terms of radiation lengths, X_0 . The cells have a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$.

In total, the electromagnetic calorimeter has 101,760 readout channels for the barrel, 62,208 channels for the end-caps, 7,808 channels for the barrel presampler, and 1,536 channels for the end-cap presamplers.

Tile Calorimeter

The tile calorimeter performs hadronic calorimetry in the region $|\eta| < 1.7$, occupying the radial range $2.28 \text{ m} < R < 4.25 \text{ m}$. Divided into a 5.8 m-long central barrel and two 2.6 m-long extended barrels on either side, it uses steel as the absorber and scintillating tiles as the active material. The total depth is approximately 7.4λ .

Each of the three calorimeter sections is divided azimuthally into 64 modules, each spanning $\Delta\phi = 5.625$. A module is shown schematically in figure ???. The outer edge of a module is a steel girder which houses the tile calorimeter readout electronics and also provides flux return for the solenoidal magnetic field. The body of the module is a self-supporting structure built from steel absorber plates, with 4 mm-thick spacer plates glued in staggered fashion to 5 mm-thick master plates. The staggered spacing creates the gaps into which the scintillating tiles are inserted, with a steel-to-scintillator volume ratio of approximately 4.7 : 1.

The scintillating tiles use polystyrene as the base material, in which ionizing particles induce the production of ultraviolet light. The polystyrene is doped with wavelength-shifting

fluors, 1.5% PTP and 0.044% POPOP, which convert the scintillation light into the visible light. The tiles measure 3 mm thick, and vary between 97 mm-187 mm in radial length and 200 mm-400 mm in azimuthal length. A plastic sleeve envelops each tile, both for the protection of the tile during installation and also to improve the scintillation light collection efficiency due to a reflectivity of $\sim 95\%$. At the tile edges, wavelength-shifting fibers transmit the light to photomultiplier tubes housed in the girder. The grouping of the fibers determines the readout cell size, as shown in figure ??; the cells measure $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the transverse direction, and are segmented into three layers longitudinally corresponding to depths of approximately 1.5λ , 4.1λ , and 1.8λ .

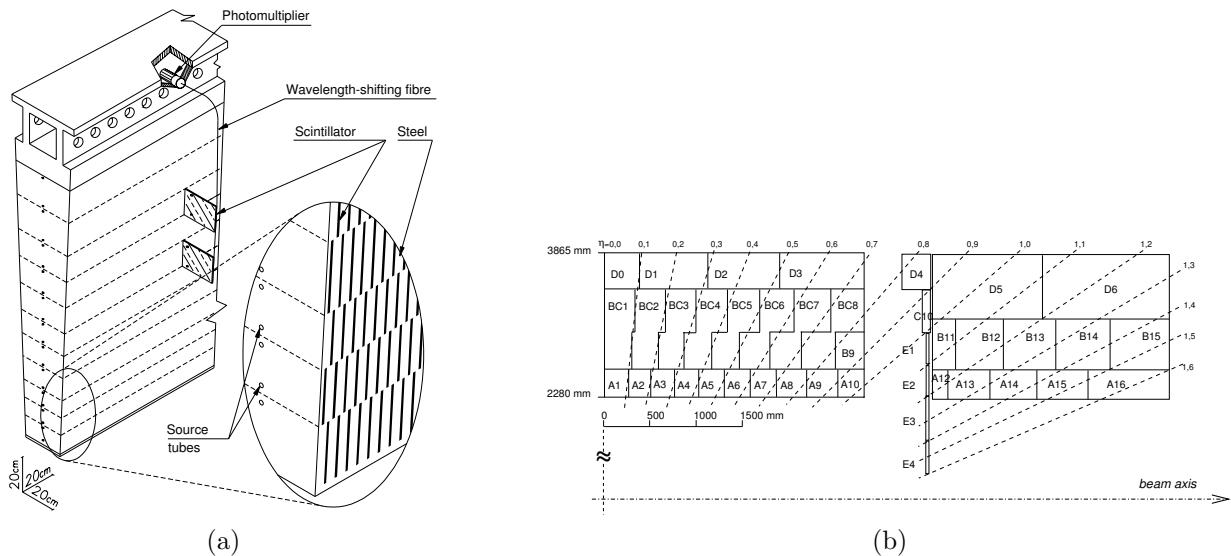


Figure 3.21: Left: Schematic view of a tile calorimeter module, showing assembly of the steel absorber, the inserted scintillator tiles, and the readout of lighting through wavelength-shifting fibers to the photomultiplier tubes. Right: Layout of the tile calorimeter cells, defined by the grouping of the readout fibers connecting the scintillating tiles to the photomultiplier tubes.

Hadronic End-Cap Calorimeters

Outside the tile calorimeter, two hadronic end-cap calorimeters (HEC) perform hadronic calorimetry in the pseudorapidity range $1.5 < |\eta| < 3.2$. The HECs use copper and LAr arranged into a flat plate geometry, as shown in figure ???. Each HEC consists of two wheels (HEC1 and HEC2), which are further divided into two longitudinal segments. The plates have an inner radius of 372 mm or 475 mm depending on z , and an outer radius of 2030 mm. HEC1 contains 24 copper plates, each 24 mm thick and spaced by 8.5 mm, plus a 12.5 mm-thick front plate. HEC2 contains 16 copper plates measuring 50 mm in thickness, with a

25 mm-thick front plate. The sampling fractions are thus 4.4% for HEC1 and 2.2% for HEC2.

The gaps contain 3 electrodes spaced by 1.8 mm, the outer two of which supply a nominal operating voltage of 1800 V, and the middle of which performs the readout. The readout pads provide a transverse segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$, and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $2.5 < |\eta| < 3.2$. In total, 5,632 channels are read from the HEC.

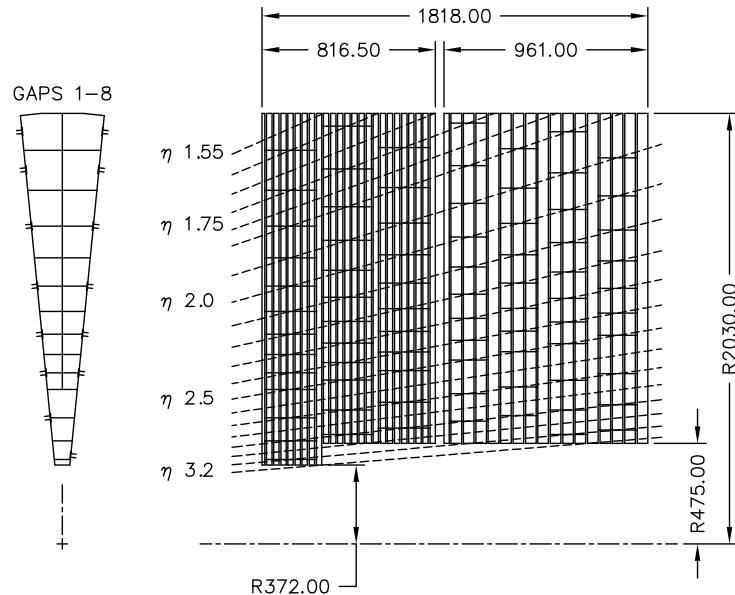


Figure 3.22: Schematic $R - \phi$ (left) and $R - z$ (right) views of the hadronic end-cap calorimeters.

Forward Calorimeter

The forward calorimeter (FCal) performs both electromagnetic and hadronic calorimetry in the very forward region, $3.2 < |\eta| < 4.9$, and also shields the muon system from high particle fluxes. On each side, the FCal is divided into three layers measuring 45 cm in depth, an electromagnetic layer (FCal1) and two hadronic layers (FCal2 and FCal 3). All three layers use LAr as the active medium, with very thin gaps due to avoid ion buildup due to the high particle flux. FCal1 uses copper absorbers to optimize the resolution and heat removal. FCal2 and FCal3 use tungsten absorbers, which provides good containment and reduces the lateral spread of hadronic showers. Finally, a passive brass plug behind FCal3 provides shielding for the muon system. The layout of the three layers and shielding plug is shown in figure ??.

FCal1 consists of stacked copper plates with 0.27 mm LAr gaps. The electrodes occupy 12,260 holes drilled through the copper plates, and consist of a copper rod coaxial with a copper tube, as shown in figure ??.. FCal2 and FCal3 consist of two 2.35 cm-thick copper

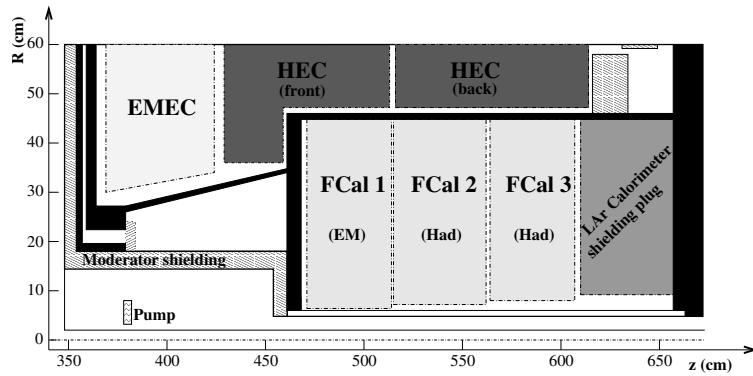


Figure 3.23: Schematic diagram showing the layout of the forward calorimeter in the R - z plane.

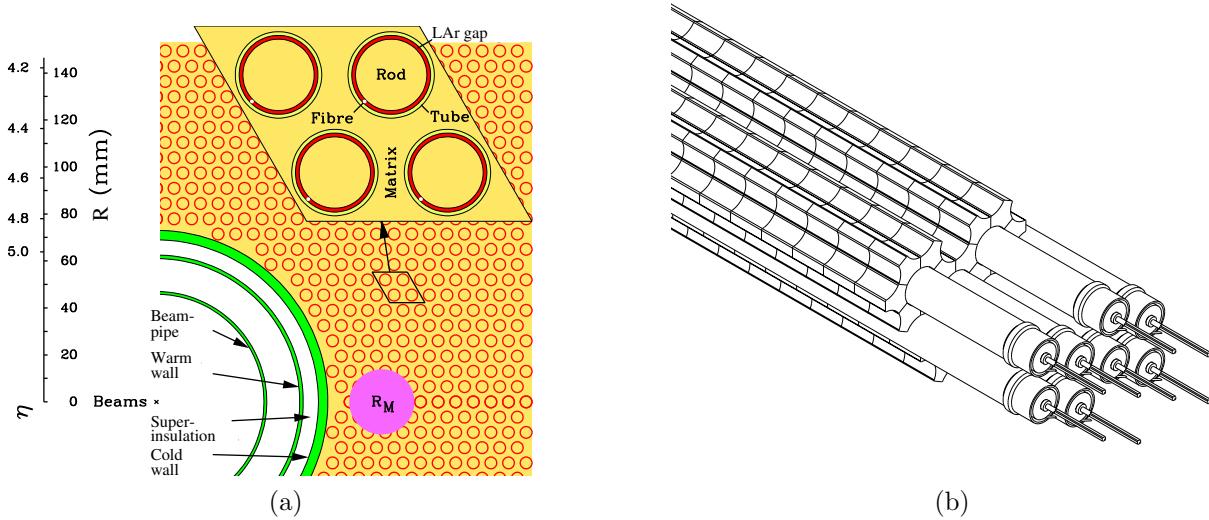


Figure 3.24: Left: the electrode structure of FCal1, showing the matrix of copper tubes and rods. The Molière radius, R_M , is shown for reference. Right: the absorber matrix in FCal2 and FCal3, made from tungsten rods and copper tubes.

end-plates spanned by an array electrodes similar to those in FCal1, except with tungsten rods, as shown in figure ???. The arrays contain 10,200 and 8,224 electrodes in FCal2 and FCal3, respectively. In total, the three layers, FCal1, FCal2, and FCal3, contain 1008, 500, and 254 readout channels, respectively, and constitute 208.1 radiation lengths and 9.94 interaction lengths.

Muon Spectrometer

The muon spectrometer detects charged particles that penetrate the barrel and end-cap calorimeters. It forms the exterior of the ATLAS detector, occupying the volume of the barrel toroid with four large wheels on each side of the interaction point. Using a combination of several technologies, the muon spectrometer performs precision measurements of muon momenta in the region $|\eta| < 2.7$, with a design resolution of 10% for 1 TeV tracks, and also provides triggering on muons for $|\eta| < 2.4$.

The structure of the muon spectrometer is shown in figures ?? and ???. The barrel systems are arranged in three layers with radii of approximately 5 m, 7.5 m, and 10 m, mounted inside and on the eight coils of the barrel toroid magnet with approximate octagonal symmetry. The end-cap systems form four large wheels, located before and after the end-cap toroid magnets at positions of $|z| \approx 7.4$ m, 10.8 m, 14 m, and 21.5 m.

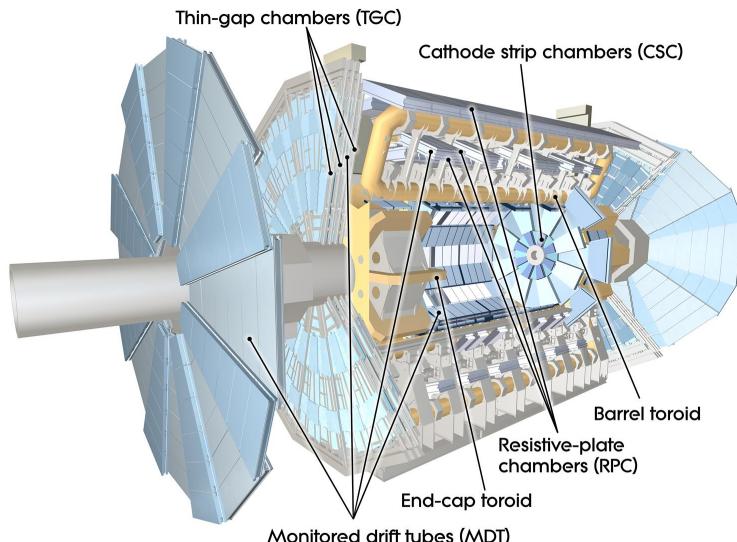


Figure 3.25: 3D model of the ATLAS muon spectrometer. The four different types of detector (MDTs, CSCs, RPCs, and TGCs) and the toroid magnets are shown.

Four different technologies are used for precision measurements and triggering, depending on the location in the detector. Monitored drift tubes (MDTs) perform the precision measurements over most of the volume with $|\eta| < 2.7$. Cathode strip chambers (CSCs) replace the MDTs in the innermost wheel for $2.0 < |\eta| < 2.7$, to handle the high particle flux. Triggering is performed by resistive plate chambers (RPCs) in the barrel ($|\eta| < 1.05$) and thin gap chambers (TGCs) in the end-caps. The trigger chambers also supplement the MDT position measurements with the hit coordinate in the non-bending plane.

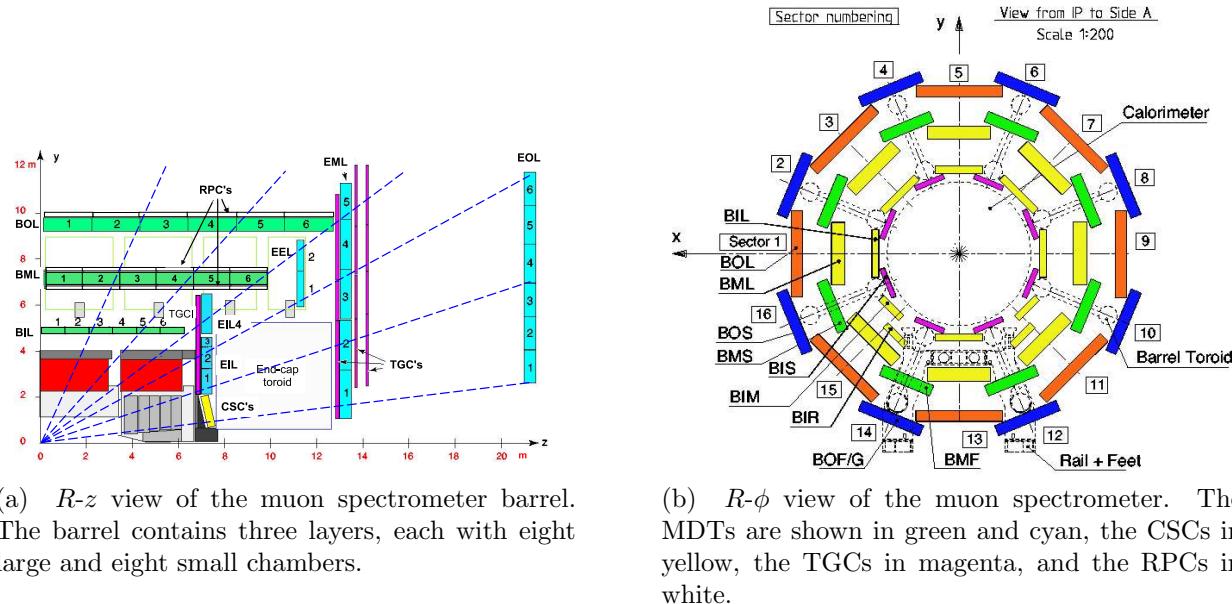


Figure 3.26: Layout of the chambers of the muon spectrometer.

Monitored Drift Tubes

The muon spectrometer contains 1088 MDT chambers, covering a total area of $\sim 550 \text{ m}^2$. The majority of the MDT chambers are rectangular in the barrel and trapezoidal in the end-caps, and are laid out to optimize the solid angle coverage under the constraints of the magnet coils and other structures in the vicinity of the toroid magnet. Each chamber contains several layers of drift tubes, as shown in figure ???. The innermost chambers are divided into two groups of tube layers, called *multilayers*, each with four layers of tubes, while the remaining chambers have two groups of three layers of tubes.

The aluminum drift tubes have a radius of 29.97 mm, with a 50 μm -diameter tungsten-rhenium wire forming the anode along the axis of the tube. A cross section of a tube is shown in figure ???. The tubes are filled with a mixture of argon and CO₂ (93:7) with a high pressure of 3 bar, and are operated with a voltage of 3080 V. The average resolution of a single drift tube is 80 μm , and the combined resolution of a chamber is 35 μm . To achieve this precision, the position of the drift tubes and wires must be known to less than 30 μm . Four optical alignment rays inside each MDT chamber continuously monitor the geometry of the chamber with a precision of a few microns, shown in figure ???. An inter-chamber optical alignment network monitors the relative positions of chambers relative to their neighbors with a precision of approximately 20 μm . The optical alignment is supplemented with track-based alignment algorithms to align the chambers with poor or absent connection to the optical network, the end-caps with respect to the barrel, and the muon spectrometer with respect to the inner detector.

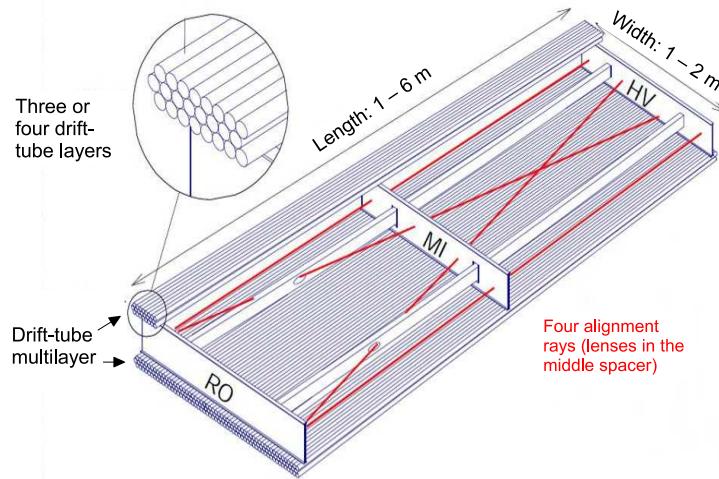


Figure 3.27: The mechanical structure of a MDT chamber. Two multilayers, each consisting of three or four rows of drift tubes, are separated by aluminum spaceers. Four optical alignment rays continuously monitor the geometry of the chamber.

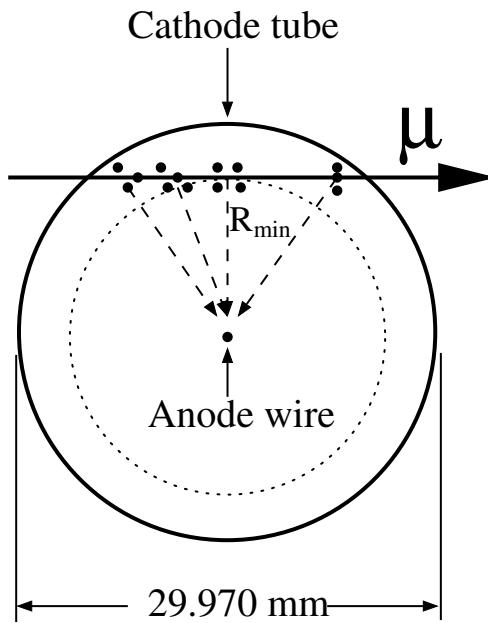


Figure 3.28: Cross section of a drift tube, showing the cathode tube, the anode wire, and an illustration of a muon ionizing the gas as it traverses the tube.

Cathode Strip Chambers

The rate limit for the safe operation of the MDTs is about 150 Hz/cm^2 , which is exceeded for $|\eta| > 2$ in the first layer of the end-cap at $|z| \approx 7 \text{ m}$. Accordingly, CSCs are used in this volume of the detector, which can operate up to rates of about 1 kHz/cm^2 . Eight large and eight small trapezoidal chambers give full coverage in ϕ , as shown in figure ??.

The CSCs are multiwire proportional chambers, with parallel wires running in the radial direction and the two cathodes finely segmented in perpendicular directions to provide measurements in both the η and ϕ directions. The perpendicular cathode segmentation provides position measurements in two dimensions, which helps distinguish hits from nearby particles. In the bending direction, the segmentation corresponds to a readout pitch of 5.31 mm and 5.56 mm for the large and small chambers, respectively. The track coordinate is determined from a relative measurement of the charge induced on 3-5 adjacent strips at the peak of the charge distribution, shown schematically in figure ???. The resolution, dominated by electronic noise in the pre-amplifiers and by the spread of charge along the anode wire due inclined tracks, delta electrons, or a Lorentz force along the wire, is roughly $60 \mu\text{m}$ per CSC plane. In the non-bending direction, a coarser segmentation leads to a resolution of 5 mm.

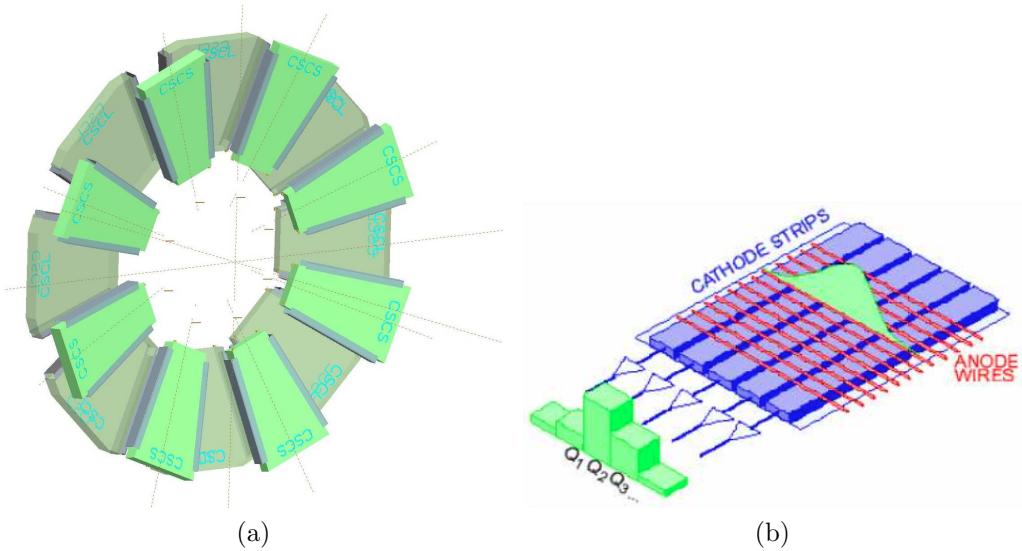


Figure 3.29: Left: Layout of the eight large and eight small CSC chambers. Right: Schematic view of the CSC anode wires and perpendicular cathode strips, showing the deposition of charge from a track on several adjacent strips.

Resistive Plate Chambers

Muon triggering in the barrel ($|\eta| < 1.05$) is performed by 544 RPCs, which exhibit good spatial and timing resolution and an adequate rate capability of $\sim 1 \text{ kHz/cm}^2$. The RPCs form three layers, or *stations*, with two on either side of the middle MDT layer and the

third on the inner or outer side of the outer MDT layer, for the small and large sectors, respectively. Each station has two independent layers, each providing a measurement of η and ϕ , resulting in six possible measurements for muons passing through three stations.

An RPC unit consists of two sets of two parallel resistive plates (2 mm-thick phenolic-melaminic plastic laminate) separated by 2 mm by insulating spacers, shown in figure ???. The interior is filled with a gas mixture of $C_2H_2F_4/Iso-C_4H_{10}/SF_6$ (94.7/5/0.3). With an electric field of 4.9 kV/mm, muons traversing the gas induce an avalanche towards the anode, which is read out via capacitive coupling to 25–35 mm-wide copper strips on the exterior face of the RPC. The RPCs achieve a spatial resolution of approximately 10 mm in z and ϕ , a timing resolution of 1.5 ns, and a detection efficiency of $\sim 98\%$.

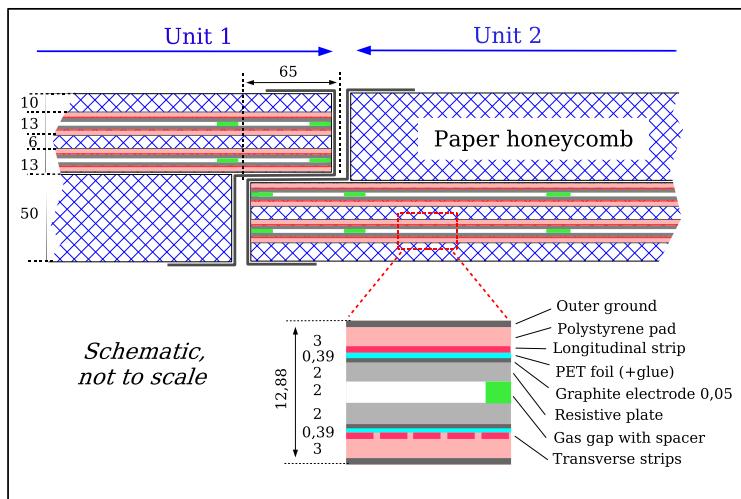


Figure 3.30: Cross section of an RPC, showing two units joined to form a single chamber. Each unit has two sets of two resistive plates (grey) separated by insulating spacers (green). The readout strips (magenta) are on the opposite side of the resistive plates from the gas gap. Outside the detecting elements, the volume of the RPC chamber is filled with paper honeycomb. The dimensions are given in millimeters.

Thin Gap Chambers

TGCs provide muon triggering in the end-caps, due to their good timing resolution and high rate capability. The TGCs also provide an azimuthal coordinate measurement, which complements the MDT measurement in the radial direction. Nine TGC disks are installed in total: a doublet mounted near the inner MDT end-cap layer, and a triplet and two doublets near the middle MDT end-cap layer. The disks are divided into two concentric annuli, one covering $1.05 \leq |\eta| \leq 1.92$ and the other covering $1.92 \leq |\eta| \leq 2.4$.

The TGCs are multiwire proportional chambers filled with a gas mixture of CO_2 and n -pentane. The cathode planes are 1.6 mm-thick FR4 (Flame Resistant 4) plates, with the

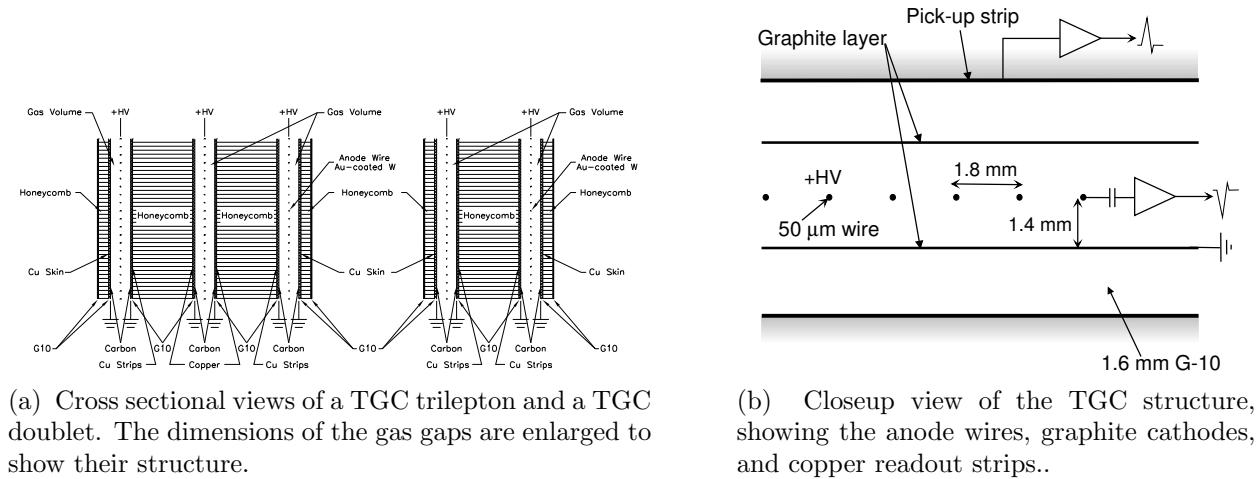


Figure 3.31: The construction of the TGCs.

interior faces coated with graphite and exterior faces cladded with copper. The triplet TGCs have three layers of wires and two layers of copper readout strips, while the doublet TGCs have two layers each of wires and strips, as shown in figure ???. The radial coordinate is measured by the anode wire groups and the azimuthal coordinate by the radial cathode strips. The chambers are “thin” in that the wire to cathode distance, 1.4 mm, is shorter than the wire-to-wire distance of 1.8 mm, as shown in figure ???. The anode wires have a diameter of $50 \mu\text{m}$ and are operated at a potential of 2.9 kV. The small wire-to-wire distance and high electric field near the anode wires contribute to a good timing resolution of 4 ns. The spatial resolution, 2–6 mm in R and 3–7 mm in ϕ , is determined by the ganging of readouts: due to the direction of the magnetic field, achieving the required momentum resolution at fixed transverse momentum requires finer resolution at larger pseudorapidities.

Trigger and Data Acquisition

During a typical data-taking run during 2012, the LHC beam contains 1377 bunches with a typical bunch spacing of $50 \mu\text{s}$, giving an event rate of $\sim 20 \text{ MHz}$. In contrast, the ATLAS data acquisition system records events at roughly 400 Hz. The filtering of events is performed by the three-level ATLAS triggering system. The first level, L1, is implemented in hardware, and reduces the event rate to less than 75 kHz. The second level, L2, is implemented in software, and reduces the event rate to less than 3.5 kHz using regions of interest (RoIs) identified by the L1 trigger. The final level of the trigger, the Event Filter, is also implemented in software, and reduce the event rate to less than 400 Hz using the full event information.

The L1 trigger reduces the event rate from 20 MHz to 75 kHz using a limited subset of the event data. While the L1 trigger is processing an event, the full event data is stored in

buffers on the detector. Due to the limited buffer size, the L1 latency must be less than $2.5\ \mu\text{s}$, of which $1\ \mu\text{s}$ is used by the cable propagation time. Therefore, the trigger is implemented in custom hardware processors. The block diagram for the trigger is shown in figure ???. The inputs to the trigger include the RPC and TGC detectors in the muon spectrometer and all of the calorimeter system. Two primary systems, the L1 Calorimeter Trigger and the L1 Muon Trigger, select events based on the presence of high- E_{T} objects or significant total event activity. These include high- p_{T} muons, electron and photons, jets, hadronically decaying tau leptons, large missing transverse energy, and large total transverse energy.

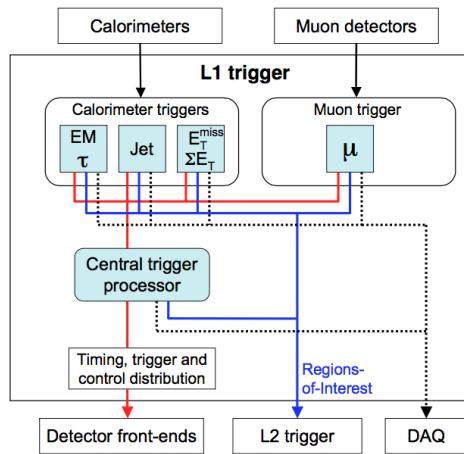


Figure 3.32: Block diagram of the L1 trigger system.

The L1 Calorimeter Trigger uses about 7,000 analogue trigger towers with a reduced granularity, $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for most of the detector. Two subsystems run in parallel: the cluster processor identifies electrons/photon and hadronically decay tau lepton candidates with E_{T} above a programmable set of thresholds, while the jet/energy-sum processor identifies jets and calculates the total scalar transverse energy and $E_{\text{T}}^{\text{miss}}$ using $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ blocks of calorimeter cells. Isolation cuts can be applied as well, limiting the energy allowed in the towers surrounding the object of interest and, in the case of electrons/photon, in the hadronic towers behind the electromagnetic towers.

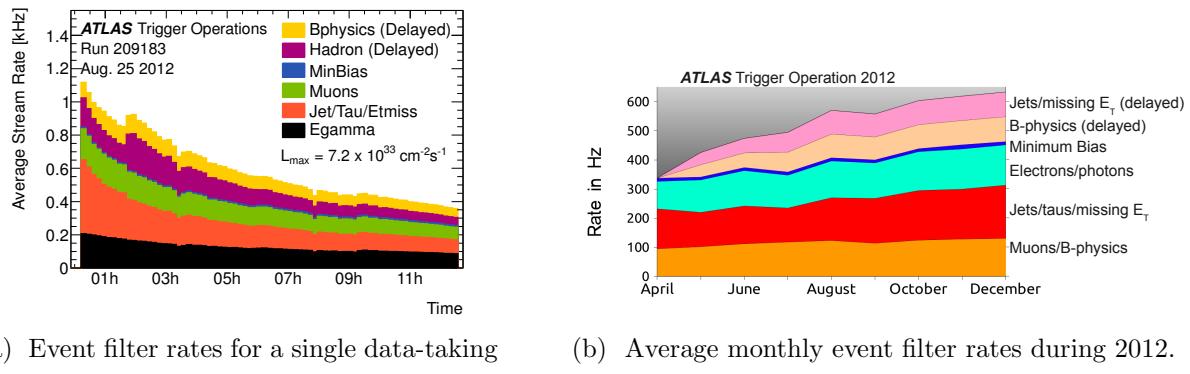
The L1 Muon Trigger searches for a coincidence of hits in consecutive muon trigger stations within a *road*, roughly corresponding to the path of a muon from the interaction point through the detector. The p_{T} threshold is encoded in the width of the road, with a narrower road corresponding to a higher p_{T} threshold. For low- p_{T} muons in the RPCs, the algorithm begins with hits in the second RPC doublet, called the *pivot plane*. The trigger requires a hit in the first RPC doublet, within the road defined by the interaction point and the hit in the pivot plane. The high- p_{T} algorithm is similar, requiring a hit along the road in the third RPC doublet as well as the first two. In the end-caps, the pivot plane is established by the outermost layer of TGCs, with the road corresponding to the path of an infinite-momentum muon originating from the interaction point. To reject backgrounds from

random coincidences, stricter requirements are imposed on the number of coincident hits in the TGC doublets and triplets.

The muon and calorimeter triggers pass their decisions along with the corresponding data to the Central Trigger Processor (CTP). The CTP communicates the total trigger decision to the front ends on the detector, including special triggers such as random triggers or the minimum bias trigger based on scintillator counters. In the event of a passed trigger, the event data is passed to the L2 trigger. The CTP also manages the *luminosity blocks*, an index representing the time at which an event was recorded with coarse (~ 1 minutes) granularity. A luminosity block is the shortest time interval for which the integrated luminosity can be determined such that the total uncertainty is dominated by systematic effects, rather than limited statistics. In the event of a detector failure, this index allows the rejection of a minimal set of affected events.

The L2 trigger reduces the event rate from a maximum of 75 kHz from the L1 trigger to less than 3.5 kHz. The trigger is implemented in software, and uses only the subset of the event data within the RoIs identified by the L1 trigger, typically 1%-2% of the total. After a successful L1 trigger, the data corresponding to the RoIs, stored in the detector-specific front end electronics, are accumulated in the ROI builder via 1574 readout links. The ROI builder combines the 1574 event fragments into a single data structure, which is passed to the L2 processing farm. At each step of a given trigger algorithm targeting particular signatures, only the relevant RoIs are analyzed. If no signatures remain valid, the event is rejected. Hence the full ROI information is transferred only for events which pass the L2 selection criteria. The typical L2 latency is about 40 ms.

Finally, the event filter reduces the event rate to the final \sim 400 Hz read out from the detector to disk. Also implemented in software, it runs the same algorithms as used in the offline event reconstruction using the full event information. Based on the objects used for triggering, the triggered events are categorized into one or more *data streams*. The data streams and other data computed by the event filter are appended to the event data, which are then transferred to CERN’s central data-recording facility for storage. An example of the event filter rates from the various streams during a single data-taking run is shown in figure ???. The monthly average rates of each stream during 2012 are shown in figure ??.



(a) Event filter rates for a single data-taking run.

(b) Average monthly event filter rates during 2012.

Table 3.5: The rates of recording events from the event filter for different data streams.

Chapter 4

Luminosity Measurement

4.1 Run I Luminosity

The ATLAS Run I data were recorded from 2011 to 2012, with a center-of-mass collision energy of $\sqrt{s} = 7\text{ TeV}$ in 2011 and $\sqrt{s} = 8\text{ TeV}$ during 2012. Integrated luminosities of $L = \int \mathcal{L} dt = 5.46\text{ fb}^{-1}$, and 22.8 fb^{-1} were delivered by the LHC in the two years, of which 5.08 fb^{-1} and 21.3 fb^{-1} were recorded by the ATLAS detector. The recorded luminosity accounts for the data acquisition inefficiency, as well as the *warm start* period, an interval of several minutes after the LHC has stable beams during which the tracking detectors are ramped to high voltage and the pixel detector preamplifiers are turned on. After masking data recorded while one or more detector subsystems were not functioning properly, 4.57 fb^{-1} and 20.3 fb^{-1} of data are considered ready for physics analyses. The volume of delivered, recorded, and physics-ready data are shown as a function of time in figure ??.

During Run I, the LHC was typically operated with 1368 bunches with a bunch spacing of 50 ns. The peak instantaneous luminosity reached values as high as $\mathcal{L} \sim 8 \times 10^{33}\text{ cm}^{-2}\text{ s}^{-1}$, corresponding to a peak average pileup value of $\mu \sim 37$. The distribution of pileup values in 2011 and 2012 data are shown in figure ??.

4.2 Measurement Overview

The instantaneous luminosity at a pp collider is given by [Grafstrom:2015go, Aad:2013ucp]:

$$\mathcal{L} = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}}, \quad (4.1)$$

where R_{inel} is the rate of inelastic pp collisions, and σ_{inel} is the pp inelastic cross section. For a storage ring with a revolution frequency f_r and n_b colliding bunch pairs, the instan-

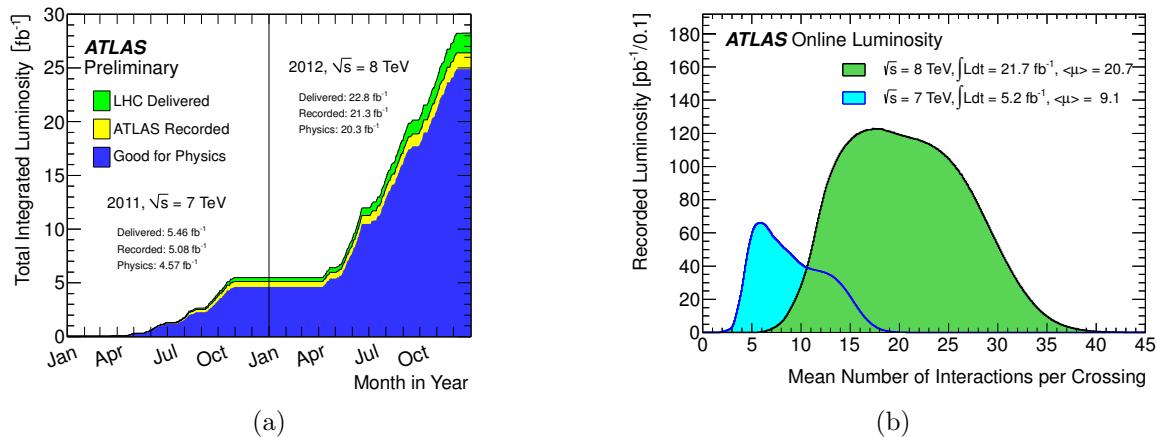


Figure 4.1: Left: The cumulative delivered, recorded, and physics-ready integrated luminosity versus time in 2011-2012. Right: Distribution of the number of interactions per bunch crossing in events recorded in 2011-2012.

taneous luminosity can be written in terms of the average number of inelastic pp collisions per bunch cross, μ , as

$$\mathcal{L} = \frac{\mu f_r n_b}{\sigma_{\text{inel}}}. \quad (4.2)$$

The instantaneous luminosity is measured by ATLAS using several detectors and algorithms, which have some efficiency ϵ to detect a pp interaction and measure the *visible* number of interactions per bunch crossing, $\mu_{\text{vis}} = \epsilon \mu$. Defining the visible cross section to be $\sigma_{\text{vis}} \equiv \epsilon \sigma_{\text{vis}}$, the instantaneous luminosity as measured by a particular detector is:

$$\mathcal{L} = \frac{\mu_{\text{vis}} f_r n_b}{\sigma_{\text{vis}}}. \quad (4.3)$$

The visible cross section is a calibration constant for a particular detector, which is determined during dedicated calibration runs in which the luminosity is determined directly from the geometry of the beams. The calibration procedure is described in section ??.

4.3 Luminosity Detectors

ATLAS performs many redundant luminosity measurements using several detectors. The detectors fall into two categories. *Event counting* detectors have a binary response, returning a 0 or 1 depending on whether a bunch crossing satisfies a set of criteria defined to detect an inelastic pp collision. Such detectors essentially measure $p(0; \mu)$, the probability that an event falls in the zero bin of a Poisson distribution, from which the mean μ be calculated. *Hit counting* detectors, on the other hand, count some quantity proportional to the number

of interactions in a given bunch crossing, such as the number of particles identified by a particular detector subsystem. A hit counting measurement typically yields more information about an event at the cost of additional systematic uncertainties.

Ideally, a luminosity detector exhibits the following features:

- The efficiency of the detector should be insensitive to pileup. The visible cross sections are typically measured at $\mu \approx \mathcal{O}(1)$, while a typical data-taking run has $10 \lesssim \mu \lesssim 50$; it is essential that the visible cross section be constant across this range of pileup values.
- The efficiency of the detector should also be constant over long timescales.
- The response of the detector and the readout should fast enough to provide a bunch-by-bunch luminosity measurement. As the LHC bunches are not identical, with different numbers of protons and emittances, it is useful to measure luminosity for each colliding bunch pair, as often as once every 25 ns. The colliding bunch pairs are labeled by the bunch crossing identification number, or BCID; consecutive BCIDs are separated by 25 ns.
- The efficiency should be high enough to yield sufficient statistics. The data are used in increments as short as 20 s, so the statistics collected over this time scale should be high enough that the total uncertainty is dominated by systematic effects. On the other hand, for event counting detectors, the efficiency should not be so high that the detector is saturated; the uncertainty on μ is large if $p(0; \mu)$ is too close to 0 or 1.
- The backgrounds should be low and understandable. Detectors can be sensitive to a wide range of phenomena aside from pp collisions, which should not be counted as luminosity. For example, some detectors observe a phenomenon called *afterglow*, a small amount of activity in the BCIDs immediately following a collision likely due to photons from nuclear deexcitations in the detector material. The afterglow background is proportional to the luminosity in the colliding BCIDs, and decays away with several time constants. Collisions between a beam and residual gas in the beam pipe, called *beam-gas interactions*, can also contribute a low level of background, and is estimated by observing non-colliding bunches passing through the interaction region.

The central value of the luminosity measurement is determined by the *beam conditions monitor* (BCM) [Cindro:2008zz], which fulfills most of these desired criteria. The BCM consists of eight diamond-based particle detectors, four on each side of the interaction point at $|z| = 184$ cm and $|\eta| = 4.2$. The detectors have a physical cross section of approximately 1 cm^2 and are arranged in a cross pattern, with two readouts corresponding to the vertical and horizontal pairs. The design purpose of the BCM is to monitor backgrounds and to trigger a beam dump if beam losses towards the inner detector become too high; accordingly, the detector has a very fast readout, and can provide a bunch-by-bunch luminosity measurement with a time resolution of ~ 0.7 ns. The luminosity is measured using event counting, and

can be measured using any combination of the readouts. The configurations used require hits in either the vertical pair (BCMV) or the horizontal pair (BCM_H), and either coincident hits on both sides of the interaction point (AND) or a single hit on either side (OR). The small size of the readouts leads to an efficiency of approximately 7% in the OR configuration, which allows for sufficient statistics without saturating the detector.

LUCID (LUminosity measurement using Cherenkov Integrating Detector) [TheATLASCollaboration] provides a supplementary, bunch-by-bunch, event counting-based luminosity measurement. The detector consists of two sets of sixteen Cherenkov detectors surrounding the beampipe at $z = \pm 17$ m, occupying the pseudorapidity range $5.6 < |\eta| < 6.0$. The Cherenkov detectors are polished aluminum tubes filled with C₄F₁₀ gas. The Cherenkov photons induced by charged particles traversing the gas are collected by photomultiplier tubes (PMTs) located at the far end of the tubes. Additional Cherenkov photons are produced in the quartz window separating the tube volume from the PMT. In this configuration, the typical single-particle yield is 60-70 photoelectrons due to photons created in the gas, and about 40 photoelectrons due to the quartz window. A hit is recorded if the PMT signal exceeds a preset threshold, corresponding to about 15 photoelectrons. However, the higher instantaneous luminosities due to the 50 ns bunch spacing led to saturation of the detectors, and hence on 30 July 2011, the gas was removed from the Cherenkov tubes to reduce the efficiency. The removal of the gas also improved the detector's stability and linearity with respect to pileup. Relative comparisons with other detectors were used to reestablish the detector calibration.

Several detector subsystems nominally designed for physics object reconstruction are also used for hit counting-based luminosity measurements. Algorithms using the tile calorimeter and the forward calorimeters (see section ??) determine the luminosity from detector currents proportional to the total particle flux in small regions of the calorimeters. The tile calorimeter algorithm monitors the PMT currents corresponding to a few selected cells near $|\eta| = 1.25$, where the highest sensitivity to changes in the luminosity is observed. Similarly, the forward calorimeter algorithm monitors the currents in the high voltage lines. In both cases, the detector response is too slow to provide a bunch-by-bunch measurement, and the current response is not sensitive to the low instantaneous luminosities during the dedicated calibration runs, requiring the calibration to be set using a relative comparison with LUCID. On the other hand, the detectors exhibit good linearity of response with pileup, and good short-term stability.

Finally, algorithms using the inner detector measure the luminosity by counting reconstructed tracks and vertices¹. The use of these higher-level objects confers certain benefits, such as low background rates and good long-term stability, at the cost of increased computational requirements, which limit the data rate to $\mathcal{O}(100\text{ Hz})$. Further, pileup effects can be significant; in particular, the efficiency of the vertex counting algorithm varies by as much as 30% up to pileup values of $\mu = 30$, limiting its utility to data with low pileup.

¹Pixel cluster counting, used by the CMS experiment, was also considered, but was not commissioned due to the difficulty of the background subtraction.

4.4 Luminosity Calibration: van der Meer Scans

The visible cross sections for the detectors and algorithms described in section ?? are calibrated during dedicated runs called van der Meer (vdM) scans [vanderMeer:1968ud]. The calibration procedure uses the definition of luminosity in terms of the beam parameters, given for a single colliding bunch by:

$$\mathcal{L} = f_r n_1 n_2 K \int \rho_1(x, y, z, t) \rho_2(x, y, z, t) dx dy dz dt, \quad (4.4)$$

where f_r is the revolution frequency, $n_{1,2}$ are the number of particles in the colliding bunches, and $\rho_{1,2}(x, y, z, t)$ are the time- and position-dependent particle density distribution, normalized so that $\int \rho_{1,2}(x, y, z, t) dx dy dz = 1$. K is a kinematic factor,

$$K = \sqrt{(\vec{v}_1 - \vec{v}_2)^2 - \frac{(\vec{v}_1 \times \vec{v}_2)^2}{c^2}}, \quad (4.5)$$

which, in the limit $|\vec{v}_{1,2}| \rightarrow c$, reduces to $2c \cos \alpha$, where α is the crossing angle between the beams. To simplify the current discussion, the crossing angle is assumed to be zero, and the bunch densities are assumed to be functions of x , y , and $z \pm ct$, i.e. that the transverse bunch profiles are constant over the duration of the collision². The luminosity can then be expressed as:

$$\mathcal{L} = f_r n_1 n_2 \int \hat{\rho}_1(x, y) \hat{\rho}_2(x, y) dx dy, \quad (4.6)$$

where $\hat{\rho}_{1,2}(x, y)$ are the transverse particle densities, normalized to unity. Under the further assumption that the transverse particle densities factorize in the horizontal and vertical directions, $\hat{\rho}(x, y) = \hat{\rho}_x(x) \hat{\rho}_y(y)$, where $\hat{\rho}_x(x)$ and $\hat{\rho}_y(y)$ are also normalized to unity, the luminosity can be written as:

$$\mathcal{L} = f_r n_1 n_2 \Omega_x(\hat{\rho}_{x1}, \hat{\rho}_{x2}) \Omega_y(\hat{\rho}_{y1}, \hat{\rho}_{y2}), \quad (4.7)$$

where $\Omega_x(\hat{\rho}_{x1}, \hat{\rho}_{x2}) = \int \hat{\rho}_{x1}(x) \hat{\rho}_{x2}(x) dx$, and similarly for the y direction. As first proposed by van der Meer [vdm], the $\Omega_{x,y}$ parameters can be determined by measuring the interaction rate as a function of transverse beam displacement, $R_{x,y}(\delta)$. Without loss of generality, assume that beam 2 is displaced by δ in the x direction, while beam 1 is held fixed. Then, $R_x(\delta) = k \int \rho_{x1}(x) \rho_{x2}(x - \delta) dx$ for some constant k , and

²Specifically, the *hourglass effect* is neglected. The collisions occur in a drift space, where the beams are typically focused, or squeezed, onto the interaction point. The transverse size of the beam in direction i as a function of z is given by $\sigma_i^2(z) = \epsilon_i \beta_i^* \left(1 + \frac{(z - z_i^w)^2}{\beta_i^{*2}}\right)$, where ϵ_i is the transverse emittance, z_i^w is location of the optical waist, and β_i^* is the betatron function at $z = z_i^w$. The effect is significant when $\beta_i^* \lesssim \sigma_z$; during the vdM scans, $\sigma_z \approx 50$ mm, while $\beta^* = 1.5$ m-11 m, and hence the hourglass effect is neglected.

$$\frac{R_x(0)}{\int R_x(\delta) d\delta} = \frac{k \int \rho_{x1}(x) \rho_{x2}(x) dx}{k \int \int \rho_{x1}(x) \rho_{x2}(x - \delta) dx d\delta} \quad (4.8)$$

$$= \int \rho_{x1}(x) \rho_{x2}(x) dx \quad (4.9)$$

$$= \Omega_x(\hat{\rho}_{x1}, \hat{\rho}_{x2}) \quad (4.10)$$

For convenience, define $\Sigma_{x,y}$ to be the characteristic widths of $R_{x,y}(\delta)$, given by:

$$\Sigma_{x,y} = \frac{1}{\sqrt{2\pi}} \frac{\int R_{x,y}(\delta) d\delta}{R_{x,y}(0)}. \quad (4.11)$$

For Gaussian beams, $\Sigma_{x,y}$ correspond to the Gaussian width of $R_{x,y}(\delta)$. Finally, the luminosity is given by

$$\mathcal{L} = \frac{f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y}. \quad (4.12)$$

Equating this with the luminosity defined in equation ??, the visible cross section for a given detector and algorithm is given by

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\text{MAX}} \frac{2\pi \Sigma_x \Sigma_y}{n_1 n_2}, \quad (4.13)$$

where $\mu_{\text{vis}}^{\text{MAX}}$ is the visible interaction rate per bunch crossing at the maximum of the scan curve, $R(\delta)$. The numbers of particles per bunch, $n_{1,2}$, are measured by the LHC Bunch Current Normalization Working Group using bunch current transformers (BCTs) [BCNWG].

2011 Luminosity Calibration

As an example, the 2011 pp calibration is derived from two pairs of scans in the x - and y -directions, performed during the same LHC fill on 15 May 2011 [Aad:2013ucp]. The beams had 14 colliding bunch pairs, $\sim 0.8 \times 10^{11}$ protons per bunch, $\beta^* = 1.5$ m, and a crossing angle of $\alpha = 240 \mu\text{rad}$. The resulting transverse beam size was approximately $\sigma_x \approx \sigma_y \approx 40 \mu\text{m}$, and the peak average number of interactions per cross with head-on collisions was $\mu \approx 2.3$. The scan was performed in 25 equal steps over a displacement range of $\delta = \pm 233 \mu\text{m}$.

An example scan curve showing $\mu_{\text{vis}}^{\text{sp}} \equiv \mu_{\text{vis}}/(n_1 n_2)$ as a function of transverse beam separation for the BCMV_OR algorithm is shown in figure ???. The vdm scan curve is fitted with a Gaussian plus a constant, which is used as $R_{x,y}(\delta)$ to calculate $\Sigma_{x,y}$ in equation ???. $\mu_{\text{vis}}^{\text{MAX}}$ is determined from the peak of the fitted function. The measured σ_{vis} values for both scans and all 14 colliding bunch pairs are shown in figure ???. The luminosity-weighted mean σ_{vis} is taken as the central value, while the scatter of the 28 measurements, which is not consistent

with statistical variation, is taken as a systematic uncertainty on the reproducibility of the measurement.

The visible cross sections for several algorithms using during 2011 are shown in table ??, along with the efficiency assuming a total inelastic cross section of $\sigma_{\text{inel}} = (71.34 \pm 0.90) \text{ mb}$ [TheATLASCo].

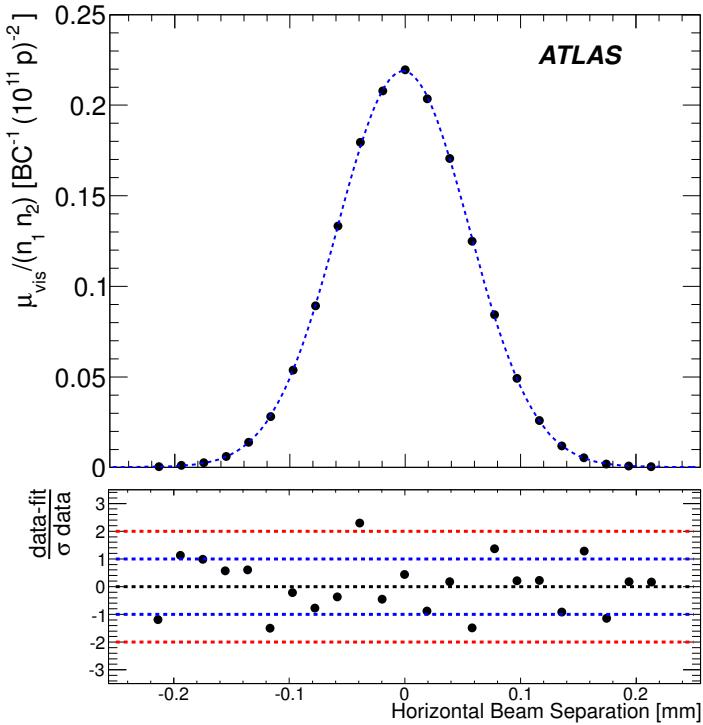


Figure 4.2: An example of a scan curve from the van der Meer scans on 15 May 2011, showing the specific visible interaction rate $\mu_{\text{vis}}^{\text{sp}} \equiv \mu_{\text{vis}}/(n_1 n_2)$ versus transverse beam separation in the x direction for a single scan and BCID. The data are fitted with Gaussian plus constant. The bottom plot shows the residual deviation of the data from the fit, divided by the uncertainty on the data.

4.5 Systematic Uncertainties

The systematic uncertainty on the 2011 and 2012 luminosity measurements are 1.8% and 2.8%³, respectively. No single source of uncertainty dominates the total; rather, the uncertainty is due to many sources, each contributing less than 1%. The sources of uncertainty on

³Preliminary uncertainty. The final uncertainty is expected to be below 2%.

Algorithm	σ_{vis} (2011)	$\frac{\sigma_{\text{vis}}}{\sigma_{\text{inel}}}$
BCM_VOR	4.82 ± 0.07	0.068
BCM_HOR	4.78 ± 0.07	0.067
BCM_VAND	0.142 ± 0.002	0.002
BCM_HAND	0.140 ± 0.002	0.002
LUCID_OR	43.3 ± 0.7	0.607
LUCID_AND	13.7 ± 0.2	0.192

Table 4.1: Visible cross sections for pp luminosity measurement algorithms used in 2011. The efficiency of detecting a pp collision is also shown for $\sigma_{\text{inel}} = 71.34 \text{ mb}$.

the 2011 luminosity measurement are shown in table ??, divided into uncertainties on the visible cross section measurement during vDM scans and uncertainties on the measurement performed over the course of data taking.

The combined uncertainty on the visible cross sections due to the vDM calibration procedure is 1.5%. The largest source of uncertainty is due to emittance growth and non-reproducibility, reflected in the scatter between BCIDs and between scans in figure ?? . Other significant sources, each roughly 0.5%, include beam-beam effects, where the two colliding beams deflect each other away from the nominal transverse separation; transverse correlations which violate the assumption that the transverse particle densities factorize as $\rho(x, y) = \rho_x(x)\rho_y(y)$; pileup dependence; and the measurement of the number of protons per bunch, $n_{1,2}$.

Uncertainties on the luminosity measurement over the 2011 data-taking period total 0.9%. These uncertainties are dominated by variations in the detector calibrations during 2011, quantified using relative comparisons between algorithms across the entire year, and pileup dependence, quantified using relative comparisons between algorithms at different pileup values. The relative comparisons are shown in figure ??.

4.6 Vertex-Based Luminosity Measurement

Primary vertices are points consistent with being the origin of a set of tracks reconstructed by the inner detector, nominally due to inelastic pp interactions. The reconstruction of tracks and vertices is described in section ?? . Vertex counting [PaganGriso:2013wn] is an appealing luminosity measurement technique for a number of reasons. The backgrounds are very low, and can be controlled by requiring a minimum number of tracks per vertex, chosen here to be five tracks with $p_T > 400 \text{ MeV}$. Further, the vertex reconstruction efficiency is expected to be stable throughout the data taking period. However, the technique has two significant drawbacks. First, the data is collected through the standard ATLAS data acquisition system, limiting the event rate during normal physics runs to $\mathcal{O}(100 \text{ Hz})$. Depending on the trigger used, a correction for the trigger deadtime may also be necessary.

Source	Uncertainty
Bunch population product ($n_1 n_2$)	0.5%
Beam centering	0.10%
Beam position jitter	0.30%
Emittance growth/non-reproducibility	0.67% ⊕ 0.55%
Bunch-to-bunch σ_{vis} consistency	0.28%
Fit model	0.31%
Background subtraction	0.29%
Length scale calibration	0.30%
Absolute ID length scale	0.30%
Beam-beam effects	0.50%
Transverse correlations	0.50%
Pileup dependence	0.50%
Afterglow correction	0.2%
BCM Stability	0.2%
Long-term consistency	0.7%
Pileup dependence	0.5%
Total	1.8%

vdM Calibration subtotal=1.5%
Uncertainties from a single
measurement of σ_{vis}

\mathcal{L} measurement subtotal=0.9%
Uncertainties evaluated from
all 2011 physics runs

Table 4.2: Systematic uncertainties on the integrated luminosity measured in 2011 using the BCM_VOR algorithm. The uncertainties are separated into uncertainties on the visible cross section (1.5%) and measurement uncertainties over the year (0.9%).

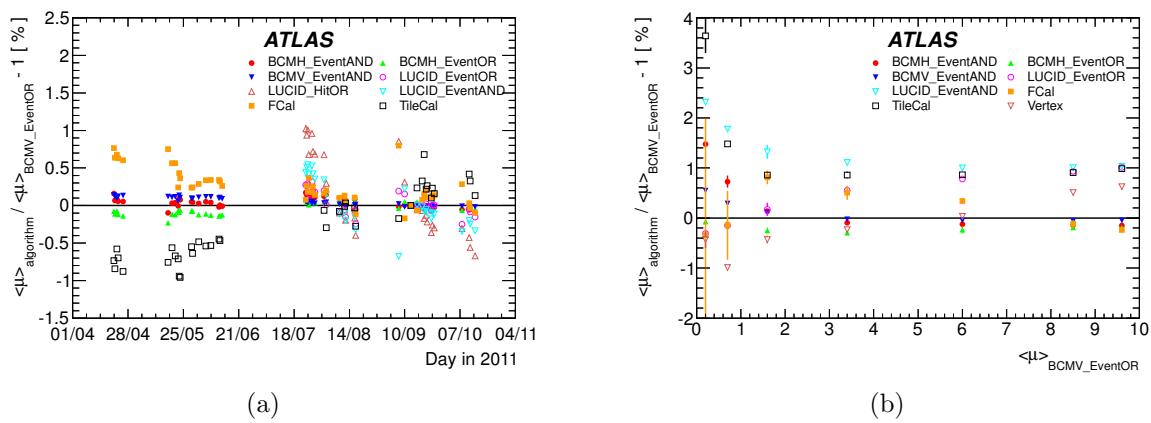


Figure 4.3: Left: Comparison of luminosities measured by different algorithms as a function of time in 2011. Right: Comparison of luminosities measured by different algorithms as a function of the average number of interactions per bunch cross, μ , as measured by BCM_VOR. The data were taken during a single run by separating the beams in the transverse direction, similar to a van der Meer scan.

Second, the efficiency of the vertex reconstruction algorithm is significantly nonlinear with pileup, with the efficiency decreasing by $\sim 30\%$ between pileup values of $\mu = 1$ and $\mu = 30$.

Vertex Counting Method

The data for the inner detector-based luminosity measurements are collected through the standard ATLAS data acquisition system. Events are collected with either a random trigger or a trigger requiring hits in the minimum bias trigger scintillators (MBTS), two discs of scintillators mounted on the inner surfaces of the LAr end-cap cryostats ($2.12 < |\eta| < 3.85$). The random trigger records a fraction of all bunch crossings in a specific set of BCIDs, enabling an unbiased measurement of μ_{vis} at the cost of lower statistics. The statistics collected worsen at low pileup values, where a large fraction of the triggered events do not contain a collision; hence, the random trigger is used mostly at high pileup. The MBTS trigger requires hits MBTS, usually at least two. The trigger inefficiency, prescale, and deadtime must be taken into account. For vertices with at least five tracks, the inefficiency is negligible. Both triggers can select a specific set of BCIDs, which is essential for special runs where a bunch-by-bunch measurement is necessary, e.g. the van der Meer scans (see section ??).

Vertices are reconstructed with two settings: the default reconstruction settings based on tracks with $p_T > 400 \text{ MeV}$, and tighter settings (VtxLumi) based on tracks with $p_T > 900 \text{ MeV}$ with stricter requirements on the track quality. The tighter settings were introduced in 2012 to reduce the computational requirements associated with collecting $\mathcal{O}(100 \text{ Hz})$ of inner detector data over the entire year. The studies described below are based on the 2011 data, where the vertex-based luminosity measurement is only performed on special runs, and use the default reconstruction settings.

The single-bunch instantaneous luminosity is calculated as follows. For the random trigger, the visible interaction rate is $\mu_{\text{vis}} = \frac{N_{\text{vtx}}}{N_{\text{evt}}}$, where N_{vtx} is the number of vertices recorded in N_{evt} triggered events. For the MBTS trigger, $\mu_{\text{vis}} = \frac{N_{\text{vtx}}}{f_r t}$, where N_{vtx} is the number of vertices recorded after correcting for the prescale and deadtime, $f_r = 11\,245.5 \text{ Hz}$ is the LHC revolution frequency, and t is the duration of the measurement. μ_{vis} is then corrected for pileup effects, described below in section ??, and finally the luminosity is given by equation ??.

Pileup Effects

The vertex reconstruction efficiency is strongly affected by three pileup-related phenomena:

- Vertex masking: a pp interaction fails to be reconstructed as a vertex because some or all of its tracks are used by an earlier vertex in the iterative reconstruction algorithm. This is the dominant pileup effect, which causes a large drop in the reconstruction efficiency at high pileup.

- Fake vertices: a vertex passes the cut on the minimum number of tracks due to acquiring tracks from another nearby pp interaction. This is a subdominant but significant effect, which causes a small increase in the reconstruction efficiency at high pileup.
- Split vertices: a single interaction is reconstructed as two separate vertices. This is a significant effect when considering vertices with two or more tracks, but is negligible for vertices with at least five tracks.

Corrections are derived for vertex masking and fake vertices. The fake vertex correction is derived using truth matching in minimum bias Monte Carlo simulation. The simulation sample was generated using PYTHIA 8, with tune A2M [[pythia](#)]. For a given cut on the minimum number of tracks per vertex, m , a reconstructed vertex is labelled as fake if less than m of its tracks are matched to charged particles originating from the same generated pp interaction. The average number of reconstructed vertices labelled as fake by Monte Carlo truth matching is shown as a function of pileup in figure ???. The fake fractions show a significant dependence on μ and on the minimum number of tracks per vertex.

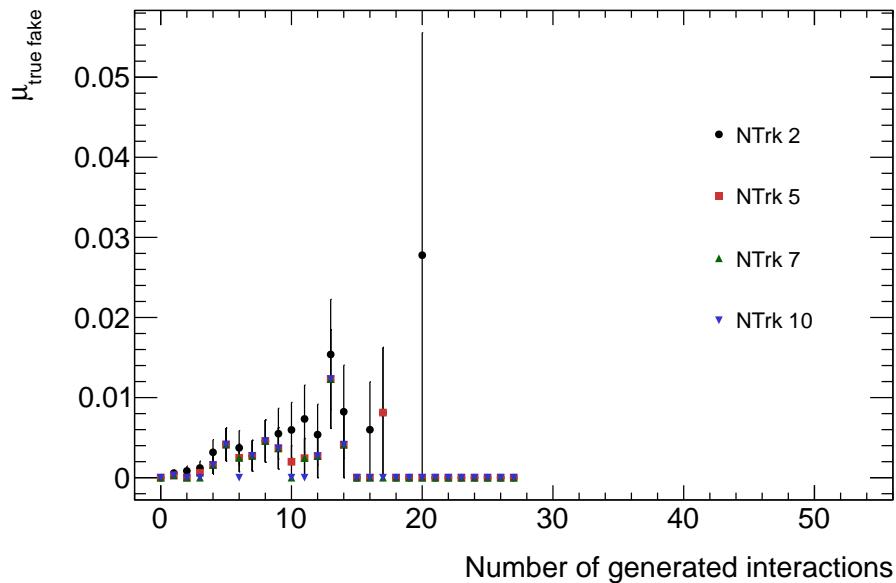


Figure 4.4: Average number of reconstructed vertices without a matched truth interaction, as a function of number of generated interactions per event. The contribution is below 0.1%.

A correction is derived for vertex masking using the distribution of longitudinal distances, Δz , between pairs of vertices in the same event. In the absence of masking, if the interaction region has a Gaussian longitudinal profile with width σ_z , then the Δz distribution would be a Gaussian with width $\sigma_z \times \sqrt{2}$. Masking manifests as an absence of reconstructed pairs near $\Delta z = 0$, as shown in figure ??.

The correction is derived in a data-driven way as follows.

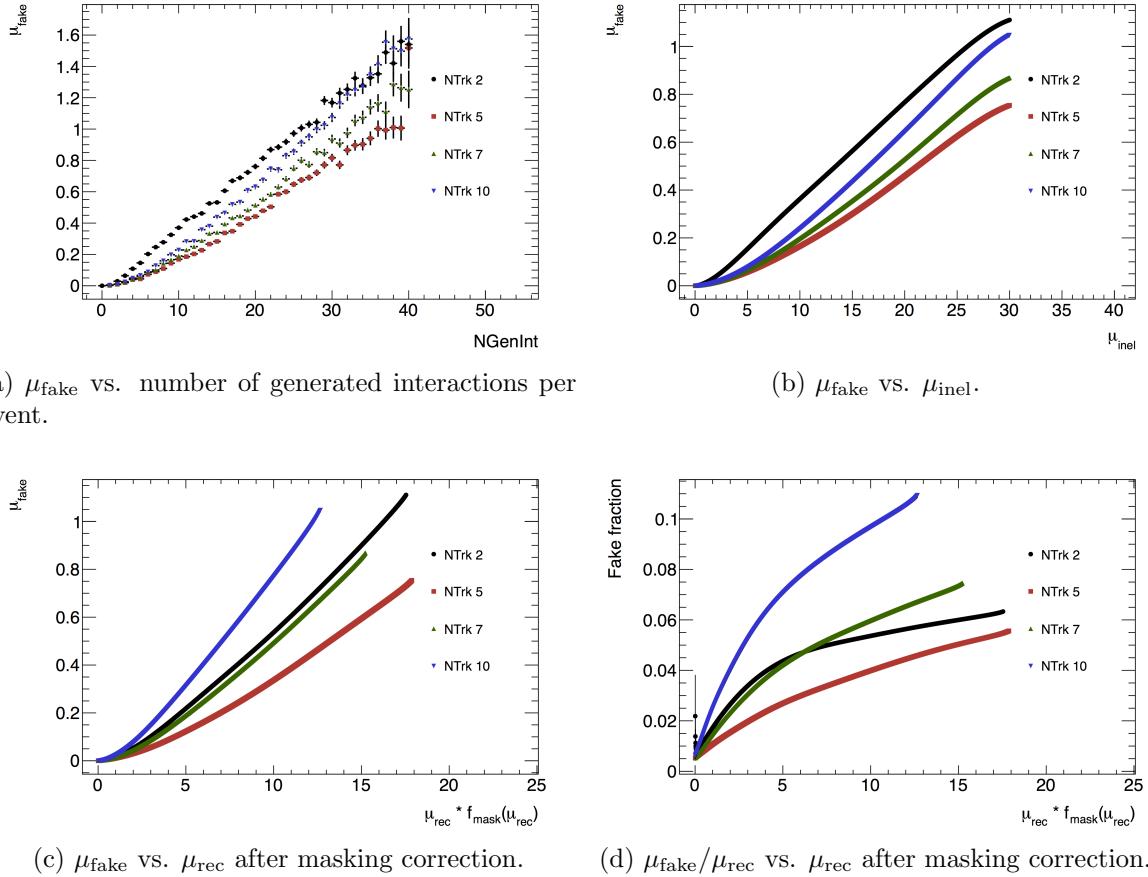


Figure 4.5: Fraction of reconstructed vertices labelled as fake by Monte Carlo truth matching, as a function of pileup.

1. In a data sample with exactly two interactions per event, the number of masked pairs is equal to the number of masked vertices. Using data taken at low pileup values and selecting events with exactly two reconstructed vertices, we calculate the 2-vertex masking probability $p_{\text{mask}}(\Delta z)$, i.e. the probability that only one of two vertices separated by Δz is reconstructed. This function is assumed to be a universal property of the vertexing algorithm, independent of μ .

Specifically, the expected Δz distribution in the absence of masking, $f_{\text{exp}}(\Delta z)$, is derived by randomly sampling pairs of points from the observed z -distribution of reconstructed vertices. Using $f_{\text{exp}}(\Delta z)$ as a template, the *observed* Δz distribution, $f_{\text{obs}}(\Delta z)$, is fitted in the range $30 \text{ mm} \leq |\Delta z| \leq 300 \text{ mm}$, where vertex masking is negligible. Finally, the masking probability is defined as:

$$p_{\text{mask}}(\Delta z) = \frac{f_{\text{exp}}(\Delta z) - f_{\text{obs}}(\Delta z)}{f_{\text{exp}}(\Delta z)} \quad (4.14)$$

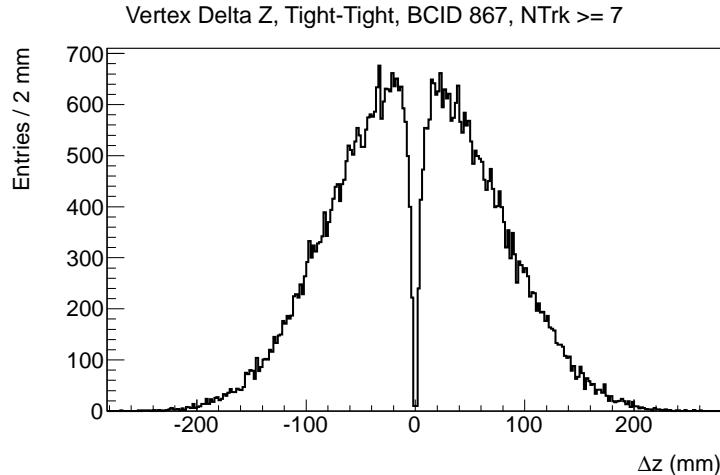


Figure 4.6: Δz distribution between pairs of vertices in the same event, with data from the May Van der Meer scan (ATLAS run 182013).

The masking probability functions derived from minimum bias Monte Carlo and for low- μ data are shown in figures ?? and ??.

2. To derive a correction for a particular data sample at arbitrary μ , an *expected* Δz distribution, $f_{\text{exp}}(\Delta z)$, is again derived by randomly sample pairs of vertices from the observed primary vertex z -distribution. Then, the total probability p_{mask} that given any two tight vertices, only one is reconstructed can be computed:

$$p_{\text{mask}} = \int_{-\infty}^{\infty} p_{\text{mask}}(\Delta z) f_{\text{exp}}(\Delta z) d(\Delta z). \quad (4.15)$$

3. The total masking probability p_{mask} is used to generate a map between the number of reconstructible vertices per event, N_{vis} , and the average number of reconstructed vertices per event, μ_{rec} , as follows. Label the generated vertices v_i , $1 \leq i \leq N_{\text{gen}}$, in the order in which the iterative vertexing algorithm reconstructs the vertices; similarly, let p_i , $1 \leq i \leq N_{\text{vis}}$, be the probability that vertex v_i is reconstructed. Proceeding vertex-by-vertex, the p_i follow a recursion relation:

$$p_1 = 1 \quad (4.16)$$

$$p_2 = p_1 \times (1 - p_{\text{mask}}) \quad (4.17)$$

$$\dots \quad (4.18)$$

$$p_k = \prod_{i=1}^{k-1} (p_i \times (1 - p_{\text{mask}}) + (1 - p_i) \times 1) \quad (4.19)$$

$$= \prod_{i=1}^{k-1} (1 - p_i p_{\text{mask}}) \quad (4.20)$$

$$= p_{n-1} \times (1 - p_{n-1} p_{\text{mask}}) \quad (4.21)$$

The average number of reconstructed vertices is then:

$$\langle N_{\text{rec}} \rangle = \sum_{i=1}^{N_{\text{vis}}} p_i. \quad (4.22)$$

4. Finally, a map is computed between the average number of reconstructible vertices, μ_{vis} , and the average number of reconstructed vertices, $\mu_{\text{rec}}(\mu_{\text{vis}})$, by convoluting $\mu_{\text{rec}}(N_{\text{vis}})$ with a Poisson distribution:

$$\langle \mu_{\text{rec}} \rangle(\mu_{\text{vis}}) = \sum_{N_{\text{vis}}=0}^{\infty} P(N_{\text{vis}}; \mu_{\text{vis}}) \mu_{\text{rec}}(N_{\text{vis}}). \quad (4.23)$$

Vertex-Based Luminosity Measurements

Due to the low event rate available during physics runs, vertex counting was used to measure luminosity only in three special runs during 2011, where inner detector data was recorded by a special high-rate data stream. The data stream reads out events at $\mathcal{O}(10 \text{ kHz})$, recording only the inner detector from a small number of bunch crossings, typically less than four. The three runs are the vdM calibration run in May 2011, the pileup scan in September 2011 shown in figure ??, and a high- β^* run with a single bunch used to measure the total pp cross section using the ALFA detector.

Van der Meer Scan

An example scan curve from the May 2011 vdM scan is shown in figure ???. The trigger used to collect the data required hits in the minimum bias trigger scintillators (MBTS), and selected 3 of the 14 colliding bunch pairs. The peak interaction rate during the scan

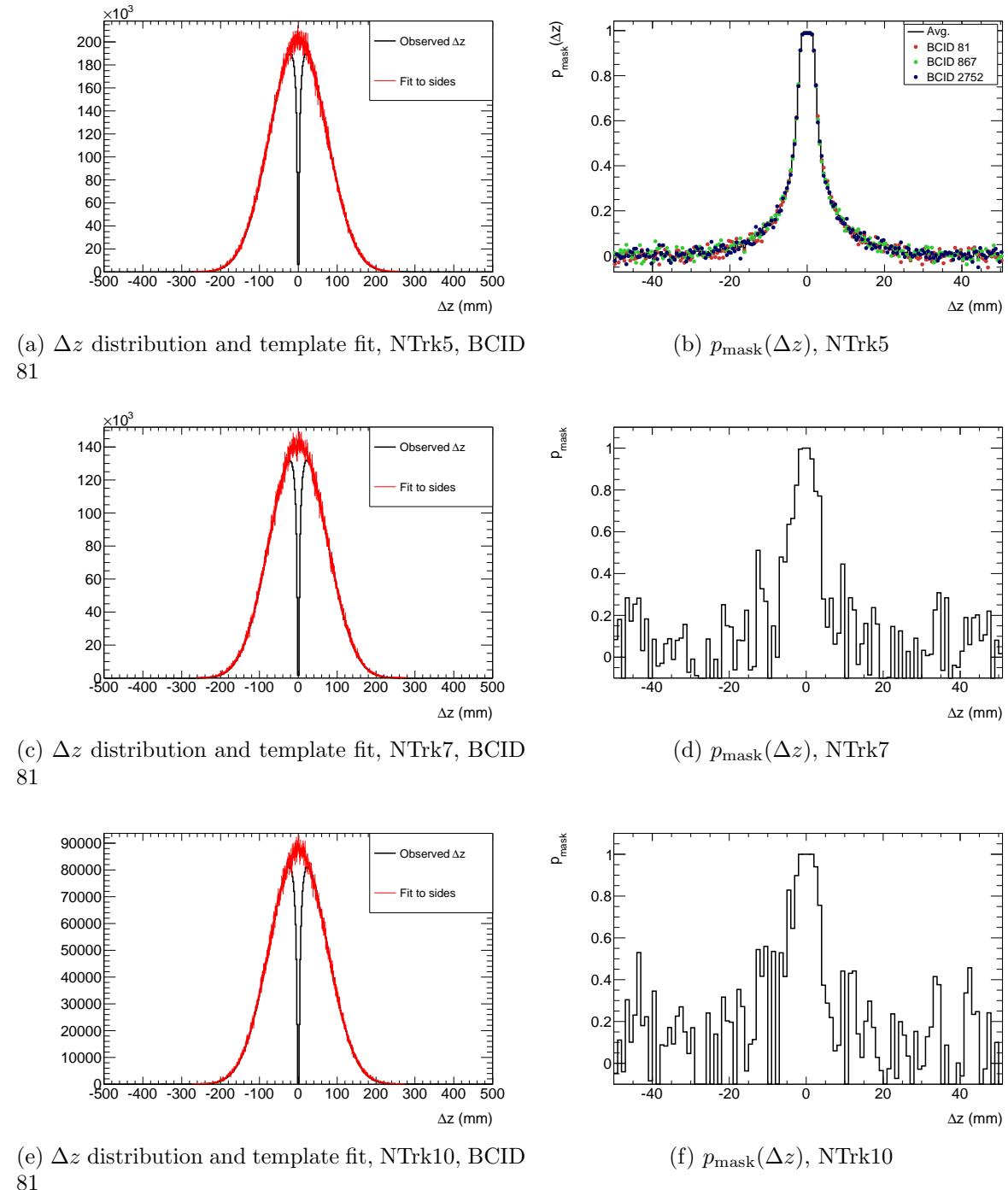


Figure 4.7: Calibration of the masking correction method on data, using run 182013. Left: Δz distributions and template fits using expected Δz distributions in the range $30 \text{ mm} \leq \Delta z \leq 300 \text{ mm}$. Right: pairwise vertex masking probability as a function of the longitudinal distance Δz between the vertices.

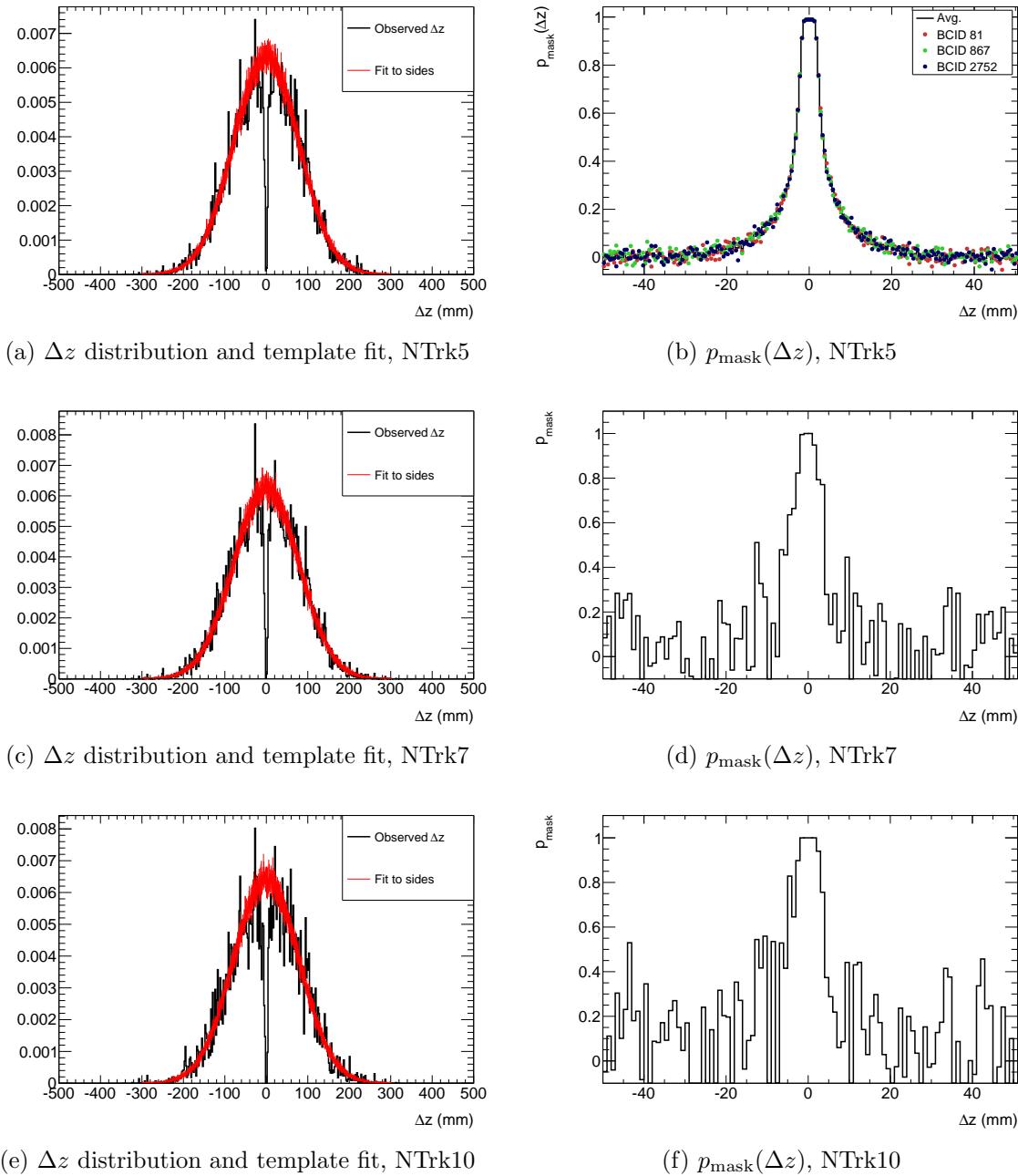


Figure 4.8: Calibration of the masking correction method on simulation, using 8 TeV minimum bias Monte Carlo (pythia 8, tune A2M). Left: Δz distributions and template fits using expected Δz distributions in the range $30 \text{ mm} \leq \Delta z \leq 300 \text{ mm}$. Right: pairwise vertex masking probability as a function of the longitudinal distance Δz between the vertices.

was $\mu \sim 1.97$. The data are corrected for vertex masking and fakes, with the correction factor reaching up to 3% at the peak of the scan curves. Following the protocol described in section ??, the visible cross section is determined to be (38.50 ± 0.12) mb, where the uncertainty reflects the RMS spread between the three bunch crossings and the two scans.

Pileup Scan

The September 2011 pileup scan is used to derive systematic uncertainties due to the non-linear response of algorithms with respect to the number of interactions per bunch crossing. The scan was performed at the end of a physics run, displacing the beams in the transverse direction to obtain a sample of data at over a pileup range of $0.02 \lesssim \mu \lesssim 10$ with otherwise identical conditions. The data was triggered by a random trigger selecting events from two bunch crossings, BCIDs 200 and 999. The data are corrected for vertex masking and fakes, with the correction reaching up to 10% as shown in figure ?? with respect to BCM_VOR. A comparison of the luminosity measurements between vertex counting and BCM_VOR, shown in figure ??, exhibits a slope of about 0.1% per unit of μ .

ALFA Run

Finally, vertex counting provides a reliable luminosity measurement for the 2011 ALFA run, a special run used for a measurement of the total pp cross section [TheATLASCollaboration:2014joa]. The beams contained a single colliding bunch pair with $\beta^* = 90$ m and $\mu \sim 0.03$, in order to measure the scattering angle of elastic pp collisions. A special luminosity analysis is required to address the very low instantaneous luminosity of $\mathcal{L} \sim 5 \times 10^{27}$ cm $^{-2}$ s $^{-1}$, about six orders of magnitude lower than a typical physics fill. The calorimeter methods are unusable due to a lack of sensitivity. For BCM and LUCID, the backgrounds at low instantaneous luminosity have a different composition: afterglow is negligible with a single colliding bunch, but beam-gas interactions can be significant, resulting in an extra 0.2% systematic uncertainty. The low-pileup conditions are ideal for vertex counting, eliminating the need to perform pileup corrections. The requirement of at least five tracks with $p_T > 400$ MeV per vertex suppresses the beam-gas backgrounds. The data were recorded by a random trigger at approximately 1 kHz.

A comparison of luminosity measurements from BCM, LUCID, and vertex counting is shown in figure ???. To be consistent with the primary pp luminosity measurement, the central value is taken from BCM_VOR; vertex counting shows agreement with this value to within 0.5%.

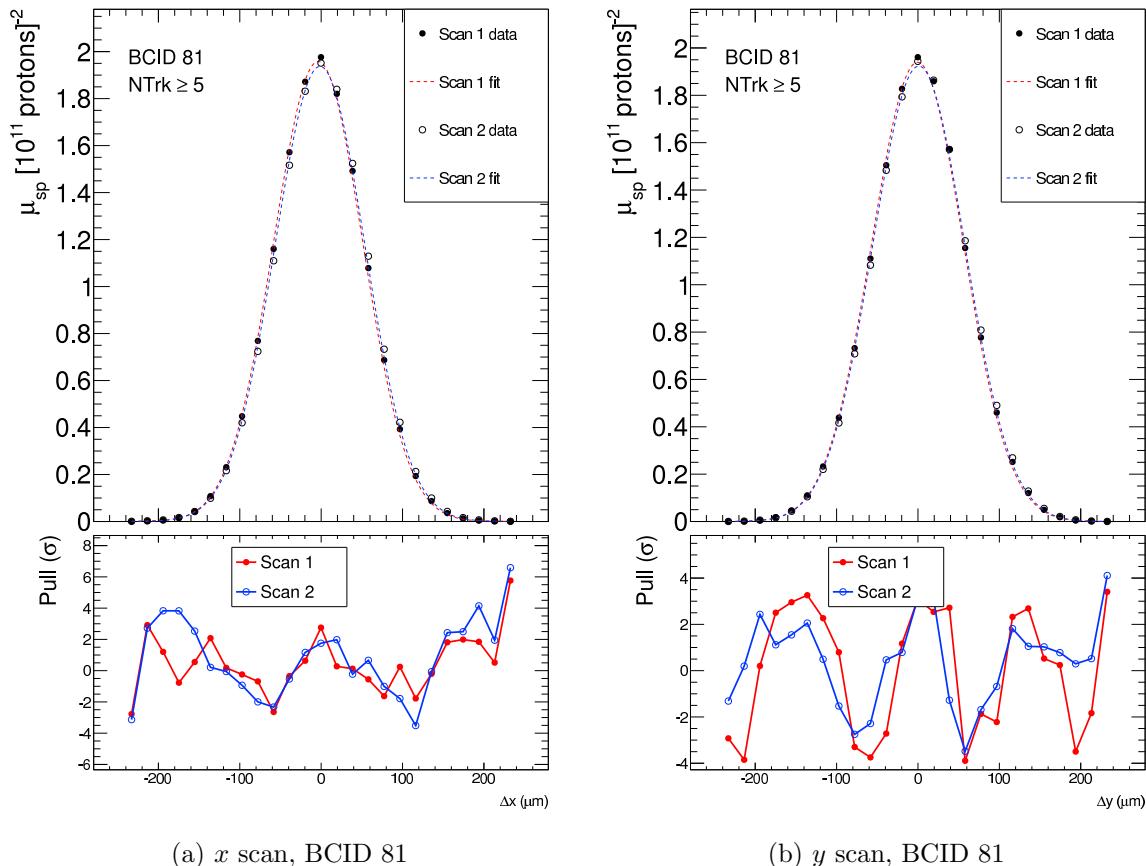


Figure 4.9: $\mu_{\text{vis}}^{(sp)}$ vs. beam separation, with single gaussian plus constant fits, and pulls.

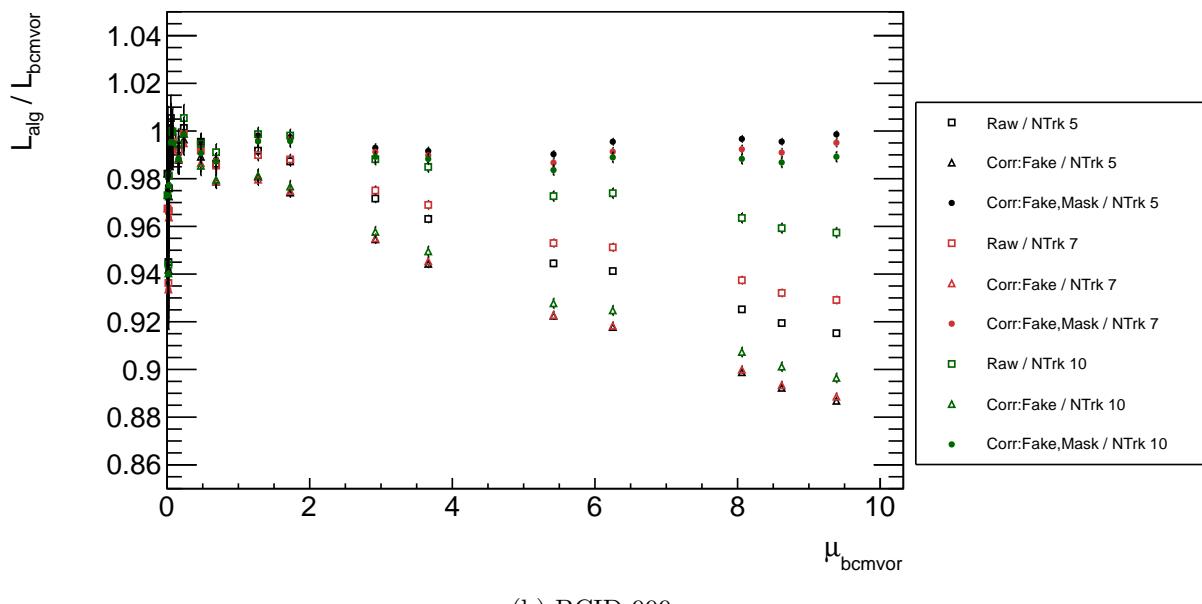
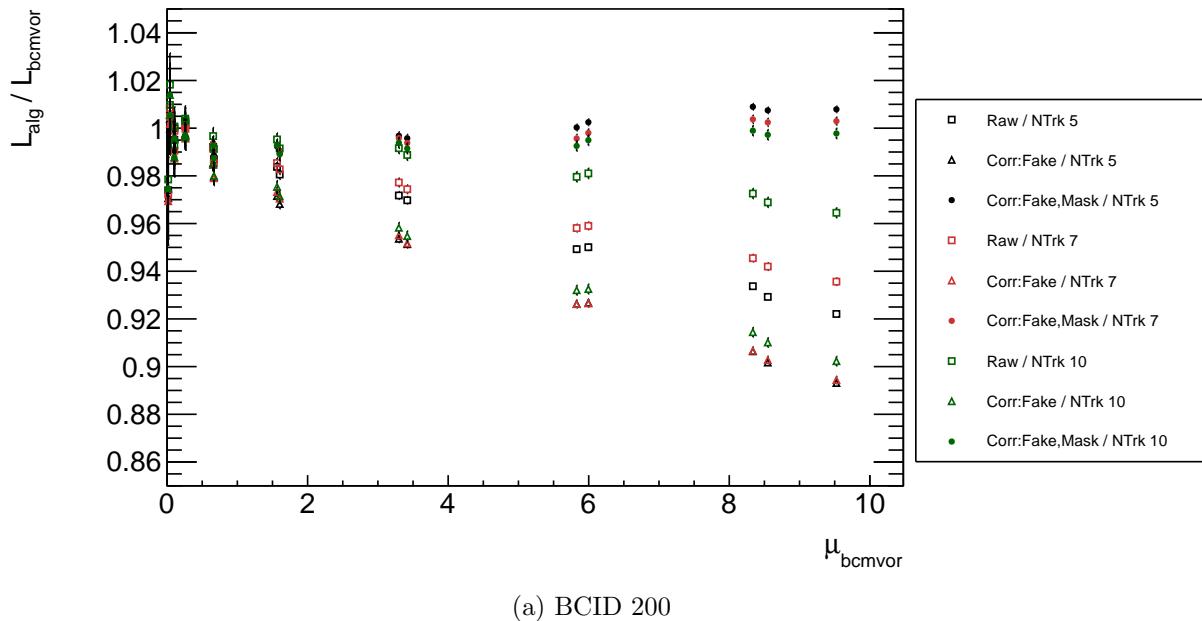


Figure 4.10: Ratio of luminosity values from vertexing and BCM_VOR, shown with each successive pileup correction applied. Fakes are subtracted first, and masking is corrected second.

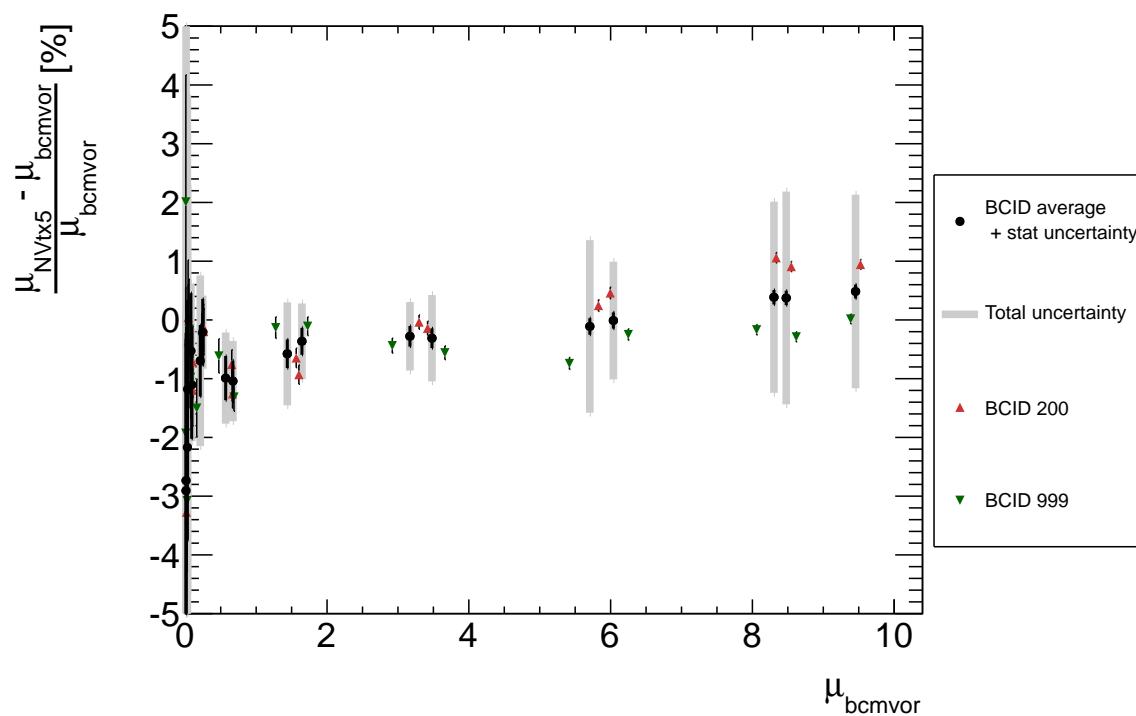


Figure 4.11: Percent difference between luminosities measured by vertex counting and BCM_VOR. The central values are taken to be the average between BCIDs 200 and 999.

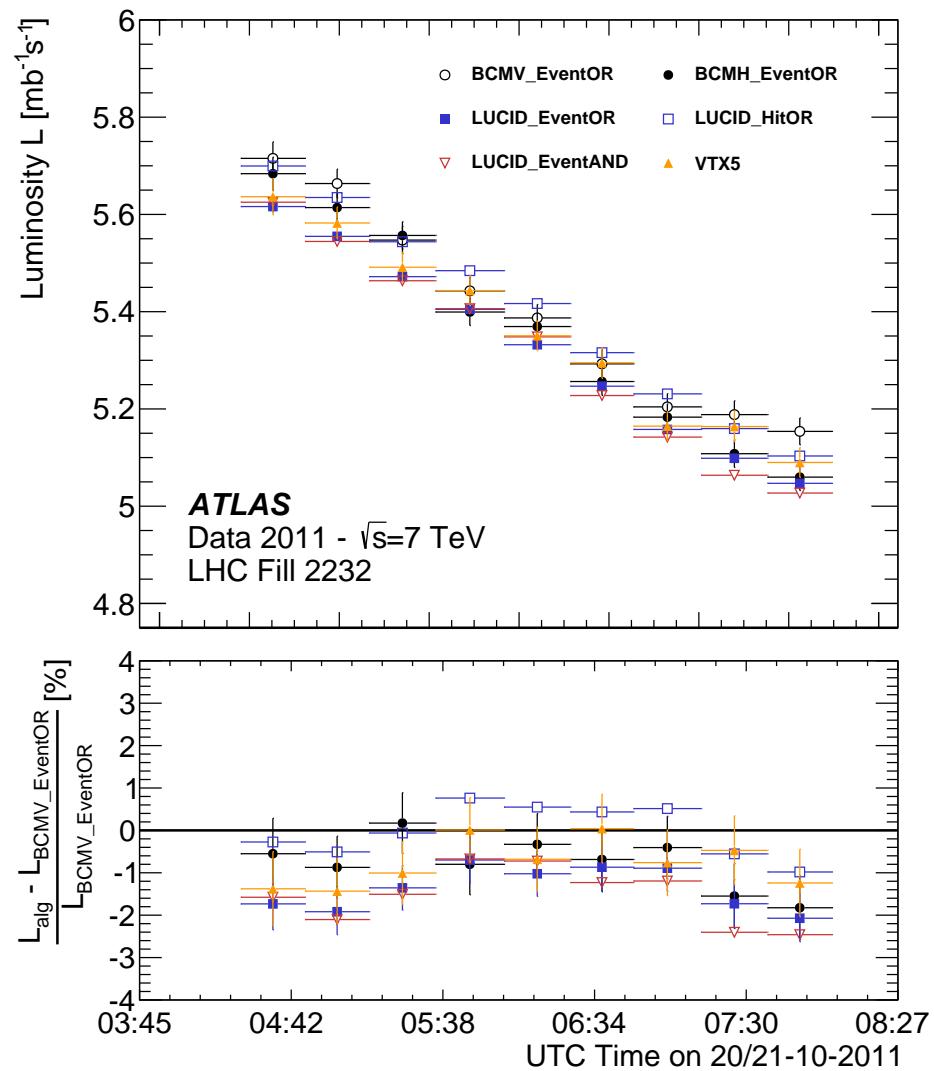


Figure 4.12: Luminosities measured by BCM, LUCID, and vertex counting during the $\beta^* = 90$ m ALFA run in October 2011.

Chapter 5

Event Reconstruction

Once recorded, the events are reconstructed using a wide variety of algorithms designed to identify the products of collisions at the center of the detector. The algorithms transform the raw data read out from the detector – hits in the inner detector silicon layers and TRT straws, energy deposits in calorimeter cells, and hits in the muon stations – into a list of physics objects and their energies or momenta. This section describes the techniques used to identify the objects used in the analyses described in chapters ?? and ??.

5.1 Track and Vertex Reconstruction

Tracks, nominally due to charged particles traversing the inner detector, are reconstructed using a series of algorithms based on hits in inner detector [**Cornelissen:2007vba**, **TheATLASCollaboration:2012tja**]. The primary “inside-out” algorithm begins by constructing three-dimensional space points associated with the hits in the pixel and SCT layers. Track seeds are formed from sets of three space points in the first four layers of the inner detector (three pixel layers and the innermost SCT layer), constrained to be consistent with a track originating from the interaction region. The track seeds are extended through the remaining SCT layers using a combinatorial Kalman filter. After screening the track candidates to reduce random coincidences and ambiguities from very close tracks, the tracks are extended through the TRT. Finally, the track is refitted using all of the associated hits.

A complementary “outside-in” algorithm identifies tracks with fewer or zero silicon hits, which can arise from hadron decays or photon conversions in the silicon layers, or from track seeds removed during the ambiguity resolution. The algorithm starts from TRT segments, and adds silicon hits proceeding inwards.

Vertices are points consistent with being the origin of a set of tracks. Primary vertices are vertices nominally due to pp collisions, while secondary vertices arise from other processes like hadron decays or photon conversions. The primary vertex reconstruction algorithm reconstructs vertices one by one, alternately finding a vertex seed and then fitting the corresponding tracks [**TheATLASCollaboration:2010wy**, **TheATLASCollaboration:2012tja**].

For most applications, the tracks are required to have $p_T > 400 \text{ MeV}$, no missing hits along the track in the pixel layers¹ (*holes*), and at most two holes in the SCT layers. The seed finding identifies the global maximum in the distribution of z coordinates of the tracks. The vertex position is then determined by an adaptive vertex fitting algorithm [Waltenberger:2007hz], a χ^2 -based fit which suppresses the contribution from outlier tracks. Tracks whose impact parameter is inconsistent by more than 7σ with the vertex position are then reused to seed a new vertex, until no further seeds can be found.

For most applications, the vertex reconstruction is performed twice. After reconstructing an initial set of primary vertices, the position and size of the interaction region, or *beam spot*, is determined from a fit to the spatial distribution of vertices. The vertices are then reconstructed a second time using the beam spot as a constraint for seed finding and track fitting².

5.2 Electrons

The signature of an electron is an energy deposit in the electromagnetic LAr calorimeter and a track pointing at the energy deposit [TheATLASCollaboration:2014fu, TheATLASCollaboration:2014vra]. The electron reconstruction algorithm begins by searching for clusters of energy in the calorimeter, based on a grid of $N_\eta \times N_\phi = 200 \times 256$ towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The tower energy is the sum of the cell energies in all longitudinal layers within the tower. Energy deposits with $E_T > 2.5 \text{ GeV}$ within a 3×5 window of towers form the seeds for both electrons and photons.

After passing loose shower shape requirements, the electron reconstruction algorithm searches for a track within a cone of radius $\Delta R = 0.3$ around the cluster barycenter. Two hypotheses are used for track pattern recognition and fitting: the standard pion hypothesis and an electron hypothesis that allows for larger energy losses due to bremsstrahlung. The cluster and track are required to satisfy one of the following two criteria:

- The barycenter of the cluster and the track extrapolation to the middle layer of the LAr calorimeter satisfy $\Delta\phi < 0.2$ in the direction of track bending or $\Delta\phi < 0.05$ in the other direction.
- The barycenter of the cluster and the track extrapolation to the middle layer of the LAr calorimeter, after rescaling the track momentum to the energy of the cluster, satisfy $\Delta\phi < 0.1$ in the direction of track bending or $\Delta\phi < 0.05$ in the other direction. This recovers low-energy electrons that potentially lose a significant fraction of their energy before reaching the calorimeter.

¹Excluding missing hits due to an inactive detector element.

²For the luminosity measurement using vertices, the beam spot constraint is not applied in order to avoid biases due to changes in the size of the interaction region

For tracks with at least four silicon hits, the track and the cluster must also satisfy $|\Delta\eta| < 0.05$. Finally, the cluster and track are rebuilt using algorithms optimized for measurement of the electron properties. The cluster is rebuilt sequentially in all four layers, using an area of 3×7 layer-2 cells in the barrel or 5×5 layer-2 cells in the end-caps. The tracks of electron candidates are refit using an optimized electron track filter based on the Gaussian Sum Filter (GSF) algorithm [**TheATLASCollaboration:2012vr**]. The GSF track and the cluster must satisfy tighter spatial matching criteria: $\Delta\phi < 0.1$ in the direction of track bending, or $\Delta\phi < 0.05$ in the opposite direction. GSF tracks with less than four silicon hits are required to satisfy even tighter criteria: $|\Delta\eta| < 0.35$ or 0.2 in the TRT barrel or end-cap, and $\Delta\phi < 0.03$ in the direction of track bending or $\Delta\phi < 0.02$ in the other direction.

Identification

Further requirements can be imposed on electron candidates to suppress backgrounds from sources like misidentified hadronic jets, photon conversions, and electrons from hadron decays. Three increasingly stringent sets of cuts are defined, called loose, medium, and tight³. The analyses described in chapters ?? and ?? use the tight cuts to select signal electrons, and the medium and loose cuts to derive data-driven background estimates. The loose set of cuts impose requirements on the shower shape in the first and second calorimeter layers, the quality of the track, the fraction of energy in hadronic calorimeter cells behind the LAr cells, and the spatial match of the track and the calorimeter cluster. The medium set of cuts consist of more stringent versions of the loose cuts, and additionally require a small track impact parameter with respect to the primary vertex, a minimum number of high-threshold TRT hits associated with the track, and a hit in the innermost pixel layer. The tight set of cuts again consist of more stringent versions of the medium cuts, with an additional cut on the ratio of the cluster energy to the track momentum (E/p) and a veto of candidates associated to a photon conversion vertex.

Efficiency Measurements

The efficiency to detect an electron can be factorized into several components. The efficiency to detect a cluster in the electromagnetic calorimeter is very high, above 99% for electrons with $E_T = 15$ GeV and 99.9% for $E_T = 45$ GeV. The reconstruction efficiency, covering the matching of a good-quality track to the cluster, and the identification efficiencies, covering the identification cuts with respect to reconstructed electrons, are measured using tag-and-probe techniques targeting $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ events [**TheATLASCollaboration:2014vz**]. One electron, the tag, is required to satisfy strict selection criteria, while the second electron, the probe, is used for efficiency measurements. The combined reconstruction and identification efficiencies are shown in figure ???. The reconstruction efficiency makes up

³An alternative method of identification using a likelihood-based multivariate method has been developed, but is not used in this dissertation

1%-5% of the total efficiency loss for electrons with $E_T < 20$ GeV, and less than 1% for electrons with $E_T > 80$ GeV. The efficiencies are computed for both data and simulation, and the ratio between the two is used to correct the efficiency in simulation.

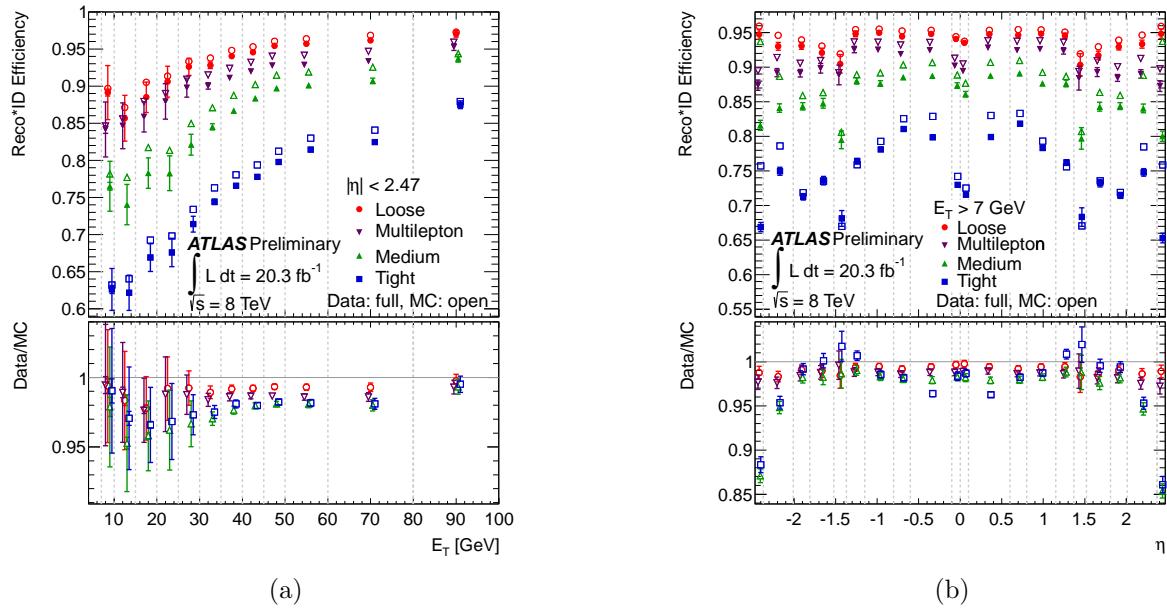


Figure 5.1: The combined reconstruction and identification efficiencies with respect to electrons detected as a cluster in the electromagnetic calorimeter, shown as a function of E_T (left) and η (right).

Energy and Momentum Measurement

For electron candidates with at least four silicon hits, the energy of the electron is taken from the calorimeter measurement, while the trajectory is taken from the GSF track. Candidates with fewer silicon hits are not used in this dissertation. The energy resolution of the calorimeter is parametrized as:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (5.1)$$

where a , b , and c are the sampling, noise, and constant terms, respectively, and generally vary with η . The design value of the sampling term is $a \approx 9 - 10\%$ in the central region, and worsens at higher pseudorapidities due to the increased material in front of the calorimeters.

The noise term, due to electronics noise and pileup effects, is approximately $b = 350 \times \cosh \eta$ MeV. The constant term dominates the resolution at high energies, with a design value of $c = 0.7\%$.

As the ATLAS calorimeters are non-compensating, the energy measurement is calibrated using a multivariate algorithm trained on single-electron simulation to determine the most probable electron energy. The method takes into account differences between data and simulation in the energy scales of each longitudinal layer and other detector effects not modeled in simulation.

After the initial simulation-based calibration, the electron energy scale and resolution are determined using $Z \rightarrow ee$ events. Residual differences in the energy scale between data and simulation are parametrized as:

$$E^{\text{data}} = E^{\text{simulation}}(1 + \alpha_i), \quad (5.2)$$

where the α_i quantify the energy scale difference in bins of pseudorapidity. The difference in energy resolution is derived assuming that the $Z \rightarrow ee$ invariant mass distribution is well-modeled up to a Gaussian constant term,

$$\left(\frac{\sigma_E}{E}\right)^{\text{data}} = \left(\frac{\sigma_E}{E}\right)^{\text{simulation}} \oplus c_i, \quad (5.3)$$

where i again denotes bins in pseudorapidity. Histograms of the $Z \rightarrow ee$ invariant mass distribution in simulation are then produced for a range of α_i and c_i values, and the optimal values are determined using a χ^2 minimization with respect to data. The energy scale corrections are found to be $\alpha_i \sim -1\%$ in the barrel, $\sim 1\%-2\%$ in the barrel/end-cap transition region, and up to $\sim -4\%$ in the end-caps. The resolution corrections are symmetrized about $\eta = 0$, with values of $c_i \approx 0.8\%$ in the barrel, $\sim 3.3\%$ in the barrel/end-cap transition region, and $\sim 0.5\%-2.5\%$ in the end-caps. The energy resolution is shown as a function of E_T in figure ??.

Sources of systematic uncertainty on the energy scale include the intrinsic accuracy of the $Z \rightarrow ee$ method, determined from closure tests where known scale variations are injected into simulation, calorimeter gain and pedestal dependence, data-simulation differences in the calibration of the calorimeter layers, and the modeling of material in front of and within the calorimeter. The total uncertainty ranges from 0.03%-0.22% for $E_T = 40$ GeV and 0.27%-2.25% for $E_T = 200$ GeV, with larger uncertainties in the pseudorapidity range $1.37 < |\eta| < 1.82$ corresponding to the transition region between the barrel and the end-cap.

The systematic uncertainty on the energy resolution is less than 10% for electrons with $E_T < 50$ GeV, and asymptotically approaches $\sim 40\%$ for high E_T . At low energies, the pileup contribution to the noise term, b , dominates the uncertainty; at higher energy, uncertainty is due to a mix of the sampling term, a , the pileup contribution to b , the modeling of material, and the intrinsic accuracy of the $Z \rightarrow ee$ method.

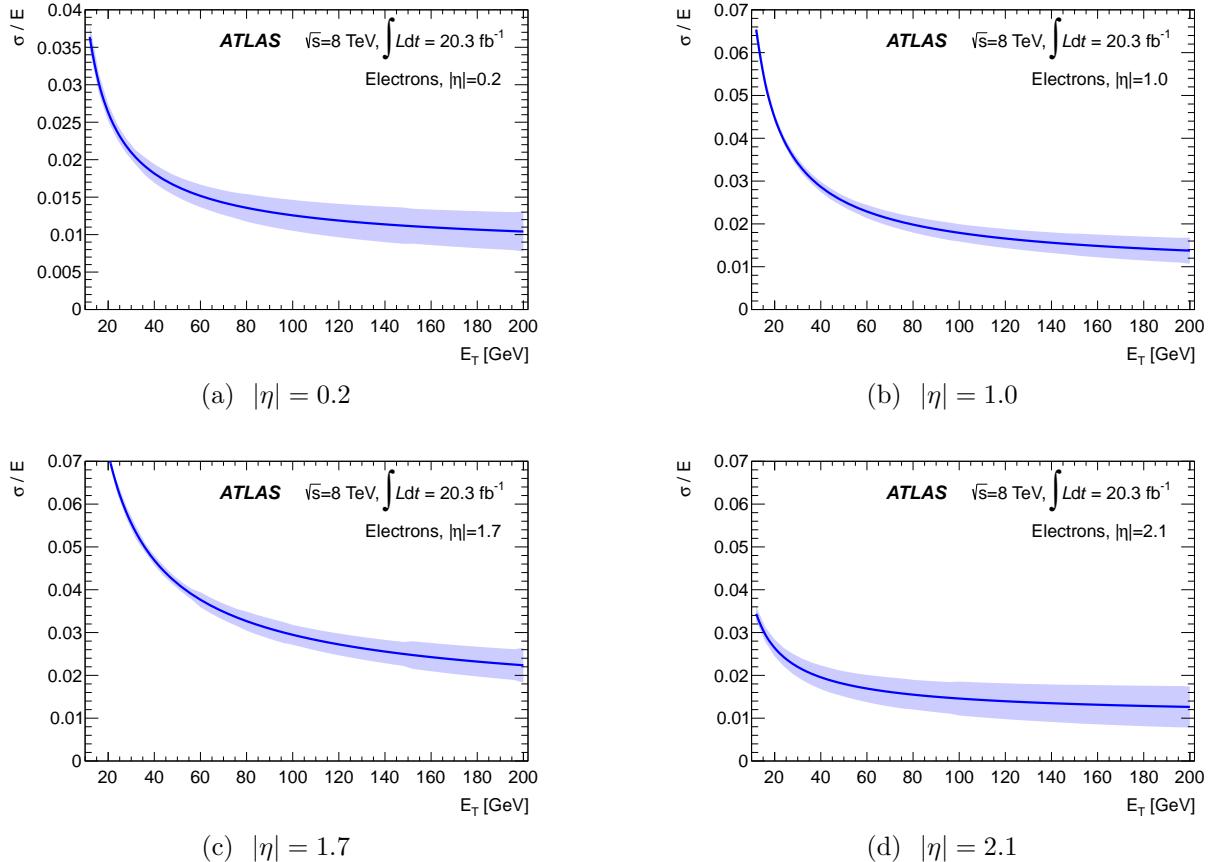


Figure 5.2: Electron energy resolution as a function of E_T for various values of $|\eta|$. The shaded band shows the uncertainty on the resolution [TheATLASCollaboration:2014gf].

5.3 Muons

Muons are identified by matching tracks in the muon spectrometer to tracks in the inner detector. The requirements differ based on the instrumentation available in the vicinity of the muon candidate. The analyses described here use *combined* (CB) muons, consisting of matched tracks reconstructed independently in the inner detector and the muon spectrometer. The muon momentum is determined from a statistical combination of the two track's parameters and their corresponding covariance matrices. Combined muons have the highest purity, but suffer from a loss of acceptance near $\eta \sim 0$, where the muon spectrometer has gaps to accommodate services for the inner detector and calorimeters, and $1.1 < \eta < 1.3$, where some trajectories only pass through one muon station due to incomplete installation. The remaining categories are *standalone* (SA) muons, consisting of a track only in the muon spectrometer; *segment-tagged* (ST) muons, consisting of an inner detector track and one

or more track segments in the MDT or CSC chambers; and *calorimeter-tagged* (CaloTag) muons, consisting of an inner detector track matched to a calorimeter energy deposit consistent with the passage of a muon. These categories recover efficiency in regions of the detector with less instrumentation at the cost of lower muon purity, and are not used in this dissertation.

For all categories of muons, the inner detector track is required to have at least 1 pixel hit, at least 5 SCT hits, at most 2 pixel or SCT holes, and at least 9 TRT hits for $0.1 < |\eta| < 1.9$. Energy losses in the calorimeter due to ionization, bremsstrahlung, and electron pair production must also be taken into account.

Efficiency Measurements

The efficiency of the reconstructing CB muons is measured as:

$$\epsilon(\text{CB}) = \epsilon(\text{CB}|\text{ID})\epsilon(\text{ID}|\text{MS}), \quad (5.4)$$

where $\epsilon(\text{CB}|\text{ID})$ is the probability that a muon reconstructed as an inner detector track is also reconstructed as a CB muon, and $\epsilon(\text{ID}|\text{MS}) \approx \epsilon(\text{ID})$ is the probability that a muon with a track in the muon spectrometer, i.e. a CB or SA muon, is also reconstructed as an inner detector track. The latter approximation is made because $\epsilon(\text{ID})$ is not directly accessible in data.

The efficiencies are measured using tag-and-probe techniques similar to those described in section ??, targeting $Z \rightarrow \mu\mu$ and $J/\Psi \rightarrow \mu\mu$ events. In this case, the tag muon is required to be a CB muon, and the probe muon is a CB or SA muon in the case of measuring $\epsilon(\text{ID}|\text{MS})$, and a CaloTag muon for the measurement of $\epsilon(\text{CB}|\text{ID})$. The efficiencies for all types of muon are shown in figure ???. CB muons have an efficiency of greater than 97% in most of the pseudorapidity range, except for significant inefficiencies due to gaps in the muon spectrometer in the ranges $|\eta| < 0.1$ and $1.1 < \eta < 1.3$. The measured efficiencies in data and simulation agree to within $\sim 2\%$, and the ratios in each pseudorapidity bin are used as scale factors to correct the efficiency in simulation. The systematic uncertainty on the scale factors is in the range 0.1-0.3%, rising near $|\eta| \sim 0$ and $|\eta| \sim 2.5$.

Energy Scale and Resolution

The muon momenta in simulation are scaled and smeared to match the momentum scale and resolution in data. The corrections are derived in bins of η and ϕ , with boundaries chosen to minimize the variation of the correction in each bin, and are applied separately to the transverse momenta measured by the inner detector (ID) and the muon spectrometer (MS). Specifically, the correction is implemented as:

$$p_T^{\text{Cor},\text{Det}} = \frac{p_T^{\text{MC},\text{Det}} + \sum_{n=0}^1 s_n^{\text{Det}(\eta,\phi)}(p_T^{\text{MC},\text{Det}})^n}{1 + \sum_{m=0}^2 \Delta r_m^{\text{Det}}(\eta,\phi)(p_T^{\text{MC},\text{Det}})^{m-1} g_m}, \quad (5.5)$$

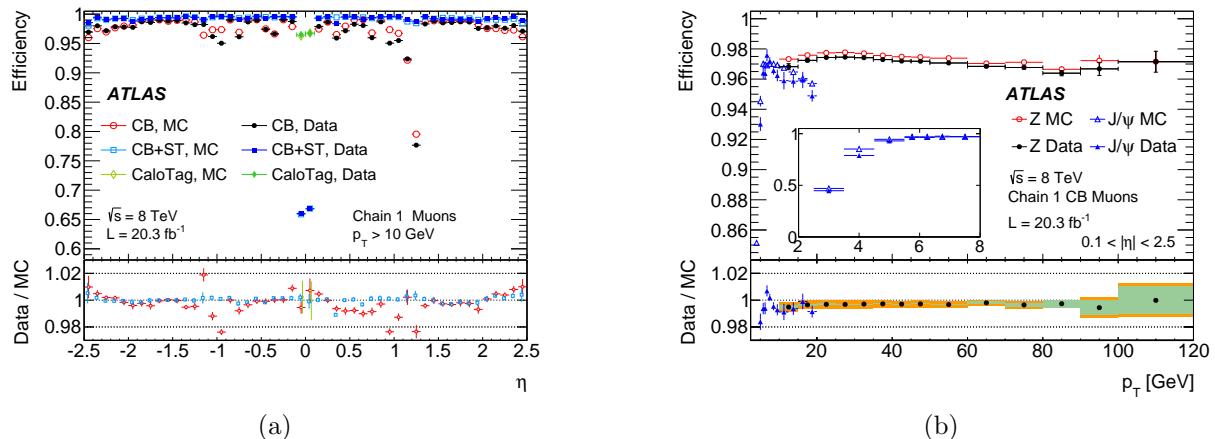


Figure 5.3: Muon reconstruction efficiencies as a function of η for muons with $p_T > 10$ GeV (left), and as a function of p_T for muons with $0.1 < |\eta| < 2.5$ (right). The uncertainty bars on the points indicate statistical uncertainties. The bottom plots show the ratio between the measured and simulated efficiencies, with the combination of statistical and systematic uncertainties indicated by the uncertainty bars.

where $\text{Det}=\text{ID}$ or MS , the $\Delta r_m^{\text{Det}}(\eta, \phi)$ parametrize the momentum resolution smearing, the $s_n^{\text{Det}}(\eta, \phi)$ parametrize the scale corrections, and the g_m are normally-distributed random variable with mean 0 and width 1^4 . The constant scale correction term, s_0^{MS} , accounts for the difference between data and simulation in the energy lost by muons before reaching the muon spectrometer. s_0^{ID} is set to zero due to the negligible energy loss before the inner detector. The linear terms, $s_1^{\text{ID,MS}}$, models discrepancies between data and simulation in the magnetic field integral and the radial dimension of the detector. The resolution corrections Δr_m^{Det} represent deviations from the resolution in data, which is parametrized empirically as:

$$\frac{\sigma(p_T)}{p_T} = \frac{r_0}{p_T} \oplus r_1 \oplus r2 \cdot p_T, \quad (5.6)$$

where the r_0 term describes energy lost by muons as they traverse the material of the detector, the r_1 term describes multiple scattering, magnetic field inhomogeneities, and local radial displacements, and the r_2 term describes intrinsic resolution effects due to the spatial resolution of the hit measurements and residual misalignment.

The corrections are derived from $J/\Psi \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$, and $Z \rightarrow \mu\mu$ events using a template maximum likelihood fit, similar to that described in section ???. The effect of the corrections on the invariant mass distribution of $Z \rightarrow \mu\mu$ events is shown in figure ???, along

⁴Note that this equation does not apply to cases where the resolution is better than that in simulation. In these cases, the resolution difference is included in the positive ID and MS variations, and the effect of the positive variation on the physical observables is symmetrized about the nominal value.

with the total systematic uncertainty. The muon resolution as measured in Z , Υ , and J/Ψ events is shown in figure ??.

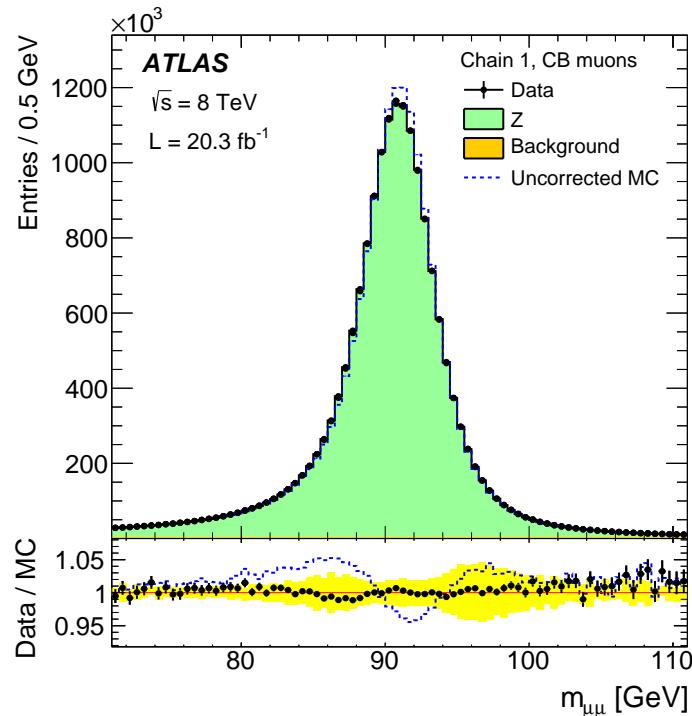
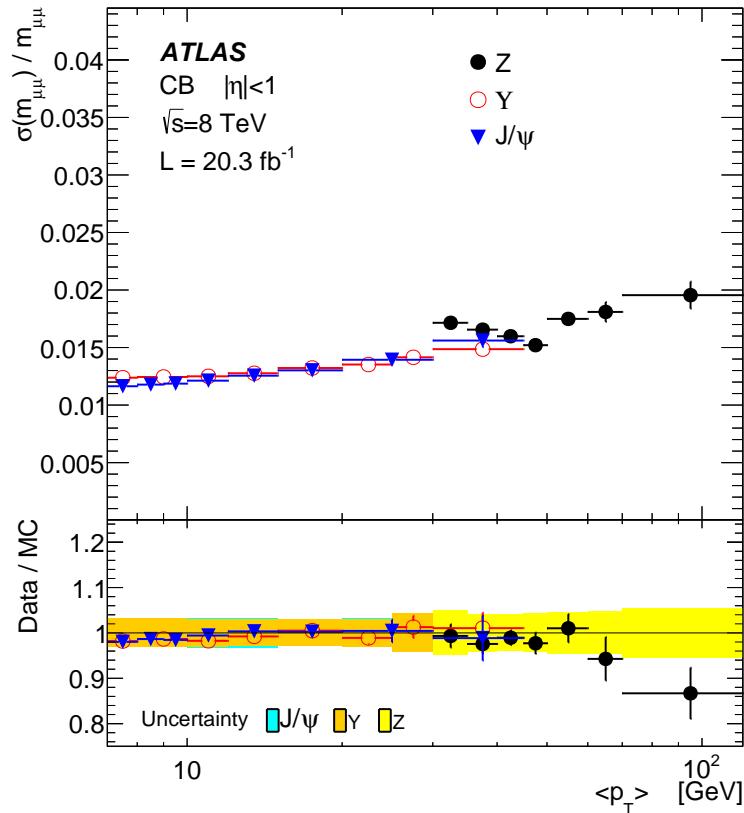
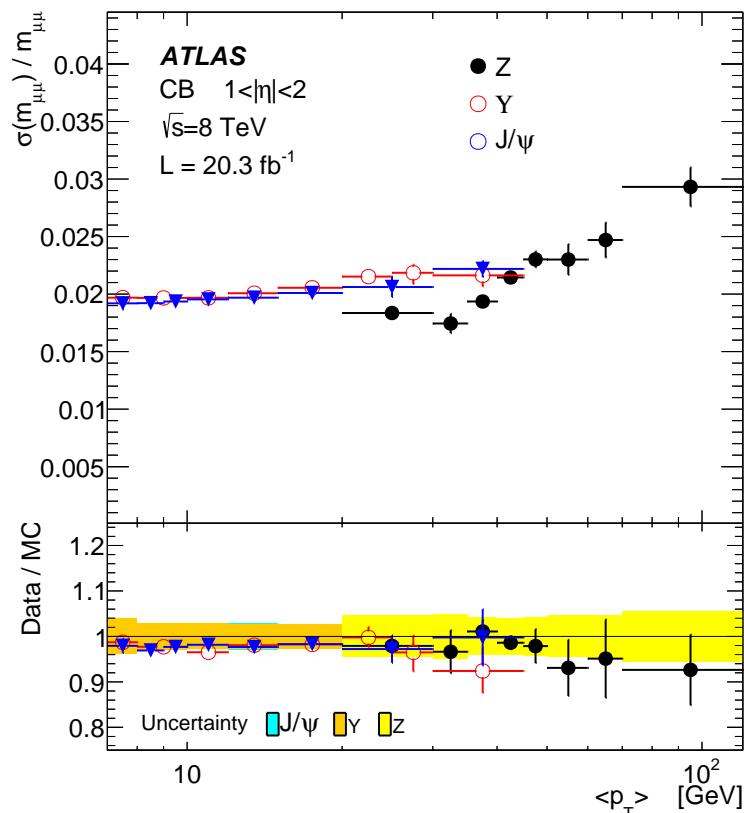


Figure 5.4: Dimuon invariant mass distribution of $Z \rightarrow \mu\mu$ events, using CB muons. The top panel shows the invariant mass distribution for data (black points), expected $Z \rightarrow \mu\mu$ signal from simulation (green), and expected background from simulation (yellow). The simulated background and signal have the muon momentum corrections applied, and their sum is normalized to the data. The dashed histogram shows the background plus signal without muon momentum corrections.

5.4 τ Leptons

The signature of τ leptons is significantly more complex than electrons and muons due to the fact that they decay to a diverse set of final states. Tau leptons have a proper decay length of $87\mu\text{m}$, and therefore typically decay before reaching the active layers of the detector. The leading decay modes are shown in table ???. 35.2% of τ leptons decay to an electron or muon plus two neutrinos. The remaining 64.8% decay to hadrons plus a neutrino. The hadronic decay modes contain one charged pion in 72% of the decays, and three charged pions in 22% of the decays; the majority of the remainder contain one or more kaons. The hadronic decay modes also frequently contain neutral pions, with 78% containing at least one neutral pion.

(a) $|\eta| < 1$ (b) $1 < |\eta| < 2$

Decay	Branching Fraction [%]
$e^- \bar{\nu}_e \nu_\tau$	17.83 ± 0.04
$\mu^- \bar{\nu}_\mu \nu_\tau$	17.41 ± 0.04
$\pi^- \nu_\tau$	10.83 ± 0.06
$\pi^- \pi^0 \nu_\tau$	25.52 ± 0.09
$\pi^- \pi^0 \pi^0 \nu_\tau$	9.30 ± 0.11
$\pi^- \pi^+ \pi^- \nu_\tau$	9.31 ± 0.06
$\pi^- \pi^+ \pi^- \pi^0 \nu_\tau$	4.62 ± 0.06

Table 5.1: Leading branching fractions of the τ lepton to final states with leptons or pions. Most of the remaining decays are to final states with kaons. From [pdg].

In this dissertation, no effort is made to identify or reconstruct leptonic τ decays (τ_{lep}), regarding them only as electrons or muons plus missing transverse energy (see section ??). Hadronic τ decays (τ_{had}) are identified using their visible decay products, namely the neutral and charged hadrons, which are collectively called $\tau_{\text{had-vis}}$. The signature consists of a narrow jet with one or three tracks, called one-prong or three-prong decays, respectively. Up to two $\pi^0 \rightarrow \gamma\gamma$ decays are also included. The reconstruction and identification proceeds as follows [TheATLASCollaboration:2014tga]:

- Jets are reconstructed using the anti- k_t algorithm with a distance parameter of $R = 0.4$, built from TopoClusters calibrated with a local hadronic calibration. Jets with $p_T > 10$ GeV and $|\eta| < 2.5$ are used as seeds for τ lepton candidates.
- For each jet seed, the τ vertex is chosen to be the primary vertex with the greatest $\sum p_T$ of tracks in a cone of radius $\Delta R = 0.2$ around the jet seed. This vertex defines the $\tau_{\text{had-vis}}$ direction, i.e. is used to determine the η and ϕ of the τ candidate.
- π^0 candidates, consisting of a pair clusters within $\Delta R < 0.2$ of the τ candidate, are identified using a BDT-based algorithm. Up to two π^0 s are considered.
- A multivariate algorithm based on boosted decision trees (BDTs) is used to discriminate hadronic τ decays from the backgrounds, primarily due to jets with low track multiplicity. The BDTs use many input variables describing the energy cluster, the spatial arrangement and energy of the tracks, and the neutral pions. Separate BDTs are trained for one-prong and three-prong τ decays, using simulated $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$, and $Z' \rightarrow \tau\tau$ decays for signal and collision data samples for the background. Three working points with different identification efficiencies are defined, called loose, medium, and tight. The performance of the identification algorithms is shown in figure ??, along with correction factors applied to simulated samples to equalize the efficiencies in simulation and real data. The correction factors are derived by measuring the efficiencies in data and simulation, using a tag-and-probe method targeting $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ events.

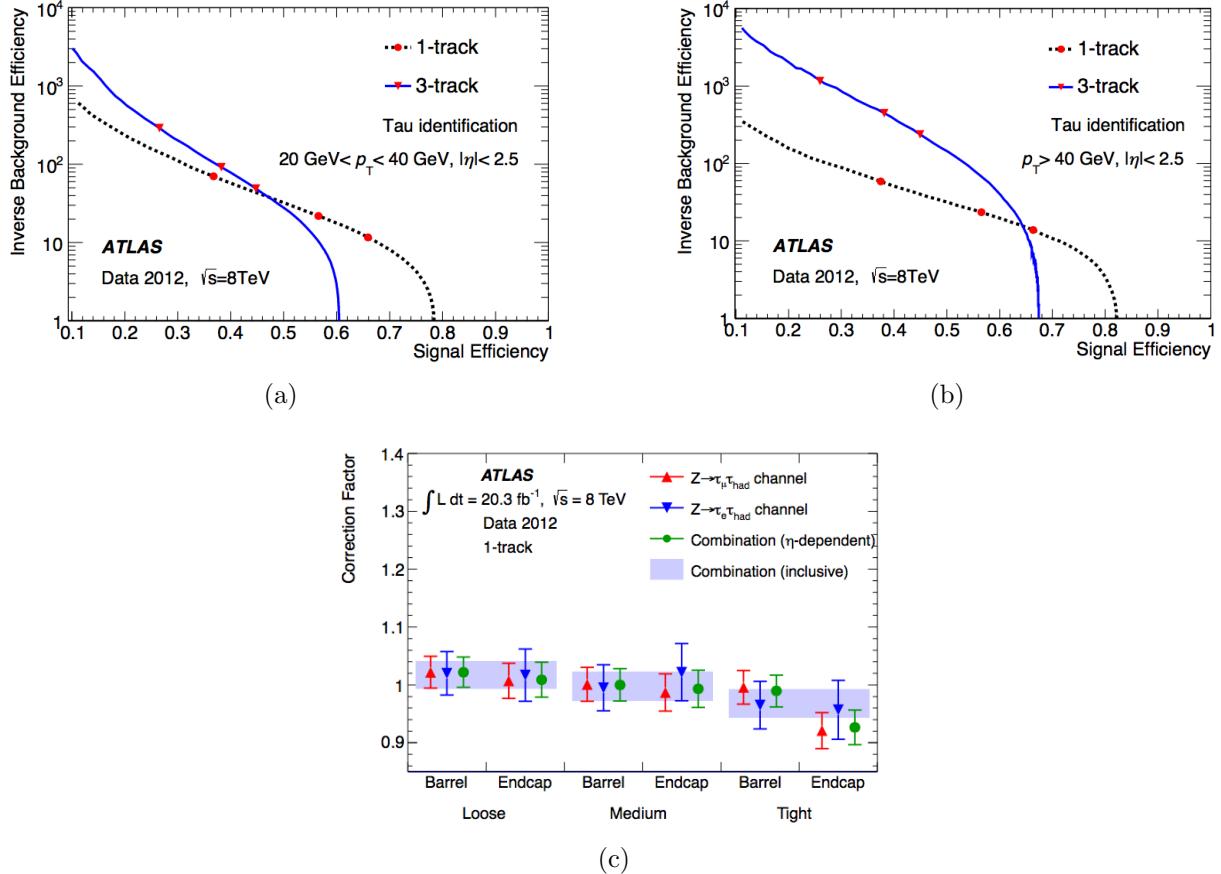


Figure 5.6: Top: inverse background efficiency (rejection power) versus signal efficiency for the BDT-based offline τ identification. Low p_T τ s ($20 \text{ GeV} < p_T < 40 \text{ GeV}$) are shown at left, and high p_T τ s ($p_T > 40 \text{ GeV}$) are shown at right. The background consists of simulated multijet events, while the signal consists of simulated Z , W , and Z' events decaying to τ leptons. Bottom: Correction factors applied to simulation to equalize the efficiency to that measured in Z tag-and-probe data.

- An additional BDT-based algorithm rejects one-prong τ candidates consistent with an electron. The most powerful discriminating variables are the ratio of high- to low-threshold TRT hits on the track and the ratio of energies deposited in the electromagnetic and hadronic calorimeters. The electron rejection power versus the τ efficiency, derived from simulated $Z \rightarrow ee$ events, is shown in figure ??.

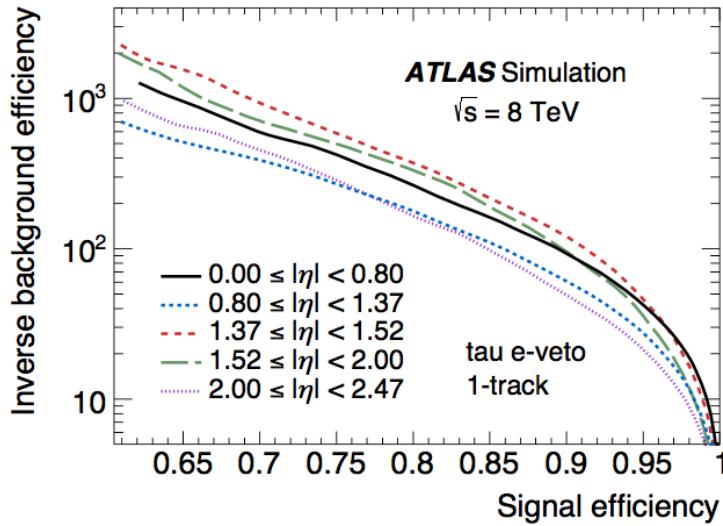


Figure 5.7: Electron rejection power versus 1-track τ_{had} efficiency for the electron rejection BDT.

Energy Scale and Resolution

The hadronic τ reconstruction and identification are based on calorimeter cells calibrated at the local hadronic scale. Several corrections are applied to correct the energy to a τ -specific energy scale (TES). First, corrections are derived using simulated $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$, and $Z' \rightarrow \tau\tau$ events, generated with PYTHIA8. These corrections are determined as a function of the reconstructed $\tau_{\text{had-vis}}$ energy and pseudorapidity based on the medium identification working point. A small pseudorapidity correction, reaching up to $|\Delta\eta| = 0.01$, corrects a bias due to underestimated cluster energies in poorly-instrumented regions of the calorimeter. To account for pileup, 90–420 MeV per additional reconstructed vertex is subtracted from the reconstructed $\tau_{\text{had-vis}}$ energy, depending on η . The simulated $\tau_{\text{had-vis}}$ energy resolution after these corrections is shown in figure ??, and ranges between about 20% at very low energy to about 5% above a few hundred GeV.

Finally, data-driven corrections and systematic uncertainties on the τ energy scale are derived using a *deconvolution* method [ref-49-from-tau-paper]. The method combines the systematic uncertainties on the single-particle response of the calorimeters based on the well-known branching fractions of hadronically decaying τ s. A TES correction of about 1% is determined. Including additional systematic uncertainties covering the detector modeling, pileup, non-closure of the calibration method, and the hadronic shower model, the total TES uncertainty is 2 – 3% for one-prong decays and 2 – 4% for three-prong decays.

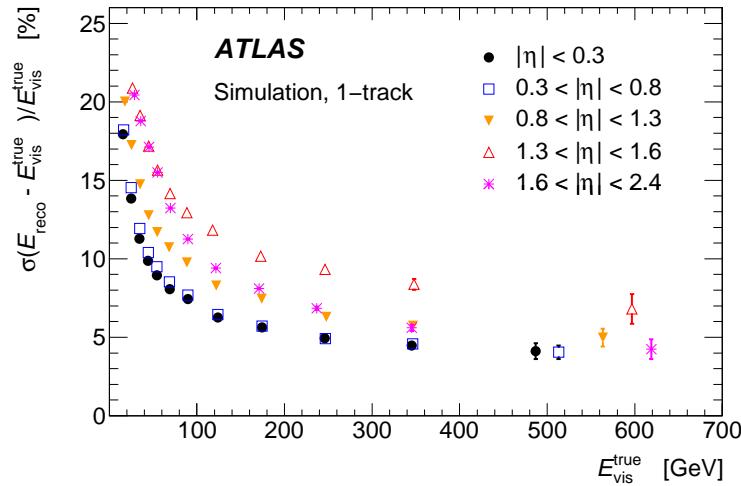


Figure 5.8: Energy resolution for hadronically decaying τ leptons with one associated track in various pseudorapidity regions. The resolution is the standard deviation of a Gaussian fit to the distribution of $\frac{E_{\text{reco}} - E_{\text{true-vis}}}{E_{\text{true-vis}}}$ in bins of $E_{\text{true-vis}}$ and $|\eta_{\text{true-vis}}|$.

5.5 Jets

Jets play an important complementary role to leptons in the analyses described in the following chapters. The new physics scenarios considered often produce jets in addition to the three required leptons. If the new particles are colored, as in strongly produced supersymmetry, then high- p_T jets can be produced in cascade decays; alternatively, if the new particles decay via the weak interaction, then the decays will often contain hadronically decaying weak bosons. From the background point of view, due to the colored initial state in pp collisions, jets are produced copiously at the LHC. Despite the stringent lepton identification cuts, jets misidentified as leptons or containing semileptonic decays can constitute a significant source of backgrounds to trilepton final states.

The reconstruction of jets in the calorimeter is based on topological clusters of energy [TheATLASCollaboration:2011ks, TheATLASCollaboration:2015ds]. The reconstruction steps are shown in ???. Clusters are formed by grouping together calorimeter cells based on their signal-to-noise ratio, S/N , where the N includes electronic noise and contributions from pileup interactions. Cells with $S/N > 4$ form the cluster seeds, which are then expanded to include all connected cells with $S/N > 2$. Finally, cells along the perimeter with $S/N > 0$ are added to the cluster. The cluster energies are then calibrated according to two complementary scales, leading to two classes of jets: the electromagnetic (EM) scale assumes that the cell energy is due to an electromagnetic shower, while the local cell signal weighting (LCW) method classifies energy deposits as electromagnetic or hadronic in origin.

Jet finding algorithms then group the calibrated clusters into collections of jets. The algorithm used in this dissertation is the anti- k_t algorithm with a distance parameter of

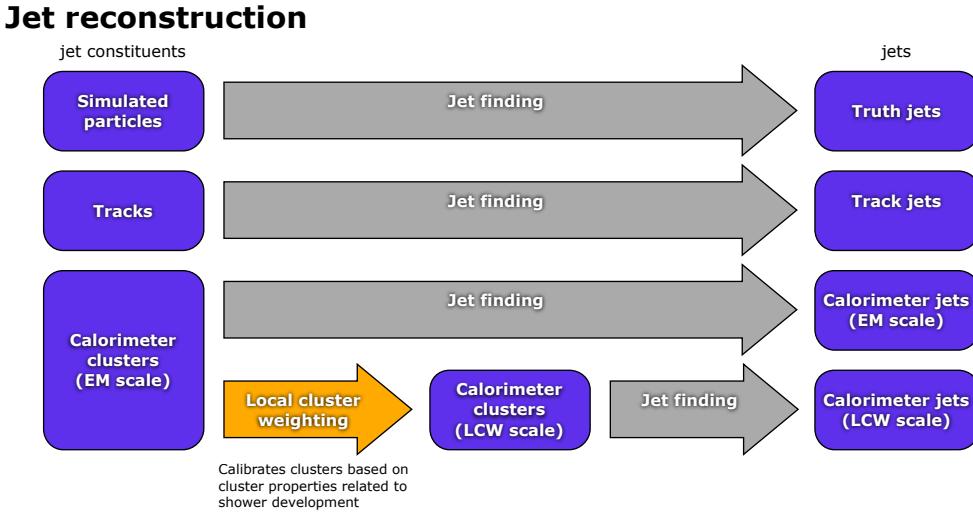


Figure 5.9: Overview of the jet reconstruction. The bottom two rows show the reconstruction of calorimeter jets, while the top two rows show jets built from truth particles in simulation and from tracks.

$R = 0.4$, implemented in FASTJET [Cacciari:2008gp, Cacciari:2011ma]. A distance measure between objects, $d_{i,j}$, is defined as

$$d_{i,j} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (5.7)$$

$$d_{i,B} = k_{t,i}^{2p} \quad (5.8)$$

where $k_{t,i}$ is the transverse momentum of object i , $\Delta_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is the distance in rapidity ($y_{i,j}$) and azimuthal angle ($\phi_{i,j}$), and $p = -1$. $d_{i,B}$ denotes the distance between objects and the beam. The algorithm combines objects sequentially by considering the smallest distance in the event. If the smallest distance is a $d_{i,j}$, then objects i and j are combined. If the smallest distance is a $d_{i,B}$, then object i is classified as a jet and removed from the object list. The algorithm continues until there are no objects left.

Loose quality cuts are applied to reject events with jets due to non-collision sources, such as beam-gas collisions between protons in one beam and the residual gas in the beam pipe, beam-halo events due to collisions with upstream collimators, muons from cosmic rays, and calorimeter noise. The cuts have an efficiency of above 99.8% for retaining real jets.

Energy Scale and Resolution

The jet energy is calibrated in several steps, as shown in figure ???. First, pileup and origin corrections are applied. Energy contributions from pileup interactions in the same or nearby bunch crossings are subtracted based on simulation, using the number of reconstructed

vertices in the event, the expected average number of interactions per crossing, the jet area, and the median p_T density of jets in the event. The geometry of the jet is corrected to point from the primary event vertex, rather than the nominal center of the ATLAS detector.

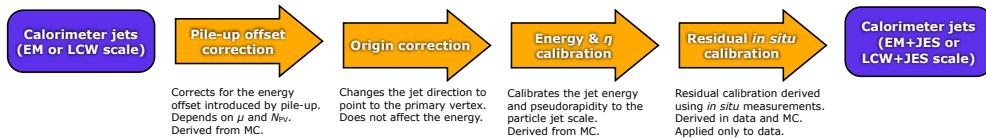


Figure 5.10: Overview of the jet calibration scheme.

The energy and pseudorapidity of the jet are initially calibrated based on the relationship between jets reconstructed from calorimeter clusters and jets reconstructed from truth particles in simulation. Then, *in situ* corrections are applied to the calibrated jets to correct for effects not described by the initial simulation-based calibration. The corrections are based on balancing the transverse momenta of jets against other well-calibrated objects. Jets in the central region, with detector rapidity $|\eta_{\text{det}}| < 1.2$, are calibrated against photons and leptonically decaying Z bosons. Jets with high transverse momentum, $p_T > 210 \text{ GeV}$, are also calibrated using multijet events, balancing the high- p_T jet against several low- p_T jets. Jets with $1.2 < |\eta_{\text{det}}| < 2.8$ are calibrated against central jets with $|\eta_{\text{det}}| < 0.8$; due to a lack of statistics with which to derive the calibration, more forward jets, with $\pm 2.8 < \eta_{\text{det}} < 4.5$, are assigned the same calibration as jets with $\eta_{\text{det}} = \pm 2.8$. The *in situ* calibrations shift the jet energy in data down by about 2% for $p_T < 100 \text{ GeV}$, and decreases with increasing p_T .

A large number of systematic uncertainties are assigned to the *in situ* calibration procedure, related to the modeling of the detector and physics processes in simulation and to the *in situ* methods themselves. The uncertainties are summarized in figure ???. Additionally, systematic uncertainties related to the pileup correction, high- p_T jets, jet flavor, and differences simulation settings are assigned where appropriate.

Pileup Suppression

Jets due to pileup interactions can be suppressed using the track information, at the cost of introducing some pileup-dependent jet reconstruction inefficiency. For a given jet and primary vertex, the *jet vertex fraction* (JVF) is defined as the ratio of the $\sum p_T$ of tracks associated with the jet and matched to the primary vertex to the $\sum p_T$ of all tracks associated to the jet, shown schematically in figure ???. Specifically, letting \mathcal{T} be the set of tracks associated to the jet, the JVF is defined as:

$$\text{JVF} = \frac{\sum_{i \in \mathcal{T}} p_{T,i} \Theta_V(i)}{\sum_{i \in \text{jet}} p_{T,i}}, \quad (5.9)$$

where $\Theta_V(i) = 1$ if the track i is matched to the primary vertex and 0 otherwise. If \mathcal{T} is empty, then $\text{JVF} \equiv -1$. In this dissertation, the primary vertex is chosen to have the highest $\sum p_T^2$ of tracks, and $\text{JVF} > 0.5$ is required for jets with $20 \text{ GeV} < p_T < 50 \text{ GeV}$.

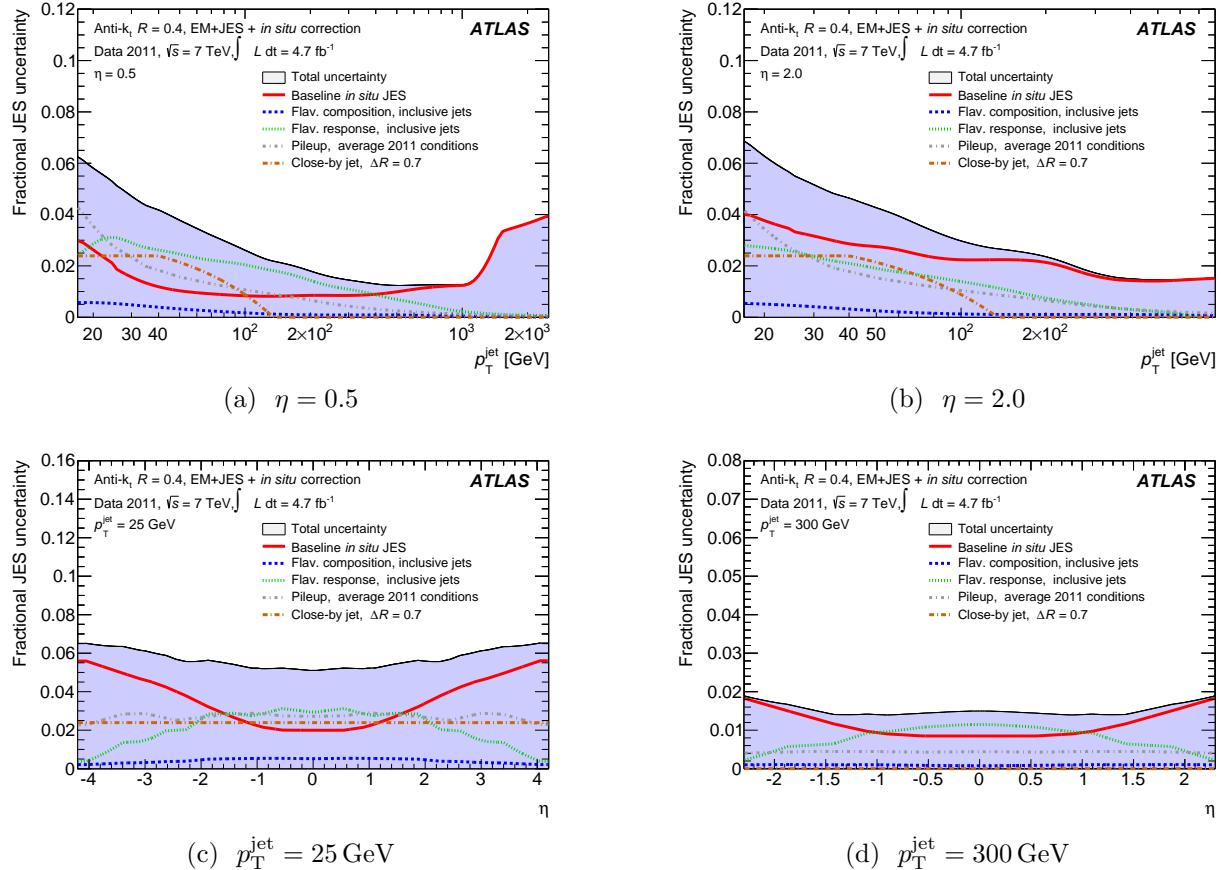


Figure 5.11: Sample-dependent fractional jet energy scale systematic uncertainty as a function of p_T^{jet} (top) and η (bottom). The jets are reconstruction using the anti- k_{T} algorithm with a radius parameter of $R = 0.4$ from clusters at the LCW scale, and are calibrated as described in the text.

b-tagging

The model-independent trilepton analysis (chapter ??) also makes use of the tagging of jets due to b quarks. Jets containing b quarks have several features which distinguish them from jets due to light quarks or gluons. B hadrons have proper decay lengths of approximate 0.5mm, long enough to be observed in the form of tracks or vertices reconstructed away from the primary interaction point. They also have large masses compared to other hadrons, leading to wider jets with more particles.

The algorithm used to tag b -jets is an artificial neural network called the MV1 algorithm [MV1]. The neural network uses three simpler likelihood-based algorithms as inputs [TheATLASCollaboration:2009ut, TheATLASCollaboration:2011wh].

- **IP3D**: Uses the transverse and longitudinal impact parameters of the tracks associated

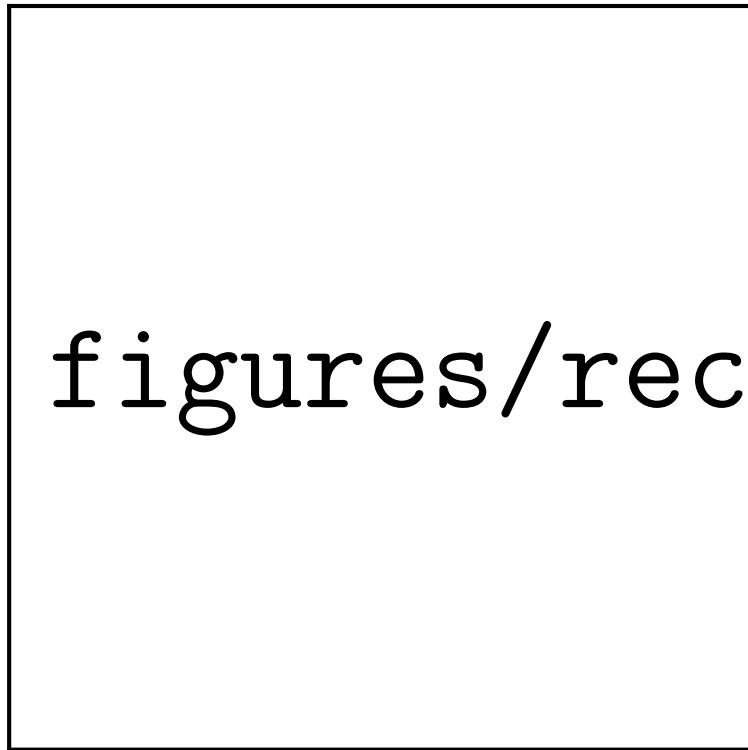


Figure 5.12: Schematic representation of the jet vertex fractions in the case of two jets and two primary vertices.

with a jet.

- **SV1:** Attempts to reconstruct a secondary vertex from the tracks associated with a jet. The most sensitive variable is the decay length significance between the secondary vertex and the primary event vertex. Additionally, the algorithm uses the invariant mass of tracks associated with the secondary vertex, the ratio of the energy of tracks assigned to the vertex to the energy of all tracks within the jet, and the number of two-track vertex candidates within the jet.
- **JetFitter:** Constructs a line connecting the primary vertex with one or more points associated with b - or c -hadron decays using a Kalman filter. The algorithm makes use of variables similar to the SV1 algorithm, along with the flight length significance between decays. Note that the algorithm does not require secondary vertices, allowing for the identification of decays with only one track.

The MV1 algorithm is trained on simulated data, using b -jets as signal and light quark jets as background, and returns a tag weight for each jet. The tag weight is used to establish working points with a given signal efficiency and background rejection power, as shown in figure ???. The model-independent triplepton analysis uses the working point with 80% signal

efficiency. The performance of the algorithm as a function of transverse momentum is shown in figure ??.

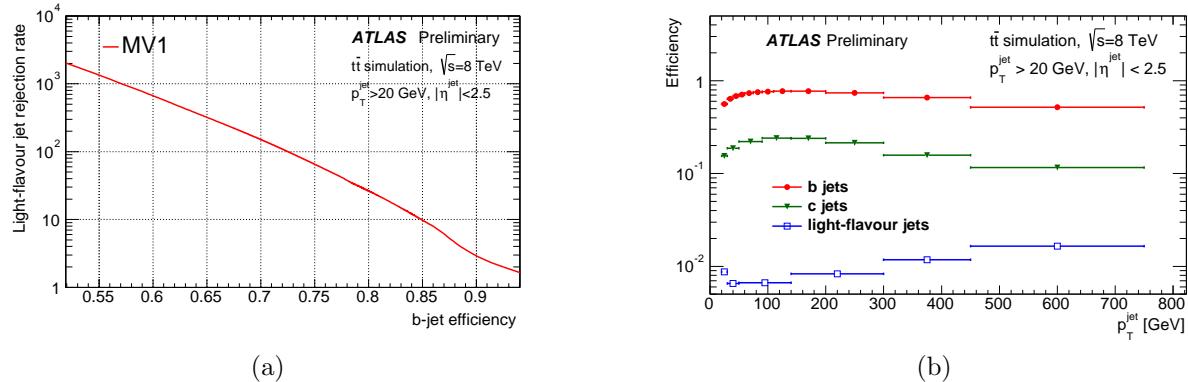


Figure 5.13: Performance of the MV1 b -tagging algorithm. The inclusive signal efficiency versus background rejection is shown at left. The signal and background efficiencies at the 70% working point are shown as a function of p_T at right.

5.6 Invisible Particles

Neutrinos interact only via the weak interaction, and hence escape the detector without interacting with any of the detector components. The same is true for any new stable, neutral, and colorless particles, such as the lightest supersymmetric particle in R parity-conserving scenarios or sterile neutrinos. The presence of such particles can only be inferred through the overall imbalance of the transverse momentum of the other visible collision products, E_T^{miss} ⁵.

The missing transverse momentum is defined as the negative vector sum of the visible objects in the event [[TheATLASCollaboration:2012jy](#), [TheATLASCollaboration:2013ua](#)]. The energies are mostly determined from calorimeter measurements, except for muons. To avoid using calorimeter energy deposits multiple times, the algorithm assigns energy to deposits to a single object according to a strict order: electrons are identified first, followed by photons, hadronically decaying τ leptons, jets, and muons. Depending on the analysis requirements, the energy of a given physics object can be determined using the full reconstruction algorithm, or the energy deposits can be calibrated to the EM or LCW scales. Finally, topological clusters and tracks not assigned to physics objects are included in E_T^{miss} calculation as the so-called soft term. The missing transverse energy is given by:

$$E_{x(y)}^{\text{miss}} = - \left(E_{x(y)}^e + E_{x(y)}^\gamma + E_{x(y)}^\tau + E_{x(y)}^{\text{jets}} + E_{x(y)}^\mu + E_{x(y)}^{\text{soft}} \right), \quad (5.10)$$

⁵The longitudinal momentum imbalance, E_z^{miss} , is not useful in pp collisions due to the unknown longitudinal momenta of the initial colliding partons.

where each term on the right represents the total momentum of the reconstructed objects in the x or y directions.

The measured $\mathbf{E}_T^{\text{miss}}$ receives contributions from sources besides invisible particles, including calorimeter noise, particles falling in insensitive regions of the detector, energy mismeasurements, and pileup interactions. These effects are mitigable to varying degrees. Particles falling outside the detector acceptance contribute irreducibly to the $\mathbf{E}_T^{\text{miss}}$ resolution. For physics objects like electrons or jets, the noise and pileup contributions are suppressed by the reconstruction algorithms and identification cuts. To mitigate the noise and pileup contributions to the soft term, the cells are grouped into topological clusters. Rejecting events where the $\mathbf{E}_T^{\text{miss}}$ is parallel or antiparallel to the physics objects can mitigate the impact of energy mismeasurements.

5.7 Object Selection

The analyses described in this dissertation search for energetic leptons and jets produced the decays of new heavy particles. The leptons and jets typically have large transverse momentum and are well-separated from other objects in the event, due to the large difference in the mass scale between the final decay products (leptons or hadrons with mass $m \lesssim 10 \text{ GeV}$) and the parents (W/Z bosons or the new particles themselves, with mass $m \gtrsim 80 \text{ GeV}$). The new particles typically have short lifetimes, and hence the leptons and jets are produced promptly at the location of the initial proton-proton interaction, consistent with originating from the selected primary vertex. These properties can be used to suppress backgrounds due to Standard Model processes and detector effects, described in chapter ??.

Leptons

The events used in this dissertation are required to have at least three reconstructed electrons, muons, or hadronically decaying τ leptons, as described in section ???. A summary of the lepton selections is shown in tables ?? and ???. Leptons are required to satisfy the following requirements:

- **Transverse momentum:** Electrons and muons must have $p_T > 15 \text{ GeV}$, while hadronically decaying τ leptons must have $p_T > 20 \text{ GeV}$. The transverse momentum cut is driven by the availability of triggers with which to perform the data-driven reducible background estimate, described in section ??.
- **Geometrical acceptance:** Electrons are required to have $|\eta| < 2.47$, excluding the transition region $1.37 < |\eta| < 1.52$ between the barrel and end-cap calorimeters. Muons and τ leptons are required to have $|\eta| < 2.5$.
- **Particle identification:** To suppress the reducible backgrounds, the leptons must satisfy strict requirements related to particle identification, as described in section ??.

Electrons candidates must satisfy the tight++ set of identification cuts. Electrons are neglected if they fall in a region affected by the presence of a dead front end board in the first or second sampling layer, a dead high voltage supply, or a masked cell in the core. Muons are required to be *combined*, with associated hits in the inner detector and muon spectrometer. Specifically, associated inner detector track must have:

- A B-layer hit, unless the muon passes through a deactivated region of the B-layer.
- ≥ 1 pixel hit and ≥ 5 SCT hits, including any deactivated sensors along the trajectory.
- ≤ 2 total missing hits in the pixel and SCT, excluding deactivated sensors along the trajectory.
- A successful extension into the TRT, with ≥ 6 TRT hits, of which $< 90\%$ are classified as outlier hits.

Finally, τ leptons must satisfy the **BDT-tight** selection criteria.

- **Impact parameter:** The inner detector track associated with electrons and muons must be consistent with originating from the event primary vertex. The transverse impact parameter significance, defined as the transverse impact parameter d_0 divided by its uncertainty σ_{d_0} , is required to satisfy $\frac{d_0}{\sigma_{d_0}} < 3$. Similarly, the longitudinal impact parameter z_0 is required to satisfy $z_0 \sin \theta < 0.5$ mm. These requirements suppress leptons from semileptonic heavy flavor decays.
- **Isolation:** To further reduce the impact of non-prompt and misidentified leptons, the leptons are required to be isolated from other activity in the event. The cuts on electrons and muons are similar, and limit the amount of nearby activity as measured by inner detector tracks and calorimeter energy deposits:

- For both electrons and muons, a cut is applied on `ptcone30`, the sum of transverse momenta of tracks associated to the same primary vertex as the lepton within a cone of $\Delta R < 0.3$.
- For muons, a cut is applied on `Etcone30`, the scalar sum of transverse energies of calorimeter cells within $\Delta R < 3.0$ of the muon track.
- For electrons, a cut is applied on `TopoEtcone30`, the sum of topological calorimeter clusters within a cone of $\Delta R < 3.0$. The use of topological clusters reduces the impact of pileup and out-of-cone leakage.

All isolation variables are required to be less than 10% of the lepton transverse momentum for leptons with $p_T < 100$ GeV, and less than $10 \text{ GeV} + 0.01 \times p_T$ for leptons with $p_T \geq 100$ GeV.

Cut	Electrons	Muons
Object ID	Tight++	Combined Tight
Leading (trigger) E_T/p_T	$E_T > 26 \text{ GeV}$	$p_T > 26 \text{ GeV}$
Subleading E_T/p_T	$E_T > 15 \text{ GeV}$	$p_T > 15 \text{ GeV}$
Trigger Acceptance	$(\eta < 2.47) \&\& !(1.37 < \eta < 1.52)$	$ \eta < 2.4$
Acceptance	$(\eta < 2.47) \&\& !(1.37 < \eta < 1.52)$	$ \eta < 2.5$
Calo. Isolation ($E_T/p_T < 100 \text{ GeV}$)	$\text{TopoEtcone30} < 0.1 \times E_T$	$\text{Etcone30} < 0.1 \times p_T$
Calo. Isolation ($E_T/p_T > 100 \text{ GeV}$)	$\text{TopoEtcone30} < 10 \text{ GeV} + 0.01 \times E_T$	$\text{Etcone30} < 10 \text{ GeV} + 0.01 \times p_T$
Track Isolation ($E_T/p_T < 100 \text{ GeV}$)	$\text{ptcone30} < 0.1 \times E_T$	$\text{ptcone30} < 0.1 \times p_T$
Track Isolation ($E_T/p_T > 100 \text{ GeV}$)	$\text{ptcone30} < 10 \text{ GeV} + 0.01 \times E_T$	$\text{ptcone30} < 10 \text{ GeV} + 0.01 \times p_T$
Track d_0	$\frac{d_0}{\sigma_{d_0}} < 3$	$\frac{d_0}{\sigma_{d_0}} < 3$
Track z_0	$z_0 \sin \theta < 0.5 \text{ mm}$	$z_0 \sin \theta < 0.5 \text{ mm}$

Table 5.2: Electron and muon selection criteria.

Cut	Taus	
Object ID	BDT	Tight
p_T	$p_T > 26 \text{ GeV}$	
Acceptance	$ \eta < 2.5$	

Table 5.3: Tau selection criteria.

Jets and Missing Transverse Energy

Jets are required to have $p_T > 30 \text{ GeV}$, in order to limit the presence of pileup jets. For the geometrical acceptance, jets must lie in the range $|\eta| < 4.5$, so that the jet falls within instrumented regions of the detector. Pileup jets are additionally suppressed with a cut on the jet vertex fraction: the $\sum p_T$ of tracks within the jet cone and associated with the selected primary vertex must be at least 50% of the $\sum p_T$ of all tracks within the jet cone [jvf].

Jets consistent with originating from the decay of a b -hadron are identified using the MV1 algorithm [MV1], with an efficiency of 80%.

The missing transverse momentum, E_T^{miss} , is calculated using the MET_Egamma10NoTau_RefFinal algorithm. Calorimeter cells associated with electrons or photons with $p_T > 10 \text{ GeV}$ are calibrated specifically to that object; cells associated with τ leptons are not calibrated as τ leptons due to the change in the energy calibration for the objects used in the data-driven reducible background estimate ??.

Overlap Removal

Objects are frequently reconstructed as multiple objects; for example, a muon with a hard bremsstrahlung emission might be reconstructed as a muon, an electron, and a jet. In order to resolve ambiguities, the following overlap removal procedure is applied:

- If $\Delta R(e, e) < 0.1$, remove lower p_T electron, to avoid “a potential bias in the simulation of the reconstruction efficiency for two real, close-by same-flavour leptons” [Adams:2014wx].

- If $\Delta R(e, \text{jet}) < 0.2$, remove jet. This addresses the ambiguity between electrons and jets.
- If $0.2 < \Delta R(\text{jet}, e) < 0.4$ AND $p_T(\text{jet}) > 30 \text{ GeV} + 0.05 * p_T(e)$, remove electron. This reduces the reducible electron backgrounds.
- If $\Delta R(\mu, e) < 0.1$, remove electron. This addresses cases where a muon radiates a hard photon, which is then identified as an electron.
- If $\Delta R(\mu, \text{jet}) < 0.1$, and:

$$\begin{aligned} p_T^{\text{jet}} < 0.5 p_T^\mu & : p_T^\mu < 200 \text{ GeV, or} \\ p_T^{\text{jet}} < 100 \text{ GeV} & : p_T^\mu \geq 200 \text{ GeV,} \end{aligned} \quad (5.11)$$

remove the jet. This is intended to reduce efficiency loss (in the next bullet point) from jets induced by muons at high muon p_T .

- If $\Delta R(\text{jet}, \mu) < 0.3$, remove muon. This reduces the reducible muon backgrounds.

Chapter 6

Background Estimation

The relevant Standard Model processes contributing to multilepton final states are diboson production (WZ , ZZ), production of a top quark pair in association with an weak gauge boson ($t\bar{t} + V$), and triboson production ($VVV^{(*)}$, where $V = W$ or Z). Examples tree-level Feynman diagrams of these processes are shown in figure ???. These backgrounds, called *prompt* backgrounds, are estimated using Monte Carlo (MC) simulation, as described in section ???. Significant backgrounds also arise from processes where at least one reconstructed lepton is due to the semileptonic decay of a hadron, the misidentification of a jet, or the asymmetric conversion of a photon in the detector; such backgrounds are called *reducible* backgrounds. These backgrounds are estimated using either MC simulation or a data-driven technique called the *fake factor* method, which is described in section ??.

6.1 Prompt Backgrounds

The prompt backgrounds are estimated using Monte Carlo simulation. The hard-scattering processes are modeled by dedicated event generators, possibly including the emission of additional partons. Additional QCD radiation is modeled using a parton shower. The detector response to the simulated events is simulated with the ATLAS simulation framework [**atlas-simulation-framework**] using the GEANT4 toolkit [**geant**]. Pileup is included by overlaying simulated minimum-bias interactions from PYTHIA on the hard scattering event. Simulated events are assigned weights to reproduce the observed pileup distributions in data, and also to account for small differences in the trigger, reconstruction, and identification efficiencies between simulation and data.

The details of the modeling of each sample are described below. The generator, parton shower, PDF set, underlying event tune, and accuracy of theoretical cross section for samples are summarized in table ??.

- SHERPA is used to model WW , WZ , and ZZ production. Both bosons in the events decay leptonically. Up to three jets are included in the matrix element. An important feature of SHERPA is that it accurately models the $W + \gamma^*$ and $Z + \gamma^*$ contributions

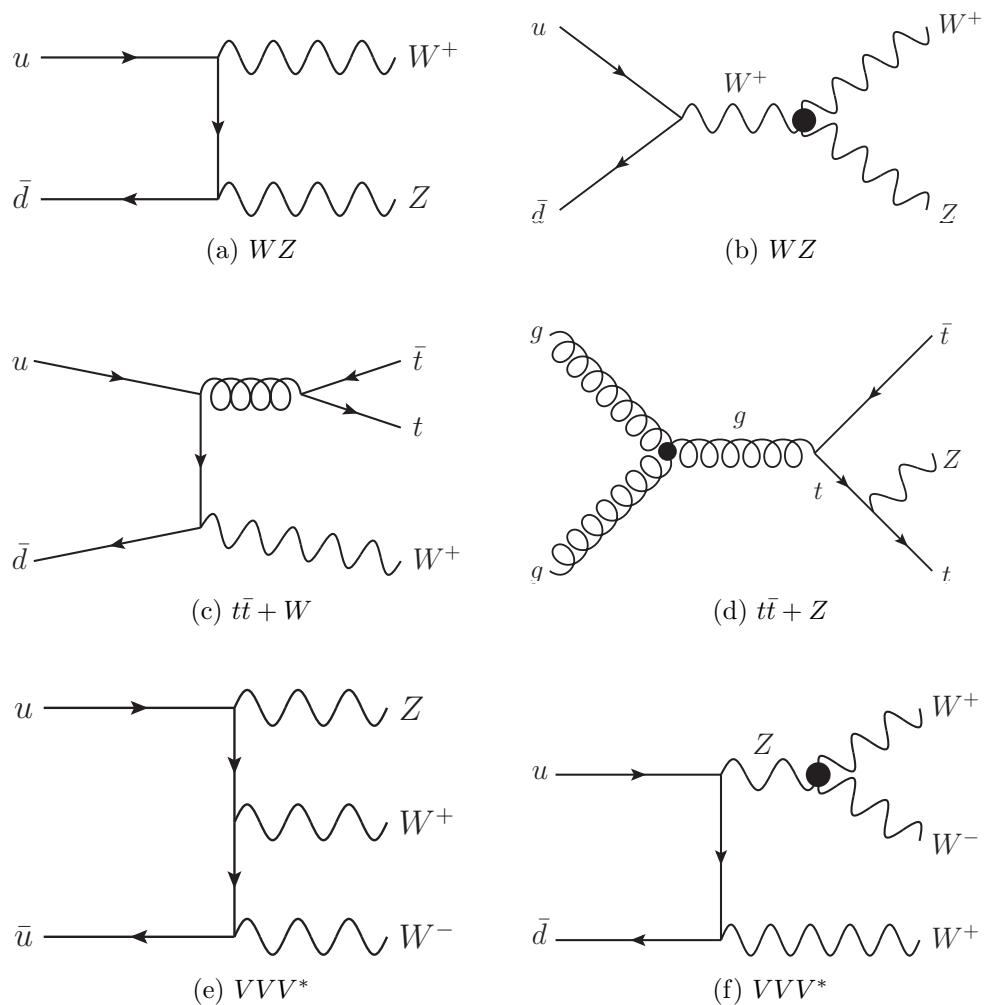


Figure 6.1: Example tree-level Feynman diagrams of Standard Model processes leading to trilepton final states.

Process	Generator	Parton shower and hadr.	PDF set	UE tune	Cross section
WZ	SHERPA1.4.3	SHERPA	CT10	SHERPA	NLO
ZZ	SHERPA1.4.5	SHERPA	CT10	SHERPA	NLO
$t\bar{t} + W/Z$	MADGRAPH 5.1.3.33	PYTHIA 6.426	CTEQ6L1	AUET2B	NLO
$VVV^{(*)}$	MADGRAPH 5.1.3.33	PYTHIA 6.426	CTEQ6L1	AUET2B	LO
$Z + \gamma$	SHERPA	SHERPA	CT10	SHERPA	LO

Table 6.1: Summary of the primary background MC samples used in this analysis. The generator, parton shower and hadronization, PDF, underlying event tune, and the order of the cross section calculation are shown for each sample.

down to very low γ^* masses; for electron decays, a cut of $m(ee) > 100$ MeV is applied, while for muon and tau decays, SHERPA naturally cuts off the divergence. To increase the statistics in the phase space relevant for this analysis, the WZ samples require at least two leptons have $p_T > 5$ GeV. Finally, the WZ sample also treats the b and c quarks as massive, which improves the modeling of heavy flavor jets at the cost of increased computation time.

- SHERPA is also used to model $Z + \gamma$ production, where the photon converts asymmetrically in the detector and is reconstructed as an electron. The events are reweighted to account for observed mismodeling in the conversion rate, described in section ??.
- $t\bar{t} + V$ production is modeled with MADGRAPH, with PYTHIA for the parton shower.
- $WWW^{(*)}$, $ZWW^{(*)}$, and $ZZZ^{(*)}$ are modeled using MADGRAPH, with PYTHIA for the parton shower. Their contributions to all the signal regions are negligible.

6.2 Reducible Backgrounds

The reducible backgrounds encompass a variety of processes in which a reconstructed lepton arises due to something away from the “hard scatter” interaction. The backgrounds are estimated using simulation or a data-driven technique, depending on the source. The contribution from $Z + \gamma$, where the photon converts asymmetrically and is reconstructed as an electron, is estimated from Monte Carlo, with scale factors applied to account for an observed overestimation in the photon-to-electron conversion rate (see section ??). Other reducible contributions, such as $W/Z + \text{jets}$ or $t\bar{t}$ where a jet fakes an electron or decays semileptonically, are estimated using the fake factor method as implemented in the 8 TeV model-independent multilepton analysis [DeViveiros:201670929].

Fake Factor Method

The fake factor method estimates the reducible backgrounds in each signal region by characterizing the fake or non-prompt leptons in terms of quantities sensitive to the fake process, such as isolation, impact parameter, or particle identification cuts. Two orthogonal sets of leptons are defined: *numerator* leptons (N) satisfy the nominal signal lepton selection criteria, while *denominator* leptons (D) satisfy most of the nominal selection criteria, except with inverted requirements on quantities sensitive to the fake process. The reducible background is estimated in a data-driven way from a control region consisting of events with a mix of numerator and denominator leptons, together with a parametrization of the relationship between numerators and denominators.

The relationship between numerator and denominator objects is called the *fake factor*, f , defined as the ratio of the number of fake leptons satisfying the numerator criteria to those satisfying the denominator criteria. The fake factor is measured in a control region enriched in fake leptons. The success of the method depends largely on the extrapolation of the f from the measurement control region to the signal regions. To capture the dependence on the event kinematics, f can be measured as a function of various lepton or event variables; in this dissertation, the fake factors are all measured in bins of lepton p_T and η .

Once f has been measured, the reducible background is determined as follows. In signal events with three or more leptons, any subset of the leptons could be real or fake. For example, an event might contain two real leptons from a Drell-Yan process plus a fake third lepton. Such an event is labeled $\ell_1^R \ell_2^R \ell_3^F$, indicating the “truth” classification of the three leptons as either real (R) or fake (F) (the ordering of the letters corresponds to some canonical order of the leptons, such as p_T ordering). On the other hand, if an event contained, say, $W \rightarrow l\nu$ plus two fake leptons, the event would be labeled $\ell_1^R \ell_2^F \ell_3^F$ (or $\ell_1^F \ell_2^R \ell_3^F$ or $\ell_1^F \ell_2^F \ell_3^R$).

The quantity we desire to determine is the number of events containing three real leptons, $n_{\ell_1^R \ell_2^R \ell_3^R}$. The quantity actually measured in a signal region is the number of events containing three numerator objects, $n_{\ell_1^N \ell_2^N \ell_3^N}$. Any of these numerators could be real or fake, so the sample can be decomposed as:

$$n_{\ell_1^N \ell_2^N \ell_3^N} = n_{\ell_1^R \ell_2^R \ell_3^R} + n_{\ell_1^R \ell_2^R \ell_3^F} + n_{\ell_1^R \ell_2^F \ell_3^R} + n_{\ell_1^R \ell_2^F \ell_3^F} + \dots \quad (6.1)$$

$$+ n_{\ell_1^F \ell_2^R \ell_3^R} + n_{\ell_1^F \ell_2^R \ell_3^F} + n_{\ell_1^F \ell_2^F \ell_3^R} + n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.2)$$

The reducible background prediction is $n_{\ell_1^N \ell_2^N \ell_3^N} - n_{\ell_1^R \ell_2^R \ell_3^R}$, the number of signal events where at least one lepton is fake. To determine the other terms, we use events with one or more denominator leptons. For example, consider $\ell_1^D \ell_2^N \ell_3^N$ events, where the first lepton is a denominator and the remaining leptons are numerators. Assuming that the denominator lepton is always a fake lepton, the number of $\ell_1^D \ell_2^N \ell_3^N$ events, each weighted by the fake factor f corresponding to the denominator lepton (represented schematically by $n_{\ell_1^D \ell_2^N \ell_3^N} f_1$), equals the number of $\ell_1^N \ell_2^N \ell_3^N$ events where the first lepton is fake:

$$n_{\ell_1^D \ell_2^N \ell_3^N} f_1 = n_{\ell_1^F \ell_2^R \ell_3^R} + n_{\ell_1^F \ell_2^R \ell_3^F} + n_{\ell_1^F \ell_2^F \ell_3^R} + n_{\ell_1^F \ell_2^F \ell_3^F}. \quad (6.3)$$

Similarly, the remaining permutations of numerators and denominators yield:

$$n_{\ell_1^N \ell_2^D \ell_3^N} f_2 = n_{\ell_1^R \ell_2^F \ell_3^R} + n_{\ell_1^R \ell_2^F \ell_3^F} + n_{\ell_1^F \ell_2^R \ell_3^R} + n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.4)$$

$$n_{\ell_1^N \ell_2^N \ell_3^D} f_3 = n_{\ell_1^R \ell_2^R \ell_3^F} + n_{\ell_1^R \ell_2^F \ell_3^F} + n_{\ell_1^F \ell_2^R \ell_3^F} + n_{\ell_1^F \ell_2^F \ell_3^R} \quad (6.5)$$

$$n_{\ell_1^D \ell_2^D \ell_3^N} f_1 f_2 = n_{\ell_1^F \ell_2^F \ell_3^R} + n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.6)$$

$$n_{\ell_1^D \ell_2^N \ell_3^D} f_1 f_3 = n_{\ell_1^F \ell_2^R \ell_3^F} + n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.7)$$

$$n_{\ell_1^N \ell_2^D \ell_3^D} f_2 f_3 = n_{\ell_1^R \ell_2^F \ell_3^F} + n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.8)$$

$$n_{\ell_1^D \ell_2^D \ell_3^D} f_1 f_2 f_3 = n_{\ell_1^F \ell_2^F \ell_3^F} \quad (6.9)$$

These equations contain eight equations and eight unknowns, so the system can be solved for the reducible background prediction:

$$\ell_1^N \ell_2^N \ell_3^N - \ell_1^R \ell_2^R \ell_3^R = (\ell_1^N \ell_2^N \ell_3^D f_3 + \ell_1^N \ell_2^D \ell_3^N f_2 + \ell_1^D \ell_2^N \ell_3^N f_1) \quad (6.10)$$

$$- (\ell_1^N \ell_2^D \ell_3^D f_2 f_3 + \ell_1^D \ell_2^N \ell_3^D f_1 f_3 + \ell_1^D \ell_2^D \ell_3^N f_1 f_2) \quad (6.11)$$

$$+ \ell_1^D \ell_2^D \ell_3^D f_1 f_2 f_3 \quad (6.12)$$

Note that throughout this method, we have assumed that the leptons used for the measurement of f and the denominator leptons in trilepton events are always fake leptons. In practice, real leptons contaminate both of these samples, and is accounted for using simulation, where the lepton can be classified as real or fake using the truth record of the event.

We now discuss the measurement of the fake factors f , separately for electrons, muons, and tau leptons.

Electron Fake Factors

The background estimation for fake electrons targets the reducible contribution from two sources: semileptonic heavy flavor decays and misidentified light hadrons. The electron denominator objects are defined by inverting either the IsEM electron identification criteria¹ or the track impact parameter cut (exclusive OR), as shown in table ??; the electron must pass all the remaining cuts listed in table ???. Additionally, an inefficiency is observed for loose offline electrons with respect to the loose electron triggers. This is due to the lack of

¹The parameter space between medium++ and loose++, rather than tight++ and medium++, is used for two reasons. First, for the model-independent trilepton analysis [DeViveiros:2016zq], electrons between medium++ and loose++ were used to define a validation region to test the fake factor method. Second, requiring the electrons to fail medium++ reduces the prompt contamination, which is nonetheless still quite large.

Gaussian sum filter (GSF) tracking [ATLAS-CONF-2012-047] at the trigger level. The inefficiency is mitigated by imposing the `tight++` requirement on the $\Delta\eta$ and $\Delta\phi$ between the electron track and associated calorimeter cluster, with the goal of cutting out electrons with large amounts of bremsstrahlung, whose tracks are not reconstructed by the non-GSF algorithm in the trigger.

Criteria	Numerator	Denominator
IsEM ID	<code>tight++</code>	<code>!medium++ && loose++</code>
Impact Parameter Significance	$\frac{ d_0 }{\sigma_{d_0}} < 3$	$3 < \frac{ d_0 }{\sigma_{d_0}} < 10$

Table 6.2: Electron denominator definitions. The denominators are taken to be an exclusive OR combination of the two selection inversions. Additionally, denominator objects must pass the tight requirement on the $\Delta\eta$ and $\Delta\phi$ between the track and the cluster.

The fake factors are measured in a control sample of single-electron events, using the entire 20.3 fb^{-1} 2012 dataset. The triggers used to collect events are listed in table ??; photon triggers are used where available ($p_T > 24 \text{ GeV}$), and loose electron triggers are used otherwise ($15 \text{ GeV} < p_T < 24 \text{ GeV}$).

p_T range (GeV)	Trigger	Average 2012 Prescale
15-17	EF_e5_loose0	56080.5
17-24	EF_e15vh_loose0	1549.7
24-45	EF_g20_loose	4412.6
45-65	EF_g40_loose	348.3
65-85	EF_g60_loose	80.9
85-105	EF_g80_loose	28.5
105-125	EF_g100_loose	13.0
125-210	EF_g120_loose	1.0
>210	EF_g200_etcut	1.0

Table 6.3: Trigger used to collect electron numerator and denominator objects.

Events are required to have $m_T < 40 \text{ GeV}$ and $E_T^{\text{miss}} < 40 \text{ GeV}$ to suppress contamination from single- W production, where m_T is the transverse mass of the electron and missing transverse energy in the event. The events are also required to have exactly one electron, numerator or denominator, in order to suppress prompt contamination from $Z \rightarrow ll$ events. The electrons are required to be trigger-matched to the trigger used to collect the event in the relevant p_T range. The residual prompt contamination, comprised mostly of W and Z events with smaller contributions from Drell-Yan, $t\bar{t}$ and single- t , is subtracted using Monte Carlo. The numerator and denominator event yields, as well as the predicted prompt contamination, are shown in figure ???. The prompt contamination consists primarily of W and Z events. The relative size of the prompt contamination is quite large for numerator

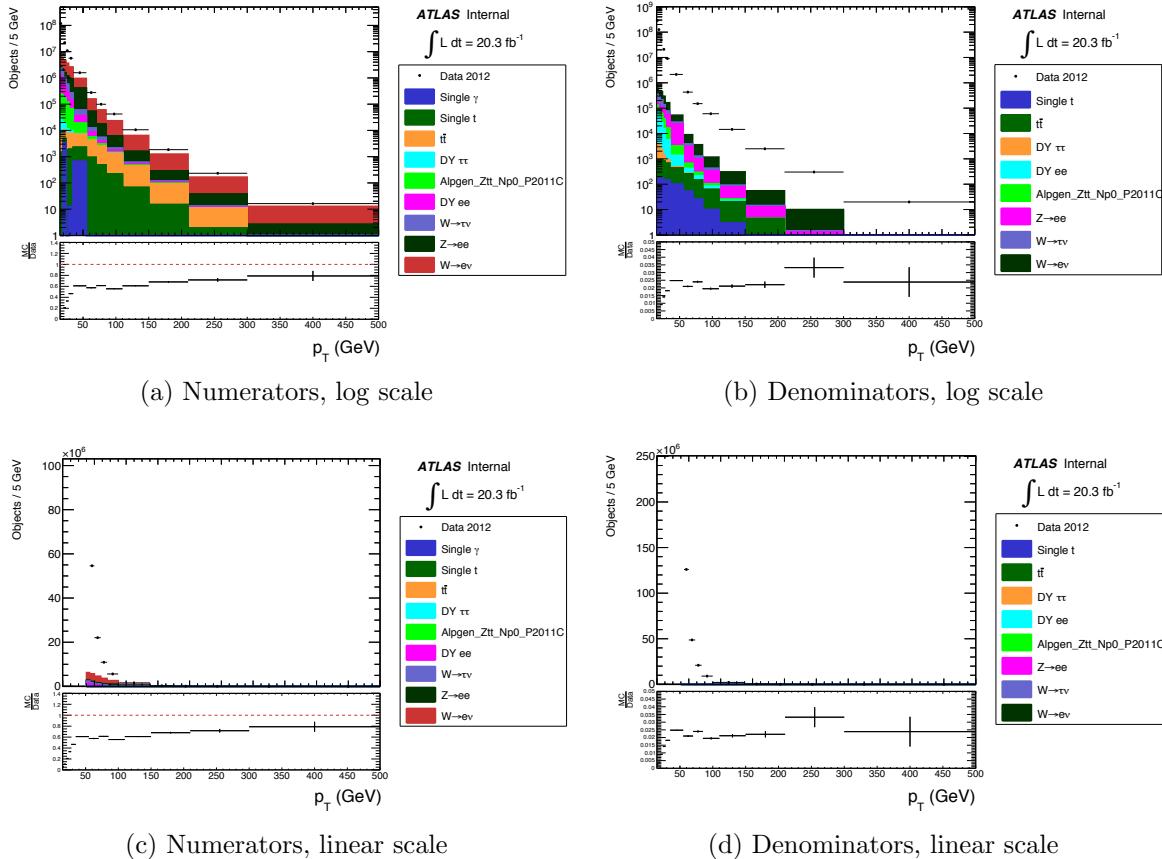


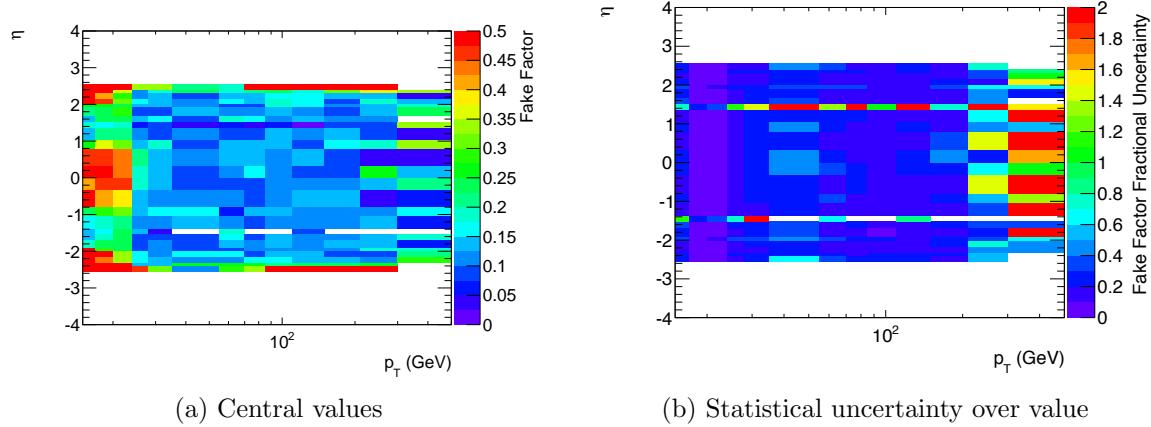
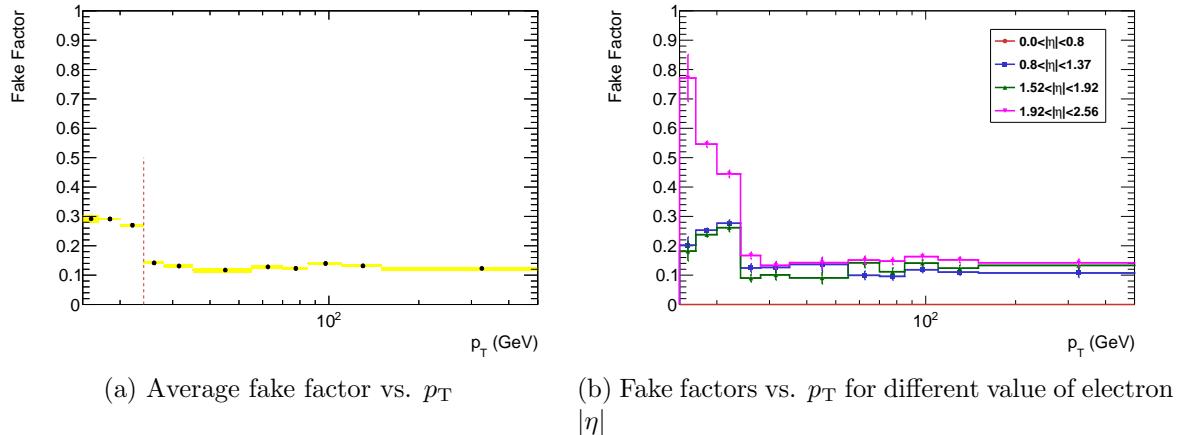
Figure 6.2: Numerator and denominator electron object counts. The data sample consists of all single-electron events in the 2012 dataset, with cuts to reduce prompt contamination as described in the text. The markers represent object counts from 2012 data, and the colored histograms indicate the prompt subtractions estimated from Monte Carlo.

objects, reaching up to $\sim 60\%$ for $p_T \sim 50$ GeV, despite the cuts intended to reduce the W contribution. The fake factors are binned two-dimensionally in p_T and η , shown in figure ???. The p_T dependence of the fake factors is shown in figure ??, for the inclusive sample (left) and for various $|\eta|$ ranges (right).

To help clarify the origin of the structure in η observed at low p_T , the numerator and denominator counts are also shown versus η for $p_T < 24$ GeV in figure ???. The numerator counts are relative flat versus η in the central region, and grow for $|\eta| \gtrsim 2$. Conversely, the denominators exhibit a significant increase near the barrel-endcap overlap region, and also for $|\eta| \gtrsim 2$.

The following sources of systematic uncertainty are considered. Again, we refer to [DeViveiros:167092] for more details.

- **Prompt subtraction:** The presence of real, prompt leptons from Standard Model

Figure 6.3: Electron fake factors parametrized in p_T and η .Figure 6.4: Electron fake factors projected in p_T . The denominator requirements are different below and above 24 GeV, where the triggers switch from EF_eXX to EF_gXX; below, the additional requirements on $\Delta\eta_1$ and $\Delta\phi_2$ cause a large drop in the denominators, and a large increase in the fake factor values.

processes in the sample used to measure the fake factors is accounted for using Monte Carlo simulation. Uncertainties on the simulated samples include luminosity; cross section uncertainties; and reconstruction, trigger, and identification efficiency scale factors. These lead to a maximum uncertainty of about 20% where the prompt subtraction is largest.

- **Trigger efficiency correction:** As mentioned previously, an inefficiency is observed in the loose electron triggers for offline `loose++` electrons. This is due to the lack of GSF tracking in the trigger. For the fake factor derivation, this affects electrons in the

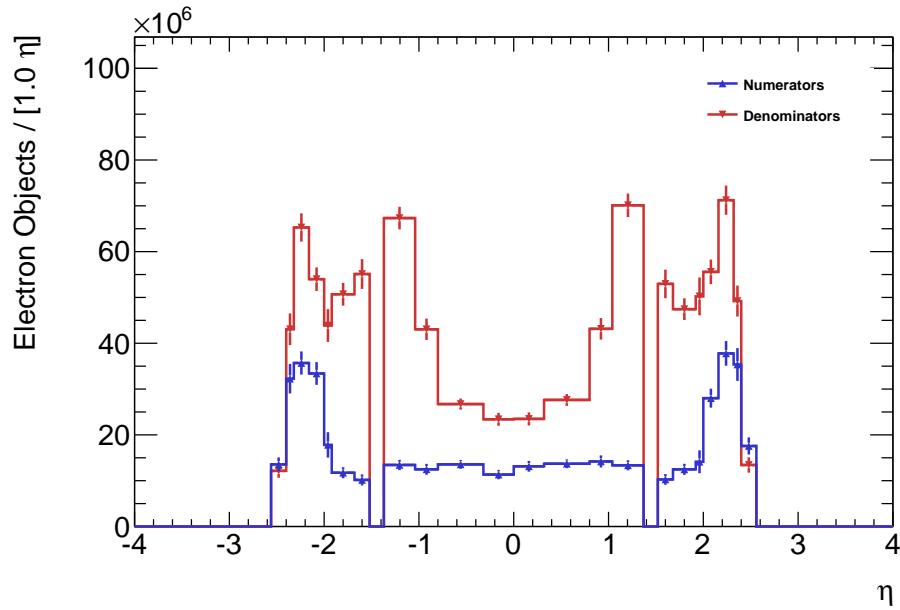


Figure 6.5: Electron numerator and denominator object counts versus η for $p_T < 24$ GeV.

range $15 \text{ GeV} < p_T < 24 \text{ GeV}$, where photon triggers are not available. Imposing the `tight++` cut on the track-cluster matching (the $\Delta\eta$ and $\Delta\phi$ between the electron track and calorimeter cluster) mitigates most, but not all, of the inefficiency, by cutting out electrons with large amounts of bremsstrahlung whose track are not reconstructed in the trigger. Based on a comparison of loose electron and photon triggers in the range $24 \text{ GeV} < p_T < 85 \text{ GeV}$, a correction of about 8% is applied to loose electron-triggered events, and the same value is taken as systematic uncertainty.

- **Extrapolation to signal region:** Two systematic uncertainties are assigned to account for bias due to the extrapolation of fake factors from the control region to the signal region. First, the cuts on m_T and E_T^{miss} are varied from $< 40 \text{ GeV}$ to $< 25 \text{ GeV}$ and $< 55 \text{ GeV}$. A p_T -dependent systematic uncertainty of up to 15% is assigned. Second, Monte Carlo-based truth studies indicate that the fake factor values are quite different for heavy- and light-flavor jets, so a difference in heavy flavor fraction between the control and signal regions will bias the fake factors. The effect of this is estimated using a $t\bar{t}$ Monte Carlo sample, and a flat systematic uncertainty of 20% is assigned. See appendix ?? for more information.

The systematic and total uncertainties on the fake factors are shown as a function of p_T in figure ??.

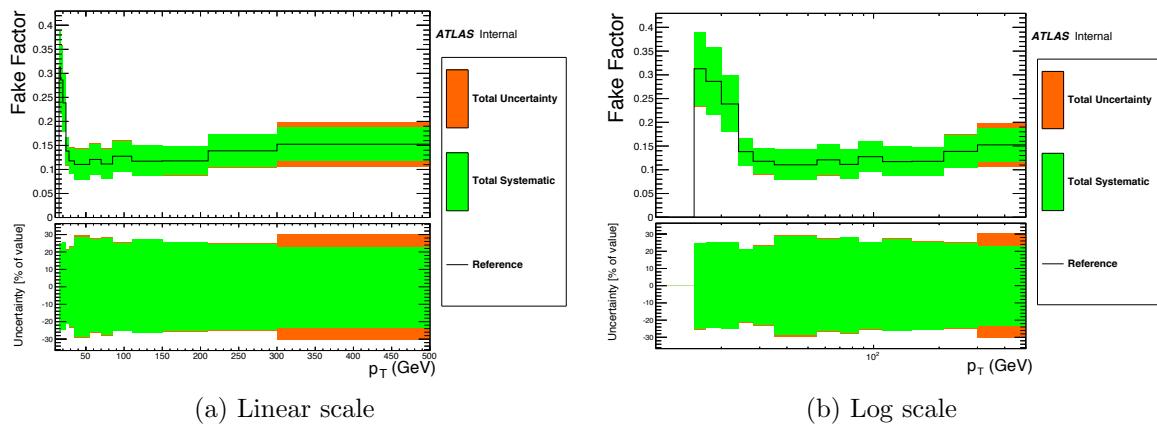


Figure 6.6: Electron fake factors vs. p_T with systematic and total uncertainties. The statistical uncertainty includes both the data and prompt subtraction Monte Carlo statistics.

Muon Fake Factors

The muon fake factor method follows that used in the same-sign dilepton search in 2011 [[LouiseThesis](#), [SSInternal](#)]. The method targets non-prompt muons from semileptonic heavy flavor decays, punch-through, and decays-in-flight of long-lived mesons by inverting the isolation requirements. Specifically, the denominator muons are defined as follows:

- Pass all numerator muon requirements in table ??, except the requirements on `Etcone30`, `ptcone30`, and $\frac{d_0}{\sigma_{d_0}}$.
- Loosen impact parameter cut:

$$\left| \frac{d_0}{\sigma_{d_0}} \right| < 10 \quad (6.13)$$

- Invert isolation:

$$\text{Etcone30, ptcone30} > \begin{cases} 0.15p_T & : p_T < 100 \text{ GeV} \\ 15 + 0.01p_T \text{ GeV} & : p_T > 100 \text{ GeV} \end{cases} \quad (6.14)$$

$$\frac{\text{Etcone30}}{p_T} < 2.0 \quad (6.15)$$

$$\frac{\text{ptcone30}}{p_T} < 2.0 \quad (6.16)$$

$$(6.17)$$

- If $p_T < 40$ GeV, apply the same overlap requirement as the signal regions, removing the muon if $\Delta R(\mu, \text{jet}) < 0.3$. This overlap requirement is not applied for muons with $p_T > 40$ GeV, which increases the statistical precision at the expense of additional systematic uncertainty. This is denoted by “dR” or “non-dR” below, for example in figure ??.

The muon fake factors are measured in a same-sign dimuon sample, triggered by the `EF_2mu13` trigger. The use of same-sign muons suppresses the prompt contamination from Z/γ^* events. The measurement uses only muons with large track impact parameter significance, $\left| \frac{d_0}{\sigma_{d_0}} \right| > 3$, to obtain a sample enriched in non-prompt muons (if both muons satisfy this requirement, then both are counted in the measurement). An extrapolation factor is derived from Monte Carlo to account for the fact that the signal region requires $\left| \frac{d_0}{\sigma_{d_0}} \right| < 3$, as detailed below.

Two sets of fake factors are measured, depending on the jet activity in the event. In the following, jets are required to have $p_T > 30$ GeV, and be separated from muons with $\Delta R(\mu, \text{jet}) > 0.3$.

- **Inclusive:** Applied to events with zero jets. The measurement uses the entire same-sign dimuon sample.

- **Two-Jet:** Applied to events with one or more jet. The measurement uses same-sign dimuon events with at least two jets with $p_T > 30$ GeV.

Fake muons from the two-jet sample are expected to come primarily from $W + \text{jets}$ and $t\bar{t}$ processes, while the inclusive sample also includes contributions from $b\bar{b}$. Figure ?? shows the p_T distributions of numerator and denominator muons in the measurement sample along with the expected prompt contributions.

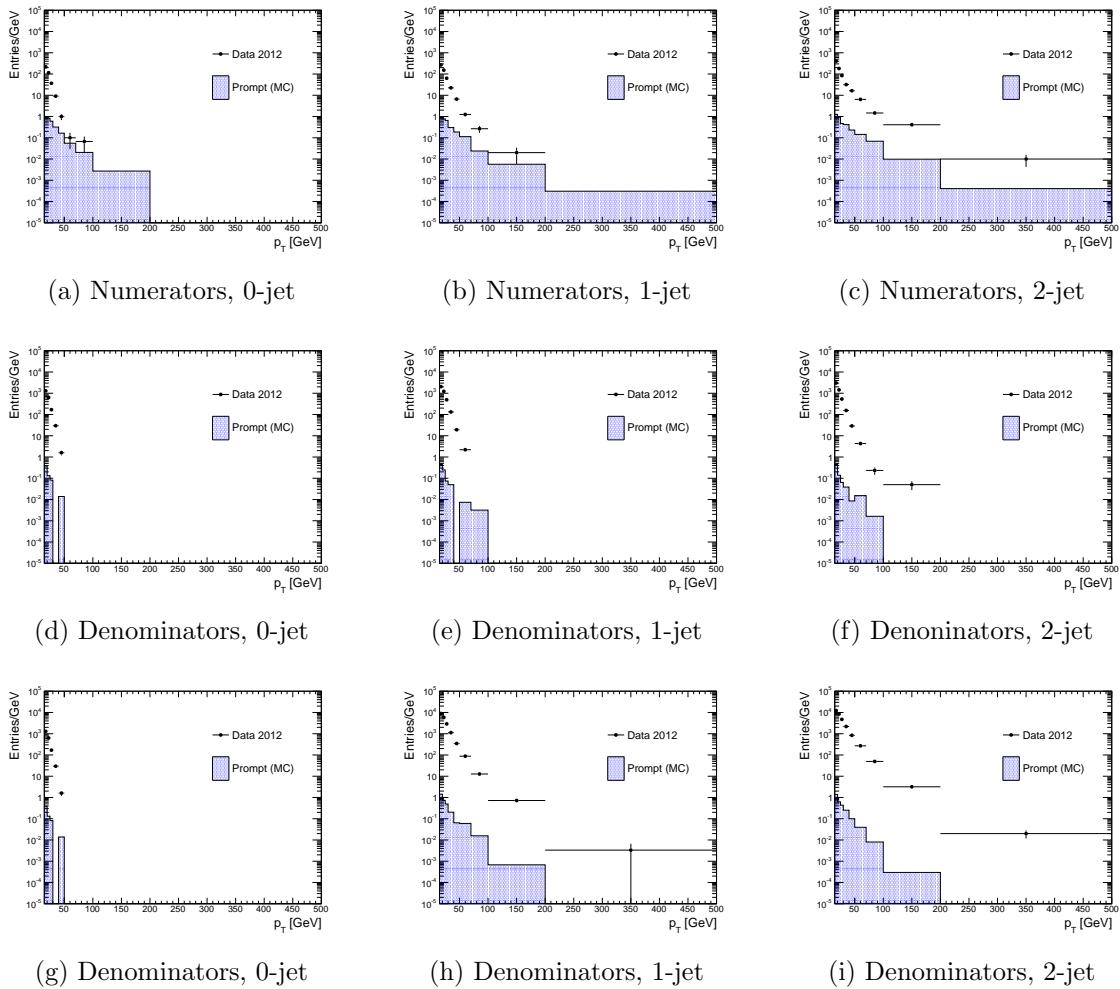


Figure 6.7: p_T spectrum of muons used in fake factor measurement. The left plots show events with zero jets, the middle plots show events with one jet, and the right plots show events with two or more jets. All numerator events require both numerator and denominator be separated from a jet by $\Delta R > 0.3$; the same requirement is applied to denominators in the second row of plots, while the third row shows denominators that are not required to be isolated from nearby jets.

The extrapolation factor from the measurement control region, with $|\frac{d_0}{\sigma_{d_0}}| > 3$, to the signal region, with $|\frac{d_0}{\sigma_{d_0}}| < 3$, is derived from various Monte Carlo samples. The extrapolation factor is simply the ratio of fake factors derived in Monte Carlo using the control region cut ($|\frac{d_0}{\sigma_{d_0}}| > 3$) to those using the signal region cut ($|\frac{d_0}{\sigma_{d_0}}| < 3$). The central value is taken from the POWHEG $t\bar{t}$ sample, using all truth-level non-prompt muons, as shown in figure ???. A systematic uncertainty is assigned by comparing with other samples (MC@NLO $t\bar{t}$ and PYTHIA $b\bar{b}$ and $c\bar{c}$), and using only same-sign dimuon events in these samples.

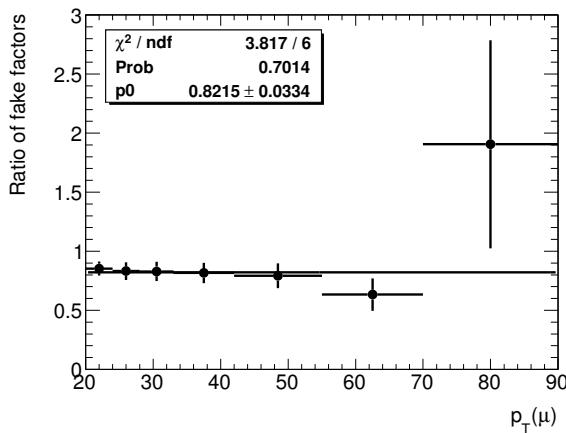


Figure 6.8: Ratio between the fake factors for muons with $|d_0|/\sigma(d_0) > 3$ compared to fake factor with nominal numerator and denominator definitions. These fake factors are derived from a POWHEG $t\bar{t}$ sample.

The fake factors are parametrized one-dimensionally in p_T and η , as there are insufficient statistics to do a full two-dimensional parametrization. The fake factor is computed as:

$$f(p_T, \eta) = \frac{f(p_T) \times f(\eta)}{\langle f \rangle} \quad (6.18)$$

where $\langle f \rangle$ is the total average fake factor. The measured fake factors are shown in figures ?? and ??.

The sources of systematic uncertainty considered are listed below, and the fractional systematic uncertainty is shown in figure ??.

- **Prompt subtraction:** The normalization of the simulated prompt subtraction samples is varied by $\pm 10\%$, leading to a systematic uncertainty of 1-6%.
- **Topological dependence:** The difference between the inclusive and two-jet fake factors is taken as a systematic uncertainty. The uncertainty is symmetrized, using the full difference as both upward and downward uncertainty, and ranges from 3% to 36%.

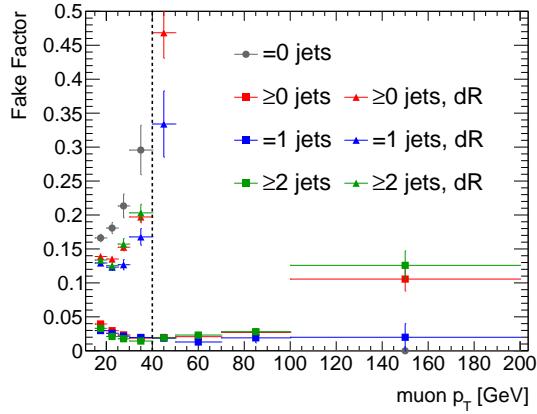


Figure 6.9: Muon fake factors as a function of p_T . Two sets of fake factors are plotted, with and without the “dR” requirement. The vertical dashed line at 40 GeV indicates the point at which the fake factors switch from the “dR” points, where the denominators are required to be separated from nearby jets, to the lower, non-”dR” points where the jet-isolation requirement is dropped to improve the statistics.

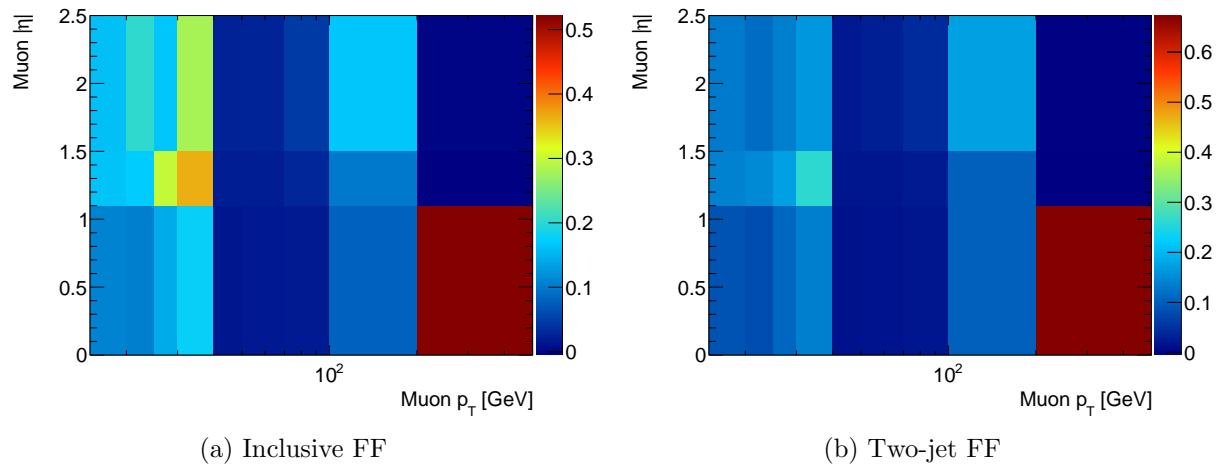


Figure 6.10: Muon fake factors as functions of p_T and $|\eta|$. The left plot shows fake factors measured in the inclusive control sample and applied to events with zero jets. The right plot shows fake factors measured in events with two jets, and applied to events with at least one jet.

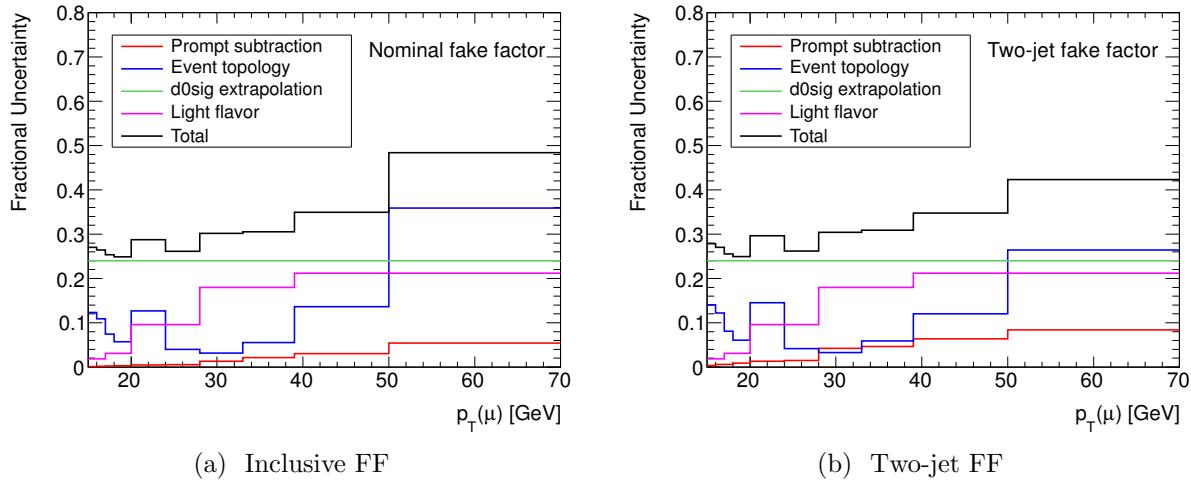


Figure 6.11: Systematic uncertainties on muon fake factor as a function of $p_T(\mu)$. The left plot shows the uncertainties for the inclusive fake factor, while the right shows the uncertainty for the two-jet fake factor.

- **Dependence on d_0 significance:** As mentioned previously, the extrapolation factor is derived in a number of different Monte Carlo samples. The largest deviation of 24% is taken as a systematic uncertainty.
 - **Light flavor fraction:** As with the electron fake factors, the fake factor values are quite different for muons originating from light flavor (LF) sources (π/K decay or punch-through) versus heavy flavor (HF) decays. The systematic uncertainty is described in detail in [DeViveiros:1670929]; in brief, it uses the difference in momenta measured by the inner detector and muon spectrometer as a discriminant between HF and LF fakes, estimates the difference in HF/LF fraction between the control and signal regions, and finally uses HF and LF fake factors measured in Monte Carlo to estimate the effect of the discrepancy in HF/LF fraction. A systematic uncertainty of 2% to 21% is assigned.

6.3 Application of Fake Factors

Figure ?? shows the fake factor-weighted reducible events in the signal regions, as given by equation ???. The estimated prompt subtractions due to $Z + \gamma$, WZ , ZZ , $t\bar{t} + V$, and $VVV^{(*)}$ events (i.e. the contributions to events with denominator objects due to these irreducible background sources, rather than the intended sources of reducible background) are shown in the colored histograms. The reducible background prediction is the difference between the data points and the prompt subtraction histograms. In cases where a bin is negative after ap-

plying the prompt subtraction, the contents of that bin are set to zero, and the histogram is scaled to preserve the overall normalization. For the 4L, $Z + e$ signal region, the overall normalization is negative and consistent with zero, so the background is neglected in this region.

The uncertainty on the background prediction is the combination of two components: the systematic uncertainties on the electron and muon fake factors, and the statistical uncertainties due to finite events with denominator objects. As described in section ??, the fake factor uncertainties constitute their own component, while the statistical uncertainties are contained in the “Monte Carlo statistics” component.

Tau Lepton Fakes

The detector signature of hadronically decaying tau leptons, consisting of a jet with one or more associated tracks, is not as distinctive as the signature of electrons and muons. Differentiating between jet, which are copiously produced in proton-proton collisions, and hadronically decaying tau leptons is difficult, and accordingly the reducible backgrounds in the 2L+ τ_{had} categories tend to be much larger than those in the 3L categories.

Lots of work to do here, to avoid copying and pasting PO’s work!

As taus can decay hadronically, their detector signature is not as distinctive as the one associated with electrons and muons. Therefore, jets, which are produced in abundance in proton-proton collisions, are the main source of tau fakes, leading to large expected fake-rates. As the tau identification algorithms use multi-variate discriminants based on calorimeter shower shapes and tracking information, the fake efficiency is strongly dependent on the jet fragmentation. This in turn introduces a dependency of the fake-rates on the parton initiating the jet (gluon, light-quark or heavy-quark).

Consequently, the fake-factors are less sensitive to the jet flavour if a tighter efficiency working point is chosen for the denominator (as some shower shape requirements are already applied). However, it is important to still choose a working point that is loose enough so that the denominator statistics are sufficiently large. A study of the flavour-dependency of the fake-factors was performed in simulation samples. For each jet faking a τ , the largest p_T truth-particle present inside of the τ reconstruction cone was taken as an estimator of the parton initiating the jet. Different denominator selections were chosen as a function of the existing loose identification working point. Denominators failing the **BDT-Medium** selection while having a **BDTScore** larger than 0.7, 0.8, 0.9 or $1.0 \times \text{BDTScore}_{\text{Loose}}$ were all studied (where **BDTScore**_{Loose} is the p_T dependent cut on the **BDTScore** used by the τ identification algorithm). Illustratively, each sequential cut tightening denominator score reduces the statistics by a factor of roughly 2. In the end, the minimal threshold of $0.9 \times \text{BDTScore}_{\text{Loose}}$ was chosen as it offered a greatly reduced sensitivity to the flavour of the initiating parton in the jet, while still offering large enough statistics for the sample. Further tightening of the cut provided little improvement in flavour dependence. Using a p_T -dependent denominator requirement ensures that the prompt efficiency is the same for all denominator taus, while lowering the fake-factor dependent on p_T .

The fake-factors are nominally parameterized as a function of the p_T and $|\eta|$ of the tau candidate, binned in 2 dimensions to preserve correlations. An additional scaling correction is applied as a function of the highest MV1 b -tagging weight of any jet in the event. The former two target dependencies associated with the jet shower shapes as a function of kinematics and detector homogeneity. The latter is taken to account for the heavy-flavour component.

Fake Factor Measurement

It is also important to measure the fake-factors in a sample which is expected to have a similar flavour composition as the signal regions used (the majority of trilepton events with a fake tau are expected to come from electroweak production with an associated jet faking the tau, for example $Z + \text{jets}$ events). This is particularly important as no observable can be used to directly and accurately classify each fake taus as potentially gluon- or quark-initiated.

While the tau fake-factors were measured in the past from a $\gamma + \text{jets}$ sample, the approach chosen for this iteration of the analysis is different. The low p_T photon triggers were heavily prescaled during the 2012 run, limiting the statistics available to measure fake-factors at low p_T values. Instead, a $W + \text{jets}$ sample is used due to its high statistics and similar final-state kinematics. A leading numerator muon is required to be reconstructed as a tag (along with the associated trigger requirement), and the fake-factors are then measured from the probe tau candidates. No other requirements are made on the events. Alternatively, selection requirements could be made to reject the main source of τ_{had} prompts ($Z \rightarrow \tau_\mu \tau_{\text{had}}$), but such requirements can potentially introduce a bias on the fake-factors (for example, introducing a same-sign requirement is expected to drastically reduce the quark-initiated jet contribution). Instead, the contribution from prompt τ_{had} is subtracted directly from Monte Carlo simulations. The two non-negligible contributions come from $Z + \text{jets}$ and $t\bar{t}$ events.

Figure ?? shows sample data distributions for the numerator and denominator tau candidates, along with the prompt contribution which must be subtracted when calculating the fake-factors. Figure ?? shows the obtained fake-factors.

Systematic Uncertainties

Various studies have been carried out to estimate the systematic uncertainties associated with the tau fake estimates. The main considered sources of uncertainties are the following:

- Uncertainties on the MC prompt estimates
- Uncertainties associated with the binning choice
- Flavour dependence of the fake-factors (gluon, light-quarks, heavy-quarks)

For the prompt estimates, the normalization of the MC samples have been fluctuated by their associated uncertainties (as listed in ??). This results in variations between 5%

and 17%, being the largest in the 30-40 GeV p_T range (where the prompt subtraction is the largest due to the Z sample kinematics).

To estimate the effect of the binning on the fake estimates, a closure test is used. The fake-factors are re-applied to the same sample they were derived from. The results of this test are shown in figure ???. Excellent agreement is seen, and a systematic of 5% is attributed to this effect.

Finally, to estimate the sensitivity of the fake-factors to the flavour composition of the sample used to derive them, the $t\bar{t}$ validation region described in ?? is used. This sample was found in MC studies to have a largely different flavour composition than the nominal $W + \text{jets}$ sample, and can therefore be used to put a conservative systematic on this effect. A flat variation of 25% is found to bracket this effect conservatively across all observables.

As a cross-check, the same effect can be probed less directly by tuning the requirements of the $W + \text{jets}$ sample. The selection requirements applied to the $W + \text{jets}$ sample will directly affect the kinematics and flavour composition of the sample, and lead to differences in the derived fake-factors. While there is no direct way to measure how different the compositions are between the $W + \text{jets}$ control sample and the signal regions, it is possible to approximate how sensitive the fake-factors are to such changes. Some selection requirements aimed at rejecting $Z + \text{jets}$ events are applied to derive different sets of fake-factors, inspired by [**zttnote**, **tauidnote**]. The following requirements are applied separately:

- $\cos \Delta\phi(\mu, E_{\text{T}}^{\text{miss}}) + \cos \Delta\phi(\tau, E_{\text{T}}^{\text{miss}}) < -0.15$
- $m_T^\mu > 70 \text{ GeV}$
- $m_Z^{\text{vis.}} < 42 \text{ GeV}$ or $m_Z^{\text{vis.}} > 82 \text{ GeV}$
- $p_T^\mu > 35 \text{ GeV}$

The resulting sets of fake-factors are shown in figure ???. The variations are no larger than 12%, and are therefore bracketed by the 25% estimate from the $t\bar{t}$ sample.

In the end, the systematics are found to vary between 25% and 30%, being largest where the prompt subtraction is most significant.

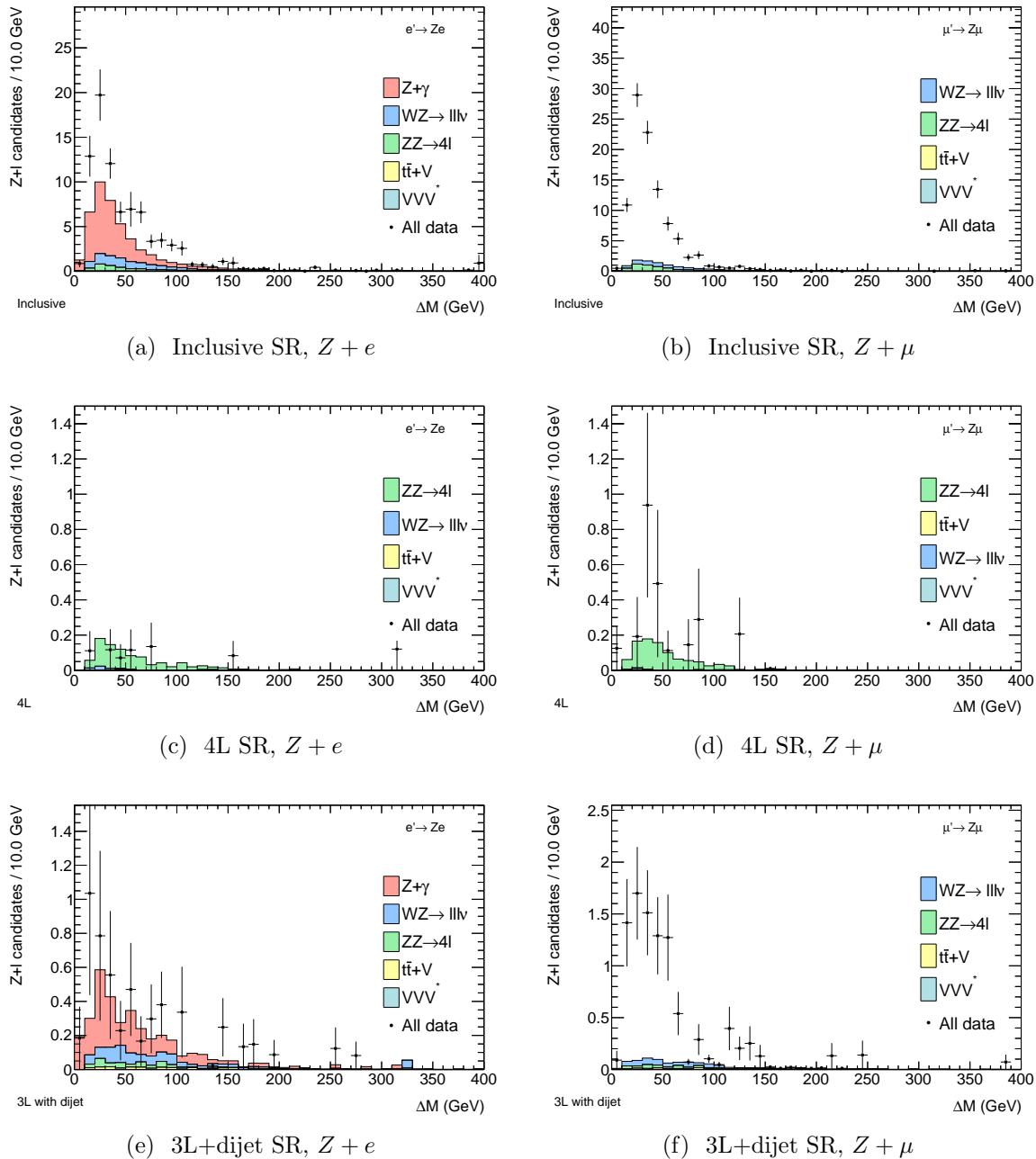


Figure 6.12: Fake factor-weighted data events and prompt subtractions. The reducible background prediction is the difference between the data points and the prompt subtraction colored histograms.

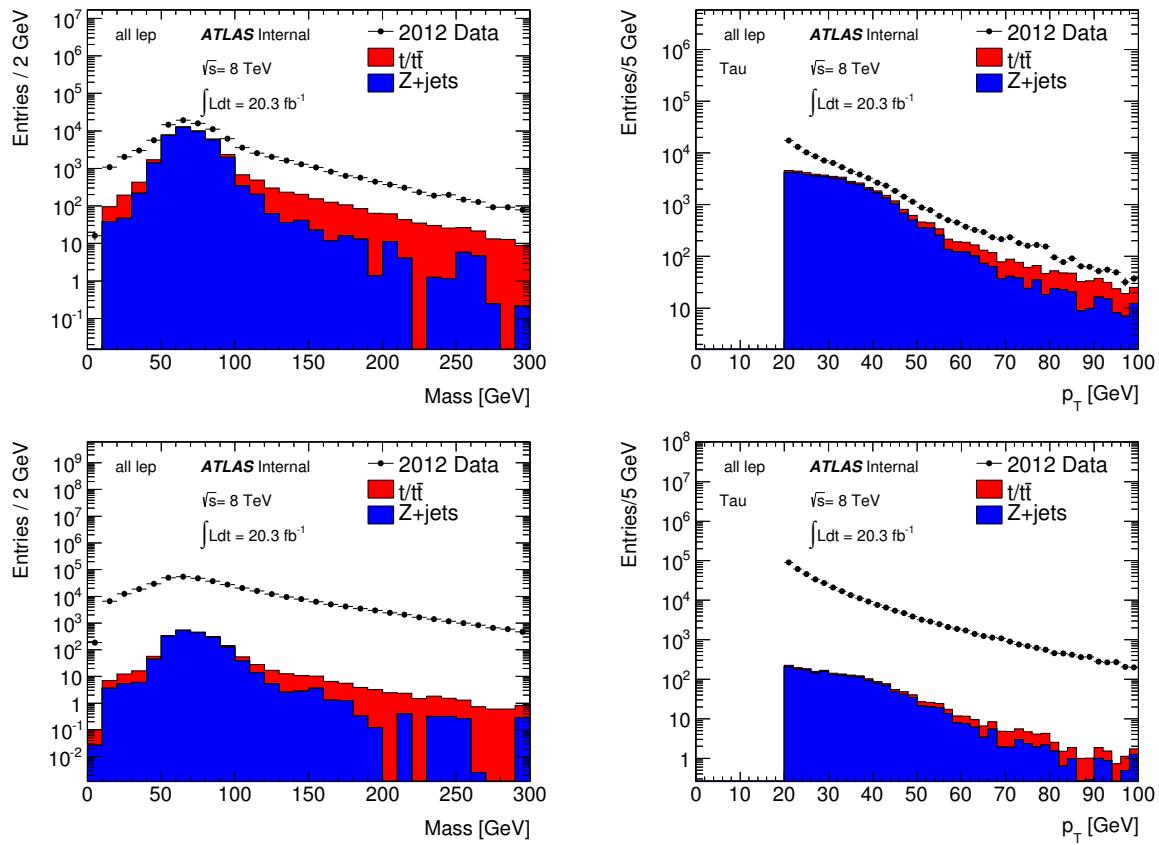


Figure 6.13: Invariant mass of the muon and tau pair (left) and p_T of the tau candidate (right), for the numerator (top) and denominator selections (down). The prompt contribution estimated from MC simulations is shown along with the measured data.

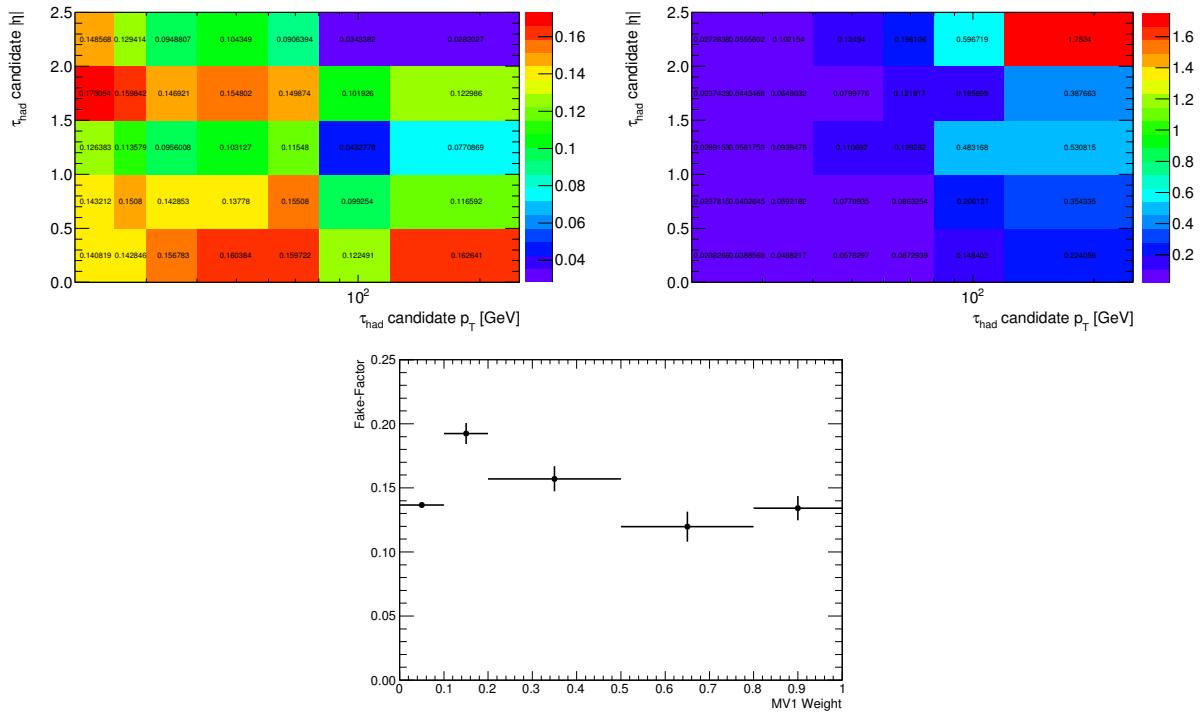


Figure 6.14: Fake-factors binned in p_T and $|\eta|$ (left) and their associated statistical uncertainty (right). The extra scaling as a function of MV1 weight is also shown (bottom). These fake-factors were derived in $W + \text{jets}$ events. The prompt contribution has been subtracted using MC predictions.

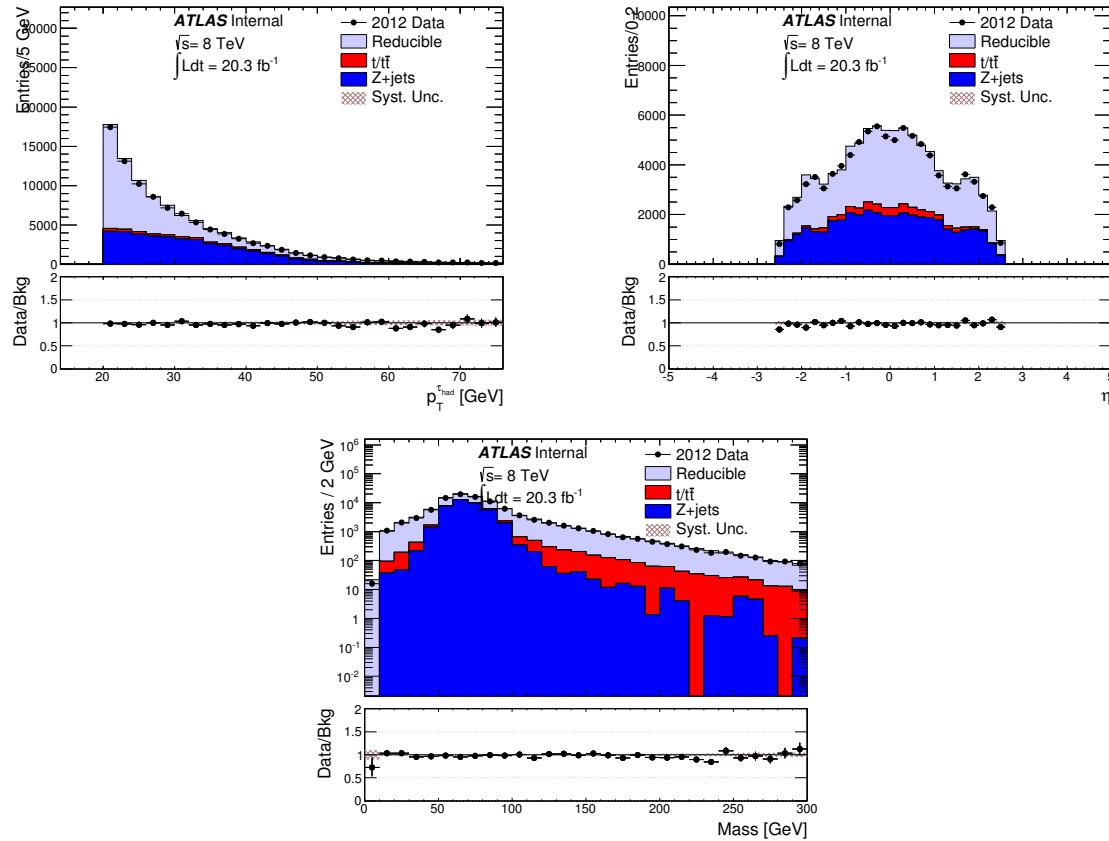


Figure 6.15: Distributions of the τ p_T (left), τ η (right), and invariant mass of the two leptons in the event (bottom).

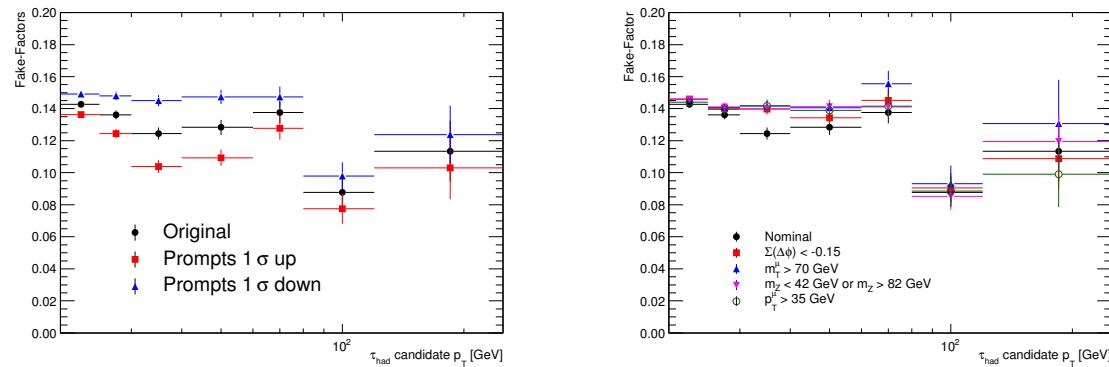


Figure 6.16: Dependency of the fake-factors on the prompt normalization (left) and the W event requirements based on the cuts to reject Z events (right).

Chapter 7

Model Independent Trilepton Search

Events containing three or more leptons are useful probes of phenomena beyond the Standard Model. On one hand, the expected Standard Model backgrounds are typically small; depending on the kinematic requirements on the three leptons, such events arise dominantly from diboson production (WZ , ZZ), or from event where one or more leptons arises from misidentified or semileptonically decaying jets. On the other hand, the production of three or more leptons is predicted by many models of phenomena beyond the Standard Model, as described in section ???. This dissertation presents two such searches: a model-independent search for non-resonant production of three or more leptons in many signal regions, and a signature-driven search for resonance trilepton production in the context of heavy leptons.

This chapter presents a search for physics beyond the Standard Model using events containing three or more leptons. The search uses 20.3 fb^{-1} of pp collision data taken at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector. Many signal regions are defined based on the properties of the leptons, jets, and overall momentum imbalance of the event, with the goal of being broadly sensitive to the non-resonant production of trilepton final states by phenomena beyond the Standard Model.

7.1 Event Selection

This section describes the selection of events containing at least three leptons in both pp collision data and in the Monte Carlo simulation samples.

Triggering

Collisions events for this analysis are triggered using the unprescaled single-electron or single-muon triggers with the lowest transverse momentum thresholds. At least one of the following triggers must have fired:

- `EF_e24vhi_medium1`: One electron with $p_T > 24 \text{ GeV}$. The electron must satisfy cuts similar to the medium++ identification criteria at the trigger level, an isolation

requirement of $\frac{p_T^{\text{cone}20}}{p_T} < 0.1$, and cuts on the leakage into the hadronic calorimeter.

- **EF_e60_medium1**: One electron with $p_T > 60$ GeV. The electron must also satisfy the medium identification cuts, but the isolation and leakage requirements are removed.
- **EF_mu24i_tight**: One muon with $p_T > 24$ GeV, satisfying an isolation requirement of $\frac{p_T^{\text{cone}20}}{p_T} < 0.12$.
- **EF_mu36_tight**: One muon with $p_T > 36$ GeV, with the isolation requirement removed.

The higher-threshold triggers without isolation requirements recover efficiency at higher p_T . Triggered events are required to have an offline lepton matched to the trigger object within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$. To avoid trigger turn-on effects near the p_T threshold, the offline lepton must have $p_T > 26$ GeV. Additionally, trigger-matched muon must have $|\eta| < 2.4$ to avoid uninstrumented regions of the detector.

Trilepton Event Selection

After successful triggering and overlap removal, events are required to have at least three selected leptons, of which at most one is a hadronically decaying tau lepton. The primary event vertex, chosen as the reconstructed vertex with the highest $\sum p_T^2$ of tracks, must have at least three tracks. Finally, events are rejected if they contain “bad jets” not associated to real energy deposits in the calorimeters due to pp collisions, i.e. from electronics problems or cosmic rays [**jet-cleaning**].

7.2 Analysis Strategy

The analysis defines a large number of non-exclusive signal regions, designed to target new physics models and to compartmentalize the expected backgrounds. First, the events are divided into 3×2 categories as follows. First, the events are divided into three categories based on the properties of any opposite-sign, same-flavor (OSSF) lepton pairs in the event:

- **on- Z** events, containing an opposite-sign, same-flavor lepton pair consistent with the decay of a Z boson, with invariant mass within 20 GeV of m_Z ;
- **off- Z , OSSF** events, containing an opposite-sign, same-flavor pair but vetoing on- Z events; and
- **off- Z , mixed** events, containing no opposite-sign, same-flavor pairs.

The on- Z category also includes events containing three leptons (two of which form a same-flavor, opposite-sign pair) with invariant mass within 20 GeV of m_Z , to include events

where, for example, a photon from final state radiation converts and is reconstructed as a prompt electron.

Next, the events are further divided into two categories based on the number of electron or muon candidates in the event:

- **3L** events, containing at least three electrons or muons, and
- **2L+ τ_{had}** events, containing exactly two electrons or muons and a hadronically decaying tau lepton.

After dividing the events into these six exclusive categories, many signal regions are defined based on the lower bound in various kinematic variables. An ordering is imposed on the leptons for the sake of disambiguation: in the 3L category, the leptons are ordered by p_T , while in the 2L category, the electrons or muons are ordered by p_T , and the τ_{had} is the third lepton. The variables used to define the signal regions are:

- H_T^{leptons} : the scalar sum of the transverse momenta of the leading three leptons. Events containing new particles with masses significantly greater than m_W or m_Z will typically have larger H_T^{leptons} than the Standard Model backgrounds.
- Minimum p_T^ℓ : the p_T of the softest of the leading three leptons. As with H_T^{leptons} , the p_T of leptons produced in the decays of heavy particles will tend to be larger than those from the expected Standard Model backgrounds.
- H_T^{jets} : the scalar sum of the transverse momenta of all selected jets in the event. This variable is sensitive to the strong production of new physics where several leptons are produced in the decays of heavy particles, such as the gluino pair production described in section ???. Conversely, the Standard Model WZ and ZZ backgrounds are weakly produced, and have softer H_T^{jets} distributions.
- E_T^{miss} : the magnitude of the missing transverse momentum in the event. In models of new physics, leptons are often produced with neutrinos in leptonic W decays, or with new invisible particles, such as the stable neutralinos in many models of R -parity conserving SUSY. Requiring large E_T^{miss} also suppress backgrounds due to $Z+\text{jets}$, where the jet decays semileptonically or is misidentified as a lepton.
- m_{eff} : the scalar sum of H_T^{jets} , E_T^{miss} , and the p_T of all identified leptons in the event. As with H_T^{leptons} by itself, multilepton production due to the decays of heavy particles will typically have a harder m_{eff} distribution than the Standard Model backgrounds.
- m_T^W : for events in the on- Z categories, the transverse mass of the missing transverse momentum, \vec{p}_T^{miss} , and the highest- p_T lepton not associated with a Z boson candidate, defined as:

$$m_T^W = \sqrt{2\vec{p}_T^\ell |\vec{p}_T^{\text{miss}}| (1 - \cos(\Delta\phi))}, \quad (7.1)$$

Variable	Meaning					
H_T^{jets}	Σp_T of all jets in the event					
m_T^W	Transverse mass of W -boson candidate (on- Z events only)					
Variable	Meaning	Lower Bounds [GeV]		Additional Requirements		
H_T^{leptons}	Σp_T of leading three leptons	0	200	500	800	
Min. p_T^ℓ	p_T of softest (third) lepton	0	50	100	150	
E_T^{miss}	Missing transverse momentum	0	100	200	300	$H_T^{\text{jets}} < 150$ GeV
E_T^{miss}		0	100	200	300	$H_T^{\text{jets}} \geq 150$ GeV
m_{eff}		0	600	1000	1500	
m_{eff}	All transverse activity	0	600	1200		$E_T^{\text{miss}} \geq 100$ GeV
m_{eff}		0	600	1200		$m_T^W \geq 100$ GeV, on- Z
Variable	Meaning	Lower Bounds				
b -tags	Number of b -tagged jets	1	2			

Table 7.1: Kinematic signal regions defined in the analysis.

where $\Delta\phi$ is the azimuthal angle between the transverse momentum of the lepton, \vec{p}_T^ℓ , and the missing transverse momentum, \vec{p}_T^{miss} .

- $N_{b\text{-tags}}$, the number of b -tagged jets. New physics scenarios related to the hierarchy problem (section ??) often couple preferentially to the third generation, due to the dominant effect of the top quark in the running of the Higgs mass.

The signal regions are defined in table ???. The signal regions use one of H_T^{leptons} , the minimum p_T^ℓ , E_T^{miss} , m_{eff} , and n_b as binning variables. H_T^{jets} , E_T^{miss} , and m_T^W are used to impose additional requirements on the signal regions. In total, 138 signal regions are defined.

7.3 Systematic Uncertainties

Systematic uncertainties are assigned to the signal and background predictions to account for possible modelling inaccuracies. The sources of uncertainty considered are:

- Uncertainties on the reducible background estimates from the data-driven estimation method, as described in section ???. The uncertainties ranges between 20% to 30% for the electron fake factors, 25% to 50% for the muon fake factors, and 25% to 30% for the tau lepton fake factors.
- The signal and background samples must be normalized to the integrated luminosity of the data. The weight assigned to each event is:

$$w = \frac{\mathcal{L}\sigma_{\text{process}}}{N_{\text{sim}}}, \quad (7.2)$$

Sample	Uncertainty	Reference
$Z + \text{jets}$	10%	
$t\bar{t} + V$	22%	[ttV]
ZZ	4.3%	App. ??
WZ	7.6%	App. ??
$t\bar{t}$	10%	[TopXS]
single- t	10%	[atlas-topmc]

Table 7.2: Systematic uncertainties used for the MC samples that contribute to the background estimates.

where \mathcal{L} is the integrated luminosity of the data, σ_{process} is the cross section of the simulated process, and N_{sim} is the number of simulated events. The integrated luminosity carries an uncertainty of 2.8%, as described in section ???. The cross sections are taken from calculations with uncertainties due to various sources, including the fixed order of the calculation, the choice of renormalization and factorization scales, and the PDFs. The uncertainties are shown in table ??.

- Events are also weighted to account for differences between data and Monte Carlo simulation in the lepton trigger, reconstruction, and identification efficiencies. The scale factors and associated uncertainties are provided by the relevant ATLAS combined performance groups. The electron and muon scale factors are close to unity, with uncertainties in the range 1-5%. The scale factors for the tau lepton identification efficiency range from 94-96% for the **BDTTight** working point, and carry uncertainties between 2.0-2.2% [**TauUncertainties**].
- The lepton energy scales carry uncertainty which has a small effect on the signal regions due to leptons or events near E_T or p_T thresholds. The uncertainties are also provided by the combined performance groups.
- The jet energies are scaled and smeared according to recommendations from the **JetEtMiss** combined performance group [**JES**, **JER**]. The uncertainty affects the H_T^{jets} and m_{eff} distributions, and also the lepton-jet overlap removal procedure.
- The lepton and jet energy uncertainties are also propagated to the missing transverse energy, E_T^{miss} , using the **MissingETUtility** package [**METUtility**].
- The efficiencies of the b -tagging algorithm carry uncertainties which are parameterized based on matching the identified b -jets to truth b -jets, as recommended by the **flavor tagging** combined performance group.
- Finally, the Monte Carlo samples carry statistical uncertainty due to simulating a finite number of events.

The dominant sources of uncertainty depend on the signal region, but are generally due to Monte Carlo statistical uncertainties, the fake factor uncertainites, and theoretical cross section uncertainties.

7.4 Background Validation

The background estimates are verified in several dedicated validation regions, which are orthogonal to the signal regions. Events containing two leptons are used to test the selection of prompt leptons and their corresponding scale factors and energy corrections. The reducible backgrounds are tested in $t\bar{t}$ validation regions, where the events are required to have two same-sign leptons to target semileptonic decays where one lepton arises from a reducible process. Finally, the fake factor procedure is validation in events with modified lepton selections, where the signal leptons are “between” the numerator and denominator definitions used in the nominal reducible background estimate.

Dilepton Validation Regions

Three dilepton validation regions target the three flavors of leptons: ee , $\mu\mu$, and $\mu\tau_{had}$. The ee and $\mu\mu$ regions require an opposite-sign lepton pair of the appropriate flavor. The invariant mass distributions of the dilepton system in these regions are shown in figure ???. Some disagreement is observed in the ee invariant mass distribution; this discrepancy is covered by the electron energy scale systematic uncertainties.

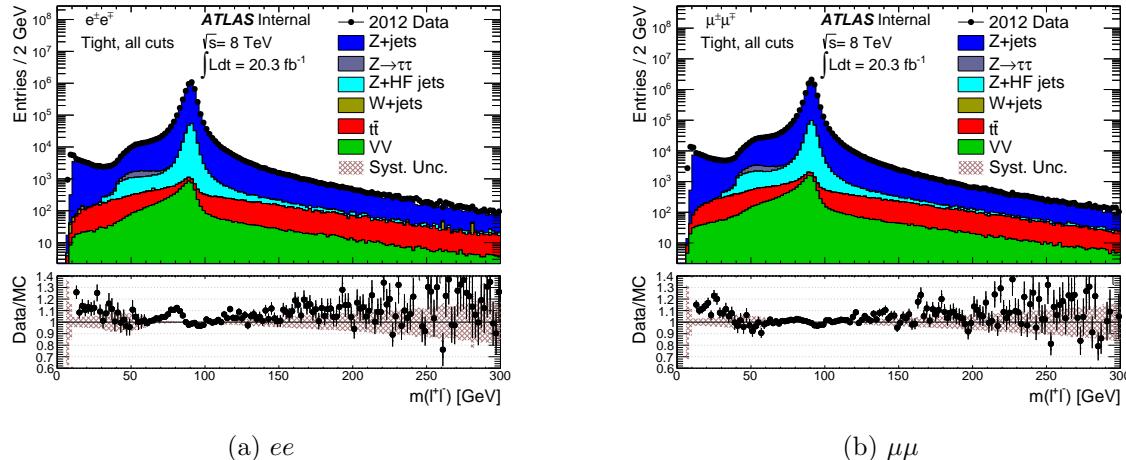


Figure 7.1: Dilepton invariant mass distributions for the ee (left) and $\mu\mu$ (right) validation regions.

The ee and $\mu\mu$ validation regions are also used to generate scale factors to account for efficiency differences between simulation and data when applying cuts on lepton isolation or

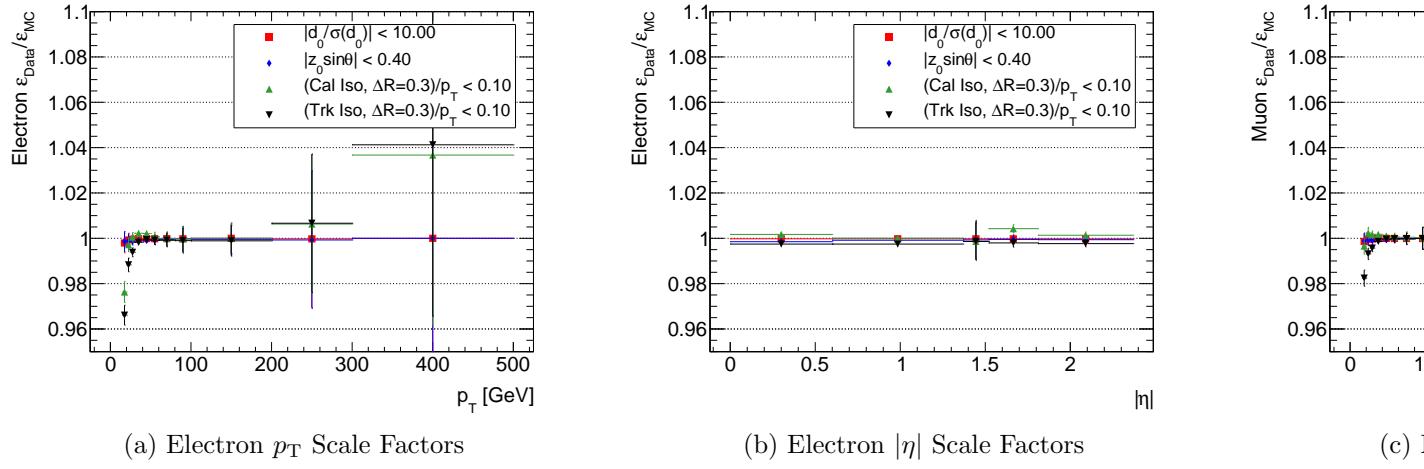


Figure 7.2: Scale factors for electrons and muons as functions of p_T (left) and $|\eta|$ (right).

impact parameter. The scale factors are computed from events with a dilepton pair with invariant mass within 10 GeV of m_Z , and are shown in figure ???. The scale factors are mostly consistent with unity to within 0.5%, except for isolation requirements on the low- p_T leptons, where the scale factors deviate from unity by up to 2%.

The $\mu\tau_{had}$ region applies addition cuts to reduce the large contribution from fake tau leptons in $W+jets$ events:

- $\cos \Delta\phi(\mu, E_T^{\text{miss}}) + \cos \Delta\phi(\tau, E_T^{\text{miss}}) > -0.15$,
- $\Delta\phi(\mu, \tau_{had}) > 2.4$,
- $m_T^\mu < 50$ GeV,
- $42 < m_Z^{\text{vis.}} < 82$ GeV, and
- $p_T^\mu < 40$ GeV.

The p_T and η distributions of the μ and τ_{had} are shown in figure ??.

$t\bar{t}$ Validation Regions

In signal regions that veto Z bosons and require large H_T^{jets} or E_T^{miss} , a large reducible background component is expected from $t\bar{t}$ events, where both W bosons decay leptonically and a third lepton arises from a misidentified jet or semileptonic heavy flavor decay. This process is tested in the $t\bar{t}$ validation regions, which require at least one b -tagged jet with $p_T > 30$ GeV, exactly two electrons or muons with the same charge, and $H_T^{\text{jets}} < 500$ GeV. The requirement that the leptons have the same sign vetoes dilepton $t\bar{t}$ events with two prompt

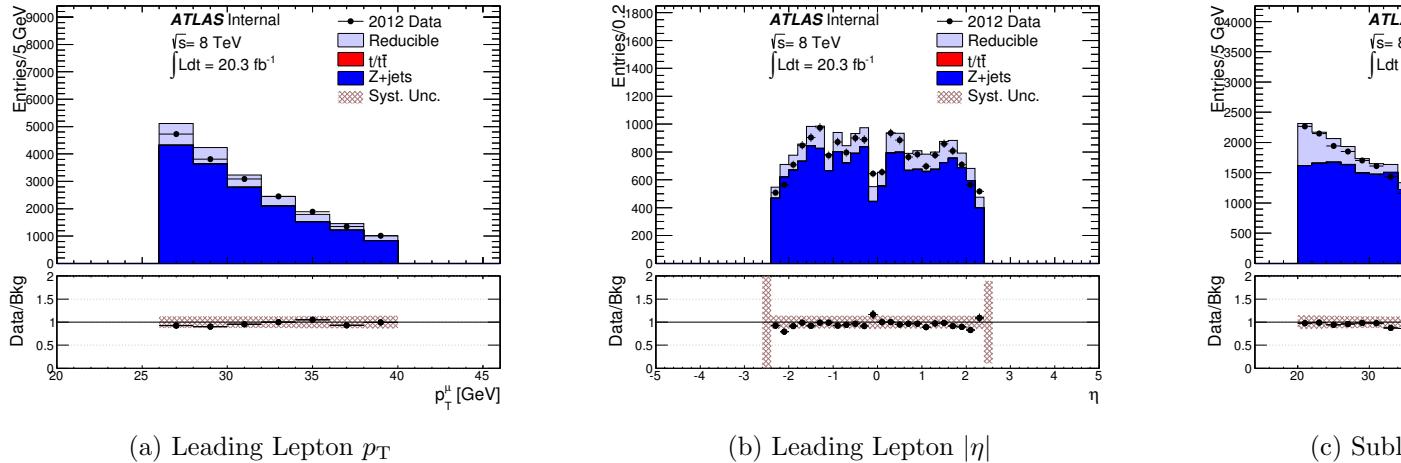


Figure 7.3: p_T and η distribution of the muon and the hadronically decaying tau lepton in the $\mu\tau_{\text{had}}$ validation region.

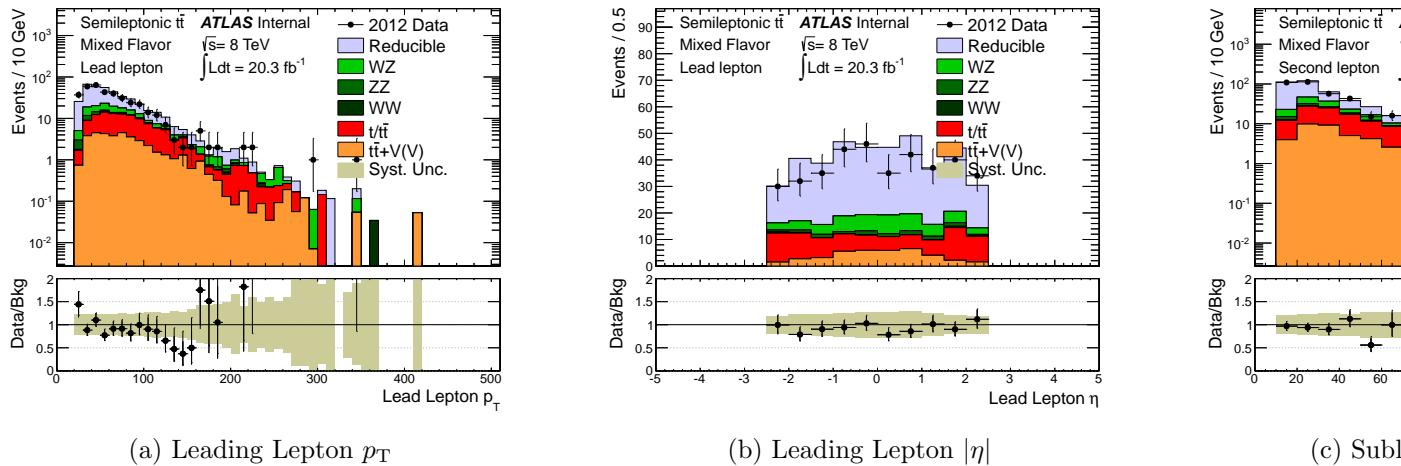


Figure 7.4: $t\bar{t}$ validation regions for all flavor (e, μ) combinations.

leptons from W decays, while the H_T^{jets} requirement reduces the potential contamination from new BSM phenomena. Note that events containing hadronically decaying tau leptons are used to derive a systematic uncertainty on the tau fake factors due to the flavor composition of the events (section ??), and so are not used as a validation region.

The p_T and η distributions of the leading and subleading leptons are shown in figure ??.

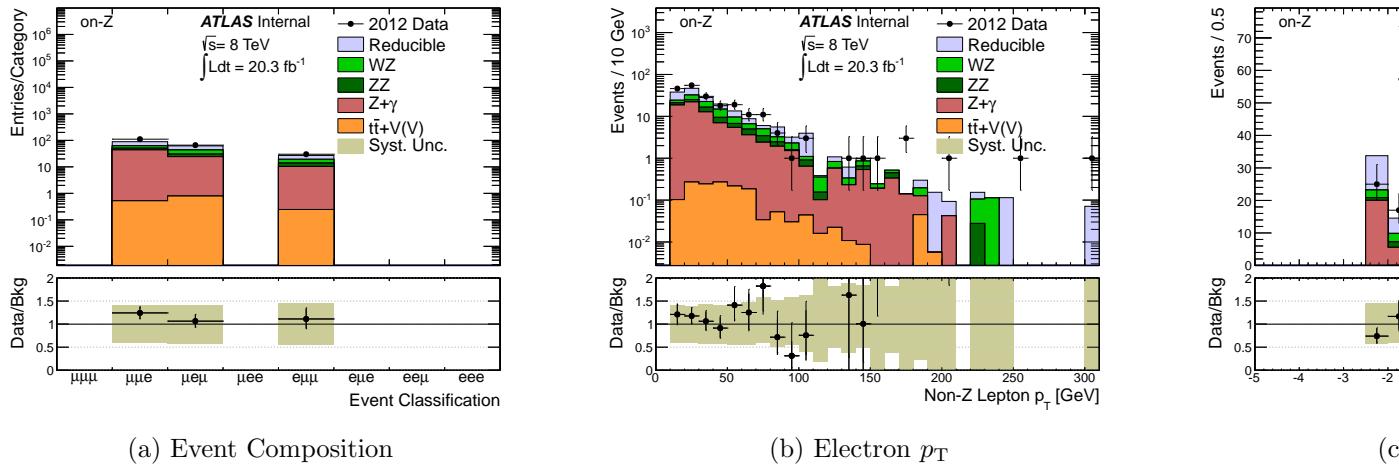


Figure 7.5: Expected backgrounds in the intermediate-electron validation region, using the nominal background estimation techniques.

Intermediate Fake Factor Validation Regions

The fake factor method is further validated in regions containing events with two signal (numerator) leptons and one “intermediate” lepton, which fulfills selection criteria looser than the numerator criteria but tighter than the denominator criteria. Separate fake factors are derived for intermediate leptons; these are defined schematically as $\langle \text{intermediate} \rangle / \langle \text{denominator} \rangle$, rather than $\langle \text{numerator} \rangle / \langle \text{denominator} \rangle$. Specifically, the intermediate leptons satisfy most of the numerator criteria except for:

- **Electrons:** the electron must pass the `medium++` cuts, but fail the `tight++` cuts.
 - **Muons:** the candidate must have either its track or calorimeter transverse isolation energy satisfy $0.10 < \text{iso}/p_T < 0.15$.
 - **Taus:** the tau candidate must pass the `Medium-BDT` identification requirement, but fail the `tight-BDT` identification requirement.

The first two regions, targeting reducible electrons and muons, require an OSSF pair of electrons or muons, plus a third intermediate electron or muon of different flavor. The events are largely due to $Z \rightarrow ee$ or $Z \rightarrow \mu\mu$ events, with an additional lepton that is either prompt and fails the signal lepton criteria, or is due to a reducible process.

The events containing an intermediate τ_{had} are much more abundant than intermediate electrons or muons, so three τ_{had} validation regions are defined, mirroring the signal region categories. Events are required to contain two electrons or muons plus an intermediate τ_{had} , and are separated into on- Z , off- Z /OSSF, and off- Z /mixed categories, depending on the two electrons or muons. The p_T and η of the intermediate τ_{had} is shown for each category in figure ??.

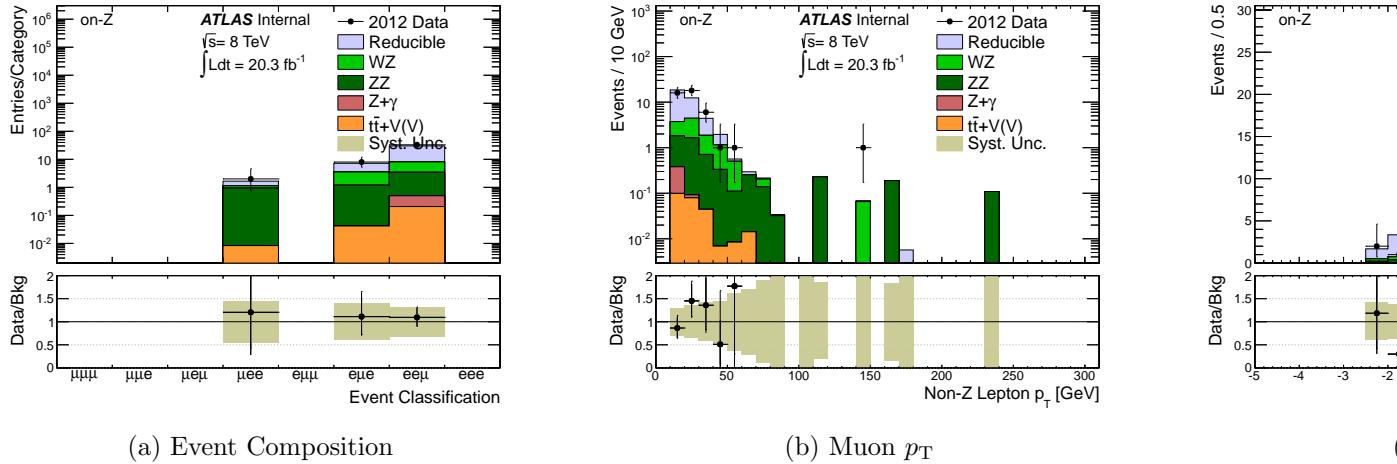


Figure 7.6: Expected backgrounds in the intermediate-muon validation region, using the nominal background estimation techniques.

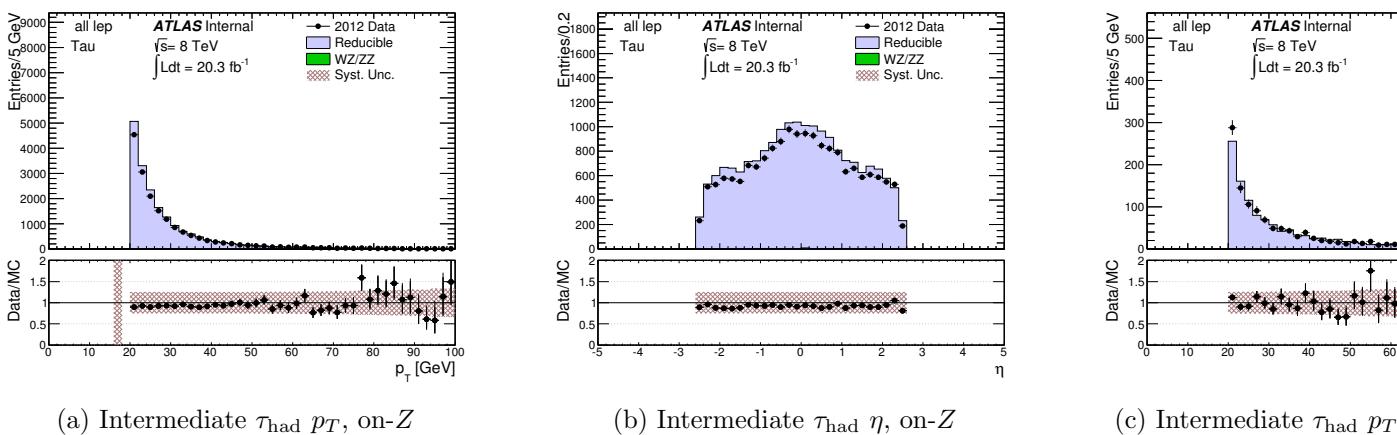


Figure 7.7: p_T and η distributions of the intermediate τ_{had} in the on- Z , off- Z /OSSF, and off- Z /mixed intermediate validation regions.

7.5 Results and Limits

The results of the search as organized as follows. For each category, the distributions of the various kinematic variables (H_T^{leptons} , H_T^{jets} , m_{eff} , E_T^{miss} , m_T^W , N_{jets} , $N_{b\text{-tags}}$, and the p_T of the third lepton) are produced (where applicable). Then, in each of the 138 signal regions, the number of observed and expected events are compared, and, absent any significant deviations, 95% confidence level (CL) upper limits are derived on the number of events from new physics, N_{95} , and the corresponding *visible cross section*, σ_{95}^{vis} , defined as

$$\sigma_{95}^{\text{vis}} = \frac{N_{95}}{\mathcal{L}}, \quad (7.3)$$

where $\mathcal{L} = 20.3 \text{ fb}^{-1}$ is the integrated luminosity of the data.

The plots for the H_T^{leptons} signal regions are shown here for demonstrative purposes. The full set of plots is available at [[model-independent-webpage](#)]. The H_T^{leptons} distributions are shown in figure ???. The observed and expected event counts in the three corresponding signal regions ($H_T^{\text{leptons}} > 200 \text{ GeV}$, $H_T^{\text{leptons}} > 500 \text{ GeV}$, and $H_T^{\text{leptons}} > 800 \text{ GeV}$) are shown in table ???. Finally, the upper limits on σ_{95}^{vis} are shown in figure ???.

The deviations between the observed and expected in all 138 signal regions are shown in figure ???, in units of the total uncertainty on the background expectation.

Figure 7.8: H_T^{leptons} distributions in the six categories.

Table 7.3: Observed and expected event counts in the H_T^{leptons} signal regions, along with the inclusive counts for the entire category.

7.6 Interpretations

The 95% CL upper limits on trilepton event production from new physics derived in section ?? provide a useful tool with which to confront models of new physics producing trilepton final states. In order to compare the predictions of a model, i.e. a set of simulated events at

Figure 7.9: 95% CL upper limits on the visible cross section of trilepton event production from new physics, σ_{95}^{vis} , in each H_T^{leptons} signal region.

Figure 7.10: Deviations between the observed event counts and the background expectations in each signal region, in units of the total uncertainty on the background prediction.

particle level from an event generator, per-lepton fiducial efficiencies are provided to quantify approximately the efficiency of reconstructing and selecting fiducial leptons at truth level. The definition of fiducial truth leptons is as follows:

- **Transverse momentum:** Electrons and muons are required to have $p_T > 10$ GeV, while tau leptons are required to have $p_T > 15$ GeV.
- **Pseudorapidity:** The same pseudorapidity cuts as the reconstructed signal leptons are applied. Electrons must have $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$, and muons and tau leptons must have $|\eta| < 2.5$.
- **Isolation:** For electrons and muons, the sum of the transverse momenta of all charged particles with $p_T > 1$ GeV within a cone of radius $\Delta R = 0.3$ of the lepton, denoted `truth_ptcone30`, must satisfy $\text{truth_ptcone30}/p_T < 0.15$. Similarly, the `truth_Etcone30`, defined as the sum of all stable, visible particles within a cone of radius $\Delta R = 0.3$, is required to satisfy $\text{truth_Etcone30}/p_T < 0.15$.
- **Origin and Decay:** Leptons must originate from the hard scattering interaction (as opposed to the interaction with the detector as simulated by GEANT4), and not arise from the decay of a hadron. Electrons and muons are also required to be stable; tau leptons are required to decay hadronically.

The efficiencies are derived from a Monte Carlo simulation sample
Efficiencies. Doubly charged higgs interpretation?

Chapter 8

Trilepton Resonance Search

The sensitivity of a search using events with many electrons or muons can be enhanced significantly if the leptons are produced resonantly, as the full reconstructible final states allows the use of a mass constraint. Searches for resonance dilepton production have a rich history, including the discoveries of the J/ψ [**jpsi1**, **jpsi2**], the Υ [**upsilon**], and the Z boson [**zua1**]. In 2012, the resonant four-lepton channel proved instrumental in the discovery of the Higgs boson [**ATLAS-higgs**, **CMS-higgs**]. At the LHC, resonant dilepton searches have placed strong constraints on a variety of new physics scenarios, such as new gauge bosons [**CMS-dilepton**, **zprime`8TeV**] and doubly-charged scalars [**doubly-charged-higgs**].

This chapter presents a search for the resonant production of three leptons using 20.3 fb^{-1} of pp collision data at $\sqrt{s} = 8 \text{ TeV}$. Trilepton resonances have been used previously at lower energies to place constraints on lepton flavor violation in muon and tau lepton decays [**lll-flavor-decay-papers**]. This analysis targets high-mass heavy leptons, L^\pm , decaying to three leptons via an intermediate, on-shell Z boson, $L^\pm \rightarrow Z + \ell \rightarrow \ell\ell\ell$. The leptons are required to be electrons or muons, $\ell = e, \mu$; in this case, the final state is fully reconstructible, and the trilepton mass constraint is a power handle with which to discriminate the signal from the background.

The chapter is organized as follows. Section ?? describes the models of new physics used to motivate the analysis. Section ?? describes the search strategy, including the event selection and the identification of trilepton resonance candidates. The background processes, including the estimation method, are described in section ?? . Section ?? lists the sources of systematic uncertainty on the signal and background estimates. The background estimates are validated in section ?? . Finally, the results and interpretations are shown in sections ?? and ??.

8.1 Signal Models

The search is motivated by two models of phenomena beyond the Standard Model: the type III neutrino seesaw model [Biggio:1368793] and extra generations of vector-like leptons [Martin:2009it], discussed earlier in sections ?? and ???. Both models propose heavy, charged, and colorless fermions which are made unstable by mixing with Standard Model leptons. The type III seesaw model also includes a neutral heavy lepton, N^0 , with the same mass as the L^\pm .

The new particles are pair produced via gauge interactions, $q\bar{q} \rightarrow L^\pm L^\mp$ or $q\bar{q}' \rightarrow L^\pm N^0$, as shown in figure ???. The production cross sections depend on how the new particles couple to $SU(2)_L \times U(1)_Y$: the type III seesaw fermions transform in the adjoint representation of $SU(2)_L$ and have zero hypercharge, $(\mathbf{3}, 0)$, while the heavy lepton in the generic vector-like lepton model inherits its gauge couplings from its $SU(5)$ multiplet, $(1, -1)$. The different gauge couplings, plus the various production modes, lead to significant differences in production rates. The production cross sections of the two models are shown in figure ??.

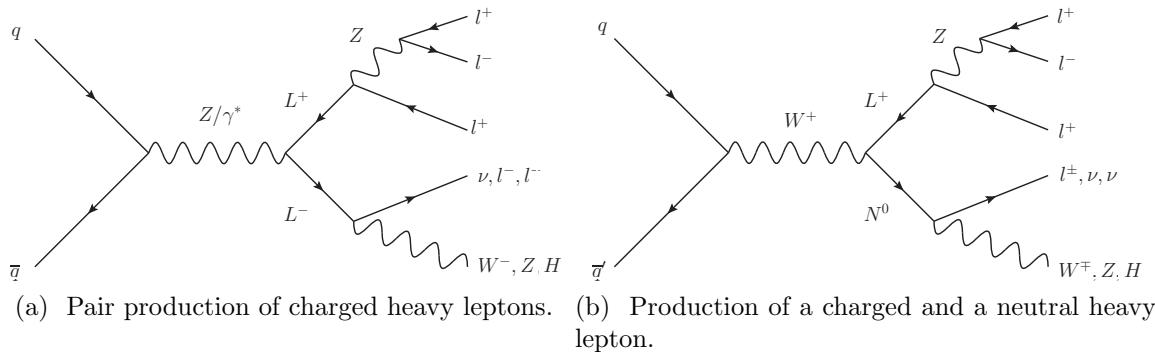


Figure 8.1: Production and decay of new heavy leptons to final states with a trilepton resonance.

The heavy leptons are made unstable by introducing mixing terms with the Standard Model leptons. For example, consider a single extra generation of fermions transforming in the adjoint representation of $SU(2)_L$, as in the type III seesaw model: $\Sigma \equiv \begin{pmatrix} N^0/\sqrt{2} & \Sigma^+ \\ \Sigma^- & -N^0/\sqrt{2} \end{pmatrix}$ [Abada:Biggio:2011ja]. The Lagrangian contains Yukawa terms mixing the heavy leptons with Standard Model leptons:

$$-\mathcal{L} \ni \sum_{l=e,\mu,\tau} \sqrt{2}\phi^0 \bar{\Psi} Y_{L^\pm l} l_L + \text{h.c.}, \quad (8.1)$$

where $\Psi \equiv \Sigma_R^{+c} + \Sigma_R^-$ is a Dirac spinor representing the four charged degrees of freedom, $\phi \equiv (\phi^+, \phi^0)^T \equiv (\phi^+, (v + H + i\eta)/\sqrt{2})^T$ is the Higgs doublet, $Y_{L^\pm l}$ are Yukawa couplings,

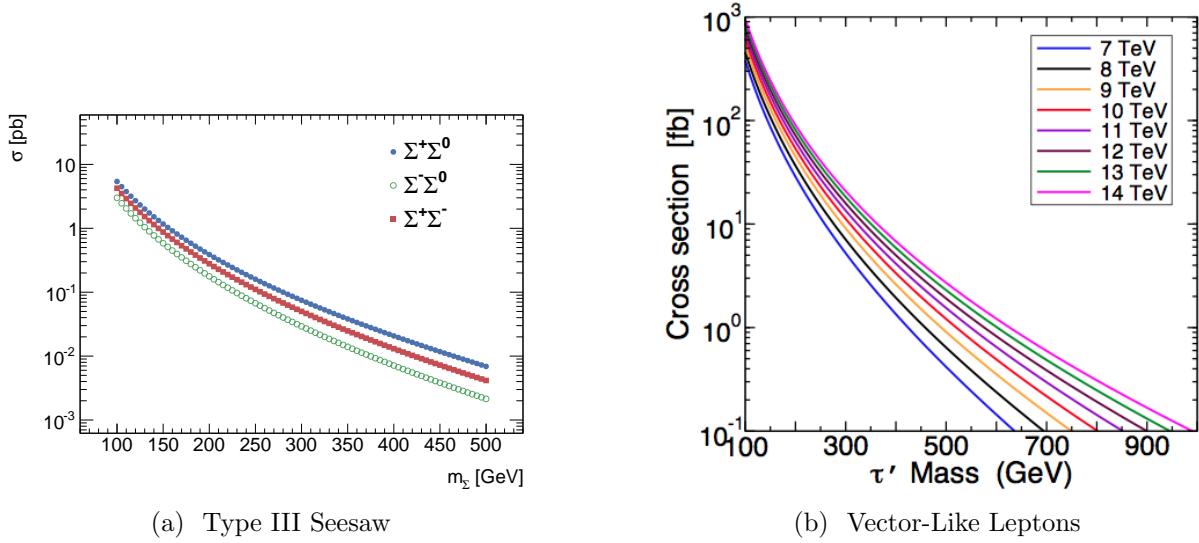


Figure 8.2: Production cross sections at $\sqrt{s} = 8$ TeV for heavy lepton pair production for the type III seesaw model (left) and the vector-like leptons model (right).

and l_L is a Standard Model lepton. After electroweak symmetry breaking, the mass matrices take the form

$$-\mathcal{L} \ni \sum_{l=e,\mu,\tau} (\bar{l}_R \quad \bar{\Psi}_R) \begin{pmatrix} m_l & 0 \\ Y_{L^\pm l} v & M_{L^\pm} \end{pmatrix} \begin{pmatrix} l_L \\ \Psi_L \end{pmatrix} + (\bar{l}_L \quad \bar{\Psi}_L) \begin{pmatrix} m_l & Y_{L^\pm l}^\dagger v \\ 0 & M_{L^\pm} \end{pmatrix} \begin{pmatrix} l_R \\ \Psi_r \end{pmatrix}. \quad (8.2)$$

Diagonalizing the mass matrices leads to off-diagonal terms in the gauge interactions, with couplings proportional to the mixing parameters $V_{\alpha\Sigma} = \frac{v}{\sqrt{2}} M_{L^\pm}^{-1} Y_{\alpha\Sigma}$. These couplings enable the decay of the heavy leptons to a boson (W , Z , or h) plus a Standard Model lepton or neutrino, with partial widths given by:

$$\Gamma(N^0 \rightarrow l_\alpha^- W^+) = \Gamma(N^0 \rightarrow l_\alpha^+ W^-) = \frac{g^2}{64\pi} |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_W^2} \left(1 - \frac{M_W^2}{M_{L^\pm}^2}\right)^2 \left(1 + 2\frac{M_W^2}{M_{L^\pm}^2}\right), \quad (8.3)$$

$$\sum_l \Gamma(N^0 \rightarrow \nu_l Z) = \frac{g^2}{64\pi c_W^2} \sum_\alpha |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_Z^2} \left(1 - \frac{M_Z^2}{M_{L^\pm}^2}\right)^2 \left(1 + 2\frac{M_Z^2}{M_{L^\pm}^2}\right), \quad (8.4)$$

$$\sum_l \Gamma(N^0 \rightarrow \nu_l H) = \frac{g^2}{64\pi} \sum_\alpha |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_W^2} \left(1 - \frac{M_H^2}{M_{L^\pm}^2}\right)^2, \quad (8.5)$$

$$\sum_l \Gamma(\Sigma^+ \rightarrow \nu_l W^+) = \frac{g^2}{32\pi} \sum_\alpha |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_W^2} \left(1 - \frac{M_W^2}{M_{L^\pm}^2}\right)^2 \left(1 + 2\frac{M_W^2}{M_{L^\pm}^2}\right), \quad (8.6)$$

$$\Gamma(\Sigma^+ \rightarrow l_\alpha^+ Z) = \frac{g^2}{64\pi c_W^2} |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_Z^2} \left(1 - \frac{M_Z^2}{M_{L^\pm}^2}\right)^2 \left(1 + 2\frac{M_Z^2}{M_{L^\pm}^2}\right), \quad (8.7)$$

$$\Gamma(\Sigma^+ \rightarrow l_\alpha^+ H) = \frac{g^2}{64\pi} |V_{\alpha\Sigma}|^2 \frac{M_{L^\pm}^3}{M_H^2} \left(1 - \frac{M_H^2}{M_{L^\pm}^2}\right)^2. \quad (8.8)$$

where α and l run over Standard Model lepton and neutrino generations. The branching fractions of the charged and neutrino heavy leptons are shown as function of m_{L^\pm} in figure ???. Note that these branching fractions are common to all the models considered in this analysis (aside from the possibility of only having a charged heavy lepton).

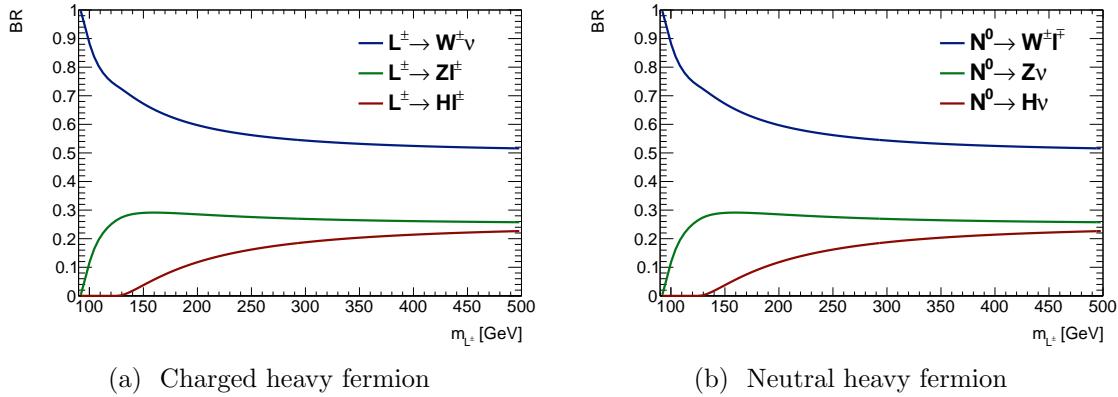


Figure 8.3: Branching ratios of a heavy lepton decaying via mixing with Standard Model leptons [Biggio:2011ja].

Constraints on the mixing parameters V_i can be derived from precision measurements of the Z width, and constraints on products of the mixing parameters from flavor vio-

lation experiments like $\mu \rightarrow e\gamma$ [**Abada:2008ea**, **Abada:2007ux**, **delAguila:2008pw**, **Altmannshofer:2013zba**]:

$$|V_{eL}| < 5.5 \times 10^{-2} \quad (8.9)$$

$$|V_{\mu L}| < 6.3 \times 10^{-2} \quad (8.10)$$

$$|V_{\tau L}| < 6.3 \times 10^{-2} \quad (8.11)$$

$$|V_{eL}V_{\mu L}| < 1.7 \times 10^{-7} \quad (8.12)$$

$$|V_{eL}V_{\tau L}| < 4.2 \times 10^{-4} \quad (8.13)$$

$$|V_{\mu L}V_{\tau L}| < 4.9 \times 10^{-4} \quad (8.14)$$

Signal Monte Carlo Samples

Signal events are generated using Monte Carlo simulation, at eleven mass hypotheses in the range $100 \text{ GeV} \leq m_{L^\pm} \leq 400 \text{ GeV}$ for the vector-like leptons model and ten mass hypothesis in the range $100 \text{ GeV} \leq m_{L^\pm, N^0} \leq 500 \text{ GeV}$ for the type III seesaw model. The samples are summarized in table ???. The events are generated with MADGRAPH 4.5.2 and 5.2.2.1 [**madgraph**], for the vector-like leptons and type III seesaw models, respectively, using the CTEQ6L1[**ct6l1**] parton distribution functions (PDF) and the AU2 underlying event tune [**AU2**]. Showering is performed with PYTHIA 8 [**Sjostrand:2007gs**]. Decays of the heavy leptons in the vector-like leptons model are performed using BRIDGE [**bridge**], while decays in the type III seesaw samples are performed by MADGRAPH. The cross sections for both samples are calculated at leading order (LO) in QCD, and are shown in table ??.

For the type III seesaw model, the charged and neutral heavy leptons are generated with identical masses. The simulated decay modes are listed in table ???. The mixing parameters are chosen to be $V_e = 0.055$, $V_\mu = 0.063$, and $V_\tau = 0$, yielding a mix of ee , $e\mu$, and $\mu\mu$ decays. For the vector-like leptons model, one heavy lepton is required to decay via $L^\pm \rightarrow Z(\ell\ell)\ell$, where $\ell = e, \mu, \tau$. The two heavy leptons are constrained to decay to the same flavor of Standard Model lepton, which is chosen with equal probability to be an electron, muon, or tau. For both models, the events are divided and reweighted to correspond to electron-only or muon-only mixing scenarios.

The type III seesaw samples were simulated with no intrinsic width for the Z bosons. As a result, the signal peaks are slightly narrower than would be expected if the Z width were correctly included. As described below in section ???, the limit setting is performed by fitting the Δm distributions with analytical functions. In order to incorporate the Z width into the signal peaks, the peak width parameters of the vector-like lepton peaks are used for the type III seesaw model.

Mass [GeV]	Dataset ID	Cross Section [pb]	Filter Efficiency	Number of Events
Type III Seesaw				
100	191030	1.273	0.1018	100,000
120	191031	2.138	0.1628	100,000
160	191032	0.853	0.1808	100,000
200	191033	0.346	0.2012	100,000
250	191034	0.135	0.2096	100,000
300	191035	0.0604	0.2178	50,000
350	191036	0.02969	0.2337	50,000
400	191037	0.01566	0.2409	50,000
450	191038	0.008733	0.2456	50,000
500	191039	0.00504	0.2493	50,000
Vector-Like Leptons				
100	180579	0.378	0.0233	100,000
110	180580	0.264	0.0412	100,000
120	180581	0.193	0.0498	100,000
130	180582	0.142	0.0548	100,000
140	180583	0.106	0.0569	100,000
160	180584	0.0645	0.0579	100,000
180	180585	0.0407	0.0575	100,000
200	180586	0.0265	0.0567	100,000
250	180587	0.0104	0.0548	100,000
300	180588	0.00457	0.0536	100,000
400	180589	0.00115	0.0520	100,000

Table 8.1: Summary of signal Monte Carlo samples. Ten mass points are simulated for the type III seesaw model, and eleven for the vector-like leptons model.

8.2 Search Strategy

Event Selection and Heavy Lepton Reconstruction

The event selection is very similar to that used in the model-independent analysis, described in section ???. In particular, the same definitions are used for leptons and jets, in order to make use of the same fake factor method to estimate the reducible backgrounds. Events containing a heavy lepton candidate are selected as follows.

- Events are triggered using the lowest-threshold unprescaled single-electron and single-muon triggers, `EF_e24vhi_medium1`, `EF_e60_medium1`, `EF_mu24i_tight`, and `EF_mu36_tight`.
- The primary event vertex must have at least three tracks with $p_T > 400$ MeV.

Production Mode	Decay 1	Decay 2
$pp \rightarrow N^0 \Sigma^+$	$N^0 \rightarrow l^+ W^-$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow N^0 \Sigma^+$	$N^0 \rightarrow l^- W^+$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow N^0 \Sigma^+$	$N^0 \rightarrow \nu_l Z$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow N^0 \Sigma^+$	$N^0 \rightarrow \nu_l H$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow N^0 \Sigma^-$	$N^0 \rightarrow l^- W^+$	$\Sigma^- \rightarrow l^- Z$
$pp \rightarrow N^0 \Sigma^-$	$N^0 \rightarrow l^+ W^-$	$\Sigma^- \rightarrow l^- Z$
$pp \rightarrow N^0 \Sigma^-$	$N^0 \rightarrow \nu_l Z$	$\Sigma^- \rightarrow l^- Z$
$pp \rightarrow N^0 \Sigma^-$	$N^0 \rightarrow \nu_l H$	$\Sigma^- \rightarrow l^- Z$
$pp \rightarrow \Sigma^- \Sigma^+$	$\Sigma^- \rightarrow l^- Z$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow \Sigma^- \Sigma^+$	$\Sigma^- \rightarrow \nu_l W^-$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow \Sigma^- \Sigma^+$	$\Sigma^- \rightarrow l^- H$	$\Sigma^+ \rightarrow l^+ Z$
$pp \rightarrow \Sigma^- \Sigma^+$	$\Sigma^- \rightarrow l^- Z$	$\Sigma^+ \rightarrow \nu_l W^+$
$pp \rightarrow \Sigma^- \Sigma^+$	$\Sigma^- \rightarrow l^- Z$	$\Sigma^+ \rightarrow l^+ H$

Table 8.2: Decays simulated for the type III seesaw signal samples.

- Events are required to have at least three electrons or muons (eee , $ee\mu$, $\mu\mu e$, $\mu\mu\mu$) satisfying the lepton cuts (table ??).
- Events must have one Z candidate, consisting of a same-flavor, opposite-sign pair of leptons with $|m_{l+l-} - m_Z| < 10$ GeV. The Z mass is taken to be $m_Z = 91.1876$ GeV [PhysRevD.86.010001].
- Four-lepton events with two leptonic Z candidates are rejected to suppresses background from ZZ events. The efficiency loss for signal events is low, in part due to the additional factor of the leptonic branching ratio of the Z ; see tables ?? and ??.
- If an event contains more than three leptons, a unique trilepton candidate in each event is chosen as follows:
 - Choose the same-flavor, opposite-sign (SFOS) pair with invariant mass closest to m_Z .
 - Choose the third (“off- Z ”) lepton to be the closest in ΔR to the reconstructed Z four-momentum.
- For low heavy lepton mass hypotheses, $m_{L^\pm} \lesssim 200$ GeV, the Z and the third lepton will tend to be collimated. The expected background and signal ΔR distributions for selected trilepton candidate are shown in figure ?? . Imposing a cut on the $\Delta R(Z, l_3)$ of the trilepton candidates improves the signal significance, and also provides a useful control region defined by inverting the cut.

The cut on $\Delta R(Z, l_3)$ is chosen by running the cut-and-count based limit setting, described in section ?? , with different values of the ΔR cut. For cut values of $\Delta R <$

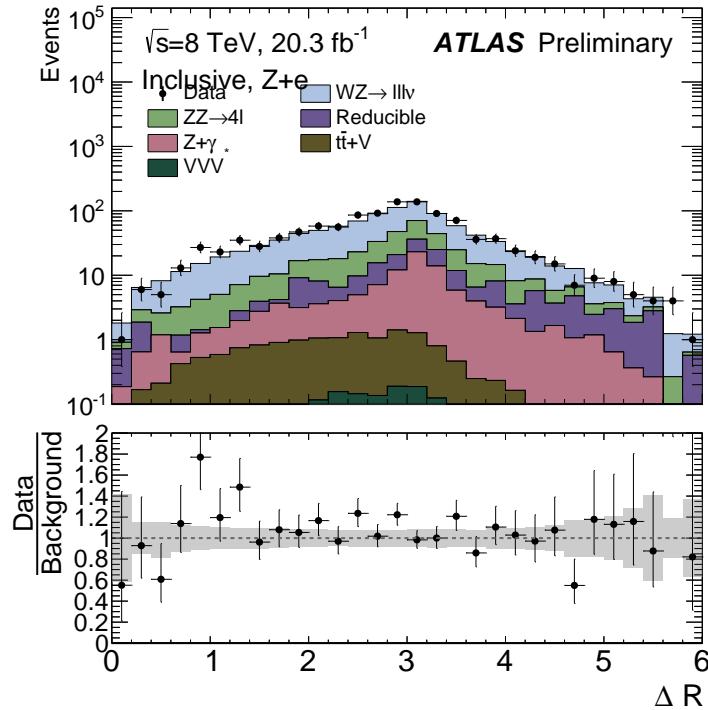


Figure 8.4: Expected $\Delta R(Z, l_3)$ distributions for background and signal in the inclusive signal region.

2.0, 2.25, 2.5, 2.75, 3.0, 3.14, 3.5, and 4.0, the expected exclusions were determined using `mclimits`, based on the expected signal and background yields in narrow mass windows around the signal mass hypotheses. For simplicity, the same $\Delta R(Z, l_3)$ cut is applied to all categories; no significant gain is observed by optimizing each category individually. The expected exclusions for different $\Delta R(Z, l_3)$ cuts are shown in figure ???. Smaller values of the cut ($\Delta R(Z, l_3) \lesssim 2.0$) perform similarly to more stringent cuts at low signal mass hypotheses, but decline at higher mass hypotheses. Cuts above $\Delta R(Z, l_3) \lesssim 3.0$ perform approximately equally. Based on these studies, and also to use the inverse of this cut to define a control region, a cut value of $\Delta R(Z, l_3) < 3.0$ is chosen.

- The variable of interest used to apply the trilepton mass constraint is $\Delta m \equiv m_{3\ell} - m_{\ell^+\ell^-}$, where the invariant mass of the two leptons associated with a Z boson decay is subtracted by the trilepton mass. This method reduces the impact of the lepton energy or momentum resolutions on the width of the resonance peak, at the cost of introducing some width due to the intrinsic width of the Z , $\Gamma_Z = 2.4952 \pm 0.0023$ GeV [pdg]. Over most of the mass range, this variable reduces the width of the signal peaks, as shown in figure ??.

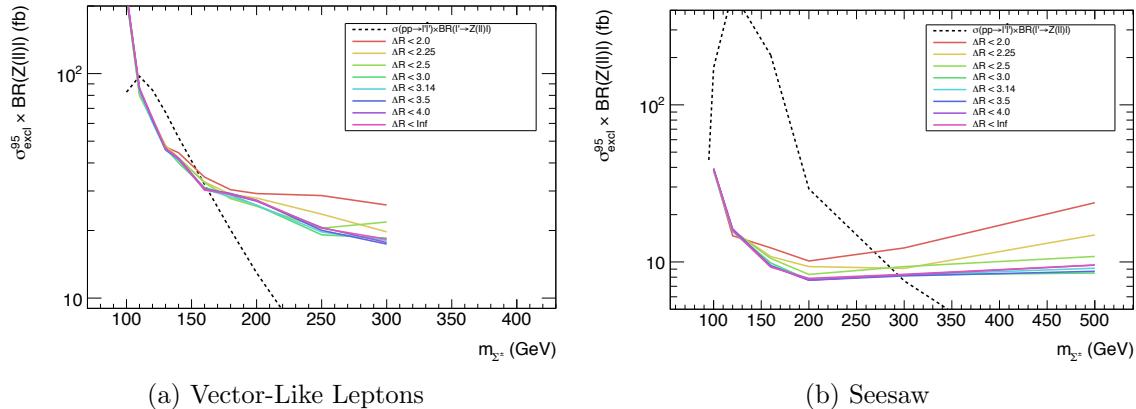


Figure 8.5: Expected upper limits on cross section times branching ratio to final states with least one trilepton resonant decay for different cuts on $\Delta R(Z, l_3)$.

m3l deltam comparison plot or table

Figure 8.6: Comparison of the full-width, half-maximum of the signal peaks between the Δm and m_{3l} variables.

After identifying the trilepton candidates, six mutually exclusive signal regions are defined. First, to target the flavor structure of the heavy lepton mixing, events are divided into two flavor channels, $Z + e$ and $Z + \mu$, depending on the flavor the off- Z lepton. Second, as the heavy leptons are pair produced, signal events will contain additional activity due to the decay of the second heavy lepton. Events are divided into three exclusive categories based on other activity in the event:

- 4ℓ : Event contains a fourth lepton, either an electron or a muon, passing the normal lepton criteria.
- $3\ell + jj$: Event contains exactly three leptons, and a dijet with $60.385 \text{ GeV} < m_{jj} < 150.9 \text{ GeV}$.
- $3\ell - \text{only}$: All other events, i.e. events with three leptons and no jet pair satisfying the invariant mass requirement.

The division into the three signal regions targeting the opposite side decay is motivated in figure ??, which shows the fraction of events with various activity from the decay of the other heavy lepton. The additional activity includes extra leptons and neutrinos, and jets from $W/Z/h$ decays. The requirement of either a fourth lepton or a hadronically decaying boson is very efficient on signal: if the second heavy fermion is charged, then every decay satisfies this requirement, while if the second heavy fermion is neutral, then every decay except for $\Sigma^0 \rightarrow Z\nu \rightarrow \nu\nu\nu$ and decays to Higgs bosons satisfies this requirement. Categorization

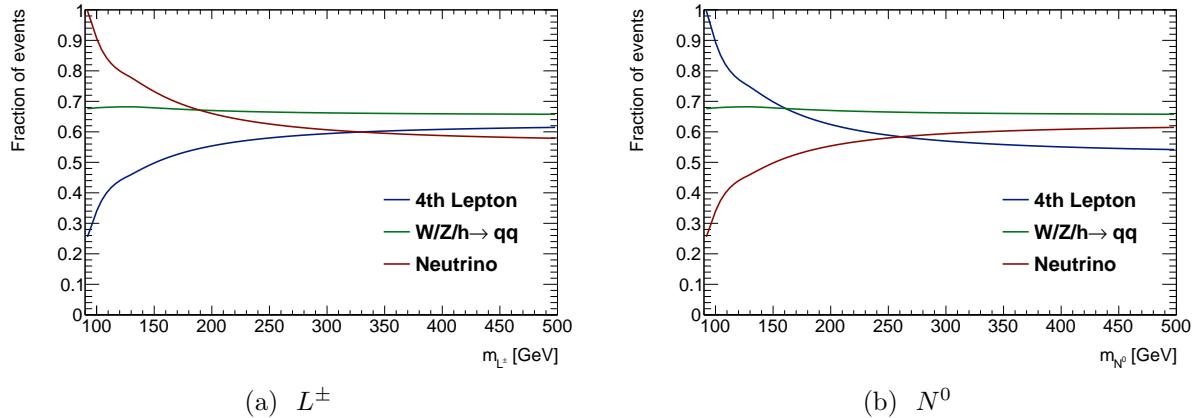


Figure 8.7: Fraction of events with various activity on the opposite side of the event, for $L^\pm\Sigma^\mp$ events at left, and $L^\pm\Sigma^0$ events at right. The left plot, showing the opposite side activity for a $L^\pm\Sigma^\mp$ final state, is identical for the seesaw and vector-like lepton model.

targeting neutrinos is less effective at separating signal from the large WZ background (see figures ??-??).

In some cases, it is useful to consider the events without dividing into these three categories, but rather only the flavor of the off- Z lepton. This is referred to as the “Inclusive” signal region below.

Performance on Fiducial Signal Events

The performance of the cuts and heavy lepton candidate selection is shown in tables ?? and ?? in terms of efficiency on fiducial events. The final efficiency for each signal region is shown as a function of signal mass m_Σ in figure ?? . The study is performed separately for the $Z + e$ and $Z + \mu$ flavor channels. The fiducial selection requires the event to have three truth-level leptons satisfying $p_T > 15$ GeV and $|\eta| < 2.5$, of which two form a Z candidate with $|m_{l^+l^-} - m_Z| < 10$ GeV, the third lepton is of the correct flavor, and the trilepton mass satisfies $|m_{3l} - m_{L^\pm}| < 5$ GeV.

Note that cut imposing a flavor requirement on the reconstructed off- Z lepton has a different efficiency between the vector-like lepton sample and the seesaw sample. This is a consequence of the initial flavor content of the samples: the vector-like lepton samples are divided into two samples with 100% branching fraction to either bachelor electron or bachelor muon, while the seesaw samples is a single mixed sample containing both electron and muon decays.

Process	Preselection	Bachelor e	$ m_{l+l-} - m_Z < 10$ GeV	$!2Z$	$\Delta R < 3.0$	$4l$	$3l + jj$	Else
VLL, 100 GeV	0.348 ± 0.008	0.318 ± 0.008	0.306 ± 0.008	0.303 ± 0.008	0.298 ± 0.008	0.054 ± 0.004	0.096 ± 0.005	0.147 ± 0.006
VLL, 110 GeV	0.478 ± 0.006	0.435 ± 0.006	0.421 ± 0.006	0.417 ± 0.006	0.401 ± 0.006	0.076 ± 0.003	0.11 ± 0.004	0.215 ± 0.005
VLL, 120 GeV	0.535 ± 0.006	0.491 ± 0.006	0.467 ± 0.006	0.457 ± 0.006	0.424 ± 0.006	0.104 ± 0.004	0.117 ± 0.004	0.204 ± 0.005
VLL, 130 GeV	0.604 ± 0.006	0.539 ± 0.006	0.509 ± 0.006	0.494 ± 0.006	0.445 ± 0.006	0.125 ± 0.004	0.127 ± 0.004	0.194 ± 0.005
VLL, 140 GeV	0.643 ± 0.006	0.578 ± 0.006	0.559 ± 0.006	0.538 ± 0.006	0.491 ± 0.006	0.153 ± 0.004	0.164 ± 0.004	0.174 ± 0.005
VLL, 160 GeV	0.663 ± 0.006	0.593 ± 0.006	0.572 ± 0.006	0.552 ± 0.006	0.499 ± 0.006	0.175 ± 0.005	0.148 ± 0.004	0.176 ± 0.005
VLL, 180 GeV	0.698 ± 0.006	0.62 ± 0.006	0.594 ± 0.006	0.573 ± 0.007	0.505 ± 0.007	0.197 ± 0.005	0.163 ± 0.005	0.144 ± 0.005
VLL, 200 GeV	0.707 ± 0.006	0.624 ± 0.007	0.596 ± 0.007	0.577 ± 0.007	0.52 ± 0.007	0.209 ± 0.006	0.167 ± 0.005	0.145 ± 0.005
VLL, 250 GeV	0.751 ± 0.007	0.654 ± 0.007	0.626 ± 0.007	0.615 ± 0.007	0.548 ± 0.008	0.245 ± 0.007	0.152 ± 0.005	0.151 ± 0.005
VLL, 300 GeV	0.771 ± 0.007	0.662 ± 0.008	0.63 ± 0.008	0.625 ± 0.008	0.547 ± 0.008	0.258 ± 0.007	0.169 ± 0.006	0.121 ± 0.005
VLL, 400 GeV	0.776 ± 0.007	0.669 ± 0.008	0.642 ± 0.008	0.641 ± 0.008	0.562 ± 0.008	0.277 ± 0.007	0.181 ± 0.006	0.103 ± 0.005
Seesaw, 100 GeV	0.381 ± 0.006	0.351 ± 0.005	0.336 ± 0.005	0.334 ± 0.005	0.308 ± 0.005	0.08 ± 0.003	0.096 ± 0.003	0.131 ± 0.004
Seesaw, 120 GeV	0.517 ± 0.004	0.474 ± 0.004	0.45 ± 0.004	0.44 ± 0.004	0.403 ± 0.004	0.128 ± 0.003	0.128 ± 0.003	0.148 ± 0.003
Seesaw, 160 GeV	0.611 ± 0.005	0.548 ± 0.005	0.532 ± 0.005	0.516 ± 0.005	0.464 ± 0.005	0.167 ± 0.004	0.141 ± 0.003	0.157 ± 0.003
Seesaw, 200 GeV	0.657 ± 0.005	0.591 ± 0.005	0.574 ± 0.005	0.559 ± 0.005	0.5 ± 0.005	0.203 ± 0.004	0.146 ± 0.003	0.152 ± 0.004
Seesaw, 250 GeV	0.671 ± 0.007	0.605 ± 0.007	0.592 ± 0.007	0.582 ± 0.007	0.505 ± 0.007	0.221 ± 0.006	0.14 ± 0.005	0.144 ± 0.005
Seesaw, 300 GeV	0.729 ± 0.007	0.649 ± 0.007	0.633 ± 0.007	0.626 ± 0.007	0.541 ± 0.008	0.23 ± 0.006	0.167 ± 0.006	0.144 ± 0.005
Seesaw, 350 GeV	0.72 ± 0.007	0.651 ± 0.008	0.638 ± 0.008	0.635 ± 0.008	0.553 ± 0.008	0.251 ± 0.007	0.157 ± 0.006	0.146 ± 0.006
Seesaw, 400 GeV	0.717 ± 0.007	0.641 ± 0.008	0.627 ± 0.008	0.623 ± 0.008	0.529 ± 0.008	0.243 ± 0.007	0.15 ± 0.006	0.136 ± 0.005
Seesaw, 450 GeV	0.762 ± 0.008	0.67 ± 0.009	0.66 ± 0.009	0.657 ± 0.009	0.554 ± 0.009	0.267 ± 0.008	0.184 ± 0.007	0.103 ± 0.006
Seesaw, 500 GeV	0.72 ± 0.007	0.636 ± 0.008	0.611 ± 0.008	0.608 ± 0.008	0.525 ± 0.008	0.24 ± 0.007	0.151 ± 0.006	0.134 ± 0.005

Table 8.3: Cut efficiencies for fiducial events at various stages in the cutflow for the $Z + e$ signal regions. Only statistical uncertainties due to finite Monte Carlo statistics are shown. The preselection cut requires three selected leptons, with one same-flavor opposite-sign pair, as well as the general event selection cuts listed above.

8.3 Systematic Uncertainties

The signal predictions and background estimates are assigned systematic uncertainties from several sources. In approximate order of significance, these are:

- **Diboson shape uncertainty:** A systematic uncertainty is assigned to the modeling of the diboson backgrounds by comparing SHERPA to POWHEG. SHERPA is used as a central value, with uncertainty given by the symmetric difference between SHERPA and POWHEG. As described below in section ??, the shapes for the 4L and Rest signal regions are taken from the inclusive signal region, so the corresponding uncertainties are derived only from the inclusive and 3L+dijet signal regions. The comparison for the WZ background is shown in figure ???. The implementation of this uncertainty in the fit-based limit section is discussed in more detail in section ??; briefly, the uncertainty is included as a gaussian-distributed nuisance parameter interpolating between the two fits estimated from the two generators.

In the case of the ZZ background, this procedure is complicated by the fact that the POWHEG sample is filtered to remove events with same-flavor, opposite-sign lepton pairs with $m_{l+l-} < 4$ GeV (mll4). In order to compare the samples in a common phase space, the mll4 filter is emulated on the SHERPA sample by rejected events that contain a same-flavor, opposite-sign pair of truth leptons with status==3, with $m_{l+l-} < 4$ GeV. SHERPA, SHERPA with mll4 filter, and POWHEG are compared in figure ???. The events from the POWHEG sample are weighted using the ratio of SHERPA to SHERPA+mll4. The final comparison of background shapes is shown in figure ??.

Process	Preselection	Bachelor μ	$ m_{l+l-} - m_Z < 10 \text{ GeV}$	$!2Z$	$\Delta R < 3.0$	$4l$	$3l + jj$	Else
VLL, 100 GeV	0.52 ± 0.007	0.508 ± 0.007	0.472 ± 0.007	0.461 ± 0.007	0.456 ± 0.007	0.064 ± 0.004	0.14 ± 0.005	0.252 ± 0.006
VLL, 110 GeV	0.61 ± 0.004	0.593 ± 0.004	0.568 ± 0.004	0.554 ± 0.004	0.536 ± 0.004	0.119 ± 0.003	0.142 ± 0.003	0.275 ± 0.004
VLL, 120 GeV	0.67 ± 0.003	0.648 ± 0.004	0.617 ± 0.004	0.596 ± 0.004	0.561 ± 0.004	0.149 ± 0.003	0.139 ± 0.003	0.273 ± 0.003
VLL, 130 GeV	0.702 ± 0.003	0.676 ± 0.003	0.642 ± 0.003	0.614 ± 0.003	0.569 ± 0.003	0.163 ± 0.003	0.143 ± 0.002	0.263 ± 0.003
VLL, 140 GeV	0.723 ± 0.003	0.701 ± 0.003	0.667 ± 0.003	0.634 ± 0.003	0.585 ± 0.003	0.173 ± 0.003	0.156 ± 0.002	0.256 ± 0.003
VLL, 160 GeV	0.752 ± 0.003	0.727 ± 0.003	0.685 ± 0.003	0.65 ± 0.003	0.592 ± 0.003	0.212 ± 0.003	0.156 ± 0.002	0.223 ± 0.003
VLL, 180 GeV	0.77 ± 0.003	0.746 ± 0.003	0.69 ± 0.003	0.659 ± 0.003	0.593 ± 0.003	0.227 ± 0.003	0.161 ± 0.002	0.205 ± 0.003
VLL, 200 GeV	0.779 ± 0.003	0.751 ± 0.003	0.69 ± 0.003	0.66 ± 0.003	0.584 ± 0.003	0.239 ± 0.003	0.161 ± 0.002	0.184 ± 0.002
VLL, 250 GeV	0.786 ± 0.003	0.752 ± 0.003	0.676 ± 0.003	0.657 ± 0.003	0.579 ± 0.003	0.268 ± 0.003	0.162 ± 0.002	0.149 ± 0.002
VLL, 300 GeV	0.789 ± 0.002	0.76 ± 0.003	0.68 ± 0.003	0.667 ± 0.003	0.578 ± 0.003	0.288 ± 0.003	0.163 ± 0.002	0.127 ± 0.002
VLL, 400 GeV	0.787 ± 0.002	0.753 ± 0.003	0.665 ± 0.003	0.656 ± 0.003	0.563 ± 0.003	0.292 ± 0.003	0.166 ± 0.002	0.106 ± 0.002
Seesaw, 100 GeV	0.535 ± 0.006	0.521 ± 0.006	0.484 ± 0.006	0.471 ± 0.006	0.46 ± 0.006	0.192 ± 0.005	0.121 ± 0.004	0.147 ± 0.004
Seesaw, 120 GeV	0.677 ± 0.004	0.66 ± 0.004	0.613 ± 0.004	0.589 ± 0.004	0.546 ± 0.004	0.212 ± 0.004	0.119 ± 0.003	0.215 ± 0.004
Seesaw, 160 GeV	0.745 ± 0.004	0.723 ± 0.004	0.685 ± 0.004	0.658 ± 0.004	0.595 ± 0.005	0.252 ± 0.004	0.14 ± 0.003	0.202 ± 0.004
Seesaw, 200 GeV	0.743 ± 0.004	0.717 ± 0.005	0.668 ± 0.005	0.651 ± 0.005	0.588 ± 0.005	0.263 ± 0.004	0.152 ± 0.004	0.173 ± 0.004
Seesaw, 250 GeV	0.761 ± 0.006	0.735 ± 0.006	0.673 ± 0.007	0.662 ± 0.007	0.585 ± 0.007	0.284 ± 0.006	0.144 ± 0.005	0.157 ± 0.005
Seesaw, 300 GeV	0.785 ± 0.006	0.764 ± 0.006	0.695 ± 0.007	0.688 ± 0.007	0.601 ± 0.007	0.295 ± 0.007	0.171 ± 0.006	0.135 ± 0.005
Seesaw, 350 GeV	0.78 ± 0.006	0.757 ± 0.007	0.688 ± 0.007	0.681 ± 0.007	0.59 ± 0.008	0.294 ± 0.007	0.155 ± 0.006	0.141 ± 0.005
Seesaw, 400 GeV	0.753 ± 0.007	0.727 ± 0.007	0.642 ± 0.008	0.634 ± 0.008	0.544 ± 0.008	0.277 ± 0.007	0.15 ± 0.006	0.118 ± 0.005
Seesaw, 450 GeV	0.762 ± 0.008	0.738 ± 0.008	0.658 ± 0.008	0.653 ± 0.009	0.555 ± 0.009	0.285 ± 0.008	0.152 ± 0.006	0.119 ± 0.006
Seesaw, 500 GeV	0.755 ± 0.007	0.729 ± 0.007	0.654 ± 0.008	0.65 ± 0.008	0.551 ± 0.008	0.275 ± 0.007	0.155 ± 0.006	0.122 ± 0.005

Table 8.4: Cut efficiencies for fiducial events at various stages in the cutflow for the $Z + \mu$ signal regions. Only statistical uncertainties due to finite Monte Carlo statistics are shown. The preselection cut requires three selected leptons, with one same-flavor opposite-sign pair, as well as the general event selection cuts listed above.

- **Monte Carlo statistics and Fit Parameter Uncertainties:** as described below in section ??, the limit setting is performed using fits with analytical functions to the Monte Carlo background predictions. Hence the uncertainty due to the finite statistics of the background samples is represented in the uncertainties on the fit parameters, returned by the fitting code. This uncertainty is discussed in more detail in section ??.
- **Monte Carlo sample normalizations:** The Monte Carlo samples used for the irreducible background prediction are normalized to the measured luminosity of the data using theoretical cross sections. The uncertainty on the luminosity is 2.8% [luminosity]. The uncertainties on the cross sections for WZ , ZZ , and $t\bar{t} + V$ are listed in table ???. The WZ and ZZ uncertainties are taken from [DeViveiros:1670929], and are based off a comparison of SHERPA with VBFNLO. A cross section uncertainty is not assigned to the $Z + \gamma$ backgrounds; instead, as described next, a large uncertainty is assigned due to applying scale factors to correct the simulated rate of photon conversions. Similarly, an uncertainty is not assigned for the $VVV^{(*)}$ cross section, due to its small contribution to the signal regions.

In the fit-based limit setting procedure, the uncertainty is implemented as a constraint on the normalization of the corresponding sample: the normalizations of the dominant diboson samples are allowed to float within the quadrature sum of all the normalization uncertainties. The normalizations of the small $t\bar{t} + V$ and $VVV^{(*)}$ backgrounds are fixed due to their comparatively small magnitude, which the fit is unable to resolve.

- **Charge flip scale factors:** The rate of trilepton events due to $Z + \gamma$, where the

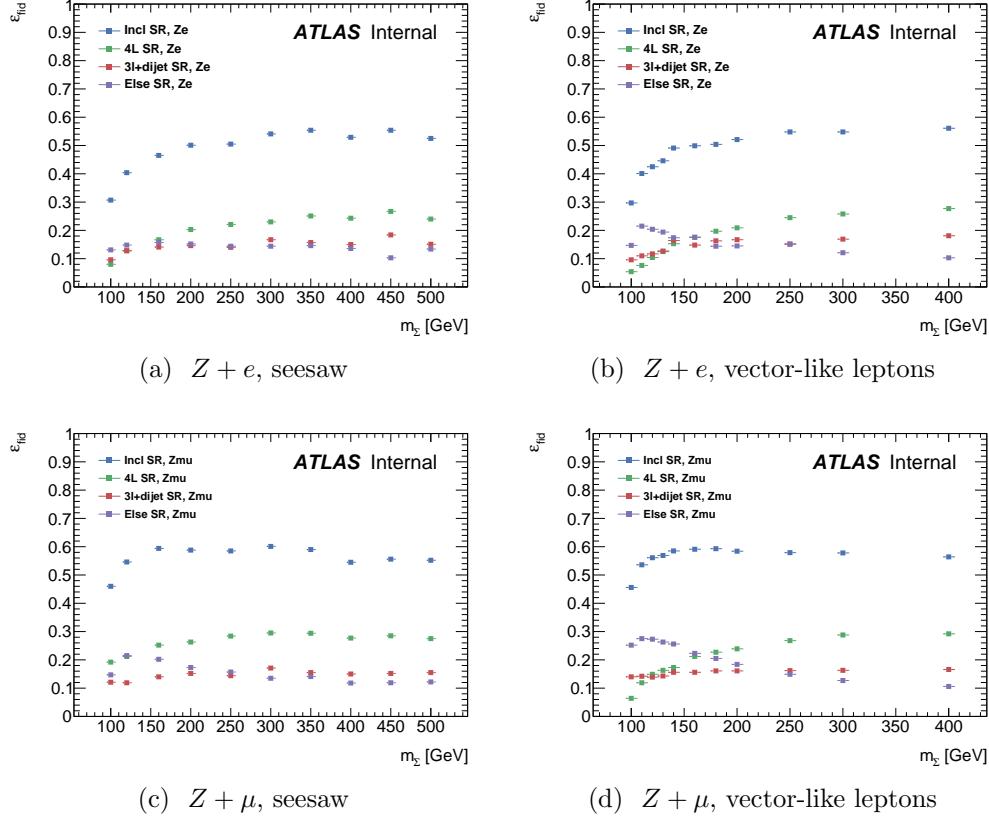


Figure 8.8: Efficiencies of cuts defining each signal region on fiducial $Z(l\bar{l})l$ events at each simulated mass point. For the $Z + e$ signal regions, the branching fraction is set to be $BR(\Sigma \rightarrow X + e/\nu_e) = 100\%$; similarly, for the $Z + \mu$ signal regions, the branching fraction is set to be $BR(\Sigma \rightarrow X + \mu/\nu_\mu) = 100\%$.

Process	Cross Section Uncertainty	Reference
WZ	7.6%	[DeViveiros:1670929]
ZZ	4.3%	[DeViveiros:1670929]
$t\bar{t} + V$	22%	[ttV]

Table 8.5: Cross sections and uncertainties for the Monte Carlo samples used for irreducible background estimation.

photon converts asymmetrically and is reconstructed as an electron, is observed to be overestimated in Monte Carlo. Scale factors are applied following the charge flip likelihood estimation method from the same-sign dilepton analysis [DeViveiros:1670929], and an uncertainty of 30% is assigned to this background sample.

- **Fake factor method:** The systematic uncertainty on the fake factors is described in

sections ?? and ?? . They range from 20%-30% for electrons and 25%-50% for muons.

- **Luminosity:** The uncertainty on the integrated luminosity in 2012 is 2.8%.
- **Lepton scale factors:** Scale factors are applied to equalize the lepton trigger, reconstruction, and identification efficiencies between data and simulation. The corrections are implemented with official performance group tools:

`MuonEfficiencyCorrections` [`MuonEfficiencyScaleFactors`] and

`TElectronEfficiencyCorrectionTool` [`ElectronEfficiencyScaleFactors`].

Systematic uncertainties on the scale factors are applied according to the official recommendations.

- **Electron energy corrections:** To improve agreement in electron energy between data and simulation, the energy is scaled in data and smeared in Monte Carlo according to the recommendations at [`ElectronEnergyCorrectionsGEO20`, `ElectronEnergyCorrectionsGEO21`]. The corrections and corresponding systematic uncertainties are provided by the `ElectronPhotonFourMomentumCorrection` package.
- **Muon mommentum corrections:** The muon momenta are smeared in simulation according to the recommendations at [`MuonMomentumCorrections`]. The corrections are derived by “comparing the reconstructed muon momentum in experimental and simulated data,” using J/ψ and Z decays.
- **Jet energy scale and resolution:** Systematic uncertainties on the jet energy scale and resolution are applied following the recommendations from the Jet/ETMiss group [`JetEtmissRecommendations`], and are taken from the package `JetUncertainties` 00-08-07. The jet energy uncertainties have a small impact on this analysis, only affecting the normalization in the 3L+dijet signal region when the dijet mass is pushed in or out of the window $m_W - 20 \text{ GeV} < m_{jj} < m_h + 25 \text{ GeV}$.

Figures ?? and ?? show the fractional uncertainty due to each source of uncertainty in 20 GeV bins for each signal region, as well as the total systematic uncertainty and expected statistical uncertainty. Note that the uncertainties shown are bin-by-bin uncertainties on the Monte Carlo predictions, and do not necessarily correspond to the final uncertainty after fitting the background shapes; for example, after fitting, the Monte Carlo statistical uncertainties are contained in the uncertainties on the fit parameters. Further, the Monte Carlo statistical uncertainties are overrepresented, as they are also contained in the WZ/ZZ shape uncertainties derived from the comparison of two statistically independent samples. Finally, due to the fact that the shapes for the 4L and 3L+dijet signal regions are taken from the inclusive signal region, the shape uncertainties derived in these categories are not used in the analysis.

The impact of each uncertainty in each signal region, in terms of fractional uncertainty on the total normalization, is shown for background in figure ??, and some example signal points in figures ??–??. The complete set of tables of signal systematics is in appendix ??.

	$Z + e$				$Z + \mu$			
	Total	$4l$ SR	$3l + jj$ SR	$3l$ -only SR	Total	$4l$ SR	$3l + jj$ SR	$3l$ -only SR
σ_{ZZ}	0.9	3.9	0.9	0.8	0.7	3.8	0.4	0.6
σ_{WZ}	4.9	0.1	4.6	5.1	5.3	—	4.9	5.6
σ_{ttV}	0.4	1.5	2.9	0.1	0.4	0.9	2.9	0.1
Luminosity	2.6	2.8	2.7	2.6	2.5	2.6	2.4	2.5
γ conv. SFs	2.0	—	1.2	2.1	—	—	—	—
ℓ efficiency	1.6	1.8	1.6	1.6	0.9	0.9	1.0	0.9
e reducible SFs	1.7	—	0.5	1.9	0.1	0.9	0.1	0.1
μ reducible SFs	0.3	—	0.3	0.3	3.4	0.6	4.3	3.5
JES/JER	0.1	$^{+0.2}_{-0.0}$	3.3	0.4	0.1	0.2	3.2	0.5
LES/LER	0.6	$^{+1.0}_{-0.3}$	$^{+0.3}_{-1.2}$	$^{+0.7}_{-0.4}$	0.1	0.1	0.2	0.1
MC Statistics	2.4	5.0	4.3	2.6	1.2	3.5	2.8	1.2
Total	6.9	7.4	8.4	7.1	7.0	6.0	8.7	7.2

Table 8.6: The impact of significant sources of uncertainty on the background prediction in each signal region, in terms of percent of the total background normalization.

8.4 Background Validation

For each flavor channel, four validation regions are defined to test the background predictions from the Monte Carlo simulation samples and the fake factor procedure. Each region starts from the set of events with three selected electrons or muons, two of which form an OSSF pair of leptons.

The high ΔR region contains events with large separation between the Z candidate momentum and the off- Z lepton momentum, $\Delta R(Z, \ell_3) > 3$. Events are also required to have exactly three leptons, and to satisfy $\Delta m < 200$ GeV – m_Z , to limit signal contamination from hypotheses with larger m_{L^\pm} where the Z and the off- Z lepton are less collimated. The region has a similar background composition to the signal regions with three leptons, and tests the WZ , ZZ , and reducible background predictions.

	$Z + e$				$Z + \mu$			
	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR
Luminosity	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
l scale factors	1.7	1.7	1.7	1.6	1.1	1.1	1.1	1.0
MC Statistics	2.1	4.5	3.7	3.0	1.4	2.7	2.7	2.2
JES/JER	0.0	0.2	3.6	3.0	0.2	0.3	3.0	2.2
LES/LER	0.1	0.4	0.1	0.1	0.1	0.2	0.2	0.2
Total	3.9	5.6	6.1	5.3	3.3	4.0	5.1	4.3

Table 8.7: The impact of different sources of systematic uncertainty on the signal prediction for the type III seesaw model with $m_\Sigma = 160$ GeV, in terms of percent of the total signal normalization.

	$Z + e$				$Z + \mu$			
	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR
Luminosity	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
l scale factors	1.8	1.7	1.8	1.9	1.2	1.2	1.2	1.2
MC Statistics	3.0	5.7	4.9	4.7	2.2	3.7	3.8	3.6
JES/JER	0.1	—	1.6	1.7	0.2	0.3	2.0	2.3
LES/LER	0.4	0.1	1.0	0.4	0.3	0.0	0.9	0.3
Total	4.5	6.5	6.2	6.0	3.8	4.8	5.3	5.3

Table 8.8: The impact of different sources of systematic uncertainty on the signal prediction for the type III seesaw model with $m_\Sigma = 300$ GeV, in terms of percent of the total signal normalization.

	$Z + e$				$Z + \mu$			
	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR	Incl SR	$4l$ SR	$3l + jj$ SR	$3l$ -only SR
Luminosity	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
l scale factors	1.8	1.8	1.8	2.0	1.3	1.4	1.3	1.3
MC Statistics	3.0	5.5	4.8	4.8	2.4	4.0	3.9	4.1
JES/JER	—	—	1.7	1.9	0.3	0.5	1.8	2.5
LES/LER	0.4	0.5	2.1	1.7	0.4	1.6	2.8	2.2
Total	4.5	6.4	6.4	6.4	3.9	5.4	6.0	6.1

Table 8.9: The impact of different sources of systematic uncertainty on the signal prediction for the type III seesaw model with $m_\Sigma = 500$ GeV, in terms of percent of the total signal normalization.

The \underline{ZZ} region contains events with two OSSF lepton pairs with invariant mass within 10 GeV of m_Z . The region tests the ZZ background prediction.

The off-Z region contains events with exactly three leptons, with an OSSF pair within 50 GeV of m_Z , but no OSSF pairs with invariant mass within 20 GeV of m_Z . As the event does not contain a Z candidate, the trilepton selection is modified: the highest mass OSSF pair is selected, and the third lepton is chosen to be the remaining lepton. In the $Z + e$ flavor channel, this region tests the $Z + \gamma$ background prediction. In the $Z + \mu$ channel, the region is dominated by ZZ , where one lepton is not selected, and the other two leptons originate from an off-shell Z^*/γ^* .

The \underline{WZ} region contains events with exactly three leptons, two of which form an OSSF pair within 10 GeV of m_Z , and the third of which satisfies $40 \text{ GeV} < m_T^W < 90 \text{ GeV}$. Additionally, the validation region requires $40 \text{ GeV} < E_T^{\text{miss}} < 100 \text{ GeV}$ and zero jets, which suppresses signal contamination.

The definition of each validation region and the backgrounds tested are summarized in table ???. The total number of observed and predicted events in each region is shown in table ??, and the corresponding Δm distributions are shown in figure ???. Good agreement is seen in most validation regions, with the observed and predicted normalizations agreeing to better than 1.5σ . A deficit of data with respect to the background prediction is seen in the off- Z , $Z + \mu$ validation region, corresponding to 2.3σ . The region is dominated by contributions from ZZ , where only three leptons pass the selection requirements and no same-flavour, opposite-sign lepton pair is reconstructed with invariant mass within 20 GeV of m_Z .

Control Region	Definition	Background Tested
High ΔR	$\Delta R(Z, l_3) > 3.0$ $\Delta m < 200 \text{ GeV} - m_Z$	WZ , ZZ , and reducible
Off Z	Reject events with $ m_u - m_Z < 20 \text{ GeV}$	$Z + \gamma$ ($Z + e$ events) and ZZ ($Z + \mu$ events)
Two Z	Require 2 Z candidates ($ m_u - m_Z < 10 \text{ GeV}$)	ZZ
WZ	3 leptons, 0 jets, $40 \text{ GeV} < m_T^W < 90 \text{ GeV}$, $40 \text{ GeV} < E_T^{\text{miss}} < 100 \text{ GeV}$	WZ

Table 8.10: Definitions and targeted backgrounds of the four validation regions.

Flavor Channel	Control Region	Data	Background Prediction	$\frac{\text{Data} - \text{Bkgd}}{\sigma_{\text{bkgd}}}$
$Z + e$	High ΔR	239	$239.16 \pm 15.47^{\text{(stat)}} \pm 13.66^{\text{(syst)}}$	-0.01
$Z + e$	OffZ	360	$348.81 \pm 18.68^{\text{(stat)}} \pm 44.21^{\text{(syst)}}$	0.23
$Z + e$	TwoZ	39	$37.29 \pm 6.11^{\text{(stat)}} \pm 2.14^{\text{(syst)}}$	0.26
$Z + e$	WZ	140	$133.47 \pm 11.56^{\text{(stat)}} \pm 10.29^{\text{(syst)}}$	0.42
$Z + \mu$	High ΔR	302	$301.00 \pm 17.35^{\text{(stat)}} \pm 12.33^{\text{(syst)}}$	0.05
$Z + \mu$	OffZ	163	$200.25 \pm 14.16^{\text{(stat)}} \pm 7.77^{\text{(syst)}}$	-2.31
$Z + \mu$	TwoZ	74	$63.4 \pm 7.97^{\text{(stat)}} \pm 3.43^{\text{(syst)}}$	1.22
$Z + \mu$	WZ	222	$192.96 \pm 13.90^{\text{(stat)}} \pm 14.42^{\text{(syst)}}$	1.45

Table 8.11: Summary of number of events observed and predicted for each control region.

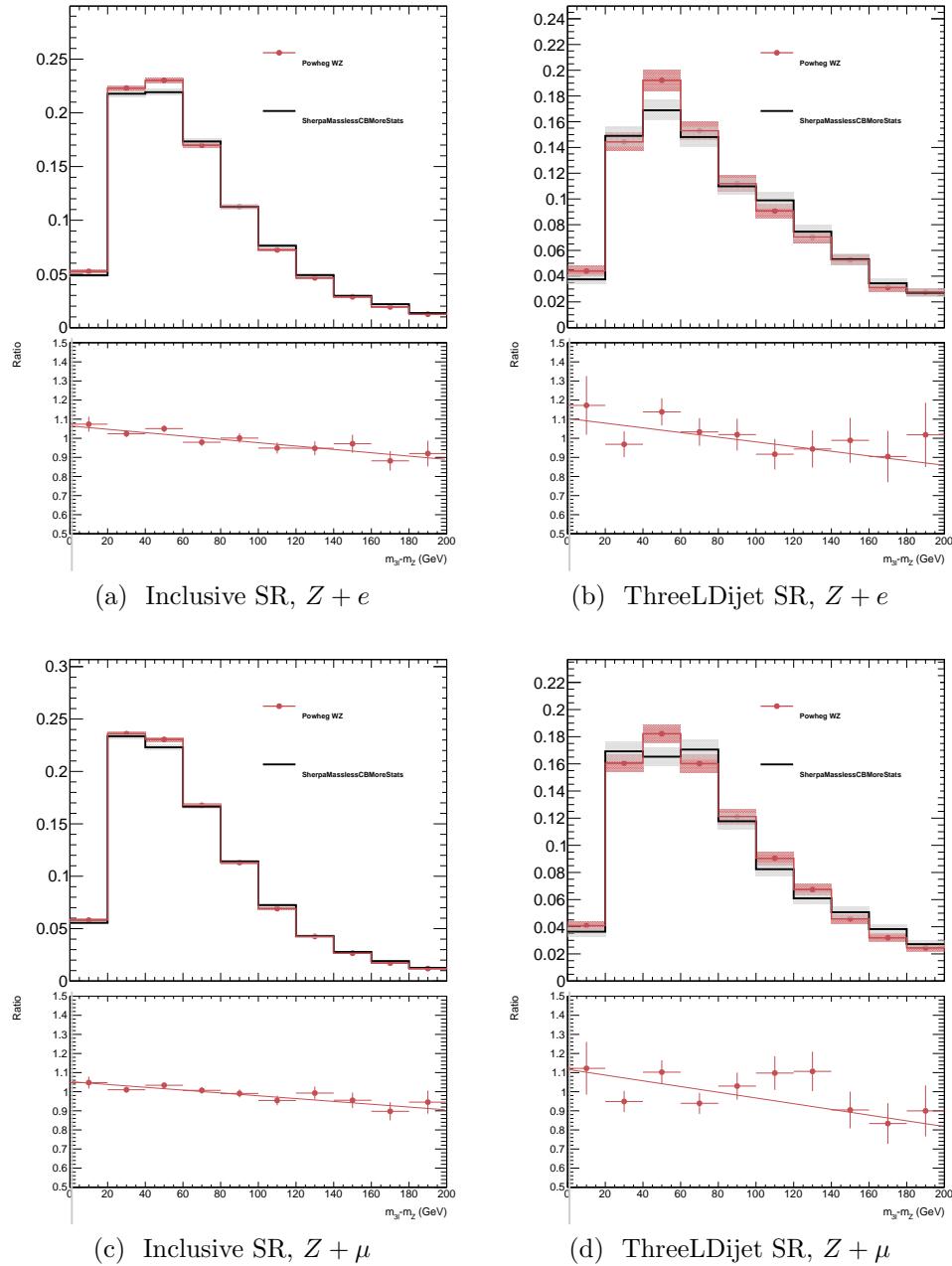


Figure 8.9: Comparison of the Δm distributions between POWHEG and SHERPA for the WZ backgrounds. The 4L signal region is omitted due to the negligible number of WZ events with four leptons.

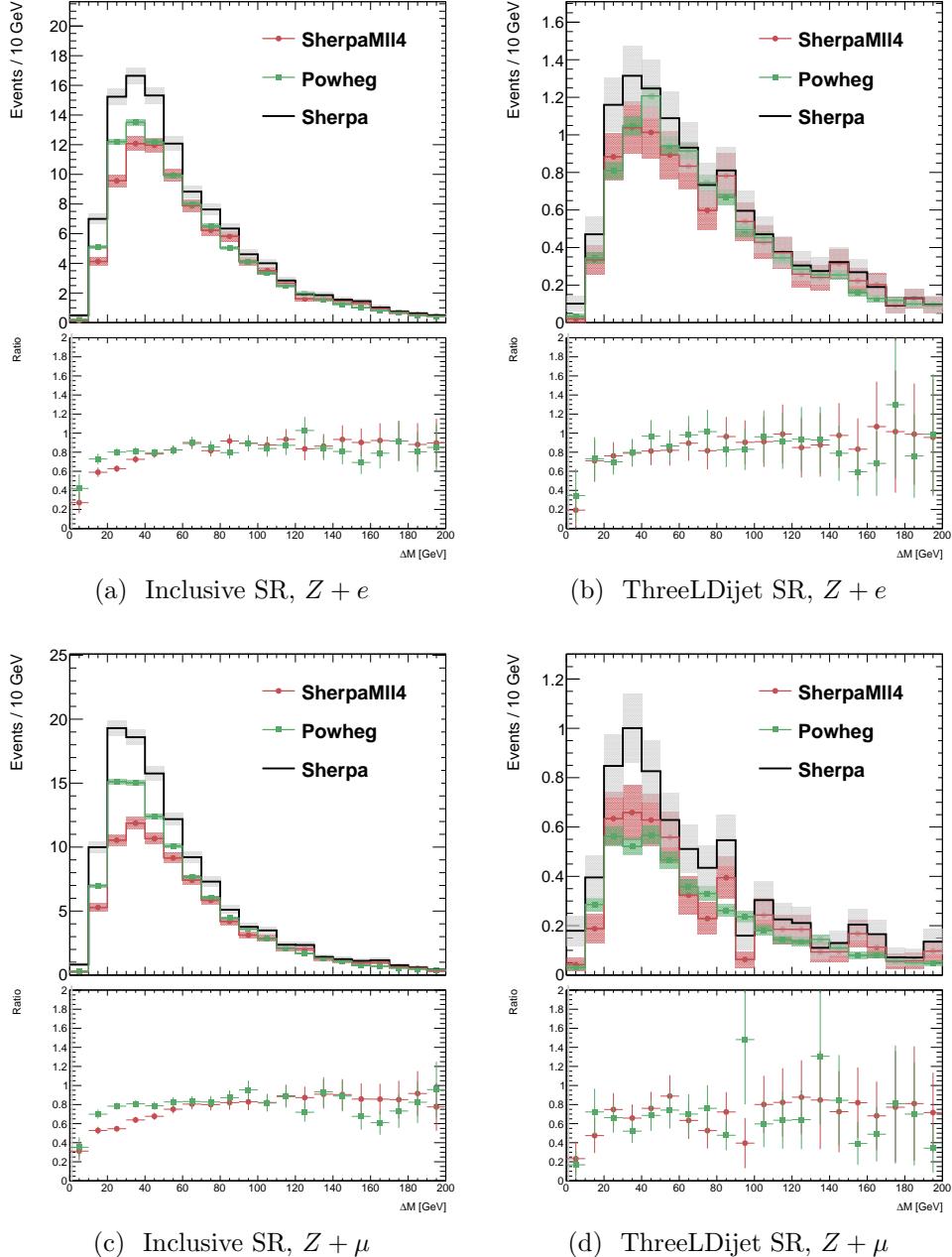


Figure 8.10: Comparison of SHERPA, SHERPA with mll4 filter, and POWHEG.

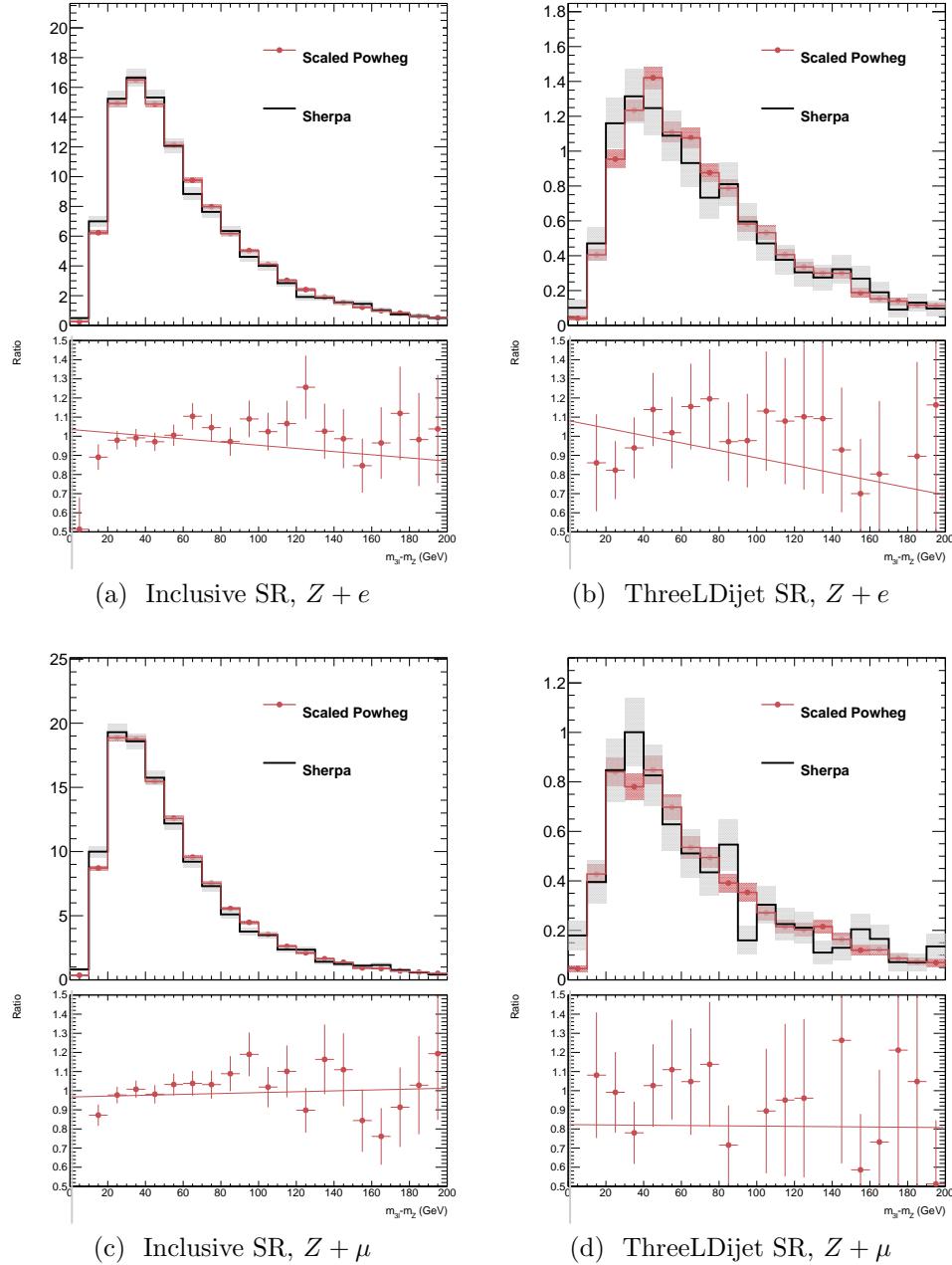


Figure 8.11: Comparison of the Δm distributions between POWHEG and SHERPA for the ZZ backgrounds. The POWHEG events are weighted to account for the mll4 filter.

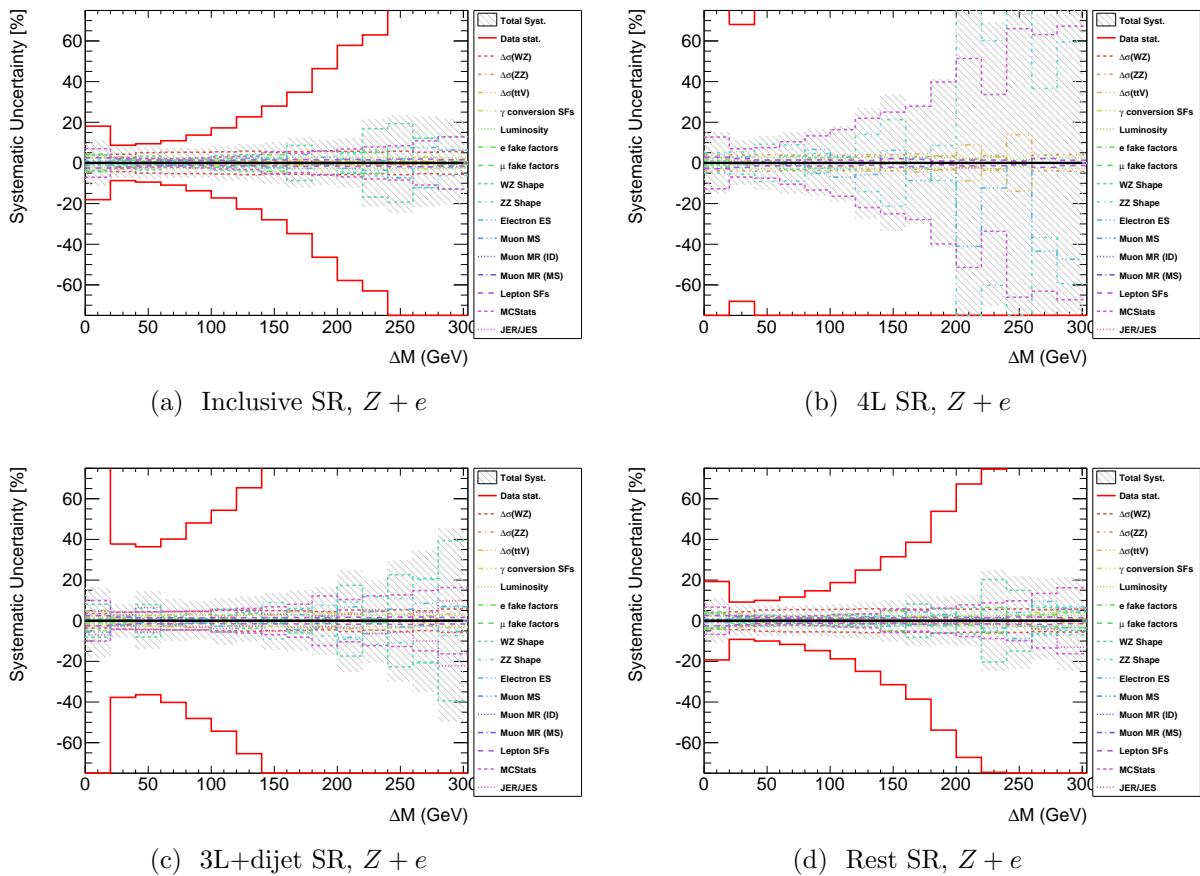


Figure 8.12: Systematics summary plots for each signal region in the $Z + e$ flavor channel. The contribution from each source of systematic uncertainty is shown in 20 GeV bins, along with the total systematic uncertainty and the expected statistical uncertainty. Note that these uncertainties reflect bin-by-bin uncertainties on the Monte Carlo predictions, and do not necessarily correspond to the final uncertainty after fitting the background shapes.

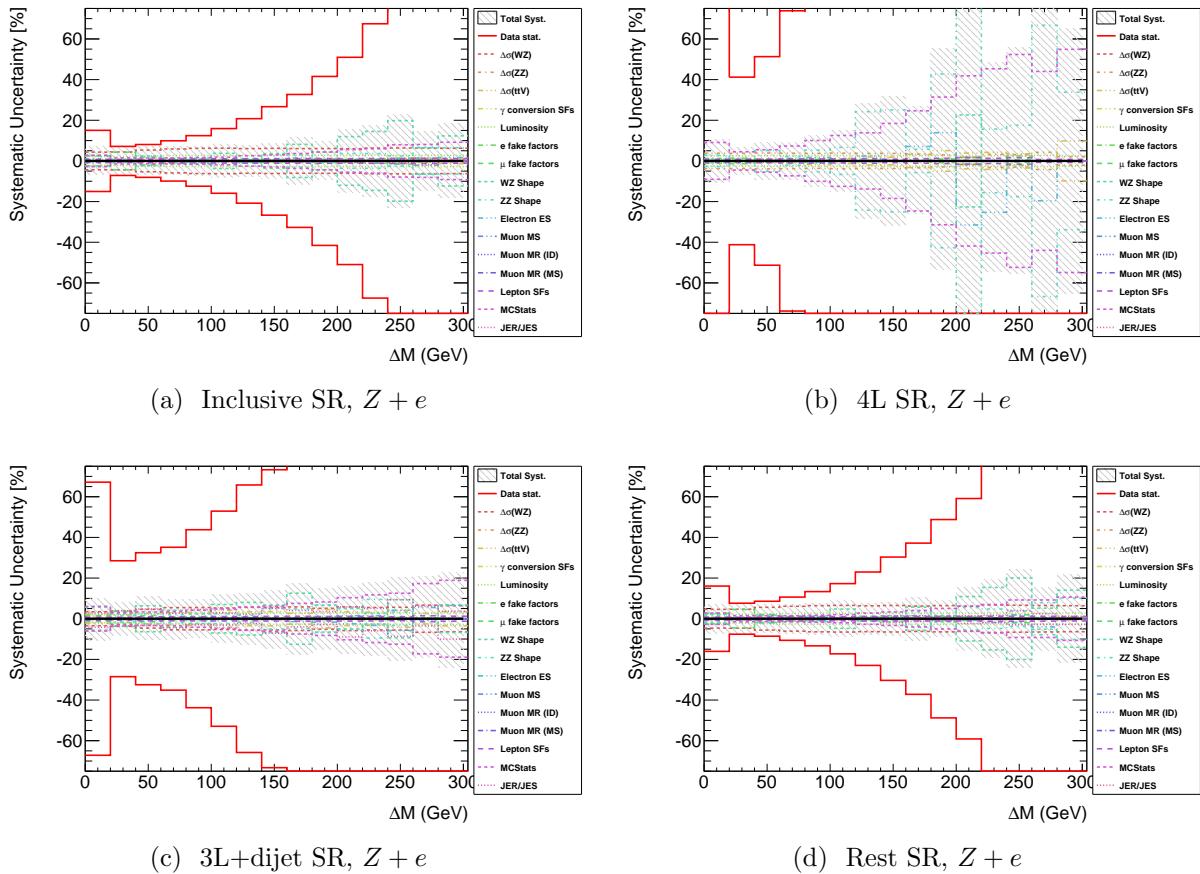


Figure 8.13: Systematics summary plots for each signal region in the $Z + \mu$ flavor channel. The contribution from each source of systematic uncertainty is shown in 20 GeV bins, along with the total systematic uncertainty and the expected statistical uncertainty. Note that these uncertainties reflect bin-by-bin uncertainties on the Monte Carlo predictions, and do not necessarily correspond to the final uncertainty after fitting the background shapes.

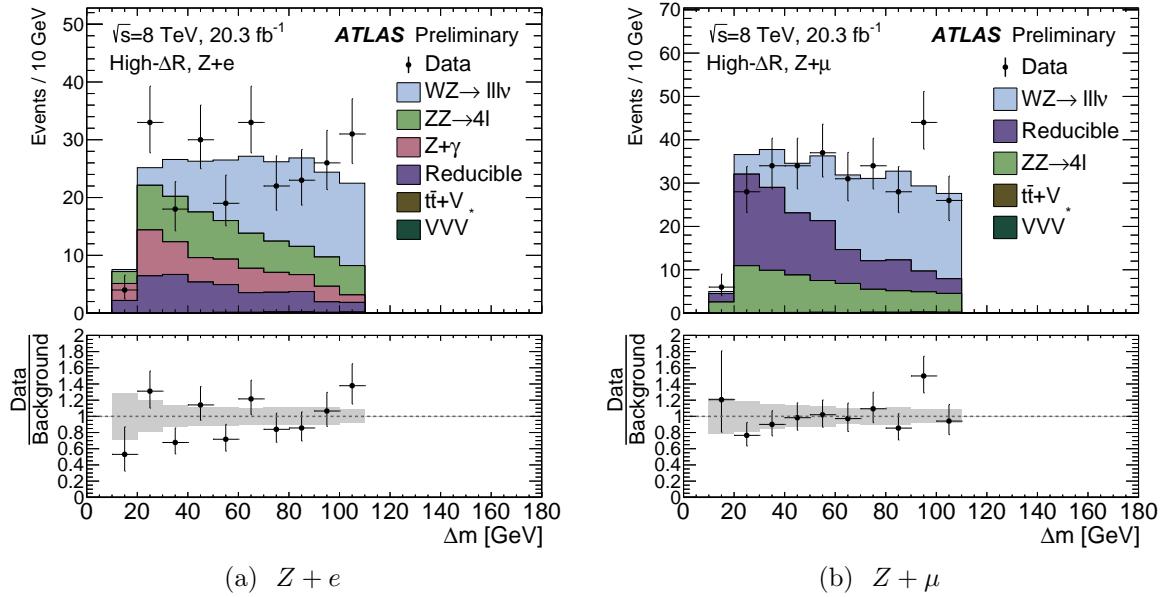


Figure 8.14: Δm distributions for the high ΔR validation regions.

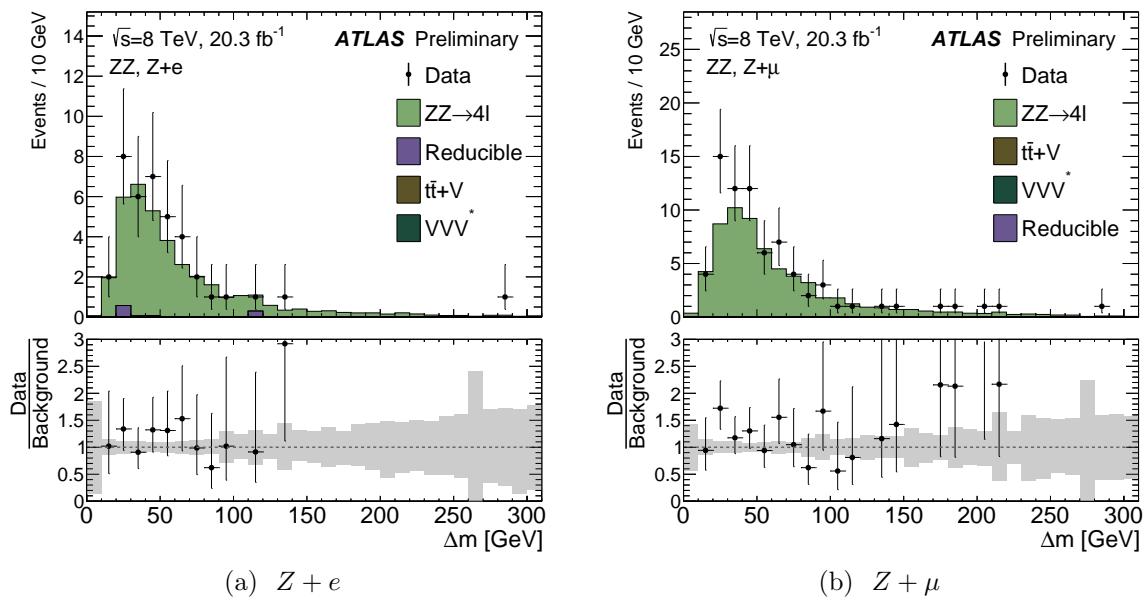


Figure 8.15: Δm distributions for the ZZ validation regions.

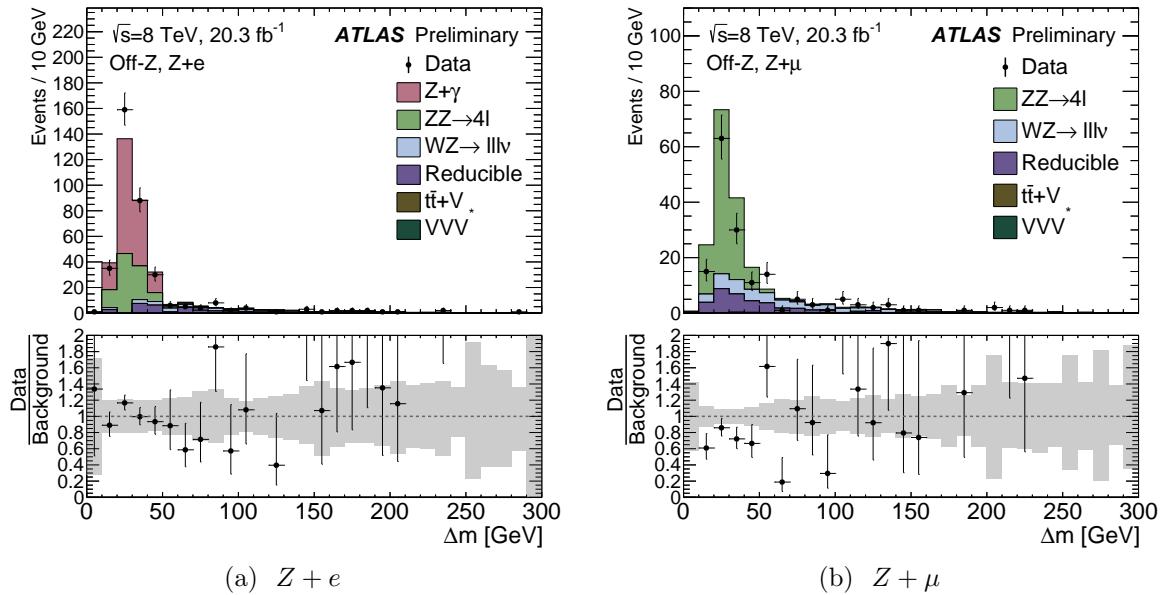


Figure 8.16: Δm distributions for the off- Z validation regions.

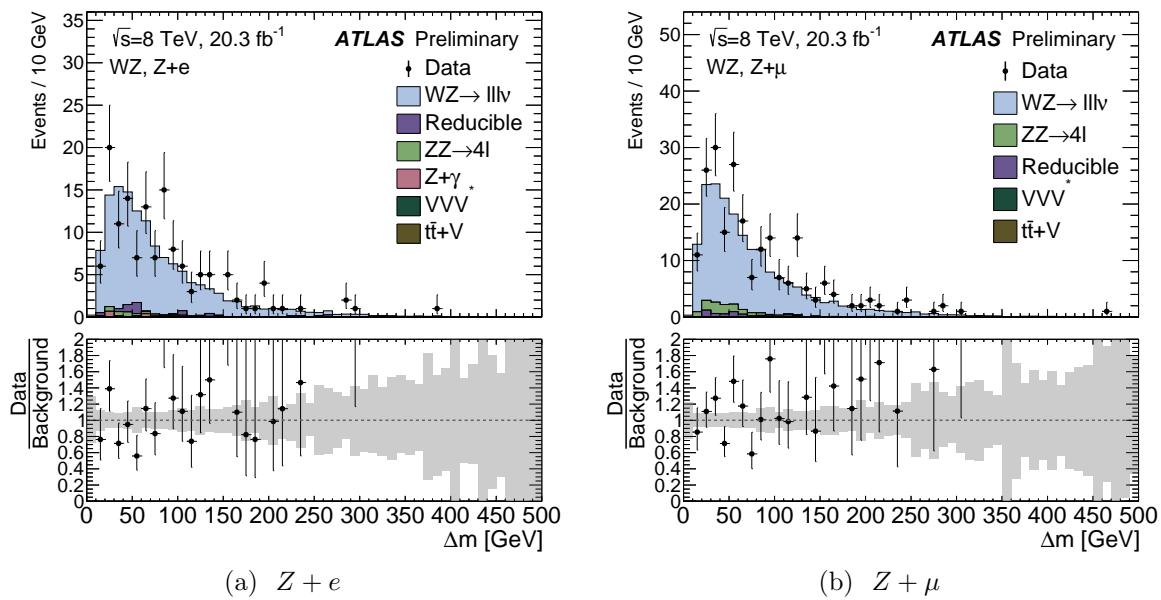


Figure 8.17: Δm distributions for the WZ validation regions.

8.5 Signal and Background Fit Model

The numbers of signal and background candidate events in data are determined from an unbinned maximum-likelihood fit of a combination of signal and background models to the Δm distributions in each signal region. The signal and background processes are modelled with analytical probability density functions (p.d.f.s). The parameters of the p.d.f.s are determined from fits to the Δm distributions predicted by simulation or the fake factor procedure. Then, in each flavor channel, the data is fitted with the combined signal and background model, performed simultaneously on the three categories. In each signal region, the normalization of the dominant background (ZZ or WZ) is a free parameter in the fit. The normalizations of all other backgrounds are constrained to fluctuate according to Gaussian probability distributions with mean and width values equal to the estimates and the total uncertainties before fitting. The uncertainties on the p.d.f. parameters are incorporated as Gaussian-distributed nuisance parameters.

Signal Modeling

The Δm distributions for the type III seesaw and vector-like leptons scenarios contain two distinct pieces: a peak associated with the correct identification of the three leptons due to the heavy lepton decay, and a broader component due to cases where the trilepton candidate does not originate from a heavy lepton decay. Accordingly, the distributions are modeled using the sum of two functions: a Voigtian for the peak and a Landau function for the combinatorial component. The signal parametrization is given by the following expression:

$$\mathcal{S}(m_\Sigma) = f_V F_V(\Delta m; \Gamma_V, m_V, \sigma_V) + (1 - f_V) F_L(\Delta m; \sigma_L, m_L), \quad (8.15)$$

$$(8.16)$$

where f_V denotes the fraction of events in the Voigtian; Γ_V , m_V and σ_V are the width, mean and gaussian smearing term of the Voigtian; and σ_L and m_L are the width and mean of the Landau distribution.

The Δm distributions at each simulated mass point for both the type III seesaw and the vector-like leptons models are fitted with the Voigtian plus Landau function, separately for each flavor channel and category. The fits for the 120 GeV mass hypothesis of the type III seesaw model are shown both for Δm and m_{3l} , along with the corresponding pulls and residuals of the simulated data with respect to the fit, in figures ??, ??, and ???. The same plots, but enlarged around the signal peak, are shown in figure ??.

Tables ??, ??, ??, and ?? show the fitted parameters for the inclusive signal regions for each of the two flavor channels and signal models. As mentioned in section ??, the type III seesaw samples were simulated without intrinsic Z width, and hence the width parameters vector-like leptons are used instead. Plots showing the fits to the Δm distributions are shown in appendix ??.

Signal hypotheses at intermediate mass points between the simulated values are obtained by linearly interpolating the fit parameters determined at the nearest simulated points above and below. To validate the linear interpolation between the mass points, a closure test was performed comparing the fit parameters determined at a simulated mass point with the values obtained from a linear interpolation between the adjacent simulated points. The resulting fit for the 160 GeV mass point is shown in Fig. ???. Further test and comparisons between the signal shapes of the two models are shown in appendix ??.

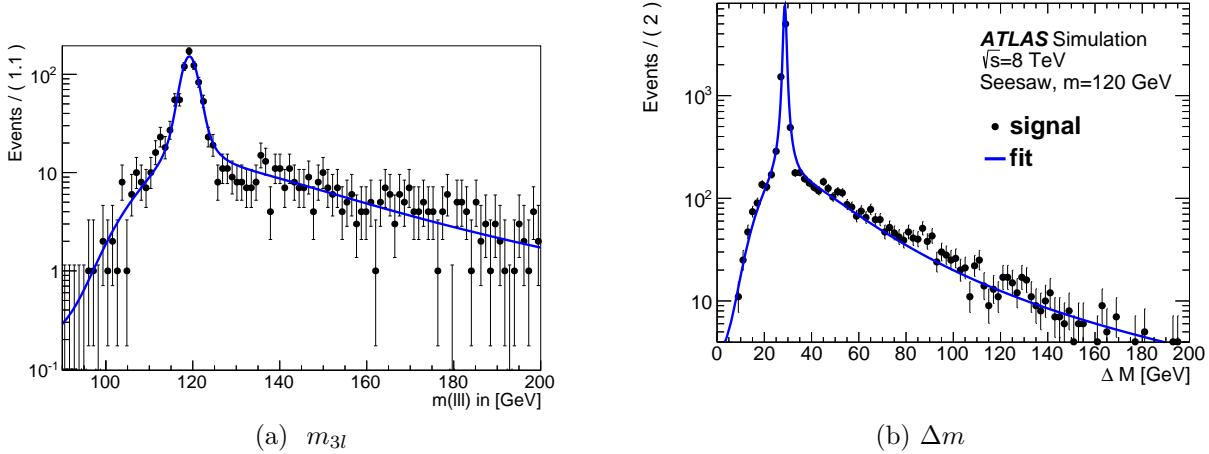


Figure 8.18: m_{3l} and Δm distributions for the 120 GeV signal mass point fitted the sum of a Voigtian and a Landau PDF. The fit is an unbinned maximum likelihood fit; the data is binned only for presentation.

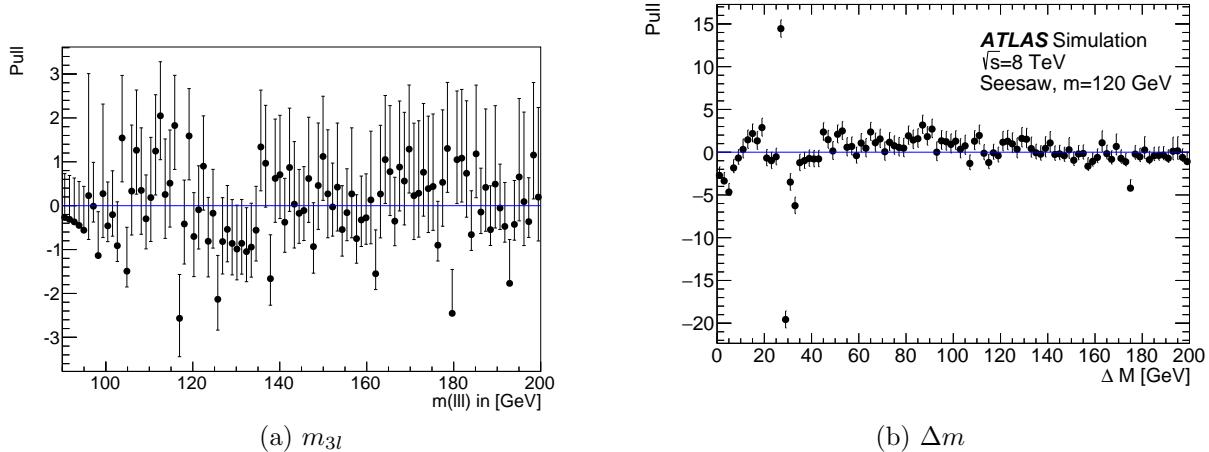


Figure 8.19: Pulls of the simulated data with respect to the fit for m_{3l} and Δm at the 120 GeV signal mass point.

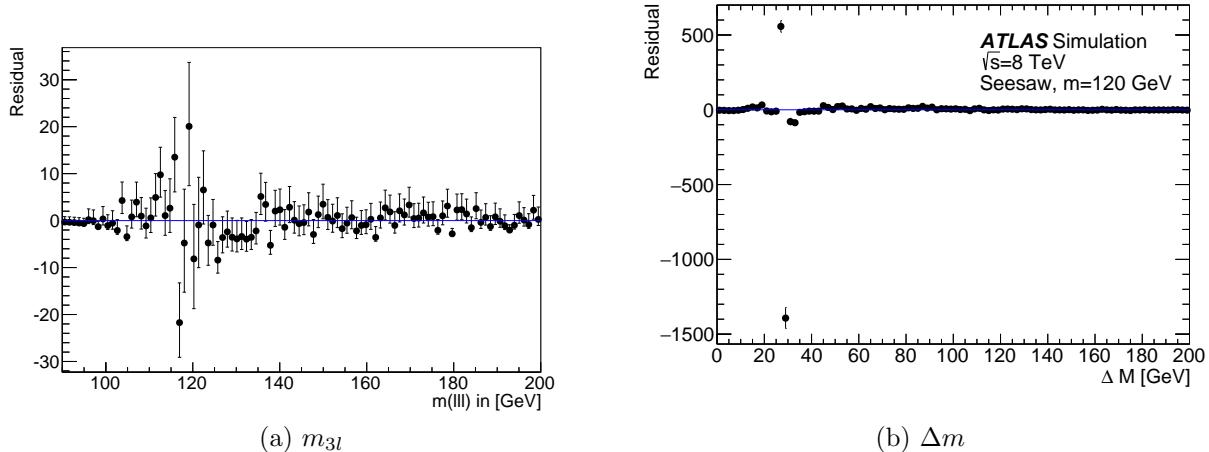


Figure 8.20: Residuals of the simulated data with respect to the fit for m_{3l} and Δm at the

Background Fits

The dominant diboson backgrounds are modeled using a Bukin function, a 5-parameter function designed to model asymmetric peaks:

$$\mathcal{P}(x; x_p, \sigma_p, \xi, \rho) = A_p \exp \left[\frac{\xi \sqrt{\xi^2 + 1} (x - x_1) \sqrt{2 \log 2}}{\sigma_p (\sqrt{\xi^2 + 1} - \xi)^2 \log(\sqrt{\xi^2 + 1} + \xi)} + \rho \left(\frac{x - x_i}{p_p - x_i} \right)^2 - \log 2 \right], \quad (8.17)$$

where $\rho = \rho_1$ and $x = x_i$ for $x < x_1$, and $\rho = \rho_2$ and $x_i = x_2$ for $x \leq x_2$. The function describes the diboson Δm distribution well, in particular successfully modeling the turn-on region at low values of Δm . However, the fit parameters are strongly correlated, with some pairs exceeding 99% correlation. To reduce the number of free parameters, two parameters, σ_p and ξ_p , are constrained to be linear functions of x_p using pseudoexperiments. 100 toy datasets are generated from the fitted Bukin function, and the Bukin function fit is repeated on each dataset. The scatter plots of σ_p versus x_p and ξ_p versus x_p are shown in figure ??, along with the linear least squares fits used to constrain σ_p and ξ_p .

The results of the 3-parameter fits are shown with the decorrelated eigenvariations of the fit parameters in figures ?? and ??.

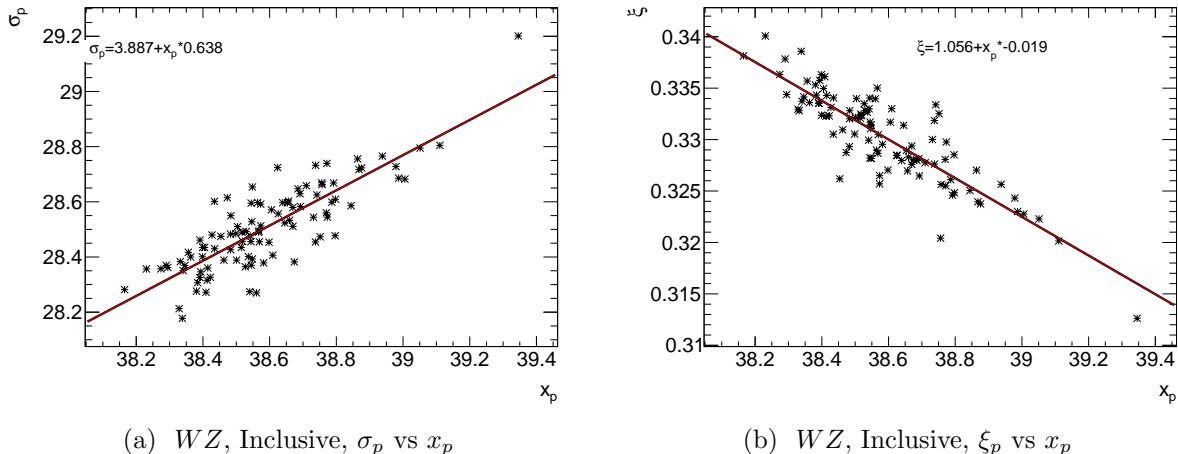


Figure 8.21: Parameter correlation for 100 toy experiments, including the linear fit for the inclusive WZ region.

The shapes of the diboson Δm distributions in the categorized signal regions are compared to the inclusive signal regions using Kolmogorov-Smirnov tests, shown in figures ?? and ?? and tables ?? and ???. The shapes of the Δm distributions in the 4ℓ and 3ℓ – only signal regions are consistent with the inclusive regions; therefore, to take advantage of the larger statistics in the inclusive signal region, **the shape from the inclusive signal re-**

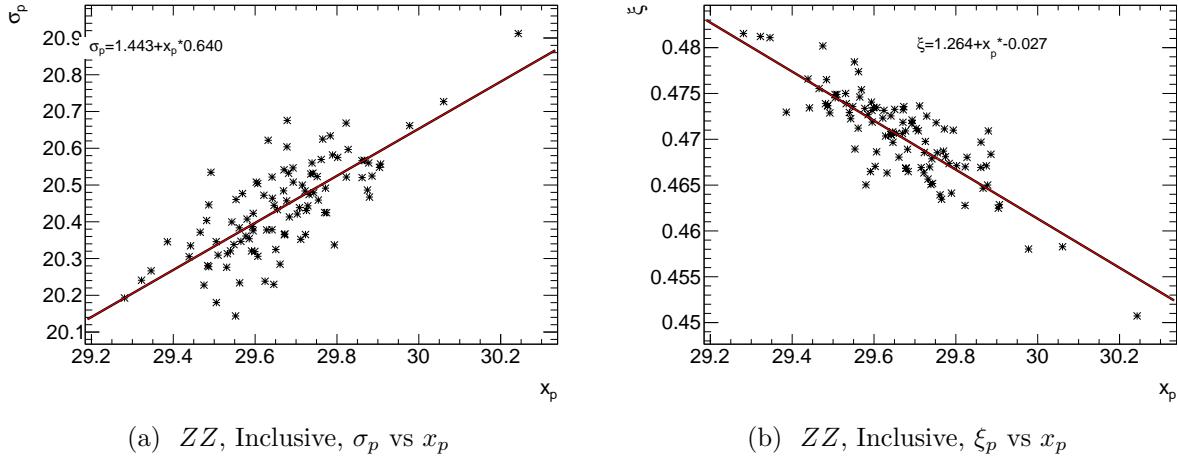


Figure 8.22: Parameter correlation for 100 toy experiments, including the linear fit for the inclusive ZZ region.

gion is used to model the 4ℓ and 3ℓ – only signal regions. The final fits are shown in figures ?? and ??.

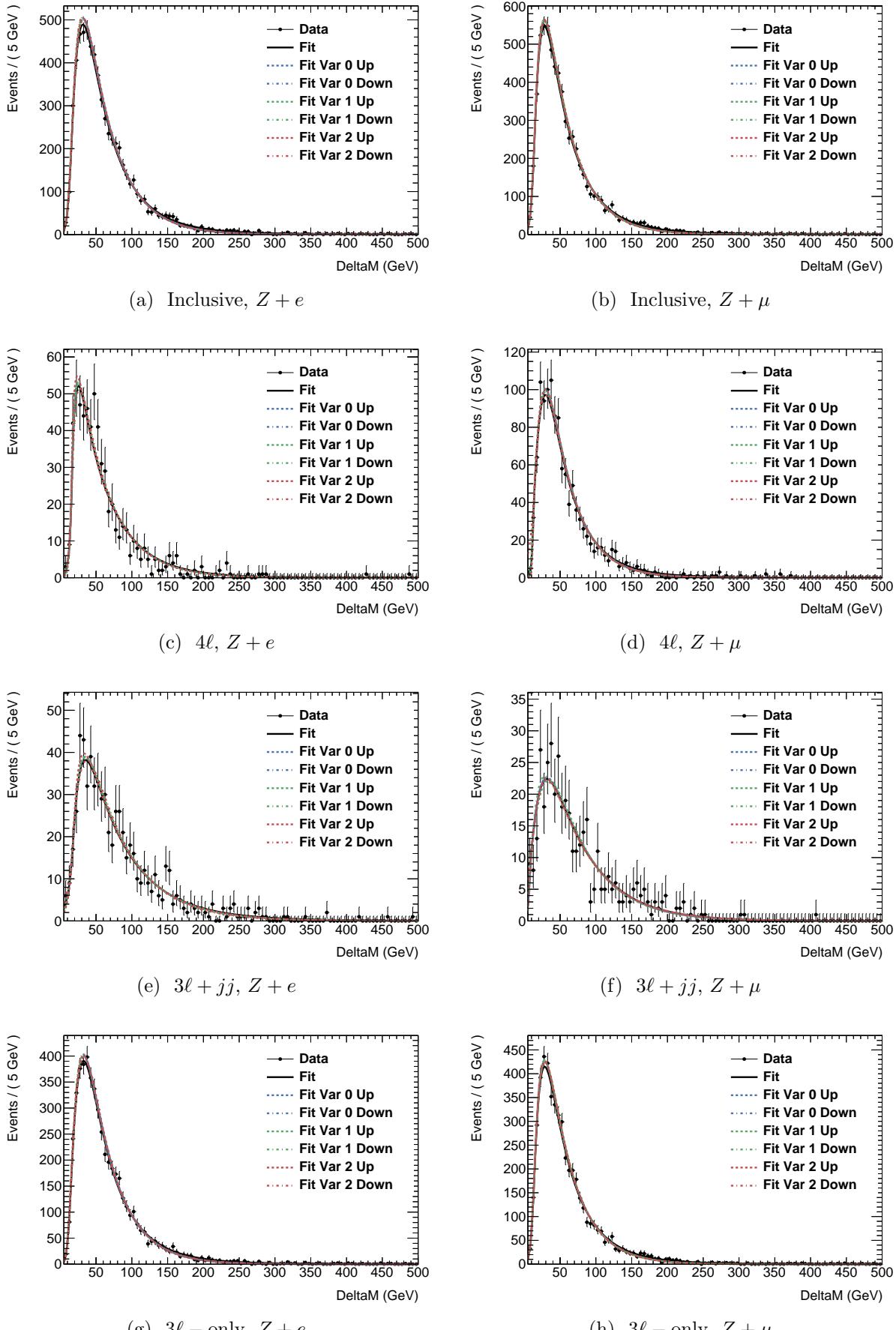
Signal Region	KS probability	D
$4\ell, Z + e$	0.4109	0.060
$3\ell + jj, Z + e$	0.001	0.122
3ℓ – only, $Z + e$	0.782	0.019
$4\ell, Z + \mu$	0.520	0.043
$3\ell + jj, Z + \mu$	0.130	0.098
3ℓ – only, $Z + \mu$	0.946	0.016

Table 8.12: Results of KS tests comparing the ZZ Δm shapes between the categorized signal regions and the inclusive signal regions.

Signal Region	KS probability	D
$3\ell + jj, Z + e$	0.0	0.165
3ℓ – only, $Z + e$	0.012	0.016
$3\ell + jj, Z + \mu$	0.0	0.152
3ℓ – only, $Z + \mu$	0.0349	0.012

Table 8.13: Results of KS tests comparing the WZ Δm shapes between the categorized signal regions and the inclusive signal regions.

The most important remaining backgrounds are due to reducible processes and $Z(l\bar{l}) + \gamma$. The $Z(l\bar{l}) + \gamma$ background is only significant in the $Z + e$ signal regions. Due to the limited



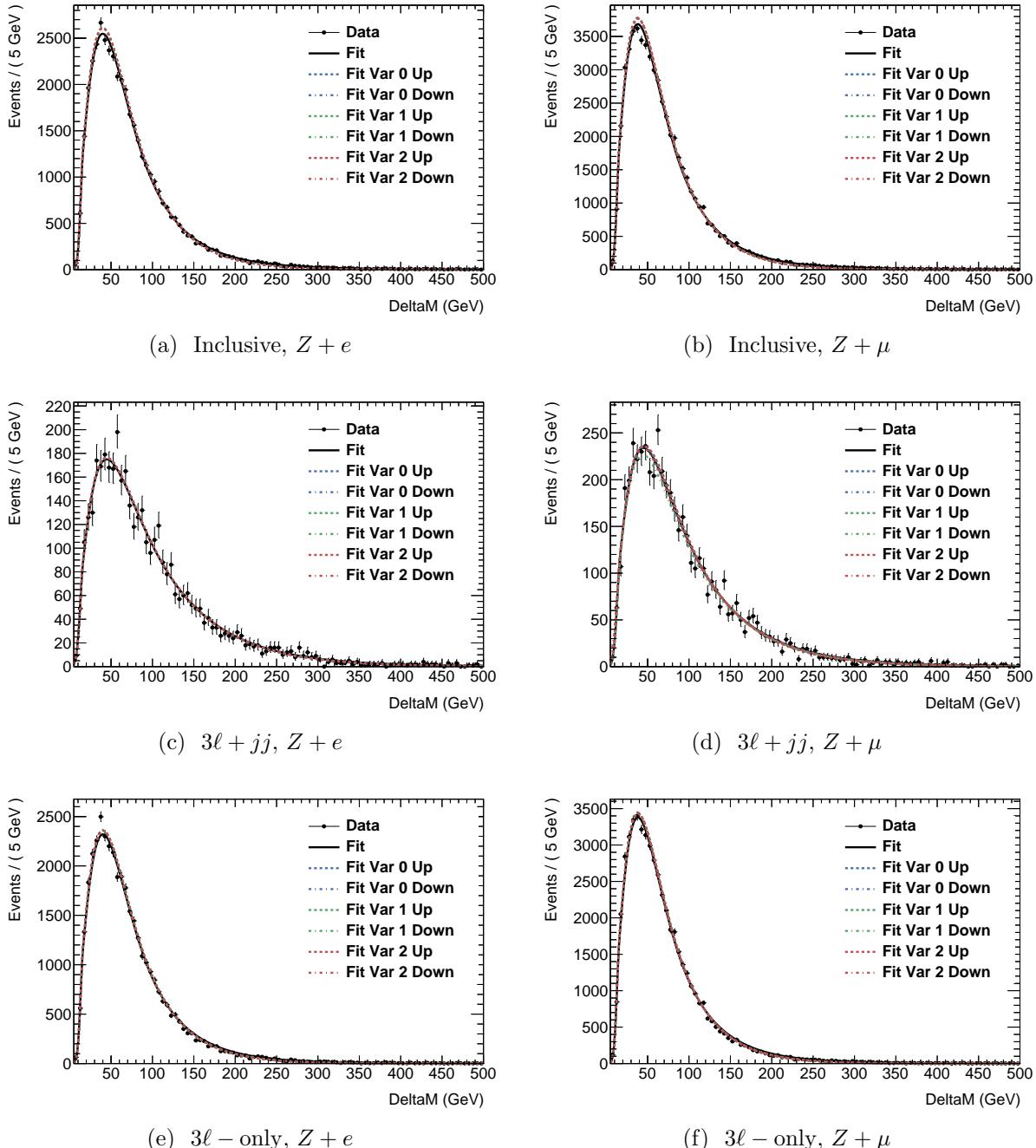


Figure 8.24: Restricted 3-parameter Bukin fits to the WZ Δm distributions in each signal region. The central value is shown in black, and the $1-\sigma$ eigenvariations of the fit parameters are shown in the colored dashed lines.

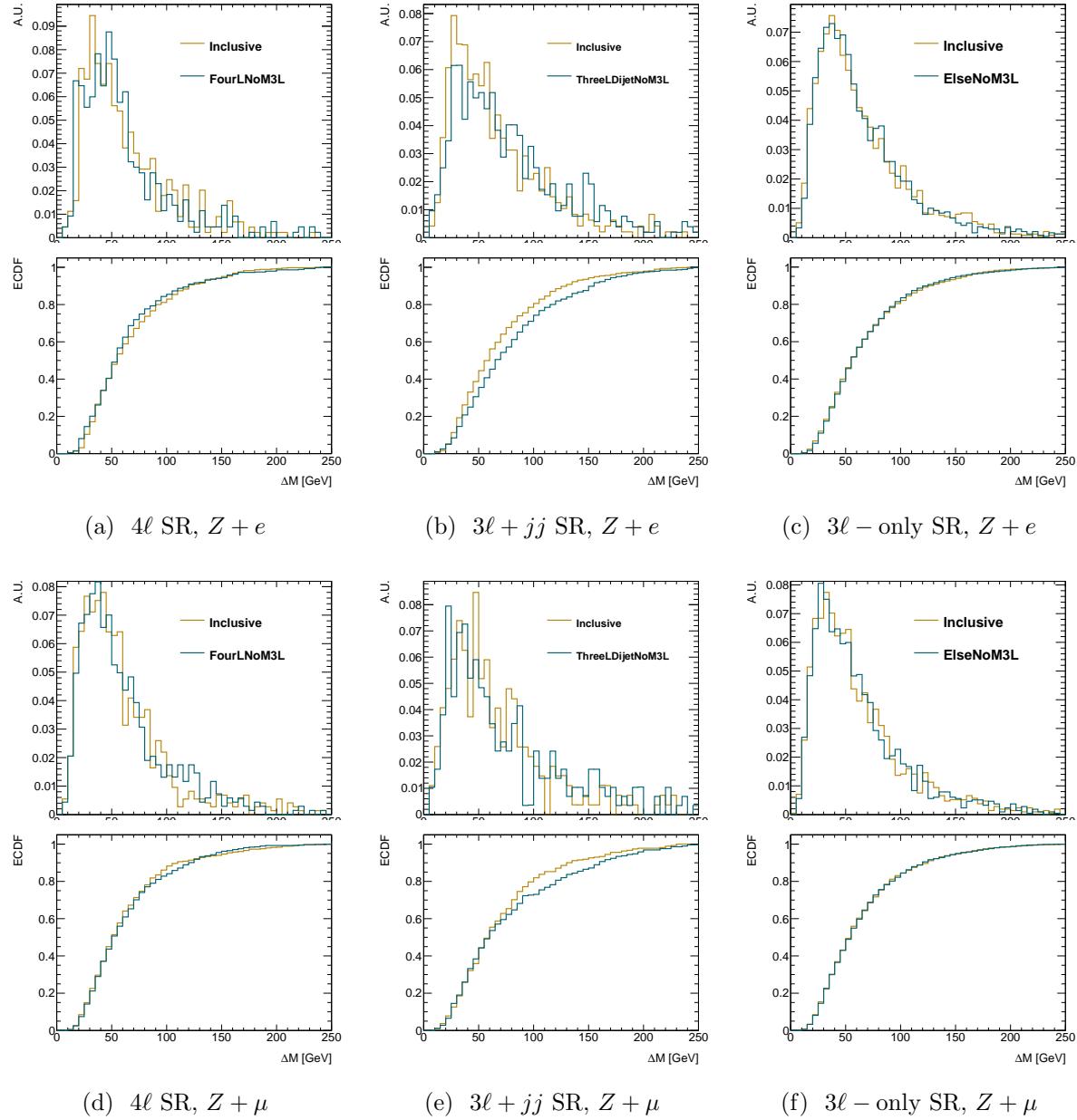


Figure 8.25: Comparison of $ZZ \Delta m$ shapes between the inclusive signal region and the categorized signal regions, with empirical distribution functions.

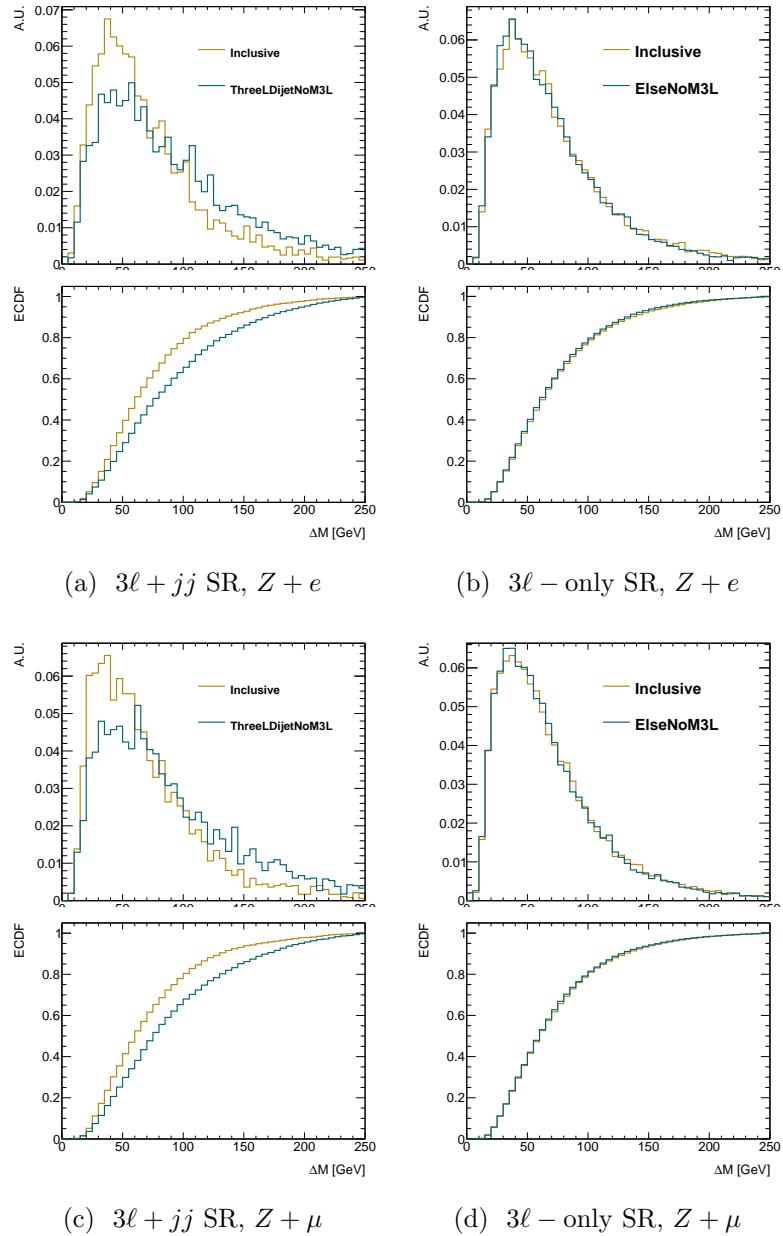


Figure 8.26: Comparison of WZ Δm shapes between the inclusive signal region and the categorized signal regions, with empirical distribution functions. The 4ℓ signal regions are not shown due to the lack of WZ events with four leptons.

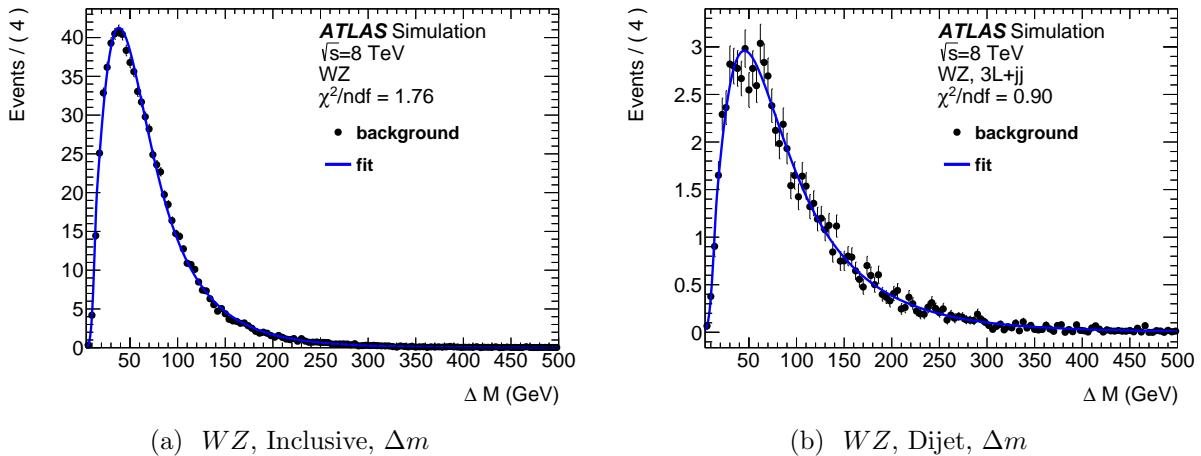


Figure 8.27: Δm distributions for WZ and Bukin function fits for the inclusive and the $3\ell + jj$ signal regions. The separate fits for the $Z + e$ and $Z + \mu$ flavor channels can be found in Appendix ??.

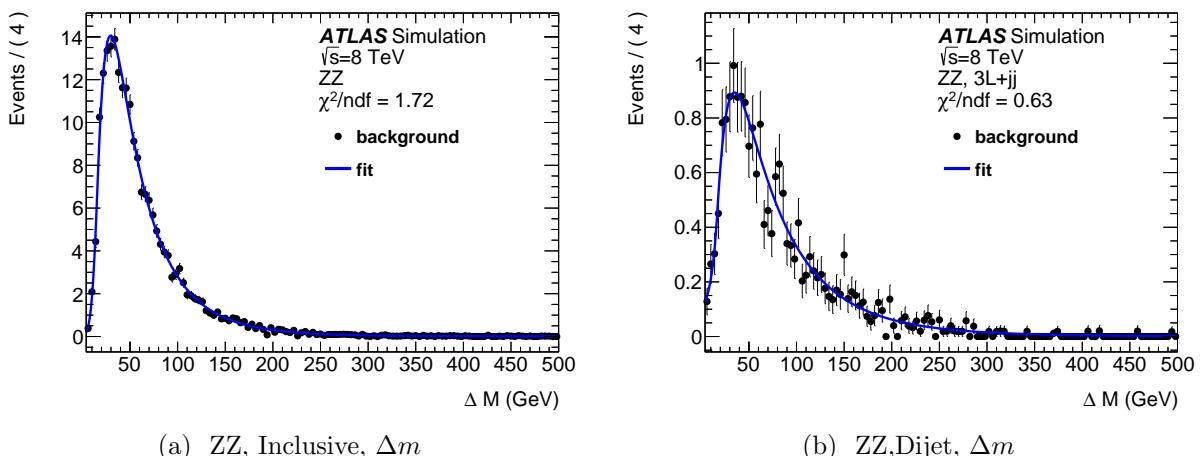


Figure 8.28: Final Δm distributions for ZZ and Bukin function fits for the inclusive and the $3\ell + jj$ signal regions. The separate flavour fits can be found in Appendix ??.

statistics from the fake factor estimate, the individual categories are combined into a single inclusive distribution for electron and muons. The reducible backgrounds are fitted with Landau distributions, shown in figure ???. The individual regions are then normalized to the expectations from the individual categories ¹.

The $Z(l\bar{l}) + \gamma$ background is only significant in the $3\ell -$ only, $Z + e$ signal region. This background is modeled with the sum of a Landau and a Gaussian, shown in figure ???. The remaining small contributions from $t\bar{t} + V$ and triboson processes are modeled together with a Landau function, shown in figure ??.

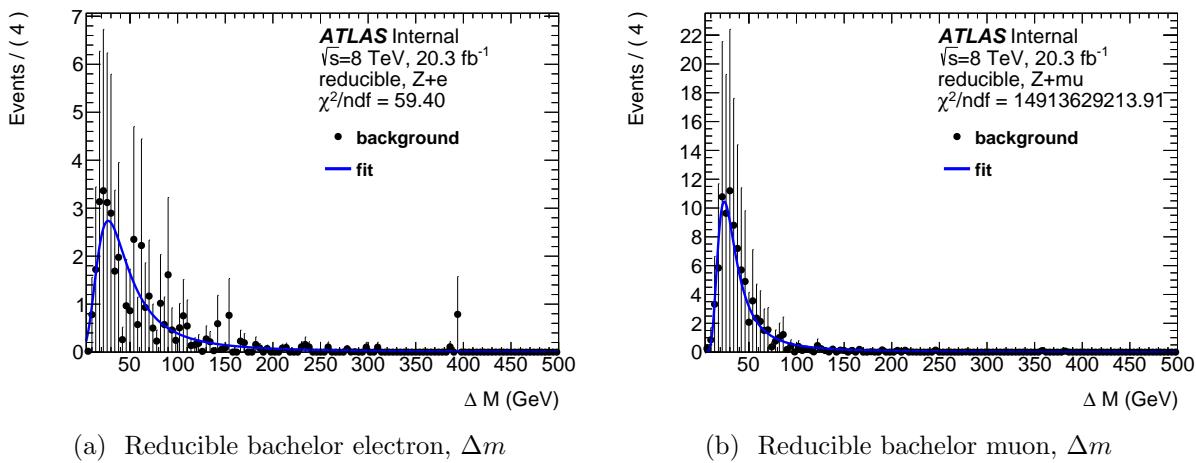


Figure 8.29: The shape of the reducible background is taken from the estimates from the fake factor method. The distribution is parametrized with a Landau for both bachelor lepton flavors. Due to the limited statistics in the categorized signal regions, the shape is determined in the inclusive signal region, and only the normalization varies in the categorized signal regions.

¹In categories where the fake factor method predicts overall normalization that is negative but consistent with zero within statistical uncertainties, which can occur due to the prompt subtraction, the normalization is set to zero.

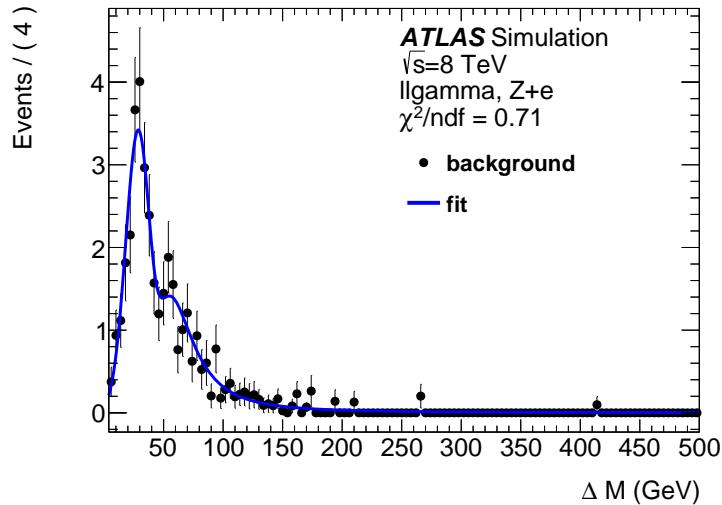


Figure 8.30: The Δm distribution for the $Z + \gamma$ background, fitted with the sum of a Landau and a Gaussian. This background is significant for final states with a bachelor electron.

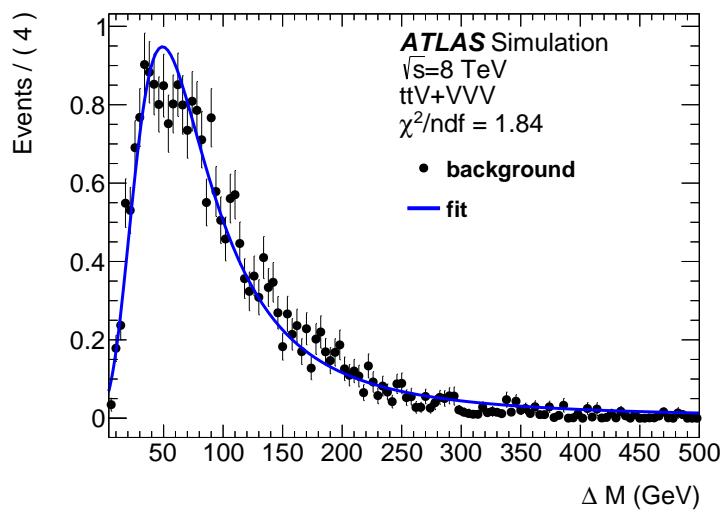


Figure 8.31: The Δm distribution for the combined $t\bar{t} + V, VVV^{(*)}$ background, fitted with the sum of a Landau and a Gaussian. This background is significant for final states with a bachelor electron.

8.6 Results

The total number of observed events each signal region is shown in table ???. The predicted backgrounds before and after performing the unbinned maximum-likelihood fit in the background-only hypothesis are also shown. The Δm distributions for the pre-fit background estimates and the data are shown in figure ??, with examples signals from the vector-like leptons model with $m_{L^\pm} = 140$ GeV and the type III seesaw model with $m_{L^\pm} = 300$ GeV also shown. The data agree with the background expectation in all cases, and no clear peak indicating resonant trilepton production is seen in any of the signal regions.

Table 8.14: Observed and expected number of events in the six signal regions, before and after the combined unbinned maximum-likelihood fit. The pre-fit uncertainties represent the total systematic uncertainties on the background estimates. The post-fit uncertainties are determined by the maximum-likelihood fit.

Process	$Z + e$			$Z + \mu$		
	4l SR	3l + jj SR	3l-only SR	4l SR	3l + jj SR	3l-only SR
Before combined background-only fit						
ZZ	10.9 ± 0.6	11.7 ± 0.8	91 ± 5	21.4 ± 1.1	7.5 ± 0.6	90 ± 5
WZ	0.08 ± 0.01	35.3 ± 3.1	337 ± 28	—	46 ± 4	480 ± 40
$Z + \gamma$	—	2.3 ± 0.8	35 ± 11	—	—	—
Reducible	—	1.6 ± 0.5	38 ± 14	1.5 ± 0.3	8.8 ± 3.0	79 ± 22
$t\bar{t} + V, VVV^{(*)}$	1.2 ± 0.2	7.8 ± 1.7	2.3 ± 0.4	1.5 ± 0.2	9.5 ± 2.1	3.3 ± 0.5
Total Background	12.2 ± 0.7	59 ± 4	504 ± 34	24.4 ± 1.2	72 ± 6	650 ± 50
After combined background-only fit						
ZZ	15 ± 4	13.4 ± 2.3	107 ± 9	22 ± 5	10.1 ± 1.6	88 ± 8
WZ	0.08 ± 0.03	39 ± 6	393 ± 28	0.02 ± 0.02	56 ± 9	460 ± 40
$Z + \gamma$	—	2.2 ± 0.8	34 ± 11	—	—	—
Reducible	—	1.8 ± 1.2	37 ± 13	1.8 ± 0.9	10.2 ± 2.8	92 ± 24
$t\bar{t} + V, VVV^{(*)}$	1.1 ± 0.3	7.5 ± 1.7	2.5 ± 0.6	1.5 ± 0.4	9.1 ± 2.1	3.3 ± 0.8
Total Background	16 ± 4	64 ± 7	574 ± 34	25 ± 5	85 ± 10	640 ± 40
Data	16	64	573	25	86	651

Good agreement is seen between the pre-fit and post-fit normalizations for the 4ℓ and $3\ell + jj$ categories in the $Z + \mu$ flavour channel. The largest change in normalization due to the fit is in the 4ℓ category for the $Z + e$ flavour channel, where the fitted ZZ normalization exceeds the prediction by 35%. The WZ and ZZ normalizations increase by roughly 15%

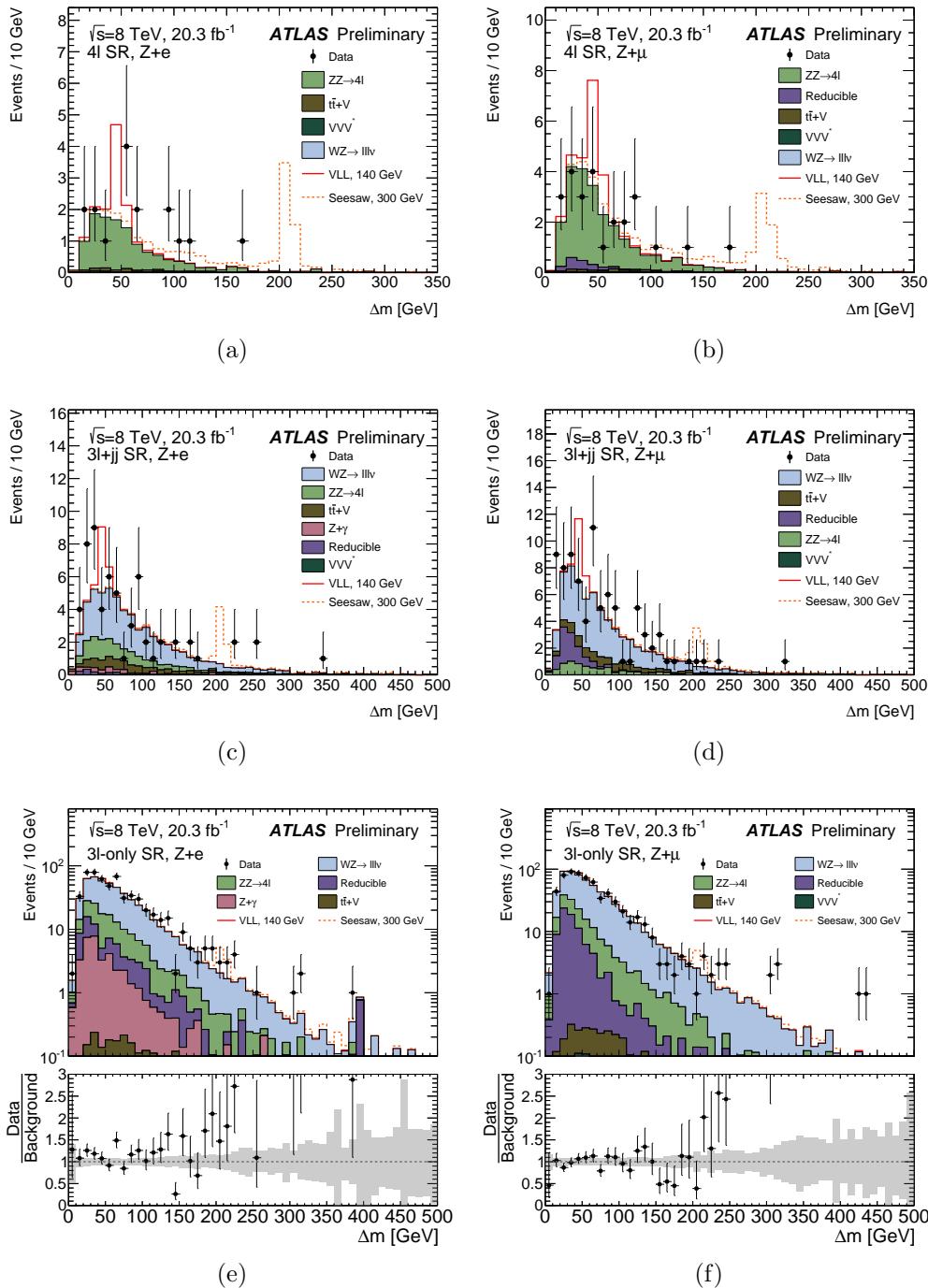


Figure 8.32: The $\Delta m = m_{3\ell} - m_{\ell^+\ell^-}$ distributions for the 4ℓ (top), $3\ell + jj$ (middle), and $3\ell -$ only (bottom) categories, divided into the $Z + e$ (left) and $Z + \mu$ (right) flavour channels. The observed data are shown as black points, while the pre-fit background expectations are shown in the coloured histograms. Also shown are examples for signal contributions for a 140 GeV L^\pm in the vector-like leptons model and a 300 GeV L^\pm in the type III seesaw model. The error bars on the data points represent statistical uncertainties, and the shaded bands represent the systematic uncertainties on the background predictions.

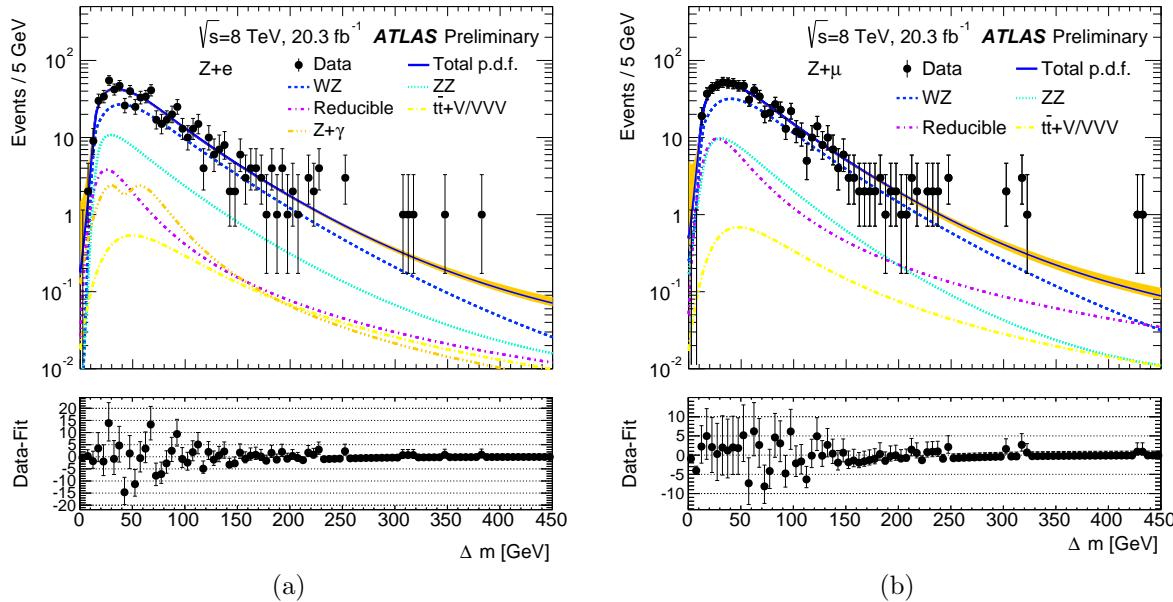


Figure 8.33: Projections onto the Δm variable of the background-only unbinned maximum-likelihood fits, shown superimposed on the data with the three categories in each flavour channel added together. The $Z + e$ flavour channel is shown in (a), and the $Z + \mu$ channel is shown in (b). The contributions of the separate background components to the total background-only fit are also shown. The error bars on the data points represent statistical uncertainties. Good agreement is observed between the background model and the data.

in the $3\ell + jj$ and $3\ell -$ only categories in the $Z + e$ flavour channel, and 30% in the $3\ell + jj$ category in the $Z + \mu$ flavour channel. The projections of the fit results in the background-only hypothesis are shown in figure ?? for the combination of the three categories in each flavour channel.

8.7 Interpretation

The data are well described by the combined fit to the three categories in each flavour channel. The consistency of the data with the background-only hypothesis is evaluated by scanning the local p_0 -value for the Δm distribution in 3 GeV intervals for signal mass hypotheses in the range $100 - 400$ GeV for the vector-like leptons model, and $100 - 500$ GeV for the seesaw model, using the unbinned maximum-likelihood fit described in section ?? with the signal strength set to zero. The agreement is expressed in terms of the p_0 -value, the probability that, assuming the background-only hypothesis is true, an experiment would yield at least as many events as observed in the current measurement. The p_0 values are calculated using the frequentist hypothesis test based on the profile likelihood ratio test

statistic and approximated with asymptotic formulae [**asimov**]. The minimum p_0 -value is $p_0 = 0.02$ at a mass of 183 GeV for the $Z + e$ flavour channel, and $p_0 = 0.05$ at a mass of 109 GeV for the $Z + \mu$ flavour channel.

Limits on Models

Since no significant excess above the background expectation is observed, the fit model is used to derive 95% confidence level (CL) exclusion limits on the heavy lepton pair-production cross section, σ , using the CL_s method [**CLs**]. The limits are shown for the vector-like leptons model in figure ??, and for the type III seesaw model in figure ??, evaluated in the same 3 GeV intervals as the p_0 -values. The vector-like leptons model is excluded for electron-only mixing in the heavy lepton mass ranges 129–144 GeV and 163–176 GeV, with an expected exclusion in the range 109–152 GeV. The corresponding observed (expected) exclusion for the muon-only mixing scenario is 114–153 GeV and 160–168 GeV (105–167 GeV). The significantly higher production cross sections for the type III seesaw model lead to an observed (expected) exclusion in the electron-only mixing scenario in the heavy lepton mass range 100–430 GeV (100–436 GeV). For the muon-only mixing scenario, the observed exclusion is in the ranges 100–401 GeV and 419–468 GeV, while the expected exclusion is 100–419 GeV.

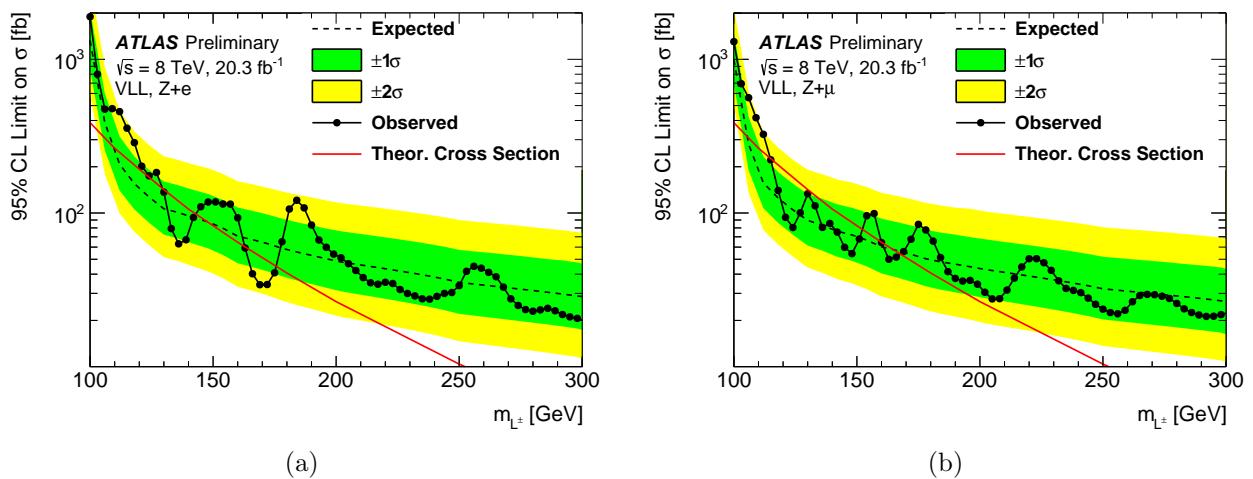


Figure 8.34: 95% CL upper limits on the vector-like lepton cross section. The left (right) plot shows the limits assuming 100% branching fraction to e/ν_e (μ/ν_μ). The solid line shows the observed limit. The dashed line shows the median expected limit for a background-only hypothesis, with green and yellow bands indicating the expected fluctuations at the $\pm 1\sigma$ and $\pm 2\sigma$ levels. The limit is evaluated in 3 GeV intervals.

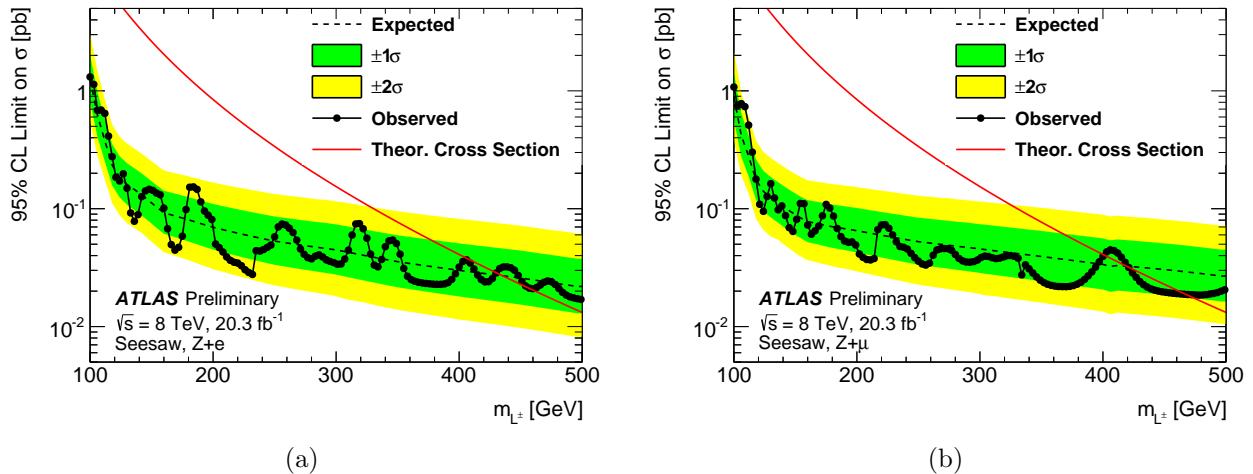


Figure 8.35: 95% CL upper limits on the type III seesaw production cross section. The left (right) plot shows the limits assuming 100% branching fraction to e/ν_e (μ/ν_μ). The solid line shows the observed limit. The dashed line shows the median expected limit for a background-only hypothesis dataset, with green and yellow bands indicating the expected fluctuations at the $\pm 1\sigma$ and $\pm 2\sigma$ levels. The limit is evaluated in 3 GeV intervals.

Model-Independent Limits

The constraints shown in figures ?? and ?? are relevant to the specific VLL and type III seesaw models considered, and are not necessarily applicable to other scenarios predicting trilepton resonances with an intermediate Z boson. A more model-independent observable is the *visible cross section*, σ_{vis} , defined as the number of observed events with $Z + \ell$ -induced trilepton resonances for a given resonance mass divided by the integrated luminosity of the data sample, 20.3 fb^{-1} . The 95% CL upper limits on σ_{vis} , denoted σ_{95}^{vis} , are derived from a fit to each flavour channel with $f_V = 1$, i.e. using only the peak component of the signal. The results for the two flavour channels, derived from the inclusive event selection without dividing the events into the three categories, are shown in figure ??.

The limits on σ_{vis} can be used to test specific models after taking into account the model's acceptance with respect to a fiducial volume, \mathcal{A} , and reconstruction and selection efficiency of events within the fiducial volume, ϵ_{fid} . The 95% CL upper limit on the cross section for the model is given by:

$$\sigma_{95} = \frac{\sigma_{95}^{\text{vis}}}{\mathcal{A} \times \epsilon_{\text{fid}}} \quad (8.18)$$

The acceptance \mathcal{A} is defined as the probability for generated signal events to lie within a fiducial volume defined by the kinematics of the generated leptons. The leptons are considered at *particle level*, i.e. after parton shower and hadronization and with lifetimes longer than 10^{-11} s, and are *dressed*, including the contributions from radiated photons within a

cone of $\Delta R = 0.1$. The fiducial volume requires that events contain an L^\pm decaying to a prompt electron or muon and a Z boson that then decays to electrons or muons. The three leptons from the L^\pm decay are required to have $p_T > 15$ GeV and lie within $|\eta| < 2.5$, with at least one lepton satisfying $p_T > 26$ GeV. Two of the leptons must form a same-flavour opposite-sign pair with a mass within 10 GeV of m_Z , and the Z boson and the off- Z lepton must be separated by $\Delta R < 3$. The events are divided into flavour channels according to the flavour of the off- Z lepton. For the VLL and type III seesaw models used in this analysis, the acceptance of events containing an $L^\pm \rightarrow Z(\ell\ell)\ell$ decay to fall within the fiducial volume is in the range 60%–65% for most of the mass range, decreasing at higher masses due to the cut on the ΔR between the Z boson and the off- Z lepton. The acceptance decreases at low masses due to the lepton p_T requirement, reaching 30%–35% at $m_{L^\pm} = 100$ GeV.

For type III seesaw and VLL events within the fiducial volume, ϵ_{fid} ranges from 20% to 49% if the other heavy lepton decays to a neutrino and a W , Z , or H boson. If the other heavy lepton decays to an electron or a muon, the efficiency is 10%–20% lower, due to the increased probability of incorrectly selecting the off- Z lepton. The event selection efficiencies for the type III seesaw model in scenarios where the second heavy lepton decays to a W boson are shown in figure ?? as a function of m_{L^\pm} ; the efficiencies for scenarios where the second heavy lepton decays to a Z or H boson and for the vector-like leptons model are consistent with these efficiencies within the statistical uncertainties.

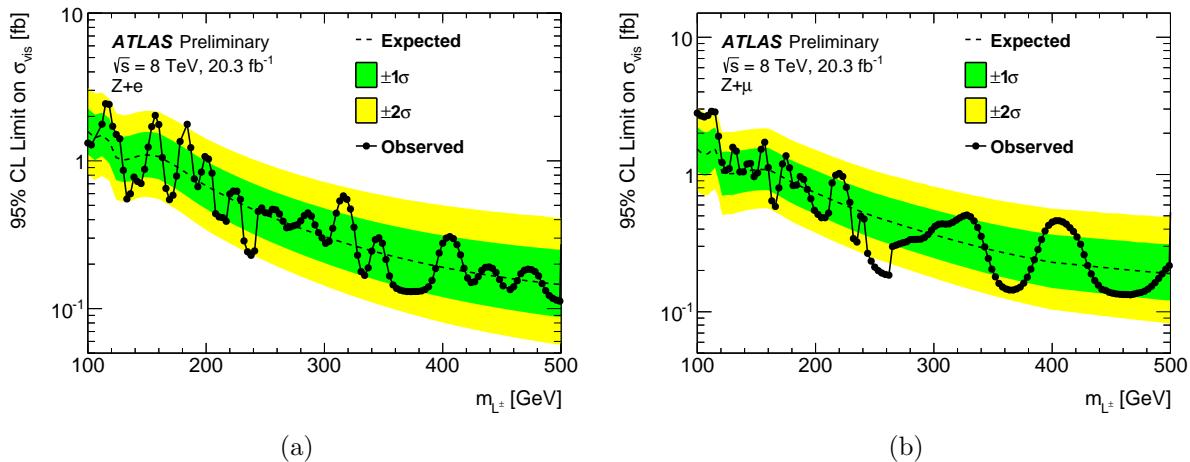


Figure 8.36: Upper limits at 95% CL on σ_{vis} for the $Z + e$ (left) and $Z + \mu$ (right) flavour channels, derived without dividing events into the three categories. The limits are evaluated in 3 GeV intervals.

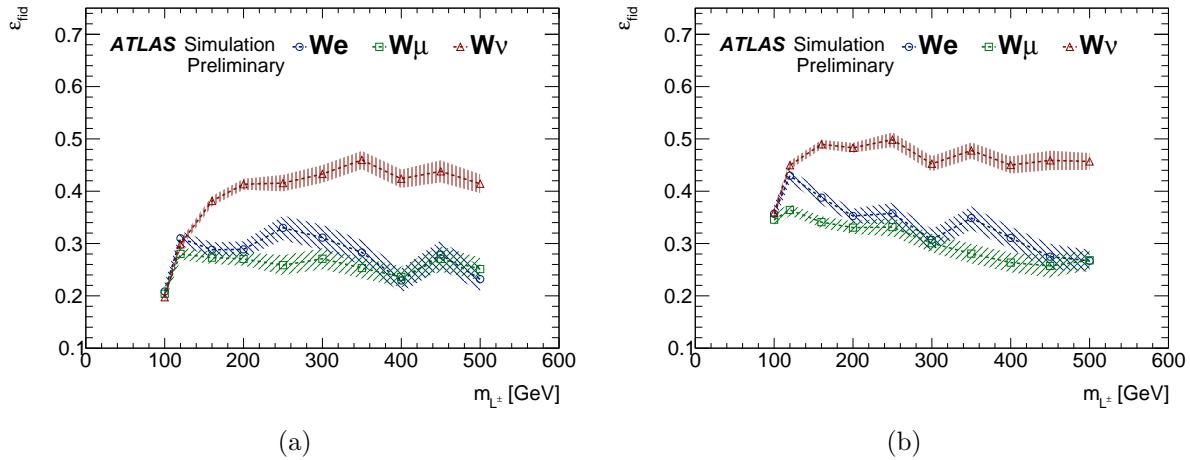


Figure 8.37: Efficiencies for reconstructing and correctly identifying the $L^\pm \rightarrow Z(\ell\ell)\ell^\pm$ decay in events within the fiducial volume for the type III seesaw model. The left (right) plot shows the efficiencies for events containing a $L^\pm \rightarrow Z(\ell\ell)e$ ($L^\pm \rightarrow Z(\ell\ell)\mu$) decay. The decay of the second heavy lepton is specified in the legend. The shaded bands show the statistical uncertainty.

Chapter 9

Conclusion