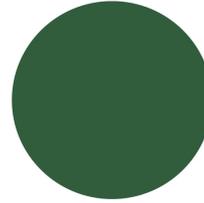


DET NATURVIDENSKABELIGE  
FAKULTET

NIELS BOHR INSTITUTTET

KØBENHAVNS UNIVERSITET



---

# Probing the nature of the Higgs Boson

A study of the Higgs spin and parity through  
the  $ZZ^* \rightarrow 4l$  final state at the ATLAS Experiment

Lars Egholm Pedersen

This thesis has been submitted to the Ph.D. school of the faculty of  
science at the University of Copenhagen

# About this thesis

The work presented this thesis aim to summarise the research I have been involved in, during my PhD studies from September 2012 to August 2015 at the Niels Bohr Institute of the University of Copenhagen. The research has been in association with the ATLAS experiment at CERN. The thesis was written in spring and summer of 2015. The typesetting and layout has been made in L<sup>A</sup>T<sub>E</sub>X.

*Principal Supervisor:*

Troels C. Petersen

*Dissertation committee:*

Stefania Xella

Pierre Savard

Yosef Nir

*Lars Egholm Pedersen*

Signed:

---

Date:

---

## Abstract

A series of studies are presented of the Higgs-like resonance that was discovered in July 2012. The studies focus on measurements of the resonance spin and parity using the full dataset collected by the ATLAS experiment during the LHC Run-1.

Alternative spin models are compared to the Standard Model Higgs. Spin-2 models are described as graviton-like particles in an effective field theory approach. Several assumptions on the boson coupling to the QCD sector are tested, as is the analysis dependency on transverse momenta cutoffs. The Standard Model is favoured over all alternative models and except for the negative parity spin-2 and a spin-2 boson that does not couple to the quark sector, each are excluded with 95% confidence level. Notably, the spin-2 universal couplings model is assigned a  $CL_S$  value of 0.0009. Spin-1 models are tested in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel and excluded in favour of the Standard Model. The resonance decay to  $\gamma\gamma$  also disfavours the possibility of spin-1.

The spin-0 boson coupling to the heavy vector bosons is described in an effective field theory approach with three components, the Standard Model and a CP-even BSM and CP-odd sector. Assuming only one coupling to be present at a time, the Standard Model ( $J^P = 0^+$ ) is tested against the CP-even BSM ( $J^P = 0_h^+$ ) and CP-odd ( $J^P = 0^-$ ) model. The Standard Model is favoured in data and the two alternate models are assigned  $CL_S$  values of 0.10 (CP-even) and 0.054 (CP-odd).

The analysis is extended by allowing a mixture of the three spin-0 components, first assuming only one BSM coupling to be present at a time. Discriminative power between models is obtained in a Matrix-Element Observable approach. The  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel is combined with the  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  analysis. The effective BSM cross section contributions to the resonance is found to be limited to the ranges:

$$f_{g2} \cdot \cos \phi_{g2} \in [-0.16, 0.12] \quad f_{g4} \cdot \cos \phi_{g4} \in [-0.41, 0.090]$$

at 95% confidence level. The data is thus found to be in agreement with the Standard Model. In a final extension of the analysis the possibility of the simultaneous presence of both BSM sectors is studied.

## Resumé

En serie af studier vil blive præsenteret af den Higgs-lignende resonans, hvis eksistens blev påvist i juli 2012. Disse vil bruge det samlede datasæt der blev indsamlet af ATLAS detektoren under LHC Run-I til at lave målinger af resonansens spin og paritet. Alternative spin modeller vil blive sammenlignet med Standard Model Higgs partiklen. Spin-2 modeller er beskrevet som graviton lignende partikler i en effektiv felt teori. Adskillige antagelser omkring dens koblinger til QCD sektoren vil blive testet, ligeledes vil analysens sensitivitet til transverse impuls begrænsninger blive testet. Standard Modellen er fundet til at være den fortrukne teori over alle alternative modeller og med undtagelse af den negative spin-2 hypotese er disse udelukket med 95% konfidensniveau. Den universelle spin-2 koblingsmodel er specielt tilskrevet en  $CL_S$  værdi på 0.0009. Spin-1 modeller er testet i  $H \rightarrow ZZ^* \rightarrow 4\ell$  henfaldet og ekskluderet til fordel for Standard Modellen. Resonansens henfald til  $\gamma\gamma$  udelukker også muligheden for spin-1 som en følge af Landau-Yang teoremet.

Spin-0 bosonens kobling til massive vektor bosoner er beskrevet i en effektiv felt teori med tre komponenter, Standard Modellen og en CP-lige BSM og CP-ulige sektor. Under antagelse af at en kobling er tilstede ad gangen vil Standard Modellen ( $J^P = 0^+$ ) blive testet mod den CP-lige alternative model ( $J^P = 0_h^+$ ) og den CP-ulige model ( $J^P = 0^-$ ). Standard Modellen er fortrukket i data og de to alternative modeller er tilskrevet  $CL_S$  værdier på 0.10 (CP-lige) og 0.054 (CP-ulige).

Analysen vil blive udvidet ved at tillade en blanding af de tre spin-0 komponenter. I første omgang vil kun en BSM kobling blive inkluderet ad gangen. Ved at bruge observable baseret på matrix elementer vil der blive skabt separation mellem modeller.  $H \rightarrow ZZ^* \rightarrow 4\ell$  henfaldet vil blive kombineret med  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  analysen. Det effektive bidrag fra BSM tværsnit til resonansen er fundet til at være begrænset til intervallerne:

$$f_{g2} \cdot \cos \phi_{g2} \in [-0.16, 0.12] \quad f_{g4} \cdot \cos \phi_{g4} \in [-0.41, 0.090]$$

med 95% statistisk sikkerhed. Data er derfor konkluderet til at være i overensstemmelse med Standard Model forudsigelser. I en sidste udvidelse af analysen vil muligheden for den samtidige eksistens af begge BSM sektor blive undersøgt.

# Contents

<b>About this thesis</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>viii</b>
<b>Introduction</b>	<b>ix</b>
<b>Thesis Outline</b>	<b>xi</b>
<b>1 Theory</b>	<b>1</b>
1.1 The Standard Model	6
1.1.1 The Standard Model Lagrangian	6
1.1.2 Amplitude Calculations	9
1.1.3 General Considerations	11
1.1.4 Proton Collision Kinematics	15
1.2 The Higgs Mechanism	19
1.2.1 Spontaneous Symmetry Breaking	19
1.2.2 Intermezzo: Predictions, Discovery and Property Measurements	21
1.2.3 Higgs Characterisation with Effective Field Theories	27
1.2.4 Spin-0 Models	28
1.2.5 Spin-1 and Spin-2 Models	32
1.3 Monte Carlo Simulation and Validation	35
1.3.1 Matrix Element Based Reweighting	36
1.4 Beyond the Standard Model	40
<b>2 Experiment</b>	<b>43</b>
2.1 The Large Hadron Collider	44
2.2 The Atlas Experiment	49
2.2.1 The ATLAS Coordinate Frame	49
2.2.2 The ATLAS Inner Detector	51
2.2.3 Calorimeter System	55
2.2.4 The Muon Spectrometer	57
2.2.5 Trigger Systems	60
2.3 Primary Vertex Reconstruction in Run-2	61

---

2.4	Lepton Identification for the $H \rightarrow ZZ^* \rightarrow 4\ell$ Analysis	64
<b>3</b>	<b>Analysis</b>	<b>69</b>
3.1	Introduction	72
3.2	The $H \rightarrow ZZ^* \rightarrow 4\ell$ Signal Region	73
3.2.1	Event Selection	73
3.2.2	Reducible backgrounds	74
3.2.3	Expectations	76
3.2.4	Systematic Uncertainties	77
3.3	Fixed Hypothesis Tests	81
3.3.1	Distinguishing Models with Boosted Decision Trees	81
3.3.2	PDF Construction	84
3.3.3	Statistical Treatment	85
3.3.4	Systematic Uncertainties	86
3.3.5	Summary of the BDT strategy	88
3.3.6	Results	88
3.3.7	Previous iterations of the $H \rightarrow ZZ^* \rightarrow 4\ell$ Fixed Hypothesis Tests	90
3.4	The Tensor Structure Fit	93
3.4.1	Coupling Ratios and Final State Observable Variations	93
3.4.2	The Matrix Element Observable Fit	94
3.4.3	PDF Construction	96
3.4.4	Statistical Procedure	104
3.4.5	Asymptotic Test Statistic Behaviour	107
3.4.6	Closure Tests	109
3.4.7	Systematic Uncertainties	111
3.4.8	Compatibility Studies, The Nine Dimensional Fit	117
3.4.9	Summary of the Tensor Structure Fit	119
3.4.10	Results	120
3.5	The $ZZ^*$ and $WW^*$ Combination	122
3.5.1	The $WW^*$ Final State	122
3.5.2	Correlating Systematic Uncertainties	123
3.5.3	Results	124
3.6	Simultaneous Fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	127
3.6.1	Results	130
3.7	Outlook: High Luminosity Prospects	134
3.7.1	Matrix Element Observables as a Probe of the $HZZ$ Vertex	134
3.7.2	Angular asymmetries as a Probe of the $HZZ$ Vertex	139
<b>4</b>	<b>Conclusions</b>	<b>143</b>
4.1	Summary	144
4.2	Concluding Remarks	146
<b>Appendix</b>		<b>147</b>
A	Statistical Considerations	148

---

A.1	Construction and Interpretation of Statistical Models	148
A.2	Fast Binned Kernel Density Estimators	155
A.3	PDF Transformations	160
A.4	Comment on MVAs	163
B	Fixed Hypothesis Test Appendix	165
C	Tensor Structure Fit Appendix	166
D	Comments on Correlation Scheme	170
<b>Abbreviations</b>		<b>187</b>
<b>Symbols</b>		<b>188</b>

# Acknowledgements

I would like to start this work by acknowledging its dependence on the many other people that have helped it come to life.

First I would like to say thank you to the high energy physics group at the Niels Bohr Institute. For providing fruitful discussion, for providing an excellent work environment and not the least for a social setting in which doing science is a pleasure. Special mentions goes to Almut for help with BibTex.

Next I would like to acknowledge the people of the ATLAS Higgs Spin/CP group. It has been some fruitful years and the interaction with you has in every way made this work possible. I would like to thank you Kirill for the work we have done together. You have been exemplary as an external supervisor to give insight and to put me in the right spot at the right time.

Above all would like to thank you Troels as my supervisor. For excellent counselling and giving me enough freedom to progress in my own way and not the least for taking me as a student. It has been three good years.

# Introduction

The summer of 2012 proved to be the point in time, where the Higgs boson would finally be discovered. The discovery both marked the culmination of decades of research and the onset of dedicated studies of its properties. The question changed from: *'Does it exist?'* to *'Is it the one predicted by the Standard Model?'*

The Higgs mechanism serves the Standard Model by enabling it to be a gauge invariant theory of massive particles. Beside conserving gauge invariance, it furthermore resolves problems of massive vector boson scattering that, without spontaneous symmetry breaking, would leave the perturbative regime at the TeV scale.

The properties of the Standard Model Higgs boson has proved to make it exceedingly difficult to discover. The Higgs introduction to the Standard Model does not offer a direct prediction of its mass and only by assuming the absence of new physics can the mass indirectly be predicted to be  $\sim 100$  GeV.

The LEP experiments had ruled out the lower end of this spectrum prior to the LHC era. In 2010 the scene was set for searches in the broad energy range that is possible with proton collisions produced by the LHC.

Due to the low production probability of the Higgs boson and its preference to decay to final state topologies that are indistinguishable from other mechanisms in proton collisions, the rare but clean  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  proved to be the most valuable for discovery.

The 2012 discovery thus demonstrated the existence of a resonance with production probability and decay to the electroweak vector bosons in agreement with predictions. Beside the large uncertainty in these properties, a multitude of different parameters were still to be constrained in order to conclude that the Higgs-like boson is indeed *The* Higgs boson.

Predictions of its coupling to the other massive particles are available to be tested. The Higgs self-coupling is fully determined in the Standard Model and provides an interesting but difficult probe. The Higgs boson is furthermore predicted to be spin-0 and CP-even, which will be the focus of the work presented in this thesis.

The analysis structure is outlined as follows: The Higgs sector of the Standard Model Lagrangian is removed ( $SM - H$ ) and generic boson models ( $J$ ) described with effective field theories are added in its place:

$$\mathcal{L}_{HC,J} = \mathcal{L}_{SM-H} + \mathcal{L}_J$$

Several scenarios will be tested for the  $J$  substitute. First, the Standard Model Higgs is tested against alternative spin-0, spin-1 and spin-2 models. The alternative spin-0 models consist of a CP-even BSM Higgs and a CP-odd pseudo scalar.

Following this, the spin-0 sector is examined by allowing mixtures between the three Lagrangian parts that constitute the Standard Model and the two spin-0 BSM terms. Deviations from the Standard Model are parameterised in BSM to SM coupling ratios of the three spin-0

---

terms:  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  (CP-Even) and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  (CP-Odd).

The analysis mainly focuses on the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channel. This specific channel has the advantage over  $H \rightarrow \gamma\gamma$  that the polarisation of the intermediate  $Z$  bosons affects the angular distributions of the final state leptons. It also has a much higher signal to background ratio. The  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel distinguishes itself over the  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  by having its final state fully reconstructable with the absence of neutrinos. This both give a cleaner probe of the massive vector boson polarisation and a higher signal to background ratio. The requirement of the two  $Z$  bosons to simultaneous decay to either electrons or muons make the probability of this decay type exceedingly rare on the other hand.

The kinematics of the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay can in general be described with eight parameters for a given Higgs mass. The high dimensionality of the final state kinematics offers considerable challenges when describing the decay through its interaction with the detector. Substantial focus of the analysis will be given to how multivariate techniques are adopted to give valid predictions.

# Thesis Outline

The thesis is divided into the three chapters: “Theory”, “Experiment” and “Analysis”. Being a work of physics measurements the focus will be on the last part.

The Theory section will start out with an overview of the Standard Model of particle physics and discuss some of its properties and challenges. Following this, the Higgs mechanism will be motivated and introduced. Its predicted properties and its discovery will be described.

The theoretical framework that is used for the analysis will be described after the general introduction. Here, an effective field theory approach is used to describe generic spin-0, spin-1 and spin-2 bosons. Finally, some “Beyond the Standard Model” physics scenarios will be mentioned.

The Experiment section describes the setup that is used to probe the Standard Model predictions. First, the design features of the Large Hadron Collider (LHC) are motivated. After the LHC introduction, the ATLAS detector will be described, where the individual sub-detectors and their properties are outlined. An algorithm for primary vertex reconstruction in ATLAS to be used in the next years of data taking is moreover described.

Finally, details on electron and muon reconstruction and identification are given. These will be the main objects used in the analysis.

The analysis aims to present studies of how knowledge of the Spin/CP nature of the Higgs boson is obtained. The chapter starts with an overview of the event selection that is imposed on the 2011 and 2012 datasets. A description of the hypothesis tests to determine if data favours the SM or some fixed alternative follows the event selection.

The next section describes the analysis of the BSM to SM coupling ratios  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The first part will follow the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. A short overview of the dedicated  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  analysis is given and its combination with the  $H \rightarrow ZZ^* \rightarrow 4\ell$  results.

The last part presents the results of a simultaneous fit of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The results are interpreted as limiting the effective BSM contribution to the combined boson cross section. Finally, an outlook will be given. The prospects of measuring the  $HVV$  tensor structure are presented for both a  $300 \text{ fb}^{-1}$  and a  $3000 \text{ fb}^{-1}$  dataset. The first study shows a dedicated ATLAS analysis and the second is given in a more general setting.

A list of symbols and abbreviations that are commonly used throughout the thesis can be found on the last two pages.

This is a thesis describing work done in a large collaboration, meaning my own contributions will naturally intertwine with the work of others. I have as much as possible tried to make the analysis text reflect my own contributions.

However, as anybody who writes a Ph.D thesis on experimental physics measurements will find: A surprisingly large amount of your time is spent trying to figure out, why your histograms look slightly different from what you think they should look like. Some considerations have been given to the readability and content of the text.

---

To clarify my own involvement, I have included a time-line of, where I have spent my time the last three years:

*Fixed hypothesis tests for the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis:*

The very first project I got involved in. During the fall of 2012 until the summer of 2013 I focused on the exclusion of alternative spin and parity models in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channel. The analysis contributed to the results that were presented at the HCP-2012[1] and Moriond-2013[2] conferences, and finally led to the “*Evidence for the spin-0 nature of the Higgs boson using ATLAS data*” publication[3].

Up until the HCP conference I was mainly involved in the preparation of discriminant distribution and construction of analysis workspaces. In the beginning of 2013 I updated the analysis framework and further optimised the multivariate separation to alternative models. Following the publication I created the first implementation of the two-dimensional approach that was subsequently used.

*HVV tensor structure analysis:*

During the summer of 2013 I got involved in the studies of the *HVV* tensor structure. The first analysis that was done was a prospects studies presented at the ECFA-2013 conference[4]. I was involved in the section called the “Matrix Element observable fit” and creating a fast simulation approximation of the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis.

Following these initial studies, the work turned to the Run-I analysis. As will be described, further development was done and a slight change of models are used. Being an actual data analysis instead of a prospect studies, the level of scrutiny is naturally also much higher. I was largely responsible for the development and running of the “Matrix Element observable fit” approach from the fall of 2013 to its completion in early 2015.

Following the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis, I was involved in the combination of the *HZZ* and *HWW* tensor structure analyses, which were finally published in the paper: “*Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*”[5].

As a side project to the main analysis, I also participated in a prospects study of the feasibility of measuring the Higgs CP-nature using angular asymmetries[6]. The study will summarily be described in the outlook section.

*Primary vertex reconstruction:*

From the beginning of 2013 to the spring of 2015 I was beside the Higgs analysis also a contributor to the ATLAS primary vertex reconstruction group. The work here mainly focused on the implementation and testing of a seeding algorithm that is more robust, when faced with the higher pile-up conditions that is expected in LHC Run-II. For this reason the vertex reconstruction algorithm is also described in more detail in the “Experiment” chapter than other parts of the ATLAS reconstruction that are of equal importance.

---

Finally, a couple of comments on the text: Natural units ( $c = \hbar = 1$ ) are assumed throughout the thesis. When describing the ATLAS detector and the LHC, SI units are sometimes used but the shift should be apparent from context.

When describing distances the metric tensor is always assumed to have positive time component and negative spatial components.

The  $H \rightarrow ZZ^* \rightarrow 4e, 4\mu, 2e2\mu$  decays are collectively named “Higgs to four lepton”. In this context only electrons and muons are referred to as leptons although the tau should also fall into this category. Taus are experimentally very different from muons and electrons and therefore are not included in the analysis. Throughout the analysis section, it can be assumed that “lepton” ( $\ell$ ) solely is a synonym for electrons ( $e$ ) and muons ( $\mu$ ).



# Chapter 1

## Theory

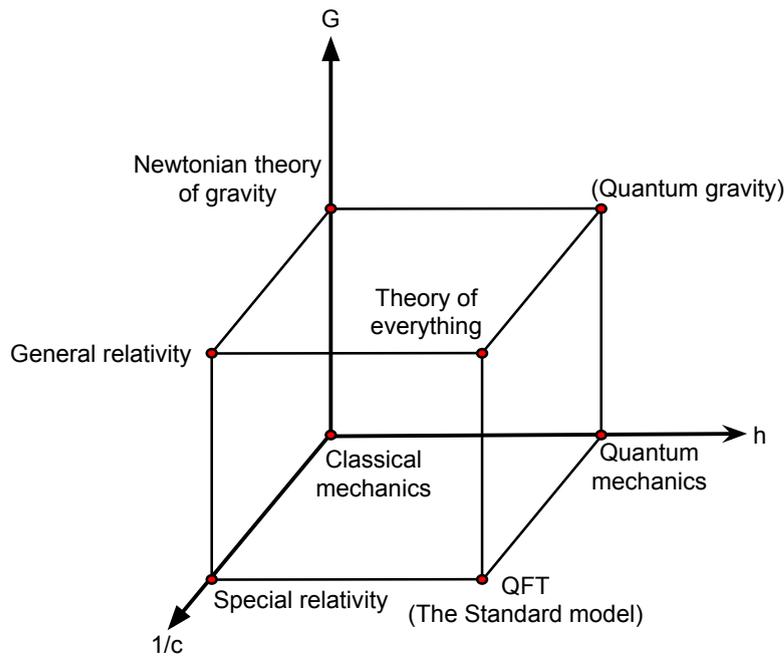


FIGURE 1.1: The Physics Cube. The physical models describing fundamental interactions are here divided into three regimes: Models describing physics in the presence of large bodies represented by the gravitational constant  $G$ . The description of physics at relativistic speeds, symbolised here by the speed of light  $c$  and models describing microscopic interactions represented by Planck's constant  $h$ . It is the grand aim to find a common description of all three areas.

The general theme of this work will be to probe possible deviations from the predicted behaviour of the Higgs boson. In this section the theoretical models that will be examined are motivated and described. First, a very general introduction to the world of particle physics is given.

The Standard Model is subsequently described with its particles and interactions, before the Higgs mechanism is finally introduced. In order to examine if the observed particle is indeed the Standard Model Higgs boson, a framework for characterising deviations is described in the same section.

Figure 1.1 summarises the expanse that physics of fundamental interactions considers. Three regimes are here represented by three constants of nature; the gravitational constant  $G$ , the speed of light  $1/c^1$  and Planck's constant  $h$ . It is possible to imagine a world, where each constant is zero. This would be the classical world of Galilean transformations and Newton's three laws of motion. Imagining one constant present at a time, three different physics models are obtained. Historically the first of these to be described is physics including gravitational attraction ( $G \neq 0$ ). The classical theory of gravitation is found in this corner. The two other corners describe a world where either the speed of light ( $c$ ) or Planck's constant ( $h$ ) are present, described with special relativity and quantum mechanics respectively.

One major goal in the development of physics theories is the attempt to create models that

<sup>1</sup> $\frac{1}{c}$  is used since Newtonian physics is equivalent to the speed of light being infinite rather than 0

combine the three fields. Two successful combinations have so far been made. First the combination of special relativity in the presence of gravity results in the general theory of relativity. By combining special relativity and quantum mechanics on the other hand, quantum field theories (QFT) emerge. Dissimilar to general relativity, QFTs are much more ambiguous and can be thought of more as a mathematical framework in which theories can be described. Unfortunately it can be demonstrated that it is not possible to directly describe gravity in this way. The dominant QFT is the Standard Model of particle physics, which this thesis is a probe of.

The last combination of two constants would give a description of quantum gravity. There are technical challenges that makes this combination unappealing and no successful description has so far been made. It is the impression of most that a model that is able to describe both quantum mechanics and gravity would necessarily also include special relativity.

One of the major problems of this corner is the weakness of gravity. Experimental probes of gravity at the quantum scale would e.g. have to take into account that the electromagnetic forces between two particles is many orders of magnitude larger than the gravitational forces. As an example, assuming classical gravitational attraction, two electrons would repel each other  $\sim 10^{39}$  times stronger due to their charges than they would attract each other due to their gravitational masses. Experiments testing this would have to probe unrealistic small effects. On the other hand without experimental guidance to the microscopic nature of gravity it is problematic to make concrete predictions.

One could on the other hand imagine a theory that encompass all three and would successfully reduce to special relativity, classical gravity and quantum mechanics in the correct limits. There does not exist many candidates to models with this property at the moment. Probably, the most known candidate is String Theory. Many however have the impression, that the most likely way to progress is to first probe the Standard Model and figure out, where and how it breaks down. The fault of a model will often give indications of what is needed to extend it. Optimistically this could eventually lead to a Theory of Everything.

Historically the development of particle physics has happened as a co-development between theoretical predictions and experimental discoveries. It is known that the Standard Model as it exist today cannot be a complete description of nature. On the other hand it is complete in the sense that all the predicted particles have now been discovered. New experimental discoveries can essentially come from two directions: Either by discovery of new particles or by measuring deviations from the predicted Standard Model parameter values. This thesis will describe an attempt of the latter type.

## The Field Point of View

Quantum Mechanics offers an approach to calculate the probability of a given outcome ( $\Phi_O$ ) given a set of initial conditions ( $\Phi_I$ ), and the Hamiltonian ( $\hat{H}$ ) describing the system. This is

normally expressed as an amplitude whose magnitude is interpreted as a probability:

$$U(t) = \langle \Psi_O | \hat{H} | \Psi_I \rangle$$

In the context of special relativity it is natural to ask how a free particle propagates through space. This will indicate an area where quantum mechanics as first formulated fails, and thereby motivates the necessity of a field viewpoint<sup>2</sup>. The free particle amplitude reads:

$$U(t) = \langle \vec{x} | e^{-iHt} | \vec{x}_0 \rangle = \left( \frac{m}{2\pi i t} \right)^{3/2} e^{im(\vec{x}-\vec{x}_0)^2/2t}$$

This is nonzero for all values of  $\vec{x}$ , which means the theory predicts that a particle should be able to propagate faster than the speed of light in contradiction of special relativity<sup>3</sup>.

From analytical mechanics it is known that systems will follow a path over the Lagrangian such that its time integral (the action) is at an extremum:

$$S = \int L dt = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi),$$

for some Lagrangian density  $\mathcal{L}$ . The Lagrangian is a function of the field  $\phi$  and its derivative. The extremum will usually be a minimum. Following the extremal path is equivalent to satisfy the Euler-Lagrange equation:

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

The Lagrangian formalism is well motivated since all expressions are explicitly Lorentz Invariant [7, p.36].

The complementary Hamiltonian formalism for point particles defines the conjugate momentum  $p \equiv \partial L / \partial \dot{q}$  for each dynamic variable  $q$ . Imagining a field as a set of point particles:  $p(x) \sim \pi(x) d^3x$  it is natural to draw the field equivalent

$$p \equiv \frac{\partial L}{\partial \dot{q}} \rightarrow \pi(x) \equiv \frac{\partial L}{\partial \dot{\phi}(x)},$$

for a spacial field  $\phi$ . In a similar approach to the first quantisation, the fields can now be reinterpreted as operators. By imposing commutation relations, a quantum mechanical description is developed:

$$[q_i, p_j] = i\delta_{ij} \rightarrow [\phi(x), \pi(y)] = i\delta^{(3)}(x - y),$$

where the other combinations ( $q$  with  $q$ , ...) commute.

As the most simple example, it is useful to examine the Klein-Gordon field:

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

<sup>2</sup>For a more detailed derivation, see: [7, Sect 2.]. The content presented in this subsection can throughout be attributed to this source.

<sup>3</sup>To be stringent, what contradicts special relativity is that there is a particle type that can travel both slower and faster than  $c$ . The expression here also assumes  $E = p^2/2m$  rather than the correct relativistic  $E = \sqrt{p^2 + m^2}$  for brevity. The conclusion would however be the same as described in the reference.

The equations of motion are derived by imposing the Euler-Lagrange equation:

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

These equations turn out to be the same as for a Harmonic Oscillator. The solutions to the Harmonic Oscillator is given by applying creation operators to the ground state, or annihilation operators to an excited state. In the Klein-Gordon case the excited states are interpreted as particles.

Returning to the question whether the new approach preserve causality. The amplitude for a particle to propagate from  $y$  to  $x$  is:

$$D(x - y) = \langle 0 | \phi(x) \phi(y) | 0 \rangle$$

It is a necessity of Lorentz invariance that any measurement at  $x$  cannot influence a measurement at  $y$  if their separation is space-like. If the commutator between the field at  $x$  and  $y$  vanishes when  $(x - y)^2 < 0$ , the two do not influence each other. Further calculations reveals that:

$$[\phi(x), \phi(y)] = D(x - y) - D(y - x) = 0, \text{ for } (x - y)^2 < 0$$

The equality to zero can be set since the  $D(x - y)$  terms are independently Lorentz invariant, such that the transformation  $x - y \rightarrow -(x - y)$  imposed on the second term guarantees this to be true. Note also that had the separation been time-like  $(x - y)^2 > 0$ , no such transformation would have existed (the causality of two events is fixed if it is possible to send information from one to the other).

The introduction here has been given in a very general setting. A more concrete approach is necessary to obtain predictions that can be used in an experimental setting. The next section will turn to the Standard Model of particle physics by taking the vantage point at the description of a free fermion.

## 1.1 The Standard Model

### 1.1.1 The Standard Model Lagrangian

Consider the Lagrangian describing a free Dirac fermion<sup>4</sup>:

$$\mathcal{L}_0 = i\bar{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - m\bar{\psi}(x)\psi(x)$$

and introduce the global  $U(1)$  transformation:

$$\psi(x) \rightarrow \psi'(x) = e^{iQ\theta}\psi(x)$$

It can be seen that  $\mathcal{L}_0$  is invariant under this type of transformation. The phase  $\theta$  does not have any physical meaning.

If the transformation however is generalised to a local transformation,  $\theta \rightarrow \theta(x)$ , an extra term will emerge from deriving  $\theta$  w.r.t.  $x$ . The Lagrangian is not invariant any more.

Intuitively it seems unphysical that once a phase without physical interpretation is chosen it must be the same for all space. The ‘‘Gauge Principle’’ states that the  $U(1)$  transformation should hold locally. This can be achieved by modifying the Lagrangian with an addition of a new spin-1 field that transforms in the following way:

$$A_\mu(x) \rightarrow A'_\mu(x) \equiv A_\mu + \frac{1}{e}\partial_\mu\theta$$

The problematic derivative from the local gauge transformation can now be replaced by the covariant derivative:

$$D_\mu\psi(x) \equiv [\partial_\mu - ieQA_\mu(x)]\psi(x) \Rightarrow D_\mu\psi(x) \rightarrow (D_\mu\psi(x))' = e^{iQ\theta}D_\mu\psi(x)$$

Local gauge invariance of the Lagrangian is achieved. If  $A_\mu$  has to be a true propagating field, it is necessary to add a kinetic term:

$$\mathcal{L}_{Kin} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}(x), \quad F_{\mu\nu} \equiv \partial_\nu A_\mu - \partial_\mu A_\nu$$

$A_\mu$  can then be identified as the photon. The Lagrangian density of Quantum Electro Dynamics thus emerges:

$$\mathcal{L}_{QED} = i\bar{\psi}(x)\gamma^\mu D_\mu\psi(x) - m\bar{\psi}(x)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}(x)$$

Note that the introduction of the field  $A_\mu$  is very similar to the Classical description of a free particle, with charge number  $Q$ , being introduced to an electromagnetic potential:

$$P \rightarrow P - eQA$$

---

<sup>4</sup>See [8] for a more detailed description of the next section.

The same covariant derivative follows informally with the canonical quantisation of promoting the momentum to the derivative operator:

$$P \rightarrow i\partial \text{ (free)} \quad P - eQA \rightarrow i(\partial_\mu - ieQA_\mu) = iD_\mu \text{ (EM)}$$

The introduction of weak interactions is more involved. It is known experimentally that fermions come in several flavours with different properties for left and right-handed fields. The theory should therefore have doublets. Moreover, the existence of three extra massive gauge bosons ( $W^\pm$ ,  $Z$ ) needs to be coupled to the fermion interactions. The simplest group with doublet representation is  $SU(2)$ . If the theory also has to describe the EM interactions from above it seem natural to consider:

$$G \equiv SU(2)_L \otimes U(1)_Y$$

In the following, only the lepton sector will be considered but the discussion holds equally for quarks as well. The derivation starts by defining the following fields:

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \psi_2(x) = \nu_{eR}, \quad \psi_3(x) = e_R^-$$

As above, it is useful to first consider the free Lagrangian<sup>5</sup>:

$$\mathcal{L}_0 = \sum_{j=1}^3 i\bar{\psi}(x)_j \gamma^\mu \partial_\mu \psi_j(x),$$

and introduce the new  $SU(2)_L$  transformation:

$$U_L \equiv e^{i\frac{\sigma_i}{2}\alpha^i}, \quad \text{for } i = 1, 2, 3$$

with Pauli matrices  $\sigma_i$ . It can be demonstrated that the Lagrangian is again invariant under the global transformation:

$$\begin{aligned} \psi(x)_1 &\rightarrow \psi_1'(x) = e^{iy_1\beta} U_L \psi_1(x), \\ \psi(x)_2 &\rightarrow \psi_2'(x) = e^{iy_2\beta} \psi_2(x), \\ \psi(x)_3 &\rightarrow \psi_3'(x) = e^{iy_3\beta} \psi_3(x) \end{aligned}$$

Following the same logic as above, the theory is required to be invariant under local gauge transformations, where the phases  $\alpha^i$  and  $\beta$  are made position dependent.

Since there are now four gauge parameters, four gauge bosons are needed to correct the

---

<sup>5</sup>As opposed to before, all particles are throughout the next assumed to be massless.

Lagrangian. The new covariant derivatives are as follows:

$$\begin{aligned} D_\mu \psi(x)_1 &\equiv [\partial_\mu - ig\tilde{W}_\mu(x) - ig'y_1 B_\mu(x)]\psi_1(x) , & \tilde{W}_\mu(x) &\equiv \frac{\sigma_i}{2} W_\mu^i(x) \\ D_\mu \psi(x)_2 &\equiv [\partial_\mu - ig'y_2 B_\mu(x)]\psi_2(x) , \\ D_\mu \psi(x)_3 &\equiv [\partial_\mu - ig'y_3 B_\mu(x)]\psi_3(x) \end{aligned}$$

Local gauge invariance requires the fields to transform as:

$$\begin{aligned} B_\mu(x) &\rightarrow B'_\mu(x) = B_\mu(x) + \frac{1}{g'}\partial_\mu\beta(x) \\ \tilde{W}_\mu(x) &\rightarrow \tilde{W}'_\mu(x) = U_L(x)\tilde{W}_\mu U_L^\dagger(x) - \frac{i}{g}\partial_\mu U_L(x)U^\dagger(x) \end{aligned}$$

Again a gauge invariant kinetic term is added, resulting in the Electroweak Lagrangian density for interactions of massless fermion fields and four massless vector bosons:

$$\mathcal{L} = \sum_{j=1}^3 i\bar{\psi}(x)_j \gamma^\mu D_\mu \psi(x)_j - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} \quad (1.1)$$

The field strengths  $B_{\mu\nu}$  and  $W_{\mu\nu}^i$  can e.g. be found in [8, Eq. (13.14),(13.16)].

It is not clear from looking at Eq.1.1 how the physical fields  $\gamma$ ,  $W^\pm$  and  $Z$  are described. By dismantling the expression, the charged current bosons can be realised to be described by the superpositions:  $W_\mu \equiv (W_\mu^1 + iW_\mu^2)/\sqrt{2}$  and  $W_\mu^\dagger \equiv (W_\mu^1 - iW_\mu^2)/\sqrt{2}$ .

The photon and  $Z$  on the other hand emerges as a rotation of the remaining two bosons<sup>6</sup>:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \equiv \Theta_W \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}$$

where  $\Theta_W$  describes the rotation matrix of angle  $\theta_W$ . If the Lagrangian is expressed in terms of the rotated fields it, will contain an  $A_\mu$  factor:  $g\frac{\sigma_3}{2} \sin(\theta_W) + g'y_j \cos(\theta_W)$ . If QED is to be reproduced by the theory, it must be required that:

$$g \sin(\theta_W) = g' \cos(\theta_W) = e , \quad Y = Q - T_3 , \quad T_3 = \frac{\sigma_3}{2} ,$$

where  $Q$  is the Electromagnetic charge operator:

$$Q_1 = \begin{pmatrix} Q_{u/\nu} & 0 \\ 0 & Q_{d/e} \end{pmatrix} , \quad Q_2 = Q_{u/\nu} , \quad Q_3 = Q_{d/e}$$

The first part relates the theory to the Electroweak interaction. The second part reveals fermion hypercharges in terms of the electric charge and weak isospin number. A unified framework that is able to describe electromagnetic and weak interactions has been found.

<sup>6</sup>For now the fact that the  $Z$  and  $W$  are massless in the theory is ignored. This will be addressed in section 1.2

It is worthwhile mentioning that the kinetic term that was added to Eq.1.1 also contain the terms for the gauge boson self-interaction:  $(Z/\gamma)W^+W^-$ ,  $W^+W^-W^+W^-$  and  $(Z/\gamma)(Z/\gamma)W^+W^-$ .

Without going into detail, the QCD Lagrangian follows a similar derivation with slight modification. In order to explain the multitude of different mesons ( $q\bar{q}$ ) and baryons ( $qqq$ ) observed in experiments, it is necessary that each quark species have three different strong charges, referred to as colour. Again one can start with the free Lagrangian for quarks with three different colours and check that it is invariant under global  $SU(3)_C$  transformations in colour space. The symmetry group has eight generators giving as many invariant gauge parameters. To upgrade the Lagrangian to local gauge invariance, eight vector bosons are introduced to create the covariant derivative. The bosons are identified as gluons. Some algebra moreover reveals that the gluons have self interactions such that vertices of the type  $ggg$  and  $gggg$  are allowed.

### 1.1.2 Amplitude Calculations

A rough outline of the Standard Model Lagrangian has been given. A framework for drawing concrete predictions out of the model is still needed. Classically the field would evolve in a way that minimises the action. The quantum mechanical equivalent however does not allow a definite path.

It is possible however to interpret the QM path as the classical path with higher order additions. The QM description instead corresponds to a sum over all possible paths weighted by the path amplitude. In principle it would be possible to solve the Euler Lagrange equations to make predictions. In practice this is impossible to do for anything other than the simplest scenarios. A different approach is needed.

The most well known approach to amplitude calculations can be attributed to Richard Feynman, with the method of Feynman diagrams.

The recipe starts by dividing the Lagrangian into its interacting and free parts. By dividing the Lagrangian, it is possible to describe scenarios where particles propagate from infinity, interact and propagate to infinity again. The interacting part will then contain all the relevant physics. Notice how this approximation resembles what is done in collider experiments.

The Matrix Element from scatter theory will be used to describe the interacting part. If interaction strengths are assumed small, it is possible to estimate the Matrix Element by perturbation. This is a good approximation for QED and Electroweak models. Due to the asymptotic freedom of QCD however, perturbation is only valid at high momentum transfers. This is also the reason why the method performs well in describing hard interactions in collision experiments but has difficulties describing e.g. the internal mechanics of a proton.

A different approach has been derived to describe low-energy behaviour: Calculations are simplified by dividing space into a lattice. Accurate predictions can be made by examining the predicted behaviour as the lattice spacing tends to zero.

Returning again to the Feynman rules. The Lagrangian is divided into its kinetic terms, identified by two field components, and interaction terms which are identified by three or more field components. The kinetic terms describe how the fields propagate, while particle scattering is described by the interaction terms. The matrix element calculation is in the next step translated to the construction of a set of Feynman diagrams.

The recipe is outlined as follows:

- Identify the initial and final state particles. Each is represented by an ingoing or outgoing line to the diagram.
- Connect via vertices and internal lines (propagators) all ingoing and outgoing lines. Each vertex will be associated with a coupling strength.
- Momentum is conserved at all vertices such that the sum of incoming and outgoing is the same.

Each component in the diagram correspond to a factor of the amplitude calculation. Their total product result in  $i\mathcal{M}$ .

Since it is assumed that the interactions are weak, such that perturbation is valid, the specific diagram contribution to  $\mathcal{M}$  will be ordered in the number of vertices. An example of the leading order diagrams contributing to  $q\bar{q} \rightarrow ZZ$  production is seen in Figure 1.2.

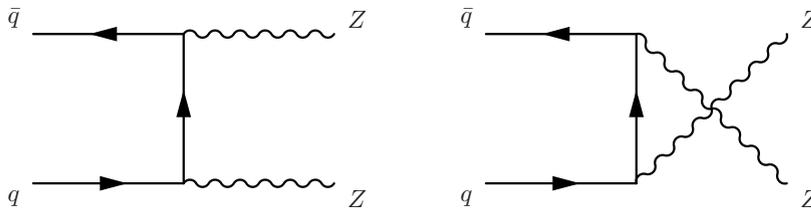
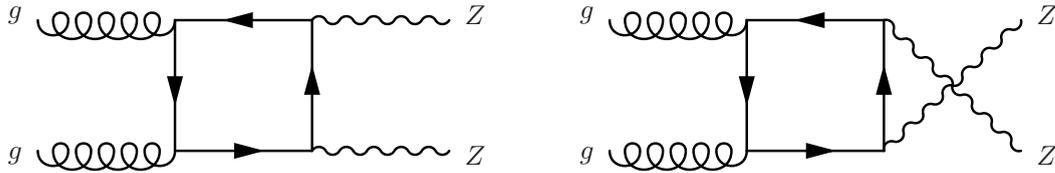


FIGURE 1.2: Leading order contributing processes of  $q\bar{q} \rightarrow ZZ$ .

The recipe however does not constrain the momentum of internal lines in loops. It is necessary to integrate over all possible momenta to get the final prediction. This is for example the case for the leading order diagrams for  $gg \rightarrow ZZ$  production, Figure 1.3. Note the difference to the  $q\bar{q}$  case. The Standard Model does not allow vertices of the type  $gqZ$ , such that the leading order process now has to be constructed with four vertices. The momentum flowing through the quark box is not constrained.

Theories will tend to diverge as the internal propagator momenta goes to infinity. The theory is said to be ultraviolet divergent.

Some mechanism is needed to counteract the divergence such that predictions of the theory are well defined. One could imagine a cutoff scheme, where momenta are confined to be below some energy-scale  $\Lambda$ . If  $\Lambda$  is chosen to be above the relevant energies of the study, it could be possible to make valid predictions. By introducing a maximal energy, the theory is said to be

FIGURE 1.3: Leading order contributions processes of  $gg \rightarrow ZZ$ .

regularised. It is however undesirable to have a model that only makes physical predictions by introducing an arbitrary scale. The introduction of a maximal energy is also equivalent to choosing a minimal length scale. Since QFTs are defined in a continuous space, this also reveals a problematic feature of the approach.

Renormalisation offers another solution to the divergency problem. The renormalisation procedure is closely related to how the coupling constants of the theory behaves. When renormalising, the coupling constants are interpreted as having a bare value and a physical value, where the latter is the one that would be measured in experiments. This way the physical value can be modified to be a function of energy such that it cancels the divergent behaviour. The physical value of a coupling  $g'$  is now a function of its bare value  $g$ , the momentum transfer  $Q$  and possibly an energy scale  $\Lambda$ :

$$g' = f(g, \Lambda, Q)$$

By fixing  $g$  and  $\Lambda$  the running of the coupling can be measured at a given  $Q$ . Once the physical coupling constant has been measured, the coupling can be extrapolated to other scales.

This method of concealing divergent behaviour might seem rather arbitrary at first glance. It is however actually measured in real life, where e.g. the electromagnetic coupling is  $\alpha_{QED} \simeq 1/137$  when measured at low energy while it changes to  $\alpha_{QED} \simeq 1/128$  when probed at the electroweak scale [9, Ch.10].

One can also interpret this phenomenon as the charged particle having a bare value, but will polarise the sea of virtual particles that surrounds it in the vacuum. As a result, the bare value is screened by the polarised vacuum when probing at some length scale.

### 1.1.3 General Considerations

The text will now turn to a more pragmatic point of view. Figure 1.4 offers an overview of the particles found in the Standard Model. These are thought to be the elementary constituents of matter and interactions<sup>7</sup>. A particle is characterised by a set of different properties, including mass, charge and spin.

The list of particles can be divided into two groups, fermions (spin-1/2) and bosons (spin-0 or spin-1) obeying Fermi or Bose-Einstein statistics, respectively. The fermions are the constituents of matter, while the spin-1 bosons are force carriers. The only observed spin-0 particle is the Higgs boson which is believed to give the other particles masses. The three

<sup>7</sup>Excitations of what is described as elementary particles, which would imply inner structure is searched for. Limits on e.g. the mass of excited electrons and muons are at the current status  $\sim 2$  TeV[11]

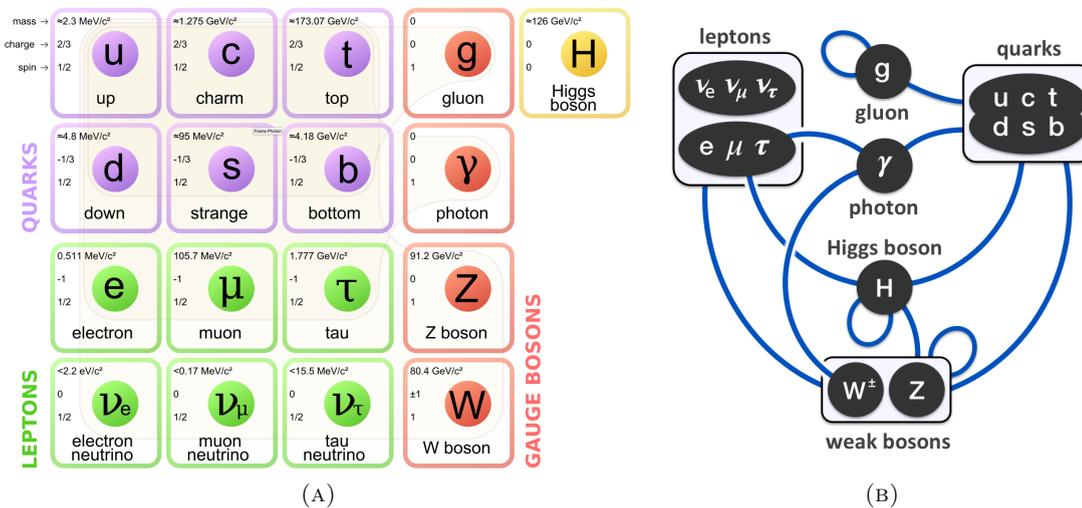


FIGURE 1.4: (A): Elementary particles in the Standard Model. (B) shows which of the Standard Model particles are able to interact[10]

forces the Standard Model describes are the strong, weak and electromagnetic, mediated by the gluon, W/Z bosons and photon. The forces carriers are collectively known as vector bosons. The matter constituents are divided into quarks and leptons. All particles also have an anti-partner, which has the same mass but opposite sign on conserved quantum numbers. A more detailed table of the particle names, masses and charges can be found in Table 1.1.

## Bosons

The gauge bosons describe the way particles interact by transmitting the three forces that the Standard Model describes, the weak, the strong and the electromagnetic.

The electromagnetic force is carried by the photon ( $\gamma$ ) and couples to all particles that have electric charge. The photon is a massless particle giving the electromagnetic force infinite range. Like the other gauge bosons, it is a spin-1 particle, but since it is massless its spin has only two possible projections instead of three.

The weak force is transmitted by the two weak charged currents ( $W^\pm$ ) and the weak neutral current ( $Z^0$ ), similarly to the photon interacting with everything that has charge, these couple to weak hyper charge. In Figure 1.4a, the top row of each type of fermion has isospin  $-1/2$  while the second has  $1/2$ . The force conserves weak isospin. Examples of processes obeying this, can e.g. be seen in the beta-minus decay of atoms, where a down quark is converted to an up quark, effectively changing the neutron to a proton:  $n \rightarrow p + e^- + \bar{\nu}_e$

The massiveness of the weak force bosons limits them to very small ranges. The W bosons have, as an example, a mass around 80 GeV, which gives the force a characteristic range of  $\frac{1}{M_W} \sim 10^{-18}$  meters. For this reason it is not seen in many everyday interactions. At interactions with energies around the weak scale ( $\sim 100$  GeV), it unifies with the electromagnetic force giving the two interactions similar characteristics.

The bosons responsible for the strong force are the eight gluons. The strong force differs from the other two in that it has three distinct charges, red, green and blue. The eight different

Name	Symbol	Mass [MeV]	EM Charge	Weak Isospin	Colour charge
<b>leptons</b> (spin 1/2)					
Electron	e	0.5485	-1	-1/2,0	
Muon	$\mu$	113.4	-1	-1/2,0	
Tau	$\tau$	1776	-1	-1/2,0	
Electron Neutrino	$\nu_e$	$< 2 \cdot 10^{-6}$		1/2	
Muon Neutrino	$\nu_\mu$	$< 0.19$		1/2	
Tau Neutrino	$\nu_\tau$	$< 18.2$		1/2	
<b>Quarks</b> (spin 1/2)					
Up	u	$\sim 2.3$	2/3	1/2,0	R/G/B
Down	d	$\sim 4.8$	-1/3	-1/2,0	R/G/B
Charm	c	1275	2/3	1/2,0	R/G/B
Strange	s	$\sim 95$	-1/3	-1/2,0	R/G/B
Top	t	$173 \cdot 10^3$	2/3	1/2,0	R/G/B
Bottom	b	4180	-1/3	-1/2,0	R/G/B
<b>Gauge Bosons</b> (spin 1)					
Photon	$\gamma$	$< 1 \cdot 10^{-18}$		*	
Weak charged current	$W^\pm$	$80.34 \cdot 10^3$	$\pm 1$	$\pm 1$	
Weak neutral current	$Z^0$	$91.18 \cdot 10^3$		*	
Gluon	g	$\sim 0$			Octet
Higgs Boson (spin 0)	$H^0$	$\sim 125.09 \cdot 10^3$			

TABLE 1.1: Particles of the Standard Model and their properties. Mass values are taken from [9] where uncertainties have been omitted. The Higgs Boson mass is from [12], the gluon mass is a theoretical prediction[9]. The weak hypercharge ( $Y_W$ ), the EM ( $Q$ ) charge and the weak isospin ( $T_3$ ) are related:  $Y_W = 2(Q - T_3)$ . \*The photon and weak neutral current are a rotation of the  $B_\mu$  and  $W_\mu^3$  fields with weak isospin 0 and 1 respectively.

gluons correspond to the possible superpositions of the colour charges.

The strong force is moreover confined. This means that there cannot exist free particles with a colour charge different from zero. The strong force is responsible for binding quarks together in hadrons of which the proton and the neutron are well known examples.

The Higgs particle is the last known fundamental boson. It offers a way to explain the masses of the Electroweak bosons,  $\gamma$ ,  $W^\pm$  and  $Z^0$ . By introducing new couplings into the Standard Model it is also able to give a non-predictive description of fermions masses<sup>8</sup>.

## Leptons

The left-handed leptons are grouped in three generations, each containing two particles. The first consist of the electron ( $e^-$ ) and the electron neutrino ( $\nu_e$ ). The following generations contain the muon ( $\mu^-$ ) and tau ( $\tau^-$ ) and their respective neutrinos ( $\nu_\mu$ ,  $\nu_\tau$ ). The right-handed leptons only include the electron, muon and tau.

In principle the Standard Model is not restricted to solely contain three generations ( $\nu_e, \nu_\mu, \nu_\tau$ ). A precise measurement of the  $Z$  decay-width however establish a preference of three neutrino

<sup>8</sup>Non-predictive because it is necessary to introduce a coupling per mass. It is however possible to imagine that it couples to fermions, but not in proportion to their masses as would be the case for the Standard Model Higgs.

families [13].

The electron, muon and tau have electric charges of -1, spin 1/2 and weak isospin of -1/2. They both couple to the electromagnetic and weak force. The neutrinos have hypercharge 1/2 and only couple to the weak force. The difference between the electron, muon and tau is their masses, as can be seen in Table 1.1 and their lepton number. The lepton number is a conserved quantity, preventing e.g. muons to decay to electrons without the emission of neutrinos.

It has however been established that neutrinos oscillate, meaning they will have flavour transitions in flight. The transition rate is dependent on flight length and neutrino energy<sup>9</sup>.

Neutrino oscillations are described by introducing lepton-number states of the neutrinos that are (non-trivial) superposition of at least three mass eigenstates. Neutrino oscillations can only occur if at least two of the states are massive.

The transition rate can be shown to be dependent on the squared mass difference. The oscillation rates can therefore not be used as a direct measure of their masses. Two of the mass eigenstates have been determined to have a squared mass difference of  $\Delta m_{12}^2 \sim 10^{-4} \text{ eV}^2$ , while the third have a squared difference to the mean of the other two of  $\Delta m^2 \sim 10^{-3} \text{ eV}^2$  [9, Tab. 14.7].

Direct measurements of the neutrino mass scale can be tried in different ways. A model independent way is to study the kinematics of  $\beta$ -decays. Other more model dependent measurements include the delay between neutrinos and light from super-novae ( $< 5.7 \text{ eV}$ [15],  $< 5.8 \text{ eV}$ [16]) and the neutrino mass influence on the background radiation power spectrum ( $< 0.5 \text{ eV}$ [17]).

## Quarks

The quarks together with the leptons form all known forms of matter. They differ from the leptons in several ways: They do not have lepton numbers, they have either charge numbers of 2/3 and -1/3. The quarks are also fermions with spin 1/2. The quarks again consists of three families, containing the up (u) and down (d) quark, the charm (c) and strange (s) and the top (t) and bottom (b) quarks, where the first mentioned have charge 2/3 and the second -1/3, Table 1.1. The quark masses span three orders of magnitude.

Due to the colour confinement of the strong force, the quarks tend to cluster into quark anti-quark pairs (mesons) or groups of three (baryons). Any quark composition is commonly called a hadron. Observation of resonances that contain four[18] and five[19] quarks have been made. The up and down quarks are the constituents of the proton (uud) and neutron (udd), which together with the electron forms the stable matter observed in everyday life.

A complicating factor when colliding protons, as is the case with the LHC, is that it is far from sufficient to describe the proton as three quarks. A sea of gluons bind the quarks together and confines them in the proton. Gluons moreover not only couple to quarks, but also to themselves. This means that there is a certain probability for a gluon to not only split into a virtual

<sup>9</sup>For a comprehensive description of neutrino oscillations, see [9, Ch. 14]. An overview of direct neutrino mass measurements is given in [14].

quark anti-quark pair but also into a virtual gluon-gluon pair from quantum fluctuations.

## Production Cross Sections

In the context of collision experiments it is useful to consider how probable the creation of different particles are. Figure 1.5 show the production probability for different types of Standard Model processes as measured in ATLAS compared to theoretical predictions. In general it can be seen that proton collisions will mostly consist of QCD final states. The production of the heavy electroweak bosons and diboson production are found at much lower probabilities. The Standard Model is accurately able to predict cross sections over many orders of magnitude. The figure is included here, but will be used as a reference for back of the envelope estimates a number of times throughout the text.

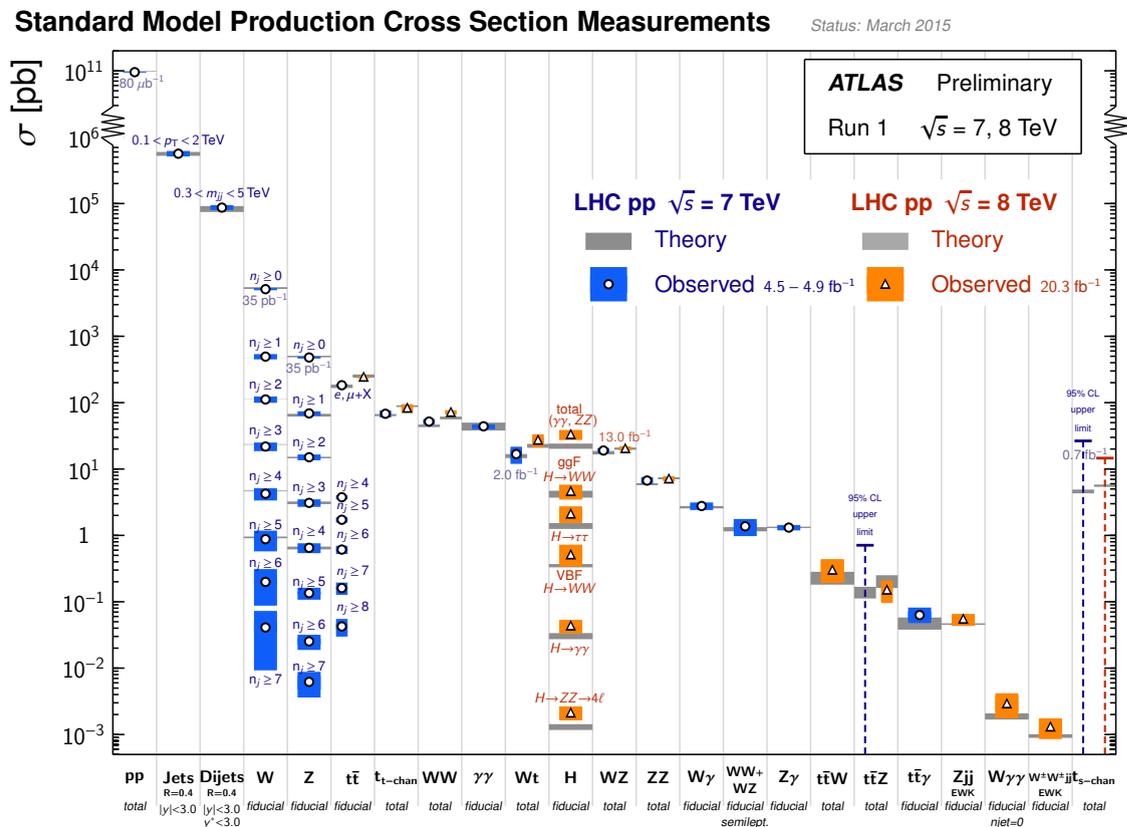


FIGURE 1.5: Cross sections for various Standard Model processes. The figure shows both theoretical predictions and the ATLAS measurements[20].

### 1.1.4 Proton Collision Kinematics

The Standard Model offers a recipe to calculate the amplitude of an outcome given incoming particles of certain momenta. The full cross section is given by integrating over the final state phase spaces.

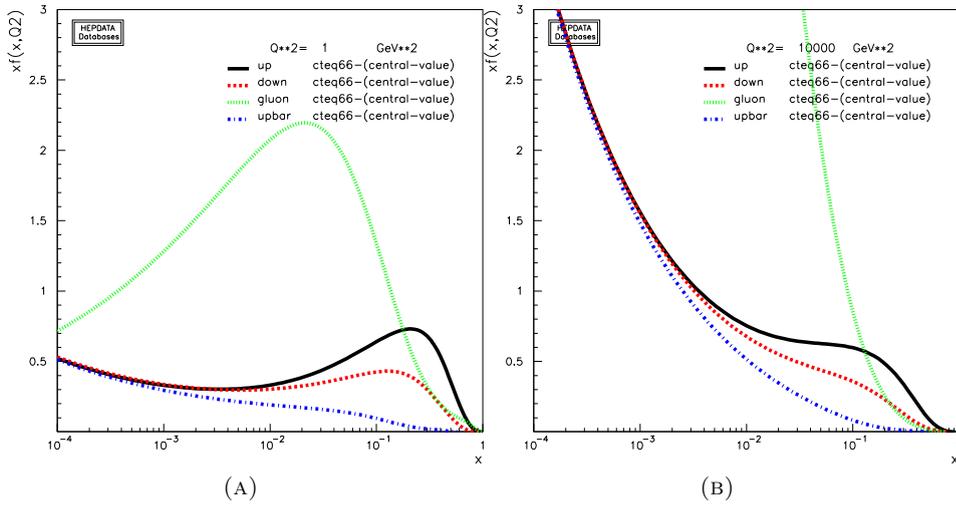


FIGURE 1.6: (A) Parton distribution function (PDF) for a proton probed at  $Q^2 = 1 \text{ GeV}^2$ . Partons are represented as:  $u$  (black),  $d$  (red),  $\bar{u}$  (blue) and  $g$  (green). Notice that the valence quark ( $u, d$ ) tend to carry  $1/3$  of the momentum for high  $x$  values. This correspond to the 'classical' proton picture. (B): PDF of a proton probed at the weak energy scale ( $Q^2 = 10^4 \text{ GeV}^2$ ). At values of  $x \lesssim 0.2$  gluons totally dominates the parton distribution function. PDFs are generated by [21].

In hadron collider scenarios, the incoming particles will be quarks or gluons, but it is clear that for a proton with constant momentum its constituents will not have fixed momenta. The probability of 'drawing' a parton out of a proton with a certain fraction of the total momentum ( $x$ ) given the scale at which the proton is probed ( $Q$ ) will therefore have to be included in the cross section calculation. For a centre of mass energy  $\sqrt{s}$  and parton collision energy  $M$ :

$$x_1 x_2 = M^2/s ,$$

where for the two partons:

$$x_i = \frac{P_{|| \text{ parton}}}{P_{|| \text{ proton}}}$$

Taking the structure of the proton into account, the measured cross section for a differential cross section  $d\hat{\sigma}$  is:

$$d\sigma \sim \int_{x_1 x_2 = M^2/s} dx_1 dx_2 f^A(x_1) f^B(x_2) d\hat{\sigma} ,$$

where e.g.  $f^A(x)$  describes the probability to draw a parton of momentum fraction  $x$  out of proton A.

It is of high importance in experimental particle physics to know the parton density functions  $f(x)$  to make precise predictions. Collaborative efforts of the determination are described in [9, Ch. 19], where MSTW[22, 23], NNPDF[24] and CT(EQ)[25] can among others be mentioned. The Monte Carlo used in the analysis is based on the CT(EQ) distribution functions.

Figure 1.6 shown the distributions for two different momentum transfers;  $Q^2 = 1 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$  (LHC scale). With this in mind it is not completely unreasonable to think of the LHC as mainly a gluon collider. A schematic picture of a proton is shown in Figure 1.7.

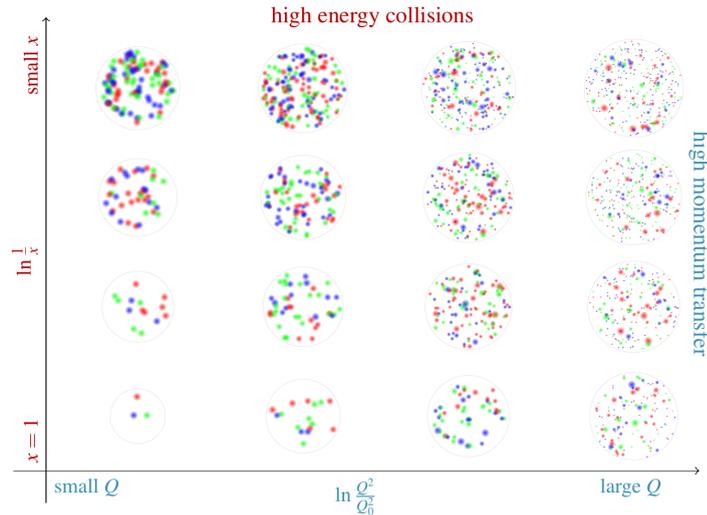


FIGURE 1.7: Schematic drawing of proton interactions at different levels of probing energy and momentum fractions. In the bottom left corner a proton is probed with low energy (small  $Q$ ) while a significant part of its momentum is contained in the observed partons ( $x \simeq 1$ ). At lower  $x$  values the image gets more distorted as the probability to draw a gluon increases. At higher momentum transfers the virtual particles become increasingly dominating[26].

Another useful consideration of proton collision kinematics is to examine the free single particle behaviour. If the phasespace does not have any preference for Cartesian Momentum components, the differential phasespace volume is[27, App. C]:

$$d^4P \delta(E^2 - p^2 - m^2) = p_T \cdot dp_T d\phi dy, \quad dy = dp_{||}/E,$$

where the differential rapidity  $dy$  is introduced. With a slight rewriting it can be seen that this conforms with the standard expression for rapidity used in particle physics:

$$y = \frac{1}{2} \ln \left( \frac{E + p_{||}}{E - p_{||}} \right) = \ln \left( \frac{E + p_{||}}{m_T} \right)$$

The differential form follows from  $dy = (dy/dp_{||})dp_{||}$  and  $E = \sqrt{p_{||}^2 + m_T^2}$ .

If the phasespace does not have any preference it follows that  $y$  and  $\phi$  are approximately flat. The large QCD backgrounds should then conveniently be uniformly distributed in these coordinates. Of course this cannot hold true to infinite rapidities, so something else will have to constrain their kinematics.

Another set of useful identities are:

$$E = m_T \cosh y \quad p_{||} = m_T \sinh y$$

By inserting in the above equation for  $y$  the expressions can be seen to be consistent.

The new identities can be used to re-write the momentum fractions of the two protons. Introducing the difference in fractions:  $x = x_1 - x_2$  :

$$x = \frac{m_T \sinh y}{P} = \frac{2m_T \sinh y}{\sqrt{s}}$$

The limit of rapidity is found when  $x$  tends to one:

$$x \simeq \frac{M}{\sqrt{s}} e^y \Rightarrow \Delta y \simeq 2 \ln \frac{\sqrt{s}}{M}$$

The particle rapidity will moreover asymptotically approximate its pseudo-rapidity  $\eta$  when  $E \gg m$ .

Based on these considerations it is expected that the low mass particle distributions will be approximately flat in  $\eta$  in the  $\Delta y$  interval from above. This e.g. explains why it is useful for particle isolation to use windows of  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ .

## 1.2 The Higgs Mechanism

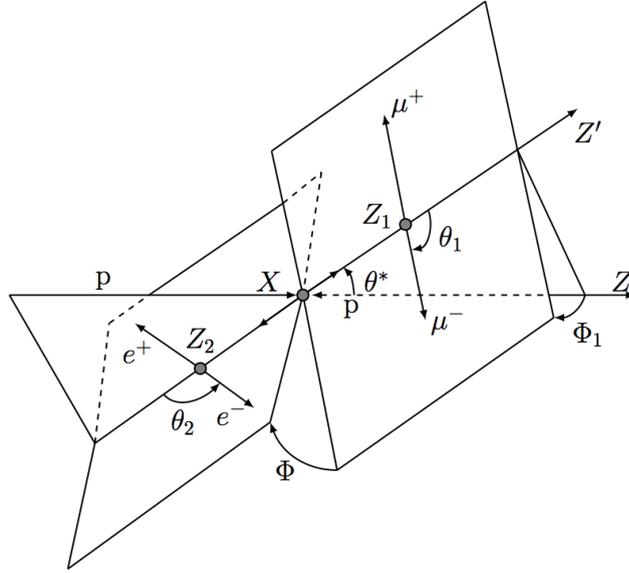


FIGURE 1.8: Decay diagram of the  $H \rightarrow ZZ^* \rightarrow 4\ell$  final state, which will be the main focus of this thesis[5]. Angle definitions inspired by [28].

### 1.2.1 Spontaneous Symmetry Breaking

The Electroweak Lagrangian was introduced in Section 1.1.1. The derivation concluded by having a model with massless bosons. This is acceptable for the photon field but not the weak bosons<sup>10</sup>. It was not possible directly to add mass terms to the Lagrangian since this will break the local gauge symmetry.

As a simple example of spontaneous symmetry breaking, a Lagrangian, which is both invariant under a group ( $G$ ) of transformations and has a degenerate set of states with minimal energy can be examined.

Take for instance the complex scalar field  $\phi(x)$  with Lagrangian:

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi), \quad V(\phi) = \mu^2 \phi^\dagger \phi + h \cdot (\phi^\dagger \phi)^2 \quad (1.2)$$

$\mathcal{L}$  is invariant under the transformation:

$$\phi(x) \rightarrow e^{-i\theta} \phi(x)$$

It must be required, for a potential of this type, that  $h \geq 0$  such that it is bounded from below and has a ground state.  $\mu^2$  can on the other hand be chosen both to be negative and positive. The positive choice will give the potential a trivial  $\phi = 0$  minimum and correspond to the introduction of a massive scalar particle.

The negative choice is interesting for the discussion of spontaneous symmetry breaking. The

<sup>10</sup>The following derivation can be attributed to [8], where more detail can also be found.

minimum is now non-zero as will satisfy:

$$|\phi_0| = \sqrt{\frac{-\mu^2}{2h}} \equiv \frac{v}{\sqrt{2}} > 0, \quad V(\phi_0) = -\frac{h}{4}v^2$$

It is clear that the choice of the complex phase of  $\phi_0$  is free and will lead to the same energy state. The Lagrangian has in other words a degenerate minimum. By making a choice (e.g.  $\theta = 0$ ) the symmetry is said to be spontaneously broken. To see how a massive state enters into the Lagrangian, it is useful to parameterise the field around the ground state:

$$\phi(x) \equiv \frac{1}{\sqrt{2}}[v + \psi_1(x) + i\psi_2(x)]$$

The new parameterisation can be put back into the potential of the original Lagrangian:

$$V(\phi) = V(\phi_0) - \mu^2\psi_1^2 + hv\psi_1(\psi_1^2 + \psi_2^2) + \frac{h}{4}(\psi_1^2 + \psi_2^2)^2$$

$\psi_1$  can be identified with a massive field ( $m^2 = -2\mu^2$ ), and can be understood as the excitation in the direction of the potential that cost energy.  $\psi_2$  on the other hand corresponds to excitations in the degenerate direction. Since these excitations do not cost energy it is a massless state, a Goldstone boson.

If the Lagrangian is invariant under the continuous symmetry  $G$ , but vacuum is only invariant under subgroup  $H$  in  $G$  there exist as many massless spin-0 particles as broken generators, i.e. in  $G$  but not in  $H$ [29].

Returning to the more physically relevant case of Electroweak interactions, the trick is now to introduce a doublet of complex scalar fields:

$$\phi(x) \equiv \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}$$

Similar to before, the scalar Lagrangian now reads:

$$\mathcal{L} = (D_\mu\phi)^\dagger D^\mu\phi - \mu^2\phi^\dagger\phi - h(\phi^\dagger\phi)^2, \quad (h > 0, \mu^2 < 0)$$

The new Lagrangian is invariant under local gauge transformations with the covariant derivative:

$$D^\mu\phi = \left[ \partial^\mu - ig\tilde{W}^\mu - ig'y_\phi B^\mu \right] \phi, \quad y_\phi = Q_\phi - T_3 = \frac{1}{2}$$

where the choice of hypercharge is made such that the photon will not couple to  $\phi^{(0)}$  and  $\phi^+$  acquires the correct electric charge.

The ground state obtains the degenerate minimum:

$$|\langle 0|\phi^{(0)}|0\rangle| = \sqrt{\frac{-2\mu^2}{2h}} = \frac{v}{\sqrt{2}}$$

The symmetry group is again broken by choosing a ground state.

By having the photon not couple,  $U(1)_{QED}$  remains a true symmetry of the vacuum. Considering the difference in the number of symmetries of the vacuum state and the Lagrangian, there will exist three massless particles. The scalar doublet is parameterised in the general form:

$$\phi(x) = \frac{1}{\sqrt{2}} e^{i \frac{\sigma_i}{2} \theta^i(x)} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

containing the three real fields  $\theta^i$  and the real  $H(x)$ . It can be demonstrated that by the local gauge invariance of  $\mathcal{L}$  it is possible to rotate away any dependency on  $\sigma_i$ , which are identified as the three massless Goldstone bosons. If the gauge  $\theta^-(x) = 0$  is chosen, the field  $\phi$  can be put into the original Lagrangian. The kinetic term are then of the form:

$$\theta^-(x) = 0 : (D_\mu \phi)^\dagger D^\mu \phi = \frac{1}{2} \partial_\mu H \partial^\mu H + (v + H)^2 \left( \frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right)$$

It is seen that the vacuum expectation value of the introduced boson has generated quadratic (mass) terms of  $W^\pm$  and  $Z$ . The expression moreover reveals that  $M_Z \cdot \cos \theta_W = M_W = \frac{1}{2} v g$ .

## Fermion Masses

The derivation of the Electroweak Lagrangian in Section 1.1.1 explicitly assumed the fermions to be massless. This was deliberate since introducing a mass term of the form  $\mathcal{L} = -m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)$  would break local gauge invariance. The scalar doublet introduced above can however also generate fermion masses by introducing gauge invariant terms with structure[8]:

$$\mathcal{L} = c (\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + h.c.$$

After the spontaneous symmetry breaking this term will take the form:

$$\mathcal{L} = \frac{1}{\sqrt{2}} (v + H) (c \bar{e} e) , \quad m_e = -c \frac{v}{\sqrt{2}}$$

Notice that this does not directly constrain the theory since the fermion mass is just replaced by the parameter ( $c$ ). It does however provide a gauge invariant mass term. Even though the individual fermion masses are not predicted, the coupling to fermions is proportional to their mass. The assumption can be tested by measuring the relative branching ratio of Higgs to fermion decays.

## 1.2.2 Intermezzo: Predictions, Discovery and Property Measurements

The derivation of the Electroweak unification predicts that there exist a massive scalar boson with properties:

- Spin-0
- Positive parity

- Coupling to  $W^\pm$ ,  $Z$  and the massive fermions
- Mass around the Electroweak scale

The introduction does however not include a concrete prediction of the mass. It is however possible from the knowledge of the masses of heavy particles to make indirect predictions. The derivation concluded by predicting that:

$$M_W = M_Z \cos \theta_W$$

The weak mixing angle  $\theta_W$  (at scale  $\sim m_W$ ) can be measured independently through e.g. the forward backward asymmetry of  $Z/\gamma^*$  production in  $p\bar{p}$  collisions[30][31].

The predicted and measured  $W$  mass can then be compared, but without further considerations this turns out to be a bad match<sup>11</sup>.

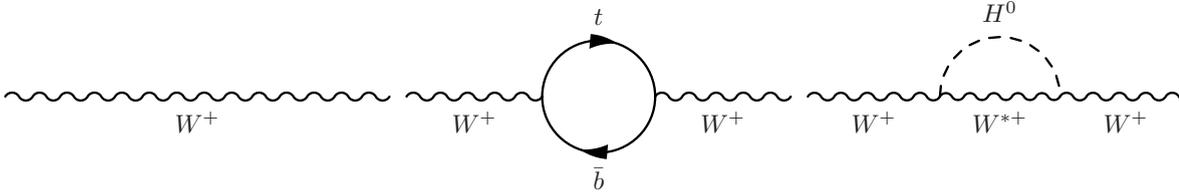


FIGURE 1.9: Left: Propagator of a  $W$  boson. Middle: Largest single loop correction to the  $W$  propagator ( $t\bar{b}$ ). Right: Loop correction to the  $W$  propagator by a Higgs boson.

From QFT it is known that the propagator of the  $W$  boson takes the form  $\sim \frac{1}{p^2 + M^2}$ . As illustrated in Figure 1.9, the propagator will have higher order corrections, effectively giving it an operational mass rather than a bare mass. The fermion and boson propagators are in the large momenta limit ( $p \gg M$ )  $1/p$  and  $1/p^2$ , respectively. The mass correction to the  $W$  is given by the correction propagators integrated over all possible momenta:

$$\delta M_{\text{corr}}^2 \sim m_{\text{fermion}}^2 \quad \delta M_{\text{corr}}^2 \sim \ln m_{\text{boson}} \quad (1.3)$$

Due to charge conservation, the loop corrections to the  $Z$  and  $W$  propagator will however be different, such that for  $Z$  bosons, the loop will contain  $t\bar{t}$  while e.g.  $W^+$  will contain  $t\bar{b}$ , as the largest fermion contribution. The very precise knowledge of the  $Z$  mass, the fine structure constant  $\alpha$ , the Fermi Coupling constant  $G_F$  and the weak mixing angle  $\cos \theta_W$ [32][9] makes it possible to calculate the effective  $W$  mass up to a correction that is dominated by the top and Higgs mass:

$$M_W^2 = M_Z^2 \cos^2 \theta_W (1 + \delta), \quad (1.4)$$

where the top and Higgs corrections are:

$$\delta_t \sim \frac{3\alpha_W}{16\pi} \left( \frac{m_t}{M_W} \right)^2 \quad \delta_H \sim -\frac{11\alpha_W \tan^2 \theta_W}{24\pi} \ln \frac{M_H}{M_W} \quad (1.5)$$

<sup>11</sup>The following calculations will follow the argumentation described in [27, Ch. 4.8], equations can be attributed to this source.

The equation can be inverted to obtain a Higgs mass prediction as function of the  $W$  and top mass. The fit results are shown in Figure 1.10, where it is seen that a mass of 125 GeV is around  $1\sigma$  from the best fit value. The calculations make the assumption that the Standard Model including the Higgs boson describe all particles with relevant couplings. If however heavy undiscovered particles exist, they could also contribute to the correction loop diagrams, making the predicted Higgs mass biased. With the Higgs discovery and subsequent mass measurement, very precise knowledge of the  $W$  mass could lead to indications of new physics. Unfortunately the logarithmic dependence of the correction to the  $W$  mass makes it difficult to improve the constraint.

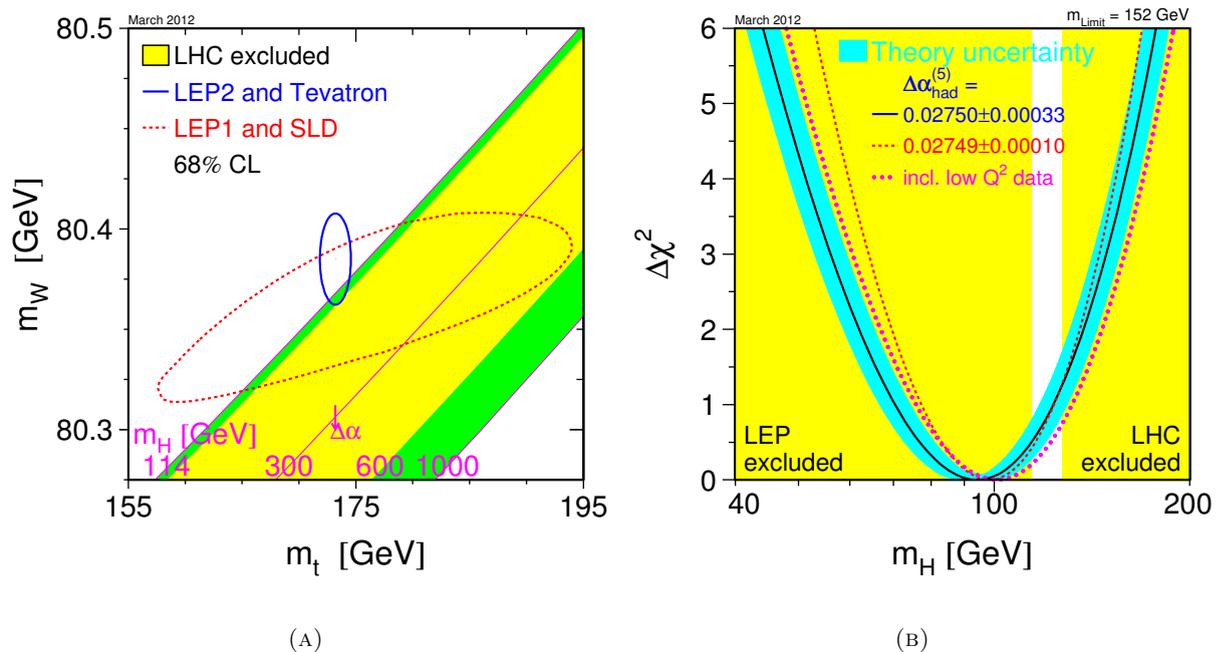


FIGURE 1.10: (A): Compatibility of different Higgs mass models to the  $W$  and top mass from LEP2 and Tevatron. (B): Indirect Higgs mass prediction from combined precision measurements together with LEP and LHC exclusions. Figures shows the status in March 2012[32].

## Unitarity of $WW$ scattering

So far, the Higgs mechanism has been introduced as a way to give the massive vector bosons and fermions masses. Another important feature of the Higgs mechanism becomes apparent when trying to describe the high energy behaviour of  $WW$  scattering without the Higgs boson<sup>12</sup>. A massive gauge boson at rest ( $k^\mu = (m, 0, 0, 0)$ ) will have a linear polarisation vector that is a combination of the three orthogonal unit vectors:

$$(0, 1, 0, 0) \quad (0, 0, 1, 0) \quad (0, 0, 0, 1)$$

<sup>12</sup>The following paraphrases [7, Ch. 21.2], where more details can be found.

If the particle is now boosted along the third component axis, its momentum vector read will read:  $k^\mu = (E, 0, 0, k)$ . The orthogonal projections of the polarisation vector will be left unchanged. The polarisation vectors however have to satisfy:

$$\epsilon^\mu k_\mu = 0 \quad \epsilon^\mu \epsilon_\mu = -1 ,$$

resulting in a longitudinal component that changes in the following way:

$$\epsilon_L^\mu(k) = \left( \frac{k}{m}, 0, 0, \frac{E}{m} \right)$$

From this it can be seen that for large momenta the longitudinal polarisation vector will approximately be parallel to  $k^\mu$ :

$$\epsilon_L^\mu(k) = \frac{k^\mu}{m} + \mathcal{O}\left(\frac{m}{E}\right)$$

The polarisation vector can now be used to examine the differential cross section of  $WW$  scattering in e.g. a lepton-positron collider:

$$\frac{d\sigma}{d\cos\theta}(e^+e^- \rightarrow W^+W^-) \sim \frac{4\pi\alpha^2}{4s} \cdot |\epsilon(k_+) \cdot \epsilon(k_-)|^2 ,$$

where  $k_\pm$  are the momenta of the outgoing  $W$  bosons. The transverse component of this is well-behaved but the longitudinal can be realised is problematic. Using the approximation for the longitudinal component from above, the cross section is given by:

$$\epsilon(k_+) \cdot \epsilon(k_-) \simeq \frac{k_+ \cdot k_-}{m_W^2} \simeq \frac{s}{4m_W^2}$$

for large momenta ( $s \gg m_W^2$ ). This will clearly lead to problems at some point. In fact inserting the relevant values indicate that unitarity will be violated at  $\sqrt{s} \simeq 1$  TeV.

Of course violation of unitarity must be wrong. In fact the calculation above also only include the tree-level diagram. It could be imagined that higher order diagrams will cancel this behaviour, but this leads to different unpleasanties. The higher order corrections would have to be increasingly important in order to restore valid predictions. In other words the electroweak theory would leave the perturbative regime and start becoming "strongly" coupled at the TeV scale.

Another solution could be found if there exist different tree-level diagrams which cancels the divergent behaviour in just such a way that the  $W$  bosons would couple weakly up to very high energies.

Fortunately the introduction of the Higgs boson actually provides such a mechanism.

The Higgs mechanism was accompanied by three unphysical Goldstone bosons. The Goldstone boson equivalence theorem[33] derives that at high energies, the longitudinal  $W$  bosons should have equal cross section to the production of scalar Goldstone bosons. The scalar particles do not have the divergent behaviour however. By giving the  $W$ 's mass through spontaneous symmetry breaking the problematic divergence conveniently cancels.

## The Higgs Discovery

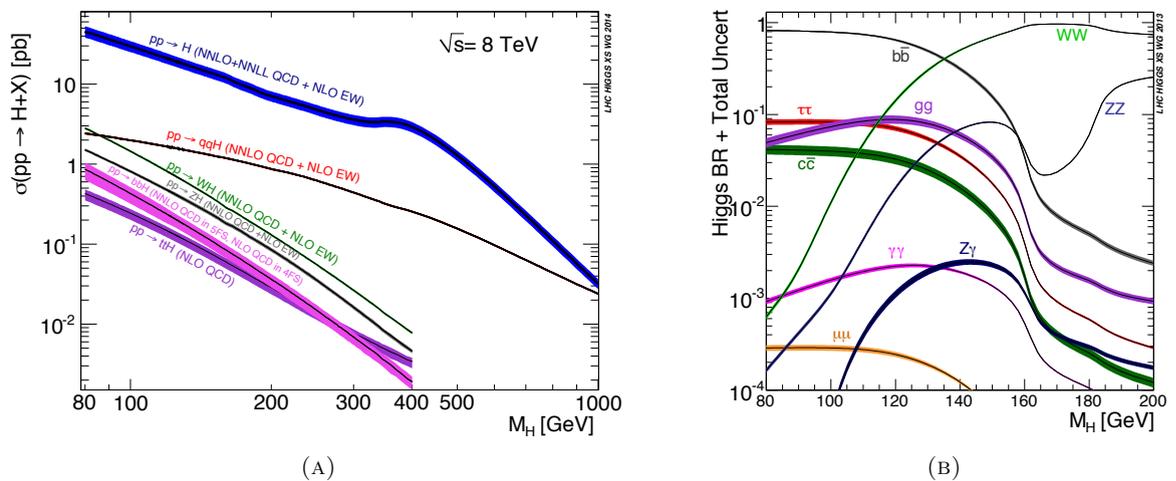


FIGURE 1.11: (A) Cross sections of a Standard Model Higgs Boson produced with proton-proton collisions at 8 TeV. (B) Branching ratios for a Higgs boson with a mass of 125 GeV[34].

The introduction of the Higgs mechanism reveals some particle properties, its spin, parity and branching ratios. The mass is predicted indirectly under assumptions that there does not exist new coupled particles and is required to be below the TeV scale. Unfortunately the lack of knowledge of the mass together with its coupling to other particles make it exceedingly hard to observe directly at hadron colliders. Since accelerators only collide light particles, as a result of the heavier being unstable, the probability of creating a Higgs boson is very small. Figure 1.11a show the dominant Higgs production modes for LHC collisions at 8 TeV, gluon fusion, vector boson fusion and associated production. For a Higgs with a mass of 125 GeV, produced at 8 TeV, these are 19.27 pb (ggF), 1.578 pb (VBF), 0.7046, 0.4153 pb (WH, ZH) and 0.1293 pb (ttH)[34]. To put these numbers into perspective the total LHC cross section is around 0.1 b and the  $Z$  cross section around 30 nb (fiducial), see Fig. 1.5. Besides the minuscule probability to produce a Higgs another challenge emerges from its branching ratios. A hadron collider like the LHC further introduce complexity since the most dominant Higgs decay channels are totally overwhelmed by hadronic backgrounds. Figure 1.11b shows that a Higgs with a mass of 125 GeV will dominantly decay to  $b\bar{b}$ ,  $WW$  and  $gg$ . The  $WW$  diboson branch is again dominated by hadronic decays.

In the end, the most sensitive channels in the low mass region are the  $\gamma\gamma$  and the leptonic final states of  $W^+W^{*-}$  and  $ZZ^*$ . The  $W^+W^{*-}$  final state has the advantage that it by far has the highest branching ratio, but has the disadvantage that the branching ratio for both  $W$  bosons to decay to either an electron or a muon is  $\sim 4\%$ [9]. The leptonic  $W$  decay is besides this made more complicated by the emission of a neutrino.

For the discussion of this thesis it is worth remembering that the discovery is mostly based on observing an excess of events in a specific window of invariant mass. The discovery demonstrated that the new particle state's production cross section and branching ratio to the diboson

channels were in agreement with Standard Model Higgs predictions[35][36]. Despite these being compatible there are still plenty of predictions that needs to be tested to conclude that the observed excess is the result of Standard Model Higgs production.

## Property Measurements

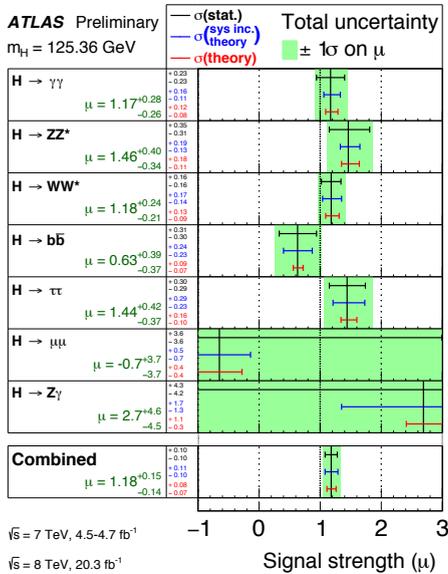


FIGURE 1.12: Summary of Higgs signal strength measurements by the ATLAS collaboration[37]. The signal strength  $\mu$  is defined as the ratio of observed number of events to Standard Model expectations.

decay channels[40] under the assumption that known physics constitute all decay channels. It is finally also of high interest to establish that the boson couples to fermions and that the couplings are in agreement with SM predictions.

The ATLAS and CMS collaborations have established strong evidence that the decay to a pair of  $\tau$  leptons exist, each with a signal more than three standard deviations above background[41][42]. The best fitting signal strength is moreover found to be in agreement with Standard Model expectations, albeit with large uncertainties.

Direct evidence for its coupling to the quark sector is moreover starting to collect<sup>13</sup>. The strongest channel is  $b\bar{b}$ , both because of its large branching ratio (Fig. 1.11b) but also because  $b$  quarks have displaced vertices making them distinguishable over other QCD backgrounds. The LHC Run-1 dataset yielded that the  $VH \rightarrow b\bar{b}$  signal is around  $2\sigma$  over background[43][44]. The measured signal strengths are found to have uncertainties of 80% (ATLAS) and 50%

<sup>13</sup>The decay to  $\gamma\gamma$  is in the Standard Model dominated by a top loop. A diphoton decay rate in agreement with SM can be considered indirect evidence for quark coupling

One of the most important measurements to be performed is the invariant mass determination. The production rates and branching ratios for instance are mass dependent. Although most measurements are limited by statistics at the moment, the invariant mass is ultimately necessary to make precise predictions. The invariant mass compatibility between the high resolution channels  $\gamma\gamma$  and  $ZZ^* \rightarrow 4\ell$  can also be tested. The ATLAS and CMS collaborations recently published the combined measurement[12]. It is found that everything agree within  $1\sigma$ .

The natural decay width as function of mass is also determined in the Standard Model. For a Higgs mass of 125 GeV the on-shell width is 4.1 MeV[34]. A direct measurement of the width by examining the  $m_H$  spectrum would be dominated by the experimental momentum resolution. It is however possible to determine the width in off-shell  $ZZ$ ,  $WW$  decays. The current 95%CL upper limits are found to be  $\simeq 5$  times SM predictions[38]. In an analogous analysis CMS achieved similar results[39]. The width can moreover be determined through the relative branching ratios to different

(CMS) making it difficult to draw final conclusions.

Figure 1.12 summarises the signal strength measurements in ATLAS, where it is seen that all observations are in agreement with SM expectations. It is worth noting that most property measurements so far are limited by statistics making the start of the LHC Run-II an interesting period to study Higgs Physics.

### 1.2.3 Higgs Characterisation with Effective Field Theories

Effective Field Theories (EFT) are used as a method to obtain approximations to the low energy behaviour of undescribed physics. As an example of the approach, consider the charged current weak interaction between two fermions:

$$\frac{g^2}{8} \bar{\psi} \gamma_\mu (1 - \gamma_5) \psi \frac{1}{q^2 - m_w^2} \bar{\psi} \gamma^{\mu\nu} (1 - \gamma_5) \psi ,$$

with  $W$  propagator  $1/(q^2 - m_W^2)$  for momentum transfer  $q$ . Expanding the propagator around  $m_W^2$  gives the series:

$$\frac{1}{q^2 - m_w^2} = -\frac{1}{m_W^2} \left[ 1 + \frac{q^2}{m_W^2} + \left( \frac{q^2}{m_W^2} \right)^2 + \dots \right] ,$$

for  $q^2 < m_W^2$ . If low energy phenomena are examined ( $q \ll m_w$ ) only the first term of the expansion will contribute significantly to the propagator. Inserting the series lowest order into the interaction again:

$$\frac{G_F}{\sqrt{2}} \bar{\psi} \gamma_\mu (1 - \gamma_5) \psi \bar{\psi} \gamma^{\mu\nu} (1 - \gamma_5) \psi , \quad G_F = \frac{\sqrt{2}}{8} \frac{\sqrt{g^2}}{m_W^2} \simeq 1.7 \cdot 10^{-5} \text{ GeV}^{-2} ,$$

revealing the Fermi interaction.

Informally this is similar to a 'zoomed out' version of the original interaction. Taking into account the full EW structure, a process like  $e^- \nu_e \rightarrow e^- \nu_e$  would have a leading order diagram with two incoming fermions connected by a  $W$  propagator to the two outgoing fermions. By 'zooming out' the  $W$  propagator diminishes such that the process look like a four fermion interaction.

Some general properties of the approach can be concluded by dimensional analysis. The action  $S$  is dimensionless. Since  $\int dx^\mu$  has dimension -4, the operators of the Lagrangian density must necessarily be of dimension 4. Here, dimensionality is used to express powers of energy:  $m = [E] \rightarrow d = 1$ ,  $x = [E^{-1}] \rightarrow d = -1$  in natural units<sup>14</sup>.

Dimensional analysis indicate that that the operator expansion must be suppressed by a dimensional quantity. Usually in the context of EFT the scale of new physics  $\Lambda$  is used. In the example above it is the EW scale,  $\Lambda = m_W$ .

Choosing a high energy scale has the advantage that as long as the theory is applied only

<sup>14</sup> $\partial_\mu$  and bosonic fields has  $d = 1$ . Fermionic fields have  $d = 3/2$ . The numbers here are only valid for 4D spacetime.

below the  $\Lambda$  scale the expansion is ordered. More precision can be achieved by adding more terms.

The EFT approach thus gives a recipe for describing new physics phenomena in a compact way if it is assumed that no new physics exist below a certain energy level.

An EFT description of the new bosonic state is derived in “*A framework for Higgs Characterization*”[45]. As above it is assumed that the boson ( $X(J^P)$ ) does not couple to any unknown particles below a scale  $\Lambda$ .  $\Lambda$  is for reference chosen to be 1 TeV, when the model is used in the analysis. The characterisation model divides the Standard Model Lagrangian into two terms:

$$\mathcal{L}_{HC,J} = \mathcal{L}_{SM-H} + \mathcal{L}_J$$

where the first term on the right-hand side refers to the Standard Model without the Higgs boson contribution and the second describes the kinetic and interaction terms of a generic boson  $X(J^P)$  with spin  $J$  and parity  $p$ .

### 1.2.4 Spin-0 Models

When constructing the spin-0 Lagrangian, it is of interest that the Standard Model can easily be reproduced. The model should moreover include all interactions that are generated by gauge-invariant dimension-6 operators above the electroweak scale. The model is expected to include a CP-odd state  $0^-$  as predicted by e.g. many SUSY theories and allow for CP-Mixing between the SM and CP-odd state.

A possible Lagrangian of this type is [45, Comments inserted by author]:

$$\begin{aligned} \mathcal{L}_0^V &= \{c_\alpha \kappa_{SM} [\frac{1}{2} g_{HZZ} Z_\mu Z^\mu + g_{HWW} W_\mu^+ W^{-\mu}] \} && \text{SM ZZ/WW vertices} \\ \hline & -\frac{1}{4} \left[ \cos(\alpha) \kappa_{H\gamma\gamma} g_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + \sin(\alpha) \kappa_{A\gamma\gamma} g_{A\gamma\gamma} A_{\mu\nu} \tilde{A}^{\mu\nu} \right] && \text{H}\gamma\gamma \text{ Vertex} \\ & -\frac{1}{2} \left[ \cos(\alpha) \kappa_{HZ\gamma} g_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + \sin(\alpha) \kappa_{AZ\gamma} g_{AZ\gamma} Z_{\mu\nu} \tilde{A}^{\mu\nu} \right] && \text{HZ}\gamma \text{ Vertex} \\ & -\frac{1}{4} \left[ \cos(\alpha) \kappa_{Hgg} g_{Hgg} G_{\mu\nu}^a G^{a,\mu\nu} + \sin(\alpha) \kappa_{Agg} g_{Agg} G_{\mu\nu}^a \tilde{G}^{a,\mu\nu} \right] && \text{Hgg Vertex} \\ & \quad -\frac{1}{4} \frac{1}{\Lambda} \left[ \cos(\alpha) \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + \sin(\alpha) \kappa_{AZZ} Z_{\mu\nu} \tilde{Z}^{\mu\nu} \right] && \text{BSM HZZ Vertex} \\ & -\frac{1}{2} \frac{1}{\Lambda} \left[ \cos(\alpha) \kappa_{HWW} W_{\mu\nu}^+ W^{-\mu\nu} + \sin(\alpha) \kappa_{AWW} W_{\mu\nu}^+ \tilde{W}^{-\mu\nu} \right] && \text{BSM HWW Vertex} \\ & \quad \overline{\text{CP - Even terms}} \qquad \qquad \qquad \overline{\text{CP - Odd terms}} \\ \hline & -\frac{1}{\Lambda} \cos(\alpha) \left[ \kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial_\mu Z^{\mu\nu} + (\kappa_{H\partial W} W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c.) \right] X_0 && \text{Contact Terms} \end{aligned}$$

The reduced field strength tensors and dual tensor of the Lagrangian are listed in [45, Eq. 2.5,2.6,2.7].

All Standard Model processes are included as a subset of the interaction.

The mixing angle  $\alpha$  describes the degree of CP-mixing. The  $g_i$ s are chosen in such a way that the an effective Standard Model is reproduced if the SM  $\kappa_i$ s are set to 1, the BSM  $\kappa_i$ s are 0 and  $\alpha = 0$ . In the context of the dimension analysis above, it is worthwhile noting that

$d(g_{WW,ZZ}) = 1$  while the others have  $d = -1$ , such that all terms end up with dimension 4, see [45, Tab. 2].

For the analysis of this thesis, only the coupling to  $ZZ$  and  $WW$  are considered. The Lagrangian reduces to:

$$\begin{aligned} \mathcal{L}_0^V = & \left\{ \cos(\alpha)\kappa_{SM} \left[ \frac{1}{2}g_{HZZ}Z_\mu Z^\mu + g_{HWW}W_\mu^+ W^{-\mu} \right] \right. \\ & - \frac{1}{4} \frac{1}{\Lambda} \left[ \cos(\alpha)\kappa_{HZZ}Z_{\mu\nu}Z^{\mu\nu} + \sin(\alpha)\kappa_{AZZ}Z_{\mu\nu}\tilde{Z}^{\mu\nu} \right] \\ & \left. - \frac{1}{2} \frac{1}{\Lambda} \left[ \cos(\alpha)\kappa_{HWW}W_{\mu\nu}^+ W^{-\mu\nu} + \sin(\alpha)\kappa_{AWW}W_{\mu\nu}^+ \tilde{W}^{-\mu\nu} \right] \right\} X_0 \end{aligned} \quad (1.6)$$

There are now three different terms for either  $ZZ$  and  $WW$ . If only one were present at a time, the  $X(J^P)$  addition to the Standard Model Lagrangian correspond to a Standard Model Higgs ( $\kappa_{SM}$ -term), a CP-Even BSM boson ( $\kappa_{HZZ}$ -term) and the CP-odd pseudo scalar ( $\kappa_{AVV}$ -term). These three separate models are denoted  $0^+$ ,  $0_h^+$  and  $0^-$  in the following.

The analysis will consider two separate scenarios for the spin-0 analysis. The first scenario examines if the data can exclude the SM in favour of either  $X(J^P) = 0_h^+$  or  $0^-$ , or conversely if either of the two BSM hypotheses are excluded in favour of  $0^+$ .

In the second scenario the Lagrangian is allowed to contain multiple terms at the same time. The mixed Lagrangian will be parameterised in BSM/SM coupling ratios:

$$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} \quad \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$$

where  $\tilde{\kappa}$  is used as a short hand notation for:

$$\tilde{\kappa}_{AVV} = \frac{1}{4} \frac{\nu}{\Lambda} \kappa_{AVV} \quad \tilde{\kappa}_{HVV} = \frac{1}{4} \frac{\nu}{\Lambda} \kappa_{HVV} \quad (1.7)$$

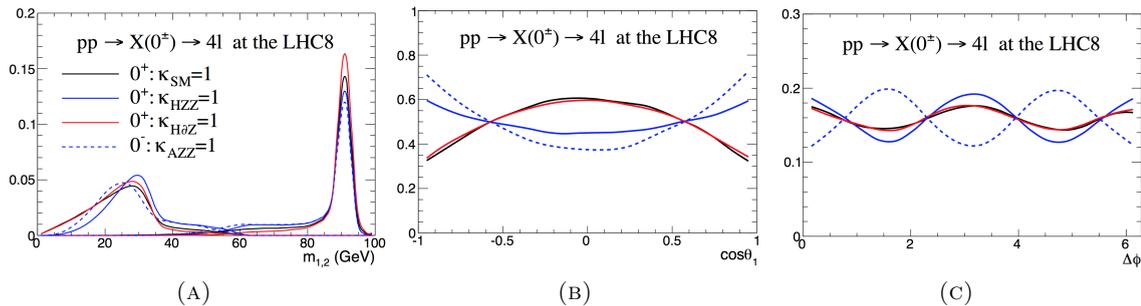
The ratio of couplings are convenient to use instead of a direct measure of the couplings. As was for instance seen in Figure 1.12, both  $WW$  and  $ZZ$  has measured a higher signal strength than SM expectations. If the  $ZZ$  and  $WW$  couplings are also changed in the Lagrangian the rate of observed  $ZZ$  and  $WW$  can also deviate from expectations. It could be feared that by including the signal strength as a model discriminant the best fitting model would be affected by the signal strength instead of the very fine alterations of final state distributions that will be examined.

By using the ratio of couplings instead, an overall normalisation factor can be put outside of the Lagrangian to swallow any deviation, corresponding to letting the signal strength float freely when fitting the models. In the high statistics limits it could be possible to also include the observed rate of events.

Table 1.2 summarises the different models that will be examined. Examples of how final state distributions are affected by different coupling configurations can moreover be seen in Figure 1.8. The different choices of couplings for the spin-0 models in Table 1.2 will not affect the boson's production mechanism. The final state observables that are related to production:  $\Phi_1$  and  $\cos\theta^*$  or equivalently  $p_T$  and  $\eta$  will have the same distribution for the different spin-0 models. The production mechanisms can instead be used to distinguish the boson from different background types in a model independent way.

Model	$J^P$	Coupling value				Coupling Ratio	
		$\kappa_{\text{SM}}$	$\kappa_{AVV}$	$\kappa_{HVV}$	$\alpha$	$\frac{\tilde{\kappa}_{AVV}}{\kappa_{\text{SM}}} \tan \alpha$	$\frac{\tilde{\kappa}_{HVV}}{\kappa_{\text{SM}}}$
Standard Model	$0^+$	1	0	0	0	0	0
BSM Spin-0 CP-odd	$0^-$	0	1	0	$\pi/2$	' $\infty$ '	0
BSM Spin-0 CP-even	$0_h^+$	0	0	1	0	0	' $\infty$ '
Mixed SM, BSM CP-odd	$(0^+, 0^-)$	$\neq 0$	$\neq 0$	0	$\neq 0, \pi/2$	$[-10, 10]$	0
Mixed SM, BSM CP-even	$(0^+, 0_h^+)$	$\neq 0$	0	$\neq 0$	0	0	$[-10, 10]$

TABLE 1.2: Summary table of the spin-0 models that will be studied in this thesis.

FIGURE 1.13: Angular distributions of spin-0 models. Angles are defined as in Figure 1.8. (A):  $m_Z$ , (B):  $\cos \theta_1$ , (C):  $\Delta\phi$ . The distributions are shown for four different coupling configurations[45].

### The $g_i$ parameterisation

A different approach to describe BSM spin-0 models is described in [46][47, Sect. 11.4.2]. Instead of using an EFT approach, the amplitude describing a spin-0 particle  $X$  of arbitrary parity and two spin-1 gauge bosons is used:

$$A(X_{j=0} \rightarrow VV) = \frac{1}{\nu} (g_1 m_V^2 \epsilon_1^* \epsilon_2^* + g_2 f_{\mu\nu}^{*(1)} f^{*(1),\mu\nu} + g_4 f_{\mu\nu}^{*(1)} \tilde{f}^{*(1),\mu\nu})$$

The notation  $g_1$ ,  $g_2$  and  $g_4$  has been adopted to describe the SM, scalar BSM and pseudo-scalar coupling of the HVV scattering amplitude. The Standard model is retrieved if only  $g_1$  is present. The BSM scalar and pseudo scalar contributions are parameterised by  $g_2$  and  $g_4$  respectively.

The approach also include the possibility of having complex couplings, a matter that will be addressed in the outlook section. Complex couplings might arise from light particles loop contributions. The main analysis, using the EFT approach, will however not consider this. It has been demonstrated that the known light particles would not create sizeable complex contributions to the couplings[6].

Considering only the  $WW$  and  $ZZ$  decay channels the different parameterisations can be

related as follows:

$$\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}} = \Re(g_2)/g_1 \quad \frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha = \Re(g_4)/g_1$$

The coupling ratios can moreover be translated to effective cross-section fractions. The effective cross-section contribution from BSM is easier to understand and as an added benefit also allows for a direct comparison to published results[48]. The effective cross-section fraction is defined as follows:

$$f_{g_i} = \frac{|g_i|^2 \sigma_i}{|g_1|^2 \sigma_1 + |g_2|^2 \sigma_2 + |g_4|^2 \sigma_4}, \quad \phi_i = \arg\left(\frac{g_i}{g_1}\right). \quad (1.8)$$

The  $\sigma_i$ s are the cross-section obtained in the presence of only  $g_i$ . If only one BSM contribution is assumed at a time, Eq. 1.8 can be rewritten in the following way:

$$f_{g_i} = \frac{r_{i1}^2}{1 + r_{i1}^2}; \quad (i = 2, 4),$$

such that a direct translation to the 'kappa' framework can be made by choosing:

$$r_{21}^2 = \frac{\sigma_{HV V}}{\sigma_{SM}} \left(\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}\right)^2, \quad \text{and} \quad r_{41}^2 = \frac{\sigma_{AV V}}{\sigma_{SM}} \left(\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}}\right)^2 \tan^2 \alpha.$$

Similar to the  $\sigma_i$ s in Eq. 1.8;  $\sigma_{SM}$ ,  $\sigma_{HV V}$  and  $\sigma_{AV V}$  are the cross sections obtained when only the single corresponding coupling contributes to the Lagrangian. The cross sections are estimated with the MADGRAPH5\_aMC@NLO generator[49] and their ratios found to be  $\sigma_{HV V}/\sigma_{SM} = 0.349$  and  $\sigma_{AV V}/\sigma_{SM} = 0.143$ . Figure 1.14 shows how the cross section contribution evolves as function of coupling ratio.

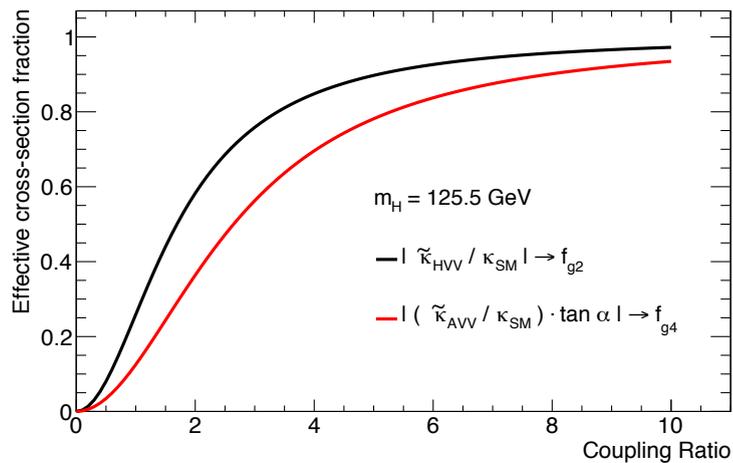


FIGURE 1.14: Conversion between coupling ratios  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$  and cross section fraction  $f_{g_2}$  and  $f_{g_4}$ .

By inverting Eq. 1.8 the coupling ratios can also be described in terms of effective cross section fractions:

$$r_{i1} = \frac{f_{g_i}}{1 - f_{g_1} - f_{g_2}}$$

for  $i = 1, 2$  and  $r_{i1}$  as above. If the assumption of only one BSM coupling at a time is removed the description is somewhat more complex, as the BSM and SM cross section fractions have to add up to one. The  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  coupling for a  $f_{g2} = 0.5$  boson will e.g. depend on where the remaining 50% of the cross section come from. The translation of cross section fractions to the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  coupling ratios is illustrated in Figure 1.15.

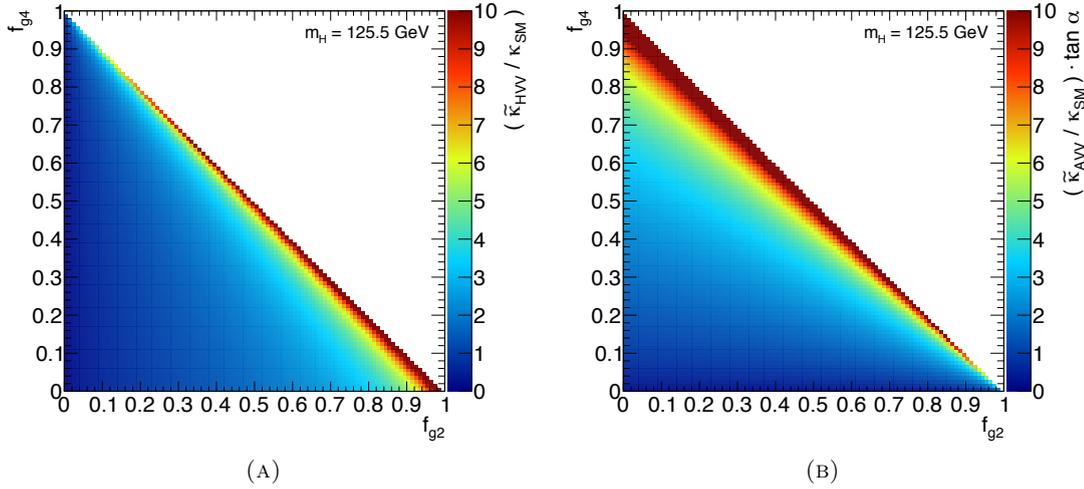


FIGURE 1.15: Conversion between cross section fractions and coupling ratios. (A)  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  as function of  $f_{g2}$  and  $f_{g4}$ . (B)  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  as function of  $f_{g2}$  and  $f_{g4}$ .

### 1.2.5 Spin-1 and Spin-2 Models

Spin-1 models will also be tested in data, using the parameterisation described in [50].

The spin-1 parameterisation that will be analysed follows a similar structure as the ' $g_i$ ' parameterisation from above. The scattering amplitude of a spin-1 bosons interaction with two  $Z$  bosons is given in Eq. 4 of the same source. Both a positive and negative parity spin-1 boson can be described with this approach.

The two parity models will be tested against the Standard Model but it is however important to mention that the spin-1 sector is of smaller physical interest. The Landau-Yang theorem states that a spin-1 particle cannot decay to a pair of massless vector bosons[51][52]. The spin-1 scenario is ruled out by the  $\gamma\gamma$  observation and disfavoured by production rates being compatible with gluon-fusion. If it is (reasonably) assumed that it is the same boson that decays to  $\gamma\gamma$  and  $ZZ$ , the  $ZZ$  spin-1 scenario is also ruled out.

The spin-2 models that are tested in the analysis are also described by EFT. The spin-2 models that will be examined couple to the fermion and boson energy-momentum tensor similar to a Graviton. For a colour, weak and electromagnetic singlet spin-2 resonance, the interaction is

uniquely determined[45]:

$$\mathcal{L}_2 = -\frac{1}{\Lambda} \left[ \sum_{V=Z,W,\gamma,g} \kappa_V \mathcal{T}_{\mu\nu}^V X^{\mu\nu} + \sum_{f=q,l} \kappa_f \mathcal{T}_{\mu\nu}^f X^{\mu\nu} \right], \quad (1.9)$$

where a coupling is introduced for each boson and fermion. Some assumptions have to be made on the couplings to reduce the number of models to a feasible size.

The couplings to the  $\gamma$ ,  $W$  and  $Z$  vector bosons are assumed to be 1. The choice is not of particular interest since it is known that the boson couplings exist due to the observed decays. Moreover, the relative magnitude is of minor relevance since the analysed final states do not mix. The signal strengths for each channel are allowed to float for similar reasons to what was described in the spin-0 discussion, making the analysis independent of the particular choice of these coupling values.

The choice of fermion and gluon couplings is of more importance. This will affect how large a portion of the production is due to gluons or quarks. Intuitively the relative magnitude of quark or gluon production will have an effect on the resonance  $p_T$  spectrum, see e.g. Figure 1.6.

Three models are considered in the analysis. The universal couplings model assumes that quark and gluon couplings are equal. A model where the boson does not couple to quarks and one where the quark coupling is twice that of the gluon coupling are also considered.

The non-universal coupling models predicts a hard  $p_T$  spectrum. The extremum of  $\kappa_q = 2\kappa_g$  is chosen since a harder  $p_T$  spectrum than what this will produce would have shown up in fiducial cross section measurements[5]. The EW spin-2 production moreover has been estimated to be negligible ( $\sigma_{EW}/\sigma_{QCD} \simeq 3 \cdot 10^{-4}$ [5]) and will not be considered.

A transverse momentum cutoff is finally imposed in order to prevent the analyses from using discriminative power from an energy regime where an EFT with  $\Lambda = 1$  GeV is invalid. The cutoff is chosen as 300 GeV and the results are validated against a conservative cutoff of 125 GeV. Table 1.3 summarises the spin-2 models that will be considered. Final state distributions for spin-2 models with different quark and gluon coupling configurations can be found in Figure 1.16. A conversion table between the scattering amplitude parameterisation and the effective field theory parameterisation for both spin-0 and spin-2 models is given in [45, Tab. 3].

Values of spin-2 quark and gluon couplings		$p_T^X$ selections (GeV)	
$\kappa_q = \kappa_g$	Universal couplings	–	–
$\kappa_q = 0$	Low light-quark fraction	< 300	< 125
$\kappa_q = 2\kappa_g$	Low gluon fraction	< 300	< 125

TABLE 1.3: Spin-2 models studied in this thesis. Table courtesy of [5].

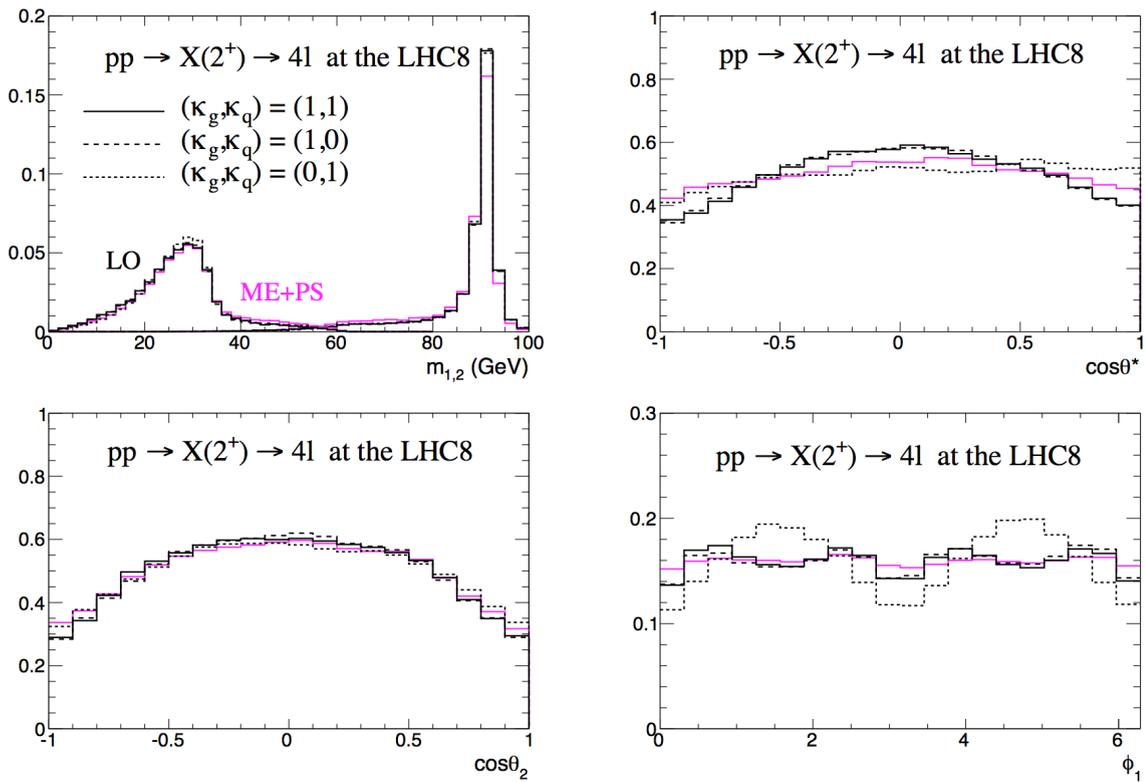


FIGURE 1.16: Normalised final state distributions for spin-2 models with different coupling configurations. ME+PS indicate the presence of extra QCD radiation compared to leading order[45].

### 1.3 Monte Carlo Simulation and Validation

Several Monte Carlo generators are used to get the best possible description of signal and background models.

#### Spin-0 models

As described in Section 1.2.4 only  $WW$  and  $ZZ$  BSM contributions are considered to the spin-0 Lagrangian. The Higgs is primarily created through gluon-fusion, meaning that the production is decoupled from the assumed model. The production of the models that are examined can adequately be described by tools that have been developed to simulate Standard Model Higgs production. The analysis uses the POWHEG-Box to simulate the hard scattering of  $gg \rightarrow H$  at next to leading order[53].

The boson decay is described with JHU[46]. The generator is able to describe the decay topology for the various models that will be examined; SM, BSM-scalar, pseudo-scalar and mixed models thereof. The JHU generator moreover describes final state interference in the  $4e$  and  $4\mu$  final states.

The generators are interfaced with PYTHIA6 to describe initial and final state radiation, hadronisation and multi-parton interactions[54]. The current analysis only considers gluon-fusion and not the rarer vector boson fusion and associated production.

The Higgs  $p_T$  is after production reweighted to the Standard Model next-to-next-to-leading-order and next-to-next-to-leading-logarithms predictions given by the HRES2.1 program[55][56].

#### Spin-2 models

The spin-2 models predict different couplings to gluons and quarks than the Standard Model, meaning that the POWHEG-BOX can no longer be utilised. Instead, both the production and decay of spin-2 models are described by the MADGRAPH5\_aMC@NLO generator[49]. Like above, initial and final state radiation, hadronisation and multi-parton interactions are described by interfacing with PYTHIA6.

#### $ZZ^*$ -continuum

The  $ZZ^*$ -continuum constitute the largest background of the analysis. The  $qq \rightarrow ZZ^*$  production is again described at next-to-leading order precision using the POWHEG-BOX. The smaller  $gg \rightarrow ZZ^*$  background processes is generated with GG2ZZ[57]. The continuum is generated in a window around the signal mass region:  $100 \text{ GeV} < m_{ZZ} < 150 \text{ GeV}$ . Initial and final state radiation, hadronisation and multi-parton interactions are described by interfacing with PYTHIA6.

In all cases only decays to the four final states;  $4e$ ,  $4\mu$ ,  $2e2\mu$  and  $2\mu2e$  are generated. All samples are moreover generated with CT10 parton distribution functions[25].

A summary of the Monte Carlo samples and generators that are used in the analysis is given in Table 1.4. The JHU generator that is used to produce the BSM spin-0 models uses the  $g_i$  parameterisation. The generator values are therefore stated in this scheme. Complex couplings

are also present in several of the samples but will mostly not be considered. Throughout the main analysis, the complex contributions to the couplings are removed with matrix element based reweighting as described below.

<b>Signal Processes</b>	
Process type	Generator
$g_1 = 1, g_2 = 0, g_4 = 0$ (SM)	POWHEG-Box JHU PYTHIA
$g_1 = 0, g_2 = 1, g_4 = 0$ ( $0_h^+$ )	POWHEG-Box JHU PYTHIA
$g_1 = 0, g_2 = 0, g_4 = 1$ ( $0^-$ )	POWHEG-Box JHU PYTHIA
$g_1 = 1, g_2 = 1 + i, g_4 = 1 + i$	POWHEG-Box JHU PYTHIA
$g_1 = 1, g_2 = 0, g_4 = 2 + 2i$	POWHEG-Box JHU PYTHIA
$g_1 = 1, g_2 = 1 + i, g_4 = 0$	POWHEG-Box JHU PYTHIA
$g_1 = 0, g_2 = 1, g_4 = 1$	POWHEG-Box JHU PYTHIA
$g_1 = 1, g_2 = 1, g_4 = 1$	POWHEG-Box JHU PYTHIA
Spin-2, Universal couplings	MADGRAPH5_aMC@NLO PYTHIA
Spin-2, Low quark fraction	MADGRAPH5_aMC@NLO PYTHIA
Spin-2, Low gluon fraction	MADGRAPH5_aMC@NLO PYTHIA

<b>Background processes</b>	
Process type	Generator
$gg \rightarrow ZZ^*$	GG2ZZ PYTHIA
$q\bar{q} \rightarrow ZZ^*$	POWHEG-Box PYTHIA

TABLE 1.4: Monte Carlo samples and generators. All processes are only generated in the four lepton final state at both 7 TeV and 8 TeV. Other backgrounds are estimated with data driven techniques.

### 1.3.1 Matrix Element Based Reweighting

The tensor structure fit that will be described in Section 3.4 examines a continuum of coupling ratios. It will not be possible to generate fully simulated Monte Carlo samples at all values that are examined, or even at enough points to do a reasonable interpolation.

Matrix Element based reweighting is instead used to get fully simulated Monte Carlo descriptions at closely located coupling ratios. The method works as follows:

First, a large Monte Carlo sample is generated at some coupling configuration, here  $g_1 = 1$ ,  $g_2 = 1 + i$ ,  $g_4 = 1 + i$ . The sample will in the following be referred to as the ‘‘Source’’ sample. To get a description at a ‘‘Target’’ coupling configuration each event is weighted in such a way that they combined will reproduce the correct distributions. The matrix element event weight is defined as:

$$w = \frac{|ME|^2(\theta_{\text{Target}})}{|ME|^2(\theta_{\text{Source}})} \quad (1.10)$$

The Matrix elements are calculated with JHUGENMELA.v4.2.1[46] which also uses the scattering amplitude parameterisation (‘‘ $g_i$ ’’).

The matrix element will take into account the full decay information for each event. It does however not contain any information on the production mechanisms of the boson. The reweighting method can therefore only be used to reproduce decay topologies. The BSM tensor structure models that will be examined in the analysis also only differ in decay topologies as mentioned in Section 1.2.4 meaning that no loss of generality is suffered from this.

The weight is calculated using the event information at generator level while the observable distributions are calculated after full detector simulation. The reweighting procedure will moreover reduce the statistical power of a sample such that the generated number of events will have to be significantly higher than for backgrounds and closure samples. A sample of 3000000 events are generated.

Matrix element based reweighting is used to get a description at a grid of  $\Re(g_2)/g_1$  and  $\Re(g_4)/g_1$  couplings. A separation of 0.25 between each point from -10 to 10 is used (81 points in either dimension). This way all integer values of couplings will lie at centres of bins, while having a symmetric range around zero.

Two tests are presented as examples of the validation tests that have been performed of the reweighting procedure. The Source sample is reweighted first to describe the Standard Model and second to describe a BSM boson where  $g_1 = 1, g_2 = 1, g_4 = 1$ .

Samples at these coupling values have also been generated directly. The predicted distributions from reweighting can as a result be compared to the prediction from a direct Monte Carlo generation.

The comparisons can be found in Figures 1.17 and 1.18, where the five different final state observables are shown. Beside any visual difference, the distributions are compared in two different ways. First the bin-by-bin difference in units of the statistical uncertainty ('the pulls') is calculated. The pull distributions will follow a unit Gaussian if the two distributions are sampled from the same parent distribution. The pull distribution will however not necessarily tell if there are systematic tendencies, but only indicate differences in the average behaviour. It could e.g. be imagined that the reweighting would underestimate in the left side and overestimate in the right side of some distribution. To rule out this possibility the probability of obtaining the Kolmogorov-Smirnov (KS) statistic of the comparison is also calculated. The distribution of probabilities is uniform if the two distributions are from the same parent distribution, such that small values indicate poor agreement. The reweighting procedure is concluded to give a good description after considering the figures.

It is worthwhile pointing out that the reweighting procedure will not produce statistically independent samples however. Large statistical fluctuations present at one coupling configuration will also be present in others.

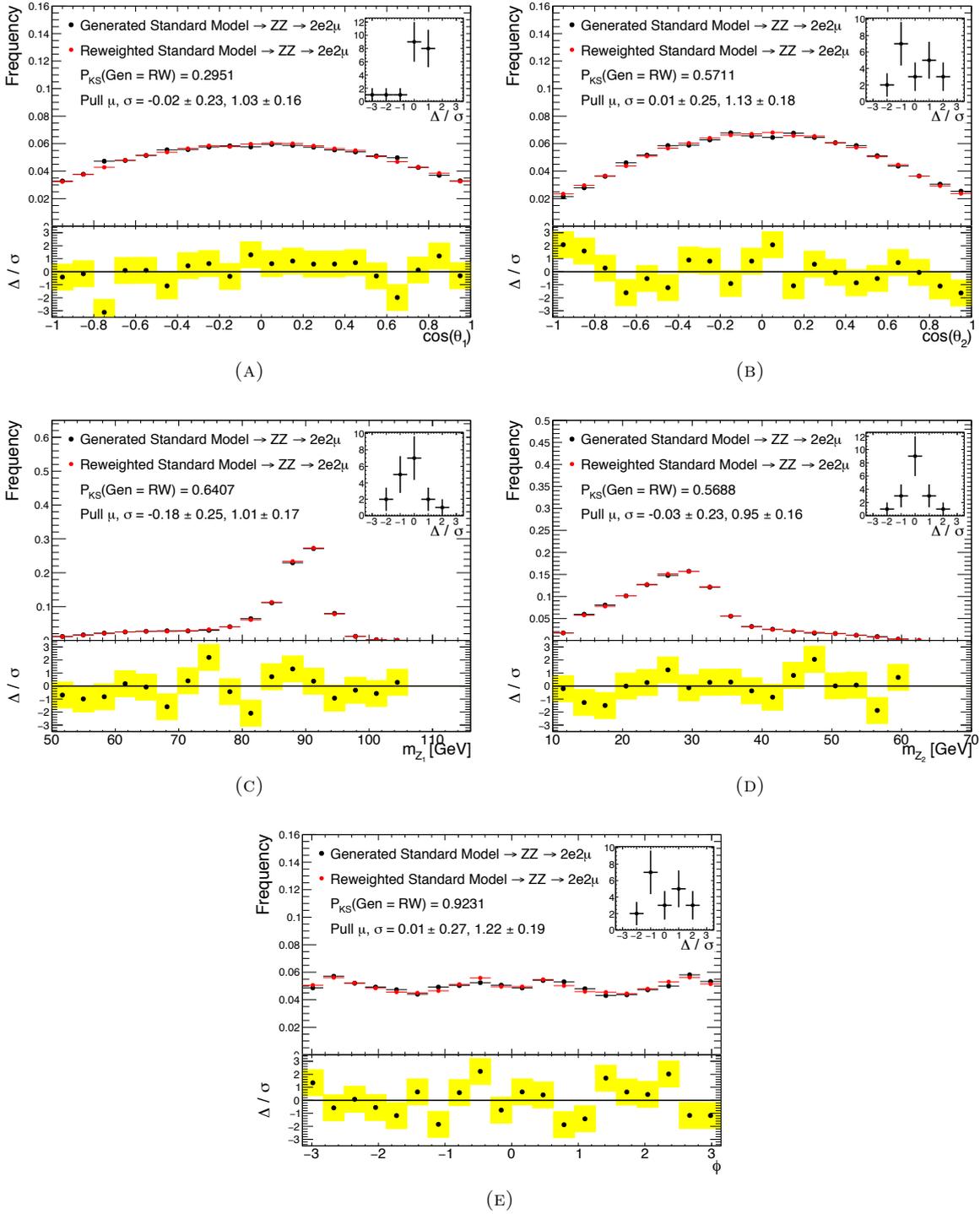


FIGURE 1.17: Reweighting validation. The large MC sample with coupling configuration ( $g_1 = 1, g_2 = 1 + i, g_4 = 1 + i$ ) is reweighted to the Standard Model. The distributions from reweighting (red) are compared to Standard Model MC (black).  $\cos\theta_1$  (A),  $\cos\theta_2$  (B),  $m_{12}$  (C),  $m_{34}$  (D),  $\phi$  (E). The observable distributions are shown in the centre of each plot, bin pulls in the bottom and pull distribution in top right.

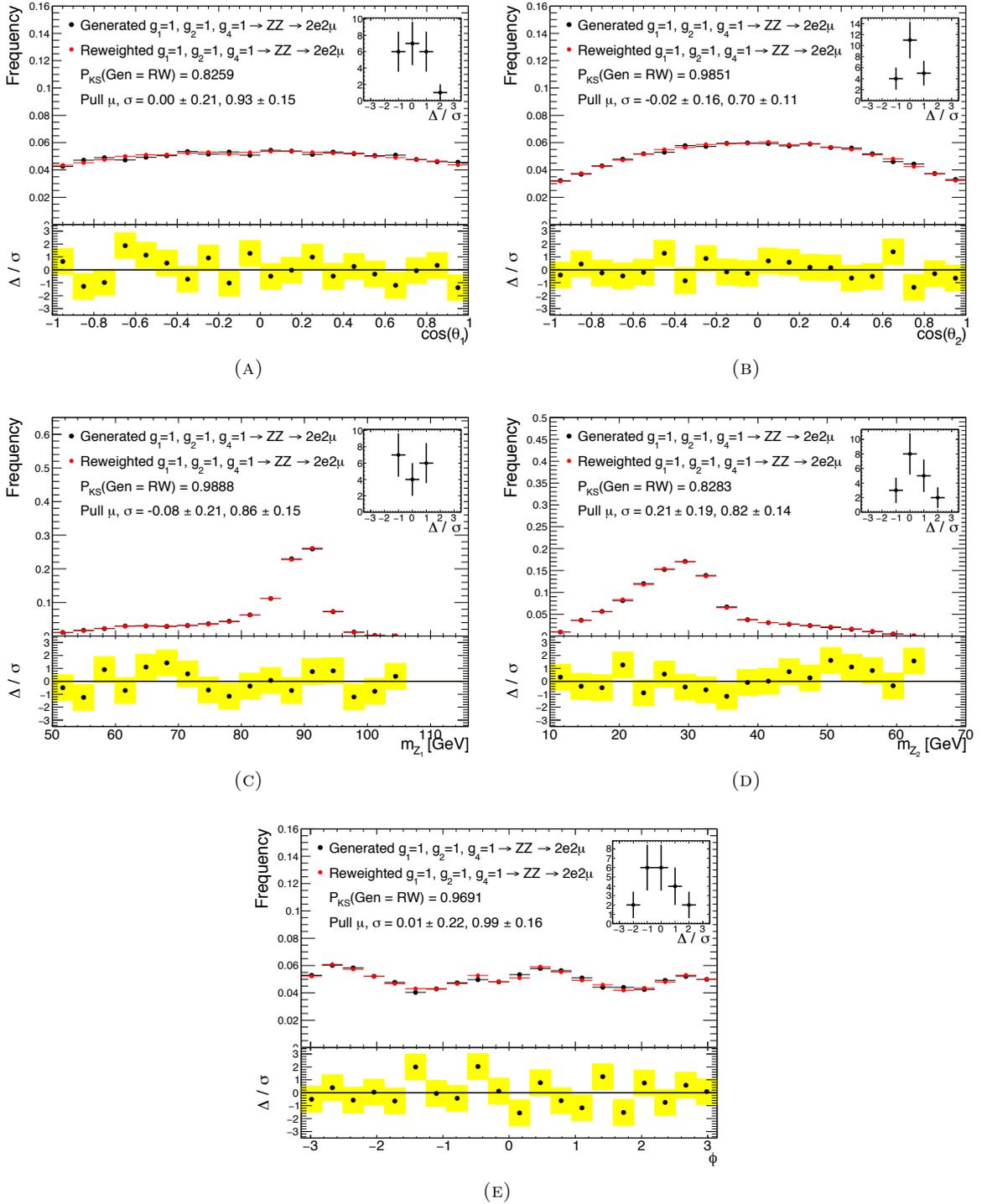


FIGURE 1.18: Reweighting validation. The large MC sample with coupling configuration  $(g_1 = 1, g_2 = 1 + i, g_4 = 1 + i)$  is reweighted to a BSM boson with  $(g_1 = 1, g_2 = 1, g_4 = 1)$ . The distributions from reweighting (red) are compared to the test BSM MC (black).  $\cos\theta_1$  (A),  $\cos\theta_2$  (B),  $m_{12}$  (C),  $m_{34}$  (D),  $\phi$  (E). The Observable Distributions Are shown in the centre of each plot, bin pulls in the bottom and pull distribution in top right.

## 1.4 Beyond the Standard Model

A few comments are finally given on the limitations of the Standard Model before the text turns to the experiment chapter<sup>15</sup>.

The problems of the Standard Model can in general be divided into unexplained phenomena and unnatural predictions.

Starting with the unexplained phenomena, it is natural to ask why the Standard Model has so many free parameters and features that are simply introduced by hand. The SM, without neutrino oscillations and Higgs mechanism, has 19 free parameters[58]. These include three gauge couplings, six quark masses and three charged lepton masses. One can also ask why the charged fermion masses are so vastly different, spanning nearly 6 orders of magnitude ( $m_e = 0.511 \text{ MeV}$ ,  $m_t = 173 \text{ GeV}$ [9]).

When neutrino oscillations are included it is furthermore necessary to introduce another three masses, three mixing angles and three CP-violating phases. The fermions were also postulated to exist in three generations without further explanation. It seems unsatisfactory that a fundamental theory would have this much freedom.

Another type of unexplained phenomena can be seen on larger scales. The Standard Model has a natural particle anti-particle symmetry. Particles and anti-particles appear in the model in pairs with similar properties. Another natural question to ask is why there is such a large preference for matter? The Standard Model does not contain strong enough CP-violation to be able to explain why there would be this asymmetry if the universe contained equal amounts of matter and anti-matter at the time of the Big Bang.

Cosmological observations moreover indicate that the visible amount of matter is only able to account for around 5% of the total energy in the universe. The remaining 95% is a combination of dark matter ( $\sim 25\%$ ) and dark energy ( $\sim 70\%$ [60]). The necessity of dark matter is e.g. seen in the angular rotation of galaxies, which would have been much smaller if only gravity of the visible matter pulled it together[61]. Furthermore, the so called ‘‘Bullet Cluster’’ has been observed, where galaxy clusters collide on an astronomical scale[62]. It was observed that the visible parts interacted with each other and were slowed down, while gravitational lensing indicated that the majority of the mass passed right through, leaving a discrepancy between the visible centre of mass and the gravitational centre of mass.

Another undescribed phenomenon that is evident at macroscopic scales is gravity. Gravity is extremely weak compared to the other forces at particle level. This makes it impossible to directly measure in today’s collider experiments. Comparing the energy scales of gravity  $M_{Planck}$  to e.g. the weak scale  $\nu = 246 \text{ GeV}$  also raises the question why gravity is so different from the other forces.

At a more technical level, if gravity were to be introduced as a quantum field theory with

<sup>15</sup>For more elaboration, see [58], [7, Ch.22,5], [59], which is largely the basis for the description here.

Feynman diagrams, it would have coupling constant with inverse energy dimensionality making it non-renormalisable[7].

The Hierarchy problem is on the other hand found in the category of unnatural predictions. The problem also relates to the difference in the weak and Planck scales. It can be demonstrated that in the Standard Model, the quantum corrections to the Higgs mass  $m_H$  and consequently to  $m_W$  are quadratically divergent[58]:

$$\delta m_H^2 \simeq \mathcal{O}\left(\frac{\alpha}{\pi}\right) \Lambda^2,$$

where  $\Lambda$  is the scale at which new physics appear. If it is imagined that the Standard Model would hold all the way up to the Planck scale, the radiative corrections would be 36 orders of magnitude larger than the physical value of  $m_H$  itself.

In principle there is nothing wrong with this from a mathematical point of view. A tree level value of  $m_H$  can be introduced that is almost the exact opposite (1 in  $10^{36}$ ) of the correction such that the measured mass will come out correctly. This kind of fine tuning seems highly arbitrary and unsatisfactory.

One possible solution to the Hierarchy problem is to introduce a particle symmetry such that each fermion has a bosonic partner and vice versa called 'Super Symmetry'. The radiative corrections for the Higgs mass for fermions and bosons come with opposite sign. In the presence of this kind of symmetry the correction instead reads:

$$\delta m_H^2 \simeq \mathcal{O}\left(\frac{\alpha}{\pi}\right) (m_B^2 - m_F^2)$$

If the super symmetric partners are found at the TeV scale the unnaturalness would naturally disappear[58]. Another motivating feature of Super Symmetry is that, if the lightest super-symmetric particle is stable, it would be a natural candidate for all the dark matter that is observed.

One incarnation of SUSY, the Minimal Supersymmetric Standard Model, contain the spectrum of a two-Higgs Double Model<sup>16</sup>. The model predicts the existence of five physical Higgs bosons after spontaneous breaking of the electroweak symmetry: A charged Higgs pair ( $H^\pm$ ), one CP-odd scalar ( $A$ ) and two CP-even states ( $H$  and  $h$ ). An appealing feature of this is that the model could provide strong enough CP-violation to give an explanation of the cosmic matter-antimatter asymmetry.

Super Symmetry is one of several possibilities of Standard Model extensions that preserve naturalness. The modifications that aim to restore naturalness to the Higgs sector can be divided into two types: The Higgs is fundamental and some other mechanism removes the need for finetuning or the observed Higgs is in fact a composite particle.

In composite models, the Higgs boson is described as a bound state of some new dynamic which become strong at the weak scale. The 'Little Higgs models' are examples of such a description[63][64]. In them the Higgs doublet is identified as a Nambu-Goldstone boson analogous to how the pions appear in QCD.

<sup>16</sup> A comprehensive description of BSM in the context of Higgs physics is given in [9, Ch. 11, Sect. 5].



## Chapter 2

# Experiment

The models that will be examined have now been established and need to be tested. The following chapter will focus on the experimental setup.

First an overview of the Large Hadron Collider is given, which is used to accelerate and collide protons. The chapter will continue with a description of the ATLAS detector with mentions of each of its subdetector constituents. The section will focus on how particles are identified and how their kinematics are measured. Finally, details on lepton identification and reconstruction will be given in context of the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis.

## 2.1 The Large Hadron Collider

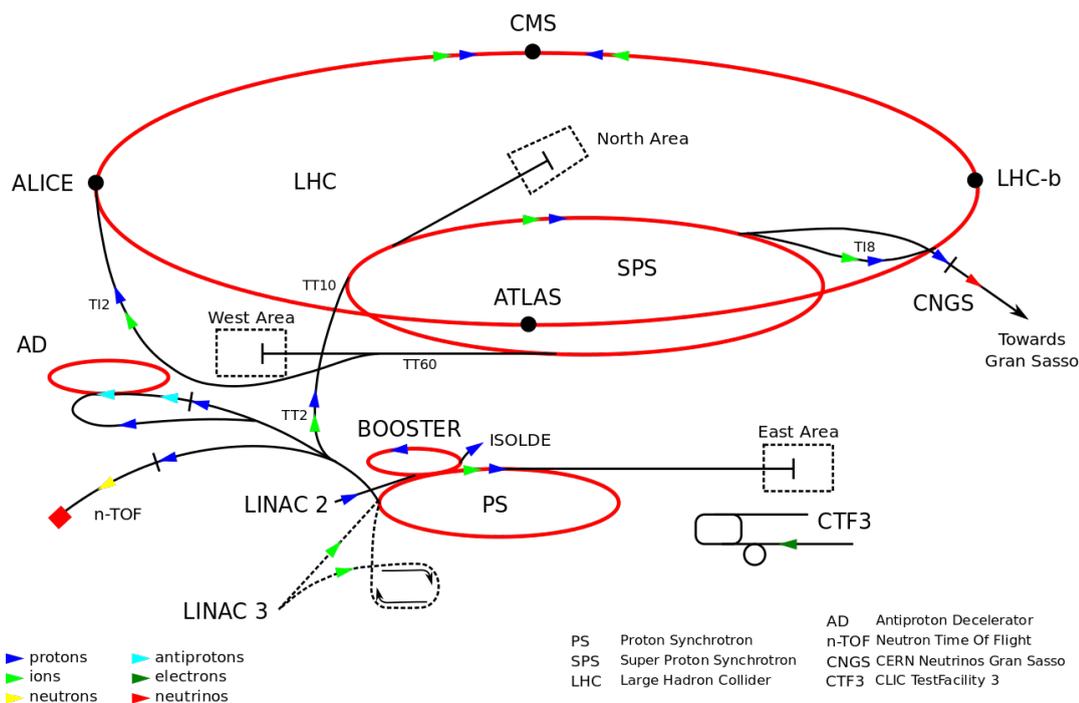


FIGURE 2.1: Map of the accelerator and detector complex found at CERN[65].

The Large Hadron Collider (LHC) is the world's biggest particle accelerator. LHC is the final step of the CERN (Organisation Européenne pour la Recherche Nucléaire) accelerator complex, situated in Switzerland and France. The accelerating ring of the LHC has a 27 km circumference.

Figure 2.1 show the CERN accelerator complex and the route protons take before colliding at e.g. the ATLAS site. After having been separated from its electron in the hydrogen state, the protons are in sequence accelerated through the LINAC-2, PSB<sup>1</sup>, PS, SPS and LHC accelerators. Each of the accelerators before LHC serves in this context to accelerate the protons to the energy level where the next accelerator starts functioning. The energy reached by each step is: 50 MeV (LINAC-2), 1.4 GeV (PSB), 25 GeV (PS) and 450 GeV (SPS). The LHC is

<sup>1</sup>The Proton Synchrotron Booster is labelled as BOOSTER in the figure.

finally able to boost the individual protons from 450 GeV up to a maximum of 7 TeV<sup>2</sup>.

The accelerator has not accelerated protons to its maximal energy yet, but in 2010 ran with a centre of mass energy of 900 GeV, 7 TeV in 2011 and 8 TeV in 2012<sup>3</sup>. The first tests at 13 TeV have started in the summer 2015. LHC can furthermore perform heavy ion collisions using lead. In lead collisions the design energy per nucleon is 2.76 TeV<sup>4</sup>.

Four large experiments are found at the LHC ring: ATLAS, ALICE, CMS and LHCb. ATLAS and CMS are general purpose detectors optimised for proton collisions. The ALICE detector on the other hand focuses mostly on Pb-Pb collisions. It is designed to reconstruct a higher track multiplicity at higher resolution compared to the ATLAS and CMS experiments. Higher resolution comes at the cost of a slower response. The LHCb experiment is mainly build close to the beam-line in one direction from the collision point. It is able to reconstruct tracks in the very forward region, between 10 to 300 mRad in the non-bending plane[66]. The LHCb design is among other optimal for heavy flavour physics.

The focus in the coming sections will be on the ATLAS detector, but first some considerations on the accelerator design are given.

Two of the main design features of an accelerator are its centre of mass energy ( $\sqrt{s}$ ) and its instantaneous Luminosity ( $\mathcal{L}$ ).

The cross section of Higgs production in proton collisions for instance increases rapidly with centre of mass energy. The gluon and vector boson fusion cross sections increased by a factor of 1.3 when going from 7 TeV to 8 TeV while an increase from 8 TeV to 14 TeV will enhance the Higgs cross section by another 2.7[34]. Moreover, searches for heavy BSM particles require that the centre of mass energy is larger than its mass. These two examples alone serve as good motivation to want a high centre of mass energy.

LHC is mainly built of two types of components, bending and focusing magnets control the beam shape in the transverse direction while radio frequency cavities accelerate and limit the longitudinal proton bunch dispersion. The LHC consist of eight arcs with eight straight sections in between. The eight straight sections contain the experiments, beam dumps, cleaning and RF cavities. The arcs consist of focusing and bending magnets.

The basic mechanics of the magnetic system can be understood by investigating the Lorentz force different types of magnets exert on charged particles. Imagining a charged particle travelling in the  $z$  direction. If the particle hits a magnetic field that only points in the  $y$  direction, it will experience a force that pushes it into a circular path. This type of field can be created locally with a dipole magnet. The LHC uses 1232 14.3 m dipole magnets with a field strength of 8.3 T to achieve this[67]. The maximal proton energy the bending magnets are able to maintain can be estimated from these numbers. A charged particle (charge  $q$ ) traversing a magnetic field (strength  $B$ ) in a circulation motion (radius  $\rho$ ) will have momentum given by[68,

<sup>2</sup>Increasing the energy by more than an order of magnitude constitute here a gain of  $\sim 650$  m/s in speed, 2.7 m/s from  $c$ .

<sup>3</sup>Colliding particles travelling in each direction with ( $E_p$ ) gives a centre of mass energy  $\sqrt{s} = 2E_p$

<sup>4</sup>7 TeV scaled by the number of charged to total number of nucleons in lead (here the stable 209 ion):  $7 \text{ TeV} \frac{82p}{209(n+p)}$

Eq. 4.1]:

$$p = qB\rho \quad \Rightarrow \quad p \simeq 0.2998 \frac{\text{GeV}}{\text{T} \cdot \text{m}} B \cdot \rho, \quad (2.1)$$

where the latter assumes unit charge. The expected energy emerges by inserting the size, number and strength of the LHC dipole magnets from above:

$$p = \frac{0.2998}{2 \cdot \pi} \frac{\text{GeV}}{\text{T} \cdot \text{m}} \cdot 8.3 \text{ T} \cdot 14.3 \frac{\text{m}}{\text{dipole}} \cdot 1232 \text{ dipoles} = 7.0 \text{ TeV} \quad (2.2)$$

Beam focusing is necessary since the protons will repel each other and consequently not be exactly aligned. Consider a quadrupole magnet with a field of the form:

$$B = (B \cdot x, B \cdot y, 0)$$

Assuming a short interaction with the magnet such that the particle continues to travel parallel to  $z$  while in the magnetic field gives equations of motion[68, Eq. 4.17]:

$$\frac{d^2x}{dz^2} \propto -x \quad \frac{d^2y}{dz^2} \propto y$$

The acceleration in the transverse  $x$  direction is proportional to negative  $x$ , meaning it has a sinusoidal solution. The  $y$  acceleration is on the other hand proportional to positive  $y$  giving it an exponentially increasing solution. This magnet type can then be used to focus in the  $x$  direction at the cost of defocusing in the  $y$  direction. A focused beam can be obtained by consecutive focusing in  $x$  and  $y$ .

The focusing will not be perfect and higher order beam corrections are needed in real life. The LHC uses cells consisting of dipoles to bend particles towards the LHC centre, quadrupoles to focus the beam and higher multipoles to correct the focusing.

The radio frequency cavities (RF) serves to accelerate and focus the protons in the direction of flight. When protons traverse the cavity, they will enter an oscillating electric field. The oscillations are matched to the proton speed such that each bunch will constantly have a decrease in the EM potential in the forward direction. The dispersion is reduced since protons behind the bunch will feel a stronger field than the average proton in the bunch. The ones in front will likewise feel a weaker field than average.

For a 20 min ramp the protons will on average receive a push of 485 keV per revolution. Contrarily the energy loss due to synchrotron radiation is around 7 keV at 7 TeV[67, Ch. 6].

The other mentioned design criteria, instantaneous Luminosity, is defined as the conversion factor between how probably an event is ( $\sigma$ ) and how often it occurs ( $R$ ):

$$R = \mathcal{L} \cdot \sigma, \quad \mathcal{L} = \frac{N^2 f}{A}, \quad (2.3)$$

The rightmost expression assumes two equal size proton bunches with population  $N$ , colliding at frequency  $f$  with an effective beam-beam cross section  $A$ .

A naive estimation of the beam-beam cross section can give a good approximation. Assuming the beams have a RMS transverse size of  $\sigma_{x,y} = 16.6 \mu\text{m}$ [67, Ch. 1] the instantaneous luminosity will be[9, Ch. 29]:

$$\mathcal{L} \simeq \frac{N^2}{4\pi \cdot t \cdot \sigma_x \cdot \sigma_y} = \frac{(10^{11})^2}{4\pi \cdot 50 \text{ ns} \cdot (16.7)^2 \mu\text{m}^2} = 7.5 \cdot 10^{33} \text{ m}^{-1}\text{cm}^{-2} \quad (2.4)$$

In good agreement with the actual LHC instantaneous luminosity, see e.g. Figure 2.2a.

Usually analyses will use datasets which have been collected over a period of time. The integrated luminosity is the equivalent conversion between observed number of events and production cross section:

$$L = \int dt \mathcal{L} ,$$

with units  $\text{m}^{-2}$ . The length of each run and how many of those are performed determines the accumulated integrated luminosity. A run is typically  $\sim 10$  hours. By then the bunches are depleted due to beam-beam and gas interactions. The total integrated luminosity from the data collection of 2010, 2011 and 2012 is shown in Figure 2.2c.

In order to avoid confusion by high powers of ten it is customary to use units of 'barns', where  $1 \text{ b} = 10^{-24} \text{ cm}^{-2}$ . The rise in integrated luminosity between 2010 and 2012 can be ascribed to an increase in the average number of interactions per bunch crossing,  $\langle \mu \rangle$ , see Fig. 2.2b.

It is apparent from the definition of the integrated luminosity that it is an essential parameter in cross section measurements and searches. Luminosity uncertainties will directly translate to uncertainties in these types of analyses. A precise luminosity measurement is however difficult to perform.

The simplest type of measurement would be to compare the observed number of events and the cross section of a well known process. This type of analysis is, as an example performed on the Drell-Yan cross section for the 7 TeV data sample with  $m_{\ell\ell} > 116 \text{ GeV}$ [70]. The systematic error is estimated to be greater than 4% in all categories. Both theoretical and experimental uncertainties are considered a systematic uncertainty.

Another method of luminosity measurements is to analyse the elastic scattering of protons during specialised runs. The scattered protons are detected at low angles far from the collision point. The relative luminosity uncertainty is found to be  $\delta\mathcal{L}/\mathcal{L} = 1.8\%$  for the same dataset as the Drell-Yan cross section is measured for[71].

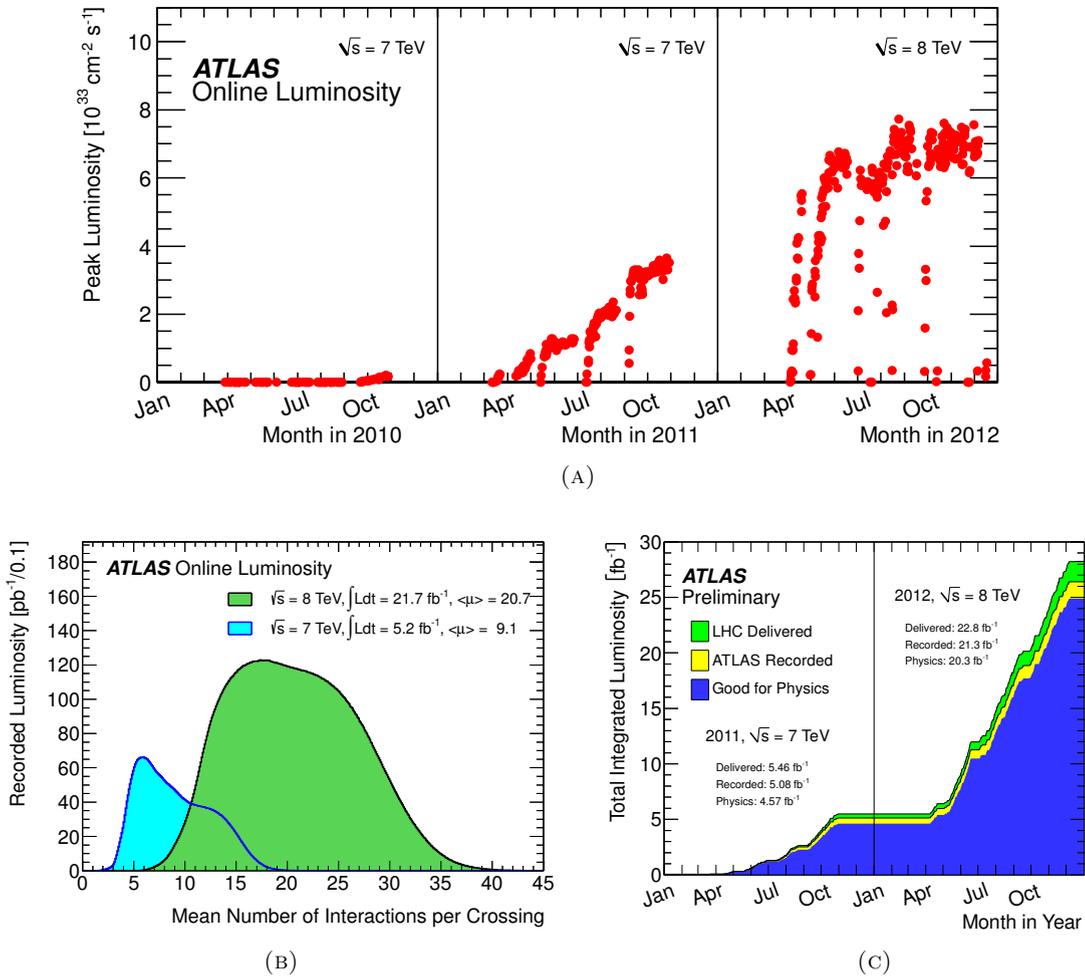


FIGURE 2.2: (A): Recorded instantaneous luminosity in ATLAS in 2010-2012. The increase in collisions per time can be explained by the increase in interactions per bunch crossing seen in (B), where the distributions for 2011 and 2012 are seen. The integrated luminosity is found in the (C)[69].

## 2.2 The Atlas Experiment

The ATLAS detector is one of the two general purpose LHC experiments. This section will describe some of its main design features. By general purpose is meant that it is able to perform a wide variety of physics measurements in both proton and heavy ion collisions.

Listing the major criteria it has to fulfil:

- Decide 40.000.000 times per second if a bunch crossing is considered worth storing.
- Reconstruct particle tracks and measure their momenta with high resolution.
- Measure particle energy.
- Identify particle type by its EM and strong interaction with matter.
- Do this in very dense environments where it is necessary to reconstruct >20 interaction vertices or >500 tracks per bunch crossing[72].
- Have full azimuthal coverage and as high polar coverage as possible<sup>5</sup>.

How these considerable challenges are met will be outlined in the following section. Figure 2.3 illustrates how particle identification, momentum and energy measurements are performed.

The detector is constructed as co-axial cylinders around the beam-axis. Closest to the centre is the inner detector, consisting of the Pixel/SCT detectors and the Transition Radiation Tracker (TRT). These, being submerged in a magnetic field are able to measure charged particle momenta. Outside of the magnet system are the EM and Hadronic calorimeter systems, designed to measure energy depositions by electrons/photons and hadrons. The muon spectrometer is the outermost layer, which is used to identify muons and improve their measured momentum resolution. Table 2.1 summarises the individual parts and their purpose.

A word on the coordinate frame that is used will be given before the following section describes the design and functionality of each of the subdetector systems.

### 2.2.1 The ATLAS Coordinate Frame

A cylindrical coordinate frame is used to take full advantage of the symmetries of the experiment. The frame is defined by:

$$r = \sqrt{x^2 + y^2} \quad \tan(\phi) = \frac{y}{x} \quad \eta = -\ln \left( \tan \left( \frac{\arctan(r/z)}{2} \right) \right)$$

The conversion is visualised in Figure 2.4. The Cartesian coordinate frame is defined such that  $x$  points to the LHC centre,  $y$  upwards and  $z$  in the beam direction towards Geneva.

<sup>5</sup>For a spherical coordinate frame where the beam-line goes through the polar axis

Type	Purpose
<b>Inner Detector System</b>	
Pixel detector	Measures bending of charged particles and primary/displaced vertices.
Silicon strip detector	Layers of silicon micro strips are used to seed tracking and measure bending of charged particles.
TRT	Gas filled tubes measure transverse track curvature. Transition radiation is created and used for electron identification.
<b>Calorimetry</b>	
LAr	Central region has high granularity giving optimal conditions for $e/\gamma$ measurements. Uses a presampler to correct for up-stream energy-loss. Higher pseudorapidity regions are covered by the LAr hadronic end-cap and LAr forward calorimeter.
Tile	Scintillating tile calorimeter outside central LAr. Improves especially hadronic containment and energy resolution.
<b>Muon spectrometer</b>	Measures bending of muons by a toroidal field in the outermost part of ATLAS.
<b>Trigger</b>	Indicates when an event topology is considered interesting. Different levels operate at different frequencies.

TABLE 2.1: Names and purposes of the ATLAS sub detectors and trigger

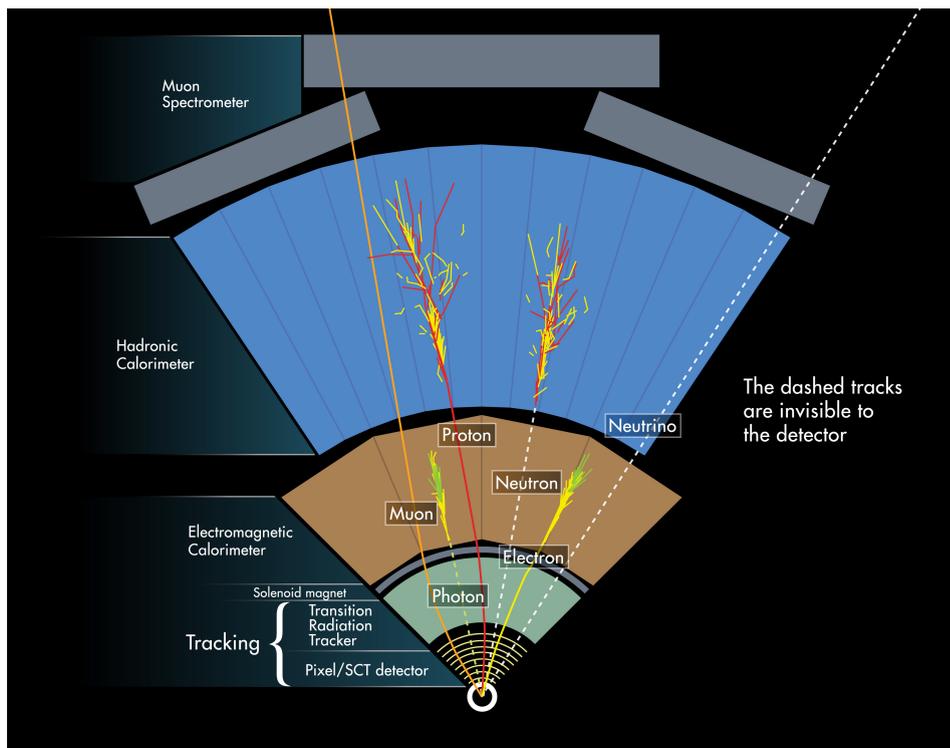


FIGURE 2.3: Depiction of how particles are identified using the ATLAS detector. After colliding in the bottom of the figure, the particle will in order traverse the Inner Detector, the Electromagnetic Calorimeter, The Hadronic Calorimeter and the Muon Spectrometer. The Inner detector is submerged in a magnetic field and is used to reconstruct charged particle trajectories. The calorimeters are used measure particle energy. The figure is somewhat simplified as hadrons will e.g. deposit energy in the E-Cal as well. **ATLAS Experiment**

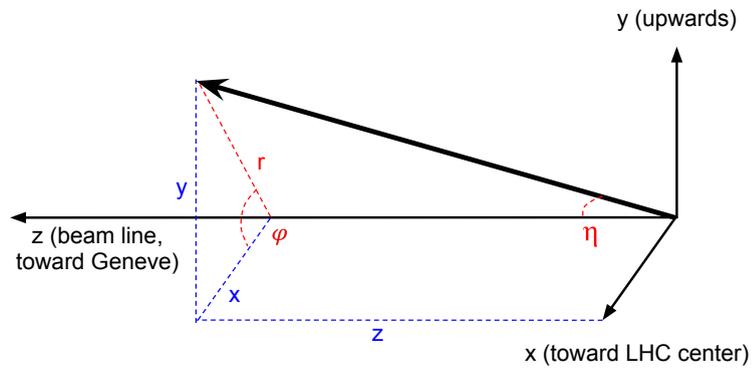


FIGURE 2.4: Figure showing the Cartesian and the cylindrical coordinate frame used in the ATLAS detector.

Since the collisions and detector are symmetrical around the beam axis, it is natural to use the azimuthal angle,  $\phi$ . In Section 1.1.4 it was moreover argued that the low momentum transfer background would tend to be distributed uniformly in rapidity. It is however as a rule of thumb difficult to measure energy/momentum and easy to measure angles. In the limit where  $E \gg m$  the pseudorapidity ( $\eta$ ) asymptotically approximates the rapidity ( $y$ ) making it the preferable choice.

The reason it is desirable to use the transverse length rather than the radial length can among other be found in the detector design. The detector has full coverage in  $\phi$  but not  $\eta$ . This means that measured momentum and energy is only conserved in the transverse plane.

### 2.2.2 The ATLAS Inner Detector

The Inner detector is designed to be hermetic in the transverse plane and be able to create robust tracking recognition. It is able to reconstruct primary, secondary vertices and tracks with high resolution. It moreover has electron identification for  $|\eta| < 2$ . The design is an optimisation between having as little material to influence particle trajectories and as high resolution as possible. The Inner Detector uses two types of detector design; a silicon sensor based tracking system and a Transition Radiation Tracker.

The silicon based tracking system consist of a pixel and a silicon microstrip tracker. Figure 2.5 depicts the inner detector system. The trackers are divided into a barrel region and end-caps. The barrel region is formed as a cylinder, and has a number of layers at constant distance from the beam axis. The end-caps consist of layers with constant  $z$ -coordinates, i.e. disks in the transverse plane. The discs are used to measure higher pseudorapidity regions.

The Pixel and SCT detectors are semiconductor or solid-state detectors. The general principle relies on charged particles creating electron-hole pairs in the material which can be collected by

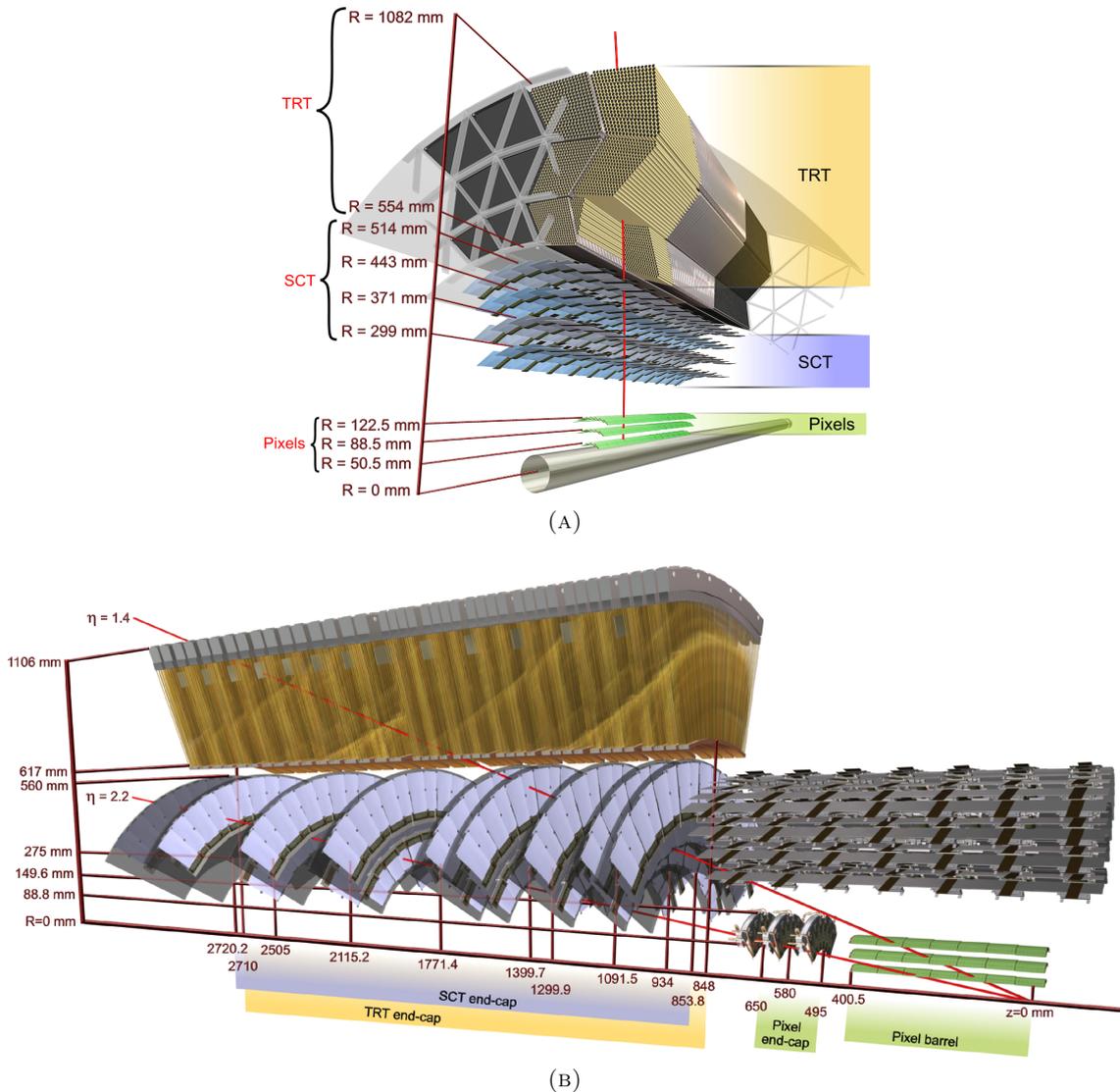


FIGURE 2.5: (A) Front view of the ATLAS Inner Detector. The Pixels, SCT and TRT systems are separately highlighted. (B) Side view of the same figure [73].

an external electric field<sup>6</sup>. Semiconducting detectors have properties that make them ideal for tracking conditions close to the collision point. They have small ionisation energies compared to gas detectors, or equivalently give a large signal when hit by an ionising particle. The drift time of electrons and holes in the semiconductor is low, resulting in fast detector response. Semi-conducting detectors are moreover possible to produce very small giving a high position resolution. A drawback of using solid-state detectors is their tendency to deteriorate from radiation damage.

The material solid-state detectors consist of (silicon in ATLAS) will have electrons in the outermost atomic shell of its crystal structure with a special energy band structure. The structure contains a valence band, where electrons are bound to an atom and a conduction band with free electrons. The two band are separated by an energy gap.

The semiconductor is said to be doped, by adding small impurities to the crystal structure

<sup>6</sup> A more detailed description of semiconductor particle detectors can be found in [74, Ch. 10]. Information here can also be attributed to this source.

with either more (*n*-type) or less (*p*-type) valence electrons than silicon has. The sensitive element of the detector is created by 'pressing' e.g. a *p*-type semiconductor and an *n*-type material together. The excess of holes and electrons in each of the parts will diffuse to the other. After stability is reached, an electric field is created across the junction. If an ionising particle traverse the junction the charge deposit will quickly flow out of the potential, leaving a signal.

The ATLAS pixel detector consist of 1744 sensors with 47232 silicon pixels on each<sup>7</sup>. For the most part, the pixels have a size of  $50 \times 400 \mu\text{m}^2$ . The closeness of the Pixel detector to the interaction point together with its high resolution enables it to measure interaction vertices very precisely.

The SCT consist of 15912 sensors with 768, 12 cm active strips per sensor. Each layer of the SCT is able to register a particle twice, as it consist of two strips which form a small angle between them. In this way they are able to both give  $\eta$  and  $\phi$  measurements. Their direction is a compromise between resolution in the  $\phi$  and  $\eta$  direction. The  $\phi$  direction weighs more heavily as a higher  $\phi$ -resolution translates to a higher  $p_T$  resolution.

The description of the ATLAS silicon detectors is based on [75, Ch. 4.1] where numbers are also taken from.

The geometry of the Transitional radiation tracker (TRT) is similar to the silicon based detectors, with a barrel region and two end-cap regions, see Tab. 2.2. Instead of pixels or strips, the TRT consists of a large number of gas filled tubes. Each tube, functioning as a cathode, has an anode wire in the centre. The straws serve as a proportional gas chamber such that if the gas inside the tubes is ionised, the free electrons will start accelerating towards the centre, knocking more electrons free on its way and in doing so starting a cascade. The charge collected at the anode serves as the detector output. The output strength of each straw tube is a measure of how many electrons hit the anode.

The straws are covered in thin multilayer films. When a charged particle passes through a barrier between two media with different dielectric properties, it will emit radiation. Starting e.g. in a medium of low dielectric constant, the medium will experience small polarisation effects and the charged particle's electric field will have a large spacial extent. If the particle passes to a medium of high dielectric constant, which thereby has a larger polarisation effect, the extent of the electric field will be reduced. The sudden redistribution of charge in the medium will give rise to transition radiation.

The radiation emitted by a charged particle passing through a single surface is [68, Eq. 12.5]:

$$W = \frac{2}{3} \alpha \omega_p \gamma ,$$

where  $\omega_p$  is the plasma frequency of the medium. The radiation is centred in a cone around the charged particle with deflection angle:  $\theta \sim 1/\gamma$ .

The radiation will ionise the gas inside the tubes and start cascades. Since the amount of ionisation from transition radiation is dependent on the traversing particles  $\gamma$ -factor it can be used for particle identification. Given the same momentum, an electron e.g. has a  $\sim 200$  times

<sup>7</sup>The sensors have 46080 readout channels since some of the pixels are clustered.

higher  $\gamma$ -factor than a pion.

Figure 2.6 show how the TRT output is used for electron identification. The signal strength collected by the anode is divided into time bins and classified as either over threshold or over high threshold.

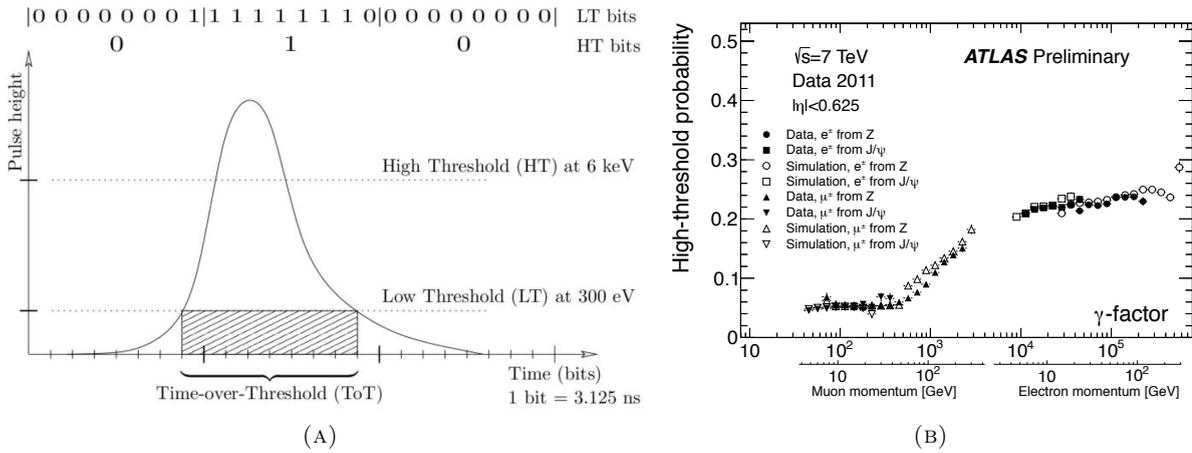


FIGURE 2.6: ATLAS TRT electron identification. (A) TRT digitisation, the signal strength is divided into time bins where two thresholds are defined, low and high[76]. (B) The probability to obtain a high threshold hit is dependent on the  $\gamma$ -factor of the traversing particle[77].

The TRT straws are 4 mm in diameter which can be considered a compromise between the cascade drift time and the number of ionisation clusters. During the LHC Run-1 a gas mixture of Xe, CO<sub>2</sub> and O<sub>2</sub> was used[78].

While the tubes are somewhat larger than the pixels and strips in the silicon detectors, they contain less material, resulting in less particle interaction.

In the barrel region there are three layers with 22, 38 and 52 tubes respectively, while the end-caps have three layers with 8 tubes each [75, Tab. 4.3].

The barrel in total contains 52544 straws, while the endcaps contain 12880. The barrel straws are 72 cm in length, joined in the end to give a barrel length of 144 cm. The end-cap straw length is 37 cm[78]. For this reason the TRT is not able to measure the  $z$  component of a traversing particle in the barrel region accurately, and likewise for the transverse component in the end-cap discs. Tracking has to be combined with the other parts of the detector. Beside particle ID, the use of the TRT improves momentum resolution and give possibility of tracking at greater distances from the collision point.

Elaborate tracking algorithms combine hits from ionising particles into tracks. A description of an early version is e.g. given in [75, Sect. 10.2.1].

The Inner Detector is submerged in a 2 T solenoidal magnetic field parallel to the beam axis, such that charged particles will bend in the transverse plane. The Lorentz force acting on a charged particle combined with the radius of curvature make it possible to calculate its transverse momentum.

The sources of the  $p_T$  error can be divided into two. The first is attributed to interaction with

Part	radius [mm]	length [mm]	Aprox. $\eta$ cove.	R- $\phi$ acc. [ $\mu\text{m}$ ]
<b>Pixel</b>				
3 barrel layers	50.5 - 122.5	0 - 400.5	0.0-2.0	10
$2 \times 3$ end-cap disks	88.0 - 149.6	495-650	2.0-2.5	10
<b>SCT</b>				
4 barrel layers	299 - 514	0 - 749	0.0-1.5	17
$2 \times 9$ end-cap disks	275 - 560	839 - 2735	1.5-2.5	17
<b>TRT</b>				
3 barrel layers	554 - 1082	0 - 780	0.0-1.0	130
$3 \times 3$ end-cap disks	617 - 1106	827 - 2744	1.0-2.0	( $z$ - $\phi$ )130

TABLE 2.2: Geometry of the inner detector. [75, table 1.2, 4.3, p.6-7] Listed are the radius, length, approximate  $\eta$  coverage and characteristic accuracy of the individual parts.

matter and is mostly relevant at low momenta. The second type stems from mis-alignment, the uncertainty of the macroscopic size of the detector pixels and finite knowledge of the magnetic field strength, which dominate the resolution at high momenta. The design expectation on the relative error on transverse momentum is [75, Tab. 1.1]:

$$\sigma_{p_T}/p_T = 0.05\% \cdot p_T \otimes 1\% \quad [\text{GeV}] , \quad (2.5)$$

where the constant term stems from material interaction and the term proportional to  $p_T$  comes from resolution. The geometrical features of the inner detector are listed in Table 2.2.

### 2.2.3 Calorimeter System

The ATLAS calorimeter consist of several subsystems. The electromagnetic, the hadronic endcap and the forward calorimeters are sampling calorimeters with liquid argon (LAr) as the active medium. The calorimeter has been designed with an accordion geometry. The geometry has been found to optimise energy resolution in all directions while leaving no cracks [75, Sect. 5.2.1]. A scintillator tile calorimeter is situated in the barrel region outside of the electromagnetic.

The calorimeter system has a high pseudorapidity coverage, extending up to  $|\eta| \simeq 4.9$ . It is designed to have high containment of electromagnetic and hadronic showers and minimise punch-through to the muon system. The system has more than 22 radiation lengths in its full coverage and around 10 interaction lengths of the active calorimeter, ensuring both high resolution of shower energies and missing transverse energy. A drawing of the system can be found in Figure 2.7, and the number of interaction lengths a particle will meet when traversing the calorimeter system is shown in Figure 2.8.

The LAr electromagnetic calorimeter is optimised for high resolution  $e/\gamma$  measurements. It consists of a barrel region  $|\eta| < 1.475$  and an endcap region  $1.375 < |\eta| < 3.2$ , each housed in a crystal. The electromagnetic calorimeter shares vacuum vessel with the central solenoid magnet system to avoid having two vacuum barriers and reduce the material budget [75, Ch. 5.2].

An electron entering a material will deflect and emit bremsstrahlung with cross section  $\propto$

$\frac{1}{M_e^2}$ [74, P.38]. Bremsstrahlung will initially be the dominant effect at LHC relevant energies. Collision effects will start to become relevant as the electron emits energy. The total energy loss can be approximated by the collision loss and the loss due to bremsstrahlung. The energy at which the two terms are equal is usually referred to as the critical energy and is typically  $\sim 100$  MeV[74, Tab. 2.2].

Different mechanisms will govern the energy depositions of photons, since they are chargeless. At high energies, the energy loss is dominated by pair production of electrons and positrons and at lower energies Compton scattering. The pair production is related to bremsstrahlung such that for  $E_\gamma \gg m_e$  the mean path a photon will have before pair production is[74, Eq.2.123]:

$$\lambda_{\text{pair}} \simeq \frac{9}{7} L_{\text{rad}}$$

Imagining for instance a high energy photon entering into a material. The photon will on average after  $\lambda_{\text{pair}}$  create an electron positron pair, each with  $E_\gamma/2$ . After two radiation lengths these will on average both have emitted one photon due to Bremsstrahlung. Initially the average number of particles will as a result grow exponentially as a function of length traversed in the medium, while the energy of each particle will decrease exponentially. A particle cascade is created until the lower energy processes take over at the critical energy and stop the particles completely.

On average, the number of particles that are created will be  $N \simeq E_I/Ec$ , where  $E_I$  is the initial energy. If the material is interlayered by detectors to count the number of particles, an energy measurement of the original particle can be obtained. The above description is of course a simplification and Monte Carlo techniques are used in real life.

The LAr-EM calorimeter uses lead to stop particles. Its innermost layer has high granularity to achieve high directional resolution. A presampler is situated in front of the calorimeter in the  $|\eta| < 1.8$  region to account for upstream energy loss. The presampler also uses LAr as its active medium

The Scintillator Tile calorimeter consists of a barrel region ( $|\eta| < 1.0$ ) and an extended barrel  $0.8 < |\eta| < 1.7$ . The calorimeter consists of 64 azimuthal pieces. The Tile Cal uses steel to stop particles and samples with scintillators. The barrel consists of three layers with 1.5, 4.1 and 1.8 interaction lengths and three layers of extended barrel with 1.5, 2.6 and 3.3 interaction lengths[75, Sect.5.3.1].

The Hadronic Endcap is built of wheels on each side of the detector. It extends between pseudorapidities of 1.5 and 3.2 such that it has an overlap with both the tile and forward calorimeters. It is made of two segments of 32 wedge shaped modules resulting in four layers per end cap. The HEC uses copper plates and liquid argon[75, Sect.5.3.2].

The forward calorimeter is integrated in the endcap crystal and is designed for very high radiation environments. The FCal has three modules, where the first has copper layers and the two outer use tungsten. The first uses copper to be optimised for electromagnetic interactions[75, Sect.5.3.3].

The processes involved in hadronic showers include multi-particle production of e.g. pions and secondary particle ionisation. Since it is likely that  $\pi^0$ s will also be produced, which almost

instantaneously decay to two photons, the hadronic shower will also be mixed with an electromagnetic, increasing the complexity of the process.

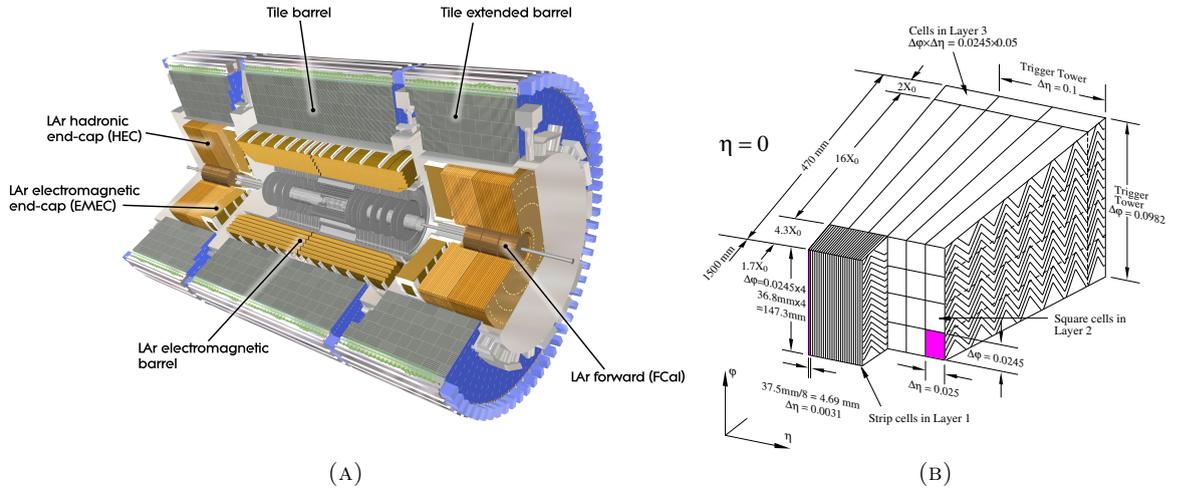


FIGURE 2.7: (A): Drawing of the ATLAS calorimeter system with individual parts labelled. (B): Accordion geometry of the LAr calorimeter in the ATLAS detector [75] **ATLAS Experiment** ©2012 CERN.

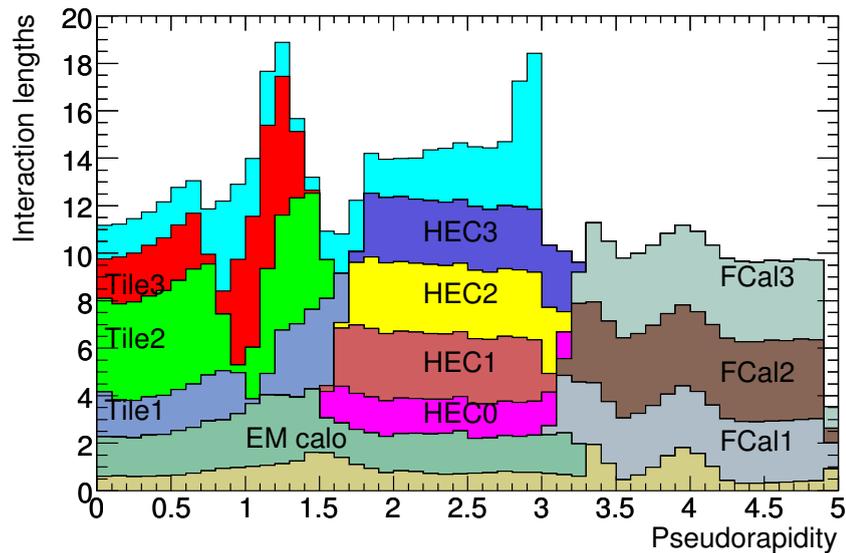


FIGURE 2.8: Integrated interaction length as function of pseudorapidity of the ATLAS calorimeter system [75].

### 2.2.4 The Muon Spectrometer

The Muon Spectrometer is the outermost component of the ATLAS detector. It is almost exclusively muons that are able to reach the central parts of the spectrometer, making both the muon selection efficiency and background rejection high. The spectrometer's immense size and

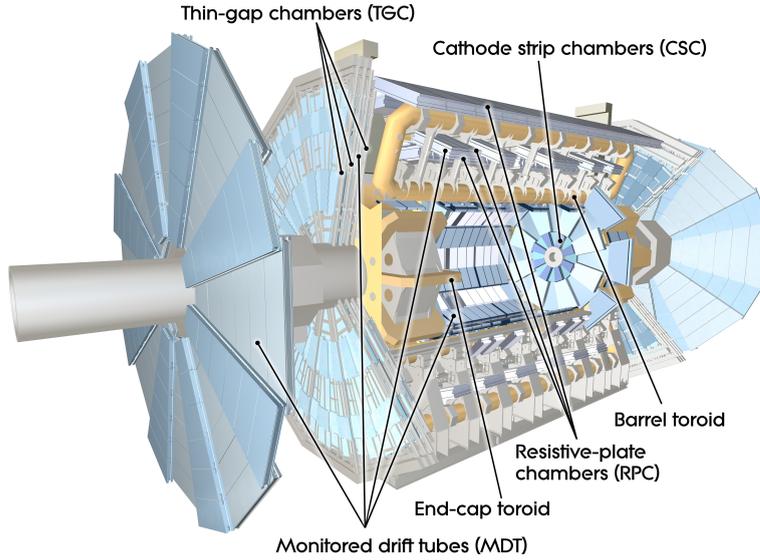


FIGURE 2.9: The Muonic Spectrometer of the ATLAS detector. **ATLAS Experiment**  
 © 2014 CERN.

high bending power yield a high momentum resolution. The relative  $p_T$  resolution is estimated on a  $Z \rightarrow \mu\mu$  sample to be 1.7% for central muons with low  $p_T$  and 4% for high  $|\eta|$  and  $p_T = 100$  GeV. The resolution is based on a combination of ID and MS, where it is expected that the MS will dominate the measurement at high  $p_T$  [79].

The muon spectrometer consists of four separate detector elements and a superconducting air-core toroidal magnet system. The detector elements provide both triggering and high precision tracking.

The magnet system is divided into a barrel region  $|\eta| < 1.4$  with bending power ( $\int Bdl$ ) between 1.5 and 5.5 Tm and two end-cap magnets in the forward  $1.6 < |\eta| < 2.7$  region with bending power between 1 and 7.5 Tm [75, Sect. 1.4.1]. Each of the three toroids consists of symmetric coils around the beamline. Bending between the two regions is supplied by a mix of the two fields. The bending of charged particles in a toroidal field will be different at the two ends of the detector, meaning the reconstruction efficiency is also slightly different in the two ends.

The detector elements consist of the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC) and the Thin Gap Chamber (TGC). Their coverage and primary function are listed in Table 2.3.

The muon spectrometer is constructed of two parts, joined at  $\eta = 0$ . The spectrometer as a result of this has a 'crack region' at very small pseudorapidities, where particles are not measured. A depiction of the MS layout can be seen in Figure 2.9.

The MDT system [75, Sect. 6.3] provides precision measurements over a large  $\eta$  range. The elements rely on 29.970 mm tubes filled with a mixture of Ar and CO<sub>2</sub>, serving as gaseous ionisation detectors. The barrel region consists of three layers of chambers while the endcap has two. The wedges that constitute the MDT system are slightly overlapping to avoid crack regions.

The CSC [75, Sect. 6.4] serves as the other system for high precision measurements in the

Part		Function	Aprox. $ \eta $ cov.
Monitored drift tube	MDT	Precision tracker	0.0 - 2.7 (first layer to 2.0)
Cathode strip chambers	CSC	Precision tracker	2.0 - 2.7
Resistive plate chambers	RPC	Trigger, second coordinate	0.0 - 1.05
Thin gap chambers	TGC	Trigger, second coordinate	1.05 - 2.7 (trigger to 2.4)

TABLE 2.3: Coverage, names and function of the muon spectrometer subsystems[75, Tab. 1.4].

muon spectrometer. It is situated as the innermost layer of the barrel region in front on the MDT. It is optimised to have high spacial and time precision and double track resolution. The CSC is able to perform with a track density of 1000 HZ/cm<sup>2</sup>, while the design density of the MDT is 150 HZ/cm<sup>2</sup>. This makes it optimal for the innermost layer of the forward region.

The CSC is a multiwire proportional chamber, where a series of parallel anode wires are held between two cathode plates. The cathode plates of the CSC are segmented. One of the plates is segmented orthogonal to the wires and the other parallel to. The multiwire proportional chamber will, except very close to the wires, have an electric field with a constant direction orthogonal to the wires. When a particle ionises the gas, the ions will drift along the electric field direction. By segmenting the cathodes it possible to obtain a coordinate of the particle path.

The MDT system consist of 1150 chambers with a combined 354000 readout channels. The CSC consist of 32 chambers with 30700 readout channels. The MDT can give 20 measurements per track in both the barrel and endcap region, while the CSC give four in the forward region. The muon spectrometer trigger systems consist of the RPC [75, Sect. 6.7] and TGC [75, Sect. 6.8]. The MS trigger serves to give bunch crossing ID, a well defined  $p_T$  threshold and muon coordinate in the orthogonal direction to the inner detector trigger. The MS has triggering coverage for  $|\eta| < 2.4$ . The RPC cover the barrel region while the TGC cover the endcap. The RPC consist of three concentric cylindrical layers around the beam pipe. Its individual component are gas detectors build of two  $23 \times 35$  mm plates with a separation of 2 mm. Each station consist of two layers both giving measurements in  $\eta$  and  $\phi$ . The RPC is thus able to give six measurements per track.

Particles that trigger a coincidence in the two innermost layers of the RPC will have transverse momenta between 6 and 9 GeV, while three coincidences can be used for variable trigger thresholds in the 9 to 35 GeV range.

The TGC is a multiwire proportional chamber with 1.4 mm wire to cathode spacing and 1.8 mm wire to wire spacing. The layout ensures fast operation, making it suitable for the forward region. The middle layer of the MDT in the forward region is supplemented with seven layer of TGC and the inner layer with two. Beside serving as a muon trigger, the TGC also give a second (azimuthal) coordinate to the MDT measurement in the radial direction.

### 2.2.5 Trigger Systems

During the LHC Run-1, the ATLAS computing system was able to write events to disc at a rate  $\sim 1$  kHz, see Figure 2.10b. With an event rate during Run-I of 20 MHz it is clearly necessary to perform an online selection. The ATLAS trigger system is designed to filter events using fast electronics and later fast computing. Comparing the total  $pp$  cross section to e.g. electroweak production of  $W$  and  $Z$  bosons in Figure 1.5 reveals a difference of six orders of magnitude. For Higgs production it is  $\simeq 10$  orders. The trigger system has to be able to identify interesting event types and reject the overwhelming backgrounds<sup>8</sup>.

The trigger system is divided in two levels, level-1 (L1) and level-2 (L2) which are followed by a subsequent event filter (EF). The level one trigger identifies particles with high transverse momenta or energy. An example of level one triggers and their rates are shown in Figure 2.10 for a proton run in 2012. The  $e/\gamma$  and muon trigger rates are 20 and 7 kHz in the beginning of the run. Their rates decrease as the proton bunches are depleted.

The L2 trigger later uses all available information from the detector to evaluate if an interesting event has occurred. This is possible because the L1 trigger has reduced the regions of interest to approximately 2% [75, p.15]. The level two trigger reduces the examined event rate by more than an order of magnitude and  $\lesssim 10^4$  Hz is passed on to the event filter which finally reduces the rate to  $\lesssim 1$  kHz. The raw data at this point contains an amount of information equivalent to 1.3Mb/event [75, p.14].

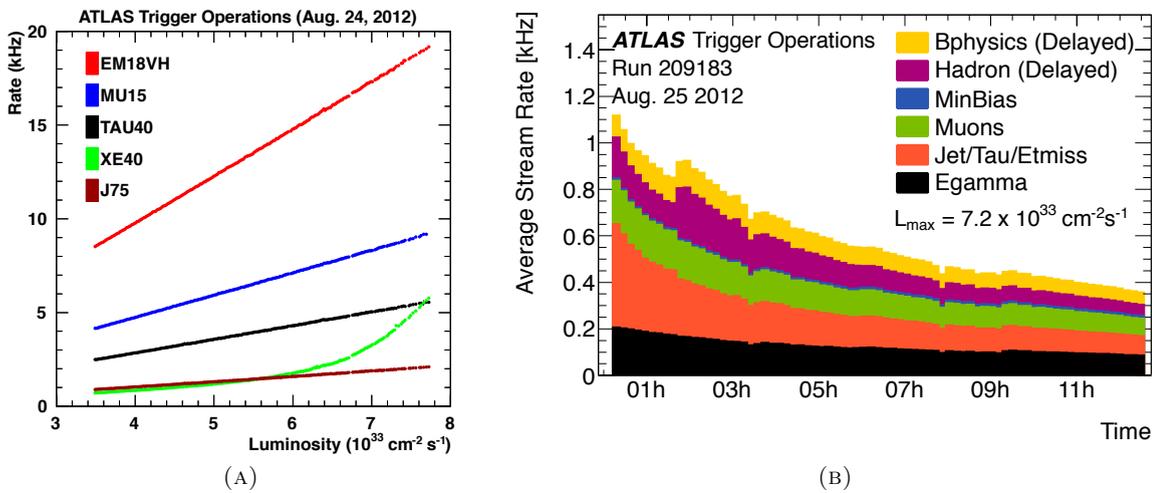


FIGURE 2.10: (A): Level-1 rates for single object triggers at  $7.8 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  (peak luminosity). The different types of triggers that are displayed are: Electron or photon with  $E_T > 18$  GeV (EM18VH), muon with  $p_T > 15$  GeV with coincidence in three stations of the barrel or endcap (MU15), hadronically decaying tau with  $E_T > 40$  GeV (TAU40), missing  $E_T > 40$  GeV (XE40) and Jet with  $E_T > 75$  GeV (J75). (B): Event filter recording rates for a run with peak luminosity of  $7.2 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  [80].

<sup>8</sup>Interesting for physics searches and rare processes measurements.

## 2.3 Primary Vertex Reconstruction in Run-2

During the LHC Run-1 an iterative algorithm was used to reconstruct primary vertices with high success. The pileup during Run-2 is however expected to increase significantly. Using simulations it is found that the iterative process loses accuracy with increasing  $\langle\mu\rangle$ .

The iterative vertex algorithm was mainly optimised to have a high efficiency and avoid vertex splitting, where a single vertex is reconstructed as two. When pileup increases, it becomes more and more important to avoid merging vertices where two are reconstructed as a single vertex. It is found that by iteratively seeding and fitting vertices there will be a tendency for this to happen.

The iterative vertex algorithm can be divided into the following steps:

1. The  $z_0$  values of the tracks' closest approach to the beamspot are scanned using a  $z$ -window. The highest density area is used as the first vertex seed.
2. The vertex is fitted using the tracks within the  $z$ -window. If a track is compatible with the fitted vertex, it is removed from further consideration.
3. The process is repeated for the remaining tracks.

The iterative algorithm finds 81% and 100% efficiency for minimum bias and  $t\bar{t}$  simulation respectively. The probability to split a vertex in the two scenarios is  $\simeq 0.04\%$ .

The problems arise under high pileup conditions. If the tracks from two different vertices are close enough to each other, they may be fitted to a single vertex. It is likely that tracks from both vertices will be consistent with the fitted vertex and therefore discarded from further consideration. If two vertices are merged there is no possibility to recover the mistake at a later state.

Instead, by simultaneously creating all seeds and thereby allowing the tracks to be globally assigned to their nearest seed this problem can partially be avoided. The following will describe an algorithm that follows this approach.

The description will outline [81] where the first ATLAS results using the technique are described. The numbers reported in this section can be attributed to this source. The ATLAS implementation of the new algorithm is to a high degree inspired by [82].

The imaging algorithm can be summarised as follows:

1. A three-dimensional histogram is defined around the beamspot with dimensions; 4, 4, 400 mm in  $x$ ,  $y$  and  $z$  respectively. The tests described here use two different settings for the number of  $z$  bins: 1024 (low  $z$ -resolution) and 2028 (high  $z$ -resolution).
2. All tracks are projected into the histogram volume such that the bin contents are proportional to the track path length they contain. The tracks are assumed to be linear such that their path is given by their perigee parameters. An example of this can be seen in figure 2.11a.

3. A Discrete Fourier Transform is performed on the histogram.
4. A Colcher and Defrise filter function is modified to the acceptance of the ATLAS inner detector[83][84]. It corresponds to the inverse Fourier transform of the detectors angular acceptance.
5. A four-term Blackman-Harris window filter is created[83]. The window filter smooths high frequency variations in the original image. A cutoff frequency is specified individually for x, y and z. This for instance reduces the probability for a single track to create a false seed.
6. The transformed track image is multiplied by the filters. An inverse Discrete Fourier Transform returns a histogram in position space. The filtered track density can be found in figure 2.11b.
7. A clustering algorithm identifies good seed candidates in the filtered image. A simplistic clustering is used for what is shown here. The filtered density is projected onto the z-axis. All local maxima above a threshold are considered seed candidates. If the local minima between two seed candidates is more than 90% of the smallest maximum, the seeds are merged.
8. Each track is associated with its closest seed before the vertices are fitted

The imaging method yields an 80-82% and 100% efficiency for minimum bias and  $t\bar{t}$  simulation respectively. The probability to split vertices is between 0.2% (min-bias, low z-resolution) and 9.5% ( $t\bar{t}$ , high z-resolution). Clearly some optimisation is still needed.

More advanced clustering algorithms could possibly reduce the splitting rate drastically. It is not given however, that the imaging method will outperform the iterative algorithm in terms of splitting probability.

Turning now to the problem of vertex merging. Figure 2.12a demonstrates an area where the imaging approach performs well. It shows the number of correctly reconstructed vertex pairs as function of the z-distance between them ( $\Delta z$ ). As  $\Delta z$  decreases fewer and fewer pairs are reconstructed. The imaging approach is seen to have a narrower gap and therefore less prone to merge vertices.

The probability to merge a vertex can be realised is quadratic in pileup. This means that the number of reconstructed vertices as function of pileup will take the form[81, Eq.1]:

$$N_{\text{vtx}}^{\text{reco}} = c_0 + \epsilon\mu(1 - m\mu) ,$$

where  $\epsilon$  describes the reconstruction efficiency and  $m$  describes the loss term due to merging. Figure 2.12b shows  $N_{\text{vtx}}^{\text{reco}}$  for three different setups; iterative and imaging with different z-resolution. A perfect vertex reconstruction would have  $\epsilon = 1$ ,  $m = 0$ . The high resolution imaging method is seen to have a smaller quadratic loss term when going to high pileup conditions. At a relevant pileup of 40, the iterative method is estimated to lose 32% of vertices, while only 20% are lost with the imaging vertex algorithm.

The algorithm shows promising results and will be interesting to further optimise for the high pileup conditions of the LHC Run-II.

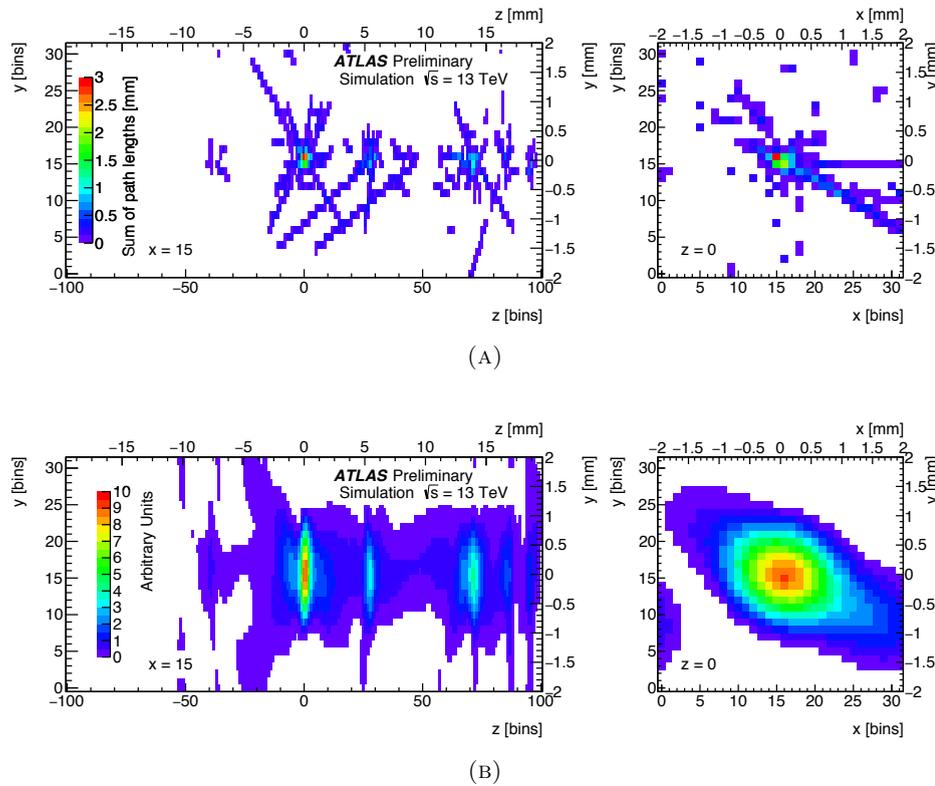


FIGURE 2.11: Two-dimensional projections of the 3D histograms used in the Imaging vertex algorithm. The histograms show a single simulated  $t\bar{t}$  event. (A) The tracks from the simulated events are projected into the volume such that the bin content correspond to the total track length the bin contains. (B) The same histograms after the image filtering process described in the text has been performed. Left: The  $yz$  ( $x=15$ ) slice. Right: The  $xz$  slice ( $y=0$ ) [81].

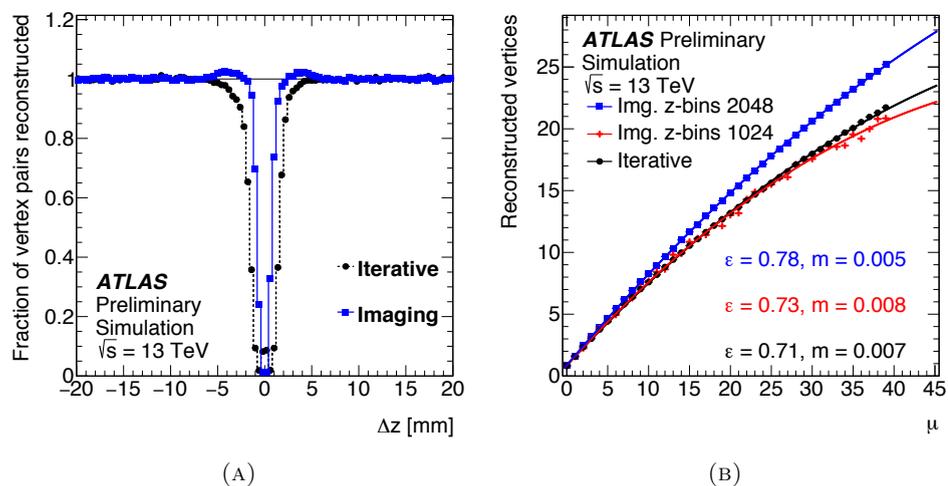


FIGURE 2.12: (A) Distance between reconstructed vertex pairs in  $z$ . The gap-width is a measure of how sensitive the approach is to merging. The imaging method uses the high  $z$ -resolution configuration. (B) Number of reconstructed vertices as function of pile-up. The iterative approach is compared to the imaging algorithm with high and low  $z$ -resolution [81].

## 2.4 Lepton Identification for the $H \rightarrow ZZ^* \rightarrow 4\ell$ Analysis

The analysis will mainly focus on the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay chain. Some detail will here be given on how electrons and muons are identified with the ATLAS detector.

The methods here are specifically tuned to the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. The description should however serve as an illustration on how detector output is utilised and cover many features, as  $e, \mu$  reconstruction is based on the trigger system, the inner detector, the calorimeter and the muon spectrometer.

More details can be found in [85]. The numbers reported here can also be attributed to this source.

As a start it is worth looking at the characteristic reconstruction efficiencies for electrons and muons in data and Monte Carlo, shown in Figure 2.13. There is in general greater loss in the electron reconstruction, where efficiency drops especially in the low transverse energy region. Muons overall have very high efficiencies, but the muon spectrometer crack can be seen to lower the efficiency for small  $|\eta|$  values.

Events that potentially contain  $H \rightarrow ZZ^* \rightarrow 4\ell$  decays are identified at the trigger level as the first step. Due to the different background types of electrons and muons, different  $p_T$  selections are required. The trigger either identify a single lepton with high  $p_T$  or two leptons with lower  $p_T$ -requirement. The triggers that are used to select events in the 2011 and 2012 datasets are found in Table 2.4. Some explanation of the ATLAS trigger nomenclature is necessary:

Starting with the 2011 dataset. The trigger names start with EF indicating that the selection is done at the Event Filter level. The relevant lepton, e or mu for electrons and muons respectively and the  $E_T, p_T$  selection is listed next. The trigger is indicated with a postfix to be either loose or medium referring to a quality requirement. For electrons the requirement is imposed on the EM shower shape and in the medium case the relative energy deposit in the electromagnetic sampler. Finally, the MG for muons indicate the reconstruction algorithm "Muon Girl". The trigger requirements tightens during 2011 due to the increase in instantaneous luminosity, see Figure 2.2a.

The rest of the indicators are: i: lepton is isolated such that surrounding tracks within a cone of  $\Delta R = 0.2$  do not have more than 10% of the lepton  $p_T$ . h: trigger is seeded from a L1 trigger with hadronic veto. v: trigger is seeded from a L1 trigger with variable threshold. T: trigger is seeded from L1 trigger with tight selection criteria. For muons in the 2012 selection there are also EFFS (Event Filter Full Scan) for the di-lepton selection. The Event Filter will in this case first find a 18 GeV muon and subsequently scan the entire event for a muon with  $p_T > 8$  GeV. L2StarB indicate a software framework to do pattern recognition at the L2.

In the subsequent analysis it is later required that the triggering leptons should be in the set of leptons that form the Higgs-quadruple. For the di-lepton trigger both leptons are required to be used.

The 2012 trigger performance is estimated with a tag-and-probe approach on  $Z \rightarrow \ell\ell$  events in data. The efficiencies for the different decay types for a Higgs mass of  $m_{4\ell} = 130$  GeV are

Period	Final State	Single-lepton	Di-lepton
2012	4e	e24vhi_medium1, e60_medium1	2e12Tvh_loose1, (L2StarB, data only)
	4 $\mu$	mu24i_tight, mu36_tight	2mu13, mu18_mu8_EFFS
	2e2 $\mu$	4e, 4 $\mu$ , e12Tvh_medium1_mu8,	e24vhi_loose1_mu8
2011, B-I	4e	EF_mu18_MG	EF_2mu10_loose
	4 $\mu$	EF_e20_medium	EF_2e12_medium
2011, J	4e	EF_mu18_MG_medium	EF_2mu10_loose
	4 $\mu$	EF_e20_medium	EF_2e12_medium
2011, K	4e	EF_mu18_MG_medium	EF_2mu10_loose
	4 $\mu$	EF_e22_medium	EF_2e12T_medium
2011, L,M	4e	EF_mu18_MG_medium	EF_2mu10_loose
	4 $\mu$	EF_e22vh_medium1	EF_2e12Tvh_medium
2011, All	2e2 $\mu$	4e, 4 $\mu$	4e, 4 $\mu$ , EF_e10_medium_mu6

TABLE 2.4: Trigger menu for  $H \rightarrow ZZ^* \rightarrow 4\ell$  candidate selection. Trigger names separated by a comma is to be understood as a logical 'or' between the two.

found to be:

$$\epsilon_{4\mu} = 97.6\% \quad \epsilon_{2e2\mu} = 97.3\% \quad \epsilon_{4e} = 99.7\%$$

The Monte Carlo efficiencies are moreover corrected by applying a scalefactor to the expected signal yield based on this. Following the trigger, more stringent requirements are imposed that utilise the full detector information.

Starting with electrons. The main backgrounds are conversion photons and QCD particles (mostly pions) interacting in the EM calorimeter. The influence of conversion photons can largely be reduced by requiring that the track delivers a hit to the innermost pixel layer, the b-layer. It is necessary to analyse the electron candidate signature in the TRT and calorimeters to reduce QCD backgrounds. For the 2012 data-sample, the analysis use a likelihood based discriminant:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad \mathcal{L}(x) = \prod_i P_i(x_i)$$

The individual *pdfs* consist of various electron/background separating observables. The full set is listed in Table 2.5. The distributions are derived from data before being applied to the Higgs analysis.

The hadronic leakage tests if energy from the electromagnetic shower leaks into the hadronic calorimeter. Hadronic leakage will tend to happen for heavier QCD particle (pions) but not for electrons. For this reason the relative energy deposit in the last EM-Cal layer will also have discriminating power.

If an electron interacts in the EM-Cal it will tend to have a narrow signature, while QCD jets for instance will be less colimated. By studying the relative energy deposited in the highest cell to the surroundings, it is possible to further reduce backgrounds. A pion will for instance also tend to traverse longer into the calorimeter before interacting than an electron. The relative energy in the first layer to the total energy is a measure of this.

The inner detector tracking can furthermore be used to separate photons from electrons. It should be noted however that the inner detector hits are not used in the likelihood discriminant but are enforced by rectangular cuts. This ensures that all electron candidates used in the

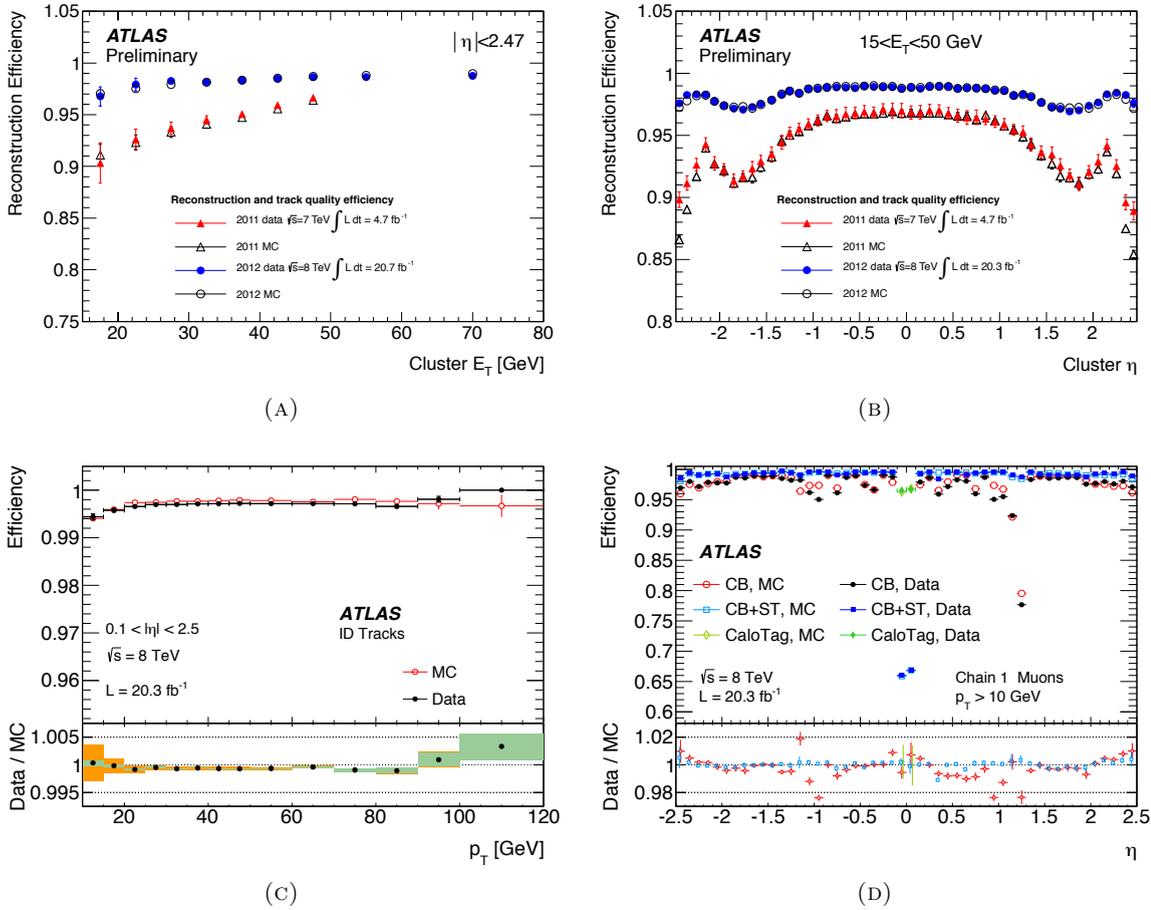


FIGURE 2.13: ATLAS electron and muon reconstruction efficiencies as function of transverse energy(momentum) and pseudorapidity. (A,B) Electron efficiencies for data and expectations from Monte Carlo for the 7 TeV and 8 TeV data samples. (C,D) Muon efficiencies for data and Monte Carlo here only shown for the 8 TeV sample. (C) Efficiency of inner detector tracks from muons. (D) Efficiencies for different reconstruction methods: CB: Combined. CB+ST Combined and Segmented Tagged. CaloTag: Muons with  $\eta$  in the spectrometer crack region which are identified in the calorimeter[86][79].

analysis have a well measured four momentum. To improve the track quality, the inner detector track is required to match the electromagnetic shower-shape. The electrons are reconstructed with a Gaussian sum fitter that improves the electron  $E/p$  to be  $\sim 1$  by including photon emissions from brehmstrahlung.

The energy measurement for most electrons will be based on the calorimeter deposit. For low energy electrons ( $E_e < 30$  GeV) the measurement is however based on a combination of EM calorimeter and tracking due to the lower calorimeter resolution.

The 2011 electron identification uses the same parameters as the 2012, but only imposes rectangular cuts instead of the likelihood approach.

The muon identification is somewhat simpler. In general ATLAS uses different identification criteria for muons:

- Combined muons: Reconstruction uses combined information from muon spectrometer and inner detector.

Type	Description
Hadronic leakage	Ratio of $E_T$ in the first hCal layer to the EM cluster for $ \eta  < 0.8$ or $ \eta  > 1.37$ . Ratio of $E_T$ in hCal to the EM cluster for $0.8 <  \eta  < 1.37$ .
Third layer of EM calorimeter	Ratio of energy deposit in the third ECal layer to total energy.
Middle layer of EM calorimeter	Lateral shower width: $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$ . $E_i$ and $\eta_i$ are the energy and pseudorapidity of cell $i$ in a window of $3 \times 5$ cells. Ratio of the energy in $3 \times 3$ cells over the energy in $3 \times 7$ cells centred at the cluster and energy ratio of $3 \times 7$ to $7 \times 7$ cells.
Strip layer of EM calorimeter	Shower width, $\sqrt{(\sum E_i (i - i_{\max})^2)(\sum E_i)}$ . The index runs over a window of around 20 strips in $\eta$ ; $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$ . Ratio of largest to second largest energy deposit in the cluster divided by their sum. Strip energy deposit to total EM energy.
<i>Track quality</i>	Number of B-layer hits discriminates against photon conversions. Number of hits in the pixel detector. Combined number of hits in pixel and SCT hits. Relative momentum loss in the inner detector.
TRT	Total number of TRT hits. Fraction of TRT hits with high threshold signature.
Track-cluster matching	$\Delta\eta$ between cluster position in EM strip layer and expectation from track. $\Delta\phi$ between cluster position in EM middle layer and expectation from track. Track momentum is re-scaled to EM cluster energy.

TABLE 2.5: Electron Identification Criteria. Track quality requirements are applied directly instead of entering the likelihood model.

- Stand-alone muons: Reconstruction is solely based on parameters measured in the muon spectrometer. Stand-alone muons are especially useful outside the inner detector coverage of  $|\eta| < 2.5$
- Segment tagged muons: Reconstruction is performed in the inner detector and track parameters are taken from there. The track is associated with segments in the precision muon chambers.
- Calorimeter tagged muons: Reconstruction is performed in the inner detector. The muon track is compatible with a minimum ionising particle signature in the calorimeter. The calorimeter reconstruction is used for muons hitting the spectrometer crack at  $|\eta| < 0.1$ .

The analysis will mainly use STACO muons (Segment TAgged, COmbined muons), to get the optimal purity and quality at the cost of some efficiency, but allow at most one calorimeter of stand-alone muon per lepton quadruple.

Moreover, the inner detector selection listed in Table 2.6 is imposed.

ID Hit requirements 2011	
	At least one B-layer hit if expected
ID pixel and SCT	More than one pixel hit. Passing an inactive sensor counts as hits More than five SCT hits. Passing an inactive sensor counts as hits At most three pixel and SCT holes
TRT Hits if $ \eta  < 1.9$	More than five hits, where number of outliers make out at most 90%
TRT Hits if $ \eta  \geq 1.9$	If more than five hits, require number of outliers make out at most 90%
ID Hit requirements 2012	
	At least one pixel hit. Passing an inactive sensor counts as hits
ID pixel and SCT	More than four SCT hits. Passing an inactive sensor counts as hits At most three pixel and SCT holes
TRT Hits if $0.1 <  \eta  \leq 1.9$	More than five hits, where number of outliers make out at most 90%

TABLE 2.6: Inner detector requirements imposed on muons for the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis.

## Chapter 3

# Analysis

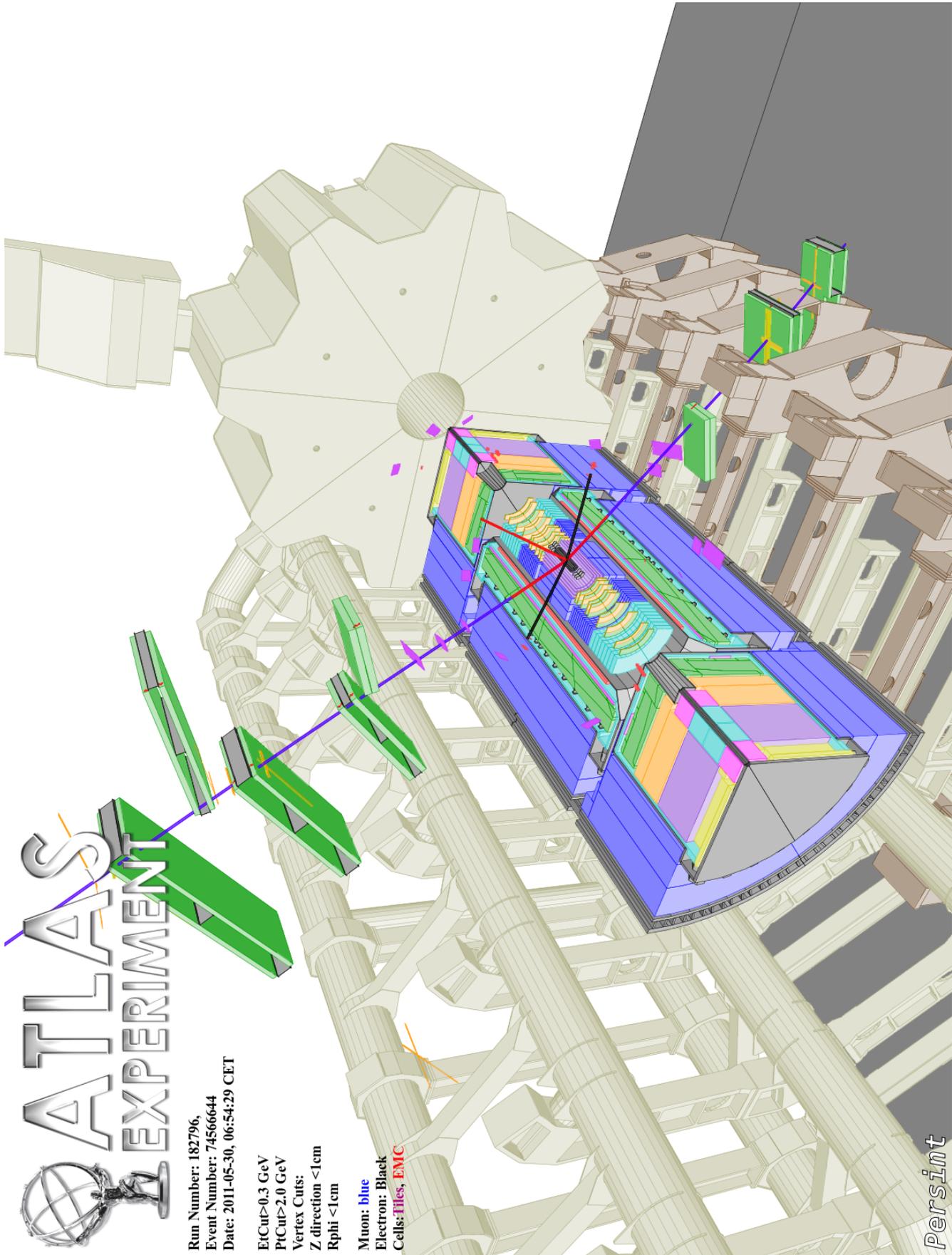


FIGURE 3.1:  $H \rightarrow ZZ^* \rightarrow 2e2\mu$  candidate,  $m_{4\ell} = 124.3$  GeV,  $m_{ee} = 76.8$  GeV,  $m_{\mu\mu} = 45.7$  GeV [87], ATLAS Experiment © 2014 CERN.

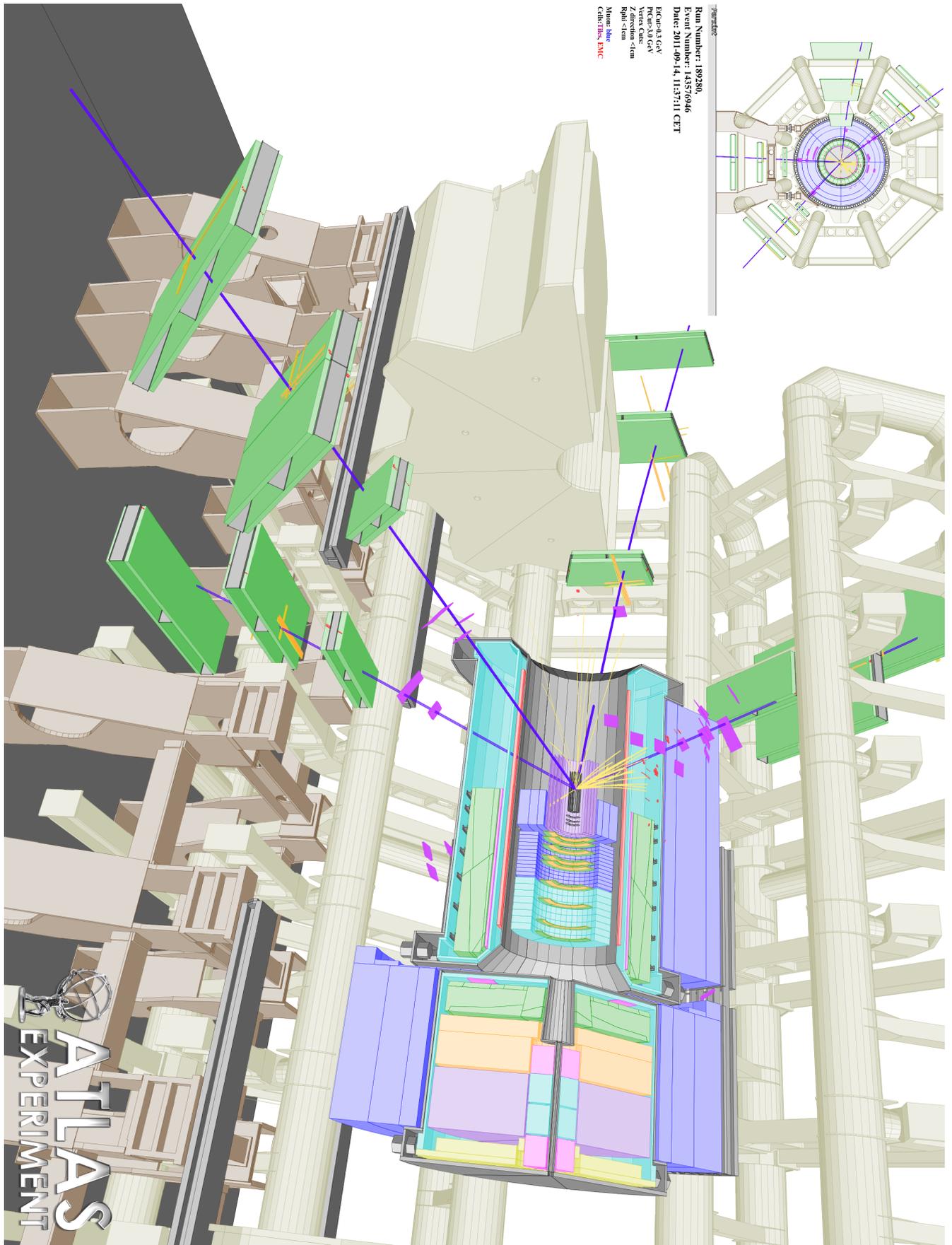


FIGURE 3.2:  $H \rightarrow ZZ^* \rightarrow 4\mu$  candidate,  $m_{4\ell} = 124.6$  GeV,  $m_{\mu\mu} = 89.7$  GeV, 24.6 GeV[87], ATLAS Experiment © 2014 CERN.

## 3.1 Introduction

The four-lepton final state can in general be described with eight parameters for a fixed Higgs mass: Four lepton four-vectors give sixteen degrees of freedom. Since each of these are real particles, they are bounded to lie on their invariant mass shell resulting in twelve degrees of freedom. Constraining the full system to the Higgs four-vector results in eight.

Theoretical descriptions exist of how probable it is for any of the Standard or alternative models to fall into any region of this phasespace. In a perfect world, the analysis likelihood could readily be constructed from this. In real life however the decay will be observed through a detector, giving rise to resolution and acceptance effects. The quest is to find a description that convolutes the theoretical predictions with detector behaviour in a computationally manageable way.

It is in principle possible to simulate the behaviour of the detector when faced with four-lepton production to a very high precision. Simulations however work on an event-by-event basis. The simulation is costly, but the dimensionality of the phase space is more problematic. Imagining a "brute-force" analysis where each dimension of the phase space is divided into 10 regions would result in a total of  $10^8$  regions. Even if it was possible to choose the phasespace such that the events would be reasonably uniformly distributed in it, it would be close to impossible to populate such a phasespace with enough statistics.

The analyser is thus faced with a choice. Either find a simplified description of the detector behaviour or perform a dimensional reduction. The mapping to the reduced phasespace will have to be defined such that practically all relevant information is conserved.

The analyses described in this section will be of the latter type. The hypothesis tests will use boosted decision trees to distinguish models and the tensor structure fit will make use of an approach based on matrix element observables.

The analysis description starts with the event selection, i.e. a description of how  $H \rightarrow ZZ^* \rightarrow 4\ell$  events are identified, how many are expected and how many are observed.

The fixed hypothesis tests introduce the analysis that aim to tests the SM against various pure BSM scenarios. The focus will be on tests against the BSM spin-0 and spin-2 models that were introduced in Section 1.2.3.

Finally, the Tensor Structure fit will be described. The section will focus on the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis, but also include its combination with  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ . The section will conclude with some high-luminosity prospects studies.

A description of how models are constructed and how statistical inference from them is made is given in App. A. The tools that will be used throughout the analysis are described in this part as well. The appendix is self contained and can be read as a whole or used as a reference for the analysis sections.

## 3.2 The $H \rightarrow ZZ^* \rightarrow 4l$ Signal Region

### 3.2.1 Event Selection

A  $H \rightarrow ZZ^* \rightarrow 4l$  optimised selection is imposed after the lepton reconstruction described in Section 2.4. The selection follows what is described as the inclusive selection in [85].

First, events passing the trigger described in the previous section are required to be recorded when any relevant detector system in ATLAS was fully operational. Events are required to have at least one vertex with associated track  $p_T > 400$  GeV. The primary vertex is chosen to be the one with the highest summed  $p_T$ . Electrons that are chosen to possibly constitute a Higgs candidate are required to have a transverse impact parameter less than 10 mm from the primary vertex. Muons have a stricter requirement of 1 mm due to contamination from cosmic-rays.

Kinematic requirements are imposed on the individual electron and muons. The electrons are required to be reconstructable in the inner detector by having a pseudo-rapidity  $|\eta| < 2.47$ . Moreover, in order to reduce the large QCD background and improve calorimeter  $E_T$  resolution a minimum of 7 GeV transverse energy is required.

The muons are divided into three categories of different types of reconstruction, see Section 2.4 for specification. The segmented tagged or combined muons are required to be reconstructable in the muon spectrometer and have transverse momentum of 6 GeV. The momentum again ensures a reduction of low energy backgrounds. Muons reconstructed by the calorimeter in the crack region have higher requirements on transverse momentum as a result of the higher background acceptance. Finally, stand-alone muons are also accepted if they fall outside of the inner detector acceptance.

At most one calorimeter or stand-alone muon is accepted at a time due to their lower purity. The leptons are required to be isolated to reduce contamination from QCD. The specific values used in the selection can be found in Tab. 3.1.

For electrons the surrounding energy density is subtracted from the cluster to remove contributions from pile-up. The muon energy deposit in the calorimeter is estimated from cells with energy  $3.4\sigma$  above the quadrature sum of the expected electronic and pile-up noise. If another lepton that passes the selection is found inside the isolation cone, its energy will be subtracted when calculating the summed energy.

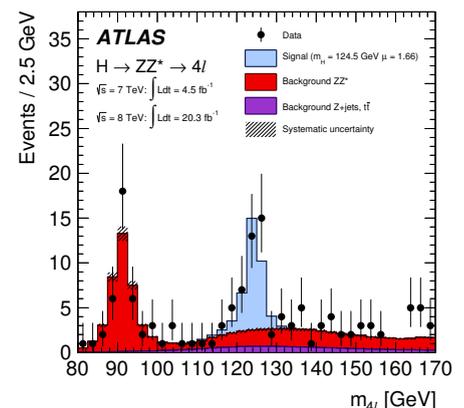


FIGURE 3.3: Four lepton mass distribution after event selection. The analyses use a [115 GeV, 130 GeV] mass window [88].

Following the individual lepton selection, a quadruple is defined as a pair of same-flavour opposite-sign leptons:  $4e$ ,  $4\mu$ ,  $2e2\mu$ ,  $2\mu2e$ . The  $4e$  and  $4\mu$  events have two possible ways to pair the oppositely charged leptons. The leading lepton pair is defined to be the combination with invariant mass closest to the nominal  $m_Z$ . If more than four leptons pass the requirements, the second pair is chosen as the next same-flavour opposite sign pair with mass closest to  $m_Z$ . The quadruple of leptons is restricted by imposing transverse momentum requirements on the leading leptons. Lepton pair invariant masses are furthermore restricted to make the probability that they stem from  $Z$  bosons higher. A  $J/\Psi$  veto finally is imposed together with calorimeter/track isolation and impact parameter cuts.

In the case that enough pairs pass the event such that it is possible to reconstruct several Higgs bosons, only one is chosen. The choice is based on the expected signal yield of the four different final states:  $N_{4\mu} > N_{2e2\mu} > N_{2\mu2e} > N_{4e}$ .

Final state photon radiation recovery is attempted after the selection. Good theoretical descriptions of the  $Z \rightarrow \ell\ell + \gamma$  radiation are available. Two different types of recovery are tried, collinear and non-collinear. The collinear looks for a photon cluster within  $\Delta R < 0.15$  with transverse energy  $E_T > 1.5$  GeV. The collinear recovery only applies to muons. If no collinear photon is found, non-collinear photons are searched for. The non-collinear looks for photons outside  $\Delta R > 0.15$  with  $E_T > 10$  GeV. The non-collinear recovery applies both to electrons and muons. In all instances it is required that the photons are isolated.

Collinear corrections are applied if  $66 \text{ GeV} < m_{\mu\mu} < 89 \text{ GeV}$  and  $m_{\mu\mu\gamma} < 100 \text{ GeV}$ . The non-collinear correction applies to events with  $m_{\ell\ell} < 81 \text{ GeV}$  and  $m_{\ell\ell\gamma} < 100 \text{ GeV}$ . The collinear recovery is estimated to have a photon efficiency of 0.7 with purity of 0.85, while the non-collinear has a photon efficiency of 0.6 but with high purity of 0.95.

The  $H \rightarrow ZZ^* \rightarrow 4l$  analysis uses a signal region defined by the mass window  $m_{4l} \in [115 \text{ GeV}, 130 \text{ GeV}]$ . Table 3.1 summarises the different criteria that are imposed on data to identify Higgs candidates. The invariant mass distribution for selected events together with signal and background expectations can be found in Figure 3.3.

### 3.2.2 Reducible backgrounds

Beside the  $ZZ^*$ -continuum, a number of different process-types will be able to emulate the four lepton signal. The  $ZZ^*$  background is considered irreducible in the sense that the final state objects are identical to that of the signal; four isolated leptons<sup>1</sup>. Contrary to the continuum are the irreducible backgrounds where other types of physics objects will emulate the behaviour of the isolated leptons. This could for electrons be jets and for muons be the muonic decay of the heavy flavour  $b$  quark and  $t\bar{t}$  production.

It is in general difficult to identify all different production mechanisms that will be able to fake the final state behaviour. Another problem exist in generating enough MC to fully describe the background. Take for instance the  $Z \rightarrow ll + \geq 2$  jets, where the additional jets fake

<sup>1</sup>Irreducible is a bit stronger than reality since production and decay observable distributions will differ between Higgs and  $ZZ^*$ -continuum production. It is possible to determine the favoured production mechanism as will be demonstrated in Section 3.3

<b>Lepton selection</b>	
Electrons	Identification as in Sect.2.4 and $E_T > 7$ GeV, $ \eta  < 2.47$
Muons	At most one calo or stand-alone muon. Others should be combined or segmented tagged, as in Sect.2.4. STACO: $p_T > 6$ GeV, $ \eta  < 2.7$ . Calo: $p_T > 15$ GeV, $ \eta  < 0.1$ . Stand-alone: $p_T > 6$ GeV, $2.5 <  \eta  < 2.7$ . $\Delta R > 0.2$ from any segment tagged.
<b>Event Selection</b>	
Kinematic Selection	A candidate consist of one quadruple of two same-flavor, opposite-charge leptons pairs. Three leading lepton $p_T$ of minimum 20, 15, and 10 GeV respectively. Leading di-lepton mass $m_{12}$ defined to be the pair with an invariant closest to the on-shell $Z$ mass. Select quadruple with pair masses closest to the on-shell $Z$ mass. Leading di-lepton mass: $50 \text{ GeV} < m_{12} < 106 \text{ GeV}$ . Sub-leading di-lepton mass: $12 \text{ GeV} < m_{12} < 106 \text{ GeV}$ . Discard candidate if a same-flavor opposite-sign lepton pair exist with $m_{ll} < 5 \text{ GeV}$ . $\Delta R(l, l') > 0.10(0.20)$ for same (different) flavour candidate leptons.
Isolation	Apply isolation to all quadruple leptons, removing contributions from any other in the quadruple. Lepton track: $\sum p_T/p_T < 0.15$ within $\Delta R = 0.2$ . Electron Calorimeter: $\sum E_T/E_T < 0.30$ within $\Delta R = 0.2$ . Muon Calorimeter: $\sum E_T/E_T < 0.30$ within $\Delta R = 0.2$ . Stand-alone muon Calorimeter: $\sum E_T/E_T < 0.15$ within $\Delta R = 0.2$ .
Impact parameter	$d_0/\sigma_{d_0} < 6.5$ for all quadruple electrons. $d_0/\sigma_{d_0} < 3.5$ for all quadruple muons.

TABLE 3.1:  $4\ell$  quadruple candidate requirements.  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ .

the third and fourth electron. The process has a cross section  $\sim 1$  pb (see e.g. Figure 1.5) and  $\text{BR}(Z \rightarrow ee, \mu\mu) = 6.6\%$ [9]. As will be demonstrated below, a total of 1.04 events of this category is expected to pass the selection and fall in the mass window in the 2012 dataset. As a back-of-the-envelope estimation of the efficiency, imagine all of these coming from  $Z \rightarrow ll + \geq 2\text{Jets}$ . The resulting efficiency is  $\simeq 0.005\%$ . In other words, by blindly generating Monte Carlo only one event in 20000 would be useful for further analysis. Millions would be necessary for this single background type.

It is instead chosen to estimate the reducible background behaviour from control regions.

In general the most likely process type to emulate a  $H \rightarrow ZZ^* \rightarrow 4\ell$  event will be of the type  $Z + \ell\ell$  where two additional objects fake the sub-leading leptons. The reducible backgrounds

are divided into  $Z \rightarrow 2\ell + 2e$  and  $Z \rightarrow 2\ell + 2\mu$  as a consequence of their very different identification mechanisms.

Control regions are defined by inverting or relaxing selection criteria. As an example, by inverting the impact parameter significance requirement and isolation criteria for the subleading muons the  $\ell\ell\mu\mu$  category will contain  $b$ -quarks due to their macroscopic flight distance.

The expected number of events in the signal region is determined from transfer factors that are estimated using Monte Carlo samples. Three different approaches for both  $Z + ee$  and  $Z + \mu\mu$  have been constructed to estimate the reducible backgrounds where one is used for the nominal description and the two others are used to describe the systematic uncertainty of the approach. A full description can be found in [85, Sect. 6].

### 3.2.3 Expectations

The number of Higgs bosons that are expected to be produced, decay to the four-lepton final state and finally be reconstructed/identified is given in Table 3.2. A total of 18 signal events are expected to be observed. Not much compared to the total number of Higgs produced by LHC during Run-I ( $\sim 500000$ [34]). Luckily the background processes that pass the selection are equally rare. The post-selection expected and observed number of events are found in Table 3.3. The signal and  $ZZ^*$  columns are derived from Monte Carlo simulations, while the reducible backgrounds are extrapolated from control regions. The 18.0 signal events are compared to 10.4 expected events from the  $ZZ^*$ -continuum and 4.2 for reducible backgrounds. The numbers take into account only events falling in the signal mass window,  $115 \text{ GeV} \leq m_{4\ell} \leq 130 \text{ GeV}$ .

$\sqrt{s}$	FS	$\sigma_H [\text{pb}^{-1}]$	$N_{\text{Run-I}}$	$\text{BR}(H \rightarrow ZZ^* \rightarrow \text{FS}) [10^{-5}]$	$N_{\text{Run-I}}$	Eff.	$N_{\text{Run-I}}$
7 TeV	4e	17.1	$77.1 \cdot 10^3$	3.42	2.64	17.0%	0.45
	4 $\mu$			3.37	2.60	39.2%	1.02
	2e2 $\mu$			3.11	2.40	26.7%	0.64
	2 $\mu$ 2e			3.11	2.40	20.0%	0.47
8 TeV	4e	21.8	$443 \cdot 10^3$	3.42	15.1	19.3%	2.91
	4 $\mu$			3.37	14.9	39.0%	5.81
	2e2 $\mu$			3.11	13.8	27.0%	3.72
	2 $\mu$ 2e			3.11	13.8	21.2%	3.00
total			$520 \cdot 10^3$		67.5		18.0

TABLE 3.2: Signal cross sections and efficiencies. Cross sections and branching ratios from [34]. Production include gluon fusion (88%), vector boson fusion (7.2(7.1)% for 8 TeV(7 TeV) and associated production. A Higgs mass of 125.5 GeV is assumed. Note that the branching ratios of 2e2 $\mu$  and 2 $\mu$ 2e are assumed the same, while not completely accurate, it is close enough for this purpose. Efficiencies are derived from [34] and [5].

### Spin/CP Discriminating Variables

The expected and observed final state distributions are examined after the event selection is imposed. Figure 3.4 show the distributions related to production and the kinematic discriminant  $KD$ , while Figure 3.5 show decay variables. The reducible backgrounds and the

	SM Signal	$ZZ^*$	$t\bar{t}, Z + \text{jets}$	Total expected	Observed
$\sqrt{s} = 7 \text{ TeV}$					
$4\mu$	$1.02 \pm 0.10$	$0.65 \pm 0.03$	$0.14 \pm 0.06$	$1.81 \pm 0.12$	3
$2\mu 2e$	$0.47 \pm 0.05$	$0.29 \pm 0.02$	$0.53 \pm 0.12$	$1.29 \pm 0.13$	1
$2e 2\mu$	$0.64 \pm 0.06$	$0.45 \pm 0.02$	$0.13 \pm 0.05$	$1.22 \pm 0.08$	2
$4e$	$0.45 \pm 0.04$	$0.26 \pm 0.02$	$0.59 \pm 0.12$	$1.30 \pm 0.13$	2
Total	$2.58 \pm 0.25$	$1.65 \pm 0.09$	$1.39 \pm 0.26$	$5.62 \pm 0.37$	8
$\sqrt{s} = 8 \text{ TeV}$					
$4\mu$	$5.81 \pm 0.58$	$3.36 \pm 0.17$	$0.97 \pm 0.18$	$10.14 \pm 0.63$	13
$2\mu 2e$	$3.00 \pm 0.30$	$1.59 \pm 0.10$	$0.52 \pm 0.12$	$5.11 \pm 0.34$	8
$2e 2\mu$	$3.72 \pm 0.37$	$2.33 \pm 0.11$	$0.84 \pm 0.14$	$6.89 \pm 0.41$	9
$4e$	$2.91 \pm 0.29$	$1.44 \pm 0.09$	$0.52 \pm 0.11$	$4.87 \pm 0.32$	7
Total	$15.4 \pm 1.5$	$8.72 \pm 0.47$	$2.85 \pm 0.39$	$27.0 \pm 1.6$	37

TABLE 3.3: Expected signal and background yields together with number of observed events in data in the  $115 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$  signal region. The number of signal events is given for a Higgs boson mass of  $125.5 \text{ GeV}$ [5].

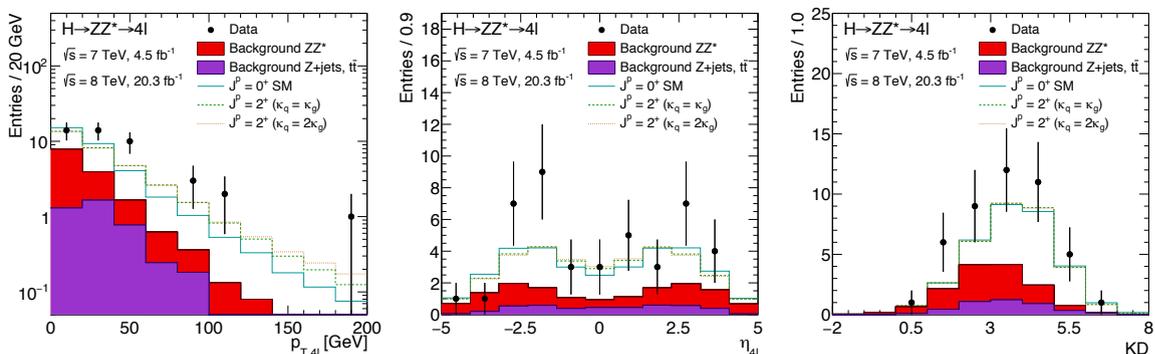


FIGURE 3.4: Expected and observed production distributions after event selection for events falling in the signal mass range  $115 \text{ GeV} \leq m_{4\ell} \leq 130 \text{ GeV}$ . The expected contributions from a Standard Model Higgs and Spin-2 bosons with different quark/gluon couplings are shown separately. Figures show from left to right  $p_T$ ,  $\eta$  and the kinematic discriminant  $KD$ .

$ZZ^*$ -continuum are shown together with SM and alternative signal distributions. The Standard Model is compared to alternative spin-2 models in the first set of figures and compared to the  $0^-$  pseudo-scalar in the second set. The kinematic discriminant will be described in Section 3.3.

It is clear that it is not possible to tell by eye which model is preferred from the figures alone. More advanced methods will have to be used that take into account the full correlation between the observables.

### 3.2.4 Systematic Uncertainties

The expected distributions and normalisations will be associated with uncertainties. Partially due to the finite statistics size and partially due to imperfect knowledge. The latter type

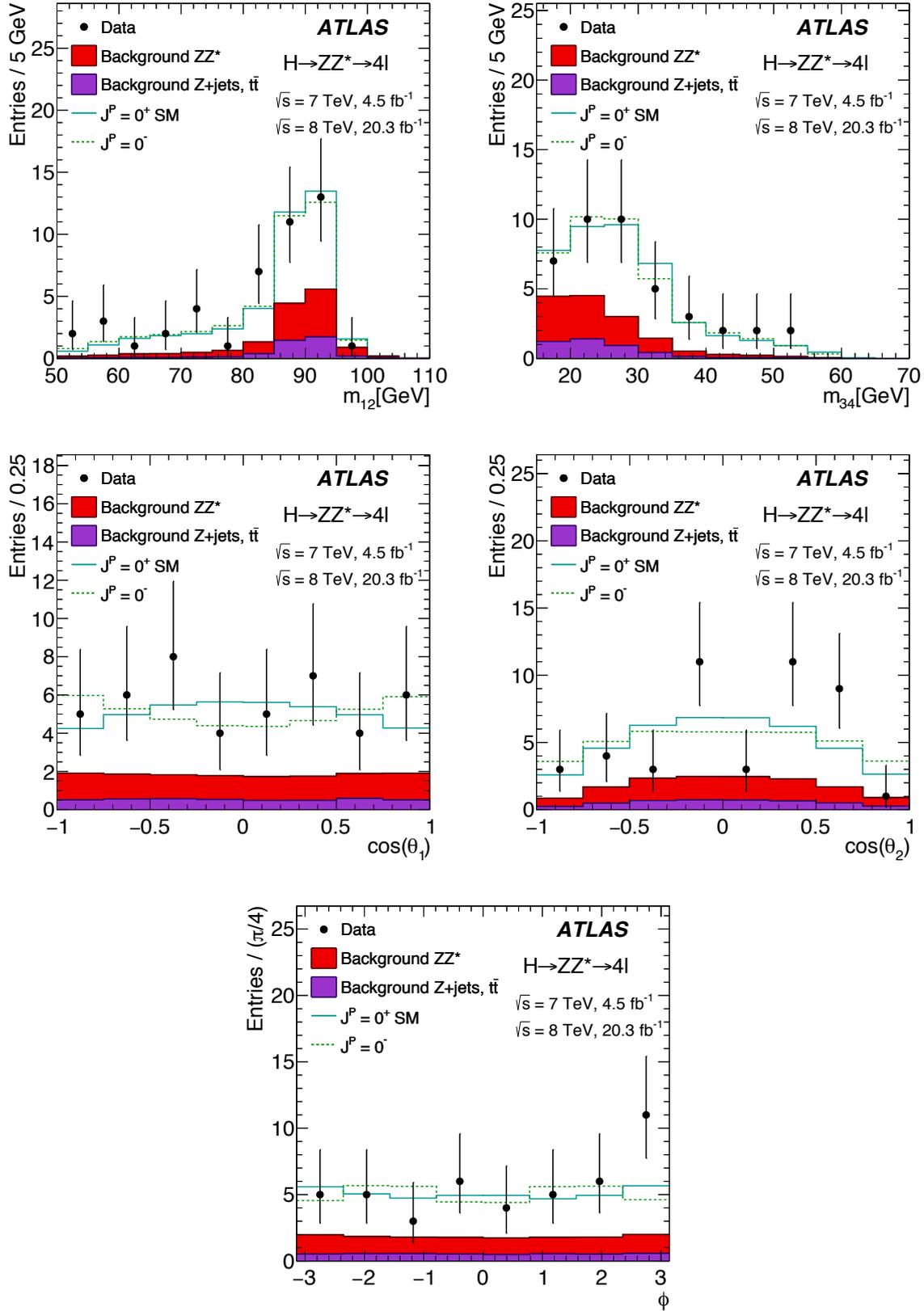


FIGURE 3.5: Expected and observed angular distributions after event selection for candidates falling in the signal mass range:  $115 \text{ GeV} \leq m_{4\ell} \leq 130 \text{ GeV}$ . The expected contributions from a Standard Model and pseudo-scalar boson are shown separately. The figure show in order,  $m_{12}$ ,  $m_{34}$  (top),  $\cos \theta_1$ ,  $\cos \theta_2$  (centre),  $\phi$  (bottom)[5].

that can lead to systematic uncertainties will be described here. The individual analyses will also have uncertainties associated with them from the way they are constructed, but the vast majority will fit into two categories; theoretical and experimental uncertainties. The different types of effects that can influence the analyses are collected here. The ATLAS collaboration uses a common naming scheme for systematic uncertainties, which will be used in the analyses descriptions. The systematic uncertainties are here listed according to this scheme and fall into the following categories:

### Theoretical:

**alpha\_BR\_VV:** The  $H \rightarrow VV$  branching ratio. A 4% uncertainty is assigned to the signal normalisation for  $H \rightarrow ZZ^* \rightarrow 4\ell$ .

**alpha\_pdf\_Higgs\_ggH:** Parton density uncertainty for signal production. Affects the signal normalisation with a variability of 7%.

**alpha\_pdf\_qq:** Parton density uncertainty for  $ZZ^*$  production. Normalisation uncertainty of 4%.

**alpha\_QCDscale\_Higgs\_ggH:** Renormalisation and factorisation scale of signal production. Normalisation is allowed to vary 8% down and 7% up.

**alpha\_QCDscale\_VV:** Renormalisation and factorisation scale of background production. The effect is estimated to be 5% at  $\sqrt{s} = 7$  TeV and 3% at  $\sqrt{s} = 8$  TeV of the normalisation.

### Experimental:

**alpha\_Lumi\_2011, alpha\_Lumi\_2012:** Uncertainty in the luminosity collected by ATLAS for the 2011 and 2012 datasets respectively. The uncertainty will affect the signal and background normalisation in a correlated way. The uncertainties are estimated to be 1.8% for  $\sqrt{s} = 7$  TeV and 2.8% for  $\sqrt{s} = 8$  TeV.

**alpha\_ATLAS\_MU\_\*:** Uncertainties associated with muons. The uncertainties include efficiency, identification, resolution and trigger performance. Four nuisance parameters are assigned to the spectrometer and one to the trigger. The uncertainties affect both shape and normalisation.

**alpha\_ATLAS\_EM\_\*:** Uncertainties associated with the electromagnetic calorimeter describing the energy scale and resolution. The uncertainties take among other into account pileup effects and material uncertainties. 21 nuisance parameters are assigned affecting both shape and normalisation.

**alpha\_ATLAS\_EL:** Six uncertainties related to electron reconstruction and identification affecting both signal and normalisation variations.

### Analysis:

`alpha_ATLAS_norm_SF_H41_Z_11ee_2011,2012` Reducible background normalisation uncertainty for the  $4e$  and  $2\mu 2e$  final states. The following uncertainties are assigned,  $4e$ -2011: 2.0%,  $2\mu 2e$ -2011: 2.3%,  $4e$ -2012: 2.1%,  $2\mu 2e$ -2012: 2.3%.

`alpha_ATLAS_norm_SF_H41_Zbb_11mumu_2011,2012` Reducible background normalisation uncertainty for the  $4\mu$  and  $2e 2\mu$  final states. The following uncertainties are assigned,  $4\mu$ -2011: 4.3%,  $2e 2\mu$ -2011: 3.8%,  $4\mu$ -2012: 1.9%,  $2e 2\mu$ -2012: 1.7%.

`alpha_ATLAS_shape_SF_H41_Z_11ee_2011,2012` Uncertainty in the reducible background distributions of the  $4e$  and  $2\mu 2e$  final states.

`alpha_ATLAS_shape_SF_H41_Zbb_11mumu_2011,2012` Uncertainty in the reducible background distributions of the  $4\mu$  and  $2e 2\mu$  final states.

`alpha_ATLAS.*` Analysis specific uncertainties will be described in their individual sections.

The systematic uncertainties are included in the likelihood model of each analysis that will be presented. Normalisation uncertainties are included by pdf-scaling. Shape uncertainties are described by pdf-interpolation between a nominal distribution and two templates that are estimated to represent the one standard deviation variation in either direction of the parameter. A Gaussian constraint term is moreover multiplied on the likelihood to reflect the preference of the nominal parameter value and limiting the model from straying far from expectations. Denoting the parameter of interest  $\alpha$ , constraint term  $C$  and set of nuisance parameters  $\theta$  the likelihood becomes:

$$\mathcal{L}(\alpha) \rightarrow \mathcal{L}(\alpha, \theta) \cdot C(\theta), \quad C_i = e^{-\theta_i^2/2},$$

where  $\theta = \pm 1$  correspond to variations described above. A separate constraint term is added for each nuisance parameter.

An effect of constraining the likelihood this way is, that if an unimportant parameter is included the uncertainty on the fitted parameter will be 1. The reason for this is if the parameter is unimportant for the analysis, the unconstrained likelihood is not going to vary with it. The constrained likelihood on the other hand will be maximised at the nominal nuisance parameter value due to the Gaussian constraint term. The uncertainty on the parameter is given by how much it should be changed before the log-likelihood ratio has grown by one:

$$1 = -2 \ln \lambda = -2 \ln \frac{\mathcal{L}(\alpha, \pm \sigma_\theta) \cdot C(\pm \sigma_\theta)}{\mathcal{L}(\hat{\alpha}, 0) \cdot C(0)} \simeq -2 \ln e^{-(\pm \sigma_\theta^2)/2} \Rightarrow \theta = \pm 1$$

If the nuisance parameter fits to zero with unit uncertainty it can be taken as indication that it is unimportant for the analysis.

### 3.3 Fixed Hypothesis Tests

The models that will be examined have been motivated and described, the experiment run and selection of events imposed. It is now time to determine if the observations are in agreement with the Standard Model or something more exotic.

The first part of the analysis focuses on the fixed hypothesis tests, where the Standard Model is tested against alternative BSM scenarios one by one. Keeping in mind the description of models in Section 1.2.3, the tests can be outlined as follows:

First the Standard Model is tested against other spin-0 models; the positive parity, higher dimensional operator boson  $0_h^+$  and the pseudo-scalar  $0^-$ , Table 1.2. Following the spin-0 tests, the Standard Model is tested against spin-2 models with different QCD coupling configurations, see Table 1.3. The spin-0 and spin-2 tests will constitute the main result of the fixed hypothesis section. Two tests of a spin-1 CP-even and CP-odd boson using the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay chain will also shortly be mentioned. The resonance decay to two photons however rules out this possibility.

From the way the models were defined in Section 1.2.4 only the five observables describing the decay will be influenced by the specific spin-0 model. Spin-2 (and spin-1) production will however in general differ due to the different QCD couplings. The section will describe how final state kinematics can be used to separate models<sup>2</sup>.

#### 3.3.1 Distinguishing Models with Boosted Decision Trees

For each hypothesis test, two Boosted Decision Trees (BDT) are used. One is optimised to separate the two  $J^p$  hypotheses under question and the other optimised to separate a Standard Model Higgs Boson ( $J^p = 0^+$ ) from the  $ZZ^*$ -continuum background. This single background discriminant is used for all hypothesis tests. The dependency between the angular and background discriminants will be described using two-dimensional templates. The final state observables used for the angular and kinematic BDT discriminants are as follows:

- $0^+$  vs.  $0^-$  and  $0^+$  vs.  $0_h^+$  separation:  $m_{12}, m_{34}, \cos(\theta_1), \cos(\theta_2), \phi$ .
- $0^+$  vs.  $2^+$  separation:  $m_{12}, m_{34}, \cos(\theta_1), \cos(\theta_2), \phi, \cos(\theta^*), \phi_1$ .
- $0^+$  vs.  $ZZ^*$  separation:  $m_{4l}, p_{T-4l}, \eta_{4l}, KD$ .

where  $m_{4l}$ ,  $p_{T-4l}$ , and  $\eta_{4l}$  are the mass, transverse momentum, and pseudorapidity of the four-lepton system respectively. The observables are defined in Figure 1.8.

<sup>2</sup>On a personal note, I have chosen to divide this section into two parts. First, the analysis in its final version is described which is followed by a short description of the earlier implementation. During my Ph.D. studies I took part in the full completion of the first version but changed to focus on the tensor structure fit that will be described later, after the initial stages of the second iteration. During these I was involved in the BDT optimisation, shift to using two-dimensional KDE templates and testing the statistical procedure that will be described. After I switched to the tensor structure analysis I also wrote the documentation of the second iteration.

The  $KD$ , an abbreviation of Kinematic Discriminant, is defined by the ratio of Higgs and  $ZZ^*$ -continuum matrix elements:

$$KD = \log \left( \frac{ME_H^2}{ME_{ZZ}^2} \right)$$

The Matrix elements (ME) are calculated at leading order using MADGRAPH[49]. The matrix element take into account the full final state structure and offers extra discriminative power against the background.

The final state observables that the boosted decision trees are trained on are summarised in Table 3.5. The distributions used in the different trainings are seen in Figure 3.5 for the  $0^+$  and  $0^-$  states and Figure 3.4 for the  $0^+$  state and different  $2^+$  models. The figures do not do the problem full justice as separation also arises from parameter correlations.

The training is performed using the TMVA tool for multivariate analysis[89]. Other types of multivariate approaches have been tested, like the Neural Network and Fishers Linear Discriminant. The BDT approach is found to give the optimal result, as it is able to describe the non-linear correlations between the final state observables without introducing significant overtraining.

The  $J^P$  discriminating BDTs are constructed from between 200 and 800 trees, each having a maximal depth of 3. The background discriminant is trained with between 300 and 900 trees with a maximal depth of 3 to 5. In all cases, the BDTs are optimised using gradient boosting and bagging re-sampling as described in [89, Sect. 7]. A description of how BDT separation is obtained is moreover given in App. A.4. Finally, a detailed list of the setup can be found in Tab. 3.4. All configurations are chosen to give the maximal separation without creating overtrained discriminants.

Separated Hypotheses	$N_{\text{Trees}}$	Shrink.	MNP	S/B	BSF	$N_{\text{Cuts}}$	Max depth
$J^P = 0^+, 0^-$	200	0.10	0.5	1	0.6	20	3
$J^P = 0^+, 0_h^+$	800	0.03	0.6	1	0.8	20	3
$J^P = 0^+, 2_m^+$	200	0.10	0.5	1	0.6	20	3
$J^P = 0^+, ZZ\text{-continuum}$							
4e	300	0.03	0.6	1.5	0.6	20	3
4 $\mu$	900	0.05	0.6	1.5	0.6	20	5
2e2 $\mu$	300	0.03	0.6	1.5	0.6	20	3
2 $\mu$ 2e	300	0.02	0.6	1.5	0.6	20	3

TABLE 3.4: Parameters used for the BDT training. The definitions of the options are listed in Table 22, 23 and 24 of [89]. The names used here are translated to the source as follows:  $N_{\text{Trees}}$ : NTrees, Shrink.: Shrinkage, MNP: MinNodeSize, S/B: SigToBkgFraction, BSF: BaggedSampleFraction,  $N_{\text{Cuts}}$ : nCuts, Max depth: MaxDepth. All forests of trees are moreover boosted using bagging re-sampling.

For each separation, the MC samples of the two hypotheses are divided into two sub-samples of equal size, whose entries are chosen at random. The BDT training is performed on one of the partitions, the training sample. To ensure that the discriminant training did not optimise separation on differences caused by statistical fluctuations, the resulting shape is compared

to the second half of the data samples. To evaluate the effect of overtraining the resulting ROC-integral is computed for both the test sample and training sample.

The ROC-curve (Receiver Operating Characteristic) describes the connection between an observable's signal efficiency and background rejection power. In other words, what the rate of type-I error is for any given type-II error.

Here the ROC-curve is defined such that an integral of 1.0 indicates perfect separation while 0.5 indicates random selection between the two samples:

$$\text{ROC} : 1 - \epsilon_{\text{bkg}} \text{ vs } \epsilon_{\text{sig}}$$

If the curve is defined in this way the integral can be interpreted as the average probability of correctly classifying a randomly selected event. It thus provides a good measure of the separation power. For each separation a  $\sim 1\%$  difference or below is found between the test and training samples, see Table 3.5.

The discriminating distributions that will be used in this analysis are shown in Figure 3.6. By combining information from the different observable distributions a preference for SM can be seen.

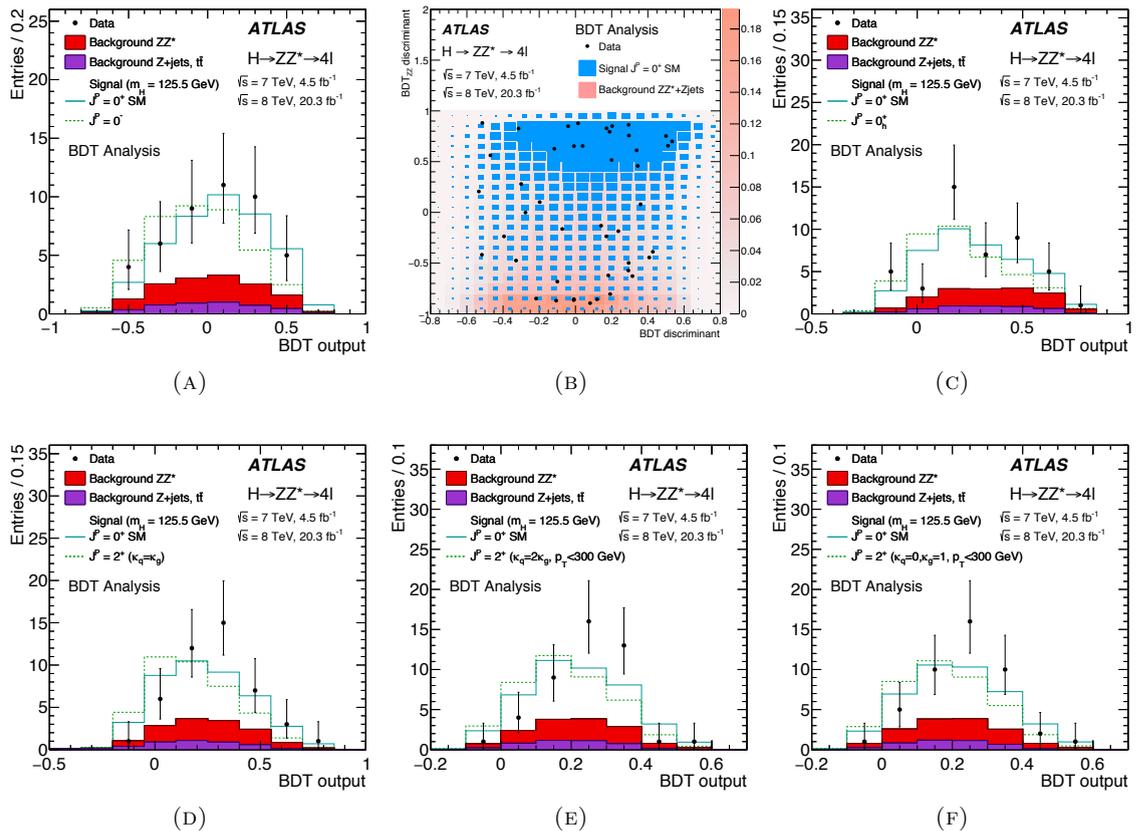


FIGURE 3.6: BDT discriminant distributions for data, backgrounds and signal models. The top row show the spin-0 models: (A)  $0^+$  vs.  $0^-$  separation. (B) correlation between  $0^+$  vs.  $0^-$  and background discriminant. (C)  $0^+$  vs.  $0_h^+$ . Bottom row shows spin-2 discriminants: (D) Universal coupling model ( $\kappa_q = \kappa_g$ ). (E) small gluon coupling. (F) No quark coupling. The non-universal spin-2 distributions have a transverse momentum cut-off of 300 GeV[5].

Separation	Training parameters	$\int \text{ROC}_{\text{Train}}$	$\int \text{ROC}_{\text{Test}}$	Difference
$\sqrt{s} = 8 \text{ TeV}$				
$J^P = 0^+, 0^-$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi$	0.681	0.677	0.6 %
$J^P = 0^+, 0_h^+$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi$	0.630	0.624	1.0 %
$J^P = 0^+, 2^+$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi, \cos(\theta^*), \Phi_1$	0.629	0.622	1.1 %
$J^P = 0^+, ZZ^*$				
4e		0.887	0.884	0.3 %
4 $\mu$	$m_{4l}, p_{T-4l}, \eta_{4l}, \text{KD}$	0.891	0.886	0.6 %
2e2 $\mu$		0.887	0.884	0.3 %
2 $\mu$ 2e		0.885	0.882	0.3 %
$\sqrt{s} = 7 \text{ TeV}$				
$J^P = 0^+, 0^-$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi$	0.683	0.680	0.4%
$J^P = 0^+, 0_h^+$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi$	0.629	0.628	0.2%
$J^P = 0^+, 2^+$	$\cos(\theta_1), \cos(\theta_2), m_1, m_2, \Phi, \cos(\theta^*), \Phi_1$	0.627	0.625	0.3%
$J^P = 0^+, ZZ^*$				
4e		0.878	0.875	0.3%
4 $\mu$	$m_{4l}, p_{T-4l}, \eta_{4l}, \text{KD}$	0.890	0.881	0.1%
2e2 $\mu$		0.876	0.878	0.2%
2 $\mu$ 2e		0.873	0.876	0.3%

TABLE 3.5: BDT summary table. The columns show in consecutive order the two models for any training, the final state observables used as training parameters and the resulting ROC integral for the sample the training has been performed on and a statistical independent sample of equal size. The last column show the difference in ROC-integral. The difference indicates the size of overtraining and is seen to be at most at the percentage level.  $2^+$  numbers are from the  $\kappa_q = \kappa_g$  model.

### 3.3.2 PDF Construction

For any given signal model and background, the discriminant separating the two spin-parity states and the background discriminant is correlated to some degree. It is therefore necessary to construct two-dimensional *pdfs* that are able to describe any non-trivial correlation to its full degree.

For each hypothesis test, the  $J^P$ -discriminant and background separating BDT distributions are described with a two-dimensional histogram of suitable binning ( $40 \times 40$  bins). The density of each model is estimated with simulated data. By construction, the BDT values are confined to the range  $[-1, 1]$ .

Populating a histogram of 1600 bins with simulated data would however require an unreasonably large amount of statistics. The *pdfs* are instead estimated from the histograms with a Gaussian Multivariate Kernel Density Estimation (KDE) using the RooNDKeysPDFs package[90]. A description of the KDE method is given in App. A.2. The multivariate KDE is performed with an adaptive bandwidth. Different choices of KDE bandwidths ( $h$  in Eq. A.5)

have been examined, where it was found that scaling the nominal  $h$  (Eq. A.6) with 0.35 gives the optimal description. The same procedure is done for signal,  $ZZ^*$ -continuum and reducible background processes separately.

All data are divided into the four final states  $4e$ ,  $4\mu$ ,  $2e2\mu$ , and  $2\mu2e$  and centre of mass energies  $\sqrt{s} = 7, 8$  TeV before having their parent distribution estimated with the KDE method. The KDE is afterwards projected into the final histogram. As an example, the hypothesis test separating the positive and negative spin-0 parity states will be based on eight *pdfs* for each of the  $J^p = 0^+$ ,  $J^p = 0^-$ ,  $ZZ^*$ -continuum, and reducible background process types. These are all subsequently created in versions reflecting the systematic changes described below.

### 3.3.3 Statistical Treatment

In the fixed hypothesis analysis, pairs of different spin and parity models are tested against each other. A maximum likelihood fit is performed on the events falling in the signal mass region. The likelihood model for any given  $J^p$  hypothesis is of the form:

$$\mathcal{L}(\bar{N}|J^p, \mu, \theta) = \prod_j^{N_{\text{chan}}} \prod_i^{N_{\text{bin}}} P(N_{i,j} | \mu_j \cdot S_{i,j}^{(j^p)}(\theta) + B_{i,j}(\theta)) \cdot \mathcal{C}(\theta), \quad (3.1)$$

where  $P$  is a Poissonian distribution evaluated at the observed  $N_{i,j}$  events in observable bin  $i$ , final state  $j$  depending on flavor of the four leptons and centre of mass energy. The expectation values are symbolised by  $\mu \cdot S$  and  $B$ . The  $S_{i,j}$  ( $B_{i,j}$ ) value describes the signal (background) probability density function evaluated in observable bin  $i$ , final state  $j$ , scaled by the integral number of signal (background) events. The background is a sum of the  $ZZ^*$ -continuum and reducible backgrounds.

The nuisance parameter  $\mu$ , the signal strength, have been emphasised from the remaining set,  $\theta$ . In the maximal likelihood fit the signal strength is freely allowed to change, modifying the signal normalisation. The background normalisations are on the other hand contained  $\theta$ . The likelihood is moreover multiplied by a Gaussian constraint term  $\mathcal{C}(\theta)$  to constrain the systematic variations to their expected ranges.

The ratio of profiled likelihoods is used as test statistic. It reads, as an example, for the hypothesis test of  $J^p = 0^+$  and  $J^p = 0^-$ :

$$-2 \ln \Lambda = -2 \ln \frac{\mathcal{L}(J^p = 0^+, \hat{\mu}_{0^+}, \hat{\theta}_{0^+})}{\mathcal{L}(J^p = 0^-, \hat{\mu}_{0^-}, \hat{\theta}_{0^-})}, \quad (3.2)$$

where  $\hat{\mu}$  and  $\hat{\theta}$  symbolises the maximum likelihood estimator of  $\mu$  and  $\theta$  for the individual likelihoods. The test statistic for all other hypotheses are constructed similarly.

A large ensemble of Monte Carlo pseudo experiments is used to estimate the expected test statistic distribution of the different hypotheses. The  $p_0$  value for any hypothesis follows directly as the tail integral of the distribution starting at the data test statistic value. Expectations for different signal models are calculated using the median value of the test statistic distributions.

### 3.3.4 Systematic Uncertainties

Dominant systematic uncertainties for hypothesis test	
Nuisance parameter	Average effect on separation
Higgs Mass uncertainty, $\sigma_m = 500$ MeV <code>alpha_ATLAS_Higgs_mass</code>	2.77%
Muon spectrometer resolution <code>alpha_ATLAS_MU_MS_RES_MS</code>	0.87%
Reducible background normalisation in the $2e2\mu + 4\mu$ final states <code>alpha_norm_SF_H41_Zbb_11mumu_2012</code>	0.78%
Constant term of EM calorimeter resolution <code>alpha_EM_mRes_CT</code>	0.71%
$q\bar{q}$ parton density function for $ZZ^*$ <code>alpha_pdf_qq</code>	0.64%
Reducible background normalisation in the $4e + 2\mu 2e$ final states <code>alpha_ATLAS_norm_SF_H41_Z_11ee_2012</code>	0.55%
Background renormalisation and factorisation QCD scale <code>alpha_QCDscale_VV</code>	0.48%
Sampling term of EM calorimeter resolution <code>alpha_EM_mRes_ST</code>	0.46%
Luminosity uncertainty <code>alpha_ATLAS_LUMI_2012</code>	0.45%

TABLE 3.6: Dominant systematic uncertainties for the hypothesis test analysis and effect of each on the  $J^P = 0^+$  vs  $0^-$  separation. The effect of each nuisance parameter is calculated as the difference between the result when including all systematic uncertainties, and when each is shifted one standard deviation to either side. The average effect is calculated from the absolute value of the upward and downward shift. A threshold of 0.45% is chosen as to where a systematic uncertainty has a potential impact on the result. The full list of systematic uncertainties can be found in Table B.1, Appendix B.

Systematic effects are described by either generating a dedicated Monte Carlo sample or assigning a set of event weights to the nominal sample. The variations are created for each systematic uncertainty separately. The procedure described above of filling two dimensional histograms and performing a KDE is repeated for each systematic change that affect distributions. The shapes corresponding to systematic changes are finally added to the likelihood model before the fitting procedure. The majority of the systematic uncertainties in the BDT spin and parity analysis are equivalent to those outlined in Section 3.2.4.

In addition to these, the construction of the BDT shape will add uncertainties. The additional systematic changes related to the BDT discriminant that have been considered are as follows:

**BDT Overtraining :**

Finite statistic in the BDT training may potentially have an impact on the MVA output. The MVA can optimise separation on statistical fluctuations present in the training sample that are absent or different in the testing sample.

In order to assess the magnitude of this effect, the expected separations between spin and parity hypotheses are compared for the combined training and test sample (as in the analysis) and for the separate samples. No difference is observed.

This is also confirmed by the insignificant difference in ROC integrals presented in Tab. 3.5

### Signal mass modelling :

`alpha_ATLAS_Higgs_mass`

Mass resolution and mis-modelling of the reconstructed Higgs boson mass has been taken as an additional source of systematic uncertainty. It is worth noting that the four lepton mass is the strongest separating parameter used in the background discriminant. The mass and the background discriminant are therefore correlated. A shift of 500 MeV in the signal four lepton mass is applied to take this uncertainty into account. The change in observable distributions are found to influence the results while changes in the signal yield do not affect the result.

In general, any systematic uncertainty that only affects the combined signal normalisation can be absorbed in the signal strength. Only distribution changes are considered for the final estimates.

### Kernel Density Estimation :

`alpha_ATLAS_rho_signal`, `alpha_ATLAS_rho_Redbkg`, `alpha_ATLAS_rho_ZZ`

It is likely that different choices of bandwidth of the KDE would have resulted in equally valid descriptions while at the same time showing slight shape differences. To estimate the uncertainty introduced by the KDE procedure the bandwidth is changed from the nominal 0.35 to 0.30 and 0.40. The same procedure is applied to signal,  $ZZ^*$ -continuum and reducible background shapes. A separate nuisance parameter is assigned to each.

The effect of each nuisance parameter is subsequently estimated. This is done by calculating the profiled likelihood ratio between the  $J^p = 0^+$  and  $J^p = 0^-$  hypotheses when all nuisance parameters are included and when a single one is shifted 1 standard deviation to either side. The likelihood ratio value is for this purpose interpreted as a significance.

The difference in significance when including all nuisance parameters and when shifting to  $\pm 1\sigma$  is used as a measure of relevance of each systematic uncertainty. It is of great interest to rank the systematic uncertainties, both because it is important to demonstrate that the impact follows expectations, but also such that any irrelevant parameters can be discarded. The computing time for the final significance should be mentioned is strongly dependent on the number of nuisance parameters.

A threshold of 0.45% is chosen as to where a parameter has potential relevance for the analysis. The systematic uncertainties that are found above this threshold for the hypothesis

test analysis are summarised in Table 3.6 while the full list can be found in Table B.1 of Appendix B.

### 3.3.5 Summary of the BDT strategy

- At the first stage of the analysis, the full event selection is applied on all the signal Monte Carlo samples for different spin and parity models and to the  $ZZ^*$  background samples simulated in the  $100 \text{ GeV} < m_{4\ell} < 150 \text{ GeV}$  region. The reducible backgrounds are estimated from control-regions in data. The observables sensitive to the spin and parity of the underlying resonance are reconstructed for all the samples in the signal mass region:  $115 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$ . The obtained set of observables is used to train a set of BDT discriminants. A single background discriminant, used for all hypothesis tests, is constructed to discriminate against the  $ZZ^*$  background continuum. A  $J^P$  discriminant is trained to distinguish between one pair of spin and parity states. Discriminants to distinguish  $J^P = 0^+$  from  $J^P = 0^-, 0_h^+$  and  $2^+$  are constructed. Several models for QCD couplings are examined for the  $2^+$  state. The distributions of values of  $J^P$  discriminants applied to the signal and background samples are later used to compare to the observed data as described below.
- At the second stage, the simulated signal and background distributions for the angular and kinematic discriminants are described using two-dimensional histograms. All non-trivial observable correlations can be taken into account in this way. The underlying distributions are subsequently described using a Gaussian KDE to reduce the influence of statistical fluctuations. The procedure is done separately for all eight channels:  $4e$ ,  $4\mu$ ,  $2e2\mu$  and  $2\mu2e$ , each in versions of the two centre of mass energies 7 TeV and 8 TeV. A separate histogram is created for each systematic uncertainty affecting distributions.
- In the final stage, pairs of signal hypotheses are compared to each other. A large set of pseudo-experiments are performed, with a maximal likelihood fit done separately for each of the two signal models. The ratio of profiled likelihoods is used as a test statistic. The test statistic distribution is created for each of the two hypotheses. The expected exclusion, under the assumption of each signal model, is obtained from the tail integral of the test statistic distribution starting at the median of the alternative distribution. The test statistic is finally calculated in data and the observed exclusion of each of the tested models follow from the tail integral of its test statistic distribution.

### 3.3.6 Results

Pseudo experiments are finally created to derive the exclusion of each model. For each model in each hypothesis test, the test statistic of Eq. 3.2 is calculated for the pseudo experiments. First, expected results are calculated and later compared to observations. Presented in Table 3.7 are the  $p_0$ -values of expected and observed exclusion for various signal hypothesis with the combined 7 TeV and 8 TeV data-samples. The table present both tests of  $J^P = 0^-, 0_h^+$  and  $2^+$

assuming a Standard Model Higgs (Alt columns) and tests of the Standard Model Higgs under the alternative hypotheses (Null column). The probabilities to obtain the test statistic value under each hypothesis are labelled  $p_0$ . The  $p_0$  values are also translated to the corresponding number of Gaussian sigmas and to  $CL_s$  values ( $p_0(H_{Alt})/(1 - p_0(H_{Null}))$ ), see e.g. App. A.1 for a description).

The values are derived from MC toy distributions and as such have a related statistical uncertainty. The statistical uncertainty on the significances corresponds to at most 0.04 sigma and equivalently is of the size of the last reported digit of the  $p_0$ -values.

Data is seen to be in agreement with Standard Model expectations while the alternative tested models are excluded with high confidence. The assumed QCD coupling is seen to affect the observed result where the absence of quark couplings is seen to be the least excluded. The two spin-0 BSM hypothesis are assigned  $CL_s$  of 0.054 (CP-odd) and 0.10 (CP-even). The potential mixing of these to the SM Higgs will be the subject of the following section.

An equivalent analysis and its combination with the  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  and  $H \rightarrow \gamma\gamma$  decays is moreover described in [5].

$\sqrt{s} = 7 + 8 \text{ TeV}$						
Hypotheses	$p_0(J_{Null=SM}^P)$	$p_0(J_{Alt}^P)$		$\sigma(J_{Alt}^P)$		$CL_s(J_{Alt}^P)$
Null, Alt	Observed	Expected	Observed	Expected	Observed	Observed
$J^P = 0^+, 0^-$	0.78	0.001	0.012	3.10	2.26	0.054
$J^P = 0^+, 0_h^+$	0.63	0.016	0.037	2.14	1.78	0.10
-----						
$J^P = 0^+, 2^+$ with:						
$\kappa_q = \kappa_g$	0.23	0.009	0.0007	2.36	3.2	0.0009
$\kappa_q = 2\kappa_g, p_T < 125 \text{ GeV}$	0.38	0.036	0.017	1.80	2.12	0.027
$\kappa_q = 2\kappa_g, p_T < 300 \text{ GeV}$	0.23	0.018	0.0023	2.09	2.84	0.0030
$\kappa_q = 0, p_T < 125 \text{ GeV}$	0.49	0.040	0.037	1.75	1.78	0.071
$\kappa_q = 0, p_T < 300 \text{ GeV}$	0.58	0.018	0.030	2.10	1.88	0.071

TABLE 3.7: The expected and observed exclusion for the combined 7 TeV and 8 TeV dataset of different alternative spin and parity models for the hypothesis test analysis. The exclusions are given in terms of  $p_0$ -values and corresponding number of Gaussian sigmas. The Alt  $p_0$  is corrected for the SM  $p_0$  to give the  $CL_s$ . The stated values take into account all systematic uncertainties that are evaluated to be significant (impact  $\geq 0.45\%$ ). When calculating the expected numbers, all nuisance parameters have been profiled to the observed data.

### 3.3.7 Previous iterations of the $H \rightarrow ZZ^* \rightarrow 4\ell$ Fixed Hypothesis Tests

The results presented so far constitute the second iteration of the ATLAS Higgs spin/parity hypothesis tests. In this section the first iteration will be presented. The results were published in [3] with more details outlined in [2].

The analyses structure for the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay chain is much the same. Some minor differences are however worth mentioning:

- The first iteration analyses did not use the  $BDT_{ZZ}$  discriminant to suppress backgrounds, nor were distributions described with KDEs. Two signal regions are instead defined: A high signal to background fraction;  $m_{4\ell} \in [121 \text{ GeV}, 127 \text{ GeV}]$  and a low signal to background region;  $m_{4\ell} \in [115 \text{ GeV}, 121 \text{ GeV}] \cup [127 \text{ GeV}, 130 \text{ GeV}]$ .
- An algorithm is structured such that after each BDT is trained, Monte Carlo simulation of signal and backgrounds are filled into one-dimensional histograms. To ensure that there are no empty bins present, the maximal range where all templates are populated is found. While simpler, the approach is somewhat less sensitive due to the weaker background suppression.

Following this, the construction of the likelihood model and test statistic is the same.

A larger variety of models are moreover examined. Both the CP-even and CP-odd spin-1 boson are tested, see Section 1.2.3. The spin-2 minimal coupling model  $2_m^+$  that is included in the results is equivalent to the previous  $2^+$  with  $\kappa_q = \kappa_g$ . A negative parity spin-2 model is moreover tested. Finally, different fractions of quark and gluon cross sections for the positive parity spin-2 are examined.

The expected and observed discriminant distributions for testing  $0^-$ ,  $1^+$  and  $2^+$  against SM can be found in Figure 3.8. Monte Carlo pseudo experiments are created and the likelihood is maximised under each hypothesis, to calculate the ratio of profiled likelihoods. The distribution for each hypothesis can be found in the same figure.

The  $p_0$ -values, significances and  $CL_s$  values are listed in Table 3.8. The Standard Model is again concluded to be preferred over all alternative hypotheses.

One will notice when examining the table for hypothesis scan of different spin-2  $q\bar{q}$  fractions, that the test statistic also has a tendency to fall in the tail of the Standard Model distribution. The spin-2 models are more excluded than the Standard Model predicts. The SM mean and one, two  $\sigma$  bands for different fractions can be found in Figure 3.7b. To figure out where the deviation from Standard Model prediction originates from the initial decay distributions are examined. Figure 3.7a show the  $m_{34}$  distribution. A deviation in data is seen for low masses. It should be underlined that the tests are performed on the same data and the BDT optimisation is partly performed on the same Monte Carlo such that deviations should in no way be considered independent. None of the measurements moreover fall outside of the  $2\sigma$

band.

The deviation disappears as seen in Figure 3.5 after the updated selection described in Section 3.2

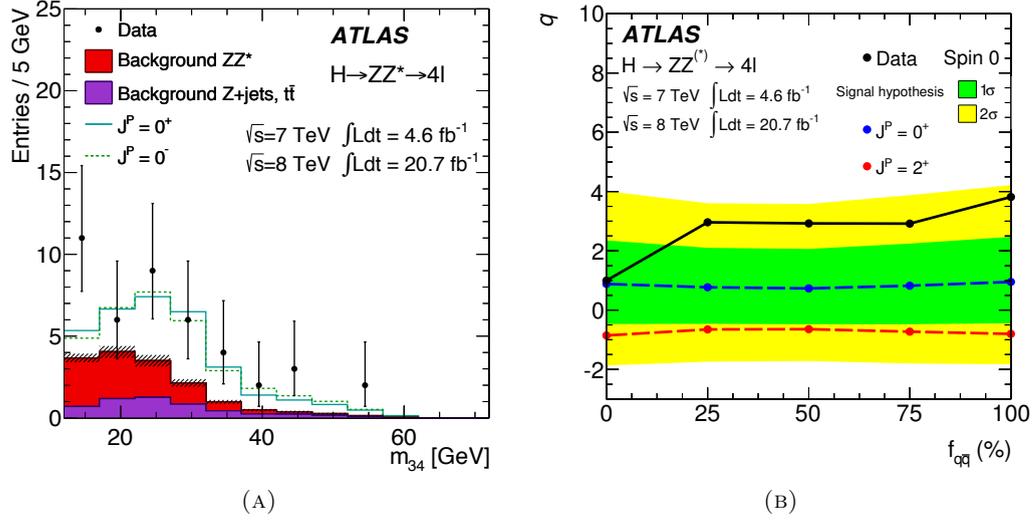


FIGURE 3.7: (A)  $m_{34}$  distribution of the “first iteration” event selection. Note the deviation in  $m_{34}$  for low values compared to Figure 3.5. (B) Test statistic distributions for different  $q\bar{q}$  fraction of spin-2 production. The Standard Model is preferred in all instances[3][2].

$\sqrt{s} = 7 + 8$ TeV						
Hypotheses	$p_0(J_{\text{Null=SM}}^P)$	$p_0(J_{\text{Alt}}^P)$		$\sigma(J_{\text{Alt}}^P)$		$\text{CL}_S(J_{\text{Alt}}^P)$
Null, Alt	Observed	Expected	Observed	Expected	Observed	Observed
$J^P = 0^+, 0^-$	0.31	0.0037	0.015	2.7	2.2	0.022
$J^P = 0^+, 1^+$	0.55	0.0016	0.0010	2.9	3.1	0.0020
$J^P = 0^+, 1^-$	0.15	0.0038	0.051	2.7	1.6	0.060
$J^P = 0^+, 2_m^+$	0.53	0.092	0.079	1.3	1.4	0.168
$J^P = 0^+, 2^-$	0.034	0.0053	0.025	2.6	2.0	0.258
$J^P = 0^+, 2^+(f_{q\bar{q}} = 100\%)$	0.962	0.082	0.001	1.4	3.1	0.026
$J^P = 0^+, 2^+(f_{q\bar{q}} = 75\%)$	0.923	0.099	0.003	1.3	2.7	0.039
$J^P = 0^+, 2^+(f_{q\bar{q}} = 50\%)$	0.943	0.113	0.002	1.2	2.9	0.035
$J^P = 0^+, 2^+(f_{q\bar{q}} = 25\%)$	0.944	0.107	0.002	1.2	2.9	0.036
$J^P = 0^+, 2^+(f_{q\bar{q}} = 0\%)$	0.532	0.092	0.079	1.3	1.4	0.169

TABLE 3.8: The expected and observed exclusion for the combined 7 TeV and 8 TeV dataset of different alternative spin and parity models for the “first iteration” hypothesis test analysis. The exclusions are given in terms of  $p_0$ -values and corresponding number of Gaussian sigmas. The Alt  $p_0$  is corrected for the SM  $p_0$  to give the  $\text{CL}_S$ . When calculating the expected numbers, all nuisance parameters have been profiled to the observed data[3][2].

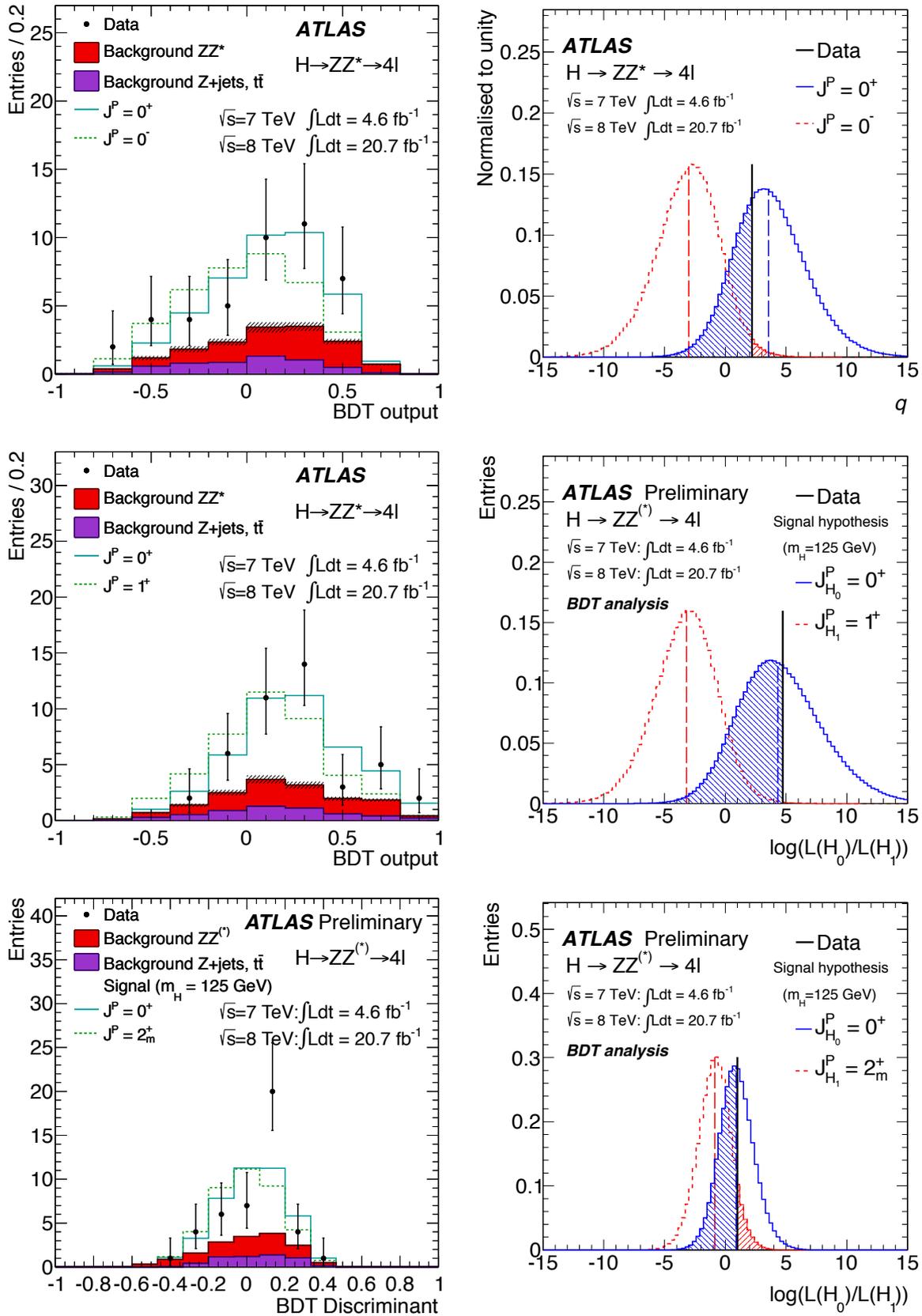


FIGURE 3.8: BDT discriminant (left) and test statistic distributions (right) for various models. Top:  $0^+$  vs  $0^-$ . Centre:  $0^+$  vs  $1^+$ . Bottom:  $0^+$  vs  $2_m^+$ . The Standard Model is favoured in all instances [\[3\]](#) [\[2\]](#).

### 3.4 The Tensor Structure Fit

In this section, the analysis adopted to examine mixed spin-0 models is presented. The different Higgs models are described in Section 1.2.4. In the fixed hypothesis tests of the previous section only spin-0 models with one coupling ( $\kappa_{SM}$ ,  $\kappa_{AVV}$ ,  $\kappa_{HVV}$ ) were considered. This section will generalise by examining mixed models, where two couplings are simultaneously present; ( $\kappa_{SM}$ ,  $\kappa_{HVV}$ ) and ( $\kappa_{SM}$ ,  $\kappa_{AVV}$ ). The latter combination will also take into account the CP-mixing angle  $\alpha$ .

The section will focus on how discriminants based on matrix element observables can be used to distinguish the examined signal models. This part will start with showing the variation of the angular distributions and  $Z$  masses for different coupling configurations in Section 3.4.1. In section 3.4.2 the Matrix Element Observable fit is described together with an evaluation of systematic uncertainties and the expected exclusion limits as estimated from MC simulation. It should be noted that in several places of the text there are references to complex coupling ratios. The possibility of imaginary contributions to the coupling ratios are not studied in this analysis. Monte Carlo with the presence of complex couplings was however generated before this choice was made. In all instances any imaginary contribution is always removed by Matrix Element based reweighting. A prospects study of complex coupling measurements will be described for high luminosity projections in Section 3.7.

#### 3.4.1 Coupling Ratios and Final State Observable Variations

The Lagrangian density that was introduced to describe a generic spin-0 boson contained three sectors. By examining shape variations, the analysis will measure which couplings of the density are necessary to describe the observed resonance behaviour.

It is known that different coupling configurations will lead to different production cross sections. By only studying coupling ratios and letting deviations in the observed number of events be absorbed in a free signal strength, the analysis removes this dependency. With more statistics and experience in performing the tensor structure fit, it could be useful to include fitted normalisation as well. For now the possibility of separating models based on normalisation will be ignored.

By solely examining shape variations it will only be necessary to measure the relative contributions of the different elements of the Lagrangian. The presented results will therefore be given as the ratio of BSM to SM couplings:

$$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}, \quad \tilde{\kappa}_{AVV} = \frac{1}{4} \frac{\nu}{\Lambda} \kappa_{AVV}$$

$$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha, \quad \tilde{\kappa}_{HVV} = \frac{1}{4} \frac{\nu}{\Lambda} \kappa_{HVV}$$

where:

$\kappa_{AVV}$  : Higher dimensional operator contribution.

$\kappa_{HVV}$  : Pseudo scalar contribution.

$\nu/\Lambda$  : Higgs field expectation value over EFT scale ( $\simeq 0.246$ ).

Figure 3.9 show the distributions of decay related final state observables in the signal mass range ( $115 \text{ GeV} \leq m_{A\ell} \leq 130 \text{ GeV}$ ), taking detector effects into account. The angular distributions and Z masses are shown for the SM Higgs together with four BSM scenarios;  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} \pm 1$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha \pm 5$ .

The analyses presented in this section also use production related observables to suppress backgrounds. The EFT expansion around SM that will be used only include the  $ZZ$  and  $WW$  vertices. The analysis solely consider the gluon fusion production mechanism such that distributions related to production are unaffected by the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ ,  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  composition.

Like the hypotheses tests, the tensor structure analysis use the selection described in Section 3.2.1.

### 3.4.2 The Matrix Element Observable Fit

The method of the Matrix Element observable fit is based on modelling the distributions of final state observables at different coupling ratios using Monte Carlo. The observables sensitive to the presence of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  considered in the analysis correspond to the first and second order Optimal Observables for a BSM amplitude with three component structure. Optimal Observables are described in [91–93]. The choice of observables will be discussed in Sect. 3.4.3. The observables are used for conditional  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  fits respectively. A kinematic BDT discriminant trained to suppress backgrounds is moreover used as an additional observable in all fits.

The Matrix Element Observable (ME-Obs) fit section is outlined as follows. In 3.4.3 a detailed description is given of how a  $pdf$  is constructed that will enable a measurement of the  $HVV$  tensor structure. The section will outline how matrix element based reweighting, KDEs and  $pdf$  interpolation is combined to finally arrive at a model that is continuous in the examined coupling ratios. The final likelihood model is defined in Section 3.4.4. The section is followed by a discussion on how the exclusion intervals at 95% CL are set and why this is justified.

In Section. 3.4.6 the tensor structure fit is demonstrated to be able to close on SM and alternative models that cover the examined intervals of coupling ratios under various conditions. Sect. 3.4.7 describes how systematic uncertainties are taken into account and present an estimate of how significant each source will be to the final result. Finally, the observed 95% CL limits are set and compared to expectations.

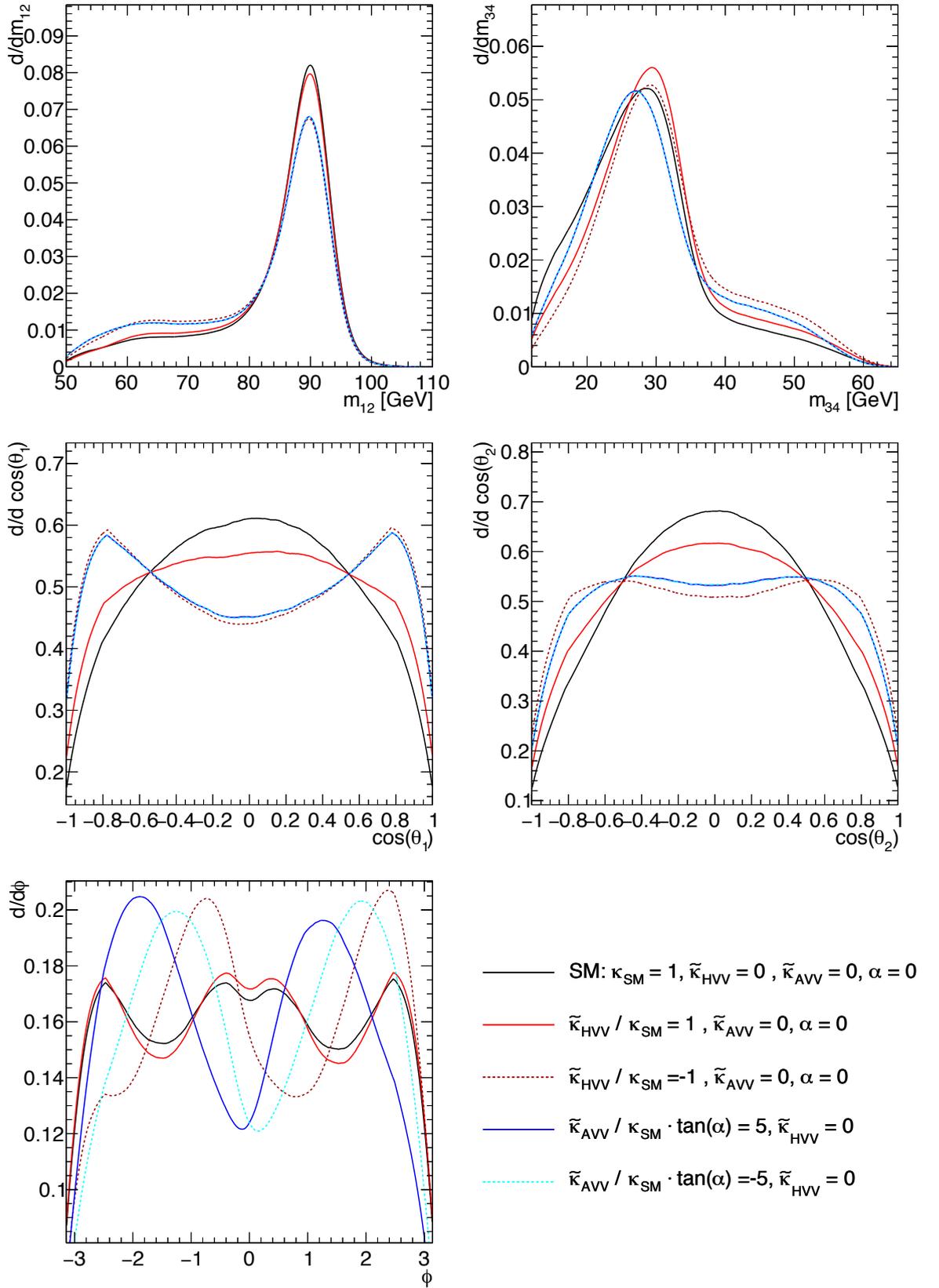


FIGURE 3.9: Distributions of decay observables for events passing selection and falling in the signal mass range:  $115 \text{ GeV} \leq m_{A\ell} \leq 130 \text{ GeV}$ . The expected contributions from the signal models,  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}} = \pm 1, \frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha = \pm 5$  and SM are presented.  $m_{12}, m_{34}$  (top).  $\cos \theta_1, \cos \theta_2$  (centre).  $\phi$  (bottom)

### 3.4.3 Pdf Construction

#### Observables

The coupling ratio sensitive observables used in the ME-Obs analysis are defined in the following paragraph. The technical implementation is based on the JHU matrix element calculator JHUGENMELA.v4.2.1[46] while the background discrimination is based on the TMVA BDT tool[89].

In the following, the shorthand notation  $ME(\theta)$  is used to describe the matrix element of a coupling ratio  $\theta$  ( $\theta = \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  or  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ ). The matrix element dependency on the particles four-momenta is left implicit.

To shortly motivate the form of the matrix element observables, it is useful to consider the EFT Lagrangian described in Section 1.2.4. Imagining only a single CP-even or CP-odd BSM contribution, the matrix element is of the form:

$$ME(\theta) \propto ME(\text{SM}) + \theta \cdot ME(\text{BSM}) ,$$

Where BSM refers to the pure, non-Standard Model term.

Squaring this, reveals an expression that is proportional to the differential probability of the process:

$$|ME|^2(\theta) \propto |ME|^2(\text{SM}) + \theta^2 \cdot |ME|^2(\text{BSM}) + \theta \cdot 2\mathcal{R}(ME(\text{SM}) \cdot ME^*(\text{BSM}))$$

The JHUGENMELA matrix element calculator does not contain the interference term:  $\mathcal{R}(ME(\text{SM}) \cdot ME^*(\text{BSM}))$ . If one however examines the expression for a coupling ratio of 1 it is seen that:

$$|ME|^2(\theta = 1) \propto |ME|^2(\text{SM}) + |ME|^2(\text{BSM}) + 2\mathcal{R}(ME(\text{SM}) \cdot ME^*(\text{BSM}))$$

giving an expression for the interference term solely described by squared matrix elements. Inserting into the expression for the squared matrix element for an arbitrary  $\theta$  and dividing by the squared Standard Model matrix element reveals:

$$\frac{|ME|^2(\theta)}{|ME|^2(\text{SM})} \propto 1 + \theta \cdot \left( \frac{|ME|^2(\theta = 1) - |ME|^2(\text{BSM})}{|ME|^2(\text{SM})} - 1 \right) + \theta^2 \cdot \frac{|ME|^2(\text{BSM})}{|ME|^2(\text{SM})}$$

The first and second order observables are defined as the two components:

$$O_1(\theta) = \frac{|ME|^2(\theta = 1) - |ME|^2(\text{BSM})}{|ME|^2(\text{SM})} \quad O_2(\theta) = \frac{|ME|^2(\text{BSM})}{|ME|^2(\text{SM})} \quad (3.3)$$

$O_i(\theta)$  is here used to symbolise the observable that is used to measure  $\theta = \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  or  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . By inserting the expression of the EFT coupling ratios, the observables defined in Eq. 3.4 emerges. The predicted distributions as observed by the ATLAS detector can be found in

Figure 3.10.

$$\begin{aligned}
O_1(\tilde{\kappa}_{HV V}) &= \frac{2\Re(ME(\kappa_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0) \cdot ME(\kappa_{HV V} \neq 0; \kappa_{SM}, \kappa_{AV V}=0; \alpha=0))}{|ME(\kappa_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0)|^2}, \\
O_2(\tilde{\kappa}_{HV V}) &= \frac{|ME(\kappa_{HV V} \neq 0; \kappa_{SM}, \kappa_{AV V}=0; \alpha=0)|^2}{|ME(\kappa_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0)|^2}, \\
O_1(\tilde{\kappa}_{AV V}, \alpha) &= \frac{2\Re(ME(\kappa_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0) \cdot ME(\kappa_{AV V} \neq 0; \kappa_{SM}, \kappa_{HV V}=0; \alpha=\pi/2))}{|ME(\tilde{\kappa}_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0)|^2}, \\
O_2(\tilde{\kappa}_{AV V}, \alpha) &= \frac{|ME(\kappa_{AV V} \neq 0; \kappa_{SM}, \kappa_{HV V}=0; \alpha=\pi/2)|^2}{|ME(\kappa_{SM} \neq 0; \kappa_{HV V}, \kappa_{AV V}=0; \alpha=0)|^2}.
\end{aligned} \tag{3.4}$$

The first two observables  $O_1(\tilde{\kappa}_{HV V})$  and  $O_2(\tilde{\kappa}_{HV V})$  are defined to have sensitivity to  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  and similarly  $O_1(\tilde{\kappa}_{AV V}, \alpha)$  and  $O_2(\tilde{\kappa}_{AV V}, \alpha)$  are defined with sensitivity to  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$ .

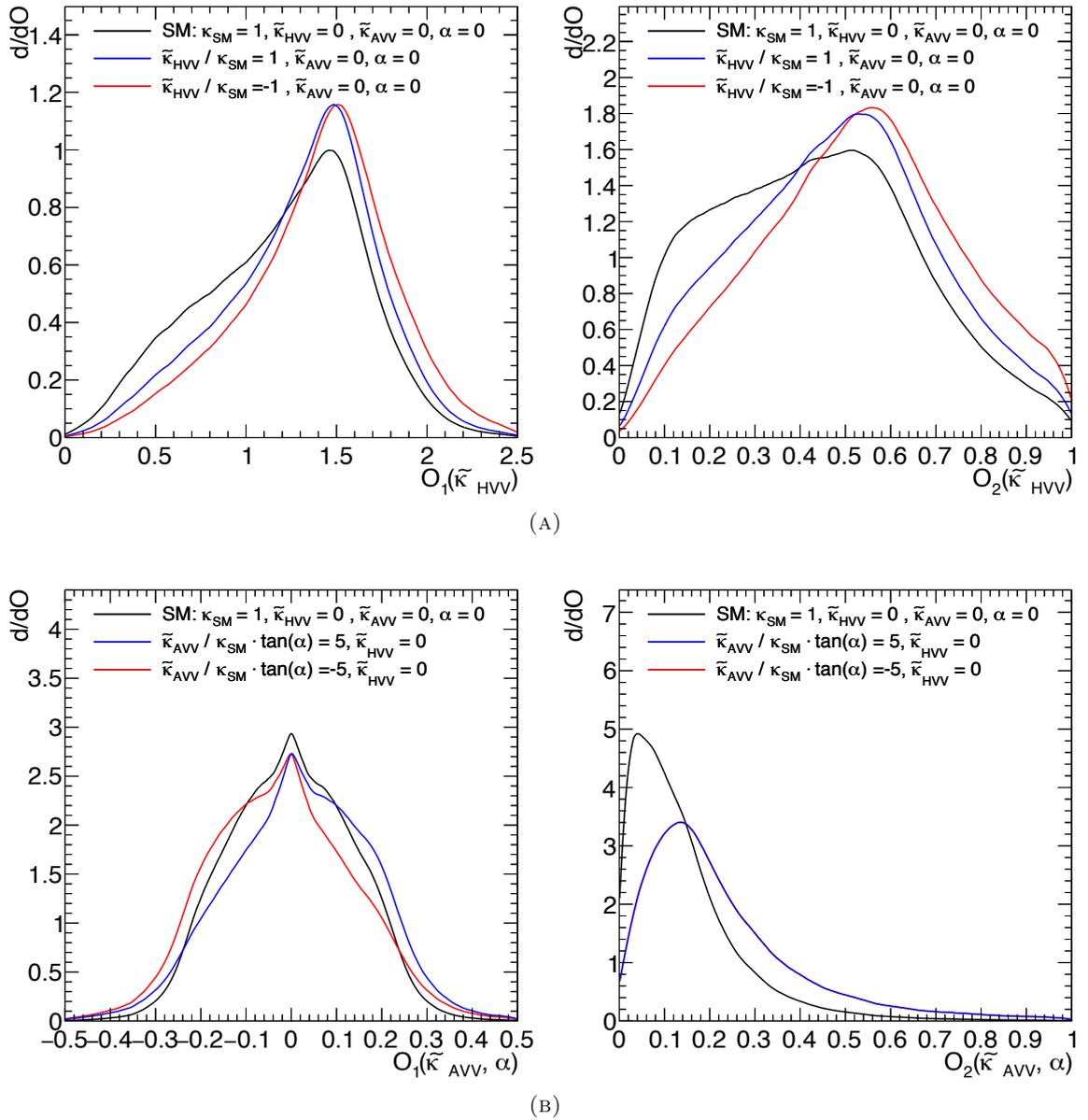


FIGURE 3.10: Distributions of Matrix Element Observables passing event selection. The expected contributions from the signal models,  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}} = \pm 1$ ,  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha = \pm 5$  and SM are presented. (A) Observables used in the  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  fit. (B)  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$  observables.

A Boosted decision tree is trained to help discriminate against backgrounds. The training is performed per final state, using 600 trees but otherwise keeps all default settings from the TMVA package [89, Tab. 21,22]. To be able to easier describe the multidimensional final state histogram it is chosen to only use production related parameters. This ensures that the background discriminant is as uncorrelated with the CP sensitive observables as possible and independent of the particular coupling ratio of a given signal model. The used parameters are as follows:

$$(m_{4l}, \quad p_{T-4l}, \quad \eta_{4l}, \quad \cos(\theta^*), \quad \Phi_1) \rightarrow BDT_{ZZ}$$

The training is only performed on events falling in the signal mass region  $115 \text{ GeV} < m_{4l} < 130 \text{ GeV}$ . This BDT discriminant will be denoted hereafter as the  $BDT_{ZZ}$ . The training is optimised to separate the signal model with coupling configuration  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 1+i$ ,  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 1+i$  from the  $qq \rightarrow ZZ^*$  background. The Monte Carlo samples used are described in Section 1.3. The specific signal model was chosen due to it having the largest MC sample.

Figure. 3.15 e.g. demonstrates that the  $BDT_{ZZ}$  distribution is not influenced by the choice of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  signal model. The background discriminant is finally divided into four bins before being used in the analysis.

The observable distributions are described using multi dimensional histograms. It can be difficult to populate these using simulated data, which motivate the use of KDE. As is customary, a Gaussian Kernel is chosen. To get the simplest possible input before the KDE step the observables are transformed to pseudo-Gaussian shapes. The derivation of the transformation is described in App. A.3.

The transformation was specifically defined such that the 1D projection of each observable is a Gaussian distribution for the Standard Model. A consequence of the "conservation of probability" that is used to derive the transformation is that the resulting likelihood function per construction is unchanged. Choosing the Standard Model to be the particular model that is exactly Gaussian will not bias the final result. Given that the distributions are described with histograms, the transformation can also be thought of as a motivated choice of binning. The  $BDT_{ZZ}$  distribution is moreover defined to be uniform for the Standard Model using a similar approach.

A simple fit of the  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  coupling ratio is performed on a simulated Standard Model Higgs to check that the pre- and post-transformed likelihoods are identical. The original and transformed distributions together with their respective likelihood curves can be found in Figure 3.11, where the likelihood values are seen to be identical. The transformed observable  $O$  will be denoted  $TO$  in the following.

Finally, it was found that the observable distributions related to  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  are to a very high degree linearly correlated. To get as easy a description as possible, the observables are decorrelated using  $TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV})$  and  $TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})$  instead<sup>3</sup>. Again,

<sup>3</sup>More precisely:  $0.5 \cdot (TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV}))$  and  $7.5 \cdot (TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV}))$  are used, such that the distributions will remain defined in the same interval. The figures do not include this in the notation since it is already plenty cluttered.

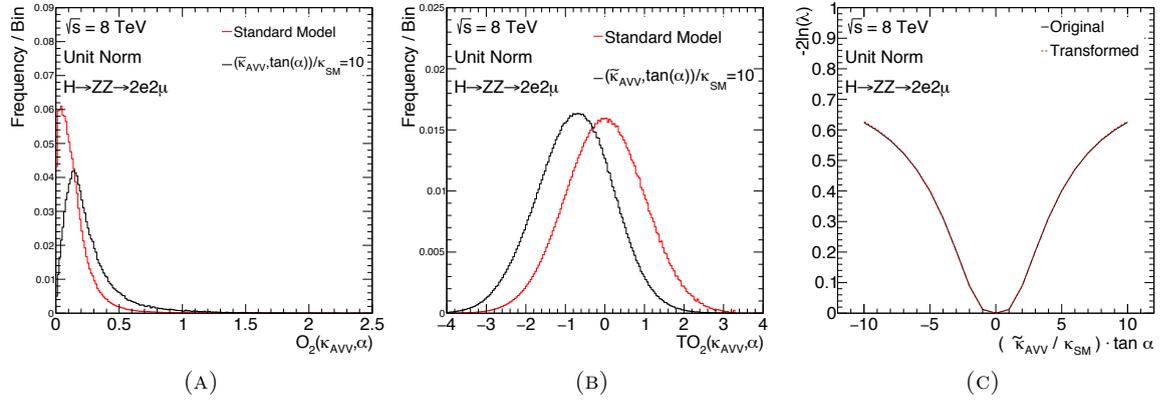


FIGURE 3.11: Transformation validation. Here applied on the  $2e2\mu$  final state of Monte Carlo generated at 8 TeV. The transformation is applied to  $O_2(\tilde{\kappa}_{AVV}, \alpha)$ . (A) Original distributions, for SM and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 10$ . (B) Transformed distribution for SM and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 10$ . (C) Likelihood scan for a test sample using original and transformed distributions.

the histogram templates will contain the same information. Using simulation however, it would have been difficult to populate areas where  $TO_1(\tilde{\kappa}_{HVV})$  is at its left tail while  $TO_2(\tilde{\kappa}_{HVV})$  is at its right and vice versa. The areas correspond to the upper left or lower right histogram corners of Figure 3.12a.

The observable distributions are filled into 3-dimensional histograms containing the three relevant observables for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  separately. For the two coupling ratios the histograms have dimensions:

$$\begin{aligned} \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} &: \{ TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV}), TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV}), BDT_{ZZ} \} \\ \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha &: \{ TO_1(\tilde{\kappa}_{AVV}, \alpha), TO_2(\tilde{\kappa}_{AVV}, \alpha), BDT_{ZZ} \} \end{aligned}$$

A binning of  $\{11, 11, 4\}$  is chosen for the  $\{x, y, z\}$  dimensions of the histograms. It was found that it is most vital for the analysis to have as detailed description of the CP-sensitive observables as possible while the final result is only weakly sensitive to the choice of binning in  $BDT_{ZZ}$ .

To help the reader visualise the multidimensional *pdfs*, the 2D projections of the 3D histograms are shown in Figure 3.12 for different coupling ratios. Only a weak dependency between the observables and the background discriminant is seen.

## Rewighting procedure

The reweighting follows what is described in Section 1.3.1. A large MC sample is generated with both SM and BSM components. To acquire MC with different coupling compositions, an event-by-event weight is calculated from the ratio of target to source squared matrix elements. The reweighting is done in 81 steps of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  from -10 to 10. This way there will exist MC description in steps of 0.25, where SM at 0 is specifically included. Section 1.3.1

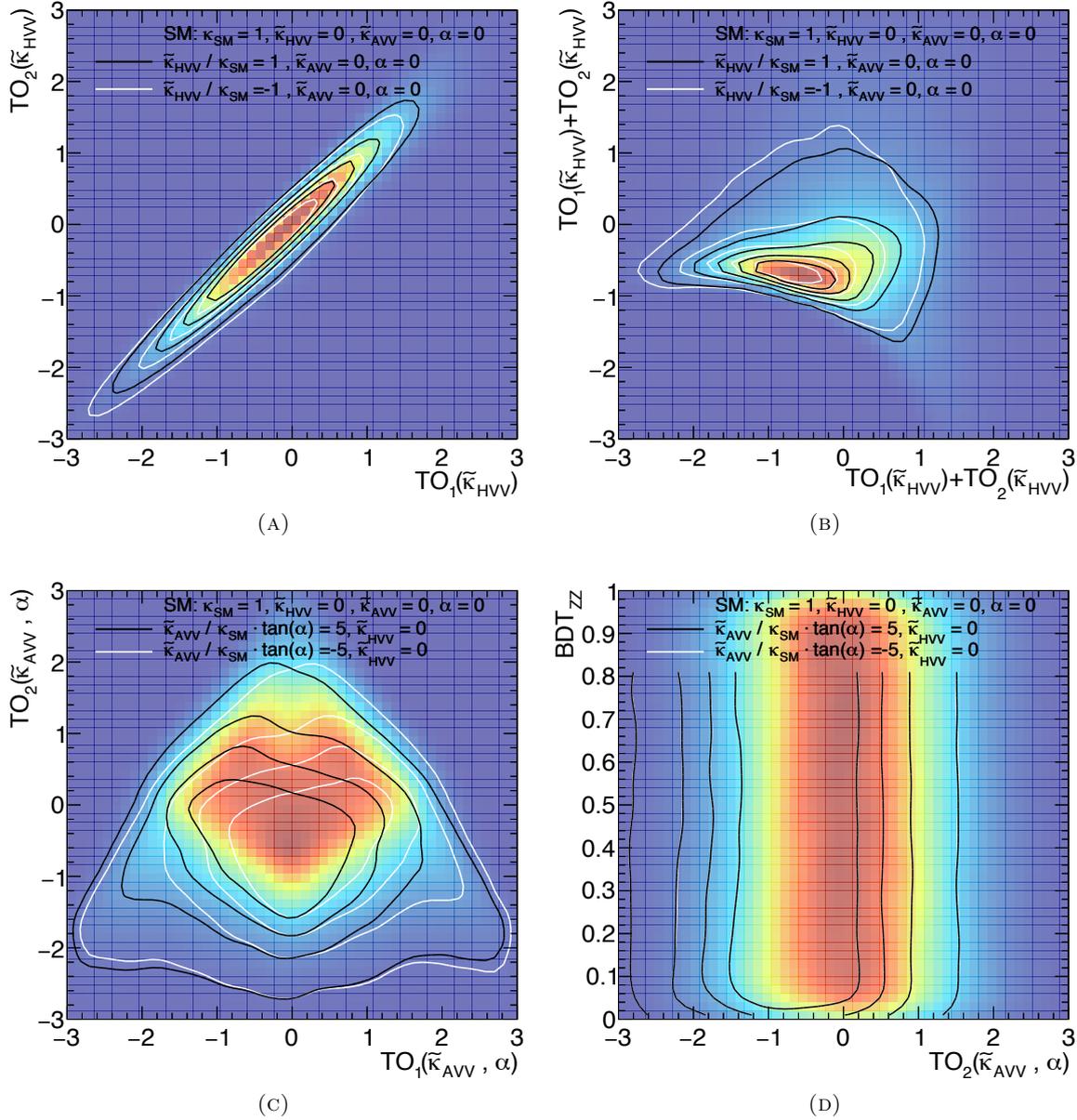


FIGURE 3.12: Two dimensional observable distributions. (A)  $TO_1(\tilde{\kappa}_{HVV})$  vs.  $TO_2(\tilde{\kappa}_{HVV})$ , (B)  $TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})$  vs.  $TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV})$ . (C)  $O_1(\tilde{\kappa}_{AVV}, \alpha)$  vs.  $O_2(\tilde{\kappa}_{AVV}, \alpha)$ . (D)  $O_2(\tilde{\kappa}_{AVV}, \alpha)$  vs.  $BDT_{ZZ}$ . The coloured background represent the Standard Model, while the lines indicate 10%, 30%, 50%, 70% and 90% of the SM maximum for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = \pm 1$  (A,B) and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = \pm 5$  (C,D). Additional figures showing the correlation to  $BDT_{ZZ}$  can be found in Figure C.11 of App. C.

moreover demonstrated that the reweighting could accurately predict final state observable distributions by comparing to independent Monte Carlo samples.

The reweighting validation is repeated here on the Matrix Element Observables and  $BDT_{ZZ}$  to ensure the procedure also accurately describes observables derived from the full final state kinematics. The procedure is repeated for the three-dimensional observables to take any possible dependency into account.

An example of the reweighting validation is shown in Figure 3.13. Here, the source sample is

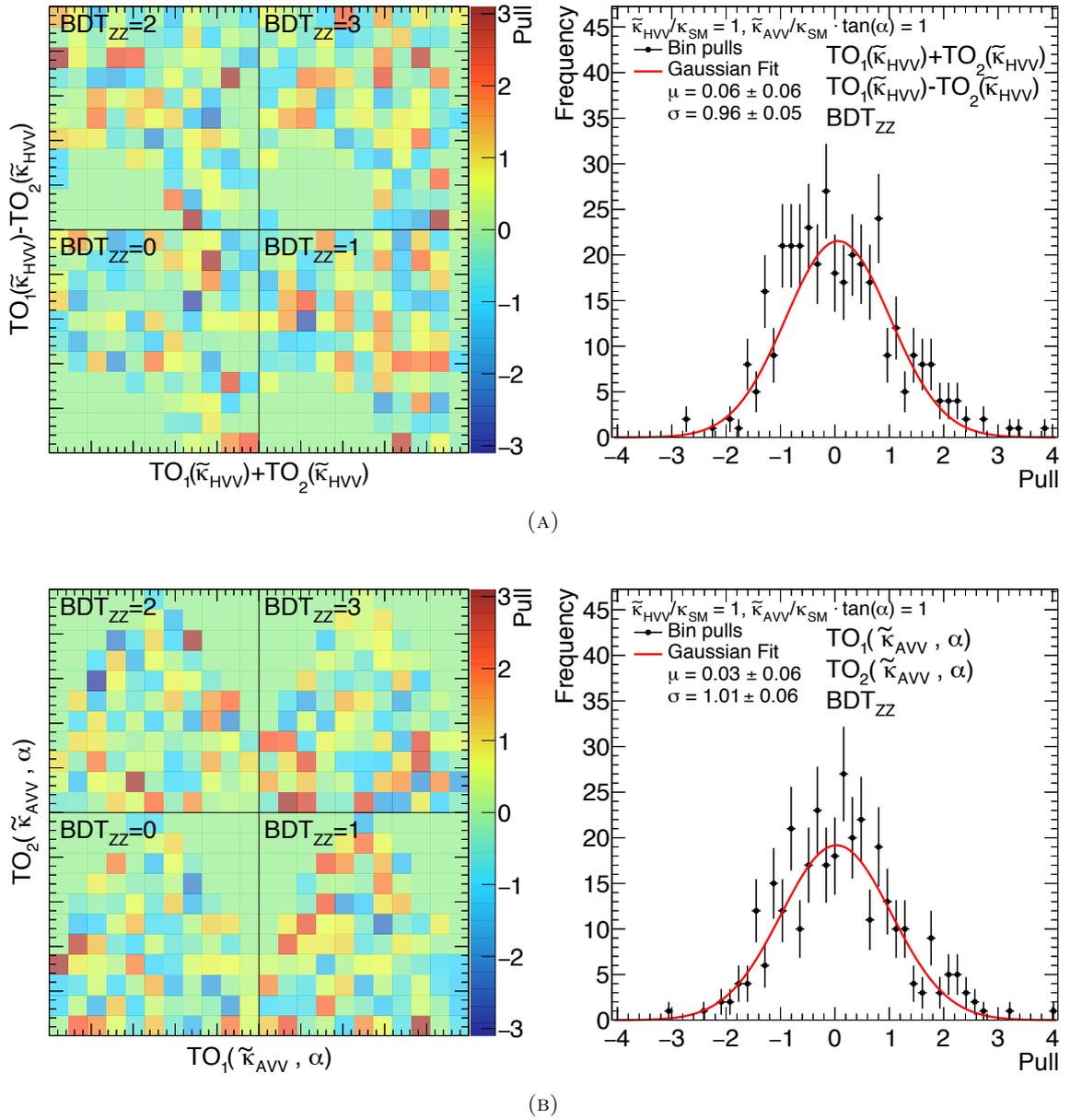


FIGURE 3.13: Reweighting validation for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observables. Left figures show the individual pulls in the four different  $BDT_{ZZ}$  bins. The right figures show the pull distributions. (A) Observables used in  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  fit. (B) Observables used in  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  fit. An additional check can be found in Figure C.12 of App. C.

reweighted to  $\left(\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 1, \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 1\right)$  and compared to distributions obtained from MC directly generated at this coupling configuration. The figure shows a good agreement between the two samples for the different observables that will be used.

### Kernel Density Estimation

A Gaussian Kernel Density Estimator (KDE) is used to get a description of the 3D  $pdfs$  that is less influenced by statistical fluctuations. This will also help to get a description of the tail

behaviour of the observable distributions such that the final likelihood will not contain empty bins.

It is customary to use the RooNDKeysPdf[90] package but it was found that for this specific analysis it could result in a problematic processing time. A binned Kernel Density Estimator was developed instead. The binned KDE can be realised is equivalent to the convolution of two vectors which can be performed using 'Fast Fourier Transform' algorithms due to the Convolution Theorem. A full description of the algorithm is given in App. A.2.

The histograms use a binning of  $\{11, 11, 4\}$ . Motivated by the small dependence of the matrix element observables to  $BDT_{ZZ}$ , it is chosen to perform the KDE separately on the four  $BDT_{ZZ}$  bins and subsequently stack the 2D histograms to retrieve the final 3D histogram. The procedure was done such that the normalisation of each of the  $BDT_{ZZ}$  bins of the final KDE histogram is identical to that of the input histogram. To validate the approach, the resulting KDE was compared to "unsmoothed" histograms. The cross checks are presented in Figure 3.14 for the shapes used in the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$   $\tan \alpha$  measurements.

Special attention is made to the reducible background shapes due to them having low statistics. Some approximations are made to get a better description. The background types that fall into each of the two categories are mainly dictated by the low  $p_T$  lepton pair. It was assumed that the  $4e$  and  $2\mu 2e$  final states can be described with the same  $pdf$ , and likewise for the  $4\mu$  and  $2e 2\mu$  final states. The matrix element observable distributions for the different  $BDT_{ZZ}$  bins are moreover assumed to be the same for reducible backgrounds such that their  $pdf$ s are described by a two times one dimensional shape. A KDE validation for the reducible backgrounds is included in App. C, Fig. C.13.

## Pdf Interpolation

The steps described above will result in a description of the observable histograms for fixed values of either  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  or  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$   $\tan \alpha$ . It is however of interest to get a continuous description of the distributions as function of coupling ratios.

A "vertical, linear"  $pdf$  interpolation is used. In essence this is a bin-by-bin interpolation between the input histograms. The  $pdf$  at coupling  $\kappa$  will take the value:

$$pdf(O|\kappa) = \frac{1}{\kappa_i - \kappa_j} [(\kappa - \kappa_j) \cdot pdf_{\kappa=i}(O) - (\kappa - \kappa_i) \cdot pdf_{\kappa=j}(O)] ,$$

where:

$pdf(O|\kappa)$  is the interpolated  $pdf$  for coupling  $\kappa$  evaluated at observable  $O$ .

$\kappa_i, \kappa_j$  are the two closest templates to  $\kappa$ .

$pdf_{\kappa=i,j}(O)$  are the input templates at  $\kappa_i, \kappa_j$  evaluated at observable  $O$ .

The interpolation is done between signal models using the 81 templates from the reweighting procedure. The continuous signal model is subsequently added to the  $ZZ^*$ -continuum and reducible backgrounds to create the final  $pdf$ .

The same interpolation type is also used to describe variations between systematic uncertainties that affect observable shapes. The systematic shape variations use only three input

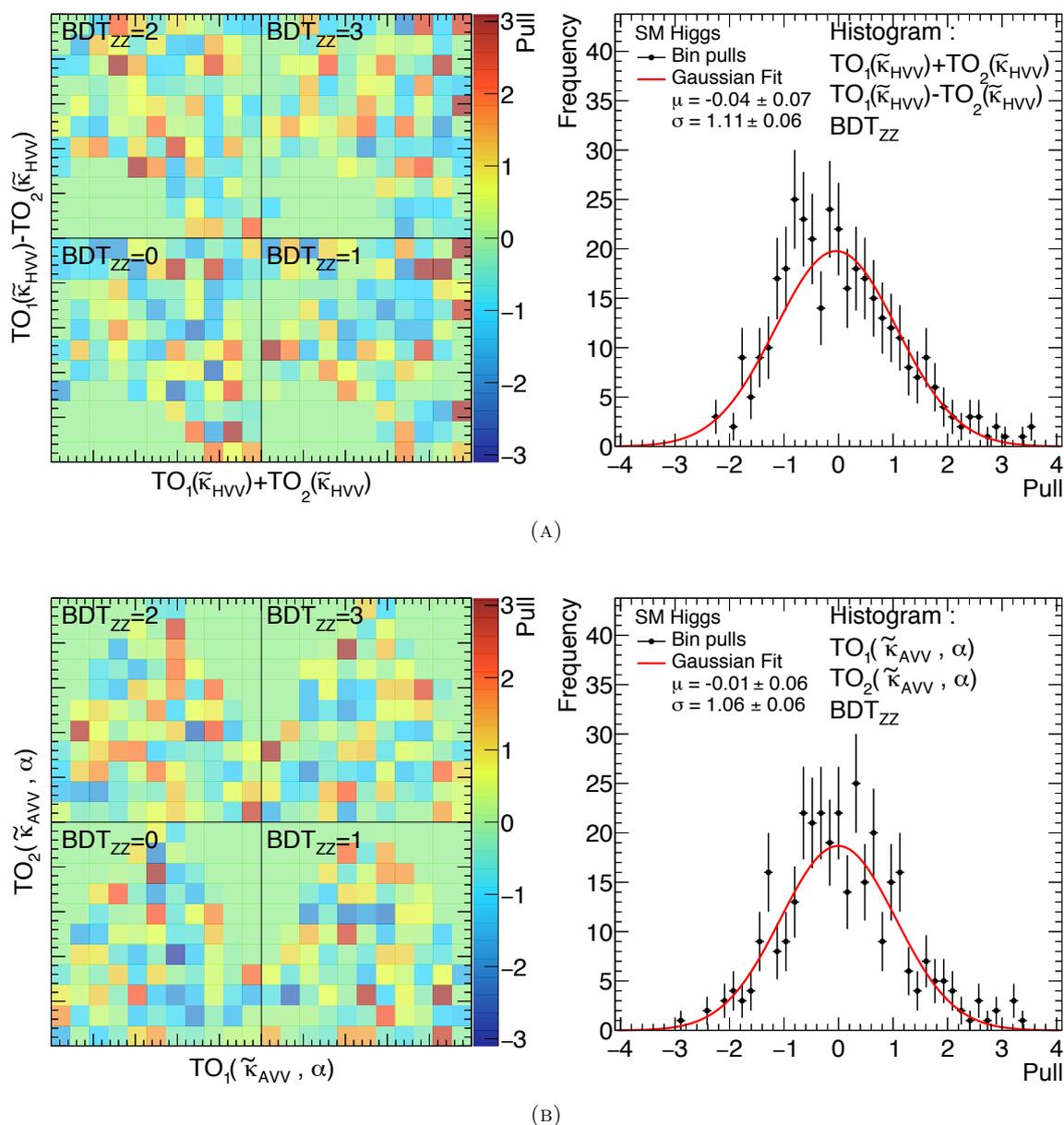


FIGURE 3.14: Kernel Density Estimation validation for  $\tilde{\kappa}_{HVV}^{\tilde{\kappa}_{SM}}$  and  $\tilde{\kappa}_{AVV}^{\tilde{\kappa}_{SM}}$   $\tan \alpha$  observables. Left figures show the individual pulls in the four different  $BDT_{ZZ}$  bins. The right figures show the pull distributions. (A) Observables used in  $\tilde{\kappa}_{HVV}^{\tilde{\kappa}_{SM}}$  fit. (B) Observables used in  $\tilde{\kappa}_{AVV}^{\tilde{\kappa}_{SM}}$   $\tan \alpha$  fit. A cross check for the reducible backgrounds can be found in Figure C.13 of App. C.

templates however; the nominal and the one  $\sigma$  up and down variations. The technical implementation of systematic variations is described [94, Sect. 4.1].

A linear interpolation model is chosen by virtue of it being the absolute simplest *pdf*-morphing method. More advanced methods are found to be unstable when applied to a large number of input histograms. Linear interpolation does however suffer from some disadvantages. The most notable is that the *pdf* is not continuously differentiable at the input points. The likelihood curve can as a result have kinks if the *pdf* is changing rapidly for parameter changes. The limits and maximum likelihood estimator of the coupling ratios that are examined in this analysis are derived from scanning the likelihood function over the coupling ratios. The kinks

will therefore not affect the minimisation procedure as could have been the case for a standard minimisation algorithm.

In the case of systematic uncertainties, it will be important to make sure that the fitted nuisance parameters do not stray far from the one standard deviation variations. In extreme cases the linear interpolation model could e.g. predict negative *pdf* values. As will be demonstrated in Section 3.4.7, the analysis is in the end only to a very small degree affected by systematic uncertainties.

The final observable distributions for backgrounds and various signal models can be found in Figure 3.15. The  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observables are not rotated and behaves as expected;  $TO_1(\tilde{\kappa}_{AVV}, \alpha)$  has sign sensitivity while  $TO_2(\tilde{\kappa}_{AVV}, \alpha)$  is more sensitive but not able to distinguish the  $\pm 5$  models. The  $BDT_{ZZ}$  background discriminant is seen to be identically distributed for the presented signal models.

### 3.4.4 Statistical Procedure

To create the final results, data collected at  $\sqrt{s} = 7$  and 8 TeV and in different final states are treated as independent channels. The combined Likelihood function reads:

$$\mathcal{L} \left( \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha, \theta \mid \Omega \right) = \prod_{\sqrt{s}, \text{FS}} \prod_{i \in TO} P \left[ \Omega_i \mid TO_{i, \text{sig}} \left( \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha, \theta \right) + TO_{i, \text{bkg}}(\theta) \right] \cdot \mathcal{C}(\theta), \quad (3.5)$$

where:

$\theta$  Set of nuisance parameters, containing systematic uncertainties and signal strengths.

$\Omega$  Observed or simulated data.

$\sqrt{s}, \text{FS}$  Centre of mass energy and final state.

$TO$  Transformed matrix element observables. The likelihood uses three dimensional signal or background templates.

$P(a|b)$  Poissonian distribution with expectation  $b$  evaluated at  $a$ .

A constraint term is included in the likelihoods similar to what was done for the hypothesis test.

The profiled likelihood ratio test statistic is finally used to establish the excluded regions:

$$-2 \ln \lambda \left( \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha \right) = -2 \ln \frac{\mathcal{L}_{\text{Free}}(\hat{\theta})}{\mathcal{L}_{\text{MLE}}(\hat{\theta})},$$

Where 'Free' and 'MLE' denotes whether the likelihood is a function of, or evaluated at the maximum likelihood estimator of the coupling parameters.

For each measurement, an Asimov dataset [95] normalised to the expected event yield is produced to calculate the expected results. The test statistic behaviour will be examined in Section 3.4.5.

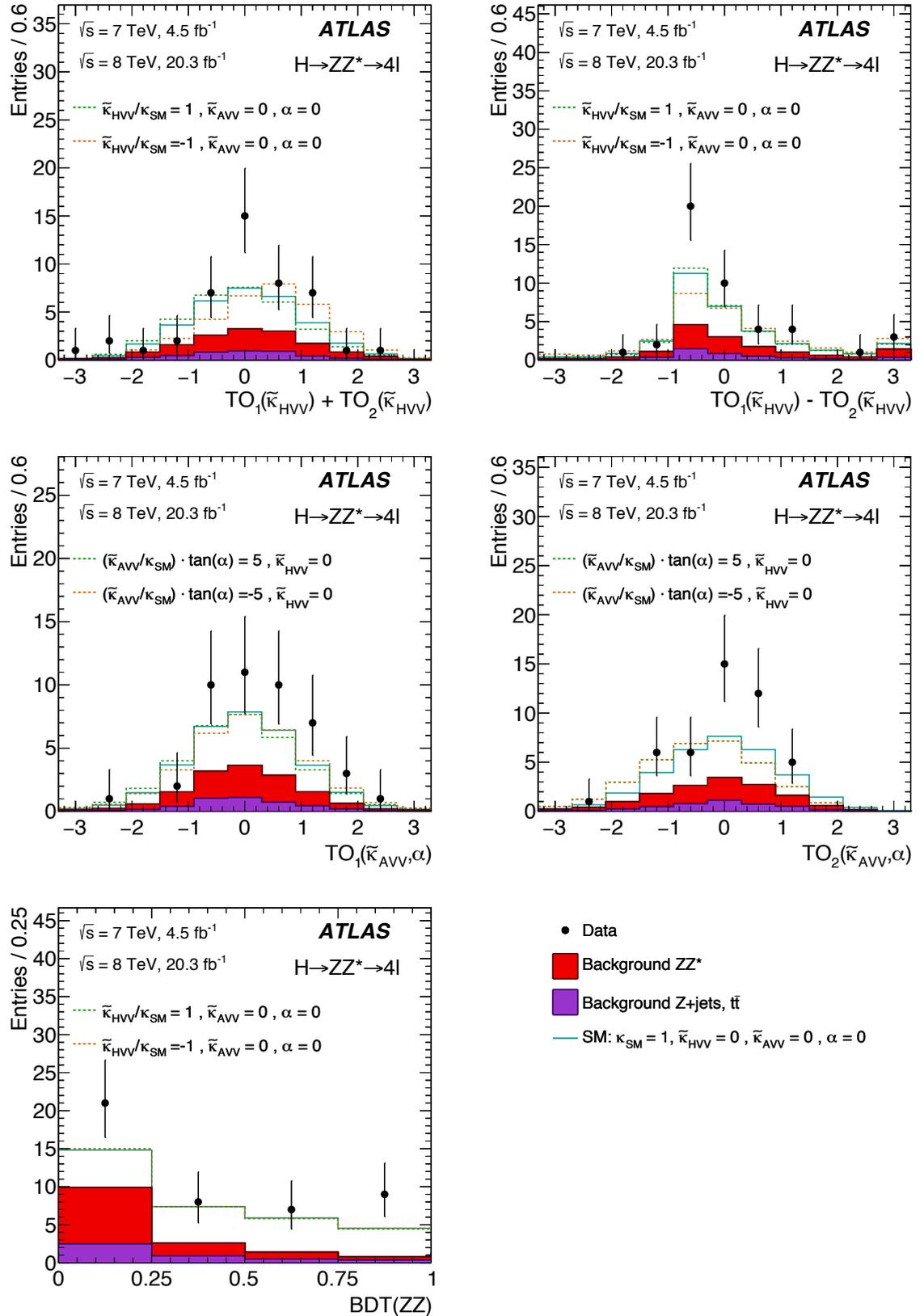


FIGURE 3.15: Matrix element observables:  $TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV})$ ,  $TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})$  (top),  $TO_1(\tilde{\kappa}_{AVV}, \alpha)$ ,  $TO_2(\tilde{\kappa}_{AVV}, \alpha)$  (centre) and background discriminant  $BDT_{ZZ}$  (bottom). The distributions show a stack of reducible backgrounds,  $ZZ^*$ -continuum and different signal models. Figures published in [5].

## Expected results

The likelihood model and test statistic that will be used to measure coupling ratios have now been established. The expected Standard Model results are presented before turning to the further tests that will evaluate the validity of the approach.

The likelihood model is constructed and a nominal Standard Model Asimov sample is created. Nominal refers to all parameters are taken at their expected value ( $\mu = 1$ ,  $m_{4\ell} = 125.5$  GeV,  $\theta = 0$ ). The systematic uncertainties that are included in the likelihood will be described in Section 3.4.7.

The combined and per-final-state likelihood curves are found in Figure 3.16 for both  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The combined likelihood is approximately the sum of the individual final state curves. The final states share systematic uncertainties however such that the equality is not exact. The amount of sensitivity derived from each final state follows their individual event yields ( $N_{4\mu} > N_{2e2\mu} > N_{2\mu2e} > N_{4e}$ ). The best fit values and excluded regions are listed in Table 3.9.

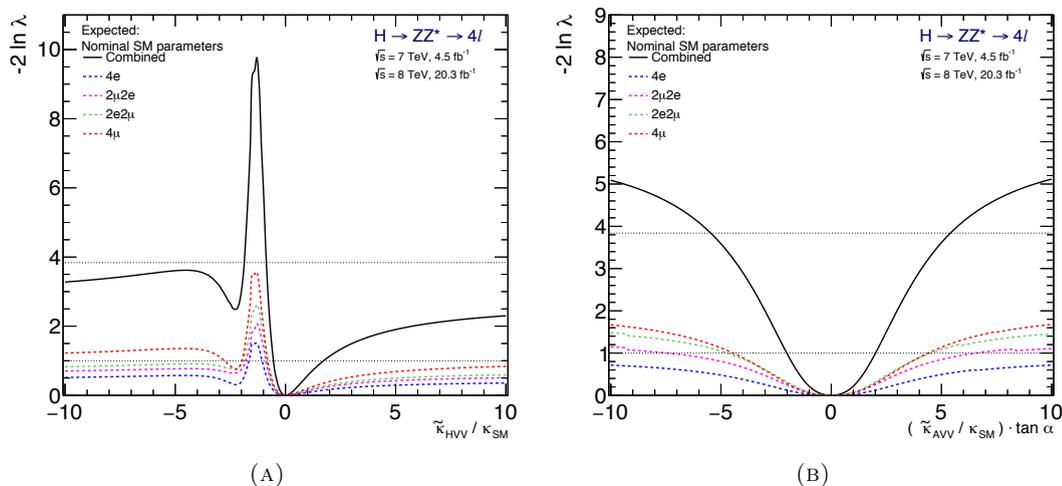


FIGURE 3.16: Expected results of the tensor structure fit over coupling ratios  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  (A) and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  (B). The fit is performed on a Standard Model Asimov sample generated with all nuisance parameters set at their nominal expected values.

Nominal Standard Model expectations: $H \rightarrow ZZ^* \rightarrow 4\ell$		
Coupling ratio	Best-fit value	95% CL Exclusion Regions
$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$	0.0	$[-1.84, -0.89]$
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	0.0	$(\infty, -5.45] \cup [5.45, \infty)$

TABLE 3.9: Expected best-fit values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  and 95% CL excluded regions obtained in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. The expected values are estimated for an Asimov data sample generated with nominal nuisance parameters. The  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV samples are combined.

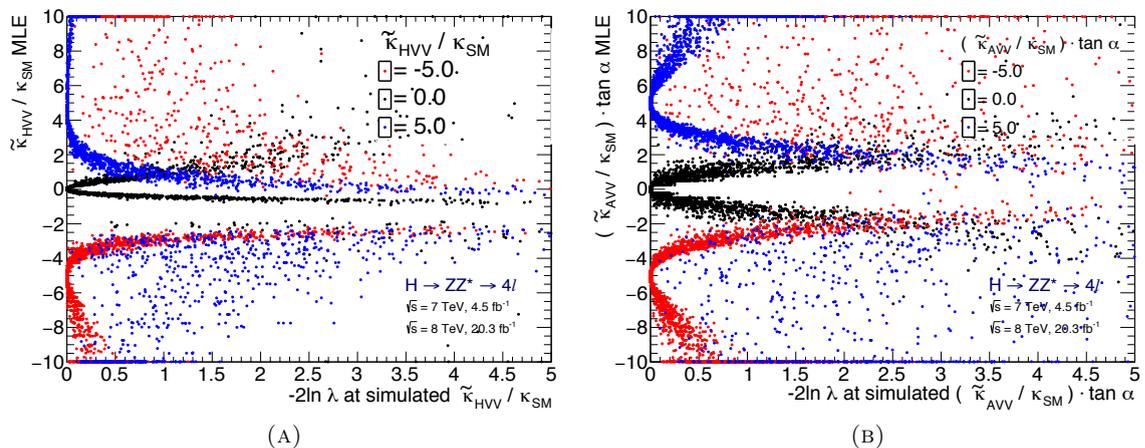


FIGURE 3.17: Fit results for Monte Carlo pseudo experiments. The figure shows the Maximum Likelihood Estimator for coupling ratio vs. the profiled likelihood ratio  $-2 \ln \lambda$  evaluated at the true generated value. (A)  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  generated at  $-5, 0$  (SM),  $5$ . (B)  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$  generated at  $-5, 0$  (SM),  $5$ .

### 3.4.5 Asymptotic Test Statistic Behaviour

When establishing that regions where  $-2 \ln \lambda > 3.84$  can be excluded at more than 95% CL, it is explicitly assumed that  $-2 \ln \lambda$  follows a  $\chi_1^2$  distribution when evaluated at the true parameter value[96]. This is however only precise to  $\mathcal{O}(1/\sqrt{N})$ , where  $N$  is the observed number of events. The exact function is analysis dependent.

MC is generated from the model to establish if it is possible to ignore the  $1/\sqrt{N}$  term. Samples with coupling ratios of  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}} = -5, 0, 5$  and  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha = -5, 0, 5$  are generated and used as input for the fit. The datasets are fitted with a no-systematic-uncertainty model, where only the signal strengths for the 7 TeV ( $\mu_{2011}$ ) and 8 TeV ( $\mu_{2012}$ ) samples are allowed to float freely. As will be demonstrated in Section 3.4.7, systematic uncertainties have very little impact on the results, such that the "no-systematics" sampling will give an accurate description of the true distribution.

Prior to generating the pseudo samples,  $\mu_{2011}$ ,  $\mu_{2012}$  are fitted to the observed data and sampling is performed according to their best fit values to get the most accurate description of data behaviour.

Figure 3.17 shows the dependence between  $-2 \ln \lambda$  and the best fitting coupling value for the sampled distribution. The figure demonstrates that the model will tend to find minima close to the true coupling value. Moreover, a higher sensitivity is seen for the Standard Model than the  $\pm 5$  distributions. Models are parameterised as coupling ratios and the ability to distinguish between them solely based on shape information. The sensitivity is therefore expected to flatten when moving away from SM such that it is not possible to distinguish between models of very large coupling ratios.

It is also seen that for the  $\pm 5$  models it is possible to fit outside of the defined analysis range. In principle, the fit should extend to infinite coupling ratios, corresponding to pure BSM

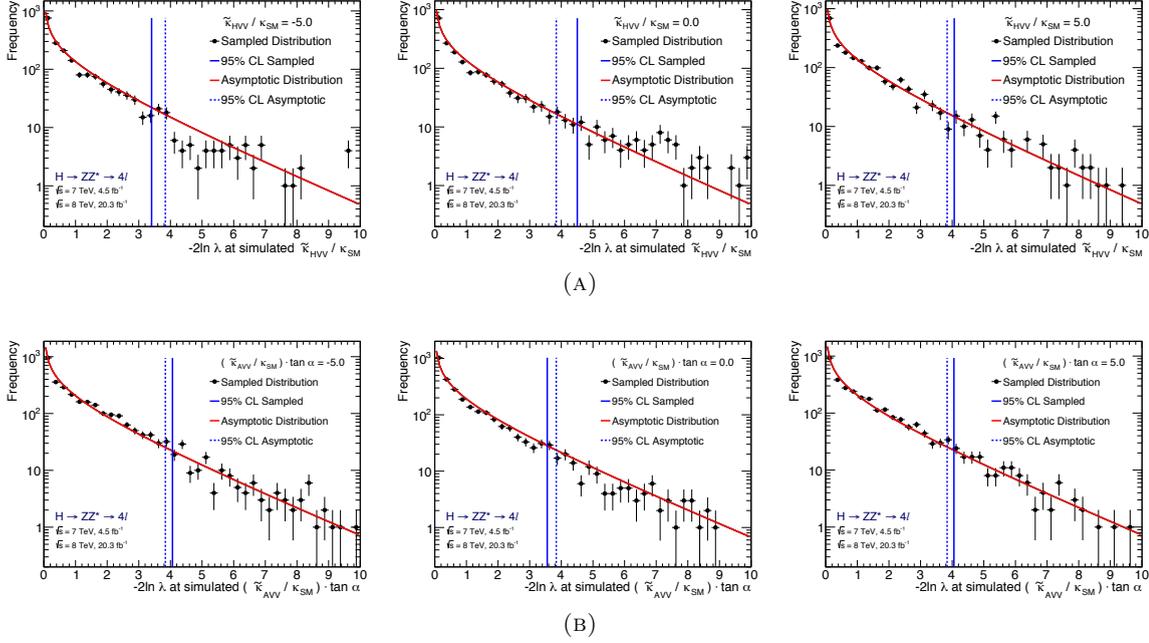


FIGURE 3.18: Sampled  $-2 \ln \lambda$  distributions for: (A)  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  generated at  $-5, 0$  (SM),  $5$ . (B)  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  generated at  $-5, 0$  (SM),  $5$ . The figures show the predicted  $\chi^2_{\text{NDF}=1}$  distribution in red. The analysis establishes 95% CL intervals in regions where  $-2 \ln \lambda > 3.84$ . The full blue line indicate the point where 95% of the pseudo experiments fell to the left of, while the dashed line show the assumed limit.

scenarios<sup>4</sup>. Based on the results of the hypothesis tests, it is reasoned that this will not be of great concern. Although used different, the hypothesis tests are based on the same shape information as the tensor structure fit. If large BSM contributions to the effective Lagrangian exists, it seems unlikely that the hypothesis test would show such a great compatibility with the Standard Model.

The  $-2 \ln \lambda$  distributions can be derived by projecting the distributions of Figure 3.17 onto the x-axis. In the following, the fits that converge outside of range are not considered in order to examine the distribution of well defined test statistic values. The  $-2 \ln \lambda$  distributions can be found in Figure 3.18. A good agreement between the sampled and asymptotic distributions is in general seen.

The pull distributions are examined to further understand the fit behaviour. Normally the pull is calculated as the difference between the observed and true parameter in units of the observed uncertainty:

$$\text{Pull} = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}},$$

where:

$\hat{\theta}$  : Observed parameter

<sup>4</sup>One could change parameterisation to the effective BSM cross sections contribution;  $f_{g2}, f_{g4}$  from 0 to 1 for instance.

$\theta_0$  : True parameter

$\sigma_{\hat{\theta}}$  : Gaussian uncertainty on the observed parameter

It will be useful here to rewrite this expression slightly. In standard measurements, the distance between each unit of uncertainty  $\sigma_{\hat{\theta}}$  will be the same<sup>5</sup>. This corresponds to the likelihood curve forming a parabolic function. If, as is the case here, the likelihood curve is not parabolic, it can no longer be assumed that e.g. the distance to one  $\sigma_{\hat{\theta}}$  is half of the distance to two  $\sigma_{\hat{\theta}}$ . Assuming asymptotic behaviour, the uncertainty can however be estimated as<sup>6</sup>:

$$\sigma_{\theta} = \frac{|\hat{\theta} - \theta_0|}{\sqrt{-2 \ln \lambda(\theta_0)}} \quad \Rightarrow \quad \text{Pull} = \frac{\hat{\theta} - \theta_0}{|\hat{\theta} - \theta_0|} \sqrt{-2 \ln \lambda(\theta_0)}$$

The pull will contain the same information as the  $-2 \ln \lambda$  distribution but moreover distinguish preferences to obtain values higher or lower than the true. The distributions for the different sampling series are shown in Figure 3.19. The Standard Model distributions form a unit Gaussian, while the BSM distributions are asymmetric. It follows intuition that the BSM samples cannot form a Gaussian tail in the direction away from the Standard Model. The analysis has poor separation power between models  $\gtrsim |5|$  as can e.g. be seen from the expected results in Figure 3.16, where the curves tend to flatten out for high values.

Another effect is seen when the fits converge to the wrong sign of the coupling ratio. In general the model will have better sensitivity to the magnitude of the coupling rather than the sign, meaning there is a chance that the fit will converge to the wrong sign. In this case the pull distribution cannot be symmetrical around zero.

The analysis is concluded to exhibit asymptotic behaviour for best fitting models with coupling ratios in  $[-5.0, 5.0]$ . If the best fit value is observed at higher ratios in data, special care should be made to ensure that the test statistic is interpreted accurately. A best fit to a coupling ratio with large BSM contributions is also expected to reveal a degenerate minimum, where there is poor exclusion to the opposite signed coupling.

With the knowledge from the hypothesis tests, it is reasoned to be unlikely that the Lagrangian would contain large BSM contributions.

### 3.4.6 Closure Tests

A series of tests are performed to establish the analysis ability to locate minima over the examined range of couplings. In each case an Asimov sample is created under various conditions and subsequently fitted. To make the tests as stringent as possible, the Asimov dataset is created from different Monte Carlo samples than the one used to create the *pdf*. The dataset is here called "independent Asimov".

Unlike the true Asimov, which per construction exactly reproduces all parameters, the independent can be used to check for analysis biases. These could for instance be the reweighting

<sup>5</sup> $\sigma_{\hat{\theta}}$  is here used as the distance along the parameter axis one has to go before the true model would result in the observed with frequency less than 31.7% ( $1\sigma$ ), 4.6% ( $2\sigma$ ), etc.

<sup>6</sup>Note that the non-parabolic likelihood can still have its  $-2 \ln \lambda$  approximate a  $\chi^2$  distribution.

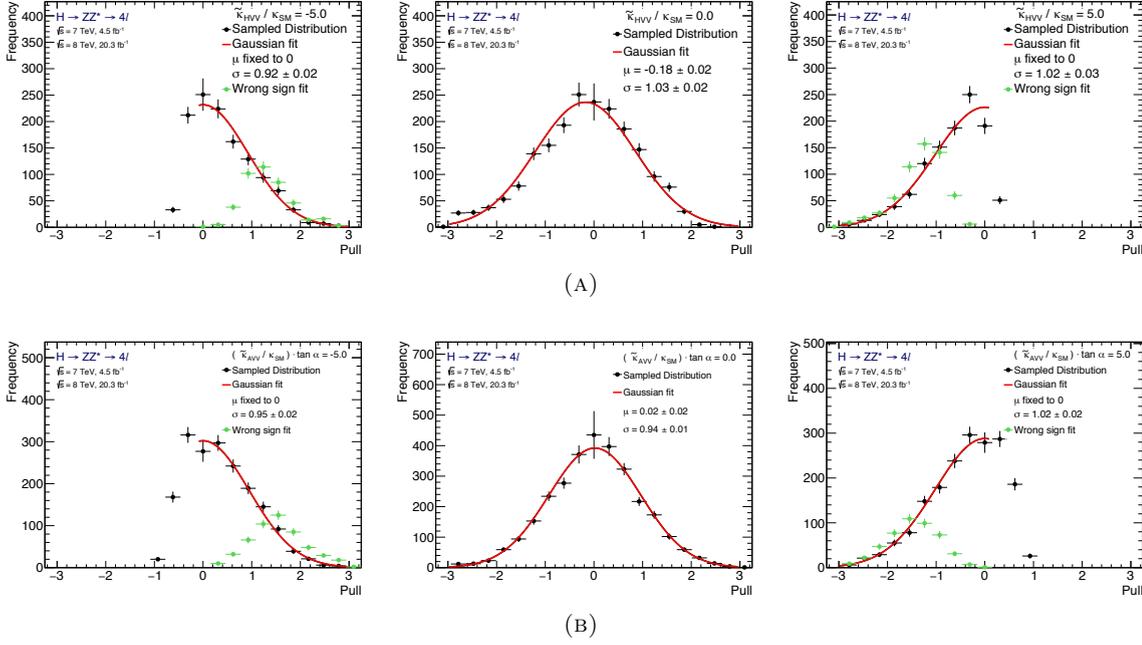


FIGURE 3.19: Sampled pull distributions. (A)  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  generated at  $-5, 0$  (SM),  $5$ . (B)  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  generated at  $-5, 0$  (SM),  $5$ . A positive value indicates that the observed parameter overestimates the true. Each distribution has been fitted to a Gaussian. In the Standard Model case the fit covers the full range. A one sided fit is performed on the BSM samples. The side that is not fitted corresponds to the side where the fit can go out of range. Fits that have converged to the wrong sign are indicated by green.

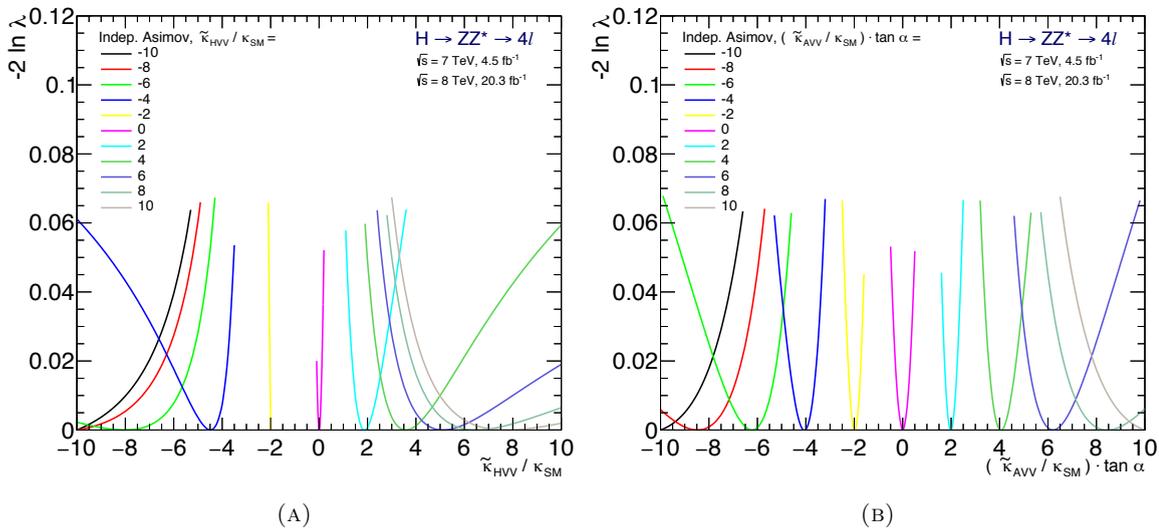


FIGURE 3.20: Closure tests performed on independent Asimov samples. (A):  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  models ranging from  $-10$  to  $10$ . (B)  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  models ranging from  $-10$  to  $10$ .

procedure, the KDE or calculation errors in general. The independent dataset will however also have statistical fluctuations compared to the Monte Carlo used to construct the  $pdf$ . Asimov-data is created in steps of two in  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  from -10 to 10. The data is created with an integrated luminosity corresponding to the LHC Run-1 dataset. Figure 3.20 show the combined fit to the different samples where it is seen that the analysis closes near the generated values. Some differences are seen for the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  analysis, but the fit in no place show a  $-2 \ln \lambda$  value of more than 0.01 at the true generated parameter value. To understand the size of the bias it can be mentioned that the likelihood values of an Asimov sample fit scales linearly with assumed statistics size. The expected biases thus correspond to approximately  $1\sigma$  for a sample size of 100·Run-1, equivalent to the total integrated luminosity expected to be collected with LHC and HL-LHC.

### 3.4.7 Systematic Uncertainties

The majority of the systematic uncertainties of the tensor structure fit are equivalent to those described before. Uncertainties related to theoretical predictions, detector effects and integrated luminosity are the same as for the hypothesis tests. In addition to these, systematic effects are introduced from the observable shape construction.

The procedure described above of filling three dimensional histograms and performing a KDE, is repeated for each systematic change separately. The shapes corresponding to systematic changes are finally added to the likelihood model before each fit. The systematic changes related to the ME-Obs fit that is not described in the common section are:

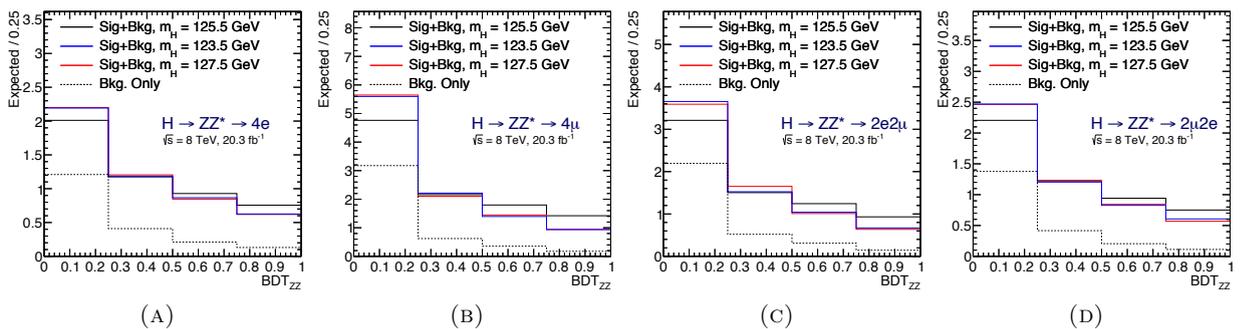


FIGURE 3.21:  $BDT_{ZZ}$  distributions for various mass models. The figures show from left to right the four different final states,  $4e$ ,  $4\mu$ ,  $2e2\mu$ ,  $2\mu2e$ . The signal models are stacked on top of the background. A shift away from 125.5 GeV makes the signal more background like.

#### Mass mis-modelling :

##### alpha\_ATLAS\_Higgs\_mass

Mass resolution and mis-modelling of the reconstructed Higgs boson mass have been taken as an additional source of systematic uncertainty. The four lepton mass is the strongest separating parameter used in the background discriminant, meaning that the mass and the background discriminant will be correlated. A nominal value of 125.5 GeV

was chosen based on the combined ATLAS  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  mass fit [88]<sup>7</sup>. The parameter has not been constrained and is freely allowed to float in all fits. Even though the uncertainty on the Higgs mass is known in the four lepton final state, this number is not used since it is based on the same data that the tensor structure fit utilises. The mass enters in the Matrix Element Observable fit indirectly through the background discriminant however and the method is as a result only weakly able to measure it. It was found using toy Monte Carlo samples that templates situated at  $\pm 2$  GeV would be able to cover most fits. Again, a linear interpolation model is used. The  $BDT_{ZZ}$  distributions for various mass models, for the four different final states can be found in Fig. 3.21. A shift away from the nominal 125.5 GeV will tend to make the signal model more background-like.

### KDE procedure :

`alpha_ATLAS_rho_signal`, `alpha_ATLAS_rho_Redbkg`, `alpha_ATLAS_rho_ZZ`

It is likely that different choices of bandwidth used in the kernel density estimation would have resulted in equally valid descriptions while at the same time showing slight shape differences. To estimate the uncertainty introduced by the kernel density estimation, the bandwidth scale was changed from the nominal 0.16 to 0.13 and 0.20 for signal and  $ZZ^*$ -continuum. Similarly, a shift from the nominal 0.5 of the reducible background to 0.4 and 0.6 was chosen such that in both cases the systematic shifts are approximately 20% of the nominal. A separate nuisance parameter is assigned for each case.

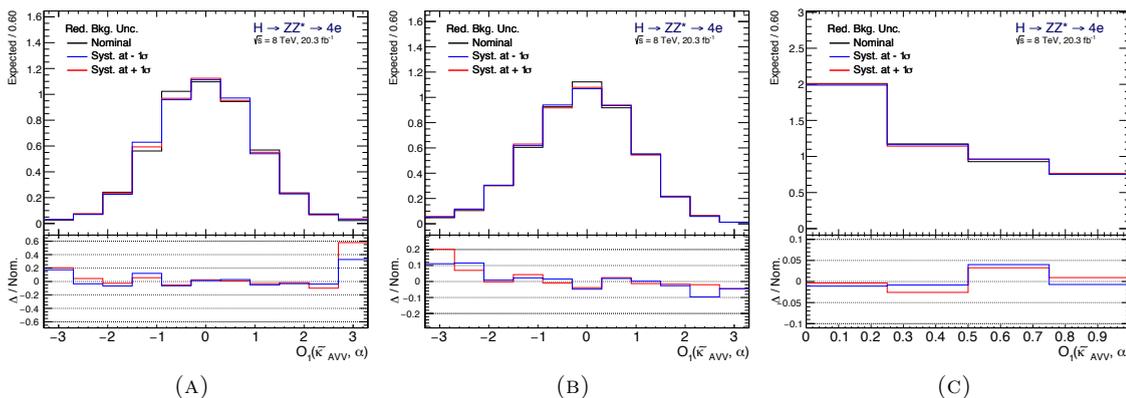


FIGURE 3.22: Impact on  $TO_1(\tilde{\kappa}_{AVV}, \alpha)$  (A),  $TO_2(\tilde{\kappa}_{AVV}, \alpha)$  (B) and  $BDT_{ZZ}$  (C) from the soft electron reducible background uncertainty `alpha_ATLAS_shape_SF_H41_Z_11ee_2012`. The shapes shown here correspond to the combined signal and background  $4e$ , 8 TeV model. The relative difference ( $\Delta/\text{Nom}$ ) is defined such that an upward shift compared to the nominal distribution is positive. The uncertainty is seen to affect the low populated regions.

The potential effect of the systematic uncertainties are evaluated such that irrelevant parameters can be discarded. The impact of each nuisance parameter is estimated in the following. This is done by examining the profiled likelihood ratio between  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 0$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 5$

<sup>7</sup>Rounded to nearest 500 MeV.

Dominant systematic uncertainties for tensor structure analysis

Nuisance parameter	Average effect
$\frac{\tilde{\kappa}_{HVY}}{\kappa_{SM}}$ fit	
Reducible background shape in the $2\mu 2e + 4e$ final states in 2012 alpha_ATLAS_shape_SF_H41_Z_11ee_2012	2.30%
Reducible background shape in the $2\mu 2e + 4e$ final states in 2011 alpha_ATLAS_shape_SF_H41_Z_11ee_2011	1.41%
Reducible background normalisation in the $2\mu 2e + 4e$ final states in 2012 alpha_ATLAS_norm_SF_H41_Z_11ee_2012	1.26%
Reducible background normalisation in the $2e2\mu + 4\mu$ final states in 2012 alpha_ATLAS_norm_SF_H41_Zbb_11mumu_2012	0.95%
Muon spectrometer resolution alpha_ATLAS_MU_MS	0.72%
Reducible background shape in the $2e2\mu + 4\mu$ final states in 2012 alpha_ATLAS_shape_SF_H41_Zbb_11mumu_2012	0.61%
$q\bar{q}$ parton density function for $ZZ^*$ alpha_pdf_qq	0.58%
EM calorimeter energy scale alpha_ATLAS_EM_ES_Z	0.52%
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ tan $\alpha$ fit	
Reducible background shape in the $2\mu 2e + 4e$ final states in 2012 alpha_ATLAS_shape_SF_H41_Z_11ee_2012	1.17%
Reducible background normalisation in the $2\mu 2e + 4e$ final states in 2012 alpha_ATLAS_norm_SF_H41_Z_11ee_2012	0.97%
Reducible background normalisation in the $2e2\mu + 4\mu$ final states in 2012 alpha_ATLAS_norm_SF_H41_Zbb_11mumu_2012	0.89%
$q\bar{q}$ parton density function alpha_pdf_qq	0.88%
Reducible background shape in the $2\mu 2e + 4e$ final states in 2011 alpha_ATLAS_shape_SF_H41_Z_11ee_2011	0.74%
Background renormalisation and factorisation QCD scale alpha_QCDscale_VV	0.72%
$\sqrt{s} = 8$ TeV luminosity uncertainty alpha_ATLAS_LUMI_2012	0.57%

TABLE 3.10: Dominant systematic uncertainties of the tensor structure fit. The effect is evaluated as the differences in  $-2\ln\lambda(-0.8)$  for  $\frac{\tilde{\kappa}_{HVY}}{\kappa_{SM}}$  and  $-2\ln\lambda(5)$  for  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan\alpha$  when shifting each nuisance parameter. The differences are calculated for a Standard Model Asimov sample. The full list can be found in Tables C.2 and C.3 of Appendix C.

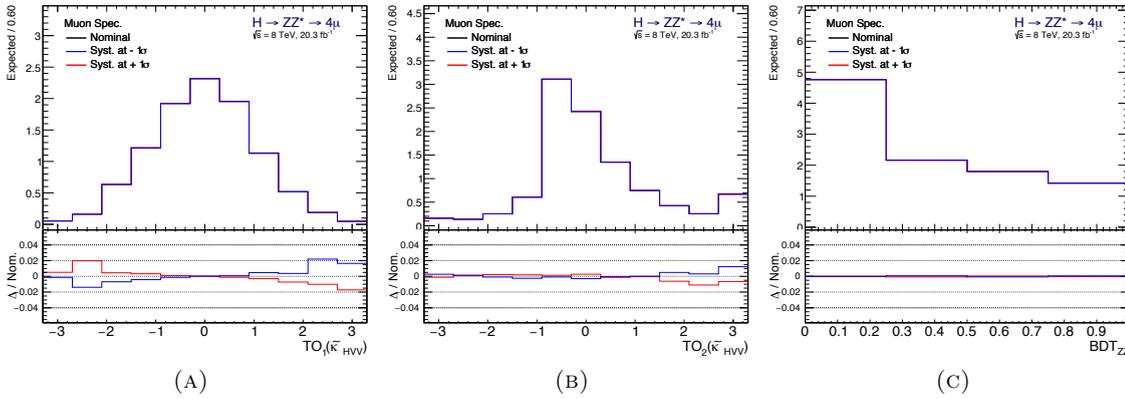


FIGURE 3.23: Impact on  $TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV})$  (A),  $TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})$  (B) and  $BDT_{ZZ}$  (C) from the muon spectrometer resolution model, `alpha_ATLAS_MU_MS`. The shapes shown here correspond to the combined signal and background  $4\mu$ , 8 TeV model. The relative difference ( $\Delta/\text{Nom}$ ) is defined such that an upward shift compared to the nominal distribution is positive. The uncertainty is seen to affect the low populated regions.

assuming a Standard Model Higgs boson. The difference is calculated under nominal circumstances and when each nuisance parameter is shifted. In a similar approach the  $-2 \ln \lambda$  impact is calculated between  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 0$  and  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = -0.8$  for the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  fit. Evaluating the impact of the systematic uncertainties using these points was motivated by the fact that it is approximately where it is expected that the 95% CL limits are going to be found.

An ideal estimation would evaluate the impact based on where the exclusion regions are found and not the likelihood value, since this will be the actual result. It is however very computationally expensive to calculate this number to the same level of precision as the size of the systematic effects. This is partly, as will be seen, because the analysis is insensitive to systematic uncertainties, but also an effect of how the fit is performed. The profiled likelihood ratio is calculated for a number of parameter values. In the approach chosen here it is necessary to perform this calculation two times for each nuisance parameter.

If on the other hand the limit were to be used, it would be necessary to perform the calculation in a very finely spaced grid to examine sub-percent effects.

The log likelihood ratio is interpreted as a significance and the difference in significance is subsequently used as a measure of relevance of each systematic uncertainty. A threshold of 0.5 % is chosen as to where a parameter is relevant for the analysis.

The list of nuisance parameters that are found to potentially influence the log-likelihood ratio more than 0.5% is listed in Table 3.10. The complete lists can be found in Tables C.2 and C.3 of Appendix C. The Higgs mass uncertainty is moreover included in all models.

Two examples of systematics that were found to be influential are given in Figure 3.22 and 3.23. The figures show the uncertainty on the soft electron reducible background ( $l\bar{l}ee$ ) and the muon spectrometer resolution uncertainty. The two systematics are both of the shape variation type. The effects are in general observed to be small.

The nuisance parameters profiles are finally examined as function of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  to understand their behaviour. Figures 3.24 and 3.25 show the best fitting value and corresponding uncertainty of each nuisance parameter that is estimated to have a potential effect. The

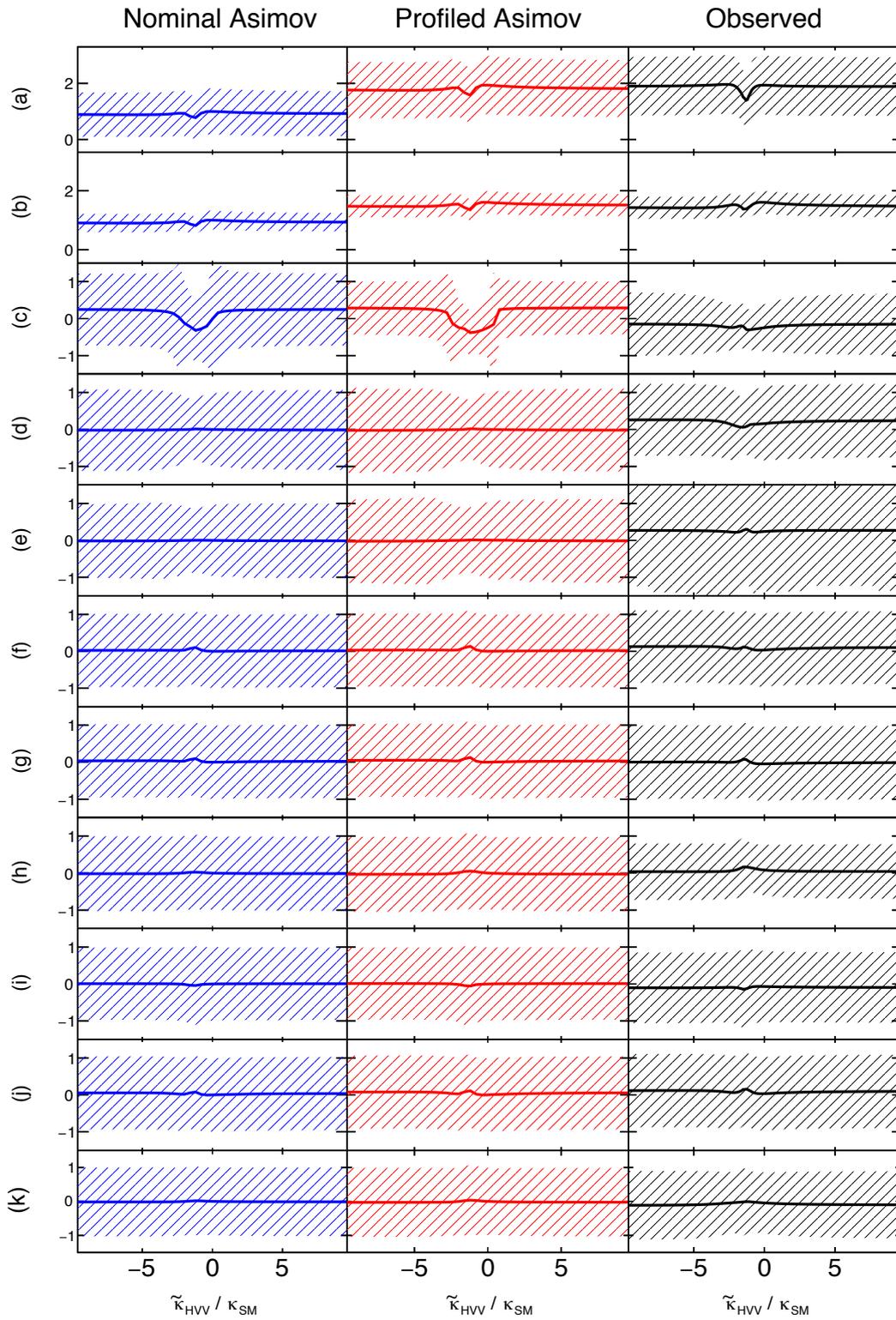


FIGURE 3.24: Best fitting value and uncertainty of each nuisance parameter in the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  scan.

(a)	mu_2011	(g)	alpha_ATLAS_norm_SF_H41_Zbb_11mumu_2012
(b)	mu_2012	(h)	alpha_ATLAS_MU_MS
(c)	alpha_ATLAS_Higgs_mass	(i)	alpha_ATLAS_shape_SF_H41_Zbb_11mumu_2012
(d)	alpha_ATLAS_shape_SF_H41_Z_11ee_2012	(j)	alpha_pdf_qq
(e)	alpha_ATLAS_shape_SF_H41_Z_11ee_2011	(k)	alpha_ATLAS_EM_ES_Z
(f)	alpha_ATLAS_norm_SF_H41_Z_11ee_2012		

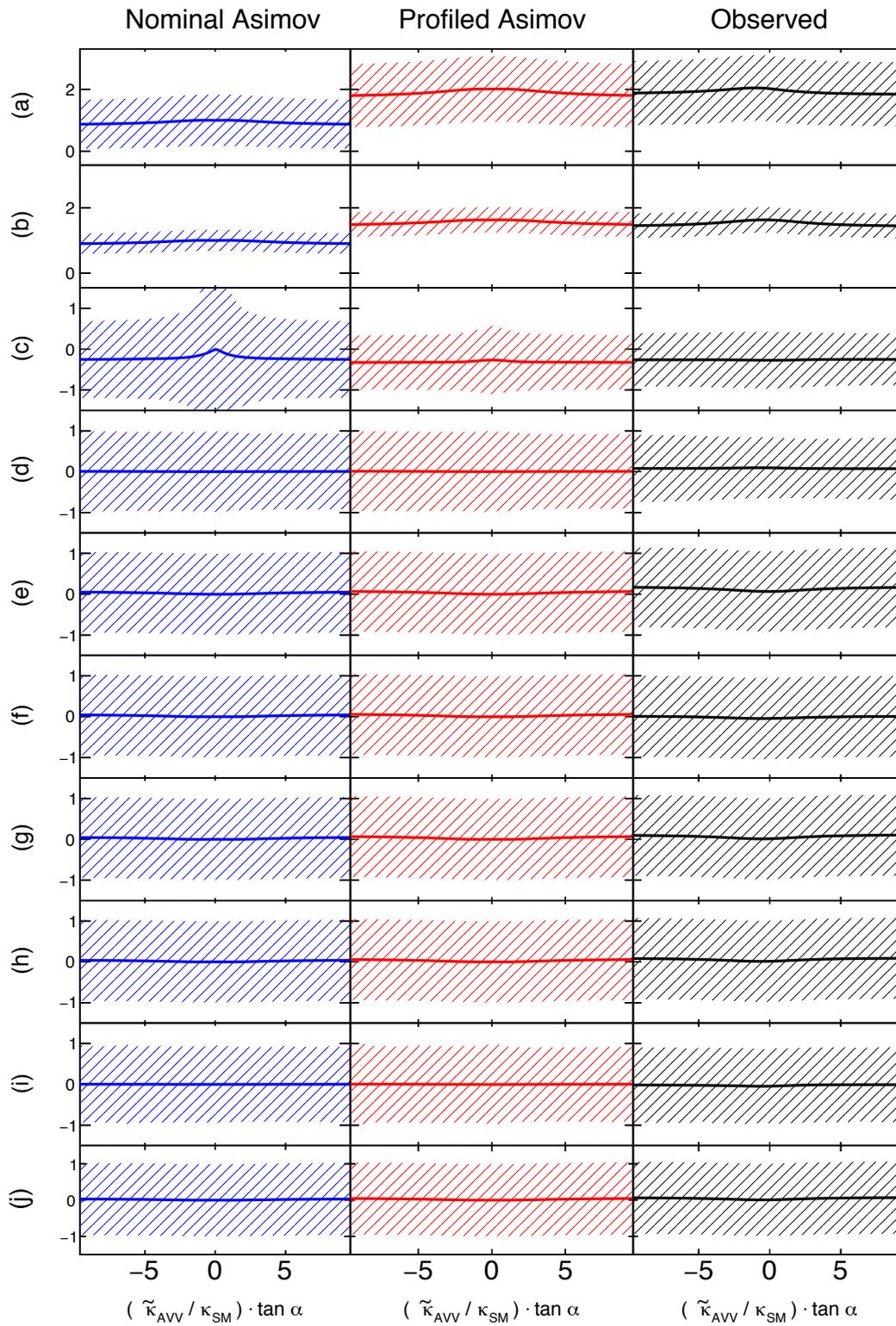


FIGURE 3.25: Best fitting value and uncertainty of each nuisance parameter in the  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  scan.

(a) mu_2011	(f) alpha_ATLAS_norm_SF_H41_Zbb_11mumu_2012
(b) mu_2012	(g) alpha_pdf_qq
(c) alpha_ATLAS_Higgs_mass	(h) alpha_QCDscale_VV
(d) alpha_ATLAS_shape_SF_H41_Z_11ee_2012	(i) alpha_ATLAS_shape_SF_H41_Z_11ee_2011
(e) alpha_ATLAS_norm_SF_H41_Z_11ee_2012	(j) alpha_ATLAS_LUMI_2012

three columns show the outcome when the fit is applied to different samples. The 'Nominal Asimov' sample is generated with expected parameter values; 0 for systematic uncertainties and 1 for signal strengths. The 'Profiled Asimov' sample is a Standard Model sample generated with all nuisance parameters fitted to data. The 'Observed' column show the outcome of the fit when applied to the actual data. The 'Profiled Asimov' best fitting nuisance parameter values at  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 0$  or  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 0$  are per construction the same as for the 'Observed' column.

The top three rows show the 7 TeV, 8 TeV signal strengths and Higgs mass respectively. For the nominal Asimov sample, it can be seen that each of these will tend to pull the signal distribution down when moving away from the Standard Model value, effectively making the combined model more background like. The remaining nuisance parameters have small variations over the fitted range and uncertainties close to one, indicating small influence.

### 3.4.8 Compatibility Studies, The Nine Dimensional Fit

A different approach was tried for the tensor structure analysis: The Nine-Dimensional fit<sup>8</sup>. Only a superficial description of the approach can be given here, since the Author has not been directly involved in the work. The approach is outlined as follows:

The analytical squared matrix element of the  $HZZ$  decay is calculated as function of the five decay observables;  $m_{12}$ ,  $m_{34}$ ,  $\cos \theta_1$ ,  $\cos \theta_2$  and  $\phi$  given coupling ratios  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The matrix element is calculated at parton-level. By normalising the function, a *pdf* is given which contain all information of the  $HZZ$  tensor structure.

The *pdf* does however not take into account detector effects. A factorised detector model is created by studying which observable corrections are correlated.

The production observables  $p_{T,4l}$ ,  $\eta_{4l}$ ,  $\cos \theta^*$  and four lepton mass  $m_{4l}$  are estimated from simulation as well and added as factors to discriminate against backgrounds. The complete model reads:

$$\begin{aligned} pdf = & f_{\text{Parton}}(m_{12}, m_{34}, \cos \theta_1, \cos \theta_2, \phi | \kappa) \\ & \cdot h_{\text{corr}}(m_{12}, m_{34} | \kappa) \cdot h_{\text{corr}}(\cos \theta_1, \cos \theta_2, \phi | \kappa) \\ & \cdot h(m_{4l}) \cdot h(\cos \theta^*) \cdot h(p_{T,4l}, \eta_{4l}), \end{aligned}$$

where:

f : Analytical function.

h : Histogram with Monte Carlo distribution.

corr : Detector correction function.

$\kappa$  : Composition of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ .

<sup>8</sup> Or Nine Dimensional Matrix Element Method (9DMEM).

The background distributions are based on a factorised model of histogram templates. The production observables are grouped as for the signal, while the decay observables are included as  $(\phi)$ ,  $(m_{12}, m_{34})$  and  $(\cos \theta_1, \cos \theta_2)$  factors.

The approach demonstrates similar abilities to the ME-Obs approach in closure, asymptotic behaviour, sensitivity to systematic uncertainties and ability to exclude models. The analysis is mentioned here, because the two methods; The Matrix Element Observable and the Nine Dimensional Fit are used as independent cross checks of one and the other. For the final validation, a set of 300 Monte Carlo samples is created. The samples are generated with Standard Model expectations and signal strength from observed data. Finally, the MC samples are created in such a way that they will have to pass through each analysis in the same fashion as the observed data.

The outcome of the checks is presented in Figure 3.26. The best fitting values of the two

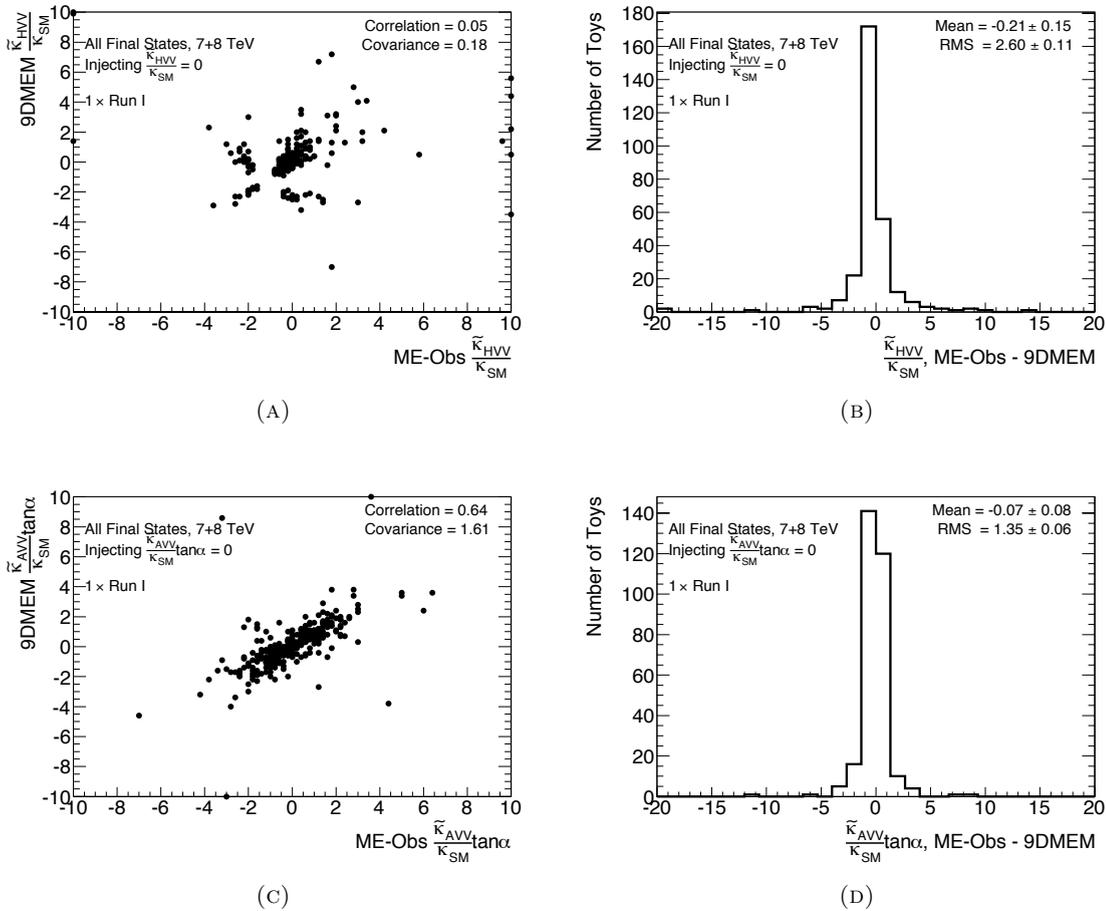


FIGURE 3.26: Compatibility checks between Matrix Element Observable fit and Nine Dimensional Fit. A set of 300 MC samples is generated with expectations profiled to data. Figures show the maximum likelihood estimator of the coupling ratios for the two different analyses.

(A,B)  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  correlation and difference. (C,D)  $\frac{\tilde{\kappa}_{AVV} \tan \alpha}{\kappa_{SM}}$  correlation and difference.

methods are to a high degree compatible. The structure of the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  correlation can e.g. be understood by examining Figure 3.16a. It is highly unlikely that a Standard Model Higgs will be fitted to  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = -1$ . On the other hand  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = -2.5$  is much more compatible with the

Standard Model. There is hence a chance that one method will fit to one side of the peak and the other method to the opposite side.

The two independent approaches are seen to be in good agreement, thereby validating them individually.

### 3.4.9 Summary of the Tensor Structure Fit

- At the first stage of the analysis, an event selection is applied to a Monte Carlo sample with coupling parameters  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 1 + i$ ,  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 1 + i$ , to the  $ZZ^*$  background samples simulated in the  $100 \text{ GeV} < m_{4\ell} < 150 \text{ GeV}$  region and to reducible backgrounds estimated from control regions in data. A signal region is defined as events falling in the mass window:  $115 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$ . Observables are defined by Matrix element ratios. The observables are sensitive to the coupling ratios  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  of the underlying resonance. A BDT discriminant is also trained on parameters related to Higgs production in order to suppress backgrounds.
- In the second step, Matrix Element based reweighting is used to obtain a description of signal models over the examined intervals of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The approach is validated by comparing the resulting distributions to those of independent Monte Carlo samples. A step size of 0.25 is used in the  $[-10, 10]$  interval of each parameter, resulting in 81 points on each axis.
- In the next stage, the simulated signal and background distributions for the matrix element observables and background discriminant are described using three-dimensional histograms. All non-trivial correlations can be taken into account in this way. The parent distributions are subsequently estimated using a Gaussian KDE. The KDE helps to reduce the influence of statistical fluctuations and improve description of tail distributions. The procedure is performed separately for all 4 final states considered in the analysis:  $4e$ ,  $4\mu$ ,  $2e2\mu$  and  $2\mu 2e$ , each of which are divided into the two centre of mass energies. A separate histogram is created for each systematic uncertainty affecting shape variations. Piece-wise linear *pdf* interpolation is used to acquire a continuous description of the three-dimensional histograms in  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ ,  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  and all nuisance parameters.
- In the final stage of the analysis, the profiled likelihood ratio is used to establish the intervals of coupling ratios that are excluded at 95% CL. Asymptotic behaviour of the test statistic distribution is assumed. The expected Standard Model sensitivities are based on Asimov data samples.

### 3.4.10 Results

The analysis is finally applied to the observed Run-1 data of  $4.5 \text{ fb}^{-1}$  taken at  $\sqrt{s} = 7 \text{ TeV}$  and  $20.3 \text{ fb}^{-1}$  taken at  $\sqrt{s} = 8 \text{ TeV}$ . In order to get the most valid comparison to Standard Model expectations, a fit is performed on data, where SM is assumed but all nuisance parameters are free. Figures 3.27a and 3.27b show the expected and observed results for the combined likelihood model and the individual final states. The fits include all systematic uncertainties that are found to potentially be influential. The 2011 and 2012 signal strengths and the Higgs mass are freely floating, while all others are under a Gaussian constraint.

Examining the figures, a good agreement between Standard Model expectations and observations is seen. The individual final states all find a best fitting value close to the Standard Model.

The best fitting values and excluded regions are listed in Table 3.11. As described in Eq. 1.8 of Section 1.2.4 the results can be re-interpreted as limiting the effective BSM cross section contribution. The results are also presented according to this scheme.

A comparison between expected and observed likelihood curves are seen in Figure 3.27c.

It is concluded that the observations are in agreement with Standard Model predictions.

Coupling ratio	Best-fit value		95% CL Exclusion Regions	
	Observed	SM Expectation	Observed	
$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$	-0.2	$(-\infty, -0.75] \cup [6.95, \infty)$	$(-\infty, -0.75] \cup [2.45, \infty)$	
$f_{g2} \cdot \cos \phi_{g2}$	-0.01	$[-1, -0.16] \cup [0.94, 1]$	$[-1, -0.16] \cup [0.68, 1]$	
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	-0.8	$(-\infty, -2.95] \cup [2.95, \infty)$	$(-\infty, -2.85] \cup [0.95, \infty)$	
$f_{g4} \cdot \cos \phi_{g4}$	-0.08	$[-1, -0.56] \cup [0.56, 1]$	$[-1, -0.54] \cup [0.11, 1]$	

TABLE 3.11: Expected and observed best-fit values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  and 95% CL excluded regions obtained in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. The expected values are estimated for an Asimov sample generated with signal strength and systematic uncertainties as measured in data. The data for  $\sqrt{s} = 7 \text{ TeV}$  and  $\sqrt{s} = 8 \text{ TeV}$  are combined. The results are also given in effective coupling ratio fractions  $f_{g2}$ ,  $f_{g4}$ . The angle  $\phi_{gi}$  is either 0 for positive coupling ratios or  $\pi$  for negative. Results published in [5].

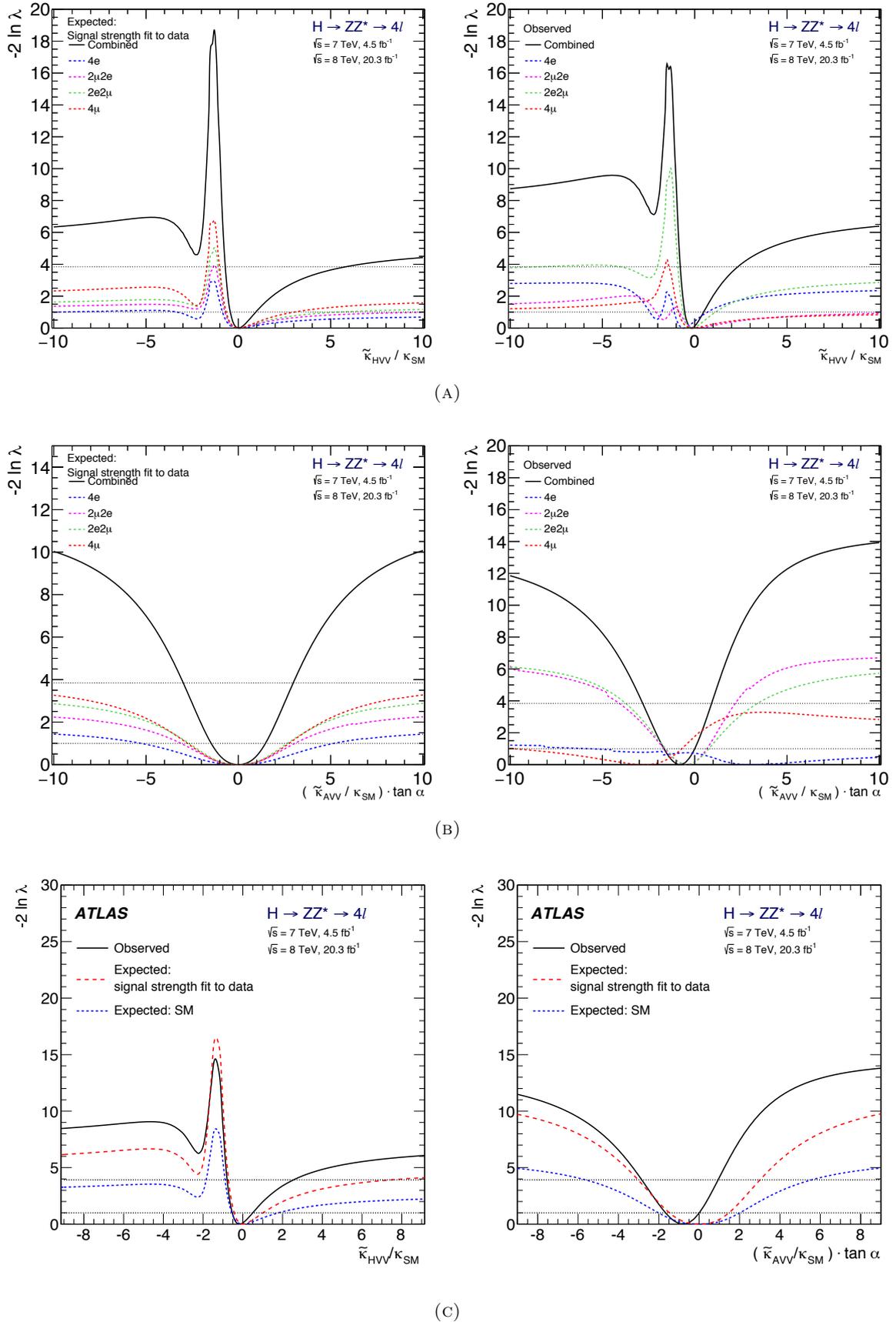


FIGURE 3.27: Expected and observed fit results per final state and their combination. The likelihood curves represent the outcome for the LHC Run-1 dataset. The expected curves are derived from a Standard Model Asimov sample where nuisance parameters have been fitted to data. (A)  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  SM expectations and observations per final state. (B)  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  SM expectations and observations per final state. (C) Expected and observed fit results of the tensor structure analysis. ATLAS figures published in [5].

### 3.5 The $ZZ^*$ and $WW^*$ Combination

Up until now, only  $ZZ$  couplings to the observed resonance have been studied. To access the complete information that can be derived from the LHC Run-1 dataset, the  $WW$  channel will now be included. Since there are no overlapping candidates between the  $WW$  and  $ZZ$  channels, the combination can be achieved by considering the two decay types as separate channels of the likelihood. This is similar to how the different lepton final states and centre of mass energies are treated as separate channels of the  $ZZ$  likelihood in Equation 3.5. Disregarding systematic uncertainties it would be a simple matter of performing the separate analyses and adding their likelihood functions together. The two channels however share systematic uncertainties from theoretical descriptions, luminosity uncertainties and because both consider leptonic final states. The challenge of the combination is thus to identify the common systematics and correlate these. First a short outline of the  $WW$  analysis is given:

#### 3.5.1 The $WW^*$ Final State

The main difference to the  $ZZ^*$  channel is also one of the main challenges, the leptonic decay of  $W$  bosons contain a neutrino. It follows from the Higgs mass and  $p_T$ -spectrum that the  $W$  bosons will not be very boosted. This means that the direction the neutrinos will travel in cannot be approximated and the final state is as a result not fully recoverable. It will therefore not be directly possible to use Matrix Element Observables as was done in the  $ZZ^*$  analysis. An MVA approach is instead opted for, using BDTs.

The  $WW^*$  decay also have significantly higher backgrounds as compared to  $ZZ^*$ . Both because of the increased difficulty in rejecting backgrounds using missing transverse energy compared to using a pair of reconstructed leptons and because the invariant mass peak is smeared over a much larger mass range than for the  $ZZ^*$  analysis. Since only two of the four leptons are possible to reconstruct, the analysis solely focus on the  $e\nu\mu\nu$  channel. The other channels are found to have overwhelming backgrounds. By using e.g.  $e\nu e\nu$  final states the analysis would for instance be prone to accept  $Z + X$  types of backgrounds. The  $WW^*$  analysis moreover only considers the zero jet category of the 8 TeV dataset.

The structure follows much the same logic as what was described in the fixed hypothesis test section: Two BDTs are trained, one separating a SM Higgs from an alternative spin-0 model ( $BDT_{CP}$ ) and one separating a SM Higgs from backgrounds ( $BDT_0$ ). The parameters used to train the  $BDT$ s are as follows:

$$\begin{aligned}
 BDT_0 (0^+ \text{ vs. bkg}) & : p_T^{\ell\ell}, m_{\ell\ell}, \Delta\phi_{\ell\ell}, p_T^{miss} \\
 BDT_{CP} (0^+ \text{ vs. } 0_h^+) & : p_T^{\ell\ell}, m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_T \\
 BDT_{CP} (0^+ \text{ vs. } 0^-) & : m_{\ell\ell}, E_{\ell\nu\nu}, \Delta p_T, \Delta\phi_{\ell\ell} ,
 \end{aligned}$$

where:

$p_T^{\ell\ell}$  : Transverse component of lepton momentum sum.

$m_{\ell\ell}$  : Invariant mass of the lepton momentum sum.

$\Delta\phi_{\ell\ell}$  : Azimuthal angle between the leptons.

$p_T^{miss}$  : Missing transverse energy.

$m_T$  : Transverse mass of leptons and missing energy.

$E_{\ell\nu\nu}$  :  $p_T^{\ell_1} - 0.5 \cdot p_T^{\ell_2} + 0.5 p_T^{miss}$  for transverse lepton momenta  $p_T^{\ell_1}$  and  $p_T^{\ell_2}$

The distributions for various relevant processes can be found in Figure 3.28.

The  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  fit will be based on the BDT trained against  $0_h^+$  and similarly the BDT trained against  $0^-$  will be used in the  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan\alpha$  fit.

The analysis uses two-dimensional templates to describe possible correlations between  $BDT_{CP}$  and  $BDT_0$ . An alternative approach to the KDE is chosen to ensure that there do not exist empty bins. Instead, the 2D histograms are unrolled to a 1D distribution, such that the new bin numbering  $i_{CP,0}$  is given by:

$$i_{CP,0} = j_{CP} + N_{CP} \cdot k_0 ,$$

where  $N_{CP}$  is the number of  $BDT_{CP}$  bins and  $j, k$  are the indices of  $BDT_{CP}$  and  $BDT_0$ , respectively. Any empty bin is merged with its neighbour after the unrolling. The subsequent approach follows the same structure as described for the  $ZZ^*$  analysis: Construct the likelihood model, maximise the likelihood as function of the coupling ratios and use the profiled likelihood ratio as test statistic. A comprehensive description of the analysis is given in [97].

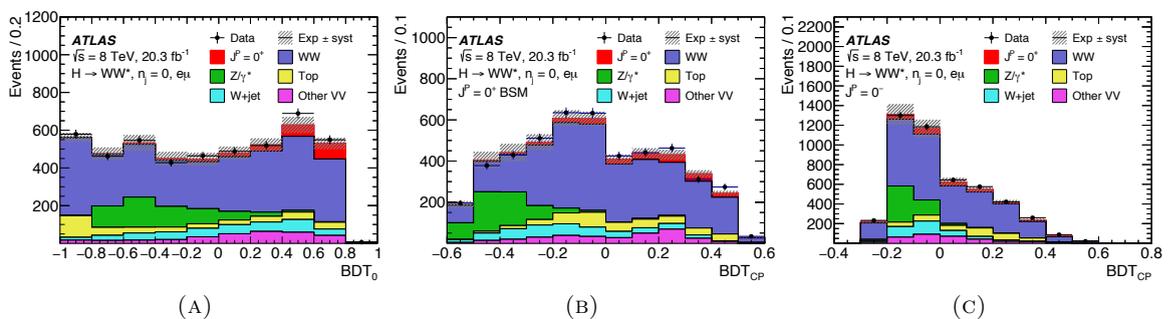


FIGURE 3.28: BDT discriminant used in the  $WW^*$  analysis. (A)  $BDT_0$  optimised to separate a SM Higgs and backgrounds. (B)  $BDT_{CP}$  optimised to separate a SM Higgs from the CP-Even spin-0 boson  $0_h^+$  (C)  $BDT_{CP}$  optimised to separate a SM Higgs from the pseudo-scalar  $0^-$  [97].

### 3.5.2 Correlating Systematic Uncertainties

The next step of the combination is to identify which systematic uncertainties the two analyses share. The combined likelihood will be the product of the two individual ones, but where the

shared systematics are correlated and the constraint term only contain the minimal set of unique uncertainties:

$$\mathcal{L}_{\text{Comb}} = \mathcal{L}_{ZZ} \cdot \mathcal{L}_{WW} \cdot \mathcal{C}(\theta_{\text{Unique}})$$

The analysis of systematic uncertainties that was conducted for the  $ZZ$  analysis demonstrated two things: The analysis is at the current level of available data completely dominated by statistical uncertainties and a large part of the relevant systematic uncertainties are related to the description of reducible backgrounds which, do not relate to the  $WW$  analysis. In the end, after the pruning of the  $ZZ$  systematic uncertainties the only ones shared between the two decay channels are:

`alpha_Lumi_2012`: Uncertainty in the luminosity collected by ATLAS for the 2012 dataset.

`alpha_pdf_qq`: Parton density uncertainty for  $VV^*$  production.

`alpha_QCDscale_VV`: Renormalisation and factorisation scale of background production.

In conclusion it is the uncertainty in background process production and the collected dataset size that will simultaneously affect the two analysis.

### 3.5.3 Results

The first result that is presented will examine how well the  $ZZ$  and  $WW$  analyses agree. Looking at the Lagrangian density of Eq. 1.6, a different coupling has been assigned to the  $ZZ$  and  $WW$  sectors. Having constructed the combined likelihood, it is possible to perform a fit on the separate BSM/SM coupling ratios for the individual decay channels. In principle a deviation in the best fitting value between the two would violate custodial symmetry. The result is presented in Figure 3.29, where it is seen that the fitted results are in agreement with the Standard Model expectations.

The peculiar shape of the  $\frac{\tilde{\kappa}_{HWW}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{HZZ}}{\kappa_{SM}}$  landscape can be understood from the behaviour of the individual results, see of Figure 3.30b. The  $WW$  channel has a degenerate minimum at  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = -1$  and  $-0.5$ , where  $-1$  is slightly preferred. This correspond to the two minima that are found in the vertical direction. The scan also more or less has a constant test statistic value of 5 at large BSM fractions, corresponding to the vertical bar of Figure 3.29.

The separate fit result indicate good agreement between the two decay-channels. It is well motivated both theoretically and by observations to combined the two coupling ratios:

$$\left( \frac{\tilde{\kappa}_{HZZ}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{HWW}}{\kappa_{SM}} \right) \rightarrow \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} \quad \left( \frac{\tilde{\kappa}_{AZZ}}{\kappa_{SM}} \tan \alpha, \frac{\tilde{\kappa}_{AWW}}{\kappa_{SM}} \tan \alpha \right) \rightarrow \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$$

The fits are repeated to create the final results. Note that the combined parameter results will appear significantly more constraining that what is indicated by Figure 3.29. This is both because the  $ZZ$  and  $WW$  information is combined, but also because the parameter combination will create a different interpretation of the profiled likelihood test statistic. The 68%

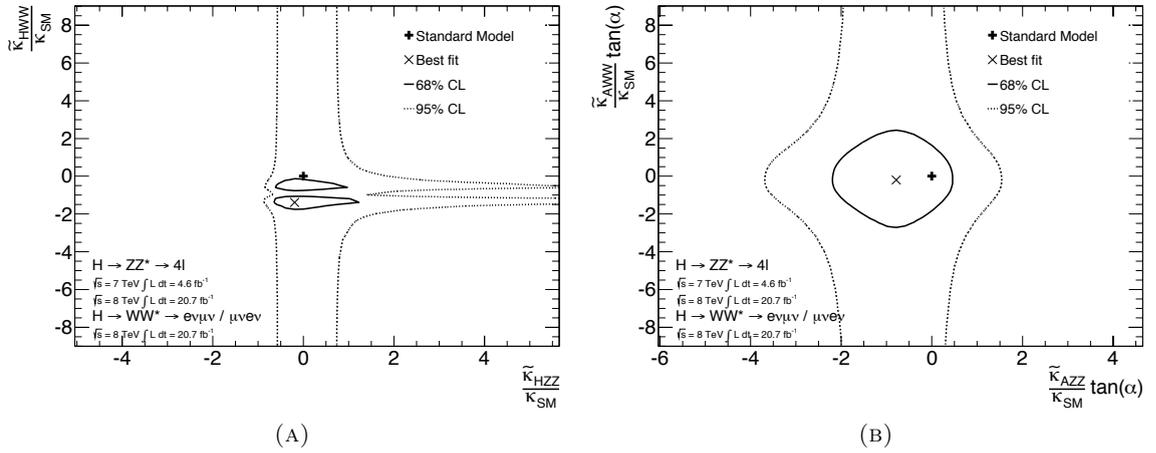


FIGURE 3.29: Simultaneous fit of  $\frac{\tilde{\kappa}_{HZZ}}{\kappa_{SM}}, \frac{\tilde{\kappa}_{HWW}}{\kappa_{SM}}$  (A) and  $\frac{\tilde{\kappa}_{AZZ}}{\kappa_{SM}} \tan \alpha, \frac{\tilde{\kappa}_{AWW}}{\kappa_{SM}} \tan \alpha$  (B). The contours are meant to be interpreted as everything that is within the solid line are within the 68% CL and everything within the dashed line is within the 95% CL. (A) has a degenerate minimum in the  $WW$  direction.

and 95% CL are set assuming asymptotic behaviour which means that the test statistic distribution follows a  $\chi^2_N$  distribution with degrees of freedom equal to the dimensionality of the parameter of interest. Combining the two parameters will make the dimension go from two to one and as a result the 95% CL is found where the likelihood crosses 3.84 instead of 5.99. The results are presented in Table 3.12. The limits are again reinterpreted as effective BSM cross section contributions. The combined and individual likelihood curves are found in Figure 3.30. Everything is again concluded to be in agreement with Standard Model expectations.

Coupling ratio	Best-fit value	95% CL Exclusion Regions	
Combined $ZZ^*, WW^*$	Observed	SM Expectation	Observed
$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$	-0.48	$(-\infty, -0.55] \cup [4.80, \infty)$	$(-\infty, -0.73] \cup [0.63, \infty)$
$f_{g2} \cdot \cos \phi_{g2}$	-0.074	$[-1, -0.096] \cup [0.89, 1]$	$[-1, -0.16] \cup [0.12, 1]$
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	-0.68	$(-\infty, -2.33] \cup [2.30, \infty)$	$(-\infty, -2.18] \cup [0.83, \infty)$
$f_{g4} \cdot \cos \phi_{g4}$	-0.062	$[-1, -0.44] \cup [0.43, 1]$	$[-1, -0.41] \cup [0.090, 1]$

TABLE 3.12: Expected and observed best-fit values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  and regions excluded at 95% CL obtained in the combined  $H \rightarrow ZZ^* \rightarrow 4l$  and  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  analysis. The expected values are estimated for an Asimov sample generated with the nuisance parameters found in data. The data for  $\sqrt{s} = 7$  TeV ( $ZZ^*$  only) and  $\sqrt{s} = 8$  TeV are combined. The results are also given in effective coupling ratio fractions  $f_{g2}, f_{g4}$ . The angle  $\phi_{gi}$  is either 0 for positive coupling ratios or  $\pi$  for negative. Results published in [5].

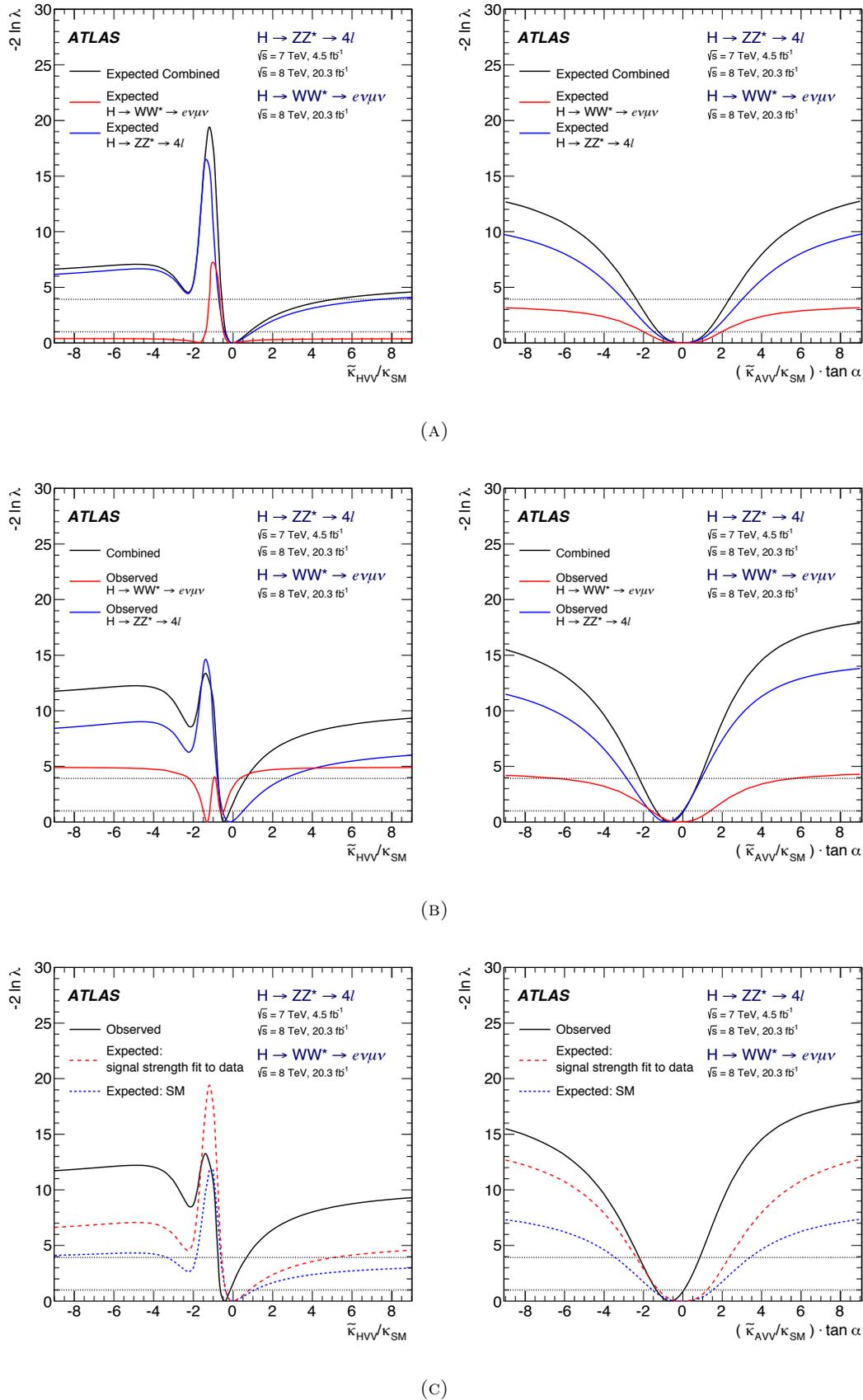


FIGURE 3.30: Expected and observed results of the  $HVV$  tensor structure analysis. Likelihood curves obtained for  $H \rightarrow ZZ^* \rightarrow 4l$ ,  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  and their combination. The data for  $\sqrt{s} = 7$  TeV (only  $ZZ^*$ ) and  $\sqrt{s} = 8$  TeV are combined. Left column show  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  scans while right column show  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  results. (A) Individual and combined expectations. (B) Individual and combined observations. (C) Expected and Observed comparison. Figures published in [5].

### 3.6 Simultaneous Fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$

The analyses so far have assumed that data could be described by having at most one BSM contribution to the effective Lagrangian. Beside it being the simplest approach, there is no real theoretical motivation to make this assumption. A more ideal approach is to allow the coexistence of all three Lagrangian terms ( $\kappa_{SM}$ ,  $\kappa_{HVV}$ ,  $\kappa_{AVV}$ ). The full inclusion will make it possible to perform a simultaneous fit of both coupling ratios, or perform the likelihood scan of one over the profile of the other<sup>9</sup>. The following section will examine how a simultaneous fit can be constructed using the Matrix Element Observable approach of fitting coupling ratios.

Ignoring how much Monte Carlo is available or reasonable to produce, the simplest possible extension would be to create five-dimension histograms with the four Matrix Element Observables and background discriminant. Using the same (conservative) binning as is done for the one-parameter analysis, the histogram would consist of  $\simeq 60000$  bins. It would be virtually impossible to populate such a histogram with simulated events.

If on the other hand it is possible to disregard some of the observable correlations, a factorised model can be used. Since the observables have been defined with separate use; two to measure  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ ; two to measure  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  and one to discriminate backgrounds, a natural factorisation scheme would be:

$$\begin{aligned} pdf(\vec{O}) &= pdf(TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV}), TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})) \cdot \\ &pdf(TO_1(\tilde{\kappa}_{AVV}, \alpha), TO_2(\tilde{\kappa}_{AVV}, \alpha)) \cdot pdf(BDT_{ZZ}) \end{aligned} \quad (3.6)$$

In the factorised scheme it will then at most be necessary to describe two dimensional distributions.

Even more than for the one-parameter analysis, it will be necessary to assume asymptotic test statistic behaviour. In that case, the test statistic will follow a  $\chi^2_{\text{Dim}=2}$  distribution and limits can readily be set. One necessary criterion for the assumption is however that the factorised observables are not correlated<sup>10</sup>.

Figure 3.31 shows the dependency between the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observables. A correlation between  $TO_2(\tilde{\kappa}_{AVV}, \alpha)$  and  $TO_1(\tilde{\kappa}_{HVV}) \pm TO_2(\tilde{\kappa}_{HVV})$  is seen. The dependency is suspected to propagate to the test statistic distribution but it is however difficult to assess how severely the correlation will affect the result solely from looking at the observable distributions. To estimate how much the factorisation scheme will overconstrain the fitted parameters, a series of tests are performed:

The *pdfs* that are assumed to be uncorrelated in Eq. 3.6 are examined two by two. In each case, a two-dimensional factorised model is constructed from a  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and a  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observable. The likelihood contour of the factorised model is compared to the model that takes into account their dependency. In the first test for instance, the likelihood contours for the two

<sup>9</sup>Profiling here means that the likelihood is minimised with respect to the coupling ratio that is not scanned over. In other words the coupling ratio that is not scanned over is treated as a nuisance parameter.

<sup>10</sup>A short discussion of the problems of factorising with correlated observables is given in App. A.3.

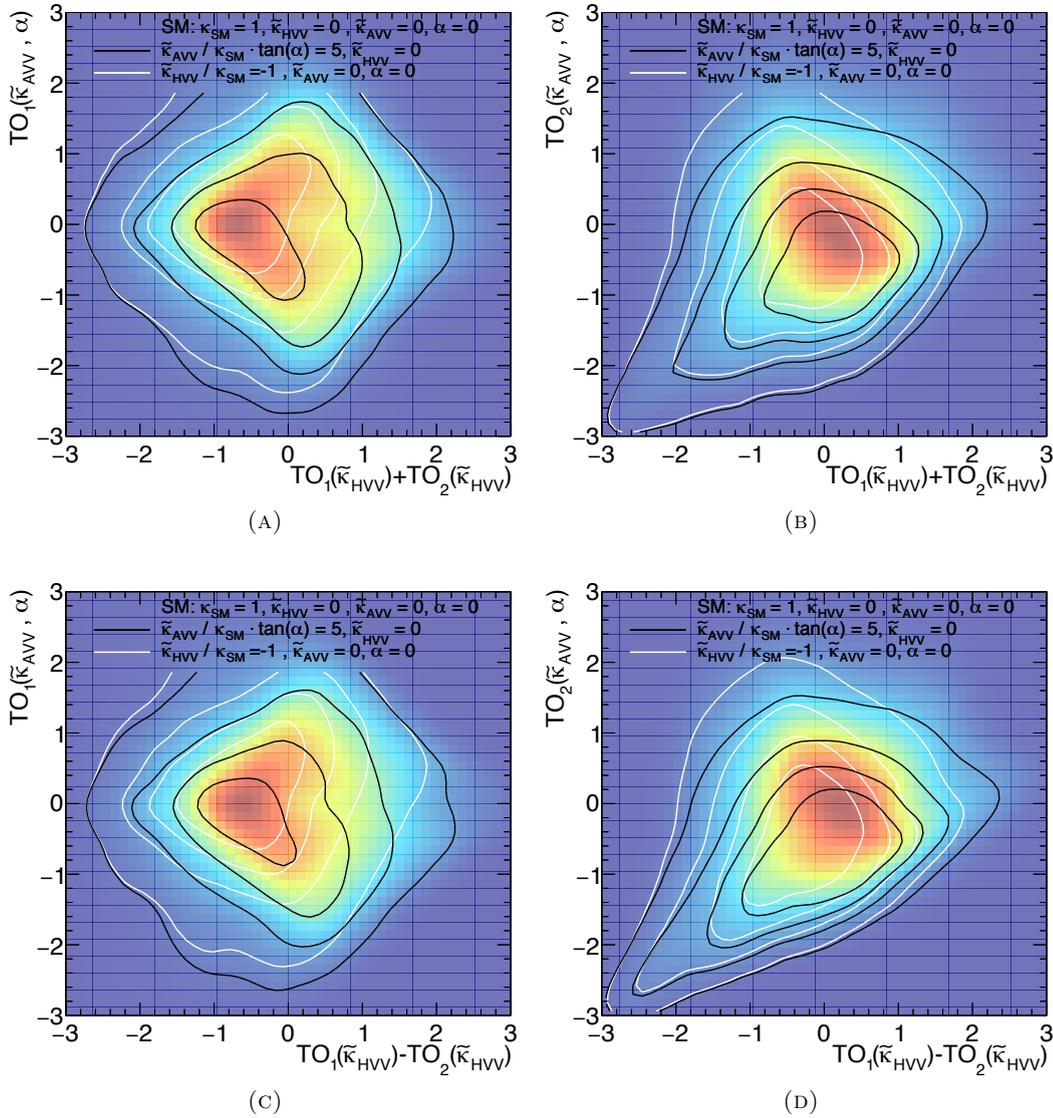


FIGURE 3.31: Correlation between  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$  observables. (A,B):  $TO_1(\tilde{\kappa}_{HV V}) + TO_2(\tilde{\kappa}_{HV V})$  vs.  $TO_1(\tilde{\kappa}_{AV V}, \alpha)$  and  $O_2(\tilde{\kappa}_{AV V}, \alpha)$ . (C,D):  $TO_1(\tilde{\kappa}_{HV V}) - TO_2(\tilde{\kappa}_{HV V})$  vs.  $TO_1(\tilde{\kappa}_{AV V}, \alpha)$  and  $O_2(\tilde{\kappa}_{AV V}, \alpha)$ . A dependency to  $TO_2(\tilde{\kappa}_{AV V}, \alpha)$  is seen.

models:

$$\begin{aligned}
 pdf(\bar{O})_{\text{Full}} &= pdf(TO_1(\tilde{\kappa}_{HV V}) + TO_2(\tilde{\kappa}_{HV V}), TO_1(\tilde{\kappa}_{AV V}, \alpha)) \cdot pdf(BDT_{ZZ}) \\
 pdf(\bar{O})_{\text{Fact.}} &= pdf(TO_1(\tilde{\kappa}_{HV V}) + TO_2(\tilde{\kappa}_{HV V})) \cdot pdf(TO_1(\tilde{\kappa}_{AV V}, \alpha)) \cdot pdf(BDT_{ZZ}),
 \end{aligned}$$

are compared. The test will show if the fitted parameters will be overconstrained by assuming that the two  $pdfs$  are independent.

The tests are continued for each of the observable distributions that are assumed to be independent. It should be noted here, that the background discriminant is assumed to be independent from the other observables. The assumption is motivated theoretically by the fact that the boosted decision tree is only trained on Higgs production related parameters

while the models that are examined are explicitly constructed to be independent of the boson's production mechanism (gluon fusion). The correlation is moreover shown in Figure 3.12d and C.11. The outcome of the observable dependency tests is shown in Figure 3.32.

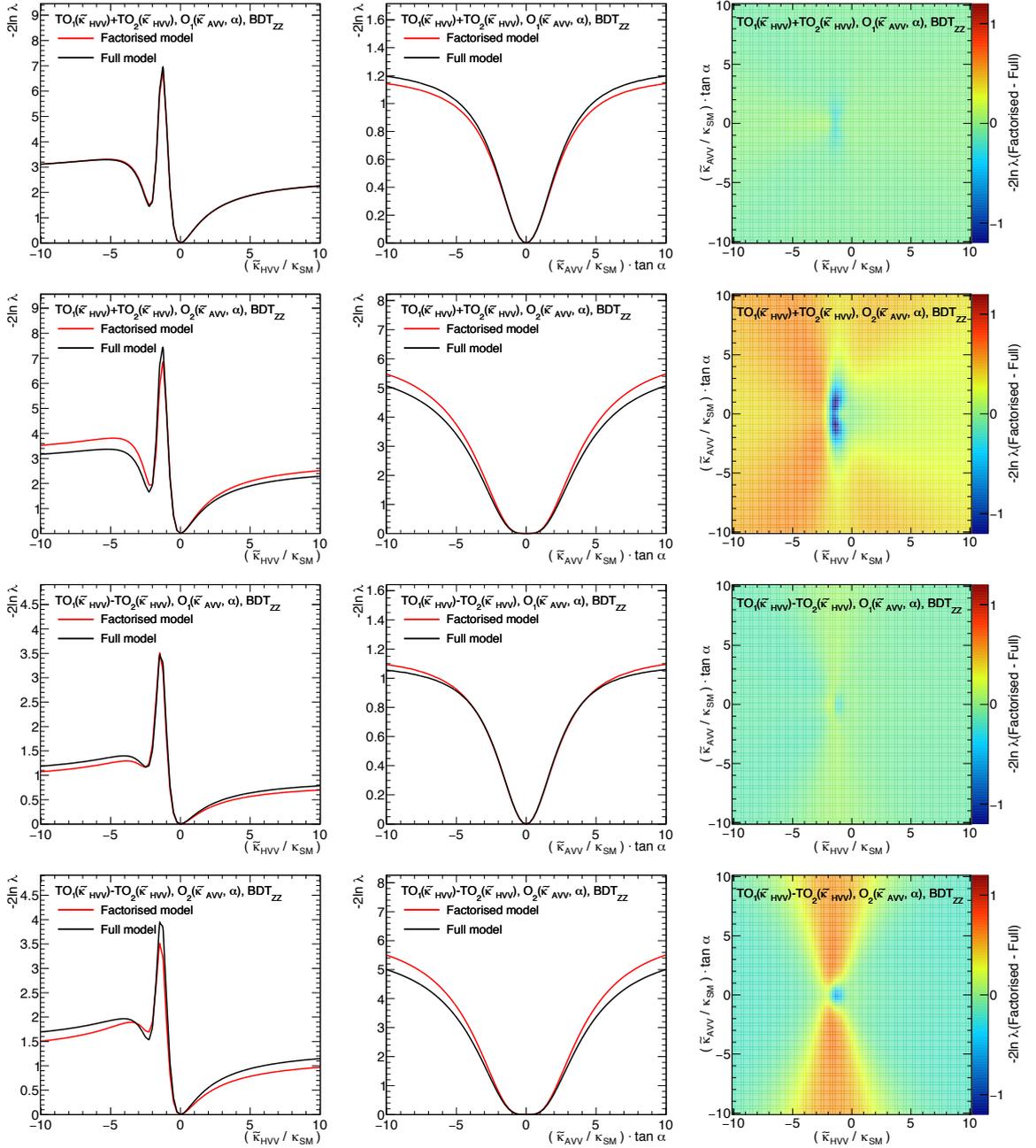


FIGURE 3.32: Likelihood contours obtained with a full two dimensional observable description and a factorised model. In all cases the background discriminant  $BDT_{ZZ}$  is added as a separate factor. A simultaneous scan over  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  is performed. The left column show the scan over  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  assuming  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  to be zero. The centre column show the  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  scan when  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  is assumed to be zero, while the rightmost column show the difference between the factorised and full model. A red colour indicate that the factorised model will overestimate the likelihood. The four rows show the four different combinations of  $pdfs$  that are assumed independent.

Assuming  $O_2(\tilde{\kappa}_{AVV}, \alpha)$  to be independent of the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  observables (row two and four) give slightly higher likelihood values as expected. The effect is stronger at coupling ratios far from the Standard Model.

The difference is seen to be around 0.5 in the most severe cases. To give an idea of how much the fitted parameter coverage changes, it can be mentioned that the tail integral of a  $\chi_2^2$  distribution from 5.5 is 0.063 compared to the 0.05 at 5.99, where limits are usually set. Likewise, the tail integral of a  $\chi_1^2$  from 3.3 is 0.07 compared to the 0.05 from 3.84 where limits are usually set. The discrepancy is concluded to be within acceptable deviations, but it is noted that the 95% CL of the following section should be considered approximate or conversely should be interpreted as  $\lesssim$  95% CL.

Further attempts at decorrelating the observables is outlined in Appendix D. The appendix describes the use of a transformation matrix that result in a set of observables with covariance given by the identity matrix.

### 3.6.1 Results

The likelihood model is created similarly to the one-dimensional tensor structure fit. For technical reasons it is not possible to perform *pdf*-interpolation between models as in the one-parameter case. Instead, the likelihood values are calculated in a grid and the intermediate likelihood values are estimated with linear interpolation. The grid is chosen as  $81 \times 81$  points from  $-10$  to  $10$  in either direction.

The likelihood model is here constructed without systematic uncertainties. It was demonstrated in Section 3.4.7 that it is not expected that the result will change in any noticeable way from this. The computational demand of fitting in the two-dimensional plane with systematic uncertainties is however problematic.

A model is constructed where the background normalisations are allowed to float unconstrained in order to assess how much the result would change by including systematic uncertainties. The relative normalisation of the individual backgrounds is kept constant such that only one extra nuisance parameter is added. The limits set on  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  for a nominal Standard Model Asimov sample was observed to change from 2.5 to 2.6.

The analysis is applied to the observed Run-1 data of  $4.5 \text{ fb}^{-1}$  taken at  $\sqrt{s} = 7 \text{ TeV}$  and  $20.3 \text{ fb}^{-1}$  taken at  $\sqrt{s} = 8 \text{ TeV}$ . In order to get the most valid comparison to Standard Model expectations, a fit is performed on a Asimov sample where SM is assumed but the model signal strength is taken from data. The first set of results are presented in Figure 3.33. The observed and expected contours are compared. Stronger limits than expected are observed for positive  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  as in the one-dimensional case. The contour is seen to be flat in this area such that relatively small changes in likelihood values will result in large changes of the limit.

The one dimensional results are moreover examined. Both, where each parameter is fitted, while the other is assumed to be zero, and where the other is minimised (profiled). A kink in the profiled likelihood curve is seen at around  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 3$  as a result of a degenerate minimum in  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ . The  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  scan will be minimised for positive values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  in the beginning and when reaching  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha \simeq 3$  will suddenly change to be minimised in the

valley seen in the upper left corner of the plane.

The profile is in some areas observed to be a stronger discriminant than when fixing the alternate parameter to zero. This can happen if the two-dimensional fit is not minimised at zero either. The expected and observed limits and best fitting values found in the figures are presented in Table 3.13.

Figure 3.34 show the results when they are re-interpreted as limiting the effective contributions to the combined cross section in terms of  $f_{g2}$  and  $f_{g4}$ . The effective cross section results are created by interpolating the likelihood values found in the coupling ratio fit. The one parameter limits can directly be calculated from Table 3.33, while the 'profile' limits are dependent on where the minimum of the alternate parameter is found. The profile values can as such be understood as limiting one BSM contribution independent on whether the remaining cross section stems from the Standard Model or the alternate BSM sector. The corresponding best fitting values and limits are listed in Table 3.14.

The results are concluded to be in agreement with the Standard Model.

Coupling ratio	Best-fit value	$\lesssim 95\%$ CL Exclusion Regions	
		SM Expectation	Observed
$H \rightarrow ZZ^* \rightarrow 4\ell$	Observed		
$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ with $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = 0$	-0.30	$(-\infty, -0.67] \cup [3.7, \infty)$	$(-\infty, -0.87] \cup [0.97, \infty)$
$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ with $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ profiled	-0.50	$(-\infty, -0.77] \cup [3.7, \infty)$	$(-\infty, -0.97] \cup [1.0, \infty)$
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ with $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = 0$	-0.74	$(-\infty, -2.5] \cup [2.5, \infty)$	$(-\infty, -2.3] \cup [0.92, \infty)$
$\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ with $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ profiled	-0.50	$(-\infty, -3.4] \cup [3.3, \infty)$	$(-\infty, -2.1] \cup [0.82, \infty)$

TABLE 3.13: Observed best fitting values and limits for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ . The expected and observed limits are given both when the alternate parameter is assumed zero and when profiled. Values correspond to the full LHC Run-I dataset.

Contribution to cross section	Best-fit value	$\lesssim 95\%$ CL Exclusion Regions	
		SM Expectation	Observed
$H \rightarrow ZZ^* \rightarrow 4\ell$	Observed		
$f_{g2} \cdot \cos \phi_{g2}$ w. $f_{g4} \cdot \cos \phi_{g4} = 0$	-0.030	$[-1, -0.14] \cup [0.83, 1]$	$[-1, -0.21] \cup [0.25, 1]$
$f_{g2} \cdot \cos \phi_{g2}$ w. $f_{g4} \cdot \cos \phi_{g4}$ profiled	-0.074	$[-1, -0.16] \cup [0.83, 1]$	$[-1, -0.24] \cup [0.24, 1]$
$f_{g4} \cdot \cos \phi_{g4}$ w. $f_{g2} \cdot \cos \phi_{g2} = 0$	-0.069	$[-1, -0.47] \cup [0.47, 1]$	$[-1, -0.43] \cup [0.11, 1]$
$f_{g4} \cdot \cos \phi_{g4}$ w. $f_{g2} \cdot \cos \phi_{g2}$ profiled	-0.035	$[-1, -0.52] \cup [0.52, 1]$	$[-1, -0.39] \cup [0.082, 1]$

TABLE 3.14: Observed best fitting values and limits for  $f_{g2} \cdot \cos \phi_{g2}$  and  $f_{g4} \cdot \cos \phi_{g4}$ . The expected and observed limits are given both when the alternate parameter is assumed zero and when profiled. Values correspond to the full LHC Run-I dataset.

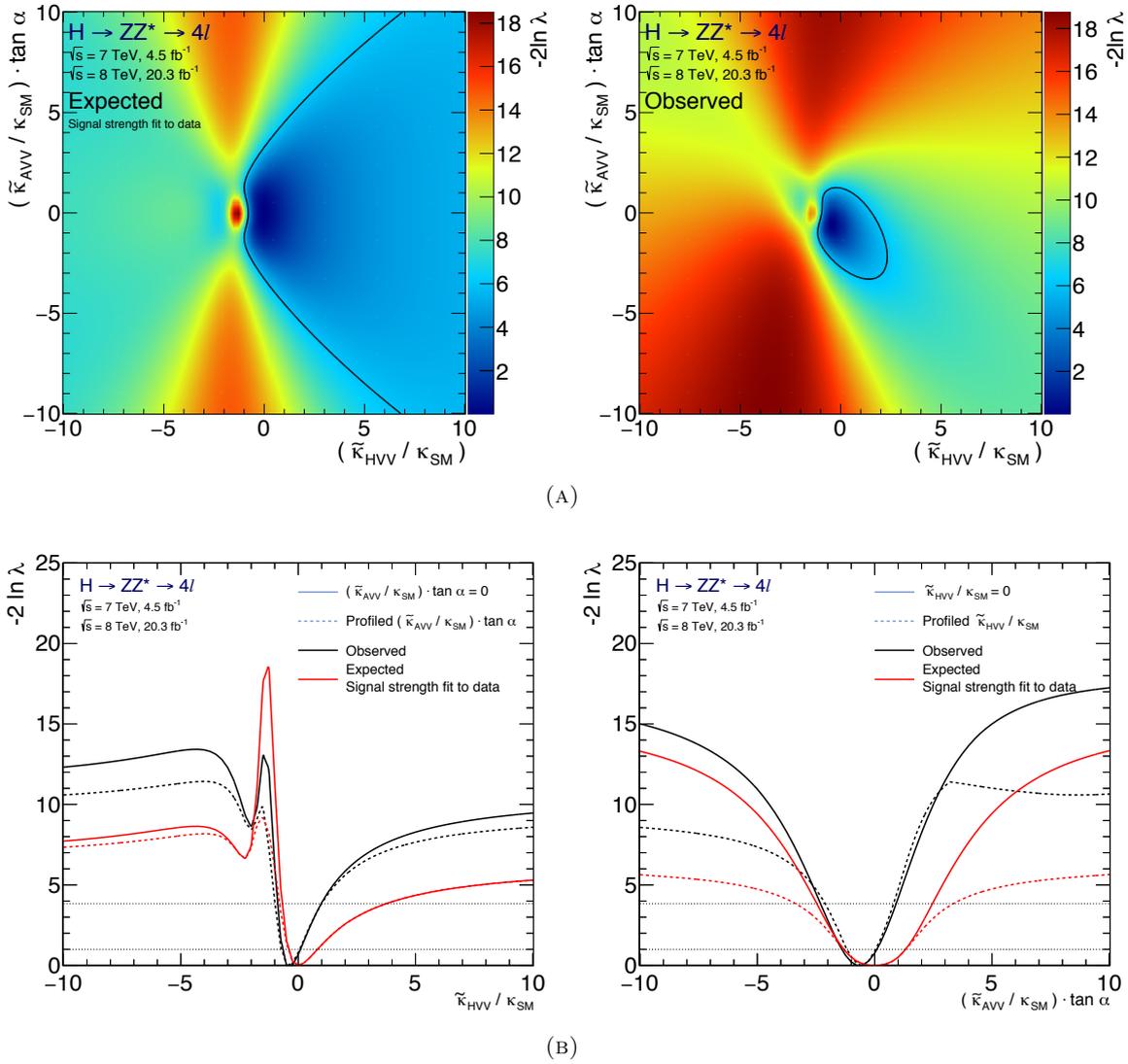


FIGURE 3.33: Simultaneous fit of  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$ . Expected and observed likelihood landscapes (A) and likelihood curves (B). The alternate parameter is both assumed zero and profiled in the single parameter scans. The figures show the analysis outcome using the full LHC Run-I dataset.

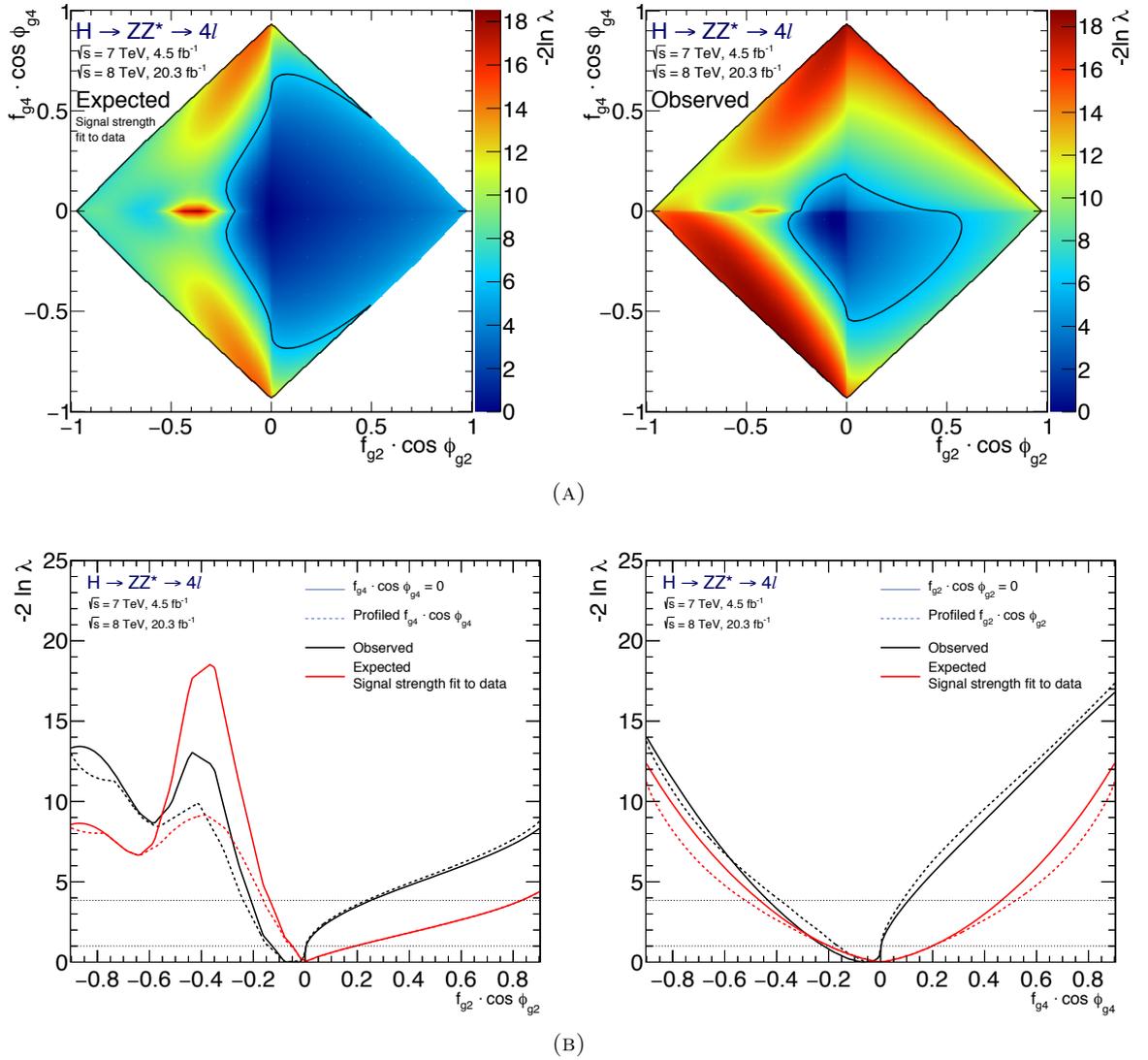


FIGURE 3.34: Expected and observed likelihood landscapes of  $f_{g2} \cdot \cos \phi_{g2}$  vs.  $f_{g4} \cdot \cos \phi_{g4}$  (A). The values are created by interpolating Figure 3.33. (B) Likelihood curves for the single parameter scan and the profiled scan. The figures show the analysis outcome using the full LHC Run-I dataset.

## 3.7 Outlook: High Luminosity Prospects

The future prospects of the tensor structure analysis is examined for high luminosity projections. Benchmark luminosities of  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$  are chosen, corresponding to the expected data-size after the LHC and HL-LHC.

### 3.7.1 Matrix Element Observables as a Probe of the $HZZ$ Vertex

The analysis described in this section follows with a few modifications the analysis structure of the previous parts<sup>11</sup>.

The structure of the  $HZZ$  vertex is here described in coupling ratios  $g_2/g_1$  and  $g_4/g_1$  following [46]. The BSM couplings  $g_2$  and  $g_4$  are in general allowed to be complex. The parameterisation can however readily be compared to the parameterisation used in previous sections:

$$\Re(g_2)/g_1 = \frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} \quad \Re(g_4)/g_1 = \frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$$

### Samples

A large Monte Carlo sample is generated with coupling composition: ( $g_1 = 1$ ,  $g_2 = 1 + i$ ,  $g_4 = 1 + i$ ) with an assumed boson mass of 125.5 GeV. The signal sample is generated using the JHU 3.1.8 Monte Carlo generator. Matrix element based reweighting is applied to get descriptions of other coupling composition using JHUGENMELA[46]. A number of BSM samples are independently generated to demonstrate the ability to close on different signal models.

MADGRAPH5[49] is used to generate a  $ZZ^*$ -continuum background sample around the signal mass region:  $100 \text{ GeV} < m_{4l} < 150 \text{ GeV}$ . The shape of the reducible backgrounds is assumed to be the same as  $ZZ^*$  but with a relative event yield of 50%.

The behaviour of the ATLAS detector is described by a simplified model. The efficiencies and resolutions of lepton reconstruction is described in [98]. Finally, a kinematic selection similar to what is described in 3.2.1 is applied to all samples. The expected event yield after the selection is listed in Table 3.15.

A set of systematic uncertainties describing variations in the expected number of events is moreover included. The list and sizes can be found in Table 3.16. The systematics will be included in the analysis likelihood model as scaling factors under Gaussian constraints.

<sup>11</sup>This analysis is historically the first to be performed. This is reflected by the inclusion of complex coupling ratios and a less developed use of matrix element observables. The analysis functioned as a feasibility study on which the Run-I analysis described above is based on. Results published in [4].

	Signal	Backgrounds
	$\sqrt{s} = 14 \text{ TeV}, 3000 \text{ fb}^{-1}$	
$4\mu$	1186	641
$2\mu 2e$	867	431
$2e 2\mu$	1035	574
$4e$	871	474
Total	3959	2120

TABLE 3.15: Expected signal and background yields for high luminosity projections, in the  $115 \text{ GeV} < m_{4\ell} < 130 \text{ GeV}$  signal region. The number of signal events is given for a Higgs boson mass of  $125.5 \text{ GeV}$ [4, Tab.2]. The  $300 \text{ fb}^{-1}$  numbers are assumed to be 1/10th of the above.

Uncertainty	Size
Luminosity uncertainty. Correlated between signal and backgrounds.	3%
Lepton reconstruction efficiency. Separate yield uncertainty for signal and backgrounds	5%
$ZZ^*$ and reducible background event yield for $300 \text{ fb}^{-1}$	9.4%
$ZZ^*$ and reducible background event yield for $3000 \text{ fb}^{-1}$	7.4%

TABLE 3.16: Assumed systematic uncertainties for high luminosity tensor structure fits[4].

### Likelihood model construction

The analysis makes use of the following set of Matrix Element observables:

$$\begin{aligned}
O(|g_2|/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = 0, g_4 = 0)|^2}{|ME(g_1 = 0, g_2 = 1, g_4 = 0)|^2} \\
O(\Re(g_2)/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = -1 + i, g_4 = 0)|^2}{|ME(g_1 = 1, g_2 = 1 + i, g_4 = 0)|^2} \\
O(\Im(g_2)/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = 1 - i, g_4 = 0)|^2}{|ME(g_1 = 1, g_2 = 1 + i, g_4 = 0)|^2} \\
O(|g_4|/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = 0, g_4 = 0)|^2}{|ME(g_1 = 0, g_2 = 0, g_4 = 1)|^2} \\
O(\Re(g_4)/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = 0, g_4 = -2 + 2i)|^2}{|ME(g_1 = 1, g_2 = 0, g_4 = 2 + 2i)|^2} \\
O(\Im(g_4)/g_1) &= \ln \frac{|ME(g_1 = 1, g_2 = 0, g_4 = 2 - 2i)|^2}{|ME(g_1 = 1, g_2 = 0, g_4 = 2 + 2i)|^2}
\end{aligned}$$

The observables are chosen to be optimal in measuring different elements of the complex coupling structure. Three observables are defined for each of the two planes. The first is constructed to be sensitive to the coupling magnitude, the second with sensitivity to the real part and the last with sensitivity to the imaginary part.

The magnitude sensitive observables  $O(|g_2|/g_1)$  and  $O(|g_4|/g_1)$  are the same as used in the Run-I analysis;  $O_2(\tilde{\kappa}_{HVV})$  and  $O_2(\tilde{\kappa}_{AVV}, \alpha)$ . The two other observables are chosen such that

the first ratio of matrix elements has difference in the real part of  $g_i$  while the other ratio differs in the imaginary part.

A Boosted Decision Tree ( $BDT_{ZZ}$ ) is moreover trained on production observables to discriminate against backgrounds. Two dimensional templates are finally created where each of the six observables is separately described with their correlation to  $BDT_{ZZ}$ .

The observable distributions for coupling ratios in the  $g_2/g_1$  and  $g_4/g_1$  planes are found with matrix element based reweighting applied to the large ( $g_1 = 1$ ,  $g_2 = 1 + i$ ,  $g_4 = 1 + i$ ) sample. 101 points from -5 to 5 are chosen for each of the dimensions. The distributions are described individually for each of the four final states.

The analysis likelihood is constructed similarly to the description of Section 3.4.4 but is here a function of both the real and imaginary part of the coupling ratios. The likelihood is calculated at the coupling values chosen as targets of the matrix element based reweighting. The points are fairly close, so an adequate description can be found by interpolating likelihood values rather than *pdf* values.

It is finally found to be difficult to describe the correlation between the observable distributions. A simplistic solution is instead adopted by using the test statistic:

$$-2 \ln \lambda(g_\alpha/g_1) = \max \left\{ -2 \ln \frac{\mathcal{L}_{O, \text{Free}}(\theta)}{\mathcal{L}_{O, \text{MLE}}(\hat{\theta})} \mid O = O(|g_\alpha|/g_1), O(\Re(g_\alpha)/g_1), O(\Im(g_\alpha)/g_1) \right\},$$

such that a profiled likelihood ratio is created for each of the three observables and limits are based on the most restrictive.

## Closure and results

The model is applied to various BSM samples to demonstrate its ability to accurately locate minima in the complex  $g_2/g_1$  and  $g_4/g_1$  planes. Figure 3.35 show the 2.23 and 5.99  $-2 \ln \lambda$  contours corresponding to regions excluded at 68% and 95% CL. The contours are shown for each of the three observables. The coupling composition is accurately found in each of the four examples. The three observables are seen to have different sensitivities as expected from their construction.

The analysis is finally applied on Standard Model Asimov samples to get the  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$  expectations. The obtained contours can be found in Figure 3.36.

The interval studied in the Run-I analysis corresponds to taking the likelihood slice along the x-axis of the figures. The expected intervals are listed in Table 3.17. The limits set on the real part of the coupling are presented in two ways. The first constructs the likelihood curve for the real component while profiling the imaginary. In the presence of complex coupling ratios this would be the most accurate way to describe limits. If complex coupling ratios however are disregarded, it no longer make sense to create the profile. The likelihood curve is instead in the second approach read directly along the x-axis.

Examining the structure of the  $g_4/g_1$  fits it can be realised that the two values are the same. The  $300 \text{ fb}^{-1}$  exclusion potential is created by the modulus sensitive observable and does not have any phase structure. In other words, the landscape is circular around the Standard Model

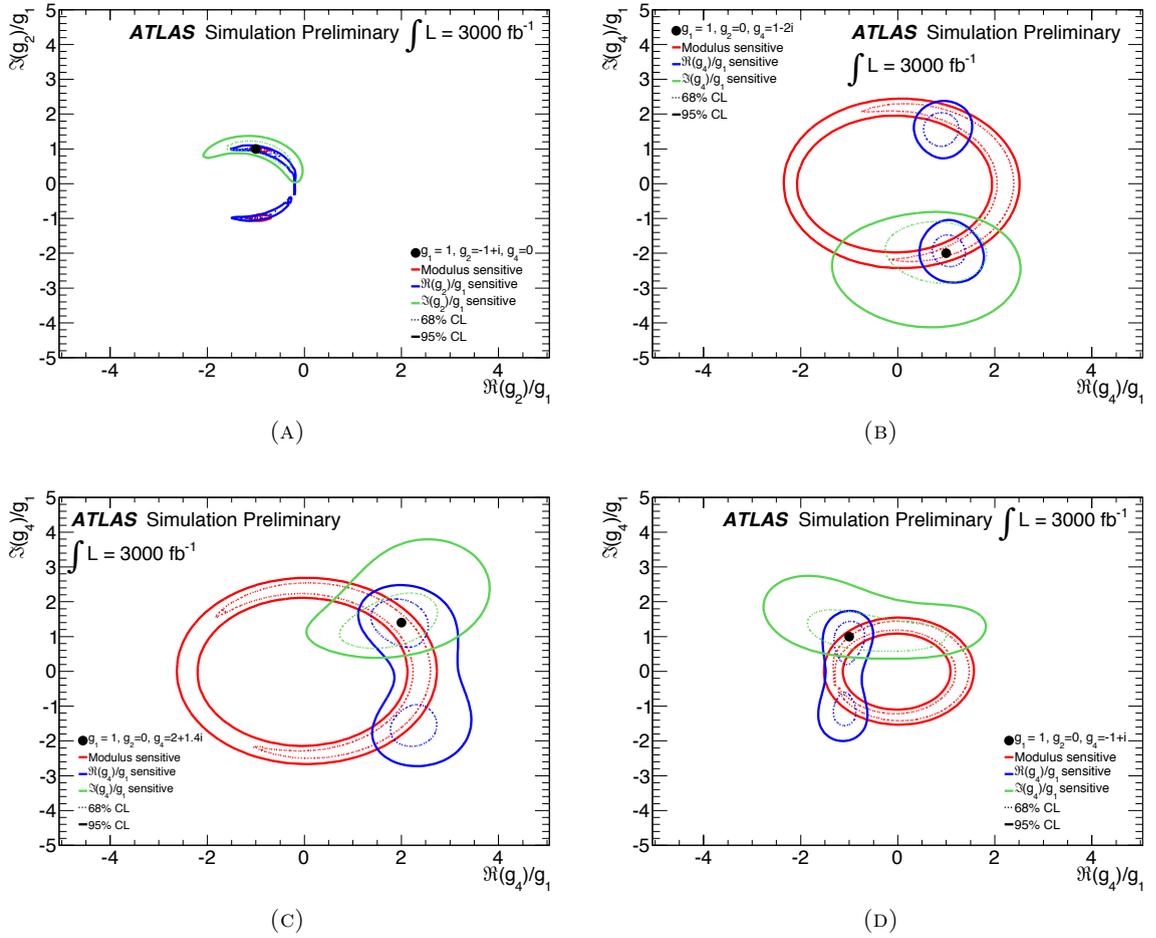


FIGURE 3.35: High luminosity closure tests. Likelihood contours for Asimov samples with coupling compositions  $(g_1 = 1, g_2 = -1 + i, g_4 = 0)$  (A),  $(g_1 = 1, g_2 = 0, g_4 = 1 - 2i)$  (B),  $(g_1 = 1, g_2 = 0, g_4 = 2 + 1.4i)$  (C) and  $(g_1 = 1, g_2 = 0, g_4 = -1 + i)$  (D)[4].

at zero and the real axis as a result contain the minimum as function of real coupling ratio. The sensitivity of the  $3000 \text{ fb}^{-1}$  scan is derived from the real-sensitive observable. The limits are seen to be close to constant as function of complex ratio.

Comparing the results to the expectations of the Run-I dataset of Table 3.9 reveals a significant increase in sensitivity. The analysis is to a very high degree limited by statistics and will keep gaining sensitivity as integrated luminosity is collected. It is expected by the end of LHC that it will be possible to exclude or measure  $\sim 10\%$  contributions from the CP-odd and 2% to 5% contributions from the CP-even BSM sector. Likewise, for the HL-LHC limits these numbers are expected to reduce to  $\sim 1\%$  (CP-even) and 0.2% to 0.4% (CP-odd) which will effectively move the study of the  $HVV$  vertex structure into the category of precision measurements.

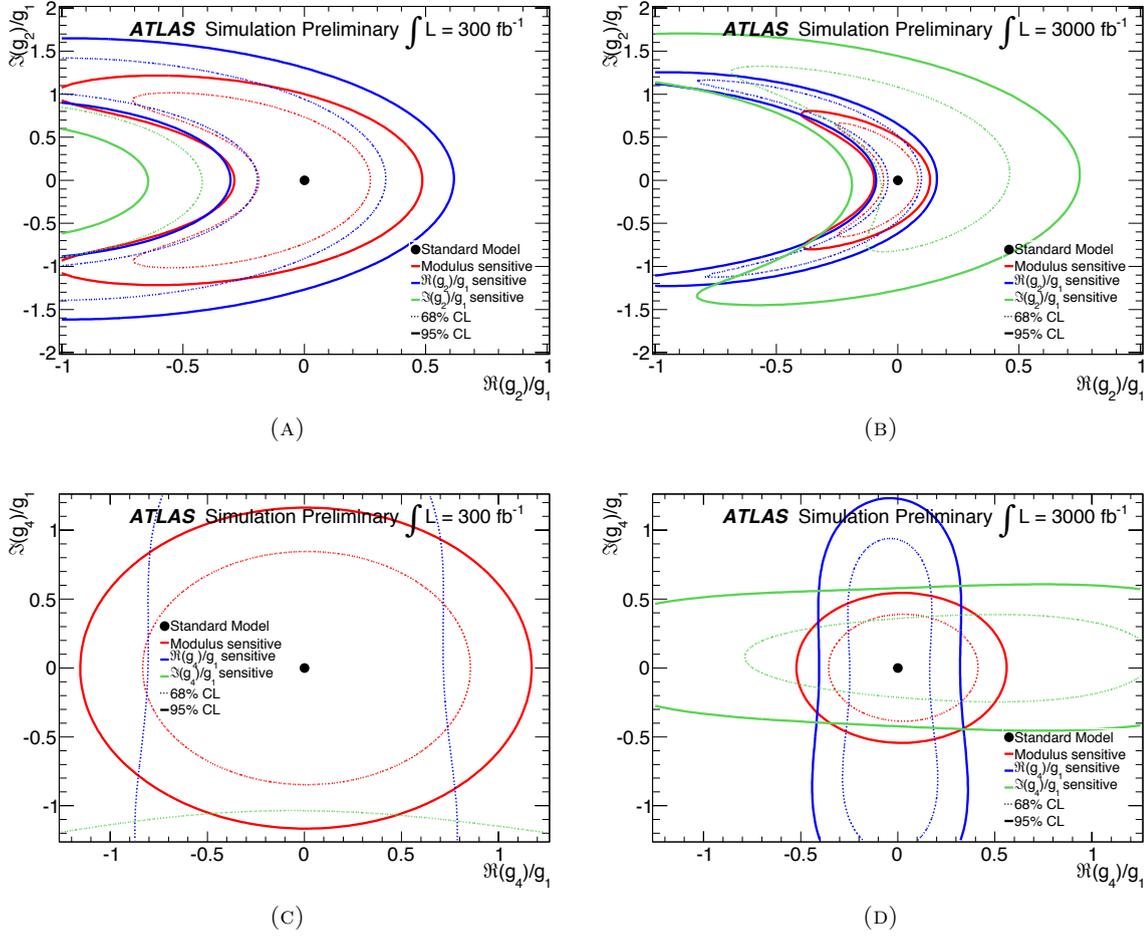


FIGURE 3.36: High Luminosity coupling ratio fits assuming a Standard Model Higgs. (A,B) likelihood contours for the  $g_2/g_1$  plane for 300 and 3000  $\text{fb}^{-1}$ . (C,D) likelihood contours for the  $g_4/g_1$  plane for 300 and 3000  $\text{fb}^{-1}$  [4].

Coupling ratio $H \rightarrow ZZ^* \rightarrow 4\ell$	95% CL Exclusion Regions	
	Expected 300 $\text{fb}^{-1}$	Expected 3000 $\text{fb}^{-1}$
$\frac{\kappa_{HVV}}{\kappa_{SM}}$ (complex profile)	$(-\infty, -0.88] \cup [0.38, \infty)$	$(-\infty, -0.33] \cup [0.11, \infty)$
$f_{g2} \cdot \cos \phi_{g2}$ (complex profile)	$[-1, -0.21] \cup [0.048, 1]$	$[-1, -0.037] \cup [0.0042, 1]$
$\frac{\kappa_{HVV}}{\kappa_{SM}}$ (real-only)	$(-\infty, -0.25] \cup [0.38, \infty)$	$(-\infty, -0.075] \cup [0.11, \infty)$
$f_{g2} \cdot \cos \phi_{g2}$ (real-only)	$[-1, -0.021] \cup [0.048, 1]$	$[-1, -0.002] \cup [0.0042, 1]$
$\frac{\kappa_{AVV}}{\kappa_{SM}} \tan \alpha$	$(-\infty, -1.01] \cup [1.01, \infty)$	$(-\infty, -0.34] \cup [0.26, \infty)$
$f_{g4} \cdot \cos \phi_{g4}$	$[-1, -0.13] \cup [0.13, 1]$	$[-1, -0.016] \cup [0.0096, 1]$

TABLE 3.17: Expected exclusion regions at 300  $\text{fb}^{-1}$  and 3000  $\text{fb}^{-1}$ . The values are both derived from profiling the complex contribution and assuming real-only couplings. The two are the same in the  $g_4/g_1$  plane. The profiled values correspond to the  $\Re$  row of [4, Tab. 6]. Note that the effective cross section fractions  $f_{g2}, f_{g4}$  differ from the paper, due to updated predictions of  $\sigma_{HVV}/\sigma_{SM}$  and  $\sigma_{AVV}/\sigma_{SM}$ . Values used here are as in Section 1.2.4.

### 3.7.2 Angular asymmetries as a Probe of the $HZZ$ Vertex

A separate study was performed which focused on the prospects of fitting the  $HZZ$  tensor structure using angular asymmetries[6]. The parameterisation of mixed models is similar to the effective field theory approach described in Section 1.2.4. The analysis focuses on CP-violating contributions to SM and therefore only considers  $\kappa_{SM}$  and  $\kappa_{AZZ}$  couplings.

CP-mixing is parameterised in terms of the mixing angle  $\alpha$ . As such, it is necessary to make assumptions on the relationship between the two couplings.  $\kappa_{SM}$  is fixed to 1 in order to reproduce SM when  $\cos \alpha = 1$ .  $\kappa_{AZZ}$  is chosen to have the value that results in equal cross section to the Standard Model,  $\kappa_{AZZ} = 28.6$  or equivalently  $\tilde{\kappa}_{AZZ} = 1.76$ <sup>12</sup>.

Monte Carlo is generated with the MADGRAPH5\_aMC@NLO generator[49] and interfaced with a CMS/ATLAS like detector simulation generated with PGS[99]. Expected event yields are similar to [4]:  $N_S = 1.32L$ ,  $N_B = 0.71L$  where  $L$  is the integrated luminosity.

So far, either Matrix Element Observables or the direct use of the final state observables have been described. The study here examines a different set of discriminants proposed in [100]. The angular distributions are defined to be asymmetric in the presence of BSM contributions to the  $HZZ$  vertex. Six observables are constructed from the final state kinematics. Here, only the  $O_4$  observable will be mentioned since it turns out (together with  $O_5$ ) to be the strongest probe of CP-mixing.

In the following, the notation  $\vec{p}_{iV}$  is used to describe the lepton three momenta. They are enumerated such that 1 and 2 come from the  $Z$  boson with the highest mass and 1 is the particle while 2 is the anti-particle. Equivalently, 3 is used to symbolise the particle and 4 the anti-particle. The  $H, Z$  index denotes whether the momentum is measured in the  $Z$  or  $H$  frame of reference. The observable  $O_4$  is defined as:

$$O_4 = \frac{[(\vec{p}_{3H} \times \vec{p}_{4H}) \cdot \vec{p}_{1H}][(\vec{p}_{3H} \times \vec{p}_{4H}) \cdot (\vec{p}_{1H} \times \vec{p}_{2H})]}{|\vec{p}_{3H} + \vec{p}_{4H}|^2 |\vec{p}_{1H} + \vec{p}_{2H}| |\vec{p}_{3Z} - \vec{p}_{4Z}|^2 |\vec{p}_{1Z} - \vec{p}_{2Z}|^2 / 16},$$

The distribution of  $O_4$  for a CP-even and CP-violating boson can be found in Figure 3.37a.

The presence of CP-mixing is measured in two complementary approaches. First the CP-mixing is measured in terms of the observable asymmetry, with corresponding significance:

$$A = \frac{N(O > 0) - N(O < 0)}{N(O > 0) + N(O < 0)} \quad S = A \frac{N_S}{\sqrt{N}},$$

where  $N_S$  and  $N$  refer to the expected number of signal events and total number of events. The observables are conveniently defined such that the  $ZZ^*$ -continuum background is more or less symmetric as in Figure 3.37a.

The significance of the asymmetry thus tells how compatible an observation is with the Standard Model. Conversely, a lack of observed symmetry can be interpreted as excluding possible mixing angles. The excluded mixing angles assuming a SM Higgs for  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$  are listed in the 'Asymmetry' row of Table 3.18.

Next, the method of creating *pdfs* using Monte Carlo templates, constructing a likelihood

<sup>12</sup> The paper also studies different coupling scenarios between  $\tilde{\kappa}_{AZZ}/1.76 = 0.6$  to  $1.4$

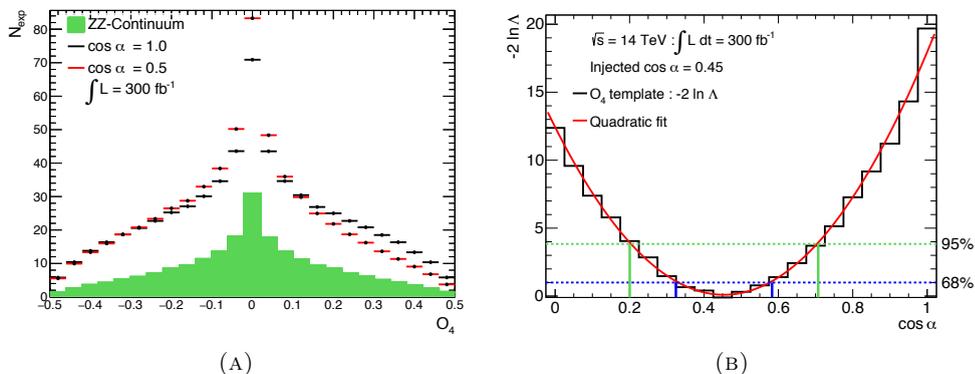


FIGURE 3.37: (A) Expected  $O_4$  distribution for the  $ZZ^*$  continuum, the Standard Model and a CP-violating  $\cos \alpha = 0.5$  model. (B) Likelihood exclusion based on the  $O_4$  distribution assuming a mixed model of  $\cos \alpha = 0.45$ . Both figures assume an integrated luminosity of  $300 \text{ fb}^{-1}$ [6].

function and using this to derive limits is tried. The approach follows the same structure as has been described in the previous sections.

An example of an expected likelihood contour assuming  $\cos \alpha = 0.45$  and an integrated luminosity of  $300 \text{ fb}^{-1}$  is shown in Figure 3.37b. Here it is demonstrated that the  $O_4$  observable would limit  $\cos \alpha$  to be between 0.2 and 0.7 if the true value is 0.45 (i.e. strong CP-violation). The procedure is repeated assuming mixing angles between  $\cos \alpha = 0$  and 1. Figure 3.38 show which regions will be excluded based on the  $O_4$  distribution as a function of the true mixing angle. The expected limits assuming a Standard Model Higgs ( $\cos \alpha = 1$ ) are listed in the ' $O_4$  Likelihood' row of Table 3.18.

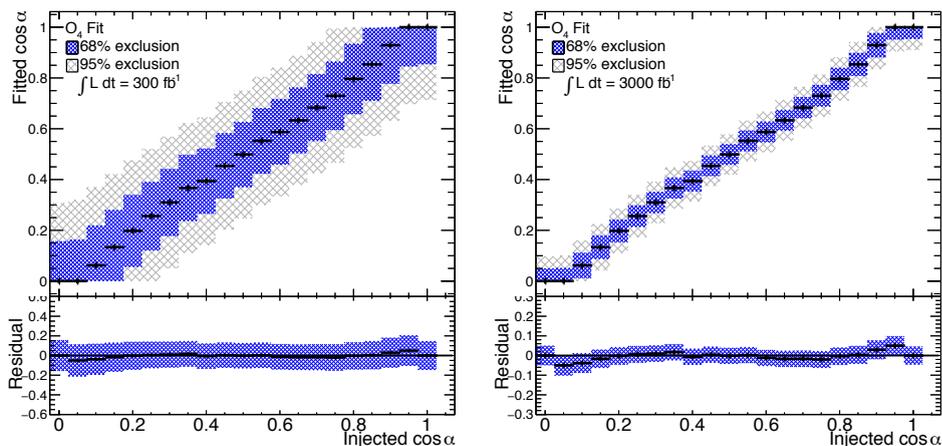


FIGURE 3.38: Expected exclusion regions as function of  $\cos \alpha$  for  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$ . The assumed model is shown on the  $x$  axis with corresponding limits on the  $y$  axis. The white areas are excluded at 95% confidence level[6].

The limits obtained in the two approaches are seen to be fairly similar. The 'Asymmetry' approach is not able to distinguish the SM Higgs from a completely CP-odd model, which is

reflected in the lack of sensitivity to  $\cos \alpha \simeq 0$ . The results shown here are slightly weaker than the matrix element observable approach presented in the previous section. The analysis did however only use one observable and combining several would potentially lead to stronger limits. A more primitive detector simulation is also used so comparison between the two methods should be done with caution.

The approach does however reflect the interesting idea that the CP nature of the Higgs boson can be probed by measuring asymmetries rather than through complex model building. As such, its use could be valuable to study in more detail, by using more realistic simulation, by combining several observables and by obtaining better understanding of how the asymmetry should optimally be used as a test statistic.

$H \rightarrow ZZ^* \rightarrow 4\ell$	95% CL Exclusion Regions			
	Expected 300 fb <sup>-1</sup>		Expected 3000 fb <sup>-1</sup>	
Approach	$\cos \alpha$	$f_{g4}$	$\cos \alpha$	$f_{g4}$
$O_4$ Asymmetry	[0.34,0.79]	0.21	[0.089,0.97]	0.028
$O_4$ Likelihood	[0,0.71]	0.30	[0, 0.91]	0.084

TABLE 3.18: Expected high luminosity limits on the CP-mixing angle using angular asymmetries. Only the limit of  $f_{g4}$  corresponding to  $\cos \alpha \simeq 1$  is given for the Asymmetry approach since the high BSM fraction has already been excluded by current observations.



## Chapter 4

# Conclusions

## 4.1 Summary

The introduction of the Higgs particle to the Standard Model of particle physics reveals a number of appealing features. It enables a description of a gauge invariant theory of massive particles and remedies the divergent behaviour of massive vector boson scattering.

The discovery of the Higgs boson marked the culmination of decades of dedicated research at the massive collaborative efforts at the LEP, Tevatron and LHC experiments.

The resonance discovery however left many questions unanswered. In an effort to test the predictions offered by the Standard Model, a series of studies have been conducted here on its spin and CP-nature. The common outline of those is to remove the Higgs sector of the Standard Model Lagrangian and substitute several alternatives to determine which is favoured by data.

The substitutes are divided into alternative spin models and alternative spin-0 models. Spin-2 alternatives are described as graviton-like particles in an effective field theory approach. The QCD couplings of the spin-2 models is not determined in theory and several assumptions are tried.

The spin-1 hypothesis is also tested in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channel but has little theoretical motivation since it is disfavoured by the resonance decay to  $\gamma\gamma$  due to the Landau-Yang theorem[51][52]. The agreement between the production cross section of the boson and the Standard Model production of Higgs through gluon fusion also disfavours this hypothesis.

Discrimination between hypotheses is obtained in a multivariate approach using boosted decision trees. The BDT approach allows for a compact description of the eight-dimensional final state of  $H \rightarrow ZZ^* \rightarrow 4\ell$  that takes into account relevant parameter correlations. A BDT is used to separate alternative signal models and another is used to separate a Standard Model Higgs from the  $ZZ^*$ -continuum background. It is found, with the exception of the negative parity spin-2 and a spin-2 boson that does not couple to the quark sector, that all alternative spin hypotheses excluded at the 95% CL. Notably, the spin-2 universal couplings model is assigned a  $CL_S$  value of 0.0009.

The spin-0 sector is described in an effective field theory approach with three component structure. Assuming only a single sector present at a time, the three components correspond to the Standard Model, a CP-even BSM boson and CP-odd boson. Using the same multivariate approach as for the spin hypothesis tests the Standard Model if found to be favoured by data and the alternative CP-even ( $J^P = 0_h^+$ ) and CP-odd ( $J^P = 0^-$ ) models are assigned  $CL_S$  values of 0.10 (CP-even) and 0.054 (CP-odd).

The analysis is extended by allowing the simultaneous presence of the SM coupling ( $\kappa_{SM}$ ) and one BSM coupling ( $\kappa_{HVV}$ ) or ( $\kappa_{AVV}$ ) where  $\kappa_{AVV}$  is accompanied by the CP-mixing angle  $\alpha$ . The  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis uses a Matrix Element Observable approach that is constructed to take into account the full decay kinematics of the process. The matrix elements observables are extended by a BDT background discriminant. A combination of matrix element based reweighting, kernel density estimation and *pdf*-interpolation is used to acquire a *pdf* that is continuous in the examined coupling ratios. Extensive testing against independent Monte Carlo simulation is used to validate the approach.

The  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel is finally combined with the  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  analysis. The channels are found to give similar results in agreement with custodial symmetry. By combining their information, the effective BSM cross section contribution to the resonance production is found to be limited to the ranges:

$$f_{g2} \cdot \cos \phi_{g2} \in [-0.16, 0.12] \quad f_{g4} \cdot \cos \phi_{g4} \in [-0.41, 0.090]$$

at 95% confidence level. Data is thus in agreement with Standard Model behaviour.

The spin-0 analysis concludes with a final extension, where all three components of the EFT Lagrangian are simultaneously studied. A factorised model of the Matrix Element Observables is developed and the simultaneous fit reveals data to be in agreement with the Standard Model. A summary of the results is presented in Figure 4.1.

A prospect study of the use of Matrix Element Observables in the high luminosity limit is conducted. It is demonstrated that for an integrated luminosity of  $300 \text{ fb}^{-1}$  it will be possible to probe 10% contributions from the CP-odd sector to the combined cross section and 2% to 5% from the CP-even BSM sector. At  $3000 \text{ fb}^{-1}$  these numbers are demonstrated to reduce to 1% (CP-even) and 0.2% to 0.4% (CP-odd).

A separate analysis of the use of angular observables as a probe for the  $HVV$  vertex is also performed. The analysis concludes that it will be possible to accurately measure CP-mixing in the Higgs sector by studying observable asymmetries. The results are demonstrated to be of a similar precision as those obtained from a full likelihood model.

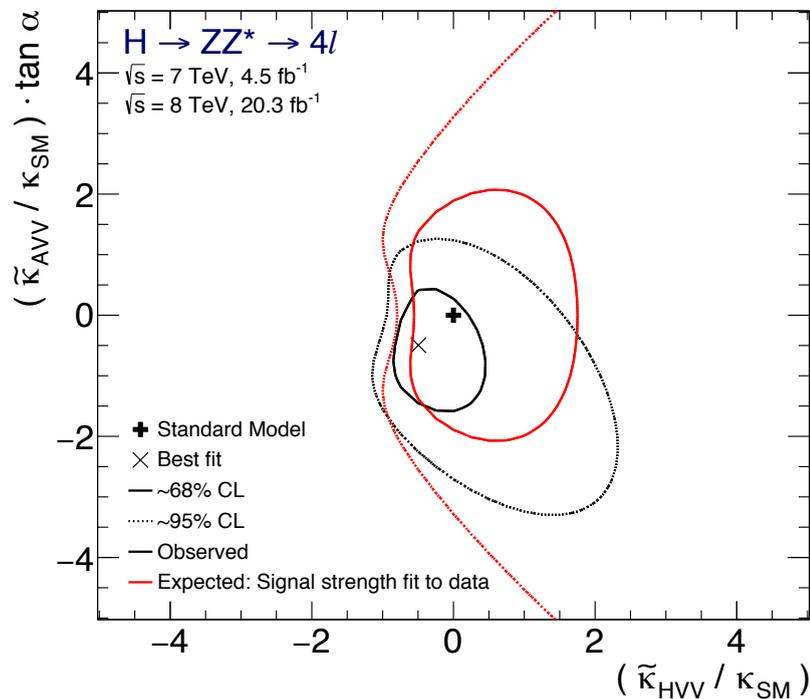


FIGURE 4.1: Simultaneous fit of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  performed with the Matrix Element Observable approach. The figure shows the expected and observed 68% and 95% confidence contours obtained for the dataset collected by ATLAS during the LHC Run-I. The fit takes into account the complete structure of the spin-0 EFT Lagrangian and serves as the final result of this thesis. Observations are found to be in agreement with Standard Model behaviour.

## 4.2 Concluding Remarks

The Standard Model has proved to accurately describe the observed data. It is concluded the observed resonance must be spin-0 and not completely CP-odd based on the alternatives that have been tested in both ATLAS and CMS[5][48].

The results of the tensor structure analysis are however limited by statistics. The analysis so far has only examined the heavy vector boson decays of a Higgs-like boson produced with gluon-fusion. With more data it will be possible to include the less probable production mechanisms and more final states. One of the first extensions could be the inclusion of vector boson fusion (VBF). The VBF production mode is accompanied by the emission of two hard jets in the forward direction. The kinematic behaviour of the two jets can be demonstrated to be affected by the level of CP-mixing exhibited by the Higgs particle. By having the heavy vector bosons produce the Higgs, the CP-nature of the  $HVV$  interaction can be examined in all different decay channels.

The analysis approach presented here will likely have to be modified if more complexity is added to the model. A factorisation scheme had to be used when examining  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AV V}}{\kappa_{SM}} \tan \alpha$  at the same time. As more features are added it will be necessary to include more observables and soon it will be simpler to utilise the final state kinematics directly.

On a personal note, I think this is the promising way to progress in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. The leptonic final states are very pure, meaning it feasible to construct a precise effective detector model. If the transfer functions were to be based directly on the lepton four-momenta instead of the derived kinematic distributions only a handful of parameters would be needed ( $p_{T,\ell}, \eta_\ell, \phi_\ell, \dots$ ). It would surely be a challenge to construct such a model, especially if the jets of VBF were to be described, but if successful the analysis would be protected against the exponential increase in complexity that has lured in every corner of the studies so far.

If not the main, then one of the largest challenges of the analysis that have been presented here proved to be the optimisation between how the most powerful conclusions are obtained and what is feasible to do. The current technological limitations of how much information can be used to describe the coupling sensitive *pdf* for instance limited the analysis from using a fine binning of the matrix element observables. Moreover, the reweighting procedure is computationally heavy and realistically can only be performed with distributed computing. When many iterations are necessary before the final results can be produced, this proves a limiting factor timewise. Further development of the analysis tools is of high priority when updated results are to be made with the LHC Run-II data.

Another interesting development of the tensor structure analysis has recently been developed. It demonstrates that it is not necessary to perform reweighting as described in the text. Instead of creating a template for each of the examined coupling configurations it is possible to mix SM and BSM templates directly, if their interference is accurately described.

The Higgs analysis of the LHC Run-I all but closes the question of whether the resonance can be described with a pure BSM spin/CP model. The tensor structure analysis has only seen its very first results and will continue to serve as an interesting probe of the Higgs sector in the years to come.

# Appendix

## A Statistical Considerations

This section will aim to explain some of the 'whys' and 'hows' of the more technical aspects of the analysis. The setting will be held very general and should not solely be considered applicable to the  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis. It has been chosen to include this in a separate section to not break the flow of the actual analysis text, such that the focus there can be held as much as possible on physics. Moreover, many of the aspects will be used repeatedly in different sections.

For interested readers, the following text should be self-contained. Otherwise, it can be used as a reference when reading the analysis sections.

### A.1 Construction and Interpretation of Statistical Models

One of the most fundamental features of any analysis is the probability density function (*pdf*). The *pdf* describes what the probability for an observation of some observable is: What is e.g. the probability that a Higgs boson produced by LHC collisions has a transverse momentum greater than 100 GeV?

In the context of physics measurements this is often not the type of question you want answered however. It is likely more interesting to know what the probability of the observed behaviour is given some model<sup>1</sup>.

The *pdf* is changed to be model dependent,  $pdf(x) \rightarrow pdf(x|\alpha)$ . In the context of discovery, the probability that the background only hypothesis will produce data is e.g. chosen to be  $\sim 10^{-6}$ .

The interesting part is thus to study how well data is described by *pdf*s for different values of  $\alpha$ . In order to examine this, it is natural to introduce a new function of the parameter ( $\alpha$ ) given observation ( $x$ ). The function is called the likelihood function.

For particle physics in particular, it is most likely the case that a measurement is identically repeated independently a number of times (Independent and Identically Distributed, IID). The *pdf* for each of these is identical and the likelihood thereby takes the form:

$$L(\alpha|x) = pdf(x|\alpha) \xrightarrow{IID} L(\alpha|x) = \prod_i pdf(x_i|\alpha)$$

The favoured  $\alpha$  is obtained by maximising the likelihood. Typically, if only one model is favoured, the maximum is found when  $\frac{\partial L}{\partial \alpha} = 0$  and  $\frac{\partial^2 L}{\partial \alpha^2} < 0$ . Moreover, as will be the theme of the following text, since the exclusion of models is based on how much the likelihood changes when moving away from the optimal  $\alpha$ , the discriminative power of the analysis is contained in the curvature of the likelihood around the maximum:  $\frac{\partial^2 L}{\partial L^2}$ .

This part will present a general discussion of the statistical approaches used in the analyses. It is the aim to argue that the approaches follows naturally from the Neyman-Pearson (NP)

---

<sup>1</sup>The Bayesian probability of a model given data will not be considered here.

lemma[101]:

Imagine two separate hypotheses  $H_0$ ,  $H_1$ , where  $H_0$  is the true model. In practice this is of course not know.

The goal is now to find a measure that ensures the maximum purity for any chosen efficiency. In other words, the analyser chooses that it is acceptable to have rejection of the true  $H_0$  with probability  $\eta$  (Type-II). The measure should ensure that the false  $H_1$  is accepted as rarely as possible (Type-I).

The NP-lemma states that the most powerful test of this type is given by the likelihood ratio:

$$\Lambda = \frac{L(H_0|x)}{L(H_1|x)} \leq \beta, \quad P(\Lambda \leq \beta|H_0) = \eta \quad (\text{A.1})$$

I.e. choosing an  $\eta$  for the analysis and requiring that  $\Lambda < \beta$  will ensure the highest possible purity.

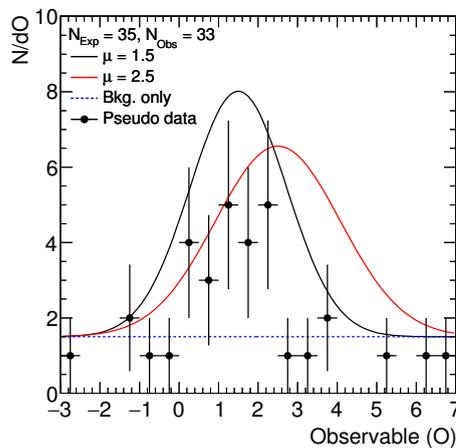


FIGURE A.2: Example model used in text. The model consist of a constant background and a signal with variable mean ( $\mu$ ). The signal is a Gaussian where the width scales as  $\sqrt{\mu}$ . Data is divided in 20 bins of 0.5.

In order to make the following description more concrete, a toy model is introduced. It consists of a Gaussian signal with mean  $\mu$  and width  $\sqrt{\mu}$  and a constant background. The *pdf* normalised to the expected signal and background yield is given by:

$$pdf(x) = \frac{N_s}{\sqrt{2\mu}} e^{-\frac{(x-\mu)^2}{\mu}} + \frac{N_{bkg}}{10}, \quad x \in [-3.0, 7.0]$$

The signal and background normalisations are chosen to be 20 and 15 events, respectively. The distributions of two different signal models together with an example of how data might look like can be found in Figure A.2.

The choice of this particular *pdf* is completely arbitrary. It was simply chosen because it is possible to construct the different types of tests that will be described here from it. The width of the signal distribution is chosen to change with the distribution mean to visualise how sensitivities will change as an effect of this. As such, the models resemble a simple version of something that could be used to measure a resonance mass in particle physics.

## Simple hypothesis tests

In the most basic scenario, only two models are considered,  $H_0$  and  $H_1$ <sup>2</sup>.

The likelihood ratio can be used to determine the favoured of the two hypotheses. In order to do this, it is necessary to know the test statistic distribution of the two models. Although it is sometimes possible to predict these distributions directly from the form of the likelihood it will often not be the case. Instead, the distributions can be derived from more empirical principles. The analyses of this thesis that fall into the 'simple hypothesis tests' category are for instance based on Monte Carlo pseudo experiments.

The general idea is to first assume the hypothesis  $H_0$  to be true. A Monte Carlo dataset is then constructed from its *pdf* and normalisations. The number of MC events simulated from the *pdf* can in relevant scenarios be drawn from a Poissonian distribution to take into account the expected statistical fluctuations.

The likelihood ratio test statistic ( $\Lambda$ ) is calculated and the procedure is repeated a large number of times. Finally, the data value is calculated and the probability for  $H_0$  to exhibit the data behaviour ( $p_0$ ) follows from the tail integral of the  $H_0$  test statistics distribution. The process can be repeated for  $H_1$  in order to find its  $p_0$ -value.

Using the toy *pdf* from above, the two models  $\mu = 1.5$  and  $\mu = 2.5$  are chosen as  $H_0$  and  $H_1$  respectively. The 'Observed' data is constructed from  $\mu = 1.5$  as well.

Note that it is customary to use  $-2\ln\Lambda$  instead of  $\Lambda$ . For now it can be seen as a way to avoid dealing with computation with really small numbers. The approach is illustrated in Figure A.3a. In order to demonstrate the effectiveness of the likelihood ratio, the test statistic is compared to another choice. Figure A.3b show the same procedure using the fitted  $\mu$  value. Using  $\hat{\mu}$  as test statistics gives weaker results as expected.

In an ideal world, the data test statistic will fall close to the median of the log likelihood ratio distribution of one hypothesis and in the tail of the other. In a blinded analysis, this should however not be assumed.

As an example assume  $H_0$  is the Standard Model Higgs Boson,  $H_1$  is a Super Symmetric parity odd Higgs, but what is actually observed is a spin-2 Graviton-like state. Both  $H_0$  and  $H_1$  will fit the model poorly. If one only consider exclusion of the Super Symmetric  $H_1$  by its  $p_0$  value it can be excluded over the Standard Model, even though this model also fits the data poorly. To avoid this problem before unblinding the data, exclusions of models can be based on  $CL_s$  instead:

$$CL_s(H_1) = \frac{p_0(H_1)}{1 - p_0(H_0)}, \quad (\text{A.2})$$

The  $CL_s$  is constructed such that it will be inconclusive if both models fit poorly and approximate  $p_0(H_1)$  if  $H_0$  is true. The  $CL_s$  approach should be said does not mend the problem if very high separation between the two models is expected and the data falls in between the two hypothesis distributions.

---

<sup>2</sup>'Simple' here refers to the assumption of only two hypotheses.

In most real scenarios the likelihood will not only be a function of the parameter of interest  $\alpha$ . It will also contain parameters that e.g. describe uncertainties in theoretical predictions and lack of knowledge of the experiment behaviour called nuisance parameters. In actual measurements these will mostly consist of systematics uncertainties.

The standard approach is to modify the likelihood ratio as:

$$\Lambda = \frac{L(H_0|x)}{L(H_1|x)} \rightarrow \frac{L(H_0, \hat{\theta}_{H_0}|x)}{L(H_1, \hat{\theta}_{H_1}|x)}$$

such that the likelihood is maximised with respect to all nuisance parameters for either hypothesis separately. This type of test statistic is commonly referred to as the 'ratio of profiled likelihoods'.

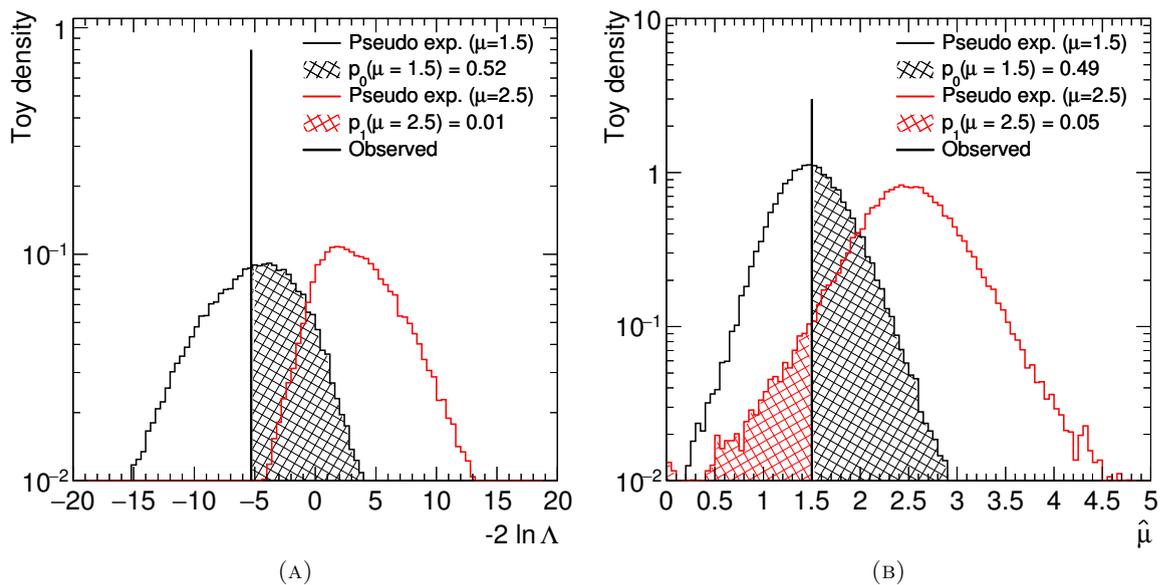


FIGURE A.3: Hypothesis test result using the likelihood ratio test statistics (A) and the maximum likelihood estimator  $\hat{\mu}$  (B). The models that are used are described in the text.

## Composite model tests

In a more general setting it will not be possible to assume that the observed data is drawn from one of two hypotheses. A composite model is needed instead of the simple hypothesis models.

As an example, it is more interesting to figure out what the best fitting Higgs mass and its corresponding uncertainty is rather than figuring out which of the 124.5 GeV and 126 GeV mass models best describe data.

The composite model is used to establish excluded regions of a continuum of models in some parameter space  $\Omega$ :

$$L(\alpha|x), \quad \alpha = H_0, H_1 \quad \rightarrow \quad L(\alpha|x), \quad \alpha \in \Omega$$

At first glance one could imagine that the same likelihood ratio procedure as for the simple hypothesis tests could be carried out by creating pseudo experiments and testing values of  $\alpha_i$  against each other in some finely spaced grid of  $\Omega$ . Aside from the possible computational problems of this approach it is not clear how the results would be interpreted. The goal is to design an optimal approach to establish which regions that would lead to the observed data with probability less than  $\eta$ .

If the NP-lemma is for now ignored, one possible solution could be to choose another test statistic that does not compare models, such that its distribution can be evaluated locally. Like above the fitted  $\hat{\mu}$  is an example of this approach.

Although possible here, it will be assumed that the distribution of means for the model can not be calculated a-priori but will have to be derived from pseudo experiments. This will be the case for most analyses.

The test statistic distribution can as such be calculated for each model and compared to data. The approach is illustrated in Figure A.4a, where the Gaussian mean is used to derive limits on the toy model from before. A pseudo-data  $\mu$  (generated at  $\mu = 1.5$ ) is fitted and its value compared to the distributions of the different models. It follows directly which models would result in an observed mean with a probability less than any chosen threshold.

When using a simple test statistic like the signal mean it is intuitive to define which models are close and far away from the observed data. An 'ordering principle' has been chosen.

The approach of choosing a "local" test statistic and deriving coverage belts, as was exemplified with the signal mean above, is usually referred to as the Neyman Construction[101].

Returning again to the likelihood. Instead of considering a fixed number of hypotheses, the likelihood is assumed to be a continuous function of the parameter of interest  $\alpha$ .

If the likelihood is calculated for each  $\alpha$  the best fitting model is found at the maximum. In the case the likelihood also contains nuisance parameters, it can instead be maximised w.r.t. these for each  $\alpha$ .

Considering that the discussion started out with an attempt to generalise the simple hypothesis likelihood ratio to a composite model, it seems natural to compare the non-optimal  $\alpha$  models to the best fitting:

$$-2 \ln \lambda(\alpha) = -2 \ln \frac{L(\alpha, \hat{\theta})}{L(\hat{\alpha}, \hat{\theta})} \quad (\text{A.3})$$

The test statistic is usually referred to as the profiled likelihood ratio. The use of  $-2 \ln$  will become apparent shortly.

The profiled likelihood ratio does not compare hypotheses in the same sense as the ratio of profiled likelihoods does, since the best fitting model does not need to be assumed. Given the NP-lemma, it is well motivated to use the profiled likelihood ratio as an ordering principle in the Neyman-Construction. This then gives an optimal method for determining excluded regions based on pseudo experiments[102]<sup>3</sup>.

The likelihood approach is applied to the toy example. Pseudo experiments are generated for a series of  $\alpha$  models to establish the test statistic distribution for each. The likelihood ratio is

<sup>3</sup>In this section the ordering is introduced as a way to obtain the most restrictive exclusion regions. This is relevant when performing measurements on established signals as will be the case for the analysis. The paper introduces the ordering as a way to get correct coverage when observing small or negative signals.

calculated for the “observed” data as a function of  $\alpha$  by which the excluded regions follows. The outcome is shown next to the limits set by sampling the mean, see Figure A.4. Similarly to the hypothesis tests, where a better exclusion than the alternative  $\mu$  method was obtained, a more restrictive interval is found by the likelihood ratio ordering than the  $\hat{\mu}$  ordering.

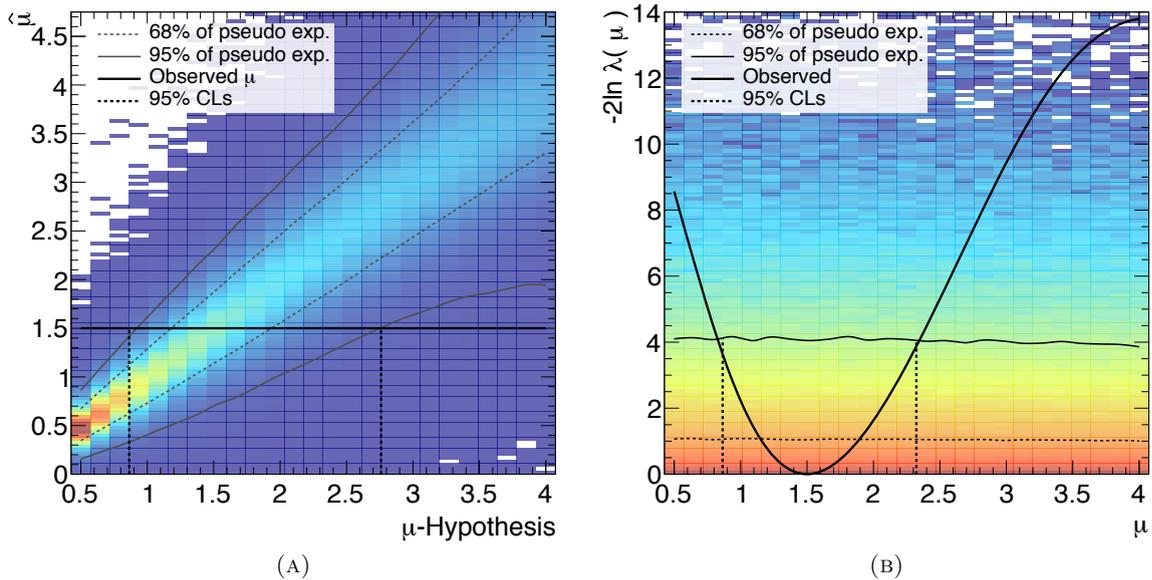


FIGURE A.4: Examples of coverage belts based on different test statistics. (A):  $\hat{\mu}$ , (B): The profiled likelihood ratio. In both figures a large number of pseudo experiments have been generated for each point on the x-axis. The test statistic value of each toy is seen on the y-axis. In the  $\mu$  case the observed  $\hat{\mu}$  is indicated by a full horizontal line. The intervals which produce such a  $\mu$  in less than 5% of cases are indicated by the dashed vertical lines. For the profiled likelihood ratio, the observed test statistic varies as function of tested  $\mu$ . The excluded regions are again found when the pseudo experiments produce the observed test statistic in less than 5% of cases.

## Asymptotic behaviour

One important distinction between the  $\hat{\mu}$  test statistic and likelihood ratio in Figure A.4, is seemingly that the profiled likelihood distributions do not change between models.

In fact, it can be proven that under regular conditions and given enough statistics this will always be the case for composite models[96]. According to Wilk’s theorem, the test statistic distribution is given by the  $\chi^2_{NDF}$  pdf for a parameter of interest with dimensionality of  $NDF$ :

$$f(-2 \ln \lambda(\alpha)) = \chi^2_{\text{Dim}(\alpha)}(-2 \ln \lambda(\alpha)) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad (\text{A.4})$$

Figure A.5a shows the test statistic values of the toy model generated at  $\mu = 2.5$ , corresponding to a vertical slice ( $\mu = 2.5$ ) of the sampled distribution.

A very good agreement between the distribution of pseudo experiments and their asymptotic approximation is found. The probability of the  $\mu = 2.5$  model would produce the observed results is given by the tail integral, starting at the test statistic value. This means that it is

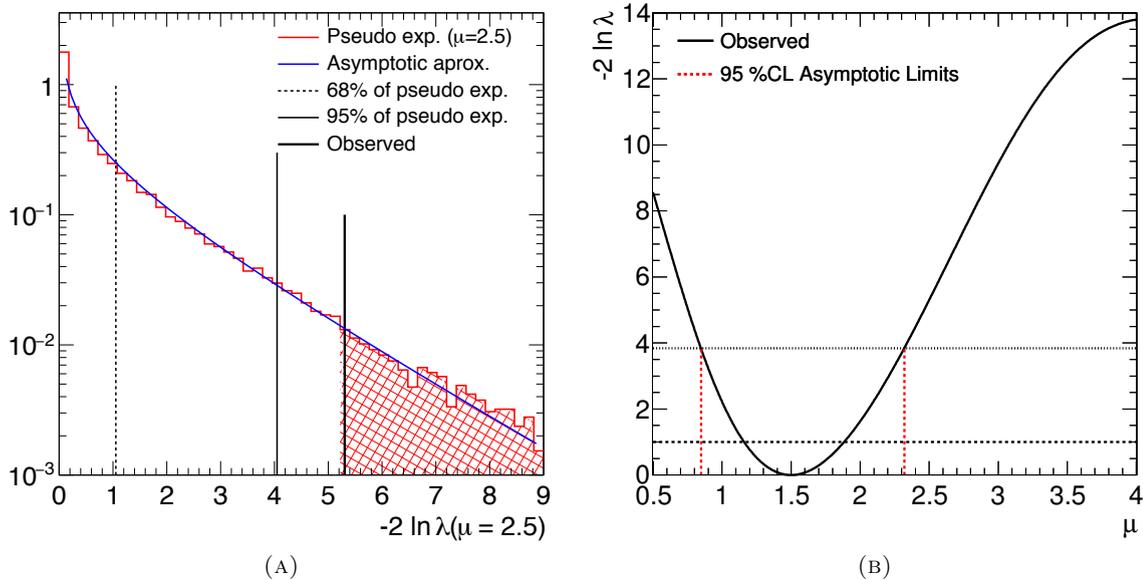


FIGURE A.5: (A): Profiled likelihood test statistic distribution for the toy model. Pseudo experiments are generated at  $\mu = 2.5$ . The test statistic is evaluated at the same value as the pseudo experiments are generated at. The distribution is compared to its asymptotic approximation. The probability that the  $\mu = 2.5$  model would produce the observed results is indicated by the tail integral. (B): Limits based on the profiled likelihood ratio assuming asymptotic behaviour.

possible to determine if a model  $\alpha_0$  is excluded at confidence level  $\eta$  by evaluating if:

$$P(\chi_{\text{Dim}(\alpha)}^2 > -2 \ln \lambda(\alpha_0)) > \eta$$

Some common values are worth mentioning as reference: Intervals with  $-2 \ln \lambda > 3.84$  are excluded at 95% CL for a one dimensional  $\alpha$ . The equivalent 95% CL regions are found above 5.99 for a two dimensional  $\alpha$ .

In summary, combining the knowledge of the NP-lemma, likelihood ordering and Wilks theorem, a recipe emerges for establishing excluded parameter regions:

- Construct the likelihood from the analysis specific *pdf*
- Calculate the profiled likelihood as function of the parameter of interest  $\alpha$  over its allowed space  $\Omega$
- Regions where  $P(\chi_{\text{Dim}(\alpha)}^2 > -2 \ln \lambda(\alpha_0)) > \eta$ , are excluded at confidence level  $\eta$ .

The asymptotic likelihood limits are set as shown in Figure A.5b. A very good agreement is seen when comparing to the results obtained by sampling. The asymptotic behaviour of the test statistic distribution is very appealing since it does not require the generation and fitting of a large number of pseudo experiments.

## A.2 Fast Binned Kernel Density Estimators

This section will shortly discuss a solution to a challenge often occurring in analysis: There does not exist an analytical description of the observable *pdf* or it is difficult to find a fitting function.

This happens to often be the case when dealing with fully simulated Monte Carlo. The underlying physics is described well, but it is close to impossible to fold prediction with the complex behaviour of the detector.

Ideally one could simulate enough Monte Carlo to be able to describe the distributions with template histograms. Monte Carlo statistics are however often a limiting factor such that the templates will suffer from large statistical uncertainties and give less reliable predictions. This is especially the case if multi-dimensional discriminants are used. All other things being equal, the number of bins grows exponentially with dimensionality. This means that the characteristic statistics size needed to populate the template will grow exponentially as well.

A possible solution is to use Kernel Density Estimators (KDE). The general idea is to describe each data point with a 'kernel', such that the sample of events is exchanged with a sample of kernel functions.

Most often a Gaussian distribution is used. One could interpret this as assigning an uncertainty to the observable value and create a *pdf* from the summed Gaussians. In the one-dimensional case, for a dataset  $X$  and kernel function  $\phi_h$  the KDE will read:

$$\hat{f}(x) = \frac{\alpha}{n} \sum_{i=1}^n \phi_h(x - X_i) \quad \rightarrow \quad \hat{f}(x)_{\text{Gauss}} = \frac{1}{hn} \sum_{i=1}^n e^{-\frac{(x-X_i)^2}{h}}, \quad (\text{A.5})$$

where the common Gaussian case has been written explicitly in the left expression. The parameter  $\alpha$  in the first expression is kernel specific and there to ensure the *pdf* is normalised. Moreover, a bandwidth parameter  $h$  is introduced. In the Gaussian case this is the distribution width, describing how much each data point is smeared. In general this will be a kernel specific parameter that is chosen such that the *pdf* will describe data optimally.

Things like the statistics size and rapid fluctuations in the distribution will influence what the optimal choice of  $h$  is. In general the best value is unknown and can be tuned to the analysis in question.

It can be demonstrated that in the large statistics limit of normally distributed data, the mean integrated squared error is minimised when[90]:

$$h = \left(\frac{4}{3}\right)^{\frac{1}{5}} \sigma n^{-\frac{1}{5}} \quad (\text{A.6})$$

And thus often serves as a good starting point for optimisation.

For rapidly changing distributions it is often beneficial to introduce an adaptive bandwidth. A good adaptive bandwidth will be narrow in regions with high statistics and wide in low statistics regions. A customary choice scales the bandwidth with the non-adaptive estimate:

$$h_i = \frac{h}{\sqrt{f_{NA}(X_i)}}$$

The adaptive  $h_i$  is seen to have the desired properties.

The simple case described above generalises to more observable dimensions. In the Gaussian case for instance, the two dimensional expression emerges from the following substitution:  $(x - X_i)^2 \rightarrow (x - X_i)^2 + (y - Y_i)^2$ .

If the distributions are linearly correlated, a better description of the dataset can be found by decorrelating the observable[90].

After the kernels have been calculated, the *pdf* can subsequently be described by a histogram. To estimate the density in each bin of the histogram the distance to each data point needs to be known. For each bin and data point a Gaussian function furthermore needs to be evaluated<sup>4</sup>. In the adaptive case this needs to be done twice, making the procedure potentially slow for large datasets.

A standard KDE approach will thus construct a Gaussian function for each datum. The underlying *pdf* that the data is sampled from is estimated by constructing a histogram with bin content given by the summed kernels.

A faster but less accurate approach is to first fill the dataset into the template histogram and subsequently perform the KDE. The algorithm that will be described here is largely motivated by [103].

Creating the histogram as the first step is equivalent to performing the KDE on a weighted dataset of  $g$  entries with separation equal to the histogram bin width. The KDE expression for a uniform binning reduces to:

$$\hat{f}(x_i) = \frac{1}{n} \sum_{i=1}^n c_i \phi_h(x - x_i) \rightarrow \frac{1}{g} \sum_{i=1}^g c_i \phi_h(|i - j|\delta), \quad (\text{A.7})$$

with histogram bin width  $\delta = x_2 - x_1$ , and  $c_i$  describing the histogram bin content.

The expression can also be generalised to more dimensions as above.

The binned expression has some desirable features compared to the unbinned approach. Before, the Gaussian function had to be evaluated  $\mathcal{O}(N_{\text{Data}} \cdot N_{\text{Bins}})$  times. If the bin width  $\delta$  is constant in the binned approach it only has to be evaluated  $\mathcal{O}(N_{\text{Bins}})$  times<sup>5</sup>.

While the binned approach already seems like the faster choice, further optimisation is still possible. If one defines the “data” and “kernel” vectors ( $C$  and  $K$ ) as arrays containing the histogram bin content and the Gaussian kernel values respectively:

$$\begin{aligned} C &= (c_1, c_2, \dots, c_g) \\ K &= \frac{1}{n} (\phi_h(0), \phi_h(\delta), \dots, \phi_h((g-1)\delta)), \end{aligned}$$

the expression for  $\hat{f}(x_i)$  from above looks very similar to the convolution:  $C * K$ . This means that  $\hat{f}(x_i)$  can be found using the Discrete Fourier Transform (DFT) and the Convolution

<sup>4</sup>If the data-point is several bandwidths away from the bin it will not really influence the result and can be ignored

<sup>5</sup>Intuitively, it seems that it should be  $\mathcal{O}(N_{\text{Bins}}^2)$  corresponding to one evaluation for each function argument  $i$  and sum index  $j$ . For uniform binning the arguments entering the Gaussian function are however solely specified by  $\delta|i - j|$  of which there are  $N_{\text{Bins}}$  of.

Theorem:

$$\hat{f}(x_i) = \text{DFT}^{-1} [ \text{DFT}(C) \cdot \text{DFT}(K) ] (x_i)$$

The convolution theorem applied on the discrete  $C$  and  $K$  vectors is however only applicable if  $C$  is periodic. In real life the distribution that is approximated will likely not have periodic behaviour making it a bad assumption. As a consequence, if the method is blindly applied, the very left of the distribution will affect how the very right part looks like. To remove this type of over-wrapping effects, the vectors can be padded with empty bins: A number of 0s are appended to the beginning and end of  $C$  such that the overwrapping effects only applies to the 'zero-regions'. The convolution can be applied, after which the zero bins are removed to get the final estimate. Some considerations are needed to figure out how this is optimally done.

A large fraction of the terms in the kernel vector  $K$  will most likely be negligible. For the Gaussian case, if only less-than- $5\sigma$  contributions are considered, any entry  $\phi_h(i\delta)$  in  $K$  where  $|i\delta/h| \geq 5$  can be set to zero.

To reduce notation, define  $l$  as the largest integer less than  $5h/\delta$ .

$$l = \min(g - 1, \text{floor}(5h/\delta))$$

In the definition of  $l$ , it is furthermore required that  $l$  should be less than the actual histogram size  $g$ . To remove overwrapping effects, the  $C$  vector then as a minimum need to have a length of  $g + l + 1$ .

An FFT algorithm will be used for the Fourier transformation. The method can be demonstrated to have a complexity of  $\mathcal{O}(n \log n)$  for all powers of two<sup>6</sup>. A choice of padding-size that both has no overwrapping effect but also ensures a fast computation is thus given by the smallest power of two that is greater than the minimally required size of  $C$ :

$$p = 2^{\text{ceil}(\log_2(g+l+1))}$$

The  $C$  vector thus has the form:

$$C = (c_1, c_2, \dots, c_g, \underbrace{0, \dots, 0}_{p-g})$$

Finally the kernel vector can be constructed. The only contributing terms (e.g. greater-than  $5\sigma$ ) will be at the beginning and end of  $K$ . In order for it to have the same size as  $C$ ,  $K$  must have the form:

$$K = \frac{1}{n} (\phi_h(0), \phi_h(\delta), \dots, \phi_h(l\delta), \underbrace{0, \dots, 0}_{p-2l-1}, \phi_h(l\delta), \dots, \phi_h(\delta))$$

The  $C * K$  convolution is ultimately performed and the zero bins removed from  $C$  to obtain the final result. An example of its use can be seen in Figure A.6.

<sup>6</sup>To be stringent it has  $\mathcal{O}(n \log n)$  complexity for all highly composite numbers of which all powers of two are a subset[103].

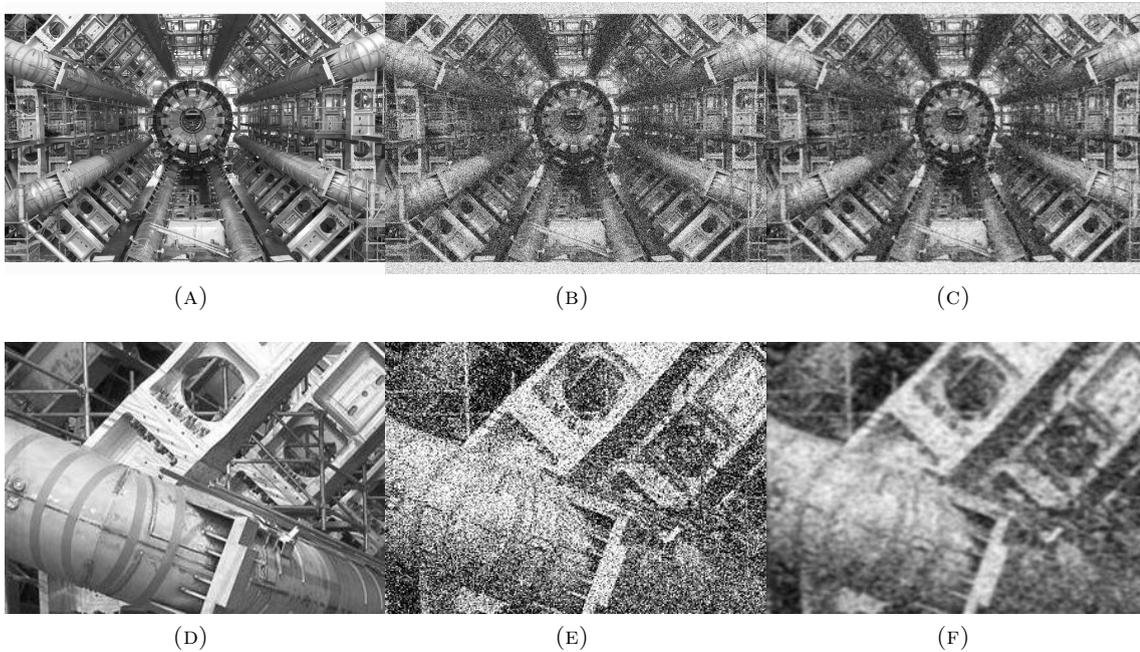


FIGURE A.6: KDE applied to a black and white photograph. The black/white photograph is described as a histogram with 256 possible entries per bin corresponding to different levels of light intensity. Figure (A) shows a well known photograph. In (B) a random noise has been added. (C) shows the KDE when applied to the noisy image. The bottom row figures show a cutout of the top left corner. The KDE diminishes the statistical fluctuations but has difficulty describing sharp contrasts.

Summarising the KDE approach:

- The general multidimensional KDE is simplified by binning the data distribution first and performing the KDE at histogram level. This is similar to convoluting a vector containing the histogram values with a 'kernel vector'.
- In order to ignore insignificant terms the vectors are modified such that their convolution will only include products over a given threshold.
- A discrete Fourier transform is performed on both vector separately. Their vector product is calculated and an inverse Fourier transform is performed on the result.

The computing time of the approach is tested on an example distribution of  $N$  events in two dimensions and compared to an unbinned approach. The characteristic evaluation time for the two methods is seen in Figure A.7. The original approach is polynomial in data-size while the binned is constant.

The binned KDE is however a simplification with less accuracy than the original approach. The use can only be found if the procedure has to be repeated a large number of times. As a reference, the analysis use 11 bins in either dimension and has a general statistics size of  $\sim 1.5 \cdot 10^5$  events available.

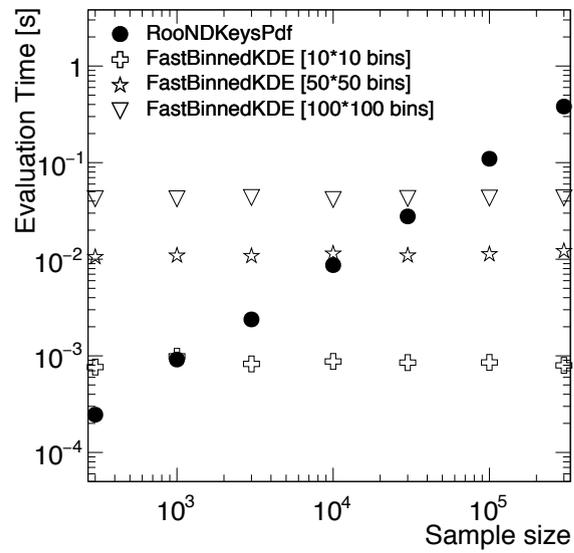


FIGURE A.7: Evaluation time of KDE methods as function of sample size and binning. The algorithms are applied on a two dimensional distribution.

### A.3 PDF Transformations

#### Transformation to known distribution

As was mentioned above, it can in general be difficult to figure out a suitable function to describe the data distributions. As such, it is often useful to transform the *pdf* to something that is easier to describe.

If a transformation is found that will make the analysis simpler, it should be bijective and constructed in such a way that it will not bias the result or throw away information.

When applying a Gaussian KDE for instance, it can easily be imagined that it is simpler to get a valid description of a distribution that is Gaussian-like rather than something more complex. To make the discussion more concrete, assume a *pdf* of some observable  $x$ :  $f(x)$ . We want to find a function  $y(x)$  such that  $y$  is distributed according to some chosen distribution,  $g(y)$ . In order to satisfy the requirement that a measurement using either  $x$  or  $y$  will yield the exact same result, it is necessary that the probability of any  $y$  observation  $y_0$  must be the same as the probability of the original  $x_0$  observation:

$$P(x_0 - \epsilon_x < x_0 < x_0 + \epsilon_x) = P(y_0 - \epsilon_y < y_0 < y_0 + \epsilon_y), \quad y_0 = y(x_0), \quad y \pm \epsilon_y = y(x_0 \pm \epsilon_x)$$

Integrating  $f$  and  $g$  over the  $\epsilon$  intervals describes these probabilities. Notice that a consequence of this is that the requirement also guarantees that any likelihood function constructed from either  $f$  or  $g$  will give the same result.

If the  $\epsilon$ -intervals tend to zero the following differential equation emerges:

$$f(x)dx = g(y)dy \Rightarrow \frac{dy}{dx} = \frac{f(x)}{g(y)} \quad 7$$

By choosing a  $g(y)$  it is possible to find the transformation by solving a differential equation. The two examples that are used in the analysis are transformation to a uniform distribution and transformation to a normal distribution.

The transformation to a uniform distribution is the simplest to find. If the interval  $[0, 1]$  is chosen for  $g$ ,  $g(y) = 1$ , and:

$$\frac{dy}{dx} = f(x) \Rightarrow y(x) = \int_{-\infty}^x dx' f(x') = \text{cdf}(x).$$

The cumulative density function of  $x$  is distributed uniformly as expected.

Turning now to the case of transformations to normal distributions. It will most likely be difficult to solve the differential equation directly. A trick can be used however: If  $x$  is first transformed to a uniform distribution and subsequently to the normal distribution it will only be necessary to solve the more simple:

$$\frac{dy}{dx} = \frac{1}{g(y)} = \frac{1}{\frac{1}{\sqrt{2}}e^{-y(x)^2/2}} = \sqrt{2}e^{y(x)^2/2}$$

---

<sup>7</sup>The sign choice of  $dy/dx$  is arbitrary. Here positive is chosen such that a greater  $x$  will result in a greater  $y$ .

A solution can be guessed based on the inverse error function:

$$\frac{d(\operatorname{erf}^{-1}(x))}{dx} = \frac{1}{2} \sqrt{\pi} e^{|\operatorname{erf}^{-1}(x)|^2} \rightarrow y(x) = \sqrt{2} \cdot \operatorname{erf}^{-1}(2x),$$

meaning that the transformation that will make any distribution normal is:

$$y(x) = \sqrt{2} \cdot \operatorname{erf}^{-1}(2 \cdot \operatorname{cdf}(x))$$

An example, where the two transformations are applied, can be found in Figure A.8.

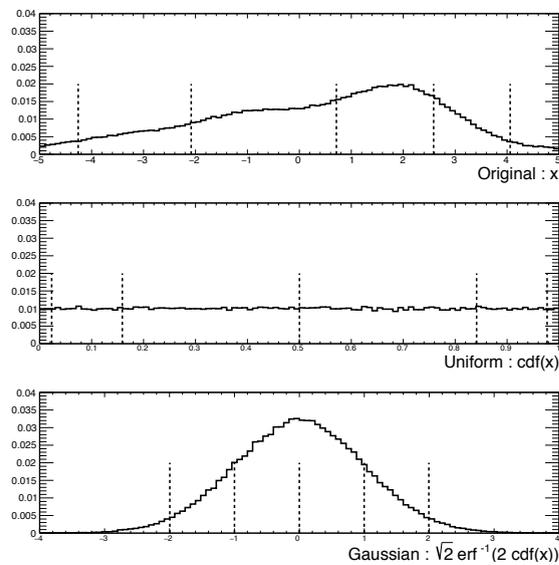


FIGURE A.8: Example of *pdf* transformations. In the top figure an arbitrary distribution is selected. In the centre and bottom figures, the same dataset is shown after the Uniform and Gaussian transformations have been applied. The dashed lines indicate where 2.3%, 15.9%, 50%, 84.1 and 97% of the distribution integral is contained to the left of, corresponding to the Gaussian standard deviations. The entries going into each interval are the same for the three distributions.

## Decorrelating observables

As has been mentioned several times, it is in general problematic if the dimensionality of the observable distribution grows too high. The characteristic amount of data that is required to describe the distributions will grow exponentially.

Another approach to reduce the dimensionality of the problem in question is to construct factorised *pdfs*.

Consider an analysis that uses two observables  $x$  and  $y$  with *pdf*  $f(x, y)$ . If the two outcome of  $x$  and  $y$  are independent,  $f(x, y)$  can be written as:

$$f(x, y) = g(x) \cdot h(y)$$

Where  $g$  and  $h$  describe the densities of  $x$  and  $y$ . Since these are individually one-dimensional, the full problem now has been reduced to also being one-dimensional.

In real life it is seldom the case that the observables entering into a measurement are completely independent by construction.

If, in the example above,  $pdf(x)$  and  $pdf(y)$  are linearly correlated, the same information will loosely speaking be used more than once in the likelihood. Assuming asymptotic behaviour of the likelihood ratio will in this scenario tend to overconstrain the parameters.

To see why this is the case it is useful to consider the example with the highest possible correlation; where  $x$  and  $y$  are the same observable. It is furthermore assumed that the individual distribution  $g(x)$  is unbiased and has asymptotic behaviour.

The  $pdf$  is now  $f(x) = g^2(x)$  and the likelihood ratio for  $f$  follows:

$$-2 \ln \lambda_f(\alpha) = -2 \ln \left( \frac{\prod g^2(x_i|\alpha)}{\prod g^2(x_i|\hat{\alpha})} \right) = 2 \cdot (-2) \ln \left( \frac{\prod g(x_i|\alpha)}{\prod g(x_i|\hat{\alpha})} \right) = 2 \cdot (-2 \ln \lambda_g)$$

It is clear that if  $g$  is unbiased in the sense that it on average will give a maximum at the true  $\alpha$  value so will  $f$ , meaning that in principle  $f$  should still be a valid  $pdf$  to use.

If standard asymptotic behaviour of the test statistic for  $f$  is assumed however, the “asymptotic 95% CL” limits will no longer cover the true value 95% of the time. In the example above this assumption will yield limits that do not have correct coverage since the actual test statistic distribution is twice as wide as assumed.

It is possible to remove any linear correlation from a dataset. Suppose a data vector  $X$  consist of  $p$  observables, each having been measured  $n$  times.  $X$  has covariance matrix  $C_X$  with elements:

$$(C_X)_{ij} = \frac{1}{n_{\text{data}}} \sum_{\text{data}} (x_i - \mu_i)(x_j - \mu_j) \quad , i, j = 1, \dots, p$$

where  $\mu_i$  denotes the observable mean and the summation index over the data columns have been made implicit. In the following  $\mu$ , without index will be used to describe the vector of means. It can be demonstrated that the dataset matrix:

$$Z = C_X^{-\frac{1}{2}}(X - \mu) \quad ,$$

has identity covariance matrix. To see why this is the case, it is useful to construct the intermediate data-vector  $Y = X - \mu$  with zero mean. The covariance matrix  $C_X$  is in terms of  $Y$  given by:

$$C_X = E(YY^T)$$

The covariance matrix  $C_Z$  for  $Z$  follows similarly:

$$C_Z = E(ZZ^T) = E(C_X^{-\frac{1}{2}}YY^T(C_X^{-\frac{1}{2}})^T) = C_X^{-\frac{1}{2}}E[YY^T](C_X^{-\frac{1}{2}})^T = C_X^{-\frac{1}{2}}C_X C_X^{-\frac{1}{2}} = I$$

The linear correlation has been removed in  $Z$ . The linear decorrelation is a good starting point when constructing factorised  $pdfs$ .

#### A.4 Comment on MVAs

The  $H \rightarrow ZZ^* \rightarrow 4\ell$  final state can be described with eight parameters. Ideally, it would be possible to construct a *pdf* of as many dimensions which would accurately describe any model that will be examined.

This will be difficult to do and as such multi-variate methods are developed. As has been the theme of large parts of this section, it will be the goal to perform a dimensional reduction such that models are simpler to describe.

One of the most well known methods that is used several times during the analysis is the Boosted Decision Tree (BDT). A short description on how it works will be given here to clarify its use.

A decision tree is an intuitive method of dividing an observable space. The structure of one is exemplified in Figure A.9. The general idea is to start with scanning one of the observables to the point where a cut would result in the highest signal and background separation. The space is now divided into two boxes; one with high signal content and one with high background content.

The scan can now be performed again on these individually whereby smaller regions with higher purity is found. In principle the process could continue until a large number of boxes of very high purity are found. In reality, since the training is performed on finite statistics, it will just be optimised on the statistical fluctuations present in the dataset, the tree is said to be overtrained.

A different approach would be to train a multitude of different, simple trees and average their result. Each tree is a weak classifier, but their combination can have high separation power.

A natural question is: If each tree is constructed to have maximal separation power, why will they not end up being identical?

A way to get around this problem is to perform boosting. The algorithm will start by training the first tree, which will misclassify a number of events. These are in the next training tree assigned a higher weight when calculating the signal to background separation. The next tree is hence optimised to perform separation on events where the first failed.

Different boosting methods exist, where the adaptive boosting is likely the most popular choice<sup>8</sup>. Here the misclassified events are weighted by a common boost weight  $\alpha$ , derived from the misclassification rate, *err*, of the previous tree:

$$\alpha = \frac{1 - \text{err}}{\text{err}}, \quad \text{err} < 0.5$$

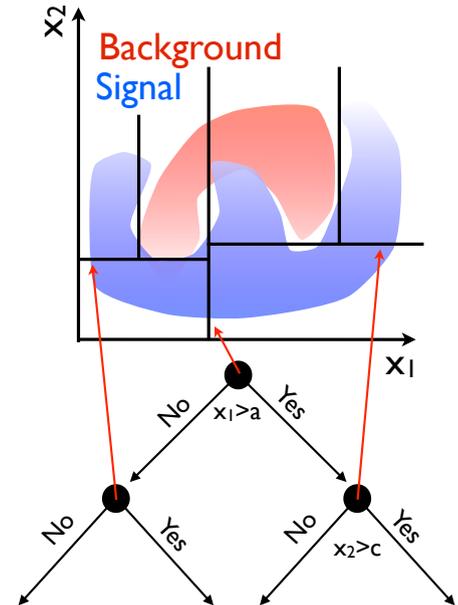


FIGURE A.9: Decision tree for a thought up process.

<sup>8</sup>See [89, Sect. 7.1,8.12] for a more detailed description. The text here is also mainly derived from this source.

The individual tree will predict events to be either signal or background like. Let  $h(x)$  denote the result of an individual tree such that  $+1$  indicates signal and  $-1$  indicates background. The boosted decision tree will predict the signal-likeness of  $x$  to be:

$$y_{\text{Boost}}(x) = \frac{1}{N_{\text{Trees}}} \sum_{i \in \text{Trees}} \ln(\alpha_i) \cdot h_i(x)$$

Other approaches exist than adaptive boosting, where only Bagging-resampling will be mentioned here. Bagging is not strictly a Boosting algorithm since it does not reweight events. Instead, each weak classifier is trained on a subsample of the training sample. Choosing the subsample at random will make statistical fluctuations even out when averaging the trained trees.

A large number of BDT parameters can be set to optimise the analysis in question. The TMVA package describes these in [89, Tab. 22,23,24]. The options for instance include how many trees there will be trained, the maximal allowed complexity of each tree, how large the bagging subsample is and what the maximal purity of a node is allowed to be.

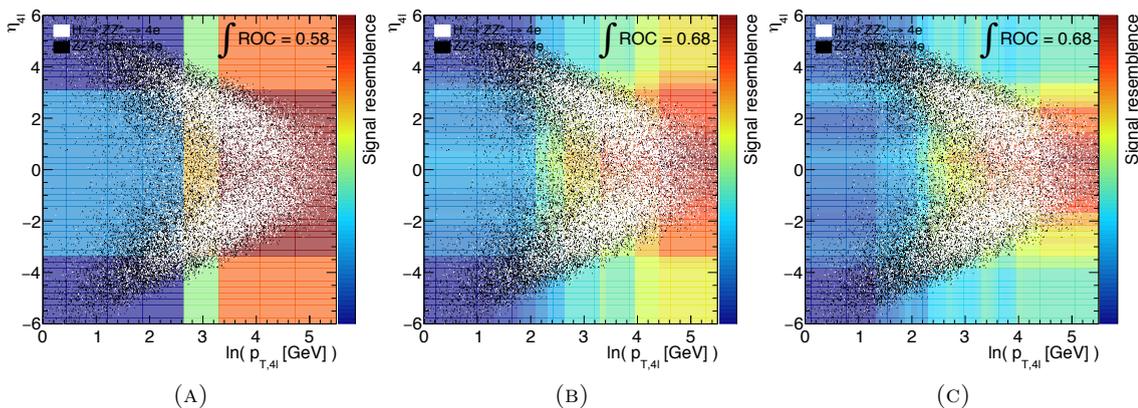


FIGURE A.10:  $p_{T,4l}$  and  $\eta_{4l}$  distributions for  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $ZZ^*-\text{continuum} \rightarrow 4\ell$ . The logarithm of the  $p_{T,4l}$  distribution is shown to make visualisation easier. The distributions are overlaid BDT classifiers of different complexity. (A):  $N_{\text{Trees}} = 3$ , (B):  $N_{\text{Trees}} = 30$ , (C):  $N_{\text{Trees}} = 300$ . Red indicates the BDT predicts events to be more signal-like. The training did not achieve better predictive power by going from 30 to 300 trees.

A BDT is trained to classify whether an event is more likely to be a Higgs or a background to illustrate the process. More specifically, it is trained to separate the event types  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $ZZ^*-\text{continuum} \rightarrow 4\ell$ , based on their transverse momenta and pseudorapidities. The distributions are shown in Figure A.10, where three classifiers of increasing complexity have been made. When adding more trees, the BDT will divide the observable space into a finer grid with better predictive power. At some complexity the BDT becomes saturated with trees and more complexity will not give more separation.

## B Fixed Hypothesis Test Appendix

NP name	Down Pull		Up pull		average effect [%]
	Significance	difference [%]	Significance	difference [%]	
<b>alpha_ATLAS_Higgs_mass</b>	2.26	-3.65	2.31	-1.90	2.77
<b>alpha_ATLAS_MU_MS_RES_MS</b>	2.36	0.49	2.32	-1.24	0.87
<b>alpha_ATLAS_norm_SF_H4L_Zbb_llmumu_2012</b>	2.37	0.81	2.33	-0.75	0.78
<b>alpha_ATLAS_EM_mRes_CT</b>	2.33	-0.66	2.33	-0.75	0.71
<b>alpha_pdf_qq</b>	2.37	0.66	2.34	-0.62	0.64
<b>alpha_ATLAS_norm_SF_H4L_Z_llee_2012</b>	2.36	0.57	2.34	-0.53	0.55
<b>alpha_QCDscale_VV</b>	2.36	0.49	2.34	-0.47	0.48
<b>alpha_ATLAS_EM_mRes_ST</b>	2.34	-0.53	2.34	-0.39	0.46
<b>alpha_ATLAS_LUMI_2012</b>	2.36	0.45	2.34	-0.44	0.45
alpha_ATLAS_rho_signal	2.34	-0.45	2.36	0.37	0.41
alpha_ATLAS_EM_mRes_MatID	2.34	-0.35	2.34	-0.45	0.40
alpha_ATLAS_MU_MS	2.35	-0.17	2.34	-0.52	0.34
alpha_ATLAS_MU_MS_RES_ID	2.35	0.18	2.34	-0.50	0.34
alpha_ATLAS_EM_mRes_PU	2.34	-0.36	2.34	-0.26	0.31
alpha_ATLAS_EM_mRes_MatCryo	2.34	-0.23	2.34	-0.30	0.26
alpha_ATLAS_EM_ES_Z	2.35	-0.03	2.34	-0.45	0.24
alpha_ATLAS_shape_SF_H4LEL_EFF_ISOIP_2012	2.36	0.24	2.34	-0.24	0.24
alpha_ATLAS_shape_SF_H4L_Z_llee_2012	2.36	0.35	2.35	0.11	0.23
alpha_ATLAS_EL_2012_REC_Low	2.35	0.19	2.35	-0.19	0.19
alpha_ATLAS_EM_LArElecUnconv_Barrel	2.35	-0.12	2.34	-0.26	0.19
alpha_ATLAS_EM_L2Gain	2.35	-0.10	2.34	-0.24	0.17
alpha_ATLAS_EM_MatCryo_Barrel	2.34	-0.25	2.35	-0.07	0.16
alpha_ATLAS_shape_SF_H4L_Zbb_llmumu_2012	2.35	0.05	2.36	0.25	0.15
alpha_ATLAS_EM_mRes_MatCalo	2.35	-0.12	2.35	-0.14	0.13
alpha_ATLAS_EL_2012_ID_low	2.35	0.12	2.35	-0.13	0.12
alpha_ATLAS_EM_ES_Momentum	2.35	-0.07	2.35	-0.17	0.12
alpha_ATLAS_MU_EFF	2.35	0.12	2.35	-0.12	0.12
alpha_ATLAS_EL_2012_IDST_high	2.35	0.11	2.35	-0.11	0.11
alpha_ATLAS_EM_mRes_MatGap	2.35	-0.09	2.35	-0.11	0.10
alpha_ATLAS_EL_2012_ST_10	2.35	0.10	2.35	-0.10	0.10
alpha_ATLAS_EM_LArCalib_Barrel	2.35	-0.06	2.35	-0.12	0.09
alpha_ATLAS_EM_PS_Barrel	2.35	-0.03	2.35	-0.14	0.09
alpha_ATLAS_EM_MatID_2	2.35	-0.11	2.35	-0.06	0.09
alpha_ATLAS_EM_Pedestal	2.35	-0.01	2.35	-0.14	0.08
alpha_ATLAS_rho_Redbg	2.35	-0.06	2.35	0.09	0.07
alpha_ATLAS_EL_2012_ST_15	2.35	0.07	2.35	-0.07	0.07
alpha_ATLAS_EM_LArElecUnconv_EC	2.35	-0.05	2.35	-0.09	0.07
alpha_ATLAS_EL_2012_REC_high	2.35	0.07	2.35	-0.07	0.07
alpha_ATLAS_MU_2012_TRIG	2.35	0.07	2.35	-0.07	0.07
alpha_ATLAS_EM_MatCryo_EC	2.35	-0.07	2.35	-0.04	0.06
alpha_ATLAS_EM_MatID_3	2.35	-0.06	2.35	-0.04	0.05
alpha_ATLAS_EM_MatCalo_Barrel	2.35	-0.09	2.35	0.01	0.05
alpha_ATLAS_EM_MatID_1	2.35	-0.04	2.35	-0.05	0.04
alpha_ATLAS_EM_LArCalib_EC	2.35	0.00	2.35	-0.08	0.04
alpha_ATLAS_EM_L1Gain	2.35	0.00	2.35	-0.08	0.04
alpha_ATLAS_EM_LArUnconvCalib_Barrel	2.35	-0.06	2.35	-0.01	0.04
alpha_ATLAS_EM_Geant4	2.35	-0.05	2.35	-0.02	0.04
alpha_ATLAS_EM_S12_Barrel	2.35	-0.01	2.35	-0.06	0.03
alpha_ATLAS_EM_LArElecCalib	2.35	-0.04	2.35	0.02	0.03
alpha_ATLAS_EM_S12_EC	2.35	-0.01	2.35	-0.05	0.03
alpha_ATLAS_EM_MatID_4	2.35	-0.03	2.35	0.01	0.02
alpha_ATLAS_EM_PS_EC	2.35	0.00	2.35	-0.03	0.02
alpha_ATLAS_rho_ZZ	2.35	0.01	2.35	0.02	0.02
alpha_ATLAS_EM_MatCalo_EC	2.35	-0.02	2.35	-0.01	0.02
alpha_ATLAS_EM_LArUnconvCalib_EC	2.35	-0.01	2.35	-0.01	0.01
alpha_ATLAS_EL_2012_TRIG	2.35	0.00	2.35	0.00	0.00
alpha_ATLAS_BR_VV	2.35	0.00	2.35	0.00	0.00
alpha_ATLAS_EL_2012_ST_7	2.35	0.00	2.35	0.00	0.00
alpha_pdf_Higgs_ggH	2.35	0.00	2.35	0.00	0.00
alpha_QCDscale_Higgs_ggH	2.35	0.00	2.35	0.00	0.00
alpha_ATLAS_EM_ConvRadius	2.35	0.00	2.35	0.00	0.00
alpha_ATLAS_EM_ConvEfficiency	2.35	0.00	2.35	0.00	0.00
alpha_ATLAS_EM_ConvFakeRate	2.35	0.00	2.35	0.00	0.00
Including all	2.35	0.00	2.35	0.00	0.00

TABLE B.1: Effect of each systematic uncertainty on the  $J^p = 0^+$  vs.  $0^-$  separation. The effect of each nuisance parameter is calculated as the difference between the result when including all systematic uncertainties (last line), and when each is shifted one standard deviation to either side. The average effect is calculated from the absolute value of the upward and downward shift. A threshold of 0.45% was chosen as to where a systematic uncertainty has an actual impact on the result. These are indicated with bold script in the table.

## C Tensor Structure Fit Appendix

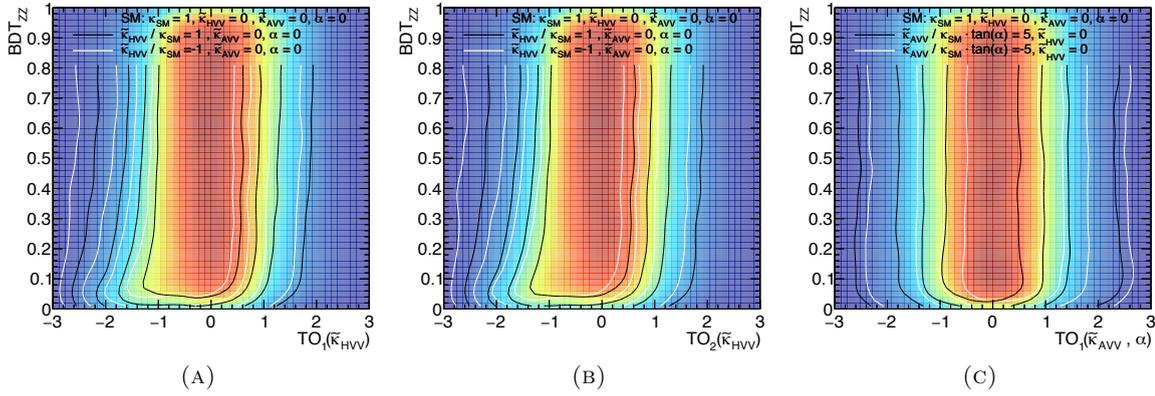


FIGURE C.11: Two dimensional observable distributions. (A)  $TO_1(\tilde{\kappa}_{HVV}) - TO_2(\tilde{\kappa}_{HVV})$  vs.  $BDT_{ZZ}$ , (B)  $TO_1(\tilde{\kappa}_{HVV}) + TO_2(\tilde{\kappa}_{HVV})$  vs.  $BDT_{ZZ}$ , (C)  $O_1(\tilde{\kappa}_{AVV}, \alpha)$  vs.  $BDT_{ZZ}$ . The coloured background represent the Standard Model, while the lines indicate 10%, 30%, 50%, 70% and 90% of the SM maximum for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = \pm 1$  (A,B) and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha = \pm 5$  (C).

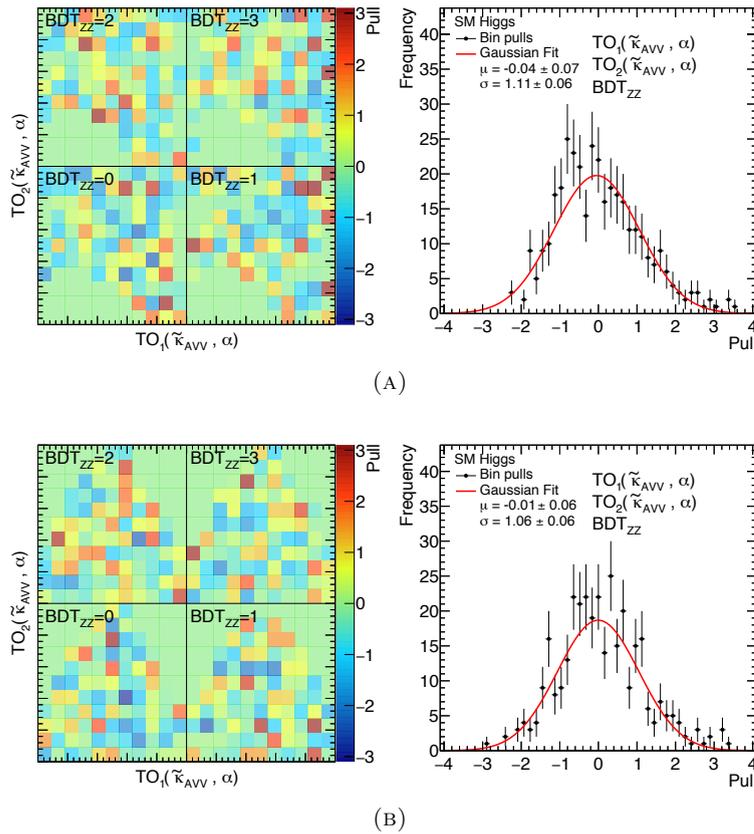
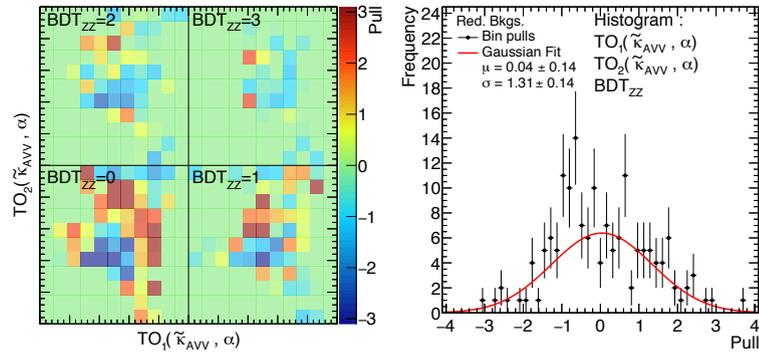
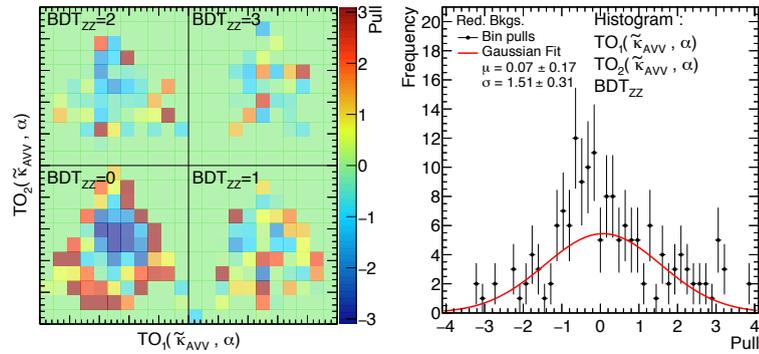


FIGURE C.12: Reweighting validation for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observables. Left figures show the individual pulls in the four different  $BDT_{ZZ}$  bins. The right figures show the pull distributions. Standard Model shapes for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  observables (A) and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  observables (B).



(A)



(B)

FIGURE C.13: Kernel Density Estimation validation for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$   $\tan \alpha$  observables. Left figures show the individual pulls in the four different  $BDT_{ZZ}$  bins. The right figures show the pull distributions. (A) Reducible background shape for  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  observables. (B): Reducible background shape for  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$   $\tan \alpha$  distributions.

NP name	Down Pull		Up pull		average effect [%]
	Significance	Difference [%]	Significance	Difference [%]	
<b>alpha_ATLAS_shape_SF_H4l_Z_ll ee_2012</b>	1.59	2.70	1.57	1.90	2.30
<b>alpha_ATLAS_shape_SF_H4l_Z_ll ee_2011</b>	1.57	1.59	1.56	1.22	1.41
<b>alpha_ATLAS_norm_SF_H4l_Z_ll ee_2012</b>	1.56	1.30	1.52	-1.22	1.26
<b>alpha_ATLAS_norm_SF_H4l_Zbb_llmumu_2012</b>	1.56	0.96	1.53	-0.93	0.95
<b>alpha_ATLAS_MU_MS</b>	1.55	0.48	1.53	-0.95	0.72
<b>alpha_ATLAS_shape_SF_H4l_Zbb_llmumu_2012</b>	1.52	-1.21	1.54	-0.01	0.61
<b>alpha_pdf_qq</b>	1.55	0.60	1.54	-0.55	0.58
<b>alpha_ATLAS_EM_ES_Z</b>	1.55	0.33	1.53	-0.70	0.52
alpha_QCDscale_VV	1.55	0.48	1.54	-0.44	0.46
alpha_ATLAS_norm_SF_H4l_Z_ll ee_2011	1.55	0.47	1.54	-0.35	0.41
alpha_ATLAS_LUMI_2012	1.55	0.40	1.54	-0.38	0.39
alpha_ATLAS_norm_SF_H4l_Zbb_llmumu_2011	1.55	0.33	1.54	-0.31	0.32
alpha_ATLAS_EM_ES_Momentum	1.54	-0.24	1.54	-0.36	0.30
alpha_ATLAS_EM_MatCryo_Barrel	1.54	-0.44	1.54	-0.15	0.29
alpha_ATLAS_EM_MatID_3	1.54	-0.35	1.54	-0.24	0.29
alpha_ATLAS_EM_mRes_MatGap	1.54	-0.31	1.54	-0.27	0.29
alpha_ATLAS_EM_PS_Barrel	1.54	-0.25	1.54	-0.33	0.29
alpha_ATLAS_EM_LArUnconvCalib_EC	1.54	-0.30	1.54	-0.27	0.28
alpha_ATLAS_EM_MatCalo_Barrel	1.54	-0.31	1.54	-0.25	0.28
alpha_ATLAS_EM_MatID_4	1.54	-0.29	1.54	-0.27	0.28
alpha_ATLAS_EM_LArCalib_Barrel	1.54	-0.23	1.54	-0.33	0.28
alpha_ATLAS_EM_S12_EC	1.54	-0.25	1.54	-0.31	0.28
alpha_ATLAS_EM_LArUnconvCalib_Barrel	1.54	-0.31	1.54	-0.25	0.28
alpha_ATLAS_EM_LArElecUnconv_Barrel	1.54	-0.14	1.54	-0.41	0.28
alpha_ATLAS_EM_LArElecCalib	1.54	-0.30	1.54	-0.25	0.28
alpha_ATLAS_EM_LArElecUnconv_EC	1.54	-0.17	1.54	-0.39	0.28
alpha_ATLAS_EM_LArCalib_EC	1.54	-0.23	1.54	-0.32	0.28
alpha_ATLAS_EM_MatID_1	1.54	-0.29	1.54	-0.26	0.28
alpha_ATLAS_EM_MatCalo_EC	1.54	-0.28	1.54	-0.26	0.27
alpha_ATLAS_EM_S12_Barrel	1.54	-0.23	1.54	-0.32	0.27
alpha_ATLAS_EM_ConvRadius	1.54	-0.27	1.54	-0.27	0.27
alpha_ATLAS_EM_PS_EC	1.54	-0.25	1.54	-0.29	0.27
alpha_ATLAS_EM_ConvFakeRate	1.54	-0.27	1.54	-0.27	0.27
alpha_ATLAS_EM_ConvEfficiency	1.54	-0.27	1.54	-0.27	0.27
alpha_ATLAS_EM_L1Gain	1.54	-0.24	1.54	-0.30	0.27
alpha_ATLAS_EM_mRes_MatCalo	1.54	-0.25	1.54	-0.28	0.27
alpha_ATLAS_EM_Geant4	1.54	-0.28	1.54	-0.25	0.27
alpha_ATLAS_EM_MatID_2	1.54	-0.34	1.54	-0.20	0.27
alpha_ATLAS_EM_Pedestal	1.54	-0.21	1.54	-0.31	0.26
alpha_ATLAS_EM_MatCryo_EC	1.54	-0.38	1.54	-0.14	0.26
alpha_ATLAS_EM_L2Gain	1.54	-0.31	1.54	-0.20	0.25
alpha_ATLAS_EM_mRes_MatCryo	1.54	-0.20	1.54	-0.24	0.22
alpha_ATLAS_MU_EFF	1.55	0.21	1.54	-0.20	0.21
alpha_ATLAS_EM_mRes_MatID	1.54	-0.15	1.54	-0.23	0.19
alpha_ATLAS_MU_MS_RES_ID	1.54	-0.16	1.54	-0.19	0.18
alpha_ATLAS_EM_mRes_ST	1.54	-0.17	1.54	-0.19	0.18
alpha_ATLAS_EM_mRes_PU	1.54	-0.18	1.54	-0.18	0.18
alpha_ATLAS_EL_2012_REC_Low	1.55	0.17	1.54	-0.17	0.17
alpha_ATLAS_shape_SF_H4l_EL_EFF_ISOIP_2012	1.55	0.16	1.54	-0.16	0.16
alpha_ATLAS_MU_MS_RES_MS	1.54	-0.26	1.54	-0.03	0.14
alpha_ATLAS_shape_SF_H4l_Zbb_llmumu_2011	1.54	-0.21	1.54	-0.01	0.11
alpha_ATLAS_EL_2012_IDST_high	1.55	0.11	1.54	-0.11	0.11
alpha_QCDscale_Higgs_ggH	1.54	0.00	1.54	-0.16	0.08
alpha_pdf_Higgs_ggH	1.54	0.00	1.54	-0.16	0.08
alpha_ATLAS_EL_2012_ID_low	1.54	0.05	1.54	-0.05	0.05
alpha_ATLAS_EM_mRes_CT	1.54	-0.02	1.54	-0.06	0.04
alpha_ATLAS_EL_2012_ST_7	1.54	0.04	1.54	-0.04	0.04
alpha_ATLAS_EL_2012_ST_10	1.54	0.03	1.54	-0.03	0.03
alpha_ATLAS_rho_signal	1.54	-0.02	1.54	0.00	0.01
alpha_ATLAS_EL_2012_ST_15	1.54	0.01	1.54	-0.01	0.01
alpha_ATLAS_rho_Redbkg	1.54	0.02	1.54	0.00	0.01
alpha_ATLAS_EL_2012_REC_high	1.54	-0.01	1.54	0.01	0.01
alpha_ATLAS_rho_ZZ	1.54	0.00	1.54	-0.00	0.00
alpha_BR_VV	1.54	0.00	1.54	0.00	0.00
alpha_ATLAS_MU_2012_TRIG	1.54	0.00	1.54	0.00	0.00

TABLE C.2: Effect of each systematic uncertainty on the separation of the Standard Model from the  $\frac{\tilde{\kappa}_{HV V}}{\kappa_{SM}} = -0.8$  model. The effect of each nuisance parameter is calculated as the difference between the result when including all systematic uncertainties, and when each is shifted one standard deviation to either side. The average effect is calculated from the absolute value of the upward and downward shift. A threshold of 0.5 % (indicated by bold font) is chosen as to where a systematic uncertainty has a potential impact on the result.

NP name	Down Pull		Up pull		average effect [%]
	Significance	Difference [%]	Significance	Difference [%]	
<b>alpha_ATLAS_shape_SF_H4L_Z_llcc_2012</b>	1.86	1.32	1.85	1.03	1.17
<b>alpha_ATLAS_norm_SF_H4L_Z_llcc_2012</b>	1.85	1.01	1.82	-0.93	0.97
<b>alpha_ATLAS_norm_SF_H4L_Zbb_llmumu_2012</b>	1.85	0.92	1.82	-0.86	0.89
<b>alpha_pdf_qq</b>	1.85	0.92	1.82	-0.85	0.88
<b>alpha_ATLAS_shape_SF_H4L_Z_llcc_2011</b>	1.85	0.91	1.84	0.57	0.74
<b>alpha_QCDscale_VV</b>	1.85	0.76	1.82	-0.69	0.72
<b>alpha_ATLAS_LUMI_2012</b>	1.85	0.60	1.82	-0.55	0.57
alpha_ATLAS_norm_SF_H4L_Z_llcc_2011	1.84	0.50	1.83	-0.34	0.42
alpha_ATLAS_EL_2012_REC_Low	1.84	0.38	1.83	-0.35	0.36
alpha_ATLAS_MU_MS	1.84	0.24	1.83	-0.43	0.34
alpha_ATLAS_norm_SF_H4L_Zbb_llmumu_2011	1.84	0.36	1.83	-0.31	0.33
alpha_ATLAS_shape_SF_H4L_EL_EFF_ISOIP_2012	1.84	0.33	1.83	-0.32	0.33
alpha_ATLAS_MU_EFF	1.84	0.28	1.83	-0.27	0.28
alpha_ATLAS_shape_SF_H4L_Zbb_llmumu_2012	1.83	-0.07	1.84	0.37	0.22
alpha_ATLAS_EM_ES_Z	1.84	0.08	1.83	-0.27	0.17
alpha_ATLAS_EL_2012_IDST_high	1.84	0.15	1.83	-0.13	0.14
alpha_ATLAS_EM_mRes_ST	1.83	-0.09	1.83	-0.16	0.12
alpha_ATLAS_EL_2012_ID_Low	1.84	0.11	1.83	-0.11	0.11
alpha_ATLAS_EM_mRes_PU	1.83	-0.08	1.83	-0.15	0.11
alpha_ATLAS_EM_MatID_3	1.83	-0.13	1.83	-0.08	0.11
alpha_ATLAS_EM_mRes_MatGap	1.83	-0.08	1.83	-0.13	0.10
alpha_ATLAS_EM_mRes_CT	1.83	0.02	1.83	-0.19	0.10
alpha_ATLAS_MU_MS_RES_MS	1.83	-0.03	1.83	-0.17	0.10
alpha_ATLAS_EM_LArElecUnconv_Barrel	1.83	-0.03	1.83	-0.17	0.10
alpha_ATLAS_EM_PS_EC	1.83	-0.08	1.83	-0.12	0.10
alpha_ATLAS_MU_MS_RES_ID	1.83	-0.01	1.83	-0.19	0.10
alpha_ATLAS_EM_MatCalo_EC	1.83	-0.11	1.83	-0.08	0.10
alpha_ATLAS_EM_MatCryo_Barrel	1.83	-0.17	1.83	-0.02	0.10
alpha_ATLAS_EM_LArUnconvCalib_EC	1.83	-0.11	1.83	-0.08	0.10
alpha_ATLAS_EM_L1Gain	1.83	-0.08	1.83	-0.11	0.10
alpha_ATLAS_EM_mRes_MatCalo	1.83	-0.09	1.83	-0.10	0.10
alpha_ATLAS_EM_mRes_MatID	1.83	-0.05	1.83	-0.14	0.09
alpha_ATLAS_EM_L2Gain	1.83	-0.12	1.83	-0.07	0.09
alpha_ATLAS_EM_Pedestal	1.83	-0.04	1.83	-0.14	0.09
alpha_ATLAS_EM_MatID_4	1.83	-0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_ConvFakeRate	1.83	-0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_ConvRadius	1.83	-0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_ConvEfficiency	1.83	-0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_LArElecCalib	1.83	-0.11	1.83	-0.07	0.09
alpha_ATLAS_EM_MatID_1	1.83	-0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_PS_Barrel	1.83	-0.06	1.83	-0.12	0.09
alpha_ATLAS_EM_S12_EC	1.83	-0.07	1.83	-0.11	0.09
alpha_ATLAS_EL_2012_ST_7	1.84	0.09	1.83	-0.09	0.09
alpha_ATLAS_EM_S12_Barrel	1.83	-0.08	1.83	-0.10	0.09
alpha_ATLAS_EM_LArCalib_EC	1.83	-0.06	1.83	-0.11	0.09
alpha_ATLAS_EM_LArCalib_Barrel	1.83	-0.06	1.83	-0.11	0.09
alpha_ATLAS_EM_ES_Momentum	1.83	-0.05	1.83	-0.12	0.09
alpha_ATLAS_EM_LArElecUnconv_EC	1.83	-0.04	1.83	-0.13	0.09
alpha_ATLAS_EM_Geant4	1.83	-0.10	1.83	-0.07	0.08
alpha_ATLAS_EM_MatCryo_EC	1.83	-0.14	1.83	-0.02	0.08
alpha_ATLAS_EM_mRes_MatCryo	1.83	-0.08	1.83	-0.08	0.08
alpha_ATLAS_EM_MatCalo_Barrel	1.83	-0.10	1.83	-0.06	0.08
alpha_ATLAS_EM_LArUnconvCalib_Barrel	1.83	-0.09	1.83	-0.06	0.08
alpha_ATLAS_EM_MatID_2	1.83	-0.11	1.83	-0.03	0.07
alpha_ATLAS_EL_2012_ST_15	1.84	0.03	1.83	-0.03	0.03
alpha_ATLAS_shape_SF_H4L_Zbb_llmumu_2011	1.83	-0.01	1.84	0.05	0.03
alpha_ATLAS_EL_2012_ST_10	1.83	0.03	1.83	-0.03	0.03
alpha_ATLAS_EL_2012_REC_high	1.83	0.01	1.83	-0.01	0.01
alpha_BR_VV	1.83	0.00	1.83	0.02	0.01
alpha_pdf_Higgs_ggH	1.83	0.00	1.83	0.02	0.01
alpha_QCDscale_Higgs_ggH	1.83	0.00	1.83	0.02	0.01
alpha_ATLAS_rho_Redbkg	1.83	-0.01	1.83	-0.01	0.01
alpha_ATLAS_rho_signal	1.83	-0.01	1.83	-0.00	0.00
alpha_ATLAS_rho_ZZ	1.83	-0.00	1.83	0.00	0.00
alpha_ATLAS_MU_2012_TRIG	1.83	0.00	1.83	0.00	0.00

TABLE C.3: Effect of each systematic uncertainty on the separation of the Standard Model from the  $\tilde{\kappa}_{AVV}^{\tilde{\kappa}_{SM}} \tan \alpha = 5$  model. The effect of each nuisance parameter is calculated as the difference between the result when including all systematic uncertainties, and when each is shifted one standard deviation to either side. The average effect is calculated from the absolute value of the upward and downward shift. A threshold of 0.5 % (indicated by bold font) was chosen as to where a systematic uncertainty has a potential impact on the result.

## D Comments on Correlation Scheme

The factorisation scheme that was used in the simultaneous fit of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  and  $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$  was found to have a slight tendency to overconstrain the fitted parameters. This appendix examines the possibility of decorrelating the observables to further reduce this effect.

The decorrelation of the  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$  observables that was described in the main section is for now disregarded as is the Gaussian transformation. Returning to the original observable definitions, the covariance matrix  $S$  of the four observables is calculated and the uncorrelated set is obtained<sup>9</sup>:

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \\ O_4 \end{pmatrix} = \begin{pmatrix} 5.3 & -6.1 & 0.0 & 1.0 \\ -6.1 & 13 & 0.0 & 2.9 \\ 0.0 & 0.0 & 7.2 & 0.0 \\ 1.0 & -2.9 & 0.0 & 8.4 \end{pmatrix} \begin{pmatrix} O_1(\tilde{\kappa}_{HVV}) - \mu_{O_1(\tilde{\kappa}_{HVV})} \\ O_2(\tilde{\kappa}_{HVV}) - \mu_{O_2(\tilde{\kappa}_{HVV})} \\ O_1(\tilde{\kappa}_{AVV}, \alpha) - \mu_{O_1(\tilde{\kappa}_{AVV}, \alpha)} \\ O_2(\tilde{\kappa}_{AVV}, \alpha) - \mu_{O_2(\tilde{\kappa}_{AVV}, \alpha)} \end{pmatrix}$$

Since the uncorrelated observables will be a mix of the previous and do not have the same physical interpretation as before, it is chosen simply to denote them  $O_1$  to  $O_4$ . Considering the correlation figures (3.12 and 3.31), the matrix follows intuition nicely. The  $O_1(\tilde{\kappa}_{HVV})$  and  $O_2(\tilde{\kappa}_{HVV})$  are highly linearly correlated. The new  $O_1$  and  $O_2$  is a mix of the two with opposite sign. The  $O_1(\tilde{\kappa}_{HVV})$  and  $O_2(\tilde{\kappa}_{HVV})$  observables were moreover demonstrated to be correlated with  $O_2(\tilde{\kappa}_{AVV}, \alpha)$  while independent of  $O_1(\tilde{\kappa}_{AVV}, \alpha)$  which is also reflected in the matrix by having contributions for the first but not the second. The decorrelation scheme has here been derived for the Standard Model distributions.

In principle, following the same arguments as for the *pdf* transformations and the fact that the rotation is bijective, it should not be possible to bias the results from basing the decorrelating matrix on SM distributions. It is however important to realise that this is only true until the factorisation step. It could be imagined that the observable correlations would differ at different coupling ratios. Rotating such that the Standard Model is uncorrelated and constructing the factorisation scheme based on this, could in this case bias the results.

The Gaussian transformation is applied to the *pdfs* again, to give an easier input to the KDE procedure. The final decorrelated distributions can for the Standard Model be seen in Figure D.14. Some structure is seen in the distributions after the decorrelating rotation. To establish if it is possible to construct a reliable factorisation scheme, the same set of tests as for the main results is repeated.

The model construction follows the same path as before. The template distributions are constructed for two centre of mass energies and four final states. The description at different coupling ratios is obtained with matrix element based reweighting. The distributions are finally estimated using KDEs before being added to the likelihood.

There are now six different correlations that can be tested to be independent. The factorisation tests are shown in Figures D.15 and D.16. The approach show promising results.

<sup>9</sup>The decorrelation procedure is in the following performed separately on the four different final state due to their different sensitivities. The numbers here correspond to the  $2e2\mu$  final state. A description of how the rotation matrix is obtained can be found in Section A.3

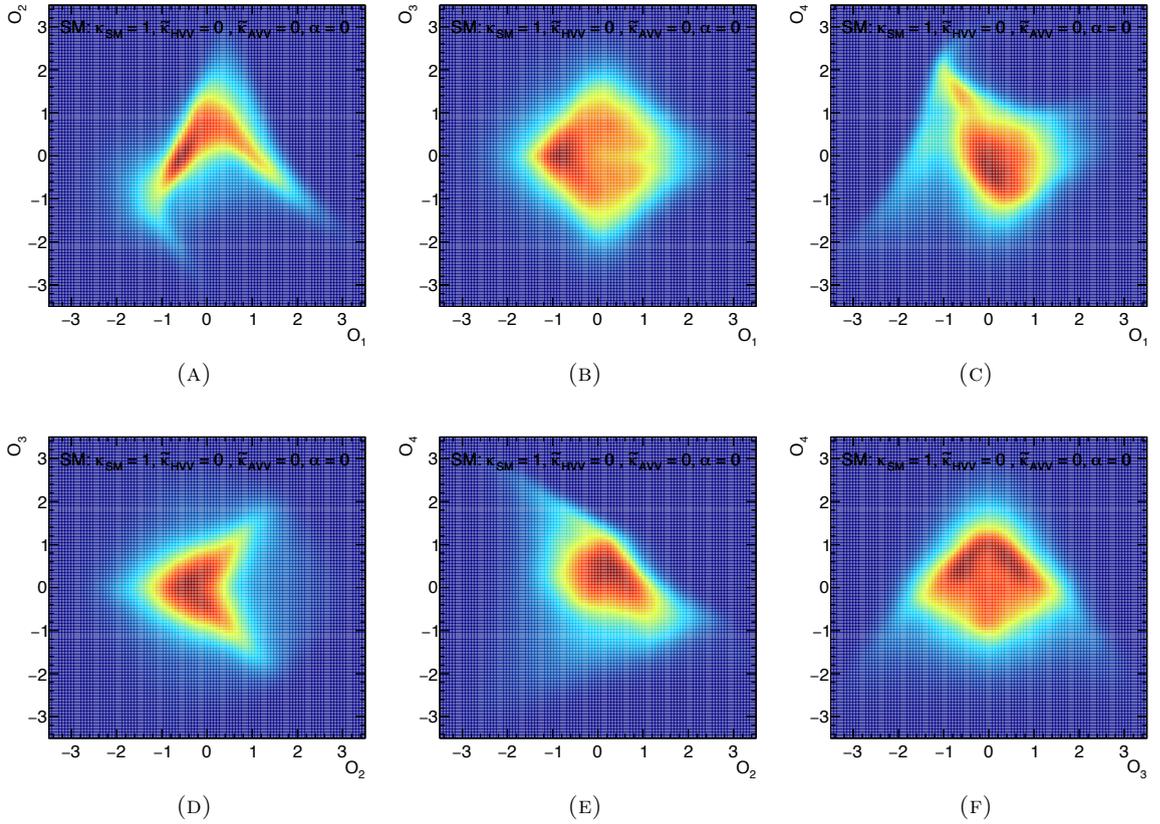


FIGURE D.14: Two dimensional projections of decorrelated observables. (A,B,C) show the dependency between  $O_1$  and  $O_2$ ,  $O_3$  and  $O_4$ . (D,E) show  $O_2$  vs.  $O_3$  and  $O_4$ .  $O_3$  vs.  $O_4$  is seen in (F). The linear dependency present in the original distributions is effectively removed.

The linear dependency of  $O_1$  to the other observables is reduced such that the factorised model will slightly underestimate the sensitivity. Some issues are seen in the second set of correlation test. It is observed that there are some problems with the fit for the factorised model for negative values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ . Most likely an error has been made when constructing the input templates.

To check whether the decorrelating matrix would change for different choices of input model, the correlation matrix ( $\rho$ ) and mean vector ( $\mu$ ) of the original observables are calculated<sup>10</sup>. The values are calculated for the SM and a BSM with  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}} = -5$ . The mean vector and correlation matrix for the Standard Model are:

$$\mu(SM)_{\text{Original}} = \begin{pmatrix} 1.2 \\ 0.45 \\ 0.0 \\ 0.14 \end{pmatrix} \quad \rho(SM)_{\text{Original}} = \begin{pmatrix} 1.0 & 0.95 & 0.0 & 0.28 \\ 0.95 & 1.0 & 0.0 & 0.42 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.28 & 0.42 & 0.0 & 1.0 \end{pmatrix},$$

<sup>10</sup>The decorrelation is based on the covariance matrix rather than the correlation matrix. The correlation matrix is presented since it is easier to interpret the numbers

while they take the following values for the BSM scenario:

$$\mu(BSM)_{\text{Original}} = \begin{pmatrix} 1.4 \\ 0.61 \\ 0.0 \\ 0.20 \end{pmatrix} \quad \rho(BSM)_{\text{Original}} = \begin{pmatrix} 1.0 & 0.93 & 0.0 & 0.38 \\ 0.93 & 1.0 & 0.0 & 0.59 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.38 & 0.59 & 0.0 & 1.0 \end{pmatrix}$$

The structure of the correlation and means are seen to change depending on the input model. This means that the decorrelating matrix will also depend on which model is used as vantage point.

The decorrelation procedure does however give promising results. With more work it should be possible to understand the weird likelihood behaviour at negative values of  $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ . More studies are needed on how the matrix structure changes with different models before a final approach can be defined. If the analysis is to be repeated when more statistics is collected, a decorrelation of the observable distributions could prove an interesting extension.

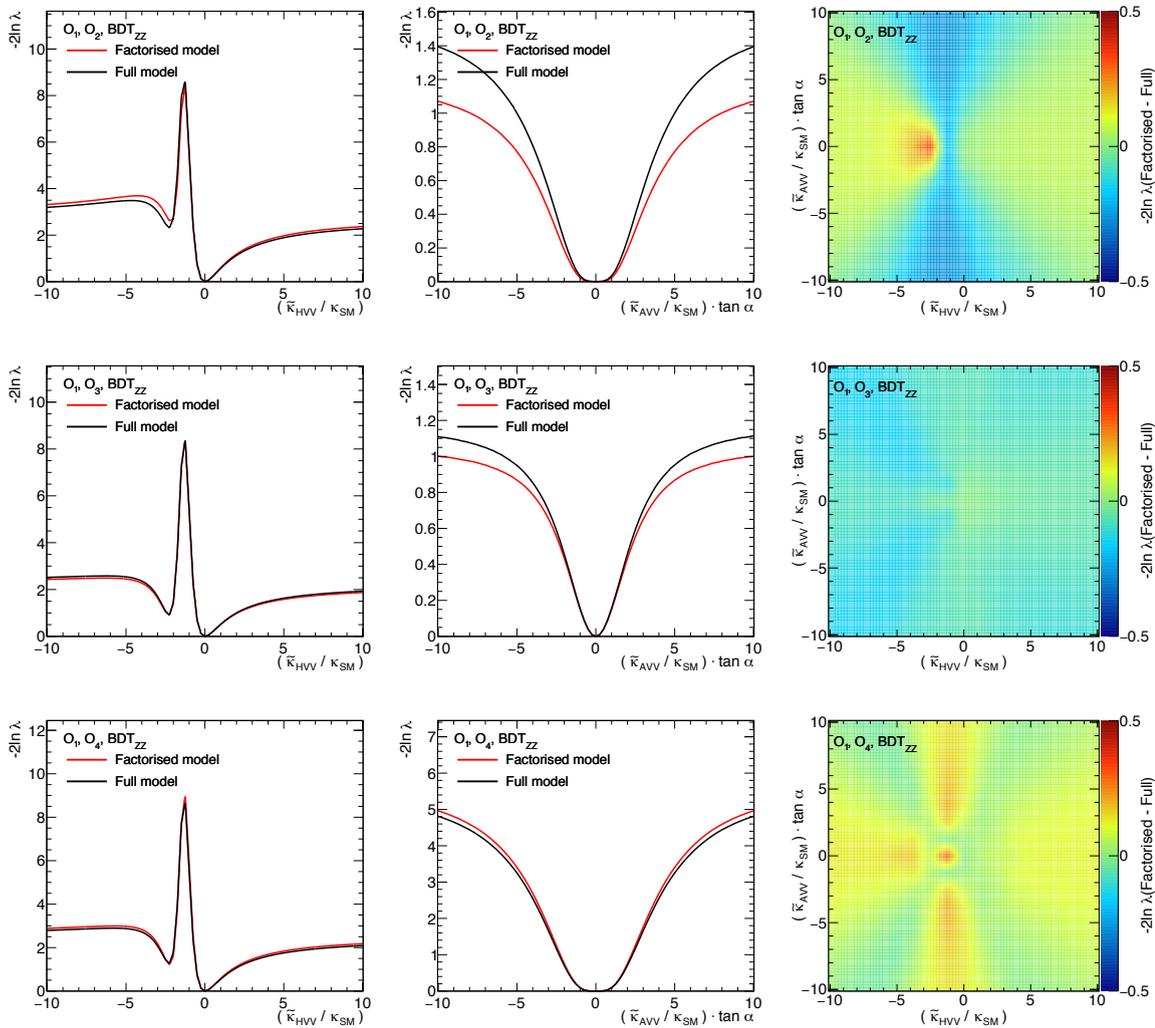


FIGURE D.15: Factorisation scheme test with decorrelated observables. The outline is similar to Figure 3.32. Correlation test for  $(O_1, O_2)$ ,  $(O_1, O_3)$ , and  $(O_1, O_4)$ .

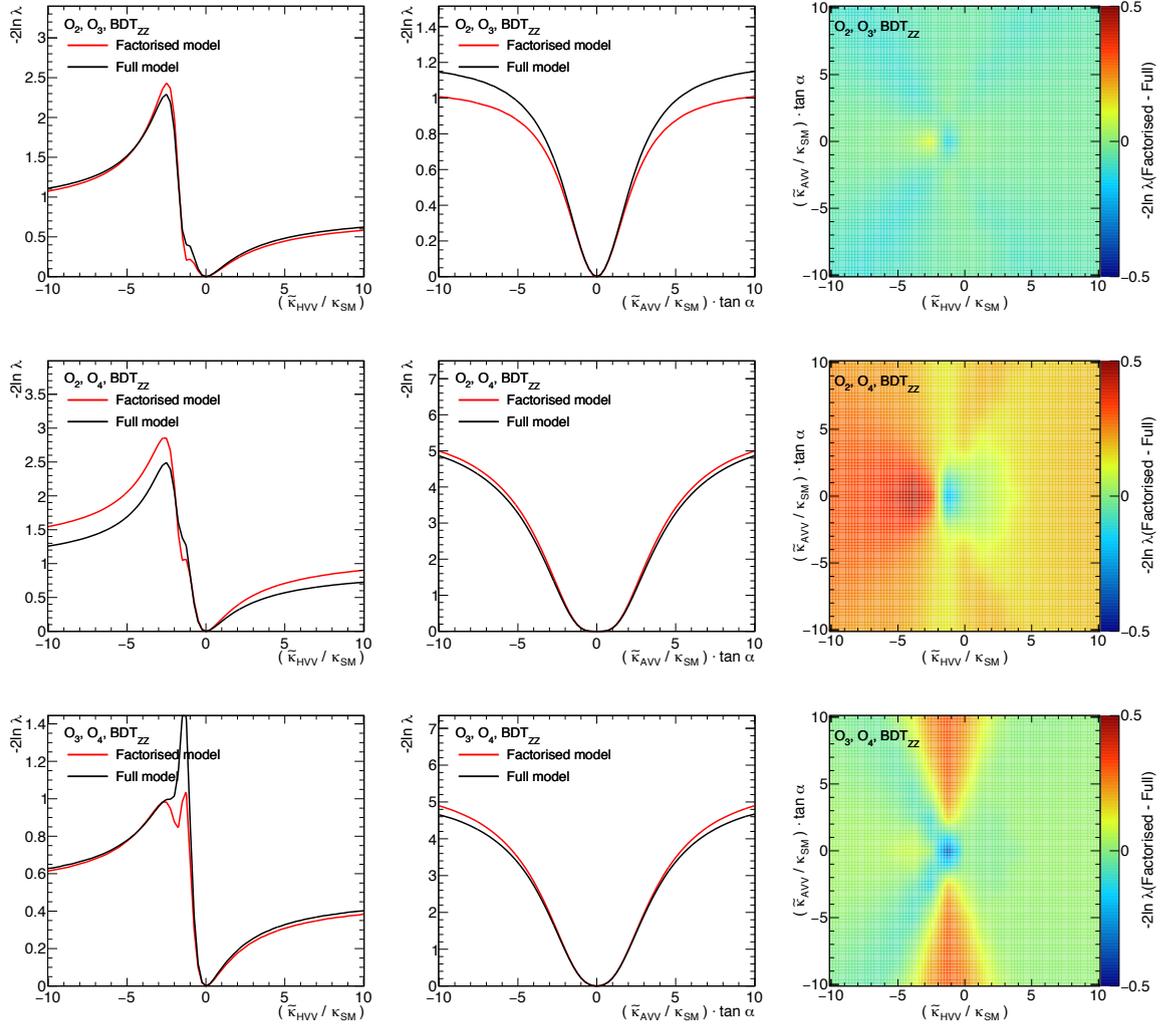


FIGURE D.16: Factorisation scheme test with decorrelated observables. The outline is similar to Figure 3.32. Correlation test for  $(O_2, O_3)$ ,  $(O_2, O_4)$ , and  $(O_3, O_4)$ .



# Bibliography

- [1] ATLAS Collaboration, *Observation of an excess of events in the search for the Standard Model Higgs boson in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel with the ATLAS detector*, ATLAS-CONF-2012-169 (2012), URL: <http://cds.cern.ch/record/1499628>.
- [2] ATLAS Collaboration, *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using  $25 \text{ fb}^{-1}$  of proton-proton collision data*, ATLAS-CONF-2013-013 (2013), URL: <http://cds.cern.ch/record/1523699>.
- [3] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, *Phys. Lett. B* **726**.1–3 (2013) 120–144, arXiv: 1307.1432 [hep-ex].
- [4] ATLAS Collaboration, *Prospects for measurements of tensor structure of the  $HZZ$  vertex in  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay with ATLAS detector*, ATL-PHYS-PUB-2013-013 (2013), URL: <http://cds.cern.ch/record/1611123>.
- [5] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, CERN-PH-EP-2015-114 (2015), URL: <http://cds.cern.ch/record/2026126>.
- [6] N. Belyaev et al., *Angular asymmetries as a probe for anomalous contributions to  $HZZ$  vertex at the LHC*, *Phys. Rev. D* **91** (11 2015) 115014, arXiv: 1502.03045 [hep-ph].
- [7] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, 1995, ISBN: 0-201-50307-2.
- [8] A. Pich, *The Standard Model of Electroweak Interactions*, Proceedings, 4th CERN-CLAF School of High-Energy Physics, CERN (2007), arXiv: 0705.4264 [hep-ph].
- [9] K. Olive et al., *Review of Particle Physics*, *Chin. Phys. C* **38** (2014) 090001, URL: <http://pdg.lbl.gov>.
- [10] MissMJ and E. Drexler, *Wikipedia - Standard Model*, retrieved Jun 19 2015”, URL: [http://en.wikipedia.org/wiki/Standard\\_Model](http://en.wikipedia.org/wiki/Standard_Model).

- [11] ATLAS Collaboration, *Search for excited leptons in proton-proton collisions at  $\sqrt{s}=7$  TeV with the ATLAS detector*, *Phys. Rev. D* **85** (7 2012) 072003, arXiv: [1201.3293 \[hep-ex\]](#).
- [12] ATLAS Collaboration and CMS Collaboration, *Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (19 2015) 191803, arXiv: [1503.07589 \[hep-ex\]](#).
- [13] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak and Heavy Flavour Groups, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (5–6 2006) 257–454, arXiv: [hep-ex/0509008 \[hep-ex\]](#).
- [14] G. Drexlin et al., *Current Direct Neutrino Mass Experiments*, *Adv. High Energy Phys.* **2013** (2013), Article ID 293986, arXiv: [1307.0101 \[physics.ins-det\]](#).
- [15] T. J. Loredo and D. Q. Lamb, *Bayesian analysis of neutrinos observed from supernova SN 1987A*, *Phys. Rev. D* **65** (6 2002) 063002, arXiv: [astro-ph/0107260 \[astro-ph\]](#).
- [16] G. Pagliaroli, F. Rossi-Torres and F. Vissani, *Neutrino mass bound in the standard scenario for supernova electronic antineutrino emission*, *Astropart. Phys.* **33** (5–6 2010) 287–291, arXiv: [1002.3349 \[hep-ph\]](#).
- [17] K. Abazajian et al., *Cosmological and astrophysical neutrino mass measurements*, *Astropart. Phys.* **35** (4 2011) 177–184, arXiv: [1103.5083 \[astro-ph.CO\]](#).
- [18] LHCb Collaboration, *Observation of the Resonant Character of the  $Z(4430)^-$  State*, *Phys. Rev. Lett.* **112** (22 2014) 222002, arXiv: [1404.1903 \[hep-ex\]](#).
- [19] LHCb Collaboration, *Observation of  $J/\psi p$  resonances consistent with pentaquark states in  $\Lambda_b^0 \rightarrow J/\psi K^- p$  decays*, *Phys. Rev. Lett.* **115** (7 2015) 072001, arXiv: [1507.03414 \[hep-ex\]](#).
- [20] ATLAS Collaboration, *Summary plots from the ATLAS Standard Model physics group*, retrieved Jun 18 2015, URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/>.
- [21] HepData team, *The Durham HepData Project*, retrieved Jun 18 2015, URL: <http://hepdata.cedar.ac.uk/pdf/pdf3.html>.
- [22] A. D. Martin et al., *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2 2009) 189–285, arXiv: [0901.0002 \[hep-ph\]](#).
- [23] A. D. Martin et al., *Extended Parameterisations for MSTW PDFs and their effect on Lepton Charge Asymmetry from W Decays*, *Eur. Phys. J. C* **73** (2 2013) 2318, arXiv: [1211.1215 \[hep-ph\]](#).

- [24] R. D. Ball et al., *Parton distributions with LHC data*,  
*Nucl. Phys. B* **867** (2 2013) 244–289, arXiv: 1207.1303 [hep-ph].
- [25] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*,  
*Phys. Rev. D* **89** (3 2014) 033009, arXiv: 1302.6246 [hep-ph].
- [26] David Z, *What's inside a proton?*, retrived July 25th 2015,  
URL: <http://physics.stackexchange.com>.
- [27] D. Green, *High  $P_T$  Physics at Hadron Colliders*, Cambridge University Press, 2005.
- [28] J. C. Collins and D. E. Soper,  
*Angular distribution of dileptons in high-energy hadron collisions*,  
*Phys. Rev. D* **16** (7 1977) 2219–2225, doi:10.1103/PhysRevD.16.2219.
- [29] J. Goldstone, *Field theories with “superconductor” solutions*,  
*Nuovo Cimento* **19** (1 1960) 154–164, URL: <http://cds.cern.ch/record/343400>.
- [30] D0 Collaboration,  
*Measurement of the Effective Weak Mixing Angle in  $p\bar{p} \rightarrow Z/\gamma^* \rightarrow e^+e^-$  Events*,  
*Phys. Rev. Lett.* **115** (4 2015) 041801, arXiv: 1408.5016 [hep-ex].
- [31] CDF Collaboration, *Indirect measurement of  $\sin^2\theta_W$  (or  $M_W$ ) using  $\mu^+\mu^-$  pairs from  $\gamma^*/Z$  bosons produced in  $p\bar{p}$  collisions at a center-of-momentum energy of 1.96 TeV*,  
*Phys. Rev. D* **89** (7 2014) 072005, arXiv: 1402.2239 [hep-ex].
- [32] LEP/TEVATRON Electroweak Working Group,  
*LEP/TEV EW WG Plots for Winter 2012*, retrieved Jun 21 2015,  
URL: <http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2012/>.
- [33] B. W. Lee, C. Quigg and H. B. Thacker,  
*Weak interactions at very high energies: The role of the Higgs-boson mass*,  
*Phys. Rev. D* **16** (5 1977) 1519–1531, doi:10.1103/PhysRevD.16.1519.
- [34] LHC Higgs Cross Section Working Group,  
*Higgs cross sections and branching ratios - Latest plots*, retrieved Jun 21 2015,  
URL: <http://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [35] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*,  
*Phys. Lett. B* **716** (1 2012) 1–29, arXiv: 1207.7214 [hep-ex].
- [36] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (1 2012) 30–61,  
arXiv: 1207.7235 [hep-ex].
- [37] ATLAS Collaboration,  
*Measurements of the Higgs boson production and decay rates and coupling strengths using  $pp$  collision data at  $\sqrt{s} = 7$  and 8 TeV in the ATLAS experiment*,  
ATLAS-CONF-2015-007 (2015), URL: <http://cds.cern.ch/record/2002212>.
- [38] ATLAS Collaboration, *Constraints on the off-shell Higgs boson signal strength in the high-mass  $ZZ$  and  $WW$  final states with the ATLAS detector*,  
*Eur. Phys. J. C* **75**, 335 (7 2015), arXiv: 1503.01060 [hep-ex].

- [39] CMS Collaboration, *Constraints on the Higgs boson width from off-shell production and decay to  $ZZ$  to  $llll$  and  $llvv$* , CMS-PAS-HIG-14-002 (2014), URL: <http://cds.cern.ch/record/1670066>.
- [40] V. Barger, M. Ishida and W.-Y. Keung, *Total Width of 125 GeV Higgs Boson*, *Phys. Rev. Lett.* **108** (26 2012) 261801, arXiv: 1203.3456 [hep-ph].
- [41] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, *JHEP* **2015**, 117 (4 2015), arXiv: 1501.04943 [hep-ex].
- [42] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of  $\tau$  leptons*, *JHEP* **2014**, 104 (5 2014), arXiv: 1401.5041 [hep-ex].
- [43] ATLAS Collaboration, *Search for the  $b\bar{b}$  decay of the Standard Model Higgs boson in associated ( $W/Z$ ) $H$  production with the ATLAS detector*, *JHEP* **2015**, 69 (1 2015), arXiv: 1409.6212 [hep-ex].
- [44] CMS Collaboration, *Search for the standard model Higgs boson produced in association with a  $W$  or a  $Z$  boson and decaying to bottom quarks*, *Phys. Rev. D* **89** (1 2014) 012003, arXiv: 1310.3687 [hep-ex].
- [45] P. Artoisenet et al., *A framework for Higgs characterisation*, *JHEP* **2013**, 43 (11 2013), arXiv: 1306.6464 [hep-ph].
- [46] I. Anderson et al., *Constraining anomalous HVV interactions at proton and lepton colliders*, *Phys. Rev. D* **89** (3 2014) 035007, arXiv: 1309.4819 [hep-ph].
- [47] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, CERN (2013), arXiv: 1307.1347 [hep-ph].
- [48] CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, *Phys. Rev. D* **92** (1 2015) 012004, arXiv: 1411.3441 [hep-ex].
- [49] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **2014**, 79 (7 2014), arXiv: 1405.0301 [hep-ph].
- [50] Y. Gao et al., *Spin determination of single-produced resonances at hadron colliders*, *Phys. Rev. D* **81** (7 2010) 075022, arXiv: 1001.3396 [hep-ph].
- [51] L. D. Landau, *The moment of a 2-photon system*, *Dokl. Akad. Nauk. USSR* **60** (1948) 207.
- [52] C. N. Yang, *Selection Rules for the Dematerialization of a Particle into Two Photons*, *Phys. Rev.* **77** (2 1950) 242–245, doi:10.1103/PhysRev.77.242.
- [53] S. Alioli et al., *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, *JHEP* **2009** (04 2009) 002, arXiv: 0812.0578 [hep-ph].

- [54] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **2006** (05 2006) 026, arXiv: [hep-ph/0603175](#) [[hep-ph](#)].
- [55] D. de Florian et al., *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, *JHEP* **2011**, 64 (11 2011), arXiv: [1109.2109](#) [[hep-ph](#)].
- [56] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*, *JHEP* **2013**, 129 (9 2013), arXiv: [1306.4581](#) [[hep-ph](#)].
- [57] T. Binoth, N. Kauer and P. Mertsch, *Gluon-induced QCD corrections to  $pp \rightarrow ZZ \rightarrow l\bar{l}l\bar{l}$*  (2008), arXiv: [0807.0024](#) [[hep-ph](#)].
- [58] J. Ellis, *Limits of the Standard Model* (2002), arXiv: [hep-ph/0211168](#) [[hep-ph](#)].
- [59] S. P. Martin, *A Supersymmetry primer*, *Adv. Ser. Direct. High Energy Phys.* **21** (2010) 1–153, arXiv: [hep-ph/9709356](#) [[hep-ph](#)].
- [60] Planck Collaboration, *Planck 2013 results. I. Overview of products and scientific results*, *Astron. Astrophys.* **571** (2014) A1, arXiv: [1303.5062](#) [[astro-ph.CO](#)].
- [61] V. C. Rubin and W. K. Ford, Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophys. J.* **159** (1970) 379–403, doi:[10.1086/150317](#).
- [62] M. Markevitch et al., *Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657–56*, *Astrophys. J.* **606** (2 2004) 819, arXiv: [astro-ph/0309303](#) [[astro-ph](#)].
- [63] N. Arkani-Hamed et al., *The Littlest Higgs*, *JHEP* **2002** (07 2002) 034, arXiv: [hep-ph/0206021](#) [[hep-ph](#)].
- [64] N. Arkani-Hamed, A. G. Cohen and H. Georgi, *Electroweak symmetry breaking from dimensional deconstruction*, *Phys. Lett. B* **513**.1–2 (2001) 232–240, arXiv: [hep-ph/0105239](#) [[hep-ph](#)].
- [65] Forthhommel, *Map of the CERN accelerator complex*, <http://en.wikipedia.org/wiki/CERN>, Source: CERN.
- [66] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (08 2008) S08005, URL: <http://cds.cern.ch/record/1129809>.
- [67] O. S. Brüning et al., *LHC Design Report*, CERN-2004-003-V-1, Geneva: CERN, 2004, URL: <http://cds.cern.ch/record/782076>.
- [68] R. Fernow, *Introduction to experimental particle physics*, Press Syndicate of the University of Cambridge, 1986, ISBN: 0-521-37940-7.

- [69] ATLAS Collaboration, *ATLAS EXPERIMENT - Public Results: Integrated luminosity summary plots for 2011-2012 data taking*, retrieved Jun 25 2015, URL: [http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Integrated\\_luminosity\\_summary\\_pl](http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Integrated_luminosity_summary_pl).
- [70] ATLAS Collaboration, *Measurement of the high-mass Drell–Yan differential cross-section in  $pp$  collisions at with the ATLAS detector*, *Phys. Lett. B* **725** (4–5 2013) 223–242, arXiv: 1305.4192 [hep-ex].
- [71] ATLAS Collaboration, *Improved luminosity determination in  $pp$  collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **73**, 2518 (8 2013), arXiv: 1302.4393 [hep-ex].
- [72] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042 (2012), URL: <http://cds.cern.ch/record/1435196>.
- [73] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, ATLAS-CONF-2014-047 (2014), URL: <http://cds.cern.ch/record/1741021>.
- [74] W. Leo, *Techniques for Nuclear and Particle Physics Experiments*, Springer-Verlag, 1994, ISBN: 3-540-57280-5.
- [75] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (08 2008) S08003, URL: <http://cds.cern.ch/record/1129811>.
- [76] P. Wagner, *Performance of the ATLAS Transition Radiation Tracker Readout with High Energy Collisions at the LHC*, ATL-INDET-PROC-2011-040 (2011), URL: <http://cds.cern.ch/record/1403142>.
- [77] ATLAS Collaboration, *Basic ATLAS TRT performance studies of Run 1*, ATL-INDET-PUB-2014-001 (2014), URL: <http://cds.cern.ch/record/1669603>.
- [78] ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*, *JINST* **3** (02 2008) P02013, URL: <http://cds.cern.ch/record/1094549>.
- [79] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, *Eur. Phys. J. C* **74**, 3130 (11 2014), arXiv: 1407.3935 [hep-ex].
- [80] ATLAS Collaboration, *ATLAS EXPERIMENT - Public Results: Trigger Operation Public Results*, retrieved Jun 18 2015, URL: <http://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults>.
- [81] ATLAS Collaboration, *An imaging algorithm for vertex reconstruction for ATLAS Run-2*, ATL-PHYS-PUB-2015-008 (2015), URL: <http://cds.cern.ch/record/2008700>.

- [82] S. Hageböck and E. von Toerne, *Medical Imaging Inspired Vertex Reconstruction at LHC*, *J. Phys. Conf. Ser.* **396** (2 2012) 022021, doi:10.1088/1742-6596/396/2/022021.
- [83] J. G. Colsher, *Fully-three-dimensional positron emission tomography*, *Phys. Med. Biol.* **25** (1 1980) 103, doi:10.1088/0031-9155/25/1/010.
- [84] M. Defrise, D. W. Townsend and R. Clack, *Three-dimensional image reconstruction from complete projections*, *Phys. Med. Biol.* **34** (5 1989) 573, doi:10.1088/0031-9155/34/5/002.
- [85] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, *Phys. Rev. D* **91** (1 2015) 012006, arXiv: 1408.5191 [hep-ex].
- [86] ATLAS Collaboration, *ATLAS EXPERIMENT - Public Results: Electron Efficiency Measurements for 2012 and 2011 Data*, retrieved Jun 25 2015, URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20130926/ATL-COM-PHYS-2013-1287/index.html>.
- [87] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the decay channel  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  with  $4.8 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 7 \text{ TeV}$* , ATLAS-CONF-2011-162 (2011), URL: <http://cds.cern.ch/record/1406357>.
- [88] ATLAS Collaboration, *Measurement of the Higgs boson mass from the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, *Phys. Rev. D* **90** (5 2014) 052004, arXiv: 1406.3827 [hep-ex].
- [89] P. Speckmayer et al., *The toolkit for multivariate data analysis, TMVA 4*, *J. Phys. Conf. Ser.* **219** (3 2010) 032057, doi:10.1088/1742-6596/219/3/032057.
- [90] K. Cranmer, *Kernel estimation in high-energy physics*, *Comput. Phys. Commun.* **136** (3 2001) 198–207, arXiv: hep-ex/0011057 [hep-ex].
- [91] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form factors of the top quark via  $e^+e^- \rightarrow t\bar{t}$* , *Phys. Rev. D* **45** (7 1992) 2405–2413, doi:10.1103/PhysRevD.45.2405.
- [92] M. Davier et al., *The optimal method for the measurement of tau polarization*, *Phys. Lett. B* **306** (3–4 1993) 411–417, doi:10.1016/0370-2693(93)90101-M.
- [93] M. Diehl, O. Nachtmann and F. Nagel, *Triple gauge couplings in polarised  $e^-e^+ \rightarrow W^-W^+$  and their measurement using optimal observables*, *Eur. Phys. J. C* **27** (3 2003) 375–397, arXiv: hep-ph/0209229 [hep-ph].
- [94] K. Cranmer et al., *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016 (2012), URL: <http://cds.cern.ch/record/1456844>.

- [95] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71**, 1554 (2 2011), Erratum: *Eur. Phys. J. C* **73**, 2501 (07 2013), arXiv: 1007.1727 [physics.data-an].
- [96] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (1 1938) 60–62, doi:10.1214/aoms/1177732360.
- [97] ATLAS Collaboration, *Determination of spin and parity of the Higgs boson in the  $WW^* \rightarrow e\nu\mu\nu$  decay channel with the ATLAS detector*, *Eur. Phys. J. C* **75**, 231 (5 2015), arXiv: 1503.03643 [hep-ex].
- [98] ATLAS Collaboration, *Performance assumptions based on full simulation for an upgraded ATLAS detector at a High-Luminosity LHC*, ATL-PHYS-PUB-2013-009 (2013), URL: <http://cds.cern.ch/record/1604420>.
- [99] J. Conway et al., *PGS - Pretty Good Simulation of high energy collisions*, URL: <http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm>.
- [100] R. M. Godbole, D. J. Miller and M. M. Mühlleitner, *Aspects of CP violation in the HZZ coupling at the LHC*, *JHEP* **2007** (12 2007) 031, arXiv: 0708.0458 [hep-ph].
- [101] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, *Phil. Trans. Roy. Soc. Lond. A* **231** (1933) 289–337, URL: <http://www.jstor.org/stable/91247>.
- [102] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, *Phys. Rev. D* **57** (7 1998) 3873–3889, arXiv: physics/9711021 [physics.data-an].
- [103] M. P. Wand, *Fast Computation of Multivariate Kernel Estimators*, *J. Comput. Graph. Stat.* **3** (4 1994) 433–445, doi:10.1080/10618600.1994.10474656.

# List of Figures

1.1	Physics cube	2
1.2	Leading order contributing processes of $q\bar{q} \rightarrow ZZ$	10
1.3	Leading order contributions processes of $gg \rightarrow ZZ$	11
1.4	Standard Model particles and interactions	12
1.5	ATLAS measurement of Standard Model processes cross sections	15
1.6	Parton density function	16
1.7	Parton density function	17
1.8	Decay diagram of the $H \rightarrow ZZ^* \rightarrow 4\ell$ final state	19
1.9	Propagator loop correction	22
1.10	Indirect Higgs mass predictions	23
1.11	Higgs Production Cross Sections	25
1.12	Measured Higgs Coupling Summary	26
1.13	Angular Distributions of spin-0 models	30
1.14	Conversion between coupling ratio to cross section fraction	31
1.15	Conversion between cross section fractions and coupling ratios	32
1.16	Final state distributions for spin-2 models with different coupling configurations	34
1.17	Standard Model reweighting validation	38
1.18	BSM reweighting validation	39
2.1	CERN complex	44
2.2	LHC delivered instantaneous luminosity, mean number of interactions and ATLAS recorded efficiency	48
2.3	ATLAS particle identification scheme	50
2.4	ATLAS coordinate frame	51
2.5	ATLAS Inner Detector	52
2.6	ATLAS TRT electron identification	54
2.7	ATLAS calorimeters	57
2.8	Radiation length of the ATLAS calorimeters	57
2.9	ATLAS Muonic Spectrometer	58
2.10	L1 and EF trigger rates	60
2.11	Track and filtered track projections	63
2.12	Primary vertex merge probability and reconstruction efficiency	63
2.13	ATLAS electron and muon reconstruction efficiencies	66
3.1	$H \rightarrow ZZ^* \rightarrow 2e2\mu$ event display	70
3.2	$H \rightarrow ZZ^* \rightarrow 4\mu$ event display	71

3.3	Four lepton mass distribution	73
3.4	Expected and observed production distributions	77
3.5	Expected and observed angular distributions	78
3.6	BDT Discriminants for examined models	83
3.7	$m_{34}$ distribution and spin-2 $q\bar{q}$ scan	91
3.8	BDT discriminant and test statistic distributions	92
3.9	Angular distributions sensitive to $HZZ$ tensor structure	95
3.10	Matrix Element Observables	97
3.11	Validation of transformation method	99
3.12	Two dimensional observable distributions	100
3.13	Reweighting validation for $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ observables	101
3.14	Kernel Density Estimation validation for $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ observables	103
3.15	Tensor structure fit observables and background discriminant	105
3.16	Expected Standard Model Results with nominal parameters	106
3.17	Maximum likelihood estimator distributions	107
3.18	Sampled $-2 \ln \lambda$ distributions	108
3.19	Sampled pull distributions	110
3.20	Closure tests	110
3.21	Background discriminant distributions for different mass models	111
3.22	Reducible background shape uncertainty	112
3.23	Muon spectrometer resolution uncertainty	114
3.24	$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ nuisance parameter profiles	115
3.25	$\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ nuisance parameter profiles	116
3.26	Compatibility checks between Matrix Element Observable fit and Nine Dimensional Fit	118
3.27	Expected and observed results of the tensor structure analysis	121
3.28	BDT discriminants used in the $WW^*$ analysis	123
3.29	Simultaneous fit of $\frac{\tilde{\kappa}_{HZZ}}{\kappa_{SM}}$ , $\frac{\tilde{\kappa}_{HWW}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AZZ}}{\kappa_{SM}} \tan \alpha$ , $\frac{\tilde{\kappa}_{AWW}}{\kappa_{SM}} \tan \alpha$	125
3.30	Expected and observed results of the combined tensor structure analysis	126
3.31	Correlation between $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$ observables	128
3.32	Factorisation scheme test	129
3.33	Simultaneous fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	132
3.34	Translation of results to $f_{gi}$ -scheme	133
3.35	High Luminosity BSM closure	137
3.36	High Luminosity SM fit	138
3.37	$O_4$ distribution and log likelihood fit	140
3.38	Exclusion regions from $O_4$ discriminant	140
4.1	Simultaneous fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	145
A.2	An example model	149
A.3	Simple hypothesis test statistic distributions	151
A.4	Coverage belts of different test statistics	153
A.5	Profiled likelihood test statistic distribution and $-2 \ln \lambda$ curve	154

A.6	Example use of Kernel Density Estimator	158
A.7	Evaluation time of KDE methods as function of sample size	159
A.8	<i>pdf</i> transformation example	161
A.9	Decision tree	163
A.10	BDT classification visual	164
C.11	Two dimensional observable distributions	166
C.12	Rewighting validation for $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ $\tan \alpha$ observables	166
C.13	Kernel Density Estimation validation for $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ $\tan \alpha$ observables	167
D.14	Two dimensional projections of decorrelated observables	171
D.15	Factorisation scheme test with decorrelated observables - I	172
D.16	Factorisation scheme test with decorrelated observables - II	173

## List of Tables

1.1	Particles and properties of the Standard Model	13
1.2	Summary table of spin-0 models	30
1.3	Summary table of Spin-2 models	33
1.4	Monte Carlo samples and generators	36
2.1	ATLAS subdetectors	50
2.2	Geometry of the inner detector	55
2.3	Coverage of the muon spectrometer	59
2.4	Trigger menu for $H \rightarrow ZZ^* \rightarrow 4\ell$ candidate selection	65
2.5	Electron ID Criteria	67
2.6	Inner detector requirements for muons	68
3.1	Quadruple Selection	75
3.2	Signal cross sections and efficiencies	76
3.3	Event yield expectations	77
3.4	BDT training summary	82
3.5	BDT training settings and performance	84
3.6	Dominant systematic uncertainties for hypothesis test analysis	86
3.7	Expected and observed hypothesis test results	89
3.8	Expected and observed hypothesis test results	91
3.9	Expected best-fit values of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ $\tan \alpha$ and 95% CL excluded regions	106
3.10	Dominant Systematic uncertainties for Tensor structure fits	113
3.11	Expected and observed best-fit values of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ $\tan \alpha$ and 95% CL excluded regions	120
3.12	Expected and observed best-fit values of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}}$ $\tan \alpha$ and 95% CL excluded regions	125

3.13	Simultaneous fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	131
3.14	Simultaneous fit of $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$ and $\frac{\tilde{\kappa}_{AVV}}{\kappa_{SM}} \tan \alpha$	131
3.15	High Luminosity event yield expectations	135
3.16	Assumed systematic uncertainties for high luminosity tensor structure fits	135
3.17	Expected high luminosity exclusion regions	138
3.18	Expected high luminosity exclusion regions using angular asymmetries	141
B.1	Full list of hypothesis test systematics	165
C.2	Full list of tensor structure fit systematics, $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$	168
C.3	Full list of tensor structure fit systematics, $\frac{\tilde{\kappa}_{HVV}}{\kappa_{SM}}$	169

# Abbreviations

9DMEM	Nine Dimensional Matrix Element Method (fit based directly on final state distributions)
BDT	Boosted Decision Tree
BSM	Beyond the Standard Model
CL	Confidence Level
EFT	Effective Field Theory
EW	Electro Weak
FS	Final State
KD	Kinematic Discriminant (ratio of matrix elements of $H$ and $ZZ^*$ )
KDE	Kernel Density Estimator (Estimation)
LHC	Large Hadron Collider
MC	Monte Carlo
ME	Matrix Element
ME-Obs	Matrix Element Observable
SM	Standard Model
SUSY	Super Symmetry (Symmetric)
PDF	Parton Distribution Function
<i>pdf</i>	Probability Density Function
QCD	Quantum Chromo Dynamics
QED	Quantum Electro Dynamics
ROC	Receiver Operating Characteristic

# Symbols

## Model parameters

$VV$	Used as substitute for $ZZ$ and $WW$
$k_{HVV}$	Coupling of CP-even BSM Spin-0 Model
$k_{AVV}$	Coupling of CP-odd BSM spin-0 Model
$\tilde{k}_{HVV}$	$k_{HVV}$ scaled by $\nu/4\Lambda$
$\tilde{k}_{AVV}$	$k_{AVV}$ scaled by $\nu/4\Lambda$
$\frac{\tilde{k}_{HVV}}{\kappa_{SM}}$	Parameterisation used to describe CP-even BSM to SM coupling ratios
$\frac{\tilde{k}_{AVV}}{\kappa_{SM}} \tan \alpha$	Parameterisation used to describe CP-odd BSM to SM coupling ratios
$0^+$	The Standard Model Higgs
$0_h^+$	Spin-0 Model obtained by having $k_{HVV}$ as the only non-zero coupling
$0^-$	Spin-0 Model obtained by having $k_{AVV}$ as the only non-zero coupling
$\alpha$	Mixing angle between CP-even and CP-odd sectors of Lagrangian
$\Re(g_2)/g_1$	Equivalent parameterisation to $\frac{\tilde{k}_{HVV}}{\kappa_{SM}}$
$\Re(g_4)/g_1$	Equivalent parameterisation to $\frac{\tilde{k}_{AVV}}{\kappa_{SM}} \tan \alpha$
$f_{g2}$	Effective contribution of $0_h^+$ to the combined cross section
$\cos \phi_{g2}$	Sign of BSM coupling ( $\kappa_{HVV}$ ) in $f_{g2}$
$f_{g4}$	Effective contribution of $0^-$ to the combined cross section
$\cos \phi_{g4}$	Sign of BSM coupling ( $\kappa_{AVV}$ ) in $f_{g4}$

## Observables

$O_1(\tilde{k}_{HVV}), O_2(\tilde{k}_{HVV})$	Matrix Element observables that are used to measure $\frac{\tilde{k}_{HVV}}{\kappa_{SM}}$
$O_1(\tilde{k}_{AVV}, \alpha), O_2(\tilde{k}_{AVV}, \alpha)$	Matrix Element observables that are used to measure $\frac{\tilde{k}_{AVV}}{\kappa_{SM}} \tan \alpha$
$TO(*)$	Transformed Matrix Element Observable
$BDT_{ZZ}$	BDT trained to separate a SM Higgs from the $ZZ^*$ continuum

## Miscellaneous

<code>alpha_ATLAS_*</code>	Systematic uncertainty
<code>mu_2011, mu_2012</code>	Signal strength observed in the 2011 and 2012 sample