

Search for the Higgs boson in $H \rightarrow \mu^+ \mu^-$ decays with
the ATLAS detector at the LHC

Suche nach dem Higgs Boson in $H \rightarrow \mu^+ \mu^-$ Zerfällen
mit dem ATLAS Detektor am LHC



Masterarbeit an der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von

Friedrich G. Hönig

geboren am 06. Mai 1988 in Weißenburg i. Bay.

27. März 2013

Gutachterin: Prof. Dr. Dorothee Schaile

Zusammenfassung

Das Standardmodell (SM) war in den letzten Jahrzehnten die erfolgreichste Theorie, praktisch alle Phänomene der Teilchenphysik zu beschreiben. Alle Materieteilchen und Austauschteilchen, die vom SM vorhergesagt werden, sind gefunden und experimentell untersucht worden. Der letzte fehlende Baustein war das Higgs-Boson, ein Resultat des Mechanismus, der für die Masse aller massiven Teilchen verantwortlich ist. Der nächste experimentelle Schritt ist nun, die Eigenschaften dieses Bosons zu vermessen. Eine dieser Eigenschaften ist die Kopplung an Fermionen wie das Myon. Diese Analyse untersucht den direkten Zerfall von Higgs-Bosonen in Myon-Antimyon-Paare. Eine Suche nach diesem Zerfall wurde anhand von 20.7 fb^{-1} Proton-Proton-Kollisionsdaten durchgeführt, die von dem ATLAS Detektor am LHC im Jahr 2012 bei einer Schwerpunktsenergie von $\sqrt{s} = 8 \text{ TeV}$ aufgenommen wurden. Das invariante Massenspektrum von Di-Myon-Ereignissen in diesen Daten wurde untersucht mittels eines kombinierten Fits einer Hintergrund- und einer Signal-Parametrisierung. Kein signifikanter Überschuss konnte festgestellt werden. Ausschlussgrenzen an die Signalstärke wurden berechnet.

Abstract

For the last decades, the Standard Model (SM) has been most successful in describing virtually all phenomena of particle physics. All matter particles and all force carriers predicted by the SM have been found and studied experimentally. The last missing piece was the Higgs boson, a result of the mechanism that generates the mass of all massive particles. A candidate boson has now been found by the ATLAS and CMS collaborations. The next experimental step is the measurement of its properties. This includes the coupling to fermions such as muons. The direct decay of Higgs bosons into muon-antimuon pairs is studied in this analysis. A search for this decay is performed using 20.7 fb^{-1} of $\sqrt{s} = 8 \text{ TeV}$ proton-proton collision data recorded by the ATLAS detector at the LHC in 2012. The invariant mass spectrum of dimuon events is examined using a combined fit of a background and a signal parametrization to this data. No significant excess has been found. Exclusion limits on the strength of this signal have been calculated.

“Everything is a mathematical
trick, except what you measure
in the lab.”

— Prof. Dr. Armin Scrinzi in a
Quantum Mechanics lecture

Contents

1	Introduction	1
2	Theoretical Background	2
2.1	The Standard Model	2
2.2	Quantum Chromo Dynamics	4
2.3	Electroweak Theory and BEH Mechanism	4
3	The Large Hadron Collider	6
3.1	Physics at the LHC	6
3.1.1	Hadronization and Jets	8
3.2	Higgs Production	10
3.3	Higgs Decays	11
4	Background Processes	15
4.1	Z + Jets	15
4.2	Leptonic Top Decays	15
4.3	Diboson Production	15
4.4	W+Jets	16
4.5	The ATLAS Experiment	17
4.5.1	Coordinate System	17
4.5.2	Magnet Systems	18
4.5.3	Inner Detector	18
4.5.4	Electromagnetic Calorimeter	19
4.5.5	Hadronic Calorimeter	20
4.5.6	Muon Spectrometer	20
4.6	Trigger and Data Acquisition	22
4.7	Grid computing	23
5	Data and Simulated Samples	24
5.1	Data	24
5.2	Monte Carlo Simulation	24
5.2.1	Signal Samples	25
5.2.2	Background Samples	25
5.3	Corrections to Monte Carlo	27
5.3.1	Muon Momentum Corrections	27
5.3.2	Scale Factors	27
5.4	Jet Reconstruction	28
5.5	Muon Reconstruction	29
5.6	Missing Transverse Energy	30

5.7	Blinding	31
6	Event Selection	32
6.1	Muon Trigger	32
6.2	Selection Cuts	32
6.3	Higgs $\rightarrow \tau^+\tau^-$ Contamination	36
6.4	W+Jets Estimation	38
7	Control Plots	39
8	Signal and Background Description	42
8.1	Signal Modeling	42
8.2	Background Modeling	45
9	Systematic Uncertainties	48
9.1	Luminosity Uncertainty	48
9.2	Pile-Up Reweighting	48
9.3	Muon-related Uncertainties	48
9.4	Production Cross-Section Uncertainty	51
9.5	Uncertainty on the Branching Ratio	51
9.6	Signal Acceptance Uncertainty	53
9.6.1	Renormalization and Factorization Scales	53
9.6.2	PDF Uncertainties	55
9.7	Initial and Final State Radiation	56
9.8	Summary of Theoretical Acceptance Uncertainties	58
10	Statistics	59
10.1	Hypothesis Testing	59
10.2	Exclusion Limit	61
10.3	Likelihood	61
11	Results	63
12	Conclusions and Outlook	66

1 Introduction

The Standard Model of particle physics is the theory that describes best almost all currently known phenomena in particle physics. It offers a description of matter in the form of fermions and all forces, except gravity, in the form of bosons, that mediate the electromagnetic, weak and strong interactions. All predicted particles of the Standard Model have been found, including the tau and the top quark. The one last missing piece was the Higgs boson. The Higgs field, of which the Higgs boson is its manifestation as a particle in its own right, has been introduced to explain the fact that the $W^\pm$ and Z bosons, the carriers of the weak interaction, have mass.

Searches at the Large Electron Positron Collider LEP, which combine measurements from the ALEPH, DELPHI, L3 and OPAL experiments, could exclude Higgs boson masses of $m_H > 114.4$ GeV at the 95% confidence level [1]. After LEP was shut down and replaced by the Large Hadron Collider LHC, searches for the Higgs boson began at the ATLAS and CMS experiments, where high Higgs boson masses could be excluded.

A Higgs-like boson with a mass of about 125 GeV has now been found at both the ATLAS [2] and the CMS [3] experiments at the LHC. So far, decays to pairs of photons, $W^\pm$, and Z bosons have been detected. The next step is to examine this new boson and to check, if it behaves as the Standard Model (SM) of particle physics predicts. Early property measurements of the new boson favor a SM Higgs boson, for example by excluding spin 2^+ in favor of the SM prediction of spin 0^+ [4].

Other important properties are the couplings to other elementary particles. One potentially accessible decay of the Higgs boson is into muon pairs. While according to the SM the branching ratio of $H \rightarrow \mu^+\mu^-$ is 2.2×10^{-4} [5], some models like the Minimal Supersymmetric Standard Model (MSSM) require multiple Higgs bosons with enhanced branching ratios [6][7].

This thesis describes the search for $H \rightarrow \mu^+\mu^-$ decays, using data taken by the ATLAS experiment at the LHC. The invariant mass spectrum of di-muon events has been studied in search of a resonance caused by Higgs bosons decaying into pairs of muons.

2 Theoretical Background

The following chapter will give a short overview of the theoretical background of the search for $H \rightarrow \mu^+ \mu^-$ decays.

2.1 The Standard Model

The Standard Model of particle physics (SM) describes matter and its interactions by electromagnetic, weak and strong forces [8]. In the SM, matter consists of three generations of fermions, whereas the interactions are mediated by the exchange of bosons. Each generation of fermions contains an up- and a down-type quark, a charged lepton and a neutral lepton-neutrino. All these fermions are experimentally shown to be massive. The neutrinos, while being extremely light, must have a finite mass, as shown by the observation of neutrino oscillations [9]. This pattern is repeated three times with increasing particle masses in each generation, as shown in table 1. For every fermion, there exists an anti-fermion with opposite charge.

	fermion	charge	mass
1st generation	up quark u	$+\frac{2}{3}$	$2.3^{+0.7}_{-0.5}$ MeV
	down quark d	$-\frac{1}{3}$	$4.8^{+0.7}_{-0.3}$ MeV
	electron e	-1	0.5110 MeV
	electron neutrino ν_e	0	< 2 eV
2nd generation	charm quark c	$+\frac{2}{3}$	1.275 ± 0.025 GeV
	strange quark s	$-\frac{1}{3}$	95 ± 5 MeV
	muon μ	-1	105.7 MeV
	muon neutrino ν_μ	0	< 0.19 MeV
3rd generation	top quark t	$+\frac{2}{3}$	$173.5 \pm 0.6 \pm 0.8$ GeV
	bottom quark b	$-\frac{1}{3}$	4.18 ± 0.03 GeV
	tau τ	-1	1776.82 ± 0.16 MeV
	tau neutrino ν_τ	0	< 18.2 MeV

Table 1: Overview of fermions in the Standard Model. Masses are taken from [10]

The electromagnetic and weak the interaction are described together as the electroweak theory. In it the electromagnetic force is mediated by the massless photon γ . The weak interaction, responsible for nuclear decays, manifests as the massive Z^0 boson and the massive $W^\pm$ bosons. Massless gluons mediate the strong force which holds together atomic nuclei and the

nucleons themselves. An overview of these force carriers is given in table 2. Additionally, the SM predicts a Higgs boson, which is involved in the mechanism that accounts for the masses of the electroweak bosons.

	force	charge	mass [GeV]
gluon g	strong	0	0 [theory]
photon γ	electromagnetic	0	0 [theory]
$W^\pm$	electroweak	± 1	80.385 ± 0.015 [10]
Z	electroweak	0	91.1876 ± 0.0021 [10]

Table 2: Overview of gauge bosons in the Standard Model

All these forces are described by quantum field theories, specifically the electroweak (EW) theory and the quantum chromo dynamics (QCD) for strong interactions. Fields are described using their Lagrangian density $\mathcal{L}$, which is a function of the fields ϕ , the derivatives of the fields with respect to space-time coordinates $\frac{\partial \phi}{\partial x_\mu}$ and the coordinates x_μ themselves.

The dynamics of the fields are then determined by the Euler-Lagrange equation

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial (\partial \phi / \partial x^\mu)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (1)$$

For a spin 0 field for example, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^2, \quad (2)$$

which results after applying the Euler-Lagrange equation in the Klein-Gordon equation

$$(\partial_\mu \partial^\mu + m^2)\phi = 0. \quad (3)$$

The SM is a so called Gauge-theory. Its interactions between particles are constructed by requiring local gauge invariance. This means, the Lagrangian has to be invariant under the transformation

$$\psi \rightarrow e^{-i\alpha^i(x)f^i} \cdot \psi, \quad (4)$$

where $\alpha^i(x)$ is an arbitrary function of the space-time coordinate and f^i are generators of a symmetry group, which has been chosen to be enforced. For the SM, this group is $U(1)_Y \otimes SU(2)_L \otimes SU(3)_C$.

2.2 Quantum Chromo Dynamics

The $SU(3)_C$ symmetry yields QCD, the theory of strong interactions. The index C refers to the charge of strong interaction, the so called color. Every colored fermion, the so called quarks, carries one of the three colors red, blue or green. Gluons are the spin 1 exchange particles of the strong interaction. Each gluon is charged with a color and an anticolor. Quarks have never been observed as free particles; they rather form bound states, called hadrons. There are two known types of hadrons, distinguished by their number of quarks. Either a quark and an antiquark are bound together as a meson, or three quarks (antiquarks) are bound as a baryon (antibaryon). This is explained by the concept of confinement. A colored particle can never exist on its own, but must be bound in a colorless state. In mesons, this is accomplished by a quark with one color, and the other with the respective anticolor. In baryons, a quark of each color together gives a colorless state (“red + blue + green = white”).

2.3 Electroweak Theory and BEH Mechanism

The electroweak theory [11], developed by Glashow, Weinberg and Salam, is based on the $U(1)_Y \otimes SU(2)_L$ symmetry group. The index Y stands for the charge of the weak interaction, the so called weak hypercharge. It is a combination of the electrical charge Q and the third component of the weak isospin T_3 .

$$Y = 2(Q - T_3) \quad (5)$$

It yields four massless gauge fields, one called B_μ from the $U(1)_Y$ and three called W_μ^i . The Brout-Englert-Higgs [12][13][14] mechanism causes the so called electroweak symmetry breaking.

A mass term for the Z^0 and $W^\pm$, written into the Lagrangian of the Standard Model, would break the local gauge symmetry. Therefore, the mass of these bosons has to be generated by some other process. The BEH mechanism proposes a solution to this problem by introducing another complex scalar field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (6)$$

with hypercharge $Y = 1$. It is described by the Lagrangian

$$\mathcal{L} = |D_\mu \phi|^2 - V(\phi) \quad (7)$$

with the potential term

$$V(\phi) = \mu^2 |\phi^\dagger \phi| + \lambda (|\phi^\dagger \phi|)^2 \quad (8)$$

and the covariant derivative

$$D_\mu = \partial_\mu + i\frac{g}{2}\vec{\tau} \cdot \vec{W}_\mu + i\frac{g'}{2}B_\mu Y. \quad (9)$$

For $\mu^2 < 0$ this potential has its minimum not at $\phi = 0$, but at a finite value called the vacuum expectation value of Φ . From the set of possible minima one chooses

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (10)$$

with $v = \sqrt{-\frac{\mu^2}{\lambda}}$. Since quantum field theory is a perturbative theory, one expands the Lagrangian around this minimum. New terms appear that can be interpreted as masses for the $W^\pm$ and Z bosons.

The four fields are rewritten as their mass eigenstates,

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad (11)$$

$$Z_\mu = -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \quad (12)$$

$$W_\mu^+ = \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2) \quad (13)$$

$$W_\mu^- = \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2), \quad (14)$$

with the Z^0 and $W^\pm$ now having acquired mass, while the photon, corresponding to the A_μ field, stays massless.

Fermion masses are generated by a Yukawa interaction with the Higgs field.

3 The Large Hadron Collider

The Large Hadron Collider (LHC) is a synchrotron proton-proton accelerator and collider, located at the European Laboratory for Particle Physics CERN. With its 27 km circumference it is the largest of its type currently in existence.

It was built in the tunnel formerly used by the Large Electron-Positron Collider (LEP), with up to $\sqrt{s} = 209$ GeV center of mass energy the most powerful lepton collider to date. LEP was used from 1989 to 2000, after which it was dismantled to make room for the LHC. Although LEP was a very useful tool for precision measurements, to reach higher $\sqrt{s}$ where new physics, including the Higgs boson, was suspected, heavier particles like protons had to be accelerated. From 2012 to the beginning of 2013, the center of mass energy was 8 TeV with plans to upgrade to 14 TeV in the future.

3.1 Physics at the LHC

At the LHC, bunches of protons are accelerated and subsequently brought to collision. Before protons enter the LHC ring, they are accelerated by a number of pre-accelerators. Free protons are produced by ionizing hydrogen. They are first accelerated by a linear accelerator called LINAC2. They are then brought to increasingly high energies by first the Proton Synchrotron Booster, the Proton Synchrotron and finally the Super Proton Synchrotron, after being injected into the main LHC ring. An overview over the CERN accelerator complex is shown in figure 1.

The expected number of proton-proton events per unit time at a collider is $f = \sigma \cdot \mathcal{L}$. Here $\mathcal{L}$ is the so called instant luminosity. It is the number of proton-proton encounters per unit time and unit area at the collision point. In ATLAS, it is defined as [16]

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y}, \quad (15)$$

where n_b is the number of proton bunches and f_r the revolution frequency of these bunches in the LHC. n_1 and n_2 are the number of protons in each colliding bunch. Σ_x and Σ_y are the width of the beam as measured in van-der-Meer scans.

The time integrated luminosity $L = \int dt \cdot \mathcal{L}$ for which the detector was running is a measure of the total amount of recorded data. For the $\sqrt{s} = 8$ TeV run, about 20.7fb^{-1} of luminosity usable for this analysis was recorded by ATLAS. The luminosity delivered by the LHC is shown in figure 2.

[illegible]

European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

The cross section σ gives the probability of a certain process happening when the two proton beams cross at a certain energy. It is given in units of area, which can be interpreted as an effective area of the proton that can be hit. When different processes can occur, their individual cross sections add up to a total cross section. An overview of proton-proton cross sections is given in figure 3.

7

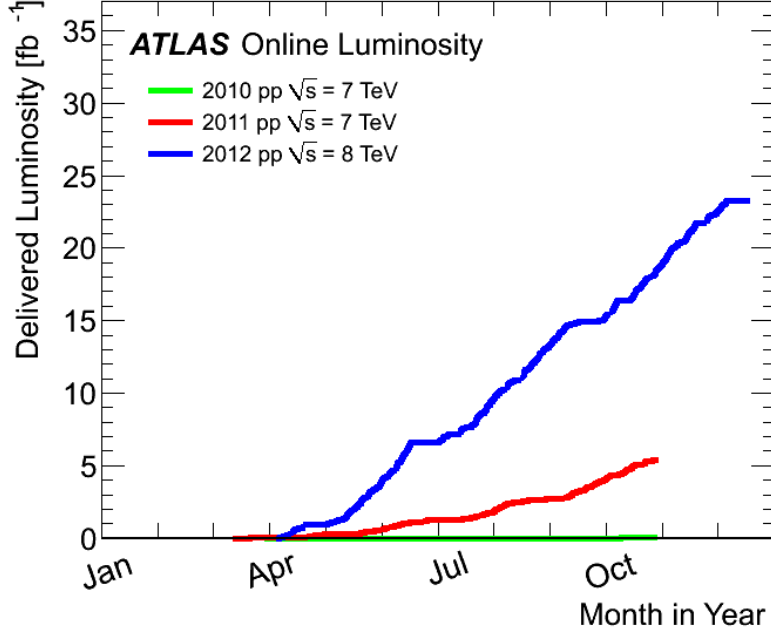


Figure 2: Integrated luminosity versus time delivered to ATLAS, shown for the years 2010 (green), 2011 (red) and 2012 (blue).

3.1.1 Hadronization and Jets

The principle of confinement states that quarks and gluons cannot exist as free particles outside of hadrons. Partons that emerge from parton-parton interaction as they do at inelastic proton-proton scattering will therefore form hadrons by creating quark antiquark pairs until all final particles are colorless. This is called hadronization. These hadrons and their decay products will emerge from the interaction point in roughly the same direction as the original parton. This collection of particles containing most of the energy and momentum of the original parton is called a jet.

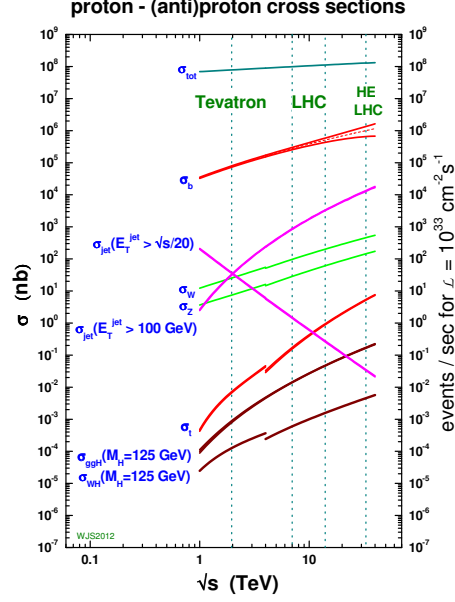


Figure 3: Standard Model process cross sections at hadron colliders as a function of center of mass energy.

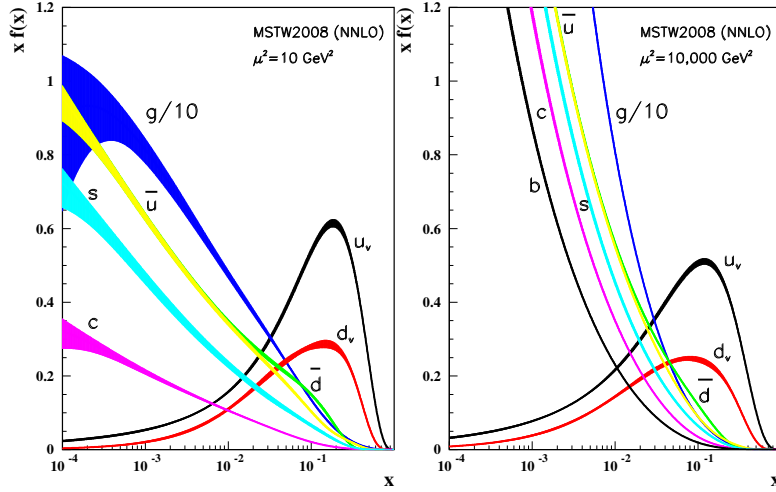


Figure 4: Distribution of x time the unpolarized parton distributions $f(x)$ at scales $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10000 \text{ GeV}^2$ (right). f can be one of the valence quarks u_v or d_v , one of the sea quarks $\bar{u}$, $\bar{d}$, s , c , b , or a gluon. The NNLO MSTW2008 parameterization is used.

3.2 Higgs Production

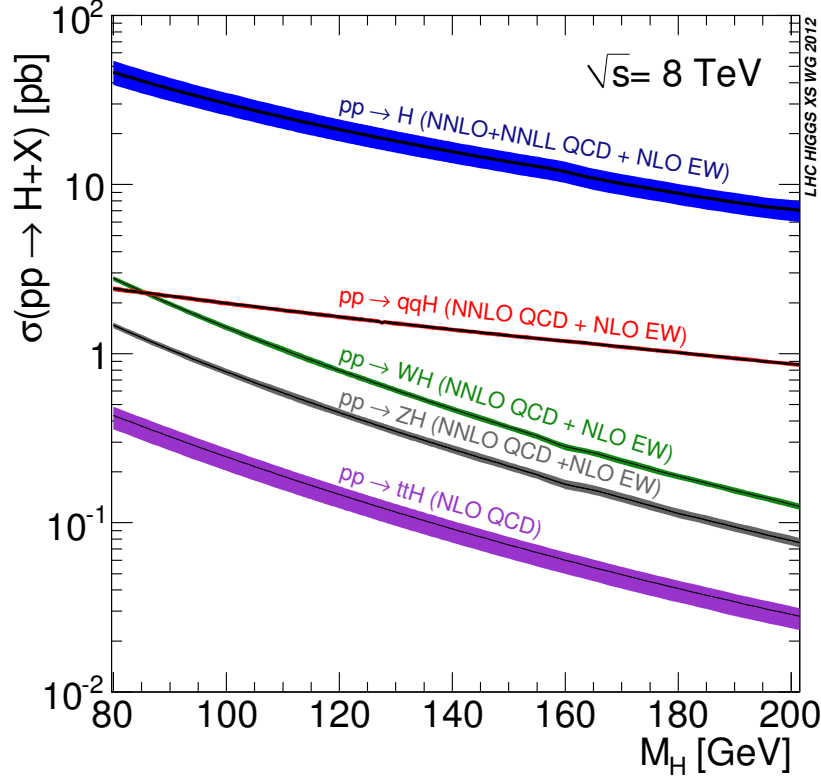


Figure 5: Higgs boson production cross sections at the LHC at a center of mass energy of 8 TeV as a function of Higgs mass M_H

The production cross section for Standard Model Higgs boson is shown in figure 5. There are a number of different production channels as depicted in figure 6. The dominant channel is the gluon gluon fusion process, in which two gluons create a Higgs boson via a fermion loop. Higgs bosons can also be created by vector boson fusion. In this process, two quarks from the initial state protons each radiate a $W^\pm$ or Z boson, which then fuse to create a Higgs boson. These quarks then create jets which are only slightly deflected off the beam axis, called forward jets. Another production is the so called Higgs-strahlung, where a Higgs boson is radiated off a Z or $W^\pm$ boson.

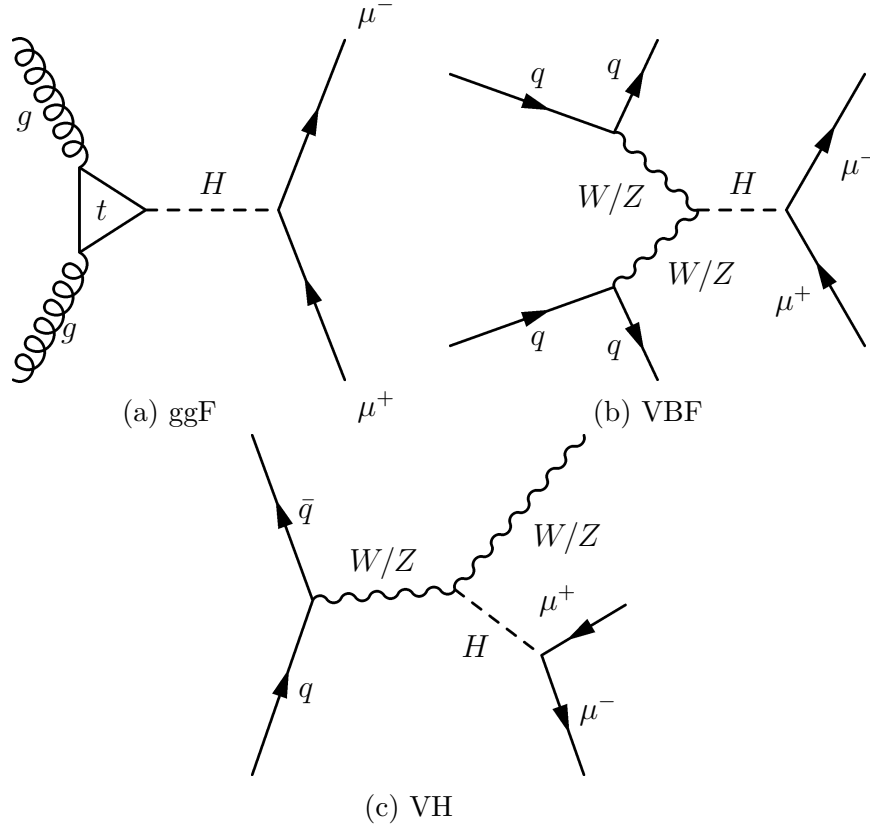


Figure 6: The dominant Higgs production modes at the LHC: gluon gluon fusion (ggF), vector boson fusion (VBF) and production in association with a Z or W boson (VH)

3.3 Higgs Decays

Once produced, the Higgs bosons decay into a variety of decay products with certain probabilities, called branching ratios. Figure 7 shows predictions of them as a function of assumed Higgs mass. For most of the major decay channels, there exist dedicated search groups in both ATLAS and CMS. As of the writing of this thesis, Higgs decays have been found in the $H \rightarrow WW$, $H \rightarrow ZZ$ and $H \rightarrow \gamma\gamma$ channels. Efforts are ongoing for $H \rightarrow bb$, $H \rightarrow \tau\tau$ and other decay channels.

The Higgs boson is also predicted to decay into muon-antimuon pairs. The SM prediction for the $H \rightarrow \mu^+\mu^-$ branching ratio in the studied range from 110 to 150 GeV varies between 2.8×10^{-4} and 6.5×10^{-5} [5]. For an assumed Higgs mass of $m_H \approx 125$ GeV, we expect about 10^2 events in the examined dataset, where the Higgs boson decays into $\mu^+\mu^-$ pairs. Such events are characterized by two oppositely charged muons with high transverse mo-

mentum. Due to the very short lifetime of the Higgs boson, both muons originate from the so called primary vertex. Since no neutrinos are involved, no missing transverse energy is expected. The signal process shows a narrow peak in the spectrum of the invariant mass of this di-muon system, centered around the Higgs mass. The width of this peak is dominated by detector resolution effects, due to the small natural width of the Higgs boson, as seen in figure 8. The di-muon system tends to have a large transverse momentum $p_T^{\mu\mu}$.

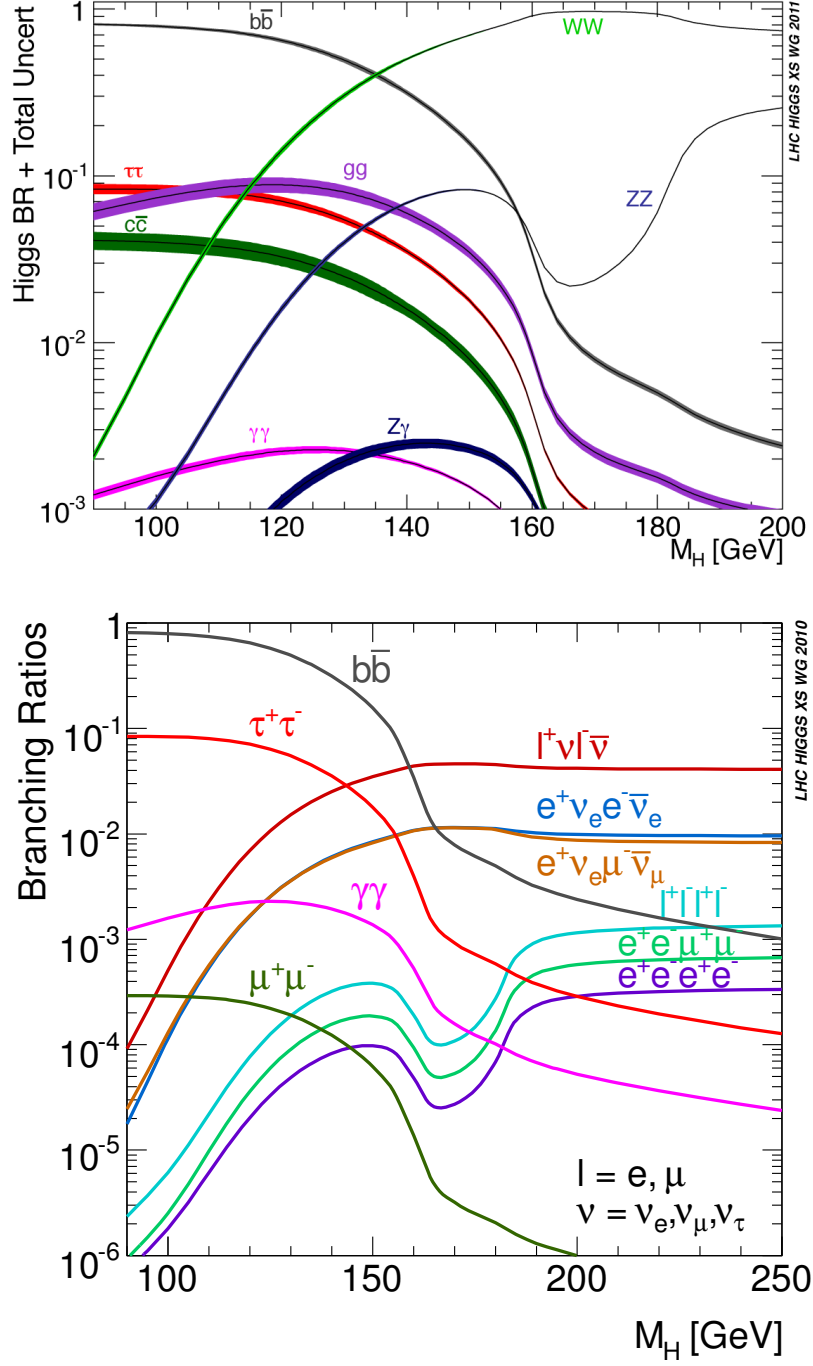


Figure 7: Branching ratios of Higgs decays as a function of Higgs mass. In the top plot, all major decay channels are shown. In the bottom plot, only final states with two or four leptons are included.

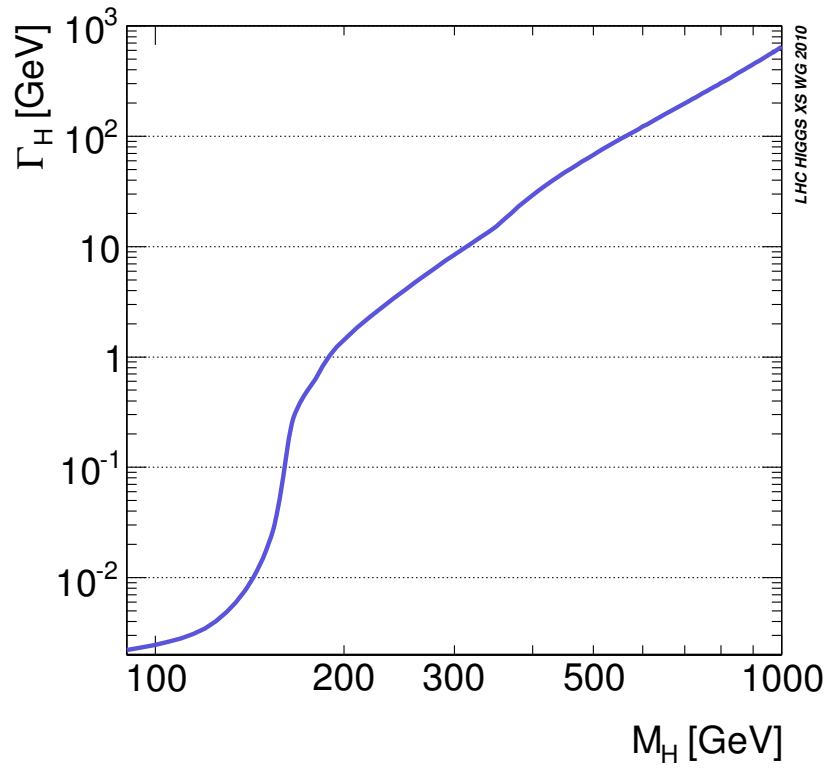


Figure 8: Total decay width of the Higgs boson, depending on Higgs mass M_H

4 Background Processes

4.1 Z + Jets

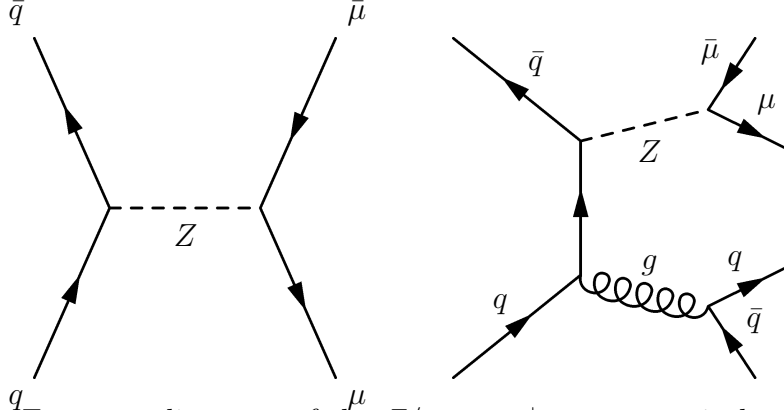


Figure 9: Feynman diagrams of the $Z/\gamma^* \rightarrow \mu^+\mu^-$ process in leading order (left) and in association with jets (right)

The most important background of this search is the $Z/\gamma^* \rightarrow \mu^+\mu^-$ Drell-Yan process, possibly with associated jets, as depicted in figure 9. Quarks in the initial state annihilate over a photon or Z boson, which then can decay into a $\mu^+\mu^-$ pair. The $m_{\mu\mu}$ spectrum is continuously falling, except for a resonance at the value of the Z mass, around 91.2 GeV. Around the Z -peak, the exchange is dominated by the Z boson. Z +jets events are characterized by two oppositely charged muons and no missing transverse energy. Events of this process therefore have a very similar signature to the signal signature. The $p_T^{\mu\mu}$ of Z +jets events tends to lower values as opposed to the Higgs events.

4.2 Leptonic Top Decays

The second largest background is the top production as depicted in figure 10. Top quarks decay almost exclusively into a W boson and a bottom quark. The W bosons then can decay into muons and neutrinos. Since the top quarks in $t\bar{t}$ -events are oppositely charged, the muons are as well. Energy carried away by the neutrinos causes missing transverse energy in the event. Top quarks can also be created without a second top, which has a significantly lower cross section, however.

4.3 Diboson Production

In events with two bosons created, two oppositely charged muons can occur in the final state. The largest diboson contribution comes from the W^+W^-

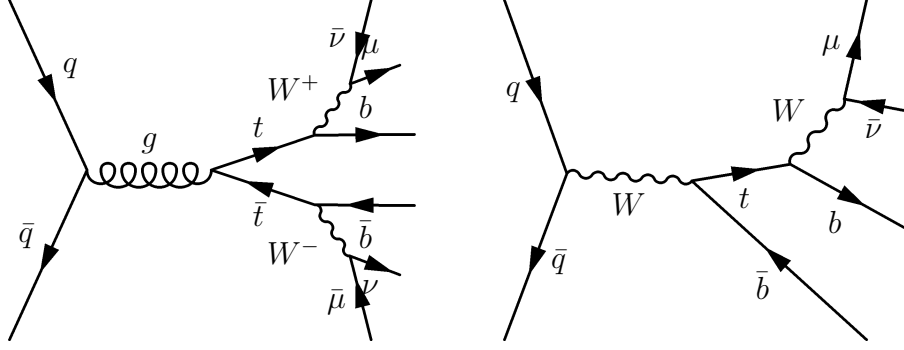


Figure 10: Diagrams of $t\bar{t}$ -pair production (left), single top production and subsequent decays into leptons

process. Other diboson background processes included are $W^\pm Z$, ZZ , $Z\gamma$ and $W^\pm\gamma$.

4.4 W+Jets

Events in which a $W^\pm$ boson is created in association with jets, which is a process with a high cross section, pose another background. Jets may be misinterpreted as leptons, so that a muon from the decay of the $W^\pm$ together with the fake muon give the appearance of a $\mu^+\mu^-$ pair.

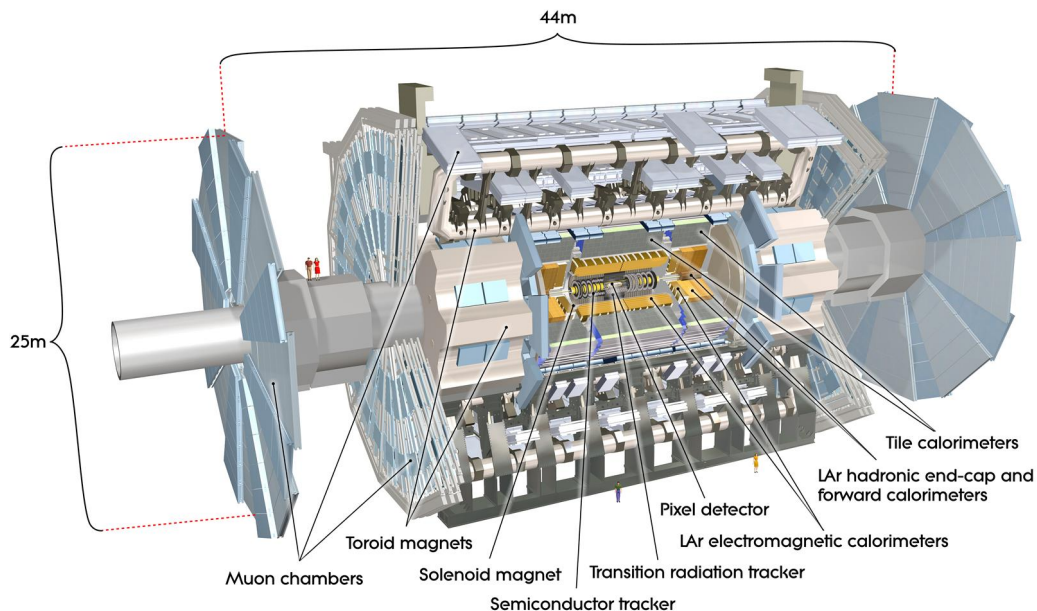


Figure 11: Computer generated image[17] of the whole ATLAS detector, showing its dimensions and subdetectors.

4.5 The ATLAS Experiment

The ATLAS¹ detector is a general purpose detector for hadron collisions[18]. It consists of a number of different subdetectors, which are layered around the interaction point, as seen in Figure 11.

4.5.1 Coordinate System

ATLAS uses a right-handed coordinate system with the nominal interaction point as the origin. The beam direction is defined as the z-axis, the x-axis as pointing to the center of the LHC ring and the y-axis upwards. The azimuthal angle Φ is defined in the transverse x-y-plane, with $\Phi = 0$ pointing to the x-axis. Transverse quantities such as the transverse momentum p_T are measured in this plane. From the polar angle θ , which is the angle from the z-axis, one defines the pseudorapidity $\eta = -\ln \tan(\theta/2)$.

¹A Toroidal LHC ApparatuS

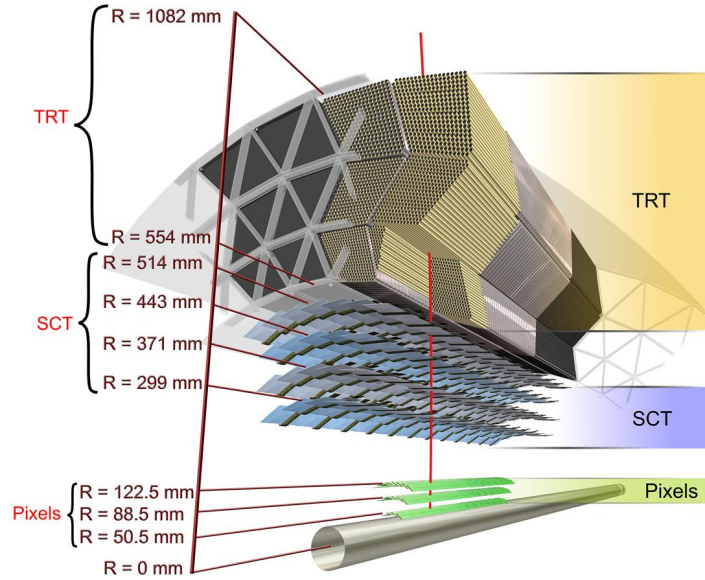


Figure 12: Drawing of the components of the Inner Detector.

4.5.2 Magnet Systems

Electrically charged particles are deflected in magnetic fields by the Lorentz force depending on their momenta and charge. Therefore, these properties can be measured by observing the curvature of tracks crossing a magnetic field. The ATLAS detector uses a solenoid surrounding the inner detector and toroid magnets for the muon spectrometer, one around the barrel, and two at the endcaps. Both use superconducting coils, cooled to 4.5 K.

4.5.3 Inner Detector

The inner detector (ID) consists of three subdetectors. Their schematic layout is shown in Figure 12. It measures the tracks of charged particles.

- The **pixel detector** consists of three layers of silicone pixel sensors. There are three cylindrical layers surrounding the beam axis and three discs of sensors on both endcaps, covering tracks with high η up to 2.5. A charged particle transversing the silicone pixel creates electron-hole pairs in the p-n junction, which can be detected by the readout

electronics. The pixel detector has about 80 million readout channels.

- The **semi conductor tracker (SCT)** surrounds the pixel detector with four layers around the barrel and nine disks each on both endcaps. In contrast to the pixel detector the SCT uses strips, not pixels, of silicone. To obtain information in the z-direction, two layers are rotated slightly with respect to each other.
- The **transition radiation tracker (TRT)** is the outermost sub-detector of the ID. It is comprised of 4 mm long straw tubes. In the barrel region, they are 144 cm long and aligned to the z-axis. In the endcap region, 37 cm tubes are oriented radially outward.

4.5.4 Electromagnetic Calorimeter

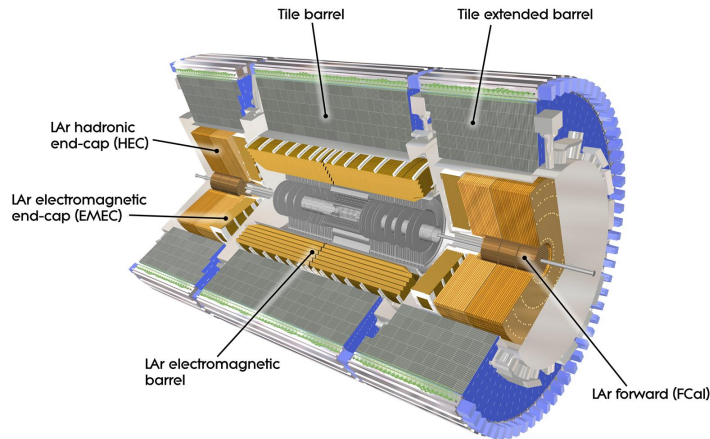


Figure 13: The ATLAS calorimeter system[19].

The electromagnetic calorimeter (ECAL) surrounds the inner detector in the barrel region and on the two endcaps. It is used to measure the energy of electromagnetically interacting particles like photons and electrons. It consists of an accordion shaped lead absorbers with gaps filled with liquid argon (LAr) as the active material.

Embedded between the lead absorbers are copper readout electrodes, to which a high voltage is applied, which leads to an electric field between them. Incoming particles create electromagnetic showers in the lead absorbers, which then ionize argon atoms. Due to the electric field the so

ionized electrons then drift to copper readout electrodes to be measured. The resolution can be roughly parametrized as $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$.

4.5.5 Hadronic Calorimeter

The hadronic calorimeter (HCAL) encompasses the ECAL. It is used to measure the energy of hadrons using the strong interaction. Hadrons interacting with the nuclei of the absorber produce showers which then get detected in the active material. The barrel region up to $|\eta| < 1$ and the extended barrel region in the range $0.8 < \eta < 1.7$ is covered by the tile calorimeter. Steel absorbers and scintillating tiles form a sampling calorimeter. At $\eta = 0$ it is 9.7 absorption lengths thick. The light produced in the scintillating tiles is read out by wavelength shifting fibers.

On the endcaps behind the electromagnetic endcap calorimeters are the LAr hadronic endcap calorimeters, extending between $1.5 < |\eta| < 3.2$. They use liquid argon as active and copper as absorber material. Near the beam pipe, in a range $3.1 < |\eta| < 4.9$, there are the LAr forward calorimeters, designed to measure jets with very high rapidity values. Those use again liquid argon as active material, but copper and tungsten as absorbers.

All above hadronic calorimeters are designed to overlap and cover as much of the solid angle around the interaction point as possible. This allows for good measurements of all produced particles and missing transverse energy.

4.5.6 Muon Spectrometer

The only particles to reach outside the ID, ECAL and HCAL, apart from neutrinos, which cannot be detected by ATLAS, are muons. The Muon Spectrometer (MS), as depicted in figure 14, is able to accurately determine p_T , η and ϕ those muons[21]. Three layers of muon chambers are placed around the beam axis in the barrel region, and three layers of chambers vertically at the endcaps. Monitored Drift Tubes (MDTs) are installed to cover a wide range of η , while Cathode Strip Chambers (CSC) with higher granularity are put close to the interaction point at high η .

MDTs are aluminum tubes with a diameter of 30 mm, filled with a non-flammable gas. Their walls have a thickness of 400 μm . They vary in length from 70 cm to 630 cm. In the center of the tube, a wire made from tungsten and rhenium, plated with gold, is stretched. It is held at a high electric potential. Passing muons create ion/electron pairs in the gas. The electrons drift towards the wire. Near the wire the electric field accelerates them, creating charge avalanches which are high enough to be measured. From the drift time, the spacial coordinate of the muon path is calculated with a

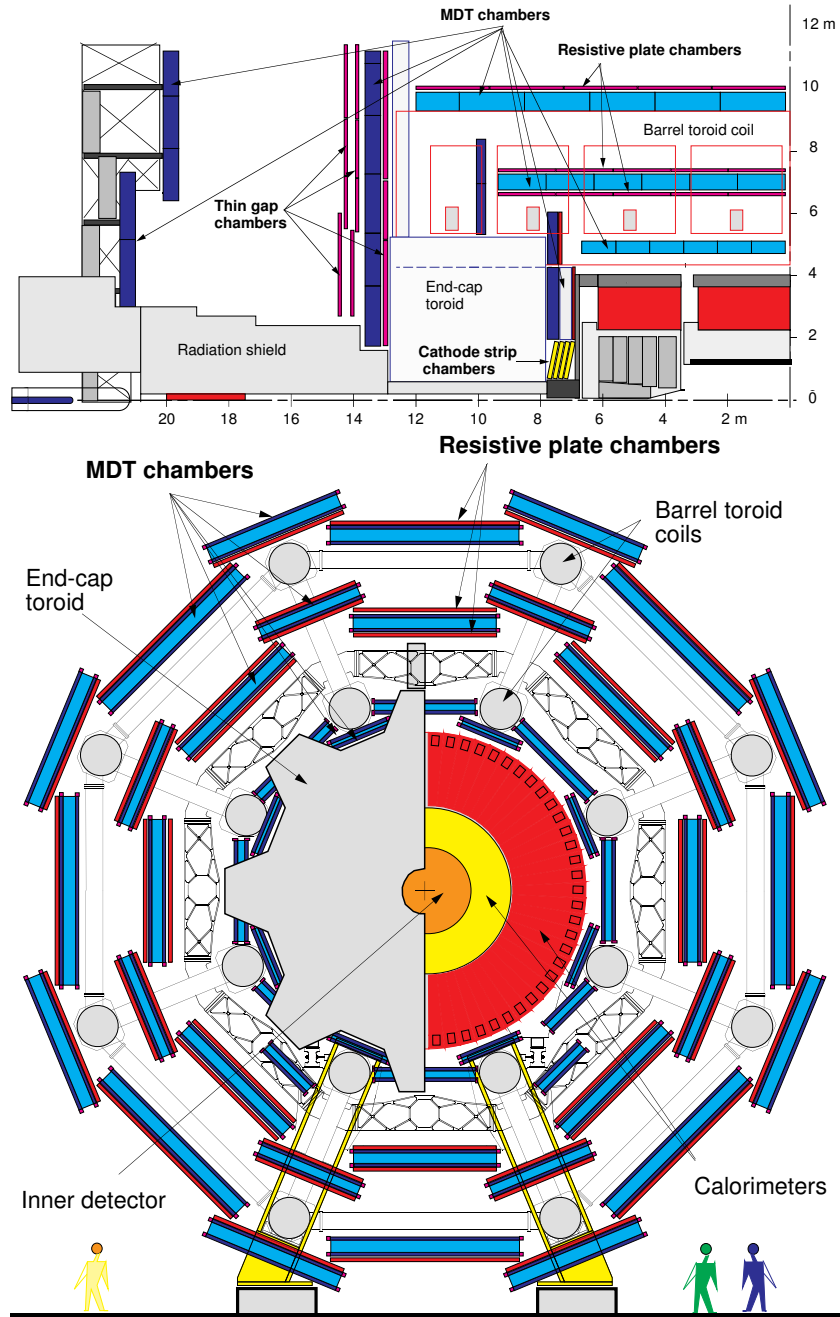


Figure 14: The ATLAS muon subsystems[20]. Shown is the view from the side (top) and in the transverse plane (bottom).

resolution of typically $80\text{ }\mu\text{m}$.

In the area where $2 < |\eta| < 2.7$, the count rate exceeds the maximum rate of MDTs, so Cathode Strip Chambers are used. They are multiwire proportional chambers with cathodes segmented into strips, with the wires oriented radially outward from the beam axis. The CSCs reach a resolution of up to $60\text{ }\mu\text{m}$.

Those two systems provide precise measurements for the muon tracks. However, they are too slow to be used for triggering and lack a second spacial coordinate. For that, fast Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) are employed in the barrel and endcap regions, respectively.

4.6 Trigger and Data Acquisition

Running the LHC at design settings causes about 40 million bunch-crossings inside the ATLAS detector per second. In each of those bunch-crossings multiple proton-proton interactions called events occur. Recording all those events would exceed today's technical capabilities. However, the cross-sections and therefore event rates of the interesting physics processes the LHC and ATLAS were designed to study are very small compared to the overall event rate. Therefore a multi-stage filtering system called "triggers" is in place to reduce the recording rate to events of interest only.

The Level-1 (L1) trigger, the first step, is hardware based. It uses a subset of the detector to look for signatures of interesting events, such as leptons with high transverse momenta and can handle a maximum rate of 100kHz. The geometric location of those objects in the detector are provided as Regions of Interest (RoI) to the Level-2 (L2) trigger.

The L2 trigger is implemented in software. Based on the data in the RoIs it confirms or rejects the findings of the L1 trigger, reducing the event rate to under 3.5kHz.

The Event Filter (EF) runs on a server farm near the detector. It uses offline reconstruction algorithms on the full detector information to reduce the event rate further down to around 200 Hz. These events are then recorded to tape in several data streams. These streams contain all events according to the class of triggers they fired. The egamma and muons streams, which are used in this analysis, hold all events that fired any electromagnetic or muon trigger. For some common events it would be unreasonable to record all of them. In order to get a representative sample of them, some triggers are prescaled. Prescaling a trigger by n means recording only one in every n events that pass said trigger.

4.7 Grid computing

The ATLAS experiment produces an enormous amount of data that has to be stored and processed. Similarly, a large amount of Monte Carlo events have to be simulated. This requires tremendous resources, both in computing power and storage capacity, which would easily overwhelm a typical data center.

In order to master this challenge, ATLAS embraces the Grid paradigm [22]. The Grid is a collection of data centers or “sites“, typically at universities and research facilities, distributed worldwide and connected via the Internet. The ATLAS computing Grid is designed hierarchically in levels called Tiers. The CERN site, where all recorded data from the ATLAS detector is initially input, is Tier0. From there, the data is replicated to the eleven privileged sites of Tier1. There the data is stored long term and reprocessed as needed. Each Tier1 site has a number of additional Tier2 sites associated with it, which contribute computing power for simulating Monte Carlo and processing data. This collection of Tier2 sites around a Tier1 is called a cloud. For example, a computing cluster at the Leibniz-Rechenzentrum (LRZ) is the Tier2 site of the LMU University Of Munich. It belongs to the German “de” cloud around the Grid Computing Centre Karlsruhe (GridKa).

Authorized users of the Grid can submit requests to perform a certain task on a specified dataset, so called jobs. This allows for a quick completion of computing intensive tasks. Most stages of this analysis have been performed as jobs in the Grid.

5 Data and Simulated Samples

5.1 Data

This analysis uses the full dataset recorded from the 2012 LHC run at $\sqrt{s} = 8$ TeV. After data quality requirements are applied, this corresponds to approximately 20.7fb^{-1} . Data from the muon and egamma stream are combined and duplicate events removed.

5.2 Monte Carlo Simulation

To understand what different physical processes in the experiment look like, samples of these processes are simulated using Monte Carlo (MC) generators. Different aspects of proton-proton collisions are simulated by a number of different software packages. First, the main parton-parton interaction is calculated by an event generator. This involves evaluating complex integrals without analytical solutions. Approximate solutions can be obtained by using methods which employ pseudo-random number generators, so called Monte Carlo Integration. Because of confinement, produced free partons then have to be hadronized into bound states. This process is called showering. After that, some produced particles can decay. This too has to be simulated. The response of the ATLAS detector to a simulated event is then computed using a virtual computer model with the Geant4 [23] software. Signal and background processes have been simulated using Monte Carlo event generators.

- **Pythia** [24] is a general purpose generator. It can simulate a wide variation of processes. It includes its own parton showering, which can be used by other generators as well. A lot of the following generators use this capability.
- **Alpgen** [25] is a leading order generator that computes matrix elements for a fixed number of partons in the final state. It is therefore especially suited for processes with a high multiplicity of jets in the final state.
- **MC@NLO** [26] is a generator for hard processes. It calculates matrix elements in next to leading order. MC@NLO needs an external program for hadronization.
- **gg2WW** [27] and **gg2ZZ** are specialized generators for di-boson generation from gluon fusion.

- **AcerMC** [28] is used to simulate some Standard Model processes that are not covered by the other generators, in the case of this analysis the t-channel single top background. It can only calculate matrix elements in leading order. For showering, it can use different generators such as Pythia.
- **POWHEG BOX** [29] is a generator that can in principle generate any next-to-leading-order process. It implements a number of Standard Model processes like di-boson and single-top production. It is also capable of simulating Higgs production by gluon-gluon fusion [30] and vector boson fusion [31]. It can use a variety of other tools for showering, including Pythia.

5.2.1 Signal Samples

The signal samples are simulations of Standard Model Higgs boson production with subsequent decay into muon pairs. The three dominant production modes as depicted in figure 6 are simulated: Gluon gluon fusion, denoted ggF, vector boson fusion, denoted VBF, and production in association of a W or Z boson, called VH.

The ggF and VBF samples have been simulated using POWHEG for the hard process and Pythia8 for the showering. To describe the parton distribution functions of the proton, the CT10 [32] PDF set is used with the ATLAS Underlying Event Tune, AU2 [33]. To agree with the predicted distribution, the ggF Higgs boson pT spectrum in POWHEG is reweighted. The WH and ZH processes are simulated using Pythia8. The AU2 and the CTEQ6 [34] PDF set are used for these samples. To test different mass hypotheses, all signal samples have been simulated for a number of assumed Higgs masses in 5 GeV steps from 100 GeV to 150 GeV.

5.2.2 Background Samples

In this analysis, the background distribution is extracted from a fit to the measured data. The simulated background samples are used to verify the quality of the used fit function.

The dominant contribution stems from the $Z/\gamma^* \rightarrow \mu^+\mu^-$ Drell-Yan process. Samples from two different MC generators are used for cross checks. As the default sample a sample that has been simulated using the same generator setup as for the ggF and VBF signal has been used because of its high statistics. A sample with cut on $m_{\mu\mu} > 60$ GeV on generator level, containing the Z-peak, is used. A cut on the Monte Carlo level Z-boson mass is then applied to exclude the region $120 \text{ GeV} < m_{\mu\mu} < 250 \text{ GeV}$. This

region is filled again with high statistics samples that were produced with the invariant mass binned in the region from 120 GeV to 180 GeV and in the region from 180 GeV to 250 GeV. This procedure helps to reduce the statistical error in the high mass tail of the $m_{\mu\mu}$ spectrum, since most of the events of the inclusive sample are concentrated near the Z-peak.

Alternative samples of Z-Drell-Yan $\rightarrow \mu^+\mu^-$ in association with jets were produced with the Alpgen generator. Since Alpgen is specialized for multi-parton final states, the Data-MC agreement of the jet multiplicity is better when this sample is used, as seen in figure 15.

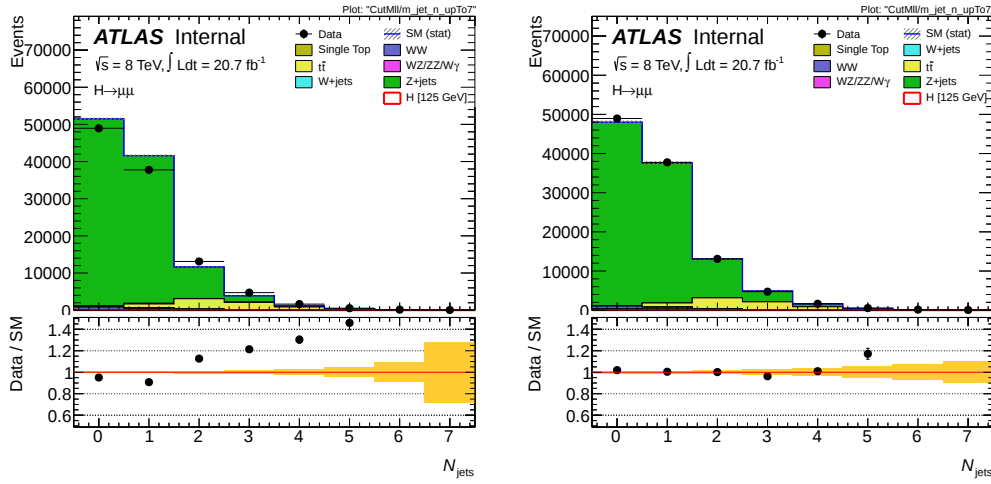


Figure 15: Jet multiplicity for POWHEG (left) and Alpgen (right) generated Z/γ^* background. The lower part of each plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the statistical uncertainty. The signal is shown in red for $m_H = 125$ GeV.

However, this analysis does not depend critically on jets, so the high statistics sample is used instead as the primary sample, and the Alpgen sample only for cross checks.

For the top pair backgrounds, samples generated with MC@NLO are used. In this samples a filter is applied to consider only events where both W bosons which were created in the top decay, decay leptonically. Single top events were generated using MC@NLO and AcerMC, depending on their production channel. The di-boson backgrounds $q\bar{q} \rightarrow WW/WW/ZZ$ were simulated using the same setup as for the ggF and VBF signal processes. The $q\bar{q} \rightarrow W\gamma$ events were generated using Alpgen and Madgraph.

5.3 Corrections to Monte Carlo

5.3.1 Muon Momentum Corrections

The muon momentum resolution as measured in data is worse than estimated in the simulations. To bring them into agreement, the momentum of muons in MC is corrected by a process called smearing. Additionally, a muon momentum scale is applied to correct the different momentum scales in MC and data. This process is described in detail in [35].

5.3.2 Scale Factors

The triggers in simulation do not behave exactly like the real triggers. In order to align MC and data, scale factors are introduced as

$$\text{scalefactor} = \frac{1 - \prod_{n=1}^N (1 - \epsilon_{Data,n})}{1 - \prod_{n=1}^N (1 - \epsilon_{MC,n})}, \quad (16)$$

where N is the number of objects which satisfy the object selection criteria, and $\epsilon_{Data,n}$ and $\epsilon_{MC,n}$ are the trigger efficiencies as determined in data and Monte Carlo, respectively. The package TrigMuonEfficiency is used to calculate these scale factors. They are then used to reweight the MC to get the desired agreement.

5.4 Jet Reconstruction

Jets, as mentioned in section 3.1.1, are collections of hadrons which travel in roughly the same direction, originating from a scattered parton. In the detector, they are registered as energy deposits in narrow cones.

Software algorithms are then used to construct a set of jets from the energy deposits in the calorimeters. The Anti- k_t Jet Algorithm [36] uses an iterative method to combine calorimeter objects into jets. It defines a distance between two objects i and j as follows:

$$d_{ij} = \min(k_{T,i}^{-2}, k_{T,j}^{-2}) \cdot \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \quad (17)$$

and the distance between object i and the beam as

$$d_{i,B} = k_{T,i}^{-2}. \quad (18)$$

where $k_{T,i}$ is the transverse momentum, y_i the rapidity and ϕ_i the azimuthal angle of object i . In each iteration, the distances are calculated. If the smallest distance is one of the d_{ij} , then the two objects i and j corresponding to this d_{ij} are combined by adding their four-vector. In the other case, if the minimum distance is one of the $d_{i,B}$, the object is not mergeable and taken from the lists of objects to be merged to the list of output jets. This process is repeated until all objects have been assigned to jets.

In this analysis, the free parameter R of this algorithm is chosen to be 0.4. Jets are selected if they are in the region of $|\eta| < 4.5$.

5.5 Muon Reconstruction

In ATLAS muons can be identified using two reconstruction methods called “MuID” and “STACO” [37]. Both implement three ways of muon reconstruction.

- **Standalone** muons use only data from the muon spectrometer. Muon tracks are fitted and then extrapolated back to the inner detector, including the effects from the inhomogeneous magnetic field and the energy loss in the calorimeters.
- **Combined** muons are obtained by matching the extrapolated track of a standalone muon to a track in the inner detector. Both tracks are then fitted together as a combined track from hits in both detectors.
- **Tagged** muons are identified by extrapolating tracks from the inner detector to hits in the muon spectrometers in areas where coverage in the MS is incomplete.

This analysis uses STACO (“STAtistical COmbination”) combined muons with tight quality requirements on the number of hits in the pixel detector, SCT and TRT in the inner detector. Muons that overlap with jets within a cone distance of 0.3 are removed. Selected muons are required to be isolated, which means that the energy and transverse momentum of all other particles in a cone of radius $\Delta R < 0.3$ around them (p_T^{cone30} and E_T^{cone30} respectively) must not be too large. Additionally, cuts on the impact parameters d_0 and z_0 suppress muons not originating from the primary vertex. Events with a significance of the transverse impact parameter $\frac{d_0}{\sigma_{d_0}} > 3$, with σ_{d_0} being the error of the d_0 measurement, are rejected. Table 3 shows the cut values on these parameters.

Cut	Value
E_T^{cone30}/p_T	$E_T^{cone30}/p_T < 0.014 \cdot p_T[\text{GeV}] - 0.15$ and $E_T^{cone30}/p_T < 0.20$
p_T^{cone30}/p_T	$E_T^{cone30}/p_T < 0.01 \cdot p_T[\text{GeV}] - 0.105$ and $p_T^{cone30}/p_T < 0.15$
d_0 significance	< 3.0
$z_0 \sin \theta$	$< 1.0 \text{ mm}$

Table 3: Cuts on the muon isolation and impact parameters.

5.6 Missing Transverse Energy

The momentum of the colliding protons is almost completely oriented along the z-axis of the ATLAS detector. Due to conservation of four-momentum, the momenta of all outgoing particles in the plane transverse to the z-axis have to sum up to zero. However, neutrinos and some hypothesized new particles as predicted by theories like Super Symmetry cannot be detected by ATLAS. They therefore lead to so called missing transverse energy or E_T^{miss} . Apparent E_T^{miss} will also be caused by detector effects such as limited energy resolution and defective calorimeter cells. This is illustrated in figure 16, where the signal and the Z+jets background show a small but finite E_T^{miss} , whereas backgrounds that include neutrinos in the final state like the $t\bar{t}$ -production show larger missing transverse energy.

Missing energy is only calculated in the transverse plane, because the momentum of the interacting partons along the z-axis is not known. For them, only a probabilistic description of their momentum fraction in the proton is known, as described in section 3.1.

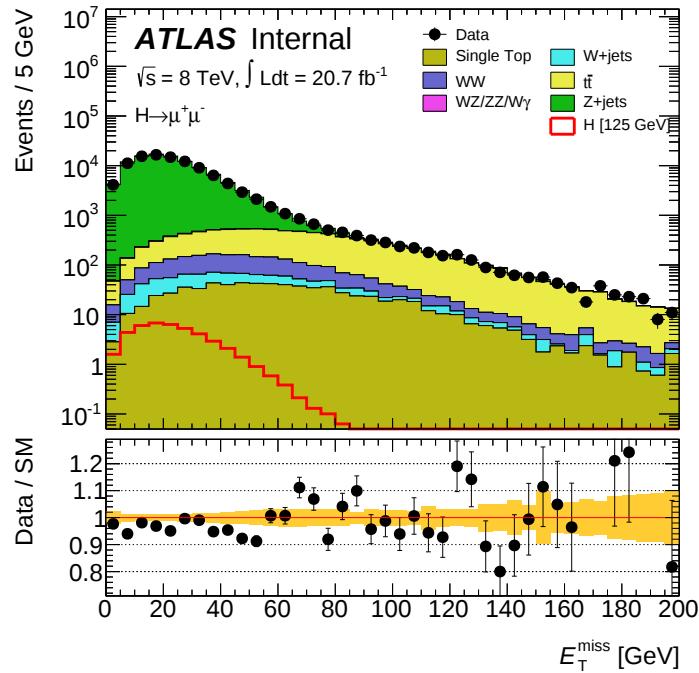


Figure 16: Missing transverse energy E_T^{miss} in data and Monte Carlo. The lower part of the plot shows the ratio between the data and the background expectation from MC, with the yellow band indicating the statistical uncertainty. The signal is shown in red for $m_H = 125$ GeV.

5.7 Blinding

The search for a faint signal in a large background is susceptible to bias stemming from the choice of the details of the statistical methods used. Features in the data which in reality result from a statistical fluctuation, may be interpreted by the experimentalist as a true signal. The analysis can then be fine-tuned to enhance the apparent significance of this feature. While this can be a deliberate action by the experimentalist, it can also be caused by a subconscious desire to “optimize” one’s results. This effect is well known in many fields of science.

The only effective method to counteract this is what is called blinding. It involves developing the complete analysis before looking at the relevant data. This prevents the data from influencing the researcher in his or her choice of methods. After the analysis has been fixed, the data is unblinded, which means it is now used to derive statistical results.

This strategy is used in ATLAS searches for the reasons described above. In the $H \rightarrow \mu^+\mu^-$ search the $m_{\mu\mu}$ spectrum was blinded by filtering out all events in the search window of $105 \text{ GeV} < m_{\mu\mu} < 160 \text{ GeV}$ in the high- $p_T^{\mu\mu}$ region. Only when selection cuts, signal and background parametrization and limit procedure were fixed, all events were examined. This technique is called the *hidden signal box* method [38].

The analysis was then optimized using Monte Carlo simulated samples for signal and background, as well as the low- $p_T^{\mu\mu}$ control region, which was assumed to contain almost no signal contamination.

6 Event Selection

To be selected, events need to have a primary vertex, consistent to the beam spot position that has at least three particle tracks originating from it. A number of quality requirements on the event have to be fulfilled.

6.1 Muon Trigger

For this analysis the EF_mu24i_tight and EF_mu36_tight triggers are used. In these names, EF stands for the Event Filter, the author of the triggers, the number after mu for muon for the p_T -threshold of the trigger in GeV, and the tight for triggers seeded by L1 trigger with a 15 GeV p_T -threshold.

The i in EF_mu24i_tight stands for isolation of the online muon. It is required to fulfill the relative track isolation condition that $\Sigma p_T/p_T(\mu) < 0.12$. Here, Σp_T is the scalar sum of all transverse momenta of tracks with $p_T > 1$ GeV, detected in an ID cone of $\Delta R = 0.2$ around the candidate muon. The absolute difference of z_0 between the EF ID track and the EF muon track must be smaller than 6 mm, which is applied to the above sum.

Those events are considered in which one or both of these two triggers have fired. Further it is required that one of the offline reconstructed muons was what fired the triggers. One of the muons has to match one of the trigger objects within a radius of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.15$. This is called trigger matching.

6.2 Selection Cuts

Candidate events must contain exactly two muons with opposite charge. No other muons are allowed. The muon with the higher p_T , the so called leading muon, has to have a $p_T > 25$ GeV, while the other, the subleading one, needs a $p_T > 15$ GeV. A cut on the transverse momentum of the di-muon-system $p_T^{\mu\mu} > 15$ GeV defines the signal region. The cuts on leading and subleading muon p_T and $p_T^{\mu\mu}$ have been optimized using a 3d scan on cuts on these variables. The $m_{\mu\mu}$ -range between 105 GeV and 160 GeV is then studied.

Due to the better resolution of the detector for low values of η , the signal gets split into two resolution categories based on the pseudorapidity of the muons. This is done to increase sensitivity. Events where both muons lie in the area where $|\eta| < 1$ are put into the “central” category, the rest in the “non-central” category. The cut of $\eta = 1$ was chosen by two methods. First, optimizing the combined significance from both categories. The significance in each category defined as the poisson significance

$$\sigma_i = \sqrt{2((s_i + b_i) \cdot \ln(1 + s_i/b_i) - s_i)}, \quad (19)$$

where i stands for the central and non-central regions. s_i and b_i are the number of signal and background Monte Carlo events in a window of ± 5 GeV around the assumed Higgs mass. In this window, over 90% of signal events are contained. These per-category significances are then combined using two different formulas,

$$\sigma_{comb}^{simple} = \frac{s_1 + s_2}{\sqrt{2}} \quad (20)$$

and

$$\sigma_{comb}^{quad} = \sqrt{s_1^2 + s_2^2} \quad (21)$$

Figure 17 shows that both ways of combining yield a maximum significance near a cut of $|\eta| = 1$.

The other way to judge the optimal cut value is to consider the width of the signal function, fitted to signal events in varying η regions. As seen in figure 18, the width of the signal peak increases sharply for $|\eta| > 1$. This agrees with measurements of the di-muon mass resolution done on $Z \rightarrow \mu^+ \mu^-$ events as seen in figure 19.

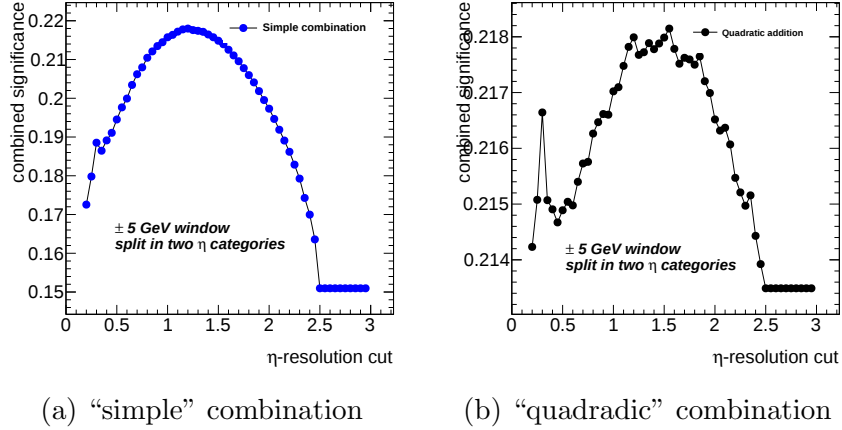


Figure 17: Dependence of the combined significance on the $|\eta|$ cut for the resolution categories.

The Higgs boson in the signal tends to a high transverse momentum. The reversal of the $p_T^{\mu\mu}$ -cut, which exploits this fact, defines a control region. Using data from this low- $p_T^{\mu\mu}$ control region the background fits were tested prior to unblinding. As described in section 5.7, it is crucial to obey the blinding procedure in order not to bias the final result. An overview of expected and observed events after each cut is given in table 4.

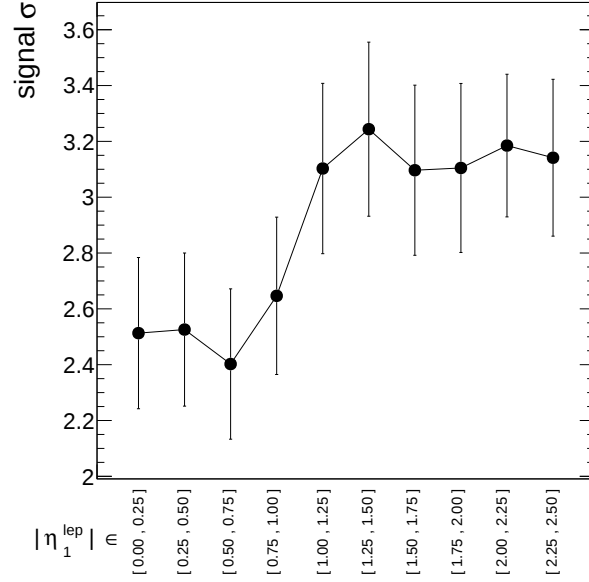


Figure 18: Width of the signal pdf, fitted to signal MC in various $|\eta|$ ranges.

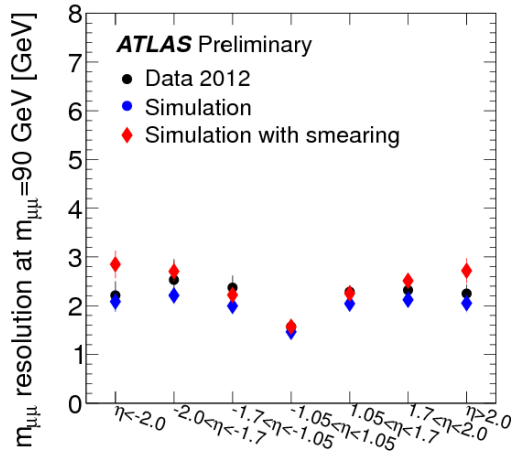


Figure 19: Resolution of the di-muon invariant mass at $m_{\mu\mu} = 90$ GeV, taken by combining measurements from inner detector and muon spectrometer [39].

	Sig. [125 GeV]	WW	WZ/ZZ/W γ	$t\bar{t}$	Single Top	Z+jets	W+jets	Total Bkg.	Observed	Data/MC
mc weights	57.0 $\pm$ 0.3	6710 $\pm$ 20	4605 $\pm$ 15	32127 $\pm$ 64	3246 $\pm$ 24	8931802 $\pm$ 5519	14760 $\pm$ 172	8993250 $\pm$ 5523	9205567 $\pm$ 3034	1.02 $\pm$ 0.00
OS leptons	56.8 $\pm$ 0.3	4280 $\pm$ 16	3275 $\pm$ 10	20963 $\pm$ 51	2167 $\pm$ 19	8720767 $\pm$ 5459	10890 $\pm$ 165	8762344 $\pm$ 5461	8593165 $\pm$ 2931	0.98 $\pm$ 0.00
lead lepton	56.8 $\pm$ 0.3	4280 $\pm$ 16	3275 $\pm$ 10	20963 $\pm$ 51	2167 $\pm$ 19	8720767 $\pm$ 5459	10890 $\pm$ 165	8762344 $\pm$ 5461	8593165 $\pm$ 2931	0.98 $\pm$ 0.00
sublead lepton	56.4 $\pm$ 0.3	4207 $\pm$ 16	3194 $\pm$ 10	20504 $\pm$ 50	2113 $\pm$ 19	8639590 $\pm$ 5432	10890 $\pm$ 165	8680499 $\pm$ 5435	8514523 $\pm$ 2918	0.98 $\pm$ 0.00
$P_T^{\ell\ell} > 15$ GeV	44.1 $\pm$ 0.2	3732 $\pm$ 15	2734 $\pm$ 10	19684 $\pm$ 49	2033 $\pm$ 19	3608084 $\pm$ 3508	5322 $\pm$ 117	3641590 $\pm$ 3510	3505195 $\pm$ 1872	0.96 $\pm$ 0.00
$108 < m_{\ell\ell}[\text{GeV}] < 160$	43.5 $\pm$ 0.2	1254 $\pm$ 9	156 $\pm$ 3	6970 $\pm$ 29	723 $\pm$ 11	100843 $\pm$ 490	350 $\pm$ 21	110295 $\pm$ 492	106900 $\pm$ 327	0.97 $\pm$ 0.01
Central region	12.7 $\pm$ 0.1	227 $\pm$ 4	29 $\pm$ 1	1978 $\pm$ 15	208 $\pm$ 6	22111 $\pm$ 229	57 $\pm$ 9	24611 $\pm$ 230	23458 $\pm$ 153	0.95 $\pm$ 0.01
Non central region	30.8 $\pm$ 0.2	1026 $\pm$ 8	127 $\pm$ 3	4992 $\pm$ 25	515 $\pm$ 10	78731 $\pm$ 433	293 $\pm$ 19	85684 $\pm$ 434	83442 $\pm$ 289	0.97 $\pm$ 0.01
$ M_H - m_{\ell\ell} \leq 5\text{GeV}$	37.7 $\pm$ 0.2	250 $\pm$ 4	30 $\pm$ 1	1374 $\pm$ 13	151 $\pm$ 5	15806 $\pm$ 124.91	59 $\pm$ 9	17669 $\pm$ 126	17442 $\pm$ 132	0.99 $\pm$ 0.01
$P_T^{\ell\ell} < 15$ GeV	12.4 $\pm$ 0.1	474 $\pm$ 5	460 $\pm$ 3	820 $\pm$ 10	81 $\pm$ 4	5031505 $\pm$ 41481	5568 $\pm$ 116	5038909 $\pm$ 4150	5009328 $\pm$ 2238	0.99 $\pm$ 0.00

Table 4: The expected numbers of background events at after selecting two muons at different stages in the cut-flow. The observed numbers of events in data are also given in the second last column and the ratio Data over MC in the last column. The background estimates are entirely from MC. The top half of the table shows the event yields of the successive cuts, while the lower half shows the event yields per line in different control regions after the $m_{\mu\mu}$ requirement. Only statistical uncertainties are shown in this table.

6.3 Higgs $\rightarrow \tau^+\tau^-$ Contamination

Tau leptons can decay as $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$. Therefore, a $H \rightarrow \tau^+\tau^-$ event, where both taus decay into muons, will have two oppositely charged muons with a high transverse momentum. This is very similar to the signal signature of the $H \rightarrow \mu^+\mu^-$ search. However, the tau decay involves neutrinos, which carry away energy and momentum. The narrow peak in the $m_{\mu\mu}$ distribution, which is the final sign for the $H \rightarrow \mu^+\mu^-$ signal, is therefore not present in the $H \rightarrow \tau^+\tau^-$ process.

This can be seen in figure 20: In the signal region of $m_{\mu\mu} > 105$ GeV almost no $H \rightarrow \tau^+\tau^-$ events are expected; they are shifted to lower $m_{\mu\mu}$ values. Table 5 shows a comparison cutflow which shows that the expected number of $H \rightarrow \tau^+\tau^-$ events in the signal region is much smaller than one. This is why contributions from this channel can safely be ignored, and this analysis is orthogonal to the $H \rightarrow \tau^+\tau^-$ searches of other groups.

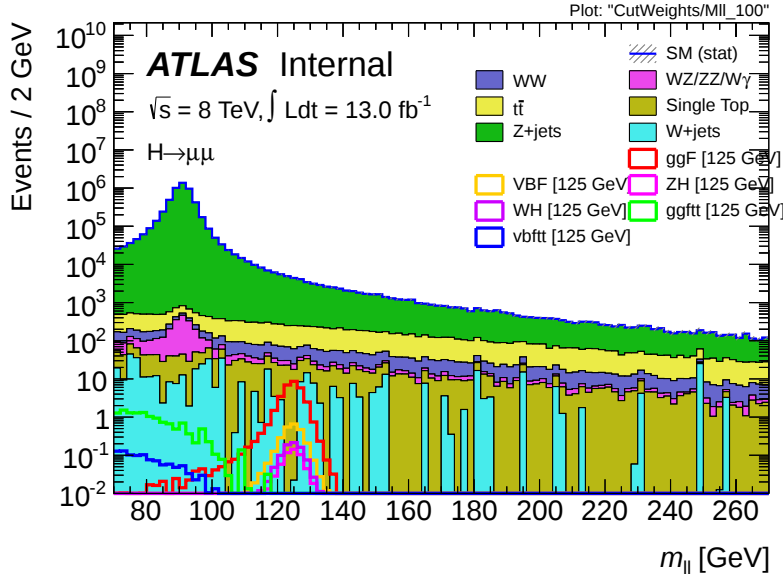


Figure 20: The distributions invariant mass $m_{\mu\mu}$ for the di-muon selection with all event selection requirements. The MC is scaled to only a partial dataset of 13fb^{-1} . The signal is shown for $m_H = 125$ GeV.

	ggF	WH	ZH	vbf	VBF tautau	ggF tautau
mc weights	32.0 ± 0.2	0.94 ± 0.01	0.57 ± 0.01	2.60 ± 0.01	3.6 ± 0.1	37 ± 1
OS leptons	32.0 ± 0.2	0.92 ± 0.01	0.57 ± 0.01	2.60 ± 0.01	3.6 ± 0.1	37 ± 1
lead lepton $p_T > 25$ GeV	32.0 ± 0.2	0.92 ± 0.01	0.57 ± 0.01	2.60 ± 0.01	3.6 ± 0.1	37 ± 1
sublead lepton $p_T > 15$ GeV	31.8 ± 0.2	0.91 ± 0.01	0.56 ± 0.01	2.57 ± 0.01	2.98 ± 0.09	31 ± 1
$P_T^{\ell\ell} > 15$ GeV	22.6 ± 0.1	0.87 ± 0.01	0.54 ± 0.01	2.49 ± 0.01	2.69 ± 0.08	21.2 ± 0.8
$105 < m_{\ell\ell} [\text{GeV}] < 160$	22.2 ± 0.1	0.85 ± 0.01	0.53 ± 0.01	2.46 ± 0.01	0.03 ± 0.01	0.15 ± 0.07
Central region	12.4 ± 0.1	0.43 ± 0.01	0.28 ± 0.00	1.43 ± 0.01	0.02 ± 0.01	0.07 ± 0.05
Non central region	9.82 ± 0.09	0.42 ± 0.01	0.25 ± 0.00	1.02 ± 0.01	0.01 ± 0.01	0.08 ± 0.05
$ M_H - m_{\ell\ell} \leq 2.5 \text{ GeV}$	13.2 ± 0.1	0.47 ± 0.01	0.31 ± 0.00	1.41 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
$P_T^{\ell\ell} < 15$ GeV	9.2 ± 0.1	0.03 ± 0.00	0.02 ± 0.00	0.08 ± 0.00	0.29 ± 0.03	10.1 ± 0.6

Table 5: The expected numbers of signal events after selecting two muons at different stages in the cut-flow. The two rightmost columns list the $H \rightarrow \tau\tau$ event yields. Only statistical uncertainties are shown in this table. Numbers are scaled to a partial dataset of 13fb^{-1} . The signal is shown for $m_H = 125$ GeV.

6.4 W+Jets Estimation

Estimating the W+jets background in the signal region is difficult due to the lack of sufficient MC statistics in this region. Therefore a data-driven method [40] is used. To avoid unphysical artefacts in the $m_{\mu\mu}$ distribution near the Z-peak, the estimation is carried out in a same-sign control region, where both muons have the same charge. This control region shows (figure 21) data/MC agreement within statistical errors for most values of $m_{\mu\mu}$. Systematic errors have not been evaluated in this analysis. In the $H \rightarrow WW$ analysis [41], which uses a similar selection, the systematic uncertainty on the W+Jets background is given as approximately 30%. The W+jets estimation is then scaled by $r = 2.8 \pm 0.1$, which is the ratio of the Monte Carlo prediction in the opposite-sign over same-sign regions, and used as the final W+jets estimation. This accounts for the different event rates in both regions.

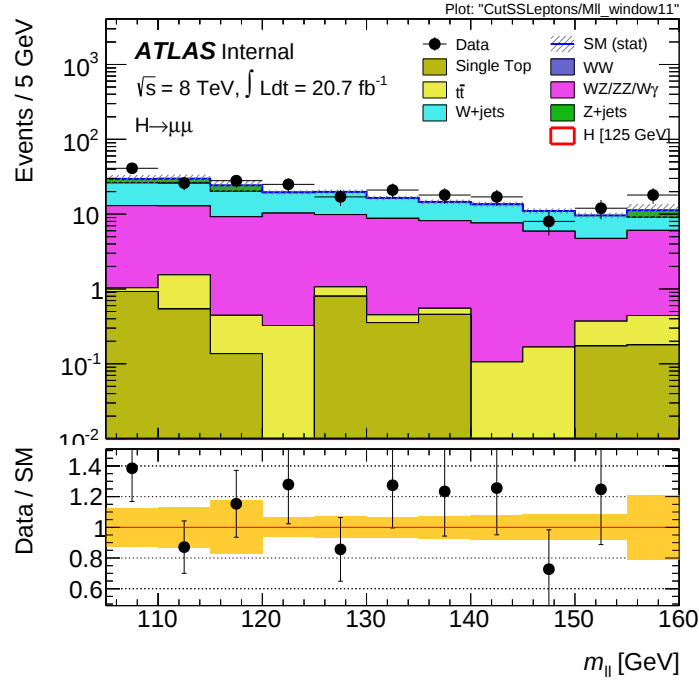


Figure 21: Distribution of the dimuon invariant mass $m_{\mu\mu}$ in the same sign control region. Only statistical errors are shown.

7 Control Plots

In order to assess whether the background is well understood, several kinematic variables are compared in data and Monte Carlo. In all following plots, the yellow band represents the statistical errors of the Monte Carlo samples. The distribution of the pseudorapidity and transverse momentum of leading and subleading muon after the selection cuts is shown in figures 22 and 23, respectively. The opening angle in the transverse plane is depicted in figure 24.

The final $m_{\mu\mu}$ distributions per resolution category from which the results are obtained are shown in figure 25.

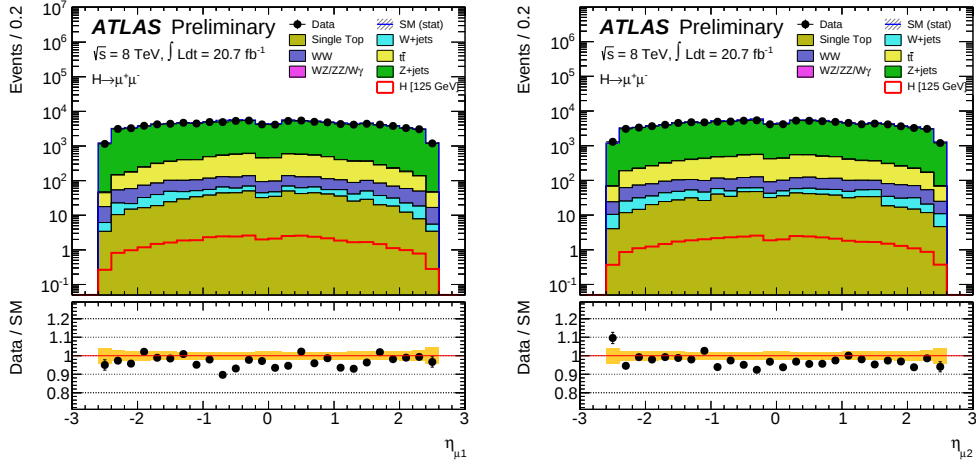


Figure 22: Pseudorapidity η of leading and subleading muons in data and Monte Carlo.

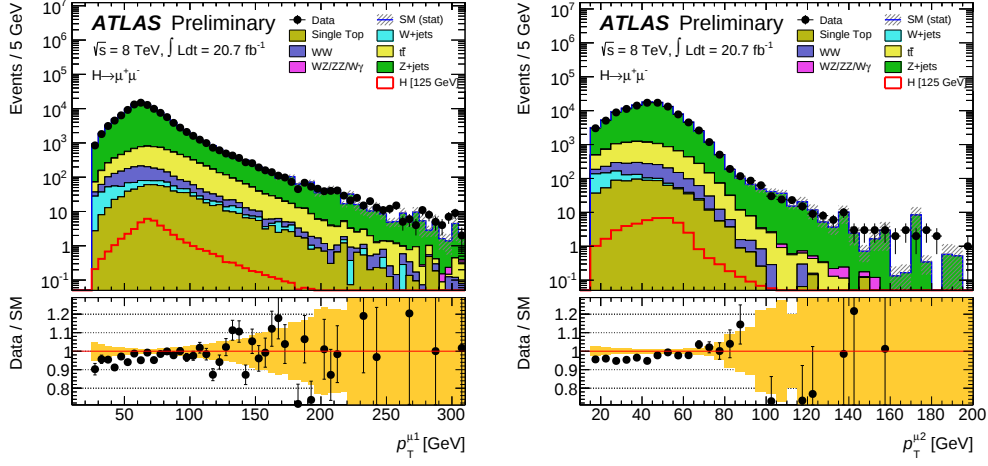


Figure 23: Transverse momentum p_T of leading and subleading muons in data and Monte Carlo.

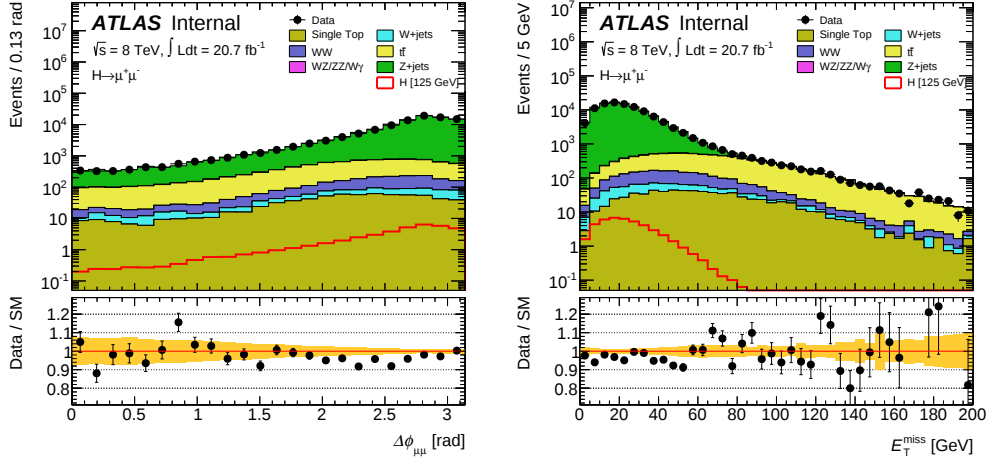


Figure 24: Transverse opening angle $\Delta\phi_{\mu\mu}$ and missing transverse energy E_T^{miss} in data and Monte Carlo.

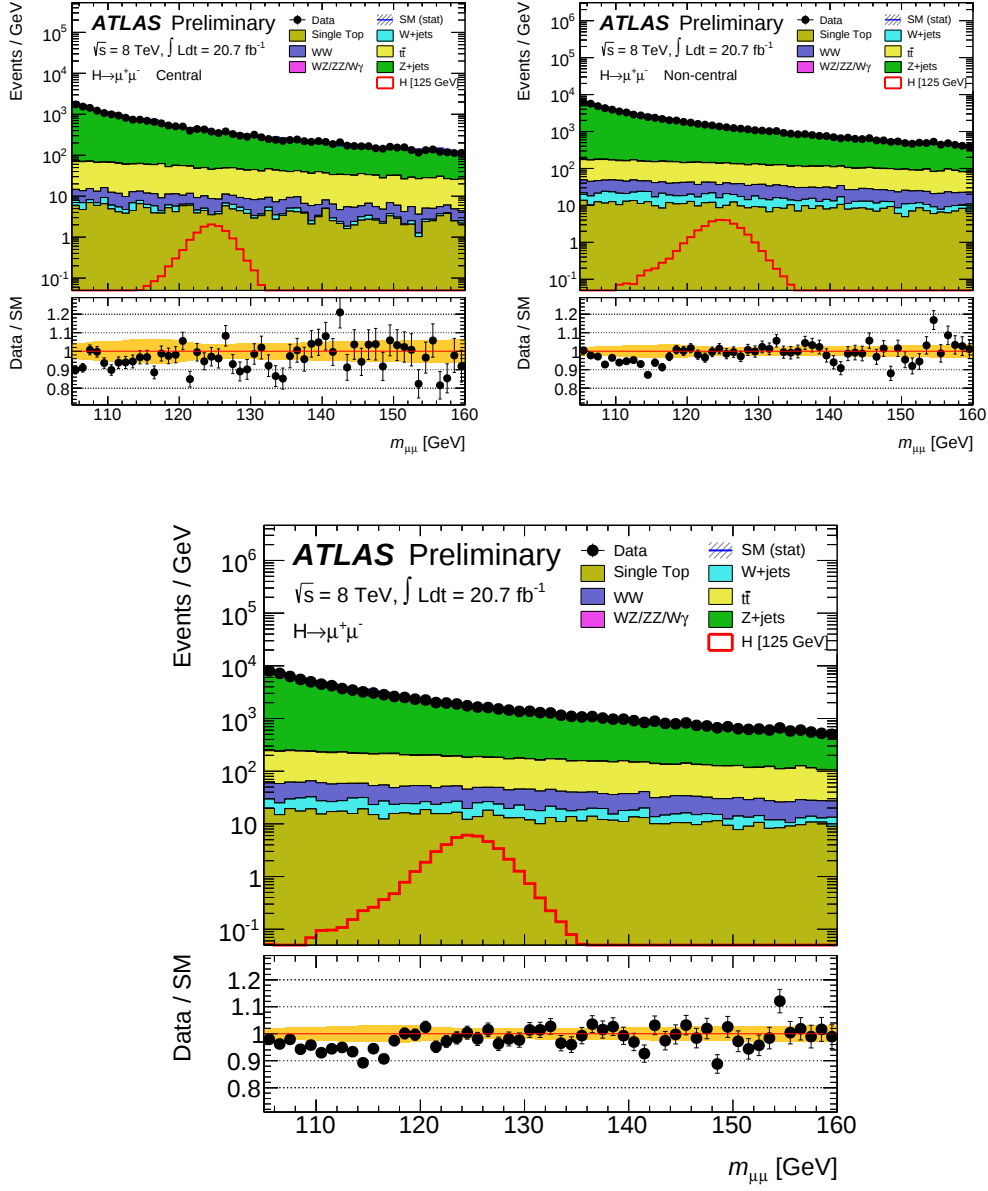


Figure 25: Invariant mass distribution in the central (left) and non-central (right) resolution category. The bottom plot shows both categories combined.

8 Signal and Background Description

8.1 Signal Modeling

The signal signature is a peak in the measured invariant di-muon mass distribution. The shape of this peak has been studied in Monte Carlo. It is modeled using an analytic function. There is a tail in the distribution to lower invariant masses, caused by final state photon radiation from the muons. To describe this asymmetry, a simple Gaussian (GS) does not suffice. This has been addressed by adding a second function. Two such functions have been tested, the Landau and the Crystal Ball (CB) distributions, the later of which was chosen for its fit stability. So the probability density function for the signal model reads

$$P_S = f_{CB} \cdot CP(x, m, \sigma_{CB}, \alpha, n) + (1 - f_{CB}) \cdot GS(x, m, \sigma_G), \quad (22)$$

where x is the running di-muon mass, σ_{CB} and σ_G the widths of Gaussian and Crystal Ball, and m the peak value of both. To increase the stability of the fit, the fraction parameter f_{CB} is fixed to 0.9, and n to 2.0.

This function is then fitted to the $m_{\mu\mu}$ -spectra obtained from Monte Carlo in both resolution categories. This is shown for the sample with $m_H = 125$ GeV in figure 26. The process is repeated for all 11 available signal samples, which have been produced in 5 GeV steps from $m_H = 100$ GeV to $m_H = 150$ GeV. In principle the $H \rightarrow \mu^+ \mu^-$ channel allows for a better Higgs mass resolution than this 5 GeV. In order to generate Higgs mass hypotheses in between these points, interpolation is used. All fitted parameters from above are independently interpolated as a function of m_H by fitting them with first order polynomials. The total number of events is fitted with a second order polynomial. This allows for constructing a signal pdf for an arbitrary m_H -hypothesis, as illustrated in figure 29.

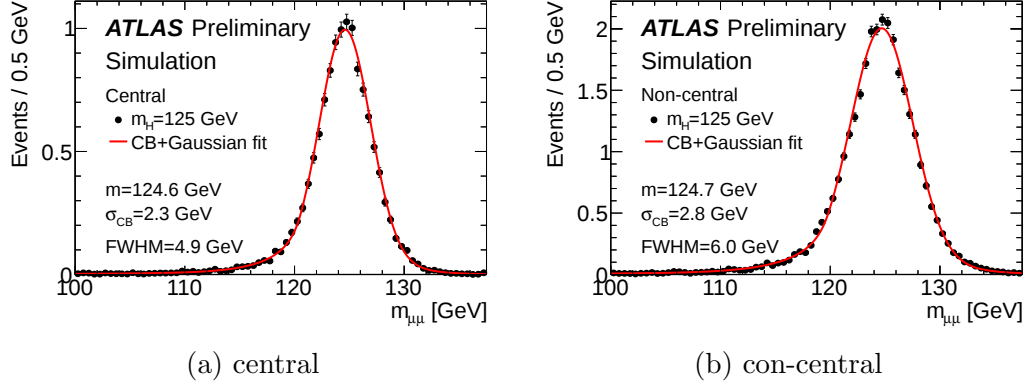


Figure 26: Fit of the signal model to the $m_{\mu\mu}$ spectrum of the $m_H = 125$ GeV MC sample

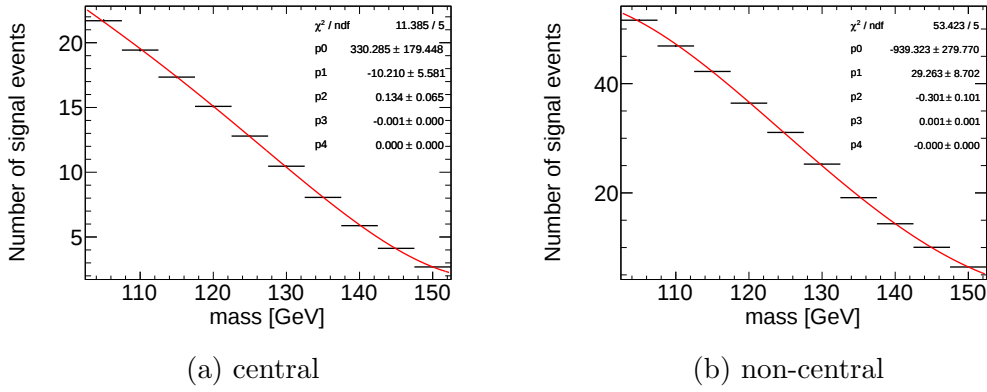


Figure 27: Number of fitted signal events in central and non-central region for each m_H signal sample (black) and fit of second order polynomial (red)

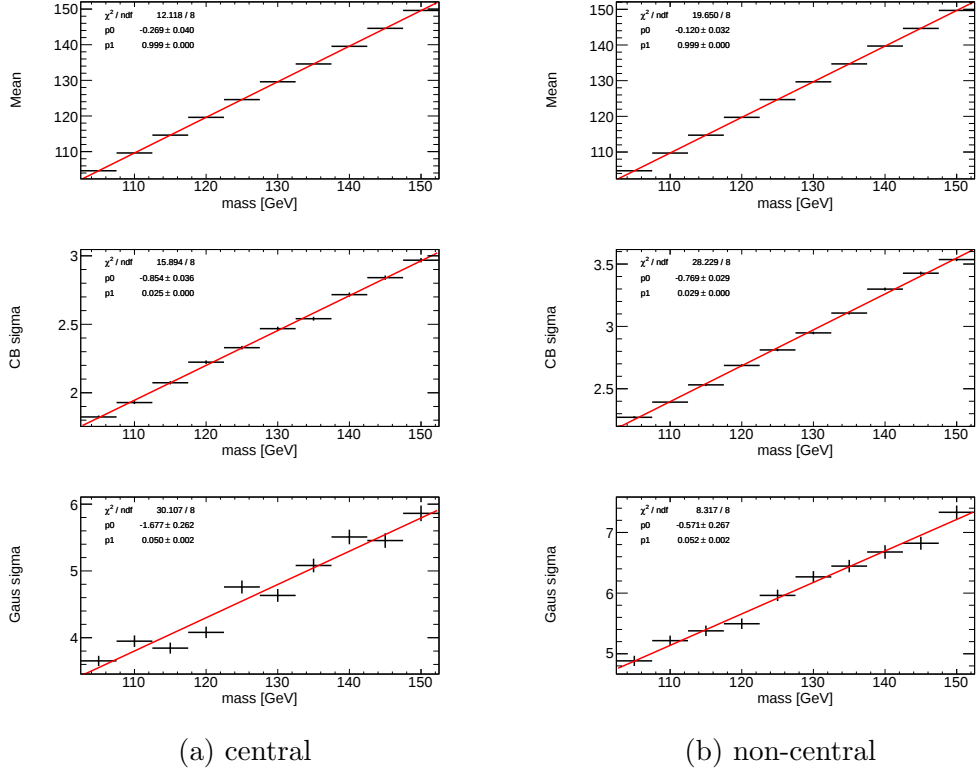


Figure 28: Fit parameters m (top), σ_{CB} (middle) and σ_{GS} (bottom) of the signal model for each signal sample (black) and linear interpolation (red)

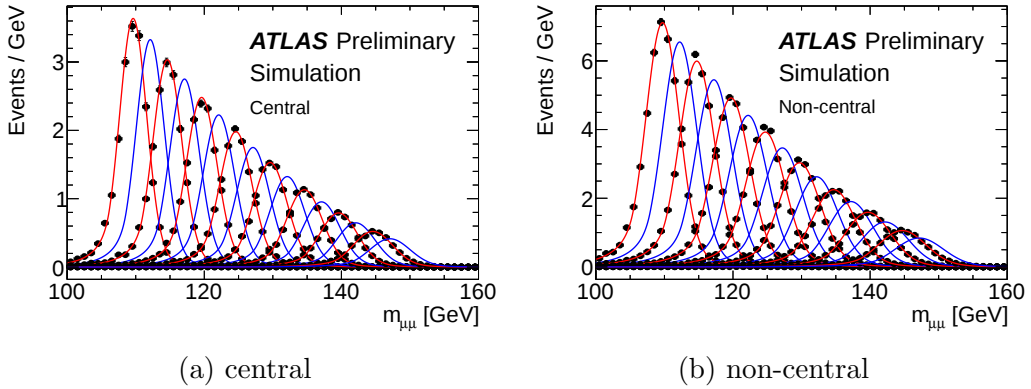


Figure 29: The $m_{\mu\mu}$ distribution for different m_H . Black dots are from Monte Carlo Higgs samples. In red are the fitted signal pdfs as shown in figure 26 for a single sample. The blue curves are the interpolated signal parameterizations shown for selected values of m_H

8.2 Background Modeling

Modeling the background is crucial. A bias in the background parametrization will lead to erroneous results in the end, such as a fake signal where none is actually present. In this analysis, the background pdf has been chosen as a combination of a Breit-Wigner (BW) function together with an exponential pdf. It can be motivated from the main backgrounds, where the BW function reproduces the peaking contributions, mainly from the Z-resonance, and the exponential function emulates the falling, non-peaking other backgrounds such as $t\bar{t}$ and WW.

$$P_{BG} = f_{BW} \cdot BW(x, M_Z, \Gamma_Z) + (1 - f_{BW}) \cdot P(e^{B \cdot x}), \quad (23)$$

where the width Γ_Z and mass M_Z are fixed to their world average measured values [42]. As for the signal pdf, x denotes the running di-muon invariant mass. The parameter B of the exponential function and the f_{BW} fraction are left floating. Before unblinding, this function has been checked against the $m_{\mu\mu}$ spectrum of the data in the low- $p_T^{\mu\mu}$ control region. It has also been checked against Monte Carlo in both the signal region as well as in the low- $p_T^{\mu\mu}$ control region. To assess how well the model describes each distribution, the pulls between model and distribution are examined. The pull of a bin is defined as the difference between model function and value of the histogram, divided by the error of the histogram value. If the model describes the histogram, the pulls should be distributed as a normal distribution with a width of 1 centered around the origin. The distribution of $m_{\mu\mu}$ in MC and fit of the background model, together with their pull distribution, are shown in figure 30. No significant biases in the background description have been observed.

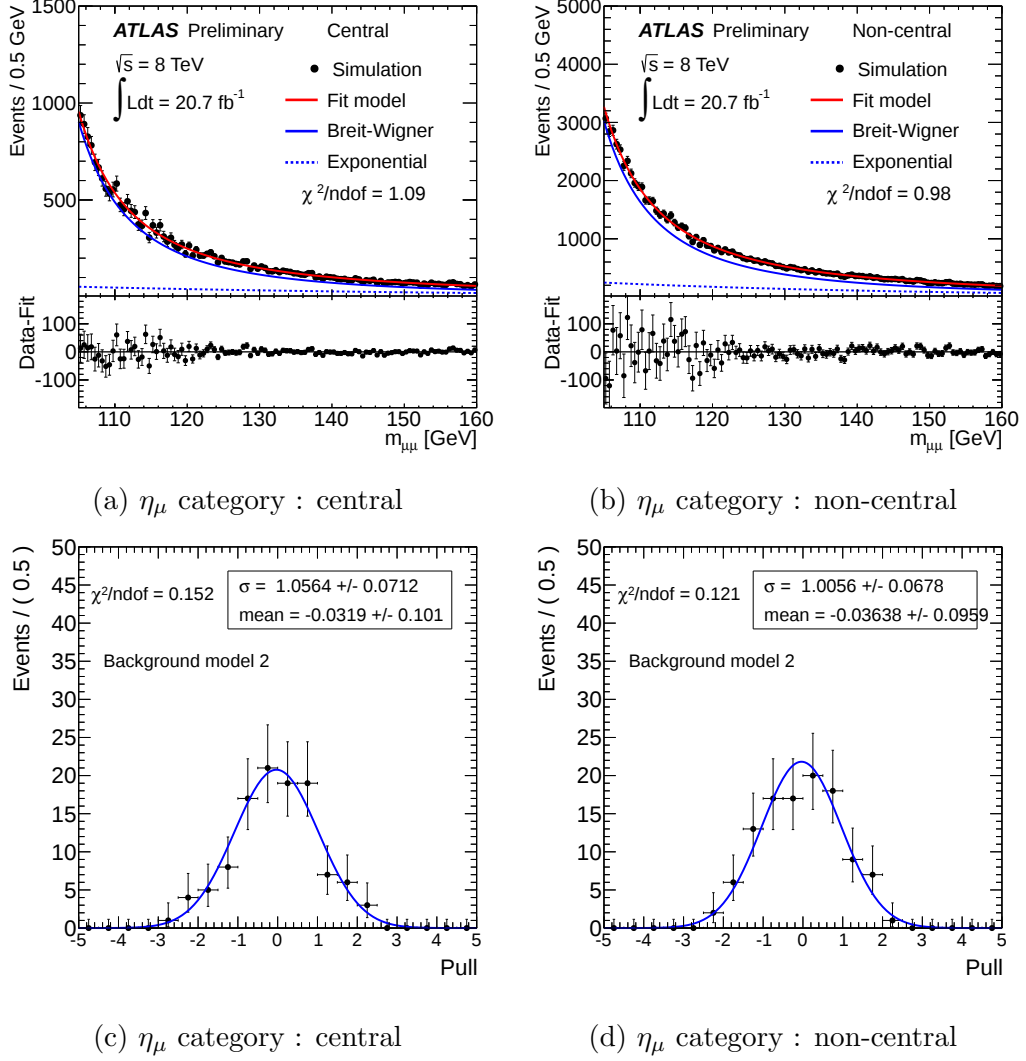


Figure 30: Monte Carlo prediction of invariant mass $m_{\mu\mu}$ spectrum in the central (left) and non-central (right) signal regions and fit of the background model. Beneath each plot the corresponding pull distribution between MC and fit is shown.

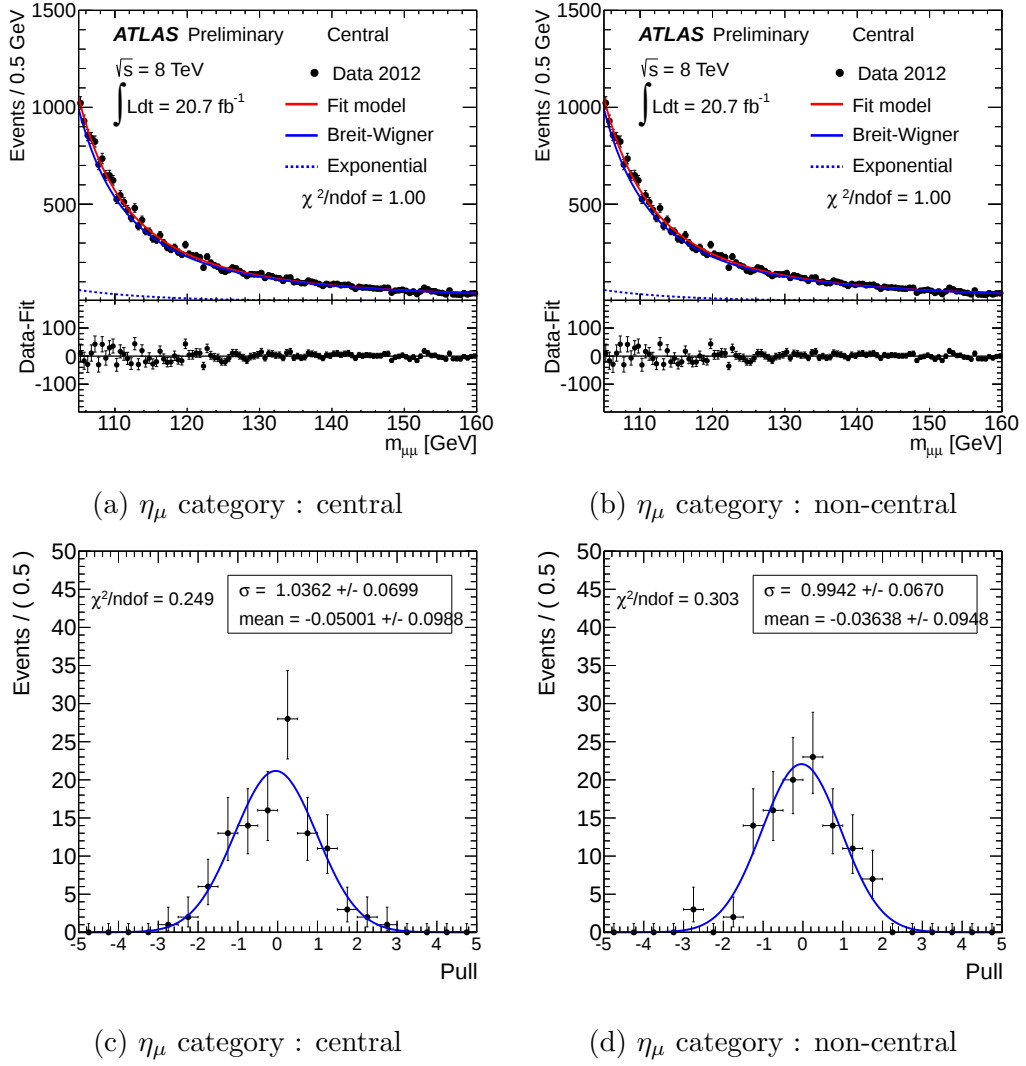


Figure 31: Invariant mass $m_{\mu\mu}$ spectrum in the central (left) and non-central (right) $p_T^{\mu\mu}$ -control regions and fit of the background model. Underneath each plot the corresponding pull distribution between data and fit is shown.

9 Systematic Uncertainties

A number of systematic experimental effects on the signal have been investigated.

9.1 Luminosity Uncertainty

The uncertainty on the integrated luminosity is $\pm 3.6\%$. It is derived, following the same methodology as that detailed in Ref. [43], from a preliminary calibration of the luminosity scale derived from beam-separation scans performed in April 2012.

9.2 Pile-Up Reweighting

To get a better agreement between the mean number of reconstructed primary vertices $\langle \mu \rangle$ in Monte Carlo and in Data, it is rescaled before applying the pileup-reweighting. The systematic effects of the μ rescaling have been estimated by varying the rescale factor from the nominal 0.9 up to 1.0 and down to 0.8. These variations are designated MuRescaleUp and MuRescaleDown respectively.

9.3 Muon-related Uncertainties

To study the systematic errors in the reconstruction of the muons, the analysis is re-run on the signal MC multiple times with muon related parameters changed. Each time a single parameter is changed up or down by one standard deviation. The muon momentum resolution is varied for the ID and MS separately, since both contribute to the uncertainty. The systematic variations for the inner detector are labeled IDUP and IDLOW, for the muon spectrometer MSUP and MSDLOW.

For the trigger scale factors, the event weights have been modified to reflect the uncertainty of the muon trigger scale factors. Those variations are called lepTriggerSFup and lepTriggerSFdown. The variations regarding the muon isolation are called lepIDup and lepIDdown. Figure 32 shows the $m_{\mu\mu}$ distribution for the nominal analysis and for each systematic variation for the example of the $m_H = 125$ GeV signal. There appear to be no significant changes in the shape of the signal. The number of expected signal events changes only very slightly, as seen in table 6.

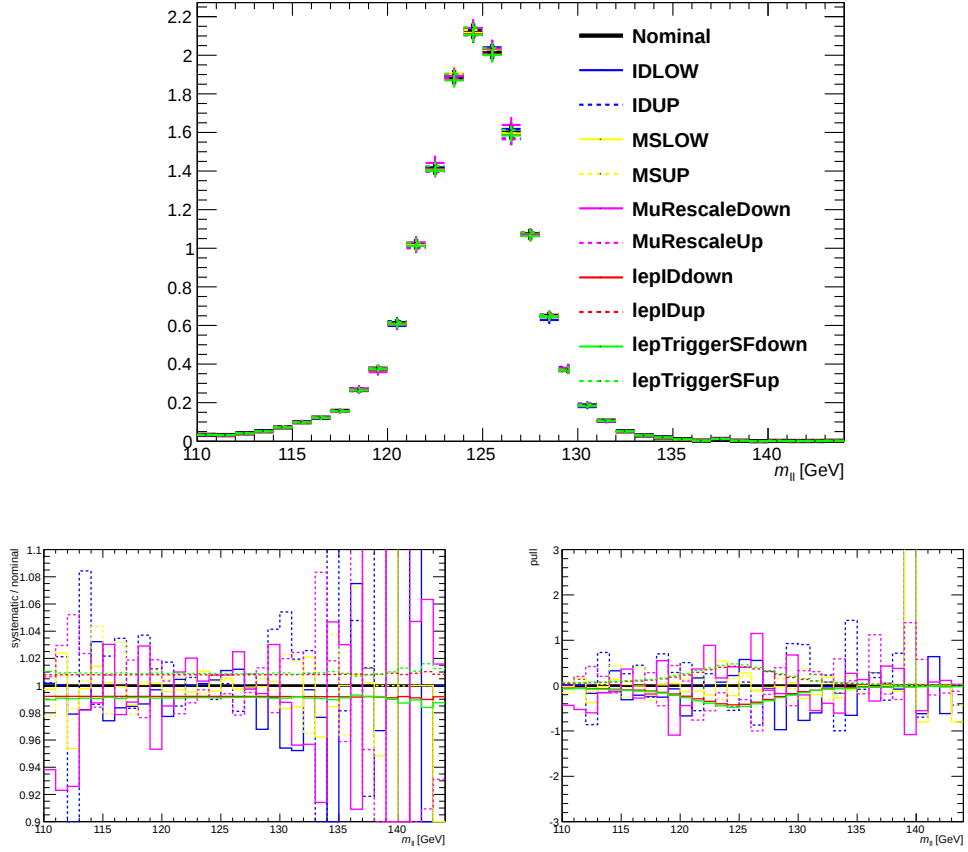


Figure 32: The $m_{\mu\mu}$ distribution of the signal MC, here shown for $m_H = 125$ GeV, created with different systematic experimental variations. Each color represents a different variation. The top image shows the spectra themselves, the bottom left image is the ratio of the variation over the nominal spectrum, and the bottom right image shows the distribution of the pulls, defined as the difference from the nominal distribution over the statistical error per bin.

	Nominal	IDLOW	IDUP	MSLOW	MSUP	MuRescaleDown	MuRescaleUp	lepIDdown	lepDup	lepTriggerSFdown	lepTriggerSFup
mc weights	35.51	+0.00%	-0.01%	-0.00%	-0.00%	-0.33%	+0.41%	+0.81%	-0.82%	+0.83%	-0.83%
OS leptons	35.41	+0.00%	-0.01%	-0.00%	-0.00%	-0.33%	+0.42%	+0.81%	-0.82%	+0.83%	-0.83%
lead lepton $p_T > 25$ GeV	35.41	+0.00%	-0.01%	-0.00%	-0.00%	-0.33%	+0.42%	+0.81%	-0.82%	+0.83%	-0.83%
sublead lepton $p_T > 15$ GeV	35.16	+0.00%	-0.01%	-0.00%	-0.01%	-0.32%	+0.40%	+0.81%	-0.81%	+0.81%	-0.81%
$P_T^{\ell\ell} > 15$ GeV	27.50	+0.04%	-0.01%	-0.00%	+0.01%	-0.27%	+0.43%	+0.81%	-0.81%	+0.83%	-0.83%
$105 < m_{\ell\ell} [\text{GeV}] < 160$ GeV	27.13	+0.05%	-0.00%	+0.01%	+0.00%	-0.29%	+0.42%	+0.81%	-0.81%	+0.83%	-0.83%
Central region	7.97	+0.05%	-0.06%	+0.01%	+0.01%	-0.39%	+0.65%	+0.79%	-0.79%	+0.94%	-0.94%
Non central region	19.16	+0.05%	+0.02%	+0.00%	+0.00%	-0.25%	+0.33%	+0.82%	-0.82%	+0.78%	-0.78%
$ M_H - m_{\ell\ell} \leq 5 \text{ GeV}$	23.30	-0.17%	+0.26%	-0.14%	+0.17%	-0.46%	+0.56%	+0.81%	-0.81%	+0.83%	-0.83%
$P_T^{\ell\ell} < 15$ GeV	7.66	-0.13%	-0.02%	+0.00%	-0.05%	-0.48%	+0.29%	+0.80%	-0.80%	+0.75%	-0.75%

Table 6: The expected numbers of events for a nominal signal with $m_H=125$ GeV after selecting two muons at different stages in the cut-flow and the relative deviation to the nominal signal yield for different systematic variations.

9.4 Production Cross-Section Uncertainty

Theoretical predictions for the production cross-sections have been calculated for the SM Higgs by the LHC cross section working group. They can be found in the CERN Yellow Report 2 [5]. The cross sections of the gluon fusion and vector boson fusion processes, together with their relative theoretical uncertainties, are shown in table 7 for the LHC at a center of mass energy $\sqrt{s} = 8$ TeV. The uncertainty of the parton density function of the gluons is propagated into the theoretical prediction of the gluon fusion process cross-section.

$m_H[\text{GeV}]$	gluon fusion			vector boson fusion		
	$\sigma[\text{pb}]$	uncertainty [%]		$\sigma[\text{pb}]$	uncertainty [%]	
		up	down		up	down
110	25.04	+15.3	-14.9	1.809	+2.7	-3.0
115	22.96	+15.0	-14.9	1.729	+2.7	-3.0
120	21.13	+14.8	-14.8	1.649	+2.8	-3.0
125	19.52	+14.7	-14.7	1.578	+2.8	-3.0
130	18.07	+14.6	-14.6	1.511	+2.8	-2.9
135	16.79	+14.4	-14.7	1.448	+2.8	-2.9
140	15.63	+14.3	-14.5	1.389	+2.7	-2.9
145	14.59	+14.1	-14.4	1.333	+2.8	-2.8
150	13.65	+14.1	-14.4	1.280	+2.8	-2.9

Table 7: SM Higgs production cross sections at 8 TeV

9.5 Uncertainty on the Branching Ratio

The CERN Yellow Report [5] also contains theoretical predictions for the Higgs decay branching ratios. In table 8 the branching ratio to $\mu^+\mu^-$ pairs is shown with their uncertainties.

m_H [GeV]	$\Gamma(H \rightarrow \mu^+ \mu^-)$	uncertainty [%]	
		up	down
110	2.76×10^{-4}	+7.0	-6.8
115	2.63×10^{-4}	+6.7	-6.6
120	2.44×10^{-4}	+6.4	-6.3
125	2.20×10^{-4}	+6.0	-5.9
130	1.90×10^{-4}	+5.5	-5.4
135	1.55×10^{-4}	+5.0	-4.9
140	1.22×10^{-4}	+3.7	-3.8
145	9.06×10^{-5}	+3.4	-3.4
150	6.19×10^{-5}	+3.1	-3.2

Table 8: SM Higgs branching ratios $\Gamma(H \rightarrow \mu^+ \mu^-)$

9.6 Signal Acceptance Uncertainty

The signal acceptance A is defined as the fraction of signal events that pass the analysis cuts.

$$A = \frac{N_{passed}}{N_{total}} \quad (24)$$

To estimate the theoretical uncertainty on A , the ggF signal has been re-simulated with a number of key parameters varied. This simulation uses the same Monte Carlo generator as the nominal analysis, POWHEG for matrix element generation in combination with Pythia8 for showering. Detector effects were not included, so all kinematic variables are considered on generator or “truth” level. The nominal analysis is then replicated as closely as possible using the truth level variables.

9.6.1 Renormalization and Factorization Scales

Since QED calculations are perturbative, they have to be truncated at some point. To account for the uncertainties on higher order contributions, renormalization and factorization scales are used. Due to them being almost 100% correlated, they are treated as a single variation with both being scaled up and down by a certain factor. The renormalization scale is scaled by 0.9 and 1.1, while the factorization scale is scaled by 0.5 and 2.0 at the same time. Figure 33 compares the $m_{\mu\mu}$ and $p_T^{\mu\mu}$ spectrum of the systematic variation to the nominal MC. No shape differences are apparent. Table 33 shows the relative difference in the acceptance.

category	ren/fac scale up	ren/fac scale down
central	0.25 ± 0.48	-2.08 ± 0.47
non-central	-0.03 ± 0.29	0.66 ± 0.29
combined	0.06 ± 0.18	-0.27 ± 0.18

Table 9: Relative acceptance difference from varying the renormalization/factorization scales up and down. All values in %. Statistical errors are given.

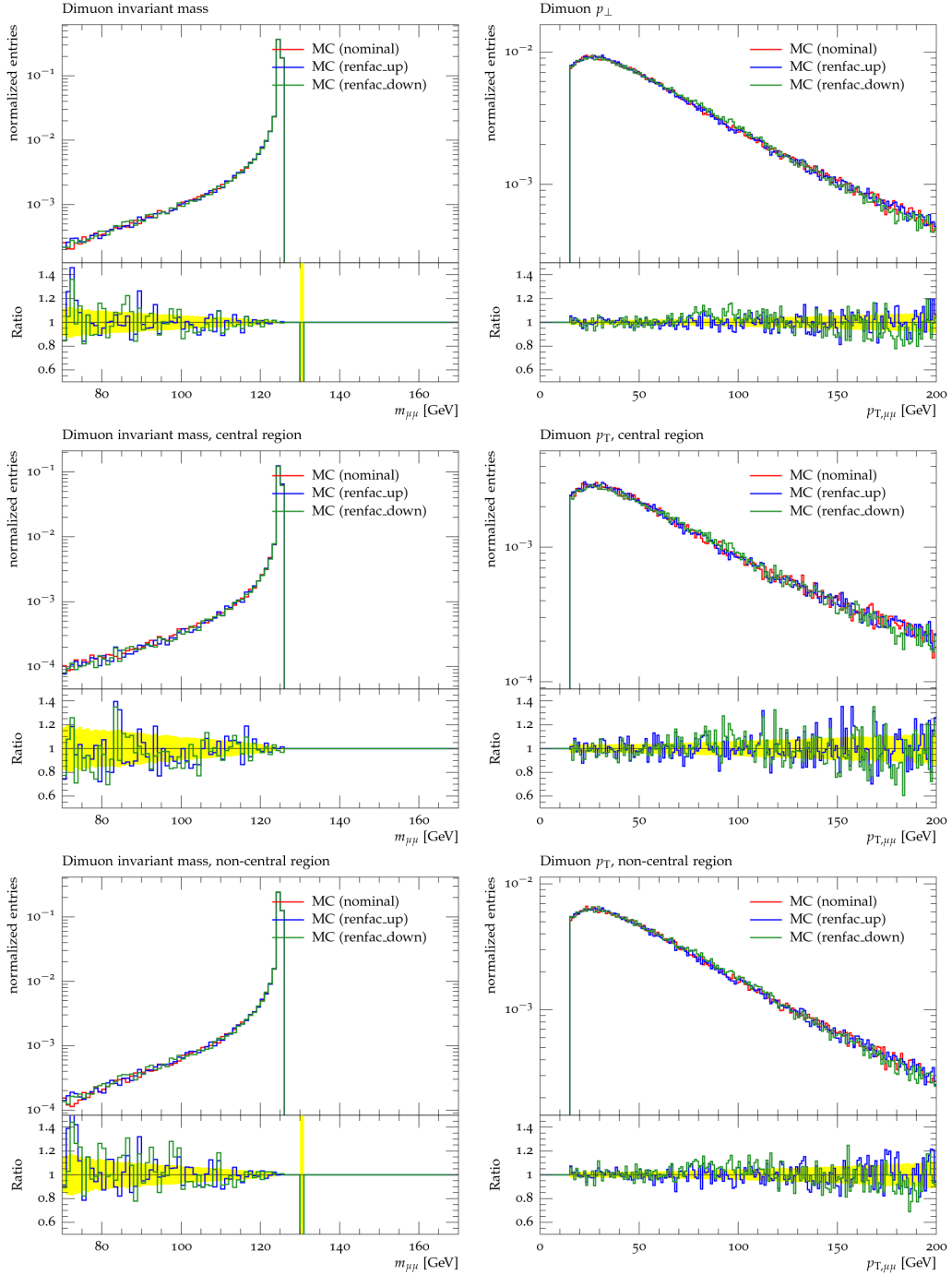


Figure 33: The di-muon invariant mass $m_{\mu\mu}$ (left column) and di-muon transverse momentum $p_{T,\mu\mu}$ (right column) at generator level. The red curve shows the nominal distribution. The blue curve shows the variation of renormalization and factorization scaled up and the green curve down. The top row represents both resolution categories together. The middle row is the central and the bottom row the non-central region. All plots are normalized to unity. The statistical uncertainty on the nominal distribution is shown as the yellow band.

9.6.2 PDF Uncertainties

Further uncertainties from the parton distribution function used are considered. As mentioned in section 5.2.1, the ggF signal was created using the CT10 PDF set. It contains a set of 26 eigenvectors, plus a set of 52 error PDFs, where each eigenvector was varied up and down within its 90 % confidence level. For each of this variations, events were generated and the difference in signal acceptance relative to the nominal PDF calculated. The resulting distribution of A was then fitted with a Gaussian as shown in figure 34 for each category. The width of this fit is taken as the uncertainty in signal acceptance due to PDF uncertainty. Since the widths in the non-central and combined categories are limited by their statistical error, they are not propagated into the final acceptance systematics.

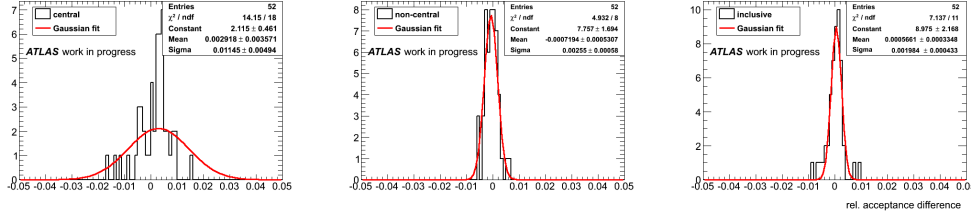


Figure 34: Distribution of the relative change in acceptance in varying each of the 52 error PDF eigenvectors. The distribution is shown for the central (left), non-central (middle) and combined categories (right).

category	$\sigma_{Gaussian}[\%]$
central region	1.15 ± 0.48
non-central region	0.25 ± 0.29
inclusive	0.20 ± 0.18

Table 10: Gaussian width of the distribution of relative difference in acceptance per category due to PDF uncertainties.

9.7 Initial and Final State Radiation

In the Monte Carlo generation, initial (isr) and final state (fsr) radiation is included. The key parameter that controls the amount of this radiation are the coupling strengths of the respective interactions. For QCD interactions this is the strong coupling constant α_S , for electromagnetic interactions the fine structure constant α_{EM} . To estimate the acceptance uncertainty due to initial and final state radiation, α_{EM} is scaled by $\pm 5\%$ and α_S by $\pm 20\%$. Since the algorithms in Pythia for initial and final state radiation are different, each variation is considered separately. Tables 11 and 12 show the relative difference in acceptance for isr and fsr, respectively.

category	isr up	isr down
central	2.47 ± 0.48	-3.35 ± 0.46
non-central	0.75 ± 0.29	-2.01 ± 0.28
combined	1.33 ± 0.18	-2.47 ± 0.18

Table 11: Relative acceptance difference from varying the initial state radiation up and down. All values in %. Statistical errors are given.

category	fsr up	fsr down
central	-0.33 ± 0.47	0.06 ± 0.48
non-central	-0.41 ± 0.29	0.05 ± 0.29
combined	-0.38 ± 0.18	0.05 ± 0.18

Table 12: Relative acceptance difference from varying the final state radiation up and down. All values in %. Statistical errors are given.

Figure 35 shows the distribution of $p_T^{\mu\mu}$ and $m_{\mu\mu}$ for varied isr and fsr. A shape difference in the $p_T^{\mu\mu}$ distribution can be seen in the isr variation. However, no additional shape uncertainty is included because the $p_T^{\mu\mu}$ variable is only used to distinguish signal region from control region by a cut, which uses no shape information.

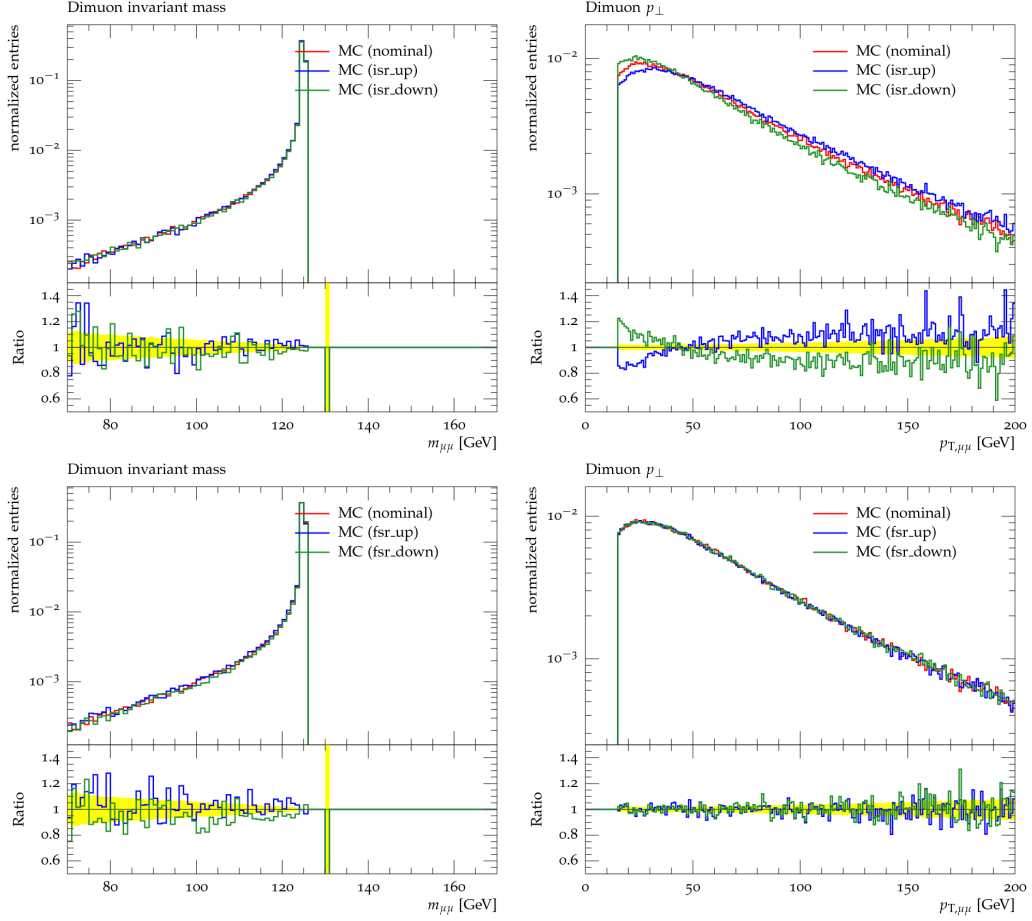


Figure 35: Distributions of $m_{\mu\mu}$ (left column) and $p_T^{\mu\mu}$ (right column). Variations with radiation scaled up are blue and down green. The top row shows variation in initial state radiation and the bottom row variation of final state radiation. The statistical uncertainty on the nominal distribution is shown as the yellow band.

9.8 Summary of Theoretical Acceptance Uncertainties

The individual uncertainties are combined into a total theoretical uncertainty on the acceptance by adding them in quadrature. Uncertainties in upwards and downwards direction are combined separately. Only contributions which are statistically significant, i.e. those which are larger than their statistical error, are considered. Table 13 shows the total uncertainty per category. These uncertainties are included in the final limit calculations as described in section 10.2.

category	downward uncertainty	upward uncertainty
central	-4.11	+2.72
non-central	-2.01	+1.00
combined	-2.47	+1.33

Table 13: Relative acceptance difference from varying the final state radiation up and down. All values in %. Statistical errors are given.

10 Statistics

10.1 Hypothesis Testing

The search for the $H \rightarrow \mu^+ \mu^-$ signal can be formulated as an hypothesis test. Here, the null or background only hypothesis describes the case where only the known SM backgrounds exist, but no signal is present. The alternative or signal+background hypothesis is that both the background and the signal we are looking for exist. The modified frequentist method or CL_S method [44] is used to quantify which hypotheses are favored or excluded.

All hypotheses that are considered in this analysis can be expressed as $\mu s + b$. Here, μ is the so called signal strength parameter. The background only hypothesis corresponds to $\mu = 0$, whereas $\mu = 1$ corresponds to the signal+background hypothesis for a SM Higgs boson decay.

To estimate the level of agreement or disagreement of a certain hypothesis with the measured data, a test statistic $\tilde{q}_\mu$ is introduced. It is a function of data, expected signal and background events, the model parameters, in this case the signal strength μ . It is constructed in such a way that a higher value of $\tilde{q}_\mu$ indicates a higher incompatibility between data and the chosen μ -hypothesis.

The p-value for a given hypothesis is then calculated to quantify this disagreement. It is defined as the probability under the assumption of this hypothesis that the test statistic is as high as or higher than the one observed. This is illustrated in figure 36 (a). To calculate a p-value one therefore integrates the probability distribution for all values of $\tilde{q}_\mu \geq \tilde{q}_{\mu,obs}$

$$p_\mu = \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu, \quad (25)$$

where $\tilde{q}_{\mu,obs}$ is the value of the test statistic observed from data and denotes the probability distribution function of $\tilde{q}_\mu$ under the assumption of signal strength μ . This pdf has traditionally been probed by Monte Carlo pseudo-experiments, where repeatedly Poisson distributions of $m_{\mu\mu}$ according to the hypothesis have to be generated and their corresponding test statistic computed. However, this is computationally expensive. Asymptotic formulae [45] exist for some choices of the test statistic, which describe its probability distribution, given a certain μ .

For this analysis, the following form of log-likelihood ratio is used:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \\ -2 \ln \frac{L(0, \hat{\theta}(0))}{L(\hat{\mu}, \hat{\theta}(\mu))} & \hat{\mu} < 0 \end{cases} \quad (26)$$

The likelihood L is based on the combined fit of the signal model, as described in section 8.1, and background model, as described in section 8.2, to the $m_{\mu\mu}$ spectrum. The specific form of L is discussed further down in section 10.3. In the formula of the test statistic, θ is a collection of parameters which are used in the fit, but are not of interest themselves, the so called nuisance parameters. All systematic uncertainties as described in section 9 are treated as nuisance parameters. $\hat{\mu}$ and $\hat{\theta}$ are the best fit values for μ and θ floating free in the fit. $\hat{\theta}$ is calculated by a fit with a fixed μ , but floating θ .

In particle physics it is common to convert the p-value into an equivalent significance Z . The local significance of a deviation from the null hypothesis is calculated from p_0 , the corresponding p-value. It is defined in such a way that a standard Gaussian distributed variable, found Z standard deviations above its mean, has an upper-tail probability equal to p . It is therefore calculated as

$$Z = \Phi^{-1}(1 - p_0), \quad (27)$$

where Φ^{-1} is the inverse of the cumulative distribution of a standard Gaussian distribution. This relationship is illustrated in figure 36 (b). To claim discovery it is customary to require a significance of at least $Z > 5$.

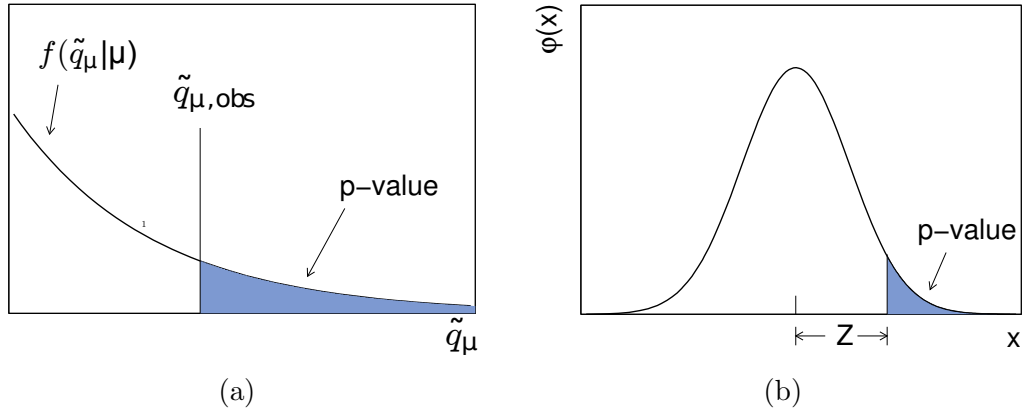


Figure 36: (a) Illustration of the p-value obtained from observed data using a test statistic $\tilde{q}_\mu$. (b) Relationship between p-value and local significance Z as derived from a standard Gaussian distribution $\varphi(x) = (1/\sqrt{2\pi}) \exp(-x^2/2)$. Images adapted from [45].

10.2 Exclusion Limit

The confidence in the signal+background hypothesis ($\mu = 1$) is given by the probability that the test-statistic is less or equal than the observed value $\tilde{q}_{\mu,obs}$.

$$CL_{s+b} = \int_{-\infty}^{\tilde{q}_{\mu,obs}} f(\tilde{q}_{s+b}|s+b)d\tilde{q}_{s+b}, \quad (28)$$

High values of CL_{s+b} indicate a good compatibility with the signal+background hypothesis. The confidence in the background only hypothesis is defined analogously as

$$CL_b = \int_{-\infty}^{\tilde{q}_{\mu,obs}} f(\tilde{q}_b|b)d\tilde{q}_b. \quad (29)$$

Finally, the modified frequentist confidence level is defined as

$$CL_s = \frac{CL_{s+b}}{CL_b}. \quad (30)$$

The signal hypothesis will be considered excluded at the confidence level CL, typically 95%, if

$$1 - CL_s \leq CL. \quad (31)$$

Now, CL_s can be calculated not only for the fixed value of $\mu = 1$, which corresponds to the signal+background hypothesis, but also for arbitrary signal strengths in a $\mu s+b$ model. For a given CL, there will be a largest μ that can be excluded at this confidence level. This is called the upper exclusion limit on the signal strength.

10.3 Likelihood

Since the data is split into two resolution categories, the complete likelihood is written as the product of the category likelihoods

$$L\left(\mu, \boldsymbol{\theta} = \bigcup_{c=1}^{n_{cat}} \boldsymbol{\theta}\right) = \prod_{c=1}^{n_{cat}} L_c(\mu, \boldsymbol{\theta}_c), \quad (32)$$

where c runs over the categories "central" and "non-central", and $\boldsymbol{\theta}_c$ are the nuisance parameters in each categories. The category likelihood itself is then the product of the per-bin likelihoods of the $m_{\mu\mu}$ histograms.

$$L_c(\mu, \boldsymbol{\theta}_c) = \prod_{k=1}^{N_c} \mathcal{L}_c(m_{\mu\mu}(k); \mu, \boldsymbol{\theta}_c) \quad (33)$$

Here k indexes the bin number, whereas $m_{\mu\mu}(k)$ is the value of $m_{\mu\mu}$ in the bin k . The individual bin likelihood is defined as follows:

$$\mathcal{L}_c(m_{\mu\mu}; \mu, \boldsymbol{\theta}_c) = N_{signal,c}(\mu, \boldsymbol{\theta}_c^{norm}) f_{signal,c}(m_{\mu\mu}) + N_{bkg,c} f_{bkg,c}(m_{\mu\mu}; \boldsymbol{\theta}_c^{bkg}) \quad (34)$$

$N_{signal,c}$ and $N_{bkg,c}$ are the numbers of signal and background events in this category, respectively, extracted from the fits. $f_{signal,c}$ and $f_{bkg,c}$ are the per-category probability density functions of the signal and background. $\boldsymbol{\theta}_c^{norm}$ and $\boldsymbol{\theta}_c^{bkg}$ are the nuisance parameters relevant to signal normalization and background parametrization.

11 Results

Figure 37 shows the fit of the background model to the data in the signal regions. There is no visible bias in the description of the background using this model.

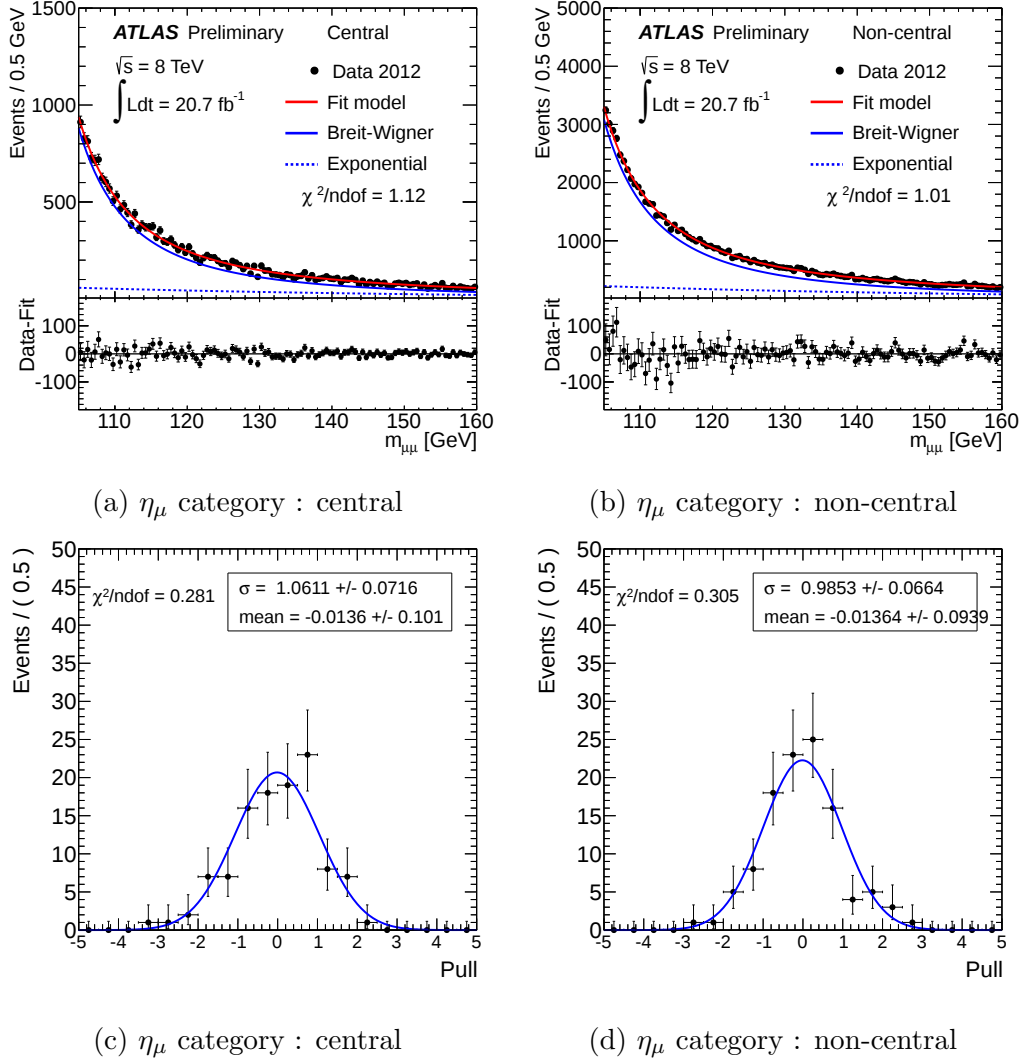


Figure 37: Top row: di-muon invariant mass spectrum in signal categories central (left) and non-central (right). Bottom row: corresponding pull distributions.

Using the procedure described in section 10.2, upper exclusion limits on the signal strength parameter μ have been calculated. For each considered masspoint between 110 GeV and 150 GeV, a signal pdf as described in sec-

tion 8.1 has been constructed. From the simultaneous fit of the signal and background model the limit is then obtained per mass point. The resulting graph in figure 38 shows these limits depending on m_H . Figure 39 shows the p_0 value and corresponding significance. Table 14 summarizes the limits for selected values of m_H . These results have been published in an ATLAS conference note [46].

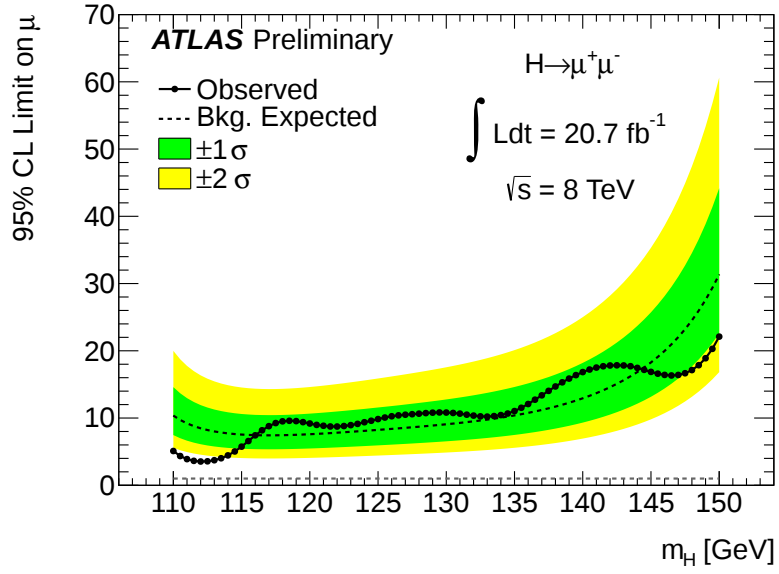


Figure 38: Expected and observed 95% CL_s upper limits on the signal strength as a function of the hypothesized Higgs mass m_H . Green and yellow bands show 1σ and 2σ error bands, respectively.

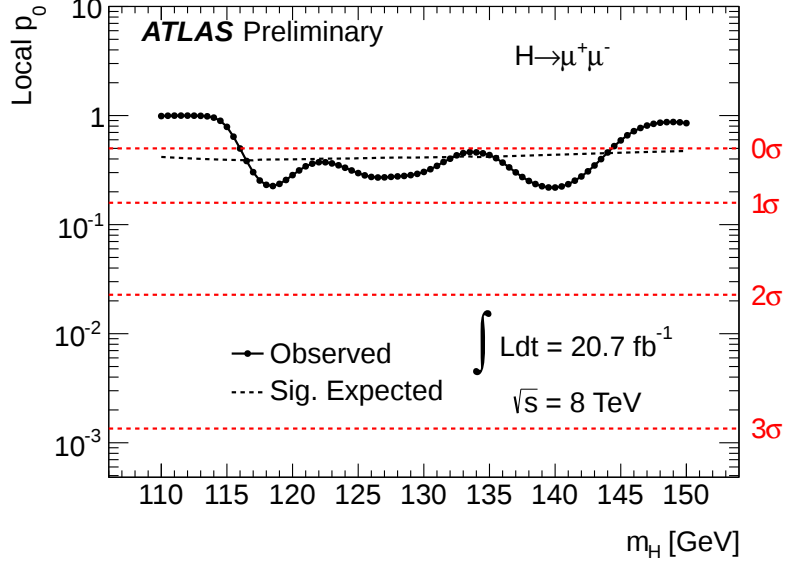


Figure 39: Expected (gray dashed line) and observed (black dots) p_0 value as a function of m_H . The dashed red lines indicate significance Z of corresponding p_0 values.

m_H [GeV]	observed limits	exp. median	exp. +2 σ	exp. +1 σ	exp. -1 σ	exp. -2 σ
110	5.1	10.4	20.0	14.6	7.5	5.6
115	5.7	7.5	14.5	10.6	5.4	4.1
120	9.2	7.6	14.6	10.7	5.5	4.1
125	9.8	8.2	15.9	11.6	5.9	4.4
130	10.8	9.1	17.5	12.8	6.5	4.9
135	11.0	10.4	20.1	14.6	7.5	5.6
140	16.9	12.9	25.0	18.2	9.3	6.9
145	16.9	18.3	35.3	25.7	13.2	9.8
150	22.1	31.4	60.6	44.2	22.6	16.8

Table 14: Observed and expected 95% CL upper limits on Higgs boson production cross section times the $H \rightarrow \mu^+ \mu^-$ branching ratio relative to the SM prediction together with the upper and lower 1 and 2 σ uncertainties for different values of m_H using 20.7 fb $^{-1}$.

12 Conclusions and Outlook

A possible decay of the Higgs boson has been studied. For this, a search for a Higgs resonance in the di-muon spectrum has been carried out in the range from 110 to 150 GeV. No significant excess has been found. The results of the statistical analysis is consistent with Standard Model backgrounds. The current sensitivity in this channel is low for a SM Higgs boson. A non-SM $H \rightarrow \mu^+\mu^-$ signal, which would be about an order of magnitude greater than the SM prediction, would be visible in this analysis. At an assumed Higgs boson mass of $m_H = 125$ GeV exclusion limits of 9.8 times the SM signal strength with an expected limit of 8.2 have been established in this analysis.

Previously, the $H \rightarrow \mu^+\mu^-$ channel has been studied in the context of the Minimal Supersymmetric Standard Model. Searches for neutral MSSM Higgs bosons can be interpreted as generic searches for Higgs like bosons, and their exclusion limits on cross section times branching ratio can be compared to SM predictions. Doing so, Ref [47] gives the previous exclusion limits as 30 from ATLAS and 34 from CMS, derived using the full data taken at center of mass energy $\sqrt{s} = 7$ TeV in the year 2011 by each experiment. Improved exclusion limits help to evaluate different proposed theories and extensions of the Standard Model, such as the MSSM with multiple Higgs bosons and theories with modified Yukawa couplings.

Future opportunities for improvements for this channel exist. This search centered around the signature of the dominating gluon-gluon fusion process, which is two high- p_T muons. Other production channels have additional features which could be exploited. The vector boson fusion process, second in strength to the ggF process, is characterized by two additional jets. They are the remnants of the quarks in the initial state, which radiated the $W^\pm$ or Z bosons that fused into the Higgs boson. These are thus slightly deflected from the beam axis. Therefore, in the VBF channel two jets with very high rapidity are expected, which could be used to suppress the dominant $Z/\gamma^* \rightarrow \mu^+\mu^-$ background. This channel is however severely statistically limited, but has high potential for a high luminosity upgraded LHC.

References

- [1] R. Barate *et al.*, “Search for the standard model Higgs boson at LEP,” *Phys.Lett.*, vol. B565, pp. 61–75, 2003.
- [2] T. A. Collaboration, “Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lh.” *Phys.Lett.* B716 (2012) 1-29, 2012.
- [3] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys.Lett.*, vol. B716, pp. 30–61, 2012.
- [4] “Study of the spin properties of the Higgs-like particle in the $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ channel with 21 fb⁻¹ of $\sqrt{s} = 8$ TeV data collected with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2013-031, CERN, Geneva, Mar 2013.
- [5] S. Dittmaier *et al.*, “Handbook of lh cross sections: 2. differential distributions,” Tech. Rep. arXiv:1201.3084. CERN-2012-002, Geneva, Jan 2012.
- [6] M. Gomez-Bock, M. Mondragon, M. Muhlleitner, M. Spira, and P. Zerwas, “Concepts of Electroweak Symmetry Breaking and Higgs Physics,” pp. 177–238, 2007.
- [7] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model,” *Phys.Rept.*, vol. 459, pp. 1–241, 2008.
- [8] C. Burgess and G. Moore, “The standard model: A primer,” 2007.
- [9] M. Gonzalez-Garcia and Y. Nir, “Neutrino masses and mixing: Evidence and implications,” *Rev.Mod.Phys.*, vol. 75, pp. 345–402, 2003.
- [10] J. Beringer *et al.*, “Review of Particle Physics (RPP),” *Phys.Rev.*, vol. D86, p. 010001, 2012.
- [11] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model,” *Phys.Rept.*, vol. 457, pp. 1–216, 2008.
- [12] G. Guralnik, C. Hagen, and T. Kibble, “Global conservation laws and massless particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

- [13] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.
- [14] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.
- [15] Christiane Lefèvre, “The CERN accelerator complex.” Dec 2008.
- [16] “Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011,” 2011.
- [17] J. Pequeno, “Computer generated image of the whole atlas detector.” Mar 2008.
- [18] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003, 2008.
- [19] J. Pequeno, “Computer generated image of the atlas calorimeter.” Mar 2008.
- [20] J. Pequeno, “Computer generated image of the atlas muons subsystem.” Mar 2008.
- [21] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS, Geneva: CERN, 1997. distribution.
- [22] E. Duckeck, G. *et al.*, “ATLAS computing: Technical design report,” 2005.
- [23] S. Agostinelli *et al.*, “Geant 4, a simulation toolkit,” *Nucl. Instrum. Meth.*, vol. A506, p. 250, 2003.
- [24] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput.Phys.Commun.*, vol. 178, pp. 852–867, 2008.
- [25] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, “ALPGEN, a generator for hard multiparton processes in hadronic collisions,” *JHEP*, vol. 0307, p. 001, 2003.
- [26] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations,” *JHEP*, vol. 0206, p. 029, 2002.
- [27] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, “Gluon-induced W-boson pair production at the LHC,” *JHEP*, vol. 0612, p. 046, 2006.

- [28] B. P. Kersevan and E. Richter-Was, “The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5,” 2004.
- [29] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX,” *JHEP*, vol. 1006, p. 043, 2010.
- [30] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG,” *JHEP*, vol. 0904, p. 002, 2009.
- [31] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG,” *JHEP*, vol. 1002, p. 037, 2010.
- [32] M. Guzzi, P. Nadolsky, E. Berger, H.-L. Lai, F. Olness, *et al.*, “CT10 parton distributions and other developments in the global QCD analysis,” 2011.
- [33] “Summary of atlas pythia 8 tunes,” Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug 2012.
- [34] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables,” vol. D78, p. 013004, 2008.
- [35] “Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010,” 2011.
- [36] M. Cacciari, G. P. Salam, and G. Soyez, “Anti-KT jet clustering Algorithm,” *JHEP*, vol. 0804, p. 063, 2008.
- [37] B. Resende, “Muon identification algorithms in ATLAS,” *PoS*, vol. EPS-HEP2009, p. 431, 2009.
- [38] A. Roodman, “Blind analysis in particle physics,” *eConf*, vol. C030908, p. TUIT001, 2003.
- [39] “Plots of muon performance in 2012 data,” Tech. Rep. ATLAS-CONF-2012-716, CERN, Geneva, June 2012.
- [40] H→WW working group, “Background estimates in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis with 13.0 fb^{-1} of data collected with the ATLAS detector at $\sqrt{s} = 8 \text{ TeV}$,” *ATL-COM-PHYS-2012-1516*, p. .

- [41] “Measurements of the properties of the higgs-like boson in the $ww^{(*)} \rightarrow \ell\nu\ell\nu$ decay channel with the atlas detector using 25 fb^{-1} of proton-proton collision data,” Tech. Rep. ATLAS-CONF-2013-030, CERN, Geneva, Mar 2013.
- [42] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, “Precision Electroweak Measurements on the Z Resonance,” *Phys. Rept.*, vol. 427, p. 257, 2006.
- [43] G. Aad *et al.*, “Improved luminosity determination in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ using the ATLAS detector at the LHC,” 2013.
- [44] A. L. Read, “Presentation of search results: the cls technique,” *J. Phys. G*, vol. 28, p. 2693, 2002.
- [45] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” 2010.
- [46] “Search for a Standard Model Higgs boson in $H \rightarrow \mu\mu$ decays with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2013-010, CERN, Geneva, Mar 2013.
- [47] A. Dery, A. Efrati, Y. Hochberg, and Y. Nir, “What if $BR(h \rightarrow \mu\mu)/BR(h \rightarrow \tau\tau)$ does not equal m_μ^2/m_τ^2 ?,” 2013.

Danksagung

Ich möchte mich bedanken bei

- Prof. Dr. Dorothee Schaile für die Möglichkeit dieser Masterarbeit an ihrem Lehrstuhl und für ihre Ratschläge und Anregungen
- Dr. Johannes Elmsheuser für seine stets engagierte Betreuung
- Meinen Bürokollegen Bonnie Chow und Christian Meineck
- Den Spezis Philipp Heimpel and Florian Rättich
- Allen Mitspielern beim täglichen Tischtennis
- Und allen anderen Mitgliedern des Lehrstuhls

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den 27. März 2013

Friedrich G. Hönig