Inclusive Search for the SM Higgs Boson in the $H \rightarrow \gamma\gamma$ channel at the LHC

Serguei Ganjour

CEA-Saclay/IRFU, 91191 Gif-sur-Yvette, France
E-mail: Serguei.Ganjour@cea.fr

Abstract. A prospective for the inclusive search of the Standard Model Higgs boson in the decay channel $H \rightarrow \gamma\gamma$ is presented with the CMS experiment at the LHC. The analysis relies on a strategy to determine the background characteristics and systematics from data. The strategy is applied to a Monte Carlo model of the QCD background, with full simulation of the detector response. The discrimination between signal and background exploits information on photon isolation and kinematics. The resolution for the reconstructed Higgs boson mass profits from the excellent energy resolution of the CMS crystal calorimeter. A discovery significance above 5 sigma is expected at integrated LHC luminosities below 30 fb$^{-1}$ for Higgs boson masses below 140 GeV/c$^2$.

Keywords: LHC, CMS, Higgs
PACS: 14.80.Bn

INTRODUCTION

Since the beginning of the LHC project the $H \rightarrow \gamma\gamma$ channel is considered as a major discovery channel of Higgs particles [1] at masses between LEP limit 114.4 GeV/c$^2$ [2] and about 140 GeV/c$^2$. Indeed, despite a high rate of the $H \rightarrow b\bar{b}$ decays at low mass, it remains out of our interest due to large QCD background and low mass resolution of the di-jet state. Since the $H \rightarrow \gamma\gamma$ decay involves virtual loops, it has relatively small branching fraction of about $10^{-3}$. However, the signal has a clean signature with two high $E_T$ isolated photons. Due to an excellent resolution of the CMS electromagnetic calorimeter (ECAL), it can be identified as a narrow peak at a Higgs boson mass on the top of a continuous background.

We investigated two analysis methods for the $H \rightarrow \gamma\gamma$ searches at CMS [3, 4]. In addition to the conventional cut-based analysis, we report the optimized discovery oriented technique based on a multivariate optimization. The latter exploits a difference in the signal and the background kinematics. These studies use a full CMS detector simulation program assuming $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ machine luminosity and including collision effects such as minimum bias and underling events in the simulation model.

RATES AND CROSS SECTIONS

The inclusive search of the Higgs boson implies any production mechanism. In proton-proton collisions at LHC energy about 80% of the Higgs bosons are produced in the gluon fusion reaction, while the rest are produced in association with either $q\bar{q}$ pairs.
(WVB or $t\bar{t}$ fusion) or vector bosons. Table 1 presents the cross sections and $H \rightarrow \gamma\gamma$ branching fractions for the different Higgs boson mass [6].

**TABLE 1.** Signal NLO cross sections and branching ratios.

| $M_H$ (GeV/c$^2$) | 115 | 120 | 130 | 140 | 150 |
|-------------------|-----|-----|-----|-----|-----|
| $\sigma_{gg}$ fusion (pb) | 39.2 | 36.4 | 31.6 | 27.7 | 24.5 |
| $\sigma_{WVB}$ fusion (pb) | 4.7 | 4.5 | 4.1 | 3.8 | 3.6 |
| $\sigma_{WH, ZH, t\bar{t}H}$ (pb) | 3.8 | 3.3 | 2.6 | 2.1 | 1.7 |
| Total $\sigma$ (pb) | 47.6 | 44.2 | 38.3 | 33.6 | 29.7 |
| $\mathcal{B}(H \rightarrow \gamma\gamma)$, % | 0.21 | 0.22 | 0.22 | 0.20 | 0.14 |
| $\sigma \times \mathcal{B}$ (fb) | 99.3 | 97.5 | 86.0 | 65.5 | 41.5 |

We consider two sorts of the background processes. The “irreducible” background has two real high $E_T$ isolated photons. Such a signature can be produced by both quark-antiquark annihilation (“born”) and gluon-gluon fusion (“box”) as well as quark- gluon Compton scattering with isolated bremsstrahlung processes. We estimate the total differential rate of “irreducible” backgrounds about 100 fb/GeV/c$^2$ at 120 GeV/c$^2$ mass. Thus, we require about 1 GeV/c$^2$ two-photon mass resolution for a powerful discrimination of the signal (see Table 1).

The dominant QCD processes like $\gamma$+jet and di-jets may lead to the fake photons induced by neutral hadrons $\pi^0$ or $\eta$ and produced in the jet fragmentation processes. The “reducible” background has at least one non-isolated photon. The PYTHIA LO cross sections of the background processes are presented in Table 2. To compute the background yields at NLO cross sections we apply the K-factors summarized in Table 3. Thus, a jet suppression has to be better then $10^{-3}$ to diminish its rate at the level of the “irreducible” background.

**DETECTOR**

The conceptual design of the CMS detector [5] exploits the conventional layout of the particle detector at hadron colliders. A dedicated electromagnetic calorimeter is required for a powerful search of the Standard Model (SM) Higgs boson. Indeed, the $H \rightarrow \gamma\gamma$ decay was employed as a benchmark channel in the design of the ECAL. It comprises 75848 lead tungstate (PWO$_4$) crystals placed in two pseudo-rapidity regions: the barrel ($|\eta| < 1.479$) and the endcap ($1.479 < |\eta| < 3.0$). The ECAL endcap is equipped by a preshower (ES) system for $\pi^0$ rejection. Due to low Moliere radius (2.19 cm) about

**TABLE 2.** LO cross sections for backgrounds.

| Process         | $p_T$ (GeV/c) | $\sigma_{LO}$ (pb) |
|-----------------|---------------|---------------------|
| $pp \rightarrow \gamma\gamma$ (born) | $> 25$ | 82 |
| $pp \rightarrow \gamma\gamma$ (box) | $> 25$ | 82 |
| $pp \rightarrow \gamma$+jet | $> 30$ | $5 \times 10^4$ |
| $pp \rightarrow$ jets | $> 50$ | $2.8 \times 10^7$ |
| Drell Yan ee    | $-$ | $4 \times 10^3$ |
TABLE 3. Background K-factors applied for the PYTHIA cross section

| Process | K-factor |
|---------|----------|
| pp → γγ (born) at the test-beam. pp → γγ (box) | 1.5 |
| pp → γ+jet (2 prompt) | 1.72 |
| pp → γ+jet (1 prompt) | 1 |
| pp → jets | 1 |

80% of the shower energy is deposited in one crystal having $2.2 \times 2.2$ cm$^2$ front face size. Such a high granularity allows us a powerful identification of isolated photons.

![Figure 1](image1.png)

**FIGURE 1.** Energy resolution of the barrel supermodule.

Figure 1 shows the energy resolution of the ECAL barrel obtained with an electron test-beam. The stochastic ($S$), the noise ($N$) and the constant ($C$) terms are obtained by a fit with the following function:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2. \quad (1)$$

Clustering algorithm accounts strong magnetic field and presence of material in front of the calorimeter. Measured energy can be computed as

$$E_{e,\gamma} = F \times \sum_{clusters} Gc_i A_i, \quad (2)$$

where $F$ is a correction function, $Gc_i$ is a calibration factor expressed as a product of a global absolute scale and intercalibration constants, respectively. $A_i$ is a signal
amplitude, in ADC counts. Sum of the incident energy in $5 \times 5$ crystals provide the best energy estimation for the unconverted photons ($F = 1$). The dedicated procedure of ECAL calibration is foreseen with the physics events such as $W \rightarrow e\nu$, $Z \rightarrow e^+e^-$ and $\pi^0 \rightarrow \gamma\gamma$.

**ANALYSIS AND RESULTS**

Higgs decays into two-photon state are selected with extremely high efficiency both Level-1 (99.7%) and High Level Triggers (88.4%). Due to tighter kinematical and isolation criteria applied in the analysis, no additional signal restriction by trigger were observed.

Two-photon candidates are selected within fiducial volume $|\eta| < 2.5$, and $p_T^\gamma > 40, 35 \text{ GeV}/c$. Since fake photons produced in the jets are accompanied by additional energetic particles, we require no tracks with $p_T > 1.5 \text{ GeV}/c$ inside a cone around photon with a radius $\Delta R < 0.3$. In addition, we demand $\sum E_T < 6(3) \text{ GeV}/c$ in the ECAL barrel (endcap) in a cone $0.06 < \Delta R < 0.35$ and $\sum E_T < 6(5) \text{ GeV}/c$ in the HCAL barrel (endcap) in a cone $\Delta R < 0.3$. The energy response of individual crystals were smeared according to calibration precision expected for an integrated luminosity of $10 \text{ fb}^{-1}$ and obtained with $W \rightarrow e\nu$ control sample.

Due to longitudinal spread of the interaction vertices ($\sim 50 \text{ mm}$) the di-photon mass is smeared by about $1.5 \text{ GeV}/c^2$. Indeed, Higgs bosons are produced in association with tracks from underlying events, initial state gluon radiation and associative particles $qqH$, $WH$, $ZH$. These tracks allow us to identify the interaction vertex, in about 80% of cases, and correct momenta of the photons.

Figure 2 shows the di-photon mass distribution for the signal (scaled by a factor 10) and the background events after the selection. For a 120 GeV/c$^2$ Higgs boson, the signal efficiency of 30% yields to 29.3 observed Higgs events per inverse femtobarn, while the total background is 178 fb/GeV/c$^2$ as shown in Table 4. Despite asymmetric lineshape about 60% of the signal remains within $\pm 1 \text{ GeV}/c^2$ mass window.

| Process                  | $M_H$, (GeV/c$^2$) |
|--------------------------|--------------------|
|                          | 115 | 120 | 130 | 140 |
| $pp \rightarrow \gamma\gamma$ (born) | 48  | 44  | 36  | 29  |
| $pp \rightarrow \gamma\gamma$ (box)    | 36  | 31  | 23  | 16  |
| $pp \rightarrow \gamma+\text{jet (2prompt)}$ | 43  | 40  | 32  | 26  |
| $pp \rightarrow \gamma+\text{jet (1prompt)}$ | 40  | 34  | 22  | 19  |
| $pp \rightarrow \text{jets}$             | 29  | 27  | 20  | 18  |
| Drell Yan ee               | 2   | 2   | 1   | 1   |
| **Total background**       | 203 | 178 | 134 | 109 |

Including of new variables or cuts optimization do not lead to the further improvement of the signal over background ratio ($s/b$). However, we find the $R_9$, fraction of the super-cluster energy deposited a small $3 \times 3$ crystal area, as a powerful discriminator of a $\pi^0$. Indeed, high value of $R_9$ readily identifies converted photons and automatically selects
against $\pi^0$. The converted category remains background enriched. However, the use of a track information would allow further rejection of the background.

To improve signal significance, we split the analyzed sample into categories with different $s/b$. Asides single category, we consider four categories, 2 categories where both photons are detected in the barrel and 2 where at least one photon is in the endcap. Each of these 2 categories are splitted for high ($>0.93$) and low ($<0.93$) values of $R_9$. Then, we form twelve categories sample based on 3 ranges in $R_9$ (0.9, 0.948) and 4 pseudo-rapidity regions in $|\eta|$ (0.9, 1.4, 2.1). Table 5 summarizes the integrated luminosity needed to discover or to exclude Higgs boson for the different event splitting. Confidence level is calculated with a frequentistic approach using a log-likelihood ratio (LLR) estimator. These outcomes are computed for many possible pseudo-experiments for a hypothesis when the signal exists and that it does not.

TABLE 5. Integrated luminosity required for observation or exclusion of the Higgs boson with a mass of 120 GeV/c$^2$.

|        | 5\sigma | 3\sigma | 95% C.L. excl. |
|--------|---------|---------|----------------|
| 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) |

The optimized analysis exploits six categories, 3 categories where both photons are detected in the barrel and 3 where at least one photon is in the endcap. These 3 categories are defined according to measured $R_9$, as for the cut-based analysis splitted into 12 categories. Loose cuts are applied for the isolation variables which are used as a neural network prior. Fixed cut for the optimized $NN_{iso}$ output variable slightly improves a background rejection. Asides two-photon mass, four other kinematical variables can be
used for further discrimination of the signal. They are transverse energy of each photon $E_T^{\gamma 1, \gamma 2}$, $\eta$ difference between two photons $|\eta_1 - \eta_2|$ and longitudinal momentum of the photon pair $P_L^{\gamma \gamma}$. These variables are combined with $NN_{iso}$ for the further neural network optimization. To reject the background rate a neural net is trained with the mass side-band events and results into $NN_{kin}$ output variable. Due to negligible correlation between invariant mass and $NN_{kin}$, the $s/b$ expectation can be estimated for each event as

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

Finally, the events are binned according to the $s/b$ estimate. Then, we exploit log($s/b$) distribution to compute the confidence level with the LLR estimator.

**FIGURE 3.** Integrated luminosity required for a 5\(\sigma\) discovery as a function of the Higgs mass.

Figure 3 shows the integrated luminosity needed for a 5\(\sigma\) discovery. For a 120 GeV/c\(^2\) Higgs boson, such a discovery can occur with about 7.7 fb\(^{-1}\) recorded data using an optimized technique, while it is required about 22.8 fb\(^{-1}\) of integrated luminosity using conventional cut-based analysis technique (see Table 5). We address major systematic errors. The dominant contribution is due to background subtraction. It is evaluated from the uncertainty of the fit function in the mass side-bands.
CONCLUSIONS

The SM Higgs boson can be readily discovered with an integrated luminosity less than 30 fb$^{-1}$, if its mass is below 140 GeV/c$^2$. We investigated the standard cut-based and discovery oriented optimized analysis techniques. The best achievement corresponds to a 120 GeV/c$^2$ Higgs boson where 7.7 fb$^{-1}$ of a recorded data is required for the 5$\sigma$ discovery. The analysis strategy implies data driven methods of the background subtraction using a side-band mass regions. There is a potential room for improvement by adding the track information to identify converted photons and ECAL preshower to reject $\pi^0$.

REFERENCES

1. CMS Collaboration, “The Compact Muon Solenoid Technical Proposal”, CERN/LHCC 1994-38
2. LEP Working Group for Higgs Boson Searches Collaboration, R. Barate et al, Phys.Lett. B 565 61
3. CMS Collaboration, “The CMS Physics Technical Design Report, Volume II”, CERN/LHCC, 2006-021
4. CMS Collaboration, M. Pieri et al, CMS-NOTE-2006/112
5. CMS Collaboration, “The CMS Physics Technical Design Report, Volume I”, CERN/LHCC, 2006-001
6. M. Spira, arXiv: [hep-ph/9510347]; A. Djouadi, J. Kalinowski and M. Spira, Comput.Phys.Commun. 108 (1998), 56 [arXiv: hep-ph/9704448]
7. T. Binoth et al, arXiv: hep-ph/0005194; T. Binoth et al, arXiv: hep-ph/0204316; T. Binoth et al, arXiv: hep-ph/0203064; T. Binoth et al, arXiv: hep-ph/0403100; Z. Bern et al, Phys.Rev. D 66 074018 (2002)