Title
Discovery of the Higgs Boson Decaying to Two Photons /

Author
Palmer, Christopher Allan

Publication Date
2014

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Discovery of the Higgs Boson Decaying to Two Photons

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Physics

by

Christopher Allan Palmer

Committee in charge:

Professor James Branson, Chair
Professor Michael Anderson
Professor Farhat Beg
Professor Sanjoy Dasgupta
Professor Oleg Shpyrko
Professor Frank Würthwein

2014

The dissertation of Christopher Allan Palmer is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2014

DEDICATION

To my grandmother, my mother, and Kate
for their support, inspiration, encouragement, and love.

EPIGRAPH

We are at the very beginning of time for the human race.
It is not unreasonable that we grapple with problems.
But there are tens of thousands of years in the future.
Our responsibility is to do what we can, learn what we
can, improve the solutions, and pass them on.

Richard Feynman, 1955

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The colossal scale of the LHC reflects the grand amount of human investment in its development and construction. Moreover, the wealth of knowledge garnered over the past several generations has brought us to today's studies of LHC data. We are all indebted to the scientists of the past century, but immediately I am thoroughly in debt to the LHC and the CMS Collaboration. Being part of CMS has allowed me to touch a piece of history and for that I will be forever grateful.

I am extremely thankful to Jim Branson and Marco Pieri whose insights and dogged work ethic are illustrations of their intellects and loyal natures. They truly do not expect more from others than they expect from themselves. Matteo Sani and Andre Holzner are wonderful collaborators from whom I have learned much. Finally, Elizabeth Dusinberre Sudano, Sean Simon, Ryan Kelley and Matt LeBourgeois provided me with support very early when I struggled the most. Thank you all for your help and your patience.

To my many collaborators at CERN I wish to express great thanks. Without you this analysis would not have been the success that it was. To Matthew Kenzie, Martina Malberti, Nick Wardle, Chiara Rovelli, Fabian Stöckli, Doug Berry, Xiaohang Quan, Rishi Patel, Roberta Volpe, Shervin Nourbakhsh, Mingming Yang, Josh Bendavid, Francesco Micheli, Aniello Spiezia, Pasquale Musella, Chris Seez, Paolo Meridiani, Daniele Del Re, Susan Gascon-Shotkin, Nancy Marinelli, Vladimir Rekovic, and Serguei Ganjour, thank you.

My time in San Diego could not have been the same without my cohort and especially my housemates: Alex, Ben, Colin and John. All the late nights struggling with problem sets, talking about politics and playing poker will stay with me forever.

Geneva attracts a wonderful cross section of people from all over the world. The friendships I made there are very meaningful to me because each one of them has a curious mind and a wealth of perspective. Abdul, Amy, Bene, Carly, Charlie, Dan, Dave, Gloria, Jonny, Lidia and Rohan, your company and conversation stimulate my mind and make me laugh. Though our intersection in Geneva may have been brief, our friendships will be enduring.

Today it is difficult to imagine life without my lovely Kate. Meeting in Geneva was a chance event that bound us together almost instantly. Your love, support, and patience kept me alive and (mostly) sane during the craziness of the discovery and while writing this document. I will endeavour to show you my earnest love and thankfulness everyday for the rest of our lives.

Finally, I thank my family. My grandparents went to great lengths to ensure my education was not disrupted even during very turbulent and emotionally difficult times. I can never repay them but I hope to live a life worthy of their struggle and sacrifice. Mom, you taught me as an article of faith that hard work opens the doors of opportunity and that doing my homework could take me anywhere in the world. It turns out you were right and those larger-than-life promises have all become reality. Without you I would not be here, but because of you I dreamed without borders. Thank you for your inspiration.

Chapter 4, in part, is a reprint of the material as it appears in "Search for the standard model Higgs boson decaying into two photons in *pp* collisions at $\sqrt{s} = 7$ TeV", Physics Letters B 710 (Feb, 2012) 403-425. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Physics Letters B 716 (2012) 30-61. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in "Search for a fermiophobic Higgs boson in *pp* collisions at $\sqrt{s} = 7$ TeV", JHEP 1209 (2012) 111. The dissertation author was the primary investigator and author of this paper.

Chapter 5, Chapter 6, and Chapter 7, in full, has been prepared for publication (expected to EPJC): "Observation of the diphoton decay of the Higgs boson and measurement of its properties". The dissertation author was the primary investigator and author of this paper.

Chapter 8 is a reprint of the material as it appears in "Search for the standard model Higgs boson in the dimuon decay channel in *pp* collisions at $\sqrt{s} =$ 7 and 8 TeV", *CMS Physics Analysis Summary* CMS-PAS-HIG-13-007 (2013). The dissertation author was the primary investigator and author of this paper.

VITA

2007	B. S. in Astronomy and Mathematics *magna cum laude*, University of Southern California
2008	M. S. in Physics, University of California, San Diego
2014	Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

CMS Collaboration, “Observation of the diphoton decay of the Higgs boson and measurement of its properties”, Preparing for publication, expected in *EPJC*.

CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Physics Letters B* 716 (2012) 3061.

CMS Collaboration, “Search for a fermiophobic Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* 1209 (2012) 111.

CMS Collaboration, “Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s} = 7$ TeV”, *Physics Letters B* 710 (Feb, 2012) 403-425.

ABSTRACT OF THE DISSERTATION

Discovery of the Higgs Boson Decaying to Two Photons

by

Christopher Allan Palmer

Doctor of Philosophy in Physics

University of California, San Diego, 2014

Professor James Branson, Chair

The Standard Model (SM) of particle physics fundamentally relies on the existence of the Higgs boson. This massive particle is a relic of the underlying and hidden Higgs field, whose transformation into the Higgs boson provides mass to weak bosons and all massive fermions in the SM. This particle has been long-sought and finally using data from proton-proton collisions at the LHC, CMS and ATLAS experiments have discovered a particle which is compatible with the SM Higgs boson. Presented here is the develeopment of one of the discovery channels, H → $\gamma\gamma$, and the final H → $\gamma\gamma$ analysis and results using the full luminosity of the LHC Run 1 dataset ~25 fb^{-1} at 7 or 8 TeV center of mass energy. The observed (expected) significance of this di-photon excess in the final analysis is 5.7σ (5.2σ) with a measured signal strength of $\sigma/\sigma_{SM} = 1.14^{+0.26}_{-0.23}$. The mass of this Higgs boson is not predicted by the SM. Using the H → $\gamma\gamma$ channel, $\mathrm{M_H}$ is measured to be $124.70^{+0.35}_{-0.34}$ GeV. Other measured quantities are presented including the signal strength modifiers of different production mechanisms and spin hypothesis tests between spin-0 and spin-2 models. Searches for this Higgs boson decaying to the di-muon and di-electron states are presented. No excess is observed and universal lepton decays of this particle are therefore ruled out, supporting the SM Higgs boson interpretation. In addition, relevant searches, observations and measurements from CMS that characterize this particle are presented.

Chapter 1

Introduction

The advancements in quantum field theory during the mid-twentieth century led to the formulation of the Standard Model (SM) of particle physics. Experimental physicists have been observing interactions predicted by the SM with stunning levels of agreement since its inception. At the heart of the SM is the Higgs field. Via the Brout-Englert-Higgs (BEH) Mechanism the Higgs field becomes hidden leaving the massive Higgs boson to mediate its interactions. This Higgs boson has been sought and not found until very recently.

This particle is infrequently produced and its mass is not predicted but its existence is essential to the SM and many of its observed predictions. Primarily the breaking of the Higgs field symmetry provides essential and observed mass to the weak gauge bosons, and furthermore gives mass to fermions via Yukawa coupling. The Large Hadron Collider (LHC) is not the first collider to attempt to produce the Higgs bosons with high energy collisions of particles, but it is the latest and most powerful.

With the very high energy proton-proton collisions from the LHC, CMS and ATLAS experiments have been able to observe a new particle which is consistent in measurable couplings, signal strength in detectable final states and spin-parity with a SM Higgs boson. This search will be motivated with a description of the SM in Chapter 2. Afterward the LHC accelerator and CMS detectors will be described in Chapter 3.

The evolution of the $\mathrm{H} \to \gamma\gamma$ analysis within CMS follows in Chapter 4.

This analysis has been the author's primary work since late 2010 and its development from that time until present (Summer 2014) will be narrated. The author's contributions to four publications (three regarding the SM Higgs boson searches/observation and one on the exclusion of the Fermiophobic Higgs boson) will be highlighted here as well.

The following three chapters describe the final CMS Run 1 data analysis of the $\mathrm{H} \rightarrow \gamma\gamma$ observation. Chapter 5 discusses the preliminary inputs to the analysis (features of the data, description of the simulation, reference list of variables, etc.). Two analyses are fully described in Chapter 6: the simple Cut-based Analysis and the more optimized MultiVariate Analysis. Finally in Chapter 7, results will be shown including the more than 5σ observation of the Higgs boson decaying to two photons.

In addition to contributing to the search in the $\mathrm{H} \rightarrow \gamma\gamma$ channel, the author has also fully performed searches in $\mu^+\mu^-$ and $\mathrm{e}^+\mathrm{e}^-$ channels for the discovered particle. Chapter 8 briefly reviews this analysis. Finally, to highlight the Higgs bosons' measured properties, other relevant CMS single channel analyses and results will be described in Chapter 9.

Chapter 2

Standard Model and the Higgs Boson

The Standard Model (SM) of particle physics is a quantum field theory of electromagnetic, strong and weak interactions communicated among all known particles. The only known fundamental interaction that is not part of this theory is gravity. However, gravity is very weak in comparison to the other interactions until the density of matter becomes very large. Moreover, high energy physics has not reached such a density (or energy scale) and so the SM has been incredibly successful in predicting numerous physical results without its inclusion.

Like all quantum field theories, the SM respects Einstein's theory of special relativity and treats individual particles as local excitations of fields which permeate the entirety of space. Particle interactions are mediated by gauge particles, which are spin-1 bosons carrying energy and momentum between spin-$\frac{1}{2}$ fermions (or composite states built from these fermions). The spin-$\frac{1}{2}$ fermions are divided into quarks and leptons, where the primary distinction between the two is that quarks interact via the strong nuclear force and leptons do not.

Today we know three generations of quarks and leptons, which have the same quantum numbers except for mass, by which quarks and leptons are (separately) organized from least (1st generation) to greatest (3rd generation). There are two types of quarks: up-type (u, c, t) with electric charge $+\frac{2}{3}$ and down-type (d, s, b) electric change of $-\frac{1}{3}$. There are two types of leptons: electrons, muons, and tauons with charge -1 and associated electromagnetically chargeless neutrinos which are considered massless in the SM although observations of neutrino oscillation predicts non-zero mass differences [3].

There are four classes of spin-1 gauge bosons which mediate the SM forces. The photon mediates the electromagnetic force. Gluons mediate the strong force. Unlike gluons and photons, Z and $W^{\pm}$ bosons are massive, and they mediate the weak force. Finally, the long-sought Higgs boson, whose discovery and characterization is the topic of this thesis, is a spin-0 particle which provides mass for weak bosons as well as the spin-$\frac{1}{2}$ fermions via electroweak spontaneous symmetry breaking. A summary of these observed particles and some of their properties is given in Figure 2.1.

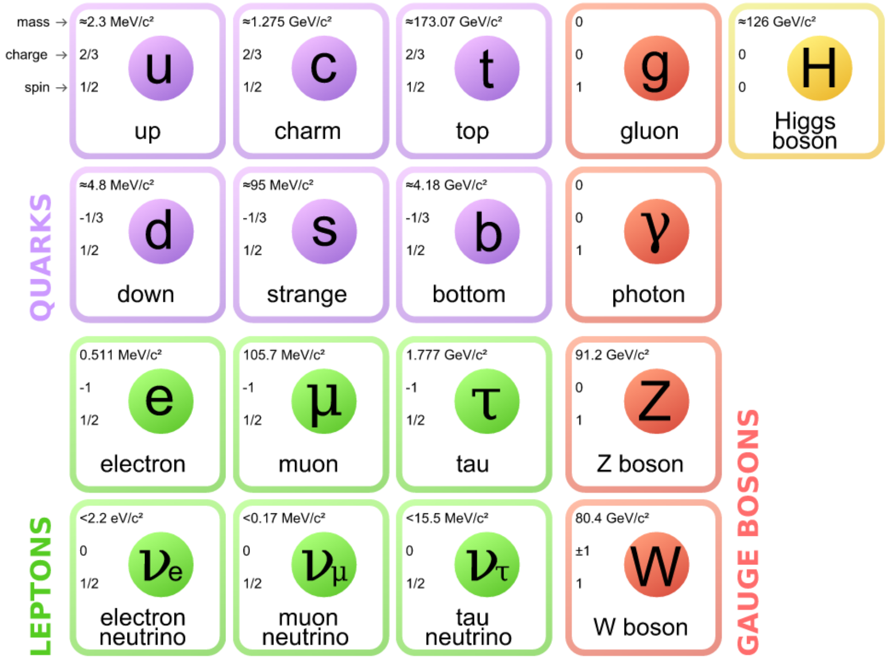


Figure 2.1: Standard Model overview. Above is the summary of all currently observed elementary particles and the gauge bosons which mediate their interactions. The current best measurements from the Particle Data Group for mass as well as charge and intrinsic spin are listed within each particle's box [4].

Quantum field theories can be formulated by understanding the underlying symmetry of an interaction and formulating a locally gauge invariant Lagrangian. Local gauge invariance is simply a formal requirement that the physics embedded in the Lagrangian will not change under transformations of the underlying symmetry even if fields are transformed separately at each point in space yet with the same "type" of transformation.

From such a Lagrangian the equations of motion and Feynman rules can be derived [5]. For quantum electrodynamics (QED) and quantum chromodynmaics (QCD) fermion mass terms were incorporated into the Lagrangians from their inception. Massless spin-$\frac{1}{2}$ particles Lagrangians will have the following, natural form:

$$\mathscr{L}_{QFT} = \bar{\psi}\left(i\gamma^{\mu}D_{\mu}\right)\psi - \frac{1}{4}K_{\mu\nu}K^{\mu\nu} \tag{2.0.0.1}$$

where D_{μ} is the covariant derivative, $K_{\mu\nu}$ is the field tensor, and γ^{μ} are the Dirac gamma matrices. Covariant derivatives have the form $D_{\mu} = \partial_{\mu} - i\sum_{j}^{\text{Generators}} c_j G_{j,\mu}$, where $c_j G_{j,\mu}$ is a constant times a generator of the symmetry group. This is a necessary extension of a normal derivative to maintain local gauge invariance.

The first term describes the kinetic energy of the fermions and their interaction with the gauge boson. However, the covariant derivative contains a term with the gauge field and that term describes the interaction between the fermions and the field. The second term is the kinetic term of the field. Both terms must be invariant under the local transformations of the Lagrangian's symmetry. This will become more clear in the following section (2.1) describing Quantum Electrodynamics.

2.1 Quantum Electrodynamics

QED is the quantum field theory which describes how particles with electric charge interact via the transfer of photons (QED's gauge boson). The symmetry of QED is U(1), and therefore, QED's Lagrangian must be invariant under local gauge transformations of U(1). That is, the spin-$\frac{1}{2}$ fermions and the gauge boson must transform as:

$$\begin{aligned} \psi_{\mu}(x) &\rightarrow e^{-i\Gamma(x)}\psi_{\mu}(x) \\ A_{\mu}(x) &\rightarrow A_{\mu}(x) - \frac{1}{e}\partial_{\mu}\Gamma(x) \end{aligned} \tag{2.1.0.2}$$

The Lagrangian can be constructed beginning with Dirac's equation and modifying the derivative such that it satisfies the invariance requirement above. In order to maintain local gauge invariance, the derivative δ_μ must be replaced by the covariant derivative D_μ, which has explicit dependence on the gauge field, A_μ. The QED Lagrangian is:

$$\mathscr{L}_{QED} = \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1.0.3}$$

where the covariant derivative and field tensor are $D_\mu = \partial_\mu - ieA_\mu$ and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. e is the magnitude of the charge of the electron. Due to the introduction of the covariant derivative, the Lagrangian has a term for the coupling of the A_μ and ψ fields which describes the interaction of charged particles and the photon.

Feynman rules can be derived using the Lagrangian above. Using these rules Feynman diagrams can be easily transformed into matrix elements, which are in turn used to calculate interaction cross sections and decay rates. Depicted in Figure 2.2 is the Feynman diagram of the QED interaction vertex, which is derived from the interaction term, $\bar{\psi}\left(i\gamma^\mu\left(-ieA_\mu\right)\right)\psi$. For each QED interaction vertex in a Feynman diagram there is a factor of $ie\gamma^\mu$ in the matrix element. The unitless fine structure constant, $\alpha = \frac{e^2}{4\pi}$, is the conventional and natural replacement for e in computed cross section calculations.

In general to determine the exact probability of any interaction, all Feynman diagrams with the same initial and final states must be summed in the matrix element. However, in QED the diagrams with the fewest number of vertices contribute the most to matrix element. Every QED vertex in the Feynman diagram contributes a factor of $\alpha \approx \frac{1}{137}$. Thus, more complex diagrams with many vertices will contribute less in general than simpler diagrams to the total cross section or decay rate because they will have more factors of α. QED is perturbative in α because higher-order diagrams need only to be taken into account for increased precision.

QED has been stunningly successful in accurately and very precisely predicting physical results. One of the most well measured predictions is the deviation of the anomalous magnetic moment of the electron from the value of two, predicted

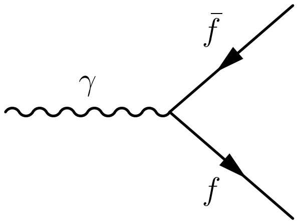


Figure 2.2: Feynman diagrams for quantum electrodynamics. In matrix element calculations the factor of $ie\gamma^{\mu}$ is used wherever there is a vertex.

by Dirac's equation. The most recent calculation to $O(\alpha^4)$ [6] matches with experimental measurements [7] up to 10 signficant figures!

2.2 Quantum Chromodynamics

While quarks do have electric charge their dominant interaction is via the strong force. Quarks carry color charges (red, blue, greeen) and do not typically exist in free single quake states. "Colorless" bound states of quarks (hadrons) consist of two quarks (mesons), three quarks (baryons) or four quarks (the newly confirmed tetraquark state) [8, 9]. Quarks within mesons have opposite color charges while quarks within hadrons have one of each three color charges for a net "white" charge. Tetraquarks have two sets of color/anti-color pairs of quarks. The nonexistence of single free quarks is known as confinement and stems from the fact that the attractive potential between quarks increases at "large" distances.

The color symmetry of QCD is SU(3), and therefore there are eight generators of transformations as opposed to one for QED. Each of these independent generators corresponds to a gauge boson, G_{μ}^{a} where $a \in \{1, 2, ..., 8\}$, whose inclusion into the covariant derivative maintains invariance under SU(3) transformations. These gauge bosons are gluons. The quarks and gluons are required to transform in the following way in order for the forthcoming QCD Lagrangian to be invariant under SU(3) symmetry:

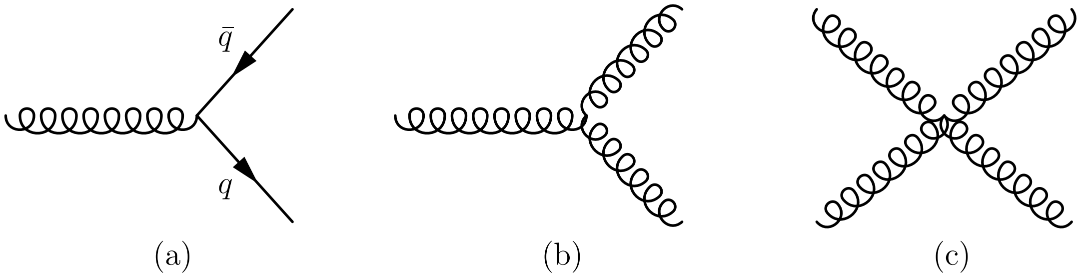
$$\begin{aligned} q_i\left(x\right) &\rightarrow e^{i\alpha_a(x)T_a} q_i\left(x\right) \\ G_{\mu}^{a}\left(x\right) &\rightarrow G_{\mu}^{a}\left(x\right) - \frac{1}{g}\partial_{\mu}\alpha^{a}\left(x\right) + f_{abc}\alpha_b\left(x\right)G_{\mu}^{c}\left(x\right) \end{aligned} \tag{2.2.0.4}$$

$\alpha_a(x)$ are the eight components of the local gauge transformations. T_a are the generators of the SU(3) rotations. f_{abc} are the SU(3) structure constants, and g is the strong coupling constant. Comparing the transformation of the gluons with the photon of QED in Equation 2.1.0.2, there is an additional term of $f_{abc}\alpha_b(x)\,G^c_\mu(x)$. This term is necessary to maintain gauge invariance in the presence of non-Abelian gauge transformations ($[T_a, T_b] = i f_{abc} T_c$).

The QCD Lagrangian which respects the SU(3) transformations above is:

$$\mathscr{L}_{QCD} = \bar{q}\left(i\gamma^\mu D_\mu - m\right) q - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \qquad (2.2.0.5)$$

where $D_\mu = \partial_\mu + igT_a G^a_\mu$ and $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f_{abc} G^b_\mu G^c_\nu$. The extra term in $G^a_{\mu\nu}$ also had to be introduced because of the non-Abelian nature of SU(3). The expansion of $G^a_{\mu\nu}G^{\mu\nu}_a$ reveals QCD interactions beyond quark-gluon interaction, which arises from the covariant derivative just as in QED (depicted in the Feynman diagram in Figure 2.3a). Indeed, the gluons interact with other gluons via three and four gluon vertices (as in the Feynman diagrams in Figure 2.3b and Figure 2.3c). These interactions have no QED anology and arise purely from the addition of terms required due to the non-Abelian nature of SU(3).



(a) (b) (c)

Figure 2.3: Feynman diagrams for quantum chromodynamics. Depicted here are the interaction vertices of QCD. (a) is the analogy of the QED vertex, while (b) and (c) depict the three and four gluon interactions which arise from the non-Abelian nature of SU(3).

In following QED as an anology if g, or better $\alpha_s = \frac{g^2}{4\pi}$, is small, then QCD will be perturbative. However, as mentioned previously, the attractive force between quarks is strong at "large" distances leading to quark confinement–that is, bound states called hadrons. These large distance, low momentum transfer, Q, interactions among quarks have large α_s. Indeed, α_s is dependent on Q and when the exchange becomes larger (i.e. interaction occurs at shorter distances) $\alpha_s(Q)$ becomes smaller. Therefore QCD is perturbative at large momentum exchanges. This effect is called *asymptotic freedom* and it is due to the so-called *running* of the coupling constant, which is merely the rather strong dependence of the strong coupling on energy scale. The Particle Data Group has summarized experimental measurements of the α_s in Figure 2.4 [4].

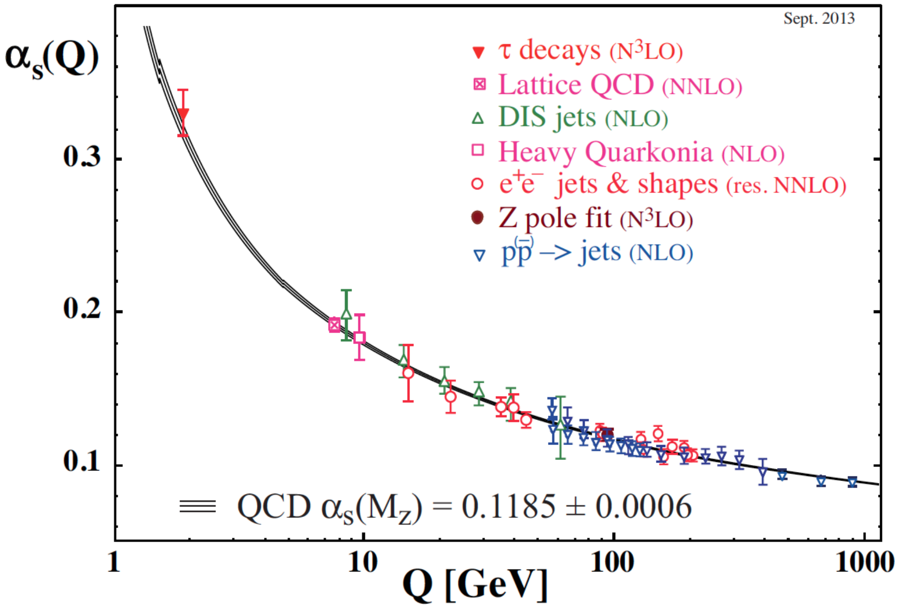


Figure 2.4: The running of the strong coupling. Numerous experimental results are overlaid on the curve which depicts the functional form of the coupling as a function of energy exchange [4].

2.3 Electroweak Theory

Weak decays of nuclei were discovered at the end of the 19th century. Nearly 40 years after the first observations of β^- decay Enrico Fermi postulated an effective field theory utilizing a four particle interaction and assuming the existence of the neutrino to successfully describe this decay [10]. Fermi's framework also describes the only later discovered muon decay effectively as well. However, even though Fermi's formulation was an excellent first framework for weak interactions, it could not fully describe weak interactions at high energies.

Given the extraordinary success of QED in the 1950s, efforts to find symmetries and write gauge invariant Lagrangians were quite expansive and fruitful throughout the 1950's and 1960's. Several different theoretical techniques needed to be developed before a final combination could be made to build a coherent theory for weak interactions.

Moreover, this coherent theory required the unification of the electromagnetic and weak interactions into a single theory requiring massive spin-1 bosons to mediate. This was suggested by Salam and Ward and independently by Glashow in the late 1950's [11,12]. However, the inclusion of massive bosons into any gauge theory is very problematic. If mass terms are naively included for the gauge particles, then local invariance is not possible. Even ignoring this issue, the theory will be unrenormalizable because the propagator goes like $i\frac{q_\mu q_\nu}{q^2 M^2}$ in the high momentum limit and different cut-offs are needed for different orders of diagrams. It was not until 1964 when the Brout-Englert-Higgs (BEH) Mechanism made use of spontaneous symmetry breaking and yet avoided any prediction of unobserved massless (Nambu-Goldstone) bosons [13–15] that it was possible to include massive gauge bosons seamlessly into the unified electroweak gauge theory. Weinberg made the proper adjustments to the previous electroweak theories utilizing the BEH mechanism in 1967 [16].

2.3.1 Spontaneous Symmetry Breaking

In the late 1950's the topic of spontaneous symmetry breaking was hotly debated in the context of inclusion into quantum field theories. Spontaneous symmetry breaking arises from fields whose Lagrangians are invariant under a symmetry but whose description near minima do not respect that symmetry. To make this clear observe the following potential of a complex scalar field, ϕ:

$$\begin{aligned} \mathscr{L} &= -\partial^\mu \phi^\dagger \partial_\mu \phi \ - V(\phi) \\ V(\phi) &= \mu^2 \phi^\dagger \phi + \frac{1}{4}\lambda \left(\phi^\dagger \phi\right)^2 \end{aligned} \tag{2.3.1.1}$$

This Lagrangian is invariant under U(1) symmetry, i.e. under the following transformations of the field $\phi\left(x\right)$:

$$\phi\left(x\right) \rightarrow e^{-i\alpha}\phi\left(x\right) \tag{2.3.1.2}$$

Observe in Figure 2.5 the potential of the Lagrangian, $V\left(\phi\right)$ where $\lambda > 0$ and crucially $\mu^2 < 0$, is plotted. It is easily observed that the minima of $V\left(\phi\right)$ lie on a circle, and therefore the ground states of the potential can be parametrized as $\frac{1}{\sqrt{2}}ve^{i\theta}$ where θ is arbitrary and v is the *vacuum expection value*, which is $v = \sqrt{\frac{-4\mu^2}{\lambda}}$

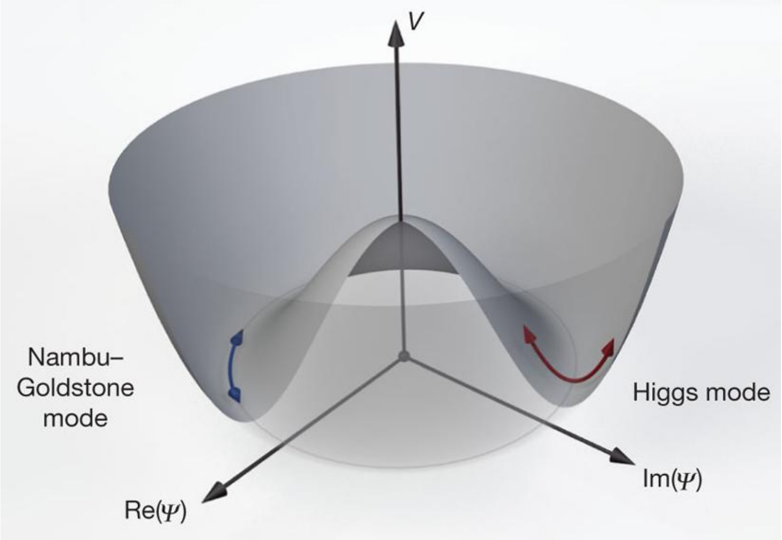


Figure 2.5: Higgs field potential. This is a depiction of a potential which allows spontaneous symmetry breaking. The symmetry of the potential is not a symmetry in the ground state [17].

In the following, the field phi is expanded around the true (stable) vacuum $\phi = \frac{1}{\sqrt{2}}\nu e^{i\theta}$ rather than $\phi = 0$. Any ground state may be chosen (which is how and why the symmetry is broken) and so $\theta = 0$ is a convenient choice. Any perturbation from this state can be parametrized as a combination of the following two types. The first type of perturbation is in the direction of the continuous minima and the other is out of the minimum. This convenient parameterization is:

$$\langle 0|\,\phi\left(x\right)|0\rangle = \frac{1}{\sqrt{2}}\left(v + \rho\left(x\right)\right)e^{i\chi\left(x\right)} \tag{2.3.1.3}$$

Finally evaluating the Lagrangian of Equation 2.3.1.1 as a function of ρ and χ instead of ϕ a Lagrangian which does not respect the U(1) symmetry is revealed:

$$\mathscr{L}' = -\frac{1}{2}\partial^{\mu}\rho\partial_{\mu}\rho - \frac{1}{2}\left(v+\rho\right)\partial^{\mu}\chi\partial_{\mu}\chi - \frac{1}{16}\lambda\rho^4 - \frac{1}{2}\sqrt{-\lambda\mu^2}\rho^3 + \mu^2\rho^2 + \frac{\mu^4}{\lambda} \quad (2.3.1.4)$$

ρ can be associated to a massive boson with $\frac{1}{2}M_{\rho}^2 = -\mu^2$, which will later be associated to a *Higgs boson*. However, given that there is only a kinetic term for χ, it can be identified as a massless state, which is known as a Nambu-Goldstone boson (NGB). $\mathscr{L}'$ of Equation 2.3.1.4 is not locally gauge invariant in terms of ρ as $\mathscr{L}$ of Equation 2.3.1.1 is with respect to ϕ. This is what it means for the symmetry to be "broken" in the ground-state.

The existence of these massless NGBs was well-known. Indeed it was proven–and conjectured several years beforehand–that if any symmetry was spontaneously broken that for each degree of freedom there must be a NGB [18]. The prediction of such particles was experimentally very problematic given the lack of observation. Therefore, before the introduction of the BEH mechanism, the assumed existence of massless bosons as a direct and unavoidable consequence of spontaneous symmetry breaking made its incorporation into a fundamental theory of particle physics appear very unlikely.

2.3.2 Brout-Englert-Higgs Mechanism

Schwinger postulated in 1962 that Goldstone's theorem may not strictly hold in all cases, and indeed he asked the very suggestive question "would the absence of a massless particle imply the existence of a stable, unit spin particle of nonzero mass?" [19]. This is indeed the solution to the puzzle of the missing NGBs. It was not until the summer of 1964 that Brout and Englert [13], Peter Higgs [14], and Guralnik, Hagen and Kibble [15] each independently submitted and published papers stating in detail how to do so.

The BEH mechanism is introduced by adding the kinetic term of a gauge field to a Lagrangian like Equation 2.3.1.1 and ensuring local gauge invariance by replacing the derivative with a covariant derivative. Such a Lagrangian can be written as:

$$\mathscr{L} = -\left(D^{\mu}\phi\right)^{\dagger} D_{\mu}\phi - \mu^{2}\phi^{\dagger}\phi - \frac{1}{4}\lambda\left(\phi^{\dagger}\phi\right)^{2} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.3.2.1}$$

where the covariant derivative is $D_{\mu} = \partial_{\mu} - igA_{\mu}$, the field tensor is $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$, and g is the coupling constant.

The kinetic term ($\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$) of the field A_{μ} is locally gauge invariant and independent of ϕ (and so not very interesting). However, the kinetic term of ϕ, ($-\left(D^{\mu}\phi\right)^{\dagger} D_{\mu}\phi$), now has dependence on A_{μ} via D_{μ}. Evaluating the kinetic term when ϕ is near a ground state reveals the key results of the BEH mechanism. However, first the NGB, $\chi\left(x\right)$, must be addressed. With the gauge field, A^{μ}, comes one degree of gauge freedom. The *unitary gauge* uses this freedom to cancel $\chi\left(x\right)$ of Equation 2.3.1.3. ϕ can be parametrized as follows In this gauge:

$$\langle 0|\,\phi\left(x\right)|0\rangle = \frac{1}{\sqrt{2}}\left(v + \rho\left(x\right)\right) \tag{2.3.2.2}$$

Placing Equation 2.3.2.2 into the kinetic term yields:

$$\begin{aligned}\left(D^{\mu}\phi\right)^{\dagger} D_{\mu}\phi &= \frac{1}{2}\left(\partial^{\mu}\rho + ig\left(v+\rho\right)A^{\mu}\right)\left(\partial_{\mu}\rho - ig\left(v+\rho\right)A_{\mu}\right)\\ &= \frac{1}{2}\left(\partial^{\mu}\rho\partial_{\mu}\rho\right) + \frac{1}{2}g^{2}\left(v+\rho\right)^{2}A^{\mu}A_{\mu}\end{aligned} \tag{2.3.2.3}$$

Expanding the last term yields a mass term for A^{μ} where $M_A = gv$.

To summarize, the unitary gauge illuminates that the BEH mechanism breaks the group symmetry of ϕ in the ground state of ϕ without creating massless bosons but allows gauge bosons to acquire mass.

2.3.3 Glashow-Salam-Weinberg Electroweak Theory

The BEH mechanism allows the breaking of the symmetry in electroweak theory. The symmetry of electroweak theory is SU(2)$_L$×U(1)$_Y$ where L denotes the symmetry of left-handed weak isospin doublets and Y denotes the symmetry of weak hypercharge singlets. Their formal definition and relation to the electromagnetic charge will be described in Section 2.3.4. Note that parity is observably violated in the lepton sector and therefore, left and right-handed particles are treated separately. The unbroken Lagrangian respecting these symmetries is:

$$\begin{aligned}
\mathscr{L} &= -\left(D^{\mu}\phi\right)^{\dagger} D_{\mu}\phi - \mu^{2}\phi^{\dagger}\phi - \frac{1}{4}\lambda\left(\phi^{\dagger}\phi\right)^{2} - \frac{1}{4}F^{a\mu\nu}F^{a}_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\
D_{\mu} &= \partial_{\mu} - i[g_2 W^{a}_{\mu}T^{a} + g_1 B_{\mu}Y]
\end{aligned} \tag{2.3.3.1}$$

where T^a are the generators of SU(2), $F^{a}_{\mu\nu} = \partial_{\mu}W^{a}_{\nu} - \partial_{\nu}W^{a}_{\mu} + g_2[\vec{W} \times \vec{W}]^{a}$ and $B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}$. Choosing a 2×2 representation of the SU(2) generators (half Pauli matrices), ϕ can be formed as a complex scalar doublet.

There are four scalar components to the *Higgs scalar*. Three degrees of gauge freedom are used to eliminate NGBs and create a convenient form of the Higgs scalar field near its ground state:

$$\phi\left(x\right) = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_0\left(x\right) + i\phi_1\left(x\right) \\ \phi_2\left(x\right) + i\phi_3\left(x\right)\end{pmatrix} \rightarrow \left\langle 0\right|\phi\left(x\right)\left|0\right\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}v + H\left(x\right) \\ 0\end{pmatrix} \tag{2.3.3.2}$$

where $H\left(x\right)$ is the observable *Higgs boson.* The covariant derivative is now a 2×2 matrix:

$$D_{\mu} = \begin{pmatrix}\partial_{\mu} & 0 \\ 0 & \partial_{\mu}\end{pmatrix} + \frac{g_2}{2}\begin{pmatrix}W^{3}_{\mu} - \frac{g_1}{g_2}B_{\mu} & W^{1}_{\mu} - iW^{2}_{\mu} \\ W^{1}_{\mu} + iW^{2}_{\mu} & -W^{3}_{\mu} - \frac{g_1}{g_2}B_{\mu}\end{pmatrix} \tag{2.3.3.3}$$

The $\vec{W}$ and B fields correspond to the SU(2)$_L$ and U(1)$_Y$ gauge fields respectively. However, they are not mass eigenstates, which are given by the following combinations:

$$\begin{aligned}
W^{\pm}_{\mu} &= \frac{1}{\sqrt{2}}\left(W^{1}_{\mu} \mp iW^{2}_{\mu}\right) \\
Z_{\mu} &= \cos\left(\theta_W\right)W^{3}_{\mu} - \sin\left(\theta_W\right)B_{\mu} \\
A_{\mu} &= \sin\left(\theta_W\right)W^{3}_{\mu} + \cos\left(\theta_W\right)B_{\mu} \\
\theta_W &= \tan^{-1}\left(\frac{g_1}{g_2}\right)
\end{aligned} \tag{2.3.3.4}$$

where θ_W is the weak mixing (Weinberg) angle. The fields above correspond to the W$^{\pm}$, Z bosons and the photon A.

Recall from Equation 2.3.2.3 that the terms from which the gauge bosons derive mass involve only the gauge bosons in D_{μ} and v. Therefore, the following component of the Lagrangian is all that is needed to determine the masses of $W^{\pm}$ and Z:

$$\begin{aligned} &-\frac{1}{8}g_2^2 \begin{pmatrix} v & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\cos(\theta_W)}Z_\mu & \sqrt{2}W_\mu^+ \\ \sqrt{2}W_\mu^- & \dots \end{pmatrix}^2 \begin{pmatrix} v \\ 0 \end{pmatrix} \\ &= -\frac{1}{8}v^2 g_2^2 \left(\frac{1}{\cos(\theta_W)} Z^\mu Z_\mu + 2W^{+\mu}W_\mu^- \right) \\ &= -\frac{1}{2}M_Z^2 Z^\mu Z_\mu - M_W^2 W^{+\mu} W_\mu^- \end{aligned} \tag{2.3.3.5}$$

where $M_W = \frac{1}{2}g_2 v$ and $M_Z = \frac{1}{2\cos(\theta_W)}g_2 v$. The field, A_μ, aquires no mass and is associated to the electromagnetic field.

At the time of Weinberg's seminal paper, *A Model of Leptons*, which first outlined this theory, the existence of massive, charged W bosons had already been nearly accepted, but "the only unequivocal new predictions made by this model have to do with the couplings of the neutral intermediate meson Z_μ" [16]. In 1973 the Gargamelle experiment at CERN discovered weak neutral-current interactions via neutrino scattering [20], which was the first indirect evidence that supported the existence of such a particle. Ten years later UA1 and UA2 discovered these particles directly and measured the masses using leptonic decay products [21–23]. Today's best measurements of the mass of these bosons summarized by the Particle Data Group [4] are:

$$\begin{aligned} M_W &= 80.385 \pm 0.015\text{GeV} \\ M_Z &= 91.1876 \pm 0.0021\text{GeV} \end{aligned} \tag{2.3.3.6}$$

In addition to the non-zero of the W and Z bosons, numerous interactions have been predicted both among the gauge bosons and with spin-$\frac{1}{2}$ fermions. The interactions among the bosons can be determined by further expansion of the Lagrangian in Equation 2.3.3.1 in terms of the ground state of the Higgs field (Equation 2.3.3.2) and the mixed fields (Equation 2.3.3.4). A summary of these interactions with Feynman diagrams is given in Figure 2.6.

These interactions are derived from the following portion of the expanded Lagrangian. The mass terms for the W and Z bosons are included as well for

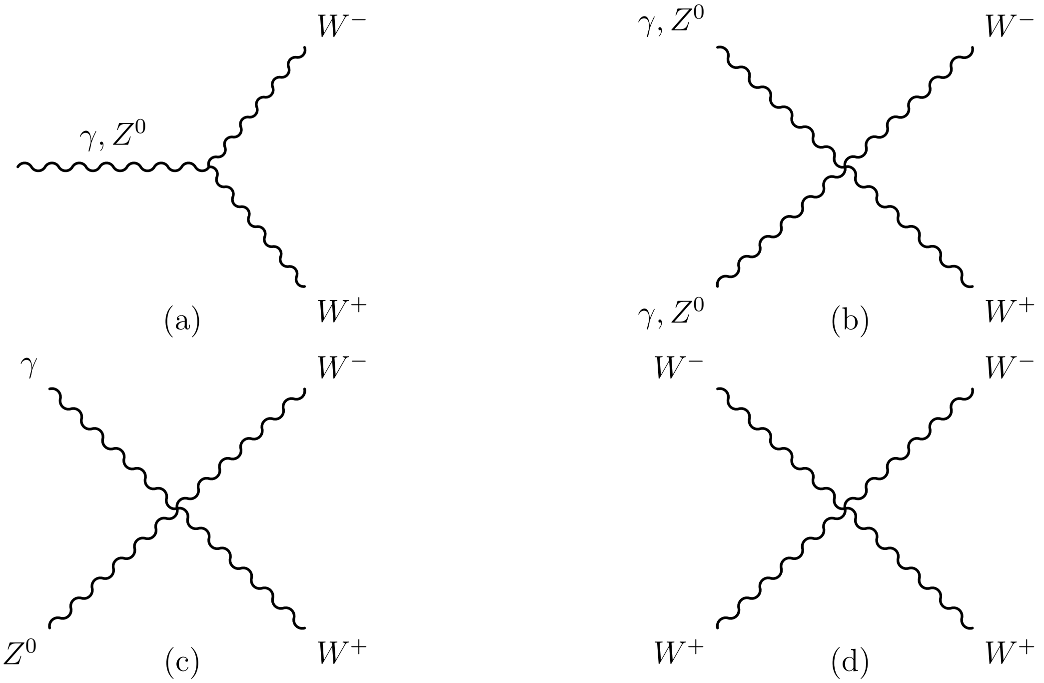


Figure 2.6: Feynman diagrams for interactions among the electroweak gauge bosons.

convenience.

$$\begin{aligned}\mathscr{L}_{Higgs} = &-\left(\frac{1}{2}M_Z^2 Z^\mu Z_\mu + M_W^2 W^{+\mu} W_\mu^-\right)\left(1+v^{-1}H\right)^2 \\ &-\frac{1}{2}\partial^\mu H \partial_\mu H - \frac{1}{2}m_H^2 H^2 - \frac{1}{2}m_H^2 v^{-1} H^3 - \frac{1}{8}m_H^2 v^{-2} H^4\end{aligned} \tag{2.3.3.7}$$

The interactions of the Higgs boson with the gauge bosons and itself are summarized in the Feynman diagrams in Figure 2.7

2.3.4 Electroweak Interactions of Quarks and Leptons

The weak interactions of spin-$\frac{1}{2}$ fermions can be understood in terms of conservation of the weak isospin, T_3, and hypercharge, Y. Weak isospin and hypercharge are related to electric charge in the following way, $Q = T_3 + \frac{1}{2}Y$.

W$^\pm$ bosons have $T_3 = \pm 1$ and $Y = 0$. Left-handed e^-, μ^-, and τ^- have $T_3 = -\frac{1}{2}$ and $Y = -1$. Neutrinos have $T_3 = \frac{1}{2}$ and $Y = -1$. Right-handed fermions have $T_3 = 0$ and so do not interact with Ws. Neutrinos are only left-handed (whereas anti-neutrinos are only right-handed). In the SM neutrinos are massless.

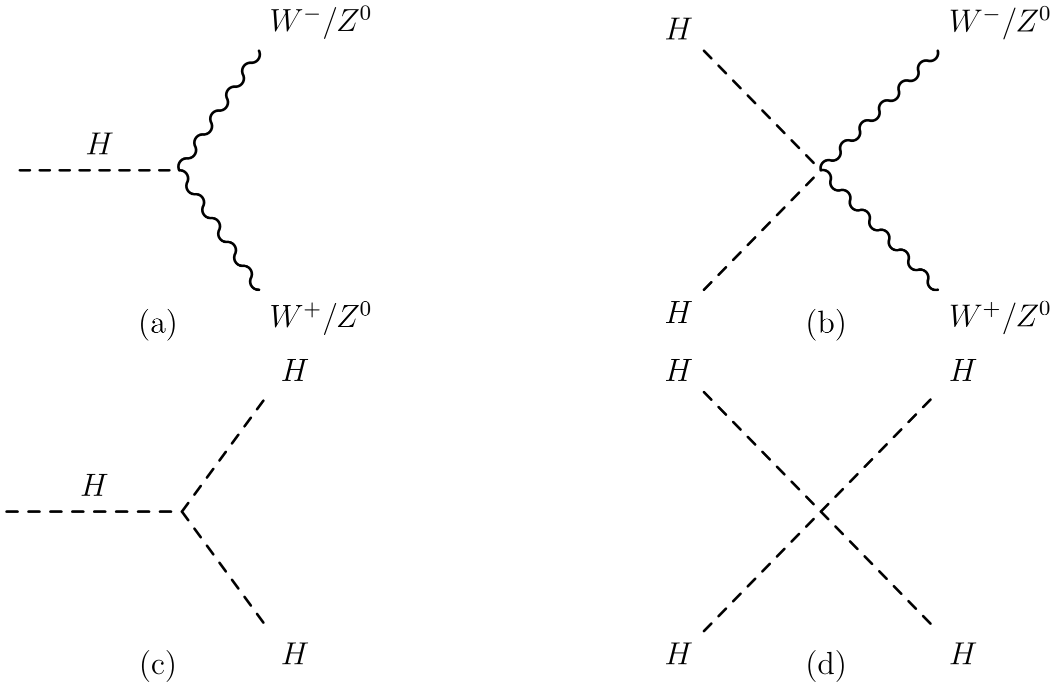


Figure 2.7: Feynman diagrams of Higgs boson interactions with bosons.

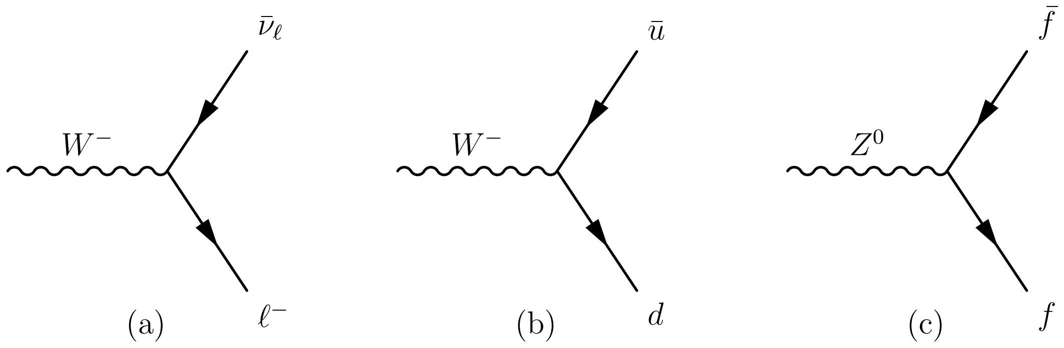
Charged leptons and neutrinos exist in the following left-handed doublets, $\Psi_{\ell L}$, and right-handed singlets, $\psi_{\ell R}$:

$$\begin{aligned} \Psi_{\ell L} &= \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \\ \psi_{\ell R} &= e^-_R, \mu^-_R, \tau^-_R \end{aligned} \tag{2.3.4.1}$$

Likewise for quarks there are left-handed weak isospin doublets, Ψ_{QL}, where up-type quarks have $T_3 = \frac{1}{2}$ and down-type have $T_3 = -\frac{1}{2}$, and right-handed singlets, ψ_{uR} and ψ_{dR}.

$$\begin{aligned} \Psi_{QL} &= \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \\ \psi_{uR} &= u_R, c_R, t_R \\ \psi_{dR} &= d_R, s_R, b_R \end{aligned} \tag{2.3.4.2}$$

The types of interactions of the Z with fermions are very much like photon interactions. However, fermions interacting with Z are not required to have electromagnetic charge which means that Z interacts with neutrinos whereas photons do not. Figure 2.8 summarizes the interactions of Z and W bosons with fermions with Feynman diagrams. Note, the photon couples the same way to left and right handed fermions while the Z does not.



(a) (b) (c)

Figure 2.8: Feynman diagrams for the electroweak interactions of the spin-$\frac{1}{2}$ fermions with the W and Z bosons.

2.4 Yukawa Coupling and Fermion Mass

The SM includes coupling of the Higgs field to massive fermions as a mechanism to produce mass terms for these fermions in the SM Lagrangian. The unbroken Lagrangian with Yukawa coupling between the Higgs field, ϕ, and the fermions is:

$$\mathscr{L}_{Yukawa} = -\overline{\Psi}_{\ell L} G_\ell \phi \psi_{\ell R} - \overline{\Psi}_{\ell Q} G_u \left(i\sigma_2\right) \phi \psi_{uR} - \overline{\Psi}_{\ell Q} G_d \phi \psi_{dR} + \text{hermitian conjugate} \tag{2.4.0.3}$$

where G_ℓ, G_u and G_d are complex (diagonalizable) 3×3 matrices and $(i\sigma_2)$ is used to make the charge-conjugate of ϕ. After diagonalizing the G matrices, breaking the Higgs field symmetry and utilizing the unitary gauge, the Lagrangian becomes:

$$\mathscr{L}_{Yukawa} = -\sum_{f}^{\text{Fermions}} \frac{G_f}{\sqrt{2}} \left(\bar{\psi}_{Lf}\psi_{Rf} + \bar{\psi}_{Rf}\psi_{Lf}\right)(v + H) \tag{2.4.0.4}$$

where the mass of each fermion is $m_f = \frac{G_f v}{\sqrt{2}}$. The coupling of fermions with the Higgs boson is predicted to be proportional to the mass of each fermion (m_f/v to be precise). With the discovery of the Higgs boson at LHC, this prediction can be tested in the coming years with direct measurements of Fermion coupling to the Higgs boson.

2.5 Limits on the SM Higgs boson Mass from Previous Searches

The Large Electron-Positron (LEP) collider at CERN provided electron-positron collisions to four experiments (ALEPH, DELPHI, OPAL, and L3) [24]. During Run 1 these experiments provided very precise results regarding the Z^0 and $W^{+/-}$ boson masses by colliding at or near the center of mass energy ($\sqrt{s}$) of the Z^0. During Run 2 (1995-2000) experiments at LEP undertook extensive SM Higgs boson searches with higher energy ($189 < \sqrt{s} < 209$ GeV) collision data. Each experiment analyzed approximately or greater than 600 pb^{-1} for a combined total of 2.5 fb^{-1} [25]. About 20% of this data was collected in 2000 with $\sqrt{s} \geq 206$ GeV which substantially pushed the boundaries of their Higgs search forward in $\mathrm{M_H}$.

The primary production mode of theoretical Higgs at LEP was via Higgsstrahlung from Z bosons (as in Figure 4.2c except with electron-positron annihilation instead of quark-anti-quark). The most frequent decay (70-76%) of Higgs in LEP's search range was $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$. The analyses were therefore designed to require two b jets and either two charged leptons, missing energy or two additional jets (depending on the Z decay mode). The results of LEP's searches were 95% confidence level $\mathrm{CL_S}$ exclusion limits of a Higgs boson with any mass less than 114.4 GeV. Figure 2.9 highlights the final combined results of LEP experiments.

The CDF and D0 experiments at the Tevatron at Fermilab recorded proton-anti-proton collisions from 1983 to 2011 with up to $\sqrt{s} = 1.96$ TeV. CDF and D0 have extensive physics programs, in which they are prominently known for discovering the top quark, t, in 1995 [26, 27]. They also have performed extensive SM Higgs boson searches with their Run II dataset. During the summer of 2011 when LHC was collecting its first data relevant to its SM Higgs search, CDF and D0 had analyzed 8.2 and 8.6 fb^{-1} of data and set 95% confidence level exclusions limits, ruling out a SM Higgs boson with mass in the range $100 < \mathrm{M_H} < 108$ GeV

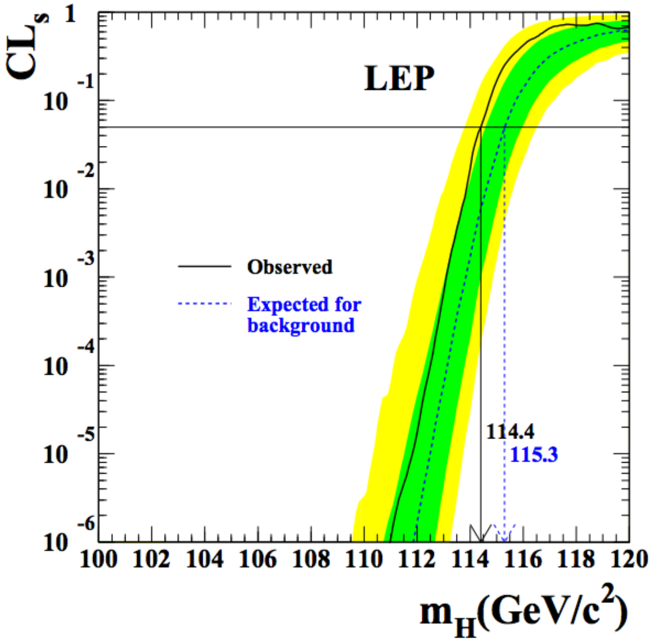


Figure 2.9: LEP Higgs' mass exclusion limits. The final upper limit on the mass of the Higgs boson is 114.4 GeV with an expectation of ruling out any SM Higgs bosons up to 115.3 GeV at the 95% confidence level [25].

and $156 < \mathrm{M_H} < 177$ GeV [28]. This result is shown in Figure 2.10. However, at the time of the Higgs discovery [29] when CDF and D0 combined their full datasets of $\sim$10 fb^{-1} each, Tevatron set final limits to $100 < \mathrm{M_H} < 103$ GeV and $147 < \mathrm{M_H} < 180$ GeV and an excess of approximately 2.5σ level was detected near 120 GeV [30].

In addition to direct searches for the SM Higgs, precision measurements of W width, Γ_W, and t mass, M_t, (along with other SM parameters) allow the SM Higgs mass to be constrained. Electroweak precision measurements are sensitive to the logarithm of the Higgs mass through higher order corrections (i.e. the Higgs appears in loops of diagrams related to these precision measurements). The constraint on the Higgs mass from Electroweak precision data at the time just before LHC started to produce collisions at 7 TeV center of mass energy is shown in Figure 2.11 [31]. Three different likelihood profiles (each based on slightly different parameter inputs) are shown as a function of $\mathrm{M_H}$. All favor a relatively light SM Higgs boson.

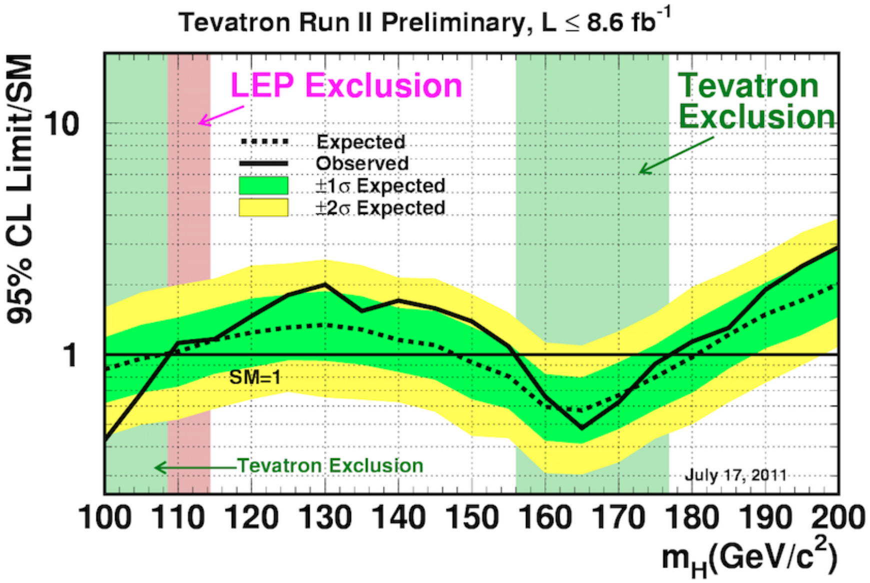


Figure 2.10: Tevatron and LEP exclusion limits in July 2011. At the time of the LHC's first $\sqrt{s} = 7$ TeV collisions this was the state of previous direct searches for SM Higgs bosons.

2.6 The Higgs boson and the Hierarchy Problem

The hierarchy problem arises from the large difference in energy scale between electroweak physics, $O\left(10^2\right)$ GeV, and expected new physics, $O\left(10^{16} - 10^{19}\right)$ GeV (depending on where one expects new physics–from Grand Unified Theory (GUT) scale to Planck Mass (m_{Pl}) scale). In particular, loop corrections to the Higgs propagator introduce corrections to the Higgs mass squared which are proportional to the new physics scale, Λ, squared. If the Higgs boson couples to fermion, f, then the first loop correction is [32]:

$$\mathrm{M}_{\mathrm{H}}^2 \approx \mathrm{M}_{\mathrm{H},0}^2 + \frac{\lambda_f^2}{8\pi^2} N_C^f \Lambda^2 \tag{2.6.0.5}$$

where $\mathrm{M}_{\mathrm{H},0}$ is the bare Higgs mass, λ_f is the Yukawa coupling constant, and N_C^f is the number of color copies of f.

The existence of any term like the second above is rather unsettling and cannot be fully correct. Several theories have been developed to eliminate this ludiciously enormous "correction" to a number which is at least 28 orders of magnitude smaller than the correction itself!

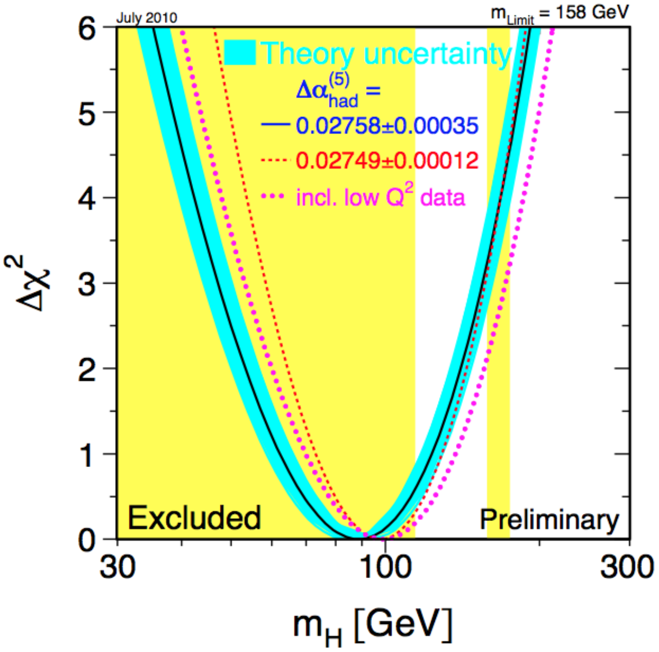


Figure 2.11: Electroweak precision constraints on Higgs' mass. The black line is the $\Delta\chi^2$ as M_H is floated. The blue band is the uncertainty of the fit from theory and the yellow background represents exclusions from LEP and Tevatron searches [31].

One particularly elegant solution is Supersymmetry (SUSY). SUSY postulates superpartner particles for all SM particles. Given that they are unobserved as of now, they must have high mass. Superpartners couple to the Higgs boson as well. Moreover, the propagator diagrams involving superpartners cancel their SM counter parts and so these large correction terms are canceled. There are several other very nice features of this theory. The lightest supersymmetric partner (LSP) is predicted to be stable if R-parity is conserved and only weakly interacting. Therefore, the LSP is a dark matter candidate. In addition, SUSY fits nicely into a paradigm of grand unification of the three interactions (i.e. strong, weak and electromagnetic) at the expected scale of $O\left(10^{16}\right)$ GeV.

Although LHC's Run 1 searches have not found evidence for SUSY, the mass of the Higgs ($M_H \sim 125$ GeV) is favorable to some SUSY models which have not yet been excluded. LHC's Run 2 searches may well discover SUSY.

Another prominent solution is extra dimensions where gravity is much stronger at microscopic scales than the extrapolation from what we know from macroscopic experiments. At small scales, where distances are near or less than the size of the extra dimensions, the $1/r^2$ gravitational force law is modified to a $1/r^{2+\delta}$, where δ is the number of extra dimensions. The Planck mass,m_{Pl}, would be reduced significantly, eliminating the hierarchy problem by making the corrections small. CMS has constrained many such models with Run 1 data.

2.7 Other Limitations of the SM

While the SM can be used to predict types of interactions and cross section of final states with unprecedented accuracy, there are 19 free parameters of the theory, which are not predicted by any underlying theory. The masses of the fermions to the Higgs particle are great examples (and they are 9 of the 19 parameters). The couplings are predicted and are proportional to the mass, but this only answers the question "how do these particles interact with the Higgs?" and not "why do these particles interact in this way?". That is, the question of why the various fermions have different masses is not addressed at all in the SM. It seems only natural to ask if a deeper theory exists from which these constants are derived.

The SM also omits many known phenomena. The most glaring omission is the lack of a quantum field theory (or any other description) of gravity. In any scenario where the density of massive objects is large enough both gravity and the interactions of the SM are simultaneously relevant and must be understood in a single theory. The discovery of such a theory would have implications in the understanding of black holes and of the brief moments immediately after the big bang (at the very least).

In addition to the gaping omission of gravity within its framework, there are particle phenomena which are simply not predicted by the SM. The SM presumes that neutrinos are massless. However, data from the past 10 years has established that neutrinos oscillate between flavors as a function of time [3]. The currently accepted theory is that mass eigenstates of neutrinos are not the same as the flavor eigenstates. The probability of observing flavor j when flavor i was produced depends notably on distance L between the production and detection site and the differences between the squares of the neutrino masses. By combining measurements of such transition probabilities from experiments with different distances L one can infer the differences of the neutrino mass squares. These oscillations imply that at least two of the three flavors of neutrinos are massive.

Astronomy has produced observations of startling and unpredicted phenomena beyond the explanation of the SM as well. Dark matter, which does not interact electromagnetically, constitutes most of the massive universe [33]. Many dark matter candidate particles beyond the SM have been hypothesized and searched for but none have been discovered and verified. Also, the analysis of data from supernova observations in the late 1990s gave the first evidence that the cosmological constant was sufficiently large for the universe to be expanding at an accelerating rate [34]. The unaccounted for energy which pushes the universe apart is known as dark energy. At this time there are several hypothetical models regarding dark energy.

The inadequacy of the SM to explain the universe fully is actually very exciting. The SM was an enormous intellectual breakthrough of the mid-20th century, yet there are many fundamental discoveries still to be made and likely many strange physical phenomena yet to be observed. The author hopes that the future of physical exploration is filled with unexpected observations which will ultimately lead to the deeper understanding of how the universe works at a fundamental level.

Chapter 3

The LHC Accelerator and CMS Detector

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is a 26.7 km circumference particle accelerator. It is the largest ever constructed. Utilizing magnetic fields produced by super-conducting materials, protons (and lead ions) are circulated around this ring at speeds very near the speed of light in two parallel beamlines, in which particles are accelerated in opposite directions. LHC is between 50-150 meters underground in France and Switzerland at the European Organization for Nuclear Physics (CERN). CERN envisioned and commissioned the LHC in 1994. Non-member states, including the United States, made additional commitments to construct hardware for the LHC experiments in 1995 and 1996 [35].

3.1.1 LHC Experiments

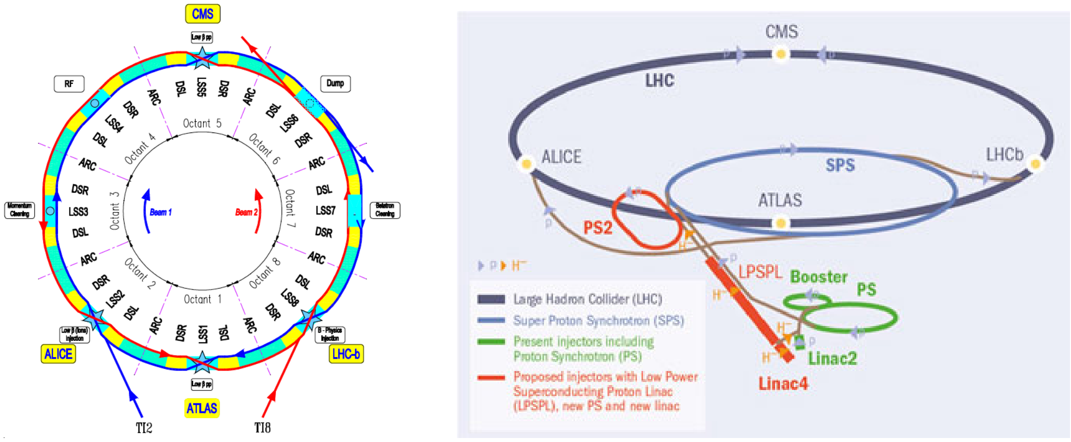
There are seven detectors constructed to record collisions from particles delivered by the LHC. A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS) are two of these detectors and they are general purpose detectors. These detectors are designed to detect many types of particles over a wide energy range. One of the primary focuses of both detectors has been to discover the previously described and long-sought Higgs Boson. Since the discovery [29], elucidation of the newly discovered Higgs Boson's properties have been a key continuing goal.

Large Hadron Collider beauty (LHCb) is a b-physics experiment which is concerned with detecting CP violation in the bottom quark sector. It is a specialized, forward detector. The fourth large experiment is A Large Ion Collider Experiment (ALICE) which is designed to record collision of lead ions in order to investigate quark-gluon plasma. In addition there are three additional smaller experiments. Two are concerned with far forward physics. Large Hadron Collider forward (LHCf) is two detectors located 140 meters from the either side of ATLAS. Its goal is to study neutral pions (π^0) produced by proton-proton collisions. The other is TOTal Elastic and diffractive cross section Measurement (TOTEM) which is located 14 meters from the interaction point at CMS. It is designed to measure total cross section, elastic scattering and diffraction dissociation. Finally, the Monopole and Exotics Detector at the LHC is located in the same cavern as LHCb and its purpose is to search for exotic and rare long-lived particles including magnetic monopoles. Figure 3.1a highlights the configuration of the experiments with respect to the beamline.

3.1.2 LHC Design

The LHC accelerator does not accelerate protons from zero momentum, but rather protons are injected from a series of smaller, ever increasing accelerators. Protons are first injected from Linac 2 (a linear accelerator) into the Proton Synchrotron Booster (PSB) with an energy of 50 MeV. The PSB accelerates protons to 1.4 GeV and injects into the Proton Synchrotron (PS). The PS accelerates the proton to 26 GeV before injecting into the Super Proton Synchrotron (SPS). The SPS accelerates the protons to 450 GeV before injecting into the LHC. This injection chain is diagramed in Figure 3.1b [35].

The LHC accelerator has many parts which have important and necessary roles in accelerating protons up to 4 TeV. There are 1232 superconducting dipole magnets made of Nb-Ti. The dipoles are used to steer the protons around bends in the LHC. Superfluid helium is used to cool the Nb-Ti to 1.9 K in the dipole

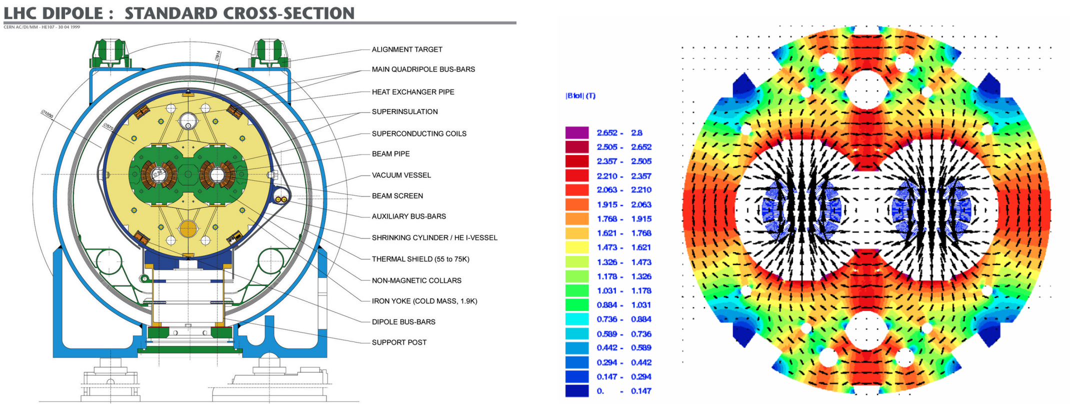


(a) (b)

Figure 3.1: Schematics of LHC beamlines and experiments. Figure (a) is a schematic of the LHC's beams and experiments. Figure (b) depicts the current LHC injection chain (from Linac2 to PSB to PS to SPS and finally to LHC) along with a proposed new injected for SLHC.

magnets. In order to compactly fit both beamlines of protons into the previous LEP tunnel, an impressive dual bore design was implemented. The key feature is that the magnetic dipole fields for the two beams of protons need to be in opposite directions. A cross section of the dipole is depicted in Figure 3.2. The maximum designed field strength inside the dipole is 8.33 T which would properly steer protons with 7 TeV of momentum. However, given the accident on September 19th, 2008 [36], which resulted from a quenched dipole magnet, the runs in 2010 and 2011 accelerated protons only up to 3.5 TeV. This corresponds to a magnetic field of 4.17 T. In 2012, protons were accelerated to 4 TeV, which corresponds to a field strength of 4.76 T.

There are also 78 superconducting quadrupole insertion magnets [38] which are also made of Nb-Ti. These magnets are used to focus the bunches of protons. In most of the quadrupole magnets a weaker target for magnetic field strength allows for a slightly higher temperature of 4.5 K. A diagram of a quadrupole magnet can be found in Figure 3.3 [39]. In addition there approximately 3800 single aperture and 1000 twin aperture corrector magnets that are used for orbital corrections [35].



(a) (b)

Figure 3.2: LHC dipole magnets. In (a) there is a schematic of LHC dipole magnetics used to bend the trajectory of bunches of protons [37]. In (b) there is a schematic of the magnetic flux through the cross section along with magnetic field lines through the cross section. From the field lines in (b) it is clear that the strength of the field is strongest inside the beamline and that the magnetic fields within the two beamlines are opposite.

3.1.3 LHC Beam and Operations

The accelerator is designed to accelerate 2808 bunches of $\sim 1.15 \times 10^{11}$ protons at a center-of-mass (COM) of 14 TeV [35]. However, due to the accident in September 2008 many of the design expectations have yet to been achieved because of a rightfully conservative approach to machine usage.

The primary variable that is used to quantify how many collisions are occurring (i.e. how much data is being collected) is instantaneous luminosity or simply luminosity. It is defined as

$$L = \frac{N^2 k_b f \gamma}{4\pi \varepsilon_n \beta^*} F$$

where

- N is the number of protons per bunch
- k_b the number of bunches
- f the revolution frequency
- γ the relativistic factor

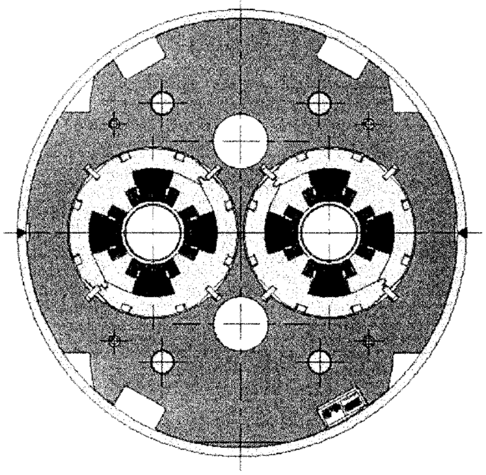


Figure 3.3: LHC quadrupole schematic.

- ε_n the normalized transverse emittance
- β^* the value of the betatron function at the interaction point
- F the reduction factor caused by the crossing angle (~0.9 in the LHC)

Over the course of three years of data-taking (2010-2012) many of these parameters evolved in order to record more data. In Table 3.1 important beam parameters have been highlighted and ranges have been listed to show the development of the LHC's proton beam during 2011 and 2012. One can see in Figure 3.4, where the instantaneous luminosity as a function of time is plotted, that the greatest instantaneous luminosity of $7.67 \times 10^{33} cm^{-2} s^{-1}$ was achieved at CMS in 2012.

Table 3.1: LHC 2011 and 2012 beam parameters [40].

Year	COM	γ	k_b	β^*	Bunch Spacing	Delivered L_{Int}
2011	7 TeV	3.86	228 - 1380	1 - 1.5	25 ns	6.1 fb^{-1}
2012	8 TeV	4.38	624 - 1380	0.6	50 ns	23.3 fb^{-1}

Integrated luminosity is simply the time integral of instantaneous luminosity: $L_{Int} = \int L(t)dt$. The number of events expected for any process is directly proportional to the integrated luminosity and so this quantity is the most relevant measure of how much data the LHC as delivered and CMS has recorded. In Figure 3.5, one can see the amount of data recorded by CMS through 2010, 2011 and 2012.

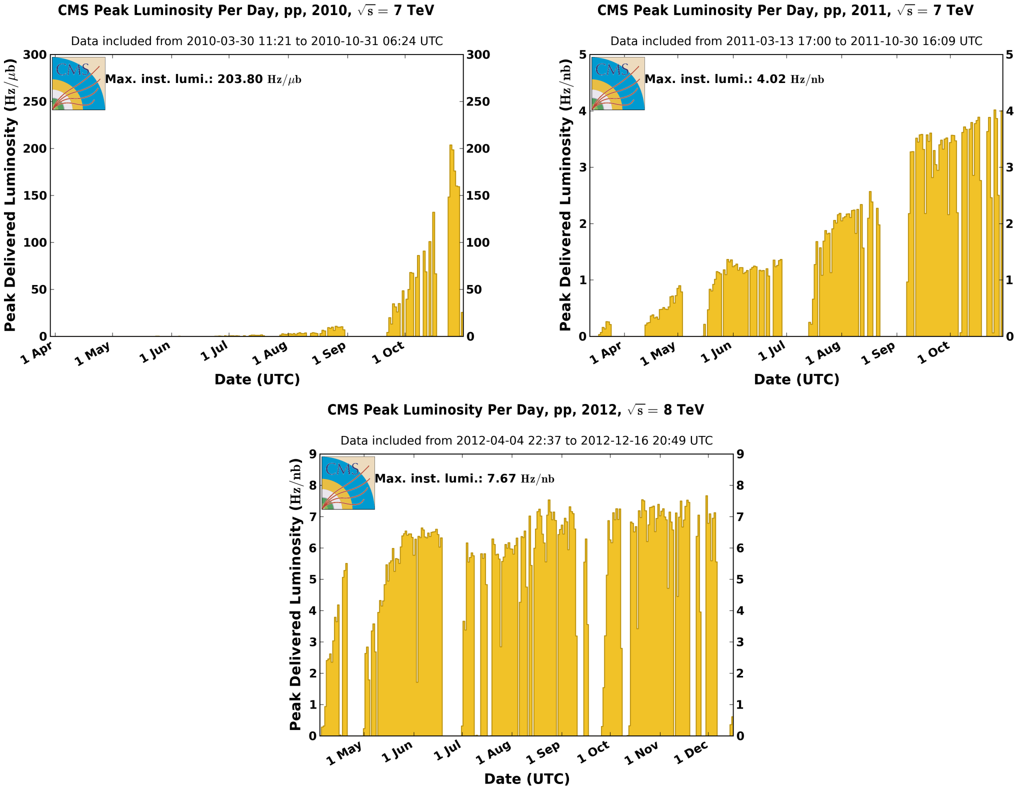


Figure 3.4: CMS's instantaneous luminosity during LHC Run 1. The figures above show the maximum instantaneous luminosity per day over the course of 2010, 2011, and 2012, respectively, as measured by CMS [41].

3.2 Compact Muon Solenoid (CMS) Detector

CMS is a general-purpose, hermetic detector which is designed to detect a variety of particles over a large energy range (from a few GeV up to the TeV scale for muons, electrons, and photons). In order to achieve this goal CMS is composed of many different subdetectors which are systematically layered and specialized for detecting specific types of signatures or particles. In addition to the subdetectors a superconducting solenoid that bathes the inner subdetectors in a uniform magnetic field of 3.8 T is a central component to CMS.

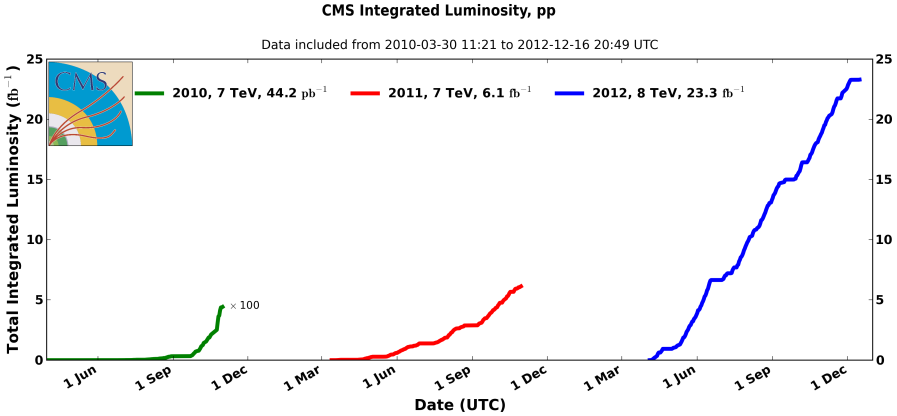


Figure 3.5: CMS's integrated luminosity during LHC Run 1. The development of delivered proton-proton collisions from 2010 to 2012 is shown. During 2010 only a small amount of data was delivered and so it is scaled by a factor of 100 to be visible on the same scale.

3.2.1 Coordinate System

Standard coordinates are defined by CMS for the convenience of the collaboration. The origin is defined as the nominal collision point. The x-direction is toward the center of the LHC ring. The y-direction is vertically upward, and therefore, since $\hat{x} \times \hat{y} = \hat{z}$, the z-direction is toward the Jura mountains, away from Lac Léman, and counter-clockwise as viewed from above LHC.

Cartesian coordinates are often very inconvenient, and so cylindrical coordinates are defined and more commonly used (as will be the case throughout this text). The azimuthal angle, ϕ, is measured in the x-y plane starting from the positive x-axis and increasing in the direction of the positive y-axis. The polar angle, θ, is measured from the positive z-axis. However, pseudorapidity, defined as $\eta = -\ln\tan(\theta/2)$, is generally preferred to θ as a coordinate because rapidity differences are Lorentz invariant and for high energy (or massless) particles rapidity and pseudorapidity are effectively identical. The final coordinate to define is the radial coordinate, $\rho = \sqrt{x^2 + y^2}$.

3.2.2 The Magnet

CMS's solenoid is a 220 t, four-layered, aluminum-reinforced, Rutherford coil made of niobium-titanium. NbTi is a superconducting material whose critical temperature, T_c, is 9.25 K in absence of a magnetic field. In the presence of high magnetic field the critical temperature is reduced ($T_c(B = 4.6T) = 7.35K$). In addition, there is a maximum temperature at which the current can flow unimpeded, T_{cs}. CMS was designed with a target of $T_{cs} = 6.5K$ so that currents of nearly 20 kA would only be one third of the critical current allowed. Therefore, this 12.5 m length and 6.3 m diameter solenoid is operated at a temperature of 4.5 K, well below this T_{cs} [42].

This solenoid has an inductance of 14.2 H, and when it is operating with current of 18,160 A to produce a central magnetic field of 3.8 T, 2.3 GJ of energy are stored in the solenoid. Its central purpose is to bend the trajectories of charged particles (particularly muons) up to the TeV scale in order to determine the charge to momentum ratio and thereafter infer the momentum of the identified particles.

Originally the magnet was designed and successfully tested to produce a field of 4 T. However, in order to increase the longevity of the magnet a field of 3.8 T was chosen to use in practice. This reduction in field strength does not impact the physics reach of CMS.

Outside of the solenoid there are three layers of steel return yokes, which are used to guide the magnetic field lines near the exterior of the solenoid. The purpose is to maintain a fairly strong magnetic field throughout the muon chambers. In the barrel the iron yoke is composed of five three-layered dodecagonal barrel wheels, which are 300 mm, 630 mm and 630 mm thick (from innermost to outermost) and 13 m long in total. In each endcap the yoke consists of three disks. They are 600 mm, 600 mm, and 250 mm thick (from innermost to outermost). Altogether these return yokes weigh 10,000 t.

The field throughout CMS was modeled and error was estimated using comic ray muons [43]. The overall uncertainty in the model's magnetic field was determined to be 3% in the three central barrel steel wheels and 8% in the outer barrel steel wheels. In Figure 3.6 a color map of the field strength with field lines are shown for the simulations of CMS's magnetic field through the interior of the detector and the return yokes.

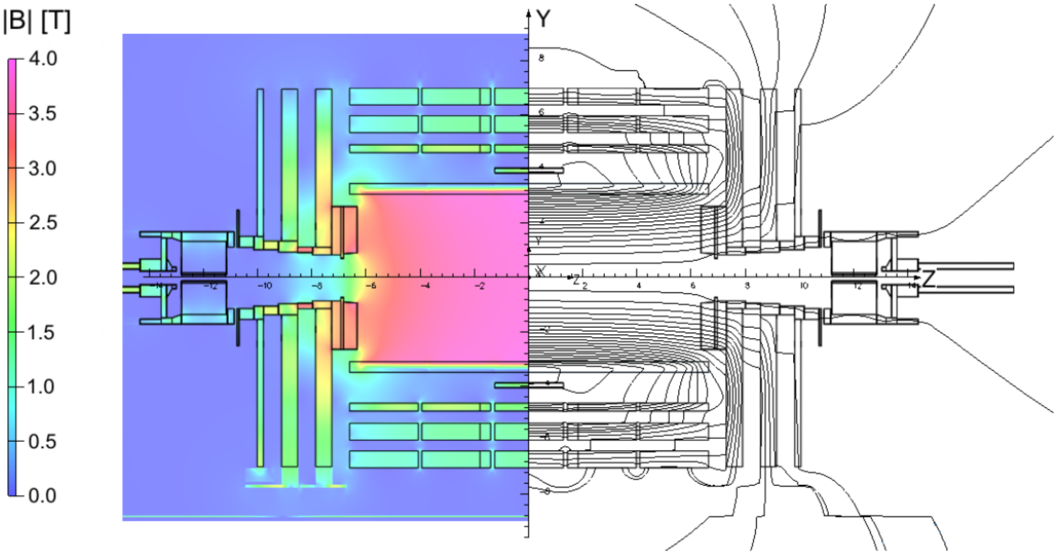


Figure 3.6: CMS magnetic field map. The magnetic field within and around CMS is modeled using TOSCA. On the left there is a color map of magnetic field strength with the central value of 3.8 T through the central cross section of CMS. On the right the field lines are displayed. From both it is clear that the iron yokes return the field lines back to the solenoid [43].

3.2.3 Tracker System

CMS's silicon tracker is the largest of its kind to be constructed. The purpose of this detector is to reconstruct the trajectories of charged particles which are produced by high energy proton-proton interactions. Using these trajectories CMS reconstructs the initial position of the collision or identifies secondary vertices (typically from bottom quark or τ decays). This is particularly challenging because (as envisioned) in 2012 the average number of interactions per bunch crossing was 21. Having many interactions overlapping in a single recorded event is called pile-up (PU), and it presents many challenges for tracking and for physics analysis.

Recording the tracks of particles with transverse momentum of 1 GeV or more requires different detection density because the density of particles decreases with ρ. At LHC nominal running conditions (luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$) at $\rho = 4$

cm a hit rate density of 1 $\mathrm{MHz/mm^2}$ is expected. When $\rho = 22$ cm the hit rate density is considerably reduced to 60 $\mathrm{kHz/mm^2}$. At $\rho = 115$ cm the hit rate density is a further order of magnitude reduced to 3 $\mathrm{kHz/mm^2}$.

There is about 200 $\mathrm{m^2}$ of active silicon area in the detector. A cross sectional view of the layout of all the silicon strips and inner pixel detector can be found in Figure 3.7. Notice that there are three layers of barrel plus endcaps in the tracker. Given the high hit density close to the interaction point a high granularity pixel detector is the innermost tracker detector from about 4 cm to 11 cm in ρ. The outer layers are the outer tracker, which is fully composed of silicon strips. The tracker is 5.8 m long with a diameter of 2.5 m.

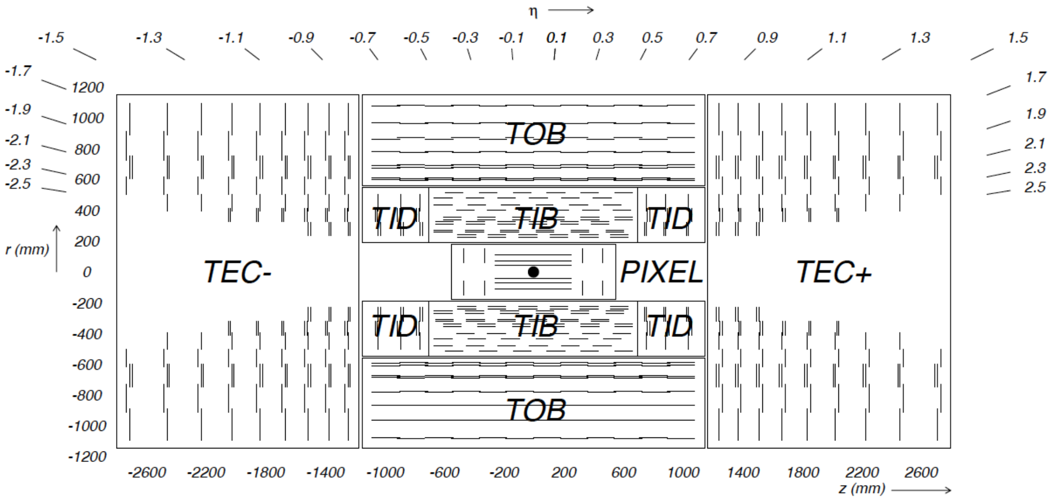
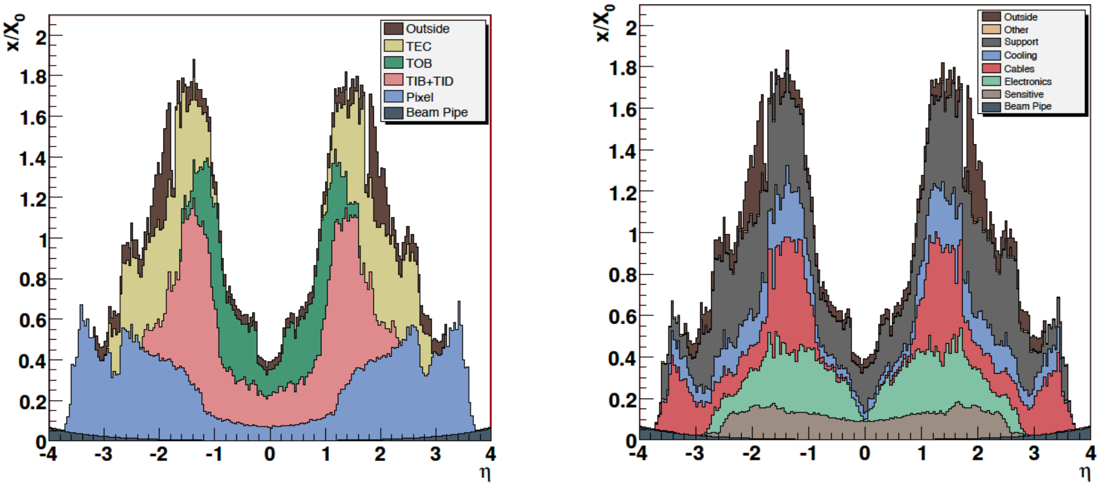


Figure 3.7: Cross sectional view of the CMS tracker along the beamline of the detector. Closest to the interaction point is the pixel detector surrounded by the tracker inner barrel (and disks). Beyond these are the tracker outer barrel and tracker endcaps.

Total power dissipation inside the tracker is expected to be 60 kW. In order to cool the tracker to near or below -10°C, tetradecafluorohexane C_6F_{14} is circulated at a rate up to 77 $\mathrm{m^3/hr}$ down to a temperature of -35°C. This corresponds to a cooling capacity of up to 128 kW.

The greatest challenge in producing such a high granularity, multilayered, silicon tracker is to minimize the amount of material inside the tracker while providing essential cooling, cabling and other infrastructure. The implications for analysis is that with more radiation lengths there is a greater probability of bremsstrahlung in electrons and nuclear interactions for photons, reducing measured energy and increasing (electron) pair production. In Figure 3.8 one can see several interesting features of the material placed in the tracker. The amount of material within a pseudorapidity of 1 is less than 1 radiation length. With a pseudorapidity of between 1 and 2.5 there is 1-2 radiations lengths.

Since the Electromagnetic Calorimeter is the next layer of the detector, more material in the tracker implies worse energy resolution for photons and electrons. The specific implication for the $H \rightarrow \gamma\gamma$ search is that the central barrel has the highest resolution photons, and therefore, the data from the central barrel is most sensitive in detecting the $H \rightarrow \gamma\gamma$ decay.



(a) (b)

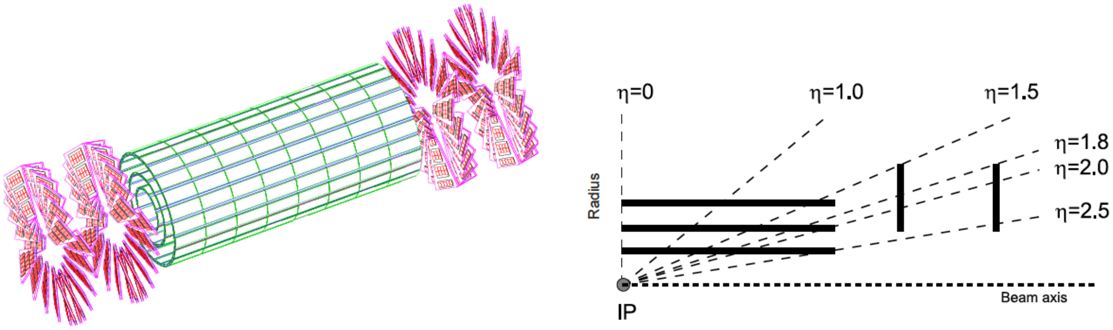
Figure 3.8: Tracker material inside CMS detector in radiation lengths as a function of pseudorapidity. The two plots show different breakdowns of the same material. In (a) the material is broken down into tracker subdetectors. In (b) the material is broken down by functional component.

3.2.3.1 Pixel Detector

The pixel detector is the innermost CMS sub-detector. In Figure 3.9a there is a three dimensional model of the pixel detector. Notice there are three layers in the pixel barrel and two disks on either end. The pixel barrels have mean radii of

4.4, 7.3, and 10.2 cm with a length of 57.0 cm. The endcap disks are 34.5 and 46.5 cm from the origin on either end with an inner radius of 6 cm and an outer radius of 15 cm. These disks are made of 2 half-cylinders with 12 blades, each tilted by 20°, in each half-cylinder. In Figure 3.9b the resulting pseudorapidity range of $-2.5 < \eta < 2.5$ is depicted.

There are approximately 48 million pixel detectors in the barrel layers and 18 million in the endcap disks. Each pixel detector has a $\rho - \phi$ by z size of $100 \times 150\ \mu\text{m}^2$. The spatial resolution of hits is between 15-20 μm in both $\rho - \phi$ and z directions because of signal interpolation between neighboring pixels. This resolution and position allows the pixel detector the ability to distinguish secondary vertices from bottom quark and τ decays with few tracks as well as to identify primary and PU vertices precisely.



(a) (b)

Figure 3.9: Schematics of CMS inner pixel detector. In (a) there is a 3-dimensional schematic of the pixel detector. One can see the three layers in the barrel and two pixel disks at either end. Figure (b) shows the physical end of the pixel detector at $\eta = 2.5$.

3.2.3.2 Silicon Strip Tracker

Outside of the pixel detector there is a silicon strip detector which is composed of 15,148 detector modules. Each of these modules either has a single 320 μm thick silicon sensor or two 500 μm thick silicon sensors. Altogether these modules contain a total of 24,244 sensors and approximately 9.3 million silicon strips.

The strip detector fills a space in ρ from 20 cm to 116 cm and has four large pieces: the Tracker Inner Barrel (TIB), the Tracker Inner Disks (TID), the Tracker Outer Barrel (TOB), and the Tracker Endcaps (TEC).

The TIB is made of four concentric cylinders with radii of of 255.0 mm, 339.0 mm, 418.5 mm, and 498.0 mm. Each cylinder extends from -700 mm to 700 mm in z-direction. The inner two cylinders use double-layered thick (500 μm) silicon sensors with a pitch of 80 μm. The outer two cylinders use single-layered thinner (320 μm) silicon sensors with a pitch of 120 μm. The TID is made of three identical disks either side of the TIB from 800 mm to 900 mm from the origin. Each has covers a radial distance from 200 mm to 500 mm. Together the TIB and TID cover the pseudorapidity range of -2.5 to 2.5.

Surrounding the TIB and TID is the TOB, which is constructed as a single "wheel" structure. Within it there are 6 detector cylinders. Their average radii are 608, 692, 780, 868, 965, and 1080 mm. The wheel is 2180 mm long with an inner radius of 555 mm and an outer radii of 1160 mm. The structure is held together using 688 rods which are also used as a pathway for cooling. By positioning the rods ±16 mm away from their center of mass positions, fiducial holes in ϕ are avoided.

Beyond the TIB-TID-TOB, the TEC extends from 220 mm $< \rho <$ 1135 mm and 1240 mm $<$ |z| $<$ 2820 mm for TEC+ and TEC-, where the sign denotes the sign of the z direction. Each TEC is made of nine disks, which contain up to seven silicon detector rings. These rings contain radial detectors with pitch varying from 97 μm to 184 μm. The sensors in the inner four rings are 320 μm thick and the sensors in the outer three rings are 500 μm thick. Rings 1, 2, and 5 have a second layer of sensors at an angle of 100 mrad which provides a measure of z-coordinate. Disks 1-3 have all seven rings. Disks 4-6 lack the first ring. Disks 7 and 8 lack the first two rings, and disk 9 only has rings 4 to 7.

3.2.4 Electromagnetic Calorimeter (ECAL)

The ECAL's main function is to induce electromagnetic showers with high energy electrons and photons and then determine the energy of those objects. For the $H \to \gamma\gamma$ search the ECAL's precision and accuracy are fundamental and indeed the ECAL was designed with this narrow resonance search in mind.

This ECAL is a hermetic homogeneous calorimeter made of 75,848 lead tungstate ($PbWO_4$) crystals. The ECAL is composed of a barrel with 61,200 crystals, endcaps at either end with 7,324 crystals in each, and layered lead-silicon preshower detectors in front of the endcaps. Figure 3.10 depicts an overview of these primary components of the ECAL.

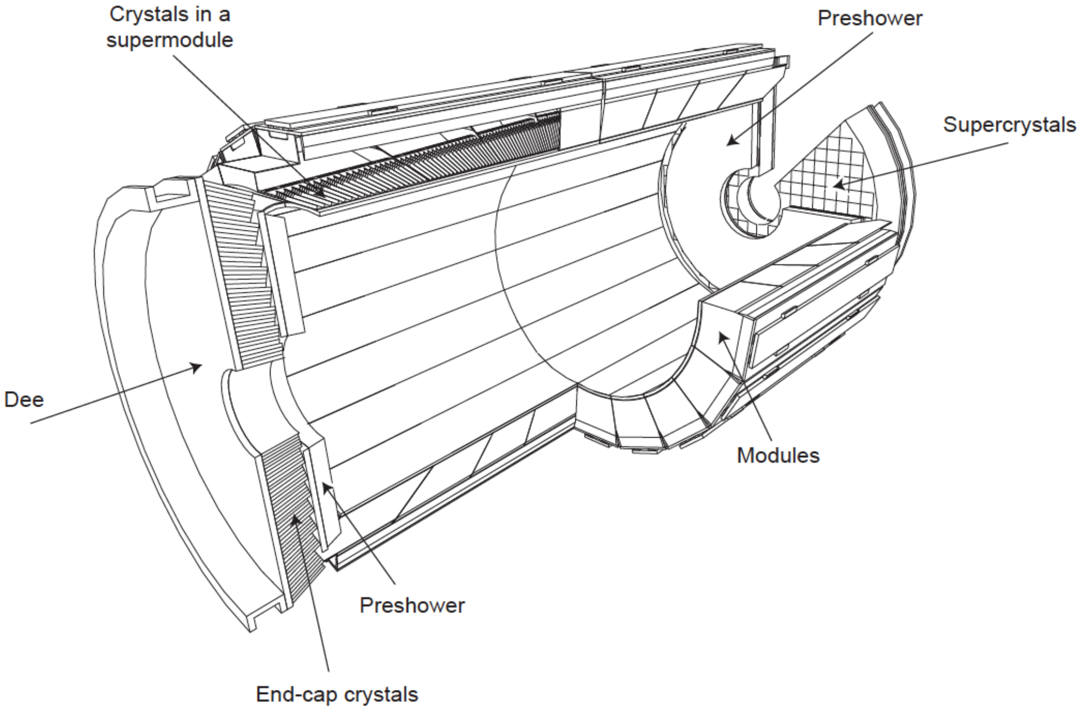


Figure 3.10: CMS electromagnetic calorimeter. This schematic of the ECAL depicts the barrel, preshower and the endcaps and highlights the supermodules of the barrel and the "Dees" of the endcaps.

Lead tungstate was chosen as the ECAL crystal because it has a short radiation length (0.89 cm), small Molière (showering) radius (2.2 cm) and because the scintillation decay time of the crystals is on the order of the bunch space of the LHC (i.e. about 80% of the light is emitted within 25 ns). The scintillation light has a broad maximum in the range of 420-430 nm (blue-green) and is absorbed by photodetectors attached at the back of the crystals. About 4.5 photoelectrons are absorbed per MeV by the photodetectors throughout the ECAL. The light output varies with temperature and so the temperature of the crystals is maintained near the nominal temperature of 18°C with an error of ±0.05°C by circulating water through each supermodule independently.

The crystals in the barrel are 23.0 cm long or 25.8 X_0 (radiation lengths). The front face (tracker side) of EB crystals is 22 × 22 mm^2 and the rear face (at the photodetector) is 26 × 26 mm^2. This corresponds to an $\eta - \phi$ cross section of approximately 0.0174×0.0174. The front faces of the EB crystals are 129 cm from the beamline throughout the EB, and their total volume is 8.14 mm^3. The EB covers $|\eta| < 1.479$.

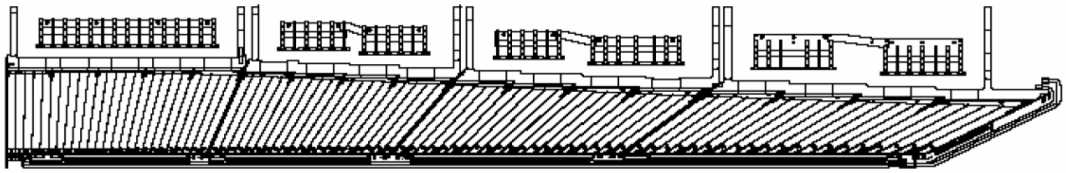
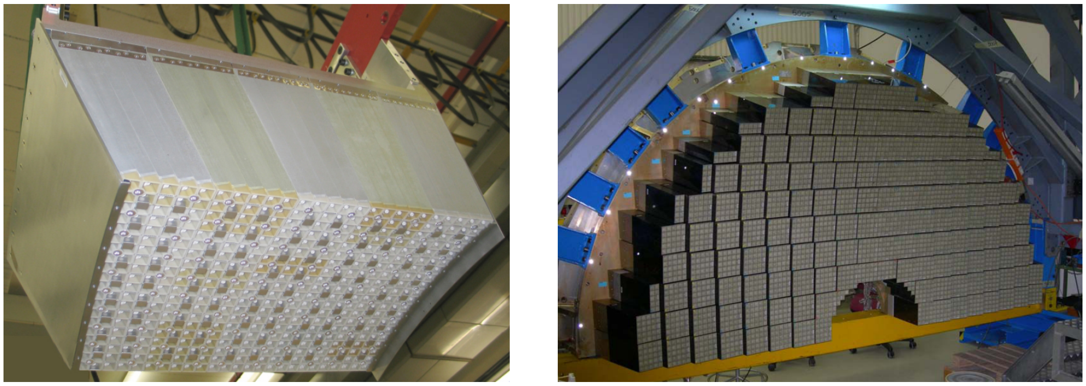


Figure 3.11: Slice view of an ECAL supermodule. Four modules within an ECAL supermodule are visible. At higher values of η the crystals are tilted more and more to point toward the interaction region [44].

The barrel crystals are grouped into 18 supermodules which cover a region in ϕ of 20° for positive and negative values of η for a total of 36 supermodules. Within each supermodule there are 1700 crystals which are divided into four modules of 400 or 500 crystals. Figure 3.11 depicts a slice of a supermodule and in Figure 3.12a there is a photograph of a single module.

The endcaps (EE) are located 315.4 cm on either side of the interaction point. Each EE is split into two pieces, which are called "Dees". Each Dee contains 3,662 crystals. The crystals are grouped in 5×5 supercrystals with the exception of 18 partial supermodules which are at the inner and outer edges of the Dees. In Figure 3.12b there is an image of a Dee attached to the HCAL. Each EE crystal is 220 mm long, which is 24.7 X_0 (radiation lengths). The front face (tracker side) of EE crystals is 28.62 × 28.62 mm^2 and the rear face (at the photodetector) is 30 × 30 mm^2. The volume of the EE crystals is 2.90 mm^3 and covers $1.479 < |\eta| < 3.0$.

Different photodetectors were chosen to capture the light from the ECAL crystals in the barrel and endcaps because of different magnetic field strengths and different levels of expected radiation damage. For the EB avalanche photodiodes



(a) (b)

Figure 3.12: ECAL modules. A single EB module is photographed in (a). In (b) there is an EE Dee attached to the HCAL. Altogether there are 144 EB modules and 4 EE Dees that compose the ECAL.

(APDs) are employed to collect scintillation light, where as vacuum phototriodes (VPTs) are used for the EE.

In the EB two Hamamatsu type S8148 reverse structure APDs are attached to each crystal as depicted in Figure 3.13a. Each has an active area of 5×5 mm^2. The APDs are operated at a gain of 50. The gain is heavily dependent on the voltage, and its stability is therefore extremely important for the APDs consistency in amplification. A custom high voltage (HV) system was made in collaboration with the CAEN Company. The HV is placed 120 m from the detector to avoid radiation. In order to stabilize the voltage to the order of 10s of mV the HV measures and corrects applied voltage. Applied voltage required for a gain of 50 is 340-430 V.

Due to significantly larger expected radiation in the EE, silicon APDs were disfavored. VPTs consisting of a single gain drop (600 V to 800 V) were developed especially for use in CMS's 4 T magnetic field. The VPTs have a diameter of 25 mm and an active area of approximately 280 mm^2. An EE crystal with a VPT attached is depicted in Figure 3.13b.

In front of both EE there are preshower units (ES) with alternating layers of lead and silicon. The objective of ES is to induce electromagnetic showering in photons and electrons before reaching the EE thereby distinguishing jets with

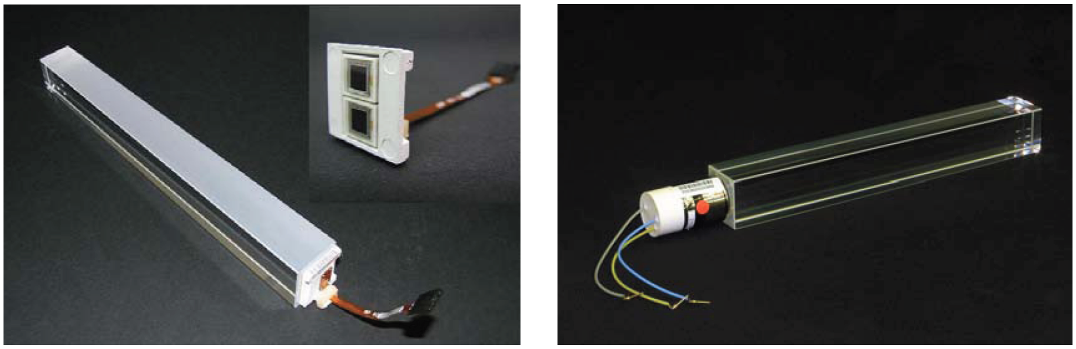

(a) (b)

Figure 3.13: Individual ECAL crystals. Figure (a) shows two APDs attached to an EB crystal. Figure (b) shows a single, circular VPT glued to the rear of an EE crystal.

high energy π^0s from photons. The ES covers 1.653< $|\eta|$ <2.6 and is 20 cm thick.

Before reaching the ES there is $2X_0$ of material. The ES itself is two silicon sensors sandwiching a $1X_0$ thick lead slab. The ES is made in two Dees similar to the EE and with the same orientation. The silicon detectors are 320 μm thick and have an area of 63×63 mm^2 (active area - 61×61 mm^2). Each sensor has 32 strips, which corresponds to a pitch of 1.9 mm. The two layers of silicon strips are orthogonal.

For photons and electrons with energies below 500 GeV, where the leakage from the rear starts to become appreciable, the resolution can be parameterized as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \qquad (3.2.4.1)$$

where S is a stochastic term, N is a noise term and C is a constant. The stochastic term takes photon statistics and event-by-event variation in lateral containment. The noise term represents electronics, digitization and PU noise. In the EB, radiation damage to the APDs increased so-called "dark current" during the 2012 data collection. The constant term is caused by non-uniformity in longitudinal, intercalibration errors and leakage of energy at the back of the crystals.

There is transparency loss due to radiation during active use in the crystals. During downtime the crystals recover some transparency. Because the response of the crystals varies it must be monitored continuously. A laser emitting light with a wavelength 440 nm, which is near the maximum for the scintillation light, is used to measure the relative response. A full scan of the ECAL requires 30 minutes.

The calibration of the ECAL is performed *in situ* with real physics events. The individual crystals are inter-calibrated using $\pi^0 \to \gamma\gamma$ $\eta \to \gamma\gamma$ events as well as electrons in events with W bosons. Minimum bias events can be used to calibrate crystals in rings of η by assuming ϕ symmetry. The energy scale of the ECAL is determined using $Z^0 \to \mathrm{e}^+\mathrm{e}^-$ events. The final calibration is validated on $Z^0 \to \mathrm{e}^+\mathrm{e}^-$ events. This process is described more fully in Section 5.5.1.

3.2.5 Hadronic Calorimeter (HCAL)

The HCAL is a sampling calorimeter made of steel and brass layers that incite hadronic showers. Between these dense metals are tile, which use the wavelength shifting fiber concept, to bring out the light. The HCAL is composed of four subdetectors: the HCAL barrel (HB), the HCAL endcap (HE), the outer HCAL (HO), and the forward HCAL (HF). Together these detectors cover a range in η up to ~5.0. Figure 3.14 shows an overview of the detectors with a cross section image through the beamline.

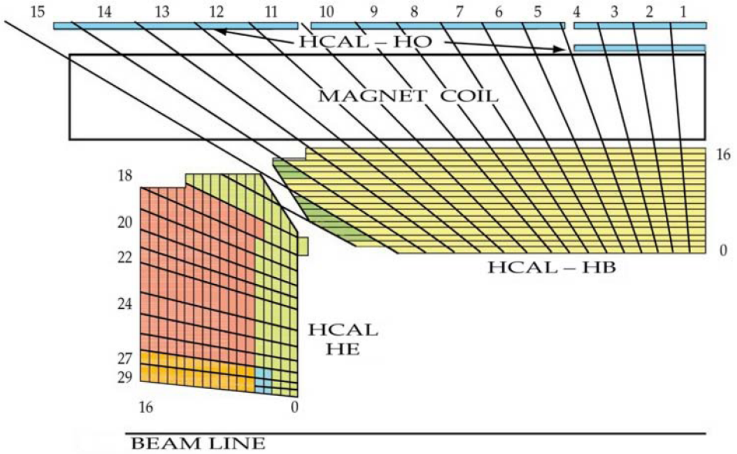
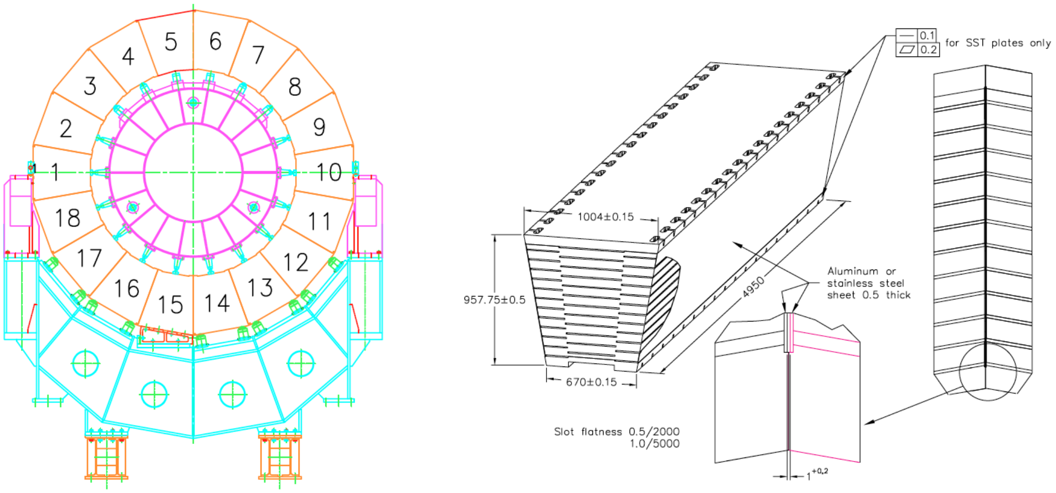


Figure 3.14: ρ-z cross sectional view of CMS hadronic calorimeter.

The HB fits between the ECAL outer edge ($\rho = 1.77$ m) and the inner edge of the solenoid ($\rho = 2.95$ m) and covers $|\eta| < 1.3$. It consists of 36 identical azimuthal wedges in two half-barrels as depicted in Figure 3.15. Each wedge is

20° in ϕ. Each wedge has four ϕ segments. From the inside of the wedge out, the absorbing layers are: a 40 mm steel plate, eight layers of 50.5 mm (C26000 cartridge, 70% Cu, 30% Zn) brass plates, six 56.5 mm brass plates and a 75 mm steel plate. At $\eta = 0$ this is 5.82 interaction lengths (λ_I). At the edge of the HB where $\eta = 1.3$ the effective length is $10.6\lambda_I$ and between the effective length varies as $\frac{1}{\sin(\theta)}$. The ECAL is an additional $1.1\lambda_I$ of material inside the HCAL.



(a) (b)

Figure 3.15: HCAL barrel wedges and individual wedge. Figure (a) is a longitudinal cross section of the HB. The wedges are labeled from 1 to 18 with wedge 1 located on the positive x-axis. Figure (b) is a single wedge in more detail.

Between the HB active layers are plastic scintillators. Just before the first layer of steel there is a 9 mm Bicron BC408 scintillator. All other scintillators in the wedge are Kuraray SCSN81 plastic. All are 3.7 mm thick except for the layer beyond the back steel plate which is 9 mm. In total the HCAL has ~70,000 scintillating tiles. The scintillated light from these tiles is collected by wavelength shifting, double-clad, 0.94 mm diameter fiber (Kuraray Y-11). The wavelength shifting fibers (WLS) are spliced onto clear fibers that finally deliver the light to hybrid photodiodes (HPD).

The material in the central barrel is insufficient to guarantee that all hadronic material will be absorbed within the solenoid. The solenoid itself provides an additional $\frac{1.4}{\sin(\theta)}\lambda_I$ of material. The HO was designed to absorb additional hadrons that reach beyond the solenoid. Recall the return yoke outside the solenoid is five steel barrels (see 3.2.2) and each is 2.536 m wide. The HO consists of five scintillator rings of the same longitudinal configuration at ρ =4.07 m. In the region of the central ring there is the least amount of material and an additional scintillator ring is placed at ρ =3.82 m. Between the two scintillators there is a 19.5 cm iron ring. With this additional material there is a depth of at least 11.8 λ_I everywhere except in the barrel-endcap transition region.

The HE covers a large range of pseudorapidity ($1.3 < |\eta| < 3.0$) and must withstand an extremely high particle flux ($\sim$34% of final state particles). In addition the magnetic field in the region of the HE is still very strong. Because C26000 cartridge brass is non-magnetic and has good physical and mechanical properties, it was chosen to be used in the HE and throughout the HCAL. The HE is attached to the iron return yokes, and the EE Dees are attached directly to the HE. The HE itself is a series of 79 mm brass plates with 9 mm gaps for the scintillators. Different layers are staggered so that there are no gaps in ϕ. All the scintillators are 3.7 mm SCSN81 except for layer 0, which is 9 mm Bicron BC408. Tower 18 has an additional layer (-1) to compensate for the towers 16 and 17 being cutoff so that photodetectors and front-end electronics may be installed there. These features can be seen in Figure 3.16. The granularity of the HE is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| > 1.6$

The HF experiences a very high particle flux because it is very close to the beamline, covering $2.9 < |\eta| < 5.2$. On average 760 GeV of energy is deposited in the HF per proton-proton interaction. The particle flux of above 10^{11}cm^{-2} is not uniform but rapidly increasing as a function of rapidity. Quartz fibers were chosen as the active material because of their resilience against radiation damage. Cherenkov light is produced by particles passing the Cherenkov energy threshold, which makes the HF most sensitive to electromagnetic radiation.

The calorimeter of the HF is a cylinder of steel absorber with 5 mm grooves for the quartz fibers. Half of these fibers run the entire length (165 cm $\approx 10\lambda_I$) and the other half start 22 cm from the front. These two sets of fibers are readout separately and serve to distinguish photon and electron showers. The HF has an

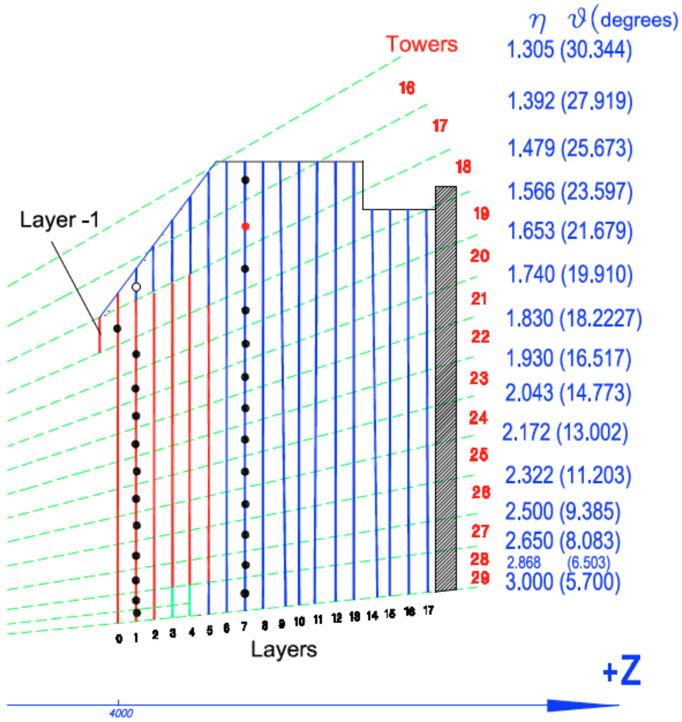


Figure 3.16: Longitudinal view of the HE.

outer radius of 130.0 cm and an inner radius of 12.5 cm. The cylinder is made of 18, 20° in ϕ wedges. There is one such cylinder at $z = 11.2$ m and another at $z = -11.2$ m. The entire HF is encased in 40 cm of steel, 40 cm of concrete and 5 cm of polyethylene. The granularity of the HF is $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$.

Measurements from the HCAL are not directly used in the $H \rightarrow \gamma\gamma$ analysis except to veto objects reconstructed as photons with large ratios of HCAL energy to ECAL energy. However, in many of the exclusive channels, which probe the production mechanisms of the Higgs, hadronic jets are required. The most significant of these is the di-jet tag targeting the vector boson fusion Higgs production. Jets from the HB, HE and very forward jets in the HF are all utilized in that analysis.

3.2.6 Muon Chambers

Detection of high resolution muons is a fundamental component of many physics searches in CMS. The most important decay, which this detector was designed to detect, is the $H \rightarrow ZZ^{(*)}$ decay to muons. Excellent resolution is needed and delivered by the muon chambers in this channel to identify the mass of the sought boson with great confidence. In the $H \rightarrow \gamma\gamma$ search, the resolution of the muon chambers are not utilized, but the detection of one or two high-p_T muons with two photons is used to probe Higgs production via higgsstrahlung from a W or Z boson where the vector boson decays leptonically with muons.

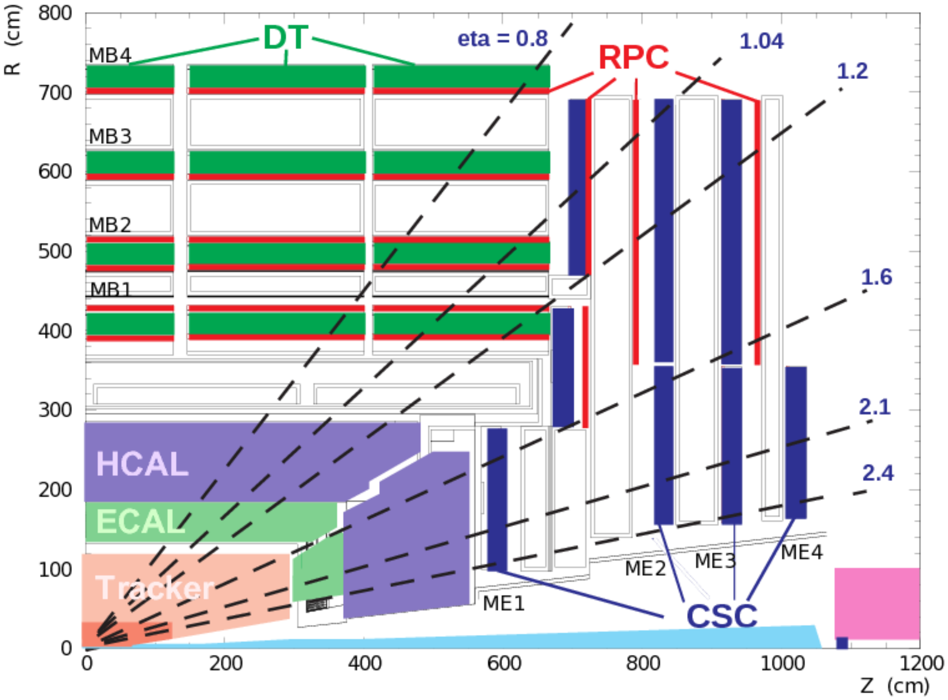


Figure 3.17: ρ-z view of CMS, highlighting the muon systems.

The design and layout of the muon systems are heavily impacted by the required iron return yokes and the varying magnetic field from the solenoid. For this reason the muon chambers are integrated with the return yokes. The muon chambers are composed of three independent muon detectors: drift tube systems (DT) in the barrel, cathode strip chambers (CSC) in the endcaps, and resistive plate chambers (RPC) throughout the detector. Together, these detectors, shown in Figure 3.17, have 95-99% muon detection efficiency over $|\eta < 2.4$ except where there are gaps between some sub-detectors (i.e. $|\eta|$ of 0.25, 0.8 and 1.2). Because 25,000 m^2 of detection planes are needed for the muon chambers, materials must be inexpensive, reliable, and robust.

3.2.6.1 Drift Tube Systems (DT)

Given the low occupancy in the barrel and relatively weak magnetic field outside the return yokes, traditional drift tube systems (DT) are employed. The muon barrel is four concentric cylinders, called stations, covering $|\eta| < 1.2$ and located between 4 m and 7.5 m from the beamline. The middle two stations are between layers of the return yoke while the inner and outer stations are within and outside the return yoke, respectively. The DT hits are used like tracker hits in practice. The inner three stations contain 60 drift tubes while the outer has 70. Inside these DTs are a gaseous mixture of Ar(85%) / CO_2(15%) and 172,000 gold-plated steel anode wires of length between 2 m and 3 m. Most wires are 2.4 m.

In the inner three stations there are 12 DT chambers. There are 14 DT chambers in the outer station. The DT chambers can be seen in Figure 3.18. Each DT chamber is made of super-layers (SL) which are composed of four staggered layers of drift cells. The DT chambers in the inner three DT chambers have three SLs. The outer two SLs are parallel to the beamline and measure position in ρ-ϕ while the inner layer's wires are orthogonal to the beamline and measures z position. The outer stations only have two DT chambers and the wires in both are parallel to the beamline.

3.2.6.2 Cathode Strip Chambers (CSC)

In the endcap DTs are not ideal because of higher, variable-strength magnetic field and because of higher rates. The CSCs are multiwire proportional chambers with six anode wire planes placed between seven cathode panels. These CSCs are trapezoidal wedges filled with $40\%Ar+50\%CO_2+10\%CF_4$. The wedges are 10° or 20° wide and one is depicted in Figure 3.19, which make up four layers of discs that cover $0.9 < |\eta| < 2.4$. During the 2010-2012 data collection there were 468 such wedges in CMS.

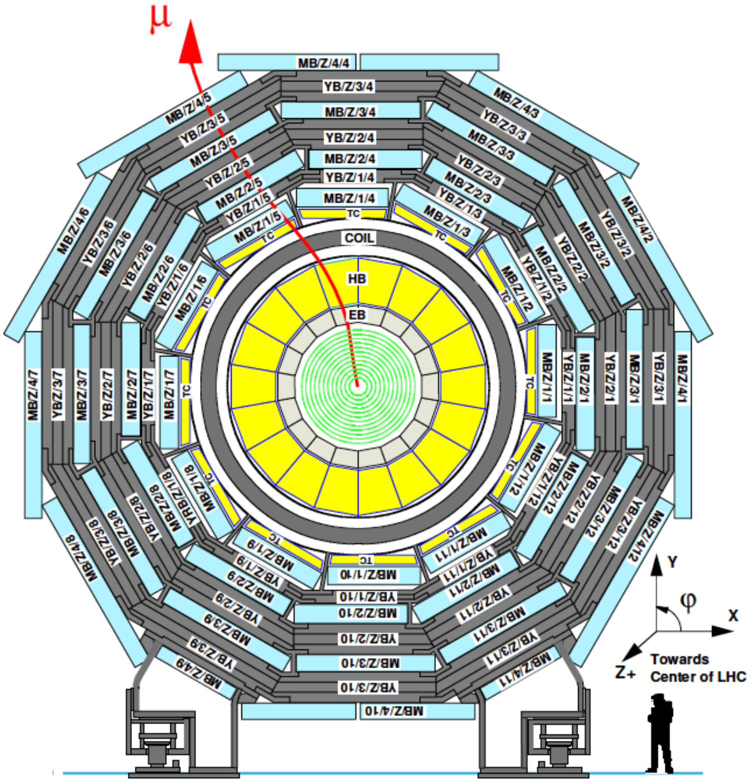


Figure 3.18: ρ-ϕ view of DT systems.

3.2.6.3 Resistive Plate Chambers (RPC)

In addition to the DTs and CSCs, there are RPCs in the barrel and endcaps which cover an area of $|\eta| < 1.6$ and can be seen in red in Figure 3.17. The RPCs have limited spatial resolution, but they are very fast and much shorter than the LHC bunch spacing. For this reason they are used in L1 triggering.

The RPCs are gas-filled parallel-plate detectors. They are filled with a non-flammable mixture of 96.2% R134a ($C_2H_2F_4$), 3.5% iC_4H_{10} and 0.3% SF_6. In the barrel there are six cylinders RPCs. There is one on either side of the first DT stations. In the outer two DT stations there are single RPCs that are next to and within the DTs.

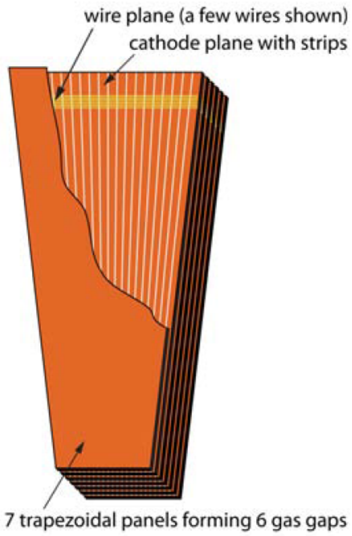


Figure 3.19: Diagram of Cathode Strip Chamber component wedge. A CSC wedge contains six multiwire anode planes sandwiched between seven cathode planes.

3.2.7 Trigger and Data Acquisition System

The bunch space between proton bunches, and therefore between collisions, was at minimum 25 ns in 2011 and 2012. This corresponds to an event rate of 40 MHz. Because each event requires approximately 1MB to record, it is not possible or at all practical to record each event due to limitations in disk space. Moreover, the hardware buffers are large enough to temporarily store all the events while any single event is being processed. Luckily, many events will not be relevant for physics analysis and need not be recorded.

There is a two step process for reducing the 40 MHz rate to a more reasonable and manageable rate of 300 Hz. First, there are Level-1 (L1) Triggers which reduce the rate to approximately 100 kHz, and after passing a L1 trigger path High-Level Triggers (HLTs) perform more complex calculations and further reduce the rate of recording events to the desired rate of 300 Hz.

The L1 is a collection of field-programmable gate arrays (FPGAs), as well as ASICs and programmable lookup tables, which can be re-programmed. Indeed, between 2011 and 2012, a significant re-design of the L1s was necessary because of rather large changes in the typical pile-up per event. The L1 must analyze events quickly, and therefore coarsely, because of a limited pipeline for events to pass through. Therefore, latency time (from bunch-crossing to L1 decision) is limited to 3.2 μs. The logic of the L1 trigger systems is to collect local Trigger Primitive Generators, which are in term made into regional objects (e.g. electrons, muons, photons, jets). These are sorted by p_T by the Global Muon Trigger and Global Calorimeter Trigger, and the relevant objects are passed to the Global Trigger. The Global Trigger decides whether an event should be fully processed and evaluated by the HLT. This process is depicted in Figure 3.20.

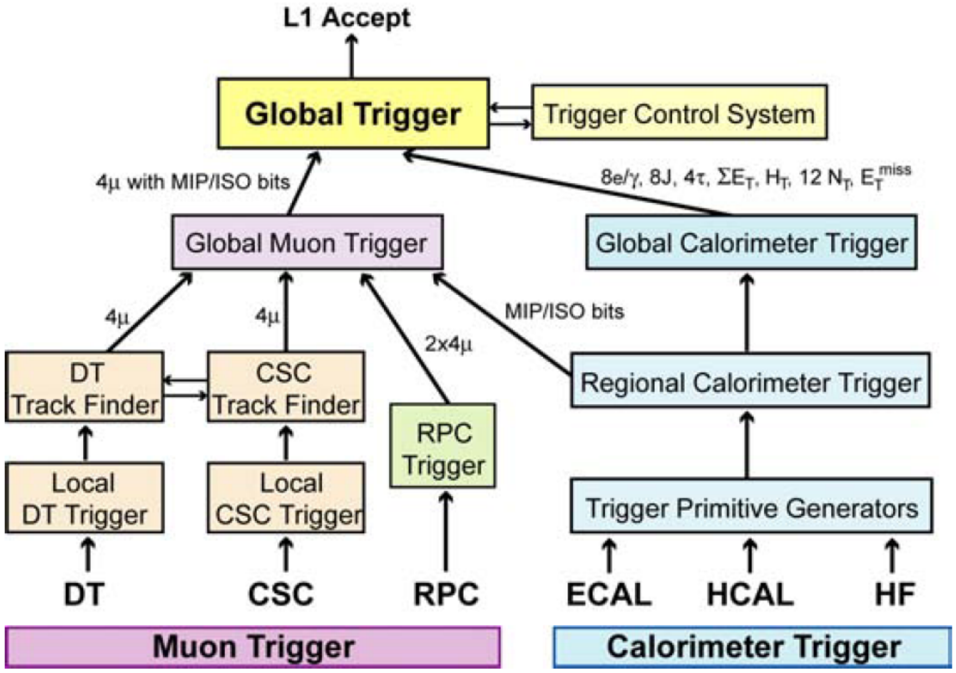


Figure 3.20: Overview of L1 architecture.

The HLT is part of the Data Acquisition (DAQ) system. Specifically the HLT is C++ code that is running on a computing farm known as the Event Builder. The DAQ is able and required to take events at a rate of 100 kHz, reconstruct them using fast, approximate algorithms, and finally apply complex filers via the HLT code to select and output events at the rate of 300 Hz. A visual depiction of this process is found in Figure 3.21. Each HLT path is designed for a specific purpose or analysis. If an event passes any HLT path, then the event is fully reconstructed and passed into one or more primary datasets based on the paths which the event passes.

Some HLT paths are “pre-scaled”, meaning that only a fraction of events which pass the L1 and HLT paths are retained for analysis. As instantaneous luminosity increased through Run 1 looser, higher rate HLTs needed to be pre-scaled in order to keep the overall rate below 300 Hz. All the HLTs used for the analyses presented here are not pre-scaled.

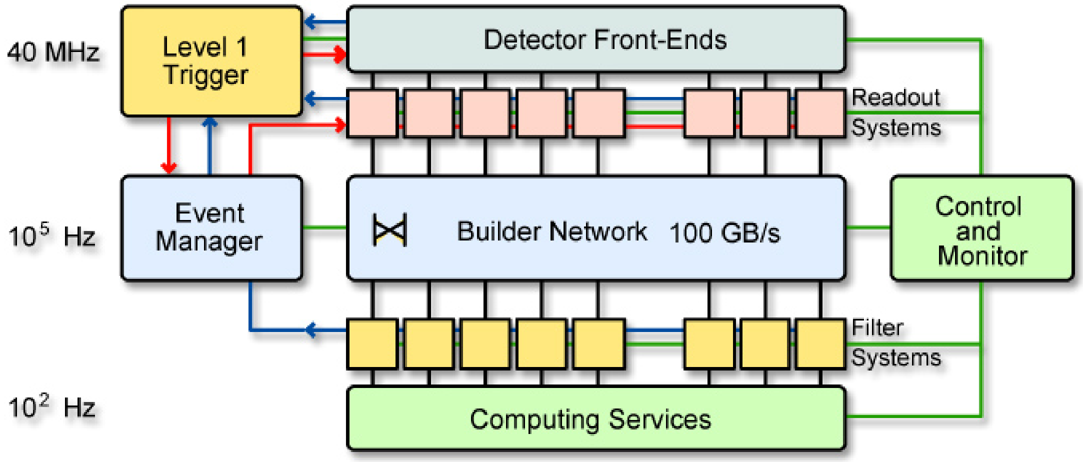


Figure 3.21: Overview of Data Acquisition architecture.

Chapter 4

The Development of CMS's $\mathrm{H} \to \gamma\gamma$ Analysis

A SM Higgs boson with mass less than 150 GeV is favored by electroweak precision measurements (see Figure 2.11) and a light Higgs boson with mass of at most 135 GeV is an important feature of some supersymmetry models [45]. The $\mathrm{H} \to \gamma\gamma$ channel was considered to be vital for the low-mass SM Higgs search, particularly for Higgs masses below 120 GeV where ZZ and WW channels are suppressed. Since the $\mathrm{H} \to \gamma\gamma$ channel was expected to have a relatively low signal to background ratio (S/B), it was important to make every reasonable attempt to get good performance from the detector and from the analysis. While the expected number of signal events is rather large, the amount of background in the di-photon final state is much larger. The strategy envisioned to estimate the background was to use a fit to the data, which is primarily background. In this way the background estimate is not sensitive to any deficiencies in MC simulation and is highly accurate.

One way to reduce background was to improve the mass resolution of the detector–reducing the amount of relevant background under a narrow mass peak. To this end, CMS chose a very high resolution ECAL design with expected photon energy resolution of about 1% for photons from the $\mathrm{H} \to \gamma\gamma$ decay. Thus, the analysis strategy was to look for a narrow mass peak above background which is determined from the data.

Since the adoption of this ECAL design, the expected amount of tracker material inside of the ECAL has increased dramatically, particularly for $|\eta| > 0.5$. Thus, photon energy resolution in the detector is quite good at low $|\eta|$, worse in the barrel at higher $|\eta|$, and even worse in the endcaps. With the actual detector, both the photon energy resolution and S/B are very non-uniform over the detector. In addition, photons that convert and start to shower in the tracker material have substantially worse expected resolution and S/B than photons identified to be unconverted. Since analysis performance is still very important and in view of the channel's high background nature, events with good resolution and good expected S/B must be optimally used but worse resolution and lower S/B events cannot be ignored.

The UCSD group proposed the following analysis [46], which became CMS's benchmark $\mathrm{H} \to \gamma\gamma$ analysis in CMS's Physics Technical Design Report (PTDR) [47]. At that time, it was already observed that events with one or more photon in the endcaps had significantly worse resolution and contributed much less to analysis sensitivity than originally envisioned. It was also seen that R_9–a shower shape variable measuring the fraction of energy in the central nine crystals–could be used as a simple conversion tag[1]. For $R_9 < 0.95$ photons generally have worse resolution and events containing low R_9 photons have low S/B. Thus, in the PTDR $\mathrm{H} \to \gamma\gamma$ analysis, it was proposed to categorize photons and classify events in $|\eta|$ and R_9. With this classification scheme, good resolution and high S/B were highly correlated and only a small number of event classes were needed. The PTDR also presents an "optimized" $\mathrm{H} \to \gamma\gamma$ analysis utilizing a neural network to improve the analysis performance.

Therefore, since the original formulation of the PTDR, these analysis principles have been important:

- Improve photon energy resolution wherever possible for the purpose of obtaining excellent mass resolution and estimating it.
- Determine the background from fits to the data with negligible bias.

---

[1] R_9 is defined precisely in Section 5.8.2.

- Divide the events into classes that have similar mass resolution and similar S/B to extract the most information from the best events and yet keep the signal acceptance very high to optimize analysis performance.
- Use the $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ process to make corrections to the mass resolution and energy scale.

With these guiding principles in place our first $\mathrm{H} \rightarrow \gamma\gamma$ analysis applied to Run 1 data was a cut-based analysis that is fundamentally based on fits to the data to estimate background. This original analysis has undergone only one major update since it was first developed.

Throughout 2011 and 2012 CMS has engaged in a hunt for further optimization in the search for the Higgs boson. In the process many aspects of the $\mathrm{H} \rightarrow \gamma\gamma$ analysis were competitively challenged by several different strategies. In addition, high S/B event classes utilizing additional final state objects from VBF, VH and $\mathrm{t\bar{t}H}$ production modes were devised. The event classes targeting VBF and VH production modes also add sensitivity to Higgs boson couplings to the weak bosons. Many of the author's contributions have been in the development of di-jet tags targeting VBF production and lepton tags targeting VH production.

Three major publications have been produced describing the $\mathrm{H} \rightarrow \gamma\gamma$ analysis and one publication describing the final analysis of the full Run 1 dataset is undergoing final editing at this time. Each publication has marked a different milestone in analyzed CMS data and an increasing level of complexity and sensitivity in the analysis. The first publication describes the first results using the entire 7 TeV dataset and shows the first hints of the Higgs boson with a 3.1σ excess near 125 GeV [48]. The first analysis was a robust cut-based analysis with an additional di-jet event class. A similar version of this analysis remains a cornerstone analysis of CMS today. In the summer of 2012 a publication describing the discovery of a Higgs boson candidate near 125 GeV was made. This publication introduced a more optimized analysis using Boosted Decision Trees (BDTs) to classify events for inclusive event classes [29]. The discovery was achieved by combining five search channels with exesses in the $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ channels driving the observation.

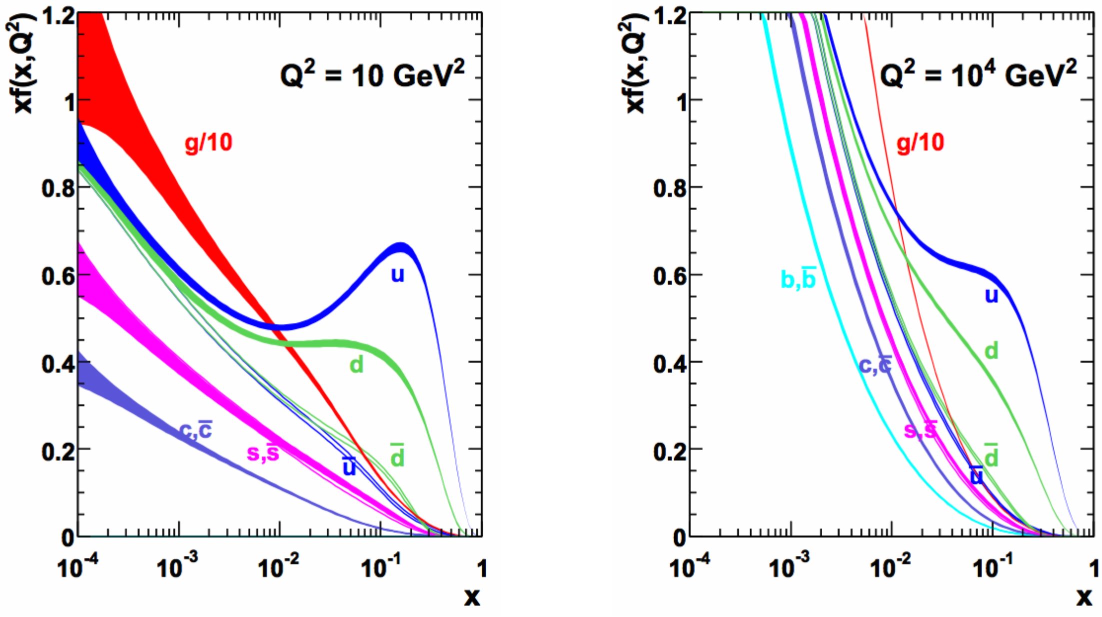
An additional publication [49] explores the interpretation of the Fermiophobic Higgs boson model with only the 2011 data. This model assumes that the Higgs boson does not couple to fermions or that such coupling is very weak. This analysis includes a di-jet tag for VBF production as well as muon and electron tags for VH production that I developed. In combination with a di-jet tag for VBF production in the H → WW channel the excess at 125 GeV was excluded at the 95% confidence level.

The final publication is fully described within this dissertation. It utilizes the full Run 1 dataset with final calibrations. This analysis achieves an independent observation in the $H \to \gamma\gamma$ channel. Many measurements have been made with this finalized dataset including the mass of this Higgs boson, various parameterizations of the couplings of this particle, and the spin-parity compatibility of the observation. This analysis has been finalized and paper editing has reached its final stages.

This chapter highlights the contributions of the author to all four publications as well as important improvements with respect to previous analyses.

4.1 Higgs Boson Production at LHC

LHC is a proton (and heavy ion) accelerator and therefore to understand the search for a Higgs boson the contents of the proton must be briefly explored. The proton contains three “valence” quarks: two up quarks and one down quark. Each has a different color to compose a colorless baryon with electric charge +1. These quarks are immersed in a pool of gluons and “sea” quarks. The “sea” is made of quark, anti-quark pairs which are continuously created and destroyed through gluon splitting and recombination. Partons generally refer to the internal components of hadrons, (i.e. gluons and quarks). These partons are characterized by the fraction of longitudinal momentum, x, and momentum transfer, Q, in an interaction with another external particle. Parton distribution functions, *pdf*s, describe the probability to find a parton at momentum fraction x in an interaction with momentum transfer Q. Deep inelastic scattering experiments are used to probe the contents of the proton in order to map *pdf*s denoted by $f\left(x, Q^2\right)$. Figure 4.1 summarizes the *pdf*s as a function of x for all relevant partons for two values of Q^2 [50].



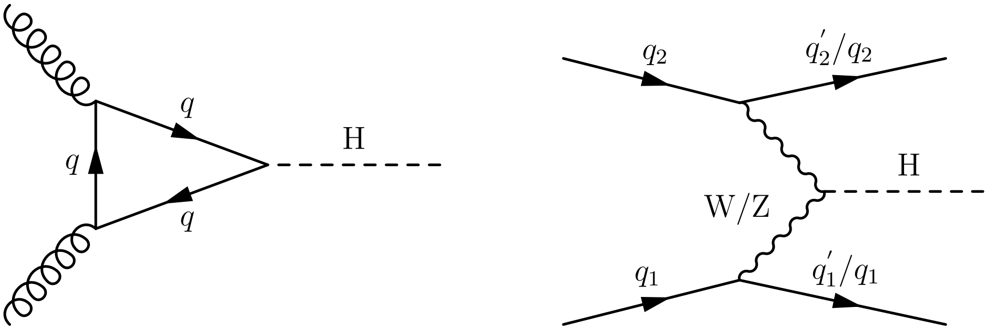
(a) (b)

Figure 4.1: Probability density functions of quarks and gluons in protons. $xf\left(x, Q^2\right)$ for quarks and gluons are plotted as a function of longitudinal momentum, x, for two different values of Q^2 [50].

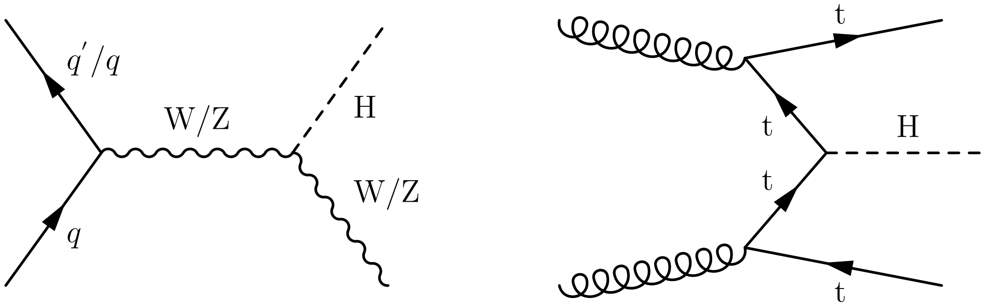
The core of a collision is described by the scattering of gluons and quarks. The primary production mode of Higgs bosons at the LHC is expected to be via Gluon Fusion (GF), whose Feynman diagram is depicted in Figure 4.2a. The other predominant production modes are Vector Boson Fusion (VBF), Higgsstrahlung (VH), and $\mathrm{t\bar{t}H}$. These states have other final state objects which can be selected in addition to the decay products of the Higgs as can be seen from their Feynman diagrams in Figure 4.2.

The cross sections of these processes depend on $\mathrm{M_H}$. The cross sections with theoretical uncertainty are shown in Figure 4.3 as a function of $\mathrm{M_H}$ for 7 TeV and 8 TeV proton-proton collisions. The values of these production cross sections can be found in Appendix A.

Once a Higgs boson is created there are many possible decay channels. The

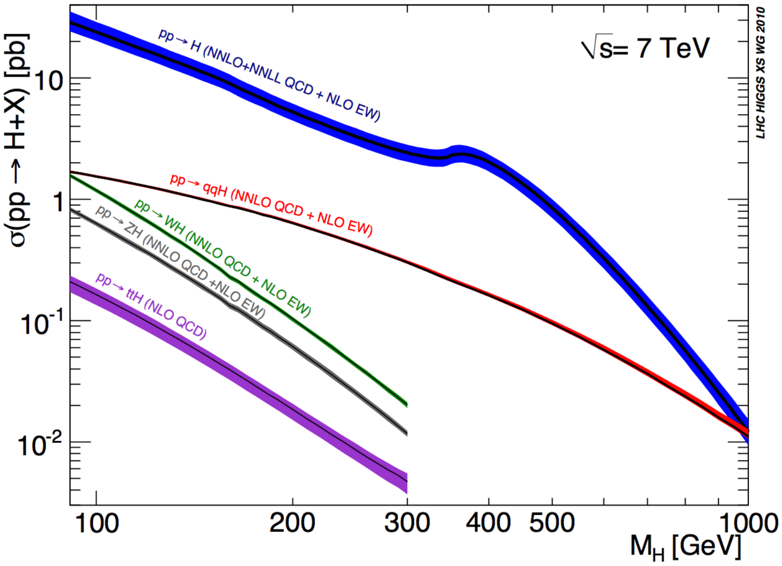


(a) Gluon Fusion (b) Vector Boson Fusion

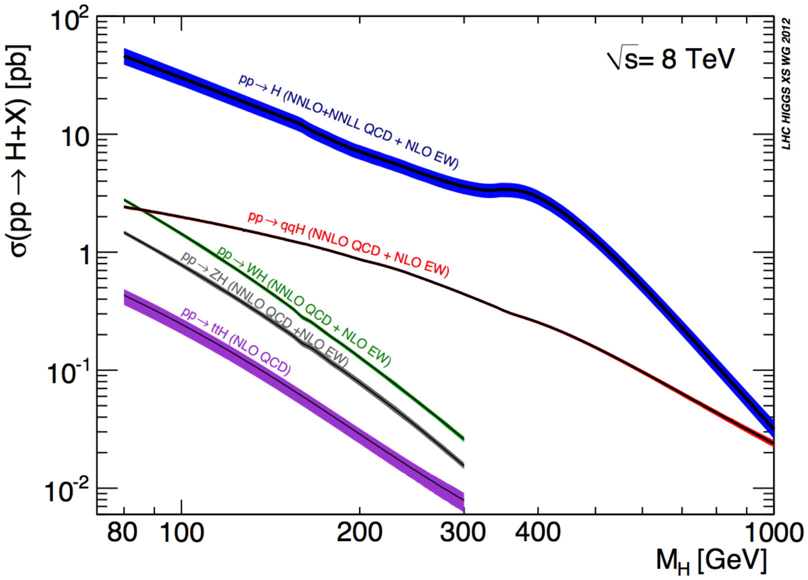


(c) Higgsstrahlung (d) $t\bar{t}$ Production

Figure 4.2: Higgs production Feynman diagrams at LHC. These are the Feynman Diagrams for the four Higgs production mechanisms of expected significance at the LHC. In (a) there is the primary diagram for production which is gluon fusion via a fermion loop. In (b) the diagram for Vector Boson Fusion is depicted. Higgs produced via associated production with a vector boson is shown in (c), and the smallest relevant contributions come from Higgs produced in association with two t quarks in (d).



(a) $\sqrt{s} = 7$ TeV



(b) $\sqrt{s} = 8$ TeV

Figure 4.3: Standard Model Higgs cross sections for 7 and 8 TeV proton-proton collisions at LHC as a function of M_H.

Higgs boson can also decay to two photons or one photon and one Z boson via fermion (mainly t) loops or a W boson loop as in Figure 4.4. The probability of decaying to a channel (known as its branching fraction or ratio) is also a function of M_H. Also, given that the Higgs boson's coupling to fermions is proportional to fermion mass, heavier fermion decays, which are kinematically allowed, are more frequent. Various channels are relevant for low values of M_H (110 − 150 GeV) which can be seen in Figure 4.5a. $H \to \gamma\gamma$ and $H \to ZZ \to 4\ell$ are among these channels, which have special significance because of their high mass resolution. For high values of M_H, decays to WW and ZZ are the most relevant as can be seen from Figure 4.5b.

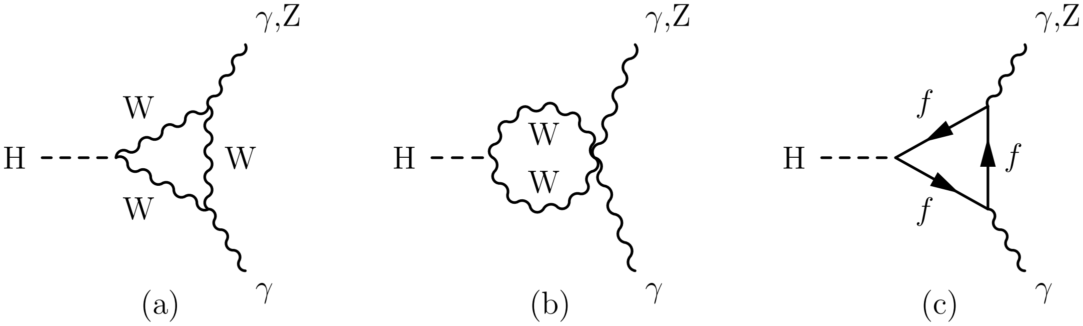
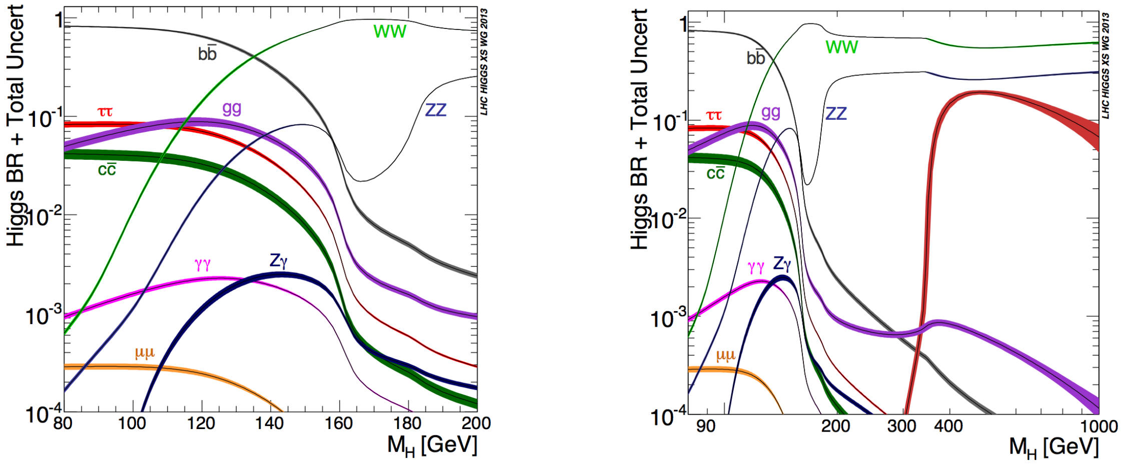


Figure 4.4: Leading order Feynman diagrams of $H \to \gamma\gamma$ decays. These Feynman diagrams can also result in photon plus Z final states.

4.2 First $H \to \gamma\gamma$ Analysis and Its Publication Using Full 2011 7 TeV Dataset

At the end of October 2011 the last proton-proton collisions were recorded for the year and after six months of debate and discussion it was time to produce the first results. Problematically we had several analyses available, but a consensus emerged around using $H \to \gamma\gamma$ analysis like that in the PTDR as our "baseline" analysis strategy. The idea of a mass-dependent BDT-driven analysis had been hotly debated (eventually becoming a cross-check analysis), but at that time, a simple analysis was preferred as the primary analysis of the CMS $H \to \gamma\gamma$ group.



(a) (b)

Figure 4.5: Standard Model Higgs decay factions as a function of $\mathrm{M_H}$. In (a) the various low mass Higgs decays are apparent, whereas in (b) it is clear that for any SM Higgs with $\mathrm{M_H} > 200$ GeV decays to ZZ and WW are dominant. To compare the $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ channels the double leptonic decay of ZZ must be taken into account using the branching ratio of $\mathrm{Z} \rightarrow \mathrm{e^+e^-}$ and $\mathrm{Z} \rightarrow \mu^+\mu^-$. The double leptonic decay (only to electrons or muons) is only 0.45% of all $\mathrm{H} \rightarrow \mathrm{ZZ}$ events. Fortunately, $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ has very excellent mass resolution and extremely low background, but its rate is small compared to $\mathrm{H} \rightarrow \gamma\gamma$.

This “baseline” was based on fits to the data for background estimates (just like both final analyses described later) and the selection uses a cut-based identification very similar to that used in the final Cut-based Analysis described in Section 6.2.1. Indeed, the event classification ($2|\eta| \times 2R_9$ classes) is nearly the same as that analysis as well[2]. The Cut-based Analysis has become the standard to which all other analyses must be compared.

In addition, a single di-jet tag event class targeting VBF production–developed by me in collaboration with a few others in the summer of 2011–was an important component of the analysis. It improved the analysis exclusion sensitivity by 10% and made it possible to measure the Higgs boson’s coupling to weak bosons. This tag paved the way for several exclusive tags targeting VBF, VH and $\mathrm{t\bar{t}H}$ production channels. This analysis also had a great impact in an alternative Higgs model, in which the Higgs boson only interacts and gives mass to vector bosons. This is called the Fermiophobic Higgs model (described in Section 4.3).

4.2.1 Cut-based Strategy and Optimization

The choice of fitting the data and using those fits to estimate the background is a fundamental aspect of this analysis. Beyond that the selection of the photons and classification of di-photon events are arguably the next most important analysis elements. The photon identification (ID) used for the first publication is a *Cuts in Categories* (CiC) photon ID [51][3].

The idea is fairly simple: construct categories of photons that are very different in resolution and S/B and set cuts in the same variables for each category independently. Each category uses the same figure of merit (S/B) for the cut-setting but arrives at different thresholds in the different categories. Resolution is not taken into account during the cut-setting. That is, all events in a large invariant mass window (uniform for all categories of photons) are used for background. This is done for several different values of S/B although only one is used in the $\mathrm{H} \rightarrow \gamma\gamma$ analysis.

---

[2] Di-photon p_T classification adds much less sensitivity with the 7 TeV dataset alone so it was not implemented at that time.

[3] This citation is for an electron CiC ID, but the methodology is the same.

Fundamental to the CiC cut-setting procedure are $n-1$ plots of background to signal (B/S) estimates in each of the variables of the ID for each category[4]. An example is depicted in Figure 4.6. In $n-1$ plots, all other cuts are applied except for the cut on the variable being plotted. That variable is mapped into another variable such that the signal distribution is rigorously flat. The background is plotted as a function of this transformed variable. A model to the background is then fit. Afterward the cut can be quite accurately set at the value for which all events have B/S less than a predetermined value of B/S.

This procedure is performed for all variables in their respectively $n-1$ plots. That collection of cuts is the selection used in the next iteration of $n-1$ plots. After several iterations the values of the cuts change very little and the procedure ends. The final values will be the cuts associated to that S/B. To obtain multiple working points, the S/B threshold is increased and the final values of all variables from the previous working point are used as initial values. Thresholds are only allowed to tighten or stay the same in all variables. We define many different working points and name them “Loose”, “Medium”, “Tight” and so on as thresholds tighten and purity increases. The CiC selection working point used in this analysis is the so-called “Super Tight” selection.

The variables used in this optimization are very similar to those used in Section 6.2.1 except for one important feature. The isolation used in this analysis was detector-based, not yet particle flow based (see Section 5.7 for a particle flow description). Particle flow algorithms were being developed at the time and were unavailable for the very first CiC ID. Particle flow isolation was known to have better performance in general and work better in high pile-up events than detector-based isolation. However, since pile-up in 2011 was relatively low (see Figure 5.2), detector-based isolation was acceptable[5].

The classification of events was constructed using the photons' R_9 and $|\eta|$ as well. Events, in which both photons are in the best photon categories (both in

---

[4] This procedure is independent of the number of categories. For simplicity one could imagine this being done with only one category.

[5] The expected pile-up in 2012 was always so large that detector-based isolation was less effective.

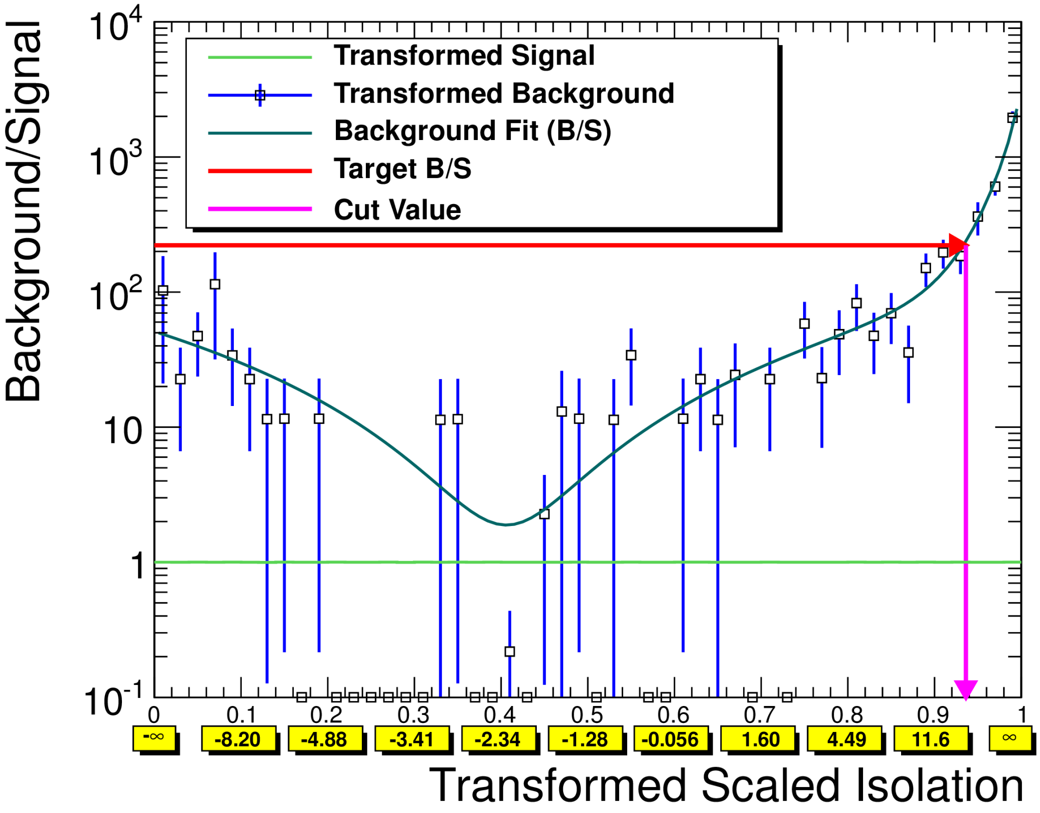
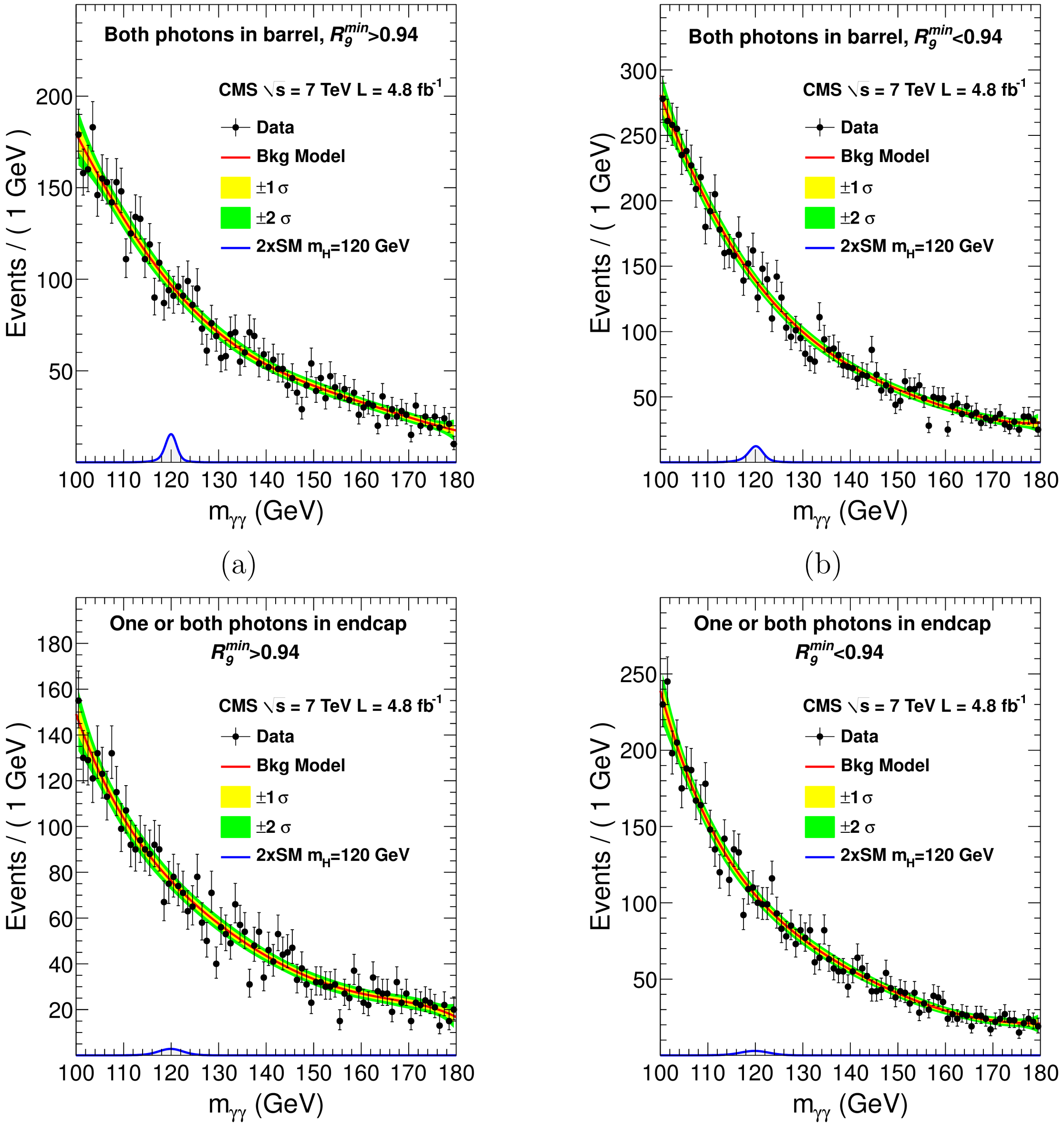


Figure 4.6: $n-1$ plot example. Variable selections are applied to the other $n-1$ variables. The variable plotted above, which is $\mathrm{PFIsoSum}_{\mathrm{ChosenVtx}}$, has had no selection applied on it. The variable has been transformed so that the signal is strictly flat, which is evident as the signal is plotted in the green histogram. The boxes in yellow show the mapping back to the original variable. The signal is normalized to the number of bins so that signal is flat at one. The background is plotted in this transformed, normalized variable and then a polynomial is fit to the background for the purpose of smoothing. The fit function is effectively the background to signal ratio (B/S) as a function of the transformed variable. For each CiC working point a fixed, pre-determined B/S value (in red) is selected. The cut value is set by determining the value of the transformed variable (in pink) at which the B/S function is equal to the target value of B/S. This is done for each variable.



(a)

(b)

(c)

(d)

Figure 4.7: First publications inclusive invariant mass distributions. Events have been selected requiring only two photons passing the Super Tight CiC photon ID and classified into η-R_9 event classes.

EB and both high R_9), make up the most sensitive event class. The next most sensitive event class is for both photons in the EB but one or both photons having converted ($R_9 < 0.94$). Two similar event classes are defined for the case when either selected photon is found in the EE. Figure 4.7 shows the selected data with background fits and signal models with $\mathrm{M_H} = 120$ GeV for these event classes.

4.2.2 Di-jet Tag Targeting VBF

The typical VBF signature can be seen in Figure 4.8. It shows two forward jets ($\eta_1 = -2.022$ and $\eta_2 = 1.860$) with two high p_T, central barrel photons.

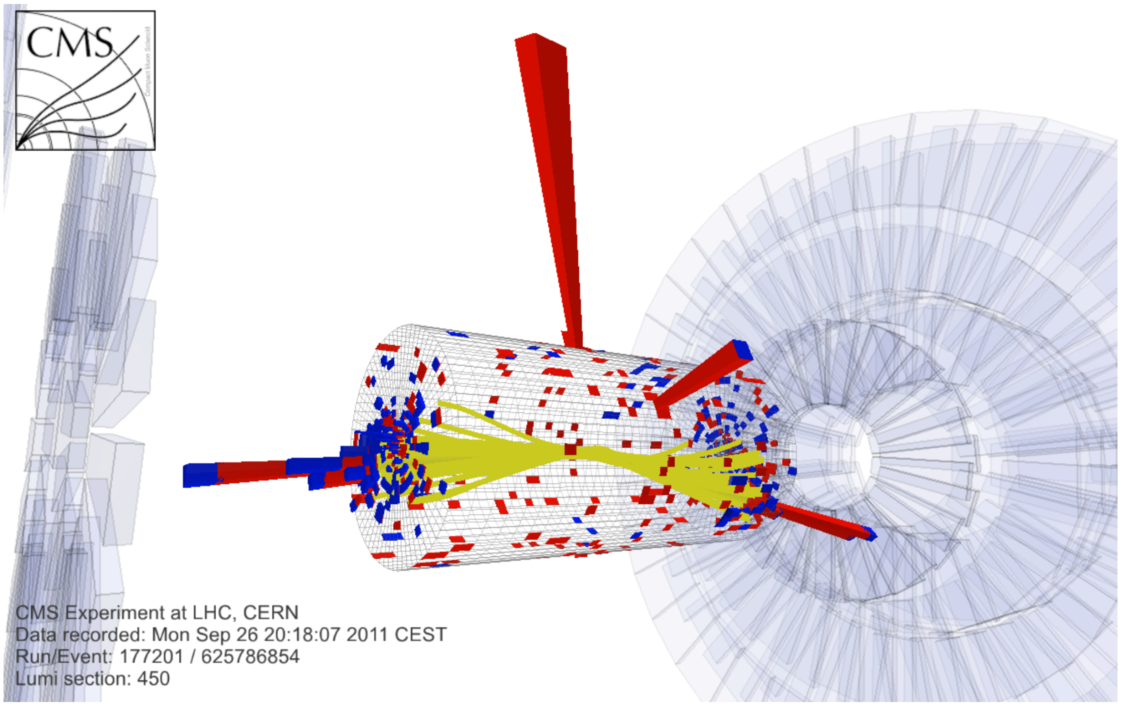
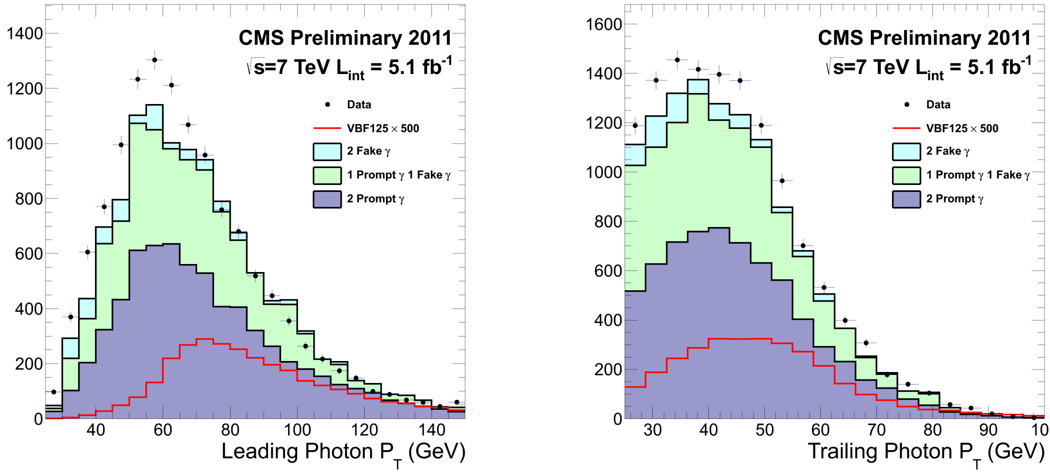


Figure 4.8: Di-photon plus di-jet candidate event. This 3D event display contains a 2011 data event, which is a di-jet tagged event with $m_{\gamma\gamma} = 121.9$ GeV. This event contains two well-isolation central barrel photons with very high p_T (193.9 and 78.0 GeV). The tracks in yellow are part of two very high p_T jets which are well separated in η. The di-jet invariant mass is 1460 GeV.

The selection of VBF events requires a careful study of jet and di-jet kinematic variables as well as a re-evaluation of photon kinematics. Since most SM Higgs bosons are produced at the LHC via Gluon Fusion, the inclusive selection of photons mainly targets these photons. However, VBF-produced Higgs bosons have on average greater di-photon p_T than Gluon Fusion produced Higgs bosons. Therefore, the distribution of p_T for the leading photon–that is the photon with larger p_T–is shifted to greater p_T. Cutting tighter on the leading photon is an excellent way to take advantage of large di-photon p_T kinematics. On the other hand, the trailing photon's p_T distribution is wider and peaks at lower p_T. Thus, the cut must be loosened to keep high efficiency. The preference for more asymmetric cuts is generally true for all exclusive production channel tags for VBF, VH and t$\bar{\text{t}}$H processes because the Higgs boson has greater p_T in all of these production modes. All exclusive channel tags take advantage of this feature in the final analysis as can be seen in Section 6.4.



(a) (b)

Figure 4.9: Photon p_T after minimal di-photon+di-jet selection. The kinematics of boosted H $\rightarrow \gamma\gamma$ events encourage even more asymmetric cuts on the photons' p_T than in the inclusive selection. In (a) the p_T distribution of 7 TeV data and background for the leading photon peak at much lower p_T than the signal in red. The distribution of the trailing photon's p_T in (b) shows less separation between signal and background.

Figure 4.9 shows the p_T distributions of the two photons after applying the CiC ID and minimally selecting di-jet events in 7 TeV data, background MC and

VBF signal MC with $\mathrm{M_H} = 125$ GeV[6]. The VBF signal is scaled by a factor 500 to show features more easily. The leading photon's p_T selection is actually $p_T/m_{\gamma\gamma}$ (at the threshold of 55/120) to avoid deforming the typical mass distribution shape. That is, a tight, uniform cut on leading p_T will remove much more background at lower values of $m_{\gamma\gamma}$. This can cause the mass distribution to be flatter at low $m_{\gamma\gamma}$ or actually create a broad, artificial peak in the background. This significantly complicates the background fits to the data. Adjusting the threshold relative to the mass (looser at lower mass, tighter at higher mass) generally maintains the expected background shape, which is always decreasing with positive concavity. The trailing photon's p_T cut in this analysis is at 25 GeV, which is the minimum that it can be before trigger inefficiency reduces signal efficiency.

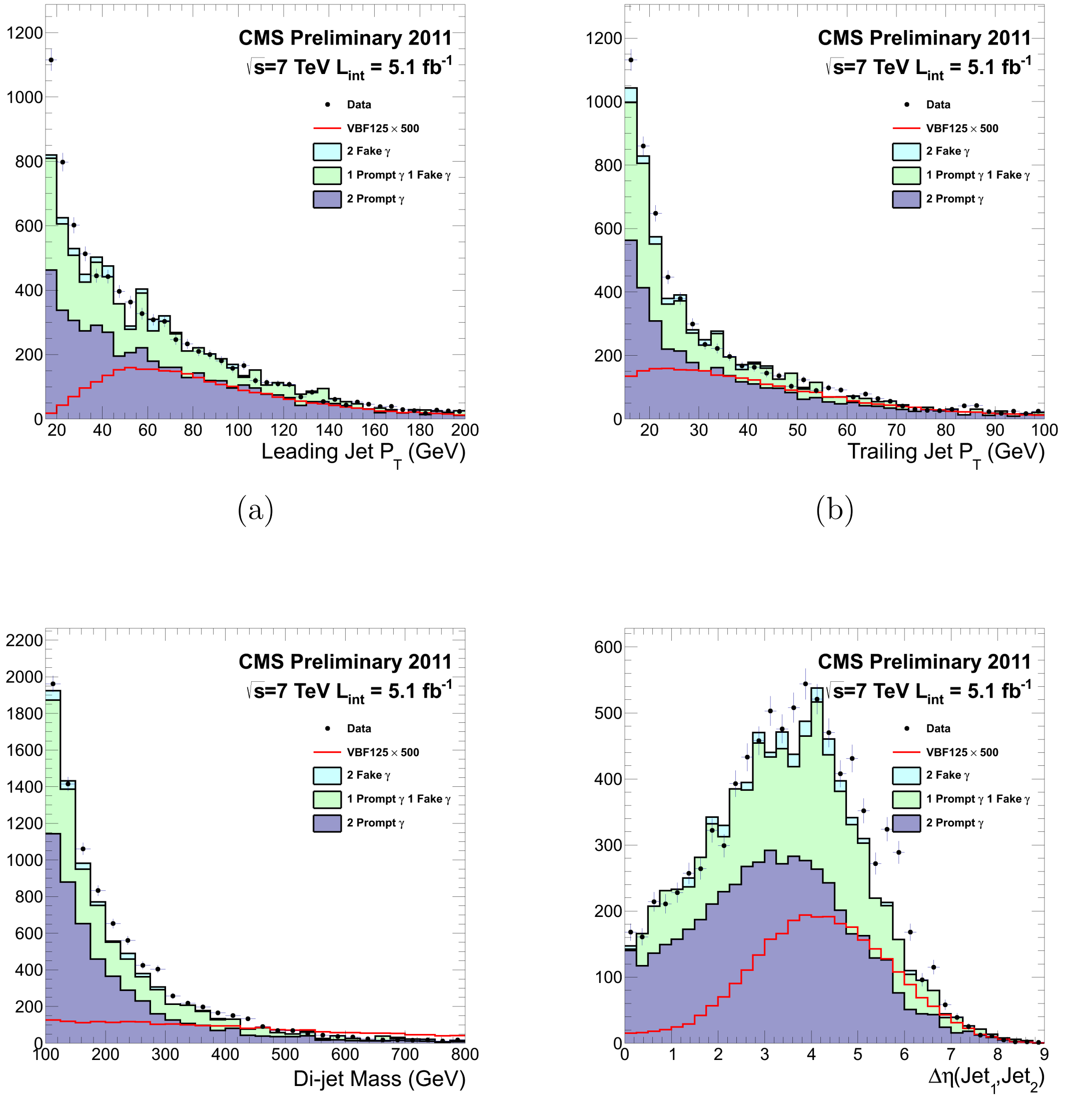
After two photons pass all the CiC ID and p_T (or $p_T/m_{\gamma\gamma}$) cuts, the di-jet selection begins. The jets must have $\Delta\mathrm{R}\,(\gamma, \mathrm{J}) > 0.5$ from both photons to avoid selecting photons as jets. Figure 4.10 shows the di-jet p_T distributions, di-jet invariant mass and $\Delta\eta$ between the two jets after minimal jet p_T requirements. Since the jets originate from separate protons, they tend to have high separation in η. Therefore, $\Delta\eta$ between the two jets is a very good variable for distinguishing between signal-like and background-like events. The cut value of 3.5 was selected. The di-jet invariant mass, M_{JJ}, is the most discriminating variable but also tight selection on M_{JJ} reduces the signal efficiency. Therefore, the cut value of 350 GeV was selected as the compromise between S/B and signal efficiency. The jet p_T requirements are 30 and 20 GeV.

Two variables are used to ensure the compatibility of the di-jet and di-photon systems. The average η of the di-photon and di-jet should be compatible. The so-called Zeppenfeld variable, $\eta_{\gamma\gamma} - \frac{\eta_{J_1}+\eta_{J_2}}{2}$, is a measure of the η compatibility that peaks near 0. The cut value was selected to be 2.5.

The $\Delta\phi\,(\gamma\gamma, \mathrm{JJ})$ variable is the azimuthal difference between the di-photon and di-jet composite objects. It should peak near π because of the Higgs' recoil against the quarks. The width near π is due to the radiative QCD corrections (i.e. gluons emitted from initial state partons and final state quarks). This variable was

---

[6] These plots are produced with the final 7 TeV data and MC for the purpose of this dissertation.



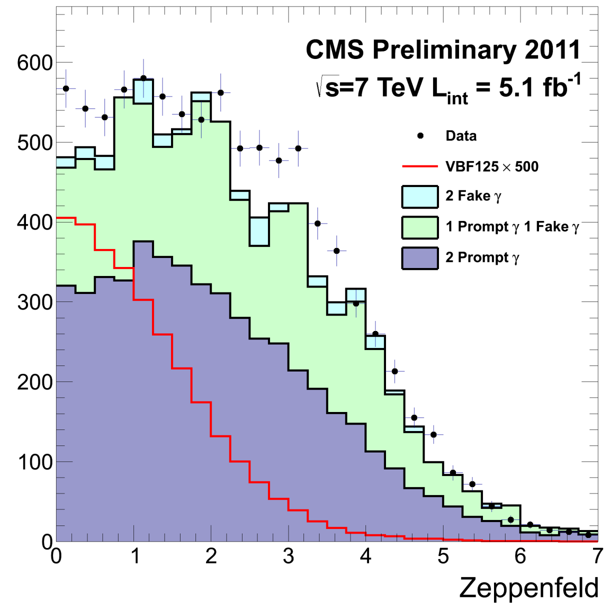
(a)

(b)

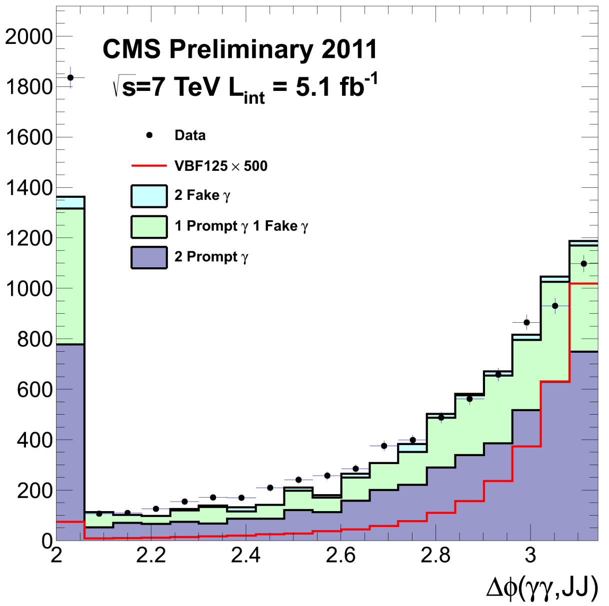
(c)

(d)

Figure 4.10: Di-jet kinematic variables. These are the primary di-jet kinimatics used to separate signal from background in data.



(a)



(b)

Figure 4.11: Di-jet and di-photon compatibility variables. The Zeppenfeld variable (a) peaks near 0 showing di-photon and di-jet compatibility in η. The difference in azimuth (b) represents di-photon (Higgs) recoiling against the di-jet (quarks). The underflow is represented in (b) in the first bin to show further background rejected with this selection.

first proposed by the author and later extensively used in CMS and also in ATLAS. Loose selection on this variable was set requiring $|\Delta\phi| > 2.6$. In addition, this variable rejects di-jet events from pile-up because those jets should have essentially random orientation with respect to the di-photon. Figure 4.11 shows both of these variables after full di-photon selection and minimal di-jet selection.

Figure 4.12 shows the di-jet tagged event class with the full selection applied. Near 124 GeV there is a visible and locally significant excess above the expected background. Figure 4.13 shows the primary statistical results of the publication. The $\mathrm{CL_S}$ limits show that the expected exclusion sensitivity is between 1.4 and 2.4 times σ_{SM} depending on $\mathrm{M_H}$ hypothesis. More importantly in Figure 4.13b the local excess is quantified as 3.1σ. In hindsight, this is the first hint of the Higgs boson to be discovered. At the time the look-elsewhere-effect [52] in the search range of 110-150 GeV reduced this significance to 1.8σ. The final line of the abstract of this publication deferred interpretation to future data, suggestively stating, "More data are required to ascertain the origin of this excess."

4.2.3 SM Higgs Combination with Full 7 TeV Dataset

Various other Higgs boson decay channels were sought by CMS in 2011 using the full 2011 dataset. At the beginning of 2012, the first combination of all relevant channels was published by CMS. CMS's high mass searches ($\mathrm{M_H} > 200$ GeV) relied on several $\mathrm{H} \rightarrow \mathrm{ZZ}$ channels, where one Z boson decays to two electrons or muons and the other Z boson has two electrons, two muons, two jets or very high E_T^{Missing} indicative of two neutrinos. The $\mathrm{H} \rightarrow \mathrm{WW}$ channel where both W bosons decay to an electron or a muon and high E_T^{Missing} from two neutrinos was quite important in the mass range $150 < \mathrm{M_H} < 200$ where WW is the dominant Higgs boson decay channel. With these analyses CMS ruled out the possibility of a SM Higgs boson with mass $127 < \mathrm{M_H} < 600$ GeV at the 95% confidence level [53]. Figure 4.14 depicts these exclusion limits.

These findings also show the first hint of an excess of events in the region around 124 GeV. Driven by an excess in the $\mathrm{H} \rightarrow \gamma\gamma$ channel a combined local significance of 3.1σ was determined by CMS. Figure 4.15 shows the observed sig-

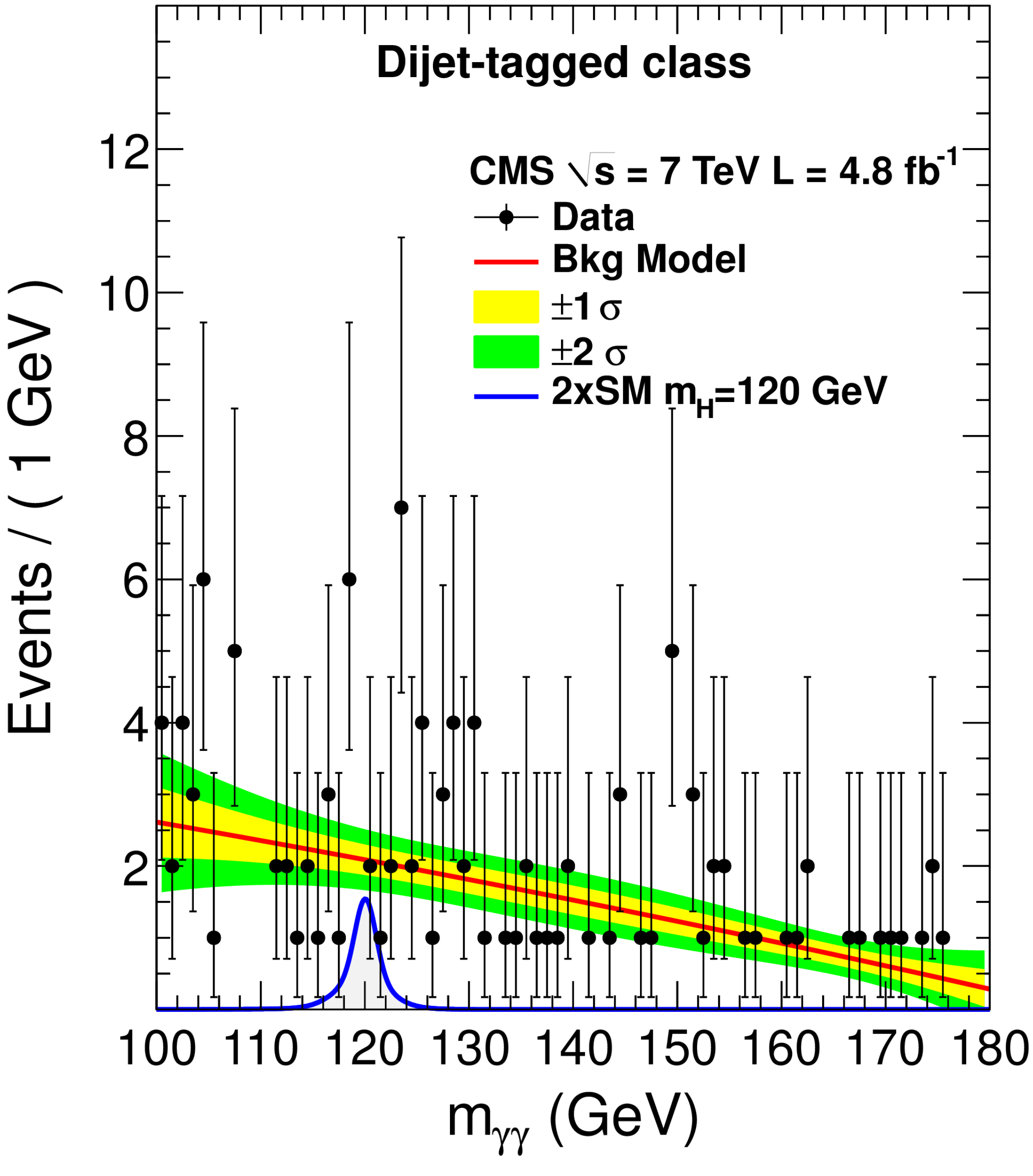
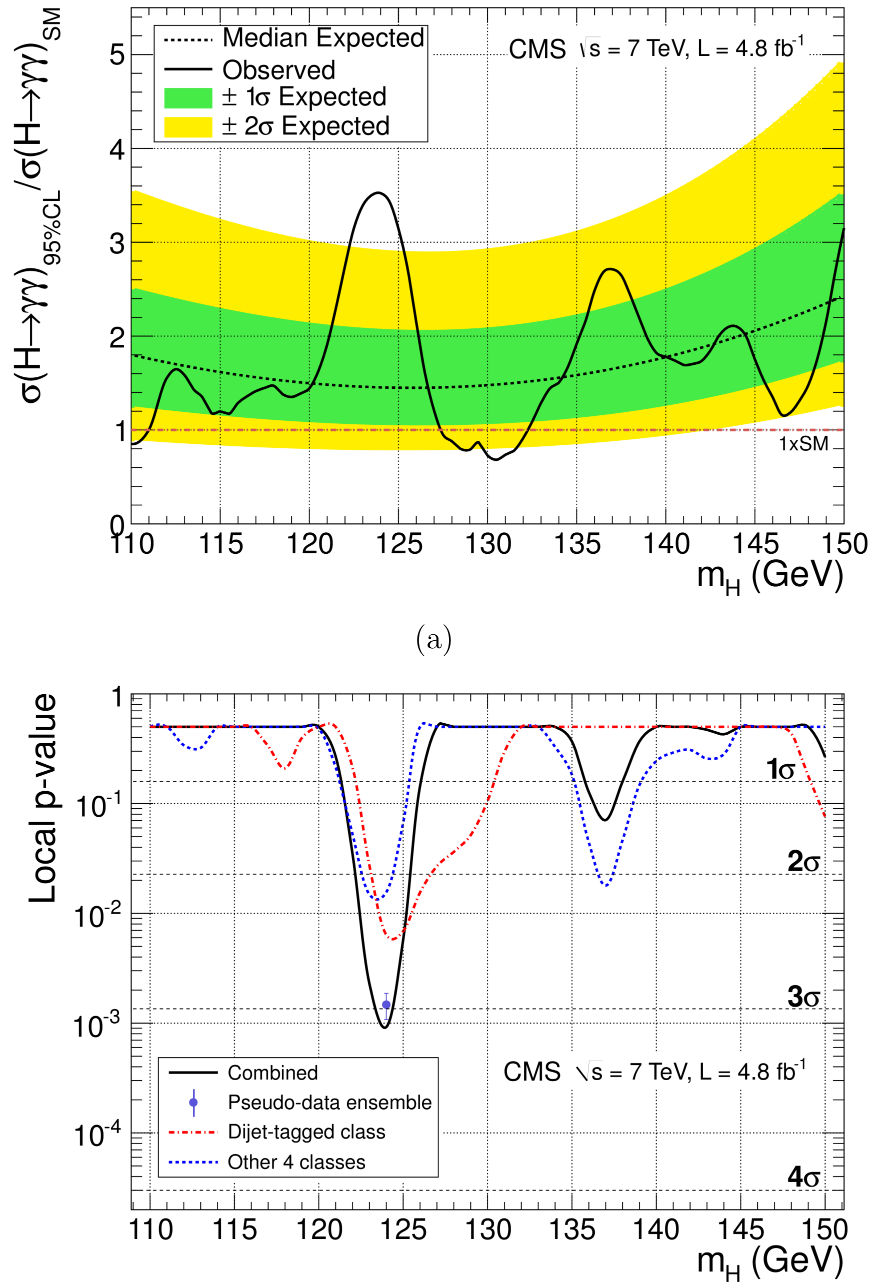


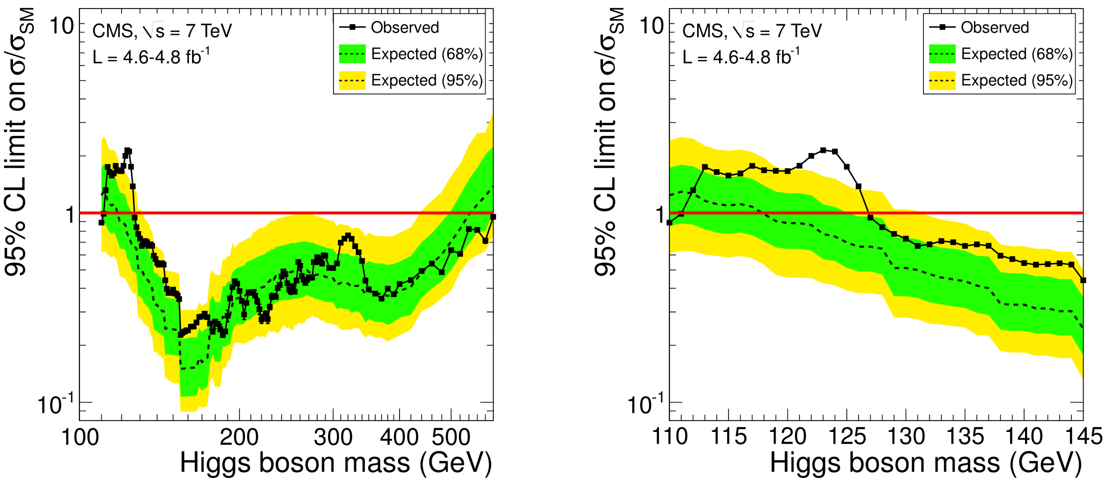
Figure 4.12: Di-jet event class of first H → $\gamma\gamma$ publication. An excess of events in near 124 GeV has a local significance of 2.4σ.



(a)

(b)

Figure 4.13: CL_S exclusion limits and local excess significance of first H → $\gamma\gamma$ publication. The analysis is not yet sensitive to the SM Higgs boson cross section. The first published exclusion limits of the SM Higgs in the H → $\gamma\gamma$ channel from CMS are determined in (a). However, more significantly the first quantification of the 3.1σ local excess near 124 GeV is published. This excess, including the large excess in the di-jet channel, has been been a key feature of every CMS H → $\gamma\gamma$ analysis published.



(a) (b)

Figure 4.14: CMS combined SM Higgs boson search exclusion limits with full 7 TeV dataset. The 95% CL upper limits on the signal strength for the SM Higgs boson hypothesis as function of the Higgs boson mass.

nificance of each channel and of the combination of all channels. ATLAS published similar results and found an excess of events near 126 GeV with local significance 3.5σ [54]. Although the local significance of both experiments at approximately the same mass is quite compelling, the global significance is diminished because the original search range was very large search range. The application of the look-elsewhere-effect [52] reduces the global significance of these excesses and both experiments deferred to the analysis of future (2012) data to determine the nature of the excess near 125 GeV.

4.3 Search for a Fermiophobic Higgs Boson with 2011 Data

The Fermiophobic Higgs model assumes that the Higgs boson either does not interact with or interacts weakly with fermions. Essentially this removes Gluon Fusion (Figure 4.2a) and t$\bar{\text{t}}$H (Figure 4.2d) production modes. Although the total Higgs boson production rate is reduced by a factor of 10, VBF and VH production are unchanged. In addition, branching ratios change because of the absence of

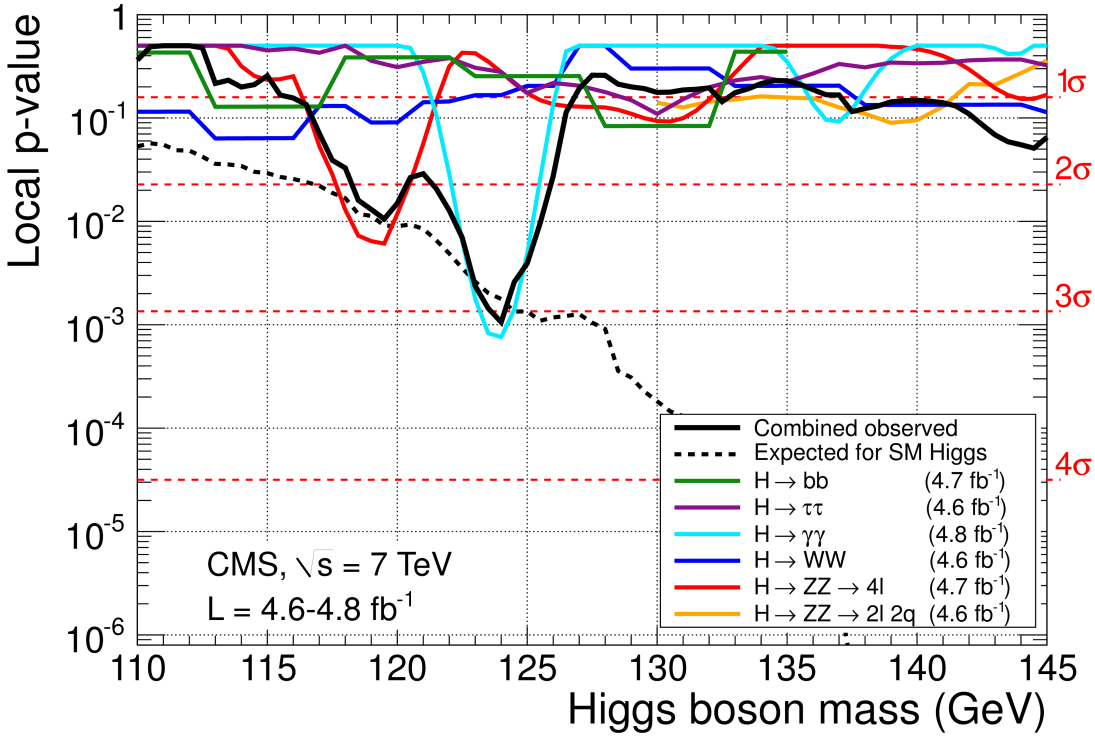


Figure 4.15: CMS low mass Higgs boson search combination p-values with full 7 TeV dataset. The combined observed local p-value as a function of M_H in the range $110-145$ GeV is the solid black, while the expected p-value of a SM Higgs boson is shown as a dashed black line. Individual channels are shown as colored solid lines. The excess in the $H \to \gamma\gamma$ channel is the primary contributor to the combination's local significance of 3.1σ near 124 GeV. Considering the look-elsewhere-effect in the search range above, the excess has a global significance of 2.1σ.

fermionic Higgs boson couplings and because the intermediate charged loop between the Higgs boson and the photons can no longer contain fermions but only the W boson (see Figure 4.4). At low mass the H $\rightarrow \gamma\gamma$ branching ratio increases substantially. For example, at M_H = 125 GeV, the branching ratio is 6.7 times larger for a Fermiophobic Higgs boson compared to a SM Higgs boson. Figure 4.16 shows the Fermiophobic and SM production cross section times branching ratios for allowed Fermiophobic decays.

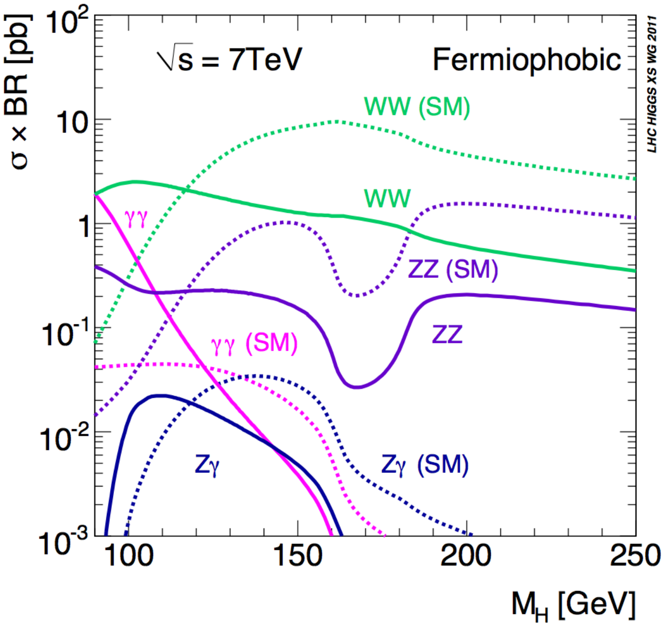


Figure 4.16: Fermiophobic and SM Higgs model cross section times branching ratios of Fermiophobic Higgs boson decays. The only allowed couplings are to Z and W bosons and so there are only a few final states. For low values of M_H H $\rightarrow \gamma\gamma$ is very sensitive because of the very large amount of expected signal.

$\sigma \times$BR in the two photon channel is much larger at lower mass until near 120 GeV where the SM and the Fermiophobic models have equal $\sigma \times$BR. The H $\rightarrow \gamma\gamma$ analysis is very sensitive to a Fermiophobic Higgs boson for this reason. Moreover, since all of the Fermiophobic Higgs bosons would be produced via VBF or VH, the di-jet tag for VBF and muon/electron tags for VH are very sensitive components of a Fermiophobic Higgs boson search. Whereas in a low mass SM Higgs boson search, lepton tags were completely insensitve with the 2011 luminosity.

4.3.1 Fermiophobic $\mathrm{H} \rightarrow \gamma\gamma$ Analysis with Di-jet and Lepton Tags

The same di-jet tag described in Section 4.2.2 is extremely sensitive to a Fermiophobic Higgs boson. Therefore, between late 2011 and the summer of 2012 it was a serious consideration that the excess in the $\mathrm{H} \rightarrow \gamma\gamma$ di-jet event class indicated the Fermiophobic character of the Higgs boson. Moreover, although this tag appears in publication combined with an inclusive analysis and lepton tags, it achieves significant lower limits to the Fermiophobic Higgs boson mass by itself. The excess made the observed limits on the mass of a Fermiophobic Higgs boson much weaker than the expected near 124 GeV[7] as shown in Figure 4.17. Using only the di-jet tag, the observed (expected) lower limit on the Fermiophobic Higgs mass at the 95% confidence level is 122 (128) GeV.

In addition, because of the large amounts of VH production in the Fermiophobic model, muon and electron tags targeting leptonic decays of associated W and Z contribute a significant improvement to this analysis.[8]

The muon tag was fairly straight-forward to develop: simply select a well-identified muon in addition to two photons. The electron tag posed a more interesting problem because of electrons' ability to fake photons. That is, Drell-Yan events leak into the electron tag when there are actually two electrons and one photon in the event. One of the electrons is tagged as the electron and the other is selected as a photon. Indeed, the problem is slightly more insidious than it first appears because electrons fake photons at a higher rate in data than in MC. Therefore, features of fake photons from electrons are not as obvious in MC.

I worked on analysis solutions to rejecting fake photons from electrons in 2011 and 2012. I proposed several techniques, which are still used to reject Drell-Yan events in electron tagged events. Two main strategies for rejecting these events were studied and implemented.

---

[7] Recall that these exclusion limits were derived before the Higgs boson's discovery and therefore the Higgs boson is not a part of the known processes in the background.

[8] The SM analysis was insensitive to this channel with the available luminosity and so the lepton tags were only added to the Fermiophobic analysis at that time. However, the lepton tags are part of the final SM analysis and many fundamental aspects of the analysis were developed by me for the purpose of this Fermiophobic analysis.

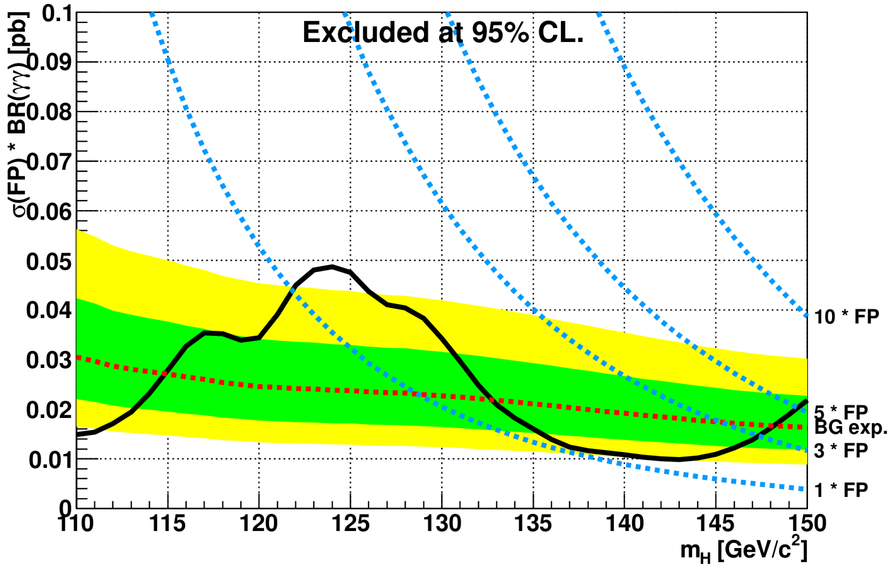


Figure 4.17: Fermiophobic Higgs model exclusion limits in $\sigma \times \mathrm{BR}(\mathrm{H} \rightarrow \gamma\gamma)$ using only the di-jet tag analysis. The dashed blue curves are multipliers (1, 3, 5 and 10) of $\sigma \times$BR for Fermiophobic Higgs models. The 95% confidence level on the lower limit of Fermiophobic Higgs boson mass is expected at 128 GeV where the dashed red line (median exclusion expectation) meets the dashed blue $1\sigma \times$BR line. However, the excess of events near 124 GeV makes the observed exclusion (solid black line) weaker–only at 122 GeV. Green and yellow bands are the 1σ and 2σ exclusion bands.

The first strategy was to reject photons that are actually electrons using stronger electron veto requirements. The primary issue here is that in reducing the fake rate via tighter selection signal efficiency decreases, and in a channel with relatively little signal, very tight cuts can render that tag insensitive. After comparing the impact of several vetoes, one was clearly better than the others.

Missing hits in the electron's track in the inner most layers of the pixel tracker can indicate that the electron originated from a photon conversion. Conversely, a reconstructed electron with pixel hits throughout all the innermost layers is much more likely to be a real electron from the primary interaction. These electrons may start to shower in the tracker and emit photons. Rejecting photons near these electrons ($\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 1.0$) with no early missing hits significantly reduces the Drell-Yan background with a minimal impact on signal efficiency.

The second strategy involved treating each photon as if it were an electron and computing the invariant mass of the tagged electron+photon. It was observed that the trailing p_T photon with the tagged electron often reconstructed an invariant mass close to the mass of the Z boson. A window around the Z mass was selected to veto these events, which reduced this background significantly.

Finally, a problem that is common to both lepton tags is the selection of final state radiation (FSR) photons emitted from leptons. That is, when the high p_T electron or muon emits a photon, it is incorrect to select it as one of the two photons from the Higgs boson. With a selection strategy that begins with diphoton selection followed by lepton tag, signal events will fail the tag because an FSR photon will be selected as a Higgs boson decay product and the lepton will be too close to it. With this selection logic, signal efficiency suffers. Inverting the selection methodology (that is, selecting the muon or electron before the photons) helps with this problem. Seeking first the muon or electron and removing any photon near the selected lepton from the list of selectable photons reduces FSR photon selection and improves signal efficiency in lepton tagged events.

The invariant mass distributions of selected events in the di-jet tag and lepton tags are shown in Figure 4.18. The background models are from polynomial

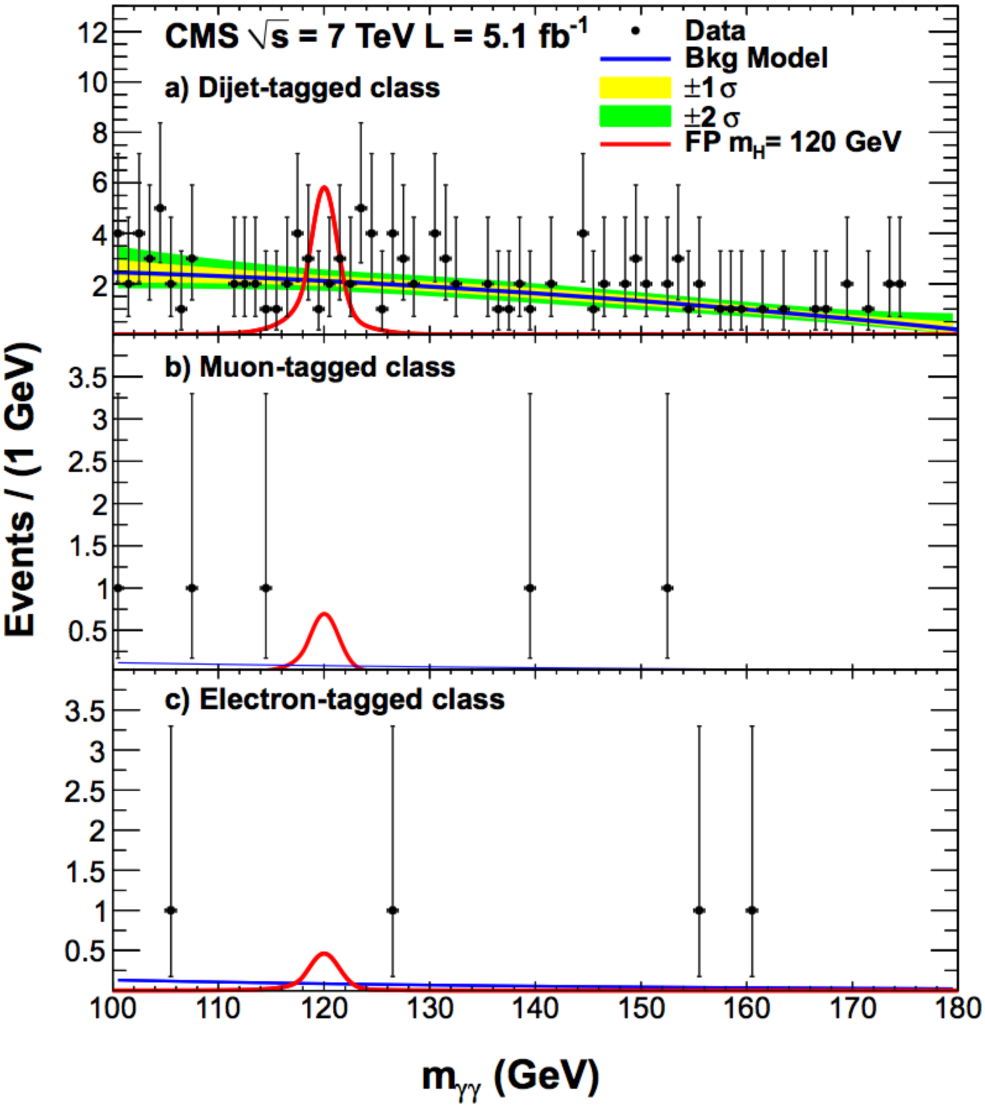


Figure 4.18: Invariant mass distributions of di-jet and lepton tag classes of 7 TeV Fermiophobic analysis. Models based on simulated signal with $M_H = 120$ GeV is shown in red. In blue is the background model from polynomial fits to the data in each event class. High S/B near 120 GeV is clear in each event class.

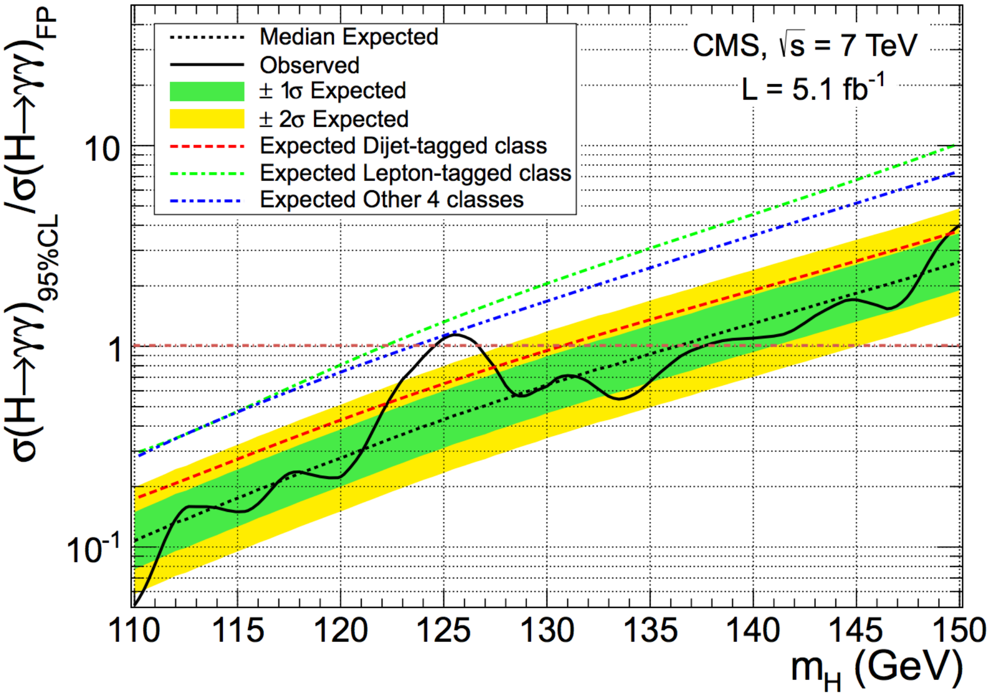


Figure 4.19: Exclusion limit on the cross section of a Fermiophobic Higgs boson decaying into two photons as a function of M_H. The observed and expected limits are shown for the combined $H \to \gamma\gamma$ analysis in black. The expected limits for the di-jet tag, combined lepton tags and untagged inclusive event classes are also shown in red, green and blue, respectively. Even though the entire 2σ band is expected to be excluded at the 95% confidence level, the di-jet excess weakens the observed exclusion limits so that it is not excluded.

fits to the data. The S/B is very high for $M_H < 125$ GeV in all these channels. In addition, a two-dimensional analysis of $m_{\gamma\gamma}$ and $p_T/m_{\gamma\gamma}$ was developed for inclusively selected events. Signal di-photons have much greater p_T on average with respect to Gluon Fusion production, which is the SM majority production mode. This analysis takes advantage of this feature.

Figure 4.19 shows the sensitivity of the di-jet tag, combined lepton tags and inclusive analysis as well as the combined $H \to \gamma\gamma$ limits. The expected exclusion limit at 95% CL covers the mass range between 110-136.5 GeV. The data excludes ranges from 110-124.5 GeV and 127-137.5 GeV. The small gap in the exclusion range is caused by the excess in the di-jet channel at 125 GeV. Its local significance in this analysis is 2.7σ.

4.3.2 Fermiophobic Combination Results

A combination of $H \to \gamma\gamma$, $H \to WW$, and $H \to ZZ$ Fermiophobic Higgs analyses was made public at the time of the Higgs boson discovery and later published [49]. The $H \to \gamma\gamma$ analysis is the most sensitive single analysis in this combination up to ~ 135 GeV, after which the expected signal yield drops very quickly (see Figure 4.16).

The $H \to WW$ and $H \to ZZ$ analyses are not dedicated Fermiophobic searches but reinterpretations of SM Higgs searches. The $H \to ZZ$ analysis contributes very little because the SM analysis (at that time) had very little dependence on Higgs' p_T and did not include significant VBF or VH production tags. However, the $H \to WW$ analysis had sensitivity near or below the Fermiophobic Higgs cross section from 100 to near 200 GeV in M_H. This sensitivity is primarily from the $H \to WW \to 2\ell 2\nu$ + di-jet tagged event class (targeting VBF production). Figure 4.20 shows the observed and expected exclusion at the 95% confidence level for each analysis and for the final combination of the three analyses. With this combination a Fermiophobic Higgs boson is excluded at the 95% confidence level in the mass range of $110 - 194$ GeV. Most significantly a Fermiophobic Higgs boson was excluded at the 95% confidence level around 125 GeV.

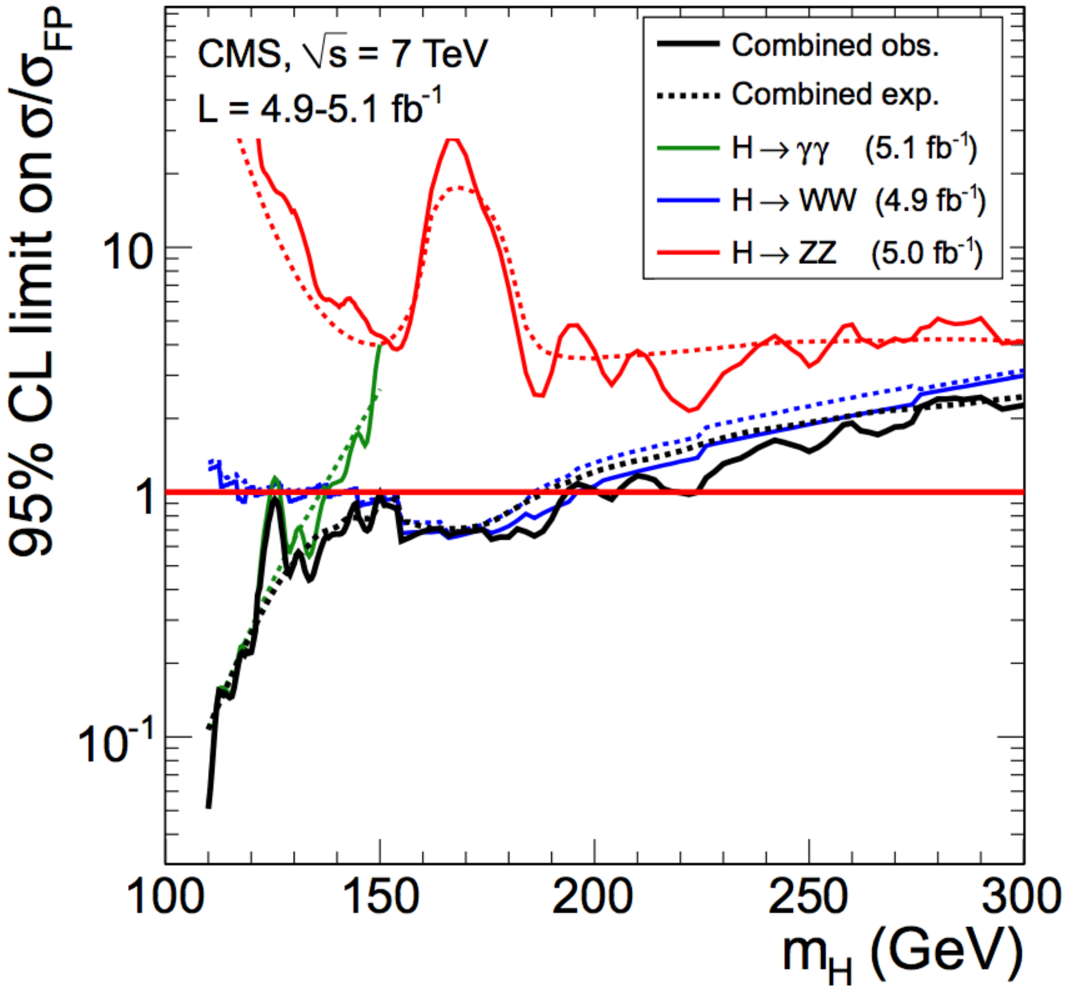


Figure 4.20: Combined Fermiophobic exclusion limits. The observed and expected 95% CL upper limits on the signal strength as a function of the fermiophobic Higgs boson mass for the three explored Higgs boson decay modes and their combination. The expected limits are shown as dashed lines and the observed limits are solid lines.

4.4 SM MultiVariate Analysis Introduction and the Discovery of a Higgs Boson

Between the publication of the 7 TeV data analysis and the summer of 2012, several aspects of the analysis began to mature and were incorporated into CMS's $\mathrm{H} \to \gamma\gamma$ group's analysis. Moreover, as the analysis developed, the SM and the Fermiophobic analyses specialized into two independent analyses. The SM analysis added an analysis with a highly optimized photon ID and event classification scheme using Boosted Decision Trees (BDTs). This is referred to as the MultiVariate Analysis, and its final version is described in some detail in Section 6.3. The Fermiophobic analysis (described in the previous section) developed differently because of physics differences in the Higgs' kinematics due to Higgs bosons being produced solely from VBF and VH processes. Both analyses benefited from improved calibration of the ECAL done between the publications.

During the international physics conferences of winter 2012, results of the SM and Fermiophobic analyses were updated and presented [55, 56]. For the discovery analysis, the SM analysis of the winter conference was used and a parallel analysis was produced for the new 8 TeV data with nearly identical methodology. Within this section, my contributions, along with some important notes and results, to the SM Higgs boson candidate discovery will be described.

4.4.1 SM Higgs Boson Discovery

The MultiVariate Analysis (MVA) was first publicly released in February of 2012 and became the primary $\mathrm{H} \to \gamma\gamma$ analysis in CMS. The background estimate is still from fits to the data, but this analysis selects very loosely on photons and uses a di-photon BDT to classify events. The photon selection in the di-jet tag had to be modified slightly to fit into this analysis. The minimum trailing photon p_T selection had to be made marginally tighter to $p_T > \frac{1}{4}m_{\gamma\gamma}$ in order to fit into the training selection of the di-photon BDT. Near $m_{\gamma\gamma} \sim 125$ GeV this is effectively an increase in the p_T cut to ~ 31 GeV. Also, the photon selection was replaced with a cut on the di-photon BDT, which tends to select photons with high photon

ID BDT values as they are inputs to the di-photon BDT.

In addition, the di-jet tag applied to the 2012 data was updated in two ways. First, jets were required to pass a cut-based pile-up rejection pre-selection (described in Section 6.1.1). This was very important because with greater pile-up (and no pile-up rejection), the signal efficiency decreased and the background efficiency increased significantly. Also, utilizing the same variables as before, tight and loose di-jet tags were produced (described in Section 6.2.3). The main variable upon which these event classes differ is in the di-jet invariant mass selection.

In summary, the $\mathrm{H} \rightarrow \gamma\gamma$ analysis used for the discovery was almost the same for the 5.1 fb^{-1} of 7 TeV data collected in 2011 and 5.3 fb^{-1} of 8 TeV data collected in early 2012. For the inclusive analysis both datasets used a di-photon BDT to classify four event classes in each dataset. This analysis was about 15% more sensitive than the cut-based analysis for inclusive event classes[9]. One di-jet event class was implemented in the 7 TeV analysis and two di-jet event classes were implemented in the 8 TeV analysis. This is a total of 11 event classes in the $\mathrm{H} \rightarrow \gamma\gamma$ discovery analysis.

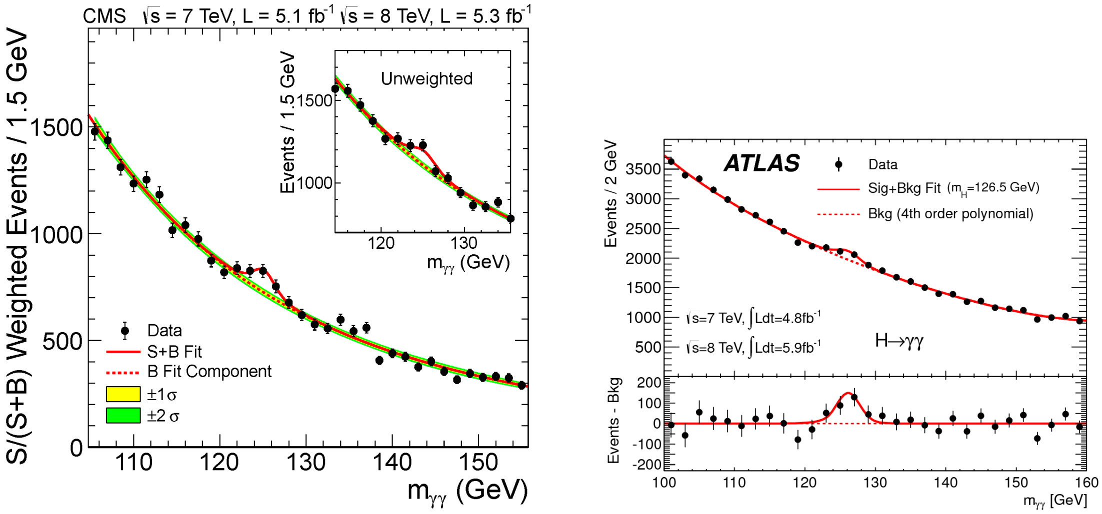
$\mathrm{H} \rightarrow \gamma\gamma$ is one of five main channels that contributes to the low mass SM Higgs boson search performed by CMS. Figure 4.5a shows that the variety of decay channels available for a Higgs boson with $\mathrm{M_H} = 125$ GeV. Although the $\mathrm{H} \rightarrow \gamma\gamma$ channel has a much lower branching ratio than $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ and $\mathrm{H} \rightarrow \tau^+\tau^-$, it is a very significant channel because of its excellent mass resolution. $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ also has excellent mass resolution when the Z bosons decay to muons or electrons. However, only 0.45% of $\mathrm{H} \rightarrow \mathrm{ZZ}$ events will have both Z bosons decaying to electrons or muons. Thus, the $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ channel has relatively low expected yield.

The $\mathrm{H} \rightarrow \mathrm{WW}$ channel is also a sensitive channel when the W bosons decay to electrons, muons or one of each and with a neutrino for each W. Unfortunately, since the final state cannot be fully reconstructed, the resolution is very poor. As previously mentioned $\mathrm{H} \rightarrow \mathrm{b\bar{b}}$ and $\mathrm{H} \rightarrow \tau^+\tau^-$ have relatively high branching ratios. However, neither was sensitive to a SM Higgs boson alone at the time because of the poor mass resolution and very large backgrounds in both channels.

---

[9] The di-jet tag event classes' sensitivity was less improved.

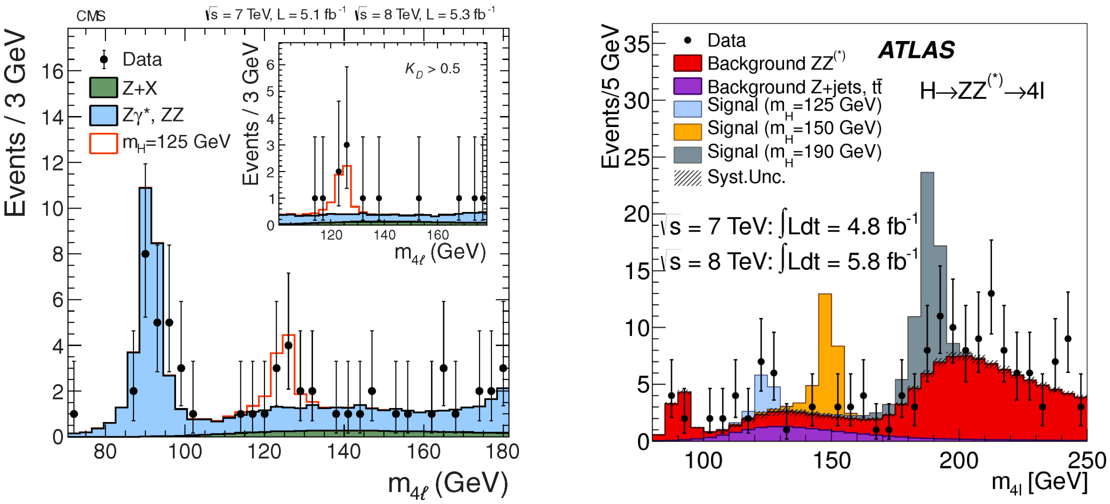
On July 4th, 2012 having analyzed the entire 2011 dataset ($\sim$5 fb^{-1}) and a significant portion of the 2012 data (5-6 fb^{-1}) the spokespersons of CMS and ATLAS announced the independent discoveries of a new particle near 125 GeV in the search for a SM Higgs boson [57]. Both experiments found 4-4.5σ excesses in the $\mathrm{H} \rightarrow \gamma\gamma$ channel, depicted in Figure 4.21, and $\sim 3\sigma$ excesses in the $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ channel, depicted in Figure 4.22. When CMS combined all five of the primary search channels, a 5σ observation of a new particle compatible with a SM Higgs boson was achieved. ATLAS also independently observed this particle at a similar mass with 5σ significance using the combination of their $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ analyses. ATLAS performed a similar, five-channel combination in their published analysis [58]. The significance of their observation increased to nearly 6σ because of excess observed in the $\mathrm{H} \rightarrow \mathrm{WW}$ channel. Figure 4.23 depicts the significance of each experiment's published five-channel combination as a function of $\mathrm{M_H}$ with the maximum significance of $\geq 5\sigma$ achieved near 125-126 GeV.



(a) CMS (b) ATLAS

Figure 4.21: 4th of July, 2012 mass spectra in the $\mathrm{H} \rightarrow \gamma\gamma$ analyses. A clear peak around 125 GeV is visible for both experiments, contributing evidence of the production of Higgs bosons at this mass.

CMS also presented the measurement of the mass during the announcement of the discovery. The $\mathrm{H} \rightarrow \gamma\gamma$ channel dominated the precision of the measurement. At that time there were two main systematic uncertainties in the mass

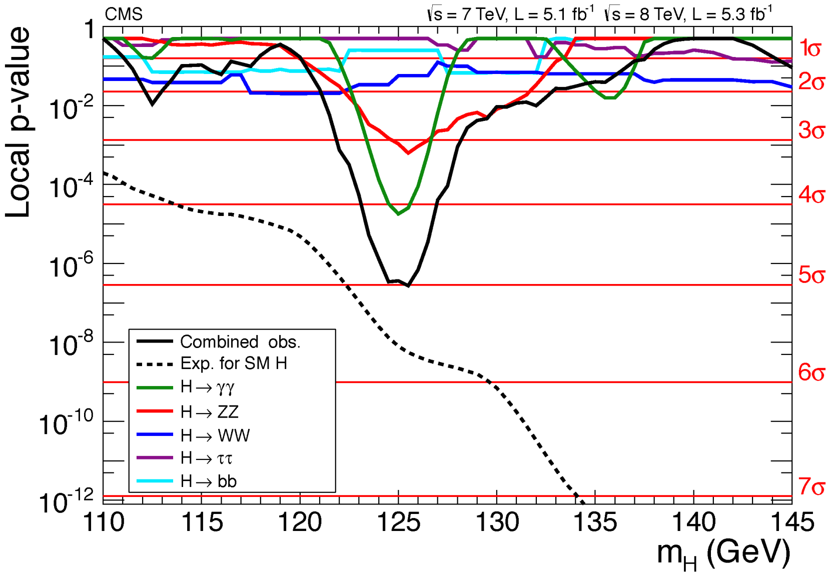


(a) CMS

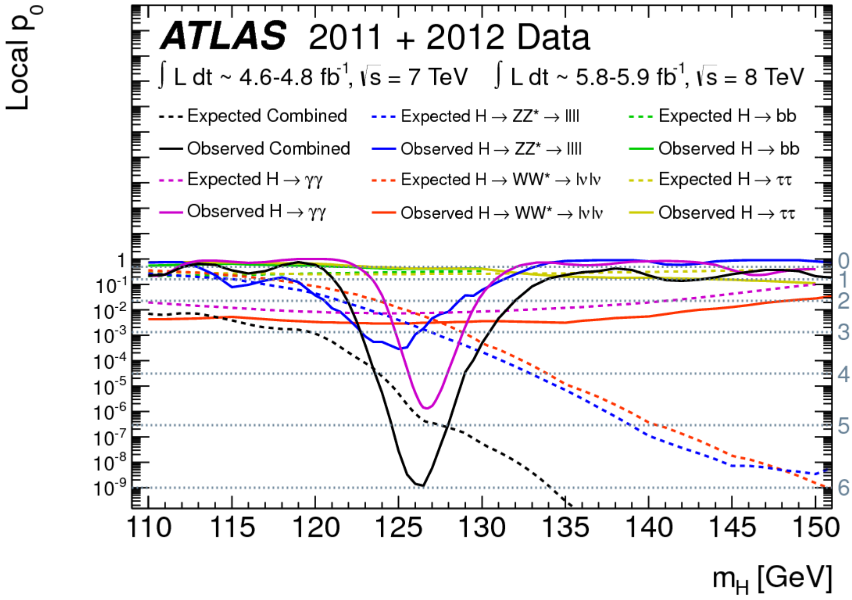
(b) ATLAS

Figure 4.22: 4th of July, 2012 mass spectra in the H → ZZ → 4ℓ analyses. A clear peak around 125 GeV is visible for both experiments, contributing evidence of the production of Higgs bosons at this mass.

measurement in the H → $\gamma\gamma$ analysis. Both arise from the fact that Z → e^+e^- events are used to calibrate the energy scale. The first source of systematic uncertainty is due to the differences in modeling of photon and electron showering in the MC, which was accentuated by somewhat larger amounts of material in the tracker than in standard MC simulations. The second major source of systematic uncertainty is due to possible non-linearity in the extrapolation from the Z peak (91.2 GeV) to the energy scale of the measurement (125 GeV). Together these uncertainties account for 0.5% scale uncertainty. The H → ZZ → 4ℓ analysis assigned an uncertainty of 0.4% due to calibration of energy and momentum for leptonic decays of Z bosons to electrons and muons, respectively. The result of the measurement was $M_H = 125.3 \pm 0.4(\text{stat.}) \pm 0.5(\text{syst.})$ GeV. In Figure 4.24, the 68% confidence level contours of the H → $\gamma\gamma$ untagged, H → $\gamma\gamma$ di-jet, H → ZZ → 4ℓ and combination are shown on a two dimensional scan of signal strength versus mass.



(a) CMS



(b) ATLAS

Figure 4.23: Published five-channel combined discovery significance in CMS and ATLAS. Plotted here is the probability of the background to produce an excess as significant or more significant than that observed in the data (see Section 7.4 for further details). A probability of 1 in 3.5 million corresponds to a one-sided tail of a gaussian distribution at 5σ. This is the convention for the probability needed to claim the discovery of a new particle over already known processes. This degree of certainty was achieved by CMS and ATLAS independently when the Higgs boson near 125 GeV was discovered.

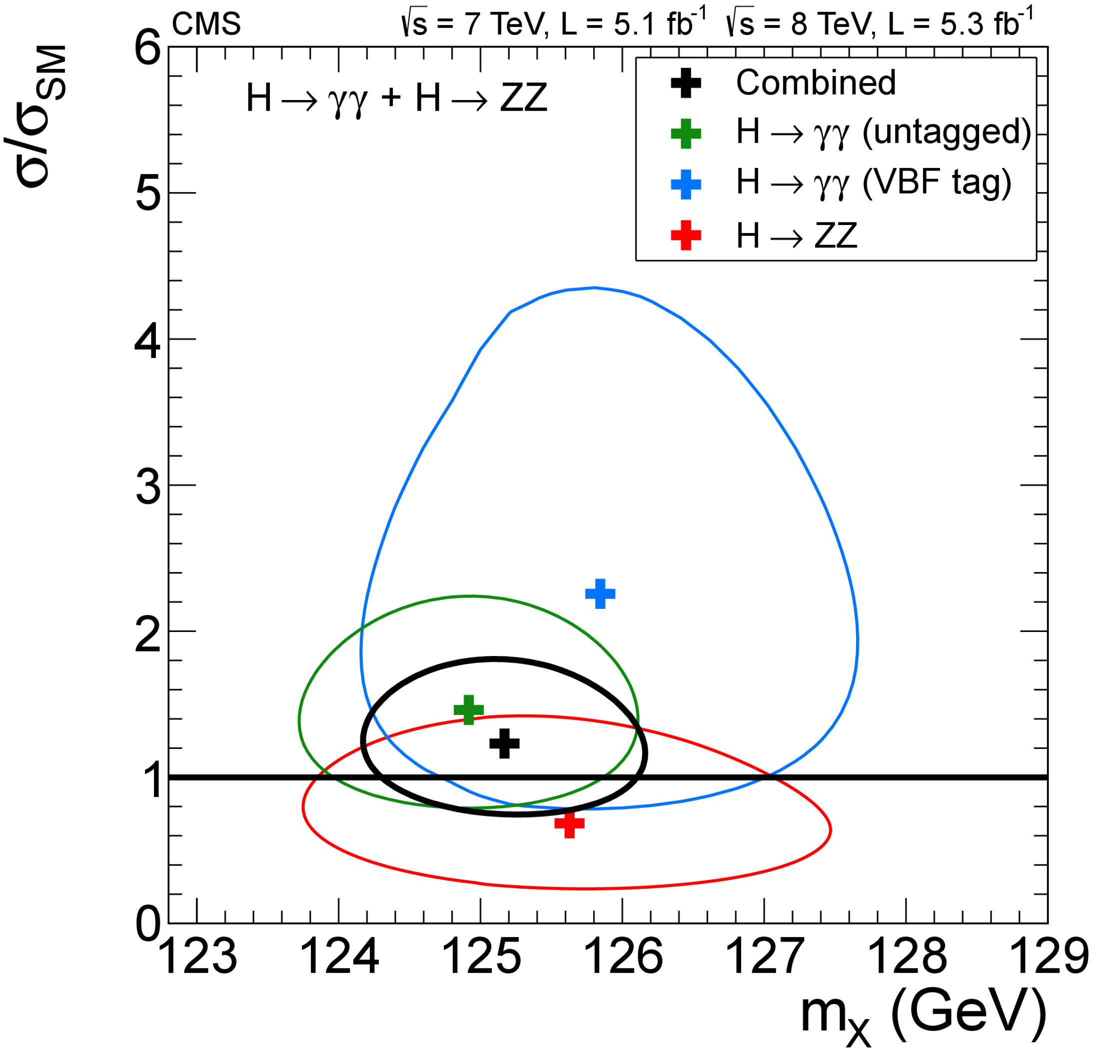


Figure 4.24: CMS's 4th of July 2012 mass measurement. The 68% CL contours for the signal strength versus the boson mass for the untagged $\gamma\gamma$, $\gamma\gamma$ with VBF-like di-jet, 4ℓ, and their combination. The symbol σ/σ_{SM} denotes the production cross section times the relevant branching fractions, relative to the SM expectation. In this combination, the relative signal strengths for the three decay modes are constrained by the expectations for the SM Higgs boson.

4.5 Comments on the Final Analysis with Run 1 Data

The upcoming H → $\gamma\gamma$ final analysis publication on CMS's Run 1 dataset, which is 5.1 fb^{-1} of 7 TeV data and 19.7 fb^{-1} of 8 TeV data, includes many updates and changes (both large and small) to the Cut-based Analysis and the MultiVariate Analysis. The calibration for this analysis is likely the most fundamental component of the analysis which has been re-derived using all calibration data from the 2011 and 2012 data. Without proper calibration this analysis cannot succeed.

I have contributed in numerous ways to these analyses. Besides numerous fundamental validation studies and analysis of the compatibility of results from different analyses and previous results, I worked on the design of the final di-jet analysis. The question of how to optimally combine di-photon and di-jet information has always been a pressing one since the introduction of the first di-photon BDT. The cut-based di-jet analyses all began with CiC selection on photons for their design. For these tags to be used in the MultiVariate Analysis some arbitrary adjustments and choices were made. These choices were not wrong but could have made the analysis slightly sub-optimal.

More optimal use of di-jet variables came with the development of the kinematic di-jet BDT, which simply uses the cut-based di-jet variables (plus di-photon p_T) as inputs. However, there was still the question of how to optimally make use of di-photon kinematics. The di-photon selection was still slightly ad hoc with reasonable choice of using the minimum di-photon BDT threshold for the untagged event classes.

However, during the summer of 2013 I performed a simple study showing that the di-photon+di-jet information could be used in a more optimal manner. Three main options were proposed: 1) a two-dimensional analysis in the di-photon and di-jet invariant mass plane in di-photon resolution bins[10]; 2) a two-dimensional classification of events using di-photon and di-jet BDTs; and 3) classification using a combined BDT with inputs of only the di-photon BDT, di-jet BDT, and di-

---

[10] This is now a cross-check analysis and its results are consistent with the other analyses.

photon p_T scaled by invariant mass. The third option was proposed by me and my advisor, Prof. James Branson. In the end it was found that our combined BDT performed the best (only slightly better than option 2 above).

Moreover, given that the resolution of the best di-photon events in the final analysis is very good and the best inclusive event classes have similar S/B to the looser di-jet tag classes (compare Untagged 0, Di-jet Tag 1 and Di-jet Tag 2 in Figure 6.16a), it became a question of where the di-jet events were optimally used. That is, the result of optimally using looser di-jet events (which typically have very good di-photon resolution) could have been to produce only one tight di-jet event class and allowing the rest of the events to fall into Untagged 0 or Untagged 1. For this reason we optimized the boundaries of the untagged and di-jet event classes simultaneously. Three di-jet event classes are still meaningful and because we optimized the event classes simultaneously, we are confident in that.

Chapter 4 is a reprint of the materials as it appears in "Search for the standard model Higgs boson decaying into two photons in *pp* collisions at $\sqrt{s} = 7$ TeV", Physics Letters B 710 (Feb, 2012) 403-425; "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Physics Letters B 716 (2012) 30-61; and "Search for a fermiophobic Higgs boson in *pp* collisions at $\sqrt{s} = 7$ TeV", JHEP 1209 (2012) 111. The dissertation author was the primary investigator and author of these publications.

Chapter 5

Final CMS $\mathrm{H} \to \gamma\gamma$ Analysis with Run 1 Data Ingredients

The $\mathrm{H} \to \gamma\gamma$ search is at its core a search for a small peak over a smoothly declining background in invariant mass. The fundamental strategy employed in the two analyses, which will be described in this and the following two chapters, is to fit the data with functions that are able to parametrize the background well and use those fits to estimate the background underneath a resonant peak whose shape is characterized by models derived from MC.

Excellent invariant mass resolution is therefore fundamental. Given that the invariant mass of a composite object constructed from two photons is:

$$\begin{aligned} m_{\gamma\gamma}^2 =& (E_1 + E_2)^2 - |\vec{p}_1 + \vec{p}_2|^2 \\ =& 2E_{\gamma_1}E_{\gamma_2}\left(1 - \cos\left(\theta_{12}\right)\right) \end{aligned} \tag{5.0.0.1}$$

accurate energy measurement and vertex determination are foundational elements of great di-photon mass resolution. Moreover, the energy of photons and the vertex of any di-photon pair are determined in the same way for both forthcoming analyses (the MultiVariate Analysis and the Cut-based Analysis).

The goal of this chapter is to describe common inputs, common procedures and numerous definitions needed for CMS's $\mathrm{H} \to \gamma\gamma$ analyses. For most readers this chapter can be considered as reference.

5.1 Datasets

Run 1 of the LHC has yielded tremendously important data. The 2011 run at 7 TeV provided 5.1 fb^{-1} of data and gave LHC the experience and confidence to increase the center of mass collision energy to 8 TeV in 2012. A total of 19.7 fb^{-1} of data was collected and deemed appropriate for analysis by CMS in 2012. With the 2011 and 2012 data a Higgs Boson with many Standard Model properties (positive parity [59], spin-0 [60], evidence of fermion decay [61]) was discovered [29].

In 2012 along with greater physics potential came more challenges. In 2012 there was much more pile-up, averaging 20 interactions per bunch crossing, than in 2011, which averaged twelve interactions per bunch crossing. Also, the APDs in the ECAL barrel experienced significantly larger dark current from radiation damage than expected. A summary of PU and integrated luminosity as a function of run period is shown in Table 5.1. The 2011 and 2012 data used for this analysis have been reconstructed with the same software and are analyzed in parallel with the same strategy.

Table 5.1: Integrated luminosity and pile-up summary for Run 1 data.

Run period	Int. lumi. (fb^{-1})	$< PU >$
2011 RunA+B	5.1	12
2012 RunA+B	5.3	19
2012 RunC	7.1	20
2012 RunD	7.3	21

5.2 Monte Carlo (MC) Simulation

Accurate simulation of background and signal using Monte Carlo techniques is paramount for designing and executing this analysis. Creating MC is generally a two-step process. First, the physics of proton-proton collisions must be simulated from collision to stable final products. There are numerous generators available [62]. Different generators and strategies have been deployed for different physics processes, which will be described later. The second step is to model the interaction and development of particles traversing the detector. Geant4 [63] is used to this end. This is particularly important for electrons and photons because as they pass through the tracker they tend to start electromagnetically shower. Accurate showering simulation is important for this analysis and where there are known inadequacies systematic uncertainties should be and are assigned.

The final H $\rightarrow \gamma\gamma$ analysis does not require MC for background estimations. However, MC is necessary for the design of the analysis. In particular, the background MC is used to study which variables may be used and how to use those variables to effectively distinguish background-like data from signal-like data. In Table 5.2 and Table 5.3 there are a list of the background MC samples produced for the 2012 and 2011 data analysis, respectively.

Table 5.2: 8 TeV Monte Carlo simulation summary including their production cross-section (including filter efficiency), equivalent simulated integrated lumiosity of the sample and weighting needed per event to scale the sample to 19700 pb^{-1}. Event weighting near or less than 1 is ideal. In the case of the QCD samples the MC event weighting is large enough to cause difficulties during analysis design.

Process	Generator	Kinematic Generator Cuts	σ (pb)	Equivalent L_{Int} (pb^{-1})	Per event weight
$\gamma\gamma$+Jets	Sherpa	$m_{\gamma\gamma} > 60$ GeV	120.4	119,880	0.164
γ+Jets	Pythia	$20 < \hat{p}_T < 40$ GeV	150.7	39,200	0.503
γ+Jets	Pythia	$\hat{p}_T > 40$ GeV	476.5	12,500	1.576
QCD	Pythia	$30 < \hat{p}_T < 40$ GeV	12,200	500	39.4
QCD	Pythia	$\hat{p}_T > 40$ GeV	51,700	189	104.
Drell-Yan→ $\ell\ell$	Madgraph	$M_{\ell\ell} > 50$ GeV	35,300	10,300	1.91

Table 5.3: 7 TeV Monte Carlo simulation summary including their production cross-section (including filter efficiency), equivalent simulated integrated lumiosity of the sample and weighting needed per event to scale the sample to 5100 pb^{-1}.

Process	Generator	Kinematic Generator Cuts	σ (pb)	Equivalent L_{Int} (pb^{-1})	Per event weight
$\gamma\gamma$+Jets	Madgraph	$m_{\gamma\gamma} > 60$ GeV	66.59	1.45×10^4	0.35
$\gamma\gamma$ (box)	Pythia	$25 < \hat{p}_T < 250$ GeV	12.37	3.44×10^4	0.15
$\gamma\gamma$ (box)	Pythia	$\hat{p}_T > 250$ GeV	0.000208	2.12×10^9	2.41×10^{-6}
γ+Jets	Pythia	$\hat{p}_T > 20$ GeV	493.44	1.12×10^4	0.46
QCD	Pythia	$30 < \hat{p}_T < 40$ GeV	9,614	632	8.1
QCD	Pythia	$\hat{p}_T > 40$ GeV	40,392	996	5.1
Drell-Yan→ $\ell\ell$	Powheg-Pythia		2475.0	1.27×10^5	0.040

The Sherpa-generated [64] di-photon sample contains all of the irreducible background (i.e. with two real photons) that is simulated for the 8 TeV analysis. The Feynman diagrams associated to this sample are the Born, which is from quark-antiquark annihilation and whose leading-order diagram is depicted in Figure 5.1a, with up to three additional jets, and the box diagram, which is depicted in Figure 5.1b. These diagrams contribute most of the real di-photon spectrum to the measured di-photon continuum. In the 7 TeV analysis the di-photon madgraph [65] sample produces the Born di-photon spectrum and there are additional box diagram samples produced with Pythia [66], which is only leading-order.

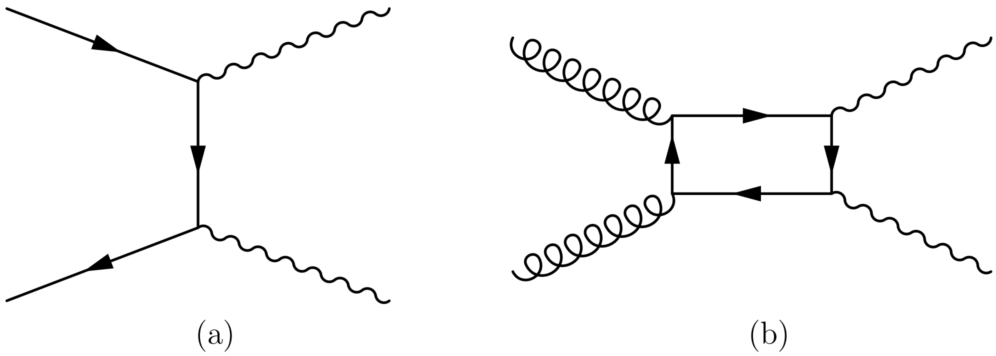


Figure 5.1: Leading-order Feynman diagrams for real (two photon) background. The Born diagram (a) and this box diagram (b) are the primary sources of irreducible di-photon background.

All other samples are reducible backgrounds, meaning that one or both of the objects commonly reconstructed as a photon is not actually a photon (or a "fake" photon). These fakes are typically jets but can be electrons where the track is not reconstructed as well. Therefore, samples with one or two jets were produced with Pythia [66], and a large drell-yan sample containing the di-electron final state with up to three jets was produced utilizing the Madgraph generator [65].

In the the γ+Jets and QCD samples there is a generator level filter, which allows events with two generated objects which can produce fake photons to proceed to the second step of simulation. It is inevitable that some events with fakes are lost due to inefficiency in the filter. In addition, Pythia is a leading-order generator and calculations have been performed to determine the increase in cross section to next-to-leading-order (NLO) Feynman diagrams. These two combined effects lead to an increase in cross section, called the "k-factor", of 1.3.

The one real, one fake components of the γ+Jets and QCD samples need improvement[1]. These samples are particularly deficient after two photons and two jets are selected. After these requirements the MC yields are typically 1.5 to 3 times lower (depending on tightness of selection) than the corresponding yield in data. A Sherpa-generated γ+Jets MC sample is being developed to replace these samples.

Table 5.4 contains the Standard Model cross section times branch ratio for Higgs Boson theoretical production in 2011 and 2012 with $\mathrm{M_H} = 125\mathrm{GeV}$[2]. The Feynman Diagrams corresponding to these production mechanisms are depicted in Figure 4.2. The process that produces the most Higgs Bosons is Gluon Fusion (GF) via fermion loop. Although this is the most prevalent production mechanism, the other mechanisms provide additional objects, such as jets, b jets via t decays, or leptons from W/Z decay, which can be tagged. In Appendix A the production cross section of a Standard Model Higgs Boson and the branching ratios of the $\mathrm{H} \to \gamma\gamma$ process are written in detail for $\mathrm{M_H}$ from 90 to 150 GeV in 5 GeV steps for 7 TeV and 8 TeV proton-proton collisions.

Table 5.4: $\mathrm{H_{SM}} \to \gamma\gamma$ theoretical yields with $\mathrm{M_H} = 125$ GeV. Samples are generated with Powheg or Pythia.

Production Mechanism	Generator	$\sigma\times$BR (pb) 7 TeV	2011 Higgs Production	$\sigma\times$BR (pb) 8 TeV	2012 Higgs Production
Gluon Fusion	Powheg	0.0345	176	0.0439	680
VBF	Powheg	0.00279	14.2	0.00360	71.0
Via W	Pythia	0.001319	6.73	0.001606	31.6
Via Z	Pythia	0.000764	3.90	0.000947	18.7
$\mathrm{t\bar{t}H}$	Pythia	0.0001968	1.00	0.000295	5.81
Total		0.0396	202	0.0503	992

5.2.1 Pile-up Re-weighting

The MC generation is done with some estimate of the expected pileup. However, the actual distribution of pile-up in the data is slightly different than

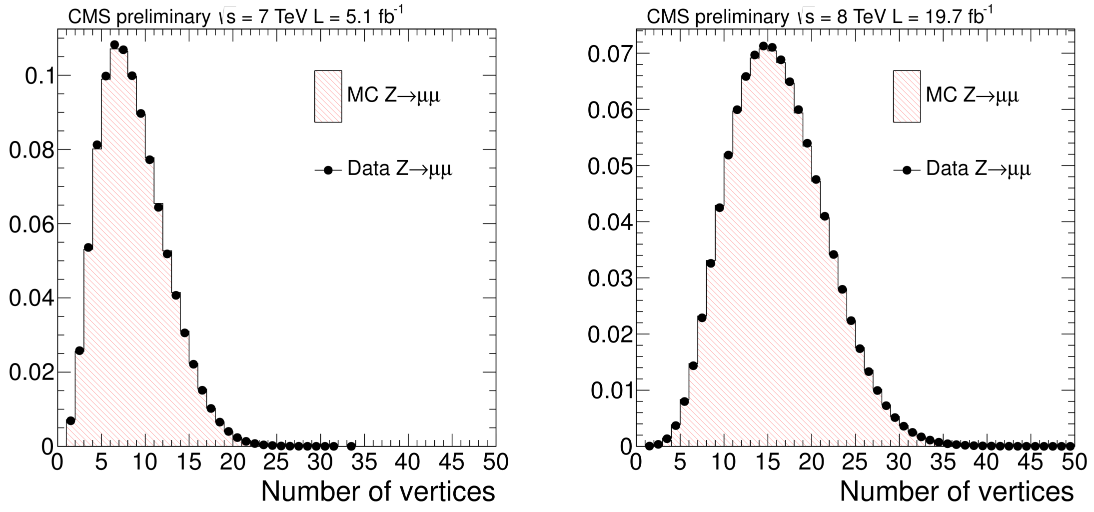
---

[1] The event weighting of the 8 TeV QCD samples was problematically large. However, this is a separate issue.

[2] Given the integrated luminosity and the collision energy.

that in the MC. The most pile-up interactions are not very interesting but frequent proton-proton interactions called minimum bias. Thus, actual pile-up is determined with instantaneous luminosity measurements as a function of time and measurements of minimum bias cross sections (68 mb for $\sqrt{s} = 7$ TeV and 69.4 mb for $\sqrt{s} = 8$ TeV). MC events are re-weighted such that the final distribution of the number of interactions per event match the measurement from data.

The pile-up re-weighting is validated using $Z \to \mu^+\mu^-$ samples, and observing data to MC agreement in number of reconstructed vertices. These distributions are shown for 2011 and 2012 data in Figure 5.2.

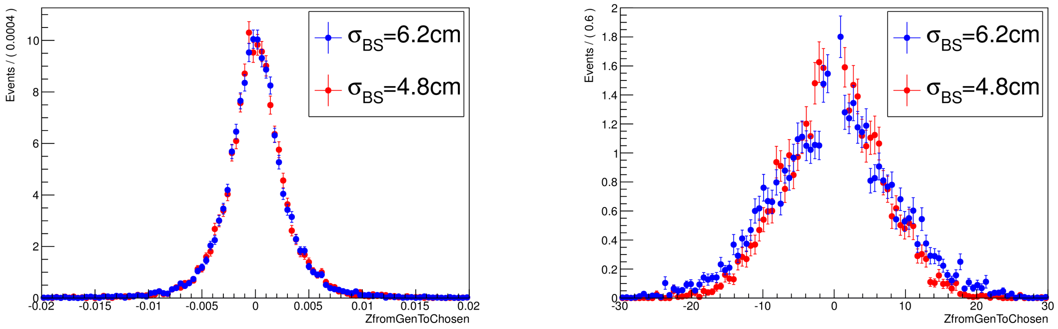


(a) 7 TeV - 2011 (b) 8 TeV - 2012

Figure 5.2: Number of reconstructed vertices in $Z \to \mu^+\mu^-$ data and MC with MC reweighting to match estimates of pile-up in each dataset for 7 TeV (a) and 8 TeV (b) datasets.

5.2.2 Beamspot Re-weighting

The beamspot width (σ_{BS}) is the gaussian width of the position of the bunch crossing within CMS. The width of this area depends on numerous LHC beam parameters. Interaction vertices are generally found in or near this area. In this analysis σ_{BS} directly impacts resolution in signal models when the wrong vertex is selected because for wider σ_{BS} the wrong vertex is farther from the correct vertex.



(a) $\Delta Z_{\text{ChosenToGen}} < 0.1$ cm (b) $\Delta Z_{\text{ChosenToGen}} > 0.1$ cm

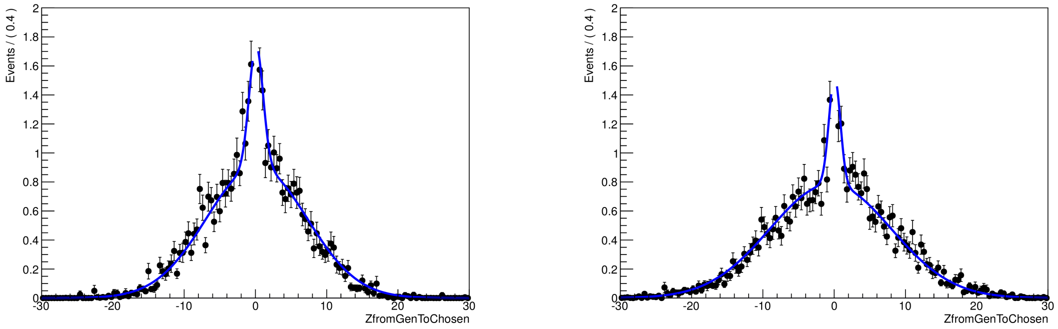
Figure 5.3: Uncorrected beamspot position in MC samples with different generated σ_{BS}. Distribution of selected vertex's position in z from the beamspot's position in z, $\Delta Z_{\text{ChosenToGen}}$ for MC signal samples with $\sigma_{\text{BS}} = 6.2$ cm and $\sigma_{\text{BS}} = 4.8$ cm: (a) for $\Delta Z_{\text{ChosenToGen}} < 0.1$ cm and (b) for $\Delta Z_{\text{ChosenToGen}} > 0.1$ cm.

The 8 TeV MC was generated assuming a wider beamspot width ($\sigma_{\text{BS}} \sim 6.2$ cm) than the actual gaussian beamspot width in 2012 data ($\sigma_{\text{BS}} \sim 5.0$ cm). Rather than reproduce all new MC a few samples of Gluon Fusion were produced with a narrower beamspot width ($\sigma_{\text{BS}} \sim 4.8$ cm) in order to determine if there were any effects to correct. It was determined that signal resolution was 2-3% worse in the nominal MC and that it should be corrected.

By comparing the distributions of the distance in Z between the selected reconstructed vertex and the generated vertex ($\Delta Z_{\text{ChosenToGen}}$) in both MC samples as in Figure 5.3, it was determined that when the selected vertex is found within 0.1 cm of the generated vertex there is very good agreement both in shape and yield. Thus, for the case of correct vertex selection (Figure 5.4a) no corrections are necessary.

However, when the selected vertex is beyond 0.1 cm of the generated vertex the shapes of $\Delta Z_{\text{ChosenToGen}}$ are considerably different in width (Figure 5.4b). Several methods of correction were compared and re-weighting to the distribution of $\Delta Z_{\text{ChosenToGen}}$ in the narrower beamspot was found to be the simplest and most successful way to recover the resolution loss.

There are two components to the $\Delta Z_{\text{ChosenToGen}}$ distribution: one with a width of ∼1 cm, and a broader component dependent on σ_{BS}. Reweighting is



(a) $\sigma_{\mathrm{BS}} = 4.8$ cm

(b) $\sigma_{\mathrm{BS}} = 5.8$ cm

Figure 5.4: Double gaussian parametrization for beamspot re-weighting of the $\Delta Z_{\mathrm{ChosenToGen}}$ distribution. Figure 5.4a shows the double gaussian fit of the narrower, more accurate signal MC, while Figure 5.4b shows the standard signal MC. The difference in the width of the wider gaussian is visually apparent.

performed as a function of $\Delta Z_{\mathrm{ChosenToGen}}$ from the ratio of double Gaussian fits to each sample, shown in Fig. 5.4. The re-weighting restores the mass resolution to nearly the level of the more realistic beamspot sample.

This re-weighting is validated in $\mathrm{Z} \to \mu^+\mu^-$ data and MC. The comparison of $\Delta Z_{\mathrm{ChosenToGen}}$ in data to MC before re-weighting (red) and MC after re-weighting (green) is shown in Figure 5.5. There is very good agreement in $\Delta Z_{\mathrm{ChosenToGen}}$ after re-weighting.

No beamspot correction is needed or applied for the 7 TeV MC.

5.3 Triggers

Two trigger strategies have been designed to efficiently collect $\mathrm{H} \to \gamma\gamma$ events. Both require the presence of a L1 seed from the ECAL. The L1 seed is either a single deposit of energy in the ECAL or two such deposits with asymmetric thresholds. The HLTs all have asymmetric thresholds on the transverse energy (E_T) of two super-clusters. One set of HLTs apply loose cuts on detector-based isolation and shower-shape variables, while the complementary set of HLTs check for the presences of two super-clusters with very narrow shower-shape. In particular, events seeded by the L1 where super-clusters with R_9–the energy in

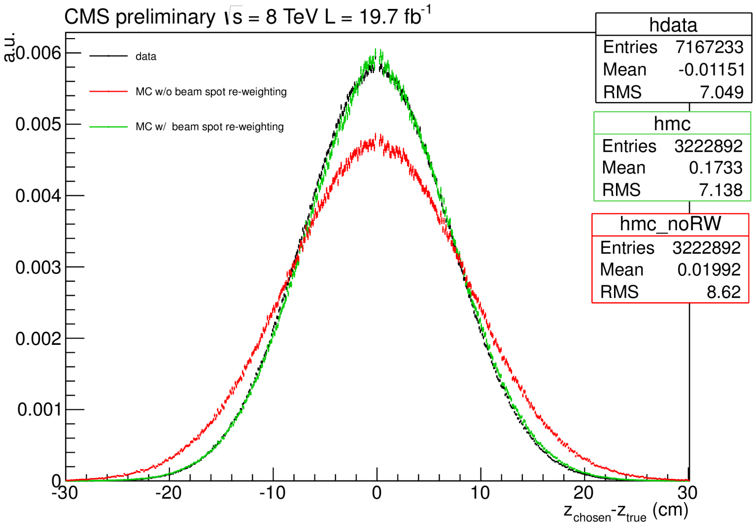


Figure 5.5: Beamspot re-weighting validation with Z $\rightarrow \mu^+\mu^-$ events. Distribution of $\Delta Z_{\mathrm{ChosenToGen}}$ between the chosen vertex and the true vertex for data (black), MC without re-weighting (red), and MC after beamspot re-weighting (green) in Z $\rightarrow \mu^+\mu^-$ events.

3×3 crystals centered at the maximum energy crystal divided by the energy in the super-cluster–greater than 0.85 in 2012 are selected regardless of any other isolation or shower-shape variables. The R_9 threshold was 0.8 or 0.9 in 2011 depending on the run period. In addition there are hybrid HLTs in which one super-cluster is required to be loosely isolated and the other is required to have high R_9. In addition since 2011 Run B a loose invariant mass cut on the composite particle formed from the triggering super-clusters is required (either 60 or 70 GeV depending on run period). A detailed list of HLTs and corresponding L1 seeds are listed in Appendix B.

In order to correct for differences between MC and data, it is very important to measure the efficiency of the triggers with data and correct the efficiency of the signal MC to match that of the data. Here $Z \rightarrow ee$ data events are used with electrons treated like photons. Since there are independent, uncorrelated triggers for double electron events, the efficiency of the H $\rightarrow \gamma\gamma$ can be measured with these events after similar selection. In Table 5.5 the results of these measurements in the 2012 data are summarized.

The trigger efficiency for the 2011 data is known to be nearly 100% from previous measurements. A conservative systematic of 1% is applied for both 2011 and 2012 datasets.

5.4 Photon Reconstruction

Photons are reconstructed from energy clusters in the ECAL. Collections of clusters make up super-clusters and these super-clusters are sorted into photons and electrons. Since the geometric layouts of the EB and EE are different, the clustering algorithms must be different. The clusters themselves are built from reconstructed hits or RecHits, which refer to the single crystal calibrated energy estimates.

The clustering algorithms are designed to deal with photons that shower in the tracker. The electrons from the shower bend in the ϕ direction due to the axial magnetic field. Thus, the goal of these algorithms is to make individual clusters and collect groups of clusters within a relatively narrow band in η which originate from the same photon.

The EB crystals are in a grid of $\eta \times \phi$ and the clustering algorithm is the Hybrid method. This method searches for an un-clustered RecHit with E_T greater than 1 GeV. When such a RecHit is found, it is designated the seed and clustering can begin. A 5×1 (in $\eta \times \phi$) group of crystals centered around the seed is the first domino of the cluster. Crystals with the same η and within 17 crystals on either side in ϕ are checked for 5×1 dominos with E_T greater than 0.1 GeV. Passing dominos are grouped with adjacent dominos and centered around local maxima. If the local maximum crystal's E_T is less than 0.35, then the group of dominos is removed. Next to the remaining groups of 5×1 dominos, 3×1 dominos centered

Table 5.5: 2012 trigger efficiency. L1 and combined HLT efficiencies for events passing 2012 selections.

	L1_DoubleEG_13_7	L1_SingleEG_22	OR of All HLTs
MVA > -0.78	99.75±0.01%	97.14±0.02%	99.43 ± 0.03
CiC Supertight	99.83±0.01%	97.81±0.02%	99.78 ± 0.03

on the same η are attached to the clustered 5×1 dominos as long as $E_T > 0$. Each 5×1 set of dominos with 3×1 domino wings is a cluster and the entire group is a super-cluster. An example of the result of the Hybrid algorithm is depicted in Figure 5.6.

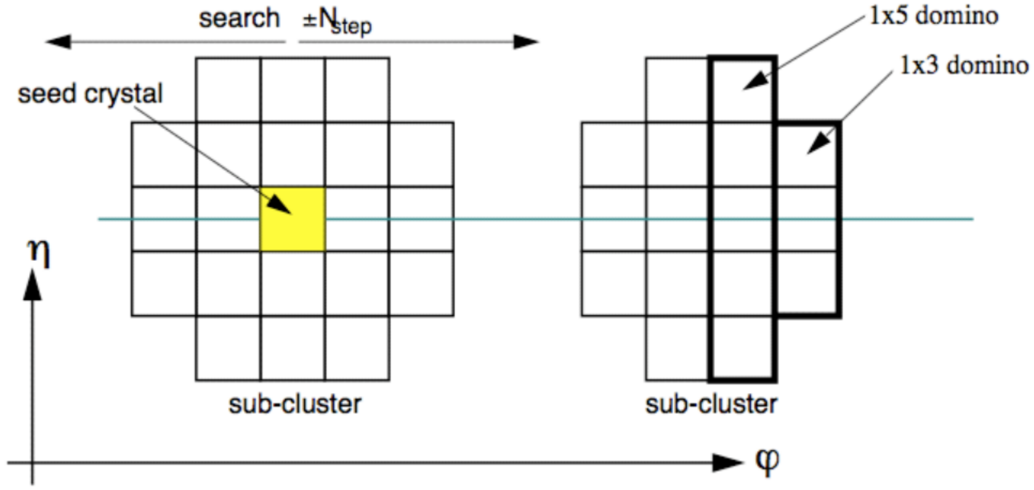


Figure 5.6: Hybrid ECAL clustering and photon reconstruction diagram.

The EE crystals are arranged in an x×y grid. Since this grid is Cartesian, the symmetries of η and ϕ cannot be utilized directly as a function of the crystals and a different clustering algorithm must be employed. The Multi5×5 algorithm works in the following way. Unclustered RecHits are checked to have E_T of 0.18 GeV or greater. If the RecHit in question passes this threshold, then it is checked to be the maximum energy deposit amongst the four crystals immediately adjacent to its faces. If it is the maximum energy crystal, then a basic cluster of 5×5 crystals is formed with it at the center. The outer 16 crystals in the 5×5 matrix are allowed to seed for other basic clusters. The idea is that in the EE there are many low energy clusters close together or overlapping because of the high levels of bremsstrahlung. No crystal may belong to two clusters. An example of two overlapping clusters is depicted in Figure 5.7.

After the EE clustering has finished, the clusters are ordered in E_T. The clusters which have E_T larger than 1.0 GeV can be the seed of a super-cluster. Starting with the highest E_T clusters rectangular windows in $\eta\times\phi$ are made around

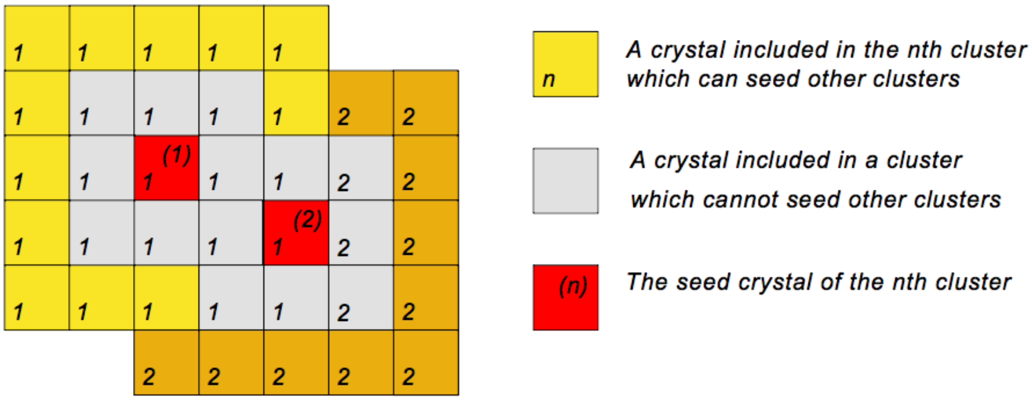


Figure 5.7: ECAL endcap clustering schematic. Here is an exmple of two overlapping clusters formed by the Multi×5 algorithm. Notice that each crystal is only assigned to one cluster.

the cluster. The window is wide in ϕ and collects clusters to make super-clusters.

5.4.1 Converted Photons

Given that there is 0.4-1.8 X_0 of material inside the ECAL (depending on η, see Figure 3.8) it is not surprising that a large number of di-photon events have at least one photon that has converted to an electron-positron pair before reaching the ECAL. Accurate reconstruction of converted photons is helpful to photon resolution via vertex pointing. The conversion tracks point back to the vertex and that information is used to improve efficiency in vertex selection and therefore mass resolution. In nearly one fourth of reconstructed events at least one photon is associated to a conversion pair.

Tracks are preselected from electron and general track collections with the criteria that there are more than four tracker hits per track and that the track's quality is minimally sufficient (i.e. $\chi^2 < 10$). Only opposite-signed track pairs, which follow the preceding selection, are considered. Due to the massless nature of the photon the decay products are very collinear. This signature distinguishes photon conversions from mis-constructed tracks and weakly-decaying massive mesons particles. To exploit this signature pairs are selected when $|\cot\theta_{Track1} - \cot\theta_{Track2}| < 0.1$. In addition, the two tracks must be compatible with each other by originating within 5 cm in the z-direction and by requiring that the tracks be parallel inside the tracker volume. Finally, the tracks are re-fit with the constraint that they originate from the vertex the tracks point to. The fit must result in compatibility better than $\chi^2 < 10^{-6}$ and the final fitted tracks must have a composite $P_T > 10\text{GeV}$.

Super-clusters are matched to conversions by passing a tight requirement of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.1$. The η and ϕ for both the conversion and the super-cluster for the ΔR computation are with respect to the re-fitted conversion vertex. If more than one conversion is within 0.1 of the super-cluster, then the conversion with minimum ΔR is selected.

5.5 Photon Energy

The intrinsic width of a Higgs boson is very narrow, $O\,(\text{1-10 MeV})$ [67, 68]. However, the measured resolution of such a resonance is limited by the nature of proton-proton collisions. CMS has designed its ECAL to reconstruct photons, such that di-photon resolution is ideally $O\,(\text{GeV})$ in the central barrel.

Detection of a Higgs boson using the two photon channel depends on the ability of our detector to achieve this goal of accurately and precisely measuring the energy of the photons. Energy resolution is therefore very important to this analysis. The mass of a Higgs boson is an explicit function of the energy of the photons as seen in Equation 5.0.0.1. Indeed, the error per event in the mass from measurement uncertainty is:

$$\left(\frac{\Delta m_{\gamma\gamma}}{m_{\gamma\gamma}}\right)^2 = \frac{1}{4}\left(\left(\frac{\Delta E_1}{E_1}\right)^2 + \left(\frac{\Delta E_2}{E_2}\right)^2 + \left(\frac{\Delta\theta_{12}}{\tan\left(\theta_{12}/2\right)}\right)^2\right) \tag{5.5.0.1}$$

There are a number of factors that impact photon resolution but arguably the most significant is the amount of material that photons travel through before reaching the ECAL. Therefore the position in the detector (particularly the value of $|\eta|$) heavily and generally impacts photon resolution. In addition, shower-shape variables are indicative of how much a photon has showered. These indicators are used to correct (on average) single photon energy using a semi-parametric energy regression.

5.5.1 Photon Energy Reconstruction

A concerted effort has been made to deliver photon energies to analysts by studying photons and electrons as objects and not as part of any physics analysis. The energy of a "raw" (i.e. object reconstruction performed globally in CMS) super-cluster is given by Equation 5.5.1.1.

$$E_{e,\gamma} = F_{e,\gamma}\left(\eta,\phi\right) \cdot \left[G \cdot \sum_{i}\left[S_i\left(t\right) \cdot C_i \cdot A_i\right] + E_{ES}\right] \tag{5.5.1.1}$$

$F_{e,\gamma}\left(\eta,\phi\right)$ corrects for geometry dependence, which is different for photons and electrons. G is global energy scale and it is derived from measurements of $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ events. The scales in the EB and the EE are set separately. $S_i\left(t\right)$ corrects for time dependent transparency-loss. The response of individual crystals is measured with the blue laser system installed in the detector. C_i is the intercalibration constant derived using three methods for calibration: 1) enforcing ϕ-symmetry in large minimum bias event samples; 2) correcting the central value in $\eta \rightarrow \gamma\gamma$ and $\pi_0 \rightarrow \gamma\gamma$ events and 3) using E/p measurements from electrons. Figure 5.9 shows the level of resolution from these methods of intercalibration for 2011 and 2012 data. A_i is the pulse amplitude converted by front-end detectors from analogue signals from the APDs and VPTs. The A_i is the recorded signal that must be converted into an energy. E_{ES} is the preshower energy, which is only non-zero when $1.65 < |\eta| < 2.6$.

5.5.2 Energy Regression

In order to achieve more optimal photon resolution a multivariate regression corrects raw super-cluster energy for local shower containment, energy lost due to early showering and extra energy from pile-up. The training and application machinery for this technique (GBRLikelihood) is built on top of RooFit, with parts of the algorithm related to the Boosted Decision Trees based originally on the TMVA implementation [71]. For the training the prompt (real) photons and

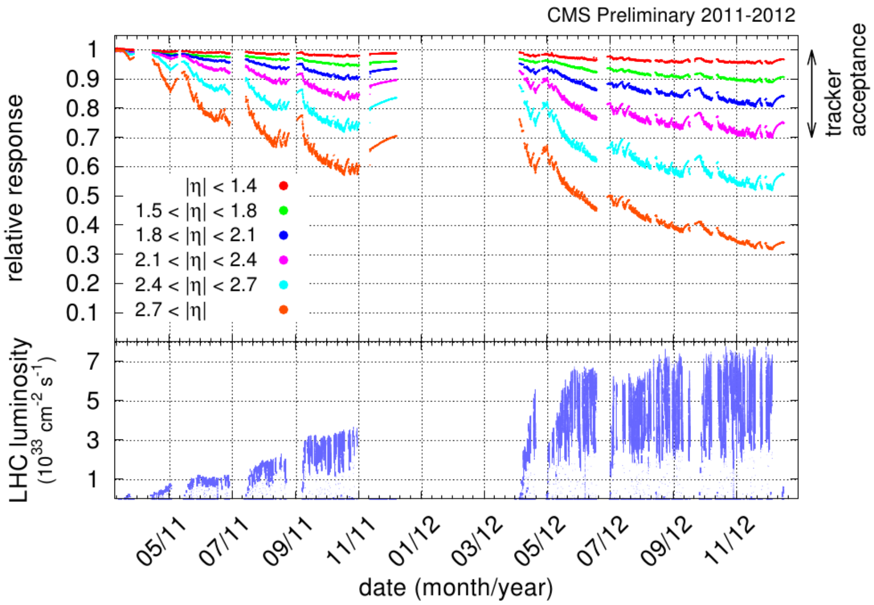
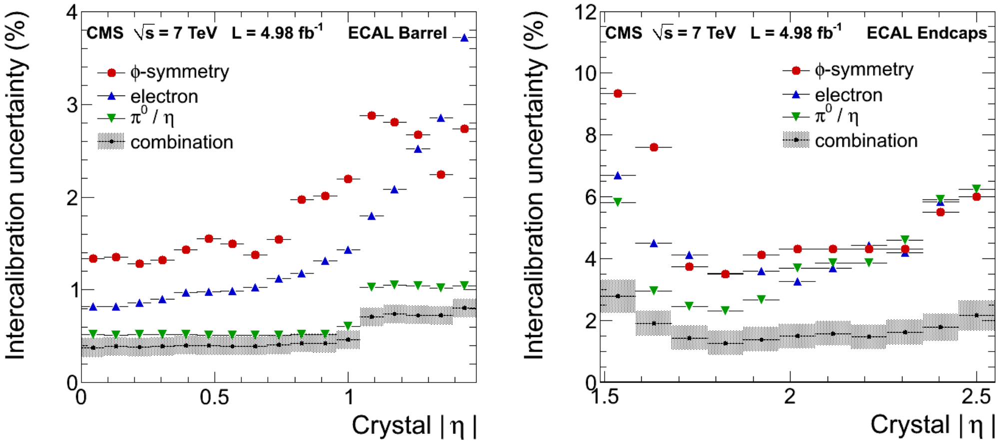


Figure 5.8: ECAL laser corrections. The measured response to calibration lasers is depicted above for crystals group by position in $|\eta|$ as a function of time for 2011 and 2012. Below is the instantaneous luminosity with identical time scale. The response drops most for higher values of $|\eta|$ and at times with highest instantaneous luminosity [69].

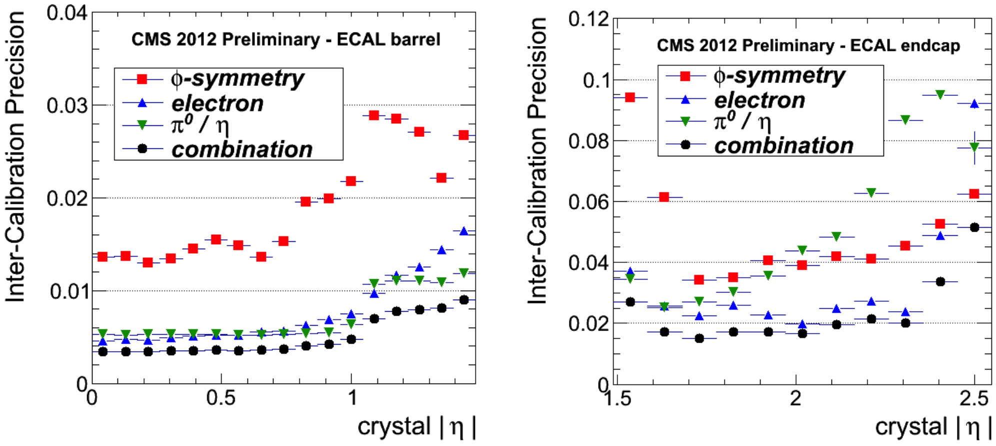
the fake photons are taken from the γ+jets MC samples. A different regression is produced for 2011 and 2012 data/MC.

Correcting the energy response, which is the super-cluster energy (plus the ES energy in the EE) divided by the generator photon's true energy, is the objective of the regression. Inputs to the regression are the super-cluster's coordinates (η, ϕ), various shower-shape variables, information about the local coordinates of some basic clusters and event information related to pile-up. The following shower shape variables are included: R_9, the ratio of 5×5 crystal energy to the super-cluster energy, the width in η and ϕ of the super-cluster, the number of basic clusters and the ratio of hadronic energy from the HCAL to the super-cluster energy. For super-clusters in the EE the ratio of the energy in the ES to the super-cluster energy is also included.

Various pieces of information regarding the basic cluster seed are also given as inputs. These include the relative energy and position of the seed, the local



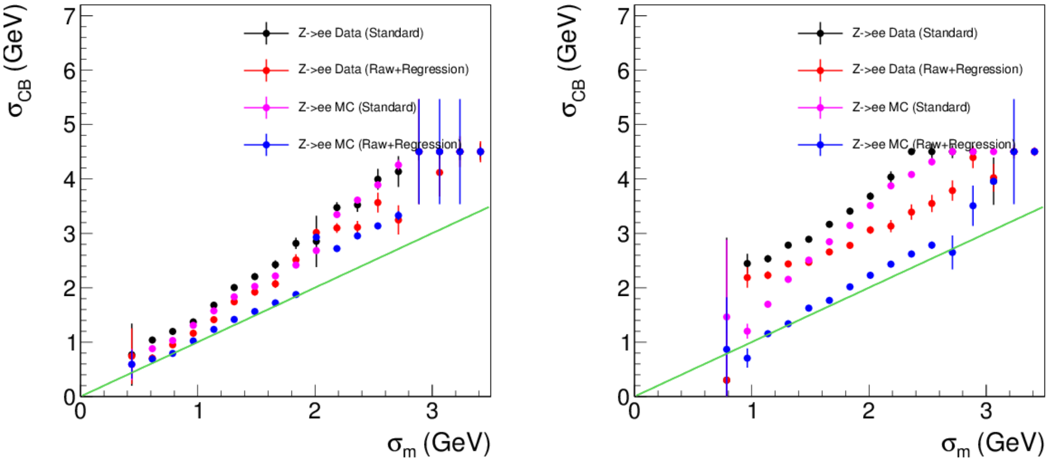
(a) 2011 Inter-calibration precision versus $|\eta|$.



(b) 2012 Inter-calibration precision versus $|\eta|$.

Figure 5.9: ECAL inter-calibration precision versus $|\eta|$. The 2011 [70] and 2012 [69] inter-calibration results yield very similar precision. When $|\eta| < 1.0$ the precision is ∼0.5%.

covariance matrix, and a number of crystal energy ratios. These variables are used by the regression to correct for global containment (i.e. missing clusters and energy absorbed in the detector prior to reaching the ECAL). In the EB, the seeding basic cluster's relative location and absolute location in terms of crystal indices in the (η, ϕ) grid are important for local containment corrections (i.e. light lost in crystals, energy lost between modules or crystals). For pile-up corrections, the number of reconstructed vertices and the average energy density, ρ, of the event are included. In the EE, the addition of the ϕ coordinates improve resolution in the MC but not in the data and therefore, all ϕ coordinates are removed for the regression in the EE.



(a) Both electrons in EB (b) At least one electron in EE

Figure 5.10: Expected versus measured resolution with regressed energy. This is a comparison of measured Z mass resolution, σ_{CB}, with expected resolution, σ_m, for data and MC. The regression energy provides better mass resolution for data and MC. The mass resolution of data and MC are not the same and therefore the MC must be smeared to match the resolution in the data.

The regression itself employs a novel semi-parametric method for training. The target is the ratio of the generator photon's true energy to super-cluster energy (plus the ES energy in the EE), to easily extract the correction to super-cluster energy. The shape of the target distribution is parametrized by a double Crystal Ball function [72], which is a gaussian core with exponential tails on either side. The tail with super-cluster energy greater than true energy are over-corrected by calibrated reconstruction, which is generally due to pile-up. The other tail represent the loss of energy due to containment.

The parameters of the double Crystal Ball are in principle very different for different categories of photons. Instead of modeling these parameters of as a function of energy and shower-shape, they are instead regressed. This allows the shape of the double Crystal Ball to vary without constraint for all types of photons. Using this regression the value of the correction and the width of the distribution, σ_{CB}, can be estimated per event. The correction is applied to the super-cluster energies in the analysis and σ_{CB} is used as a partial estimate of photon energy resolution.

$\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ events are used to validate this regression and its methodology. Although the photon regression can be applied to electrons, there are differences in showering behavoir between photons and electrons. Thus, to avoid applying the photon regression on electrons an electron regression is trained in parallel on $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ MC. Figure 5.10 shows the estimated mass resolution with standard electron super-cluster energy and regressed energy in data and MC. The regressed mass resolution is closer to expected than the standard mass resolution for both data and MC. It has also been found that the regressed mass scale and resolution is nearly flat as a function of number of reconstructed vertices, which is indicative of pile-up, whereas this is not the case with the standard calibrated super-cluster energy.

5.5.3 Energy Smearing and Scale Corrections from $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ Measurements

Despite all the MC-based efforts highlighted above there are still discrepancies between data and MC. One example can be seen in Figure 5.10 where the resolution in the data is worse than it is in the MC for both standard energies and regressed energies. This final procedure corrects relatively minor scale differences in $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ peak position between data and MC due to imperfect laser corrections as a function of time and determines the amount of additional random smearing needed for the MC resolution to match that of the data in $\eta \times R_9$ bins generally and in $\eta \times R_9 \times E_T$ bins for the 2012 data in the EB. An electron energy regression is applied when determining these corrections (as opposed to the H $\rightarrow \gamma\gamma$ photon energy regression).

The final corrections to data and MC are made in a two step process (a three step process for the 2012 data in the EB). The first step, which is the called the fit method, performs fits of the M_{ee} spectrum of data and MC in bins of $|\eta|$, two in the EB and two in the EE, and in bins of range run, because the corrections are time-dependent. The fits use a convolution of a Breit-Wigner and a crystal-ball (gaussian with one-sided exponential tail). The mass and width parameters of the Breit-Wigner are fixed to PDG [4] values ($M_Z = 91.188\text{GeV}$ and $\Gamma_Z = 2.495\text{GeV}$) while the crystal-ball parameters are allowed to float. The scale corrections are defined as:

$$\Delta P = \frac{\Delta M_{\text{Data}} - \Delta M_{\text{MC}}}{M_Z} \tag{5.5.3.1}$$

The second step is not categorized by run range, but adds a further categorization in R_9 because both the scale and resolution differences between data and MC depend on shower-shape. In this step, the smearing method, the MC invariant mass shapes are used as the PDF for a maximum likelihood fit. This PDF is distorted as the individual electron energies are smeared by gaussians with central values of $1 + \Delta P$ (second scale shift) and with width of $\Delta\sigma$. For each $\eta \times R_9$ category there is an uncorrelated ΔP and $\Delta\sigma$, and for n electron categories there are $\frac{1}{2}n\,(n+1)$ distinct di-electron categories. These $2n$ variables are fit to data to maximize the likelihood.

The final scale corrections for the 7 TeV data and the 8 TeV EE data are a product of the fit method scale corrections and the smaller smearing method scale corrections. They are in bins of $\eta \times R_9 \times$time. The smearings derived in the smearing method are listed in Table 5.6.

Given that the 8 TeV EB data samples are very large, binning in E_T for further corrections was possible and preferable. The smearing method is again applied on the already corrected electron energies but now with E_T dependence and in E_T bins. For this fit the smearing of the electrons' energies are now constants

Table 5.6: Additional MC smearing in η-R_9 categories. These are the additional constant smearings derived by $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events and applied to $\mathrm{H} \to \gamma\gamma$ photons.

Category	2011 $\Delta\sigma$ (%)	2012 $\Delta\sigma$ (%)
EB, $\lvert\eta\rvert < 1$, $R9 > 0.94$	0.68 ± 0.04	0.72 ± 0.03
EB, $\lvert\eta\rvert < 1$, $R9 < 0.94$	0.96 ± 0.03	0.83 ± 0.02
EB, $\lvert\eta\rvert > 1$, $R9 > 0.94$	1.01 ± 0.14	1.07 ± 0.09
EB, $\lvert\eta\rvert > 1$, $R9 < 0.94$	1.85 ± 0.04	1.87 ± 0.02
EE, $\lvert\eta\rvert < 2$, $R9 > 0.94$	1.58 ± 0.18	1.56 ± 0.08
EE, $\lvert\eta\rvert < 2$, $R9 < 0.94$	1.85 ± 0.07	1.98 ± 0.04
EE, $\lvert\eta\rvert > 2$, $R9 > 0.94$	2.01 ± 0.06	1.87 ± 0.04
EE, $\lvert\eta\rvert > 2$, $R9 < 0.94$	1.83 ± 0.09	1.94 ± 0.05

in quadrature with an E_T-dependent term which is:

$$\Delta\sigma = \frac{\Delta S}{\sqrt{E_T}} \oplus \Delta C \tag{5.5.3.2}$$

There is a large and unknown correlation between the two terms in the smearing. To determine the appropriate ratio of the two a transformation into polar coordinates is performed as follows:

$$\begin{aligned} \Delta C &= \rho \cdot \sin\phi \\ \Delta S &= \rho \cdot \sqrt{< E_T >} \cdot \cos\phi \\ \sqrt{< E_T >} &= \frac{\Delta C|_{\Delta S=0}}{\Delta S|_{\Delta C=0}} \end{aligned} \tag{5.5.3.3}$$

$\sqrt{< E_T >}$ is the square root of the average E_T for the category. It is derived by imposing the condition that the additional smearing is the same for wholly E_T-dependent and wholly constant smearings.

This parametrization allows the same likelihood maximization to be performed in each electron category, now in 20 bins of $\eta \times R_9 \times E_T$, using $\Delta P, \rho$, and ϕ. The best fit results with the ρ, ϕ parametrization are shown in Table 5.7, while Table 5.8 shows the results as the additional constant and E_T-dependent terms.

In addition to the E_T-dependent smearings, further scale corrections in E_T bins have been derived. The residual corrections are shown in Figure 5.11.

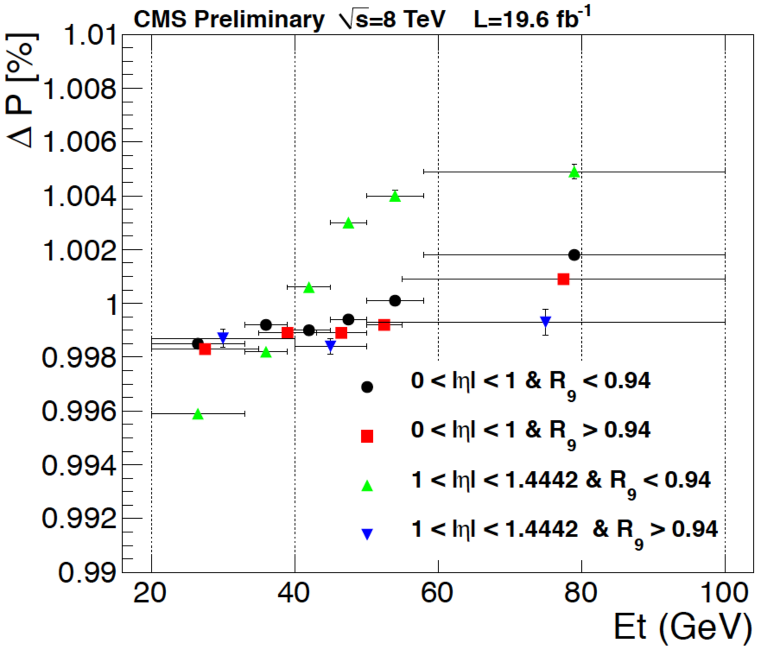


Figure 5.11: Residual energy scale corrections in E_T bins on top of the second step energy scale corrections.

Table 5.7: ρ-ϕ parametrization of residual E_T-dependent corrections for 8 TeV MC. Best fit values for ρ and ϕ parameters of additional single photon smearings. $\phi = 0$ corresponds to the residual smearing being completely E_T-dependent.

Category	$\rho[\%]$	$\phi[rad]$	$\sqrt{< E_T >}$
EB, $\|\eta\| < 1$, $R_9 > 0.94$	0.74 ± 0.07	0.00 ± 0.16	6.60
EB, $\|\eta\| < 1$, $R_9 < 0.94$	0.77 ± 0.06	0.00 ± 0.16	6.73
EB, $\|\eta\| > 1$, $R_9 > 0.94$	1.12 ± 0.13	0.00 ± 0.22	6.52
EB, $\|\eta\| > 1$, $R_9 < 0.94$	1.26 ± 0.10	0.00 ± 0.07	6.73

Table 5.8: Constant plus E_T-dependent parametrization of residual corrections for 8 TeV MC. Additional smearings for 8 TeV MC in terms of additional constant term (ΔC) and an additional E_T-dependent term ΔS.

Category	$\Delta C[\%]$			$\Delta S[\%]$		
	best fit	mean	std. dev.	best fit	mean	std. dev.
EB, $\|\eta\| < 1$, $R_9 > 0.94$	0.00	0.06	0.04	4.88	4.80	0.26
EB, $\|\eta\| < 1$, $R_9 < 0.94$	0.00	0.06	0.04	5.18	5.29	0.29
EB, $\|\eta\| > 1$, $R_9 > 0.94$	0.00	0.19	0.14	7.31	7.03	0.79
EB, $\|\eta\| > 1$, $R_9 < 0.94$	0.00	0.05	0.04	8.48	9.31	0.42

5.6 Vertex Selection

The accurate selection of the vertex is very important for the H $\rightarrow \gamma\gamma$ analysis because the position of the vertex impacts the measured opening angle between the photons. Equation 5.5.0.1 shows the dependence of opening angle in the per-event mass uncertainty. Efforts have been made to utilize all relevant and useful information as inputs to a BDT whose output is used as a discriminator amongst the many reconstructed vertices.

5.6.1 Vertex BDT Inputs

Events with two photons often come from interactions with few high p_T tracks leading to the photons, except in the case of photon conversion to electron-positron pairs. However, there is generally more hadronic activity, and therefore charged tracks, from hard interactions like those producing a Higgs Boson than in minimum bias events, which most pile-up interactions are. Thus, while the sum of vertex-associated tracks' transverse momentum can still be indicative of the origin of the photons, it is insufficient to be used alone in determining the vertex. The following variables along with some conversion information when available are used as inputs to the BDT:

- Sum $P_T^2 = \sum\limits_{\text{Tracks}} |\vec{P}_T^{\text{Track}}|^2$
- P_T balance $= -\sum\limits_{\text{Tracks}} \left(\vec{P}_T^{\text{Track}} \cdot \frac{\vec{P}_T^{\gamma\gamma}}{|\vec{P}_T^{\gamma\gamma}|} \right)$
- P_T asymmetry $= \left(\frac{|\sum\limits_{\text{Tracks}} \vec{P}_T^{\text{Track}}| - |\vec{P}_T^{\gamma\gamma}|}{|\sum\limits_{\text{Tracks}} \vec{P}_T^{\text{Track}}| + |\vec{P}_T^{\gamma\gamma}|} \right)$

In addition to these variables when conversions are found and associated to super-clusters the tracks can be used to project back to beamline. There are two methods for doing this projection: one uses only the conversion tracks and the other uses the super-cluster position to project through the origin of the conversion. Both methods are shown in Figure 5.12.

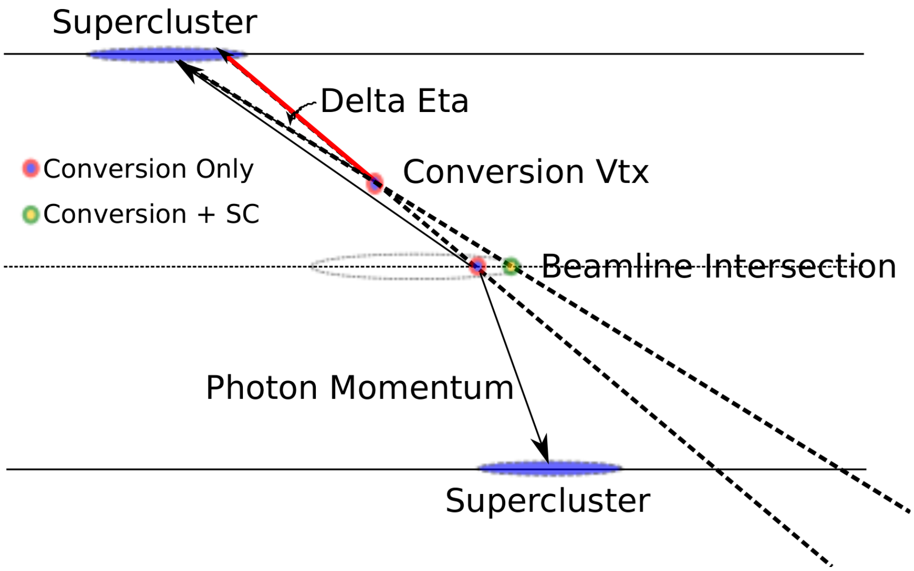


Figure 5.12: Converted photon vertex finding schematic. Depicted here are two methods for determining the vertex using conversion tracks. For tracks originating deeper in the tracker system the vertex location is projected back to the beamline using the conversion tracks alone, whereas if the tracks originate later (i.e. there are few tracker hits) the vertex is found projecting from the super-cluster to the origin of the conversion tracks.

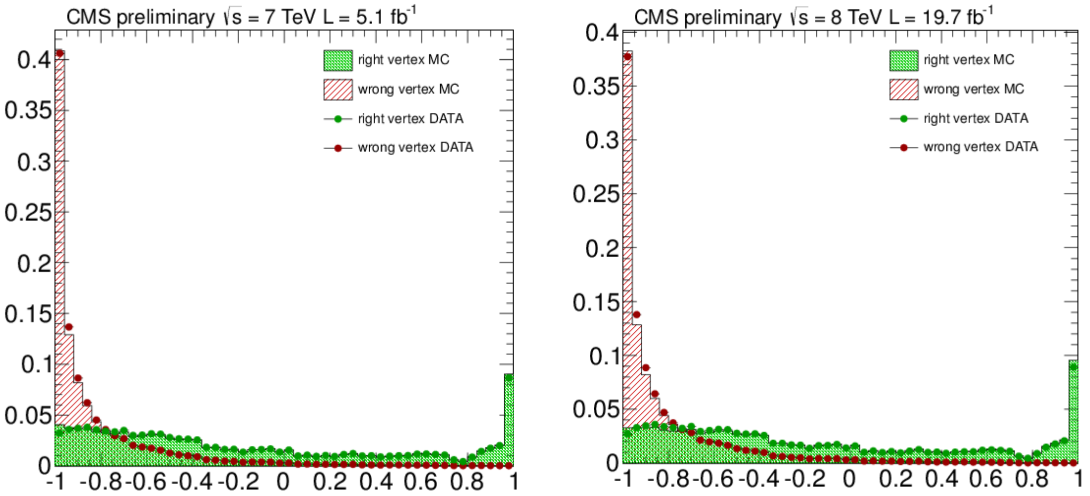
When at least one conversion is present an additional variable is utilized by the BDT described above. That is:

$$\text{pull}_{\text{conv}} = \frac{|z_{\text{vertex}} - z_{\text{conv}}|}{\sigma_{\text{conv}}} \tag{5.6.1.1}$$

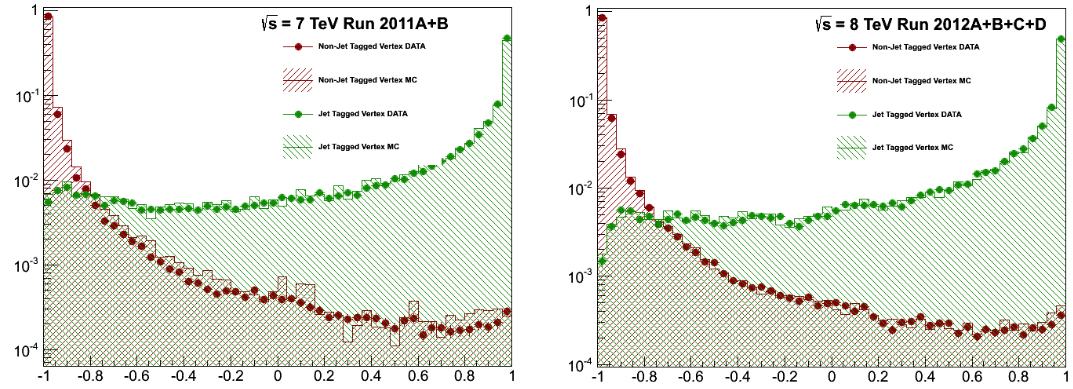
where z_{conv} is the z-position found using the methods described above and in Figure 5.12 and σ_{conv} is the resolution of the conversion position as measured in data. The BDT is trained using $\text{H} \to \gamma\gamma$ events with $\text{M}_\text{H} = 120$ GeV. The generator-matched vertex is used as signal for training and all other vertices are used as background.

5.6.2 Vertex BDT Output and Validation

Validation of BDT input variables and BDT output is performed using $\text{Z} \to \mu^+\mu^-$ events in data and MC for the conversion-free scenario. The μ tracks are removed and the vertices are reconstructed without them, which creates an analogous situation with $\text{H} \to \gamma\gamma$ events. With di-muon events the true vertex is known in data and in MC, and using that information one can observe extremely consistent behavior between the data and the MC as a function of the BDT output when the algorithm is both correct and incorrect. Figure 5.13a and Figure 5.13b displays this agreement for both 7 TeV and 8 TeV data and MC.



(a) 7 TeV - Z $\rightarrow \mu^+\mu^-$ (b) 8 TeV - Z $\rightarrow \mu^+\mu^-$



(c) 7 TeV - γ+jet (d) 8 TeV - γ+jet

Figure 5.13: Vertx BDT output in muon trackless reconstruction of Z $\rightarrow \mu^+\mu^-$ events in data and MC. These plots shown very nice data and MC agreement for right and wrong vertex choices in Z $\rightarrow \mu^+\mu^-$ events with the μ's tracks removed [Figures (a) and (b)] and γ+jet events with the tracks near the jet removed [Figures (c) and (d)] as a function of BDT output for 7 TeV and 8 TeV datasets. These assessments validate the BDT with and without the presence of conversions.

The Z $\rightarrow \mu^+\mu^-$ validation of the vertex BDT only validates it when conversions are absent. To validate its behavior with conversions, events with one converted photon and one jet are analyzed. The tracks associated to the jet gives the vertex of the photon in data and MC. Both the jet and the photon must have $P_T > 30$ GeV. To mimic the appearance of two isolated photons, the tracks within $\Delta R < 0.5$ of the jet are omitted in computing the other inputs of the vertex BDT. Again there is very good agreement between data and MC. The comparison is shown in Figure 5.13c and Figure 5.13d.

These validations provide confidence that the BDT output will be very similar for data and MC. This assurance is important for H $\rightarrow \gamma\gamma$ signal MC where there is not currently a large data set with which direct comparisons of resonant di-photons can be performed. Finally after the BDT is fully validated the efficiency as a function of the BDT in H $\rightarrow \gamma\gamma$ events can be thoughtfully evaluated. Figure 5.14 shows this efficiency with and without the use of conversions as a function of Higgs' p_T. Although the gaussian resolution of vertices which are reconstructed and chosen correctly is of the order $\sim 20\ \mu$m, the mass resolution is not significantly impacted until the vertex is misplaced by about 1 cm. Thus, the vertex is deemed correct if the selected vertex is within 1 cm in the z-direction of the generator vertex. For low p_T Higgs the efficiency of the vertex BDT is lowest. The use of conversions helps the most here. Overall, the integrated efficiency of the vertex BDT is 75% in H $\rightarrow \gamma\gamma$ events.

5.6.3 Per-Event Vertex Probability

Being able to estimate the per-event probability of the choosing the correct vertex is useful for making mass resolution estimates given that choosing the correct vertex greatly improves $m_{\gamma\gamma}$ resolution. To this end an additional BDT on top of the vertex BDT has been trained and using a simple linear transformation, it gives accurate probabilities of the chosen vertex's correctness. This section details its derivation and in Section 6.3 its use within the main analysis will be detailed. The following are inputs to this BDT:

- di-photon p_T
- number of reconstructed vertices

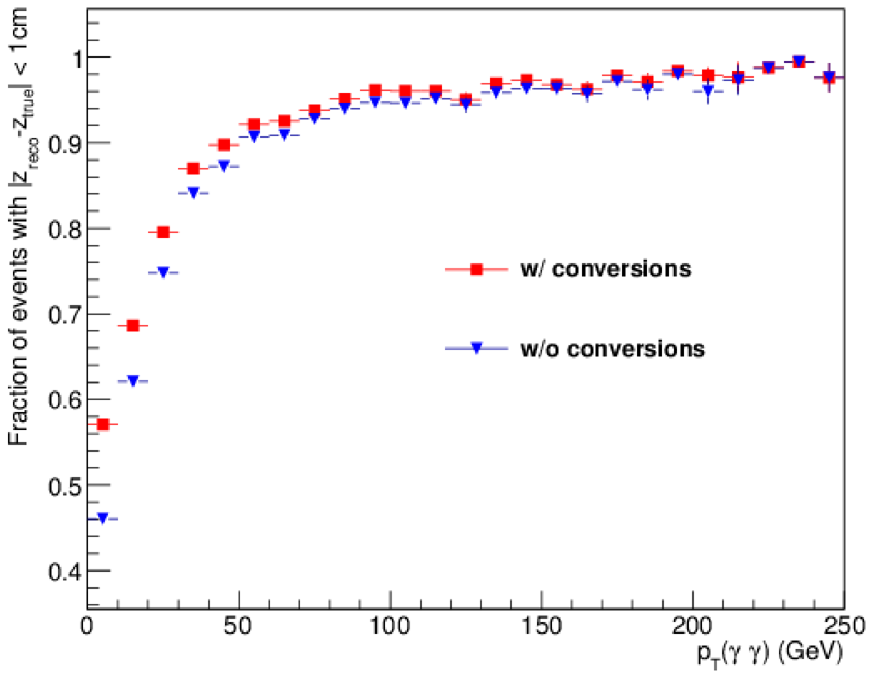


Figure 5.14: Vertex selection efficiency. The efficiency of the vertex selection within 1 cm of MC truth as a function of the Higgs' p_T. The use of conversions significantly helps in the low p_T range.

- per-vertex BDT values for the top three vertices
- Δz between the 1st and 2nd ranked vertices
- Δz between the 1st and 3rd ranked vertices
- number of photons associated to conversions

Again the task of validating the vertex probability BDT goes to Z $\rightarrow \mu^+\mu^-$ events for the events without conversions and to γ+jet events for events with conversions. A similar suite of comparisons is performed on the input variables and the output. The output BDT is shown in Figure 5.15 for 2011 and 2012, data and MC for both of these validations. The overall behavior of the vertex probability BDT in data is in good agreement with the prediction from MC.

In order to estimate the vertex identification on an event-by-event basis, a linear transformation is performed on the BDT. The comparison between the true vertex identification efficiency and the average estimated vertex probability is shown in Figure 5.16 as a function of the reconstructed di-photon p_T and of the number of reconstructed vertices. These estimations are accurate within 3-5%.

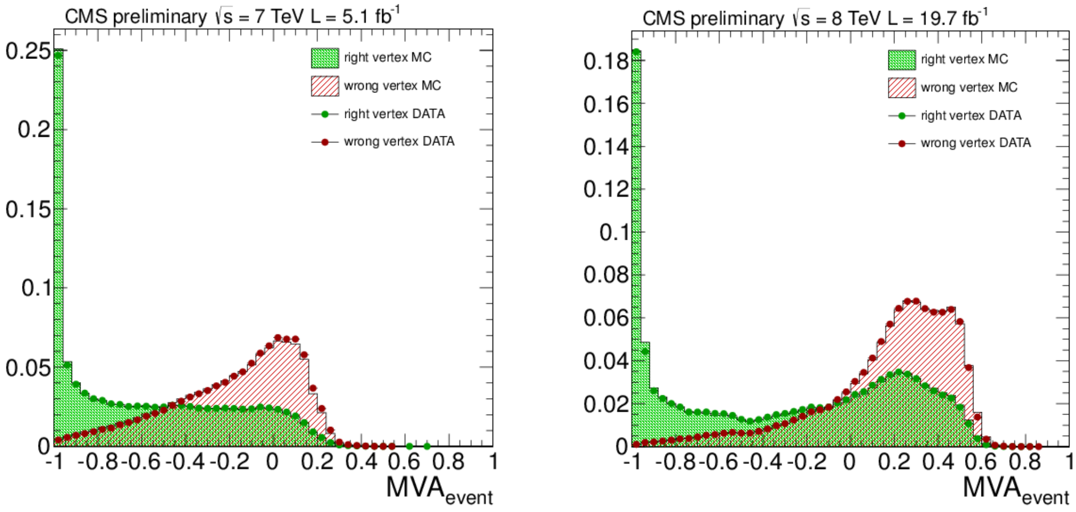
5.7 Particle Flow (PF) Event Description

The Particle Flow (PF) objects and algorithms are utilized for photon isolation variables, jet reconstruction, and E_T^{Missing} estimations in the $\text{H} \to \gamma\gamma$ analysis. PF is a step toward a general event description algorithm for organizing and identifying all individual particles including photons, electrons, muons, as well various and numerous low p_T neutral and charged hadrons. The PF algorithm uses information from all portions of the detector to construct this collection of particles [73].

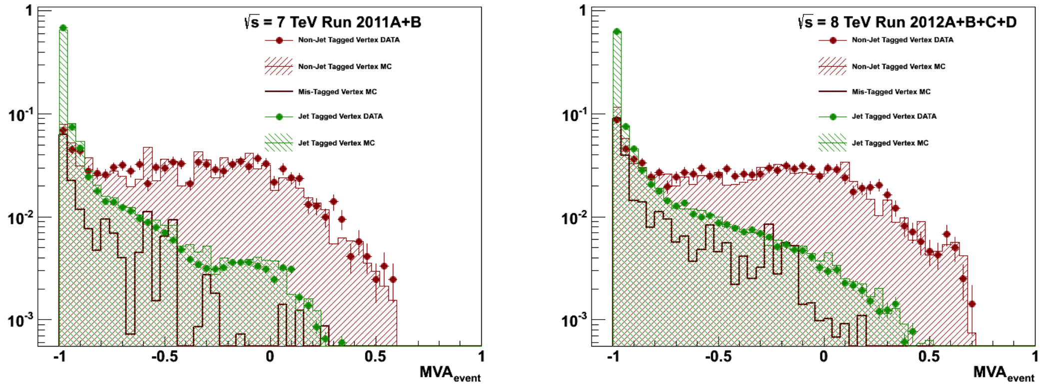
The first step in this process is a highly efficient, iterative tracking algorithm with a low fake rate on the order of a percent. In the first iteration very tight constraints are put on pixel and tracker hits as particle tracks are reconstructed. The hits associated with the most unambiguous tracks are removed from the following iterations and ever looser requirements are placed on the tracks to be added as PF candidates. In the fourth and fifth iterations the constraint on the position of the origin of the tracks is removed so that tracks from secondary vertices and photon conversions can be determined. Most of these tracks are from charged hadrons and this algorithm is more than 90% efficient in reconstructing them.

After the tracking is done, clusters of energy in the ECAL and HCAL are created. The aim of the clustering is to detect and measure the energy of photons and neutral hadrons so that they do not deteriorate the energy resolution in the charged hadrons. In addition, electron clusters will have Bremsstrahlung photons associated to them. A final goal for the clustering algorithm is to measure the energy of deposits to be associated with charged hadrons. In some cases these measurements will be nearly the complete measurement of charged hadron p_T.

Once the PF clusters and PF tracks are constructed they are linked together to form reconstructed PF candidate particles. Typically particles will have one to

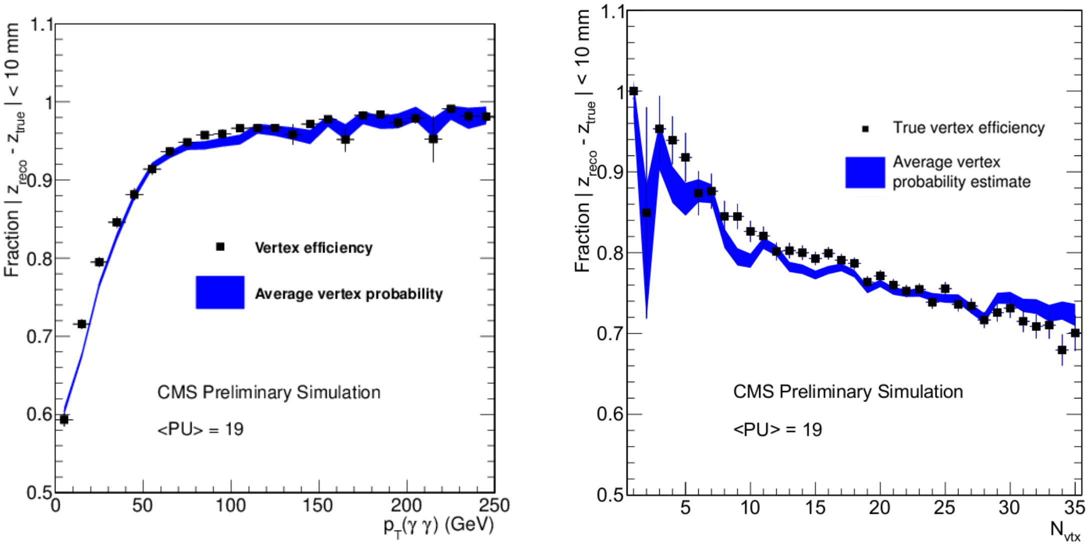


(a) 7 TeV - Z $\rightarrow \mu^+\mu^-$ (b) 8 TeV - Z $\rightarrow \mu^+\mu^-$



(c) 7 TeV - γ+jet (d) 8 TeV - γ+jet

Figure 5.15: Vertex probability BDT output in muon trackless reconstruction of Z $\rightarrow \mu^+\mu^-$ events in data and MC. These plots shown very nice data and MC agreement for right and wrong vertex choices in Z $\rightarrow \mu^+\mu^-$ events with the μ's tracks removed [Figures (a) and (b)] and γ+jet events with the tracks near the jet removed [Figures (c) and (d)] as a function of the vertex probability BDT output for 7 TeV and 8 TeV datasets.



(a) (b)

Figure 5.16: Vertex probability compared to actual efficiency as a function of Higgs' p_T in (a) and number of vertices in (b).

three sub-detector elements which must be linked. For example, the elements required in the linking for PF charged hadron are a PF track, PF ECAL cluster and PF HCAL cluster. For linking to occur the extrapolated position of the track must pass through (or within one cell of) the clusters. For PF HCAL clusters the extrapolated track position must pass through the cluster at a depth of $1\lambda_I$ (one hadronic interaction length).

In an attempt to recover Bremsstrahlung photons from the track of its parent electron, tangents to pre-selected PF electron tracks at each tracker layer are used to point to PF ECAL clusters. If a tangent trajectory passes through a PF ECAL cluster, then the track and the cluster are linked. Energy deposits from ES, ECAL, and HCAL which are very close in (η, ϕ) are linked as well. The final linking step is between muon tracks in the outer chambers and tracker. The PF track and the muon track are re-fit together and if the χ^2 of the fit is low enough linking occurs. If there are multiple tracks that produce acceptable χ^2, then only the track with the minimal χ^2 is linked to the muon track.

After all the linking has occurred particle identification can be done. Each particle type has its own identification criteria. Once particles are identified all of their associated elements are removed from the selection pool. Muons are the first particles to be selected and removed. After muons, electrons, charged hadrons, photons and neutral hadrons are identified (in that order). All PF elements may only be used in one object.

Once all of these individual particles are re-constructed, PF jets and PF taus are constructed from groups of various PF particles, and PF E_T^{Missing} is computed from the negative vectorial sum of all PF objects' transverse momentum. Both PF jets and PF E_T^{Missing} are used in events classes tagging jets and E_T^{Missing}. The resolution in the jets and E_T^{Missing} are better than their detector-based counterparts because PF objects are generally less impacted by pile-up. Moreover, the PF algorithm allows for much better vertex association of tracks and clusters. For this reason isolation based on PF objects is performed in this analysis.

5.8 Defining Key Variables

Throughout this analysis photon, jet and other event variables are used to sort and select events. These variables are used to select events beyond thresholds as in the photon pre-selection, cut-based identification, pile-up jet pre-selection, di-jet tag for VBF, and other tags. In other cases these variables will be used as inputs for BDTs which will be used for classification or cut on instead of these variables directly.

This section is primarily written for reference.

5.8.1 Event Variables

The following variables are related to pile-up.

- $\mathrm{N}_{\mathrm{Vtx}}$
  - The number of reconstructed vertices. This number is most often less than the number of pile-up interactions due to close vertices merging and reconstruction inefficiency.
- ρ

  - This is the per-event energy density from pile-up as measured using jet areas as computed by the FastJet package [74].

5.8.2 Photon Variables

There are two main types of photon variables: shower-shape and isolation. Shower-shape variables generally help to distinguish real photons from fake photons. In addition, for real photons shower-shape variables measure if and to what degree a photon has converted in the tracker. Isolation variables measure energetic activity near but outside of the super-cluster. Large values can indicate that the photon is a fake from a jet or that there are large amounts of pile-up activity in that region. For either reason non-isolated photons are undesirable.

- R_9
  - Ratio of 3×3 crystals around the seed crystal and super-cluster energy
- $\sigma_{\mathrm{i}\eta\mathrm{i}\eta} = \sqrt{\frac{\sum(\eta_i - \bar{\eta})^2 w_i}{\sum(w_i)}}$ where $\bar{\eta} = \frac{\sum(\eta_i) w_i}{\sum(w_i)}$ and $w_i = \max\left(0, 4.7 + \log\left(\frac{E_i}{E_{5\times5}}\right)\right)$
  - The η-width in the 5×5 crystal surrounding the maximum energy crystal in the super-cluster's basic cluster seed.
- $\mathrm{cov}_{\mathrm{i}\eta\mathrm{i}\phi} = \frac{\sum(\eta_i - \bar{\eta})\left(\phi_i - \bar{\phi}\right) w_i}{\sum(w_i)}$
  - The $\eta - \phi$ covariance of the 5×5 matrix around the basic cluster seed.
- $\sigma_{\eta,\mathrm{SC}}$
  - The energy weighted standard deviation of single crystal η for crystals composing a super-cluster.
- $\sigma_{\phi,\mathrm{SC}}$
  - The energy weighted standard deviation of single crystal ϕ for crystals composing a super-cluster.
- $\mathrm{E}_{2\times2}/\mathrm{E}_{5\times5}$

  - $E_{2\times2}$ is the energy in the 2x2 crystals with maximum energy containing the crystal with maximum energy.
  - $E_{5\times5}$ is the energy in the 5x5 crystals centered at the crystal with maximum energy.

- σ_{RR}
  - The standard deviation of the shower spread in the x and y planes of the ES.

Isolation variables serve to distinguish photons from jets with a different strategy. These variables seek out extra activity outside of the photon, which is the super-cluster. In the absence of many pile-up interactions photons may be very well isolated and energy deposits near the super-clusters in the ECAL and HCAL may be indicative of a reconstructed super-cluster not being a photon. However, the level of pile-up experienced in 2012 makes detector-based isolations very detrimental to signal efficiency unless they are used very loosely. The Particle Flow algorithm described in Section 5.7 has been designed to efficiently link signatures of particles from various sub-detectors. The resulting PF collection can be used to construct isolation variables which are relatively independent of pile-up.

- PFCharged
  - $\sum\limits_{\Delta R<\Delta R_{Max}} \left(P_{T,i}^{\text{Charged}}\right)$ where ΔR_{Max} is typically 0.2 or 0.3.
  - Only PF charged hadrons associated to the chosen vertex (see Section 5.6 for choice algorithm) are considered in this scalar sum of transverse momentum.
  - PF charged hadrons within $\Delta R < 0.2$ are summed. However, those within $\Delta R < 0.02$ are vetoed.
- PFChargedWorst
  - $\sum\limits_{\Delta R<\Delta R_{Max}} \left(P_{T,i}^{\text{Charged}}\right)$ where ΔR_{Max} is typically 0.3 or 0.4.

  - Only PF charged hadrons associated to the vertex with maximum isolation are considered in this scalar sum of transverse momentum.
  - PF charged hadrons within $\Delta R < \Delta R_{Max}$ are summed. However, those within $\Delta R < 0.02$ are vetoed.

- PFPhoton
  - $\sum\limits_{\Delta R<\Delta R_{Max}} \left(P_{T,i}^{\text{Photon}}\right)$ where ΔR_{Max} is typically 0.3 or 0.4.
  - Transverse momentums of all PF photons within $\Delta R < \Delta R_{Max}$ of the photon are summed.
  - There is no association with any vertex.

- PFNeutral
  - $\sum\limits_{\Delta R<\Delta R_{Max}} \left(P_{T,i}^{\text{Neutral}}\right)$ where ΔR_{Max} is typically 0.3 or 0.4.
  - Transverse momentums of all PF neutral hadrons within $\Delta R < \Delta R_{Max}$ of the photon are summed.
  - There is no association with any vertex.

- $\text{PFIsoSum}_{\text{ChosenVtx}} = (\text{PFCharged03} + \text{PFPhoton03} + 2.5 - 0.09 * \rho)\,\frac{50\text{GeV}}{E_{T,\gamma}}$
  - The ρ subtraction removes average pile-up activity.

- $\text{PFIsoSum}_{\text{WorstVtx}} = (\text{PFCharged04Worst} + \text{PFPhoton04} + 2.5 - 0.23 * \rho)\,\frac{50\text{GeV}}{E_{T,\gamma}}$
  - The ρ subtraction is larger to account for a larger cone and for a bias toward greater activity given the selection of the worst isolation.

- ECALIso
  - The scalar sum of RecHits' transverse energy inside a cone of $\Delta R < 0.3$.
  - Crystals within 3.5 crystals from the central crystal are omitted.
  - Crystals within 2.5 crystals from the super-cluster's η are omitted.

- HCALIso

  - The scalar sum of HCAL RecHits' transverse energy inside a cone of $\Delta R < 0.3$.
  - The inner footprint within $\Delta R < 0.015$ is ommitted from the sum.
- TrkIso
  - The scalar sum of tracks' transverse momentum inside a cone of $\Delta R < 0.3$.
  - The tracks within $\Delta R < 0.02$ of the super-cluster are ommitted from the sum.
- H/E
  - Ratio of hadronic energy behind a super-cluster and super-cluster energy.

5.8.3 Jet Variables

The following variables are used to reject jets from pile-up.

- $\beta^* = \frac{\sum_{\text{Jet tracks not compatible with chosen vertex}} P_{T,\text{Trk}}}{\sum_{\text{All Jet Tracks}} P_{T,\text{Trk}}}$
  - This is the fraction of the transverse momentum sum of jet tracks which are from vertices other than the selected one. When this value is large, the jet is incompatible with the selected vertex.
- $\overline{\Delta R^2} = \frac{\sum_{\text{Particles in Jet}} \left(P^2_{T,\text{Trk}} (\Delta R(\text{Trk, Jet Center}))^2 \right)}{\sum_{\text{Particles in Jet}} P^2_{T,\text{Trk}}}$
  - $\overline{\Delta R^2}$ is a measure of how wide a jet is. Wider jets tend to be from close jets which have merged.

Recalling the Feynman diagram for Vector Boson Fusion (see Figure 4.2b) one can see that there are two quarks in the final state. These quarks produce high p_T jets which are forward. They typically have high separation because the two quarks originate from different protons. These topological features lead to a large invariant mass, M_{JJ}, and large separation in η, $\Delta\eta_{JJ}$. In addition, the di-jet system should also have topological features relative to the di-photon system; those being similar values of η and opposite azimuthal position reflecting the recoil of the di-jet against the Higgs Boson. These di-jet variables are used in the di-jet tags to be described in Section 6.2.3 and Section 6.3.4.

- M_{JJ}
  - Di-jet invariant mass. Powerful discriminant between signal and background. Since the VBF signature has two high p_T forward jets, the energies and the opening-angle are both large, yielding a high invariant mass.
- $\Delta\eta_{JJ} = \eta_{J_1} - \eta_{J_2}$
  - Separation in η of the two selected jets. High separation between the jets is expected and required.
- $Z = \eta_{\gamma\gamma} - \frac{\eta_{J_1}+\eta_{J_2}}{2}$
  - The Zeppenfeld variable, defined in [75]. This variable represents compatibility between the di-photon and the di-jet systems in η. It is expected that the average pseudorapidity of the jets will be comparable to the pseudorapidity of the di-photon. Thus, for the compatible di-photon plus di-jet events Z will be close to 0.
- $\Delta\phi = |\phi_{\gamma\gamma} - \phi_{JJ}|$
  - The difference in azimuthal angle between the di-photon and di-jet systems, which is less than or equal to π.
  - This variable represents compatibility between the di-photon and the di-jet systems in ϕ. The di-photon and the di-jet are expected to be back-to-back in signal, yielding a peaking $|\Delta\phi| \sim \pi$.
  - Very large theory uncertainty in this variable [76] near π for the significant 2 jet signal contribution from Gluon Fusion production moti-

vates the blunting of this variable by using in training and in practice $\min(\Delta\phi, \pi - 0.2)$ in the di-jet kinematic BDT.

5.8.4 Lepton Variables

Defined here are isolation and identification variables used for tagging muons in the VH Leptonic Tight and Loose Tags (see Section 6.4).

- $\text{PFIsoSum}_\mu = \frac{\text{PFCharged04}_\mu + \max(0, \text{PFPhoton04}_\mu + \text{PFNeutral04}_\mu - \Delta B)}{p_T^\mu}$
  - PFCharged04_μ, PFPhoton04_μ, and PFNeutral04_μ are defined analogously to PF candidate sums defined in Section 5.8.2 for photons except here there is no inner veto cone.
  - ΔB is a per-event pile-up correction.
- Muon Tight ID
  - Number of pixel hits > 0
  - Track fit quality, $\chi^2/\text{n.d.f} < 10$
  - Number of muon hits > 0
  - Number of matched muon stations > 1
  - Number of tracker layers > 5
  - d_0 w.r.t. selected vertex < 0.2 cm
  - d_z w.r.t. selected vertex < 0.5 cm
- Full muon selection
  - $\text{PFIsoSum}_\mu < 0.2$
  - Muon Tight ID
  - $|\eta_\mu| < 2.4$

Defined here are isolation and identification variables used for tagging electrons in the VH Leptonic Tight and Loose Tags (see Section 6.4).

- $\text{PFIsoSum}_{\text{Ele}} = \frac{\text{PFCharged03}_{\text{Ele}} + \max(0, \text{PFPhoton03}_{\text{Ele}} + \text{PFNeutral03}_{\text{Ele}} - A_{\text{Eff}} * \rho)}{p_T^{\text{Ele}}}$
  - $\text{PFCharged03}_{\text{Ele}}$ is defined analogously to PFCharged03 in Section 5.8.2 for photons with an inner veto of $\Delta R < 0.015$ in the EE and none in the EB.
  - $\text{PFPhoton03}_{\text{Ele}}$ is defined analogously to PFPhoton03 defined in Section 5.8.2 for photons with an inner veto of $\Delta R < 0.08$ in the EE and none in the EB.
  - $\text{PFNeutral03}_{\text{Ele}}$ is defined analogously to PFNeutral03 defined in Section 5.8.2 for photons with no inner veto cone.
  - A_{Eff} is a scaling factor for the ρ correction which is dependent on $|\eta|$.
- MVA_{Ele}
  - BDT with various electron shower-shape variable inputs.
  - Very similar in concept to photon ID BDT in that it also discriminates between real electrons and fakes from jets.
  - Trained with low p_T electrons in particular.
- $N_{\text{MissingHits}}$
  - The number of active tracker layers where the electron's track does not have a corresponding hit.
- $\widetilde{\chi^2}_{\text{Ele}} = \chi^2/\text{n.d.f}$ - track fit quality
- $d_{0,\text{Ele}}$ - radial distance at track's point of closest approach to selected vertex
- $d_{z,\text{Ele}}$ - distance in z at track's point of closest approach to selected vertex
- Full electron selection
  - $\text{PFIsoSum}_{\text{Ele}} < 0.15$
  - $\text{MVA}_{\text{Ele}} > 0.9$
  - $N_{\text{MissingHits}} < 2$

- $\widetilde{\chi^2}_{\mathrm{Ele}} < 10^{-6}$
- $d_{0,\mathrm{Ele}} < 0.02$ cm
- $d_{z,\mathrm{Ele}} < 0.2$ cm
- $|\eta_{SC}| < 1.4442$ or $1.566 < |\eta_{SC}| < 2.5$

Chapter 5, in full, has been prepared for publication submission to EPJC: "Observation of the diphoton decay of the Higgs boson and measurement of its properties". The dissertation author was the primary investigator and author of this paper.

Chapter 6

Final CMS $\mathrm{H} \to \gamma\gamma$ Analysis with Run 1 Data Description

Grouping events with similar mass resolution is often an optimal classification scheme. In general classes of event with different signal to background (S/B) ratios should be analyzed in distinct event classes to take advantage of the best classes while still utilizing others. Events with better resolution typically have better S/B and so in both analyses most events are classified using variables which are indicative of mass resolutions.

In addition, production of a Higgs Boson via VBF, VH, or $\mathrm{t\bar{t}H}$ results in additional final state objects. Searching for and selecting these objects (e.g. two jets, a muon, an electron or high E_T^{Missing}) in addition to two photons creates very high S/B event classes because the background is significantly reduced. The most significant of these tags is the di-jet tag targeting the VBF production of a Higgs Boson decaying to two photons.

The forthcoming two analyses use the same basic strategy, which is to fit the data to estimate the background. Likewise, both analyses use exactly the same methodology for determining the energy of the photons and for selecting the vertex. However, the selection of events and classification strategies differ significantly. The first analysis consists of simple, robust cuts on photons and jets and classifies events based on basic photon and di-photon variables. It will be referred to as the Cut-based Analysis. The second is a more fully optimized analysis, which relies heavily on Boosted Decision Trees (BDTs) to classify events based on photon and jet information after a very loose pre-selection. It will be referred to as the MultiVariate Analysis or MVA.

This chapter describes the selection requirements and the classification schemes of both analyses. In addition signal and background models are described and a full review of systematic uncertainties is detailed.

6.1 Photon and Jet Pre-selections

For all events in each of the analyses there must be two photons passing the following pre-selection. This pre-selection is imposed such that the data and MC cover the same phase space. In particular, the MC does not have the HLT applied explicitly but rather the pre-selection uses nearly equivalent offline variables for the HLT variables. In addition, there are cuts that mimic the generator filter for the QCD and γ+jet Pythia samples for the data. All of these cuts have negligible impact on the signal MC. The pre-selection cuts are summarized in Table 6.1.

Table 6.1: Photon preselection cuts. The variables below are defined in Section 5.8.2.

Variables	EB $R_9 > 0.9$	EB $R_9 \leq 0.9$	EE $R_9 > 0.9$	EE $R_9 \leq 0.9$
$\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$	0.014	0.014	0.034	0.034
H/E	0.082	0.075	0.075	0.075
PFCharged02	4 GeV	4 GeV	4 GeV	4 GeV
HCALIso03$-0.05 \times p_T$	50 GeV	4 GeV	50 GeV	4 GeV
TrkIso03$-0.05 \times p_T$	50 GeV	4 GeV	50 GeV	4 GeV

Applied to each photon is also a conversion-safe electron veto. Any supercluster that is matched to an electron without any missing hits (aside from non-functioning pixels and strips) in the tracker is rejected as a photon because of the very high likelihood that it is an electron. This is always applied except in the case of $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ validation where the electron veto is inverted in order to explicitly select a control sample with no overlap with the selection.

In addition to selection on shower-shape and isolation there are kinematic requirements on the photons and on $m_{\gamma\gamma}$. The leading p_T photon and sub-leading p_T photon must pass the following cuts respectively, $p_T/m_{\gamma\gamma} > 1/3$ and $p_T/m_{\gamma\gamma} > 1/4$. For most cut-based exclusive channels the sub-leading photon's $p_T/m_{\gamma\gamma}$ cut is relaxed to $p_T > 25$ GeV, which is the loosest that the cut can be and still be fully within the acceptance of the HLT. Indeed for some channels the truly optimal cut may be even looser. In addition to all of these requirements, if $m_{\gamma\gamma}$ is not in the range 100-180 GeV, the event will be dropped. This range is needed for fits which give unbiased background estimates in the range of 110-150 GeV.

6.1.1 Jet Pre-selection

Not all events are required to have jets but in categories where jets are present, the following pre-selections are applied.

In 2012 pile-up became a very important issue for jets. Because of the increased pile-up (an average of 20 interactions per recorded event) there are more hard jets from the pile-up interactions. In addition there are more closely overlapping jets, which tend to be constructed as very wide single jets. To combat these symptoms of pile-up many variables and algorithms were tested and the most powerful variables to filter out these pile-up jets were employed. β^* and $\overline{\Delta R^2}$, defined in Section 5.8.3, were found to be most useful. β^* gives a measure of how compatible the jet in question is with the selected vertex. Near the end of the tracker coverage β^* cannot be used due to lack of tracks. Different cuts on these two variables are applied in the different regions of the detector. Table 6.2 lists all jet pile-up rejection pre-selection cuts.

Table 6.2: Pile-up jet preselection cuts. Notice for higher values of $|\eta|$ cuts on β^* are not applied because of the lack of tracks.

$\lvert\eta_{\text{Jet}}\rvert$	β^*	$\overline{\Delta R^2}$
$\lvert\eta\rvert < 2.5$	$< 0.2 \log N_{vtx} - 0.64$	< 0.06
$2.5 < \lvert\eta\rvert < 2.75$	$< 0.3 \log N_{vtx} - 0.64$	< 0.05
$2.75 < \lvert\eta\rvert < 3$	-	< 0.05
$3 < \lvert\eta\rvert < 4.7$	-	< 0.055

Given the difficulty of correcting jet energy in the most forward region of the HF, jet pseudorapidity is limited to $|\eta| < 4.7$. Before jet pre-selection the two photons are selected. Jets are required to have $\Delta R > 0.5$ between the jet in question and either of the photons. For event classes where two jets are required, the two jets with the highest p_T satisfying all of the above requirements are selected. In the di-jet tags targeting VBF production, the leading jet must have $P_T^{\mathrm{J}_1} > 30$ GeV, the sub-leading jet must have $P_T^{\mathrm{J}_2} > 20$ GeV, and the di-jet must have M_{JJ} >250 GeV.

6.2 Cut-based Analysis

The purpose of the cut-based analysis is to provide a simple, robust analysis whose basis is cuts on straight-forward photon identification variables. The energies of the photons are determined using the methods described in Section 5.5 and the vertex is selected using the methods described in Section 5.6. In addition, events must pass the preselection detailed in Section 6.1.

6.2.1 Cut-based Photon Identification (ID)

The photon identification is optimized in $2|\eta| \times 2R_9$ categories and photons from the transition region between the EB and EE ($1.4442 < |\eta| < 1.566$) are excluded from this analysis. This photon ID is called Cuts in Categories 4 - Particle Flow (CiC4PF) because there are separate cuts on particle flow isolation in four photon categories. The η categories are simply EB and EE, requiring the photon to have $|\eta| < 1.4442$ and $1.566 < |\eta| < 2.5$, respectively, because the detectors are very different. The R_9 categories are $R_9 > 0.94$ and $R_9 \leq 0.94$ to roughly divide converted and unconverted photons.

The variables used for the CiC4PF are a mixture of shower-shape and isolation variables. The Cuts in Categories (CiC) method sets the cut values by iterating on distributions which estimate signal to background (S/B) ratios for each variable as a function of that variable. These distributions require all of the previous iterations cuts except the one on that variable. These types of plots are thus called $(n-1)$ plots. A particular and arbitrary S/B value is chosen and the cut value for each variable is set at the value where that variable's S/B estimation reaches the chosen S/B value. The program iterates until the cuts become stable. More detail on this procedure, including an annotated $(n-1)$ plot example, is given in Section 4.2.1

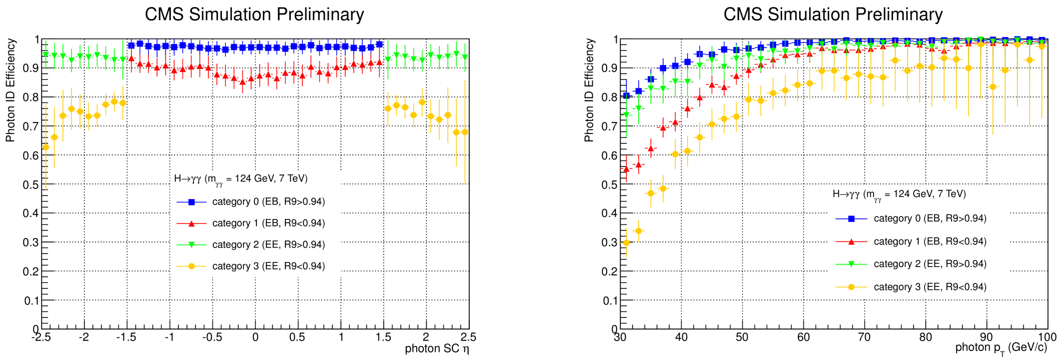
This method is applied separately to events in the four categories of η-R_9 listed above. The split between EB and EE is because there are many differences in S/B and in variable shape in the different detector regions. High R_9 photons have narrower shower-shapes because they tend to be from photons that do not shower in the tracker. Therefore, the splitting in R_9 is used to separate showering and non-showering photons which is necessary given the different distributions. In the final cuts, all of the thresholds are tighter in the low R_9 photon categories which indicates that to achieve the same S/B non-showering photons must be more tightly selected.

Table 6.3: Cut-based identification selection cuts for the CiC4PF SuperTight cut-level.

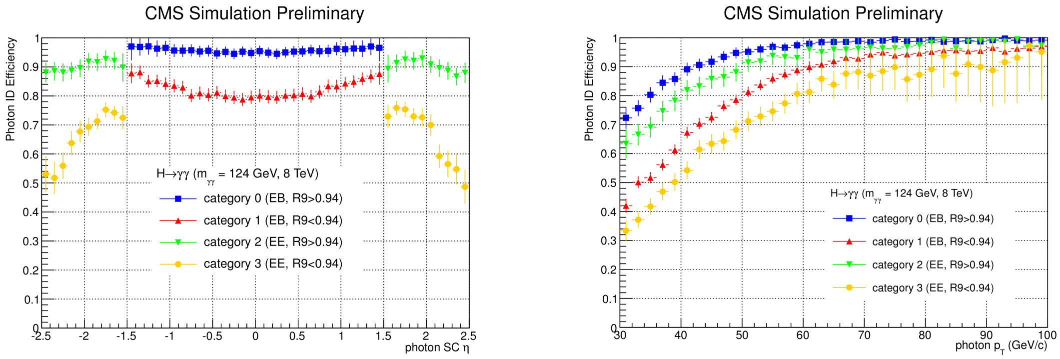
	Barrel		Endcap	
	$R_9 > 0.94$	$R_9 \leq 0.94$	$R_9 > 0.94$	$R_9 \leq 0.94$
$\text{PFIsoSum}_{\text{ChosenVtx}}$	6	4.7	5.6	3.6
$\text{PFIsoSum}_{\text{WorstVtx}}$	10	6.5	5.6	4.4
PFCharged03	3.8	2.5	3.1	2.2
$\sigma_{\text{i}\eta\text{i}\eta}$	0.0108	0.0102	0.028	0.028
H/E	0.124	0.092	0.142	0.063
R_9	0.94	0.298	0.94	0.24

There are numerous different values of S/B chosen and the result is numerous CiC working points. It is required in the cut-setting program that no cut values be looser in a tighter working point than in a looser working point. However, the cut value can remain the same. The level chosen for this analysis is the so-called "SuperTight" working point.

Since the CiC working points are set in S/B, signal efficiency is not necessarily uniform between categories or as a function of p_T and η. This is shown in plots of signal efficiency for 7 TeV and 8 TeV Higgs MC in Figure 6.1. Alterna-



(a) 7 TeV - η (b) 7 TeV - p_T



(c) 8 TeV - η (d) 8 TeV - p_T

Figure 6.1: CiC4PF efficiency for signal photons $M_H = 124$ GeV for 7 TeV (a), (b) and 8 TeV (c), (d) simulations as a function of pseudorapidity and p_T in each of the four η-R_9 photon categories.

tively stated, to achieve the same S/B in all photon categories some cuts must be tighter in the different categories. This is clear when one examines the cut values detailed in Table 6.3.

6.2.2 Inclusive Di-photon Classification

The classification for this analysis is based on simple photon and di-photon variables which are indicative of differing levels of photon resolution and S/B. Given that photon resolution is best in the EB, the first level of classification selects events where both photons are in the EB. The next level of classification reflects the different levels of background for photons with narrow and wide shower-shape variables, namely R_9. When both photons have $R_9 > 0.94$ (the same boundary as the CiC4PF R_9 factorization), they are selected for preferred event classes.

The final level of classification is based on the $p_{T,\gamma\gamma}/m_{\gamma\gamma}$ of the di-photon system. Here a 5% improvement in expected limits is found when only considering the inclusive event classes. This improvement mainly comes from the significantly reduced background in the higher $p_{T,\gamma\gamma}/m_{\gamma\gamma}$ category. In addition, the higher $p_{T,\gamma\gamma}/m_{\gamma\gamma}$ event classes have much higher fractions of expected signal events from Higgs production mechanisms that require coupling to vector bosons. This improves our ability to measure the Higgs' coupling to vector bosons.

To summarize, there are eight inclusive event classes constructed using by dividing di-photons into EB and EE, high and low R_9, and high and low $p_{T,\gamma\gamma}/m_{\gamma\gamma}$. These event classes are used for 7 TeV and 8 TeV datasets.

6.2.3 Cut-based Di-jet Tag

There are numerous additional production mechanism signature tags which are applied in this analysis and to the more complex main analysis, but the one that contributes most significantly to expected sensitivity is the di-jet tag targeting the VBF production mechanism. A simple, two-category di-jet analysis utilizing the di-jet pre-selection described in Section 6.1.1 is employed and summarized in Table 6.4 along with the previous incarnation of this analysis (the single category di-jet tag) for reference.

Table 6.4: Cut-based di-jet tags for VBF production. The two category cut-based di-jet tag (labeled Tight and Loose) are listed as well as the previously published [48] single category di-jet for comparison.

Variable	Tight	Loose	Single Category
$p_T^{\gamma_1}/m_{\gamma\gamma}$	> 0.5	> 0.5	$> 55/120$
$p_T^{\gamma_2}$	> 25 GeV	> 25 GeV	> 25 GeV
$p_T^{\mathrm{J}_1}$	> 30 GeV	> 30 GeV	> 30 GeV
$p_T^{\mathrm{J}_2}$	> 30 GeV	> 20 GeV	> 20 GeV
M_{JJ}	> 500 GeV	> 250 GeV	> 350 GeV
$\lvert\Delta\eta_{JJ}\rvert$	> 3.0	> 3.0	> 3.5
$\lvert Z\rvert$	< 2.5	< 2.5	< 2.5
$\lvert\Delta\phi(\gamma\gamma, JJ)\rvert$	> 2.6	> 2.6	> 2.6

The most powerful variable is M_{JJ} which is the most significant difference between the tight and loose di-jet categories. Indeed, M_{JJ} is the variable which is most significantly changed between the early publication's single category and the two category analysis presented here [48].

Additional tags in the cut-based analysis will be briefly described in Section 6.4 and results for this analysis will be forthcoming in Section 7.

6.3 MultiVariate Analysis (MVA)

Fundamentally this analysis is very similar to the cut-based analysis. Both rely on the same key ingredients: fits to the data for background estimates, same procedure for photon energy determination, inclusive classification related to mass resolution and exclusive classification based on production mode signatures. However, instead of using cuts on variables for photon ID and simple inclusive classification, Boosted Decision Trees (BDTs) are used for both. The advantage in optimization is significant as the BDT utilizes the correlations among the many variables.

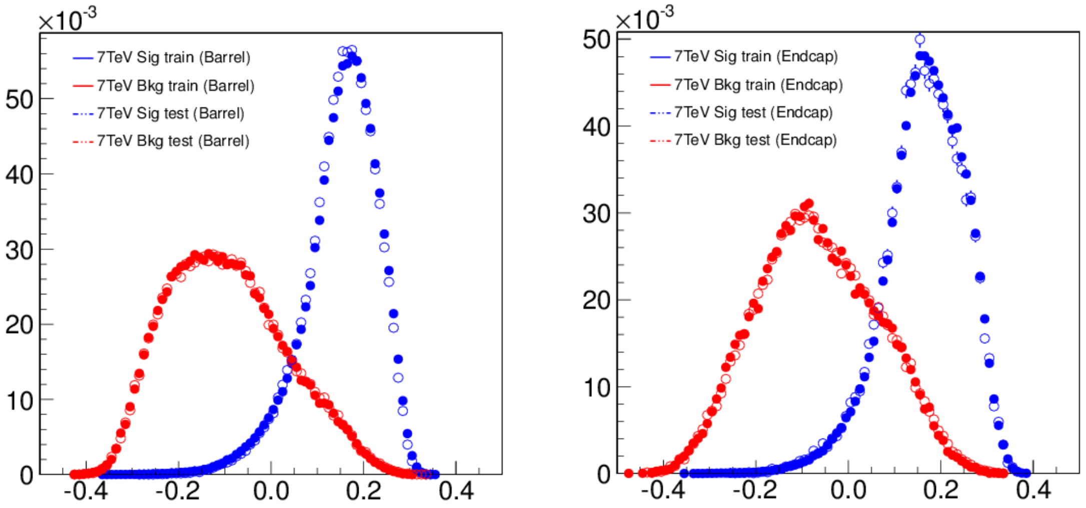
6.3.1 Photon ID BDT

The photon ID BDT discriminates between real photons and "fake" photons from jets. High BDT values are more likely to be photons and low values are more likely to be jets. For the training of this BDT, Pythia γ+jet samples were used for separate trainings at $\sqrt{(s)} = 7$ TeV and 8 TeV and for EB and EE for a total of four BDTs. All prompt photons from the γ+jet samples were used as the signal for training and half of the jets from each event were used as the background for training. For testing the output of the photon ID, photons from $\mathrm{H} \to \gamma\gamma$ via gluon fusion samples (at 121 GeV for 7 TeV and at 124 GeV for 8 Tev) were used. The testing background sample is the other half of the jets from the γ+jet samples. In Figure 6.2 the testing and training samples for signal and background for 7 TeV and 8 TeV ID BDT outputs are overlaid. There is no noticable difference in the training and testing samples ID BDT outputs in signal or background, nor has any other evidence of over-training been found.

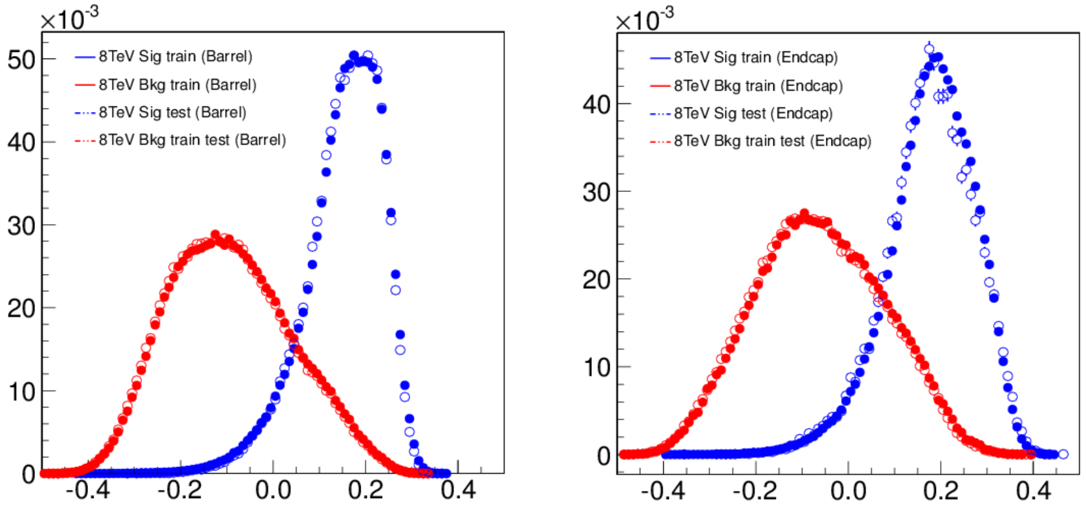
The photon ID BDT uses both shower-shape and isolation variables, just as the CiC4PF does. However, more shower-shape variables are included in the training. In a cut-based analysis these additional variables would not be discriminating enough to warrant an additional cut, but a BDT can use its correlation with other variables and therefore become useful to this photon ID. These are the shower-shape variables included in the BDT training: $\sigma_{\mathrm{i}\eta\mathrm{i}\eta}$, $\mathrm{cov}_{\mathrm{i}\eta\mathrm{i}\phi}$, R_9, $\sigma_{\eta,\mathrm{SC}}$, $\sigma_{\phi,\mathrm{SC}}$, $\mathrm{E}_{2\times2}/\mathrm{E}_{5\times5}$, and σ_{RR} (for EE).

In addition, the individual components of isolations are included separately instead of adding isolation sums and subtracting a pile-up correction. The BDT will combine these variables in an effective way without them being explicitly combined beforehand and again the BDT takes advantage of the correlations among these variables. These are the isolation variables included for the training: PFPhoton03, PFCharged03, PFCharged03Worst, and the average pile-up energy density, ρ. See Section 5.8.2 for definitions of all variables. In addition, the super-cluster's raw energy, E_{Raw}, and η_{SC} are also included.

Photon p_T independence was an explicit goal in this training for the purpose of mass independence in the di-photon BDT. To make the BDT training as flat



(a) 7 TeV - EB (b) 7 TeV - EE



(c) 8 TeV - EB (d) 8 TeV - EE

Figure 6.2: Photon identifcation BDT output. These are four photon ID BDTs (7 TeV, 8 TeV)×(EB, EE). For each there is a comparison of the training sample with an equal-sized test sample for signal (prompt photons from γ+jet MC) and background (reconstructed photons from jets from γ+jet MC). The matching between the training and testing samples shows that there is no over-training.

as possible in photon p_T and η_{SC}, the signal prompt photons are re-weighted to match the spectrum of the reconstructed photons from jets used as background. In Figure 6.3 and Figure 6.4 the 7 TeV and 8 TeV photon ID efficiencies as a function of p_T are plotted for signal and background. Overlaid with them is the previous version of the photon ID. In the current version of the photon ID there is no significant loss in efficiency at low p_T.

6.3.2 Di-photon Mass Resolution Estimates

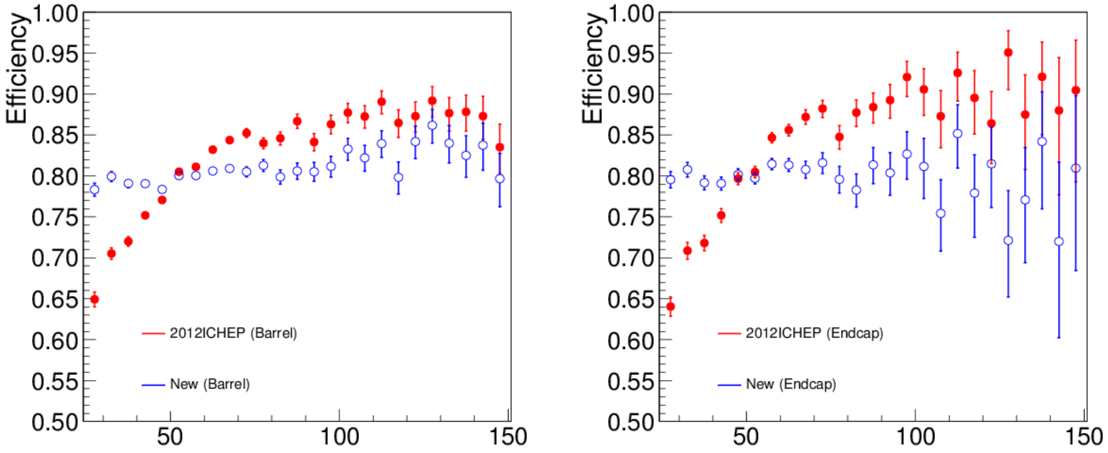
Estimates of mass resolutions are very important inputs to the di-photon BDT, which will be described in Section 6.3.3. There are two cases, for which the resolution will be very different, and the resolutions must be calculated separately.

The more straight-forward case is for the correct choice of the di-photon's vertex. In this scenario the uncertainty on the opening angle between the two photons is negligible in comparison to the error in energy measurement. Thus, Equation 5.5.0.1 simplifies to:

$$\left(\frac{\sigma_{m_{\gamma\gamma}\text{Right Vtx}}}{m_{\gamma\gamma}}\right) = \frac{1}{2}\sqrt{\left(\left(\frac{\sigma_{E_1}}{E_1}\right)^2 + \left(\frac{\sigma_{E_2}}{E_2}\right)^2\right)} \tag{6.3.2.1}$$

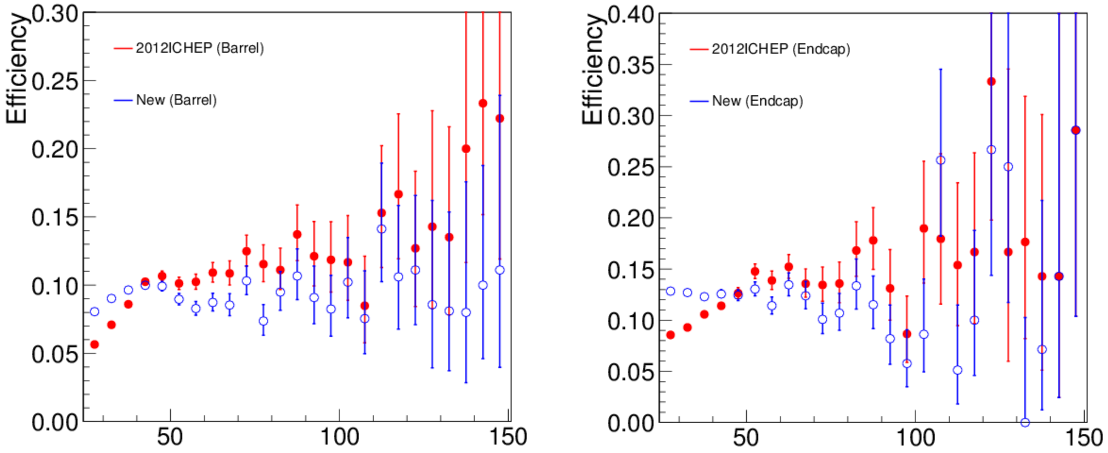
σ_{E_1} and σ_{E_2} are primarily estimated from σ_{CB} derived from the energy regression described in Section 5.5.2. However, the additional energy smearing needed to make Z → e^+e^- MC match the data (described Section 5.5.3) is also needed here for a full estimate of the energy of the photons. The two resolution estimates need only be added in quadrature.

In the case where the vertex is incorrectly identified, the error on the opening angle can be substantial and must be taken into account. In this case the uncertainty in the opening angle is in principle a function of the coordinates of the super-clusters and the vertex. However, the z-component of the vertex position has a dominant uncertainty under the assumption that the vertex could be anywhere in the beamspot, which has a gaussian width of $\sigma_{\mathrm{BS}} = 5.8$ cm for 2011 data and $\sigma_{\mathrm{BS}} = 5.0$ cm for 2012 data. The uncertainty in z_{Vtx}, ($\sigma_{z,\mathrm{Vtx}}$), can be modeled as with a gaussian with width of $\sqrt{2} * \sigma_{BS}$. For simplicity this width is used for



(a) Signal p_T in EB

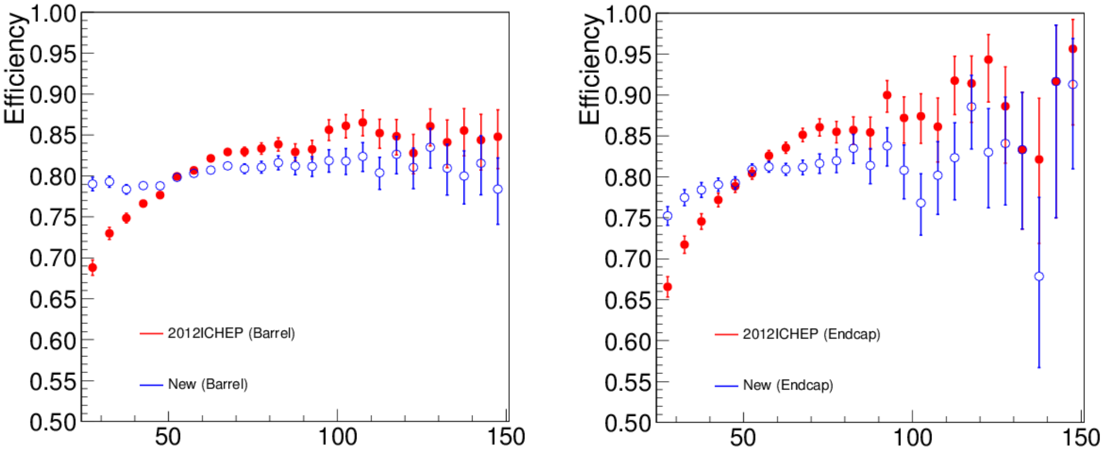
(b) Signal p_T in EE



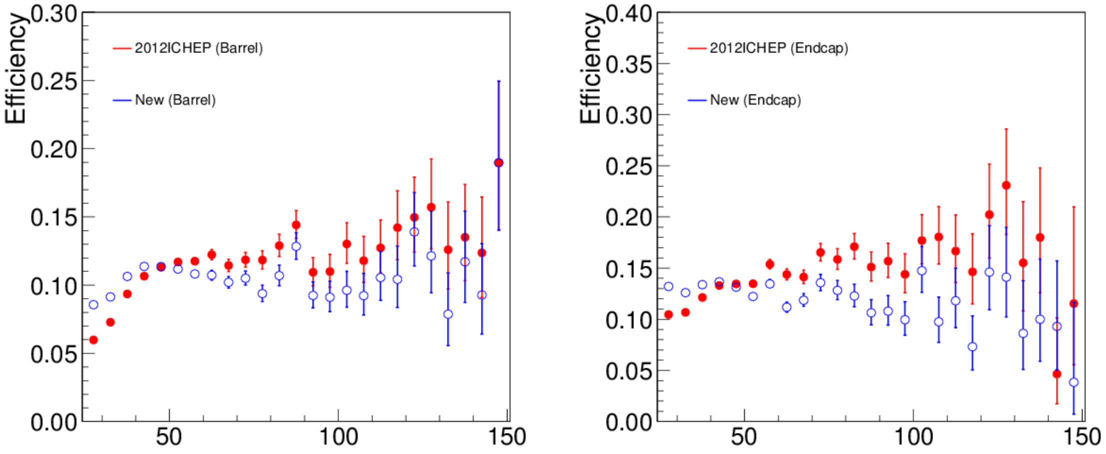
(c) Background p_T in EB

(d) Background p_T in EE

Figure 6.3: 7 TeV photon identification BDT efficiency versus p_T. Plotted in (a) and (b) are the efficiencies of the current ID BDT (in blue) versus the previous version of the BDT (in Red) for H $\rightarrow \gamma\gamma$ events at 7 TeV and $m_{\gamma\gamma} = 121$ GeV as a function of photon p_T. In (c) and (d) are the efficiencies of the testing background samples (jets from γ+jet samples). The working point selected for this comparison is 80% overall signal efficiency.



(a) Signal p_T in EB (b) Signal p_T in EE



(c) Background p_T in EB (d) Background p_T in EE

Figure 6.4: 8 TeV photon identification BDT efficiency versus p_T. Plotted in (a) and (b) are the efficiencies of the current ID BDT (in blue) versus the previous version of the BDT (in Red) for H $\rightarrow \gamma\gamma$ events at 8 TeV and $m_{\gamma\gamma} = 124$ GeV as a function of photon p_T. In (c) and (d) are the efficiencies of the testing background samples (jets from γ+jet samples). The working point selected for this comparison is 80% overall signal efficiency.

$\sigma_{z,\mathrm{Vtx}}$. Given that the uncertainty in the opening angle is principally from this, Equation 5.5.0.1 simplifies to:

$$\begin{aligned}\left(\frac{\sigma_{m_{\gamma\gamma}\text{Wrong Vtx}}}{m_{\gamma\gamma}}\right)^2 &= \frac{1}{4}\left(\frac{\sigma_{E_1}}{E_1}\right)^2 + \frac{1}{4}\left(\frac{\sigma_{E_2}}{E_2}\right)^2 + \left(\frac{1}{m_{\gamma\gamma}}\frac{\partial m_{\gamma\gamma}}{\partial z_{\mathrm{Vtx}}}\sigma_{z\mathrm{Vtx}}\right)^2 \\ &= \left(\frac{\sigma_{m_{\gamma\gamma}\text{Right Vtx}}}{m_{\gamma\gamma}}\right)^2 + \left(\frac{1}{m_{\gamma\gamma}}\frac{\partial m_{\gamma\gamma}}{\partial z_{\mathrm{Vtx}}}\sigma_{z,\mathrm{Vtx}}\right)^2,\end{aligned} \tag{6.3.2.2}$$

6.3.3 Di-photon BDT

The di-photon BDT is designed to transform photon kinematics, mass resolution and photon ID into a single variable which can be used to categorize di-photon events. In order to achieve this goal the variables listed in Table 6.5 are included in the training.

Table 6.5: Di-photon BDT inputs summary. Variables included in the di-photon BDT training. Where variables are normalized by $m_{\gamma\gamma}$, it is for the purpose of avoiding mass-dependence.

Photon ID	BDT_{ID,γ_1} BDT_{ID,γ_2}	Section 6.3.1
Resolution Estimates	$\sigma_{m_{\gamma\gamma}\text{Right Vtx}}/m_{\gamma\gamma}$ $\sigma_{m_{\gamma\gamma}\text{Wrong Vtx}}/m_{\gamma\gamma}$	Section 6.3.2
	$\mathrm{Prob}_{\mathrm{Vtx}}$	Section 5.6.3
Kinematics	$p_{T,\gamma_1}/m_{\gamma\gamma}$ $p_{T,\gamma_2}/m_{\gamma\gamma}$ η_{γ_1} η_{γ_2} $\cos(\phi_{\gamma_1}-\phi_{\gamma_2})$	

The di-photon BDT is trained using the TMVA package [71] with the following samples. For signal MC, all four production mechanisms with $\mathrm{M_H}$ = 123 GeV for 8TeV (and at $\mathrm{M_H}$ = 121 GeV for 7 TeV's separate training). All even events from MC samples listed in Table 5.2 for 8 TeV (Table 5.3 for 7 TeV) are used as background in the training. Odd events are left for preliminary over-training testing and for event class boundary optimization (see Section 6.3.6). The photon pre-selection, described in Section 6.1 is applied to both photons including the electron veto and the $p_T/m_{\gamma\gamma}$ cuts. If there is more than one pair of photons that pass this selection, then the pair with the greatest value of $p_{T,1} + p_{T,2}$ is selected. In addition, there is a cut on the photon ID of $BDT > -0.2$. This cut removes almost exclusively background as can be verified from Figure 6.3 and Figure 6.4.

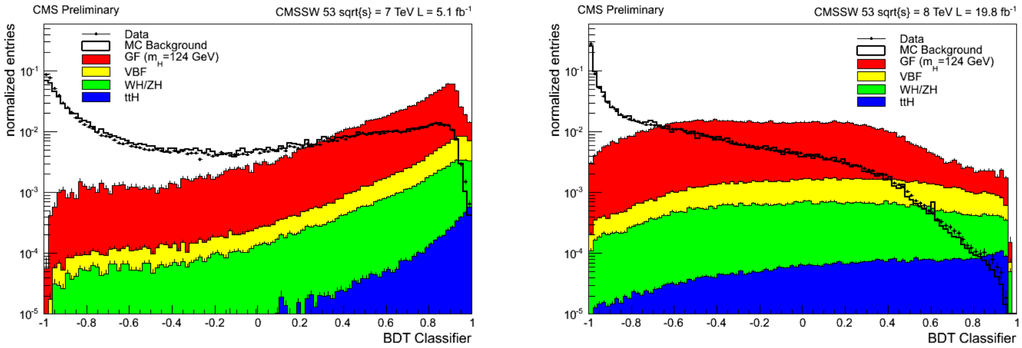
Events are weighted according to the cross-section times luminosity weight. In addition, to provide the BDT with the information that higher resolution events yield higher S/B separation, it is necessary to weight signal events inversely to their mass resolution. To account for the probability for correctly assigning the vertex, the following additional weighting is added:

$$w_{sig} = \frac{\text{Prob}_{\text{Vtx}}}{\sigma_{m_{\gamma\gamma}\text{Right Vtx}}/m_{\gamma\gamma}} + \frac{1 - \text{Prob}_{\text{Vtx}}}{\sigma_{m_{\gamma\gamma}\text{Wrong Vtx}}/m_{\gamma\gamma}} \tag{6.3.3.1}$$

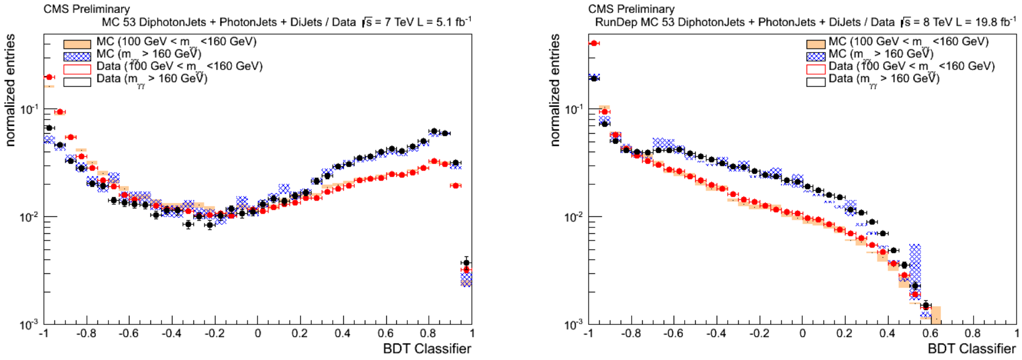
Figure 6.5 shows the BDT output shape for data, background MC, and signal MC. The same strategies are employed for the two trainings. However, the shapes of the BDTs are different because of differences in the relative weighting of the various MC backgrounds. In the 8 TeV training the weight of the QCD events had to be reduced by a factor of 15 in order to avoid very large differences in event weights among samples. The large difference in Sherpa di-photon and QCD di-jet weights is the main reason for this reduction in the 8 TeV QCD weight. In addition, the primary cause of the large difference in the somewhat arbitrary shape of the 7 TeV and 8 TeV BDT outputs is the large difference between ratio of the QCD and di-photon samples' event weights in the 7 TeV and 8 TeV MC. However, when the outputs of the two di-photon BDTs are transformed such that signal is flat as a function of the output as in Figure 6.6 very similar features can be seen.

6.3.4 Kinematic Di-jet BDT

A kinematic di-jet BDT has been developed for the purpose of discriminating VBF-like di-jet events from background using primarily di-jet kinematics. This di-jet BDT is a natural extension of the cut-based analysis reported in Section 6.2.3 as the inputs for this BDT are essentially the same variables used there with the important addition of $p_T^{\gamma\gamma}/m_{\gamma\gamma}$. For clarity's sake the inputs are: $p_T^{\gamma_1}/m_{\gamma\gamma}$, $p_T^{\gamma_2}/m_{\gamma\gamma}$, $p_T^{\text{J}_1}$, $p_T^{\text{J}_2}$, $|\Delta\eta_{JJ}|$, $|Z|$, M_{JJ}, $|\Delta\phi(JJ, \gamma\gamma)|$, and $p_T^{\gamma\gamma}/m_{\gamma\gamma}$.

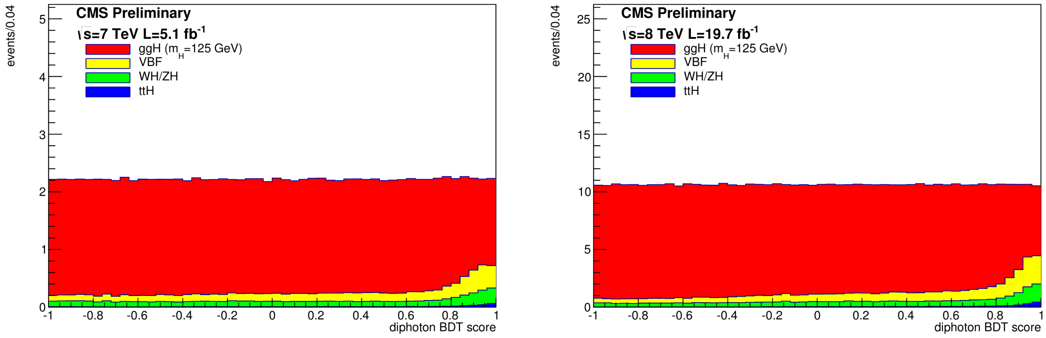


(a) 7 TeV (b) 8 TeV

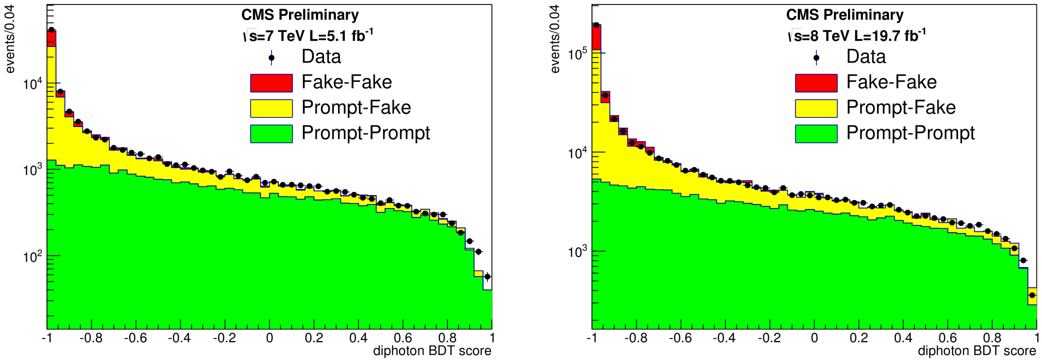


(c) 7 TeV (d) 8 TeV

Figure 6.5: Di-photon BDT output. In (a) the 7 TeV di-photon BDT for signal MC, background MC, and data are shown. All are normalized to unity. The corresponding 8 TeV plot is in (b). (c) and (d) show the data/MC comparisons in the high and low mass regions for the BDT output.



(a) 7 TeV - Signal MC (b) 8 TeV - Signal MC



(c) 7 TeV - Background MC & Data (d) 8 TeV - Background MC & Data

Figure 6.6: Flat signal transformed di-photon BDT output. In these plots the di-photon BDT output has been transformed such that the signal MC is flat as a function of the output after full pre-selection. The behavior is similar for 7 TeV (left) and 8 TeV (right) trainings for signal MC (above) and background MC with data (below). Similar features can be seen in signal and background MC.

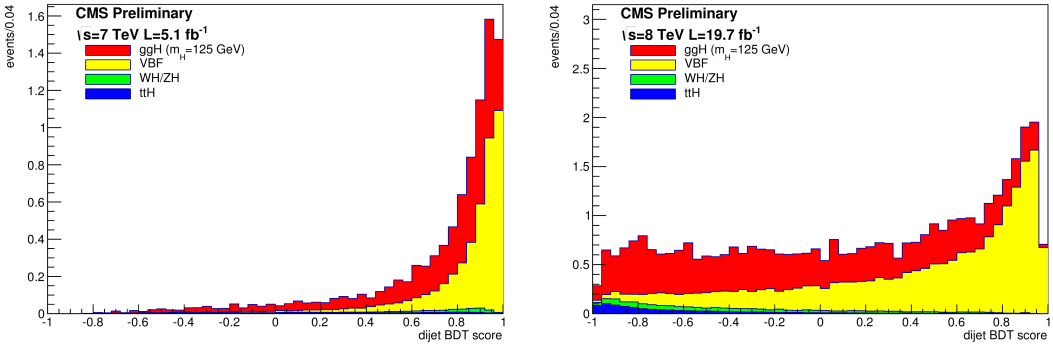
Training this BDT is a challenge because the full pre-selection leaves only a few high weight QCD events to represent most of the remaining background. The solution is to relax most of the kinematic pre-selection cuts for the training. In addition, not cutting on the photon ID BDTs provides more events for training. Indeed since the di-jet BDT inputs are fully kinematic, the quality of the photons is not highly relevant to the training. VBF signal MC with $m_{\gamma\gamma} = 123$ GeV is used for both 7 TeV and 8 TeV trainings where as $m_{\gamma\gamma} = 125$ GeV samples are used for testing. All background MC samples are used in the training but only even events so that the odd events can be used for testing and for categorization optimization.

In addition, since there is a significant yield of di-photon+di-jet signal from Gluon Fusion, Gluon Fusion produced H $\rightarrow \gamma\gamma$ at corresponding masses are included with the backgrounds in the training. The weight of the Gluon Fusion is artificially increased up to the highest level possible without impacting background efficiency. This scaling weight is experimentally determined to be 200 times the cross section times branching ratio weight of the Gluon Fusion samples. Doing the training in this fashion reduces the Gluon Fusion selection by 5-10% and does not impact the background rejection as measured with MC or data.

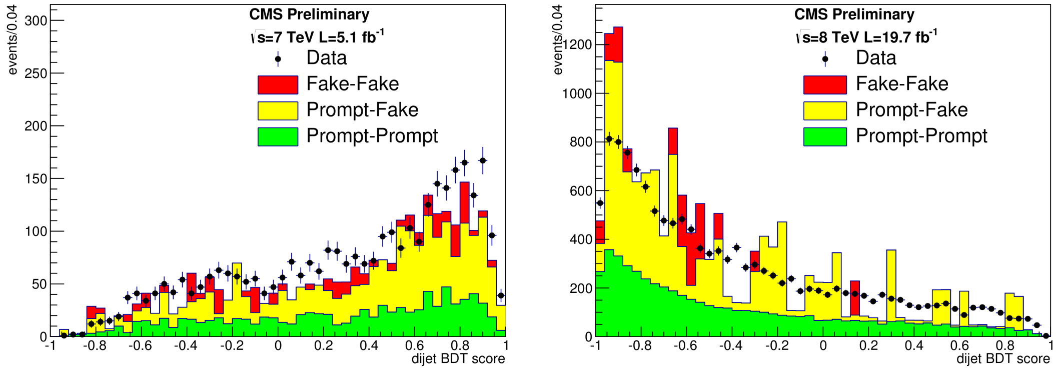
For testing purposes the full pre-selection of photons and jets is required in MC. In Figure 6.7, the shapes of kinematic di-jet BDT are shown for signal MC, background MC, and data. Again the shapes of the two are considerably different because of the different ratios of di-photon background weights to QCD background weights in the training. However, Figure 6.8 shows the kinematic di-jet output transformed such that the VBF signal is flat, and the behavior for signal MC, background MC and data is similar between the 7 TeV and 8 TeV BDTs.

6.3.5 Combined BDT

The question of how to optimally cut on photons and jets simultaneously is complicated and has a variety of solutions. The solution presented has been chosen because it was more optimal than other solutions. In particular, rectangular categorization in the di-photon BDT versus di-jet BDT 2-dimensional plane was explicitly tested and found to be slightly less sensitive. In addition, this method is

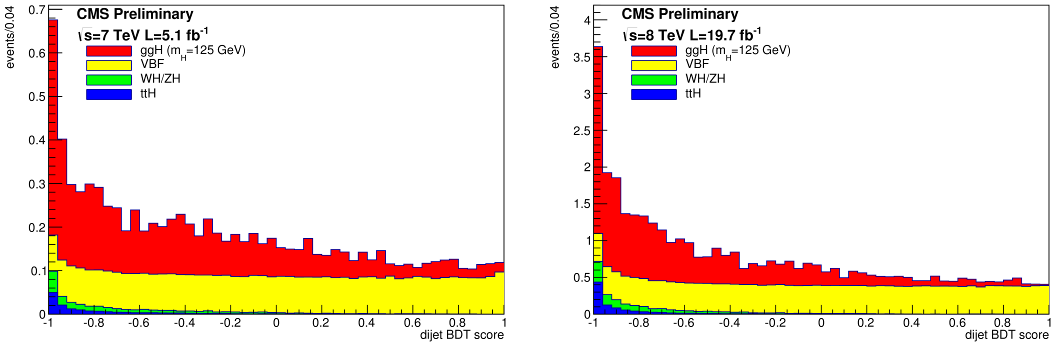


(a) 7 TeV - Signal MC (b) 8 TeV - Signal MC

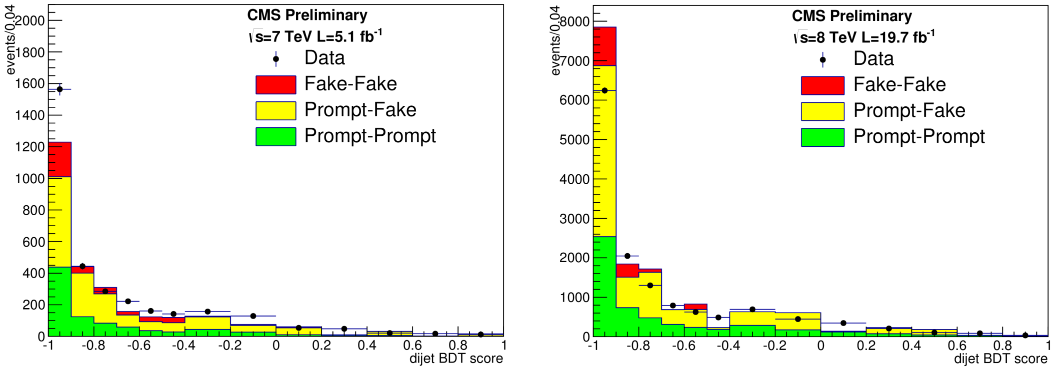


(c) 7 TeV - Background MC & Data (d) 8 TeV - Background MC & Data

Figure 6.7: Kinematic di-jet BDT output is shown when the full pre-selection is applied to signal MC (above) and background MC with data (below) for the 7 TeV (left) and 8 TeV (right) trainings.



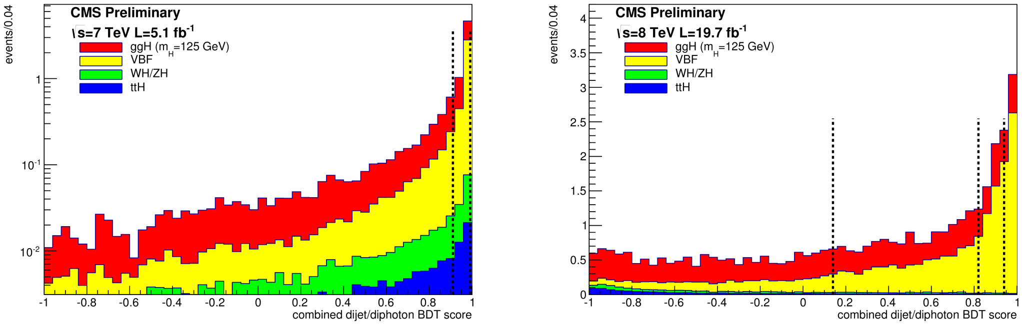
(a) 7 TeV - Signal MC (b) 8 TeV - Signal MC



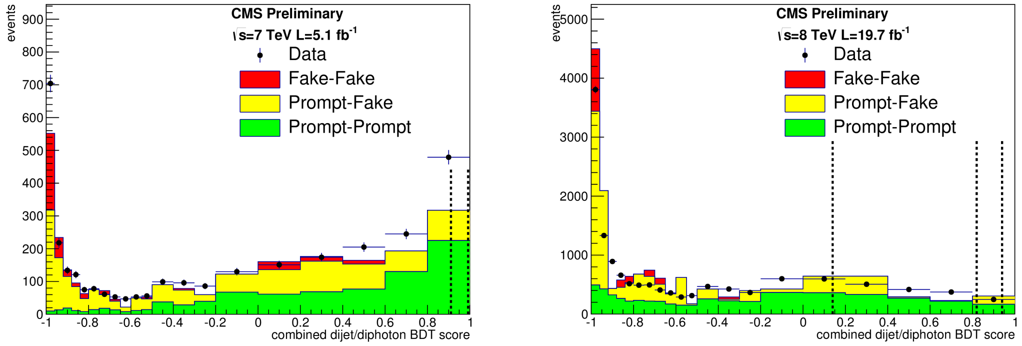
(c) 7 TeV - Background MC & Data (d) 8 TeV - Background MC & Data

Figure 6.8: Flat VBF signal transformed kinematic di-jet BDT output. In these plots the kinematic di-jet BDT output has been transformed such that the VBF signal is flat as a function of the output after full pre-selection of photons and jets. The behavior is similar for 7 TeV (left) and 8 TeV (right) trainings for signal MC (above) and background MC with data (below).

careful to not explicitly correlate the photon ID BDT or mass resolution estimates with di-jet kinematics because all these variables require large systematics and the correlations maybe not be trustworthy. To this end, an addition combined BDT is trained with only three inputs: the di-photon BDT, the di-jet BDT and $p_T^{\gamma\gamma}/m_{\gamma\gamma}$. The output of the combined BDT is shown in Figure 6.9 for signal MC and background MC with data. Figure 6.10 depicts the very visibly lower background efficiency of the combined BDT with respect to the Di-jet BDT alone.

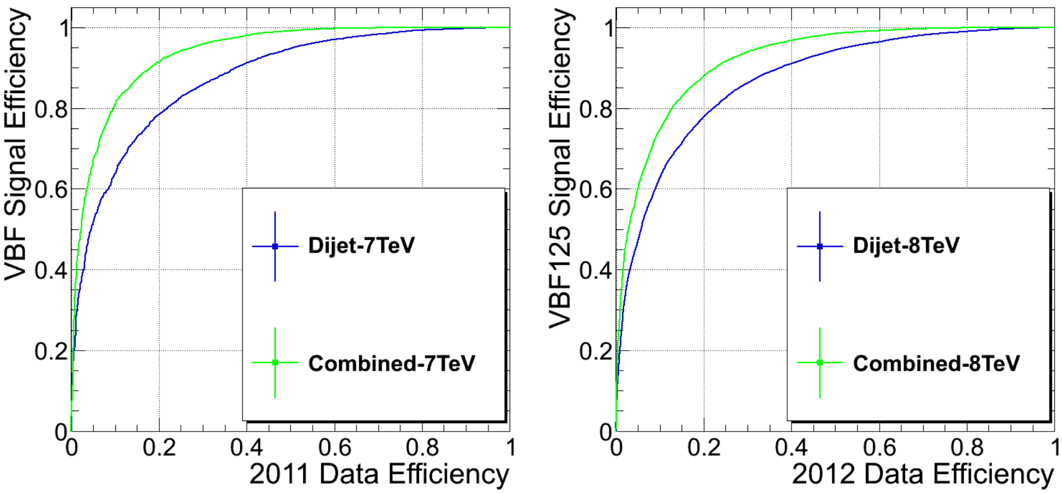


(a) 7 TeV - Signal MC (b) 8 TeV - Signal MC



(c) 7 TeV - Background MC & Data (d) 8 TeV - Background MC & Data

Figure 6.9: Combined di-photon/di-jet BDT output is shown when the full preselection is applied to signal MC (above) and background MC with data (below) for the 7 TeV (left) and 8 TeV (right) trainings. The dashed lines are the category boundaries described in Section 6.3.6.



(a) 7 TeV (b) 8 TeV

Figure 6.10: Signal efficiency versus background efficiency for kinematic di-jet BDT and combined BDT measured by data after full pre-selection for 7 TeV training (a) and 8 TeV training (b).

6.3.6 Di-jet and Inclusive Event Classification

The high score region of the di-photon BDT is generally characterized with high $p_T^{\gamma\gamma}/m_{\gamma\gamma}$ because these events have great expected resolution and narrow shower-shapes. Given that Higgs' production via all production mechanisms except Gluon Fusion have additional final state objects, against which they recoil, it is natural that the di-photon BDT score of non-Gluon Fusion events are disproportionately high as can be seen in Figure 6.6a and in Figure 6.6b. For this reason most VBF signal events will be in the very important high score inclusive categories when not utilized by the di-jet tag. Thus, it is very important to optimize the boundaries of the combined BDT for di-jet classes and di-photon BDT for inclusive classes simultaneously so that VBF events can be optimally used.

Events passing di-jet boundaries are placed in di-jet classes and events failing the loosest di-jet class boundary are used in inclusive categories. The optimization is performed on the error of $\mu_{\mathrm{VBF+VH}}$ ($\delta\mu_{\mathrm{VBF+VH}}$) for di-jet classes and the error on the overall μ ($\delta\mu$) for the inclusive classes. For the 8 TeV analysis three di-jet classes and five inclusive classes were found to be optimal. Given the lesser amount of data in the 7 TeV dataset two di-jet classes and four inclusive classes are optimized. The boundaries typically separate groups of events with large differences in S/B. Adding additional boundaries improves the sensitivity of the analysis by less than 1%.

The optimization is performed on MC because optimization on data may select boundaries at under-fluctuations in data or consider signal in the data as background. Either scenario would bias the final statistical analysis. Only odd background MC events are used for the boundary optimization because the even events were used in the BDT trainings. The signal MC used for the optimization is $m_{\gamma\gamma} = 124$ GeV for 8 TeV and $m_{\gamma\gamma} = 121$ GeV for 7 TeV. In addition, high weight events are smeared using an adaptive bandwidth Gaussian kernel estimator for the QCD background. Otherwise the optimization procedure would act to simply isolate these events and fail to create sensible boundaries.

Table 6.6: Event class boundaries for inclusive and combined BDTs. These are the optimized event class boundaries for the inclusive and di-jet classes for both 7 TeV and 8 TeV datasets.

7 TeV Inclusive Classes	0	1	2	3
Diphoton-MVA Boundary	0.93	0.85	0.70	0.19

7 TeV Dijet Classes	0	1
Combined-MVA Boundary	0.992	0.911

8 TeV Inclusive Classes	0	1	2	3	4
Diphoton-MVA Boundary	0.76	0.36	0.00	-0.42	-0.78

8 TeV Dijet Classes	0	1	2
Combined-MVA Boundary	0.94	0.82	0.14

The procedure is to cyclically iterate over each boundary until the boundaries converge. For most values of the di-photon and di-jet BDTs, steps of 0.01 are used, but in cases where the purity of events changes more rapidly steps of 0.001 are considered. In each iteration a power-law fit of the MC background is performed, and a signal model histogram is constructed using signal MC. $\delta\mu$ is computed using the profile likelihood uncertainty computed on a signal+background representative pseudo-data simulation. The background yield and power law parameter are profiled as well. In order to compute $\delta\mu_{\mathrm{VBF+VH}}$ the signal strength of the additional production mechanisms, $\delta\mu_{\mathrm{GF+t\bar{t}H}}$, is profiled.

The 7 TeV background MC is significantly more discrepant compared to data after di-jet selection than the 8 TeV MC. In addition, there are many fewer MC events. Therefore, the 7 TeV di-jet boundaries are taken from matching the efficiency times acceptance for the tighter two 8 TeV di-jet categories.

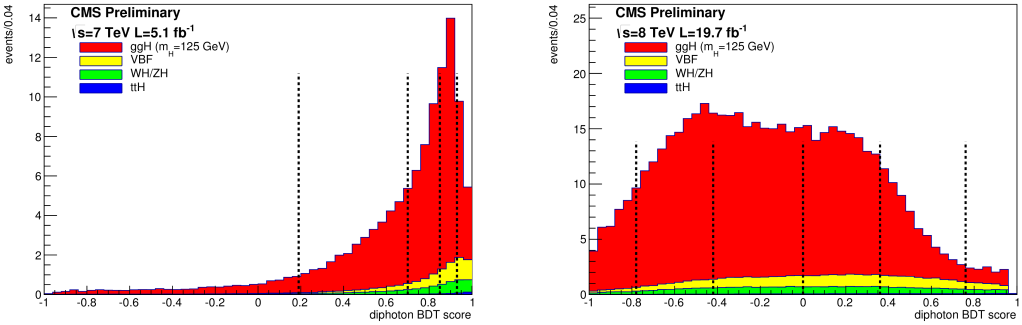
The final boundaries are listed in Table 6.6 and are graphically shown in Figure 6.9 for the di-jet classes in combined BDT and in Figure 6.11.

6.4 Additional Tags

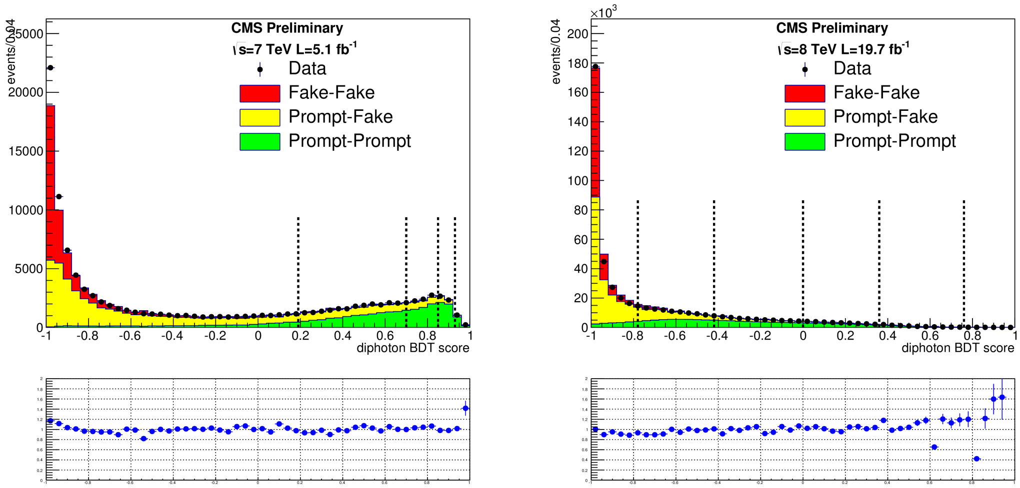
In addition to the di-jet event classes for VBF production and inclusive event classes, there are six (five) additional event classes in the 8 TeV (7 TeV) analysis. Four of these are intended to select the decay products of vector bosons which produce Higgs Bosons via Higgsstrahlung (see Figure 4.2c). Three of these four VH tags are for leptonic decays of W/Z bosons. One of these tags is straightforwardly for large E_T^{Missing}, while the other two require the presences of at least one muon or electron. Further classification is performed on the presence of a second same-flavor lepton and the amount of measured E_T^{Missing}. The fourth VH event class targets the hadronic decay of the vector boson where two jets are required.

The other two event classes are intended to select the decay products of $\mathrm{t\bar{t}}$ produced in association with $t\bar{t}$ Higgs production (see Figure 4.2d). These two event classes are combined into a single event class in the 7 TeV analysis because one would have too few events for proper analysis. There are corresponding event classes in the Cut-Based and MultiVariate Analyses.

In all of these tags more asymmetric cuts on the photons are preferable because the signal di-photons are more boosted. Higher thresholds on the leading p_T photon significantly reduce background whereas lower thresholds on the trailing p_T photon increase acceptance. In the cut-based analysis the lower p_T cut is reduced to 25 GeV for all of these tags, whereas the threshold in the MultiVariate



(a) 7 TeV - Signal MC (b) 8 TeV - Signal MC



(c) 7 TeV - Background MC & Data (d) 8 TeV - Background MC & Data

Figure 6.11: Final di-photon BDT boundaries overlaying di-photon BDT. The dashed lines are the category boundaries of the di-photon BDT optimized simultaneously with the combined BDT for di-jet classes.

Analysis must remain $p_T/m_{\gamma\gamma} > 1/4$ to match the preselection of the di-photon BDT training.

Table 6.7: Optimized di-photon selection for additional tags in event class boundaries and leading $p_T/m_{\gamma\gamma}$ thresholds for the VH and t$\bar{\text{t}}$H tags.

Tag	$p_T^{\text{Lead}}/m_{\gamma\gamma}$	7 TeV Di-photon BDT	8 TeV Di-photon BDT
VH Leptonic Tight	45/120	0.1	-0.6
VH Leptonic Loose	45/120	0.1	-0.6
VH E_T^{Missing}	45/120	0.8	0.0
VH Hadronic	60/120	0.6	0.2
t$\bar{\text{t}}$H Leptonic	60/120	0.6	-0.2
t$\bar{\text{t}}$H Hadronic	60/120	0.6	-0.6

These tags have generally been optimized on top of the inclusive cut-based analysis. When the tags are moved into the MultiVariate Analysis, the strategy of photon selection is changed to a cut on the di-photon BDT. The di-photon BDT cut is optimized separately for each category. That optimization must be done using control samples in data for background because even with loose preselection of these tags there are very few MC events. Moreover, optimizing with the data using the signal selection could bias the results. Therefore, a control sample is defined for each and translated into an estimate of background with the signal selection. With these background estimates and signal yields with full selection, $\delta\mu_{\text{VBF+VH}}$ is estimated as a function of di-photon BDT. The minimum value for each tag is selected as the threshold.

This procedure is performed in the 8 TeV analysis only. The 7 TeV di-photon BDT thresholds are taken by matching signal efficiency times acceptance with those thresholds derived for the 8 TeV analysis. Table 6.7 summarizes the $p_T/m_{\gamma\gamma}$ thresholds and di- photon BDT thresholds for the 7 TeV and 8 TeV analyses.

6.4.1 VH Leptonic Tight Tag

This tag requires the presences of an electron or a muon with $p_T > 20$ GeV passing full selection criteria as described in Section 5.8.4, and corrected $E_T^{\text{Missing}} >$

45 GeV which is indicative of the leptonic decay of a W boson. Section 6.4.3 describes the E_T^{Missing} corrections. The presences of two same-flavor leptons with $p_T > 10$ GeV with invariant mass between 70 and 110 GeV, which is the decay signature of a Z boson, is an alternative selection for this event class.

Unlike other event classes leptons are selected before photons. If a qualifying lepton (or two) is found, only photons beyond $\Delta R > 1.0$ of the lepton ($\Delta R > 0.5$ when there are two leptons) are available for selection. The purpose of this selection algorithm is to avoid selecting photons which are final state radiation from the lepton. In the case where the lepton is an electron this algorithm aids in avoiding selecting the photons which are mis-identified electrons. Another safeguard against selecting photons which are electrons themselves or associated Bremsstrahlung is to reject photons which are within $\Delta R > 1.0$ of any electron's track as long as there are no missing hits.

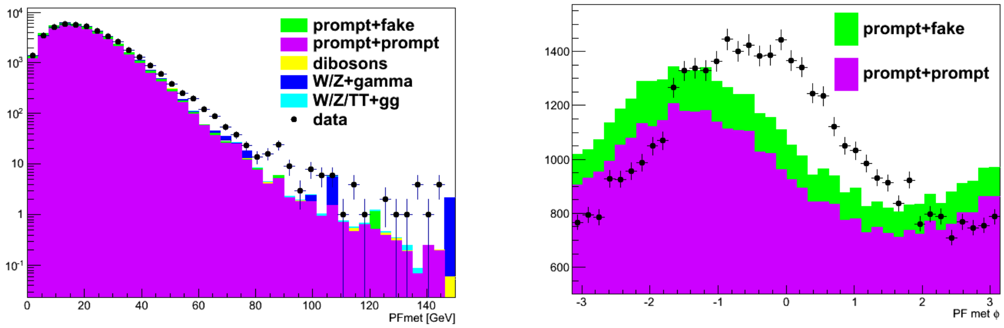
In order to reduce the contribution of $\mathrm{t\bar{t}H}$ events jets within $|\eta| < 2.4$, sufficiently far ($\Delta R > 0.5$) from both photons and lepton(s) and with $p_T > 20$ GeV are counted. If there are more than three such jets the event is vetoed. Removing $\mathrm{t\bar{t}H}$ events improves the expected measurement of $\mu_{\text{VBF+VH}}$.

6.4.2 VH Leptonic Loose Tag

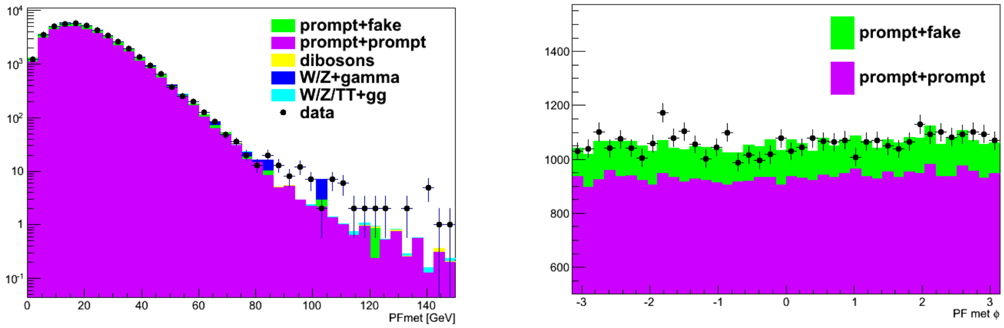
This tag is the complement of the single lepton selection in VH Leptonic Tight Tag described in Section 6.4.1. The primary difference is that $E_T^{\text{Missing}} < 45$ GeV. Additionally, some lepton-photon cuts are tightened. In particular, the ΔR requirement between the lepton and the photon is tightened ($\Delta R > 1.0$), and for electrons an additional requirement is added on the invariant mass of the electron with either photon to further reject $\mathrm{Z} \rightarrow \mathrm{e^+e^-}$ events with mis-identified electrons. That is, $|\mathrm{M}_{e\gamma} - \mathrm{M_Z}| > 10$ GeV is demanded of the electron with either photon.

6.4.3 VH E_T^{Missing} Tag

Neutrinos cannot be detected inside CMS. However, their presence can be inferred by the lack of of net-zero balance in transverse momentum. Due to unsimulated noise, mis-alignment between sub-detectors and inaccuracies in jet energy reconstruction the E_T^{Missing} resolution in data is significantly worse than in MC. This is corrected in MC by smearing the jet energies and then re-computing the E_T^{Missing}. In addition the ϕ-direction of the E_T^{Missing} is not flat in ϕ and must be corrected for data and MC separately. The corrections are derived for x and y components of E_T^{Missing} as a function of total transverse energy for data and MC. In Figure 6.12 E_T^{Missing} and its direction are plotted before and after corrections.



(a) Uncorrected E_T^{Missing} (b) Uncorrected $\phi_{E_T^{\text{Missing}}}$



(c) Uncorrected E_T^{Missing} (d) Corrected $\phi_{E_T^{\text{Missing}}}$

Figure 6.12: E_T^{Missing} corrections. These are the E_T^{Missing} distributions and its direction before (above) and after (below) corrections are applied on 8 TeV data and MC.

After E_T^{Missing} is corrected, analysis level cuts can be applied. In addition to $E_T^{\text{Missing}} > 70$ GeV there are two further topological cuts. Requiring $|\phi_{\gamma\gamma} - \phi_{E_T^{\text{Missing}}}| > 2.1$ essentially requires that the di-photon and the E_T^{Missing} (i.e. the Higgs and the vector boson) are back-to-back. It is a loose cut that does not impact

VH signal efficiency but reduces Higgs events from gluon fusion and background which are flatter (or uncorrelated) in this variable. The increased purity of the VH fraction of the expected signal in this tag decreases the expected $\delta\mu_{\mathrm{VBF+VH}}$ measurements.

Finally, in control samples with high E_T^{Missing} an overall excess in data compared to MC was found. After investigation it was found that in events with a high p_T jet that is back-to-back with the di-photon, there is significant fake E_T^{Missing}. This feature is much larger in data than MC this caused the data/MC discrepancy. Adding the requirement that $|\phi_{\gamma\gamma} - \phi_{\mathrm{Jet}^{\mathrm{Lead}}}| < 2.7$ when the leading jet's $p_T > 50$ GeV largely eliminates this background (both in data and MC) while reducing the signal efficiency by less than 5%.

6.4.4 VH Di-jet Tag

This tag intends to select associated W and Z bosons which decay hadronically to two quark jets. Since hadronic resolution is wider than the difference between W and Z masses ($\sim$ 11 GeV) both hadronic decays are selected together in a wide window of 60-120 GeV. The jets are pre-selected in precisely the same manner as in the di-jet tag for the VBF (see Section 6.1.1).

Three variables thresholds have been optimized at various signal efficiencies using TMVA. The final selection was chosen minimizing $\sigma_{\mu_{\mathrm{VBF+VH}}}$ as a function of working points from TMVA. $p_T^{\gamma\gamma}/m_{\gamma\gamma} > 130/120$ is a tight cut which again takes advantage of the recoil of the Higgs against the vector boson. The presences of two jets with $p_T > 40$ GeV within $|\eta| < 2.4$ is required for further background separation. Finally, the third requirement necessitates the computation of the pre-Higgsstrahlung vector boson composite particle from the selected di-jet and di-photon (i.e. V^* in $V^* \rightarrow V$+H) and then Lorentz boosting into its frame of reference. θ^* is the angle between the direction of V^* in the lab frame and either of the decay products in V^*s rest frame. The distribution of $\cos(\theta^*)$ in signal is very flat. Whereas background events are much more likely to be back-to-back leading to a peaking distribution in $\cos(\theta^*)$ near -1 and 1. The optimized cut is $|\cos(\theta^*)| < 0.5$

6.4.5 $\mathrm{t\bar{t}H}$ Leptonic Tag

Figure 4.2d depicts the Feynman diagram of the $\mathrm{t\bar{t}H}$ production. The decay of the top quarks in the final state have very unique signatures which can be used to produce this and the following high S/B event class. Given the near absolute top quark weak decay to bottom quarks ($t \rightarrow \mathrm{W}^{+}b$) there are two W boson decays and two b decays. This event class seeks out a muon or an electron from the leptonic decay of a W (i.e. $\mathrm{t\bar{t}} \rightarrow b\bar{\ell}\nu\bar{b}\ell\bar{\nu}$ or $\mathrm{t\bar{t}} \rightarrow b\bar{\ell}\nu\bar{b}q\bar{q}'$). The selection of the muon or electron is the same as the tight lepton tag as well as the extra selection on photons against final state radiation and fake electrons (see Section 6.4.1). The lepton p_T must be greater than 20 GeV.

Selected events must have at least two jets with $p_T > 25$ GeV which are well-separated from the two photons ($\Delta R > 0.5$). At least one of those jets must originate from a b quark. To identify jets originating from the hadronization of bottom quarks the Combined Secondary Vertex (CSV) b-tagging algorithm is employed [77]. The algorithm identifies jets from b decays by identifying their displaced decay vertex. The working point selected provides an efficiency for b jets of about 70% and a misidentification probability for jets from light quarks and gluons of about 1%.

In the 7 TeV data there are too few events for proper statistical analysis of this tag. The selection is still used but this event class is merged with the $\mathrm{t\bar{t}H}$ Hadronic Tag.

6.4.6 $\mathrm{t\bar{t}H}$ Hadronic Tag

The target for this tag is events containing two photons from a Higgs' decay as well as fully hadronic decays of W bosons from two t quarks (i.e. $\mathrm{t\bar{t}} \rightarrow bq_1\bar{q}_2\bar{b}q_3\bar{q}_4$). Selected events must have at least five jets with $p_T > 25$ GeV which are well-separated from the two photons ($\Delta R > 0.5$). At least one of those jets must originate from a b quark and no well-identified leptons may be present.

6.4.7 Event Classification

All events are only used once within this analysis. Therefore a priority system must be implemented for the situation where a single event passes more than one of the several event types which are being scanned. The priority is generally that event classes with higher S/B are filled first. Table 6.8 lists the order of the selection of event classes.

Table 6.8: Priority ordering of event classes.

Event Class
$\mathrm{t\bar{t}H}$ Leptonic
VH Leptonic Tight
VH Leptonic Loose
VBF Di-jet
VH E_T^{Missing}
$\mathrm{t\bar{t}H}$ Hadronic
VH Di-jet
Untagged

6.5 Signal Models

6.5.1 Efficiency Corrections to Signal Events

Producing MC which replicates important features of physical objects in data is very important to the H → $\gamma\gamma$ analysis. Great care has been taken to produce signal MC in which the noise in the ECAL and pile-up are modeled as a function of run period. Although these recent improvements provide better agreement between data and MC in isolation, shower-shape and other identification variables than ever before, there are still some discrepancies which cannot be easily corrected from first principles. However, these differences are not catastrophic and their largest potential impact is of a few percent difference in the signal efficiency. Therefore, using independent samples in data and MC various selection efficiencies are compared, and those comparisons are used to correct the efficiency of signal MC.

Copious decays of Z bosons to electrons are the most frequent independent samples which are used to derive efficiency corrections in this analysis. Electrons are treated like photons and the *Tag and Probe* method is used to measure the efficiency of various cuts (e.g. single photon pre-selection cuts in Table 6.1). This is done in data and MC and the ratio of efficiency cuts can be used to correct signal efficiency MC.

In the *Tag and Probe* method one electron (the tag) is selected with tight cuts to ensure that a real electron is selected. The second electron (the probe) is first selected with a p_T cut and a tag+probe invariant mass selection near the Z mass peak is required. This procedure delivers a fairly pure sample of $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events with hardly any selection bias on the probe. After the probe is selected, the full selection is applied to the probe and the passing rate is the measured efficiency. This is done in the same fashion for data and MC.

When the probe fails the full selection there is often a component of those failing events which are fake electrons. That background must be estimated and subtracted since only the real probe electrons are analogous to photons of the signal.

Each of the cut-based selections applied in either the Cut-based Analysis or MultiVariate Analysis needs an efficiency correction. The photon pre-selection (described in Section 6.1) efficiency is measured in $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events without the electron veto in η-R_9 categories. The R_9 boundary matches the pre-selection's at 0.9. The corrections are of the order 1% in every η-R_9 category for both 7 TeV and 8 TeV datasets. The measurements and the corrections are listed in Table 6.9.

The efficiency correction of the electron veto is measured with $\mathrm{Z} \to \mu^+\mu^-\gamma$ events. The tag in this case is two well identified muons with $p_T > 10.5$ GeV and $\mathrm{M}_{\mu^+\mu^-} > 35$ GeV. The probe is a photon passing all pre-selection cuts except the electron veto, within $\Delta R < 0.8$ of one of the muons, and with $p_T > 20$ GeV. In addition, the photon may not be reconstructed as an electron as well. Finally, the three body invariant mass is restricted to $60 < \mathrm{M}_{\mu^+\mu^-\gamma} < 120$ GeV. The efficiencies of applying the electron veto in data and MC and the ratios used as corrections to 8 TeV signal efficiency are reported in Table 6.10. The sample sizes are small in

Table 6.9: Pre-selection photon identification efficiencies measured in the 4 photon categories using tag and probe with Z → e^+e^- events (for all cuts except electron rejection).

	DATA	MC	Data/MC	
	Efficiency	Effciency	Ratio	Error
7 TeV				
EB; R_9 >0.90	0.9872	0.9908	0.996	0.003
EB; R_9 <0.90	0.9619	0.9670	0.995	0.006
EE; R_9 >0.90	0.9906	0.9824	1.008	0.009
EE; R_9 <0.90	0.9606	0.9560	1.005	0.018
8 TeV				
EB; R_9 >0.90	0.9879	0.9864	0.999	0.003
EB; R_9 <0.90	0.9566	0.9610	0.995	0.006
EE; R_9 >0.90	0.9838	0.9789	1.005	0.009
EE; R_9 <0.90	0.9545	0.9445	1.011	0.018

the 7 TeV dataset and the corrections derived are consistent with 1.

Table 6.10: Efficiency of the conversion-safe electron veto, measured in the four photon categories using tag and probe with Z → $\mu^+\mu^-\gamma$ events. The data to MC ratio efficiency corrections are also depicted. The efficiency is built using for the denominator the number of photons passing all cuts except the electron veto, and for the numerator the number of photons passing all cuts including the electron veto.

	DATA	MC	Data/MC	
8 TeV	Efficiency	Efficiency	Ratio	Error
EB; R_9 >0.94	0.9984	0.9991	0.9994	0.0004
EB; R_9 <0.94	0.9867	0.9930	0.9937	0.0014
EE; R_9 >0.94	0.9893	0.9938	0.9955	0.0020
EE; R_9 <0.94	0.9639	0.9738	0.9899	0.0045

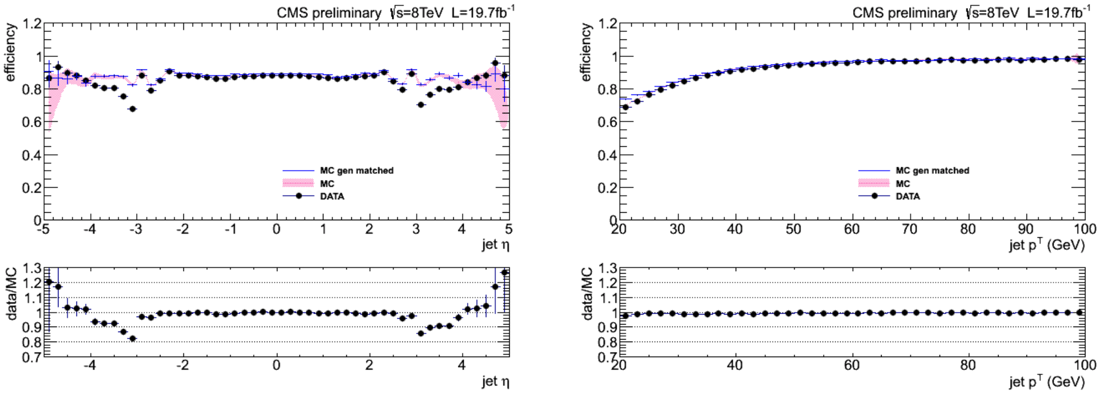
The tag and probe efficiencies of the CiC4PF (described in Section 6.2.1) along with their ratios are reported in Table 6.11. These efficiencies are measured using Z → e^+e^- events and the electron veto is omitted.

In addition for events in the 8 TeV analysis that include jets a conservative efficiency correction is applied as a function of jet η and p_T to correct the signal efficiency for data to MC differences in the pile-up jet ID variables. Figure 6.13 shows the data and MC jet η and p_T distributions for Z → $\mu^+\mu^-$ with 2 jets events.

Table 6.11: Cut-based photon identification efficiencies measured in the 4 photon categories using tag and probe with Z $\rightarrow$ e^+e^- events (for all cuts except electron rejection).

	DATA	MC	Data/MC	
	Efficiency	Effciency	Ratio	Error
7 TeV				
EB; R_9 >0.94	0.9580	0.9591	0.999	0.003
EB; R_9 <0.94	0.8560	0.8498	1.007	0.009
EE; R_9 >0.94	0.9170	0.9092	1.009	0.007
EE; R_9 <0.94	0.6674	0.6611	1.010	0.019
8 TeV				
EB; R_9 >0.94	0.9192	0.9176	1.002	0.003
EB; R_9 <0.94	0.7665	0.7537	1.017	0.008
EE; R_9 >0.94	0.8142	0.7938	1.026	0.006
EE; R_9 <0.94	0.5459	0.5312	1.028	0.013

The ratio of data efficiency to MC efficiency is used as the correction as long as the ratio is less than 1. In effect the MC yield is never scaled up. Since there is no jet ID applied to jets in the 7 TeV dataset, there is no correction in that dataset.



(a) η^{Jet} (b) p_T^{Jet}

Figure 6.13: Pile-up jet preselection efficiency. The data to MC comparison of PU jet identification efficiency above is used to correct the efficiency of the signal MC as a function of η^{Jet} and p_T^{Jet}.

6.5.2 Parametric Models

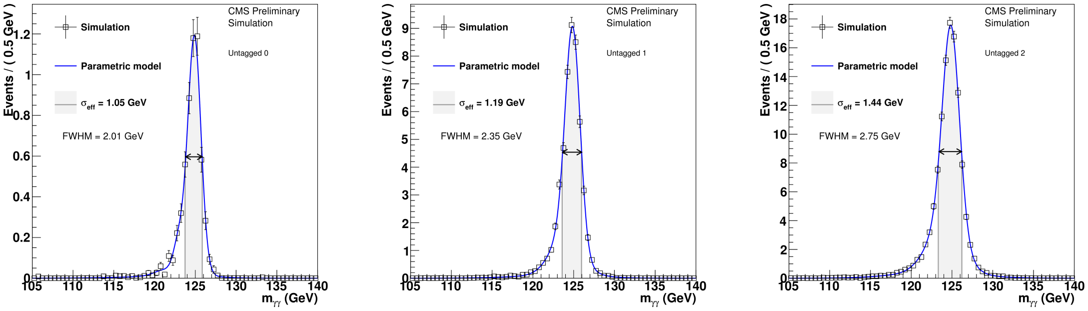
Signal MC is produced for each production mechanism in 5 GeV steps in $\mathrm{M_H}$ from 110 to 150 GeV, which is the search range of the analysis. In order to analyze the data for arbitrary values $\mathrm{M_H}$ in this range, parametric models are produced for each production mechanism in every event class as a function of $\mathrm{M_H}$.

After full selection is applied to each MC sample with all re-weighting and efficiency corrections, the signal models are constructed for $\mathrm{M_H} = 125$ GeV by fitting two to seven gaussians for each production mechanism in each event class. The number of gaussians depends on the resolution of the model and features of the low mass tail. The parameters of the corresponding parameters for all other masses are initialized to those from the $\mathrm{M_H} = 125$ GeV models. They are then re-fit with moderate constraints. The parameters of the model are interpolated between the 5 GeV steps in MC. In this way continous functions of the parameters are produced.

The parametric signal models at $\mathrm{M_H} = 125$ GeV for the 14 events classes at 8 TeV of the MultiVariate Analysis are shown in Figure 6.14, and Figure 6.15. The white box points in these plots are the selected, re-weighted and corrected MC. The smooth blue curve is full model, which is the sum of several gaussians. In each event class the four production mechanisms are summed. In grey under the model is the narrowest region containing 68.3% of the signal model. Half of the width of this area defines σ_{eff} which is one metric for resolution in the event classes. Another metric is the full width at half max (FWHM), which is the horizontal double arrow black line drawn at half the maximum value of the model.

Both σ_{eff} and FWHM are written in each plot but the values are not necessary to see the qualitative difference in expected signal resolution among the event classes. Among the untagged classes and in the di-jet classes the resolution of the models goes from very good to worse as one goes from earlier event class to later. The progression of resolution in the models is based on the resolution dependence in the di-photon BDT and combined BDT.

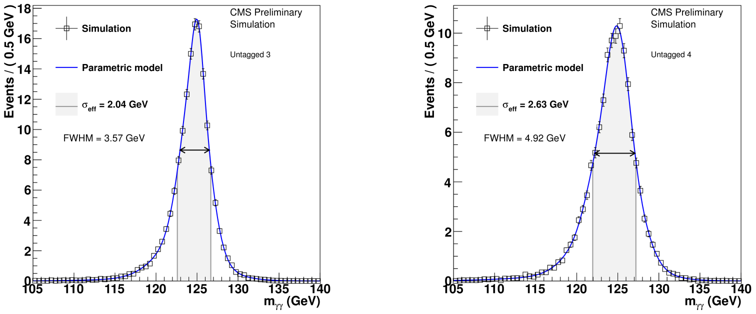
The corresponding 11 MVA event classes signal models at 7 TeV and all 31 signal models for the Cut-based Analysis can be found in Appendix C.



(a) Untagged 0
8 TeV - MVA

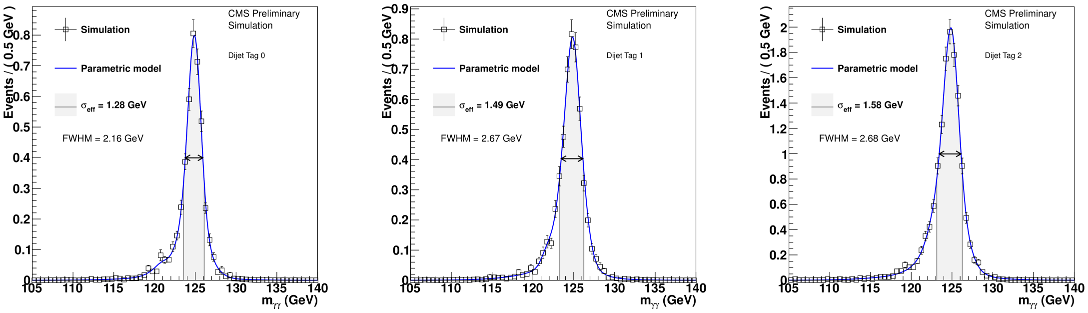
(b) Untagged 1
8 TeV - MVA

(c) Untagged 2
8 TeV - MVA



(d) Untagged 3
8 TeV - MVA

(e) Untagged 4
8 TeV - MVA

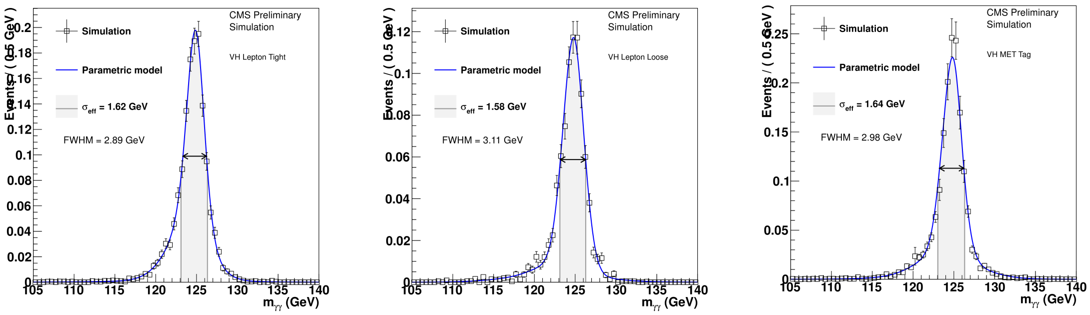


(f) Di-jet 0
8 TeV - MVA

(g) Di-jet 1
8 TeV - MVA

(h) Di-jet 2
8 TeV - MVA

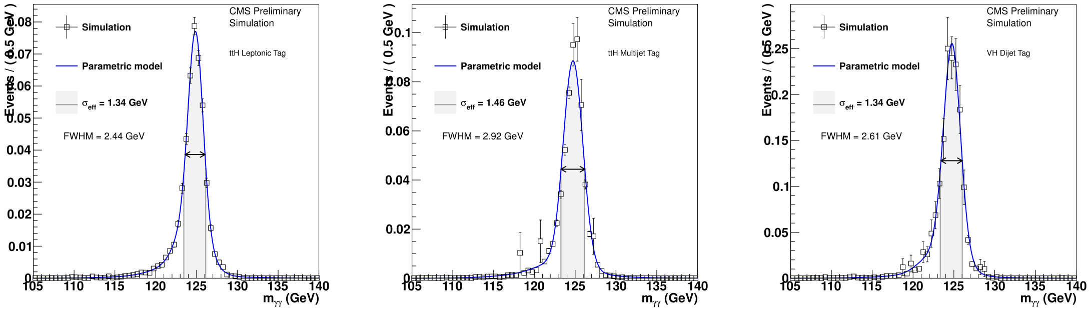
Figure 6.14: MVA 8 TeV inclusive and di-jet event class signal models.



(a) VH Leptonic Tight
8 TeV - MVA

(b) VH Leptonic Loose
8 TeV - MVA

(c) VH E_T^{Missing}
8 TeV - MVA



(d) $\mathrm{t\bar{t}H}$ Leptonic
8 TeV - MVA

(e) $\mathrm{t\bar{t}H}$ Hadronic
8 TeV - MVA

(f) VH Di-jet
8 TeV - MVA

Figure 6.15: MVA 8 TeV additional tag event class signal models.

6.5.3 Signal Event Class Characterization

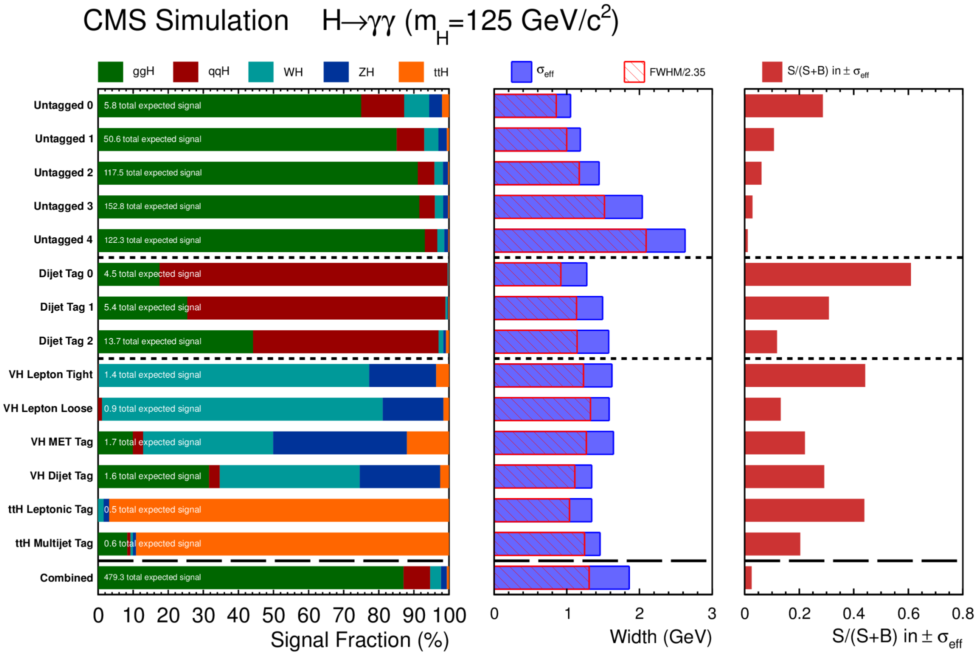
Figure 6.16 exposes many features of the event classes for both the MultiVariate and Cut-based analyses. There are the color-coded signal composition (green is Gluon Fusion, red is VBF, light blue is WH, blue is ZH and orange is $\mathrm{t\bar{t}H}$), signal resolution (σ_{eff} and FWHM/2.35) and the S/(S+B) estimated within $\pm\sigma_{\mathrm{eff}}$ for each class. These figures are derived from the information in Table C.1 and Table C.2.

In the untagged event classes of the MVA there is a smooth progression from greatest to lowest S/(S+B) from Untagged 0 to 4. The S/(S+B) is apparently anti-correlated with signal resolution and the reason for that is two-fold. Firstly, photons from detector regions with better resolution have better S/(S+B) because excellent resolution is correlated with great photon identification and isolation. In addtion, narrower (better) resolution implies less background under the signal peak.

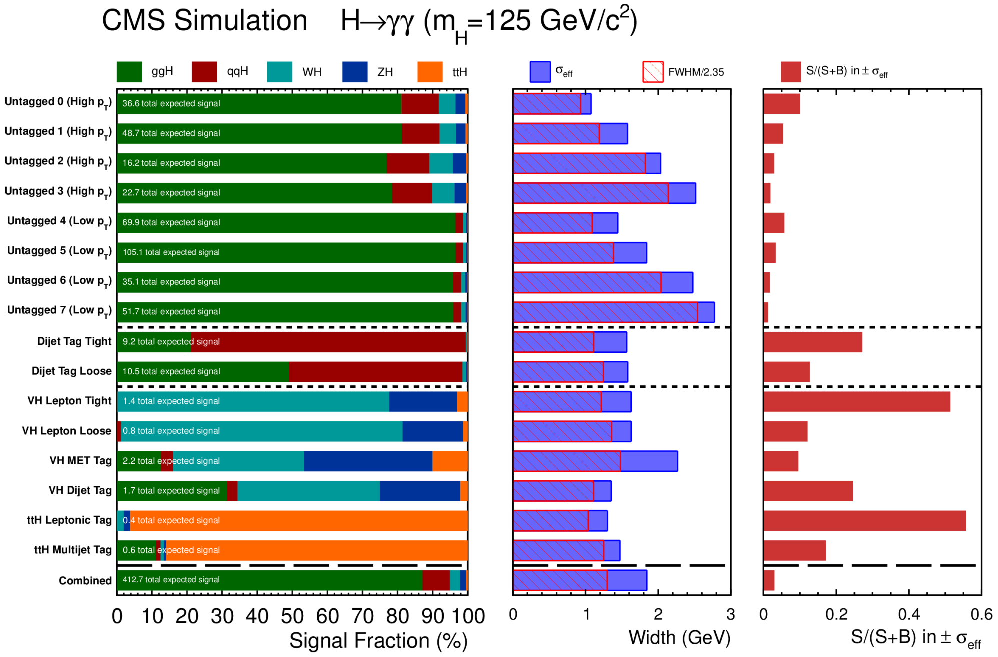
The signal composition in the untagged classes in both analyses is dominantly Gluon Fusion. However, in the MVA there is a higher fraction of events from other production mechanisms in the better event classes. This is an indication of the high degree of correlation between the di-photon BDT and di-photon p_T. Moreover, all production mechanisms besides Gluon Fusion tend to have higher di-photon p_T because the Higgs recoils against other objects. Likewise, in the Cut-based analysis the high $p_T/m_{\gamma\gamma}$ event classes have a high fraction ($\sim 20\%$) of non-Gluon Fusion events, whereas the low $p_T/m_{\gamma\gamma}$ classes are much purer, $\sim 95\%$ Gluon Fusion.

All the tagged event classes have larger S/(S+B) than untagged event classes. The only exception is Untagged 0 in the MVA which has exceptionally high S/(S+B) because of the tightness of selection and extremely good resolution. The composition of these channels matches very well the intent of the tags. For example, $\mathrm{t\bar{t}H}$ tags select mainly $\mathrm{t\bar{t}H}$ signal, and the lepton tags select mainly signal produced in association with a Z or a W boson.

While the di-jet tags and E_T^{Missing} tag also select the intended signal, there is a significant contamination in signal purity from Gluon Fusion. This is most signifi-



(a) 8 TeV - MVA



(b) 8 TeV - Cut-based

Figure 6.16: Graphical representation of 8 TeV event classes' signal composition, signal width and S/(S+B) estimates for the MVA (a) and Cut-based Analysis (b) signal MC with $M_H = 125$ GeV.

cant in the di-jet targeting VBF. Because the di-jet component of the Gluon Fusion MC has larger systematic uncertainties, minimizing this component is preferable. To a great degree this feature is automatic as tighter VBF-like di-jet selection reduces the fraction of Gluon Fusion selected, but in addition Gluon Fusion MC is used as a background in the training of the kinematic di-jet BDT.

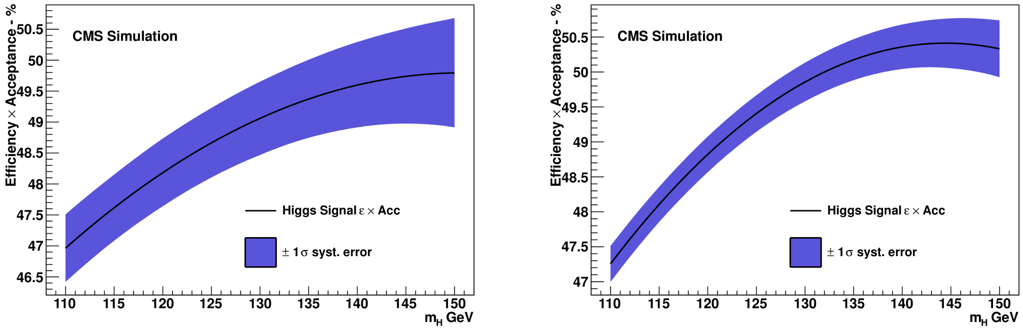
The mass resolution indicators show that the di-jet tag in the Cut-based Analysis is much less sensitive to the mass resolution than the MVA. That is, among the three MVA di-jet for VBF event classes the tightest class has noticeably better resolution. The resolution sensitivity comes from the combined BDT's di-photon BDT input. In addition, the cut-based di-jet tag allows photons to come from EB or EE and does not classify based on R_9, which is where the resolution sensitivity of the Cut-based Analysis originates.

The efficiency × acceptance varies as a function of $\mathrm{M_H}$ for both the Cut-based and MultiVariate Analyses as Figure 6.17 shows. The bands are produced by estimating the variation in signal efficiency due to systematic errors (both above and below the nominal value) described in Section 6.7. At $\mathrm{M_H} = 125$ GeV the MVA is 48-49% efficient. The Cut-based Analysis is ∼42% efficient in the 8 TeV analysis and ∼46% in the 7 TeV analysis. The difference in efficiency between the two datasets is due to the effect of greater pile-up in 2012. The same photon identification cuts are applied in both analyses.

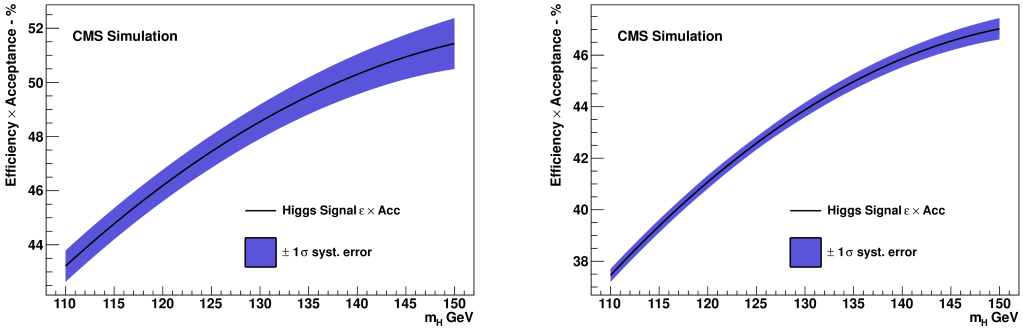
Overall Figure 6.16 shows that these event classes (and the analysis in general) perform as expected. In particular, the signal resolution progression in the untagged classes, the signal composition of each event class and the high S/(S+B) in the tagged classes validate that the analysis' optimization goals have been achieved.

6.6 Background Estimation

The background estimate used in the two analyses describe is with a fit to the data in the range of $100 < m_{\gamma\gamma} < 180$ GeV in each event class. Given that even in the highest S/B event class the selection in this wide range is dominated by background, no adjustments need to be made for the case where signal is present



(a) 7 TeV - MultiVariate Analysis (b) 8 TeV - MultiVariate Analysis



(c) 7 TeV - Cut-based (d) 8 TeV - Cut-based

Figure 6.17: Signal efficiency as a function of M_H in the MultiVariate Analysis for 7 TeV (a) and 8 TeV (b). The corresponding Cut-based Analysis plots are (c) and (d).

in the data.

The parametrized shape of the background invariant mass in each event class is not known a priori. Therefore, a strategy that minimizes potential bias estimations from imperfect background modeling is used. In the discovery paper [29] the strategy was to fit the data with polynomials with orders high enough (typically 3-5) that uncertainty from the fit is at least five times larger than the average estimated bias. Potential bias is estimated from polynomial fits to pseudo-data generated from various different "truth functions". The following functions are those truth models:

- Polynomials with Bernstein basis
  - $N\mathrm{Ber}\,(x) = \sum\limits_{i=1}^{i=N} \beta_i b_{(i,N)}$ where $b_{(i,N)} = \begin{pmatrix} N \\ i \end{pmatrix} x^i (1-x)^{N-i}$
- Power-law series
  - $N\mathrm{Pow}\,(x) = \sum\limits_{i=1}^{i=N} \beta_i x^{-\alpha_i}$
- Exponential series
  - $N\mathrm{Exp}\,(x) = \sum\limits_{i=1}^{i=N} \beta_i e^{\alpha_i x}$
- Laurent series
  - $N\mathrm{Lau}\,(x) = \sum\limits_{i=1}^{i=N} \beta_i x^{-4+\sum\limits_{j=1}^{i}(-1)^j (j-1)}$

For each event class a truth function of each type is fit to the data of that event class. This fit is performed without observing the data in the search range to avoid the possibility of selection bias. The order of the selected truth functions for each family of functions is selected by increasing the order until the $(N+1)$ order function has acceptable and similar $\tilde{\chi}^2$ to the data. These truth functions are then used to produce the pseudo-data toys which are tested for potential bias from polynomial fits. First and second order fits are almost always rejected because they do not fit the data well and result in large estimates of bias.

The previously described method for background estimation is considered very conservative and with the size of the 2012 dataset even larger order polynomials may have been necessary. The study of how to minimize background parametrization uncertainty and potential bias has been a central topic of research within CMS's Higgs to two photons group. Since the discovery both the method for determining potential bias and the strategy to fit the background have been re-evaluated.

Previously potential bias was estimated as the deviation of the median signal strength among background-only pseudo-data toys. This method is difficult to use in event classes with few events, and therefore, a more traditional approach of measuring bias with the pull, defined as $< (\mu_{\text{FromToy}} - \mu_{\text{True}}) / \sigma_{\mu} >$, is used now with a threshold of 0.14, beyond which potential bias is deemed too large. The final background models (instead of various truth models) are used to produce pseudo-data toys with no signal and with SM signal injection.

For the background models themselves a novel approach is used. In this method twice the negative logarithms of the likelihood[1] (2NLL), corrected with the addition of the number of degrees of freedom, N_i, in the model for all "acceptable" models, are considered as a function of the signal strength, μ. For all values of μ the model which has the minimal 2NLL is selected. 2NLL can be written in terms of the traditional likelihood function, $\mathcal{L}$, in the following way:

$$2NLL = \min\left(-2\ln\mathcal{L}_i + N_i\right) \tag{6.6.0.1}$$

This makes the width of the 2NLL broader than just a single model. In practice there are usually only one or two models that minimize the 2NLL as a function of μ in any given event class. In Figure 6.18 an exemplifying schematic is shown of two functions making the envelope in a single event class.

The same family of functions from above are used for background parameterizations in this analysis. The order of each function type is selected using the fit quality, $\chi^2_{N\rightarrow(N+1)}$, which is determined in the following way:

$$\chi^2_{N\rightarrow(N+1)} := 2(NLL_N - NLL_{N+1}) \tag{6.6.0.2}$$

---

[1] The normal likelihood is defined in Section 7.1.

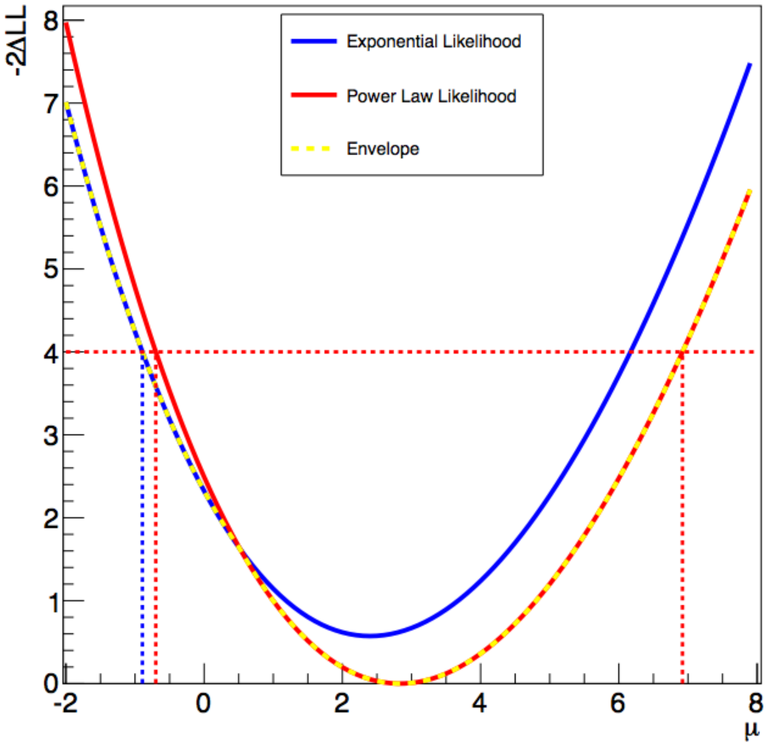


Figure 6.18: Envelope example graphic. For a fictitious example event class twice the negative logarithms of the likelihood (2NLL) is plotted for two background models. The two background models are fit to pseudo-data distributions, which are not shown here, for various values of μ. 2NLL of a power law (exponential function) with a penalty of the number of parameters is plotted in red (blue). The resulting envelope of the 2NLL function is the yellow dashed curve. The horizontal, dashed red line between the envelope is the resulting 2σ uncertainty.

$\chi^2_{N\to(N+1)}$ is used to decide whether or not order N is acceptable or if order $N+1$ fits the data *significantly better*. If order $N+1$ is not at least 5% better than N, then order N is used. This test is performed in the following manner:

$$p\left(\chi^2 \geq \chi^2_{N\to(N+1)}\right) < 0.05. \tag{6.6.0.3}$$

Even though the width of the envelope of the several selected functions is broader than that of any individual functions, the collection of models have fewer degrees of freedom compared to conservative polynomials with approximately the same estimate of potential bias. Thus, the envelope method retains the feature of minimal bias while reducing the uncertainty derived from the fit parameters. The uncertainty in expected signal strength (σ_μ) is reduced by $\sim$ 10% compared to polynomials.

6.7 Systematic Errors

In general, the effect of systematic errors in this analysis are treated in two ways: 1) as nuisance parameters embedded in the parametric models (shape systematics) or 2) as additional functions with floating parameters included as penalties in the likelihood function (efficiency systematics). Most systematic errors are treated as efficiency systematics. The technical implementation of these systematics are further discussed in the Section 7 and in this section the origin and derivation of the systematic error estimations will be discussed.

6.7.1 Photon Energy Systematics

Determining the energy of the photons is among the most vital and fundamental aspects of this analysis. Without accurate determination of the photons the analysis simply cannot work. Moreover, photons cannot all be treated identically because showering photons are more prone to energy containment problems. Obtaining accurate energy for different categories of photons is paramount to the success of the analysis. If the energy scale in multiple event classes do not "line up", then it is possible to be statistically insensitive to signal in data.

Moreover, understanding systematic uncertainties of photon energy and implementing those uncertainties within the signal models as nuisance parameters makes the models more flexible. That is, if the energy of a particular category of photons is systematically offset, then the nuisance parameter for that category, given enough flexibility, can morph the shape of the signal model to fit it during statistical analysis.

Understanding and testing the underlying assumptions in deriving corrections for photon energy is how the systematics are derived. Most of these assumptions are related to the necessity of using $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ events to derive energy scale corrections and to determine extra smearing needed for MC resolution to match the data.

The first energy scale and resolution uncertainties are from the method used to obtain energy scale corrections and additional MC smearing. They are estimated in the same η-R_9 categories, in which corrections are derived (see Section 5.5.3). A large source of systematic uncertainty in this procedure is from the difference in photons and electrons. This is estimated by re-weighting the electrons in η and R_9 such that $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ data and MC match the $\mathrm{H} \rightarrow \gamma\gamma$ MC at $\mathrm{M_H} = 90$ GeV in those variables. The entire procedure for deriving scale and resolution corrections is repeated and the differences between these corrections and the nominal analysis corrections are taken as the systematic errors in each photon category. Also, the standard energy regression applied to electrons is trained on $\mathrm{Z} \rightarrow \mathrm{e}^+\mathrm{e}^-$ MC. An alternative energy regression, which is trained on $\mathrm{H} \rightarrow \gamma\gamma$ MC instead, is applied to test the dependence and estimate the deviation from nominal as the systematic uncertainty in the same way.

Other sources of scale and smear uncertainty include selection bias from the p_T and identification cuts as well as the $\mathrm{M_{ee}}$ selection range. Systematic uncertainties are estimated by varying those thresholds before deriving alternate corrections. All these uncertainties are all summed in quadrature to produce the full uncertainty. These are implemented into the signal models as eight, partially correlated nuisance parameters. The estimated uncertainties are shown in Table 6.12 along with a summary of all of the shape systematics.

6.7.2 Further Energy Scale Systematics

In addition, once a signal is found in data, its mass becomes one of the most important measurements to be made. In an effort to ascertain the uncertainty in the mass measurement many ideas related to energy scale uncertainty have been investigated. Several additional energy scale systematics have been estimated in the course of these investigations. Only the few which are relevant are discussed here.

One major energy scale systematic uncertainty originates from the extrapolation of energy corrections derived from and tested at the Z peak (91.2 GeV) to the scale of the mass measurement ($\sim$ 125 GeV). That is, it could be that

Table 6.12: Signal model embedded scale and resolution shape systematic uncertainties. These systematics are treated within the signal models in both the Cut-based and MultiVariate Analyses so that the shapes can appropriately morph to fit signal in data. Values listed are for 8 TeV dataset and 7 TeV are very similar.

Shape systematics		**Uncertainty**	
Per photon category		**Barrel**	**Endcap**
	$\lvert\eta\rvert$ sub-categorization	< 1.0, > 1.0	< 2.0, > 2.0
Energy resolution	$R_9 > 0.94$	0.05%, 0.10%	0.07%, 0.03%
	$R_9 < 0.94$	0.05%, 0.09%	0.09%, 0.06%
Energy scale	$R_9 > 0.94$	0.03%, 0.07%	0.06%, 0.02%
	$R_9 < 0.94$	0.01%, 0.05%	0.10%, 0.08%
	$\lvert\eta\rvert$ sub-categorization	< 0.8, 0.8-1.0, > 1.0	
Photon energy scale from	$R_9 > 0.94$	0.04%, 0.12%, 0.34%	0.34%
material mismodelling	$R_9 < 0.94$	0.04%, 0.06%, 0.22%	0.22%
		7 TeV dataset	**8 TeV dataset**
	$\lvert\eta\rvert$ sub-categorization	< 1.0, > 1.0	< 1.0, > 1.0
Smearing mixing angle	$R_9 > 0.94$	50%, 50%	10%, 14%
(Barrel only)	$R_9 < 0.94$	50%, 50%	10%, 4%
Per event			
Energy scale non-linearity		0.1% 0.2% for MVA Untagged 0 20% correlation between 7 and 8 TeV	
Light collection		0.02%	
Geant4 showering		0.05%	

energy corrections needed for high p_T photons (and electrons) may not be simply proportional to the energy.

Two methods have been developed to study the energy scale non-linearity. The first compares the ratio of energy to momentum (E/p) in electrons from W and Z bosons as a function of E_T. The second examines invariant mass stability in data and MC of boosted Z boson events in bins of $H_T = E_{T,1} + E_{T,2}$. Both methods yield compatible results of at most a 0.1% effect in both analyses, except in the very high p_T MVA Untagged 0 event class where the effect is at most 0.2%. The ratio of data to MC energy responses of E/p and $\mathrm{M_{ee}}$ are shown in Figure 6.19 for the single electrons and di-electrons from Z bosons respectively. It should be noted that the use of residual E_T-dependent scale corrections in the 8 TeV analysis greatly reduces the non-linearity uncertainty.

While the energy scale corrections are derived on different MC and with

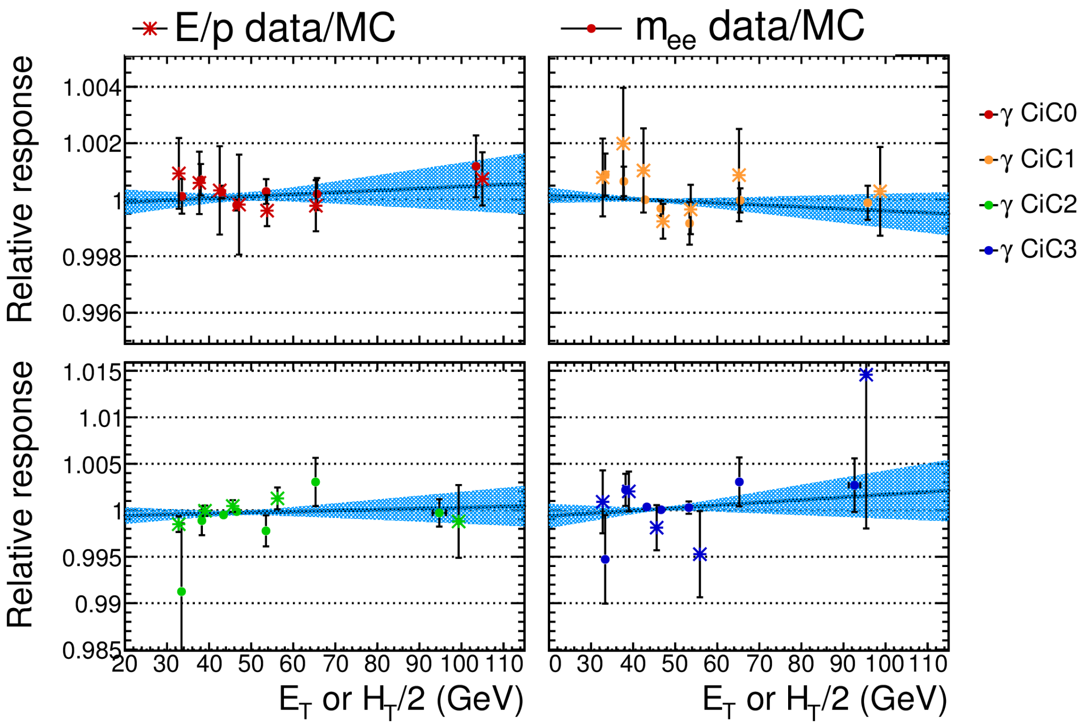


Figure 6.19: Non-linearity of energy scale in electrons. Each graph shows the data to MC ratio of single electron E/p (stars) in E_T bins or of di-electron m_{ee} (dots) in $H_T/2$ bins in the four photon identification categories.

different regressions in the 7 TeV and 8 TeV analyses, the methodology is exactly the same. Therefore, this systematic error should not be 100% correlated but it has been estimated to be up to 20% correlated. That level of correlation is conservatively assumed.

Nearly all other systematics related to energy scale emanate from differences in showering properties of electrons and photons. The largest of these is due to imperfect estimations in the amount of tracker material used in standard geometry during Geant's simulations. Energy scale corrections derived using electrons can be applied to electrons ignoring this issue. However, since showering properties are slightly different for electrons and photons, corrections derived using electrons should not be identical to those derived using photons. Thus, correcting photons with electron-derived data/MC corrections is not completely accurate.

The detector has been found to have approximately 5-10% (20%) more material in the tracker within the inner ECAL barrel (within the ECAL outer barrel and endcaps) than in the standard detector geometry used for simulation. $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ MC with extra material were developed to estimate this systematic uncertainty. The impact of this extra material is determined by comparing the “double-difference” in the mass of $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ samples with extra material MC and with nominal material MC as in:

$$\frac{1}{\mathrm{M_H}}\left([m_{\gamma\gamma}\,(\mathrm{Extra}) - \mathrm{M_{ee}}\,(\mathrm{Extra})] - [m_{\gamma\gamma}\,(\mathrm{Nom}) - \mathrm{M_{ee}}\,(\mathrm{Nom})]\right) \tag{6.7.2.1}$$

Uncertainties in η-R_9 photon categories have been derived and shown in Table 6.12. These uncertainties are all fully correlated.

A further difference between data and MC due to electron-photon differences is the modeling of the variation of the fraction of scintillation light reaching the photodetector as a function of the longitudinal depth in the crystal at which it was emitted. Uniformity of light collection in the ECAL crystals was achieved by depolishing the front face (tracker side) of each barrel crystal, but an uncertainty on the degree of uniformity achieved remains. In addition, the uniformity is modified by the radiation-induced loss of transparency of the crystals. The effect of the uncertainty, including the effect of radiation-induced transparency loss, has been simulated. It results in a difference in the energy scale between electrons and unconverted photons which is not present in the standard simulation. The resulting uncertainty on the global energy scale is 0.02%.

Finally, a small uncertainty of about 0.05% on photon energy scale is assessed for imperfect electromagnetic showering in Geant. A simulation made with an improved shower model, using the Seltzer-Berger model [78] for the bremsstrahlung energy spectrum (sub-GeV electron showering) changes the energy scale for both electrons and photons. The much smaller changes in the difference between electron and photon energy scales, although mostly consistent with zero, are interpreted as a limitation on our knowledge of the correct simulation of the showers.

6.7.3 Photon Efficiency Systematics

Further systematics are efficiency uncertainties. These systematics are summarized in Table 6.13. The systematics on the integrated luminosity, trigger and theoretical branching ratio of the Higgs are fully correlated across all event classes and all production mechanisms because they uniformly impact all signal. The production cross section uncertainties are fully correlated over all event classes but only to the associated production mechanism. The per photon efficiencies must be propagated through to di-photons per event class. In the case of the photon ID efficiency all event classes will be positively correlated because they imply removal or addition of signal.

Table 6.13: Summary of efficiency systematics. Sources of systematic uncertainty on the signal accounted for in the analysis, and applicable to events in all classes. Numbers given are for 8 TeV dataset and 7 TeV are very similar.

Efficiency systematic uncertainties		**Uncertainty**	
Per photon category		**Barrel**	**Endcap**
Photon identification efficiency		1.0%	2.6%
Photon $R_9 > 0.94$ efficiency (**Cut-based only**)		2.3%	5.5%
Photon identification BDT (**MVA only**) (Effect of up to 4.3% event class migration.)		±0.01 (shape shift)	
Energy resolution estimation (**MVA only**) (Effect of up to 8.1% event class migration.)		±10% (scaling)	
Per event		**7 TeV dataset**	**8 TeV dataset**
Integrated luminosity		2.2%	2.5%
Vertex finding efficiency		0.2%	0.2%
Trigger efficiency		1.0%	1.0%
Production cross sections		**PDF**	**Scale**
7 TeV:	Gluon fusion	+7.6% -7.1%	+7.1% -7.8%
	Vector boson fusion	+2.5% -2.1%	0.3%
	WH production	2.6%	0.9%
	ZH production	2.7%	2.9%
	$\mathrm{t\bar{t}H}$ production	8.1%	+3.2% -9.3%
8 TeV:	Gluon fusion	+7.5% -6.9%	+7.2% -7.8%
	Vector boson fusion	+2.6% -2.8%	0.2%
	WH production	2.3%	1.0%
	ZH production	2.5%	3.1%
	$\mathrm{t\bar{t}H}$ production	8.1%	+3.8% -9.3%
Branching fraction		+5.0% -4.9%	

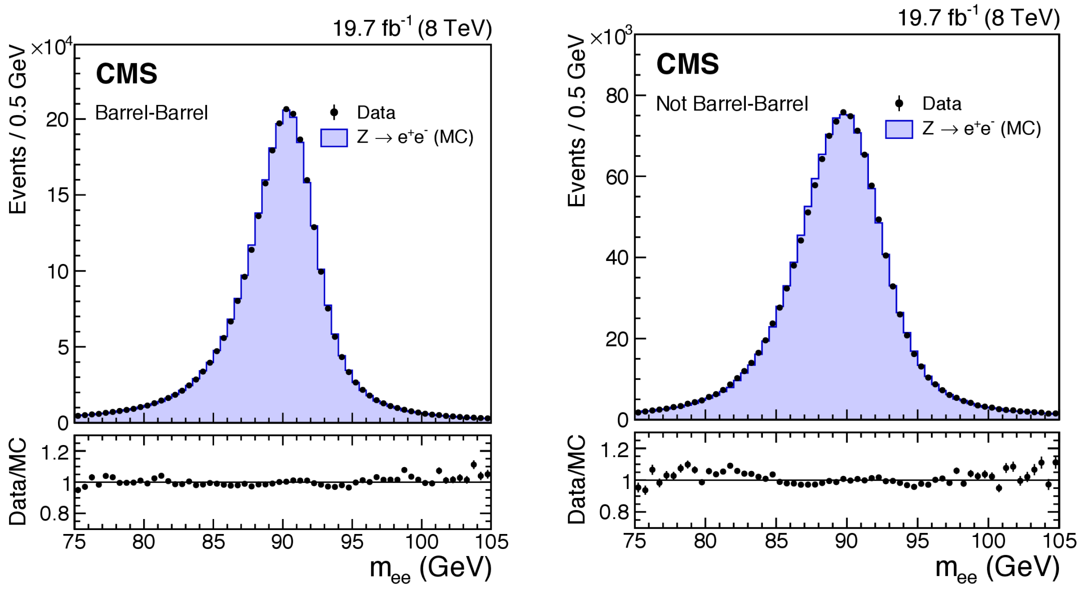
6.7.4 Di-photon BDT Systematics

To validate the background MC used in the training of the di-photon BDT, two regions of data are used for full selection comparisons without classification. The two are created by selecting 100-160 GeV and $>$ 160 GeV in data and MC. The latter should be completely signal free. The comparison of the input variables is in Appendix D. The BDT output is compared in Figure 6.5c and Figure 6.5d. The agreement in all of these plots is reasonably good. However, since the background estimates are taken from fits to the data, it is not strictly necessary to have excellent agreement between data and MC background samples.

On the other hand, validation of the signal is paramount to the success of the analysis because it is necessary to use very detailed features of the signal MC in the training and in the signal modeling. In order to validate these BDTs for signal-like events, $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ events in data and MC are again utilized. Electrons are treated like photons and electron veto is inverted. In addition the kinematic selection is adjusted to better match the kinematics of the Z boson. That is, the $p_T/m_{\gamma\gamma}$ thresholds are changed to p_T thresholds and reduced to the analysis minima of 33.3 GeV and 25 GeV for leading and sub-leading (in p_T) electrons. The invariant mass range selected is 70-120 GeV–targeting the Z boson peak. 70 GeV is the tightest $m_{\gamma\gamma}$ cut among the HLTs required for any $\mathrm{H} \rightarrow \gamma\gamma$ analysis. The invariant mass of the selected 2012 $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ events are shown in Figure 6.20 for low and high number of reconstructed vertices. The p_T distribution of the $\mathrm{Z} \rightarrow \mathrm{e}^{+}\mathrm{e}^{-}$ MC is re-weighted to match the data.

Several of the input shower-shape variables to the ID BDT and the energy regression have disagreements between data and MC. These differences were not corrected in MC before the trainings because the disagreements were smaller than in the past and because a generous systematic error is applied to the shape of both the ID BDT and the estimates of single photon energy resolution. The systematic applied to the ID BDT is determined by adding and subtracting 0.01 to the MC BDT value. This systematic covers the deviation from perfect data to MC agreement except in a few small regions as can be seen with the red hashes in Figure 6.21. The systematic applied to the estimation of the energy resolution is

±10%. The single photon resolution estimates with red hashed, systematic bands are shown in Figure 6.22.



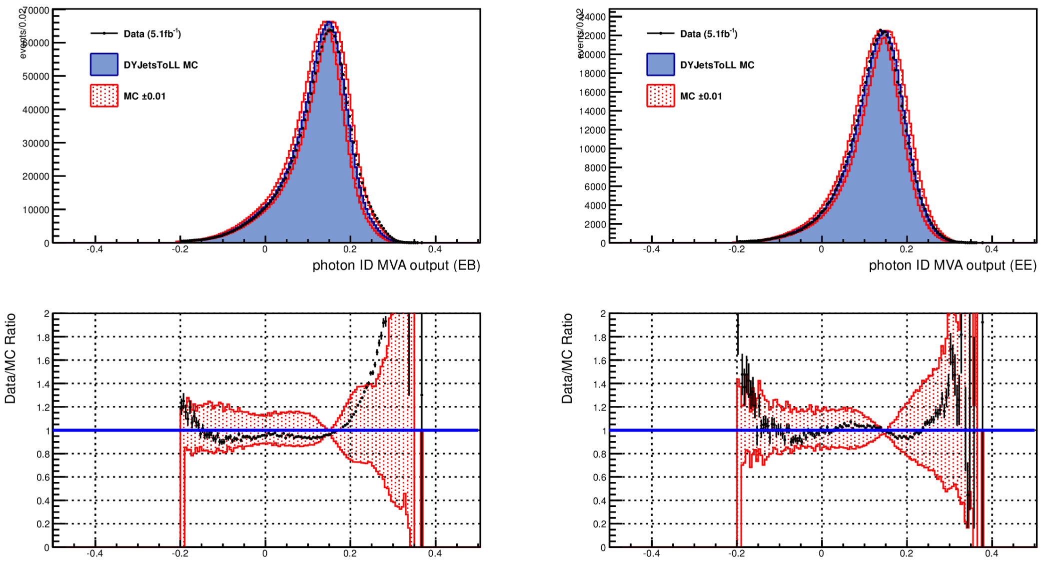
(a) (b)

Figure 6.20: Invariant mass of 8 TeV $Z \to e^{+}e^{-}$ control sample in data (points), and in simulated events (histogram), where the electrons are reconstructed as photons, and the full set of photon corrections and smearings are applied. The comparison is shown for (a) events with both showers in the barrel, and (b) the remaining events. This control sample is used for validation of di-photon BDT systematic uncertainties and estimation of energy scale non-linearity systematic uncertainty.

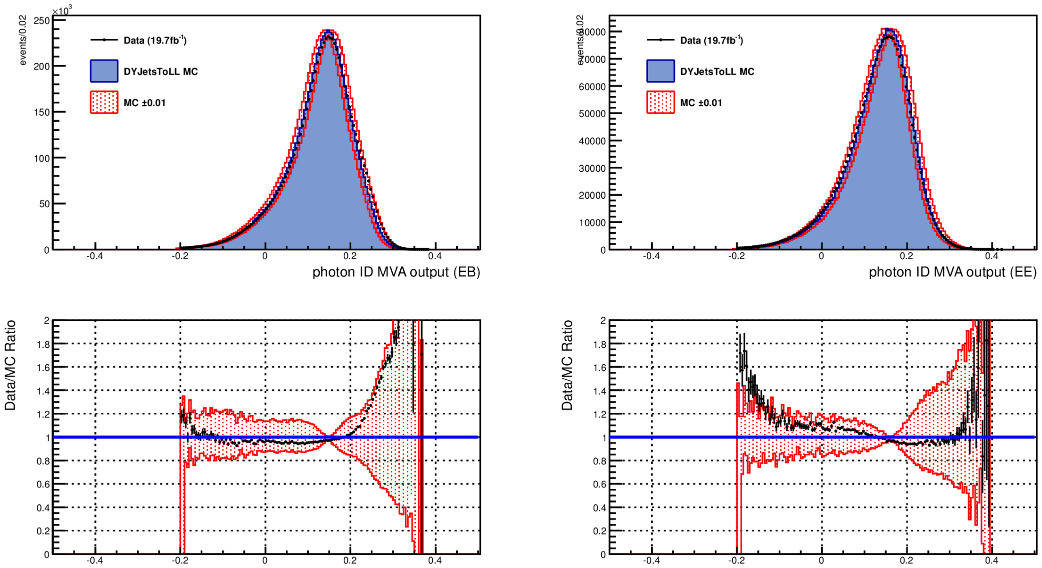
These systematics are the largest for the inputs to the di-photon BDT. When both are propagated to the di-photon BDT the resulting error bands sufficiently cover the deviation in data to MC ratio from 1 as depicted in Figure 6.23. These plots decomposed into categories of photon regions (i.e. EB-EB, EB-EE, and EE-EE) are shown Figure D.5.

6.7.5 Jet Systematics

As is true with all other event classes, it is preferable, but not strictly required, that agreement between data and background MC be very good. Indeed, in Figure 6.7c and Figure 6.7d the data agrees generally with the shape of the

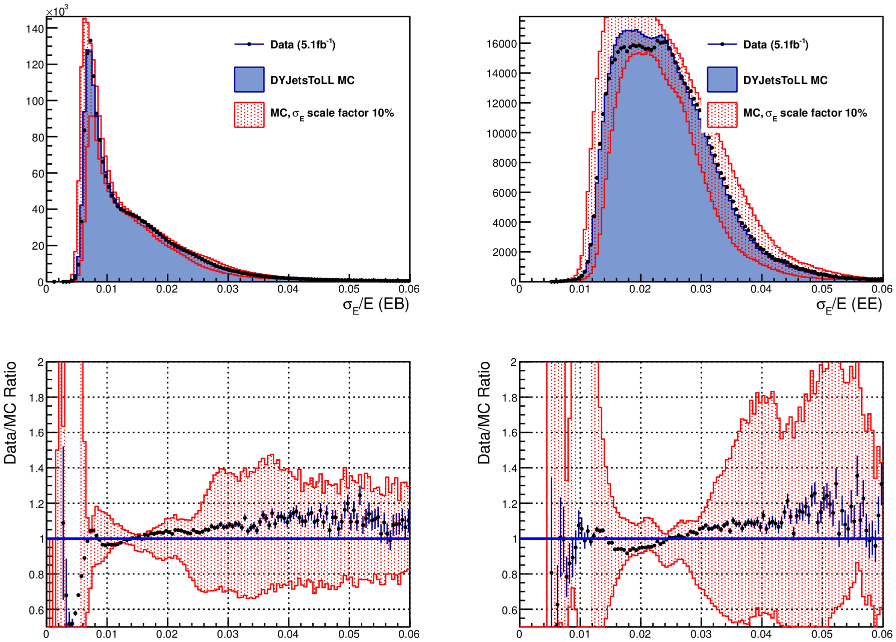


(a) 7 TeV

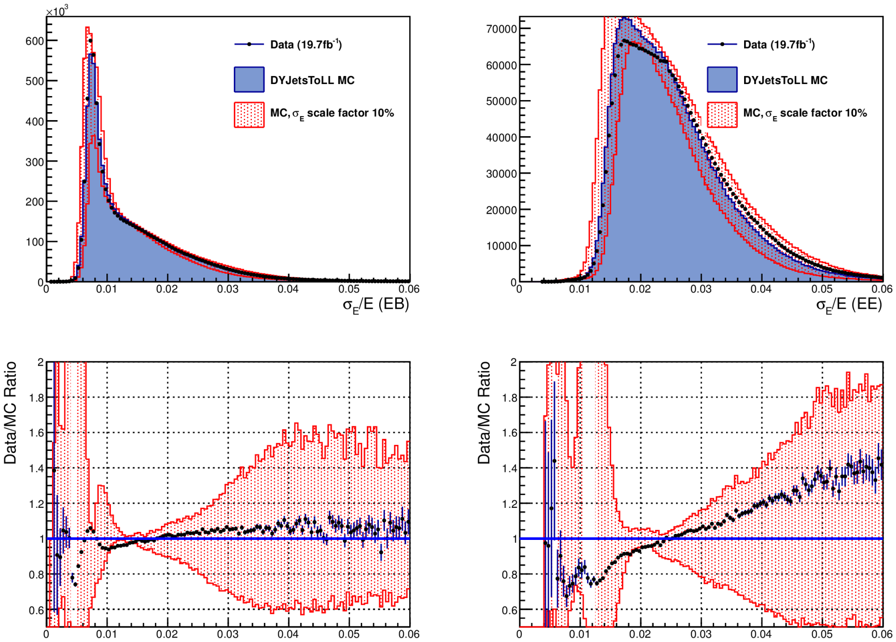


(b) 8 TeV

Figure 6.21: Photon ID BDT in Z → e^+e^- control sample with systematic uncertainty. Photon ID BDT with plus and minus 0.01 score systematics applied for EB (left) and EE (right) for 7 TeV analysis (a) and 8 TeV analysis (b).

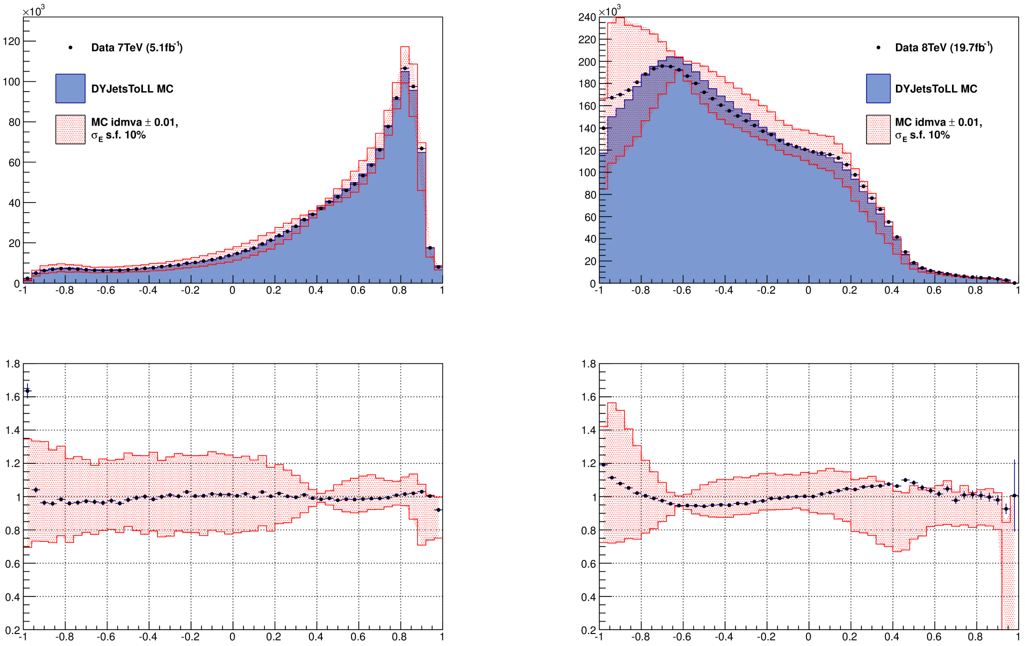


(a) 7 TeV



(b) 8 TeV

Figure 6.22: Resolution estimates in $Z \to e^+e^-$ control sample with systematics uncertainty. The single photon energy resolution estimates for EB (left) and EE (right) with ±10% systematic error are shown in both the 7 TeV analysis (a) and the 8 TeV analysis (b).



(a) 7 TeV

(b) 8 TeV

Figure 6.23: Di-photon BDT in $Z \rightarrow e^+e^-$ control sample with systematic uncertainty in data and MC with the systematics for the photon IDs and single photon energy resolution propagated to the di-photon BDT. Below shows the data to MC ratio of the plot above. In (a) is the validation for the 7 TeV analysis and in (b) for the 8 TeV analysis. The uncertainty bands generously cover the deviation from 1 in the ratio of data to MC almost everywhere.

background MC, but the finer details do not because of the high weight QCD events and a general deficit of MC at high score. No systematic is necessary to compensate for the difference because the background estimate is taken from unbiased fits to the data.

However, the signal efficiency and shape are again very important. All uncertainties affect the analysis causing events to migrate both among VBF di-jet event classes and between the di-jet classes and inclusive classes. The systematic uncertainties propagated to the analysis are factorized in accordingly. They are summarized for the 8 TeV MultiVariate Analysis in Table 6.14 and 8 TeV Cut-based Analysis in Table 6.15. The 7 TeV systematics are very similar except there is no jet pile-up ID for the 7 TeV analysis and therefore no systematic for it.

Uncertainty on the jet energy scale (JES) and the jet energy resolution (JER) are applied to the signal to estimate their effect on signal efficiency. These uncertainties are determined centrally in CMS by a dedicated Physics Object Group. The methodology is outlined in [79] although the actual jet energy corrections and associated uncertainties are updated using 2011 and 2012 data.

The $\Delta\phi$ variable used in the analysis makes the use of an additional jet veto unnecessary as both essentially remove QCD background where there is uncorrelated di-photon and multi-jet activity. This fact makes this variable sensitive to QCD scale uncertainties. This topic is thoroughly discussed in *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties* [76], which prescribes methodology for assessing the impact of the associated uncertainty. Even though the impact of this uncertainty is limited by replacing $\Delta\phi$ with $\min(\pi - 0.2, \Delta\phi)$, the QCD scale is still the dominant uncertainty on the VBF di-jet categories.

6.7.6 Other Systematics

Additional systematics are estimated for the other event classes and are listed in Table 6.16. Most of these systematics are based on identification efficiencies derived centrally in CMS and applied onto various objects in these event classes.

Table 6.14: 8 TeV MVA di-jet systematics. Summary of systematics uncertainties on jet selection in the 8 TeV analysis. The systematics in the 7 TeV MVA di-jet analysis are very similar but without the migration between the tighter two categories and the loosest.

	All VBF categories		Migration $(5+6) \rightarrow 7$		Migration $5 \rightarrow 6$	
Jet veto, QCD scale	30%		14%		5%	
	GF	VBF	GF	VBF	GF	VBF
Underlying event	3%	1%	1%	2%	1%	2%
JEC & JER	10%	4%	6%	1%	5%	1%
Jet Pile-up ID			GF	VBF	GF	VBF
Di-jet Event Class 0			3%	3%	2.5%	1%
Di-jet Event Class 1			3%	3.5%	2.5%	1%
Di-jet Event Class 2			4%	4%	2.5%	1%

Table 6.15: Cut-based di-jet systematics. Summary of systematics uncertainties on jet selection in the cut-based selection.

	All VBF categories		Migration Tight→Loose	
Jet veto, QCD scale	25%		5%	
	GF	VBF	GF	VBF
Underlying event	1%	1%	2%	1%
JEC & JER	11%	3%	2%	2%
Jet Pile-up ID	GF	VBF	GF	VBF
Di-jet Tight	3%	3%	2.5%	1%
Di-jet Loose	5.5%	5.5%	4.5%	2.5%

Table 6.16: Additional tag event class systematics. Sources of systematic uncertainty on the selection efficiency of signals events, from each of the Higgs production processes in the production-signature tagged event classes.

Event class	**Uncertainty**				
	GF	**VBF**	**WH**	**ZH**	$\mathbf{t\bar{t}}$H
VH Leptonic Tight	–	–	1.2%	1.0%	1.1%
VH Leptonic Loose	–	–	1.9%	1.6%	1.2%
VH E_T^{Missing}	4%	4%	2.6%	2.1%	4%
VH Di-jet	1%	1%	0.9%	0.9%	0.9%
$t\bar{t}$H Leptonic	–	–	–	–	1.5%
$t\bar{t}$H Hadronic	41%	–	–	–	1%

Chapter 6, in full, has been prepared for publication submission to EPJC: “Observation of the diphoton decay of the Higgs boson and measurement of its properties”. The dissertation author was the primary investigator and author of this paper.

Chapter 7

Final CMS $\mathrm{H} \to \gamma\gamma$ Analysis with Run 1 Data Results

This chapter reviews the statistical results of the analyses described in the preceding two chapters. The first several sections show the results of the Cut-based Analysis and MultiVariate Analysis side-by-side while later results are only shown for the MVA or are from a specialized analysis (e.g. the spin analysis).

7.1 The Likelihood Function

The full statistical framework for analysis in CMS and ATLAS is detailed in [80]. This section will outline this methodology and define the likelihood function which is used to derive all results.

Statistical analysis begins simply with the prediction that a certain number of events, $\mathrm{N}_{\mathrm{Exp}}$, will be selected:

$$\mathrm{N}_{\mathrm{Exp}} = \mu S + B \tag{7.1.0.1}$$

where B is the number of expected background events, S is the number of expected signal events, and μ is a signal strength multiplier. S is typically expected events from all production mechanisms summed together[1]. The natural variation among

---

[1] If there are N types of signal, then μS is expanded to $\sum_{i=1}^{N} \mu_i S_i$. An important example of this is the measurement of $\mu_{\mathrm{GF+t\bar{t}H}}$, $\mu_{\mathrm{VBF+VH}}$.

number of observed events, $\mathrm{N_{Obs}}$, is a poisson function of $\mathrm{N_{Exp}}$. That is, the probability or likelihood of observing $\mathrm{N_{Obs}}$ given an expectation of $\mathrm{N_{Exp}}$ is:

$$Prob\left(\mathrm{N_{Obs}}\right) = \frac{(\mu S + B)^{\mathrm{N_{Obs}}}}{\mathrm{N_{Obs}}!} e^{-(\mu S + B)} \tag{7.1.0.2}$$

Ignoring systematics for the moment, analyses that count the number of events in many event bins base their analysis on the product of the probabilities above. That is,

$$Prob\left(n_1, n_2, ...n_k\right) = \prod_{j=1}^{j=k} Prob(n_j) = \prod_{j=1}^{j=k} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \tag{7.1.0.3}$$

where n_j, s_j, and b_j are the number observed, signal expected and background expected in bin j of k total bins. Formally this is called the *likelihood*, and all results rely fundamentally on likelihood fits to the data under background-only ($\mu = 0$) and signal-plus-background ($\mu > 0$) assumptions.

In the $\mathrm{H} \rightarrow \gamma\gamma$ analysis, a shape analysis is performed with $m_{\gamma\gamma}$-parametrized functions for signal (fits from MC) and background (fits to the data) so instead of Equation 7.1.0.3, the likelihood for each event class is:

$$\mathcal{L}\left(\mathrm{data}|\mu\right)_{\mathrm{EvClass}} = \frac{1}{\mathrm{N_{Obs}}} \prod_{i \in \mathrm{Events}} \left(\mu S f_s\left(m_{\gamma\gamma}\left(i\right); \mathrm{M_H}\right) + B f_b\left(m_{\gamma\gamma}\left(i\right)\right)\right) e^{-(\mu S + B)} \tag{7.1.0.4}$$

where $f_s\left(m_{\gamma\gamma}\right)$ and $f_b\left(m_{\gamma\gamma}\right)$ are unity normalized *pdf*s (probability density functions) of signal and background models respectively parametrized as a function di-photon invariant mass. The likelihood above is only for one event class and the full likelihood for this analysis (without signal systematics or background envelope penalties) is simply the product of the likelihood in each event class. Notice the likelihood is also a function of the theoretical Higgs Boson mass, $\mathrm{M_H}$, as well because of the dependence of $\mathrm{M_H}$ in the signal models.

For the signal systematics the two types of systematics (shape and efficiency) are implemented differently. Shape systematics for resolution and scale are embedded in the signal model as nuisance parameters, $\vec{\theta}$. They have gaussian parametrization around the central values, $\vec{\tilde{\theta}}$, found from the signal model fits after all efficiency corrections and energy corrections have been applied as in:

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\theta-\tilde{\theta})^2}{2\sigma^2}} \tag{7.1.0.5}$$

where the means of these gaussians, $\vec{\sigma}$, are derived from systematic studies and statistical limitation of measurements. These studies and values are described in Section 6.7.

Signal efficiency systematics are almost always implemented as so-called log-normal functions of the form:

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}\ln(\kappa)} e^{-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln(\kappa))^2}} \frac{1}{\theta} \tag{7.1.0.6}$$

The log-normal functional form avoids the undesirable negative range which is unphysical for a yield multiplier.

It should be noted that systematics are typically correlated among some events classes and some signal production mechanisms. To show the full signal model with systematics the models must be decomposed finally into production mechanisms and so the full signal model for the jth event class is:

$$\sum_{k \in ProdMech} S_{jk} f_{s,jk}\left(m_{\gamma\gamma}|\mathrm{M_H}, \vec{\theta}\right) \prod_{l} \rho_{jkl}(\theta_{jkl}) \tag{7.1.0.7}$$

With the full likelihood function described, the method of optimizing the likelihood can be addressed. Let us denote the full likelihood including nuisance parameters as $\mathcal{L}\left(\mathrm{data}|\mu, \mathrm{M_H}, \vec{\theta}\right)$. First, nuisance parameters ($\vec{\theta}$) are "profiled" fixing the value of quantity (or quantities) to be measured (μ, $\mathrm{M_H}$ or both). The likelihood is maximized[2] floating $\vec{\theta}$ for *many* fixed values of μ. At each fixed value of μ there will be a maximum in the vector space of $\vec{\theta}$, denoted $\hat{\vec{\theta}}(\mu)$. From here the profile-likelihood function can be defined as maximum likelihood of all $\hat{\vec{\theta}}(\mu)$:

$$\mathcal{L}_{\mathrm{Pro}}(\mu) = \max_{\hat{\vec{\theta}}(\mu')} \mathcal{L}\left(\mathrm{data}, \hat{\vec{\theta}}(\mu')\,|\mu\right) \tag{7.1.0.8}$$

A fixed value of $\mathrm{M_H}$ is assumed for simplicity. This profile-likelihood function is what is maximized in the determination of all final results

---

[2] In practice 2NLL is minimized.

The envelope method, described in Section 6.6, employed in this analysis applies a penalty to twice the negative logarithm of the likelihood (2NLL) of the number of parameters to be fit in the full model. The envelope method can be thought of as profiling over all background models.

7.2 Selected Data and Measured Signal Strength

The data fits shown here are profile-likelihood optimized with the mass fixed to the most significant[3] mass of $\mathrm{M_H} = 124.7$ GeV and μ floated and measured. This best value of $\mathrm{M_H}$ is determined by profiling the mass along with the nuisance parameters during the mass measurement to be described in Section 7.5. In Figure 7.1, Figure 7.2, Figure 7.3, and Figure 7.4 are the signal plus background fits overlaid by data in the 25 event classes of the MultiVariate Analysis. In Figure 7.5, Figure 7.6, Figure 7.7, and Figure 7.8 are the signal plus background fits overlaid by data in the 31 event classes of the Cut-based Analysis. Each plot contains one and two-sigma error bands of the background portion of the fit. The red curves are the signal-plus-background function. Accompanying each data with full model plot, there is the background-subtracted data.

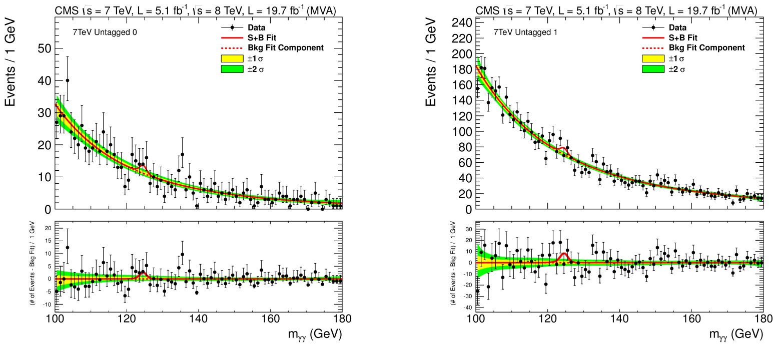
In most event classes signal in data is not extremely visible (with the notable exception of the 7 TeV tight di-jet event class in both analyses). However, when all event classes are summed the signal becomes apparent. Figure 7.9a and Figure 7.9b show the sum of all the event classes in data with the signal-plus-background profile-likelihood fit overlaid for the MultiVariate and Cut-based analyses.

Another way to combine the different event classes is to combine them by weighting with the S/(S+B) of that event class. In this way event classes with large background will be de-emphasized to make the signal more apparent. This interpretation fairly reflects the relative importance of events in the likelihood and is shown for both analyses in Figure 7.9c and Figure 7.9d. The weighted plots are not used to extract any results and are primarily for visualization purposes.

From the profile likelihood fits of each analysis the Standard Model signal

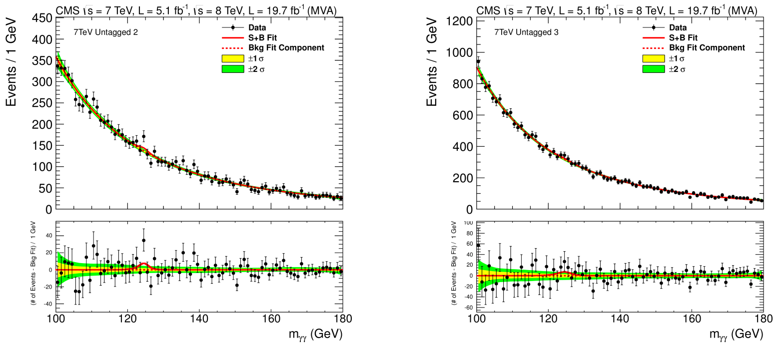
---

[3] Significance is defined in Section 7.4.



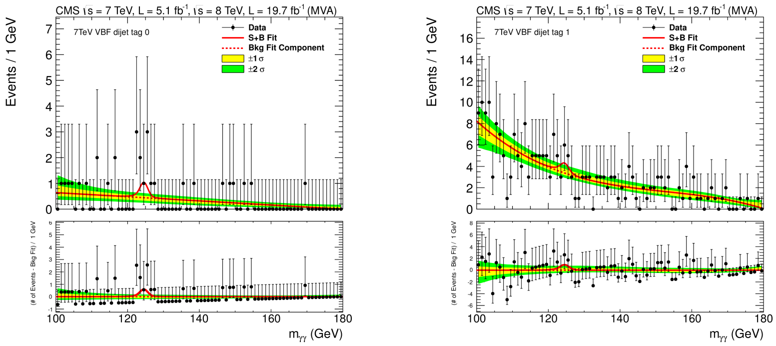
(a) Untagged 0
7 TeV - MVA

(b) Untagged 1
7 TeV - MVA



(c) Untagged 2
7 TeV - MVA

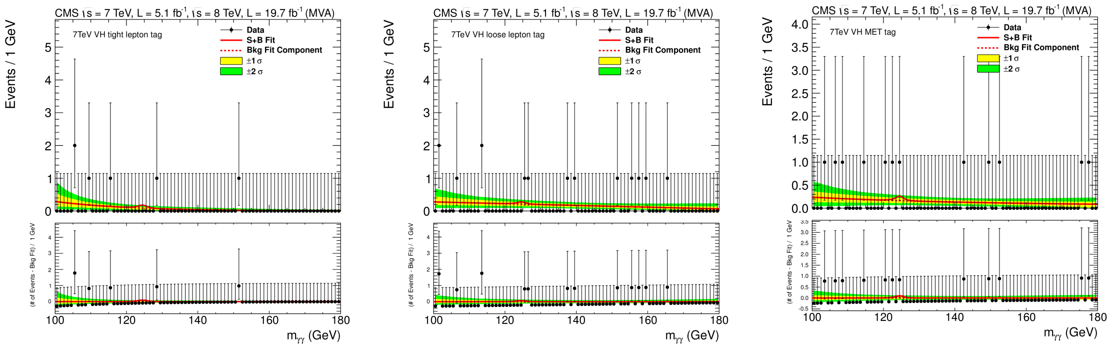
(d) Untagged 3
7 TeV - MVA



(e) Di-jet 0
7 TeV - MVA

(f) Di-jet 1
7 TeV - MVA

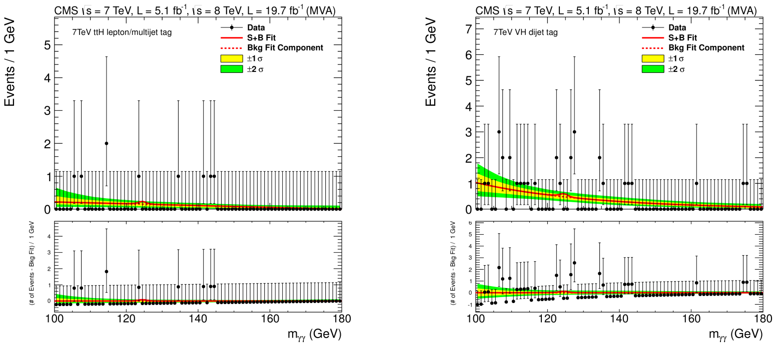
Figure 7.1: 7 TeV MVA inclusive and di-jet event classes' selected data with full model fits.



(a) VH Leptonic Tight
7 TeV - MVA

(b) VH Leptonic Loose
7 TeV - MVA

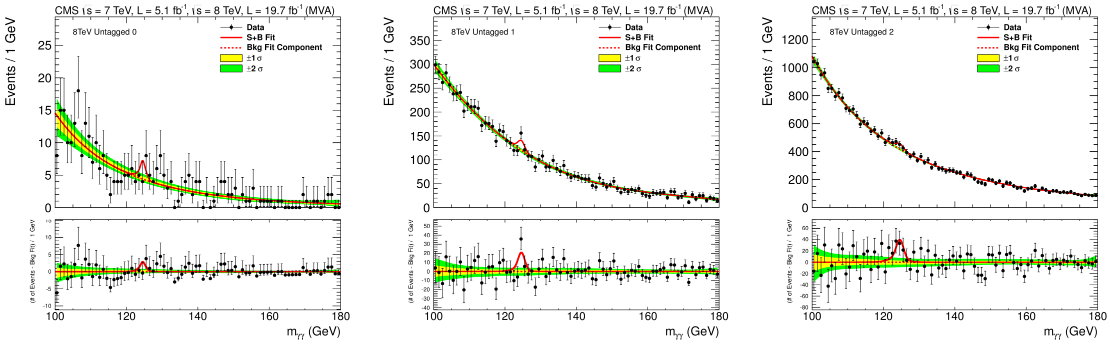
(c) VH E_T^{Missing}
7 TeV - MVA



(d) $t\bar{t}$H Combined
7 TeV - MVA

(e) VH Di-jet
7 TeV - MVA

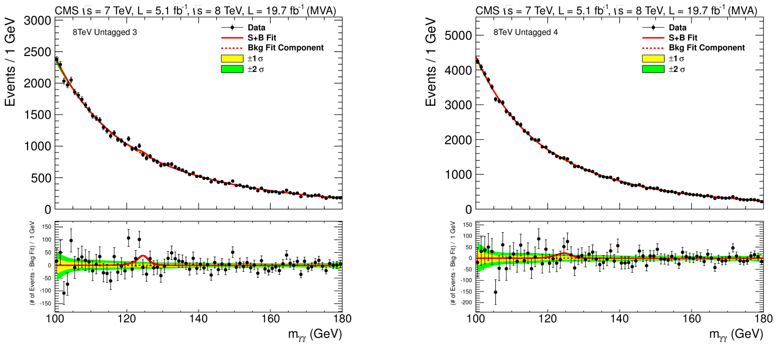
Figure 7.2: 7 TeV MVA additional tag event classes' selected data with full model fits.



(a) Untagged 0
8 TeV - MVA

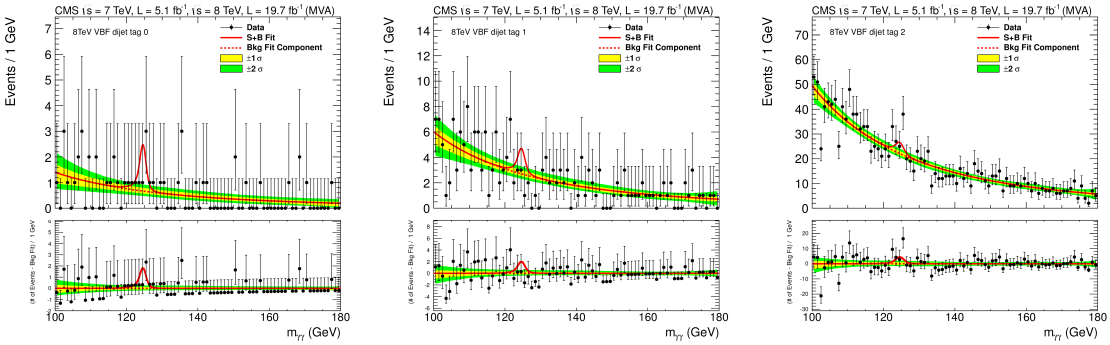
(b) Untagged 1
8 TeV - MVA

(c) Untagged 2
8 TeV - MVA



(d) Untagged 3
8 TeV - MVA

(e) Untagged 4
8 TeV - MVA

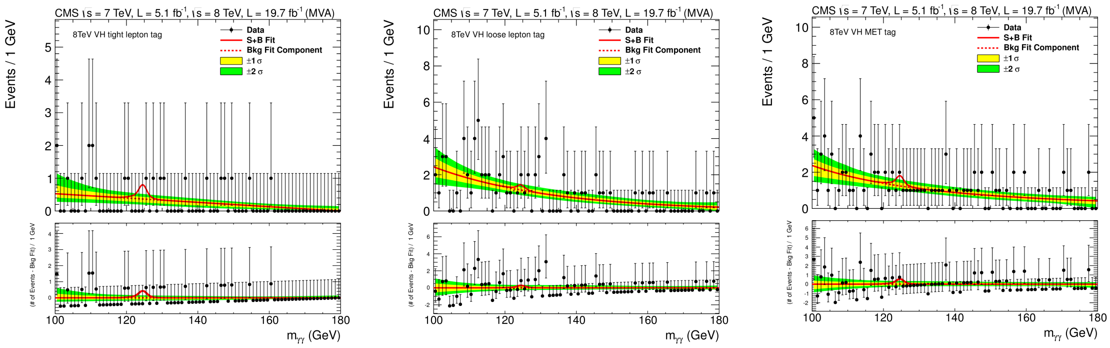


(f) Di-jet 0
8 TeV - MVA

(g) Di-jet 1
8 TeV - MVA

(h) Di-jet 2
8 TeV - MVA

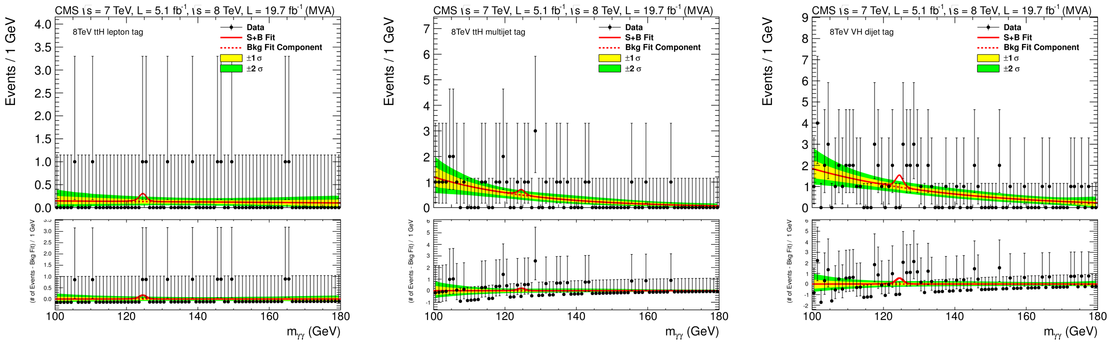
Figure 7.3: 8 TeV MVA inclusive and di-jet event classes' selected data with full model fits.



(a) VH Leptonic Tight 8 TeV - MVA

(b) VH Leptonic Loose 8 TeV - MVA

(c) VH E_T^{Missing} 8 TeV - MVA

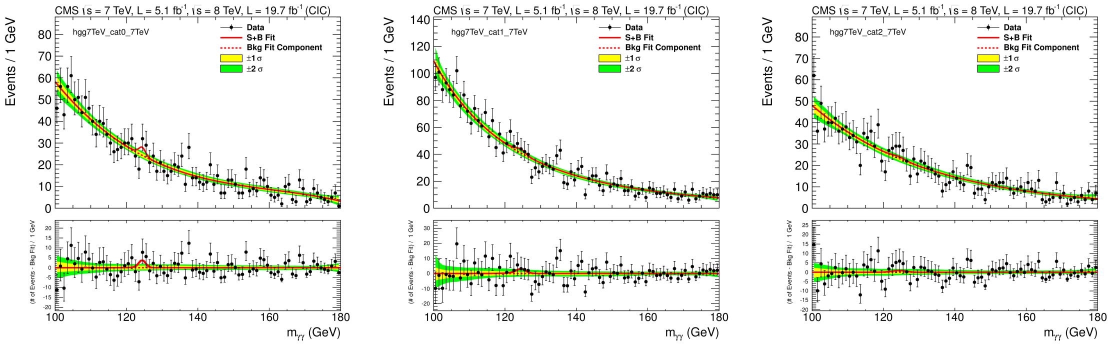


(d) $t\bar{t}$H Leptonic 8 TeV - MVA

(e) $t\bar{t}$H Hadronic 8 TeV - MVA

(f) VH Di-jet 8 TeV - MVA

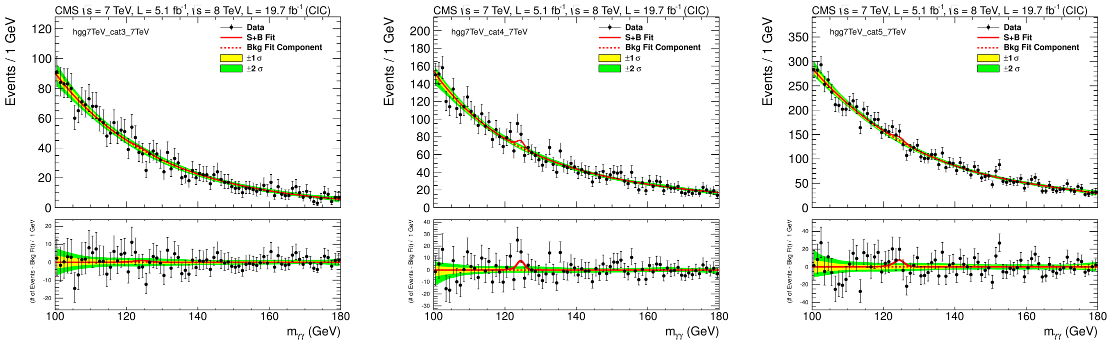
Figure 7.4: 8 TeV MVA additional tag event classes' selected data with full model fits.



(a)
EB-EB, High R_9, High p_T
7 TeV - Cut-based

(b)
EB-EB, Low R_9, High p_T
7 TeV - Cut-based

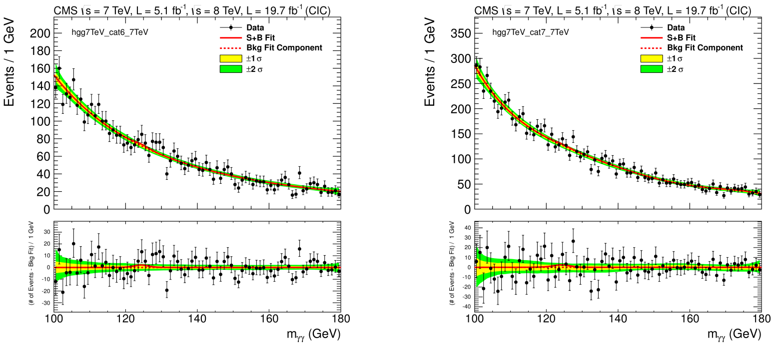
(c)
Mixed, High R_9, High p_T
7 TeV - Cut-based



(d)
Mixed, Low R_9, High p_T
7 TeV - Cut-based

(e)
EB-EB, High R_9, Low p_T
7 TeV - Cut-based

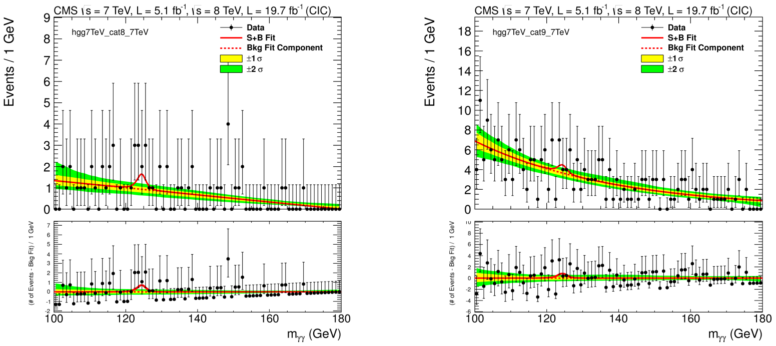
(f)
EB-EB, Low R_9, Low p_T
7 TeV - Cut-based



(g)
Mixed, High R_9, Low p_T
7 TeV - Cut-based

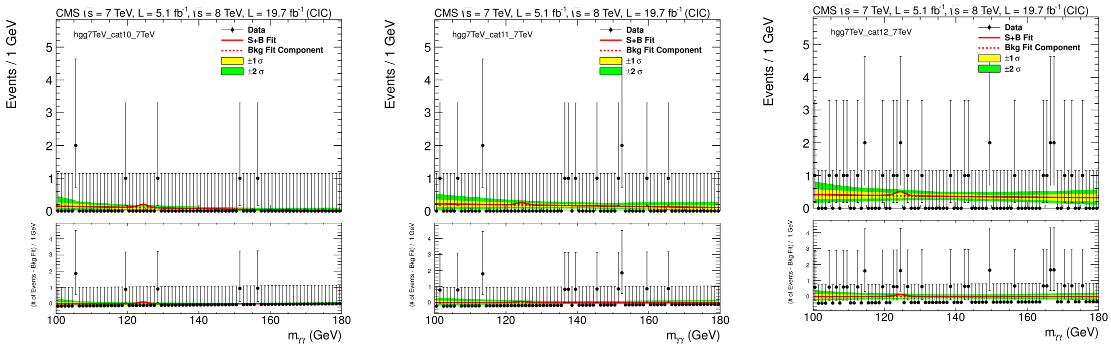
(h)
Mixed, Low R_9, Low p_T
7 TeV - Cut-based

Figure 7.5: 7 TeV cut-based inclusive event classes' selected data with full model fits.



(a) Di-jet 0
7 TeV - Cut-based

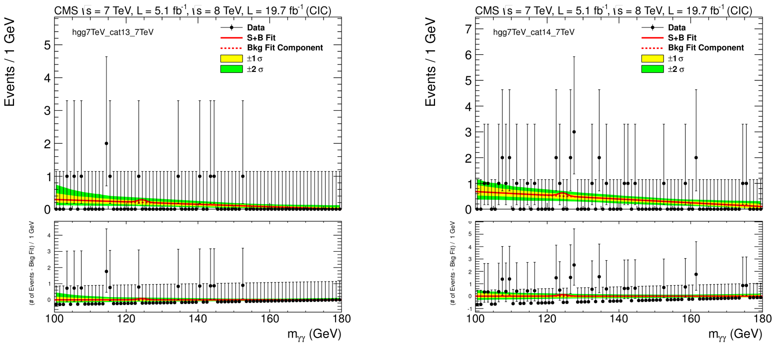
(b) Di-jet 1
7 TeV - Cut-based



(c) VH Leptonic Tight
7 TeV - Cut-based

(d) VH Leptonic Loose
7 TeV - Cut-based

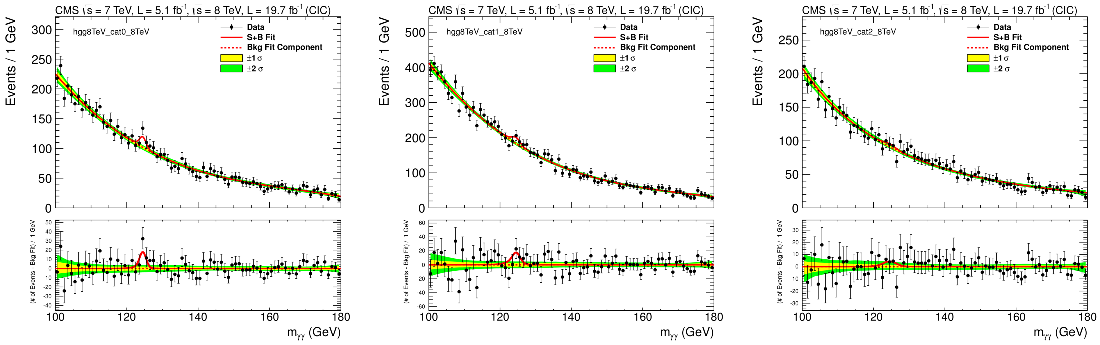
(e) VH E_T^{Missing}
7 TeV - Cut-based



(f) $\mathrm{t\bar{t}H}$ Combined
7 TeV - Cut-based

(g) VH Di-jet
7 TeV - Cut-based

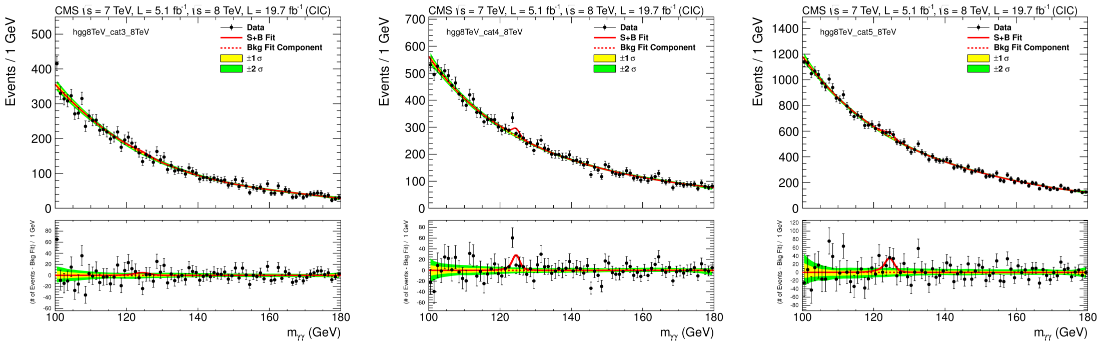
Figure 7.6: 7 TeV cut-based additional tag event classes' selected data with full model fits.



(a)
EB-EB, High R_9, High p_T
8 TeV - Cut-based

(b)
EB-EB, Low R_9, High p_T
8 TeV - Cut-based

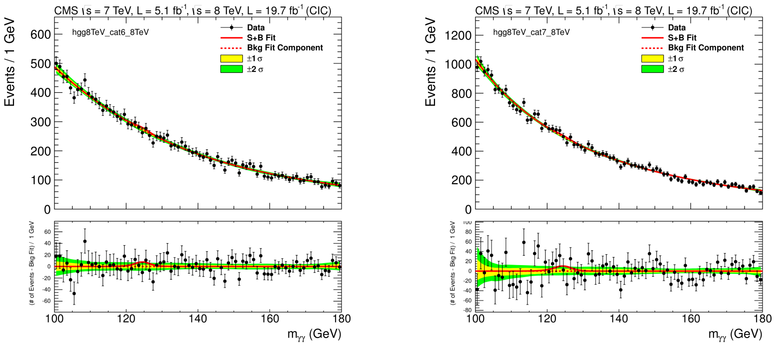
(c)
Mixed, High R_9, High p_T
8 TeV - Cut-based



(d)
Mixed, Low R_9, High p_T
8 TeV - Cut-based

(e)
EB-EB, High R_9, Low p_T
8 TeV - Cut-based

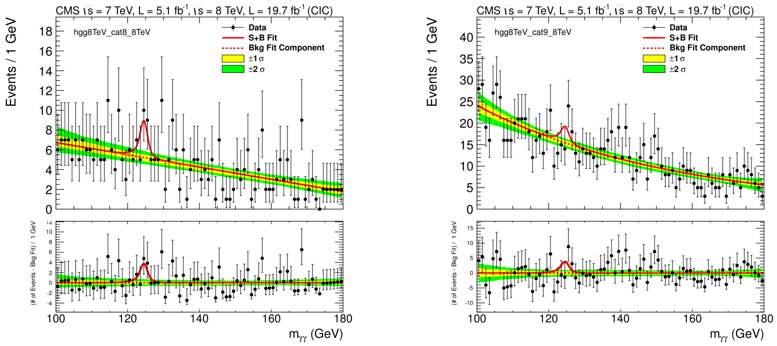
(f)
EB-EB, Low R_9, Low p_T
8 TeV - Cut-based



(g)
Mixed, High R_9, Low p_T
8 TeV - Cut-based

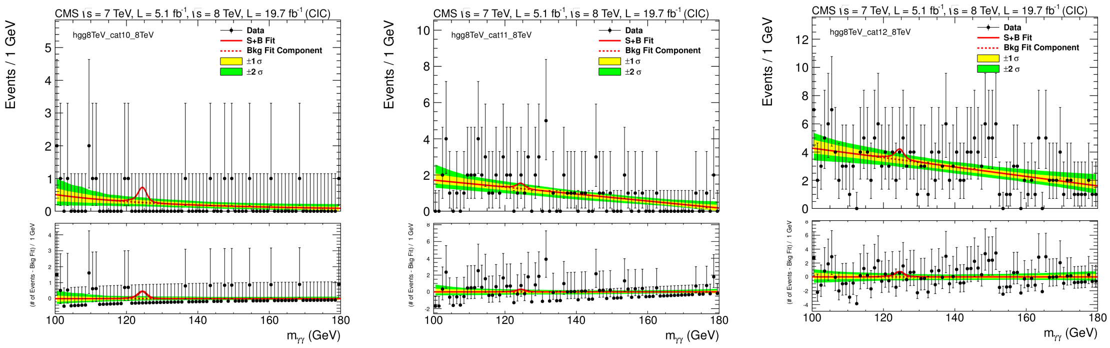
(h)
Mixed, Low R_9, Low p_T
8 TeV - Cut-based

Figure 7.7: 8 TeV cut-based inclusive event classes' selected data with full model fits.



(a) Di-jet 0
8 TeV - Cut-based

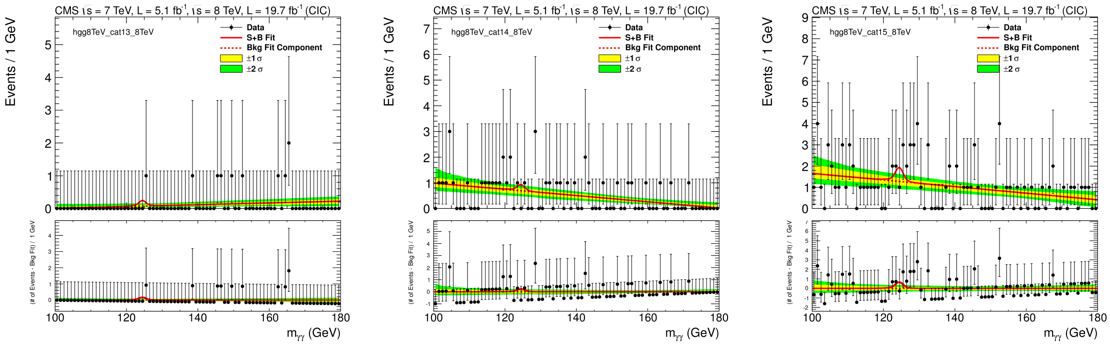
(b) Di-jet 1
8 TeV - Cut-based



(c) VH Leptonic Tight
8 TeV - Cut-based

(d) VH Leptonic Loose
8 TeV - Cut-based

(e) VH E_T^{Missing}
8 TeV - Cut-based

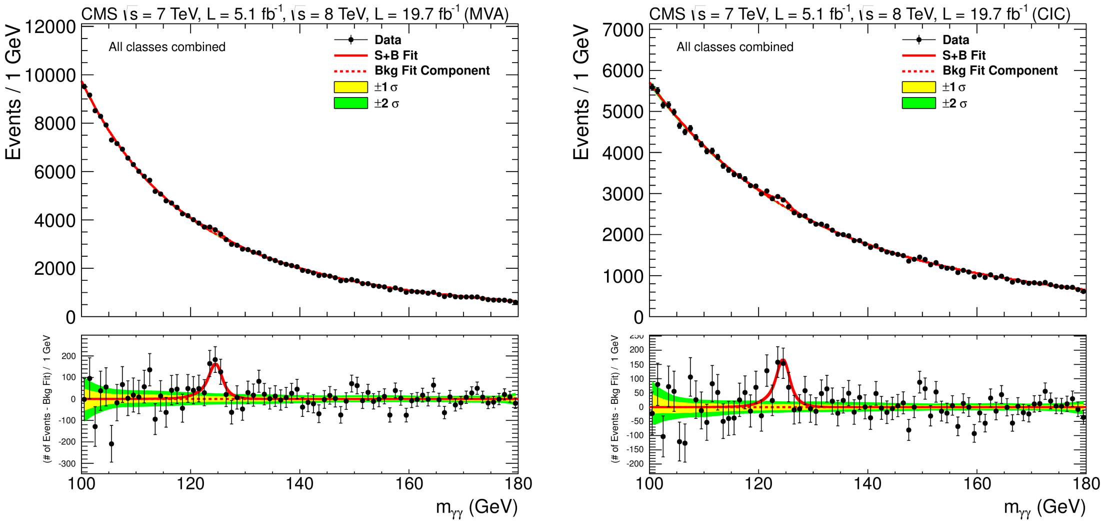


(f) $\mathrm{t\bar{t}H}$ Leptonic
8 TeV - Cut-based

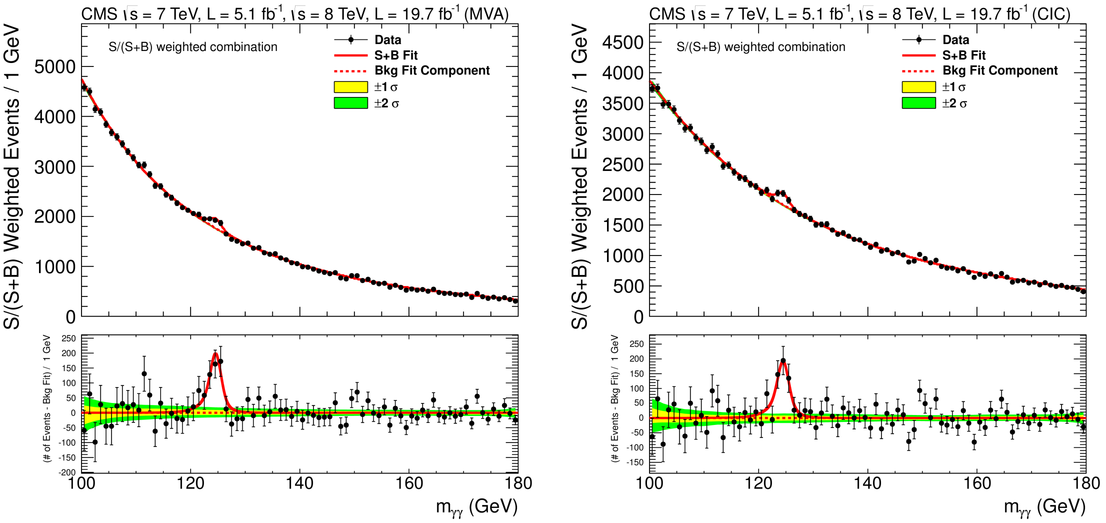
(g) $\mathrm{t\bar{t}H}$ Hadronic
8 TeV - Cut-based

(h) VH Di-jet
8 TeV - Cut-based

Figure 7.8: 8 TeV cut-based additional tag event classes' selected data with full model fits.



(a) MVA - Unweighted (b) Cut-based - Unweighted



(c) MVA - Weighted (d) Cut-based - Weighted

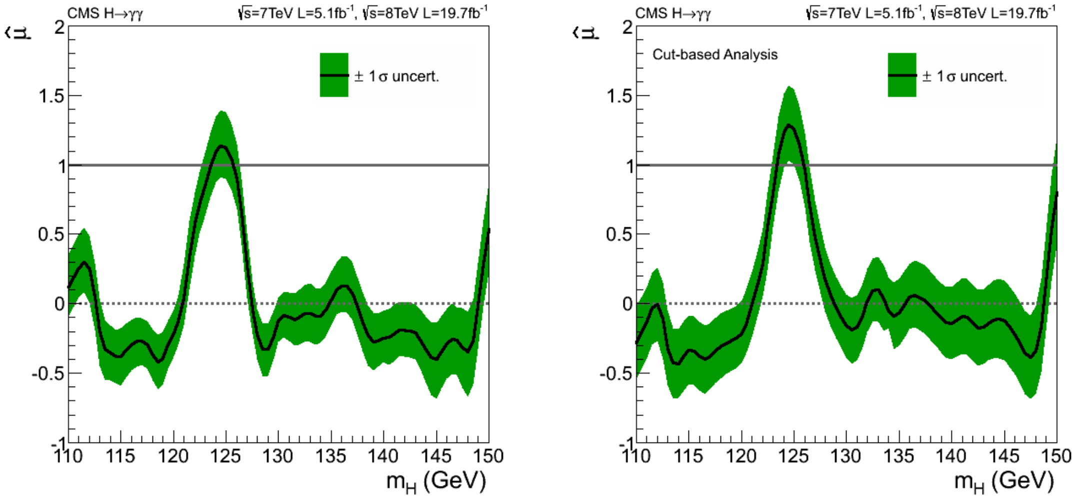
Figure 7.9: Combined invariant mass distributions. In (a) and (b) the event classes and signal plus background fits are summed directly. In (c) and (d) the event classes are weighted by the S/(S+B) and summed. They are normalized such that the total number of signal events in preserved. Each plot has an accompanying plot below of the data with the background subtracted leaving the signal. The signal is visible in both analyses.

strength is extracted. These results are summarized in Table 7.1. The observed signal strength that is compatible to the Standard Model near or below the 1σ level.

Table 7.1: Observed signal strength summary in each dataset and the combined for both the analyses are listed below.

	MultiVariate Analysis		Cut-based Analysis	
Dataset	$\hat{M}_H$ (GeV)	$\hat{\mu}$	$\hat{M}_H$ (GeV)	$\hat{\mu}$
7 TeV	124.2	$2.22^{+0.62}_{-0.55}$	124.1	$2.42^{+0.70}_{-0.62}$
8 TeV	124.9	$0.90^{+0.26}_{-0.23}$	124.8	$1.07^{+0.29}_{-0.27}$
7 + 8 Tev	124.7	$1.14^{+0.26}_{-0.23}$	124.6	$1.29^{+0.29}_{-0.26}$

Signal strength as a function of M_H is shown in Figure 7.10. The green band corresponds to the $\pm 1\sigma$ error bands on the measured value of $\hat{\mu}$ for both analyses. The analyses agree very well throughout the search range. In addition to the clear near SM-like signal strength near 125 GeV, both analyses observe an excess of events just outside of the search range with significance just over 2σ at 151 GeV.



(a) MVA (b) Cut-based

Figure 7.10: Best-fit signal strength, $\hat{\mu}$, as a function of the mass hypothesis, M_H. The results are shown for the MultiVariate Analysis (a) and for the Cut-based Analysis in (b).

7.3 Exclusion Limits

Exclusion limits and significance of observed excess are both based on the likelihood ratio of signal plus background and background-only. In each case an appropriate test statistic, $\tilde{q}_\mu = -2\ln\left(\frac{\mathcal{L}_1}{\mathcal{L}_2}\right)$, is prepared and evaluated for both assumptions. In the H → $\gamma\gamma$ analysis these evaluations are done at fixed $\mathrm{M_H}$ at reasonable intervals throughout the range of 110-150 GeV. The scan is finer near the Higgs Boson near 125 GeV.

Exclusion limits at 95% confidence level are typical standard to determine the absence of a given process at a given signal strength, μ. The test statistic used for to determine exclusion limits in CMS and ATLAS is:

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}\left(data|\mu,\hat{\theta}_\mu\right)}{\mathcal{L}\left(data|\hat{\mu},\hat{\theta}\right)}, \text{ where } 0 \leq \hat{\mu} \leq \mu. \tag{7.3.0.9}$$

$\hat{\mu}$ is constrained to be ≥ 0 to be physical, and it is not allowed to exceed the given value of μ so that over fluctuations are not penalized. $\hat{\theta}_\mu$ is the set of floating parameters in the fit. For numerous, fixed values of μ pseudo-data toys of signal plus background are produced and $\tilde{q}_\mu$ is evaluated for each as if it is data. This is also done with the background-only assumption. The unity-normalized *pdf*s of $\tilde{q}_\mu$, $f\left(\tilde{q}_\mu\right)$, and $\tilde{q}_0$, $f\left(\tilde{q}_0\right)$, are used to determine the confidence level of signal at that μ in the following manner.

The confidence level of signal plus background ($\mathrm{CL_{S+B}}$) and background-only ($\mathrm{CL_B}$) must be measured before finally arriving at the desired ($\mathrm{CL_S}$). This is done by evaluating the test statistic using the data, $\tilde{q}_\mu^{obs}$. The $\mathrm{CL_{S+B}}$ is defined as the probability of observing a less compatible (larger) value of the test statistic.

$$\mathrm{CL_{S+B}} = \int_{\tilde{q}_\mu^{\mathrm{obs}}}^{\infty} f\left(\tilde{q}_\mu\right) d\tilde{q}_\mu \tag{7.3.0.10}$$

$\mathrm{CL_{S+B}}$ in principle could be used to set exclusion limits on signal strength since it is the probability of observing anything less compatible with the signal plus background hypothesis. However, when there is little or no signal combined with a downward background fluctuation, there may be strong exclusions down to $\mu = 0$ [81].

Therefore, the $\mathrm{CL_S}$ method was defined ($\mathrm{CL_S} = \mathrm{CL_{S+B}}/\mathrm{CL_B}$) to avoid making strong exclusions in the case where there is very little sensitivity. $\mathrm{CL_B}$ is analogously defined to $\mathrm{CL_{S+B}}$:

$$\mathrm{CL_B} = \int_{\tilde{q}_\mu^{\mathrm{obs}}}^{\infty} f\left(\tilde{q}_0\right) d\tilde{q}_0 \tag{7.3.0.11}$$

With the computation of $\mathrm{CL}_\mathrm{B}^\mu$ and $\mathrm{CL}_\mathrm{S+B}^\mu$ at fixed μ, $\mathrm{CL}_\mathrm{S}^\mu$ is computed. If $\mathrm{CL}_\mathrm{S}^\mu \leq \alpha$, then the signal strength of μ is excluded at the $1-\alpha$ level. The minimum value of μ, for which $\mathrm{CL}_\mathrm{S}^\mu \leq 0.5$, is the 95% CL exclusion, $\mu^{95\%\mathrm{CL}}$.

This would require a rather CPU intensive exercise to produce $f\left(\tilde{q}_\mu\right)$ at all μs necessary to determine $\mu^{95\%\mathrm{CL}}$. This must be done separately for all values of $\mathrm{M_H}$ which are tested. However, methods have been developed and validated to determine $f\left(\tilde{q}_\mu\right)$ without creating all of the CPU intense pseudo-data toys [82].

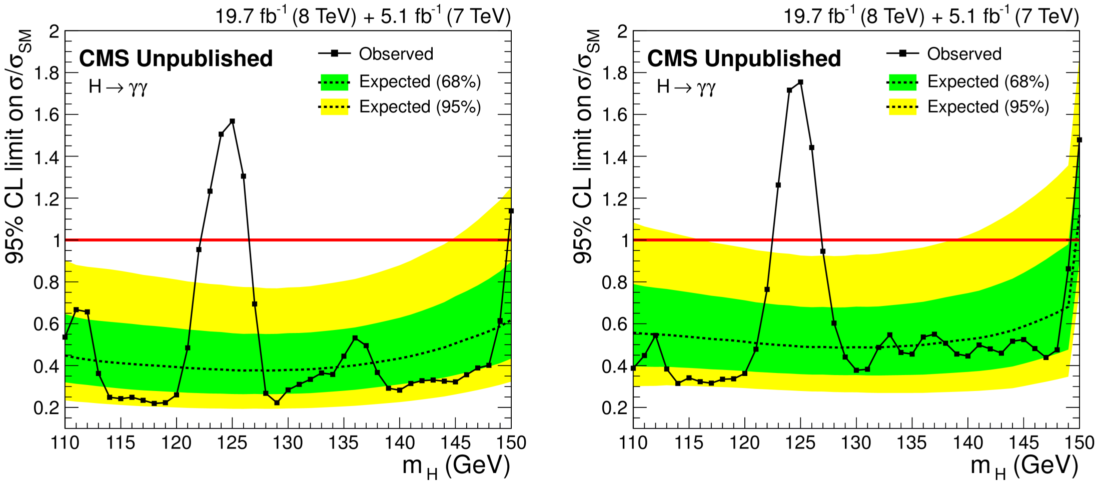
The results of the $\mathrm{CL_S}$ 95% CL exclusion limits are reported in Figure 7.11a and Figure 7.11b for the MultiVariate Analysis and Cut-based Analysis respectively. Values of $\mathrm{M_H}$ are tested every 1 GeV in the range of 110-150 GeV. The presence of the Higgs Boson near 125 GeV can be seen in these figures, where the Standard Model Higgs Boson is excluded at the 95% CL for all values of $\mathrm{M_H}$ except between 122-127 GeV and between 149-150 GeV which is near a more than 2σ excess at 151 GeV, see Figure 7.9.

7.4 Excess Significance

An additional test statistic can be used to determine the significance of an observed excess. The following test statistic examines the compatibility of the background-only hypothesis with the background plus any fitted signal strength:

$$q_0 = -2\ln\frac{\mathcal{L}\left(data|0,\hat{\hat{\theta}}_0\right)}{\mathcal{L}\left(data|\hat{\mu},\hat{\theta}\right)}, \text{ where } 0 \leq \hat{\mu}. \tag{7.4.0.12}$$

Again, the unity-normalized *pdf* of q_0, $f\left(q_0\right)$, can be made from pseudo-data toys under the background-only assumption. With $f\left(q_0\right)$ the probability, p_0, of observ-



(a) MVA (b) Cut-based

Figure 7.11: Exclusion limits for H → $\gamma\gamma$ analyses. In (a) and (b) show the 95% confidence level exclusion limits in μ as a function of mass in black for the two analyses. The red line at $\mu = 1$ is the Standard Model Higgs Boson cross section. The dotted black line is the expected exclusion limits and the green and yellow bands are the 1σ and 2σ bands respectively. All masses of a Standard Model Higgs Boson are excluded except between $122 - 127$ GeV where a Higgs Boson has been observed and between $149 - 150$ GeV where there is an excess of events just beyond 150 GeV.

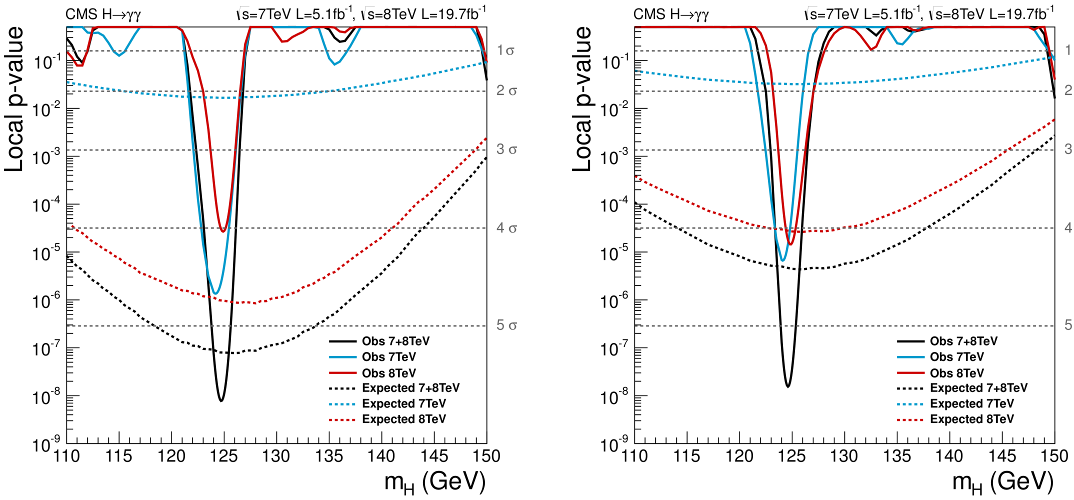
ing an excess with q_0^{obs} or less compatible by integrating:

$$p_0 = \int_{q_0^{\text{obs}}}^{\infty} f\left(q_0\right) dq_0 \qquad (7.4.0.13)$$

This probability is converted into the number of standard deviations, Z, needed in a gaussian distribution with width of one and mean of zero to produce a one-sided tail with the same probability. That is,

$$p_0 = \int_{z=Z}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{-z^2}{2}} dz \qquad (7.4.0.14)$$

The convention within particle physics is to require Z=5, corresponding to $p_0 = 2.8 \times 10^{-7}$, in order to claim the discovery of a process over the background-only assumption.



(a) MVA

(b) Cut-based

Figure 7.12: Excess significance for H → $\gamma\gamma$ analyses. In (a) and (b) the expected and observed significances of the MultiVariate and Cut-based Analyses are shown for the full Run 1 dataset in black. The datasets for 2011 and 2012 are shown separately in blue and red respectively. The MultiVariate Analysis reports a combined significance of 5.7σ (5.2σ expected) while Cut-based analysis reports 5.5σ (4.4σ expected).

A new particle was discovered in July 2012 [29] by both the CMS and ATLAS collaborations independently in their searches for a Standard Model Higgs Boson. Significant excesses in two channels (H → $\gamma\gamma$ and H → ZZ → 4ℓ) were observed and combined to reach $\geq 5\sigma$ significance in both collaborations' datasets.

In Figure 7.11a and Figure 7.11b, the significance in $\mathrm{H} \rightarrow \gamma\gamma$ channel using the final calibration and analysis as described in this dissertation are shown (for the MultiVariate Analysis and Cut-based Analysis respectively). The MultiVariate Analysis reports a significance of 5.7σ particle excitation to two photons at $\mathrm{M_H} = 124.7\mathrm{GeV}$ where only 5.2σ is expected for a Standard Model Higgs Boson. The Cut-based Analysis reports 5.5σ at $\mathrm{M_H} = 124.6\mathrm{GeV}$ where only 4.4σ is expected. Both analyses confirm the independent observation of this Standard Model-like Higgs Boson.

7.5 Mass Measurement

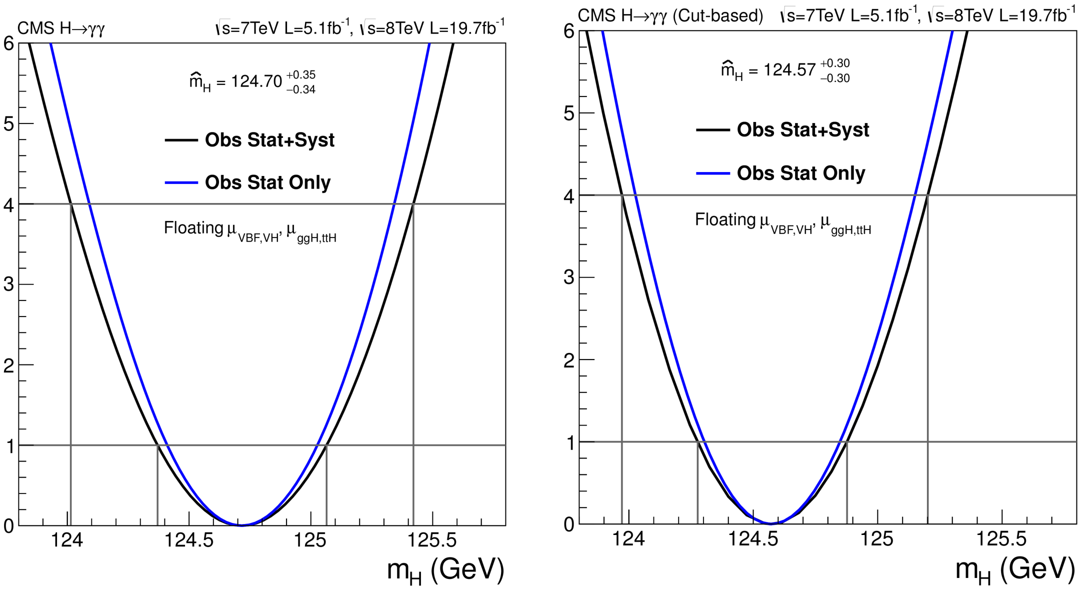
The high resolution nature of the $\mathrm{H} \rightarrow \gamma\gamma$ final state presents the opportunity for a relatively high precision measurement of $\mathrm{M_H}$. Since the July 2012 observation, lengthy and painstaking work has been done both to improve the sensitivity of the analysis and to reduce systematic errors related to the overall energy scale. Considerable effort has gone into the development of, for example, the development of E_T-dependent energy corrections which have reduced the systematic uncertainty on the non-linearity of the energy scale.

Likewise, many sources of systematic uncertainty have been evaluated for their impact on energy scale. In Section 6.7.2 several systematic uncertainties on the energy scale have been described. These systematic uncertainties, along with the scale and resolution correction uncertainties, are the primary contributors to the error in the mass measurement. Table 7.2 qualifies the amount of uncertainty in GeV due to each of these with all the minor uncertainties combined.

The best fit mass is determined by fitting $\mathrm{M_H}$ after profiling the nuisance parameters. A slightly modified likelihood is used to make the measurement less model dependent. Instead of using a single signal strength multiplier, μ, two multipliers are used. One is the signal strength of gluon fusion plus $\mathrm{t\bar{t}H}$ production, $\mu_{\mathrm{GF+t\bar{t}H}}$, and the other is the signal strength of VBF and VH production combined, $\mu_{\mathrm{VBF+VH}}$. $\mu_{\mathrm{GF+t\bar{t}H}}$ and $\mu_{\mathrm{VBF+VH}}$ are profiled with the nuisance parameters. In Figure 7.13 the resulting scans in 2NLLR show the 1σ and 2σ errors where 2NLLR=1

Table 7.2: Mass measurement systematic uncertainties. The magnitude of the uncertainty on the best fit mass broken down by each relevant systematic uncertainty. These numbers have been obtained by quadratic subtraction of the statistical uncertainty.

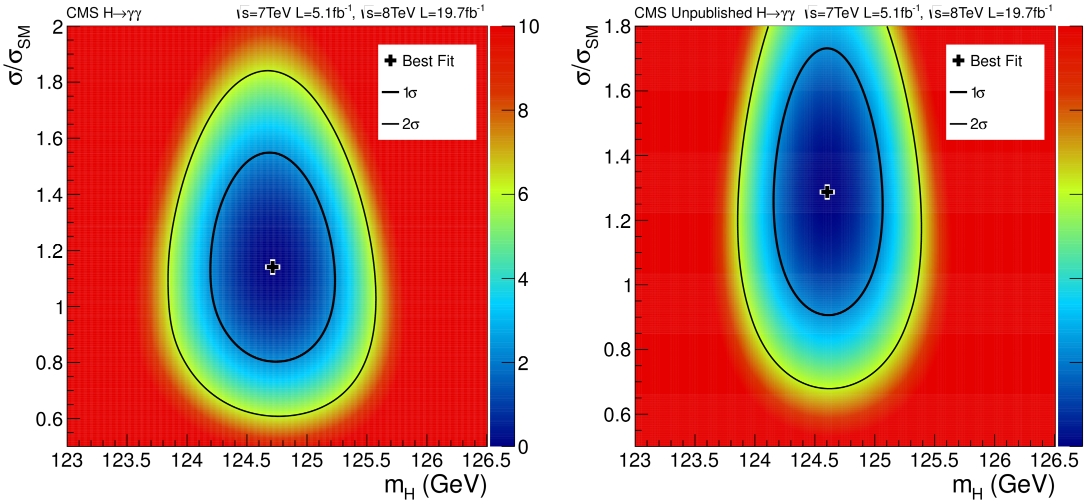
Source of uncertainty	**Uncertainty on $\widehat{\mathrm{M_H}}$ (GeV)**
Imperfect simulation of electron-photon differences	0.10
Linearity of energy scale	0.10
Energy scale calibration and resolution	0.04
Other	0.04
All systematic uncertainties on the signal model	0.15
Statistical	0.31
Total	**0.34**



(a) MVA (b) Cut-based

Figure 7.13: Mass measurement from H → $\gamma\gamma$ analyses. These are the 1-dimensional scans of $\mathrm{M_H}$ where $\mu_{\mathrm{GF+t\bar{t}H}}$ and $\mu_{\mathrm{VBF+VH}}$ are profiled parameters instead of μ alone.

and 2NLLR=4, respectively. The mass measurements are $124.70^{+0.35}_{0.34}$ GeV for the MultiVariate Analysis and 124.57 ± 0.30 GeV for the Cut-based Analysis.



(a) MVA

(b) Cut-based

Figure 7.14: Scan of M_H and μ with H $\rightarrow \gamma\gamma$ analyses. These are the 2-dimensional scans of M_H and μ in 2NLL.

Alternatively, the mass is measured as the minimum in 2NLLR in the 2-dimensional plane of M_H and μ, the overall signal strength multiplier. The masses at the minima in the MultiVariate and Cut-based Analyses are very similar to the results from the 1-dimensional scans. The results are shown in Figure 7.14.

7.6 Signal Strength by Channel and Production Mechanism

In the MultiVariate Analysis the signal strength is measured in various interesting sub-components. The signal strength of each event class is measured in Figure 7.15 and overlaid on the overall signal strength of the analysis. Numerous event classes are specialized for extracting signal events produced via specific production mechanisms. In Figure 7.16a the SM signal strength is measured in groups of event classes which intend to tag a specific production mechanism. The untagged channels are primarily for Higgs bosons produced via Gluon Fusion because in these events there are no additional taggable features in the leading-order Feynman diagrams.

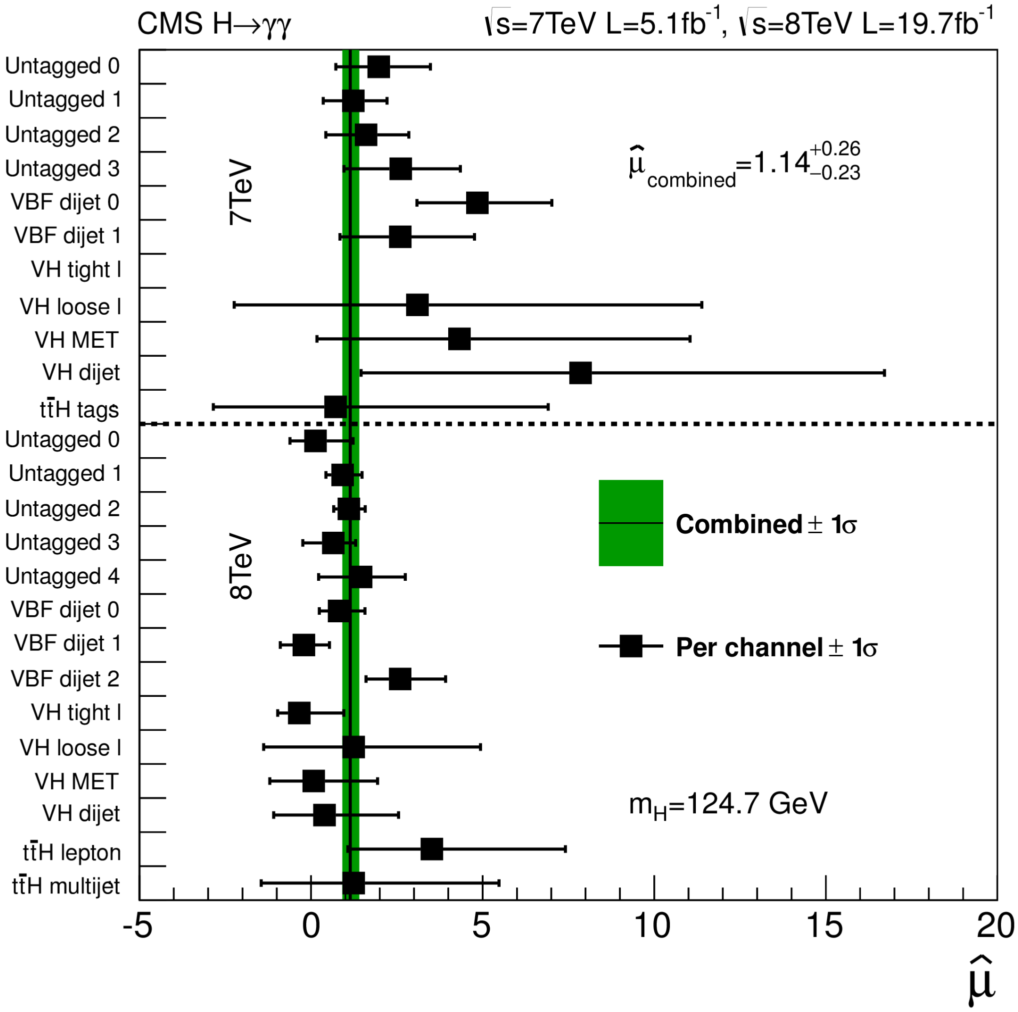
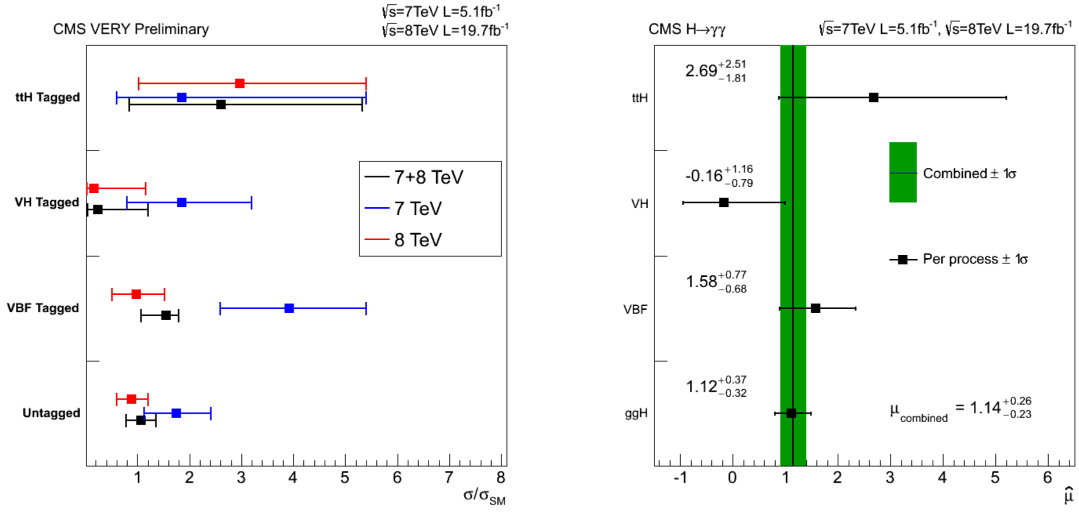


Figure 7.15: Signal strength per event class in the MultiVariate Analysis. The signal strengths of all 25 event classes at $M_H = 124.7$ GeV are marked with black squares with horizontal error bars. They lay on top of the overall signal strength drawn as a vertical black line with the error visualized as the green bar.

Given that these event classes are not 100% pure in signal selection (see Figure 6.16a), a more precise method for discerning the signal strength of each production mechanism is to allow all four production mechanisms to float and profile the mass with the nuisance parameters. This is shown in Figure 7.16b. The value of the profiled mass is $\hat{M}_H = 124.68$ GeV which is very similar to all other mass measurements.



(a) Topological scan by event classes (b) Four floating signal strengths

Figure 7.16: Topological and per production mechanism signal strength. In (a) $\sigma/\sigma_{\mathrm{SM}}$ is measured in each group of event classes targeting a production mechanism for 7 TeV in blue, 8 TeV in red and the combined dataset in black. In (b) a signal strength for each of the four production mechanisms is floated and fitted within the likelihood while the mass is profiled. (b) is the more fundamental measurement, but a high degree of correlation between (a) and (b)can be observed.

VH and VBF production channels both fundamentally probe the coupling of the Higgs Boson to the vector bosons, and so measuring their combined signal strength, $\mu_{\mathrm{VBF+VH}}$, with small uncertainty has been a primary goal of this analysis. Likewise, Gluon Fusion and $\mathrm{t\bar{t}H}$ production both probe the fermion coupling of the Higgs Boson, and the sensible combination of their signal strengths, $\mu_{\mathrm{GF+t\bar{t}H}}$, can be measured. These two complementary signal strengths are measured in two ways. One is to do a 2-dimensional scan of $\mu_{\mathrm{GF+t\bar{t}H}}$ and $\mu_{\mathrm{VBF+VH}}$ in the 2NLLR which in shown in Figure 7.17 where the minimal point in the scan is $(\mu_{\mathrm{GF+t\bar{t}H}}, \mu_{\mathrm{VBF+VH}}) = (1.08, 1.22)$.

Alternatively, $\mu_{\mathrm{GF+t\bar{t}H}}$ and $\mu_{\mathrm{VBF+VH}}$ can be measured independently by profiling one while measuring the other. Figure 7.18 shows like 1-dimensional scans of $\mu_{\mathrm{GF+t\bar{t}H}}$ with $\mu_{\mathrm{VBF+VH}}$ profiled and $\mu_{\mathrm{VBF+VH}}$ with $\mu_{\mathrm{GF+t\bar{t}H}}$ profiled. The minima with 1σ uncertainty (where 2NLLR $= 1$) are $\mu_{\mathrm{GF+t\bar{t}H}} = 1.13^{+0.37}_{-0.31}$ and

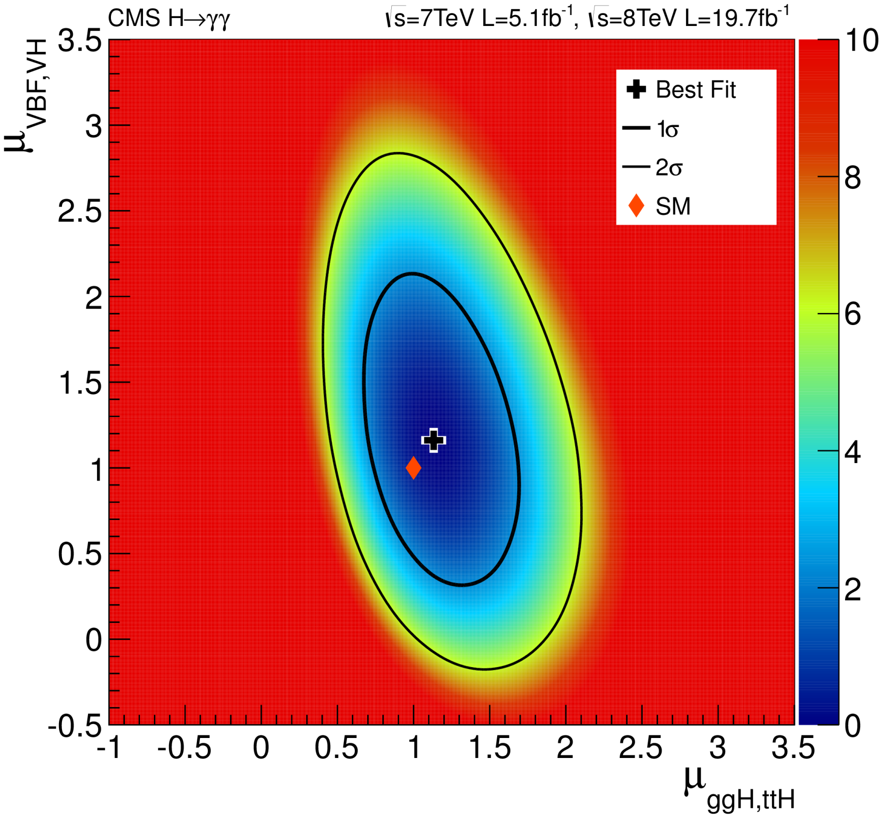
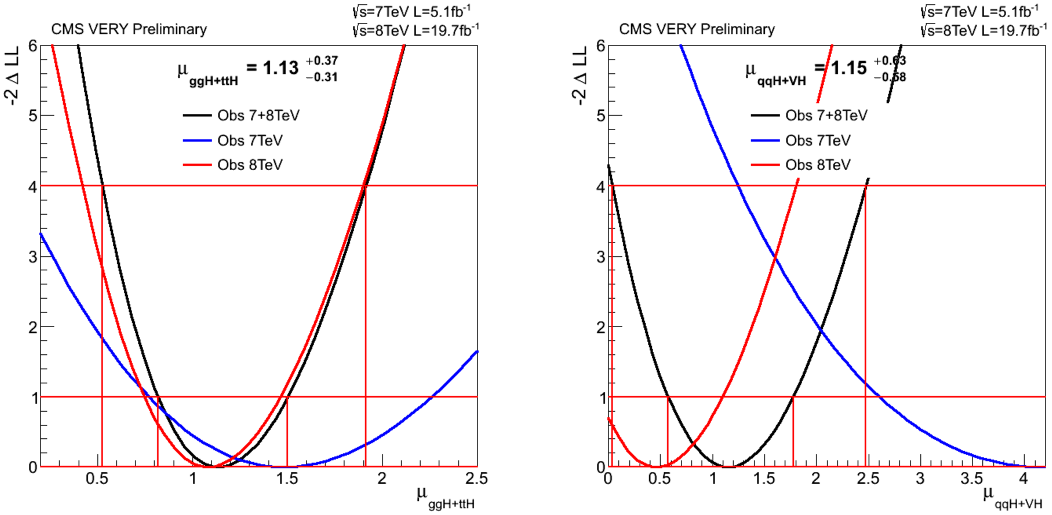


Figure 7.17: Likelihood scan of $(\mu_{\mathrm{GF+t\bar{t}H}}, \mu_{\mathrm{VBF+VH}})$. 1σ and 2σ curves are black solid and dashed curves respectively. The minimal point in the scan is $(\mu_{\mathrm{GF+t\bar{t}H}}, \mu_{\mathrm{VBF+VH}}) = (1.08, 1.22)$.

$$\mu_{\mathrm{VBF+VH}} = 1.15^{+0.63}_{-0.58}$$



(a) $\mu_{\mathrm{GF+t\bar{t}H}}$ (b) $\mu_{\mathrm{VBF+VH}}$

Figure 7.18: One dimensional scans of $\mu_{\mathrm{GF+t\bar{t}H}}$ and $\mu_{\mathrm{VBF+VH}}$. These are the 2NLLR scans of $\mu_{\mathrm{GF+t\bar{t}H}}$ with $\mu_{\mathrm{VBF+VH}}$ profiled (a) and $\mu_{\mathrm{VBF+VH}}$ with $\mu_{\mathrm{GF+t\bar{t}H}}$ profiled (b). The results with only the 7 TeV dataset are in blue, with only the 8 TeV dataset in red, and the combined dataset in black.

7.7 Coupling Measurements

The MultiVariate Analysis can also be used to measure couplings of the Higgs boson. One fundamental parameterization of the Higgs boson's coupling is to use a single multiplier for all SM fermionic couplings, κ_{f}, and to likewise use a single multiplier, κ_{V}, for both SM vector boson couplings to the Higgs boson. The $\mathrm{H} \rightarrow \gamma\gamma$ channel is special in that its decay involves W loops and charged fermion loops (see Figure 4.4). These Feynman diagrams result in opposite signed couplings in the decay width of $\mathrm{H} \rightarrow \gamma\gamma$. This leads to the apparent correlation of κ_{f} and κ_{V} in the 2-dimensional scan shown in Figure 7.19.

Processes with large contributions from loop diagrams can be sensitive to new physics. That is, if an unobserved particle can mediate the loop as well, then there could be significant deviation from the expected cross section or decay width when measured. With this motivation in mind effective Higgs boson coupling to

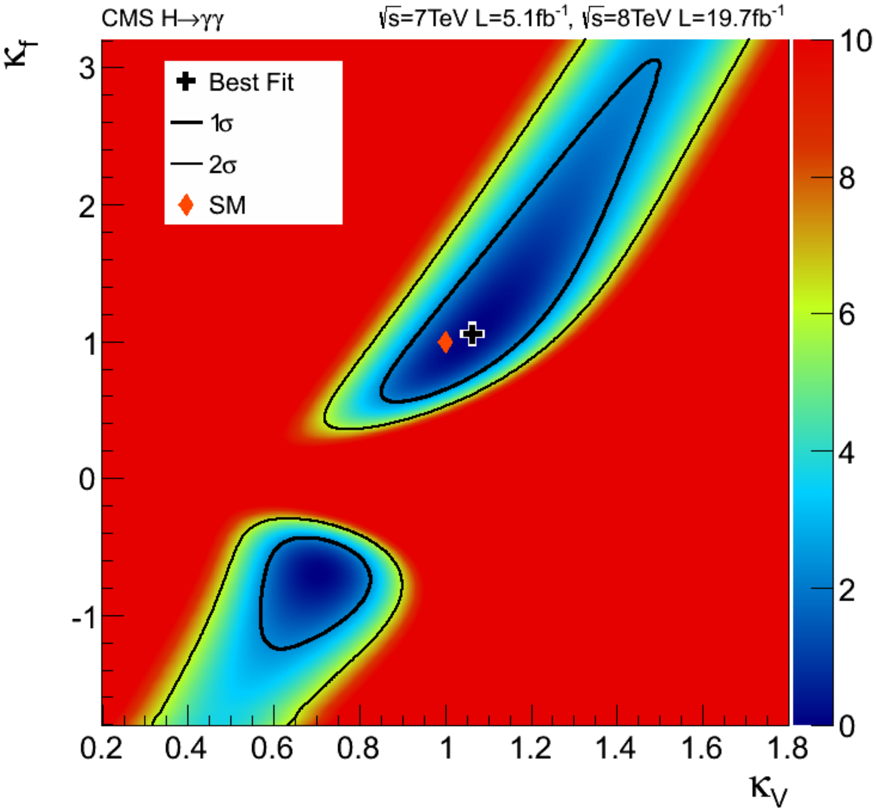


Figure 7.19: Two dimensional scan of κ_f,κ_V couplings with MVA. The coupling modiers of the Higgs boson to vector bosons (κ_V) and to fermions (κ_f) are shown above. This scan is performed with $\mathrm{M_H}$ = 124.7 GeV. Profiling the mass results in an extremely similar result.

gluons via the quark loop in Gluon Fusion production (see Figure 4.2a), κ_{g}, and effective Higgs boson coupling to photons, κ_γ, via the H $\rightarrow \gamma\gamma$ decay processes (see Figure 4.4) compose an interesting parameterization of the Higgs boson's couplings. All other Higgs boson couplings are fixed to their SM values for the 2-dimensional scan of κ_{g},κ_γ shown in Figure 7.20. Given that the cross section times branching ratio of the H $\rightarrow \gamma\gamma$ process produced via Gluon Fusion (the dominant SM production) is proportional to $\kappa_{\mathrm{g}}^2\kappa_\gamma^2$, one should expect the inverse behavior of κ_{g} and κ_γ in the minimum of the scan.

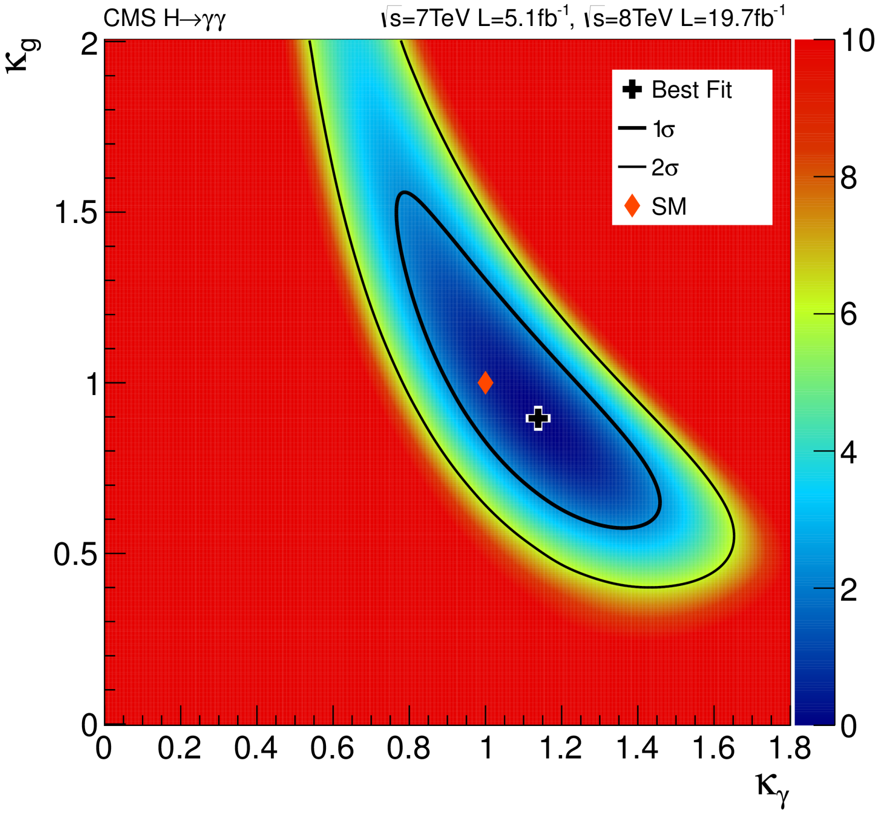


Figure 7.20: Two dimensional scan of κ_{g},κ_γ couplings with MVA. Effective couplings of the Higgs boson to the Gluon Fusion quark loop (κ_{g}) and to the charged fermion/W loop in H $\rightarrow \gamma\gamma$ decay (κ_γ) are scanned. Other tree level couplings are assumed to be 1. These scans are performed with $M_{\mathrm{H}} = 124.7$ GeV. Profiling the mass results in an extremely similar result.

7.8 Decay Width

It is possible to set a limit on the width of the observed signal, albeit a limit far in excess of the SM expectation. To accommodate the natural width of the Higgs Boson the signal width is assumed to be negligible as compared to the detector resolution. The signal resolution is replaced by an analytic convolution of a Breit-Wigner distribution (modeling a non-zero decay width) with a Gaussian distribution (modeling the detector resolution).

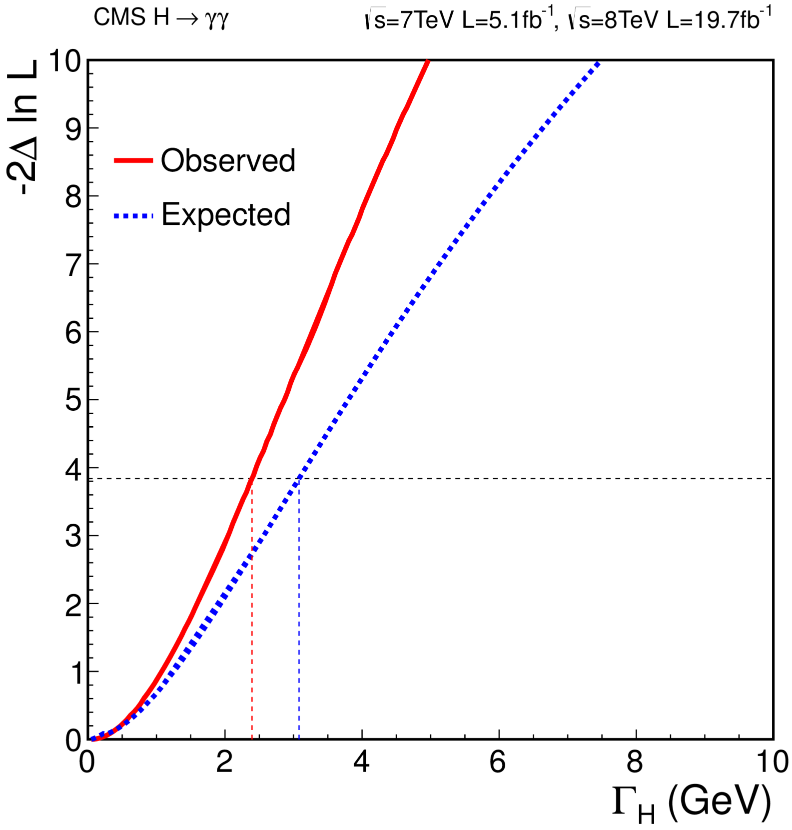
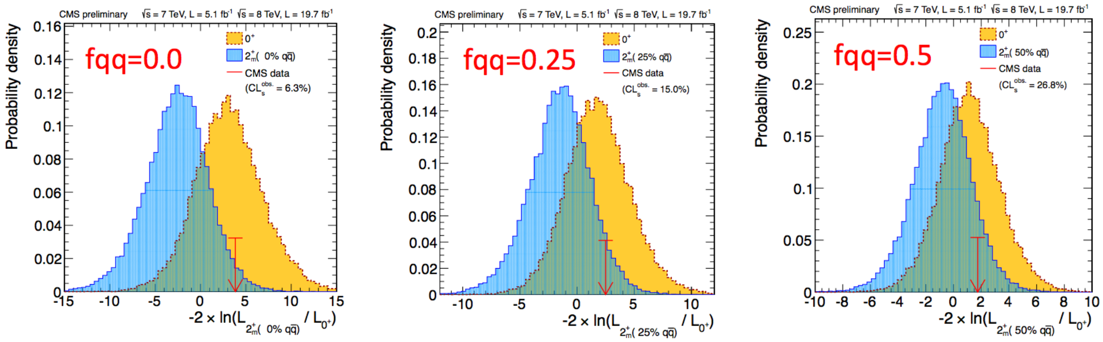


Figure 7.21: Higgs decay width scan and limit. Scan of the negative-log-likelihood as a function of the Higgs decay width. The observed (expected) upper limit on the width is found to be 2.4 (3.1) GeV at the 95% confidence level.

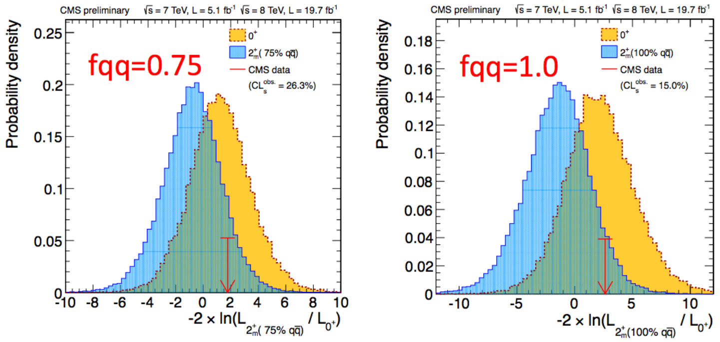
Figure 7.21 shows a scan of 2NLLR with profiled mass as a function of the observed states decay width for the combined 7 and 8 TeV dataset. The upper limit on the width is calculated using $\mathrm{CL_S}$ method described above. The observed (expected) upper limit is found to be 2.4 (3.1) GeV at 95% confidence level.

7.9 Spin Compatibility

Quantum selection rules derived by Landau [83] and Yang [84] disallow odd spin particle decay to two photons. Therefore, this Higgs Boson with excitation in the di-photon channel cannot be a spin 1 particle. However, an analysis has been developed to compare the compatibility of our data with the Standard Model's spin 0^+ (where + denotes parity) and with the spin 2^+_m hypothesis of the minimally coupling graviton [85]. The graviton is allowed to be produced by Gluon Fusion (GF) or by quark-antiquark annihilation ($q\bar{q}$). The fraction of $q\bar{q}$, $f_{q\bar{q}}$, is taken as unknown and varied from 0-100%.



(a) $f_{q\bar{q}} = 0.00$ (b) $f_{q\bar{q}} = 0.25$ (c) $f_{q\bar{q}} = 0.50$



(d) $f_{q\bar{q}} = 0.75$ (e) $f_{q\bar{q}} = 1.00$

Figure 7.22: SM versus spin-2 pseudo-data toys. The distribution of the test statistic for pseudo-data toys generated under the Standard Model hypothesis (orange) and the graviton-like spin-2 hypotheses (blue) with the fraction of quark-antiquark production ($f_{q\bar{q}}$) out of gluon fusion plus quark-antiquark production varied from from 0.00 to 1.00 in steps of 0.25. The observed value in the data is shown as the red arrow.

The distribution of the cosine of the scattering angle in the Collins-Soper frame is highly discriminating among different signal models [86]. The Collins-Soper frame defines the central axis to be that which bisects the acute angle between the proton beam directions in the rest frame of the di-photon. The scattering angle is between the Collins-Soper axis and either photon. In terms of the photons' energies and momenta that is:

$$\cos(\theta^*_{\mathrm{CS}}) = 2 \times \frac{E_2 p_{z1} - E_1 p_{z2}}{m_{\gamma\gamma}\sqrt{(m_{\gamma\gamma})^2 + (P_{T,\gamma\gamma})^2}} \tag{7.9.0.15}$$

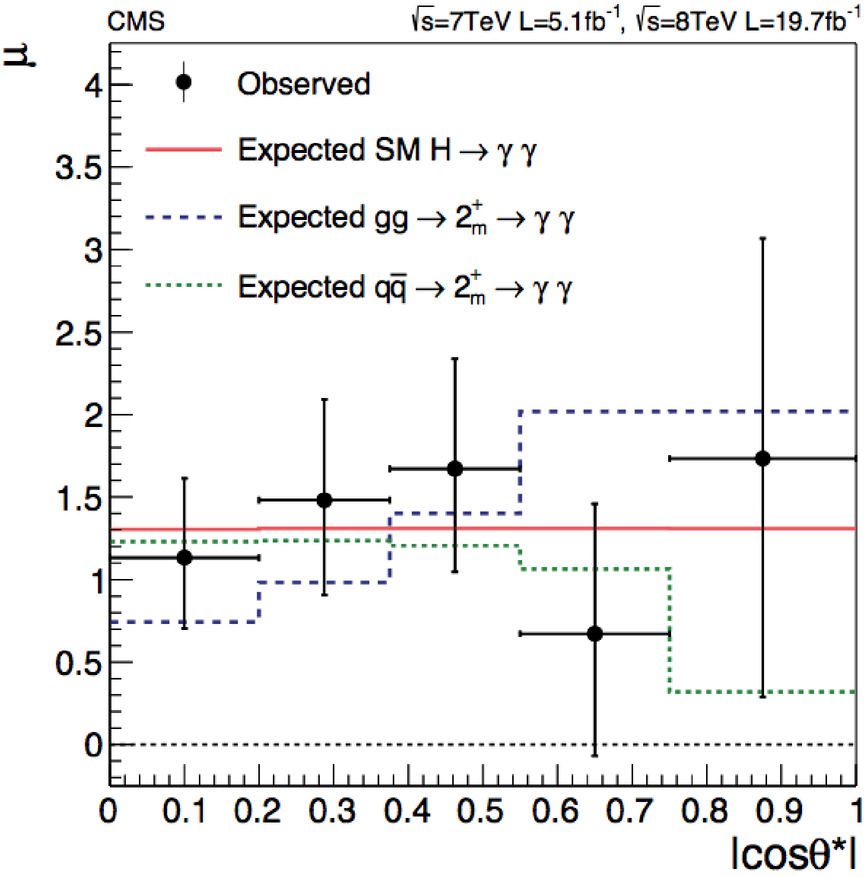


Figure 7.23: SM and various spin-2 model yields in $\cos(\theta^*_{\mathrm{CS}})$ bins. The Standard Model extracted signal yield as a function of $\cos(\theta^*_{\mathrm{CS}})$ for spin-0 expectation (red line), spin-2 expectation with gluon-fusion production only (blue line), the spin-2 expectation with quark-antiquark annihilation production only (green line), the spin-2 expectation with $f_{q\bar{q}} = 0.50$ (magenta line) and the observation (black points).

This analysis is performed with the cut-based photon ID and event classification is performed in the same inclusive η-R_9 di-photon classes. In addition, within each η-R_9 bin there are five optimized bins in $\cos(\theta^*_{\mathrm{CS}})$ for a total of 20 event classes for the 7 TeV and 8 TeV datasets. The background models are polynomials and the signal models are derived in the same fashion as the primary analysis.

Pseudo-data toy experiments are simulated for the Standard Model hypothesis and the graviton hypothesis with various $f_{q\bar{q}}$ (0.00, 0.25, 0.50, 0.75 and 1.00). Twice the negative logarithm of the likelihood ratio (2NLLR) of the two hypotheses $(-2\ln\frac{\mathcal{L}_{2^+_m+\mathrm{Bkg}}}{L_{0^++\mathrm{Bkg}}})$ is evaluated for all of these toys. Then 2NLLR is evaluated for each $f_{q\bar{q}}$ assumption and the $\mathrm{CL_S}$ probability is evaluated for exclusion.

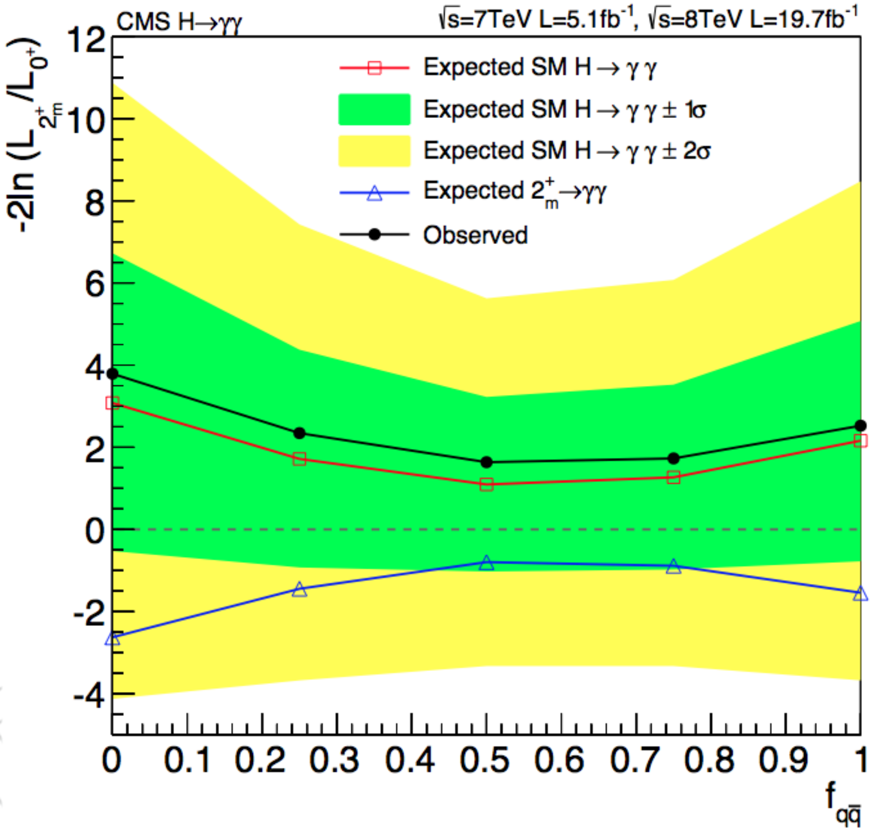


Figure 7.24: H → $\gamma\gamma$ spin compatibility results. The distribution of the test statistic for pseudo experiments thrown under the Standard Model, spin-0 hypothesis (red) and the *graviton-like*, spin-2 hypothesis (blue) as a function of $f_{q\bar{q}}$ production relative to gluon fusion production. The observed distribution in the data is shown by the black points.

Figure 7.22 shows the distribution of 2NLLR for Standard Model versus 2^+_m

toys with observations in red. The observed Standard Model signal strength in bins of $\cos(\theta^*_{\mathrm{CS}})$ is shown in Figure 7.23. Given that the observed signal strength in the first three $\cos(\theta^*_{\mathrm{CS}})$ bins are all more SM-like than in all $f_{q\bar{q}}$ the results as a function of $f_{q\bar{q}}$ are expected to be correlated. Figure 7.24 shows the $\mathrm{CL_S}$ separation between the 2^+_m and 0^+ models as a function of $f_{q\bar{q}}$ and the results appear more compatible with the Standard Model regardless of $f_{q\bar{q}}$.

Table 7.3: $\mathrm{CL_S}$ exclusion confidence level spin-2 models with various values of $f_{q\bar{q}}$.

$f_{q\bar{q}}$	Prob(SM>Obs)	Prob(ALT>obs)	$\mathrm{CL_s}$	Exclusion Confidence Level
0.00	0.409	0.026	0.064	0.936
0.25	0.394	0.059	0.151	0.849
0.50	0.382	0.102	0.267	0.733
0.75	0.397	0.104	0.261	0.739
1.00	0.429	0.064	0.150	0.850

The $\mathrm{CL_S}$ separation at each point is detailed in Table 7.3. The exclusion confidence level is nearly 94% for the purely gluon fusion assumption and 85% for the purely quark-antiquark fusion assumption. No exclusions are possible but the Standard Model spin-0 assumption is favored by the observations for all $f_{q\bar{q}}$.

7.10 Second Higgs Boson Search

To search for a possible additional Higgs-boson-like state in the mass range $110 - 150$ GeV, the observed signal around 125 GeV is added to the background model and its mass and signal strength are allowed to vary in the fit. An additional, independent, signal mode is introduced as a second Higgs boson, for which the exclusion limits are calculated using the same approach of Section 7.3. The resulting exclusion limit is shown in Figure 7.25. Once sufficiently away from the already observed Higgs boson at 124.7 GeV, the same limit is obtained as when searching for a single SM Higgs boson. The shading indicates a window with a width of 10 GeV, centred at the best-fit mass, where the expected sensitivity to a second Higgs boson is severely degraded due to the presence of the already observed state.

There are only two small regions (111 – 112 GeV and near 150 GeV) that are not excluded at the 95% confidence level in the search range and away from the mass of the Higgs boson.

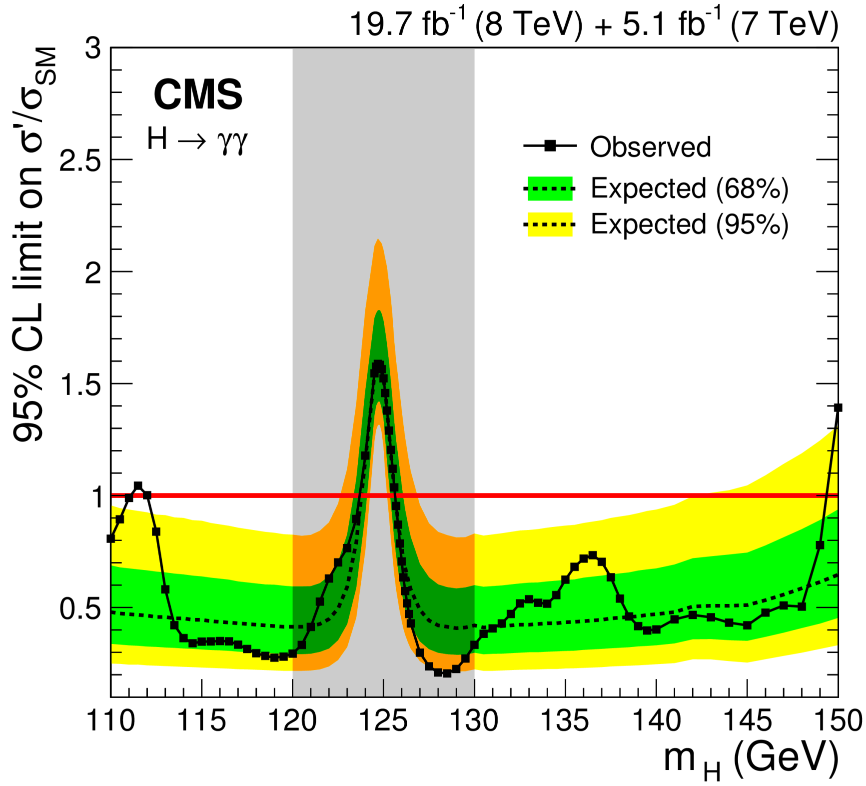
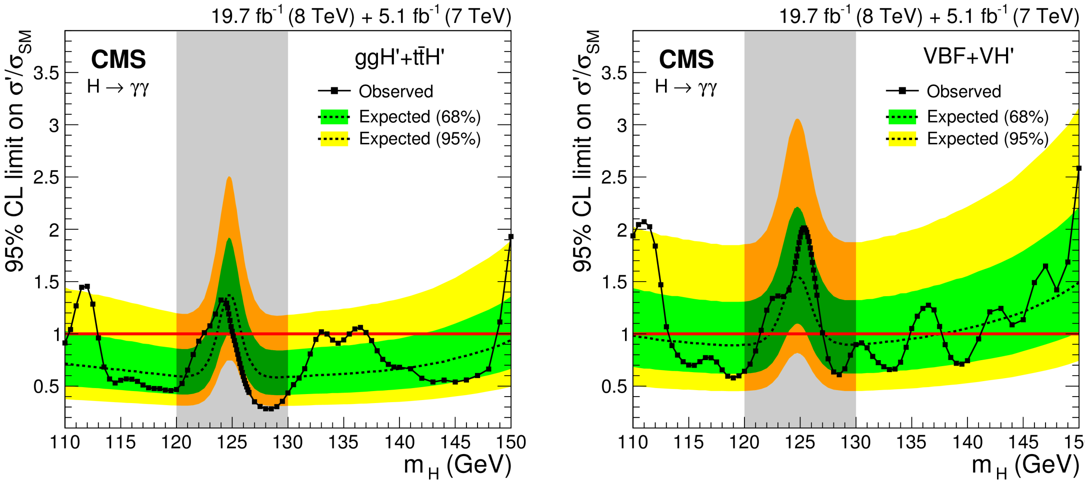


Figure 7.25: Exclusion limit on the signal strength, $\frac{\sigma}{\sigma_{\text{SM}}}$, for a second SM Higgs boson. The observed state at 125 GeV is included as part of the background for this search. The shading indicates a window with a width of 10 GeV, centred at the best-fit mass, where the expected sensitivity to a second Higgs boson is severely degraded due to the presence of the already observed state. In most of the search range a second Higgs boson with SM signal strength is excluded at the 95% confidence level. However, an excess of events beyond the search window leads to the weak exclusion limits near 150 GeV and in the range of 111 – 112 GeV the 95% confidence level exclusion is just above the SM cross section times branching ratio.

A further particular case of interest is when the second state is produced only via couplings to fermions or only via couplings to vector bosons. Figure 7.26 shows the exclusion limits obtained when the observed signal near 125 GeV is added to the background model and its mass and signal strength are allowed to vary in the t, and an additional state produced (Figure 7.26a) only by the Gluon Fusion and $t\bar{t}H$ processes, or (Figure 7.26b) only by the VBF and VH processes. The limits are given in terms of the SM cross section times branching fraction for those processes. Even for the VBF and VH processes, which have lower cross sections, an additional state with SM-like signal strength is excluded or disfavoured over much of the mass range.



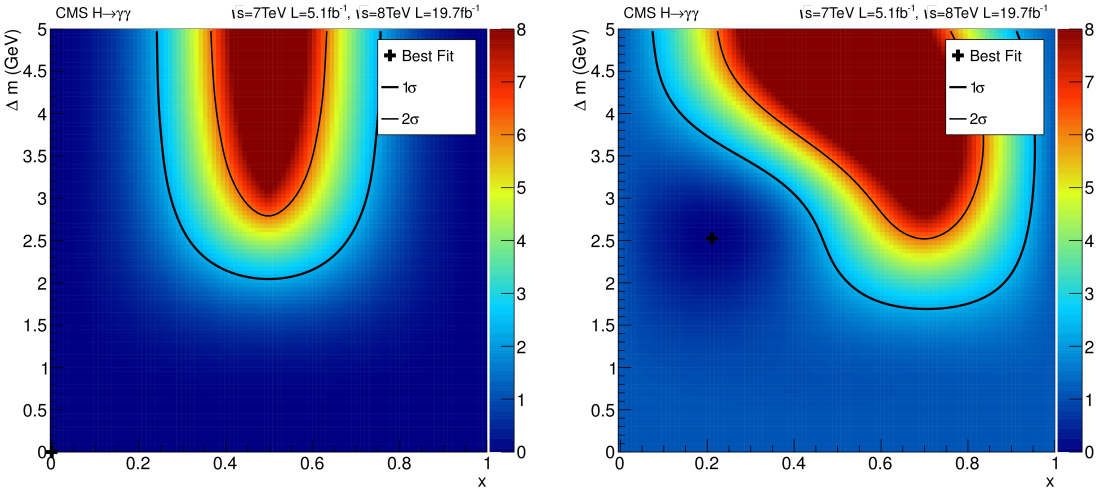
(a) GF + $t\bar{t}H$ (b) VBF + VH

Figure 7.26: Search for a second SM Higgs boson produced only via Gluon Fusion+$t\bar{t}H$ and only via VBF+VH. Exclusion limit on the signal strength, $\frac{\sigma}{\sigma_{\mathrm{SM}}}$, for a second SM Higgs boson produced only via Gluon Fusion+$t\bar{t}H$, (a), and only via VBF+VH, (b). The observed state at 125 GeV is included as part of the background for these searches. The shading indicates a window with a width of 10 GeV, centred at the best-t mass, where the expected sensitivity to a second Higgs boson is severely degraded due to the presence of the already observed state.

7.11 Degenerate Higgs Boson Search

The shaded regions in Figure 7.25 and Figure 7.26, where the expected sensitivity to a second Higgs boson is severely degraded due to the presence of the already observed state, are probed by a dedicated search using the high resolution of the diphoton channel to provide sensitivity to a pair of states separated by only a few GeV. The signal model is re-parameterized with two signals, having masses

$\mathrm{M_H}$ and $\mathrm{M_H} + \Delta\mathrm{m}$. The relative strengths of the two signals, parameterized by the variable x, are allowed to vary such that the two signals are modulated by μx and $\mu(1-x)$ respectively. μ is the total signal strength and x is the fraction of signal contained in the state lower in mass. A two-dimensional scan of $\Delta\mathrm{m}$ and x is obtained, while allowing both $\mathrm{M_H}$ and μ to vary as free parameters in the fit. Figure 7.27 shows the expected (upper plot) and observed (lower plot) 2NLL in the (x, $\Delta\mathrm{m}$) plane. Sensitivity is expected in regions where $\Delta\mathrm{m}$ is close to or greater than the experimental mass resolution and where the two signal strengths are similar. The black cross shows the best-fit value, and the lines correspond to the 1σ and 2σ uncertainty contours for the SM (i.e. a single state). It can be seen that a region of the parameter space is disfavoured at more than 2σ: where the ratio of the signal strengths is between 0.2 and 0.8 and the mass difference is greater than values ranging between 2.5 and 4 GeV depending on the ratio of the signal strengths. The somewhat asymmetrical shape of the excluded region and the position of the best-fit value, are a reflection of the slightly asymmetrical mass peak seen in Figure 7.9c, also reflected in the figures showing the local p-value, and exclusion limit as a function of $\mathrm{M_H}$.



(a) Expected

(b) Observed

Figure 7.27: Expected and observed results of two, near mass-degenerate Higgs bosons search. Above are the maps of the values of twice the negative logrithm of the likelihood ratio, $q\,(x, \Delta\mathrm{m})$, for the near mass degenerate states parametereized by x (the fraction of signal in the lower mass state) and Δm (the mass difference between the states). The black cross shows the best-fit value, and the lines correspond to the 1σ and 2σ uncertainty contours for the SM (signal state) expection in Figure 7.27a and the observation in Figure 7.27b. The observed minimum is at (0.2, 2.5 GeV).

Chapter 7, in full, has been prepared for publication submission to EPJC: "Observation of the diphoton decay of the Higgs boson and measurement of its properties". The dissertation author was the primary investigator and author of this paper.

Chapter 8

Searches for the Higgs boson near 125 GeV Decaying to Two Muons and Two Electrons

With the discovery of a new boson consistent with a SM Higgs boson, the measurement of its properties are of paramount importance. The expected SM decay branching ratio of a Higgs boson into muon (electron) pairs is small (tiny) as the SM predicts the couplings of a SM Higgs boson to fermions to scale proportionally to the lepton mass. Therefore, the branching ratio is proportional to the mass squared.

While the search in the $\tau^+\tau^-$ decay channel is already sensitive to the SM cross section, the SM $\mu^+\mu^-$ channel is not expected to be sensitive with the current integrated luminosity. The branching ratio into e^+e^- is expected to be many orders of magnitude smaller, which means an observation of this decay would require the new boson to have an interpretation beyond the SM. At the same time, the high trigger efficiency, high lepton identification efficiency and very good mass resolution of muons and electrons make the e^+e^- and $\mu^+\mu^-$ analyses much more sensitive than the $\tau^+\tau^-$ for equal cross section times branching ratio. Observed branching ratios in $\mu^+\mu^-$ and e^+e^- channels much smaller than that expected for $H \to \tau^+\tau^-$ is evidence for lepton-non-universality. This is an expected feature of a SM Higgs boson (i.e. this adds to the evidence that the new particle observed is the Standard

Model Higgs boson).

Many features of these analyses are very similar to the H → $\gamma\gamma$ analysis previously described. Therefore, the descriptions to follow are in some cases very brief and not described in detail.

.

8.1 Analysis Strategy

The search for the Higgs boson decaying to e^+e^- and $\mu^+\mu^-$ is very similar to the H → $\gamma\gamma$ search. A narrow mass peak on a smooth background defines this search strategy as well. The background in these searches is dominated by the irreducible Drell-Yan tail above the Z boson resonance. The remaining background originates mainly from $t\bar{t}$ and WW. All reducible backgrounds with at least one fake or non-isolated lepton are highly reduced by lepton identification cuts.

To improve the search sensitivity different event classes were used. The first event classes are defined with two di-jet tag classes of H → $\gamma\gamma$'s Cut-based Analysis (described in Section 6.2.3). These event classes aim to select events Higgs produced via the Vector Boson Fusion (VBF). The remaining events (those which pass neither of the di-jet tag selections) are put into two inclusive event classes depending on whether both selected leptons were reconstructed in the central detector region or not. The boundary in $|\eta|$ between barrel and endcaps is different between electrons and muons, which reflects the geometry of the ECAL, tracker and muon detectors.

8.2 Datasets and Monte Carlo

For the e^+e^- ($\mu^+\mu^-$) analysis 19.6 (19.7) fb^{-1} of 8 TeV data are used. These correspond to the total luminosity of the 2012 dataset.

Signal MC has been generated by the author privately [1] to produce these

---

[1] Most MC is produced centrally in CMS, but for this specialized analysis no samples were available. The author had to make necessary modifications to existing files in order to produce appropriate H → e^+e^- and H → $\mu^+\mu^-$ samples.

results using Powheg [88,89] interfaced with Pythia [66]. The average pile-up of the generated samples is 22.7 interactions per event. Samples have been generated for $\mathrm{M_H} = 120$, 125, 130, 135, and 140 GeV. The branching ratios used for $\mathrm{H} \rightarrow \mu^+\mu^-$ signal are taken from LHC's Higgs cross section working group [76]. The $\mathrm{H} \rightarrow \mathrm{e}^+\mathrm{e}^-$ cross sections are derived by scaling the $\mathrm{H} \rightarrow \tau^+\tau^-$ branching ratios by $\left(\frac{m_e}{m_\tau}\right)^2$. Higgs produced via $\mathrm{t\bar{t}H}$ production were not included for this analysis.

Background MC is used only to compare different lepton identification working points and to choose optimally amongst them. Like the $\mathrm{H} \rightarrow \gamma\gamma$ analysis the background MC is not used in the final statistical analysis but rather fits to the data are used to make background estimates. Drell-Yan MC was produced using Madgraph at next to leading order in α_s. All other samples are produced with Pythia 6.4 at leading order.

8.3 Trigger

Two types of triggers are used for both electrons and muons. The first type selects events in which one lepton is very well identified and passes a p_T threshold. This threshold varies as a function of data taking period with higher thresholds used as the instantaneous luminosity. The thresholds increased mainly during early 2012 after which the instantaneous luminosity was fairly consistent. For electrons this threshold varied from 17 to 27 GeV and for muons from 17 to 24 GeV. The other trigger strategy requires two leptons (of the same flavor) with asymmetric p_T thresholds of 17 and 8 GeV. The identification requirements are less strict for these di-lepton triggers.

8.4 Lepton Selection

Electrons are required to be inside the ECAL fiducial region $|\eta| < 1.4442$ (barrel) or $1.566 < |\eta| < 2.5$ (endcaps). Muons are divided into central barrel, $|\eta| < 0.8$, and beyond, $0.8 < |\eta| < 2.4$, where Drift Tube coverage is only partial.

All leptons are required to have $p_T > 25$ GeV. Two leptons of the same type and with opposite charge are selected. Cut-based working points derived by CMS's dedicated Physics Object Group (POG) for electron and photons [90] and for muons [91] are used in the final selection. Only lepton pairs with invariant mass between 110-180 GeV are relevant for this analysis as this is the mass range which is fit for background estimates.

8.5 Energy/Momentum Corrections

The electron energy is determined using an energy regression originally intended for the H $\rightarrow$ ZZ $\rightarrow 4\ell$ analysis but is still very satisfactory for this purpose [92]. Muon momentum is corrected using standard CMS corrections to eliminate dependence on ϕ which results from relative misalignment between the tracker and muon chambers. Z $\rightarrow \mu^+\mu^-$ events are used in data and MC to check the agreement after corrections are applied. Corrections for jets are the same as for H $\rightarrow \gamma\gamma$.

8.6 Classification

In each analysis the same classification scheme is used. After requiring a di-lepton pair to pass all of the identification, isolation, p_T, and invariant mass selections, the event is checked for a di-jet pair matching the tight or loose criteria in Section 6.2.3. The jet pile-up identification pre-selection is applied to jets as well.

If the event does not pass the di-jet selection, then the events are classified based on the location of the leptons in the detector. One class contains events with both leptons in the barrel ($|\eta| < 1.4442$ for electrons, $|\eta| < 0.8$ for muons), while the other contains events with at least one of the selected leptons in the endcaps.

8.7 Background Modeling

Several background models for the fit to the data were proposed. Given that the dominant background for $\mathrm{H} \rightarrow \mu^+\mu^-$ and $\mathrm{H} \rightarrow \mathrm{e}^+\mathrm{e}^-$ is irreducible Drell-Yan, several physically-motivated models (mainly numerous variants of Breit-Wigner distributions) were attempted. The final model used to derive this preliminary result is an exponential function divided by a second-order polynomial:

$$\frac{\exp\left(\mathrm{p}_1 * \mathrm{M}_{\ell\ell}\right)}{\left(\mathrm{M}_{\ell\ell} - \mathrm{p}_2\right)^2} \tag{8.7.0.1}$$

This hybrid function is an amalgamation of the exponential shape of Drell-Yan continuum and the Breit-Wigner shape of the Z resonance (in the denominator).

After this result was made public, it was found that there is significant potential bias in the search range because of the few number of parameters in the fit function and the high number of the selected events. An upcoming publication features an updated version of these results, which uses a modeling strategy that greatly reduces the potential bias. The impact of the new background treatment increases the observed and expected $\sigma\times\mathrm{BR}(\mathrm{H} \rightarrow \ell\ell)$ 95% condifence level limits by a factor of 2-4 (i.e. the sensitivity is decreased). However, the limits are still strict enough to uphold the primary results quoted below. That is, universal lepton decay of the Higgs boson is still excluded.

8.8 Systematics

Systematic uncertainties have been determined for a Higgs boson signal with a mass of 125 GeV at $\sqrt{s} = 8$ TeV. The efficiency scale factors related to lepton identification, isolation, and trigger have been applied to each signal MC sample to correct discrepancies between data and MC. In order to evaluate the systematic uncertainty related to the extraction of the efficiencies, the scale factors are increased or decreased by ±0.5%. This systematic uncertainty is associated with the efficiency of the signal model and found to be of the order of 2.5%.

Jet energy scale, resolution and MC choice of jet simulation parameters impact di-jet tagging efficiency. These factors have been evaluated and systematics of the order of 30% and 10% are applied to Gluon Fusion and VBF produced Higgs samples.

The uncertainty in the amount of total luminosity was 4.4%[2] for this analysis. The uncertainty on the amount of pile-up is evaluated by varying the total inelastic cross-section by ±5%. The total uncertainty on signal yields is found to be sub-percent.

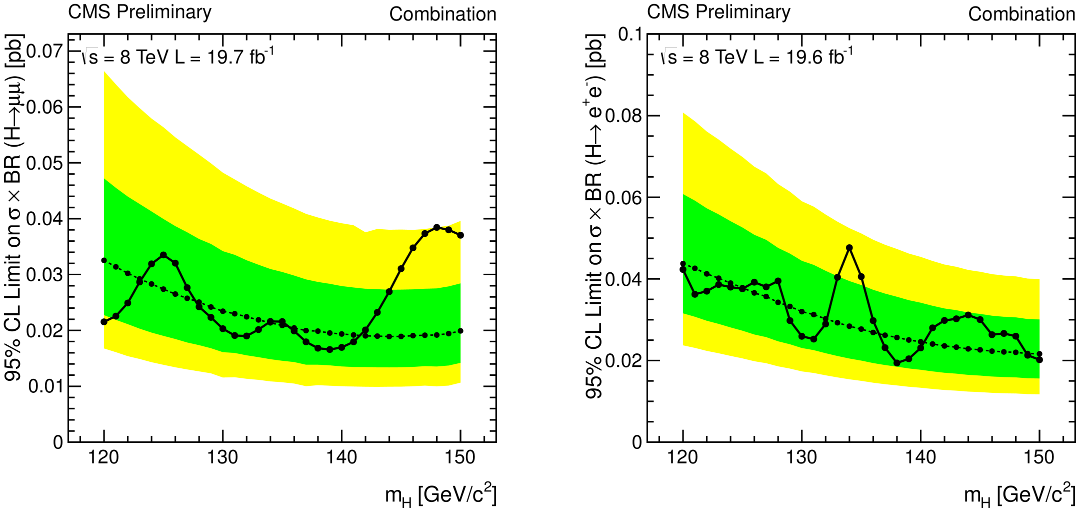
8.9 Results and Discussion

No significant excess above the background has been found in either channel and therefore 95% confidence level $\mathrm{CL_S}$ exclusion limits are set on the $\sigma \times \mathrm{BR}$ in each channel. Figure 8.1 shows these limits in the search range of 120 to 150 GeV. At 125 GeV, $\sigma \times \mathrm{BR}\,(\mathrm{H} \rightarrow \mathrm{e^+e^-}) < 0.038$ pb at 95% confidence level, which is precisely the median expected limit in the background-only assumption. The corresponding limit for $\mathrm{H} \rightarrow \mu^+\mu^-$ is $\sigma \times \mathrm{BR}\,(\mathrm{H} \rightarrow \mu^+\mu^-) < 0.034$ pb whereas the expected limit is 0.027 pb.

The best fit to the data in the $\mathrm{H} \rightarrow \tau^+\tau^-$ channel is $\sigma_{SM} \times \mathrm{BR}\,(\mathrm{H} \rightarrow \tau^+\tau^-) =$ 1.40 pb at 8 TeV. CMS's $\mathrm{H} \rightarrow \tau^+\tau^-$ analysis is sensitive to $0.55\frac{\sigma}{\sigma_{SM}}$ [93] which corresponds roughly to an expected limit in $\sigma \times \mathrm{BR}\,(\mathrm{H} \rightarrow \tau^+\tau^-)$ of about 0.8 pb. Thus, this analysis provides an exclusion limit far beyond the expected sensitivity of the $\mathrm{H} \rightarrow \tau^+\tau^-$ analysis. In addition, given CMS's $\mathrm{H} \rightarrow \tau^+\tau^-$ result showing 3.2σ evidence of the Higgs boson at 125 GeV with a signal strength of 0.78 ± 0.27 corresponding to $\sigma \times \mathrm{BR}\,(\mathrm{H} \rightarrow \tau^+\tau^-) = 1.09$ pb at 8 TeV, these results favor the interpretation of the boson at 125 GeV to be a Higgs boson and strongly rule out the possibility of universal lepton decay at a confidence level much greater than 95%.

---

[2] This result is from an early public result. Since that time CMS's luminosity measurement has improved considerably.



(a) $\mu^+\mu^-$ (b) $\mathrm{e}^+\mathrm{e}^-$

Figure 8.1: $\sigma_{SM} \times$ BR limits for $\mathrm{H} \to \mu^+\mu^-$ and $\mathrm{H} \to \mathrm{e}^+\mathrm{e}^-$ searches, which are the primary results of the first searches for the Higgs boson at 125 GeV decaying to $\mu^+\mu^-$ and $\mathrm{e}^+\mathrm{e}^-$ are shown. These limits are very strict in comparison to the expected $\sigma_{SM} \times \mathrm{BR}\,(\mathrm{H} \to \tau^+\tau^-)$ of $O\,(\mathrm{pb})$.

Chapter 8 is a reprint of the material as it appears in CMS Collaboration, “Search for the standard model Higgs boson in the dimuon decay channel in *pp* collisions at $\sqrt{s} = 7$ and 8 TeV”, *CMS Physics Analysis Summary* CMS-HIG-PAS-13-007 (2013). The dissertation author was the primary investigator and author of this paper.

Chapter 9

Further Results of the Higgs Boson at 125 GeV

CMS and ATLAS have both utilized LHC's Run 1 dataset to make strong statements about the existence and properties of what appears to be a SM Higgs boson with $\mathrm{M_H} \sim 125$ GeV. A summary of relevant results and measurements of the Higgs Boson at 125 GeV will follow with the CMS dataset. ATLAS has also produced similar results including a recent combination of all channels with full Run 1 luminosity [94], which will however not be discussed in detail here.

9.1 CMS $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ in Run 1

The $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ channel is an important tool in determining many of the properties of the Higgs boson at 125 GeV. Foremost, the low-background, high-resolution nature of this channel makes it extremely important in determining the existence and mass of such a state. The invariant mass of the four-body system (4μ, $2\mu 2\mathrm{e}$, or 4e) for selected events in CMS's Run 1 final analysis is shown in Figure 9.1 for data and MC including a simulated Higgs boson with $\mathrm{M_H} = 126$ GeV [95]. The visual agreement between data and prediction is striking, and indeed utilizing only this invariant mass distribution a 5.0σ-significance observation is made (5.6σ is expected).

However, the extremely well-measured final state leptons are rich with the

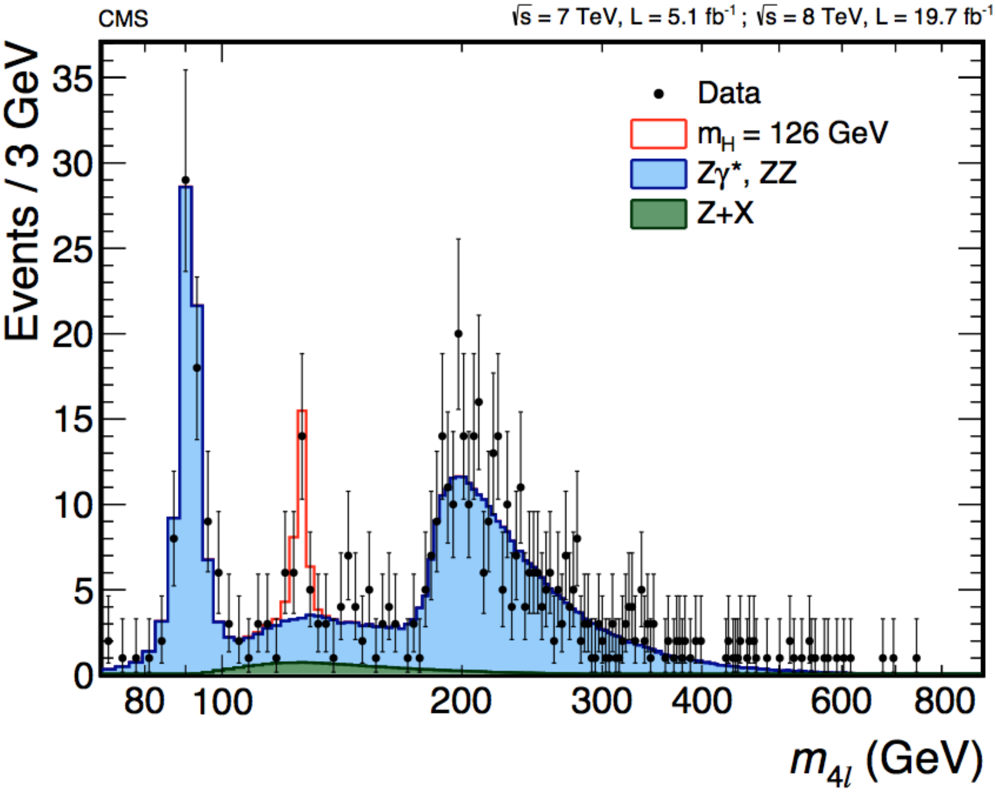


Figure 9.1: Final Run 1 H → ZZ → 4ℓ invariant mass in CMS [95]. The combined invariant mass of all 4ℓ channels (4μ, 2μ2e, and 4e) for selected events. All MC samples are normalized to the SM expectations and data is visualized as black points with error bars [95].

kinematics of the Z bosons and therefore reconstucted Higgs boson candidate decay kinematics. Five decay angles (see Figure 9.2) and the two masses of the reconstructed Z bosons depend on the spin and parity of parent particle, which is spin-0 and positive parity for a SM Higgs boson. Therefore, this information can be exploited to distinguish between background and signal as well as between different spin-parity models. To this end a kinematic discriminant, $\mathcal{D}^{\mathrm{kin}}_{\mathrm{bkg}}$, has been developed for the primary analysis whose purpose is to distinguish between SM Higgs decay kinematics and background ZZ kinematics.

A likelihood function, based on three variables $m_{4\ell}$, $\mathcal{D}^{\mathrm{kin}}_{\mathrm{bkg}}$ and $p_{T,4\ell}$ (or a dijet discriminator $\mathcal{D}_{\mathrm{jet}}$ for the classes designed to select vector boson fusion events) is used to quantify the significance of the signal above the expected background and measure the signal strength. The observed significance of the signal is 6.8σ

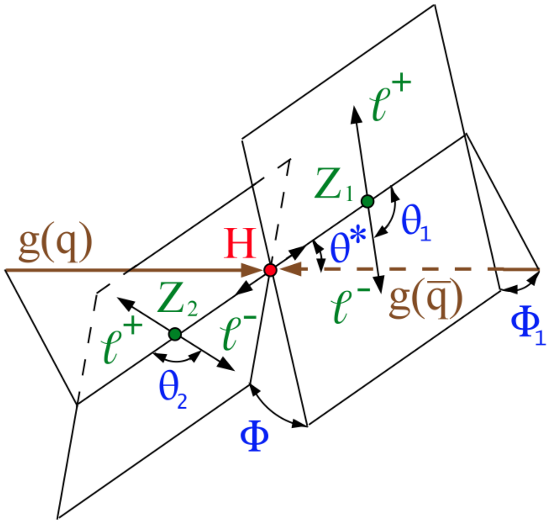


Figure 9.2: H → ZZ → 4ℓ decay kinematics schematic. With the precise measurement of the electrons and muons in the final state, the Z bosons can be reconstructed and the full decay of Higgs candidates can be reconstructed.

(with 6.7σ expected) at 125.7 GeV. This likelihood function is also used to measure the signal strength of $\mu = 0.93^{+0.26}_{-0.23}(\text{stat.})^{+0.13}_{-0.09}(\text{syst.})$ at the best fit mass of 125.6 GeV.

A similar 3-dimensional likelihood is used to measure the mass. However, instead of $p_{T,4\ell}$ or $\mathcal{D}_{\text{jet}}$, the per-event mass resolution estimate ($\mathcal{D}_{\text{m}}$), which is estimated from lepton resolution, is used with $m_{4\ell}$ and $\mathcal{D}^{\text{kin}}_{\text{bkg}}$ to build the likelihood function. The best fit mass is measured to be $\text{M}_{\text{H}} = 125.6 \pm 0.4(\text{stat.}) \pm 0.2(\text{syst.})$.

Since information about spin and parity is contained in the decay kinematics, the H → ZZ → 4ℓ channel is ideally suited to test alternative spin-parity hypotheses. Just as the kinematic discriminant, $\mathcal{D}^{\text{kin}}_{\text{bkg}}$, is used for background and SM Higgs discrimination, numerous $\mathcal{D}_{J^P}$ have been developed for alternative model discremination and incorporated into analogous likelihood functions, $\mathcal{L}_{J^P}$. To distinguish between different hypothesis, test statistics of the form $-2\ln\left(\frac{\mathcal{L}_{J^P}}{\mathcal{L}_{0^+}}\right)$ have been created. SM and alternative model pseudo-data are produced to evaluate the expected and observed confidence levels.

Figure 9.3 shows one, two and three standard deviation bands for the SM and alternative models for each alternative model with the observed test statistic in black. It is visually apparent that the SM, 0^+ assumption is strongly favored in almost all of the comparisons, and moreover, all spin-1 models are ruled out at the 99% confidence level, while all spin-2 models are ruled out at the 95% confidence level.

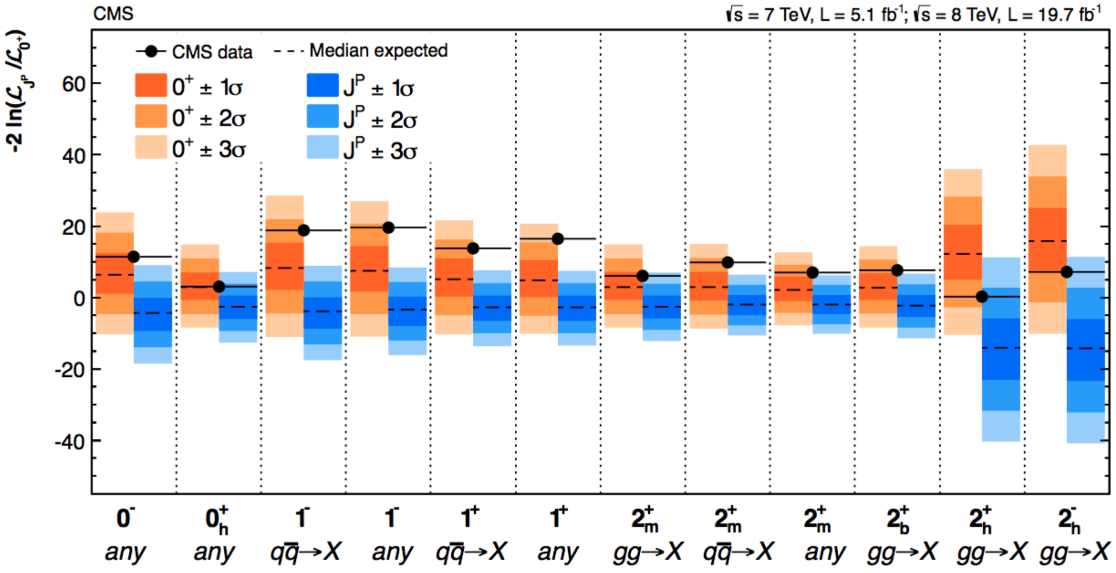


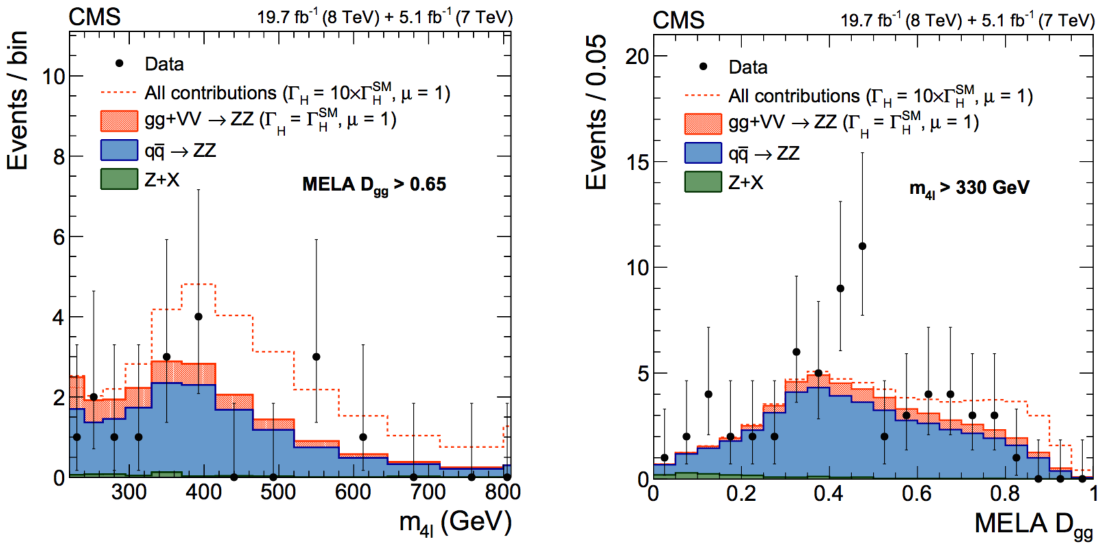
Figure 9.3: $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ spin-parity compatibility results. Many alternative spin-parity (J^P) models of a Higgs boson have been compared to the SM prediction of 0^+ using selected data with $121.5 < m_{4\ell} < 130.5$ GeV. Above are the results of pseudo-data toy studies comparing SM toys with alternative spin-parity toys using the same statistics test. The observed statistics test is in black points over the SM (orange) and alternate (blue) toy distributions. All alternate models are disfavored compared to the SM. Indeed, all spin-1 models are ruled out at the 99% confidence level, while all spin-2 models are ruled out at the 95% confidence level [95].

Finally, two measurements of the intrinsic width of the Higgs boson have been carried out using this channel. The first is a direct measurement from the width of the signal resonance. This measurement is carried out by adding the Higgs' natural width, Γ_{H}, into the signal model[1]. 95% confidence level upper limit is set at 3.4 GeV for the width whereas 2.8 GeV is the expected upper limit. The $\mathrm{H} \to \gamma\gamma$ analysis has a very similar result detailed in Section 7.8. The corresponding observed (expected) upper limit is 2.4 (3.1) GeV. Both of these measurements are far from the SM width, $\Gamma_{\mathrm{H}}^{\mathrm{SM}} = 4.15$ MeV, because of limited experimental resolution.

---

[1] The same likelihood function used for the mass measurement is used for the width measurement.

However, a more strict upper limit on the intrinsic width can be determined using the ratio of on-shell and off-shell H $\rightarrow$ ZZ cross section measurements [96]. Using the H $\rightarrow$ ZZ $\rightarrow$ 4ℓ and H $\rightarrow$ ZZ $\rightarrow$ $2\ell 2\nu$ analyses using CMS's full Run 1 dataset [97]. For large values of Γ_{H}, high mass ZZ yields increase substantially because of greater off-shell yields of H $\rightarrow$ ZZ and despite negative interference with the ZZ continuum. In addition to the substantial difference in the yields of selected high mass ZZ events, the kinematics of the ZZ decay can be further utilized when the Z bosons can be fully reconstructed as is the case in four lepton events. A kinematic discriminant, $\mathcal{D}_{gg}$, is constructed with the objective of distinguishing gg $\rightarrow$ ZZ events (including signal, background and interference) from $q\bar{q}$ $\rightarrow$ ZZ, which is primarily background. These two variables are shown for data, background MC and two different Γ_{H} signal MC samples in Figure 9.4.



(a) (b)

Figure 9.4: H $\rightarrow$ ZZ $\rightarrow$ 4ℓ off-shell Higgs width variables. Figure (a) shows the high mass contribution of a large width Higgs on top of background MC with data overlaid. The kinematic discriminant, $\mathcal{D}_{gg}$, is shown in Figure (b). In each plot the variable not shown has a relatively tight selection applied to show the signal-like region of the plotted variable. Both variables show large separation between signal models with different assumed values of Γ_{H} [97].

In the H $\rightarrow$ ZZ $\rightarrow$ $2\ell 2\nu$ channel, the decay kinematics are incomplete because of the presence of multiple neutrinos. Therefore, the fact that the yields are significantly higher is the sole source of discrimination. Therefore, the discriminating variables are transverse mass, m_T (shown in Figure 9.5), and E_T^{Missing}. However, this channel still has similar sensitivity to that of the H → ZZ → 4ℓ width measurement without using a more sophisticated kinematic discriminant. The on-shell portion of this measurement is taken from the H → ZZ → 4ℓ observation.

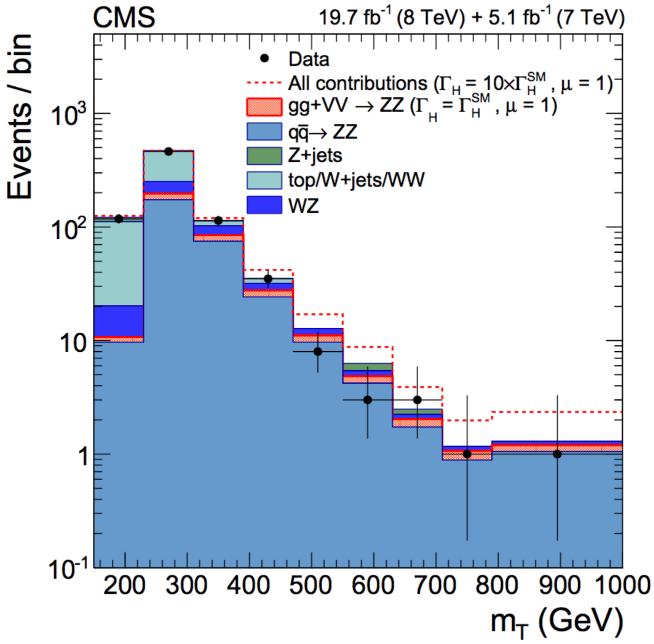


Figure 9.5: Distribution of the transverse mass in the H → ZZ → $2\ell 2\nu$ channel. Points represent the data, filled histograms the expected contributions from the backgrounds, and from the gluon fusion (gg) and vector boson fusion (VV) SM processes (including the Higgs-mediated contributions). The dashed line corresponds to the total expected yield for a Higgs boson width 10 times the SM width. Gluon Fusion and VBF signal strengths are fixed to their SM values [97].

The observed 95% confidence level upper limits on the Higgs width are both 33 MeV in both the H → ZZ → 4ℓ and H → ZZ → $2\ell 2\nu$ channels individually, whereas the expected upper limits are 42 MeV and 44 MeV for H → ZZ → 4ℓ and H → ZZ → $2\ell 2\nu$ channels respectively. The combined observed (expected) 95% confidence level limit is 22 MeV (33 MeV). These upper limits (observed and expected) are shown in Figure 9.6. This constraint is much better than what one would naively expect from hadron collision data and two orders of magnitude smaller than the direct on-shell measurement.

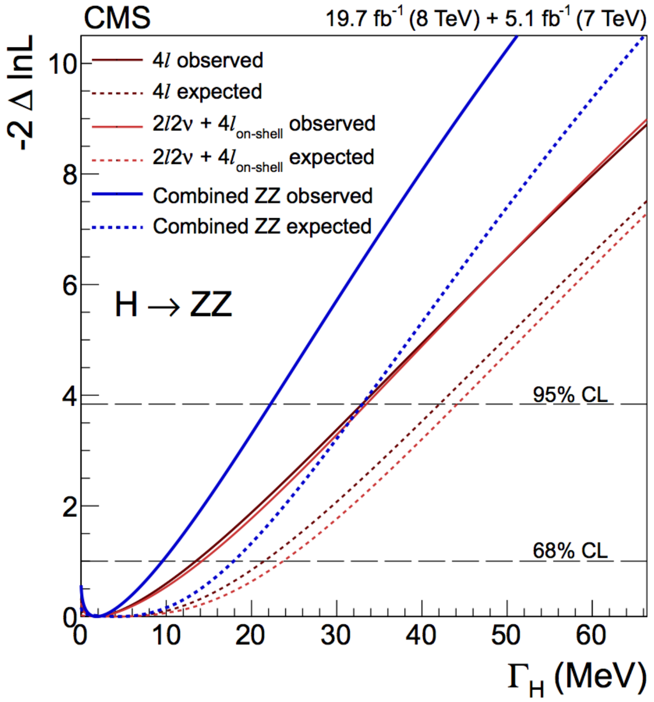


Figure 9.6: Off-shell constraint of the Higgs width. This is a more strict constraint of the Higgs width using the H → ZZ → 4ℓ and H → ZZ → $2\ell 2\nu$ than the [97].

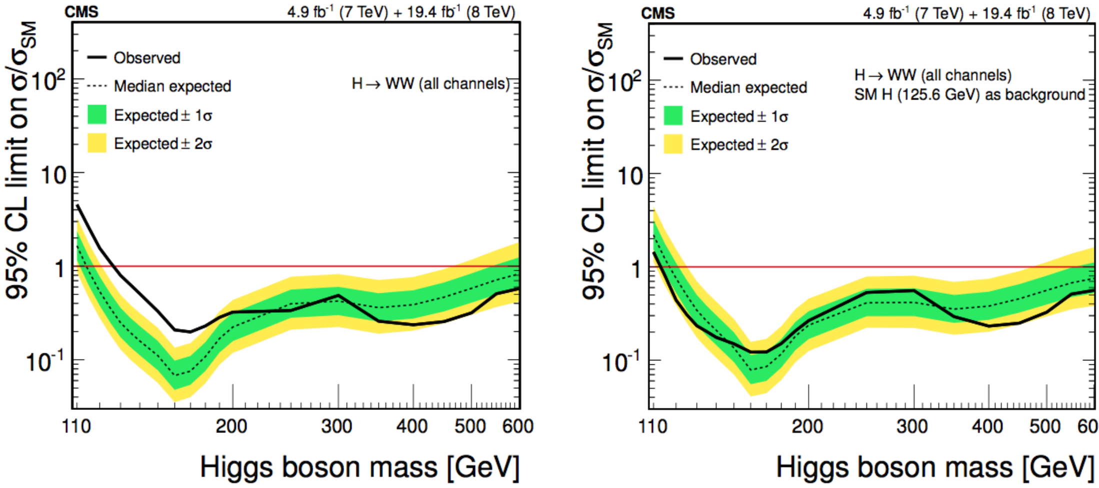
9.2 Low Resolution Channels and CMS Summary

The remaining channels which are sensitive to a Higgs boson at 125 GeV have much lower mass resolution than the H → ZZ → 4ℓ and H → $\gamma\gamma$ channels. Therefore M_H is not directly measurable and any signal detected is found as an excess over a wide range in mass.

9.2.1 Bosonic Channels

The H → WW channel has a relatively high branching ratio over a wide range of M_H. Both W bosons decaying to either an electron plus a neutrino or a muon plus a neutrino, which accounts for 4.5% of H → WW events, constitutes the most powerful sub-channel. However, the presence of multiple final state neutrinos makes it impossible to reconstruct the invariant mass of the Higgs boson. Figure 9.7a shows 95% confidence level exclusion limits in this H → WW sub-channel [98]. This channel is able to exclude a SM Higgs boson $127 < M_H < 600$

GeV at 95% confidence level. In addition, there is a clear, broad excess observed at low mass. Figure 9.7b shows that the observed exclusion limits fall within expectations if a SM Higgs boson with $M_H = 125.6$ GeV is included in the background. Assuming there is a SM Higgs boson with this mass, a signal strength of $0.72^{+0.20}_{-0.18}$ is measured with an observed (expected) signficance of 4.3σ (5.8σ).



(a) (b)

Figure 9.7: H → WW search results. Figure (a) contains the expected and observed exclusion limits relative to the SM cross section as a function M_H for the H → WW search in CMS. All masses in the search range are excluded except in the low mass range where there is a large excess. Figure (b) shows the expected and observed exclusion limits when a SM Higgs boson with $M_H = 125.6$ GeV treated as a background. The observation falls within expectations [98].

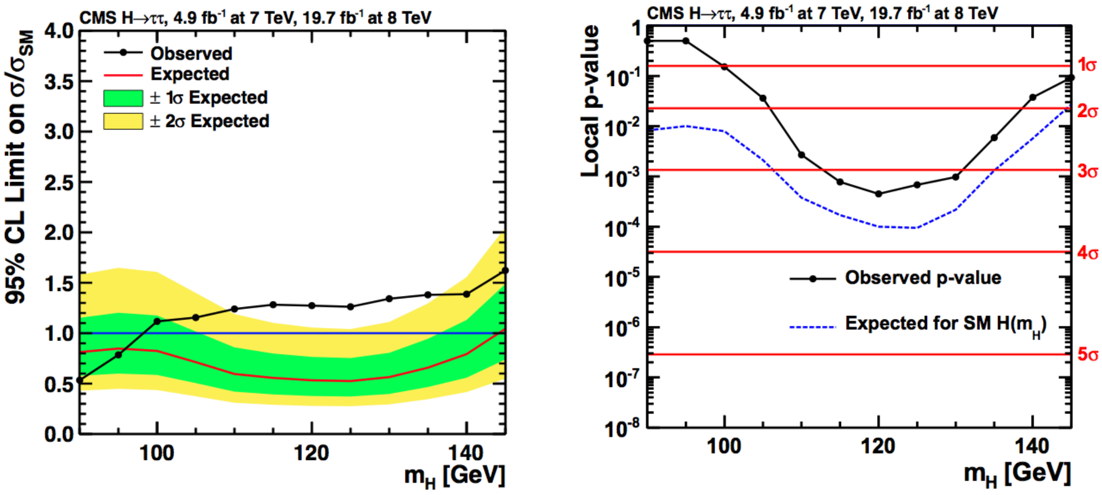
9.2.2 Fermionic Channels

In the $H \rightarrow \tau^+\tau^-$ channel [93] there is broad excess which is visible in both the exclusion limits (Figure 9.8a) and the local pvalues as a function of M_H (Figure 9.8b). The excess is has a signficance of 3.2σ (3.7σ expected) at $M_H = 125$ GeV. Moreover, the excess broadly impacts the significance over a wide mass range in a fashion very similar to that expected of a SM Higgs boson with $M_H = 125$ GeV.

Likewise, in the $H \rightarrow b\bar{b}$ channel there is a broad excess that weakens the exclusion limits, which is shown in Figure 9.9 [99]. When signal with $M_H = 125$

GeV is injected, the exclusion limits with signal plus background are comparable to the observed limits.

These excesses are compatible with the expected signal from a Higgs boson corresponding to that which has been independently observed in H → $\gamma\gamma$ and H → ZZ → 4ℓ channels and have been interpreted as such. Most importantly evidence in these channels represent couplings to fermions which is a fundamental prediction of the SM Higgs boson. Combining these two channels at 125 GeV yields 3.8σ-significance evidence (4.4σ expected) of fermion coupling with a combined signal strength of $\sigma/\sigma_{\mathrm{SM}} = 0.83 \pm 0.24$ [61].



(a) (b)

Figure 9.8: H → $\tau^+\tau^-$ search results. Figure (a) contains the expected and observed exclusion limits relative to the SM cross section as a function of $\mathrm{M_H}$ for the H → $\tau^+\tau^-$ search in CMS. Figure (b) shows local pvalues of the data as well as pvalues expected for a SM Higgs boson with $\mathrm{M_H} = 125$ GeV as a function of mass [93].

Table 9.1 summarizes the mass measurements, signal strengths ($\sigma/\sigma_{\mathrm{SM}}$), observed and expected significances for the five main channels that are sensitive to the SM Higgs boson with $\mathrm{M_H}$ 125 GeV with CMS Run 1 data. At the time of writing a CMS combination including all final results is being prepared but has not been completed with all updated results. This is because the final H → $\gamma\gamma$ analysis results have been completed only very recently.

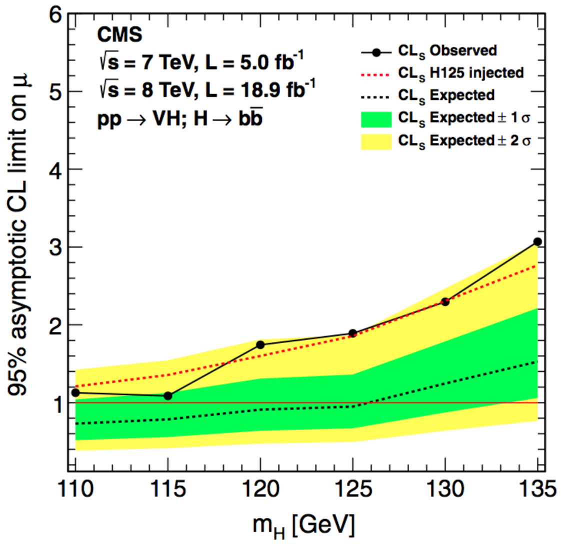


Figure 9.9: H → b$\bar{\text{b}}$ search results. This figure contains the expected and observed exclusion limits relative to the SM cross section as a function M_H for the H → b$\bar{\text{b}}$ search in CMS. The red dashed line shows the expected exclusion limits when a SM Higgs boson with $\text{M}_\text{H} = 125$ GeV is injected. The observed limits are compatible with the presence of a SM Higgs boson with $\text{M}_\text{H} = 125$ GeV [99].

Table 9.1: Summary of CMS's five main searches are summarized below. In the low resolution channels a mass is assumed for measurements of signal strength and significance because it cannot be measured directly.

Decay Channel	Measured Mass (GeV)	Assumed Mass (GeV)	Signal Strength	Observed (Expected) Significance
H → b$\bar{\text{b}}$	–	125	1.0 ± 0.5	2.1σ (2.1σ)
$\text{H} \to \tau^+\tau^-$	–	125	0.78 ± 0.27	3.2σ (3.7σ)
H → WW	–	125.6	$0.72^{+0.20}_{-0.18}$	4.3σ (5.8σ)
$\text{H} \to \text{ZZ} \to 4\ell$	125.6 ± 0.5	–	$0.93^{+0.29}_{-0.25}$	6.8σ (6.7σ)
$\text{H} \to \gamma\gamma$	$124.69^{+0.35}_{-0.34}$	–	$1.14^{+0.26}_{-0.23}$	5.7σ (5.2σ)

Chapter 10

Conclusions

The search and observation of the Higgs boson in the two photon decay channel have been described in detail. All measurements with the two photon observation are in favor of or consistent with the SM Higgs boson. The mass of this Higgs boson, which has no prediction, is measured to be $124.70^{+0.35}_{-0.34}$ GeV. The signal strength is measured to be $\sigma/\sigma_{SM} = 1.14^{+0.26}_{-0.23}$, which is very well in agreement with SM expectations. The SM assumption of spin-parity, 0^+, is favored compared to minimally coupling spin-2 models, but spin-2 cannot be excluded at the 95% confidence level.

In addition, the searches for di-muon and di-electron decays of this Higgs boson have also been described. The absence of any excess is completely expected in a SM interpretation of the observed particle. However, these searches are more than sensitive to the cross section times branching ratio of the SM $\mathrm{H} \to \tau^+\tau^-$ expectation. Given the evidence of $\mathrm{H} \to \tau^+\tau^-$ decays at roughly that cross section times branching ratio, these analyses have been able to eliminate the possibility of universal lepton decay by this particle.

Moreover, 3.8σ significant evidence of direct fermionic decay with $\mathrm{H} \to \tau^+\tau^-$ and $\mathrm{H} \to \mathrm{b\bar{b}}$ channels' combination re-enforces a SM interpretation. Likewise, the $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ analysis in CMS has provided the very important result of strong exclusions of all tested non-0^+ models, which again shows experimental preference for this particle being the SM Higgs boson. Indeed, none of the measurements from LHC Run 1 (from CMS or ATLAS) have been found to contradict or disfavor a SM interpretation of this particle.

However, there are several SM Higgs features to still be sought. The most fundamental feature of a Higgs boson (SM or not) yet to be observed is the property of self-coupling. Three- and four-Higgs couplings are fundamental to the BEH mechanism but LHC's Run 1 datasets should not produce measurable amounts of self-coupling events nor have any excesses been observed. Likewise, the precision in measurements of couplings to fermions, which should be proportional to the particle mass in the SM, will improve with further LHC data.

LHC's Run 2 dataset will pose many challenges (even greater pile-up, further radiation damage, etc) but it also has great potential for exciting further analysis of the Higgs boson at 125 GeV. Moreover, with the potentially large amounts of data expected, much more physics beyond the SM can be tested and probed including supersymmetry. Yet the most tantalizing outcome is a clear, yet completely unpredicted observation requiring new physics theory.

Appendix A

Signal Cross Sections at 7 TeV and 8 TeV

Table A.1: SM Higgs cross sections at 8 TeV (pb) for different production mechanisms and H → $\gamma\gamma$ branching ratios for different Higgs masses.

M_H (GeV)	Gluon Fusion	Vector Boson Fusion	W → WH, Z → ZH	$t\bar{t}H$	Branching Fraction
90	36.2	2.19	1.99, 1.09	0.32	1.22×10^{-3}
95	32.7	2.08	1.70, 0.938	0.28	1.39×10^{-3}
100	29.7	1.98	1.45, 0.810	0.24	1.58×10^{-3}
105	27.0	1.89	1.24, 0.702	0.21	1.77×10^{-3}
110	24.7	1.80	1.07, 0.613	0.19	1.95×10^{-3}
115	22.7	1.73	0.927, 0.536	0.17	2.11×10^{-3}
120	20.9	1.65	0.805, 0.471	0.15	2.23×10^{-3}
125	19.3	1.58	0.705, 0.415	0.13	2.28×10^{-3}
130	17.9	1.51	0.617, 0.367	0.12	2.24×10^{-3}
135	16.6	1.44	0.542, 0.326	0.10	2.12×10^{-3}
140	15.4	1.39	0.477, 0.290	0.09	1.93×10^{-3}
145	14.5	1.33	0.422, 0.258	0.08	1.67×10^{-3}
150	13.6	1.28	0.373, 0.231	0.07	1.36×10^{-3}

Table A.2: SM Higgs cross-sections at 7 TeV (pb) for different production mechanisms and H → $\gamma\gamma$ branching ratios for different Higgs masses.

M_H (GeV)	Gluon Fusion	Vector Boson Fusion	W → WH, Z → ZH	$t\bar{t}H$	Branching Fraction
90	29.03	1.723	1.654, 0.8959	0.2162	1.22×10^{-3}
95	26.10	1.639	1.404, 0.7678	0.1880	1.39×10^{-3}
100	23.64	1.557	1.195, 0.6616	0.1637	1.58×10^{-3}
105	21.45	1.478	1.029, 0.5724	0.1432	1.77×10^{-3}
110	19.56	1.410	0.8847, 0.4978	0.1257	1.95×10^{-3}
115	17.89	1.344	0.7626, 0.4345	0.1105	2.11×10^{-3}
120	16.43	1.279	0.6617, 0.3808	0.09758	2.23×10^{-3}
125	15.13	1.222	0.5785, 0.3351	0.08632	2.28×10^{-3}
130	13.98	1.168	0.5059, 0.2957	0.07660	2.24×10^{-3}
135	12.95	1.117	0.4431, 0.2616	0.06816	2.12×10^{-3}
140	12.02	1.069	0.3839, 0.2322	0.06079	1.93×10^{-3}
145	11.24	1.023	0.3437, 0.2068	0.05429	1.67×10^{-3}
150	10.51	0.980	0.3034, 0.1842	0.04867	1.36×10^{-3}

Appendix B

2011 and 2012 Trigger Tables

Table B.1: Trigger paths employed in 2011A data-taking.

Trigger Path	Prescale	L1 Seed
Runs 165970–173198		
HLT_Photon26_CaloIdL_IsoVL_Photon18_CaloIdL_IsoVL_v	1	L1_SingleEG15
HLT_Photon26_CaloIdL_IsoVL_Photon18_R9Id_v	1	L1_SingleEG15
HLT_Photon26_R9Id_Photon18_CaloIdL_IsoVL_v	1	L1_SnigleEG15
HLT_Photon26_R9Id_Photon18_R9Id_v	1	L1_SingleEG15
HLT_Photon26_CaloIdL_IsoVL_Photon18_v	20–120	L1_SingleEG15
HLT_Photon26_Photon18_v	180–1100	L1_SingleEG15
HLT_Photon36_CaloIdL_IsoVL_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_CaloId(L)_IsoVL_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_CaloIdL_IsoVL_Photon22_v	7–45	L1_SingleEG20
HLT_Photon36_IsoVL_Photon22_v	5–30	L1_SingleEG20
Runs 161216–165633		
HLT_Photon26_CaloIdL_IsoVL_Photon18_CaloIdL_IsoVL_v	1	L1_SingleEG12
HLT_Photon26_CaloIdL_IsoVL_Photon18_R9Id_v	1	L1_SingleEG12
HLT_Photon26_R9Id_Photon18_CaloIdL_IsoVL_v	1	L1_SingleEG12
HLT_Photon20_R9Id_Photon18_R9Id_v	1	L1_SingleEG12
HLT_Photon26_CaloIdL_IsoVL_Photon18_v	1	L1_SingleEG12
HLT_Photon26_Photon18_v	10–200	L1_SingleEG12
HLT_Photon20_CaloIdVL_IsoL_v	250–5000	L1_SingleEG12
Runs 160404–161176		
HLT_Photon26_CaloIdL_IsoVL_Photon18_CaloIdL_IsoVL_v	1	L1_SingleEG12
HLT_Photon26_IsoVL_Photon18_IsoVL_v	1	L1_SingleEG12
HLT_Photon26_CaloIdL_IsoVL_Photon18_v	1	L1_SingleEG12
HLT_Photon26_IsoVL_Photon18_v	1	L1_SingleEG12
HLT_Photon26_Photon18_v	1–30	L1_SingleEG12

Table B.2: Trigger paths employed in 2011B data-taking.

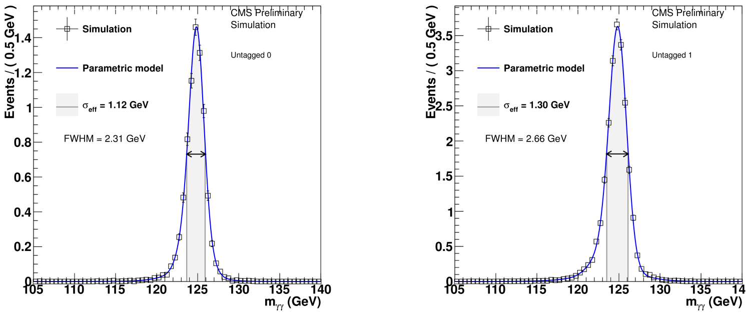
Trigger Path	Prescale	L1 Seed
Runs 178420–180252		
HLT_Photon26_CaloIdXL_IsoXL_Photon18_CaloIdXL_IsoXL_Mass60_v	1	L1_DoubleEG_12_5
HLT_Photon26_CaloIdXL_IsoXL_Photon18_R9IdT_Mass60_v	1	L1_DoubleEG_12_5
HLT_Photon26_R9IdT_Photon18_CaloIdXL_IsoXL_Mass60_v	1	L1_DoubleEG_12_5
HLT_Photon26_R9IdT_Photon18_R9IdT_Mass60_v	1	L1_DoubleEG_12_5
HLT_Photon26_CaloIdXL_IsoXL_Photon18_v	110–190	L1_DoubleEG_12_5
HLT_Photon26_Photon18_v	340–600	L1_DoubleEG_12_5
HLT_Photon36_CaloIdL_IsoVL_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_CaloIdL_IsoVL_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_CaloIdL_IsoVL_Photon22_v	45–75	L1_SingleEG20
HLT_Photon36_Photon22_v	180–300	L1_SingleEG20
Runs 173236–178380		
HLT_Photon26_CaloIdXL_IsoXL_Photon18_CaloIdXL_IsoXL_v	1	L1_DoubleEG_12_5
HLT_Photon26_CaloIdXL_IsoXL_Photon18_R9Id_v	1	L1_DoubleEG_12_5
HLT_Photon26_R9Id_Photon18_CaloIdXL_IsoXL_v	1	L1_DoubleEG_12_5
HLT_Photon26_R9Id_Photon18_R9Id_v	1	L1_DoubleEG_12_5
HLT_Photon26_CaloIdXL_IsoXL_Photon18_v	20–190	L1_DoubleEG_12_5
HLT_Photon26_Photon18_v	180–1740	L1_DoubleEG_12_5
HLT_Photon36_CaloIdL_IsoVL_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_CaloIdL_IsoVL_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_CaloIdL_IsoVL_v	1	L1_SingleEG20
HLT_Photon36_R9Id_Photon22_R9Id_v	1	L1_SingleEG20
HLT_Photon36_CaloIdL_IsoVL_Photon22_v	6–65	L1_SingleEG20
HLT_Photon36_Photon22_v	75–750	L1_SingleEG20

Table B.3: Trigger paths employed in 2012 data-taking.

Trigger Path	Prescale	L1 Seed
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass60	1	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass70	1	L1_DoubleEG_13_7
HLT_Photon26_CaloId10_Iso50_Photon18_CaloId10_Iso50_Mass60	1	L1_DoubleEG_13_7
HLT_Photon26_CaloId10_Iso50_Photon18_R9Id85_Mass60	1	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_Photon18_CaloId10_Iso50_Mass60	1	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_Photon18_R9Id85_Mass60	1	L1_DoubleEG_13_7
HLT_Photon26_Photon18	600-1600	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18	1000-1600	L1_DoubleEG_13_7
HLT_Photon36_CaloId10_Iso50_Photon22_CaloId10_Iso50	1	L1_SingleEG_22
HLT_Photon36_CaloId10_Iso50_Photon22_R9Id85	1	L1_SingleEG_22
HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon22_R9Id85_OR_CaloId10_Iso50	1	L1_SingleEG_22
HLT_Photon36_R9Id85_Photon22_CaloId10_Iso50	1	L1_SingleEG_22
HLT_Photon36_R9Id85_Photon22_R9Id85	1	L1_SingleEG_22
HLT_Photon36_Photon22	300-800	L1_SingleEG_22
HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon22	300-500	L1_SingleEG_22

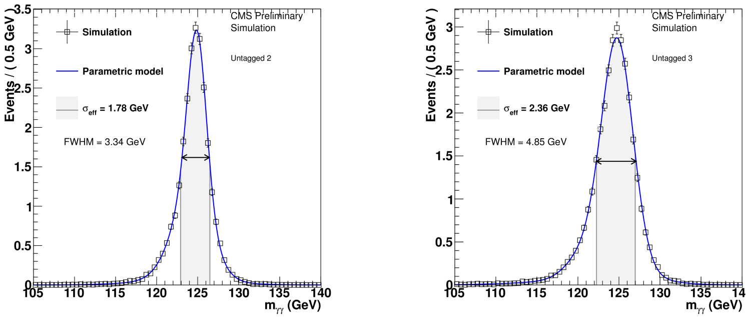
Appendix C

Signal Model Reference Information



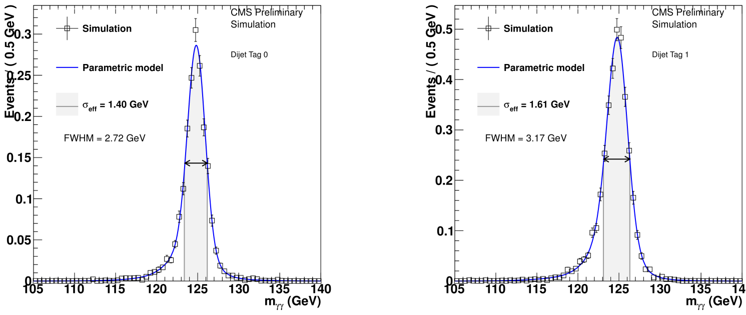
(a) Untagged 0
7 TeV - MVA

(b) Untagged 1
7 TeV - MVA



(c) Untagged 2
7 TeV - MVA

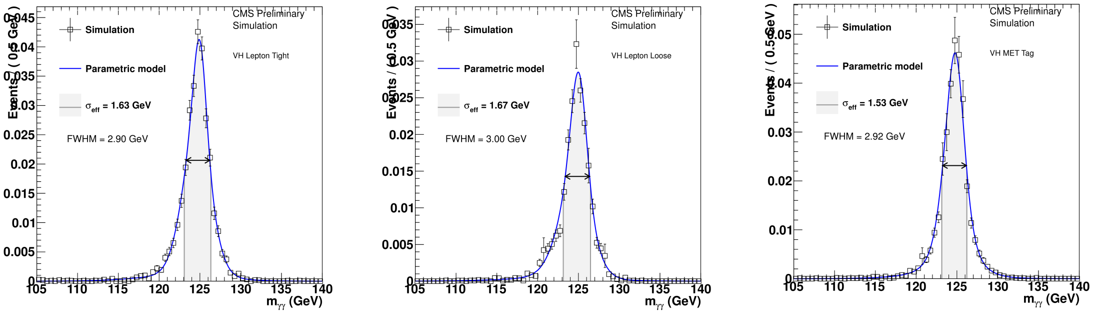
(d) Untagged 3
7 TeV - MVA



(e) Di-jet 0
7 TeV - MVA

(f) Di-jet 1
7 TeV - MVA

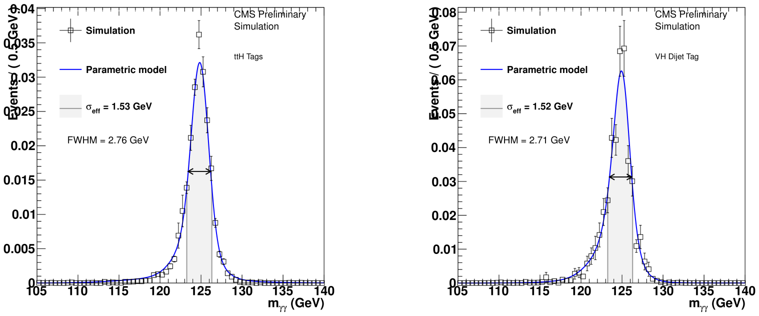
Figure C.1: MVA 7 TeV inclusive and di-jet event class signal models.



(a) VH Leptonic Tight 7 TeV - MVA

(b) VH Leptonic Loose 7 TeV - MVA

(c) VH E_T^{Missing} 7 TeV - MVA



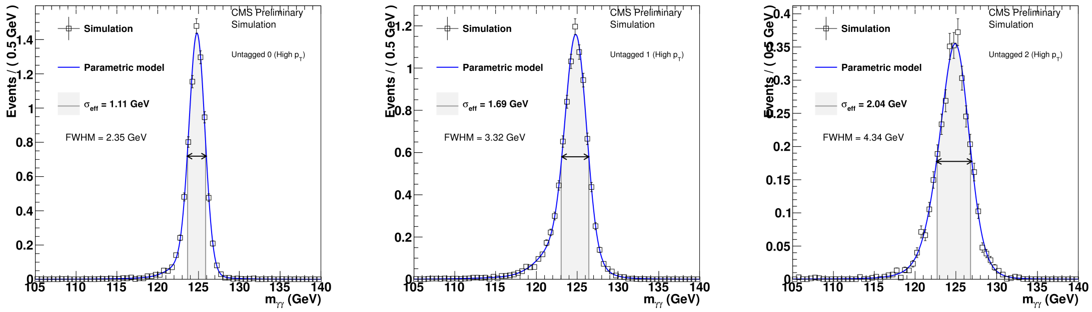
(d) $\mathrm{t\bar{t}H}$ Combined 7 TeV - MVA

(e) VH Di-jet 7 TeV - MVA

Figure C.2: MVA 7 TeV additional tag event class signal models.

Table C.1: Expected events, signal composition and signal width for MVA event classes with a mass hypothesis of M_H = 125 GeV for the 7 TeV and 8 TeV Multi-Variate Analysis.

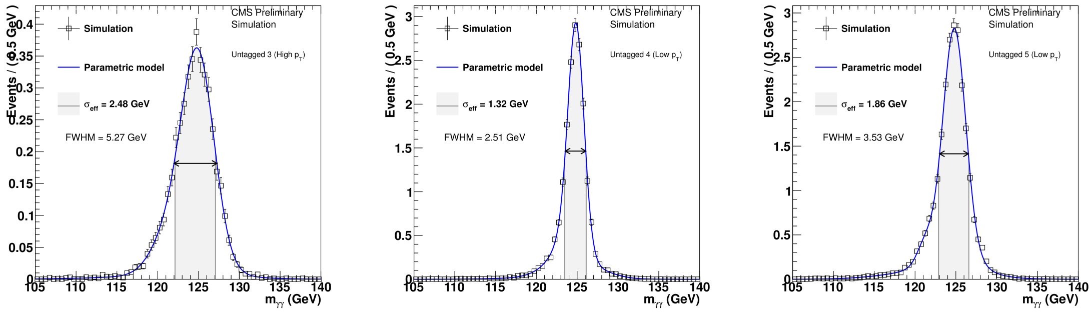
Event classes		SM Higgs boson expected signal (M_H =125 GeV)							
		Total	GF	VBF	WH	ZH	$t\bar{t}H$	σ_{eff} (GeV)	$\frac{\mathrm{FWHM}}{2.35(\mathrm{GeV})}$
7 TeV 5.1fb^{-1}	Untagged 0	7.8	84.9%	7.4%	4.5%	2.6%	0.6%	1.12	0.98
	Untagged 1	22.5	91.8%	4.3%	2.4%	1.4%	0.2%	1.30	1.13
	Untagged 2	26.4	91.7%	4.3%	2.5%	1.4%	0.2%	1.78	1.42
	Untagged 3	32.3	91.7%	4.2%	2.5%	1.4%	0.2%	2.36	2.06
	Dijet Tag 0	1.9	29.0%	70.5%	0.3%	0.2%	0.1%	1.40	1.16
	Dijet Tag 1	3.7	49.2%	48.9%	1.0%	0.6%	0.3%	1.61	1.35
	VH Lepton Tight	0.3	0.0%	0.0%	77.2%	20.6%	2.2%	1.63	1.23
	VH Lepton Loose	0.2	4.8%	1.1%	78.2%	15.0%	0.9%	1.67	1.28
	VH MET Tag	0.3	12.1%	1.1%	38.2%	41.0%	7.6%	1.53	1.24
	VH Dijet Tag	0.4	40.4%	2.3%	35.6%	20.0%	1.7%	1.52	1.15
	ttH Tags	0.2	6.3%	1.0%	2.2%	1.3%	89.2%	1.53	1.17
	Combined 7 TeV	96.0	87.1%	7.4%	3.2%	1.7%	0.5%	1.75	1.35
8 TeV 19.7fb^{-1}	Untagged 0	5.8	75.0%	12.3%	7.1%	3.7%	2.0%	1.05	0.86
	Untagged 1	50.6	85.1%	7.9%	4.0%	2.4%	0.6%	1.19	1.00
	Untagged 2	117.5	91.1%	4.7%	2.4%	1.4%	0.3%	1.44	1.17
	Untagged 3	152.8	91.6%	4.4%	2.4%	1.4%	0.2%	2.04	1.52
	Untagged 4	122.3	93.1%	3.5%	2.0%	1.1%	0.2%	2.63	2.09
	Dijet Tag 0	4.5	17.6%	82.1%	0.2%	0.1%	0.1%	1.28	0.92
	Dijet Tag 1	5.4	25.4%	73.6%	0.6%	0.2%	0.2%	1.49	1.13
	Dijet Tag 2	13.7	44.2%	52.9%	1.4%	0.8%	0.8%	1.58	1.14
	VH Lepton Tight	1.4	0.0%	0.2%	77.1%	19.1%	3.7%	1.62	1.23
	VH Lepton Loose	0.9	0.0%	1.1%	80.0%	17.3%	1.6%	1.58	1.33
	VH MET Tag	1.7	9.9%	2.9%	37.0%	38.1%	12.0%	1.64	1.27
	VH Dijet Tag	1.6	31.6%	3.1%	39.9%	22.9%	2.5%	1.34	1.11
	ttH Leptonic Tag	0.5	0.0%	0.0%	1.6%	1.6%	96.8%	1.34	1.04
	ttH Multijet Tag	0.6	8.3%	0.9%	0.7%	0.9%	89.2%	1.46	1.24
	Combined 8 TeV	479.3	87.1%	7.6%	3.1%	1.7%	0.6%	1.86	1.31



(a)
EB-EB, High R_9, High p_T
7 TeV - Cut-based

(b)
EB-EB, Low R_9, High p_T
7 TeV - Cut-based

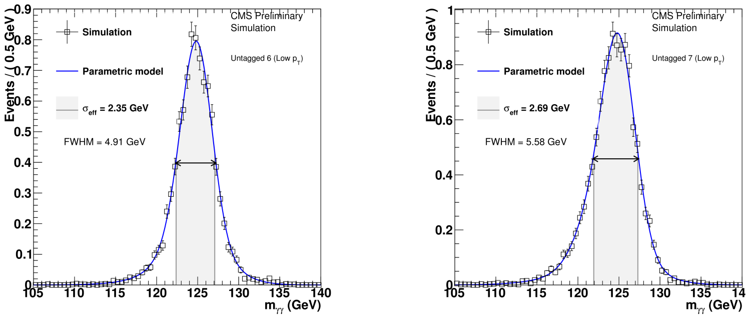
(c)
Mixed, High R_9, High p_T
7 TeV - Cut-based



(d)
Mixed, Low R_9, High p_T
7 TeV - Cut-based

(e)
EB-EB, High R_9, Low p_T
7 TeV - Cut-based

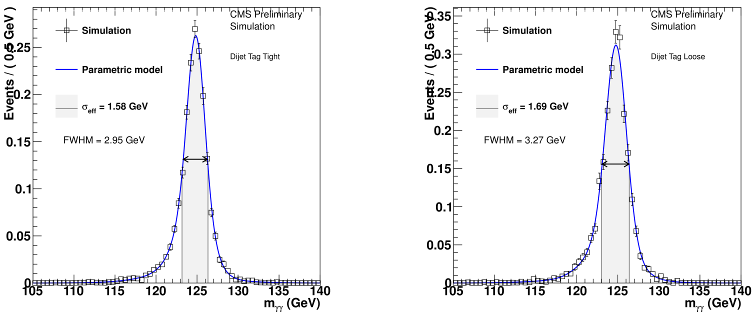
(f)
EB-EB, Low R_9, Low p_T
7 TeV - Cut-based



(g)
Mixed, High R_9, Low p_T
7 TeV - Cut-based

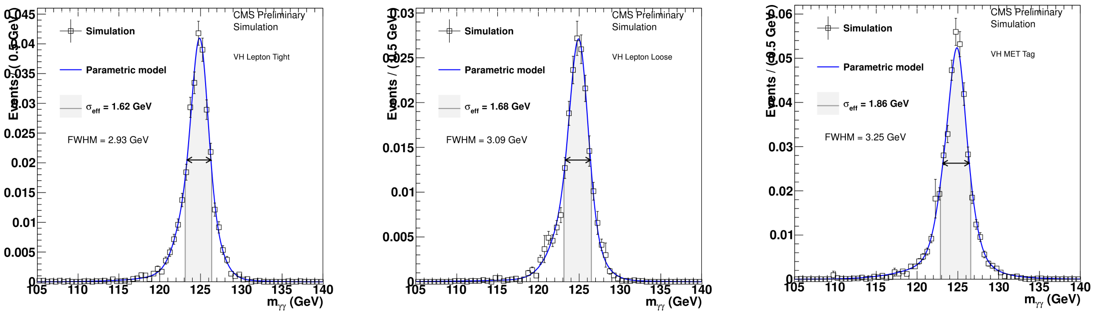
(h)
Mixed, Low R_9, Low p_T
7 TeV - Cut-based

Figure C.3: Cut-based 7 TeV inclusive event class signal models.



(a) Di-jet 0
7 TeV - Cut-based

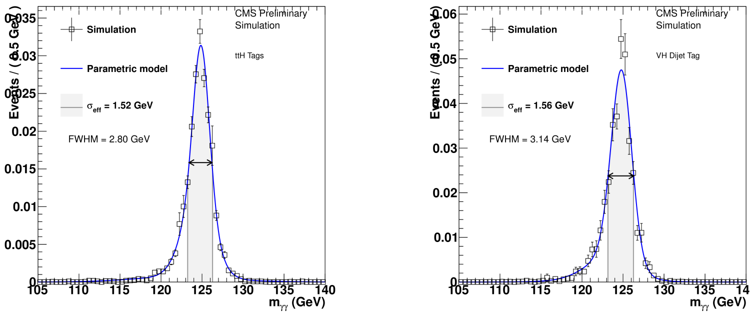
(b) Di-jet 1
7 TeV - Cut-based



(c) VH Leptonic Tight
7 TeV - Cut-based

(d) VH Leptonic Loose
7 TeV - Cut-based

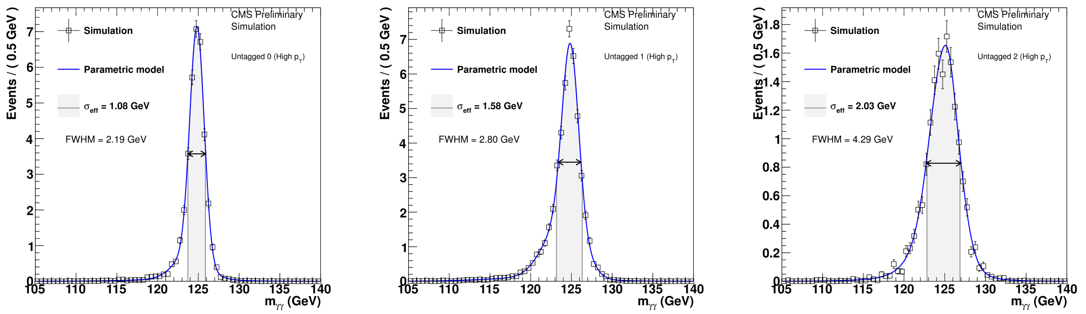
(e) VH E_T^{Missing}
7 TeV - Cut-based



(f) $t\bar{t}$H Combined
7 TeV - Cut-based

(g) VH Di-jet
7 TeV - Cut-based

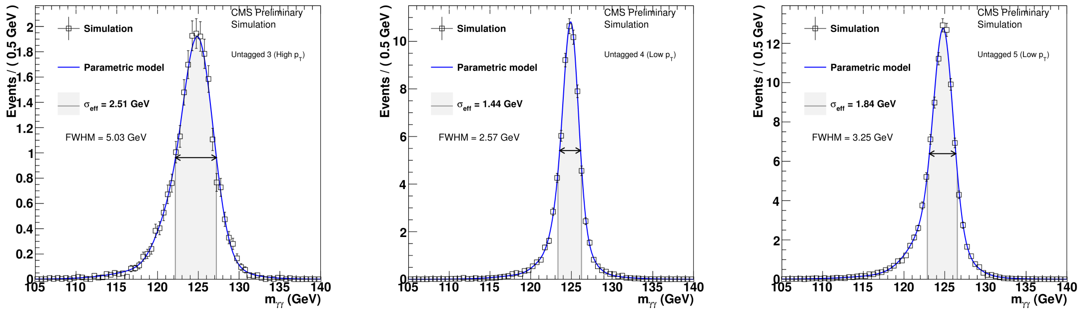
Figure C.4: Cut-based 7 TeV additional tag event class signal models.



(a)
EB-EB, High R_9, High p_T
8 TeV - Cut-based

(b)
EB-EB, Low R_9, High p_T
8 TeV - Cut-based

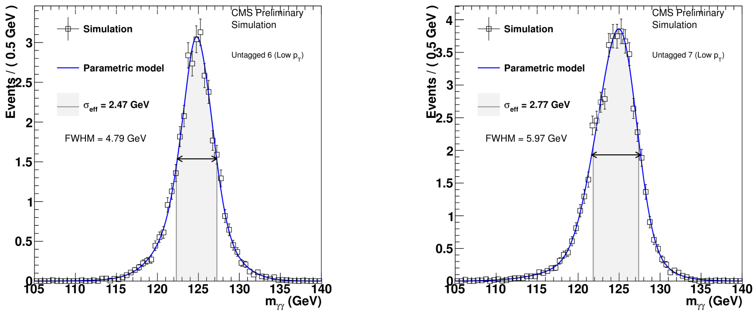
(c)
Mixed, High R_9, High p_T
8 TeV - Cut-based



(d)
Mixed, Low R_9, High p_T
8 TeV - Cut-based

(e)
EB-EB, High R_9, Low p_T
8 TeV - Cut-based

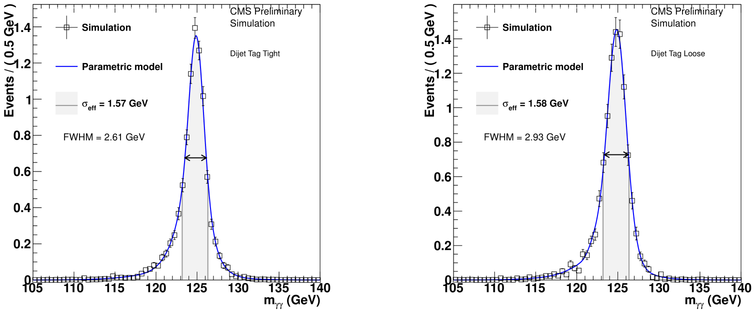
(f)
EB-EB, Low R_9, Low p_T
8 TeV - Cut-based



(g)
Mixed, High R_9, Low p_T
8 TeV - Cut-based

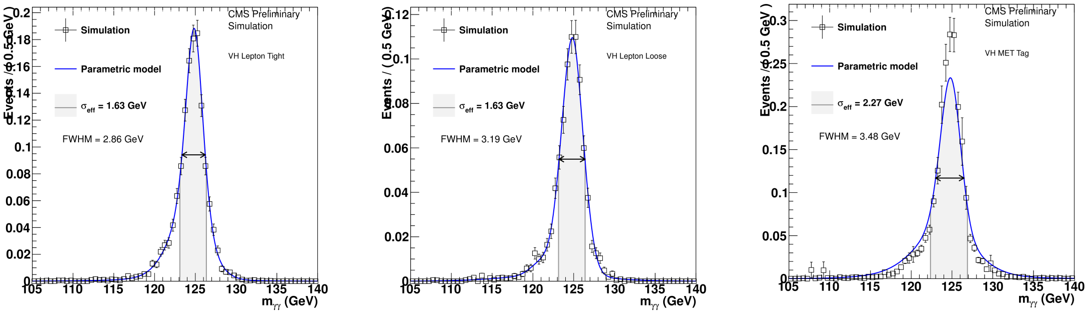
(h)
Mixed, Low R_9, Low p_T
8 TeV - Cut-based

Figure C.5: Cut-based 8 TeV inclusive tag event class signal models.



(a) Di-jet 0
8 TeV - Cut-based

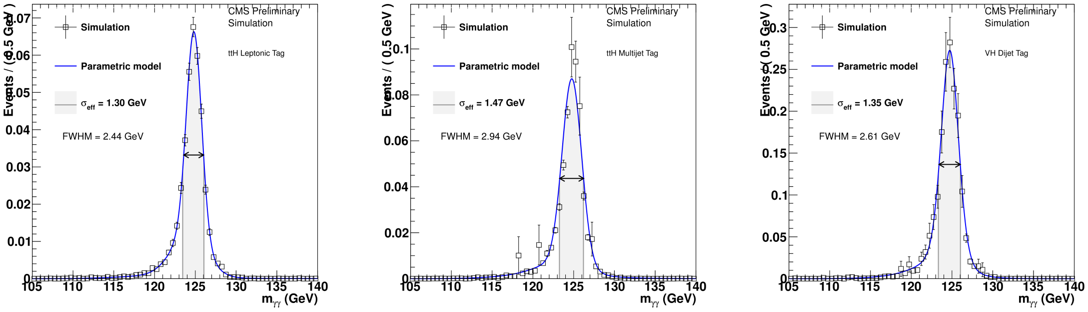
(b) Di-jet 1
8 TeV - Cut-based



(c) VH Leptonic Tight
8 TeV - Cut-based

(d) VH Leptonic Loose
8 TeV - Cut-based

(e) VH E_T^{Missing}
8 TeV - Cut-based



(f) $t\bar{t}$H Leptonic
8 TeV - Cut-based

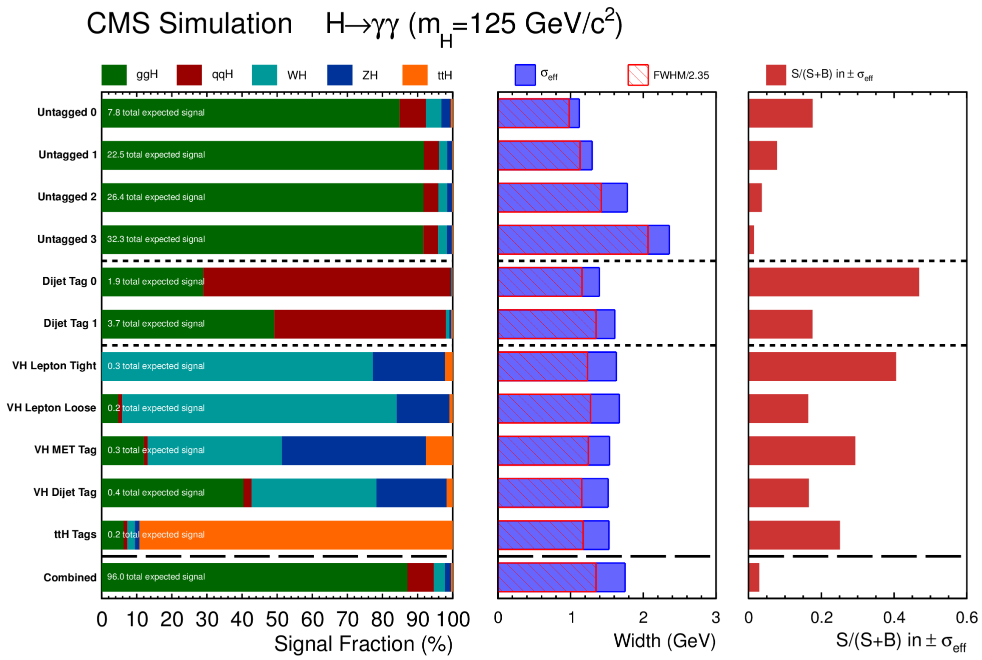
(g) $t\bar{t}$H Hadronic
8 TeV - Cut-based

(h) VH Di-jet
8 TeV - Cut-based

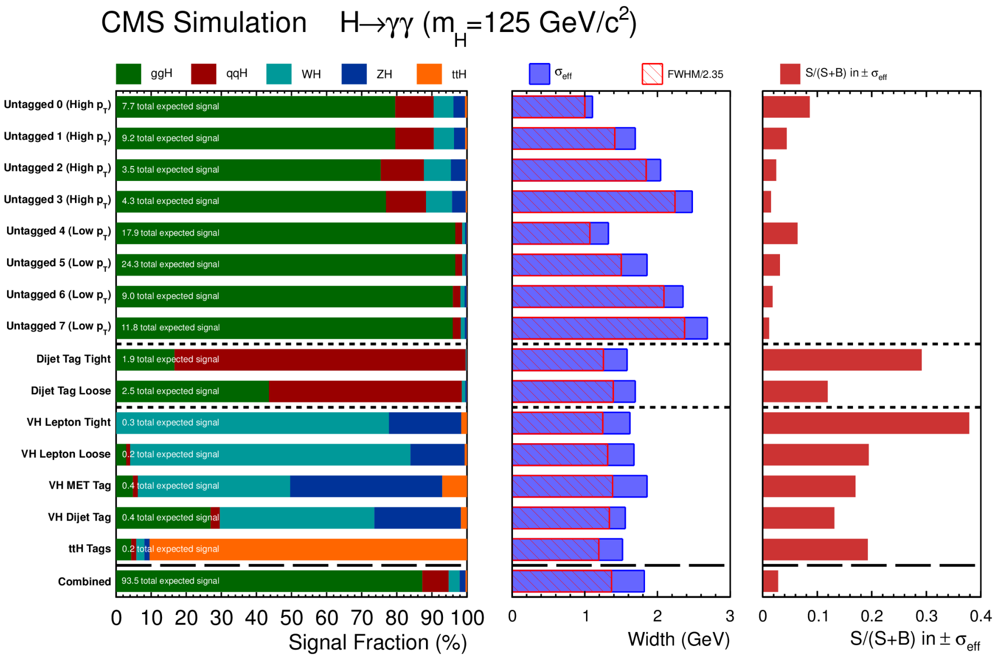
Figure C.6: Cut-based 8 TeV additional tag event class signal models.

Table C.2: Expected events, signal composition and signal width for MVA event classes with a mass hypothesis of $\mathrm{M_H}$ = 125 GeV for the 7 TeV and 8 TeV Cut-based Analysis.

Event classes		SM Higgs boson expected signal ($\mathrm{M_H}$ =125 GeV)							
		Total	GF	VBF	WH	ZH	$\mathrm{t\bar{t}H}$	σ_{eff} (GeV)	$\frac{\mathrm{FWHM}}{2.35(\mathrm{GeV})}$
7 TeV 5.1fb^{-1}	Untagged 0 (High p_T)	7.7	79.6%	11.0%	5.7%	3.3%	0.6%	1.11	1.00
	Untagged 1 (High p_T)	9.2	79.6%	11.0%	5.7%	3.2%	0.5%	1.69	1.41
	Untagged 2 (High p_T)	3.5	75.5%	12.3%	7.7%	4.2%	0.4%	2.04	1.84
	Untagged 3 (High p_T)	4.3	76.9%	11.4%	7.4%	3.9%	0.4%	2.48	2.24
	Untagged 4 (Low p_T)	17.9	96.7%	2.0%	0.8%	0.5%	0.0%	1.32	1.07
	Untagged 5 (Low p_T)	24.3	96.7%	2.0%	0.8%	0.5%	0.0%	1.86	1.50
	Untagged 6 (Low p_T)	9.0	96.0%	2.1%	1.2%	0.6%	0.0%	2.35	2.09
	Untagged 7 (Low p_T)	11.8	95.9%	2.3%	1.1%	0.6%	0.0%	2.69	2.37
	Dijet Tag Tight	1.9	16.7%	83.0%	0.2%	0.1%	0.0%	1.58	1.26
	Dijet Tag Loose	2.5	43.6%	55.0%	1.0%	0.5%	0.0%	1.69	1.39
	VH Lepton Tight	0.3	0.0%	0.0%	77.7%	20.6%	1.6%	1.62	1.25
	VH Lepton Loose	0.2	2.9%	1.1%	79.9%	15.5%	0.6%	1.68	1.31
	VH MET Tag	0.4	4.8%	1.4%	43.4%	43.4%	7.0%	1.86	1.38
	VH Dijet Tag	0.4	27.0%	2.6%	44.0%	24.7%	1.7%	1.56	1.34
	ttH Tags	0.2	4.3%	1.5%	2.3%	1.5%	90.4%	1.52	1.19
	Combined 7 TeV	93.5	87.3%	7.5%	3.1%	1.7%	0.4%	1.82	1.37
8 TeV 19.7fb^{-1}	Untagged 0 (High p_T)	36.6	81.1%	10.7%	4.7%	2.8%	0.7%	1.08	0.93
	Untagged 1 (High p_T)	48.7	81.2%	10.8%	4.7%	2.7%	0.6%	1.58	1.19
	Untagged 2 (High p_T)	16.2	77.0%	12.2%	6.7%	3.7%	0.5%	2.03	1.82
	Untagged 3 (High p_T)	22.7	78.5%	11.3%	6.3%	3.3%	0.4%	2.51	2.14
	Untagged 4 (Low p_T)	69.9	96.5%	2.1%	0.8%	0.4%	0.1%	1.44	1.09
	Untagged 5 (Low p_T)	105.1	96.6%	2.1%	0.8%	0.5%	0.1%	1.84	1.38
	Untagged 6 (Low p_T)	35.1	95.8%	2.3%	1.1%	0.6%	0.0%	2.47	2.04
	Untagged 7 (Low p_T)	51.7	95.9%	2.3%	1.2%	0.6%	0.0%	2.77	2.54
	Dijet Tag Tight	9.2	21.2%	78.2%	0.3%	0.1%	0.1%	1.57	1.11
	Dijet Tag Loose	10.5	49.0%	49.4%	1.0%	0.5%	0.0%	1.58	1.25
	VH Lepton Tight	1.4	0.0%	0.2%	77.5%	19.3%	3.1%	1.63	1.22
	VH Lepton Loose	0.8	0.0%	1.2%	80.3%	17.2%	1.3%	1.63	1.36
	VH MET Tag	2.2	12.6%	3.4%	37.3%	36.6%	10.0%	2.27	1.48
	VH Dijet Tag	1.7	31.4%	3.0%	40.5%	23.0%	2.1%	1.35	1.11
	ttH Leptonic Tag	0.4	0.0%	0.0%	2.0%	1.8%	96.2%	1.30	1.04
	ttH Multijet Tag	0.6	11.2%	1.3%	1.0%	0.6%	85.9%	1.47	1.25
	Combined 8 TeV	412.7	87.1%	7.7%	3.0%	1.6%	0.5%	1.84	1.30



(a) 7 TeV - MultiVariate Analysis

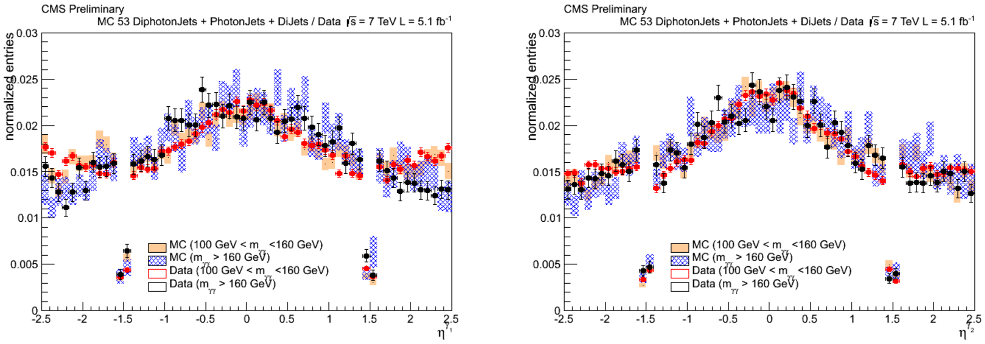


(b) 7 TeV - Cut-based

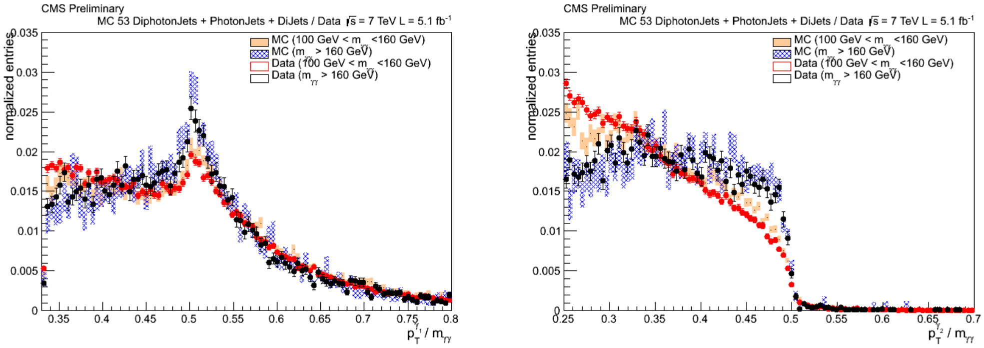
Figure C.7: Graphical representation of 7 TeV event classes' signal composition, signal width and S/(S+B) estimates for event classes in the MVA (a) and Cut-based Analysis (b) signal MC with $M_H = 125$ GeV.

Appendix D

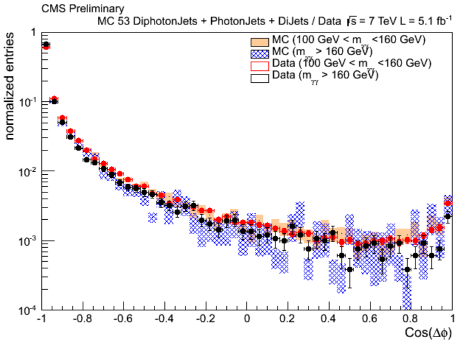
Di-photon BDT Validation Plots



(a) η_1 (b) η_2

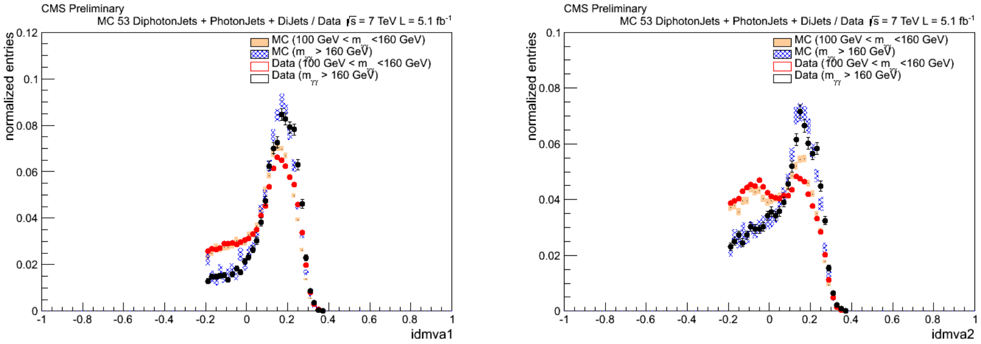


(c) $p_{T,1}/m_{\gamma\gamma}$ (d) $p_{T,2}/m_{\gamma\gamma}$

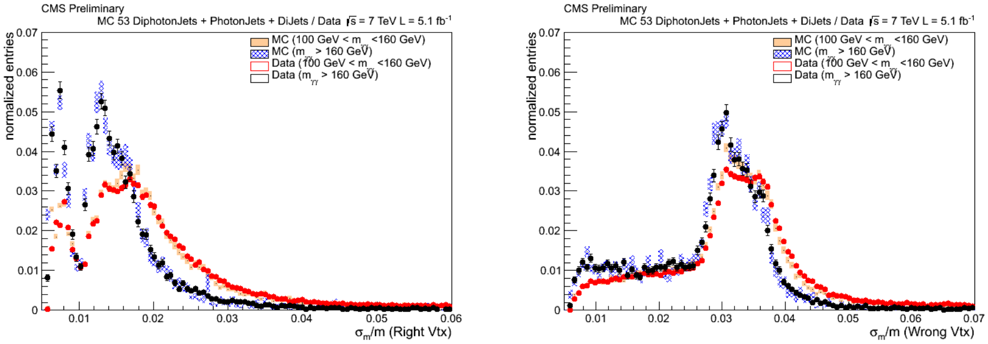


(e) $\cos(\Delta\phi)$

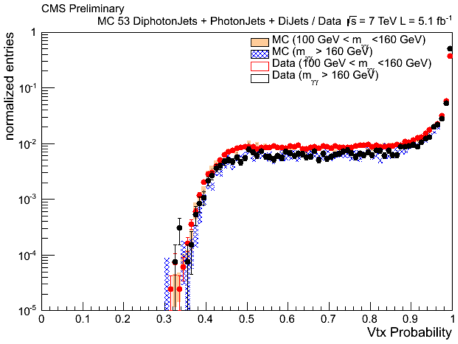
Figure D.1: Kinematic di-photon BDT inputs' validation for 7 TeV analysis.



(a) ID BDT γ_1 (b) ID BDT γ_2

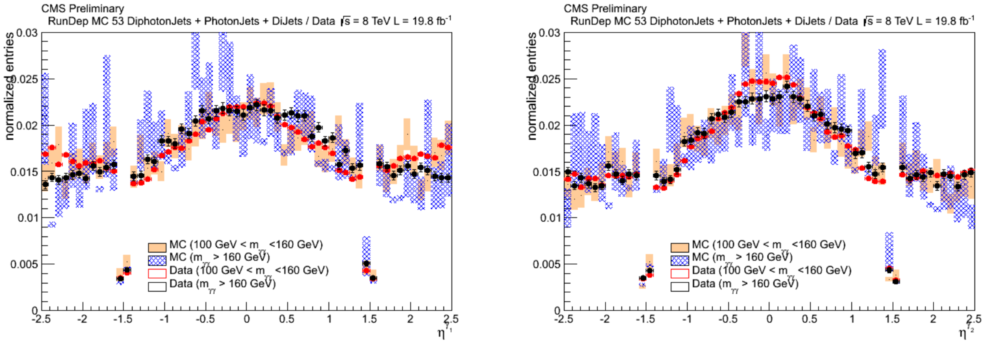


(c) $\sigma_{m,\text{Right}}/m_{\gamma\gamma}$ (d) $\sigma_{m,\text{Wrong}}/m_{\gamma\gamma}$

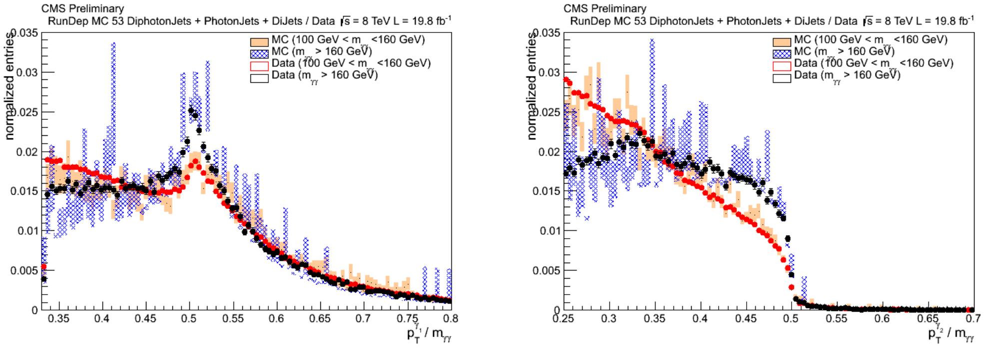


(e) Vertex Probability

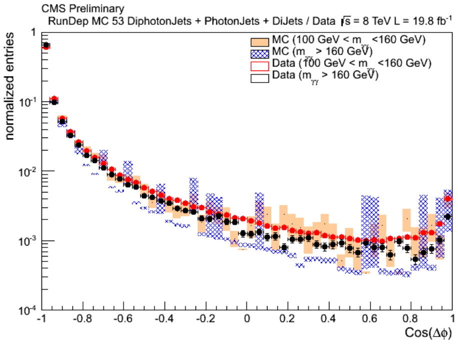
Figure D.2: Other di-photon BDT inputs' validation for 7 TeV analysis.



(a) η_1 (b) η_2

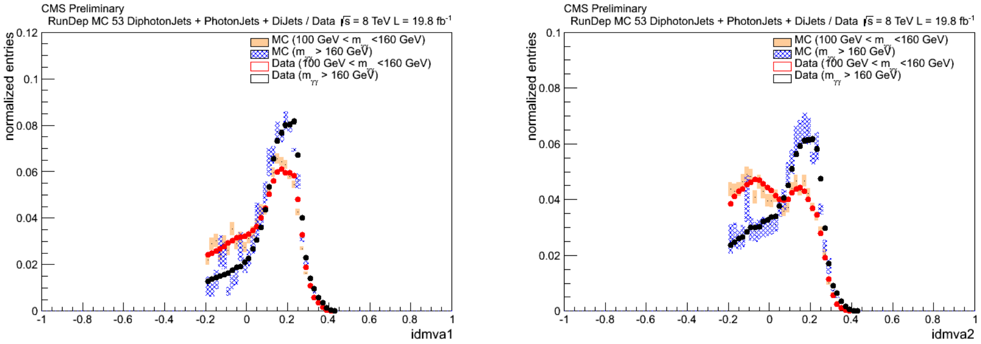


(c) $p_{T,1}/m_{\gamma\gamma}$ (d) $p_{T,2}/m_{\gamma\gamma}$

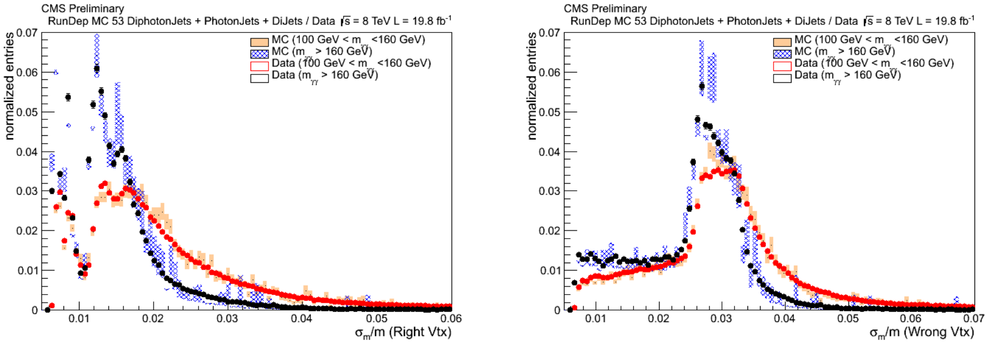


(e) $\cos{(\Delta\phi)}$

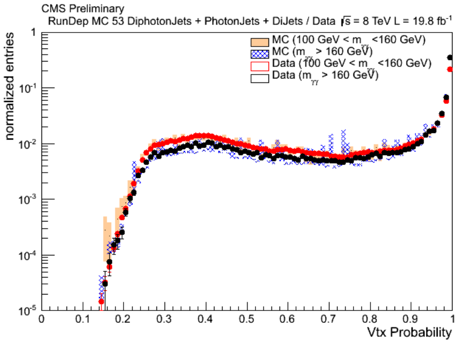
Figure D.3: Kinematic di-photon BDT inputs' validation for 8 TeV analysis.



(a) ID BDT γ_1 (b) ID BDT γ_2

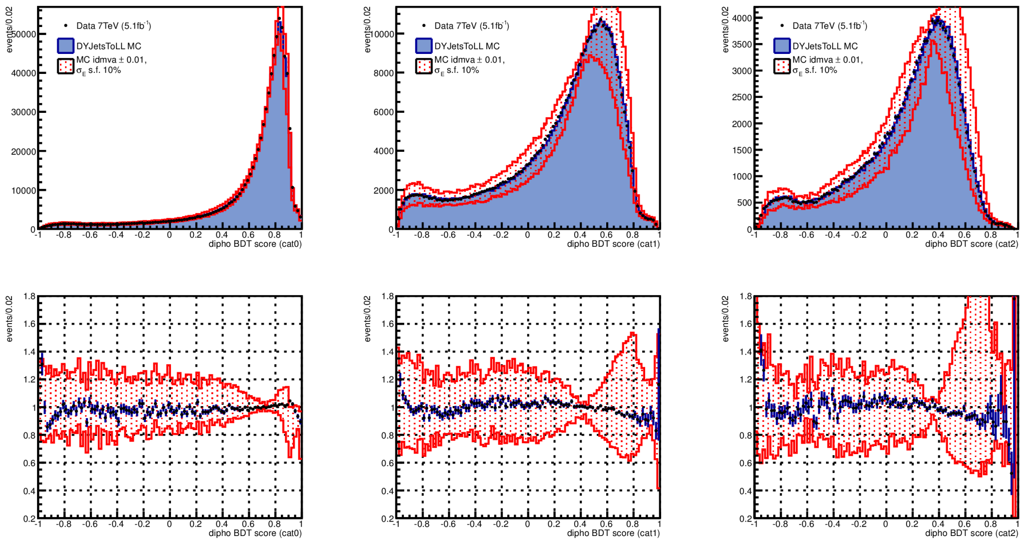


(c) $\sigma_{m\text{Right}}/m_{\gamma\gamma}$ (d) $\sigma_{m\text{Wrong}}/m_{\gamma\gamma}$

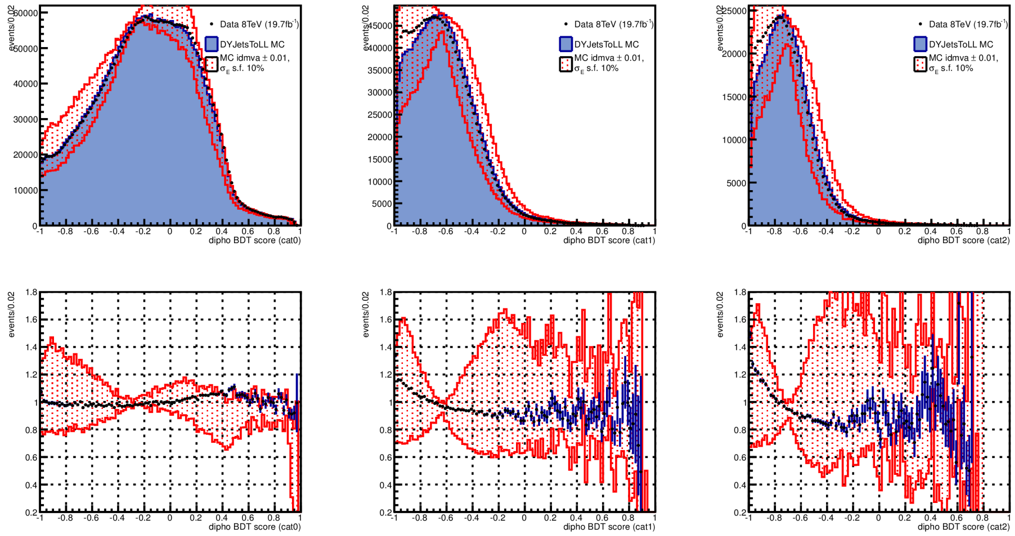


(e) Vertex Probability

Figure D.4: Other di-photon BDT inputs' validation for 8 TeV analysis.



(a) 7 TeV



(b) 8 TeV

Figure D.5: Above is the di-photon BDT output for $Z \rightarrow e^+e^-$ in data and MC with the systematics for the photon ID's and single photon energy resolution propagated to the di-photon BDT. Below shows the data to MC ratio of the plot above. The three columns are events split into categories of both electrons in the EB, one EB and one EE, and both in the EE. In (a) is the validation for the 7 TeV analysis in and in (b) for the 8 TeV analysis. The uncertainty bands generously cover the deviation from 1 in the ratio of data to MC almost everywhere.

Bibliography

[1] F. Halzen and A. D. Martin, "Quarks and Leptons: An Introductory Course in Modern Particle Physics". John Wiley & Sons, 1984.

[2] M. Srednicki, "Quantum Field Theory". Cambridge University Press, 2007.

[3] R. McKeown and P. Vogel, "Neutrino masses and oscillations: triumphs and challenges", *Physics Reports* **394** (2004), no. 6, 315 – 356, arXiv:hep-ph/0402025. doi:10.1016/j.physrep.2004.01.003.

[4] J. Beringer et al., "Review of Particle Physics (RPP)", *Phys.Rev.* **D86** (2012 and 2013 partial update for the 2014 edition) 010001. doi:10.1103/PhysRevD.86.010001.

[5] J. C. Romao and J. P. Silva, "A resource for signs and Feynman diagrams of the Standard Model", *Int.J.Mod.Phys.* **A27** (2012) 1230025, arXiv:1209.6213. doi:10.1142/S0217751X12300256.

[6] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, "Revised value of the eighth-order QED contribution to the anomalous magnetic moment of the electron", *Phys. Rev. D* **77** (Mar, 2008) 053012. doi:10.1103/PhysRevD.77.053012.

[7] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse, "Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment", *Phys. Rev. A* **83** (May, 2011) 052122. doi:10.1103/PhysRevA.83.052122.

[8] Belle Collaboration, "Observation of a Resonancelike Structure in the $\pi^{+-}\psi'$ Mass Distribution in Exclusive $B \to K\pi^{+-}\psi'$ Decays", *Phys. Rev. Lett.* **100** (Apr, 2008) 142001, arXiv:0708.1790. doi:10.1103/PhysRevLett.100.142001.

[9] LHCb Collaboration, "Observation of the resonant character of the $Z(4430)^-$ state", *LHCb Preprint* **LHCB-PAPER-2014-014** (2014) arXiv:1404.1903.

[10] F. L. Wilson, “Fermi’s Theory of Beta Decay”, *American Journal of Physics* **36** (1968), no. 12, 1150–1160. doi:10.1119/1.1974382.

[11] A. Salam and J. Ward, “Weak and electromagnetic interactions”, *Il Nuovo Cimento Series 10* **11** (1959), no. 4, 568–577. doi:10.1007/BF02726525.

[12] S. L. Glashow, “Partial-symmetries of weak interactions”, *Nuclear Physics* **22** (1961), no. 4, 579 – 588. doi:10.1016/0029-5582(61)90469-2.

[13] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323. doi:10.1103/PhysRevLett.13.321.

[14] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. doi:10.1103/PhysRevLett.13.508.

[15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587. doi:10.1103/PhysRevLett.13.585.

[16] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.

[17] F. R. Villatoro, “Higgs potential with Goldstone mode and Higgs mode”, Jul, 2012.

[18] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127** (Aug, 1962) 965–970. doi:10.1103/PhysRev.127.965.

[19] J. Schwinger, “Gauge Invariance and Mass”, *Phys. Rev.* **125** (Jan, 1962) 397–398. doi:10.1103/PhysRev.125.397.

[20] Gargamelle Collaboration, “Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment”, *Physics Letters B* **46** (1973), no. 1, 138 – 140. doi:10.1016/0370-2693(73)90499-1.

[21] UA1 Collaboration, “Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV”, *Physics Letters B* **122** (1983), no. 1, 103 – 116. doi:10.1016/0370-2693(83)91177-2.

[22] UA1 Collaboration, “Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider”, *Physics Letters B* **126** (1983), no. 5, 398 – 410. doi:10.1016/0370-2693(83)90188-0.

[23] UA2 Collaboration, “Evidence for $Z_0 \to e^+e^-$ at the CERN $\bar{p}p$ collider”, *Physics Letters B* **129** (1983) 130 – 140. doi:10.1016/0370-2693(83)90744-X.

[24] R. W. Assmann, M. Lamont, and S. Myers, “A Brief History of the LEP Collider”, *Nucl. Phys. B, Proc. Suppl.* **109** (Apr, 2002) 17–31. doi:10.1016/S0920-5632(02)90005-8.

[25] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, and the LEP Working Group for Higgs Boson Searches, “Search for the Standard Model Higgs boson at LEP”, *Physics Letters B* **565** (2003) 61 – 75, arXiv:hep-ex/0306033v1. doi:10.1016/S0370-2693(03)00614-2.

[26] CDF Collaboration, “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”, *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631, arXiv:hep-ex/9503002v2. doi:10.1103/PhysRevLett.74.2626.

[27] D0 Collaboration, “Search for High Mass Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV”, *Phys. Rev. Lett.* **74** (Mar, 1995) 2422–2426, arXiv:hep-ex/9411001. doi:10.1103/PhysRevLett.74.2422.

[28] Tevatron New Phenomina and Higgs Working Group, CDF Collaboration, D0 Collaboration, “Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data”, arXiv:1107.5518.

[29] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys.Lett.* **B716** (2012) 30–61, arXiv:1207.7235. doi:10.1016/j.physletb.2012.08.021.

[30] D. C. Tevatron New Physics Higgs Working Group, CDF Collaboration, “Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data”, arXiv:1207.0449.

[31] ALEPH Collaboration, CDF Collaboration, D0 Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak and Heavy Flavour Groups, “Precision Electroweak Measurements and Constraints on the Standard Model”, arXiv:1012.2367.

[32] J. L. Feng, “Naturalness and the Status of Supersymmetry”, *Ann.Rev.Nucl.Part.Sci.* **63** (2013) 351–382, arXiv:1302.6587. doi:10.1146/annurev-nucl-102010-130447.

[33] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: evidence, candidates and constraints”, *Physics Reports* **405** (2005), no. 5 - 6, 279 – 390, arXiv:hep-ph/0404175v2. doi:10.1016/j.physrep.2004.08.031.

[34] S. Perlmutter and the Supernova Cosmology Project, “Measurements of Ω and Λ from 42 High-Redshift Supernovae”, *The Astrophysical Journal* **517** (1999) arXiv:astro-ph/9812133. doi:10.1086/307221.

[35] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, “LHC Design Report vol. 1: the LHC Main Ring”, *CERN Report* **CERN-2004-003-V-1** (2004).

[36] T. Taylor, “Incident at the CERN LHC: The event, the repair and lessons to be drawn”, *TEION KOGAKU* **45** (2010) 344–351. doi:10.2221/jcsj.45.344.

[37] AC Team, “Diagram of an LHC dipole magnet. Schéma d’un aimant dipôle du LHC”, June, 1999.

[38] N. C. Lasheras, W. V. Delsolaro, A. Siemko, R. Ostojic, and G. Kirby, “Quench Performance of the LHC Insertion Magnets”, *IEEE Trans. Appl. Supercond.* **19** (2009), no. 3 (part 2), 1172–1175. doi:10.1109/TASC.2009.2017708.

[39] R. Ostojic, “The LHC Insertion Magnets”, *IEEE Trans. Appl. Supercond.* **12** (March, 2002) 196–201. doi:10.1109/TASC.2002.1018382.

[40] R. Pinget and A. Macpherson, “LHC Performance and Statistics”. http://cern.ch/lhc-statistics.

[41] CMS Collaboration, “CMS Luminosity - Public Results”. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[42] CMS Collaboration, “The CMS magnet project: Technical Design Report”, *Scientific Committee Paper* **CERN-LHCC-97-010; CMS-TDR-1** (1997).

[43] CMS Collaboration, “Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays”, *JINST* **5** (Oct, 2009) T03021. 37 p, arXiv:0910.5530. doi:10.1088/1748-0221/5/03/T03021.

[44] W. Bertl, K. Deiters, Q. Ingram, D. Renker, T. Sakhelashvili, and E. Frlez, “Feasibility of Intercalibration of CMS ECAL Supermodules with Cosmic Rays”, *Eur.Phys.J.* **41** no. 2, 11–17. doi:10.1140/epjcd/s2005-02-007-y.

[45] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model”, *Phys.Rept.* **459** (2008) 1–241, arXiv:hep-ph/0503173. doi:10.1016/j.physrep.2007.10.005.

[46] M. Pieri, S. Bhattacharya, I. Fisk, J. Letts, V. Litvin and J. Branson, “Inclusive Search for the Higgs Boson in the $H \to \gamma\gamma$ Channel”,.

[47] CMS Collaboration, “CMS Physics Technical Design Report, Volume II: Physics Performance”, *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 6, 995. doi:10.1088/0954-3899/34/6/S01.

[48] CMS Collaboration, “Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s}$=7 TeV”, *Phys. Lett. B* **710** (Feb, 2012) 403–425, arXiv:1202.1487. doi:10.1016/j.physletb.2012.03.003.

[49] CMS Collaboration, “Search for a fermiophobic Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **1209** (2012) 111, arXiv:1207.1130. doi:10.1007/JHEP09(2012)111.

[50] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt, “Parton distributions for the LHC”, *TEur.Phys.J.* **C63** (2009) 189–285, arXiv:0901.0002. doi:10.1140/epjc/s10052-009-1072-5.

[51] M. Sani, J. Branson, M. Pieri, “”Cuts in Categories” (CiC) Electron Identification”, Jul, 2011.

[52] E. Gross and O. Vitells, “Trial factors for the look elsewhere effect in high energy physics”, *Eur.Phys.J.* **70** (2010), no. 1-2, 525–530. doi:10.1140/epjc/s10052-010-1470-8.

[53] CMS Collaboration, “Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys.Lett.* **B710** (2012) 26–48, arXiv:1202.1488. doi:10.1016/j.physletb.2012.02.064.

[54] ATLAS Collaboration, “Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC”, *Phys.Lett.* **B710** (2012) 49–66, arXiv:1202.1408. doi:10.1016/j.physletb.2012.02.044.

[55] CMS Collaboration, “A search using multivariate techniques for a standard model Higgs boson decaying into two photons”, *CMS Physics Analysis Summary* **CMS-PAS-HIG-12-001** (2012).

[56] CMS Collaboration, “Search for the fermiophobic model Higgs boson decaying into two photons”, *CMS Physics Analysis Summary* **CMS-PAS-HIG-12-002** (2012).

[57] CERN Video Productions, “4th July 2012, Seminar at CERN. Update on the Higgs Boson searches at the LHC.”. http://cds.cern.ch/record/1459513, July, 2012.

[58] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”,

Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.

[59] CMS Collaboration, “Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs”, *Phys. Rev. Lett.* **110** (Feb, 2013) 081803. doi:10.1103/PhysRevLett.110.081803.

[60] ATLAS Collaboration, “Evidence for the spin-0 nature of the Higgs boson using ATLAS data”, *Phys.Lett.* **B726** (2013) 120–144, arXiv:1307.1432. doi:10.1016/j.physletb.2013.08.026.

[61] CMS Collaboration, “Evidence for the direct decay of the 125 GeV Higgs boson to fermions”, *Nature Physics* (2014) arXiv:1401.6527.

[62] M. D. et al., “Les Houches guidebook to Monte Carlo generators for hadron collider physics”, arXiv:hep-ph/0403045.

[63] J. Allison et al., “Geant4 developments and applications”, *IEEE Transactions on Nuclear Science* **53** (Feb, 2006) 270–278. doi:10.1109/TNS.2006.869826.

[64] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1”, *JHEP* **0902** (2009) 007, arXiv:0811.4622. doi:10.1088/1126-6708/2009/02/007.

[65] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, Michel, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, “MadGraph/MadEvent v4: The New Web Generation”, *JHEP* **0709** (2007) 028, arXiv:0706.2334. doi:10.1088/1126-6708/2007/09/028.

[66] T. Sjostrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175. doi:10.1088/1126-6708/2006/05/026.

[67] N. Kauer and G. Passarino, “Inadequacy of zero-width approximation for a light Higgs boson signal”, *JHEP* **1208** (2012) 116, arXiv:1206.4803. doi:10.1007/JHEP08(2012)116.

[68] N. Kauer, “Inadequacy of zero-width approximation for a light Higgs boson signal”, *Mod.Phys.Lett.* **A28** (2013) 1330015, arXiv:1305.2092. doi:10.1142/S0217732313300152.

[69] CMS Collaboration, “2012 ECAL detector performance plots”, *CMS Detector Performance Summary* **CMS-DP-2013-007** (Mar, 2013).

[70] CMS Collaboration, “Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV”, *JINST* **8** (June, 2013) P09009. doi:10.1088/1748-0221/8/09/P09009.

[71] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, “TMVA: Toolkit for Multivariate Data Analysis”, *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[72] J. E. Gaiser, “Appendix-F Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime”, *Ph.D. Thesis* **SLAC-R-255** (1982).

[73] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (April, 2009).

[74] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys.Lett.* **B659** (2008) 119–126, arXiv:0707.1378. doi:10.1016/j.physletb.2007.09.077.

[75] D. Rainwater, R. Szalapski, and D. Zeppenfeld, “Probing color-singlet exchange in Z + two jet events at the CERN LHC”, *Phys. Rev. D* **54** (Dec, 1996) 6680–6689. doi:10.1103/PhysRevD.54.6680.

[76] S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka (editors) and LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group”, *CERN Report* **CERN-2013-004** (2013), no. arXiv:1307.1347. CERN-2013-004, arXiv:1307.1347.

[77] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013), no. 04, P04013, arXiv:1211.4462. doi:10.1088/1748-0221/8/04/P04013.

[78] S. Seltzer and M. Berger, “Transmission and reflection of electrons by foils”, *Nuclear Instruments and Methods* **119** (1974), no. 0, 157 – 176. doi:http://dx.doi.org/10.1016/0029-554X(74)90747-2.

[79] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *JINST* **6** (2011) P11002, arXiv:1107.4277. doi:10.1088/1748-0221/6/11/P11002.

[80] The ATLAS and CMS Collaborations and The LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, *CMS Note* **CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11** (Aug, 2011).

[81] A. L. Read, “Modified frequentist analysis of search results (the CL_s method)”, 2000.

[82] G. Cowan, K. Cranmer, E. Gross et al., “Asymptotic formulae for likelihood-based tests of new physics”, *Eur.Phys.J.* **C71** (2011) 1554, arXiv:1007.1727. doi:10.1140/epjc/s10052-011-1554-0.

[83] L. D. Landau, “On the angular momentum of a two-photon system”, *Dokl. Akad. Nauk. Ser. Fiz.* **60** (1948) 207.

[84] C. N. Yang, “Selection Rules for the Dematerialization of a Particle Into Two Photons”, *Phys. Rev.* **77** (1950) 242. doi:10.1103/PhysRev.77.242.

[85] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, and N. V. Tran, “Spin determination of single-produced resonances at hadron colliders”, *Phys. Rev. D* **81** (2010) 242–245, arXiv:1001.3396.

[86] J. C. Collins and D. E. Soper, “Angular distribution of dileptons in high-energy hadron collisions”, *Phys. Rev. D.* **16** (1977) 2219–2225. doi:10.1103/PhysRevD.16.2219.

[87] CMS Collaboration, “Search for the standard model Higgs boson in the dimuon decay channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *CMS Physics Analysis Summary* **CMS-PAS-HIG-13-007** (2013).

[88] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, *JHEP* **04** (2009) 002, arXiv:0812.0578. doi:10.1088/1126-6708/2009/04/002.

[89] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”, *JHEP* **02** (2010) 037, arXiv:0911.5299. doi:10.1007/JHEP02(2010)037.

[90] CMS Collaboration, “Electron performance with 19.6 fb^{-1} of data collected at $\sqrt{s} = 8$ TeV with the CMS detector.”, *CMS Detector Performance Summary* **CMS-DP-2013-003** (Mar, 2013).

[91] CMS Collaboration, “Single Muon efficiencies in 2012 Data”, *CMS Detector Performance Summary* **CMS-DP-2013-003** (Mar, 2013).

[92] CMS Collaboration, “Properties of the Higgs-like boson in the decay $H \to ZZ \to 4\ell$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *CMS Physics Analysis Summary* **CMS-PAS-HIG-13-002** (2013).

[93] CMS Collaboration, “Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons”, *JHEP* (2014) arXiv:1401.5041.

[94] ATLAS Collaboration, “Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data”, *ATLAS Conference Note* **ATLAS-CONF-2011-163** (Mar, 2014).

[95] CMS Collaboration, “Measurement of the properties of a Higgs boson in the four-lepton final state”, *Phys. Rev. D* (2014) arXiv:1312.5353.

[96] F. Caola and K. Melnikov, “Constraining the Higgs boson width with ZZ production at the LHC”, *Phys.Rev.* **D88** (2013) 054024, arXiv:1307.4935. doi:10.1103/PhysRevD.88.054024.

[97] CMS Collaboration, “Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs”, *Submitted to Physics Letters B* (2014) arXiv:1405.3455.

[98] CMS Collaboration, “Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states”, *JHEP* **01** (2014) 1–86, arXiv:1312.1129. doi:10.1007/JHEP01(2014)096.

[99] CMS Collaboration, “Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks”, *Phys. Rev. D* **89** (Jan, 2014) 012003, arXiv:1310.3687. doi:10.1103/PhysRevD.89.012003.