Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector

Rhys Roberts

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

School of Physics and Astronomy

2018

Contents

Word count: 41002

List of Figures

List of Tables

Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering

Rhys Roberts
School of Physics and Astronomy
2018

Abstract

This thesis presents analyses performed with proton-proton collision data collected with the ATLAS detector in 2015 and 2016 at $\sqrt{s} = 13$ TeV.

The measurement of so-called non-factorisation effects in the calibration of luminosity is presented. An estimate of the correction, and systematic uncertainty, due to non-factorisation on the calibration is calculated by measuring the proton bunch density profiles from the variation in the luminosity and distribution of reconstructed vertices during beam separation scans. The correction is applied to the calibrations of the total luminosity collected in both 2015 and 2016.

A novel multivariate algorithm designed to reject non-prompt leptons (produced from the decays of b- and c-quarks) is presented, utilising information from nearby tracks to discriminate from prompt leptons (produced from W, Z and H boson decays). This algorithm is used to reject non-prompt backgrounds in the search for the associated production of a top quark pair and a Higgs boson ($t\bar{t}H$) in multilepton final states with 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV data. Multilepton states refer to the Higgs boson decaying into pairs of W bosons, Z bosons or τ leptons.

The combination of the multilepton analysis with the other search analyses of $t\bar{t}H$ production in which the Higgs decays to pairs of photons, b-quarks and $ZZ \to 4\ell$ is also shown. The measured value of the signal strength of $t\bar{t}H$ production in data is $\mu(t\bar{t}H) = 1.2 \pm 0.3$, corresponding to an observed (expected) discovery significance of 4.2σ (3.8σ) and constituting evidence for the $t\bar{t}H$ production mode.

Declaration and copyright statement

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Rhys Roberts

Copyright statement

[1] See http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420
[2] See http://www.library.manchester.ac.uk/about/regulations/

Acknowledgements

Massive thanks go to both Prof. Terry Wyatt and Dr. Rustem Ospanov for their excellent supervision during my degree. To Terry, for offering me a chance to carry on research in particle physics after my fourth year master's project and for always having deep insight into any (of the many) problems I had. And to Rustem, for stepping up as a supervisor when all the other ATLAS analyses were taken and for always being available for a discussion or giving guidance: I hope he would agree that our collaboration over the last few years was very fruitful.

I would like to thank all the members of the Manchester particle physics group for being so friendly, too large to name individually. The last few years would not have been so enjoyable without you.

I would like to thank my parents, Mary and Dave, for all their support over the years and to my brother, Aron, for reminding me life is not just for living indoors.

And lastly I would like to say a huge thank you to Naomi. Over the last four years she has always been there to support me through the inevitable ups and downs of a postgraduate degree. I will always be grateful for her love during this time.

Preface

The author of this thesis is a member of the ATLAS collaboration. The particular contributions of the author in the analyses detailed are as follows.

In Chapter 6, the author was responsible for the determination of the non-factorisation in beam separation scans with the luminous region method in 2015 and 2016, under the supervision of Prof. Terry Wyatt. This work was done as part of the author's ATLAS qualification task, with the work continuing past qualification. The analysis built on work previously undertaken by Dr. Sam Webb and Mr. Yaadav Bhauruth. The entire non-factorisation correction and systematic uncertainty on the calibration quantity, σ_{vis}, in 2016 was determined by the author, and the non-factorisation correction to σ_{vis} in 2015 was determined in collaboration with the correction determined from the coupled model method by Dr. Miguel Ignacio Arratia Munoz. The framework within which the luminous region correction was determined was entirely developed by the author.

The author was responsible for the development of the non-prompt BDT, presented in Chapter 7. The work was performed under the supervision of, and in collaboration with, Dr. Rustem Ospanov. This involved the full research and development of the novel algorithm; the production and maintenance of the code to deploy the algorithm in ATLAS derivations; the validation of the modelling of the input variables and discriminant in data; and the presentation, and subsequent approval, of the novel method to the egamma, muon, flavour tagging and isolation forum ATLAS combined performance groups.

The search for $t\bar{t}H$ production in multilepton final states, presented in Chapter 8, was performed within an analysis group. The author was responsible for the optimisation of tight lepton definitions used in the analysis. The optimisation included defining the pre-selection regions for the most sensitive 2ℓSS and 3ℓ channels. The author then provided a calibration of the optimised muon working point for prompt muons for use in the analysis. The author also undertook cross-checks on the mis-modelling of the non-prompt BDT due to vertex density and pileup effects.

Chapter **1**

Introduction

The discovery of the Higgs boson by the ATLAS and CMS collaborations in 2012 [1, 2] was a major milestone in the validation of the Standard Model (SM) of particle physics. Ever since, the properties of the Higgs boson have been under scrutiny to determine whether it really is the particle predicted by the SM. Many beyond the SM (BSM) theories predict modifications to the properties of the SM Higgs boson in some way. Significant deviations of the properties from the SM values could indicate new physics.

The main topic of this thesis is the search for the Higgs boson produced in association with two top quarks, $t\bar{t}H$. This production mechanism is rare in the SM and, until recently, has yet to have been observed. Particles in the SM are predicted to acquire mass via the Higgs mechanism and the top quark is (by far) the most massive particle in the theory. Therefore the measurement of $t\bar{t}H$ will go some way to help identify the role of the Higgs boson in this phenomenon.

Particular emphasis is put on $t\bar{t}H$ production in the cases where the Higgs decays to pairs of W bosons, Z bosons or τ leptons, commonly referred to as multilepton final states. The light leptons in these decays (electrons or muons) are prompt leptons, produced from the very fast decays of the bosons and the slower decay of the τ. However, there can be backgrounds from events containing non-prompt light leptons, in which the lepton is produced from the decay of longer lived particles such as B hadrons. In some cases, the properties of non-prompt leptons are similar to the prompt leptons. To this end, a novel multivariate algorithm to reject non-prompt leptons has been developed. The algorithm exploits information from tracks nearby a lepton to establish if the lepton was produced from a source with a non-zero lifetime. This algorithm is used to improve the sensitivity in the search for $t\bar{t}H$ production in multilepton decays with 36.1 fb^{-1} of proton-proton collision detector at a centre-of-mass energy $\sqrt{s} = 13$ TeV with the ATLAS detector.

This thesis is structured as follows. Chapter 2 gives an overview of the Standard Model of particle physics. Hadron collider physics is discussed in Chapter 3. The Large Hadron Collider (LHC) accelerator and the subcomponents of the ATLAS detector are described in Chapter 4, and the reconstruction and identification of particles from the subcomponents are described in Chapter 5. In Chapter 6, the concept of the luminosity of an accelerator is introduced. How the luminosity is measured and calibrated at ATLAS is presented, focusing on the measurement of a correction, and corresponding uncertainty, applied to the calibration due to so-called non-factorisation effects. Chapter 7 describes the novel multivariate method developed to reject non-prompt leptons at ATLAS. The application of this method in the context of the search for $t\bar{t}H$ in multilepton final states is discussed in Chapter 8. In Chapter 9, the full combination of the $t\bar{t}H$ search analyses, targeting different Higgs decays, is shown. Finally, the concluding remarks are given in Chapter 10.

Chapter **2**

The Standard Model of particle physics

The Standard Model of particle physics is a quantum field theory that describes the interactions of matter with three of the four fundamental forces in nature: the electromagnetic (EM), weak and strong forces. The theory is formulated from an amalgam of theoretical arguments and experimental constraints, being one of the most thoroughly and successfully tested theories in physics [3]. The underlying principle of the model is the local gauge symmetry of the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$, with the non-abelian $SU(3)_c$ and $SU(2)_L$ groups describing quantum chromodynamics (QCD) and the electroweak sector respectively, and the abelian $U(1)_Y$ group describing the hypercharge sector in which quantum electrodynamics (QED) is embedded.

The elementary particles in the theory are categorised into two classes by the value of the spin of the particle: the matter sector consisting of half-integer spin particles called fermions and the force-carrying sector consisting of integer spin particles called bosons. The fundamental forces are described by the interaction of the corresponding gauge (spin 1) boson with the set or a subset of the fermions. The fermions and bosons are summarised in Table 2.1, with the corresponding values of spin, electric charge and mass listed in Table 2.2.

The fermions are categorised into two basic types called leptons and quarks. Each group consists of six particles related in pairs forming three generations. Each fermion has an associated anti-particle with the same mass but opposite quantum numbers that make up anti-matter. The leptons consist of the electron (e), the muon (μ) and the tau (τ) each with a corresponding neutrino (ν). The electron, muon and tau all have an electric charge and mass (increasing respectively), whereas the neutrinos are neutrally charged and have very small mass. The quarks consist of the up (u), down (d), strange (s), charm (c), bottom (b) and top (t). Each have mass, increasing with generation, and each are electrically charged. Quarks also contain colour charge corresponding to the coupling of the strong interaction.

All fermions interact via the weak interaction, the charged fermions interact via the electromagnetic interaction and only the quarks interact via the strong interaction. The gauge bosons mediating these interactions are the photon, γ, for the electromagnetic force; the gluon, g, for the strong interaction; and the W^+, W^- and Z bosons for the weak interaction. In addition to the gauge bosons there is a scalar (spin 0) boson in the theory, the Higgs boson, H, that gives mass to both bosons and fermions via a spontaneous symmetry breaking of the electroweak interaction. The discovery of this particle was a major test of the SM and a testament to the model's predictive power.

However, the SM is not a complete theory of the universe. It is does not describe the fourth (and weakest) force, gravity; it does not explain the observation of the large asymmetry of matter over antimatter; and does not explain the indirect observation of dark matter. Therefore measurements of SM processes are vital to determine any discrepancies between experiment and theory in which these problems may reside.

Fermions			
Leptons	$\begin{pmatrix}\nu_e \\ e\end{pmatrix}$	$\begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}$	$\begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}$
Quarks	$\begin{pmatrix}u \\ d\end{pmatrix}$	$\begin{pmatrix}c \\ s\end{pmatrix}$	$\begin{pmatrix}t \\ b\end{pmatrix}$
Bosons			
$g \quad \gamma \quad W^+ \quad W^- \quad Z \quad H$			

Table 2.1: A summary of the known matter elementary particles in the SM. The red, blue and black coloured brackets encompass the fermions that interact with the gauge bosons via the strong, electromagnetic and weak forces respectively.

Particle	Symbol	Spin	Electric charge	Mass
electron	e^-	1/2	-1	511.0 keV
electron neutrino	ν_e	1/2	0	< 2 eV
muon	μ^-	1/2	-1	105.7 MeV
muon neutrino	ν_μ	1/2	0	< 2 eV
tau	τ^-	1/2	-1	1.776 GeV
tau neutrino	ν_τ	1/2	0	< 2 eV
up	u	1/2	2/3	$2.2^{+0.6}_{-0.4}$ MeV
down	d	1/2	$-1/3$	$4.7^{+0.5}_{-0.4}$ MeV
strange	s	1/2	$-1/3$	96^{+8}_{-4} MeV
charm	c	1/2	2/3	1.28 ± 0.03 GeV
bottom	b	1/2	$-1/3$	$4.18^{+0.04}_{-0.03}$ GeV
top	t	1/2	2/3	173.21 ± 0.87 GeV
photon	γ	1	0	0
gluon	g	1	0	0
W	W	1	± 1	80.385 ± 0.015 GeV
Z	Z	1	0	91.1876 ± 0.0021 GeV
Higgs	H	0	0	125.09 ± 0.24 GeV

Table 2.2: A summary of spin, electric charge and mass of the known matter elementary particles in the SM [3]. The spin and electric charges are those stated in the SM. The quoted masses are those measured by experiment, except for u, d, s, c, b, γ and g where the theoretically calculated values are given. The unlisted anti-matter particles are assumed to have the same mass and spin as the matter particles but with opposite quantum number.

2.1 Local gauge symmetry

The principle of symmetry is the building block of the SM. The observational Lorentz symmetry principle states that the laws of physics do not change for an observer in different reference frames. The symmetry implies that the product of two spacetime four-vectors, x^μ and y_ν,

$$x^\mu y_\nu = \eta_{\mu\nu} x^\mu y^\nu, \tag{2.1}$$

with $\eta_{\mu\nu}$ being the Minkowski metric, is invariant under transformations to the four-vectors such as

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu, \tag{2.2}$$

where Λ^μ_ν is a Lorentz transform. This transformation is said to be global, in that it does not depend on the position in spacetime. A local transformation is said to be one that does depend on the position in spacetime. The interactions between the gauge bosons and the fermions are added to the theory by requiring a local gauge symmetry to the corresponding gauge group.

2.1.1 Quantum electrodynamics

The principle of gauge symmetry is most easily understood by applying transformations to the Dirac Lagrangian, defining the dynamics of a free spin-$\frac{1}{2}$ field, $\psi(x)$,

$$\mathcal{L}_{\text{Dirac}} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{2.3}$$

where m is the mass of the spinor field, γ^{μ} the 4×4 Dirac matrices and $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ the Dirac adjoint. If we apply a global transformation to the spinor field,

$$\psi \rightarrow \psi^{'} = e^{i\alpha}\psi, \tag{2.4}$$

where α is a real number and independent of the spacetime position, x, we see that the Lagrangian is invariant. However, under a *local* gauge transformation

$$\psi \rightarrow \psi^{'} = e^{ie\alpha(x)}\psi, \tag{2.5}$$

with the phase α now depending on x and adding an additional coupling constant e, the Lagrangian is no longer locally gauge invariant

$$\begin{aligned}
\mathcal{L}_{\text{Dirac}} &\rightarrow i\bar{\psi}e^{-ie\alpha(x)}\gamma^{\mu}\partial_{\mu}(e^{ie\alpha(x)}\psi) - m\bar{\psi}e^{-ie\alpha(x)}e^{ie\alpha(x)}\psi \\
&= i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - \bar{\psi}\gamma^{\mu}\psi e(\partial_{\mu}\alpha(x)) - m\bar{\psi}\psi \\
&= \mathcal{L}_{\text{Dirac}} - e(\partial_{\mu}\alpha(x))\bar{\psi}\gamma^{\mu}\psi
\end{aligned} \tag{2.6}$$

This problem is resolved by introducing the covariant derivative, $\mathcal{D}_{\mu}$,

$$\mathcal{D}_{\mu} = \partial_{\mu} + ieA_{\mu}, \tag{2.7}$$

where A_{μ} is a newly introduced vector field, and making the substitution $\partial_{\mu} \rightarrow \mathcal{D}_{\mu}$. $\mathcal{D}_{\mu}\psi$ is required to transform as ψ to ensure local gauge invariance and hence the newly introduced vector gauge field, A_{μ}, must transform as

$$A_{\mu} \rightarrow A_{\mu}^{'} = A_{\mu} + \partial_{\mu}\alpha(x) \tag{2.8}$$

The only other locally gauge invariant term involving A_{μ} that can be introduced to the Lagrangian is the kinetic term, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, where $F^{\mu\nu}$ is the field strength tensor given by the commutator of covariant derivatives,

$$F_{\mu\nu} = -\frac{i}{e}[\mathcal{D}_{\mu}, \mathcal{D}_{\nu}] = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \tag{2.9}$$

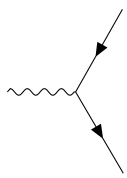
One notices that if the A_{μ} mass term, $\frac{m^2}{2}A_{\mu}A^{\mu}$, is added to the Lagrangian then local gauge invariance is not retained and hence A_{μ} is forced to be massless, $m = 0$.

Adding all the pieces to the Lagrangian results in $\mathcal{L}_{QED}$,

$$\begin{aligned}
\mathcal{L}_{QED} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\gamma^{\mu}(\partial_{\mu} + ieA_{\mu})\psi(x) - m\bar{\psi}\psi \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - e\gamma^{\mu}\bar{\psi}\psi A_{\mu} - m\bar{\psi}\psi.
\end{aligned} \tag{2.10}$$

The new term in the Lagrangian in Equation 2.10, $-e\gamma^{\mu}\bar{\psi}\psi A_{\mu}$, is an interaction term implying a coupling of the spinor field, ψ, with the vector field, A_{μ}, with coupling strength, e. The Feynman rule for this interaction is given by

A_{μ}



$$= ie\gamma_{\mu}. \tag{2.11}$$

A_μ is in fact the electromagnetic potential and e the electric charge of the fermion field. Hence, by simply requiring local gauge invariance in the theory, we have introduced a new massless vector field that corresponds to, under a more complete electroweak treatment (see Chapter 2.3.2), the electromagnetic force mediator; the photon.

2.1.2 Non-abelian gauge symmetry

In Section 2.1.1, the local gauge transformations can actually be considered to be one dimensional unitary matrix transformations, U. The group of these unitary matrices is $U(1)$ and the local gauge symmetry is an abelian (commutative) $U(1)$ gauge invariance. The idea of group symmetry is trivial for the one dimensional case but becomes intrinsic in more complicated group symmetries.

For the non-abelian $SU(N)$ groups, a local gauge transformation is of the form

$$\begin{aligned} \psi \to \psi' &= U(x)\psi \\ &= e^{-ig\alpha^i(x)T^i}\psi \end{aligned} \tag{2.12}$$

where T^i are the N^2-1 generators of the $SU(N)$ group and g is a coupling constant. $U(x)$ is of the form of a unitary $N \times N$ matrix with unit determinant, defined as the fundamental representation. The generators, T^i, satisfy the Lie algebra

$$[T^i, T^j] = if^{ijk}T^k \tag{2.13}$$

with f^{ijk} being the structure constants of the group. Analogously to Equation 2.10, a covariant derivative is defined to ensure local gauge invariance under $SU(N)$ transformations

$$\mathcal{D}_\mu = \partial_\mu + igT^iA^i_\mu \tag{2.14}$$

with the gauge fields, A^i_μ, being linear combinations of the generators, T^i. As with the abelian case, $\mathcal{D}_\mu\psi$ is required to transform like ψ and thus A_μ must transform like

$$A_\mu \to A'_\mu = U(x)A_\mu U^{-1}(x) + \frac{i}{g}U(x)[\partial_\mu U^{-1}(x)] \tag{2.15}$$

to retain the symmetry. The field strength tensor is found to be

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]^i \\ &= \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + gf^{ijk}A^j_\mu A^k_\nu \end{aligned} \tag{2.16}$$

The final term is due to the non-abelian nature of $SU(N)$ and is responsible for self (and higher order) couplings of the gauge fields A^i_μ.

2.1.3 Quantum chromodynamics and $SU(3)_c$ symmetry

Quantum chromodynamics (QCD) is the theory that governs the dynamics of particles that couple via the strong force. QCD has a $SU(3)_c$ local gauge symmetry, with c denoting colour. There are eight generators, $T_a = \frac{1}{2}\lambda^a$ with λ^a the Gell-Mann matrices, which define the transformations

$$\psi \to \psi' = e^{-ig_s\alpha^a(x)T^a}\psi, \tag{2.17}$$

with a new coupling constant for the strong force, g_s, being introduced. The generators obey the commutation relation

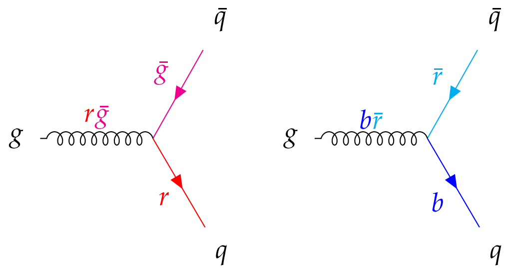
$$[T^a, T^b] = if^{abc}T^c \tag{2.18}$$

defining the Lie algebra of the group discussed in Section 2.1.2.

The quark fields carry *colour* charge and transform as a triplet in the fundamental representation,

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}. \tag{2.19}$$

The r, b and g indices correspond to the specific colour charge of the quark; called (somewhat arbitrarily) red, blue and green. The eight new gauge fields required to retain local gauge symmetry belong to the gluon, g. The gluon also has colour charge but in a different way to the quarks. Gluons have both colour and anticolour ($r\bar{g}$, $b\bar{r}$ etc.) and act by converting the colour of two quarks that couple with it, conserving colour at the vertex[1]:



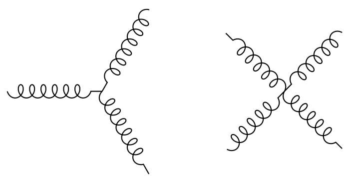
As in Section 2.1.2, we define a covariant derivative that ensures local gauge invariance under $SU(3)_c$ transformations

$$\mathcal{D}_\mu = \partial_\mu + ig_s \frac{1}{2}\lambda^a G^a_\mu, \tag{2.20}$$

with the G^a_μ corresponding to the eight gluon fields. If the QCD kinetic term is evaluated with the newly defined $F^a_{\mu\nu}$

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu, \tag{2.21}$$

we find that self and quartic couplings of gluons are allowed by the theory:



These self couplings contribute to the two distinct features of QCD; asymptotic freedom and confinement, discussed below.

The strong coupling constant is commonly written as

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{2.22}$$

It is found that, in higher orders of perturbation theory, the strong coupling α_s depends on the energy scale of interactions, Q^2. Comparisons of theoretical calculations and measurements of this running behaviour can be seen in Figure 2.1. The scale dependence can be seen clearly in the one-loop approximation of α_s,

$$\alpha_s(Q^2) \approx \frac{1}{\beta_0 \ln \frac{Q^2}{\Lambda^2}}. \tag{2.23}$$

[1] In fact the colour charge representation of the gluons is not quite as simple as shown here; the different *colours* are in a superposition of states, such as $(r\bar{g} + \bar{r}g)/\sqrt{2}$.

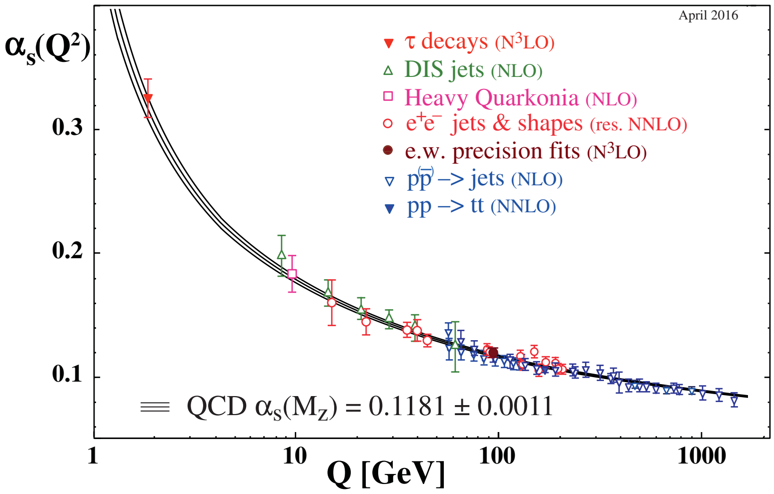


Figure 2.1: Measurements of the value of the strong coupling constant, α_S, as a function of momentum transfer, Q [3]. The respective order of QCD perturbation theory used in the extraction of α_s is indicated in brackets. The order of calculations is introduced in Chapter 3.

Here β_0 is the first order perturbative constant such that

$$\beta_0 = \frac{33 - 2N_f}{12\pi} \tag{2.24}$$

with N_f the number of quark flavours that can appear in the loop. The QCD scale, Λ, is introduced and is typically set to the energy scale at which α_s starts to get very large ($\Lambda \sim$ 200 MeV). So at high energy scales we have that α_s is small which gives quarks *free* behaviour at high energies (asymptotic freedom), and at low energies α_s is large leading to so-called confinement; the observation that individual free quarks are not observed and instead are detected in strongly bound colourless states called hadrons.

2.1.4 The weak interaction and $SU(2)_L$ symmetry

The method for constructing gauge invariant Lagrangians for QED and QCD can now be turned to the weak interaction, which has a $SU(2)$ gauge symmetry. The charged weak interaction is known to only couple to left-handed leptons, as measured by parity violation measurements [4, 5]. Left or right-handedness refers to the chirality of the particle. The left or right-handed projections of a spin-$\frac{1}{2}$ field are given by the operators

$$P_R = \frac{1}{2}(1+\gamma^5), \qquad P_L = \frac{1}{2}(1-\gamma^5),$$

acting on the spinor, $\psi = \psi_R + \psi_L$,

$$\psi_{R,L} = P_{R,L}\psi. \tag{2.25}$$

To align with the observation of parity violation, only the left-handed neutrino is allowed to interact via the weak force. Hence left-handed fields are made to transform as weak isospin doublets in $SU(2)_L$, and the right-handed fields as scalars. For example the first generation left-handed lepton and quark fields are given as

$$\psi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \qquad \psi_L = \begin{pmatrix} u \\ d' \end{pmatrix},$$

with the notation for the quark doublets discussed further in Section 2.4. In this form there is no coupling of right-handed neutrinos in the weak interaction by construction.

We can introduce the weak isospin quantum number, I, analogously to charge, Q, for the electromagnetic interaction. Left-handed isospin doublets have $I = \frac{1}{2}$ with the upper and lower members of the third component of isospin, I_3, satisfying $I_3 = \pm\frac{1}{2}$. The right-handed fermions thus have $I = I_3 = 0$.

$SU(2)$ gauge transformations to the isospin fields are of the form

$$\psi \to e^{ig\alpha^i(x)T^i}\psi, \tag{2.26}$$

with the generators $T^i = \frac{1}{2}\sigma^i$ and σ^i being the three Pauli spin matrices. A new coupling constant for the weak interaction is also introduced, g. The antisymmetric Levi-Civita tensor, ϵ^{ijk}, defines the structure constants of the group and the covariant derivative that enables gauge invariance for $SU(2)_L$ is given by

$$\mathcal{D}_\mu = \partial_\mu + ig\frac{1}{2}\sigma^i W^i = \partial_\mu + ig\frac{1}{2}\begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix}, \tag{2.27}$$

with three new gauge fields, W^i, being introduced.

2.2 The electroweak interaction

A $SU(2)_L \times U(1)_Y$ symmetry represents the unified electroweak theory [6, 7, 8]. It is based on the conservation of hypercharge, Y, related to the third component of weak isospin, I_3, and electrical charge, Q, by

$$Y = 2(Q - I_3). \tag{2.28}$$

The introduction of the hypercharge quantum number is essential to the unification of the two different interactions; there are no right-handed fermion interactions in the weak $SU(2)_L$ sector, but there are in the electromagnetic theory. The generator of the $U(1)_Y$ group is now hypercharge, and the covariant derivative for the electroweak interaction can be defined

$$\mathcal{D}_\mu = \partial_\mu + ig'\frac{1}{2}YB_\mu + ig\frac{1}{2}\sigma^i W^i, \tag{2.29}$$

with a new gauge field B_μ being introduced to ensure $U(1)$ gauge invariance and defining a new coupling constant for $U(1)_Y$, g'. The relation of g' to elementary electrical charge, e, is discussed in Section 2.3.2.

Now we have a unified electroweak interaction. So far we have been forced to set the mass of the gauge fields in the theory to zero to ensure the Lagrangian remains gauge invariant. However, we know the $W^\pm$ and Z are *massive* bosons. It should also be noted that the fermion mass term discussed as a component of the Dirac Lagrangian, $m\bar{\psi}\psi$, is no longer gauge invariant in $SU(2)_L$ due to the mixing of left- and right-handed fermion fields. Both gauge bosons and fermions now need to gain mass in some way.

2.3 The Higgs mechanism

So far we have added gauge bosons to the theory by requiring local gauge invariance of the Lagrangian. However, we know the typical mass terms for the gauge bosons in the Lagrangian,

$\frac{1}{2}m_A^2 A_\mu A^\mu$, cannot be explicitly added since these terms break local gauge symmetry. This problem in the SM is solved by spontaneously breaking the local electroweak gauge symmetry with the inclusion of a new complex scalar field [9, 10, 11]. This same mechanism produces mass terms for fermions in a natural way.

2.3.1 Spontaneous symmetry breaking with a complex scalar field

The complex scalar field, $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, has Lagrangian

$$\begin{aligned}\mathcal{L} &= (\partial_\mu \phi)^*(\partial^\mu \phi) - V(\phi)\\ &= (\partial_\mu \phi)^*(\partial^\mu \phi) - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2.\end{aligned} \tag{2.30}$$

The parameters μ and λ are chosen so that $\mu^2 < 0$ and $\lambda > 0$, giving the potential, $V(\phi)$, a *Mexican hat* shape, as shown in Figure 2.2. This potential now has degenerate minima along the

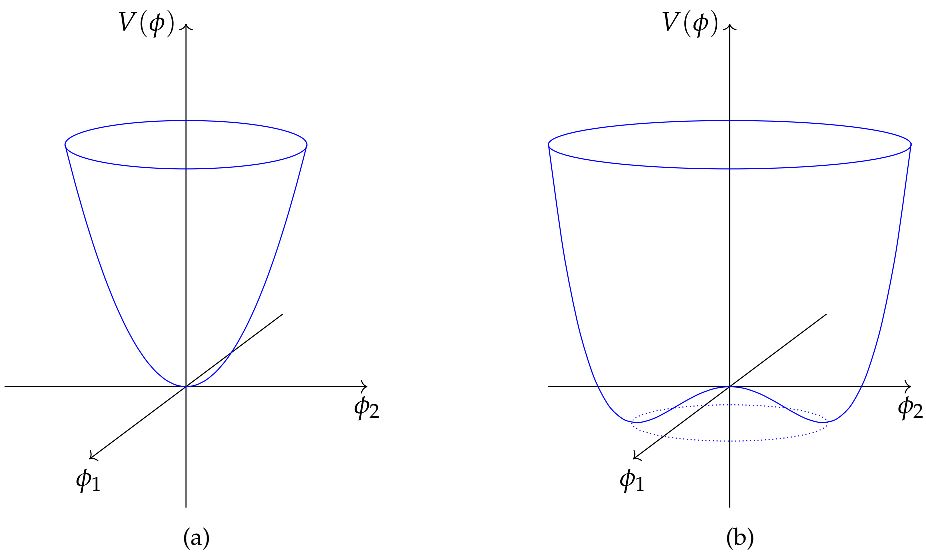


Figure 2.2: A diagram illustrating the *Mexican hat* potential of the Higgs field in terms of the real (ϕ_1) and the imaginary (ϕ_2) components of the complex scalar field, ϕ, for $\lambda > 0$ and (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$.

complex circle, described by

$$\phi^*\phi = \phi_1^2 + \phi_2^2 = \frac{-\mu^2}{\lambda} = v^2, \tag{2.31}$$

with v the now non-zero vacuum expectation value of the scalar field corresponding to the value of ϕ at the new minima of $V(\phi)$. Note that there is only one minimum at $\phi_1 = \phi_2 = 0$ if μ was required to satisfy $\mu^2 > 0$, as seen in Figure 2.2a.

One notices that the Lagrangian in Equation 2.30 has a global $U(1)$ gauge symmetry for transformations such as $\phi \to \phi' = e^{-i\alpha}\phi$. If we choose an arbitrary vacuum state from the complex circle of minima defined by the new potential, so that $\phi \neq 0$, and apply a global gauge transformation, then ϕ is transformed to a new vacuum state on the circle of minima which *spontaneously* breaks the $U(1)$ gauge symmetry. The Lagrangian itself remains gauge invariant. This process is known as spontaneous symmetry breaking.

The result of such symmetry breaking in the theory can be studied by exciting a chosen vacuum state. Choosing the real minimum, $(\phi_1, \phi_2) = (v, 0)$, for the vacuum gives

$$\phi(x) = \frac{v + \eta(x) + i\rho(x)}{\sqrt{2}}, \tag{2.32}$$

with η and ρ being new fields reflecting the deviation from the true ground state. The Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho)(\partial^\mu \rho) + \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) + \mu^2 \eta^2 - \lambda(\eta\rho^2 + \eta^3) - \frac{\lambda}{2}\eta^2\rho^2 - \frac{\lambda}{4}\eta^4 - \frac{\lambda}{4}\rho^4. \tag{2.33}$$

Note that now the Lagrangian contains a massive scalar field, η, with mass

$$-\frac{1}{2}m_\eta^2 = \mu^2 \Rightarrow m_\eta = \sqrt{-2\mu^2}, \tag{2.34}$$

and a massless scalar field, ρ (due to the vanishing ρ^2 term). The ρ is an unphysical Goldstone boson that is theorised to appear whenever a continuous global symmetry is spontaneously broken.

If one instead considers a local gauge transformation of ϕ then the generic gauge field A_μ, that is introduced to retain the local gauge symmetry as discussed, itself gains a mass term through the interaction with the massive scalar field. The method by which the $W^\pm$ and Z bosons gain their mass through this mechanism is discussed below.

2.3.2 The Standard Model Higgs mechanism

In the SM the local electroweak gauge symmetry is spontantaneously broken. The scalar field that does this is the Higgs field which is written as an $SU(2)_L$ isospin doublet,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.35}$$

containing the charged, ϕ^+, and neutral, ϕ^0, complex scalar fields. The Lagrangian now has the form

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2, \tag{2.36}$$

with $\mathcal{D}_\mu$ the covariant derivative for the electroweak $SU(2)_L \times U(1)_Y$ symmetry. The potential $V(\phi)$ has minima at $\phi^\dagger \phi = v^2$, and thus a vacuum state can be chosen:

$$\begin{aligned} &\phi_1 = \phi_2 = \phi_4 = 0, \\ &\phi_3 = v. \end{aligned} \tag{2.37}$$

Notice that only the real component of the neutral scalar field is given a non-zero value. The physical Higgs field, $H(x)$, can now be introduced by expanding around this vacuum state

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.38}$$

This choice of real vacuum state is called the unitary gauge. Using Equation 2.28 we can state $Y = 1$ for the Higgs doublet and thus, by evaluating the $(\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi)$ term in the Lagrangian in Equation 2.36, we find there are two mass terms for the gauge fields,

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= \frac{1}{8}v^2 g^2((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{8}v^2(g' B_\mu - g W_\mu^3)^2 \\ &= \frac{1}{4}v^2 g^2 W_\mu^+ W^{-\mu} + \frac{1}{8}v^2(g' B_\mu - g W_\mu^3)^2, \end{aligned} \tag{2.39}$$

using the relation,

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp iW^2), \tag{2.40}$$

in the second line. The first term in Equation 2.39 is quadratic in $W^{\pm}$ and hence is the expected mass term. M_W is thus given by

$$m_W^2 = \frac{1}{4}v^2 g^2 \Rightarrow m_W = \sqrt{\frac{v^2 g_2^2}{4}} = \frac{vg}{2}. \tag{2.41}$$

with the usual factor $\frac{1}{2}$ emitted due to the mass term describing the two W bosons. The second term in Equation 2.39 contains the two neutral B_μ and W_μ^3 fields. If one introduces the so-called weak mixing angle, θ_W,

$$g\sin\theta_W = g'\cos\theta_W = e, \tag{2.42}$$

then the B_μ and W_μ^3 fields can be mixed into the observable Z and photon fields, Z_μ and A_μ:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \tag{2.43}$$

In this form only a mass term for the Z_μ field is retained,

$$m_Z = \frac{gv}{2\cos\theta_W}, \tag{2.44}$$

and the mass of A_μ is zero. The photon field is also forced to only couple to charge, as observations require.

Three of the four degrees of freedom of the Higgs field are used to give mass to the weak bosons. The final degree of freedom also gives the excitation of the Higgs field, the Higgs boson, a mass.

2.3.3 The Yukawa interaction

The Higgs mechanism also provides gauge invariant mass terms in the Lagrangian for fermions. The standard mass term in the Dirac Lagrangian is

$$-m\bar{\psi}\psi = -m\bar{\psi}_L\psi_R - m\bar{\psi}_R\psi_L \tag{2.45}$$

This is not invariant under $SU(2)_L$ gauge symmetry. Instead mass terms can be produced from the Higgs mechanism by specifying a coupling, y, between fermions and the Higgs which is $SU(2)_L$ invariant:

$$\mathcal{L}_{\text{Yukawa}} = -y(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\bar{\phi}\psi_L). \tag{2.46}$$

This coupling is known as the Yukawa coupling. This term not only describes the interaction of the Higgs field with the fermions but also a fermion mass term when the Higgs field has a non-zero expectation value, as shown in Equation 2.37. Taking the first generation lepton doublet as an example,

$$\begin{aligned}
\mathcal{L}_e &= -y_e[\begin{pmatrix} \bar{\nu} & \bar{e} \end{pmatrix}_L \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} e_R + \bar{e}_R \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & v+H \end{pmatrix} \begin{pmatrix} \nu \\ e \end{pmatrix}_L] \\
&= -\frac{y_e(v+H)}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) \\
&= -\frac{y_e(v+H)}{\sqrt{2}}\bar{e}e \\
&= -\frac{y_e v}{\sqrt{2}}\bar{e}e - \frac{y_e}{\sqrt{2}}H\bar{e}e.
\end{aligned} \tag{2.47}$$

The first term is the mass term for the electron, $m_e = \frac{y_e v}{\sqrt{2}}$, and the second the Higgs-electron interaction term. One notices that the coupling of the Higgs with a lepton pair is proportional to the mass of the lepton and that, in the unitary gauge, the neutrino does not interact with the Higgs or gain a mass as required, due to the charged upper element of ϕ being chosen as zero.

2.4 Electroweak quark sector

The Higgs mechanism can also be used to give mass to the quarks. Analogously to the leptons, we can construct $SU(2)_L$ isospin doublets

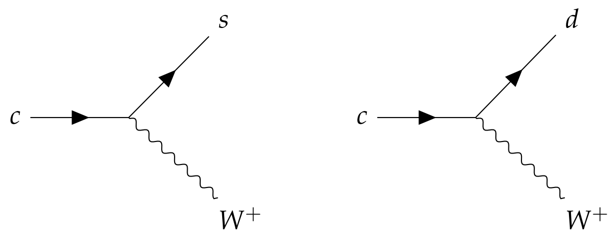
$$\psi_L = \begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} c \\ s' \end{pmatrix} = \begin{pmatrix} t \\ b' \end{pmatrix}. \tag{2.48}$$

The d', s' and b' correspond to the weak eigenstates of the d, s and b quarks, which are themselves in mass eigenstate form. These mass eigenstates do not correspond directly to the weak eigenstates that are concurrent in electroweak interactions. The difference between mass and weak eigenstate is analogous to neutrinos; this is the mechanism that drives flavour-changing oscillations [12].

The mass eigenstates can be rotated into the weak eigenstate basis via the CKM matrix, V_{CKM},

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.49}$$

with the square modulus of the CKM $V_{qq'}$ elements, $|V_{qq'}|^2$, describing the probability of the transition $q \to q'$ of two different quark flavours q, q'. This results in the possibility of inter-generation flavour changing quark decays. For example, there is an approximate 5% probability (from $|V_{cd}|^2$) that a c-quark will decay to a d-quark (right), rather than the expected s-quark (left):



The above processes are flavour changing charged current processes, since the processes are mediated by the charged W boson. It should be noted that so-called flavour changing neutral current processes are not observed (at tree level) in nature and are forbidden in the SM.

The quark isospin doublets are acted on in the same way as Equation 2.47 to give mass to the down-type quarks. The up-type quarks require the conjugate Higgs doublet, ϕ_c, to spontaneously break the gauge symmetry,

$$\phi_c = \begin{pmatrix} \bar{\phi}^0 \\ -\phi^- \end{pmatrix}, \tag{2.50}$$

which results in the required form for up-type masses in the unitary gauge

$$\phi_c = \frac{1}{\sqrt{2}} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix}. \tag{2.51}$$

Hence, the masses of the quarks, m_q, both up- and down-type, relate to the corresponding Yukawa coupling, y_q, in the same way as the leptons:

$$m_q = \frac{y_q v}{\sqrt{2}}. \tag{2.52}$$

This is the final piece of the puzzle. By requiring a local $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge symmetry and introducing a heavy scalar boson to spontaneously break this symmetry, we have arrived at the Standard Model of particle physics.

Chapter 3
The physics of hadron colliders

This thesis focuses on the analysis of proton-proton (pp) collisions measured with the ATLAS detector. The study of these collisions allows one to test the predictions of the SM. However, from the quark model we know that protons themselves are not elementary particles; they are composite particles called hadrons (specifically baryons), made up of quarks and gluons. The kinematics of these composite particles are more complicated than the bound *partons* themselves. The physics of the collisions of these partons, and by extension the proton, is discussed below.

3.1 Cross sections

A particularly important quantity to test the validity of the SM is the cross section (σ) of a particular process. The cross section is a quantum-mechanical probability for a particular particle interaction to take place. The total inclusive cross section for such a process can be measured and the cross section can in addition be measured differentially; as a function of kinematic properties of the final state. This allows for a plethora of theoretical calculations to be validated experimentally.

The differential scattering cross section, $d\sigma$, for two incoming particles, i, resulting in some final state, f, is given by

$$d\sigma = \frac{|\mathcal{M}_{fi}|^2}{F} d\Phi_N, \tag{3.1}$$

where $\mathcal{M}_{fi}$ is the Lorentz-invariant matrix element governing the probability of the transition $i \to f$, F is a flux factor and $d\Phi_N$ a Lorentz invariant phase space factor for the N final state particles [13]. The $d\Phi_N$ can be written as

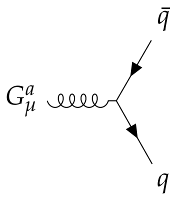
$$d\Phi_N = (2\pi)^4 \delta^4(p_1 + p_2 - \sum_i^N p_i) \prod_i^N \frac{dp_i^4}{(2\pi)^4} (2\pi)\delta(p_i^2 - m_i^2), \tag{3.2}$$

where $p_{1,2}$ and p_i are the initial and final state particle 4-momentum respectively and m_i is the mass of a final state particle. The first delta function ensures conservation of momentum between the initial and final states and the second to ensure produced particles are *on-shell*; they satisfy the classical equations of motion.

It follows that cross section calculations typically involve integrals with a large number of dimensions. They can thus be extremely computationally expensive to compute analytically. These integrals are estimated using Monte Carlo (MC) simulation, described in Section 3.2.

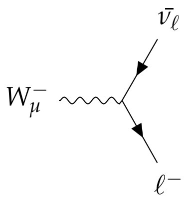
3.1.1 Matrix elements

The matrix element, $\mathcal{M}$, is calculated using interaction terms determined from Feynman diagrams for the interesting process. The interaction term of an electron-positron pair with a photon is given in Equation 2.11 above. Vertices for a quark-gluon interaction



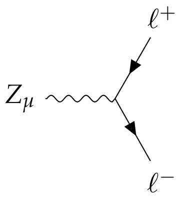
$$= i g_s \gamma_\mu \frac{1}{2} \lambda^a, \tag{3.3}$$

and leptons, ℓ, interacting with a W boson



$$= -i \frac{g}{2\sqrt{2}} \gamma_\mu (1 - \gamma^5), \tag{3.4}$$

and a Z boson



$$= i \frac{g}{4\cos\theta_W} \gamma_\mu (1 - 4\sin^2\theta_W - \gamma^5), \tag{3.5}$$

are given with their corresponding interaction terms. These terms are used in the calculation of the matrix element for the corresponding process. It is possible for different diagrams to contribute to the same process (i.e. same final state) and thus interaction terms must be summed to obtain the full matrix element. The addition of multiple processes can result in constructive or destructive quantum-mechanical interference.

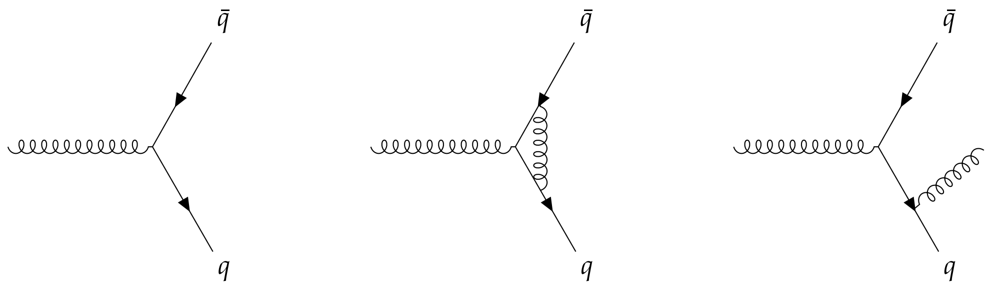
3.1.2 Next-to-leading order corrections

So far the matrix elements are calculated from Feynman diagrams at *tree level*; at the first order of perturbation theory. The interaction terms are taking the first relevant term in the perturbative series in the relative coupling constant. In fact, Feynman diagrams with higher order terms of the coupling constant can contribute to the cross section in question. These higher order terms can come in the form of loop diagrams or real emissions of partons. A tree level, or leading-order (LO), and two next-to-leading order (NLO) QCD Feynman diagrams for a gluon splitting into a pair of quarks ($g \rightarrow q\bar{q}$) are shown in Figure 3.1.

Divergences appear when calculating NLO cross sections in QCD. These occur when a gluon is emitted collinearly from a quark or at very low energy (infrared divergences) and at high energy (ultraviolet divergences). The ultraviolet divergences are mitigated using renormalisation procedures. The infrared divergences are worked around by applying infrared subtraction algorithms that aim to cancel the virtual and real emission terms, in Figures 3.1b and 3.1c respectively.

3.1.3 Cross sections at a hadron collider

So far only cross sections for incoming elementary particles has been discussed. Protons are however not elementary particles.



(a) (b) (c)

Figure 3.1: Feynman diagrams for $t\bar{t}$ at (a) tree level and (b,c) NLO, where (b) shows a virtual emission of a gluon and (c) a real emission.

The differential cross section for a process with two initial protons, p_1 and p_2, to some final state X is defined as the sum of partonic cross sections, $\hat{\sigma}_{q_1 q_2 \to X}$,

$$d\sigma_{p_1 p_2 \to X} = \int dx_1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1, \mu_F^2) f_{q_2}(x_2, \mu_F^2) d\hat{\sigma}_{q_1 q_2 \to X}(x_1 x_2 s, \mu_R, \mu_F). \tag{3.6}$$

The $f_q(x, \mu_F^2)$ are parton distribution functions (PDFs). These PDFs describe the probability density that parton q, carrying momentum fraction x, of hadron p, at scale μ_F^2, enters the hard scatter process. Hard refers to the amount of transferred energy between the partons and thus events with large momentum in the plane transverse to the collision are hard.

Parton distribution functions are determined from different types of experimental data, typically including DIS (deep inelastic scattering) data from the electron-proton HERA collider. They are parameterised as a function of x for a reference scale Q_0^2. The evolution of the PDF to different scales, Q^2, are calculated using the DGLAP equations [14, 15, 16] and model the changes in momenta of different species of parton due to gluon emissions and splittings.

The MSTW NLO PDF sets are shown in Figure 3.2, at scales of $Q^2 = 10\,\text{GeV}^2$ and $Q^2 = 10^4\,\text{GeV}^2$. Roughly half the total momentum of the proton at any one time is from gluons, with the other half typically being from up and down quarks, as one would expect. The probability of heavier quarks contributing to the PDF increases with Q^2. The CTEQ [17] and NNPDF [18] PDF sets are those commonly used in simulation at the LHC.

The scale that PDFs are defined at is called the factorisation scale, μ_F^2. This scale is set such as to separate long and short distance interactions. The calculation of the cross section is then *factorised*; perturbation theory is used to calculate the high energy (small α_s) partonic cross sections, and the soft processes (large α_s) that cannot be calculated in perturbation theory are enveloped into the experimentally determined PDF; thus considered to be part of the hadronic structure. The choice of factorisation scale is somewhat arbitrary and introduces a theoretical uncertainty in the calculation of the cross section. This uncertainty can be deduced by modifying the factorisation scale and observing the difference in calculated cross section.

3.2 Monte Carlo simulation

Calculating cross sections is a difficult task. A typical event can contain hundreds of particles in the final state, and thus the resulting integrals contain a large number of dimensions. Monte Carlo (MC) techniques in event generation are ideally suited to the task. A simple description

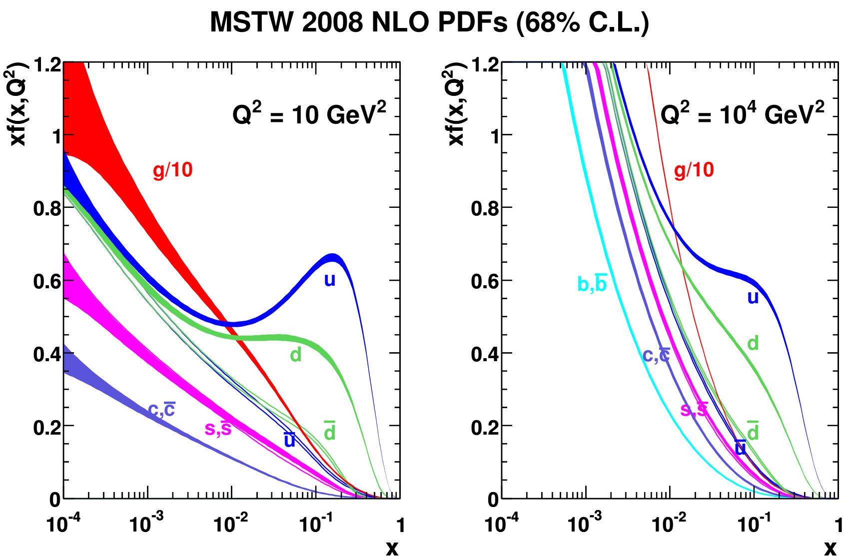


Figure 3.2: Parton distribution functions (PDFs) $xf(x, Q^2)$ as a function of parton momentum fraction, x, at scales $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right) [19].

of MC simulation is given in Section 3.2.1, the event generation procedure in Section 3.2.2 and event generators used in ATLAS MC simulation in Section 3.2.3.

3.2.1 Simple description

Suppose there is a variable, I, which is calculated by integrating a function f over an arbitrary variable x in the range x_2 to x_1 in one dimension,

$$I = \int_{x_1}^{x_2} f(x)dx. \tag{3.7}$$

The Monte Carlo technique estimates this integral by a finite sum

$$I \approx I_{MC} = (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{3.8}$$

for N generated *events*, with I_{MC} the estimate of I. Taking the limit $N \to \infty$ results in the convergence of I_{MC} to the true value I, $I_{MC} \to I$.

It can be shown, using the central limit theorem, that

$$I \approx I_{MC} \pm \sqrt{\frac{\sigma^2(f)}{N}} \tag{3.9}$$

with $\sigma^2(f)$ the variance of the function f. The average error on I_{MC}, $\sigma_{I_{MC}}$, is thus given by

$$\sigma_{I_{MC}} = \frac{\sigma(f)}{\sqrt{N}}, \tag{3.10}$$

with $\sigma(f)$ the corresponding standard deviation of $f(x)$. That is the error in the MC estimate of the integral scales by $1/\sqrt{N}$. This formalism is trivially extended to n-dimensional integrals, with the error still scaling by $1/\sqrt{N}$ [20].

3.2.2 Particle event generation

The MC event generation in hadronic physics is split up into a number of steps.

3.2.2.1 Matrix elements

Firstly, as discussed above, the matrix element amplitudes are calculated using perturbation theory at the desired level, with the phase space integration being performed by MC techniques. Typically matrix elements are calculated at LO or NLO. The procedure for LO diagrams contributing to the process is generally automated with modern MC event generators. Some event generators automate the NLO calculations but otherwise these are performed manually. Typically physics processes in this thesis are modelled and simulated at NLO accuracy.

3.2.2.2 Parton shower

The Feynman diagrams that are used to calculate matrix element amplitudes are just one part of the picture at a hadron collider. One needs to understand how these partons evolve. A final state parton would never be measured as such, but instead as a number of clustered colourless composite hadrons, called a jet.

The coloured partons involved in the hard process radiate partons sequentially until observable low energy final state hadrons are produced. This *showering* can be modelled by QCD perturbation theory from the high momenta of the hard process down to the momentum scale of confinement, where the theory breaks down and becomes no longer valid. However, the amplitude calculations for these high multiplicity events become extremely complex and *parton shower* algorithms are used to augment the matrix element calculations. These parton shower algorithms treat the radiation of additional partons, including the effect of soft and collinear emitted partons, that would otherwise cause divergences in the integrals. *Matching* is required between the matrix element and parton shower calculations to avoid double counting of parton radiation in regions of phase space to which both can contribute.

3.2.2.3 Hadronisation

We know that individual partons are not observed in nature, but combine to form hadrons. The parton shower evolves the hard scatter to partons at low energies, but cannot describe how these partons are bound into hadrons due to the break down of perturbation theory at large α_s. These low energy partons are converted to the colourless hadrons using hadronisation models, utilising Lorentz invariance to model how low energy QCD evolves. The common models used by modern MC event generators are the Lund string model [21, 22] and the cluster model [23].

3.2.2.4 The underlying event

The underlying event is defined as any additional activity in an event not specifically from the hard scatter partons. The other partons within the protons that are not involved in the hard scatter process (the *spectators*) still have a high probability of interaction with one another, separate from the hard scatter. Such events are typically soft but can modify the colour flow in an event in such a way as to modify final state predictions. One such modification could be the multiplicity of hadronic jets within an event.

3.2.2.5 Pile-up

Typically bunches of hadrons (in this case protons) at hadron colliders are accelerated and brought together for collision. Another effect similar to the underlying event is *pile-up*: additional (usually soft) proton-proton collisions producing activity in the event at the same time as the hard collision. The pile-up can be in-time or out of-time, referring to whether the soft collision occurred in the same bunch crossing as the hard scatter of interest or another respectively. These events are simulated separately from the hard scatter as inelastic proton-proton collisions, or minimum bias events. They are then later overlaid on the simulated hard scatter.

3.2.3 Event generators

There are a number of MC event generators that generate events for the different stages discussed above. The main general purpose MC event generators are PYTHIA [24, 25, 26], HERWIG [27, 28] and SHERPA [29]. As discussed above, all these generators can automatically calculate LO matrix elements for the process in question. They then run through the full chain of generation; calculating phase space integrations and then applying parton shower algorithms, hadronisation and underlying event models. These generators are sometimes called shower MC generators, since they model the full event.

If NLO accuracy is requested then NLO MC event generators are also available, such as POWHEG [30, 31] and MC@NLO [32]. The NLO matrix elements produced from these generators, only containing bare partons, are matched to the general purpose generator of choice which models the remaining event evolution. In fact, the general purpose SHERPA generator can also automatically generate NLO matrix elements and match them to its internal parton shower algorithm [33].

So far events have only been produced at *particle level*. When comparing simulated MC events with real data from an experiment, one needs to understand how the raw particle level events are affected by their passage through the detector. Typically the generated MC events above are passed through the GEANT4 [34] simulation toolkit to model such effects. The detector response is then modelled by a *digitisation* stage before particle reconstruction is performed.

Chapter 4

The ATLAS detector at the Large Hadron Collider

The ATLAS experiment [35] is a general purpose particle detector situated in the Large Hadron Collider (LHC) at CERN near Geneva, a particle accelerator built across the French-Swiss border [36].

The LHC and surrounding accelerator complex is discussed in Section 4.1 and each component of the ATLAS detector in Section 4.2.

4.1 The Large Hadron Collider

The LHC is a 27 km circumference circular synchotron particle accelerator, capable of colliding particles with a centre-of-mass energy of up to $\sqrt{s} = 14$ TeV. The nominal physics program at the LHC utilises proton-proton collisions, but also has the capability of colliding heavy ions[1]. Only proton-proton collisions are considered in this thesis.

The full CERN accelerator complex is shown in Figure 4.1.

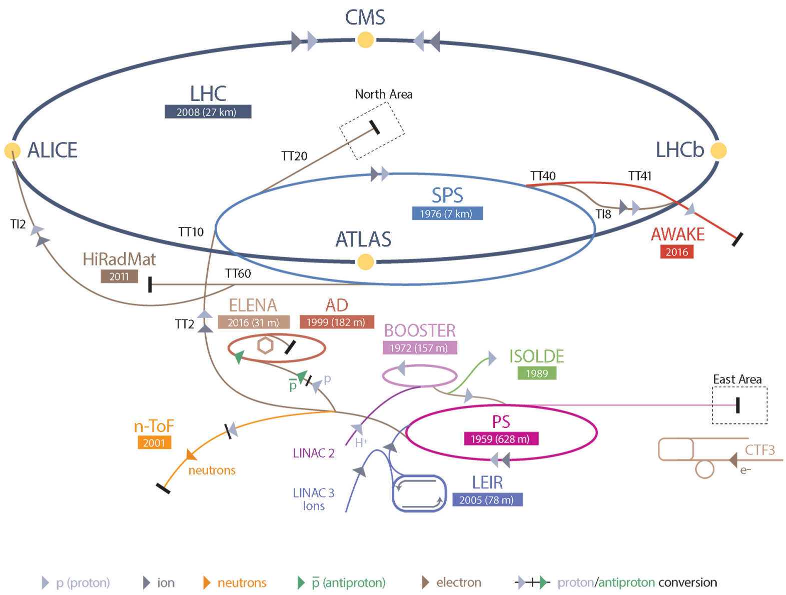


Figure 4.1: The CERN accelerator complex [37].

The LHC ring is not entirely circular; it contains eight 530 m straight sections with arcs in between each. The ring itself contains two beam pipes for the counter rotating protons, with the beams being brought together for collision and detection by four major experiments in the straight sections. The other general purpose particle physics experiment on the ring is the CMS experiment [38]. The ATLAS and CMS experiments are labelled "general purpose" since they are designed to cover as much solid angle around their collision points as possible to ensure

[1]Such as lead (Pb) ions and, in 2017, intermediately heavy xenon (Xe) ions.

maximum detection of particles from interesting physics processes. They were primarily built to detect the Higgs boson, which was discovered in 2012 [1, 2], and to search for any BSM physics. The other two main experiments on the LHC ring are the LHCb experiment [39], a forward detector designed for precision heavy-flavour quark measurements, and the ALICE experiment [40], designed to study quark-gluon plasma in heavy ion collisions. Three smaller experiments exist on the ring located near to three of the four main collision points; the LHCf experiment [41], designed to measure neutral particles emitted in the forward regions of LHC collisions; the MoEDAL experiment [42], designed to search for magnetic monopoles amongst other exotica; and the TOTEM experiment [43], designed to study diffractive scattering and to measure the total elastic proton-proton cross section at the LHC.

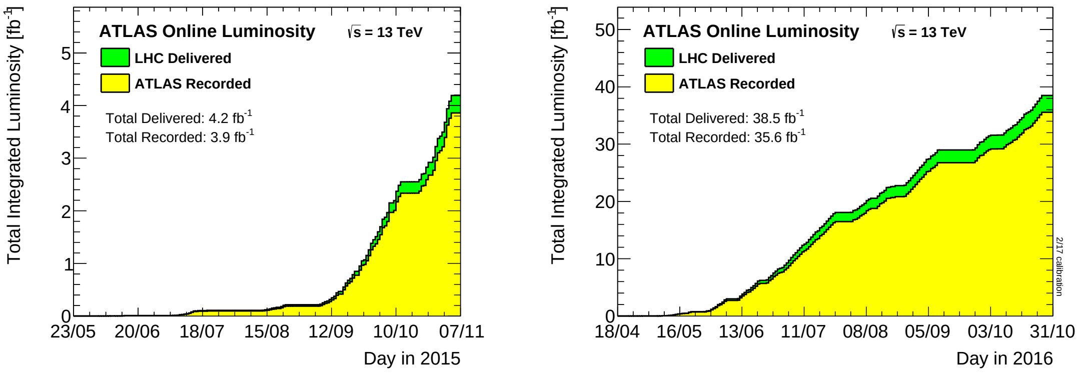
A combination of linear and smaller synchotron accelerators, known as the LHC injection chain, consecutively increase the energy of the proton beams provided to the LHC. The final accelerator in the chain is the Super Proton Synchotron (SPS), which injects 450 GeV protons into two of the straight sections of the LHC ring. The SPS itself has a rich physics program associated with it, with the discoveries of the W and Z bosons at the UA1 and UA2 experiments [44, 45, 46, 47] and the discovery of charge-parity (CP) violation in neutral kaon decays at the NA48 experiment [48] particular highlights . The beams are captured and accelerated up to the maximal proton energy of 7 TeV in 400 MHz superconducting radio-frequency (RF) cavities. The cavities sort the beam into *buckets* in 2.5 ns spacing; one in ten bunches are nominally filled with protons giving a bunch spacing of 25 ns. This allows for a maximum of 3564 bunches in the LHC at any one time, which is decreased to 2808 due to operational limitations.

The proton beams are bent in the arcs of the LHC ring by 1252 superconducting dipole magnets, capable of producing a magnetic field of up to 8.3 T. The dipole magnet system provides oppositely directed magnetic fields for each beam pipe, allowing for the counter-rotating positively charged beams. Quadrupole magnets are used to squeeze and focus[2] the beams to provide more proton-proton collisions at each of the collision points. Even higher evenly dimensional pole magnets (multipoles) are used to correct beam orbit distortions.

A typical proton bunch contains approximately 10^{11} protons. During *physics* runs the LHC delivers the proton bunches in *trains*, with a bunch crossing rate of 40 MHz. The amount of proton-proton collision data that is delivered by the LHC and recorded with the ATLAS detector is quantified by a quantity called luminosity, $\mathcal{L}$, characterising the instantaneous rate of proton collisions. The total amount of proton-proton data recorded is measured in integrated luminosity, $\int \mathcal{L}dt$, typically in units of inverse femtobarns, fb^{-1}. The total cumulative data delivered by the LHC and recorded by ATLAS is shown in Figure 4.2 for 2015 and 2016. The definition, measurement and calibration of luminosity is discussed in detail in Chapter 6.

There are three run periods foreseen for the current accelerator: Run-1 spanning 2010 to 2012 at $\sqrt{s} = 7$ TeV (2010 and 2011) and $\sqrt{s} = 8$ TeV (2012), Run-2 spanning 2015 to 2018 at $\sqrt{s} = 13$ TeV and Run-3 spanning 2021 to 2023 at $\sqrt{s} = 14$ TeV. This thesis will solely focus on data collected in 2015 and 2016 at $\sqrt{s} = 13$ TeV.

[2] A single quadrupole magnet would focus the beams in one plane and defocus in the other. However if two quadrupoles are employed with their focusing directions orthogonal to one another, a net focusing can be achieved.



(a) (b)

Figure 4.2: The cumulative integrated luminosity, $\int \mathcal{L}dt$, delivered by the LHC (green) and recorded by the ATLAS detector (yellow) for the (a) 2015 and (b) 2016 data taking periods. The difference in value respects the inefficiency of data acquisition by the ATLAS detector and is at the level of 7 to 8% for both years.

4.2 The ATLAS detector

ATLAS is a 7000 tonne, approximately cylindrical detector with a length of 44 m and diameter 25 m. A cut-away view of the detector is shown in Figure 4.3, with the main detector and system components labelled. ATLAS comprises of three main detector systems.

Nearest to the beam pipe is the inner detector, used for reconstructing tracks of charged particles in the high radiation region near the interaction point. The inner detector consists of three main sub-systems: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). Each subdetector has fine granularity to accurately track the position of charged particles as they pass through the inner detector. The inner detector is surrounded by a solenoid magnet, designed to produce an axial magnetic field to bend the charged particles and allow for momentum measurement. The inner detector and the solenoid magnet system is further discussed in Section 4.2.2.

Encompassing the inner detector are the electromagnetic (EM) and hadronic calorimeter systems, designed to measure the energies of incident particles by total absorption. The EM calorimeter is designed to accurately measure the deposited energy of electrons and photons and the hadronic calorimeter is designed to measure the deposited energies of hadronic jets. The calorimeters are described in detail in Section 4.2.3.

The final detector system, furthest from the beam-pipe, is the muon spectrometer (MS). A toroidal magnet system intertwines with the MS, allowing for precise measurements of muon momentum. The MS and the toroidal magnet system are discussed further in 4.2.4.

4.2.1 Coordinate system and nomenclature

A right-handed coordinate system is employed by ATLAS [35]. The z-axis is defined as parallel to the proton beam line that travels longitudinally through the cylindrical detector. Proton beam 1 travels clockwise around the LHC ring (passing from positive to negative z through ATLAS) and beam 2 counter-clockwise (negative to positive z). ATLAS is symmetrical in the plane transverse to the beam and hence polar coordinates are used. The positive x-direction is defined as pointing towards the middle of the LHC ring and the azimuthal angle, ϕ, is defined

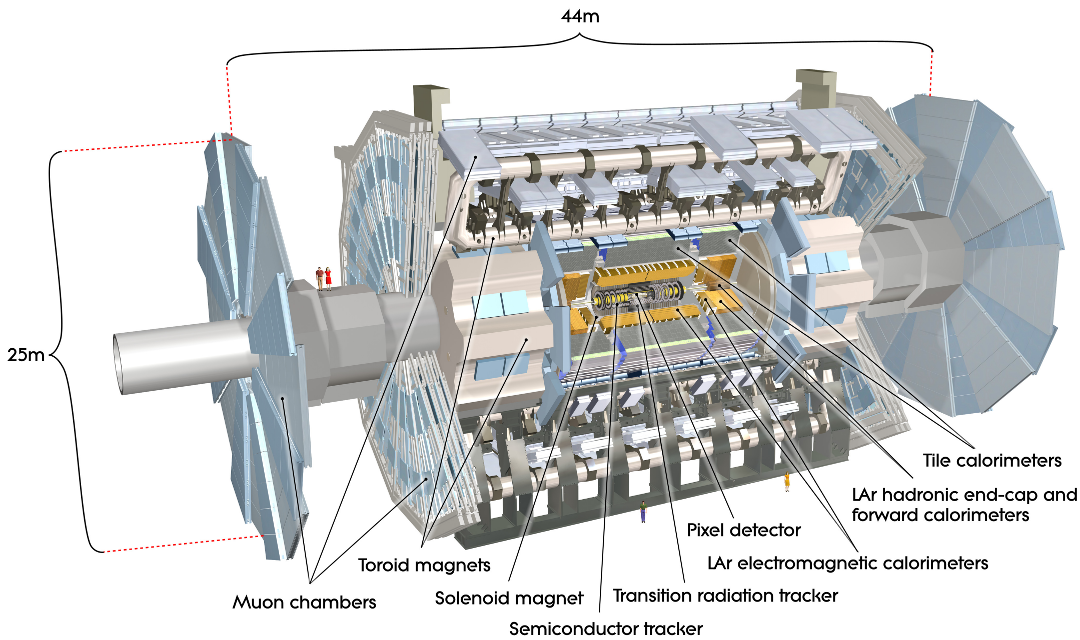


Figure 4.3: The ATLAS detector [35].

as the angle measured from positive-x in the transverse plane. The polar angle, θ, is defined as the angle measured from positive-z.

Variables are typically measured in the transverse plane in ATLAS, due to hard scatter collisions producing particles with large transverse momentum, p_T, and energy, E_T. The transverse momentum of a particle is related to the total momentum, p, by

$$p_T = p\sin\theta, \tag{4.1}$$

and transverse energy is defined as

$$E_T = \sqrt{p_T^2 + m^2}, \tag{4.2}$$

where m is the mass of the particle. A missing transverse momentum, $\mathbf{p}_T^{\text{miss}}$, is also defined that typifies the imbalance of p_T measured in an event. It is given as the negative vector sum of the transverse momentum of all reconstructed particles

$$\mathbf{p}_T^{\text{miss}} = -\sum \mathbf{p}_T^{\text{reco}}. \tag{4.3}$$

The missing transverse energy, E_T^{miss}, can subsequently be defined as the absolute value of missing transverse momentum

$$E_T^{\text{miss}} = |\mathbf{p}_T^{\text{miss}}|. \tag{4.4}$$

Signatures of E_T^{miss} in an event can be used to indirectly measure the momentum of particles that do not interact in the detector, such as neutrinos. Large values of E_T^{miss} could also be an indication of new physics and thus such signatures are used in BSM search analyses at ATLAS.

Another useful property is the invariant mass of a system, m, which can be defined from the total energy and momentum of that system:

$$m = \sqrt{(\sum E)^2 - (\sum \mathbf{p})^2}. \tag{4.5}$$

The invariant mass is determined from quantities which are conserved during a decay and is thus Lorentz invariant. If one calculates the invariant mass using the energy and momentum of the decay products of a single particle then the invariant mass of that system is equal to the mass of the particle that decayed. It is thus a very useful property to use in the search for new particle resonances.

Typically the systems of particles produced from colliding partons with a difference in momentum are boosted longitudinally. The difference in polar angle, $\Delta\theta$, between two particles is not Lorentz invariant under such boosts. This problem can be solved by defining the rapidity, y,

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{4.6}$$

with E and p_z the total energy and longitudinal momentum of a particle. Rapidity has the desired property that the rapidity difference, Δy, of two particles is Lorentz invariant for boosts along the longitudinal z-axis. Taking the massless limit results in the definition of pseudorapidity, η:

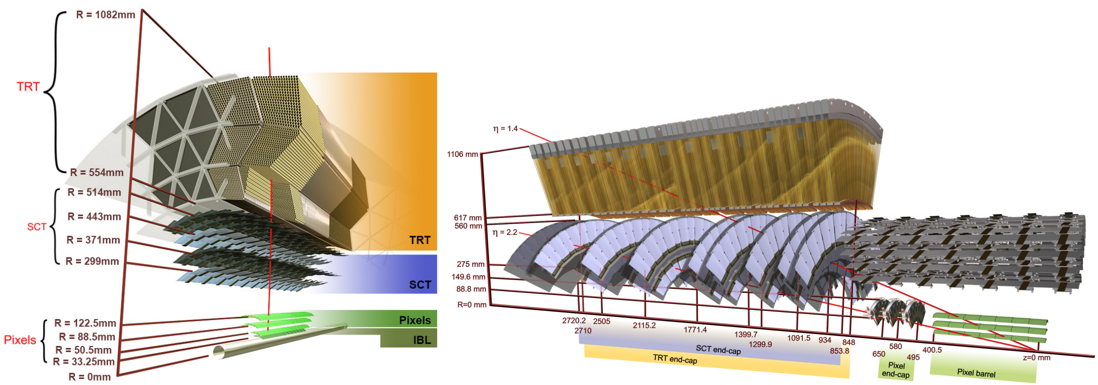
$$\eta = -\ln(\tan\frac{\theta}{2}). \tag{4.7}$$

This massless limit is a good approximation due to the high energy particles produced by the LHC collisions, such that $E \gg m$. This Lorentz invariant property of $\Delta\eta$ leads to η being the preferred choice for polar coordinate over θ. Now the azimuthal and polar coordinates are defined, one can also define ΔR, a radius of a cone in η-ϕ space,

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{4.8}$$

4.2.2 The inner detector

The inner detector [49] covers the pseudorapidity range $|\eta| < 2.5$, with complete coverage in azimuth. Each subsection of the ID is composed of a barrel region, arranged in cylinders around the interaction point at $z = 0$, and two endcaps, arranged in disks perpendicular to the beam, either side of the barrel in the forward regions [35]. Schematic diagrams of the barrel and endcap ID are shown in Figures 4.4a and 4.4b respectively.



(a) (b)

Figure 4.4: The ATLAS inner detector in (a) the barrel region [50] and (b) the endcap region [35]. The transverse radius R_ϕ and longitudinal position z for each sub-component is shown. The IBL is not shown in (b).

The inner detector is itself surrounded by a solenoid magnet [51], designed to produce an axial magnetic field of 2 T. This magnetic field bends the tracks of the charged particles, with the radius of curvature allowing for the measurement of particle momentum.

4.2.2.1 The pixel detector

The pixel detector [52] consists of layers of silicon pixel modules; four parallel to the beam-pipe in the barrel region ($|\eta| < 1.7$) and three transverse in each endcap ($1.7 < |\eta| < 2.5$). Doped silicon is a semiconductor; when charged particles pass through the material *electron-hole* pairs are produced. A bias voltage is applied across the detector, causing the charge to drift to a readout where it is measured. If the charge collected reaches a pre-set threshold then a particle hit is recorded.

In Run-1, three barrel layers were present in the pixel detector; the B-layer, layer-1 and layer-2 at transverse radii of $R_\phi = 50.5$ mm, $R_\phi = 88.5$ mm and $R_\phi = 122.5$ mm respectively. Each pixel module has a typical pixel size $(R_\phi, z) = (50, 400)$ μm, providing a spatial resolution of around 10 μm in the transverse plane and 115 μm in the longitudinal (radial) planes for the barrel (endcaps).

During the long shutdown between Run-1 and Run-2 (LS1) an additional pixel layer, named the insertable B-layer (IBL), was inserted close to the beam-pipe at $R_\phi = 33.2$ mm [50]. The IBL consists of a mix of planar and 3D silicon sensors with pixel size $(50, 250)$ μm. The inclusion of this new layer led to an improvement in tracking and secondary vertex resolution with respect to Run-1 [53].

4.2.2.2 The semiconductor tracker

The SCT surrounds the pixel detector and consists of four layers of single sided silicon microstrip detectors in the barrel ($|\eta| < 1.4$) and nine layers in the endcap ($1.4 < |\eta| < 2.5$). The SCT modules in the barrel consist of four silicon microstrip detectors, with two detectors each glued back to back with a small 40 mrad stereo angle separating them. This allows for spatial resolution in the plane parallel to the beam, as well as transverse [54]. In the endcaps the modules are constructed the same way as in the barrel, but have variable size dependent on the longitudinal position. The SCT provides 17 μm spatial resolution in the transverse plane and a 580 μm in the longitudinal (radial) directions for the barrel (endcaps).

4.2.2.3 The transition radiation tracker

The TRT is a straw-tracker detector consisting of 2 mm radius tubes containing a mixture of gases: 70% Xe, 27% CO_2 and 3% O_2 [35]. Incident charged particles are identified by the ionisation of the gas and, similarly to the silicon detectors, charge is measured at a readout in the centre of the straw by applying a bias voltage. Straws in the barrel ($|\eta| < 0.7$) are aligned parallel to the beam, whilst in the endcaps ($0.7 < |\eta| < 2.0$) the straws are arranged radially in wheels, with eighteen wheels per endcap. The TRT provides a transverse spatial resolution of 30 μm.

A polypropylene radiator is placed between the straws; when a relativistic charged particle passes the boundary between the polypropylene and the straw, it radiates photons. This process is known as transition radiation and hence the subdetector name. The photons are absorbed by and ionise the Xe gas, causing a larger total ionisation rate than purely from a single incident particle. This mechanism is used to distinguish tracks from electrons and pions (the most common charged hadron), since only electrons with energy up to approximately 100 GeV produce enough transition radiation for the gas plus transition radiation current threshold to be reached. At energies greater than 100 GeV, the gas plus transition radiation current for pions also reaches this defined threshold and the two particle types can no longer be separated [55].

4.2.3 The calorimeter systems

Sampling calorimeters are composed of two types of material: the first a material that causes the incident particle to shower into a large number of secondary particles and the second a material to absorb and measure the deposited energy of the showered particles. This is a multi-stage process with many alternating layers of material to ensure full absorption of the showered energy. An illustration of the ATLAS calorimeter systems [56] is shown in Figure 4.5.

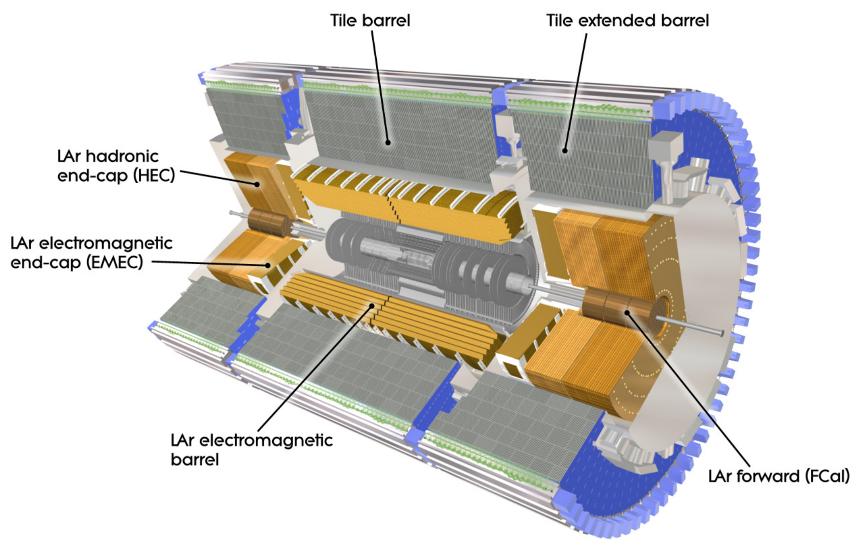


Figure 4.5: The ATLAS EM and hadronic calorimeter systems [35].

4.2.3.1 The electromagnetic calorimeter

The EM calorimeter [57] is composed of lead absorbers in liquid argon (LAr). When incident upon the lead, electrons radiate photons (bremsstrahlung) which subsequently convert into electron-positron pairs. The resulting electrons (positrons) undergo the same process. This recursive showering ensures that the energy of incident electrons and photons are maximally deposited in the calorimeter. The electrons and positrons in the shower ionise the LAr and the corresponding current is collected at readouts. A radiation length is defined, X_0, that denotes the distance travelled by the electron in which its energy is reduced via bremsstrahlung by a factor e.

The calorimeter systems, like the inner detector, are comprised of a barrel and endcap regions. The EM calorimeter has a barrel spanning $|\eta| < 1.475$ and two endcaps in the region, $1.375 < |\eta| < 3.2$. In the pseudorapidity range, $1.37 < |\eta| < 1.52$, the measured energy resolution is poorer than other regions of phase space due to the barrel-endcap transition. A LAr forward calorimeter (FCal) [58], also used to measure hadrons, resides in the region $3.1 < |\eta| < 4.9$. FCal is further discussed in Section 4.2.3.2.

A single module in the EM barrel is shown in Figure 4.6. The module consists of three layers, positioned in an accordion geometry. The first layer consists of strip cells, with granularity ($\Delta\eta$, $\Delta\phi$) = (0.0031, 0.0982). The fine segmentation in η of the first layer allows for measurements of close-by electromagnetic objects, especially useful in the measurement of boosted $\pi^0 \rightarrow \gamma\gamma$ decays [3]. The second layer consists of square cells with coarser transverse granularity than the first, ($\Delta\eta$, $\Delta\phi$) = (0.025, 0.0245) and the third layer coarser granularity again in η, ($\Delta\eta$, $\Delta\phi$) = (0.05, 0.0245). The large depth of the second layer (16 X_0) allows for the measurement of the majority of the energy of incident particles and the third layer provides a measurement to correct the energy overlap in the subsequent hadronic calorimeter away from the interaction point.

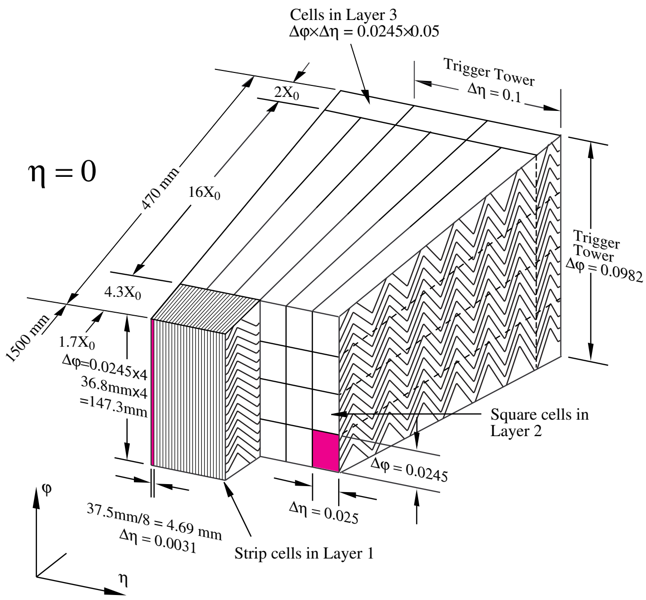


Figure 4.6: A sketch of a barrel module of the electromagnetic calorimeter [35]. X_0 refers to the radiation length.

4.2.3.2 The hadronic calorimeter

The hadronic calorimeter [59], also known as the tile calorimeter (TileCal), contains scintillating tiles as the active material with steel absorbers. Instead of X_0, a similar interaction length, λ, is defined for hadronic interactions. Each scintillating tile is connected to a photomultiplier tube (PMT) that measures the incoming signal.

The tile calorimeter extends to $|\eta| < 1.7$, with the tile barrel in the region $|\eta| < 1.0$ and the two extended barrels in $0.8 < |\eta| < 1.7$, split azimuthally into 64 wedges. The tile calorimeter contains three layers, each with different interaction lengths. The total length of all three layers is approximately 7λ in both the barrel and endcaps, depending on η. The granularity of the tiles are ($\Delta\eta$, $\Delta\phi$) = (0.1, 0.1) in the first two layers and ($\Delta\eta$, $\Delta\phi$) = (0.2, 0.1) in the third.

Forward measurements are made by hadronic end-caps (HEC) in the region $1.5 < |\eta| < 3.5$, utilising copper absorbers in LAr. The HEC are placed behind the EM calorimeter endcaps and, like the EM endcaps, η dependent granularity; for $1.5 < |\eta| < 2.5$, ($\Delta\eta$, $\Delta\phi$) = (0.1, 0.1) and for $2.5 < |\eta| < 3.5$, ($\Delta\eta$, $\Delta\phi$) = (0.2, 0.2).

The hadronic calorimeter finally ends with the forward calorimeter, which also contains three layers of modules. The first layer utilises copper, as with the HECs, which is used to measure electromagnetic showers. The second and third layers contain tungsten absorbers, for measuring hadronic interactions.

4.2.4 The muon spectrometer

The muon spectrometer [60] measures the momenta of charged particles that penetrate the calorimeter systems. In general, the only charged particles that make it so far are muons due to their minimum ionising nature.

The muon spectrometer is composed of four types of subdetector: the monitored drift tubes (MDTs) and the cathode strip chambers (CSCs) provide spatial measurements up to $|\eta| < 2.7$, with the resistive plate chambers (RPCs) and the thin gap chambers (TGCs) providing tracking information for use in the trigger system, to $|\eta| < 2.4$.

Three large superconducting air-core toroidal magnets [61] intertwine the muon spectrometer; one in the barrel region and two endcaps. The measurement of muon momentum is based on the magnetic deflection in this system, designed to bend the muons in the η direction. The barrel toroid [62] consists of eight coils, providing a non-uniform magnetic field of roughly 0.5 T in the region $|\eta| < 1.4$. The endcap toroids [63] also consist of eight coils, providing a magnetic field of approximately 1.0 T in the region $1.6 < |\eta| < 2.7$. In the intermediate region the magnetic field is provided by both the barrel and the endcap toroids and is thus reduced.

The muon spectrometer and toroid systems are shown in Figure 4.7.

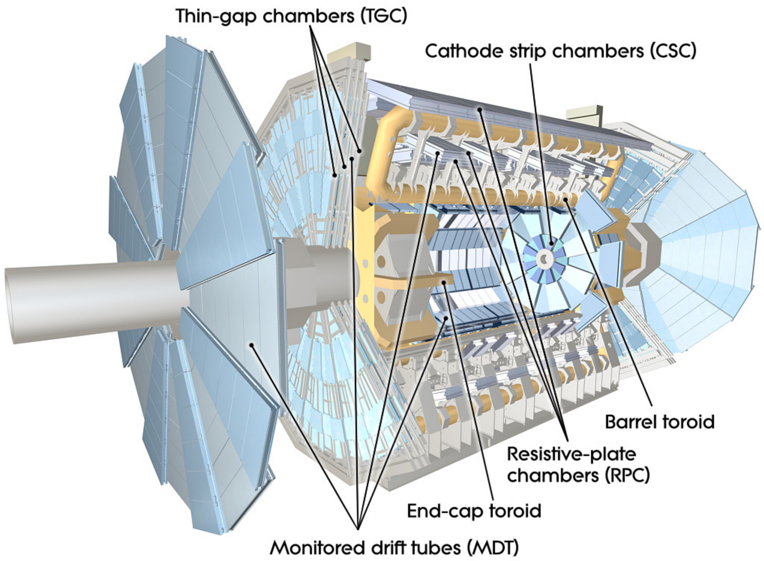


Figure 4.7: The ATLAS muon spectrometer system [35].

4.2.4.1 Precision tracking

Measurements of muon tracking are provided by the MDT chambers. An MDT chamber consists of three to eight layers of pressurised drift tubes, depending on the position of the chamber. Muons passing through the chamber ionise the mixture of CO_2 and Ar, with the resultant charge collected at the centre of the tube. Each chamber results in a spatial resolution of approximately 35 μm in the bending plane. There are three layers of MDT chambers; the second and third layers, furthest away from the beam pipe, have range $0 < |\eta| < 2.7$, spanning both barrel and endcap. The first layer is in the range $0 < |\eta| < 2.0$.

The track density is highest in the forward region, $2.0 < |\eta| < 2.7$. In this region, the first layer of MDTs are replaced by the CSCs. The CSCs are multi-wire proportional chambers. Two sets of cathode strips are aligned parallel and orthogonal to the wires in the chambers, allowing for reading of both η and ϕ coordinates of the incoming muon due to charge induced on the wires. The CSCs have a higher rate capability than the MDTs, with a time resolution of 7 ns. This time resolution allows for an additional measurement of the spatial resolution in the transverse plane from charge drift times, providing resolutions of approximately 40 μm in the η direction and 5 mm in ϕ.

4.2.4.2 Trigger chambers

The RPCs and TGCs used for triggering are found in the $|\eta| < 1.05$ and $1.05 < |\eta| < 2.4$ regions respectively. The total time for a muon to provide a hit in a chamber and the resulting signal to drift and be passed to the trigger system is typically only 15 to 25 ns after passage. This fast triggering allows for individual bunch crossings to be identified.

The RPCs consist of two parallel plates, with a uniform electric field and gas in between. Incident muons ionise the gas and a signal current is read out. The TGCs are multi-wire proportional chambers, like the CSCs, and work in a similar manner. The trigger chambers provide measurements of the muon track in both the bending (η) and non-bending (ϕ) planes; the RPCs provide a spatial resolution of 10 mm in both planes and the TGCs 2 to 7 mm in η and 3 to 7 mm in ϕ.

4.2.5 Trigger and data acquisition system

The LHC delivers bunches of protons at a rate of 40 MHz (25 ns bunch crossing). The number of protons in each bunch is large, resulting in a rate of inelastic collisions of approximately 1 GHz. The majority of these collisions do not have the "interesting" features of high-energy physics, such as high p_T leptons or photons to name two. To this end a trigger system is employed, where only events with specific features are saved.

The ATLAS trigger system [64] is a combination of hardware and software and consists of two levels. The Level-1 (L1) trigger is a hardware trigger using information from the calorimeters and the muon spectrometer. A trigger menu is run on each bunch crossing defined by asking if specific event signatures are seen. The L1 trigger reduces the event rate to approximately 85 kHz and hands over to the second trigger tier.

The second level is the software-based high-level trigger (HLT). The event signatures that triggered L1 are modified into regions of interest (RoIs). This segments the detector into (η, ϕ) and employs fast but detailed reconstruction algorithms to determine whether the L1-triggered event should be retained. This reduces the rate to roughly 1 kHz; if the event passes the trigger it is saved to disk for use in offline analysis.

Chapter **5**

Particle reconstruction and identification

Final state particles produced in proton-proton collisions leave traces of their passage in the various subdetectors of ATLAS. The reconstruction and identification of such particles are required to isolate interesting event topologies.

This chapter describes the reconstruction and identification of tracks and vertices (Section 5.1), electrons (Section 5.2), muons (Section 5.3), light lepton isolation (Section 5.4), jets (Section 5.5) and hadronic taus (Section 5.6) that are used to define an event. Other final state particles and properties such as photon reconstruction/identification and missing transverse energy are omitted due to a lack of use of such objects in this thesis.

5.1 Tracks and vertices

Inner detector tracks are reconstructed by combining information from the pixel, SCT and TRT subdetectors with a sequence of algorithms. The *inside-out* algorithm [65] seeds tracks from three initial silicon hits in the pixel and SCT subdetectors. Further silicon hits are added to the track away from the interaction point with a Kalman filter [66], and the tracks are then extended into the TRT.

The inside-out algorithm is used to reconstruct primary tracks[1]. Due to the 2 T strength of the solenoid magnet, only tracks with $p_T > 400$ MeV are reconstructed with the inside-out algorithm; tracks with p_T less than this do not fully traverse the inner detector. The second stage in track reconstruction involves *back-tracking*: segments from the TRT are extended back into the silicon detectors with hits being consecutively added. This method is used to reconstruct secondary tracks; tracks produced by the interactions of the primaries. For example, the tracks from photons that convert in the inner detector may be typically reconstructed with the back-tracking technique due to the absence of inner pixel hits.

After tracks are reconstructed, vertex candidates can then be defined. A vertex is the intersection of two or more inner detector tracks. The *primary* vertex in an event is that with the largest $\sum p_T^2$ of tracks associated with the vertex[2]. Tracks are associated to vertices in a two stage process: track association to vertex candidates, called *vertex finding*, and reconstruction of vertex position, called *vertex fitting* [67]. The algorithm is iterative: vertex candidates are reconstructed with a seed position, the best position is determined with a fit, incompatible tracks are removed and the procedure is repeated.

The impact parameter of the track is an important quantity, especially in the reconstruction of leptons. The transverse impact parameter of a track, d_0, is the closest transverse distance of the track to the primary vertex and the longitudinal impact parameter, z_0, is defined as the longitudinal distance between the point on the track at which d_0 is determined and the primary

[1] Primary in this instance is defined as particles with lifetime greater than 3×10^{-11} s directly produced in a proton-proton interaction or from decays or interactions of particles with lifetime less than 3×10^{-11} s [65].

[2] Confusingly all the reconstructed vertices in a bunch crossing can also be called primary vertices, to distinguish them from secondary vertices that originate from the decays of the particles produced in the proton-proton collisions.

vertex. Typically $z_0 \sin\theta$ is quoted as the longitudinal impact parameter. A diagram illustrating the definition of the transverse impact parameter is shown in Figure 5.1.

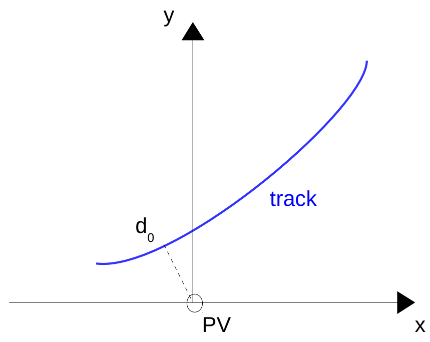


Figure 5.1: A schematic diagram defining the transverse impact parameter of a track, d_0, with respect to the primary vertex (PV) of an event in the transverse plane. The longitudinal impact parameter is the longitudinal distance between the point on the track where d_0 is calculated and the primary vertex.

Another important quantity is the transverse impact parameter significance, d_0/σ_{d_0}, with uncertainty on d_0, σ_{d_0}. These quantities can be used for distinguishing prompt and non-prompt particles; those that originate from the primary vertex and those that originate from secondary (and tertiary) vertices respectively. z_0 requirements can also be used to reject particles associated with pile-up collisions.

5.2 Electrons

The electron reconstruction (Section 5.2.1), identification (Section 5.2.2), energy calibration (Section 5.2.3) and triggering (Section 5.2.4) are discussed below. Only electrons reconstructed with $|\eta| < 2.47$ are considered in this thesis.

5.2.1 Electron reconstruction

The reconstruction of an electron candidate starts from a deposit of energy in the EM calorimeter, called a cluster. The EM calorimeter is firstly divided into cells of size $(\Delta\eta, \Delta\phi) = (0.025, 0.025)$ [68]. Energy deposits (in the transverse plane) of greater than 2.5 GeV are searched for in *windows* of size (3, 5), in units of cell size. If this criterion is met, then the cluster is matched to any inner detector tracks that point at it. A track is matched to the cluster if the impact of the track with respect to the cluster centre is within $|\Delta\eta| < 0.05$ and $|\Delta\phi| < 0.1$. If the cluster is on the opposite side to the bend of the track then the $\Delta\phi$ selection is tightened to $|\Delta\phi| < 0.05$. This asymmetric transverse acceptance is due to electron electromagnetic radiation loss, called bremsstrahlung. Only one track is matched to the electron: if there are more than one candidate track then the track containing the best mixture of silicon hits and closest ΔR to the cluster is the one chosen. If no track is found that matches with the cluster then the electron candidate is discarded and is then considered to be a photon candidate. A schematic diagram of the detectors used in the electron reconstruction chain is shown in Figure 5.2.

The efficiency of the track and cluster matching reconstructions for electron candidates are measured in $Z \rightarrow ee$ events with the tag and probe method. The tag and probe method is used to determine selection efficiencies for prompt leptons and is discussed in detail in Chapter 7.

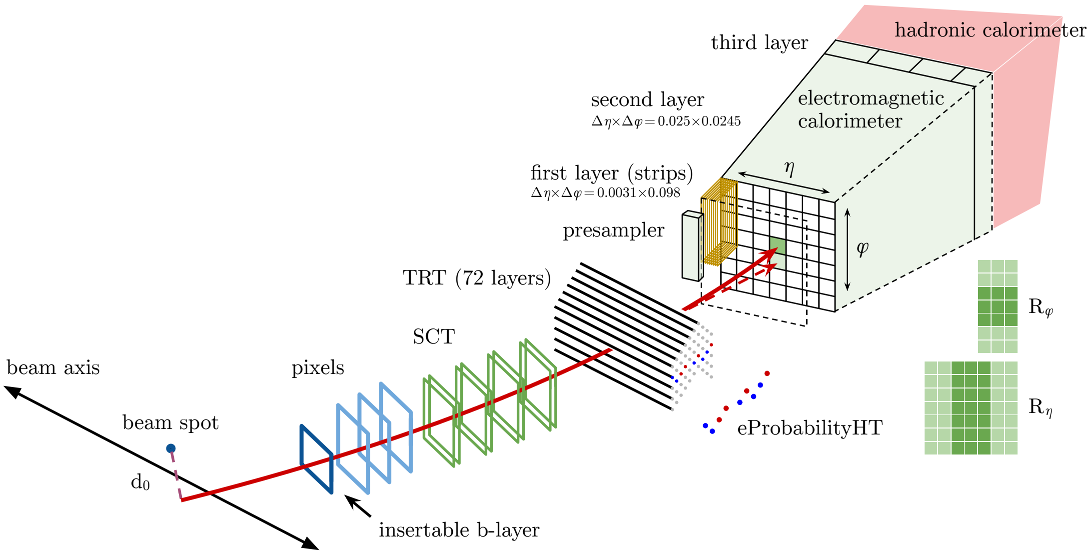


Figure 5.2: A schematic diagram of the reconstruction of electron candidates. R_ϕ (R_η) refers to the ratio of the energy in 3×3 (3×7) cells over the energy in 3×7 (7×7) cells centred at the electron cluster position and eProbabilityHT refers to the likelihood probability based on transition radiation in the TRT. These are used as discriminating variables in the electron likelihood identification requirements (see Section 5.2.2) [68].

The reconstruction efficiency is roughly 99% for central and/or high E_T electrons and 97% for forward and/or low E_T electrons.

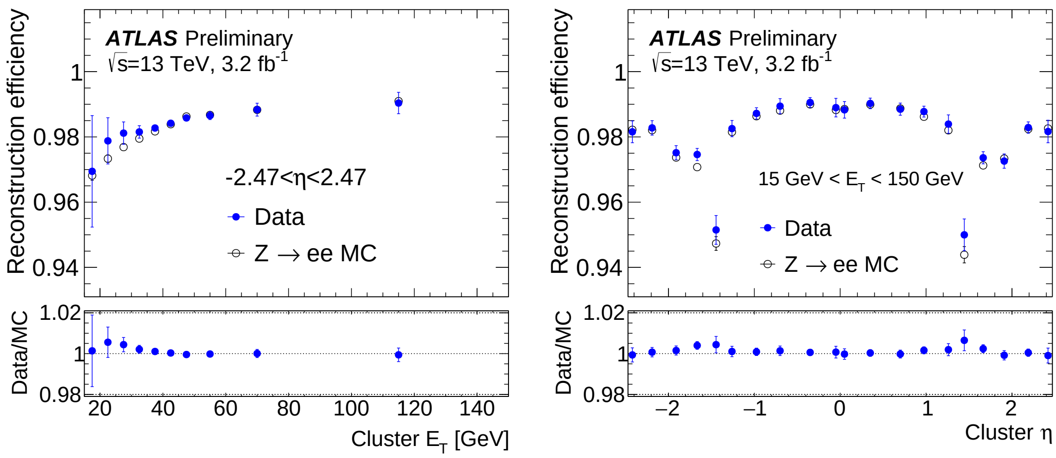


Figure 5.3: The measured reconstruction efficiencies as a function of (a) E_T and (b) η for $15 < E_T < 150$ GeV for the 2015 dataset. The uncertainties are statistical plus systematic [68].

5.2.2 Electron identification

Once electron candidates are reconstructed, the *true* electrons need to be identified. A large proportion of the electron candidates are actually jets and thus multivariate techniques are used to reject such wrongly reconstructed electrons. Both a cut-based and likelihood discriminant are used to identify electrons at ATLAS. Only electrons chosen using likelihood (LH) working points are used in this thesis.

Three working points are defined for the likelihood method: `LooseLH`, `MediumLH` and `TightLH`. A number of discriminating variables are chosen, defining electron-like track parameters, shower shape and track-cluster matching. The probability density functions (PDFs) of variable, i, for signal, $P_{S,i}$, and background, $P_{B,i}$, are used to build likelihoods and a discriminant, $d_{\mathcal{L}}$,

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \qquad \mathcal{L}_S(\mathbf{x}) = \prod_{i=1}^{N} P_{S,i}(x_i), \qquad \mathcal{L}_B(\mathbf{x}) = \prod_{i=1}^{N} P_{B,i}(x_i), \tag{5.1}$$

where $\mathcal{L}_{S,B}$ are the respective likelihoods for signal and background as a function of the variable set, $\mathbf{x}$. This discriminant allows for calculating a probability of the candidate being a real electron (or not). For each working point the same variables are used to compute the LH discriminant, but the working points are constructed so that `MediumLH` is a subset of `LooseLH`, and `TightLH` is a subset of `MediumLH`. For `LooseLH` electrons the signal identification efficiency for candidates with $E_T > 20$ GeV are 92% to 97%, increasing with E_T, and the background identification efficiencies are 0.8% to 0.3% decreasing with E_T. For `TightLH` the signal efficiencies are 78% to 91%, and the background efficiencies 0.35% to 0.1% [68]. The efficiencies are shown as a function of the number of reconstructed vertices in Figure 5.4.

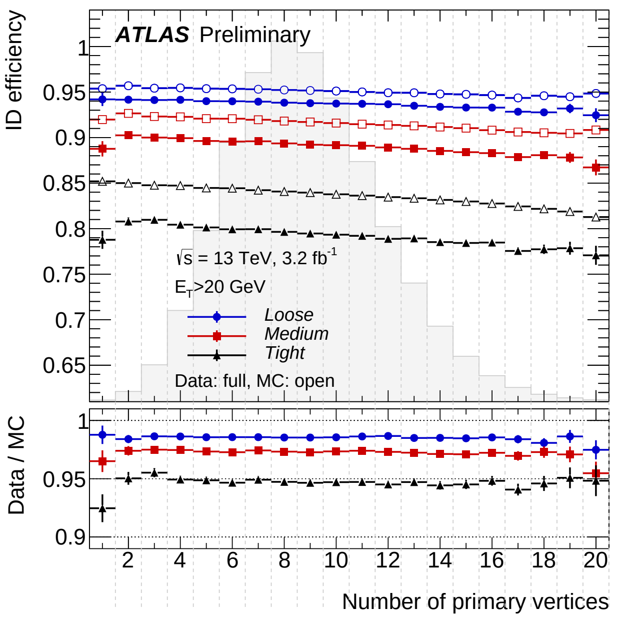


Figure 5.4: Efficiencies for the three different LH working points as a function of the number of reconstructed vertices. The distribution of the number of reconstructed vertices in the selected data events is shown in grey [68].

5.2.3 Electron energy calibration

The energy of an electron candidate is built from energies of clusters of cell size (3, 7) and (5, 5) for the EM barrel and endcaps respectively, which are calibrated using multivariate techniques based on simulation. The electron energy is corrected for estimated energies deposited in the dead material in front of the EM calorimeter, after the EM calorimeter and outside the cluster in the EM calorimeter. The material passage and intercalibration of the different LAr calorimeter layers are also used in the calibration procedure [69].

Any further electron energy miscalibrations are determined by comparing the dielectron invariant mass for $Z \to ee$ events in data and MC. An energy scale, α, is defined to correct the energy in MC to that of data

$$E_{\text{data}} = E_{\text{MC}}(1+\alpha). \tag{5.2}$$

This scale results in increasing or decreasing the invariant mass of the two electrons.

The electron energy resolution, σ, is parameterised by

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{5.3}$$

where parameters a, b and c are the sampling, noise and constant terms respectively and $\oplus$ denotes a sum in quadrature. A correction to the energy resolution is also derived by assuming that the resolution in MC models data well up to the constant term c. Both c and α are determined in bins of (η_i, η_j) of the two electrons with both corrections and uncertainty on the corrections below 0.1%. The electron energy scale uncertainty is determined to be 0.05% to 0.2% for $E_T \approx 40$ GeV and 0.4% to 1.1% for $E_T \approx 10$ GeV. The difference in uncertainty for each value of E_T depends on the region in the detector; specifically how much material the electron passes through.

5.2.4 Electron trigger

All triggers aiming to save events with electrons use L1 calorimeter information as a seed. Signals are recorded in *towers* of the EM and hadronic calorimeters of size $(\Delta\eta, \Delta\phi) = (0.4, 0.4)$. E_T thresholds can be set as a function of η to take into account different energy responses in different regions of the calorimeters.

Fast tracking information is included in the HLT step to reconstruct electron candidates online. Tracks with a minimum p_T of 1 GeV are required to match clusters within $\Delta R < 0.2$. These early electron candidates are then subject to precise algorithms designed to be similar (but faster) than the offline algorithms. This electron reconstruction results in early background rejections allowing for the event rate to be reduced to a reasonable level. The event rate is reduced further by requiring the electrons pass one of the identification working points discussed above. These working points are designed to be slightly looser than the the corresponding offline working points to ensure minimal efficiency losses.

5.3 Muons

The muon reconstruction (Section 5.3.1), identification (Section 5.3.2), momentum calibration (Section 5.3.3) and trigger (Section 5.3.4) are discussed below.

5.3.1 Muon reconstruction

Muon reconstruction combines information from tracks in the inner detector and the muon spectrometer to form a muon candidate [70].

Tracks in the muon spectrometer are started by searching for hit patterns in each chamber, which are then formed into sequential segments. A Hough transform [71] is used to search for hits in the curvature plane of the toroidal magnetic field in the MDT, where a straight line fit is used to produce segments. The transform is robust against the misidentification of hadrons in the MS. In the CSC segments, the MS track is formed by requiring a loose association to the luminous region.

The track candidates are then evolved by combining segments of different layers of the MS. A combinatorial search starts by using middle layer segments as seeds and then the search is extended to the inner and outer layers of the MS, matching segments by their relative positions and angles. At least two segments are needed to reconstruct an MS track in the barrel, except for the transition region between the barrel and endcap where only one single quality segment is needed. An iterative χ^2 fitting procedure is employed to assign hits to tracks.

There are four different types of reconstructed muon candidates used in ATLAS: *combined*, *segment-tagged*, *calorimeter-tagged* and *extrapolated* muons. The most precise reconstruction of muon candidates are the *combined* muons, containing tracks in the inner detector and the muon spectrometer that are matched in a global χ^2 fit procedure. The fitting procedure allows for modification of the number of hits in the MS track if the fit is improved. *Segment-tagged* muons are candidates with an inner detector track that match only a single segment in the MS. *Calorimeter-tagged* muons are candidates with an inner detector track that is matched to a "muon-like" calorimeter energy deposit. Calorimeter-tagged muons are common in the $|\eta| < 0.1$ region of the ATLAS detector where equipment to power the inner detector and the calorimeters block the MS. Finally, *extrapolated* muons are candidates with an MS track that is only loosely positioned relative to the luminous region, with no matched inner detector track. Extrapolated muons are used for the forward acceptance region of the MS beyond the inner detector, $2.5 < |\eta| < 2.7$.

5.3.2 Muon identification

Muons are the most cleanly identified particles in the ATLAS detector. There is only a very small background contribution from punch-through hadrons that manage to fully traverse the calorimeters. However, the background definition of muons at ATLAS is extended to include "true" muons that are produced in the decays of light flavour hadrons, such as pions and kaons. The identification is based on the rejection of these decays. It should be noted that muons produced in heavy flavour decays (from b and c quarks) are not included in this background definition.

A muon that originates from a charged hadron in-flight typically decays in the inner detector and the reconstructed track will have a *kink* detailing the position of the decay vertex. The momentum matching of the inner detector tracks with the muon spectrometer track is then likely to be poor. Four muon identification working points are provided: `Loose`, `Medium`, `Tight` and `HighPt`. The `Loose` criterion is designed to maximise reconstruction efficiency for good-quality muon tracks. All the different muon types are used, with roughly 97.5% being combined muons. The combined muons are required to have more than three hits in the MDT layers except for $|\eta| < 0.1$ where only one layer and no more than one hole in the MDT is allowed. The ID and MS track momentum are required to be compatible to reject the aforementioned charged light hadron decays. The definitions for the other identification working points can be found in [70]. Only `Loose` muons are used in this thesis. The signal efficiency for `Loose` muons with $20 < p_T < 100$ GeV is 98.1% with background (muons from light hadrons) efficiency of 0.76%. The efficiency as a function of η is shown for $Z \to \mu\mu$ decays in Figure 5.5. The `Medium` selection is the same as the `Loose` selection for $|\eta| > 0.1$.

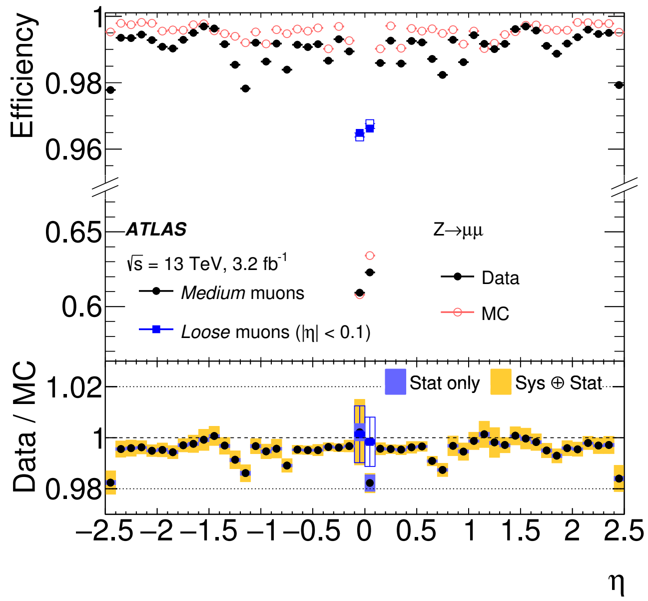


Figure 5.5: The muon reconstruction efficiency as a function of η for the `Medium` identification working point in $Z \to \mu\mu$ events in MC and 2015 data for muons with $p_T > 10$ GeV. The efficiency of the `Loose` selection in the region $|\eta| < 0.1$ is shown where the `Loose` and `Medium` selections differ [70].

5.3.3 Muon momentum calibration

The muon momentum scale is determined from the muon track. Only combined muons are used to determine the correction. The momenta of the tracks from the inner detector and muon spectrometer are calibrated separately and the corrected momentum to a (combined) muon is a weighted average of each. The corrected transverse momentum of each track (either in the inner detector or muon spectrometer) takes the nominal value of transverse momentum and modifies it by momentum scale and momentum resolution corrections.

Differences in momentum scale between data and simulation are typically due to mismodelling of the magnetic field and energy loss of the muon in the calorimeter and other material before passing through to the muon spectrometer. The muon momentum resolution is parameterised as

$$\frac{\sigma}{p_T} = \frac{r_0}{p_T} \oplus r_1 \oplus r_2 p_T. \tag{5.4}$$

The r_0 term mainly accounts for energy loss in traversed material, r_1 for multiple scattering and magnetic field homogeneities and r_2 for detector misalignment effects [70]. The momentum scale and resolution corrections are determined by comparing the invariant mass peaks of the J/ψ and Z resonances in data and simulation. Momentum scale and resolution corrections are typically below 0.1% and 1% respectively. The effects of the fitted corrections are shown in Figure 5.6 for $Z \to \mu\mu$ events in 2015 data.

The uncertainty in the momentum scale varies from 0.05% to 0.3% for $|\eta| \approx 0.1$ and $|\eta| \approx 2.5$ respectively. The relative muon p_T resolution derived from the resolution of the dimuon invariant mass is roughly 2% for Z decays.

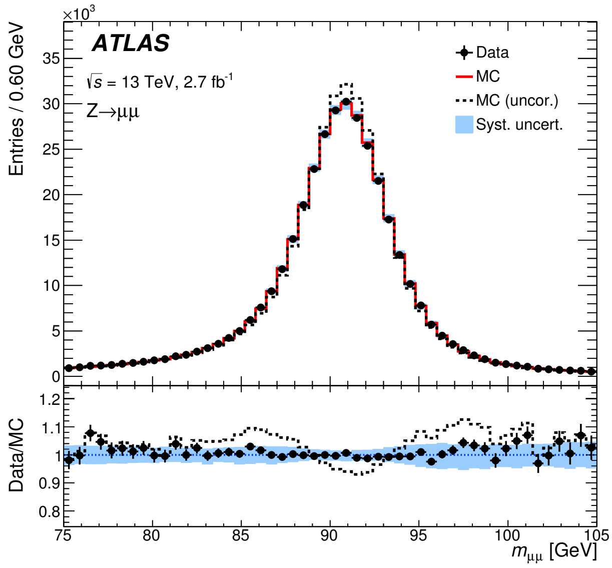


Figure 5.6: The dimuon invariant mass distribution of $Z \to \mu\mu$ combined muon candidate events for 2.7 fb^{-1} of 2015 data and simulation. The points show the data, and the continuous line corresponds to the simulation with the MC momentum corrections applied while the dashed lines show the simulation when no correction is applied. The residuals refer to the data points with respect to the corrected MC [70].

5.3.4 Muon trigger

Muon triggers are seeded by L1 candidates from the muon spectrometer trigger chambers described above. The L1 candidates seed the HLT where information from the MDTs are added to construct a simple muon spectrometer track. The track is then matched (if possible) to the inner detector and, for muons with large enough p_T, an online track reconstruction algorithm is performed. Muon triggers can also combine information about identification and isolation at this stage to reduce the event rate.

5.4 Lepton isolation

In addition to the identification criteria described above, most analyses require light leptons to fulfill isolation requirements. The isolation variables measure the detector activity around a lepton candidate. This allows to discriminate between prompt leptons (from heavy boson decays, such as from W, Z or H bosons) and other non-isolated lepton candidates. For electrons, these non-isolated candidates could originate from converted photons produced in hadron decays, electrons from heavy flavour hadron decays, and misidentified light hadrons [68]. For muons, non-isolated cases are typically from semileptonic light and heavy flavour decays [70]. Two lepton isolation variables are defined to help categorise light leptons: calorimeter isolation and track isolation.

For the calorimeter isolation, topological cluster transverse energies (at the EM scale) are summed within a cone of $\Delta R = 0.2$ around the candidate lepton. Other calorimeter isolation variables with cones of $\Delta R = 0.3$ and $\Delta R = 0.4$ are also used. The lepton cluster energy itself is subtracted from the isolation and the extra contribution from pileup and the underlying event activity is corrected for on an event-by-event basis using an ambient energy density technique, detailed in [72]. This correction allows for the calorimeter isolation to be negative in certain cases.

The track isolation is defined as the sum of transverse momenta of tracks within a variable cone around the lepton, dependent on p_T. The typical cone size is $\Delta R = \min(0.2, 10\text{ GeV}/p_T)$; that is for $p_T < 50$ GeV the cone size is $\Delta R = 0.2$ and $\Delta R = 10\text{ GeV}/p_T$ for $p_T > 50$ GeV. Again the minimum $\Delta R = 0.2$ can be swapped for isolation variables with radius parameters $\Delta R = 0.3$ or $\Delta R = 0.4$. Tracks used in the calculation must have $p_T > 1$ GeV, a number of silicon hits and $|z_0 \sin\theta| < 3$ mm.

Further corrections are applied to electron isolation variables; for the calorimeter isolation the electron energy leakage outside the cluster is corrected for, and for the track isolation all electron-associated (including from brehmsstrahlung) tracks are subtracted.

5.5 Jets

Jets are tricky objects. The main aim of defining a jet is to estimate the evolution of the partons that took part in a hard scatter and therefore a crucial aspect of defining a jet is that the hadronisation process has little effect on the final result. In this way, good jet algorithms account for collinear and soft emissions; if a parton undergoes a collinear split then the resulting jet is the same as if it had not split, and similarly for soft emission.

The most commonly used jet algorithm in ATLAS is the anti-k_t jet clustering algorithm [73]. The algorithm clusters objects iteratively within a radius parameter, R. The "distance" measures

$$d_{ij} = \min\left(\frac{1}{k_{Ti}^2}, \frac{1}{k_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{k_{Ti}^2} \tag{5.5}$$

are defined, where $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ for rapidity, y, azimuthal angle, ϕ, and object transverse momentum k_T. The algorithm compares the distance between the objects, d_{ij}, with the distance between object i, and the beam, B, d_{iB}. If d_{ij} is smallest, i and j are combined otherwise i is said to be a jet and is removed from the list of objects. This procedure is then performed iteratively. The inverse nature of the object transverse momentum results in harder objects being merged into jets first. Because of this, the jet boundary is insensitive to soft particles and are thus regularly shaped. Generally a radius parameter of $R = 0.4$ is used to collect objects, and only anti-k_t jets with this radius parameter are considered in this thesis hereafter.

The reconstruction, identification and energy calibration of jets at ATLAS are discussed in Sections 5.5.1, 5.5.2 and 5.5.3 below. Algorithms developed to tag b-jets are discussed in Section 5.5.4.

5.5.1 Jet reconstruction

Jets typically deposit most of their energy in the electromagnetic calorimeters. These include π_0 hadrons decaying to two photons and softer jets. However, higher p_T jets manage to deposit some of their energy in the hadronic calorimeters. A three-dimensional topological clustering (topo-clusters) of the individual calorimeter cell energy deposits is performed [74], being built from neighbouring calorimeter cells containing large energy deposits with respect to the expected level from noise and pile-up [75]. The topo-clusters are then used as input in the anti-k_t jet algorithm with $R = 0.4$, as mentioned above. It should be noted that objects other than topo-clusters are used to cluster jets. *Track jets* cluster tracks (with $p_T > 400$ MeV) to reconstruct jet objects, and are used extensively in Chapter 7.

5.5.2 Jet identification

One wishes to reconstruct jets from proton-proton collisions from the primary vertex in an event. Background jets can be produced from beam induced backgrounds [76], where protons upstream from the collision point can collide with something (such as beam gas particles) and produce high energy muons. These muons have the capability of depositing large amounts of energy in the calorimeters and to be mistakenly reconstructed as a jet. Other jet backgrounds include high energy muons from cosmic rays and calorimeter noise [77]. Quality criteria are applied to jets to suppress these backgrounds. The number of background jets in ATLAS are extremely small and negligible to most analyses.

Further to rejecting background jets, the so-called jet vertex tagger (JVT) [78] suppresses jets from pile-up interactions. It is designed to provide a stable jet efficiency as a function of the number of vertices in an event, N_{vtx}. The tagger is a multivariate combination of two track based variables used to suppress pile-up jets.

5.5.3 Jet energy calibration

The energy of the topo-clusters are that measured as the energy deposited due to electromagnetic interactions within the calorimeter material. This does not involve any hadronic energy showering. The true jet energy scale calibration therefore requires multiple steps.

Firstly an origin correction points the clustered jet to the primary vertex. Then pile-up corrections are applied that remove any excess energy deposited (due to both in-time and out-of-time pile-up) and any dependence on the number of vertices in the event. The jet energy, currently at the EM scale, is then corrected for in a MC-based calibration to the particle-level energy scale. Finally an in-situ calibration, characterising the differences between MC and data, is applied to jets only in data. Further details of the procedure can be found in [75]. The uncertainty in the jet energy scale is at a relative level of 4.5% for jets with $p_T = 20$ GeV and decreasing to 1% at 200 GeV, with fair stability as a function of jet $|\eta|$.

5.5.4 *b*-jet tagging

The flavour of a jet reconstructed at ATLAS is, in general, ambiguous. That is, the jet could have been initiated by any flavour quark or even a gluon. However, large efforts are undertaken to identify b-jets that are interesting objects to study for many analyses. The method for determining b-jets is discussed below.

A number of different algorithms are used to identify b-jets. B-hadrons (containing b-quarks) have significant lifetimes and thus decay some distance from the primary vertex, dependent on the boost of the system. This lifetime allows for discrimination from light jets. The main algorithms used to determine if a jet is a b-jet are impact parameter based, secondary vertex based and multi-vertex based.

Since the b-quark decays away from the primary vertex, the charged particle tracks originating from the b-quark secondary vertex are likely to have large impact parameter (which is calculated from the primary vertex). Two algorithms utilise this effect, both using the signed transverse impact parameter and one additionally the signed longitudinal impact parameter. The sign is determined using the jet direction: the impact parameter is positive if the track intersects the jet axis in front of the primary vertex. Tracks originating from b-quarks typically have positive impact parameter, with light flavour jets roughly sign-symmetric. This is illustrated in tracks from jets in simulated $t\bar{t}$ events in Figure 5.7 [79].

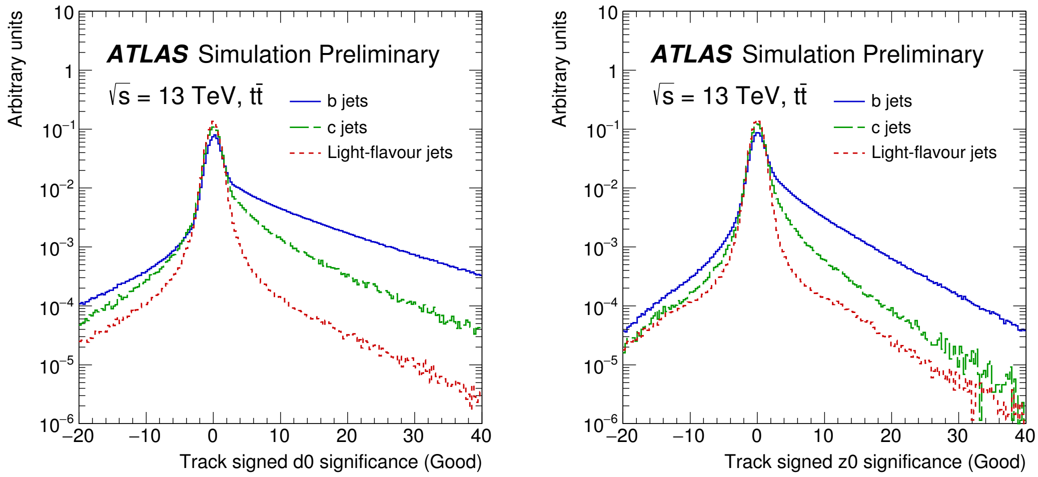


Figure 5.7: The signed (a) transverse and (b) longitudinal impact parameter significance of *Good* tracks in $t\bar{t}$ events for b-, c- and light flavour jets [79]. *Good* tracks refer to tracks that pass requirements on the number of hits in the inner detector.

Tracks are categorised dependent on the number of pixel hits and expected hits in the IBL and BL and the number of silicon hits. PDFs are obtained from the signed impact parameters and are used to calculate ratios of the b-, c- and light flavour jet probabilities. A log-likelihood ratio discriminant is used,

$$\mathrm{IPXD} = \sum_{i=1}^{N} \ln\left(\frac{p_b}{p_u}\right), \tag{5.6}$$

where N is the number of tracks collected around the jet and $p_{b,u}$ the PDFs for the b- and light jet hypotheses, dependent on the track quality. The discriminant is shown for b-, c- and light jets in Figure 5.8.

A secondary vertex algorithm, SV1, attempts to reconstruct a displaced secondary vertex in the jet. All track pairs within a jet are tested to produce a vertex and such vertices are rejected if they are likely to originate from long lived light quark jets (such as K_s^0 or Λ^0), photon conversions or hadronic interactions [79]. The resulting vertices are combined into an inclusive secondary vertex.

A further decay chain algorithm, called JetFitter [80], attempts to reconstruct the full $b \to c$ chain within a jet. A Kalman filter is applied to determine the line of flight of the b-quark hadron. The b- and c-quark hadron decay vertices are assumed to lie along this line and therefore all charged particles from said vertices intersect this line. This allows for the possibility of either the b- or c-vertices to be resolved from a single charged particle track. Requiring a $b \to c$ cascade topology greatly reduces the light jet fake rate and results in complementary information to the robust SV1 algorithm.

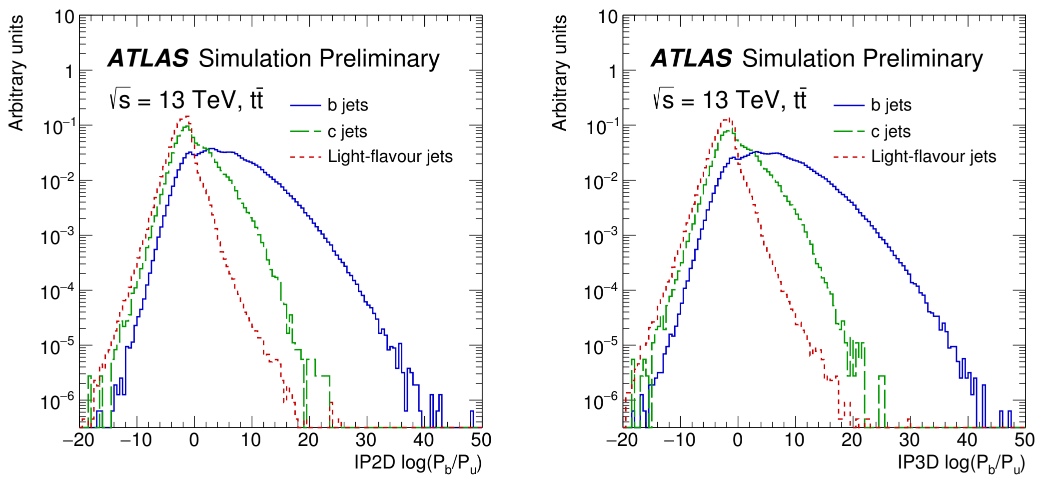


Figure 5.8: The (a) IP2D and (b) IP3D log-likelihood ratio for b-, c- and light flavour jets [79].

All algorithms discussed above are combined into a boosted decision tree (BDT) algorithm, called MV2c10. MV2c10 uses both impact parameter/vertexing algorithmic information and kinematic information. In the training of the algorithm, the p_T and $|\eta|$ of b-jets are reweighted to resemble the light jet background. The background sample contains 93% light jets and 7% c-jets, since the majority of analyses require greater light jet rejection than c-jet. A dedicated algorithm to separate b-jets and c-jets is also developed for use by those where this assumption is not true.

The distribution of the MV2c10 discriminant for b-, c- and light jets is shown in Figure 5.9, along with the background rejection factors as a function of b-jet efficiency. Four working points are defined for use in tagging b-jets, described by a one dimensional cut to the MV2c10 distribution corresponding to b-jet efficiencies of 85%, 77%, 70% and 60%. Each of these working points are calibrated by deriving scale factors for simulated events to correct for the disagreement observed in the modelling of the MV2c10 distribution when compared to data [81]. In addition to b-jet calibration, the mistag rates of light flavour jets are also calibrated by determining the efficiency of light flavour jets that pass a *negative-tag* algorithm, which reverses some of the criteria used in the nominal identification algorithm [82].

5.6 Hadronic taus

Taus can decay hadronically or leptonically, roughly in 64.7% and 36.3% of cases respectively [3]. There is very little information to discriminate a $W/Z \to \tau \to \ell$ decay from a $W/Z \to \ell$ decay. Typically the hadronic activity from hadronic tau decays are used to reconstruct tau leptons at ATLAS [83].

Tau candidates are seeded by jets with a local hadronic calibration (LC). Tau candidates are vetoed in the $1.37 < |\eta| < 1.52$ crack region. A vertexing procedure is applied to reconstruct the tau vertex; the vertex with the largest fraction of momentum from tracks within $\Delta R < 0.2$ of the jet seed. The tracks must have $p_T > 1$ GeV and contain silicon hit requirements. If a track is close to the tau vertex (satisfying $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm) then the track is associated to the tau; if within $\Delta R < 0.2$ a core track or an isolation track if within $0.2 < \Delta R < 0.4$. Only tau

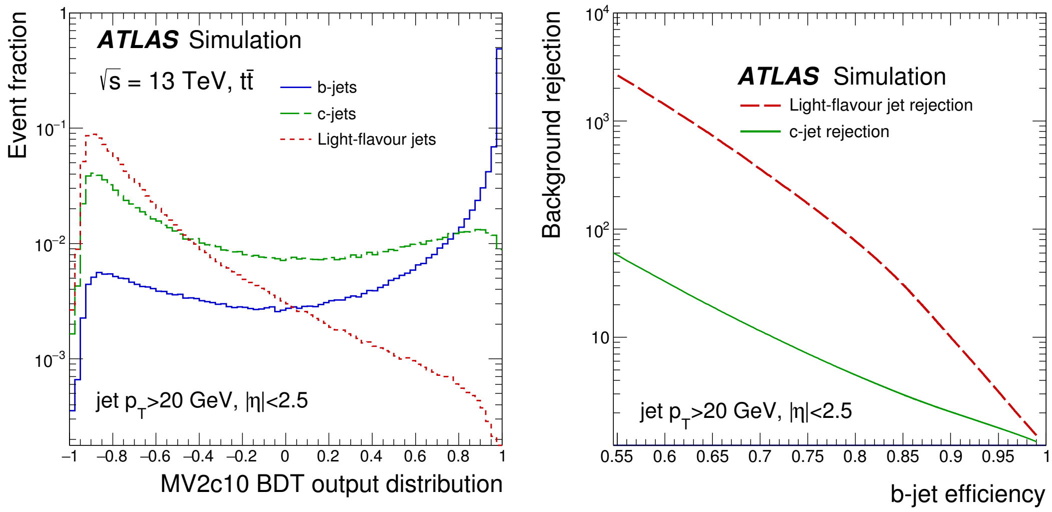


Figure 5.9: (a) the MV2c10 BDT distribution for b-, c- and light jets in simulated $t\bar{t}$ events and (b) the c- and light jet rejection factors as a function of b-jet tagging efficiency of the MV2c10 algorithm [81].

candidates with one or three associated tracks are used, corresponding to one prong or three prong tau decays [3].

The tau energy scale is calibrated in order to correct energy measured in the detector to that of the *true* value of the decay products in MC. The baseline correction calibrates the tau by applying a LC-calibrated sum of the energy of the topo-clusters within $\Delta R < 0.2$ of the tau and accounting for the energy deposited by pile-up. This correction is scaled by a detector response determined from MC. The degradation of this calibration at low p_T results in a multivariate based calibration also being applied. The multivariate calibration aims to reconstruct the charged and neutral pion decays around the tau. The improvement in energy resolution with respect to the baseline correction is due to the track momentum measurement of the charged pions from the inner detector. More details are in [83].

Since hadronic taus are actually jets, the largest background comes from hadronic activity. In order to reject jets that resemble taus, boosted decision trees (BDTs) containing calorimeter- and track-based variables are trained to reject quark and gluon jets. The distributions of the BDT scores in $Z \to \tau\tau$ events are shown in Figure 5.10. The BDT algorithm is discussed further in Chapter 7.

`Loose`, `Medium` and `Tight` identification working points are defined by targeting flat efficiencies, independent of the tau p_T, by applying a number of cuts on the jet BDT score as a function of tau p_T. For one prong taus, the target efficiencies are 60%, 55% and 45% respectively; for three prong, 50%, 40% and 30%. In addition to the jets, one-prong taus can look very similar to an electron candidate. To reduce this background, any reconstructed one-prong taus within $\Delta R < 0.4$ of a reconstructed electron are vetoed if the electron passes a `VeryLooseLH` electron LH working point. The electron likelihood score is very loose and specifically tuned to provide 95% efficiency for hadronic taus when vetoed. The electron likelihood score is not for use in identifying electrons in an analysis.

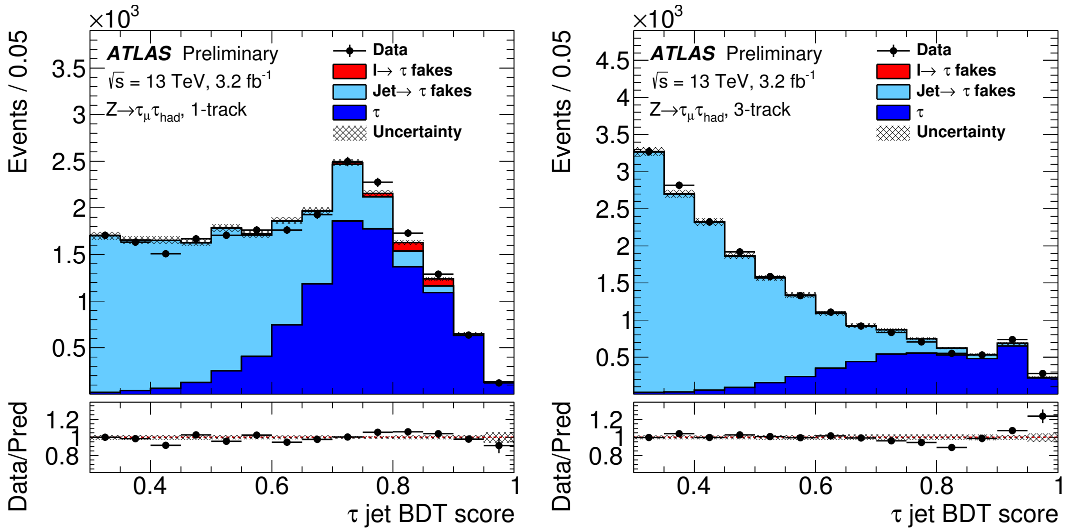


Figure 5.10: The jet BDT distribution for one track (left) and three track (right) τ_{had} candidates. The uncertainty band contains only the statistical uncertainty [83].

Chapter 6

Non-factorisation effects in the calibration of luminosity

Accurate measurements of luminosity at hadron colliders are vital for precision measurements in many physics analyses. This chapter discusses how the luminosity is calculated (Section 6.1), measured (Section 6.2) and calibrated (Section 6.3) in ATLAS. A particular emphasis is put on the determination of a correction and the corresponding systematic uncertainty due to so-called non-factorisation effects in the calibration of luminosity in van der Meer (vdM) scans in Section 6.4.

6.1 Measurements of luminosity in a hadron collider

The instantaneous luminosity of a single proton-proton bunch pair, $\mathcal{L}_b$, colliding head on (with no crossing angle) in a hadron collider is given by

$$\mathcal{L}_b = f_r n_1 n_2 \int \hat{\rho}_1(x,y)\hat{\rho}_2(x,y)dxdy, \tag{6.1}$$

where f_r is the LHC bunch revolution frequency, $n_{1,2}$ the number of protons in each proton bunch and $\hat{\rho}_{1,2}$ the normalised transverse particle beam densities [84]. Measuring the transverse beam densities is non-trivial and, in general, assumptions need to be made to determine the absolute scale of luminosity. The method to determine this scale for the ATLAS detector is discussed in Section 6.3.

The instantaneous luminosity of a bunch can also be defined as

$$\mathcal{L}_b = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu f_r}{\sigma_{\text{inel}}} \tag{6.2}$$

with R_{inel} the rate of inelastic proton-proton collisions. R_{inel} can be expressed in terms of the product of f_r and the pile-up parameter μ, the average number of inelastic interactions per bunch crossing [85]. The total instantaneous luminosity is therefore the sum of $\mathcal{L}_b$ over all bunches.

The true value of the number of inelastic interactions per crossing is not directly measurable by detectors in ATLAS but a value *visible* to a specific detector, μ_{vis}, is which is related to $\mathcal{L}_b$ by

$$\mathcal{L}_b = \frac{\mu_{\text{vis}} f_r}{\sigma_{\text{vis}}}, \tag{6.3}$$

with σ_{vis} the *visible* proton-proton inelastic cross section, related to σ_{inel} by an efficiency ϵ, $\sigma_{\text{inel}} = \epsilon\sigma_{\text{vis}}$. σ_{vis} is a calibration quantity that is determined for each luminosity detector in ATLAS during van der Meer scans [86].

6.2 Measurements of luminosity with the ATLAS detector

A number of different luminosity detectors are employed by ATLAS. Each detector typically measures more than one observable used to determine μ_{vis}.

In standard physics running many proton bunches are provided by the LHC for collision. Each bunch crossing in ATLAS is given a unique identification, called a BCID number. The main luminosity detectors that can determine μ_{vis} from a single BCID are the Lucid [87] and the BCM [88] detectors. Both detectors are situated in the forward (and backward) regions of ATLAS. Lucid consists of photomultiplier tubes (PMTs) that measure Cherenkov light produced when particles traverse quartz fibre windows. BCM consists of diamond sensors that are ionised from the passage of charged particles. The main luminosity algorithms used by ATLAS to quantify the total luminosity delivered and collected in 2015 and 2016 use data collected by the Lucid detector.

Both Lucid and BCM have a number of different algorithms used to determine the visible rate, μ_{vis}. These are in two categories: event counting and hit counting algorithms. Event counting algorithms measure μ_{vis} by counting the number of bunch crossings with at least one hit in the detector. They are further categorised into EventOR and EventAND algorithms; EventOR corresponding to at least one hit on either side of the interaction point, and EventAND a hit on both sides. EventOR algorithms are the nominal event counting algorithms used in ATLAS. From Poisson statistics one can determine the probability, P, for such events

$$P = 1 - e^{-\mu_{\text{vis}}} \approx \frac{N_{\text{hit}}}{N_{\text{all}}}, \tag{6.4}$$

with N_{hit} the number of bunch crossings with at least one hit compared to all bunch crossings, N_{all}. Therefore μ_{vis} can be calculated:

$$\mu_{\text{vis}} = -\ln(1 - \frac{N_{\text{hit}}}{N_{\text{all}}}). \tag{6.5}$$

One notices a problem with this algorithm however; if every bunch crossing has a hit ($N_{\text{hit}} = N_{\text{all}}$) then the algorithm saturates due to the logarithm of zero. This is, in fact, what happened due to the large pileup during 2016 data taking. In these cases the event counting algorithm can be swapped for a hit counting algorithm, that measures the number of hits in a detector per BCID in a similar way.

Other luminosity detectors can measure the total luminosity but are not sensitive to measurements of luminosity for individual BCIDs. The main methods used to do this are the measurements of the current drawn by the PMT readouts of the tile calorimeter (Tile) and the counting of tracks in the inner detector (track counting). The basic premise behind both methods is that the luminosity is proportional to the number of particles passing through the detector. These bunch integrating detectors are not calibrated by the vdM procedure which require bunch-by-bunch measurements for best accuracy, but are very useful as a reference in determining the long term stability of the luminosity as measured by Lucid.

6.3 The calibration of luminosity with the ATLAS detector

Luminosity is calibrated in ATLAS using information from beam separation (or vdM) scans [89]. The process was pioneered by Simon van der Meer at the ISR accelerator at CERN [86].

6.3.1 The van der Meer method

The van der Meer method allows for a measurement of the beam overlap integral in Equation 6.1 without measuring the transverse beam densities. If one imagines beams colliding with a slight offset in the horizontal direction x, δ, then the observed inelastic collision rate would be

$$\mu_{\text{vis}}(\delta) = \epsilon \int \hat{\rho}_1(x+\delta)\hat{\rho}_2(x)dx, \tag{6.6}$$

where ϵ is a constant referring to the efficiency of measuring μ with a specific detector[1]. If one *scans* the beams totally, then the integral over the full phase space can be made

$$\begin{aligned} \int \mu_{\text{vis}}(\delta)d\delta &= \epsilon \int_\delta \int_x \hat{\rho}_1(x+\delta)\hat{\rho}_2(x)dxd\delta \\ &= \epsilon \int_x \Big(\int_\delta \hat{\rho}_1(x+\delta)d\delta \Big) \hat{\rho}_2(x)dx \\ &= \epsilon \end{aligned} \tag{6.7}$$

with the last line using the definition that the bunch density profiles are normalised (the integral of the beam profile over the full phase space is unity). Taking the ratio of the maximal value of μ_{vis} to that of the full integral over δ allows one to define the convolved beam size, Σ_x,

$$\frac{\mu_{\text{vis}}^{\delta=0}}{\int \mu_{\text{vis}}(\delta)d\delta} = \int \hat{\rho}_1(x)\hat{\rho}_2(x)dx = \frac{1}{\sqrt{2\pi}\Sigma_x}. \tag{6.8}$$

The method can be extended to two dimensions, now in x and y:

$$\frac{\mu_{\text{vis}}^{\delta_x,\delta_y=0}}{\int \mu_{\text{vis}}(\delta_x,\delta_y)d\delta_x d\delta_y} = \int \hat{\rho}_1(x,y)\hat{\rho}_2(x,y)dxdy = \frac{1}{2\pi\Sigma_x\Sigma_y}. \tag{6.9}$$

Time constraints on the use of the beam in ATLAS at the LHC (the LHC can only conduct beam separation scans in a single collision point at any one time) means that a grid measurement of all points in the (x, y) phase space is not possible. Instead two beam separation scans are typically conducted, one spanning the horizontal direction, x, and one vertical, y, in a number of increments (typically 25 increments in 2015 and 2016). To calibrate the luminosity with this method requires the assumption that the beam profiles factorise into x and y components,

$$\int \hat{\rho}_1(x,y)\hat{\rho}_2(x,y)dxdy = \int \hat{\rho}_1(x)\hat{\rho}_2(x)dx \int \hat{\rho}_1(y)\hat{\rho}_2(y)dy. \tag{6.10}$$

The extent to which this assumption is valid in 2015 and 2016 (the degree of *non-factorisation* is discussed in Section 6.4.

Assuming factorisation and using the results from Equations 6.1 and 6.9 gives the expression for the instantaneous luminosity of a colliding bunch,

$$\mathcal{L}_b = \frac{f_r n_1 n_2}{2\pi\Sigma_x\Sigma_y}. \tag{6.11}$$

This can be combined with Equation 6.3 to allow for a calibration of the detector used to measure μ_{vis},

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\delta=0}\frac{2\pi\Sigma_x\Sigma_y}{n_1 n_2}. \tag{6.12}$$

6.3.2 Non-factorisation

So-called non-factorisation effects are measured at ATLAS one of two ways. The first, named the *coupled model* method, compares coupled and uncoupled fits to the luminosity profiles of a

[1] Here beam 1 is shifted by δ with beam 2 remaining in its nominal position, where in fact both beams are shifted by $\delta/2$ in beam separation scans.

luminometer where the x-y factorisation is turned on and off in the two fit models. The second, named the *luminous region* method, determines the position and width of the luminous region, along with the total rate (luminosity), by determining single beam density profiles.

Each method has benefits and detriments. A study of the non-factorisation using the luminous region method in the 2015 and 2016 proton-proton vdM sessions at ATLAS is discussed in Section 6.4.

6.4 Non-factorisation determination with the luminous region method

Determining non-factorisation effects in the calibration of luminosity from vdM scans requires some understanding of the underlying distributions of the proton beams. The coupled model method simultaneously studies luminosity, or collision rate, profiles in the horizontal and vertical scans. Such collision rates are actually produced from the convolution of the two beams. General fit functions, such as the sum of two Gaussians (double Gaussian) or Gaussian-times-polynomial, are used to determine this convolution. The coupled model method therefore determines a non-factorisation correction and uncertainty agnostic to the individual beam profiles themselves.

In the luminous region method the understanding of the individual beam profiles is more fundamental. The luminous region, or beam spot, can also be defined by the convolution of the two individual beam profiles, as with luminosity. How the beam spot evolves during a vdM scan, as well as how the rate evolves, is used to determine density profiles for each beam. When these profiles are known the degree of non-factorisation is trivial to determine. A detailed analysis of the non-factorisation determined with the luminous region method for the 2015 and 2016 vdM scan periods is discussed below.

6.4.1 Types of beam separation scan

Different areas of phase space are explored in the ATLAS vdM program to aid in measuring any non-factorisation effects. There are three types of beam separation scan that are employed to this end.

The nominal scan type is a *centred* scan pair. The beams are separated either horizontally or vertically. Both beams are separated by the same amount in opposite directions. The data from these scans are nominally used to calibrate luminosity at ATLAS. The two other scan types are special scans employed purely to determine non-factorisation effects.

The first special scan type is an *offset* (or off-axis) scan pair. The format is similar to the centred scan, except that the beams are offset in the orthogonal transverse direction to the scanning direction. For example, the horizontal offset scans in 2015 and 2016 were performed with a $\sim$300 μm separation offset in the vertical direction, and vice-versa for the vertical scan. These scans explore the tails of the beams, where one might expect non-Gaussian effects to be important.

The second special scan type is a (single) *diagonal* scan. In this case the beams are separated both horizontally and vertically, and between each scan step both horizontal and vertical separations are changed by the same amount. This scan explores the region between the centred scans, and again provides complementary information to both the centred and offset scans. A simple diagram describing the positions of the individual beams in each of the three vdM scan types is shown in Figure 6.1.

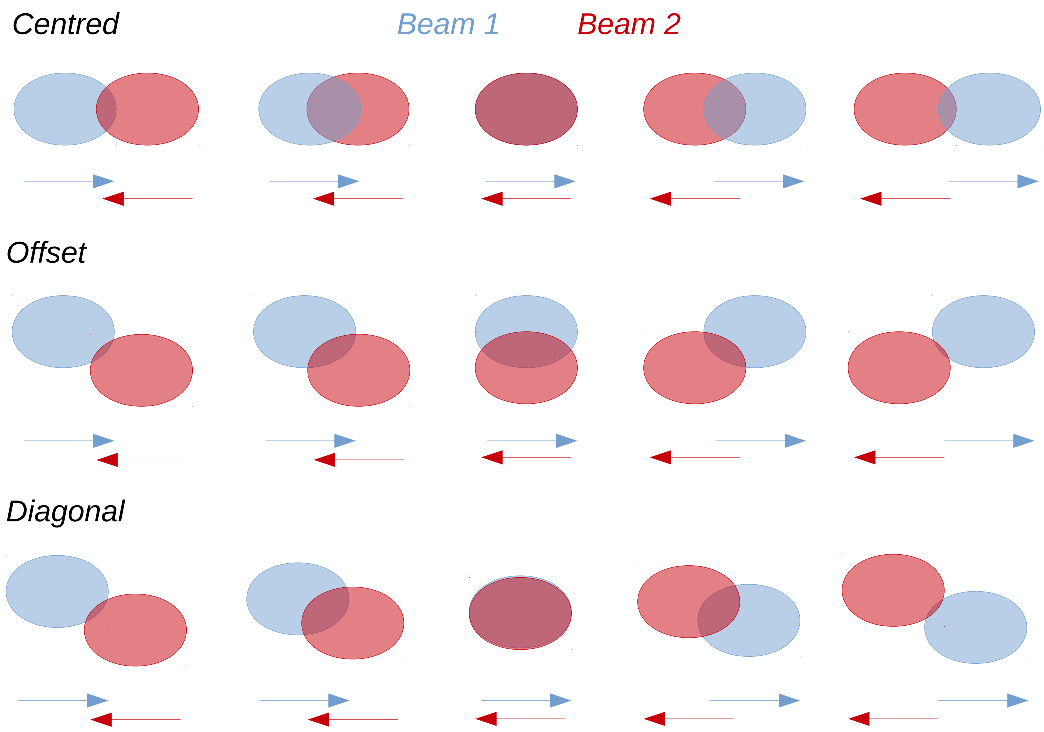


Figure 6.1: A schematic diagram showing the positions of each beam during a single beam separation scan with five steps for each of the beam separation scan types. The beams are being separated in the transverse plane in ATLAS and have slightly different approximate widths in the x and y directions.

6.4.2 Beam spot data

The luminous region is the region in which proton-proton collisions occur in ATLAS during LHC running. The evolution of the beam spot in a vdM scan can be used to determine the underlying distributions of the beams and are thus used to determine any present non-factorisation effects. The characteristics of the beam spot are modelled from reconstructed vertex data and the resulting information, such as the modelled three dimensional mean positions and widths, are used as input to the non-factorisation analysis.

The beam spot is modelled by a three dimensional Gaussian maximum likelihood fit to the spatial distribution of vertices in the luminous region over a period of time [90]. From this Gaussian fit the three dimensional mean position and mean widths of the beam spot (in x, y and z) can be taken from the means and widths of the fitted Gaussian. In a vdM scan the beam spot fit is performed from the vertices produced at each scan step. The evolution of the positions and widths of the beam spot during the scan can then be used to determine proton beam densities.

The beam spot fit takes into account the difference between the expected and actual vertex resolution with a dimensionless scale factor, k. This scale factor is typically allowed to float in the beam spot fit.

6.4.3 Luminosity data

Any luminosity algorithm could be used in the luminous region analysis in conjunction with beam spot data, with the caveat of providing a good statistical accuracy.

The luminosity data and algorithm chosen is from the Lucid detector, called LucidEvtOR. It is used for both the 2015 and 2016 luminous region analyses. The algorithm is an EventOR algorithm using data from the Lucid detector, providing an excellent statistical accuracy of μ_{vis} at all scan steps in the vdM scan, including the outermost scan steps where only the edges of the beams are convolved.

6.4.4 Fit model

One needs to determine a fit model for each proton beam density profile in the fit. A good first guess is a three dimensional Gaussian, as in the beam spot analysis. Functions to describe the proton beam density profiles including the Gaussian distribution are motivated by the central limit theorem. However, it was found that this fit model did not well describe the data in 2011 and 2012 [91], due to the model only generating linear predictions of the movement of the luminous centroid in the scan. Non-linear movements were seen. It is also a known fact that the beams in a hadron machine are not always fully Gaussian because of slow diffusion processes[2] [92].

The most successful fit model when the luminous region measurement was performed on the 2012 vdM scans was the double Gaussian model: the sum of two Gaussians [93, 94]. Both the single and double Gaussian models are useful to determine single beam density profiles since the convolution of two Gaussians are themselves Gaussians and the integral can thus be calculating analytically. The double Gaussian model consists of four convolutions compared to the single convolution of the single Gaussian model.

6.4.5 Evolution of the luminosity and beam spot observables

The luminosity and beam spot observables can be calculated from the density profiles of each beam. The luminosity density, L, at some beam separation (δ_x, δ_y) as a function of $\mathbf{x}$ is proportional to the time integral of the product of the profiles [92]

$$\begin{aligned} L(\mathbf{x}, \delta_x, \delta_y) &\propto \int \rho_1(\mathbf{x}, t, \delta_x, \delta_y)\rho_2(\mathbf{x}, t, \delta_x, \delta_y)dt \\ &= \rho_1(x, y)\rho_2(x, y) \int \rho_1(z, t)\rho_2(z, t)dt. \end{aligned} \tag{6.13}$$

The second line factorises the beams dependent on time and requires that there is no (or minimal) beam crossing angle. Only the transverse dependence is discussed in this thesis and the solution to the time integral is neglected, but can be found in [93] and [94].

The luminosity, $\mathcal{L}$, itself is thus

$$\mathcal{L}(\delta_x, \delta_y) = \int L d^3\mathbf{x}. \tag{6.14}$$

The beam spot centroid position is given by the expectation value of $\mathbf{x}$, $\langle \mathbf{x} \rangle$,

$$\langle \mathbf{x} \rangle(\delta_x, \delta_y) = \frac{1}{\mathcal{L}} \int \mathbf{x}\rho_1\rho_2 d^3\mathbf{x}dt, \tag{6.15}$$

[2] This is not the case for lepton machines where synchotron radiation damps such slow diffusion process causing non-Gaussian profiles.

and the beam spot width by the standard deviation of the luminosity density, σ. The variance, s, is given by

$$s(\delta_x, \delta_y) = \langle \mathbf{x}^2 \rangle - \langle \mathbf{x} \rangle^2 = \frac{1}{\mathcal{L}} \int \mathbf{x}^2 \rho_1 \rho_2 d^3\mathbf{x} dt - \langle \mathbf{x} \rangle^2, \tag{6.16}$$

and hence the standard deviation is

$$\sigma(\delta_x, \delta_y) = \sqrt{s}. \tag{6.17}$$

Given a model to describe the individual proton beam densities allows one to measure such observables and compare them to data.

6.4.6 Gaussian beam models

As mentioned, models incorporating Gaussian functions are motivated for use in the description of proton bunch densities. How a single Gaussian model (Section 6.4.6.1) and multi-Gaussian models (Section 6.4.6.2) can describe the luminous region is discussed below.

6.4.6.1 Single Gaussian model

A normalised, three dimensional Gaussian as a function of the position vector, $\mathbf{x}$, is given by

$$\mathcal{N}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}|}} \exp\Big(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \Big), \tag{6.18}$$

where $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ the covariance matrix,

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \kappa\sigma_x\sigma_y & 0 \\ \kappa\sigma_x\sigma_y & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{pmatrix}, \tag{6.19}$$

and κ is the x-y correlation coefficient. The x-z and y-z correlations are set to zero due to the zero beam crossing angle condition. If there was a non-negligible beam crossing angle then these terms would be needed in the model.

If one parameterises each beam as a Gaussian then the luminous region is the product of two Gaussians which is also a Gaussian. This can be determined analytically.

The product of two Gaussians $\mathcal{N}(\mathbf{x}, \boldsymbol{\mu_1}, \boldsymbol{\Sigma_1})$ and $\mathcal{N}(\mathbf{x}, \boldsymbol{\mu_2}, \boldsymbol{\Sigma_2})$ is:

$$\begin{aligned} \mathcal{N}(\mathbf{x}, \boldsymbol{\mu_1}, \boldsymbol{\Sigma_1})\mathcal{N}(\mathbf{x}, \boldsymbol{\mu_2}, \boldsymbol{\Sigma_2}) &= A\exp\Big(-\frac{1}{2}\big[(\mathbf{x} - \boldsymbol{\mu_1})^T \boldsymbol{\Sigma_1}^{-1}(\mathbf{x} - \boldsymbol{\mu_1}) + (\mathbf{x} - \boldsymbol{\mu_2})^T \boldsymbol{\Sigma_2}^{-1}(\mathbf{x} - \boldsymbol{\mu_2})\big]\Big) \\ &= A'\exp\Big(-\frac{1}{2}\big[\mathbf{x}^T(\boldsymbol{\Sigma_1}^{-1} + \boldsymbol{\Sigma_2}^{-1})\mathbf{x} - \mathbf{x}^T(\boldsymbol{\Sigma_1}^{-1}\boldsymbol{\mu_1} + \boldsymbol{\Sigma_2}^{-1}\boldsymbol{\mu_2}) \\ &\qquad - (\boldsymbol{\mu_1}^T\boldsymbol{\Sigma_1}^{-1} + \boldsymbol{\mu_2}^T\boldsymbol{\Sigma_2}^{-1})\mathbf{x}\big]\Big). \end{aligned} \tag{6.20}$$

The normalisation factor A is changed to A' in the second line to encompass the terms independent of $\mathbf{x}$. Now two new parameters can be defined

$$\begin{aligned} \boldsymbol{\Lambda} &= \boldsymbol{\Sigma_1}^{-1} + \boldsymbol{\Sigma_2}^{-1}, \\ \boldsymbol{\nu} &= \boldsymbol{\Sigma_1}^{-1}\boldsymbol{\mu_1} + \boldsymbol{\Sigma_2}^{-1}\boldsymbol{\mu_2}, \end{aligned} \tag{6.21}$$

so that

$$\mathcal{N}(\mathbf{x}, \boldsymbol{\mu_1}, \boldsymbol{\Sigma_1})\mathcal{N}(\mathbf{x}, \boldsymbol{\mu_2}, \boldsymbol{\Sigma_2}) = A'\exp\Big(-\frac{1}{2}\big[\mathbf{x}^T\boldsymbol{\Lambda}\mathbf{x} - \mathbf{x}^T\boldsymbol{\nu} - \boldsymbol{\nu}^T\mathbf{x}\big]\Big). \tag{6.22}$$

We can introduce a new term in the exponent to complete the square

$$\mathcal{N}(\mathbf{x},\boldsymbol{\mu_1},\boldsymbol{\Sigma_1})\mathcal{N}(\mathbf{x},\boldsymbol{\mu_2},\boldsymbol{\Sigma_2}) = A''\exp\Big(-\frac{1}{2}\big[\mathbf{x}^T\boldsymbol{\Lambda}\mathbf{x} - \mathbf{x}^T\boldsymbol{\nu} - \boldsymbol{\nu}^T\mathbf{x} + \boldsymbol{\nu}^T\boldsymbol{\Lambda}\boldsymbol{\nu}\big]\Big). \tag{6.23}$$

The extra $\boldsymbol{\nu}^T\boldsymbol{\Lambda}\boldsymbol{\nu}$ term is taken away from A' and incorporated into A''. It should be noted that $\boldsymbol{\Lambda}$ is symmetric and invertible since it is the weighted sum of two symmetric covariance matrices. Therefore new variables can be introduced:

$$\begin{aligned}\boldsymbol{\Sigma} &= \boldsymbol{\Lambda}^{-1},\\ \boldsymbol{\mu} &= \boldsymbol{\Lambda}^{-1}\boldsymbol{\nu}.\end{aligned} \tag{6.24}$$

The product then becomes

$$\begin{aligned}\mathcal{N}(\mathbf{x},\boldsymbol{\mu_1},\boldsymbol{\Sigma_1})\mathcal{N}(\mathbf{x},\boldsymbol{\mu_2},\boldsymbol{\Sigma_2}) &= A''\exp\Big(-\frac{1}{2}\big[\mathbf{x}^T\boldsymbol{\Lambda}\mathbf{x} - \mathbf{x}^T\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-1}\boldsymbol{\nu} - \boldsymbol{\nu}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}\mathbf{x} + \boldsymbol{\nu}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-1}\boldsymbol{\nu}\big]\Big)\\ &= A''\exp\Big(-\frac{1}{2}\big[\mathbf{x}^T\boldsymbol{\Lambda}\mathbf{x} - \mathbf{x}^T\boldsymbol{\Lambda}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\Lambda}\mathbf{x} + \boldsymbol{\mu}^T\boldsymbol{\Lambda}\boldsymbol{\mu}\big]\Big)\\ &= A''\exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big)\\ &= A^{\text{lum}}\mathcal{N}(\mathbf{x},\boldsymbol{\mu},\boldsymbol{\Sigma})\end{aligned} \tag{6.25}$$

with normalisation factor

$$\begin{aligned}A^{\text{lum}} &= A'' \times 2\pi\sqrt{|\boldsymbol{\Sigma}|}\\ &= \frac{2\pi\sqrt{|\boldsymbol{\Sigma}|}}{4\pi^2\sqrt{|\boldsymbol{\Sigma_1}||\boldsymbol{\Sigma_2}|}}\exp\Big(-\frac{1}{2}\big[\boldsymbol{\mu_1}^T\boldsymbol{\Sigma_1}^{-1}\boldsymbol{\mu_1} + \boldsymbol{\mu_2}^T\boldsymbol{\Sigma_2}^{-1}\boldsymbol{\mu_2} - \boldsymbol{\nu}^T\boldsymbol{\Lambda}\boldsymbol{\nu}\big]\Big)\\ &= \frac{\sqrt{|\boldsymbol{\Sigma}|}}{2\pi\sqrt{|\boldsymbol{\Sigma_1}||\boldsymbol{\Sigma_2}|}}\exp\Big(-\frac{1}{2}\big[\boldsymbol{\mu_1}^T\boldsymbol{\Sigma_1}^{-1}\boldsymbol{\mu_1} + \boldsymbol{\mu_2}^T\boldsymbol{\Sigma_2}^{-1}\boldsymbol{\mu_2} - \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\big]\Big).\end{aligned} \tag{6.26}$$

The luminous region is hence a single Gaussian from this beam model. The luminosity determined from this overlap is given by

$$\mathcal{L} = \int A^{\text{lum}}\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma}) = A^{\text{lum}}, \tag{6.27}$$

since the Gaussian model is normalised. The centroid position of the beam spot is given simply by $\boldsymbol{\mu}$ and the beam spot widths by the square root of the diagonal indices (variances) of $\boldsymbol{\Sigma}$, each in three dimensions. In this parameterisation the beam spot widths are not dependent on beam positions (and hence beam separation) and the beam spot positions are linear transformations of the individual beam positions. The beam spot widths in data are, in general, seen to vary with separation and the beam spot positions have non-linear behaviour as a function of separation. A better model is needed to describe the beam spot.

6.4.6.2 Multi-Gaussian models

The single Gaussian model above can be extended to models with multiple Gaussians. A double Gaussian model is of the form

$$G_2 = w\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma_1}) + (1-w)\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma_2}), \tag{6.28}$$

where w is a simple weight factor that is applied to couple the single Gaussians. Similarly a triple Gaussian model can be written

$$G_3 = w\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma_1}) + (1-w)[w_b\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma_2}) + (1-w_b)\mathcal{N}(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\Sigma_3})], \tag{6.29}$$

with another weight factor, w_b.

When determining the simulated beam spot, the two multi-Gaussian profiles simulating the individual beams are multiplied together, as shown above for the single Gaussian model. This results in N^2 single Gaussian products for a N-Gaussian model.

The simulated luminosity from the multi-Gaussian beam models are easily determined by summing each single Gaussian product

$$\mathcal{L} = \sum_{i=1}^{N^2} A_i^{\text{lum}}. \tag{6.30}$$

The beam spot positions and widths are estimated using the luminosity evolutions in Equations 6.15 and 6.16 respectively. They are given by [93]

$$\langle \mathbf{x} \rangle = \frac{\sum_{i=1}^{N^2} (A_i^{\text{lum}} \boldsymbol{\mu}_i)}{\mathcal{L}}, \tag{6.31}$$

and the beam spot covariance matrix, $\boldsymbol{\Sigma}_{\text{bs}}$, by

$$\boldsymbol{\Sigma}_{\text{bs}} = \frac{\sum_{i=1}^{N^2} A_i^{\text{lum}} (\boldsymbol{\Sigma}_i + \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i)}{\mathcal{L}} - \frac{\sum_{i=1}^{N^2} \sum_{j=1}^{N^2} (A_i^{\text{lum}} \boldsymbol{\mu}_i)^T (A_j^{\text{lum}} \boldsymbol{\mu}_j)}{\mathcal{L}^2}. \tag{6.32}$$

The beam spot widths are given by the square root of the variances of $\boldsymbol{\Sigma}_{\text{bs}}$, as above. Now the beam spot position in this model is given by a sum of linear terms allowing for non-linearity as a function of beam separation and the beam spot widths are allowed to be non-constant as a function of beam separation.

In the luminous region fits in this thesis a triple Gaussian model is used. This builds on the double Gaussian model used in 2012. In the double Gaussian model weight factors are applied to each Gaussian; w to Gaussian 1 and $(1-w)$ to Gaussian 2. w is limited in the range $[0.0, 1.0]$, therefore two positively weighted Gaussians are used to describe each beam. However, it was found that the transverse beam spot widths in 2015 and 2016 appear to vary with two "minima" either side of nominal beam separation. This is similar to the November scans in 2012 [93]. There is no mechanism in which the double Gaussian, incorporating two positively weighted Gaussians, can describe such behaviour. A supergaussian model was used to model the scans in November 2012 but this model cannot be integrated analytically.

The triple Gaussian model incorporates another weight factor, w_b, in the model: Gaussian 1 is weighted by w, Gaussian 2 by $(1-w)w_b$ and Gaussian 3 by $(1-w)(1-w_b)$. The w_b weight is encouraged to be negative so that Gaussian 2 is negatively weighted. This behaviour allows for good description of the beam spot widths in 2015 and 2016. In addition to weights, the x-y correlation, κ, is also incorporated into the fits. Due to statistical accuracy, only κ_1 (from Gaussian 1) is allowed to float in the triple Gaussian fit. Gaussian 1 is designed to roughly describe the central bulk of the proton density, with Gaussian 2 being negatively weighted to describe transverse width fluctuations and Gaussian 3 the broader Gaussian helping to describe the tails of the beams.

One can also incorporate crossing angles into the model, in x-z and y-z. All scans in 2015 and 2016 were performed with zero crossing angle and so these parameters are set to zero in the model.

6.4.7 Beam parameter determination

The single beam parameters are extracted from a χ^2 minimisation procedure. The χ^2 measure quantifies the difference of the data with the model, weighted by the statistical uncertainty of the data for all scan steps, i,

$$\chi^2 = \sum_i \left(\frac{\text{data}_i - \text{model}_i}{\sigma_{\text{data,i}}} \right). \tag{6.33}$$

χ^2 is calculated for all beam spot and rate observables for all scan steps in a pair of horizontal and vertical scans. The simulated observables in the model are calculated from the three dimensional time integral discussed in Section 6.4.5 above.

From the single beam parameters a non-factorisation correction, as determined from the luminous region fit method, is calculated and denoted R. The correction is defined as the ratio of the *true* luminosity, $\mathcal{L}_{\text{true}}$, with respect to the *factorised* luminosity, $\mathcal{L}_{\text{factorised}}$,

$$R = \frac{\mathcal{L}_{\text{true}}}{\mathcal{L}_{\text{factorised}}} = \frac{\int \rho_1(x,y)\rho_2(x,y)dxdy}{\int \rho_1(x)\rho_2(x)dx \int \rho_1(y)\rho_2(y)dy}. \tag{6.34}$$

The true luminosity refers to the luminosity determined from the triple Gaussian model, which aims to fully describe each proton beam and be independent of factorised effects. The factorised luminosity is determined by measuring the convolved beam widths, Σ_x and Σ_y, from the horizontal and vertical triple Gaussian luminosity profiles and using the standard luminosity calibration formula (Equation 6.11), assuming factorisation of the beams. Both values of luminosity in the ratio are purely calculated from the simulation to determine the correction. One notices then that the non-factorisation correction to μ_{vis} is $1/R$ and the correction to the luminosity simply R.

6.4.8 Corrections to the data

A number of corrections to the luminosity and beam spot data need to be performed before the luminous region fit.

6.4.8.1 Centring and background corrections

The separation of the beams during a scan are those determined by the LHC machine. In such cases one would expect the maximum of the luminosity peak to correspond to zero separation. However, this is not always the case due to the beams drifting slightly from nominal.

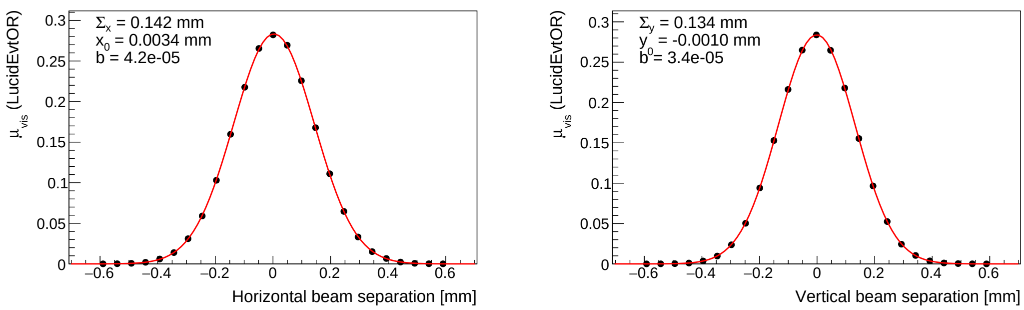
The true zero separation can be determined by taking the value of separation corresponding to the maximum value of the luminosity rate, extracted from a preliminary fit to the luminosity rate only. The same fit can be used to determine a background term, b, which is modelled as a constant. The product of a Gaussian and fourth order polynomial function is used, and is fit for both horizontal and vertical separations:

$$GP_4 = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(x-x_0)^2}{2\sigma^2}} (1 + c_2 x^2 + c_4 x^4) + b. \tag{6.35}$$

Here x represents the beam separation, x_0 the separation at the central (maximum) rate and σ the width of the Gaussian. Only even terms are used in the fourth order polynomial. One can also extract a preliminary estimate of the convolved beam width from this function:

$$\Sigma_x = \sigma(1 + c_2 + 3c_4) \tag{6.36}$$

The background correction is subtracted from the rate data and the centring correction from the separation. An example of such fits for both horizontal and vertical scans in a single BCID in 2016 is shown in Figure 6.2.



(a) (b)

Figure 6.2: LucidEvtOR luminosity profiles fit with a GP_4 function for the (a) horizontal and (b) vertical scan for BCID 1, scan I in 2016. The convolved beam widths, Σ, background terms, b, and centring corrections, x_0, y_0, are also shown.

6.4.8.2 Length scale correction

The separation as tuned by the LHC using the vdM bumps during a scan may not be the same as that ATLAS measures. This difference, or length scale, is determined by taking the ratio of the beam displacement measured by ATLAS using the average position of the luminous region to the nominal displacement entered into the accelerator control system [85]. This length scale is determined in so-called length scale scans, where each of the four bump amplitudes are calibrated and combined to define factors for horizontal and vertical separation scales.

Differences in length scale result in changing the widths of the beam: if the length scale is less than unity then the beams are less wide than in the accelerator units and vice-versa for length scale greater than unity. This results in different measurements of convolved beam width, and thus σ_{vis}. Systematic uncertainties relating to length scale are applied to the luminosity calibration.

A length scale correction is applied in the luminous region fits by scaling the values of the separation by the determined length scale product in each direction. The values used in 2015 and 2016 are shown in Table 6.1.

	Length scale product correction		Total uncertainty on correction	
	Horizontal	Vertical	Horizontal	Vertical
2015	1.0015	1.0005	± 0.0006	± 0.0006
2016	1.0020	1.0010	± 0.0010	± 0.0025

Table 6.1: The horizontal and vertical length scale correction products in 2015 and 2016, with total uncertainty.

6.4.8.3 Orbit drift correction

A transverse drift of individual beam orbits at the interaction point (IP) can affect the overlap integral of the beams and modify the non-factorisation calculated. A number of beam position monitors (BPMs) either side of the IP in the arcs of the LHC are used to measure this effect. The

BPMs are situated outside of the vdM steering dipole magnets (*bumps*) used to conduct a vdM scan and thus are agnostic to the local vdM beam separation.

The BPM system consists of an array of electrodes that measure the electric field of the beams in the arcs of the LHC ($|z| > 500$ m). The measurements allow for an extrapolation of the position of the particle beam at the IP [95]. This extrapolation is done for each beam in the horizontal and vertical directions. From these measurements a correction to the separation of the beams due to orbit drift can be applied.

Figure 6.3 shows the raw orbit drift measurements from the BPMs and a quadratic fit to the scans in the 2016 vdM session. The quadratic fit is that applied as a correction to the separation. The position of each of the beams horizontally and vertically are fixed to zero at the start of the scan and the quadratic function is fit for each scan; both horizontal and vertical. Corrections consisting of the difference of the two beam positions horizontally and vertically are applied to the separation values at each scan step determined from the vdM bumps.

In addition to modifying the separation, the absolute value of the beam spot position may change due to orbit drift. The centroid position of the beam spot changes by the average of the positions of the beams horizontally and vertically. This correction is applied to the transverse positions of the beam spot per scan step, as per the orbit drift separation correction.

6.4.8.4 Beam-beam deflection correction

Both proton beams are positively charged. When coming into collision slightly off-centre each beam receives an electromagnetic *kick* from the other beam that causes an orbit shift and distorts the nominal separation. The Basseti-Erskine formula [96] is used to analytically calculate this orbit shift using electrodynamics. This formula uses the populations of the opposing beam, the horizontal and vertical convolved beam sizes, β^* and the betatron tune. β^* typifies how *focused* the beams are at the interaction point and the betatron tune is defined as the number of betatron oscillations per revolution in a circular accelerator.

From this analytical calculation the separation at each scan point is thus modified. Typical values of beam-beam deflection in scan I, 2016 is shown in Figure 6.4 below. Maximum deflection is seen at a separation of approximately 0.2 mm. The deflection is smaller the closer to head-on the beams become and the further away they are where the electromagnetic field strength of the opposing beam is lessened.

In an offset scan, where the beams are already offset at roughly 0.3 mm from one another, the situation is similar to the in-plane corrections. However the corrections are smaller due to the beams being further away. In the offset scan there are also beam-beam deflections in the plane transverse to the scanning direction. Both cases are shown in Figure 6.5. Due to this offset, a maximum correction is seen in the out-of-plane separation when the beams are aligned in the scanning plane due to the offset in the transverse. The total separation between beam centroids is increased at any other separation value. A very simple validation of the procedure can thus be performed by comparing the beam-beam deflection at a beam separation of 0.3 mm in the centred scan, to that of no separation in the offset scan. This corresponds to the same area of the beam overlap and give very similar values when comparing Figure 6.4 with Figure 6.5.

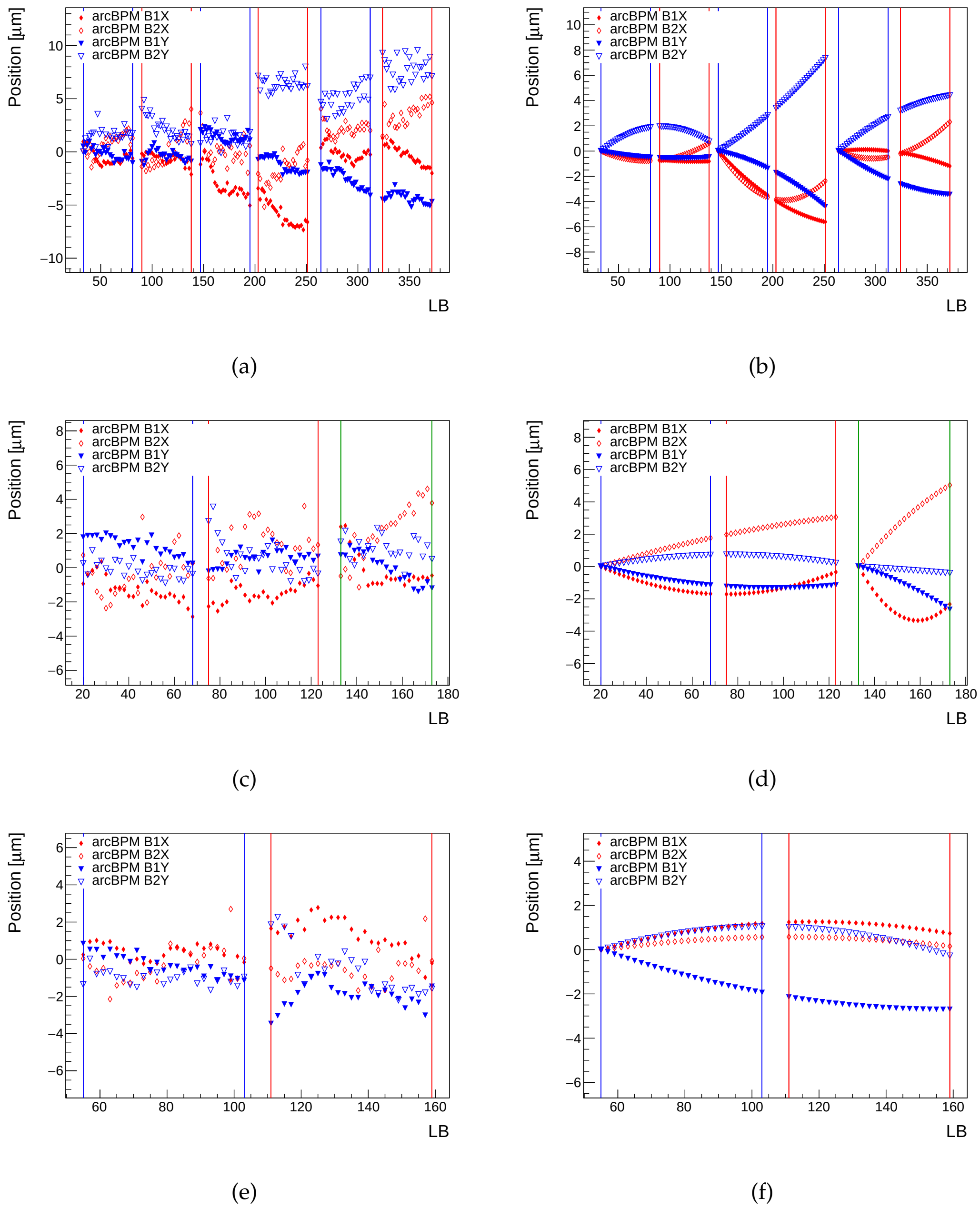
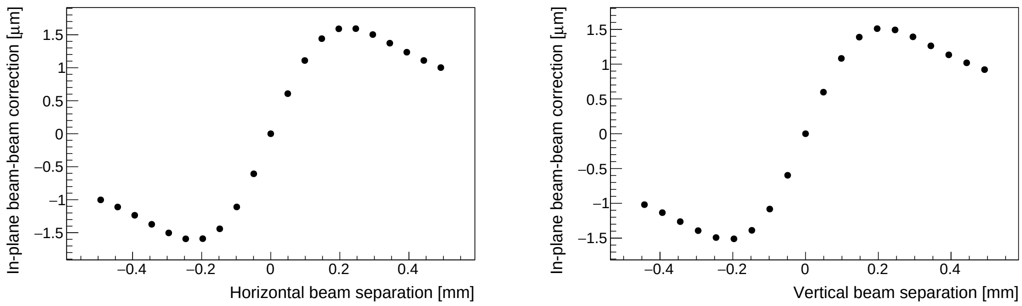


Figure 6.3: Orbit drift measurements from the BPMs for beam 1 and 2, horizontally and vertically from raw data (left) and a quadratic fit (right) in (a, b) scans I, II and III, (c, d) scans IV and V, and (e, f) scan VI in 2016. The blue lines represent a horizontal scan and the red lines a vertical scan. The single diagonal scan V is shown between green lines. Scans I, II and III and scans IV and V are run one after another in individual fills. LB (lumiblock) refers to the scan step.



(a) (b)

Figure 6.4: The beam-beam deflection corrections to the separation in the (a) horizontal and (b) vertical scans in BCID 1, scan I, 2016.

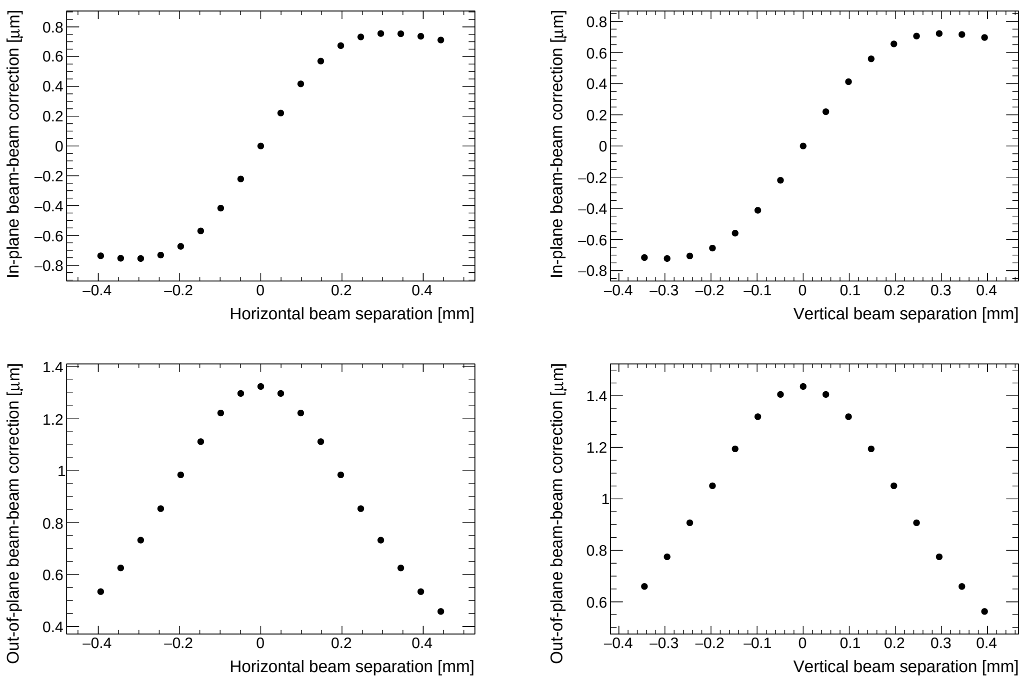


Figure 6.5: The in-plane (top) and out-of-plane (bottom) beam-beam deflection corrections to the separation in the horizontal (left) and vertical (right) scans in BCID 872, scan III, 2016.

6.4.9 Luminous region results in 2015

The results of the triple Gaussian fits in the 2015 scan sessions are discussed below. There were four scans performed in two fills in August. They are labelled I to IV.

Scans I, II and III were all performed in the same fill (Fill 4266, Run 277025, 24th-25th August). Scans I and II are centred scans and scan III is an offset scan, performed one after the other. Another centred scan IV was performed in another fill a day later, for longer term reproducibility of calibration results (Fill 4269, Run Run 277089, 26th August).

In the first three scans, five BCIDs have associated beam spot information allowing for a luminous region analysis: BCID 51, BCID 891, BCID, 1631, BCID 2451 and BCID 2674. Scan IV had a different BCID structure with four BCIDs containing beam spot information: BCID 491, BCID 571, BCID 1783 and BCID 1903.

The centred scan results are discussed in Section 6.4.9.1 and the offset scan in Section 6.4.9.2.

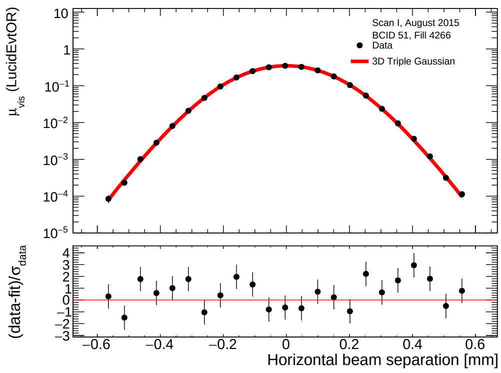
6.4.9.1 Centred scans

The single beam parameters are determined from the χ^2 minimisation procedure discussed above. The beam spot and rate observables are then determined from these simulated parameters and compared to data. This comparison for scan I, BCID 51 in 2015 is shown for the horizontal scan in Figure 6.6 and the vertical scan in Figure 6.7. A summary of the single beam parameters are shown in Table 6.2.

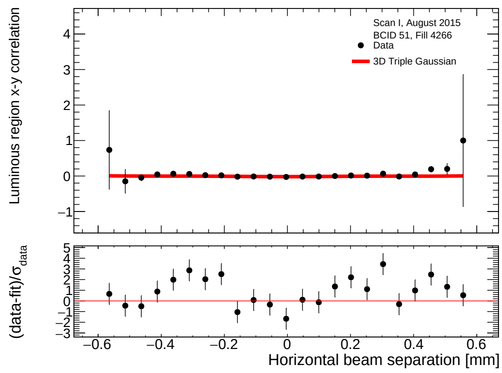
The most interesting distributions are the transverse beam spot positions and widths. If one was to see a change in horizontal beam spot width in the vertical scan (or vice-versa) then this would indicate non-factorisation effects in the beams. The wavy shape of the beam spot widths as a function of separation that requires a triple Gaussian model for best representation can be seen. The non-linearity of the beam spot positions in the data are can also be seen, and how well the triple Gaussian models these positions. Typically x-y correlation, longitudinal beam spot width and position are approximately flat with respect to separation. If one had a non-zero crossing angle then one might expect a drift of the longitudinal beam spot position from nominal.

A common problem with the triple Gaussian model in 2015 is that the luminous widths in the same scanning direction (horizontal width in the horizontal scan, vertical width in the vertical scan) are slightly overestimated by the model compared to data. This would imply that the individual widths of the single beams are also overestimated by the model. Pre-corrections are applied that result in an increase in the width of the beams in the model: a length scale greater than unity and beam-beam deflection. In fact, if both of these corrections are not applied the χ^2 per degree of freedom of the fit decreases and the transverse luminous widths are better modelled. This problem persists into 2016, discussed in Section 6.4.10.1 below.

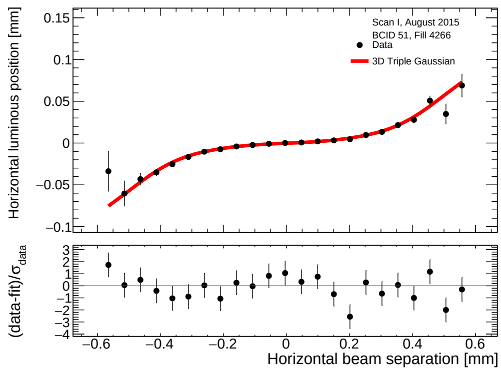
Other BCIDs in scan I and the other centred scans show similar profiles to BCID 51, shown here. A summary of the values of R and χ^2 per degree of freedom for each BCID in the centred scans in 2015 is shown in Table 6.3. The χ^2 per degree of freedom for the fits to scan IV are much worse than for scans I and II. There were large orbit drifts in scan IV and the corresponding correction is thought to not fully capture the effect. Nevertheless, the value of R for each BCID does not differ much than that of the first scans.



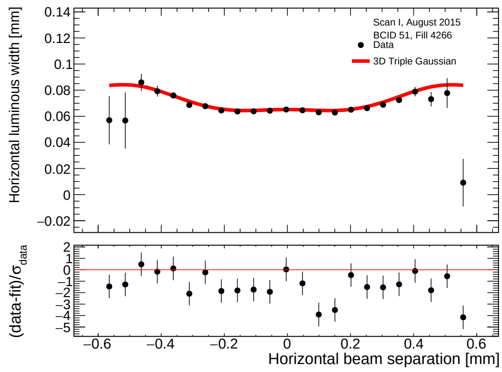
(a)



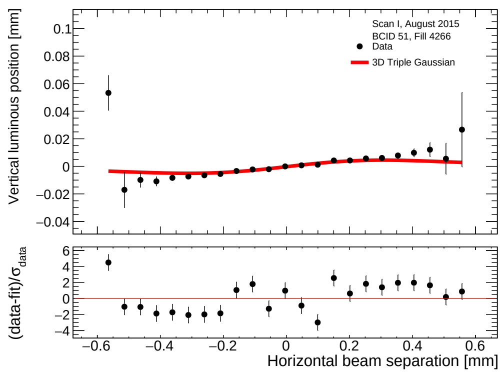
(b)



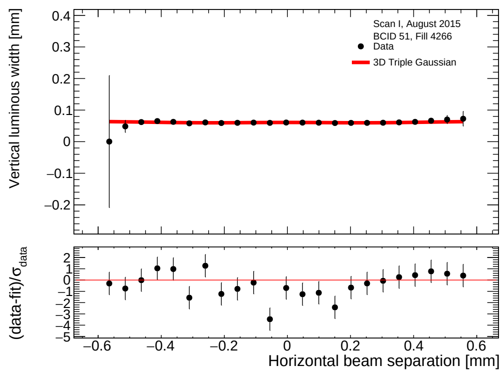
(c)



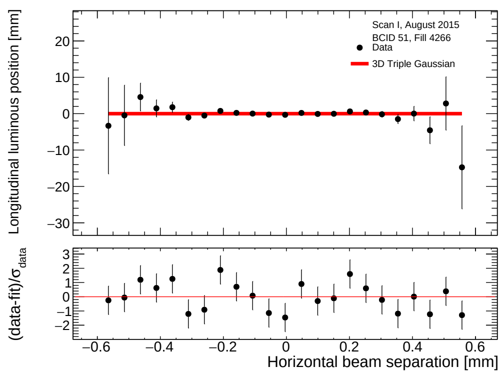
(d)



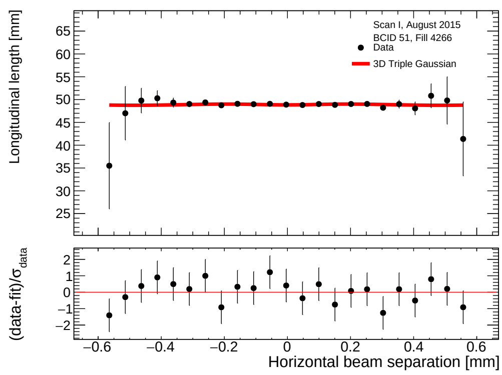
(e)



(f)



(g)



(h)

Figure 6.6: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the horizontal scan, BCID 51, scan I, 2015.

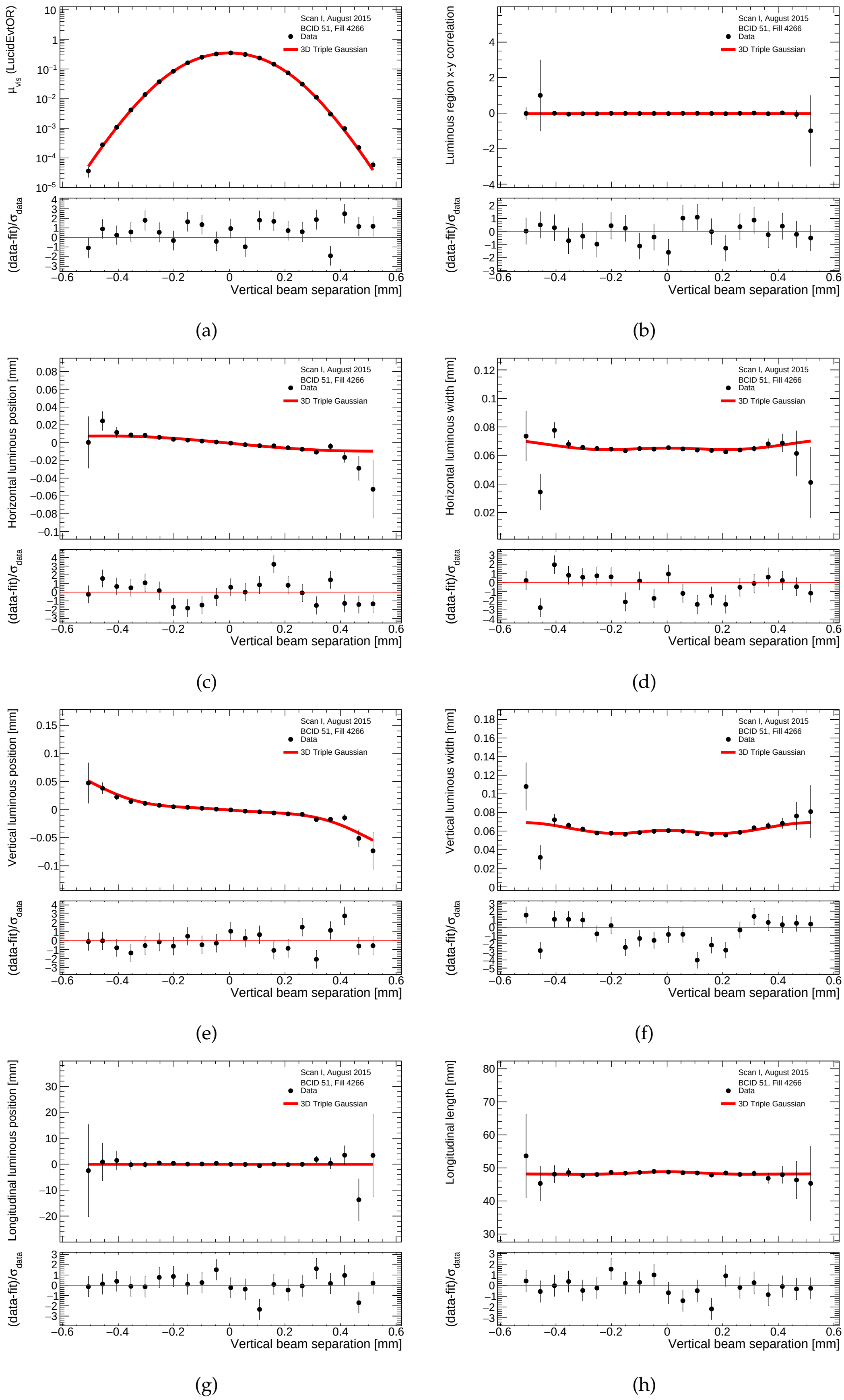


Figure 6.7: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the vertical scan, BCID 51, scan I, 2015.

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.085	0.074
$\sigma_{x,2}$ [mm]	0.059	0.030
$\sigma_{x,3}$ [mm]	0.125	0.094
$\sigma_{y,1}$ [mm]	0.079	0.077
$\sigma_{y,2}$ [mm]	0.034	0.042
$\sigma_{y,3}$ [mm]	0.106	0.081
$\sigma_{z,1}$ [mm]	90.5	4.9
$\sigma_{z,2}$ [mm]	33.6	86.0
$\sigma_{z,3}$ [mm]	89.5	42.5
κ_1	-0.043	0.170
w	0.852	0.379
w_b	-0.474	-0.084
χ^2 / d.o.f	576.6 / 322 = 1.791	
R	10.191 / 10.183 = 1.001	

Table 6.2: The single beam parameters and fit results from the triple Gaussian fit to BCID 51, scan I, 2015.

	Scan I				
	BCID 51	BCID 891	BCID 1631	BCID 2451	BCID 2674
R	1.001	0.997	1.003	0.999	1.003
χ^2/d.o.f	1.79	2.21	1.53	2.16	2.10

	Scan II				
	BCID 51	BCID 891	BCID 1631	BCID 2451	BCID 2674
R	1.002	0.998	1.003	0.999	1.001
χ^2/d.o.f	1.89	1.72	1.86	1.71	2.47

	Scan IV			
	BCID 491	BCID 571	BCID 1783	BCID 1903
R	1.002	1.004	1.002	0.998
χ^2/d.o.f	2.71	3.29	3.27	4.29

Table 6.3: Values of R and χ^2 per degree of freedom for the centred scans I (top), II (middle) and IV (bottom) in 2015.

6.4.9.2 Offset scan

The offset scan III provides information about the overlap of the edges of the beams. A displacement of $\sim$300 μm is made in the plane orthogonal to the scanning direction in each fill. The offset scan was performed directly after scan II in the same fill and hence the same beam parameters should describe scan III. Scan III can be used to constrain the beam parameters deduced purely from centred scan data by running a combined fit of the two scan sessions. This simply extends the χ^2 minimisation procedure over two scans instead of one.

The results for BCID 51 in the offset scan (in a combined fit with scan II) are shown for the horizontal scan in Figure 6.8 and the vertical scan in Figure 6.9. The agreement between the triple Gaussian model and the data is reasonable.

The comparison of the beam parameters in the centred only scan with the combined fit are shown in Table 6.4. The value of R decreases by 0.001 from the centred only result and the χ^2 per degree of freedom of the fit is slightly higher. Similar results for the four other BCIDs with beamspot data are seen.

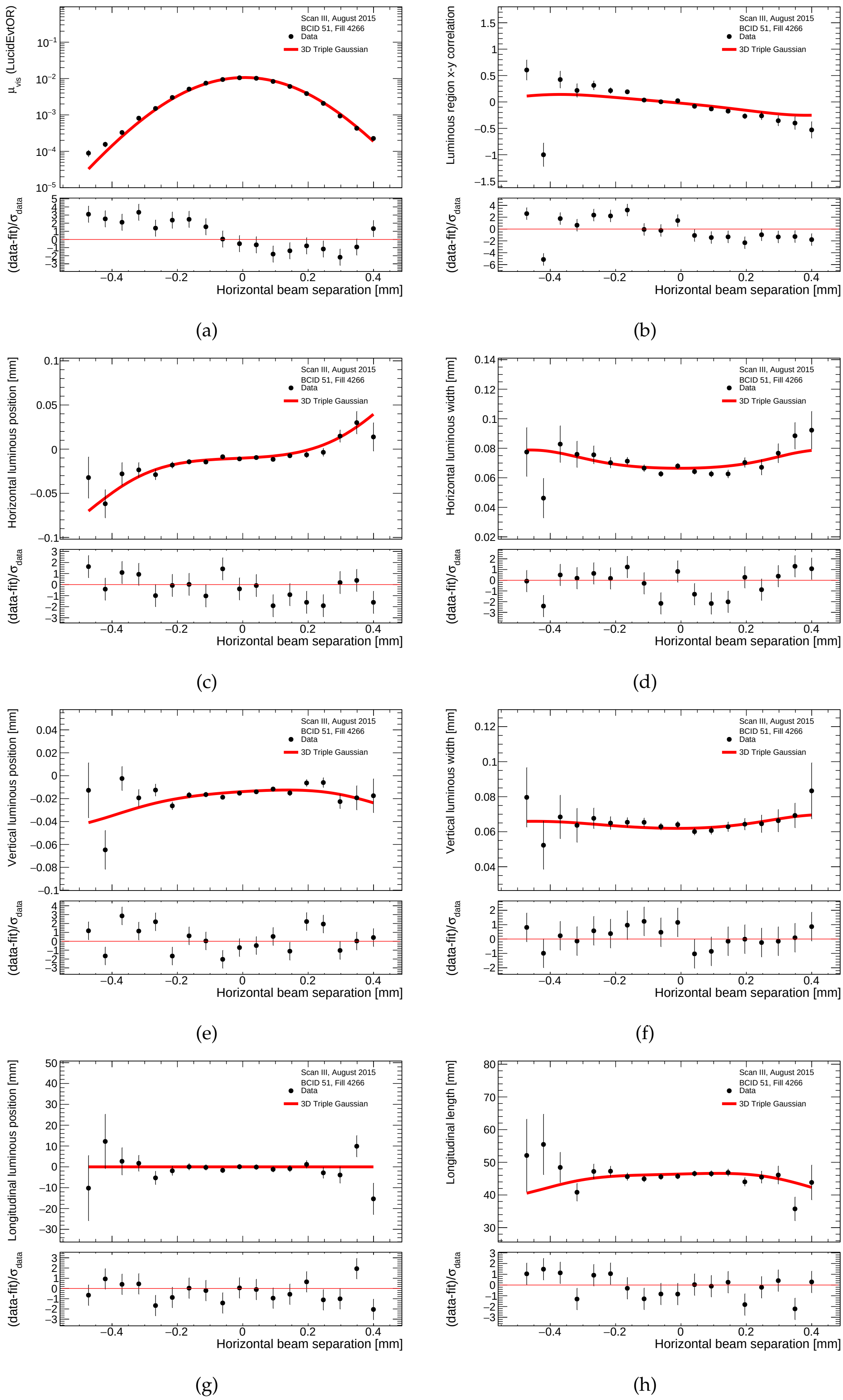


Figure 6.8: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the horizontal offset scan, BCID 51, scan III, 2015.

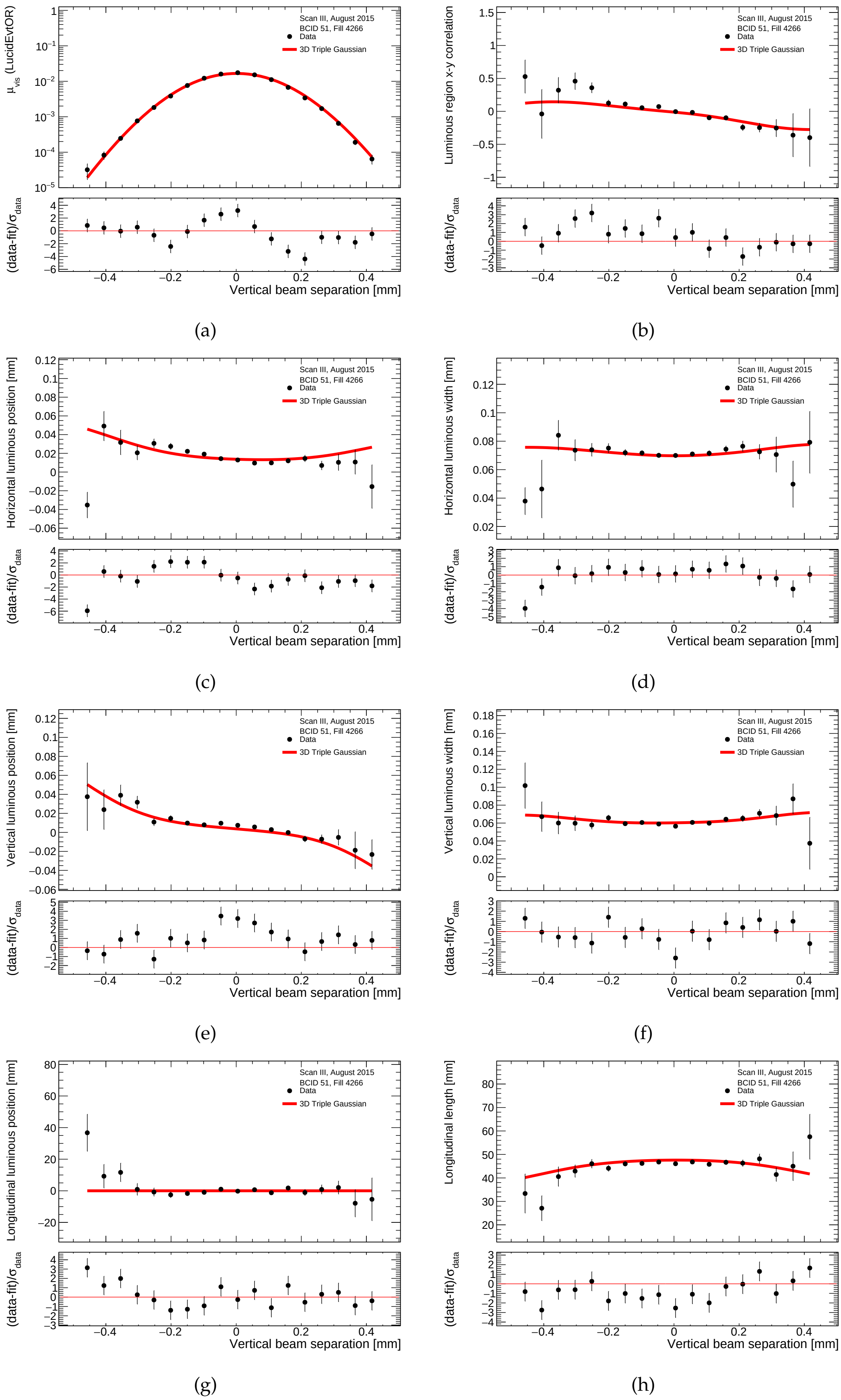


Figure 6.9: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the vertical offset scan, BCID 51, scan III, 2015.

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.084	0.081
$\sigma_{x,2}$ [mm]	0.048	0.046
$\sigma_{x,3}$ [mm]	0.120	0.103
$\sigma_{y,1}$ [mm]	0.078	0.076
$\sigma_{y,2}$ [mm]	0.022	0.055
$\sigma_{y,3}$ [mm]	0.103	0.089
$\sigma_{z,1}$ [mm]	91.9	0.5
$\sigma_{z,2}$ [mm]	18.9	95.2
$\sigma_{z,3}$ [mm]	39.3	113.1
κ_1	-0.049	0.078
w	0.796	0.868
w_b	-0.296	-0.757
χ^2 / d.o.f	608.9 / 322 = 1.891	
R	10.206 / 10.189 = 1.002	

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.087	0.084
$\sigma_{x,2}$ [mm]	0.041	0.080
$\sigma_{x,3}$ [mm]	0.130	0.086
$\sigma_{y,1}$ [mm]	0.080	0.083
$\sigma_{y,2}$ [mm]	0.018	0.078
$\sigma_{y,3}$ [mm]	0.109	0.080
$\sigma_{z,1}$ [mm]	83.8	66.1
$\sigma_{z,2}$ [mm]	66.2	107.1
$\sigma_{z,3}$ [mm]	40.8	80.5
κ_1	-0.054	0.400
w	0.902	0.154
w_b	-0.514	-0.799
χ^2 / d.o.f	1317.9 / 608 = 2.168	
R	10.186 / 10.180 = 1.001	

Table 6.4: The single beam parameters and fit results from the triple Gaussian fit to BCID 51 in scan II (left) and the combined fit of scans II and III (right) in 2015.

Two extra parameters are included in the combined offset fit: the orthogonal offset values for each scan. These parameters are allowed to float in the fit due to large discrepancies whilst using the values dialled in by the LHC. Both the floating values and nominal are shown for each BCID in Figure 6.10. The floating horizontal offsets in the vertical scan differ from LHC nominal by approximately 15 μm, but the vertical offsets agree well with nominal. The LHC machine applied a 10 μm in-plane centering correction before the horizontal scan which is not accounted for in the 315 μm horizontal offset value quoted for the vertical scan. If one compares a 325 μm horizontal offset with the fitted values then the agreement is within a few microns.

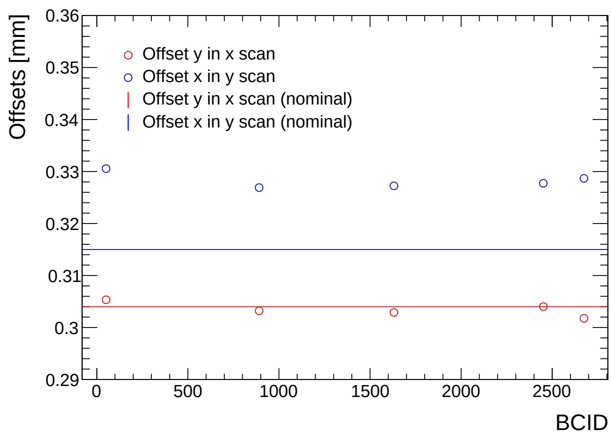


Figure 6.10: Values of the offset as the floating parameter in the fit (open circles) and as tuned by the LHC (solid lines) in the offset scan III in 2015.

A summary of the values of R and χ^2 per degree of freedom for each BCID is shown in Table 6.5.

6.4.10 Luminous region results in 2016

The results of the triple Gaussian fits in the May 2016 scan sessions are discussed below. There were six scans performed in two fills. They are labelled I to VI.

	Scan II & III				
	BCID 51	BCID 891	BCID 1631	BCID 2451	BCID 2674
R	1.001	0.995	1.001	0.993	1.001
χ^2/d.o.f	2.17	2.60	2.02	3.06	2.86

Table 6.5: Values of R and χ^2 per degree of freedom for the combined fit of the centred scan II and the offset scan III in 2015.

Scans I, II, III, IV and V were performed in a single fill (Fill 4945, Run 299390, 18th-19th May). Scans I, II and IV were centred scans. Scan III is an offset scan and Scan V a diagonal scan. A final centred scan VI was performed in another fill (Fill 4954, Run 300287, 27th May). All scans had the same BCID structure, with five BCIDs containing tracking information: BCID 1, BCID 872, BCID, 1152, BCID 1863 and BCID 2934.

The centred scan results are discussed in Section 6.4.10.1, the offset scan in Section 6.4.10.2 and the diagonal scan in Section 6.4.10.3.

6.4.10.1 Centred scans

The results for scan I, BCID 1 in 2016 are shown for the horizontal scan in Figure 6.11 and the vertical scan in Figure 6.12. A summary of the beam parameters is shown in Table 6.6.

The beam structures in 2016 are quite similar to those of 2015. The triple Gaussian model is again needed to describe the beam spot widths in the same plane as the scan separation plane, but the behaviour is harder to spot due to large uncertainties on some beam spot fit points within the tails of the scan. However the model overestimates these widths, again as with 2015. This is discussed further in Section 6.4.11.

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.110	0.093
$\sigma_{x,2}$ [mm]	0.050	0.061
$\sigma_{x,3}$ [mm]	0.057	0.080
$\sigma_{y,1}$ [mm]	0.096	0.081
$\sigma_{y,2}$ [mm]	0.039	0.060
$\sigma_{y,3}$ [mm]	0.062	0.103
$\sigma_{z,1}$ [mm]	35.0	66.6
$\sigma_{z,2}$ [mm]	66.3	22.5
$\sigma_{z,3}$ [mm]	19.9	63.3
κ_1	0.022	0.047
w	0.923	0.763
w_b	-0.800	-0.800
χ^2 / d.o.f	528.2 / 298 = 1.772	
R	8.435 / 8.393 = 1.005	

Table 6.6: The single beam parameters and fit results from the triple Gaussian fit to BCID 1, scan I, 2016.

A summary of the values of R and χ^2 per degree of freedom for each BCID in the centred scans in 2016 is shown in Table 6.7. The quality of the fits are relatively similar between scans and BCIDs, as are the values of R.

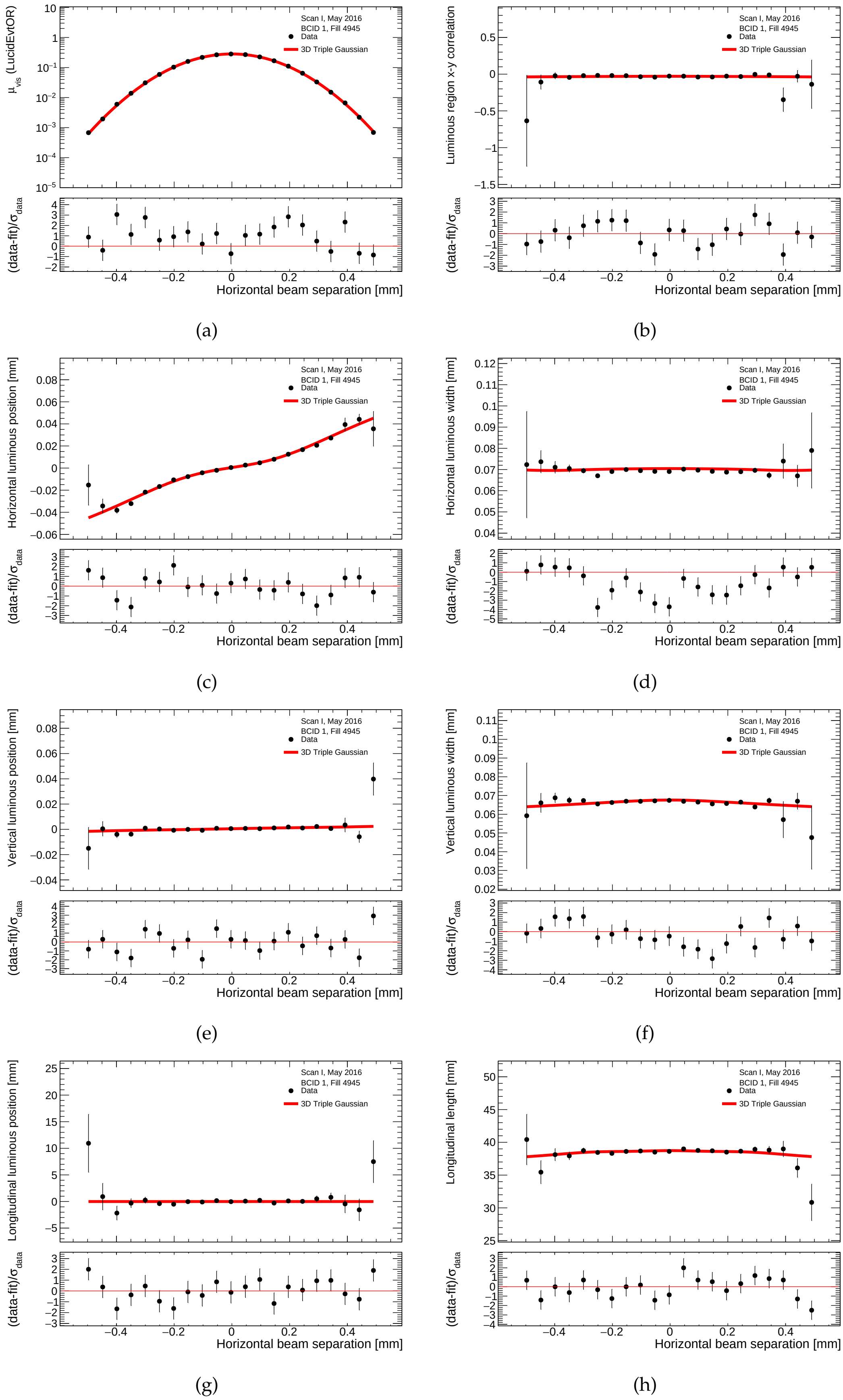


Figure 6.11: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the horizontal scan, BCID 1, scan I, 2016.

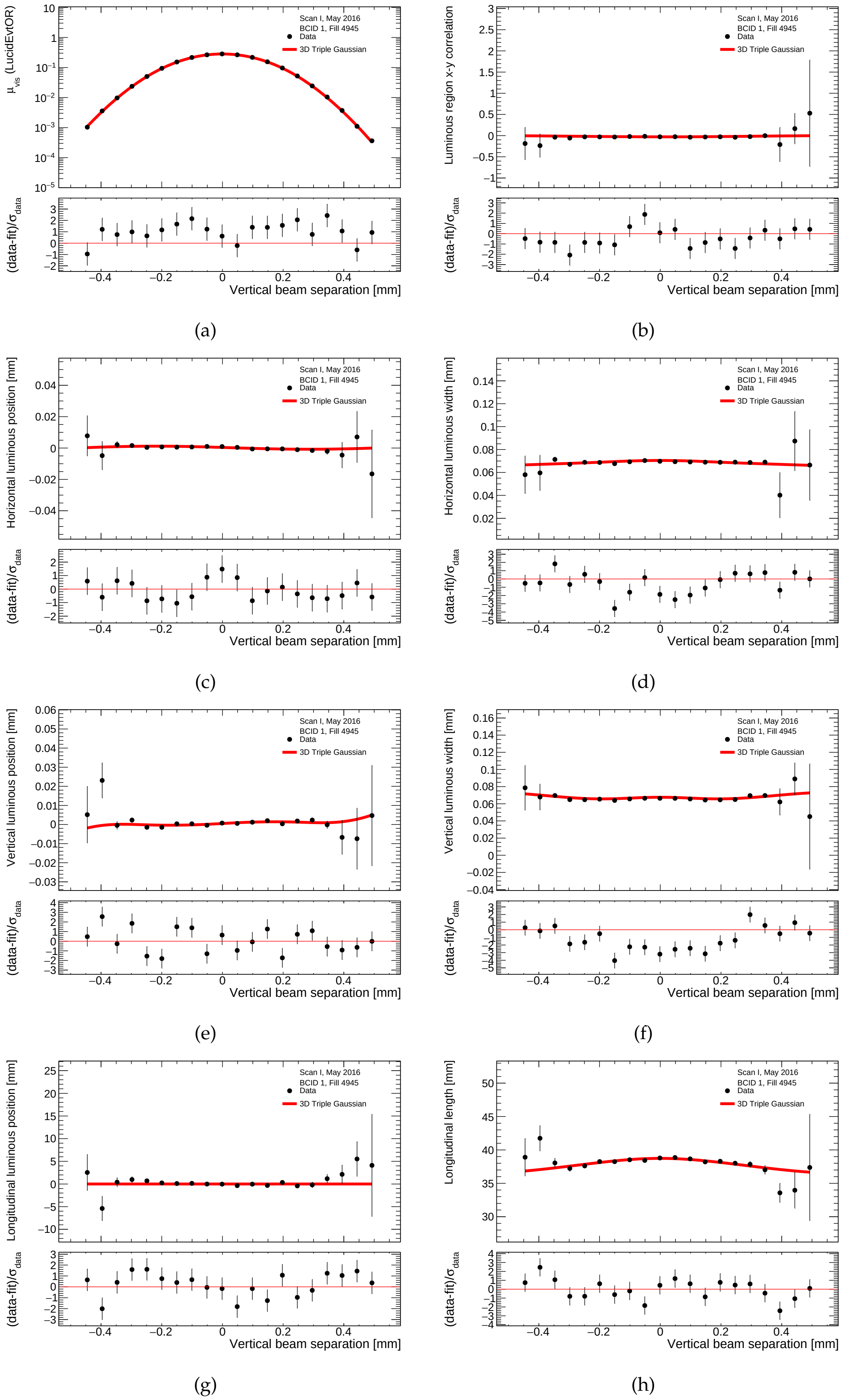


Figure 6.12: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the vertical scan, BCID 1, scan I, 2016.

	Scan I				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.005	1.008	1.013	1.003	1.009
χ^2/d.o.f	1.77	1.77	1.92	1.90	1.88

	Scan II				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.006	1.006	1.008	1.003	1.009
χ^2/d.o.f	1.49	1.32	1.97	1.78	1.92

	Scan IV				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.006	1.009	1.008	1.007	1.010
χ^2/d.o.f	2.15	2.10	2.23	1.94	1.91

	Scan VI				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.014	1.008	1.009	1.008	1.008
χ^2/d.o.f	1.52	1.82	1.71	2.05	2.27

Table 6.7: Values of R and χ^2 per degree of freedom for the centred scans I, II, IV and VI (in descending order) in 2016.

6.4.10.2 Offset scan

The offset scan III in 2016 was performed in combination with Scan II, as in 2015. Again, the model does a reasonable job at modelling the data. A comparison of the single beam parameters from the centred only and combined fits is shown in Table 6.8. The χ^2 per degree of freedom increases compared to centred only and the value of R decreases, getting closer to a value of 1, or no non-factorisation present. This behaviour is common to all BCIDs in the combined offset scan.

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.104	0.091
$\sigma_{x,2}$ [mm]	0.093	0.060
$\sigma_{x,3}$ [mm]	0.130	0.079
$\sigma_{y,1}$ [mm]	0.088	0.077
$\sigma_{y,2}$ [mm]	0.036	0.055
$\sigma_{y,3}$ [mm]	0.118	0.101
$\sigma_{z,1}$ [mm]	27.7	72.9
$\sigma_{z,2}$ [mm]	77.9	56.1
$\sigma_{z,3}$ [mm]	24.4	65.8
κ_1	0.024	0.050
w	0.928	0.744
w_b	-0.499	-0.800
χ^2 / d.o.f	442.6 / 298 = 1.485	
R	8.537 / 8.488 = 1.006	

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.099	0.082
$\sigma_{x,2}$ [mm]	0.072	0.059
$\sigma_{x,3}$ [mm]	0.129	0.098
$\sigma_{y,1}$ [mm]	0.085	0.078
$\sigma_{y,2}$ [mm]	0.046	0.055
$\sigma_{y,3}$ [mm]	0.114	0.104
$\sigma_{z,1}$ [mm]	48.3	64.1
$\sigma_{z,2}$ [mm]	26.3	64.0
$\sigma_{z,3}$ [mm]	43.8	50.7
κ_1	0.008	0.036
w	0.907	0.762
w_b	-0.800	-0.800
χ^2 / d.o.f	1148.4 / 584 = 1.966	
R	8.524 / 8.503 = 1.002	

Table 6.8: The single beam parameters and fit results from the triple Gaussian fit to BCID 1 in scan II (left) and the combined fit of scans II and III (right) in 2016.

Similarly to 2015, two extra floating offset parameters are added to the fit. The agreement with the nominal values tuned by the accelerator are within a few microns and are shown in Figure 6.13.

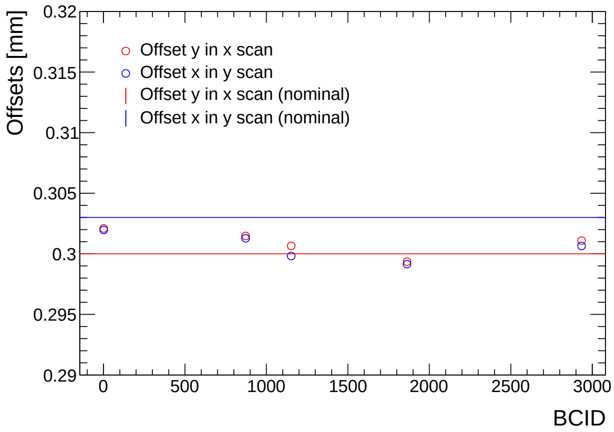


Figure 6.13: Values of the offset as the floating parameter in the fit (open circles) and as tuned by the LHC (solid lines) in the offset scan III in 2016.

The values of R and χ^2 per degree of freedom for each BCID is shown in Table 6.9.

	Scan II & III				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.002	1.005	1.002	1.002	1.004
χ^2/d.o.f	1.97	1.79	2.03	2.00	2.28

Table 6.9: Values of R and χ^2 per degree of freedom for the combined fit of the centred scan II and offset scan III in 2016.

6.4.10.3 Diagonal scan

In 2016 a diagonal scan was performed. This scan consists of 21 scan steps where, in each step, the same change in separation both vertically and horizontally is made. This scan explored parts of the beam that neither the centred nor the offset scans typically explore. Similarly to the offset scan, the diagonal scan can be used to constrain beam parameters determined in a centred scan close in time to the diagonal. In this case the centred scan IV is used which was performed immediately before the diagonal scan V.

The comparison of the triple Gaussian fit with data is shown in Figure 6.14. There is only a single diagonal scan performed, mainly to reduce the time vdM sessions take in ATLAS. The features seen are relatively similar to both the centred and offset scans in 2016 apart from the x-y correlation. The x-y correlation appears to depend on the beam separation much more than in the centred scans, with almost an inverse behaviour with respect to the horizontal and vertical luminous widths. The longitudinal luminous length also decreases at high and low beam separations relative to zero separation, more harshly than the centred scans.

A comparison of the single beam parameters from the centred only and combined diagonal fits is shown in Table 6.10. The χ^2 per degree of freedom is relatively large for the centred only scans and the value increases for the combined fit. The value of R again becomes closer to 1 for the constrained combined fit compared to centred only.

The values of R and χ^2 per degree of freedom for each BCID is shown in Table 6.11.

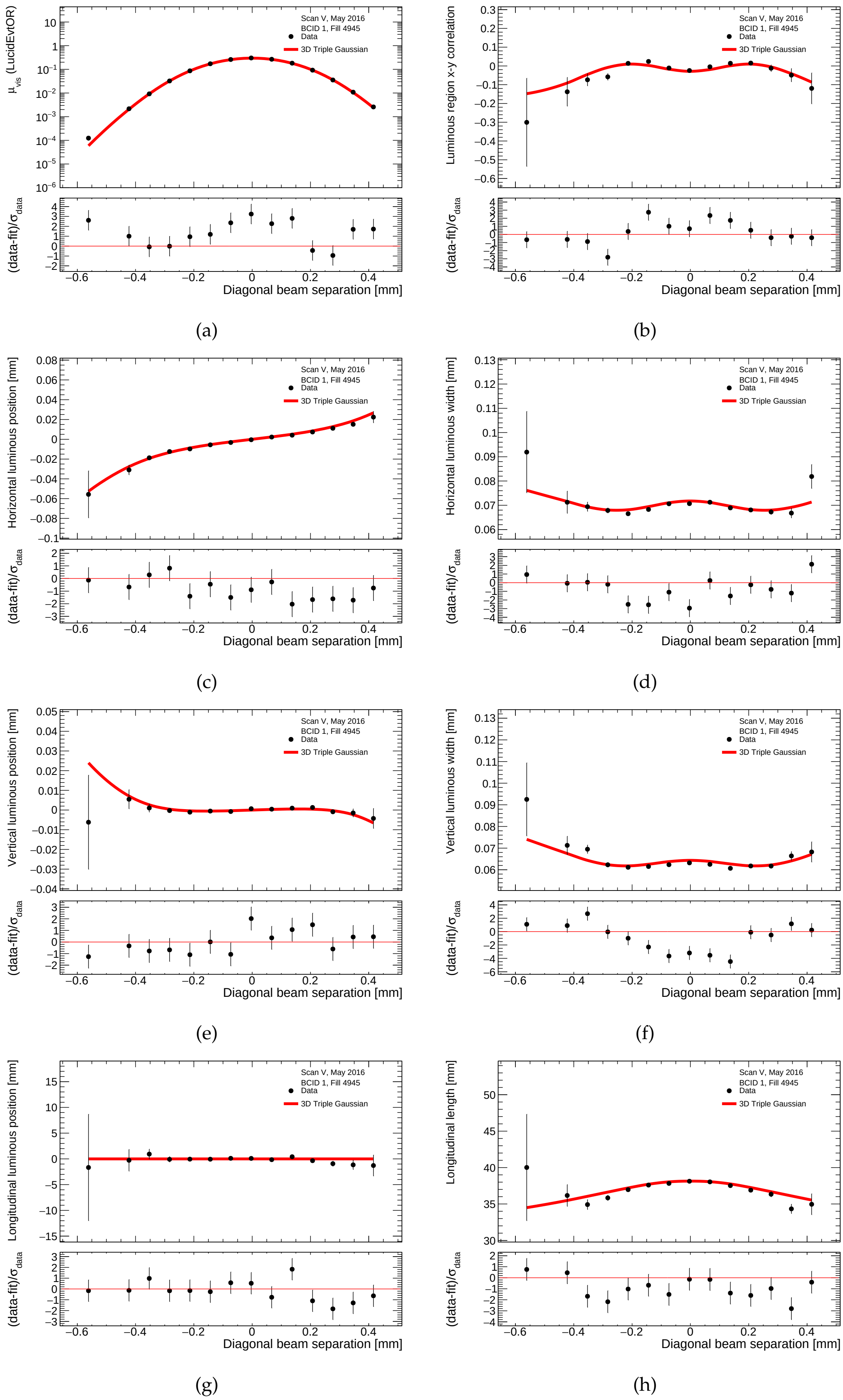


Figure 6.14: Triple Gaussian fit results for (a) μ_{vis} from the LucidEvtOR algorithm, (b) x-y correlation, (c) horizontal position, (d) horizontal width, (e) vertical position, (f) vertical width, (g) longitudinal position and (h) longitudinal width for the diagonal scan, BCID 1, scan V, 2016.

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.099	0.091
$\sigma_{x,2}$ [mm]	0.058	0.047
$\sigma_{x,3}$ [mm]	0.135	0.078
$\sigma_{y,1}$ [mm]	0.083	0.081
$\sigma_{y,2}$ [mm]	0.039	0.061
$\sigma_{y,3}$ [mm]	0.113	0.094
$\sigma_{z,1}$ [mm]	53.0	60.2
$\sigma_{z,2}$ [mm]	10.4	50.3
$\sigma_{z,3}$ [mm]	30.6	36.1
κ_1	0.033	0.034
w	0.926	0.782
w_b	-0.800	-0.745
χ^2 / d.o.f	656.9 / 306 = 2.147	
R	8.890 / 8.838 = 1.006	

Parameter	Beam 1	Beam 2
$\sigma_{x,1}$ [mm]	0.096	0.087
$\sigma_{x,2}$ [mm]	0.057	0.045
$\sigma_{x,3}$ [mm]	0.132	0.094
$\sigma_{y,1}$ [mm]	0.081	0.081
$\sigma_{y,2}$ [mm]	0.045	0.052
$\sigma_{y,3}$ [mm]	0.111	0.100
$\sigma_{z,1}$ [mm]	38.1	71.1
$\sigma_{z,2}$ [mm]	17.2	56.4
$\sigma_{z,3}$ [mm]	41.1	18.8
κ_1	0.030	0.032
w	0.899	0.849
w_b	-0.800	-0.761
χ^2 / d.o.f	947.6 / 417 = 2.272	
R	8.895 / 8.862 = 1.004	

Table 6.10: The single beam parameters and fit results from the triple Gaussian fit to BCID 1 in the centred scan IV (left) and the combined fit of scans IV and the diagonal scan V (right) in 2016.

	Scan IV & V				
	BCID 1	BCID 872	BCID 1152	BCID 1863	BCID 2934
R	1.004	1.006	1.006	1.004	1.008
χ^2/d.o.f	2.27	2.46	2.66	2.07	2.19

Table 6.11: Values of R and χ^2 per degree of freedom for the combined fit of the centred scan IV and diagonal scan V in 2016.

6.4.11 Statistical and systematic uncertainties

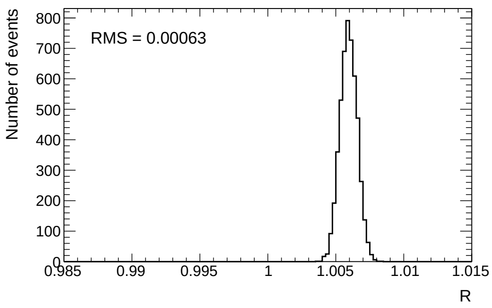
The value of R quoted for each BCID has some uncertainty associated with it, both statistical and systematic. All the uncertainties applied to R are discussed.

6.4.11.1 Statistical uncertainty on R

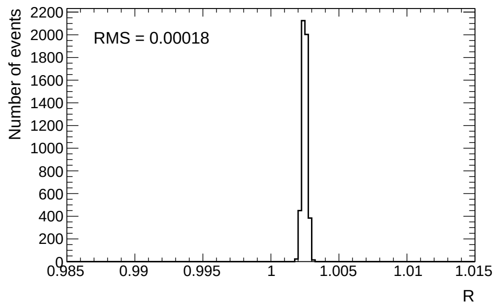
The luminosity and beam spot data used to determine the value of R contain statistical uncertainties. The triple Gaussian fit for centred scans typically contains 30 free floating parameters, with the value of each parameter containing an uncertainty from the fit. One method to determine the statistical uncertainty on R would be to produce a large number of toy datasets from the original data, modified according to the statistical uncertainty at each data point. A value of R could be calculated from fits of each toy dataset and the distribution of the values of R could be used to estimate the statistical uncertainty. However, this would be very computationally expensive since each fit takes of the order of 10 s to run nominally for each BCID.

Another, less computationally expensive, way of estimating the statistical uncertainty on R is by randomly sampling a multivariate Gaussian, with mean vector corresponding to the minimised beam parameters and covariance matrix that returned from the fitting procedure. The covariance matrix contains all correlations between parameters in the fit. Randomly sampling this Gaussian results in modified beam parameters which can then be used to calculate a new value of R. The Gaussian is typically sampled 5000 times to ensure an accurate estimate of the statistical uncertainty on R, which is taken as the root mean square (RMS) of the resulting distribution of R.

An example uncertainty is shown for a BCID in scan II, 2016 and the same BCID in the combined fit of scan II and III, 2016 in Figure 6.15. The increased statistical accuracy from the combined fit of the centred scan and the offset scan can be seen compared to the centred scan.



(a)



(b)

Figure 6.15: Statistical uncertainties for BCID 1 in (a) scan II and (b) the combined fit of scan II and III in 2016.

6.4.11.2 Vertex resolution uncertainty

As mentioned in Section 6.4.2, the three dimensional single Gaussian beam spot fit takes into account the difference between the actual and expected vertex resolution by a free floating parameter, k. This parameter, with corresponding uncertainty, can vary wildly, especially in the tails of the scans where there are a low number of vertices. The effect of changing vertex resolution on the quoted value of R is measured in scans I, II and III in the 2016 session.

The values of k with reasonable uncertainties are shown for all BCIDs in scans I, II and III in Figure 6.16. From these distributions, three choices of fixed k were chosen for new beam spot fits to determine a vertex resolution uncertainty on R: $k = 1.14$, $k = 1.16$ and $k = 1.18$.

In Figure 6.17 below, the values of R for each of the fixed k fits and the nominal floating k fit are shown for scans I, II and the combined fit of scans II and III. A maximal range of R per BCID for the different fits is of the order of 0.2% and so a conservative $\pm 0.1\%$ systematic uncertainty is applied to account for vertex resolution per BCID. This same uncertainty is applied to individual BCIDs in the 2015 session, under the assumption that the vertex resolution range is at a similar level.

6.4.11.3 Length scale uncertainty

The length scale correction is itself applied nominally to the separations in the luminous region fit. An uncertainty on the value of R due to this correction is taken by running fits with the nominal length scale plus upper length scale uncertainty and nominal minus lower uncertainty. The maximum difference in R from nominal for the two fits is taken as the length scale uncertainty on R. Typical values of R for the three different fits (including nominal correction) are shown for scans I and II in 2016 in Figure 6.18. Ranges are typically of the order of 0.1%.

Interestingly, the χ^2 per degree of freedom of the fits with nominal length scale minus lower uncertainty are much improved compared to nominal, as shown in Figure 6.19. This alludes back to the model overestimating the horizontal and vertical luminous widths. Similarly, the χ^2

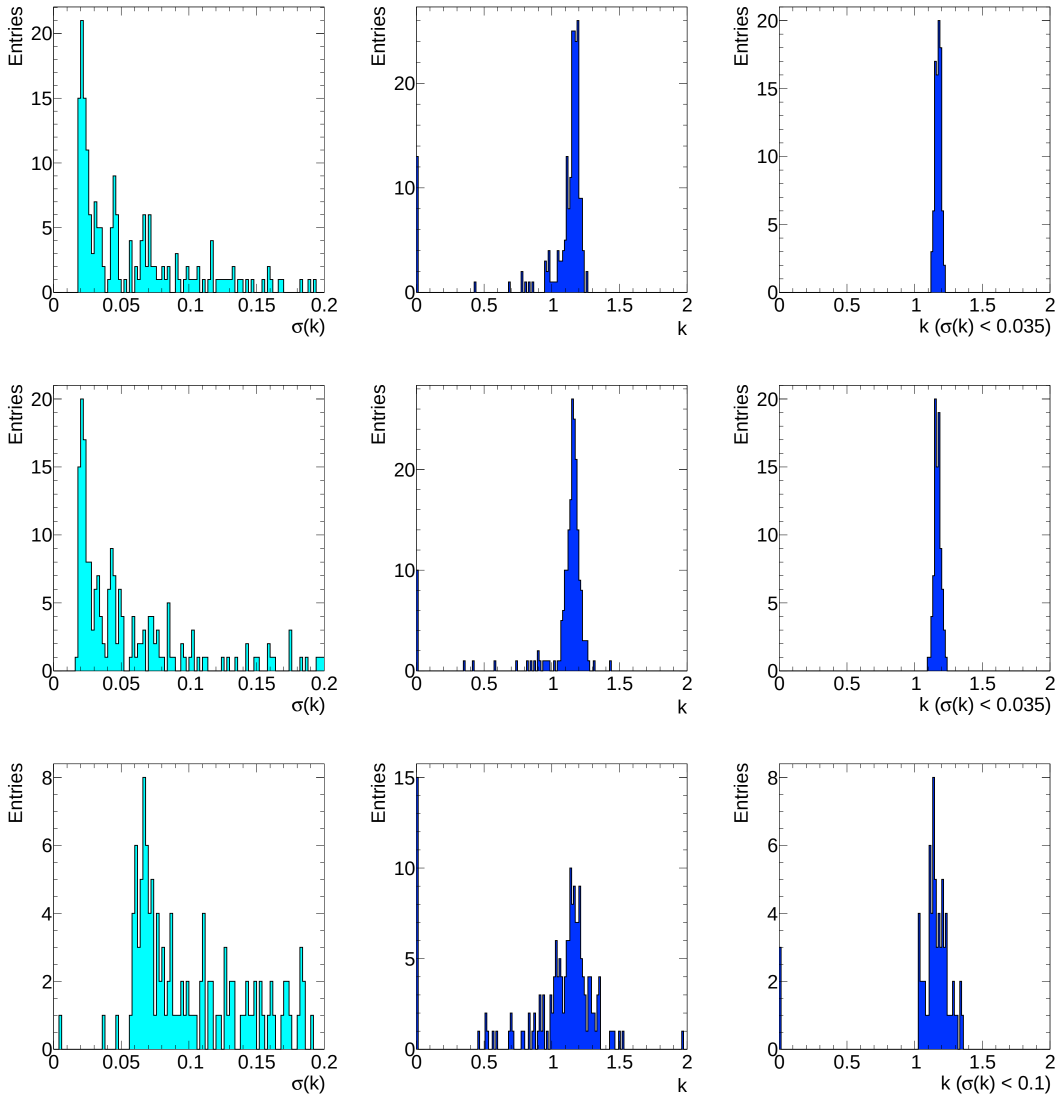
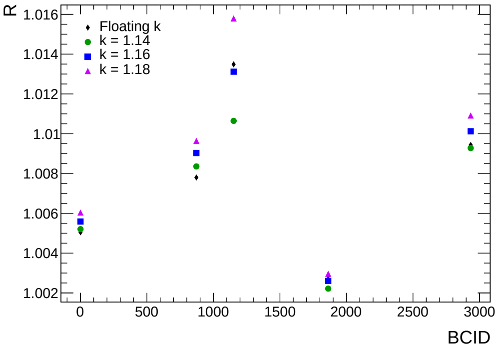
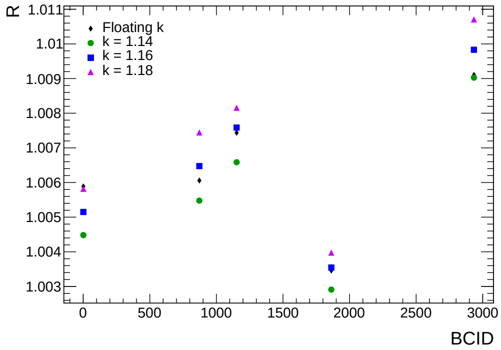


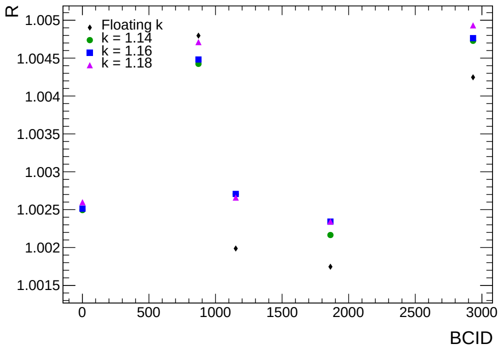
Figure 6.16: Uncertainties (left) and values (middle) of the vertex resolution factor k for scan I (top), scan II (middle) and scan III (bottom). The vertex resolution factors limited by uncertainty, $\sigma(k)$, (right) are also quoted for $\sigma(k) < 0.035$ to remove outlying values for the centred scans and $\sigma(k) < 0.1$ for the offset scan III.



(a)

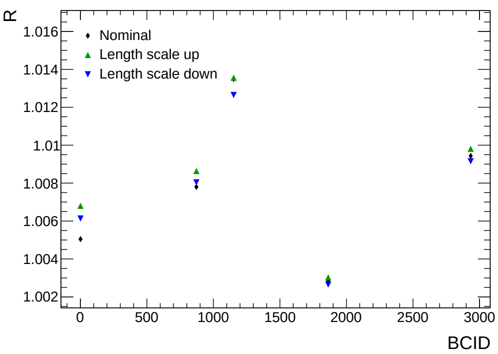


(b)

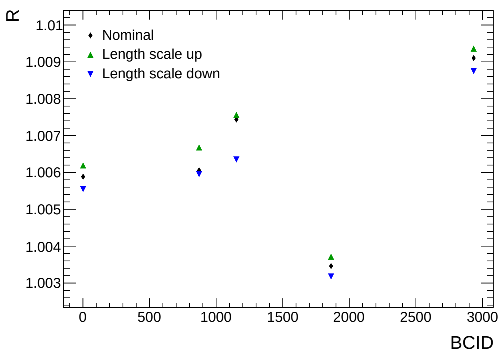


(c)

Figure 6.17: Values of R for beam spot data with fixed values of $k = 1.14$, $k = 1.16$, $k = 1.18$ and the nominal floating k for (a) scan I, (b) scan II and (c) the combined fit of scan II and III.



(a)



(b)

Figure 6.18: Values of R for separation data with the nominal and length scale variations applied for (a) scan I and (b) scan II, 2016.

per degree of freedom decreases in fits where no beam-beam deflection correction is applied, compared to nominal. The reason for this behaviour is not understood, but the uncertainty on R due to the length scale correction helps cover any discrepancies on the final non-factorisation correction due to this feature.

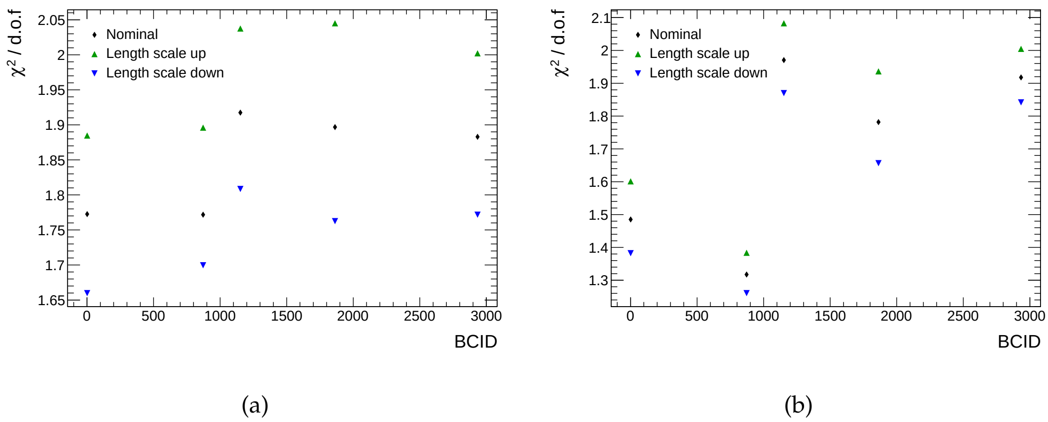


Figure 6.19: Values of χ^2 per degree of freedom for separation data with the nominal and length scale variations applied for (a) scan I and (b) scan II, 2016.

6.4.11.4 Beam model uncertainty

There is also an uncertainty associated with the beam model used to determine the non-factorisation correction. However, as discussed, the triple Gaussian model was deduced to be the best beam model to describe the 2015 and 2016 beam profiles. A somewhat artificial systematic uncertainty could be deduced by applying single and double Gaussian fits to the profiles and taking the difference from the triple Gaussian results. Applying such a systematic uncertainty would be extremely conservative since these beam profiles do not describe the data as well as the triple Gaussian. One would expect different, and less accurate, values of R. Therefore no systematic uncertainty is applied due to the beam model.

6.4.12 Non-factorisation results

6.4.12.1 2015

The results, with total uncertainty, for the 2015 scan sessions are shown in Figure 6.20. A summary of the individual uncertainties on R for each BCID in each scan is given in Table 6.12.

The total non-factorisation correction is close to unity on average in all scans. The values of non-factorisation for BCID 891 and 2451 are slightly less than unity in the centred scans I and II, and then become further from unity for the constrained scan II and III fit. This implies the tails of the beams contribute to non-factorisation more than the centred regions, and that the "no non-factorisation" assumption from the centred scans may not be totally true. However the three other BCIDs in that region are very close to $R = 1$. There is no reason why each BCID in a scan should have the same level of non-factorisation.

There is a slight tension with the non-factorisation corrections determined with the coupled model method in 2015, which show a (general) non-factorisation correction corresponding to $R > 1$ for all BCIDs. Due to this tension, no non-factorisation correction is applied to the calibrated value of σ_{vis}, with a conservative $\pm 1\%$ non-factorisation systematic uncertainty

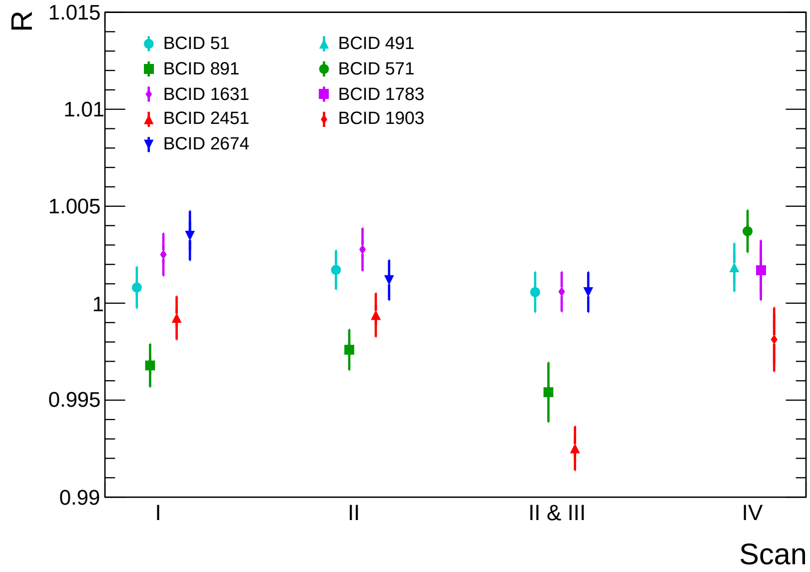


Figure 6.20: Values of R for each BCID in each scan session in 2015, with the total uncertainty shown.

Scan	BCID	R	Uncertainties				
			Statistical	Vertex resolution	Length scale	Systematic	Total
I	51	1.001	0.0004	0.0010	0.0002	0.0010	0.001
	891	0.997	0.0005	0.0010	0.0001	0.0010	0.001
	1631	1.003	0.0005	0.0010	0.0001	0.0010	0.001
	2451	0.999	0.0005	0.0010	0.0001	0.0010	0.001
	2674	1.003	0.0007	0.0010	0.0003	0.0010	0.001
II	51	1.002	0.0001	0.0010	0.0000	0.0010	0.001
	891	0.998	0.0003	0.0010	0.0000	0.0010	0.001
	1631	1.003	0.0005	0.0010	0.0001	0.0010	0.001
	2451	0.999	0.0005	0.0010	0.0000	0.0010	0.001
	2674	1.001	0.0003	0.0010	0.0000	0.0010	0.001
II & III	51	1.001	0.0003	0.0010	0.0001	0.0010	0.001
	891	0.995	0.0003	0.0010	0.0011	0.0015	0.002
	1631	1.001	0.0002	0.0010	0.0000	0.0010	0.001
	2451	0.993	0.0005	0.0010	0.0001	0.0010	0.001
	2674	1.001	0.0003	0.0010	0.0000	0.0010	0.001
IV	491	1.002	0.0007	0.0010	0.0002	0.0010	0.001
	571	1.004	0.0004	0.0010	0.0003	0.0010	0.001
	1783	1.002	0.0012	0.0010	0.0001	0.0010	0.002
	1903	0.998	0.0013	0.0010	0.0002	0.0010	0.002

Table 6.12: A summary of the values and uncertainties on R for each BCID and each scan in the 2015 vdM sessions. The associated uncertainties are broken down into individual components. The systematic and total uncertainty are also given.

applied. At the time of the study, the uncertainty on the total integrated luminosity collected in 2015 was 5% and thus the uncertainty due to non-factorisation is subdominant.

The amount of proton-proton data collected in 2015 is only roughly 10% of that collected in 2016. Reducing the size of the systematic uncertainty on the calibration of luminosity in 2016 is thus very important for precise physics measurements, such as those presented in Chapters 8 and 9.

6.4.12.2 2016

The non-factorisation results for the 2016 scan sessions are shown in Figure 6.21, with the break down of the uncertainties on said results in Table 6.13.

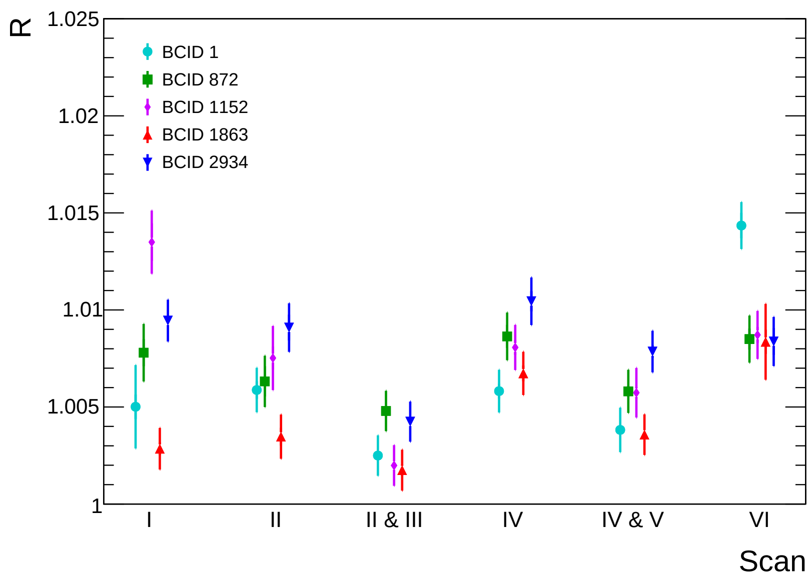


Figure 6.21: Values of R for each BCID in each scan session in 2016, with the total uncertainty shown.

The non-factorisation correction is roughly $R \approx 1.006$ on average. Only five BCIDs have tracking information, and hence beam spot data, available during a vdM scan. There are approximately 30 BCIDs which collect luminosity data. Therefore, a BCID by BCID luminous region non-factorisation correction cannot be applied. Instead a constant correction (over all BCIDs) is applied from the luminous region method.

All BCIDs in all the scans have relatively consistent values of R. The R values from the special offset (scan III) and diagonal (Scan V) scan types in combined fits all decrease compared to the centred scan used in the same fit (scans II and IV). However the difference is typically of the order 0.002, which is smaller than the total spread (in R) of the centred scans.

The flat non-factorisation correction applied to σ_{vis} is $R = 1.006 \pm 0.004$, which covers almost all BCIDs, except for outliers in scan I and VI. The application is σ_{vis}/R and hence the value of σ_{vis} is decreased due to the slight non-factorisation determined in the beams. No coupled model measurement was made in 2016 to compare with the luminous region method, but the consistent values of R over all scan types investigating different areas of the luminous region gives confidence in the analysis.

Scan	BCID	R	Uncertainties				
			Statistical	Vertex resolution	Length scale	Systematic	Total
I	1	1.005	0.0006	0.0010	0.0018	0.0021	0.002
	872	1.008	0.0007	0.0010	0.0008	0.0013	0.001
	1152	1.013	0.0010	0.0010	0.0008	0.0013	0.002
	1863	1.003	0.0003	0.0010	0.0002	0.0010	0.001
	2934	1.009	0.0002	0.0010	0.0004	0.0011	0.001
II	1	1.006	0.0005	0.0010	0.0003	0.0011	0.001
	872	1.006	0.0008	0.0010	0.0004	0.0011	0.001
	1152	1.008	0.0006	0.0010	0.0012	0.0015	0.002
	1863	1.003	0.0005	0.0010	0.0003	0.0010	0.001
	2934	1.009	0.0007	0.0010	0.0003	0.0011	0.001
II & III	1	1.002	0.0002	0.0010	0.0002	0.0010	0.001
	872	1.005	0.0002	0.0010	0.0002	0.0010	0.001
	1152	1.002	0.0003	0.0010	0.0002	0.0010	0.001
	1863	1.002	0.0003	0.0010	0.0002	0.0010	0.001
	2934	1.004	0.0002	0.0010	0.0001	0.0010	0.001
IV	1	1.006	0.0004	0.0010	0.0003	0.0010	0.001
	872	1.009	0.0006	0.0010	0.0004	0.0011	0.001
	1152	1.008	0.0004	0.0010	0.0005	0.0011	0.001
	1863	1.007	0.0004	0.0010	0.0003	0.0010	0.001
	2934	1.010	0.0005	0.0010	0.0005	0.0011	0.001
IV & V	1	1.004	0.0002	0.0010	0.0005	0.0011	0.001
	872	1.006	0.0003	0.0010	0.0004	0.0011	0.001
	1152	1.006	0.0003	0.0010	0.0007	0.0012	0.001
	1863	1.004	0.0002	0.0010	0.0002	0.0010	0.001
	2934	1.008	0.0003	0.0010	0.0002	0.0010	0.001
VI	1	1.014	0.0007	0.0010	0.0002	0.0010	0.001
	872	1.008	0.0005	0.0010	0.0004	0.0011	0.001
	1152	1.009	0.0004	0.0010	0.0006	0.0012	0.001
	1863	1.008	0.0006	0.0010	0.0016	0.0019	0.002
	2934	1.008	0.0007	0.0010	0.0003	0.0011	0.001

Table 6.13: A summary of the values and uncertainties on R for each BCID and each scan in the 2016 vdM sessions. The associated uncertainties are broken down into individual components. The systematic and total uncertainty are also given.

Chapter 7

A multivariate method to reject non-prompt leptons

The decays of W and Z bosons are commonly selected by the identification of light leptons (ℓ); one or two electrons or muons. The mean lifetime (time before decay) of the weak bosons is negligible and thus any leptons produced in the boson decay originate from the primary interaction vertex and are labelled *prompt*. ATLAS analyses using these light leptons impose strict reconstruction quality, isolation and impact parameter requirements primarily to remove *fake* or *non-prompt* leptons: fake referring to a physics object being wrongly reconstructed as a lepton and non-prompt referring to true leptons that do not originate from the primary vertex. Non-prompt leptons are produced in decays of hadrons that contain b or c quarks. Such hadrons have significant mean lifetimes that can be detected experimentally in ATLAS. These non-prompt leptons typically have a large impact parameter and are non-isolated due to the nearby jet constituents.

However, non-prompt leptons can occasionally also pass the tight selection criteria. In analyses that involve t quarks, which decay almost exclusively into a W boson and a b quark [3], non-prompt leptons from the semileptonic decay of bottom and charm hadrons can be a significant source of background events.

In this chapter, a novel multivariate method to reject such non-prompt light leptons in ATLAS is presented. The non-prompt leptons are identified using lifetime information associated with a track jet that matches the selected light lepton. This lifetime information is computed using tracks contained within the jet. Typically, lepton lifetime is determined using the impact parameter of the associated track of the lepton. Using additional reconstructed charged particle tracks increases the precision of identifying the displaced decay vertex of bottom or charm hadrons that produced a non-prompt lepton. This lifetime information is fed into a BDT, along with jet reconstruction information and lepton isolation to form an algorithm used to reject non-prompt leptons.

In Section 7.1 the boosted decision tree algorithm is discussed. In Section 7.2 the development and training of the so-called non-prompt BDTs are detailed. The performance of the algorithm in rejecting non-prompt leptons in $t\bar{t}$ simulation is shown in Section 7.3 and a validation of the BDT distributions for different particle types in data validation regions are shown in Section 7.4. The calibration of a working point of the BDT is finally discussed for prompt muons in Section 7.5.

7.1 Boosted decision trees

BDTs are a popular machine learning algorithm in high energy physics [97]. A standard approach in isolating interesting topologies is to apply *cuts* to discriminating variables to best separate a defined signal from a background. However, when such a cut is applied any events that do not pass the criteria are rejected and are thus not used further in any way. If the cut is not fully efficient then signal events can be rejected. Decision trees apply multivariate techniques to optimise cut values and to further analyse such rejected events, by sequentially applying cuts to other variables to events that both pass and fail the previous cut.

In this thesis, only two classifications are considered when building decision trees; *signal* and *background*, in this case prompt and non-prompt respectively. One has to choose certain features to build the decision tree that show discrimination between signal and background. For each feature a maximum *separation* is defined, detailing how well signal and background can be discriminated with a one dimensional cut. The tree then starts from an initial node, called the root node, which takes the variable with the best separation and applies the corresponding cut to achieve that separation. The node is then split into two *branches*: those that pass and those that fail the cut. The procedure is applied iteratively until stopping criteria are met. The stopping criteria can be a number of different requirements: if these criteria are met then the final node is turned into a *leaf*. The method for which the tree is built until there are only leaves remaining is called training. An example of a fully trained decision tree is shown in Figure 7.1.

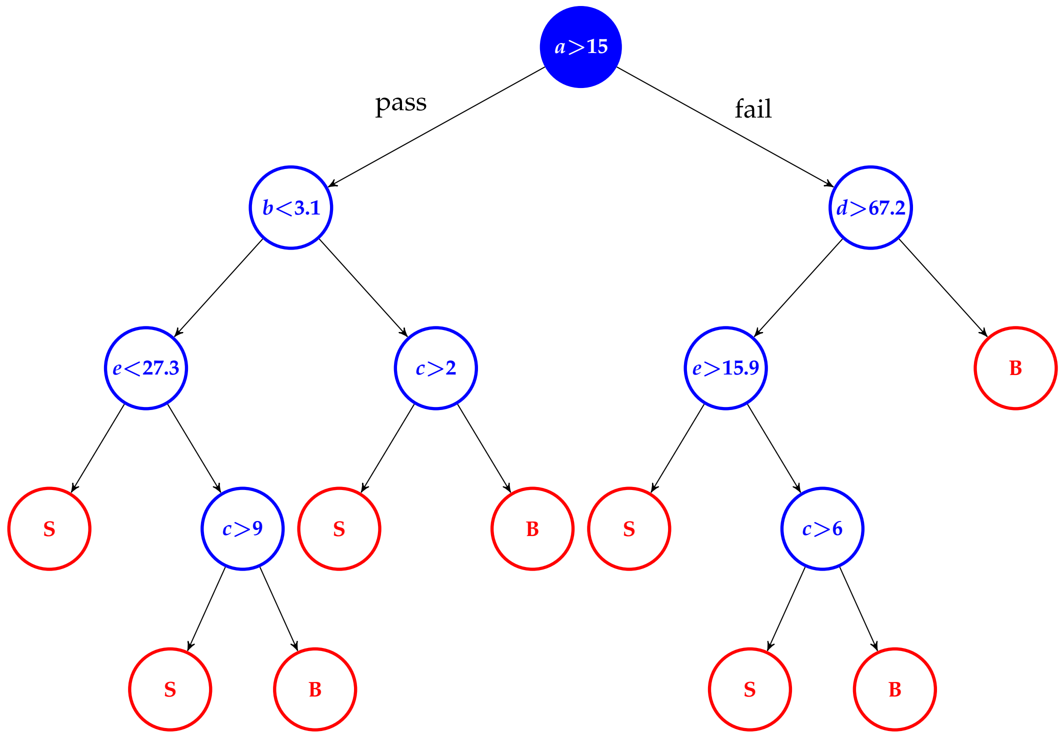


Figure 7.1: An example of a decision tree, with features a, b, c, d, and e. The tree starts at the filled blue root node. The blue circles are internal nodes containing the optimised splitting criteria for the available features. If an event passes the criteria then it follows the path of the left arrow and the right arrow for failing events. The red circles are leaves labelled as S or B, denoting if the final leaf has signal purity, p_s, greater than or less than 0.5.

The signal and background purities, p_s and p_b, are defined as

$$p_s = \frac{s}{s+b}, \qquad p_b = \frac{b}{s+b}, \tag{7.1}$$

with s the weighted number of signal events and b the weighted number of background events in a leaf. The separation criteria applied to each node to determine whether the node is a leaf or is further split is a value of the Gini index, defined as

$$G = 1 - \sum_{i=s,b} p_i^2 = 2p_s p_b = \frac{2sb}{(s+b)^2}. \tag{7.2}$$

The node is then split if the variable that minimises the misclassification rate as defined by the Gini index improves the separation at the node.

A single decision tree is likely to be unstable due to cut optimisations being done on a finite training sample. Statistical fluctuations in the sample may be picked up by the tree and such trees are said to be *overtrained*. A decision tree can be checked for overtraining by comparing the performance of the decision tree when applied to the training sample and an orthogonal testing sample. There are a number of methods of reducing the instability of such algorithms.

Trees can be pruned and averaged to negate such instabilities. Pruning refers to removing leaves with low statistics, which can be extended to whole branches. Averaging refers to taking the weighted mean of multiple decision trees. For example, the k-fold cross validation technique [98] involves splitting a training sample up into k equal subsets and training a decision tree on each subset, validating on the remaining $k-1$ samples. These trees are then averaged which minimises overtraining. A more complex "averaging" technique is *boosting*.

The idea of boosting [99, 100] is to train many versions of a weak learner (in this case a decision tree) and to combine these weak learners into a more robust algorithm. Boosting sequentially reweights training events to minimise a loss function that typifies the difference between the model and the true value. The resulting algorithm is then a weighted average of the iteratively boosted trees. This method typically greatly reduces the error rate and increases performance with respect to the single weak learner. The BDTs discussed in this thesis are gradient boosted decision trees, unless otherwise stated. Gradient boosting is discussed in the literature [101].

7.2 Non-prompt BDT training

Non-prompt leptons originate in decays of hadrons containing b or c quarks. In general, these hadrons are embedded within particle jets that contain a large number of tracks and calorimeter clusters close to the non-prompt lepton. Thus a majority of non-prompt leptons fail typical isolation selection criteria. Any non-prompt lepton that passes these criteria can do so by two means. First, the lepton may carry a large fraction of the energy of the b- or c-jet, so that the remaining jet components are not energetic enough for the reconstructed lepton to fail the imposed isolation requirements. Second, the lepton can decay in a direction far enough away from the remaining jet components that the isolation cone is not large enough to capture the energy carried by the jet. Both of these problems can in some way be mitigated by investigating the properties of the nearest track jet to the lepton and by checking whether this jet is consistent with a bottom or charm jet. Track jets are chosen for this task due to the high probability of reconstruction near to any lepton, since lepton reconstruction requires the presence of a reconstructed inner detector track. Calorimeter jets have a lower probability of reconstruction and matching, especially for the minimally ionising muons.

Figure 7.2 shows a basic schematic diagram of how a non-prompt lepton can pass isolation and impact parameter requirements. This figure illustrates how the additional information associated with the track jet (which is matched to the reconstructed lepton) can be used to identify and veto non-prompt leptons. The isolation cone may miss high p_T tracks (shown in green) from the secondary decay vertex in the non-prompt case. A reconstructed track jet nearly always includes the reconstructed lepton track; it also uses a larger radius to collect tracks than the isolation variable, which may enable an errant high p_T track to be included in the analysis. A secondary vertex can also be reconstructed within the track jet, which leads to a powerful way of discriminating between the prompt and non-prompt cases. This section focuses on tagging non-prompt leptons by using standard b-tagging techniques on track jets containing the lepton track, amongst other discriminating features.

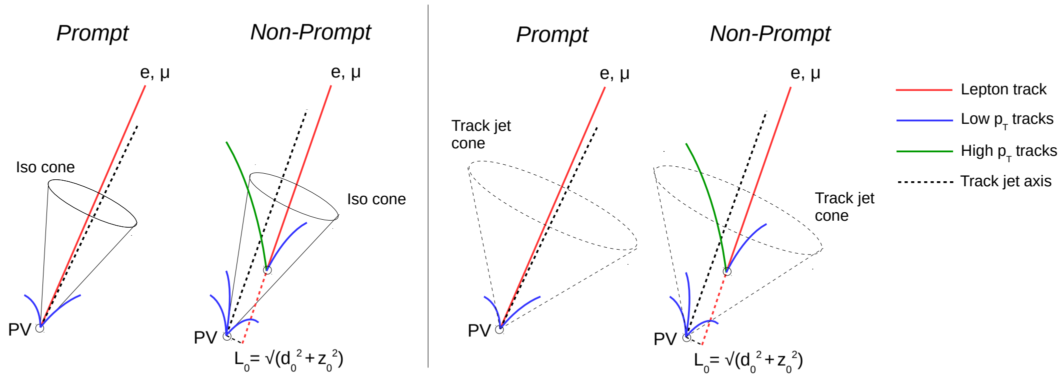


Figure 7.2: Schematic diagrams of prompt and non-prompt leptons that pass basic impact parameter and isolation cuts. Extra tracks that are missed and unused when only using isolation (left) to discriminate between prompt and non-prompt leptons compared to using track jet information (right) are highlighted. d_0 and z_0 correspond to the transverse and longitudinal impact parameters respectively, with L_0 corresponding to the secondary vertex decay length from the primary vertex (PV).

7.2.1 MC sample

A full simulation $\sqrt{s} = 13$ TeV $t\bar{t}$ MC sample was used in the training of the BDTs. The sample was produced using the POWHEG [31] NLO generator interfaced with PYTHIA 6 [24]. The CT10 [102] PDF was used in the matrix element generation and the CTEQ6L1 [103] PDF in the parton shower. The heavy flavour quark decays are modelled by EVTGEN [104]. All generated events are passed through a GEANT4 [34] full simulation of the ATLAS detector. A filter is applied to the sample so that all events with at least one leptonic W decay are retained.

7.2.2 Training selections

Search analyses with leptons at ATLAS employ tight restrictions on the p_T, isolation and impact parameters of the leptons to reduce non-prompt lepton backgrounds. Due to these tight selection criteria, the number of isolated non-prompt leptons within the $t\bar{t}$ MC sample discussed above is not large enough to train a BDT algorithm without noticeable statistical fluctuations. Thus two sets of lepton selection criteria are employed: loose and tight. The loose selection criteria are used for training the BDT (training selection). The tight selection criteria (testing selection) are used to evaluate BDT performance. The tight selections correspond to typical tight lepton selections. The BDT algorithm can also be used with any other lepton selections that may be different from those evaluated here.

7.2.2.1 Object selection

Table 7.1 lists the training and testing lepton selections.

Track jet candidates are reconstructed using the anti-k_t clustering algorithm with a radius parameter $R = 0.4$ to cluster tracks within the jet. The tracks must originate from the primary

	Training		Testing	
Object	Electron	Muon	Electron	Muon
p_T [GeV]	> 10	> 10	> 10	> 10
$\|\eta\|$	< 2.47	< 2.5	< 2.47	< 2.5
PID	`LooseLH`	`Loose`	`TightLH`	`Loose`
$\|d_0/\sigma_{d_0}\|$	< 7	< 7	< 5	< 3
$\|z_0 \sin\theta\|$ [mm]	< 2	< 2	< 0.5	< 0.5
p_TVarCone20$/p_T$	< 0.5	< 0.5	< 0.06	-
p_TVarCone30$/p_T$	-	-	-	< 0.06
E_TTopoCone20$/p_T$	< 0.5	< 0.5	< 0.06	-

Table 7.1: The lepton training and testing selections. The isolation selections for testing correspond to the `FixedCutTight(FixedCutTightTrackOnly)` isolation working points for electrons and muons respectively.

vertex and have $p_T > 400$ MeV. The selected track jets are required to have $p_T > 10$ GeV and $|\eta| < 2.5$.

7.2.2.2 Event selection

$t\bar{t}$ events with at least one training electron or one training muon candidate are used. Each reconstructed lepton is required to be matched to a same flavour truth lepton within $\Delta R < 0.05$. The selected reconstructed electrons and muons are split up according to truth level information: the lepton is labelled as prompt if the parent of the truth lepton is a W boson or a τ lepton. If the lepton parent is a τ, a further check is required to check that the parent of the τ is a W, to stop wrongly classifying $b/c \to \tau$ decays as prompt. If the lepton originates from a bottom or charm hadron it is labelled non-prompt. If the lepton has any other parent then it is discarded from the training. Such muons that are not classified this way are typically from light quark decays and for electrons from mis-reconstructed jets and photon conversions. Leptons that do not have a truth matched particle are also not considered in the training; there are fewer than 1.3% (1.7%) of such electrons (muons).

A track jet is required to be within $\Delta R < 0.4$ of a lepton. If no track jet is found near to the lepton, the lepton is excluded from the training. The fraction of leptons that do not have a track jet within $\Delta R < 0.4$ is less than 0.1%.

7.2.3 Training variables

Eight variables are chosen to train the BDT algorithm in order to discriminate between prompt and non-prompt leptons, briefly described in Table 7.2. Approximately 3000000 (3000000) prompt and 360000 (540000) non-prompt electrons (muons) are used to train the non-prompt BDTs.

The track jets that are matched to the non-prompt leptons correspond to jets initiated by b- or c-quarks and hence may contain a displaced vertex. Consequently, three of the selected variables are used to identify b-tag jets by the standard ATLAS flavour tagging algorithms [105]. These variables are IP2D, IP3D and SV1+JF N_{TrkAtVtx}.

IP2D and IP3D are impact parameter based algorithms. They use the signed transverse impact parameter of collected nearby tracks in the jet to distinguish between b, c and light jets. Likelihoods are built and the log-likelihood ratio between the b-jet and light jet hypotheses are taken. IP2D merely uses the signed transverse impact parameter when building the likelihoods, and IP3D uses both transverse and longitudinal impact parameters. These variables are hence correlated but provide complementary information. The variables are more discriminatory than purely the lepton impact parameter, due to the extra information from nearby tracks that is used. The SV1+JF N_{TrkAtVtx} variable is the addition of the number of tracks found in secondary vertices by the SV1 and JetFitter algorithms. If one of the algorithms finds a secondary vertex then the likelihood that the track jet investigated is in fact a heavy flavour jet is high. The addition of the same variable from the two algorithms maximises the chance that at least one of SV1 and JetFitter in fact find a secondary vertex. In isolated environments this variable is not very discriminatory, but is a very useful tool in less isolated cases. The higher level b-tagging variables, such as the SV1 and JetFitter log-likelihood ratios and the MV2 BDTs, are overlooked in the BDT training since these variables are trained to discriminate between heavy and light flavour jets in high-track environments encompassing calorimeter jets. The lower level variables are used instead (inputs to the higher level) that are less biased to the local environment in which the non-prompt BDT is trained.

Two variables used in the non-prompt BDT use the differing relationship between the track jet and the lepton for prompt and non-prompt; the ratio of the lepton p_T with respect to the track jet p_T and ΔR between the lepton and the track jet axis. One would expect the track jet p_T and ΔR to the lepton to be larger for non-prompt than prompt due to the expected extra tracks.

Finally three additional variables test whether the reconstructed lepton is isolated; the number of tracks collected by the track jet and the lepton track and calorimeter isolation variables.

All variables are chosen to be minimally dependent on lepton momentum. The momentum itself would be an extremely powerful variable to use in the training of the BDT. However lepton momentum is highly dependent on the kinematics of different events. The non-prompt BDT is calibrated for prompt leptons in Z events (see Section 7.5) and this calibration is used for all prompt leptons in a number of different events. If the non-prompt BDT was strongly correlated with p_T then it the validity of this calibration on other events is not clear. A similar argument can be made for lepton η.

Variable	Description
N_{track} in track jet	Number of tracks collected by the track jet
IP2 $\log(P_b/P_{\text{light}})$	Log-likelihood ratio between the b and light jet hypotheses with the IP2D algorithm
IP3 $\log(P_b/P_{\text{light}})$	Log-likelihood ratio between the b and light jet hypotheses with the IP3D algorithm
N_{TrkAtVtx} SV1 + JF	Number of tracks used in the secondary vertex found by the SV1 algorithm in addition with the number of tracks from secondary vertices found by the JetFitter algorithm with at least two tracks
$p_T^{\text{lepton}}/p_T^{\text{track jet}}$	The ratio of the lepton p_T and the track jet p_T
$\Delta R(\text{lepton, track jet})$	ΔR between the lepton and the track jet axis
$p_T\text{VarCone30}/p_T$	Lepton track isolation, with track collecting radius of $\Delta R < 0.3$
$E_T\text{TopoCone30}/p_T$	Lepton calorimeter isolation, with topological cluster collecting radius of $\Delta R < 0.3$

Table 7.2: The variables used in the training of the non-prompt BDTs.

The distributions of the eight training variables for the training electron and training muon selections are shown in Figures 7.3 and 7.4 respectively. Equivalent distributions for the tighter testing electron and testing muon selections are shown in Figures 7.5 and 7.6 respectively.

The distributions of the training selections show greater separation between prompt and non-prompt due to the very loose isolation requirements. The results of the isolation requirements in the lepton selections can be seen in the isolation distributions. Even with very tight isolation requirements, there is still discrimination between prompt and non-prompt for the testing leptons. This gives confidence that the non-prompt BDT will be able to outperform tight isolation working points since there is still discriminatory power for leptons that pass all isolation requirements.

Equivalent distributions for electrons and muons are in general similar, with one exception being the p_T ratio. For combined muons, p_T is measured from a fit of the inner detector track and muon spectrometer track. One would expect naively that the distribution would peak at unity; that is the track jet would have larger p_T than purely from the lepton track, itself included in the track jet. Actually the prompt distributions can be greater than unity, which are from the cases with a track jet with a single track nearby. In this case the comparison is from the single inner detector track from the track jet and the combined track from the muon, which can be larger. The electron distribution is different due to the different measurement of p_T for electrons. Calorimeter energy information is also used to measure the p_T and so the comparison to the inner detector track is not a like-for-like comparison[1]. The p_T ratio distributions for leptons matched to a track jet with only a single track are shown in Figure 7.7.

Another variable to note is the N_{TrkAtVtx} SV1 + JF variable. To make a vertex at least two tracks are required. One would then assume that the addition of these variables would have no values at 1. However, due to technical reasons both individual variables are initialised to -1 and thus if one algorithm finds a vertex and the other does not than a value of 1 can be achieved. This may result in a slight drop in performance compared to an initialisation at 0 but this is expected to be a negligible effect. If one of the algorithms finds a vertex it is a very powerful variable to discriminate prompt and non-prompt, due to the low fake vertex rate for prompt leptons.

The correlation matrices for the input variables that are used in the BDT are shown in Figure 7.8. Variables that are highly correlated collectively have poorer rejection power than variables with minimal correlation. As expected the impact parameter based variables are quite highly correlated (60 to 80% correlation) for both leptons and both categories. Especially for prompt leptons, where the number of tracks are less and thus the IP scores are very similar. However, BDTs can use correlated variables and this extra rejection power is non-negligible. The final BDT weight is that which is used in data and calibrated; any input variable correlations are accounted for in this calibration.

7.2.4 Training parameters

The hyper-parameters used to train the BDTs are shown in Table 7.3 below.

A small optimisation of three important hyper-parameters was performed to determine the optimum configuration setup. The number of trees, N_{trees} (600, 800, 1000), the maximum depth

[1] A fairer comparison would be to compare the electron *track* p_T with the track jet p_T. The same argument applies to muons with the inner detector track, but to a lesser extent. For technical reasons this variable was not used in the training.

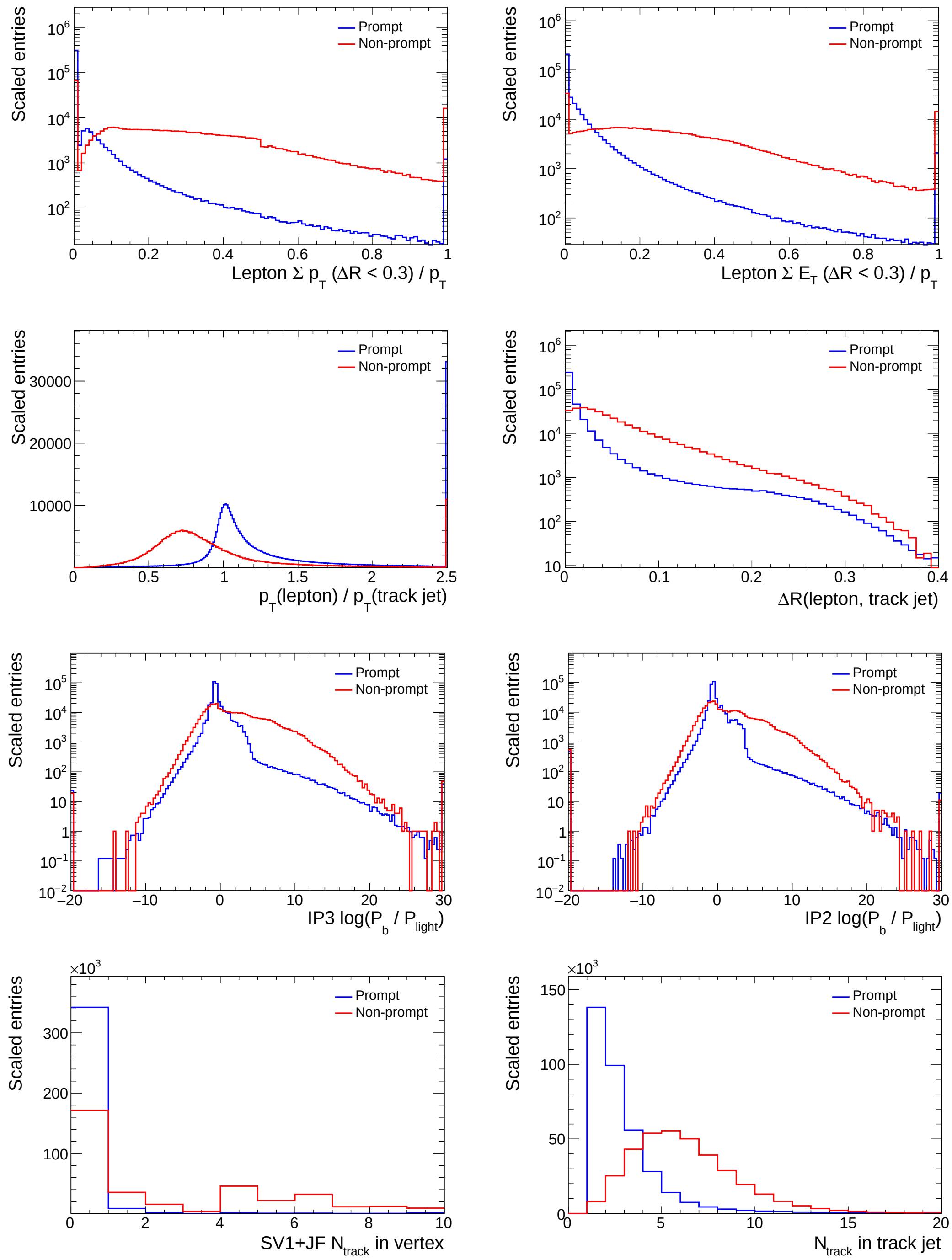


Figure 7.3: Distributions of the variables used as an input to non-prompt BDT for the electron training selection in $t\bar{t}$ simulation. The prompt distributions (in blue) are scaled to the number of non-prompt electrons (in red). Roughly 3000000 prompt and 360000 non-prompt electrons pass the training sample selection.

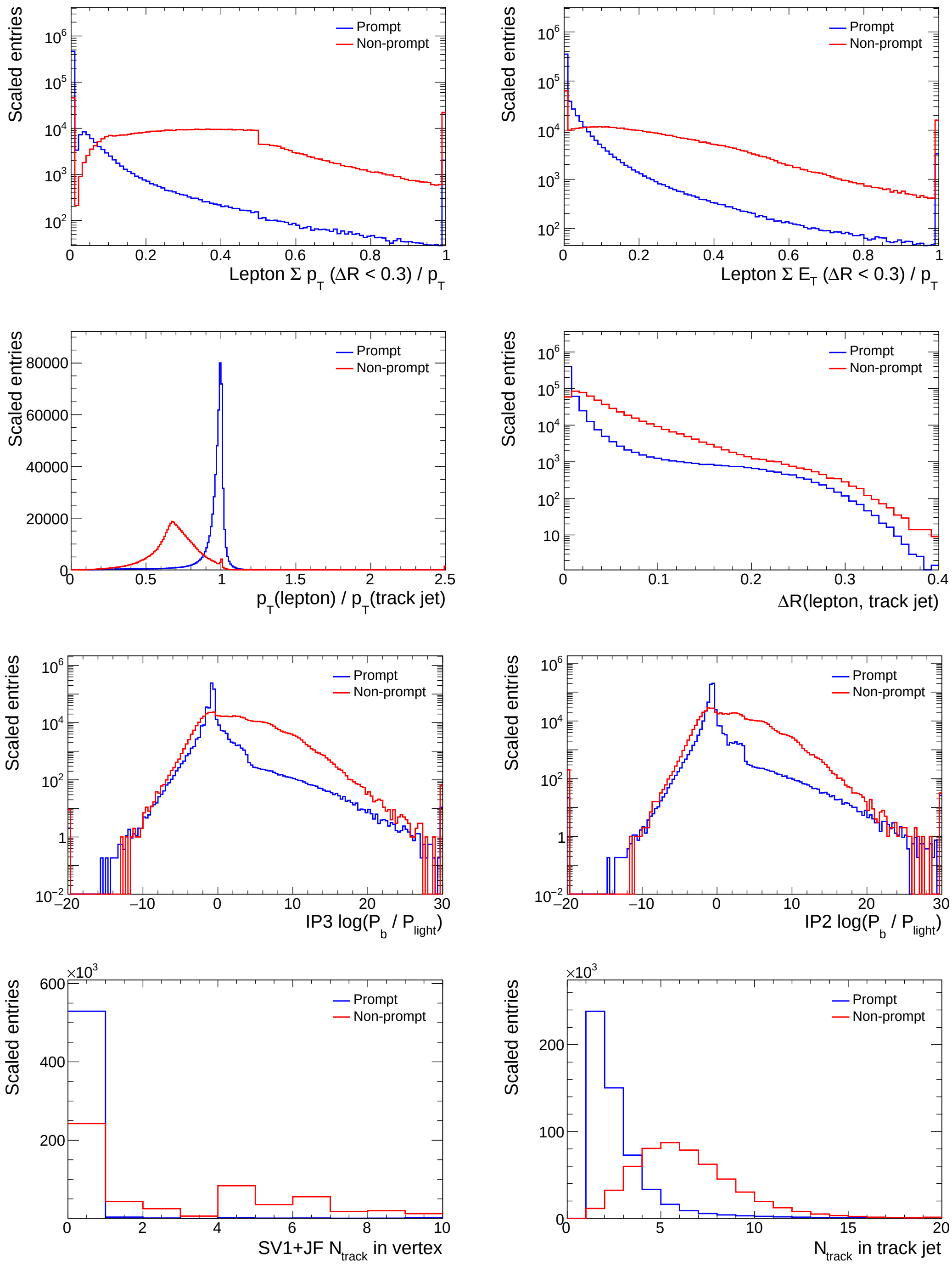


Figure 7.4: Distributions of the variables used as an input to non-prompt BDT for the muon training selection in $t\bar{t}$ simulation. The prompt distributions (in blue) are scaled to the number of non-prompt muons (in red). Roughly 3000000 prompt and 540000 non-prompt muons pass the training sample selection.

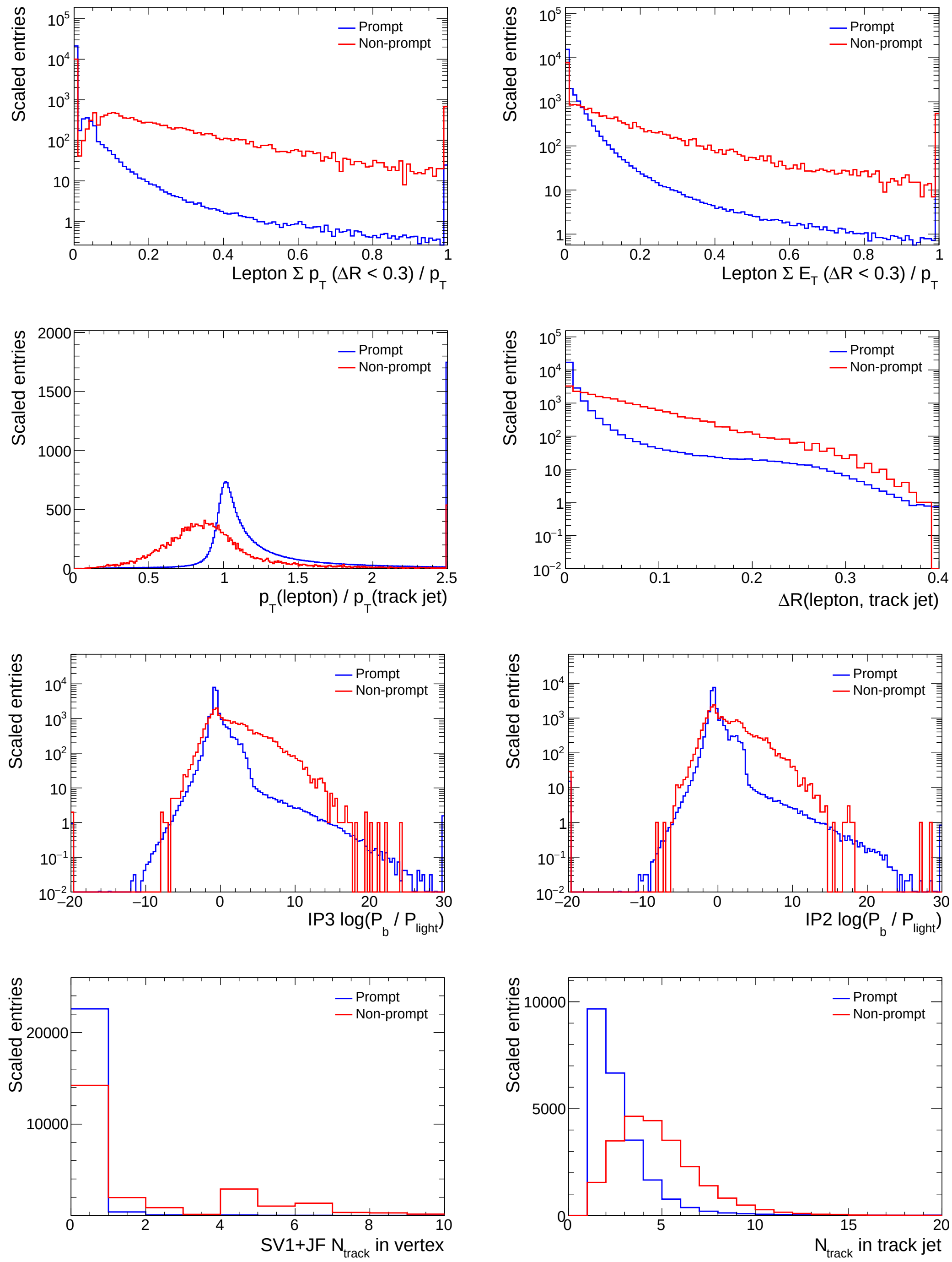


Figure 7.5: Distributions of the variables used as an input to the non-prompt BDT for the electron testing selection in $t\bar{t}$ simulation. The prompt distributions (in blue) are scaled to the number of non-prompt electrons (in red). Roughly 2250000 prompt and 23000 non-prompt electrons pass the testing sample selection.

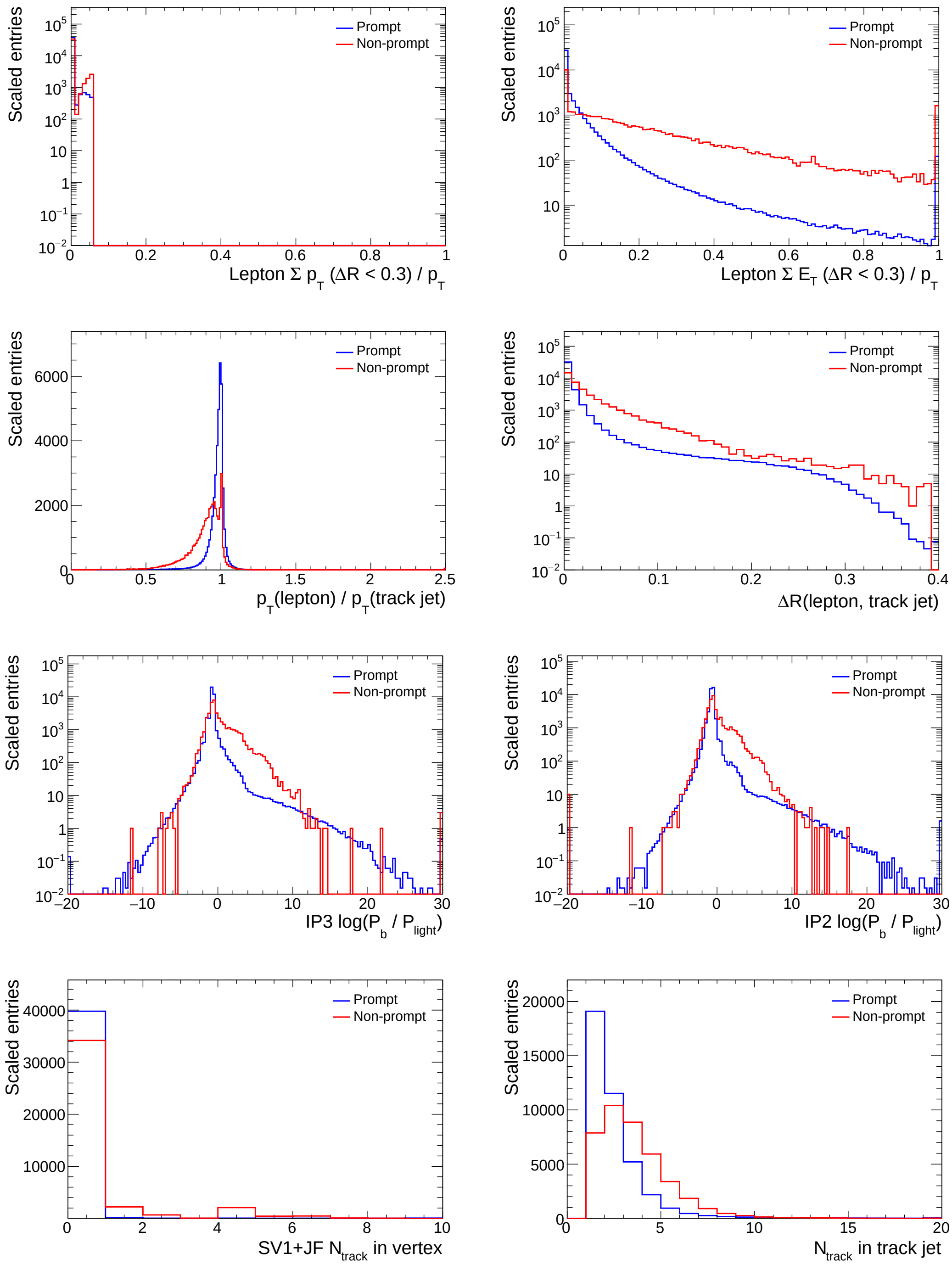
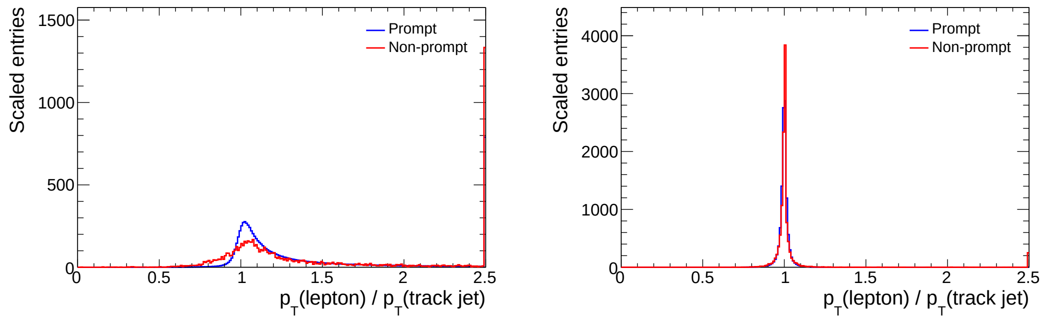


Figure 7.6: Distributions of the variables used as an input to non-prompt BDT for the muon testing selection in $t\bar{t}$ simulation. The prompt distributions (in blue) are scaled to the number of non-prompt muons (in red). Roughly 2650000 prompt and 40300 non-prompt muons pass the testing sample selection.



(a) (b)

Figure 7.7: Distributions of the $p_T^{\text{lepton}} / p_T^{\text{track jet}}$ variable for (a) electrons and (b) muons matched to a track jet only with a single track jet in $t\bar{t}$ simulation. The prompt distributions (in blue) are scaled to the number of non-prompt leptons (in red).

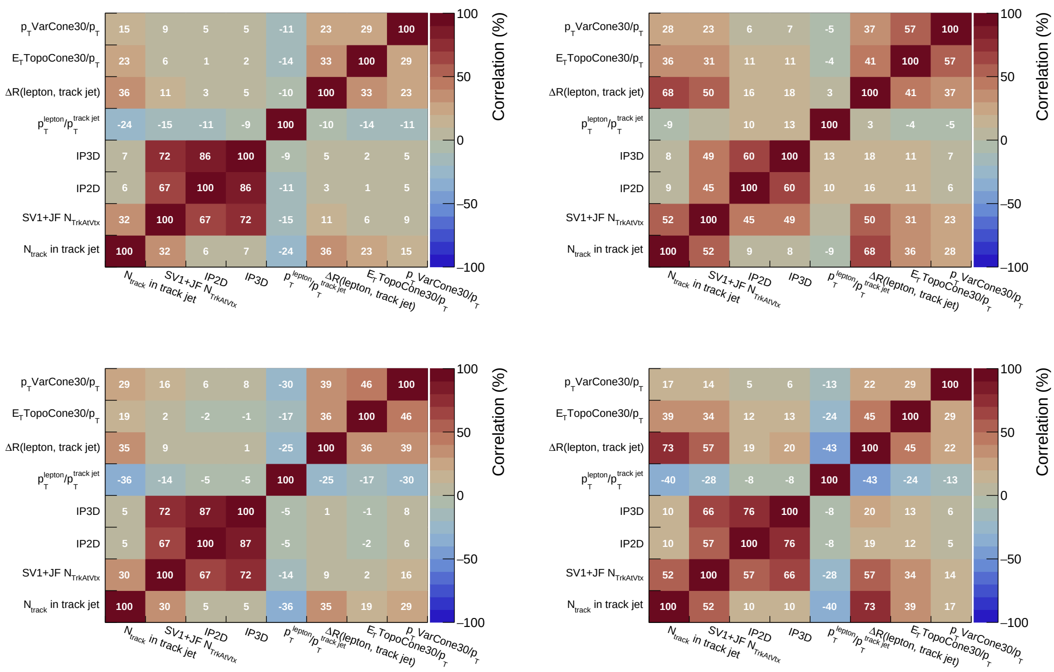
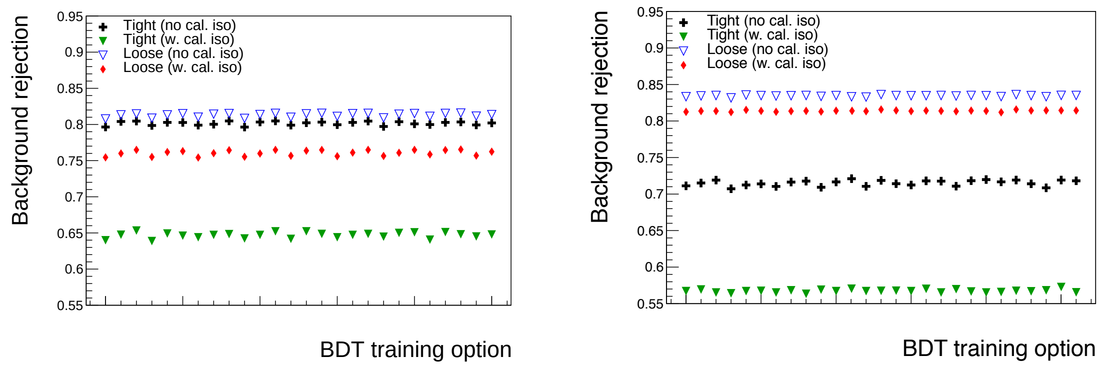


Figure 7.8: Input variable correlation matrices for prompt (left) and non-prompt (right) electron (top) and muons (bottom) used in the BDT training in $t\bar{t}$ simulation.

Parameter	Value	Description
N_{trees}	800	Number of trees, or boosting iterations, in the BDT
Minimum node size	0.05%	Minimum number of events allowed in a node size relative to total number of events
Maximum depth	9	Maximum number of cuts to reach a leaf
N_{cuts}	200	The number of points in a variable range used in finding the optimal cut in node splitting
Shrinkage	0.1	The learning rate for the gradient boost algorithm

Table 7.3: The hyper-parameters used in the non-prompt BDT training.

(5, 7, 9) and the number of cuts, N_{cuts} (100, 150, 200), were modified and BDTs were trained for each modification. A figure of merit is defined to determine the best performing configuration: the non-prompt rejection ($1-\epsilon_b$ for non-prompt efficiency, ϵ_b) of the one dimensional cut giving a 95% prompt efficiency, ϵ_s. The figure of merit for each of the different trainings and for four different lepton selections is shown in Figure 7.9. The maximum difference in figure of merit for each lepton selection is of the order of 1%. It was found that requiring a maximum depth of 9 in the training of the BDT resulted in slightly better performance than otherwise and so was retained. The maximum depth is allowed to exceed the number of variables used in the BDT due to splittings using the same variable in a branch of the tree. N_{trees} is a less important parameter and so 800 boosting iterations were chosen to reduce computational times.



(a) (b)

Figure 7.9: The non-prompt rejection, $1-\epsilon_b$, at 95% prompt efficiency for BDT trainings in $t\bar{t}$ simulation with different hyper-parameters for (a) electrons and (b) muons. The three values of maximum depth, N_{cuts} and N_{trees} are increased sequentially on the x-axis. The lepton selections are as in Table 7.1, with and without the calorimeter isolation selection applied. Loose refers to the training selection and Tight to the testing selection.

7.2.5 Training cross-validation

The $t\bar{t}$ MC sample is categorised into electrons and muons, which are then defined as prompt and non-prompt. Each set is divided into leptons used for training the BDT, and leptons used to validate the BDT training. The two sets are statistically independent and so the separation calculated from the BDT on the validation sample has no biases due to statistical fluctuations captured by the BDT from the leptons the BDT was trained on. All plots of the non-prompt BDT input variables in Section 7.2 are from leptons that were not themselves used to train the BDT (from the validation sample).

A comparison of the efficiency curves of the non-prompt BDT between the leptons used in the BDT training and the leptons used for validating the output is shown in Figure 7.10, for leptons passing the training selection. The AUC (area under the curve) values in the cross-validation differ by no more than 0.2%, indicating that there is negligible over-training.

7.3 Performance of the non-prompt BDT in simulation

This section presents the performance of the non-prompt BDT in separating prompt and non-prompt leptons using the tight (testing) selection criteria in simulated $t\bar{t}$ events.

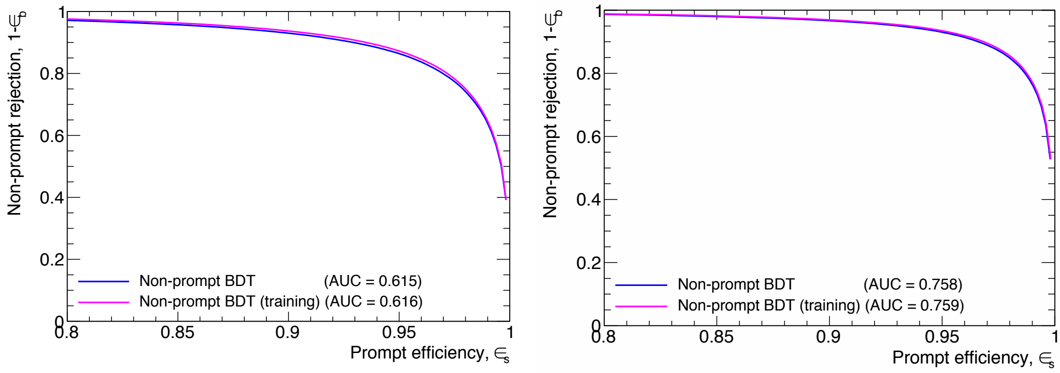


Figure 7.10: ROC curves describing the prompt lepton efficiency versus the non-prompt lepton rejection with the non-prompt BDT for (a) electrons and (b) muons passing the training lepton selections in $t\bar{t}$ simulation. Efficiency curves for the set of leptons used in the training and those used for validation are compared and the area under the curve (AUC) is shown. The difference in AUC between the two samples is 0.2% in both cases.

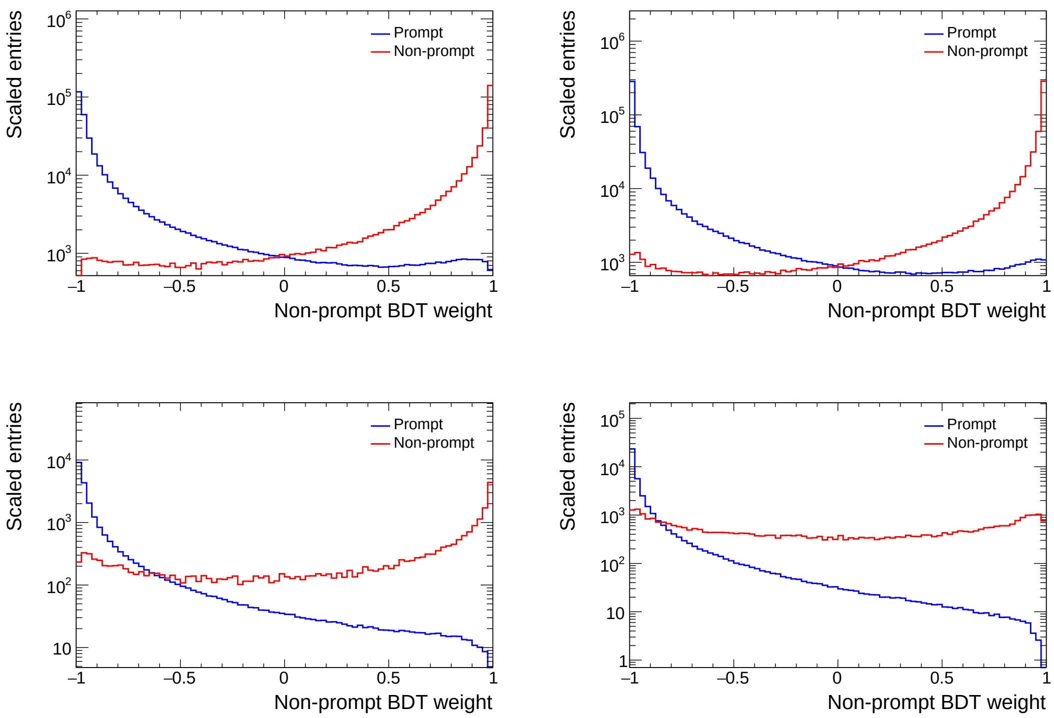
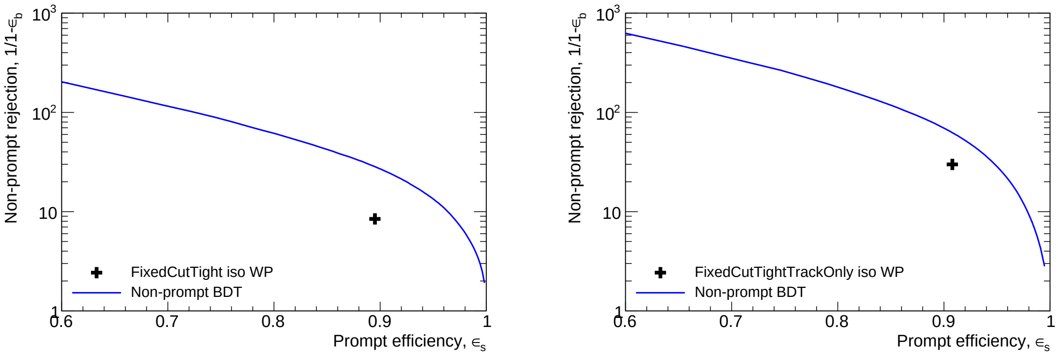


Figure 7.11: Distributions of the non-prompt BDT in $t\bar{t}$ simulation for prompt (in blue) and non-prompt (in red) training (top) and testing (bottom) electrons (left) and muons (right) with a log scale. The prompt distribution is scaled to the number of non-prompt leptons and only the cases where a lepton has a coincident track jet within $\Delta R < 0.4$ are considered.

The electron and muon BDT distributions are shown in Figure 7.11 for prompt and non-prompt training and testing leptons. The non-prompt BDT shows excellent separation between the prompt and non-prompt cases for the leptons passing the training selections. Even for the tightly isolated testing selection leptons, the BDT retains a good separation between prompt and non-prompt.

A non-prompt BDT working point is designed to totally replace the isolation lepton selection and not to apply the working point on top of tight isolation selections. Figure 7.12 compares the efficiency of selecting prompt leptons with the efficiency of rejecting non-prompt leptons for the non-prompt BDT and the best performing isolation selections. The non-prompt BDT outperforms the tightest (data calibrated) lepton isolation selections: `FixedCutTight` for electrons and `FixedCutTightTrackOnly` for muons.



(a) (b)

Figure 7.12: ROC curves describing the prompt lepton efficiency versus the non-prompt lepton rejection with the non-prompt BDT in $t\bar{t}$ simulation for (a) electrons and (b) muons. The leptons pass the testing selections apart from the isolation selection which is not applied, so that the performance of the non-prompt BDT can be studied. The tightest lepton isolation selection is also shown for reference.

Tables 7.4 and 7.5 compare the prompt efficiency and non-prompt rejection for the `isoFixedCutTight (isoFixedCutTightTrackOnly)` electron (muon) isolation working points against the non-prompt BDT working point with similar prompt efficiency. Roughly a factor 3 (2) increase in non-prompt rejection for the same prompt efficiency is achieved for electrons (muons). However, these results do not include an overlap removal procedure; a procedure in which leptons can be removed if close to jets. This procedure acts in some way like an isolation cut and so the true performance increase is to be measured with the overlap procedure applied and checked in real data. This is discussed in the context of the $t\bar{t}H$ multilepton analysis in Chapter 8.

Working point	Prompt efficiency, ϵ_s (%)	Bkg. rejection, $1-\epsilon_b$ (%)
`isoFixedCutTight`	89.5	88.2
Non-prompt BDT	89.8	96.4

Table 7.4: Comparison of the prompt efficiency and non-prompt rejection in $t\bar{t}$ simulation for the `isoFixedCutTight` working point and the non-prompt BDT for testing electrons without the application of isolation. The non-prompt BDT working point shown is designed to have roughly the same prompt efficiency as the isolation working point.

Algorithm	Prompt efficiency, ϵ_s (%)	Bkg. rejection, $1-\epsilon_b$ (%)
FixedCutTightTrackOnly	90.8	96.7
Non-prompt BDT	90.9	98.4

Table 7.5: Comparison of the prompt efficiency and non-prompt rejection in $t\bar{t}$ simulation for the `isoFixedCutTightTrackOnly` working point and the non-prompt BDT for testing muons without the application of isolation. The non-prompt BDT working point shown is designed to have roughly the same prompt efficiency as the isolation working point.

7.4 Modelling of the non-prompt BDT

So far the performance of the non-prompt BDT has only been studied in simulation. This section studies the modelling and performance of the non-prompt BDT in real data in validation regions containing both prompt and non-prompt leptons.

The lepton selections used are shown in Table 7.6 below. The lepton p_T and η cuts are dependent on the validation region and, if different to Table 7.6, are discussed below where appropriate.

Variable	Electrons	Muons
p_T	> 25 GeV	> 25 GeV
$\lvert\eta\rvert$	< 1.37 or $1.52 < \lvert\eta\rvert < 2.5$	< 2.5
$\lvert d_0\rvert/\sigma_{d_0}$	< 5	< 3
$\lvert z_0 \sin(\theta)\rvert$	< 0.5 mm	< 0.5 mm
PID	`TightLH`	`Loose`
Isolation	`isoLoose`	`isoLoose`

Table 7.6: Definition of the electron and muon selections used for data/MC comparisons in validation regions. The calorimeter crack region is vetoed for electrons.

Some regions require selections on the number of jets or *b*-jets. The jet selections are shown in Table 7.7.

Variable	Jets	*b*-jets
p_T	> 25 GeV	> 25 GeV
$\lvert\eta\rvert$	< 2.5	< 2.5
JVT	pass	pass
MV2c10	-	70%

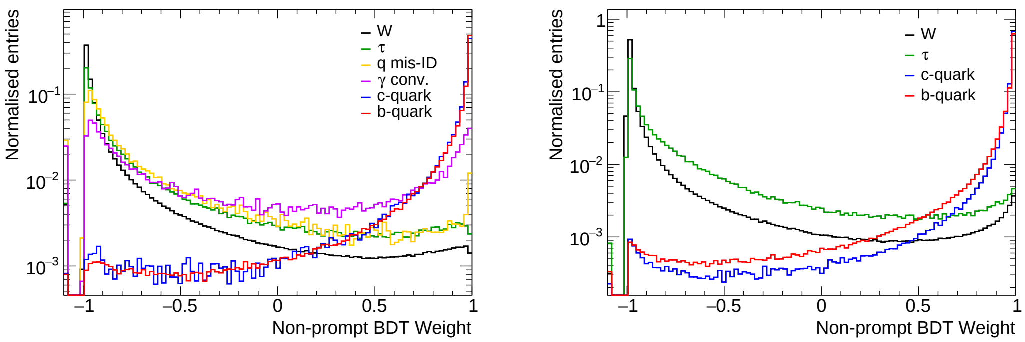
Table 7.7: Definition of the jet and *b*-jet selections used for data/MC comparisons in validation regions.

7.4.1 Truth distributions

In Sections 7.2 and 7.3 the distributions of the non-prompt BDT for prompt and non-prompt leptons is shown in simulation. In this case prompt refers to $W \to \ell$ and $W \to \tau \to \ell$, and non-prompt to leptons with parents containing either b or c-quarks.

Figure 7.13 shows the inclusive, normalised distributions from simulation for each of the above subcategories for muons and two extra categories for electrons; photon conversions and charge misidentified electrons. Photon conversions are real electrons that originate from photons that convert into an electron-positron pair in the inner detector. At least one of the electrons are then reconstructed. Charge-misidentified electrons are predominantly real electrons that bremsstrahlung in the inner detector with the resulting photon converting to an electron-positron pair and the wrong sign electron being reconstructed. There is also a charge misidentification contribution from very high energy electrons where the curvature of the track is misconstrued and the wrong charge reconstructed. Both photon conversions and charge-misidentified electrons in the $t\bar{t}$ sample predominantly originate from a $W \to e$ decay but their nature of producing a "vertex" in the inner detector result in modifications to the non-prompt BDT distributions. These cases show a larger proportion of leptons with no track jet reconstructed nearby corresponding to the BDT weights at -1.1.

The distributions of the non-prompt BDT for $\tau \to \ell$ decays also show shift towards the non-prompt region compared to $W \to \ell$ decays. τ leptons have a mean lifetime of 87 μm [3] and thus will, on average, have a larger impact parameter (and impact parameter significance) than W decays.



(a) (b)

Figure 7.13: Normalised distributions of the non-prompt BDT for different lepton types for (a) electrons and (b) muons in $t\bar{t}$ simulation. "q mis-id" refers to charge misidentification and "γ conv." to photon conversions.

7.4.2 Data distributions

So far we have only looked at how the non-prompt BDT performs in simulated data. In reality, the performance in data is more important. The modelling of the BDT weight and the input variables for prompt and non-prompt leptons in a number of different validation regions is discussed below. The modelling for prompt leptons in $Z \to \ell\ell$ and $t\bar{t}$ are discussed in Sections 7.4.2.3 and 7.4.2.4 and for non-prompt leptons in a non-prompt enriched $t\bar{t}$ region in Section 7.4.2.5. The simulation samples used to determine the modelling are the same as in Table 8.4 below.

7.4.2.1 Region definitions

Prompt and non-prompt validation regions are used to study the performance of the non-prompt BDT with data and simulation. $Z \to ee$ and $Z \to \mu\mu$ validation regions are defined to study the properties of prompt leptons. Z events decaying to leptons are very easily isolated by requiring the invariant mass of the leptons match the Z mass. They are thus used for prompt lepton calibration procedures. The calibration of non-prompt BDT working points for prompt muons is discussed in Section 7.5.

Separate $t\bar{t}$ validation regions are defined to study properties of prompt leptons from W decays and non-prompt leptons from decays of heavy quark hadrons. A validation region defined by an opposite charge sign electron-muon pair selection predominantly studies prompt leptons from $t\bar{t}$ in a different environment to Z events. The electron-muon selection is used to minimise the contribution from $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ opposite-sign events. A same-sign lepton selection is also used to study non-prompt leptons from $t\bar{t}$.

Table 7.8 describes the selection criteria for the three validation regions. The lepton definitions are as described in Table 7.6, with the exception of additional p_T and η requirements which are discussed below. An overlap removal procedure between reconstructed objects is applied, shown in Section 7.4.2.2 below.

For the non-prompt validation regions an extra loose lepton selection is defined, with a new corresponding overlap removal procedure. The standard overlap removal procedure of muons with jets removes a large number of non-prompt muons by giving precedence to jets close to muons. This is normally a useful attribute for analyses targeting prompt muons. To retain a large number of non-prompt leptons only a very loose isolation selection is applied and the jet-lepton overlap removal is reversed for muons. Full details of the procedure are discussed in Section 7.4.2.2 below.

Region	Selection criteria
$Z \rightarrow \ell^+\ell^-$	One same-flavour opposite-sign lepton pair $\lvert m(\ell^+\ell^-) - m_Z\rvert < 10$ GeV $p_T^{0,1} > 25$ GeV
2ℓ $t\bar{t}$ opposite-sign $e^{\pm}\mu^{\mp}$	One electron and one muon, with opposite charge $p_T^{0,1} > 25$ GeV $N_{\text{b-jets}} \geq 1$ $m(\ell^+\ell^-) > 40$ GeV
2ℓ same-sign	Two light leptons with same-sign charge One tight lepton and one loose lepton $p_T^0 > 25$ GeV, $p_T^1 > 15$ GeV $2 \leq N_{\text{jets}} \leq 4$, $N_{\text{b-jets}} \geq 1$ Central electrons, $\lvert\eta\rvert < 1.37$ $\lvert m(e^{\pm}e^{\pm}) - m_Z\rvert > 10$ GeV for $e^{\pm}e^{\pm}$ events

Table 7.8: Definitions of the validation regions used to investigate the modelling of the non-prompt BDTs. At least one lepton in each region must be matched to an object reconstructed by a single lepton trigger. The leptons are ordered by p_T; lepton 0 is leading and lepton 1 is subleading.

7.4.2.2 Overlap removal

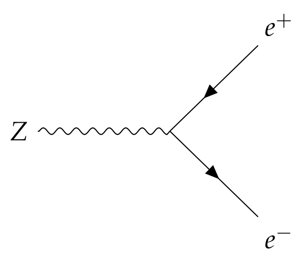
An overlap removal procedure is applied to the physics objects in this section, detailed in Table 7.9.

The removal is applied in the following order. An electron candidate within $\Delta R < 0.1$ of another electron candidate with higher p_T is removed. The remaining electron candidates that are within $\Delta R < 0.1$ of a muon are removed. If a jet candidate is within $\Delta R = 0.3$ of an electron then the jet is removed and if a muon and a jet are within $\Delta R = \min(0.4, 0.04{+}10[\text{GeV}]/p_T(\text{muon}))$ of one another, then the jet is kept and the muon is removed. The sliding overlap removal cut for muons results in high p_T muons being allowed to be nearer to jets than at low p_T.

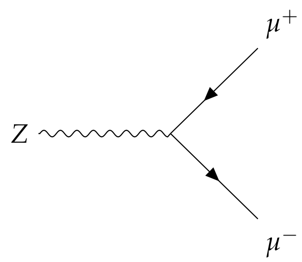
When an event has a *loose* lepton further procedures are applied, to ensure that loose leptons are not double counted with standard leptons and to retain non-prompt muons close to jets.

Keep	Remove	Condition
Electron	Electron (low p_T)	$\Delta R < 0.1$
Muon	Electron	$\Delta R < 0.1$
Electron	Jet	$\Delta R < 0.3$
Jet	Muon	$\Delta R <$min(0.4, 0.04 + 10[GeV]/p_T(muon))
Electron	Loose electron	$\Delta R < 0.1$
Muon	Loose electron	$\Delta R < 0.1$
Loose electron	Jet	$\Delta R < 0.4$
Muon	Loose muon	$\Delta R < 0.3$
Loose muon	Jet	$\Delta R < 0.4$
Loose muon	Electron	$\Delta R < 0.3$

Table 7.9: A summary of the overlap removal procedure. The procedure is applied in order. Only the procedures involving loose electrons or muons are applied if either are specified in an event.



(a)



(b)

Figure 7.14: Feynman diagrams for (a) $Z \to ee$ and (b) $Z \to \mu\mu$.

7.4.2.3 Prompt $Z \to \ell\ell$ events

Feynman diagrams of Z events decaying to two opposite sign leptons are shown in Figure 7.14. Nominally only two same-flavour opposite-sign leptons define such an event.

Figure 7.15 shows the non-prompt BDT weights in the $Z \to ee$ and $Z \to \mu\mu$ validation regions. There is large mismodelling between data and MC at values of the BDT greater than -1, with the mismodelling being worse for muons than for electrons.

The input variables to the non-prompt BDT for the $Z \to ee$ and $Z \to \mu\mu$ validation regions are shown in Figures 7.16 and 7.17 for the leading lepton respectively. Both leading and subleading leptons are prompt and thus have similar distributions. Therefore the subleading distributions are retained for brevity. The most mismodelled variables appear to be the impact parameter based variables, IP2D and IP3D. The modelling of these variables in the $Z \to \mu\mu$ region is particularly poor. Track isolation is also poorly modelled, with data showing more non-isolated prompt leptons than MC. One potential reason for this is due to vertex density mismodelling. Further details discussing the source of this mismodelling can be found in Section 7.4.3.

7.4.2.4 Prompt $e^{\mp}\mu^{\pm}$ $t\bar{t}$ events

The modelling of the non-prompt BDT for prompt leptons is also studied with the opposite-sign, opposite-flavour validation region. This region is highly dominated by prompt dilepton $t\bar{t}$ events. An example Feynman diagram of such an event is shown in Figure 7.18. Both W bosons from $t\bar{t}$ decay leptonically, with the event also nominally containing two b-jets.

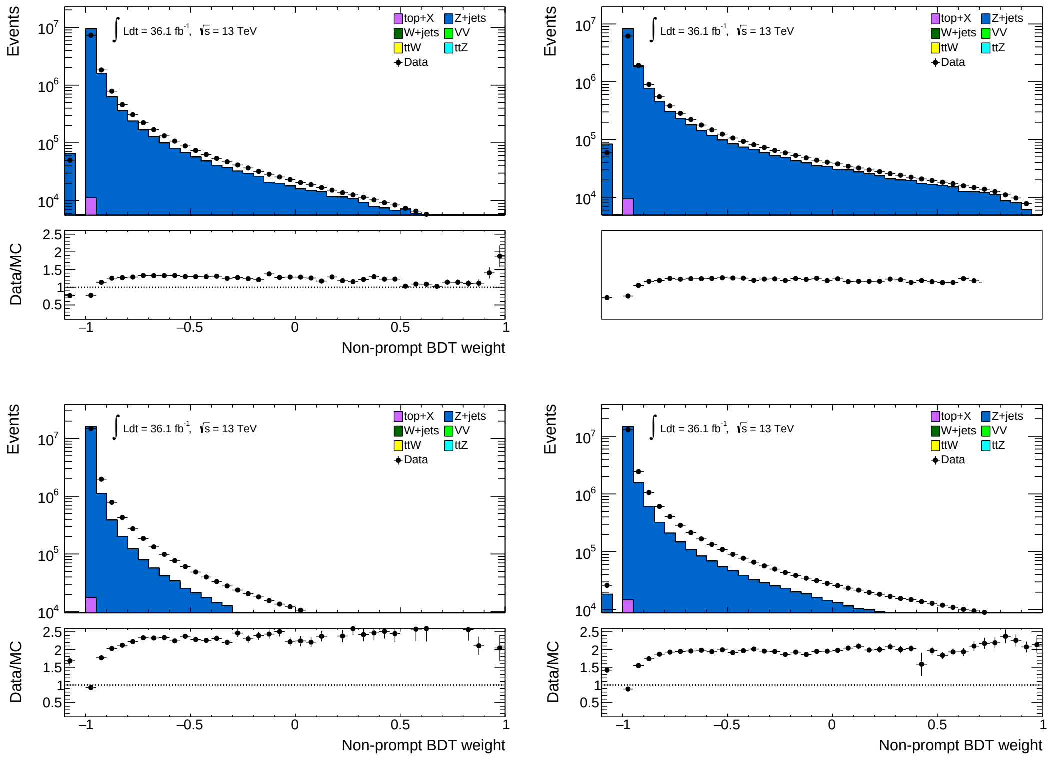


Figure 7.15: Distributions of the non-prompt BDT in data (black circles) and simulation (filled histograms) in the $Z \to ee$ (top) and $Z \to \mu\mu$ (bottom) validation regions. The left distribution is for the leading lepton and the right subleading.

The number of MC non-prompt leptons from $t\bar{t}$ in this validation region is found to be less than 0.1%. Figure 7.19 shows the non-prompt lepton distributions for the leading and subleading lepton in this region, as well as the individual electron and muon distributions. Very similar behaviour to the Z validation region is observed. The prompt lepton calibration procedure makes the assumption that calibrations for prompt leptons are valid in any event, independent on environment. The similar behaviour of the non-prompt BDTs in the quite different Z and $t\bar{t}$ environments go some way to verifying that assumption.

7.4.2.5 $t\bar{t}$ events with a non-prompt lepton

$t\bar{t}$ events also an provide an excellent non-prompt lepton validation region. Inclusively requiring two same-sign leptons generally results in one prompt lepton and one non-prompt lepton from $t\bar{t}$. Figure 7.20 is a Feynman diagram in which a $t\bar{t}$ decay can produce two same-sign negatively charged leptons. One is produced from a W and is hence prompt, and one from the semileptonic decay of a b quark and hence is non-prompt. Two same-sign positively-charged leptons can be produced from a similar diagram. This region also has a small, subdominant contribution from W+bb where the W decays leptonically and a non-prompt lepton is reconstructed from one of the b-jets. However, the final leptonic state of a prompt and a non-prompt lepton is the same as non-prompt and does not impact the modelling expected for non-prompt leptons.

The non-prompt BDT weights for the two same-sign electron and two same-sign muon validation regions are shown in Figure 7.21 and Figure 7.22 respectively. These validation regions

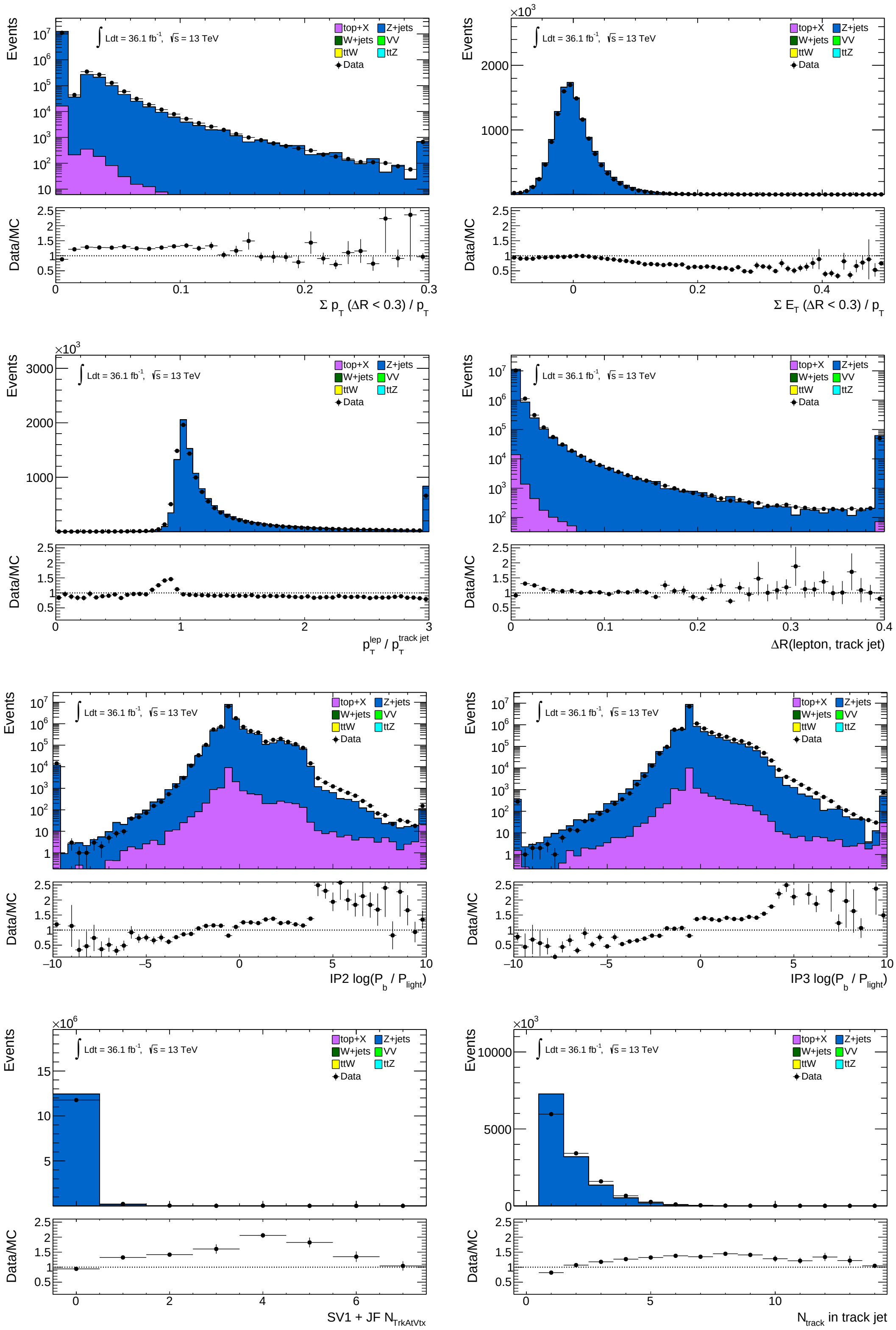


Figure 7.16: Distributions of the variables in the non-prompt BDT in data (black circles) and simulation (filled histograms) in the $Z \to ee$ validation region for the leading electron.

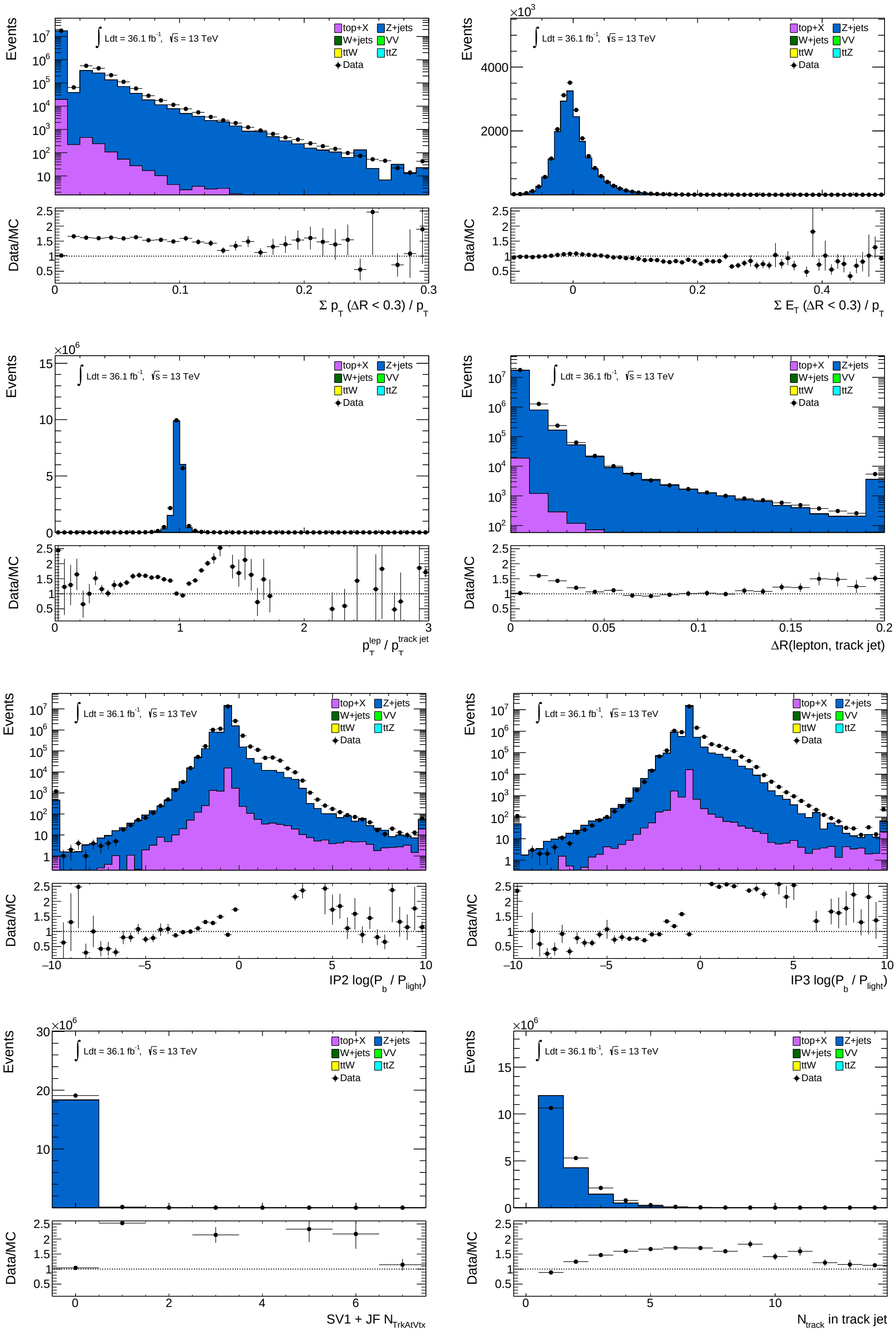


Figure 7.17: Distributions of the variables in the non-prompt BDT in data (black circles) and simulation (filled histograms) in the $Z \to \mu\mu$ validation region for the leading electron.

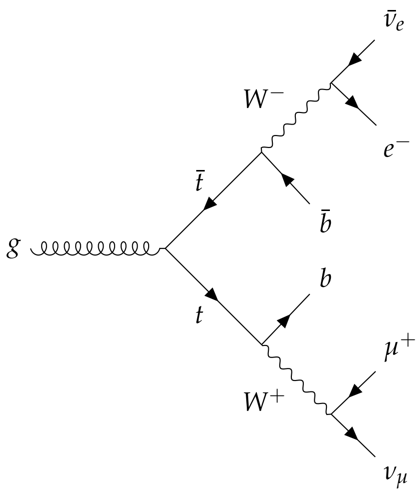


Figure 7.18: A Feynman diagram of an opposite-sign $t\bar{t}$ decay with one electron and one muon.

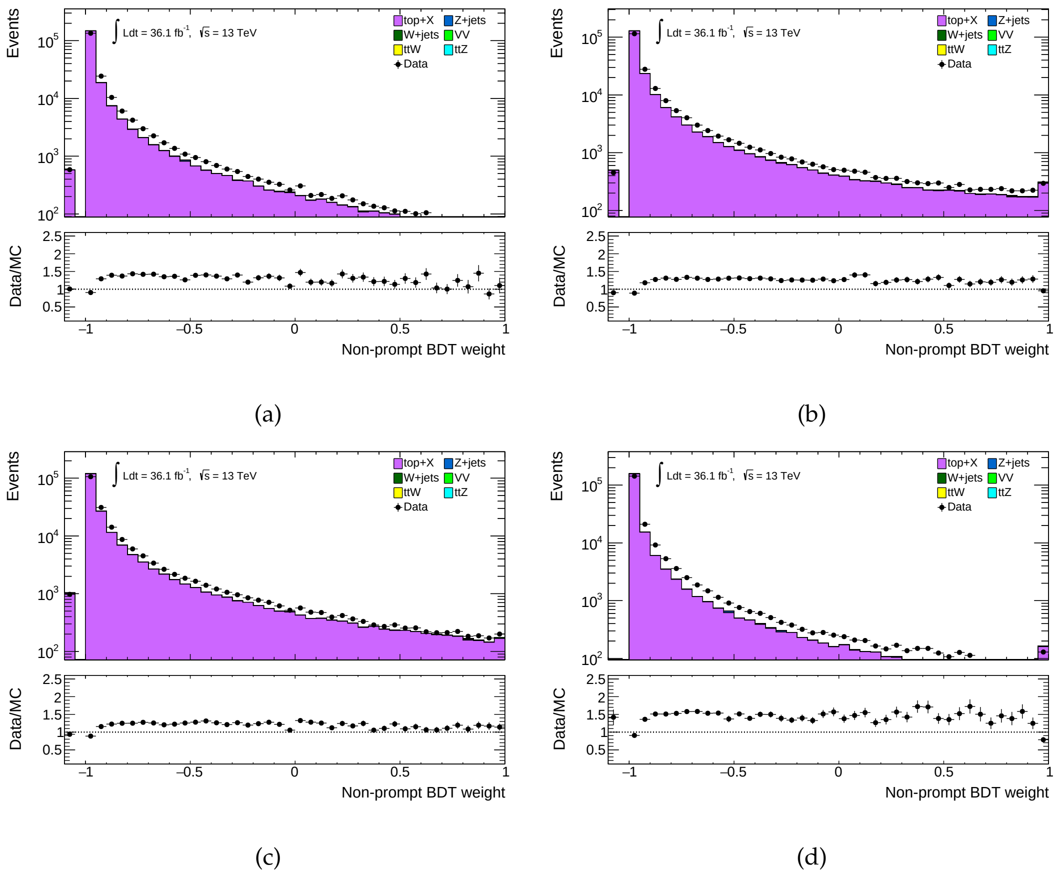


Figure 7.19: Distributions of the non-prompt BDT in data (black circles) and simulation (filled histograms) in the opposite-sign $t\bar{t}$ validation region with one electron and one muon, for (a) the leading lepton, (b) the subleading lepton, (c) the electron and (d) the muon.

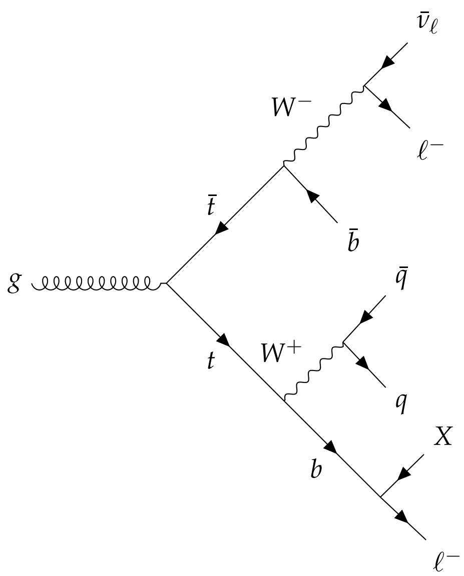


Figure 7.20: Feynman diagram of a non-prompt $t\bar{t}$ decay. X represents the remnants of the semileptonic B hadron decay produced from the b-quark. The non-prompt lepton produces a lepton with the same sign as the W^-.

reduce the subleading p_T to 15 GeV to maximise the number of non-prompt leptons.

A non-prompt lepton is embedded within a b-jet where it has to share the initial b-quark energy with the remainders of the B hadron decay and other particles created by the b-quark fragmentation process. The non-prompt leptons have a soft p_T spectrum because the underlying spectrum of the B hadron p_T is soft, and also because leptons with lower p_T values have a higher probability to satisfy isolation criteria. Since the prompt leptons are produced in W decays they carry significant boost due to the high mass of the W boson, which itself has a boost from the top quark decay. Thus, in a same-sign $t\bar{t}$ decay, the prompt lepton has a harder p_T spectrum than the non-prompt lepton. Therefore, the leading lepton is more likely to be prompt than non-prompt and vice-versa for the subleading lepton. This is reflected in Figures 7.21 and 7.22. However, there are cases where the non-prompt lepton has higher p_T than the prompt. This can be seen in the non-prompt BDT weights for the tight lepton and the loose lepton, irrespective of p_T. In this case clearly the tight lepton is prompt and the loose lepton is non-prompt.

The modelling of the non-prompt leptons is very good, for both electrons and muons. The input variable distributions for the loose lepton in the two same-sign electrons and muons are shown in Figures 7.23 and 7.24, where good agreement is also seen.

7.4.3 Prompt lepton mismodelling

From Section 7.4.2 it is clear that prompt leptons are mismodelled in MC. There are a number of reasons why. This section focuses on the contribution of mismodelling due to vertex density discrepancies between MC and data.

7.4.3.1 Pileup reweighting

Typically in ATLAS, analyses reweight the pileup, μ, in the MC to that of the data it is compared to. Pileup vertices can effect the shape of certain distributions due to differences such as a

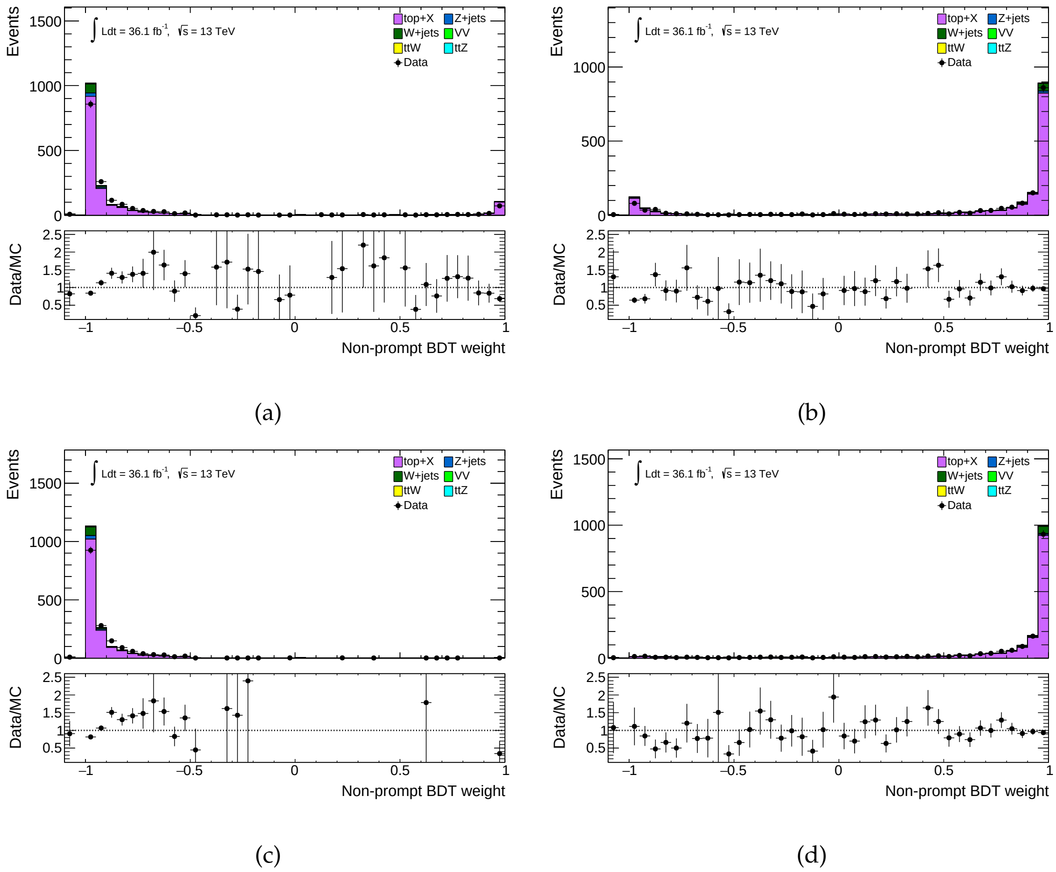


Figure 7.21: Distributions of the non-prompt BDT in data (black circles) and simulation (filled histograms) in the 2ℓ same-sign non-prompt validation region with two electrons for (a) the leading electron, (b) the subleading electron, (c) the tight electron and (d) the loose electron.

change in the track activity in the event. Comparing MC and data with different pileup for these distributions could show signs of data/MC mismodelling, even if the distribution is in fact perfectly modelled internally.

The pileup profile in MC is modelled with PYTHIA 8. A minimum bias (MinBias) tune is applied. However, this minimum bias tune is known to be too "hard". That is, simulation is best compared to data with a smaller value of pileup than data itself. A linear scaling is applied, with a scale factor of 1.09 for the MC simulation samples used in this thesis. Figure 7.25 shows the pileup distributions with and without pileup reweighting applied in the $Z \to \mu\mu$ validation region. The reweighted MC μ profile scaled by 1.09 would match that of data. All data plots in this thesis have pileup reweighting applied, unless otherwise stated.

7.4.3.2 Vertex density mismodelling in Z events

The pileup reweighting can be interchanged with reweighting to the number of vertices in an event, N_{vtx}, or the vertex density, ρ_{vtx}. Both of these reweighting schemes aim to model the underlying structure of the proton-proton collisions occurring in ATLAS.

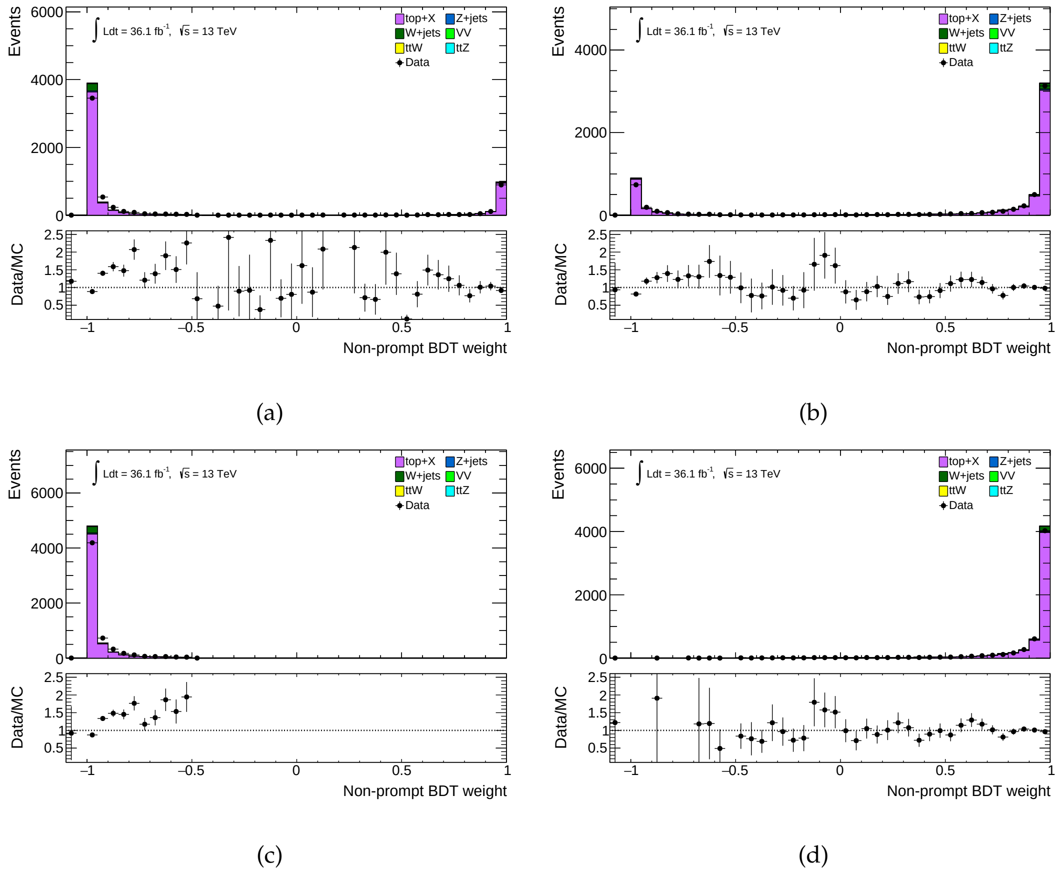


Figure 7.22: Distributions of the non-prompt BDT in data (black circles) and simulation (filled histograms) in the 2ℓ same-sign non-prompt validation region with two muons for (a) the leading muon, (b) the subleading muon, (c) the tight muon and (d) the loose muon.

The vertex density is defined as the longitudinal density of vertices:

$$\rho_{\text{vtx}} = \frac{\mu}{\sqrt{2\pi\sigma^2_{\text{beam-z}}}} e^{-\frac{1}{2}\left(\frac{z_{\text{PV}}-z_{\text{beam}}}{\sigma_{\text{beam-z}}}\right)^2}. \tag{7.3}$$

Here z_{PV} and z_{beam} are the longitudinal positions of the primary vertex and the centroid of the luminous region respectively and $\sigma_{\text{beam-z}}$ the standard deviation of the longitudinal beam position. The vertex density is hence a Gaussian distribution of the longitudinal shape of the beamspot, scaled by μ. The vertex density profile for 2015 and 2016 data in the $Z \to \mu\mu$ is shown in Figure 7.26. For values greater than $\rho_{\text{vtx}} > 0.29\ \text{mm}^{-1}$, MC does not describe data at all.

The density of the vertices longitudinally is an important effect in the mismodelling of the non-prompt BDT. The larger the vertex density, the more vertices and the more likely pileup tracks are used by the non-prompt BDT. The isolation variables used as features in the non-prompt BDT are calculated from tracks with $p_T > 1$ GeV within a longitudinal distance to the primary vertex of $|z_0 \sin\theta| < 3$ mm. The vertex density distribution for 2015 and 2016 data in Figure 7.26 is maximum at greater than $0.5\ \text{mm}^{-1}$, or 1 vertex every 2 mm on average. This implies that, on average, there will be one pileup vertex with tracks that could be used in the calculation of the isolation. Typically pileup tracks are low p_T and not necessarily near to the lepton from the primary vertex but this effect is not to be discounted.

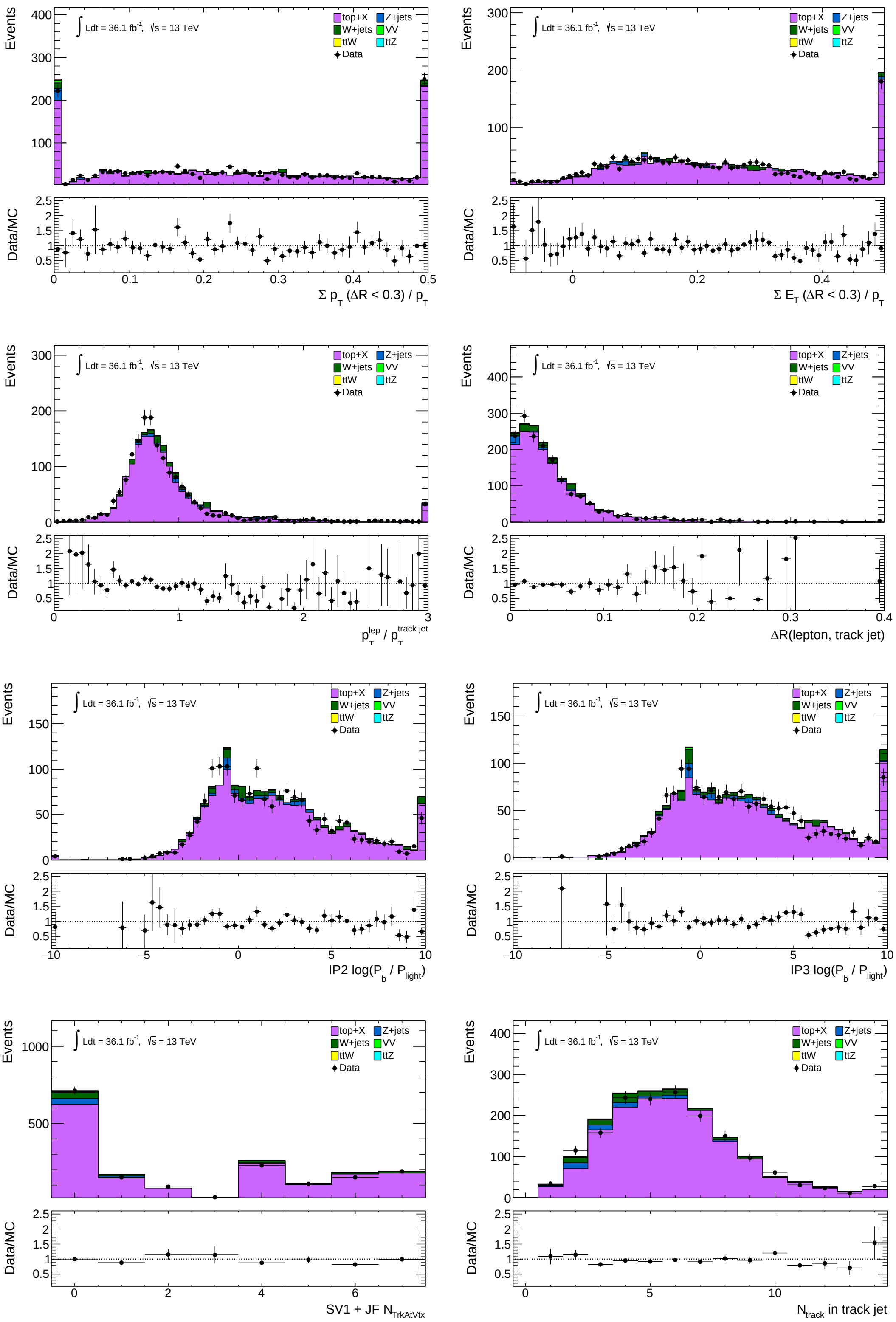


Figure 7.23: Distributions of the variables in the non-prompt BDT in data (black circles) and simulation (filled histograms) in the 2ℓ same-sign non-prompt validation region with two electrons.

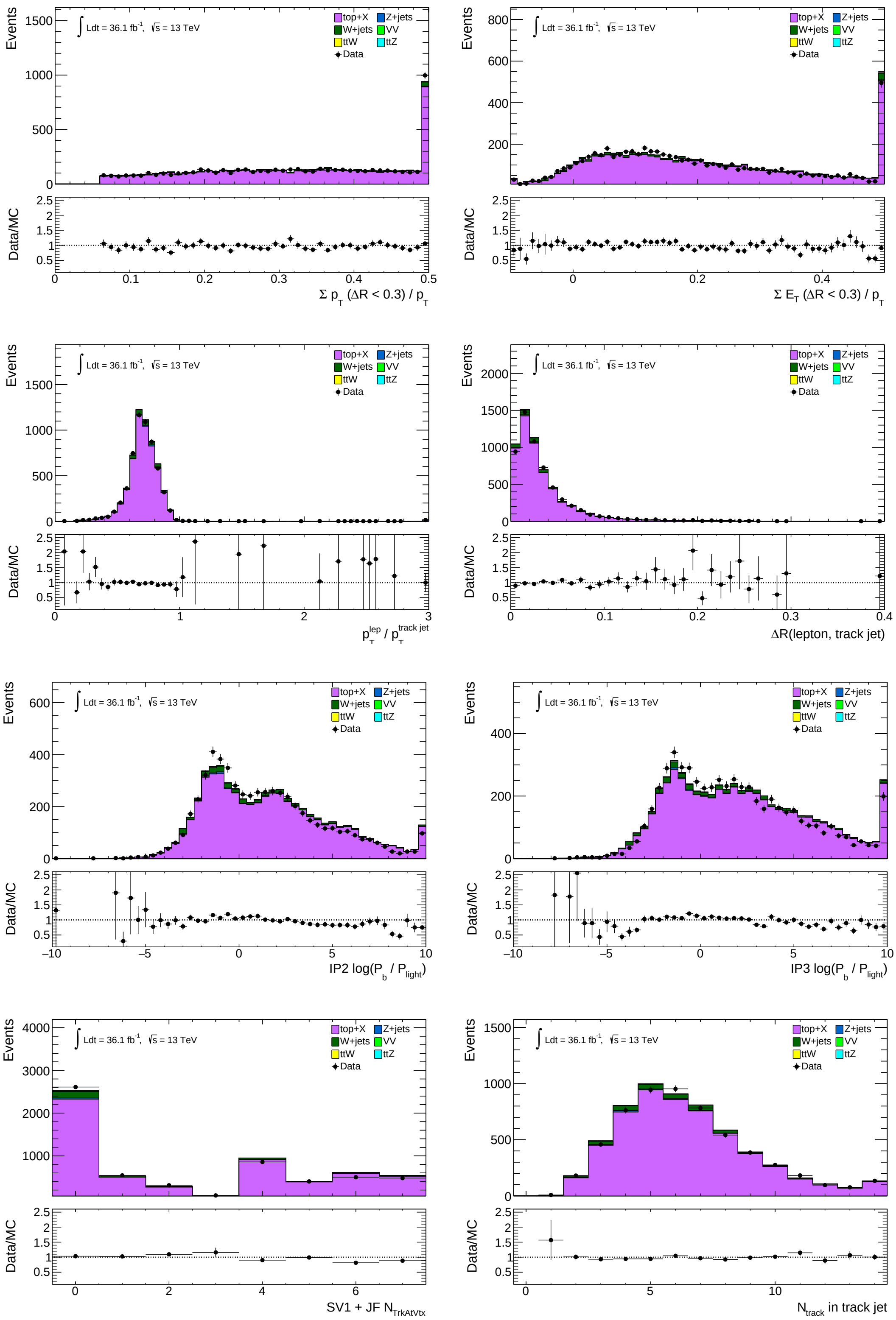


Figure 7.24: Distributions of the variables in the non-prompt BDT in data (black circles) and simulation (filled histograms) in the 2ℓ same-sign non-prompt validation region with two muons.

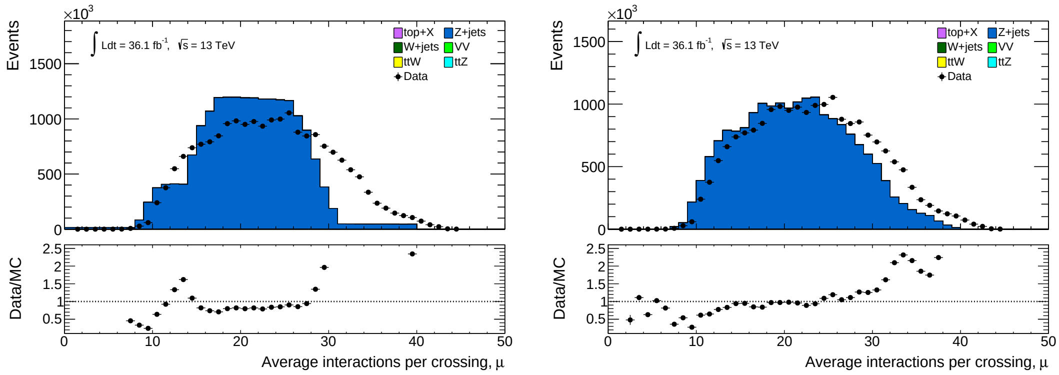


Figure 7.25: Pileup distributions in data (black circles) and simulation (filled histograms) when (a) not including and (b) including pileup reweighting in the $Z \to \mu\mu$ validation region.

For large vertex densities it is likely pileup tracks become more and more important and this itself is likely to degrade performance, as with high values of pileup.

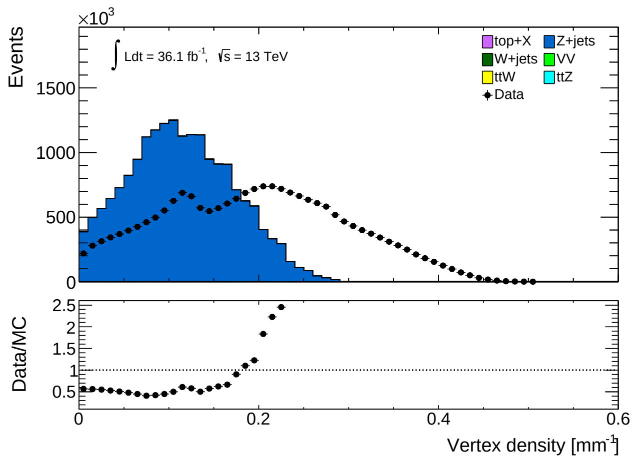


Figure 7.26: Vertex density profiles in 2015 and 2016 data (black circles) and simulation (filled histograms) in the $Z \to \mu\mu$ validation region.

The beamspot variables that are used to calculate the vertex density are shown in Figure 7.27. The difference in position between the centre of the beamspot and the primary vertex is smaller in data than in MC. For the same μ this implies a larger vertex density. The variance is also smaller in data than MC, with only a single value being used in MC due to uncertainty of how the 2016 beamspot profile would look.

One needs MC to cover all the data distribution of a variable to reweight properly. Reweighting just increases or decreases the number of events with a certain value of vertex density and cannot actually increase the vertex density itself. From Figure 7.26 it is clear that there are regions of high vertex density that MC does not cover, already giving warnings about the mismodelling of the non-prompt BDT. To this end a cut on $\rho_{\mathrm{vtx}} < 0.29\ \mathrm{mm}^{-1}$ is applied to data and simulation here, which reduces the integrated luminosity to 29.3 fb^{-1}. Vertex density weights are calculated by taking the values of the ratio of inclusive data and MC simulation and applying these weights to MC. The effect of this vertex density reweighting is shown in Figure 7.28, compared to the distribution of vertex density with pileup reweighting applied.

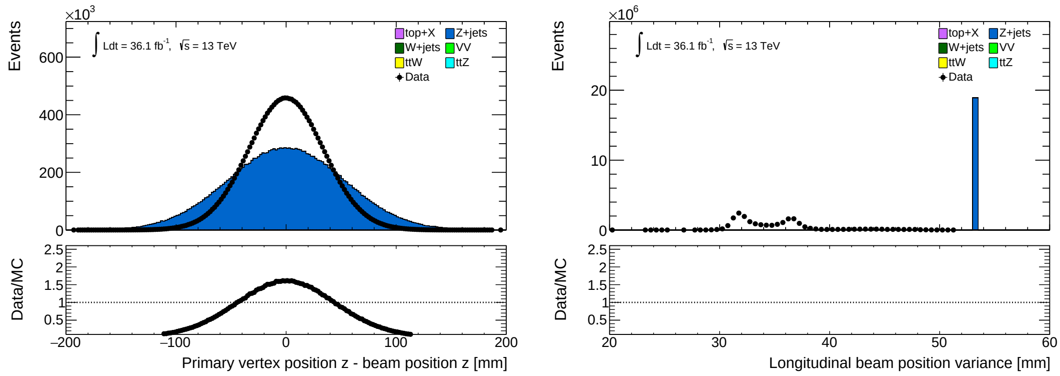


Figure 7.27: Distributions of variables used as input to the calculation of vertex density in data (black circles) and simulation (filled histograms) in the $Z \to \mu\mu$ validation region. The longitudinal beamspot variance is the square of the longitudinal beamspot position uncertainty, $\sigma^2_{\text{beam-z}}$.

The reweighting is not perfect due to the weights not being produced from the $Z \to \mu\mu$ region, but MC still describes the data well.

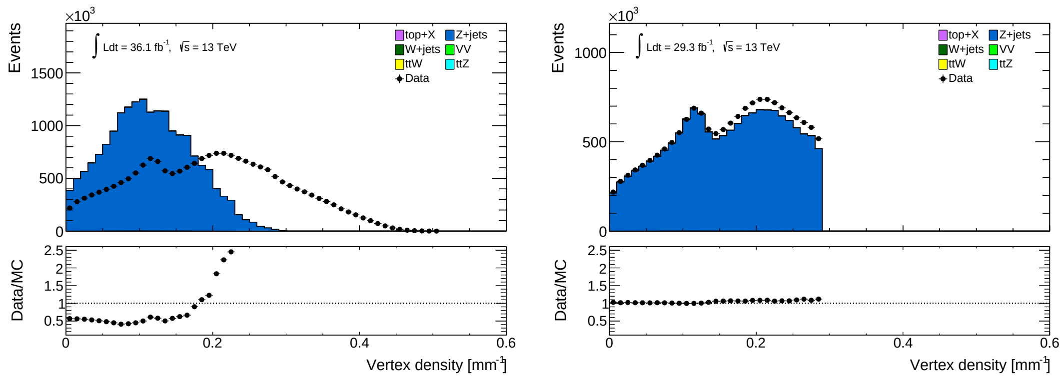


Figure 7.28: Vertex density profiles in 2015 and 2016 data (black circles) and simulation (filled histograms) with (a) pileup reweighting and (b) vertex density reweighting applied in the $Z \to \mu\mu$ validation region.

The non-prompt BDT weights for the two reweighting schemes are shown for the leading lepton in the $Z \to ee$ and $Z \to \mu\mu$ validation regions in Figure 7.29. The differences in the distributions between the leading and subleading lepton are similar in both regions. The mismodelling is reduced by half for both electrons and muons. Simply removing high ρ_{vtx} events will contribute to this difference.

The comparison of pileup and vertex density reweighting on the input variables to the non-prompt BDT are shown in Figures 7.30 and 7.31 for the leading electron in $Z \to ee$ and Figures 7.32 and 7.33 for the leading muon in $Z \to \mu\mu$. In almost all distributions the input variable modelling is improved. The exception is for calorimeter isolation, where a pileup specific correction is made. Removing events with large vertex densities changes the pileup distribution and thus the pileup correction no longer works as well at improving the modelling. Removing these events *decreases* the total inclusive pileup distribution. The improvement in modelling when removing high vertex densities and reweighting is also likely to be due to reduced pileup vertices.

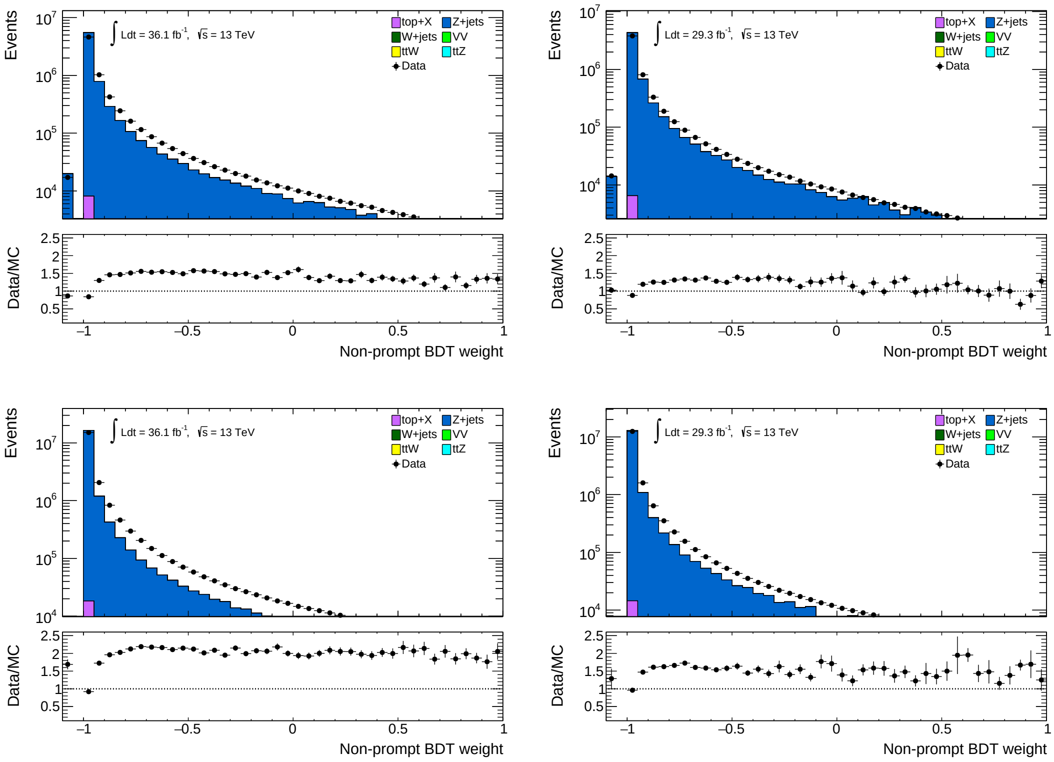


Figure 7.29: Non-prompt BDT weights in 2015 and 2016 data (black circles) and simulation (filled histograms) for the leading lepton in the $Z \to ee$ (top) and $Z \to \mu\mu$ (bottom) validation regions with pileup reweighting (left) and vertex density reweighting (right) applied.

This vertex density mismodelling issue requires the removal of almost 20% of the total data and is thus not pursued further. Pileup density reweighting, without any reduction of data, is applied to any data distributions in this thesis.

7.5 Calibrating the non-prompt BDT for prompt muons

This section focuses on the data calibration of a non-prompt BDT working point for prompt muons. The non-prompt BDT working point is the same for electrons and muons:

$$\texttt{isoLoose \&\& Non-prompt BDT} < -0.5$$

This working point is optimised for the $t\bar{t}H$ multilepton analysis. The analysis itself is discussed in Chapter 8 and the choice of working point and the optimisation procedure in Section 8.5.2.

This working point is designed to replace a traditional isolation working point. The working point includes a loose isolation requirement to a flat cut on the BDT to ensure there is no unintended non-isolated phase space that the BDT may have missed.

In Section 7.5.1 the Z tag and probe method used to calibrate prompt muons is introduced. In Section 7.5.2 data/MC isolation scale factors are determined for the non-prompt BDT working point and in Section 7.5.3 trigger scale factors are determined for the same working point.

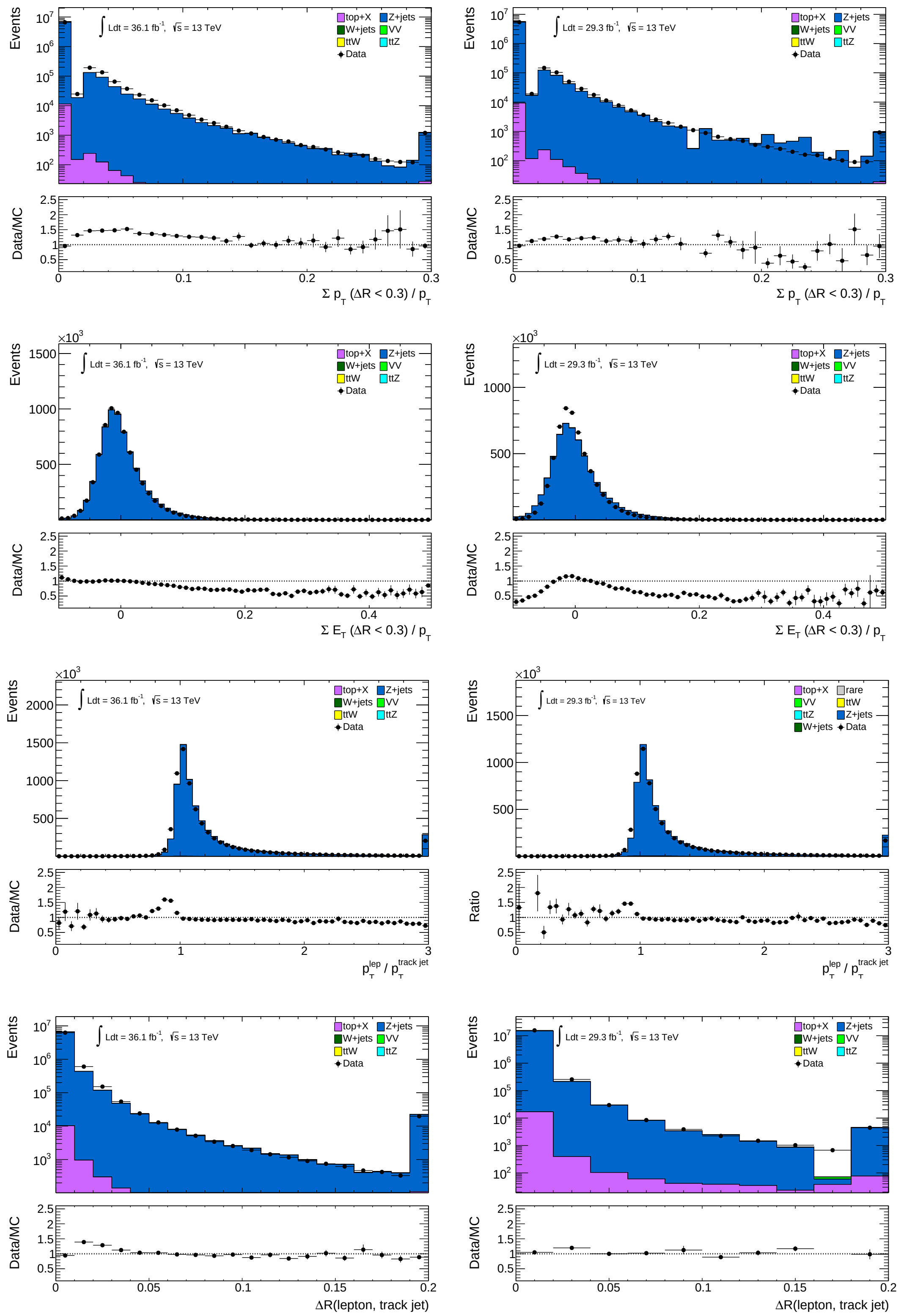


Figure 7.30: Distributions of the isolation variables, $p_T^{\text{lepton}} / p_T^{\text{track jet}}$ and ΔR(lepton, track jet) used as an input to the non-prompt BDT in data (black circles) and simulation (filled histograms) for pileup reweighted (left) and vertex density reweighted (right) for the leading lepton in the inclusive $Z \to ee$ validation region.

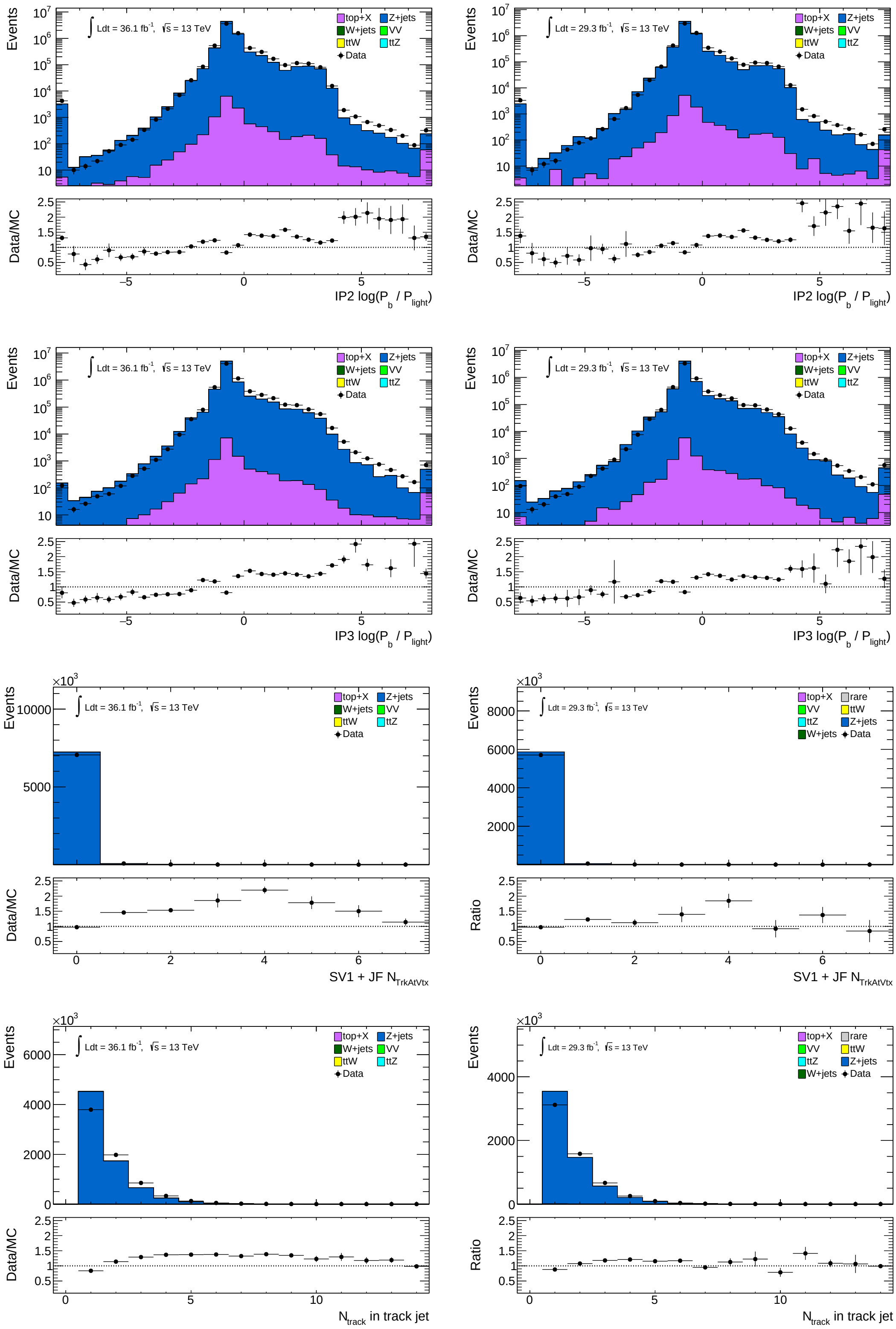


Figure 7.31: Distributions of the b-tagging variables and N_{track} used as an input to the non-prompt BDT in data (black circles) and simulation (filled histograms) for pileup reweighted (left) and vertex density reweighted (right) for the leading lepton in the inclusive $Z \to ee$ validation region.

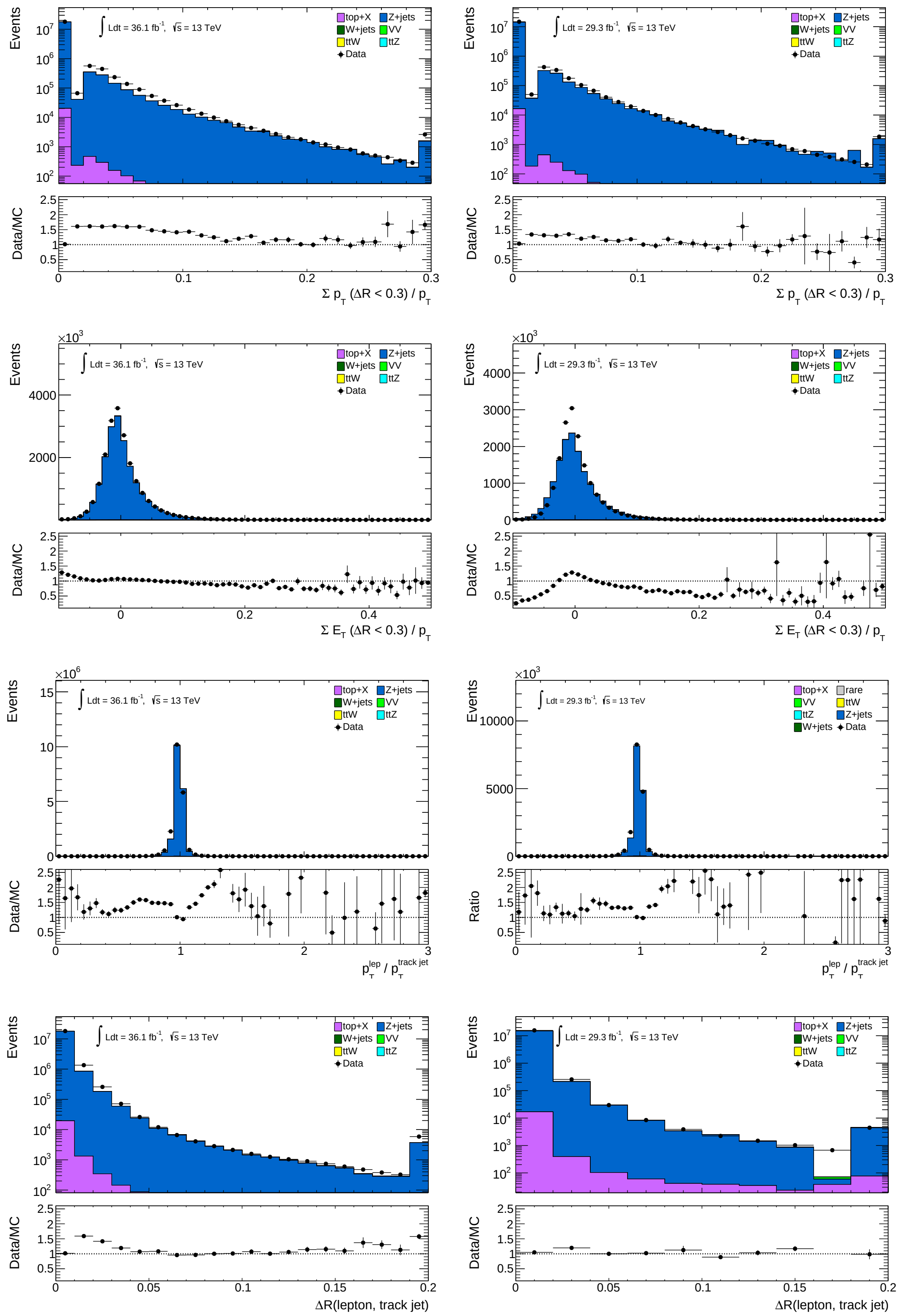


Figure 7.32: Distributions of the isolation variables, $p_T^{\text{lepton}} / p_T^{\text{track jet}}$ and ΔR(lepton, track jet) used as an input to the non-prompt BDT in data (black circles) and simulation (filled histograms) for pileup reweighted (left) and vertex density reweighted (right) for the leading lepton in the inclusive $Z \rightarrow \mu\mu$ validation region.

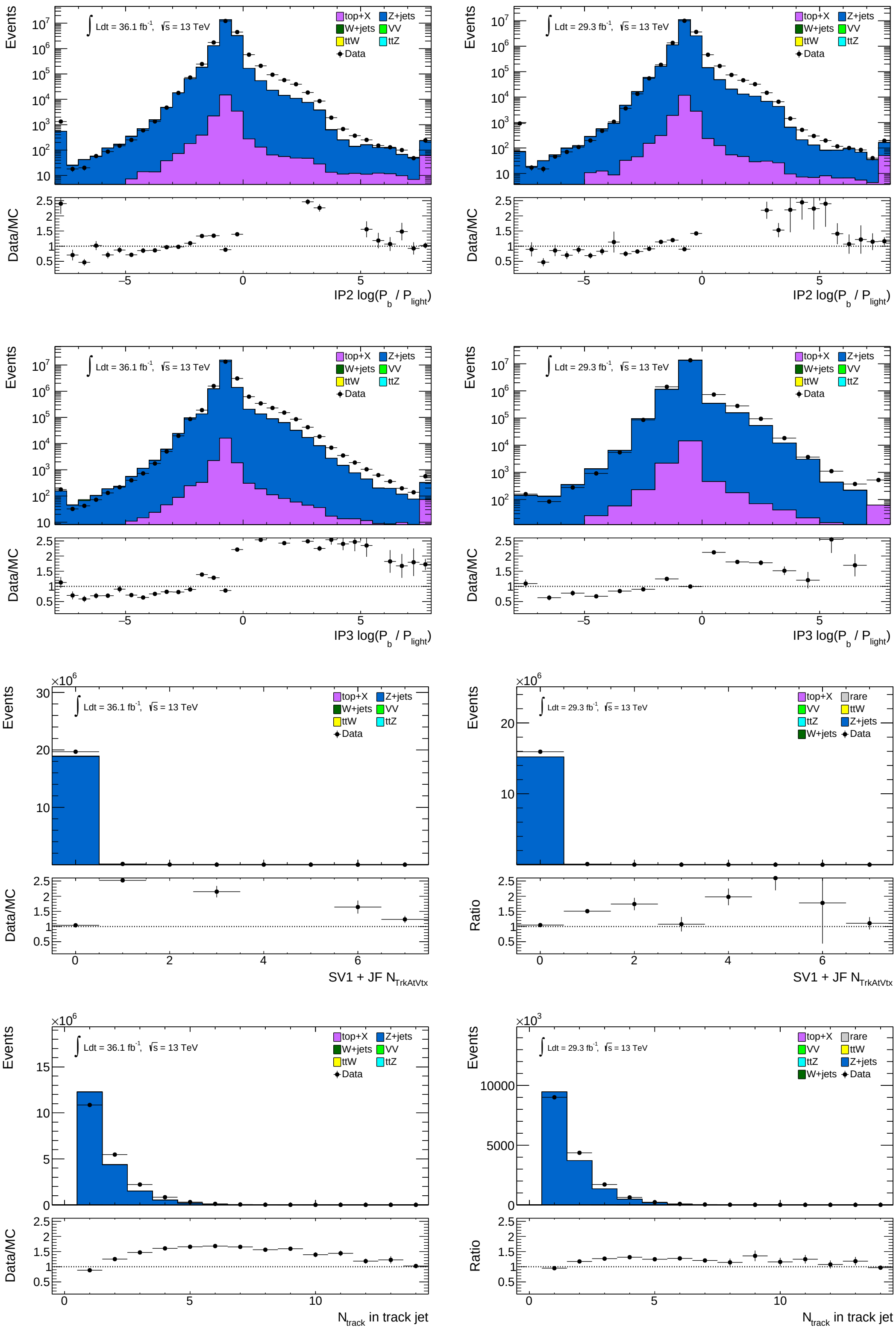


Figure 7.33: Distributions of the b-tagging variables and N_{track} used as an input to the non-prompt BDT in data (black circles) and simulation (filled histograms) for pileup reweighted (left) and vertex density reweighted (right) for the leading lepton in the inclusive $Z \to \mu\mu$ validation region.

7.5.1 The tag and probe method

The efficiency of the non-prompt BDT working point is measured using the tag and probe method. An efficiency, ϵ, expresses a probability for some object to pass a given selection

$$\epsilon = \frac{\text{number of objects passing selection}}{\text{total number of objects}}. \tag{7.4}$$

Such efficiencies are, in general, easily determined in MC. However, MC is but the best approximation of the real data and detector effects. One needs well understood and easily isolated particle events in which to compare the MC efficiency with that of data, where one knows, with some certainty, that the data events are those that are being compared to MC. Z events decaying to light leptons fit these criteria.

Z events can be selected with almost total purity by applying a Z mass window selection on the invariant mass of the two leptons produced in the decay. The method requires that one tightly selected lepton fires a single lepton trigger and is labelled the *tag* lepton. The tag lepton is required to have a very low probability of misidentification. To complete the invariant mass event selection, another same-flavour, opposite-sign, loosely selected lepton is required in the event and is labelled the *probe*. The probe is loosely selected in order to measure the efficiency of applying the tighter selection which is that being calibrated. The probe is thus an independent entity, known to be a prompt lepton, which is used to measure the efficiency in data and MC. The difference of the efficiency in data and MC is a scale factor that is applied to MC to calibrate it to that of data. It should be noted that the probe can also fire the single lepton trigger, and then the role of the two leptons is reversed and another data/MC efficiency measurement can be made.

There are not many probes with low p_T (less than 10 GeV for example) from Z events. In such cases other, lower mass resonances can be used such as the decays of J/ψ or Υ. The scale factor measurement at high and low p_T can then be merged from the different resonances. The non-prompt BDT is only designed to be used in high p_T environments and thus only Z events are used in the calibration.

The efficiency of the probe passing the tight selection can be measured as a function of different variables, both in data and MC. The difference between the two efficiencies is taken as the scale factor that is applied to MC used to calibrate prompt leptons. Typically isolation scale factors are derived as a function of p_T and/or η. Isolation scale factors for the non-prompt BDT working point are discussed below. The data/MC efficiency dependence on other angular and kinematic event variables are used to investigate extra dependencies and to determine systematic uncertainties.

7.5.2 Isolation scale factors

The total scale factor, ϵ, applied on MC for light leptons is as follows:

$$\epsilon = \epsilon_{\text{ID}} \times \epsilon_{\text{TTVA}} \times \epsilon_{\text{isolation}} \times \epsilon_{\text{trigger}}. \tag{7.5}$$

Firstly scale factors for the lepton identification working points are derived. Then scale factors for the impact parameter cuts are applied on top of the identification working points. In this way there are three impact parameter scale factors applied; those derived for each of the three identification working points. Then isolation scale factors are applied, and finally trigger scale factors are applied. One can imagine the combinatorics required to produced trigger scale factors for every possible working point. Luckily, for muons, trigger scale factors do not (in general) depend on the isolation selections applied before it. The same is not necessarily true for electrons.

The non-prompt BDT working point is designed to totally replace the standard isolation working point and hence to calibrate prompt leptons "isolation" scale factors need to be calculated. The calibration of the muon non-prompt BDT working point is discussed below. The dependence of the trigger scale factors on this working point is checked in Section 7.5.3, to determine if the trigger scale factors are non-prompt BDT independent as well as isolation independent.

7.5.2.1 Muon and event selection

Two opposite sign muons are required. They are required to pass the selections detailed in Table 7.10. The muon working point optimised for $t\bar{t}H$ multilepton uses muons passing the `Loose` identification requirement. However, muon isolation scale factors have negligible dependence on the quality requirement so `Medium` muons are required as the probe. Variations of selecting `Loose` and `Tight` probes are used to determine a systematic uncertainty on the probe quality requirement. Therefore the isolation scale factors are valid for all three muon quality requirements with standard impact parameter cuts applied. Only muons with $p_T > 10$ GeV are calibrated. The prompt leptons from $t\bar{t}H$ events rarely have such small transverse momentum and 10 GeV is thus used as the cut-off value.

	Tag	Probe
p_T [GeV]	$> 1.05 \times p_T^{\text{trigger}}$	> 10
$\lvert\eta\rvert$	< 2.5	< 2.5
PID	`Medium`	`Medium`
$\lvert d_0/\sigma_{d_0}\rvert$	< 3	< 3
$\lvert z_0 \sin\theta\rvert$ [mm]	< 0.5	< 0.5
Isolation	`FixedCutTightTrackOnly`	None

Table 7.10: The muon tag and probe selections.

Table 7.11 details the nominal event selections. Pairs of muons are selected with combined invariant mass within 10 GeV of m_Z. The tag muon is required to match the highest unprescaled single muon trigger in the time period: `HLT_mu20_iloose_LIMU15` in 2015, `HLT_mu24_ivarmedium` in early 2016 and `HLT_mu26_ivarmedium` in late 2016. The muon p_T is required to be greater than 1.05 times the trigger threshold. Additional event cuts on ΔR between the tag and the probe and ΔR between the probe muon and the nearest calorimeter jet are used.

A simple background subtraction of the number of same-sign muons in the same Z mass window is applied. The production of non-prompt muons from QCD is roughly charge symmetric (same number of positive and negatively charged non-prompt muons) so a subtraction of the same-sign non-prompt events in the Z mass window roughly equates to the number of opposite sign non-prompt events. This background subtraction dominates at low muon p_T, where one would expect an increased number of non-prompt muons.

Table 7.11 also details the selections used as systematic variations. Systematic uncertainties are applied to the isolation scale factors in bins of p_T. The size of the uncertainty is taken by determining the largest variation with respect to nominal from the up or down variations and are then symmetrised. All the systematic variations apart from the number of vertices in an event, N_{vtx}, are common to the standard muon isolation working points as discussed in [106]. The N_{vtx} variation is added as an extra systematic to this working point, due to a scale factor dependence that is shown below.

Variation	Nominal	Up	Down
Background subtraction	$1.0 \times \texttt{SC}$	$2.0 \times \texttt{SC}$	$0.5 \times \texttt{SC}$
Z mass cut	< 10 GeV	< 5 GeV	< 20 GeV
$\Delta R(\text{probe, tag})$	$\Delta R > 0.3$	$\Delta R > 0.2$	$\Delta R > 0.5$
Tag isolation	Loose	All isolation WP	
Probe quality	`Medium`	`Tight`	`Loose`
$\Delta R(\text{probe, cal.jet})$	$\Delta R > 0.4$	$\Delta R > 0.5$	$\Delta R > 0.3$
N_{vtx}	-	$N_{\text{vtx}} < 15$	$N_{\text{vtx}} \geq 15$

Table 7.11: Definition of the event and lepton selections used to isolate $Z \to \mu\mu$ decays. The up and down variations used to define the systematic uncertainties of the scale factors are also detailed. All isolation working points are determined to calculate a systematic uncertainty due to the isolation requirement on the tag muon. SC stands for "same-charge".

7.5.2.2 Scale factor dependencies

The data and MC efficiencies and efficiency scale factors of the non-prompt BDT working point for full 2015 and 2016 data are shown in Figure 7.34 below as a function of p_T, η, $\Delta R(\text{probe, cal. jet})$, N_{jet}, pileup μ and N_{vtx}.

There is some dependence on both the efficiency and the scale factor with p_T; there is a turn-on trend in the efficiency, occurring from low to high p_T. This inefficiency at low p_T is attributed loosely to pileup events. The isolation variables are relative to inverse lepton p_T, so pileup tracks collected by the isolation cone for low p_T leptons have more of an impact than for high p_T. Pileup tracks can also affect the other input variables similarly, as discussed in the context of vertex density mismodelling above. The efficiency scale factor in p_T also has a trend: the value is approximately 0.92 for $10 < p_T < 15$ GeV and averaging at 0.98 to 0.99 for higher p_T leptons. A number less than unity corresponds to the efficiency in data being lower than that of MC.

The large efficiency scale factor at low p_T can be studied by comparing the values obtained with 2015 only data and 2015+2016 data, as shown in Figure 7.35. The scale factor in the same $10 < p_T < 15$ GeV bin is approximately 0.96. Both results employ the same background subtraction, discounting the entire discrepancy being due to a non-prompt background. Likely contributors are due to the differences in both the distributions of pileup and vertex density between 2015 and 2016, with MC poorly describing the vertex density in 2016. Another indication of this is in the clear dependence of the data/MC efficiency for N_{vtx}, moving further from unity at larger values. This dependence is less clear in the distribution of pileup μ. Comparing efficiencies of N_{vtx} for 2015 and 2015+2016 data in Figure 7.36 goes some way to support this: the data/MC efficiency for 2015 is not nearly as bad as 2015+2016 for the full range of N_{vtx}, despite the low statistics at large N_{vtx} in 2015. One difference between the two datasets, hidden within the mismodelling, is the vertex density. The vertex density is, in general, largest at large N_{vtx} where the mismodelling between data and MC is seen most in 2015+2016.

There are some trends in the nominal efficiencies and the data/MC efficiency ratios with respect to other event variables. At high η there is a drop in efficiency, with similar behaviour seen

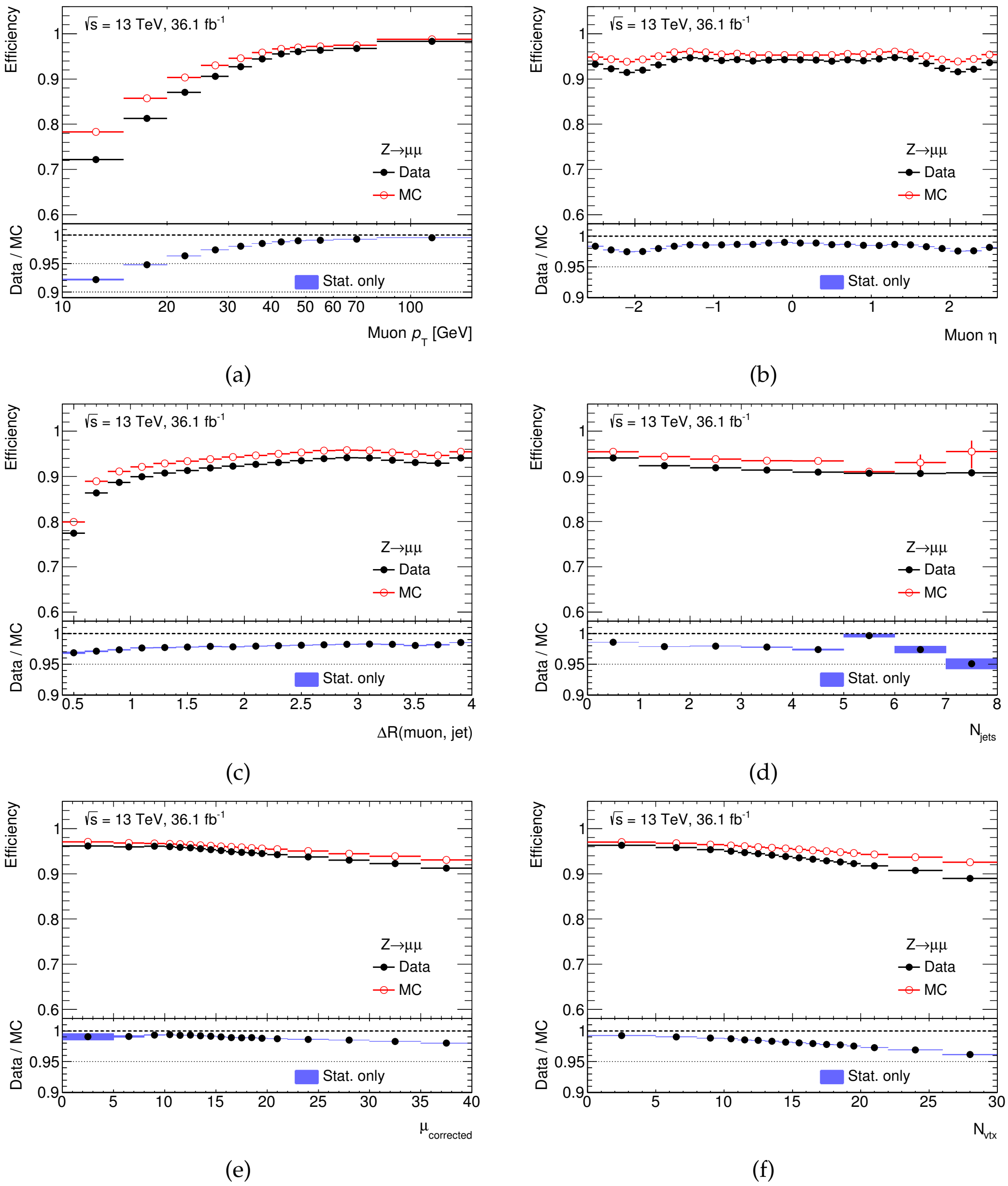
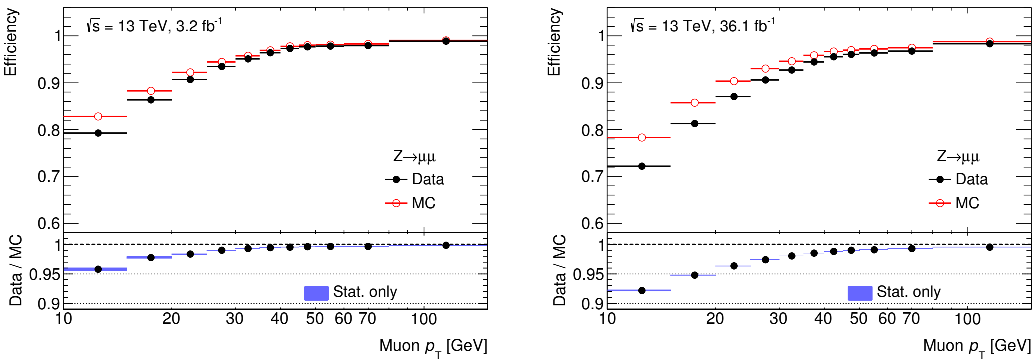


Figure 7.34: Efficiency and data/MC efficiency scale factors of the non-prompt BDT working point in $Z \to \mu\mu$ data (black circles) and simulation (red open circles) events as a function of (a) probe p_T, (b) probe η, (c) ΔR(probe, cal. jet), (d) N_{jet}, (e) pileup μ and (f) N_{vtx}. The uncertainties on the scale factors are statistical only.

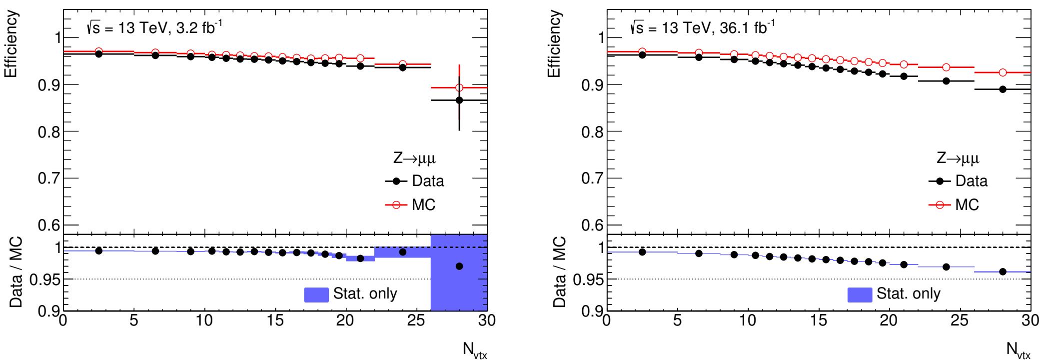
in data/MC. However, there is a correlation with the scale factor discrepancy at low p_T to account for. One also notices a drop in efficiency for low ΔR(muon, jet) and high N_{jets}. When a muon is near to a jet, the activity of the jet is likely to disrupt the chance of the muon passing the working point, since that muon will no doubt appear "non-prompt" due to the extra track activity. At large N_{jet} a similar effect is seen; the efficiency drops when there is more activity in the event compared to no jets. The data/MC efficiency is relatively flat for both variables.



(a) (b)

Figure 7.35: Efficiency and data/MC efficiency scale factors of the non-prompt BDT muon working point in $Z \to \mu\mu$ data (black circles) and simulation (red open circles) events as a function of probe p_T for (a) 2015 only and (b) 2015+2016 data.

Efficiency and data/MC efficiency scale factors of the non-prompt BDT working point in $Z \to \mu\mu$ data (black circles) and simulation (red open circles) events as a function of (a) probe p_T, (b) probe η, (c) ΔR(probe, cal. jet), (d) N_{jet}, (e) pileup μ and (f) N_{vtx} . The uncertainties on the scale factors are statistical only



(a) (b)

Figure 7.36: Efficiency and data/MC efficiency scale factors of the non-prompt BDT muon working point in $Z \to \mu\mu$ data (black circles) and simulation (red open circles) events as a function of the number of reconstructed vertices in the event for (a) 2015 only and (b) 2015+2016 data.

7.5.2.3 Scale factor systematic uncertainties

The efficiency scale factors are generally supplied as a two dimensional correction in (η, p_T), that take into account any differences in the p_T and η data/MC efficiency. Other muon isolation scale factors are delivered as a one dimensional correction in p_T, but a dependence on η is observed for the non-prompt BDT working point. Similar η dependence is also seen for the isolation scale factors for the tighest muon isolation working point: `isoFixedCutTightTrackOnly` [106]. For $p_T > 80$ GeV the more statistically accurate one dimensional scale factor in p_T is used for all bins in η due to the disappearance of the η dependence on the scale factor at such momentum values. The nominal scale factor as a function of p_T and η is shown in Figure 7.37.

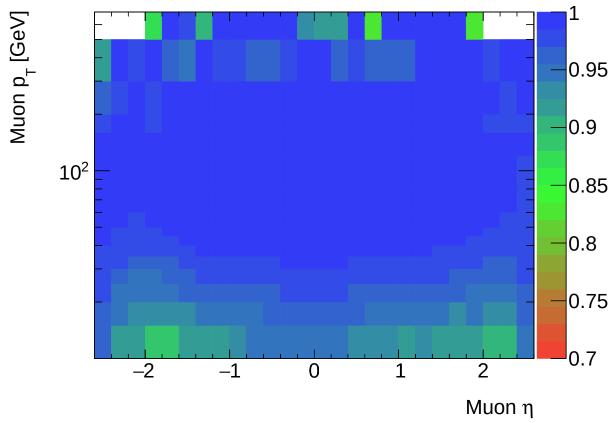


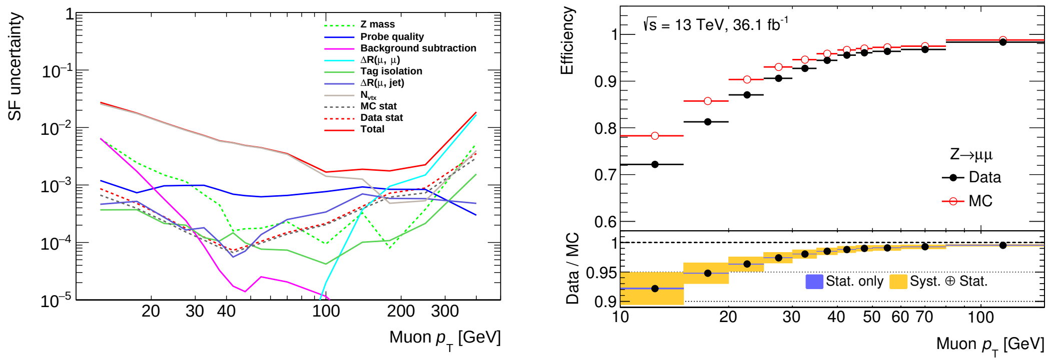
Figure 7.37: The two dimensional isolation scale factors of the non-prompt BDT working point as a function of muon η and p_T.

The data/MC efficiency dependence on other variables in Figure 7.34 is minimal, with the exception of N_{vtx} as discussed above. An extra systematic uncertainty is applied to cover this discrepancy, by taking the largest variation from nominal for scale factors with $N_{\text{vtx}} < 15$ or $N_{\text{vtx}} \geq 15$. The value 15 is chosen as it is approximately the median value in 2015+2016 data. The other systematic variations applied are shown in Table 7.11.

The deduced systematic uncertainties and their application to the data/MC efficiencies as a function of p_T are shown in Figure 7.38. The systematic attributed to N_{vtx} is the dominant uncertainty for muons with $p_T < 100$ GeV. This systematic uncertainty is conservative to cover the full discrepancy but, at a maximum level of approximately 3%, is negligible for many search analyses in which the non-prompt BDT can be used. The overall scale factor, integrated over p_T, with systematic uncertainty for prompt muons is approximately 0.98 ± 0.01.

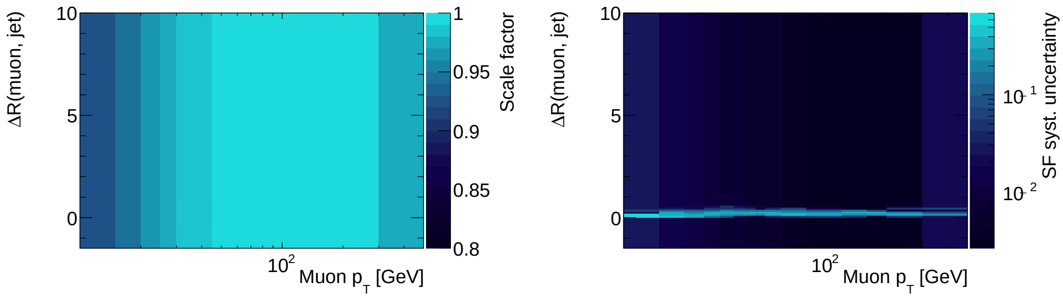
In addition to the systematic uncertainties in p_T, an extra systematic uncertainty is applied to muons within $\Delta R < 0.6$ of a calorimeter jet. This is due to a large mismodelling of the efficiency in data and MC due to the presence of these jets. This mismodelling was first seen for standard muon isolation scale factors and was applied for muons within $\Delta R < 0.4$ of a calorimeter jet [106]. The ΔR cut was extended for the non-prompt BDT to ensure full coverage of the mismodelling.

The systematic uncertainty is again binned in p_T in the same way as the nominal systematics and is simply taken as the difference in efficiency between data and MC, no matter how large that difference is. If the difference between data and MC is less than that of the nominal systematic uncertainties, then the nominal uncertainty is retained. The scale factor and associated systematic uncertainties are shown in bins of $\Delta R(\text{muon, cal. jet})$ versus p_T in Figure 7.39. Only the systematic uncertainty on the scale factor is dependent on $\Delta R(\text{muon, cal. jet})$, with the actual scale factor correction being independent.



(a) (b)

Figure 7.38: The values of the (a) systematic uncertainties broken down into each component as a function of muon p_T and (b) the efficiency comparison of data (full black circles) and MC (red open circles) in $Z \to \mu\mu$ events, where the value of the systematic uncertainties are applied to the data/MC scale factors in p_T.



(a) (b)

Figure 7.39: The (a) scale factor and (b) scale factor systematic uncertainty as a function of muon p_T and ΔR(muon, cal. jet).

7.5.2.4 Scale factor validation

Internal validation of the scale factors are performed in the same $Z \to \mu\mu$ events. Figures 7.40, 7.41 and 7.42 show this validation for η, ΔR(muon, cal. jet) and N_{vtx} respectively.

Almost perfect agreement of the scale factor corrected MC efficiencies to data are expected in p_T and η, due to the correction being applied in p_T and η. Any small differences are attributed to differences in binning between the scale factors and the plots. The scale factor corrected MC agrees well with data, with the extra N_{vtx} based systematic covering the discrepancies seen previously. The extra systematic uncertainty applied for muons with $\Delta R(\text{muon, cal. jet}) < 0.6$ can also be seen to cover such discrepancies.

7.5.3 Trigger scale factors

In general, the trigger scale factors for muons do not depend on the muon isolation selection. The difference in measured scale factor from nominal whilst applying different isolation selections to the probe muon is used as one of the systematic variations. This section studies whether the non-prompt BDT selection biases the trigger scale factors, by applying an extra "isolation" systematic variation to the probe muons.

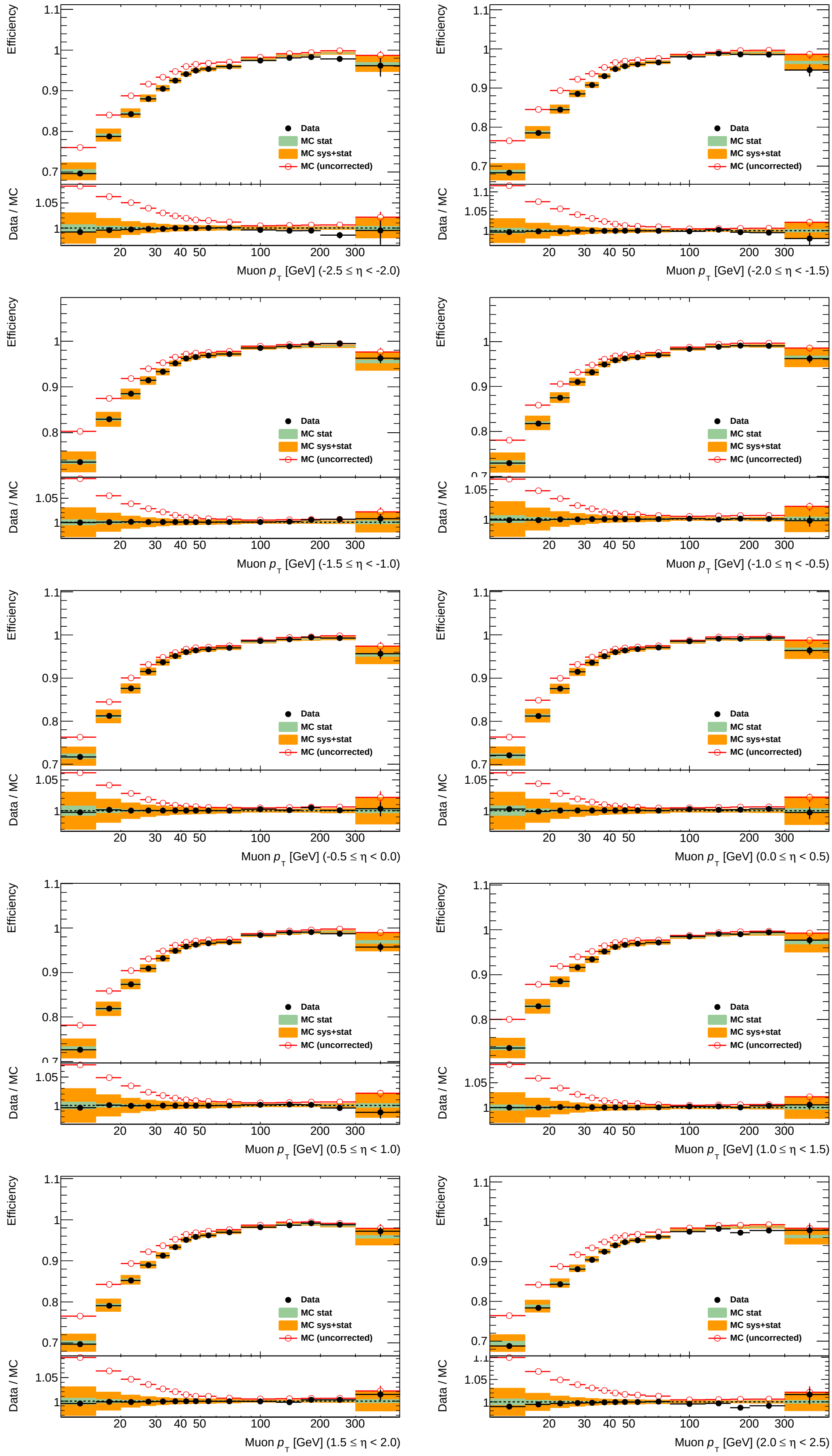


Figure 7.40: The comparison of scale factor corrected MC efficiencies, with statistical and systematic uncertainty applied, to the data efficiencies (black circles) as a function of muon p_T for a number of muon η selections. The uncorrected MC efficiencies (in red open circles) are also shown for reference.

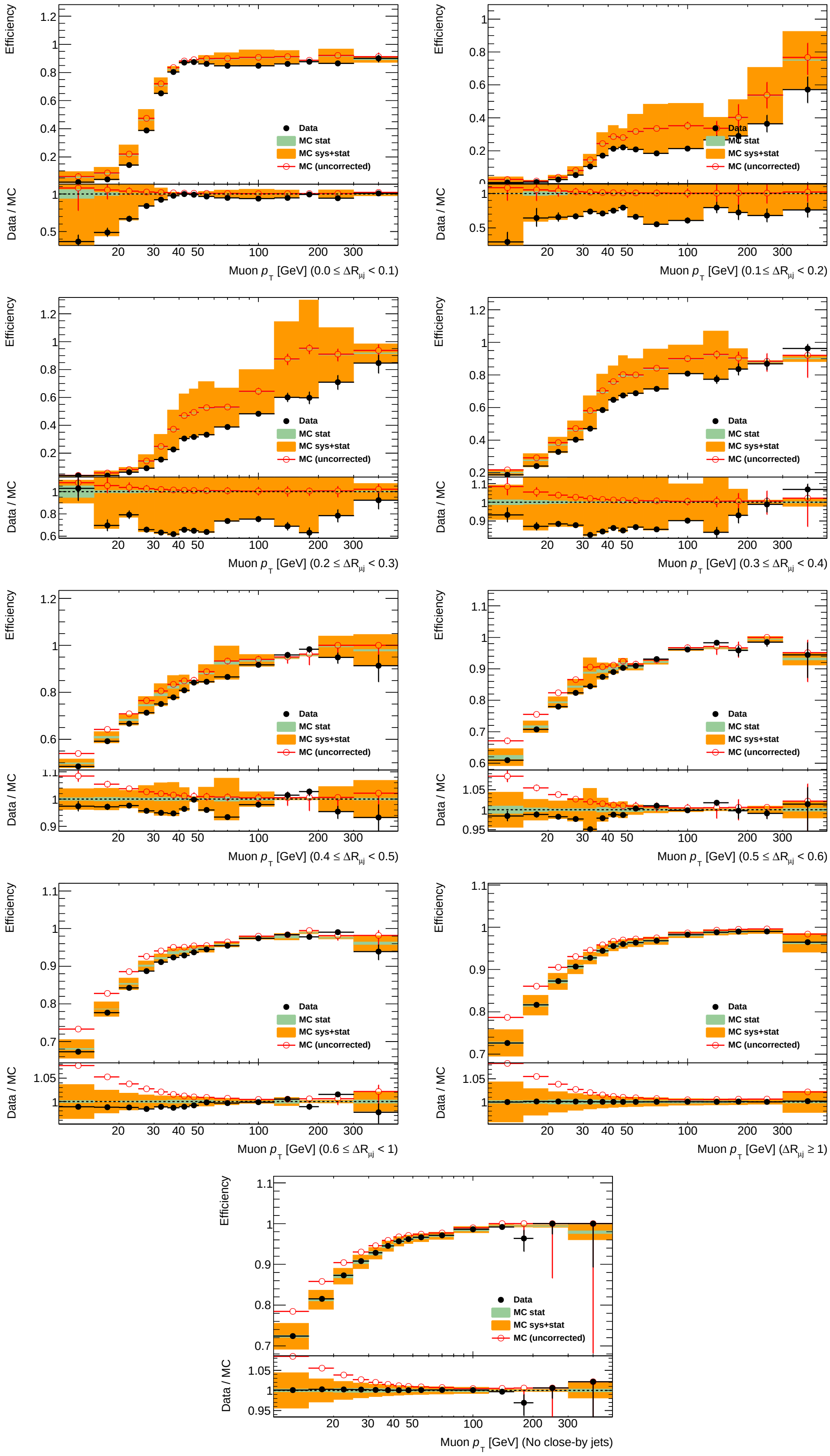


Figure 7.41: The comparison of scale factor corrected MC efficiencies, with statistical and systematic uncertainty applied, to the data efficiencies (black circles) as a function of muon p_T for a number of ΔR(muon, cal. jet) selections. The uncorrected MC efficiencies (in red open circles) are also shown for reference. The inflated systematic uncertainties for ΔR(muon, cal. jet) < 0.6 are clearly seen.

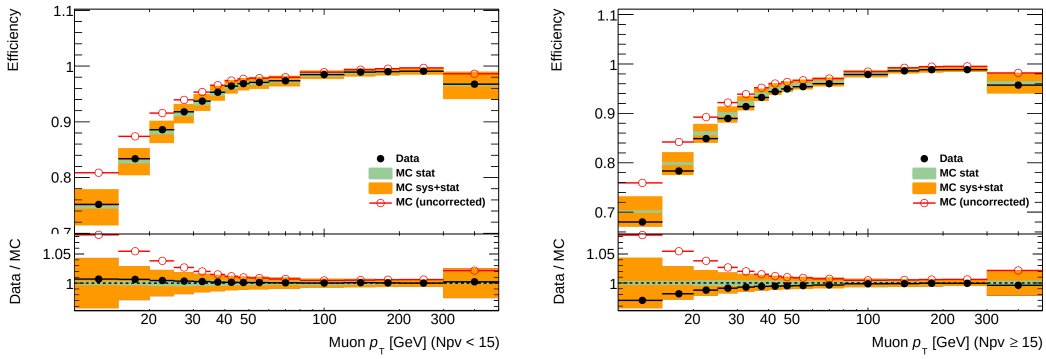


Figure 7.42: The comparison of scale factor corrected MC efficiencies, with statistical and systematic uncertainty applied, to the data efficiencies (black circles) as a function of muon p_T for (a) $N_{\text{vtx}} < 15$ and (b) $N_{\text{vtx}} > 15$. The uncorrected MC efficiencies (in red open circles) are also shown for reference.

7.5.3.1 Muon triggers

The single muon trigger used to analyse 2016 data is the logical OR of `HLT_mu26_ivarmedium` with `HLT_mu50`. In 2015 the trigger p_T threshold was lower than 2016, with the single muon trigger being the OR of `HLT_mu20_iloose_L1MU15` again with `HLT_mu50`. The scale factors for these triggers are shown below for `Loose` quality muons, along with the difference from nominal when applying the non-prompt BDT working point to the probe muon. The muon legs of the nominal dilepton triggers are also investigated for the electron-muon `HLT_e17_lhloose_nod0_mu14` trigger and the dimuon `HLT_mu22_mu8noL1` trigger in 2016, and `HLT_e17_lhloose_mu14` and `HLT_mu18_mu8noL1` in 2015. In the case of the electron-muon trigger for both 2015 and 2016, this corresponds to the `mu14` leg.

7.5.3.2 Systematic uncertainties

A flat systematic uncertainty is applied to the muon trigger scale factors, which are applied as a function of (η, ϕ) to account for differences in efficiency between data and simulation in spatial sections of the detector during different run periods. The nominal selections used to determine the scale factors and the different systematic variations compared to nominal are shown in Table 7.12 below.

The tag muon is matched to the `HLT_mu20_iloose_L1MU15` and `HLT_mu26_ivarmedium` for 2015 and 2016 respectively. The trigger p_T threshold for the logical OR of the single muon triggers is also taken as $1.05 \times 20(26)$ GeV for 2015 (2016).

The systematic variations are those used for all muon trigger scale factors. An extra systematic variation of applying the non-prompt BDT working point to the probe is applied, similarly to the isolation systematics.

7.5.3.3 Scale factor results

Tables 7.13 and 7.14 show the nominal scale factor value for `Loose` muons and the associated systematic uncertainties for the single muon triggers in 2015 and 2016 respectively. These triggers are shown for each entire year of data and split up by barrel and endcap. The scale factors calculated per data period within the year are also investigated and found to have similar results to the inclusive results. In both years, and for both barrel ($|\eta| < 1.0$) and endcap

Selection	Nominal	Variation(s)
Tag trigger	Match lowest unprescaled single muon trigger	-
Tag p_T	$p_T \geq 1.05 \times p_T^{\text{trigger}}$	-
$m(\ell\ell)$	$\lvert m(\ell\ell) - m(Z)\rvert < 10$ GeV	$\lvert m(\ell\ell) - m(Z)\rvert < 15$ GeV
Probe isolation	-	`isoGradient`
Probe p_T	$p_T > 10$ GeV	$p_T < 40$ GeV $p_T > 40$ GeV
N_{vtx}	-	$N_{\text{vtx}} \leq 11$ $N_{\text{vtx}} > 11$
Probe charge, Q	-	$Q = 1$ $Q = -1$
$\Delta\phi(\ell\ell)$	-	$\lvert\Delta\phi(\ell\ell)\rvert < \frac{\pi}{2}$
Probe IP	$\lvert d_0/\sigma_{d_0}\rvert < 3$ $\lvert z_0 \sin\theta\rvert < 2$ mm	No IP
Probe quality	`Loose`	-

Table 7.12: Definition of the event and lepton selections used to isolate $Z \to \mu\mu$ decays for calculating trigger scale factors. The systematic variations on these selections are also shown. `isoGradient` refers to an isolation working point designed to produce a gradient in the isolation efficiency from low p_T to high p_T.

($1.0 < |\eta| < 2.7$), the trigger scale factors for probe muons selected with the tight non-prompt BDT working point compared to nominal (no isolation) is within the total systematic uncertainty of the trigger scale factor itself. Slightly larger differences from nominal are seen in the barrel with respect to the endcap. The total trigger scale factor uncertainty without the non-prompt BDT systematic variation is quoted, since this is the uncertainty that is applied to muons in Chapter 8.

For completeness, Tables 7.15 and 7.16 show the results for the dilepton triggers in 2015 and 2016 respectively. The same statement from the single muon triggers applies.

Loose `HLT_mu20_iloose_L1MU15_OR_HLT_mu50` barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.8611 ± 0.110	0.000
N_{vtx} down	0.8609 ± 0.133	-0.022
N_{vtx} up	0.8595 ± 0.233	-0.177
$\Delta\phi(\ell\ell)$	0.8602 ± 0.134	-0.105
μ^+	0.8606 ± 0.155	-0.052
μ^-	0.8615 ± 0.155	0.052
p_T down	0.8622 ± 0.158	0.132
p_T up	0.8600 ± 0.152	-0.121
$m(\ell\ell)$	0.8605 ± 0.107	-0.068
no IP	0.8601 ± 0.109	-0.112
`isoGradient`	0.8627 ± 0.113	0.185
non-prompt BDT	0.8631 ± 0.111	0.241
Total		0.436
Total w/o non-prompt		0.363

Loose `HLT_mu20_iloose_L1MU15_OR_HLT_mu50` endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9760 ± 0.074	0.000
N_{vtx} down	0.9756 ± 0.091	-0.034
N_{vtx} up	0.9775 ± 0.157	0.156
$\Delta\phi(\ell\ell)$	0.9755 ± 0.091	-0.051
μ^+	0.9770 ± 0.105	0.100
μ^-	0.9750 ± 0.105	-0.099
p_T down	0.9774 ± 0.120	0.142
p_T up	0.9747 ± 0.094	-0.131
$m(\ell\ell)$	0.9756 ± 0.073	-0.038
no IP	0.9754 ± 0.074	-0.057
`isoGradient`	0.9766 ± 0.075	0.068
non-prompt BDT	0.9766 ± 0.074	0.062
Total		0.314
Total w/o non-prompt		0.307

Table 7.13: Scale factors and systematic uncertainties for the logical OR of `HLT_mu20_iloose_L1MU15` and `HLT_mu50` for 2015 data for the barrel region (left) and endcap region (right). The inclusive data/MC efficiency with statistical uncertainty and the percentage systematic uncertainty is given for each region. The total systematic uncertainties on the data/MC scale factors with and without applying the non-prompt BDT systematic variation are also shown.

Loose HLT_mu26_ivarmedium_OR_HLT_mu50 barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.8958 ± 0.043	0.000
N_{vtx} down	0.9010 ± 0.095	0.583
N_{vtx} up	0.8936 ± 0.050	-0.239
$\Delta\phi(\ell\ell)$	0.8947 ± 0.053	-0.118
μ^+	0.8965 ± 0.060	0.078
μ^-	0.8951 ± 0.061	-0.078
p_T down	0.8917 ± 0.066	-0.453
p_T up	0.8990 ± 0.056	0.359
$m(\ell\ell)$	0.8954 ± 0.042	-0.037
no IP	0.8933 ± 0.043	-0.272
isoGradient	0.8978 ± 0.043	0.230
non-prompt BDT	0.8999 ± 0.043	0.461
Total		1.048
Total w/o non-prompt		0.941

Loose HLT_mu26_ivarmedium_OR_HLT_mu50 endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9824 ± 0.029	0.000
N_{vtx} down	0.9877 ± 0.064	0.540
N_{vtx} up	0.9805 ± 0.034	-0.193
$\Delta\phi(\ell\ell)$	0.9810 ± 0.036	-0.146
μ^+	0.9831 ± 0.041	0.071
μ^-	0.9817 ± 0.041	-0.071
p_T down	0.9768 ± 0.052	-0.573
p_T up	0.9850 ± 0.035	0.265
$m(\ell\ell)$	0.9821 ± 0.029	-0.030
no IP	0.9820 ± 0.029	-0.042
isoGradient	0.9850 ± 0.029	0.265
non-prompt BDT	0.9851 ± 0.029	0.270
Total		0.951
Total w/o non-prompt		0.912

Table 7.14: Scale factors and systematic uncertainties for the logical OR of `HLT_mu26_ivarmedium` and `HLT_mu50` for 2016 data for the barrel region (left) and endcap region (right). The inclusive data/MC efficiency with statistical uncertainty and the percentage systematic uncertainty is given for each region. The total systematic uncertainties on the data/MC scale factors with and without applying the non-prompt BDT systematic variation are also shown.

Loose HLT_mu14 barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9111 ± 0.098	0.000
N_{vtx} down	0.9110 ± 0.119	-0.010
N_{vtx} up	0.9100 ± 0.208	-0.118
$\Delta\phi(\ell\ell)$	0.9105 ± 0.119	-0.073
μ^+	0.9096 ± 0.138	-0.168
μ^-	0.9127 ± 0.139	0.170
p_T down	0.9124 ± 0.138	0.139
p_T up	0.9098 ± 0.139	-0.145
$m(\ell\ell)$	0.9108 ± 0.096	-0.035
no IP	0.9109 ± 0.098	-0.027
isoGradient	0.9117 ± 0.101	0.062
non-prompt BDT	0.9123 ± 0.100	0.128
Total		0.373
Total w/o non-prompt		0.350

Loose HLT_mu14 endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9808 ± 0.064	0.000
N_{vtx} down	0.9807 ± 0.078	-0.010
N_{vtx} up	0.9818 ± 0.136	0.098
$\Delta\phi(\ell\ell)$	0.9806 ± 0.079	-0.021
μ^+	0.9819 ± 0.091	0.114
μ^-	0.9797 ± 0.091	-0.114
p_T down	0.9824 ± 0.101	0.164
p_T up	0.9794 ± 0.083	-0.140
$m(\ell\ell)$	0.9808 ± 0.063	-0.002
no IP	0.9808 ± 0.064	-0.003
isoGradient	0.9808 ± 0.066	0.003
non-prompt BDT	0.9808 ± 0.065	0.004
Total		0.288
Total w/o non-prompt		0.288

Loose HLT_mu18_mu8noL1 barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9800 ± 0.047	0.000
N_{vtx} down	0.9797 ± 0.057	-0.031
N_{vtx} up	0.9800 ± 0.101	0.006
$\Delta\phi(\ell\ell)$	0.9801 ± 0.057	0.010
μ^+	0.9804 ± 0.067	0.044
μ^-	0.9795 ± 0.067	-0.044
p_T down	0.9772 ± 0.068	-0.288
p_T up	0.9828 ± 0.065	0.288
$m(\ell\ell)$	0.9797 ± 0.046	-0.025
no IP	0.9798 ± 0.047	-0.020
isoGradient	0.9803 ± 0.049	0.036
non-prompt BDT	0.9810 ± 0.047	0.103
Total		0.428
Total w/o non-prompt		0.416

Loose HLT_mu18_mu8noL1 endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9706 ± 0.025	0.000
N_{vtx} down	0.9708 ± 0.030	0.021
N_{vtx} up	0.9700 ± 0.054	-0.057
$\Delta\phi(\ell\ell)$	0.9704 ± 0.031	-0.018
μ^+	0.9713 ± 0.035	0.076
μ^-	0.9699 ± 0.036	-0.075
p_T down	0.9735 ± 0.039	0.296
p_T up	0.9686 ± 0.033	-0.208
$m(\ell\ell)$	0.9704 ± 0.025	-0.021
no IP	0.9706 ± 0.025	0.002
isoGradient	0.9706 ± 0.026	0.004
non-prompt BDT	0.9706 ± 0.026	0.004
Total		0.383
Total w/o non-prompt		0.383

Table 7.15: Scale factors and systematic uncertainties for the `mu14 leg` of the `HLT_e17_lhloose_mu14` trigger (top) and for `HLT_mu18_mu8noL1` (bottom) for full 2015 data for the barrel region (left) and endcap region (right). The inclusive data/MC efficiency with statistical uncertainty and the percentage systematic uncertainty is given for each region. The total systematic uncertainties on the data/MC scale factors with and without applying the non-prompt BDT systematic variation are also shown.

Loose HLT_mu14 barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9339 ± 0.036	0.000
N_{vtx} down	0.9379 ± 0.079	0.431
N_{vtx} up	0.9324 ± 0.041	-0.151
$\Delta\phi(\ell\ell)$	0.9335 ± 0.043	-0.034
μ^+	0.9338 ± 0.050	-0.003
μ^-	0.9339 ± 0.051	0.003
p_T down	0.9334 ± 0.050	-0.053
p_T up	0.9344 ± 0.051	0.054
$m(\ell\ell)$	0.9336 ± 0.035	-0.031
no IP	0.9315 ± 0.035	-0.255
isoGradient	0.9334 ± 0.037	-0.052
non-prompt BDT	0.9359 ± 0.036	0.220
Total		0.577
Total w/o non-prompt		0.533

Loose HLT_mu14_RM endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9848 ± 0.023	0.000
N_{vtx} down	0.9873 ± 0.052	0.249
N_{vtx} up	0.9841 ± 0.027	-0.079
$\Delta\phi(\ell\ell)$	0.9844 ± 0.029	-0.042
μ^+	0.9852 ± 0.033	0.040
μ^-	0.9845 ± 0.033	-0.040
p_T down	0.9846 ± 0.037	-0.027
p_T up	0.9849 ± 0.030	0.004
$m(\ell\ell)$	0.9848 ± 0.023	-0.002
no IP	0.9847 ± 0.023	-0.016
isoGradient	0.9849 ± 0.024	0.003
non-prompt BDT	0.9850 ± 0.024	0.013
Total		0.272
Total w/o non-prompt		0.272

Loose HLT_mu22_mu8noL1 barrel		
barrel	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9906 ± 0.017	0.000
N_{vtx} down	0.9912 ± 0.038	0.063
N_{vtx} up	0.9904 ± 0.020	-0.017
$\Delta\phi(\ell\ell)$	0.9904 ± 0.021	-0.014
μ^+	0.9905 ± 0.024	-0.003
μ^-	0.9906 ± 0.024	0.003
p_T down	0.9896 ± 0.025	-0.102
p_T up	0.9915 ± 0.023	0.095
$m(\ell\ell)$	0.9903 ± 0.017	-0.025
no IP	0.9880 ± 0.017	-0.258
isoGradient	0.9906 ± 0.018	-0.001
non-prompt BDT	0.9928 ± 0.017	0.221
Total		0.374
Total w/o non-prompt		0.302

Loose HLT_mu22_mu8noL1 endcap		
endcap	data/MC ± stat. error	(syst-nom.)/nom.
nominal	0.9945 ± 0.007	0.000
N_{vtx} down	0.9952 ± 0.016	0.068
N_{vtx} up	0.9943 ± 0.008	-0.022
$\Delta\phi(\ell\ell)$	0.9945 ± 0.009	-0.002
μ^+	0.9946 ± 0.010	0.006
μ^-	0.9945 ± 0.010	-0.006
p_T down	0.9945 ± 0.012	-0.000
p_T up	0.9945 ± 0.009	-0.002
$m(\ell\ell)$	0.9945 ± 0.007	-0.004
no IP	0.9945 ± 0.007	-0.002
isoGradient	0.9946 ± 0.007	0.004
non-prompt BDT	0.9946 ± 0.007	0.007
Total		0.072
Total w/o non-prompt		0.072

Table 7.16: Scale factors and systematic uncertainties for the mu14 leg of the HLT_e17_lhloose_nod0_mu14 trigger (top) and for HLT_mu22_mu8noL1 (bottom) for full 2016 data for the barrel region (left) and endcap region (right). The inclusive data/MC efficiency with statistical uncertainty and the percentage systematic uncertainty is given for each region. The total systematic uncertainties on the data/MC scale factors with and without applying the non-prompt BDT systematic variation are also shown.

Chapter 8
The search for $t\bar{t}H$ in multilepton final states

The measurement of the associated production of the Higgs boson and a top quark pair ($t\bar{t}H$) is an important test of the SM, being one of two predicted Higgs production modes that can probe the top Yukawa coupling at tree level; the other being the tH mechanism. This chapter focuses on the search for this production mode in Higgs decays into pairs of W bosons, Z bosons or τ leptons, more commonly referred to as multilepton final states, with 36.1 fb^{-1} of data collected with the ATLAS detector during 2015 and 2016.

Higgs boson phenomenology is discussed in Section 8.1, focusing on production mechanisms, decays and the frameworks used to determine Higgs boson properties experimentally. The statistical model used to measure such properties of the Higgs is shown in Section 8.2. The main backgrounds to the multilepton decays of $t\bar{t}H$ and how these decays are measured is shown in Section 8.3. The preliminary multilepton analysis utilising early Run-2 data is briefly discussed in Section 8.4, with the problems and inadequacies highlighted for context for the updated measurement of this production mode in full 2015 and 2016 data, which is discussed in full in Section 8.5.

8.1 Higgs boson phenomenology

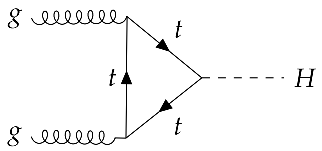
The Higgs boson can be produced in a number of different mechanisms and can decay into a large number of particle types. Such production mechanisms (Section 8.1.1) and decays (Section 8.1.2) are discussed below, with specific emphasis on top quark associated Higgs production with the Higgs decaying to pairs of W bosons, Z bosons or τ leptons.

Measurements of such events require frameworks and assumptions within which to work. The κ coupling modifier and signal strength, μ, are designed to measure the Higgs boson properties with respect to the SM values. These are discussed in Sections 8.1.3 and 8.1.4.

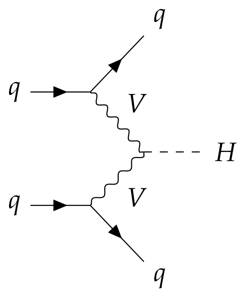
8.1.1 Higgs boson production

There are four main Higgs boson production mechanisms at the LHC. In order of decreasing total cross section and with leading order Feynman diagrams:

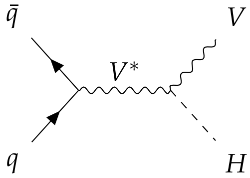
1. Gluon-gluon fusion, $gg \to H$



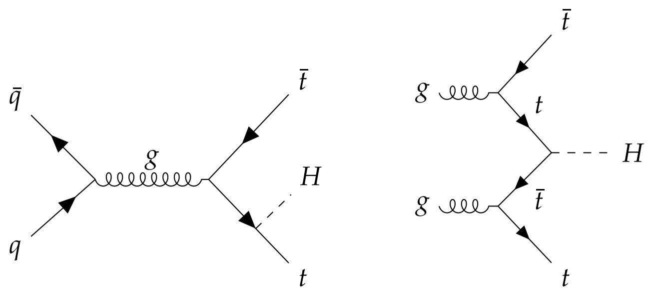
2. Vector boson fusion, $qq \to qq + VV \to qq + H$



3. Vector boson associated production (or Higgstrahlung), $q\bar{q} \to V + H$



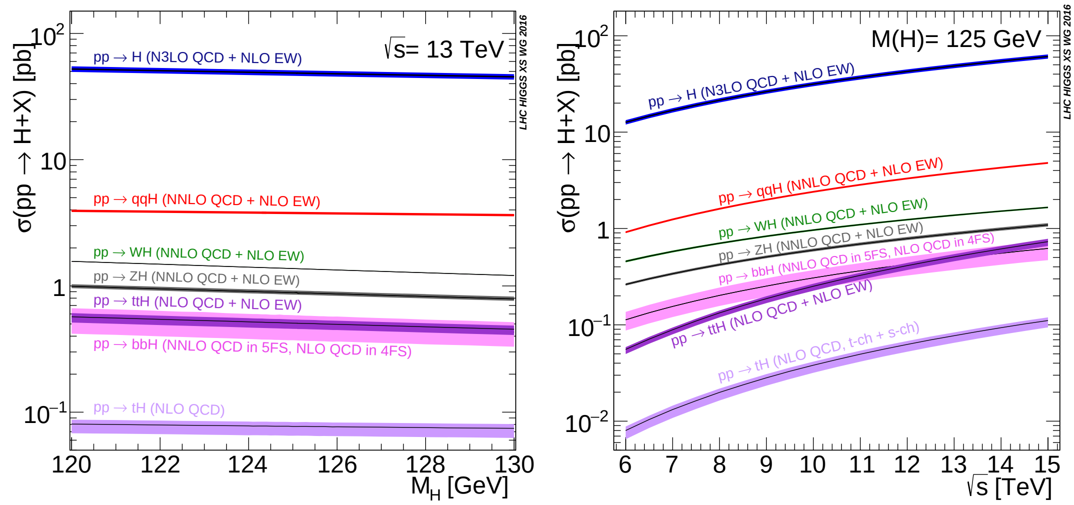
4. Top quark associated production, $q\bar{q}/gg \to t\bar{t} + H$



It should be noted that a top quark loop is shown in the production of a Higgs boson via gluon fusion. Other particles that couple via the strong force, such as a b quark, contribute to this loop but are subdominant. This top quark loop production of the Higgs results in an indirect method of measuring the top quark Yukawa coupling; indirect due to not being able to determine, experimentally, whether a top quark was present in the loop or not.

There are also two other subdominant production mechanisms: bottom quark associated production ($b\bar{b}H$) and single top associated production (tH). These mechanisms are both suppressed due to interference effects. tH provides another production measurement with which to measure the top Yukawa coupling at tree level, as with $t\bar{t}H$, but the order of magnitude decrease in production cross section relative to $t\bar{t}H$ makes this a very difficult measurement.

The calculated production cross sections as a function of Higgs mass, m_H, at $\sqrt{s} = 13$ TeV and as a function of $\sqrt{s}$ for $m_H = 125$ GeV are shown in Figure 8.1. It can be seen that the production cross section for gluon fusion is approximately two orders of magnitude larger than the cross section for $t\bar{t}H$ for a Higgs boson with $m_H = 125$ GeV. When comparing how the cross section changes as a function of $\sqrt{s}$, clearly the rate of $t\bar{t}H$ increases most rapidly due to the increase in the kinematic phase space for the production of such a heavy final state (approximately 470 GeV).



(a) (b)

Figure 8.1: Higgs boson production cross-sections as a function of (a) Higgs mass at $\sqrt{s} = 13$ TeV and (b) $\sqrt{s}$ for a Higgs with $m_H = 125$ GeV [107].

8.1.2 Higgs boson decays

The mass of the Higgs boson at $m_H = 125$ GeV results in a large number of possible decays to both fermions and bosons. The branching ratio, $\mathcal{B}$, or the fraction of decays to a specific final state, in the SM as a function of m_H at $\sqrt{s} = 13$ TeV is shown in Figure 8.2.

The dominant decay mode of the Higgs boson is $H \to b\bar{b}$ [1], occurring in 57.8% of total Higgs decays in the SM [3]. Just the flavour Evidence of $H \to bb$ has recently been presented in VH production [108], but observation of $H \to bb$ is yet to be obtained due to large QCD jet backgrounds. In top quark associated production, the nominal final state of $t\bar{t}H(H \to bb)$ is $bbbbWW$. This has a large, irreducible background of $ttbb$; a top quark pair produced in association with a bottom quark pair. This background results in a difficult measurement of $t\bar{t}H(H \to bb)$, despite the increased statistics with respect to other decays (or channels) [109].

The second most common decay mode is $t\bar{t}H(H \to WW^*)$ at 21.6% in the SM. The relatively large W mass, m_W, results in one of the W bosons being produced off-shell ($2m_W > m_H$). $H \to WW$ observations have been made by both ATLAS [110] and CMS [111] and are thus an exciting area in which to search for $t\bar{t}H$ production, in the intermediate statistical accuracy region. Leptonic decays of W bosons are used to search for this final state. Final states with $H \to \tau\tau$ ($\mathcal{B} = 6.3\%$) and $H \to ZZ$ ($\mathcal{B} = 2.6\%$) can also resemble $H \to WW$ decays, and are thus combined into one "multilepton" analysis as discussed above.

Both ATLAS and CMS actually observed the Higgs boson for the first time in the clean, resonant $H \to \gamma\gamma$ and $H \to ZZ(4\ell)$ decays [1, 2]. These decays may have small branching ratios (approximately $\mathcal{B} = 0.2\%$ and 0.01% respectively) but also have low backgrounds. With increasing statistics provided by the LHC year on year, these decay channels will become more and more dominant in future Higgs boson measurements.

[1] The highlighting of the anti-particle nature of one of the Higgs decays is suppressed herein.

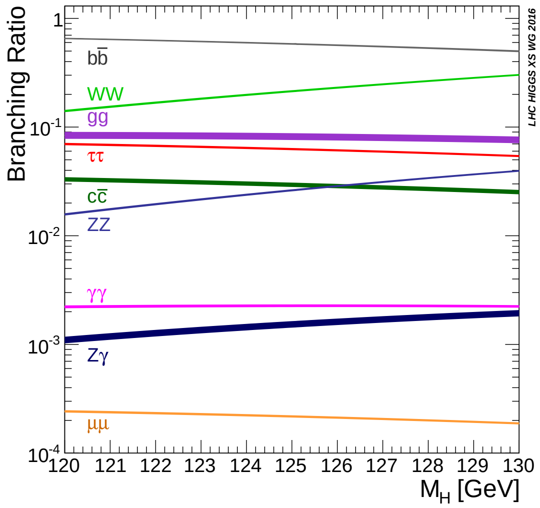


Figure 8.2: Higgs boson branching ratios as a function of m_H at $\sqrt{s} = 13$ TeV [107].

8.1.3 Higgs boson couplings

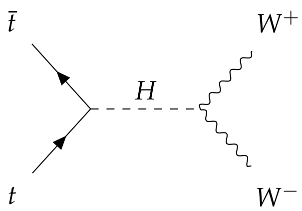
The κ-parameterisation provides a consistent framework in which the couplings of the Higgs boson can be measured. The framework makes the assumption that the measured Higgs decay particles originate from a single narrow resonance at 125 GeV, with negligible width. The Higgs boson width is calulated to be 4 MeV [3] in the SM which is approximately four orders of magnitude smaller than the value of m_H, satisying this requirement. In this case the narrow-width approximation can be used [112], resulting in

$$\sigma(ii \to H) \cdot \mathcal{B}(H \to ff) = \sigma(ii \to H) \cdot \frac{\Gamma_{ff}}{\Gamma_H}, \tag{8.1}$$

where $\sigma(ii \to H)$ is a Higgs production cross-section, $\mathcal{B}$ the branching ratio of a Higgs decay, Γ_{ff} the partial decay width of the Higgs decay and Γ_H the total Higgs decay width. Coupling scale factors κ_i can be introduced, defined so that σ and Γ scale with respect to the SM value by κ^2:

$$\frac{\sigma}{\sigma_{SM}} = \frac{\Gamma}{\Gamma_{SM}} = \kappa^2. \tag{8.2}$$

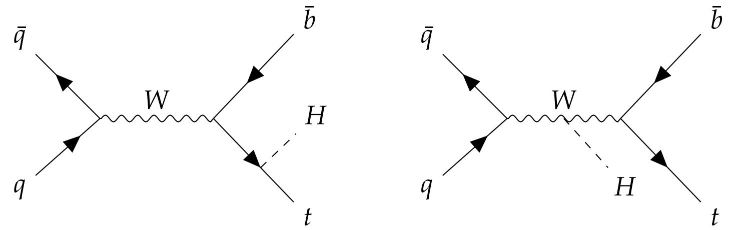
Taking the Higgs boson being produced from top quarks and decaying to W bosons as an example:



$$\begin{aligned}\sigma(tt \to H) \cdot \mathcal{B}(H \to WW) &= \kappa_t^2 \sigma_{SM}(tt \to H) \cdot \frac{\kappa_W^2 \Gamma_{WW}^{SM}}{\kappa_H^2 \Gamma_H^{SM}} \\ &= \sigma_{SM}(tt \to H) \cdot \mathcal{B}_{SM}(H \to WW) \cdot \frac{\kappa_t^2 \kappa_W^2}{\kappa_H^2}\end{aligned} \tag{8.3}$$

Here the κ_t^2 represents the modification of the Higgs-top coupling compared to the SM, κ_W^2 the modification of the Higgs-W coupling and κ_H^2 the modification to the total Higgs decay width. Note that the sign of κ is not a priori accessible due to the square nature of the factor.

The relative sign of κ can be inferred from interference effects between diagrams [113]. For example, the tH production cross section at the LHC is small; approximately an order of magnitude smaller than $t\bar{t}H$ inclusively. This is due to destructive interference between tH diagrams involving the Higgs boson being radiated by either the W boson or the top quark:



Measuring the tH process therefore results in sensitivity to the relative sign of κ_t and κ_W.

This framework is used to test whether the Higgs boson couplings comply with the predictions of the SM. If one plotted the Yukawa coupling, y, with respect to lepton mass, m_ℓ, one would expect to observe a linear behaviour for different lepton masses. If one plotted the gauge boson coupling with respct to gauge boson mass, m_V, then one would expect to observe a quadratic behaviour. The "effective" couplings are shown as a function of particle mass in Figure 8.3 for all the measured decay channels of the Higgs boson. This "effective" coupling is determined by introducing fit parameters ϵ and M [114] to the standard lepton

$$y_f = \sqrt{2}\left(\frac{m_f}{M}\right)^{1+\epsilon} \rightarrow \sqrt{2}\frac{m_f}{v}, \tag{8.4}$$

and boson couplings

$$\lambda_V = 2\left(\frac{m_V^{2(1+\epsilon)}}{M^{1+2\epsilon}}\right) \rightarrow 2\frac{m_V^2}{v}, \tag{8.5}$$

that result in the SM values for $\epsilon \rightarrow 0$ and $M \rightarrow v$, the SM Higgs expectation value. This coupling is then scaled by the relevant κ factors. It appears that nature is in good agreement with the SM, even in this simplistic framework.

One might also notice that the statistical uncertainty on the value of κ_t from Figure 8.3 is small. This is due to the extraction of κ_t from the effective coupling κ_g and κ_γ. Both processes require a loop in the production or decay and are thus parameterised as a function of the particles in the loop that interact with the Higgs [113]. A top quark loop is dominant in both κ_g and κ_γ and result in the relatively precise measurement.

In these loops, it is not known for sure which particle interacted in the loop. BSM particles could potentially contribute which would influence the measurement of κ_t, as shown in Figure 8.4. This is one reason why the measurement of $t\bar{t}H$ (and tH) is so important; the processes access the top Yukawa coupling tree level and are thus immune to such particle loop ambiguities.

8.1.4 Signal strengths

The signal strength, μ, is defined as the measured rate of a process with respect to the value of the rate from the SM. This is discussed in the context of Higgs boson production and decays below.

The signal strength can be defined for both the production process, $ii \rightarrow H$ and the Higgs decay, $H \rightarrow ff$,

$$\mu_i = \frac{\sigma(ii \rightarrow H)}{\sigma_{SM}(ii \rightarrow H)}, \qquad \mu^f = \frac{\mathcal{B}(H \rightarrow ff)}{\mathcal{B}_{SM}(H \rightarrow ff)}. \tag{8.6}$$

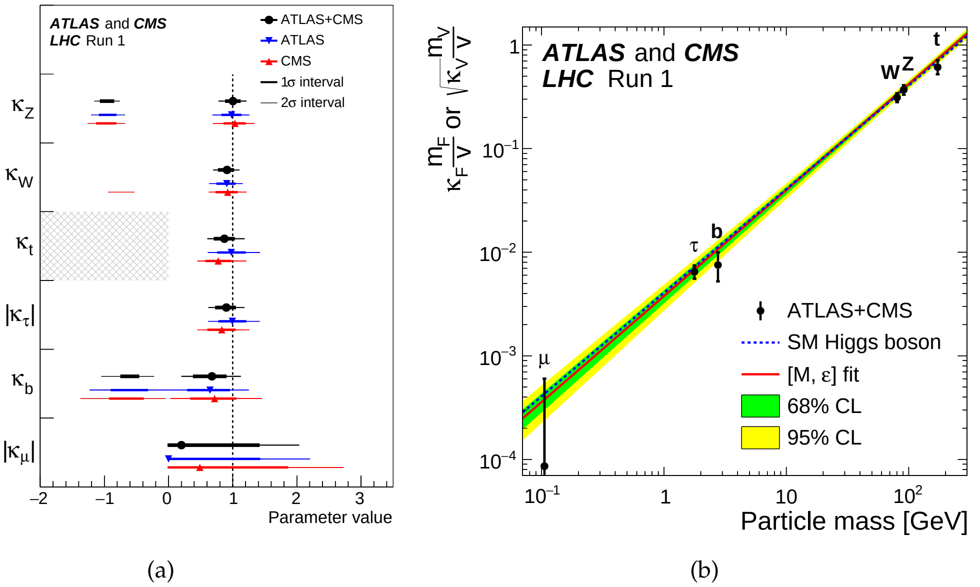


Figure 8.3: (a) The fitted κ values from the global Higgs boson coupling fit. (b) The "effective" couplings of bosons and fermions to the Higgs as measured in data. A best fit of the modified couplings, with parameters M and ϵ, are shown that result in the SM value for $M \to v$ and $\epsilon \to 0$ [113].

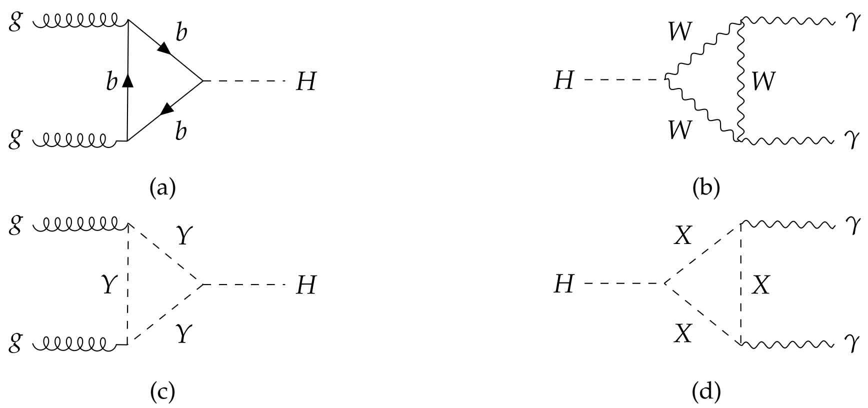
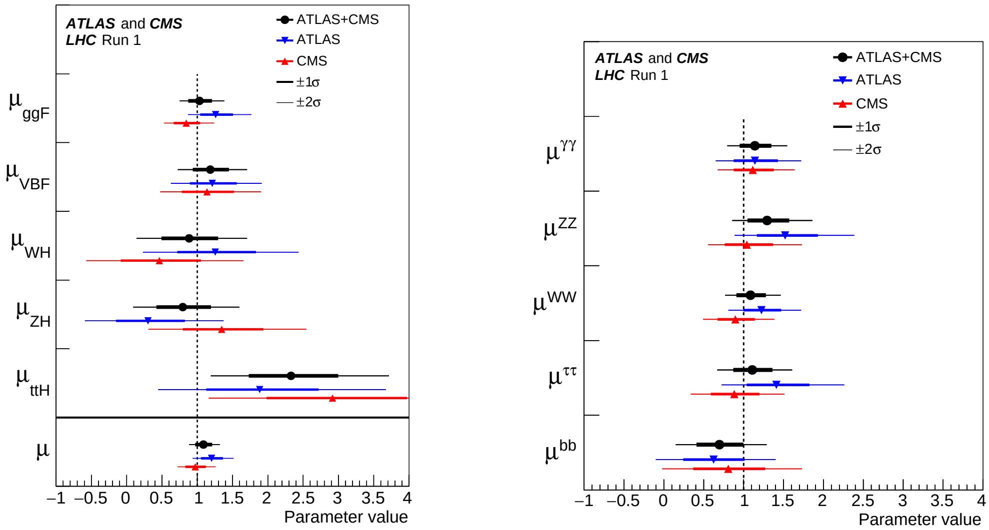


Figure 8.4: Example Feynman diagrams of decays of a Higgs to two photons and the gluon fusion production mechanism with (a) b quarks, (b) W bosons and exotica (c) Y, (d) X. The unknown exotic X is charged and couples via the weak interaction. The similar exotic Y couples via the strong interaction. These interactions would result in a modification of the Higgs branching ratios and production cross-sections compared to the SM, if present.

By definition, a measured $\mu = 1$ refers to exact agreement of data with the SM. Neither $\sigma(ii \to H)$ nor $\mathcal{B}(H \to ff)$ can be measured individually when measuring a full $ii \to H \to ff$ process. Instead the product of the production and decay fraction signal strengths is typically measured,

$$\mu_{i\to f} = \mu_i \mu_f. \tag{8.7}$$

The measured μ values for both production mechanisms and decay rates are shown in Figure 8.5 for combined ATLAS and CMS Run-1 data. For the production mechanism measurement, the value of μ for the decay rates are assumed to be the SM, $\mu_f = 1$, and vice-versa for the decay rates ($\mu_i = 1$). The agreement with the SM is excellent, albeit with some large uncertainties in places. More precise measurements in Run-2 will decrease these uncertainties and start to probe the Higgs boson properties more thoroughly.



(a) (b)

Figure 8.5: The values of μ measured for Higgs boson (a) production mechanisms and (b) decay rates from the global Higgs boson coupling fit using combined ATLAS and CMS Run-1 data [113].

8.2 Statistical model

A statistical model needs to be implemented in order to calculate the signal strength of top quark associated Higgs production in data. This process has not been observed before (statistically speaking) and therefore one must define hypotheses to test statistically.

The first hypothesis is the null hypothesis; the situation in which there is no top quark associated Higgs production. This corresponds to $\mu = 0$. The second is the signal-plus-background ($S + B$) hypothesis; the situation in which the SM correctly predicts the rate of $t\bar{t}H$. One aims to reject the null hypothesis. The validity of the null hypothesis can be determined with a test statistic; the test statistic used to measure $t\bar{t}H$ is discussed in Section 8.2.1 below.

Typically one measures a test statistic in data and a probability to observe this result within the null hypothesis can be calculated. This probability is called a p-value. The discovery significance of the result can then be calculated from the p-value. Typically the significance is stated in terms of the number of standard deviations of a Gaussian distribution, σ. In particle physics, discovery significances of 3σ correspond to *evidence* of a process and 5σ to *observation* of a given signal process.

The significance can be roughly estimated by $S/\sqrt{B}$, where S and B are the total number of signal and background events for a sufficiently large background with respect to signal. This is what is used in the optimisation of lepton selections in Section 8.5.2.

8.2.1 Profile likelihood model

The test statistic used to measure $t\bar{t}H$ is the profile log-likelihood ratio [115].

Typically one measures the signal strength in different regions of phase space, or bins. The binned likelihood, $\mathcal{L}(\mu, \boldsymbol{\theta})$, can be defined as the probability of observed data with respect to a hypothesis for a given signal strength and the set of nuisance parameters, $\boldsymbol{\theta}$ [2]. These parameters are used to parameterise the effect of each systematic uncertainty on the signal and background expectations in each region. The likelihood is constructed from a product of Poisson probability terms

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{(\mu S_i + B_i)^{N_i}}{N_i!} e^{-(\mu S_i + B_i)} \prod_{\theta_j}^{\boldsymbol{\theta}} f(\theta_j), \tag{8.8}$$

where n is the number of bins, N_i is the number of events in bin i and S_i (B_i) are the number of signal (background) events in that bin. S and B are functions of the set of nuisance parameters. The probability density function for each nuisance parameter is denoted by $f(\theta_j)$. These functions are penalty terms; they result in decreasing the likelihood when the individual nuisance parameter is shifted from nominal. Typically Gaussian functions with width of the size of the systematic uncertainty are used to model the nuisance parameters. Log-normal probability density functions are also used where the parameter must remain positive, such as uncertainties on normalisation terms.

One can determine the value of μ and set $\boldsymbol{\theta}$ that maximise the likelihood in the fit. These parameters are defined as $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. The test statistic, q_μ, is then devised from the log-likelihood ratio, Λ_μ,

$$q_\mu = -2\ln\Lambda_\mu(\mu) = -2\ln\frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{8.9}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ refers to the profiled set of nuisance parameters that maximise the likelihood for a certain value of μ. Therefore the test statistic under the null hypothesis can be determined,

$$q_0 = -2\ln\Lambda_\mu(0), \tag{8.10}$$

and the discovery significance simply as $\sqrt{q_0}$.

The test statistic is devised in such a way that the total uncertainty on μ, σ_μ, is calculated by measuring the variation in the q_μ profile as a function of μ when increased by one unit from the minimum value. The statistical uncertainty on the signal strength, σ_{stat}, is determined by fixing the nuisance parameters to their best fit values and measuring the variation in q_μ as a function of μ as above. The systematic uncertainty, σ_{syst}, is then calculated by subtracting the statistical uncertainty in quadrature with the total uncertainty:

$$\sigma_{\text{syst}} = \sqrt{\sigma_\mu^2 - \sigma_{\text{stat}}^2}. \tag{8.11}$$

[2] There are nominally 315 nuisance parameters in the $t\bar{t}H$ multilepton analysis (see Section 8.5.7).

8.3 Channels and backgrounds

A number of different channels can be defined to measure multilepton Higgs decays. Instead of measuring individual decays like $H \rightarrow WW$ or $H \rightarrow \tau\tau$, the number of light leptons and hadronic taus in the event are measured and it is those that define orthogonal channels. These channels then contain admixtures of each of the three multilepton Higgs decay modes. From hereafter light leptons may be referred to simply as "leptons" and τ leptons as hadronic taus.

Example Feynman diagrams of the multilepton $t\bar{t}H$ decays are shown in Figure 8.6. $H \rightarrow WW$ is the dominant decay mode of the multilepton decays in almost all channels, as one might expect from the large branching ratio. The nominal final state is $WWWWbb$. One can imagine selecting $t\bar{t}H$ by requiring there is a b-jet in the event and then by the leptonic decays of each W; from one to four light leptons.

Decays with only one lepton, with a large number of jets (due to the other hadronic W decays) will have large backgrounds from processes such as W+jets, semileptonic $t\bar{t}$ decays and QCD related events. Decays with two opposite-sign leptons resemble a dilepton $t\bar{t}$ decay with extra jets. The inclusive cross section for $t\bar{t}$ is roughly three orders of magnitude larger than $t\bar{t}H$ and thus such measurements of $t\bar{t}H$ would have extremely large backgrounds. However, two same-sign leptons (2ℓSS) are a good signature in which to isolate $t\bar{t}H(H \rightarrow WW)$ events. Events containing two same-sign leptons are rare in the SM. Other processes in the SM, complementary to $t\bar{t}H$, are diboson production (VV), top quark associated W boson production ($t\bar{t}W$) and top quark associated Z boson production ($t\bar{t}Z$). Other rarer processes in the SM also have this signature, and include WWW, ttt and $tttt$. Typical Feynman diagrams for $t\bar{t}W$, $t\bar{t}Z$ and VV are shown in Figures 8.7 and 8.8 below. Both $t\bar{t}Z$ and $t\bar{t}W$ nominally have b-quarks in the final state and thus have very similar topologies to $t\bar{t}H$. VV on the other hand does not have b-quarks in every final state and thus the requirement of a single b-jet in the event is a good way of rejecting diboson events. Diboson events can still have two same-sign leptons and a b-jet in an event when the two bosons are produced in association with a bb pair due to a gluon splitting from one of the incoming quarks.

One notices that all these processes contain prompt leptons, from either the W or Z boson in the specific event. A same-sign lepton event can also be achieved from a semileptonic $t\bar{t}$ decay containing a non-prompt lepton, as discussed in Section 7.4. Even when applying tight isolation selections to the leptons, this background is the dominant background in the two same-sign dilepton multilepton search in both the Run-1 and early Run-2 analyses. In same-sign events containing electrons, there is also a non-negligible contribution from charge misidentified electrons, predominantly from Z events for same-sign ee and from $t\bar{t}$ for same-sign $e\mu$. One of the electrons in these events interact with the detector and are reconstructed with the wrong sign charge. Typically these electrons will originate from a W or Z boson, depending on the background.

All the same backgrounds, except for charge misidentification, contribute to the three lepton channel (3ℓ). In the case of non-prompt leptons, a three lepton decay is overwhelmingly from a dileptonic decay of $t\bar{t}$ with one of the remaining b-quarks being reconstructed as a lepton.

The four lepton channel (4ℓ) has negligible backgrounds from $t\bar{t}$ and $t\bar{t}W$, which would require two and one non-prompt lepton(s) in the event respectively. $t\bar{t}Z$ dominates the background with some contribution from VV.

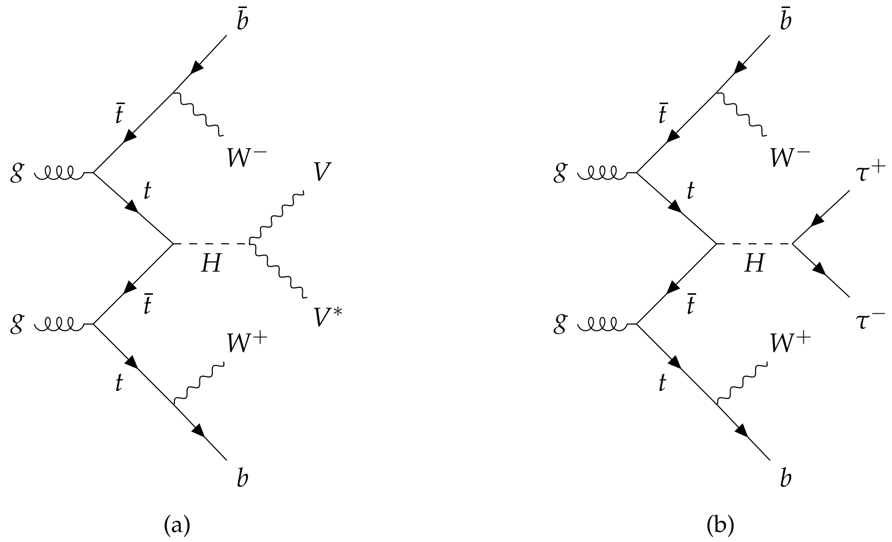


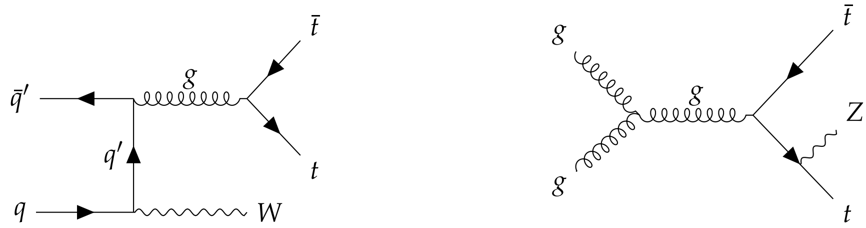
Figure 8.6: Feynman diagrams for the multilepton decays of the $t\bar{t}H$ production process. The Higgs decays to pairs of (a) bosons and (b) τ leptons. V refers to either W or Z bosons. V^* refers to one of the the bosons produced in the Higgs decay being off-shell, due to $2m_V > m_H$.

In addition to the light lepton decays, one can imagine defining channels also containing taus. In a $t\bar{t}H(H \to \tau\tau)$ decay, the final state consists of $WW\tau\tau bb$. Two channels used to search for $t\bar{t}H$ are two same-sign leptons with one hadronic tau (2ℓSS+1τ_{had}) and three leptons with one hadronic tau (3ℓ+1τ_{had}). In $H \to \tau\tau$ decays, both these decays would require one τ decaying hadronically and one τ leptonically. In the 2ℓSS+1τ_{had} decay, the light lepton produced from the τ is required to be the same sign as the leptonically decaying W. In this case the other W would decay hadronically, and in 3ℓ+1τ_{had} it would decay leptonically to ensure three light leptons in the final state. However, in both of these cases there is still a large contribution from $t\bar{t}H(H \to WW)$ with one W in the final state decaying to a hadronic tau, and the other W bosons decaying to produce either same-sign or three light leptons. Dominant backgrounds in these channels are from fake hadronic tau candidates (from a number of sources) and from $t\bar{t}Z$. The 2ℓSS+1τ_{had} channel also contains a significant non-prompt background, as with 2ℓSS.

Other channels including hadronic taus are one lepton with two hadronic taus ($1\ell + 2\tau_{\text{had}}$) and two opposite-sign leptons with one hadronic tau (2ℓOS+1τ_{had}). The $1\ell + 2\tau_{\text{had}}$ channel aims to reconstruct the $H \to \tau\tau$ decay, and including a light lepton from the decays of either of the W bosons from the remaining $t\bar{t}$. 2ℓOS+1τ_{had} aims to retrieve extra events, complementary to 2ℓSS+1τ_{had}, from $t\bar{t}H$ events decaying to two leptons and one hadronic tau. The backgrounds in both of these channels are overwhelmingly dominated by $t\bar{t}$ events containing fake hadronic taus.

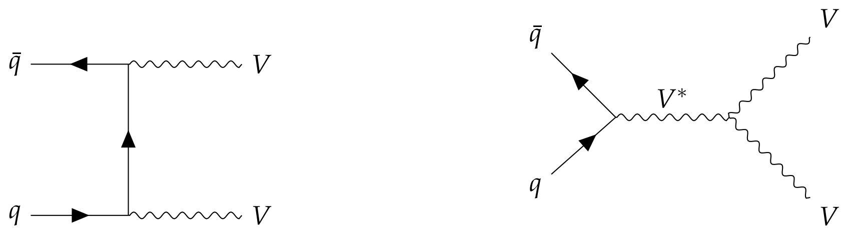
8.4 The search for $t\bar{t}H$ multilepton with 13.2 fb^{-1} of data at $\sqrt{s} = 13$ TeV

A measurement of the multilepton Higgs decays in $t\bar{t}H$ production was made with 13.2 fb^{-1} of early Run-2 data [116]. This analysis was very similar to the analysis performed using 20.3 fb^{-1} data at $\sqrt{s} = 8$ TeV from Run-1 of the LHC [117]. The analysis strategy was "cut-and-count"; signal regions were defined by applying selections on the number of leptons, hadronic taus, jets and b-jets in an event to isolate $t\bar{t}H$ decays in four channels: 2ℓSS, 2ℓSS+1τ_{had}, 3ℓ and 4ℓ.



(a) (b)

Figure 8.7: Feynman diagrams for the dominant contributions to (a) $t\bar{t}W$ and (b) $t\bar{t}Z$ production.



(a) (b)

Figure 8.8: Dominant Feynman diagrams for diboson production at the LHC in (a) the t-channel and (b) the s-channel.

The object and event selections are discussed in Sections 8.4.1. The estimation of backgrounds is shown in Section 8.4.2 and the results of the analysis in Section 8.4.3.

8.4.1 Object and event selections

Both a loose and tight lepton selection are defined. Loose leptons are used in determining the data-driven non-prompt estimates. Tight leptons are those used in the signal regions. The tight definition of the leptons are designed to minimise non-prompt contributions. Both selection criteria for light leptons are shown in Table 8.1.

Hadronic taus and jets are also used in the analysis. The hadronic tau selections are shown in Table 8.2. "Medium" hadronic tau candidates are chosen. The jet (and b-jet) selections are the same as in Table 7.7 above.

	Loose		Tight	
	Electron	Muon	Electron	Muon
p_T [GeV]	> 10	> 10	-	-
$\lvert\eta\rvert$	< 2.47	< 2.5	< 2.47	< 2.5
PID	`LooseLH`	`Loose`	`TightLH`	`Loose`
$\lvert d_0/\sigma_{d_0}\rvert$	< 5	< 3	< 5	< 3
$\lvert z_0 \sin\theta\rvert$ [mm]	< 0.5	< 0.5	< 0.5	< 0.5
Isolation	99% eff.	99% eff.	`isoFixedCutTight`	`isoFixedCutTightTrackOnly`

Table 8.1: Definitions of the loose and tight electron and muon selections. The calorimeter crack region, $1.37 < |\eta| < 1.52$, is vetoed for electrons. The p_T requirements are dependent on region for tight leptons. "99% eff." refers to the `isoLoose` working point [116].

The event selections for the four signal regions are discussed in Table 8.3 below. The reasons behind individual event selections in each of the channels are given in Section 8.5.4 for the updated analysis, which uses very similar definitions.

Variable	Hadronic tau
p_T [GeV]	> 25
$\lvert\eta\rvert$	< 2.47
N_{tracks}	1 or 3
$\sum Q_{\text{tracks}}$	± 1
Jet BDT score	`Medium`
Electron LH veto	pass

Table 8.2: Definitions of the hadronic tau selections. The calorimeter crack region, $1.37 < \lvert\eta\rvert < 1.52$, is vetoed.

Channel	Selection criteria
2ℓSS	Two tight same-sign light leptons No τ_{had} candidates $N_{\text{jets}} \geq 5$, $N_{\text{b-jets}} \geq 1$ $p_T^{0,1} > 25$ GeV Electrons with $\lvert\eta\rvert < 1.37$ $\lvert m(e^{\pm}e^{\pm}) - m_Z\rvert > 10$ GeV for ee events
2ℓSS+$1\tau_{\text{had}}$	Two tight same-sign light leptons One τ_{had} candidate, opposite charge to lepton candidates $N_{\text{jets}} \geq 4$, $N_{\text{b-jets}} \geq 1$ $p_T^0 > 25$ GeV, $p_T^1 > 15$ GeV $\lvert m(e^{\pm}e^{\pm}) - m_Z\rvert > 10$ GeV for ee events
3ℓ	Three light leptons $\sum Q(\ell) = \pm 1$ Same-sign leptons are tight, opposite sign loose $N_{\text{jets}} = 3$ and $N_{\text{b-jets}} \geq 2$ or $N_{\text{jets}} \geq 4$ and $N_{\text{b-jets}} \geq 1$ $p_T^{1,2} > 20$ GeV All SFOS pairs must satisfy $m(\ell^{\pm}\ell^{\pm}) > 12$ GeV All SFOS pairs must satisfy $\lvert m(\ell^{\pm}\ell^{\pm}) - m_Z\rvert > 10$ GeV $\lvert m(3\ell) - m_Z\rvert > 10$ GeV
4ℓ	Four light leptons Leptons pass `isoGradient` selection $\sum Q(\ell) = \pm 0$ $N_{\text{jets}} \geq 2$, $N_{\text{b-jets}} \geq 1$ All SFOS pairs must satisfy $m(\ell^{\pm}\ell^{\pm}) > 12$ GeV All SFOS pairs must satisfy $\lvert m(\ell^{\pm}\ell^{\pm}) - m_Z\rvert > 10$ GeV $100 < m(4\ell) < 250$ GeV $\lvert m(4\ell) - m_H\rvert > 5$ GeV

Table 8.3: Definitions of the signal regions used to search for $t\bar{t}H$ in multilepton final states. SFOS refers to same-flavour opposite-sign lepton pairs. At least one lepton in each region must be matched to an object reconstructed by a single lepton trigger. The leptons are ordered by p_T for the 2ℓSS regions; lepton 0 is leading and lepton 1 is subleading. The leptons are ordered differently in 3ℓ; lepton 0 is the opposite-sign lepton and leptons 1 and 2 are the same-sign pair, ordered by minimum ΔR to lepton 0 [116].

8.4.2 Background estimation

The backgrounds in each channel are estimated a number of different ways. A more thorough explanation of each background estimation technique is given in the context of the updated analysis in Section 8.5.5.

Backgrounds with prompt leptons are estimated using MC simulation. Information about the simulated samples are in [116]. In all regions $t\bar{t}V$ is the dominant prompt background, with VV subdominant. "Rare" prompt processes include tZ, tWZ, $t\bar{t}t\bar{t}$ and $t\bar{t}WW$. tH production is also considered background and is included in the "Rare" category.

There are three other main backgrounds from processes not involving prompt leptons; non-prompt leptons, charge misidentified electrons and fake hadronic taus.

Non-prompt leptons are determined from data-driven techniques using the fake factor method. An extrapolation from control regions enriched in non-prompt leptons is made to the signal region. The control regions are designed to have lower jet multiplicity than the signal region and different lepton selections. Anti-tight leptons are defined as leptons that pass the loose requirement and not the tight leptons. Events with a tight and an anti-tight lepton are therefore enriched in non-prompt. Fake factors can then be defined as the ratio of tight events to anti-tight events in low jet multiplicity in data that are then extrapolated to the signal region. The method requires that the fake factors are flat as a function of N_{jet}.

The fake factors in the 2ℓSS channel are determined for each sub-region: ee, $e\mu$ and $\mu\mu$. These fake factors are then applied to the same-sign leptons in 3ℓ. New fake factors are determined for 2ℓSS+$1\tau_{\text{had}}$, inclusively in lepton flavour due to low statistics.

Electrons with incorrect charge reconstructed are also estimated from data for same-sign ee and $e\mu$ events, in the 2ℓSS and 2ℓSS+$1\tau_{\text{had}}$ channels. The rates in data are determined by comparing opposite-sign and same-sign ee event yields, for events with dielectron invariant mass consistent with m_Z.

Fake τ_{had} candidates are estimated from data for the 2ℓSS+$1\tau_{\text{had}}$ channel. An opposite sign lepton selection with a hadronic tau is used to estimate a scale factor to apply to data. This region is dominated by a dileptonic $t\bar{t}$ decay with a fake τ_{had} candidate.

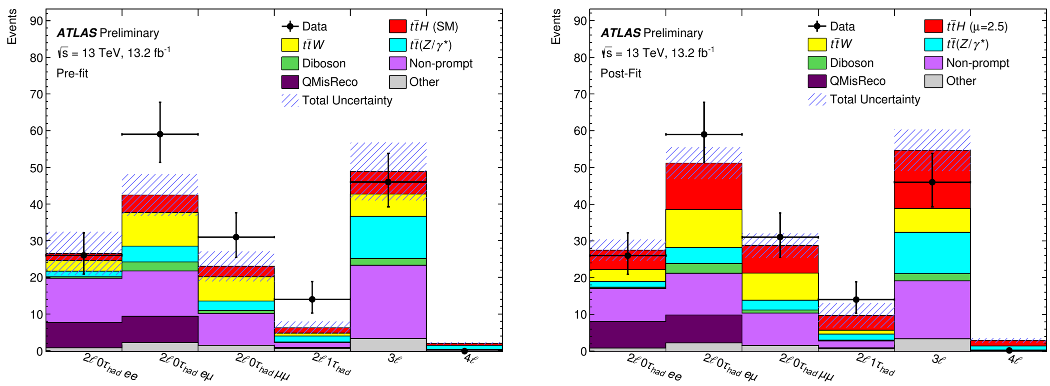
8.4.3 Results

A profile likelihood fit, as discussed in Section 8.2, is performed on the data to measure the signal strength of $t\bar{t}H$ production. The yields of the signal and backgrounds are shown for all channels, both pre-fit and post-fit, in Figure 8.9. The dominant background and uncertainty in the 2ℓSS, 2ℓSS+$1\tau_{\text{had}}$ and the 3ℓ regions are from the non-prompt contribution from $t\bar{t}$ decays. The subdominant background in the 2ℓSSee and 2ℓSS$e\mu$ regions are from charge misreconstructed electrons. Both these backgrounds are reducible.

The best-fit value of signal strength in data is

$$\mu(t\bar{t}H) = 2.5 \pm 0.7(\text{stat})^{+1.1}_{-0.9}(\text{syst}) = 2.5^{+1.3}_{-1.1}.$$

An excess over the SM value is seen, but is consistent within the (large) uncertainties. The analysis is systematically limited, with the largest systematic uncertainty coming from the estimation of the non-prompt and charge misreconstruction backgrounds. Employing the same analysis strategy moving forward with more data will clearly not increase the sensitivity greatly.



(a) (b)

Figure 8.9: The (a) pre-fit and (b) post-fit background and signal predictions and observed data yields for each signal region. For pre-fit the $t\bar{t}H$ yields correspond to the SM expectation ($\mu = 1$) and for post-fit they correspond to the best-fit signal strength value of $\mu = 2.5^{+1.3}_{-1.1}$. Charge misreconstruction backgrounds are indicated as "QMisReco" [116].

A strategy to reduce the yield and associated systematic uncertainty on the measurement of the non-prompt and charge misreconstruction background is vital for an improved analysis.

8.5 The search for $t\bar{t}H$ multilepton with 36.1 fb^{-1} of data at $\sqrt{s} = 13$ TeV

The sensitivity of the search for $t\bar{t}H$ production in multilepton final states with 13.2 fb^{-1} at $\sqrt{s} = 13$ TeV is poor due to a few main factors:

- *Non-prompt lepton yields*
  A large (relative) number of non-prompt leptons pass the tightest (calibrated) isolation requirements.
- *Charge misreconstructed electrons*
  Electrons with incorrect charge reconstruction from Z and $t\bar{t}$ events in the same-sign electron signal region are not targeted for reduction from lepton selections, since such electrons are typically isolated. An inner B-layer hit in the inner detector is used in the definition of the electron likelihood identification working points used, but this requirement does not reject this background sufficiently.
- *Cut-and-count analysis*
  A simple cut-based approach was used to define orthogonal signal regions by requiring different multiplicities of leptons, jets and b-jets. In this approach $\mu_{t\bar{t}H}$ can simply be measured by fitting a global normalisation, using the yields in each signal region. This approach can be further optimised by applying multivariate algorithms using event selections, to either define further signal-enriched regions or to maximise the separation of $t\bar{t}H$ and the dominant backgrounds in looser pre-selection regions.

These items are addressed in the analysis undertaken with full 2015 and 2016 data to maximise sensitivity to $t\bar{t}H$ production in multilepton final states, discussed below.

Firstly the data and MC samples used in the analysis are shown in Section 8.5.1. Secondly, a lepton selection optimisation study designed to reduce non-prompt and charge misreconstruction

Process	Event generator	Matrix element	Parton shower	PDF
$t\bar{t}H$	MadGraph5_aMC@NLO [119]	NLO	Pythia 8	NNPDF 3.0 NLO [120]
$t\bar{t}W$	MadGraph5_aMC@NLO	NLO	Pythia 8	NNPDF 3.0 NLO
$t\bar{t}(Z/\gamma^* \to ll)$	MadGraph5_aMC@NLO	NLO	Pythia 8	NNPDF 3.0 NLO
$t\bar{t}$	Powheg-BOX v2 [121]	NLO	Pythia 8	NNPDF 3.0 NLO
$VV(\to llXX)$	Sherpa 2.1.1	NLO	Sherpa	CT10 [102]
$Z \to ll$	Sherpa 2.2.1	NLO	Sherpa	NNPDF 3.0 NLO

Table 8.4: The configurations used for event generation of signal and background processes [118]. The PDFs listed are those used in the matrix element generation. The PDF set used for the parton shower in all samples is NNPDF 2.3 LO [122]. Samples using Pythia 8 have heavy flavour hadron decays modeled by EvtGen.

yields is shown in Section 8.5.2. The definition of objects, channels and background estimations are discussed in Sections 8.5.3, 8.5.4 and 8.5.5 respectively. Finally the systematic uncertainties and results of the analysis are discussed in Sections 8.5.6 and 8.5.7 respectively.

8.5.1 Data and simulation samples

8.5.1.1 Data sample

The data were collected by the ATLAS detector during proton-proton LHC physics running in 2015 and 2016, at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The mean number of proton-proton collisions per bunch crossing, $\langle\mu\rangle$, is 24 in the combined dataset. Only events passing operational quality requirements are used, resulting in a total integrated luminosity of 36.1 fb^{-1}.

8.5.1.2 MC samples

Simulated MC samples of signal and background processes are used. Nominally the samples are produced using a full Geant4 ATLAS detector simulation. The simulation of pileup is applied by generating inelastic proton-proton collisions with Pythia 8, which are overlaid on the generated hard scatter events. The underlying pileup distribution is then reweighted to describe that of the data, with the caveat discussed in Chapter 7. A summary of the different generators, parton showers and PDFs used to model the signal and the main background processes is shown in Table 8.4. A more complete table, including information of the generation of rarer background processes, can be found in [118].

8.5.2 Optimising tight lepton selections

Lepton definitions for the same-sign leptons in the 2ℓSS and 3ℓ signal regions in [116] are designed to maximise non-prompt rejection. However, the non-prompt background is still the dominant background in the analysis. This problem needs to be addressed to increase the sensitivity of the updated analysis with full 2015 and 2016 data.

The `isoFixedCutTight(TrackOnly)` isolation working points for electrons(muons) were used in the previous analysis. As discussed in Chapter 7, these isolation working points can be replaced with non-prompt BDT working points. Here optimisations of such working points designed to maximise the sensitivity of measuring $t\bar{t}H$ production in multilepton final states in the 2ℓSS and 3ℓ channels are shown. These channels have the largest signal statistics but also the largest non-prompt backgrounds. Reducing the number of non-prompt background events in these regions will greatly improve the sensitivity of the analysis.

Unlike the `isoFixedCutTight(TrackOnly)` working points that have a fixed $t\bar{t}H$ signal efficiency and non-prompt rejection in the corresponding signal regions, a cut on the non-prompt BDT can be tuned to improve the overall sensitivity of the analysis. For an analysis employing event MVA selections, typically a pre-selection region is defined. This region is not that used to measure μ from a normalisation, but rather a region in which multivariate selections can be used to further separate $t\bar{t}H$ events from background events. Therefore one wishes to retain maximum $t\bar{t}H$ acceptance. Both the signal-over-background (S/B) figure of merit and the significance ($S/\sqrt{B}$) are used below to determine the optimal lepton selection.

The initial signal region definitions used to optimise the non-prompt BDT working points in the 2ℓSS and 3ℓ channels are shown in Table 8.3 above. They mirror those used in the previous analysis. The optimisation of lepton working points with the non-prompt BDT involves loosening these signal region requirements, which are (in general) designed to reduce the non-prompt backgrounds. In this way $t\bar{t}H$ acceptance can be increased, with the non-prompt BDT being used to control the number of non-prompt leptons allowed in the region.

8.5.2.1 2ℓSS electron definition

Electrons in the 2ℓSS region suffer from both non-prompt and charge misreconstruction reducible backgrounds. The non-prompt BDT addresses the non-prompt background. A similar BDT algorithm, named `CFT` ("Charge Flip Tagger"), was developed to reject charge misreconstruction. Two working points are defined, designed to retain 95% and 97% of electrons with correctly assigned charge. Each working point is designed to reject charge misreconstructed electrons with approximate factors of 17 and 11 for `TightLH` electrons respectively. The 95% working point is used to determine the optimal non-prompt BDT cut for the same-sign electron region to ensure maximum charge misidentified electrons are rejected.

Table 8.5 describes the new tight electron lepton definitions, compared to the previous definitions. In the previous analysis electrons in this region were restricted to the barrel region ($|\eta| < 1.7$). This is due to the increase in likelihood of charge misreconstructed electrons in the forward regions of the detector, due to the extra material in which the electron must pass. The sensitivity using the full range of η ($|\eta| < 2.5$) for electrons is shown, with the `CFT` algorithm being used to control charge misreconstructed events allowed in by this requirement. The `isoLoose` isolation working point is used in conjunction with the non-prompt BDT cut. This cut is used as a minimum isolation baseline for the non-prompt BDT.

The electron selections discussed are applied to both electrons in the same-sign region. One might expect the sub-leading lepton to normally be non-prompt but, as shown in Chapter 7, there is a non-negligible amount of non-prompt leading leptons in same-sign regions. Applying non-prompt BDT selections to both leptons results in a further reduction of signal due to inefficiencies but this is offset by further reduction in non-prompt background.

The yields for a number of different non-prompt BDT working points are shown in four different same-sign electron event selections in Table 8.6. The effect of loosening event region cuts, such as N_{jet} and subleading lepton p_T, to increase $t\bar{t}H$ acceptance is shown. Backgrounds from $t\bar{t}$, $t\bar{t}Z$, $t\bar{t}W$, VV and Z+jets are considered in the total background, and are taken from simulation. The normalisation of the non-prompt contribution from $t\bar{t}$ is taken from MC and increased by a factor 1.5. This factor is taken from the previous analysis and corresponds roughly to the increased number of isolated non-prompt lepton data events compared to

Variable	Previous lepton selection	New lepton selection
$\lvert\eta\rvert$	< 1.37	< 1.37 & $1.52 < \lvert\eta\rvert < 2.5$
$\lvert d_0\rvert/\sigma_{d_0}$	< 5	< 5
$\lvert z_0 \sin(\theta)\rvert$ [mm]	< 0.5	< 0.5
PID	`TightLH`	`TightLH`
Isolation	`isoFixedCutTight`	`isoLoose`
CFT	-	95% WP
Non-prompt BDT	-	TBC

Table 8.5: Definitions of the previous and new electron selections. The `CFT` working point is a one dimensional cut (`CFT` > 0.067) designed to retain 95% of electrons with correct charge identification. The value of the non-prompt BDT cut used in the selection is shown below.

MC simulation. A simple 30% relative uncertainty is applied to the normalisation of $t\bar{t}$ to mimic typical uncertainties from data driven measurements of the non-prompt and charge misreconstruction backgrounds from the previous analysis. This is the only uncertainty applied to the total background; the other uncertainties on the prompt backgrounds are typically much smaller and are taken from MC simulation.

The five BDT working points, combined with the 95% `CFT` working point, can be compared to the previous tight isolation requirements. The addition of the `CFT` working point is also shown in conjunction with the tight isolation requirement for each event selection. For a 10% loss of $t\bar{t}H$ events, the `CFT` reduces the amount of $t\bar{t}$ in the nominal signal region by roughly 65%, when used with `isoFixedCutTight`. $t\bar{t}$ can further be reduced by 65% when `CFT` is used in conjunction with the non-prompt BDT < -0.5 working point. The acceptance of $t\bar{t}H$ reduces by approximately 20% but the significance increases. However, this loss of acceptance can be mitigated by loosening event selections.

With the previous lepton selections and signal region definitions, one would expect roughly 5 $t\bar{t}H$ events with 65 total background events, 48 of which from $t\bar{t}$ events (in red). If one relaxes the η range of electrons and reduces the subleading lepton p_T to 20 GeV, then we find that the non-prompt BDT < -0.5 + `CFT` working point results in roughly 6 signal events, with 40 total background, 16 from $t\bar{t}$ (in green). That is, one more signal event for a reduction of 25 total background events. The use of both algorithms greatly increases the sensitivity of the same-sign electron channel. One can further gain $t\bar{t}H$ acceptance by reducing $N_{\text{jet}} \geq 4$ but non-prompt contributions begin to increase. Event multivariate algorithms are used to further reduce these non-prompt events.

8.5.2.2 2ℓSS muon definition

Table 8.7 compares the previous and new lepton selections for muons. There is no charge misreconstruction background for muons and so only non-prompt BDT working points are determined. The yields for the previous signal region and for regions with loosened requirements are shown in Table 8.8. By loosening the subleading lepton p_T to 20 GeV, one can roughly increase the number of $t\bar{t}H$ events by 1 for roughly the same number of non-prompt background from $t\bar{t}$ when using a non-prompt BDT < -0.5 working point (in green), compared to the previous signal region and isolation selection (in red). One can see that the significance increases for ever tightening cuts on the non-prompt BDT. However, multivariate techniques are still to be employed to reduce the non-prompt background, so the non-prompt BDT < -0.5 working point is a good compromise to retain signal whilst still rejecting large amounts of

Event/lepton selection	Isolation selection	$t\bar{t}H$ (S)	Total bkg. (B)	S/B	$S/\sqrt{B}$	$S/\sqrt{B+\sigma_B^2}$	$t\bar{t}$
$N_{\text{jets}} \geq 5$, $p_T^1 > 25$ GeV, $\lvert\eta^{0,1}\rvert < 1.37$	BDT < -0.70 + CFT	3.4	16.9	0.20	0.83	0.79	4.4
	BDT < -0.60 + CFT	3.8	18.5	0.20	0.88	0.82	5.4
	BDT < -0.50 + CFT	4.0	19.9	0.20	0.89	0.82	6.3
	BDT < -0.40 + CFT	4.2	23.1	0.18	0.87	0.76	9.2
	BDT < -0.30 + CFT	4.3	24.2	0.18	0.88	0.75	10.0
	isoFixedCutTight	5.3	64.6	0.08	0.66	0.32	47.5
	isoFixedCutTight + CFT	4.8	32.9	0.15	0.84	0.60	18.5
$N_{\text{jets}} \geq 5$, $p_T^1 > 25$ GeV, Full range $\lvert\eta^{0,1}\rvert$	BDT < -0.70 + CFT	4.4	27.7	0.16	0.84	0.76	8.5
	BDT < -0.60 + CFT	4.9	30.3	0.16	0.88	0.78	10.0
	BDT < -0.50 + CFT	5.1	32.1	0.16	0.91	0.78	11.0
	BDT < -0.40 + CFT	5.4	36.5	0.15	0.90	0.73	14.5
	BDT < -0.30 + CFT	5.6	38.1	0.15	0.90	0.72	15.6
	isoFixedCutTight	7.2	145.3	0.05	0.60	0.20	112.3
	isoFixedCutTight + CFT	6.3	49.8	0.13	0.89	0.58	27.6
$N_{\text{jets}} \geq 5$, $p_T^1 > 20$ GeV, Full range $\lvert\eta^{0,1}\rvert$	BDT < -0.70 + CFT	5.3	34.7	0.15	0.91	0.76	13.0
	BDT < -0.60 + CFT	5.8	38.0	0.15	0.95	0.76	15.0
	BDT < -0.50 + CFT	6.1	40.2	0.15	0.97	0.77	16.3
	BDT < -0.40 + CFT	6.5	45.6	0.14	0.96	0.71	20.8
	BDT < -0.30 + CFT	6.6	48.9	0.14	0.95	0.67	23.5
	isoFixedCutTight	8.5	167.2	0.05	0.65	0.21	130.7
	isoFixedCutTight + CFT	7.4	64.5	0.12	0.92	0.52	39.4
$N_{\text{jets}} \geq 4$, $p_T^1 > 20$ GeV, Full range $\lvert\eta^{0,1}\rvert$	BDT < -0.70 + CFT	7.5	65.9	0.11	0.93	0.66	26.4
	BDT < -0.60 + CFT	8.1	73.2	0.11	0.95	0.64	31.1
	BDT < -0.50 + CFT	8.6	78.9	0.11	0.97	0.62	35.1
	BDT < -0.40 + CFT	9.1	91.1	0.10	0.95	0.57	42.0
	BDT < -0.30 + CFT	9.3	97.3	0.10	0.94	0.54	47.2
	isoFixedCutTight	11.6	437.8	0.03	0.56	0.11	341.5
	isoFixedCutTight + CFT	10.1	134.1	0.08	0.88	0.37	83.9

Table 8.6: Simulation yields for the same-sign electron signal region, normalised to 36.1 fb^{-1}. The top section corresponds to the event selections used in the signal region in the previous analysis. The line highlighted in red corresponds to the yields using the previous analysis' lepton and event selections. The following sections contain relaxed event selections; by increasing the electron η range, reducing subleading lepton p_T and jet multiplicity, N_{jets}. The line highlighted in green corresponds to the new lepton selection in a region with looser event selections. An `isoLoose` requirement is combined with the non-prompt BDT cuts. The normalisation of $t\bar{t}$ is taken as that from MC multiplied by a factor of 1.5 and the uncertainty on the total background, σ_B, is purely assumed to be 30% from $t\bar{t}$, with other background uncertainties ignored.

non-prompt leptons.

One sees better relative improvements in the electron selection than the muon selection. One large difference is the use of `CFT` in conjunction with the non-prompt BDT, and the "double" background of charge misidentification and non-prompt in the region. The track isolation requirement from the standard isolation working points on electrons are also looser than for muons, with a slightly smaller cone size being used. Nevertheless, the non-prompt BDT does improve the significance in the same-sign muon region and allows event selections to be relaxed without massively increasing the number of non-prompt leptons, as required.

8.5.2.3 2ℓSS lepton selections in 2ℓSS $e\mu$

The increase in sensitivity in using the non-prompt BDT "isolation" working points deduced in Sections 8.5.2.1 and 8.5.2.3 can be verified in the opposite-flavour 2ℓSS region.

Table 8.9 shows the yields in a similar way to above. The other non-prompt BDT working points are shown for reference and the `CFT` working point is used with the electron selections.

Variable	Previous lepton selection	New lepton selection
$\lvert\eta\rvert$	< 2.5	< 2.5
$\lvert d_0\rvert/\sigma_{d_0}$	< 3	< 3
$\lvert z_0 \sin(\theta)\rvert$ [mm]	< 0.5	< 0.5
PID	Loose	Loose
Isolation	`isoFixedCutTightTrackOnly`	`isoLoose`
Non-prompt BDT	-	TBC

Table 8.7: Definitions of the previous and new muon selections. The definition of the non-prompt BDT cut is shown below.

Event/lepton selection	Isolation selection	$t\bar{t}H$ (S)	Total bkg. (B)	S/B	$S/\sqrt{B}$	$S/\sqrt{B+\sigma_B^2}$	$t\bar{t}$
$N_{\text{jets}} \geq 5$, $p_T^1 > 25$ GeV	BDT < -0.70	8.4	39.1	0.22	1.35	1.29	6.4
	BDT < -0.60	8.9	41.8	0.21	1.37	1.29	7.9
	BDT < -0.50	9.2	42.0	0.22	1.42	1.32	8.7
	BDT < -0.40	9.5	43.2	0.22	1.44	1.33	9.1
	BDT < -0.30	9.7	45.2	0.21	1.44	1.32	9.8
	`isoFixedCutTightTrackOnly`	9.8	50.9	0.19	1.37	1.18	14.3
$N_{\text{jets}} \geq 5$, $p_T^1 > 20$ GeV	BDT < -0.70	9.7	46.9	0.21	1.42	1.29	10.4
	BDT < -0.60	10.3	50.4	0.20	1.45	1.28	12.4
	BDT < -0.50	10.7	52.4	0.20	1.48	1.26	14.8
	BDT < -0.40	11.0	54.0	0.20	1.50	1.27	15.7
	BDT < -0.30	11.3	56.5	0.20	1.50	1.25	16.7
	`isoFixedCutTightTrackOnly`	11.4	67.7	0.17	1.38	0.99	26.6
$N_{\text{jets}} \geq 4$, $p_T^1 > 20$ GeV	BDT < -0.70	14.0	88.6	0.16	1.49	1.18	23.9
	BDT < -0.60	14.7	98.5	0.15	1.48	1.08	31.5
	BDT < -0.50	15.3	105.2	0.15	1.49	1.01	36.9
	BDT < -0.40	15.8	111.0	0.14	1.50	0.97	41.5
	BDT < -0.30	16.1	114.7	0.14	1.50	0.94	44.3
	`isoFixedCutTightTrackOnly`	16.1	145.3	0.11	1.34	0.65	72.8

Table 8.8: Simulation yields for the same-sign muon signal region, normalised to 36.1 fb^{-1}. The top section corresponds to the event selections used in the signal region in the previous analysis. The line highlighted in red corresponds to the yields using the previous analysis' lepton and event selections. The following sections contain relaxed event selections; by reducing subleading lepton p_T and jet multiplicity, N_{jets}. The line highlighted in green corresponds to the new lepton selection in a region with looser event selections. An `isoLoose` requirement is combined with the non-prompt BDT cuts. The normalisation of $t\bar{t}$ is taken as that from MC multiplied by a factor of 1.5 and the uncertainty on the total background, σ_B, is purely assumed to be 30% from $t\bar{t}$, with other background uncertainties ignored.

Comparing the previous signal region and lepton selection (in red) to the proposed working point in the region relaxed in electron η, subleading p_T and N_{jet} (in green) one can see that there is approximately a reduction of 20 non-prompt (and charge misidentification) events for an increase of 6 signal events. The working points do an excellent job at rejecting $t\bar{t}$ events in same-sign dilepton regions.

8.5.2.4 3ℓ lepton definitions

In the 3ℓ signal region, two of the three leptons are required to be the same-sign and the other the opposite sign to the pair. Lepton 0 is defined as the opposite sign lepton, and leptons 1 and 2 the same-sign pair. One would expect one of the same-sign pair to be the non-prompt lepton in $t\bar{t}$ events passing the 3ℓ event selections; roughly 97% of the time [118]. In the same way as in 2ℓSS above, non-prompt BDT working points are only applied to the same-sign pair.

One does not expect charge misidentified electrons in 3ℓ. The use of `CFT` is not applied to the

Event/lepton selection	Isolation selection	$t\bar{t}H$ (S)	Total bkg. (B)	S/B	$S/\sqrt{B}$	$S/\sqrt{B+\sigma_B^2}$	$t\bar{t}$
$N_{\text{jets}} \geq 5$, $p_T^1 > 25$ GeV, $\lvert\eta_e\rvert < 1.37$	BDT < -0.70 (+ `CFT`)	10.7	56.5	0.19	1.42	1.26	13.2
	BDT < -0.60 (+ `CFT`)	12.6	60.8	0.21	1.61	1.40	15.0
	BDT < -0.50 (+ `CFT`)	13.1	62.7	0.21	1.65	1.42	15.7
	BDT < -0.40 (+ `CFT`)	13.7	66.2	0.21	1.68	1.40	17.9
	BDT < -0.30 (+ `CFT`)	14.0	67.4	0.21	1.71	1.42	18.4
	`isoFixedCutTight(TrackOnly)`	15.5	127.2	0.12	1.37	0.61	75.8
	`isoFixedCutTight(TrackOnly)` (+ `CFT`)	15.0	86.6	0.17	1.61	1.04	36.9
$N_{\text{jets}} \geq 5$, $p_T^1 > 25$ GeV, Full range $\lvert\eta_e\rvert$	BDT < -0.70 (+ `CFT`)	10.6	73.1	0.15	1.24	1.04	18.6
	BDT < -0.60 (+ `CFT`)	12.9	79.0	0.16	1.45	1.18	21.5
	BDT < -0.50 (+ `CFT`)	14.8	83.1	0.18	1.63	1.28	23.7
	BDT < -0.40 (+ `CFT`)	15.5	86.9	0.18	1.66	1.28	26.0
	BDT < -0.30 (+ `CFT`)	15.9	89.5	0.18	1.68	1.27	27.6
	`isoFixedCutTight(TrackOnly)`	18.2	246.2	0.07	1.16	0.33	177.9
	`isoFixedCutTight(TrackOnly)` (+ `CFT`)	17.1	113.0	0.15	1.61	0.93	49.6
$N_{\text{jets}} \geq 5$, $p_T^1 > 20$ GeV, Full range $\lvert\eta_e$	BDT < -0.70 (+ `CFT`)	10.4	66.4	0.16	1.28	1.06	18.6
	BDT < -0.60 (+ `CFT`)	12.7	71.9	0.18	1.50	1.20	21.3
	BDT < -0.50 (+ `CFT`)	15.0	75.1	0.20	1.74	1.36	23.0
	BDT < -0.40 (+ `CFT`)	15.8	80.1	0.20	1.76	1.32	26.5
	BDT < -0.30 (+ `CFT`)	15.5	82.8	0.19	1.71	1.25	28.3
	`isoFixedCutTight(TrackOnly)`	17.8	164.0	0.11	1.39	0.51	107.4
	`isoFixedCutTight(TrackOnly)` (+ `CFT`)	17.3	117.0	0.15	1.60	0.80	62.2
$N_{\text{jets}} \geq 4$, $p_T^1 > 20$ GeV, Full range $\lvert\eta_e^{0,1}\rvert$	BDT < -0.70 (+ `CFT`)	17.8	131.7	0.14	1.55	1.02	43.6
	BDT < -0.60 (+ `CFT`)	20.5	143.0	0.14	1.72	1.07	50.2
	BDT < -0.50 (+ `CFT`)	21.4	151.4	0.14	1.74	1.03	55.6
	BDT < -0.40 (+ `CFT`)	22.4	163.4	0.14	1.75	0.96	65.1
	BDT < -0.30 (+ `CFT`)	22.9	169.6	0.14	1.76	0.93	69.5
	`isoFixedCutTight(TrackOnly)`	24.7	358.4	0.07	1.30	0.31	255.9
	`isoFixedCutTight(TrackOnly)` (+ `CFT`)	23.8	243.7	0.10	1.53	0.52	144.8

Table 8.9: Simulation yields for the same-sign opposite-flavour signal region, normalised to 36.1 fb^{-1}. The top section corresponds to the event selections used in the signal region in the previous analysis. The line highlighted in red corresponds to the yields using the previous analysis' lepton and event selections. The following sections contain relaxed event selections; by increasing the electron η range, reducing subleading lepton p_T and jet multiplicity, N_{jets}. The line highlighted in green corresponds to the new lepton selection in a region with looser event selections. An `isoLoose` requirement is combined with the non-prompt BDT cuts and the 95% `CFT` working point to electrons where stated. The normalisation of $t\bar{t}$ is taken as that from MC multiplied by a factor of 1.5 and the uncertainty on the total background, σ_B, is purely assumed to be 30% from $t\bar{t}$, with other background uncertainties ignored.

electrons in the same-sign pair.

There is a large combination of flavours allowed in the 3ℓ region. To optimise the working points, only the inclusive (all flavours) region is used. Both electron and muon non-prompt BDT selections are changed simultaneously with the same value. If one takes the non-prompt BDT selection from 2ℓSS (without `CFT` for electrons), with a reduced subleading lepton p_T (in green) and compare it to the previous nominal (in red) then we can see that roughly the same $t\bar{t}H$ acceptance is retained with a reduction of roughly 40% of the non-prompt contribution from $t\bar{t}$. The significance increase is not as great as in 2ℓSS due to larger relative contributions from the prompt backgrounds.

Event/lepton selection	Isolation selection	$t\bar{t}H$ (S)	Total bkg. (B)	S/B	$S/\sqrt{B}$	$S/\sqrt{B+\sigma_B^2}$	$t\bar{t}$
$p_T^{1,2} > 20$ GeV	BDT < -0.70	15.1	58.2	0.26	1.98	1.94	5.1
	BDT < -0.60	16.2	62.4	0.26	2.05	1.99	6.6
	BDT < -0.50	16.9	65.6	0.26	2.09	2.01	7.8
	BDT < -0.40	17.5	68.3	0.26	2.12	2.02	9.1
	BDT < -0.30	18.1	70.8	0.26	2.15	2.02	10.1
	`isoFixedCutTight(TrackOnly)`	19.8	87.8	0.23	2.12	1.72	22.3
$p_T^{1,2} > 15$ GeV	BDT < -0.70	17.6	68.1	0.26	2.13	2.03	9.0
	BDT < -0.60	18.9	74.6	0.25	2.19	2.02	12.0
	BDT < -0.50	19.9	79.4	0.25	2.23	2.02	14.1
	BDT < -0.40	20.7	83.7	0.25	2.26	1.99	16.4
	BDT < -0.30	21.3	88.0	0.24	2.27	1.95	18.9
	`isoFixedCutTight(TrackOnly)`	23.1	119.0	0.19	2.11	1.33	45.0

Table 8.10: Simulation yields for the inclusive 3ℓ signal region, normalised to 36.1 fb^{-1}. The top section corresponds to the event selections used in the signal region in the previous analysis. The line highlighted in red corresponds to the yields using the previous analysis' lepton and event selections. The following section contains a relaxed event selection, by reducing subleading lepton p_T. The line highlighted in green corresponds to the new lepton selection in a region with looser event selections. An `isoLoose` requirement is combined with the non-prompt BDT cuts and both cut values are applied to both muons and electrons simultaneously. The normalisation of $t\bar{t}$ is taken as that from MC multiplied by a factor of 1.5 and the uncertainty on the total background, σ_B, is purely assumed to be 30% from $t\bar{t}$, with other background uncertainties ignored.

8.5.2.5 Lepton definition summary

Three non-prompt BDT working points are optimised to reject the non-prompt lepton (and charge misidentification) contributions in the 2ℓSS and 3ℓ regions. These working points are calibrated with the Z tag and probe method. The muon calibration procedure is discussed in Chapter 7. The two electron working points are calibrated in a similar manner [118].

The efficiency of the non-prompt BDT electron working point is shown in Figure 8.10. The overall efficiency is slightly lower (corresponding to a tighter cut) than for muons, but the scale factors are very similar. The scale factors determined from the calibration procedures are applied to the leptons hereafter.

Different selections of leptons can thus be defined. In total five selections are made, and are as shown in Table 8.11. They are loose; loose and isolated; loose, isolated and passing the non-prompt BDT; tight and very tight. They define whether the lepton passes certain isolation, identification and non-prompt BDT criteria (and `CFT` for electrons). Five separate definitions are needed to describe each type of electron used in the analysis. In fact only three muon definitions are used, with the final three selections (L*, T and T*) resulting in identical definitions. The different definitions are designed for use in the non-prompt and fake lepton estimates (see Section 8.5.5) and to maximise the significance in different regions with differing statistical accuracies.

8.5.3 Object selections

The light lepton selections are discussed in Section 8.5.2. Data/MC scale factors are applied to account for the reconstruction, identification, "isolation" and trigger selections. Jets, b-jets and hadronic tau candidates are also used in the analysis.

The jet selections are the same as in the previous analysis (see Table 7.7), as is the b-jet selection

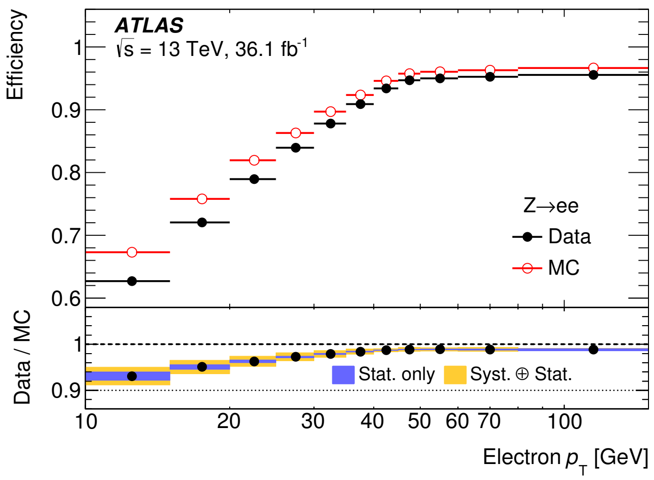


Figure 8.10: The efficiency to select electrons passing the non-prompt lepton BDT working point as a function of the lepton p_T in $Z \to ee$ events. The measurements in data are shown as full black circles and simulation are shown in open red circles. The ratio of the efficiency in data with respect to simulation with statistical (blue) and total (yellow) uncertainties is shown [118].

Lepton selections	Electrons, e					Muons, μ		
	L	L†	L*	T	T*	L	L†	L*/T/T*
`isoLoose`	✗	✓	✓	✓	✓	✗	✓	✓
Non-prompt BDT	✗	✗	✓	✓	✓	✗	✗	✓
`CFT`	✗	✗	✗	✗	✓	-		
Identification	`LooseLH`			`TightLH`		`Loose`		
$\lvert d_0 \rvert / \sigma(d_0)$	< 5					< 3		
$\lvert z_0 \sin\theta \rvert$	< 0.5 mm							

Table 8.11: Loose (L); loose and isolated (L†); loose, isolated and passing the non-prompt BDT (L*); tight (T) and very tight (T*) light lepton definitions. For muons, the L*, T and T* definitions are identical [118].

(the 70% MV2c10 working point). Scale factors to correct the simulation to the data are applied for the selections on the jet vertex tagger and for the MV2c10 selction for b-jets.

The hadronic tau selections are also the same as the previous analysis (see Table 8.2), but with a few additions. Two more tau selections are introduced in addition to the "medium" τ_{had} candidates (those passing the `Medium` identification working point); "loose" and "tight". Tight hadronic taus pass the `Tight` hadronic tau identification requirement, and loose hadronic taus pass a very loose selection on the tau jet BDT score. These taus are used in the estimates of fake hadronic taus in a similar way to the leptons. In addition to the electron likelihood veto, the contribution of hadronic taus faked by b-jets are reduced by vetoing b-tagged candidates and the tau vertex is required to match to the primary vertex in an event to reduce pile up jets faking hadronic tau candidates. Data/MC scale factors to correct for the reconstruction and identification of hadronic taus are also applied.

The objects in the analysis undergo the same overlap removal as in Table 7.9, with the addition of the removal of electrons and muons within $\Delta R < 0.2$ of a hadronic tau candidate. In the case of low p_T muons, the hadronic tau candidate is removed to better reject the number of fake hadronic taus produced from muons.

8.5.4 Channel definitions and strategies

All channels discussed in Section 8.3 are used in the search for $t\bar{t}H$ in multilepton final states. The summary of the channels defined by the multiplicity of loose leptons and hadronic taus in the final state are shown in Figure 8.11. The event selections used to define each channel are shown in Table 8.12.

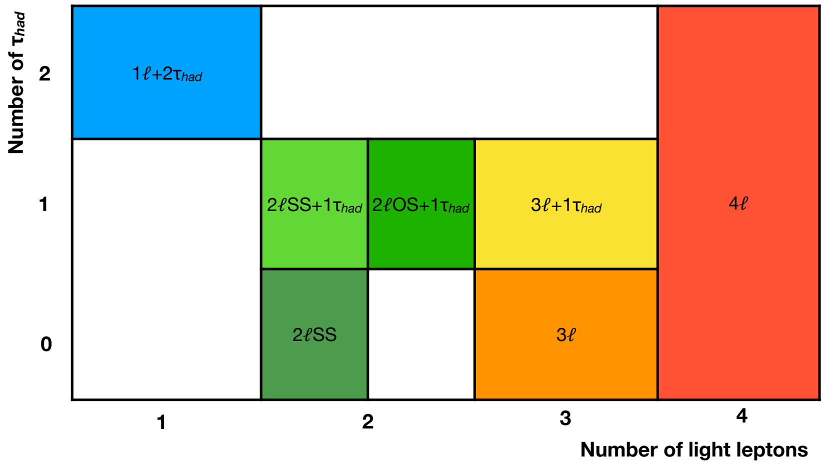


Figure 8.11: The seven orthogonal channels used in the analysis, displayed as the number of light leptons versus the number of hadronic taus [118].

The selection of each event is based, in some way, on the presence of light leptons. Both single lepton and dilepton triggers are used to record such events in data. Due to different pileup conditions, the lepton p_T thresholds differ in 2015 compared to 2016. In 2015 data, the single electron (single muon) trigger required a lepton candidate with $p_T > 24(20)$ GeV, which increased to 26 GeV in 2016 data. This increase in p_T in 2016 was chosen so that the bandwidth required to keep the trigger unprescaled was at a manageable level. The dielectron (dimuon) triggers had thresholds of 12+12 (18+8) GeV in 2015, which increased to 17+17 (22+8) GeV in 2016. The electron-muon dilepton trigger had thresholds of 17+14 GeV in both 2015 and 2016. The muon legs of each of these triggers are discussed in the context of trigger scale factors for the non-prompt BDT in Chapter 7. A combination of the single and dilepton triggers are used in all regions except 1ℓ+$2\tau_{\text{had}}$, where only a single lepton trigger can be used.

Each channel has a strategy in which to measure $t\bar{t}H$. These are stated below, with a description of the event selections.

8.5.4.1 2ℓSS

The event selections used to define the 2ℓSS region shown in Table 8.12 correspond to a pre-selection region. Two very tight same-sign light leptons are required in an event, with $p_T > 20$ GeV. Events with at least four jets are required and events with three or more b-jets are vetoed. The reason behind this b-jet veto is not based on maximising the search significance, but rather due to a large mismodelling of data and simulation in such events. To avoid potential biases in the fit these events are rejected.

To further reject the dominant $t\bar{t}$ and $t\bar{t}W$ backgrounds in this region, two BDT algorithms are separately trained to reject each background. The outputs of the two BDTs are then combined to produce a single classifier with which to measure $t\bar{t}H$.

Channel	Selection criteria
All	$N_{\text{jets}} \geq 2$ and $N_{\text{b-jets}} \geq 1$
2ℓSS	Two very tight same-sign leptons, with $p_T > 20$ GeV Zero medium τ_{had} candidates $N_{\text{jets}} \geq 4$ and $N_{\text{b-jets}} < 3$
3ℓ	Three leptons with $p_T > 10$ GeV; sum of charges must be ±1 Two same-sign leptons must be very tight, with $p_T > 15$ GeV Opposite-sign lepton must be loose, isolated and pass the non-prompt BDT Zero medium τ_{had} candidates $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - m_Z\rvert > 10$ GeV for all SFOS pairs $\lvert m(3\ell) - m_Z\rvert > 10$ GeV
4ℓ	Four leptons; sum of lepton charges must be zero Third and fourth leading leptons must be tight $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - m_Z\rvert > 10$ GeV for all SFOS pairs $\lvert m(4\ell) - m_H\rvert > 5$ GeV
Z-depleted	Zero SFOS pairs
Z-enriched	Either two or four SFOS pairs
1ℓ+$2\tau_{\text{had}}$	One tight lepton with $p_T > 27$ GeV Two medium τ_{had} candidates of opposite charge, at least one being tight $N_{\text{jets}} \geq 3$
2ℓSS+$1\tau_{\text{had}}$	Two very tight same-sign leptons, with $p_T > 15$ GeV One medium τ_{had} candidate, opposite-sign to the lepton pair $N_{\text{jets}} \geq 4$ $\lvert m(e^\pm e^\pm) - m_Z\rvert > 10$ GeV for $e^\pm e^\pm$ events
2ℓOS+$1\tau_{\text{had}}$	Two loose and isolated leptons with $p_T^0 > 25$, $p_T^1 > 15$ GeV Opposite-sign leptons One medium τ_{had} candidate $m(\ell^+\ell^-) > 12$ GeV and $\lvert m(\ell^+\ell^-) - m_Z\rvert > 10$ GeV for the SFOS pair $N_{\text{jets}} \geq 3$
3ℓ+$1\tau_{\text{had}}$	Same as 3ℓ, except for: One medium τ_{had} candidate Charge of τ_{had} must be opposite to the total charge of the leptons The two same-sign leptons must be tight, with $p_T > 10$ GeV The opposite-sign lepton must be loose and isolated

Table 8.12: Selection criteria applied to define different channels. SFOS refers to same-flavour, opposite-sign lepton pairs [118].

8.5.4.2 3ℓ

The same-sign pair of leptons in the 3ℓ channel are required to be very tight. A `CFT` requirement is used for electrons despite negligible background from charge misidentified electrons. In fact `CFT` also helps reject a photon conversion background, where a photon converts to an electron pair in the detector (and only one electron is reconstructed). Such 3ℓ candidate events typically come from the $t\bar{t}\gamma$ process. This "fake" lepton background becomes more dominant in 3ℓ since the non-prompt contribution is decreased with the use of the non-prompt BDT.

Further event selections are applied to the 3ℓ channel to reject resonances from same-flavour opposite-sign lepton pairs. The invariant mass of these lepton pairs is required to be greater than 12 GeV to reject low mass resonances decaying to light leptons, and the invariant mass is required to be greater than 10 GeV away from the Z mass. In addition to these cuts, the invariant mass of all three leptons is also required to be greater than 10 GeV away from the Z mass. This is to reject $Z \to \ell\ell\gamma^* \to \ell\ell\ell'(\ell')$ decays, where an off-shell photon decays to two leptons in which one is soft and not reconstructed. The opposite charge lepton is also required to satisfy the L* selection criteria discussed above.

From this pre-selection region a five-dimensional BDT is trained to separate $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, VV and $t\bar{t}$, using the XGBoost tree boosting system [123]. The discriminants are mapped into five categories, defining the 3ℓ signal region and four control regions for each of the targeted decays. The control regions can be used in the fit to help understand whether background processes are well modelled by simulation in data. Events that are not assigned to a category (due to the left over phase space from the multidimensional binning procedure) are typically from non-prompt leptons and are hence assigned to the $t\bar{t}$ control region. In the signal region, the $t\bar{t}H$ BDT discriminant is used in the fit.

8.5.4.3 4ℓ

In the 4ℓ channel, the charges of the four leptons are required to sum to zero; resulting in two opposite-sign lepton pairs. The third and fourth leptons (with the lowest p_T) are required to pass the tight lepton selections. The invariant masses of the same-flavour opposite-sign pairs are required to be more than 10 GeV away from the Z mass, in the same way as 3ℓ. This suppresses the dominant $t\bar{t}Z$ background. The light resonance invariant mass veto described in 3ℓ is also applied to the same-flavour opposite-sign pairs. To retain orthogonality with a dedicated $t\bar{t}H(H \to ZZ \to 4\ell)$ analysis, the invariant mass of the four leptons is required to be greater than 5 GeV from the Higgs mass.

Due to lack of statistics, the 4ℓ channel uses a cut-and-count strategy to measure the signal strength. The region is further split into two categories: Z-depleted and Z-enriched, consisting of zero same-flavour opposite-sign pairs and otherwise (either two or four same-flavour opposite-sign pairs) respectively. The main background from $t\bar{t}Z$ (mainly from off shell Z decays) typically resides in the Z-enriched category and so a BDT is trained to discriminate $t\bar{t}H$ from this background. The Z-enriched region is then defined as a single bin by applying a requirement on the discriminant.

8.5.4.4 1ℓ+$2\tau_{\text{had}}$

The 1ℓ+$2\tau_{\text{had}}$ aims to reconstruct the $H \to \tau\tau$ decay. The two medium τ_{had} candidates must be opposite-sign to one another, with at least one of them also passing the tight identification requirements. The main background of fake τ_{had} candidates from $t\bar{t}$ events is reduced by training a BDT to separate $t\bar{t}H$ signal from this background. The BDT is then binned and the shape is fitted.

8.5.4.5 2ℓSS+$1\tau_{\text{had}}$

The event selections for 2ℓSS+$1\tau_{\text{had}}$ are similar to 2ℓSS, except the light leptons must have $p_T > 15$ GeV and there is no b-jet veto. In addition, a single medium τ_{had} candidate are required to be the opposite-sign to the same-sign pair. A BDT is trained to reject the dominant $t\bar{t}$ background and the shape of the BDT is used in the fit. Due to low statistics, the BDT is trained in a region with looser event selections.

8.5.4.6 2ℓOS+$1\tau_{\text{had}}$

Two opposite-sign light leptons and one medium τ_{had} candidate is required in the 2ℓOS+$1\tau_{\text{had}}$ channel. The main backgrounds are from dilepton decays of $t\bar{t}$ and Z+jets with a fake τ_{had}

candidate. To better reject these backgrounds at least three jets are required, and the invariant mass of the light lepton pair must be further than 10 GeV from the Z mass to specifically reject the Z+jets background. Lighter resonances are rejected in similar cuts to 3ℓ and 4ℓ above. A BDT trained to reject the main remaining $t\bar{t}$ background is applied and the shape fitted.

8.5.4.7 3ℓ+$1\tau_{\text{had}}$

The event selections for 3ℓ+$1\tau_{\text{had}}$ are similar to 3ℓ, with the addition of a medium τ_{had} candidate with opposite-sign to the total charge of the light leptons. The p_T requirement on the same-sign leptons is decreased to 10 GeV, and the leptons must pass the tight selection (as opposed to the very tight selection in 3ℓ). In addition, the opposite-sign lepton must be loose and isolated (not loose, isolated and passing non-prompt BDT requirement). All of these changes compared to 3ℓ are made to increase $t\bar{t}H$ acceptance due to low statistics in this channel. Nevertheless, the (still) low statistics result in this channel utilising a simple cut-and-count strategy with the signal region discussed here.

8.5.5 Background estimations

All channels discussed above have different background sources. These can typically be split up into the type of light lepton or hadronic tau causing such backgrounds. Backgrounds consisting of prompt leptons, non-prompt and fake light leptons, charge misassigned electrons and fake hadronic taus are discussed in Sections 8.5.5.1, 8.5.5.2, 8.5.5.3 and 8.5.5.4.

8.5.5.1 Prompt backgrounds

The prompt backgrounds are irreducible backgrounds since the final state of these backgrounds are the same as the signal. In this case all leptons are prompt. One cannot use lepton identification to distinguish between the two cases.

As discussed, the largest prompt backgrounds are from $t\bar{t}V$ and VV. Rarer prompt processes included in the analysis are tZ, tW, tWZ, $t\bar{t}WW$, VVV, $t\bar{t}t$ and $t\bar{t}t\bar{t}$. The tH process is also included as a prompt background, despite giving access to the top Yukawa coupling at tree level as with $t\bar{t}H$. These backgrounds are all estimated from simulation, at NLO where available.

8.5.5.2 Non-prompt and fake light lepton backgrounds

Non-prompt and fake lepton backgrounds are reducible in that one can determine differences between the leptons produced in these events compared to $t\bar{t}H$ events. A number of different channels suffer from non-prompt and fake light lepton backgrounds.

One does not expect simulation to describe such processes perfectly. Therefore data-driven estimates are used in conjunction with simulation to determine these backgrounds. Typically the yields from control regions enriched in such backgrounds are extrapolated to the signal region. The fractional contributions of fake and non-prompt light leptons from simulation in a number of different control regions are shown in Figure 8.12. These control regions are discussed below.

One can see that for muons, typically all contributions are from non-prompt decays from b- or c-quarks. For electrons there is also a non-negligible contribution from photon conversions.

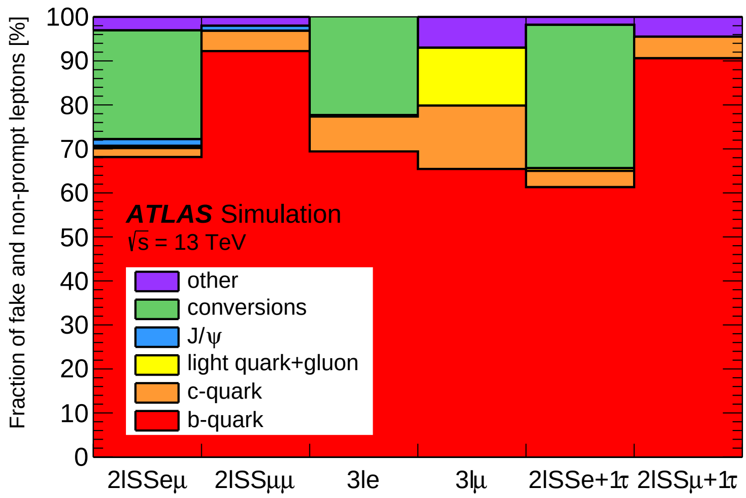


Figure 8.12: The fractional composition from simulation of non-prompt and fake leptons in the control regions used in the estimates of the non-prompt contributions. The 2ℓSS$e\mu$ and 2ℓSS$\mu\mu$ control regions are those used in the estimate for the 2ℓSS and 3ℓ channels. The control regions labelled $3\ell e$ and $3\ell\mu$ are used for the estimate in 4ℓ where the lowest p_T lepton is denoted. Those labelled 2ℓSSe+1τ_{had} and 2ℓSSμ+1τ_{had} are those used for the 2ℓSS+1τ_{had} channel. The non-prompt lepton background is separated into components from b-jets, c-jets, light and quark jets, J/ψ, photon conversions and other contributions. Other contributions include decays from pions, kaons and non-prompt taus and the cases where the reconstructed lepton cannot be assigned to a particular source [118].

In the 2ℓSS non-prompt estimate, a control region is defined by lowering the multiplicity of jets to $2 \leq N_{\text{jets}} \leq 3$ and by requiring that one lepton is loose. In this way, this region is dominated by non-prompt leptons produced from $t\bar{t}$ events. Almost exactly the same region is used for the 3ℓ non-prompt estimate, since the $t\bar{t}$ background producing the non-prompt leptons in the region is very similar except with another prompt lepton from the other top decay. The only difference is the p_T of the leptons is reduced to 15 GeV, to align with the 3ℓ signal region definition.

The *matrix method* [124] is used to determine the non-prompt estimates in these regions. Four orthogonal categories are defined, using leptons ordered in p_T: events with exactly two very tight leptons (TT), one very tight and one anti-very-tight lepton ($T\bar{T}$), one anti-very-tight and one very tight lepton ($\bar{T}T$), and two anti-very-tight leptons ($\bar{T}\bar{T}$). Anti-very-tight corresponds to leptons that pass the loose definition but not the very-tight. The total number of such events can be mapped into four regions characterised by different real and fake lepton compositions:

$$\begin{pmatrix} N^{TT} \\ N^{T\bar{T}} \\ N^{\bar{T}T} \\ N^{\bar{T}\bar{T}} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\bar{\varepsilon}_{r,2} & \varepsilon_{r,1}\bar{\varepsilon}_{f,2} & \varepsilon_{f,1}\bar{\varepsilon}_{r,2} & \varepsilon_{f,1}\bar{\varepsilon}_{f,2} \\ \bar{\varepsilon}_{r,1}\varepsilon_{r,2} & \bar{\varepsilon}_{r,1}\varepsilon_{f,2} & \bar{\varepsilon}_{f,1}\varepsilon_{r,2} & \bar{\varepsilon}_{f,1}\varepsilon_{f,2} \\ \bar{\varepsilon}_{r,1}\bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{r,1}\bar{\varepsilon}_{f,2} & \bar{\varepsilon}_{f,1}\bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{f,1}\bar{\varepsilon}_{f,2} \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}. \tag{8.12}$$

Here N^{rf} corresponds to the total number of events where the leading lepton is real and the subleading is fake, and so on. The $\bar{\varepsilon}$ correspond to the anti-very-tight efficiencies. The indexes for the ε_r and ε_f efficiencies are ranked in terms of lepton p_T and are measured separately for electrons and muons.

To obtain the number of fake leptons in the signal region, the matrix can be inverted and the total number of tight-tight events with at least one fake lepton, can be obtained as

$$N^{f}_{TT} = w_{TT}N^{TT} + w_{\bar{T}T}N^{\bar{T}T} + w_{T\bar{T}}N^{T\bar{T}} + w_{\bar{T}\bar{T}}N^{\bar{T}\bar{T}} \tag{8.13}$$

The w weights are functions of the measured prompt and non-prompt lepton efficiencies.

The tag and probe method is used to estimate the efficiencies. The real efficiencies, ϵ_r, are measured in an opposite-sign opposite-flavour control region dominated by prompt dilepton $t\bar{t}$ decays and the non-prompt/fake efficiencies, ϵ_f, in either same-sign $e\mu$ or same-sign $\mu\mu$ regions.

For the prompt efficiencies, the tag lepton is required to be very tight and matched to a single lepton trigger and the residual fake or non-prompt background is subtracted using the estimate from simulation. The prompt lepton efficiencies are parameterised as a function of p_T and are shown in Figure 8.13.

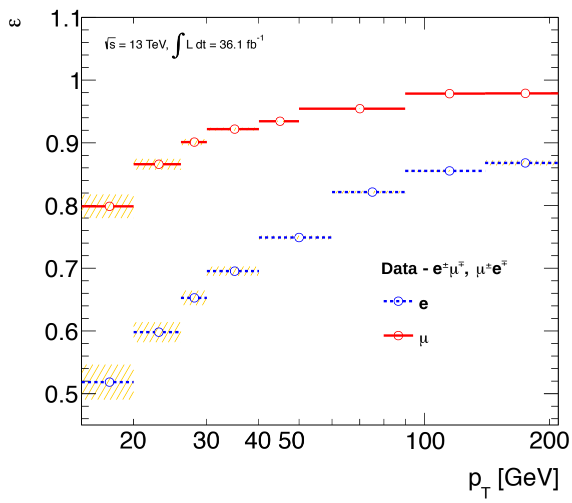
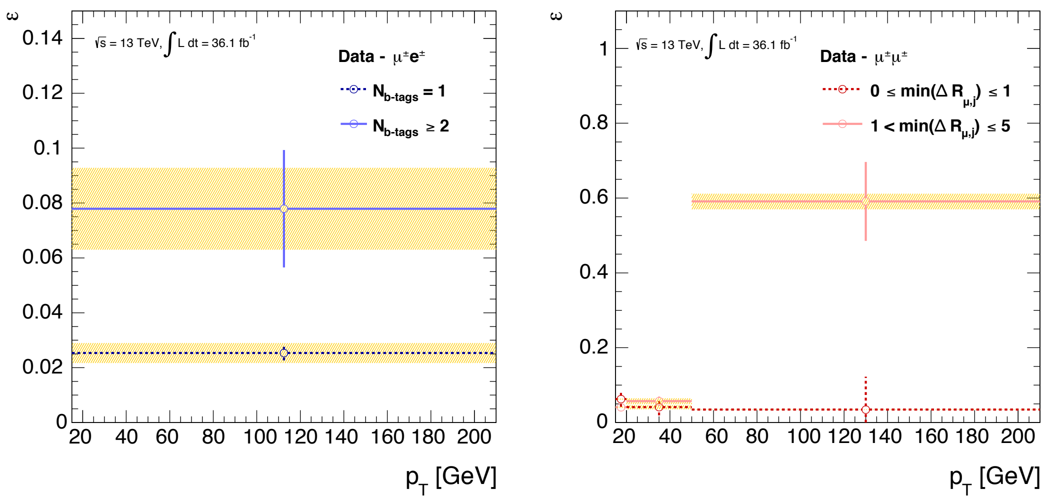


Figure 8.13: The real (prompt) electron and muon efficiencies as measured in data in an opposite-sign opposite-flavour control region. The statistical uncertainty (line) and systematic uncertainty (orange band) are shown. The systematic uncertainty includes the uncertainty due to the prompt background subtraction.

The non-prompt electron efficiency is parameterised in $N_{\text{b-jets}}$ in addition to p_T. This is due to the increase (fractionally) of photon conversions in events with $N_{\text{b-jets}} \geq 2$. The non-prompt muon efficiency is parameterised as a function of p_T and $\Delta R(\text{lepton, jet})$ due to large differences seen in the efficiencies in data. For both measurements the residual prompt background is subtracted using estimates from simulation and for electrons a data-driven charge misassignment contribution is also subtracted (see Section 8.5.5.3). They are shown in Figure 8.14.

The photon conversion fake efficiency is different (higher) than that of the non-prompt electron efficiency. Therefore when one extrapolates to the signal region using the (non-prompt dominated) fake rate the fraction of photon conversions is underestimated. To account for this a rescaling is performed, capturing the fractional change (α) of photon conversions in simulation going from the control region to the signal region:

$$f_{SR} = (1 + \alpha) f_{CR}. \tag{8.14}$$



(a) (b)

Figure 8.14: The fake/non-prompt (a) electron and (b) muon efficiencies as measured in same-sign opposite-flavour and same-sign muon control regions respectively. The statistical uncertainty (line) and systematic uncertainty (orange band) are shown. The systematic uncertainty includes the uncertainty due to the prompt background subtraction and additionally charge misassignment subtraction for electrons.

A 40% systematic uncertainty on this rescaling is estimated, with contributions from conversion and non-prompt modelling uncertainties in simulation.

The non-prompt yields are shown in Table 8.13. The total ratio of data-driven to simulation yields in 2ℓSS is roughly 1.7, which is comparable to the factor 1.5 used when optimising tight lepton definitions. The ratio increases in 3ℓ to 2.3 (inclusively). This would result in a greater increase in expected discovery significance using the non-prompt BDT in the lepton definition with respect to only using isolation than reported above.

Prediction	$\mu\mu$	$e\mu$	ee	2ℓSS
MC	32 ± 3	65 ± 6	38 ± 4	134 ± 13
MM	47 ± 14	110 ± 22	76 ± 17	233 ± 40
MM/MC	1.5 ± 0.5	1.7 ± 0.4	2.0 ± 0.5	1.7 ± 0.3

Prediction	$\ell\mu\mu$	$\ell e\mu$	ℓee	3ℓ
MC	16.4 ± 4.1	24.9 ± 5.0	11.1 ± 3.3	52.4 ± 13.6
MM	35.3 ± 5.9	62.0 ± 7.9	24.0 ± 4.9	121.3 ± 18.9
MM/MC	2.2 ± 0.7	2.5 ± 0.6	2.2 ± 0.8	2.3 ± 0.7

Table 8.13: Non-prompt estimate yields from the matrix method (MM) and from simulation, with total uncertainty, in the 2ℓSS (top) and 3ℓ (bottom) pre-selection regions.

Non-prompt (and fake) light lepton estimates are also needed in the 4ℓ and 2ℓSS+$1\tau_{\text{had}}$ regions. In the 4ℓ region a semi data-driven approach is used to determine scale factors to apply to simulation due to contributions from leptons originating from heavy flavour (containing b- or c-quarks) or light flavour jet sources. For electrons individual scale factors for heavy jet, $\lambda^{e}_{\text{heavy}}$, and light jet, $\lambda^{e}_{\text{light}}$, sources are used and for muons an overall scale factor, λ^{μ}, is used (due to limited light jet contributions). These scale factors are derived in a simultaneous fit to four control region with three loose leptons, split up by flavour: eee, $ee\mu$, $e\mu\mu$ and $\mu\mu\mu$. They are then applied to any events containing non-prompt leptons in the 4ℓ region to correct simulation with data. The derived scale factors are shown in Table 8.14.

Scale factor	Value	Statistical uncertainty	Systematic uncertainty	Total uncertainty
$\lambda^{e}_{\text{heavy}}$	1.48	0.22	0.44	0.50
$\lambda^{e}_{\text{light}}$	0.72	0.53	0.22	0.57
λ^{μ}	0.66	0.19	0.20	0.27

Table 8.14: Fake scale factors applied to events with non-prompt leptons in simulation. A 30% systematic uncertainty is determined for each scale factor by measuring the scale factor dependence as a function of lepton p_T.

In the 2ℓSS+1τ_{had} region the fake factor method is used to determine the number of non-prompt leptons. This method is similar to the matrix method, with the main difference being the fake factor method estimates the prompt contribution from simulation, as opposed to the matrix method which used data. A similar control region as that used in 2ℓSS is used to measure these factors with the addition of a medium hadronic tau candidate in the event. The treatment of conversions is also the same as in 2ℓSS.

8.5.5.3 Electron charge misidentification

The probability of an electron charge misidentification background in the 2ℓSS and 2ℓSS+1τ_{had} regions is determined from data via a likelihood-based method. The probability is parameterised in p_T and $|\eta|$. One expects charge misidentification rates to increase with $|\eta|$, since the amount of material the electron passes through has a large impact on the production of trident electrons, and to also increase with p_T due to the tracks of electrons becoming straighter at higher energies.

The likelihood method uses same-sign and opposite-sign $Z \to ee$ events in data to compute charge flip rates. Sidebands in high and low dielectron invariant mass are used to subtract backgrounds from the events used to measure the rates. The likelihood method is also validated by comparing to a truth matched method performed in $Z \to ee$ simulation. The effects of both background subtraction and likelihood validation on the charge flip rates are taken as systematic uncertainties. The rates measured in data are shown in Figure 8.15 for both electron types used in the non-prompt estimates discussed above. The decrease in the charge flip rates for the very tight electron selections is clear.

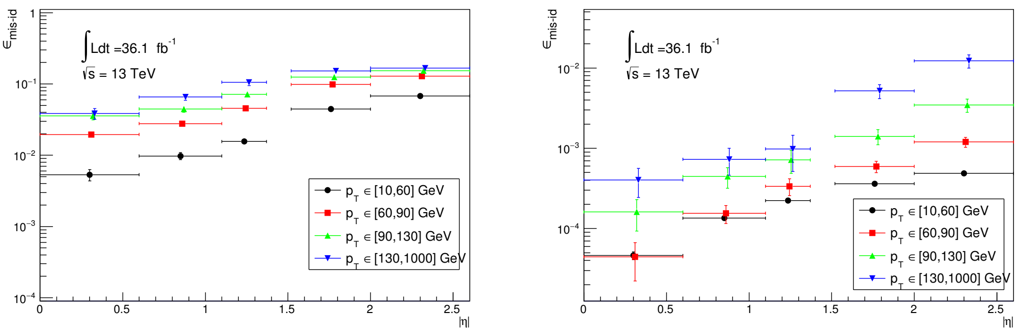
The charge misidentification event yields in the 2ℓSS and 2ℓSS+1τ_{had} regions are determined by weighting events with the same definition but with opposite-sign leptons by the measured charge-flip rate. The weighting is applied to each electron in the event. For a same-sign electron region, the weighting applied to opposite-sign electron events is as follows,

$$w = \epsilon_1(1-\epsilon_2) + \epsilon_2(1-\epsilon_1), \tag{8.15}$$

where ϵ_1 is the charge flip rate for the leading electron and ϵ_2 for the subleading. The full yields, with comparison to expected simulation yields, is shown in Table 8.15. The simulation slightly under-predicts charge flip rates compared to data.

8.5.5.4 Fake τ_{had} background

The channels with hadronic taus all suffer from backgrounds with fake τ_{had} candidates. These are usually from light-quark jets. In all cases (except 2ℓSS+1τ_{had}) the numbers of non-prompt



(a) (b)

Figure 8.15: Electron charge misidentification probabilities determined from data with the likelihood method for (a) anti-very-tight electrons (loose-but-not-very-tight) and (b) very tight electrons. The total uncertainty is shown with systematic elements from the background subtraction definition and validation of the likelihood method in simulation.

Predictions	2ℓSS (ee)	2ℓSS ($e\mu$)	2ℓSS	2ℓSS+1τ_{had}
Data-driven (DD)	18.5 $\pm$ 6.5	14.1 $\pm$ 5.0	32.6 $\pm$ 11.4	0.05 $\pm$ 0.01
MC	15.8	12.9	28.7	-
DD/MC	1.2	1.1	1.1	-

Table 8.15: Data-driven charge flip electron event yields with total uncertainty in the 2ℓSS and 2ℓSS+1τ_{had} pre-selection regions. The ratio of the data-driven to MC prediction is given. No charge flip events are present in MC for the 2ℓSS+1τ_{had} region.

light leptons are negligible and estimated from simulation.

The fake factor method is used to determine a data-driven measurement of the fake τ_{had} contribution to the 2ℓOS+1τ_{had} channel. The factors are determined in an orthogonal control region similar to the nominal 2ℓOS+1τ_{had} region, except with a veto on b-jets in the event. Another orthogonal region is defined with the exact same event selections with an "anti-medium" (loose-but-not-medium) τ_{had} candidate, which is further enriched in fake τ_{had} candidates. A ratio of the yields in the two regions define the fake factors, which can then be extrapolated to the pre-selection region with at least one b-jet in the event. The fake factors are parameterised in τ_{had} p_T, and depend negligibly on other event properties. Systematic uncertainties on the scale factors take into account the difference in the fake composition (taken from simulation) and scale factors between the different regions used in the estimate.

The fractional composition of fake τ_{had} candidates in all channels including hadronic taus are shown in Figure 8.16. The composition and origin of fake τ_{had} candidates in the 3ℓ+1τ_{had} and 2ℓSS+1τ_{had} regions are very similar to that of 2ℓOS+1τ_{had}. In both the 2ℓSS+1τ_{had} and 3ℓ+1τ_{had} regions, inverting the hadronic tau identification to determine scale factors actually results in an overlap between the 2ℓSS and 3ℓ signal regions respectively. Therefore scale factors, utilising the 2ℓOS+1τ_{had} fake factors, are applied to both regions. These scale factors are found to be independent of p_T and therefore the fake factors from 2ℓOS+1τ_{had} are scaled with a flat value. Systematic uncertainties on the scale factors are produced by comparing values in control regions enriched in $t\bar{t}$ and Z+jets, compared to the nominal region. The determined scale factor is 1.36 $\pm$ 0.16. It is only applied in the 2ℓSS+1τ_{had} region where both leptons are prompt, and the hadronic tau is fake (in simulation). These are typically from $t\bar{t}V$ events, as are the fake hadronic tau events in 3ℓ+1τ_{had}.

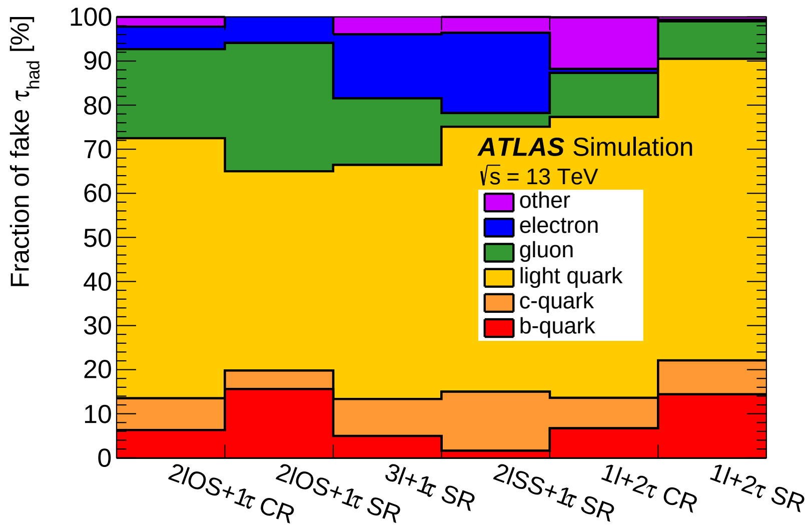


Figure 8.16: The fractional fake τ_{had} composition, shown in the control regions used to make data-driven estimates and in the signal regions of each channel. The fake τ_{had} background has been separated into components from b-jets, c-jets, light quark jets, gluon jets, electrons and other contributions. Other contributions include muons wrongly identified as hadronic taus and cases where the reconstructed hadronic tau candidate cannot be assigned to a particular source [118].

The fake hadronic tau contribution in the 1ℓ+$2\tau_{had}$ region is slightly different than the other hadronic tau channels. The dominant background is from $t\bar{t}$ production, where either one or two of the τ_{had} candidates are fake. Same-sign τ_{had} regions are used to measure a single inclusive fake factor; one with nominal tau selection and one where a single tau candidate is anti-medium. They are then extrapolated to the signal region. Systematic uncertainties are deduced from a closure test comparing the number of expected opposite-sign tau events from corrected simulation compared with data. A 30% unfolding uncertainty is applied due to this non-closure.

8.5.5.5 Summary

The total fraction of backgrounds, split up by the type of particle process for the prompt events and fake/non-prompt for the reducible processes, is shown in Figure 8.17. The relatively large non-prompt backgrounds (proportionally) in the 2ℓSS region and the 3ℓ signal region are further rejected with multivariate techniques in the final fit.

8.5.6 Systematic uncertainties

Systematic uncertainties used in the analysis are summarised in Table 8.16. Each systematic uncertainty can be defined one of three ways. Firstly, the uncertainty can be treated as a normalisation. For example, a theoretical uncertainty on a cross-section results in either an increase or decrease in the number of events from a process, which is the same as applying different normalisations. Another type of systematic uncertainty can be treated as a "shape" uncertainty; an uncertainty that does not change the total number of events but, for example, affects how the shape of a discriminant is distributed. The final type of systematic is one that combines both types above.

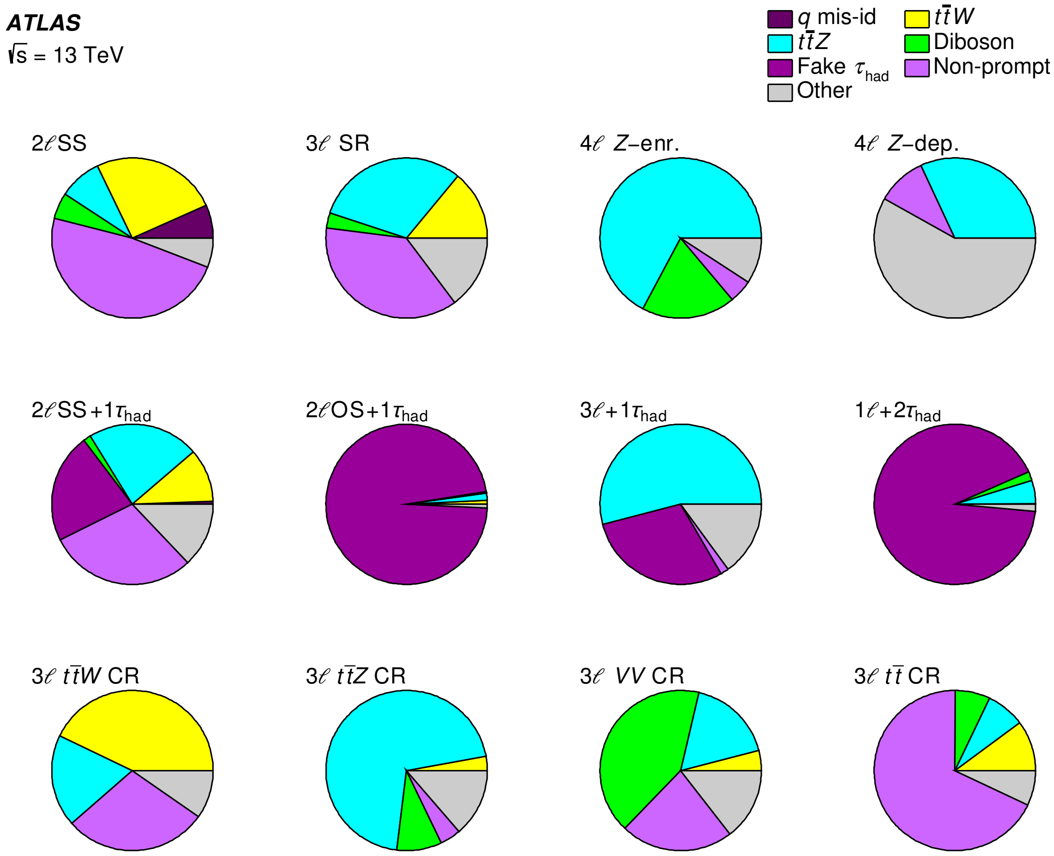


Figure 8.17: Fractional contributions of various background processes and lepton/τ_{had} types to the total predicted background in each of the channel categories. Where appropriate, the reducible processes are data-driven. Other refers to the rare processes [118].

The systematic uncertainties in the analysis can be categorised into either *experimental*, *data-driven* or *theoretical* based uncertainties. Experimental uncertainties include those related to the reconstruction of physics objects. That includes the reconstruction, identification and calibration of light leptons and hadronic taus and the energy scale and flavour tagging of jets. Other experimental uncertainties include the total uncertainty on the integrated luminosity of 2015+2016 data (2.1%) and on pileup reweighting.

The uncertainties on the estimation of the reducible backgrounds in the analysis are discussed in Section 8.5.5. For the non-prompt light lepton estimates, these typically consist of the statistical uncertainty on the number of events used in control regions, the uncertainty on measured prompt and non-prompt efficiencies, and uncertainties due to the non-closure of the method and photon conversion fractions. Uncertainties are applied to the charge misassignment estimate and due to the fake hadronic tau estimates and extrapolations.

A number of theoretical systematic uncertainties are applied. Systematic uncertainties due to theoretical values of cross sections are determined by varying the cross sections within their uncertainties in the fit. The uncertainty on the acceptance of simulated events in different channels due to uncertainties in the modelling of the parton shower and hadronisation models are determined using alternative simulation samples employing different models to that of nominal. Other theoretical uncertainties include the effect of QCD factorisation and renormalisation scales and uncertainties on PDFs and the strong coupling α_S.

Systematic uncertainty	Type	Components
Luminosity	N	1
Pileup reweighting	SN	1
Physics Objects		
Electron	SN	6
Muon	SN	15
Hadronic tau	SN	10
Jet energy scale and resolution	SN	28
Jet vertex tagger	SN	1
Jet flavour tagging	SN	126
Missing E_T	SN	3
Total (Experimental)	–	191
Data-driven non-prompt/fake leptons and charge misassignment		
Control region statistics	SN	38
Light lepton efficiencies	SN	22
Non-prompt estimates: non-closure	N	5
γ-conversion fraction, α	N	5
Fake τ_{had} estimates	N/SN	12
Electron charge misassignment	SN	1
Total (Data-driven)	–	83
$t\bar{t}H$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Parton shower and hadronisation model	SN	1
Higgs boson branching fraction	N	4
Shower tune	SN	1
$t\bar{t}W$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
$t\bar{t}Z$ modelling		
Cross section	N	2
Renormalisation and factorisation scales	S	3
Matrix-element MC event generator	SN	1
Shower tune	SN	1
Other background modelling		
Cross section	N	15
Shower tune	SN	1
Total (Signal and background modelling)	–	41
Total (Overall)	–	315

Table 8.16: Types of systematic uncertainty considered in the analysis. "N" refers to uncertainties treated as a normalisation, "S" denotes uncertainties that consider shape and "SN" refers to uncertainties that are applied both as shape and normalisation. Missing E_T is included as a systematic uncertainty due to being used in defining a control region to measure non-prompt contributions in the 4ℓ channel. This uncertainty is pruned from the final fit [118].

8.5.7 Results

A binned maximum-likelihood fit is performed on all twelve categories simultaneously to extract the signal strength. As discussed, a signal strength of $\mu = 1$ indicates the presence of top quark associated Higgs boson production, and $\mu = 0$ indicates the absence of such a production. The Higgs boson branching fractions are fixed to the SM value in the fit. In the 2ℓSS, 3ℓ, 2ℓSS+1τ_{had}, 2ℓOS+1τ_{had} and 1ℓ+2τ_{had} signal regions, the shape of a BDT discriminant is used. The BDT is binned in such a way as to maximise the sensitivity of the analysis in simulation. Only a single bin is used in the two 4ℓ signal regions and the 3ℓ+1τ_{had} signal region. In all, 32 bins are fitted simultaneously to extract μ.

The systematic uncertainties in the analysis are accounted for by applying nuisance parameters to the likelihood, as discussed in Section 8.2. Nominally, 315 nuisance parameters are included. This includes a large proportion from lepton and jet related experimental uncertainties. Many of these nuisance parameters act negligibly on the fitted value of μ; thus, a pruning procedure is applied where nuisance parameters that modify the fitted value of $\mu = 1$ less than 1% are removed, reducing the number of parameters to 230.

The expected sensitivity on the measured value of μ can be determined from an *Asimov* fit [125]. The fitting procedure discussed in Section 8.2.1 is applied to simulation which corresponds, by definition, to the SM hypothesis ($\mu = 1$). An expected discovery significance can then be quoted from the total uncertainties in the fit. The expected best fit value of μ is

$$\mu_{\mathrm{exp}} = 1.0 \pm 0.3(\mathrm{stat.}) \pm 0.3(\mathrm{syst.}) = 1.0 \pm 0.4$$

where "stat." stands for statistical uncertainty and "syst." for systematic uncertainty. This corresponds to an expected discovery significance of 2.8σ.

The best-fit value of μ in data is

$$\mu = 1.6 \pm 0.3(\mathrm{stat.})^{+0.4}_{-0.3}(\mathrm{syst.}) = 1.6^{+0.5}_{-0.4}$$

corresponding to a 4.1σ excess in the expected number of $t\bar{t}H$ events with respect to the null hypothesis, and a 1.4σ excess with respect to the SM hypothesis. The yields in all channels, both pre-fit and post-fit, are shown in Table 8.17.

The impact of a systematic uncertainty on μ is determined by calculating the difference in μ, $\Delta\mu$, from fits where the corresponding nuisance parameter is set to the fitted value plus upward uncertainty and fitted value minus downward uncertainty. All other nuisance parameters are allowed to float in this process.

The systematic uncertainties with largest impact on μ are shown in Table 8.18. The theoretical modelling of $t\bar{t}H$ and the uncertainty on jet energy scale are the largest. The uncertainty on the modelling of $t\bar{t}H$ enters into all channels and directly modifies the fitted value of μ when shifted from nominal. Many channels require a large number of reconstructed jets in the final state, so uncertainties on the energy scale of individual jets are combined. The uncertainty in the non-prompt estimates in the sensitive 2ℓSS and 3ℓ channels also factor highly.

Category	Non-prompt	Fake τ_{had}	q mis-id	$t\bar{t}W$	$t\bar{t}Z$	Diboson	Other	Total Bkgd.	$t\bar{t}H$	Observed
					Pre-fit yields					
2ℓSS	233 ± 39	–	33 ± 11	123 ± 18	41.4 ± 5.6	25 ± 15	28.4 ± 5.9	484 ± 38	42.6 ± 4.2	514
3ℓ SR	14.5 ± 4.3	–	–	5.5 ± 1.2	12.0 ± 1.8	1.2 ± 1.2	5.8 ± 1.4	39.1 ± 5.2	11.2 ± 1.6	61
3ℓ $t\bar{t}W$ CR	13.3 ± 4.3	–	–	19.9 ± 3.1	8.7 ± 1.1	< 0.2	4.53 ± 0.92	46.5 ± 5.4	4.18 ± 0.46	56
3ℓ $t\bar{t}Z$ CR	3.9 ± 2.5	–	–	2.71 ± 0.56	66 ± 11	8.4 ± 5.3	12.9 ± 4.2	93 ± 13	3.17 ± 0.41	107
3ℓ VV CR	27.7 ± 8.7	–	–	4.9 ± 1.0	21.3 ± 3.4	51 ± 30	17.9 ± 6.1	123 ± 32	1.67 ± 0.25	109
3ℓ $t\bar{t}$ CR	70 ± 17	–	–	10.5 ± 1.5	7.9 ± 1.1	7.2 ± 4.8	7.3 ± 1.9	103 ± 17	4.00 ± 0.49	85
4ℓ Z-enr.	0.11 ± 0.07	–	–	< 0.01	1.52 ± 0.23	0.43 ± 0.23	0.21 ± 0.09	2.26 ± 0.34	1.06 ± 0.14	2
4ℓ Z-dep.	0.01 ± 0.01	–	–	< 0.01	0.04 ± 0.02	< 0.01	0.06 ± 0.03	0.11 ± 0.03	0.20 ± 0.03	0
1ℓ+$2\tau_{\text{had}}$	–	65 ± 21	–	0.09 ± 0.09	3.3 ± 1.0	1.3 ± 1.0	0.98 ± 0.35	71 ± 21	4.3 ± 1.0	67
2ℓSS+$1\tau_{\text{had}}$	2.4 ± 1.4	1.80 ± 0.30	0.05 ± 0.02	0.88 ± 0.24	1.83 ± 0.37	0.12 ± 0.18	1.06 ± 0.24	8.2 ± 1.6	3.09 ± 0.46	18
2ℓOS+$1\tau_{\text{had}}$	–	756 ± 80	–	6.5 ± 1.3	11.4 ± 1.9	2.0 ± 1.3	5.8 ± 1.5	782 ± 81	14.2 ± 2.0	807
3ℓ+$1\tau_{\text{had}}$	–	0.75 ± 0.15	–	0.04 ± 0.04	1.38 ± 0.24	0.002 ± 0.002	0.38 ± 0.10	2.55 ± 0.32	1.51 ± 0.23	5
					Post-fit yields					
2ℓSS	211 ± 26	–	28.3 ± 9.4	127 ± 18	42.9 ± 5.4	20.0 ± 6.3	28.5 ± 5.7	459 ± 24	67 ± 18	514
3ℓ SR	13.2 ± 3.1	–	–	5.8 ± 1.2	12.9 ± 1.6	1.2 ± 1.1	5.9 ± 1.3	39.0 ± 4.0	17.7 ± 4.9	61
3ℓ $t\bar{t}W$ CR	11.7 ± 3.0	–	–	20.4 ± 3.0	8.9 ± 1.0	< 0.2	4.54 ± 0.88	45.6 ± 4.0	6.6 ± 1.9	56
3ℓ $t\bar{t}Z$ CR	3.5 ± 2.1	–	–	2.82 ± 0.56	70.4 ± 8.6	7.1 ± 3.0	13.6 ± 4.2	97.4 ± 8.6	5.1 ± 1.4	107
3ℓ VV CR	22.4 ± 5.7	–	–	5.05 ± 0.94	22.0 ± 3.0	39 ± 11	18.1 ± 5.9	106.8 ± 9.4	2.61 ± 0.82	109
3ℓ $t\bar{t}$ CR	56.0 ± 8.1	–	–	10.7 ± 1.4	8.1 ± 1.0	5.9 ± 2.7	7.1 ± 1.8	87.8 ± 7.9	6.3 ± 1.8	85
4ℓ Z-enr.	0.10 ± 0.07	–	–	< 0.01	1.60 ± 0.22	0.37 ± 0.15	0.22 ± 0.10	2.29 ± 0.28	1.65 ± 0.47	2
4ℓ Z-dep.	0.01 ± 0.01	–	–	< 0.01	0.04 ± 0.02	< 0.01	0.07 ± 0.03	0.11 ± 0.03	0.32 ± 0.09	0
1ℓ+$2\tau_{\text{had}}$	–	58.0 ± 6.8	–	0.11 ± 0.11	3.31 ± 0.90	0.98 ± 0.75	0.98 ± 0.33	63.4 ± 6.7	6.5 ± 2.0	67
2ℓSS+$1\tau_{\text{had}}$	1.86 ± 0.91	1.86 ± 0.27	0.05 ± 0.02	0.97 ± 0.26	1.96 ± 0.37	0.15 ± 0.20	1.09 ± 0.24	7.9 ± 1.2	5.1 ± 1.3	18
2ℓOS+$1\tau_{\text{had}}$	–	756 ± 28	–	6.6 ± 1.3	11.5 ± 1.7	1.64 ± 0.92	6.1 ± 1.5	782 ± 27	21.7 ± 5.9	807
3ℓ+$1\tau_{\text{had}}$	–	0.75 ± 0.14	–	0.04 ± 0.04	1.42 ± 0.22	0.002 ± 0.002	0.40 ± 0.10	2.61 ± 0.30	2.41 ± 0.68	5

Table 8.17: The observed yields in all channels with 36.1 fb^{-1} of data at $\sqrt{s} = 13$ TeV. The total uncertainty in the background estimates are shown. The non-prompt, fake τ_{had} and "q mis-id" (charge misidentification) estimates are data-driven. Rare processes include tZ, tW, tWZ, $t\bar{t}WW$, VVV, $t\bar{t}t$, $t\bar{t}t\bar{t}$, tH and rare top decays. The pre-fit values are quoted at the top which use the initial values of background systematic uncertainty nuisance parameters and the signal expected from the SM. The corresponding post-fit values are quoted at the bottom. The prediction and uncertainties post-fit for $t\bar{t}H$ reflect the best-fit signal strength of $\mu = 1.6\,^{+0.5}_{-0.4}$ and the uncertainty in the total background estimate is smaller than for the pre-fit values due to anticorrelations between the nuisance parameters obtained in the fit [118].

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modelling (cross section)	$+0.20$	-0.09
Jet energy scale and resolution	$+0.18$	-0.15
Non-prompt light-lepton estimates	$+0.15$	-0.13
Jet flavour tagging and τ_{had} identification	$+0.11$	-0.09
$t\bar{t}W$ modelling	$+0.10$	-0.09
$t\bar{t}Z$ modelling	$+0.08$	-0.07
Other background modelling	$+0.08$	-0.07
Luminosity	$+0.08$	-0.06
$t\bar{t}H$ modelling (acceptance)	$+0.08$	-0.04
Fake τ_{had} estimates	$+0.07$	-0.07
Other experimental uncertainties	$+0.05$	-0.04
Simulation sample size	$+0.04$	-0.04
Charge misassignment	$+0.01$	-0.01
Total systematic uncertainty	$+0.39$	-0.30

Table 8.18: A summary of the systematic uncertainties with largest impact on the fitted value of μ [118].

The best-fit value of μ for each channel can also be extracted by defining individual fitted parameters of interest. The results of the individual fits are shown in Table 8.19. All channels see an excess except for 1ℓ+$2\tau_{\mathrm{had}}$ and 4ℓ. In these cases the value of μ is negative. This corresponds to the case where there is less data than in the background-only hypothesis. No discovery significances are quoted for these cases.

The 4ℓ and 3ℓ+$1\tau_{\mathrm{had}}$ analyses are statistically limited. The other channels have roughly similar statistical and systematic uncertainties on the fitted signal strength value. The modelling of the discriminating variables used in the fit in the eight signal regions with the fitted μ applied are shown in Figures 8.18 and 8.19.

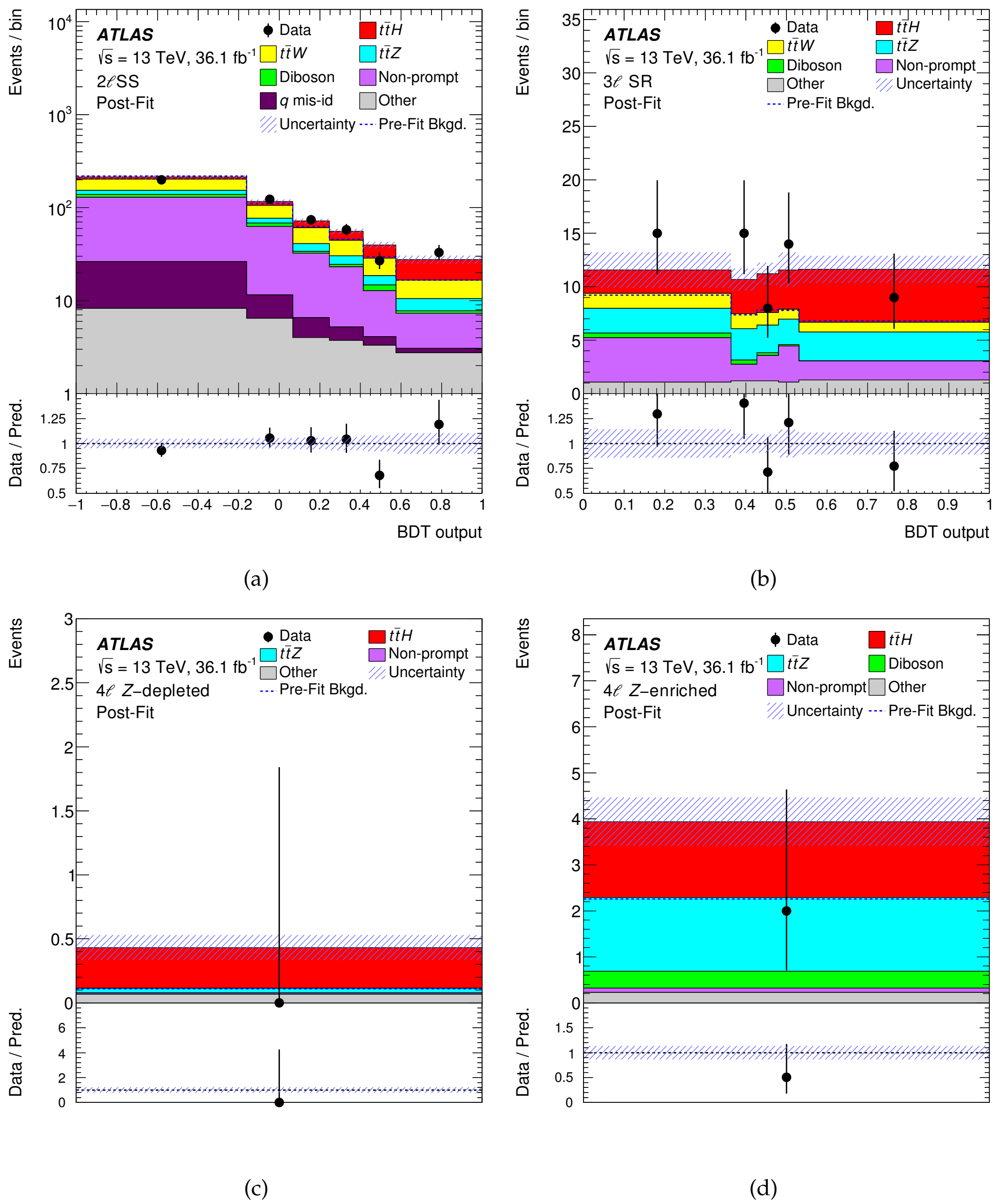


Figure 8.18: The distribution of the discriminants in data and simulation in the (a) 2ℓSS, (b) 3ℓ, (c) 4ℓ Z-depleted and (d) 4ℓ Z-enriched signal regions. The background contributions after the global fit are shown as filled histograms and the total background contribution before the fit is shown as a dashed blue histogram. The Higgs boson signal (red filled histogram) is scaled according to the results of the fit. The size of the combined statistical and systematic uncertainty in the sum of the signal and fitted background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and fitted background is also shown [118].

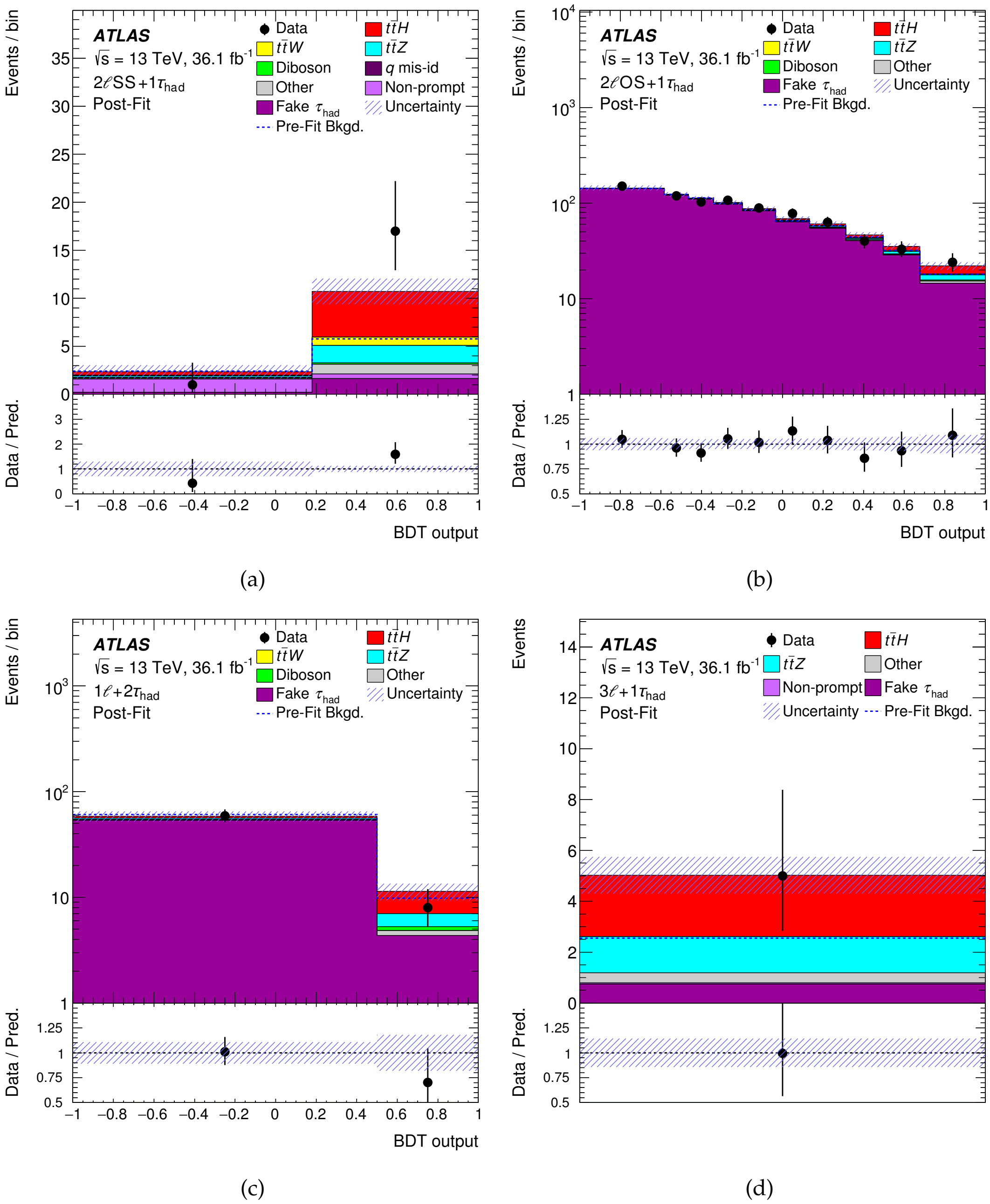


Figure 8.19: The distribution of the discriminants in data and simulation in the (a) 2ℓSS+1τ_{had}, (b) 2ℓOS+1τ_{had}, (c) 1ℓ+2τ_{had} and (d) 3ℓ+1τ_{had} signal regions. The background contributions after the global fit are shown as filled histograms and the total background contribution before the fit is shown as a dashed blue histogram. The Higgs boson signal (red filled histogram) is scaled according to the results of the fit. The size of the combined statistical and systematic uncertainty in the sum of the signal and fitted background is also shown. The ratio of the data to the sum of the signal and fitted background is also shown [118].

Channel	Best-fit μ		Significance	
	Observed	Expected	Observed	Expected
2ℓOS+$1\tau_{\text{had}}$	$1.7\,^{+1.6}_{-1.5}$ (stat.) $^{+1.4}_{-1.1}$ (syst.)	$1.0\,^{+1.5}_{-1.4}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	0.9σ	0.5σ
1ℓ+$2\tau_{\text{had}}$	$-0.6\,^{+1.1}_{-0.8}$ (stat.) $^{+1.1}_{-1.3}$ (syst.)	$1.0\,^{+1.1}_{-0.9}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	–	0.6σ
4ℓ	$-0.5\,^{+1.3}_{-0.8}$ (stat.) $^{+0.2}_{-0.3}$ (syst.)	$1.0\,^{+1.7}_{-1.2}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	–	0.8σ
3ℓ+$1\tau_{\text{had}}$	$1.6\,^{+1.7}_{-1.3}$ (stat.) $^{+0.6}_{-0.2}$ (syst.)	$1.0\,^{+1.5}_{-1.1}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	1.3σ	0.9σ
2ℓSS+$1\tau_{\text{had}}$	$3.5\,^{+1.5}_{-1.2}$ (stat.) $^{+0.9}_{-0.5}$ (syst.)	$1.0\,^{+1.1}_{-0.8}$ (stat.) $^{+0.5}_{-0.3}$ (syst.)	3.4σ	1.1σ
3ℓ	$1.8\,^{+0.6}_{-0.6}$ (stat.) $^{+0.6}_{-0.5}$ (syst.)	$1.0\,^{+0.6}_{-0.5}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	2.4σ	1.5σ
2ℓSS	$1.5\,^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	$1.0\,^{+0.4}_{-0.4}$ (stat.) $^{+0.4}_{-0.4}$ (syst.)	2.7σ	1.9σ
Combined	$1.6\,^{+0.3}_{-0.3}$ (stat.) $^{+0.4}_{-0.3}$ (syst.)	$1.0\,^{+0.3}_{-0.3}$ (stat.) $^{+0.3}_{-0.3}$ (syst.)	4.1σ	2.8σ

Table 8.19: Observed and expected best-fit values of the signal strength and discovery significance under the background-only hypothesis in all channels, arranged by expected sensitivity. No observed significance is given for channels with negative values of μ [118].

Chapter 9

Evidence for $t\bar{t}H$ production

In Chapter 8, the search for $t\bar{t}H$ production in multilepton final states was presented with full 2015 and 2016 data. As discussed, the Higgs boson can decay into other particles and these other decays can also be used to measure $t\bar{t}H$ production. In this chapter, brief overviews of the $t\bar{t}H(H \to bb)$ (Section 9.1), $t\bar{t}H(H \to \gamma\gamma)$ (Section 9.2) and the $t\bar{t}H(H \to ZZ(4\ell))$ (Section 9.3) analyses are given. In Section 9.4 the results of the combined analysis of all the Higgs decays in the search for $t\bar{t}H$ production is presented. Finally the results of a newer combined $t\bar{t}H$ analysis utilising 2017 data is shown in Section 9.5.

9.1 $t\bar{t}H(H \to bb)$

The search for $t\bar{t}H(H \to bb)$ [109] is similar to that of the multilepton search. In the $t\bar{t}H(H \to bb)$ case there are nominally four b quarks and two W bosons in the final state. The search is made difficult by the overwhelming background from $t\bar{t}$, especially from a top quark pair produced in association with a pair of b quarks. One can see that this background is irreducible and hence multivariate techniques are employed to distinguish it from $t\bar{t}H(H \to bb)$.

The leptonic decays of the top quark pair in $t\bar{t}H$ are targeted; either a semileptonic or dileptonic decay. These final states nominally have six ($bbbbjj$) and four ($bbbb$) jets in the final state respectively. Regions specified by the jet and b-jet multiplicity are used to enrich regions with $t\bar{t}H$, $t\bar{t}+\geq 2b$, $t\bar{t}+1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}+$ light jets. Pseudo-continuous b-tagging is used.

So far in this thesis, a b-jet working point defined from a one dimensional cut to the MV2c10 BDT distribution is used. Such working points that are calibrated include 60%, 70%, 77% and 85% b-jet efficiencies, each with differing light and c-jet rejections. Each working point thus has two possibilities; either a jet passes or fails the cut. The approach employed by pseudo-continuous b-tagging is to use the entire distribution of MV2c10 weights to classify jets in an event, divided into five bins defined by the supported working points. This allows for a finer segmentation and classification of both light and b-jets.

For the semileptonic channel, at least five jets are required. If there are exactly five jets then at least three of the jets are required to pass the 77% b-jet efficiency MV2c10 working point. If there are six or more jets then either two jets are required at the 60% working point or three jets at the looser 77% working point. In the dilepton channel, at least three jets are required with two passing the 77% MV2c10 working point. Both leptons are required to be opposite sign and have invariant mass at least 10 GeV away from m_Z.

Regions enriched in $t\bar{t}H$, $t\bar{t}+\geq 2b$, $t\bar{t}+1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}+$ light jets are then specified; eleven and seven such regions are defined for the semileptonic and dileptonic channels respectively. BDT discriminants are used in the $t\bar{t}H$ enriched regions to discriminate between signal and background. An additional boosted region is defined, targeted to measure boosted Higgs decays. This region searches for a semileptonic decay of $t\bar{t}H$ and therefore such boosted events are thus removed from the resolved semileptonic regions, to remain orthogonal to one another.

Due to disagreement between simulation and data, the normalisations of $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ are allowed to float in the unbinned likelihood fit when determining $\mu(t\bar{t}H)$. These normalisations are determined from the control regions discussed above. The three parameter fit results in a decrease in the sensitivity of the measured $\mu(t\bar{t}H)$ due to uncertainty between the three normalisations.

The yields of the signal and control regions pre- and post-fit are shown in Figure 9.1.

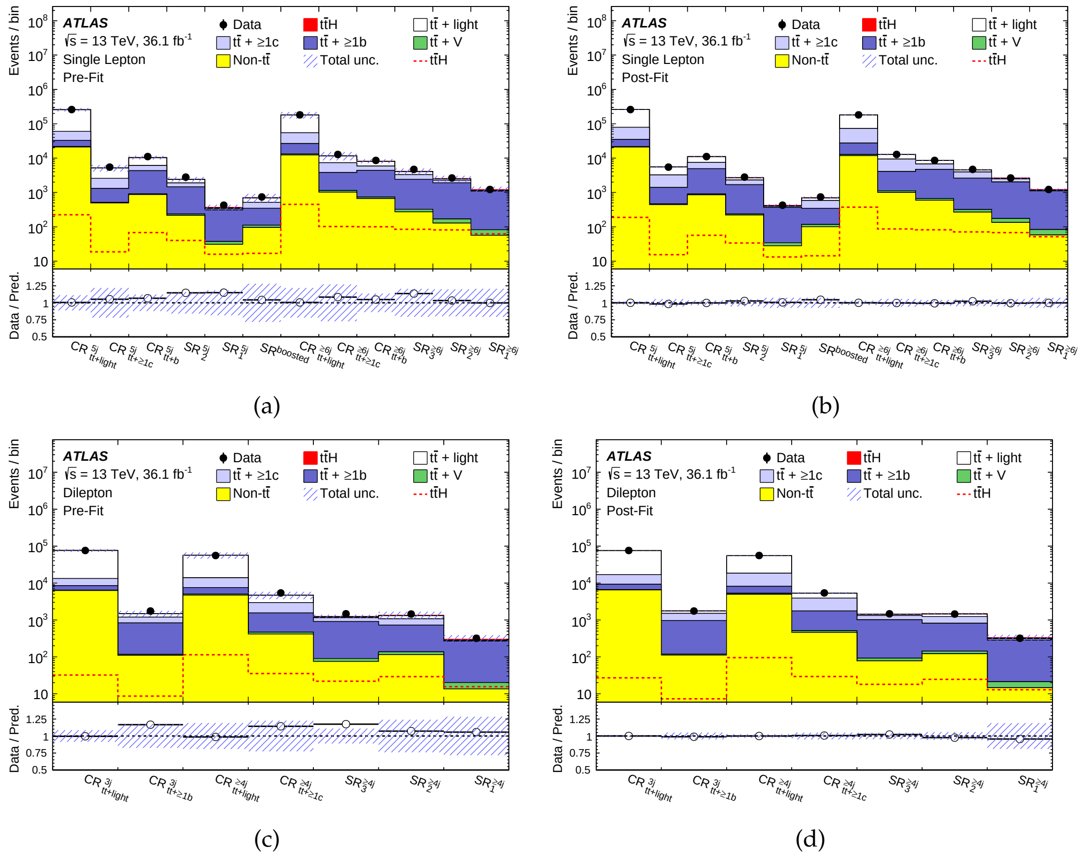


Figure 9.1: The event yields in each of the $t\bar{t}H(H \to bb)$ control and signal regions for the semilepton (top) and dilepton (bottom) channels, pre- (left) and post- (right) fit. Data are given by black circles and simulation by filled histograms. The hatched lines corresponds to the fitted uncertainty in the total prediction [109].

9.2 $t\bar{t}H(H \to \gamma\gamma)$

The resonant $H \to \gamma\gamma$ decay was one of the two "golden" channels used in the first observation of the Higgs boson. The other resonant Higgs decay used, $H \to ZZ(4\ell)$, is discussed in Section 9.3. These resonant channels have clean signatures and the ability to totally reconstruct the invariant mass of the Higgs from the final decay products. This differs from decays such as $H \to WW$ which either contains neutrinos due to leptonic W decays or jets (and thus large QCD backgrounds) from hadronic W decays.

However, both channels have relatively small Higgs decay branching ratios and thus, coupled with the small cross-section of $t\bar{t}H$ with respect to the other Higgs production mechanisms, the search for these decays are statistically limited.

The $t\bar{t}H(H \to \gamma\gamma)$ analysis [126] requires two isolated photons in the final state, with the $t\bar{t}$ pair either decaying hadronically or leptonically. In addition to $t\bar{t}H$, tH events are also targeted. In total nine categories, enriching either $t\bar{t}H$, tHq or tHW, are defined.

Three categories containing at least one prompt lepton and one b-tagged jet in the final state are defined. A $t\bar{t}H$ enriched leptonic category requires at least two jets. Two further categories targeting tH are constructed by requiring there is only one lepton in the final state. One category requires at most three central jets ($|\eta| < 2.5$) with no forward jets ($|\eta| > 2.5$) and the other requires at most four central jets with at least one forward jet.

Six hadronic categories are also constructed. The hadronic category requires five jets with one b-tagged jet. A BDT discriminant is built to reject ggH and a multijet background, and three cuts on the BDT are used to define four regions with differing $t\bar{t}H$ sensitivity. Two further categories, containing exactly four jets, with either one or two b-tagged, are defined to enrich tH events.

An unbinned likelihood fit is performed to determine $\mu(t\bar{t}H)$. The invariant mass of the two photons in the $t\bar{t}H$ enriched regions are shown in Figure 9.2.

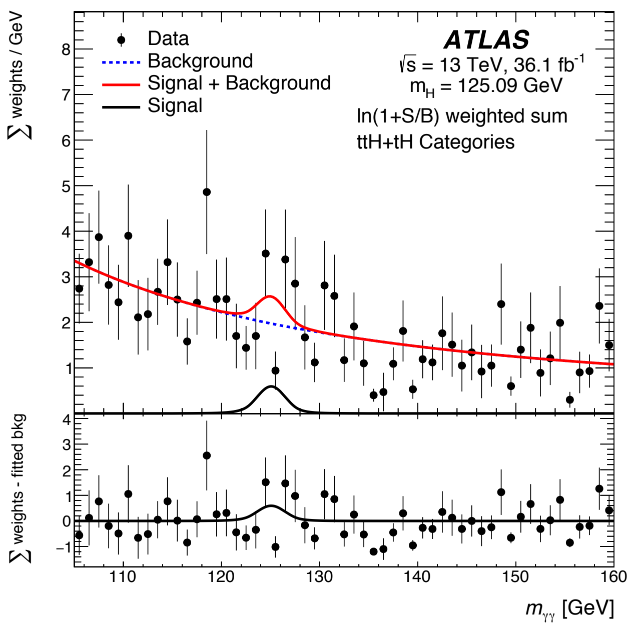


Figure 9.2: Weighted diphoton invariant mass spectra observed for the $t\bar{t}H$ enriched categories. Each event is weighted by $\ln(1 + S_{90}/B_{90})$ of the expected signal (S_{90}) and background (B_{90}) of the 90% signal quantile. The fitted signal-plus-background model is shown in red and the background component of the fit in blue [126].

9.3 $t\bar{t}H(H \to ZZ(4\ell))$

The resonant $H \to ZZ(4\ell)$ decay also results in a clean environment with which to search for $t\bar{t}H$ production.

The $t\bar{t}H(H \to ZZ(4\ell))$ analysis [127] requires four leptons with total invariant mass within approximately 5 GeV of m_H are selected. From these four leptons, two same-flavour opposite-charge pairs are required. To enrich the contribution from $t\bar{t}H$ at least one b-tagged jet, passing the 70% MV2c10 working point, is required in the event. This region is then required to contain at least four jets (including the b-tagged jet), by targeting hadronic decays of the remaining $t\bar{t}$ pair, or one additional lepton with at least two extra jets, targeting the semileptonic(dileptonic) $t\bar{t}$ decay.

Only 0.39 events were expected from simulation. No events were observed in data and thus only upper limits are set on the cross sections and signal strengths for the production mode, calculated using pseudo-experiments with the $\mathrm{CL_s}$ method [128]. The upper limit on the cross section is 0.11 pb from the expected 0.015 pb, and the upper limit on the signal strength is $\mu(t\bar{t}H) < 7.5$ at 95% confidence level.

9.4 Combination results

A combined fit of all four analyses measuring $t\bar{t}H$ production can be performed. Unlike $t\bar{t}H(H \to bb)$ or $t\bar{t}H$ multilepton, the $t\bar{t}H$ enriched regions in the resonant $t\bar{t}H(H \to \gamma\gamma)$ and $t\bar{t}H(H \to ZZ(4\ell))$ analyses contain admixtures of other Higgs production mechanisms, which are measured in data. For the combination fit these rates are fixed to the SM value, and thus different results of the fitted $\mu(t\bar{t}H)$ are used compared to the individual analyses. Also, as mentioned in the multilepton fits in Chapter 8, the tH production mechanism is fixed to the SM value.

The combined likelihood function $L(\mu, \boldsymbol{\theta})$ is obtained from the product of likelihood functions of the individual analyses. This results in approximately 500 nuisance parameters in the combined fit. Most parameters are treated as correlated across the channels except for cases where systematics are analysis specific.

The best-fit value of μ from the combined fit is

$$\mu = 1.17 \pm 0.19(\text{stat.})^{+0.27}_{-0.23}(\text{syst.}) = 1.17^{+0.33}_{-0.30}$$

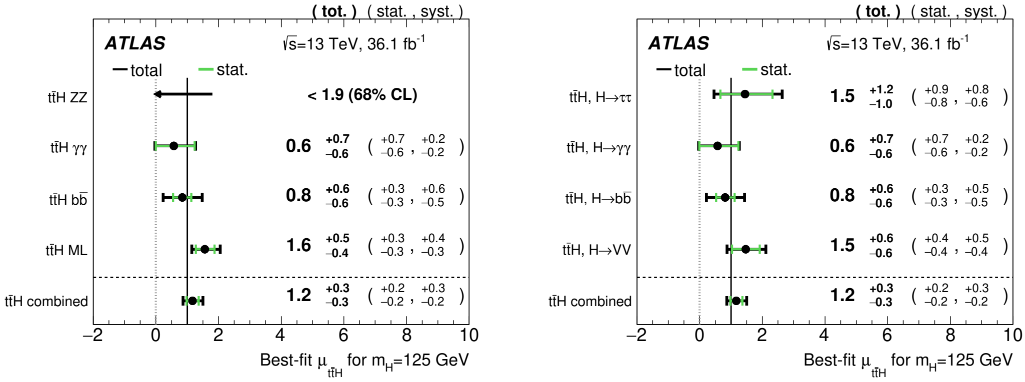
which corresponds to a discovery significance in the background-only hypothesis of 4.2σ from an expected significance of 3.8σ given the SM value. This constitutes evidence for $t\bar{t}H$ production.

The best-fit values of μ split up by analysis and Higgs boson decay is shown in Figure 9.3. All values are consistent with the SM. No candidate events are observed in the $t\bar{t}H(H \to ZZ(4\ell))$ analysis are observed so an upper limit on the value of μ at 68% CL is given.

If one assumes that the value of the best-fit rate observed is due to the SM Higgs boson, then a model-dependent extrapolation can be made to the inclusive phase space to measure the $t\bar{t}H$ cross section. The theoretical uncertainties for the $t\bar{t}H$ production cross section are removed from the best-fit rate. The measured $t\bar{t}H$ production cross section corresponding to the best fit value of μ is 590^{+160}_{-150} fb compared to the theoretical value of 507^{+35}_{-50} fb. As with the value of μ, the measured $t\bar{t}H$ cross section is consistent with the SM.

9.4.1 Coupling interpretation

The full combination of $t\bar{t}H$ analyses are sensitive to the coupling of the Higgs boson with top quarks, b-quarks, τ leptons, W and Z bosons and the effective coupling to photons. The sensitivity to the top Yukawa coupling comes from the $t\bar{t}H$ production itself, interference of W-H and t-H from tH production and the W-t interference from $H \to \gamma\gamma$. The κ parameterisation can be used to interpret whether these couplings differ significantly from the SM.



(a) (b)

Figure 9.3: Summary of (a) the measurements of μ from individual analyses and (b) the best-fit values of μ broken down by Higgs boson decay mode. The best-fit values of μ for the individual analyses on the left are extracted independently. A 68% confidence level upper limit on μ is given for $H \to ZZ(4\ell)$. The $H \to WW^*$ and and $H \to ZZ^*$ decays on the right are assumed to have the same signal-strength modification factor due to the weak sensitivity of ZZ and are shown together as VV [118].

Due to relatively poor accuracies on the individual coupling modifiers, the assumption is made that the fermion, κ_F, and boson, κ_V, couplings each scale by a common factor. This is motivated by the difference in the origin of the couplings; the couplings of the Higgs to the bosons is intrinsic within electroweak symmetry breaking and the couplings to fermions added later. κ_γ is expressed in terms of κ_F and κ_V and κ_g is fixed to κ_F, with κ_H modified accordingly. The parameterisations are taken from [112].

Figure 9.4 shows the results of a two dimensional likelihood scan in the κ_F-κ_V plane. The results agree well with the SM prediction of $\kappa_F = 1$ and $\kappa_V = 1$.

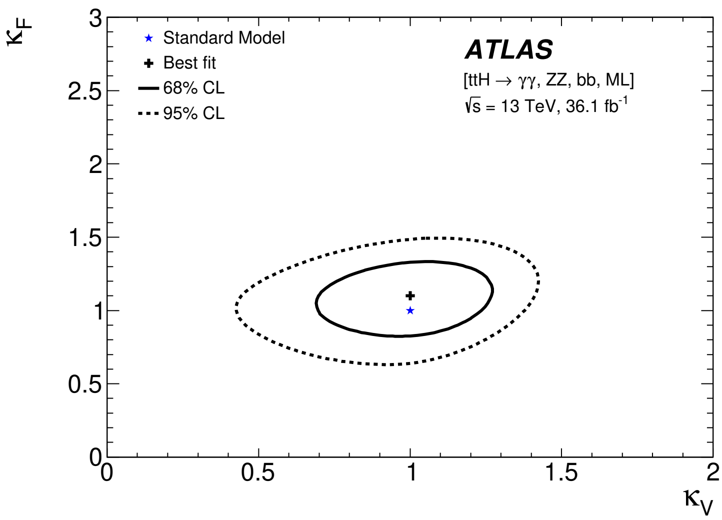
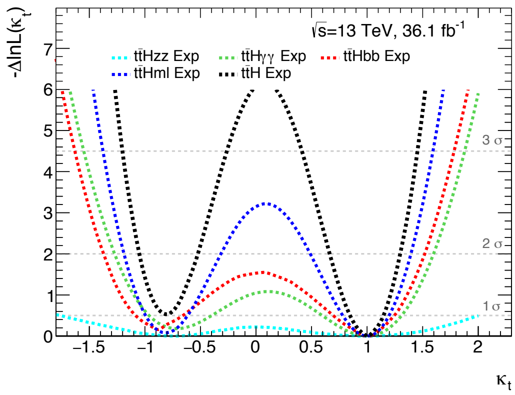


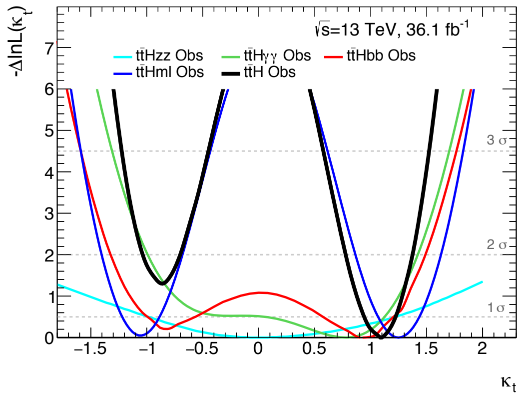
Figure 9.4: The 68% and 95% CL in the κ_F-κ_V two dimensional plane from the combination of all $t\bar{t}H$ channels [118].

κ_t can still be investigated in the combination. The expected and observed profile likelihood ratio as a function of κ_t for the combined $t\bar{t}H$ fit is shown in Figure 9.5. Only the (tree-level)

$t\bar{t}H$ and tH production cross sections are exploited: loops involving top quarks are fixed to their SM values. The observed value of κ_t starts to exclude the (BSM) $\kappa_t = -1$ case (due to tH production) but more sensitivity is needed to exclude this with confidence.



(a)



(b)

Figure 9.5: The (a) expected and (b) observed profile likelihood ratios as a function of κ_t for the combined fit. The $t\bar{t}H$ and tH modes are exploited, with all other coupling scale factors fixed to the SM value.

9.5 Observation of $t\bar{t}H$ production

An updated combined analysis of $t\bar{t}H$ production using 2017 data has recently been made with the ATLAS detector [129]. When combined with the Run-1 dataset, the observed (expected) significance in the search for top quark associated Higgs boson production with respect to the background-only hypothesis is 6.3σ (5.1σ), constituting observation of the $t\bar{t}H$ production mechanism with the ATLAS detector. This signals the end of the search analyses and the start of precise measurements of $t\bar{t}H$ and, by extension, the top quark Yukawa coupling.

Chapter **10**

Conclusions

The degree of non-factorisation of the proton beam profiles in beam separation scans is measured using luminosity and reconstructed vertex distribution data collected from the luminous collision region. A corresponding correction to the calibration quantity, σ_{vis}, is determined from this data in 2015 and 2016. The 1.0% systematic uncertainty due to non-factorisation applied to σ_{vis} in 2015 is reduced to 0.4% for 2016 data. This reduction allowed for the systematic uncertainty due to non-factorisation on the total integrated luminosity estimate to become a sub-dominant term, contributing to the precise 2.1% uncertainty quoted for the total luminosity collected in 2015 and 2016.

The development of a new multivariate method to reject non-prompt leptons is presented. The method uses information from the tracks nearby to a lepton to determine whether the lepton appears to have been produced from a source with non-zero lifetime, in addition to information about the amount of activity around the lepton. This method to reject non-prompt leptons is the first of its type with the ATLAS detector, and is shown to better reject non-prompt leptons than the standard isolation techniques often used in analyses with the ATLAS detector. As the method is designed to reject non-prompt leptons generally, any analysis can use it to improve their sensitivity. In addition to $t\bar{t}H$, the (not yet submitted) analysis measuring the $t\bar{t}W$ cross section at $\sqrt{s} = 13$ TeV is employing the technique and some searches for supersymmetry have investigated the use. Further optimisations to the algorithm are planned and a low-p_T, non-prompt lepton based BDT tagger is under development.

This multivariate method is used to reject non-prompt lepton backgrounds in the search for top quark associated Higgs boson production in multilepton final states with 36.1 fb^{-1} data collected by the ATLAS detector in 2015 and 2016. By comparing with the lepton selections from the analysis with 13.2 fb^{-1} data collected in 2015 and 2016, the sensitivity of the analysis including the non-prompt BDT and the charge-flip tagger is found to increase by roughly 20%. Further increases in the sensitivity of the analysis are achieved with the use of event multivariate algorithms. The measured signal strength of top quark associated Higgs boson production in multilepton final states was found to be $\mu = 1.6^{+0.5}_{-0.4}$, corresponding to an observed (expected) discovery significance of 4.1σ (2.8σ).

By combining the result of the multilepton analysis with the searches for $t\bar{t}H$ in $H \to \gamma\gamma$, $H \to bb$ and $H \to ZZ(4\ell)$ decays, a more precise signal strength of $\mu = 1.2 \pm 0.3$ is measured. This corresponds to a 4.2σ observed (3.8σ expected) discovery significance, constituting the first statistical evidence for $t\bar{t}H$ production with the ATLAS detector.

With increased data comes increased precision in the measurements of $t\bar{t}H$ production. Including proton-proton data collected in 2017, the $t\bar{t}H$ production process was observed with the ATLAS detector. By extrapolating to the 3000 fb^{-1} of data expected to have been collected by the ATLAS detector after the high luminosity upgrade, one expects the measurement precision of the total $t\bar{t}H$ cross section to be reduced to roughly 1%. In this regime any deviations from the SM expectation can be thoroughly tested and the search for new physics enhanced.

Bibliography

[1] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, `arXiv:1207.7214 [hep-ex]`.

[2] CMS Collaboration, S. Chatrchyan et al., *A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider*, Science **338** (2012) 1569–1575.

[3] Particle Data Group Collaboration, C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** (2016) 100001.

[4] T. D. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254–258.

[5] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1414.

[6] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[7] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[8] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377.

[9] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133.

[10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[11] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[12] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, `arXiv:hep-ex/9807003 [hep-ex]`.

[13] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course In Modern Particle Physics*. Wiley, 1984.

[14] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977) 641–653.

[15] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450.

[16] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298–318.

[17] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, `arXiv:1007.2241 [hep-ph]`.

[18] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, `arXiv:1410.8849 [hep-ph]`.

[19] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, `arXiv:0901.0002 [hep-ph]`.

[20] S. Weinzierl, *Introduction to Monte Carlo methods*, `arXiv:hep-ph/0006269 [hep-ph]`.

[21] T. Sjostrand, *Jet Fragmentation of Nearby Partons*, Nucl. Phys. **B248** (1984) 469–502.

[22] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145.

[23] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. **B238** (1984) 492–528.

[24] T. Sjostrand et al., *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175 [hep-ph]`.

[25] T. Sjostrand et al., *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[26] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, `arXiv:1410.3012 [hep-ph]`.

[27] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, `arXiv:hep-ph/0011363 [hep-ph]`.

[28] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, `arXiv:0803.0883 [hep-ph]`.

[29] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, `arXiv:0811.4622 [hep-ph]`.

[30] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, `arXiv:hep-ph/0409146 [hep-ph]`.

[31] C. Oleari, *The POWHEG-BOX*, Nucl. Phys. Proc. Suppl. **205-206** (2010) 36–41, `arXiv:1007.3893 [hep-ph]`.

[32] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, `arXiv:hep-ph/0204244 [hep-ph]`.

[33] S. Hoche, F. Krauss, M. Schonherr, and F. Siegert, *Automating the POWHEG method in Sherpa*, JHEP **04** (2011) 024, `arXiv:1008.5399 [hep-ph]`.

[34] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

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

[36] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[37] R. Garoby, *Plans for Upgrading the CERN Proton Accelerator Complex*, Journal of Physics: Conference Series **110** (2008) 2.

[38] CMS Collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

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

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

[41] LHCf Collaboration, O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.

[42] MoEDAL Collaboration, J. Pinfold et al., *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Jun, 2009. `https://cds.cern.ch/record/1181486`.

[43] TOTEM Collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[44] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. **122B** (1983) 103–116.

[45] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c*^{2} *at the CERN SPS Collider*, Phys. Lett. **126B** (1983) 398–410.

[46] UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN* $\bar{p}p$ *Collider*, Phys. Lett. **122B** (1983) 476–485.

[47] UA2 Collaboration, P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\bar{p}p$ *Collider*, Phys. Lett. **129B** (1983) 130–140.

[48] NA48 Collaboration, V. Fanti et al., *A new measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465** (1999) 335–348, `arXiv:hep-ex/9909022 [hep-ex]`.

[49] ATLAS Collaboration, R. Nickerson et al., *ATLAS inner detector: Technical design report. Vol. 2*, tech. rep., CERN, 1997. `http://cds.cern.ch/record/331064`.

[50] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, PoS **EPS-HEP2015** (2015) 261, `arXiv:1608.07850 [physics.ins-det]`.

[51] ATLAS Collaboration, *ATLAS central solenoid: Technical design report*, tech. rep., CERN, 1997. `http://cds.cern.ch/record/331067`.

[52] ATLAS Collaboration, N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical design report*, tech. rep., CERN, 1998. https://cds.cern.ch/record/381263.

[53] ATLAS Collaboration, M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013, ATLAS-TDR-19, CERN, September, 2010. https://cds.cern.ch/record/1291633.

[54] A. Ahmad et al., *The Silicon microstrip sensors of the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A578** (2007) 98–118.

[55] E. Klinkby, *Simulation of transition radiation and electron identification ability of the ATLAS TRT*, Nucl. Instrum. Meth. **A706** (2013) 79 – 82.

[56] ATLAS Collaboration, *ATLAS calorimeter performance Technical Design Report*, tech. rep., CERN, 1996. http://cds.cern.ch/record/331059.

[57] ATLAS Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, tech. rep., CERN, 1996. http://cds.cern.ch/record/331061.

[58] A. Artamonov et al., *The ATLAS forward calorimeters*, JINST **3** (2008) P02010.

[59] ATLAS Collaboration, *ATLAS tile calorimeter: Technical design report*, tech. rep., CERN, 1996. https://cds.cern.ch/record/331062.

[60] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, tech. rep., CERN, 1997. https://cds.cern.ch/record/331068.

[61] ATLAS Collaboration, *ATLAS magnet system: Technical design report*, tech. rep., CERN, 1997. https://cds.cern.ch/record/338080.

[62] ATLAS Collaboration, *ATLAS barrel toroid: Technical design report*, tech. rep., CERN, 1997. http://cds.cern.ch/record/331065.

[63] ATLAS Collaboration, *ATLAS endcap toroids: Technical design report*, tech. rep., CERN, 1997. http://cds.cern.ch/record/331066.

[64] ATLAS Collaboration, M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77** (2017) 317, arXiv:1611.09661 [hep-ex].

[65] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, Tech. Rep. ATLAS-CONF-2012-042, CERN, Geneva, 2012. https://cds.cern.ch/record/1435196.

[66] R. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, ASME. J. Basic Eng. **82** (1960) 35–45.

[67] ATLAS Collaboration, M. Aaboud et al., *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, Eur. Phys. J. **C77** (2017) 332, arXiv:1611.10235 [physics.ins-det].

[68] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2157687.

[69] ATLAS Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[70] ATLAS Collaboration, G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. **C76** (2016) 292, arXiv:1603.05598 [hep-ex].

[71] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87 – 116.

[72] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[73] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[74] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[75] ATLAS Collaboration, M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. **D96** (2017) 072002, arXiv:1703.09665 [hep-ex].

[76] ATLAS Collaboration, G. Aad et al., *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run*, JINST **8** (2013) P07004, arXiv:1303.0223 [hep-ex].

[77] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-029, CERN, Geneva, Jul, 2015. http://cds.cern.ch/record/2037702.

[78] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. http://cds.cern.ch/record/1700870.

[79] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. https://cds.cern.ch/record/2160731.

[80] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032032.

[81] ATLAS Collaboration, M. Aaboud et al., *Measurements of b-jet tagging efficiency with the ATLAS detector using* $t\bar{t}$ *events at* $\sqrt{s} = 13$ *TeV*, arXiv:1805.01845 [hep-ex].

[82] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, Tech. Rep. ATLAS-CONF-2018-006, CERN, Geneva, April, 2018. http://cds.cern.ch/record/2314418.

[83] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ TeV, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, 2017. http://cds.cern.ch/record/2261772.

[84] P. Grafström and W. Kozanecki, *Luminosity determination at proton colliders*, Prog. Part. Nucl. Phys. **81** (2015) 97–148.

[85] ATLAS Collaboration, M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** (2016) 653, `arXiv:1608.03953 [hep-ex]`.

[86] S. van der Meer, *Calibration of the Effective Beam Height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31, CERN, June, 1968. `https://cds.cern.ch/record/296752/`.

[87] O. Viazlo and A. L. Collaboration, *ATLAS LUCID detector upgrade for LHC Run 2*, Tech. Rep. ATL-FWD-PROC-2015-004, CERN, Geneva, Oct, 2015. `https://cds.cern.ch/record/2062038`.

[88] V. Cindro et al., *The ATLAS beam conditions monitor*, JINST **3** (2008) P02004.

[89] A. Bertin et al., *Calibrating the ATLAS luminosity detectors using beam separation scans*, Tech. Rep. ATL-COM-LUM-2010-021, CERN, Geneva, Jun, 2010. `https://cds.cern.ch/record/1271460`.

[90] ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, Tech. Rep. ATLAS-CONF-2010-027, CERN, Geneva, May, 2010. `https://cds.cern.ch/record/1277659`.

[91] L. Tomlinson, *An analytical determination of the parameters of the single-Gaussian model of bunch densities and their impact on the luminosity calibration by the van der Meer method*, Tech. Rep. ATL-COM-LUM-2013-009, CERN, Geneva, Mar, 2013. `https://cds.cern.ch/record/1523278`.

[92] S. M. White, H. Burkhardt, and P. Puzo, *Determination of the Absolute Luminosity at the LHC*. PhD thesis, Université Paris-Sud, 2010. `https://cds.cern.ch/record/1308187`.

[93] S. N. Webb, *Determining single beam density profiles from beam spot and luminosity data in van der Meer scans*, Tech. Rep. ATL-COM-DAPR-2014-004, CERN, Geneva, Feb, 2014. `https://cds.cern.ch/record/1662745`.

[94] S. N. Webb, *Factorisation of beams in van der Meer scans and measurements of the ϕ^*_η distribution of $Z \to e^+e^-$ events in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*. PhD thesis, The University of Manchester, June, 2015. `https://cds.cern.ch/record/2020875`.

[95] W. Kozanecki, T. Pieloni, and J. Wenninger, *Observation of Beam-beam Deflections with LHC Orbit Data*, Tech. Rep. CERN-ACC-NOTE-2013-0006, CERN, Geneva, Aug, 2013. `https://cds.cern.ch/record/1581723`.

[96] M. Bassetti and G. A. Erskine, *Closed expression for the electrical field of a two-dimensional Gaussian charge*, Tech. Rep. CERN-ISR-TH-80-06. ISR-TH-80-06, CERN, Geneva, 1980. `https://cds.cern.ch/record/122227`.

[97] Y. Coadou, *Boosted Decision Trees and Applications*, EPJ Web Conf. **55** (2013) 02004.

[98] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

[99] R. E. Schapire, *The Strength of Weak Learnability*, Mach. Learn. **5** (1990) 197–227.

[100] Y. Freund, *Boosting a Weak Learning Algorithm by Majority*, Inf. Comput. **121** (1995) 256–285.

[101] J. H. Friedman, *Stochastic Gradient Boosting*, Comput. Stat. Data Anal. **38** (2002) 367–378.

[102] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. D **89** (2014) 033009.

[103] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, `arXiv:0802.0007 [hep-ph]`.

[104] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152 – 155.

[105] V. Dao et al., *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-COM-PHYS-2015-642, CERN, Geneva, Jul, 2015. `https://cds.cern.ch/record/2032459`.

[106] D. V. D. Regie et al., *ATLAS electron, photon and muon isolation in Run 2*, Tech. Rep. ATL-COM-PHYS-2017-290, CERN, Geneva, Mar, 2017. `https://cds.cern.ch/record/2256658`.

[107] LHC Higgs Cross Section Working Group Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, `arXiv:1610.07922 [hep-ph]`.

[108] ATLAS Collaboration, M. Aaboud et al., *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP **12** (2017) 024, `arXiv:1708.03299 [hep-ex]`.

[109] ATLAS Collaboration, M. Aaboud et al., *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D97** (2018) 072016, `arXiv:1712.08895 [hep-ex]`.

[110] ATLAS Collaboration, G. Aad et al., *Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector*, Phys. Rev. **D92** (2015) 012006, `arXiv:1412.2641 [hep-ex]`.

[111] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP **06** (2013) 081, `arXiv:1303.4571 [hep-ex]`.

[112] LHC Higgs Cross Section Working Group Collaboration, J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, `arXiv:1307.1347 [hep-ph]`.

[113] ATLAS, CMS Collaboration, G. Aad et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 045, `arXiv:1606.02266 [hep-ex]`.

[114] J. Ellis and T. You, *Updated Global Analysis of Higgs Couplings*, JHEP **06** (2013) 103, arXiv:1303.3879 [hep-ph].

[115] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an], [Erratum: Eur. Phys. J. **C73** (2013) 2501].

[116] ATLAS Collaboration, G. Aad et al., *Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2016-058, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2206153.

[117] ATLAS Collaboration, G. Aad et al., *Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector*, Phys. Lett. **B749** (2015) 519–541, arXiv:1506.05988 [hep-ex].

[118] ATLAS Collaboration, M. Aaboud et al., *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. **D97** (2018) 072003, arXiv:1712.08891 [hep-ex].

[119] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[120] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[121] S. Frixione, P. Nason, and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv:0707.3088 [hep-ph].

[122] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[123] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, CoRR (2016), arXiv:1603.02754.

[124] ATLAS Collaboration, G. Aad et al., *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C71** (2011) 1577, arXiv:1012.1792 [hep-ex].

[125] I. Asimov and D. Shannon, *Franchise*. Creative Classic. Creative Education, 1989.

[126] ATLAS Collaboration, M. Aaboud et al., *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, arXiv:1802.04146 [hep-ex].

[127] ATLAS Collaboration, M. Aaboud et al., *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **03** (2018) 095, arXiv:1712.02304 [hep-ex].

[128] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693–2704.

[129] ATLAS Collaboration, M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, arXiv:1806.00425 [hep-ex].