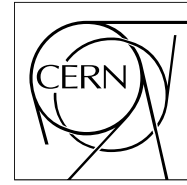


The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



9 June 2006

Inclusive Search for the Higgs Boson in the $H \rightarrow \gamma\gamma$ Channel

M. Pieri¹⁾, S. Bhattacharya²⁾, I. Fisk³⁾, J. Letts¹⁾, V. Litvin⁴⁾, J.G. Branson¹⁾

Abstract

We have carried out a detailed study of the inclusive search for the Higgs boson in the $H \rightarrow \gamma\gamma$ channel with the CMS detector at LHC. The analysis is based on full Monte Carlo simulation. Though for the moment we have only simulated data to work with, we have designed an analysis that can determine the background from real data and will thus only depend on signal Monte Carlo when data will be available. This largely reduces the systematic error and, together with the excellent energy resolution of the CMS electromagnetic calorimeter, enables a search that would otherwise be very difficult because of the large amount of expected background. We have studied both a standard cut-based search and a more optimized analysis that takes advantage of the wide range of signal/background expectations as function of the possible selection cuts. Since discovery in this channel is expected to take one or more years of LHC running, such optimized analyses should be studied to minimize the time to discovery and to assure that the experiment is competitive in this important measurement.

¹⁾ University of California San Diego, La Jolla, CA, USA

²⁾ University of Delhi, Delhi, India

³⁾ FNAL, Batavia, IL, USA

⁴⁾ California Institute of Technology, Pasadena, CA, USA

1 Introduction

The $H \rightarrow \gamma\gamma$ channel has been studied since the initial planning of the LHC and SSC as an important channel for the discovery of Higgs particles at masses between the LEP limit and about 150 GeV [1]. The signature sought in the inclusive analysis is two high E_t isolated photons with an invariant mass consistent with the Higgs mass. Discovery of a Higgs in this low mass range is more difficult than at higher masses since the dominant decay mode of the Higgs is $b\bar{b}$ which is hard to separate from background. The $\gamma\gamma$ decay mode can be identified well experimentally but has a small branching fraction of about 0.002 in the most important mass range. The signal rate is rather small compared to backgrounds with two real photons (irreducible), and backgrounds in which one or more of the photons are due to misidentified particles from jets (reducible). After the selection, for an integrated luminosity of 20 fb^{-1} and for a Higgs boson mass of 120 GeV, we expect approximately 350 signal events in a mass window of 2 GeV over 7000 background events.

The rates of the various backgrounds are quite uncertain at LHC energies. Nevertheless, it has long been understood that $H \rightarrow \gamma\gamma$ can be identified as narrow mass peak above a large background. The background rate can be determined from the region outside the peak. A low mass Higgs can then be discovered in this channel and its mass can be measured well.

In this study we present two complementary inclusive analyses for the $H \rightarrow \gamma\gamma$ channel: a standard cut based analysis and a high performance, discovery-oriented analysis, based on the method described in [2, 3]. Both are carried out with our present knowledge of the expected background, estimated with full detector simulation. When data will be available we will not make use of Monte Carlo background information and only the signal Monte Carlo will be used. We concentrate on the first years of LHC operation when the luminosity will be $2 \times 10^{33} \text{ cm}^{-2}\text{sec}^{-1}$.

We extend the idea of measuring the rate of background by using the mass regions adjoining the Higgs peak to also measure the characteristics of the background that can be used to help separate background from signal. The $H \rightarrow \gamma\gamma$ channel is particularly well suited for this technique because the signal is small compared to the expected background and can be confined to a narrow mass region thanks to the excellent photon energy and position resolution of the CMS detector [4].

This study requires a comprehensive understanding and simulation of the CMS detector. In particular, the electromagnetic calorimeter is used to make the primary measurements of photon energy and position. The tracker is used to measure the position of the interaction at which the Higgs candidate was produced. The tracker, ECAL and HCAL are used to determine if the photon candidate is well isolated. While we will be able to measure background characteristics from data, the signal must be well simulated to perform our analysis. This implies a detailed understanding of the detector performance as well as its calibration.

By using photon isolation and photon kinematic information, significant additional discrimination between signal and background can be achieved. The optimized analysis can be trained¹⁾ to use this information to discriminate between signal and background by comparing data in mass sidebands with signal Monte Carlo.

In this approach, the expected purity in terms of signal/background, corresponding to each event, can be estimated based on these quantities and each event then can be used optimally to evaluate the likelihood of a signal plus background hypothesis compared to a background-only hypothesis.

We will show that such a discovery-oriented analysis can be performed. Since we do not yet have data, we must use simulated backgrounds. With this simulation, we can calculate a benchmark performance of such an analysis. It is not our main goal to predict exactly what luminosity will be required to discover the Higgs. The rate and characteristics of the backgrounds are quite uncertain at this time and such a prediction would therefore have a large uncertainty. Rather, we study the statistical and systematic uncertainties of a future discovery based on the type of analysis proposed.

¹⁾ Training can include the optimization of cuts or variables for use in discrimination. In this study we have used a neural net, however, likelihood variables or other techniques may prove to be better in the future.

2 Higgs Boson Production and Decay

The inclusive $H \rightarrow \gamma\gamma$ analysis by its nature studies all channels in which the Higgs boson decays into two photons. Since this analysis focuses on the measurement of the photons, the different production channels are not separated by tagging with other particles. For this analysis we have simulated the largest production modes in the Standard model. The Higgs boson is produced through gluon fusion, that is the dominant process, qqH through Weak Vector Boson Fusion (WVB), associated with W or Z bosons, and associated with a $t\bar{t}$ pair. In each case we study the mode in which the Higgs decays to two photons. In general, these are isolated high E_t photons. The cross sections for the different production processes [5] and the $H \rightarrow \gamma\gamma$ branching ratios [6] are summarized in Table 1.

Table 1: Next to Leading Order cross sections for the different Higgs boson production processes and $H \rightarrow \gamma\gamma$ branching ratios.

M_H	115 GeV	120 GeV	130 GeV	140 GeV	150 GeV
σ (gg fusion)	39.2 pb	36.4 pb	31.6 pb	27.7 pb	24.5
σ (WVB fusion)	4.7 pb	4.5 pb	4.1 pb	3.8 pb	3.6
σ (WH, ZH, $t\bar{t}H$)	3.8 pb	3.3 pb	2.6 pb	2.1 pb	1.7
Total σ	47.6 pb	44.2 pb	38.3 pb	33.6 pb	29.7
$H \rightarrow \gamma\gamma$ Branching Ratio	0.00208	0.00220	0.00224	0.00195	0.00140
Inclusive $\sigma \times B.R.$	99.3 fb	97.5 fb	86.0 fb	65.5 fb	41.5 fb

While some of the channels may have additional leptons, jets or missing energy, we have not taken advantage of those signatures. Such splitting into identifiable channels will take place in the future and will certainly improve the performance of the analysis. For the moment these channels are investigated individually in other studies [7, 8]. Figure 1 shows an event display of a $H \rightarrow \gamma\gamma$ event with $M_H=120$ GeV.

3 Backgrounds

Backgrounds with two real high E_t photons are called “irreducible”, although their effect can be somewhat reduced due to kinematic differences from signal processes in which high mass particles are produced. Two photons can be produced from two gluons in the initial state through a “box diagram” or from initial quark and anti-quark annihilation. These backgrounds have reasonably small cross sections, approximately 80 pb each, and are therefore straightforward to simulate.

Backgrounds in which at least one final state jet is interpreted as a photon are called “reducible” and are much harder to simulate since jets are copiously produced at the LHC and Monte Carlo samples that correspond to order of 5-10 fb^{-1} are much too large to fully simulate. Selections at generator level have been devised in order to be able to select multi-jet and γ plus jets events that contribute to the background of the $H \rightarrow \gamma\gamma$ channel and reject events that have negligible chance of producing background to the final analysis.

For $pp \rightarrow \gamma + \text{jet}$, a loose pre-selection at generator level²⁾ was employed.

The $\gamma + \text{jet}$ sample can be viewed, from the selection point of view, as coming from two different sources: one where another photon is radiated during the fragmentation of the jet (two prompt photons), the other where there is only one prompt photon in the final state and the other photon candidate corresponds to a mis-identified jet or isolated π^0 (one prompt plus one fake photon). In both analyses described in the following we will separate these

²⁾ Electrons, positrons and photons were considered. The first step of the selection was to look for seed particles with a p_t larger than 5 GeV and $|\eta| < 2.7$. Candidate electromagnetic calorimeter trigger tower energies were then estimated by adding all the electromagnetic particles found within $|\Delta\eta| < 0.09$ and $|\Delta\phi| < 0.09$ from the seed. Trigger tower candidates within $|\Delta\eta| < 0.2$ and $|\Delta\phi| < 0.2$ of each other were then suppressed and only the candidate with the highest p_t was retained. Finally the level-1 double photon electromagnetic trigger was simulated by requiring two such candidates with transverse energies larger than p_{t1LV1} and p_{t2LV1} respectively. The cuts p_{t1LV1} and p_{t2LV1} were chosen to be 30 and 20 GeV and events have been generated with a value of $\hat{p}_\perp > 30$ GeV.

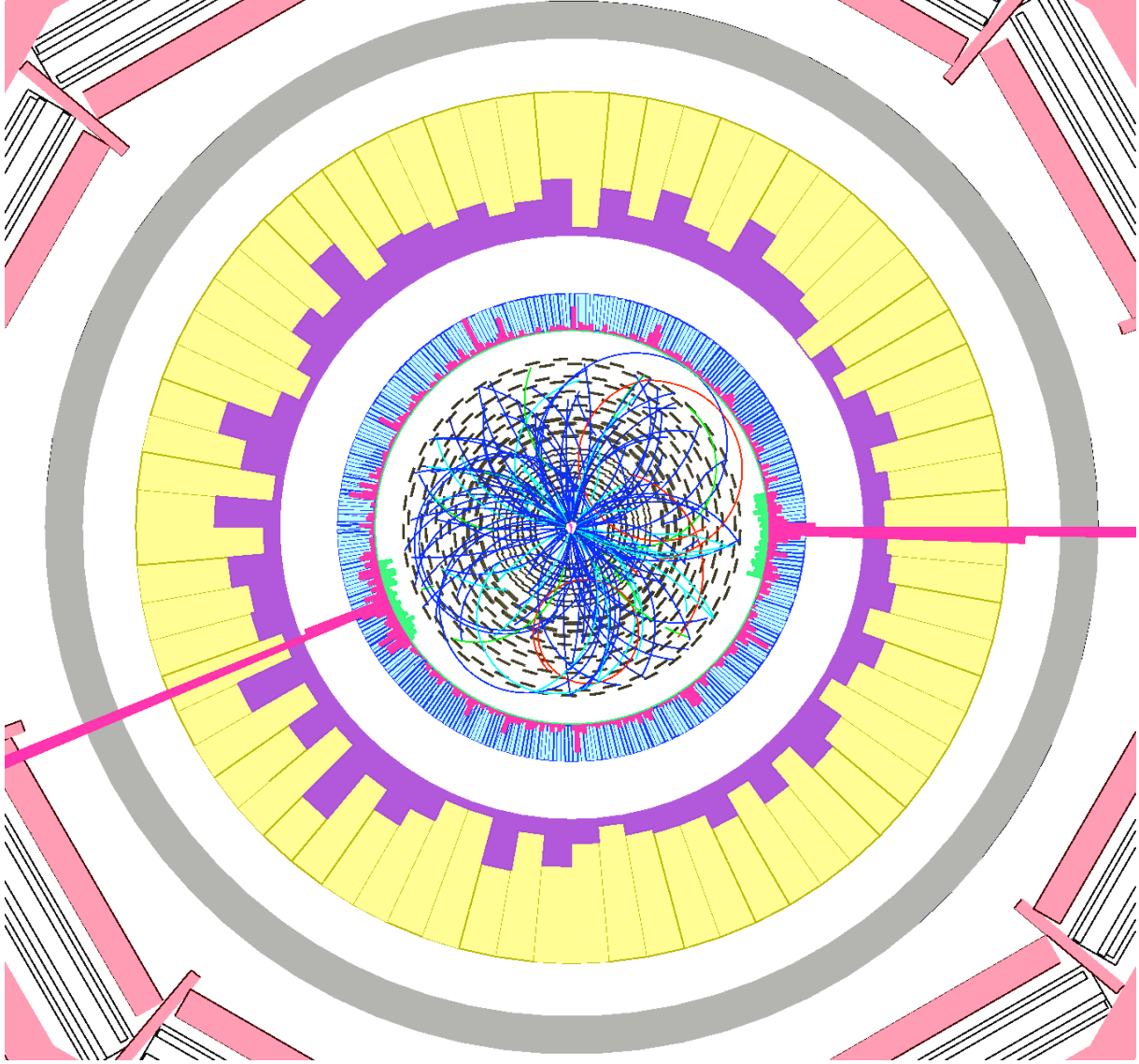


Figure 1: $H \rightarrow \gamma\gamma$ event produced in gluon fusion with $M_H=120$ GeV observed in the CMS detector.

two types of background. For the first type we have the additional requirement of a second prompt photon at generator level, with E_t larger than 10 GeV.

For $pp \rightarrow$ jets, a tighter set of cuts at the particle kinematics level was carefully studied. We devised a selection based on the generated particles aimed at selecting events which could produce two electromagnetic clusters consistent with isolated photons in the detector. In order to apply cuts on the invariant mass of the two candidates we estimate lower and upper limits to the energy of the candidates that will be reconstructed after the simulation. First of all of we look for seed candidates. We consider the following particles: γ , e^\pm , π^0 , η , η' , ρ , ω that could produce electromagnetic showers consistent with photons³⁾ in the detector. The seeds should satisfy the cuts:

- $P_{tseed} > P_{tseedmin}$

³⁾ We found that their contribution of K_{long}^0 to the events selected at analysis level is negligible while the number of events to be simulated is approximately doubled so K_{long}^0 were not included as seeds.

- $|\eta_{\text{seed}}| < \eta_{\text{seedmax}}$

In order to take into account the fact the super-clusters have a sizable spread in η and ϕ and that more than one particle can contribute to the reconstructed energy we do not use these seeds as photon candidates but we sum up all the final state γ s and e^\pm s around their direction. We define the photon candidate as the vector sum of the momenta of the γ s and e^\pm s that satisfy:

- $P_{t\gamma} > P_{t\gamma\text{min}}$
- $|\eta_\gamma| < \eta_{\gamma\text{max}}$
- $P_{te} > P_{te\text{min}}$
- $|\eta_e| < \eta_{e\text{max}}$
- $(\Delta R < \Delta R_{\gamma\text{-seed}}, \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2})$ or $(|\Delta\eta| < |\Delta\eta|_{\gamma\text{-seed}}$ and $|\Delta\phi| < |\Delta\phi|_{\gamma\text{-seed}})$

We ignore the energy that may be deposited in the electromagnetic calorimeter by charged hadrons.

In addition we sum up the γ s and e^\pm s within a narrow cone around the seed direction in order to estimate the lowest possible energy that could be reconstructed for the candidate. These should satisfy:

- $P_{t\gamma} > P_{t\gamma\text{min}}$
- $|\eta_\gamma| < \eta_{\gamma\text{max}}$
- $P_{te} > P_{te\text{min}}$
- $|\eta_e| < \eta_{e\text{max}}$
- $\Delta R < \Delta R_{\text{narrow}}$

We will refer to this energy as “narrow cone energy”. In order to evaluate the effect of tracker isolation, stable or long lived charged particles around the direction of the photon candidates are counted using the following cuts:

- $P_{t\text{tk}} > P_{t\text{tkmin}}$
- $\Delta R < \Delta R_{t\text{kmax}}$
- $|\eta\text{tk}| < \eta_{t\text{kmax}}$

This set of cuts defines a “candidate photon”. Finally we require that at least a pair of candidate photons in the event satisfy:

- $P_{t1} > P_{t1\text{min}}$
- $P_{t2} > P_{t2\text{min}}$
- $N_{t\text{kcone1}}, N_{t\text{kcone2}} \leq N_{t\text{kconemax}}$
- $N_{t\text{kcone1}} + N_{t\text{kcone2}} \leq N_{t\text{ksummax}}$
- $|\eta_1|, |\eta_2| < \eta_{\text{candmax}}$
- Invariant mass of the pair $> M_{\gamma\gamma\text{min}}$
- Invariant mass of the pair (using the narrow cone energies) $< M_{\gamma\gamma\text{max}}$

Table 2: Cuts and selected cross sections of the selections for the $pp \rightarrow \text{jets}$ and $pp \rightarrow \gamma + \text{jet}$ processes. The last line of the table, corresponding to the selection of $pp \rightarrow \gamma + \text{jet}$ is added for completeness. The generator level selection actually used in this analysis is that described in the footnote 2.

	Selection A	Selection B	Selection C	Selection B'	Selection C'	Selection D
$p_{t\text{seedmin}}$ (GeV)	5	5	5	5	5	5
η_{seedmax}	2.6	2.6	2.6	2.6	2.6	2.6
$p_{t\gamma\text{min}}$ (GeV)	0.0	0.0	0.0	0.0	0.0	0.0
$\eta_{\gamma\text{max}}$	2.8	2.8	2.8	2.8	2.8	2.8
$p_{t\text{emin}}$ (GeV)	2.0	2.0	2.0	2.0	2.0	2.0
η_{emax}	2.8	2.8	2.8	2.8	2.8	2.8
Use K_{long}^0	Yes	No	No	No	No	No
$p_{tK_{\text{long}}^0\text{min}}$ (GeV)	1.0	–	–	–	–	–
$\eta_{K_{\text{long}}^0\text{max}}$	2.8	–	–	–	–	–
$ \Delta\phi _{\gamma\text{-seed}}$	0.4	0.4	0.4	0.3	0.4	0.3
$ \Delta\eta _{\gamma\text{-seed}}$	0.15	0.15	0.15	0.12	0.15	0.12
$\Delta R_{\gamma\text{-seed}}$	0.2	0.2	0.2	0.0	0.2	0.0
ΔR_{narrow}	0.02	0.02	0.02	0.02	0.02	0.02
$p_{t\text{tkmin}}$ (GeV)	1.6	1.6	1.6	1.6	1.6	1.6
η_{tkmax}	2.2	2.2	2.2	2.2	2.2	2.2
ΔR_{tkmax}	0.2	0.2	0.2	0.2	0.2	0.2
$p_{t1\text{min}}$ (GeV)	37.5	37.5	37.5	37.5	37.5	37.5
$p_{t2\text{min}}$ (GeV)	22.5	22.5	22.5	22.5	22.5	22.5
η_{candmax}	2.6	2.6	2.6	2.6	2.6	2.6
$N_{\text{tkconemax}}$	3	2	1	2	1	0
N_{tksummax}	6	3	1	3	1	0
$M_{\gamma\gamma\text{min}}$ (GeV)	70	90	100	80	90	95
$M_{\gamma\gamma\text{max}}$ (GeV)	250	250	200	14000	14000	14000
$pp \rightarrow \text{jets}$ σ selected (pb)	1.0×10^5	1.4×10^4	4.6×10^3	2.0×10^4	4.4×10^3	1.5×10^3
$pp \rightarrow \gamma + \text{jet}$ σ selected (pb)	1.5×10^3	4.5×10^2	2.0×10^2	6.2×10^2	2.5×10^2	1.2×10^2

Various cuts were studied and we ultimately picked the selection labeled as C' in Table 2 along with a generator level cut on \hat{p}_{\perp} of 50 GeV. This led to a rejection factor of 6000 before simulation with an inefficiency of selection of 20% compared to a loose pre-selection for analysis. In fact, events rejected by this selection have rather low E_t photons and we assume that the effect of the lost events is negligible in the final optimized analysis.

All Monte Carlo samples used in the analysis are summarized in Table 3. All events, for signal and background, were generated with PYTHIA [9], simulated with the Geant [10] based CMSIM [11] or OSCAR [12], and reconstructed with ORCA version 8.7.3 [13]. Pile-up events from minimum bias interactions were added to the hard interaction, assuming a luminosity of $2 \times 10^{33} \text{ cm}^{-2}\text{sec}^{-1}$ and a beam crossing rate of 40 MHz.

In addition we apply K-factors to take into account the expected differences between the lowest order cross sections given by PYTHIA and the NLO cross sections of the different background processes [14, 15, 16, 17, 18]. The K-factors used for each background are summarized in Table 4 and are estimated to have an uncertainty of 20-30%.

Table 3: Monte Carlo samples used in the $H \rightarrow \gamma\gamma$ analysis with LO cross section from PYTHIA and total corresponding integrated luminosities of the analyzed samples.

Process	\hat{p}_\perp (GeV)	Higgs Mass (GeV)	Cross Section (pb)	Pre-selected Cross section (pb)	Events Analyzed	Int Lum. (fb $^{-1}$)
$H \rightarrow \gamma\gamma$ (gg fusion)	-	120	-	-	181K	-
$H \rightarrow \gamma\gamma$ (WB fusion)	-	120	-	-	193K	-
$H \rightarrow \gamma\gamma$ (gg fusion)	-	115–150	-	-	20K	-
$H \rightarrow \gamma\gamma$ (WB fusion)	-	115–150	-	-	20K	-
$H \rightarrow \gamma\gamma$ (WH,ZH,ttH)	-	115–150	-	-	20K	-
$pp \rightarrow \gamma\gamma$ (born)	> 25	-	82	44	920K	30
$pp \rightarrow \gamma\gamma$ (box)	> 25	-	82	31	668K	20
$pp \rightarrow \gamma + \text{jet}$	> 30	-	5×10^4	2.5×10^3	5.5M	2.2
$pp \rightarrow \text{jets}$	> 50	-	2.8×10^7	4.7×10^3	4.5M	1.0
Drell Yan ee	-	-	4×10^3	4×10^3	460K	0.1

Table 4: Background K-factors applied to PYTHIA cross sections.

$pp \rightarrow \gamma\gamma$ (born)	1.5
$pp \rightarrow \gamma\gamma$ (box)	1.2
$pp \rightarrow \gamma + \text{jet}$ (2 prompt)	1.72
$pp \rightarrow \gamma + \text{jet}$ (1 prompt+ 1 fake)	1
$pp \rightarrow \text{jets}$	1

4 Outline of the Analysis

Two analysis are presented: a cut-based analysis and a highly optimized analysis in which the expected signal to background ratio is calculated for each event. The analysis performance is determined by the efficiency and the s/b ratio, defined as the expected signal divided by the expected background in each bin of the final distribution that is used to compute the confidence level for discovery or exclusion. For the cut-based analysis this is the invariant mass distribution while for the optimized analysis the final distribution is obtained with the procedure that will be outlined in the following.

By dividing the cut-based analysis in various categories with different levels of sensitivities (s/b ratio) results improve toward those that are obtained with the optimized analysis, that is equivalent to a continuous splitting into categories. If the maximum s/b ratio in the optimized analysis is limited to that achieved in the cut-based analysis, the performances of the two analyses are nearly identical.

The optimized, discovery-oriented analysis is very interesting in the $H \rightarrow \gamma\gamma$ channel because it is useful to mitigate the effect of the large background and because the Higgs signal appears in a narrow mass peak allowing us to train the analysis on the background in the mass side-bands. Thus, the analysis will not be at all limited by the poor simulation of the background once data will be available.

5 Common Analysis Tools

In this section we describe a series of analysis tools that are common to the two analyses:

5.1 Trigger

$H \rightarrow \gamma\gamma$ events are selected with extremely high efficiency both by the Level-1 and High Level triggers that are described in details in reference [19]. Since in the analysis selection we apply tighter E_t and isolation cuts, the inefficiency due to the trigger is negligible.

5.2 Photon Reconstruction

Photons are reconstructed with the standard ECAL software⁴⁾ using the hybrid clustering algorithm in the barrel and the island algorithm in the end cap [20]. At this level the photon reconstruction efficiency is over 99.5% for photons in the region covered by the ECAL.

New energy corrections for photons were developed and improve the energy resolution compared to previous corrections developed for electrons. These energy corrections depend on the pseudo-rapidity and shower shape of the super-cluster. In particular corrections are applied as function of the shower shape variable R_9 , that will be defined in subsection 5.4, of the impact parameter within the crystal, estimated by the ratio of the energy in the highest energy crystal in a super-cluster and the energy contained in the 3×3 array centred around it, of the vicinity to super-module borders and of the pseudo-rapidity of the photon. We try to recover in this way the energy lost due to particles not reaching the electromagnetic calorimeter because of photon conversion and curvature in the magnetic field or due to the leakage in crystal and inter-module borders.

The energy resolution of reconstructed photons is excellent for photons that do not convert or that convert late in the tracker. Energy resolution deteriorates somewhat for photons that convert early in the tracker. Nevertheless, the photon energy resolution is substantially less affected by tracker material than is the electron energy resolution and the Higgs reconstruction in the calorimeter is quite reliable even for converted photons.

For signal events the energy response of the individual crystals of the ECAL have been smeared by the expected precision that will be obtained after the calibration with $W \rightarrow e\nu$ events based on an integrated luminosity of 10 fb^{-1} , as described in [21]. The precision is 0.3% in the central part on the barrel, growing up to 1.0% at the edge of the barrel and in the endcaps.

For the moment we did not use the tools that have been developed to identify and reconstruct photon conversions in the tracker [22] and we also did not use the silicon preshower detector located in front of the ECAL endcaps to perform π^0 identification. In future the inclusion of these tools will improve the performance of the analysis.

5.3 Primary Vertex Identification

The bunch length at LHC has a standard deviation of ~ 75 mm that results in a longitudinal spread of interaction vertices of ~ 53 mm. If the mean longitudinal position is used (nominal vertex), the invariant mass of a two-photon state, such as the $H \rightarrow \gamma\gamma$, is smeared by about 1.5 GeV, due to the mis-measurement of the angle between the two photons related to the uncertainty of the photon directions.

The two high E_t photons coming from the Higgs boson decay are produced in association with other tracks that may come from the underlying event and initial state gluon radiation or from the other particles produced with the Higgs boson in the case of WBF fusion, WH or ZH production and $t\bar{t}H$ production.

The charged tracks associated to the Higgs production vertex are typically harder than those coming from minimum bias interactions. Therefore the vertex can be identified by reconstructing the primary vertices in the event and selecting the one that most likely corresponds to the Higgs boson production, based on charged tracks.

The variable that we use to identify the primary vertex are:

- $P_{\text{tvtx}}^{\text{mod}}$: $\sum |P_t|$, sum of the moduli of the transverse p_t of all tracks with longitudinal impact points within 5 mm;

⁴⁾ The clustering developed to efficiently measure electrons works very well on isolated photons.

- $P_{\text{vtvx}}^{\text{vec}} : |\sum P_t|$, modulus of the sum of the transverse p_t vectors of all tracks with longitudinal impact points within 5 mm from each other;

In both cases all tracks with a p_t larger than 1 GeV and with the absolute value of the pseudo-rapidity less than 2.4 are used.

We then choose the vertex with at least 2 tracks associated and with the largest $P_{\text{vtvx}}^{\text{mod}}$. In case the vertex with largest $P_{\text{vtvx}}^{\text{mod}}$ does not have also the highest $P_{\text{vtvx}}^{\text{vec}}$ we select the first if $P_{\text{vtvx}1}^{\text{mod}}/P_{\text{vtvx}2}^{\text{mod}} > P_{\text{vtvx}1}^{\text{vec}}/P_{\text{vtvx}2}^{\text{vec}}$. If this relation is not satisfied we select the second one.

At low luminosity ($2 \times 10^{33} \text{ cm}^{-2}\text{sec}^{-1}$) we are able to identify the right vertex within 5 mm from the real one in about 81% of the cases. Clearly these results may be affected by possible deviations of the structure of the underlying events from our current predictions. Table 5 shows the efficiency of the vertex determination, computed after the selection described in section 6.1, for the Higgs signal and for the different BG samples.

Table 5: Vertex reconstruction efficiency for the different signal and background processes.

Process	Efficiency (%)
$H \rightarrow \gamma\gamma$ gluon-gluon fusion	80
$H \rightarrow \gamma\gamma$ WBF fusion	89
Gluon fusion	72
Quark annihilation	71
$\gamma + \text{jet}$ (eg03_gamjet) (2 prompt γ)	76
$\gamma + \text{jet}$ (eg03_gamjet) (1 prompt $\gamma + 1$ fake)	85
QCD hadronic jets	90

As one can expect the best results are obtained for events without real hard photons in the final state and for those with pairs of photons with large total transverse momentum that consequently have high p_t recoiling objects.

5.4 Photon Isolation

Detailed studies of photon isolation optimization have been previously carried out [23, 24]. Fake photon signals due to jets can usually be rejected by looking for additional energetic particles in a cone around the photon candidate. Charged pions and kaons can be detected in the tracker or in the calorimeter. Neutral pions and other particles decaying to photons can be detected in the ECAL. The hadron calorimeter may be important for detecting charged particles not efficiently reconstructed in the tracker, particularly at high η , or other particles like neutrons or K_{long}^0 .

The variables with the best discrimination in the tracker have been found to be:

1. the sum of track p_t for tracks with p_t above a given threshold in a cone or between two cones of size ΔR around photon candidate;
2. the angle, from the ECAL cluster, to the 1st nearest track above some p_t threshold;
3. the number of tracks above a given threshold, within a cone around the photon candidate;

In the three cases the separation between the track and photon direction is measured by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Different cone sizes have been considered and tracks have been used when their p_t exceeds a given threshold p_t^{min} .

To study isolation in the ECAL we used basic clusters and super-clusters [20] reconstructed with the Island algorithm even though the super-clusters in the barrel were reconstructed using the hybrid algorithm. Island basic clusters were found to be significantly better for this purpose. The variables with the most discriminating power in the ECAL have been found to be:

1. the sum of E_t from basic ECAL clusters, in a cone of radius ΔR around the candidate super-cluster, excluding all the energy in the super-cluster;
2. The sum of E_t from basic clusters between two cones around the photon candidate ($\Delta R_{\min} < \Delta R < \Delta R_{\max}$);
3. The simple shower shape variable R_9 which is defined as the fraction of the super-cluster energy found inside the 3 by 3 array of crystals centred around the highest energy crystal.

The only type of isolation variable that is used in the hadron calorimeter is the sum of E_t from towers in a cone of radius ΔR around the photon candidate.

In addition to these we use a neural network to define a discriminating variable NN_{isol} that combines the various isolation variables into a single result for each photon candidate. It provides a continuous variable that we can cut on to improve the signal to background ratio. We used the following isolation variables as inputs to the Neural Network:

- the angle to the first track above an E_t of 2 GeV;
- the ECAL transverse energy in a cone of radius $\Delta R < 0.3$ excluding the candidate super-cluster energy;
- the shower shape variable R_9 ;
- E_t sum in the HCAL in a cone of radius $\Delta R < 0.35$;
- the sum of E_t in the tracker in a narrow cone of $\Delta R < 0.2$.

At high photon efficiency, this output variable is significantly better than any single isolation variable, reaching an efficiency ratio of 50 at 90% photon efficiency. For the optimized analysis, we are interested in very large rejection factors as well. At 50% photon efficiency, the signal to background efficiency ratio is above 200. NN_{isol} is one of the key inputs to the optimized analysis enabling the estimation of s/b for each event.

We should also note that none of these neural net results uses any of the kinematic variables such as the super-cluster E_t . The Neural Net output for isolation can then be simply combined in an analysis with other background rejection techniques that make use of kinematic information.

5.5 Separation into Categories based on Lateral Shower Shape and Pseudo-rapidity

The shower shape variable R_9 has been shown to be very useful in discriminating between photons and jets. Because it looks in a small 3×3 crystal area inside the super-cluster it can provide information about narrow jets and even about π^0 s. Signal photons sometimes have low values of R_9 due to conversions, but there is no doubt that this provides additional isolation information about the super-cluster.

Photon conversion in the material of the tracker affects both the energy resolution for the photon and our ability to distinguish a photon from a jet. Photons with large value of R_9 either did not convert or converted late in the tracker and have good energy resolution. Photons that did convert early have lower values of R_9 and worse than optimum energy resolution.

In the following we will separate the events into categories with different R_9 so that we can take advantage of better mass resolution when we expect it, yet still use all of the events since the mass resolution varies by at most a factor of two. This separation also tends to put background events involving jets into categories with lower R_9 . Single π^0 s will also tend to give lower R_9 .

We also expect that photons detected in the endcaps will have worse energy resolution and higher background than photons detected in the barrel so that it is useful to separate events with one or more photons in the endcaps from those with both photons in the barrel.

5.6 Calculation of Confidence Levels

The confidence levels are computed by using the Log Likelihood Ratio frequentistic method, as described in [25], for a given Higgs mass hypothesis.. To evaluate how much the outcome of the experiment is signal-like we use as test-statistics the Log Likelihood Ratio (LLR) which compares the likelihood of the data to come from a background-only distribution to the likelihood to come from a signal-plus-background distribution. Given the expected signal and background distributions in the final discriminating variable (the mass for the cut-based analysis), the Likelihood \mathcal{L} of the outcome of the experiment is defined as:

$$\mathcal{L}(x) = \prod_{i=1}^n \frac{e^{-(xs_i+b_i)} (xs_i + b_i)^{n_i}}{n_i!}$$

where the index i runs over all the bins of the discriminating variable distribution, n_i is the number of observed events in bin i , the quantities s_i and b_i are the expected signal and background events in bin i and the argument x is zero for the background only hypothesis and 1 for the signal plus background hypothesis.

The Log Likelihood ratio $LLR = \log(\mathcal{L}(b)) / \log(\mathcal{L}(s + b))$ is:

$$\sum_{i=1}^n [s_i - n_i \log(1 + s_i/b_i)]$$

The LLR of the data is compared to the outcome of random experiments simulated by means of the expected signal and background distributions of the discriminating variable. A large number of “gedanken” experiments is carried out in the two hypotheses: background only (index “b”) and signal plus background, where a Higgs signal of mass m_H is added to the background, (index “s+b”).

If we define $P_b(x)$ and $P_{s+b}(x)$ as the normalized distributions of the two sets of experiments, the discovery confidence level $1 - CL_b$ and the exclusion confidence level CL_s are given by:

$$1 - CL_b = \int_{-\infty}^{LLR_{\text{obs}}} P_b(x) dx$$

$$CL_{s+b} = \int_{LLR_{\text{obs}}}^{\infty} P_b(x) dx$$

$$CL_s = CL_{s+b} / CL_b$$

where LLR_{obs} is the value of the estimator computed from the data. To evaluate the discovery and exclusion potential we use, in place of the data, the median estimators of the signal plus background and background only experiments.

A plot of the distribution of the estimator, corresponding to a median discovery significance of about 5 sigma, is shown in Figure 2. For a 5 sigma discovery, we want about 0.26 background-only trials in a million to have a more signal-like estimator than the median estimator for the $s + b$ trials. The plot also shows that significant fluctuations are possible in the data that could cause the luminosity needed for discovery to be either substantially larger or substantially smaller than that expected from the median.

5.7 Effect of Systematic Errors

When we want to include the systematic errors we simply randomize the background and signal expectation by the systematic error when generating random experiments while we keep their expectations at the nominal value for the computation of the log likelihood ratio estimator. If necessary we include the correlations among the errors on the different analysis bins. We observe that there is no effect of the signal systematic error on the median LLR of signal-plus-background experiments nor on that of background-only experiments. Of course the distribution

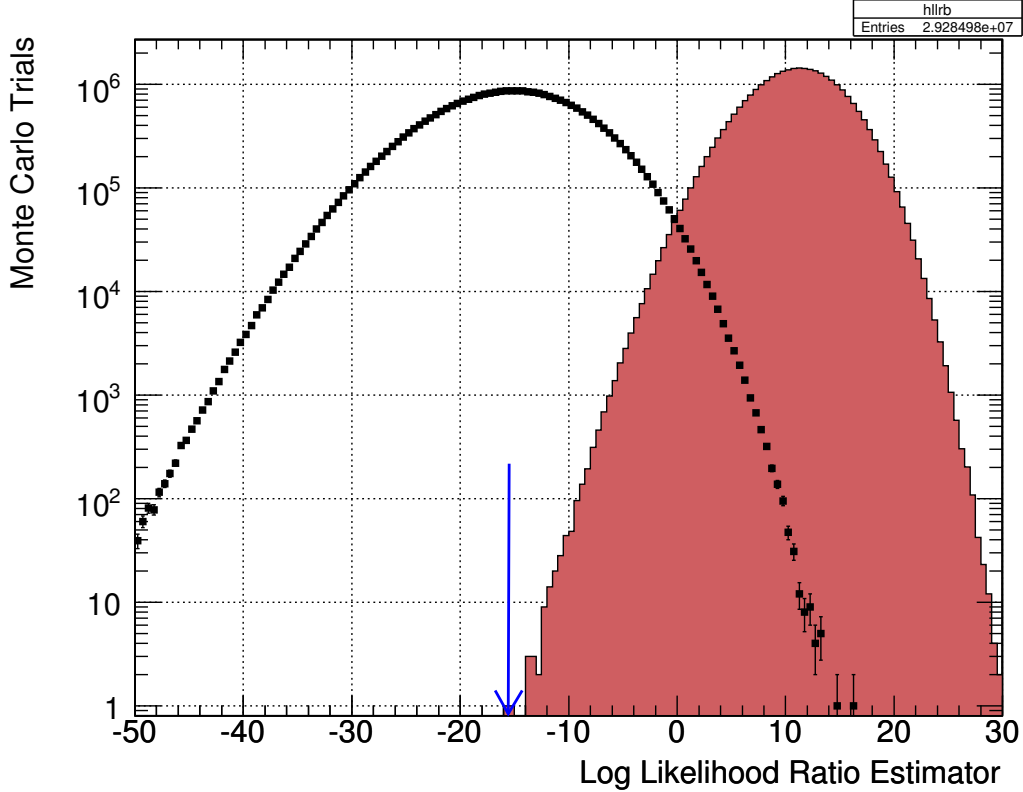


Figure 2: Distribution of background-only (solid histogram) and signal-plus-background (black squares) LLR estimators for millions of random trials used to compute the confidence level. The arrow indicates the median of the signal plus background distribution.

corresponding to the signal-plus-background experiments is enlarged by the systematic error on the signal and this makes the exclusion more difficult. On the other hand the effect of the systematic error on the background is very large, because of the very small signal over background ratio of at least part of the events selected for the confidence level calculation. The mean of the distributions is still unchanged but the widths are enlarged both for background-only experiments and for signal-plus-background experiments. The result is that both discovery and exclusion sensitivities are decreased.

6 Cut-based Analysis

In this analysis we simply search for a mass peak in the two-photon invariant mass distribution. We select events passing the double photon High Level Trigger. We then apply a tighter selection that requires two isolated high E_t photons.

6.1 Selection

We require that there are two photon candidates with pseudo-rapidity $|\eta| < 2.5$ and with transverse energies larger than 40 GeV and 35 GeV respectively and that satisfy the following isolation requirements:

- no tracks with p_t larger than 1.5 GeV/c must be present inside a cone with $\Delta R < 0.3$ around the photon candidate⁵⁾;

- the total E_t of all ECAL island basic clusters with $0.06 < \Delta R < 0.35$ around the direction of the photon candidate, regardless of whether they belong to the super-cluster or not must be less than 6 in the barrel and 3 GeV in the endcaps;
- the total transverse energies of HCAL towers within $\Delta R < 0.3$ around the photon candidate must be less than 6 GeV in the barrel and 5 GeV in the endcaps.

In order to further reduce the background that is higher when at least one of the photons is detected in the electromagnetic calorimeter endcaps and to increase the performance of the analysis in the forward region we apply additional isolation requirements when at least one of the candidates has $|\eta| > 1.4442$. For these events we require that also the candidate in the barrel satisfies the tighter isolation selection that is applied to photons in the endcaps: ECAL isolation < 3 GeV and HCAL isolation < 5 GeV.

Finally, in case that there are more than two candidates that satisfy the above requirements, the two with the highest transverse energy are used.

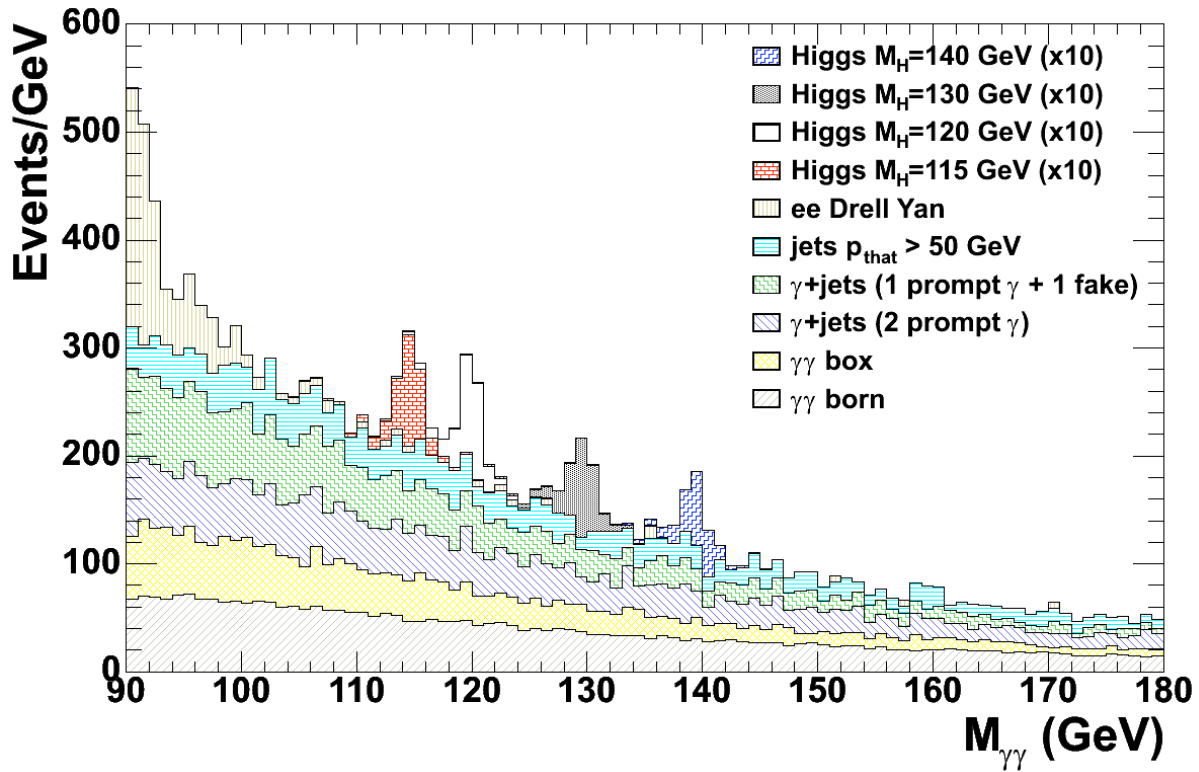


Figure 3: Di-photon invariant mass spectrum after the selection for the cut-based analysis. Events are normalized to an integrated luminosity of 1 fb^{-1} and the Higgs signal, shown for different masses, is scaled by a factor 10.

Figure 3 shows the mass distribution after the selection. The efficiency for a 120 GeV Higgs boson is 30.% and the total expected background is 178 fb/GeV. The number of expected background events for the different types of background is reported in Table 6 while the Higgs efficiency in different mass windows is reported in Table 7, the efficiency is computed using all generated signal events. As we can see from the figure, the signal contribution to the total number of events is very small, compared to the background, particularly outside the mass region under study. The background can therefore be estimated by a fit to the data mass distribution.

⁵⁾ We only consider tracks with hits in at least two layers of the silicon pixel detector, therefore converted photons are may be rejected only if they convert before the second pixel layer

Table 6: Expected background after the selection for Higgs boson masses between 115 and 150 GeV, expressed in fb/GeV

Process	$M_H=115$ GeV fb/GeV	$M_H=120$ GeV fb/GeV	$M_H=130$ GeV fb/GeV	$M_H=140$ GeV fb/GeV	$M_H=150$ GeV fb/GeV
$pp \rightarrow \gamma\gamma$ (born)	48	44	36	29	24
$pp \rightarrow \gamma\gamma$ (box)	36	31	23	16	12
$pp \rightarrow \gamma + \text{jet}$ (2 prompt)	43	40	32	26	22
$pp \rightarrow \gamma + \text{jet}$ (1 prompt and 1 fake)	40	34	22	19	14
$pp \rightarrow \text{jets}$	29	27	20	18	14
Drell-Yan ee	2	2	1	1	1
Total background	203	178	134	109	86

Table 7: Selection efficiency for the Higgs signal in different mass windows.

M_H (GeV)	Mass window ± 1 GeV	Mass window ± 1.5 GeV	Mass window ± 2.5 GeV	Mass window ± 5 GeV	Mass window Total
115	17%	21%	25%	28%	29%
120	18%	22%	26%	29%	30%
130	18%	22%	27%	31%	32%
140	18%	23%	28%	32%	34%
150	28%	24%	29%	33%	36%

The error on the background estimation comes from two sources:

- the statistical precision which decreases with the size of the mass range that is used to perform the fit;
- the systematic error related to the shape of the function that is used to fit the distribution.

It will always be impossible to know the exact functional form of the background shape and we must estimate this error by assuming a true function, simulating a distribution and then using a different function to fit the data. Clearly this error grows with the size of the mass range used. For a reasonable mass range of $\pm 10 - 20$ GeV excluding $+3$ and -5 GeV from the Higgs boson mass under study and for an integrated luminosity of 20 fb^{-1} we estimate the statistical and systematic errors to be respectively 0.4% and 0.5%. The statistical error decreases with the integrated luminosity while the systematic error is constant.

6.2 Splitting into Categories

By changing the cuts or adding new discriminating variables to this analysis we did not obtain large improvements in the sensitivity of the analysis. We can see this for example from the fact that we are not able use R_9 , that is a very powerful variable, to reject events without loosing in performance. This happens because the increase in s/b ratio does not compensate the loss in efficiency.

The only way we can improve the sensitivity of the analysis is to keep all selected events and split the sample into different categories with different s/b ratios.

We then consider three possibilities:

- one single category;

- four categories from 2 R_9^{min} ranges (R_9^{min} larger or smaller than 0.93) times 2 pseudo-rapidity regions $|\eta|^{max}$ in barrel or endcaps;
- twelve categories from 3 R_9^{min} ranges ($R_9^{min} > 0.948$, $0.9 < R_9^{min} < 0.948$ and $R_9^{min} < 0.9$) times 4 pseudo-rapidity regions ($|\eta|^{max} < 0.9$, $0.9 < |\eta|^{max} < 1.4442$, $1.4442 < |\eta|^{max} < 2.1$ and $|\eta|^{max} > 2.1$).

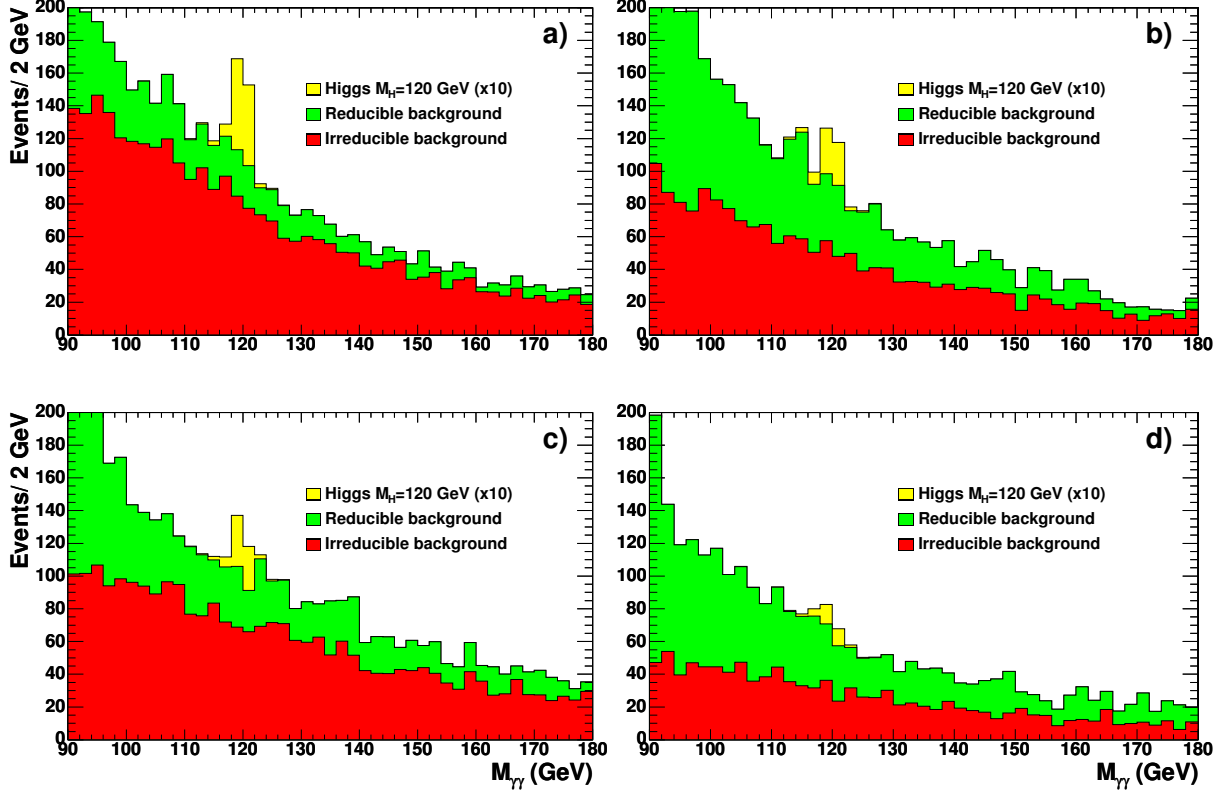


Figure 4: Invariant mass spectrum after the selection relative to the cut-based analysis with four categories defined in the text: barrel with large R_9 (a), barrel with small R_9 (b), endcaps with large R_9 (c) and endcaps with small R_9 (d). Events are normalized to an integrated luminosity of 1 fb^{-1} and the Higgs signal, shown for different masses, is scaled by a factor 10.

Figure 4 shows the mass spectrum after the selection in the four categories. We can observe how the signal over background ratio is much larger in the best category and also that the composition of the background varies between the different samples: irreducible backgrounds dominate for large R_9 and reducible backgrounds are larger for small R_9 .

Table 8 instead shows, for the 12 category analysis, the fraction of events along with the maximum s/b ratio in each category.

Table 9 shows the integrated luminosity needed to obtain 5σ discovery or 95% CL exclusion for a 120 GeV mass Higgs boson with the different analyses. The effect of the systematic errors is also shown. We want to stress again that in the three cases (one, four and twelve categories) the event selection is the same and that the differences in performance come from the splitting of the total sample in different sub-samples with different sensitivities (s/b).

Figure 5 shows the integrated luminosity needed to obtain 5σ discovery or 95% CL exclusion and the discovery sensitivity with an integrated luminosity of 30 fb^{-1} in the Higgs boson mass range between 115 and 150 GeV.

We limited ourselves to splitting by means of R_9 and pseudo-rapidity. Important further improvements could be

Table 8: Fractions of events in each of the 12 categories and maximum s/b in the mass region of 120 GeV.

	$ \eta ^{\max} < 0.9$		$0.9 < \eta ^{\max} < 1.4442$		$1.4442 < \eta ^{\max} < 2.1$		$ \eta ^{\max} > 2.1$	
	frac.	s/b	frac.	s/b	frac.	s/b	frac.	s/b
$R_9^{\min} > 0.948$	15.5%	14.7%	13.1%	9.0%	10.8%	6.1%	8.5%	4.5%
$0.9 < R_9^{\min} < 0.948$	9.4%	12.2%	6.8%	6.8%	6.7%	4.8%	2.7%	2.8%
$R_9^{\min} < 0.9$	8.3%	7.6%	11.1%	4.3%	5.4%	3.2%	1.7%	2.2%

Table 9: Integrated luminosity needed to discover or exclude the Higgs boson with mass 120 GeV with or without taking into account the systematic errors (fb^{-1}).

Analysis	5σ	5σ	3σ	3σ	95%	95%
	discovery no syst	discovery syst	evidence no syst	evidence syst	exclusion no syst	exclusion syst
counting exp.	27.4	48.7	10.0	13.2	4.5	6.5
1 category	24.5	39.5	8.9	11.5	4.1	5.8
4 categories	21.3	26.0	7.5	9.1	3.5	4.8
12 categories	19.3	22.8	7.0	8.1	3.2	4.4

obtained by splitting into categories using the transverse energies of the photons or the transverse momentum of the Higgs boson candidate. For the moment we chose not to use these variables in order to be complementary with the optimized analysis that will be discussed in the next section.

6.3 Systematic errors

We have seen that when we do not split into categories and we use the full mass distribution, the total error on the signal is approximately 0.65% and is due to the uncertainty of the function fit to the side-bands of the mass distribution, estimated to be 0.5%, plus the statistical error on the fit that is approximately 0.4% for an integrated luminosity of 20 fb^{-1} .

An error of 0.65% has a very large effect on the discovery CL when we use only one category. The reason is that a large fraction of signal events corresponds to a very low s/b (of the order of a percent). The effect can be reduced by applying a cut on the signal over background s/b . This corresponds to using events in a mass window around the analyzed mass, until s/b becomes smaller than the chosen cut. The optimal cut for this analysis is 0.02. This cut is also applied when systematic errors are not included.

When we split into categories the number of background events in each category is reduced on average by $1/N_{\text{cat}}$ on average and this increases the statistical error on the background estimation by approximately a factor $\sqrt{N_{\text{cat}}}$ and this error is completely uncorrelated between the different categories. The error related to the uncertainty of the fit function remains constant and it is also uncorrelated between the different categories because, due to the different cuts the true background shapes are different and are described by completely different functions. The total error is then less than the total error reduced by $1/N_{\text{cat}}$. This reduces the effect of the systematic error on the discovery.

The total systematic errors on the background are estimated to be at 20 fb^{-1} :

- 0.65% 1 analysis
- 0.95% 4 analysis
- 1.5 % 12 analysis

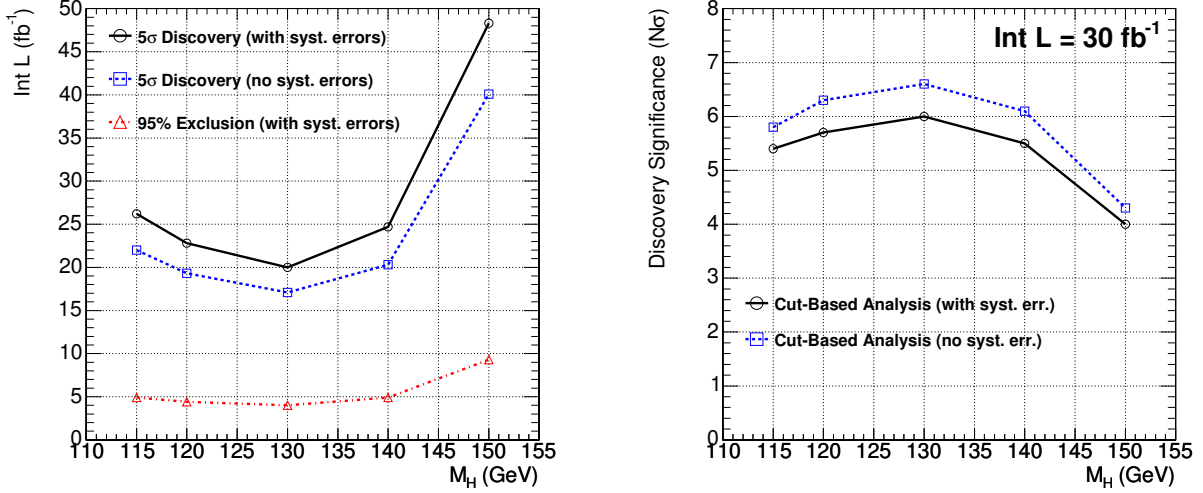


Figure 5: Integrated luminosity needed for a 5σ discovery or 95% CL exclusion (left) and discovery sensitivity with an integrated luminosity of 30 fb^{-1} (right) with the cut-based analysis with 12 categories.

and their statistical part decreases with the square root of the integrated luminosity.

The effect of the systematic error on the background estimation is also related to the s/b of the analysis. A more sensitive analysis, for which a larger part of the signal has a higher s/b ratio, is less affected by the same relative uncertainty on the background.

Clearly our current understanding of the background is affected by larger uncertainties such as: cross section, diphoton kinematic distributions and efficiency of the selection (mainly affected by jet fragmentation, pile-up and by the structure of the underlying events). All of this results in a large uncertainty on the background rate that on the other hand will be known very precisely when we will start taking data.

The systematic error on the signal, that as has been mentioned has no effect on the discovery CL, is due to the theoretical uncertainty of the cross section (+15 -12% from the scale variation and +4 -5%), to the measurement of the integrated luminosity ($\sim 5\%$), to the trigger ($\sim 1\%$), to the analysis selection (that will be measured for example with $Z \rightarrow \mu\mu\gamma$) and to the uncertainties on the photon energy resolution. Other effects that could modify our ability to discover the Higgs boson are: the uncertainties on the structure of the underlying events, that could change the efficiency of the primary vertex determination and the amount of material in the tracker before the electromagnetic calorimeter.

We have considered the effect on the performances of the analysis of an increase of 20% of the tracker material. The main effects on such change on the analysis would be:

- increase of the inefficiency of the track isolation requirements for early photon conversions, before or inside the second layer of the pixel detector.
- increase of the inefficiency of ECAL isolation cut;
- decrease of the value of R_9 for all photons that would cause a migration of events from more sensitive categories to less sensitive categories.

We estimated that such change would increase the luminosity needed to achieve a given discovery CL of approximately 6%. Given that the amount of tracker material will be known with a precision of $\sim 2\%$ [26] the related systematic error is less than 1%.

In the following we will use a conservative 20% systematic error on the signal. Its primary effect is on the exclusion limit and not discovery potential, since the signal rate is directly measured from data in case of discovery.

6.4 Results of the Cut-Based Analysis

The results are shown in Table 9 for a 120 GeV mass Higgs boson with the different analyses. We can observe how the performance increases and the effect of the error on the background estimation decreases with the number of categories.

Figure 5 shows the integrated luminosity needed for discovery and exclusion for all the mass range studied between 115 and 150 GeV when we use the 12-category analysis. We can discover the Higgs boson with mass between 115 and 140 GeV with less than 30 fb^{-1} and exclude it at 95% CL with less than 5 fb^{-1} in the same mass range.

The final thing we want to remind is that all the above results have been obtained assuming a calibration of the ECAL with the precision given in [21], that is 0.3% in the central part on the barrel, growing up to 1.0% at the edge of the barrel and in the endcaps, after having collected an integrated luminosity of 10 fb^{-1} . If instead we could calibrate the whole ECAL with a precision better than 0.5% over all the solid angle, the results would improve in such a way that approximately 10% less integrated luminosity would be needed for discovery.

7 Optimized Analysis Estimating s/b for Each Event

In the optimized analysis six categories are used, three in which both photons are in the barrel and three in which at least one photon is in an endcap. The three categories are defined, as for the cut-based analysis, to have the lowest R_9 photon candidate with $R_9 > 0.948$, $0.948 > R_9 > 0.90$ and $R_9 < 0.90$ respectively. They are numbered from 0 to 5: first the 3 barrel categories with decreasing values of R_9 then the 3 endcap categories again with decreasing values of R_9 .

7.1 Mass Distributions in Categories

The di-photon mass distributions enable the separation of signal from background. Signal peaks sharply at the Higgs mass while the backgrounds are quite smooth. This allows us to estimate the amount of background under the peak well. Events near the Higgs mass hypothesis have a large expected signal to background ratio while events outside the peak have a much lower s/b expectation.

The mass plots in the six categories are shown in Figure 6 for signal and background. The best mass resolution and the best s/b ratio in the peak is found in category 0, with high R_9 in the barrel.

7.2 Loose Selection of Events for Optimized Analysis

First of all we apply isolation requirements to photon candidates prior to the computation of the neural network isolation variables NN_{isol} :

- the transverse E_t of the photon candidates must be larger than 40 GeV and the absolute value of their pseudo-rapidity less than 2.5;
- no tracks with p_t larger than 1.5 GeV/c must be present inside a cone with $\Delta R < 0.1$ around the photon candidate;
- the total E_t of all ECAL island basic clusters with $\Delta R < 0.3$ around the photon candidate, excluding those belonging to the super-cluster itself must be less than 5 GeV;
- the total transverse energies of HCAL towers within $\Delta R < 0.35$ around the photon candidate must be less than 35 GeV;

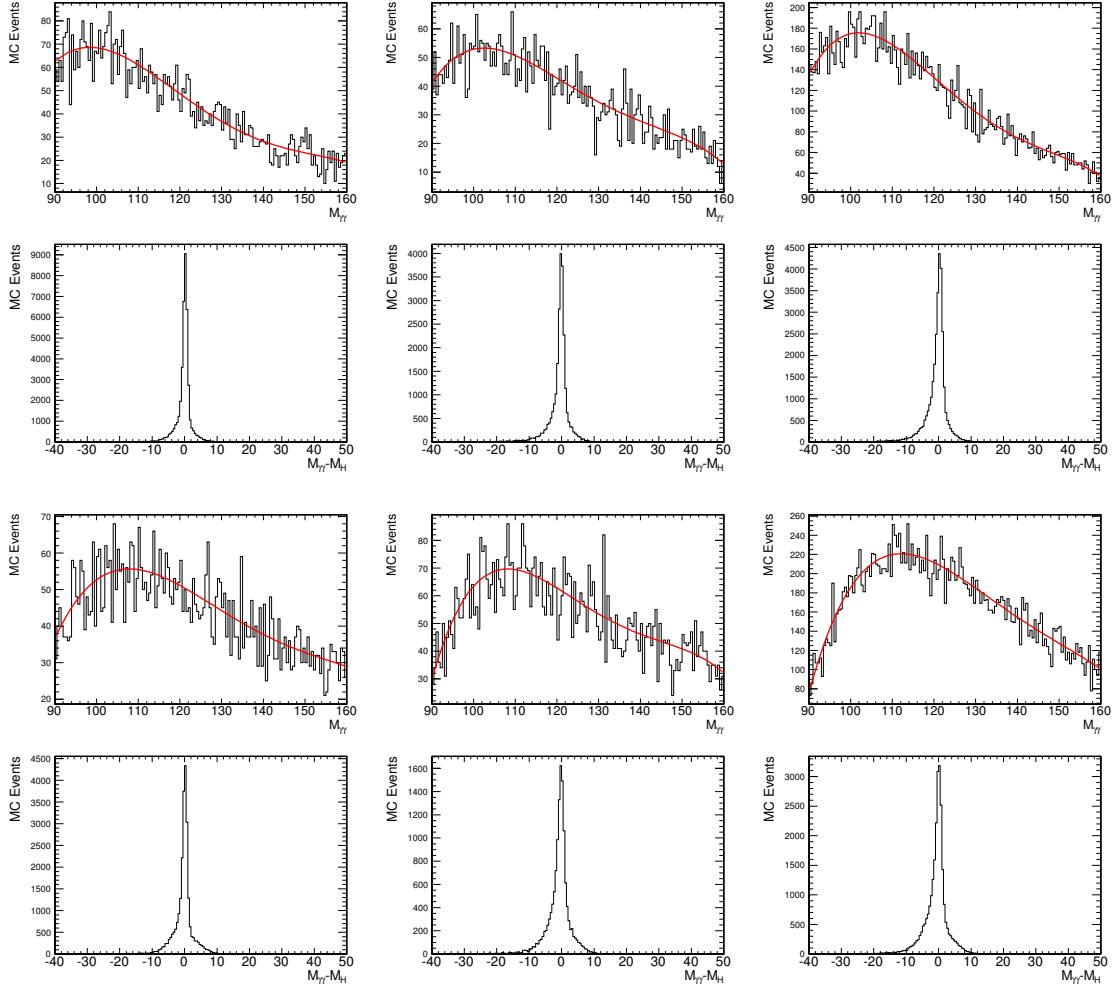


Figure 6: The distribution of di-photon mass for background and signal in the six categories with arbitrary normalization. The three barrel categories are shown on the top and the corresponding three endcap categories are shown on the bottom. The signal plots, shown for $M_H=120$ GeV, are as a function of mass difference from the true MC Higgs mass.

- the sum of the transverse momenta of charged tracks within $\Delta R < 0.2$ around the photon candidate must be less than 100 GeV.

Before optimizing the final analysis, some additional cuts are applied. These both simplify the neural network training and slightly improve the performance:

- we require that the events pass the double photon High Level Trigger;
- we require that the isolation neural net output be greater than 0.25 for both photons.

7.3 Optimized use of Kinematic Variables to Separate Signal and Background

In addition to the mass, there are kinematic differences between signal and background. In particular, the background photon rate decreases with E_t while the signal extends to higher photon E_t . The background can have a

high mass by having a large η difference. Weak Boson Fusion and associated production of a Higgs with other massive particles enhance these differences between signal and background. The large, reducible backgrounds often have photon candidates that are not well isolated.

As with the Higgs searches performed at LEP, higher performance can be achieved if we estimate the expected signal over background, s/b , for each event. This will be particularly effective if the s/b varies significantly from event to event. We find this to be the case for this analysis due to wide variations in photon isolation and photon E_t . There is also significant dependence on photon conversion and on location in the detector.

In this step of the analysis, we combine one photon isolation variable NN_{isol} for each photon, with kinematic variables to help separate signal and background. A neural net is trained to distinguish background events from the mass side-bands from signal Monte Carlo events. There is no danger of over-training since background events from the signal mass region are not used and we can easily use independent samples for the signal Monte Carlo. The input variables are devised to be insensitive to the di-photon mass so that background rejection due to the kinematics and isolation is independent of the background rejection from the mass distribution.

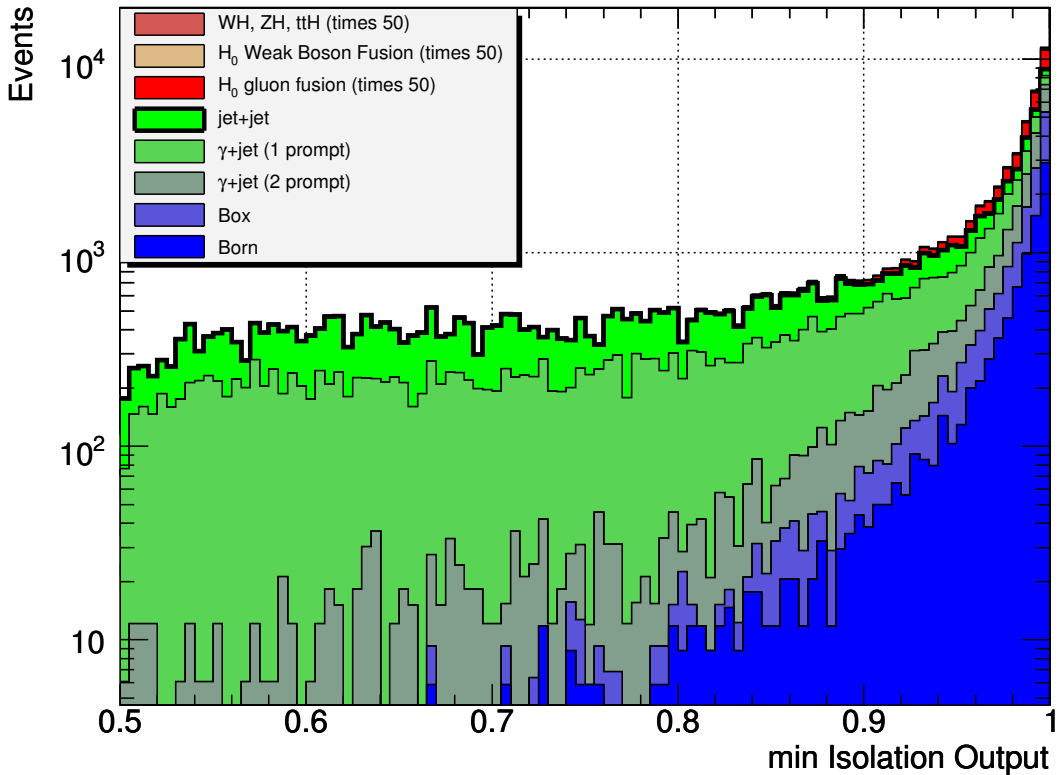


Figure 7: Distribution of the minimum value of the NN_{isol} variables of the two photon candidates. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal ($M_H=120 \text{ GeV}$) is scaled by a factor 50.

Six variables are used as inputs to a neural net. They are the isolation NN outputs NN_{isol} for the two photons, the transverse energies of the two photons, normalized to the di-photon mass, the absolute value of the rapidity difference between the two photons, and the longitudinal momentum of the photon pair:

1. $NN_{\text{isol}1}$
2. $NN_{\text{isol}2}$

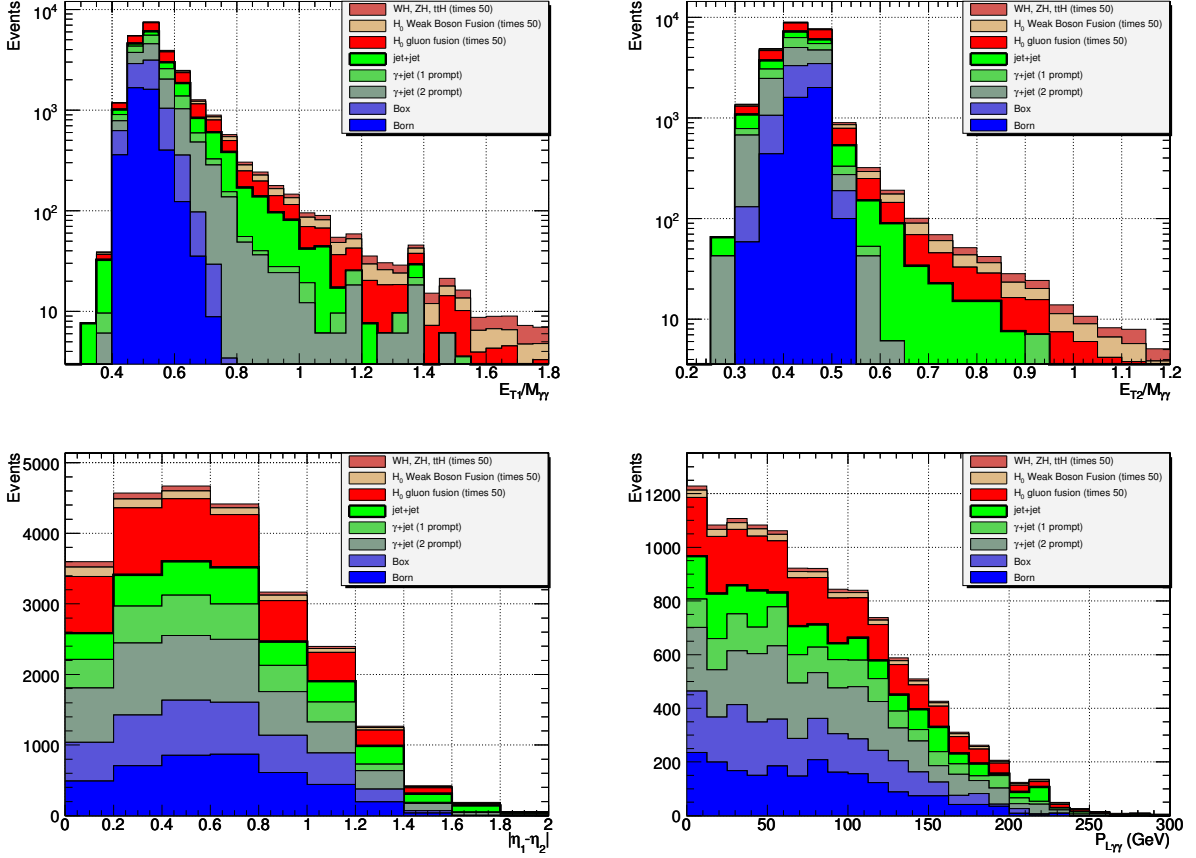


Figure 8: Distribution of the kinematic inputs to the neural network for signal and background sources. A value of the neural net output is required to be greater than 0.85. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal ($M_H=120 \text{ GeV}$) is scaled by a factor 50.

3. $E_{t1}/M_{\gamma\gamma}$
4. $E_{t2}/M_{\gamma\gamma}$
5. $|\eta_1 - \eta_2|$
6. P_L

The distributions of the input variables are shown for signal and background in Figures 7 and 8. We have not used variables or kinematic information that are likely to be highly sensitive to higher order corrections to the background simulation such as the E_t of the Higgs boson candidate or more particularly the E_t transverse to the photon direction. We also have not used information about additional jets. Certainly this type of information will ultimately be useful but may not be reliable until we have better simulations or actual data to train on.

The neural net is trained in each of the six categories independently. The net has six input nodes, twelve intermediate nodes in a single layer, and one output node. We have modified the error function from the standard to improve training toward a high signal over background region. Figure 9 shows the distributions of neural net output for signal and background in category 0. A minimum neural net output cut is applied that eliminates 1% of the signal in each category and a function is fit to the distribution above that cut. In the Figure the cut value is indicated by the end of the fit lines. These functions will be used to bin the data and to smooth the background in a limited region.

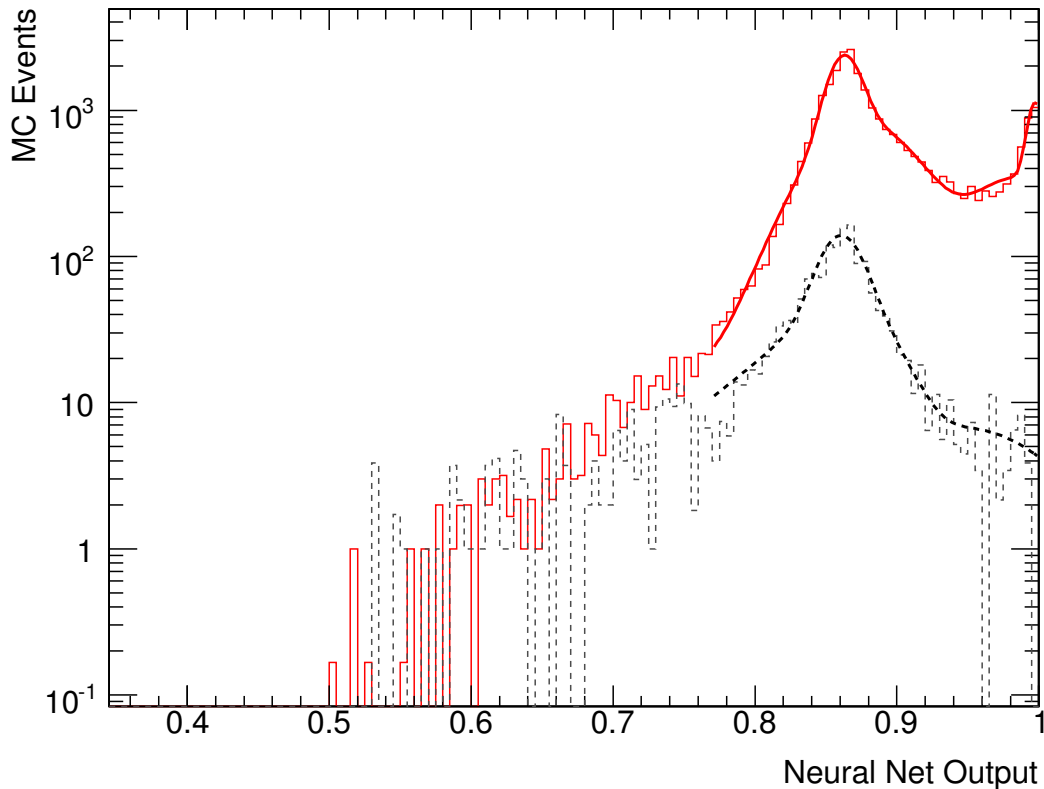


Figure 9: The distribution of kinematic neural net outputs for $M_H=120$ GeV Higgs signal (solid) and background (dashed) in category 0 with arbitrary normalization. The fit functions are also drawn.

It is useful to see the source of events in the neural net output distribution. Figure 10 plots the various signal and background sources versus the neural net output. Low NN outputs are dominated by photon candidates from jets which are not well isolated. The large peak represents both signal and background where the photon is relatively well isolated and the photon E_t is around $M_H/2$, corresponding to events with a large value of NN_{isol} . Higher photon E_t events are found in the peak near 1. There is an enhancement of the signal there for all types of signal but particularly for the WBF and associated production processes. The background there is dominated by events with at least one jet interpreted as a photon.

7.4 Estimation of Signal to Background Ratio for Each Event

In order to get the most information out of each event, we wish to estimate the signal over background for that event. We will count events at high s/b separately from events at low s/b . In the simplest analyses, cuts are applied to select only high s/b events and those are counted. such a simple analysis loses information because some of the events that are cut could contribute to the measurement and because some of the events that are accepted are not used optimally.

In our case, events in the mass peak for the Higgs mass hypothesis under consideration have high s/b expectation while events outside the peak have lower expected s/b . Similarly, events at high NN_{kin} output have higher s/b expectation. We have made the kinematics and isolation information in NN_{kin} independent of mass information

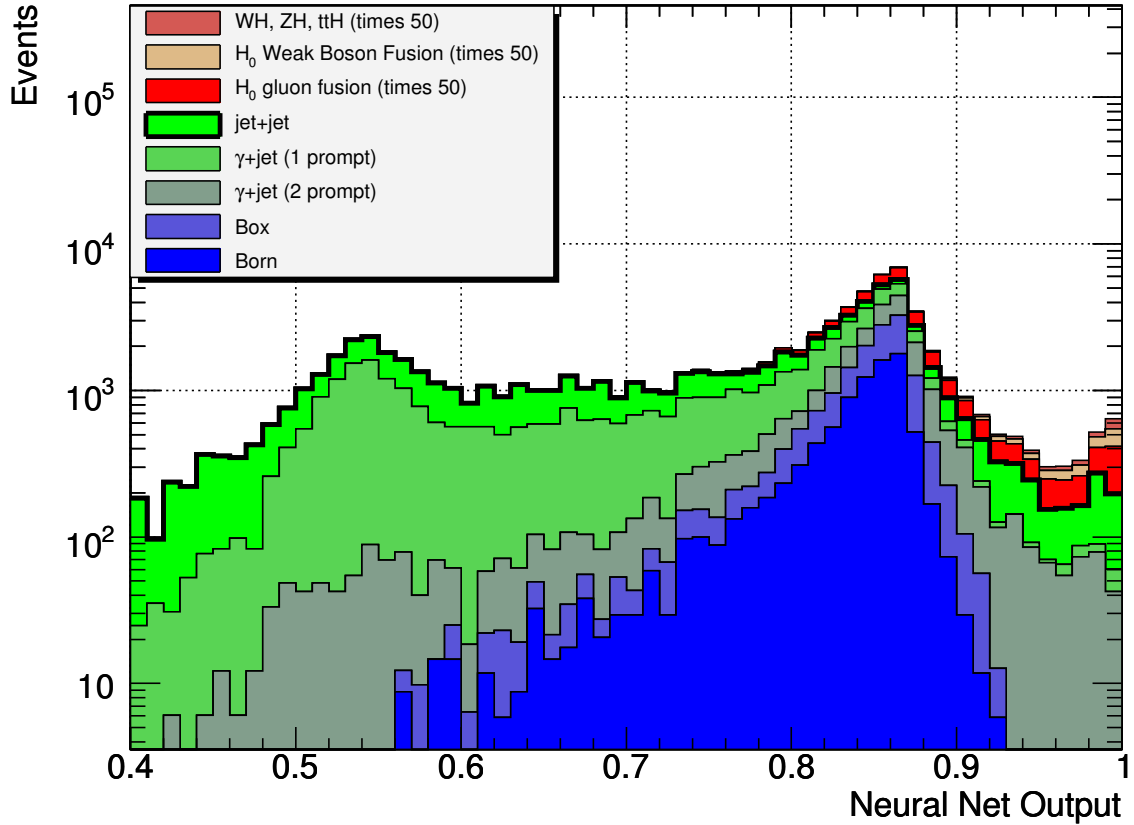


Figure 10: The neural net output for events in the barrel for each signal ($M_H=120$ GeV) and background source. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal is scaled by a factor 50.

so we can multiply the two s/b ratios together to get a good estimate⁶⁾ of the s/b expectation for the event:

$$\left(\frac{s}{b}\right)_{est.} = \left(\frac{s}{b}\right)_{mass} \times \left(\frac{s}{b}\right)_{kin}$$

We now bin signal and background in exactly the same way, according to the s/b estimate. Histograms are made in each of the six categories. We will use these histograms to calculate confidence levels that data are consistent with a background-only hypothesis or with a signal-plus-background hypothesis.

To help understand why the event by event use of an s/b estimate improves the performance of the analysis, Figure 11 shows the variation of the inverse, b/s , in category 1 by plotting the background events in a ± 8 GeV mass window versus the confidence level that the event is consistent with signal. This CL is just computed from the NN output such that the signal distribution in this plot is rigorously flat and has a value of 1. The 10% of the signal that is in the range $0 < CL_{sig} < 0.1$ has a very bad s/b while the 10% of the signal in the range $0.9 < CL_{sig} < 1$ has a good s/b which can improve by a factor of about 10 when mass information is also used.

⁶⁾ Note that this is only an estimate that is used in the same way to bin signal and background. If the estimate is bad, the performance of the analysis suffers because good s/b events are not well separated from bad ones. It is not possible for a bad estimate to make the analysis appear to perform too well. The s/b estimate need not be normalized correctly. It is only a relative number used to bin events from signal and background. The actual signal to background ratio will be computed from the binned events.

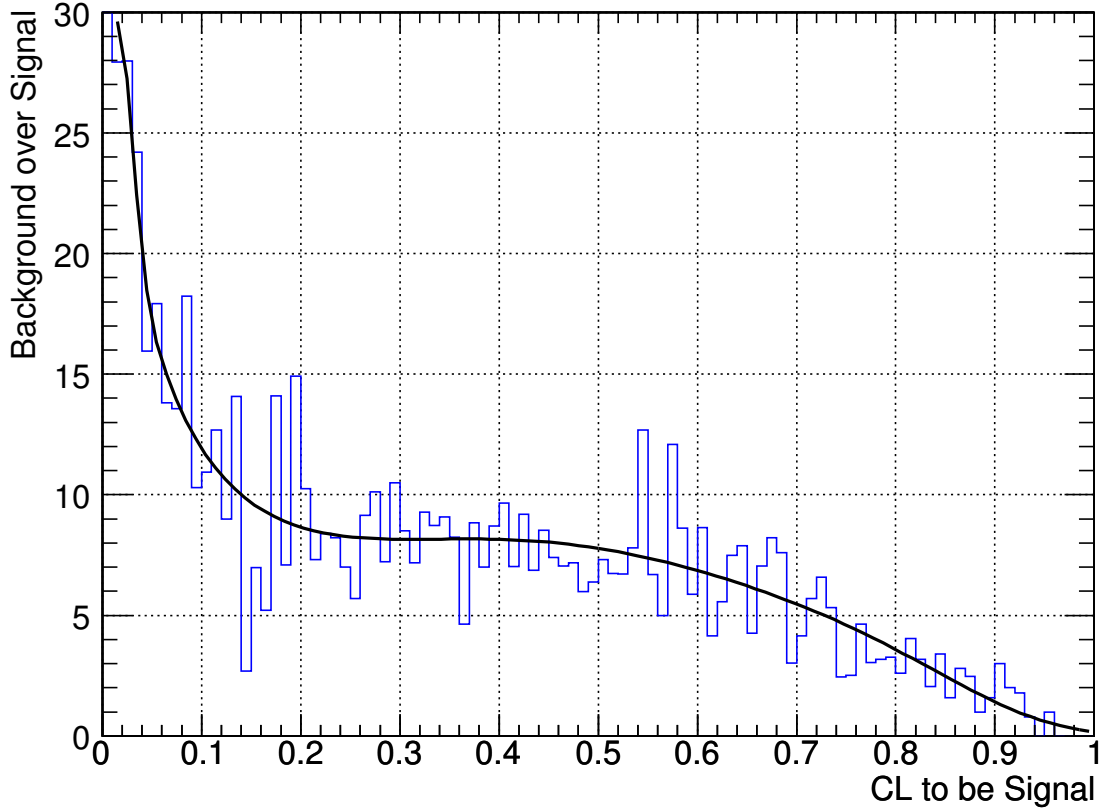


Figure 11: Background events plotted versus the confidence level, computed from the NN output for an event to be consistent with signal in category 0. The signal ($M_H=120$ GeV) is rigorously flat in this variable. The normalization is such that the y-axis gives b/s in a 16 GeV mass window.

We may also note in the figure that MC statistics for the background are low and pose a significant challenge for this analysis.

7.5 Smoothing the Background

The $H \rightarrow \gamma\gamma$ channel has the good feature that the mass is essentially independent of isolation and suitably chosen kinematic variables. Assuming that mass and NN output are uncorrelated, background can be smoothed well even in regions with fairly low statistics.

The background expectation in a bin must be reliably estimated in order to correctly calculate confidence levels. Downward fluctuations in the background estimation can have a significant impact on the CL. With the current simulation statistics, we have about one seventh of the statistics on important backgrounds that we will have in the data at the time we would discover the Higgs, therefore, problems with background estimation are even more difficult now than they will be when the experiment will be performed.

The background distributions are very regular in the mass variable. We therefore can reliably smooth the distribution in mass. We do this by spreading each event over a ± 5 GeV region according to the functions fit to the mass distribution. We could use a wider mass region but this would interfere⁷⁾ with the training of the analysis on an

⁷⁾ For now, we have erred on the side of independence of the training and analysis samples. This may not be optimal in the

independent sample in the mass side-bands.

We also smooth the background distribution in the neural net output over a region of ± 0.05 using the fit functions. It is therefore quite important that the background fit functions accurately represent the neural net distribution. In the smoothing process, we are careful to maintain the normalization of the background to high accuracy.

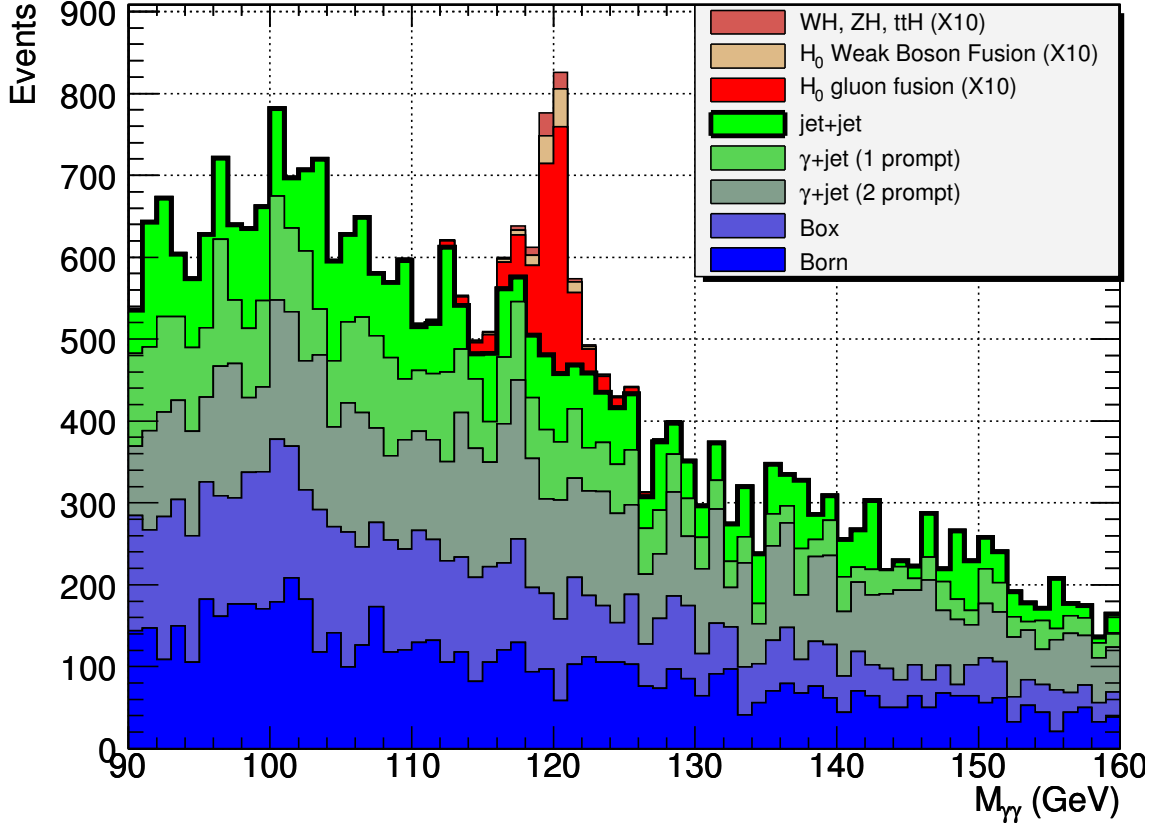


Figure 12: The di-photon mass distribution for each source for barrel events with kinematic neural net output greater than 0.85. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal ($M_H=120 \text{ GeV}$) is scaled by a factor 10.

With this two-dimensional smoothing, the background statistics are dramatically enhanced and accurate background expectations are obtained except in the regions with extremely small amounts of background. In such regions, we must combine bins until sufficient background statistics are available⁸⁾. This combination clearly reduces the sensitivity of the analysis but cannot be avoided without a more detailed understanding of the background. This will be a goal for this analysis in the future. At this time, we require at least 20 Monte Carlo background events in a bin to use the background estimate. Since our current MC statistics is about seven times lower than the statistics we expect to have in the data, significant performance improvements are expected to come from our ability of using higher s/b bins.

The mass distributions for various neural net cuts are shown to understand the signal we have and to see the importance of background smoothing in an analysis such as this.

end when a good estimation of background in very low background regions is important.

⁸⁾ If a s/b bin has too few MC background events contributing to it, we combine it with the nearest (lower s/b) bin. We continue to do this until there is sufficient statistics.

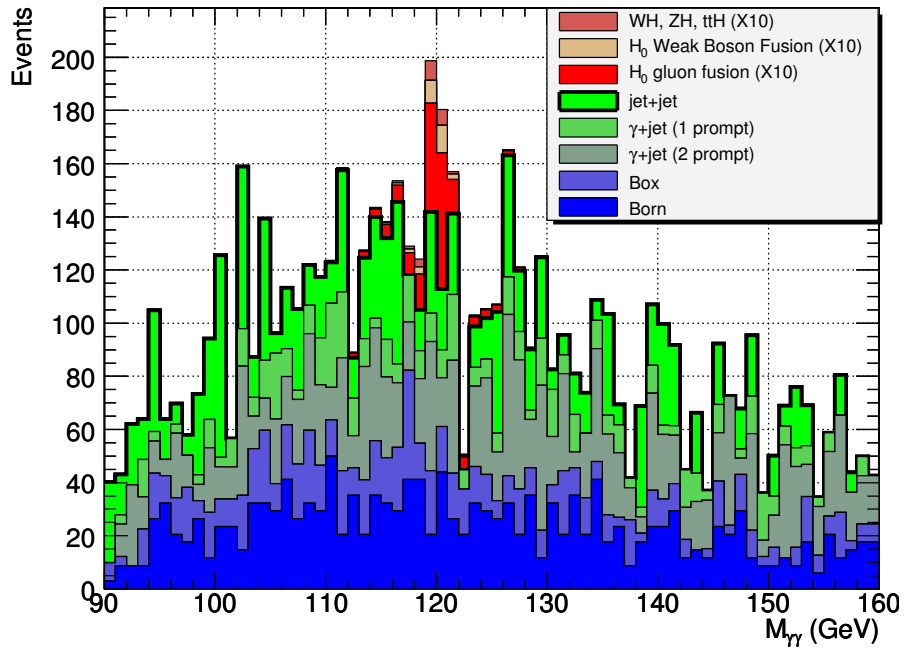
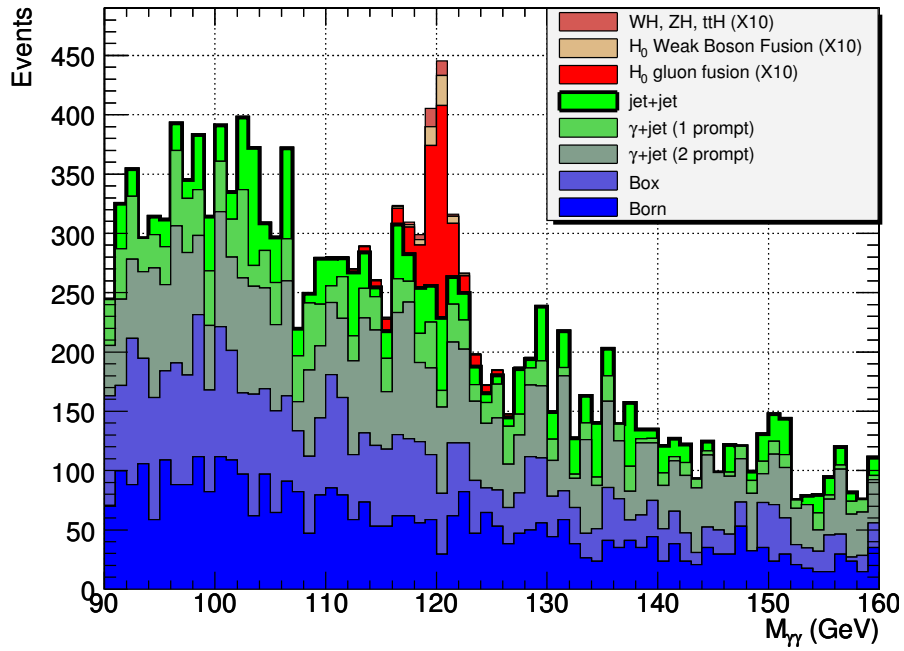


Figure 13: The di-photon mass distribution for events in the barrel (top) and for endcap events (bottom) with kinematic neural net output greater than 0.85 and $R_9 > 0.948$ for each signal ($M_H=120$ GeV) and background source. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal is scaled by a factor 10. Both the signal rate and the purity are seen to be lower for endcap events.

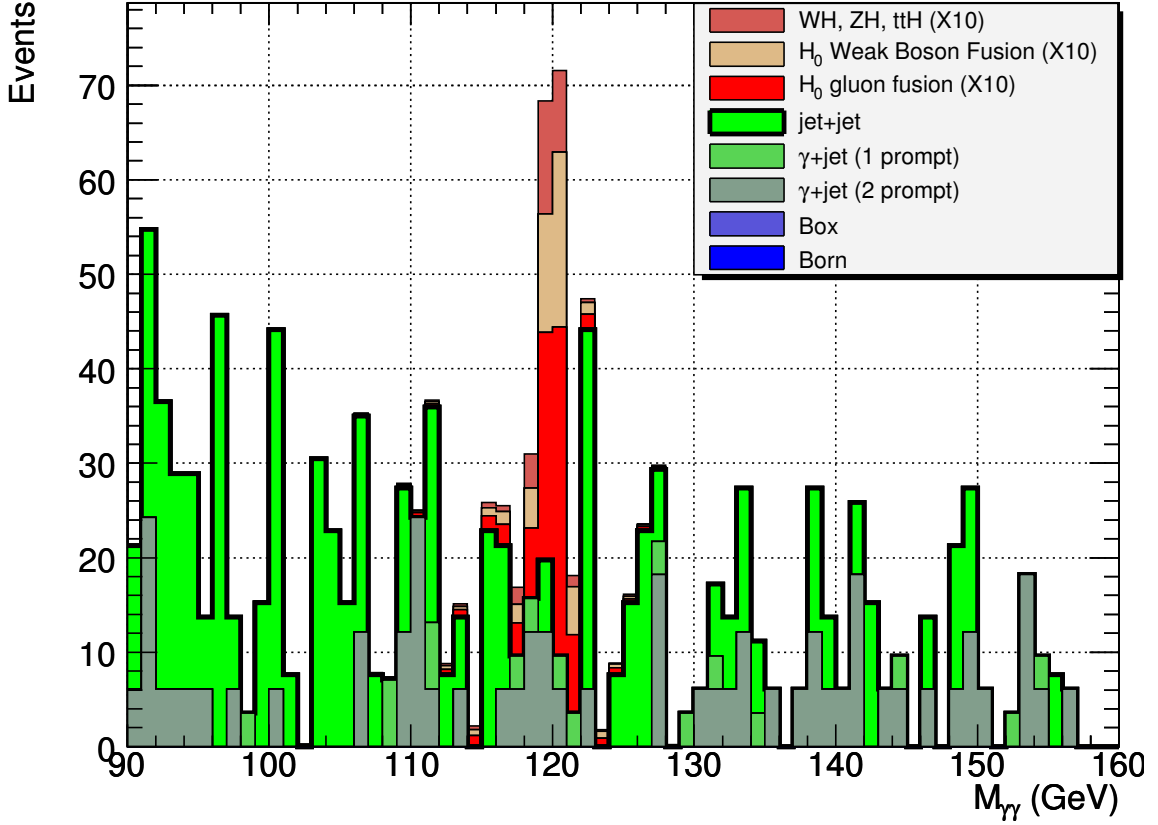


Figure 14: The di-photon mass distribution for each source for barrel events with kinematic neural net output greater than 0.97. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal ($M_H=120 \text{ GeV}$) is scaled by a factor 10.

Figure 12 shows the mass distribution for barrel events with a cut on the neural net output at 0.85. This cut simply excludes most of the obviously non-isolated candidates. One can see that all of the backgrounds are important at this level.

Figure 13 shows the barrel and endcaps mass distribution with the same cut but only for category 0 which has the most narrow photon showers in the calorimeter. We see that the total signal and signal to background are smaller in the endcaps. Figure 14 shows the mass distribution for barrel events with a tight cut at 0.97. Here we can clearly see the importance of background smoothing which has not been applied for this distribution. Figure 15 shows the mass distribution for neural net output greater than 0.97 in category 0. Again it is clear that smoothing in two dimensions is needed to get a reasonable estimate of the background. It is useful to note that even in this very high s/b region, the largest contribution to the signal is from gluon fusion and that all three signal sources are represented.

7.6 Combination of Categories into Final s/b Distribution

At this point we have signal and background binned in s/b in six categories. We could simply use these to calculate the confidence level, however, it seems most useful, in light of future plans to analyze separate channels, to combine the categories into one s/b plot. We combine the six histograms into one which can be used to calculate confidence levels. In principle, this is the same as combining results from different channels or even from different

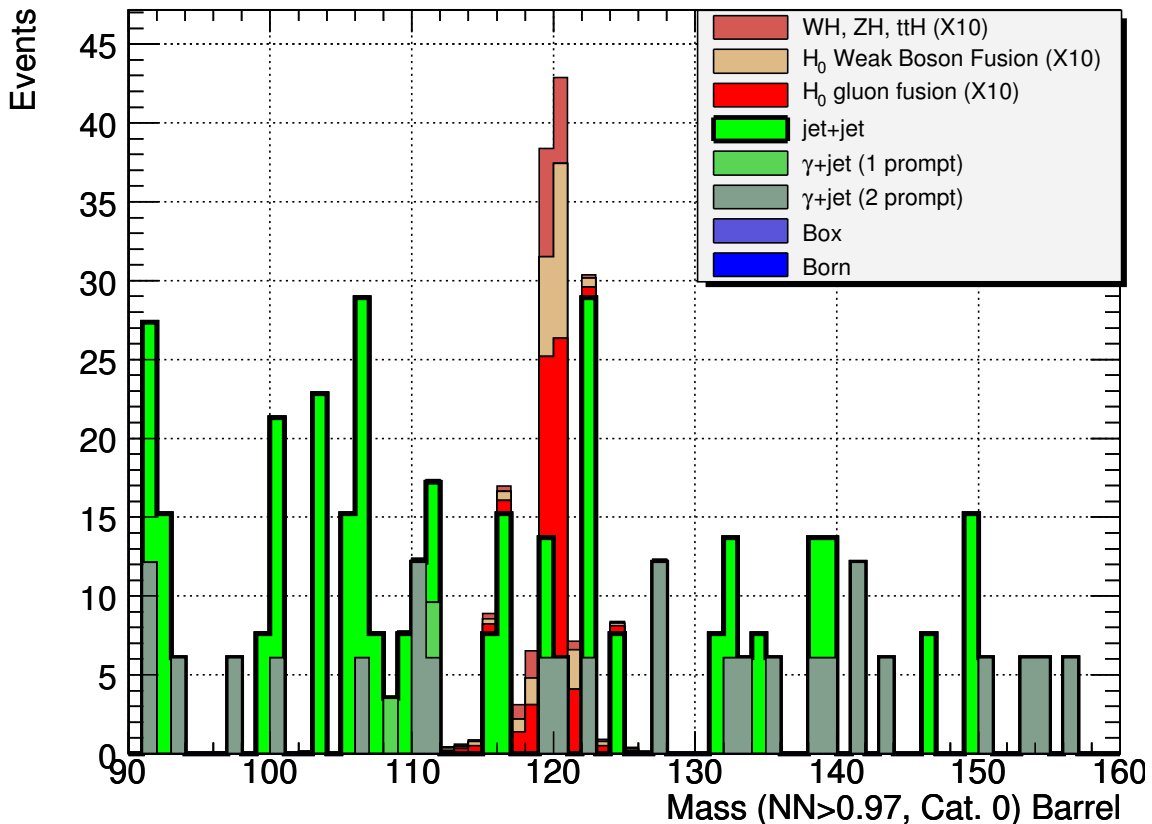


Figure 15: The di-photon mass distribution for each source for barrel events with kinematic neural net output greater than 0.97 and $R_\theta > 0.948$. Events are normalized to an integrated luminosity of 7.7 fb^{-1} and the signal ($M_H=120 \text{ GeV}$) is scaled by a factor 10.

experiments in a way that makes optimal use of all channels and does not pollute high quality channels with data of lesser purity.

We combine the six histograms into one final histogram based on the actual signal to background in each bin.

The final binning of data into s/b bins is shown in figure 16. The plot extends from very low s/b to a small number of events with $s/b > 1$.

One can estimate the relative contribution of each barrel or endcap category by comparing the total LLR with the one computed excluding one category at a time. The six categories have rather widely varying contributions to the Log Likelihood Ratio and hence to the performance of the analysis. Table 10 shows the fraction of signal compared to the fraction of the LLR for each category.

Some of the categories have fairly small effect on the analysis. This remains true after the application of systematic normalization uncertainties described below. It is clear that photons which convert cause a significant deterioration of the performance of this analysis. We hope to mitigate this somewhat by using the reconstructed conversions in the future, however, the poorer mass resolution cannot be recovered and we don't expect a big effect.

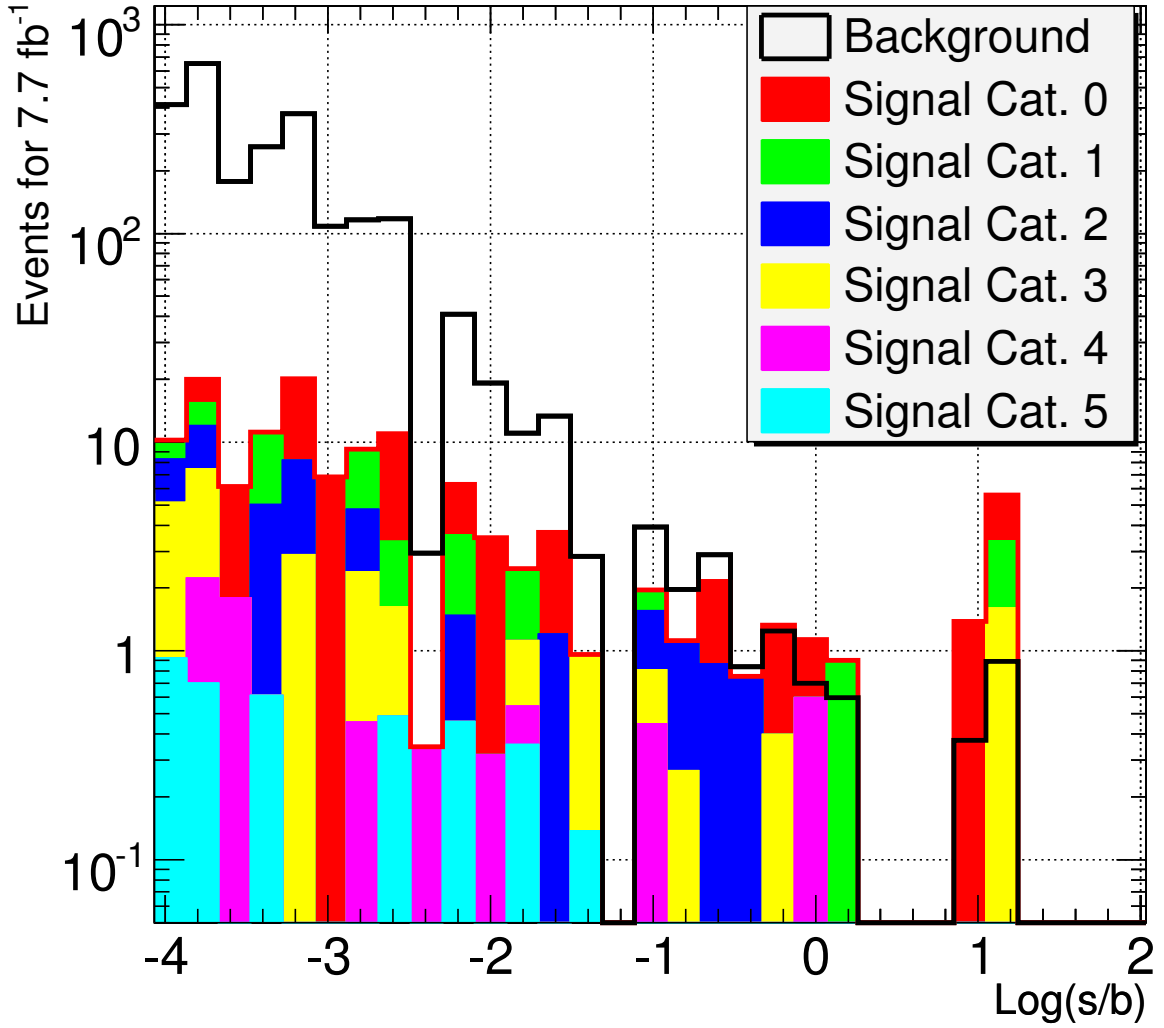


Figure 16: The final distribution of binned signal ($M_H=120$ GeV) and background in $\log(s/b)$ for an integrated luminosity of 7.7fb^{-1} . Here the Higgs signal is normalized to the integrated luminosity and the statistics benefits of the smoothing of the background. Signal and background events are added independently.

Table 10: Performance in the six categories for $M_H = 120$ GeV.

Category	Signal %	LLR %
0	27.8	48.0
1	16.1	24.8
2	21.7	11.9
3	16.6	9.7
4	9.0	4.1
5	8.8	1.5

7.7 Systematic Errors

In the optimized analysis there will also be an important systematic error on the background interpolation under the Higgs mass peak. We estimate that this will be a 1% uncertainty on the background normalization. This normalization uncertainty primarily affects the low s/b region.

Again, to reduce the effect of this shift in normalization it is best to impose a minimum cut on s/b to eliminate bins with a large amount of background and little signal. For all analyses we require $s/b > 0.02$.

The systematic error on the signal does affect the estimate of the amount of luminosity needed for a discovery.

7.8 Results of the Optimized Analysis

We have performed most of our development and studies for a Higgs mass of 120 GeV. For this mass, we find that a five sigma discovery can be made with about 7 fb^{-1} luminosity.

A 1% background normalization uncertainty corresponds to an increase of the luminosity needed for a 5 sigma discovery from 7 fb^{-1} to 7.7 fb^{-1} .

Of course there is a great deal of uncertainty in this benchmark estimate of luminosity due to our poor understanding of the backgrounds we will have when the LHC starts running, however, this is not considered a systematic error on this analysis since we propose to measure the background from the data.

Figure 17 shows the luminosity needed for a five sigma discovery and the discovery sensitivity with an integrated luminosity of 30 fb^{-1} for several Higgs masses. These should be treated as benchmark numbers at this time not precise predictions of what will happen. Nevertheless, it seems possible to discover, or at least have strong evidence for a low mass Higgs in the first good year of running.

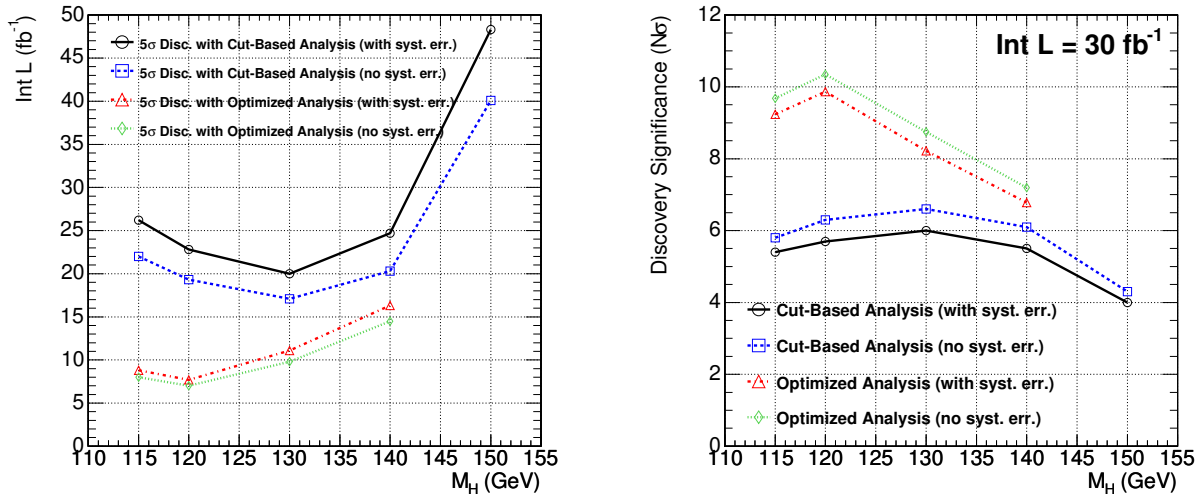


Figure 17: Integrated luminosity needed for a 5σ discovery (left) and discovery sensitivity with an integrated luminosity of 30 fb^{-1} (right) with the optimized analysis. The results from the cut-based analysis in 12 categories are also shown for comparison.

One important result can be gleaned from table 10. The net effect of events in which one or more of the photon candidates is in the endcaps is only about 15%. This may help with prioritization of calibration and analysis effort.

8 Measurement of the Higgs boson mass

If the Higgs boson will be discovered in the $H \rightarrow \gamma\gamma$ channel then we will be able to measure its mass. We have studied the mass measurements with the cut based analysis with two different methods:

- measurement from the $\Delta\text{Log}(\text{likelihood})$ using all events;
- measurement from the $\Delta\text{Log}(\text{likelihood})$ using the cut-based analysis split in 12 categories.

The expected statistical errors are shown in Table 11 for an integrated luminosity of 30 fb^{-1} . The statistical errors simply scale with $1/\sqrt{\text{Int L}}$. The errors are slightly asymmetric, due to the tail of the reconstructed Higgs mass distribution at lower masses, the positive error being approximately 10% smaller than the negative. The Table shows the average between the two.

Table 11: Expected statistical errors on the Higgs boson mass measurement for 30 fb^{-1} .

	$M_H=115 \text{ GeV}$	$M_H=120 \text{ GeV}$	$M_H=130 \text{ GeV}$	$M_H=140 \text{ GeV}$	$M_H=150 \text{ GeV}$
All events	184 MeV 0.16%	184 MeV 0.15%	201 MeV 0.15%	222 MeV 0.16%	298 MeV 0.20%
12 categories	127 MeV 0.11%	139 MeV 0.12%	129 MeV 0.10%	156 MeV 0.11%	204 MeV 0.14%

As we can see the statistical error will be 0.1 to 0.2% already with 30 fb^{-1} , when the significance of the discovery would be 5 to 6 σ with the cut based analysis. Of course this measurement will be affected by the uncertainty of the absolute scale of the photon energy measurement that will be derived for example by the measurement of the Z mass in the radiative Z decays $Z \rightarrow \mu\mu\gamma$.

9 Summary

With a standard cut-based analysis with less than 30 fb^{-1} of integrated luminosity we can discover the Higgs boson with 5σ significance between the LEP lower limit and 140 GeV. Approximately 5 fb^{-1} are needed to exclude its existence in the same mass range if the SM Higgs boson does not exist in that mass range.

We have also demonstrated that the $H \rightarrow \gamma\gamma$ channel can be used to discover a low mass Higgs with an integrated luminosity not too different from that needed for higher mass Higgs, 7.7 fb^{-1} at 120 GeV. Because of the excellent mass resolution expected in the di-photon channel, we can measure the background rate and characteristics from the data at masses away from the Higgs mass hypothesis.

We have presented here an inclusive analysis. In future we will identify the various signal channels by looking for additional jets, leptons, or missing energy. This will clearly improve the sensitivity of the analysis. For the moment these channels are investigated individually in other studies [7, 8].

Acknowledgments

We would like to thank Tiziano Camporesi and Ren Yuan Zhu for the careful reading of the manuscript and for useful comments.

References

- [1] CMS Technical Proposal, CERN/LHCC 94-38, LHCC/P1, 15 December 1994.
- [2] A. Favara and M. Pieri, Preprint DFF-278/4/1997, E-preprint hep-ex 9706016.

- [3] L3 Collaboration, M. Acciarri *et al.* Phys. Lett. B **411** (1997) 373.
- [4] CMS ECAL TDR, CERN/LHCC 97-33, CMS TDR 4, 15 December 1997.
- [5] M. Spira, “HIGLU: A program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections”, arXiv:hep-ph/9510347.
- [6] A. Djouadi, J. Kalinowski and M. Spira, “HDECAY: A program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension”, Comput. Phys. Commun. **108** (1998) 56 (arXiv:hep-ph/9704448).
- [7] M. Dubinin *et al.* CMS Note 2006/097.
- [8] M. Lethuillier *et al.* CMS Note 2006/110.
- [9] <http://www.thep.lu.se/torbjorn/Pythia.html>;
T. Sjöstrand, Computer Physics Communications **80** (1994) 74;
S. Mrenna, Computer Physics Communications **101** (1997) 292.
- [10] Geant4 Collaboration, NIM A **506** (2003) 250;
Geant3, CERN Program Library Long Writeup W5013.
- [11] CMSIM, CMS simulation and reconstruction package, <http://cmsdoc.cern.ch/cmsim/cmsim.html>
- [12] Object Oriented Simulation for CMS Analysis and Reconstruction, <http://cmsdoc.cern.ch/OSCAR/>
- [13] Object Oriented Reconstruction for CMS Analysis, <http://cmsdoc.cern.ch/orca>
- [14] T. Binoth *et al.* hep-ph/0204316, (2002).
- [15] T. Binoth *et al.* hep-ph/0403100, (2004).
- [16] T. Binoth *et al.* hep-ph/0203064, (2002).
- [17] T. Binoth, hep-ph/0005194 (2000).
- [18] Z. Bern *et al.*, Phys. Rev. **D66** 074018 (2002).
- [19] L. Agostino and M. Pieri, CMS Note 2006/078.
- [20] E. Meschi *et al.* CMS Note 2001/034.
- [21] L. Agostino *et al.* CMS Note 2006/021.
- [22] N. Marinelli, CMS Note 2006/005.
- [23] V. Litvin *et al.* CMS Note 2002/030.
- [24] M. Pieri *et al.* CMS Note 2006/007.
- [25] The LEP Collaborations: ALEPH, DELPHI, L3 and OPAL, “Lower Bound for the Standard Model Higgs Boson Mass from Combining the Results of the Four LEP Experiments”, CERN EP 98-046, April 1998.
- [26] S. Baffioni *et al.* CMS Note in preparation.