Properties of the Higgs-like state around 125 GeV in its decay to two photons at the CMS experiment at the LHC

Matthew Kenzie
Imperial College London

A dissertation submitted to Imperial College London
for the degree of Doctor of Philosophy

DRAFT

DRAFT

Abstract

Some stuff here

DRAFT

DRAFT

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. Figures from other sources are labelled as CMS and the source is referenced in the figure captions. This dissertation does not exceed the word limit for the respective Degree Committee.

Matthew Kenzie

DRAFT

DRAFT

Acknowledgements

Acknows

DRAFT

DRAFT

Preface

Blah blah blah

DRAFT

DRAFT

Contents

DRAFT

DRAFT

"For my father."

—

DRAFT

DRAFT

Chapter 1.

Introduction and Theory

"All science is either physics or stamp collecting."

— Ernest Rutherford

1.1. Description

Motivation for Higgs searches and the theory of the Standard Model. In particular, the electroweak sector and symmetry breaking.

15 pages

Some things. Ref example [1–3].

1.2. The Standard Model

The **sm!** (**sm!**) is very successful.

$$m_{\gamma\gamma} = 2\sqrt{E_{\gamma 1} E_{\gamma 2}(1 - \cos(\alpha))} \tag{1.1}$$

DRAFT

DRAFT

Chapter 2.

The CMS experiment

"You need something to open up a new door,
To show you something you seen before,
But overlooked a hundred times or more."
— Bob Dylan

2.1. The LHC

The Large Hadron Collider (LHC) is an octagonal 27km ring (large) proton-proton (hadron) particle collider. Using a multistage acceleration process two beams of protons are circulated in opposite directions at a centre-of-mass energy, $\sqrt{s}$=7 TeV (8 TeV) for data collection in 2011 (2012). The beams of protons are accelerated and circulated by electric and magnetic fields respectively. Further precision magnetic fields can control the position and intensity of the beams. There are four points around the ring where the beams can be forced to intersect producing high energy proton-proton collisions. Particle detectors are constructed around these points such that the collision can be reconstructed with the purpose of measuring physical properties and processes, calibrating the detectors with already known processes and searching for new physics. The remainder of this chapter concentrates on a description of one of these detectors, the Compact Muon Solenoid (CMS), which the author has worked on.

DRAFT

2.2. The CMS detector

The CMS detector, pictured in Fig. 2.1, is a multipurpose experiment designed for the measurement of and search for a multitude of different processes. We will primarily discuss its function as a Higgs finding machine, a more detailed description can be found in Ref. [4]. It has a cylindrical shape consisting of a barrel segment, 21.6m long, and two endcaps, 14.6m in diameter, aligned along the beam direction with its centre at the beam interaction point. The endcaps are those nearer the beam line and so the materials in these components typically have to be able to withstand higher amounts of radiation and therefore tend to have worse performance. Many of its features exploit what one would expect for measuring Higgs decays: it has almost full coverage of the area around the collision point so that nearly every particle emanating from the collision can be reconstructed, it has many complimentary subsystems (or layers) designed to measure different specific particles so that Higgs bosons can be detected through a multitude of decay modes.

For a Higgs with an intermediate mass (100 - 200 GeV) the high resolution (narrow peak) channels are $H \to ZZ* \to l^+l^-l^+l^-$[1] and $H \to \gamma\gamma$ so good energy resolution and identification of electrons and muons is desirable down to very low p_{T} ($\sim O(10\mathrm{GeV})$ as well as good resolution and identification of high energy photons.

The central design feature of CMS is the very powerful superconducting magnet which produces an axial magnetic field of 4T. The size of this field, as well as the density of the calorimeter materials, allows for a compact and economical design (much more so than its sister detector, ATLAS). Outside of the magnet lie the muon stations which also serve as a return yoke for the magnetic field. The muon chambers in the barrel consist of alternating layers of drift tubes and restive plate chambers which provide both accurate timing and hit location, in order to reconstruct muons down to low energies. In the endcap the drift tubes are replaced with cathode strip chambers. Combining information from the muon subsystem with information from the inner tracking system (described below) allows muons at CMS to be reconstructed down to $p_T < 10$ GeV with a resolution of $\sim 1\%$. The other three main subsystems at CMS, the tracking system and the calorimeters, are located inside the magnetic field.

The first layer is the tracking system which is used to reconstruct the momentum of any outgoing charged particles and to locate the primary and secondary vertices.

[1] A * denotes that one Z can be off mass shell

DRAFT

This is surrounded by the calorimeters, the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The first is a single layer of dense, transparent crystals which collects deposits of energy left by electrons and photons which shower inside the material. The second compliments this by providing a measurement of the energy deposited by hadrons (reconstructed as objects known as jets) through nuclear interactions. The HCAL is a sampling calorimeter in which the active material (plastic scintillator) is sandwiched between a dense absorbent material (brass or steel). This extends the radiation length of the calorimeter (clearly accommodating the compact design) and provides pointing information but degrades the resolution of reconstructing jets.

CMS uses a right-handed Cartesian coordinate system with the origin at the interaction point and the z-axis pointing along the beam axis. The x-axis points towards the centre of the LHC ring and the y-axis points vertically upwards. The azimuthal angle, $\phi \in [-\pi, \pi]$, is defined with respect to the x-axis in the transverse $(x-y)$ plane. The polar angle θ is measured from the z-axis. Commonly, the direction of an outgoing particle is defined by ϕ and its pseudo-rapidity η,

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \tag{2.1}$$

The LHC is capable of producing 40M bunch collisions per second although many of these are not hard interactions, the result being that the outgoing particle debris follows the beam line. A hard (and therefore interesting) collision is characterised by the amount of energy produced in the transverse $(x-y)$ plane. Therefore particles are commonly characterised by the projection of their momentum onto this plane, their transverse momentum,

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{2.2}$$

and the corresponding transverse energy, $E_T = E\sin(\theta)$.

DRAFT

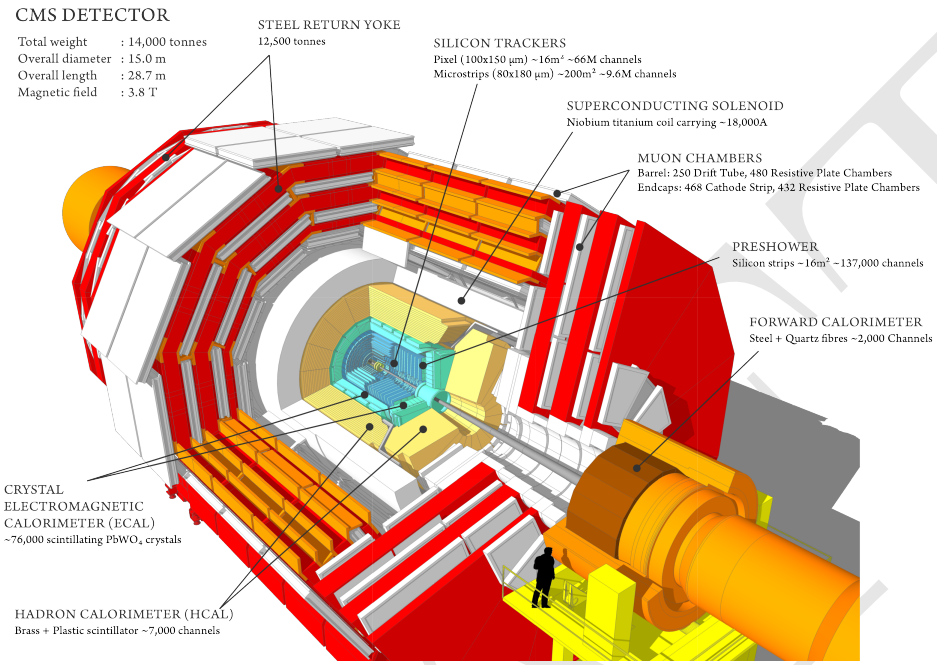

Figure 2.1.: This is a schematic representation of CMS showing the layered structure of subdetectors; tracking system, calorimeters, magnet and muon system.

2.2.1. Tracking system

During nominal LHC running conditions in 2012 there on average over 1000 particles from up to 50 overlapping *p*-*p* collisions (pileup) per bunch crossing (every 50 ns). The tracker is designed to efficiently and precisely reconstruct all charged particle trajectories, thus their position and momentum, which are known as tracks. Due to the vast number of tracks eminating from multiple vertices in typical LHC collisions the tracking material and electronics are required to have high granularity, a fast response and be radiation hard. This conflicts with another important design feature of the tracker which is the aspiration to use the minimal amount of material in order to reduce multiple scattering, bremsstrahlung, photon conversion and nuclear interactions before particles reach the calorimeters. These criteria motivate the choice of silicon throughout the CMS tracking system. The structure of the CMS tracker is shown in Figure 2.2 and consists of a central pixel detector surrounded by layers of silicon strips aligned parallel to the beam line in the barrel (TIB and TOB) and perpendicular to the beam line in the endcap (TID and TEC).

By making multiple precise measurements of tracks as they pass through the pixel and silicon layers (hits) the track trajectory can be reconstrcuted and their momentum

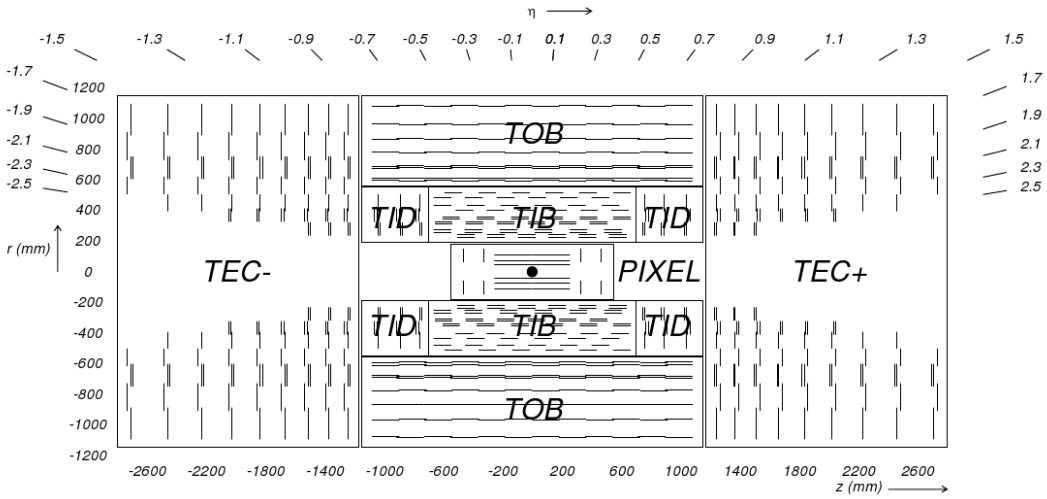

Figure 2.2.: A diagram of the CMS tracking system showing the PIXEL detector and outer silicon strip layers [4].

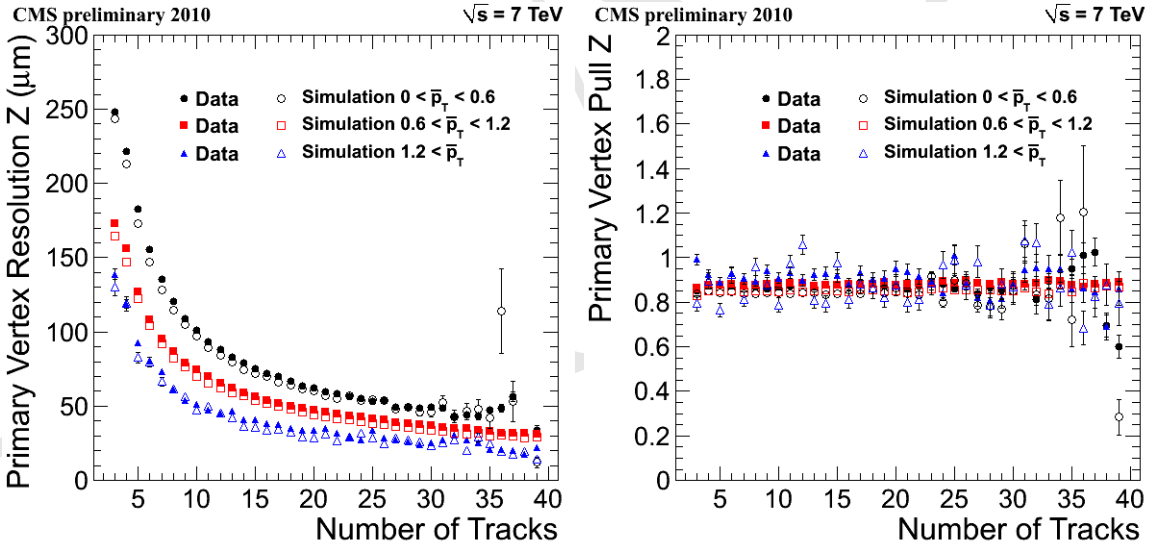

Figure 2.3.: A demonstration of the vertex position resolution in Z as a function of the number of tracks originating from that vertex. The different colors respresent three different bins in the average track p_{T} where data is shown as solid points and simulation as open points [5]

calculated using their curvature in the ϕ plane due to the axial magentic field. Tracks are grouped together (requiring that their separtion is less than 1cm in the z coordinate at the point of closest approach to the beamline) and assigned to a common point or origin (the primary vertex). The vertex resolution is driven both by the number of tracks originating from a particular vertex and how large their average p_{T} is. This is shown in Fig. 2.3 for preliminary data taken in 2010 at $\sqrt{s} = 7$ TeV.

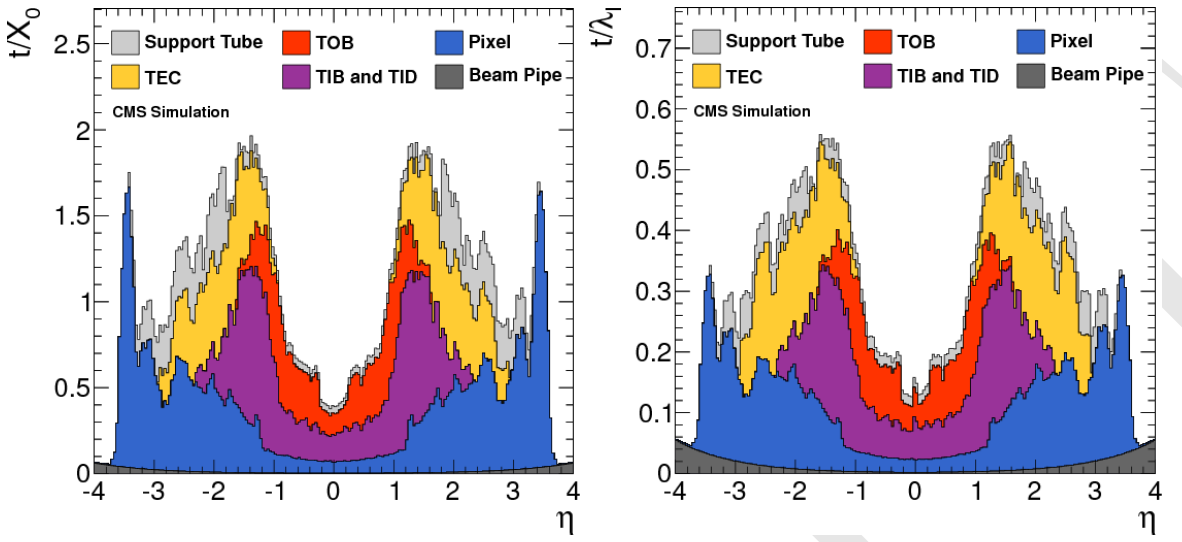

Figure 2.4.: The amount of CMS tracker material in radiation lengths (X_0) on the left and in nuclear interaction lengths (λ_I) on the right as a function of η for the different tracking subsytems.

The material budget for the tracker is shown in Fig. 2.4 which demonstrates as a function of η which subsystems of the tracker, beam pipe and servicing contrubute to material inbetween the interaction point and the calorimeters in radiation lengths (X_0) and nuclear interaction lengths (λ_I). As shown in Fig. 2.2 the tracker has full coverage in ϕ and up to $|\eta| \leq 2.5$. As we will see later the tracking system is very important for the $H \to \gamma\gamma$ search at CMS as without it locating the primary vertex would be practically impossible.

2.2.2. Electromagnetic calorimeter

The ECAL is used to reconstruct the energy of electrons and photons which deposit their energy via electromagnetic showers inside the calorimeter material. The shower inside the crystal produces photons whose total energy is proportional to the energy of the incoming particle, hence the light output from the shower can be measured by photodiodes at the back of each crystal which in turn provides a measurement for the energy of the original particle. The ECAL has full hermetic coverage of the interaction point and consists of a single layer of lead tungstate ($PbWO_4$) crystals. The crystals are laid out in a quasi-projective geometry such that they point towards the interaction point with an offset of 3° making it much less likely that a photon, or electron, will pass straight through a gap between crystals. The ECAL consists of a barrel section and two

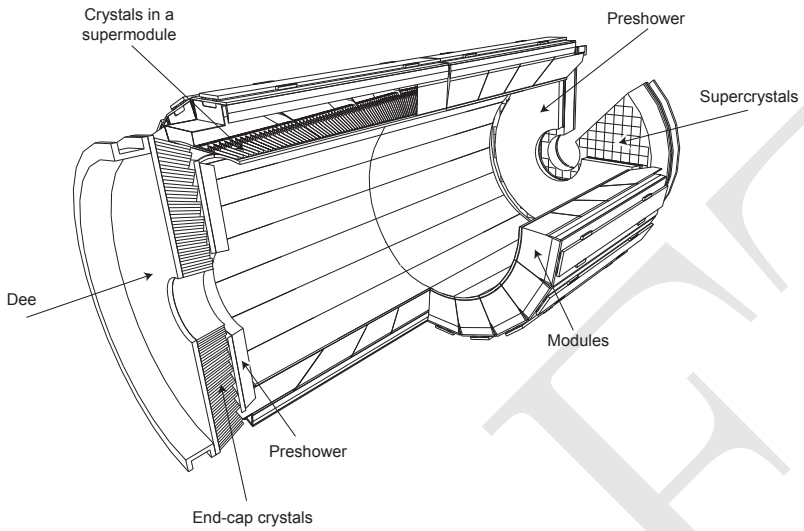

Figure 2.5.: A schematic drawing of the CMS ECAL layout showing the "modules" of crystals in the ECAL barrel and the "Dees" of crystals in the ECAL endcap [4].

endcap disks which are preceded by a preshower (to aid with π^0 rejection), a schematic is shown in Fig. 2.5. There is a fiducial region between the barrel and endcap (to prevent reconstruction of showers which overlap both subsystems) yielding an ECAL coverage of $|\eta| < 2.5$ but not in the range $1.444 < |\eta| < 1.556$ and full coveraege in ϕ. The crystals in both sections have a depth of 23cm (which amounts to 25.8 X_0) implying that practically the whole depth of the shower is contained within this single layer.

$PbWO_4$ has some attractive properties for an electromagnetic calorimeter especially when considering $H \to \gamma\gamma$ decays. In order to acheive good energy resolution it is desirable to have a design in which most of the electromagentic shower from an incoming photon or electron is contained within a single crystal, as less of the shower (and therefore energy) is lost in cracks and gaps inbetween crystals. $PbWO_4$ has a short radiation length ($X_0 = 0.89$cm) which compliments a compact design as the entire depth of the shower can be contained in a crystal which is not very long. It has a small Molière radius (1.96cm) which means that the lateral size of the shower is small. This generally means that the cross-sectional size of a crystal can be small (yielding high granularity of the detector) whilst still containing a large percentage of the shower. It has a very short scintillation time decay constant (85% of the light is collected in 25ns), in other words it is very "fast", allowing the energy in the shower to be collected and measured very quickly. This is clearly desirable in an environment like the LHC when collisions are happening up to every 25ns. The one draw back of $PbWO_4$ (apart from expense) which has crippled its use in detectors previously is its incredibly low light yield at room

temperature ($\sim$50-80 photons/MeV). This is overcome at CMS with the use of silicon avalance photodiodes (APDs) in the ECAL barrel and vacuum phototriodes (VPTs) in the endcap, which amplify the signal enough to make accurate measurements of the original photons energy.

Energy resolution

As we have seen previously (assume this formula will appear in the Introduction and can be referenced) the diphoton invariant mass is given by,

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1-\cos\alpha)}, \tag{2.3}$$

where E_1 and E_2 are the energy of the two photons and α is the angle between them. Therefore the mass resolution has terms that depend on the photon energy resolution and angular resolution,

$$\frac{\sigma_M}{M} = \frac{1}{2}\left[\frac{\sigma_{E_1}}{E_1} \oplus \frac{\sigma_{E_2}}{E_2} \oplus \frac{\sigma_\alpha}{\tan(\alpha/2)}\right], \tag{2.4}$$

where σ denotes the resolution and $\oplus$ the quadratic sum. It is therefore desirable to have both good energy resolution and good position resolution for photons (accurate position measurements at the ECAL face alongside knowledge of the primary vertex can be used to calculate the individual photons' direction and ergo the angle between them). The energy resolution is usually then further parametrised as,

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{2.5}$$

where a is the stochastic term, b the noise term and c a constant term. In order to acheive the best possible resolution all three of these terms need to be of a similar order and as small as possible. The size of these terms has been determined from test beam data in Ref. [4]. The stochastic term is driven by the material choice and detector type so cannot

DRAFT

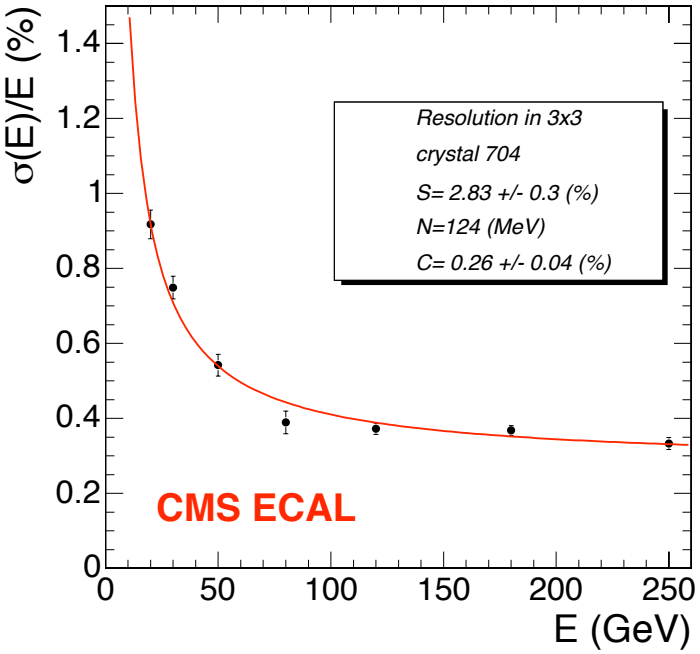

Figure 2.6.: The ECAL energy resolution, σ_E/E, as a function of electron energy measured from test beam data. The energy is measured in a 3×3 array of crystals centered on crystal of electron impact [4].

be improved once the machine is built. The main contributions to this term are lateral shower containment fluctuations, photostatistics and fluctuations in the energy deposited in the preshower absorber. For a fully active calorimeter (the CMS ECAL is not a sampling calorimeter) made of $PbWO_4$ the size of this term is good ($a = 2.8 \pm 0.3\%\text{GeV}^{\frac{1}{2}}$). The constant term, which depends on non-uniformity of longitudal light, intercalibration errors and energy leakage from the back of the crystal, can be minimzed by use of *in situ* calibration of individual crystals and amounts to $c = 0.26 \pm 0.05\%$. However the $H \to \gamma\gamma$ analysis at CMS dispenses with the intercalibration constants by using a regression technique to estimate the photon energy which is discussed in Section 3.3. The noise term which has contributions from electronics noise (including signal digitisation) and event pile-up (additional particles causing overlapping signals) is measured as $b = 126\text{MeV}$. The ECAL energy resolution, σ_E/E, as a function of electron energy is shown in Fig. 2.6 as measured from a beam test.

Transparency corrections

Due to the high particle flux present at CMS the ECAL crystals and electronics have to be radiation hard, especially in the endcap. This is another motivating factor for using $PbWO_4$ as the crystal material. The crystals are additionally doped with Nb to improve the induced absorbtion coefficient. Over time and long exposure to radiation the crystals loose their transparency, although there is considerable natural recovery during down periods. An important part of the ECAL monitoring and calibration comes in the form of transparency corrections to compensate for these losses. At regular intervals during LHC running laser pulses are injected into the crystals to measure the crystal response. Two different wavelengths of laser are used, one blue ($\lambda = 440$nm) which is very similar to the scintillation emission peak and therefore expected to be affected by transparency changes in a similar way to typical scinitillation light, and one red ($\lambda = 796$nm) which is far from the scinitillation emission peak and affected very little by the changes in transparency. Hence, by comparing the red and blue laser light response, time and η dependent corrections for crystal transparency loss can be calculated. A closure test for these corrections, in 2012 data, is shown in Fig. 2.7 which shows the ratio of electron energy (calculated from the ECAL) to electron momentum (calculated from the tracker) before and after transparency (or laser) corrections.

Preshower

The dominant source of background to high energy photon signals are neutral mesons, mainly pions (π^0), which decay into two approximately collinear photons and can therefore look very much like a single high energy photon. The ECAL endcap is preceeded by a preshower to specifically target this and provide the endcap with a higher granularity. The preshower is a sampling calorimeter which consists of two layers: a lead plate, to initiate the shower, in front of a fine grained silicon detector which has two layers of orthogonal strips. There are many other characteristics of π^0 decays which can help in differentiating them from real ("prompt") photons, these include isolation (discussed later in this Chapter in Section 2.4) and the shower shape (discussed in Chapter 4 in Sections 4.1.1 and **??**).

DRAFT

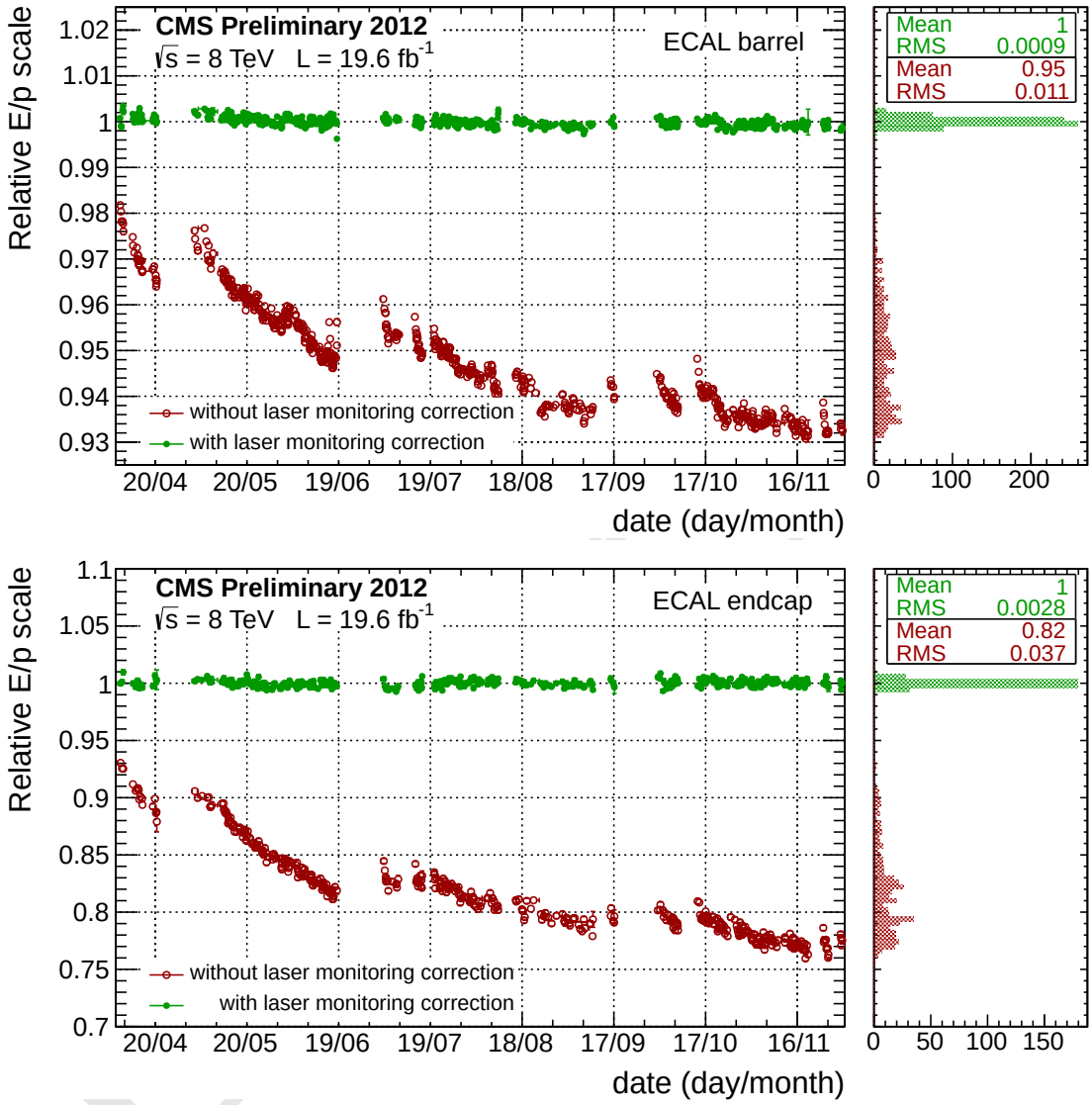

Figure 2.7.: Ratio of the electron energy, E, to the electron momentum, p, measured in the CMS barrel (top) and endcap (bottom) for 2012 data. The red (green) points show the performance before (after) the laser monitoring derived corrections [6].

Photon reconstruction

Calculating an incoming photons energy amounts to summing the energy deposited by the electromagentic shower which is initiated by the photon impact at the crystal face. Due to the presence of material in the beam pipe and tracking system (see Fig 2.4) about 60% of photons will convert into an electron-positron pair before they reach the ECAL. If this is the case the shower will spread out in ϕ due to the presence of the magnetic field; the photon converts to electrons which bend and Bremsstrauhlung radiate additional photons, and can be distributed among multiple (up to hundreds of) crystals. Consequently clustering (pattern matching) algorithms are deployed to calculate the "raw" photon energy. Corrections to this energy are subsequently applied to account for any energy loss as explained in Section 3.3. The shower will appear as a local maxima amongst a spatial neighbourhood of crystal energy deposits and so the algorithms used search first for the most energetic crystals (known as the "seed" crystals) and then extend to amass as a large a fraction as possible of the original shower energy. There are three cases to include which are i) photons which reach the ECAL without interacting with any of the intermediate material in the beam pipe and tracking system (referred to as unconverted photons), ii) photons which convert into an electron-positron pair inside the tracker and shower in the barrel, iii) photons which convert and shower in the endcap.

We will first consider the case of converted photons in the barrel. This is so similar to the case of a real electron that the identical algorithm is used for electrons as well. The method used is known as the "Hybrid" algorithm, depicted in Fig. 2.8, which makes clusters of clusters known as a "supercluster": a cluster being a set of crystals which pick up an electron or a bremsstraulunged photon and a cluster of clusters being a set of these which make up all the electrons and photons radiated from the original object. The algorithm can be described as a five step process as follows,

1. Locate the seed crystal which is the maximum energy crystal in the search region, not already in a cluster, and which must satisfy the threshold condition, $E_T > 1$ GeV.

2. Extend in η to construct a "domino" which is 1×3 crystals in $\phi \times \eta$. If the energy of the central crystal in the 1×3 domino is greater than 1 GeV then extend this to a 1×5 domino in $\phi \times \eta$.

3. Traverse along ϕ, up to a maximum of 17 crystals in both directions, adding dominoes in this way. If a domino has less energy than 0.1 GeV then it is excluded.

DRAFT

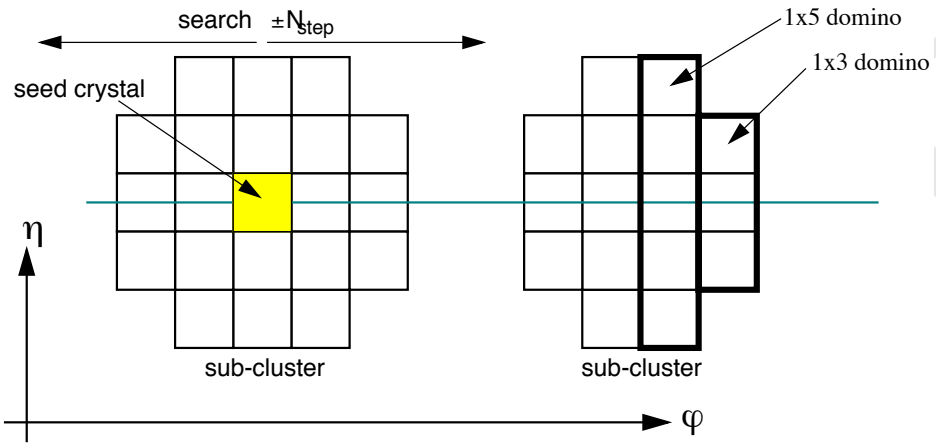

Fig. 6: Domino construction step of Hybrid algorithm

Figure 2.8.: The domino construction setup of the hybrid clustering algorithm [7].

4. Of the clusters of adjacent dominoes the seed domino (most energetic) must have energy greater than 0.35 GeV.

5. Repeat starting from Step 1 to build a new supercluster.

In this way a collection of dominoes are clusted in ϕ creating a "supercluster" of smaller clusters (which has a maximum of area of 5×35 in $\eta \times \phi$). A similar but slightly modified algorithm is used for photons and electrons in the endcap. This is known as the "Multi5×5" algorithm and proceeds,

1. Locate the seed crystal which is the maximum energy crystal in the search region, not already in a cluster, and which must satisfy the threshold condition, $E_T > 180$ MeV.

2. The 5×5 array of crystals surrounding the seed are summed to make a cluster.

3. Crystals at the edge of the cluster can seed new overlapping 5×5 clusters if they are local maxima compared to their neighbouring crystals.

4. Proceed in this way in any direction building an overlapping collection of 5×5 clusters to create a supercluster.

5. Repeat starting from Step 1 to build a new supercluster.

The final case concerns photons which reach the ECAL without converting which have a shower that is much more localised in η and ϕ, about 94% of its energy is deposited in an area of 3×3 crystals and more than 97% in an area of 5×5 crystals. This provides a definition for the conversion variable, R_9, such that,

DRAFT

$$R_9 = \frac{E_{3\times3}}{E_{SC}} \begin{cases} \text{unconverted if } R_9 \geq 0.94 \\ \text{converted otherwise.} \end{cases} \tag{2.6}$$

If a photon fufills the unconverted requirement given above (Eq. 2.6) the photon energy is reconstructed as the sum of all energy deposited in the 5×5 array of crystals which surround the most energetic crystal. It has been shown that using a fixed window for non-converting photons yields a better energy resolution than any clustering procedure [7].

The location of a supercluster is determined as the energy weighted mean position of the crystals in the supercluster. This gives a position resolution which is much smaller than the size of an individual crystal ($\sim 20 \times 20\text{mm}^2$). Values for an electron of p_{T} =35 GeV in the ECAL barrel, in the absence of pileup, are $\sigma_\eta = 1.0 \times 10^{-3}$, $\sigma_\phi = 1.6$ mrad.

Electron and photon differences

It is clear that when only considering the ECAL there is no difference between electrons and photons. Consequently at the level of the calorimetry there is no distinction between them, simply the idea of a supercluster which can apply to both. In the case of an electron, information from the tracker can be included using a Gaussian sum filter algorithm [8], where a series of compatible track hits are associated to the supercluster. This is used to provide a supplementary measurement of the electrons momentum and consequently improve the energy resolution of electrons. When considering photons for an analysis an electron veto must be applied requiring that no track hits should be found close to the interaction point near the photon direction (see Chapter 3). An important feature of supercluster reconstruction at CMS is that when all track information is ignored electrons and photons are identical. This is a principal ingredient in the $H \to \gamma\gamma$ analysis which allows data driven calibration, validation and efficieny measurements of photons using electrons (see Section 3.6).

2.2.3. Hadronic calorimeter

Surrounding the ECAL, but still inside the magnet, is a sampling HCAL which has geometric coverage up to $|\eta| < 5.0$ when including the specialised forward components.

DRAFT

It consists of alternating layers of brass plates and plastic scintillators (where in the very forward region the brass is replaced with steel). The HCAL thickness constitutes around 10-15 nuclear interaction lengths (λ_I) depending on η. Any outgoing hadrons from the interaction (of which there are many for a typical event with high E_T) get reconstructed as objects known as "jets" by amalgamating information from the tracking the system, the ECAL and the HCAL. This process is described in more detail in the particle flow section below (section 2.3).

2.2.4. Muons

Outside of the magnet lie the CMS muon chambers. These consist of alternating layers of drift tube chambers (cathode strip chambers) in the barrel (endcap) and resistive plate chambers which also act as a return for the magentic flux. The muon detector has coverage up to $|\eta| < 2.4$ and given the particularly conspicuous signature of muons (several hits in the tracker and hits in each muon station layer) the reconstruction efficiency and momentum resolution of muons is very good even down to low p_{T}. The muon resolution as a function of p_{T} is shown in the left hand plot of Fig. 2.9 for simulated data at $\sqrt{s} = 7$ TeV.

2.3. Particle flow and jets

A traditional approach to detector based particle physics is to consider the objects we measure as opposed to the underlying physics objects. These are often known as calorimeter objects, for example, a track, an electromagnetic shower or a calorimeter jet. A more modern approach is to couple information in all of the subdetector systems together to reconstruct more physical objects. For example, a charged hadron will leave a track, deposit some energy in the ECAL and deposit the rest of its energy in the HCAL. This technique of reconstruction is known as particle flow (PF)[2] and is particularly useful when considering jets. Whilst an electron, photon or muon are fairly unique looking signatures hadrons are often not. The abundant number of gluons and quarks produced in LHC collisions hadronise via the strong interaction as they travel away from the interaction. As they have typically high momentum the hadronistation occurs in a collimated fashion leaving a signature of several tracks and a hadronic cluster. The PF

[2]The offcial name inside CMS is global event description (GED) although this is rarely used

DRAFT

reconstruction algorithm can be simplistically viewed as the follwing procedure (more details given in Ref. [9]):

1. Make small clusters from each subdetector component; tracks, ECAL and HCAL clusters to create a list of unassociated objects.
2. Match tracks and clusters together and associate them to a newly reconstructed particle known as a PF candidate:
   - Tracks and clusters associated with hits in the muon chambers are tagged as *muons* and removed from the list.
   - Tracks and clusters associated with electrons, including Bremsstrauhlung photons, are tagged as *electrons* and removed from the list.
   - Tracks associated to an HCAL cluster are tagged as *charged hadrons*, assigned an energy ascertained from a weighted average of the cluster energy and track momentum and subsequently removed from the list.
   - Any excess cluster energy in the HCAL is assigned as a *neutral hadron* and removed from the list.
   - If an ECAL cluster is associated to an HCAL cluster and a track, it is assigned as a *charged hadron* with the appropriate weighted energy and removed from the list.
   - If an ECAL cluster is associated to an HCAL cluster with no track, it is assigned as either a *photon* or a *neutral hadron* depending on the HCAL to ECAL energy ratio and removed from the list.
   - Any remaining unlinked candidates are assigned as *photons* or *neutral hadrons* depending on whether they are ECAL or HCAL clusters.
3. In this way all information in the detector is used to create a list of candidates which can be any of a muon, electron, photon, charged hadron or neutral hadron.
4. These are then used to construct composite detector objects such as jets if necessary.

Particle flow jets are constructed using the anti-k_T algorithm [10]. This algorithm is both infrared and collinear (IRC) safe and preferentially clusters soft (low p_{T}) jets with hard (high p_{T}) jets to be robust in the LHC pileup conditions. These jets can then be additionally tagged as b or c (i.e. those containing a bottom or charm quark

DRAFT

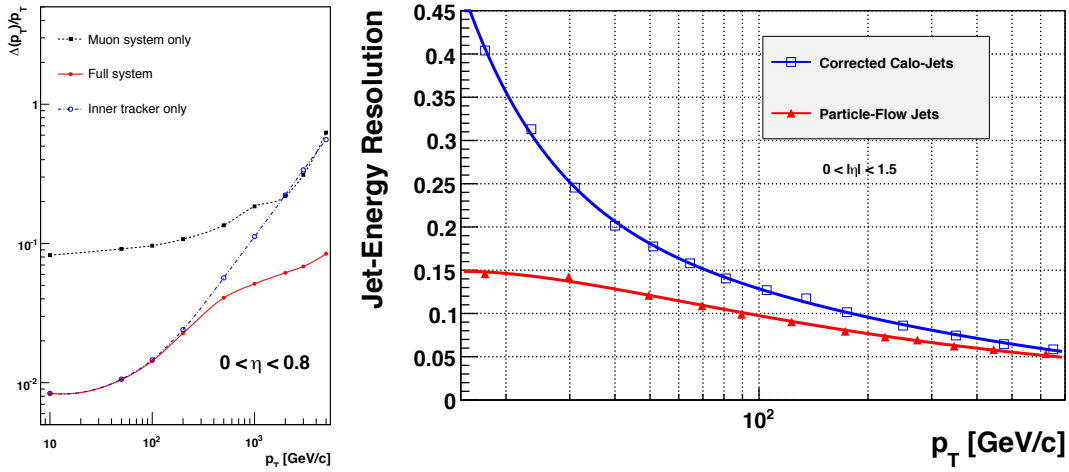

Figure 2.9.: Left: The muon p_{T} resolution as a function of p_{T} for muons in the range $0 < |\eta| < 0.8$ when using only the muon system (black), only the tracking system (blue) and using both (red) [4].
Right: The jet energy resolution as a function of jet p_{T} when using PF jets as compared to calo jets [13].

respectively) using the techniques of the type described in [11]. There are also energy corrections applied to jets to account for pileup (ρ subtraction technique as described in section 2.5), non-uniform detector response (p_{T} and η dependent corrections dervied from Monte Carlo (MC)) and data MC differences (residual p_{T} and η corrections derived from γ+jet and Z+jet samples in data), see Ref. [12] for details. The clustering, and subsequent energy correction, of jets in this way also provides a measure of the missing transverse energy ($\not\!\!E_T$), the amount of energy in an event taken away by undetectable particles such as neutrinos. A comparison of the jet energy resolution, as a function of jet p_{T}, for calorimeter jets and particle flow jets is shown in Fig. 2.9 for simulated data at $\sqrt{s} = 7$ TeV.

Should possibly include a note here stating that a PF photon is very different to the photons used in the analysis. PF is useful for tagging physics objects like tau's or b's and for isolation sums but is not necessarily the best for reconstructing well measured objects like photons, electrons and muons.

2.4. Isolation

One way of differentiating between real (prompt) photons and fakes is the use of isolation. One would expect that in the absence of pileup a real photon would be isolated, which is

to say there are no other particles (detector activity) in its vicinity. For a jet faking a photon (which is nearly always a π^0) this is not the case and one would expect the π^0 to be surrounded by additional hadronised particles (detector activity in the tracker, ECAL and HCAL). In the CMS $H \to \gamma\gamma$ analysis isolation sums are used to distinguish prompt photons from fakes. Three variables are used which consider each photons isolation relative to activity in the surrounding environment. The procedure is to create a hollow cone around the photon candidate (of outer radius ΔR_O and inner radius ΔR_I, where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) and sum the energy contained in that cone of PF candidates; charged hadrons, neutral hadrons and electrons/photons. These three variables are defined in the following way:

- **Charged hadron isolation:** Sum of charged hadron PF candidates E_T in cone of $\Delta R_O = 0.3$ and $\Delta R_I = 0.02$.
- **Neutral hadron isolation:** Sum of neutral hadron PF candidates E_T in cone of $\Delta R_O = 0.3$ and $\Delta R_I = 0.0$.
- **e/γ isolation:** Sum of e/γ PF candidates E_T in cone of $\Delta R_O = 0.3$ with a central η strip of 0.070 (0.015) removed for barrel (endcap) photons.

2.5. Pileup

There can be up to 1.1×10^{11} protons in each bunch at the LHC which results in multiple interactions (can be as many as 50 primary vertices) per bunch crossing. This effect is known as pileup and can present a challenge in finding the primary vertex and also produces additional energy in each event which originates from somewhere other than the primary vertex. To combat the latter of these two effects a technique called ρ subtraction is used to corrected the energy of jets and isolation sums for pileup. ρ is defined as a per event quantity and is computed by summing the energy in all the calorimeters and dividing by the calorimeter area and thus represents the average energy density in the detector per event.

DRAFT

Chapter 3.

Common analysis components

"I don't have a quote"

— Matthew Kenzie, 1785–1854

This thesis describes three complementary analysis regimes in the Higgs to two photons search at CMS. These differ in their photon selection, event selection, event classification (or categorisation) and statistical methods for extracting results. They are described in the following chapter (Chapter 4). However, there are many components which they share. These are detailed below.

As we have seen in Eq. 1.1, repeated below in Eq. 3.1 for convenience, the diphoton invariant mass is constructed from the two photon energies and the angle between them. Consequently, important considerations for this analysis are photon energy resolution and good opening angle resolution which is completely dominated by the vertex resolution, as the position resolution of the photons (location they hit the ECAL) is neglible in comparison. Details of how this is exploited in the analyses are given at the end of this chapter in Sections 3.3 and 3.4. The selection of events is described in the chapter after this (Chapter 4) alongside the categorisation, or binning, scheme into different classes of event which take advantage of areas of phase space which share similar signal to background ratios.

$$m_{\gamma\gamma} = \sqrt{E_1 E_2 (1 - \cos(\alpha))} \tag{3.1}$$

DRAFT

After a preliminary discussion of multivariate analysis techniques, the datasets, the triggering and the MC are discussed.

3.1. Boosted Decision Trees

MVAs are commonly used in High Energy Physics analyses to extract the maximum possible signal sensitivity in cases where the background rates are high. The advantage of MVAs is that given a set of input variables a network of sequential cuts can be built, to classify or correct events, in a multidimensional phase space to exploit differences between the signal and background in these variables and importantly in the correlations between these variables. A particular type of MVA which is used widely in this analysis is the Boosted Decision Tree (BDT). BDTs are prefered because they are more robust to the inclusion of variables which have little or no discriminating power. There are two broad types of BDT used, one is known as a regression BDT and the other as a classification BDT.

A classification BDT will, given a set of input variables, assign a value (typically between -1 and 1) to each event based on how signal like that event is. This serves to collapse all the event information into one discrimanting variable which can be used to classify differences between the signal and background. The input is provided as the probability distributions (which can be supplied as binned or unbinned data samples or as a functional form) of the background and signal for a set of "input variables". The process involves construction of a series of Decision Tree (DT)s complemented by a "boosting" step which serves to mitigate against "overtraining" on fluctuations within the training samples. This analysis chooses a particular type of decision tree boosting known as "gradient" boosting because it is more robust against outliers or mislabled data points [14].

The DT is built by applying sequential cuts to the input variables and assessing the relative signal purity, p, in the sub-sample remaining after each cut.

$$p = \frac{N_s}{N_s + N_b}, \tag{3.2}$$

DRAFT

where N_s and N_b are the sum of weights of the signal and background remaining in each sub-sample. A threshold criterion, known as the Gini index [14] $p(1-p)$, is applied to decide whether to split the sample further. The process continues and the splitting is curtailed when either the threshold or the user defined maximum tree depth (number of subsamples allowed) is reached. The value of each cut is varied such that the signal purity, p, in each sub-sample is maximised. An event is assigned a value of -1 or +1 depending on whether it falls into a sub-sample with p >0.5 or not. Clearly some fraction of events will be misclassifed where the actual number which get misclassified will depend on the discrimnatory power available from the chosen input variables. In order to reduce this effect a series of DTs are trained and each assigned a weight derived by the "boosting" process. If we assign each DT as a member of a family of M functions, $f(\vec{x};\vec{a}_m)$, which depend on the input variables, $\vec{x}$ and the set of cuts in that tree, $\vec{a}_m$. The object is to construct an overall decision tree which consists of the weighted average of each DT,

$$F(\vec{x};\vec{\beta},\vec{a}) = \sum_{m=0}^{M} \beta_m f(\vec{x},\vec{a}_m) \quad \text{where} \quad \vec{\beta} = (\beta_0,\beta_1 ... \beta_m) \tag{3.3}$$

The boosting procedue is implemented by adjusting the weights $\vec{\beta}$ in order to minimise the deviation in the loss function (Eq. 3.4) between the weighted tree response $F(\vec{x};\vec{\beta},\vec{a})$ and the true output y obtained from the training sample.

$$L(F,y) = \ln(1+e^{-2F(x)y}) \tag{3.4}$$

A common procedure when constructing a BDT to check for overtraining is to split both the background and signal into two independent samples. One is used to *train* the BDT and one is used to *test* the response of the output. Clearly one requires that both the training and independent test sample look the same in the output variable. This is usually quantified by use of a Kolmogorov-Smirnov test, which broadly speaking ascertains the probability that the training and test samples originate from the same underlying distribution [15].

DRAFT

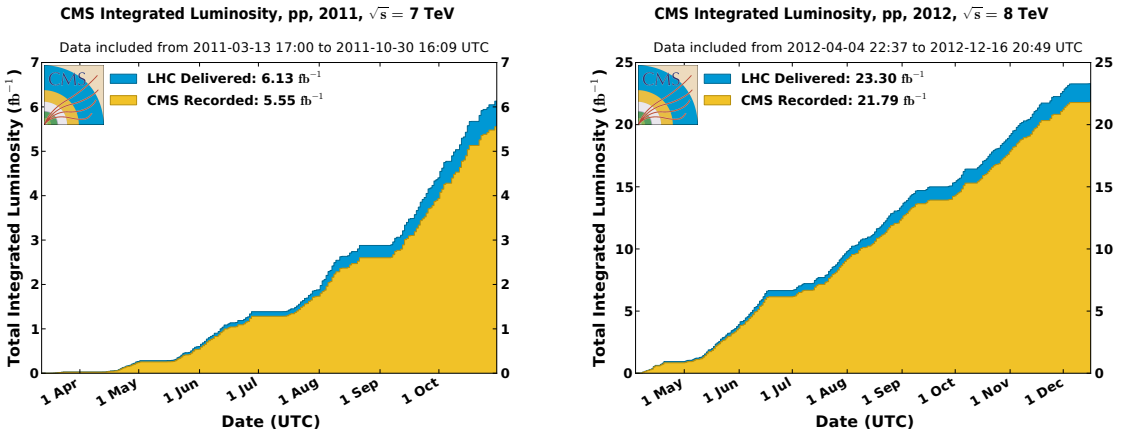

Figure 3.1.: The total integrated luminosity delived and recorded by CMS during the 2011 (left) and 2012 (right) run periods.

In this way the output of $F(\vec{x};\vec{\beta},\vec{a})$ for a classification BDT will be a "semi-continuous" output from -1 to 1 with signal events in general given a higher score than background events.

A regression BDT is used to estimate the true value of some variable given the values and correlations of several other variables. They are commonly used for correcting the energy of a particular object, for example a photon. Given a MC source of photons the "true" energy is regressed from the position, shape and raw energy of the supercluster. For regression BDTs the output $F(\vec{x};\vec{\beta},\vec{a})$ represents the estimated corrected energy and the boosting procedure targets minimising the deviation between this and the true energy in MC.

3.2. Data samples and triggering

The data consists of two indepedent samples of proton-proton collisions collected by the CMS experiment at the LHC in 2011 and 2012 with a centre-of-mass energy ($\sqrt{s}$) of 7 and 8 TeV respecivtely. The total integrated luminosity of the two samples is $5.1 fb^{-1}$ and $19.7 fb^{-1}$ in 2011 and 2012 respectively and collectively referred to as LHC Run 1. The response of the detector has changed considerably over this period and much of the variation is modelled by the MC simulation. Figure 3.1 shows the integrated luminosity delivered and recorded by CMS during LHC Run 1.

DRAFT

Events are selected for the analysis by requiring they pass an asymmetric diphoton trigger with E_T thresholds of 26 (18) and 36 (22) GeV for the leading (trailing) photon in the 2011 and 2012 runs respectively. The candidates are also required to have either a high value of R_9 or to pass a loose calorimetric identifaction and isolation requirement. High trigger efficiency is achieved by selecting photon candidates which pass either requirement. The efficiency of the trigger for the analysis preselection is 99.5%.

3.2.1. Monte Carlo

Accurate simulation of detector effects and efficiency is highly important. Knowledge of the expected Higgs signal shape is clearly essential and although the size and shape of the $m_{\gamma\gamma}$ background is entirely data driven when extracting results, simultaing the kinematics, shower shape and resolution properties of the background is essential when training the selection and optimising the binning.

As explained in Chapter 1 the two main production mechanisms for a Standard Model Higgs boson at the LHC are gluon fusion (ggH) and vector boson fusion (VBF). Typically the latter is produced at much higher Higgs p_{T} and this feature is exploited in the analysis. Consequently, it is important to model the p_{T} distribution of these two production modes accurately. The signal samples for these two processes are generated using POWHEG [16, 17] at NLO interfaced with PYTHIA [18] including a reweighting factor which matches their p_{T} spectrum to that when including the NNLO and NNLL terms. For the associated production modes, with a $W^{\pm}$,Z or t-quarks, (vector boson associated production (VH) and top quark associated production ($t\bar{t}H$)) only PYTHIA is used. The Standard Model Higgs boson cross sections and branching fractions are taken from Ref. [19]

The spin-2 graviton with minimal couplings, 2^{+}_{m}, has two production mechanisms, one via gluon-fusion (ggX) and one via quark-antiquark annihilation ($q\bar{q}X$). The graviton samples are generated using the JHU generator [20] in which the p_{T} spectrum of these samples is matched to the SM so that any bias obtained from mismodelling of the p_{T} is avoided.

The simulated background samples are used soley for cut and category optimisations and training of multivariate discriminants. The background which contains the quantum chromodynamics (QCD) continuum of prompt photons (refering back to Chapter 1 these are produced by Born and Box diagrams) is simulated using SHERPA [21] at 8TeV and

DRAFT

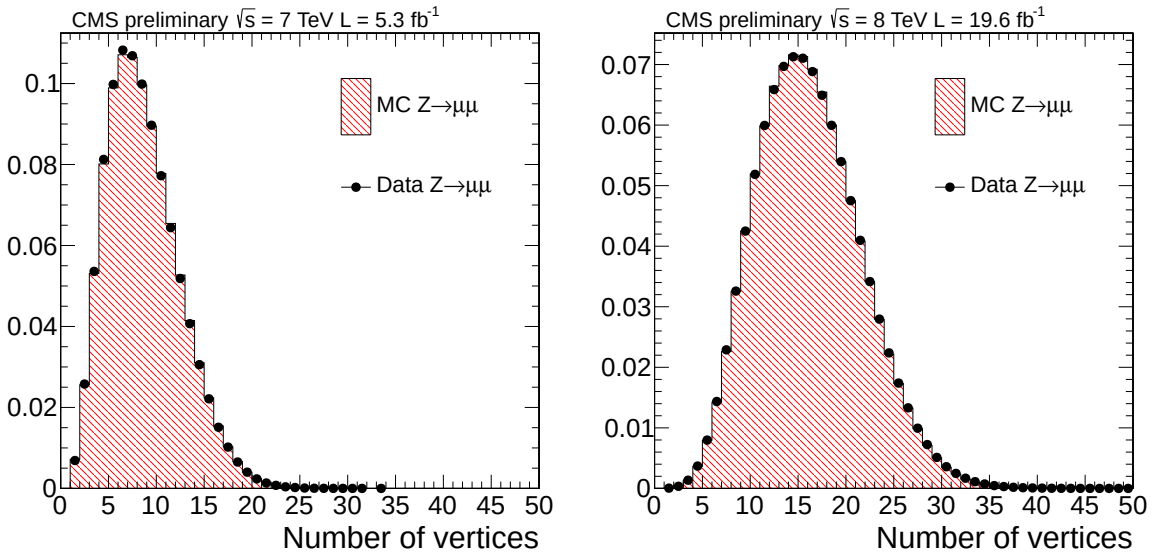

Figure 3.2.: Distribution of the number of reconstructed vertices in the 2011 (left) and 2012 (right) run periods. Calculated using the Deterministic Annealing alogirithm in [REF!] for $Z \to \mu^+\mu^-$ events in data (black dots) and MC (red histogram) after reweighting.

MadGraph [22] at 7TeV. The prompt-fake and fake-fake backgrounds, in which one or more photons are faked by a neutral meson (usually a π^0) reconstructed as a jet are generated using pythia. Samples of $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ and $Z \to \mu^+\mu^-\gamma$ used for data/MC comparisons are generated with powheg.

All of these *generator level* samples are then run through the full CMS detector simulation using Geant4 [23]. This includes the effect of overlapping vertices (pileup) and detector effects (such as noise and crystal degredation) in four bins of time (Run2011, Run2012AB, Run2012C, Run2012D).

3.2.2. Pileup and beamspot reweighting

An important difference between the simulated samples and the data which can have a large impact on the analysis is the distribution of the number of primary vertices. The *pileup* in the event effects many important analysis variables, for example photon shower shape and photon isolation as well as the diphoton invariant mass if the chosen vertex is wrong. Consequently the MC is reweighted such that the pileup distribution matches that in data. The reweighting technique is validated using $Z \to \mu^+\mu^-$ events as shown in Figure 3.2 for the 7 and 8 TeV samples respectively.

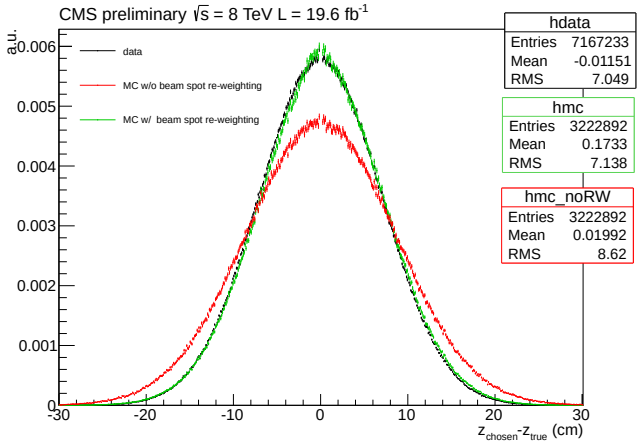

Figure 3.3.: Distribution of Δz (the distance between the chosen vertex and the true vertex in the z direction) for data (black), the MC (red) and the MC after beam spot reweighting (green) for $Z \to \mu^+\mu^-$ events.

When the chosen vertex is incorrect the mass resolution is dominated by the spread in position of the pileup vertices (known as the beamspot width, $\sigma_z^{beamspot}$). Accurate modelling of this spread is important so that the resolution of wrong vertex events in simulation matches that in data. The MC overestimates the beamspot spread by some 20% so a simple reweighting is implemented for MC events in which the wrong vertex chosen (as the effect is neglible for events in which the chosen vertex is correct) such that the distribution of the distance between the chosen vertex postion and the true vertex position, $\delta z = z_{chosen} - z_{true}$, match between data and MC. The effect with and without reweighting compared to data is shown in Figure 3.3.

3.3. Energy measurement of photons

The photon energy obtained from the supercluster sum described in Section 2.2.2 even when including the intercalibration and transperency corrections shown in Figure 2.7 does not give the most optimal resolution for the energy measurement of photons at CMS. On top of this energy (known as the *raw* supercluster energy, E_{raw}) it is also valuable to correct for additional energy losses. These arise from *bremstrauhlung*; where the photon converts in the material upstream of the ECAL and the two electron legs radiate additional photons and thus the some of the photon shower is missed, and for local containment of the shower; where some energy is lost through small gaps between ECAL crystals and larger gaps between "modules" or sections of crystals. These corrections are

obtained using a specialised regression BDT (see Sec. 3.1) trained on a MC source of prompt photons from a sample containing photons and jets. The BDT targets accurate measurements of individual photons' energy by correcting the raw supercluster energy and provides an estimate for the energy resolution of each photon given the position and shower shape of the supercluster. The training is done separately for barrel and endcap photons (as the cluster shapes look very different for these two distinct regions) and is also performed separately for the 7 and 8 TeV datasets. The following input variables are used:

- The global position of the supercluster in η and ϕ
- A collection of shower shape variables which aim at providing information on the likelihood and location of a photon conversion and the degree of showering in the material:
  - The R_9 of the supercluster (as previously described in Section 2.2.2).
  - The ratio of the 5×5 crystal energy to the raw supercluster energy (equivalent of R_{25}).
  - The energy weighted η-width and ϕ-width of the supercluster (in other words the spread of the shower).
  - The number of basic clusters.
  - The ratio of energy in the HCAL behind the supercluster to the ECAL energy of the supercluster, H/E.
  - The ratio of the preshower energy to the raw supercluster energy (endcap only).
- A collection of the seed crystal and the seed cluster variables which aim at providing information about energy lost through gaps and cracks between crystal and crystal modules:
  - The relative energy and position of the seed cluster.
  - The local energy covariance matrix.
  - Energy ratios between the seed and the 3×3 and 5×5 areas around the seed.
  - The η and ϕ index of the seed crystal and the relative position of the seed cluster to the crystal centre.

DRAFT

- Additionally the number of primary vertices and the median energy density, ρ, (see Sec. 2.5) are included to account for residual energy scale effects from pileup.

The regression is trained using an additional piece to that described in Section 3.1 whereby the target is to predict the full probability distribution of the ratio of the true energy to the raw energy, E_{true}/E_{raw}. The target is a double crystal ball distribution which consists of a Gaussian core and power law tails on either side. This can be fully parametrised by six variables, the Gaussian mean and width (μ, σ), the power parameters (n_L, n_R) and the power law tail cutoff parameters (α_L, α_R). Each of these parameters has a non-parametric dependence on the input variables, $\vec{x}$, and this is *learned* by the regression training whilst simultaneously minimising the likelihood,

$$-\ln \mathcal{L} = -\sum_{MCphotons} \ln p(E_{true}/E_{raw}|\mu(\vec{x}), \sigma(\vec{x}), \alpha_L(\vec{x}), \alpha_R(\vec{x}), n_L(\vec{x}), n_R(\vec{x})), \tag{3.5}$$

for the double crystal ball distribution, p. The most probable value for the true energy estimate of each photon is then given by,

$$E(\vec{x}, E_{raw}) = \mu(\vec{x}) E_{raw} \tag{3.6}$$

and the per-photon energy resolution is given by,

$$\frac{\sigma_E(\vec{x}, E_{raw})}{E(\vec{x}, E_{raw})} = \frac{\sigma(\vec{x})}{\mu(\vec{x})}. \tag{3.7}$$

In this way the regression predicts the full probability density of E_{true}/E_{raw} and provides an estimate of the optimal energy correction and the energy resolution per photon. A comparison of this distribution with a statistically independent MC sample is shown for the 8 TeV training in Figure 3.4. The performance of the regression compared to the default photon reconstruction is shown in Figure 3.5.

DRAFT

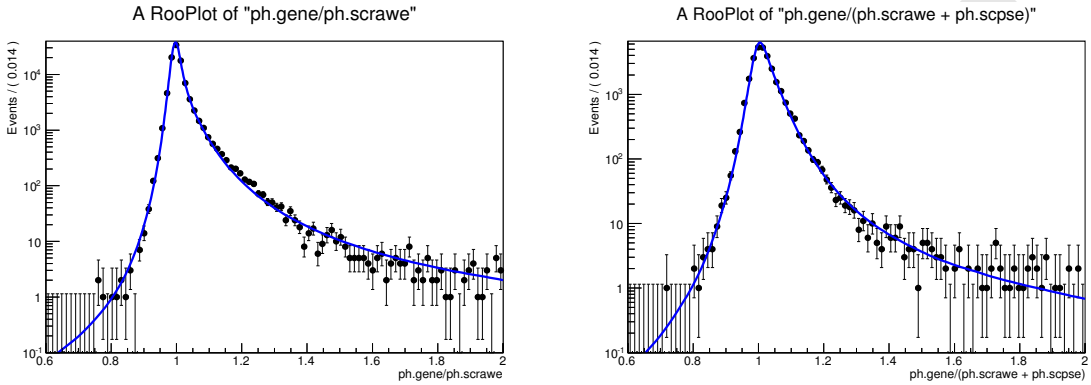

Figure 3.4.: A comparison of the predicted probability density of E_{true}/E_{raw} from the regression training (blue line) to the distribution in a statiscally independent MC sample (black points) for barrel photons (left) and endcap photons (right). PLOTS NEED TIDYING

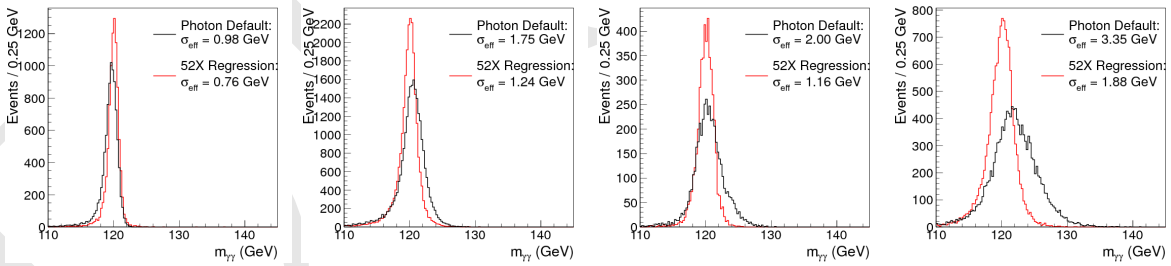

Figure 3.5.: A comparison of the regression performance compared to the photon default for the diphoton invariant mass on a MC source of Higgs decays to two photons. This is shown when both photons are in the barrel (left two plots) when both photons have $R_9 > 0.94$ (far left) and when at least one photon has $R_9 < 0.94$ (middle left). When at least one photon is in the endcap (right two plots) when both photons have $R_9 > 0.94$ (middle right) and when at least one photon has $R_9 < 0.94$ (far right). The value shown in the legend of each is the effective width, σ_{eff}, which is defined as half of the narrowest interval which contains 68.3% of the distribution. PLOTS NEED TIDYING.

3.3.1. Correcting for residual discrepancies between data and Monte Carlo

After application of the energy regression correction there are some remaining discrepancies between data and MC. These residual effects are accounted for using $Z \to e^+e^-$ events in data and simulation to correct the energy scale in the data and to apply an additional smearing term to the MC with systematic uncertainties propagated through the analysis to account for the uncertainties on these corrections.

Energy scale corrections to the data

The supercluster energy is identical for electrons and photons so by correcting the supercluster energy scale to a known source, namely the mass of the Z-boson, in dielectron decays the smaller residual energy scale effects are accounted for. This can be done several times to account for various different effects. In the first stage scale corrections are derived in bins of time (run range) and η, after applying these corrections further, much smaller, residual effects are accounted for in bins of R_9 (the size of the effect is different for converted and non-converted photons). After applying both of these a further step is taken for the 8 TeV data in the barrel to derive residual corrections in bins of E_T. Consequently the total scale correction is a product of three corrections in 59 bins of time $\times$ 4 bins in η $\times$ 2 bins in R_9 $\times$ 6 bins in E_T (for the 8TeV barrel photons alone).

The strategy for deriving these corrections is to take $Z \to e^+e^-$ events in data and MC and extract the invariant dielectron mass in the relevant bin of interest. This mass distribution is fitted with a convolution of a Breit-Wigner (designed to handle the underlying Z line shape [24]) and a Crystal-Ball which models the calorimeter resolution effects and bremstrauhlung losses in the material upstream of the ECAL. The Breit-Wigner parameters are fixed to the PDG values of $M_Z = 91.188$ GeV and $\Gamma_Z = 2.495$ GeV [24] whilst the Crystal-Ball parameters which model the detector effects are allowed to float. The scale correction, ΔE, is then defined as the relative difference between the Crystal-Ball peak in data and simulation,

$$\Delta E = \frac{m_{data} - m_{MC}}{m_Z}. \tag{3.8}$$

DRAFT

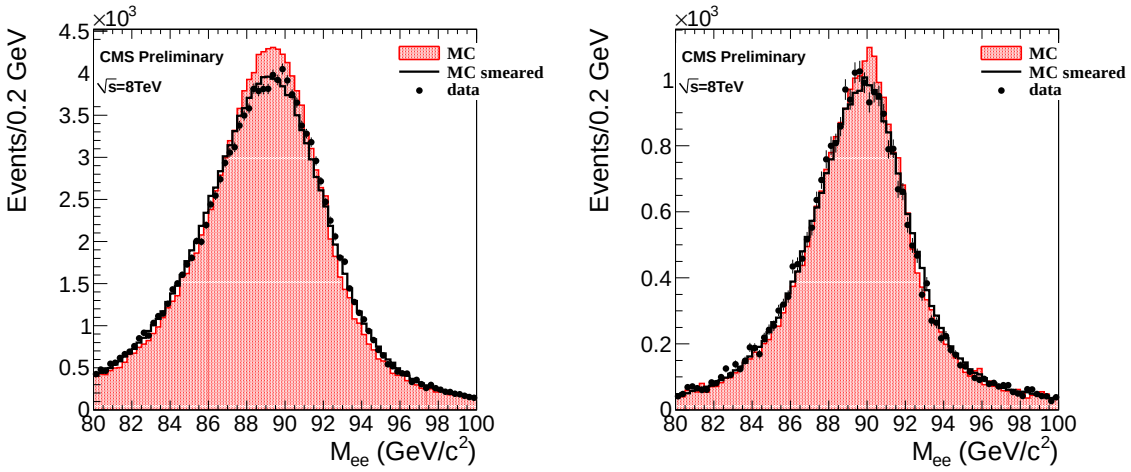

Figure 3.6.: The $Z \to e^+e^-$ invariant mass shape comparison before and after the scale and smearing corrections are applied. Shown for 8 TeV for photons with $1. < |\eta| < 1.444$ and $R_9 < 0.94$ on the left and for photons with $|eta| < 1.$ and $R_9 > 0.4$ on the right.

Energy resolution smearing the Monte-Carlo

A similar method is used to extract a smearing factor that can be applied to the MC such that the width of the invariant mass distribution in $Z \to e^+e^-$ decays matches between data and MC. This is done in 4 bins of η $\times$ 2 bins of R_9 and is parametrised as the quadratic sum of two resolution components: a constant term, ΔC, and a stochastic term, ΔS, which aims to model the expected resolution effects explained in Eq. 2.5 in Sec. 2.2.2. The smearing term, $\Delta\sigma$, is parametrised as,

$$\Delta\sigma = \frac{\Delta S}{\sqrt{E_T}} \oplus \Delta C. \tag{3.9}$$

The effect of the scale and smearing corrections is shown for the $Z \to e^+e^-$ data and MC samples in Figure 3.6 and for the analysis dataset after preselection with the electron veto inverted in Figures 3.7 and 3.8. Each of these corrections has an associated uncertainty and these uncertainties are propagated per photon through the analaysis. There are also additional uncertainties included which account for differences between electrons and photons and the difference between the Z mass scale (around 90 GeV) and the Higgs mass scale (around 125 GeV). Systematic uncertainties are described in more detail in Section. 5.4.

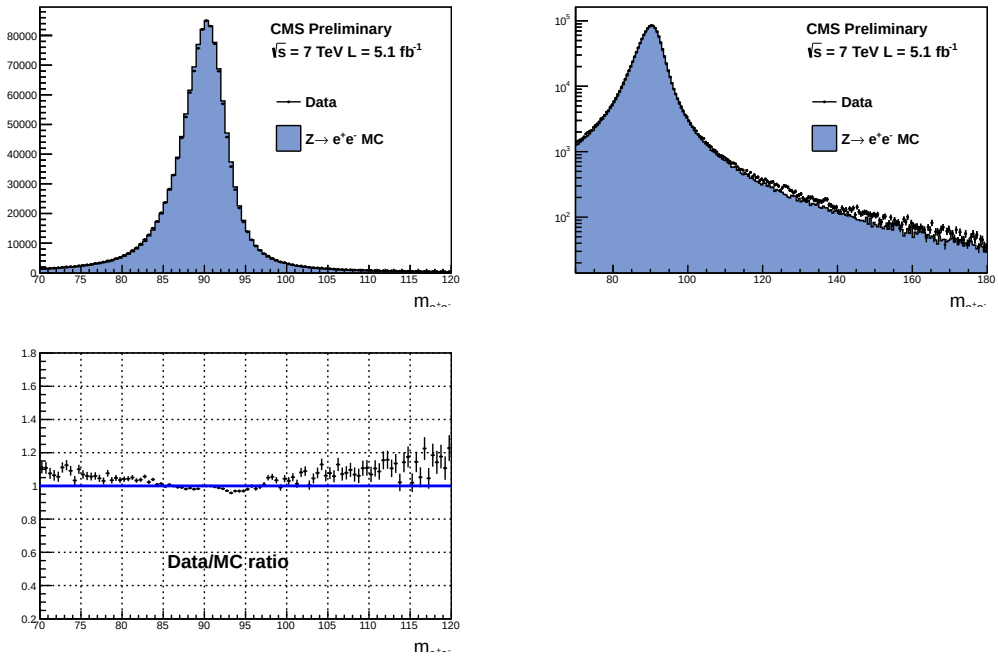

Figure 3.7.: The $Z \to e^+e^-$ invariant mass distribution at 7TeV in data (black points) and MC (blue histogram) for events which pass the analysis preselection in which the electron veto is inverted.

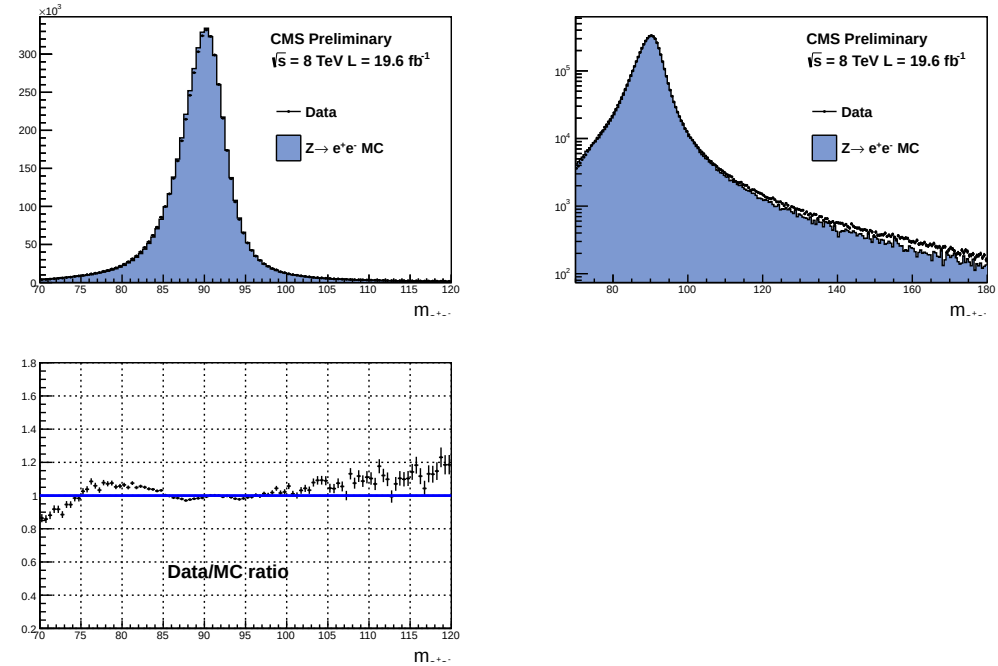

Figure 3.8.: The $Z \to e^+e^-$ invariant mass distribution at 8TeV in data (black points) and MC (blue histogram) for events which pass the analysis preselection in which the electron veto is inverted.

3.4. Vertex reconstruction

The resolution on the opening angle has a neglible effect if the correct vertex can be found within 10mm of the true interaction point. As seen in Sec. 3.2.2 the beamspot has an RMS spread of about 5 cm in the z direction and there is an average of $\sim$20 vertices per bunch crossing. Because the beam direction is along the z-axis the spread of the vertex in the x and y directions is tiny (number?) and consequently mismeasurement of the primary vertex in the x-y plane is small and has no impact on the mass resolution. By assigning the correct vertex to the diphoton pair, using other information in the tracking system, most of the mass resolution can be preserved. The method used to extract the primary vertex is a BDT which exploits the correlation between the diphoton pair and the recoiling tracks from the underlying interaction as well as additional information in the tracking system from a photon conversion pair. The output of this per vertex BDT is evaluated for each vertex in the event and the primary vertex is assigned as the one with the highest value of BDT output (i.e. the value nearest 1.). In addition, it is possible to construct another BDT whose output is proportional to the probability that the chosen vertex is the correct one (described in Sec. 3.4.1). This probability becomes a useful discrimnanting variable for the analyses later on.

The vertex BDT uses the following input variables:

- $\sum_i |\vec{p}_T^{\,i}|^2$ - the sum of the transverse momentum squared of all of the tracks which originate from this vertex, representing how hard the interaction is at this vertex.
- $\sum_i \left(\vec{p}_T^{\,i} \cdot \frac{\vec{p}_T^{\,\gamma\gamma}}{|\vec{p}_T^{\,\gamma\gamma}|}\right)$ - the sum of the magnitude of the transverse momentum of each track orginating from this vertex relative to the transverse momentum of the diphoton system, representing the recoil of the tracks to the diphoton system.
- $\left(|\sum_i \vec{p}_T^{\,i}| - \vec{p}_T^{\,\gamma\gamma}\right)/\left(|\sum_i \vec{p}_T^{\,i}| + \vec{p}_T^{\,\gamma\gamma}\right)$ - the asymmetry between the diphoton system and the other tracks originating from this vertex.
- $|z_v - z_c|/\sigma_c$ - this is added for events which contain at least one photon conversion where z_v is the z position of the vertex in question and z_c and σ_c are the estimated z position of the vertex from conversion information and its approximate error as defined below.

For events which contain at least one photon conversion, the conversion tracks and/or the conversion momentum can be used to point back to the beam line and estimate the

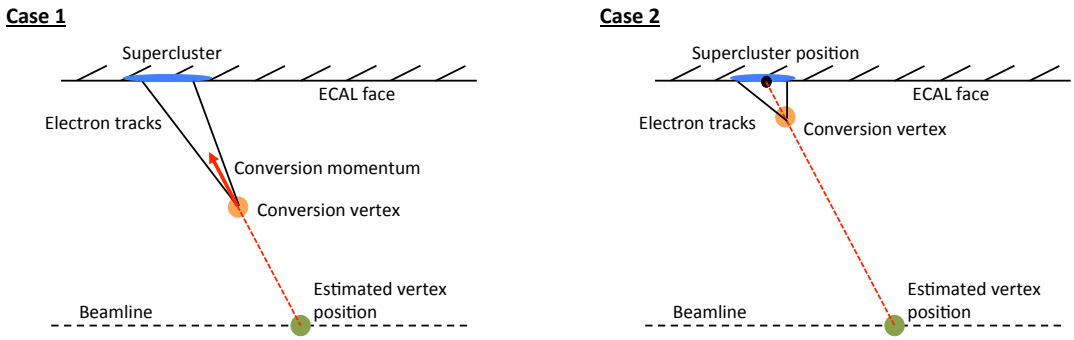

Figure 3.9.: A representation of the two methods for locating the primary vertex using photon conversion information. The left plot is for cases where the conversion occurs early enough in the tracker that the two electron tracks can be used to construct the converted pair momentum which is combined with the conversion vertex position to point back to the beam line. The right plot is for cases where the conversion occurs late in the tracker and the energy weighted supercluster position and the conversion vertex position are used to point back to the beam line.

vertex position. This can be acheived in one of two ways. In cases where the conversion occurs early, i.e. in one of the first layers of the tracking system, then the two electron legs of the conversion will leave two clean and distinct tracks. This means that the momentum of the conversion pair can be accurately reconstructed and used to point from the conversion vertex position back to the beam line and thus the nearest primary vertex. In cases where the conversion occurs late in the tracking system there are not enough track hits to accurately reconstruct the momentum of the conversion pair, however the incident position of the photon at the ECAL face is well known in this case, so the line which connects the ECAL position with the conversion vertex can be used to point back to the beam line. This is diagramtically represented in Figure 3.9 for both cases.

For Case 1 conversions the primary vertex z position is calculated as,

$$z_c = z_{conv} - r_{conv}\cot(\alpha), \tag{3.10}$$

where z_{conv} is the z position of the conversion vertex, r_{conv} is the distance of the conversion vertex from the beam line and α is the angle between the beam line and the conversion momentum.

For Case 2 conversions the primary vertex z position is calculated as,

$$z_c = \frac{z_{conv} - r_{conv}}{(r_{SC} - r_{conv})(z_{SC} - z_{conv})}, \tag{3.11}$$

where z_{conv} and z_{SC} are the z positions of the conversion vertex and supercluster respectively, and r_{conv} and r_{SC} are the distance of the conversion vertex and the supercluster from the beamline.

There are 6 regions of the tracking system (refer back to Figure 2.2). When the conversion vertex is located in one of the inner regions; Pixel Barrel, Pixel Forward, TID, the Case 1 conversion information is included in the BDT, otherwise the Case 2 conversion information is used. The resolution on the primary vertex position in conversions is estimated per tracking region by calulating the effective width [1] of the distribution of the difference between the z position of the primary vertex without conversion information and the z position of the primary vertex using conversion information alone, $\Delta z = z_v - z_c$. Consequently the fourth input variable to the BDT, shown in the list above as $|z_v - z_c|/\sigma_c$, is effectively a pull distribution for the conversion vertex. The BDT will favour vertices whose value of this variable is near zero.

The BDT is trained on a sample of $H \to \gamma\gamma$ MC events. It is tested with a statistically independent sample and further validated using $Z \to \mu^+\mu^-$ decays in data and MC. The efficiency is measured in data using the $Z \to \mu^+\mu^-$ channel where the muon tracks are removed from the BDT variables to simulate a diphoton like situation in data. The distributions of the input variables are shown for the $H \to \gamma\gamma$ training sample in Figure 3.10 (probably unnecessary). The BDT response is shown for $Z \to \mu^+\mu^-$ data and MC for both the signal (right vertex) and background (wrong vertex) in Figure 3.11. The chosen primary vertex is the one which gives the highest score BDT output. The efficiency of the vertex selection as a function of the Z p_{T} and the number of reconstructed vertices as measured in $Z \to \mu^+\mu^-$ data and MC samples is shown in Figure 3.12.

3.4.1. Estimating the per-event probability that the correct vertex is chosen

The total efficiency of assigning the correct vertex using the method described in the preceeding section is at the level of 75% during 2012 running conditions, where the

[1] Half the narrowest interval which contains 68.3% of the distribution

DRAFT

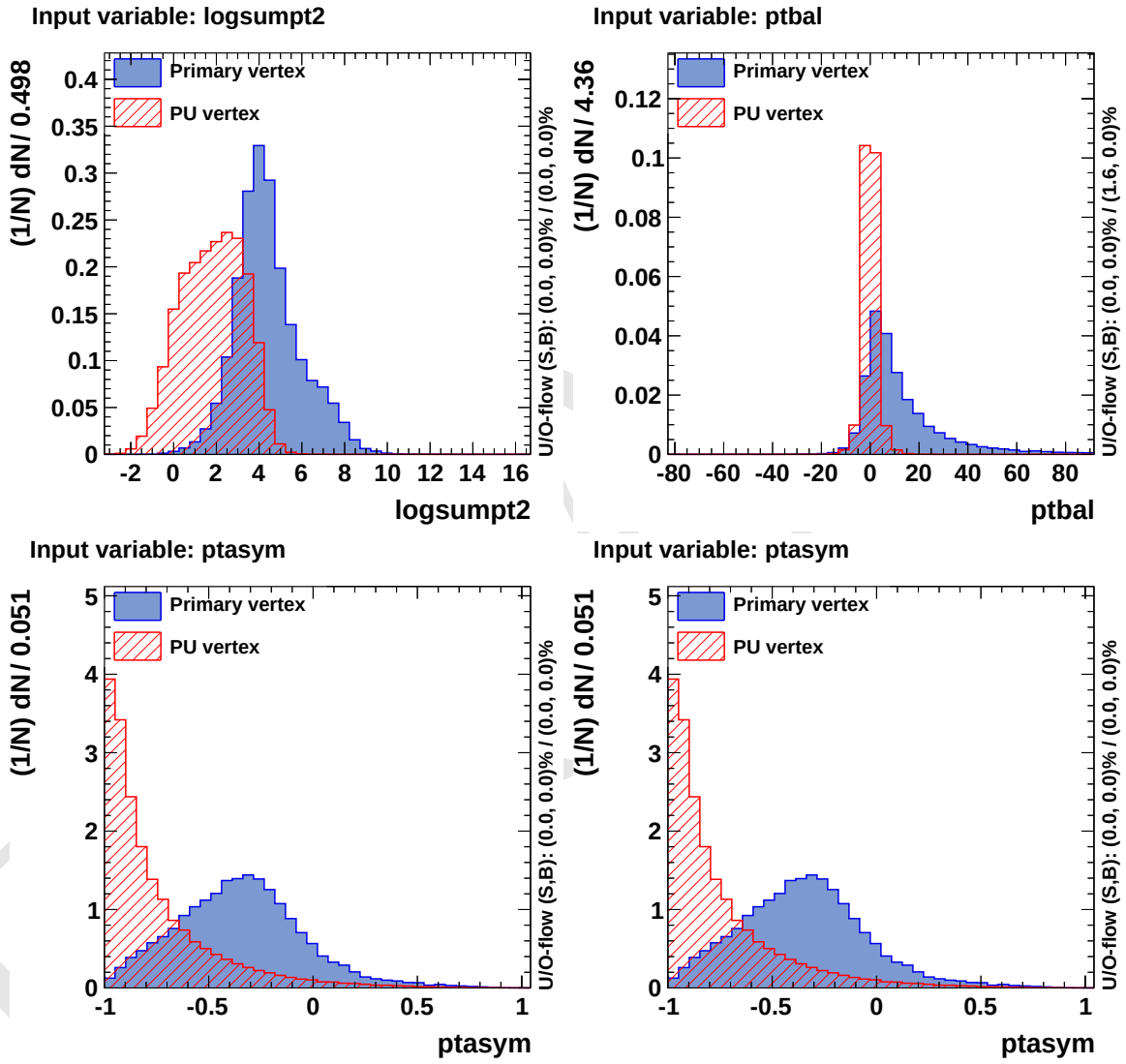

Figure 3.10.: Distributions of the input variables for the vertex BDT in the MC $H \to \gamma\gamma$ training (points) and test (filled) samples at 8TeV. Shown for the target primary vertex (blue histograms) and the background pileup vertices (red histograms). Plots need updating, tidying and labelling correctly.

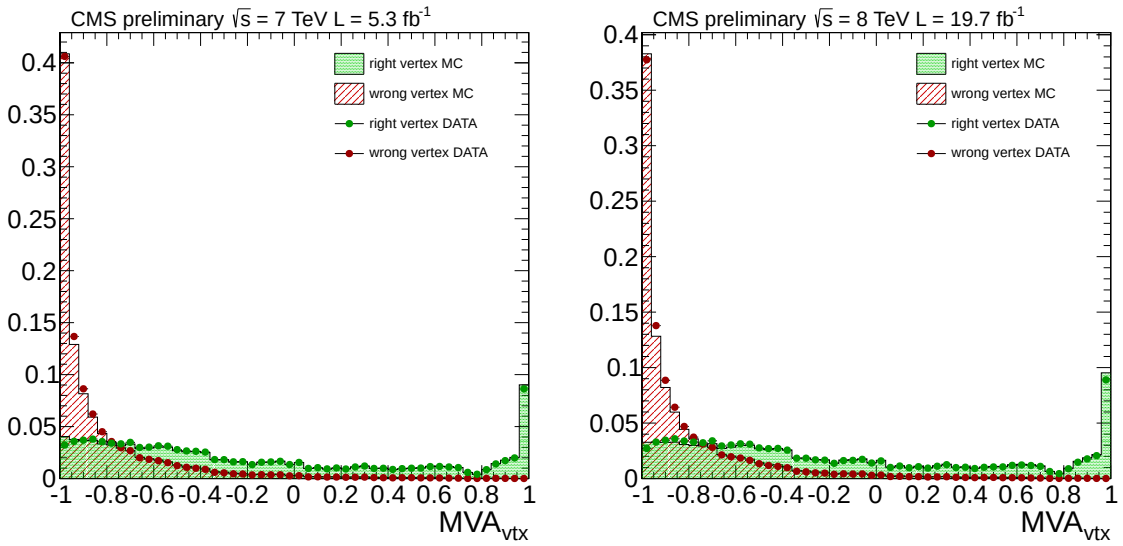

Figure 3.11.: The vertex BDT response for $Z \to \mu^+\mu^-$ events in data (points) and MC (filled) for the primary vertex (green) and the background pileup vertices (red).

correct vertex is defined as being within 10 mm of the true vertex. This means that for around 25% of preselected events the mass resolution is dominated by the vertex resolution. Consequently, it is important to ascertain the probability that the chosen vertex is the correct one. An additional specific BDT is constructed to address exactly this topic. The input variables used for this BDT are:

- The p_{T} of the diphoton system
- The number of vertices in each event
- The value of the per-vertex BDT described above
- The z distance, Δz, between the chosen vertex and the second and third choice vertices.
- The number of photon conversions used, either 0, 1 or 2.

There is a linear relation between the response of this BDT and the correct vertex efficiency (or probability). This relationship is demonstrated in Figure 3.13 and is fitted with a single parameter straight line in order to analytically obtain the per-event correct vertex probability for a given event. Figure 3.14 shows that this estimation reproduces the required vertex efficiency as a function of Higgs p_{T} and number of reconstructed vertices.

DRAFT

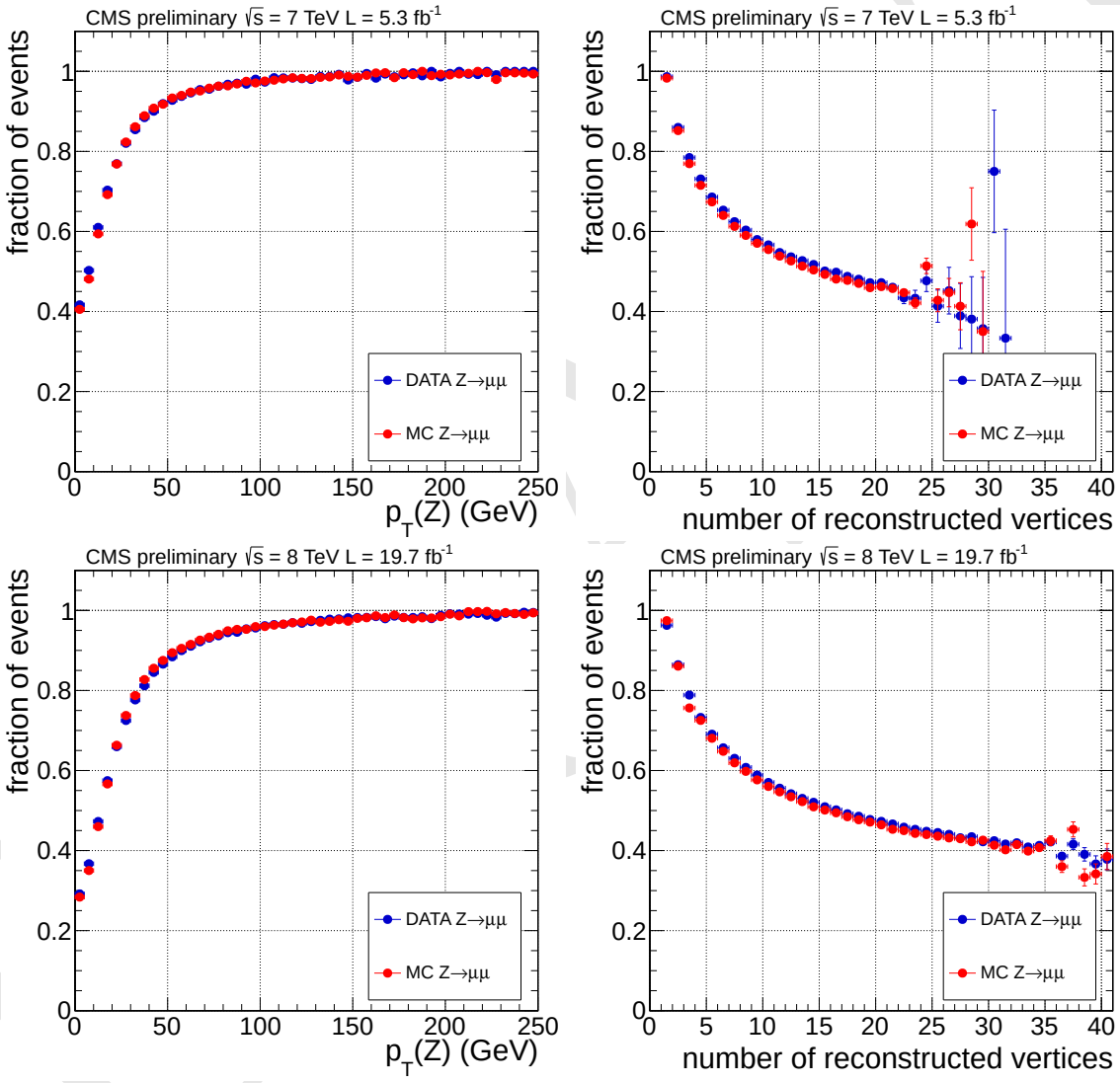

Figure 3.12.: The chosen vertex efficiency as measured in $Z \to \mu^+\mu^-$ data and MC as a function of Z p_{T} (left) and number of reconstructed vertices (right) for the 7 TeV (top row) and 8 TeV (bottom row) data samples.

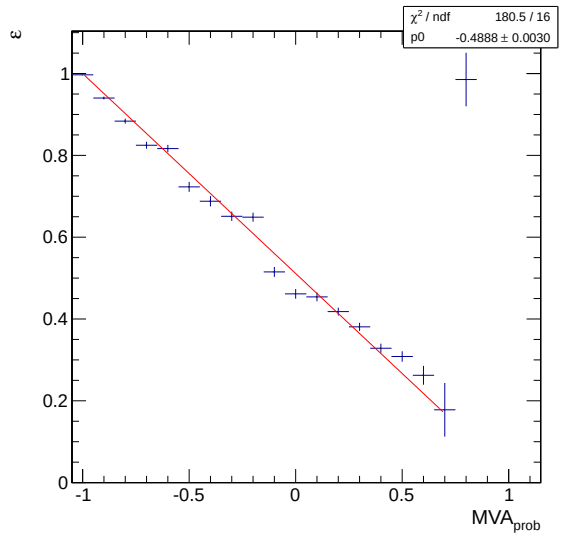

Figure 3.13.: A demonstration of the linearity relation between the per-event vertex probability BDT output distribution and the correct vertex probability

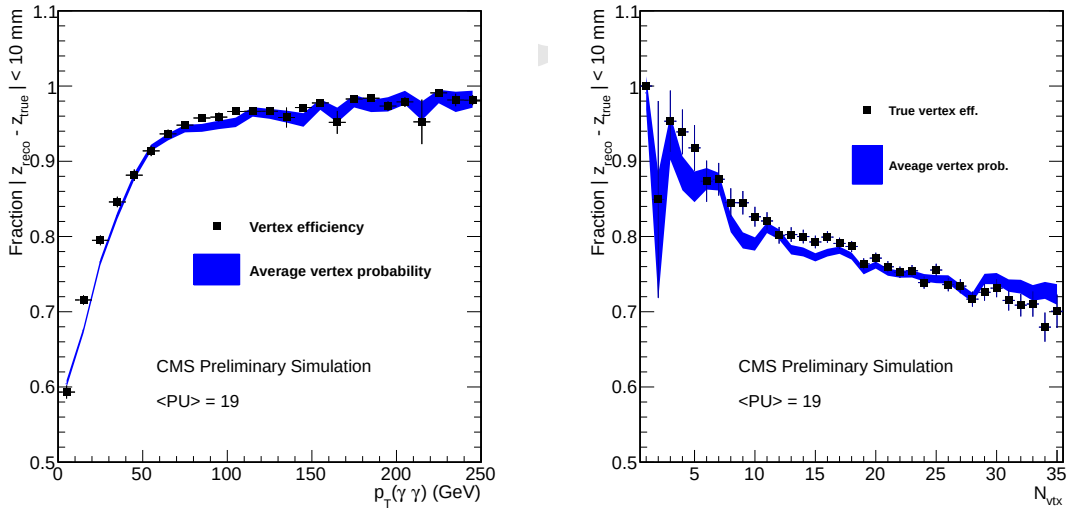

Figure 3.14.: A comparison of the true vertex efficiency (black points) and the average vertex probability (blue band) a statistically independent MC Higgs sample simulated with 2012 running conditions.

3.5. Event preselection

A simple and loose preselection is applied to all photons before entering the analysis. The preselection requirements are identical for all analysis approaches and are designed to remove some *fake* photons whilst mainting near 100% trigger efficiency. The variables used for preselection are defined below and the preselection cuts are described in Table 3.1.

- H/E - **The ratio of hadronic energy in the HCAL tower behind the supercluster to the electromagnetic energy in the supercluster**. Neutral jets which fake photons typically leave a fraction of their energy in the HCAL so there is a requirement that the value of this variable is small.

- $\sigma^2_{i\eta i\eta}$ - **The RMS spread of the shower in the eta direction**. Multiple showers of a π^0, or more than one π^0, result in a wider shower in η (as the π^0 decay product photons are separated in space). This cannot be exploited in the ϕ direction because conversion electrons get separated by the magentic field, however in η single photons, even when converted, occupy a narrow region. The separation of the two photons from a π^0 is minimal when they share the energy eqaully and given that typically $p_T >> m_\pi$ the separation is close to minimal for most. Taking the transverse plane in the barrel, the separation $d = 2Rm_\pi/p_T$ (where R is the radius of the barrel), which for $p_T = 40$GeV gives a value of d =8mm $\sim 0.006\eta$. By refering to Table **??** it is clear that the preselection requirement is quite loose.

- ISO_{ECAL} - **The total ρ corrected electromagnetic energy in a cone of radius 0.4 in (η,ϕ) around the photon candidate** - see Sec. 2.4

- ISO_{HCAL} - **The total ρ corrected hadronic energy in a cone of radius 0.4 in (η,ϕ) around the photon candidate** - see Sec. 2.4

- ISO_{Tracks} - **The total ρ corrected track energy in a cone of radius 0.4 in (η,ϕ) around the photon candidate** - see Sec. 2.4

- ISO_{PFCh} - **The total ρ corrected particle flow charged hadron energy in a cone of radius 0.4 in (η,ϕ) around the photon candidate** - see Sec. 2.4

In addition to the above an electron veto is applied to prevent contamination of the photon sample with electrons which originate from Drell Yan decays. This is acheived by removing photon candidates whose supercluster is matched to an electron track which has no missing hits in the innermost tracking region.

DRAFT

Table 3.1.: Preselection cut values.

	Barrel		Endcap	
R9	H/E	$\sigma^2_{i\eta i\eta}$	HoE	$\sigma^2_{i\eta i\eta}$
≤ 0.9	< 0.075	< 0.014	< 0.075	< 0.034
> 0.9	< 0.082	< 0.014	< 0.075	< 0.034
Both Barrel and Endcap				
R9	ISO_{ECAL}	ISO_{HCAL}	ISO_{Tracks}	ISO_{PFCh}
≤ 0.9	< 4 GeV	< 4 GeV	< 4 GeV	< 4 GeV
> 0.9	< 50 GeV	< 50 GeV	< 50 GeV	< 4 GeV

3.6. Using Z decays for validation and efficiency measurements

Whilst no known "standard candles" with high statistics exist for high p_{T} photons in the LHC environment a powerful control source for the $H \to \gamma\gamma$ decay in both data and MC is the $Z \to e^+e^-$ decay. From a detector view point electrons are very similar to photons and the Z is relatively near the relevant Higgs search range in mass. The differences between the Z and the Higgs, in both their mass and p_{T} distribution, and also the differences between electrons originating from a Z and photons originating from a Higgs are important systematic uncertainties on the Higgs mass scale and resolution. By inverting the electron veto usually applied in the preselection the Higgs to two photon analysis can be identically replicated but the very pure diphoton sample replaced with a pure dielectron sample. One additional process which can be used as a direct control for photons is the $Z \to \mu^+\mu^-\gamma$ decay although the statistics, even with the LHC luminosity, are very low. Many of the input variables used in training the BDTs and cuts are validated with both $Z \to e^+e^-$ and $Z \to \mu^+\mu^-\gamma$ data/MC comparison plots. An example is shown in Figure 3.15 for the reconstructed di-electron mass for events passing the preselection described in Table 3.1.

As previously shown (in Sec. 3.3.1) data/MC comparison of the $Z \to e^+e^-$ decay are used to derive scale corrections for the data and smearing of the MC. Discrepancies between data and MC in $Z \to e^+e^-$ decays of important analysis variables are accounted for by introducing systematic uncertainties to cover them. In addition the "tag and probe" method [25] is used on $Z \to e^+e^-$ decays to evaluate the signal efficiency for the preselection and analysis cuts. Several stages of the analysis are validated in this

DRAFT

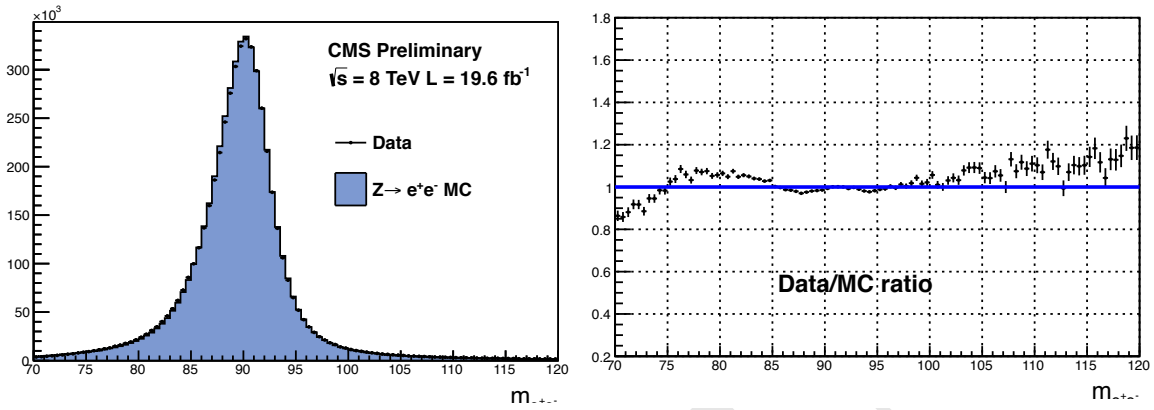

Figure 3.15.: The di-electron mass distribution after applying the preselection (described in Sec. 3.5) and the scale and smearing corrections (described in Sec. 3.3.1) whilst inverting the electron veto. The left plot shows the data at 8TeV as the black points with the Drell Yan MC sample as the blue histogram. The Data/MC ratio is shown in the right hand plot. Plot needs updating

way and where appropriate systematic uncertainties are included to account for any data/MC differences. Although the numbers and uncertainties themselves are derived from $Z \to e^+e^-$ samples (because of the much higher statistics), they are always cross checked with the $Z \to \mu^+\mu^-\gamma$ sample.

DRAFT

DRAFT

Chapter 4.

Selection and Categorisation

"Science never solves a problem without creating ten more."
— George Bernard Shaw, 1856 – 1950

This thesis presents results of three different analyses used in the Higgs to two photon search at CMS. The nominal results and properties are obtained from the so called Mass Factorized MVA (MFM) analysis, which uses multivariate methods optimised specifically to search for a Standard Model Higgs boson to select and categorise events. Two further analyses are presented as cross checks to the baseline MFM analysis. The first, the Cuts in Categories (CiC) analysis, is designed for simplicity and robustness as a cut based approach and, owing to its low level of model dependence, is used for the spin analysis. The second, the Sideband MVA (SMVA) analysis, serves to cross check the background, the most significant unknown in a search like this, and categorisation by extracting the background under the signal region from sidebands in the diphoton invariant mass spectrum.

Both the CiC and MFM analyses have a fully parametric definition of the diphoton invariant mass spectrum where the signal shape is derived from MC simulation. The SMVA uses a cut and count method.

4.1. Event selection

There are two complimentary event selections used in the three analyses. The CiC analysis uses a cut based photon selection whilst the MFM and SMVA analyses use

DRAFT

878 a series of BDTs to first select photons and then select events. The ultimate aim is
879 a selection which accepts two prompt photons (i.e. does not contain any fakes) and
880 exploits regions of phase space which have a narrow mass resolution and high signal to
881 background ratio.

882 All events must contain at least two photons which pass $p_T^{\gamma_1} > m_{\gamma\gamma}/3$ (for the leading
883 photon) and $p_T^{\gamma_2} > m_{\gamma\gamma}/4$ (for the subleading photon) and the invariant mass of the
884 diphoton pair must be in the range $100 < m_{\gamma\gamma} < 180$. In the case where there are more
885 than two photon candidates in an event, the two photons which give the highest scalar
886 sum p_{T} are chosen.

887 4.1.1. Selection using cuts in categories

888 Cuts are optimised for photons in four non-overlapping categories to take advantage of
889 the different photon energy resolution between the barrel and the endcap and between
890 converted and non-converted photons. The categories are defined as $|\eta| < 1.444$ or
891 $|\eta| > 1.556$ (i.e. either barrel or endcap) and $R_9 < 0.94$ or $R_9 \geq 0.94$ (i.e. either converted
892 or non-converted) and the cuts are chosen by targeting a specific signal to background
893 ratio (S/B). The procedure to define the cuts is as follows,

894 - A set of loose cuts are defined as the starting values.
895 - The "N-1" distribution of each cut variable is produced. This is the distribution of
896 each variable after the cuts on all other variables have been applied.
897 - A smooth curve is fitted to the distribution of the B/S ratio versus the cut variable.
898 - The cut value is chosen by evaluating the curve for the required value of S/B.
899 - Do the same for all the other variables.

900 Consequently a stable set of cut values is obtained by iterating this procedure a few
901 times. Each cut will then select events with the same purity (S/B) and thus the efficiency
902 for selected photons is maximised for the chosen purity level. As a result the cuts in
903 the endcap are much tighter than in the barrel, similarly the cuts for low R_9 photons
904 are much tighter than the cuts for high R_9 photons. The cut variables are described
905 below and the cut values shown in Table 4.1. The cut setting procedure is optimised on
906 a signal sample of $H \to \gamma\gamma$ MC (with m_H =120GeV) and a background sample of γ+jet

DRAFT

	Barrel		Endcap	
	$R_9 > 0.94$	$R_9 < 0.94$	$R_9 > 0.94$	$R_9 < 0.94$
PF isolation sum, chosen vertex	<6	<4.7	<5.6	<3.6
PF isolation sum, worst vertex	<10	<6.5	<5.6	<4.4
Charged PF isolation sum	<3.8	<2.5	<3.1	<2.2
$\sigma_{i\eta i\eta}$	<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

Table 4.1.: Photon ID selection cut values. The cuts are applied to both the leading and subleading photons.

events. The photon efficiency of the cuts for the four different classes of photon is shown in Figure 4.1 as a function of the supercluster position, η, and the photon p_{T}.

- **PF isolation sum, chosen vertex** - Sum of particle flow photon, charged and neutral hadron isolation sums defined in Sec. 2.4 where all PF candidates originate from the primary vertex selected in Sec. 3.4
- **PF isolation sum, worst vertex** - As above but where all PF candidates originate instead from the vertex which gives the largest charged hadron PF isolation sum. This add protection for cases when the primary vertex is incorrectly assigned.
- **Charged PF isolation sum** - Described above in Sec. 2.4
- $\boldsymbol{\sigma_{i\eta i\eta}}$ - Described above in Sec. 3.5
- $\boldsymbol{H/E}$ - Described above in Sec. 3.5
- $\boldsymbol{R_9}$ - Described above in Sec. 2.2.2

4.1.2. Photon ID MVA

For the multivariate approach a BDT is trained to discriminate between prompt photons and jets. The desire is to factorise the photon selection, which is required to distinguish prompt photons from neutral mesons faking photons, and the event selection, which is required to consider kinematics, resolution etc. to distinguish $H \to \gamma\gamma$ from the $pp \to \gamma\gamma, \gamma + jet, \mathrm{jet} + \mathrm{jet}$ background. The input variables used are designed specifically

DRAFT

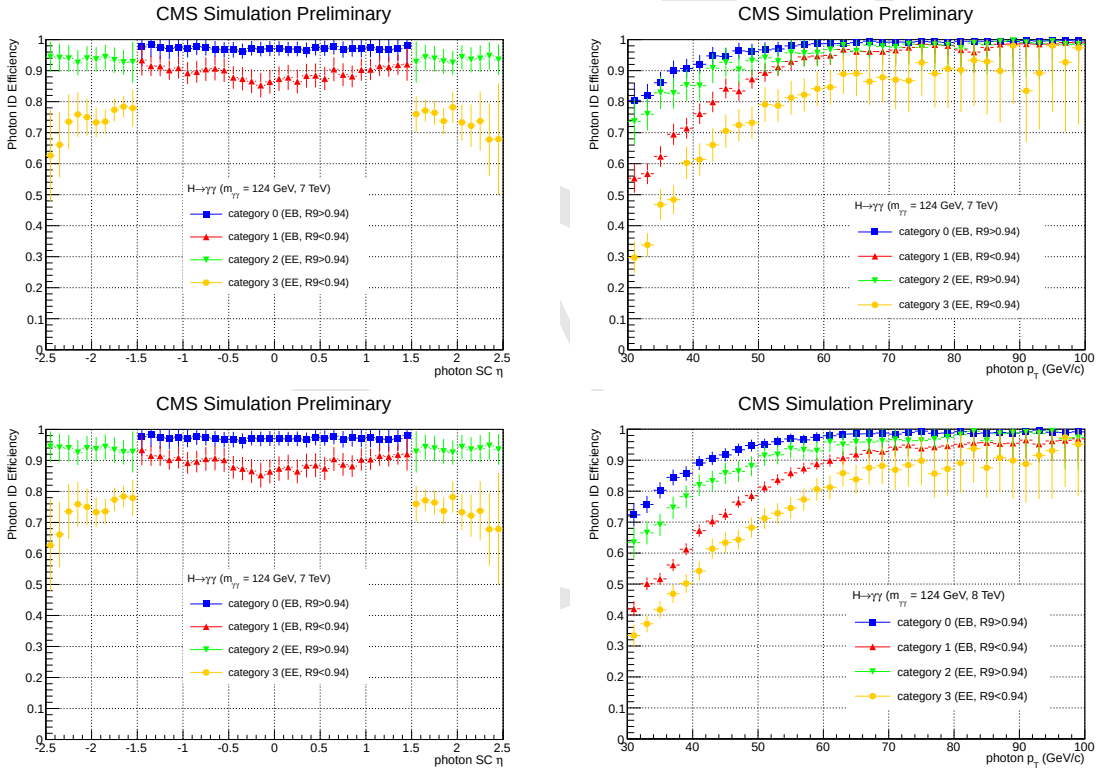

Figure 4.1.: Cut based photon ID efficiency as measured in $Z \to e^+e^-$ tag and probe. MORE DESCRIPTION

to distinguish between photons and fakes and to decouple from any event or photon kinematics which make the identification Higgs specific. Consequently, the training samples used are γ+jet samples where the identification BDT signal is the prompt γ and the background is the fake jet. The p_{T} and supercluster η distribution of the prompt photon are reweighted to match the background non-prompt photons in order to further negate any photon kinematics which the BDT could exploit. The photon ID BDT is trained separately for the barrel and endcap in both 7 and 8 TeV as these regions of phase space are so different. The result is four separate trainings in total.

The input variables aim to exploit differences in the shower shape and isolation between prompt and non-prompt photons and the correlation between these variables and the supercluster position and energy. They are,

Shower shape variables

- $\sigma_{i\eta i\eta}$ - Explained above in Sec. 3.5
- $\sigma^2_{i\eta i\phi}$ - The equivalent diagonal spread (in η,ϕ) of the shower.
- $E_{2\times2}/E_{5\times5}$ - Ratio of energy in the most energetic 2×2 cluster which contains the seed to the energy in the 5×5 cluster.
- R_9 - Explained above in Sec. 2.2.2.
- σ_η - The energy weighted standard deviation of single crystal eta within the supercluster.
- σ_ϕ - The energy weighted standard deviation of single crystal phi within the supercluster.
- σ_{xy} (for endcap only) - The standard deviation of the shower spread in the x, y planes of the preshower.

Isolation variables

- PF Photon ISO - Particle flow photon isolation sum
- PF Charged Hadron ISO (selected vertex) - Particle flow charged hadron isolation sum for candidates originating from selected vertex.
- PF Charged Hadron ISO (worst vertex) - Particle flow charged hadron isolation sum for candidates originating from the vertex with the largest isolation sum.

DRAFT

954 **Correlation variables**

955 • ρ - The median energy density in the event.

956 • η - The η position of the photon supercluster.

957 • E_{raw} - The raw energy of the photon supercluster.

958 The testing sample used to verify the output of the photon identification BDT is a
959 MC $H \to \gamma\gamma$ sample (m_H =124GeV). The photon identification BDT output (which is
960 shown for the four different trainings in Figure 4.2) provides a measure of an individual
961 photons' "quality" and is used as an input to the event level BDT (described in the next
962 section). Even so, a considerable amount of background is cut out by defining a loose
963 cut (the BDT output must be > -0.2) on the photon ID BDT output value which is
964 more than 99% signal efficient.

965 The photon ID BDT response for each photon is used as a direct input to the event
966 level MVA. Given that imperfect modeling of the detector response can result in a small
967 change in the photon ID response which has a direct impact on the event level MVA
968 response, which is used to classify events, a systematic on the photon quality is applied
969 and propagated through to the event level MVA. Validation of the photon ID BDT
970 response in the $Z \to e^+e^-$ decay is shown in Figure 4.3 with the size of the systematic
971 shown. Validation is done with the $Z \to e^+e^-$ decay on all of the input variables, these
972 distributions are shown in Appendix A (this may not be entirely necessary)

973 4.1.3. Diphoton event level MVA

974 Whilst the CiC analysis selects events based on photon identification, the MVA analysis
975 approach is to first select photons using the photon identification BDT described in the
976 section above and then pass all the relevant event information through an event level
977 BDT. The event level classifier, referred to as the diphoton BDT, is constructed to give
978 a high score to events which fulfill the following criteria:

979 1. The event kinematics should be compatible with a Higgs decay.

980 2. The event has good mass resolution.

981 3. The event contains two "high quality" photons (i.e. they have a high score from the
982 photon ID BDT).

DRAFT

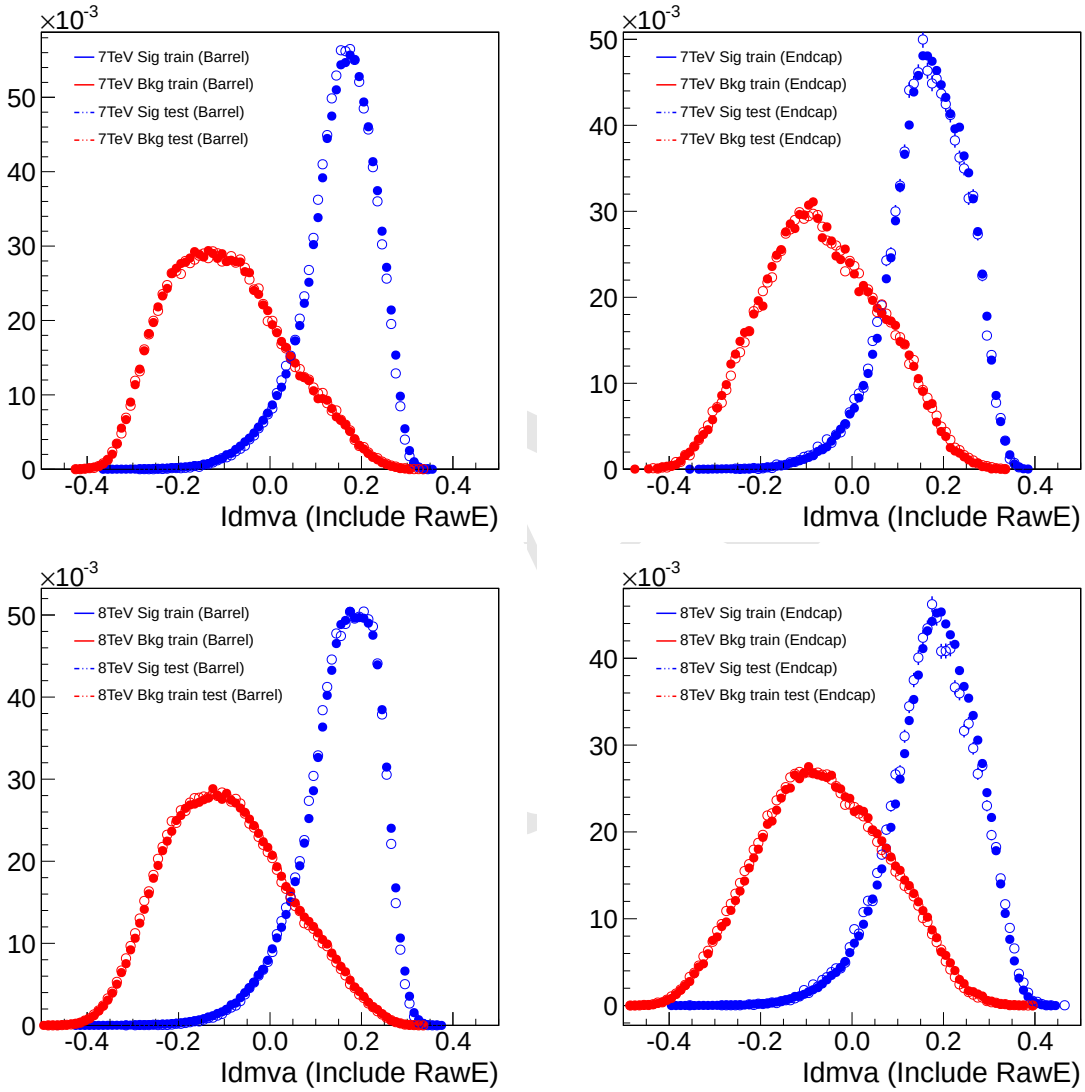

Figure 4.2.: The output distribution of the photon identification BDT for 7TeV barrel (top left), 7TeV endcap (top right), 8TeV barrel (bottom left) and 8TeV endcap (bottom right). The solid points show the γ+jet training sample distributions and the hollow points show the $H \to \gamma\gamma$ test sample distributions for prompt signal photons in blue and fake background photons in red. A cut of > -0.2 is made on all photons.

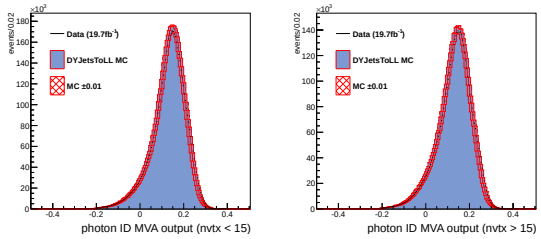

Figure 4.3.: The output distribution of the photon identification BDT for the 8 TeV training as validated by the $Z \to e^+e^-$ decay. The data is shown as the black points with the MC as the blue histogram. The systematic uncertainty on the output as applied to the MC is shown as the red band. This is literally just a place holder for the plot that I want to show.

It is highly important that the BDT is completely independent of Higgs mass and that the input variables have no, or at the least very little, dependence on the Higgs mass. This is essential to have a fair training. For example if the BDT included the Higgs mass, or a variable highly correlated with it, it would preferentially select events with this mass therefore biasing the selection towards events which have a mass near the mass of the signal used to train with. The input variables used are,

Event kinematics

- $p_T^{1(2)}/m_{\gamma\gamma}$ - The mass relative transverse momenta of each photon.
- $\eta^{1(2)}$ - The pseudorapidity of each photon.
- $\cos(\phi_1 - \phi_2)$ - The cosine of the angle between the two photons in the transverse plane. This variable reflects the p_{T} of the diphoton system (in other words the reconstructed Higgs candidate) without introducing a mass dependence.

Mass resolution

- $\sigma_m^{right}/m_{\gamma\gamma}$ - The mass resolution of the event assuming the correct primary vertex has been selected. In this case the angular resolution is neglible so the variable is caluclated using just the two photon energy resolution values as,

$$\frac{\sigma_m^{right}}{m_{\gamma\gamma}} = \frac{1}{2}\left(\frac{\sigma_{E_1}}{E_1} \oplus \frac{\sigma_{E_2}}{E_2}\right). \tag{4.1}$$

- $\sigma_m^{wrong}/m_{\gamma\gamma}$ - The mass resolution of the event assuming the wrong vertex is selected. The vertex position in z is distributed as a Gaussian with a width equivalent to $\sqrt{2}\sigma_z^{beamspot}$ and so the angular resolution σ_m^{vtx} can be analytically calculated given

the ECAL impact positions of the two photons. Consequently the wrong vertex variable is calculated as,

$$\frac{\sigma_m^{wrong}}{m_{\gamma\gamma}} = \frac{\sigma_m^{right}}{m_{\gamma\gamma}} \oplus \frac{\sigma_m^{vtx}}{m_{\gamma\gamma}}. \tag{4.2}$$

- p_{vtx} - The probability that the selected primary vertex is correct. In order to tie together the mass resolution information given the right vertex hypothesis and the wrong vertex hypothesis, the probability that the vertex is correct is used in addition.

- It is also important to specify in the training that the signal to background ratio is inversely proportional to the mass resolution. Accordingly the signal events in the training are weighted by a factor,

$$w = \frac{p_{vtx}}{\sigma_m^{right}/m_{\gamma\gamma}} + \frac{1 - p_{vtx}}{\sigma_m^{wrong}/m_{\gamma\gamma}}. \tag{4.3}$$

Photon quality

- $photID^{1(2)}$ - The photon ID BDT output value of each photon.

The training is performed separately for 7 and 8 TeV and the samples used for the signal are all of the Standard Model $H \to \gamma\gamma$ MC samples (ggH, VBF, VH, $t\bar{t}H$) appropriately weighted by cross section and the samples used for the background are the cross section weighted mixture of Standard Model backgrounds which include contributions from $pp \to \gamma\gamma$ (prompt-prompt), $pp \to \gamma + \text{jet}$ (prompt-fake) and $pp \to \text{jet} + \text{jet}$ (fake-fake) as described in Sec. 3.2.1. The training is performed on only half of the event samples (selected by even event number) so that the BDT response can be tested on the other half (selected by odd event number).

A cut is placed on the BDT response in order to remove almost all the events which contain two fake photons. The remaining events which pass this cut are categorised in coarse bins based the on the BDT response. The strategy for optimising this cut value and the category boundaries is explained in Section. 4.2. The BDT response in data, background and signal is shown in Figure 4.4. The cut values are > 0.19 for the 7 TeV training and > -0.78 for the 8 TeV training. It can be seen from this figure that the cut on the event level BDT removes nearly all of the fake-fake contribution to the background whilst the remainder consists of about 70% prompt-prompt and 30% prompt-fake.

DRAFT

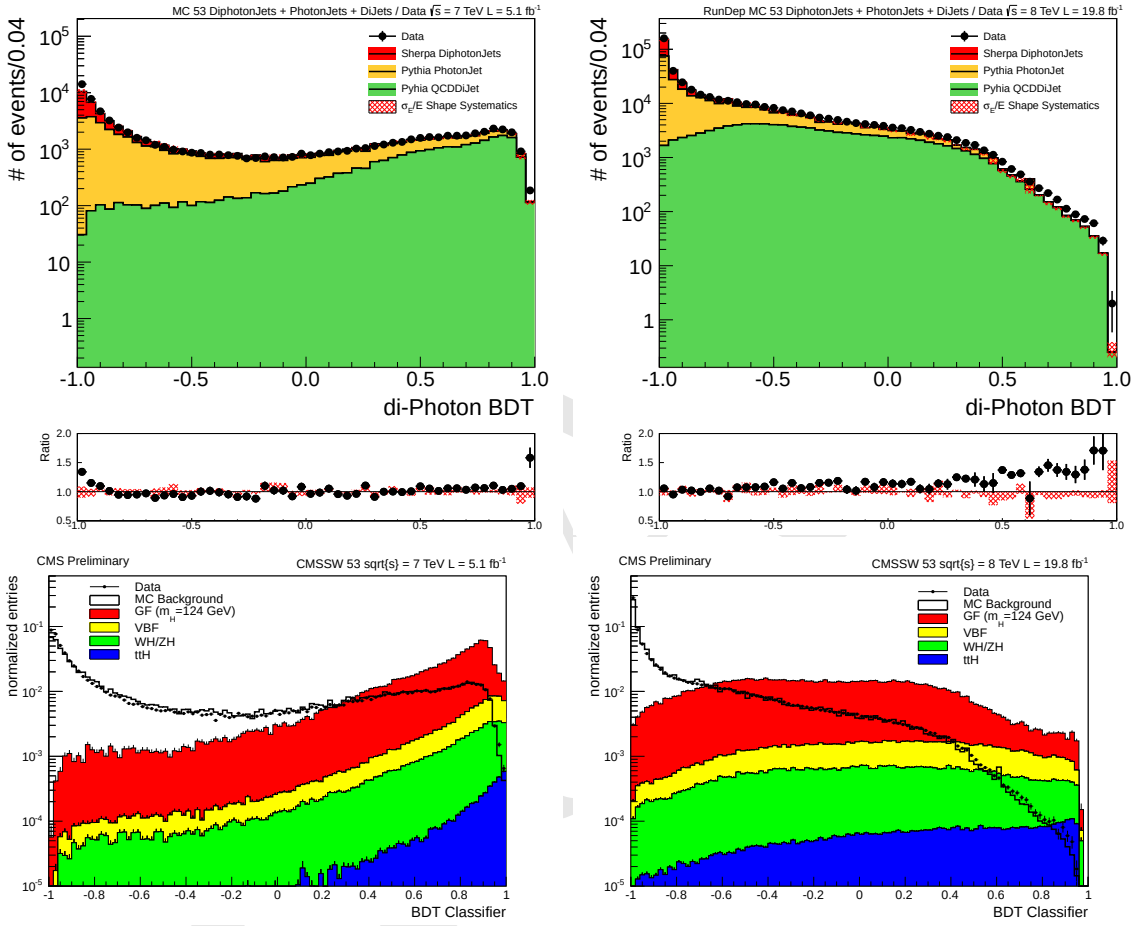

Figure 4.4.: The diphoton BDT response for the 7 TeV training (left column) and 8 TeV training (right column). The data and background distributions are shown in the plots on the top row for data in the range $100 < m_{\gamma\gamma} < 120$ GeV and $130 < m_{\gamma\gamma} < 180$ GeV (black points) and for the prompt-prompt background (green), prompt-fake background (yellow) and fake-fake background (red). The signal distributions are shown in the plots on the bottom row for gluon fusion (red), vector boson fusion (yellow), associated W, Z production (green) and associated $t\bar{t}$ production (blue) alongside the data in the range $100 < m_{\gamma\gamma} < 120$ GeV and $130 < m_{\gamma\gamma} < 180$ GeV (black points) and the total background (hollow histogram). PLOTS NEED UPDATING

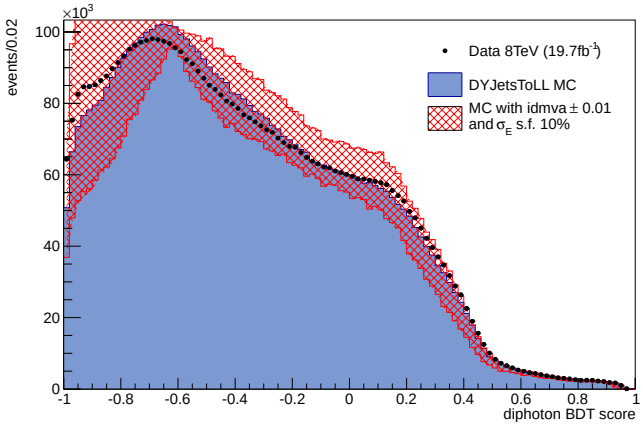

Figure 4.5.: The diphoton BDT response for the 8 TeV training in the $Z \to e^+e^-$ decay. The data is shown as the black points and the MC as the blue histogram. The systematic applied to account for variation in the BDT response from mismodelling in the photon quality response and the photon energy resolution estimate are shown as the red band. This is just a placeholder for the plot I want to show

Whilst the background in this analysis is obtained in a completely data driven way the signal model is obtained from MC. Consequently any uncertainties which effect the shape of the output distribution of this BDT in the signal result in event migrations between the final analysis categories. The input variables whose uncertainties have the largest effect on the BDT response in signal are the photon ID quality and the photon energy resolution estimate. This is because these variables have both i) relatively large uncertainties because of imperfect detector response modelling in the simulation and ii) are highly discriminative and hence can have a relatively large impact on the BDT response. As these variables both vary monotonically with the BDT response the sytematic uncertainty on them is propagated through the analysis as an event migration. Systematic uncertainties are described in more detail in Sec. 5.4. The size of this effect is shown in the BDT response validation plot with the $Z \to e^+e^-$ decay in Figure 4.5. Validation plots using the $Z \to e^+e^-$ channel for each of the event level BDT input variables is shown in Appendix B (may not be necessary).

4.2. Event categorisation

In order to exploit different regions of phase space with similar signal to background ratios the events are split into categories. Furthermore, additional categories can be

designed to enrich the selection with events characteristic of particular Higgs production modes. The VBF production mode is typically accompanied by a pair of jets with a large pseudorapidity separation, the VH production may be accompanied by a charged lepton, missing transverse energy or jets originating from the decay of the associated W or Z boson. Similarly, $t\bar{t}H$ production may be accompanied by b-quarks and or charged leptons. The predominant production mode, which accounts for about 88% of the signal, is ggH which is inclusive. The amount of signal produced from exclusive modes is approximately 8% for VBF, 4% for VH and <1% for $t\bar{t}H$. By including a series of event tags, and separating events accordingly, all four production modes of the Higgs at the LHC are harnessed in this analysis. This not only helps to increase the overall sensitivty to a Standard Model Higgs boson (given the very low background rates expected for the exclusive modes) but also significantly helps to reduce the error on measurements of the observed bosons respective couplings with fermions and bosons as the signal from the relative production modes gets split into distinct categories. The inclusive mode categorisation is done differently for each of the three analyses, a summary of all the event classifications is given in Table 4.4 at the end of the chapter.

4.2.1. Exclusive mode tagging

All events which make it to this stage will have passed the photon preselection (Sec. 3.5), as well as the basic requirement of two high p_{T} photons with $100 < m_{\gamma\gamma} < 180$ (Sec. 4.1), and either the CiC selection (Sec. 4.1.1) or the MVA selection (Sec. 4.1.3) and these make up the final event sample. They now pass through the "tagging" procedure described below in which they are organised into a set of non overlapping event classes. The tagging is done in a specific order to ensure there is no overlap between classes and the order is chosen such that preference is given to categories with a higher expected signal to background ratio. This order is shown alongside a summary of the relevant cut values in Table 4.4 at the end of the chapter. The dijet selection for the MVA analyses uses a specific combined dijet-diphoton BDT to exploit the very particular kinematic properties of VBF whereas for the CiC analysis this is left as a cut based criteria. If an event does not meet the requirements of a particular tag it is passed onto the next tag and if it fails all tag requirements it is placed in one of the inclusive categories, whose structure is described below (Secs. 4.2.2- 4.2.4), meaning that no event can now not be included in the analysis.

DRAFT

Variable	tight category	loose category
$p_T^{\gamma_1}/m_{\gamma\gamma}$	> 0.5	> 0.5
$p_T^{\gamma_2}$	> 25 GeV	> 25 GeV
$p_T^{\mathrm{j}_1}$	> 30 GeV	> 30 GeV
$p_T^{\mathrm{j}_2}$	> 30 GeV	> 20 GeV
$\lvert\Delta\eta_{j_1 j_2}\rvert$	> 3.0	> 3.0
$\lvert Z\rvert$	< 2.5	< 2.5
$M_{j_1 j_2}$	> 500 GeV	> 250 GeV
$\lvert\Delta\phi(\mathrm{jj},\gamma\gamma)\rvert$	> 2.6	> 2.6

Table 4.2.: Final selection cuts for the VBF selection. Events from the first category are removed from the second one.

Dijet tagged categories for VBF

The following variables are used to exploit the specific topolgy of the jet pairs associated to VBF Higgs production,

- $p_T^{\gamma}/m_{\gamma\gamma}$ - The mass relative transverse momenta of the leading and subleading photons,
- $p_T^{\mathrm{j}}/m_{\gamma\gamma}$ - The mass relative transverse momenta of the leading and subleading jets,
- m_{jj} - The dijet invariant mass,
- $\Delta\eta_{\mathrm{j_1 j_2}}$ - The pseduorapidity difference between the two jets.
- $Z = \eta(\gamma_1 + \gamma_2) - \big(\eta(\mathrm{j}_1) + \eta(\mathrm{j}_2)\big)/2$ - The so called *Zeppenfeld* variable [26],
- $\Delta\phi_{\mathrm{j_1 j_2}}$ - The angular difference between the two jets in the transverse plane.

Additionally the leading photon p_{T} requirement is raised to $p_T^{\gamma_1} > m_{\gamma\gamma}/2$. The energy measurement of jets in the event are calibrated to correct for detector effects [27] and additional energy in the jets from pileup is removed using the FastJet jet areas technique described in [28–30]. Jets are required to be within the pseudorapidity range, $\eta < 4.7$.

For the CiC analysis the dijet tagging is done in a cut based way and tagged events are placed into two categories as per the cuts defined in Table. 4.2.

DRAFT

For the MVA analyses the dijet tagging is done with use of two additional MVAs. The first is designed to exploit the VBF kinematic properties and the second is used to combine this information with the diphoton BDT. The input variables for the kinematic dijet BDT are identical to those used in the cut-based VBF selection where candidates are required to pass a VBF preselection of two jets with $p_T^{j1} > 30$ GeV and $p_T^{j2} > 20$ GeV and invariant mass, $m_{jj} > 75$ GeV. The signal sample used for training is the Standard Model MC with just VBF production, whilst the Standard Model gluon fusion MC is included as background along with the usual prompt-prompt, prompt-fake and fake-fake contributions. This helps produce an output in which a high score gives a very pure VBF sample.

The combined dijet-diphoton BDT has inputs of the kinematic dijet BDT output, the diphoton BDT output and $p_T^{\gamma\gamma}/m_{\gamma\gamma}$ in order to discriminate VBF from both the other signal types and the background, utilising all the information avaiable in the event. The mass relative transverse momenta of the diphoton system, $p_T^{\gamma\gamma}/m_{\gamma\gamma}$, is included because of its considerable correlation with both the dijet BDT output and the diphoton BDT output.

One finds that the background rejection for VBF is significantly improved by the use of the combined dijet-diphoton BDT whilst the VBF purity (i.e. separation from ggH) is not as good when collapsing the kinematic dijet BDT and the combined BDT into one step. Consequently the trainings are performed separately and the VBF categories are defined by picking out regions which have a high score in the combined dijet-diphoton BDT response. The optimisation procedure for deciding the category boundaries is analogous to the one used for the inclusive categories in the MFM analysis where the target is to minimise the expected uncertainty of the signal strength from the VBF process alone when moving the boundaries around. It is explained further in Section 4.2.3. At 8 TeV there are three VBF categories and at 7 TeV there are two. The output distributions for the signal and background with the cut boundaries are shown for the two dijet BDTs are shown in in Figures 4.6 and 4.7. Each successive BDT training uses statisically independent MC samples to avoid selection bias from fluctuations in the simulation.

Lepton, jet and $\not{E}_T$ tagged categories for VH

The selection for the four categories designed to tag VH production are optimised by minimising the expected uncertainty on the signal strength of this process alone. Two

DRAFT

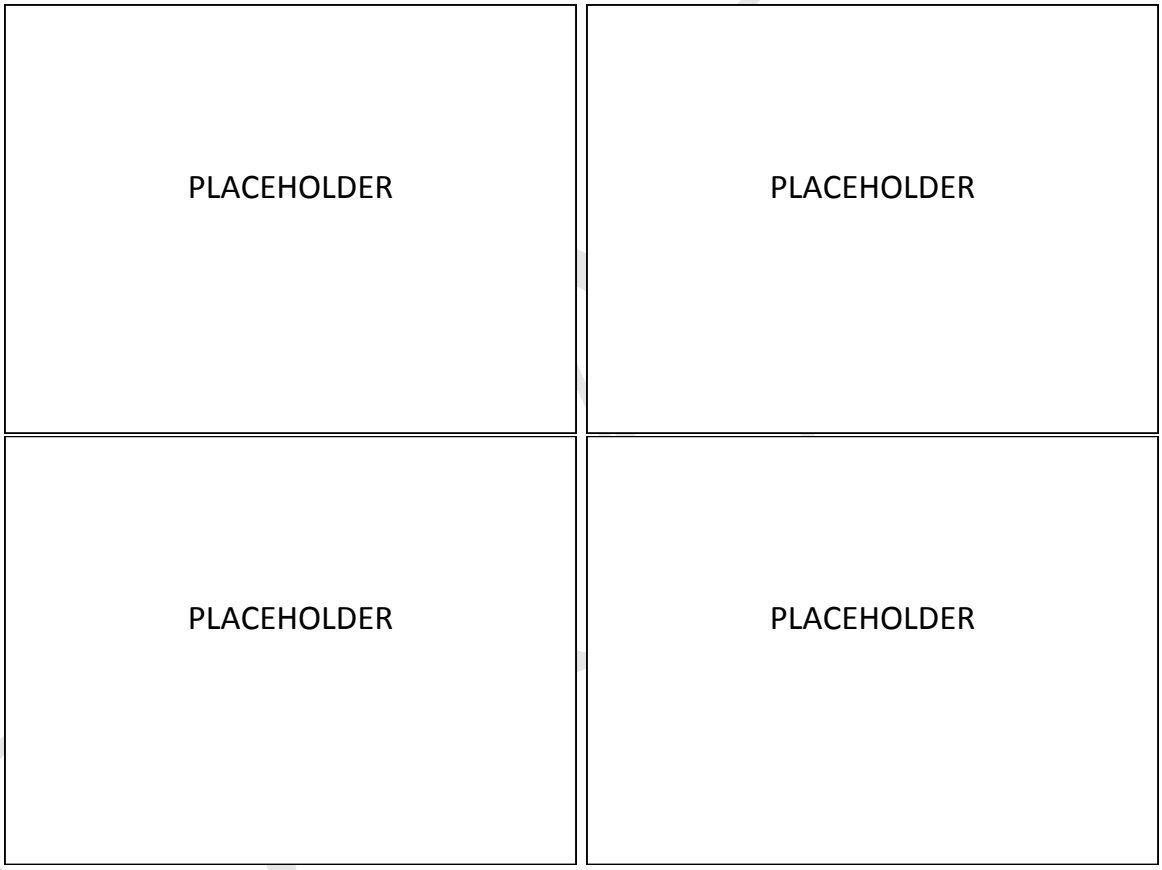

Figure 4.6.: The distributions of the kinematic dijet BDT response for the VBF signal (left) and the background (right) for the 7TeV (top row) and 8TeV (bottom row) training.

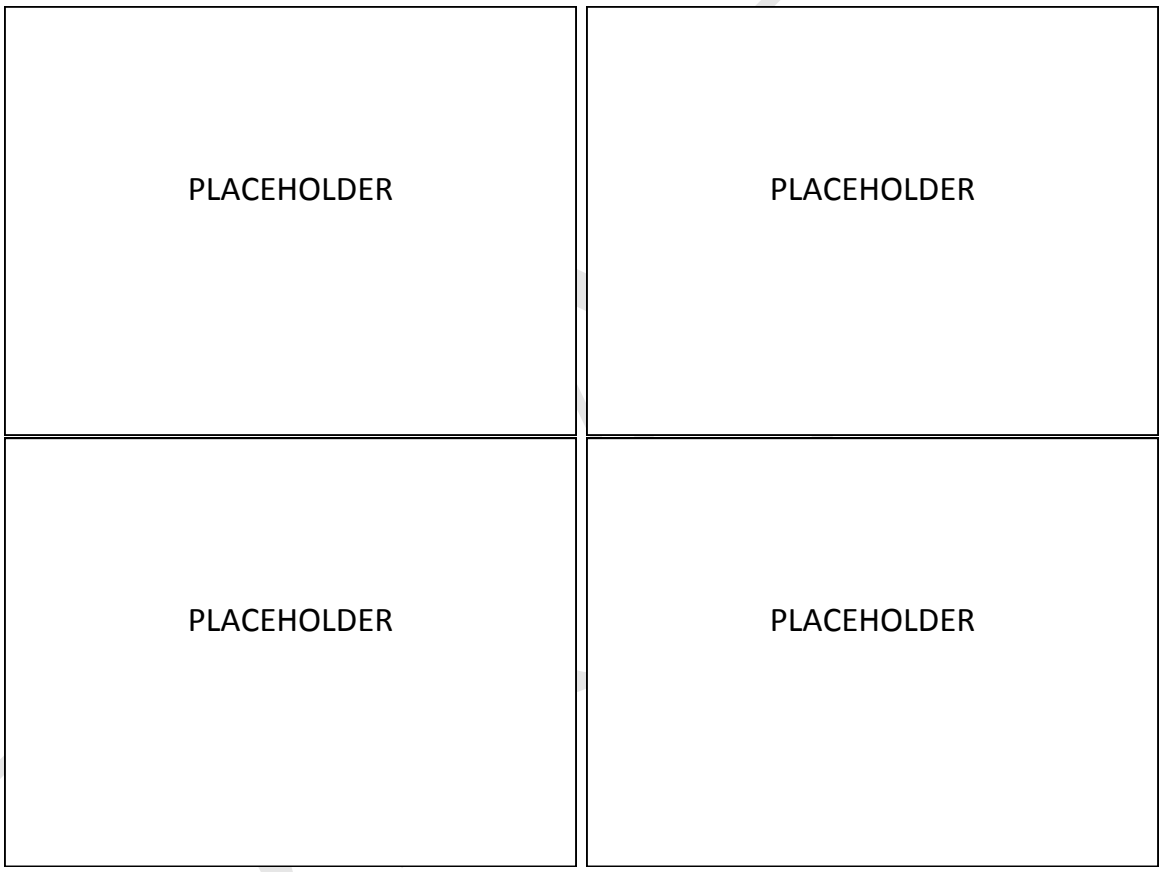

Figure 4.7.: The distributions of the combined dijet-diphoton BDT response for the VBF signal (left) and the background (right) for the 7TeV (top row) and 8TeV (bottom row) training. The red dahsed lines corresponded to the location of the category boundaries.

of the classes require at least one charged muon or electron and are split into a tight selection category and a looser selection category, the third is for events consistent with large $\not\!\!E_T$ and the fourth for events consistent with two or more jets. The leading photon cut is raised to $p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$ for the lepton and $\not\!\!E_T$ tagged categories and $p_T^{\gamma 1} > m_{\gamma\gamma}/2$ for the jet tagged category. The category requirements are as follows,

- **VH Tight *l***: The tighly selected lepton category is characterised by the signature of a leptonically decaying W or Z boson and as such requires the presence of $\not\!\!E_T$ > 35 GeV or another lepton of the same flavour and opposite charge as the first. In the first case (of a lepton + $\not\!\!E_T$) the lepton is required to have $p_T > 20$ GeV, in the latter case (of two leptons) the requirement is $p_T > 10$ GeV for both leptons whilst the invariant mass of the dilepton pair must be $70 < m_{ll} < 110$ GeV. The diphoton BDT output is required to be > 0.1 (> -0.6) for the 7 (8) TeV datasets.

- **VH Loose *l***: For the loosely selected lepton category the lepton p_{T} must satisfy $p_T > 20$ GeV. The selection requirements are designed to reduce the background from leptonic Z bosons (not associated with a Higgs) that contain initial or final state radiation faking the diphoton signal. Consequently, leptons are required to be separated by at least $\Delta R > 1.0$ from the closest photon and the invariant mass of any electron-photon pairs must be more than 10 GeV away from the Z boson mass. In addition a conversion veto is applied to the electrons to reduce the rate of misidentified photons. The diphoton BDT output is required to be > 0.1 (> -0.6) for the 7 (8) TeV datasets.

- **VH $\not\!\!E_T$ tag**: Accurate measurement and simulation of the $\not\!\!E_T$ vector has been studied in detail at CMS and a set of standard corrections (for both data and simulation) are applied [31]. The corrected $\not\!\!E_T$ is required to pass $\not\!\!E_T > 70$ GeV whilst the angular separation in the transverse plane betwen the diphoton system and the $\not\!\!E_T$ direction must pass $\Delta\phi(\gamma\gamma, \not\!\!E_T) > 2.1$, and similarly the angle between the diphoton system and the leading jet must pass $\Delta\phi(\gamma\gamma, \mathrm{jet}) < 2.7$. The diphoton BDT output is required to be > 0.8 (> 0.0) for the 7 (8) TeV datasets.

- **VH jet tag**: The event must contain at least one jet pair in which both jets have $p_T > 40$ GeV and $|\eta| < 2.4$ and have an invariant mass within the range $60 < m_{jj} < 120$ GeV. The diphoton transverse momenta must satify $p_T^{\gamma\gamma} > 13m_{\gamma\gamma}/12$. Additionally the angular correlation between the diphoton system and the dijet system from VH associated Higgs decays can be exploited. The angle, $\theta^\star$, between the diphoton direction in the diphoton-dijet rest frame and the lab frame is flat for

DRAFT

events from VH decays whereas for the background and gluon fusion produced Higgs decays the distribution peaks at $|cos(\theta^{\star})| = 1$. Consequently, there is a requirement that $|cos(\theta^{\star})| < 0.5$. The diphoton BDT output is required to be > 0.6 (> 0.2) for the 7 (8) TeV datasets.

Lepton and jet tagged categories for $t\bar{t}H$

There are two categories for tagging production from $t\bar{t}H$ decays, one which is lepton based and one which is jet based. The total fraction of signal expected from $t\bar{t}H$ is <1% so only a handful of events are expected. Consequently for the 7 TeV dataset the two categories are merged into one class. As for the VH tagged categories the cuts are optimised to minimise the expected uncertainty of the signal strength measurement of the $t\bar{t}H$ process alone.

For both classes the leading photon p_{T} cut is raised to $p_T^{\gamma 1} > m_{\gamma\gamma}/2$, all jets are required to have $p_{\mathrm{T}} > 25$ GeV and there must be at least one b-tagged jet present. The specific requirements of each category are as follows,

- **$t\bar{t}H$ multijet tag**: The requirement is at least four additional jets in the event and no lepton. The diphoton BDT output is required to be > 0.6 (> -0.2) for the 7 (8) TeV datasets.
- **$t\bar{t}H$ lepton tag**: At least one more jet in the event and one muon or electron which has $p_{\mathrm{T}} > 20$ GeV. The diphoton BDT output is required to be > 0.6 (> -0.6) for the 7 (8) TeV datasets.

4.2.2. Inclusive mode categorisation in the cut based analysis

Any event which passes the cut based photon selection described in Section 4.1.1 and does not fall into one of the exclusive categories described above is split into one of eight inclusive categories depending on the supercluster position, η, of the two photons, the conversion variable, R_9, of the two photons and the mass relative diphoton transverse momenta, $p_T/m_{\gamma\gamma}$. The categories are defined in Table 4.3.

DRAFT

	$p_T/m_{\gamma\gamma}$	Maximum η	Minimum R_9
Untagged 0	>40 GeV	<1.444	>0.94
Untagged 1			<0.94
Untagged 2		<2.5	>0.94
Untagged 3			<0.94
Untagged 4	<40 GeV	<1.444	>0.94
Untagged 5			<0.94
Untagged 6		<2.5	>0.94
Untagged 7			<0.94

Table 4.3.: The definition of the inclusive categories for the CiC analysis

4.2.3. Inclusive mode categorisation and VBF dijet categorisation in the mass factorised MVA analysis

Firstly the combined dijet-diphoton BDT score is used to define a set of VBF categories. The goal is to find the configuration of category boundaries which minimise the expected uncertainty on the signal strength for VBF production alone, allowing the number of categories and where the category boundaries lie to be completely free floating, with the additional requirement that the efficiency $\times$ acceptance of the categories matches between 7 and 8 TeV. This results in a tight VBF category for events with a very high combined dijet-diphoton BDT score, a somewhat looser category and then one or more very loose categories. One finds that dropping the loosest category has a neglible impact (<1%) on the expected uncertainty and as such the upper boundary for the loosest category is turned into a lower cut. All events which pass the VBF preselection (described in Sec. 4.2.1) and fail the lower cut are then classified somewhere in the inclusive categories defined below.

Once the VBF category boundaries have been found the same procedure is deployed using the diphoton BDT score. This time the target is to minimise the expected uncertainty on the total signal strength, allowing the number of categories and the category boundary values to be completely free floating. One finds a rather similar structure and that dropping events in the very loosest category has a neglible impact on the performance and consequently this dictates the lower cut value for the diphoton BDT.

DRAFT

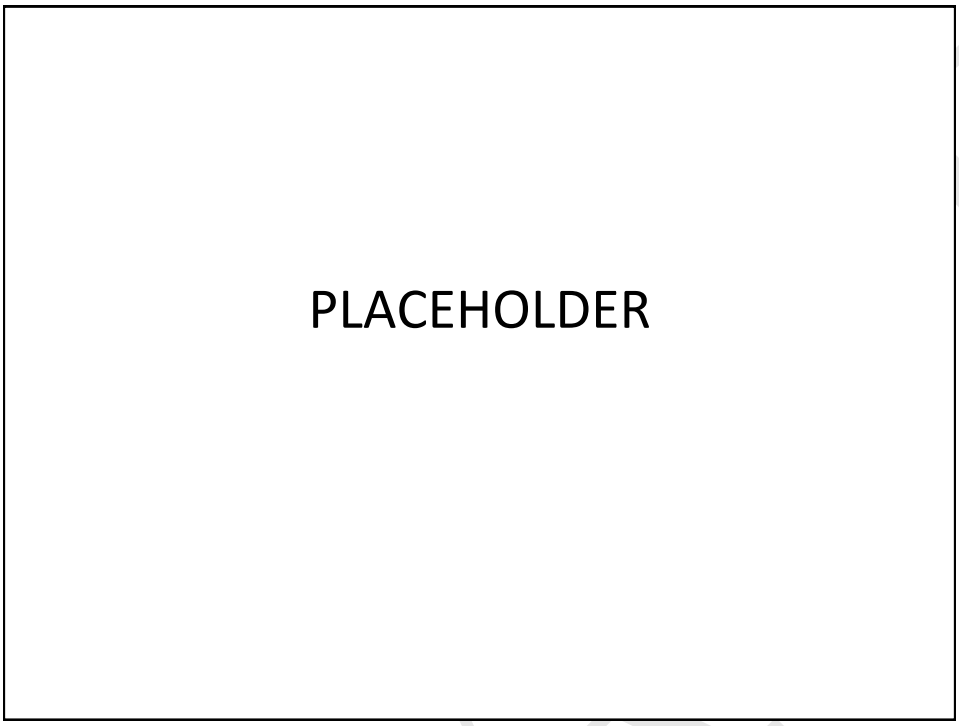

Figure 4.8.: A pictoral representation of how the VBF and inclusive category boundaries are optimised

Due to the amount of simulation MC available to do these studies the precision to which the boundary values are important is quite broad. Furthermore, the expected uncertainty minima are very shallow so the exact position of the boundaries has a very small impact on the performance of the analysis. For the 7 (8) TeV datasets there are 4 (5) inclusive categories with a lower diphoton BDT cut of 0.19 (-0.78) and 2 (3) VBF categories. The boundary optimisation procedure is pictorally represented in Figure 4.8.

4.2.4. Inclusive mode categorisation in the sideband MVA analysis

In the SMVA analysis all the exclusive categories are identical to the MFM analysis (including the VBF categories). However the inclusive categorisation is done slightly differently. In the sideband analysis the signal region is defined in a $\pm 2\%$ window around the hypothesis Higgs mass. The analysis is performed as a cut and count in the signal window over several bins. There is one bin for each exclusive category and then several more for the inclusive. The binning scheme for the inclusive events is defined as follows,

- Make two dimensional distributions of the diphoton BDT score and the distance of the invariant mass from the hypothesised Higgs mass, $\Delta m/m_H$, in the $\pm 2\%$ window for signal and background, as shown in Figure 4.9, where,

$$\Delta m/m_H = \frac{m_{\gamma\gamma} - m_H}{m_H}. \tag{4.4}$$

- Select bins by isolating regions of this 2D phase space which have similar S/B ratios and optimise the boundaries to give the maximum expected signal significance.

Clearly the most sensitive bins will be the ones which have a high diphoton BDT score and have a low value of $|\Delta m/m_H|$ (i.e. are near the signal peak). The category boundaries in this 2D plane are shown as different shades in Figure 4.10. In total there are 8 (10) inclusive bins for the 7 (8) TeV samples in the SMVA.

4.2.5. Event categorisation summary

All the categories and the tagging order are summarised in the table below (Table 4.4).

DRAFT

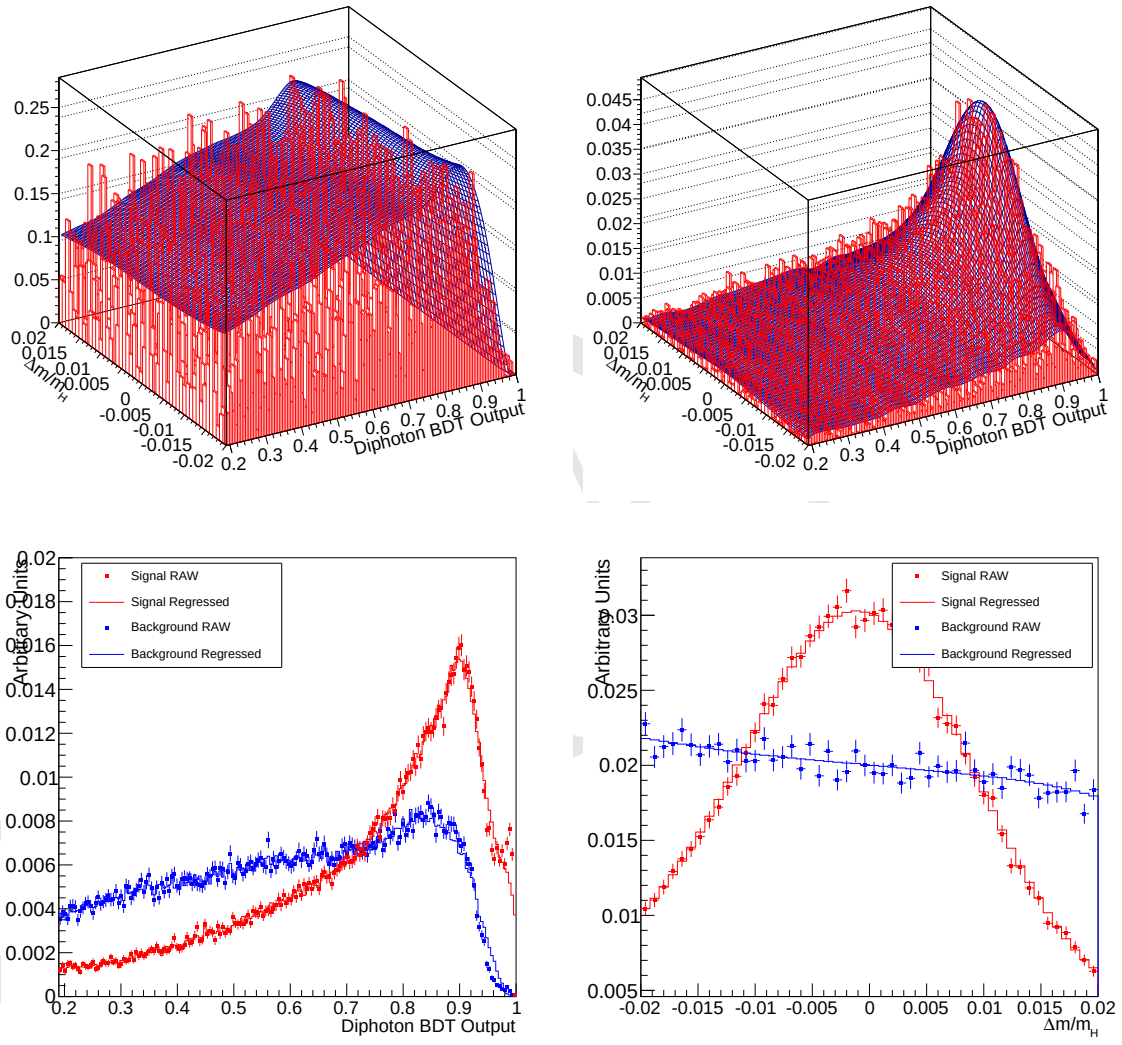

Figure 4.9.: Two dimensional distributions of the diphoton BDT output and $\Delta m/m_H$ are shown on the top row for the background (left) and signal (right) for the 7 TeV sample. The bottom rows show the projection for signal (red) and background (blue) in the two variables.

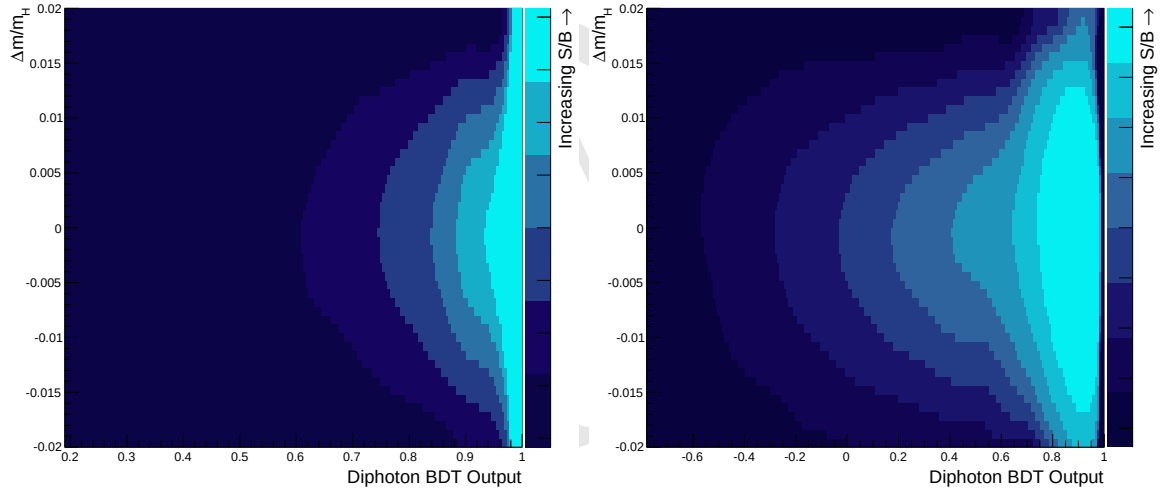

Figure 4.10.: The inclusive category bin definitions for the SMVA analysis. Shown for the 7 TeV dataset on the left and the 8 TeV dataset on the right.

Table 4.4.: The event classes at 7 and 8TeV and some of their main selection requirements. Events are tested against the selection requirements of the classes in the order they are listed here.

Label	No. of classes 7GeV	8GeV	Main requirements
$t\bar{t}H$ lepton tag	$\star$	1	$p_T^\gamma(1) > m_{\gamma\gamma}/2$ 1 b-tagged jet + 1 electron or muon **For MVAs:** diphoton BDT $> 0.6(-0.6)$ at 7(8) TeV
VH tight ℓ tag	1	1	$p_T^\gamma(1) > 45 \cdot m_{\gamma\gamma}/120$ e or μ, $p_{\mathrm{T}} > 20GeV$, and $\not{E}_T > 45GeV$ OR $2e$ or 2μ, $p_{\mathrm{T}} > 10GeV$; $70 < m_{ll} < 110GeV$ **For MVAs:** diphoton BDT $> 0.1(-0.6)$ at 7(8) TeV
VH loose ℓ tag	1	1	$p_T^\gamma(1) > 45 \cdot m_{\gamma\gamma}/120$ e or μ, $p_{\mathrm{T}} > 20GeV$ **For MVAs:** diphoton BDT $> 0.1(-0.6)$ at 7(8) TeV
VBF dijet tag	2	2/3†	$p_T^\gamma(1) > m_{\gamma\gamma}/2$ **For MVAs:** 2 jets; dijet and combined diphoton-dijet BDTs used **For CiC:** 2 jets, cut based dijet selection used
VH $\not{E}_T$ tag	1	1	$p_T^\gamma(1) > 45 \cdot m_{\gamma\gamma}/120$ $\not{E}_T > 70GeV$ **For MVAs:** diphoton BDT $> 0.8(0.0)$ at 7(8) TeV
$t\bar{t}H$ multijet tag	$\star$	1	$p_T^\gamma(1) > m_{\gamma\gamma}/2$ 1 b-tagged jet + 4 more jets **For MVAs:** diphoton BDT $> 0.6(-0.2)$ at 7(8) TeV
VH dijet tag	1	1	$p_T^\gamma(1) > 45 \cdot m_{\gamma\gamma}/120$ jet pair, $p_{\mathrm{T}} > 40GeV$ and $60 < m_{\mathrm{jj}} < 120GeV$ **For MVAs:** diphoton BDT $> 0.6(0.2)$ at 7(8) TeV
Untagged	4/8‡	5/8/10‡	The remaining events, **For MVAs:** classified using diphoton BDT **For CiC:** classified using η and R_9 of photons and $p_T^{\gamma\gamma}/m_{\gamma\gamma}$

$\star$ For the 7TeV dataset, events in the $t\bar{t}H$ lepton tag and multijet tag classes are combined, after selection, to form a single event class.

† For the CiC (MVA) analysis there are 2 (3) dijet categories at 8 TeV.

‡ For the CiC analysis there are 8 inclusive categories at 7 and 8 TeV, for the MFM there are 4 (5) categories at 7 (8) TeV and for the SMVA there are 8 (10).

DRAFT

Chapter 5.

Analysis and Results

"I'm a bus"

— Darren Burton 1987–2013

5.1. Description

This would include a detailed description of the signal and background models used and the systematics included. I would describe the statistical tools used such as hypothesis testing with a test statistic and show the results which would include exclusion limits, pvalues and SM best fit parameters (mass, signal strength, couplings etc.). One of the important considerations of the Hgg analysis is the estimation of the background. This second part of this section will concentrate on presenting an alternate analysis which takes a completely different approach to modeling the background and cross-checks the result above.

DRAFT

5.2. Signal modelling

5.2.1. Mass factorised analysis

5.2.2. Sideband analysis

5.3. Background modelling

5.3.1. Mass factorised analysis

5.3.2. Sideband analysis

5.4. Systematic Uncertainties

5.5. Statistics

5.5.1. Use of the Likelihood function as a test statistic

5.6. Results of the mass factorised analysis

5.7. Results of the sideband analysis

DRAFT

Appendix A.

Photon ID BDT input variables validation in $Z \to e^+e^-$

DRAFT

DRAFT

Appendix B.

Diphoton BDT input variables validation in $Z \to e^+e^-$

DRAFT

DRAFT

1247

DRAFT

DRAFT

Bibliography

[1] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[2] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2**, 1285 (1970).

[3] S. Willenbrock, (2004), hep-ph/0410370.

[4] The CMS Collaboration, JINST **3**, S08004 (2008).

[5] The CMS Collaboration, Eur. Phys. J. C **70**, 1165 (2010), hep-ex/1007.1988.

[6] The CMS Collaboration, (2012), CMS-DP-2013-007.

[7] E. Meschi, T. Monteiro, C. Seez, and P. Vikas, (2001), CMS-AN-2001-034.

[8] W. Adam, R. Früwirth, A. Strandlie, and T. Todorov, J. Phys. G: Nucl. Part. Phys. **31**, N9 (2005).

[9] J. Weng, (2008), hep-ex/0810.3686v1.

[10] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. **04**, 063 (2008).

[11] C. Ferro, (2012), hep-ex/1201.5292.

[12] The CMS Collaboration, J. Phys. Conf. Ser. **404**, 012014 (2012).

[13] The CMS Collaboration, (2009), CMS-PFT-09-001.

[14] A. Hoecker *et al.*, (2007), physics/0703039.

[15] F. J. Massey, JSTOR **46**, 68 (1951).

[16] S. Alioli, P. Nason, C. Oleari, and E. Re, JHEP **04**, 002 (2009).

[17] P. Nason and C. Oleari, JHEP **02**, 037 (2010).

[18] T. Sjöstrand, S. Mrenna, and P. Z. Skands, JHEP **0605**, 026 (2006).

[19] LHC Higgs Cross Section Working Group, S. Heinemeyer *et al.*, (2013), 1307.1347.

DRAFT

[20] Y. Gao *et al.*, Phys. Rev. D **81**, 242 (2010).

[21] T. Gleisberg *et al.*, JHEP **0902**, 007 (2009), 0811.4622.

[22] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, JHEP **1106**, 128 (2011), 1106.0522.

[23] GEANT4, S. Agostinelli *et al.*, Nucl. Instrum. Meth. A **506**, 250 (2003).

[24] J. Beringer *et al.*, Phys. Rev. D **86**, 010001 (2012).

[25] The CMS Collaboration, JHEP **2011**, 1 (2011).

[26] R. Szalapski and D. Zeppenfeld, Phys. Rev. D **54**, 6680 (1996), arXiv:hep-ph/9605444.

[27] CMS, S. Chatrchyan *et al.*, JINST **06**, P11002 (2011).

[28] M. Cacciari and G. P. Salam, Phys. Lett. **B659**, 119 (2008).

[29] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 005 (2008).

[30] M. Cacciari, G. P. Salam, and G. Soyez, (2011), 1111.6097.

[31] T. C. Collaboration, CMS PAS **CMS-PAS-JME-12-002** (2012).

DRAFT

List of Figures

DRAFT

DRAFT

DRAFT

DRAFT

DRAFT

DRAFT

List of Tables

DRAFT

List of Acronyms

BDT	-	Boosted Decision Tree
DT	-	Decision Tree
CMS	-	Compact Muon Solenoid
ECAL	-	electromagnetic calorimeter
HCAL	-	hadronic calorimeter
LHC	-	Large Hadron Collider
MVA	-	Multivariate Analysis
CiC	-	Cuts in Categories
MFM	-	Mass Factorized MVA
SMVA	-	Sideband MVA
$PbWO_4$	-	lead tungstate
APDs	-	avalance photodiodes
VPTs	-	vacuum phototriodes
GED	-	global event description
PF	-	particle flow
MC	-	Monte Carlo
QCD	-	quantum chromodynamics
ggH	-	gluon fusion
VBF	-	vector boson fusion

DRAFT

VH - vector boson associated production

$t\bar{t}H$ - top quark associated production

DRAFT