Discovery of the Higgs Boson, Measurements of its Production, and a Search for Higgs Boson Pair Production

James C. Saxon

a dissertation
in
Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
2014

Hugh H. Williams, Professor, Physics
Supervisor of Dissertation

Marija Drndic, Professor, Physics
Graduate Group Chairperson

Dissertation Committee
Eugene Beier, Professor, Physics
I. Joseph Kroll, Professor, Physics
Elliot Lipeles, Associate Professor, Physics
Mark Trodden, Professor, Physics
Hugh H. Williams, Professor, Physics

CERN-THESIS-2014-084
13/06/2014

Discovery of the Higgs Boson,
Measurements of its Production,
and a Search for Higgs Boson Pair Production

copyright
2014
James C. Saxon

Acknowledgements

I began working on the ATLAS experiment for the University of Pennsylvania more than eleven years ago, at the tender age of fifteen. My time on ATLAS therefore coincides not only with my formation as a physicist but also very substantially with my development as a person. I am overwhelmed to reflect on the impact that the Penn group and the ATLAS collaboration have had on me. I am grateful for the camaraderie of barbecues, dinners, Bond flics, and discussions that continued through years. I am grateful for the technical skills, the intellectual challenges, and the infectious passion that so many people have shared with me. To all those who supervised me, helped me, and taught me – thank you for inviting me to join in this adventure.

Rick Van Berg gave me my very first job at Penn, and our innumerable discussions over the intervening decade have enriched me both intellectually and personally. Brig Williams is a true gentleman – sharp as a tack, but gracious and genuinely kind. He has cultivated the Penn group with wisdom and tenacity. He has been unfailingly accessible and supportive, and led with high expectations and a long leash. I could not have asked for a better advisor. All of the professors in the ATLAS group – Elliot Lipeles, Joe Kroll, and Evelyn Thomson – assisted in supervising me, and it was a blessing to have such a collaborative team. I am also grateful to the non-ATLAS reviewers of my thesis – Mark Trodden and Gene Beier.

Mike Hance, John Alison, and Dominick Olivito set the bar for Penn students at CERN – and they set it high. Mike was incredibly generous with his phenomenal expertise with the TRT and with photons. John defined, more than anyone, the *esprit de corps* that made the group so successful. He is one of the most intense intellectuals that I have ever met, and I look forward to our continued collaboration in Chicago. Dominick turned the TRT into one of the smoothest functioning detectors on ATLAS. He personally watched over my development and put up with me with greater patience than anyone should have had to – in commissioning the TRT, refining the DAQ, and preparing physics analysis.

Special thanks goes to my other house-mates who magnanimously put up with 'a boy singing': Kurt Brendlinger and Alex Tuna at Taney, and Chris Lester and Jon Stahlman (formerly of Taney). Kurt has been a great, long-suffering friend since about the seventh grade, and hopefully for a long time to come (though perhaps with less suffering). He was also an invaluable homework buddy during my first two years at Penn. Tuna sprinted through his classes but was always happy to help me slug through. My life would have fewer cats without him. Lester is a mensch and a great thinker, and helped me hold on to scraps of sanity in the months that we overlapped at Hautains. Stahlman bore a great piece of the responsibility for the TRT during my time at CERN, and was a stalwart champion of the BBQs for several years.

The Penn Army (and honorary Penn Army) was a dynamic, fun-loving, hard-working crew, I am much the richer for having grown up with all of you, and I consider myself very fortunate. You have been incredible colleagues and wonderful friends. I look forward to watching the group continue to evolve and grow – both as current members move on to new posts and projects, and as rising students take the group to ever greater heights.

My time on the TRT stretched over its construction, commissioning, and Run I operations. Ole Røhne,

Franck Martin, Jack Fowler, Ben LeGeyt (and of course, Mike and Dominick) ensured the successful completion of the detector, and watched over me in my first summers at CERN. Christoph Rembser is a model of good leadership in large collaborations, shouldering bureaucracies, supporting a strong community, and responding to every request with great humour. Anatoli Romaniouk's long-term care for details has been fundamental to achieving and maintaining the high performance of the TRT. Andrey Loginov was always there to help clear the way for us to 'just work.' Zbyszek Hajduk, Jola Olszowska, Elzbieta Banas of the Krakow group kept the TRT electronics running and the detector safe, and supported us every step of the way. Of my home in the DAQ crew, it was a great privilege to work with Mike, Dominick, Peter Wagner, Jon Stahlman, Sarah Heim, Ximo Poveda, Paul Keener, Mitch Newcomer, Godwin Mayers, and Rick Van Berg – and now Chris Meyer, Khilesh Mistry, Bijan Haney, Tony Thompson and Leigh Schaefer.

My first analysis work was in the e/γ group. Fabrice Hubaut, Andrea Bocci, Marco Delmastro, and Marcos Jimenez provided good-natured and level-headed leadership and taught me everything that I know about photons (that I didn't learn from Mike and Brig!). It was a pleasure for me to work with Kun Liu and Evgeny Soldatov on the efficiency measurements in 2011 and 2012. Leonardo Carminati was a great ally in preparing cuts for 2012 data.

The $h \to \gamma\gamma$ working group managed to be a functional, pleasant environment for a young student, despite the enormous pressures at the time of the discovery. The group conveners played an important role in maintaining this good environment, and helped to make my time in the group exciting and enjoyable. Marc Escalier did a phenomenal job, defining the analysis in very early data, and shepherding it through an 'exciting' moment in April 2011. I am very grateful to Junichi Tanaka, Kerstin Tackmann, Krisztian Peters, Nicolas Berger, and Sandrine Laplace, who allowed me opportunities to stretch my wings and really learn to do analysis. Nicolas helped me with patience, kindness, and energy well beyond any reasonable expectations, to understand statistics.

The Orsay group gave me a home away from home in 2013, and afforded me the detailed, expert attention that has made their own group such a powerhouse. My especial thanks go to Louis Fayard, Estelle Scifo, Marc Escalier, and Nansi Andari, who welcomed me several times at LAL. The Wisconsin group were also frequent and diligent collaborators; I particularly want to thank Andrew Hard, Hongtao Yang, and Haichen Wang. In studying missing energy and preparing couplings analyses, Bruce Melado, Amanda Kruse, and Elisabeth Petit were a pleasure to work with. Within the missing energy group, my thanks go to the Milan team, and to Irene Vichou who made the $\not{E}_T$ work, and answered my endless questions to make sure that the definitions used for $h \to \gamma\gamma$ were correct and optimal.

One of my greatest pleasures in my time on ATLAS was working with Dag Gillberg and Florian Bernlochner on the differential cross sections conference note. They are fast, smart, insightful, fun, and basically just the best collaborators I could imagine working with. My thanks go to the full differential cross sections team, that produced a beautiful result on a very short time frame, and to our editorial board (Fabrice and Mike, again, as well as Mark Sutton) who helped us to get it out in a timely manner.

Turning to the measurement of Higgs boson pair production, I worked with an all-star team from Yale – Andrey Loginov, Johannes Erdmann, and Jahred Adelman. In particular, I worked very closely with Jahred for nearly a year, and he was always efficient and thoughtful – both in thinking creatively about the analysis and producing solid results, and also in shepherding it through the (sometimes sticky) politics of the ATLAS collaboration. Jahred taught me a great deal, and it was an abiding pleasure to work and chat with him; I look forward to our continued collaboration in Illinois. Leandro Nisati and our entire editorial board mustered quite a force in reviewing the analysis and pushing it through to approval.

Among physicists, my last thanks go Chicago, for inviting me to continue with them on the next leg of this journey. It is an awesome team that I am extremely excited to join!

My parents granted me tremendous independence at a young age, but supported me when I needed them, and helped me to think through and deal with my decisions. They set for me a rich example of deliberate, practical morality and decency.

My final thanks go to my Renata. I love you; I am yours.

ABSTRACT

Discovery of the Higgs Boson, Measurements of its Production, and a Search for Higgs Boson Pair Production

James C. Saxon

Hugh H. Williams

This document chronicles the discovery of the Higgs boson and early measurements in the diphoton decay channel. Particular attention is paid to photon identification, to the coupling of the Higgs to the vector bosons, and to differential cross sections of the Higgs boson. As these measurements yielded good agreement to the predictions of the Standard Model, an additional search is performed, for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state. The dataset used represents 5 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV and 20 fb^{-1} of collisions at $\sqrt{s} = 8$ TeV, recorded by the ATLAS experiment at CERN's Large Hadron Collider in 2011 and 2012.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Standard Model of particle physics is a testament to the intensity of human curiosity and an intellectual triumph. It describes three of the four known forces with astounding accuracy; it integrates a century of experiments within an alternately elegant and ad hoc theory. Yet for five decades its keystone was absent. This document describes the crowning achievement of the Standard Model: the experimental observation of the Higgs boson.

In March 1984, the European Organization for Nuclear Research (CERN) and the European Committee for Future Accelerators (ECFA) hosted a workshop in Lausanne, to consider proposals for a 'Large Hadron Collider in the LEP Tunnel.' Twenty years had already passed since the Higgs mechanism was first described [1–3]. The Standard Model described by Glashow, Weinberg, and Salam had just been vindicated by the discovery of the W and Z vector bosons by the UA1 and UA2 experiments at the Super Proton Synchrotron [4–7]. The participants in the 1984 workshop concluded that 'searching for the Higgs meson [sic.] as it appears in the standard model looks difficult,' and in particular that gluon fusion 'does not seem to be a promising mechanism' [8]. On the other hand, supersymmetry was expected to be readily accessible.

By 1992 however, when the ATLAS and CMS Collaborations submitted Letters of Intent for large, multi-purpose detectors on the LHC [9, 10], the detector design parameters and the physics case of the two experiments had swung decidedly in favor of the discovery potential of the Higgs boson in the diphoton and ZZ channels. The two experiments were approved in early 1997, and civil engineering work began shortly thereafter. The ATLAS prototype detectors were first assembled in a 2003 test beam, and the experimental cavern was completed that year, 100 m below the Franco-Swiss border. Installation of the 7000-ton, 13-story detector would continue for the next four years. LHC collisions began in September 2008, but abruptly ended nine days later, when a faulty interconnect between two superconducting magnets caused a quench and an explosion that destroyed several magnets [11]. Data collection recommenced in 2009 and accelerated through 2011 and 2012, leading to the discovery of the Higgs boson announced on July 4, 2012.

The diphoton and ZZ decay channels were the two components to the initial Higgs observation by ATLAS.[1] The diphoton channel remains one of the most-important modes for the study of its couplings. Its small branching fraction of just 0.2% is offset by a high selection efficiency of ∼40%, good mass resolution, and a relatively straightforward analysis strategy. This document chronicles the discovery of the Higgs boson in this channel, and describes early measurements of its interactions. It concentrates, naturally, on those topics to which the author made meaningful contributions:

- ▷ Chapters 2 and 3 briefly describe the theoretical context for the work, and present the Large Hadron Collider and the ATLAS detector.

[1] The $h \to WW$ channel followed a week or so later.

- ▷ Chapter 4 describes techniques for identifying photons showers and rejecting QCD backgrounds. These methods were central to the Higgs discovery by ATLAS, which is discussed in Chapter 5. That work was first published in Ref. [12].
- ▷ Following the discovery of a new particle, the task was to determine whether it behaved as the SM Higgs boson. Two paths were open: (1) to check that the particle was indeed a scalar, or (2) to measure whether its interactions with other particles scaled with their masses. The author chose the second path. Chapter 6 therefore describes efforts to select events consistent with Higgs bosons produced in association with a W or Z boson, and thereby isolate production through the WWh and ZZh vertices. This work was an important part of the early ATLAS paper on the Higgs couplings [13].
- ▷ As the early coupling and spin results agreed with the predictions of the SM within experimental error, differential cross sections of Higgs boson production were extracted, to further assess the compatibility of the new particle to the SM Higgs boson. These measurements are described in Chapter 7 based on work first released in a conference note [14], for which the author served as an editor.
- ▷ A search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ channel, is described in Chapter 8. The cross section for this process is exceedingly small in the SM, but it is enhanced, resonantly or non-resonantly, in a number of extensions to it. No significant discrepancy was found from the null hypothesis, but this channel remains promising for Run II. A paper has been released, also co-edited by the author [15].

Complementary measurements of mass, spin, and parity are summarized briefly, to present a fuller picture of the consistency of the new particle to the SM Higgs boson. Measurements from and combinations with other channels (WW, ZZ, $b\bar{b}$, $\tau\tau$) are described where appropriate.

Chapter 2

Theoretical Context

A complete presentation of the Standard Model is beyond the scope of this document, and has long been available from better-qualified authors [16,17]. The few details necessary to motivate this work are included for completeness.

2.1 The Standard Model

The Standard Model of particle physics (SM) is a theoretical description of the elementary particles that make up matter, and three of the four forces that govern their interactions. There are three families of quarks and leptons, each with a left-handed doublet and two right-handed singlets:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \; u_R, \; d_R \qquad \text{and} \qquad \begin{pmatrix} \nu \\ e^- \end{pmatrix}_L, \; e^-_R, \; \nu_R \tag{2.1}$$

(The right-handed neutrinos need not exist, but appear in some models of neutrino masses.) Each of the three forces is mediated by spin-1 bosons. The strong force (Quantum Chromodynamics, QCD) is mediated by an octet of colored gluons, electromagnetism (Quantum Electrodynamics, QED) by the photon, and the weak force by the $W^{\pm}$ and Z^0 (after symmetry breaking). Together, these forces and particles describe all familiar phenomena aside gravity.

Local gauge invariance is required for the renormalizability of the theory. The SM Lagrangian is invariant under local transformations of the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$, where c refers to color, L denotes the handedness of the $SU(2)$ coupling, and Y refers to the weak hypercharge. Local gauge invariance is preserved through the introduction of a covariant derivative, which in turn implies the existence of the spin-1 gauge bosons. The structure functions of the underlying Lie groups (the Lie Algebra), determine the form of the covariant derivative and the properties of the gauge interaction.

The SM Lagrangian may be factored into four pieces:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_f + \mathcal{L}_\varphi + \mathcal{L}_{\text{Yukawa}} \, . \tag{2.2}$$

'Gauge' includes the kinetic energies and (self-)interactions of the gauge bosons. $\mathcal{L}_f$ includes kinetic energies of the fermions and, through the covariant derivative, their interactions with the gauge bosons. Because the left-handed and right-handed fermions transform differently under $SU(2)_L$, it is impossible to directly write mass terms for them. Mass terms for the gauge bosons like $m^2 G_\mu G^\mu$ are not gauge invariant, and are similarly absent at this stage.

The bulk of this thesis focuses on tests of the second two terms in Equation 2.2: the Higgs kinetic terms and potential, $\mathcal{L}_\varphi$, and its interactions with the fermions, $\mathcal{L}_{\text{Yukawa}}$. Without these terms, the weak interaction would have infinite range, the fermions would be massless, and there would be no Higgs boson.

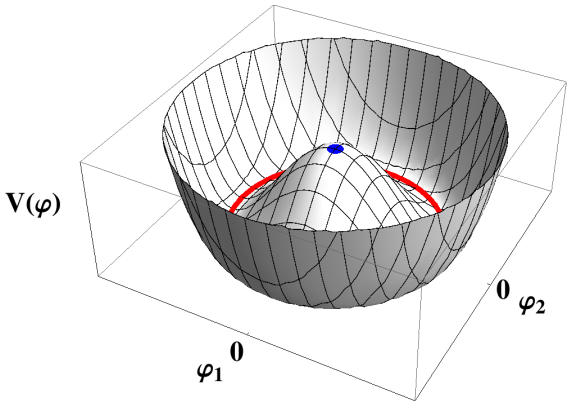


Figure 2.1: The 'Mexican hat' potential canonically illustrates the principle of a spontaneously broken symmetry.

2.2 The Higgs Mechanism

The ground state of a physical system need not preserve its intrinsic symmetries. For instance, a heated magnet is rotationally symmetric, but 'chooses' a preferred direction when cool. In the canonical example, the 'Mexican hat potential' of a complex field φ

$$V(\varphi) = \mu^2 \left(\varphi^\dagger \varphi\right) + \lambda \left(\varphi^\dagger \varphi\right)^2 \tag{2.3}$$

is rotationally symmetric, but for $\mu^2 < 0$ and $\lambda > 0$, its minimum lies in circle around the origin, with radius $\nu = \sqrt{-\mu^2/\lambda}$ (Figure 2.1). The ground state spontaneously breaks the rotational symmetry, by choosing a point on that circle.

The Higgs mechanism uses spontaneous symmetry breaking to introduce both gauge boson and fermion masses while respecting gauge invariance [1–3]. This is accomplished through the introduction of a doublet of scalar fields φ that transforms under $SU(2)_L \times U(1)_Y$, along with the (initially) symmetric potential $V(\varphi)$ (Equation 2.3). The Lagrangian $\mathcal{L}_\varphi$ takes the form

$$\mathcal{L}_\varphi = \left(D_\mu \varphi\right)^\dagger \left(D_\mu \varphi\right) - \mu^2 \left(\varphi^\dagger \varphi\right) - \lambda \left(\varphi^\dagger \varphi\right)^2 , \tag{2.4}$$

where the covariant derivative is $\left(D_\mu \varphi\right) = \left(\partial_\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}_\mu/2 + ig' B_\mu/2\right)\varphi$ and W^i and B are the gauge bosons associated to $SU(2)_L$ and $U(1)_Y$ respectively. The minimum of the potential is $\nu = \sqrt{-\mu^2/\lambda}$, and is chosen to fall along the real part of the lower component of φ:

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix} \qquad \Longrightarrow \qquad \varphi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} . \tag{2.5}$$

Using $\varphi \to \nu + h$ to expand Equation 2.4 around this minimum, and reorganizing the W^i_μ and B_μ as the familiar $W^\pm$ and Z^0, yields

$$\mathcal{L}_\varphi = \left(\partial_\mu h\right)^2/2 + m_W^2 W_\mu^+ W^{-\mu} \left(1 + \frac{h}{v}\right)^2 + m_Z^2 Z_\mu Z^\mu \left(1 + \frac{h}{v}\right)^2 + \frac{\mu^4}{4\lambda} + \mu^2 h^2 + \lambda\nu h^3 + \frac{\lambda}{4} h^4 . \tag{2.6}$$

The gauge boson masses thus appear as couplings to the condensate of the Higgs potential. The coupling of the Higgs boson with the W and Z is proportional to the square of each of their masses. The degrees of freedom from the broken symmetry that would typically manifest as Goldstone bosons are instead 'eaten' by

the longitudinal polarization of the W and Z. The combination of W^3 and B corresponding to the photon does not acquire a mass and $U(1)_{EM}$ remains as a symmetry of the theory. Note that the Higgs boson's own mass has also appeared: $m_h = \sqrt{-2\mu^2} = \sqrt{2\lambda}\nu$.

The fermion masses appear in a similar way: as a (three-point) Yukawa interaction with the vev of the Higgs field. Before symmetry breaking, this is

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{\text{families}} \left(\Gamma^u_{mn}\overline{q}_{mL}\tilde{\varphi}u_{nR} + \Gamma^d_{mn}\overline{q}_{mL}\varphi e_{nR} + \Gamma^e_{mn}\overline{\ell}_{mL}\varphi e_{nR} + \Gamma^\nu_{mn}\overline{\ell}_{mL}\tilde{\varphi}\nu_{nR}\right) \tag{2.7}$$

where the conjugate form of the Higgs field breaks to $\tilde{\varphi} = (v+h, 0)/\sqrt{2}$ and the fermion fields are the weak (not mass) eigenstates. The matrices Γ^i_{mn} are totally arbitrary, but will be proportional to a (non-diagonal) mass matrix. After symmetry breaking, these terms are again reexpressed around the vev leading to terms like $\Gamma^e_{mn}\overline{\ell}_{mL}(\nu+h)\,e_{nR}/\sqrt{2}$. Suggestively defining $M_{mn} \equiv \nu\Gamma_{mn}/\sqrt{2}$ yields instead

$$\mathcal{L}_{\text{Yukawa}} = -\overline{\ell}_{mL}M^e_{mn}\,(1 + gh/2m_W)\,e_{nR}/\sqrt{2} + \cdots . \tag{2.8}$$

Because the Higgs boson and the vev of the field are introduced together in the Yukawa interaction, the resultant masses of the fermions are proportional to the strength of their coupling with the Higgs boson.

2.3 Theoretical and Experimental Constraints

In tandem with precision electroweak and top quark measurements, the theoretical framework of the Standard Model restricted the values that the Higgs boson mass could take and provided indirect measurements of its value, before its direct observation. Both LEP and the Tevatron also searched directly for the Higgs without finding it, and set exclusions.

Three fundamental theoretical limits may be considered. A first consideration provides both a (weak) upper limit on the Higgs mass and, perhaps more importantly, a very strong reason for the Higgs to exist. If m_h is too large, or if it is removed from the theory, the amplitude for longitudinally polarized WW scattering grows linearly with s. This leads to a violation of unitarity at the TeV scale, so *something* must break electroweak symmetry.

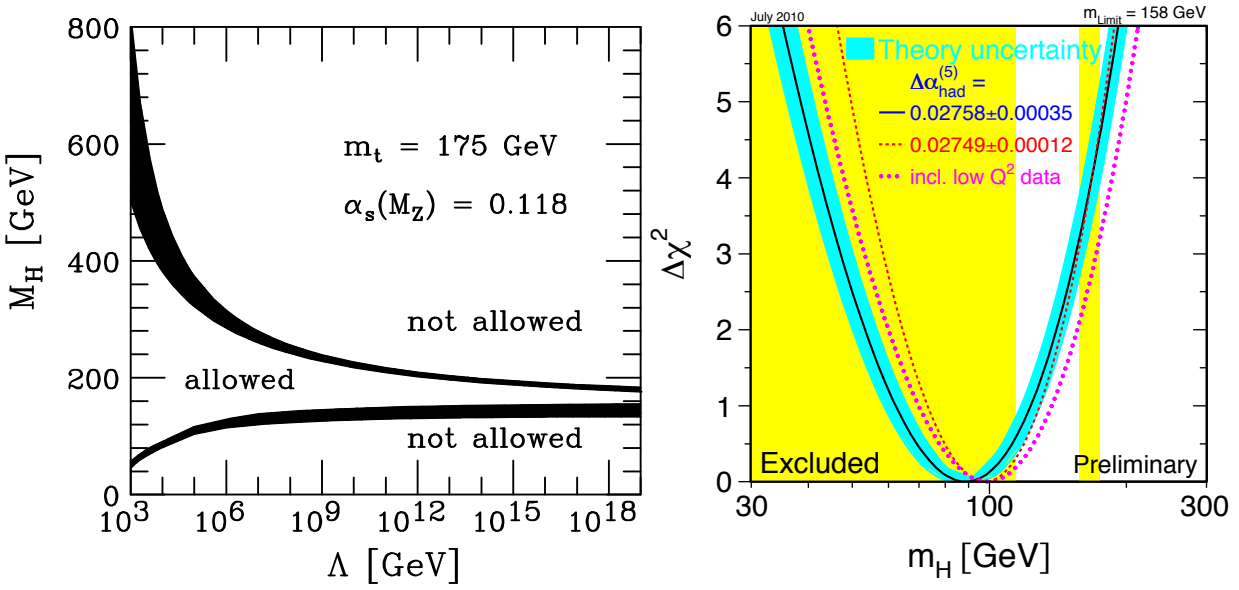
Two other limits take advantage of the running of $\lambda = m_h^2/2\nu^2$ with Q^2, and yield results that depend on an upper range of applicability of the theory, Λ. The first, triviality, refers to the 'Landau pole' that arises in $\lambda\left(Q^2\right)$ at large m_h: the self-coupling λ cannot diverge where the theory is valid. This leads to a limit of $m_h < 140$ GeV, if the theory is valid up to the Planck scale, or $m_h \lesssim 650$ GeV if the theory is valid up to $\Lambda = 1.5$ TeV [16, 18]. A lower bound comes from requiring $\lambda\left(Q^2\right) > 0$, which is necessary for the stability of the vacuum. The resultant limits range from 85 GeV at $\Lambda = 1.5$ TeV to around 115 GeV at $\Lambda = m_P$. These results are summarized in Figure 2.2a.

Radiative corrections to the W mass contain a weak (logarithmic) dependence on the Higgs mass, as do the forwards-backwards asymmetries of the Z. In the context of a global fit of the Standard Model, the Higgs mass may be calculated as $m_h = 94^{+25}_{-22}$ GeV [19], which may also be interpreted as an upper bound. Finally, direct searches at LEP and the Tevatron provided a 95% CL lower limit of $m_h > 114.4$ GeV and an exclusion band of $158 - 175$ GeV [20]. Paired with the indirect constraints, this yields an allowed region for the Higgs mass of $114 - 158$ GeV, as shown in Figure 2.2b.

The Higgs mechanism provides an elegant solution to serious theoretical problems: it provides masses to both the fermions and the vector bosons. Electroweak symmetry breaking sidesteps unitarity violation in TeV-scale W scattering. Before its discovery, the value of the Higgs mass was constrained both theoretically and through direct searches, leaving a relatively narrow window in which to search for it.

2.4 Higgs Production at the LHC

There are six principal Higgs boson production modes at the LHC. The Feynman diagrams are presented in Figure 2.3. The cross sections and uncertainties compiled by the LHC Cross Section Working Group [21–23]



(a) Triviality and Stability

(b) Direct and Indirect Limits from SLD, LEP and the Tevatron

Figure 2.2: Constraints on the Higgs boson mass before its observation.

	Single Higgs Production			Pair Production		
	Cross Section [pb]	Scale Vars [%]	PDF + α_s [%]	Cross Section [fb]	Scale Vars [%]	PDF + α_s [%]
$ggh(h)$	19.27	$^{+7.2}_{-7.8}$	+7.5/-6.9	8.16	+20.4/-16.6	+8.5/-8.3
VBF	1.578	$^{+0.2}_{-0.2}$	+2.6/-2.8	0.49	+2.3 /-2.0	+6.7/-4.4
$Wh(h)$	0.705	$^{+1.0}_{-1.0}$	+2.3/-2.3	0.21	+0.4 /-0.5	+4.3/-3.4
$Zh(h)$	0.415	$^{+3.1}_{-3.1}$	+2.5/-2.5	0.14	+3.0 /-2.2	+3.8/-3.0
$tth(h)$	0.129	$^{+3.8}_{-9.3}$	+8.1/-8.1	0.22	–	–

Table 2.1: SM production cross sections for Higgs boson (pair) production at $m_h = 125$ GeV and $\sqrt{s} = 8$ TeV [23, 24]. Fractional uncertainties from scale variations and PDFs are displayed. The single-Higgs boson uncertainties are very similar for $\sqrt{s} = 7$ TeV. The bbh production mechanism has only received attention more-recently, and was not included in these benchmark references. Its rate is expected to be approximately 1.6% of the ggh one. The branching rates to photons and b quarks used in this document are 0.00228 (±4.9%) and 0.569 (±3.3%), respectively.

are listed in Table 2.1 for convenience. The production rates are shown as a function of m_h in Figure 2.4a. The various production modes may be separated via the presence of additional objects in the final state.

- ▷ Gluon fusion (ggh) through a top quark loop accounts for 87% of Higgs boson production at $\sqrt{s} = 8$ TeV at the LHC. To first order, there are no additional objects in the final state, though higher-order corrections obviously lead to some quark and gluon radiation. The ggh production cross section is computed at next-to-next-to-leading order (NNLO) in QCD [25–27], and next-to-leading order (NLO) electroweak (EW) corrections from Refs [28–30] are applied. These results are compiled in Refs [31,32] assuming factorization between QCD and EW corrections [31, 33].

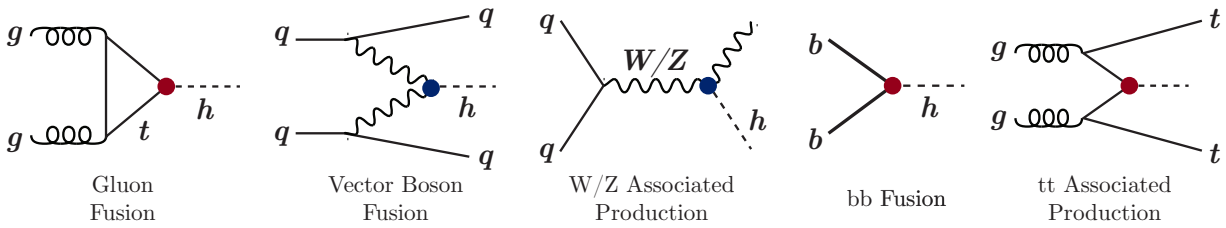
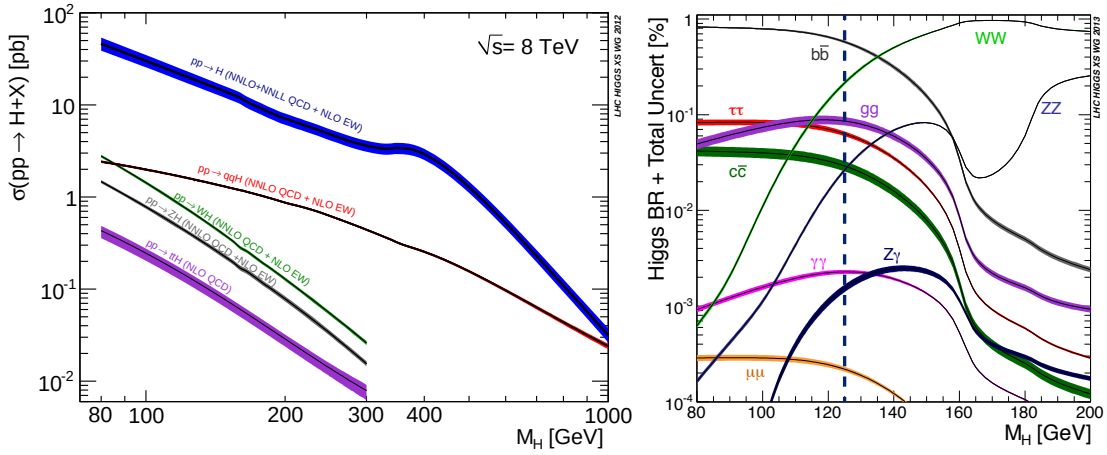


Figure 2.3: Standard Model Higgs boson production diagrams.

- ▷ Vector boson fusion production (VBF) is distinguished by the presence of two jets with large difference in pseudorapidity and, typically, very large dijet mass m_{jj}. The production cross section has been calculated with full NLO QCD and EW corrections [34–36], and approximate NNLO QCD corrections [37].
- ▷ Associated production with a W or Z boson ('Higgsstrahlung;' Wh and Zh) may be 'tagged' by the presence of two central jets with mass near m_W or m_Z ($W \to qq'$ or $Z \to q\bar{q}$), or by the presence of leptons or missing energy ($W \to \ell\nu$, $Z \to \ell\ell$, $Z \to \nu\nu$). The QCD corrections to the Wh and Zh processes have been calculated at NLO [38] and at NNLO [39]; NLO EW radiative corrections from Ref. [40] are applied.
- ▷ Associated production with top quarks (tth) allows direct access to the tth vertex (without the loop). It accounts for a very small piece of the total production at the LHC. The $t\bar{t} \to W^+bW^-\bar{b}$ decay leads to a messy final state with leptons and/or many jets, two of which are initiated by b-quarks. The full NLO QCD corrections for $t\bar{t}h$ are used [41–44].
- ▷ Bottom quark fusion ($b\bar{b}h$) is included at tree-level in the 'five flavor scheme' in which bs are explicitly included in the Parton Distribution Functions (PDFs) of the incoming protons. Otherwise, this production may be viewed as an alternative diagram for gluon fusion, with two gluon splittings as shown for $t\bar{t}h$ production in Figure 2.3. Either way, additional radiation tends to be very soft, and the final state closely resembles gluon fusion [45–47]. Bottom quark fusion has a rate 1.6% as large as ggh, and is not included in any of the analyses presented. This was an oversight and in future ATLAS analyses it will be added.

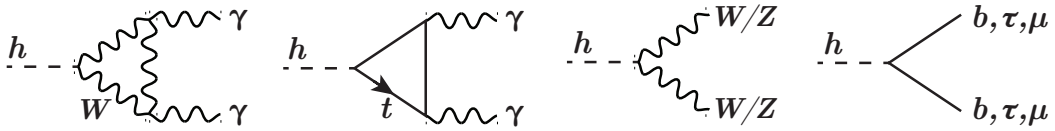
The Higgs boson couples to other particles according to their masses, as described in Section 2.2. Its branching ratios are accordingly calculated in Refs [48–50] and presented presented in Figure 2.4b. The three diboson decay modes provided the initial discovery at ATLAS:

- ▷ The diphoton mode represents the bulk of this thesis. The signal over background in this mode is around $S/B \approx 1/30$ and the branching ratio $\mathcal{B} = 0.00228$ is very small, but the rate remains competitive, thanks to a high selection efficiency. The analysis strategy is straightforward, and the mass resolution of ∼1.6 GeV is quite good. The large number of events and clean signature make this channel attractive for studies of the couplings of the Higgs boson.
- ▷ The $ZZ \to 4\ell$ mode benefits from a high $S/B \approx 3/2$ [13] and oustanding resolution. The small $Z \to \ell\ell$ branching fraction of 6.6% [51] (squared!) leads to very low statistics for this analysis, but because all four leptons are reconstructed it nevertheless has good sensitivity to spin and CP eigenvalues.
- ▷ After branching ratios and selection efficiency, the WW mode has rate comparable to that of $h \to \gamma\gamma$. Its $S/B \approx 1/8$ is better than in $\gamma\gamma$, but it has very poor mass resolution, due to the neutrinos in the final state.



(a) Cross Sections (b) Branching Ratios

Figure 2.4: Production cross sections and branching ratios for the SM Higgs boson as a function of mass, at $\sqrt{s} = 8$ TeV [23]. The observed mass of ∼125 GeV allows for a great diversity of different decay channels to be measured.



Higgs to Two Photons Tree-Level Boson Decays Fermionic Decays

Figure 2.5: Diagrams of the decay modes of the SM Higgs boson.

Because a fermiophobic particle would not be produced through gluon fusion, fermionic couplings may be inferred well before direct decays to fermions are measured. Direct observations in $h \to b\bar{b}$ and $h \to \tau^+\tau^-$ are important for measuring (a) the couplings to different quark types (top v. bottom) and leptons, which could be altered in scenarios with multiple Higgs doublets, and (b) for the long 'lever-arm' in demonstrating that the Higgs couplings run proportional to mass. These two channels have fairly large branching fractions (0.57 and 0.06), but are difficult to distinguish from very large backgrounds.

The $h \to \mu\mu$ and $h \to Z\gamma$ modes are exceptionally rare, but nevertheless interesting, for the 'full picture' of the consistency to the SM. The decay to muons is the only channel in which the coupling to second-generation fermions could realistically be observed.

2.5 Higgs Boson Pair Production

At the end of the first run of the LHC, the Higgs boson self-interaction λ_{hhh} stands as one of the great unmeasured features of the Standard Model. This coupling will be measured through the pair production of two Higgs bosons. But at $\sqrt{s} = 8$ TeV, the rate for this production is small – less than 10 fb, before branching ratios. In the meantime, Higgs boson pair production provides an important portal to new physics.

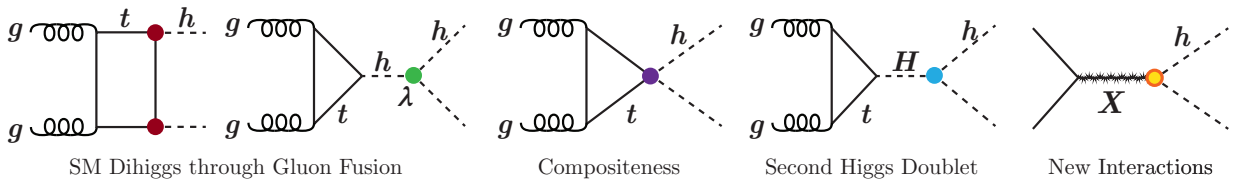


Figure 2.6: Illustrative dihiggs production diagrams in the SM and beyond.

2.5.1 Standard Model Production

SM Higgs boson pair production proceeds by the same basic diagrams as single Higgs boson production (Figure 2.3). The difference is that for each single Higgs diagram, two variants are possible: (a) two Higgs bosons may be radiated off of a quark or boson line, or (b) the Higgs boson itself may split, to pair produce. This is illustrated for the *gghh* mode, in the left two diagrams of Figure 2.6. As listed in Table 2.1, gluon fusion remains by far the dominant production mode for dihiggs boson production. The 'box diagram' interferes destructively with and overwhelms the far more-interesting self-coupling diagram. This is the reason that the self-coupling will be so difficult to measure: even after Higgs pair production is observed, it will be a long way from extracting λ_{hhh}.

Nevertheless, beginning this search affords an opportunity to develop experimental methods and understand backgrounds.

2.5.2 Production Beyond the Standard Model

The serious motivation for searching for Higgs boson pair production with Run I data is the cornucopia of extended models that enhance this rate. Perhaps most exciting is the potential for an utterly anticipated interaction! The enhancements fall into two basic categories: resonant and non-resonant production.

2.5.2.1 Resonant Production

In many BSM theories, a second Higgs doublet is introduced with $SU(2)_L \times U(1)_Y$ charge; together with the charge conjugate doublet this provides four additional degrees of freedom. This leads to four new bosons: a heavier scalar H, a pseudoscalar A, and two charged Higgs bosons, $H^{\pm}$. For convenience, three new states (H^0, A, $H^{\pm}$) are typically taken to have similar masses (an additional scale also creates some theoretical problems). Different permutations of fermion-doublet Yukawa couplings are possible. In 'Type I' 2HDMs all leptons and quarks couple to a single doublet, and in 'Type II' 2HDMs up-type quarks couple to one doublet while down-type quarks and leptons couple to the other. The properties of the models are then determined by the masses, by the ratio of the vevs of the two doublets, $\tan\beta \equiv v_1/v_2$, and by the angle α that describes the mixing of the two neutral scalars. The punchline is that if the heavier scalar has mass $m_H > 2m_h$, the cross section for $pp \to H \to hh$ may reach a few picobarns, as shown in Figure 2.7. Rates drop off again for $m_H > 2m_t$, where the branching $H \to t\bar{t}$ turns on.

Evidence that the observed Higgs boson closely matches SM predictions motivates two classes of 2HDM parameters: (1) the decoupling limit where the 'extra' bosons are very heavy and (2) the alignment limit $\cos(\beta - \alpha) = 0$ where the vev lies entirely in the neutral component of one of the doublets. Current measurements constrict 2HDMs tightly to these limits [52–54].

Many other resonant models are possible. Gravitons can decay to a pair of Higgs bosons [55], as can radions (whose field is useful for stabilizing the graviton field) [56]. Yet again, stoponium production could lead to a narrow resonance of two Higgs bosons [57]. It is easy to add another singlet to the SM or to a 2HDM; because the Higgs is a scalar, it is easy for it to mix with the singlet [58, 59]. The Higgs would then be a portal to this new sector.

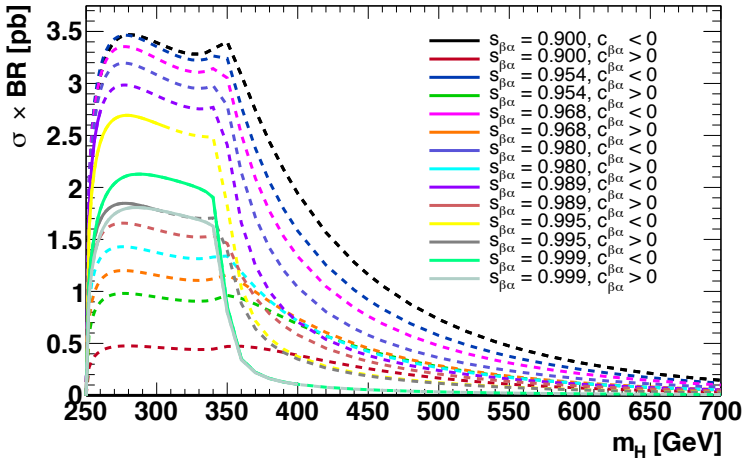


Figure 2.7: Cross section times branching ratio for a resonant dihiggs production in a Type 1 2HDM with $\tan\beta = 1$, as a function of the mass of the heavy scalar. A variety of values are shown for $\sin(\beta - \alpha)$, near the alignment limit. Models that remain viable with respect to the stability of the potential and unitarity are drawn solid, while other combinations of parameters are dashed.

2.5.2.2 Non-Resonant Production

Non-resonant enhancements to dihiggs production are also possible. Simply modifying the self-coupling λ_{hhh} – turning it off or changing the sign – can lead to modest enhancements of the $pp \to hh$ rate [24], but these would not be accessible with present data sets. In composite models, a direct (anomalous) $t\bar{t}hh$ coupling could boost dihiggs production [60], as shown in Figure 2.6. Finally, light colored scalars running in the loop could enhance the $pp \to hh$ rate [61].

Chapter 3

Experimental Apparatus

This chapter briefly describes the Large Hadron Collider and the ATLAS experiment. Vast documentation exists on these projects, so references with greater detail are given liberally and may be consulted at will. The feats recorded in those pages are the foundation upon which the entire present work is built.

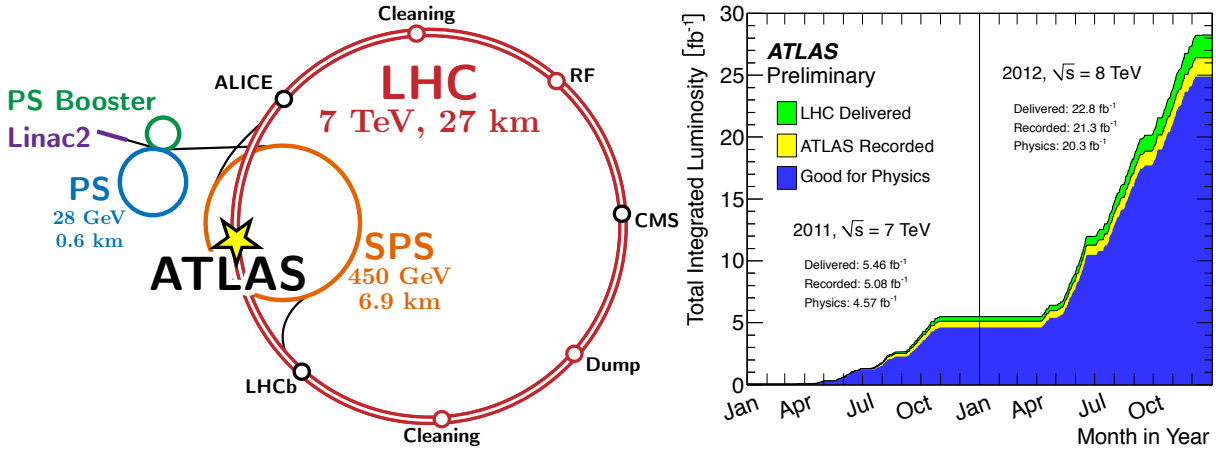
3.1 The Large Hadron Collider

The Large Hadron Collider [62–65] is a 26.7 km super-conducting accelerator designed to collide protons at a center of mass energy of $\sqrt{s} = 14$ TeV at a rate of 10 nb^{-1} per second ($10^{34}/\text{cm}^2\,\text{s}$).

The injection complex of the LHC reuses several of CERN's older accelerators [63]. The site layout is illustrated in Figure 3.1a. The acceleration chain begins with a duoplasmatron that extracts protons from hydrogen molecules: it bombards the H_2 molecules with free electrons to dissociate the valence electrons from the nuclei, and accelerates the resultant protons into the Linac2. The Linac2 focusses this beam and accelerates it to 50 MeV, delivering it to the Proton Synchrotron (PS) Booster, which accelerates the protons in turn to 1.6 GeV. The PS Booster feeds into the PS which accelerates the beam to 26 GeV, and feeds into the Super Proton Synchrotron (SPS). The SPS accelerates the protons to an energy of 450 GeV and feeds into the LHC. The full injection chain takes around 4 minutes.

The LHC itself may be divided into eight arcs, each of which has a long straight section of 528 m and two bending regions at either end. Each straight region serves as an insertion point either for an experiment or for a beam utility. Adjacent to the ATLAS [66] experiment are experimental sites for LHCb [67] and ALICE [68], while CMS [69] is installed on the opposite side of the ring. Of the four remaining insertion points, two are taken up by collimators for the beam, one is reserved for the beam dump, and the last holds the radio frequency (RF) acceleration cavities. The total energy achievable by the LHC is limited by its circumference and the field of its bending magnets. The LHC contains 1232 bending dipoles with nominal field 8.33 T, for $\sqrt{s} = 14$ TeV. Quadrupole, sextupole, and octopoles are used to focus the beam and reduce aberrations. It takes around 20 minutes to ramp the beam energy from the 450 GeV injection energy to the full energy. Due to persistent concerns over the catastrophic magnet failure mentioned in Chapter 1, the LHC was operated at $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012, instead of at its design energy of $\sqrt{s} = 14$ TeV.

Along with the beam energy, the second important parameter is the luminosity $\mathcal{L}$, which is proportional to the rate of collisions. More luminosity means more Higgs bosons. The luminosity may be expressed as the quotient of the total number of times that two protons cross paths (per second), divided by the cross sectional area A at the collision point. The number of crossings is given as $N_b^2 n_b^2 f_{\text{rev}}$, where $N_b \sim 10^{11}$ is



(a) Large Hadron Collider (b) Data Delivered to ATLAS

Figure 3.1: The CERN accelerator chain begins with a linear accelerator and three booster rings - the Proton Synchrotron Booster, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS) – that accelerate protons to 450 GeV, before injecting them into the Large Hadron Collider (LHC). The LHC accelerates the protons to a design energy of 7 TeV. Adapted from Ref. [62]. (b) The total data delivered by the LHC, and recorded and deemed usable for physics by the ATLAS experiment is displayed for 2011 and 2012, when the accelerator was operating at energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV respectively. [70]

the number of protons per bunch, $n_b = 1380$ is the number of bunches, and $f_{\text{rev}} = 26.7\ \text{km}/c \approx 11.25$ kHz is the revolution frequency of the beam.

The cross sectional area meanwhile, is $A = 4\pi\varepsilon_n\beta^*/F\gamma$. F is a geometric factor that describes the crossing at the interaction point. The emittance ε_n is the average normalized phase space occupied by the beam in momentum and position space, β^* is a measure of the transverse beam size, and γ is the Lorentz factor. The total instantaneous luminosity may thus be written,

$$\mathcal{L} = \frac{N_b^2 n_b^2 f_{\text{rev}}}{A} = \frac{N_b^2 n_b^2 f_{\text{rev}} \gamma}{4\pi\varepsilon_n\beta^*} F\,. \tag{3.1}$$

Peak luminosities in ATLAS reached around 7 nb^{-1} per second in 2012 ($7 \times 10^{33}/\text{cm}^2\,\text{s}$) [70]. The parameters used to achieve these high luminosities also resulted in as many as 30 interactions in a single crossing. Many simultaneous interactions makes for messier events, but it is worth the cost. Figure 3.1b shows the total integrated luminosity for the two years, and the steep rate in 2012. Nearly 25/fb of collisions were recorded in the two years. At peak luminosity, the LHC was producing a Higgs boson for ATLAS every seven seconds!

3.2 The ATLAS Detector

ATLAS (A Toroidal LHC ApparatuS) [66, 71–73] is a multipurpose detector with forward-backward symmetric cylindrical geometry and nearly 4π coverage in solid angle. It is composed of three cylindrical subsystems, arranged in concentric shells around the interaction point:

(1) The inner detector (ID) is a tracker immersed in an axial magnetic field that measures the tracks of charged particles, and hence their origin ('primary' or 'secondary' vertices), position, and momentum.

(2) Electromagnetic and hadronic calorimeters stop electrons, photons, and hadrons, and measure their energies. This is also important for reconstructing 'missing' energy from neutrinos.

(3) Muons interact little with the calorimeter, so their trajectories are measured a second time in a 'muon spectrometer' (MS). The MS is largely contained within the toroidal magnet system that gives ATLAS its name.

Each system typically has a 'barrel' component centered at the interaction point sandwiched between two 'endcap' components, further along the beam line.

3.2.1 Coordinate System

ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) and the z-axis directed along the beam pipe, counter-clockwise around the LHC ring if looking downwards. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upward.

The momenta of incoming partons in proton-proton collisions are not well-determined; rather, they are described by Parton Distribution Functions (PDFs) that quantify the fraction of the total momentum in each parton. Since the other partons 'escape' down the beam-pipe, momentum conservation is not manifest along the z direction. However, since the beams collide head-on in z, momentum conservation *is* apparent in x-y. Vectors – in particular, momenta – projected into the x-y plane are called 'transverse.' Cylindrical coordinates (R, φ) are used in the transverse plane, where φ is the azimuthal angle around the beam pipe.

Particle production is roughly constant as a function of the rapidity, $y \equiv \frac{1}{2} \ln \left[(E + p_z) / (E - p_z) \right]$. For massless particles (such as photons), this is equivalent to the *pseudorapidity*, which is defined in terms of the polar angle θ from $\hat{z}$ by $\eta \equiv -\ln \left[\tan (\theta/2) \right]$. The pseudorapidity is the preferred variable for describing the detector, since it is well-defined in the detector frame (independent of the particle mass).

The angular separation between two objects is typically described by $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2}$.

3.2.2 Detector Overview

3.2.2.1 Inner Detector

The ID [74–76] provides accurate reconstruction of the positions, momenta, and origins of charged particles. It is designed using two technologies. The pixel detector [77, 78] and the semiconductor tracker (SCT) [79–81] use silicon pixels and microstrips, while the Transition Radiation Tracker (TRT) [82–85] is a straw tracker with particle identification capabilities using transition radiation. The geometry of the ID is shown in Figure 3.2.

The Pixel Detector. The principle of silicon tracking is that the difference in Fermi energies on either side of an interface between p and n-type semiconductors (a diode) leads to a 'depletion region' essentially void of free charge carriers. When a charged particle passes through the semiconductor it leaves a high density of electrons and holes in its wake – typically 80 $e^- e^+$ pairs per micron of Silicon. In the depletion region, these charges do not immediately recombine. Applying an additional large voltage across the semiconductor then serves two purposes: (1) it enlarges the depletion region, allowing for a greater number of un-recombined charges and (2) it provides the electric field necessary to definitively separate and measure them.

ATLAS uses a total of 80.4 million pixels spread over 1744 sensor modules, providing an average of three 'hits' on track. The pixel size is 50 μm in $R\varphi$ by 115 μm in z and the intrinsic accuracy is 10 μm in $R\varphi$ and 115 μm in z (R) in the barrel (endcap). The modules are arranged in three concentric cylinders in the barrel ($R = 50.5$, 88.5, 122.5 mm) and three disks in each endcap ($z = 495$, 580, 650 mm), providing coverage out to $|\eta| < 2.5$.

In addition to 'standard' tracking, the pixel detector is critical for distinguishing separate vertices in events with many hard scatters, and for reconstructing the decays and secondary vertices from b-hadrons.

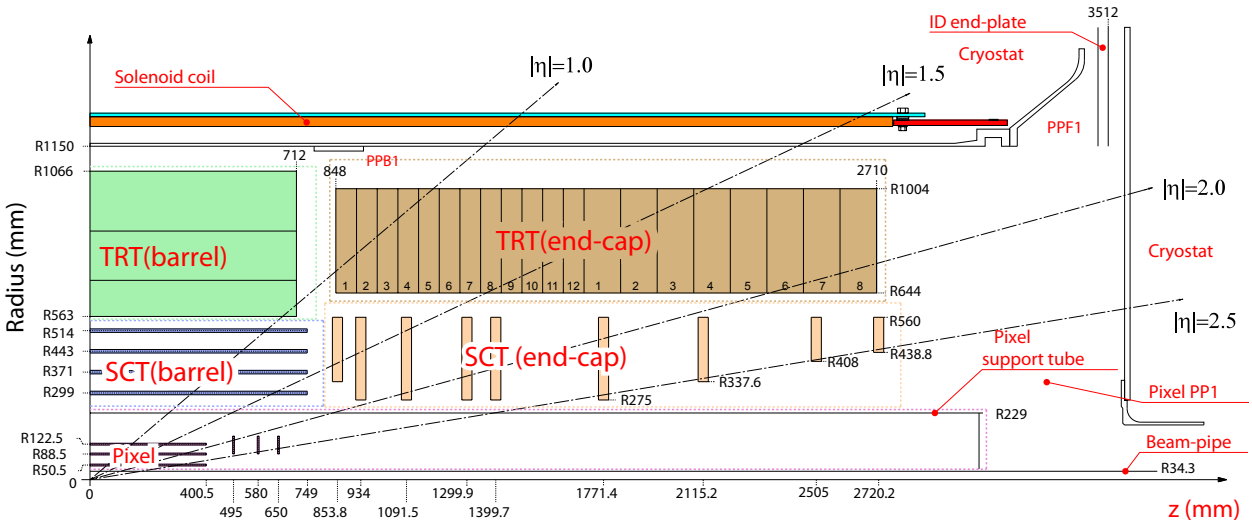


Figure 3.2: Schematic of one quadrant of the inner detector, projected in R-z. Starting from the center are the pixel detector, the semiconductor tracker, and the transition radiation tracker. Each piece piece contains one barrel and two endcaps. The inner detector extends to $|\eta| < 2.5$. [66]

The innermost 'b-layer' is also used to distinguish between electrons (which leave hits in that layer) and photons that convert into electron-positron pairs (which should not).

The Semiconductor Tracker. Beyond the pixel detector, the SCT provides tracking out to $R < 563$ mm. Each of 15912 sensors contains 768 'strips,' for a total of 6.4 million channels. Each strip is 12 cm long and has a pitch of 80 μm. Sensors are mounted on both sides of each module with an angular offset of 40 mrad; this 'stereo' measurement provides a longitudinal (radial) constraint in each layer in the barrel (endcap). The full intrinsic resolution per hit is 17×580 μm in $R\varphi \times z$ (R in the endcap). The modules are arranged in four layers in R in the barrel (299, 371, 443, 514 mm) and in nine layers in z (from 854 to 2720 mm), providing an average of 8 hits on track (4 space points) out to $|\eta| < 2.5$.

The Transition Radiation Tracker. The TRT is a straw tracker with 350 thousand channels providing semi-continuous tracking and an average of 36 hits on track. The straws have a radius of 2 mm and are constructed of polyimide and stabilized with carbon fiber. They are filled with a xenon/carbon dioxide/oxygen gas mixture (70%, 27%, 3%), and a 32 μm diameter gold-plated tungsten wire is strung down the center. The straw acts as the cathode and is held at -1500 V with respect to the wire. When a charged particle passes through a straw, it ionizes the gas. The electrons drift towards the wire, inducing an 'avalanche' (gas gain) that amplifies the signal by about 2×10^4. The timing of the leading edge of the signal is related to the radius of closest approach of the ionizing particle from the wire, and this information can be used to obtain 130 μm resolution in $R\varphi$ per straw.

The TRT provides tracking to $|\eta| < 2.0$. The barrel extends to $|z| < 712$ mm with $563 < R < 1066$ mm, and the endcap fills the volume of $644 < R < 1004$ mm and $848 < |z| < 2710$ mm. The combination of many measurements along with the much-longer lever arm, enhance the TRT's contribution to the total momentum measurement.

In addition to tracking, the TRT has extremely unusual particle ID capabilities. Polyethelene felt mats are interleaved with the straws in the barrel, and polypropylene sheets are placed between wheels (disks) of straws in the endcap. These 'radiators' induce transition radiation (TR) by incident particles, equal to

$$E = \alpha\hbar\omega_p\gamma/3 \tag{3.2}$$

where $\alpha = 1/137$ is the fine structure constant, ω_p is the plasma frequency of the radiator, and γ is the Lorentz factor. Typical TRT photons have energies of several keV – much more energy than typically left through ionization. These TRT photons are absorbed by the xenon gas and lead to a cluster of electrons that induce a large shower. Showers that exceed both a lower threshold used for tracking and a higher 'TR threshold' are flagged with a dedicated 'bit.' Because the likelihood for a TR photon to be emitted scales with γ, the fraction of hits that exceed the TR threshold can be used as flag for discriminating electrons and pions: pions are around 250 times more massive than electrons, so electrons and pions with equal momenta have very different γ factors.

Solenoid. A thin super-conducting solenoid operating at 4.5 K provides a 2 T axial magnetic field that curves the tracks of charged particles and makes it possible to measure their momenta [86].

3.2.2.2 Calorimetry

The ATLAS calorimeter [87] consists of four distinct subsystems providing electromagnetic and hadronic energy measurement to $|\eta| < 4.9$. In addition to electron, photon, and hadron energy measurements, the calorimeters and muon spectrometer (MS) make it possible to infer the presence of neutrinos (or more exotic particles) in the form of 'missing' energy in events. The emphasis in this section is given to the lead/liquid-argon electromagnetic calorimeter in the central region $|\eta| < 2.5$, because that is what is used to identify and measure photons (Chapter 4) in the context of the $h \to \gamma\gamma$ analyses (Chapters 5-8). Indeed, many of the design parameters of this detector are a direct response to the needs of this search!

Liquid Argon Electromagnetic Calorimeter. When a photon or electron enters a medium it 'showers' in a series of $\gamma \to e^+e^-$ conversions and $e^\pm \to \gamma e^\pm$ Bremsstrahlung, until the electron energy loss through Bremsstrahlung falls to the level of loss through ionization of particles in the medium. A material may be understood in terms of the radiation depth X_0, which is defined as the distance that an electron must traverse in a material to lose $1/e$ of its energy to Bremsstrahlung. For photons, it is the distance necessary for a 54% probability of conversion into an electron-positron pair. The depth of a shower and the total number of electrons produced are proportional to the energy of the incident particle, while the depth of the maximum energy deposition grows logarithmically with the energy [88].

The ATLAS electromagnetic (EM) calorimeter [89] is a sampling instrument with inactive lead absorbers interspersed with active liquid argon that samples the ionization. That ionization induces currents that are collected on Kapton electrodes. Cells in the calorimeter are defined in η by the etching process on the Kapton, and grouped in azimuth. The divisions along η allow for 'pointing' back to the interaction point by tracing between layers of the detector. Photons that do not convert in the inner detector do not leave tracks, and in this case the 'pointing' is a useful constraint on the origin of the photon. The detector is also notable for its accordion geometry (Figure 3.3) which makes it fully hermetic, and gives approximate azimuthal symmetry. It is more than $22X_0$ deep for all η.

Four layers are defined in X_0, with varying depths and $\Delta\eta$ and $\Delta\varphi$ granularities. To assist in estimating losses from material upstream of the calorimeter, a presampler is installed in front of the first layer and before some detector services. The presampler has $\Delta\eta \times \Delta\varphi$ granularity of 0.025×0.1 for $|\eta| < 1.52$ and $1.5 < |\eta| < 1.8$. It ends at $|\eta| = 1.8$, where there is less dead material and where a photon or electron of a given transverse momentum has a greater total energy ($p_\mathrm{T} = E/\cosh(\eta)$). After the presampler comes the 'strips,' which are finely segmented in $\Delta\eta$ to provide π^0 rejection (Section 4.3). They are $6X_0$ deep. For $|\eta| < 1.4$ and $1.5 < |\eta| < 1.8$ the width is just $\Delta\eta = 0.025/8$ (one eighth of a second-layer cell), while for $1.8 < |\eta| < 2.0$ and $2.0 < |\eta| < 2.4$ the granularity broadens to $0.025/6$ and $0.025/4$ respectively, since the physical space corresponding to a unit of η falls with η. At $|\eta| = 2.4$, the fine segmentation ends, and photon identification accordingly stops at $|\eta| = 2.37$. The majority of the energy deposited in the EM calorimeter is collected in the second layer, which is the deepest layer ($\gtrsim 15X_0$). Cells in the second layer of the calorimeter measure $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$. For $|\eta| < 2.5$, the third and final layer of the calorimeter allows a last

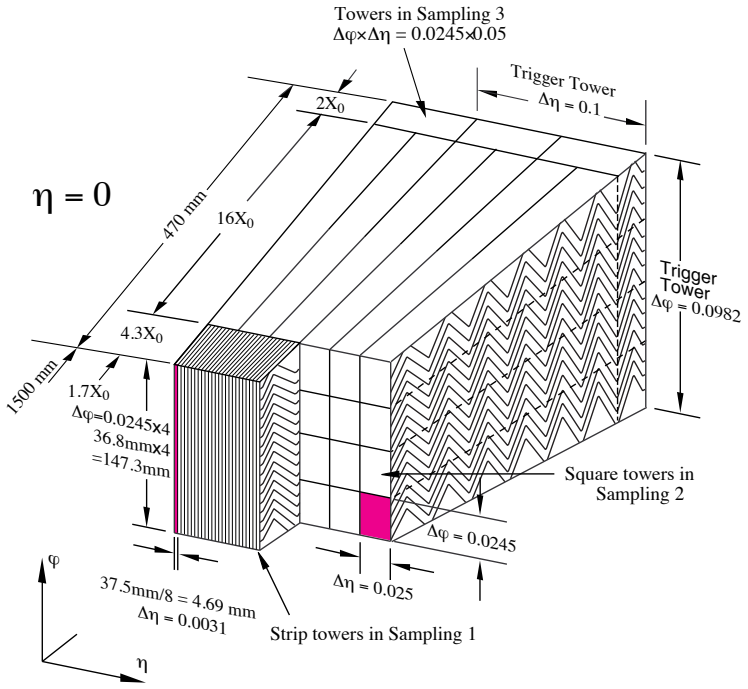


Figure 3.3: Schematic of the liquid argon calorimeter, showing the accordion geometry that avoids azimuthal cracks. The first layer is finely segmented, to discriminate photons from neutral pions. Most of the energy is deposited in the deep (16 radiation-length) second layer. [66]

measurement of the electron or photon before the hadronic calorimeter. The EM calorimeter does extend to $|\eta| < 3.2$, but without a third layer or fine segmentation in the first layer.

The combination of these elements provides outstanding performance in electron and photon identification (target jet rejection of 5000×, Section 4.3) and energy resolution (Section 4.2). The energy resolution is typically parameterized as $\delta E/E = A/\sqrt{E} \oplus B/E \oplus C$ where $A \sim 10\%$ is the sampling term, $B \sim 100$ MeV is the noise term, and $C \sim 0.5\%$ is constant term.

It is worth noting that although the EM calorimeter is designed for electrons and photons, typical jets will deposit around half of their energy there.

Tile Calorimeter. The ATLAS hadronic calorimeter [90] uses steel absorbers and active scintillators that are read out through wavelength shifting fibres that feed into photomultiplier tubes. The barrel covers $|\eta| < 1.0$ while the extended barrels covers $0.8 < |\eta| < 1.7$. The depth is measured in terms of the nuclear interaction lengths λ, which is the mean length that a particle will travel before undergoing a hard scatter. The calorimeter is more than 11λ deep everywhere which reduces 'punch-through' of hadrons into the MS far below the irreducible background from prompt and decay muons. The $\Delta\eta \times \Delta\varphi$ granularity is 0.1×0.1 in its first two layers and 0.1×0.2 in the third.

Hadronic Endcap Calorimeter. For $1.7 < |\eta| < 3.2$ liquid argon with copper absorbers are used for the hadronic calorimeter [89]. The instrument has four layers of longitudinal segmentation, and angular granularity of 0.1×0.1 at $1.5 < |\eta| < 2.5$ that broadens to 0.2×0.2 for $2.5 < |\eta| < 3.2$.

Forward Calorimeter. Wrapping around the beampipe in the forward region, the liquid argon forward calorimeter provides a final depth of 10λ at $3.1 < |\eta| < 4.9$. It consists of three layers of absorbers: first copper and then two tungsten, with $\Delta\eta \times \Delta\varphi$ granularity granularity 0.2×0.2. Since a 'size parameter' of

$\Delta R = 0.4$ is used in jet reconstruction (Section 6.1), the 'end' of the detector at $|\eta| = 4.9$ motivates the requirement that jets be contained within $|\eta| < 4.5$.

3.2.2.3 Muon Detectors.

The ATLAS muon spectrometer (MS) [91] consists of two tracking systems and two triggering systems immersed in the ∼0.5 T field of the 1300-ton toroidal magnet [92, 93] that gives ATLAS its name. The toroid bends the muons' tracks and the spectrometer is designed to measure the sagita of these tracks to reconstruct the muons' p_{T}s within a few percent. Muons are used to flag Wh and Zh production (both alone and in missing energy reconstruction) in Chapter 6, and in reconstructing b-jets in Chapter 8.

The Monitored Drift Tubes (MDTs) [94] and the Cathode Strip Chambers (CSCs) [95] are designed for precision tracking. The MDTs are effectively larger versions of the TRT's straws. The 350 thousand aluminum tubes range in length between 1.2 and 6 m; they are 30 mm in diameter with 50 μm gold-coated tungsten wires at the center. The tubes are filled with an $\mathrm{Ar} - \mathrm{CH_4} - \mathrm{N_2}$ mixture, providing gas gain of 2×10^4. After installation, structural deformations in the system are monitored (whence the name) by an optical system. In order to achieve percent-level accuracy in the momentum resolution of muon tracks, accuracy better than 50 μm is required in the muon tracks. Individual tubes enable a position measurement of 80 μm; together, they attain the design requirements [91].

In the higher-flux region at low-$|z|$ and high-$|\eta|$, the occupancy is too high for the drift tubes, so CSCs are used. These are proportional chambers with tungsten anode wires at 2600 V running perpendicular to cathode readout strips in an $\mathrm{Ar} - \mathrm{CO_2} - \mathrm{CF_4}$ gas mixture providing a gas gain of 10^4.

The requirements for the trigger systems are quite different from the tracker. It must be extremely fast, and efficiently trigger within the correct 'bunch crossing' (25 ns window). This is accomplished using Resistive Plate Chambers (RPCs) [96] and Thin Gap Chambers [97]. The RPCs are made up of pairs of large area bakelite sheets separated by a narrow gas region filled with tetrafluorethane in an electric field of 4.5 kV/mm. Discharges initiated by primary ionization electrons from tracks are read out through capacitive coupling on orthogonal (η, φ) read-out strips on the two sides of the detector. TGCs are made of effectively the same technology as the CSCs. They have gold-plated tungsten wires at 3100 V in a $\mathrm{CO_2}$ (55%) and n-pentane (45%) mixture, providing a gas amplification of 10^6. In the TGCs, both the strips and the wires are used for readout. The system is 99% efficient to trigger in one bunch-crossing. Both the RPCs and the TGCs also provide a rough position measurement in the direction orthogonal to that of the MDTs and CSCs, with an intrinsic accuracy of ∼10 mm.

3.2.2.4 Trigger System

In 2011 and 2012, the LHC operated with 1380 filled bunches and a 11 kHz orbit frequency, for a total event rate of 15 MHz. Events were written to disk at 400 Hz. The 2×10^{-5} rejection is performed using a three-level trigger implemented in hardware and software. The first level, Level 1 (L1), is implemented in hardware and reduces the total rate by a factor of 200 using coarse calorimeter trigger towers along with the RPCs and TGCs. This decision is taken within ∼2.5 μs of the event and triggers the readout of the data from the detector. The L1 also seeds the software-based Level 2 trigger with 'Regions of Interest' (ROIs) in which to reconstruct physics quantities (track and particles). The rate out of L2 is around 3.5 kHZ (a factor of 20 rejection). Finally, the Event Filter fully reconstructs events in software and provides the final factor of 10 reduction, to the output rate of 400 Hz.

Chapter 4

Photon Reconstruction and Identification

The methods for recording, reconstructing, identifying, and calibrating photons are the foundations of the search for the Higgs boson in the diphoton decay channel. The development of these methods was pursued in the context of the early SM prompt photon measurements [98] and refined for the $h \to \gamma\gamma$ search. This chapter describes these methods, with a heavy emphasis on photon identification.

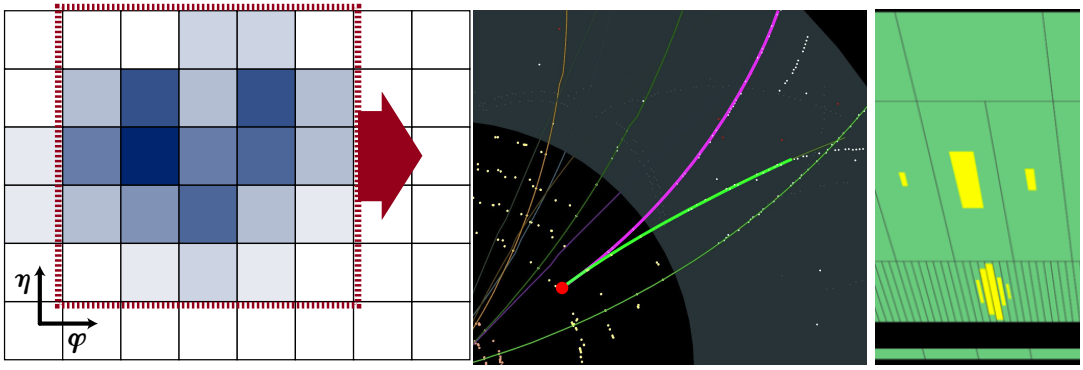
4.1 Photon Reconstruction

Photons and electrons are both triggered and reconstructed at ATLAS based on clusters in the electromagnetic calorimeter built with a 'sliding window' algorithm. In this algorithm, a rectangular window is shifted cell by cell across the calorimeter out to $|\eta| < 2.5$ (Figure 4.1a). At each position, the energy within the window is summed and local maxima that exceed 3 GeV are collected as 'preclusters.' The window dimensions of 5 cells square in $\Delta\eta \times \Delta\varphi$ results from an optimization that balances the efficiency of finding true clusters against the rate of false clusters from noise [72, 99].

At the trigger, a loose set of criteria are applied to ensure that the shapes of the shower in the calorimeter agree with the expectations for photons. These criteria will be described in Section 4.3.

Candidates without an associated track with $p_{\mathrm{T}} > 500$ MeV are classified as unconverted photons, while all clusters with associated tracks are tentatively classified as electrons. However, in the presence of matter a photon may convert into an electron-positron pair and these 'conversions' must be retrieved from the electron container. This is accomplished by searching for a conversion vertex along electron tracks, less than 0.8 m from the beam line. If two opposite-sign tracks are found with zero opening angle at their point of closest approach, the candidate is classified as a two-track conversion. Two dedicated algorithms were implemented to improve the efficiency for recovering late (vertex at $R \gtrsim 0.4$ m) or asymmetric (most energy carried by one track) conversions. First, 'back-tracking' uses calorimeter clusters as seeds, to look for tracks that originate deep in the ID. Second, 'single-track' conversions are identified by looking for a vertex along an electron track, effectively requiring that the track have no b-layer hit (but not requiring a second track). If any vertex is found, the precluster is classified as a converted photon. At the energy scale of the $h \to \gamma\gamma$ analysis, about one third of photons convert; of these, more than half have only a single associated track.

Once the precluster is classified a final cluster size is assigned, taking into account both the classification and its $|\eta|$ position in the calorimeter. Since a single unit of pseudorapidity is physically smaller in the endcap than in the barrel, clusters in the barrel measure 3 cells in $\Delta\eta$ while clusters in the endcap are allowed $\Delta\eta = 5$. The inner detector is immersed in a solenoidal field that curves the trajectories of electrons and converted photons, inducing bremsstrahlung that spreads their energy in the azimuthal direction. They are therefore allowed a broader window in $\Delta\varphi$ than unconverted photons: 7 cells instead of 5. Table 4.1 summarizes the cluster dimensions used for each type of object.



(a) Sliding Window Algorithm (b) Two-Track Conversion (c) Calo. Shower

Figure 4.1: (a) A sliding window of dimension $\Delta\eta \times \Delta\varphi = 5 \times 5$ is used to scan across the detector, for energy deposits that are used as to seed photon triggers. (b) The conversion of a photon into an electron-positron pair is shown in the inner tracker. (c) The energy deposited in the calorimeters by an unconverted photon has a single, narrow peak in the 'strip' layer of the liquid argon calorimeter, and should not leak into the hadronic calorimeter.

	Barrel	Endcap
Unconverted Photons	3×5	5×5
Converted Photons	3×7	5×5
Electrons	3×7	5×5

Table 4.1: Cluster sizes in units of one $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$ cell, for photons and electrons.

4.2 Photon Calibration

The energy of electromagnetic clusters is calibrated in three stages: (1) currents in calorimeter cells are transformed into energies, (2) energies are corrected for detector inhomogeneities and losses using a 'calibration hit' calibration, and (3) comparison of $Z \to ee$ events in data an Monte Carlo provides and an absolute scale and intercalibration.

4.2.1 Calculation of Cell Energies

The calculation of the energy in a single cell has two main factors: (a) a product $a_i\,(s_i - p)$ designed to extract a total pulse size from a series of 'samplings' (measurements) and (b) a unit conversion factor $f_{\mathrm{DAC}\to\ \mathrm{MeV}}$. Calorimeter cells are sampled at 40 MHz; if an event is triggered, five samples s_i of the current are preserved. The offset from zero is known as the pedestal p, and is first subtracted from the measurements. Each corrected sample $(s_i - p)$ the receives a factor a_i, which are known collectively as optimal filtering coefficients (OFCs). These factors (along with b_i, used for pulse timing) are calculated so as to minimize the total variance in the combined measurement, from electronic and pileup noise.

The unit conversion factor $f_{\mathrm{ADC}\to\ \mathrm{MeV}}$ is the product of four measured subfactors. First, the gain G_1 that 'converts' from analog to DAC counts is measured from fits of the electronics calibration. Next, $F_{\mathrm{DAC}\ \to\ \mu\mathrm{A}}$ and $F_{\mu\mathrm{A}\to\ \mathrm{MeV}}$ convert the electronic output to a physics quantity: $F_{\mathrm{DAC}\ \to\ \mu\mathrm{A}}$ comes from known properties of the calibration board, and $F_{\mu\mathrm{A}\to\ \mathrm{MeV}}$ is derived from Geant4 simulation and test beam measurements. Finally, differences between pulses from the calibration board or from physics are

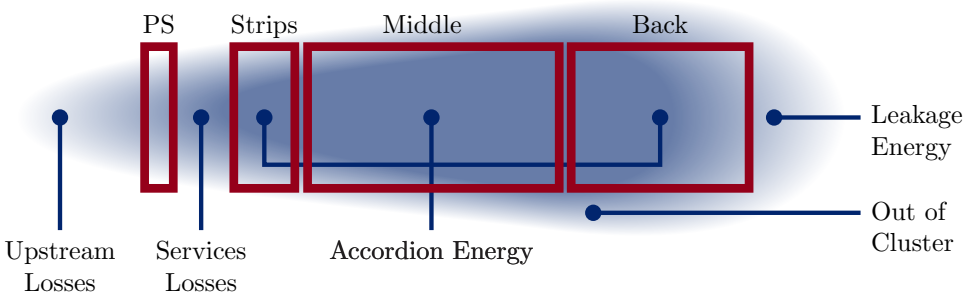


Figure 4.2: Schematic of a calorimeter cluster. The active volume of the electromagnetic calorimeter is boxed in red. The total energy of the cluster contains the sum of the energy lost before the calorimeter with the total accordion energy. Corrections must be applied for 'out of cluster' energy and 'leakage' out of the back of the calorimeter.

covered by a factor $(M_{\text{calib.}}/M_{\text{phy.}})$. In all, the energy in the cell is

$$E_{\text{cell}} = f_{\text{ADC}\to\text{ MeV}} \sum_i a_i \left(s_i - p\right) . \tag{4.1}$$

4.2.2 Corrected Cluster Energy

After individual cells are converted into 'physics energies,' they are assembled into the clusters described in Section 4.1 and displayed in Table 4.1. The total energy of the physics object includes the energy deposited both inside and outside the cluster. The energy outside the cluster is illustrated in Figure 4.2. It comprises upstream losses (before the calorimeter), leakage energy (behind the calorimeter), and lateral leakage (within the accordion but outside of the cluster window).

The energy before the calorimeter is expressed as a quadratic equation of the energy in the presampler a, b, and c that depend on both $|\eta|$ and the energy in the cluster. The energy beyond the presampler starts from the sum of the energies in each of the cells. That energy is then rescaled by the inverse of the fraction of the energy that is actually sampled by the active material, $s_{\text{acc.}}^{-1}$. Two additional factors cover the energy outside of the cluster: $(1 + f_{\text{out}})$ represents lateral leakage outside of the fixed size of the cluster while $(1 + f_{\text{leak}})$ represents leakage into the hadronic calorimeter.

A final factor $F(\eta, \varphi)$ multiplies the presampler, accordion, and leakage energies; it corrects for modulations in the response as a function of φ and η. The φ modulation arises from the regular structure in φ of the accordion geometry. The η modulation derives from the fact that a photon that hits the center of a cell will be better contained in the full cluster, while a photon towards the edge of a cell will also lose more energy at the boundary of the cluster.

The total energy may thus be expressed [72, 100, 101]

$$\begin{aligned} E_{\text{reco.}} = \Big[& a\left(E, |\eta|\right) + b\left(E, |\eta|\right) E_{\text{PS}}^{\text{clus.}} + c\left(E, |\eta|\right) \left(E_{\text{PS}}^{\text{clus.}}\right)^2 \\ & + \left(\frac{1 + f_{\text{out}}\left(X, |\eta|\right)}{s_{\text{acc.}}\left(X, |\eta|\right)}\right) \left(\sum_i E_i^{\text{clus.}}\right) \times \left(1 + f_{\text{leak}}\left(X, |\eta|\right)\right) \Big] \times F\left(\eta, \varphi\right) . \end{aligned} \tag{4.2}$$

Each of the constants (a, b, c) and functions (f_{out}, f_{leak}) are derived in a 'calibration hits' simulation, separately for single electrons, and unconverted and converted photons. This simulation makes it possible to account for all of the energy in the shower. These constants are derived with fine $|\eta|$ granularity, but are constant in φ. The $F(\eta, \varphi)$ modulation correction is also derived in simulation, but obviously is not symmetric in φ!

4.2.3 Residual Calibration from Data

The final piece in calibrating a photon is a determination of the overall scale in data. There are limited options for an absolute photons scale, so $Z \to ee$ events provide the standard candle. The Z mass is known to exquisite precision, from resonant depolarization measurements of the LEP beam [102]. Reconstructing the Z mass in data and simulation, and setting the observation to the known value, absorbs any remaining offset. The scale choices are moreover allowed to vary with $|\eta|$ which results in η-intercalibration, expressed $(1 + \alpha)$, per electron or photon. At this point, electrons and photons are treated identically.

Finally the Z lineshape in Monte Carlo is fixed to better match data, by broadening the momentum resolution through a 'smearing.'

4.2.4 Conversion Correction

A non-standard correction was developed in the $h \to \gamma\gamma$ analysis for converted photons: the radius $R_{\text{conv.}}$ of the conversion vertex affects the how much they interact with the detector. This leads to an $R_{\text{conv.}}$-dependent energy loss that is not accounted for in Equation 4.2. This dependence was measured in simulation and then corrected for.

4.3 Photon Identification

Jet production at ATLAS dwarfs photon production: the overwhelming majority of the objects reconstructed with the procedures described above are not photons or electrons, but jets. Without aggressive algorithms to select photons and reject jets, QCD dijet and photon-jet production swamp the diphoton production rate, and completely conceal the $h \to \gamma\gamma$ signal.

These backgrounds are rejected based on differences in the shapes of the showers induced by photons and jets in the electromagnetic calorimeter. Photons and electrons have narrow showers that are well-contained within the ∼20 radiation lengths of the liquid argon calorimeter. Quark- and gluon- induced showers are broader, and a significant fraction of their energy is typically deposited in the hadronic calorimeter. This Section and Section 4.4 describe the two pillars for selecting photons and rejecting jets: 'identification' algorithms (ID) that select photon-like electromagnetic showers, and 'isolation' measurements that quantify the amount of energy or number of tracks near a photon.

Two identification algorithms were used for the Higgs discovery:

(1) For 2011 data, a multivariate, neural network (NN) approach was developed to increase the selection efficiency by taking advantage of the full correlations between the variables.

(2) A cuts-based ID that uses boolean cuts on each variable, with η- and conversion-dependent values was originally used in 2011, and was completely revamped for 2012 data.

The author was responsible for both of these developments.

4.3.1 IsEM Variables

Photon identification at ATLAS is performed using ten 'isEM' variables that describe the shapes of the electromagnetic showers in the calorimeter. The isEM variables are illustrated in Figure 4.3.

Six variables are defined in the first 'strips' layer of the calorimeter:

- $f_1 \equiv E^{S1}/E^{\text{Tot.}}$ measures the fraction of the total shower energy deposited in the strips. Discrimination between the signal photons and QCD background is limited. For cuts-based ID, f_1 is used as a 'safety' check that *some* energy is deposited in the front of the calorimeter. It provides some discrimination to the NN.

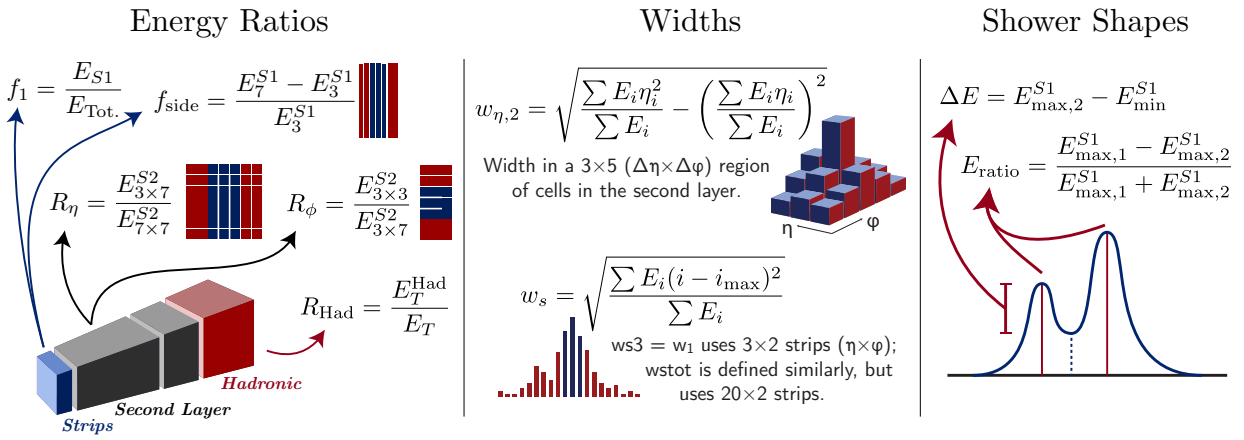


Figure 4.3: Schematic of isEM variables.

- ▷ $f_{\text{side}} \equiv \left(E_{7\times1}^{S1} - E_{3\times1}^{S1}\right)/E_{3\times1}^{S1}$ measures the fraction of the energy deposited away from the center of the shower. Signal-like showers tend to have lower f_{side} than the background.
- ▷ $w_{s,3} = w_1 \equiv \sqrt{\frac{\sum_i E_i(i-i\text{max})^2}{\sum_i E_i}}$ measures a 'width' in the strips using just the central bin and its neighbours ($i \in -1, 0, 1$). Signal showers tend to be more-narrow than the background.
- ▷ $w_{s,\text{tot.}}$ is defined identically to $w_{s,3}$, but uses 20 strips.

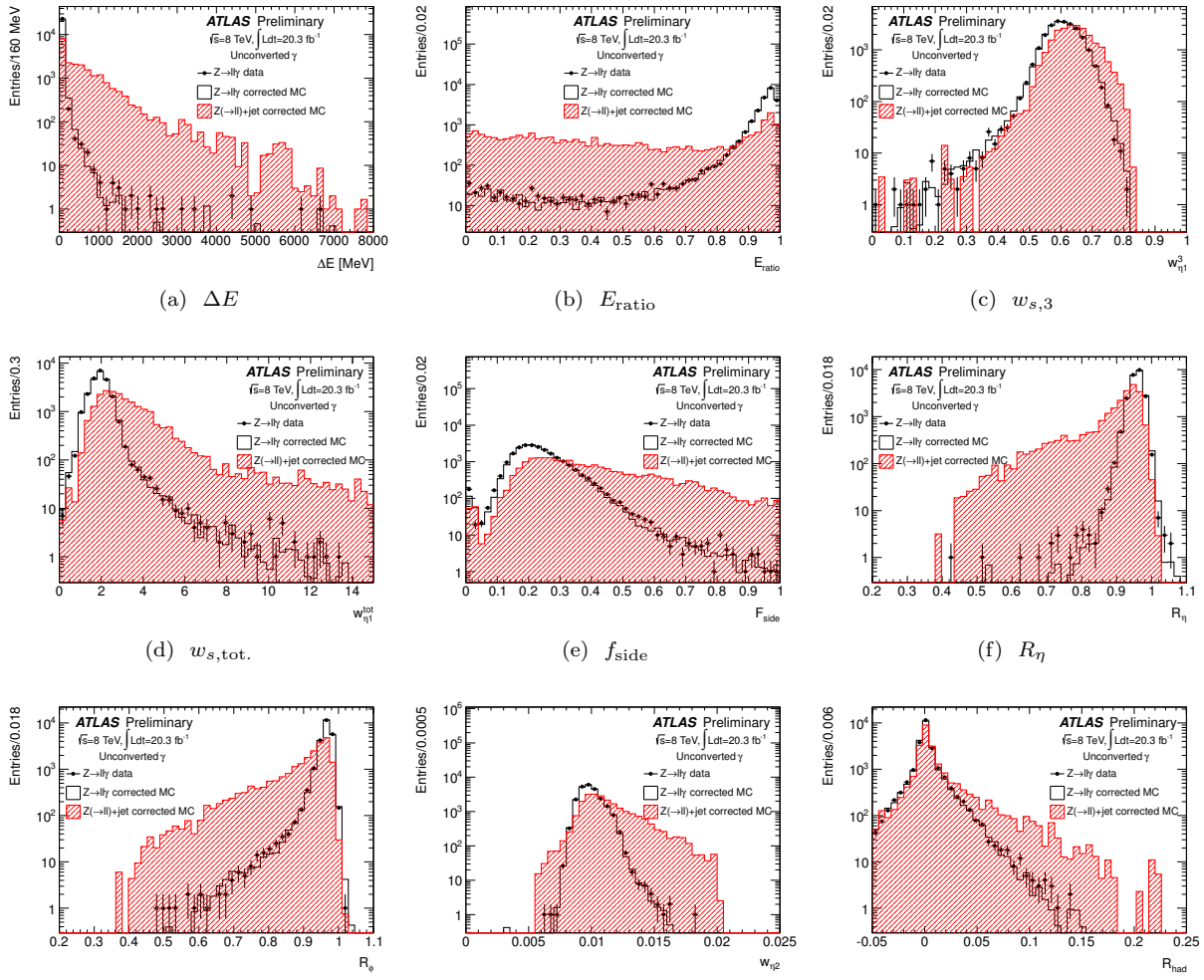
Neutral mesons that decay into photons ($\pi^0 \to \gamma\gamma$, $\rho \to \gamma\gamma$, etc.) leave clean, narrow showers in the calorimeter; since they *are* photons, their showers do not reach the hadronic calorimeter. They are distinguished from prompt photons only by a 'double peaked' signature left by the decay products in the finely-segmented strips. Two variables are defined to quantify this, for showers with two maxima:

- ▷ $\Delta E = E_{\text{max},2}^{S1} - E_{\text{min}}^{S1}$ is the height of a second maximum of the shower, over the minimum between the two maxima (see Figure 4.3). This value is small for prompt photons, since there is no meaningful second maximum.
- ▷ $E_{\text{ratio}} = \frac{E_{\text{max},1}^{S1} - E_{\text{max},2}^{S1}}{E_{\text{max},1}^{S1} + E_{\text{max},2}^{S1}}$ quantifies how evenly the energy is shared between the two maxima. It is close to one, for prompt photons.

Three variables are defined in the second layer of the calorimeter:

- ▷ $R_\eta = E_{3\times7}^{S2}/E_{7\times7}^{S2}$ takes the ratio of the energy in a $\Delta\eta \times \Delta\varphi = 3 \times 7$ set of 0.25 square cells centered on the photon cluster, over the energy in a 7×7 square. Real photons tend to be narrower, and hence have higher R_η.
- ▷ $R_\varphi = E_{3\times3}^{S2}/E_{3\times7}^{S2}$ is defined similarly to R_η, but in the φ direction. The differences between converted and unconverted photons are particularly large for this variable, since the solenoidal field separates converted photons in φ.
- ▷ $w_{\eta,2} = \sqrt{\frac{\sum_i E_i\eta_i^2}{\sum_i E_i} - \left(\frac{\sum_i E_i\eta_i^2}{\sum_i E_i}\right)}$ is a proper width defined using a $\Delta\eta \times \Delta\varphi = 3 \times 5$ cell region centered around the photon, in the second layer.

A final variable is defined using the energy in the hadronic calorimeter:



(a) ΔE (b) E_{ratio} (c) $w_{s,3}$

(d) $w_{s,\text{tot.}}$ (e) f_{side} (f) R_η

(g) R_φ (h) $w_{\eta,2}$ (i) $R_{\text{had.}}$

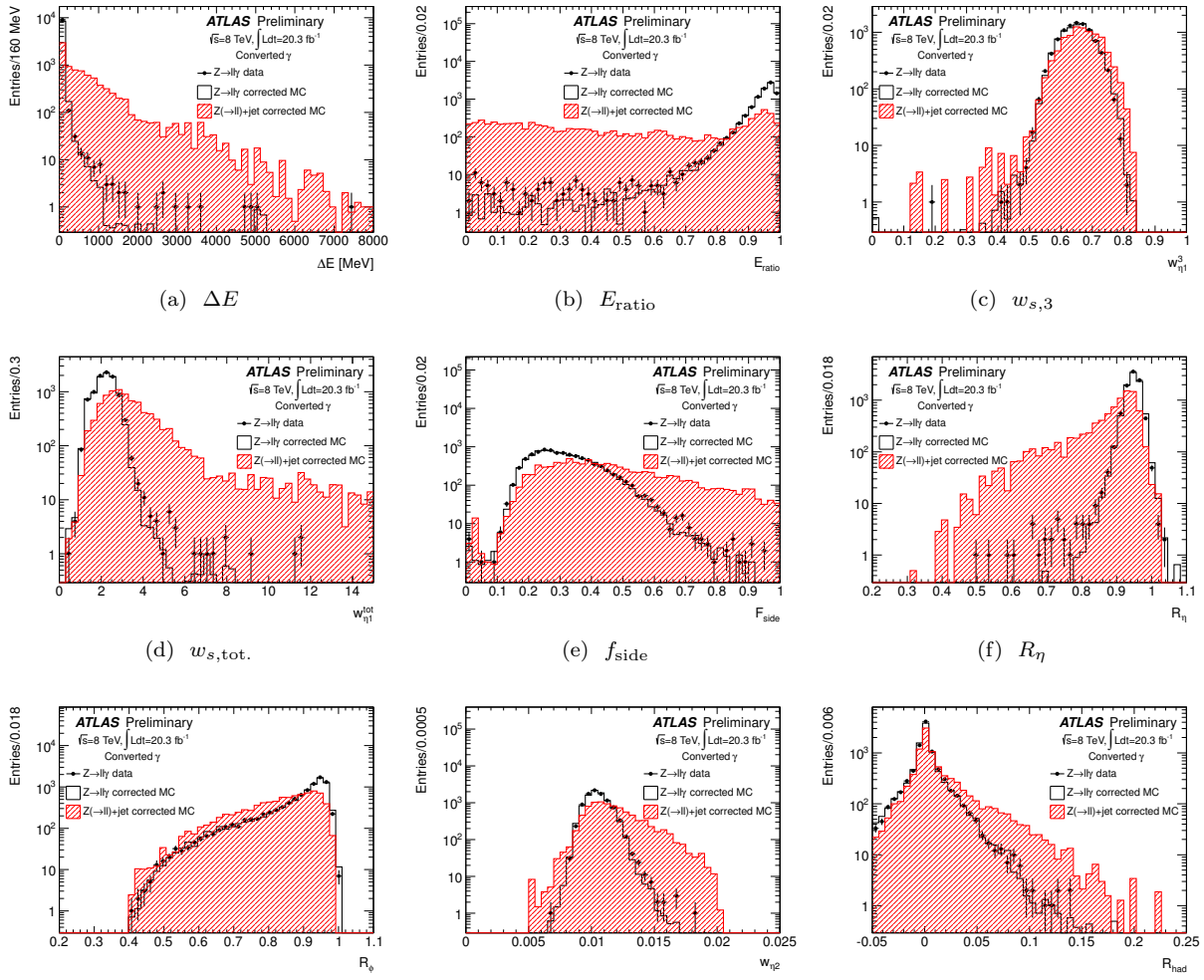
Figure 4.4: 'IsEM' variables for unconverted photons from $Z \to \ell\ell\gamma$ decays, from 2012 data.

▷ $R_{\text{had.}} = E_T^{\text{Had.}}/E_T$ is simply the ratio of the energy in a 0.24×0.24 window of the hadronic calorimeter behind the photon cluster, to the energy of the photon cluster. Since the liquid argon calorimeter is so deep ($> 20X_0$), the energy deposited in the hadronic calorimeter by real photons is centered at 0 (and can fluctuate negative). For $0.8 \leq |\eta| < 1.37$, the full energy in the hadronic calorimeter is used; elsewhere, only the first sampling layer of the hadronic calorimeter is considered and the variable is technically called '$R_{\text{Had},1}$.'

Sample distributions of the isEM variables are shown in Figure 4.4 and 4.5.

4.3.2 'Fudge Factors'

Geant4 nicely describes the photon showers' longitudinal development, but the transverse (η, φ) modelling is imperfect. Though the shapes of the isEM distributions described above are fairly well-described, they tend to be offset from the values observed in data. This has a several-percent impact on the efficien-



(a) ΔE (b) E_{ratio} (c) $w_{s,3}$

(d) $w_{s,\text{tot.}}$ (e) f_{side} (f) R_η

(g) R_φ (h) $w_{\eta,2}$ (i) $R_{\text{had.}}$

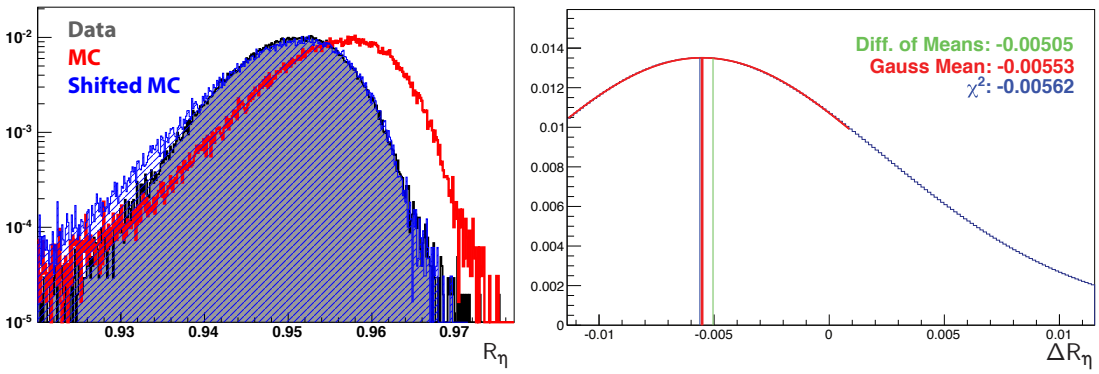
Figure 4.5: 'IsEM' variables for converted photons from $Z \to \ell\ell\gamma$ decays, from 2012 data.

cies. To correct both the input distributions and the output boolean is, to first order, simply to shift the distributions back – colloquially known as 'fudging.' The shifts themselves are known as 'fudge factors.'

In 2011, the shifts were defined as the difference of the means from data and MC for each discriminating variable (DV), in each $|\eta|$, conversion, and energy bin, i:

$$\Delta\mu^i_{\text{DV}} \equiv \langle DV^i_{\text{data}} \rangle - \langle DV^i_{\text{MC}} \rangle \ . \tag{4.3}$$

Yet the data sample to use is non-obvious. Indeed, the lack of a large, unbiased sample of pure photons is the perpetual challenge of photon efficiency optimization and measurements. It will be discussed in detail in Section 4.3.4.1. In the absence of such a sample, 2011 fudge factors were derived with a simple preselection: single photon triggers at 20, 40, 60, and 80 GeV; good quality, 'tight' photons (Section 4.3); and corrected Etcone40 isolation (Section 4.4) less than 5 GeV. This selection indeed enhances the purity in photons, but it is still heavily contaminated with jets. For the Monte Carlo, 'filtered' jet samples were therefore used, with a generator-level filter designed to emulate a loose photon trigger (JF17, JF35, JF70).



(a) Distribution pre- and post-Shift (b) Fit of Inverted χ^2

Figure 4.6: An example of the χ^2 shifting procedure with R_η, using unconverted photons with $60 < p_T \leq 80$ GeV and $1.81 < |\eta| \leq 2.01$. The data and pre-shifted MC distributions (a) are 'slid' past each other, generating a χ^2 curve that is inverted to give (b). Curve (b) is fit by a Gaussian, and the mean parameter of this fit is the derived 'fudge factor.' For most bins and variables, the agreement is good between the differences of means, the minimum 'χ^2', and the Gaussian mean.

There are two issues with the simple method defined above. First, the use of the means in Equation 4.3 makes the definition sensitive to the tails of distributions which pull 'harder' on the mean than changes in the core. The use of *tight* cuts does reject tails, but to the extent that it does so, it also dramatically sculpts the distributions. The distributions *end* at the value of the tight cut, and since this is done before applying any shifts, the fraction of the distribution that is included in the mean will be different for data and simulation. The second problem is that the composition of the sample (real photons v. jets) is not considered at all. A photon selection is used in data, but is expected to have large jet contamination; the MC simply uses jets. The composition affects the shape of the distribution and hence can impact the shifts. Nevertheless, this method was used for the 2011 'fudge factors.'

As an alternative to the 'difference of means,' a binned χ^2 difference method was implemented in the course of the 2011 and 2012 PID optimization. While the same isolation and object quality cuts are applied, the distributions used are slightly different from above: (1) 'tight' cuts are imposed only on those isEM variables that are uncorrelated from the variable under study, and (2) the purity of the data sample, in terms of real photons versus jets faking photons, is measured and emulated in Monte Carlo. With these distributions in hand, they are 'slid' past each other, and the offset that minimizes the χ^2 is taken as the shift. Technically, to derive this shift, the 'χ^2' distribution is inverted, and a Gaussian function is fit to the core of the inverted distribution. This is illustrated in Figure 4.6.

Many variants or permutations on these methods are possible. For instance, for the final 2012 fudge factors, the χ^2 method was used, but with selection criteria closer to that of 2011: simply requiring isolated photon candidates, and using the JF Monte Carlo samples instead of the purity-corrected mixture of jet and photon simulation. The agreement between these methods tends to be good.

4.3.3 Cuts-Based Identification and Trigger

Cut-based methods of photon identification use the union of a set of requirements (cuts) on the isEM variables to return a single boolean. These cuts have a substantial benefit of clarity. One can (in principle) isolate the impact of every individual cut – it is not a black box. The cuts are moreover separable: one can drop or apply cuts at will.

$\lvert\eta\rvert$		0-0.6	0.6-0.8	0.8-1.15	1.15-1.37	1.52-1.81	1.81-2.01	2.01-2.37
2011 Loose Cuts								
R_η	$>$	0.927	0.912	0.925	0.916	0.906	0.920	0.908
$w_{\eta,2}$	$<$	0.012	0.012	0.013	0.013	0.015	0.013	0.013
$R_{\text{had.}}$	$<$	0.010	0.010	0.010	0.008	0.025	0.015	0.014
2012 Loose Cut Menu								
R_η	$>$	0.8825	0.8825	0.8575	0.8875	0.8725	0.9025	0.8875
$w_{\eta,2}$	$<$	0.013	0.014	0.015	0.015	0.016	0.015	0.015
$R_{\text{had.}}$	$<$	0.02425	0.02275	0.02575	0.01975	0.02725	0.02725	0.02725

Table 4.2: Discriminant cuts for 'loose' photon identification in 2011 and 2012.

Loose, cuts-based photon identification is used in the trigger. Cuts are applied on three variables: R_η, $w_{\eta,2}$, and $R_{\text{had.}}$. These cuts are shared with the electrons. The cuts defined for both 2011 and 2012, are listed in Table 4.2. These cuts also constitute an element of the preselection of the $h \to \gamma\gamma$ analysis.

A tighter set of requirements on all ten isEM variables constitutes the final determination of 'what is a photon' in most ATLAS analyses. *Tight* cuts wered used offline for 2011 data before the introduction of the MVA method, and for all of 2012 data. The baseline cuts for 2011 are presented in Table 4.3. These served as the comparison point from which the neural network identification was optimized.

4.3.4 Neural Network Identification for 2011 Data

In 2011 and early 2012, the ATLAS Higgs boson search was statistically limited: increasing the number of identified prompt diphoton events while maintaining the rejection of photon-jet and di-jet backgrounds would directly improve the sensitivity of the analysis. This motivated the development of neural network identification. Just as with cuts-based ID, the signal is prompt photons, the background is jets, and the isEM variables are the input.

In general, multivariate techniques are designed to maximize the separation between two classes of objects by leveraging both the discriminating power of several input variables, and the correlations between them. Many algorithms and techniques have been developed to do this, which share the fundamental strategy of minimizing the error on a 'training sample' where the 'right answer' is known, by tuning the weights or parameters that define the network or decision tree. In principle, this minimization procedure 'automatically' converges to weights that provide the optimal separation between the two classes of objects. Yet multivariate techniques also raise new challenges. Because the training procedure is completely automated, the inner workings of the network are not transparent. While neural networks provide a smooth output discriminant – a 'knob' for tuning the efficiency versus rejection – they do not allow for the 'factorization' of cuts, for instance cutting only on variables from the second layer or strips of the calorimeter.

This section describes: (1) tests used to check the reliability of the isEM variables in (fudged) Monte Carlo, (2) the mechanics of the neural network, and (3) tests of systematic uncertainties on its performance.

4.3.4.1 Validation of Input (isEM) Distributions

Simulated photon decays were ultimately used for the training samples of the NN; while this sidesteps the problem of finding a pure signal sample for training, it immediately confronts the question of the validity of the MC. This thus alters the problem but does not eliminate it: rather than finding sources of pure *photons*, one must instead find sources of relatively pure isEM *distributions* to validate the description of the MC.

$Z \to \ell\ell\gamma$ Tag and Probe. For electrons and muons, $Z \to \ell\ell$ decays provide an easy, unbiased source of pure leptons. Events are selected with two like-flavor, opposite-charge leptons that together reconstruct the Z

mass. If either leg of that decay candidate can be 'tagged' by satisfying tight identification cuts, then the other is expected to be a real lepton: this provides a clean, sample to serve as a 'denominator' for efficiency measurements, as well as unbiased (e.g., isEM) distributions of pure leptons.

Things are not so simple for photons. The sole available kinematic tagging process is leptonic Z decays with final state radiation[2] ($Z \to \ell\ell\gamma$ with $m_{\ell\ell\gamma} \sim m_Z$), whose rate is far smaller than $Z \to \ell\ell$.

In validating the NN inputs, events were selected from the full 2011 dataset of 4.8/fb that satisfied the data quality criteria of the SM and W/Z working groups (SMWZ Good Run List v. 36). The selection criteria were:

- ▷ At least one single lepton or dilepton trigger. The p_{T} threshold of the lowest single muon trigger was 18 GeV, while the lowest single electron trigger was at 20 GeV. Several dielectron triggers had thresholds at $p_{\mathrm{T}} > 12$ GeV.
- ▷ At least one primary vertex with at least three tracks.
- ▷ Two opposite-charged, like flavor leptons with $p_T^{\ell} > 15$ GeV.
  - – Electrons: *Tight* identification and good object quality.
  - – Muons: Good muon quality, based on the 1.02 fb^{-1} SM WW paper [103].
- ▷ A photon candidate with $\Delta R\,(\ell, \gamma) > 0.2$ and `Etcone40_corrected` < 5 GeV.
- ▷ Fiducial $|\eta|$ cuts imposed for all particles.
- ▷ Mass Cuts: $40 < m_{\ell\ell} < 83$ GeV, $80 < m_{\ell\ell\gamma} < 96$ GeV.

For 2011 data, there were just ∼12000 selected events, which consist primarily of unconverted, soft photons. The isEM distributions from these photons are included in the Appendix. The distributions are shown with greater statistics from 2012, in Figures 4.4 and 4.5.

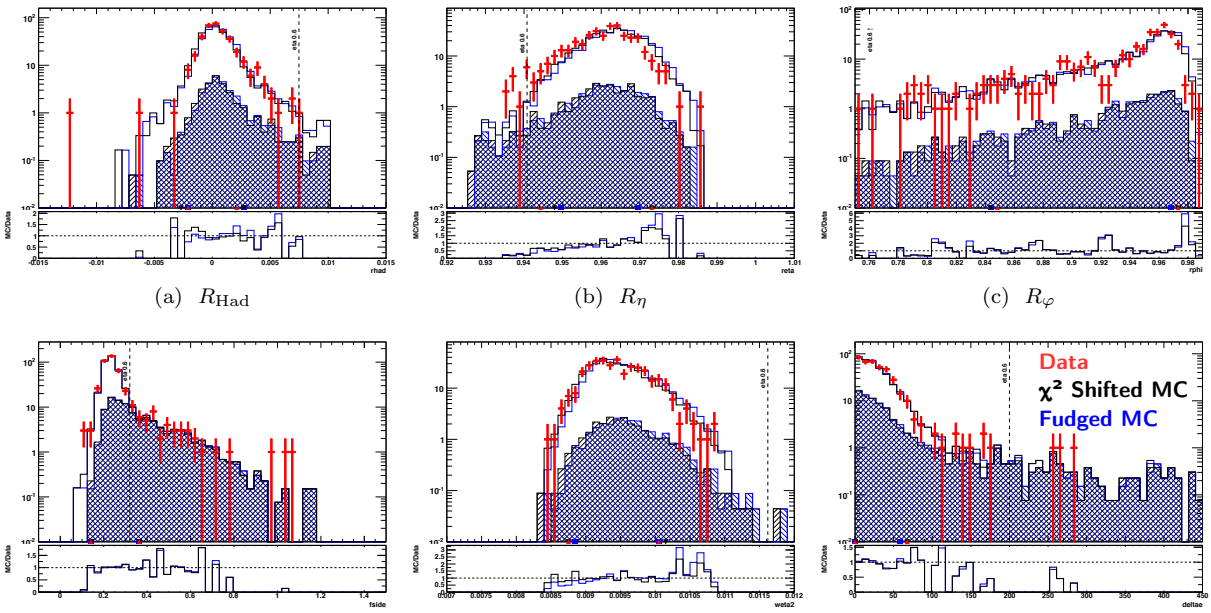
Leading (di)Photons. For events with two reconstructed photons, the higher-p_T (leading) photon is 'fairly' likely to be a real photon. This is because the detector response to hadronic activity is lower than for photons, and more of the jet activity is likely to fall outside of the small window defined for photons. This means that on average, jets reconstructed as photons will be 'missing' energy. If a photon and a jet are simply recoiling off of eachother with opposite true p_{T}s, the one with the higher measured energy is more likely to be the real photon.

To select these 'leading photons' the same preselection was used as for the 2011 $h \to \gamma\gamma$ analysis (Section 5.1). The subleading photon was required to satisfy the *tight* selection criteria, and its isolation in `Etcone40_corrected` was required to be less than 3 GeV. Since these samples were used only to validate the description of the isEM variables in simulation, for each variable i, tight cuts were also apllied to the leading photon, for all uncorrelated variables j ($|\rho_{ij}| < 0.4$).

Yet substantial jet contamination remained (∼15%). To address this, the purity in photons of the data sample was measured, using the methods described below, and this purity was matched in Monte Carlo by adding the appropriate fraction of jet MC to the photon MC. A sample of the resultant distributions are shown in Figure 4.7. Post-shifting, the Monte Carlo description of the isEM variables is good.

Other Methods. Several other methods indirectly validate the description of the inputs and the performance of the NNs. *Loose*, isolated photons with $p_T > 100$ GeV tend to be quite pure, but this is beyond the typical kinematic space of $h \to \gamma\gamma$ decays. Alternatively, the isEM distributions are defined in the same way for photons and electrons; since it is 'easy' to select $Z \to ee$ decays, these may be compared in data and MC. This method will be used as a systematic check of the efficiency measurement. Other methods for validating the results – the output instead of the inputs – are described later.

[2]Selecting initial state radiation ($m_{\ell\ell} \sim m_Z$) leads to a larger jet contamination.



(a) R_{Had} (b) R_{η} (c) R_{φ}

(d) f_{side} (e) $w_{\eta,2}$ (f) ΔE

Figure 4.7: Data is compared to Monte Carlo for the leading photon after a diphoton selection. After applying the standard $h \to \gamma\gamma$ preselection, *Tight* cuts uncorrelated to the variable in question are applied. The purity of the data sample is measured, and an appropriate fraction of jet MC is reinjected (hashed). In this example, the photons are converted and have $|\eta| < 0.6$ and $50 < p_{\mathrm{T}} < 60$ GeV. Monte Carlo is shown using bothe χ^2-shifts and 'difference of means' fudge factors.

4.3.4.2 Description of the Neural Network and its Training

Binning the Network. The kinematics (p_T, η) and conversion status of a photon strongly influence the shape of its electromagnetic shower, and the values of the isEM variables that describe it. For example: the amount of material in the detector changes as a function of η, showers become more collimated at higher energy, and only converted photons are bent by the magnetic field. To assist the neural nets in taking advantage of these features, the analysis was reproduced for 126 bins: 3 conversion categories (unconverted, single track, and two tracks), 7 $|\eta|$ bins, and 6 p_T bins:

$$p_T : [20, 25)\,, [25, 40)\,, [40, 50)\,, [50, 60)\,, [60, 80)\,, [80, 1000) \ \text{GeV}$$
$$|\eta| : [0, 0.6)\,, [0.6, 0.8)\,, [0.8, 1.15)\,, [1.15, 1.37)\,, [1.52, 1.81)\,, [1.81, 2.01)\,, [2.01, 2.37)$$

Choice of Method. An important early consideration in the implementation of a multivariate discriminant is the choice of the method itself. The Toolkit for Multivariate Analysis (TMVA) package implements a wide variety of methods, making it possible to simply 'try out' and tune many different alternatives: Binary Decision Trees (BDTs), likelihoods, multilayer perceptrons (MLP; also neural nets, NN), or binary cuts-based discriminants. The manual for that package contains a detailed description of the defaults of each method [104]. The efficiency versus rejection curves were compared for each alternative; an example of this is shown in Figure 4.8. The efficiency is defined as the fraction of loose, isolated photon candidates matched to a true photon, that satisfy the discriminant cut; rejection is the fraction of the loose, isolated photon candidates that come from non-prompt photons, and which are rejected by the cut. Perfect separation is

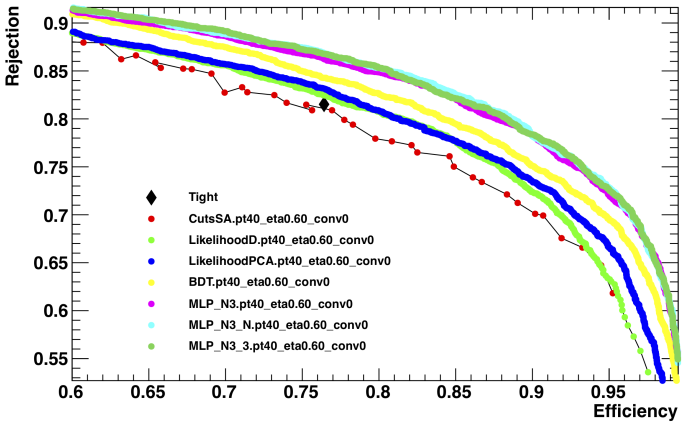


Figure 4.8: Photon efficiency v. jet rejection for nominal cut-based identification, TMVA-optimized cuts, a binary decision tree, likelihoods, and a multilayer perceptron. The efficiency and rejection are shown with a denominator of loose, isolated photon candidates.

thus at the point $(1, 1)$, so the discriminants whose curves stretch further towards the upper right corner perform better.

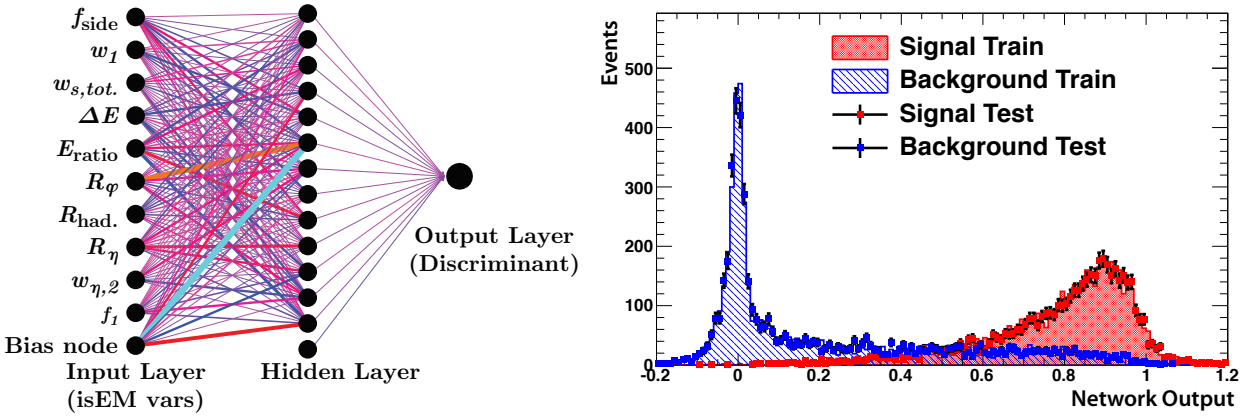
Each method was optimized to a nominal level, in an attempt at a 'fair comparison' between them. Particular attention was paid to cuts-based identification, to see that it at least matched the nominal 2011 identification (Table 4.3). The likelihood methods perform somewhat better than the cuts-based identification, but worse than the BDTs. The multilayer perceptron showed the best performance, with gains in rejection of around 5% at fixed efficiency or, vice versa, gains in efficiency of up to 10% at fixed rejection. It was selected for further optimization and study.

Multilayer Perceptron and its Training. A multilayer perceptron consists of a network of nodes with activation functions, a collection of weights for the connections between those nodes, and a method for choosing the weights. An abundant theoretical literature is available on the topic [105, 106], but the highlights are transcribed here for interest.

The net begins with N input variables x_i along with a constant 'bias' value x_0. Together these $N + 1$ *nodes* form the *input layer*. Every node in the input layer is connected to each of M nodes in the subsequent *hidden* layer, with a weight w_{ji}. This produces M weighted sums $\sum_i w_{ji}x_i$, each of which is processed by an activation function g_j. The output of these M nodes and another bias node are similarly linked with weights w_{kj} to every node in the *following* layer, and so forth. The final layer provides the discriminant. The implementation used in the present instance, which has a single hidden layer, is displayed schematically in Fig. 4.9 and may be expressed algebraically as

$$d_k = g_k \left(\sum_{j=0}^{M} w_{kj} \times g_j \left(\sum_{i=0}^{N} w_{ji}^{(0)} x_i \right) \right) . \tag{4.4}$$

Sigmoidal activation functions are generally preferred by theory, but hyperbolic tangents were used: they converge faster, and they can be mapped into sigmoids by two linear transformations, one to the input and the second to the output. (Using the tanh is thus equivalent to changing the bias weights or scaling the input.) Three layers of weights are theoretically sufficient to model any separation function, but it can be shown that with a sigmoidal activation function and a finite number of inputs and outputs, two layers of weights can approximate this to arbitrary precision.



(a) Network Schema (b) Output Discriminant

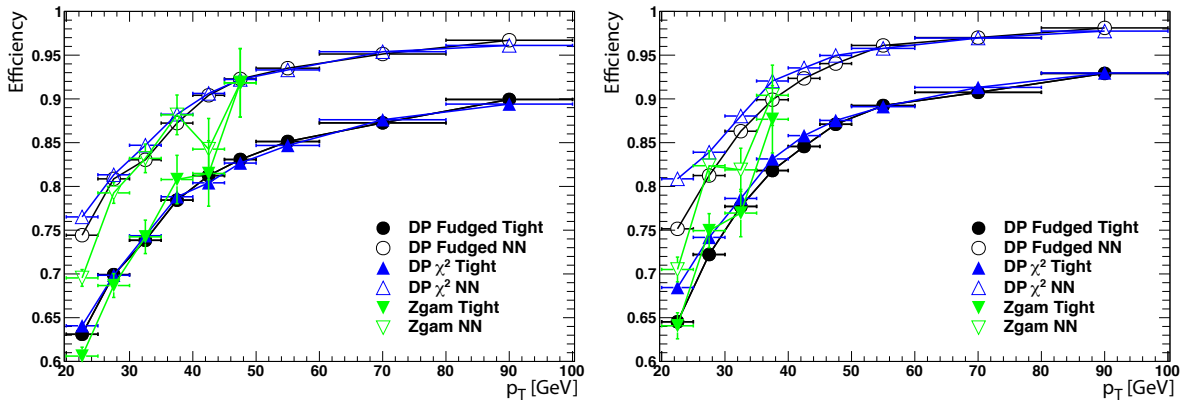
Figure 4.9: Schematic of the neural network used, with one hidden layer with $N + 3 = 13$ nodes. Every input variable is connected to each node in the hidden layer with weights denoted by the thicknesses and shades of the connections (shown here is the $20 < p_T \leq 25$ GeV, $|\eta| < 0.6$, unconverted bin). Each of these nodes feeds in turn into the final layer, with yet another weight. The discriminant output is optimized to send signal photons to 1, and background (QCD) to 0.

Various configurations were created, with either one or two hidden layers, and anywhere between 1 and $2N$ nodes per layer (with $N = 10$ isEM variables). The performance of these methods was largely equivalent, but among the best options were two hidden layers with $(N + 3, N)$ nodes or $(N + 3, 3)$ nodes, and a single hidden layer with $N + 3$ nodes. The single-layer net was selected for its simplicity.

The weights were derived using the standard back-propagation algorithm, which minimizes the error of the network with respect to its weights. The error on the network is calculated as the sum of the errors of all objects used in the training sample: the differences of the discriminant to the target values (0 for background and 1 for signal). Back-propagation is a method for determining the gradient of the total error with respect to the weights. Each iteration of the training steps along the steepest descent towards the minimum, at a learning rate of 0.02 (the TMVA default). Smaller learning rates did not significantly improve the ultimate performance of the nets. The algorithm was set to terminate after 20 cycles with no improvement (convergence).

A final concern is 'over-training' the sample – creating a network that is tuned to the fluctuations of the training sample rather than the 'true characteristics' of the signal. There are two ways for handling this problem. The first is to use an 'infinite' input sample with negligible fluctuations; the second is to define a completely separate 'test sample' and ensure that iterations on the training sample (to which the back-propagation is applied) continue to yield gains in that second sample. The nominal size for both the testing and training samples was 5000 signal photons, and 5000 background objects. Out of 126 bins, 11 had fewer events than that, the lowest at 2101 testing and 2101 training events.

Choice of Discriminant Cuts. Because the event level $h \to \gamma\gamma$ analysis did not employ multivariate techniques, the final task in implementing an identification algorithm is to define the cut – the value that the discriminant must exceed for a shower to be classed as *photon*. A number of methods were explored for this, in attempts to maximize the overall $h \to \gamma\gamma$ expected significance. Ultimately however, none were found to yield significantly better performance than simply setting the discriminant cuts to achieve the same level of rejection as the 2011 cuts menu. This simple method was therefore retained.



(a) Unconverted (b) Converted

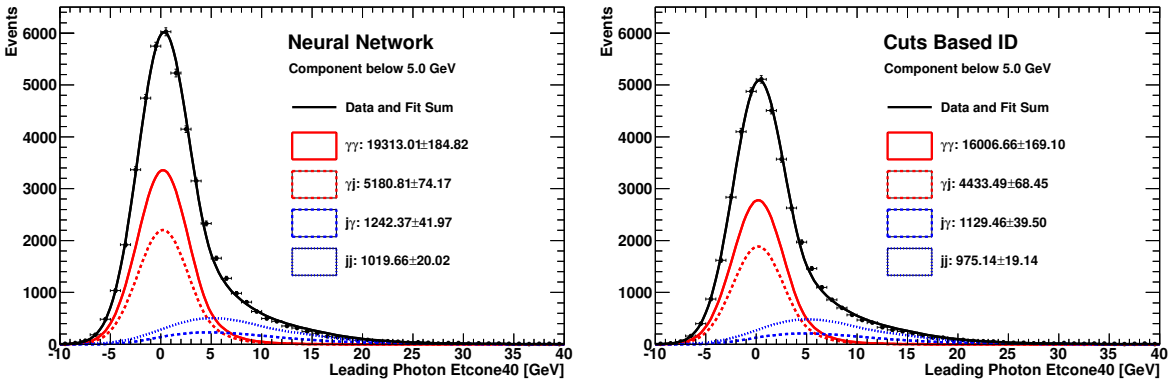
Figure 4.10: Neural net and *Tight* photon efficiencies in direct photon (DP) Monte Carlo, with the 'difference of means' or χ^2 shifting (Section 4.3.2), and in data from from $Z \to \ell\ell\gamma$ decays. Also shown are efficiencies of photons from $Z \to \ell\ell\gamma$ decays in data. The denominator are isolated photons, while the numerator imposes loose cuts along with either NN or *tight* PID.

4.3.4.3 Results

Efficiencies in Monte Carlo. The simplest check of gains from the use of the neural network is to look at the impact in Monte Carlo. Efficiencies of unconverted and converted photons are presented as a function of their p_{T} in Figure 4.10, using MC11B Direct Photon (DP) samples with generator-level p_{T} cuts at 17, 35, and 70 GeV. The denominator is container-level truth-matched, isolated photons from both bremsstrahlung and hard process, and in the numerator *loose* cuts are applied with either NN or *tight* PID (since the loose ID is applied at trigger level, this could not be removed). The isEM input variables are shifted ('fudged') using either the 'difference of means method', or the χ^2 shifts described above. This straightforward test shows good agreement between the two MC shifting methods, and large gains for the MVA identification.

Diphoton Yield and Purity in Data. Increased photon identification efficiency has a particularly large impact on the yield of the diphoton selection used for the $h \to \gamma\gamma$ analysis, since any changes to the efficiency or rejection apply to both selected objects. The parameter of interest is the number of selected diphoton events: the product of the overall yield and the purity in diphoton events. In early data, several methods were defined to extract the diphoton purity of a set events. All methods share a fundamental strategy: using the isolation as an auxiliary measurement, assumed to be uncorrelated to the identification. Very roughly, the isolation is the amount of energy or the number of tracks in a $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2}$ cone around the photon; it will be described in greater detail, in Section 4.4. If the isolation spectrum (or the efficiency of a cut) is known for true photons and background objects, one can define a simultaneous signal plus background fit (or solve a system of equations) to derive the purity.

In the following, a two-dimensional fit is performed in the calorimeter isolation of leading and subleading photons. Separate templates are defined for diphoton, photon-jet (p_{T}-ordered), jet-photon, and dijet events, and the normalizations of those components are allowed to vary in the fit. The isolation profile of the background comes from requiring at least one of ΔE, E_{ratio}, w_1 or f_{side} to fail *tight* cuts, but requiring all other variables to pass. This is often called *loose'*. This isolation profile has been shown to fit well to a Novosibirsk function. Two techniques have been used for the isolation profile of the signal. First, under the assumption that no 'true photons' have isolation above some 'large' threshold of 7 or 10 GeV, one can normalize the background template to the spectrum of tight photons above that value, and subtract it off



(a) Neural Network ID (b) Cuts Based ID

Figure 4.11: Two-dimensional isolation template fit projected along the leading photon isolation. We find comparable purities for *Tight* and neural net identification (71% and 72% respectively), but substantially higher efficiency for the neural nets.

to derive the signal component. The template used in what follows uses the second alternative. Electrons from Z decays provide a clean source of electromagnetic objects, with isolation profiles extremely similar to those of photons. Because they can be selected so purely, no background subtraction is required. However, the isolation profiles of electrons and photons are not identical – they are found to be shifted with respect to each other, in both data and Monte Carlo. When using templates derived from electrons, the peak is accordingly allowed to float in the fit. Both signal templates are described analytically by a Crystal Ball function.

Armed with the two analytical one-dimensional signal templates, the $\gamma\gamma$, γ-jet, and jet-γ templates are simply their products. The jet-jet template was a smoothed 2D adaptive kernel estimation PDF where both leading and subleading photons satisfied 'reversed-cuts.'

For the final fit, the full 2011 dataset is used, that satisfies the selection defined for the Spring 2012 version of $h \to \gamma\gamma$ analysis [107]. Figure 4.11 displays the projection along the isolation of the leading photon. The purity of the analysis improves slightly, when using the neural nets – from $71 \pm 1\%$, to $72 \pm 1\%$. The efficiency, however, jumps dramatically – by $21 \pm 1\%$. This compares to an increase of $17 \pm 1\%$ seen in Monte Carlo. This agreement is not perfect, particularly considering that the MC is totally pure in photons and should have a larger increase than data, which has fairly large (25%) contamination of the photon-jet and dijet background processes. On the other hand, the processes are different: the $h \to \gamma\gamma$ process produces more photons towards central η than the standard diphoton process. The NN gains tend to be smaller at at lower $|\eta|$.

A less-optimistic explanation is that the purity measurement is failing due to correlations between the neural net and the isolation. The difference in isolation between *(loose' and ∼NN)* and *NN* datasets is larger than between *(loose' and ∼tight)* and *tight*. This is not in itself a problem, but if the correlations are not taken into account, it may lead to an over-estimate of the purity of the sample.

4.3.4.4 Efficiencies from Data, and Systematics

Given the heavy use of Monte Carlo in deriving the neural networks, it is critical to evaluate their performance in data. To do this, the pure sample from the $Z \to \ell\ell\gamma$ described in Section 4.3.4.1 is supplemented with a 'Matrix Method' that uses a system of equations with track isolation as a second discriminant [108],

and a 'Smirnov Transform' that builds 'photon' objects out of electrons and thereby extrapolates the efficiency. More details on these methods and their statistical combination is given for tight, *cuts-based* PID in Ref. [109]. The results from this combination for the neural net is presented in Fig 4.12. The efficiencies do not exactly agree with Figure 4.10 because a different denominator is used there.

The overall systematic uncertainty for the PID efficiency is assigned based on the difference between the efficiencies from the fudged MC and the combined data-driven results. A 4% uncertainty is taken everywhere, except 5% for unconverted photons with $1.52 < |\eta| \leq 1.81$ and 7% for unconverted photons with $1.81 < |\eta| \leq 2.37$. At the time it was measured, the uncertainty on the NN ID was smaller than the *tight* PID! However, continued studies of *tight* PID that culiminated in scale factors to cover the data/MC disagreement ultimately wittled the cuts-based uncertainties lower again.

4.3.4.5 Additional Studies

Distorted Material Maps. Before data-driven comparisons were mature, the primary systematic on photon ID came from the difference in the efficiencies measured in nominal simulation using the 'best knowledge' detector geometry, and a simulation using a 'distorted' geometry with additional material. These differences are fairly consistent with respect to p_T and $|\eta|$, and are within 3% except for converted photons at high $|\eta|$. The neural net provides more consistent behavior between the two samples than the tight cuts, as shown in Figure 4.13.

Electron Efficiencies in Data and MC. Given the difficulty in deriving pure photon samples, electrons serve as useful additional check on the consistency of the efficiencies between data and Monte Carlo. All 2011 data are used, and compared to the 2011 Pythia $Z \to ee$ simulation (data set 106046). The isEM distribution of the electrons in simulation are 'fudged' using the shifts derived for converted photons, and evaluated as 'single-track conversions' on both the cuts menu and the neural nets. The efficiencies in data and Monte Carlo are binned in $|\eta|$ and included in Figure B.1 of the appendix; the critical numbers – the *differences* between the efficiencies are shown in Figure 4.14. For the neural networks, the differences are typically less than 2%, compatible with or smaller than *Tight*. In the largest $|\eta|$ bin, the differences for the neural nets are larger than those for *Tight*, but still within 10%. Within the kinematic range used for the $h \to \gamma\gamma$ analysis ($p_T^{\text{lead}} > 40$ GeV and $p_T^{\text{subl}} > 30$ GeV), the difference is always within 6%. This is interpreted as a confirmation of the reliability of the efficiency modeling of the Monte Carlo, and does not motivate any additional uncertainty.

Pileup Dependence. In part simply because its efficiency is higher than the cuts-based ID, the NN ID is seen to be less-sensitive to pileup, as illustrated in Figure 4.15.

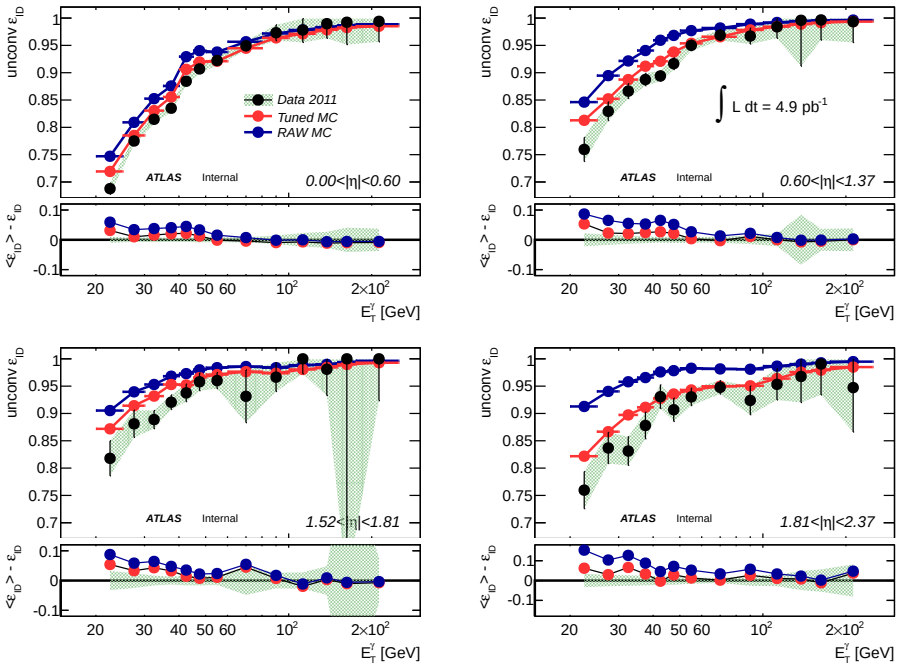
4.3.5 Cuts-Based Optimization for 2012 Data

For early 2012 data, a conservative approach to photon identification was required, that could be quickly validated and understood in the rapidly converging search for the Higgs boson. To do this, a decision was made to refine the cuts-based algorithm that had been used earlier in 2011. These cuts were needed on a very short time scale; they had to be less sensitive to pileup and more efficient than previous menus.

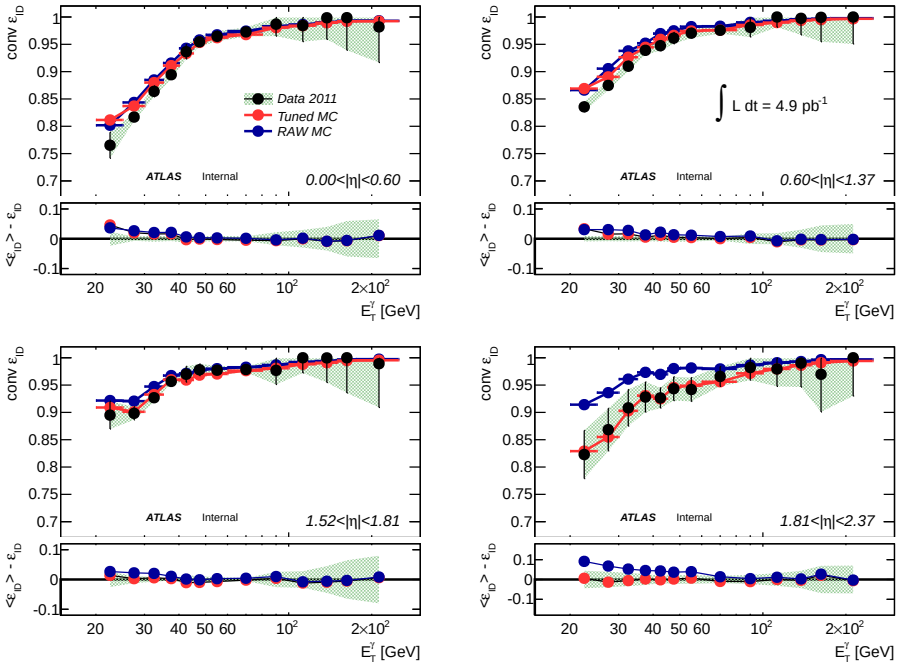
The optimization of cuts-based particle identification is dramatically different from the development of a neural network. Instead of 'trusting' the inner workings of a neural net and working to validate the inputs and outputs, the majority of the work on cuts-based ID comes from tuning the cuts one by one – in this case, 140 of them. The final result of this work described in this Section – the final menu of cuts used for 2012 – is presented in Table 4.3, along with the 2011 cuts.

4.3.5.1 Studies of Potential Systematics

At the beginning of 2012 data-taking, the increasing number of interactions per bunch crossing was a major concern. One of the preliminary steps in the optimization of a new menu was to understand how the

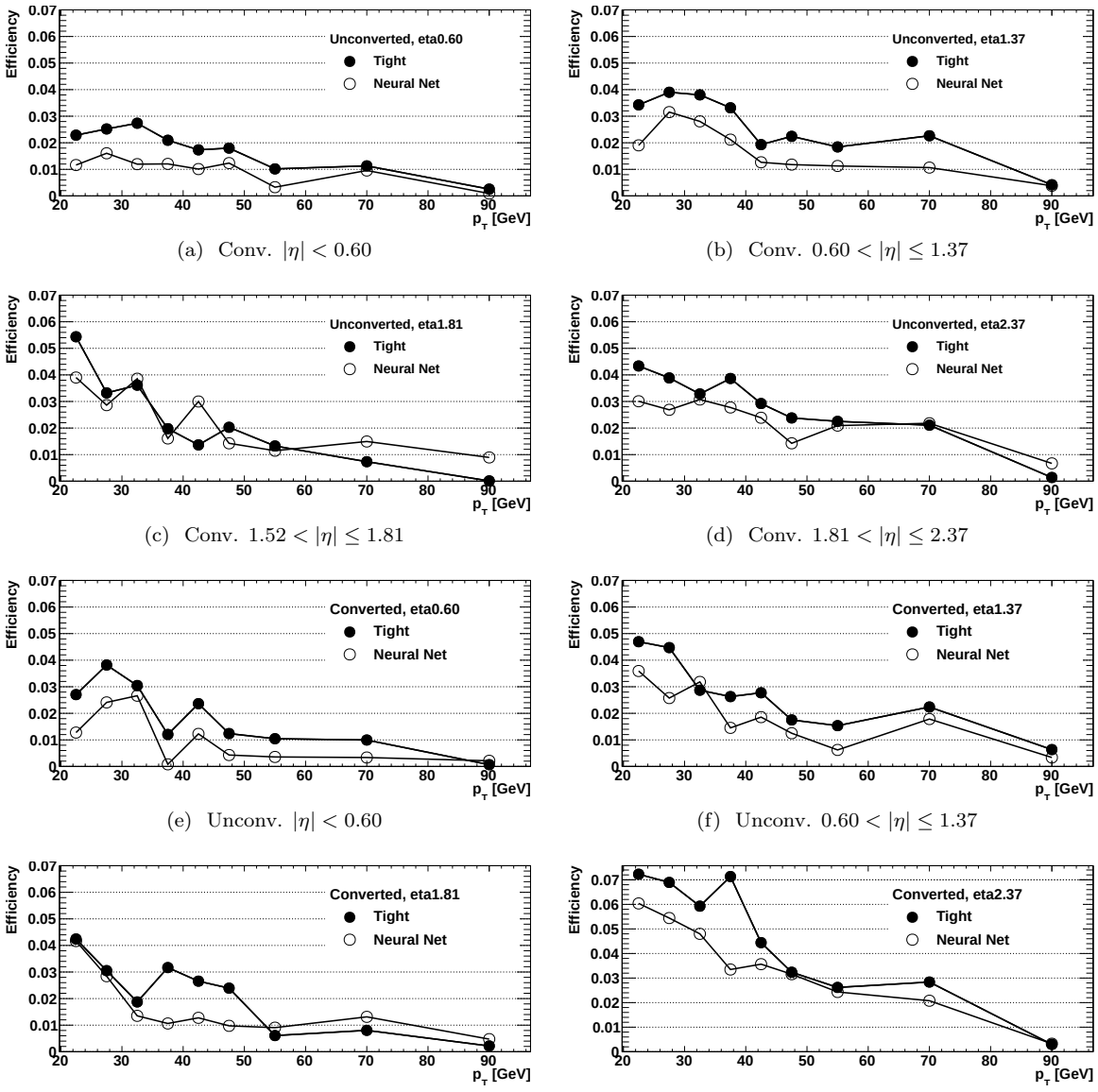


(a) Unconverted Photons



(b) Converted Photons

Figure 4.12: Neural network identification efficiencies are presented for unconverted and converted photons. The data-driven measurements use the combination of $Z \to \ell\ell\gamma$, electron extrapolation, and matrix method results described in Section 4.3.4.4. These are compared to raw Monte Carlo, and 'tuned' MC after fudging has been applied. The difference between the fudged Monte Carlo and the data-driven methods is taken as the systematic uncertainty.



(a) Conv. $|\eta| < 0.60$

(b) Conv. $0.60 < |\eta| \leq 1.37$

(c) Conv. $1.52 < |\eta| \leq 1.81$

(d) Conv. $1.81 < |\eta| \leq 2.37$

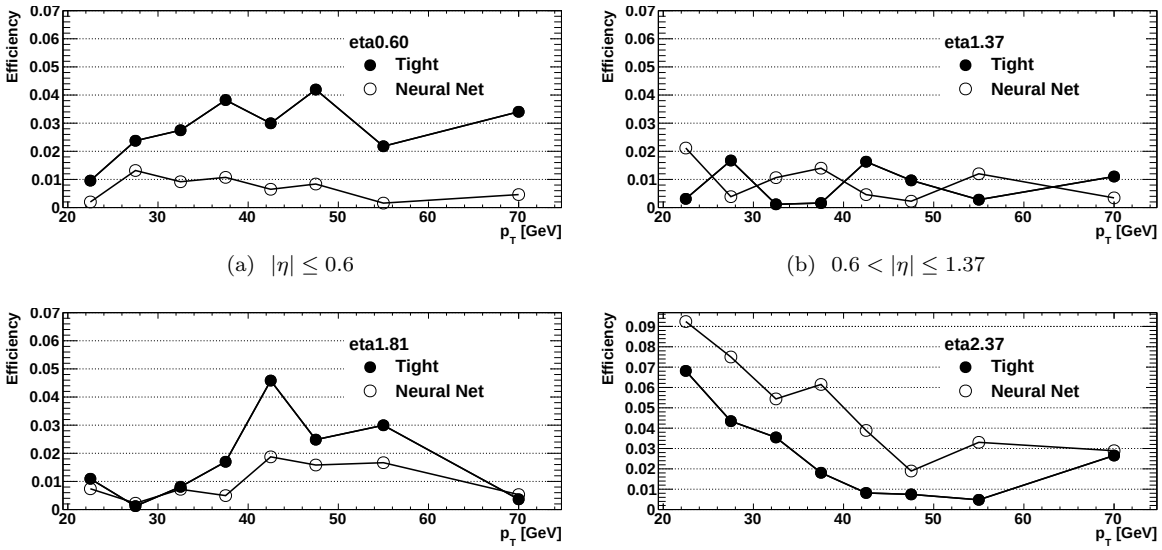
(e) Unconv. $|\eta| < 0.60$

(f) Unconv. $0.60 < |\eta| \leq 1.37$

(g) Unconv. $1.52 < |\eta| \leq 1.81$

(h) Unconv. $1.81 < |\eta| \leq 2.37$

Figure 4.13: Absolute value of the differences between the *tight* and neural network efficiencies for nominal and distorted samples. This difference was originally used as the systematic on identification algorithms. The neural net is overall less sensitive to this change than the cuts-based ID.



(a) $|\eta| \leq 0.6$ (b) $0.6 < |\eta| \leq 1.37$

(c) $1.52 < |\eta| \leq 1.81$ (d) $1.81 < |\eta| \leq 2.37$

Figure 4.14: Differences between Monte Carlo and Data for electrons evaluated on Neural Nets and the *Tight* cuts method. The NN is no more sensitive than *Tight*.

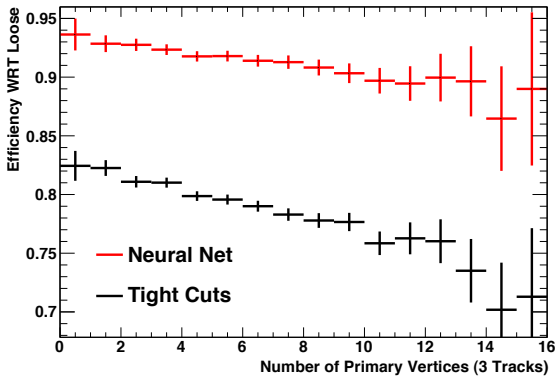


Figure 4.15: The photon identification efficiency is shown for cuts-based and NN ID, as a function of the number of vertices. The slope is shallower for NN ID, partly because it is simply more efficient overall.

$\lvert\eta\rvert$		0-0.6	0.6-0.8	0.8-1.15	1.15-1.37	1.52-1.81	1.81-2.01	2.01-2.37
2011 Tight Cut Menu: Unconverted Photons								
f_1	$>$	0.05	0.05	0.05	0.05	0.05	0.05	0.05
f_{side}	$<$	0.3492	0.4252	0.4087	0.5627	0.7920	0.3177	0.2877
$w_{s,3}$	$<$	0.6716	0.7016	0.7115	0.8245	0.7460	0.6695	0.6285
$w_{s,\text{tot.}}$	$<$	2.9608	4.4108	3.2712	3.4112	3.8282	2.4203	1.6603
ΔE	$<$	90	90	95	107	101	102	140
E_{ratio}	$>$	0.6347	0.8447	0.8273	0.8913	0.8768	0.7075	0.7775
R_η	$>$	0.9481	0.9371	0.9386	0.9426	0.9262	0.932	0.9188
R_φ	$>$	0.9543	0.9503	0.5876	0.8176	0.9280	0.9432	0.9312
$w_{\eta,2}$	$<$	0.0120	0.0120	0.0113	0.0114	0.0120	0.0130	0.0127
$R_{\text{had.}}$	$<$	0.0089	0.0070	0.0060	0.0080	0.0190	0.0150	0.0137
2011 Tight Cut Menu: Converted Photons								
f_1	$>$	0.05	0.05	0.05	0.05	0.05	0.05	0.05
f_{side}	$<$	0.3968	0.5048	0.5826	0.6096	0.6632	0.3301	0.2931
$w_{s,3}$	$<$	0.7077	0.7197	0.7632	0.7942	0.7913	0.6926	0.6646
$w_{s,\text{tot.}}$	$<$	2.8158	2.9658	2.9143	3.1643	3.7331	2.0218	1.5018
ΔE	$<$	197	197	119	83	120	76	128
E_{ratio}	$>$	0.9109	0.9139	0.8043	0.7993	0.6618	0.9183	0.9653
R_η	$>$	0.9384	0.9244	0.9252	0.9262	0.9109	0.9320	0.9150
R_φ	$>$	0.3920	0.4180	0.4851	0.4291	0.5297	0.4814	0.6944
$w_{\eta,2}$	$<$	0.0120	0.0120	0.0130	0.0130	0.0150	0.0123	0.0132
$R_{\text{had.}}$	$<$	0.0075	0.0070	0.0049	0.0080	0.0149	0.0150	0.0110
2012 Tight Cut Menu: Unconverted Photons								
f_1	$>$	0.05	0.05	0.05	0.05	0.05	0.05	0.05
f_{side}	$<$	0.28	0.33	0.38	0.425	0.42	0.255	0.24
$w_{s,3}$	$<$	0.67	0.69	0.69	0.715	0.72	0.66	0.645
$w_{s,\text{tot.}}$	$<$	3.0	3.0	3.3	3.5	3.3	2.3	2.0
ΔE	$<$	180	170	165	160	425	500	560
E_{ratio}	$>$	0.80	0.80	0.76	0.82	0.78	0.80	0.80
R_η	$>$	0.92	0.92	0.93	0.925	0.925	0.925	0.910
R_φ	$>$	0.93	0.93	0.93	0.92	0.93	0.93	0.93
$w_{\eta,2}$	$<$	0.011	0.0115	0.0115	0.0115	0.012	0.012	0.0128
$R_{\text{had.}}$	$<$	0.020	0.020	0.01975	0.01825	0.02425	0.02575	0.02325
	$<$	0.01825	0.01975	0.01525	0.01675	0.02125	0.02275	0.01975
2012 Tight Cut Menu: Converted Photons								
f_1	$>$	0.05	0.05	0.05	0.05	0.05	0.05	0.05
f_{side}	$<$	0.33	0.38	0.46	0.52	0.52	0.31	0.25
$w_{s,3}$	$<$	0.73	0.715	0.74	0.75	0.75	0.69	0.66
$w_{s,\text{tot.}}$	$<$	2.8	2.9	3.1	3.3	3.5	2.2	1.8
ΔE	$<$	160	160	120	125	350	520	525
E_{ratio}	$>$	0.85	0.85	0.80	0.78	0.82	0.86	0.88
R_η	$>$	0.92	0.9125	0.915	0.91	0.908	0.917	0.903
R_φ	$>$	0.57	0.60	0.60	0.64	0.68	0.72	0.72
$w_{\eta,2}$	$<$	0.011	0.0117	0.012	0.0120	0.0130	0.012	0.0127
$R_{\text{had.}}$	$<$	0.020	0.018	0.01975	0.018	0.02425	0.024	0.024
	$<$	0.01825	0.01975	0.01525	0.01675	0.02125	0.02275	0.01975

Table 4.3: Discriminant cuts for 'tight' cuts-based photon identification in 2011 and 2012.

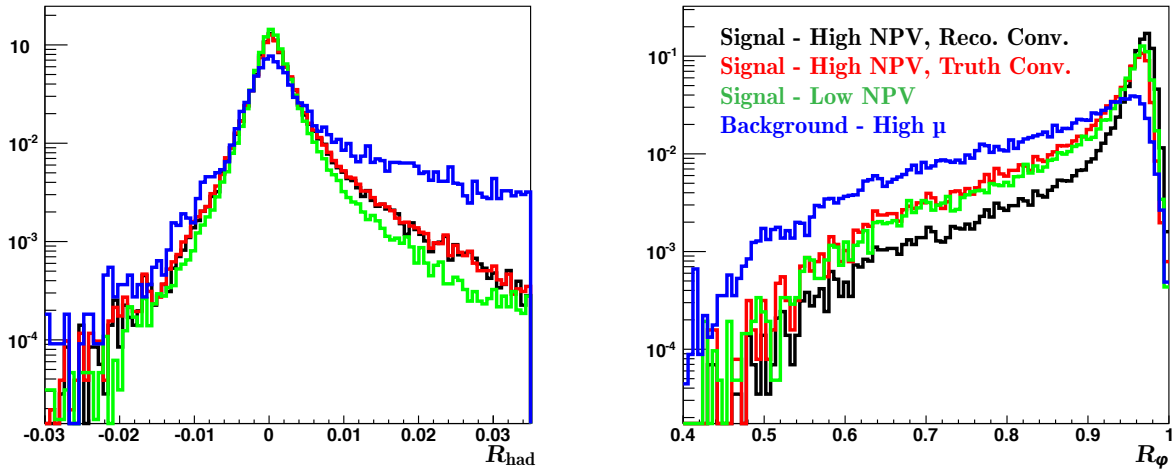


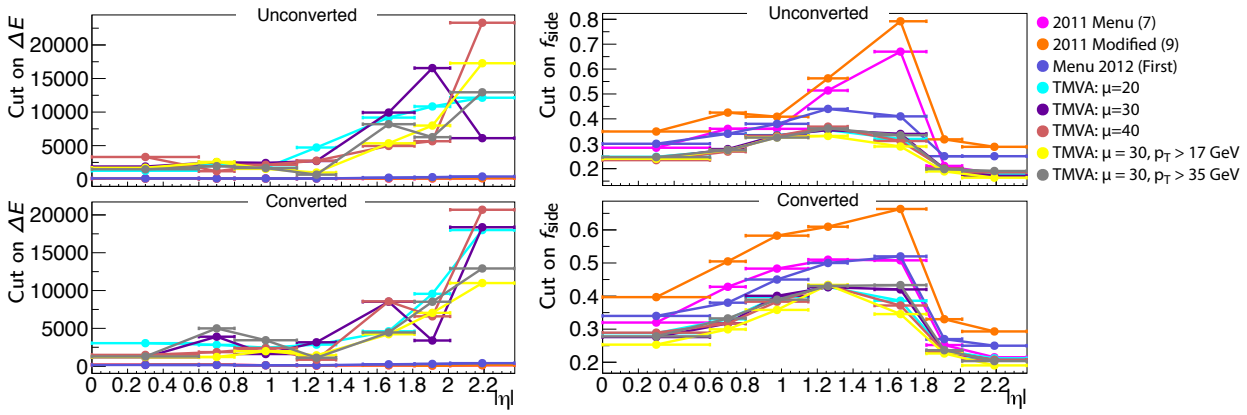
Figure 4.16: Degradation of the isEM variables with pileup. R_{φ} and $R_{\text{had.}}$ are shown for converted photon candidates with $|\eta| < 0.6$. Signal processes are divided among large and small numbers of primary vertices ($NPV > 20$, $NPV < 12$), and backgrounds are shown for comparison. R_{φ} and $R_{\text{had.}}$ are two of the variables most-heavily impacted by pileup, but for different reasons. $R_{\text{had.}}$ suffers from *real* additional energy in the hadronic calorimeter, whereas the degradation of R_{φ} arises from increased misclassification of the conversions at larger pileup. It is clear that for large numbers of primary vertices, R_{had} degrades regardless of the method used to classify the photons. R_{φ} on the other hand, only changes significantly if the reconstructed conversion classification is used – if the classification is lifted from the truth record, this effect disappears.

isEM variables were impacted by pileup, and what other parameters can affect their shapes.

Early in the investigations, two potential factors were removed from these considerations. The photon production mechanism – whether from the hard process or bremsstrahlung – did not affect the isEM shapes of isolated photons. Second, the isEM shapes were not found to correlate strongly to the photon isolation, confirming that for real photons, the small-scale shower description 'factorizes' from the larger isolation profile. This is also important because the isolation definitions have continued to evolve, and a strong-dependence would demand separate efficiency systematics or even different cuts menus for each isolation prescription.

The largest impacts on the isEM values of a photon are its $|\eta|$ position, conversion status, and transverse momentum, along with the pileup in the detector. The first two parameters are binned in the optimization, so they are of no concern. The benchmark p_{T} spectrum was set to roughly match the one used for the $h \to \gamma\gamma$ analysis. (An obvious improvement to the cuts would be to bin them in p_{T}, but in the interest of time and simplicity this was not done.) The dependence on pileup is a serious one and the cuts on variables with the largest pileup dependence were set conservatively, keeping the cuts far from the cores of the distributions.

The variables with the largest pileup dependence are R_{had}, R_{η}, and R_{φ}, but these are for different reasons. The R_{had} and R_{η} dependences might be considered a 'genuine' degradation of the isEM shape: energy not-associated to the photon, deposited at random in the calorimeter diminishes the discriminating power of the variable. On the other hand, the pileup dependence observed in R_{φ} was found to arise primarily from misclassification of photons by the reconstruction algorithm. Unconverted photons were often misclassified as converted photons, leading to a narrower distribution than expected for the conversions. The impact of the pileup *alone* was the reverse – to slightly broaden the R_{φ} width! These effects are illustrated in Figure 4.3.5.1. This misclassification was significantly mitigated by improved conversion reconstruction for 2012 ('Gaussian Sum Filtered' tracking).



(a) ΔE (b) f_{side}

Figure 4.17: The TMVA 'simulated annealing' was used to generate cuts 'automatically.' TMVA opts simply to ignore ΔE (it made heavy use of E_{ratio}, which is strongly-correlated), which is perhaps flawed 'advice.' On the other hand, it used f_{side} more heavily than past cuts menus. These automatically-generated menus were used as a landmark for the starting point of the optimization.

4.3.5.2 'Landmark' Cuts Menus

Given the size of the space to be 'scanned,' the next challenge was simply to find a starting point – cuts from which to iterate. As mentioned above, the TMVA package has a 'Simulated Annealing' routine that can quickly calculate 'optimal' cuts to select signal or reject background with a target efficiency. Cuts were averaged over five neighbouring efficiency optimizations (each set of cuts is fixed, and does not vary continuously) to reduce fluctuations. This hardly yielded a final solution: the cuts derived varied dramatically as a function of $|\eta|$ and conversion status, and they often came close to the core of the distributions of pileup-sensitive observables. Examples of the cuts suggested by TMVA are shown in Figure 4.17.

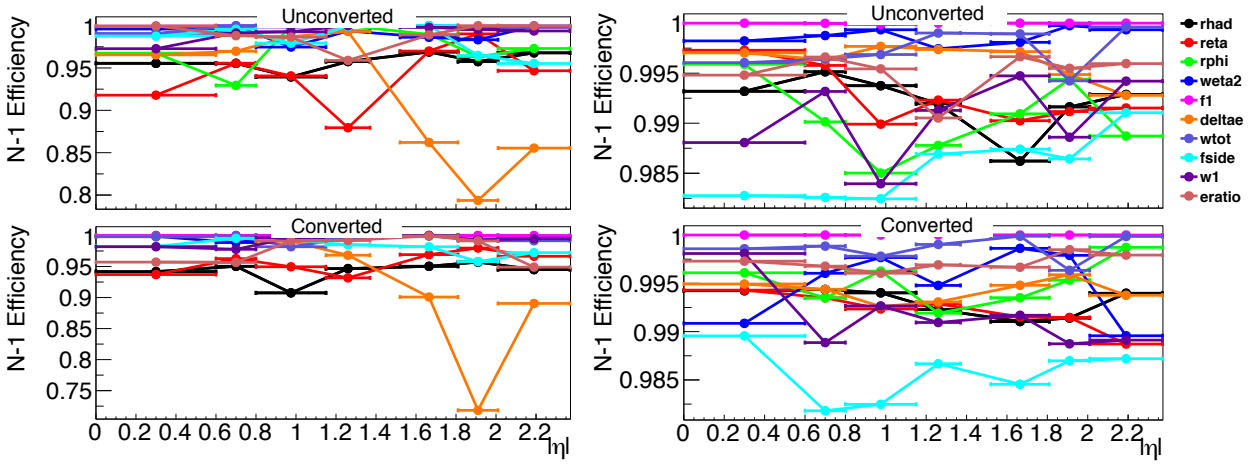
The 'quality of the advice' was also inconsistent between variables. For example, TMVA completely ignored ΔE, which was perhaps 'unwise;' on the other hand, the menus it produced relied much more heavily on f_{side} than earlier versions of the cuts, which was quite reasonable. Overall, in conjunction with the earlier cuts menus, they were a useful 'landmark' for guiding the initial optimization.

4.3.5.3 Tools for Refining Menus

As hinted above, the bulk of the work for the final 2012 menu came in painstaking refinement of each individual cut. A series of metrics were developed to guide the optimization; these fall into two categories – 'maps' of the big picture and 'microscopes' of the individual distributions. Used together, they made it possible to rapidly identify problems with the full menu, and then dive to a low-level to solve them.

The most important overall maps were the efficiencies of each single cut, either as the first (single-cut) or last ($N-1$) cut applied. This quickly presented any losses, making it easy to fix them. The final copy of this plot – after all 2012 optimization – is shown in Figure 4.18. As shown, no single cut has an $N-1$ inefficiency larger than 2%. Reducing the $N-1$ inefficiencies as far as possible was critical to removing the pileup sensitivity of the overal menu: by placing cuts far from the core of each variable, changes in the distributions – whether from pileup or from mismodelling in the simulation – had small impacts. This performance is illustrated in Figure 4.18, where the expected performance with pileup is shown.

For cuts requiring attention, 'microscopes' were prepared as in Figure 4.19. Individual distributions are shown for signal and background distributions, along with the cuts values and their efficiencies as a function



(a) 2011 Cuts Menu, on 2011 High-μ Sample

(b) 2012 Proposed Menu, on 2012 MC Signal

Figure 4.18: Efficiencies are presented for each variable listed, as the last cut applied $(N-1)$, as a function of $|\eta|$ for unconverted and converted photons. The large – and unnecessary – losses due to E_{ratio} and R_η in the 2011 menu stand out, and are quickly corrected. The extensive work came in ensuring that *no* single cut had an $(N-1)$ inefficiency of more than 2%.

of $|\eta|$, for reference. The loose cuts applied at trigger were also displayed; since the tight menu was applied on top of the loose cuts and trigger, these cuts were treated as boundaries that could not be crossed. Altogether, this put each cut in context, and made for relatively fast iterations.

Yet high efficiency and low pileup dependence are obviously incomplete metrics: they are no use at all, if the background is not also diminished. The final factor to check was whether the rejection was 'large enough' for the overall menu. An example of this is shown in Figure 4.20 for converted and unconverted photons, though in practice, this was divided along $|\eta|$ as well. This afforded a comparison of the current rejection to the rejection used in 2011. Since it was not a priori known how large the rejection would have to be for greater than around 75% of the selected events to be be real diphotons, two completely separate menus were prepared: 'Menu 2012' as a baseline and 'Menu 10' as a back-up with much larger rejection. Ultimately, the rejection of the 2012 menu was found to be sufficient.

4.3.5.4 Early Data: New Optimal Filtering Coefficients

No new data conforms exactly to expectations, so early collisions in 2012 demanded extremely fast iterations to finalize the photon ID. As described in Section 4.2.1, the energy in each calorimeter cell is calculated using 'optimal filtering coefficients' designed to minimize the combined pileup and electronic noise on the measurements. Because of a large increase in the amount of pileup, these coefficients were recalculated before the start of 2012 operations. This change in the energy calculation naturally impacted the isEM distributions.

The first order of business was to full recalculate the 'fudge factors' described in Section 4.3.2, to be able to use the Monte Carlo. Once this was done, any cuts could be refined. The most-impacted variables were w_1 and f_{side} at high-$|\eta|$; small changes were made to these cuts.

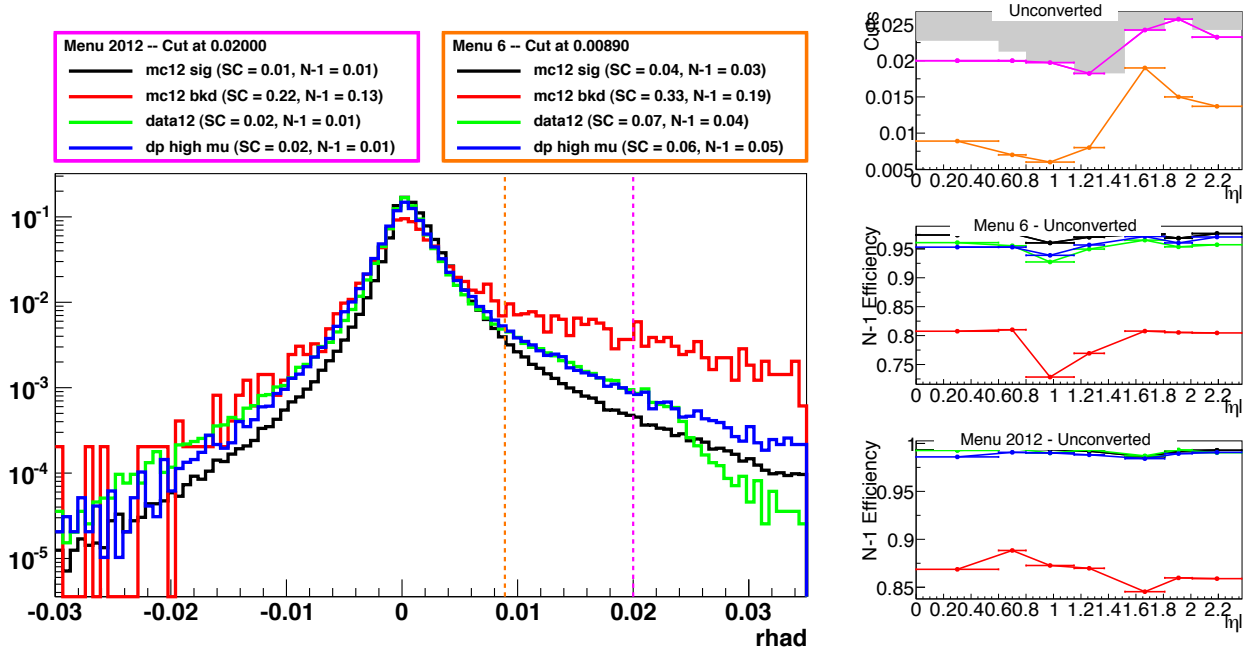


Figure 4.19: The distribution of R_{had} is shown for signal and background processes, for unconverted photon candidates with $|\eta| < 0.6$, after all other cuts. The cut values of the proposed menus are shown as well. At left, reference plots provide context of the cut values and $N-1$ efficiencies as a function of $|\eta|$. The grey shaded region of the top right plot shows the values of the loose cuts used at trigger level.

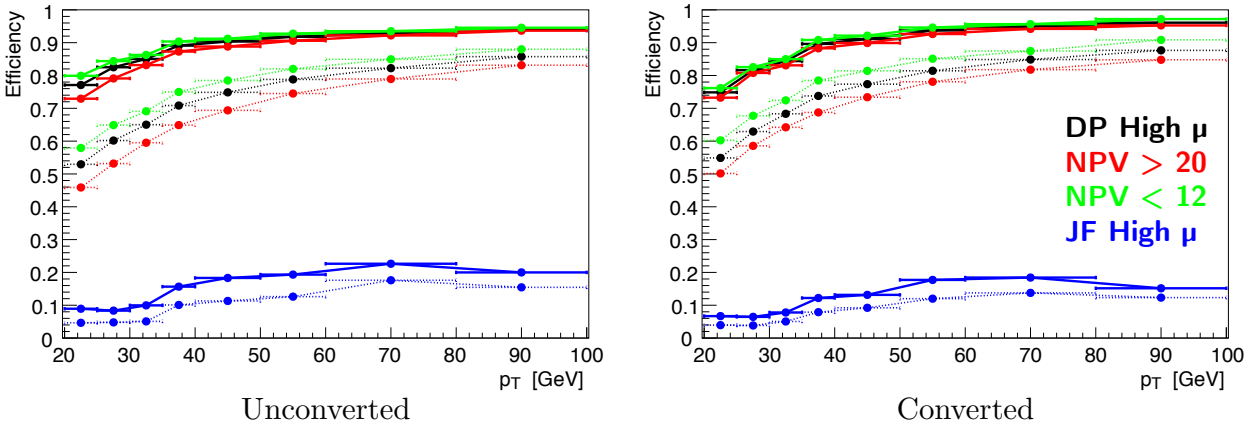


Figure 4.20: Efficiencies in simulation of the (nearly completed) cuts menu as a function of p_{T}. Solid curves show the 2012 cuts menu while dashed lines show the 2011 cuts menu. The spread between low- and high-pileup samples is much smaller for the new menu.

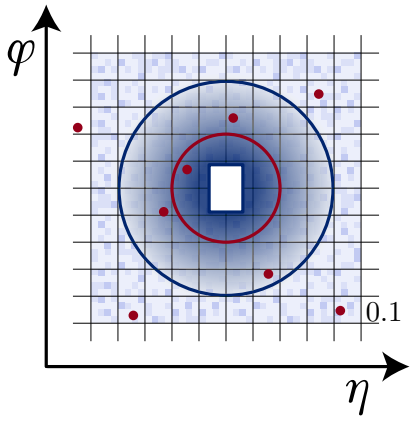


Figure 4.21: In the $h \to \gamma\gamma$ analysis, both calorimeter (blue) and track isolation (red) are used. As shown, a core around the photon must be removed for the calorimeter isolation. Leakage of the photon out of the core, and any pileup must be removed. Track isolation provides a confirmation of the calorimeter isolation from the inner tracker. Because tracks can be associated to vertices, it is inherently pileup resistant. For conversions, it is easy to remove the track(s).

4.3.6 Measurements of Identification Efficiency

Measurements of the photon identification efficiency based on the cuts described have been performed according to the methods described above in Section 4.3.4.4. Because the higher overall efficiency, the pileup dependence has found to be smaller than the earlier menus, the agreement between data and MC is better, and the systematics are smaller.

4.4 Photon Isolation

Isolation measures of the 'amount of stuff' in a cone of radius ΔR around an object – in the present case, a photon. The 'stuff' may be the number of tracks, the sum of the p_{T} of the tracks, or the calorimeter energy measured in any of several ways. Isolation requirements typically use ΔR cones with radius between 0.2 and 0.4, which is a much larger solid area than used by the identification described above; the isEM variables use boxes of sides at most $\Delta\eta \times \Delta\varphi = 0.175 \times 0.175$. The surface used by the isolation is thus a factor of 4 to 16 larger, and for real photons the behavior is understood to be uncorrelated from the isEM variables. This makes isolation an extremely powerful second discriminant, and also very important in extracting purities or making efficiency measurements, as described in Section 4.3.4.3 above.

When calculating the isolation, two quantities should be removed from the total: the photon itself as well as any noise. This section describes the calorimeter and track isolation methods used for the $h \to \gamma\gamma$ analysis. Both methods are illustrated in Figure 4.21.

4.4.1 Calorimeter Isolation

The calorimeter isolation is defined as the sum of all positive topological clusters within $\Delta R < 0.4$ of a photon. The $\Delta\eta \times \Delta\varphi = 5 \times 7$ rectangle around the photon is subtracted from this total, which corresponds to better than 95% of the photon energy.[3] Topological clusters are seeded by individual cells with energy greater than four standard deviations of the combined electronic and pileup noise, and include all cells that

[3]The Molière radius in the barrel of the liquid argon calorimeter is about 5 cm, corresponding to around 1.5 cells.

can be continuously connected to the seed, with energy greater than 2σ above 0 [99]. The use of topological clusters builds in some noise rejection.

Two corrections are applied to the initial quantity to account for the leakage of the photons energy outside the 5×7 rectangle, and pileup noise of the event. The leakage of photons outside of the 5×7 rectangle depends on both the amount of material in its path ($|\eta|$) and its energy. A leakage correction is therefore binned in $|\eta|$ and parameterized in p_{T}. The pileup noise is parameterized by reconstructing an ambient energy density for the event, which is taken as the median energy density of all the jets in the event [110]. Together, these corrections remove any dependence on pileup or the photon's energy.

4.4.2 Track Isolation

Track isolation is an appealing supplement to calorimeter isolation, partially because it provides confirmation from a completely separate system of the detector. In addition, the corrections for pileup and the photon energy are very straightforward: any tracks reconstructed as part of a (converted) photon are explicitly removed, and any tracks that do not originate from the primary vertex associated to the photons are dropped. All good quality within $\Delta R < 0.2$ of the photon, $p_{\mathrm{T}} > 1$ GeV, impact parameter $d_0 < 1.5$ mm with respect to the primary vertex, and $z_0 \sin\theta < 15$ mm are included. The only trick with diphoton events is that the primary vertex is not necessarily easy to define, since unconverted photons do not leave tracks that 'flag' the correct vertex. This particular challenge will be explored in the following chapter.

Chapter 5

Discovery of the Higgs Boson

Once the ATLAS detector was built and the LHC flooded it with collisions; after the data was recorded and processed and the photons were built and calibrated and selected – then discovering the Higgs boson in the diphoton decay mode was not such a challenging task. Indeed, the simplicity of the search in this channel is one of its fundamental strengths: it is a bump-hunt in the diphoton invariant mass spectrum (Figure 5.1). The backgrounds are (almost) entirely constrained by data. The high signal selection efficiency of around 40% compensates for the low branching ratio to two photons (0.00228), and gives hundreds of events to study in the full dataset. The energies of the photons are precisely measured, leading to a narrow diphoton mass resolution, leading to a clear resonance and an unambiguous discovery.

This chapter describes the $h \to \gamma\gamma$ search both in the context of the discovery, and as it serves as a baseline for further measurements. In cases where 'best practice' has evolved since the discovery the current methods are presented, in particular in the description of systematic uncertainties. Nevertheless, the initial discovery results are included alongside the results with the full dataset, for historical interest.

5.1 Event Selection

The analysis optimized to select events with diphoton mass $100 < m^{\gamma\gamma} < 160$ GeV. As shown in Figure 2.4b, the Higgs boson branching ratio to a pair of photons plummets for $m_h > 160$ GeV. Above 140 GeV, the ZZ and WW decay modes are already far more sensitive. Masses below 114 GeV were directly excluded by the LEP search. The region of interest is therefore $115 < m^{\gamma\gamma} < 150$ GeV, but both upper and lower sidebands are needed for the sideband fit.

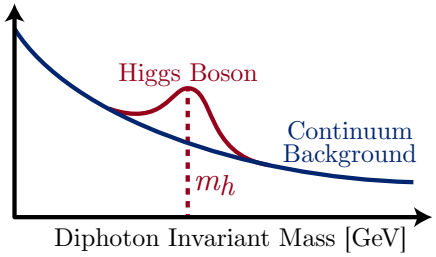


Figure 5.1: The search for the Higgs boson in the diphoton decay channel comes down to a straightforward 'bump-hunt' in the diphoton invariant mass. The backgrounds are largely constrained by data, and the signal shape is taken from simulation.

Events are recorded using a diphoton trigger with an efficiency above 99% with respect to the offline selection, requiring two energy clusters satisfying the loose photon selection criteria. In 2011, both photons were required to have $p_{\mathrm{T}} > 20$ GeV. The thresholds were raised in 2012: the leading (most energetic) cluster had to have $p_{\mathrm{T}} > 35$ GeV, while the subleading (second most energetic) was required to have $p_{\mathrm{T}} > 25$ GeV.

Offline, events are required to have passed the trigger and satisfy basic data quality requirements. The event must have at least one vertex with three or more tracks. The two highest-p_{T} loose photons are then 'preselected.' They are required to fall within the fiducial region of the detector for photons: $|\eta| < 2.37$, excluding the transition region between the barrel and endcap calorimeters, $1.37 \leq |\eta| < 1.56$ (1.52, at the time of the discovery).

At this juncture, a primary vertex (z position) must be selected for the event. The primary vertex has two important roles. (1) After 'redefining' the origin position, one can recalculate the pseudorapidity of the photons; this affects in turn the photons' p_{T}s and the combined mass. (2) Several pileup suppression techniques hinge associating tracks to vertices, and rejecting tracks or jets that do not originate from the same vertex as the hard process. However, selecting the correct vertex in diphoton events is complicated by the fact that unconverted photons obviously do not leave any track. The standard ATLAS method of choosing the primary vertex with the highest sum of p_{T}^2 of the associated tracks is therefore insufficient for these events, since the diphoton vertex may not have any hard tracks at all! A multivariate discriminant is used to select the most probable vertex, using the photon 'pointing' information (extrapolating between layers of the EM calorimeter to the beam line), along with the tracks of converted photons, the $\sum p_{\mathrm{T}}$ and $\sum p_{\mathrm{T}}^2$ of tracks associated to a vertex, and the azimuthal separation $\Delta\varphi$ between the (vector) sum of the tracks' momenta and the diphoton system. At the time of the discovery, a simple likelihood was used; a binary decision tree was developed for later analyses. Once a vertex has been selected, the photons' η and $p_{\mathrm{T}} = E/\cosh(\eta)$ are recalculated, as is their combined mass.

Kinematic cuts are applied to the updated four-vectors of the photons. At the time of the Higgs discovery and for the coupling analysis (Chapter 6), the p_{T}s of the leading and subleading photons were required to exceed 40 and 30 GeV, respectively. For the spin, differential cross section (Chapter 7), and di-Higgs (Chapter 8) analyses, a cut is placed on the ratio of the p_{T}s to the combined mass: $p_{\mathrm{T}}/m^{\gamma\gamma} > 0.35$ (0.25) for the leading (subleading) photon. These 'relative cuts' sidestep substantial deformations introduced in the $m^{\gamma\gamma}$ spectrum that arise with the 'absolute cuts' when binning the dataset in $p_{\mathrm{T}}^{\gamma\gamma}$ or $|\cos\theta^*|$.

Tight or neural network identification is imposed on the two photons, as described in the last Chapter. For 2011 and 2012 data at the initial discovery, an isolation cut of 4 GeV is placed on the topological calorimeter isolation within a cone of $\Delta R = 0.4$, of each photon candidate. For the final 2012 analyses, a cut of 6 GeV made on the topological isolation, and a cut of 2.6 GeV is placed on the sum of the p_{T}s of the tracks within a cone of $\Delta R = 0.2$ of the photon candidates.

Finally, a mass cut is applied to the selected events. For early analyses, this was $100 < m^{\gamma\gamma} < 160$ GeV, while for later analyses, the combination of the relative cuts and the trigger thresholds of 25 and 35 GeV motivated the slightly narrower window of $105 < m^{\gamma\gamma} < 160$.

The efficiency for all cuts is roughly 40% for both years, and is slightly higher for ggh and VBF than Vh and tth. In 4.8/fb at $\sqrt{s} = 7$ TeV roughly 80 events are expected from all modes (according to which analysis), while 400 events are expected from the 20.3/fb of data at $\sqrt{s} = 8$ TeV.

5.2 Simulation

The detector acceptance of and response to physics processes are simulated using Monte Carlo techniques. In this analysis, MC is used primarily to determine the shape of the signal model (Section 5.3.1), to calculate the overall rate of SM Higgs boson production, and to derive systematic uncertainties (Section 5.5). In the differential cross section analysis (Chapter 7), the signal MC is also used to derive correction factors for the unfolded cross sections. Simpler 'truth ntuples' that contain a record of generated particles but not of their interactions with the detector, are used to help validate the background model. MC may also shed light on the composition of the background processes, but this is not used in the analysis itself.

5.2.1 Signal

The Higgs boson production and decay are simulated for Higgs boson masses m_h in 5 GeV steps between 85 and 150 GeV, in order to create a parameterization of the reconstructed signal line shape as a function of m_h (Section 5.3.1). The five dominant production modes are simulated for each mass value. The normalization and factorization scales are set to the Higgs mass. The common ATLAS simulation tunes and parton distribution function (PDF) sets are used [111,112]. Parton level ggh and VBF samples are generated using POWHEG [113–115] with the CT10 PDF tune and interfaced to Pythia8 [111] to simulate the decay of the Higgs boson, showering and hadronization. For 2012, the POWHEG `HFact` parameter that controls the Sudakov form factors is tuned to match the HqT calculation that includes with finite mass effects and soft-gluon resummations up to next-to-next-to-leading logarithmic order (NNLL) [22,116–118]. Higgs bosons produced in association with a W or Z boson, or a $t\bar{t}$ pair, are generated with Pythia8 with the CTEQ6L1 PDF tune. The interaction of particles with the detector is simulated [119] using Geant4 [120]. Effects from multiple interactions in a bunch crossing (pileup) are simulated by overlaying each signal event with a variable number of simulated inelastic proton-proton collisions.

Each process is scaled to the rate predicted from the theoretical predictions presented in Table 2.1.

5.2.1.1 Corrections to the Signal Monte Carlo

Despite best efforts, simulated data may fail to reproduce characteristics of data or represent the best knowledge of a physical process. In cases where the discrepancy can be identified and understood, it can be corrected.

Fudge Factors. As described in Section 4.3.2, the photon shower development in Monte Carlo is imperfect. The distributions of the isEM shapes are seen to be shifted for simulated showers with respect to those from data. The 'fudge factor' corrections described in Section 4.3.2 represent one of the largest overall corrections to the simulation.

Energy 'Smearing.' The energy resolution of Monte Carlo broadened to match the one observed in data, by smearing the energies of electrons, to better agree with the $Z \to ee$ lineshape. The derived smearings are applied to photons.

Interference Weight. Gluon fusion Higgs boson production interferes destructively with the $gg \to \gamma\gamma$ (box) continuum background [121]. This interference depends strongly on $|\cos\theta^*|$ and peaks as it approaches 1, since the $gg \to \gamma\gamma$ amplitude has a pole there. It reduces the expected gluon fusion yield by 2.2% overall, and is applied as a reweighting of that sample.

Beam-spot position. The actual spread in z of hard process interactions follows a Gaussian of width of approximately 5 cm around the detector origin, but the value used is not consistent between data and Monte Carlo, or even, different Monte Carlo releases. This z position impacts, for instance, the precision of the photon pointing (since the segmentation of the liquid Argon calorimeter is optimized for the nominal interaction point) and the number of additional vertices in close proximity to the hard process. The z position of the generated hard processes are therefore reweighted to the actual beam parameters.

Pileup distribution. Particularly when a Monte Carlo data set is produced before recording the data, it will not reproduce the distribution of the number of interactions per bunch crossing. Even after data taking has ended and the true pileup distribution of the run is known high-pileup tails may added to the generated distribution to facilitate studies of the performance under different pileup conditions. Pileup has an important impact on tracking, calorimeter resolution, and particle identification, and it is important that the distribution be correctly modeled in MC. This distribution is therefore reweighted.

Dalitz decays. The decay of the Higgs boson to one on-shell photon and two fermions through the 'Dalitz decay' of an off-shell photon, $h \to \gamma\left(\gamma^* \to f\overline{f}\right)$, proceeds with a branching ratio roughly 10% as large as the one to two on-shell photons [122–124]. This is illustrated in Figure C.1 of the Appendix. This rate

supplements the $h \to \gamma\gamma$ fraction. In the ATLAS 2012 Monte Carlo production however, the Pythia8 `TimeShower:QEDshowerByGamma` parameter was left on, so that decays intended to be entirely on-shell in fact included the off-shell piece. These $\gamma f \overline{f}$ decays were therefore parasitic on the on-shell rate. Since only a small fraction of the these events (usually, $ee\gamma$) are reconstructed as diphoton events, this reduced the fraction of selected $h \to \gamma\gamma$ events. This led to a 5% over-estimate of the reported signal strength in $h \to \gamma\gamma$ at the time of the discovery. In more recent analyses, these simulated events are removed from the $h \to \gamma\gamma$ total. The Dalitz fraction is reweighted to the leading theoretical predictions, since some of them *are* selected as diphoton events.

Transverse momentum of gluon-fusion produced Higgs bosons. Several advanced calculations performed soft emission resummation of gluon-fusion Higgs boson production. This is important since most early versions of the analysis categorized selected events by their $p_{\mathrm{T}t}$ (p_{T} along the 'thrust' axis, Section 5.4), which is highly correlated to p_{T}, in order to improve the sensitivity. If the fraction of events in these categories is wrong, it can affect the significance and the signal strength. In 2011, the ggh p_{T} spectrum was reweighted to the HqT program; in 2012 POWHEG was tuned to match it as described above. However, that tuning was not totally successful, and future analyses will use the HRes 2.2 to calculate and reweight the p_{T} spectrum.

5.2.2 Background Simulation

The continuum background from diphoton and photon-jet events is almost entirely constrained from data, but Monte Carlo is used to evaluate the appropriateness of the analytical functions used to describe the shapes. To do this, samples are needed with negligible statistical uncertainty, as described in Section 5.3.2. Backgrounds from prompt diphoton and photon-jet processes (about 75% and 20% of the total background) are simulated using Sherpa [125], with up to three quark or gluon emissions from the matrix element using a leading order multijet merging algorithm. The jet-jet background, about 5% of the total background, is simulated with Pythia8. Hundreds of millions of events are required, so simulating the full detector acceptance and interaction is not feasible, and only the truth particle record is produced.

5.3 Modelling

The crux of the $h \to \gamma\gamma$ analysis is a signal plus background fit of the $m^{\gamma\gamma}$ shape. The analytical functions used to describe the signal and background are therefore critically important to the analysis.

5.3.1 Signal

The Higgs signal shape is modelled analytically using a Crystal Ball (CB) function plus a wide Gaussian component for tails. The CB shape is defined by

$$\begin{cases} \exp\left\{-t^2/2\right\} & t > -\alpha \\ \left(\dfrac{n}{|\alpha|}\right)^n \exp\left\{-\alpha^2/2\right\} \left(\dfrac{n}{|\alpha|} - \alpha - t\right)^{-n} & t \leq -\alpha \end{cases} \tag{5.1}$$

with $t \equiv (m_{\gamma\gamma} - \mu)/\sigma_{CB}$, and $n \equiv 10$ (in the present analysis). The position of the peak is offset from the true Higgs mass, so the mean parameter is defined as $\mu \equiv m_H - \Delta\mu$. The CB fraction of the signal model is called f_{CB}. The Gaussian's width is denoted by $\sigma_{Ga.} \equiv \kappa_{Ga.} \times \sigma_{CB}$, and its mean μ shared with the CB component.

The parameters of this fit – $\Delta\mu$, σ_{CB}, α_{CB}, $\kappa_{Ga.}$, and f_{CB} – vary smoothly with the mass of the Higgs candidate, and can be parameterized as linear functions of it. In practice, only σ_{CB} has a significant dependence on m_h. The overall normalization is a cubic function of m_h. Using these parameterizations, a single 'global signal model' can describe the shape across the entire mass space. This parameterization is necessary both (1) to test arbitrary masses before the true mass was known and (2) in evaluating the

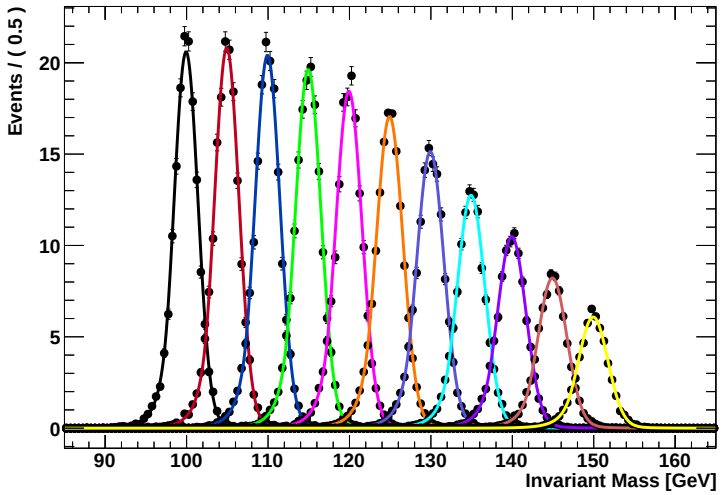


Figure 5.2: Global signal fit for the transverse momentum bin, $p_{\mathrm{T}} < 20$ GeV. A single model can be used to describe the signal shape for an arbitrary Higgs boson mass hypothesis.

background functions (which is done as a function of the mass). An example of this global fit is shown, for events with $p_{\mathrm{T}}^{\gamma\gamma} < 20$ GeV, in Figure 5.2.

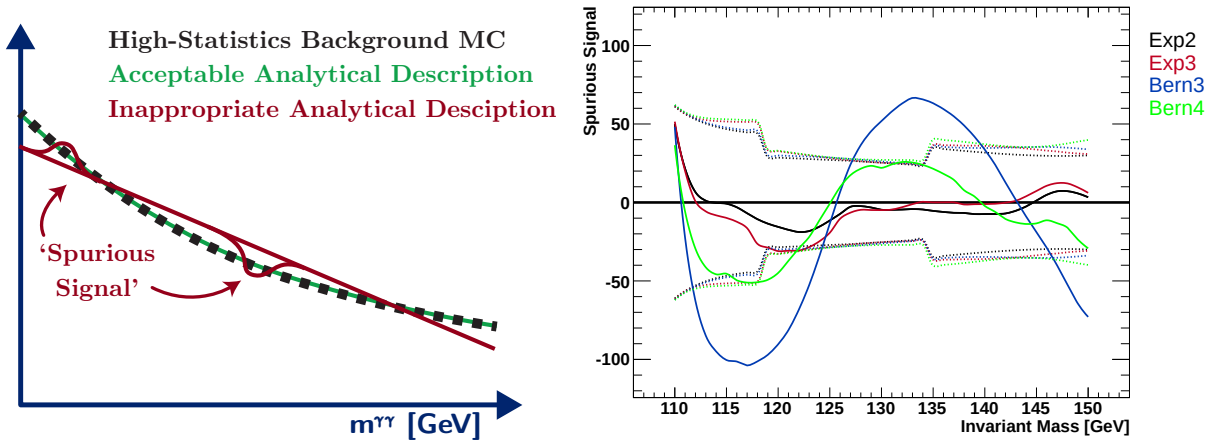
5.3.2 Background Shape and 'Spurious Signal'

The background shape in $m^{\gamma\gamma}$ is fit to the data, but it is described by an analytical function that may or may not be able to successfully reproduce the true shape. Any undershoots or overshoots due to an overly- or insufficiently-flexible background model can be fitted as a 'spurious signal' or alter the measured rate of a true signal. This is illustrated in Figure 5.3a, where fitting a straight line to a curved background leads to obvious biases. In practice, the effects are far more subtle. A procedure was defined to both (a) identify an appropriate background functions for each category and (b) quantify any residual bias.

In short, the aim is to fit a candidate background shape parameterization to the 'true shape,' or a stand-in for the same. To do this, extremely high-statistics Monte Carlo was generated, as described in Section 5.2.2. This provided a background only sample with effectively no statistical fluctuations. However, because of the size of the samples, the interactions of the particles with the detector were not simulated Instead, the photon conversion status, isolation, and identification efficiencies were all parameterized as functions of the photons' p_{T} and $|\eta|$ using fully simulated Monte Carlo. This results in 'true' background shapes against which to test models.

Simultaneous signal plus background fits are made to these shapes, for fixed signal masses from 110 to 150 GeV. Any extracted signal is 'spurious,' and the rate of spurious signal for any signal mass was required to be less than 10% of the expected SM signal yield, or 20% of the statistical uncertainty on the background.[4] Two classes of background models were tested: Bernstein polynomials of order 2 through 4, and exponentiated polynomials of the form $e^{ax+bx^2+\cdots}$. Models that failed the requirements above were rejected, while the successful model with the lowest number of degrees of freedom was selected. The largest measured bias for that model is taken as a systematic uncertainty on the yield. The measured spurious signals for a category of the Dihiggs analysis is shown as an example, in Figure 5.3b.

[4] After the Higgs boson mass was measured, these requirements were relaxed for mass hypotheses more than four standard deviations from m_h, to 20% of the signal yield and 30% of the background uncertainty.



(a) Schematic of the Derivation of Spurious Signal

(b) Example: Dihiggs Search

Figure 5.3: Spurious signal

5.4 Categorization

By dividing the data set selected in Section 5.1 into exclusive *categories*, one can isolate regions of phase space with higher $S/\sqrt{B}$ or lower uncertainties, or study properties of the production. In the time leading up to the discovery, the categorization was optimized for overall sensitivity. Later on, production mechanisms were isolated (Chapter 6) and events were binned according to kinematic variables (Chapter 7).

The energy resolution of a photon depends on whether or not it converts into a di-electron pair and where in the detector it goes – in short, what part of the detector it reaches and how it interacts with it. This directly impacts the diphoton mass resolution. The effective $S/\sqrt{B}$ depends on the number of events beneath the signal; if the signal width (resolution) is smaller, there are fewer events underneath it. The initial categorization therefore divided the dataset in exclusive regions of conversion status and $|\eta|$, as illustrated in Figure 5.4a. As shown, the core of the barrel has the best resolution, while the transition region between the barrel and endcap calorimeters has the worst resolution. Five categories were defined:

(1) Neither photon converts, and both are contained in the 'good' region of the calorimeter with $|\eta| < 0.75$.

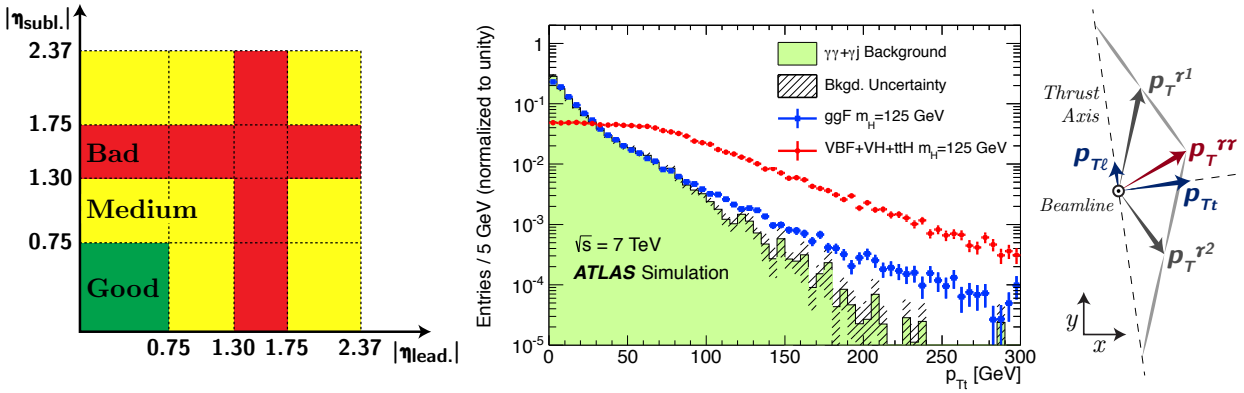
(2) Neither photon converts, but at least one of them has $|\eta| > 0.75$.

(3) At least one of the photons converts, but both are in the 'good' region of the calorimeter ($|\eta| < 0.75$).

(4) At least one of the photons converts, but neither is contained in the 'bad' region of the calorimeter ($1.3 < |\eta| < 1.75$).

(5) At least one of the photons converts, and at least one of them is contained in the 'bad' region of the calorimeter ($1.3 < |\eta| < 1.75$).

Additional discrimination between production modes and background – specifically for VBF and VH processes – is attained using the $p_{\mathrm{T}t}$ variable, as shown in Figure 5.4b. The $p_{\mathrm{T}t}$ is defined in the plane transverse to the beam, as the projection of the diphoton transverse momentum perpendicular to the 'thrust axis,' which itself is the difference of the leading and subleading photons' transverse momenta (and not actually the thrust axis). This is illustrated in Figure 5.4c. Theoretically, the $p_{\mathrm{T}t}$ has better resolution than the p_{T} [126]. Using this variable, four additional categories were defined by 'splitting' each of the first four categories above, among those events with $p_{\mathrm{T}t} \leq 60$ GeV or $p_{\mathrm{T}t} > 60$ GeV.



(a) Regions in $|\eta|$ (b) $p_{\mathrm{T}t}$ for Background and Signal Processes (c) $p_{\mathrm{T}t}$ Definition

Figure 5.4: For early results, exclusive categories were defined, using three variables: the conversion status of the photons, as well as the pseudorapidity of the two photons and their $p_{\mathrm{T}t}$. The amount of material in the detector depends on $|\eta|$ and affects how the photon energy resolution. The $p_{\mathrm{T}t}$ changes the $S/\sqrt{B}$ by enhancing, in particular, the VBF and VH production modes.

An additional category was defined to select the VBF-type topology, for the discovery. Events with at least two jets were selected (see Chapter 6) with a pseudorapidity gap between them of $\Delta\eta\,(j_1, j_2) > 2.8$, a combined mass $m_jj > 400$ GeV, and an azimuthal separation from the diphoton system $\Delta\varphi\,(\gamma\gamma, jj) > 2.6$. The VBF category was revisited, and additional 'production-mode' categories were added in the 'couplings analysis.'

5.5 Overview of Uncertainties

5.5.1 Uncertainties on the Yield

Yield uncertainties affect the total number of expected signal events. For instance, theoretical uncertainties alter the number of events produced, and uncertainties on the identification efficiency impact the fraction of events that are selected.

Theoretical Uncertainties. The theoretical uncertainties on prediction cross sections from parton distribution functions, α_s, and scale variations are included in Table 2.1 for $\sqrt{s} = 8$ TeV. The uncertainties for $\sqrt{s} = 7$ TeV are extremely similar.

Luminosity. The luminosity is evaluated using beam separation (van der Meer) scans [127] following the techniques described in Ref. [128]. The uncertainty for 2011 was 1.8%; the initial uncertainty on the first 5.9/fb of 2012 data was 3.6%, while the final uncertainty on the full dataset was 2.8%.

Trigger Efficiency. The trigger efficiency is measured to be better than 99% for both years, using 'bootstrap methods,' comparisons to Monte Carlo, and studies of tagged photons from $Z \to \ell\ell\gamma$ [129]. The initial uncertainty of 1% was reduced to 0.5%.

Photon Identification. For 2011, the neural network photon identification efficiency was based on the maximum observed discrepancy between Monte Carlo and data-driven measurements from (a) $Z \to \ell\ell\gamma$ tag and probe, (b) extrapolation from electrons, and (c) a system of equations ('matrix method') using track isolation. As described in Section 4.3.4.4, the differences were typically less than 4%, but as high as 7% for unconverted photons in the endcap. For 2012 these same methods were used to determine the efficiency; for early analyses, the difference between data and MC was again taken as the systematic: 5% for photons in

the barrel and 7% for photons in the endcaps. For later data, the data-driven measurements were used as the uncertainty, and a model of the correlations on that measurement was used to map the overall uncertainty for events with two photons; the resultant uncertainty dropped to 2.4%. This treatment was inconsistent and incomplete, however, because scale factors were not applied to the identification efficiency in MC.

Isolation. Differences in the isolation efficiency in data and MC are evaluated using 'probe' electrons from $Z \to ee$ decays. Initially, a 'shift' was determined for the difference in the peaks of the distributions observed in data and Monte Carlo, and this was used to recalculate an 'effective' cut. Ultimately, with the introduction of track isolation, simply the difference in efficiencies was used. That uncertainty depends on the number of jets in the event: 1% for events with with 0 or 1 jets, 2% for events with 2 jets, and 4% for events with 3 or more jets.

Photon Energy Scale. The photon energy scale impacts both the fitted peak position and also the number of events that satisfy the kinematic cuts. The resultant uncertainty on the yield is 0.3%.

Pileup (Early Analyses). For early analyses only, a dedicated pileup uncertainty of 4% was evaluated from the difference in the overall selection efficiency from low and high-pileup events in MC.

5.5.2 Uncertainties on the Mass Shape

As described in Section 4.2, the photon energy scale and resolution is derived from comparing of $Z \to ee$ lineshape in data and MC. The Monte Carlo is 'smeared' to match the observed resolution. Uncertainties on that smearing, and on the extrapolation from electrons to photons are propagated to the signal shape. For early analyses, the differences in the lineshape were not taken into account, and an uncertainty of around 14% was assessed. For later analyses, the uncertainties in the couplings analysis ranged as high as 22%.

5.5.3 Migration Uncertainties

'Migration' uncertainties describe the movement of event populations between categories. They are thus defined so as to preserve the total number of selected events. For the simpler categorization of the earlier analyses, all that mattered was the fraction of conversions and high-$p_{\mathrm{T}t}$ diphoton pairs. (The η resolution is quite good.) This list of uncertainties will expand dramatically for the couplings analysis in Chapter 6 where a number of additional objects are required and categories are defined.

Material Modelling: Conversions. Uncertainty over material modelling alters the fraction of photons that convert. This results in an -4% uncertainty on unconverted photon categories or $+3\%$ for categories with converted photons.

Higgs p_T. The theoretical uncertainty on the Higgs boson p_{T} is evaluated by altering the scales and PDFs used by the HqT program; this has a small impact on the low-p_{T} categories (1.2%) but larger impacts on the high-p_{T} and VBF categories ($\sim$10%).

5.6 Statistical Model and Mechanics

A statistical model is defined that analytically describes all salient observables or uncertainties on a data set in a single likelihood expression. Of the observables, one or several 'parameters of interest' are defined – most commonly, the 'signal strength' with respect to the expectation from the SM, $\mu \equiv \sigma_{\mathrm{obs.}}/\sigma_{\mathrm{exp.}}$. (The signal strength can of course be further divided: μ_{VBF}, μ_{Vh} and so forth.) The remaining observables are called 'nuisance parameters,' and denoted $\boldsymbol{\theta}$. They include both uncertainties (e.g., on the theoretical predictions, or efficiencies), and the parameters of the overall fit (the Higgs boson mass, the slope of the background shapes, etc). In a profile likelihood fit, all parameters are allowed to vary and are fit simultaneously.

The likelihood ratio is used [130],

$$\lambda(\mu) = \mathcal{L}\left(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu)\right) \Big/ \mathcal{L}\left(\hat{\mu}, \hat{\boldsymbol{\theta}}\right) \tag{5.2}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ represent the overall best fit, while $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ represents the best of the nuisance parameters for a given μ. This normalizes the overall likelihood of a given signal strength to the best possible fit for *any* strength. From the likelihood, a 'test statistic,' the log-likelihood ratio is defined:

$$t_\mu = -2\log\lambda(\mu) \ . \tag{5.3}$$

Equation 5.2 ranges between 0 for bad agreement to a hypothesis, to 1 for the most probable hypothesis. The test statistic therefore ranges between 0 (good agreement) and infinity (bad agreement).

Two important special cases are defined: the test statistics for discovery or for setting upper limits [131]. In evaluating the incompatibility to the background-only hypothesis negative fluctuations are disallowed, and the 'physical' best fit value is interpreted simple as $\mu = 0$; in this case, the test statistic is modified as

$$\tilde{q}_\mu = \begin{cases} -2\log\left(\mathcal{L}(\mu)/\mathcal{L}(0)\right) & \hat{\mu} < 0 \\ -2\log\left(\mathcal{L}(\mu)/\mathcal{L}(\hat{\mu})\right) & \hat{\mu} \geq 0 \end{cases} \tag{5.4}$$

For upper limits on the signal strength, a second modification is made: best fit values $\hat{\mu}$ *above* the tested value μ are not considered incompatible with the test value. The test statistic is then:

$$\tilde{t}_\mu = \begin{cases} -2\log\left(\mathcal{L}(\mu)/\mathcal{L}(0)\right) & \hat{\mu} < 0 \\ -2\log\left(\mathcal{L}(\mu)/\mathcal{L}(\hat{\mu})\right) & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{5.5}$$

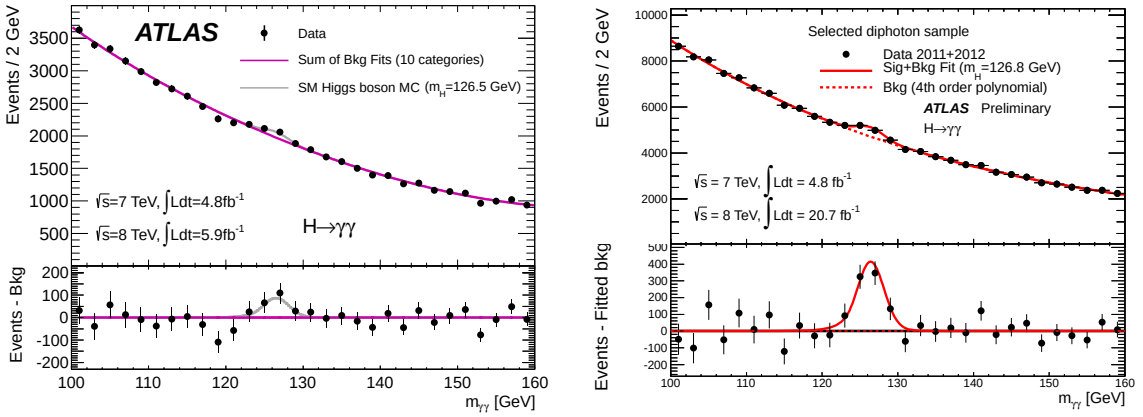
In each case the probability (p-value) is defined as the integral of the test statistic distribution, from the observed value to infinity: it is the fraction of experiments that result in a lower compatibility to the tested hypothesis. The p-value is typically transformed into a number of standard deviations via $Z \equiv \Phi^{-1}(1 - p_\mu)$. The threshold to exclude the background-only hypothesis is set at $Z > 5$ ($p_0 < 2.87 \times 10^{-7}$). The problem is: where to get the distribution of the ensemble of experiments?

Traditionally, this was done using pseudo-experiments (toys) in which the nuisance parameters/auxiliary measurements were randomized within their uncertainties, and events were generated from the template of the full statistical model. However, this procedure is phenomenally expensive computationally. For large-N datasets, 'asymptotic methods' are used to calculate the number of standard deviations without generating millions of toy datasets. The trick is to define an *Asimov dataset* that, if used as the estimator for any parameter, yields the true parameter. This is effectively a 'perfect' dataset, without any statistical fluctuations. It is then possible to derive the variance on the parameter of interest either by solving for the covariance matrix, or solving directly for σ, using the properties of the Asimov dataset and the known properties of the test statistic [131].

In most cases of the $h \to \gamma\gamma$ analysis, statistics are sufficiently large for the use of the asymptotic formulae. In particular, the approximations are valid for all categories of the couplings measurements. However, in the differential cross section analysis no such formulae are required, and the very-low statistics of the Higgs boson pair production search preclude their use for that analysis.

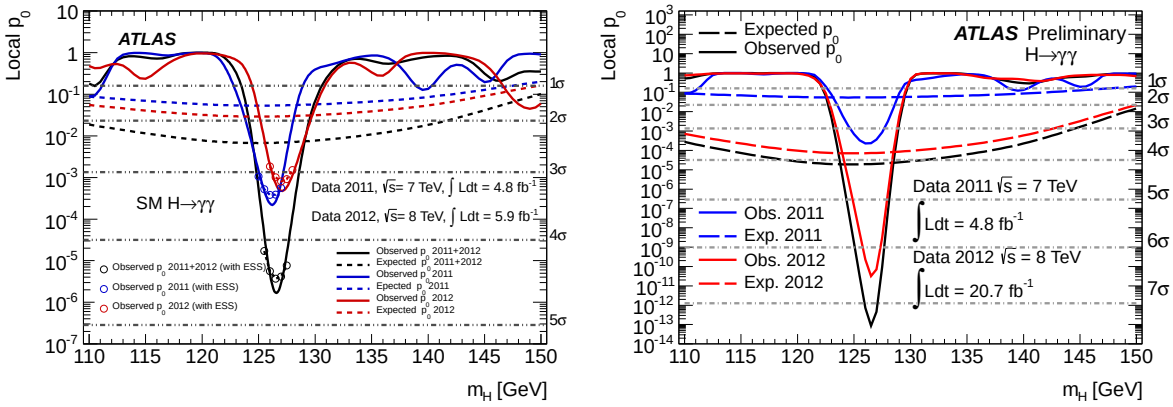
5.7 Discovery of the Higgs Boson

Nearly fifty years after the Higgs mechanism was first hypothesized, the discovery of a new boson was officially announced on July 4, 2012. With 4.8/fb of data at $\sqrt{s} = 7$ TeV and 5.9/fb of data at $\sqrt{s} = 8$ TeV. The inclusive diphoton invariant mass spectra are presented in Figure 5.5, both as used for the discovery and as observed after all 2012 data was collected. The resonance is clearly visible, even before subtracting



(a) Discovery: July 2012 Dataset (b) Full Dataset

Figure 5.5: The diphoton invariant mass spectrum is presented with 4.8/fb of data at $\sqrt{s} = 7$ TeV and 5.9/fb at $\sqrt{s} = 8$ TeV. The continuum fit is the sum of the background models from each individual category.



(a) Discovery: Observed p_0 from $h \to \gamma\gamma$ (b) Discovery in $h \to \gamma\gamma$ Alone

Figure 5.6: CL_S exclusion and the local probability of the background-only hypothesis (p_0) from the $h \to \gamma\gamma$ decay channel alone, at the time of the discovery.

the background. The compatibility to the background-only hypothesis reaches a minimum at a mass of $m_h \sim 126$ GeV. Figure 5.5 shows the p_0 in the diphoton channel. At the time of the discovery, the significance of the excess was already more than 4σ. For the full year, the significance in the $h \to \gamma\gamma$ channel alone reached 7.4σ, enough to claim discovery with just that channel.

Finally, the 'discovery plot' is presented in Figure 5.7. Combining the diphoton channel with $h \to ZZ$ and $h \to WW$, leads to an incompatibility to the background-only hypothesis at the level of 6σ. A new boson was discovered, with mass and nominal couplings consistent with the expectations of the SM Higgs boson. That boson decayed to photons, precluding a spin-1 particle. The WW analysis used the spin-0 nature of the Higgs boson to extract it from the SM WW background, and measured a rate consistent with the SM expectation. If hints of the Higgs boson in $W\,(h \to bb)$ from the Tevatron were to be taken seriously,

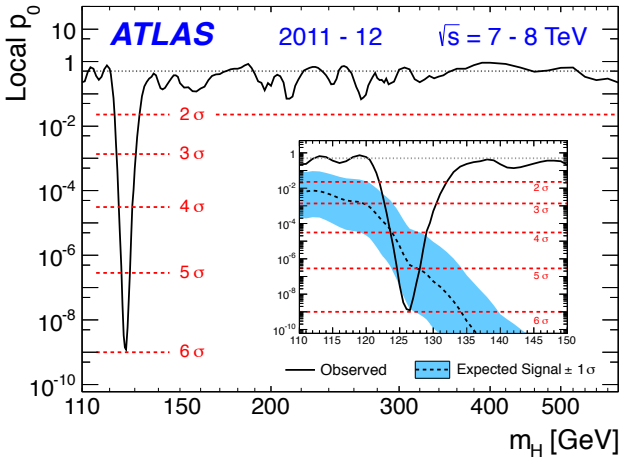


Figure 5.7: Local probability of the no-Higgs boson hypothesis, for the ATLAS discovery.

a spin-2 particle was already highly-disfavored [132, 133]. It looked like a scalar – but was it? Did it indeed interact according to the predictions of the SM Higgs boson? Were there any hints of physics beyond the Standard Model? The balance of this thesis proposes tests to these questions.

Chapter 6

Coupling Measurements

The interactions of the Higgs boson to different particles is a central part of what makes it 'the Higgs:' do the couplings run proportional to mass? To isolate and measure this, one can either (a) measure many decay channels or (b) sort events according to production mode. Of course, the two strategies complement each other and ATLAS has pursued both. Sorting events by production mode comes down to selecting additional objects in events with Higgs bosons: two forward jets with a wide rapidity gap and large combined mass for VBF; leptons, missing energy, or two central jets with $m_{jj} \sim m_Z \sim m_W$, for Vh. This chapter describes the definition of these categories in the context of the ATLAS $h \to \gamma\gamma$ analysis.

Section 6.3 briefly extends these results through the combination of measurements between decay channels. The diphoton channel is a crucial part of those combinations. Not only does it have large statistical power in all production modes, but the decay loop to diphoton pairs is sensitive to the presence of additional charged particles, and interference in $h \to \gamma\gamma$ decay loop provides the only input on the relative *sign* of the fermionic and bosonic couplings.

6.1 Object Definitions

Identifying 'additional features' of the dataset demands the introduction of new object definitions. The ATLAS 'legacy' papers that present the 'final word' on Run I data will use both 2011 and 2012 data for all categories, but in what follows the electron, muon, and missing energy definitions apply only to 2012 data.

Electrons. Electrons are reconstructed using the same fundamental methods as photons, but they form the complement to the photons: a track must be associated to the cluster, and candidates associated to conversion vertices can be vetoed. Electrons are required to fall within $|\eta| < 2.47$ and have $p_{\mathrm{T}} > 15$ GeV. They must be isolated, with the topological isolation energy (defined as for photons) less than 20% of the electron's p_{T}, and the sum of the p_{T} of the tracks within a $\Delta R = 0.2$ cone of the electron less than 15% of its p_{T}.

Muons. Muons are reconstructed for $|\eta| < 2.7$. Their p_{T} must exceed 10 GeV. The track quality must be good, and the track must be consistent with originating from the primary vertex selected for the event. The track isolation of the muon must be less than 15% of its p_{T} and its calorimeter isolation must be less than 20% of its p_{T}.

Jets. Jets are reconstructed from three-dimensional clusters in the electromagnetic and hadronic calorimeters, using the anti-k_t algorithm [134]. They are calibrated initially based on the simulated response. Corrections are applied for pileup using the jet-area method [135], and the final calibration is refined using data.

To suppress contributions of jets from pileup interactions, tracks within jets are associated to primary vertices. The Jet Vertex Fraction (JVF) is defined as the ratio of the sum of the p_{T}s of tracks from the primary vertex of interest, to the sum of the p_{T}s of *all* tracks in the jet. In 2011 data that ratio is required to exceed 0.5 while in 2012 data it is required to exceed 0.25. This cut is only applicable within the acceptance of the tracker, $|\eta| < 2.4$. For 2012, it is only used for jets with $p_{\mathrm{T}} < 50$ GeV. Because the pileup suppression 'ends' at $|\eta| = 2.4$, the minimum jet p_{T} increases from 25 GeV to 30 GeV for $|\eta| > 2.4$. All jets must satisfy $|\eta| < 4.5$.

Overlap Removal. Double-counting of objects is avoided through overlap removal. In this analysis, photons are given priority. Jets, electrons, and muons within $\Delta R < 0.4$ of a photon are vetoed. If a muon lies within $\Delta R < 0.4$ of a jet it too is vetoed, as are electrons within $\Delta R\,(e, \mathrm{jet}) < 0.2$. Analysis-specific ambiguity resolution will be discussed in Sections 6.2.1 and 6.2.2.

Missing Transverse Energy. Missing energy is reconstructed as the negative of every other object in an event, making it in a sense the most-complex object in the event The reconstruction used for the $V\,(h \to \gamma\gamma)$ search will be discussed in Section 6.2.2.

6.2 Categories

The discovery paper categorized events based on their expected $m^{\gamma\gamma}$ mass resolution, to maximize overall sensitivity. For the categories described in this Section, the signal is not 'Higgs bosons' but rather the rarer production modes: VBF and Vh.[5] In fact, contamination from ggh is particularly detrimental since the theoretical uncertainties on ggh – especially with additional radiation – are large.

6.2.1 Lepton Category

Selecting a lepton almost 'effortlessly' yields a category with very high signal purity in Vh production, by isolating leptonic W and Z decays. Yet one can do better than simply selecting leptons.

Although a higher p_{T} threshold is required of electrons than muons, far more events are selected with electrons! Computing the invariant mass of the electron with each photon and choosing the pair with smaller $|m_{e\gamma} - m_Z|$, it becomes evident that the additional contribution in the electron channel comes from $Z \to ee$ plus an additional photon, where one of the two electrons is misidentified as the second photon. This impacts only events with electrons, because while electrons fake photons quite often (around 5% of the time, for the energies and selection criteria used here), muons do not. Figure 6.1 shows that simply cutting out those events with an electron-photon pair with mass satisfying $84 < m_{e\gamma} - m_Z < 94$ GeV dramatically reduces the background. Vetoing photons that pass the electron selection requirements also helps. The impact of these requirements is presented in Table 6.1.

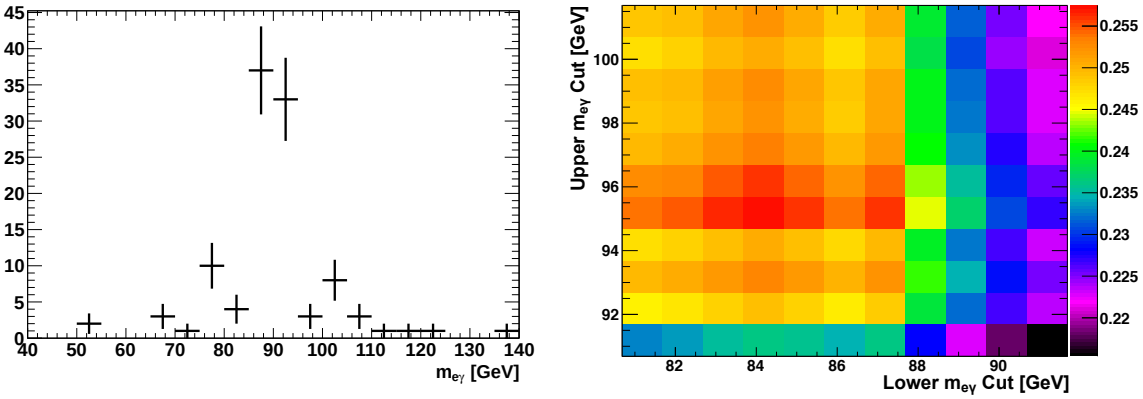
6.2.2 Missing Energy Category

Higgs bosons produced in association with a vector boson or a $t\bar{t}$ pair may also be flagged using a cut on the missing transverse energy in the event. However, the cut is not so straightforward as for leptons, because the resolution on missing energy is so broad.

6.2.2.1 Reconstruction and Alterations to the Default Definition

Missing transverse energy is calculated as the complement to the visible transverse energy in an event, assuming conservation of momentum. Two methods have been implemented to do this: a sum over locally-calibrated calorimeter clusters known as LocHadTopo; or a sum over calibrated objects, called RefFinal. This

[5] The dedicated $t\bar{t}\,(h \to \gamma\gamma)$ analysis came later [136], and will be integrated into the combined $h \to \gamma\gamma$ fit for the legacy papers. The $b\bar{b}h$ mode is scarcely separable from ggh – though it is certainly worth trying, since alterations in the bbh vertex could be more dramatic in production, than in the decay.



(a) Before m_Z Veto

(b) Rejecting Fakes

Figure 6.1: Before explicitly vetoing events with $m_{e\gamma}$ near m_Z, the events with electrons in the 'lepton' category are dominated by $Z \to ee$ events where one of the electrons fakes a photon. A simple cut around the Z mass goes far in rejecting these. At left, the category before any 'electroweak veto.' At right, the cuts on $m_{e\gamma}$ are optimized using $S/\sqrt{B}$ metric, with the expected Vh yield as signal and the data sidebands as background. The optimization is performed after rejecting photons that also pass electron quality cuts, which explains the somewhat smaller (15%) apparent gains.

definition classifies all objects in the event and applies the calibrations appropriate to each class: photons are calibrated as photons, electrons as electrons, jets as jets, etc. Each Cartesian component is thus defined:

$$E^{\text{miss}}_{x(y)} \equiv -\left(\sum E^{\gamma}_{x(y)} + \sum E^{e}_{x(y)} + \sum E^{\tau}_{x(y)} + \sum E^{\mu}_{x(y)} + \sum E^{\text{jets}}_{x(y)} + \sum E^{\text{soft jets}}_{x(y)} + \sum E^{\text{cell out}}_{x(y)}\right) \quad (6.1)$$

The Cartesian components are then combined in quadrature. This analysis uses RefFinal, for reasons elaborated below. Detailed studies of missing energy performance in 2011 and 2012 may be found in Refs [137, 138].

Three changes were made to the 'default' missing energy definition, for this analysis:

(1) *Selection Priority to Photons.* Each object enters the total sum just once. For example, once an object has been selected as an electron, it will not be double-counted as a photon – even if it is in fact a conversion track. The default method for the missing energy calculation selected electrons before photons. Because the background to be rejected was *ggh* and VBF production with *real* photons and not leptons, the first change was to choose photons first.

(2) *Object ID.* The default photon identification used the 2011 tight cuts menu, which was not used in any of the analyzed data. The $E_{\text{T}}^{\text{miss}}$ was redefined, using the current photon identification.

(3) *Calibration.* Originally, it was considered that there were so many fake photons that for generic data, a generic electromagnetic scale calibration would be more appropriate than a photon-specific tune. The most-accurate available calibration was instead used.

The impact of these alterations is discussed in Appendix D.

6.2.2.2 Confronting Pileup

One of the principal challenges in measuring and using missing energy is its broad resolution: since it is the 'opposite' of everything else in the event, its error is the sum in quadrature of every other error in

		Error on $\mu = 1.7$		VH Statistics, MET Category	
		$\delta\mu_{VH}$	$\delta\mu_{V}$		
Initial Lepton Selection		2.67	1.37		
+ Veto m_Z, BL Hits and `medium++`		2.12	1.28	Purity	Events
MET Category	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV	2.02	1.26	51%	0.66
	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV, $p_T > 70$ GeV	1.89	1.23	69%	0.55
	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV, $p_{Tt} > 70$ GeV	1.91	1.23	70%	0.39
	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV, $\not{E}_T^{\mathrm{Track}} > 70$ GeV	1.85	1.22	72%	0.55
	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV, $\Delta\varphi\left(E_{\mathrm{T}}^{\mathrm{miss}}, \gamma\gamma\right) > 3\pi/4$	1.90	1.23	76%	0.58
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Significance > 4.0	1.84	1.21	61%	1.28
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Significance > 4.5	1.75	1.19	73%	1.03
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Significance > 5.0	1.66	1.16	80%	0.86

Table 6.1: The expected error on the signal strength on the (VBF + Vh) and Vh production modes is evaluated using asymptotic formulae, and compared for various definitions of a missing energy category. The use of missing energy significance dramatically improves the measurement. The cut of $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 5 yields a category with very high purity (less than 5% ggh) and 0.9 expected events. The expected error on the $\mu_{\mathrm{VBF}+Vh}$ and μ_{Vh} signal strengths ($\delta\mu_{\mathrm{VBF}+Vh}$ and $\delta\mu_{Vh}$) is presented for each proposed $E_{\mathrm{T}}^{\mathrm{miss}}$ category, using Asimov data with $_{Vh} = 1.7$ (the then-measured value) and 13/fb of integrated luminosity.

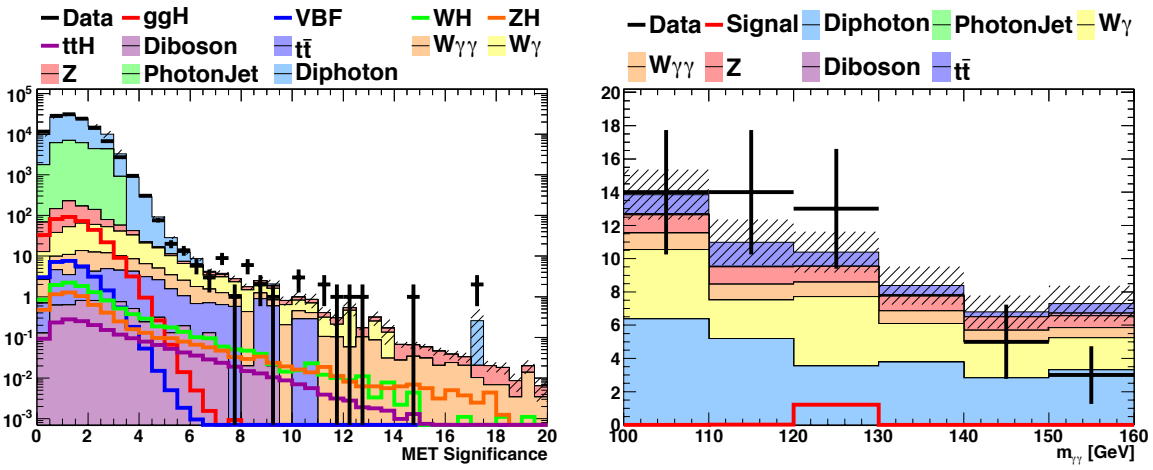
the event. One of the largest – and least well measured – contributions to the total energy in the event is the 'cell out' term, comprising all energy not associated to physics objects (photons, electrons, muons, jets, etc.). The magnitude of the cell out energy scales with pileup. This results in a 'random walk' whose 'length' is proportional to the total of the energy in the event, $\sum E_{\mathrm{T}}$. Since the average final distance covered by a random walk scales with the square root of the number of steps, the resolution on the missing energy degrades with $\sqrt{\sum E_{\mathrm{T}}}$.

One solution widely proffered by ATLAS to the degradation in $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution (referred to as the STVF algorithm) was to scale the cell out and soft jets in each event by an estimate of the fraction of the total energy in the event that came from the hard process. While the idea is appealing at first blush it does not hold up to scrutiny; it fails in particular in the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of ggh events leading to a larger contamination of the $E_{\mathrm{T}}^{\mathrm{miss}}$ category, intended for Vh and $t\bar{t}h$. The performance of this algorithm is evaluated briefly in Appendix D.2.

A simpler, time-tested method for mitigating the impact of pileup is to define missing energy 'significance,' by a cut on the missing energy over its resolution. The resolution itself is estimated by $k\sqrt{\sum E_{\mathrm{T}}}$ so that the 'significance' is $E_{\mathrm{T}}^{\mathrm{miss}}/k\sqrt{\sum E_{\mathrm{T}}}$. Based on studies of the resolution in 2011 [137] the proportionality factor k is set to 0.67, but this is irrelevant in practice since the cut on this value is also tuned, and the two values can simply be rescaled.

The impact of introducing this variable is shown in Table 6.1, where the expected error on the signal strength in the VBF and Vh production modes is compared for different definitions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ category. Considerable effort had been expended in identifying variables that could 'confirm' real missing energy: the recoil of the diphoton system $p_{\mathrm{T}}^{\gamma\gamma}$, missing momentum measured with tracks $\not{E}_T^{\mathrm{Track}}$, or the angular separation between the $\gamma\gamma$ system and the $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 6.1 indeed shows some improvement in sensitivity through these variables, but they were dwarfed by the simply replacing the cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ by the cut on its significance. Due to concerns about fits with low-statistics at the time, the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance was not combined with these variables and the single cut $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 5 was adopted for the category definition.

The distribution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance in data and Monte Carlo is shown in Figure 6.2a (though Monte Carlo is not used at all in the analysis). The invariant mass distribution of this category is shown in



(a) $E_{\mathrm{T}}^{\mathrm{miss}}$ Significance (b) Invariant Mass in the $E_{\mathrm{T}}^{\mathrm{miss}}$ Category

Figure 6.2: The missing energy 'significance' efficiently rejects non-resonant backgrounds, and the ggh and VBF production modes that have no real $E_{\mathrm{T}}^{\mathrm{miss}}$, in favor of Vh and $t\bar{t}h$. After all cuts, two types of backgrounds are roughly split: diphoton events with mismeasured $E_{\mathrm{T}}^{\mathrm{miss}}$, and $W \to e\nu\gamma$ events with a misidentified electron.

Veto	Data Eff.	Wh/Zh Eff.	$\varepsilon_S/\sqrt{\varepsilon_B}$
None	100.0%	100.0%	1.00
`medium++`	64.1%	96.0%	1.20
Pixel Hits	53.8%	81.0%	1.10
B-Layer Hits	84.6%	98.0%	1.07
B-Layer Hits or `medium++`	64.1%	95.9%	1.20
Pixel Hits or `medium++`	46.2%	80.1%	1.18

Table 6.2: Various 'electron vetoes' are proposed on top of a $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 5 cut. For instance, events could be rejected from the $E_{\mathrm{T}}^{\mathrm{miss}}$ category if either photon candidate passed `medium++` electron identification, or if it had any hits in the b-layer of the pixel detector, or any pixel hits *whatsoever*. The improvement in significance is estimated as $\varepsilon_S/\sqrt{\varepsilon_B}$, using Wh and Zh signal and backgrounds from data sidebands. The most performant configurations are `medium++` with or without b-layer hits.

Figure 6.2b. The category is populated roughly evenly between diphoton events with mismeasured $E_{\mathrm{T}}^{\mathrm{miss}}$, and $W \to e\nu\gamma$ events with a misidentified electron. The irreducible contribution from $(Z \to \nu\nu)\,\gamma\gamma$ is smaller.

The $W \to e\nu\gamma$ piece is exactly analogous to the $Z \to ee\gamma$ events seen in the lepton category, and the same need for removing electrons faking photons applies here. Table 6.2 shows several electron identification requirements, that may be inverted and applied to the photons in the $E_{\mathrm{T}}^{\mathrm{miss}}$ category. The method of reverting the standard electron identification (`medium++`) yields the best expected significance, and is employed in the analysis

		JES: Close-By	JES: Model 1	JES: η Model	JER	Soft Scale	Soft Res.
ggh	Low p_{Tt}	-0.00%	-0.00%	-0.00%	-0.00%	-0.01%	-0.00%
	High p_{Tt}	+0.00%	+0.00%	+0.00%	-0.00%	-0.00%	-0.00%
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Cat.	+7.17%	+4.54%	+5.79%	+16.34%	+59.97%	+21.10%
VBF	Low p_{Tt}	-0.00%	-0.00%	-0.00%	-0.00%	-0.01%	-0.00%
	High p_{Tt}	-0.00%	-0.00%	-0.00%	-0.00%	-0.00%	+0.00%
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Cat.	+10.45%	+6.23%	+8.59%	+10.95%	+22.40%	+9.95%
$Vh/t\bar{t}h$	Low p_{Tt}	-0.09%	-0.04%	-0.02%	-0.00%	-0.12%	+0.04%
	High p_{Tt}	+0.07%	+0.04%	-0.01%	+0.02%	-0.05%	-0.09%
	$E_{\mathrm{T}}^{\mathrm{miss}}$ Cat.	+0.51%	-0.18%	+0.30%	-0.06%	+1.01%	+0.20%

Table 6.3: Categorization migration uncertainties for the signal. The largest uncertainties on the missing energy in the couplings analysis derive from the soft terms (cells and soft jets) and jet energy scales and resolutions. Each column shows one of the uncertainties applied, and represents the fractional change in the population of three super-categories: low-p_{Tt}, high-p_{Tt}, and $E_{\mathrm{T}}^{\mathrm{miss}}$. 'Up' and 'Down' variations are symmetrized, as the average of their absolute values; the sign is copied from the 'up' variation, to preserve the direction of correlations. N.B. that no 'down' variation is considered for the jet resolution.

6.2.2.3 Systematic Uncertainties

The yield in the $E_{\mathrm{T}}^{\mathrm{miss}}$ category is extracted using a simultaneous signal plus background fit. That yield is divided in fractions of ggh, VBF, Wh, Zh, and $t\bar{t}h$. Uncertainties on the missing energy affect the population of events in each category and degrade the precision with which the measured rate can be associated to a particular production mode. These migration uncertainties preserve the total number of expected events, since they only alter the categorization.

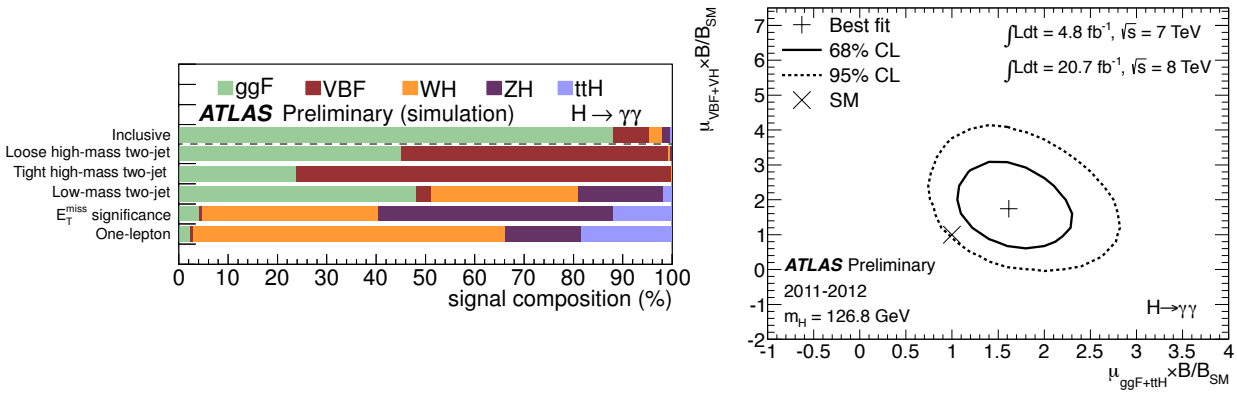
Systematic uncertainties on missing energy are evaluated by varying the resolution and scale of each component that *constitutes* the $E_{\mathrm{T}}^{\mathrm{miss}}$ within its own systematic uncertainties. The full table of all uncertainties on all components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is included in Table F.1 of Appendix F. The six uncertainties considered non-negligible are presented in Table 6.3; these are the uncertainties on the soft terms (cell out and soft jets), the jet energy resolution, and three of the jet energy scale uncertainties. The uncertainties affect primarily the ggh and VBF populations in the $E_{\mathrm{T}}^{\mathrm{miss}}$ category, since these are the tails of the distributions, and a small absolute change leads to a large fractional difference. Because the ggh and VBF fractions are so small however, these uncertainties do not pose any problem. The uncertainties are grouped into 'super-categories' which contain (a) all low-$p_{\mathrm{T}t}$ categories, (b) all high-$p_{\mathrm{T}t}$ categories, and (c) the missing energy category itself.

6.2.3 Other Categories

Two other types of categories were also defined for the couplings analyses and are described for completeness, though the author did not participate in their construction. The uncertainties on these additional categories are presented in Table G.2 of the Appendix.

6.2.3.1 Hadronic Vh.

To tag Vh production with hadronically decaying Ws and Zs, a low-mass dijet category was defined. Four requirements were imposed: (1) combined mass of the two jets must satisfy $60 < m_{jj} < 110$ GeV, (2) the pseudorapidity gap between the two jets must be $\Delta\eta\,(j,j) < 3.5$, (3) the diphoton and dijet systems



(a) Purities of Production-Mode Categories (b) Vector-Type v. Top-Type Production Modes

Figure 6.3: The composition by production mode of the five 'couplings categories' are compared to the nominal production fraction (inclusive). The lepton and missing energy categories are overwhelmingly populated by Vh, while the low-mass 2-jet has a larger contamination. The two VBF categories based on a BDT also yield respectable purities of VBF-produced events.

must have $\Delta\eta\left(\gamma\gamma, jj\right) < 1$, and (4) the $p_{\mathrm{T}t}$ of the diphoton system must exceed 70 GeV. This category is roughly evenly split between Vh and ggh. The theoretical uncertainties on $ggh + 2$jet production are large.

6.2.3.2 VBF Binary Decision Tree.

To select events consistent with Vector Boson Fusion, eight variables were combined in a binary decision tree. These variables were: m_{jj}, $\Delta\eta\left(j, j\right)$, η^{j1}, η^{j2}, $p_{\mathrm{T}t}$, $\Delta\varphi\left(\gamma\gamma, jj\right)$, $\Delta R_{\mathrm{min}}\left(\gamma, j\right)$, $\Delta\eta^{*} \equiv |\eta^{\gamma\gamma} - \left(\eta^{j1} + \eta^{j2}\right)/2|$. Two categories were defined from the discriminant output: one 'tight' category and a second 'loose' category. Systematics on the BDT itself were evaluated by comparing $(Z \to ee) + 2$jet events in data and MC. Systematics on the jet energy, underlying event, η^{*} modelling, etc. are included Table G.2. As for the hadronic Vh category, there are large uncertainties on contributions from $ggh + 2$ jets, particularly on the $\Delta\varphi\left(\gamma\gamma, jj\right)$ variable that responds to additional soft radiation.

6.2.3.3 Addendum: Recent Additions and Continued Work

Due to concerns at the time over low-statistics fits, cuts on additional variables were not applied to the $E_{\mathrm{T}}^{\mathrm{miss}}$ category, for the early papers. The potential for additional discrimination by 'confirming' the missing energy through either the boost of the diphoton system $p_{\mathrm{T}}^{\gamma\gamma}$ or its angular separation from the $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\varphi\left(\gamma\gamma, E_{\mathrm{T}}^{\mathrm{miss}}\right)$, is included in Appendix E and will be used in the ATLAS Run I 'legacy' analyses.

Explicitly identifying events with two same flavor, opposite sign leptons from a Z peak provides discrimination of the Wh and Zh processes, though the rate is exceedingly low ($\mathcal{B}\left(Z \to \ell\ell\right) = 0.066$). The search for $t\bar{t}\left(h \to \gamma\gamma\right)$ has also been released [136], and will be included in the final couplings measurements.

6.3 Results

Figure 6.3a shows the punchline of the preceding work: the categories defined have dramatically enhanced fractions of the rare production modes. The lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ categories in particular are $> 95\%$ Vh and $t\bar{t}h$. This allows for the measurement of the signal strength, splitting gluon fusion and $t\bar{t}h$ (top type couplings) from VBF and Vh associated production (vector-type couplings).

Knowing the expected signal composition in each category makes it possible to parameterize the yields of each category as a function of the signal strengths μ in each production mode: μ_{ggh}, μ_{Vh}, etc. Since the yields are measured, one can 'solve' for the signal strengths.

This is accomplished using the same fundamental techniques described in Section 5.6. Instead of solving for a single production rate, one simply uses two parameters of interest. In this case, the two parameters are the signal strengths $\mu_{ggh+t\bar{t}h}$ and $\mu_{\text{VBF}+Vh}$. Figure 6.3b shows the resultant contour. The measured value is consistent with the SM expectation at the level of 2σ. This is performed at the best fit mass of $m_h = 126.8$ GeV. In the final ATLAS results, the combined mass of 125.5 GeV [13] will be used, which will decrease the signal strength in this decay channel.

6.3.1 Combinations with other Channels

Naturally $h \to \gamma\gamma$ was not the only decay mode pursuing couplings measurements. All ATLAS Higgs measurements have been combined into an extremely powerful framework for studying each type of Higgs coupling. For instance, whereas $h \to \gamma\gamma$ is sensitive to the WWh coupling in its decay loop and in VBF or Wh associated production, the $h \to WW$ process also measures it directly in its decay. Similarly, $h \to ZZ$ provides unique power in the ZZh vertex and $h \to \tau\tau$ shows the Higgs coupling to leptons.

By parameterizing both the production and decay in every ATLAS category, powerful fits are possible. The results of these fits are presented in Figure 6.4. Unfortunately none of them hint at any non-SM feature of the dataset [139]:

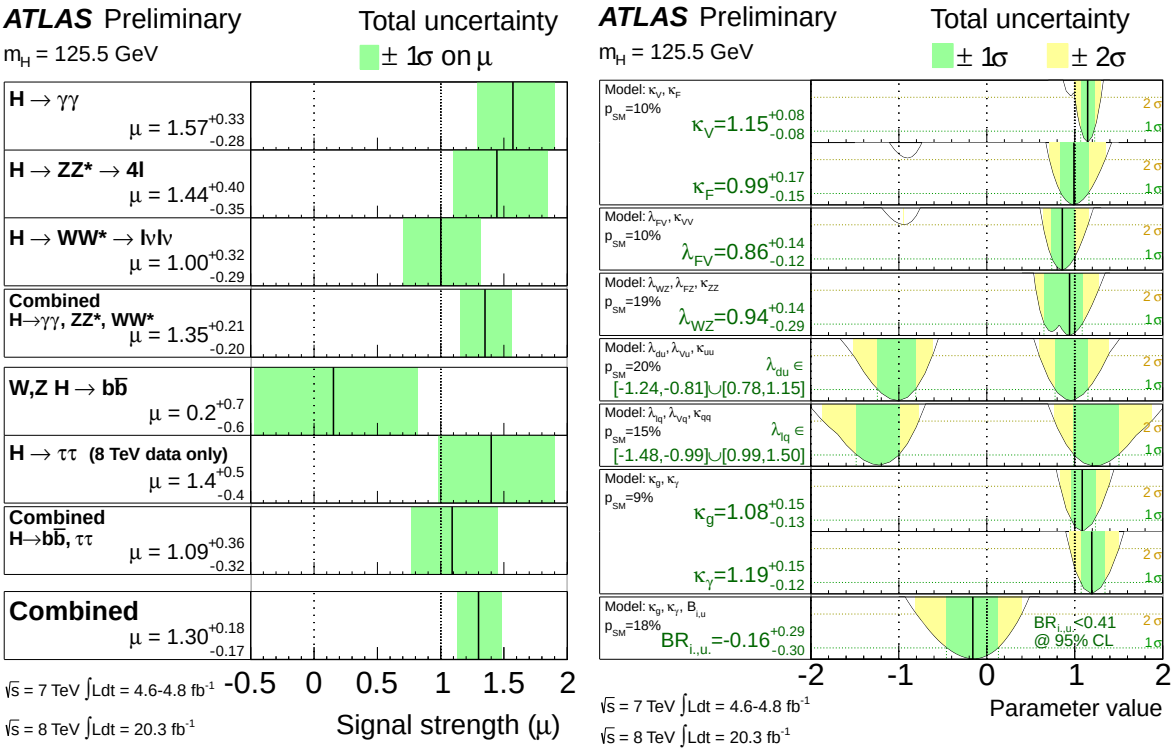
- ▷ The coupling to Ws and Zs agrees with the SM prediction, directly validating custodial symmetry in the Higgs sector.
- ▷ Couplings to up and down type quarks seem consistent with SM prediction, constraining the parameter space available for (Type II) 2HDMs.
- ▷ Allowing the effective ggh and $h\gamma\gamma$ couplings to float tests the potential for enhancements of the loops, and would be sensitive to additional electrically or strongly charged particles. The fit is consistent with the SM expectation.
- ▷ By fitting the total rate, it is possible to constrain the Higgs branching ratio to unobserved or invisible decays. The rate is consistent with 0, though the errors remain large.

The Higgs Boson, at Last

Within current uncertainties, the particle discovered couples as the Higgs boson. The spin and CP eigenvalues 0^+ are overwhelming favored over every other tested hypothesis [140]. The particle discovered indeed appears to be the SM Higgs boson.[6]

The consistency of the observed particle to the properties of the SM Higgs boson severely constrains many BSM interpretations, squeezing the parameter space for 2HDMs or additional generations. Nevertheless, the discovery opens new doors for particle physics. Even a small deviation from the SM predictions would be tremendously exciting. The phenomenology of the Higgs sector is outstandingly rich. So the adventure continues.

[6] 'If it walks like a duck and it quacks like a duck, then it is the Higgs boson.' – Konstantinos Nikolopoulos



(a) Rate by Channel

(b) Couplings Compared to SM

Figure 6.4: The combination of the couplings measurements between all decay channels shows overall and individual production rates consistent with the SM (a). Breaking down the categories according to various types of couplings – fermionic v. bosonic, Ws v. Zs, quark v. leptons, etc. – shows overall extremely good agreement to the SM in all parameters. The right-hand plot shows couplings as deviations from their values in the SM, so $\kappa_{\rm SM} \equiv 1$. The couplings to fermions and vector bosons κ_F and κ_V are consistent with 1. The ratio of the couplings to Ws and Zs $\lambda_{WZ} = \kappa_W/\kappa_Z$ is consistent with 1 (custodial symmetry). The relative couplings of up and down type quarks as well as leptons, $\lambda_{du} = \kappa_u/\kappa_d$ and $\lambda_{\ell q} = \kappa_\ell/\kappa_q$, are consistent with 1. The effective couplings to gluons (κ_g) and photons (κ_γ) are consistent with 1, and the branching to invisible or unobserved states $\mathcal{B}_{i,u}$ is consistent with 0.

Chapter 7

Differential Cross Sections

The new particle interacts as the Higgs boson; beyond all reasonable doubt, it is *a* Higgs boson – and one remarkably similar to that predicted by the SM. Differential cross sections measurements are a natural next step: they cast a broad net for deviations from SM expectations that might hint at new physics in the Higgs sector; failing any exciting hints, they nevertheless begin to confront (sometimes very-advanced) calculations with (admittedly, statistics-limited) reality. The measurements are also natural from an experimental perspective. Most of the techniques and inputs can simply be 'recycled' from the baseline and 'couplings' analyses. Yields are extracted in bins of each physical observable using simultaneous fits of the signal plus background; these yields are unfolded to the actual production cross-sections (called particle-level, or truth-level throughout) using simple correction factors.

Seven variables are extracted from the $h \to \gamma\gamma$ data set at $\sqrt{s} = 8$ TeV, and listed in Table 7.1. The motivations for the chosen variables will be explored when interpreting the results in Section 7.6. Since the same procedure was executed for each observable, the details of each measurement (shapes used in the fit, uncertainties, yields, etc.) are summarized in Tables H.1-H.7 of the Appendix.

7.1 Selection Requirements

The event and object selection requirements largely follow the ones used for the baseline and couplings analyses. There is an important addition, however. In order to unfold the cross sections, the selection

	Variable	Definition
Inclusive	$p_T^{\gamma\gamma}$	Transverse momentum of the diphoton system.
	$\lvert y^{\gamma\gamma}\rvert$	Rapidity of the diphoton system.
	$\lvert\cos\theta^*\rvert$	Helicity angle in the Collins-Soper frame.
	N_{jets}	Jet multiplicity.
	p_T^{j1}	Transverse momentum of the hardest parton emission.
2-jets	$\Delta\varphi_{jj}$	Azimuthal separation of two hardest jets.
	$p_T^{\gamma\gamma jj}$	Transvserse momentum of diphoton plus dijet system.

Table 7.1: Extracted differential cross sections.

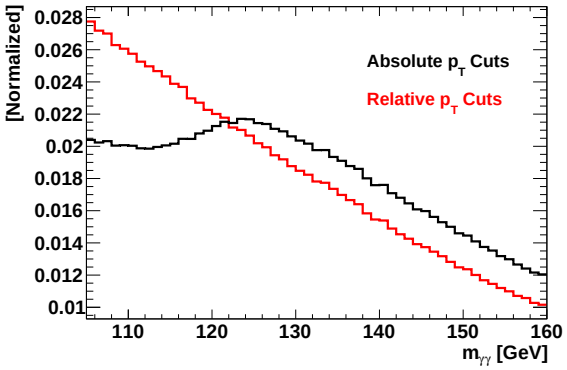


Figure 7.1: The invariant mass spectrum is presented from high-statistics background MC, for events with $80 \leq p_{\mathrm{T}}^{\gamma\gamma} < 100$ GeV. It is clear that the use of relative cuts on $p_{\mathrm{T}}/m^{\gamma\gamma}$ makes for far simpler shapes than the original/default absolute cuts on the two photons' p_{T}s.

requirements are defined twice: once for data and fully-simulated and reconstructed MC, and a second time for particle-level generated events.

7.1.1 Reconstructed Events and Data

The trigger and photon kinematic, identification, and isolation requirements are lifted from earlier $\sqrt{s} =$ 8 TeV analyses. The geometrical acceptance of the analysis is unchanged: $|\eta| < 2.37$ excluding the crack region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.56$).

Three modifications are made to the common selection:

(1) Cuts or bins of certain variables – especially $p_{\mathrm{T}}^{\gamma\gamma}$ and $|\cos\theta^*|$ – were found to induce dramatic deformations of the background shape in $m^{\gamma\gamma}$. These deformations make it difficult to fit the distributions with a small number of parameters, and lead to large uncertainties on the background parameterization ('spurious signal,' Section 5.3.2). Figure 7.1 shows the invariant mass distribution of a high-p_{T} slice of the combined diphoton, photon-jet, and dijet continuum, taken from smeared truth Monte Carlo. The absolute cuts on the p_{T}s are compared to the 'relative' cuts on the ratio of the (sub)leading photon p_{T} to the combined invariant mass of the diphoton system of $p_{\mathrm{T}}/m^{\gamma\gamma} > 0.35$ (0.25). Distributions of $m^{\gamma\gamma}$ resulting from the relative cuts are far easier to model.

(2) Because the trigger selected two photons greater than 35 (25) GeV, the lower cut on the diphoton invariant mass was raised to 105 GeV to avoid turn-on effects on the subleading photon.

(3) Rather than using an $|\eta|$-dependent p_{T} cut on the jets, a flat cut of 30 GeV was made. Further, rapidity was used in place of pseudorapidity, with a cut of $|y^j| < 4.4$.

With 20.3/fb of data in 2012, ATLAS collected 94135 events satisfying the above requirements. Signal yields expected from MC are listed in Table 7.2.

7.1.2 Truth-Level Events: Definition of the Fiducial Region

7.1.2.1 Definition of the Fiducial Region

The particle-level selection defines the fiducial region of the differential cross sections. To minimize the extrapolation made by the correction factors, this definition mirrors the reconstructed level cuts so far as possible, except that the crack region between the calorimeters is not removed. The two highest-E_{T}, isolated

	Fiducial	Fully Simulated and Selected Signal					
	Signal (MC)	Yield	ggh [%]	VBF [%]	Wh [%]	Zh [%]	$t\bar{t}h$ [%]
Inclusive	612	407	87.9	7.3	2.8	1.6	0.5
≥ 1 Jet	252	180	76.3	15.1	4.8	2.7	1.1
≥ 2 Jets	85	64	59.4	25.6	7.6	4.4	2.9

Table 7.2: Expected selected yields from the Standard Model ($\mu = 1$) for 20.3/fb of data at $\sqrt{s} = 8$ TeV with $m_h = 125$ GeV. The relative fractions selected in the five production modes, ggh, VBF, Wh, Zh, and $t\bar{t}h$ are detailed. The fiducial signal is calculated at generator level.

final state photons within $|\eta| < 2.37$ are selected. The cut of $E_{\mathrm{T}}/m_{\gamma\gamma} > 0.35$ (0.25) is applied to the (sub)leading photon as in data, and the combined mass is required to satisfy 105 GeV $< m_{\gamma\gamma} <$ 160 GeV.

7.1.2.2 Definition and Motivation of the Isolation Requirement

The particle level isolation is defined as the scalar sum of the p_{T} of all stable particles[7] within $\Delta R < 0.4$ of the photon, excluding muons and neutrinos. The cut was set at 14 GeV, to correspond approximately to the calorimetric isolation cut of 6 GeV at reconstruction level. The track isolation cut at detector level is much looser, and does not impact the derivation of this value.

The isolation requirement reduces the rate at which non-signal photons (e.g., $\pi^0 \to \gamma\gamma$) are selected, and thereby reduces the dependence of the measured cross sections on the model used to generate the unfolding corrections (Section 7.3). Without any isolation cut, the correction factors derived for ggh and $t\bar{t}h$ are quite different. In ggh production the two photons from the Higgs decay are almost always selected both at reconstruction and truth level, because there are few other objects *to* select. In $t\bar{t}h$ production however, the objects preselected at reconstructed level are less likely to be the true $h \to \gamma\gamma$ signal photons. If jets are selected, they are unlikely to satisfy the reconstructed-level isolation requirement, and if they *do* pass, the 'diphoton' pair is unlikely to sit at the Higgs mass. Even if the correct 'pair' is selected, the isolation efficiencies of true photons in ggh and $t\bar{t}h$ are different. Applying the truth-level isolation cut eliminates mis-pairing and allows the truth-level cuts better to emulate the cuts applied at reconstructed-level. Together, these lead to more-consistent correction factors (Section 7.3.1) between production modes, and smaller uncertainties (Section 7.4.3).

7.1.2.3 Jet Definition

Truth jets are defined using the anti-k_t algorithm with distance parameter $R = 0.4$, as at reconstruction level. All stable particles excluding muons and neutrinos serve as input. As at reconstruction level, jets must have $p_{\mathrm{T}} > 30$ GeV and $|y| < 4.4$.

7.2 Signal Extraction

7.2.1 Binning of the Observables

The differential cross sections are extracted bin-by-bin in each of the physical observables. A fine binning is naturally appealing since it provides greater shape information, but this must be balanced against the available statistics and the limitations of the method. The binning was chosen to allow differential measurements, with statistics sufficient for a significant measurement in each bin, using the expected signal yields from simulation. The nominal target significances were 2σ (1.5σ) per bin for the inclusive (2-jet) variables.

[7] A particle is considered stable if it has a lifetime of more than 10 ps.

To obtain reliable unfolding factors, the migrations between bins in the reconstructed distributions must be small; the target 'purity' of events reconstructed into the bin in which they were generated was set to 75%. The binnings chosen based on these considerations are displayed in the variable summary tables in Appendix H.

7.2.2 Fit Procedure and Yield Extraction

As in earlier Chapters, the keystone of this measurement is a signal plus background fit of the diphoton invariant mass spectrum. The fitting procedure used here differs from the couplings analyses, in that the signal yields there (μ value) were correlated between categories, whereas in this measurement they are independent.

An unbinned fit is performed simultaneously in all bins, for each observable. The Higgs boson mass m_h and the nuisance parameters on the signal shape and position are common among all bins for each observable. The likelihood function maximized has the form

$$\mathcal{L}\left(\boldsymbol{m}_{\gamma\gamma};\boldsymbol{\nu}^{\text{sig}},\boldsymbol{\nu}^{\text{bkg}},m_h\right)=\prod_i\left\{\frac{e^{-\nu_i}}{n_i!}\prod_j^{n_i}\left[\nu_i^{\text{sig}}\,\mathcal{S}_i\left(m_{\gamma\gamma}^j;m_h\right)+\nu_i^{\text{bkg}}\,\mathcal{B}_i\left(m_{\gamma\gamma}^j\right)\right]\right\}\times\prod_k\mathcal{G}_k \tag{7.1}$$

with ν_i^{sig} and ν_i^{bkg} being the number of signal and background events estimated in data in the i^{th} bin of the observable, $\nu_i=\nu_i^{\text{sig}}+\nu_i^{\text{bkg}}$ the mean value of the underlying Poisson distribution of the n_i events, and $m_{\gamma\gamma}^j$ is the diphoton mass for event j.

The probability density functions of the signal $\mathcal{S}_i$ and background $\mathcal{B}_i$ are defined and derived according to the methods outlined in Section 5.3. The signal PDFs depend on the Higgs boson mass m_h and on nuisance parameters from the energy resolution and scale. The term $\mathcal{G}_k$ is a function of the k^{th} nuisance parameter and implements constraints from the photon energy resolution and scale into the fit. Uncertainties that do not affect the shape of the fit, for instance the background model uncertainty or trigger, are not included at this stage. Rather, they are applied during the unfolding procedure (Section 7.3).

For observables where a set of events is not included in the measured spectrum, the un-categorized events are placed into an additional bin, which is included in the fit. These include, for example, events with $p_{\text{T}}^{\gamma\gamma}$ larger than the upper edge of the highest bin, or events with 0 or 1 jet for $\Delta\varphi_{jj}$. The events in this additional bin help to constrain the Higgs boson mass and other nuisance parameters.

The fitted yields and errors are validated using an ensemble of pseudo-experiments and the statistical component of the error on the total yield is separated using a second fit with the nuisance parameters fixed to their profiled values. Example fits are shown for the jet multiplicity binning, in Figure 7.2.

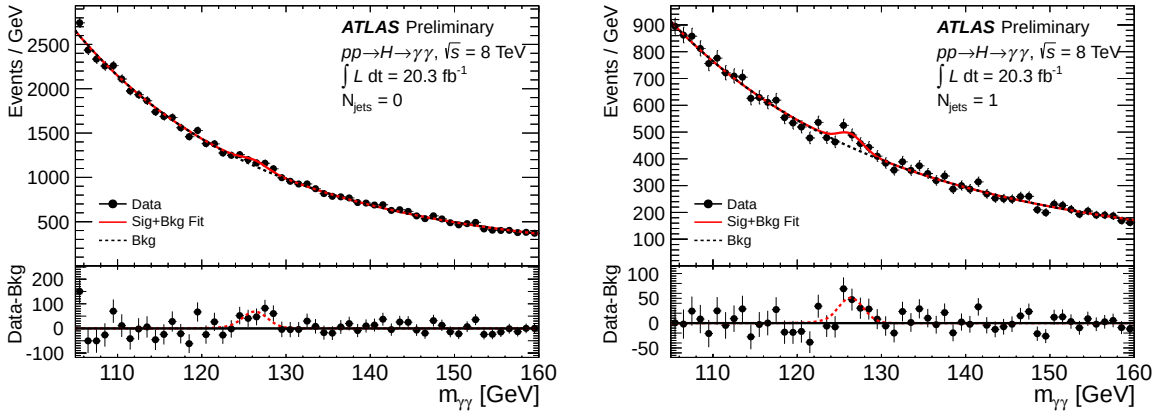
7.3 Unfolding Procedure

7.3.1 Correction Factors

The data yields extracted in the previous section are corrected for detector effects using bin-by-bin factors. These are derived as the ratio of the yields from particle level to reconstruction level from simulated Higgs boson events, according to the SM expectation listed in Table 7.2. In each bin,

$$c_i=n_i^{\text{Particle level}}\big/\,n_i^{\text{Reconstructed}} \tag{7.2}$$

is used to correct the extracted signal yield in data. This unfolding procedure corrects for all efficiencies, acceptances, and resolution effects. The correction factors range from 1.2 to 1.8, and include the extrapolation (about 20%, across all bins and observables) over the small region in $|\eta|$ excluded from reconstructed photon candidates. The correction factors for all bins are shown in Figure 7.3, and tabulated in Appendix H.



(a) $N_{\text{jets}} = 0$ (b) $N_{\text{jets}} = 2$

Figure 7.2: Examples of the fits performed in the diphoton invariant mass are shown for the first and second bin of the N_{jets} partial cross sections.

The method is formally unbiased provided that

$$c_i^{\text{MC}} = c_i^{\text{Data}} \,. \tag{7.3}$$

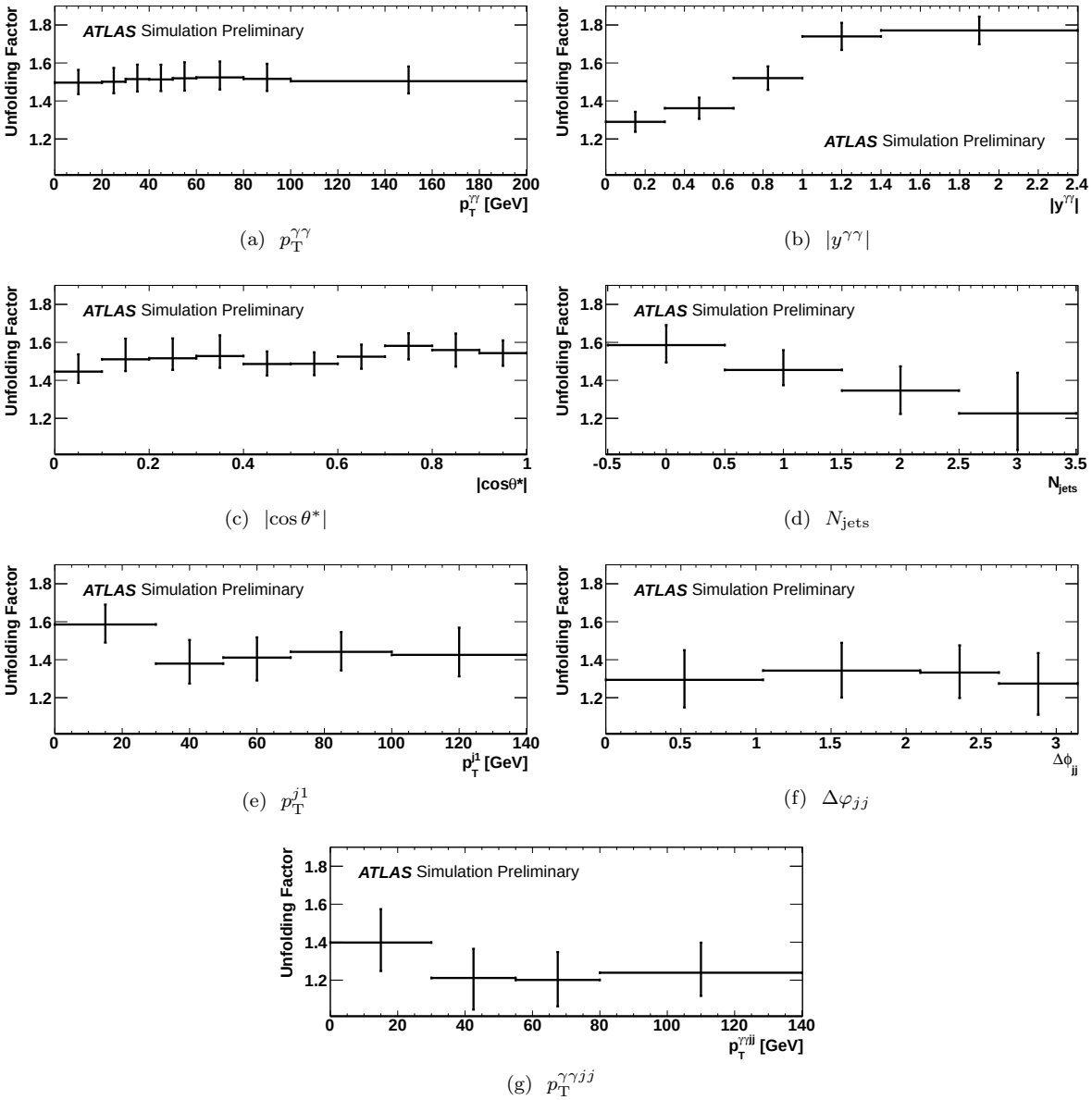
In practice, the requirement to use this method is that the 'purity' of events reconstructed into the same bin in which they were generated should not be too low. Among the measured observables, $p_{\text{T}}^{\gamma\gamma}$, $|\cos\theta^*|$, and $|y^{\gamma\gamma}|$ have very high purity ($> 87\%$); for the jet variables, it can be as low as 50%. The lower purity for the jet variables may be understood as a 'double migration': first for passing or failing the jet definition, and second for the migrations between bins of the observable. Purities for $p_{\text{T}}^{\gamma\gamma}$ and $\Delta\varphi_{jj}$ are shown in Figure 7.4. The purities and correction factors for each bin of each observable are presented in the summary tables in Appendix H. Uncertainties from the choice of the Monte Carlo are discussed in Section 7.4.3, below.

7.3.2 Alternative Method: Bayesian Unfolding

Simple correction factors are appropriate for statistics-limited analyses like the present one, but many more 'advanced' techniques are available and were considered. In particular, the Bayesian iterative unfolding method [141] of the RooUnfold package [142] was implemented. Bayesian unfolding is motivated by correcting the input distribution (prior) used to derive the correction factors. Figure 7.5 shows the extracted yields for two bins, using the Bayesian unfolding. For most bins, the agreement between the Bayesian unfolding and the correction factors is quite good and the Bayesian method typically converges very quickly. However, a few bins did not converge, but rather 'oscillated' between iterations. This is a known feature of the method, with low statistics; based on this experience, the more-robust correction factors were preserved.

7.4 Systematic Uncertainties

There are two basic classes of uncertainties in this measurement: uncertainties that impact the signal shape (mass resolution or peak position), and those that do not. The former are included in $\mathcal{G}_k$ listed in Equation 7.1, while the latter are implemented as uncertainties on the correction factors. These include both flat, overall uncertainties like the luminosity, as well as uncertainties on migrations between bins like the jet energy scale (in the p_{T}^{j1} measurement) and uncertainties on the correction factors themselves, from the choice of the model.



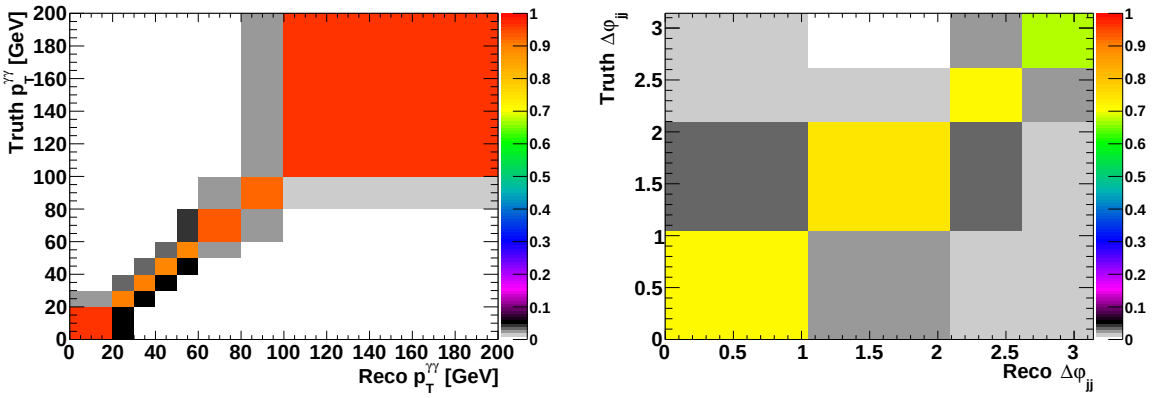
(a) $p_{\mathrm{T}}^{\gamma\gamma}$ (b) $|y^{\gamma\gamma}|$

(c) $|\cos\theta^*|$ (d) N_{jets}

(e) p_{T}^{j1} (f) $\Delta\varphi_{jj}$

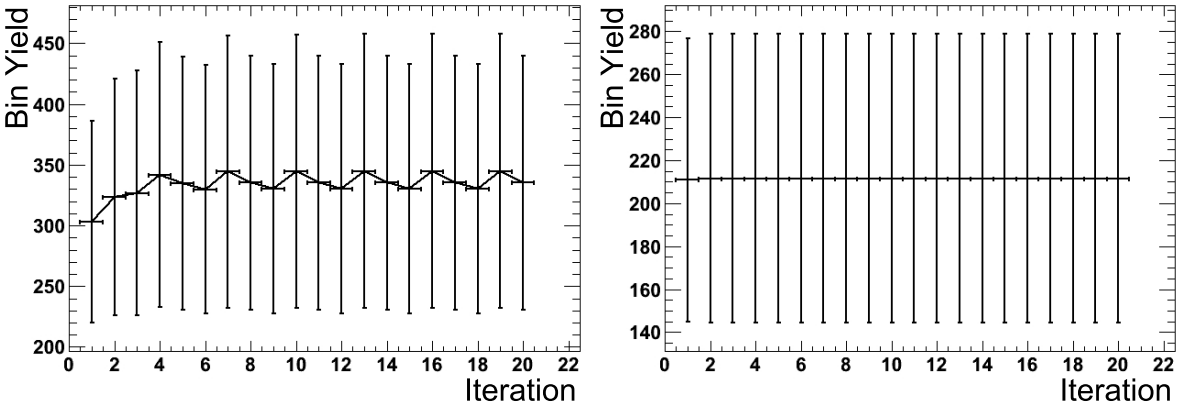
(g) $p_{\mathrm{T}}^{\gamma\gamma jj}$

Figure 7.3: Final correction factors are displayed, along with the full modelling uncertainties on the unfolding.



(a) $p_{\mathrm{T}}^{\gamma\gamma}$ (b) $\Delta\varphi_{jj}$

Figure 7.4: Purities are shown for events generated with a given $p_{\mathrm{T}}^{\gamma\gamma}$ or $\Delta\varphi_{jj}$ to be reconstructed into the correct bin. This 'purity' is quite high for $p_{\mathrm{T}}^{\gamma\gamma}$ since the photon energy resolution is good and the bin widths are large. For $\Delta\varphi_{jj}$ (and other 2-jet variables), the purities are much lower. This is because *two* large migrations contribute: the uncertainty of reconstructing the (right) jets *at all*, and the broader angular and energy resolution for jets.



(a) $20 \leq p_{\mathrm{T}}^{\gamma\gamma} < 30$ GeV (b) $|y^{\gamma\gamma}| < 0.3$

Figure 7.5: Bayesian iterative unfolding was implemented as an alternative to the simple correction factor method. While most bins quickly converge to a stable value, several bins in the Bayesian method 'oscillate' between iterations and do not converge. This motivated the use of the simpler, more robust correction factors.

7.4.1 Shape and Modelling Uncertainties

The output of the fit yields the extracted value and its uncertainty, which is overwhelmingly statistical. Uncertainties on the resolution of the signal model do affect the extracted yield; the uncertainty on both the mass resolution and the ultimate (statistical + resolution) uncertainty on the measurement in each bin is listed in the summary tables of Appendix H.

7.4.2 Uncertainties Shared with Previous Results

Many uncertainties are common with previous results; only the form of their implementation changes. The variation in yields due to uncertainties in the luminosity (2.8%), trigger (0.5%), photon identification (2.4%), and photon isolation (1-4% according to the number of jets) are all common with earlier results, and the resultant error is combined in quadrature in the total error.

Improper modelling of the jet energy scale or resolution (for example) could engender migrations between bins. This obviously does not affect the measured yields; instead, it alters the unfolding. For these cases, each variation of each uncertainty is propagated in the fully reconstructed sample, and the correction factors are recalculated from this 'altered' sample. The difference between the nominal and distorted correction factors is the uncertainty.

Three sources of uncertainties are considered for the variables that are based on jets: jet energy resolution (JER) and scale (JES), jet vertex fraction (JVF), and jets from pileup (that do not originate from the primary vertex).

- ▷ The jet energy scale and resolution uncertainties are estimated by applying shifts and smearings to the jet energy within their expected uncertainties. These shifts are derived from the data, exploiting the transverse momentum balance in $\gamma+jet$, $Z+jet$, dijet and multijet events. Discrepancies between data and MC for the jet energy scale for these measurements lead to a set of baseline uncertainties. The most important JES uncertainties for this analysis are from the η-intercalibration that particularly impacts the calibration of forward jets, and due to the unknown composition and modelling of the associated calorimeter response of quark and gluon initiated jets.
- ▷ Uncertainties from JVF modelling are quantified by varying the JVF cut up and down according to its uncertainty around the nominal cut value of 0.25. The JVF uncertainty is estimated by comparing simulation with data in $Z+$jet events and is parametrized as a function of jet p_{T} and η.
- ▷ The uncertainty associated with the modelling of jets originating from pileup interactions is evaluated by randomly subtracting a fraction of the simulated pileup jets. The fraction of pileup jets removed is estimated by comparing the data to MC ratio of jets in pile-up enriched control regions of $Z+jets$ events.

The combined impact of these uncertainties on the signal yields ranges from 3-15% according to the observable and bin, and is tabulated in the Appendix.

Migrations between bins due to the photon energy scale were found to be negligible. The contamination of the jet related observables, due to simultaneous dijet and Higgs boson production through double parton interactions (DPI) was evaluated using the expected dijet cross section with the jet requirements of the analysis, with the measured effective area parameter for hard DPI [143]. The impact was also found to be negligible, and was dropped.

7.4.3 Uncertainties on the correction factors, from the choice of model

To the extent that the physics processes and simulation model do not perfectly reproduce the data, the correction factors defined in Eq. 7.3 will be biased and model dependent. This dependence is evaluated and taken as an uncertainty on the method. The potential bias can be categorized in two aspects: sample composition and shape.

The sample composition and description affect correction factors, because different production processes or generators may have different reconstruction level efficiencies. For instance the larger jet activity in $t\bar{t}h$ events results in lower isolation efficiencies and larger correction factors than ggh. The shape of the input distributions matter, because the net migrations in a distribution will be from a more populated bin towards its less populated neighbour. The uncertainties from these considerations are evaluated by manipulating the Monte Carlo from which the correction factors are determined in four ways:

(1) Alter the composition from the SM expectation by varying the relative VBF and $Wh + Zh$ fractions up ($2\times$) and down ($0.5\times$), motivated by the 1σ contour presented in the $h \to \gamma\gamma$ coupling results [13]. Since the $t\bar{t}h$ fraction was not yet well-constrained by those results, it is varied by factors of 6 and 0, motivated by the combined up variations necessary to fully cover the observed excess for large N_{jets} in data (Figure 7.8d).

(2) Estimate the uncertainties from missing higher order perturbative corrections in ggh production, by varying the factorization and renormalization scales up ($2\times$) and down ($0.5\times$). To further test the dependence of the modelling of ggh production, replace the nominal POWHEG + Pythia8 sample with the leading order matrix element multijet merged prediction of Sherpa.

(3) Disable multi-parton interactions to test the dependence on the modelling of the underlying event.

(4) Reweight the combined SM prediction in $p_{\text{T}}^{\gamma\gamma}$ and $|y^{\gamma\gamma}|$ or p_{T}^{j1} spectra to match the unfolded results in these variables. This is effectively the first 'iteration' of Bayesian unfolding.

For each of these variations new correction factors are derived, and the shifts with respect to the nominal SM correction factor are assigned as uncertainties. These four methods are not orthogonal: changing samples and reweighting impacts the predicted spectra. The full uncertainty is taken thus as the envelope of the data-based reweighting, with the quadratic sum of the composition, scale and underlying event variations.

7.4.4 Summary of the Uncertainties

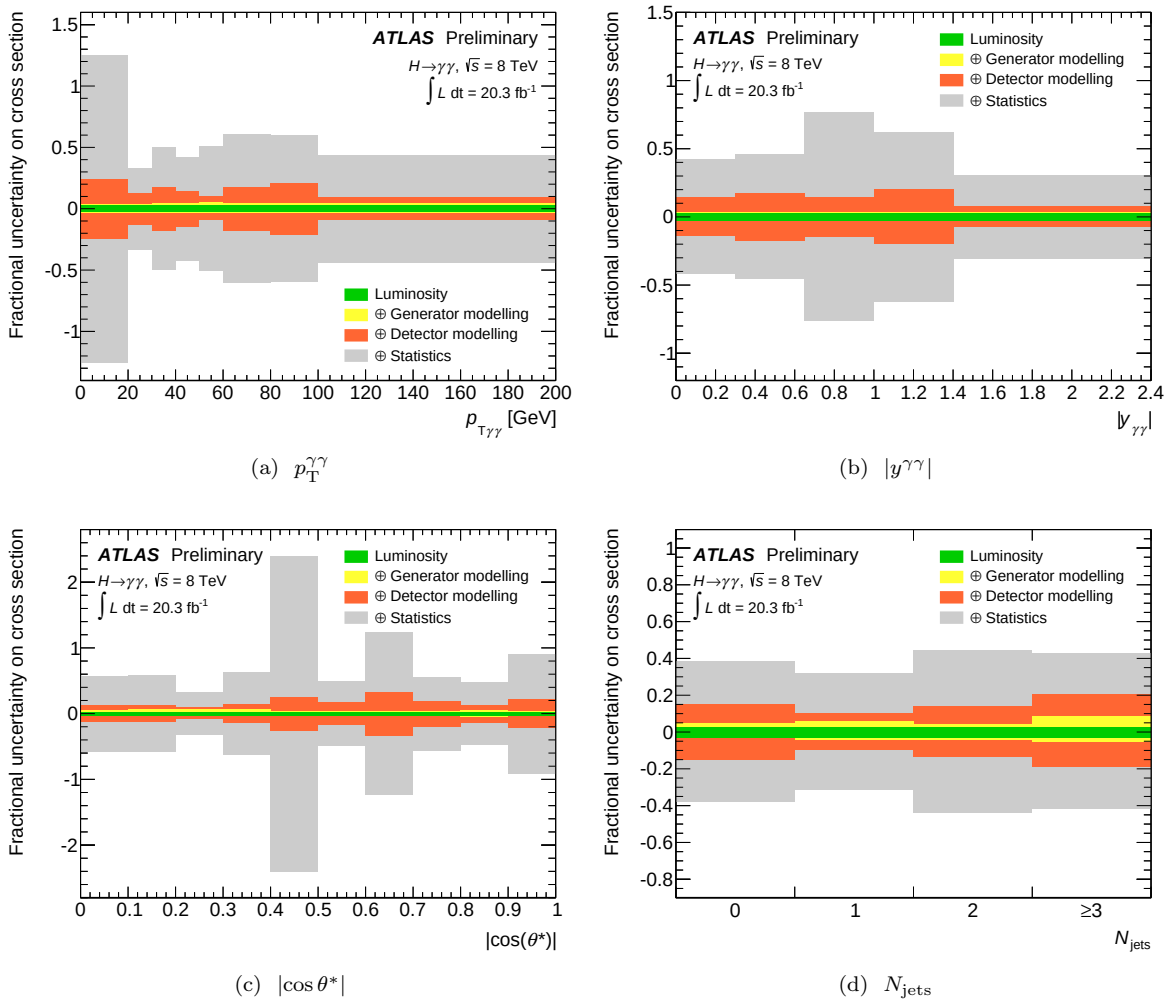
Figures 7.6 and 7.7 present the combination of all uncertainties. The statistical uncertainty (grey) dominates overwhelmingly. In practice, this means (a) that the correlations between bins of the final measurement are small and (b) that the analysis will benefit enormously with $\sigma_{13\text{ TeV}}/\sigma_{8\text{ TeV}} \approx 2$ and 100/fb expected for Run II.

7.5 Theoretical Predictions

The extracted results were compared to state of the art predictions. In all cases, the NLO prediction from POWHEG BOX is used for ggh and VBF production, with the leading order Wh/Zh and $t\bar{t}h$ contributions from Pythia8. These are the predictions used as the baseline for all ATLAS measurements (though, as detailed in Section 5.2.1.1, many corrections are made to these). For each prediction, the mass of the Higgs boson was set to the best fit value from $h \to \gamma\gamma$ available at the time of this measurement, of $m_h = 126.8$ GeV.[8] This value was consistent with the one extracted in each of the differential cross sections.

Two higher-order predictions were produced for ggh (also, with $m_h = 126.8$ GeV). Multi-scale improved NLO (MINLO, rev. 2290) for H + 1 jet proffers a procedure for smoothly merging between Higgs + N parton production and the cross section with $(N-1)$ partons and Sudakov form factors for additional radiation [145]. The events are generated using the CT10 PDF set [111,112,146] and interfaced with Pythia8 for the simulation of underlying event, showering, and hadronization. The second prediction, HRes 1.0 fully resums the logarithms, and is accurate to NNLO+NNLL. It uses the infinite top quark mass approximation; the finite quark mass treatment was released with HRes 2.0. Events are generated with the MSTW 2008 NNLO PDFs [147]. Because HRes is a parton level prediction with an inclusive treatment of the radiative

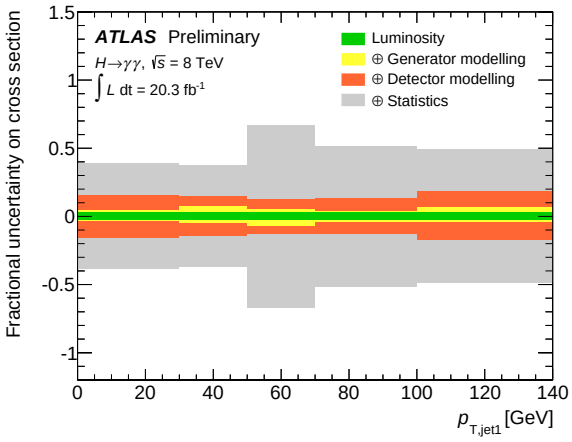
[8] The final ATLAS measurement for the Higgs mass yielded $m_h = 125.36 \pm 0.37\,(\text{stat}) \pm 0.18\,(\text{syst})$ GeV [144].



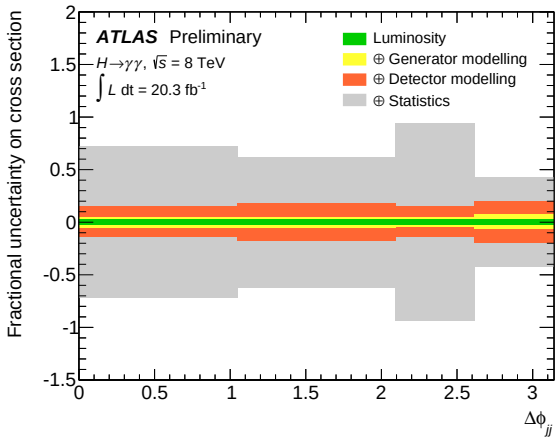
(a) $p_{\mathrm{T}}^{\gamma\gamma}$ (b) $|y^{\gamma\gamma}|$

(c) $|\cos\theta^*|$ (d) N_{jets}

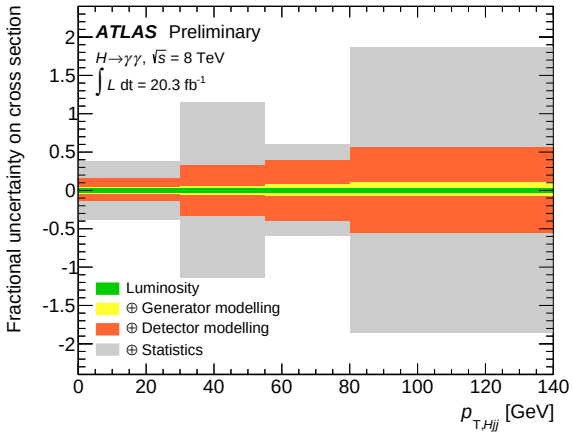
Figure 7.6: Contributions to the uncertainties on the differential cross sections are presented for $p_{\mathrm{T}}^{\gamma\gamma}$, $|y^{\gamma\gamma}|$, $|\cos\theta^*|$, and N_{jets}. The uncertainties are presented as a fraction of the extracted value. The statistical uncertainty is, by far, the dominant component.



(a) p_{T}^{j1}



(b) $\Delta\varphi_{jj}$



(c) $p_{\mathrm{T}}^{\gamma\gamma jj}$

Figure 7.7: Contributions to the uncertainties on the differential cross sections are presented for p_{T}^{j1}, $\Delta\varphi_{jj}$, and $p_{\mathrm{T}}^{\gamma\gamma jj}$. The uncertainties are presented as a fraction of the extracted value. The statistical uncertainty is, by far, the dominant component.

	N_{jets}	$p_{\text{T}}^{\gamma\gamma}$	$\lvert y^{\gamma\gamma}\rvert$	$\lvert\cos\theta^*\rvert$	p_{T}^{j1}	$\Delta\varphi_{jj}$	$p_{\text{T}}^{\gamma\gamma jj}$
POWHEG	0.54	0.55	0.38	0.69	0.79	0.42	0.50
MINLO	0.44	–	–	0.67	0.73	0.45	0.49
HRes 1.0	–	0.39	0.44	–	–	–	–

Table 7.3: Displayed are the probabilities from χ^2 tests for the agreement between the unfolded observation and the theoretical predictions, calculated with the full covariance between bins of the observables.

corrections, it is not possible to apply any isolation requirement on this sample 7.1.2. The effect on the *ggh* sample is, however, expected to be quite small.

Two additional higher-order (NNLO) predictions are presented for the p_{T}^{j1} distribution, in Appendix I.

7.5.1 Errors on Theoretical Predictions

Just as experimental errors are included in the unfolded measured cross sections, several theoretical errors apply to the calculations performed. The $h \to \gamma\gamma$ branching ratio of 0.00228 has an uncertainty of 4.9%. Scale and PDF + α_s uncertainties are evaluated according to the standard methods – the uncertainty is calculated as the linear sum of the upwards and downwards variations of the scale with the PDF eigentune variations.

7.6 Results and Interpretation

The complete results are presented in Figures 7.8 and 7.9, overlaid with the predictions described above. In interpreting the results, two 'slants' are possible: one can either ask if the observation matches the SM, in which case the MC is taken as SM; or one can probe the reliability of the MC, in which case the measurement represents the SM. It is worth noting that the overall production rate in the diphoton channel is somewhat high – $\mu = 1.57$ at the time these cross sections were derived – and this overall enhancement of the cross section is seen in all observables. All variables display moderately good agreement to the theoretical calculations, and the probabilities of the χ^2 differences (including correlations between bins) are presented in Table 7.3. Nevertheless, some interpretation is in order.

Transverse momentum. Along with the Higgs boson's mass and rapidity, the p_{T} of the diphoton system defines the full kinematics of Higgs boson production (assuming symmetry in φ). The p_{T} spectrum is of particular interest since it is sensitive to both QCD in *ggh* production, and also to other SM or BSM production modes. Overall, the distribution is slightly hard – though entirely consistent within the statistical precision. This correlates with the larger-than-expected jet multiplicity, and potentially reflects the large k-factors recently predicted for higher-order Higgs + jets production [148]. The impact of the finite quark masses (not included in HRes 1.0) has an impact of ∼10% in the high-p_{T} tail.

Rapidity. The rapidity of the Higgs boson will eventually be an interesting probe of gluon PDFs, but this is still a long ways off. One bin is slightly high, at large rapdity.

Helicity angle. The helicity angle in the Collins-Soper frame [149] is sensitive to the spin of the particle. Earlier analyses of the spin in the diphoton channel have used this variable, but correlated the yields extracted between bins for each hypothesis [140]. The extraction here is entirely model independent, and provides the unbiased observation.

Jet multiplicity. The jet multiplicity is sensitive to the production model, as well as to additional QCD radiation. The multiplicity is perhaps slightly higher than expected, in agreement with the NNLO predictions already mentioned.

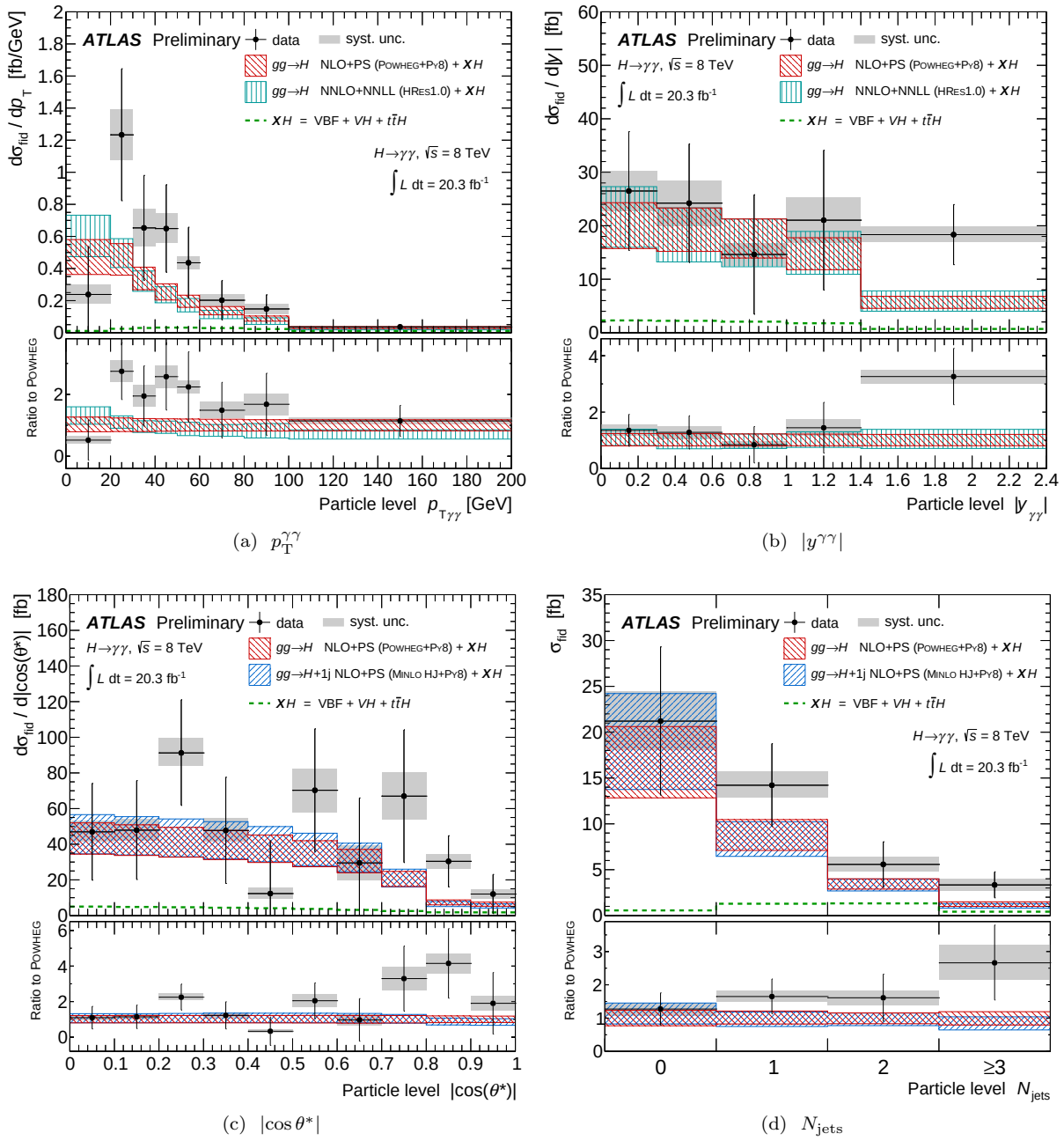
Transverse momentum of the leading jet. The hardest parton emission in Higgs boson production is primarily of interest for QCD. Allowing for the larger overall rate, it appears to be very-well predicted by the Monte Carlo. Two higher-order predictions by the JetVeto group [150] and Tackmann et al [151] are included in Appendix I. These naturally reduce the theoretical uncertainties, but they do not change the picture.

Azimuthal separation between the leading jets. The angular separation between between the two leading jets is one of the variables where the h+1 jet calculation with MINLO should improve angular description, since the second parton emission is included in the matrix element instead of just the parton shower algorithm. This is apparent in the opposite concavities of the two predictions. The most-striking feature of the distribution, however, is the $\sim 2\sigma$ excess observed in back-to-back jets. This topology resembles the simple overlay of a dijet event on a diphoton event – either through pileup or as a double parton interaction (DPI). Neither seems to be the case. The pileup hypothesis was studied by binning the dataset in low- and high-pileup halves to see if the effect was larger at high-pileup. In fact, more of the excess came from the low-pileup half of the data. The DPI rate was calculated with the inclusive dijet cross section, for the kinematic cuts of the analysis, using the effective DPI cross section of 15 mb measured by ATLAS [143]. The contribution from DPI is expected to be roughly 2% of that bin, and this effect should moreover be covered by the explicitly evaluated uncertainties on the underlying event.

With larger statistics, the ratio of the cross sections in $(\pi/3, 2\pi/3)$ over $(0, \pi/3)$ and $(2\pi/3, \pi)$ will be sensitive to the spin and CP eigenvalues of the Higgs boson, in both *ggh* and VBF production [152].

Transverse momentum of the diphoton plus dijet system. The transverse momentum of the combined diphoton + dijet system is effectively a measure of additional radiation in these events. Since VBF production has no color flow between quarks this radiation is suppressed, and the diphoton and dijet systems tend to be better balanced than in *ggh* production where there *is* color flow. The $p_{\mathrm{T}}^{\gamma\gamma jj}$ and the closely-correlated $\Delta\varphi\,(\gamma\gamma, jj)$ are therefore used in analyses of VBF production, but come with extremely large – if not divergent – uncertainties in the region of interest ($\Delta\varphi \sim \pi$). The first step towards reducing these errors is the direct measurement.

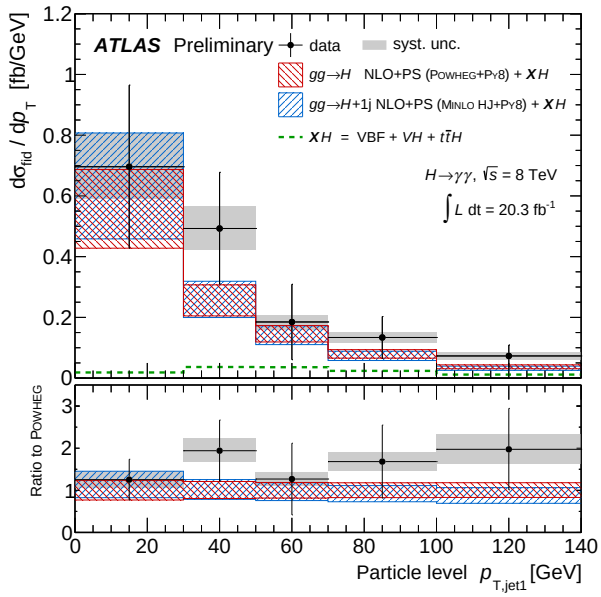
'Jet veto' efficiency. A final variable is defined as the ratio of the inclusive and exclusive partial cross sections in the number of jets, $\sigma_{N_{\mathrm{jets}}=i}/\sigma_{N_{\mathrm{jets}}\geq i}$. These ratios are calculated from the jet multiplicity spectrum, and presented in Figure 7.10. For pure gluon fusion the values would be equal for all i, since there is roughly equal probability of additional radiation at each higher order (related directly to α_s). In the present case, contamination from production modes with inherent $N_{\mathrm{jets}} > 0$ distorts that expectation.



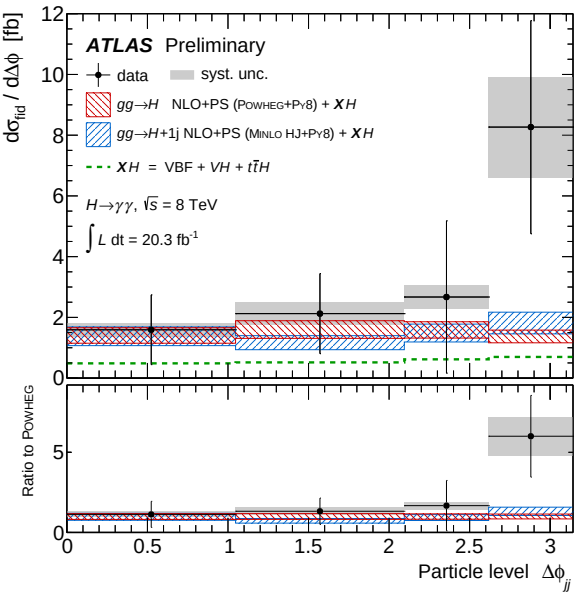
(a) $p_{\mathrm{T}}^{\gamma\gamma}$ (b) $|y^{\gamma\gamma}|$

(c) $|\cos\theta^*|$ (d) N_{jets}

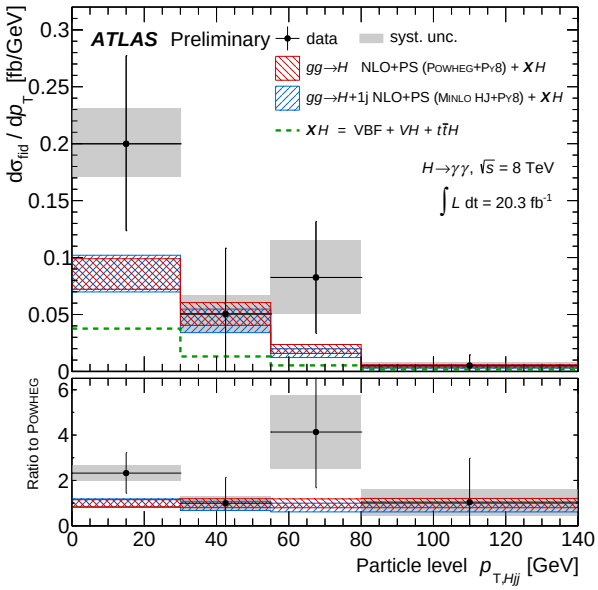
Figure 7.8: Fiducial differential cross sections of the Higgs boson are presented in $p_{\mathrm{T}}^{\gamma\gamma}$, $|y^{\gamma\gamma}|$, $|\cos\theta^*|$, and N_{jets}.



(a) p_{T}^{j1}



(b) $\Delta\varphi_{jj}$



(c) $p_{\mathrm{T}}^{\gamma\gamma jj}$

Figure 7.9: Fiducial differential cross sections of the Higgs boson are presented in p_{T}^{j1}, $\Delta\varphi_{jj}$, $p_{\mathrm{T}}^{\gamma\gamma jj}$.

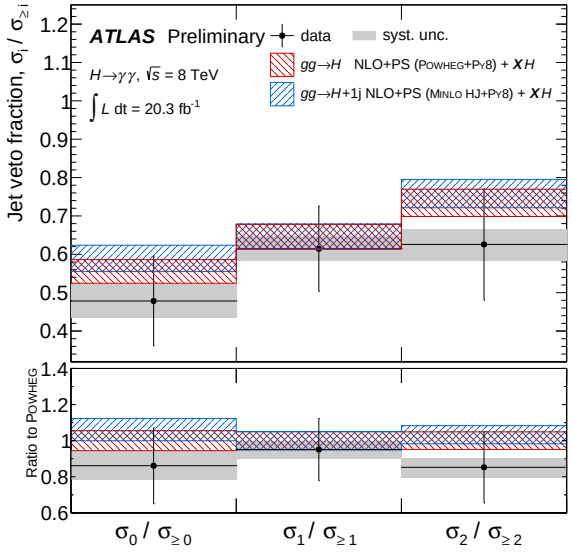


Figure 7.10: Jet veto distribution for Higgs boson production.

Chapter 8

Pair Production

The SM description of Higgs boson production is apparently vindicated in each of the tests proposed in the previous chapters of this thesis. This is impressive but also disappointing – discrepancies would be exciting! Despite the lack of immediate surprises, the Higgs sector offers a rich phenomenology and exciting prospects for new physics. This final chapter, describing a search for Higgs boson pair production, is therefore proposed to the patient reader. Both resonant and non-resonant production mechanisms are studied.

As discussed in Section 2.5, the standard motivation for this search – to measure the Higgs self-coupling, λ_{hhh} – is still a long ways off. Any signal observed in the present data would be a sign of new physics: a second Higgs doublet, additional electroweak singlets, gravitons, radions, stoponium, colored scalars, a composite Higgs boson, or something completely new.

The search is performed in the $\gamma\gamma b\bar{b}$ channel, which has long been noted as one of the most promising [153], thanks to the large $h \to b\bar{b}$ branching ratio, a clean diphoton trigger, excellent diphoton invariant mass resolution, and low backgrounds. This channel is particularly important in the search for resonances with mass m_X in the range $260 < m_X < 500$ GeV, where backgrounds and combinatorics make other channels such as $4b$ or $b\bar{b}\tau^+\tau^-$ challenging.

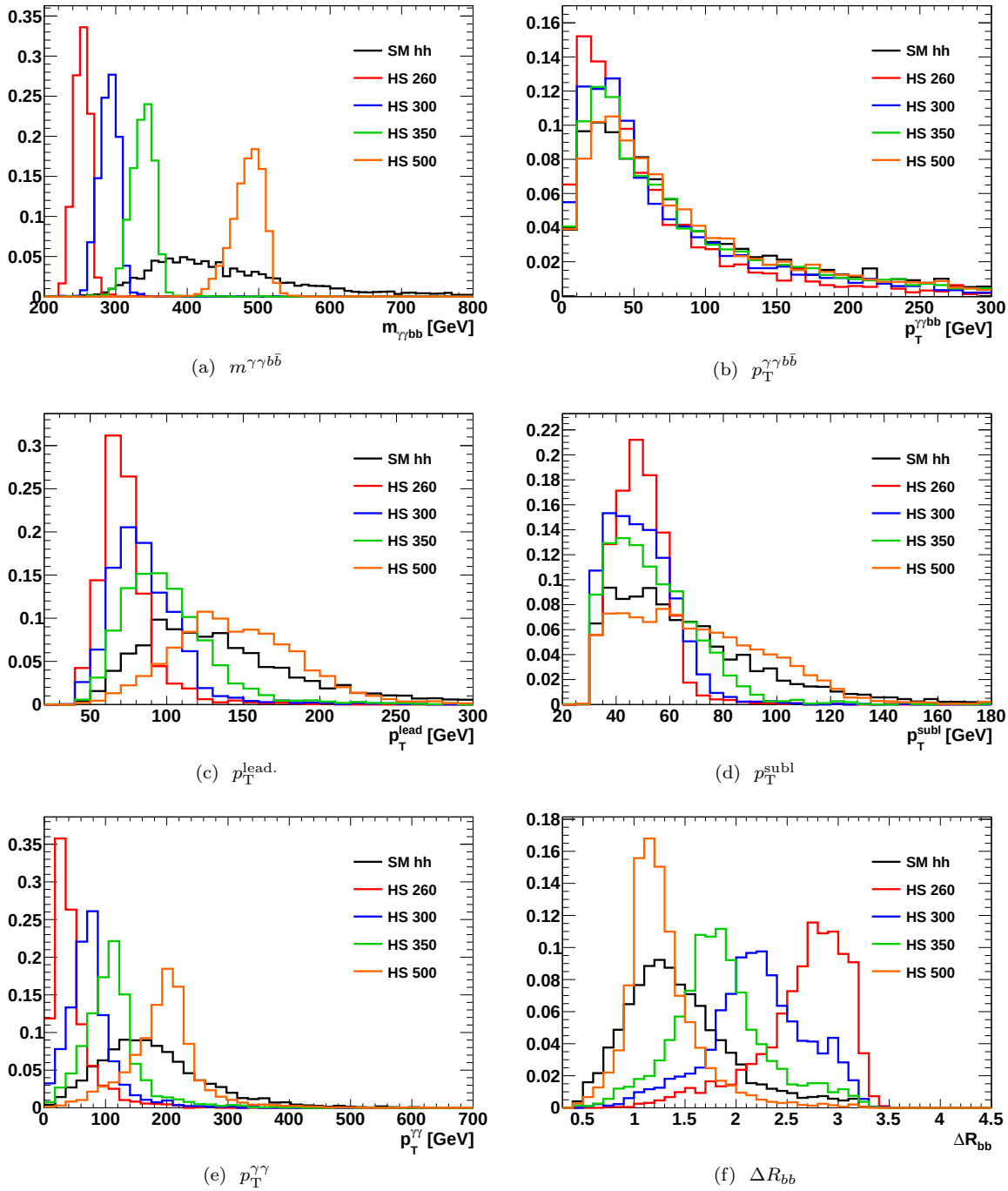
8.1 Simulated Samples

Simulation is used for processes that contain Higgs bosons: both pair production signal samples (presented here), and single Higgs boson production backgrounds (Section 5.2). Two benchmark signal models were defined. For non-resonant production, SM Higgs boson pair production was used, while the search for a resonance used a narrow-width gluon fusion-produced scalar. Figure 8.1 presents basic kinematics of the benchmark models described below.

Backgrounds that are not resonant in the diphoton mass are estimated using data-driven methods, but Monte Carlo is used as a cross-check to understand their composition.

8.1.1 SM Higgs Boson Pair Production

SM Higgs boson pair production through gluon fusion proceeds principally through two diagrams that interfere destructively: the 'box' diagram with two Higgs bosons radiated off of the top quark loop, and the 'self-coupling' diagram where the single Higgs boson splits (Figure 2.6). The Higgs pair production code from MadGraph 5 is used, with Higgs mass $m_h = 125$ GeV [154, 155]. The scale is left at the MadGraph default dynamic scale of $\sqrt{m_h^2 + p_{\mathrm{T}}^2}$, and CTEQ6L1 PDFs are used [111, 112, 156]. Events are showered, and the Higgs bosons are decayed using Pythia8. The decays are forced to proceed to pairs of b-quarks and photons, and a generator-level filter requires one of each decay. Good agreement was found when the



(a) $m^{\gamma\gamma b\bar{b}}$ (b) $p_{\mathrm{T}}^{\gamma\gamma b\bar{b}}$

(c) $p_{\mathrm{T}}^{\mathrm{lead.}}$ (d) $p_{\mathrm{T}}^{\mathrm{subl}}$

(e) $p_{\mathrm{T}}^{\gamma\gamma}$ (f) ΔR_{bb}

Figure 8.1: Basic kinematic properties are presented of the benchmark Higgs boson pair production models. Narrow resonances are shown with masses 260, 300, 350, and 500 GeV, along with SM production of two Higgs bosons (black). This difference between models is most clear in the four-object mass, $m^{\gamma\gamma b\bar{b}}$ (a). The distributions of the $\gamma\gamma$ and $b\bar{b}$ pairs are virtually identical. Comparisons to alternative models are included in Appendix J.

generated events were compared to theoretical predictions for the SM [24, 153]; a few examples of such comparisons are presented in Appendix L.

8.1.2 Narrow-Width, Gluon-Initiated Scalar

Two Higgs doublets are widespread in BSM theory. As described in Section 2.5.2.1, the four additional degrees of freedom result in four new Higgs bosons, H, A, and $H^{\pm}$. For $m_H > 2m_h$, the branching $H \to hh$ can be large or dominant – even for low $\tan\beta$ and $\cos(\beta-\alpha)$. The $H \to hh$ process is thus generic and it is strongly motivated. A gluon-fusion produced scalar resonance was therefore taken as the benchmark for the resonance analysis. In the region most interesting for 2HDMs of $2m_h < m_H < 2m_t$ and low $\cos(\beta-\alpha)$, the width of the resonance is within a few GeV – far less than the experimental resolution. The narrow width approximation (NWA) was therefore adopted, setting $\Gamma_X = 10$ MeV. Events were generated using the MadGraph 5 'Heavy Scalar' model [155]. As in the non-resonant case, the Higgs bosons were decayed and the events were showered using Pythia8.

Several alternative resonant models were considered, implemented, and compared. Pythia8 was used to generate the heavy scalar H with narrow width, through both gluon fusion and quark-initiated (predominantly VBF) production. An alternative MadGraph 2HDM implementation for a gluon-induced heavy scalar gave identical results to the one used. Finally, a Randall-Sundrum KK graviton was implemented through MadGraph [157–159].

The three gluon-induced models tested were extremely consistent. The quark-induced H production from Pythia8 was found to differ only in the transverse momenta of the individual Higgs bosons and of the four-object system, $p_{\mathrm{T}}^{\gamma\gamma b\bar{b}}$, as might be expected. The kinematic distributions of the individual photons and b-jets were found to be very similar, and the final cut efficiencies for fully simulated events were consistent within 10%. The spin-2 graviton model resulted in more-central, higher-p_{T} Higgs bosons. These features led to a higher selection efficiency for fixed p_{T} cuts. The finite width of the graviton ($\gtrsim 8$ GeV) led to substantial deformations of the $m^{\gamma\gamma b\bar{b}}$ lineshape near the kinematic turn-on at $2m_h$. Several examples of these effects are displayed for generated events smeared for detector resolution, in Appendix J.

Based on these comparisons of production mode and object properties, the results presented are considered reasonably model-independent. The caveat is that great care must be taken in interpreting models of resonances whose widths exceed a meaningful fraction of the experimental resolution.

8.1.3 Samples for Background Studies

The background composition is studied using events generated with leading order MadGraph 5. Six samples were generated, allowing for either one or two light flavor jets (j) to fake flavor-tagging or photon-identification: $bb\gamma\gamma$, $bbj\gamma$, $bbjj$, $bj\gamma\gamma$, $bjj\gamma$, and $jj\gamma\gamma$. Because of the very-low rates at which these fakes occur, and given that this simulation is only used as a cross-check and for a single uncertainty, it is prohibitively expensive and unnecessary to fully simulate these processes. The events are therefore generated at 'truth-level,' and detector resolutions and efficiencies are parameterized ('smeared'). The parameterizations are derived from fully simulated MC as a function of p_{T} and $|\eta|$. The flavor-tagging efficiencies also depend on the flavor of the true parton (light, charm, bottom, or tau). The k-factors for these processes are extremely uncertain. Available measurements of W, Z, and $t\bar{t}$ with heavy flavor all suggest a value of around 1.5 [160–163], but the agreement to data is better with $k \approx 2$. (Note that the comparison to data also incorporates any mismodelling in the naïve parameterizations.)

The $t\bar{t}$ process is studied slightly differently: fully simulated samples are used, and an electron to photon 'fake rate' measured in data is used to weight the events.

	No Corr.	p_T Corr.	**Muon Corr.**	Both
SM hh	15.2 GeV	15.1 GeV	13.4 GeV	13.2 GeV
$m_H = 260$ GeV	12.5 GeV	13.0 GeV	11.9 GeV	12.2 GeV
$m_H = 300$ GeV	13.2 GeV	13.6 GeV	12.8 GeV	13.1 GeV
$m_H = 350$ GeV	14.6 GeV	14.5 GeV	13.3 GeV	13.3 GeV
$m_H = 500$ GeV	15.6 GeV	15.5 GeV	13.8 GeV	13.7 GeV

Table 8.1: The resolution is presented for Bukin functions, fitted to the $m^{b\bar{b}}$ spectrum of the five signal models. Both jets were tagged, and their p_{T}s were required to exceed 55 GeV and 35 GeV. The functions were fit in the range $[80-150$ GeV$]$. The optimal performance across models comes by adding in the muons, but dropping the p_{T}-reco correction.

8.2 Event Selection and its Optimization

Wherever possible, the event selection follows earlier analyses in the diphoton channel. The photon selection in particular, is practically identical to the differential cross sections analysis (the only difference is an update to the photon calibration). The analysis differs only in the selection of two b-tagged jets, and cuts on the di-jet and four-body invariant masses, $m^{b\bar{b}}$ and $m^{\gamma\gamma b\bar{b}}$.

There are three types of optimizations to be made: (1) improving the reconstruction and resolution of the b-jet p_{T}s and the di-jet mass, (2) setting appropriate kinematic and flavor-tagging cuts, and (3) better constraining $m^{\gamma\gamma b\bar{b}}$.

8.2.1 Optimization

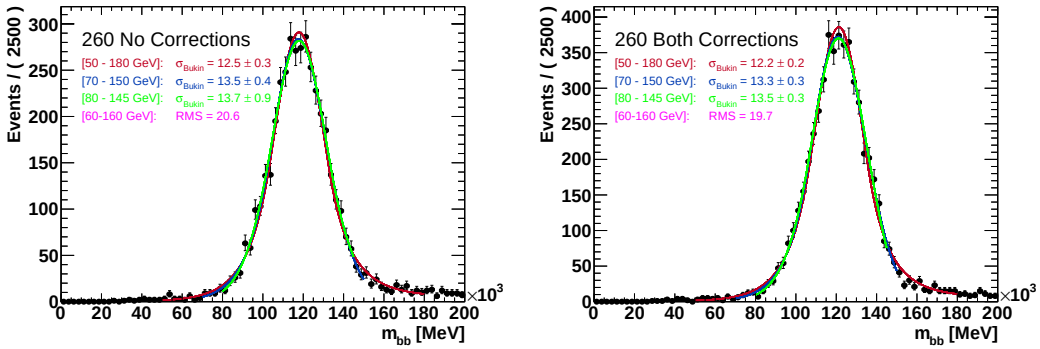
8.2.1.1 Jet Momentum Corrections

The $h \to b\bar{b}$ decay has extremely broad mass resolution, because the underlying b-jets are not precisely measured: the resolution is already quite wide for generic jets, and energy lost to neutrinos in semi-leptonic b decays only makes things worse. Following on the heels of the ATLAS $h \to b\bar{b}$ analyses [164], two improvements were proposed to the default jet reconstruction. First, if a muon is found within within $\Delta R < 0.4$ of a tagged jet, its four-vector is added to the jet's. Second, a 'p_{T}-reco.' correction was defined as the ratio of the truth-level parton p_{T} to the reconstructed p_{T}, $p_{\mathrm{T}}^{\mathrm{truth}}/p_{\mathrm{T}}^{\mathrm{reco.}}$. This ratio was extracted as a function of jet p_{T} and η from Monte Carlo. The correction is process-dependent and it was originally derived by the ATLAS $h \to b\bar{b}$ working group for Wh and Zh decays. However, to avoid re-defining the jet p_{T} for every tested hypothesis (non-resonant and resonant models, and variable resonance mass), the Wh/Zh corrections were attempted 'out-of-the-box.'

The performance of each of these methods was evaluated for each signal Monte Carlo by measuring the width of a Bukin function, fitted in various ranges, for the four permutations of corrections (both, either, neither). In performing these fits, both jets were required to be tagged at the 70% operating point, and the p_{T} of the leading (subleading) jet was required to exceed 55 (35) GeV. An example of these fits is presented in Figure 8.2. The resultant widths are tabulated in Table 8.1, where it is evident that the 'p_{T}-reco' corrections do not (universally) improve the resolution. The 'failure' of the p_{T}-reco. correction in this context is understood to follow simply from the different samples used to derive and apply the correction. Based on these results, only the muon correction is applied in the analysis.

8.2.1.2 Cuts Optimization

Three cuts must be optimized for the analysis: the p_{T} cuts on the leading and subleading jets, the mass cuts on the di-jet system, and the efficiency/working-point of the flavor-tagging algorithm.



(a) No Corrections (b) Both Corrections

(c) p_{T}-Reco. Correction (d) Muon Correction

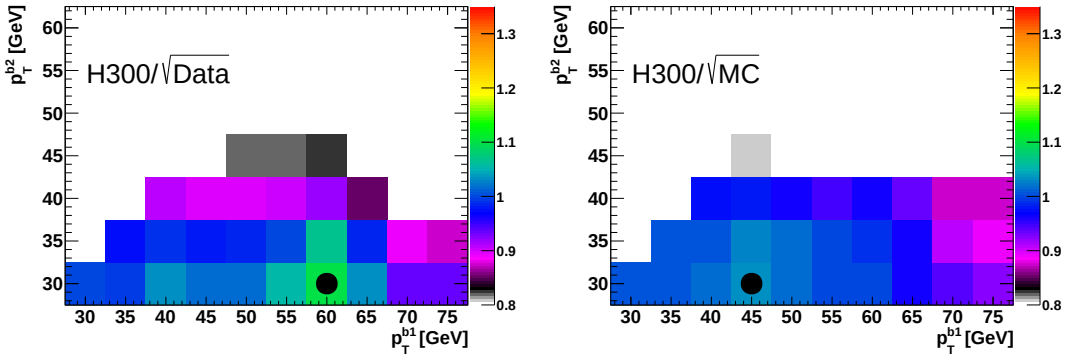
Figure 8.2: The bb mass spectrum is shown for the 260 GeV 2HDM benchmark, after various corrections to the energies of the b-jets. A Bukin polynomial is fitted to the core of these spectra, to parameterize their width.

Cuts were optimized for discovery potential, based on the $S/\sqrt{B}$ figure of merit. In the interest of simplicity of the analysis, a single set of cuts was defined for all signal models: non-resonant SM hh, and all resonance mass points.[9]

Transverse momentum of the b-jets. The optimal p_{T} cuts obviously vary with the resonance mass. The mass $m_X = 300$ GeV lies at the center of the interesting range of $2m_h - 2m_t$, and it was used as the benchmark. Because of very limited statistics in data, the impact on the background was estimated with both un-tagged events in data, and with fully simulated (MadGraph) MC $\gamma\gamma b\bar{b}$ production. A broad mass cut of $80 < m_{jj} < 140$ GeV is placed on the jet mass, since the mass and p_{T} spectra are correlated. The optimization with $m_X = 300$ GeV is shown in Figure 8.3, and $m_X = 260$ and 350 GeV, as well as SM hh are shown in Appendix K. An earlier optimization in which the p_{T}-reco. correction was still used resulted in the optimal point $p_{\mathrm{T}}^{b1} > 55$ GeV and $p_{\mathrm{T}}^{b2} > 35$ GeV. The optimal point with only the muon corrections is slightly lower ($p_{\mathrm{T}}^{b2} > 30$ GeV), but the cut value was not changed, since it was considered 'close enough,' and the analysis was already very advanced.

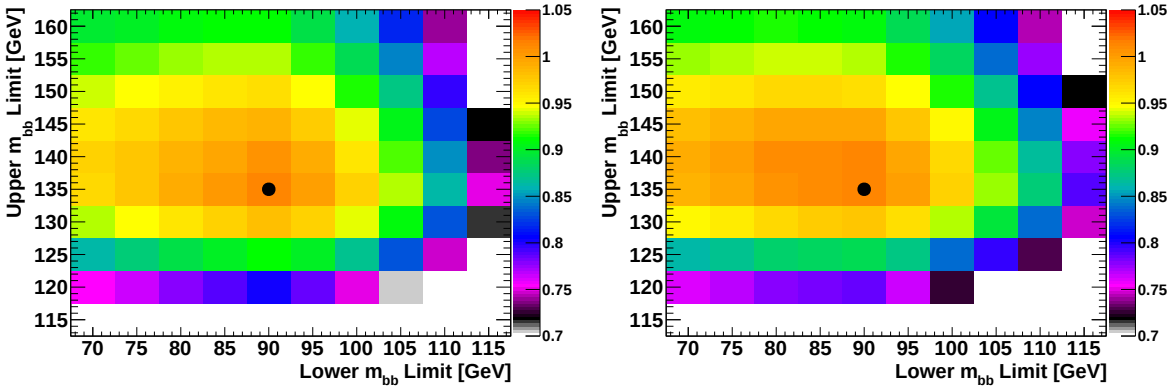
Lower and upper cuts on the di-jet mass, $m^{b\bar{b}}$. The optimization of the di-jet mass cut was performed analogously to the p_{T} optimization above. The $S/\sqrt{B}$ metric was again used, and un-tagged data was used as a stand-in for the background. In this case all models were very-consistent in the optimal points, as seen

[9] Alternative analysis strategies could allow for $m^{\gamma\gamma b\bar{b}}$-dependent p_{T} cuts. In the future, lower cuts may improve limits.



(a) Data (b) Monte Carlo

Figure 8.3: The optimization of the b p_T cuts is presented for the mass point $m_X = 300$ GeV. The background response is estimated either from un-tagged (but kinematically identical) jets in data, or from $\gamma\gamma b\bar{b}$ MC. The figure of merit $S/\sqrt{B}$ is shown as a function of the leading and subleading b p_T cut. Each plot is scaled so that the value at 55/35 GeV (the cut used) is at 1.0. The black dot in each plot shows the bin with the highest significance. Alternative signal models are included in Appendix K.



(a) 2HDM Signal: $m_H = 300$ GeV (b) SM hh Signal

Figure 8.4: The lower and upper cuts on $m^{b\bar{b}}$ are optimized, using $S/\sqrt{B}$ as a figure of merit and estimating 'B' as un-tagged events from data. The optimal cuts of $90 < m^{b\bar{b}} < 135$ GeV is slightly modified, raising the lower cut to 95 GeV to minimize contributions from $(h \to \gamma\gamma)\,(Z \to b\bar{b})$.

in Figure 8.4. The optimal cuts were found to be $90 < m^{b\bar{b}} < 135$ GeV. However, because this optimization did not take $(h \to \gamma\gamma)\,(Z \to b\bar{b})$ associated production into account, the lower cut was shifted 'by hand' to 95 GeV to reduce this contribution. (This also helps reduce unwanted $A \to Zh$ contamination in 2HDM interpretations.)

Flavor tagging working point. The selection criteria for identification of the two b-jets was optimized very early in the analysis using, very simply, the target efficiencies of the working points (60%, 70%, 80%) as the signal efficiency of each jet, and data sidebands as the background. The outcome was that the 70%-efficient working point was used for both b-jets, as shown in Table 8.2.

Tag Working Point	70+80	70+70	70+60	80+60	60+60
Data Sidebands	43	18	16	35	12
$(S/\sqrt{B})$/norm	0.74	1.00	0.91	0.70	0.90

Table 8.2: This optimization was performed before the p_{T} cuts in the analysis were quite as tight. The metric, as usual, is $S/\sqrt{B}$, using the efficiency target of the b-tagging working point for each b-jet to scale the signal yield (60%, 70%, or 80%) and the data sidebands for the background. The figure of merit is scaled so that the nominal choice is at 1.00

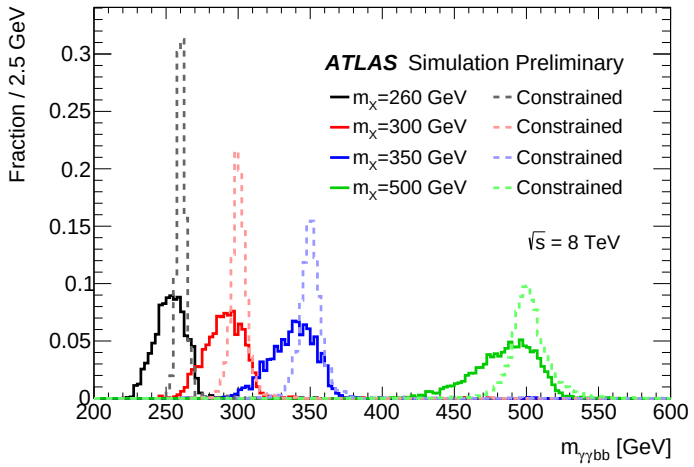


Figure 8.5: Plot of the $m^{\gamma\gamma b\bar{b}}$ distribution both before (solid) and after (dashed) a simple constraint of $m^{b\bar{b}}$ to m_h. For the latter, four-vector of the $m^{b\bar{b}}$ system is scaled by m_h over its mass, resulting in a dramatic improvement the four-body mass resolution.

8.2.1.3 Mass Constraint

Despite the efforts made to improve the b-jet momentum, by adding in the muons, the resolution of the dijet mass $m^{b\bar{b}}$ remains very large. This in turn impacts the four-body mass, $m^{\gamma\gamma b\bar{b}}$, which is used to select signal windows, in the search for a resonance (Section 8.4, below). A very simple rescaling is therefore applied to the four-vector of the bb system, multiplying it by $m_h/m^{b\bar{b}}$, according to hypothesis that the selected bb system comes from a Higgs boson decay. This dramatically improves in the $m^{\gamma\gamma b\bar{b}}$ resolution, as shown in Figure 8.5.

Additional studies showed that the impact on the background shape in $m^{\gamma\gamma b\bar{b}}$ was minimal. In future iterations of this analysis, more-complicated scalings could allow the two b-jets to vary differently, according to their resolutions.

8.2.2 Event Selection

As a result of the optimization procedures described above, the additional selection cuts (beyond those used for the standard $h \to \gamma\gamma$ analysis) are:

- ▷ The two highest-p_{T} jets with $|\eta| < 2.5$ that satisfy the 70% efficient working point of the ATLAS multi-variate tagging procedure (MV1) are selected. This procedure combines information about the

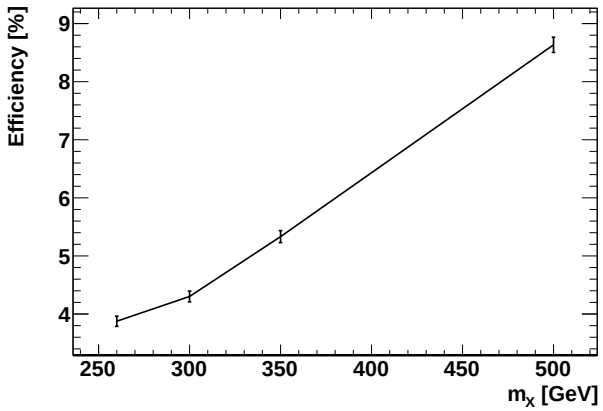


Figure 8.6: Selection efficiency for the benchmarks of the resonance analysis, from Table 8.3. This efficiency is interpolated linearly, for any un-simulated mass hypothesis.

	SM	Resonant NWA				
	Benchmark	260 GeV	300 GeV	350 GeV	500 GeV	1000 GeV
Generated	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Trigger	73.1%	72.5%	71.6%	71.8%	73.6%	81.0%
Preselection	57.3%	56.7%	56.1%	56.2%	57.7%	65.1%
Photon p_T	51.6%	51.6%	49.8%	49.2%	52.5%	62.4%
Photon Identification	45.3%	44.2%	42.8%	42.6%	46.4%	56.2%
Isolation	39.1%	33.1%	33.8%	35.9%	40.6%	47.4%
$105 < m_{\gamma\gamma} < 160$ GeV	39.0%	33.0%	33.8%	35.9%	40.5%	47.4%
2 Central Jets	33.9%	25.5%	26.9%	29.8%	36.2%	45.1%
Tagging	12.5%	8.4%	8.9%	10.0%	14.1%	19.1%
b p_T Cuts	10.1%	4.8%	5.6%	7.2%	12.0%	18.1%
$95 < m_{b\bar{b}} < 135$ GeV	7.4%	4.0%	4.3%	5.3%	8.6%	14.2%

Table 8.3: Cumulative cut efficiency/acceptance, for signal models.

impact parameter (significance) of tracks within a candidate jet, with information about explicitly reconstructed vertices of b-hadrons. [165–167]

- ▷ Because the b-tagging systematics are defined for a JVF cut of 0.5 instead of 0.25, the former value is adopted for this analysis.
- ▷ The p_{T} of the more (less) energetic jet must exceed 55 (35) GeV.
- ▷ The combined mass of the dijet system must fall between 95 and 135 GeV.

Table 8.3 presents the cumulative efficiency of the benchmark signal simulation. The total efficiency is of the order 4-10%. When testing resonance mass hypotheses that have not been fully simulated, a linear extrapolation between simulated points is used, as shown in Figure 8.6.

Both the resonance and non-resonance analyses also use events with identical kinematics requirements but without b-tagging, as 'control regions' to help constrain the shapes or extrapolations used in $m^{\gamma\gamma}$ and $m^{\gamma\gamma b\bar{b}}$. These will be discussed in Section 8.4.

Process	$\gamma\gamma b\bar{b}$	$\gamma\gamma bj$	$\gamma\gamma jj$	γbbj	γbjj	$bbjj$	$t\bar{t}$	Total
Cross Section [fb]	8.3	79.8	1.51×10^3	3.96×10^3	1.83×10^5	8.66×10^6	113	-
k-factor	2	2	2	2	2	2	1.22	-
Events Expected	0.3	0.3	0.4	0.5	0.2	0	0.3	2.1

Table 8.4: The estimated contributions from various backgrounds are estimated from smeared Monte Carlo. The $t\bar{t}$ background alone is estimated with fully simulated events, but using a paramaterized electron to photon fake rate measured in data. The expected number of events is shown for 20.3/fb of data, with $|m^{\gamma\gamma} - m_h| < 5$ GeV after the full kinematic selection.

8.3 Background Studies

8.3.1 Non-Resonant Backgrounds

This analysis uses fits to data, almost exclusively, to estimate the backgrounds from processes that do not include Higgs bosons. This 'sideband' estimate is presented, along with a study of the background composition using Monte Carlo. The MC study is not used in the analysis *per se*, but is presented for information.

8.3.1.1 Sideband Fit in Data

An exponential fit is performed to the $|m_h - m^{\gamma\gamma}| > 5$ GeV sidebands of the signal region, after all cuts defined in Section 8.2.2. Because the event count is so low (just nine events), the shape is constrained using control regions: either inverting the photon identification requirement (requiring at least one photon to fail), or selecting either 1 or 0 tagged jets. Figure 8.7 shows that the three control regions give completely consistent results of 2.2 ± 0.7 events within $|m_h - m^{\gamma\gamma}| < 5$ GeV. The error is completely dominated by the statistical error on the nine events. Two alternative background shapes (a straight line and a second-order exponential) are also tested, and presented in Appendix L. Using a flat extrapolation from the sideband into the 10 GeV window gives a prediction of $9 \times (10\text{ GeV}/45\text{ GeV}) = 2$ events, which is also very consistent.

In the resonance analysis, a smaller window is used of $\pm 2\sigma_{\gamma\gamma} \approx 6.4$ GeV, leading to a smaller total expectation.

8.3.1.2 Monte Carlo: Composition

The 'smeared' samples described in Section 8.1.3 are used to evaluate contributions from non-resonant processes with 0, 1, or 2 fakes.

Two comparisons are made to data. First, in order to gain in statistics, events with just one b-tag are used to check the kinematic distributions and the overall performance of the procedure. Examples of these comparisons are included in Figure J.4 of Appendix J, using a k-factor of 2 which is found to improve agreement with data. The kinematic distributions are reasonably well described, as shown by the Kolmogorov-Smirnov tests performed on those distributions.

The second comparison is the overall rate expected of non-resonant backgrounds in the signal region. This gives the background composition in the signal region. The cross sections and expected number events after all cuts with $|m^{\gamma\gamma} - m_h| < 5$ GeV are presented in Table 8.4. The total yield of 2.1 events agrees well with the prediction from the sideband fits in $m^{\gamma\gamma}$ of 2.2 ± 0.7 events. No single process is found to dominate. Events contribute from the fully irreducible background ($\gamma\gamma bb$) and with one or two fakes: electrons faking photons and light flavor jets faking flavor tagging, or faking photon ID.

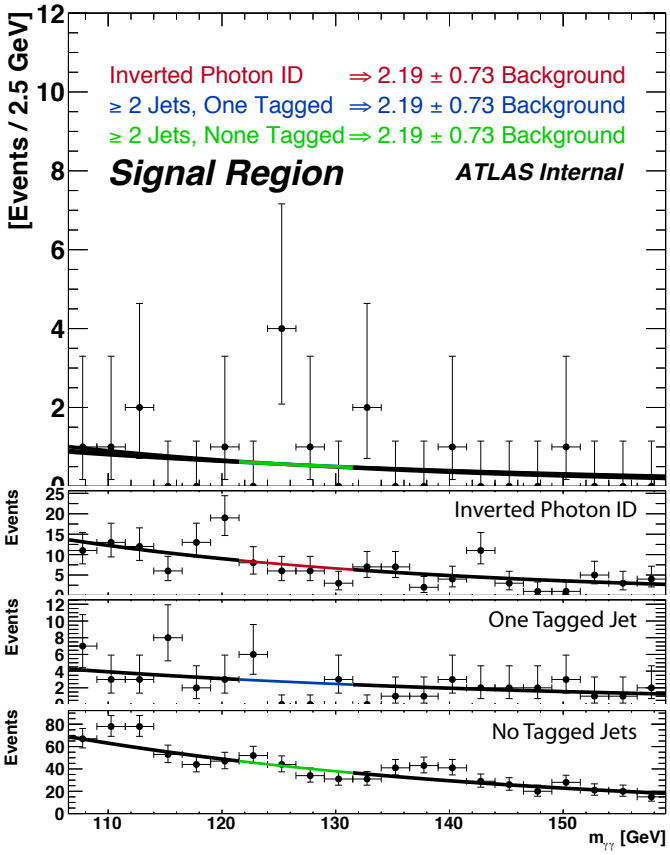


Figure 8.7: Exponential fits are performed to the $m^{\gamma\gamma}$ sidebands ($|m^{\gamma\gamma} - m_h| > 5$ GeV). The shape is obtained from several different sidebands (inverted photon identification, and 0 or 1 b-tags), while the normalization comes from the $m^{\gamma\gamma}$ sideband after all cuts. This demonstrates the insensitivity of the prediction to the control region. The background yield in the (initially, and for purposes of the fit, blinded) 10 GeV signal region is shown in the plot, and may be compared to the yields from smeared simulation. In the resonance analysis, a smaller window of $\pm 2\sigma_{\gamma\gamma}$ (6.4 GeV) is used.

8.3.2 Backgrounds from Single Higgs Boson Production

Because they are resonant in $m^{\gamma\gamma}$, the single Higgs boson backgrounds are not included in the fits presented above. This contribution is instead estimated from MC of the five production modes ggh, VBF, Wh, Zh, and $t\bar{t}h$. A total of 0.17 events are expected. The breakdown among production modes is presented in Table 8.5.

Although the production cross section for $b\bar{b}h$ is nearly three times that of $t\bar{t}h$, and the production does include real b-jets, those jets tend to be very soft. This is self-evident in the '5 flavor scheme' where the b-quarks are included in the PDF instead of arising from a gluon splitting (Figure 2.3). In that case, the two b-quarks annihilate to form the Higgs boson and there are no b-jets in the final state (as in the $t\bar{t}h$ diagram). Using theoretical predictions available at $\sqrt{s} = 2$, 7, and 14 TeV, the contribution from $b\bar{b}h$ was estimated as 0.003-0.007 events, and considered negligible [45–47, 168]. Very recently, MC methods for bbh have been released, giving a prediction of approximately 0.004 events that agrees well with this early estimate.

Process	Fraction of total
ggh	11%
VBF	2%
Wh	1%
Zh	17%
$t\bar{t}h$	69%
Total	0.17 ± 0.04 Events

Table 8.5: Predicted number and composition of SM single Higgs boson background events in the non-resonance search. The total expected SM signal from pair production of Higgs bosons is 0.04 events.

8.4 Analysis Strategy

Different strategies are employed for the non-resonance search and the resonance search. The non-resonance search uses a simultaneous signal plus background fit of the diphoton invariant mass, reminiscent of earlier $h \to \gamma\gamma$ analyses. The resonant search hinges on cuts on $m^{\gamma\gamma b\bar{b}}$ that make such fits impossible: the event count is simply too low. The search instead proceeds as a counting experiment.

8.4.1 Non-Resonant Production

An unbinned, simultaneous signal plus background fit of $m^{\gamma\gamma}$ is performed to all events passing the dijet and diphoton selections described above. This fit has three components: the signal with two Higgs bosons, the SM single Higgs boson background that is resonant in $m^{\gamma\gamma}$, and the continuum background that falls with $m^{\gamma\gamma}$.

The fit is performed simultaneously in two categories. The first category is the signal region, in which at least two jets are b-tagged. The second is a control region, containing events with fewer than two b-tags. The two classes of events are kinematically identical: in the signal region, the mass requirement and p_{T} cuts defined above must be satisfied by the two leading tagged jets, whereas in the control region, they are met by the two leading jets.

As described in Section 5.3.1, the shape of the $m^{\gamma\gamma}$ resonance is described by the sum of a Crystal Ball function and a wide Gaussian component that models the tails of the distribution. However, in this case, the signal model comes simply from a fit at $m_h = 125$ GeV, instead of simultaneously across all m_h hypotheses. A simple exponential function describes the continuum backgrounds that fall with $m^{\gamma\gamma}$. The slope of the exponential is shared in the fit between the two categories so that the control region constrains the background shape in the signal region. Figure 8.8 shows the separate diphoton mass distributions for events with ≥ 2 b-tags and events with ≤ 1 b-tags. The combined acceptance and selection efficiency for SM pair production of Higgs bosons is 7.4%.

8.4.2 Resonant Production

The search for resonant pair production begins with the same signal region of two b-tagged jets as the non-resonance search. It proceeds as a counting analysis with cuts and extrapolation factors in $m^{\gamma\gamma}$ and (constrained) $m^{\gamma\gamma b\bar{b}}$.

First, the diphoton mass is required to be within $\pm 2\sigma_{m^{\gamma\gamma}}$ of the Higgs boson mass, $m_h = 125.5$ GeV [13]. The acceptance of this requirement on background events without Higgs bosons, $\varepsilon_{m^{\gamma\gamma}}$, is measured by fitting an exponential function to the $m^{\gamma\gamma}$ sidebands for events with fewer than 2 b-tagged jets, as discussed in Section 8.3.1 and shown in Figure 8.7. For N observed events with two b-tags in the sideband ($|m^{\gamma\gamma} - m_h| > 2\sigma_{m^{\gamma\gamma}}$), the number of expected non-Higgs boson background events ($N_{m^{\gamma\gamma}}$) within $2\sigma_{m^{\gamma\gamma}}$ around m_h is

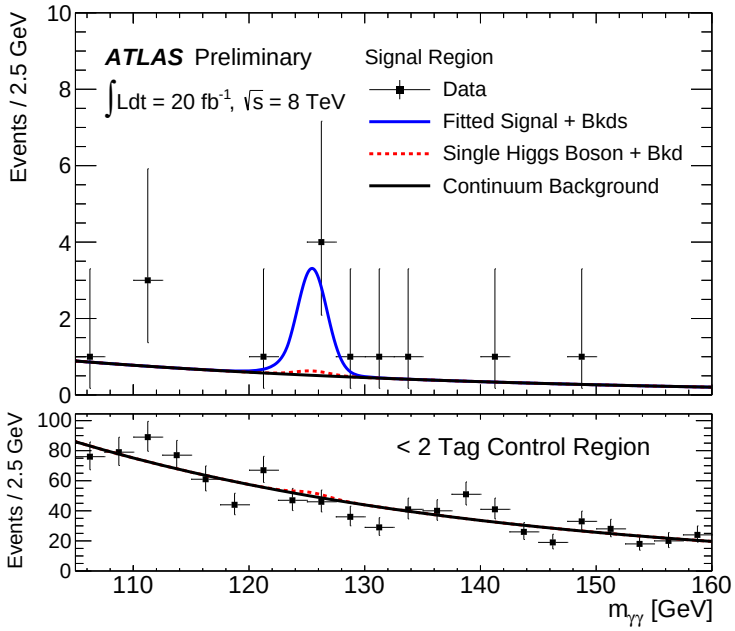


Figure 8.8: The diphoton invariant mass spectrum is shown for the signal region of the non-resonance search. The slope of the exponential function in the signal region is constrained using a simultaneous fit to the control region with fewer than two b-tags (below).

given by:

$$N_{m^{\gamma\gamma}} = N \frac{\varepsilon_{m^{\gamma\gamma}}}{1 - \varepsilon_{m^{\gamma\gamma}}}, \tag{8.1}$$

where the denominator compensates for the fact that $\varepsilon_{m^{\gamma\gamma}} = 0.13$ is derived relative to the full $m^{\gamma\gamma}$ spectrum while N contains only those events in the sidebands.

The second extrapolation is on the four-object mass $m^{\gamma\gamma b\bar{b}}$, with the constraint on the $b\bar{b}$ system discussed in Section 8.2.1.3. The smallest window containing 95% of the signal events in the narrow-width simulation is selected. The requirements vary linearly with the mass of the resonance considered, as shown in Figure 8.9. The width of the signal window varies from 17 GeV at $m_X = 260$ GeV to 60 GeV at $m_X = 500$ GeV.

The acceptance for the background to pass this cut, $\varepsilon_{m^{\gamma\gamma b\bar{b}}}$, also varies with m_X. It is measured using events in data with $|m^{\gamma\gamma} - m_h| < 2\sigma_{m^{\gamma\gamma}}$ and fewer than 2 b-tags. Studies in both data sidebands and simulation show that the shapes of $m^{\gamma\gamma b\bar{b}}$ and $m^{\gamma\gamma jj}$ agree within statistical uncertainties. Figure 8.10a shows comparisons made in the diphoton sidebands ($|m^{\gamma\gamma} - m_h| > 5$ GeV) for 0, 1, and 2 b-tags in data. Figure 8.10b shows the diphoton signal region from smeared Monte Carlo, varying the individual components contributing to $m^{\gamma\gamma b\bar{b}}$ up and down by a factor of two, as well as the expected $m^{\gamma\gamma jj}$ shape. The difference between the $m^{\gamma\gamma b\bar{b}}$ and $m^{\gamma\gamma jj}$ shapes is assigned as (a part of) the systematic on the method (Section 8.5). (This systematic is the sole point in the analysis where the smeared MC backgrounds are used.) The distribution of $m^{\gamma\gamma jj}$ in data is fitted with a Landau function, which is integrated in the signal window to obtain $\varepsilon_{m^{\gamma\gamma b\bar{b}}}$. The bottom panel of Figure 8.11 shows this fit. The value of $\varepsilon_{m^{\gamma\gamma b\bar{b}}}$ is small ($< 8\%$) at low and high m_X, and peaks at 18% for $m_X = 300$ GeV. The combined acceptance and selection efficiency for a resonance signal to pass all cuts varies from 3.8% at $m_X = 260$ GeV to 8.2% at $m_X = 500$ GeV.

The total background from sources without Higgs boson decays in the resonance analysis N_B is given by:

$$N_B = N \frac{\varepsilon_{m^{\gamma\gamma}}}{1 - \varepsilon_{m^{\gamma\gamma}}} \varepsilon_{m^{\gamma\gamma b\bar{b}}}, \tag{8.2}$$

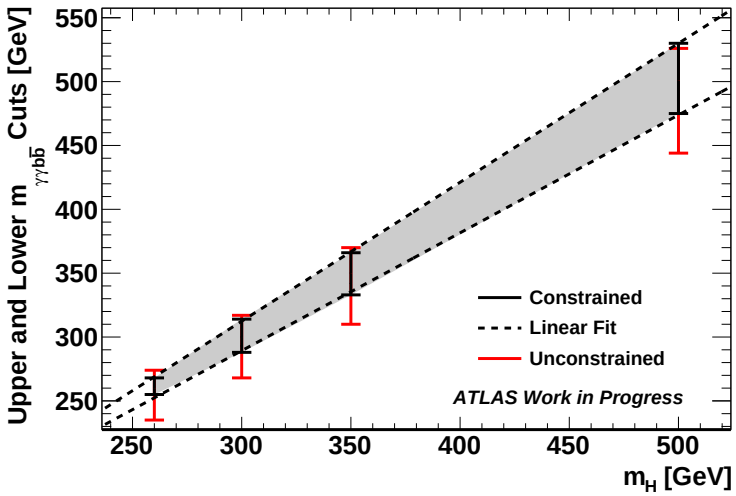
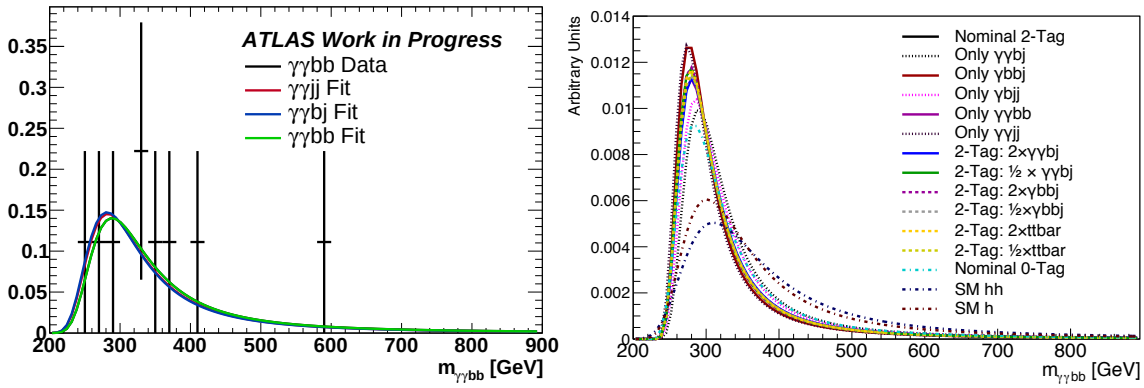


Figure 8.9: The cuts on the $m^{b\bar{b}}$-constrained values of $m^{\gamma\gamma b\bar{b}}$ made in this analysis, for the four simulated mass points. The cuts are defined to select the smallest window containing 95% of the events passing all other cuts, for each mass hypothesis. The red points show the values that would have been used, without the mass constraint. The linear extrapolation of the cuts is shown by the dashed line.



(a) Data in Diphoton Sidebands

(b) MC with Varied Composition

Figure 8.10: (a) The $m^{\gamma\gamma b\bar{b}}$ distribution is shown for data in the $m^{\gamma\gamma}$ sidebands of the signal region, and overlaid with the Landau fit. The Landau fits from samples with 1 and 2 b-tags are also shown, and the agreement in the shapes is very close between the three regions. (b) In smeared MC, the dominant background components for the 2-tag region are varied by factors of 2 and $\frac{1}{2}$, and compared as well to the control region with 0 tags. The difference between the $\gamma\gamma b\bar{b}$ and $\gamma\gamma jj$ regions in MC is one component of the uncertainty assessed on the Landau fits.

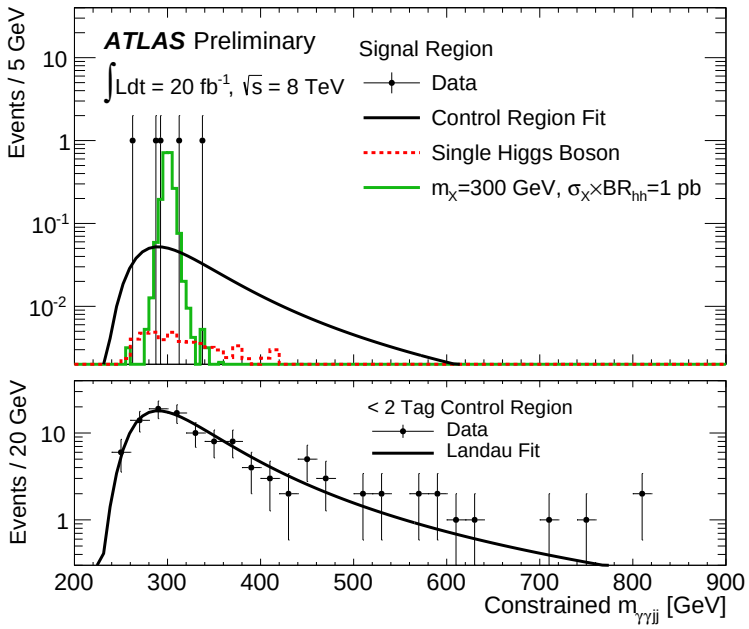


Figure 8.11: (Upper plot) The constrained four-object invariant mass, $m^{\gamma\gamma jj}$, is shown for data events satisfying all cuts and falling within the $m^{\gamma\gamma}$ signal region. The expected backgrounds from non-resonant processes and single Higgs boson production at the SM rate are also shown. (Lower plot) The shape of the continuum contribution is obtained from a fit to events with fewer than two b-tags.

where N_B and $\varepsilon_{m^{\gamma\gamma b\bar{b}}}$ are functions of m_X. Migration of events into and out of the $m^{\gamma\gamma}$ and $m^{\gamma\gamma b\bar{b}}$ signal regions are covered by uncertainties described in the following section.

Because they are not accounted for by the above $m^{\gamma\gamma}$ sideband techniques, contributions from Higgs bosons produced in association with jets (particularly with $c\bar{c}$ or $b\bar{b}$ pairs) are estimated using simulation. In the resonance search, the small rate of SM pair production is included as a background. SM cross sections and branching fractions are assumed in all cases.

8.5 Systematic Uncertainties

Particularly in the resonance analysis, systematic uncertainties are small compared with the Poisson fluctuations. For the most part, the systematic uncertainties (yet again) follow the ones used in previous analyses. In some cases the implementation of these uncertainties changes. For instance, uncertainties on the $m^{\gamma\gamma}$ resolution for the non-resonance search become uncertainties on migrations between categories in the resonance analysis. There are also completely new uncertainties, for b-jets and modelling the $m^{\gamma\gamma b\bar{b}}$ shape. All uncertainties are presented in Table 8.6.

Theoretical uncertainties on single and double Higgs boson production, from PDF and scale variations have been presented in Table 2.1 [23,24]. Following previous analyses of $t\bar{t}h$ [136] an additional uncertainty is applied for the SM production of single Higgs bosons with heavy flavor: the VBF and Wh processes, which proceed through qq diagrams, are assigned 150% uncertainties based on $W+b$ studies [169]; the ggh process, which proceeds through gg loops, is assigned a 200% uncertainty, motivated by $t\bar{t}+$ HF studies [170]. No uncertainty is applied to Zh or $t\bar{t}h$, since these processes naturally include real b-jets.

Because of the relative $p_{\mathrm{T}}/m^{\gamma\gamma}$ cuts, photon energy scale uncertainties are negligible: varying the photon p_{T}s affects the mass in the same way. The uncertainty of 13% on the diphoton mass resolution is evaluated by

varying the photon energy resolution within its uncertainties. In the non-resonance search it is implemented as an uncertainty on the width of the signal model. In the resonance analysis it is transformed into a 1.6% uncertainty on migrations into and out of the signal region. This represents the fraction of events where an upward variation of the photon resolution causes the diphoton mass to leave the $m_h \pm 2\sigma_{m^{\gamma\gamma}}$ window required for the signal region. The uncertainty on the Higgs boson mass measured by ATLAS impacts the peak position in $m^{\gamma\gamma}$ in the signal plus background fit of the non-resonance analysis, and is transformed into a 1.7% migration uncertainty in the resonance search. The uncertainty for the acceptance of the $m^{\gamma\gamma}$ cuts on non-Higgs boson backgrounds is estimated by comparing fits of $m^{\gamma\gamma}$ to data in control regions with reversed photon identification or b-tagging requirements, and using different functional forms; these additional fits are included in Figure J.4 of Appendix J. A uniform distribution is also compared. The largest deviation observed from these fits (11%) is used for all searches.

Three components contribute to the uncertainty on $\epsilon_{m^{\gamma\gamma b\bar{b}}}$, which are combined in quadrature:

(1) Limited statistics in the control region with fewer than two b-tags used for the Landau fit lead to a relative uncertainty between 3-18% that varies as a function of m_X.

(2) The $m^{\gamma\gamma b\bar{b}}$ shape for untagged jets might not exactly mirror the one for tagged jets. The signal region and control samples are compared in simulation and the relative difference on $\epsilon_{m^{\gamma\gamma b\bar{b}}}$ is taken as the uncertainty. This value varies with m_h and is always less than 30%. Cross-checks in data show that the $m^{\gamma\gamma b\bar{b}}$ and $m^{\gamma\gamma jj}$ shapes are fully consistent within the statistics available.

(3) Finally, an uncertainty of 16-30%, depending on m_X, is included to cover the fit function choice. This was evaluated via comparisons of Landau shapes to alternate functions in simulation, including Landau shapes where the width varies with $m^{\gamma\gamma b\bar{b}}$, as well as Crystal Ball functions.

Potential contamination from single Higgs boson processes in the control region is estimated to be less than 4% and is subtracted with negligible impact on the shape.

Uncertainties due to the b-tagging calibration are typically $2-4$% for both single Higgs boson and signal processes. Uncertainties due to the jet energy scale are 7% (22%) for single Higgs boson backgrounds in the non-resonance (resonant) analysis, and 1.4% (4.4%) for signal processes. Uncertainties due to the resolution on jets are 4.8% (21%) for single Higgs boson backgrounds, and 6.3% (9.3%) for signal processes. The uncertainty on the integrated luminosity is 2.8% [128].

8.6 Results and Interpretations

The combined signal plus background fit for the non-resonance analysis is shown in Figure 8.8. Within a $\pm 2\sigma_{m^{\gamma\gamma}}$ window around the Higgs boson mass, 1.5 events are expected, with 1.3 from the continuum background and 0.2 from single Higgs boson production which is dominated by $t\bar{t}h$ events. At the SM rate, 0.04 events are expected with pair-produced Higgs bosons. Five events are observed; using the full fit, this is 2.4 standard deviations from the background-only hypothesis. The 95% confidence level (CL) upper limit on the Higgs boson pair production cross section is calculated using the frequentist CL_S method [171]. Exclusions and significances are evaluated using pseudo-experiments. Assuming SM branching ratios for the light Higgs boson decays, the expected upper limit is $1.0^{+0.6}_{-0.3}$ pb; the observed limit is 2.2 pb.

For the resonance analysis the expected exclusion improves from 1.8 to 0.8 pb as a function of m_X, as shown in Figure 8.12. This behavior derives from increased event-level acceptance at larger masses. The five events selected in the $m^{\gamma\gamma}$ signal region are shown in $m^{\gamma\gamma b\bar{b}}$, in Figure 8.11. The local probability of the compatibility to the background-only hypothesis, p_0, reaches a minimum of 0.002 at $m_X = 300$ GeV, corresponding to 3.0 standard deviations. After accounting for the look-elsewhere effect [172], the global probability of such an excess occurring at any mass in the range studied is 0.019, corresponding to 2.1 standard deviations. The number of events lying within the $m^{\gamma\gamma b\bar{b}}$ window of each mass hypothesis is readily apparent in 'steps' in the exclusion plot.

Systematic uncertainty		Non-Resonance Analysis			Resonance Analysis		
		SM h Bkd	hh Signal	Cont.	SM $h(h)$ Bkd	$X \to hh$ Signal	Cont.
Trigger	[%]	1		–	1		–
Luminosity	[%]	2.8		–	2.8		–
Photon	Identification [%]	2.4		–	2.4		–
	Isolation [%]	2		–	2		–
Mass	Resolution [%]	Resolution: 13		–	Migration: 1.6		–
	Position	Value: +0.5/-0.6 GeV		–	Migration: 1.7%		–
Shape	$m^{\gamma\gamma}$ Continuum [%]	–		11	–		11
	$m^{\gamma\gamma b\bar{b}}$: Statistical [%]	–		–	–		3-18
	$m^{\gamma\gamma b\bar{b}}$: jj vs bb [%]	–		–	–		0-30
	$m^{\gamma\gamma b\bar{b}}$: Fit Model [%]	–		–	–		16-30
Jets	b-Tagging [%]	3.3	1.8	–	3.4	2.4	–
	Energy Scale [%]	6.5	1.4	–	19	3.8	–
	b-jet Energy Scale [%]	2.6	0.3	–	6.5	2.2	–
	Energy Resolution [%]	4.8	6.3	–	15	9.3	–
Theory	PDF+Scale [%]	8.4	–	–	+18/-15	–	–
	Single h+HF [%]	14	–	–	14	–	–

Table 8.6: Summary of systematic uncertainties. Values marked '–' do not apply. Except for the uncertainty on m_h in the non-resonance analysis, and the $m^{\gamma\gamma}$ resolution for that fit, all uncertainties are the percent impact on the expected yields. The impact on the yield from the 13% uncertainty on the resolution is similar to the value used in the resonance analysis: a bit under 2%. The jet energy scale includes components from various sources, including uncertainties on jets arising from b quarks. The b-tagging uncertainty includes uncertainties for efficiencies to tag jets arising from b quarks as well as jets from c quarks and light-flavor quarks.

The limits derived are juxtaposed in Fig. 8.12 with the expectation from a sample type-I 2HDM not excluded by current data with $\cos(\beta - \alpha) = -0.05$ and $\tan(\beta) = 1$. The heavy Higgs bosons are taken to be degenerate in mass, and the mass of the lightest CP-even Higgs boson is set to 125 GeV. All major production mechanisms of $H \to hh$ are considered. Cross sections and branching ratios have been calculated as discussed in Ref. [173].

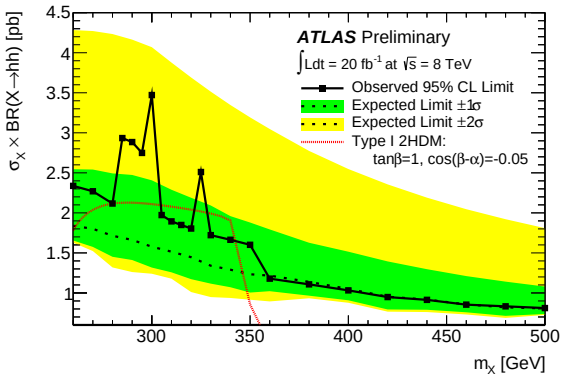


Figure 8.12: A 95% CL_S upper limit on the cross section times branching ratio of a narrow resonance decaying to pairs of Higgs bosons is presented as a function of m_X.

Chapter 9

Conclusions

The first run of the LHC is crowned by the long-awaited experimental observation of the Higgs boson. This document has presented the course of this discovery in the diphoton decay channel. Collosal efforts made in the design and construction of the ATLAS detector, and for the reconstruction, calibration, and identification of photons rendered the analysis strategy in this channel straightforward – a simultaneous fit of the continuum backgrounds and the resonant signal, in the diphoton invariant mass. The high signal selection efficiency of around 40% compensates for the low branching ratio of 0.2%, providing competitive statistics and making the channel attractive for a wide array of measurements.

The author contributed substantially to the identification of photons used in all $h \to \gamma\gamma$ analyses and in particular in the original discovery [12]. Motivated by measuring the couplings of the new particle [13], he worked to tag its production in association with a vector boson. He played a significant role in extracting its differential cross sections [14]. This work, along with measurements of the spin and CP eigenstate of the new particle, cemented the discovery of the Standard Model Higgs boson. To reach beyond the SM expectations and to begin to prepare for the coming runs of the LHC, a final search was performed for the resonant or non-resonant production of Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state [15]. A small excess over the background only expectation is not reproduced in CMS data [174, 175].

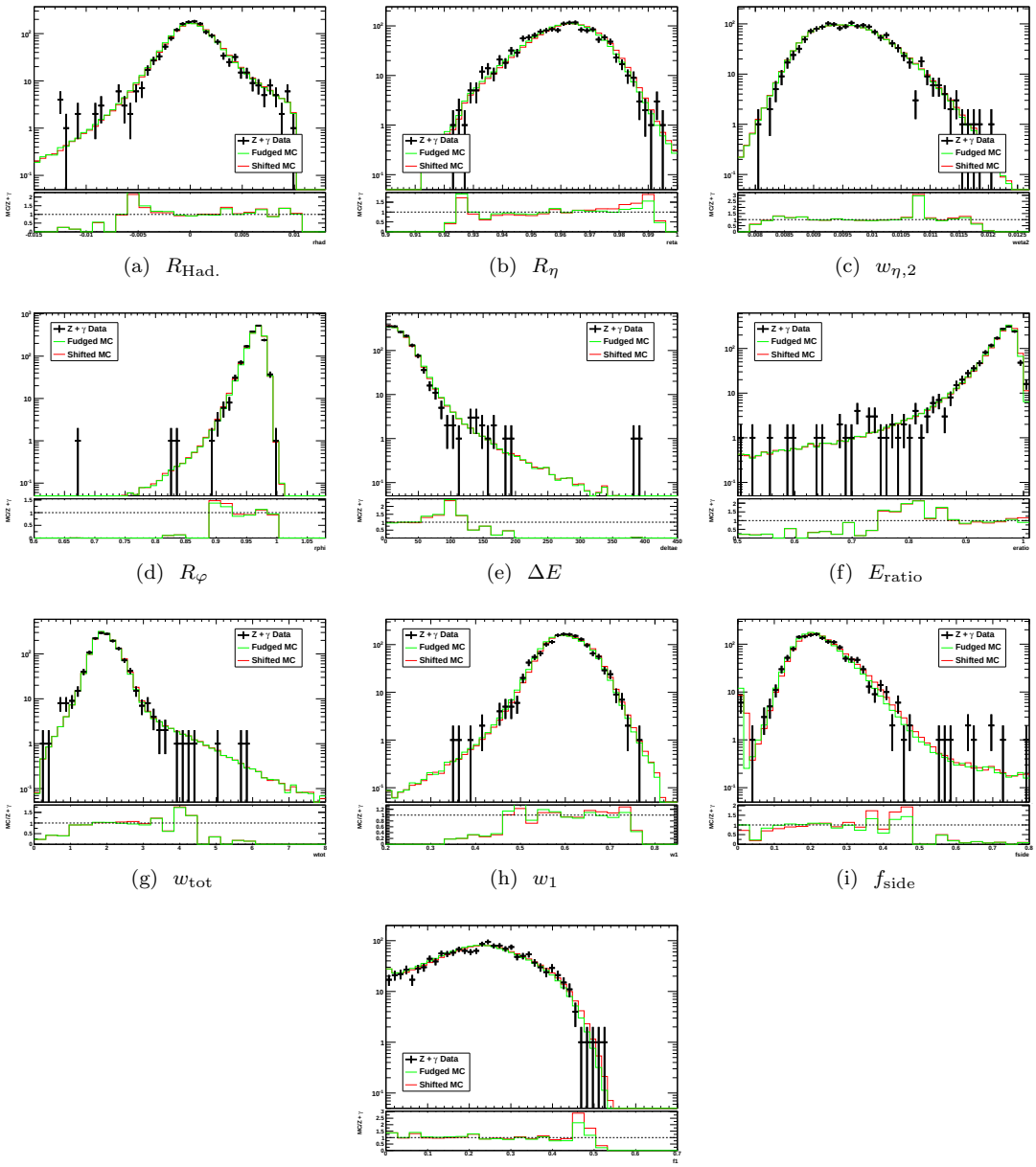
Yet with the keystone of the Standard Model finally hoisted into place, the bedrock is already fractured: the Standard Model is an incomplete theory. Precise measurements of the Cosmic Microwave Background (CMB) and of the angular momenta of galaxies, as well as many other observations, all indicate that around 80% of the total matter in the universe is non-baryonic. The CP violation observed in the Standard Model is insufficient to explain the dominance of matter over anti-matter in the universe. The mass of the Higgs boson itself is subject to radiative corrections that scale quadratically with any scale Λ of physics beyond the SM – unless protected by some new symmetry.

These questions and others motivate the coming runs of the LHC, and demand a continued broad program of searches and precision measurements. Perhaps the answers will come in the form of a simple resonance, suddenly accessible thanks to the increase in $\sqrt{s}$. Maybe hints of new physics will instead begin as whispers from rare flavor processes – from heavy states beyond the LHC reach running in loops. Yet again, the Higgs itself could open a portal to physics beyond the SM. The effective couplings to gluons or photons could hint at new colored or electrically charged states. Measurements of the $h \to \gamma\gamma$, and $h \to VV$ rates constrain the effective scale of higher-dimensional operators. Current limits on the unobserved or (non-SM) invisible width of Higgs decays still leave substantial space for new physics, so continued indirect measurements of the width through interference [176–178], direct searches for invisible decays, and clever parameterization of coupling measurements will all help to 'rout out' new physics, if it is hiding there. Direct searches for exotic decays will be ever more exciting, and there is a panoply of potential enhancements of Higgs boson pair production.

The discovery of the Higgs boson is the triumphant inauguration of the LHC program. The second run may open a new era for physics.

Appendices

A IsEM Distributions of $Z \to \ell\ell\gamma$ Radiative Decays



(a) $R_{\text{Had.}}$ (b) R_η (c) $w_{\eta,2}$

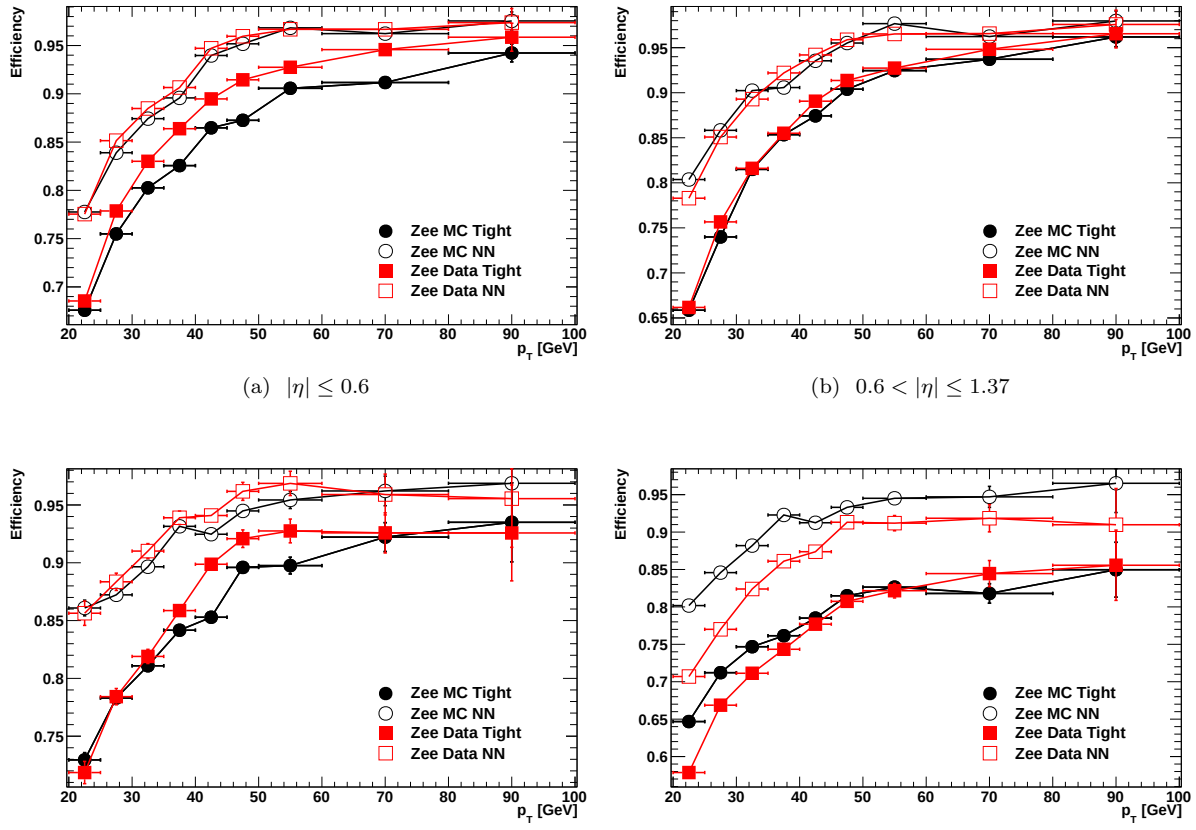
(d) R_φ (e) ΔE (f) E_{ratio}

(g) w_{tot} (h) w_1 (i) f_{side}

(j) f_1

Figure A.1: IsEM distributions for $Z \to \ell\ell\gamma$, with photons from final state radiation in 2011 data, compared to shifted direct photon Monte Carlo (unconverted, with $25 < p_T \leq 40$ GeV and $|\eta| < 1.4$).

B $Z \to ee$ Efficiencies in Data and Monte Carlo



(a) $|\eta| \leq 0.6$

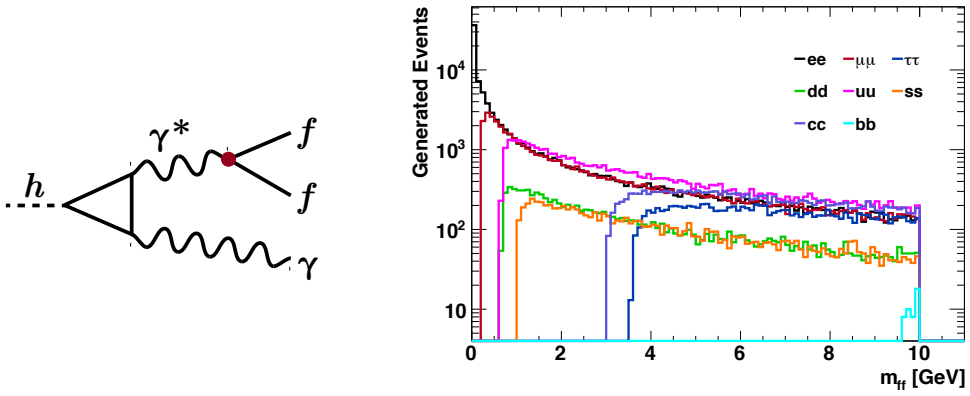
(b) $0.6 < |\eta| \leq 1.37$

(c) $1.52 < |\eta| \leq 1.81$

(d) $1.81 < |\eta| \leq 2.37$

Figure B.1: $Z \to ee$ probe electrons from data and (fudged) MC yield equivalent agreement for neural nets and tight ID.

C Higgs Decays to $f\bar{f}\gamma$



(a) Dalitz Decay

(b) Mass Spectrum of $f\bar{f}$ Pairs

Figure C.1: Higgs boson decays $h \to f\bar{f}\gamma$ proceed through an internal conversion or 'Dalitz decay' of an excited photon, at a rate approximately 10% of $h \to \gamma\gamma$. The mass spectrum of these decays shows the turn-ons at twice the dilepton mass, or the lightest meson masses, for each quark and lepton. The cut-off for this decay in PYTHIA8 is 10 GeV.

D Missing Energy Performance for Coupling Studies

D.1 Impact of the Redefinition of the Missing Energy

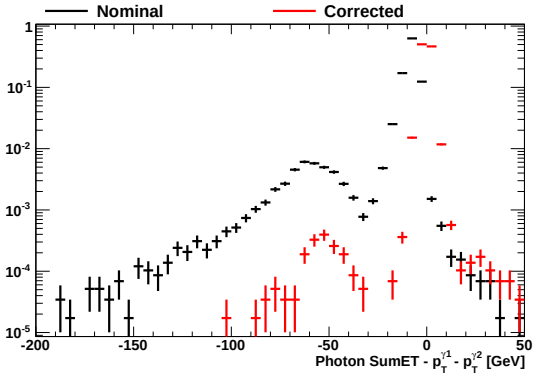


Figure D.1: The scalar sum of photon p_{T}s are compared to the value reconstructed by the $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm. Two effects are apparent. First, the fraction of events near zero increases dramatically. This is because photons are more reliably 'classified' as photons (instead of electrons) by the $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm: they do not even enter into the 'sum of photon p_{T}s' in many cases. The second change is in fact more important: the distribution of events *near* zero is much tighter. This demonstrates that the calibration of the photons closer matches the (best-knowledge) calibration used in the analysis. This is ultimately found to improve the accuracy ('linearity') of the final missing energy. Though the first effect appears more dramatic in the plot, it is less important, because objects not classed as photons are not rejected, they are merely included elsewhere.

D.2 Studies of the 'Soft Track Vertex Fraction'

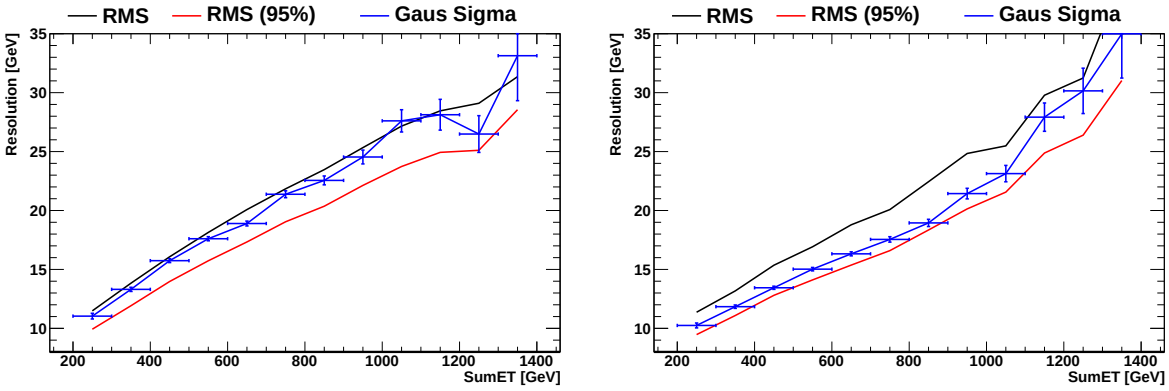
The 'Soft Track Vertex Fraction' was an algorithm, designed to improve the resolution of missing energy in events with large pileup bys reweight the soft jets and cell out terms in RefFinal by an estimate of the fractional contribution from the vertex of interest. This reweighting is defined as

$$\mathrm{STVF} \equiv \sum_{\substack{\text{Tracks,}\\ \text{Diphoton}\\ \text{Vertex}}} p_T \Big/ \sum_{\substack{\text{All}\\ \text{Tracks}}} p_T$$

The most probable value of this distribution is 0.2. In addition, soft jets are weighted or rejected according to their Jet Vertex Fraction (JVF). Because unconverted photons leave no track, there is an immediate technical challenge of ensuring that the vertex used is in fact the diphoton vertex. This is not as straightforward as, for instance $Z \to \ell\ell$ events, where the highest $\sum p_T^2$ vertex is reliably also the primary vertex.

Yet the problem with STVF runs much deeper. First, the proposal does not make logical sense: it proposes to scale local inhomogeneities by a global quantity. This procedure is very much unlike the Jet Vertex Fraction (JVF) where jets are accepted or rejected based on the fraction of their energy expected to come from the correct vertex. In that case, a decision about a jet is made based on information about that jet. For STVF, soft jets and energy deposits ('cell out') is scaled by a factor of 0.2-0.3 that has very little to do with that jet or that calorimeter cluster.

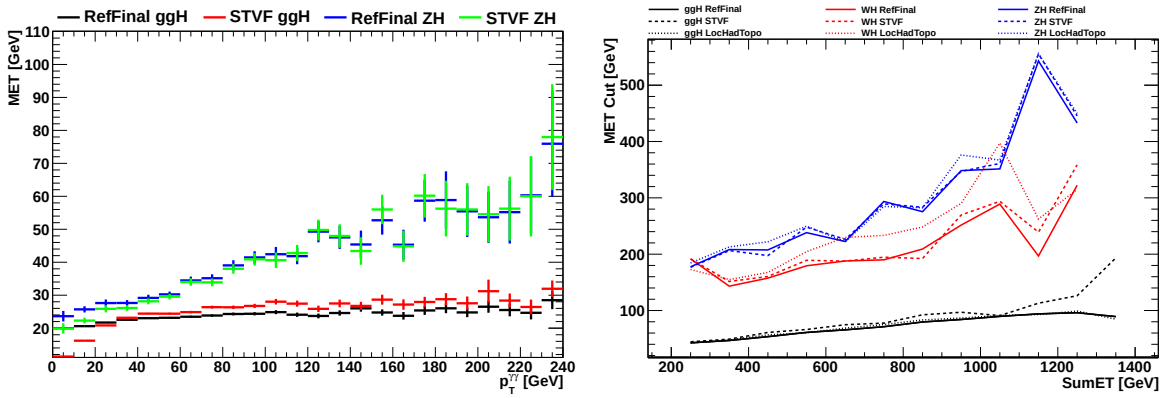
STVF effectively decreases the length of a 'walk,' but that walk is not only random. The cell out and soft jet terms were included for a reason – often, they are the 'return trip' of an excursion. In *ggh* events



(a) RefFinal (b) STVF

Figure D.2: The resolution for STVF and 'nominal' RefFinal are compared as a function of the event $\sum E_{\mathrm{T}}$, using several metrics: (1) the RMS of the E_x distribution around zero, (2) the RMS of the narrowest 95% of events, and (3) the width of a Gaussian, fit within $\pm 2\sigma$. The STVF shows a better overall resolution.

with very small $p_{\mathrm{T}}^{\gamma\gamma}$ STVF appears to perform exceptionally well, because it does not let the $E_{\mathrm{T}}^{\mathrm{miss}}$ 'wander' as far from zero. Because most events have low $p_{\mathrm{T}}^{\gamma\gamma}$, the overall resolution appears to improve with respect to the 'nominal' RefFinal algorithm. This is shown in Figure D.2, as a function of the $\sum E_{\mathrm{T}}$ of the event. However, for events with $p_{\mathrm{T}}^{\gamma\gamma} > 30$ and no real $E_{\mathrm{T}}^{\mathrm{miss}}$, the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ is actually much larger for the STVF algorithm: in those cases, the soft energy is recoiling against the diphoton system, and is very necessary to 'return to 0.' Because it has been scaled down, it is unable to accomplish this effectively. This is shown in Figure D.3a. Figure D.3b shows that this behavior leads to larger tails in ggh $E_{\mathrm{T}}^{\mathrm{miss}}$ when using the STVF algorithm. This in turn led to worse purity in Vh and $t\bar{t}h$ events, in the '$E_{\mathrm{T}}^{\mathrm{miss}}$ category.' Needless to say, STVF was not used.

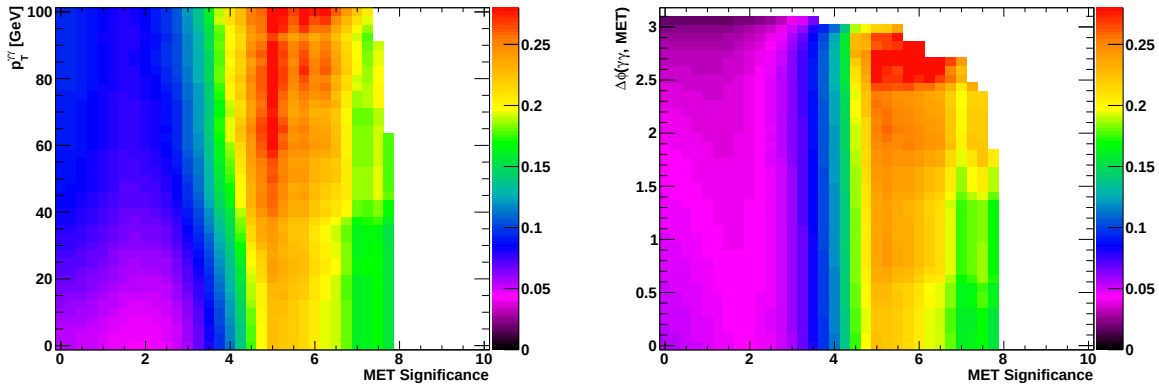


(a) Profile of $E_{\mathrm{T}}^{\mathrm{miss}}$ v. $p_{\mathrm{T}}^{\gamma\gamma}$

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$ Tails v. $\sum E_{\mathrm{T}}$

Figure D.3: Although it performs better on average, STVF performs poorly in the tails of the distribution, needed to identify Vh and $t\bar{t}h$ events in the $h \to \gamma\gamma$ analysis. For ggh events with large $p_{\mathrm{T}}^{\gamma\gamma}$, the soft terms serve to 'return' the $E_{\mathrm{T}}^{\mathrm{miss}}$ to zero. By scaling these terms down, they do not function correctly. (a) For $p_{\mathrm{T}}^{\gamma\gamma} > 30$ GeV, the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ is larger using the STVF reconstruction than the nominal reconstruction, in ggh events without real missing energy. Since the mean and peak of a Rayleigh ($E_{\mathrm{T}}^{\mathrm{miss}}$) distribution with no true $E_{\mathrm{T}}^{\mathrm{miss}}$ is proportional to the resolution, this also shows that the resolution grows with $p_{\mathrm{T}}^{\gamma\gamma}$. (b) The tails of $E_{\mathrm{T}}^{\mathrm{miss}}$ are presented for the 'nominal,' STVF, and 'LocHadTopo' $E_{\mathrm{T}}^{\mathrm{miss}}$ definitions. These 'tail' is quantified as the cut required to reject 99.5% of events. For ggh, the cuts are lowest for RefFinal: it has the smallest tail.

E Additional Variables for a $E_{\mathrm{T}}^{\mathrm{miss}}$-Only Category



(a) $\Delta\varphi\left(\gamma\gamma, E_{\mathrm{T}}^{\mathrm{miss}}\right)$ (b) $p_{\mathrm{T}}^{\gamma\gamma}$

Figure E.1: The boost of the diphoton system and its angular separation from the $E_{\mathrm{T}}^{\mathrm{miss}}$ provide 'confirmation' of real missing energy in an event. The gains from additional cuts on these variables are investigated, defining the metric $S/\sqrt{B}$ using S as the expected Vh events, and B as the data events in the sideband.

F Missing Energy Systematics for Coupling Studies

	ggH		VBFH		WH		ZH		ttH	
	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
Photon Energy Scale	-0.0%	5.8%	1.7%	3.8%	-1.3%	-0.4%	0.2%	0.9%	-0.2%	0.4%
Photon Energy Resolution	5.3%	4.1%	7.2%	6.2%	0.4%	-1.1%	-0.2%	0.2%	0.3%	-0.0%
Electron Energy Scale	0.0%	0.0%	0.0%	0.5%	0.0%	-0.4%	0.0%	0.2%	-0.0%	-0.1%
Electron Energy Resolution	0.0%	0.0%	0.4%	0.4%	-0.4%	-0.4%	0.2%	0.0%	0.0%	-0.0%
Muon Res. - Inner Detector	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%	0.0%	0.0%	0.2%	0.2%
Muon Res. - Muon Spect.	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%	0.0%	0.0%	0.2%	0.2%
Jet Energy Scale	18.2%	3.9%	35.8%	-3.8%	-0.2%	2.1%	0.2%	-1.3%	5.2%	-2.7%
– CloseBy	8.6%	5.8%	13.4%	-8.0%	0.2%	1.7%	0.4%	-0.4%	4.5%	-2.4%
– Mu	1.4%	0.7%	0.8%	4.2%	0.0%	0.0%	-0.1%	0.1%	-0.1%	-0.0%
– NPV	3.1%	2.3%	1.1%	3.6%	-0.4%	0.0%	0.1%	-0.2%	-0.1%	0.1%
– PileupTopo	4.9%	-0.0%	9.2%	0.2%	-0.4%	0.0%	0.2%	0.0%	0.4%	-0.5%
– PileupPt	1.3%	0.0%	3.0%	1.1%	0.0%	0.0%	0.0%	0.1%	-0.0%	-0.1%
– Statistical NP1	1.5%	-1.2%	4.1%	0.0%	-0.4%	0.0%	0.0%	0.1%	0.2%	-0.1%
– Statistical NP2	0.0%	1.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	-0.0%
– Statistical NP3	0.6%	-0.1%	4.1%	0.0%	-0.4%	0.0%	0.0%	0.0%	0.2%	-0.2%
– Modelling NP1	7.3%	0.1%	10.9%	-0.3%	-0.7%	0.1%	0.1%	-0.4%	0.8%	-1.0%
– Modelling NP2	-0.5%	1.8%	4.1%	-1.8%	0.0%	-0.4%	0.0%	0.0%	0.5%	-0.4%
– Modelling NP3	0.6%	1.0%	0.0%	4.1%	0.0%	0.0%	0.0%	0.0%	-0.1%	0.1%
– Modelling NP4	1.0%	0.0%	3.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	-0.1%
– Detector NP1	1.0%	-0.2%	11.1%	0.0%	-0.4%	0.0%	0.2%	0.1%	0.7%	-0.8%
– Detector NP2	0.0%	0.0%	0.0%	4.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
– Detector NP3	1.8%	0.0%	4.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	-0.0%
– Mixed NP1	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	-0.1%
– Mixed NP2	1.8%	-0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
– η Intercalib. Model	10.1%	-0.8%	16.2%	-1.4%	-0.2%	-0.5%	0.0%	-0.2%	1.0%	-0.9%
– η Intercalib. Stat/Method	1.3%	0.9%	9.2%	0.8%	-0.4%	0.0%	0.2%	0.1%	0.3%	-0.3%
– Single Particle High-p_T	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
– Pythia Rel. Non-Closure	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Jet Energy Resolution	17.1%	0.0%	12.9%	0.0%	-0.4%	0.0%	-0.4%	0.0%	0.9%	0.0%
SoftScale	94.3%	-27.3%	30.5%	-14.4%	1.1%	-0.5%	0.7%	-1.2%	0.8%	-0.3%
SoftRes	16.8%	28.5%	7.0%	9.6%	1.2%	0.6%	-0.2%	1.5%	0.6%	0.1%

Table F.1: Signal systematics for the five production modes, divided among the object uncertainties. In this table, 'ES' is Energy Scale and 'ER' is Energy Resolution. The largest uncertainties are due to jet and soft term scales and resolution, in ggH and VBF events. Events in these categories are in the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions; the cuts are thus (definitionally) at the highest point of the falling spectrum. Small changes in the number of events have a larger fractional impact. Moreover, the events that populate the tails are the most affected by changes in the scale and resolution.

G Migration Uncertainties for the Coupling Analysis

Systematic	Categories	Value by Mode [%]			
		All	ggh	VBF	$Vh/t\bar{t}h$
Material Modelling	Unconverted	-4.0			
	Converted	+3.5			
Transverse Momentum	Low $p_{\mathrm{T}t}$	+1.3			
	High $p_{\mathrm{T}t}$	-10.2			
	Tight VBF	-10.4			
	Loose VBF	-8.5			
	Vh Hadronic	-12.5			
	$E_{\mathrm{T}}^{\mathrm{miss}}$	-2.0			
	Lepton	-4.0			
Underlying Event	Tight VBF		+8.8	+2.0	+8.8
	Loose VBF		+12.8	+3.3	+12.8
	Vh Hadronic		+12.0	+3.9	+12.0
Jet Energy Scale	Low $p_{\mathrm{T}t}$		-0.1	-1.0	-0.1
	High $p_{\mathrm{T}t}$		-0.7	-1.3	+0.4
	Tight VBF		+11.8	+6.7	+20.2
	Loose VBF		+10.7	+4.0	+5.7
	Vh Hadronic		+4.7	+2.6	+1.4
	Lepton		+0.0	+0.0	-0.1
Jet Energy Resolution	Low $p_{\mathrm{T}t}$		+0.0	+0.2	+0.0
	High $p_{\mathrm{T}t}$		-0.2	+0.2	+0.6
	Tight VBF		+3.8	-1.3	+7.0
	Loose VBF		+3.4	-0.7	+1.2
	Vh Hadronic		+0.5	+3.4	-1.3
	Lepton		-0.9	-0.5	-0.1
Missing Energy	$E_{\mathrm{T}}^{\mathrm{miss}}$		+66.4	+30.7	+1.2
Jet Vertex Fraction	Tight VBF		-1.2	-0.3	-1.2
	Loose VBF		-2.3	-2.4	-2.3
η^* Modelling	Tight VBF	+7.6			
	Loose VBF	+6.2			
Dijet Angular Modelling	Tight VBF	+12.1			
	Loose VBF	+8.5			

Table G.2: Migration uncertainties for the couplings analysis. A Gaussian constraint is used for all uncertainties except the underlying event, which uses a log-normal constraint. The uncertainties on electron and muon reconstruction and energy scale/resolution are found to be negligible. The uncertainties on the jet energy scale and resolution are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ category and included in that uncertainty. Correlations between categories for a single uncertainty, are denoted by the sign of the uncertainty. If a category is not listed for a given uncertainty, the impact on that category is negligible.

H Differential Cross Sections: Variable Summary

Variable summary tables are provided for each of the 7 differential cross sections measured. These contain descriptions of the uncertainties, yields, unfoldings, and results of the full study. Several of the uncertainties are flat across observables and bins, for example the luminosity (2.8%), trigger (0.5%), and particle identification (2.4%). These are described in Sections 5.5.1 and 7.4. The destructive interference described in Section 5.2.1.1 reduces the ggh yield by approximately 2.3%. This effect is not unfolded.

As the measurements are dominated by statistical uncertainties, the bin-to-bin correlations are weak. The largest correlation is found for $p_{\mathrm{T}}^{\gamma\gamma jj}$ and is less than 20%. For non-jet variables, typical correlations are below 10%.

The expected yields in the fiducial region and after full simulation, are for $m_h = 125$ GeV (simulated with the nominal POWHEG +PYTHIA8 samples described in Section 5.2).

The 'Rate and Selection' entry includes the luminosity and trigger, and efficiency uncertainties for particle identification and isolation. 'Fiducial Region (MC)' denotes the particle level yields for SM Higgs boson production with $m_h = 125$ GeV, within the fiducial region described in Section 7.1.2, and generated with POWHEG. 'Fully Simulated (MC)' denotes the final yield of the same sample after full simulation and reconstruction, and the full requirements outlined in Sections 5.1 and 7.1.1. Details of the signal and background PDFs used in the fit are given: the width (in GeV) of the dominant Crystal Ball component of the signal PDF, and the model used to describe the background (Exp1 and Exp2 here denotes an exponential function of a polynomial of first and second order, respectively).

	N_{jets}	0	1	2	3+
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.82	1.84	1.78	1.74
	Mass Resolution Uncertainty	23%	24%	26%	24%
Background Model	Function	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	360	168	64	21
	Fully Simulated (MC)	227	115	47	17
	Data 105-160 GeV	60100	23431	7782	2822
	Data Fit Signal Yield	271	198	84	55
Yield Uncertainties	Fit: Statistical + Resolution	37%	31%	43%	39%
	Rate and Selection	3.9%	3.9%	4.2%	5.5%
	Background Modelling	6.6%	3.6%	4.9%	3.0%
	Jet Resolution + Scale	$^{+3.9\%}_{-4.2\%}$	$^{+2.7\%}_{-2.9\%}$	$^{+7.8\%}_{-7.5\%}$	$^{+14.6\%}_{-14.0\%}$
	Composition + Shape	$^{+3.7\%}_{+1.2\%}$	$^{+5.4\%}_{+2.7\%}$	$^{+3.4\%}_{+3.2\%}$	$^{+7.9\%}_{+4.5\%}$
Unfolding	Truth to Reco. 'Purity'	95%	78%	70%	70%
	Correction Factor	159%	145%	135%	123%
Results [fb]	Measured Spectrum	21	14	5.6	3.3
	Total Uncertainty	±38%	±32%	±44%	±42%

Table H.1: Variable summary for N_{jets}. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

$\lvert y^{\gamma\gamma}\rvert$		0-0.3	0.3-0.65	0.65-1.0	1.0-1.4	1.4-2.4
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.62	1.70	1.88	2.10	2.19
	Mass Resolution Uncertainty	23%	24%	28%	26%	23%
Background Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	120	135	123	119	115
	Fully Simulated (MC)	93	99	81	68	65
	Data 105-160 GeV	18227	19479	17571	19085	19773
	Data Fit Signal Yield	125	126	68	98	210
Yield Uncertainties	Fit: Statistical + Resolution	42%	45%	76%	62%	30%
	Rate and Selection	3.9%	3.9%	3.9%	3.9%	3.9%
	Background Modelling	2.8%	3.0%	7.4%	5.1%	2.7%
	Composition + Shape	$^{+1.3\%}_{+1.3\%}$	$^{+1.4\%}_{+1.4\%}$	$^{+1.3\%}_{+1.3\%}$	$^{+1.4\%}_{+1.4\%}$	$^{+1.4\%}_{+1.4\%}$
Unfolding	Truth to Reco. 'Purity'	98%	97%	97%	97%	97%
	Correction Factor	129%	136%	152%	174%	177%
Results [fb]	Measured Spectrum	26	24	15	21	18
	Total Uncertainty	$\pm$42%	$\pm$46%	$\pm$76%	$\pm$62%	$\pm$31%

Table H.2: Variable summary for $\lvert y^{\gamma\gamma}\rvert$. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

	$p_{\mathrm{T}}^{\gamma\gamma}$ [GeV]	0-20	20-30	30-40	40-50	50-60	60-80	80-100	100-200	Overflow
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.86	1.87	1.85	1.86	1.88	1.83	1.81	1.74	1.56
	Mass Resolution Uncertainty	22%	22%	22%	22%	23%	24%	24%	26%	31%
Background Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp1
Yields	Fiducial Region (MC)	196	98	73	53	40	54	33	53	11
	Fully Simulated (MC)	131	65	48	35	27	36	22	35	8
	Data 105-160 GeV	42010	16133	10841	7345	5145	6351	3036	3056	218
	Data Fit Signal Yield	65	167	87	87	58	54	40	48	6
Yield Uncertainties	Fit: Statistical + Resolution	123%	33%	50%	42%	50%	60%	59%	43%	80%
	Rate and Selection	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	–
	Background Modelling	21.7%	3.5%	4.2%	3.3%	4.5%	8.2%	6.0%	4.9%	–
	Composition + Shape	$^{+2.3\%}_{+1.2\%}$	$^{+3.0\%}_{+1.2\%}$	$^{+3.2\%}_{+2.0\%}$	$^{+3.4\%}_{+1.2\%}$	$^{+4.0\%}_{+1.9\%}$	$^{+3.9\%}_{+1.6\%}$	$^{+3.5\%}_{+1.7\%}$	$^{+3.4\%}_{+1.9\%}$	–
Unfolding	Truth to Reco. 'Purity'	96%	89%	89%	88%	88%	93%	92%	97%	–
	Correction Factor	150%	150%	152%	151%	152%	152%	152%	150%	–
Results [fb/GeV]	Measured Spectrum	0.24	1.23	0.65	0.65	0.44	0.20	0.15	0.036	–
	Total Uncertainty	±125%	±33%	±50%	±42%	±51%	±61%	±60%	±44%	–

Table H.3: Variable summary for $p_{\mathrm{T}}^{\gamma\gamma}$. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

	$\lvert\cos\theta^*\rvert$	0.0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0
Signal	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.75	1.78	1.81	1.82	1.86	1.91	1.97	2.01	1.91	1.85
Parameters	Mass Resolution Uncertainty	25%	25%	24%	24%	24%	24%	22%	21%	21%	24%
Bkd Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	88	86	84	80	76	70	62	41	13	11
	Fully Simulated (MC)	61	57	55	53	51	47	41	26	8	7
	Data 105-160 GeV	10262	9650	9746	9930	11634	12923	14491	12079	2014	1406
	Data Fit Signal Yield	66	64	122	63	17	96	39	86	40	16
Yield Uncertainties	Fit: Statistical + Resolution	57%	57%	32%	62%	239%	49%	122%	55%	47%	91%
	Rate and Selection	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%	3.9%
	Background Modelling	7.0%	6.2%	3.0%	4.4%	22.4%	6.0%	20.7%	8.5%	3.3%	7.2%
	Composition + Shape	+5.0% +1.4%	+6.1% +1.4%	+5.7% +1.2%	+6.0% +1.2%	+2.3% +1.4%	+1.2% +1.2%	+1.6% +1.6%	+1.6% +2.3%	+4.0% +4.0%	+2.0% +2.0%
Unfolding	Truth to Reco. 'Purity'	99%	98%	98%	98%	98%	98%	98%	95%	95%	97%
	Correction Factor	145%	151%	152%	153%	149%	149%	152%	158%	156%	154%
Results [fb]	Measured Spectrum	47	48	91	48	12	70	30	67	30	12
	Total Uncertainty	±58%	±58%	±32%	±62%	±240%	±49%	±123%	±56%	±47%	±91%
Interference	Correction to ggH Yield	-1.8%	-1.8%	-1.8%	-1.9%	-2.1%	-2.2%	-2.5%	-3.0%	-4.1%	-10.7%

Table H.4: Variable summary for $\lvert\cos\theta^*\rvert$. Errors presented are fractions of the extracted and unfolded yields. The large apparent variations in the scale of the uncertainties (e.g., 0.4-0.5) is due to fluctuations in the extracted yields, and not to the uncertainties themselves. See the beginning of this Appendix or Chapter 7 for full details. The interference correction discussed in Section 5.2.1.1 depends directly on $\lvert\cos\theta^*\rvert$, so the scale of this correction is presented here, for convenience.

	p_{T}^{j1} [GeV]	0-30	30-50	50-70	70-100	100-140	Overflow
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.82	1.85	1.85	1.79	1.74	1.63
	Mass Resolution Uncertainty	23%	24%	24%	24%	26%	29%
Background Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	360	106	57	44	25	20
	Fully Simulated (MC)	227	77	41	30	17	14
	Data 105-160 GeV	60100	16514	8184	5427	2400	1510
	Data Fit Signal Yield	267	145	53	56	42	39
Yield Uncertainties	Fit: Statistical + Resolution	37%	36%	66%	51%	48%	39%
	Rate and Selection	3.9%	3.9%	3.9%	3.9%	3.9%	–
	Background Modelling	6.7%	4.4%	6.5%	7.2%	4.8%	–
	Jet Resolution + Scale	$^{+3.9\%}_{-4.5\%}$	$^{+4.7\%}_{-5.3\%}$	$^{+4.8\%}_{-4.4\%}$	$^{+5.4\%}_{-5.2\%}$	$^{+7.0\%}_{-6.6\%}$	–
	Composition + Shape	$^{+3.7\%}_{+1.2\%}$	$^{+6.6\%}_{+3.9\%}$	$^{+4.4\%}_{+6.2\%}$	$^{+2.8\%}_{+2.0\%}$	$^{+6.1\%}_{+2.3\%}$	–
Unfolding	Truth to Reco. 'Purity'	95%	67%	65%	71%	73%	–
	Correction Factor	159%	138%	141%	144%	143%	–
Results [fb/GeV]	Measured Spectrum	0.70	0.49	0.18	0.13	0.073	–
	Total Uncertainty	±39%	±37%	±67%	±52%	±49%	–

Table H.5: Variable summary for p_{T}^{j1}. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

$\Delta\varphi_{jj}$		0-π/3	π/3-2π/3	2π/3-5π/6	5π/6-π	Overflow
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.70	1.77	1.82	1.84	1.86
	Mass Resolution Uncertainty	27%	26%	25%	24%	23%
Background Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	26	30	15	14	527
	Fully Simulated (MC)	20	22	11	11	342
	Data 105-160 GeV	1932	2429	2412	3831	83531
	Data Fit Signal Yield	26	34	21	69	482
Yield Uncertainties	Fit: Statistical + Resolution	71%	61%	93%	40%	26%
	Rate and Selection	4.0%	4.0%	4.0%	4.0%	–
	Background Modelling	4.0%	4.8%	9.5%	5.5%	–
	Jet Resolution + Scale	$^{+10.4\%}_{-9.6\%}$	$^{+9.0\%}_{-8.7\%}$	$^{+9.0\%}_{-8.6\%}$	$^{+9.8\%}_{-10.9\%}$	–
	Composition + Shape	$^{+4.5\%}_{+4.2\%}$	$^{+4.4\%}_{+4.2\%}$	$^{+4.0\%}_{+3.1\%}$	$^{+6.8\%}_{+5.5\%}$	–
Unfolding	Truth to Reco. 'Purity'	72%	74%	71%	69%	–
	Correction Factor	129%	134%	133%	127%	–
Results [fb]	Measured Spectrum	1.6	2.1	2.7	8.3	–
	Total Uncertainty	±72%	±62%	±94%	±42%	–

Table H.6: Variable summary for $\Delta\varphi_{jj}$. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

	$p_{\mathrm{T}}^{\gamma\gamma jj}$ [GeV]	0-30	30-55	55-80	80-140	Overflow
Signal Parameters	σ_{CB} [GeV] at $m_H = 126.8$ GeV	1.78	1.76	1.74	1.73	1.86
	Mass Resolution Uncertainty	25%	25%	26%	27%	23%
Background Model	Function	Exp2	Exp2	Exp2	Exp2	Exp2
Yields	Fiducial Region (MC)	48	22	8	5	529
	Fully Simulated (MC)	34	18	7	4	343
	Data 105-160 GeV	5524	3314	1154	533	83610
	Data Fit Signal Yield	87	21	35	5	485
Yield Uncertainties	Fit: Statistical + Resolution	36%	113%	58%	185%	27%
	Rate and Selection	4.0%	4.0%	4.0%	4.0%	–
	Background Modelling	3.9%	14.0%	3.7%	15.8%	–
	Jet Resolution + Scale	$^{+11.4\%}_{-9.3\%}$	$^{+10.9\%}_{-12.3\%}$	$^{+8.5\%}_{-9.3\%}$	$^{+5.6\%}_{-6.5\%}$	–
	Composition + Shape	$^{+3.3\%}_{+3.2\%}$	$^{+4.8\%}_{+4.4\%}$	$^{+7.7\%}_{+5.5\%}$	$^{+10.7\%}_{+6.1\%}$	–
Unfolding	Truth to Reco. 'Purity'	70%	51%	49%	61%	–
	Correction Factor	140%	121%	120%	124%	–
Results [fb/GeV]	Measured Spectrum	0.20	0.050	0.083	0.0051	–
	Total Uncertainty	±38%	±114%	±59%	±186%	–

Table H.7: Variable summary for $p_{\mathrm{T}}^{\gamma\gamma jj}$. Errors presented are fractions of the extracted and unfolded yields. See the beginning of this Appendix or Chapter 7 for full details.

I Differential Cross Sections: Alternative Theoretical Predictions

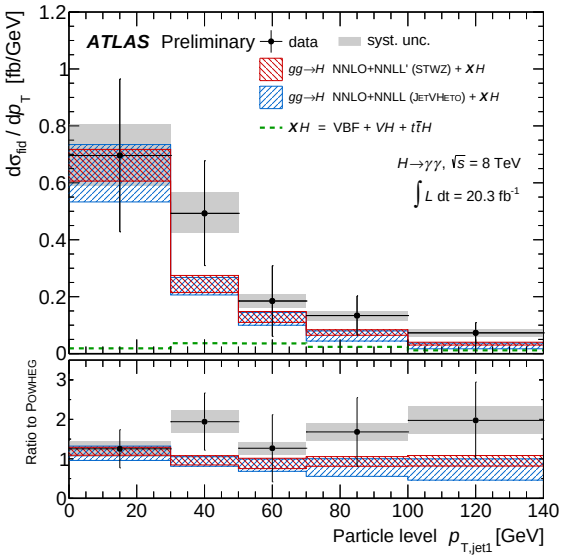
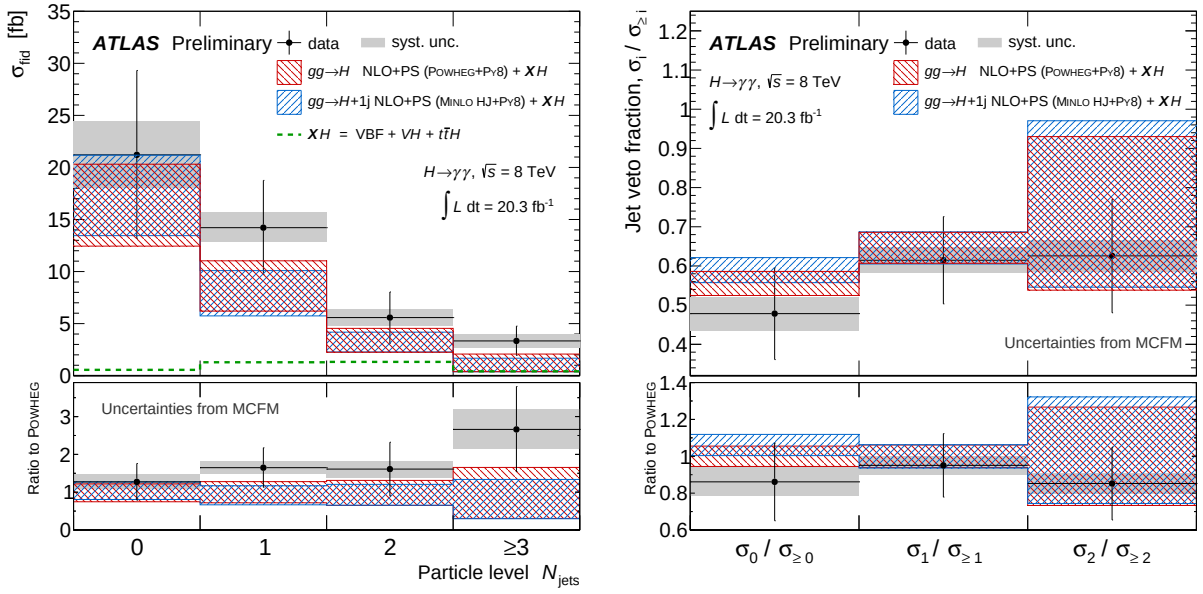


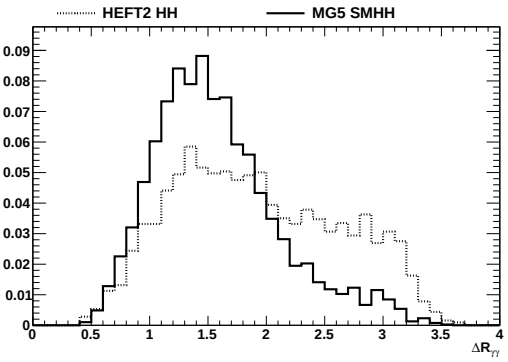
Figure I.1: The differential cross section of the leading partonic radiation from events with Higgs bosons is presented, and compared to higher-order calculations from STWZ and JetVeto.



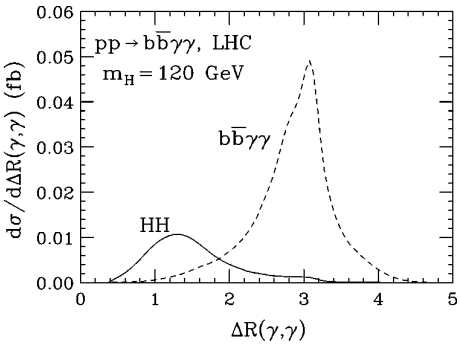
(a) N_{jets} (b) $\sigma_{N_{\text{jets}}=i}/\sigma_{N_{\text{jets}}\geq i}$

Figure I.2: Differential cross sections for N_{jets} and $\sigma_{N_{\text{jets}}=i}/\sigma_{N_{\text{jets}}\geq i}$ with theoretical uncertainties evaluated with the Stewart-Tackmann procedure.

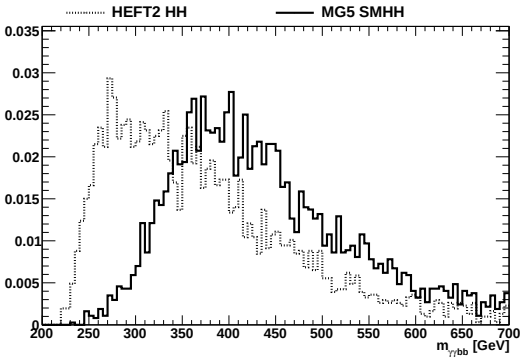
J Simulation Samples for the Higgs Pair Production Search



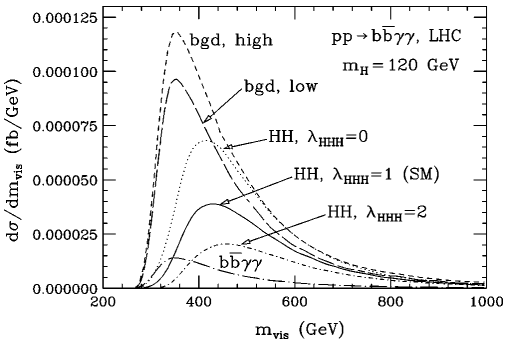
(a) MADGRAPH 5: $\Delta R(\gamma,\gamma)$



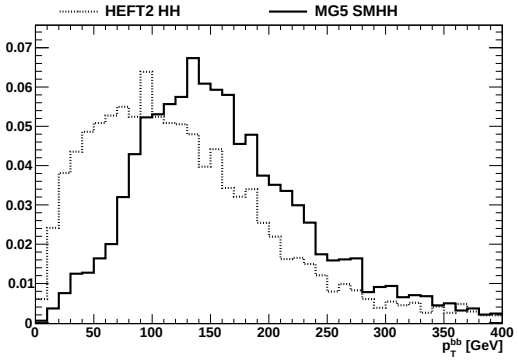
(b) Baur et al [153]: $\Delta R(\gamma,\gamma)$



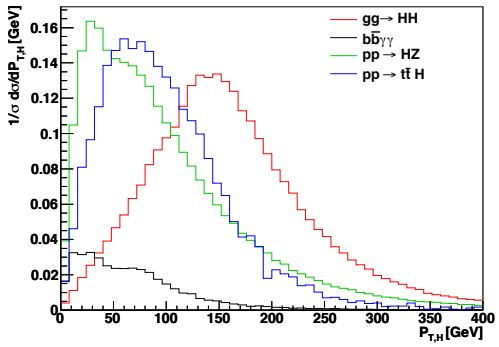
(c) MADGRAPH 5: $m^{\gamma\gamma b\bar{b}}$



(d) Baur et al: $m^{\gamma\gamma b\bar{b}}$

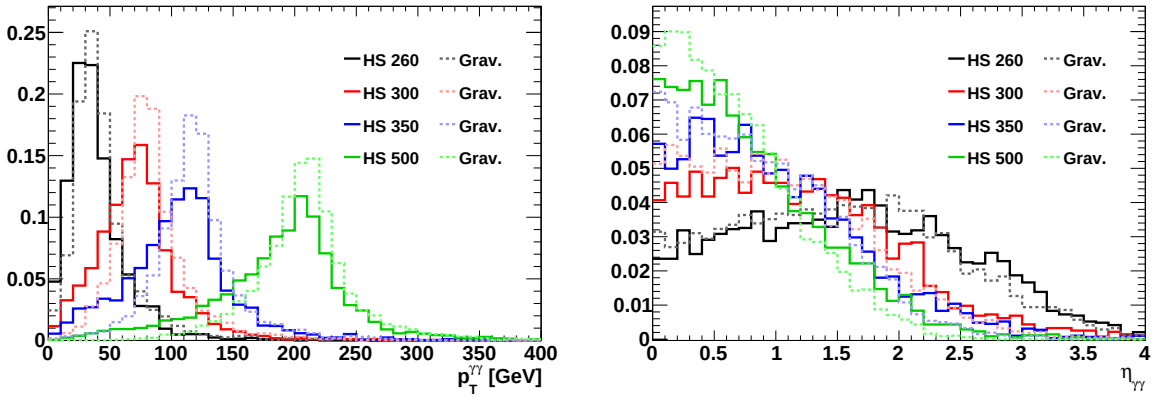


(e) MADGRAPH 5: $p_{\mathrm{T}}^{b\bar{b}}$



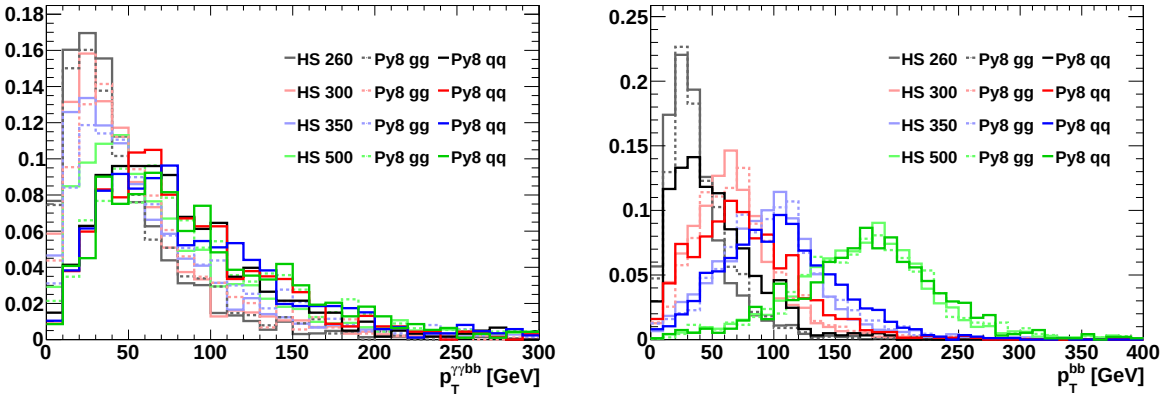
(f) Baglio et al [24]: $p_{\mathrm{T}}^{b\bar{b}}$

Figure J.1: Kinematic distributions are compared for the two available SM dihiggs predictions from Madgraph. Substantial differences are seen. Comparisons to theoretical predictions are provided [24, 153], and motivate the use of the new (MG5, SMHH) model.



(a) $p_{\mathrm{T}}^{\gamma\gamma}$ (b) $|\eta^{\gamma\gamma}|$

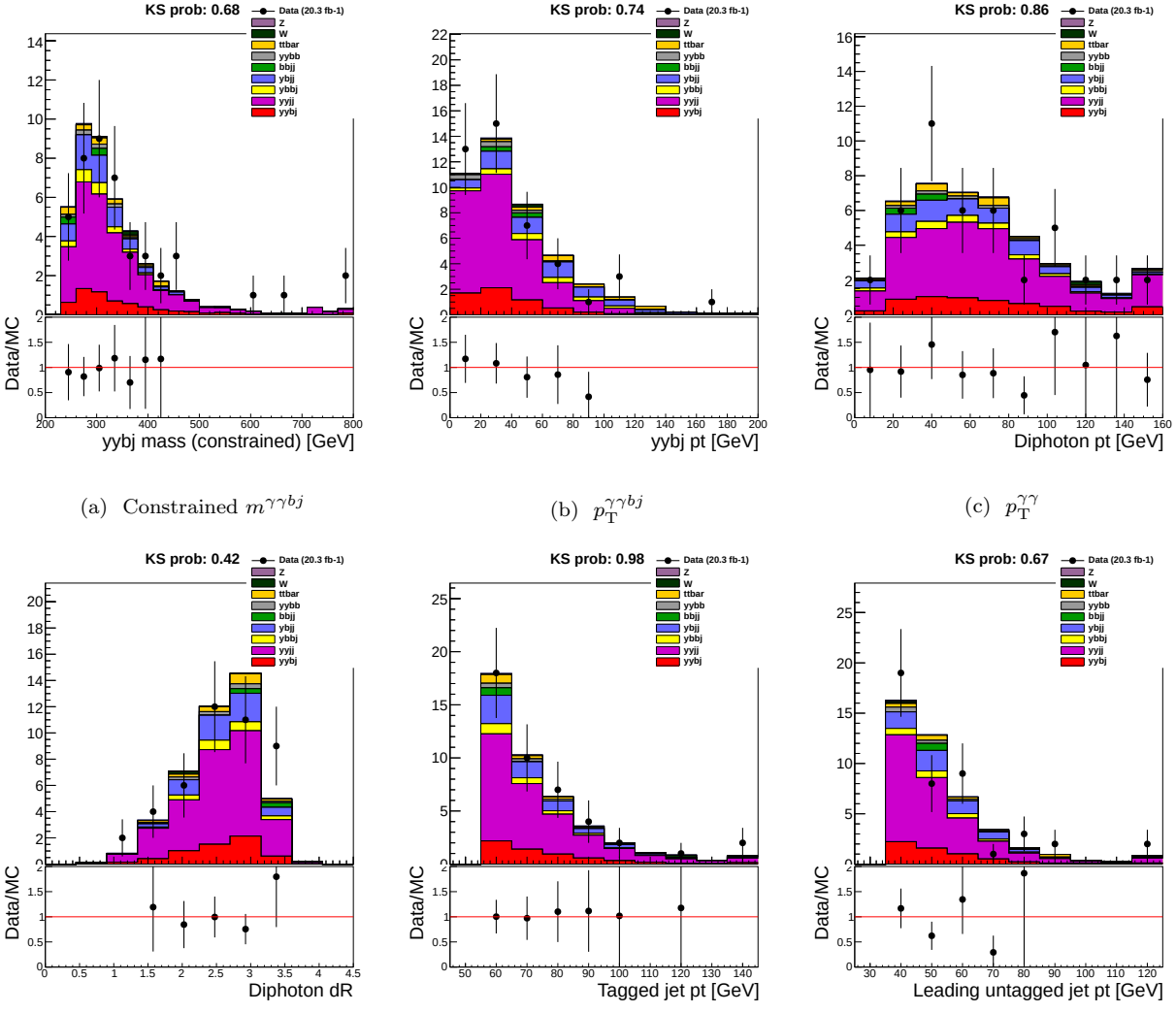
Figure J.2: The angular distributions of the Randall-Sundrum KK graviton are somewhat different from the 'benchmark' gluon-fusion model. In particular, the light Higgs bosons tend to be more central and harder than in the 2HDM benchmark.



(a) $p_{\mathrm{T}}^{\gamma\gamma b\bar{b}}$ (b) $p_{\mathrm{T}}^{b\bar{b}}$

(c) p_{T}^{b1} (d) p_{T}^{b2}

Figure J.3: Kinematic distributions are compared for qqH and ggH production in a 2HDM. Events generated with the benchmark MADGRAPH 'heavy scalar' model are compared to the PYTHIA8 predictions for gluon fusion and VBF. The profiles are overall quite similar, except in the boost of the full $\gamma\gamma b\bar{b}$ system, $p_{\mathrm{T}}^{\gamma\gamma b\bar{b}}$.

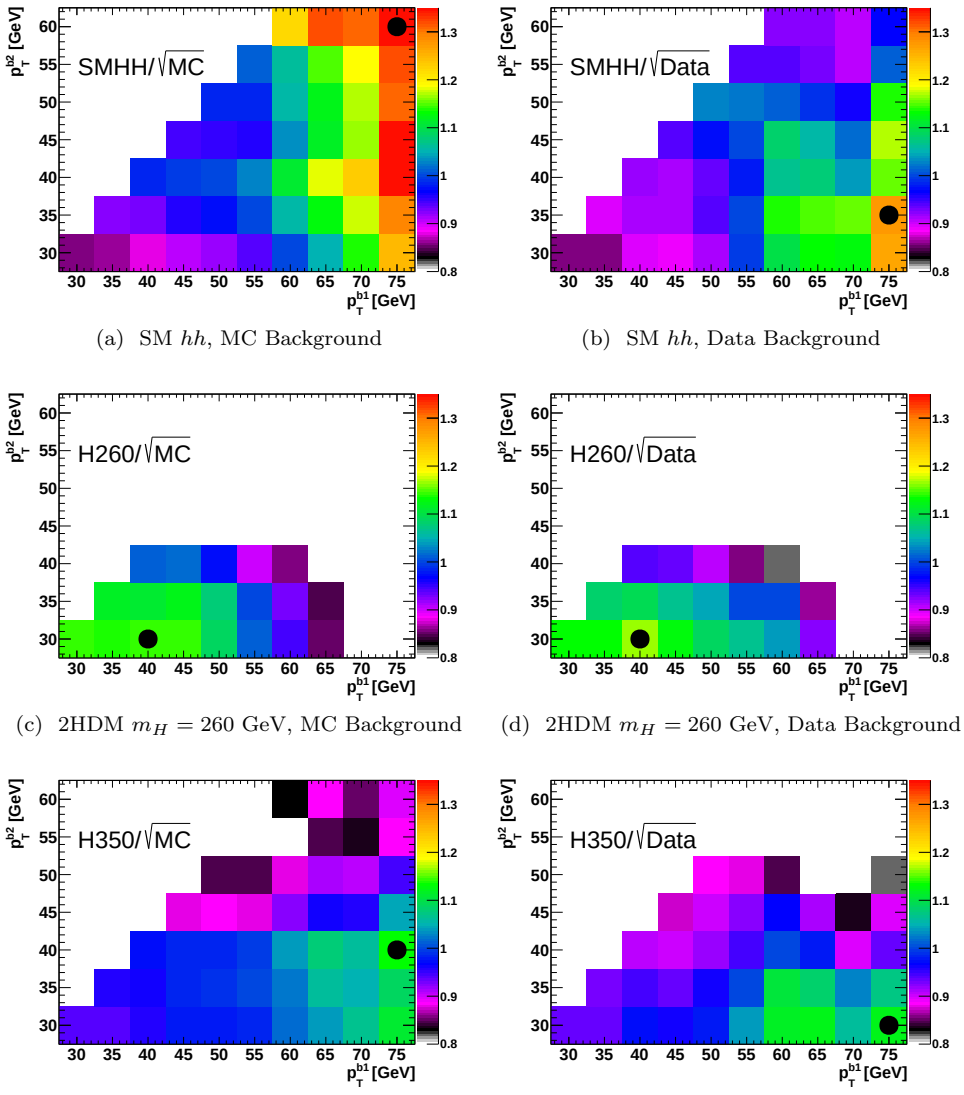


(a) Constrained $m^{\gamma\gamma bj}$ (b) $p_{\mathrm{T}}^{\gamma\gamma bj}$ (c) $p_{\mathrm{T}}^{\gamma\gamma}$

(d) $\Delta R(\gamma,\gamma)$ (e) p_{T}^{b1} (f) p_{T}^{b2}

Figure J.4: The control region with a single tag is presented from data and smeared MC, using a k-factor of 2.

K Cuts Optimization for Higgs Pair Production



(a) SM hh, MC Background

(b) SM hh, Data Background

(c) 2HDM $m_H = 260$ GeV, MC Background

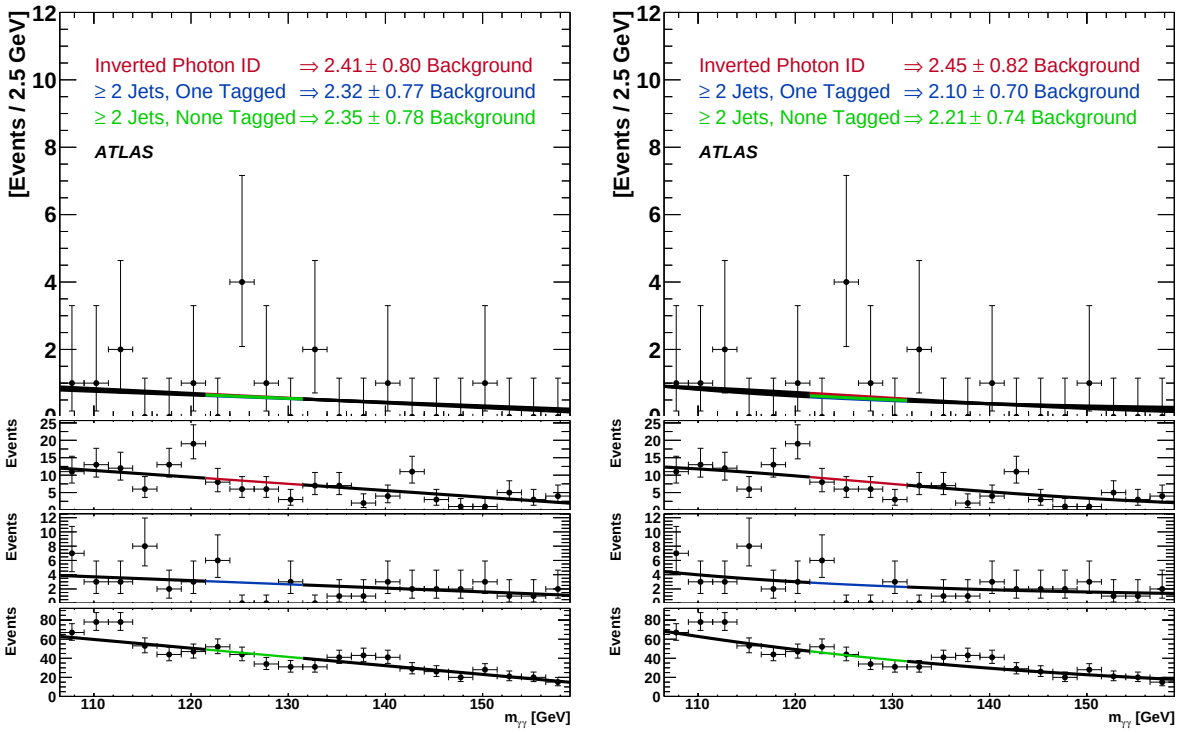
(d) 2HDM $m_H = 260$ GeV, Data Background

(e) 2HDM $m_H = 350$ GeV, MC Background

(f) 2HDM $m_H = 350$ GeV, Data Background

Figure K.1: The optimization of the cuts on the leading and subleading b-jet p_{T}s is shown for 2HDM signal models $m_H = 260$ and 350 GeV, as well as SM hh production. The $m_H = 300$ GeV mass point is shown at Figure 8.3 of the main text. The background response is estimated either from un-tagged (but kinematically identical) jets in data, or from $\gamma\gamma b\bar{b}$ MC. The figure of merit $S/\sqrt{B}$ is shown as a function of the leading and subleading b p_{T} cut. Each plot is scaled so that the value at 55/35 GeV (the cut used) is at 1.0. The black dot in each plot shows the bin with the highest significance.

L Additional Control Regions and Fits for Higgs Pair Production



(a) Linear Fit

(b) Second-order exponential fit

Figure L.1: Alternative background shapes are used to fit the $m^{\gamma\gamma}$ sidebands of the pair production signal region: (a) a linear fit, and (b) a secon-order exponential (e^{ax+bx^2}). The signal region $|m^{\gamma\gamma} - m_h| > 5$ GeV is excluded from the fit. The fit shapes are constrained by the control regions, and the normalization is constrained by the signal region, as shown.

Bibliography

[1] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321–323. 1, 2.2

[2] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132–133. 1, 2.2

[3] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508–509. 1, 2.2

[4] The UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys.Lett. **B**122 (1983) 103–116. 1

[5] The UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys.Lett. **B**122 (1983) 476–485. 1

[6] The UA1 Collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around* 95 GeV/c^2 *at the CERN SPS Collider*, Phys.Lett. **B**126 (1983) 398–410. 1

[7] The UA2 Collaboration, P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\bar{p}p$ *Collider*, Phys.Lett. **B**129 (1983) 130–140. 1

[8] M. Jacob, ed., *Large Hadron Collider in the LEP Tunnel*, CERN. CERN, Geneva, September, 1984. http://cds.cern.ch/record/154938. 1

[9] The ATLAS Collaboration, *ATLAS: letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*, tech. rep., Geneva, 1992. https://cds.cern.ch/record/291061. 1

[10] The CMS Collaboration, *Letter of intent: by the CMS Collaboration for a general purpose detector at LHC*, Tech. Rep. CERN-LHCC-92-003. LHCC-I-1, CERN, Geneva, 1992. http://cds.cern.ch/record/290808. 1

[11] M. Bajko et al., *Report of the Task Force on the Incident of 19th September 2008 at the LHC*, Tech. Rep. LHC-PROJECT-Report-1168, CERN, Geneva, Mar, 2009. http://cds.cern.ch/record/1168025/. 1

[12] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **7**16 (2012) no. 1, 1 – 29. http://www.sciencedirect.com/science/article/pii/S037026931200857X. 1, 9

[13] The ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys.Lett. **B**726 (2013) 88–119, arXiv:1307.1427 [hep-ex]. 1, 2.4, 6.3, 1, 8.4.2, 9

[14] The ATLAS Collaboration, *Differential cross sections of the Higgs boson measured in the diphoton decay channel using 8 TeV pp collisions*, Tech. Rep. ATLAS-CONF-2013-072, CERN, Geneva, Jul, 2013. http://cds.cern.ch/record/1562925. 1, 9

[15] The ATLAS Collaboration, *Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector*, arXiv:1406.5053 [hep-ex]. 1, 9

[16] P. Langacker, *The Standard Model and Beyond*. Series in High Energy Physics, Cosmology and Gravitation. Taylor & Francis, Boca Raton, 2010. 2, 2.3

[17] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*. Advanced book classics. Addison-Wesley Publishing Company, Reading, MA, 1995. http://books.google.com/books?id=i35LALN0GosC. 2

[18] T. Hambye and K. Riesselmann, *SM Higgs mass bounds from theory*, arXiv:hep-ph/9708416. 2.3

[19] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, Eur.Phys.J. **C**72 (2012) 2205, arXiv:1209.2716 [hep-ph]. 2.3

[20] The ALEPH, CDF, D0, DELPHI, L3, OPAL, and SLD Collaborations, The LEP Electroweak Working Group, The Tevatron Electroweak Working Group, and The SLD Electroweak and Heavy Flavour Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:1012.2367 [hep-ex]. 2.3

[21] The LHC Higgs Cross Section Working Group Collaboration, LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002 (CERN, Geneva, 2011) , arXiv:1101.0593 [hep-ph]. 2.4

[22] The LHC Higgs Cross Section Working Group Collaboration, LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN-2012-002 (CERN, Geneva, 2012) , arXiv:1201.3084 [hep-ph]. 2.4, 5.2.1

[23] The LHC Higgs Cross Section Working Group Collaboration, LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph]. 2.4, 2.1, 2.4, 8.5

[24] Baglio, J. and Djouadi, A. and Gröber, R. and Mühlleitner, M.M. and Quevillon, J. and others, *The measurement of the Higgs self-coupling at the LHC: theoretical status*, JHEP **1**304 (2013) 151, arXiv:1212.5581 [hep-ph]. 2.1, 2.5.2.2, 8.1.1, 8.5, J.1f, J.1

[25] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **8**8 (2002) 201801, arXiv:hep-ph/0201206 [hep-ph]. 2.4

[26] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. **B**646 (2002) 220–256, arXiv:hep-ph/0207004 [hep-ph]. 2.4

[27] V. Ravindran, J. Smith, and W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. **B**665 (2003) 325–366, arXiv:hep-ph/0302135 [hep-ph]. 2.4

[28] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, Phys. Lett. **B**595 (2004) 432–441, arXiv:hep-ph/0404071 [hep-ph]. 2.4

[29] G. Degrassi and F. Maltoni, *Two-loop electroweak corrections to Higgs production at hadron colliders*, Phys.Lett. **B**600 (2004) 255–260, arXiv:hep-ph/0407249 [hep-ph]. 2.4

[30] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two-loop electroweak corrections to Higgs production in proton-proton collisions*, arXiv:hep-ph/0610033 [hep-ph]. 2.4

[31] C. Anastasiou, R. Boughezal, and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, JHEP **0**904 (2009) 003, arXiv:0811.3458 [hep-ph]. 2.4

[32] D. de Florian and M. Grazzini, *Higgs production through gluon fusion: Updated cross sections at the Tevatron and the LHC*, Phys. Lett. **B**674 (2009) 291–294, arXiv:0901.2427 [hep-ph]. 2.4

[33] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Phys. Lett. **B**670 (2008) 12–17, arXiv:0809.1301 [hep-ph]. 2.4

[34] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC*, Phys. Rev. Lett. **9**9 (2007) 161803, arXiv:0707.0381 [hep-ph]. 2.4

[35] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, Phys. Rev. **D**77 (2008) 013002, arXiv:0710.4749 [hep-ph]. 2.4

[36] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al., *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*, Comput. Phys. Commun. **1**80 (2009) 1661–1670, arXiv:0811.4559 [hep-ph]. 2.4

[37] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, Phys. Rev. Lett. **1**05 (2010) 011801, arXiv:1003.4451 [hep-ph]. 2.4

[38] T. Han and S. Willenbrock, *QCD correction to the $pp \to WH$ and ZH total cross-sections*, Phys. Lett. **B**273 (1991) 167–172. 2.4

[39] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B**579 (2004) 149–156, arXiv:hep-ph/0307206 [hep-ph]. 2.4

[40] M. Ciccolini, S. Dittmaier, and M. Krämer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. **D**68 (2003) 073003, arXiv:hep-ph/0306234 [hep-ph]. 2.4

[41] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, et al., *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **8**7 (2001) 201805, arXiv:hep-ph/0107081 [hep-ph]. 2.4

[42] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, et al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, Nucl. Phys. **B**653 (2003) 151–203, arXiv:hep-ph/0211352 [hep-ph]. 2.4

[43] S. Dawson, L. Orr, L. Reina, and D. Wackeroth, *Next-to-leading order QCD corrections to $pp \to t\bar{t}h$ at the CERN Large Hadron Collider*, Phys. Rev. **D**67 (2003) 071503, arXiv:hep-ph/0211438 [hep-ph]. 2.4

[44] S. Dawson, C. Jackson, L. Orr, L. Reina, and D. Wackeroth, *Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections*, Phys. Rev. **D**68 (2003) 034022, arXiv:hep-ph/0305087 [hep-ph]. 2.4

[45] R. Harlander and M. Wiesemann, *Jet-veto in bottom-quark induced Higgs production at next-to-next-to-leading order*, JHEP **1**204 (2012) 066, arXiv:1111.2182 [hep-ph]. 2.4, 8.3.2

[46] S. Dawson, C. B. Jackson, L. Reina, and D. Wackeroth, *Exclusive Higgs boson production with bottom quarks at hadron colliders*, Phys. Rev. D **6**9 (Apr, 2004) 074027. http://link.aps.org/doi/10.1103/PhysRevD.69.074027. 2.4, 8.3.2

[47] S. Dittmaier, M. Kramer, and M. Spira, *Higgs radiation off bottom quarks at the Tevatron and the CERN LHC*, Phys.Rev. **D**70 (2004) 074010, arXiv:hep-ph/0309204. 2.4, 8.3.2

[48] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A Program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **1**08 (1998) 56–74, arXiv:hep-ph/9704448 [hep-ph]. 2.4

[49] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, *Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons*, Phys. Rev. **D**74 (2006) 013004, arXiv:hep-ph/0604011 [hep-ph]. 2.4

[50] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NNLO computational techniques: The cases $H \to \gamma\gamma$ and $H \to gg$*, Nucl. Phys. **B**811 (2009) 182–273, arXiv:0809.3667 [hep-ph]. 2.4

[51] The Particle Data Group Collaboration, K. Nakamura et al., *Review of particle physics*, J.Phys. **G**37 (2010) 075021. 2.4

[52] N. Craig, J. Galloway, and S. Thomas, *Searching for Signs of the Second Higgs Doublet*, arXiv:1305.2424 [hep-ph]. 2.5.2.1

[53] The ATLAS Collaboration, *Constraints on New Phenomena via Higgs Coupling Measurements with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2014-010, CERN, Geneva, Mar, 2014. http://cds.cern.ch/record/1670531. 2.5.2.1

[54] J. Baglio, O. Eberhardt, U. Nierste, and M. Wiebusch, *Benchmarks for Higgs Pair Production and Heavy Higgs Searches in the Two-Higgs-Doublet Model of Type II*, arXiv:1403.1264 [hep-ph]. 2.5.2.1

[55] Y. Tang, *Implications of LHC Searches for Massive Graviton*, JHEP **1**208 (2012) 078, arXiv:1206.6949 [hep-ph]. 2.5.2.1

[56] K. Cheung, *Phenomenology of radion in Randall-Sundrum scenario*, Phys.Rev. **D**63 (2001) 056007, arXiv:hep-ph/0009232. 2.5.2.1

[57] N. Kumar and S. P. Martin, *LHC search for di-Higgs decays of stoponium and other scalars in events with two photons and two bottom jets*, arXiv:1404.0996 [hep-ph]. 2.5.2.1

[58] M. J. Dolan, C. Englert, and M. Spannowsky, *New Physics in LHC Higgs boson pair production*, Phys.Rev. **D**87 (2013) 055002, arXiv:1210.8166 [hep-ph]. 2.5.2.1

[59] D. Curtin, R. Essig, S. Gori, P. Jaiswal, A. Katz, et al., *Exotic Decays of the 125 GeV Higgs Boson*, arXiv:1312.4992 [hep-ph]. 2.5.2.1

[60] R. Contino, M. Ghezzi, M. Moretti, G. Panico, F. Piccinini, et al., *Anomalous Couplings in Double Higgs Production*, JHEP **1**208 (2012) 154, arXiv:1205.5444 [hep-ph]. 2.5.2.2

[61] G. D. Kribs and A. Martin, *Enhanced di-Higgs Production through Light Colored Scalars*, Phys.Rev. **D**86 (2012) 095023, arXiv:1207.4496 [hep-ph]. 2.5.2.2

[62] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001. 3.1, 3.1

[63] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report: 1. The LHC Main Ring*, CERN-2004-003-V-1 (CERN, Geneva, 2004) . http://cds.cern.ch/record/782076. 3.1

[64] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report: 2. The LHC infrastructure and general services*, CERN-2004-003-V-2 (CERN, Geneva, 2004) . 3.1

[65] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report: 3. The LHC injector chain*, CERN-2004-003-V-3 (CERN, Geneva, 2004) . http://cds.cern.ch/record/823808. 3.1

[66] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. 3.1, 3.2, 3.2, 3.3

[67] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005. 3.1

[68] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002. 3.1

[69] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004. 3.1

[70] The ATLAS Collaboration, *Luminosity Public Results*, http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults, Sep., 2013. 3.1

[71] The ATLAS Collaboration, *Performance of the ATLAS Detector using First Collision Data*, JHEP **1**009 (2010) 056, arXiv:1005.5254 [hep-ex]. 3.2

[72] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512 [hep-ex]. 3.2, 4.1, 4.2.2

[73] The ATLAS Collaboration, *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur.Phys.J. **C**71 (2011) 1593, arXiv:1011.6665 [physics.ins-det]. 3.2

[74] The ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur.Phys.J. **C**70 (2010) 787–821, arXiv:1004.5293 [physics.ins-det]. 3.2.2.1

[75] The ATLAS Collaboration, S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, eds., *ATLAS Inner Detector*, vol. 1 of *Technical Design Report*. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331063. 3.2.2.1

[76] The ATLAS Collaboration, S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, eds., *ATLAS Inner Detector*, vol. 2 of *Technical Design Report*. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331064. 3.2.2.1

[77] The ATLAS Collaboration, *ATLAS pixel detector*. Technical Design Report. CERN, Geneva, 1998. http://cdsweb.cern.ch/record/381263. 3.2.2.1

[78] The ATLAS Pixel Detector Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007. 3.2.2.1

[79] The ATLAS SCT Collaboration, *The barrel modules of the ATLAS semiconductor tracker*, Nucl.Instrum.Meth. **A**568 (2006) 642–671. 3.2.2.1

[80] *The ATLAS semiconductor tracker end-cap module*, Nucl.Instrum.Meth. **A**575 (2007) 353–389. 3.2.2.1

[81] The ATLAS SCT Collaboration, *The Silicon microstrip sensors of the ATLAS semiconductor tracker*, Nucl.Instrum.Meth. **A**578 (2007) 98–118. 3.2.2.1

[82] The ATLAS TRT Collaboration, *The ATLAS TRT barrel detector*, JINST **3** (2008) P02014. 3.2.2.1

[83] The ATLAS TRT Collaboration, *The ATLAS TRT end-cap detectors*, JINST **3** (2008) P10003. 3.2.2.1

[84] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013. 3.2.2.1

[85] P. Cwetanski, F. Dittus, and R. Orava, *Straw Performance Studies and Quality Assurance for the ATLAS Transition Radiation Tracker*. PhD thesis, University of Helsinki, Helsinki, June, 2006. 3.2.2.1

[86] The ATLAS Collaboration, *ATLAS central solenoid*. Technical Design Report. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331067. 3.2.2.1

[87] The ATLAS Collaboration, *ATLAS calorimeter performance*. Technical Design Report. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331059. 3.2.2.2

[88] K. Kleinknecht, *Detectors for Particle Radiation*. Cambridge University Press, Cambridge, U.K., second ed., 1998. http://books.google.com/books?id=RI05DW7rtmMC. 3.2.2.2

[89] The ATLAS Collaboration, *ATLAS liquid-argon calorimeter*. Technical Design Report. CERN, Geneva, 1996. http://cdsweb.cern.ch/record/331061. 3.2.2.2, 3.2.2.2

[90] The ATLAS Collaboration, *ATLAS tile calorimeter*. Technical Design Report. CERN, Geneva, 1996. https://cdsweb.cern.ch/record/331062. 3.2.2.2

[91] The ATLAS Collaboration, *ATLAS muon spectrometer*. Technical Design Report. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331068. 3.2.2.3

[92] The ATLAS Collaboration, J. P. Badiou, J. Beltramelli, J. M. Baze, and J. Belorgey, eds., *ATLAS barrel toroid*. Technical Design Report. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331065. 3.2.2.3

[93] The ATLAS Collaboration, *ATLAS end-cap toroids*. Technical Design Report. CERN, Geneva, 1997. http://cdsweb.cern.ch/record/331066. 3.2.2.3

[94] F. Bauer et al., *Construction and test of MDT chambers for the ATLAS muon spectrometer*, Nucl.Instrum.Meth. **A**461 (2001) 17–20. 3.2.2.3

[95] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE Trans.Nucl.Sci. **56** (2009) 1568–1574. 3.2.2.3

[96] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon trigger in the barrel of the ATLAS experiment*, Nucl.Phys.Proc.Suppl. **158** (2006) 11–15. 3.2.2.3

[97] S. Majewski, G. Charpak, A. Breskin, and G. Mikenberg, *A Thin Multiwire Chamber Operating in the High Multiplication Mode*, Nucl.Instrum.Meth. **217** (1983) 265–271. 3.2.2.3

[98] The ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys.Rev. **D**83 (2011) 052005, arXiv:1012.4389 [hep-ex]. 4

[99] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002, CERN, Geneva, Apr, 2008. http://cds.cern.ch/record/1099735. 4.1, 4.4.1

[100] D. Banfi, L. Carminati, and L. Mandelli, *Calibration of the ATLAS electromagnetic calorimeter using calibration hits*, Tech. Rep. ATL-LARG-PUB-2007-012. ATL-COM-LARG-2007-007, CERN, Geneva, Jul, 2007. 4.2.2

[101] E. Nebot and J. Del Peso, *Electron Energy Reconstruction for the EM Endcap Calorimeter using Calibration Hits*, Tech. Rep. ATL-LARG-INT-2008-001. ATL-COM-LARG-2008-005, CERN, Geneva, May, 2008. 4.2.2

[102] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak Group, and SLD Heavy Flavour Group, *Precision electroweak measurements on the Z resonance*, Phys.Rept. **4**27 (2006) 257–454, arXiv:hep-ex/0509008 [hep-ex]. 4.2.3

[103] The ATLAS Collaboration, *Measurement of the WW cross section in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS detector and limits on anomalous gauge couplings*, Phys.Lett. **B**712 (2012) 289–308, arXiv:1203.6232 [hep-ex]. 4.3.4.1

[104] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **A**CAT (2007) 040, arXiv:physics/0703039 [PHYSICS]. 4.3.4.2

[105] G. Cowan, *Statistical Data Analysis*. Oxford University Press, Oxford, 1998. 4.3.4.2

[106] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 1995. 4.3.4.2

[107] The ATLAS Collaboration Collaboration, ATLAS Collaboration, *Search for the Standard Model Higgs boson in the diphoton decay channel with 4.9 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV with ATLAS*, PRL (2012) , arXiv:1202.1414 [hep-ex]. 4.3.4.3

[108] K. Liu, Y. Liu, and G. Marchiori, *Measurement of the identification efficiency of isolated prompt photons using the matrix method and 4.9 fb^{-1} of ATLAS data*, Tech. Rep. ATL-COM-PHYS-2012-242, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1428909. 4.3.4.4

[109] The ATLAS Collaboration, *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb1 of pp collision data collected in 2011*, Tech. Rep. ATLAS-CONF-2012-123, CERN, Geneva, Aug, 2012. http://cds.cern.ch/record/1473426. 4.3.4.4

[110] M. Hance and H. H. Williams, *Measurement of Inclusive Isolated Prompt Photon Production in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector*. PhD thesis, University of Pennsylvania., Philadelphia, July, 2011. http://cds.cern.ch/record/1367057. 4.4.1

[111] The ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-PHYS-PUB-2012-003, 2012. http://cds.cern.ch/record/1474107. 5.2.1, 7.5, 8.1.1

[112] The ATLAS Collaboration, *Further ATLAS tunes of PYTHIA6 and Pythia 8*, Tech. Rep. ATL-PHYS-PUB-2011-014, 2011. http://cds.cern.ch/record/1400677. 5.2.1, 7.5, 8.1.1

[113] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0**711 (2007) 070, arXiv:0709.2092 [hep-ph]. 5.2.1

[114] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0**904 (2009) 002, arXiv:0812.0578 [hep-ph]. 5.2.1

[115] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **1**002 (2010) 037, arXiv:0911.5299 [hep-ph]. 5.2.1

[116] S. Catani, D. de Florian, M. Grazzini, and P. Nason, *Soft gluon resummation for Higgs boson production at hadron colliders*, JHEP **0**307 (2003) 028, arXiv:hep-ph/0306211 [hep-ph]. 5.2.1

[117] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, JHEP **1**111 (2011) 064, arXiv:1109.2109 [hep-ph]. 5.2.1

[118] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to 2\gamma$, $H \to WW \to \ell\nu\ell\nu$ and $H|toZZ \to 4\ell$ decay modes*, JHEP **1**206 (2012) 132, arXiv:1203.6321 [hep-ph]. 5.2.1

[119] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur.Phys.J. **C**70 (2010) 823–874, arXiv:1005.4568 [physics.ins-det]. 5.2.1

[120] The GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A**506 (2003) 250–303. 5.2.1

[121] L. J. Dixon and M. S. Siu, *Resonance continuum interference in the diphoton Higgs signal at the LHC*, Phys.Rev.Lett. **9**0 (2003) 252001, arXiv:hep-ph/0302233 [hep-ph]. 5.2.1.1

[122] D. A. Dicus and W. W. Repko, *Dalitz Decay $H \to f\bar{f}\gamma$ as a Background for $H \to \gamma\,\gamma$*, arXiv:1402.5317 [hep-ph]. 5.2.1.1

[123] A. Firan and R. Stroynowski, *Internal conversions in Higgs decays to two photons*, Phys.Rev. **D**76 (2007) 057301, arXiv:0704.3987 [hep-ph]. 5.2.1.1

[124] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **1**78 (2008) 852–867, arXiv:0710.3820 [hep-ph]. 5.2.1.1

[125] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **0**902 (2009) 007, arXiv:0811.4622 [hep-ph]. 5.2.2

[126] M. Vesterinen and T. Wyatt, *A Novel Technique for Studying the Z Boson Transverse Momentum Distribution at Hadron Colliders*, Nucl.Instrum.Meth. **A**602 (2009) 432–437, arXiv:0807.4956 [hep-ex]. 5.4

[127] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. http://cds.cern.ch/record/296752. 5.5.1

[128] The 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 (2013) 2518, arXiv:1302.4393 [hep-ex]. 5.5.1, 8.5

[129] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C**72 (2012) 1849, arXiv:1110.1530 [hep-ex]. 5.5.1

[130] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, Phys.Rev. **D**57 (1998) 3873–3889, arXiv:physics/9711021 [physics.data-an]. 5.6

[131] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011) no. 2, 1–19. http://dx.doi.org/10.1140/epjc/s10052-011-1554-0. 5.6, 5.6

[132] J. Ellis, D. S. Hwang, V. Sanz, and T. You, *A Fast Track towards the 'Higgs' Spin and Parity*, JHEP **1211** (2012) 134, arXiv:1208.6002 [hep-ph]. 5.7

[133] J. Ellis, V. Sanz, and T. You, *Associated Production Evidence against Higgs Impostors and Anomalous Couplings*, Eur.Phys.J. **C73** (2013) 2507, arXiv:1303.0208 [hep-ph]. 5.7

[134] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph]. 6.1

[135] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys.Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph]. 6.1

[136] The ATLAS Collaboration, *Search for ttH production in the $H \to \gamma\gamma$ channel at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2013-080, CERN, Geneva, Jul, 2013. http://cds.cern.ch/record/1564319. 5, 6.2.3.3, 8.5

[137] The ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2012-101, CERN, Geneva, Jul, 2012. 6.2.2.1, 6.2.2.2

[138] The ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, Tech. Rep. ATLAS-CONF-2013-082, CERN, Geneva, Aug, 2013. 6.2.2.1

[139] The ATLAS Collaboration, *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar, 2014. http://cds.cern.ch/record/1670012. 6.3.1

[140] The ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys.Lett. **B726** (2013) 120–144, arXiv:1307.1432 [hep-ex]. 6.3.1, 7.6

[141] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl.Instrum.Meth. **A362** (1995) 487–498. 7.3.2

[142] T. Adye, *Unfolding algorithms and tests using RooUnfold*, ArXiv e-prints (May, 2011) , arXiv:1105.1160 [physics.data-an]. 7.3.2

[143] The ATLAS Collaboration, *Measurement of hard double-parton interactions in $W \to \ell\bar{\nu}_\ell$ + 2 jet events at $\sqrt{s} = 7$ TeV with the ATLAS detector*, New J.Phys. **15** (2013) 033038, arXiv:1301.6872 [hep-ex]. 7.4.2, 7.6

[144] The ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, arXiv:1406.3827 [hep-ex]. 8

[145] K. Hamilton, P. Nason, and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, JHEP **1210** (2012) 155, arXiv:1206.3572 [hep-ph]. 7.5

[146] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph]. 7.5

[147] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. **C**63 (2009) 189–285, arXiv:0901.0002 [hep-ph]. 7.5

[148] R. Boughezal, F. Caola, K. Melnikov, F. Petriello, and M. Schulze, *Higgs boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, JHEP **1**306 (2013) 072, arXiv:1302.6216 [hep-ph]. 7.6

[149] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, Phys. Rev. D **1**6 (1977) 2219–2225. 7.6

[150] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, *Higgs and Z-boson production with a jet veto*, Phys.Rev.Lett. **1**09 (2012) 202001, arXiv:1206.4998 [hep-ph]. 7.6

[151] I. W. Stewart, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Jet p_T resummation in Higgs production at NNLL'+NNLO*, Phys. Rev. D **8**9 (Mar, 2014) 054001, arXiv:1307.1808 [hep-ph]. 7.6

[152] J. R. Andersen, K. Arnold, and D. Zeppenfeld, *Azimuthal Angle Correlations for Higgs Boson plus Multi-Jet Events*, JHEP **1**006 (2010) 091, arXiv:1001.3822 [hep-ph]. 7.6

[153] U. Baur, T. Plehn, and D. L. Rainwater, *Probing the Higgs self-coupling at hadron colliders using rare decays*, Phys.Rev. **D**69 (2004) 053004, arXiv:hep-ph/0310056. 8, 8.1.1, J.1b, J.1

[154] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **1**106 (2011) 128, arXiv:1106.0522 [hep-ph]. 8.1.1

[155] F. Maltoni, Higgs Pair Production, `https://cp3.irmp.ucl.ac.be/projects/madgraph/wiki/HiggsPairProduction`, Dec., 2013. 8.1.1, 8.1.2

[156] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0**207 (2002) 012, arXiv:hep-ph/0201195 [hep-ph]. 8.1.1

[157] O. Antipin and T. Hapola, *ED Tools*, `http://cp3-origins.dk/research/units/ed-tools`, Dec., 2013. 8.1.2

[158] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, *Warped Gravitons at the LHC and Beyond*, Phys.Rev. **D**76 (2007) 036006, arXiv:hep-ph/0701186 [hep-ph]. 8.1.2

[159] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys.Rev.Lett. **8**3 (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph]. 8.1.2

[160] The ATLAS Collaboration, *Measurement of the cross section for the production of a W boson in association with b-jets in pp collisions at with the ATLAS detector* , Physics Letters B **7**07 (2012) no. 5, 418 – 437. `http://www.sciencedirect.com/science/article/pii/S0370269311015036`. 8.1.3

[161] T. Gehrmann, N. Greiner, and G. Heinrich, *Diphoton + jets at NLO*, arXiv:1311.4754 [hep-ph]. 8.1.3

[162] The ATLAS Collaboration, A. Bundock, G. Chiodini, P. Davison, D. Hellmich, G. Hesketh, S. Hillert, E. Jansen, P. Laycock, N. Orlando, M. Sanders, S. Spagnolo, M. Tibbetts, and D. Vladoiu, *Measurement of Zb and Zbb cross sections with 4.6 fb^{-1} of 7 TeV ATLAS data*, Tech. Rep. ATL-COM-PHYS-2012-1504, CERN, Geneva, Oct, 2012. `https://cds.cern.ch/record/1484442`. Internal ATLAS. 8.1.3

[163] The ATLAS Collaboration, *A study of heavy flavor quarks produced in association with top quark pairs at sqrt(s) = 7 TeV using the ATLAS detector*, arXiv:1304.6386 [hep-ex]. 8.1.3

[164] The ATLAS Collaboration, *Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2013-079, CERN, Geneva, Jul, 2013. http://cds.cern.ch/record/1563235. 8.2.1.1

[165] The ATLAS Collaboration, *b-Jet Tagging Efficiency Calibration using the System8 Method*, Tech. Rep. ATLAS-CONF-2011-143, CERN, Geneva, Oct, 2011. http://cds.cern.ch/record/1386703. 8.2.2

[166] The ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011. http://cds.cern.ch/record/1369219. 8.2.2

[167] The ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014. http://cds.cern.ch/record/1664335. 8.2.2

[168] J. M. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, *Higgs-Boson production in association with a single bottom quark*, Phys.Rev. **D**67 (2003) 095002, arXiv:hep-ph/0204093. 8.3.2

[169] The ATLAS Collaboration, *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1**306 (2013) 084, arXiv:1302.2929. 8.5

[170] The ATLAS Collaboration, *A study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, arXiv:1304.6386. 8.5

[171] A. L. Read, *Presentation of search results: The CL(s) technique*, J.Phys. **G**28 (2002) 2693–2704. 8.6

[172] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, Eur.Phys.J. **C**70 (2010) 525–530, arXiv:1005.1891 [physics.data-an]. 8.6

[173] R. Harlander, M. Mhlleitner, J. Rathsman, M. Spira, and O. Stl, *Interim recommendations for the evaluation of Higgs production cross sections and branching ratios at the LHC in the Two-Higgs-Doublet Model*, arXiv:1312.5571 [hep-ph]. 8.6

[174] The CMS Collaboration, *Search for stop and higgsino production using diphoton Higgs boson decays*, arXiv:1312.3310 [hep-ex]. 9

[175] The CMS Collaboration, *Search for resonant HH production in 2gamma+2b channel*, Tech. Rep. CMS-PAS-HIG-13-032, CERN, Geneva, 2014. http://cds.cern.ch/record/1697512. 9

[176] L. J. Dixon and Y. Li, *Bounding the Higgs Boson Width Through Interferometry*, Phys.Rev.Lett. **1**11 (2013) 111802, arXiv:1305.3854 [hep-ph]. 9

[177] F. Caola and K. Melnikov, *Constraining the Higgs boson width with ZZ production at the LHC*, Phys.Rev. **D**88 (2013) 054024, arXiv:1307.4935 [hep-ph]. 9

[178] The ATLAS Collaboration, *Projections for measurements of Higgs boson cross sections, branching ratios and coupling parameters with the ATLAS detector at a HL-LHC*, Tech. Rep. ATL-PHYS-PUB-2013-014, CERN, Geneva, Oct, 2013. http://cds.cern.ch/record/1611186. 9