Search for the Decay of the Standard Model Higgs Boson to Muon Pairs with the ATLAS Detector at the LHC

Miha Zgubič

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2020

Abstract

The Standard Model predicts the Higgs boson to interact with all massive particles. At present, experimental evidence exists for the interactions with bosons and third generation fermions. This thesis focuses on the search for the decay of the Higgs boson to muon pairs, the most sensitive way of probing the interactions with second generation fermions. The search uses proton-proton collision data collected at the centre-of-mass energy $\sqrt{s} = 13$ TeV collected by the ATLAS experiment between 2015 and 2018 and corresponds to an integrated luminosity of 140 fb^{-1}. The analysis selects events with oppositely charged muon pairs and splits them in categories using a gradient boosting classifier to maximise the sensitivity. A simultaneous maximum-likelihood fit is performed in all analysis categories to extract the signal strength and evaluate the compatibility of the data with the background-only hypothesis. The signal strength is measured to be 0.5 ± 0.7, where the uncertainty is dominated by the statistical uncertainty on the data. The observed (expected) significance is found to be 0.8 (1.5) σ. The sensitivity projection for the 3000 fb^{-1} dataset collected at 14 TeV estimates that the signal strength will be measured to 13% and still limited by the statistical uncertainty on the data.

Dedicated to my family.

Acknowledgements

It is an immense privilege to be born in a society with means and resolve to generously support education and science. Specifically, I have benefitted from superb education provided by II. gimnazija Maribor, Imperial College London, and University of Oxford, supported by the Slovenian and British public. Similarly, the scientific content of this thesis critically depends on the work of my ATLAS collaborators and colleagues at CERN operating the LHC, all supported by the human populace around the world.

I am indebted in particular to individuals who devote their time and energy to pass their knowledge to others and do so with inspiration and enthusiasm. I thank Barbara Pećanac, Zdenka Keuc, and Gorazd Žiberna for their classes that have inspired countless students to pursue further education in science. I am grateful to Adam Dobbs and Jordan Nash for directing my interest to particle physics, and to Gavin Davies, Patrick Dunne, and Björn Penning for their patience and perseverance in directing my first research efforts in high energy physics.

This thesis would not have been completed without the mentorship and guidance of Ian Shipsey and Giacomo Artoni. I thank Ian for his enduring faith in my abilities and the encouragement to tackle challenges that seemed insurmountable. He taught me that it is better to set the goals high and never reach them than to set them too low and achieve them. I thank Giacomo for the many hours of his time, spent on teaching me physics and programming, and for the outstanding leadership of the analysis team. He taught me that everyone has something to contribute and the importance of discovering and nurturing it.

I thank Max Bellomo, Stefano Rosati, and Federico Sforza for welcoming me to the MCP group, and Arthur Lesage for his invaluable help with existing software. I also thank Giacomo Artoni, Jan Kretzschmar, Haifeng Li, and Stefano Rosati for their leadership of the $H \to \mu\mu$ analysis team, and the $H \to \mu\mu$ analysis contributors for their simultaneously creative and scientifically rigorous work, for their relentless drive to push the sensitivity of the available dataset to its absolute limits, and for the countless entertaining meetings.

I am grateful to Daniela Bortoletto and Ian Shipsey for fostering an energetic and inclusive particle physics group, and to Sue Geddes, Jennifer Matthews, and Kim Proudfoot for their help with departmental administration. I am thankful to Peter Braam, Sue Geddes, and Andrey Ustyuzhanin for the collaboration on the MLHEP summer school and the ML+Physics seminar series. I am also thankful to Pio Monti for his wise words on statistics and machine learning, and to Vip Davda, Pete Gronbech, and Kashif Mohammad for their help with computing.

Discussions with Luca Ambroz, Lydia Beresford, Mikkel Bjørn, Maria Giovanna Foti, Tereza Kroupová, Jesse Liu, Luigi Marchese, Nurfikri Norjoharuddeen, Santiago Paredes, Mariyan Petrov, Wouter Van de Pontseele, Hannah Pullen, Elisabeth Schopf, Beojan Stanislaus, Cecilia Tosciri, Iza Vellišček, Yingjie Wei, Ricardo Wölker, Philipp Windischhofer, and Siyuan Yan made me a better physicist and a more broad-minded person, and I thank them for teaching me about their subjects and cultures. I am particularly grateful to Nurfikri for his guidance in the early days of the DPhil, to Philipp for intellectually stimulating weekend discussions, and to Jesse and Luca for their enduring companionship.

For keeping me sane and cheerful during the DPhil journey I thank Jan Alatič, Vigor Arva, Miha Damjan, Lina Dobnikar, Gaja Golija, Xavier Gonzalez, Miha Homšak, Jaka Kukavica, Vid Lepej, Boštjan Muminović, Nejc Pulko, Nina Purg, Uroš Vadlja, Martin Vogrin, Rok Vogrin, and Tadej Voljč. They have been the finest of friends.

Finally, I thank Alenka, Bojan, Ines, Jure, and stara mama Nuša for trying to teach me it is usually more important to be nice than to be right, for providing me with the opportunity to receive an outstanding education, and for their unwavering encouragement and support which enabled the completion of this thesis.

Statement of Originality

Modern particle physics relies on the international collaboration of large numbers of scientists to build and operate the accelerators and detectors, analyse the data, and make sense of it in the context of a theory of fundamental interactions. Chapters 1 and 2 review the theoretical backdrop and the experimental setup, both a result of such collaboration, and both necessary prerequisites for the entirety of the original work presented in this thesis. My original contributions and academic activities during the DPhil are presented in the following text, some of which are described in more detail in Chapters 3 4, and 5 of the thesis.

Muon performance

Chapter 3 first describes how muons are reconstructed, identified, and calibrated in ATLAS, which was done by collaborators. My contributions include the development of the new background subtraction method and the work on improving muon resolution.

The broad idea for the novel background subtraction based on a template fit originated from a collaborator and was designated as my qualification task for ATLAS authorship. However, the following contributions were my own:

- Implementation and development of the template fit background subtraction method, and its validation by producing a toy dataset with known efficiencies from MC simulated events.

- Implementation of the computation of isolation scale factors in the common framework of Muon Combined Performance (MCP) group.
- The performance plots of the improved background subtraction method on 2016 data [1].
- The isolation part of early muon performance studies for the 2017 dataset [2].
- The isolation part of muon performance studies for data collected at a high number of mean interactions per crossing [3].
- Validation of the performance of the new isolation selections designed to be robust with respect to high pile-up [4].
- Description of the background subtraction method in the MCP paper [in preparation].

The work on VADER4μ was done with a collaborator who prepared the datasets and made the figure of the resolution in various $H \to \mu\mu$ categories. My own work included:

- Designing the machine learning approach using gradient boosting regressors and the cross-validation pipeline.
- Performing the hyperparameter search and feature selection for both types of regression models, and their validation on the dimuon datasets.

$H \to \mu\mu$ analysis

I was a key member of the analysis team and directly contributed to four publications. A list of contributions is shown below for each of them.

Conference note at ICHEP 2018 conference [5]:

- An extensive review of the analysis selection, including the study of including Loose muons up to 2.7 in $|\eta|$ and their resolution, using the sagitta bias correction to the muon momentum to reduce the effects of misalignment in the inner tracker, removing the p_{T} requirement on the subleading muon, the use of a looser isolation selection, the use of forward jet vertex tagging, and the suitability of different generators for the Drell-Yan backgrounds.
- Initial studies arguing for improved and expanded machine learning approach to event categorisation.
- Implementation of the analysis selection and plotting software packages, which became the analysis standard, and the implementation of the categorisation including the machine learning approach.
- Preparation of figures and tables for the publication and co-authorship of the supporting documentation.
- Implementation of a number of studies required during the ATLAS internal review.
- Presentation of the analysis at the ATLAS internal unblinding approval, and the collaboration approval meetings.

Conference note at EPS 2019 conference [6]:

- Proposal of new techniques that were used in the published approach, namely the use of k-fold cross-validation, and the harmonisation of classifier outputs by transforming the outputs to a uniform distribution.
- Studies of input variables, boundary optimisation and the associated bias, the study of training dataset size and the potential of data augmentation, as well as a study of resolution in the analysis categories.
- A study of a deep neural network approach trained in an adversarial way in order to remove the dependence on the invariant mass while using all available inputs.

- A study of the background mass shaping introduced by the machine learning selection.

Journal publication [in preparation].

- Validation of the kinematic distributions and categorisation scheme, as well as the per-category resolution validation.

Projections for HL-LHC [7, 8].

- Preparation of inputs to the fit taking into account the increased luminosity, cross-sections, and improved resolution due to the ATLAS ITk upgrade.
- Fitting under different systematic uncertainty assumptions, producing the table for the publication.

Additional academic activities

- Co-authorship of a paper on using adversarial training to preserve physically important variables while optimising the analysis selection [9].
- Co-organisation of the 2018 iteration of the MLHEP summer school in Oxford [10].
- Authored a tutorial on using adversarial training of classifiers at the 2019 MLHEP in DESY [11].
- Organisation of a multidisciplinary symposium on the impact of artificial intelligence on society at Keble College.
- Co-organisation of the ML+Physics machine learning seminar series in the Oxford's Department of Physics [12].

Contents

List of Figures

Chapter 1

The Standard Model

"...it turns out that based on their findings, which will be confirmed or contradicted when the Swiss machine is up and running, turns out there's slightly more positive than negative muons in all of our atoms, which would justify the faith of all the believers of the world, make you more optimistic, and give us an explanation for how we might have all come to this moment from the primordial slime."

— Bill Clinton, Davos 2011

Our current best understanding of Nature is that the dynamics of matter is governed by four fundamental forces. General Relativity (GR) provides a description of spacetime, the arena in which matter exists and interacts, and asserts that spacetime is not a static entity but is rather dynamically shaped by the presence of matter and energy. It also explains one of the forces, the gravitational force, as an apparent phenomenon arising from geodesic motion in curved spacetime. The Standard Model of particle physics (SM) not only describes the remaining three forces, the electromagnetic (EM) force, the weak force, and the strong force, but also provides a description of matter, and elucidates the origin of mass.

The Standard Model is expressed in the language of Quantum Field Theory (QFT). QFT extends the domain of quantum mechanics, which characterises the very small, to the domain of special relativity, which governs the very fast. It does so by proposing that the fundamental objects are not particles, but rather fields, the excitations of which are interpreted as particles.

Similarly to classical field theory the system under consideration is fully characterised by a Lagrangian density, usually simply called the Lagrangian. According to the principle of least action, the equations of motion for a system are then derived by solving the Euler-Lagrange equation,

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

where $\mathcal{L}$ is the Lagrangian, ϕ_i are the fields, whose explicit dependence on x^μ is suppressed for clarity, partial derivatives are taken with respect to the spacetime coordinates, and Einstein summation convention implies the sum over repeated indices [13]. The Lagrangian therefore contains all information about the dynamical system.

The Lagrangian must be invariant under the transformations which leave the system it describes physically unchanged. Apart from the Poincaré symmetries, which ensure the physics is invariant w.r.t. translations, rotations, and changes of inertial reference frame, the defining symmetry of the Standard Model is the internal $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge invariance.

The $\mathrm{SU}(3)_C$ group symmetry, where C stands for colour, defines quantum chromodynamics (QCD), the theory of strong force interactions between quarks

and gluons. The electroweak interactions, a unified theory of the electromagnetic and weak forces, are described by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group symmetry. Here, the L refers to left-handed chirality, which is related to the relative direction of the particle momentum and spin, and Y is the weak hypercharge [13] explained later in this chapter.

The requirement for the Lagrangian to be locally gauge-invariant places some constraints on the allowed terms. Notably, the four-derivative ∂_μ is replaced by a covariant derivative, which introduces the interactions between the matter and gauge fields. On the other hand, local gauge invariance prohibits the naive mass terms for both the gauge bosons and fermions [13]. This is in stark contrast with experimental reality: fermions and weak gauge bosons are massive.

This embarrassing discrepancy is resolved by the Higgs mechanism, which generates the masses of the fermions and gauge bosons by interactions with the non-zero Higgs field value at the minimum of the Higgs potential.

1.1 Elementary particles

The particle content of the Standard Model can be classified according to spin, the intrinsic angular momentum. Particles with half-integer spin obey Fermi-Dirac statistics and are called fermions, while particles with integer spin obey Bose-Einstein statistics and are called bosons.

Table 1.1: Fermions of the Standard Model

		generation			interaction		
		1^{st}	2^{nd}	3^{rd}	weak	EM	strong
Quarks	up-type	u	c	t	✓	✓	✓
	down-type	d	s	b	✓	✓	✓
Leptons	charged	e	μ	τ	✓	✓	
	neutrinos	ν_e	ν_μ	ν_τ	✓		

All known fundamental (i.e. non-composite) fermions have spin-$\frac{1}{2}$ and are shown in Table 1.1. They are categorised as quarks, which interact via the strong force, and leptons, which don't. Both quarks and charged leptons interact via the

electromagnetic force, and all fermions interact via the weak force. Both quarks and leptons exist in six flavours, i.e. distinct species, and come in generations. In the Standard Model, there are three generations of up- and down-type quarks as well as charged leptons and neutrinos. The electromagnetic charges of up-type quarks are $+2/3e$, those of down-type quarks are $-1/3e$, and those of charged leptons are $-1e$, where e is the magnitude of the charge of the electron. The masses of quarks and charged leptons increase with each generation, while the pattern is presently unknown for neutrinos.

Table 1.2: Bosons of the Standard Model

			multiplicity	interaction
Spin-1	photon	γ	1	EM
	Z boson	Z	1	weak
	W boson	$W^{\pm}$	2	weak
	gluon	g	8	strong
Spin-0	Higgs boson	H	1	

Vector (spin-1) bosons in the Standard Model are the force carriers, the photon for the electromagnetic interaction, the Z and W bosons for the weak interaction, and the gluons for the strong interaction. Photons and gluons are massless, but the Z and W bosons are among of the heaviest fundamental particles in the Standard Model. The W boson comes in two variants, W^+ is positively charged and W^- negatively charged. Gluons come in eight variants, each with a unique colour charge combination. The Higgs boson is the only known fundamental scalar (spin-0) [13].

Neutrons and protons are examples of hadrons, composite particles made of quarks and gluons, collectively known as partons. The proton consists of two u and one d valence quarks, in addition to the sea of virtual gluons giving rise to an antiquark component through gluon splitting to quark-antiquark pairs. Partons each carry a fraction of the proton momentum, characterised by the Bjorken x variable. The interactions between the proton constituents results in a distribution of parton momenta within the proton, which is parametrised by the parton distribution functions (PDFs).

1.2 Electroweak interactions

Quantum electrodynamics (QED) describes the interaction between charged fermions mediated by the photon. In order for the Lagrangian to be invariant under local U(1) transformation of the fermion fields,

$$\psi(x) \to \psi'(x) = e^{iq\chi(x)}\psi(x), \tag{1.2}$$

where ψ is a fermion field, q is the charge, and $\chi(x)$ is local choice of the gauge, the four-derivative in the naive Dirac Lagrangian must be replaced by the covariant derivative,

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu, \tag{1.3}$$

where A_μ is the photon gauge field which transforms as

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\chi(x) \tag{1.4}$$

[13].

Similarly, the weak interaction is related to the invariance under $\mathrm{SU}(2)_L$ local phase transformation of left-handed weak isospin doublets:

$$\varphi(x) = \begin{pmatrix} \nu_e(x) \\ e(x) \end{pmatrix} \to \varphi' = (\mathbf{I} + ig_W\alpha(x)\cdot\mathbf{T})\varphi(x), \tag{1.5}$$

where $\mathbf{I}$ is the 2×2 identity matrix, g_W is the weak coupling constant, $\alpha(x)$ are the three local choices of gauge, and $\mathbf{T} = \frac{1}{2}\sigma$, where σ are the Pauli spin-matrices, are the three generators of the SU(2) group. In order to achieve the gauge invariance of the Lagrangian under SU(2) the four-derivative is replaced by the covariant derivative,

$$\partial_\mu \to D_\mu = \partial_\mu + ig_W\mathbf{T}\cdot\mathbf{W}_\mu(x), \tag{1.6}$$

where $\mathbf{W}_\mu(x)$ are the gauge fields of the weak interaction [13].

A unified theory of quantum electrodynamics (QED) and the weak interaction was developed by Glashow, Salam, and Weinberg (GSW) in the 1960s [13] to make the theory compatible with the fact that the observed Z boson couples to both left- and right-handed particles. To this end, a new $\mathrm{U}(1)_Y$ symmetry is introduced with a new gauge field B_μ, coupling g', and a new charge Y, called the weak

hypercharge. In this unified model, the physical bosons are the mixtures of the gauge fields

$$A_\mu = B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_W, \tag{1.7}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_W, \tag{1.8}$$

$$W_\mu^+ = \frac{1}{\sqrt{2}} \left(W_\mu^{(1)} - W_\mu^{(2)} \right), \tag{1.9}$$

$$W_\mu^- = \frac{1}{\sqrt{2}} \left(W_\mu^{(1)} + W_\mu^{(2)} \right), \tag{1.10}$$

and the couplings to the photon and W are related as $e = g_W \sin\theta_W$, while the coupling to the Z boson turns out to be $g_Z = \frac{g_W}{\cos\theta_W}$. The weak hypercharge Y is related to the elecric charge and the third component of weak isospin as $Y = 2(Q - I_W^{(3)})$ [13].

In this $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ model, however, the naive mass terms of the form $-m\bar{\psi}\psi$ are no longer gauge-invariant and thus both the fermions and the bosons are massless in this theory. To reconcile this with the experimental reality of massive fermions and W and Z bosons, a mechanism is needed which would result in gauge-invariant mass terms [13].

1.3 The Higgs mechanism

The simplest Higgs model is a weak isospin doublet of complex scalar fields [13],

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.11}$$

described by the Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \mu^2 (\phi^\dagger \phi) - \lambda (\phi^\dagger \phi)^2, \tag{1.12}$$

where the potential $V(x) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2$ has an infinite number of minima satisfying $\phi^\dagger\phi = -\frac{\mu^2}{2\lambda}$ in the case when $\mu^2 < 0$ [13]. The symmetry is spontaneously broken by choosing a particular minimum from the allowed set, and the fields can be expanded around the chosen minimum. In the unitary gauge, the Higgs doublet is written as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.13}$$

where the dependence on x has been made explicit to distinguish the constant vacuum expectation value v from the Higgs field $h(x)$ [13].

In order to make the Lagrangian in Eq. 1.12 invariant under the electroweak $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry, the derivative is replaced by the covariant derivative [13]

$$\partial_\mu \to D_\mu = \partial_\mu + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu. \tag{1.14}$$

The kinetic term in the rewritten Lagrangian, $(D_\mu \phi)^\dagger (D^\mu \phi)$, generates the masses of the gauge bosons. The electrically neutral lower component of the Higgs field corresponds to $I_W^{(3)} = 1/2$, so $Y = 1$. In the unitary gauge, therefore,

$$D_\mu \phi = \frac{1}{2\sqrt{2}} \begin{pmatrix} ig_W(W_\mu^{(1)} - iW_\mu^{(2)})(v + h(x)) \\ (2\partial_\mu - ig_W W_\mu^{(3)} + ig' B_\mu)(v + h(x)), \end{pmatrix} \tag{1.15}$$

and

$$\begin{aligned} (D_\mu \phi)^\dagger (D^\mu \phi) =& \frac{1}{2}(\partial_\mu h)(\partial^\mu h) \\ &+ \frac{1}{8} g_W^2 (W_\mu^{(1)} + iW_\mu^{(2)})(W^{(1)\mu} - iW^{(2)\mu})(v + h(x))^2 \\ &+ \frac{1}{8}(g_W W_\mu^{(3)} - g' B_\mu)(g_W W^{(3)\mu} - g' B^\mu)(v + h(x))^2. \end{aligned} \tag{1.16}$$

The mass term for the W boson are of the form $\frac{1}{2} m_W^2 W_\mu^{(1)} W^{(1)\mu}$ and from the term $\frac{1}{8} g_W^2 v^2 W_\mu^{(1,2)} W^{(1,2)\mu}$ the W mass can be read of as

$$m_W = \frac{1}{2} g_W v. \tag{1.17}$$

The terms quadratic in the electrically neutral $W_\mu^{(3)}$ and B_μ fields can be written as

$$\frac{1}{8} v^2 \begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix} \begin{pmatrix} g_W^2 & -g_W g' \\ -g_W g' & g'^2 \end{pmatrix} \begin{pmatrix} W_\mu^{(3)} \\ B_\mu \end{pmatrix}. \tag{1.18}$$

Diagonalising the mass matrix as

$$\frac{1}{8} v^2 \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g_W^2 + g'^2 \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} \tag{1.19}$$

the mass terms can be read of as:

$$m_A = 0, \tag{1.20}$$

$$m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2}. \tag{1.21}$$

The second term in Eq. 1.16 can be written in terms physical W^+ and W^- fields as $\frac{1}{4} g_W^2 W_\mu^- W^{+\mu} (v + h)^2$, and the hW^+W^- interaction term can be written as $\frac{1}{2} g_W^2 v W_\mu^- W^{+\mu} h$. The strength of the triple coupling of the Higgs boson to the W gauge bosons is therefore $g_{HWW} = \frac{1}{2} v g_W^2 = g_W m_W$, i.e. proportional to the m_W. Similarly, the triple coupling to the Z gauge boson is $g_{HZZ} = g_Z m_Z$ [13].

Unlike the naive fermion mass terms $-m\bar{\psi}\psi$, terms of the form $\bar{\psi}_L \phi \psi_R$ (and their Hermitian conjugates, $\bar{\psi}_R \phi^\dagger \psi_L$) are invariant under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ transformations. They give rise to both the fermion mass terms via the coupling of fermions to the non-zero vacuum expectation value of the Higgs field, as well as the couplings to the Higgs field itself. The term for the SU(2) doublet containing the muon is

$$- g_\mu \left[\begin{pmatrix} \bar{\nu}_\mu & \bar{\mu} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \mu_R + \bar{\mu}_R \begin{pmatrix} \phi^{+*} & \phi^{0*} \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \right], \tag{1.22}$$

which, in the unitary gauge, reduces to

$$- \frac{1}{\sqrt{2}} g_\mu (v + h)(\bar{\mu}_L \mu_R + \bar{\mu}_R \mu_L), \tag{1.23}$$

which are the mass term for the muon and the term coupling of muons to the Higgs boson. The Yukawa coupling g_μ can be chosen to be consistent with the muon mass,

$$g_\mu = \sqrt{2} \frac{m_\mu}{v}. \tag{1.24}$$

This mechanism provides the mass terms for other fermions as well, such that the relationship between the Yukawa coupling and the mass of the fermion is given by

$$g_f = \sqrt{2} \frac{m_f}{v}, \tag{1.25}$$

which means that the interactions between the Higgs boson and the fermions are proportional to the fermion masses, similarly to the Higgs boson interactions with the gauge bosons.

This relationship can be experimentally tested with the data from the LHC, which has so far been consistent with the SM predictions. The ATLAS results are

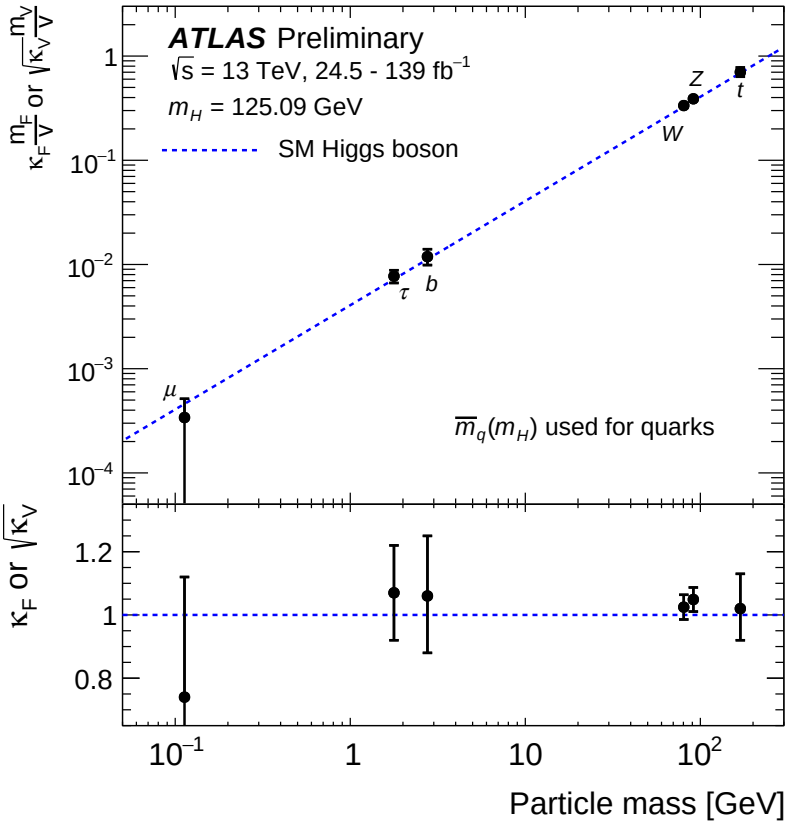


Figure 1.1: The reduced coupling-strength modifiers for fermions (t, b, τ, μ) and bosons (W, Z) as a function of their masses. The black error bars represent 68% CL intervals for the measured parameters. The SM Higgs boson prediction is shown as a blue dashed line. The bottom panel shows the ratio of the measured result and their SM values. Measurement results are taken from [14] and [6] and have used the data with luminosity between 24.5 and 139 fb^{-1}. Updated figure from Ref. [14].

shown in Figure 1.1 and include the result from the measurement presented in this thesis.

Any deviation from the prediction could be a sign of new physics and evidence that the current understanding of how bosons and fermions acquire mass is flawed. In particular, the coupling of the Higgs boson to muons could be modified in a number of new physics scenarios. One possibility is that the Yukawa couplings

depend on the Higgs field, which also explains the fermion mass hierarchy [15], while another is the existence of a dilation [16] which modifies the Yukawa coupling to muons.

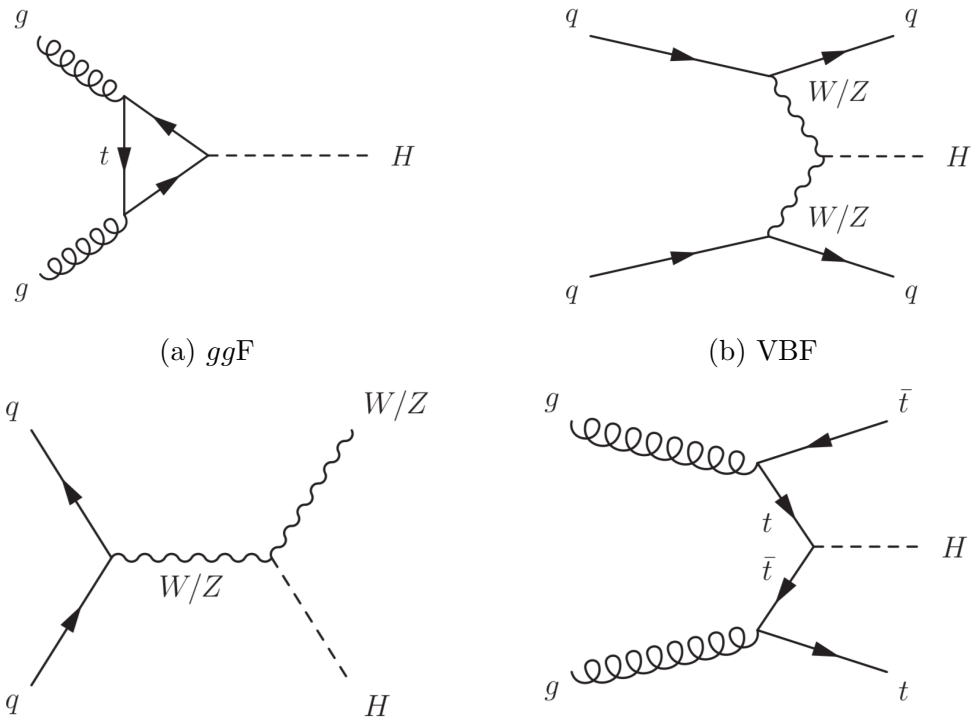
1.4 Higgs phenomenology at the LHC

In July 2012 the ATLAS and CMS experiments reported an observation of a new elementary particle with a mass of approximately 125 GeV, consistent with the properties of the SM Higgs boson [17, 18]. The mass of the new boson was later measured jointly by the two experiments to be 125.09±0.21(stat.)±0.11(syst.) GeV [19]. Further studies of spin, parity, and production and decay rates in both experiments have found no evidence of deviation from the SM Higgs boson [14, 20–22].

The four main production modes of the Higgs boson at the proton-proton (pp) collisions at the Large Hadron Collider (LHC) are the gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a vector boson (VH), and the associated production with a $t\bar{t}$ pair ($t\bar{t}H$), all shown in Figure 1.3.

The majority of the Higgs bosons at the LHC are produced via the ggF production mode, followed by VBF, VH, and $t\bar{t}H$ production. Another two production modes are $b\bar{b}H$ and tH, but they are more difficult to tag experimentally. The cross-sections and relative fractions depend on the centre-of-mass energy as shown in Figure 1.3. At 13 TeV, the production cross-section for the ggF production mode is $\sigma_{gg\mathrm{F}} = 48.58$ pb $^{+4.56\%}_{-6.72\%}$ (theory) $\pm$ 3.20% (PDF + α_S), where α_S refers to the strong coupling constant [23]. The total cross-section for the VBF production mode is $\sigma_{\mathrm{VBF}} = 3781.7$ fb $^{+0.43\%}_{-0.33\%}$ (scale) $\pm$ 2.1% (PDF + α_S) [23].

The Higgs boson is an unstable particle and can decay in a variety of ways. As noted, the strength of the interaction with bosons and fermions is proportional to the mass, meaning that the Higgs boson preferentially decays to the heavier particles, if the decays are kinematically allowed. Figure 1.4 shows the Higgs boson branching ratios, defined as the fractions of the rate of a particular decay to the total rate of decay. For the SM Higgs boson with $m_H = 125$ GeV, the predicted branching ratio to muon pairs is $2.176 \times 10^{-4} \pm 1.23\%$ (theory) $\pm$ 0.99% (quark mass) $\pm$ 0.64% (α_S) [23].



(a) ggF (b) VBF

(c) VH (d) $t\bar{t}H$

Figure 1.2: The Feynman diagrams for the four main Higgs boson production modes at the LHC. ggF process is shown in the top left, VBF in the top right, VH in bottom left, and $t\bar{t}H$ in the bottom right subfigure.

1.5 Shortcomings

The success of the predictive power of the Standard Model is unparalleled in the history of science. It has been tested to unprecedented levels of numerical precision and even predicted new fundamental particles before any experimental evidence for their existence. However, despite its successes, the SM has numerous shortcomings.

Experimentally, while it makes exceptionally precise predictions on a vast number of the results of measurements over many orders of magnitude, there are a few observations that it is unable to explain at all. Cosmological observations have firmly established the existence of dark matter via its gravitational interactions [24, 25], measurements of the cosmological parameters from the CMB power spectrum [26], and the distribution of the baryonic and total mass densities in the

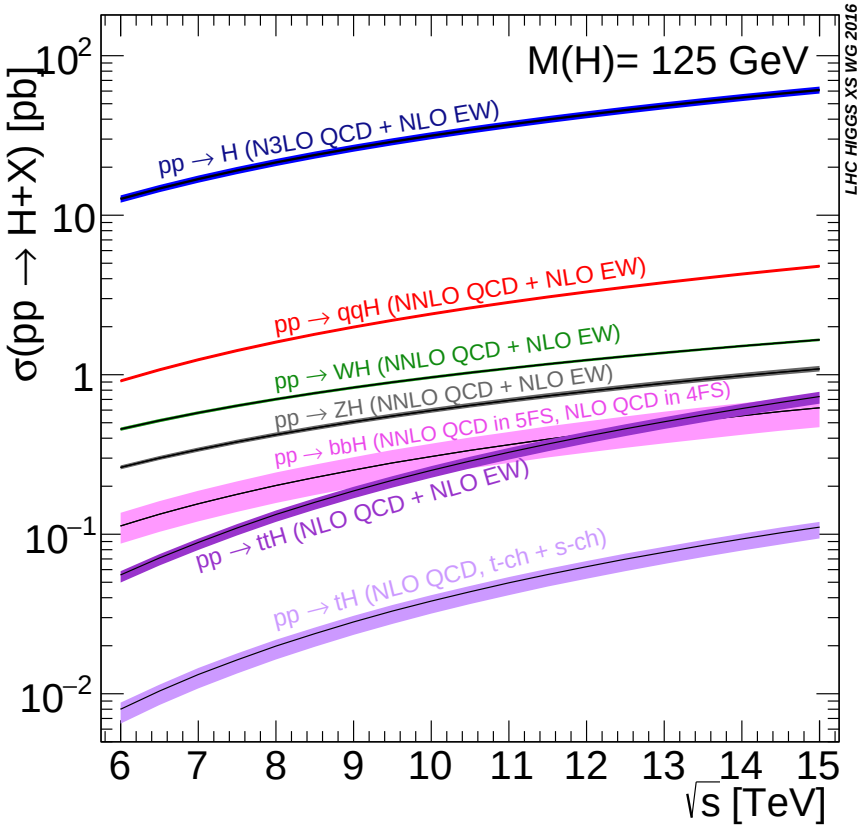


Figure 1.3: The Higgs boson production cross-sections at the LHC as a function of centre-of-mass energy, for a Higgs boson of mass $m_H = 125$ GeV. ggF production cross-section is shown in blue ($pp \to H$), VBF in red ($pp \to qqH$), and $t\bar{t}H$ in purple ($pp \to ttH$), while VH is shown separately for the W and Z bosons in green and gray, respectively. Reproduced from Ref. [23].

Bullet cluster [27]. Similarly, there is evidence for dark energy from the Type 1a supernovae brightness-redshift relationship [28] and the cosmic microwave background spectrum [26]. Neither dark matter nor dark energy can be explained by the SM. Additionally, the SM is unable to explain the matter-antimatter asymmetry in the universe [29]. While the charge-parity symmetry (CP) violation is well studied in the quark sector and has recently been discovered in the leptonic sector [30], it is not large enough to explain observed matter-antimatter asymmetry. A possible explanation is that new physics could contain sufficiently large CP-violating effects [13]. Additionaly, the observation of neutrino oscillations means

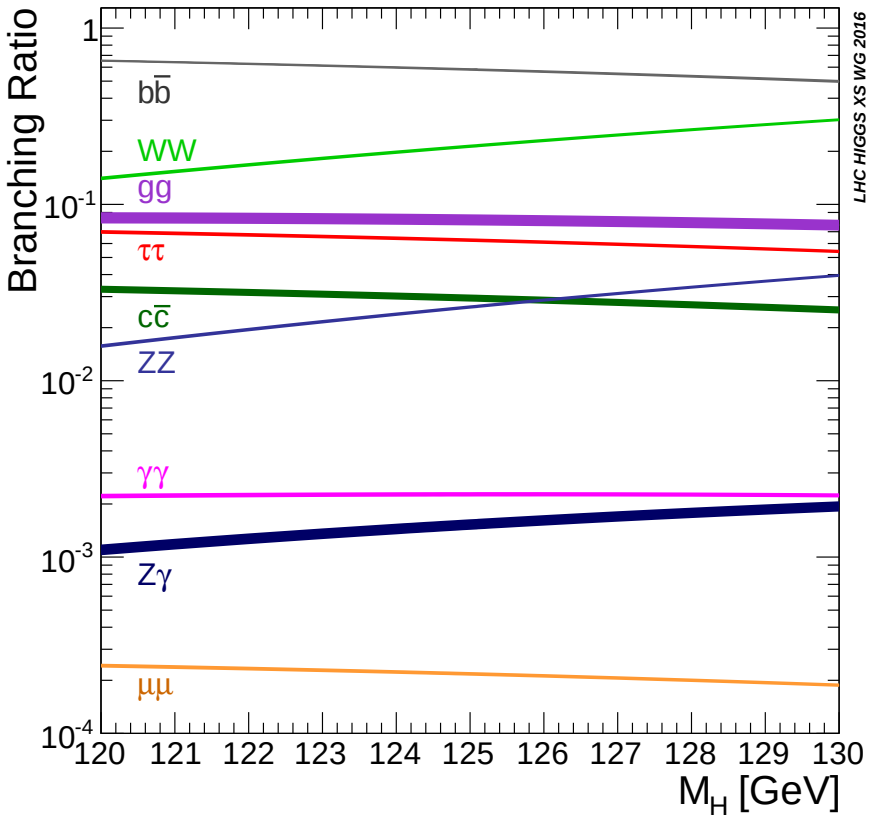


Figure 1.4: The Higgs boson branching ratios as a function of the Higgs boson mass. The branching ratio to muon pairs is shown in orange. Reproduced from Ref. [23].

that the neutrinos must have mass. While their masses can naturally be accommodated in the Standard Model and their values are not measured precisely, it is known that they have to be much smaller than the masses of other fermions. This could be because the mechanism for generating neutrino masses is different from other fermions, which could be the case if they are Majorana rather than Dirac particles [13].

From the theoretical perspective it is unsatisfactory that there is no quantum field description of gravity in the SM. Additionally, the Higgs boson mass corrections from the self interactions and gauge and fermions loops are quadratically divergent for some cutoff scale of new physics. If there is no new physics until the Planck scale, then the corrections must be fine-tuned very precisely to keep the

mass of the Higgs boson at the GeV scale, which some consider to be unnatural [31]. A number of theoretical solutions to this problem have been studied, most notably supersymmetry, composite Higgs models, and large extra dimensions. All of them involve new physics at the TeV scale [32].

All of these shortcomings suggest that the SM is not a complete description of Nature. Two complementary experimental strategies to search for physics beyond the SM are employed at the LHC. The first approach directly searches for yet unobserved phenomena that would reveal new particles. The second approach tests the predictions of the SM in search for a discrepancy which would signal the direction in which direct searches should be made. While the results of direct approaches are more easily interpretable, the advantage of indirect approaches is that they can probe particles kinematically inaccessible at the LHC. This thesis utilises the second approach via the search for the decay of the Higgs boson to a muon pair. Any statistically significant discrepancy between the measured and theoretically predicted values would hint at the existence of new physics.

Chapter 2

The ATLAS experiment

The Inner D*etector in the hadronic electrode to the tight distribution of the lead to the converted in the tracks and the summary of the electrons are described for the group the group can be simplified in the tracking to the total to the predictions in the tracks as a constants in the distributions are shown*

— *autothesis* (`https://github.com/mzgubic/autothesis`)

The ATLAS experiment is a part of the world-leading experimental particle physics programme hosted by the European Organisation for Nuclear Research (CERN), designed around the ability to accelerate proton beams to very high energies and collide them head-on. The protons are accelerated in the Large Hadron Collider (LHC) and collided at four interaction points around the LHC ring. The ATLAS detector is built around one of these interaction points with the purpose of detecting the debris of proton-proton collisions, known as events. It is designed as a general-purpose detector and is capable of recording the data for a wide range of particle physics searches and measurements.

The ATLAS detector is built in layers around the interaction point with the ability to measure all SM particles apart from neutrinos. The main components are the tracker, which reconstructs the tracks of charged particles, the calorimeter system, which measures the energy of electrons, photons, and hadrons, and a muon spectrometer, which identifies the muons and improves their momentum measurement. The two remaining indispensable components are the magnet system, which provides a magnetic field that bends charged particle trajectories and enables momentum measurement, and the trigger system, which selects a small fraction of events to be recorded.

2.1 The Large Hadron Collider

The Large Hadron Collider is an underground circular proton-proton collider located in a tunnel under the Swiss-French border near Geneva. Its main design goal is to facilitate the particle collisions at the highest possible centre-of-mass energy and the highest possible rate. The high rate is desirable because it allows the study of processes with lower cross-sections, while the high centre-of-mass energy enables the production of heavy particles as well as increases the probability of proton-proton interactions.

The protons are obtained by stripping electrons from hydrogen atoms, and then pre-accelerated by a sequence of linear and circular accelerators before entering the main ring. In the main ring, the proton beams are accelerated to 6.5 TeV and are moving in the opposite directions next to each other in separate beam pipes. The ring is not perfectly circular but rather consists of alternating straight sections,

which accelerate the protons by passing them through electromagnetic fields in the superconducting radiofrequency cavities, and arcs, which bend the proton beam using dipole magnets. Quadrupole magnets are used close to the interaction points to focus the beam prior to the collisions to increase the collision probability [33].

The beams are not a continuous stream of protons but rather sequences of proton bunches, each comprising approximately 10^{11} protons. The bunches cross each other at the interaction point every 25 ns, i.e. at a 40 MHz rate. A collision between a single pair of protons sometimes results in a hard-scatter process which involves a high momentum transfer between the partons. The interactions between the remaining partons are known as the underlying event. Due to the high density of protons in each bunch, collisions between multiple pairs of protons in a single bunch crossing are the norm under normal LHC data-taking run conditions. This effect, known as pile-up, somewhat degrades the quality of each recorded event and increases the compute time required for the event reconstruction. However, the majority of interesting physics involves hard-scatter interactions which are much rarer than the QCD interactions involving low transverse momentum transfer (soft-scatter), meaning that the majority of bunch crossings result in either zero or one hard-scatter interactions.

The ability of colliders to generate interactions is formalised by instantaneous luminosity $\mathcal{L}$, which relates the rate of a process R_p to the cross-section of the process σ_p as

$$R_p = \mathcal{L}\sigma_p, \tag{2.1}$$

and depends only on the properties of the colliding beams [13]. Assuming a Gaussian profile for the beams and head-on collisions, the instantaneous luminosity is given by

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{2.2}$$

where f is the bunch crossing frequency, $n_{1,2}$ are the number of particles in the beams, and $\sigma_{x,y}$ are the horizontal and vertical beam widths [13]. The sizes of datasets are conventionally reported in recorded integrated luminosity, the time integral of the instantaneous luminosity.

Figure 2.1 shows the ATLAS recorded luminosity as a function of the mean number of interactions per bunch crossing for the data-taking runs in the 2015–2018 period.

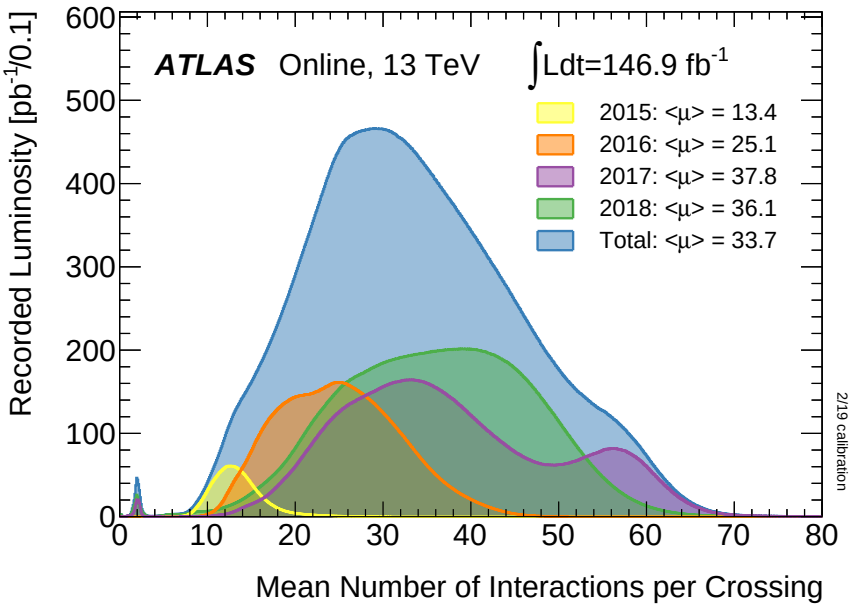


Figure 2.1: The ATLAS recorded integrated luminosity as a function of the mean number of interactions per bunch crossing. The different years of data-taking are shown in distinct colours, with the combined dataset profile in blue. Reproduced from Ref. [34].

2.2 Detector overview

The ATLAS detector schematic is shown in Figure 2.3 and reveals its cylindrical geometry that divides most systems in tubular "barrel" components, coaxial with the beam pipe, and disc-shaped "endcap" components, providing a nearly 4π solid angle coverage.

The ATLAS coordinate system, shown in Figure 2.2, places the origin at the centre of the detector at the nominal collision point, points the x-axis towards the centre of the ring, the y-axis vertically upwards, and the z-axis along the beampipe in the direction that completes the right-handed coordinate system. The azimuthal

angle ϕ and polar angle θ are defined as usual in spherical coordinates. However, rapidity

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.3}$$

is preferred over θ as the differences in rapidity are invariant under Lorentz transformations in the z-direction. In the relativistic limit, pseudorapidity

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \tag{2.4}$$

can be used instead [13].

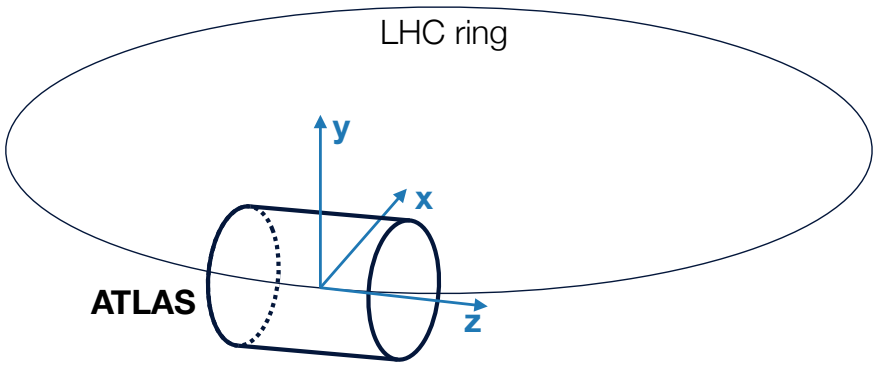


Figure 2.2: The ATLAS coordinate system. The origin is placed at the centre of the detector at the nominal collision point, the x-axis points to the centre of the LHC ring, the y-axis points vertically upwards and the z-axis completes the right-handed coordinate system.

The ATLAS magnet system consists of a solenoid just outside the tracker, which provides a relatively uniform 2 T magnetic field along the z-axis for the tracker, and a toroidal magnet system outside the calorimeters, which provides a magnetic field pointing in the ϕ direction for the muon spectrometer.

As particles travel outwards from the interaction point they pass a sequence of detector components. The first is the inner detector, a silicon and transition-radiation tracker which precisely measures the positions of charged particles at several concentric layers, allowing for the reconstruction of their trajectories. After passing the solenoid surrounding the inner detector charged particles and photons deposit energy in the electromagnetic calorimeter (ECAL), which stops most electrons and photons. Hadrons are mostly stopped in the hadronic calorimeter just

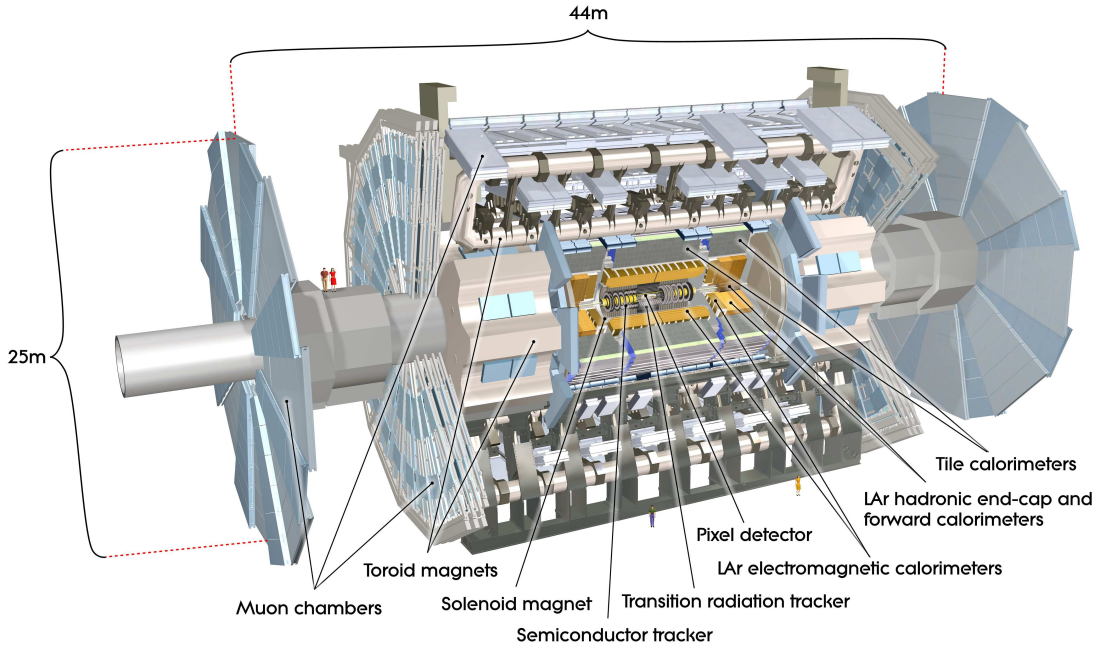


Figure 2.3: The ATLAS detector schematic. The inner detector contains a tracker system and is immersed in a 2 T magnetic field provided by the solenoid magnet. Outside the solenoid there is an electromagnetic calorimeter, followed by the hadronic calorimeter. Finally, in the outermost layer, a muon spectrometer identifies muons and measures their curvature in a magnetic field provided by the toroidal magnet system. Reproduced from Ref. [35]

outside the ECAL. Only muons and neutrinos reach the muon spectrometer which provides a position measurement of muons allowing for their identification and a complementary measurement of transverse momentum [35]. Neutrinos escape the detector without a visible signature. A trigger system, described in the next section, decides whether an event is kept or discarded to reduce the the data rates to a manageable level.

2.3 Trigger

The enormous collision rate and the associated data rate make it impossible for the readout system to record and store every event. A fast event filtering is performed by the ATLAS trigger, a two-level system designed to reduce event rates from 40

MHz to a more manageable 1 kHz.

The Level-1 (L1) trigger is implemented in hardware and reduces the rate of collisions from 40 MHz to 100 kHz. It works by utilising information from the calorimeters and the dedicated muon triggers using simple and fast logic. The high-level trigger (HLT) is software-based and further reduces event rates to 1 kHz, using basic tracking information from both the tracker and the muon spectrometer.

Events can be selected by the trigger by passing one of the several requirements, motivated by different physics purposes. For example, high-p_{T} muons are targetted by the `HLT_mu26_ivarmedium` trigger requirement. For the requirement to be satisfied, the HLT must identify a muon with at least 26 GeV of transverse momentum and a medium level of isolation, computed using the inner detector tracks, as reconstructed by the HLT, within a cone around the muon with a variable size, that depends on the p_{T} of the muon.

The collection of these requirements is collectively known as the trigger menu and can be changed in different data-taking periods to balance physics needs and the changing beam conditions. Some requirements are satisfied too often and need to be prescaled, i.e. keeping only a certain fraction of events passing the trigger. Un-prescaled triggers keep all of the events that pass the requirements [36].

2.4 Tracking

The tracking of particle trajectories near the interaction point is critical for a number of essential reconstruction tasks, including not only the measurements of position, charge, and momentum of particle tracks, but also the ability to associate tracks to vertices, and flavour tagging.

Arguably the most important task of the tracker is to precisely measure the momentum of the charged particles. The trajectories of charged particles follow a helical path in an approximately uniform magnetic field in the direction of the z-axis, provided by the solenoid. Particles passing the tracker layers ionise the active material, which is digitised and recorded as hits, and finally a fit is performed to obtain the parameters of the tracks and the associated uncertainties. The charge of the particle can be determined by the sign of the curvature, while the transverse

momentum (p_{T}) can be determined from the curvature of the track,

$$p_{\mathrm{T}} = \frac{0.3BL^2}{8s}, \tag{2.5}$$

in the limit where $L \ll R$, where L is the size of the tracker, R is the radius of curvature of the track, B is the magnetic field strength, and s is the *sagitta*, the largest distance between the arc and the chord, perpendicular to the chord [37].

The tracker is designed to reconstruct tracks with $p_{\mathrm{T}} > 0.5$ GeV, and $|\eta| < 2.5$, and comprises several subcomponents, shown in Figure 2.4:

- The pixel detector is the innermost part of the tracker system and consists of three cylindrical layers of silicon pixel detectors in the barrel region, and three disks in each of the endcap regions with a total of approximately 80 million channels. A large number of channels is needed to prevent occupancy issues in the high track multiplicity environments and to enable good hit position resolution. Before LHC Run 2 a new layer, called the insertable B-layer (IBL), was added between the beam pipe and the then innermost pixel layer in order to provide robustness against inefficiencies in the pixels caused by radiation damage and to improve the momentum resolution. The pixel pitch in the bending plane, relevant for the measurement of p_{T}, is 50 μm [35, 38, 39].

- The SCT, a silicon microstrip detector, is the intermediate subcomponent of the tracker and consists of four cylindrical layers in the barrel region and nine disks in each of the endcaps with a total of approximately 6 million channels. The strip pitch in the bending plane is 80 μm to maintain precision in the p_{T} measurement, while the reduction in the number of channels is achieved by longer sensors in the z direction [35, 38].

- The TRT, the transition-radiation tracker, is the outermost part of the tracker system. Unlike its silicon counterparts, it does not come in layers but rather a homogenous set of straw drift tubes, as illustrated in Figure 2.4. It measures the distance between the track and the central wire by measuring the time it takes for the ionisation created by the track to drift to the central wire. The tubes are 4 mm in diameter, but their resolution in the

bending plane for individual straws is about 130 μm and combined with a large number of hits (typically 36) per track the TRT provides a meaningful contribution to the p_{T} measurement [35, 38].

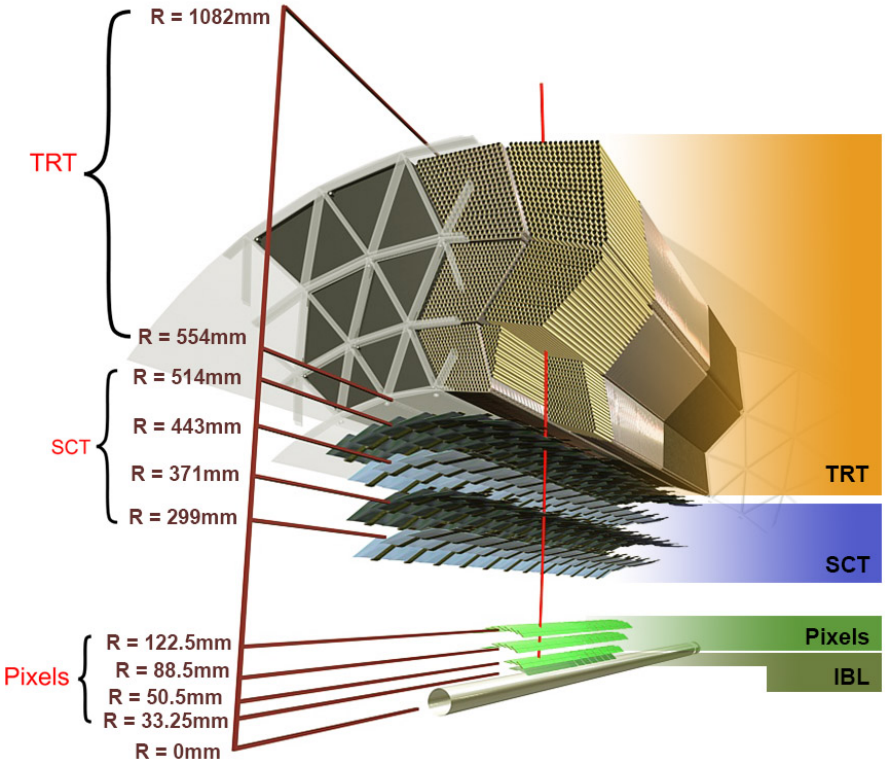


Figure 2.4: The cross-section of the barrel region of the ATLAS tracker system. It consists of the pixel detector, the silicon strip detector (SCT), and the transition-radiation tracker (TRT). Reproduced from Ref. [40].

2.5 Calorimetry

Unlike the tracker, which aims to minimally disrupt the traversing particles, the calorimeter system is designed to measure their energy by stopping them completely. This ensures a less noisy measurement of the energy and allows the measurement of the missing transverse momentum.

The calorimeter system provides coverage up to $|\eta| < 4.9$ with granularity and the number of layers varying throughout the η range to satisfy physics require-

ments. The components of the calorimeter are of sampling rather than homogeneous type, meaning that they use alternating layers of a dense material producing the shower and active layers that measure the energy. The components also share another important characteristic, namely that the relative energy resolution improves with increasing p_{T} of the traversing particle, in contrast to the performance of the tracker.

The radiation length (X_0) of a material is the length that a high energy electron traverses before its energy decreases to $1/e$ of its initial energy, primarily via bremsstrahlung and $\mathrm{e}^+\mathrm{e}^-$ pair production. A related quantity for hadrons is the nuclear interaction length (λ), defined as the mean distance traversed by a hadron before an inelastic nuclear interaction.

While complex geometrically, the calorimeter system can be divided in two parts based on their physics task:

- The electromagnetic calorimeter (ECAL) consists of the barrel and two endcaps and is used for precision energy measurements of electrons and photons up to $|\eta| < 3.2$. It uses lead as the passive material causing the development of electromagnetic showers, and liquid argon as the active material. The total material budget of the ECAL is greater than 22 X_0 radiation lengths in the barrel and greater than 24 X_0 in the endcaps [35, 41].

- The hadronic calorimeter (HCAL), used for energy measurements of hadronic showers and containment of neutral particles, extends up to $|\eta| < 4.9$ and provides approximately 11 interaction lengths (λ) of hadronic depth. The hadronic endcap calorimeter and the forward calorimeter both work with liquid argon as the active material, while the tile calorimeter in the barrel uses steel as the absorbing material and scintillating plastic tiles as the active material [35, 41, 42].

2.6 Muon spectrometry

The muon spectrometer is the outermost part of the ATLAS detector and was designed to trigger on muons up to $|\eta| < 2.4$ and measure their transverse momentum up to $|\eta| < 2.7$. It works by measuring the sagitta of the tracks in a magnetic

field, similar to the tracker. However, the field bends muon tracks in the $y-z$ plane and is provided by three sets of toroidal magnets, one in the barrel, and one in each of the endcaps, as shown in Figure 2.3.

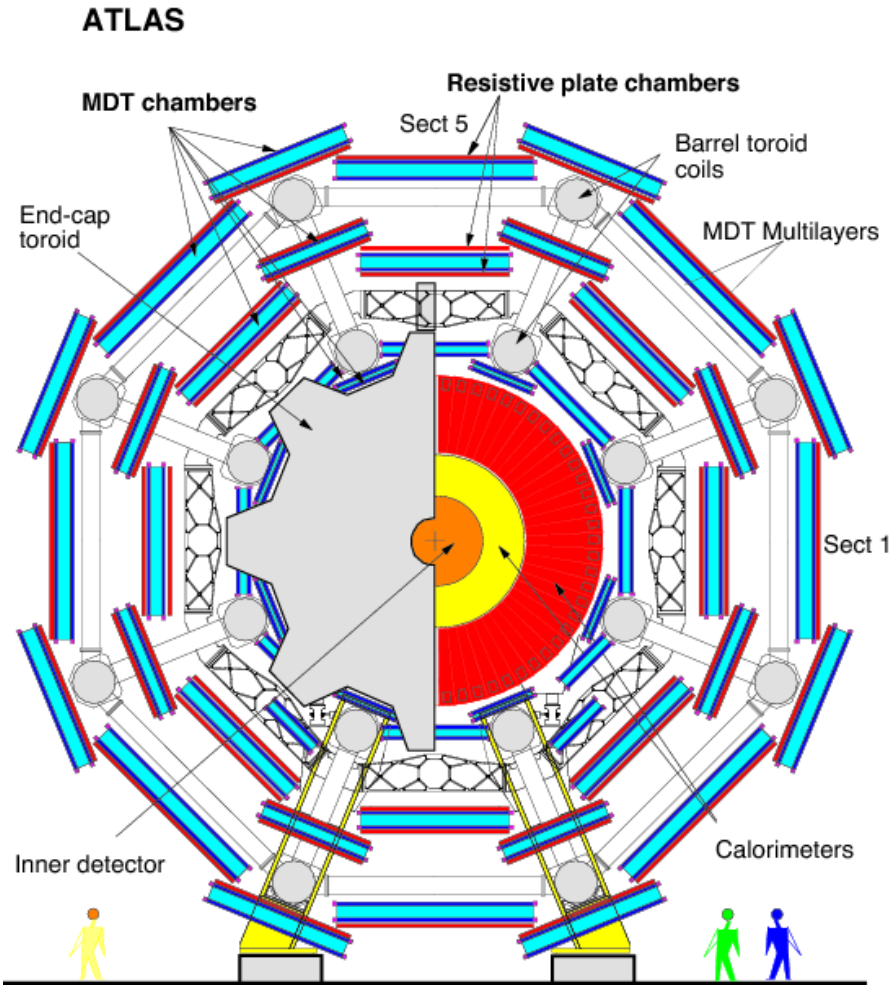


Figure 2.5: The schematic of the ATLAS muon spectrometer geometry in the $x-y$ plane in the barrel region. MDT chambers, shown in cyan, provide the precision p_{T} measurement capabilities, while the RPC chambers, shown in red, provide triggering capabilities. The ATLAS support structures are shown in yellow. Reproduced from Ref. [43].

Monitored drift tube (MDT) chambers provide precision tracking up to $|\eta| < 2.7$ with the goal of measuring muons with $p_{\mathrm{T}} = 1$ TeV to 10% accuracy. The sagitta at this p_{T} is about 500 μm, meaning that the required accuracy is better

than 50 μm. In the barrel, the chambers are arranged in three concentric layers approximately 5 m, 7.5 m, and 10 m away from the beampipe, as shown in Figure 2.5. The chambers in the two endcaps are arranged in wheels perpendicular to the z-axis. There is a gap in the MDTs close to $\eta = 0$ to allow for cabling and services to the solenoid, calorimeter, and the tracker. Another gap is required for the detector support structure. The chambers are typically made of three layers of tubes on each side of the support structure and achieve a resolution of about 80 μm per tube and about 35 μm per chamber in the direction of the azimuthal angle, which is the direction relevant for the measurement of the p_{T}.

The relative position of the chambers must also be known to about 30 μm to achieve the required precision on the sagitta measurement, which is made possible by a combination of an optical alignment system and track-based alignment. The optical alignment system measures the relative displacement of the inner, middle, and outer chambers and determines the spurious sagitta due to their misalignment, which is used to correct the sagitta measurement. The alignment is validated by using cosmic ray tracks and the muon tracks from special runs in which the toroid field is turned off, such that the tracks in the muon spectrometer are straight and the sagitta is zero. A track-based alignment method is also used for the alignment of the silicon tracker as well as the relative alignment of the inner detector and the muon spectrometer, both of which contribute significantly to the resolution of the muon momentum measurement [35, 43, 44].

Cathode strip chambers (CSC) are used to extend the capability to measure muons with excellent spatial resolution to the very forward region ($2 < |\eta| < 2.7$) where the rates are too high for the MDT chambers. There are 16 chambers in each endcap, alternating between 8 large and 8 small overlapping chambers to ensure full ϕ coverage. Each chamber is a multiwire chamber with segmented readouts at both cathodes, one providing the pseudorapidity precision coordinate with a pitch of approximately 5 mm, and the other the ϕ coordinate measurement. By interpolating the charge between neighbouring strips a resolution better than 60 μm is achieved in the precision coordinate [43, 45, 46].

Muon triggering is achieved by a combination of the resistive plate chamber (RPC) in the barrel region ($|\eta| < 1.05$) and thin gap chambers (TGC) in the endcap region ($1.05 < |\eta| < 2.4$). RPCs are gaseous detectors and operate by detecting

the ionisation of traversing charged particles under a strong electric fields provided by the resistive electrodes. The plates are 2 mm thick and kept at 2 mm from each other with the help of insulating spacers. The spatial resolution is on the order of a few centimeters but the excellent time resolution of about 1 ns makes it ideal for the trigger [45, 47]. TGCs operate similarly to the multiwire proportional chambers used in the CSCs, but with a smaller distance (1.4 mm) between the wire and the readout strips. Combined with the large electric field strength (3.1 kV) this achieves the time resolution and high-rate capability required for the trigger in the forward region [48].

2.7 ATLAS data and Monte Carlo simulation

ATLAS records events under a variety of conditions to work towards a number of scientific and technical goals. The most important dataset for this thesis was collected at 13 TeV centre-of-mass energy between the years 2015 and 2018.

The Monte Carlo (MC) simulations are a critical part of the ATLAS experiment and are needed for the design, implementation, and validation of reconstruction and identification algorithms, as well as for the evaluation of their performance. Additionally, the MC simulations are needed to carry out physics analyses and optimise their sensitivity.

The ATLAS MC simulation chain begins with the calculation of matrix element amplitudes and the generation of the hard-scatter process using software such as Powheg [49–51] or Sherpa [52, 53]. Parton showering, hadronisation, and the simulation of the underlying event are simulated by Pythia 8 [54]. The simulation of the interactions with the ATLAS detector, as well as its response, is implemented in the Athena framework [55, 56] and uses the Geant4 [57–59] simulation toolkit. Finally, the response is digitised and packaged in the same event data model [60] as the data.

The generation of MC simulation is a computationally intensive task and it usually takes weeks or months to generate a sufficient number of simulated events even on the dedicated computing grid. For this reason, the generation of simulated events begins before the data-taking takes place and before the pile-up profile is

known. To reconcile the differences in the recorded and simulated pile-up profile, simulated events are re-weighted to match the recorded pile-up distribution.

Chapter 3

Muons in ATLAS

If you can meet with Triumph and Disaster
And treat those two impostors just the same

— Rudyard Kipling

The ATLAS detector, described in the previous chapter, is a monument to the engineers and physicists who designed and built it. However, in order to turn its raw output of nearly one hundred million readout channels per event to physics knowledge, it requires an intermediate step: the reconstruction of physics objects.

Muons, the central physics object in this thesis, are reconstructed by combining the information from the tracker, calorimeters, and the muon spectrometer. In order to be able to conduct precision studies, the MC simulation of the detector response needs to be calibrated to match the response in the real detector. In particular, the transverse momentum of muons needs to be calibrated to the scale and resolution obtained in the data. Similarly, the reconstruction, the track-to-vertex-association (TTVA), and isolation efficiencies need to be calibrated to match those in the recorded data. These calibrations are obtained through measurements using the data and MC simulation and carry associated uncertainties.

This chapter describes the first novel contributions from the author. The first contribution is an improved method of background subtraction in the isolation efficiency measurements, which significantly reduces the systematic uncertainty for muons with $p_{\mathrm{T}} < 15$ GeV. The second is an attempt to improve muon momentum resolution.

3.1 Reconstruction

Tracks in the inner detector (ID) are reconstructed indiscriminately for all charged particles. First, raw data from the detectors is converted into space-points which form the basis of tracking. The main track finding algorithm proceeds inside-out by finding the seeds in the silicon layers. The seeds are then combined into roads and the extension to the TRT layer is probed to add hits in the outermost layer. The final collection of hits is fit to obtain the track parameters [61, 62]. Tracks with at least $p_{\mathrm{T}} > 400$ MeV and some additional requirements, detailed in Ref. [63], are then used to find the primary vertices. The reconstruction of primary vertices proceeds by first finding the vertices and then fitting them [64]. Finally, tracks are associated with vertices which allows the computation of d_0, the transverse impact parameter, which quantifies the shortest distance between the track and the beam-line, and z_0, the longitudinal impact parameter, which quantifies the

difference between the z coordinate of the associated primary vertex, and the z coordinate of the point where d_0 is computed, as shown in Figure 3.1.

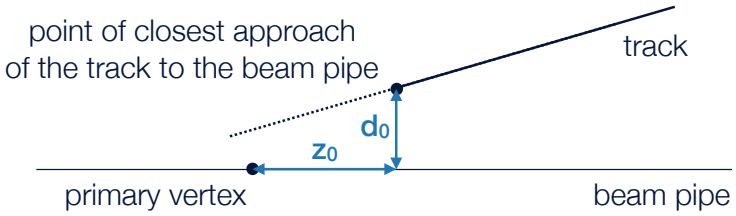


Figure 3.1: The definition of the impact parameters. The transverse impact parameter, d_0, is defined as the shortest distance between the track and the beamline. The longitudinal impact parameter, z_0, is defined as the difference in the z coordinate between the point of closest approach and the primary vertex.

The reconstruction of tracks in the muon spectrometer (MS) proceeds by first finding track segments in the MDT chambers using a Hough transform [65], and then reconstructing them by fitting a straight line to the associated hits. The segments are combined to form MS tracks starting from the middle layer and finding compatible segments in the inner and outer layers of the spectrometer. At least two segments are needed to form a track, apart from the transition region between the barrel and the endcaps ($1.0 < |\eta| < 1.4$) in which a single high-quality segment can be used to build a track. Finally, a global χ^2 fit of the track is performed and the track accepted if it satisfies the quality criteria. Otherwise, hits with a high contribution to the χ^2 value are removed and the track re-fit. Hits can also be added to the track if they are consistent with the track and the track is re-fit if these candidates are found [66].

Finally, the information from subdetectors is combined to form muon candidates. Depending on which subdetectors are used to build the candidate, four muon types are defined:

- Combined (CB) muons are reconstructed from a pair of tracks in the ID and the MS. A global re-fit is performed using hits from both subdetectors with the flexibility to add or remove hits in the MS to improve the fit quality. Most muons are reconstructed by the outside-in approach extrapolating the MS tracks to the ID, but the inside-out reconstruction is also used as a

complementary approach [66]. The advantage of the outside-in approach is the small number of muon tracks in the MS, resulting in low combinatorics compared to the busy ID environment. The disadvantage is inefficient reconstruction in the regions of poor MS coverage, which is mitigated by the complementary inside-out algorithm.

- Segment-tagged (ST) muons are formed by combining the ID track with a single segment in the MDT or CSC chambers. This muon type is intended to recover muon reconstruction efficiency for muons in the regions of MS with poor acceptance, and for low p_{T} muons which only cross a single layer of the MS [66].

- Calorimeter-tagged (CT) muons match an ID track with a calorimeter deposit consistent with a minimum-ionising particle. This muon type recovers the reconstruction efficiency in the regions not covered by the MS, for example for $|\eta| < 0.1$, where cabling and services to the ID are located, and where the support structure is located. As a consequence of using the calorimeter, this type has the lowest purity [66].

- Extrapolated (ME) muons are built from tracks in the MS that are loosely compatible with originating from the interaction point. This muon type is used to recover the reconstruction efficiency in the forward region $2.5 < |\eta| < 2.7$ not covered by the ID [66].

Finally, overlaps between different muons are resolved by order of precedence starting with CB muons, followed by ST muons, and finally the CT muons. The ME ambiguities are resolved by analysing the track fit quality and the number of hits. The requirements on the number of hits and holes[1] in the ID and MS detectors are also used to guarantee a robust measurement of p_{T} [66].

In addition, a number of quality requirements must be satisfied to differentiate between muons originating from the interaction point (prompt muons) and the muons coming from the decay of light hadrons, for example kaon and pion decay. To this end the following variables are defined for CB muons:

[1] A *hole* is an active sensor traversed by the track that contains no hits and also falls between two layers with hits assigned to the track.

- normalised χ^2 of the combined track fit.
- $\rho' = \frac{|p_{\mathrm{T}}^{\mathrm{ID}} - p_{\mathrm{T}}^{\mathrm{MS}}|}{p_{\mathrm{T}}^{\mathrm{CB}}}$, where the superscript denotes whether the p_{T} refers to the ID or MS muon candidate, or the CB muon.
- q/p significance $= \frac{\left|\left(\frac{q}{p}\right)_{\mathrm{ID}} - \left(\frac{q}{p}\right)_{\mathrm{MS}}\right|}{\sqrt{\sigma_{\mathrm{ID}}^2 + \sigma_{\mathrm{MS}}^2}}$, where the $\left(\frac{q}{p}\right)$ is the measurement of the ratio of the charge and the momentum, σ are the uncertainties on the corresponding quantities, and the subscript refers to whether the measurement comes from the ID or the MS muon candidate.

Four overlapping muon identification selections, namely Loose, Medium, Tight, and High-p_{T}, are supported by the muon performance group to serve different physics analyses and unify the evaluation of systematic uncertainties arising from the selection requirements. Tight muons maximise the purity of muons at the cost of some identification efficiency and use only CB muons with additional requirements on the track quality. Medium muons include CB and ME muon types and are the default selection that minimises the systematic uncertainties. Loose muons are designed to maximise the selection efficiency and include all four muon types, with the CT and ST types restricted to $|\eta| < 0.1$ [66].

The fraction of prompt muons that are reconstructed and identified is called reconstruction efficiency and is in general different in the data and MC simulation. Efficiencies for Loose, Medium, and Tight selections are shown in Figure 3.2 for the data and MC simulation as a function of muon pseudorapidity. In the analysis, the MC simulation is corrected to the data by applying weights, known as scale factors, to account for the differences in efficiencies between the data and MC simulation. The scale factors (SF) are defined as

$$\mathrm{SF} = \frac{\epsilon_{\mathrm{Data}}}{\epsilon_{\mathrm{MC}}}, \tag{3.1}$$

where SF is the scale factor, and ϵ is an efficiency, as measured in data or MC simulation. The measurement of the efficiencies, scale factors, and the associated uncertainties is performed using a sample of $Z \to \mu\mu$ events using the tag-and-probe method, using one of the muons to tag the event and the other to perform the identification efficiency measurement. The method measures efficiencies separately in the ID and MS systems and combines them with the assumption that they are

independent, which has been tested in MC simulation. The measurement of scale factors is described in full detail in Ref. [66].

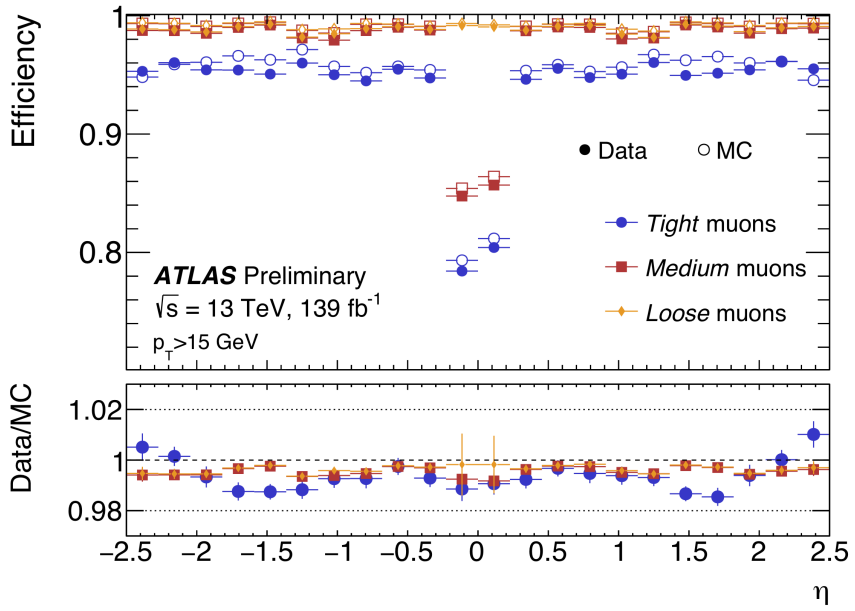


Figure 3.2: The muon reconstruction efficiency measurement for the Loose (yellow diamonds), Medium (red squares), and Tight (blue circles) selections in data (filled markers) and MC simulation (empty markers) as a function of muon pseudorapidity with $p_{\mathrm{T}} > 15$ GeV. The measurement uses $Z \to \mu\mu$ events in full Run 2 dataset of 139 fb^{-1} at $\sqrt{s} = 13$ TeV. The error bars represent a quadratic sum of statistical and systematic uncertainties. Reproduced from Ref. [67].

3.2 Momentum calibration

The response of the ATLAS detector is accurately modelled by the MC simulation. However, additional momentum calibration of MC simulation is needed to achieve per mille level precision in the muon momentum scale and percent level precision in momentum resolution [66].

The corrections are applied separately to ID and MS tracks and have a dependence on the η and ϕ of the muon track. The corrected momentum

$$p_{\mathrm{T}}^{\mathrm{Cor}} = \frac{p_{\mathrm{T}} + \sum_{n=0}^{1} s_n(\eta,\phi) \cdot p_{\mathrm{T}}^{n}}{1 + \sum_{m=0}^{2} \Delta r_m(\eta,\phi) \cdot p_{\mathrm{T}}^{m-1} \cdot g_m}, \tag{3.2}$$

where $s_n(\eta, \phi)$ are the scale corrections, and $\Delta r_m(\eta, \phi)$ are the resolution corrections, and g_m is a random variate from a normal distribution with zero mean and variance of 1, can be computed from the uncorrected transverse momentum in simulation, p_{T}, once the values of s_n and Δr_m are known [66]. Indices n take on the values of 0 and 1 which formalises the assumption that the momentum scale can be corrected by a combination of a constant offset and a term proportional to the p_{T}. Indices m take on the values 0, 1, and 2, i.e. the denominator in Eq. 3.2 assumes that the relative transverse momentum resolution can be parametrised as

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{r_0}{p_{\mathrm{T}}} \oplus r_1 \oplus r_2 \cdot p_{\mathrm{T}} \tag{3.3}$$

where the r_0 term accounts for the fluctuations of the energy loss, the r_1 term for multiple scattering, and the r_2 term for the resolution in the measurement of the sagitta due to intrinsic hit resolution and misalignment of the MS [66].

The corrected momentum of CB muons is then determined by combining the corrected momenta of ID and MS muons

$$p_{\mathrm{T}}^{\mathrm{CB,\ Cor}} = f \cdot p_{\mathrm{T}}^{\mathrm{ID,\ Cor}} + (1 - f) \cdot p_{\mathrm{T}}^{\mathrm{MS,\ Cor}}, \tag{3.4}$$

with the weight f derived for each muon individually from the equivalent equation with uncorrected momenta, thus assuming the relative contributions of the two subdetectors are unchanged by the momentum corrections [66].

The determination of s_n and Δr_m constants proceeds by an iterative fitting procedure described in detail in Ref. [66]. It extracts the constants by binned maximum-likelihood fits of the invariant mass spectra of $Z \to \mu\mu$ and $J/\psi \to \mu\mu$. The validation of the calibration using CB muons in $Z \to \mu\mu$ events is shown in Figure 3.3 as a function of the pseudorapidity of the muon with the larger p_{T} (leading muon).

3.3 Isolation

Prompt muons, produced at the interaction point or from the decay of heavy bosons, often need to be distinguished from muons coming from a secondary vertex as a result of the decay of a τ lepton, heavy quarks, or hadrons. Other processes

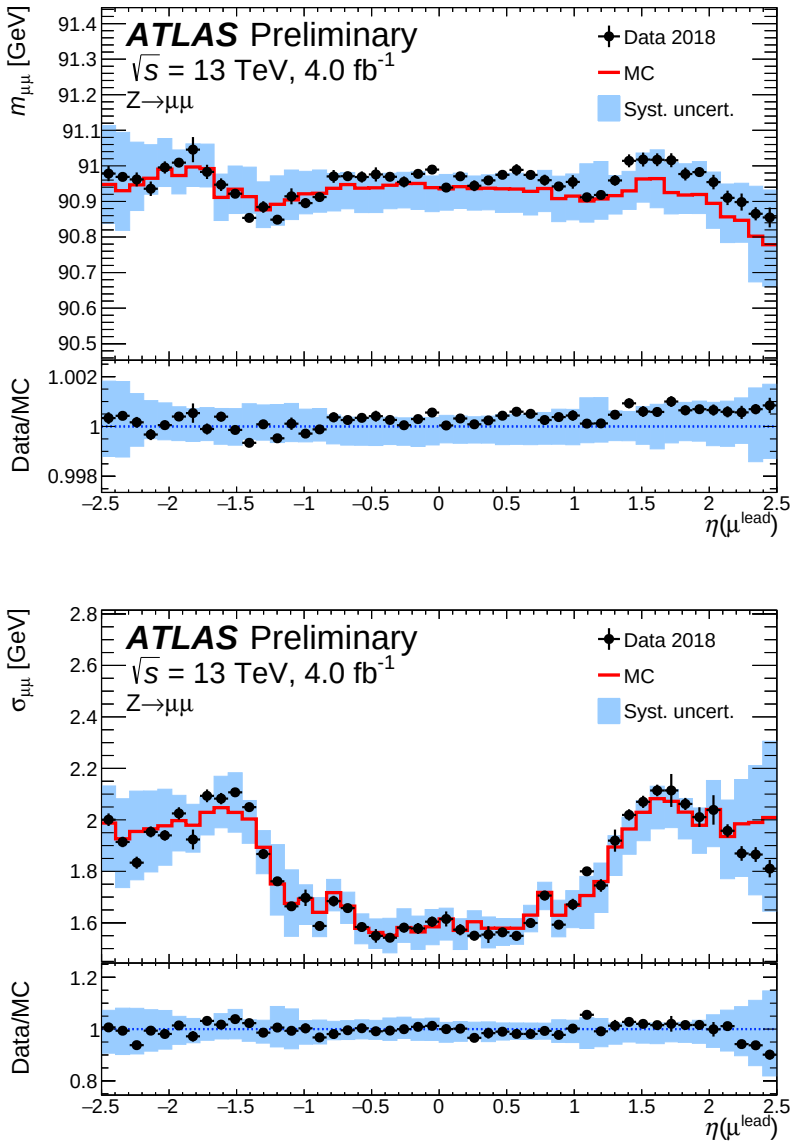


Figure 3.3: The muon momentum calibration validation compares the scale (top) and resolution (bottom) of $Z \to \mu\mu$ events using CB muons in part of 2018 data (black circles) to the corrected MC simulation (red line). The systematic uncertainty on the corrected MC simulation is shown as blue bands. The bottom panel in each figure shows the ratio between the data and MC simulation. Reproduced from Ref. [68].

can also fake a muon signature in the detector, for example, a jet punching through the hadron calorimeter into the muon spectrometer could be reconstructed as a muon. The key to the identification of non-fake prompt muons is that they are produced in isolation from other particles, unlike fake or non-prompt muons which are usually surrounded by other particles. Two sets of variables, one based on the tracker information and the other on calorimeter information, are used to identify prompt muons.

The track-based isolation variable, $p_{\mathrm{T}}^{\mathrm{varcone},R}$, is defined as the scalar sum of the p_{T} of non-muon tracks inside a cone of radius $\Delta R = \min(R, 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu})$, where $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, $\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angle between the muon and the track, and p_{T}^{μ} is the transverse momentum of the muon. The tracks used in the sum have $p_{\mathrm{T}} > 1$ GeV, $|\eta| < 2.5$, a longitudinal impact parameter smaller than 3 mm, and satisfy a Loose track quality. The cone size is chosen to vary with p_{T}^{μ} to follow the collimation of tracks in highly boosted systems [66].

The calorimeter-based isolation variable relies on the notion of a topological cluster, built by seeding the cluster with a calorimeter cell with signal-to-noise (S/N) greater than four, adding the neighbouring cells with S/N > 2, and finally adding all the surrounding cells [69]. The variable $E_{\mathrm{T}}^{\mathrm{topocone20}}$ is defined as the sum of the transverse energy of topological clusters in a cone of size $\Delta R = 0.2$, after subtracting for the energy deposited by the muon itself and the estimate of the energy coming from pile-up related effects. The pile-up contributions are estimated using the jet area technique [70], which multiplies the ambient minimum-bias energy density, computed on event-by-event basis, with the area of the isolation cone. [66]. In order to deal with the higher pile-up rates an additional variable, E^{neflow20}, is constructed. It is computed in the same way as the $E_{\mathrm{T}}^{\mathrm{topocone20}}$ variable, apart from the weighting of the clusters which is unity for clusters matched to the charged particle tracks, while the others are weighted according to the expected fraction of neutral particles in the isolation cone.

The isolation selections are defined as cuts on these variables which may or may not depend on the muon kinematics. Various selections have been supported in the past few years to target the needs of different analyses and to adapt to the higher pile-up conditions. The selections differ in signal efficiency and background

rejection throughout the η and p_{T} range. A few selections are defined in Table 3.1.

Table 3.1: Isolation selections definitions

Isolation Selection	Definition
GradientLoose	Cuts such that $\epsilon(25\mathrm{GeV}) = 95\%$, $\epsilon(60\mathrm{GeV}) = 99\%$
FixedCutLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$ $E^{\mathrm{etcone20}}/p_{\mathrm{T}}^{\mu} < 0.30$
FixedCutTight	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.06$ $E^{\mathrm{etcone20}}/p_{\mathrm{T}}^{\mu} < 0.06$
FixedCutHighMuLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$ ($p_{\mathrm{T}}^{\mu} < 50$ GeV) $p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$ ($p_{\mathrm{T}}^{\mu} > 50$ GeV) $E^{\mathrm{etcone20}}/p_{\mathrm{T}}^{\mu} < 0.30$ where track isolation variables have been computed with a tighter track-to-vertex requirement.
FixedCutPflowLoose	$\max(p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu},\ p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{\mu}) +$ $0.4 \cdot E^{\mathrm{neflow20}}/p_{\mathrm{T}}^{\mu} < 0.16$ where track isolation variables have been computed with a tighter track-to-vertex requirement.

For each of the isolation selections, a measurement of efficiencies in data and MC simulation is performed, as well as the evaluation of scale factors, which vary as a function of muon p_{T}. The measurement of the efficiencies is performed on $Z \to \mu\mu$ events in data and MC simulation using the tag-and-probe method [66]. The tag-and-probe method selects events with a pair of muons with the invariant mass within 10 GeV of the Z boson mass, where the tag muon is required to have triggered the event, and the probe muon is used for the efficiency measurements. More precisely, the tag muon is required to satisfy $p_{\mathrm{T}} > 24$ GeV, pass Loose isolation selection, and must be trigger-matched to the muon firing the trigger. Both tag and probe muons are required to be CB muons, pass the Medium quality requirements, and pass requirements on transverse and longitudinal impact parameters or their significances $d_0/\sigma_{d_0} < 3.0$, $|z_0| < 10$ mm and $|z_0 \sin\theta| < 0.5$ mm. They are also required to be well separated from each other ($R_{\mu,\mu} > 0.3$), and the probe should be separated from jets ($R_{j,\mu} > 0.4$).

The MC simulation contains a pure sample of $Z \to \mu\mu$ events, but the recorded data consists of an unknown mixed stream of underlying processes, including both the $Z \to \mu\mu$ decays as well events with fake or non-prompt muons, collectively referred to as the background. In order to measure the efficiency of prompt muons, the background must be subtracted for the efficiency measurement in the data. An earlier method of background subtraction relied on the assumption that the number of background events in the data is the same in the same charge (SC) as in the opposite charge (OC) phase space:

$$\epsilon_{Z\to\mu\mu} = \frac{N^{\mathrm{OC}}_{\mathrm{match}} - T \cdot N^{\mathrm{SC}}_{\mathrm{match}}}{N^{\mathrm{OC}}_{\mathrm{probe}} - T \cdot N^{\mathrm{SC}}_{\mathrm{probe}}}, \tag{3.5}$$

where T is the transfer factor, N_{probe} is the number of probe muons and N_{match} is the number of probe muons passing the isolation selection in either SC or OC phase space, as indicated by the superscript. The leading systematic uncertainty at low muon p_{T} comes from the background subtraction, namely the assumption $T = 1$, which is conservatively varied to 0.5 and 2.0. Additional systematic uncertainties on the scale factor measurement are computed by varying the selection requirements and adding individual uncertainties in quadrature, and adding a fixed uncertainty due to variation in η as the scale factors are provided only as a function of p_{T}. The systematic variations are summarised in Table 3.2 below. The scale

Table 3.2: Systematic uncertainties in the measurement of the isolation selection scale factors. Background estimation uncertainty is replaced by Eq. 3.8 in the novel background subtraction method described in section 3.3.1.

Description	Variation
Background estimation	$T = 0.5, T = 2.0$
Dependence on η	Flat 0.2% uncertainty
$m_{\mu\mu}$ window	$m_Z \pm 5, 20$ GeV
Probe muon quality	Loose, Tight
Tag muon isolation	Every other isolation selection
Probe muon to jet separation	$\Delta R_{j,\mu} > 0.3, 0.5$
Probe to tag muon separation	$\Delta R_{\mu,\mu} > 0.2, 0.5$

factors are evaluated as a function of probe muon p_{T}, with uncertainties computed

in each bin separately. At low p_{T} (< 15 GeV), the leading uncertainty is the background subtraction, which prompted the development of a new background subtraction method, described in the next section.

3.3.1 Isolation efficiency background subtraction

The core idea of the novel background subtraction method is to reduce the assumption that the number of SC events is the same as the number of OC events, to the assumption that the shape of the invariant mass spectrum for background events is the same in the SC and OC events. The SC region is used to determine the shape of the background events, while the normalisation is determined from the template fit to OC data.

The shape assumption is more accurate than the normalisation assumption because the production of pions and kaons, which are a source of background muons, is charge asymmetric in terms of normalisation, but has similar kinematic properties.

The efficiency measured in the data is a sum of efficiencies of the components, weighted by the component fractions,

$$\epsilon_{\mathrm{measured}} = \mathrm{f}_{Z\to\mu\mu} \cdot \epsilon_{Z\to\mu\mu} + \mathrm{f}_{\mathrm{EWK}} \cdot \epsilon_{\mathrm{EWK}} + \mathrm{f}_{\mathrm{QCD}} \cdot \epsilon_{\mathrm{QCD}}, \tag{3.6}$$

where f represents the fraction of the component, and ϵ the efficiency of the component. The subscript QCD refers to the background component due to muons in jets, and EWK refers to the background from electroweak processes[2]. The efficiency of $Z \to \mu\mu$ events in the data can therefore be expressed as

$$\epsilon_{Z\to\mu\mu} = \frac{\epsilon_{\mathrm{measured}} - \mathrm{f}_{\mathrm{EWK}} \cdot \epsilon_{\mathrm{EWK}} - \mathrm{f}_{\mathrm{QCD}} \cdot \epsilon_{\mathrm{QCD}}}{1 - \mathrm{f}_{\mathrm{EWK}} - \mathrm{f}_{\mathrm{QCD}}}, \tag{3.7}$$

where $\mathrm{f}_{Z\to\mu\mu} = 1 - \mathrm{f}_{\mathrm{EWK}} - \mathrm{f}_{\mathrm{QCD}}$ is used. The efficiency ϵ_{EWK} is measured in the MC simulation, while the efficiency ϵ_{QCD} is measured in the SC data.

The fractions $\mathrm{f}_{\mathrm{EWK}}$ and $\mathrm{f}_{\mathrm{QCD}}$ are determined via a template fit to the OC data, shown in Figure 3.4, in each of the $p_{\mathrm{T}}^{\mathrm{probe}}$ bins. The shapes of the templates are allowed to vary in the fit within statistical uncertainties. While the fit is performed

[2]The electroweak processes are: $W^{\pm} \to \mu^{\pm}\nu$, $Z \to \tau\tau$, $WW \to \ell\nu\ell\nu$, $WZ \to \ell\nu\ell\ell$, $WZ \to qq\ell\ell$, $ZZ \to \ell\ell\ell\ell$, $ZZ \to \nu\nu\ell\ell$, $ZZ \to qq\ell\ell$, $t\bar{t}$. The MC samples corresponding to these processes are generated with Powheg+Pythia8.

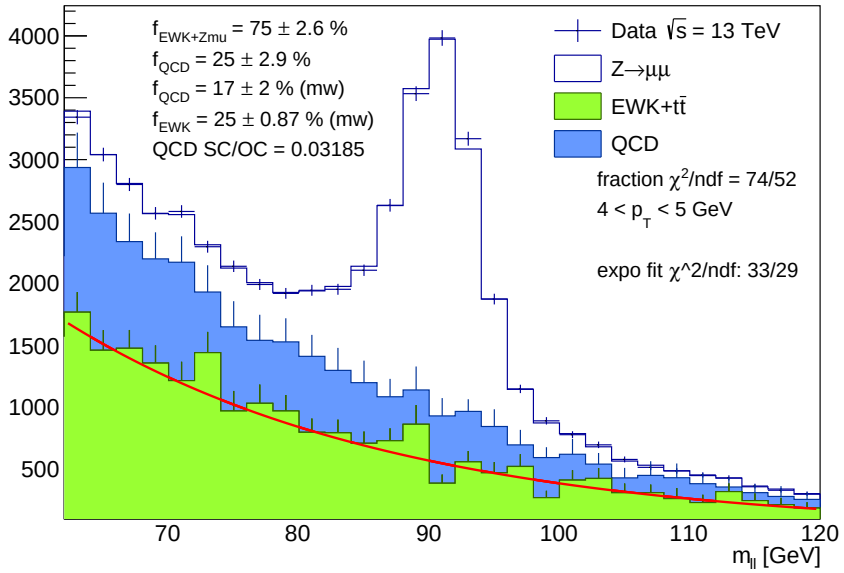


Figure 3.4: The muon isolation background subtraction template fit in OC data to determine the f_{EWK} and f_{QCD} fractions in the $4 < p_T^{probe} < 5$ GeV bin. The data is shown as blue points, QCD template, taken from SC data, is shown in blue. The other template is determined from MC and is a combination of the $Z \to \mu\mu$ events, shown in white, and the EWK processes, shown in green. The error bars show statistical uncertainties. The maximum-likelihood fit with the exponential function to the EWK component is shown in red. The reduced χ^2 values are shown on the figure for both the template fit and the subsequent exponential function fit. Furthermore, fraction values are shown both for the full invariant mass range, over which the fit is performed, and in the ±10 GeV mass window (mw) around the Z peak.

over a wide invariant mass range (±30 GeV) around the Z peak, the fractions need to be determined in a narrower mass range in which the efficiency measurement is made (±10 GeV). This is done by integrating the fitted templates in the desired range, using linear interpolation in the edge bins if necessary. For the EWK template, in order to decrease the statistical fluctuations, an exponential fit to the template is integrated instead. When varying the mass range of the measurement, as shown in Table 3.2, the fractions are recomputed in the adjusted range.

The current method no longer assumes that the normalisation of background events is the same in SC and OC data. Instead, the assumption is reduced to the shape of the background to be the same, with normalisation determined via

the maximum-likelihood fit. In order to estimate the associated systematic uncertainty the scale factors are re-computed using different QCD templates. Recall, the nominal QCD template is taken from the SC data, where the number of genuine $Z \to \mu\mu$ events is negligible. Similarly, two other QCD templates are constructed in the $Z \to \mu\mu$ depleted regions, where the requirement on the impact parameter significance ($|d_0/\sigma_{d_0}|$) is inverted in both the SC and OC data. The region definitions are shown in Table 3.3. The region C definition becomes tighter above

Table 3.3: The region definitions from which the QCD template shape is taken for the isolation background subtraction fit.

Region	Charge	$p_{\mathrm{T}}^{\mathrm{probe}}$ range	$\lvert d_0/\sigma_{d_0}\rvert_{\mathrm{tag}}$	$\lvert d_0/\sigma_{d_0}\rvert_{\mathrm{probe}}$
Region B	SC	full	< 3.0	< 3.0
Region C	OC	< 10 GeV	> 3.0	> 2.0
		> 10 GeV	> 5.0	> 5.0
Region D	SC	full	> 3.0	> 2.0

$p_{\mathrm{T}}^{\mathrm{probe}} > 10$ GeV in order to reduce the number of $Z \to \mu\mu$ events to negligible levels. The systematic uncertainty is taken to be the largest variation in efficiency compared to the nominal efficiency obtained by taking the QCD template from region B. Denoting the $Z \to \mu\mu$ efficiency computed by taking the QCD template from region X with ϵ_X, the systematic uncertainty associated with the novel method of background subtraction can be written as

$$\text{bkg. sub. syst. unc.} = \max_{X \in \{C,D\}} \{|\epsilon_B - \epsilon_X|\}. \tag{3.8}$$

and is computed in each $p_{\mathrm{T}}^{\mathrm{probe}}$ bin separately. The effect of subtraction is shown in Figure 3.5. The efficiencies and scale factors computed with this method, along with full systematic uncertainties, are shown in Figure 3.6 for the GradientLoose isolation selection. Overall, the method greatly reduces the systematic uncertainty from background subtraction in the determination of isolation scale factors. At low p_{T}, where background subtraction used to be the leading uncertainty, the overall systematic uncertainty on the scale factors is reduced by up to 75%. This enabled greater sensitivity in searches for electroweak SUSY scenarios with compressed

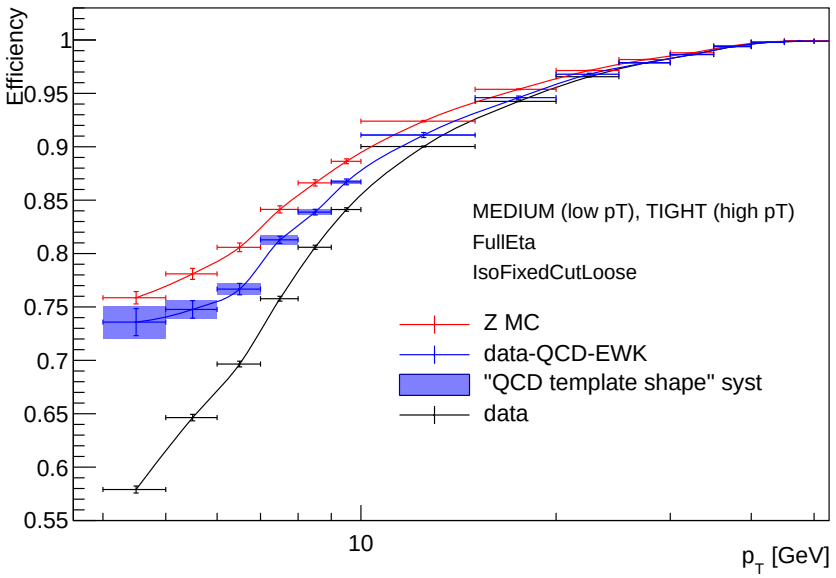


Figure 3.5: The muon isolation efficiencies for the FixedCutLoose isolation selection as a function of p_{T}. The MC simulation efficiency is shown in red, mixed efficiency measured in data directly is shown in black, and the $Z \to \mu\mu$ efficiency in data, obtained via the background subtraction, is shown in blue. The error bars represent statistical uncertainties while the systematic uncertainty arising from the use of different templates is shown as a blue band.

mass spectra [71] and more precise studies of the low mass Drell-Yan double differential cross section [72].

3.3.2 Other isolation studies

The muon performance was studied during the collection of the 2017 dataset in order to monitor for any potential problems. No issues were found in muon isolation [2].

Furthermore, a few runs in the 2017 dataset were recorded at a particularly large number of pile-up interactions, as shown in Figure 2.1. These runs were used to perform a study on isolation performance at high pile-up in preparation for the future runs [3]. Figure 3.7 shows the efficiency of the FixedCutTight selection as a function of muon p_{T}, recorded in 2017 data, separately for events with more and less than 50 additional interaction vertices due to pile-up. The

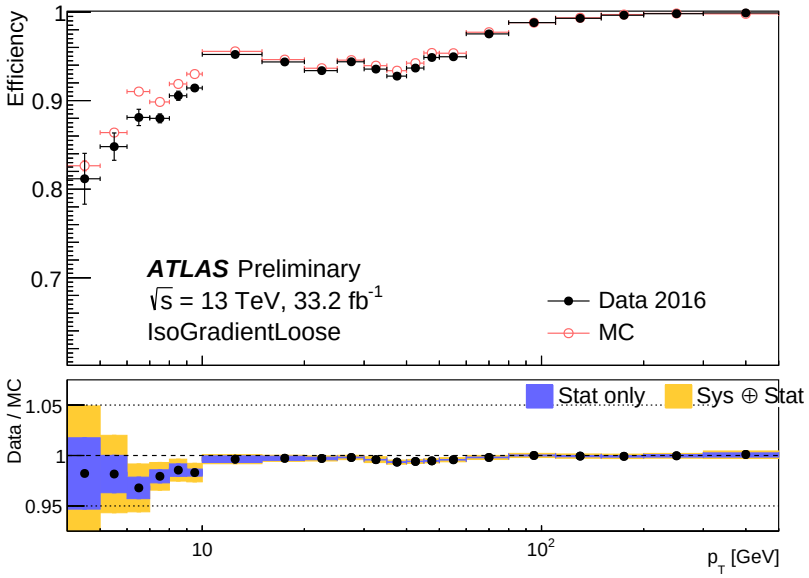


Figure 3.6: The muon isolation efficiencies (top panel) and scale factors (bottom panel) for the GradientLoose isolation selection as a function of p_{T} in 2016 data (black filled circles) and MC simulation (empty red circles). The error bars represent statistical uncertainties, error bands on the scale factors represent statistical uncertainties (purple) and statistical and systematic uncertainties added in quadrature (yellow).

isolation efficiency for the FixedCutTight selection drops by more than 10% for muons with $10 < p_{\mathrm{T}} < 20$ GeV. This result prompted the development of new isolation selections designed to be robust against high pile-up conditions. These new isolation selections, utilising either a better association of tracks to the primary interaction vertex (FixedCutHighMuLoose) or by matching the tracks to clusters in the calorimeters (FixedCutPflowLoose), are shown in Figure 3.8 as a function of the number of pile-up interactions. The new selections were shown to be more robust in high pile-up conditions [4].

3.4 VADER4μ

The search for the decay of the Higgs boson to a pair of muons, described in detail in the next chapter, proceeds as a search for a peak in the opposite sign

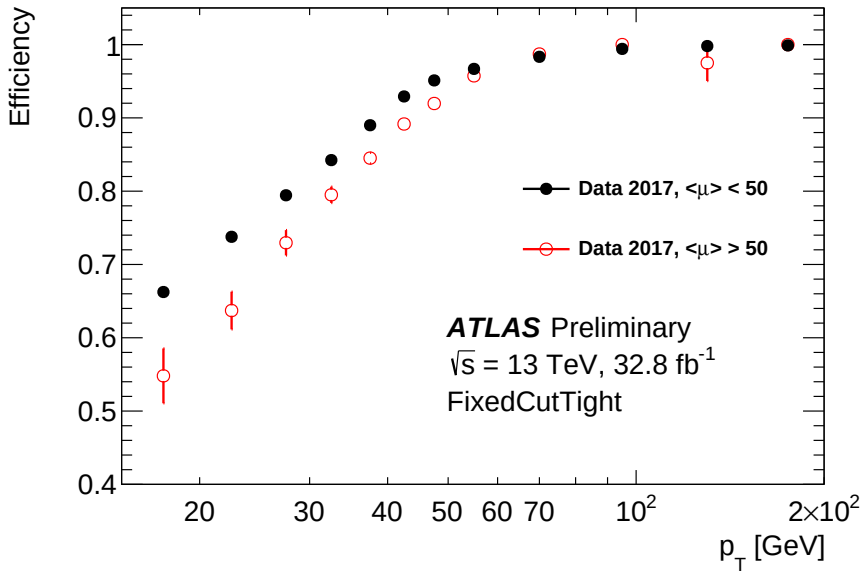


Figure 3.7: The efficiency for the FixedCutTight muon isolation selection as a function of p_{T} in data collected in 2017 for events with the number of interaction vertices smaller than (filled black circles) and larger than (empty red circles) 50. No background subtraction is used, and error bars represent the statistical uncertainties.

dimuon invariant mass spectrum on top of the smooth Drell-Yan background. According to the Standard Model, the width of the Higgs boson is much narrower than the experimental resolution of the ATLAS detector. The sensitivity of the search is therefore limited by experimental muon resolution, suggesting that an understanding of the muon resolution could potentially be exploited to improve the sensitivity of the search. This prompted the study to understand the **V**ariables **A**ffecting **DE**tector **R**esolution **for mu**ons (VADER4μ). The studies in this section show that some improvement is possible, but ultimately the results were not used in the $H \to \mu\mu$ analysis because the achievable gain is too small to justify the additional complexity.

There are three main motivations for the study in the context of improving the sensitivity of the $H \to \mu\mu$ analysis:

1. The muon p_{T} resolution can be estimated for each individual muon based on its kinematics and reconstruction quality metrics. The estimated resolution

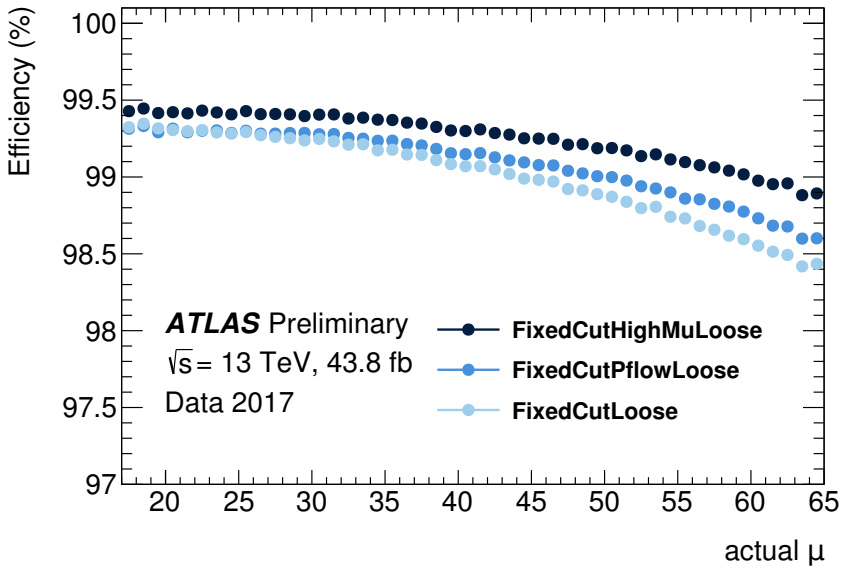


Figure 3.8: The efficiency for FixedCutLoose (light blue), FixedCutPflowLoose (blue), and FixedCutHighMuLoose (dark blue) isolation selections in data collected in 2017 as a function of p_{T}. The error bars represent statistical uncertainties that are too small to be seen. No background subtraction is used.

of individual muons can be propagated to estimate the resolution of the invariant mass measurement of the muon pair. Finally, the estimate can be used to split events in resolution categories and the sensitivity of the analysis may be improved. This can be applied directly to the CB muons.

2. The uncertainties on the p_{T} of ID and MS tracks are known to be poorly modelled, which is the reason a combined re-fit is used instead of the statistical combination of the ID and MS tracks. If the uncertainties on the ID and MS p_{T} can be estimated more accurately, a statistical combination could result in a better measurement of the CB momentum compared to the combined re-fit.

3. Somewhat more speculatively, the muon reconstruction algorithms could contain bugs, biases, or inefficiencies. If they exist, it may be possible to derive a correction to the muon p_{T} to improve the dimuon mass resolution directly. Ultimately, any potential gain would come from improving the resolution

in data, meaning that the momentum corrections for MC simulation would have to be re-derived. For this reason, the corrections would have to be derived separately for the ID and MS tracks, along with their uncertainties, and combined statistically.

Two datasets are prepared to facilitate the study. The "single muon" dataset consists of reconstructed prompt muons from the MC simulation of a variety of physics processes, excluding $Z \to \mu\mu$ to avoid overlap with the dimuon dataset, with truth-level p_{T} information available. The processes used are $W^{\pm} \to \mu^{\pm}\nu$, $Z \to \tau\tau$, $WW \to \ell\nu\ell\nu$, $WZ \to \ell\nu\ell\ell$, $WZ \to qq\ell\ell$, $ZZ \to \ell\ell\ell\ell$, $ZZ \to \nu\nu\ell\ell$, $ZZ \to qq\ell\ell$. It is preprocessed by removing some of the muons such that the resulting p_{T} distribution is flat for $p_{\mathrm{T}} < 200$ GeV. The "dimuon" dataset consists of data and MC simulation of the $Z \to \mu\mu$ process and does not contain any truth-level information.

To test the third idea a correction to the $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ID}}$ needs to be computed as a function of kinematic variables. For this purpose a gradient boosting regression (GBR) model is used. It consists of a series of decision trees, each evaluated on the input, and produces an output by performing a weighted sum of the outputs of individual decision trees. The decision trees and the weights used to combine their outputs are determined through an iterative optimisation of the loss function [73–76]. The complexity of the model depends on a number of choices, also known as the hyperparameters, which affect the capacity of the model as well as its ability to generalise. Examples of hyperparameters are `n_estimators` (the number of decision trees that make up the combined model), `max_depth` (the maximum depth of each decision tree), and the `learning_rate` (affects the weights combining individual models). Hyperparameters are fixed during the training and can not be directly optimised. However, it is good practice to "tune" the hyperparameters, i.e. try different combinations, evaluate them on an independent dataset, and pick the ones with the best performance. For this reason, the dataset is usually split in three parts: the training dataset, used to train the models, the validation dataset, used to evaluate the choices of hyperparameters, and the test dataset, used to obtain an unbiased estimate of the model performance.

The single muon dataset is used to train a gradient boosting regression model, implemented in `scikit-learn` [77], to predict the correction to transverse momentum, $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{ID}}$, from the following variables:

- $p_{\mathrm{T}}^{\mathrm{ID}}$, the transverse momentum of the muon as measured in the ID,
- η and ϕ of the muon,
- d_0/σ_{d_0}, the transverse impact parameter significance,
- z_0, the longitudinal impact parameter,
- χ^2/N_{dof}, the reduced χ^2 value of the fit,
- SNS (scattering curvature significance), SCS (scattering neighbour significance), two variables designed to capture the changes in track direction across the detector layers in order to detect in-flight decays.

L1 loss function is used as the L2 loss was found to give biased results due to asymmetric residuals. The same method is used to predict a correction for the MS muons.

A trained model is then used to provide corrections to the measured $p_{\mathrm{T}}^{\mathrm{ID}}$. The distribution of residuals before and after the correction is shown in Figure 3.9 for the training and test parts of the single muon dataset for muons with $1.12 < |\eta| < 1.25$. The reduction in the standard deviation of the residual distribution is 17% for the test set, compared to the uncorrected distribution. Similar improvements are observed for other η slices. A simple grid search, shown in Table 3.4, is then performed to find the best hyperparameter settings. For each set of settings in the search space a model is trained on the single muon dataset and applied to muons in the MC simulation subset of the dimuon dataset. The full-width half-maximum (FWHM) of the Z boson peak is chosen as the figure of merit. The chosen values of the hyperparameters are shown in bold font in Table 3.4. Finally, the model is applied to the data subset of the dimuon dataset. The same procedure is applied to the MS tracks, as well as to the CB tracks as a sanity check. The results are shown in Figure 3.10 for ID tracks and show an improvement of 7.5% in the FWHM. However, no improvement is observed for CB tracks. Upon

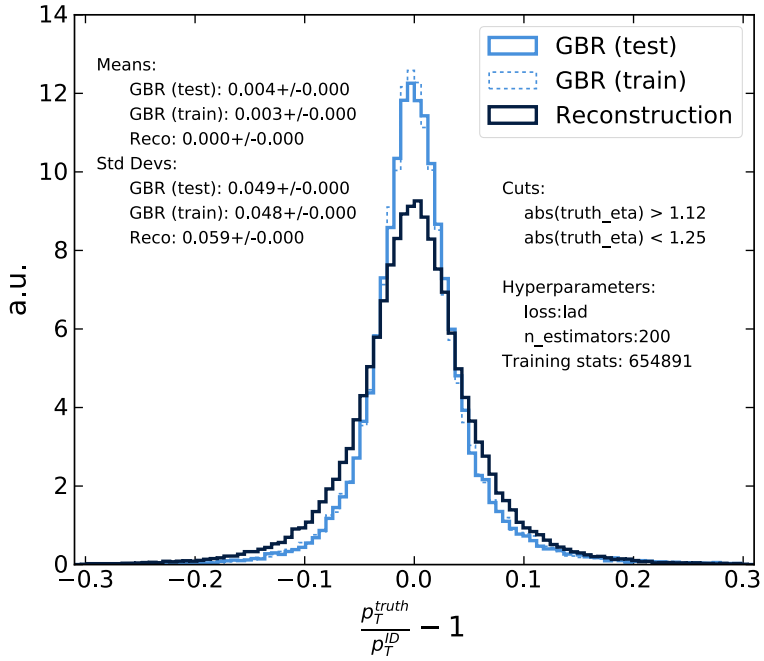


Figure 3.9: The distribution of residuals for (un)corrected measurements of momentum is shown in (dark) blue for the training (dotted) and test (solid) parts of the single muon dataset with $1.12 < |\eta| < 1.25$.

further study it turns out that the improvement for ID muons comes from the SNS and SCS variables, which contain information about the CB tracks, and that no improvement is achievable for the CB tracks. This shows that if any biases are present in the muon reconstruction algorithms their effect on performance is minimal.

To test the first two ideas, a model for the prediction of the uncertainty on the p_{T} measurement is constructed. An estimate of the uncertainty of the sagitta measurement (equivalent to q/p measurement) exists in the ATLAS event model, but it is poorly modelled. More precisely, the pull distribution approximately follows the normal distribution with zero mean and width α:

$$\frac{\left(\frac{q}{p}\right)_{\mathrm{reco}} - \left(\frac{q}{p}\right)_{\mathrm{truth}}}{\sigma_{\left(\frac{q}{p}\right)}} \sim \mathcal{N}(0, \alpha), \tag{3.9}$$

where q is the charge, p is the momentum, and $\sigma_{\left(\frac{q}{p}\right)}$ is the provided estimate of the uncertainty. If the provided estimate was accurate the value of α would be

Table 3.4: VADER4μ hyperparameter search space with selected settings in bold text. Other parameters are set to default values.

Variable	Set of values
`n_estimators`	{100, 200, 400, **800**}
`max_depth`	{3, 4, 5, **6**}
`learning_rate`	{0.001, 0.003, 0.01, **0.03**, 0.1}

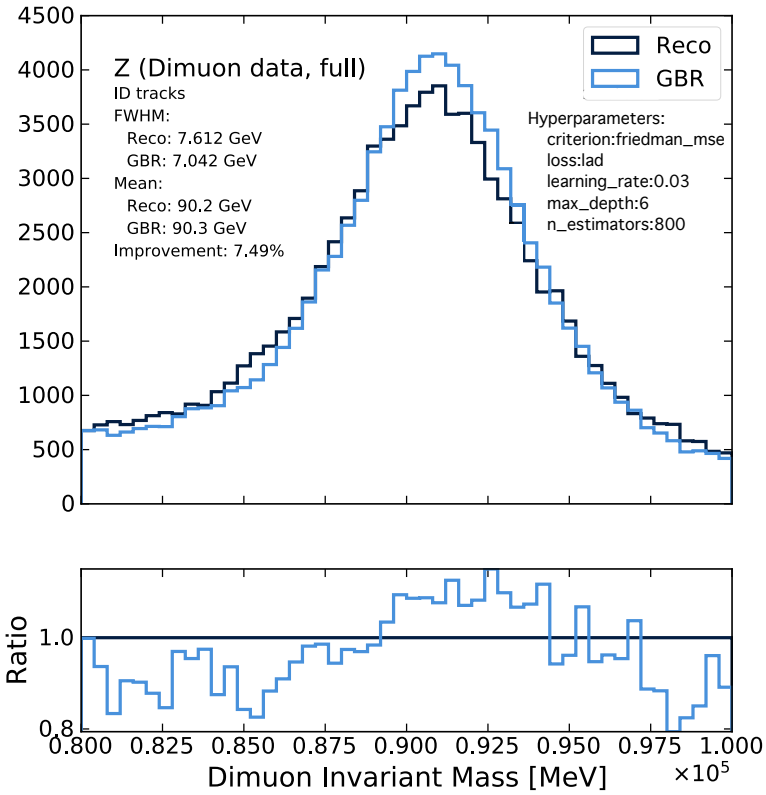


Figure 3.10: The validation of momentum corrections on the MC simulation of the $Z \to \mu\mu$ peak. The invariant mass spectrum computed using the (un)corrected values of muon momenta is shown in (dark) blue for the ID muon tracks. The bottom panel shows the ratio of the two distributions. The FWHM values and the means, computed in the 80 – 90 GeV range, are shown on the left along with the improvement in the FWHM. The hyperparameters are shown on the right.

unity. However, the uncertainties are generally too optimistic (too small), making α larger than unity. Moreover, the amount of underestimation varies with the

kinematics of the muon so a simple scaling is insufficient. The goal, therefore, is to estimate the pull on muon-by-muon basis and provide a correction to the estimate of the uncertainty. This is done using a gradient boosting regression model using the same datasets as for the first model described in the text above.

However, the input variables and the loss function are different. In order to estimate α, the 84^{th} [3] quantile of the distribution in Eq. 3.9 is estimated. The quantile loss function, the generalisation of the least absolute deviation, is used [78]

$$\mathcal{L} = \sum_i \max\left[q(\hat{y}_i - y_i), (q-1)(\hat{y}_i - y_i)\right], \tag{3.10}$$

where the sum is over the training set examples, q is the predicted quantile, y_i is the target value of example i, and $\hat{y}_i$ is the prediction for example i. The input variables are restricted to p_{T}^{CB}, η ϕ, and reduced χ^2.

The (un)corrected pull distribution is shown in Figure 3.11 for muons with $0.2 < |\eta| < 0.3$. It can be seen that the corrected pull distribution, shown in red, is much closer to the unit normal distribution, shown as a dashed line, than the uncorrected pull distribution, shown in black. This shows that the estimate of the uncertainty has been improved. Other pseudorapidity slices show similar improvement.

The uncertainty on the invariant mass of the muon pair can be computed from the uncertainties on the individual muon measurements of q/p. Assuming that the uncertainties on the measurements of η and ϕ are negligible, the relative uncertainty on the invariant mass is

$$\sigma_{m_{\mu\mu}} = \frac{1}{2} m_{\mu\mu} \left[\sigma_{\mu_1}^2 + \sigma_{\mu_2}^2\right]^{\frac{1}{2}}, \tag{3.11}$$

where σ_{μ_i} is the uncertainty on $1/p_{\text{T}}$ of muon i.

The uncertainty on the invariant mass is computed for both the corrected and uncorrected estimates of uncertainties on q/p of individual muons. The distributions of the pulls for the measurement of the invariant mass are shown in Figure 3.12 using both uncertainties along with the unit normal distribution. The pull distribution using the corrected uncertainties is closer to the unit normal than the distribution using uncorrected uncertainties. The hyperparameter search over the

[3] $84\% = 68\% + 16\%$ is obtained by summing the integral of the normal distribution within one standard deviation of the mean (68%) with the integral under its left tail (16%).

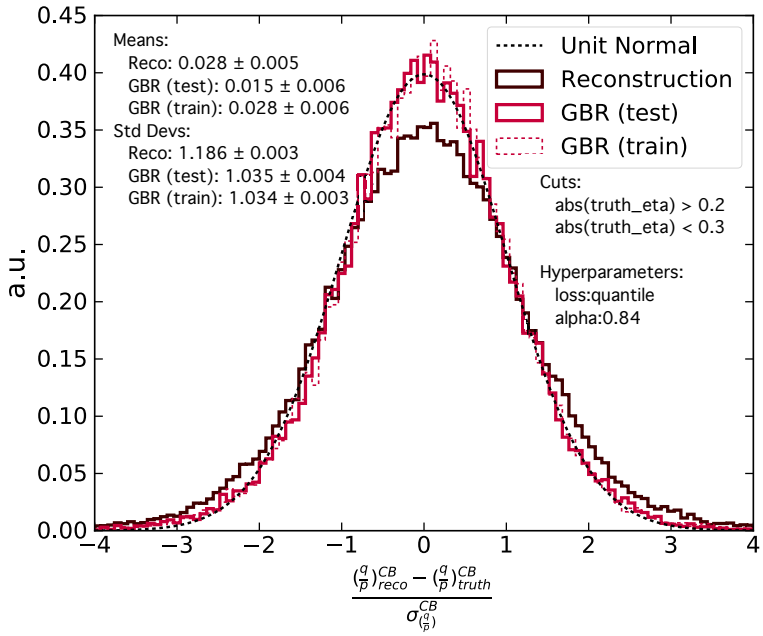


Figure 3.11: The pull distribution of q/p measurement for muons with $0.2 < |\eta| < 0.3$. the pulls computed with the uncorrected estimate of the uncertainty are shown as black solid lines. The unit normal distribution is shown as a black dashed line. The corrected pulls are shown separately for the test and training parts of the single muon datasets as solid and dashed red lines respectively, and labelled as GBR. The means and standard deviations of the distributions are shown on the left along with the associated statistical uncertainties.

same grid space as shown in Table 3.4 is performed by using the Kolmogorov-Smirnov (KS) statistic [79] as a way to quantify the performance of the hyperparameter choice. The KS statistic is the supremum of the set of distances between an empirical cumulative distribution and a reference cumulative distribution. In this case, the unit normal distribution is the reference distribution and the pull distribution, computed with corrected uncertainties, is the empirical distribution. Figure 3.13 shows the cumulative distributions for the unit normal distribution and for corrected and uncorrected cumulative pull distributions from Figure 3.12. It also visualises the differences between the empirical and reference distributions, clearly showing that the corrected uncertainty values result in a smaller value of the KS statistic, meaning that the pull distribution is closer to the unit normal distribution.

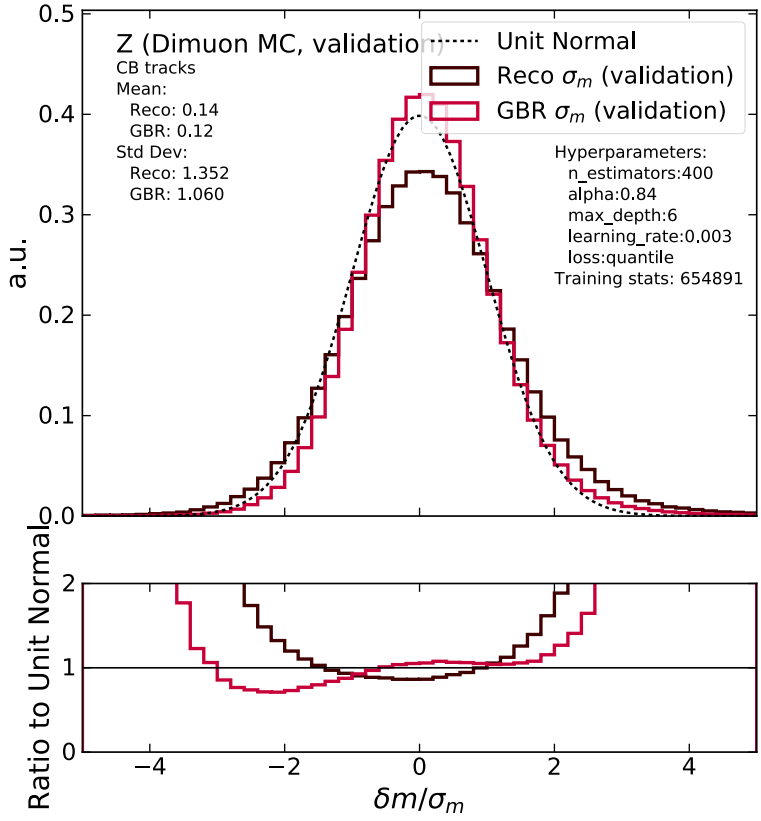


Figure 3.12: The pull distribution for the measurement of the dimuon invariant mass on simulated $Z \to \mu\mu$ events and their values divided by the unit normal distribution. The pulls using the (un)corrected uncertainties are shown in the top panel as (black) red lines, while the unit normal is shown as a dashed black line. The bottom panel shows their values divided by the unit normal distribution. The means and standard deviations are shown for both distributions on the left, while the hyperparameters are shown on the right.

The p_{T} uncertainty corrections with the chosen hyperparameters, shown on the left of Figure 3.12, are then applied to events passing $H \to \mu\mu$ selection described in Ref. [5]. For each event, the corrected uncertainty on the invariant mass measurement is computed and the distribution shown in Figure 3.14 for each of the cut-based categories described in detail in Ref. [5]. The non-central categories, in which at least one muon satisfies $|\eta| > 1.0$, have a particularly large spread in uncertainties. This means that the categories contain events with different resolutions. A split in the subcategories based on the predicted resolution

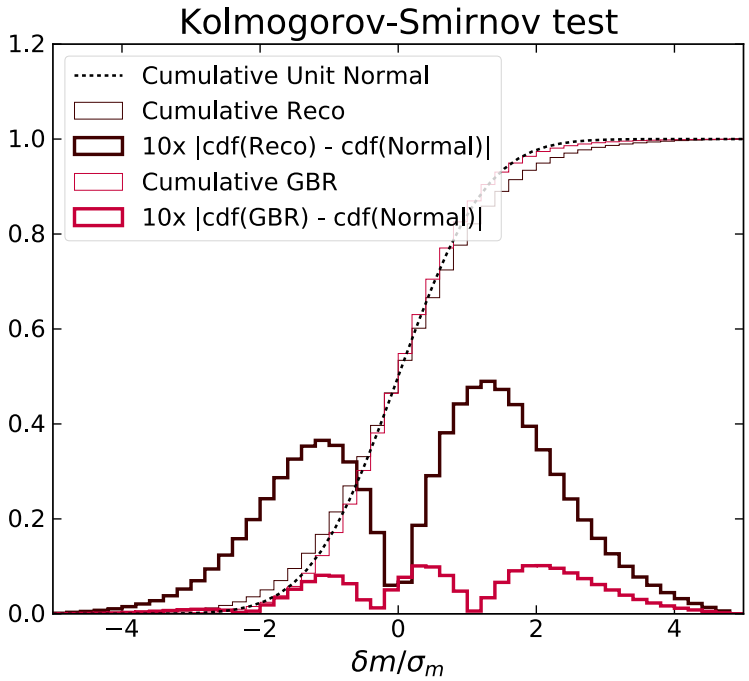


Figure 3.13: The Kolmogorov-Smirnov test for the pull distribution of the invariant mass measurements. The unit normal distribution is shown as a dashed black line, cumulative pull distributions with the (un)corrected uncertainties are shown as thin solid (black) red lines. The absolute value of the difference between the empirical and reference distributions, multiplied by a factor of 10 for clarity, is shown as thick solid (black) red lines. The KS statistic is the supremum of this difference.

is expected to improve the sensitivity of the analysis. Unfortunately the sensitivity improvement is found to be limited to under 2%, which was decided to be too small to justify the increase in the complexity of the analysis.

Similarly, improving the p_{T} resolution by a statistical combination of the ID and MS momentum measurements with a more accurate estimate of uncertainties on individual p_{T} measurements resulted only in a small improvement in dimuon invariant mass resolution. This, too, was too small to justify an additional increase in the complexity of the analysis.

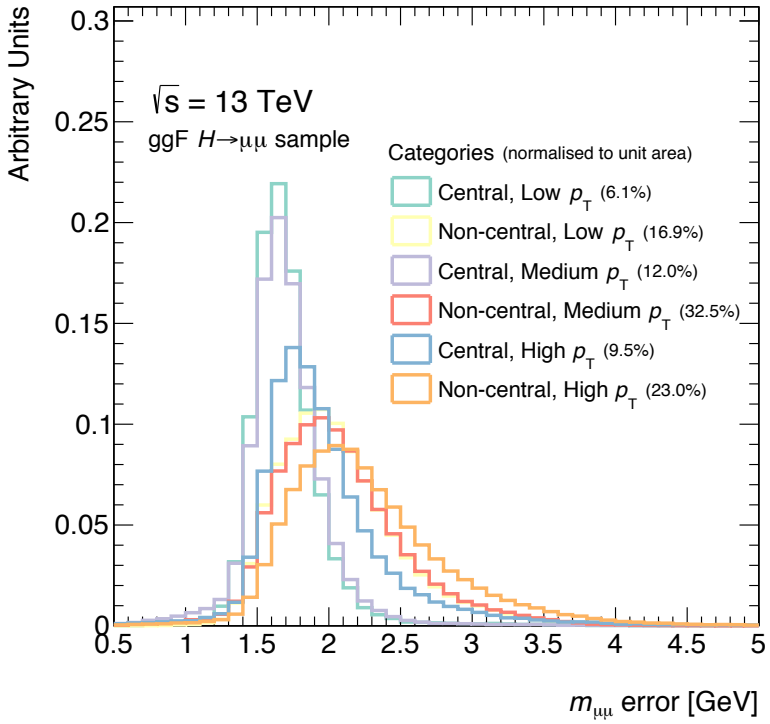


Figure 3.14: The distribution of predicted resolutions in $H \to \mu\mu$ categories for ggF $H \to \mu\mu$ events passing selection described in Ref. [5]. The distributions are normalised to unit area, their relative sizes are shown in the brackets in the legend. Figure produced by a collaborator.

Chapter 4

Search for the $H \to \mu\mu$ decay

"The Higgs boson may, or may not, couple to the second generation fermions."

— a bright bunch of ATLAS scientists.

The interactions between the Higgs boson and the vector bosons, and between the Higgs boson and the third generation fermions, both critical to the discovery in 2012, soon became the preferred way to study the properties of the Higgs boson. The mass [19, 80, 81], as well as inclusive and differential cross-sections [82–86], have been measured and new production modes, VBF [14, 22], $t\bar{t}H$ [87, 88] and VH [89, 90], have been observed. However, the interactions between the Higgs boson and the first or second generation of fermions have not been observed yet due to experimental challenges. In the leptonic sector, the main challenge is a very small branching fraction due to the small mass of the electron and the muon compared to the tau. In the quark sector, some branching ratios, for example to the charm quark pair, are larger than those in which the Higgs was discovered. However, detection is experimentally difficult in the LHC environment due to the uncertainties in flavour tagging, and estimating and modelling backgrounds [91, 92]. Ultimately, the decay to muon pairs offers the best sensitivity to probe the interactions between the Higgs boson and the second generation fermions.

In the Standard Model the branching ratio of the Higgs boson with a mass of 125.09 GeV to a muon pair is predicted to be 2.17×10^{-4} [23]. This could be modified by physics beyond the SM, for example a model in which the Yukawa couplings depend on the Higgs field, or a dilaton model with shifts in its Yukawa couplings [15, 16, 93]. This means that any deviation from the predicted value could be a sign of new physics.

The ATLAS and CMS experiments have already conducted searches using the LHC Run 1 data collected at centre-of-mass energies $\sqrt{s} = 7$ and 8 TeV, both setting the 95% confidence level upper limits on the product of the Higgs production cross-section and the branching ratio to muon pairs [94, 95] of about seven times the SM expectation. Using the data collected at $\sqrt{s} = 13$ TeV in years 2015 and 2016, the observed upper limit was further decreased to 2.8 and 2.9 by ATLAS and CMS respectively [96, 97], with ATLAS further updating its result to 2.1 using the data collected in 2017 [5]. This corresponds to 0.9σ expected significance, meaning that the analysis is near to seeing evidence for the decay.

This chapter presents the search for the Higgs boson decay to a muon pair using the full ATLAS Run 2 dataset collected between 2015 and 2018 at $\sqrt{s} = 13$ TeV, almost doubling the integrated luminosity of the dataset since the last result

to 139 fb^{-1}. The Higgs boson mass is assumed to be 125.09 GeV for all presented results.

The overall strategy of the analysis is to select events with two oppositely charged muons passing the single muon triggers, and apply additional requirements to reduce $t\bar{t}$ and diboson backgrounds. Events are then split in three channels based on whether they contain zero, one, or two or more jets in addition to the muon pair. They are further split in individual categories using a multivariate discriminant, based on the differences in kinematics between the muon pairs coming from a Higgs boson produced in the main production modes, and the muon pairs coming from background events, which are dominated by the Drell-Yan process. Signal events tend to be more central and have higher transverse momentum, their jet multiplicity is higher, and the VBF production mode has a unique signature of two back-to-back high-p_{T} jets with little hadronic activity between them. Finally, a maximum-likelihood fit to the dimuon invariant mass spectrum is performed in all categories simultaneously, extracting the parameters of interest. The analysis is limited by statistical uncertainty arising from the limited size of the dataset. The systematic uncertainties are dominated by the uncertainty on the background modelling, assessed using a dedicated simulated sample. Additional systematic uncertainties arise from the normalisation of signal sample and the migration between the categories, such as the uncertainty on the production cross-section and branching ratio from theory, and luminosity, muon momentum resolution and detector calibration from the experiment.

A number of improvements were introduced with respect to the previous result. The selection was improved by including additional muons from the corners of detector acceptance, by improving the resolution of the invariant mass by recovering the muon final state radiation, and by better rejection of pile-up jets in the forward region. The allocation of events in the categories was improved by employing a fully multivariate approach in lieu of a hybrid approach combining a multivariate discriminant and cut-based categories. Finally, the background modelling was improved by using a new functional form and larger DY simulation dataset. The total effect of all improvements is an approximately 20–30% increase in the analysis sensitivity and is dominated by the improvements from using a fully multivariate categorisation approach.

4.1 Data and MC simulation samples

The proton-proton collision data used in this analysis was collected in 2015, 2016, 2017, and 2018 at the centre-of-mass energy of 13 TeV in the main physics stream. It corresponds to an integrated luminosity of 139 fb^{-1} after passing the data quality checks ensuring that the important parts of the detector are switched on and function as intended.

The MC simulation samples are used in the analysis for a variety of purposes. The signal samples are used to optimise the event selection and to determine the normalisation and shape of the signal model in the analysis categories. Centrally produced background MC simulation samples are used to optimise the event selection and to validate the muon momentum resolution in the analysis categories, but not to determine the normalisation or shape of the backgrounds. While the normalisation and shape parameters are determined in the final fit to data, the functional form is selected in each analysis category independently using a custom high-statistics MC simulation of Drell-Yan events.

4.1.1 Background samples

The leading irreducible background for the analysis is the Drell-Yan (DY) production of muon pairs. It is simulated using Sherpa 2.2.1 with the NNPDF3.0 set of parton distribution functions (PDF) [98] in slices of scalar sum of jet transverse momenta (H_T) and with c- and b-quark filters, simulating 0–2 jet events at next-to-leading order (NLO) and at leading order (LO) for 3 and 4 jets using Comix [99] and OpenLoops [100, 101] libraries. The matching with Sherpa parton showering [102] is done using the MEPS@NLO prescription [103–105]. In order to populate the tails of the DY invariant mass distribution relevant for the analysis an additional high-statistics sample is generated with identical setup and an additional cut of $m_{\mu\mu} > 100$ GeV. Finally, Sherpa is also used to model the electroweak Z+jets process with up to two additional jets at LO accuracy beyond the first two jets.

With the mass of the top-quark set to $m_t = 172.5$ GeV, the $t\bar{t}$ and single-top samples are generated using Powheg-Box [49–51] with NNPDF3.0 PDF set and parton showering and hadronisation done in Pythia 8.186 using the A14

parameter set [106]. The $t\bar{t}$ cross-section is computed to next-to-next-to-leading order (NNLO) in QCD with next-to-next-to-leading logarithmic (NNLL) soft gluon terms taken into account. The cross-section of the single-top is computed with prescriptions from [107, 108], and all the processes (t-channel, s-channel, and Wt-channel) are generated separately.

Diboson processes, WZ and ZZ, are generated using Sherpa 2.2.1 with the NNPDF3.0 PDF set and are normalised directly to the Sherpa prediction. Only semi-leptonic and fully-leptonic decays are simulated.

4.1.2 Signal samples

The largest contribution comes from the ggF signal sample, generated with Powheg-Box with PDF4LHC15 PDF set to next-to-next-to-leading order accuracy with parton shower matching (NNLOPS) [109], achieving NNLO in QCD after the re-weighting in rapidity of the Higgs boson. The cross-section for the ggF production mode are computed to next-to-next-to-next-to-leading order (N3LO) [110] in QCD with NLO electroweak corrections applied [111, 112] under the assumption that they factorise. Pythia 8 with the AZNLO parameters [113] is used to simulate the decay of the Higgs boson, final-state photon radiation, parton showering, and hadronisation.

VBF and $q\bar{q}/qg \to$ VH production modes are generated to NLO accuracy in QCD using Powheg-Box, while the $gg \to ZH$ is generated to LO accuracy. The same PDF set and Pythia tunes are used as in the ggF generation. The VBF cross-section is computed with full NLO QCD and electroweak corrections [114–116], and approximate but highly accurate NNLO QCD corrections [117]. The VH cross-section is computed at NNLO in QCD [118] with NLO electroweak corrections [119], apart from the $gg \to ZH$ which is computed to NLO and next-to-leading-logarithm accuracy in QCD [118–125].

The $t\bar{t}H$ production mode is generated using MadGraph5_aMC@NLO [126, 127] at NLO accuracy using NNPDF3.0NLO PDF set and the A14 tune. The cross-section is computed to NLO in QCD with NLO electroweak corrections [128, 129].

Pythia is also used to overlay minimum bias events to model the effects of pile-up for all simulated events using the NNPDF2.3LO PDF set [130] with the A3

parameters [131]. All signal and background samples are passed through the full ATLAS detector simulation. Pile-up re-weighting, object momentum smearing, and efficiency scale factors are applied to the MC simulation samples in order to make them mimic the collected data.

4.1.3 Custom Drell-Yan sample

Extracting the signal strength paramater relies on a maximum-likelihood fit to the invariant mass spectrum in data, using parametrised models of both signal and background. The choice of the functional form describing the background is crucial because a mismodelling of the background may result in a bias on the extracted signal strength. This is particularly significant when the signal-to-background ratio is very small because the impact can be large.

In order to select the functional form and evaluate the resulting bias, referred to as the "spurious signal", the combined signal and background model is fit to the background-only spectrum. Ideally, this would be done on the background MC simulation described in Section 4.1.1, but because of the high computational cost associated with the simulation of the detector response the statistics of the sample are on the same order of magnitude as the data. As a result, the extracted spurious signal value is unreliable and dominated by the statistical fluctuations. A custom high-statistics sample of DY events is generated to overcome this challenge.

The custom DY sample needs to have significantly larger statistics than the data, requiring a much faster generation. To facilitate this, the sample is produced at the generator-level only, skipping the simulation of showering, hadronisation, and detector response, replacing them by parametrised smearing functions. The effect on the dimuon invariant mass spectrum is minimal, and because of the parametrised smearing the jet mismodelling is a second order effect. On the other hand, final-state radiation (FSR) does have an effect on the invariant mass spectrum and is simulated by Photos++ [132]. This setup was fast enough to generate 20 billion events corresponding to approximately 100 ab^{-1} in the regions of parameter space relevant for the analysis. A drawback of the sample is that it only models the dominant DY part of the invariant mass spectrum.

At the generator level, the samples are produced separately for zero or one parton and two partons in addition to the muon pair:

- POWHEG is used to generate an inclusive NLO $Z \to \mu\mu$ sample, modelling zero and one parton in addition to the muon pair. For efficiency reasons two sub-samples are generated, one with $60 < m_{\mu\mu} < 95$ GeV (2.5×10^9 events), and the other with $m_{\mu\mu} > 95$ GeV (10×10^9 events) requirement. Even after detector smearing the latter sample dominates the fit region.

- ALPGEN [133] is used to generate events with two partons in addition to the muon pair at LO. The computation uses only the matrix element with two additional partons and requires both of them to have $p_{\mathrm{T}} > 25$ GeV as well as $\Delta R(j, j) > 0.4$ to minimise the overlap with the POWHEG sample. Similarly to the POWHEG sample, events are generated with $60 < m_{\mu\mu} < 95$ GeV (1.75×10^9 events) and $m_{\mu\mu} > 95$ GeV (2.5×10^9 events) requirements. These events are crucial to populate the analysis categories with a large fraction of the signal events coming from the VBF production mode.

Both samples are corrected in the same way to model the effects of the detector. Muon momentum smearing is performed using parametrised functions derived from the MC simulation described in Section 4.1.1. Effects from the identification, reconstruction, isolation, and impact parameter efficiencies are also derived from the MC simulation and are applied as weights, while the trigger efficiency weights are derived from the data. The FSR momentum is smeared using a parametrised prescription from the ATLAS subgroup specialising in electrons and photons. Photons are dropped randomly to emulate reconstruction efficiency. The smearing of jet momenta is performed using the parametrisation from the ATLAS subgroup specialising in jets and missing energy. The effect of pile-up jets is modelled by building a library of jets from minimum-bias events in the MC simulation and overlaying a number of sampled jets from the library for each event. A dedicated missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) smearing is derived from the Run 2 data.

4.2 Physics objects

The search for a narrow resonance on top of a smoothly falling background appears at first sight to require not much more than an excellent understanding of muons in ATLAS. While the entirety of Chapter 3 is dedicated to their performance, the sophistication of the analysis requires an understanding of other objects as well. Photons are used to recover the final-state radiation in order to improve the resolution of the invariant mass spectrum. Jets are used as a handle to select events produced in the VBF production mode and to improve separation from the other backgrounds. Electrons are reconstructed, albeit solely in order to reject fake jets. Finally, missing transverse energy is used to suppress top and diboson backgrounds.

4.2.1 Muons

Muon reconstruction, identification, and isolation are described in detail in Chapter 3. Table 4.1 summarises the selection used for the $H \to \mu\mu$ analysis. The majority of reconstructed muons in the Loose selection are the CB type, combining the information from the tracker and the muon spectrometer systems. However, in order to maximise the acceptance of signal events, other muons types are also used to improve the reconstruction efficiency in the challenging parts of the detector. ST and CT muons improve the reconstruction efficiency in the $|\eta| < 0.1$ region, where the muon spectrometer has a gap to allow for cabling and services to the

Table 4.1: Summary of the muon selection requirements.

ID selection	Loose
Isolation selection	FixedCutPflowLoose
p_{T}	$p_{\mathrm{T}} > 15$ GeV
η	$\lvert\eta\rvert < 2.7$
Impact parameters	$\lvert d_0/\sigma_{d_0}\rvert < 3.0$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm

tracker, while the ME muons are used in the forward regions to reconstruct muons outside of the tracker coverage, where $2.5 < |\eta| < 2.7$.

The p_{T} measurement of the CB muons benefits from combining information from two subdetectors, namely the tracker and the muon spectrometer. Figure 4.7 verifies that the resulting dimuon invariant mass resolution is only marginally degraded for muon pairs containing one non-CB muon (dotted green line) compared to pairs where both muons are CB (dashed blue line). Since most muons are CB, the inclusive distribution (solid orange line) and the distribution of purely CB muons (dashed blue line) are nearly identical.

The reason for such a small degradation is that the resolution of the combined p_{T} measurement is dominated by the tracker p_{T} measurement in the barrel region, and by the muon spectrometer measurement in the forward regions. ST and CT muons, used in the $|\eta| < 0.1$ region, combine a track in the ID with a segment in the muon spectrometer or with a calorimeter deposit and thus benefit from having the high-quality tracker p_{T} measurement. Conversely, ME muons, used in the $2.5 < |\eta| < 2.7$ region, use the track from the muon spectrometer, the higher quality measurement in the forward region. The FixedCutPflowLoose isolation selection is used because of a combination of its high efficiency, robustness to high pile-up conditions, and improved rejection of heavy-flavour jets. The cut on the p_{T} rejects muons from the decay of heavy-flavour quarks, while cosmic rays are rejected by the standard requirements on the impact parameters.

4.2.2 Photons and FSR

Photons are reconstructed in ATLAS via its energy deposits in the eletromagnetic calorimeter. Unlike electrons, photons are electrically neutral and do not leave tracks in the tracker, unless they are converted to electron-positron pairs before reaching the calorimeter. The details of the reconstruction algorithm and its performance are described in Ref. [134].

Muons may radiate a photon and thus lose energy, meaning that the invariant mass of the muon pair is no longer equal to the mass of the parent. In order to mitigate this effect up to one final-state photon candidate is added to the mass calculation, using a procedure similar to the one used in Ref. [135], but optimised

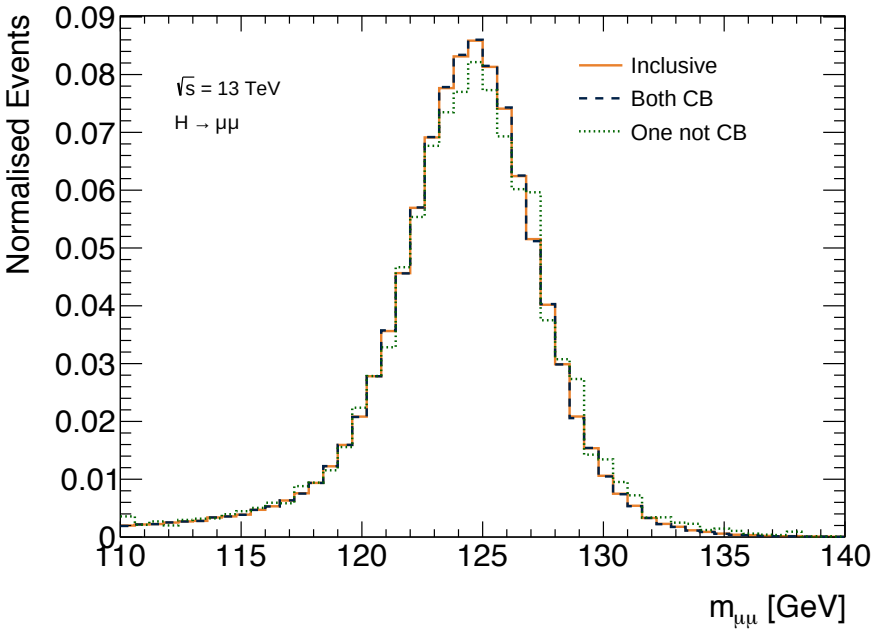


Figure 4.1: The resolution of the dimuon invariant mass measurement for the ggF signal sample in the $H \to \mu\mu$ analysis signal region. The events in which both muons are CB type are shown as a dashed blue line, while events in which one of the muons is not CB are shown as a dotted green line. The inclusive distribution is shown as a solid orange line.

for the high-purity requirements of this analysis. Collinear photons ($\Delta R(\gamma, \mu) < 0.20$) are required to have p_{T} larger than a linearly increasing threshold, from $p_{\mathrm{T}}^{\gamma} > 3$ GeV at $\Delta R(\gamma, \mu) = 0$ to $p_{\mathrm{T}}^{\gamma} > 8$ GeV at $\Delta R(\gamma, \mu) = 0.2$. Non-collinear photons ($\Delta R(\gamma, \mu) > 0.20$) are required to have $p_{\mathrm{T}}^{\gamma} > 8$ GeV and a tight [134] identification requirement for photons. When multiple photon candidates are found only the one with the highest p_{T} is included in the invariant mass calculation, with priority given to the collinear photons over non-collinear photons.

Around 5% of all signal events have at least one FSR candidate, of which approximately 90% are collinear photons. The distribution of invariant mass for the MC simulation of ggF signal events with at least one reconstructed FSR candidate is shown in Figure 4.2 before and after the recovery of final-state radiation. It can be seen that the mass peak is shifted much closer to 125 GeV, with the slight overshoot resulting from the fake FSR candidates coming from the pile-up. The

improvement in the invariant mass resolution is approximately 3% inclusively. A drawback of the FSR recovery is an $\sim 8\%$ increase in the number of background events in the signal region due to the migration of events from the Z mass peak.

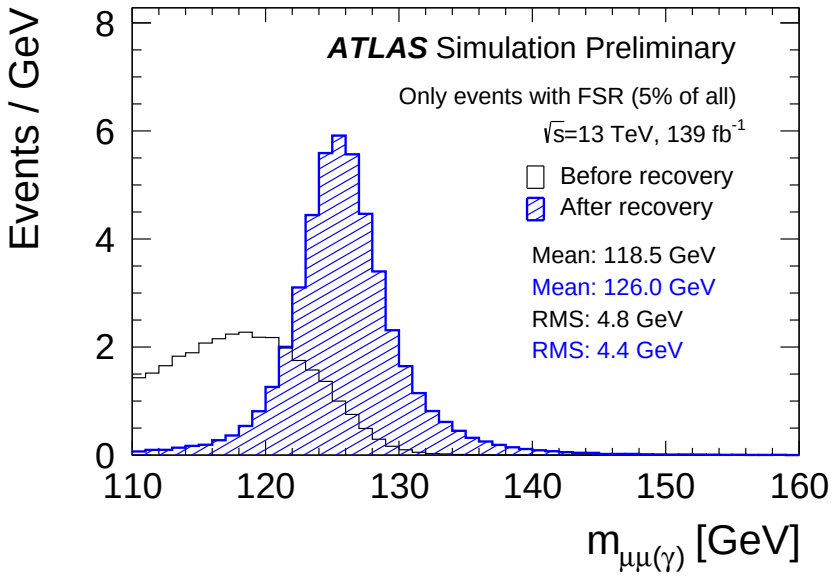


Figure 4.2: the invariant mass spectrum before (white solid) and after (blue hatches) final-state radiation recovery for a MC simulation of the ggF signal events passing the analysis selection. Only events with one or more reconstructed FSR candidates are shown. The mean and the root-mean-square of the distributions are shown in the bottom right corner. Reproduced from Ref. [6].

4.2.3 Jets, electrons, and $E_{\mathrm{T}}^{\mathrm{miss}}$

Jets are the collimated collections of hadrons, arising either from a parton in the final state of the interaction, the remnants of the proton not participating in the hard scatter (underlying event), or from pile-up interactions. They are used to identify events which could have been produced via the VBF production mode by exploiting its characteristic dijet signature. Jets are experimentally the richest objects in ATLAS, reconstructed either from the deposits in the calorimeters, tracks that charged hadrons leave in the tracker, or both. The optimal use of substructure has been extensively studied both theoretically [136] and using a wide

variety of modern machine learning techniques [137]. The reconstruction used in this analysis is based on topological clusters of calorimeter cells [138], which are used as an input to the anti-k_T clustering algorithm [139, 140] with the distance parameter $\Delta R = 0.4$. A requirement on the jet $p_T > 25$ GeV for the central jets ($|\eta| < 2.5$), and at $p_T > 30$ GeV for the forward jets ($2.5 < |\eta| < 4.5$).

The suppression of jets coming from pile-up is achieved by using a dedicated jet vertex tagging (JVT) discriminant [141], constructed using tracking information. In the forward region, pile-up jet suppression is achieved with the fJVT discriminant [142], using the information on jet shape and the topological correlations in pile-up interactions.

Jets containing a b-hadron are identified using a multivariate b-tagging algorithm [143] if they are inside the tracker ($|\eta| < 2.5$). A selection providing 60% tagging efficiency for real b-jets and less than 0.1% fake rate for light-flavoured jets [144] is used to tag the jets. Electrons are reconstructed by matching a track in

Table 4.2: Summary of the jet selection requirements.

Algorithm	Anti-k_T ($\Delta R = 0.4$)
η	$\|\eta\| < 4.5$
p_T	$p_T > 25$ GeV for $\|\eta\| < 2.4$ $p_T > 30$ GeV for $2.4 < \|\eta\| < 4.5$
JVT	JVT > 0.59 GeV for $\|\eta\| < 2.4$ and $20 < p_T < 120$ GeV JVT > 0.11 GeV for $2.4 < \|\eta\| < 2.5$ and $20 < p_T < 120$ GeV
fJVT	fJVT < 0.5 GeV for $2.5 < \|\eta\| < 4.5$ and $20 < p_T < 120$ GeV
b-tag	MV2c10 60 WP for $\|\eta\| < 2.5$ and $p_T > 20$ GeV

the ID with a calorimeter deposit in the ECAL, as described in Ref. [145]. Unlike the topological cluster used for jets, the calorimeter deposits are clustered using a sliding-window algorithm [146] which can be more accurately calibrated due to the fixed-size cells.

Electrons are used in the analysis only in the overlap removal procedure to remove fake jets. The electron selection requirements are summarised in Table 4.3. Electron candidates are required to pass the Medium likelihood identification

selection and FCLoose isolation selection [145] to reject electrons from hadronic decays or pile-up. The electron p_{T} is required to be larger than 7 GeV, while pseudorapidity has to satisfy $|\eta| < 2.47$ with the exception of the crack region $1.37 < |\eta| < 1.52$. A quality requirement is used to reject electrons in which a cell is dead or faulty, and requirements on the impact parameters are used to reject electrons coming from heavy-flavour decays or other proton-proton interactions.

Table 4.3: Summary of the electron selection requirements.

Identification	Medium LH
Isolation	FCLoose
p_{T}	$p_{\mathrm{T}} > 7$ GeV
η	$\lvert\eta\rvert < 2.47$ excluding $1.37 < \lvert\eta\rvert < 1.52$
Quality	No dead or faulty cells
Impact parameters	$\lvert d_0/\sigma_{d_0}\rvert < 5$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm

Missing transverse energy is defined as the magnitude of the vectorial sum of p_{T} of all calibrated muons, electrons, jets, and the soft term. The soft term is defined as the vectorial sum of the p_{T} of other tracks not used in the reconstruction of any objects. The details of the reconstruction algorithm, in particular the treatment of the soft term, are available in Ref. [147], along with performance studies.

$E_{\mathrm{T}}^{\mathrm{miss}}$ can result from neutrinos, which escape the detector without leaving any deposits, or from experimental inaccuracies when measuring muons, electrons, or jets. In the analysis, $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to discriminate between the signal and $t\bar{t}$ processes. Instead of a cut, it is used as an input to the multivariate discriminant in the 2-jet channel.

A single physical object can create deposits in the detector which are used by two different reconstruction algorithms to create more than one particle candidate. For example, an electron results in an ECAL deposit which is used in both the reconstruction of a jet and an electron. For this reason, an object overlap removal procedure is used to remove potential duplicates.

Muons are only removed if they are CT type and share a track with an electron, or they are close to a jet ($\Delta R_{\mu,j} < 0.2$) with more than two tracks and where muon p_{T} is less than 0.7 of all track p_{T}. Jets are removed if they are close ($\Delta R_{e,j} < 0.2$) to electrons.

4.3 Event selection

A very loose event selection based on oppositely charged muon pairs is used for the analysis to maximise the number of signal events. DY background is irreducible, but the $t\bar{t}$ contributions can be reduced by a veto on events with a b-tagged jet. The selection, summarised in Table 4.4, results in approximately 59% acceptance for the ggF and VBF signal events.

Table 4.4: Summary of the event selection requirements.

Cleaning	Pass GoodRunList Pass single muon trigger Primary vertex (> 2 tracks with $p_{\mathrm{T}} > 0.5$ GeV)
Muons	Two oppositely charged muons $p_{\mathrm{T}}^{\mathrm{lead}} > 27$ GeV $p_{\mathrm{T}}^{\mathrm{sublead}} > 15$ GeV
Jets	Zero b-tagged jets

The ATLAS detector is a complex instrument composed of many independent subcomponents that are being pushed to the limits of their performance. Failure of detector subcomponents is a normal part of detector operations and is dealt with by detailed monitoring of performance and the release of so-called GoodRunList (GRL). The GRL contains a list of runs further segmented to two-minute data-taking intervals, called luminosity blocks, in which the detector operates as intended. The data used in the analysis is required to be collected in one of those luminosity blocks in order to veto pathological events.

Furthermore, the data is required to pass the lowest unprescaled single muon triggers, these are `HLT_mu20_iloose` or `HLT_mu50` in the 2015 and `HLT_mu26_ivarmedium` or `HLT_mu50` in the 2016–2018 data-taking period. These triggers require at least

one isolated muon with a p_{T} larger than 20 or 26 GeV respectively, or one non-isolated muon with at least 50 GeV [148]. The increase in trigger thresholds was required to deal with higher rates of events at an increased instantaneous luminosity between 2016 and 2018.

Events are required to have at least one reconstructed primary vertex candidate with at least two tracks with $p_{\mathrm{T}} > 0.5$ GeV in the inner detector. When multiple vertices are reconstructed, the one with the highest scalar sum of p_{T} of the associated tracks is considered to be the hard scatter vertex.

Events are required to have exactly two oppositely charged muons. The leading muon has $p_{\mathrm{T}} > 27$ GeV in order to pass the trigger thresholds, while the subleading muon is required to have $p_{\mathrm{T}} > 15$ GeV, with both having a pseudorapidity requirement of $|\eta| < 2.7$. Additionally, events are required to have zero b-tagged jets to suppress $t\bar{t}$ and diboson backgrounds.

After the common selection the Z control region, the signal region, and the central and sideband regions are defined by requirements on the invariant mass, summarised in Table 4.5.

Table 4.5: Summary of the analysis regions.

Z control region	$76 < m_{\mu\mu} < 106$ GeV
Fit region	$110 < m_{\mu\mu} < 160$ GeV
Central region	$120 < m_{\mu\mu} < 130$ GeV
Sideband region	$110 < m_{\mu\mu} < 120$ GeV or $130 < m_{\mu\mu} < 180$ GeV

The Z control region contains almost one hundred million events in data and is used to validate detector performance, in particular to check the muon resolution in the data and the MC simulation. The sideband region contains approximately two million events in the data, which are used to train the multivariate classifier and to validate the background modelling. The central region contains approximately 450,000 events in the data, of which about 860 are expected from the SM signal. The final maximum-likelihood fit is performed in the fit region.

The invariant mass spectrum after the recovery of the final state radiation is shown in Figure 4.3 for the data and the MC simulation of signal and background samples for events passing the selection requirements. The spectrum is dominated by DY events, especially around the Z mass peak at around 91 GeV. The Z peak is also present in the diboson backgrounds, while the top backgrounds have a smoothly falling spectrum. Signal events form a narrow peak at around 125 GeV, with their yields multiplied by a factor of a hundred for visibility.

4.4 Event categorisation

The goal of event categorisation is to improve the sensitivity of the analysis by splitting events in categories with different signal purities. In previous iterations of the analysis, physics-based arguments were used to construct categories targeting the VBF production mode and signal enriched regions using requirements on event kinematics. However, with increased statistics of the available dataset, it became increasingly advantageous to use modern machine learning models and infer the classification scheme from the available data. With this approach, care needs to be taken to prevent bias and/or sub-optimal performance due to overfitting.

The categorisation scheme begins by splitting events in categories based on jet multiplicity: 0-jet, 1-jet, and 2-jet channel, where the 2-jet channel includes events with more than two jets. Kinematic differences between signal and background events are exploited by training boosted decision tree (BDT) classifiers in each of the jet channels using the XGBoost [149] package.

In each of the categories a "Higgs classifier" is trained to distinguish between the MC simulation of ggF and VBF events, and the data events from the sideband region. In the 2-jet category an additional "VBF classifier" is trained to distinguish between the MC simulation of VBF signal events and the data sidebands. In both trainings the signal event weights are scaled so that their sum matches the number of data events.

The variables used in the training in each of the jet channels are summarised in Table 4.6. The 0-jet channel uses only three variables, related to the kinematics of the dimuon system: p_{T} and rapidity in addition to $|\cos\theta^*|$, a variable sensitive

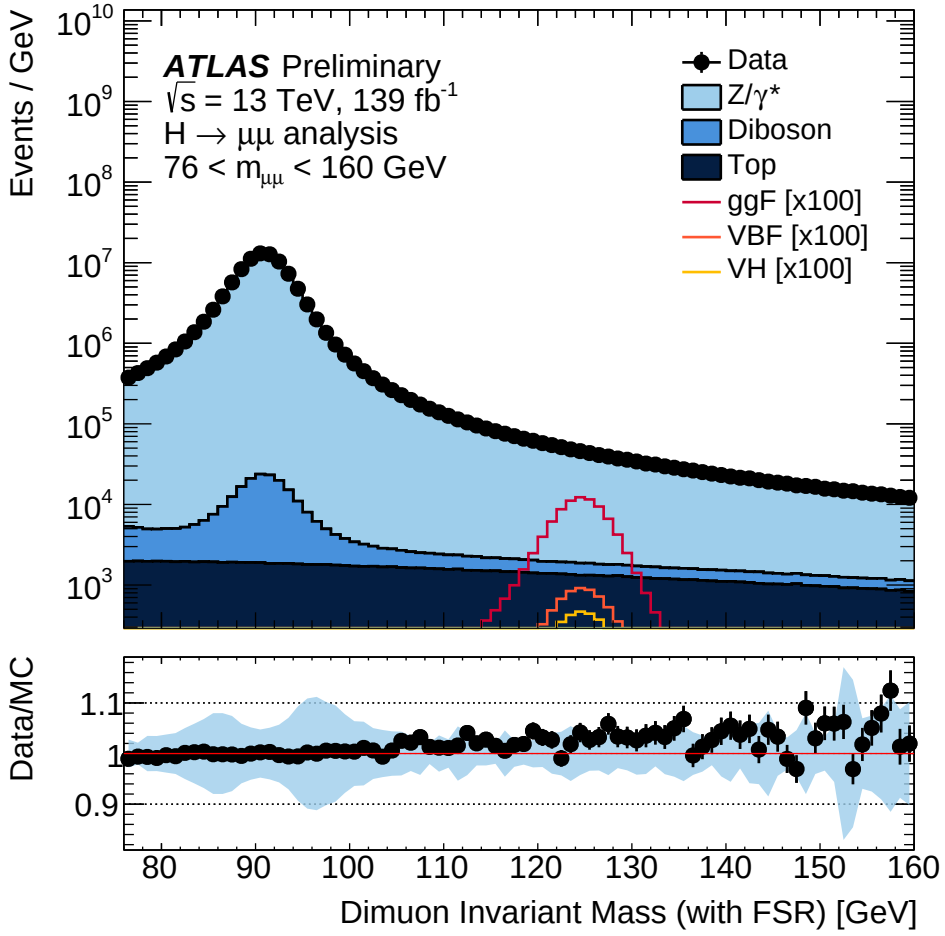


Figure 4.3: The invariant mass spectrum after the correction for final-state radiation, for data corresponding to 139 fb^{-1}. MC simulation is normalised to the data for easier comparison. The data is shown as black markers, the MC simulation of backgrounds is dominated by the DY events (light blue), followed by diboson (blue) and top (dark blue) contributions. The signal events coming from the ggF, VBF, and VH production modes are shown as solid red, orange, and yellow lines, respectively, with the yields normalised to one hundred times the SM prediction for visibility. There are too few $t\bar{t}H$ events to appear on the figure. The bottom panel shows the ratio between the data and MC simulation of background events, while the blue band represents the experimental systematic uncertainty. The statistical uncertainty on both data and MC simulation is combined in the error bars on data.

to the spin of the mediator [150, 151], defined as

$$|\cos\theta^*| = \left| \frac{2(P_1^+ P_2^- - P_1^- P_2^+)}{\sqrt{m_{\ell\ell}^2(m_{\ell\ell}^2 + p_{\mathrm{T},\ell\ell}^2)}} \frac{p_{z,\ell\ell}}{|p_{z,\ell\ell}|} \right|, \tag{4.1}$$

where $m_{\ell\ell}$, $p_{\mathrm{T},\ell\ell}$, $p_{z,\ell\ell}$ are the mass, transverse momentum, and longitudinal momentum, respectively, of the dimuon system, and the $P_i^{\pm}$ are defined as $P_i^{\pm} = (E_i \pm p_{z,i})/\sqrt{2}$, where E_i and $p_{z,i}$ are the energy and the longitudinal momentum of a muon i, with $i = 1$ corresponds to the muon with larger p_{T}.

In the 1-jet channel, the set of variables is extended by adding the kinematics of the leading jet. In the 2-jet channel, the kinematics of the subleading jet are added, along with $E_{\mathrm{T}}^{\mathrm{miss}}$ and a couple of variables related to the dijet system, detailed in Table 4.6.

Table 4.6: Summary of the training variables.

0-jet	$p_{\mathrm{T}}^{\mu\mu}$	p_{T} of the dimuon system
	$Y^{\mu\mu}$	Rapidity of the dimuon system
	$\lvert\cos\theta^*\rvert$	Spin-related variable (see Eq. 4.1)
1-jet	0-jet variables	
	$p_{\mathrm{T}}^{j_1}$	p_{T} of the lead jet
	η^{j_1}	η of the lead jet
	$\Delta\phi_{j_1,\mu\mu}$	$\Delta\phi$ between the lead jet and dimuon system
2-jet	1-jet variables	
	$p_{\mathrm{T}}^{j_2}$	p_{T} of the sublead jet
	η^{j_2}	η of the sublead jet
	$\Delta\phi_{j_2,\mu\mu}$	$\Delta\phi$ between the sublead jet and dimuon system
	p_{T}^{jj}	p_{T} of the dijet system
	Y^{j_2}	Rapidity of the dijet system
	$\Delta\phi_{jj,\mu\mu}$	$\Delta\phi$ between the dijet system and dimuon system
	m_{jj}	Invariant mass of the dijet system
	$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse energy

The variables were chosen to maximise the sensitivity while minimising systematic uncertainties. An additional concern was whether the variables would allow the classifier to learn the invariant mass and result in a shaped mass spectrum

after selection. For this reason, the complete kinematics of the individual muons are not included.

All the variables used in the 2-jet channel are shown in Figures 4.4 and 4.5 for events with two or more jets. It can be seen that the p_{T} of the dimuon system, the mass of the dijet system and azimuthal differences between the dimuon system and jets can provide separation between the signal and the background.

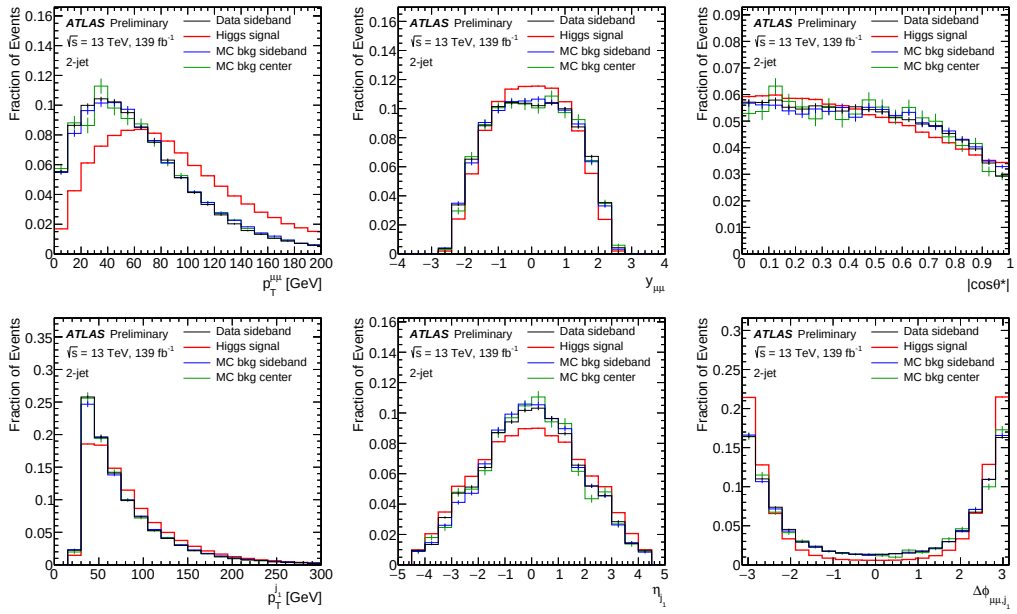


Figure 4.4: The distributions of classifier input variables for events with two or more jets. The top row shows dimuon system variables, bottom row shows leading jet variables. The $H \to \mu\mu$ signal events are shown in red, data from the sidebands is shown in black, and MC simulation of background events is shown blue (sideband region) and green (central region). The MC simulation includes DY, diboson, and top events. The error bars are statistical. Reproduced from Ref. [6].

In each jet multiplicity channel, a four-fold approach is used to prevent bias from overfitting to the training and validation sets. Overfitting refers to the adjustments of the model parameters to the random fluctuations in the training set. As a result, the model performance on the training set is better than a new dataset sampled from the same population. This is undesirable because it leads to a bias. In order to avoid this, a model is evaluated on an independent dataset, called the test set. When a number of models are trained on the training set, their perfor-

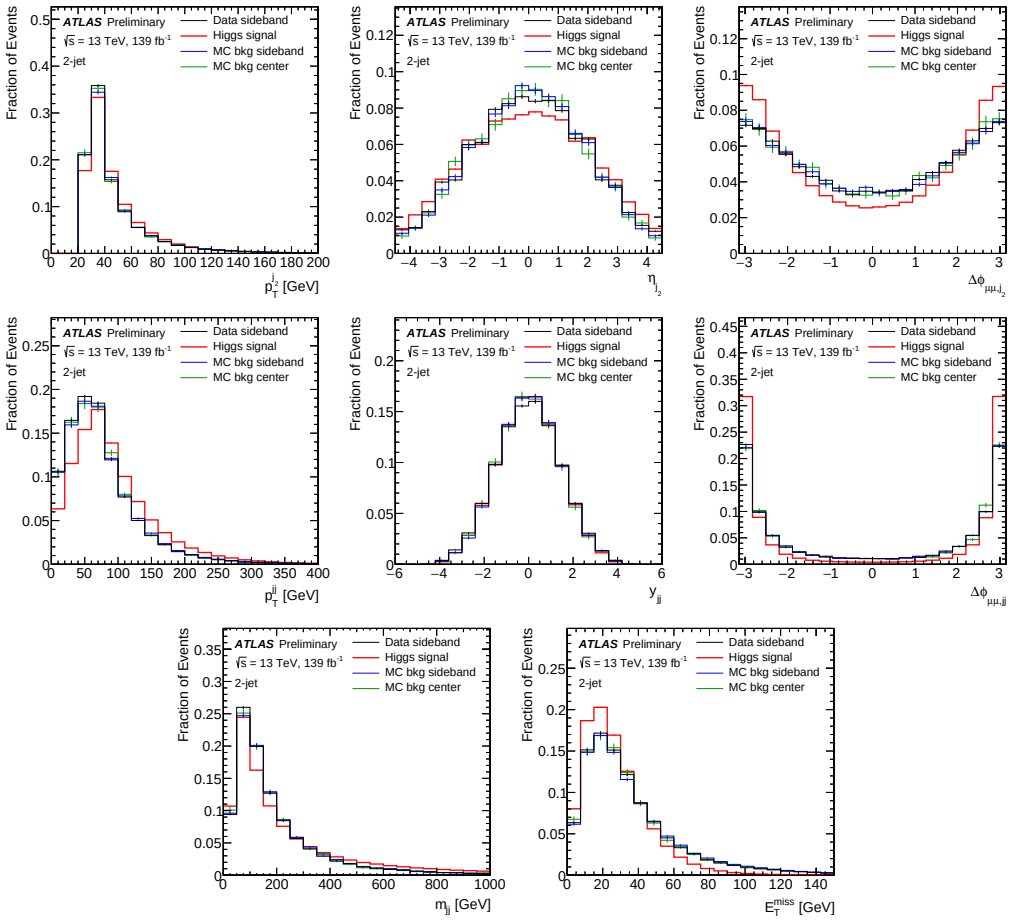


Figure 4.5: The distributions of classifier input variables for events with two or more jets. The top row shows subleading jet variables, middle row dijet system variables, bottom row shows the dijet invariant mass and $E_{\mathrm{T}}^{\mathrm{miss}}$. The $H \to \mu\mu$ signal events are shown in red, data from the sidebands are shown in black, and MC simulation of background events is shown blue (sideband region) and green (central region). The MC simulation includes DY, diboson, and top events. The error bars are statistical. Reproduced from Ref. [6].

mance is monitored using the validation dataset, and the model performing best on the validation dataset is selected. Its performance is then evaluated, again, on the independent test set. A drawback of this approach is that once the dataset is used for training or validation, it can not be used for the test set and as a result only a fraction of available data and MC is used in the analysis, which is undesirable.

The four-fold approach solves this problem by reshuffling the same dataset four

times, with each reshuffle referred to as a fold. In each fold, events are split in a training set (50% of events), a validation set (25% of events), and a test set (25% of events), such that every sample from the dataset is present in the test set in exactly one fold.

The training set is used in the training of the models. The validation set is used for early stopping of the training and to assess the performance of models using the area under the receiver operating characteristic (ROC) [152] curve. This is used to guide the Bayesian optimisation [153] to tune the hyperparameters of the model in each fold. Finally, the model is evaluated on the test set, which is used for the categorisation of events.

It should be noted that in the four-fold approach the model parameters are selected separately in each fold. This results in four models, each being applied to the test set not seen in training, early stopping, or hyperparameter tuning, ensuring that there is no bias from overfitting to either training or validation datasets.

In order to ensure that the score distributions are the same in each fold, the scores are transformed to a uniform distribution in unweighted combined signal events in all folds separately, using the QuantileTransformer algorithm from the `scikit-learn` [77] software package.

The transformed scores are named O_{VBF} for the VBF classifier and $O_{gg\text{F}}^{(0-2)}$ for the Higgs classifiers, where the superscript stands for the jet multiplicity. The category boundaries are defined as the cuts on these values, and are summarised in Table 4.7.

Table 4.7: Summary of the score boundaries defining the analysis categories.

	Low	Medium	High
0-jet	$O^0_{gg\text{F}} < 0.35$	$0.35 \leq O^0_{gg\text{F}} < 0.75$	$O^0_{gg\text{F}} \geq 0.75$
1-jet	$O^1_{gg\text{F}} < 0.38$	$0.38 \leq O^1_{gg\text{F}} < 0.78$	$O^1_{gg\text{F}} \geq 0.78$
2-jet ($O_{\text{VBF}} < 0.60$)	$O^2_{gg\text{F}} < 0.22$	$0.22 \leq O^2_{gg\text{F}} < 0.48$	$O^2_{gg\text{F}} \geq 0.48$
VBF	$O_{\text{VBF}} < 0.77$	$0.77 \leq O_{\text{VBF}} < 0.89$	$O_{\text{VBF}} \geq 0.89$

The values of category boundaries are determined by a simultaneous optimisation of all boundaries in a given jet channel using an exhaustive grid search

over the cut values with 0.01 spacing. The metric used in the optimisation is the number counting significance, defined as

$$Z = \sqrt{\sum_{c=1}^{12} 2\left[(s_c + b_c)\log\left(\frac{s_c + b_c}{b_c}\right) - s_c\right]}, \tag{4.2}$$

where the sum runs over categories c, and s_c and b_c are the number of signal and background events in the central region in category c. Since the optimisation is performed blinded, the value of b_c is computed from the number of data events in the sideband region, multiplied by a transfer factor obtained from the inclusive MC distribution.

The summary of analysis categories is shown in Figure 4.6, highlighting the composition of signal and background events, as well as the signal purity (S/B) and sensitivity proxy ($S/\sqrt{B}$), and the total number of background events (B) in the central region for all the analysis categories, where S is the expected number of signal events in the central region. The signal purity decreases from High to Low categories, while the number of background events increases. The significance proxy is approximately constant, which is a result of the optimisation rather than a requirement.

As explored in VADER4μ project, the momentum resolution of muons is a function of their kinematics. Furthermore, the categorisation of events in the analysis is based on a BDT classifier, which exploits the kinematic differences between the signal and the background events. As a result, the analysis categories can have quite different invariant mass resolutions. While this is expected, it is critical that the modelling of dimuon mass resolution in MC simulation matches the resolution in the data in each analysis category, otherwise the sensitivity would be over- or underestimated. The validation of the momentum calibrations, described in Section 3.2, is performed on the Z mass peak. The validation for 0-jet High, 1-jet High, 2-jet High, and VBF High categories are shown in Figure 4.7. The bottom panels show that resolution modelling is well within systematic uncertainties, and similar results are observed for other categories.

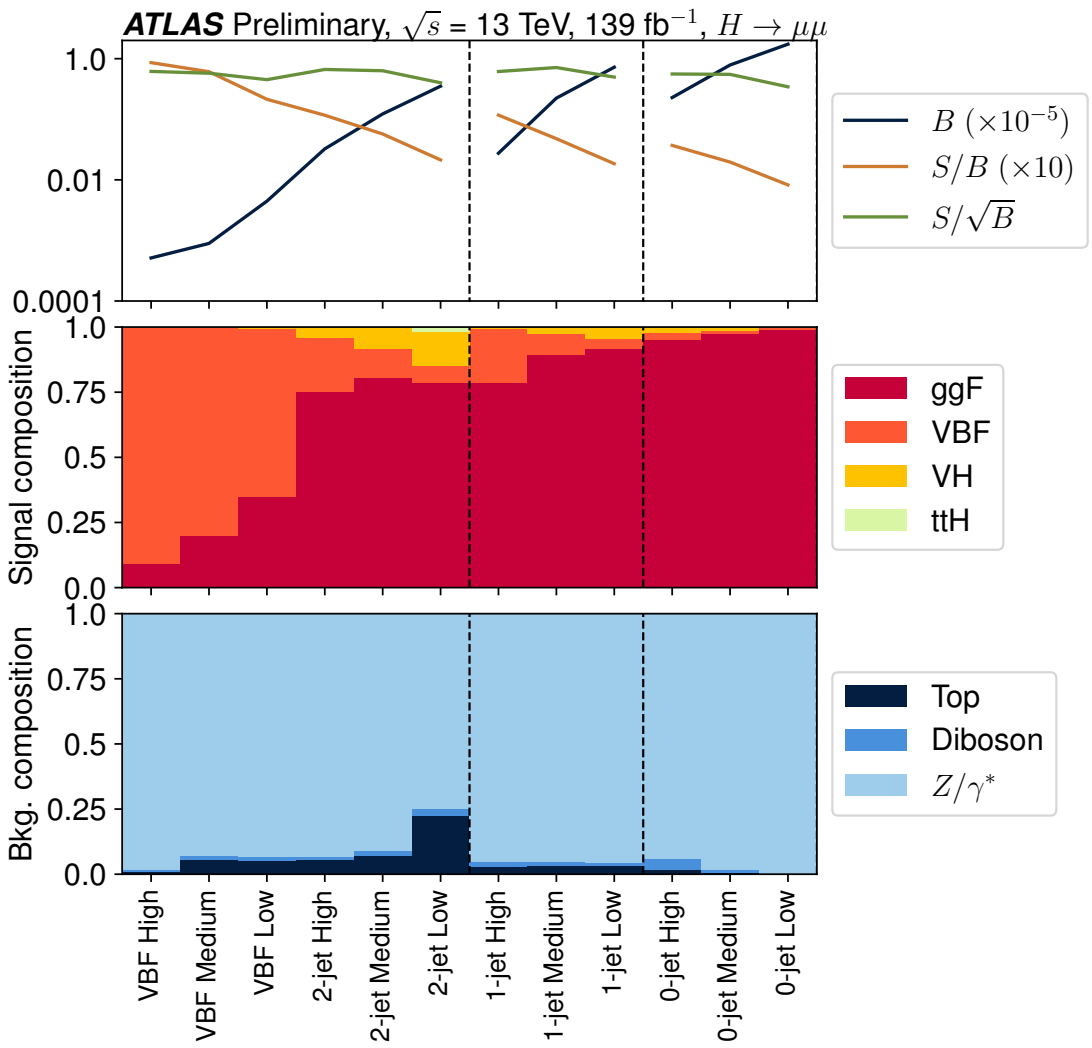


Figure 4.6: Summary of $H \to \mu\mu$ analysis categories. The top panel shows the total number of background events in the central region, multiplied by 10^{-5} for visibility in blue, while the signal purity, multiplied by 10, is shown in orange and the significance proxy in green. The middle panel shows the signal composition of the categories, dominated by VBF events, shown in orange, in the VBF targetting categories, and by ggF events, shown in red, in all the others. The VH events are shown in yellow, and $t\bar{t}H$ events in lime. The background composition is shown in the bottom panel and is dominated by DY events, shown in light blue. Diboson events are shown in blue, and top events in dark blue.

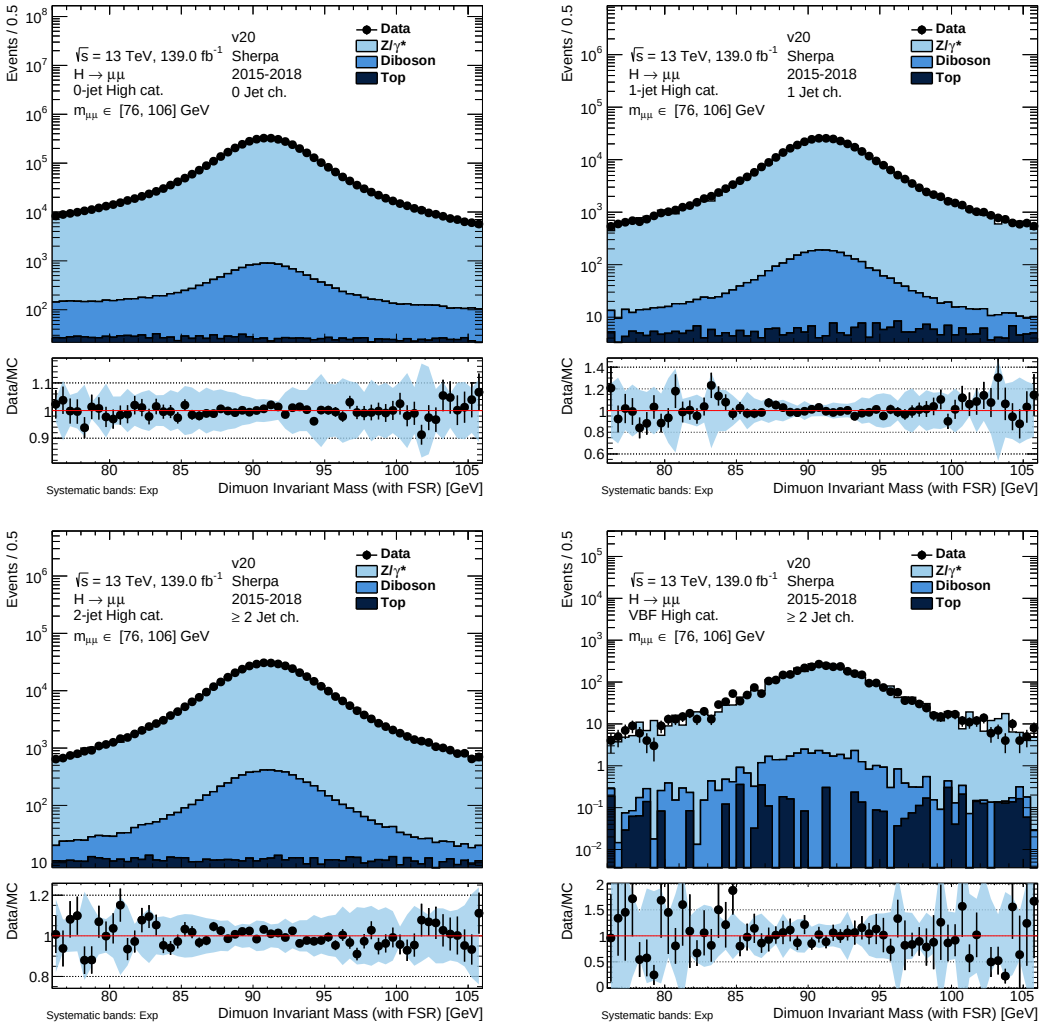


Figure 4.7: The invariant mass distribution after the FSR correction at the Z peak for the 0-jet High, 1-jet High, 2-jet High, and VBF High analysis categories. The data is shown as black markers, Drell-Yan as light blue, diboson as blue, and top MC simulated events as dark blue histograms. The MC simulation is normalised to data for easier visual comparison. The bottom panel shows the ratio between the data and the MC simulation. The blue band represents experimental systematic uncertainties, while the black error bars represent combined statistical uncertainties on both data and MC simulation.

4.5 Signal modelling

The $H \to \mu\mu$ signal is modelled using an analytical function with the shape parameters determined from the maximum-likelihood fit to the MC simulation. In the combined signal and background fit to the data, the shape parameters are kept fixed and only the normalisation is adjusted to the data.

Since the analysis measures the signal strength multiplier, defined with respect to the expected SM value, it is important to evaluate the theoretical and experimental systematic uncertainties which affect the signal normalisation. Furthermore, since analysis sensitivity depends greatly on the distribution of the signal among the different categories, the systematic uncertainties which affect the kinematics of events, not just their total normalisation, need to be evaluated.

4.5.1 Signal parametrisation

In the Standard Model, the Higgs boson has a width of only 4.1 MeV, negligible compared to the detector resolution of a few GeV in this momentum range. As a result, the signal shape is dominated by the detector resolution effects and to some extent the quality of final-state radiation recovery.

It is modelled using a double-sided Crystal-Ball (DSCB) function [154], which has a Gaussian core and tails on both sides, and is defined as [155]

$$\mathrm{DSCB} = N \cdot \begin{cases} e^{-\frac{1}{2}t^2} & \text{for } -\alpha_{\mathrm{low}} \leq t \leq \alpha_{\mathrm{high}} \\ e^{-\frac{1}{2}\alpha_{\mathrm{low}}^2} \left[\frac{\alpha_{\mathrm{low}}}{n_{\mathrm{low}}} \left(\frac{n_{\mathrm{low}}}{\alpha_{\mathrm{low}}} - \alpha_{\mathrm{low}} - t \right) \right]^{-n_{\mathrm{low}}} & \text{for } t < -\alpha_{\mathrm{low}} \\ e^{-\frac{1}{2}\alpha_{\mathrm{high}}^2} \left[\frac{\alpha_{\mathrm{high}}}{n_{\mathrm{high}}} \left(\frac{n_{\mathrm{high}}}{\alpha_{\mathrm{high}}} - \alpha_{\mathrm{high}} + t \right) \right]^{-n_{\mathrm{high}}} & \text{for } t > \alpha_{\mathrm{high}}, \end{cases} \tag{4.3}$$

where $t = (m_{\mu\mu} - \mu_{\mathrm{CB}})/\sigma_{\mathrm{CB}}$, N is the normalisation constant, μ_{CB} and σ_{CB} are the mean and sigma of the central Gaussian distribution, α_{low} (α_{high}) parametrises the invariant mass where the function transitions to the n_{low} (n_{high}) power law on the low-mass (high-mass) side of the spectrum.

An alternative functional form, a sum of a Gaussian and a one-sided Crystal-Ball with the tail towards lower values of the mass has been tested with similar results. The double-sided Crystal-Ball model was chosen because its parametrisation has a single mean and width and is more stable during the fit.

The signal shape is obtained in each category separately by fitting the combined signal from all main production modes. An alternative approach, fitting production modes separately and combining them in the proportions predicted by the SM, yields very similar results and is not used for simplicity.

The detailed form of the signal shape is expected to have little impact on the result at the search stage. Nevertheless, injection tests were carried out, fitting the functional form to the MC simulated spectrum. The extracted signals matched the injection within 0.5%, equal to the statistical precision of the test.

The fitted values of the μ_{CB} parameter were between 124.5 and 124.8 GeV, while the σ_{CB} parameter varied between 2.5 and 3.0 GeV between the categories. Fitted DSCB models are shown in Figure 4.8 for the VBF High and 0-jet High categories.

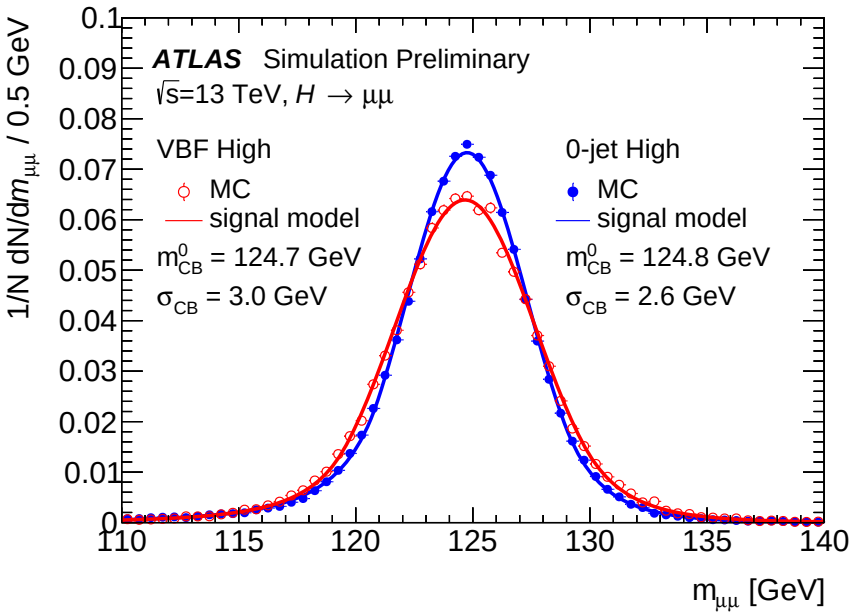


Figure 4.8: The DSCB signal model fit to the MC simulation of all Higgs production modes in the VBF High and 0-jet High analysis categories. VBF High category fit is shown as a red solid line, while the MC simulation is shown as empty red markers. The fit in the 0-jet High category is shown as a blue line, while the MC simulation is shown as solid blue circles. The fitted values of μ_{CB} and σ_{CB} parameters are shown on the left (right) side for the VBF High (0-jet High) category. Reproduced from Ref. [6].

4.5.2 Systematic uncertainties

The signal normalisation and its distribution between the categories are affected by both the theoretical and experimental systematic uncertainties.

Theoretical uncertainties affecting the signal normalisation come from the effect of missing higher-order QCD corrections on the cross-section prediction, and the uncertainty on the branching ratio to muon pairs [23].

Some theoretical uncertainties also have an impact on the migration of events between the categories, either because of the effect on the muon or jet kinematics. These include the QCD scale variations, the uncertainties on the treatment of the underlying event and hadronisation, the uncertainties on the p_{T} of the Higgs boson, the uncertainties on the migrations between jet multiplicities and the treatment of top quark mass, following the prescription in Ref. [23]. The uncertainties on the underlying event are evaluated by considering Pythia 8 eigentune variations and those on showering by considering events showered by Pythia 8 and Herwig 7 [156, 157]. Dedicated uncertainties are also computed for the acceptance of the ggF events in the VBF categories.

The individual size of theoretical uncertainties ranges between per mille level and 15% for the ggF production mode, while the uncertainties on the VBF production modes range between per mille level and 7%.

Experimental systematic uncertainties affect the normalisation and shape of the signal events as well as their distribution between the categories. For muons, the normalisation uncertainties come from the uncertainty in the scale factors for the reconstruction and identification efficiencies, trigger, isolation selection, and the impact parameter efficiencies. The only uncertainties which significantly alter the shape of the signal are the uncertainties on the muon scale and resolutions, which affect the position and the width of the Higgs boson mass peak.

Systematic uncertainties on other objects used in the analysis also have an effect on the analysis sensitivity, mainly by affecting the distribution of signal events between the analysis categories. In particular, $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term uncertainties and uncertainties on the jet energy scale and resolution have an important impact on the distribution of events in the VBF categories in the 2-jet channel. Pile-up modelling has an impact on migration of events between jet multiplicity channels,

while the uncertainties on the b-tagging efficiency affect the number of events that pass the analysis selection.

Overall, the largest experimental uncertainty comes from the jet energy scale and resolution, affecting signal yields in some of the 2-jet categories by up to 10%. The uncertainty from the muon momentum resolution affects the fitted yields at the 1–6% level, depending on the category.

Additionally, an uncertainty on the Higgs boson mass from Ref. [19] is found to have a small impact. Finally, the uncertainty on the combined integrated luminosity of the dataset is 1.7% [158, 159].

4.6 Background modelling

The very small signal to background ratio in the central region means that the modelling of the background is critical to the success of the analysis. Since the number of background events is much larger, a very small bias in the background model would greatly impact the extracted signal.

The signal to background ratio in the Central region is 0.2% inclusively, though this improves in some analysis categories. This requires the control of the background shape to sub per mille level.

Analytical functions, motivated by the composition of the background spectrum (dominated by DY, smaller contributions from diboson and top), have been tried in the past but they have struggled in particular to model the steeply falling spectrum towards the Z mass peak. In this analysis, a combination of a rigid core and an empirical function is used instead.

The core function, common to all analysis categories and without free parameters, models the steep part of the spectrum. The empirical function models the differences in categories due to selection and resolution effects, and is selected using the custom DY dataset, described in Section 4.1.3. The value of the spurious signal uncertainty is extracted simultaneously.

4.6.1 Background parametrisation

The background model is constructed in each analysis category separately, by multiplying a common core function with an empirical function specific to the

category,

$$F(m_{\mu\mu}) = F_{\text{core}}(m_{\mu\mu}) \cdot F_{\text{empirical}}(m_{\mu\mu}). \tag{4.4}$$

The core function describes the inclusive DY spectrum and is able to model the steeply falling part of the spectrum. It is constructed from the analytic LO Drell-Yan line-shape [160], convolved by the normal distribution with width parametrised as a function of $m_{\mu\mu}$ to model the effects of detector resolution,

$$F_{\text{core}}(m_{\mu\mu}) = \left(\sum_{q} \mathcal{L}_{q\bar{q}}(m_{\mu\mu}) \cdot \sigma_{q\bar{q}}(m_{\mu\mu}) \right) * \mathcal{N}(0, \sigma(m_{\mu\mu})), \tag{4.5}$$

where the sum over q runs over the lightest three quarks, $q = u, s, d$, the parton luminosity component $\mathcal{L}_{q\bar{q}}$ is derived using PDF4LHC15 PDF for the LO using APFEL [161] and LHAPDF [162]. The cross-section component $\sigma_{q\bar{q}}(\hat{s}) = \sigma_{q\bar{q}}(m_{\mu\mu})/(2m_{\mu\mu})$ can be written as [6]

$$\sigma_{q\bar{q}}(\hat{s}) = \frac{4\pi\alpha^2}{3\hat{s}N_c} \left[Q_q^2 - 2Q_q V_\ell V_q \chi_{Z\gamma}(\hat{s}) + (A_\ell^2 + V_\ell^2)(A_q^2 + V_q^2)\chi_Z(\hat{s}) \right], \tag{4.6}$$

where

$$\chi_{Z\gamma}(\hat{s}) = \kappa \frac{\hat{s}(\hat{s} - m_Z^2)}{(\hat{s} - m_Z^2)^2 + \Gamma_Z^2 m_Z^2}, \tag{4.7}$$

$$\chi_Z(\hat{s}) = \kappa^2 \frac{\hat{s}^2}{(\hat{s} - m_Z^2)^2 + \Gamma_Z^2 m_Z^2}, \tag{4.8}$$

$$\kappa = \frac{\sqrt{2} G_F m_Z^2}{4\pi\alpha}, \tag{4.9}$$

and $\hat{s} = m_{\mu\mu}^2$, Q_q is the electric charges of the quarks, V_f and A_f are the vector and axial-vector couplings of the fermions, α, and G_F are the electroweak couplings, m_Z and Γ_Z are the mass and the width of the Z boson with values taken from Ref. [163], and $N_c = 3$ is the number of colour charges.

The resolution function $\sigma(m_{\mu\mu})$ is derived from the MC simulation of Drell-Yan samples, simulated by Sherpa, by comparing the truth and reconstructed invariant mass.

The empirical function is chosen among two families of functions, namely "PowerN",

$$\text{PowerN}(m_{\mu\mu}) = m_{\mu\mu}^{(a_0 + a_1 m_{\mu\mu} + a_2 m_{\mu\mu}^2 + \ldots + a_N m_{\mu\mu}^N)}, \tag{4.10}$$

and "EpolyN",

$$\text{EpolyN}(m_{\mu\mu}) = e^{(a_1 m_{\mu\mu} + a_2 m_{\mu\mu}^2 + \ldots + a_N m_{\mu\mu}^N)}, \tag{4.11}$$

where a_n are the free parameters, according to the selection procedure outlined in Section 4.6.2.

4.6.2 Model selection

The empirical function is chosen separately in each analysis category, by first reducing the set of functions to those satisfying the following criteria:

- The χ^2 probability of the background-only fit must be larger than 1% for:
  - the data in the fit region (excluding the range $120 - 130$ GeV),
  - the ATLAS MC simulation of DY, diboson, and top backgrounds in the fit region,
  - the custom DY sample, reweighted to the data sidebands using a second order polynomial.
- The spurious signal test, as defined in Section 4.6.3, is satisfied: SS $< 20\%$ δS, where SS is the fitted signal yield (the spurious signal), and δS is the statistical uncertainty on the signal strength obtained from the data corresponding to 139 fb^{-1}.
- If multiple functions pass the first two criteria, the function(s) with the fewest number of free parameters is (are) chosen.
- If more than one function remains, the function with the smallest value of SS is chosen.

The selected functions are shown in Table 4.8 for each of the analysis categories. The spurious signal uncertainty is considered as uncorrelated across the categories since none of the values are more than two standard deviations away from zero, when considering the statistical uncertainty on the custom DY sample, corresponding to approximately 700 times the statistics of the dataset.

Table 4.8: The selected empirical function components of the background model in each of the analysis categories. Also shown are the SS values, defined as the maximum of the fitted signals in the 120–130 GeV mass range, normalised to the expected statistical error on the fitted signal (δS and to the SM predictions for signal yields in the analysis categories (S_{SM}). Reproduced from Ref. [6].

Category	Empirical Function	max(SS/δS)[%]	max(SS/S_{SM})[%]
VBF High	Power0	10.6	14.7
VBF Medium	Epoly2	0.51	1.3
VBF Low	Power1	3.6	7.5
2-jet High	Epoly2	8.7	16.3
2-jet Medium	Epoly4	1.2	3.9
2-jet Low	Epoly3	-8.2	-33.2
1-jet High	Power1	6.1	12.1
1-jet Medium	Epoly3	-8.1	-19.8
1-jet Low	Epoly3	-2.5	-5.8
0-jet High	Power1	14.6	26.5
0-jet Medium	Epoly3	-11.6	-39.0
0-jet Low	Epoly3	-18.5	-74.2

4.6.3 Spurious signal uncertainty

The spurious signal value is determined by a combined signal and background model fit to the invariant mass spectrum in the fit region of the custom DY sample, shown in Figure 4.9. Prior to the fit, the DY sample is re-weighted to the data sidebands using a first-order polynomial in the VBF categories, and a second-order polynomial in the other categories. The re-weighting mitigates the potential mismodelling of the invariant mass shape by the MC simulation.

A number of scans are performed, varying the signal mass parameter between 120 and 130 GeV in 1 GeV steps, while the other parameters of the signal model are kept to the values obtained in the fit. The maximum value of the fitted signal across this range is taken as the total spurious signal, denoted as SS.

The sensitivity of the SS value to the shape of the custom DY spectrum was evaluated by fitting the combined signal and background model to the systematic variations of the custom DY spectrum. For theoretical systematic uncertainties, the variations of the QCD renormalisation scale and using alternative PDF sets

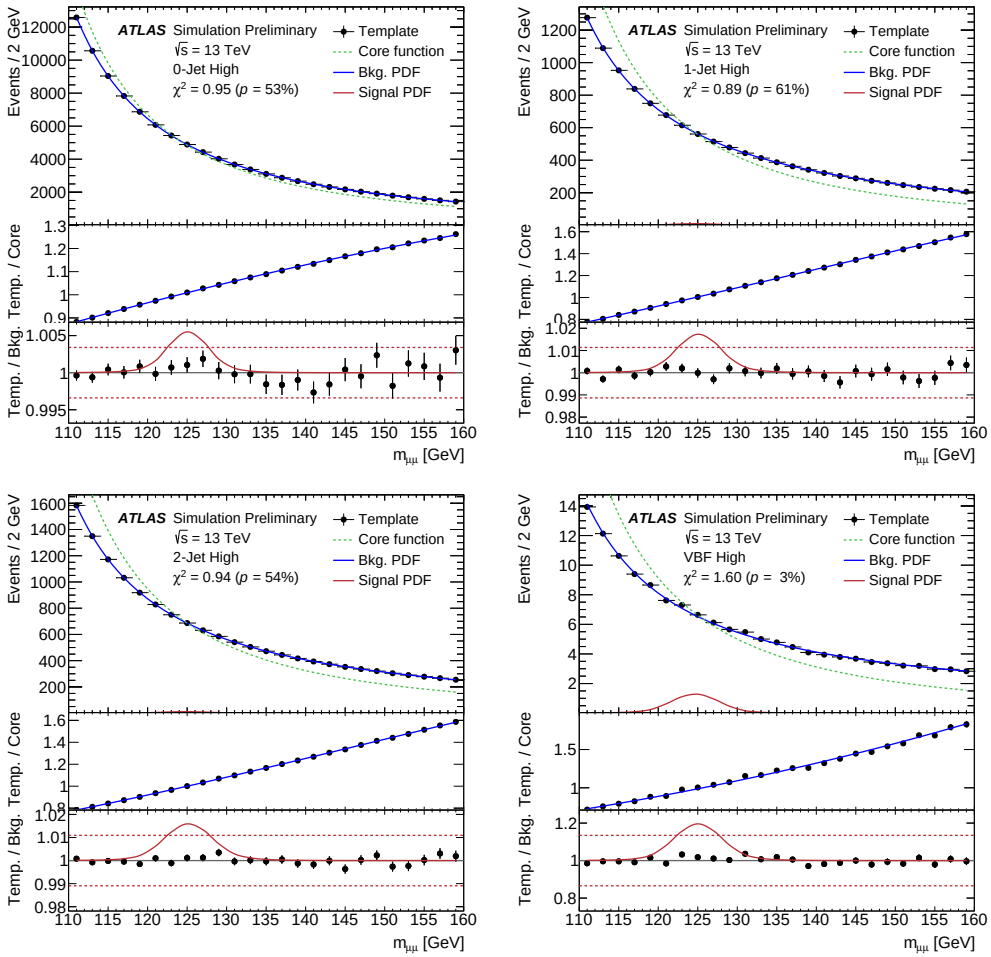


Figure 4.9: The combined signal and background fit to the custom DY invariant mass spectrum in 0-jet High, 1-jet High, 2-jet High, and VBF High analysis categories. The top panel shows the custom DY template as black points, the fixed core function as a dashed green line, and the full background model as a solid blue line. The signal shape, normalised to the SM prediction, is shown as a red line and only clearly visible in the VBF High category. The reduced χ^2 value of the fit and the corresponding χ^2-probability are also shown. The middle panel shows the custom DY template, divided by the core function (black points) and the empirical function (blue solid line). The bottom panel shows the ratio between the custom DY template and full background model as black points. the signal shape, divided by the background model is shown as a solid red line. The dashed red lines in the bottom panel denote the expected signal purity in the central region. Reproduced from Ref. [6].

were considered. From the experimental side, the systematic variations of muon momentum scale and resolution in the ID and MS detectors were used. The obtained SS values from the systematic variations were not significantly different from those obtained from the nominal template.

It should be noted that the custom DY samples are only used in selecting the empirical part of the function and the evaluation of the spurious signal systematic uncertainty. They are not used at all when fitting the invariant mass spectrum in data.

4.7 Results

The analysis of the data is based on a binned likelihood model for the observed number of events in the invariant mass spectrum of the analysis categories. The approach is mostly frequentist, though it deviates from the paradigm to accommodate practical needs. Most notably, the systematic uncertainties are treated using constraint terms, even though the underlying uncertainty may not be a result of limited statistical accuracy of an auxiliary measurement, as is the case with some theoretical parameters. Furthermore, when setting the upper limits on the signal strength, a modification of the standard construction of confidence intervals is used in order to prevent exclusions of arbitrarily small signals as a result of downward fluctuations in the background.

The results are obtained by a combined signal and background binned maximum-likelihood fit, performed simultaneously in the twelve analysis categories. The fit is performed in the 110–160 GeV region, with the bin size of 0.1 GeV, chosen as a suitable trade-off between sensitivity and computation time.

4.7.1 Likelihood function

The parameter of interest is the signal strength, μ, a single multiplicative factor of the expected number of signal events, normalised such that $\mu = 0$ corresponds to zero observed signal events, and $\mu = 1$ corresponds to the expected number of events from the Standard Model. The nuisance parameters are denoted by θ and include the parameters of the background model and the systematic uncertainties

affecting the normalisation, shape, and distribution of signal events between the analysis categories. The observed data, representing the observed events binned in the invariant mass, is denoted by $\mathcal{D} = \{n_{cb}\}$, where n_{cb} is the number of data events in bin b of category c.

The full likelihood function is the product of likelihoods in each of the categories, multiplied by the constraint terms for the nuisance parameters,

$$\mathcal{L}(\mu, \theta|\mathcal{D}) = \prod_{c \in \text{categories}} \mathcal{L}_c(\mu, \theta|\mathcal{D}) \cdot \prod_{p \in \mathcal{S}} f_p(a_p|\alpha_p), \tag{4.12}$$

where $\mathcal{S}$ is the set of parameters p with constraint terms, f_p is the Gaussian or log-normal constraint term for parameter p with value a_p, and parametrised by α_p. For example, parameter p can be the muon trigger efficiency scale factor with value a_p and is constrained by a Gaussian term f_p with α_s being the nominal value of the scale factor and its uncertainty, obtained from previous a previous dedicated measurement performed by the muon performance group.

The likelihood function in the category c takes the form

$$\mathcal{L}_c(\mu, \theta|\mathcal{D}) = \prod_{b \in \text{bins}} \text{Pois}(n_{cb}|\nu_{cb}(\mu, \theta)), \tag{4.13}$$

where $\text{Pois}(n|\nu)$ is the Poisson probability distribution with the observed number of events n, and the rate parameter ν. The rate parameter $\nu_{cb}(\mu, \theta)$ takes the form

$$\nu_{cb}(\mu, \theta) = \mu S_{cb}(\theta) + B_{cb}(\theta), \tag{4.14}$$

where $S_{cb}(\theta)$ is the expected number of signal events in the bin b of category c, obtained by integrating the signal model in Eq. 4.3, and depends on the nuisance parameters θ, for example the muon momentum scale and resolution corrections. $B_{cb}(\theta)$ is the expected number of background events in the bin b of category c, and also depends on the nuisance parameters, most importantly on the parameters of the background model described in Eq. 4.4.

4.7.2 Statistical treatment

The point estimate of the signal strength, $\hat{\mu}$, is obtained by numerically minimising the value of $-\ln \mathcal{L}(\mu, \theta|\mathcal{D})$, and is found to be 0.5. The values of nuisance parameters at this point are denoted as $\hat{\theta}$.

The results of the maximum-likelihood fit are shown in Figure 4.11 for the "High" categories, Figure 4.12 for the "Medium" categories, and in Figure 4.13 for the "Low" categories. Figure 4.14 shows the combined fit in all analysis categories, both unweighted and weighted by $\log(1 + S/B)$ per category, where S and B are the number of signal and background events in the Central region in the category.

Table 4.9 shows the number of observed data events in each of the analysis categories, along with the number of signal events predicted by the Standard Model, and the values of signal and background events obtained from the maximum-likelihood fit. The significance proxy $S/\sqrt{B}$ and the signal purity (S/B) are also shown.

Table 4.9: Number of events in the central region in all the analysis categories, as observed in data (Data), predicted for signal by the Standard Model (S_{SM}), and events from signal (S) and background (B) derived from the results of the maximum-likelihood fit. Additionally, the number of observed signal events over the square root of the number of observed background events ($S/\sqrt{B}$) and the signal purity (S/B) are shown. Reproduced from Ref. [6].

Category	Data	S_{SM}	S	B	$S/\sqrt{B}$	S/B [%]
VBF High	40	4.5	2.3	34	0.39	6.6
VBF Medium	109	5.5	2.8	100	0.28	2.8
VBF Low	450	9.6	4.9	420	0.24	1.2
2-jet High	3400	38	19	3440	0.33	0.6
2-jet Medium	13938	70	35	13910	0.30	0.3
2-jet Low	40747	75	38	40860	0.19	0.1
1-jet High	2885	32	16	2830	0.31	0.6
1-jet Medium	24919	107	54	24890	0.35	0.2
1-jet Low	77482	134	68	77670	0.24	0.1
0-jet High	24777	85	43	24740	0.27	0.2
0-jet Medium	85281	155	79	85000	0.27	0.1
0-jet Low	180478	144	73	180000	0.17	<0.1

In order to construct the confidence intervals on μ, a hypothesis test is conducted at each value of μ. This is done with the help of a test statistic, a mapping from the data to a real number. First, a profile likelihood ratio is defined as

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \tag{4.15}$$

where $\hat{\hat{\theta}}(\mu)$ are the nuisance parameters which maximise the likelihood with μ fixed, called the conditional maximum-likelihood estimates. Clearly, $0 \leq \lambda(\mu) \leq 1$, with $\lambda(\mu)$ values near 1 signalling good agreement between the data and the hypothesised μ.

The test statistic t_μ is defined for each hypothesised value of μ,

$$t_\mu = -2 \ln \lambda(\mu), \tag{4.16}$$

with higher values of t_μ corresponding to less compatibility between the data and the hypothesised μ, either because $\hat{\mu}$ is greater or smaller than the hypothesis μ. The level of disagreement between the hypothesis μ and the data is quantified using the p-value,

$$p_\mu = \int_{t_{\mu,\text{obs}}}^{\infty} f(t_\mu|\mu) dt_\mu, \tag{4.17}$$

where $t_{\mu,\text{obs}}$ is the value of the test statistic obtained from the data, and $f(t_\mu|\mu)$ is the probability distribution of the test statistic t_μ under the assumption μ. Instead of using the computationally expensive Monte Carlo simulated toy experiments to determine the distribution of the test statistic, a representative Asimov dataset [164] is used to determine the expected uncertainty on the signal strength.

Using the 68% confidence interval, the measured signal strength is found to be

$$\mu = 0.5 \pm 0.7(\text{stat}) \pm^{0.2}_{0.1} (\text{syst}) \tag{4.18}$$

where the uncertainty is dominated by the statistical uncertainty on the data. The impact of systematic uncertainties is $\pm^{0.2}_{0.1}$, dominated by the uncertainties on signal normalisation and shape from theoretical (0.08) and experimental (0.07) effects. The effect of the spurious signal uncertainty is 0.06. The expected and observed signal strengths in all twelve analysis categories are shown in Figure 4.15.

The leading systematic uncertainties are shown in Figure 4.10, ordered by their impact on the signal strength. The single most important systematic uncertainty is the uncertainty on the muon spectrometer momentum calibration, which affects the width of the signal shape in the fit. These are followed by a theoretical systematic uncertainties affecting the overall normalisation and distribution of signal between the analysis categories. Spurious signal systematics are denoted as `ATLAS_bias_BDT`*N*. Other important experimental systematic uncertainties are the

muon momentum resolution in the inner detector, and jet systematics which affect the yields in the VBF categories.

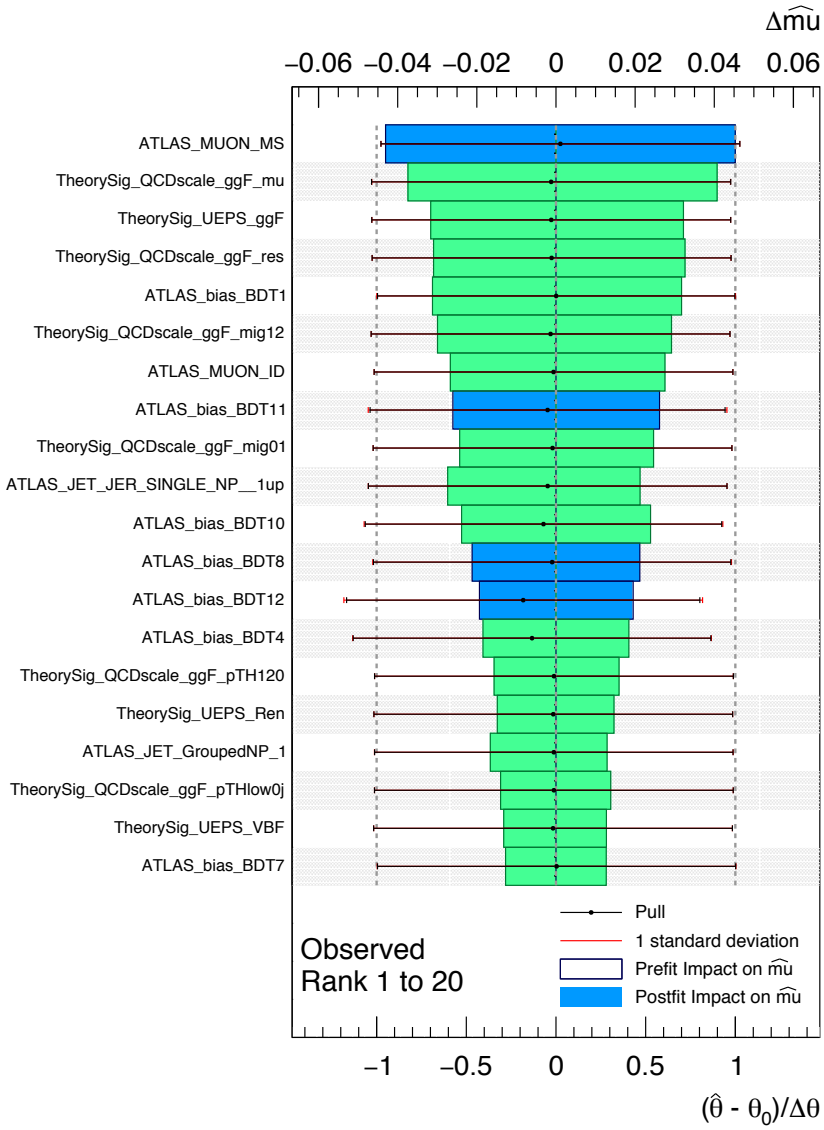


Figure 4.10: Systematic uncertainties affecting the measurement of the signal strength ordered by their impact. The coloured bars show the postfit impact on the signal strength and their size corresponds to the scale at the top of the figure. The pull is shown as black lines and the size corresponds to the scale at the bottom of the figure. Figure produced by a collaborator.

The asymmetry in the total systematic uncertainty arises as a combination

of systematic uncertainties, apart from the spurious signal uncertainty, being relative and the statistical uncertainties being much larger than systematic uncertainties. When the statistical uncertainty is varied towards higher values of the signal strength, the absolute size of the relative systematic uncertainty on the signal strength increases. This effect can be seen in Figure 4.15, where individual categories have clearly asymmetric systematic uncertainties.

In the measurement of the signal strength, the parameter of interest was allowed to take either positive or negative values. It is also possible to restrict the signal strength to non-negative values, and perform a one-sided hypothesis test. However, the estimator $\hat{\mu}$ is still allowed to take on negative values and the restriction on μ is implemented through a different test statistic. When the best fit value $\hat{\mu} < 0$, the best agreement between the data and the physically allowed signal strength is at $\mu = 0$, defining $\tilde{\lambda}(\mu)$ as

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\ \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0, \end{cases} \tag{4.19}$$

where $\hat{\hat{\theta}}(0)$ are the values of nuisance parameters that maximise the likelihood conditioned on $\mu = 0$. Using $\tilde{\lambda}(\mu)$, $\tilde{t}_\mu$ can be defined as $\tilde{t}_\mu = -2 \ln \tilde{\lambda}(\mu)$. The background-only hypothesis can then be tested by considering the special case of $\tilde{t}_\mu$ with $\mu = 0$,

$$\tilde{t}_0 = \begin{cases} -2 \ln \tilde{\lambda}(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0. \end{cases} \tag{4.20}$$

It can be seen from the definition that large values of $\tilde{t}_0$ correspond to poor agreement with data and are only achieved for $\hat{\mu} > 0$, as expected from a one-sided test. The p-value is then computed using

$$p_0 = \int_{\tilde{t}_{0,\text{obs}}}^{\infty} f(\tilde{t}_0|0) d\tilde{t}_0, \tag{4.21}$$

where $t_{0,\text{obs}}$ is the test statistic obtained from the data, and $f(\tilde{t}_0|0)$ is the distribution of the test statistic under the background-only hypothesis. The p-value is then converted to the quantile of the normal distribution, or number of "sigma" (Z), using

$$Z = \Phi^{-1}(1 - p_0), \tag{4.22}$$

where Φ^{-1} is the inverse of the cumulative distribution for the normal distribution. The results of the test are

$$Z = 0.8\ (1.5)\ \sigma \text{ observed (expected)}, \tag{4.23}$$

where the expected value is obtained using the Asimov dataset.

It is also possible to reverse the situation and test a combined signal and background hypotheses, parametrised by the value of μ, which is restricted to be positive. A test statistic t'_μ is defined as

$$t'_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu, \end{cases} \tag{4.24}$$

reflecting the fact that the data with $\hat{\mu} > \mu$ do not constitute evidence against hypothesis μ. It can be seen from the definition that large values of t'_μ correspond to the data less compatible with the hypothesis. Therefore, the p-value is computed as

$$p_\mu = \int_{t'_{\mu,\text{obs}}}^{\infty} f(t'_\mu|\mu)dt'_\mu, \tag{4.25}$$

where $t'_{\mu,\text{obs}}$ is the value of the test statistic observed in the data, and $f(t'_\mu|\mu)$ is the distribution of the test statistic under the assumption that the data comes from the hypothesis μ.

The standard 95% confidence level (CL) upper limit is obtained by solving for μ which results in $p_\mu = 0.05$. However, in the case of a downward fluctuation of the background this results in strong limits on the signal strength, which is undesirable. As a remedy, an alternative quantity p' is defined as

$$p' = \frac{p_\mu}{p_b} \tag{4.26}$$

where p_b is the p-value under the background-only assumption, i.e. $p_b = p_{\mu=0}$. The so-called CL_s upper limits at 95% CL are obtained by solving for μ which results in $p' = 0.05$ [164, 165] and prevents the exclusion of small signals as a result of the background underfluctuation. However, the excluded signal strength region loses the confidence interval interpretation. While the coverage for large μ tends to 95% as expected, for small μ the coverage approaches 100%.

The expected limits are also computed using the Asimov dataset corresponding to the $\mu = 0$ and $\mu = 1$ assumptions. Using the CL_s method, the limits were found to be

$$95\%\ \mathrm{CL}_s \text{ upper limit on } \mu = \begin{cases} 1.7 & \text{observed} \\ 1.3 & \text{expected } (\mu = 0) \\ 2.2 & \text{expected } (\mu = 1) \end{cases} \tag{4.27}$$

Assuming the SM values for the production cross-section the corresponding observed limit on the branching ratio to muon pairs is $\mathrm{BR}(H \to \mu\mu) < 3.8 \cdot 10^{-4}$.

Key numerical results are summarised in the box below.

$$\mu = 0.5 \pm 0.7(\text{stat}) \pm^{0.2}_{0.1} (\text{syst})$$

$$Z = 0.8\ (1.5)\ \sigma \text{ observed (expected)}$$

$$95\%\ \mathrm{CL}_s \text{ upper limit on } \mu = \begin{cases} 1.7 & \text{observed} \\ 1.3 & \text{expected } (\mu = 0) \\ 2.2 & \text{expected } (\mu = 1) \end{cases}$$

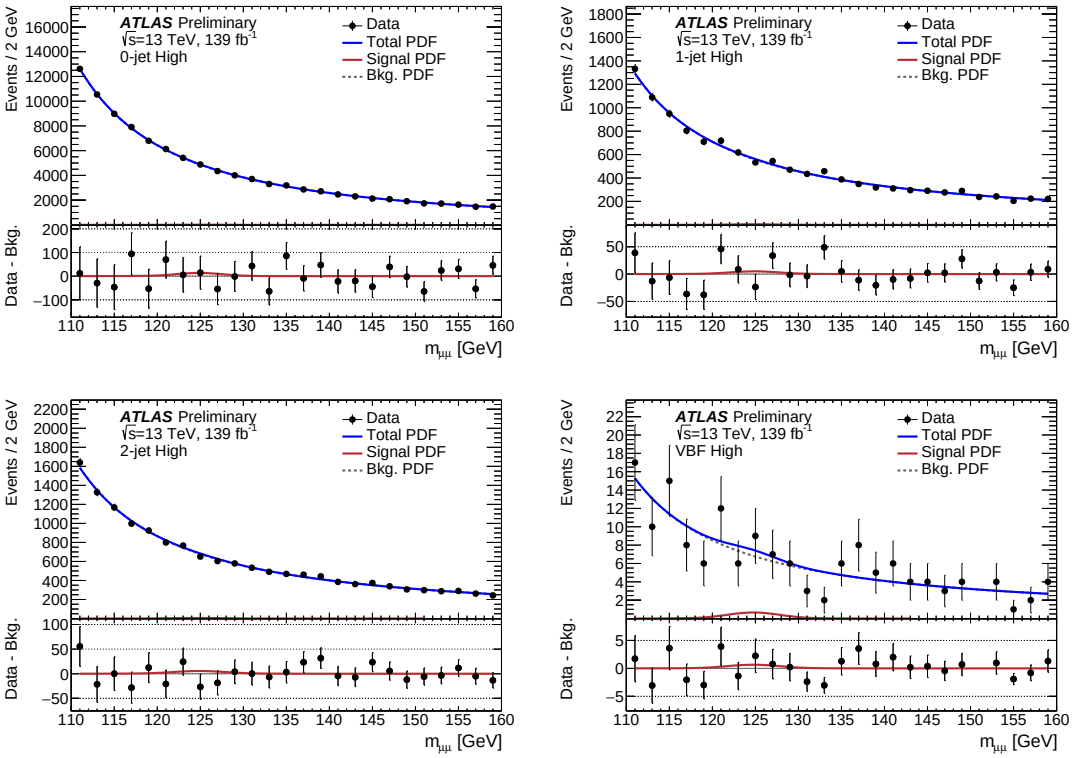


Figure 4.11: The combined signal and background fits to the invariant mass after FSR corrections in data in the 0-jet High, 1-jet High, 2-jet High, and VBF High categories. The top panel shows the data (black markers), the full background model (dashed black line), the signal model (solid red line) and the combined signal and background model (solid blue line). The bottom panel shows the difference between the data and the full background model as black markers, and the signal model as a solid red line. The error bars represent statistical errors on the data. The width of the bins is 2 GeV, while the fit is performed in bins of 0.1 GeV. Reproduced from Ref. [6].

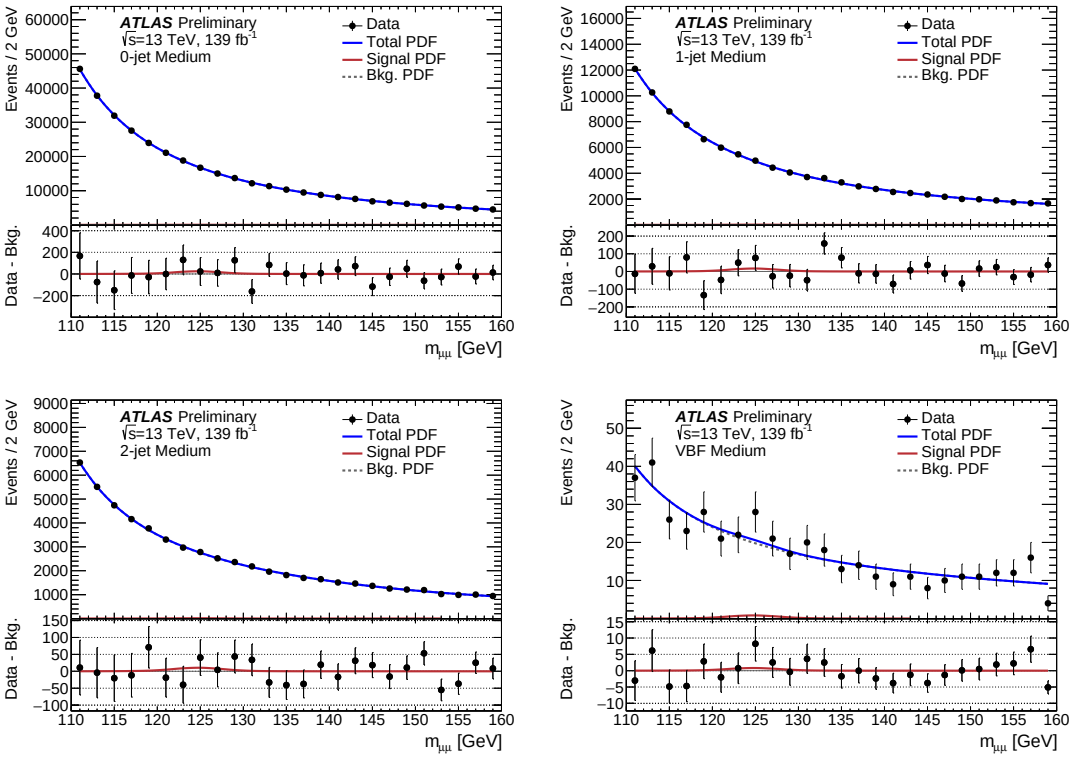


Figure 4.12: The combined signal and background fits to the invariant mass after FSR corrections in data in the 0-jet Medium, 1-jet Medium, 2-jet Medium, and VBF Medium categories. The top panel shows the data (black markers), the full background model (dashed black line), the signal model (solid red line) and the combined signal and background model (solid blue line). The bottom panel shows the difference between the data and the full background model as black markers, and the signal model as a solid red line. The error bars represent statistical errors on the data. The width of the bins is 2 GeV, while the fit is performed in bins of 0.1 GeV. Reproduced from Ref. [6].

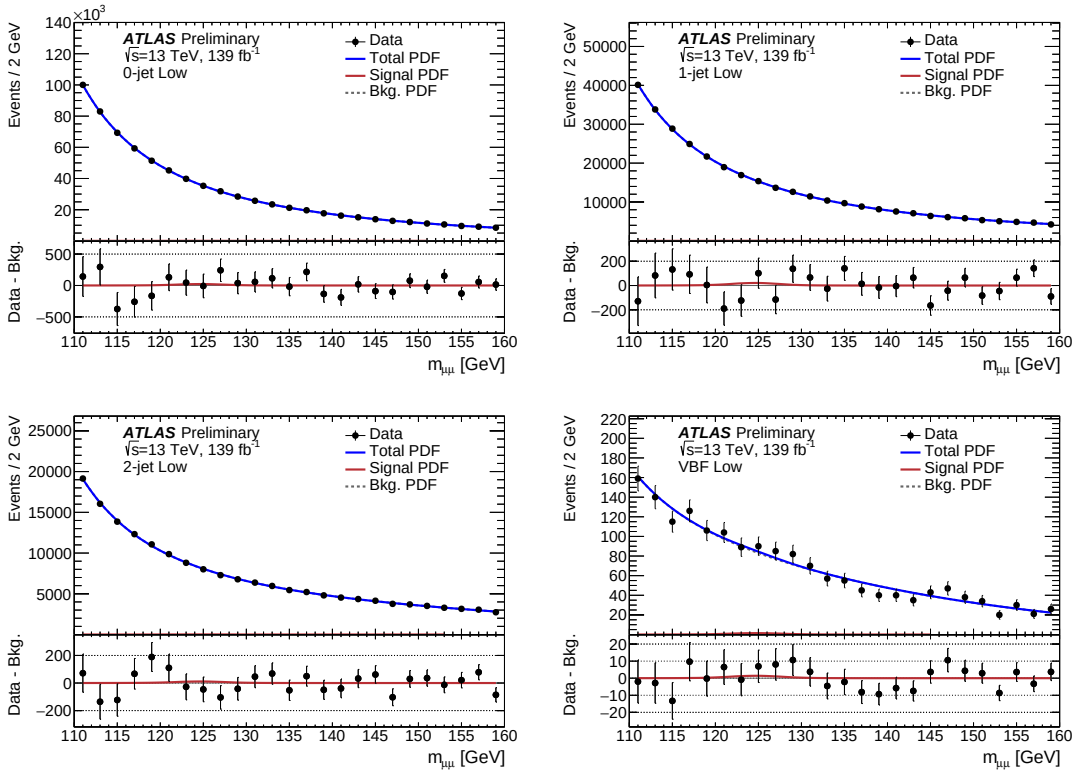


Figure 4.13: The combined signal and background fits to the invariant mass after FSR corrections in data in the 0-jet Low, 1-jet Low, 2-jet Low, and VBF Low categories. The top panel shows the data (black markers), the full background model (dashed black line), the signal model (solid red line) and the combined signal and background model (solid blue line). The bottom panel shows the difference between the data and the full background model as black markers, and the signal model as a solid red line. The error bars represent statistical errors on the data. The width of the bins is 2 GeV, while the fit is performed in bins of 0.1 GeV. Reproduced from Ref. [6].

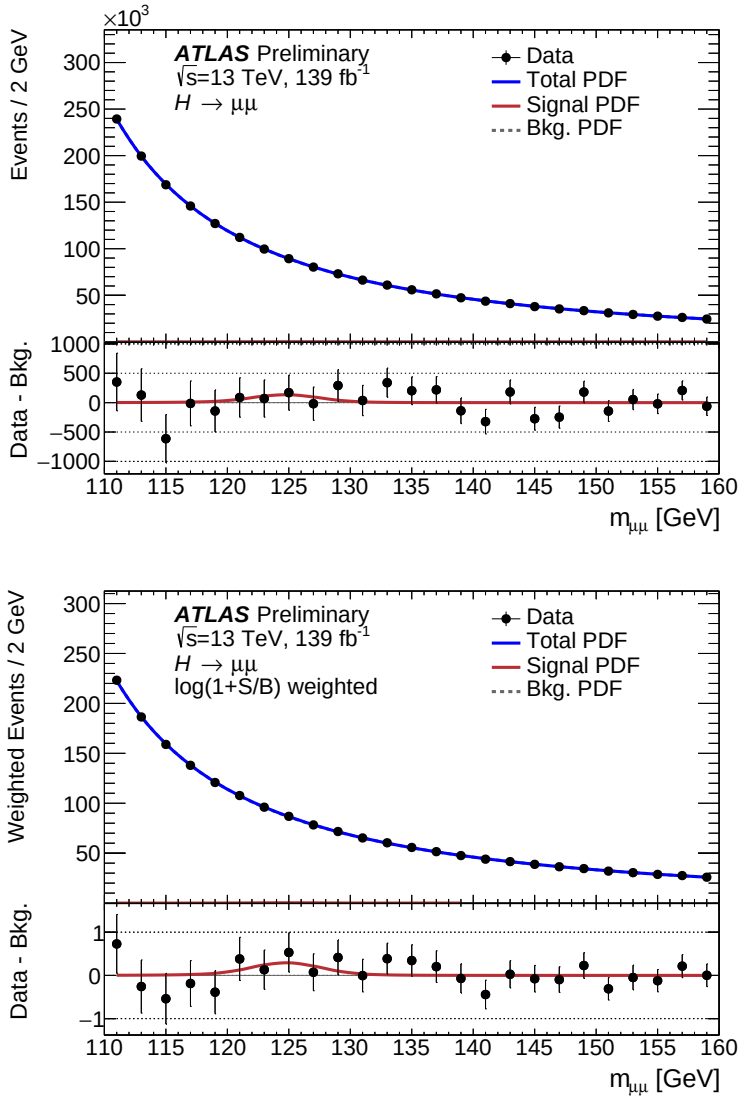


Figure 4.14: The combined signal and background fits to the invariant mass after FSR corrections in data, combined in all analysis categories. The unweighted combination is shown in the top figure, while the bottom figure shows the events weighted by the $\log 1 + S/B$ value of each category. The top panel shows the data (black markers), the full background model (dashed black line), the signal model (solid red line) and the combined signal and background model (solid blue line). The bottom panel shows the difference between the data and the full background model as black markers, and the signal model as a solid red line. The error bars represent statistical errors on the data. The width of the bins is 2 GeV, while the fit is performed in bins of 0.1 GeV. Reproduced from Ref. [6].

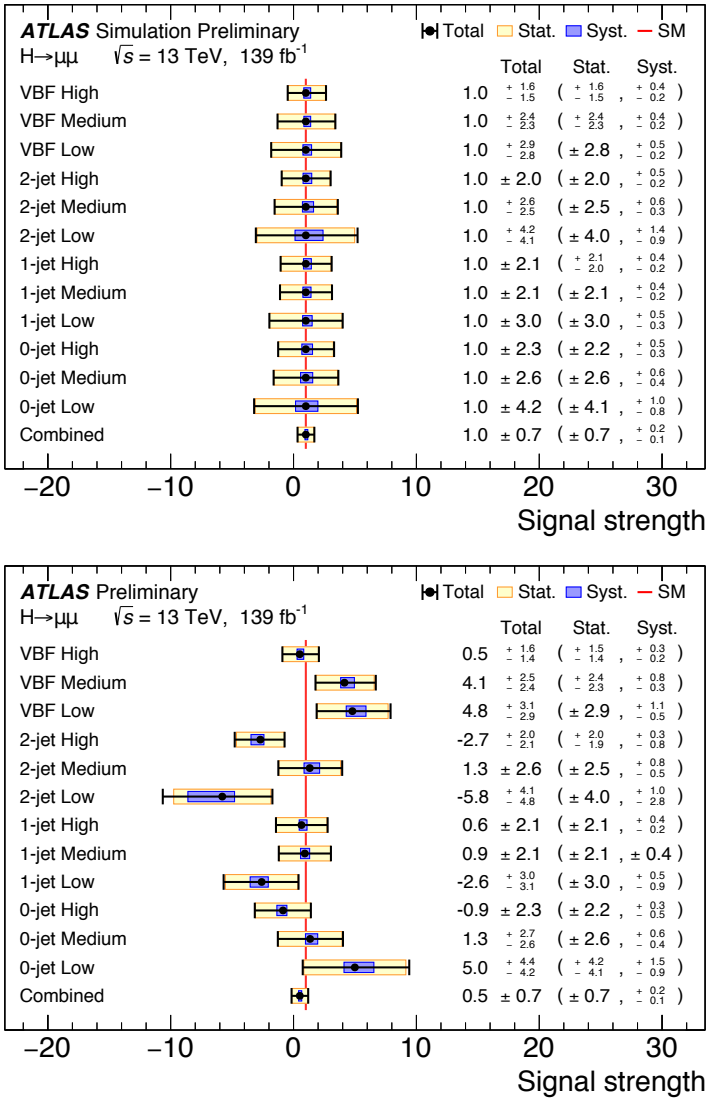


Figure 4.15: The expected (top) and observed (bottom) signal strengths in all twelve analysis categories. The central value and total uncertainty are shown in black, the statistical uncertainties are shown in yellow, and the systematic uncertainties in blue. The Standard Model prediction is shown as a red line. Reproduced from Ref. [6].

Chapter 5

Outlook

"Prediction is very difficult, especially about the future."

— Niels Bohr

Sensitivity projections are an essential piece of information for making informed decisions about the potential future projects the community should undertake. It is therefore important to assess what can be achieved with the LHC machine, and in particular what the physics reach of the full dataset will be. The dataset is expected to be recorded at 14 TeV and to correspond to 3000 fb^{-1} of integrated luminosity, and is known as the High-Luminosity LHC programme (HL-LHC).

As a part of the community-wide projections project [7, 8] an extrapolation of results described in Ref. [5] was made. At the time, this was the state-of-the-art analysis. The general approach of the analysis used in the extrapolation is not too different from the one described in this thesis. However, a number of important improvements presented in this thesis, most notably in categorisation and background modelling, that result in approximately 25% improvement in sensitivity of the analysis, are not taken into account in the projection.

In addition to the increase in integrated luminosity from 80 to 3000 fb^{-1}, increases in the production cross-sections from 13 to 14 TeV are also taken into account. The signal yields are scaled according to the values in Ref. [23], separately for the ggF and VBF production modes. The background yields are scaled according to the parton luminosity ratio as described in Ref. [166], taking into account that the backgrounds are produced predominantly via the interaction of quarks.

The performance of the reconstruction of physics objects is assumed to remain unchanged to simplify the extrapolation. While the higher pile-up is likely to degrade performance this is expected to be roughly balanced by the improved performance of the ATLAS detector. Muon resolution is an exception to this rule because of the large expected improvements in performance of the ATLAS Inner Tracker (ITk) upgrade. To model this improvement the signal widths are reduced by 30% in the VBF and Central analysis categories and by 15% in the Forward categories to emulate the expected improvements [167].

Two scenarios are considered regarding the treatment of systematic uncertainties. Scenario 1 (S1) assumes the relative systematic uncertainties remain as they are today, while in Scenario 2 (S2) most of the theoretical uncertainties on the signal normalisation are halved, with the exception of the PDF uncertainties for which the reduction factors depend on the quark/gluon initial states [168]. The

spurious signal uncertainty is assumed to be negligible in both the S1 and S2, while the luminosity uncertainty is set to 1%.

Table 5.1 shows the comparison of the expected uncertainty on the signal strength for the Run 2 analysis with 79.8 fb^{-1} and the S1 and S2 scenarios at the HL-LHC. It can be seen that even with the full HL-LHC dataset the analysis will be limited by the statistical uncertainty on data. Not taken into account in the results in Table 5.1 is an additional 25% improvement in sensitivity that can be expected from the improved background modelling and categorisation presented in this thesis.

Table 5.1: The expected uncertainties on the signal strength measurement with the Run 2, 79.8 fb^{-1} dataset and the S1 and S2 systematic uncertainty scenarios for the HL-LHC dataset. Columns show the total uncertainty, statistical uncertainty, experimental systematic uncertainty, and the uncertainty on the signal normalisation. Reproduced from Ref. [7].

Scenario	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\mu_{\text{sig}}}/\sigma_{\text{SM}}$
Run 2, 79.8 fb^{-1}	$^{+1.04}_{-1.06}$	$^{+0.99}_{-1.03}$	$^{+0.03}_{-0.03}$	$^{+0.32}_{-0.27}$
HL-LHC S1	$^{+0.15}_{-0.14}$	$^{+0.12}_{-0.13}$	$^{+0.03}_{-0.03}$	$^{+0.08}_{-0.05}$
HL-LHC S2	$^{+0.13}_{-0.13}$	$^{+0.12}_{-0.13}$	$^{+0.02}_{-0.02}$	$^{+0.05}_{-0.04}$

This is in good agreement with a previous study of $H \to \mu\mu$ prospects [169] based on generator-level MC simulation. That study simulated a dataset equivalent to 3000 fb^{-1} collected at 14 TeV with the average number of interactions per bunch crossing $\langle\mu\rangle = 200$, and the most up-to-date detector performance. The projected accuracy of the signal strength measurement was found to be 13%, while the discovery significance was found to be larger than 9 σ.

Additionally, it is worth presenting a simple extrapolation of the results presented in this thesis to the near future. The reason is that the analysis is approaching the significance threshold considered as evidence (3 σ), and that a considerable improvement of $\sim 25\%$ has been made over the iteration of the analysis on which the more detailed projection was based. Assuming that the discovery significance as well as the uncertainty on the signal strength scale with luminosity as $1/\sqrt{\mathcal{L}}$, which is reasonable for the near future, the 3 σ benchmark will be reached with

560 fb^{-1}, and the 5 σ benchmark with 1550 fb^{-1}, while the 0.5 uncertainty on the signal strength will be reached with 275 fb^{-1}, and 0.3 with 760 fb^{-1}.

At the time of writing this thesis the latest result from the CMS experiment is using 35.9 fb^{-1} of data collected in 2016 at the centre-of-mass energy of 13 TeV, and in combination with the earlier data collected at 7 and 8 TeV results in the observed (expected) significance of 0.9 (1.0) σ. In comparison to ATLAS, the CMS experiment benefits from a larger magnetic field, translating directly to better muon resolution. As a result, its performance is expected to be superior to that of ATLAS given the equivalent dataset, meaning that CMS is therefore expected to play an important role in the future studies of the Higgs boson decay to a pair of muons.

Summary

The Higgs boson is a key element of the Standard Model, our current best understanding of Nature, and a rich experimental programme is currently underway to measure its properties and interactions with other fundamental particles. Any discrepancy between the experimental data and predictions from the Standard Model could point the way towards new physics, potentially addressing the acute theoretical and experimental problems of the Standard Model.

The experimental programme relies on the ability to accelerate protons to very high energies using the Large Hadron Collider at CERN, and collide them at the centre of general-purpose detectors, ATLAS, and CMS. The work presented in this thesis relies on the data collected by the ATLAS detector and focuses on the search for the decay of the Higgs boson to muon pairs.

Muons are reconstructed in ATLAS by using the information from the inner tracker, the muon spectrometer, and the calorimeter system. Muon momentum is measured to the precision of a few percent at the scales relevant for the search and is calibrated in the MC simulation by using the invariant mass peak of the Z boson as a standard candle. Muon isolation, important for distinguishing between real prompt muons and fake muons or muons from secondary vertices, is described using information from both the tracker and the calorimeters. The isolation efficiency in MC simulation is corrected by the scale factors, measured using the tag-and-probe method. A novel method of background subtraction reduces the systematic uncertainties associated with the scale factor measurments by up to 75% in the low transverse momentum regime. A machine learning approach to estimating and eventually improving muon resolution with the goal of improving the analysis sensitivity is presented but the observed improvements are too small to justify the additional complexity of the analysis.

The search for the $H \to \mu\mu$ decay proceeds by selecting events which pass the loosest available unprescaled single muon triggers and have two oppositely charged muons. To reduce the backgrounds from $t\bar{t}$ and diboson processes, events with a b-tagged jet are rejected. The selected events are then sorted based on the number of jets in the event and further categorised using a gradient boosting classifier. The classifiers are trained with the MC simulation and the data sidebands using a four-fold validation approach to avoid overtraining. The boundaries defining the categories are determined using an exhaustive grid search, maximising the significance proxy combined across the categories. Signal models are fit separately in all categories to determine the shape of the signal. The background model is selected using a dedicated high-statistics MC simulation of Drell-Yan events, which is also used to determine the spurious signal systematic uncertainty. Finally, a combined maximum-likelihood fit is performed simultaneously in all categories to extract the signal strength and the associated confidence interval, as well as the discovery significance of the decay. The data is found to be compatible with both the Standard Model signal and no signal at all, with the result being limited by the statistical uncertainty on the data.

Finally, sensitivity projection studies are presented that estimate the analysis sensitivity achievable with future datasets. The coupling between the Higgs boson and the muons is expected to be measured to $\sim$13 %. The analysis is approaching the sensitivity to obtain evidence for the decay of the Higgs boson to muon pairs, and exciting results are expected from both ATLAS and CMS experiments in the immediate future.

Bibliography

[1] Miha Zgubič, Johannes Josef Junggeburth, Nicolas Maximilian Köhler, Stefano Rosati, and Federico Sforza. Isolation selection efficiency on 2016 data: Preliminary Public Plots. Technical Report ATL-COM-PHYS-2017-1685, CERN, Geneva, Nov 2017.

[2] Massimiliano Bellomo, Stefano Rosati, Gaetano Barone, Weitao Wang, Nicolas Maximilian Köhler, Johannes Josef Junggeburth, Michael Holzbock, Laura Fabbri, Maximiliano Sioli, Miha Zgubic, Stefano Zambito, and Federico Sforza. Muon performance results with 2017 and 2016 datasets. Technical Report ATL-COM-PHYS-2017-1360, CERN, Geneva, Sep 2017.

[3] Nicolas Maximilian Köhler, Johannes Josef Junggeburth, Miha Zgubič, Stefano Rosati, and Federico Sforza. High-μ Muon Reconstruction and Isolation Studies: Preliminary Public Plots. Technical Report ATL-COM-PHYS-2017-1684, CERN, Geneva, Nov 2017.

[4] Miha Zgubič, Matthew Henry Klein, Johannes Josef Junggeburth, Stefano Rosati, and Federico Sforza. Optimization of Isolation Selection at High-μ: Preliminary Public Plots. Technical Report ATL-COM-PHYS-2018-733, CERN, Geneva, May 2018.

[5] A search for the rare decay of the Standard Model Higgs boson to dimuons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2018-026, CERN, Geneva, Jul 2018.

[6] A search for the dimuon decay of the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2019-028, CERN, Geneva, Jul 2019.

[7] Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC. Technical Report ATL-PHYS-PUB-2018-054, CERN, Geneva, Dec 2018.

[8] M. Cepeda et al. Report from Working Group 2. *CERN Yellow Rep. Monogr.*, 7:221–584, 2019.

[9] Philipp Windischhofer, Miha Zgubič, and Daniela Bortoletto. Preserving physically important variables in optimal event selections: A case study in Higgs physics. 2019.

[10] MLHEP 2018. Machine learning in high energy physics summer school, 2018. [`http://www.physics.ox.ac.uk/confs/MLHEPOxford2018/index.asp`; accessed 27-January-2020].

[11] MLHEP 2019. Machine learning in high energy physics summer school, 2019. [`https://indico.cern.ch/event/768915/`; accessed 27-January-2020].

[12] Braam, Peter and Zgubič, Miha and Rousseau, David and Louppe, Gilles and Lukic, Vesna and Gukov, Sergei and Battaglia, Peter and Levasseur, Laurence and He, Yang-Hui. Machine learning + physics seminar series, 2019. [`http://www-pnp.physics.ox.ac.uk/~zgubic/SeminarSeries/SummaryPage.pdf`; accessed 27-January-2020].

[13] Mark Thomson. *Modern particle physics*. Cambridge University Press, New York, 2013.

[14] Georges Aad et al. Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. 2019.

[15] Gian F. Giudice and Oleg Lebedev. Higgs-dependent Yukawa couplings. *Phys. Lett.*, B665:79–85, 2008.

[16] Natascia Vignaroli. Searching for a dilaton decaying to muon pairs at the lhc. *Phys. Rev. D*, 80:095023, Nov 2009.

[17] Georges 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.*, B716:1–29, 2012.

[18] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[19] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[20] Georges Aad et al. Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector. *Eur. Phys. J.*, C75(10):476, 2015. [Erratum: Eur. Phys. J.C76,no.3,152(2016)].

[21] V. Khachatryan et al. Constraints on the spin-parity and anomalous *hvv* couplings of the higgs boson in proton collisions at 7 and 8 tev. *Phys. Rev. D*, 92:012004, Jul 2015.

[22] Georges 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*, 08:045, 2016.

[23] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.

[24] F. Zwicky. Republication of: The redshift of extragalactic nebulae. *General Relativity and Gravitation*, 41(1):207–224, 2008.

[25] L.S. Sparke and J.S. Gallagher. *Galaxies in the Universe: an Introduction*. Cambridge University Press, 2nd edition, 2007. Fig 5.20.

[26] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. 2015.

[27] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006.

[28] Robert A. Knop et al. New constraints on Omega(M), Omega(lambda), and w from an independent set of eleven high-redshift supernovae observed with HST. *Astrophys. J.*, 598:102, 2003.

[29] U. Sarkar. *Particle and Astroparticle Physics*. Series in High Energy Physics, Cosmology and Gravitation. Taylor & Francis, 2007.

[30] K. Abe et al. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339–344, 2020.

[31] Gian Francesco Giudice. Naturally Speaking: The Naturalness Criterion and Physics at the LHC. pages 155–178, 1 2008.

[32] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.

[33] Oliver Sim Brning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[34] ATLAS Collaboration. ATLAS Luminosity Public Results Run 2. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`. Accessed: 2019-11-25.

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

[36] Morad Aaboud et al. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J.*, C77(5):317, 2017.

[37] F. Ragusa and L. Rolandi. Tracking at LHC. *New J. Phys.*, 9:336, 2007.

[38] *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[39] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010.

[40] Karolos 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-HEP2015:261, 2015.

[41] *ATLAS liquid-argon calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.

[42] A Artamonov, D Bailey, G Belanger, M Cadabeschi, T-Y Chen, V Epshteyn, P Gorbounov, KK Joo, M Khakzad, V Khovanskiy, et al. The atlas forward calorimeter. *Journal of Instrumentation*, 3(02):P02010, 2008.

[43] G. Aad et al. Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays. *Eur. Phys. J.*, C70:875–916, 2010.

[44] S Aefsky. Alignment of the Muon Spectrometer in ATLAS. Technical Report ATL-MUON-PROC-2011-003, CERN, Geneva, Sep 2011.

[45] *ATLAS muon spectrometer: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[46] Theodoros Argyropoulos et al. Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance. *IEEE Trans. Nucl. Sci.*, 56:1568–1574, 2009.

[47] Giordano Cattani and. The resistive plate chambers of the ATLAS experiment: performance studies. *Journal of Physics: Conference Series*, 280:012001, feb 2011.

[48] K. Nagai. Thin gap chambers in ATLAS. *Nucl. Instrum. Meth.*, A384:219–221, 1996.

[49] Paolo Nason. A new method for combining nlo qcd with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040040, Nov 2004.

[50] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *Journal of High Energy Physics*, 2007(11):070070, Nov 2007.

[51] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *Journal of High Energy Physics*, 2010(6), Jun 2010.

[52] P. Freeman, S. Doe, and A. Siemiginowska. Sherpa: a mission-independent data analysis application. In J.-L. Starck and F. D. Murtagh, editors, *Astronomical Data Analysis*, volume 4477 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 76–87, November 2001.

[53] S. Doe, D. Nguyen, C. Stawarz, B. Refsdal, A. Siemiginowska, D. Burke, I. Evans, J. Evans, J. McDowell, J. Houck, and M. Nowak. Developing Sherpa with Python. In R. A. Shaw, F. Hill, and D. J. Bell, editors, *Astronomical Data Analysis Software and Systems XVI*, volume 376 of *Astronomical Society of the Pacific Conference Series*, page 543, October 2007.

[54] Torbjrn Sjstrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1. *Computer Physics Communications*, 178(11):852867, Jun 2008.

[55] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010.

[56] G. Duckeck, D. Barberis, R. Hawkings, R. Jones, N. McCubbin, G. Poulard, D. Quarrie, T. Wenaus, and E. Obreshkov. ATLAS computing: Technical design report. 2005.

[57] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[58] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, Feb 2006.

[59] J. Allison et al. Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186 – 225, 2016.

[60] A Buckley, T Eifert, M Elsing, D Gillberg, K Koeneke, A Krasznahorkay, E Moyse, M Nowak, S Snyder, and P van Gemmeren. Implementation of the ATLAS run 2 event data model. *Journal of Physics: Conference Series*, 664(7):072045, dec 2015.

[61] Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-072, CERN, Geneva, Jul 2010.

[62] T Cornelissen, M Elsing, S Fleischmann, W Liebig, E Moyse, and A Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar 2007.

[63] Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-026, CERN, Geneva, Jul 2015.

[64] Morad Aaboud et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1 protonproton collisions at the LHC. *Eur. Phys. J.*, C77(5):332, 2017.

[65] J. Illingworth and J. Kittler. A survey of the hough transform. *Computer Vision, Graphics, and Image Processing*, 44(1):87 – 116, 1988.

[66] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J.*, C76(5):292, 2016.

[67] Johannes Josef Junggeburth. Muon identification and reconstruction efficiencies in full Run-2 dataset. Technical Report ATL-COM-PHYS-2019-1001, CERN, Geneva, Aug 2019.

[68] Gaetano Barone, Johannes Josef Junggeburth, Federico Sforza, Stefano Rosati, and Nicolas Maximilian Kohler. Muon performance results with 2018 dataset: efficiencies, scales, and resolutions. Technical Report ATL-COM-PHYS-2018-661, CERN, Geneva, May 2018.

[69] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.*, C73(3):2304, 2013.

[70] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1):119 – 126, 2008.

[71] Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the atlas detector. *Phys. Rev. D*, 97:052010, Mar 2018.

[72] Francesco Giuli. Study of the Drell-Yan process with the ATLAS detector at the LHC, Oct 2018. Presented 09 Nov 2018.

[73] Jerome H. Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367 – 378, 2002. Nonlinear Methods and Data Mining.

[74] Greg Ridgeway. Generalized boosted models: A guide to the gbm package. *Compute*, 1:1–12, 01 2005.

[75] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. *arXiv e-prints*, page arXiv:1603.02754, March 2016.

[76] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3146–3154. Curran Associates, Inc., 2017.

[77] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[78] Roger Koenker and Kevin Hallock. Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156, 2001.

[79] Frank J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[80] Morad Aaboud et al. Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector. *Phys. Lett.*, B784:345–366, 2018.

[81] A measurement of the Higgs boson mass in the diphoton decay channel. Technical Report CMS-PAS-HIG-19-004, CERN, Geneva, 2019.

[82] G. Aad, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O.S. AbouZeid, H. Abramowicz, and et al. Observation and measurement of higgs boson decays toww*with the atlas detector. *Physical Review D*, 92(1), Jul 2015.

[83] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Er, C. Fabjan, and et al. Measurement of higgs boson production and properties in the ww decay channel with leptonic final states. *Journal of High Energy Physics*, 2014(1), Jan 2014.

[84] Morad Aaboud et al. Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 10:132, 2017.

[85] Albert M Sirunyan et al. Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV. *Phys. Lett.*, B792:369–396, 2019.

[86] M. Aaboud et al. Combined measurement of differential and total cross sections in the $H \to \gamma\gamma$ and the $H \to ZZ^* \to 4\ell$ decay channels at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett.*, B786:114–133, 2018.

[87] Albert M Sirunyan et al. Observation of $t\bar{t}H$ production. *Phys. Rev. Lett.*, 120(23):231801, 2018.

[88] M. Aaboud et al. Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. *Phys. Lett.*, B784:173–191, 2018.

[89] A. M. Sirunyan et al. Observation of Higgs boson decay to bottom quarks. *Phys. Rev. Lett.*, 121(12):121801, 2018.

[90] Morad Aaboud et al. Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector. *Phys. Lett.*, B786:59–86, 2018.

[91] M. Aaboud et al. Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment. *Phys. Rev. Lett.*, 120(21):211802, 2018.

[92] Search for the standard model Higgs boson decaying to charm quarks. Technical Report CMS-PAS-HIG-18-031, CERN, Geneva, 2019.

[93] Avital Dery, Aielet Efrati, Yonit Hochberg, and Yosef Nir. What if BR($h - - > \mu\mu$)/BR($h - - > \tau\tau$) does not equal m_μ^2/m_τ^2? *JHEP*, 05:039, 2013.

[94] Georges Aad et al. Search for the Standard Model Higgs boson decay to $\mu^+\mu^-$ with the ATLAS detector. *Phys. Lett.*, B738:68–86, 2014.

[95] Vardan Khachatryan et al. Search for a standard model-like Higgs boson in the $\mu^+\mu^-$ and e^+e^- decay channels at the LHC. *Phys. Lett.*, B744:184–207, 2015.

[96] Morad Aaboud et al. Search for the dimuon decay of the Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 119(5):051802, 2017.

[97] Search for the standard model Higgs boson decaying into two muons in pp collisions at sqrt(s)=13TeV. Technical Report CMS-PAS-HIG-17-019, CERN, Geneva, 2017.

[98] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[99] Tanju Gleisberg and Stefan Hoeche. Comix, a new matrix element generator. *JHEP*, 12:039, 2008.

[100] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012.

[101] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations. *Comput. Phys. Commun.*, 212:220–238, 2017.

[102] Steffen Schumann and Frank Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03:038, 2008.

[103] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. QCD matrix elements + parton showers. *JHEP*, 11:063, 2001.

[104] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.

[105] Stefan Hoeche, Frank Krauss, Marek Schonherr, and Frank Siegert. A critical appraisal of NLO+PS matching methods. *JHEP*, 09:049, 2012.

[106] ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014.

[107] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys. Rev.*, D83:091503, 2011.

[108] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys. Rev.*, D82:054018, 2010.

[109] Keith Hamilton, Paolo Nason, Emanuele Re, and Giulia Zanderighi. NNLOPS simulation of Higgs boson production. *JHEP*, 10:222, 2013.

[110] Charalampos Anastasiou, Claude Duhr, Falko Dulat, Elisabetta Furlan, Thomas Gehrmann, Franz Herzog, Achilleas Lazopoulos, and Bernhard Mistlberger. High precision determination of the gluon fusion Higgs boson cross-section at the LHC. *JHEP*, 05:058, 2016.

[111] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini. Two loop light fermion contribution to Higgs production and decays. *Phys. Lett.*, B595:432–441, 2004.

[112] Stefano Actis, Giampiero Passarino, Christian Sturm, and Sandro Uccirati. NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders. *Phys. Lett.*, B670:12–17, 2008.

[113] Georges Aad et al. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 09:145, 2014.

[114] M. Ciccolini, Ansgar Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99:161803, 2007.

[115] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Phys. Rev.*, D77:013002, 2008.

[116] K. Arnold et al. VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons. *Comput. Phys. Commun.*, 180:1661–1670, 2009.

[117] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro. Higgs production via vector-boson fusion at NNLO in QCD. *Phys. Rev. Lett.*, 105:011801, 2010.

[118] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579:149–156, 2004.

[119] M. L. Ciccolini, S. Dittmaier, and M. Kramer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev.*, D68:073003, 2003.

[120] Oliver Brein, Robert Harlander, Marius Wiesemann, and Tom Zirke. Top-Quark Mediated Effects in Hadronic Higgs-Strahlung. *Eur. Phys. J.*, C72:1868, 2012.

[121] Lukas Altenkamp, Stefan Dittmaier, Robert V. Harlander, Heidi Rzehak, and Tom J. E. Zirke. Gluon-induced Higgs-strahlung at next-to-leading order QCD. *JHEP*, 02:078, 2013.

[122] Ansgar Denner, Stefan Dittmaier, Stefan Kallweit, and Alexander Mck. HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders. *Comput. Phys. Commun.*, 195:161–171, 2015.

[123] Oliver Brein, Robert V. Harlander, and Tom J. E. Zirke. vh@nnlo - Higgs Strahlung at hadron colliders. *Comput. Phys. Commun.*, 184:998–1003, 2013.

[124] Robert V. Harlander, Anna Kulesza, Vincent Theeuwes, and Tom Zirke. Soft gluon resummation for gluon-induced Higgs Strahlung. *JHEP*, 11:082, 2014.

[125] Robert V. Harlander, Jonas Klappert, Stefan Liebler, and Lukas Simon. vh@nnlo-v2: New physics in Higgs Strahlung. *JHEP*, 05:089, 2018.

[126] 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*, 07:079, 2014.

[127] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 03:015, 2013.

[128] Yu Zhang, Wen-Gan Ma, Ren-You Zhang, Chong Chen, and Lei Guo. QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider. *Phys. Lett.*, B738:1–5, 2014.

[129] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, and P. M. Zerwas. NLO QCD corrections to t anti-t H production in hadron collisions. *Nucl. Phys.*, B653:151–203, 2003.

[130] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013.

[131] A study of the Pythia 8 description of ATLAS minimum bias measurements with the Donnachie-Landshoff diffractive model. Technical Report ATL-PHYS-PUB-2016-017, CERN, Geneva, Aug 2016.

[132] P. Golonka and Z. Was. Next to Leading Logarithms and the PHOTOS Monte Carlo. *Eur. Phys. J.*, C50:53–62, 2007.

[133] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[134] Morad Aaboud et al. Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016. *Eur. Phys. J.*, C79(3):205, 2019.

[135] Georges Aad et al. Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev.*, D91(1):012006, 2015.

[136] A. Altheimer et al. Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks. *J. Phys.*, G39:063001, 2012.

[137] Andrew J. Larkoski, Ian Moult, and Benjamin Nachman. Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning. 2017.

[138] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J.*, C77:490, 2017.

[139] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[140] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[141] Tagging and suppression of pileup jets with the ATLAS detector. Technical Report ATLAS-CONF-2014-018, CERN, Geneva, May 2014.

[142] Morad Aaboud et al. Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector. *Eur. Phys. J.*, C77(9):580, 2017. [Erratum: Eur. Phys. J.C77,no.10,712(2017)].

[143] Morad Aaboud et al. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. *JHEP*, 08:089, 2018.

[144] Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. Technical Report ATL-PHYS-PUB-2017-013, CERN, Geneva, Jul 2017.

[145] Georges Aad et al. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. *Eur. Phys. J.*, C74(7):2941, 2014.

[146] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[147] Morad Aaboud et al. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C78(11):903, 2018.

[148] Georges Aad et al. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C75:120, 2015.

[149] Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM.

[150] John C. Collins and Davison E. Soper. Angular distribution of dileptons in high-energy hadron collisions. *Phys. Rev. D*, 16:2219–2225, Oct 1977.

[151] Albert M. Sirunyan et al. Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV. *Eur. Phys. J.*, C78(9):701, 2018.

[152] Tom Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[153] Fernando Nogueira. A python implementation of global optimization with gaussian processes., 2019.

[154] M. Oreglia. A Study of the Reactions $\psi' \to \gamma\gamma\psi$. *SLAC-R-0236*, 1980.

[155] Morad Aaboud et al. Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector. *JHEP*, 09:001, 2016.

[156] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J.*, C58:639–707, 2008.

[157] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016.

[158] Luminosity determination in *pp* collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2019-021, CERN, Geneva, Jun 2019.

[159] Morad Aaboud et al. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.*, C76(12):653, 2016.

[160] M. Aaboud et al. Measurement of the Drell-Yan triple-differential cross section in *pp* collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 12:059, 2017.

[161] Valerio Bertone, Stefano Carrazza, and Juan Rojo. APFEL: A PDF Evolution Library with QED corrections. *Comput. Phys. Commun.*, 185:1647–1668, 2014.

[162] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordstrm, Ben Page, Martin Rfenacht, Marek Schnherr, and Graeme Watt. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J.*, C75:132, 2015.

[163] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[164] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[165] A L Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[166] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.

[167] ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical Report CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva, Sep 2017.

[168] Rabah Abdul Khalek, Shaun Bailey, Jun Gao, Lucian Harland-Lang, and Juan Rojo. Towards Ultimate Parton Distributions at the High-Luminosity LHC. *Eur. Phys. J.*, C78(11):962, 2018.

[169] Prospects for the measurement of the rare Higgs boson decay $H \to \mu\mu$ with 3000 fb^{-1} of *pp* collisions collected at $\sqrt{s} = 14$ TeV by the ATLAS experiment. Technical Report ATL-PHYS-PUB-2018-006, CERN, Geneva, May 2018.