



SAPIENZA  
UNIVERSITÀ DI ROMA

# Searching for new physics in the Higgs sector: the $H \rightarrow \gamma\gamma$ channel at CMS

Scuola di dottorato Vito Volterra  
Dottorato di Ricerca in Fisica – XXVII Ciclo

Candidate  
Francesco Micheli  
ID number 1172095

Thesis Advisor  
Dr. Paolo Meridiani


A thesis submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Physics  
December 2014

Thesis defended on January 27, 2015  
in front of a Board of Examiners composed by:  
Prof. Anna Di Ciaccio (chairman)  
Prof. Nadia Pastrone  
Prof. Valeria Ferrari

---

**Searching for new physics in the Higgs sector: the  $H \rightarrow \gamma\gamma$  channel at CMS**

Ph.D. thesis. Sapienza – University of Rome

 2014 Francesco Micheli.

Released under a Creative Commons *Attribution-Non Commercial-Share Alike* License

This thesis has been typeset by L<sup>A</sup>T<sub>E</sub>X and the Sapthesis class.

Version: December 16, 2014

Author's email: francesco.micheli@cern.ch

*Indugiarono un po' a considerare le orme che i partiti avevano lasciato e poi mossero gli occhi intorno e in alto. C'era da restare accecati a voler fissare là dove il cielo d'un azzurro di maggio si saldava alla cresta delle colline, di tutto nude fuorché di neve cristallizzata. Una irresistibile attrazione veniva, col barbaglio, da quella linea: sembrava essere la frontiera del mondo, da lassù potersi fare un tuffo senza fine.*

***Beppe Fenoglio***

*My practicality consists in this: in the knowledge that if you beat your head against the wall it is your head which breaks and not the wall ... that is my strength, my only strength.*

***Antonio Gramsci***



# Contents

<b>Introduction</b>	<b>vii</b>
<b>1 The Role of the Higgs Boson in the Standard Model and Beyond</b>	<b>1</b>
1.1 The Standard Model and the Higgs Mechanism . . . . .	1
1.2 Higgs Production Modes at LHC . . . . .	3
1.2.1 The diphoton Decay Channel . . . . .	4
1.3 Experimental Searches of the Higgs Boson . . . . .	6
1.3.1 Searches for the Higgs boson at LEP and Tevatron . . . . .	6
1.3.2 The Higgs Boson Discovery at LHC . . . . .	7
1.4 Standard Model limitations . . . . .	7
1.5 New Physics in the Higgs Sector . . . . .	9
1.5.1 The effective Lagrangian for BSM models . . . . .	9
1.5.2 BSM models: 2HDM and VLQ . . . . .	11
1.6 Coupling structure of a Higgs-like particle . . . . .	12
1.6.1 Benchmark parametrizations . . . . .	14
1.7 Analysis strategy for BSM searches in $H \rightarrow \gamma\gamma$ channel . . . . .	15
<b>2 The CMS Experiment at the LHC</b>	<b>17</b>
2.1 The Large Hadron Collider . . . . .	17
2.2 The CMS Experiment . . . . .	18
2.2.1 Magnet . . . . .	22
2.2.2 The Silicon Tracker . . . . .	23
2.2.3 The Electromagnetic Calorimeter . . . . .	24
2.2.3.1 The crystals . . . . .	25
2.2.3.2 Geometry . . . . .	26
2.2.3.3 Energy resolution . . . . .	27
2.2.4 The Hadronic Calorimeter . . . . .	27
2.2.5 The Muon System . . . . .	28
2.2.6 The Trigger System . . . . .	28
2.2.7 CMS DAQ and Monte Carlo simulations . . . . .	29
<b>3 Reconstruction and Identification of Physics Objects</b>	<b>31</b>
3.1 Photons . . . . .	31
3.1.1 Photon reconstruction . . . . .	31
3.1.2 Photon Energy . . . . .	32
3.1.3 Diphoton triggers and photon preselection . . . . .	34

3.1.4	Cut-based photon identification . . . . .	37
3.1.5	Photon Identification MVA . . . . .	38
3.2	Jets . . . . .	40
3.3	Leptons . . . . .	41
3.3.1	Muons . . . . .	41
3.3.2	Electrons . . . . .	42
<b>4</b>	<b>The <math>H \rightarrow \gamma\gamma</math> SM Analysis and Coupling Measurements</b>	<b>43</b>
4.1	Introduction . . . . .	43
4.2	Data Samples and simulated events . . . . .	43
4.2.1	Diphoton vertex assignment . . . . .	44
4.2.2	Event Classification overview . . . . .	45
4.2.3	Multivariate event classifier . . . . .	47
4.3	Exclusive Categories Tagging for Associate Production Modes . . . . .	48
4.3.1	VBF Tag . . . . .	50
4.3.2	Tags for VH selection . . . . .	51
4.3.2.1	VH Leptonic Tag . . . . .	52
4.3.2.2	VH $E_T^{\text{miss}}$ Tag . . . . .	52
4.3.2.3	VH Hadronic Tag . . . . .	52
4.3.3	Tags for ttH selection . . . . .	53
4.3.4	Signal Model . . . . .	54
4.4	Fit Methodology . . . . .	56
4.5	Systematic Uncertainties . . . . .	65
4.6	Results . . . . .	69
4.6.1	Standard Model Analysis Results . . . . .	69
4.6.1.1	Couplings measurement . . . . .	72
4.6.2	Combination with Other Channels . . . . .	73
4.6.2.1	ttH Combination with Other Channels . . . . .	80
4.7	Summary and conclusions . . . . .	82
<b>5</b>	<b>Direct searches beyond Standard Model in the Higgs sector</b>	<b>83</b>
5.1	Associate Production of a Single Top Quark and a Higgs boson . . . . .	83
5.1.1	Analysis strategy . . . . .	85
5.1.2	Event Selection . . . . .	86
5.1.3	Background Estimation . . . . .	88
5.1.4	Systematic Uncertainties . . . . .	90
5.1.5	Results . . . . .	91
5.2	Search for Pair Production of Heavy Vector-Like Partners of the Top Quark . . . . .	93
5.2.1	Analysis Strategy . . . . .	93
5.2.2	Event Selection . . . . .	95
5.2.3	Systematic Uncertainties . . . . .	96
5.2.4	Results . . . . .	96
5.2.5	Combination with Other channels . . . . .	100
5.2.6	Summary of T searches . . . . .	103
<b>6</b>	<b>Conclusions</b>	<b>109</b>

# Introduction

The Standard Model of elementary particles is the theory describing interactions between the fundamental constituents of matter. This elegant theoretical model, developed in the late sixties, has been verified at a high precision level by different experiments over the last decades.

The last elusive particle predicted by the Standard Model, the Higgs boson, has been discovered in 2012 by experiments operating at the *Large Hadron Collider* (LHC), a proton-proton accelerator located at the CERN laboratories in Geneva. This particle has a central role in the theory, since it is responsible, through the Higgs mechanism, of the mass generation of fundamental particles. Its discovery is an important milestone in high-energy physics and a confirmation of the robustness of the actual description of fundamental laws of Nature.

Despite this great success, the Standard Model cannot be seen as a definitive, complete theory of fundamental interactions. First of all, it does not incorporate gravity, one of the four fundamental forces. Moreover, the Standard Model only describes the ordinary matter, which constitutes about 5% of the whole universe, failing to describe experimentally observed phenomena like dark matter and dark energy. From a mathematical point of view, the actual theory is also considered inelegant and implausible for the high number of free parameters and the so-called *fine tuning* problem, related to the Higgs boson mass.

Several theories beyond Standard Model (BSM) have been elaborated to overcome these limitations and include the Standard Model in a more general theoretical framework able to solve these open issues. Most of these models predict new particles, new types of interactions and a modified Higgs sector. These effects are expected to appear at the TeV energy scale and thus could be accessible at LHC.

The recent discovery of the Higgs boson, therefore, can be seen as a portal to access the BSM sector and a brand-new opportunity to prove or deny different models. Two different approaches can be followed to achieve this goal: indirect constraints from precision measurements of the properties of the newly discovered particle or direct searches for new particles predicted by BSM theories.

This thesis describes indirect and direct experimental searches in the Higgs sector exploiting the decay of this particle in a pair of photons ( $H \rightarrow \gamma\gamma$ ). The analyses presented here use data collected by the *Compact Muon Solenoid* (CMS) detector, analyzing the proton-proton collisions produced by the LHC. The full dataset corresponds to  $5.1 \text{ fb}^{-1}$  of data collected at  $\sqrt{s} = 7 \text{ TeV}$  during 2011 and  $19.7 \text{ fb}^{-1}$  collected at  $\sqrt{s} = 8 \text{ TeV}$  during 2012.

Chapter 1 gives a theoretical overview of Higgs boson physics. The role of this particle in the Standard Model and in alternative BSM models are reviewed.

Moreover, it summarizes past experimental searches and explains the characteristics of the diphoton decay channel, which is the chosen one for all the different analyses presented.

Chapter 2 offers a detailed description of the experimental apparatus. After a brief introduction on the LHC, the various subdetectors composing the Compact Muon Solenoid experiment are introduced. Particular attention is given to the description of the electromagnetic calorimeter, given its importance in photon reconstruction.

Identification and reconstruction of physics objects are described in chapter 3. The main focus of the chapter is photon reconstruction, which has been designed specifically for the  $H \rightarrow \gamma\gamma$  decay channel.

Chapter 4 presents the measurements of the Higgs boson properties in the diphoton decay channel. The whole analysis, including object reconstruction and statistical methods, is reviewed. The original contribution of this thesis to the  $H \rightarrow \gamma\gamma$  analysis concerns the creation of exclusive categories tagging associate production mechanisms, in particular Higgs boson production in association with vector bosons (W or Z) or a pair of top quarks ( $t\bar{t}H$ ). The design and optimization of these categories have been the main topic of the first years of my PhD. For this reason, particular emphasis is given to the description of these categories. The inclusion of these production modes in the analysis, gave the possibility to precisely measure the Higgs boson couplings to vector bosons and top quarks. Therefore, these results can be used to indirectly constrain the phase space of BSM models described in chapter 1.

Chapter 5 describes direct searches of BSM particles with the Higgs boson in their decay chain. The study of these processes was a natural extension of the work performed in  $t\bar{t}H$  analysis at the beginning of my PhD. Indeed, these processes present a similar final state: two high energy photons and top quarks. Therefore, I could exploit, in BSM searches with different topologies, the experimental techniques and the tools developed for  $t\bar{t}H$  analysis. Two different searches are described: single top associate Higgs production and heavy vector-like partners of the top quark. The first process can highlight anomalous top-Higgs coupling while the second one could be an evidence of a new strong sector of heavy quarks, useful to solve the hierarchy problem.



## Chapter 1

# The Role of the Higgs Boson in the Standard Model and Beyond

The Higgs boson was the last missing piece of the Standard Model. Its discovery, at LHC experiments, is one of the main achievements in the history of particle physics. This breakthrough represents a starting point for an extended characterization of the electroweak symmetry breaking mechanism and a brand-new opportunity to spot new physics effects using the Higgs sector.

This chapter provides a theoretical overview of Higgs boson physics. The role of this particle in the Standard Model and in possible extended theoretical frameworks will be reviewed. Specific attention will be given to the description of the theories which inspired the experimental searches reported in this thesis.

### 1.1 The Standard Model and the Higgs Mechanism

The Standard Model (SM) is the theoretical model currently used to describe the interaction of fundamental particles. It is an elegant and successful theory developed by Glashow, Weinberg and Salam in the late sixties [1–3] and verified experimentally, during the years, at a high precision level.

The Standard Model is a quantum field theory which models the interaction of elementary particles through electromagnetic, weak and strong interactions. It includes elementary fermions, which are fundamental constituents of the matter, and gauge bosons which are the force carriers mediating the fundamental interactions. Elementary fermions have spin  $\frac{1}{2}$  and are organized in three families of quarks (carrying color charge and therefore subject to strong interaction) and other three families of leptons (interacting only through electroweak forces). Gauge bosons have spin 1 and are of three kinds, according to the interaction that they carry: photons, for electromagnetic interaction, W and Z bosons, for the weak interaction and gluons, for the strong one.

Standard Model theory has the advantage to describe in a unified mathematical structure three of the four fundamental interactions (gravity is not included). Its Lagrangian is invariant under transformations of the symmetry group

$SU(3)_C \times SU(2)_L \times U(1)_Y$ .

This Lagrangian ( $\mathcal{L}_{SM}$ ) can be split in two contribution, according to the different nature of the forces:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW}$$

$\mathcal{L}_{QCD}$ , where QCD stands for quantum chromodynamics, describes interactions between quarks and gluons, with  $SU(3)$  symmetry.  $\mathcal{L}_{EW}$ , where EW stands for electroweak is instead invariant under transformations of the group  $SU(2)_L \times U(1)_Y$  and describes interactions between fermions and EW gauge bosons, i.e. photons, W and Z bosons.

In a gauge theory, if a symmetry is exactly conserved, all particles should be massless, contradicting experimental observations. In the SM this problem is solved through the electroweak symmetry breaking (EWSB) mechanism [4]. This mechanism for mass generation has been proposed independently by Higgs, Brout, Englert and other theorists nearly fifty years ago [4–6].

If a Lagrangian is invariant under a transformation also its non-degenerate eigenstate will reflect that symmetry. On the other side, if an eigenstate of an invariant Lagrangian is degenerate, it could be non-symmetric under the action of the same symmetry. If the ground state is degenerate a unique choice for the state of minimum energy is not possible. One of the possible groundstates, which are all asymmetric, must be chosen and this procedure is known as *spontaneous symmetry breaking mechanism*.

In the SM, this mechanism is realized by the introduction of a scalar isospin doublet, the Higgs field  $\Phi$ . The simplest lagrangian for an autointeracting scalar field has the form:

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi)$$

where  $D_\mu$  is the covariant derivative and the potential  $V(\Phi)$  is given by:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.1)$$

Requiring  $\lambda > 0$  the energy spectrum is bounded below. Therefore ground states exist. Imposing  $\mu^2 < 0$  a degenerate ground state appears, since the potential has its minimum in correspondence of:

$$\Phi^\dagger \Phi = -\frac{\mu^2}{2\lambda}$$

This means that there is an infinite number of ground states satisfying Eq. 1.1 lying on a circle around the local maximum of the potential, i.e. the origin. The electroweak gauge symmetry is, thus, spontaneously broken by the choice of one of the possible ground states, which breaks the rotational symmetry of the Lagrangian.

The expansion of the Lagrangian around the minimum of the potential, using perturbation theory, leads to quadratic terms for the bosons fields. These terms correspond to the mass terms for the  $W^\pm$  and  $Z$  bosons. In this way, the gauge bosons mediating the weak interaction get mass. In order to have a massless photon, in this procedure,  $SU(2)_L \times U(1)_Y$  is spontaneously broken conserving the electric charge symmetry group  $U(1)_Y$ .

Fermions get mass through Yukawa interaction terms, which couple the left-handed component of a fermionic doublet ( $\psi_L$ ) and the right-handed singlet ( $d_R$  or  $u_R$  respectively for down and up quarks) to the Higgs field ( $\Phi$ ). These terms will have the form:

$$g^d \bar{\psi}^L \Phi d^R + g^u \bar{\psi}^L \Phi u^R + \text{h.c.}$$

Writing down these terms for the three families of quarks, the mass Lagrangian for quarks is obtained:

$$\mathcal{L}_{q\Phi} = \sum_{ij} g_{ij}^d \bar{\psi}_i^L \Phi d_j^R + g_{ij}^u \bar{\psi}_i^L \Phi u_j^R + \text{h.c.} \quad (1.2)$$

The mass terms in 1.2 are not diagonal in the  $u$  and  $d$  fields. Diagonalizing and substituting the vacuum expectation value of the Higgs boson in the equation, the mass of the physical fields is obtained. The physical fields, i.e. the mass eigenstate, are linked to the eigenstate of electroweak interactions ( $u$  and  $d$  in equation 1.2) by the Cabibbo-Kobayashi-Maskawa matrix [7, 8]. In this way the mass of fermions and weak decays between different quark families are naturally included in the theory.

Given the presence of these interaction terms in the Lagrangian, Higgs production and decay will proceed through different processes, whose strength is determined by the coupling of that particular interaction. In particular, the SM Higgs boson couplings to fundamental fermions are proportional to the fermion mass  $m_f$ , while couplings to bosons are proportional to the squares of the boson mass  $m_V$  [9]. Consequently, the Higgs dominant production and decay mechanisms will be those involving  $W^\pm$ ,  $Z$  bosons and the third generation of quarks and leptons.

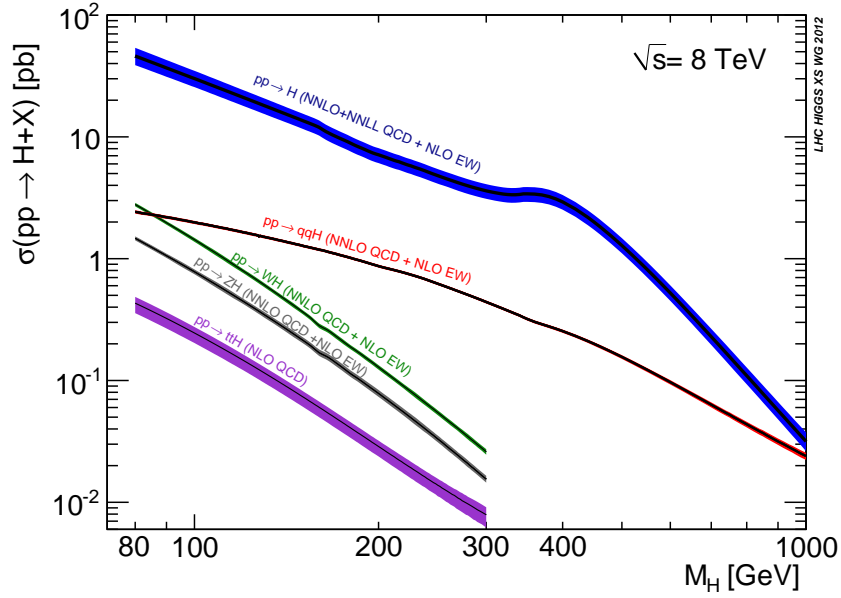
The mass of the Higgs boson  $m_H$  is a free parameter of the theory and should be determined experimentally by direct searches. Searches for the Higgs boson particle at colliders started at the  $e^+e^-$  collider LEP at CERN (operating from 1989 to 2000) [10] and continued at the Tevatron  $p\bar{p}$  collider at Fermilab until its shutdown in 2011 [11]. The ultimate searches for the Higgs boson have been performed at the LHC over the last years and the discovery of a new particle with 125 GeV mass, compatible with the SM Higgs boson, has been announced on the 4<sup>th</sup> of July 2012 [12, 13]. These searches are described in more details in section 1.3.

The challenge of experimental physicists in the next years is to characterize the nature of this particle, understanding if it is exactly the Standard Model Higgs or if there are additional effects beyond the SM theory to be considered.

## 1.2 Higgs Production Modes at LHC

The Large Hadron Collider (LHC) is a proton-proton accelerator operative at CERN since 2009. It operated at a center of mass energy of 7 and 8 TeV and will start again taking data in 2015 at an energy of 13 or 14 TeV. One of the aim that led to the construction of LHC is the discovery of the Higgs boson. More details about LHC will be given in the following chapter. We describe here the production and the decay mechanisms of the Higgs particle at a hadron collider of such energies.

In proton-proton collisions at  $\sqrt{s} = 7-8$  TeV four main production mechanisms are predicted by the SM. The cross sections at 8 TeV for different production mechanisms are shown in Figure 1.3 as a function of the Higgs mass.



**Figure 1.1.** SM Higgs boson production cross sections for pp collisions at 8 TeV, including theoretical uncertainties. Different production modes are shown in different colors: gluon-gluon fusion (blue), vector boson fusion (red), associate production with a W boson (green) or a Z boson (brown), top-pair associate production (violet). ggH cross section is computed at NNLL precision while VBF, WH and ZH at NNLO and ttH at NLO precision.

The one with the largest cross section (one order of magnitude greater than other mechanisms) is gluon-gluon fusion (ggH) where no additional particles, at leading order, are present in the final state. It is followed by vector boson fusion (VBF) which corresponds to the process  $pp \rightarrow qqH$ , associate production with a Z or a W boson (VH) or a top-pair ( $t\bar{t}H$ ). The diagrams corresponding to these production modes are shown in figure 1.2.

The branching ratios for the most relevant decay channels for a SM Higgs boson, as a function of its mass, are shown in Figure 1.3. The numerical values for the branching ratios for a 125 GeV Higgs boson are reported in Table 1.1.

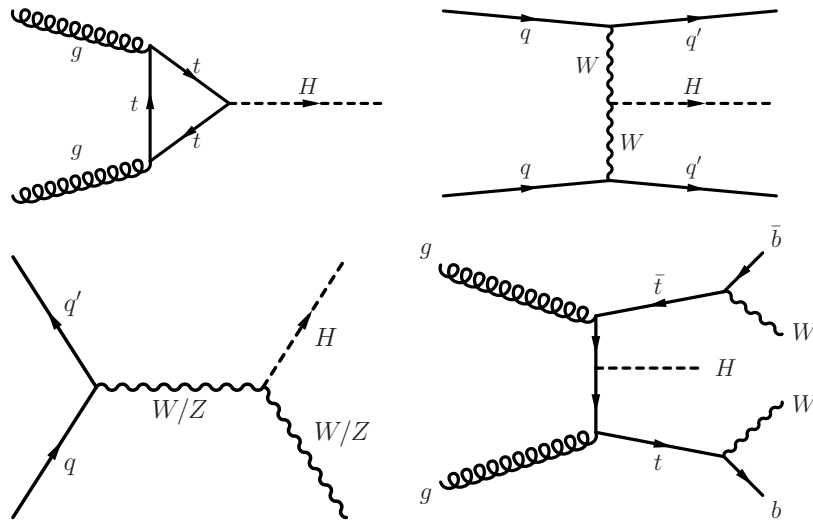
The analysis presented in this thesis focuses on the  $\gamma\gamma$  decay channel.

$H \rightarrow b\bar{b}$	$H \rightarrow WW$	$H \rightarrow gg$	$H \rightarrow ZZ$	$H \rightarrow c\bar{c}$	$H \rightarrow \gamma\gamma$	$H \rightarrow Z\gamma$
0.58	0.21	0.086	0.026	0.029	0.023	0.015

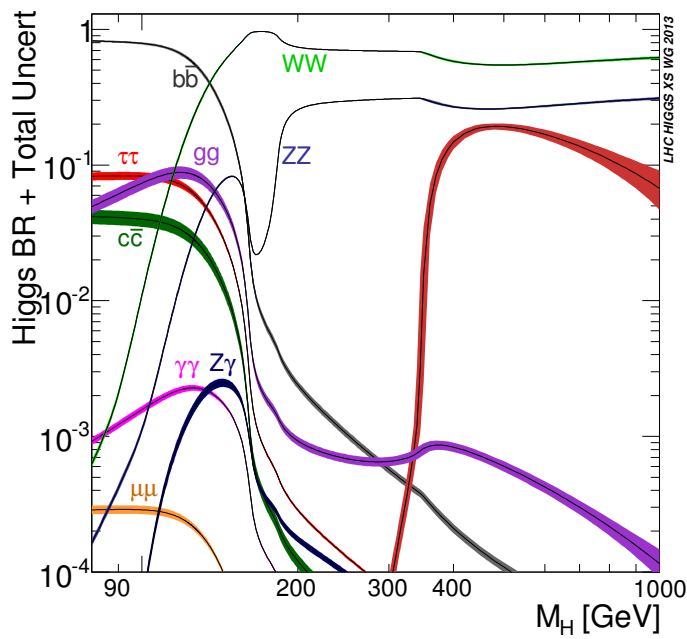
**Table 1.1.** Branching ratios for a 125 GeV Standard Model Higgs boson

### 1.2.1 The diphoton Decay Channel

Photons are massless particles, so they do not directly couple to the Higgs boson. However, the  $H \rightarrow \gamma\gamma$  decay can still happen through loop processes involving massive particles. The dominant Feynman diagrams at the lowest possible order are



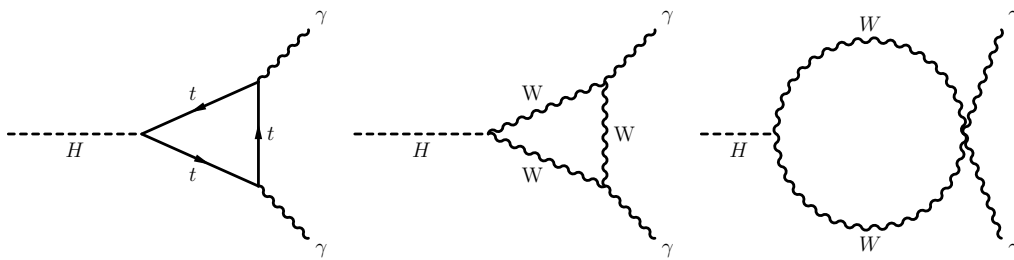
**Figure 1.2.** Feynman diagrams for Higgs boson production: ggH (top left) VBF (top right), VH (bottom left) and ttH (bottom right).



**Figure 1.3.** Standard Model Higgs boson decay branching ratios

shown in Figure 1.4. Due to their large coupling to the Higgs boson, the Feynman diagrams involving top and  $W$  loops are the most relevant.

Since it is a higher order process, the branching ratio is low compared to other



**Figure 1.4.** Main decay mechanisms of the Higgs boson in the diphoton channel

type of decays. Using perturbation theory the decay width can be computed:

$$\Gamma(H \rightarrow \gamma\gamma) \sim \frac{G_F \alpha_{em}^2 m_H^3}{128\pi^3 \sqrt{2}} \left| 7 - \frac{16}{9} + \dots \right|^2 \quad (1.3)$$

where  $G_F$  is the Fermi constant and  $\alpha_{em}$  the fine-structure constant. From 1.3 the branching ratio can be obtained. For  $m_H=125$  GeV it is:

$$\mathcal{B}(H \rightarrow \gamma\gamma) = 2.28 \times 10^{-3}$$

## 1.3 Experimental Searches of the Higgs Boson

### 1.3.1 Searches for the Higgs boson at LEP and Tevatron

The Large Electron-Positron Collider (LEP) was a circular collider operating at CERN laboratories in Geneva from 1989 to 2000. At  $e^+e^-$  colliders the main production mechanism for the Higgs boson is the Higgs-strahlung,  $e^+e^- \rightarrow HZ$ . During the LEP 1 phase, when the collider was operating at  $\sqrt{s}$  close to the mass of the  $Z$  boson, the data collected by the four experiments (ALEPH, DELPHI, L3 and OPAL) allowed to set a lower bound on the Higgs boson mass up to 65 GeV [14].

Precision electroweak measurements, on the  $Z$  boson at LEP, favored a relatively low mass for the Higgs boson [15].

During the LEP 2 phase the collision energy was raised up to 209 GeV and all possible decay modes were used to discover a SM Higgs boson. The combination of all the LEP data yields a 95% C.L. exclusion of a Higgs boson with a mass up to 114.4 GeV [10]. A small excess in data was seen around 115 GeV, but not sufficient to claim for a discovery, therefore LEP was shut down to start the construction of the LHC.

The Tevatron [16] was a circular  $p\bar{p}$  accelerator and collider operating at the Fermi National Accelerator Laboratory, in the United States. It operated from 1987 to 2011 colliding particles at  $\sqrt{s} = 1.96$  TeV. Unlike LEP, since the Tevatron was a hadronic collider, the main production processes for the Higgs boson were gluon-gluon fusion and, secondly, associate production with  $W$  or  $Z$  bosons. All searches, targeting different production modes and decays, from CDF and D0 experiments were combined together to reach better sensitivity. The last results of Tevatron

exclude two ranges of the Higgs mass: from 100 to 106 GeV and between 147 and 179 GeV. An excess of data, coming mainly from  $H \rightarrow b\bar{b}$  searches, was seen in the range 115-135 GeV, with a local significance of  $\sim 3\sigma$  [17].

### 1.3.2 The Higgs Boson Discovery at LHC

On the 4<sup>th</sup> July 2012 the ATLAS and CMS experiments announced, in a seminar at CERN, the discovery of a new particle, with a mass around 125 GeV, compatible with the SM Higgs boson. The discovery was obtained analyzing the datasets collected at 7 and 8 TeV by ATLAS and CMS, exploiting different experimental and analysis techniques. Among the different decay channels the discovery was driven by the high-sensitivity ones, i.e.  $H \rightarrow ZZ \rightarrow 4l$  and  $H \rightarrow \gamma\gamma$ . Indeed, these channels play a major role for a light Higgs, since they give the possibility to reconstruct the mass. To achieve the sensitivity needed to discover this new particle also other decay channels, i.e.  $bb$ ,  $WW$  and  $\tau\tau$  were used in the combination.

Since then, the two collaborations released different publications investigating in more detail the properties of the newly discovered particle. Some of these analyses (mainly those related to couplings measurements) will be discussed in details in this thesis. Until now no striking deviation from SM predictions has been highlighted. The most updated and precise results for the Higgs boson mass, for ATLAS and CMS are at the moment:

$$m_H = 125.36 \pm 0.37(stat) \pm 0.18(syst)$$

$$m_H = 125.03^{+0.26}_{-0.27}(stat)^{+0.13}_{-0.15}(syst)$$

## 1.4 Standard Model limitations

After the discovery of the Higgs boson and the precise measurement of its mass, all the ingredients of the Standard Model theory have been experimentally established. However the Standard Model is far from being a complete theory describing Nature for some intrinsic limitations, which have to be solved if one wants to extend the SM into a “Theory of everything”. As already said, gravity is not included in a coherent theoretical framework with other forces. Moreover, the SM does not give an explanation for important physics phenomena experimentally observed like dark matter, dark energy and the large difference among the values of the masses of particles, and of the coupling constants.

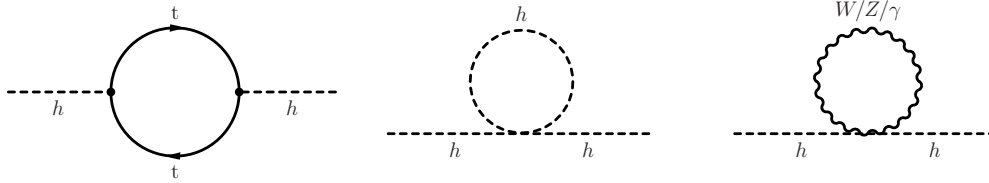
One of these unsolved questions, the so-called hierarchy problem, is just related to the Higgs sector, and specifically to the Higgs boson mass. All particles in Standard Model, including the Higgs boson, have a *bare mass* which is the mass obtained from the quantum propagator at the lowest order. This is not anyway the *physical mass*, i.e. the mass that can be measured experimentally, because radiative corrections at higher orders coming from loops have to be considered. The renormalization process relates the properties of the physical quantities (mass, charge...) to those of the bare particles, introducing suitable cut-off parameters in considering higher order corrections.

Without going into details, the renormalization procedure corrects the squared bare mass ( $m_0$ ) of the Higgs boson with an extra term, including higher order corrections,  $\delta m_H^2$  to obtain the physical mass  $m_H$ :

$$m_H^2 = m_0^2 - \delta m_H^2 \quad (1.4)$$

$\delta m_H^2$  includes all contributions from radiative corrections to the H propagator. The main one for Higgs boson are those involving top quarks, vector bosons and the Higgs itself.

The corresponding Feynman diagrams are shown in figure 1.5.



**Figure 1.5.** Main divergent contributions to the Higgs mass predicted by the Standard Model

The integrals corresponding to the amplitude of these processes are divergent, so a cutoff parameter  $\Lambda$  is introduced. This parameter represents the energy up to which the Standard Model can be considered still valid. In principle, one can assume that the Standard Model is valid up to the Planck scale; indeed that is the energy at which gravitational effects cannot be neglected. With this assumption  $\Lambda$  is of the order of  $\sim 10^{19}$  GeV. The full calculation gives that  $\delta m_H^2$  is proportional to  $\Lambda^2$  which means:

$$\delta m_H^2 \propto \Lambda^2 \sim 10^{38} \text{ GeV}^2$$

Since  $m_H$  is 125 GeV ( $\sim 10^2$ ) equation 1.4 can be rewritten like:

$$10^4 \text{ GeV}^2 \sim m_0^2 - \Lambda^2 \sim m_0^2 - 10^{38} \text{ GeV}^2$$

which means that  $m_0^2$  is of the same order of  $\Lambda^2$  ( $10^{38}$ ) and that these two terms cancel with a very high precision to obtain the value of the physical mass. This mathematical problem, known as *fine tuning*, does not invalidate the theory, which is still consistent. Anyway it seems an unnatural and implausible coincidence that  $m_0^2$  cancels all the loop contributions up to this astonishing precision.

The choice of  $\Lambda$  made in the previous calculation is somehow arbitrary because it is based on the assumption that the Standard Model is still valid up to the greatest possible energy, the Planck scale. If a lower  $\Lambda$  is chosen, the cancellation is tuned to an acceptable level. If, for instance,  $\Lambda \sim 1$  TeV is chosen, the hierarchy problem is completely solved since the cancellation is of the order of one over ten, which seems a natural and acceptable value.

In some way if one accepts the fine-tuning argument, New Physics phenomena at the TeV scale are expected, since at energy higher than  $\Lambda$  the Standard Model is not valid anymore. There are several models, beyond Standard Model (BSM), that



introduce new heavy particles around the TeV scale to solve the hierarchy problem. The Higgs sector can be seen, therefore, as a portal for accessing new unknown heavy particles.

## 1.5 New Physics in the Higgs Sector

Since the SM is verified experimentally at a great precision level, any new theory considers it as a starting point to develop further models. Given the theoretical limitations of the Standard Model, many theories have been built, in the last thirty years, to include the SM theory in a wider and more general framework. Now that a Higgs boson has been discovered, particular attention should be given to theories predicting new effects involving this particle. These models are very different in the implementation but, generally speaking, they predict the presence of new fermions or bosons, solving the hierarchy problem. These new particles, entering in the loops involving the Higgs production and decay, balance the divergent radiative corrections to the Higgs mass. In this way, a more natural theory without the need of fine tuning is possible. The interaction of these new particles would lead to a modification of the Higgs properties (cross section, branching ratios, total width...) with respect to the SM prediction.

The goal of this thesis is to exploit the interplay between direct and indirect searches to discover or exclude new physics phenomena. Some theories, indeed, predict new states already accessible by the experiments; a direct search is therefore needed to confirm or exclude their existence. On the other hand indirect constraints, obtained measuring possible deviation from the SM prediction, are important to reduce the available phase space for theoretical models. We focus on the diphoton decay channel, one of the most important for the discovery, which plays a major role also in many BSM scenarios. The approach in these experimental searches is as model independent as possible, focussing on measuring quantities which can be interpreted in a meaningful way in different classes of theories. We give here a brief review of the main classes of models that can take advantage from the direct and indirect searches presented in this thesis.

### 1.5.1 The effective Lagrangian for BSM models

In the theoretical models that extend the SM, it is considered as an effective theory, applicable up to a certain scale  $\Lambda$ . A new field theory above that scale is therefore needed. Given the lack of direct observation of new states, that could favor a specific theoretical model with respect to another one, a general approach can be held to describe Higgs properties in a BSM framework.

Since the predictions of the SM Lagrangian are verified experimentally at a great precision, any BSM field theory should reduce to the SM at low energy and therefore its gauge group should contain  $SU(3) \times SU(2) \times U(1)$ . The most general  $SU(3) \times SU(2) \times U(1)$ -invariant Lagrangian can be written expanding in higher-dimensional operators suppressed by powers of  $\Lambda$  [18, 19] as in equation 1.5

$$\mathcal{L} = \mathcal{L}_{SM}^{(4)} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} = \mathcal{L}_{SM}^{(4)} + \frac{1}{\Lambda} \sum_i c_i^{(5)} O_i^{(5)} + \frac{1}{\Lambda^2} \sum_i c_i^{(6)} O_i^{(6)} + \mathcal{O}\left(\frac{1}{\Lambda^3}\right) \quad (1.5)$$

where  $\mathcal{L}_{SM}$  is the Standard Model Lagrangian, described in section 1.1, which contains operators of dimension 4 or less. The additional terms  $\mathcal{L}^{(n)}$  contain dimension- $n$  operators ( $O_i^n$ ), with coupling constants  $c_i$ . They parametrize possible deviations from the SM caused by New Physics phenomena. No assumptions on the form of New Physics is made and contributions from operators with a dimension higher than 6 are neglected since not relevant, given the current experimental precision. This model only assumes that the newly discovered particle around 125 GeV is a CP-even weak doublet and it preserves baryon and lepton numbers. A full description of the effective Lagrangian can be found in [19], only relevant operators modifying the phenomenology of the Higgs sector are described here.

$\mathcal{L}^{(5)}$  is not relevant for the Higgs physics, since related to neutrino masses.  $\mathcal{L}^{(6)}$ , instead, contains 6-dimensional operators that could induce modifications in the Higgs coupling potentially accessible at experiments. Instead of considering  $\mathcal{L}^{(6)}$  in its entirety, further simplifications can be made ignoring operators inducing big deviations strongly disfavoured by previous experimental constraints or inducing very small deviations, not detectable by the LHC experiments. The simplified Lagrangian involving relevant operators for the Higgs sector can be therefore written as:

$$\begin{aligned} \mathcal{L}_{sim}^{(6)} = & \frac{h}{v} (2c_v m_W^2 W_\mu^+ W^{\mu-} + c_v m_Z^2 Z_\mu Z^\mu \\ & - c_u \sum_{q=u,c,t} m_q \bar{q}q - c_d \sum_{q=d,s,b} m_q \bar{q}q - c_l \sum_{q=e,\mu,\tau} m_l \bar{l}l \\ & + \frac{1}{4} c_{gg} G_{\mu\nu}^\alpha G^{\alpha\mu\nu} - \frac{1}{4} c_{\gamma\gamma} \gamma_{\mu\nu} \gamma^{\mu\nu} - \frac{1}{2} c_{WW} W_{\mu\nu}^+ W^{-\mu\nu} \\ & - \frac{1}{4} c_{ZZ} Z_{\mu\nu} Z^{\mu\nu} - \frac{1}{2} c_{Z\gamma} \gamma_{\mu\nu} Z^{\mu\nu}) \end{aligned} \quad (1.6)$$

where  $h$  is the higgs field,  $v$  its vacuum expectation value and the 9  $c$  parameters represent the couplings for the different processes considered. If custodial symmetry is imposed, the  $c_{WW}$  and  $c_{ZZ}$  are constrained by  $c_{\gamma\gamma}$  and  $c_{Z\gamma}$  and therefore  $\mathcal{L}_{sim}^{(6)}$  contains 7 free parameters. The SM Lagrangian for the Higgs boson is reobtained in the limit:

$$c_v = c_u = c_d = c_l = 1; c_{gg} = c_{\gamma\gamma} = c_{Z\gamma} = 0$$

In Section 1.1 it has been shown that in the SM the mass of fermions and bosons is the consequence of the interaction with the Higgs field and that the value of the couplings of the Higgs boson to other particles can be predicted accurately in the Standard Model. Their magnitude, proportional to the masses of the particles, directly influences the production and decay rates of the Higgs boson. Equation 1.6 shows that, in this effective modelling, if new physics exists it should be observable as modified couplings of the Higgs boson. Some of the most important theoretical BSM models involving the Higgs boson will be reviewed in the next section.

Measuring the couplings at experiments is, therefore, of particular interest to prove or deny SM theoretical predictions. Additional properties of the Higgs boson, like spin or parity, can be tested to probe the existence of new physics. In this thesis we will focus only on couplings. In the next chapter we will present a general parametrization of Higgs couplings, reflecting the versatility of the effective approach for BSM physics presented here.

### 1.5.2 BSM models: 2HDM and VLQ

The discovery of a Higgs boson is a good opportunity to search for a possible extended EWSB sector. Indeed, several theoretical models addressing the hierarchy problem, like Supersymmetry [20] or composite Higgs models [21, 22] predict the existence of additional Higgs scalars.

The underlying idea of some of these models is to add a second Higgs doublet to the one predicted by the SM. They are therefore known as 2 Higgs doublet models (2HDM). With two doublets the Higgs potential  $V(\Phi)$  in equation 1.1 becomes:

$$\begin{aligned}
 V(\Phi_1, \Phi_2) = & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - [m_{12}^2 \Phi_1^\dagger \Phi_2 + \text{h.c.}] \\
 & + \frac{1}{2} \lambda_1 (\Phi_1^\dagger \Phi_1)^2 + \frac{1}{2} \lambda_2 (\Phi_2^\dagger \Phi_2)^2 + \lambda_3 (\Phi_1^\dagger \Phi_1) (\Phi_2^\dagger \Phi_2) + \lambda_4 (\Phi_1^\dagger \Phi_2) (\Phi_2^\dagger \Phi_1) \\
 & + \left[ \frac{1}{2} \lambda_5 (\Phi_1^\dagger \Phi_2)^2 + \lambda_6 (\Phi_1^\dagger \Phi_1) (\Phi_1^\dagger \Phi_2) + \lambda_7 (\Phi_2^\dagger \Phi_2) (\Phi_1^\dagger \Phi_2) + \text{h.c.} \right]
 \end{aligned} \tag{1.7}$$

where  $\Phi_1$  and  $\Phi_2$  are the two Higgs doublets. The minimization of this potential leads to five physical states: two neutral CP-even scalar ( $h$  and  $H$ ), one neutral CP-odd pseudoscalar ( $A$ ) and two charged scalars ( $H^+$  and  $H^-$ ). All the different couplings  $\lambda_i$  can be expressed as a function of two quantities: the ratio of the vacuum expectation values of  $\Phi_{1,2}^0$  parametrized by  $\tan \beta = |\langle \Phi_2^0 \rangle / \langle \Phi_1^0 \rangle|$  and  $\alpha$ , which is the mixing angle between  $h$  and  $H$ . The angles  $\alpha$  and  $\beta$  completely determine the couplings of a Higgs boson to a pair of fermions or gauge bosons, as well as the couplings between two Higgses and a single gauge boson.

In the so-called *alignment limit*, i.e.  $\cos(\alpha - \beta) = 0$ , the couplings of the lightest Higgs boson  $h$  are exactly the same as the SM Higgs. Identifying  $h$  with the observed SM Higgs, deviations in production and decay rates of the SM-like Higgs boson, can be directly parametrized in terms of  $\alpha$  and  $\beta$ . Therefore, limits on the masses of the four heavier Higgs bosons can be inferred.

Depending on different assumptions on couplings between fermions and the two doublets, 2HDM can be divided in four types:

- Type I: all fermions couple to one doublet
- Type II: up-type quarks couple to one doublet while down type quarks and leptons to the other one
- Type III: quarks couple to one doublet and leptons to the other one
- Type IV: up-type quarks and leptons couple to one doublet while down-type quarks to the other one

For instance, Minimal Supersymmetric Standard Model (MSSM) is a theoretical model of type II. Typically signals of type III and IV resemble respectively to those of type I and II. The phenomenological panorama is quite different depending on theoretical assumptions. However, an accurate measurement, with the current dataset, could constrain 2HDM models at a level of 20% of deviation from the alignment limit for both type I and type II models. This will be of paramount

importance to suggest which may be the most probable new phenomena to be directly observed at higher energy collisions.

Another class of models, called here vector-like quarks (VLQ) models, instead, requires the presence of fermionic partners of the top quark. In this kind of theories the radiative corrections of the top quark to the Higgs mass are balanced by contributions of new fermions with the same spin of the top but with vectorial couplings. This means that the left- and right-handed chiralities of the top partners have the same color and electroweak quantum numbers.

The most studied scenarios predicting VLQ can be divided in three classes:

- *Composite Higgs* [23,24]: QCD light scalars, like pions and kaons, are predicted without problems of fine tuning, because, instead of being elementary like the SM Higgs, they are composite states. Indeed, pions are pseudo-Goldstone bosons, emerging from spontaneous breaking of the chiral symmetry of QCD by quark condensation. Assuming the existence of a new, QCD-like, strong sector at the TeV scale, the Higgs boson can be seen as the pseudo-Goldstone boson of this new spontaneously broken symmetry. The compositeness of the Higgs would therefore solve elegantly the hierarchy problem.
- *Little Higgs* [25–27]: These models are quite similar to the composite Higgs ones. The peculiarity of these models is that the Higgs mass is protected by the existence of a global additional symmetry explicitly broken. The Higgs is the Goldstone boson related to this symmetry and its mass emerges from the explicit symmetry breaking.
- *Extra dimensions* [28]: These models predict the existence of additional dimensions. The difference between the Higgs mass and the Planck scale is therefore explained immersing the Higgs in this higher dimensional space, where the divergences of the Higgs boson mass corrections are less severe.

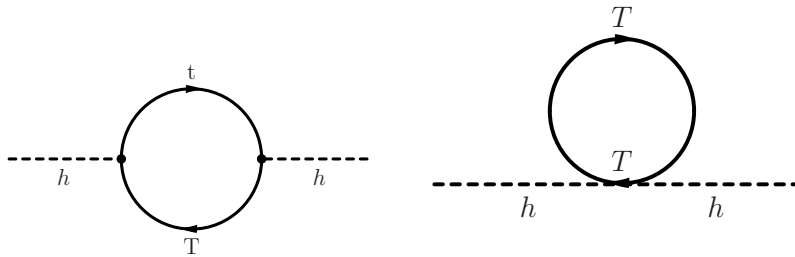
Regardless the different theoretical assumption, the top partner contribution can be included in the effective Lagrangian with additional terms of the form:

$$\mathcal{L}_T = \lambda_t H t \bar{t} + \lambda_T H T \bar{t} + \frac{\lambda'_T}{2m_T} H H \bar{T} T + h.c.$$

where  $H$  is the Higgs boson field,  $T$  is the generic top partner field and  $t$  is the SM top quark field. These terms cause corrective radiations to the Higgs mass, with loops involving new particles, like those shown in the diagrams of figure 1.6. The couplings of the interactions are such that the new loop diagrams lead to a natural cancellation of divergences that stabilizes the Higgs boson mass. The mass of top partners is predicted to be not so heavy, below the TeV, therefore direct searches at LHC could spot their existence, with the current amount of data.

## 1.6 Coupling structure of a Higgs-like particle

As seen in Paragraph 1.5.1, a disagreement between theory and experiments for the Higgs couplings could be a hint of physics beyond the SM, signalling the presence of new particles, new gauge symmetries or a different mechanism to give mass to



**Figure 1.6.** Additional radiative corrections to the Higgs boson mass predicted by VLQ theories

bosons or fermions. Both the Higgs boson production cross section and branching ratios are sensitive to its couplings to fermions and bosons.

In this thesis a phenomenological effective framework, suitable for the effective Lagrangian description, will be used to explore the coupling structure. This approach has the advantage to define observables directly interpretable as deviations from the predicted SM values without assuming a specific alternative framework.

In particular, the method used in this thesis, described in detail in [29], computes all deviations from the SM predictions assuming that there is a SM-like Higgs boson of 125 GeV which spontaneously breaks electroweak symmetry. No additional assumptions on new BSM states, such as extra Higgs bosons or families of fermions, are made. Starting from the picture given by past experimental results three additional simplification can be made without loss of generality:

- The signals observed in different search channels originate from a single narrow resonance with a mass near 125 GeV. This excludes the possibility of overlapping resonances in this mass region.
- The width of the 125 GeV Higgs boson is neglected, since assumed to be not significantly different from the SM one (few MeV). Hence the predicted signal cross-section times branching ratio can be decomposed in:

$$(\sigma \cdot \mathcal{B})(ii \rightarrow H \rightarrow ff) = \frac{\sigma_{ii} \cdot \Gamma_{ff}}{\Gamma_H}$$

where  $\sigma_{ii}$  is the production cross section through the initial state  $ii$ ,  $\Gamma_{ff}$  the partial decay width into the final state  $ff$  and  $\Gamma_H$  the total width of the Higgs boson.

This assumption, which is an excellent approximation for a light SM-like Higgs, is expected to be valid also for a wide range of BSM models compatible with current experimental results.

- The tensor structure of the SM is preserved. This implies, in particular, that the observed state is CP-even.

With these assumptions, scale factors can be defined to take into account possible deviations from the expected values. The cross sections and partial decay widths

are dressed with scale factors  $k_i$ . They are defined in such a way that  $\sigma_{ii}$  and  $\Gamma_{ii}$ , associated to the SM particle  $i$ , scale with a factor  $k_i^2$  with respect to the corresponding SM prediction.

Using this scheme, for example, the cross section for the process  $gg \rightarrow H \rightarrow \gamma\gamma$  can be written as:

$$(\sigma \cdot \mathcal{B})(gg \rightarrow H \rightarrow \gamma\gamma) = \sigma_{SM} \cdot \mathcal{B}_{SM}(H \rightarrow \gamma\gamma) \cdot \frac{k_g^2 \cdot k_\gamma^2}{k_H^2}$$

These scale factors in some cases should be intended as effective couplings which describe loop-induced processes, such as  $k_\gamma^2$  for the  $H \rightarrow \gamma\gamma$  process and  $k_g^2$  for  $gg \rightarrow H$ . They can be expressed (not considering possible new BSM particles contributing to the loop) as a function of more elementary quantities, as in these relationships:

$$k_g^2 = \frac{k_t^2 \cdot \sigma_{ggH}^{tt} + k_b^2 \cdot \sigma_{ggH}^{bb} + k_t k_b \sigma_{ggH}^{tb}}{\sigma_{ggH}^{tt} + \sigma_{ggH}^{bb} + \sigma_{ggH}^{tb}}$$

$$k_\gamma^2 = \frac{\sum_{i,j} k_i k_j \cdot \Gamma_{\gamma\gamma}^{ij}}{\sum_{i,j} \Gamma_{\gamma\gamma}^{ij}}$$

where  $\sigma_{ggH}^{ij}$  and  $\Gamma_{\gamma\gamma}^{ij}$  are the SM values, computable from theory [30–32] and the indexes  $i, j$  in the second equation correspond to all relevant contributions in the  $H \rightarrow \gamma\gamma$  loop, i.e.  $b, t, \tau$  and  $W$ . With this parametrization SM predictions are recovered for  $k_i = 1$ .

Results on different quantities can be extracted from fits to the data, using statistical methods such as likelihood ratio, where the  $k_j$  are treated as parameters of interest or nuisance parameters, depending on the case. This effective simplified model is general and versatile and allows to perform different benchmark parametrizations to investigate several aspects of the Higgs boson nature.

### 1.6.1 Benchmark parametrizations

The simplest interesting quantity to measure is the overall signal strength, usually called  $\mu$ , defined as  $\mu = \frac{\sigma}{\sigma_{SM}}$ . In the simplified model described in the previous section, this corresponds to use a single scale factor  $k^2$ , that applies to all production and decay modes. Regarding the individual production and decay scale factors to test, the current amount of data is not sufficient to measure all of them individually. Anyway, a set of useful benchmark models to test particular aspects of the Higgs properties can be provided and are listed in the following. Experimental results for these quantities will be given in chapter 4 and 5.

#### Signal strength splitting production processes

Different production modes can be tested separately. The four main Higgs boson production mechanisms (described in section 1.2) can be associated with either fermion couplings ( $ggH$  and  $t\bar{t}H$ ) or vector bosons couplings (VBF and VH). Using dedicated analyses targeting different production modes (see chapter 4) they can be decoupled measuring strength modifiers for these two different kind of processes:  $\mu_{ggH+t\bar{t}H}$  and  $\mu_{VBF+VH}$ .

## Couplings to vector bosons and fermions

An interesting benchmark model can be built simply considering a single scale factor,  $k_f$ , for coupling to fermions and another one,  $k_V$ , for coupling to bosons. The  $H \rightarrow \gamma\gamma$  width ( $\Gamma_{\gamma\gamma}$ ) is the only sensitive at leading order to the relative sign of  $k_f$  and  $k_V$ ; all other partial widths scale as  $k_f^2$  or  $k_V^2$ .

## BSM particles

The presence of new particles of high mass, even if not directly observed, can induce modifications in the cross-sections and branching ratios. In particular, BSM particles usually can enter in loops for production and decay processes, such as  $gg \rightarrow H$  and  $H \rightarrow \gamma\gamma$ . The two interesting parameters in this benchmark model will be, therefore, the effective scale factors  $k_g$  and  $k_\gamma$ .

## Coupling to top quarks

The top-Higgs coupling is an interesting quantity for theoretical reasons. Indeed, it is widely believed that the large mass of the top quark could be a hint of a special role of this particle in the EWSB mechanism. Moreover, as seen in previous paragraphs, states involving top quarks are the most promising in BSM searches.

The more straightforward way to measure this coupling in the SM is to study the associated production of a Higgs boson and a top pair ( $t\bar{t}H$ ). This production mechanism is sensitive only to the magnitude of the top-Higgs coupling but not to its sign since the possible diagrams at tree level are of the same kind of the one shown in figure 1.2 (bottom right). On the other hand, studying single top associate production ( $tHq$ ) in the diphoton decay channel allows to test more exotic scenarios with negative top-Higgs coupling. In this thesis both processes are used; more details about  $t\bar{t}H$  and  $tHq$  analysis will be given, respectively, in chapter 4 and 5.

## 1.7 Analysis strategy for BSM searches in $H \rightarrow \gamma\gamma$ channel

As explained in this chapter, the Higgs sector is an interesting field to search for new physics. The aim of this thesis is to shed light on the Higgs sector following the two paths of direct and indirect BSM particles searches to fully exploit CMS discovery potential. The diphoton decay channel is promising for both kinds of searches, given the possibility to fully reconstruct the final state and the good signal over background ratio.

Regarding indirect searches, possible deviations of the Higgs boson couplings from the values predicted by the SM are investigated. The results can be interpreted in different classes of BSM models, such as the 2HDM or VLQ scenarios described here. For this purpose, an exclusive analysis for SM Higgs boson is presented in chapter 4. Particular attention is given to the description of exclusive categories targeting different production modes, exploiting the sensitivity to different couplings.

Regarding direct searches, associate production of a single top quark and a Higgs boson and vector-like partners decaying in a top quark and a Higgs are studied.

These searches are described in chapter 5. They both use tools developed for the SM analysis (especially for categories tagging  $t\bar{t}H$  production) even if a different optimization is devised according to the different final state.



## Chapter 2

# The CMS Experiment at the LHC

This chapter is dedicated to the description of the Compact Muon Solenoid (CMS), the experimental apparatus used for all the physics searches presented in the following chapters. After a short introduction on the CERN LHC, all the subdetectors of CMS will be described, paying particular attention to the ones which are more relevant for the searches described in this thesis.

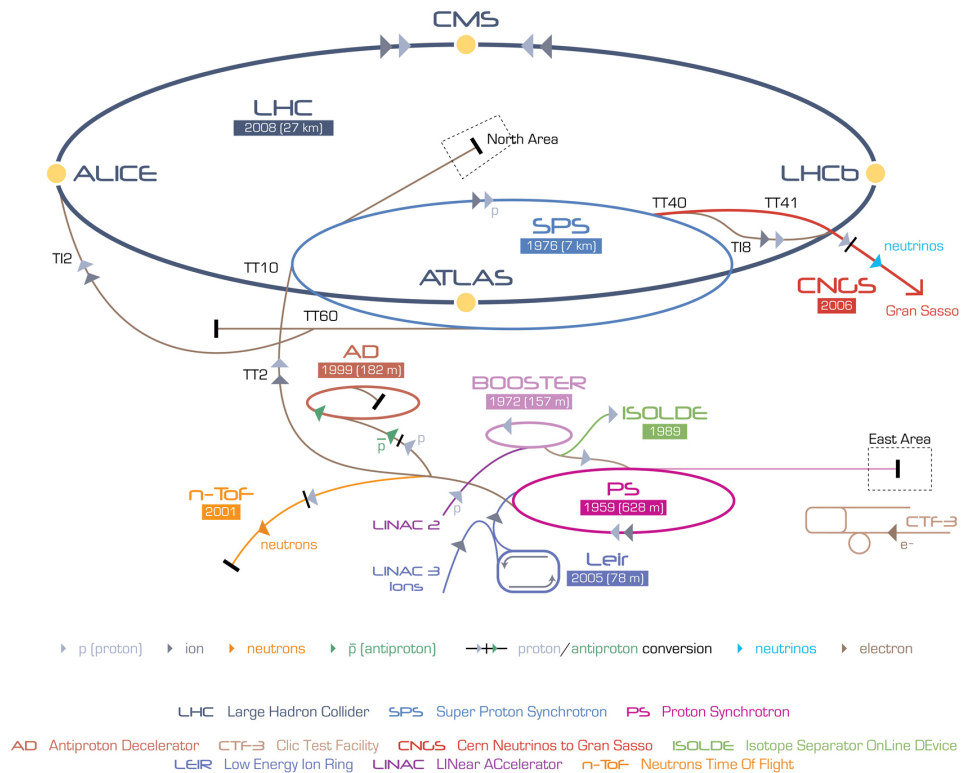
### 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most energetic particle accelerator. It is located at CERN (*European Organization for Nuclear Research*) laboratories in Geneva, Switzerland. It is the last of a series of linear and circular accelerators used to accelerate two high energy proton beams. The beams contain up to 2800 bunches of protons, with a nominal number of  $10^{11}$  protons each. They have a small transverse spread (about  $15 \mu\text{m}$ ) and are about 7.5 cm long in the beam direction. The minimal bunch time separation is 25 ns. The LHC is also used to collide beams of heavy ions, but this configuration is not relevant for this thesis and will not be described.

The whole CERN accelerator complex is shown in figure 2.1. Protons are extracted from a hydrogen source and accelerated up to 50 MeV in the linear collider LINAC2, then up to 1.4 GeV in the PS (*Proton Synchrotron*) and finally the SPS (*Super Proton Synchrotron*) injects them into the LHC at 450 GeV.

LHC consists of a 27 km long circular tunnel, previously hosting the LEP collider, about 100 meters underground. It is equipped with 1232 superconducting dipole magnets, placed in the curved sections of the tunnel, and boost structures. Since magnets operate in a superconducting state, a cooling system of liquid helium is used to cool them down to 1.9 K. LHC is capable to produce collisions at a center of mass energy of up to 14 TeV and a maximal instantaneous luminosity of  $10^{34}\text{cm}^{-2}\text{s}^{-1}$ . Acceleration cavities are located in one of the linear sections.

A summary of the main LHC technical parameters, with design values, is given in table 2.1. LHC delivered collisions at 7 TeV for the whole 2011 and at 8 TeV during 2012 with a 50 ns bunch time separation.



**Figure 2.1.** Schematic view of CERN accelerator scheme

The LHC can cross the two beams in four interaction points, where four main experiments are located:

- ALICE (*A Large Ion Collider Experiment*) [33]
- ATLAS (*A Toroidal LHC ApparatuS*) [34]
- CMS (*Compact Muon Solenoid*) [35]
- LHCb (*Large Hadron Collider beauty*) [36]

ATLAS and CMS are two high-luminosity, general purpose experiments, built to fully reconstruct the largest possible number of physics processes. ALICE is an experiment dedicated to the study of heavy ion collisions, while LHCb is specialized in b-quark physics.

The data presented in this thesis are collected by the Compact Muon Solenoid experiment and a complete description of it will be given in the following.

## 2.2 The CMS Experiment

CMS is a general purpose detector designed to reconstruct a wide range of particles and phenomena produced by collisions in the LHC. The Higgs boson search was one of the main physics goals and it drove the design philosophy of the detector [37].

**Table 2.1.** Summary of the main LHC technical parameters

Circumference [km]	27
Number of magnet dipoles	1232
Dipolar magnetic field [T]	8.33
Radiofrequency [MHz]	400
Maximal number of bunches	2808
Magnet temperature [K]	1.9
Maximal beam energy [TeV]	7
Maximal luminosity [ $\text{cm}^{-2}\text{s}^{-1}$ ]	$10^{34}$
Protons per bunch	$1.05 \cdot 10^{11}$
Bunch spacing [m]	7.48
Minimal bunch time separation [ns]	25
Bunch length [cm]	7.5
Bunch transverse size [ $\mu\text{m}$ ]	15
Crossing angle [rad]	$2 \cdot 10^{-4}$
Beam lifetime [h]	7
Luminosity lifetime [h]	10

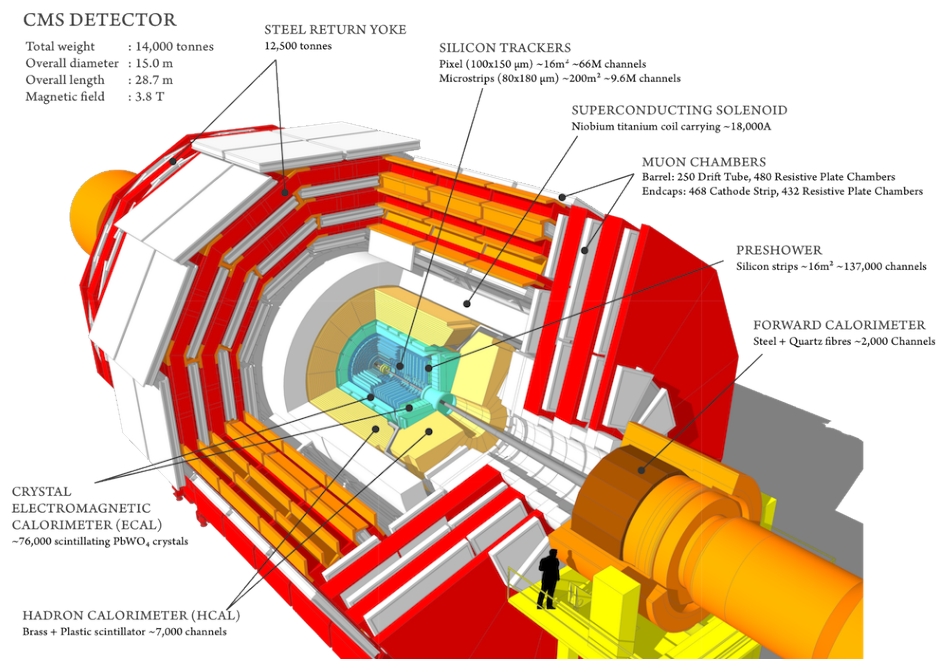
CMS consists of a detector made of different layers, exploiting different detection techniques to achieve the best energy and momentum measurement for each type of particle.

The main requirements in CMS design were:

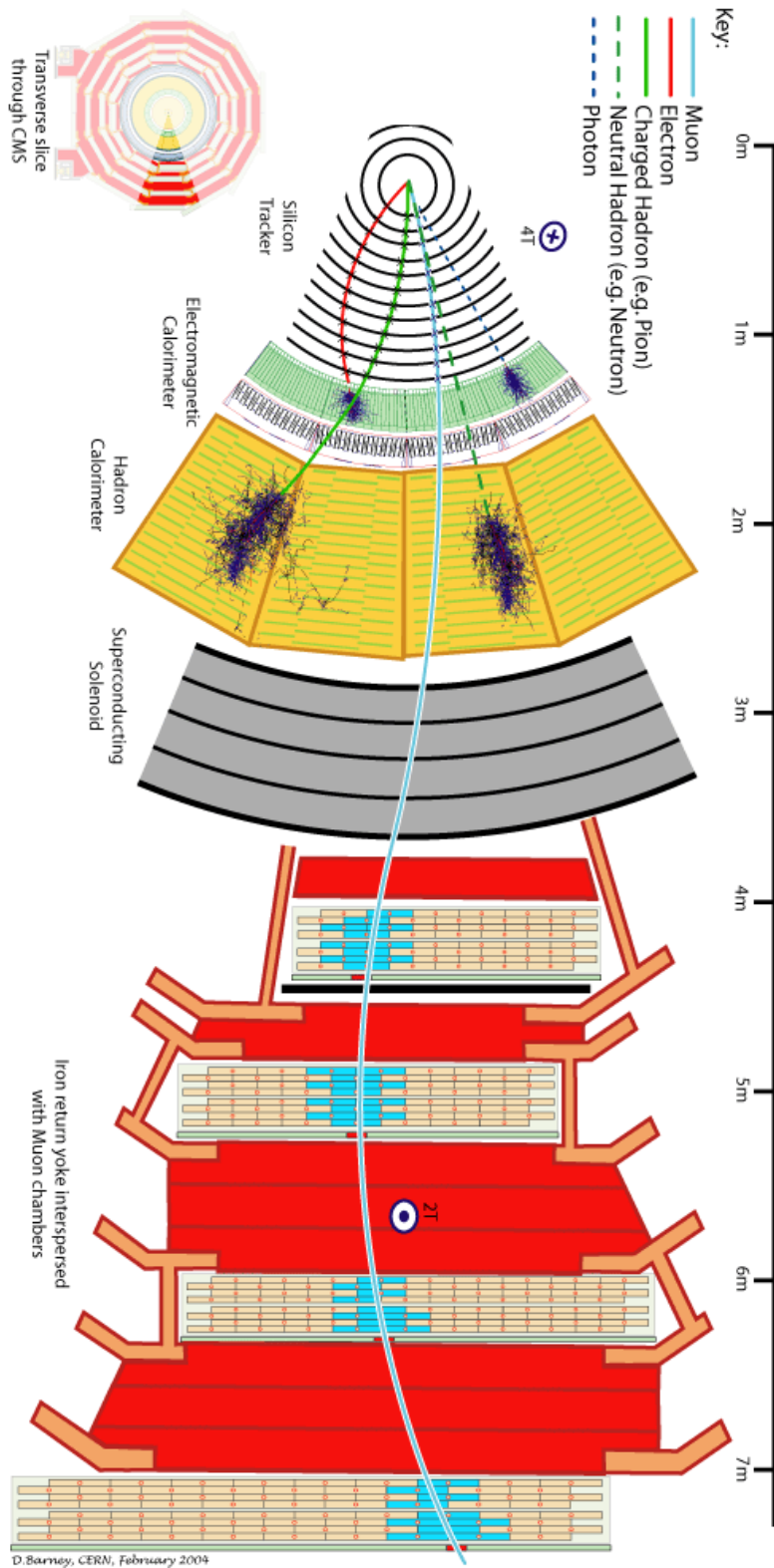
- a high performance muon system, capable to measure accurately the transverse momentum of this kind of particles up to 1 TeV. This has led to the presence of a large superconducting solenoidal magnet producing an intense magnetic field of 4T. This allows a compact muon spectrometer with high performances in track and charge measurement;
- a high resolution electromagnetic calorimeter to detect and measure electrons and photons, compatible with the presence of the magnet;
- an inner tracking system allowing accurate momentum measurements for charged particles;
- an hermetic hadronic calorimeter for good measurement of jets and missing transverse energy.

A section of the CMS detector is shown in Figure 2.2. It has a cylindrical structure, is 21 m long, 15 m wide and 15 m high, for a total weight of about 12500 tons. It is made of a central section, coaxial to the beam axis (the *barrel*), closed by two hermetic discs orthogonal to the beam (the *endcaps*).

A transverse slice of the detector is shown in figure 2.3. The detector is a sequence of different layers to stop, track or measure all different particles from the interaction point. Moving outwards from the beam, emerging particles find the silicon tracker, the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL), the superconducting solenoid and the iron return yoke interspersed with muon chambers.



**Figure 2.2.** Schematic view of the CMS detector



**Figure 2.3.** Schematic view of a transverse slice of CMS detector. The sketch reports traces left in subdetectors by different types of particles.

CMS uses a right-handed cartesian coordinate system, with the origin at the nominal interaction point at the center of the detector. The  $x$ -axis points to the centre of the LHC, the  $y$ -axis points upwards, perpendicular to the LHC plane, and the  $z$ -axis along the anticlockwise-beam direction.

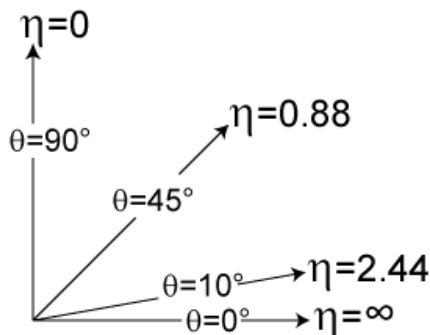
Given the cylindrical system, it is also useful to define some variables in a pseudo-angular reference system. The polar angle,  $\theta$ , is measured from the positive  $z$ -axis and the azimuthal angle,  $\phi$ , lies in the  $x$ - $y$  plane. Transverse energy and momentum ( $p_T$  and  $E_T$ ), i.e. the particle's energy ( $E$ ) and momentum ( $p$ ) in the transverse plane, are therefore defined as  $p_T = p \cdot \sin(\theta)$  and  $E_T = E \cdot \sin(\theta)$ . A frequently-used variable, at hadron colliders, is the *pseudorapidity*, defined as:

$$\eta = -\ln \tan \frac{\theta}{2}$$

For high energies, intervals of pseudorapidity are invariant under Lorentz boosts along the beam direction. Since in hadron colliders the boost of the colliding partons in the lab frame is unknown, pseudorapidity is preferred to the polar angle  $\theta$ . Values of pseudorapidity corresponding to given  $\theta$  values are reported in figure 2.4.

Distance between particles with two particles is usually computed using  $\Delta R$ , defined as the distance in the  $\phi$ - $\eta$  plane:  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ .

In the following a complete description of all subdetectors of CMS will be given paying particular attention to the ECAL which plays a major role in the searches described in this thesis.



**Figure 2.4.** Values of pseudorapidity in correspondence of several values of the  $\theta$  angle.

### 2.2.1 Magnet

A strong magnetic field in particle experiments is needed to bend particles, to achieve the best possible resolution in the muon detection system. The CMS choice was to have a relatively small solenoid producing an intense field. This innovative feature has conditioned the design of all other subdetectors.

The superconducting magnet [38] is a 13 m long cylinder of Niobium-Titanium, with a diameter of 5.9 m. It provides a uniform magnetic field of 3.8 T at its center, carrying a current of 18 kA. The magnet flux is returned via a 1.5 m thick saturated iron yoke which is also instrumented with four stations of muon chambers.

### 2.2.2 The Silicon Tracker

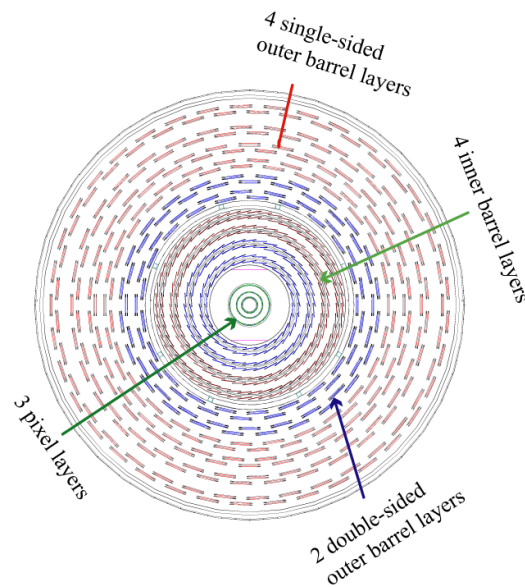
The silicon tracker design was driven by the necessity to track particles with a good precision in a high multiplicity, highly-radioactive environment.

It consists of a detector entirely made of silicon [39]. It is the first example in high-energy physics of a tracking system completely built using this technology. It allows to measure few points for each track (from 10 to 14) with an excellent resolution (about  $10 \mu\text{m}$ ).

The tracker system is composed of three regions, with lower granularity moving outwards from the interaction point:

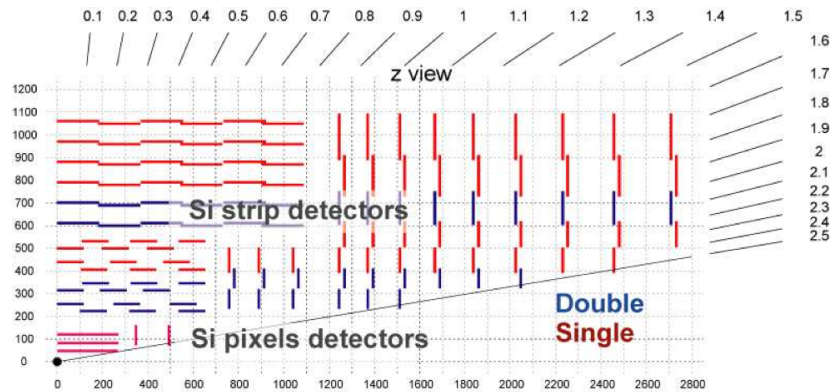
- Close to the interaction vertex, where the particle flux is the highest, a detector made of *pixels* is used. The size of a pixel is about  $100 \mu\text{m} \times 150 \mu\text{m}$ .
- The intermediate region is made of layers of *microstrips* with a minimum cell size of  $10 \text{ cm} \times 80 \mu\text{m}$ .
- In the outermost region the particle flux is low enough to adopt larger strips with a maximum cell size of  $25 \text{ cm} \times 80 \mu\text{m}$ .

Pixels are arranged in a 3-layer barrel, with 2-layer endcaps on each side. Silicon microstrips cover a very large volume up to the calorimetric system. There are approximately 10 million strips arranged in 10 cylindrical layers, in the barrel region, and in 9 disks in each of the two endcaps. In figure 2.5 a section of the barrel region is shown, highlighting the different layers. Figure 2.6 shows instead a longitudinal section of the tracking system.

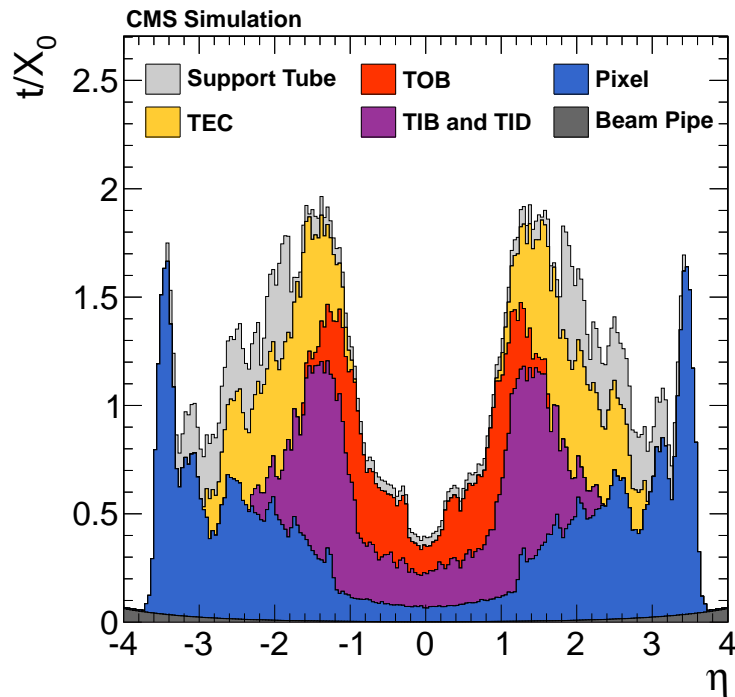


**Figure 2.5.** Transverse section of the tracker barrel.

The material budget of the tracker as a function of the pseudorapidity is shown in figure 2.7. The tracker adds up to less than half a radiation length in the center of the barrel, increasing to a maximum of about  $1.8 X_0$  around  $|\eta| = 1.5$ .



**Figure 2.6.** Schematic longitudinal section of one quarter of the CMS silicon detector. The nominal interaction point is at the bottom-left corner. Distances are marked in millimeters on the left and bottom axes, and pseudorapidity values are shown on the top and right borders.



**Figure 2.7.** Silicon tracker material budget as a function of pseudorapidity, expressed in units of radiation lengths ( $X_0$ ). Different material categories are shown: beam pipe, pixels, different parts of the strip system (TIB, TID, TOB and TEC) and support tubes.

### 2.2.3 The Electromagnetic Calorimeter

The CMS electromagnetic calorimeter (ECAL) has been designed in order to obtain an excellent resolution in the measurement of electrons or photons [40]. In particular,



the benchmark for its design was the  $H \rightarrow \gamma\gamma$  decay channel, fundamental for the discovery of a low-mass Higgs boson. The chosen solution was an hermetic, homogeneous electromagnetic calorimeter, made of 75848 lead tungstate scintillating crystals of lead tungstate ( $\text{PbWO}_4$ ).

### 2.2.3.1 The crystals

$\text{PbWO}_4$  characteristics are summarized in Table 2.2.

The choice of this crystal has been driven by several factors:

- its short radiation length allows the construction of a compact calorimeter. Given the limited space for the calorimetric system imposed by the magnet, this is of primary importance for CMS;
- its small Molière radius ensures lateral shower containment and, therefore, high granularity. High granularity is needed for  $\pi^0-\gamma$  separation and angular resolution;
- its very fast light emission process makes it suitable for the LHC where bunch crossings are interspaced by only 25-50 ns.

Parameter	Value
$X_0$	0.89 cm
$R_M$	2.2 cm
Light yield	100 $\gamma/\text{MeV}$
% of light emitted in 25 ns	80%

**Table 2.2.** Main characteristic of  $\text{PbWO}_4$  crystals.

The  $\text{PbWO}_4$  characteristics allow to build compact crystals: 25 cm of length correspond to 28  $X_0$  and the Molière radius is only 2.2 cm. A crystal of these dimensions ensures an excellent containment of the electromagnetic showers up to about 1.5 TeV. The light yield of  $\text{PbWO}_4$  is low compared to other scintillating materials and necessitates the use of specific photodetectors, suitable to operate in the high-intensity magnetic field.

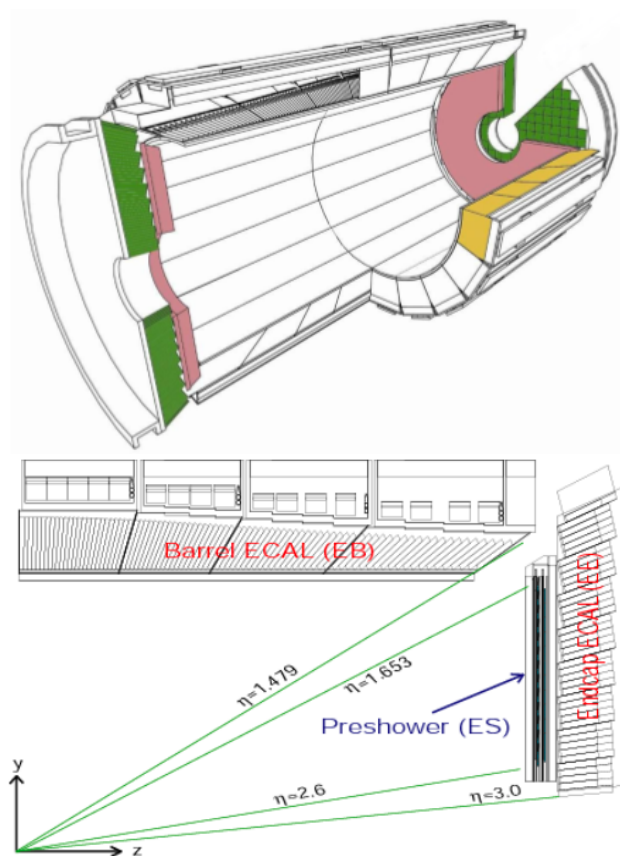
Avalanche photodiodes (APD) are used in the barrel. These photodetectors are made of semiconducting silicon with a strong electric field. When an electron is knocked out from an atom by scintillation light, it is accelerated in the electric field, knocking out electrons from the other atoms. Therefore, APDs are able to produce a very high signal in a short time and are suitable for the low light yield of lead tungstate. They are operated at a gain of 50.

In the endcaps, where radiation is too high for silicon photodiodes, vacuum photodiodes (VPT) are used. VPTs contain three electrodes within a vacuum tube: one of them releases an electron when light strikes, one works as anode producing several electrons that are all then accelerated to the third electrode (the dynode), releasing a second batch of electrons. As in the barrel case, the signal is amplified and digitized and sent along optic fibers to the upper level readout.

### 2.2.3.2 Geometry

A scheme of ECAL geometry is shown in figure 2.8. Roughly 80% of crystals (61200) are used in the central barrel (EB), covering the region  $|\eta| < 1.479$ . The barrel is made of 36 identical supermodules covering half of the barrel length. Crystals have a trapezoidal shape and a longitudinal dimension of  $25.8 X_0$ . Crystals are organized in a semi-projective geometry, forming a  $3^\circ$  angle with respect to the line that connects them to the nominal interaction point. This avoids that photons fall in the separation zone of two crystals. The barrel has a granularity of 360 crystals in the  $\phi$ -direction and  $2 \times 85$  crystals in the  $\eta$ -direction.

Two endcaps (EE) are placed at a distance of 3 m from the nominal interaction point, covering up to  $|\eta| = 3$ . They are made of crystals with a length of  $24.7 X_0$ . The endcaps are equipped with a preshower detector (ES), covering the region  $1.7 < |\eta| < 2.6$ . It is a two-layer sampling calorimeter made of lead and silicon strips detectors. The thickness of the two lead absorbers is respectively  $2 X_0$  and  $1 X_0$ . The preshower is used to obtain a better spatial resolution in the endcaps, needed to separate  $\pi^0$  from  $\gamma$ .



**Figure 2.8.** Schematic view of ECAL. On the top a tridimensional view is shown, different colors correspond to different parts: the barrel (yellow), the endcaps (green) and the preshower (pink). On the bottom a longitudinal section of a quarter of ECAL is shown.

### 2.2.3.3 Energy resolution

The energy resolution of a homogeneous calorimeter for electrons and photons of energy  $E$  can be parametrized as a squared sum of three terms:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.1)$$

where  $E$  is the energy expressed in GeV.

The stochastic term ( $S$ ) depends on fluctuations of the number of detected photons, the noise term ( $N$ ) is the term due to the electronics noise and the constant term ( $C$ ) depends on lateral containment, non uniformity of response and intercalibration.

The values of these parameters have been measured at a beam test [41] at a single-crystal level and were found to be:

$$S = 2.8\% \text{GeV}^{\frac{1}{2}} \quad N = 124 \text{ MeV} \quad C = 0.3\%$$

The energy resolution for photons with  $E_T \approx 60$  GeV varies between 1.1% and 2.6% over the solid angle of the ECAL barrel, and from 2.2% to 5% in the endcaps [42]. The ECAL energy resolution for electrons with  $E_T \approx 45$  GeV from  $Z \rightarrow ee$  decays is better than 2% in the central region of the ECAL barrel ( $|\eta| < 0.8$ ), and is between 2% and 5% elsewhere. For low-bremsstrahlung electrons, where 94% or more of their energy is contained within a  $3 \times 3$  array of crystals, the energy resolution improves to 1.5% for  $|\eta| < 0.8$  [42].

### 2.2.4 The Hadronic Calorimeter

The hadronic calorimeter (HCAL) [43] measures the energy of hadrons, such as pions and kaons. It is a crucial element in jet reconstruction, together with ECAL, and it is used to indirectly determine the energy of undetected particles, such as neutrinos.

The hadronic calorimeter is made of a central part (HCAL), covering the pseudorapidity range up to  $|\eta| < 3$ , and a forward part (HF) for the high pseudorapidity region ( $3 < |\eta| < 5$ ).

HCAL is a sampling calorimeter, with brass (organized in layers of 3.7 mm) used as absorber and plastic scintillators as active material. Special optic fibers are used to collect the light and photodetectors to amplify the signal. In HF quartz fibers, interspersed by layers of iron absorber, are used as active material. HCAL has a total thickness of 7 interaction lengths ( $\lambda_0$ ). Since this depth is not sufficient to ensure complete containment of hadronic showers, an additional layer of one  $\lambda_0$  has been added outside the magnet, to increase energy resolution.

The energy resolution of HCAL for pions is parametrized as:

$$\frac{\sigma(E)}{E} = \frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 8\%$$

while the resolution of the ECAL-HCAL combined system is:

$$\frac{\sigma(E)}{E} = \frac{84.7\%}{\sqrt{E(\text{GeV})}} \oplus 8\%$$

### 2.2.5 The Muon System

Muons are the only charged particles which are able to pass the calorimetric system without being absorbed. CMS has a system of muon chambers placed inside the iron of the magnet return yoke [44], where the magnetic field is about 1.5 T. Muon identification is done using these chambers, while  $p_T$  measurement is done combining chambers information with tracker measurements. A sketch of the muon system is shown in figure 2.9. As other detectors, it is subdivided in a barrel ( $|\eta| < 1.2$ ) and two endcaps ( $1.2 < |\eta| < 2.4$ ).

Different experimental techniques in different regions of the detector are used:

- Drift tubes (DT): This kind of detectors is used in the central part of the muon system. Each chamber is made of twelve 4-cm-wide tubes containing a stretched wire within a gas volume.
- Cathode strip chambers (CSC): In the endcap region, where particle multiplicity is higher, arrays of anode wires, crossed with cathode strips, within a gas volume are used for muon detection.
- Resistive Plate Chambers (RPC): Both barrel and endcaps are equipped with this fast gaseous detectors. They consist of two parallel plates separated by a gas volume. Their excellent time resolution (3 ns) makes them suitable to be used also as fast high-efficiency triggers.

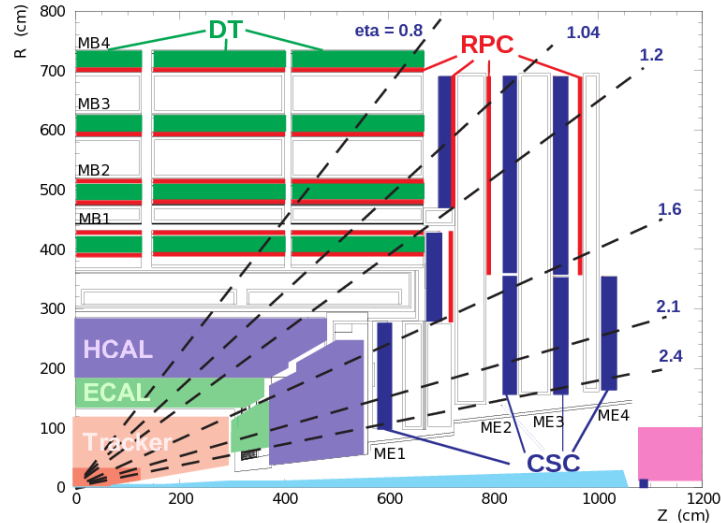


Figure 2.9. Transverse view of a quarter of the CMS muon system.

### 2.2.6 The Trigger System

When LHC is performing at its peak one billion of proton interactions per second takes place inside CMS detector. Since the typical raw event size is 1 MB, not all

these events can be recorded. Moreover, saving all events would not be useful since most of them are soft collisions, not interesting for the CMS physics program.

A “trigger system” is therefore needed in order to rapidly select potentially interesting events. The aim of the trigger system [45] is to lower the event rate from  $10^9$  Hz to the manageable level of 100 Hz. The Trigger system is structured in two levels: a Level-1 Trigger (L1) and a high-level trigger (HLT). The L1 system is made of a series of hardware processors able to do simple logical operations directly on detector signals. This allows fast decisions (maximum decision time is  $3.2 \mu s$ ) and the rate is reduced to 50-100 kHz.

Events passing L1 trigger are analyzed by HLT, which is a software system implemented in a computer farm made of about one thousand of processor. It takes decisions analyzing the reconstructed quantities for a given object. As an example, triggers selecting diphoton events for the analysis presented in this thesis, require the presence of two photons with a transverse momentum greater than a certain threshold and good isolation in the tracker and in the calorimetric system. The HLT system reduces the output rate to about 100 Hz.

### 2.2.7 CMS DAQ and Monte Carlo simulations

Once triggered, the events are written to disks and transferred to the CERN computer centre, where they are reconstructed and distributed to the physicists for the analysis.

Scientists around the world are able to access the CMS data using the GRID, a network of well coordinated computing centres. These preserve and distribute the data, producing also simulated events for CMS.

In order to tune the analysis strategies and to understand subtle detector effects, CMS makes use of Monte Carlo simulation of events. Particles distributions are first generated according to theory predictions, then a full description of the CMS detector response takes place. This simulation is based on the GEANT4 software which describes in detail interacting processes, like electromagnetic and hadronic showering. GEANT4 models the geometry of the different CMS subdetectors, the field map of the solenoid and the electronic readout. Simulated events are written to disk in a format similar to real data events, then are reconstructed and analyzed with the same software chain.

The analyses presented here use simulated events to estimate the number of expected signal events and to tune and optimize the different event selections. Several generators are used depending on the process, and will be described in detail in Sec. 4.2, 5.1.1 and 5.2.1. Since in all analyses two photons are required, the main background processes, i.e. prompt diphoton production ( $\gamma\gamma$ ), events with one real photon and one jet misidentified as photon ( $\gamma + \text{jets}$ ) and events with two jets (jet-jet) are used to optimize photon and electron identification and reconstruction described in the next chapter. The agreement between data and simulations after reconstruction is checked using Z production in the electron decay channel ( $Z \rightarrow e^+e^-$ ).



## Chapter 3

# Reconstruction and Identification of Physics Objects

The analyses described in this thesis present a great variety of final states. The presence of two photons, coming from the Higgs boson, is always required and the presence of additional particles is also used to target the different topologies. In this chapter the variables and the algorithms used for the reconstruction and identification of physics objects will be described. Given the key role of photons, their reconstruction and identification criteria will be thoroughly discussed.

The following sections describe how these objects are defined:

- photons,
- jets (including b-jets),
- electrons,
- muons.

### 3.1 Photons

#### 3.1.1 Photon reconstruction

Photon reconstruction is a key element for analyses involving the diphoton decay of the Higgs boson. Indeed, the aim of these analyses is to reconstruct the diphoton invariant mass and identify the narrow Higgs boson mass peak over a continuum, irreducible background spectrum. A good resolution in reconstructing the energy of the photons and the position of the interaction vertex is therefore crucial in  $H \rightarrow \gamma\gamma$ . Therefore the photon reconstruction and identification presented here is aimed at having a good resolution in the wide  $p_T$  range (from  $\sim 25$  GeV up to hundreds of GeV) typical of the  $H \rightarrow \gamma\gamma$  process.

An electromagnetic shower from a photon or an electron is spread over several ECAL crystals. The presence of material in front of the ECAL (corresponding to  $1-2 X_0$  depending on the  $\eta$  region) causes conversion of photons and bremsstrahlung from electrons and positrons. The strong magnetic field of the experiment tends to spread the radiated energy along  $\phi$  within the tracker volume.

Dynamic clustering algorithms [42] are used to sum together energy deposits in crystals belonging to the same electromagnetic shower and to recover the radiated energy. They proceed first with the formation of “basic clusters”, corresponding to local maxima of energy deposits. These are then merged to form a “supercluster” (SC), which is extended in  $\phi$ , to recover the radiated energy. In the barrel region, clustering is performed with the Hybrid algorithm which forms superclusters merging five-crystal-wide strips in  $\eta$  around the locally most energetic crystal, with a variable extension in  $\Phi$  (up to 35 crystals wide). In the endcap region the Multi5x5 algorithm is used. Matrices of 5x5 crystals, which may partially overlap and are centered on the locally most energetic crystal, are merged if lie in a narrow  $\Phi$  road.

The photon candidates are collected within the ECAL fiducial region  $|\eta| < 2.5$ , excluding the barrel-endcap transition region  $1.44 < |\eta| < 1.57$  where performances are suboptimal.

### 3.1.2 Photon Energy

The photon energy is computed using signals recorded from ECAL crystals forming the supercluster. In the region covered by the preshower detector, signals recorded in it are also considered. Calorimeter signals are corrected and calibrated for several instrumental effects.

Different components are needed to determine photon energy:

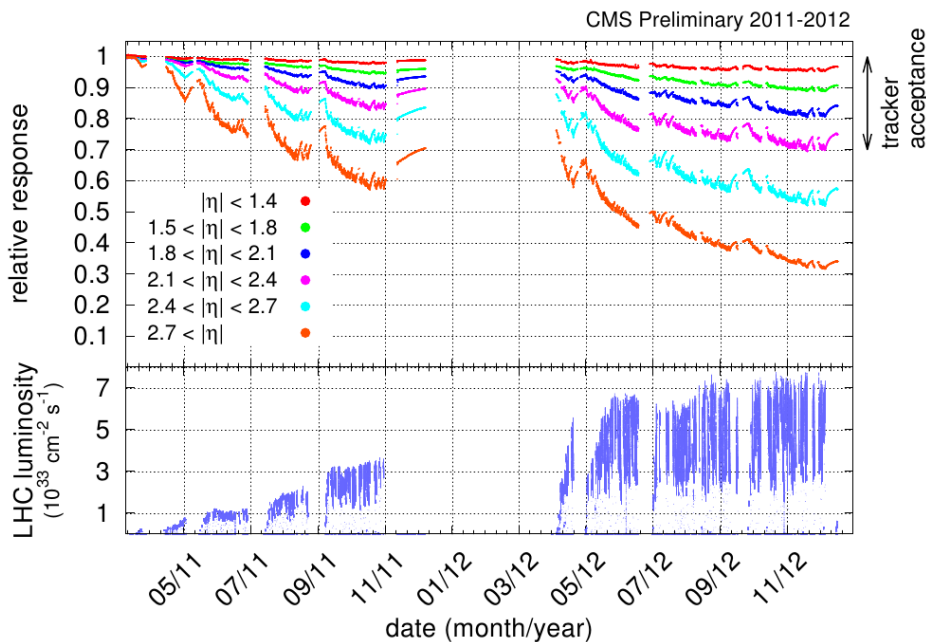
- Transparency loss correction
- Channel intercalibration
- Containment, showering and pileup corrections
- Energy scale and resolution corrections

The loss of crystal transparency during the data-taking, due to radiation, causes a variation in response. This effect is corrected using measurements of the crystal transparency performed via a laser monitoring system which checks the response of each single crystal approximately every 40 minutes. Crystal transparency decreases during proton collisions, while  $\text{PbWO}_4$  crystals spontaneously recover transparency during interfill periods [46]. The response change observed is of the order of a few percent in the barrel, while it reaches up to 25% in the most forward endcap regions, as can be seen in the top plot of figure 3.1. The plot also reports instantaneous luminosity delivered during the data-taking.

Intercalibration between different channels is achieved exploiting the  $\Phi$ -symmetry of the energy flow, the mass constraint on the energy of two photons in  $\pi^0$  and  $\eta$  decays and the momentum constraint on the energy of electrons from vector bosons decays.

A multivariate regression technique is used to derive corrections for the containment of the shower in the crystals and effects related to the material upstream ECAL and pileup. The photon energy response distribution can be parametrized by a Gaussian core and two power law tails; the regression is used to provide a per-photon estimate of the parameters of this function, which can be translated in the best estimate for the true energy and resolution.





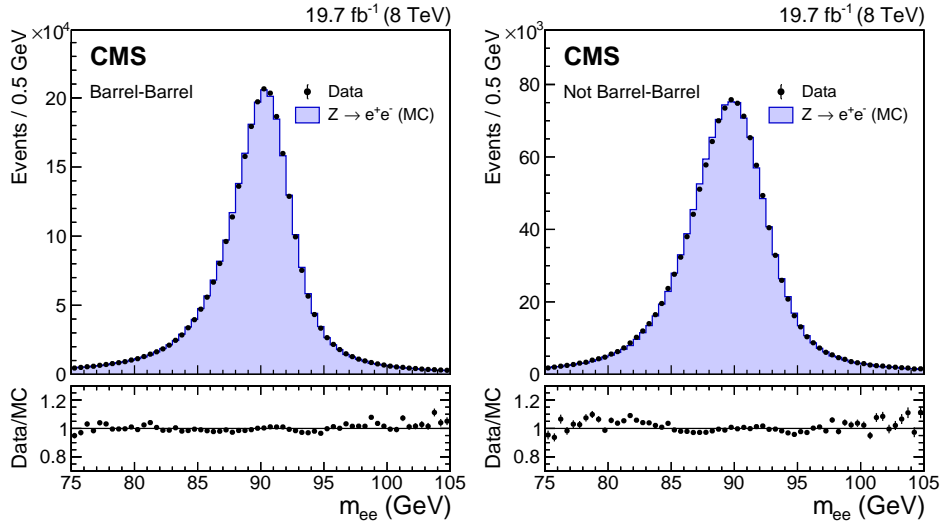
**Figure 3.1.** Crystal relative response in different  $\eta$  ranges for the full data-taking period. LHC instantaneous luminosity is also reported in the bottom plot.

The input variables to the regression are related to shower shape, like  $R_9$ , which is defined as the energy sum of a  $3 \times 3$  crystal matrix centered on the most energetic crystal in the supercluster divided by the energy of the supercluster. Other useful shower shape variables used in the regression are the energy weighted width of the supercluster in  $\eta$  ( $\sigma_{i\eta i\eta}$ ) and the ratio between the hadronic energy behind the supercluster and the electromagnetic energy (H/E). The number of vertexes and the median energy density  $\rho$  are also used to take into account pileup effects. A full description of transparency, intercalibration and additional corrections can be found in [42].

After these corrections,  $Z \rightarrow e^+e^-$  events are used to correct residuals observed discrepancies between data and simulation. A multistep procedure is used to correct the energy scale in data, and to determine a smearing term to be applied to showers in simulated events to reproduce correctly the resolution observed. The energy scale in data is equalized, using  $Z \rightarrow e^+e^-$  events in which electrons are reconstructed as photons. Since data-MC difference is  $\eta$  and time dependent, these corrections are derived in bins of  $\eta$  using 8 epochs in the 7 TeV dataset and 51 epochs in the 8 TeV one. After this, photon energy resolution in simulated events is made more realistic adding a Gaussian smearing to match the  $Z \rightarrow e^+e^-$  energy resolution measured in data. This smearing term is derived in bins of  $\eta$  and  $R_9$ , since converted (low  $R_9$ ) and unconverted (high  $R_9$ ) photons need different corrections. A sufficient number of  $Z \rightarrow e^+e^-$  events is present for 8 TeV dataset to apply further residual scale and resolution corrections for electrons in the barrel. This is done in 20 bins defined by ranges in  $\eta$ ,  $R_9$  and  $E_T$ . This allows to consider also an energy dependence in

correction factors.

Figure 3.2 shows the invariant mass distribution for  $Z \rightarrow e^+e^-$  events in the 8 TeV data and simulated events in which electron showers are reconstructed as photons. The full set of corrections to data and smearings to simulated energies described here are applied. An excellent agreement between data and simulation is observed after all these corrections. Small discrepancies in the low-tail of the distribution for endcap electrons are related to non-Gaussian effects in energy loss.



**Figure 3.2.** Invariant mass of dielectron pairs in  $Z \rightarrow e^+e^-$  events in the 8 TeV dataset (points) and simulations (full histogram) in which electrons are reconstructed as photons. The comparison is shown dividing events with both shower in the barrel (left) and events with at least one shower in the endcap.

### 3.1.3 Diphoton triggers and photon preselection

Since in all the analyses presented here at least two photons are required in the final state, all events are required to pass diphoton triggers. Different requirements with asymmetric transverse energy thresholds and two different photon selections are used at HLT level. One selection requires a loose calorimetric identification, based on the shape of electromagnetic shower, and loose isolation requirements. The other requires a high value for the shape variable  $R_9$ , without any further requirement on the photon candidate. Since these two selections have a complementary nature, the photon candidate is required to pass at least one of the two, in order to preserve high efficiency on signal events. Two different trigger threshold configurations are used depending on the data taking period: one requiring  $E_T > 26$  (18) GeV for the leading (trailing) photon and another one requiring  $E_T > 36$  (22) GeV. The trigger efficiency for events entering the analysis is 99.4%. The full list of trigger paths used in the 2012 dataset is given in table 3.1

A loose preselection is then applied on photons entering the analysis. The aim of this preselection is to remove jets misreconstructed as photons in events passing

**Table 3.1.** HLT paths employed in 2012 data-taking for the analyses presented in this thesis. HLT paths and corresponding L1 triggers are shown.

Trigger Path	L1 Seed
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass60	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass70	L1_DoubleEG_13_7
HLT_Photon26_CaloId10_Iso50_Photon18_CaloId10_Iso50_Mass60	L1_DoubleEG_13_7
HLT_Photon26_CaloId10_Iso50_Photon18_R9Id85_Mass60	L1_DoubleEG_13_7
HLT_Photon26_CaloId10_Iso50_Photon18_R9Id85_Mass60	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_Photon18_CaloId10_Iso50_Mass60	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_Photon18_R9Id85_Mass60	L1_DoubleEG_13_7
HLT_Photon26_Photon18	L1_DoubleEG_13_7
HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18	L1_DoubleEG_13_7
HLT_Photon36_CaloId10_Iso50_Photon22_CaloId10_Iso50	L1_SingleEG_22
HLT_Photon36_CaloId10_Iso50_Photon22_R9Id85	L1_SingleEG_22
HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon22_R9Id85_OR_CaloId10_Iso50	L1_SingleEG_22
HLT_Photon36_R9Id85_Photon22_CaloId10_Iso50	L1_SingleEG_22
HLT_Photon36_R9Id85_Photon22_R9Id85	L1_SingleEG_22
HLT_Photon36_Photon22	L1_SingleEG_22
HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon22	L1_SingleEG_22

the HLT requirements. Typically these photon candidates arise in events with high energy neutral mesons that take a substantial fraction of the jet  $p_T$ . The preselection criteria are similar to, but slightly more stringent than, the trigger requirements.

These consist of:

- $p_T^{\gamma^1} > 33$  GeV and  $p_T^{\gamma^2} > 25$  GeV, where  $p_T^{\gamma^1}$  and  $p_T^{\gamma^2}$  are the transverse momenta of the leading (in  $p_T$ ) and subleading photon, respectively,
- a selection on the hadronic leakage of the shower, measured as the ratio of energy in HCAL cells behind the supercluster to the energy in the supercluster (H/E),
- a loose selection based on isolation and shower shape,
- an electron veto, which removes the photon candidate if its supercluster is matched to an electron track with no missing hits in the innermost tracker layers, thus excluding almost all  $Z \rightarrow e^+e^-$  events.

The selection requirements are applied with different stringency in four categories, in  $\eta$  and  $R_9$ , defined to match the different selections used in the trigger. The efficiency of the photon preselection is measured in data using a “tag-and-probe” technique [47]. The efficiency of all preselection criteria, except the electron veto requirement, is measured using  $Z \rightarrow e^+e^-$  events. The efficiency for photons to satisfy the electron veto requirement is measured using  $Z \rightarrow \mu\mu\gamma$  events, where the photon is produced by final-state radiation, which provide a more than 99% pure source of prompt photons. The ratio of the photon efficiency measured in data to that found in simulated  $Z \rightarrow e^+e^-$  events,  $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$ , is consistent with unity in all categories. For photons in simulated Higgs boson events the efficiency of these preselection criteria ranges from 92% to 99% depending on  $\eta$  and  $R_9$  and it is reported in table 3.2.

**Table 3.2.** Photon preselection efficiencies for both the 7 and 8 TeV datasets measured for  $Z \rightarrow e^+e^-$  events, where the electrons are reconstructed as photons, in four photon categories. The statistical uncertainties in the efficiencies found in simulated events are negligible.

Preselection category	$\epsilon_{\text{data}}$ (%)	$\epsilon_{\text{MC}}$ (%)	$\epsilon_{\text{data}}/\epsilon_{\text{MC}}$
7 TeV dataset			
Barrel; $R_9 > 0.90$	$98.7 \pm 0.3$	99.1	$0.996 \pm 0.003$
Barrel; $R_9 < 0.90$	$96.2 \pm 0.5$	96.7	$0.995 \pm 0.006$
Endcap; $R_9 > 0.90$	$99.1 \pm 0.9$	98.2	$1.008 \pm 0.009$
Endcap; $R_9 < 0.90$	$96.1 \pm 1.5$	95.6	$1.005 \pm 0.018$
8 TeV dataset			
Barrel; $R_9 > 0.90$	$98.8 \pm 0.3$	98.6	$0.999 \pm 0.003$
Barrel; $R_9 < 0.90$	$95.7 \pm 0.6$	96.1	$0.995 \pm 0.006$
Endcap; $R_9 > 0.90$	$98.4 \pm 0.9$	97.9	$1.005 \pm 0.009$
Endcap; $R_9 < 0.90$	$95.5 \pm 1.7$	94.5	$1.011 \pm 0.018$

### 3.1.4 Cut-based photon identification

In this thesis two different sets of photon identification criteria are used: a cut-based photon identification and a multivariate one. The first is a selection that uses simple cuts on shower shape and isolation to separate true photons from fake ones; the second is a multivariate selection trained on  $H \rightarrow \gamma\gamma$  SM signal. For BSM searches, presented in chapter 5, the cut-based one is used, since it is more model-independent, while for SM Higgs searches the MVA one is preferred to measure Higgs properties with the best possible sensitivity.

The cut-based photon identification criteria (CiC4PF) are described in [48]. Photon identification is performed dividing photons into four mutually exclusive categories depending on whether the photon is in the barrel or endcap, and on whether the shower is narrow ( $R_9 > 0.94$ ).

The four event classes are:

- Both photons are in the barrel and have  $R_9 > 0.94$ ,
- Both photons are in the barrel and at least one of them fails the requirement  $R_9 > 0.94$ ,
- At least one photon is in the endcap and both photons have  $R_9 > 0.94$ ,
- At least one photon is in the endcap and at least one of them fails the requirement  $R_9 > 0.94$ .

The discriminating variables used are a subset of the ones that will be used in the multivariate identification and cut values are optimized separately for the four categories. In table 3.3 a summary of the variables and the cut values for each category are shown.

For photon isolation, energy contributions computed with the particle-flow (PF) algorithm [49, 50] are used. The particle-flow event reconstruction consists in reconstructing and identifying each single particle with an optimized combination of all subdetector information. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of three contributions: the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy.

PF isolation sum is defined as:

$$\sum PFChIso + \max(0., PFPhotIso + PFNeuIso - A_{eff} \times \rho)$$

where  $PFChIso$ ,  $PFPhotIso$  and  $PFNeuIso$  are respectively the energy contribution, computed with the PF algorithm, of charged hadrons, photons and neutral

hadrons in a cone of  $\Delta R = 0.3$  around the photon direction. As can be seen from Table 3.3 the PF isolation sum is computed for two different vertex hypotheses: the “chosen vertex” which is the photon vertex chosen by the PF algorithm and the “worst vertex”, which is the one with the highest value of Charged PF Isolation sum (*PFChIso*). The variable  $\rho$  represents the medium energy density of the event and  $A_{eff}$  measures the effective area covered by the cone, subtracting from the isolation cone the inner region.

Photon identification efficiency on ggH events as a function of  $\eta$  and  $p_T$  are shown for Higgs events in figure 3.3. Efficiency as a function of  $\eta$  is lower in the endcaps given the bigger dimension of the crystals in the  $(\eta-\phi)$  space. Efficiency is higher at high  $p_T$  as imposed by isolation cuts which scale as a function of  $p_T$ .

Photons with high  $R_9$  are in general unconverted and have a better energy resolution and a lower contamination from misidentified jets. Similarly, photons in the barrel have better resolution, therefore the first category is the one with higher energy resolution. Performances and consistency between data and simulation for CiC4PF selection are checked using  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-\gamma$  events. Even if this selection is less sensitive with respect to MVA identification, good performances in Higgs searches for signal and background separation and good agreement between data and simulation are observed.

**Table 3.3.** Photon ID selection cut values. The cuts are applied to both the leading and subleading photons.

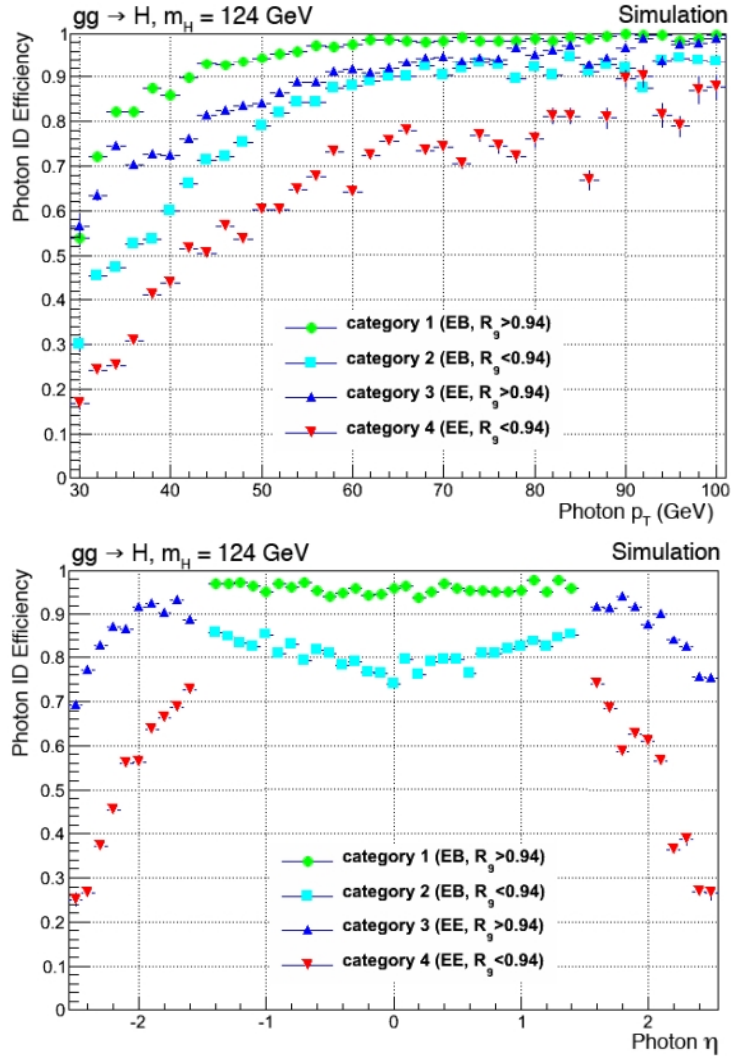
	barrel		endcap	
	$R_9 > 0.94$	$R_9 < 0.94$	$R_9 > 0.94$	$R_9 < 0.94$
PF isolation sum, chosen vertex	6 GeV	4.7 GeV	5.6 GeV	3.6 GeV
PF isolation sum worst vertex	10 GeV	6.5 GeV	5.6 GeV	4.4 GeV
Charged PF isolation sum	3.8 GeV	2.5 GeV	3.1 GeV	2.2 GeV
$\sigma_{i\eta i\eta}$	0.0108	0.0102	0.028	0.028
H/E	0.124	0.092	0.142	0.063
$R_9$	0.94	0.298	0.94	0.24

### 3.1.5 Photon Identification MVA

A Boosted Decision Tree (BDT), a multivariate analysis method implemented in the TMVA software [51], is trained to separate prompt photons from non-prompt ones. Non-prompt photons are primarily decay products of neutral mesons ( $\pi^0$ ,  $\eta$ ) from jet fragmentation. This BDT is trained using Monte Carlo simulated events with diphoton,  $\gamma + \text{jets}$  and QCD events as background and the Higgs boson process as signal.

The input variables to this BDT can be divided in different categories:

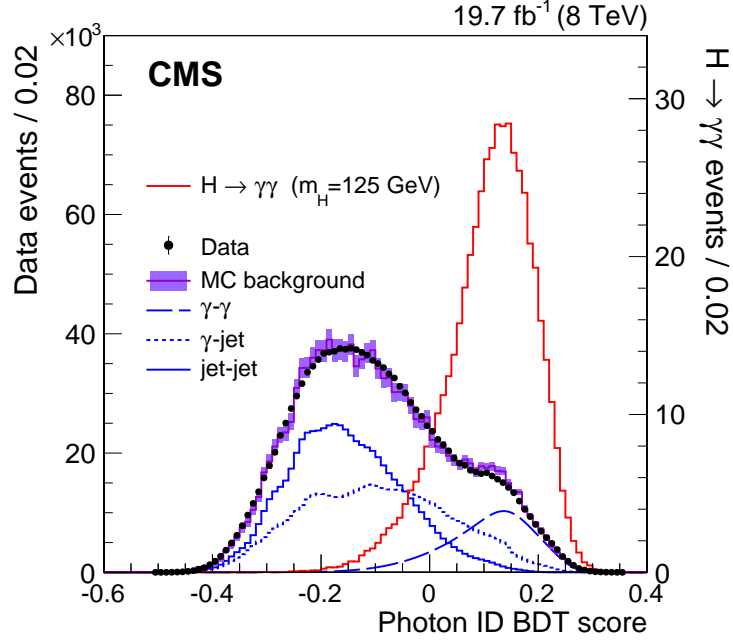
- Lateral shower shape variables: variables related to shower topology, like  $R_9$ , or the shower spread in the preshower detector (where it is present). Agreement between data and simulation is checked with  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-\gamma$  events.



**Figure 3.3.** Efficiency of cut-based photon identification as a function of  $p_T$  and  $\eta$  on  $ggH$  events. Different colors correspond to different categories of the selection.

- Isolation variables: all variables are based on the particle-flow algorithm and use sums of  $p_T$  of photons and charged hadrons, within regions of  $\Delta R < 0.3$  around the candidate photon, where  $\Delta R = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}$ .
- The median energy density per unit area in the event,  $\rho$ : this variable is used to take into account pileup dependence of the photon isolation variables.
- The pseudorapidity and the energy of the supercluster corresponding to the candidate photon.

In figure 3.4 the photon identification BDT score of the lower scoring photon in diphoton pairs is shown. In this plot events from the 8 TeV dataset in the range  $100 < m_{\gamma\gamma} < 180$  GeV, passing preselection, are considered. A good agreement between data and simulated events used for the BDT training can be observed. For



**Figure 3.4.** Photon identification BDT score of the lower-scoring photon of diphoton pairs in the range  $100 < m_{\gamma\gamma} < 180$  GeV for the 8 TeV dataset. Different contributions are shown: data (points), simulated background events (histogram with shaded error bands for statistical uncertainty) and simulated Higgs boson signal events (tall histogram on the right corresponding to the right-hand vertical axis).

photon ID, the agreement between data and simulation is checked using electrons from  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-\gamma$  decays and diphoton events with  $m_{\gamma\gamma} > 160$  GeV.

### 3.2 Jets

Given the QCD process called *color confinement*, single colored particles, like quarks and gluons, lose their color charge forming colorless configuration. This process leads to *hadronization* of partons into collimated jets of hadrons, collinear to the original direction of the particle. The aim of jet reconstruction is to reconstruct coherently the group of particles combining all the experimental information coming from the tracker and the calorimetric system.

Jets are reconstructed by clustering charged and neutral particles with the anti- $k_T$  algorithm [52], using a distance parameter of 0.5. All the particles are reconstructed using the particle-flow algorithm, which globally combines measurements from single sub-detectors to construct jets [50]. Jets are required to have at least two PF constituents and more than 1% of their energy in both the electromagnetic and hadronic components to reject jets arising from instrumental effects. Leptons are included in the clustering, but a  $\Delta R > 0.5$  from the jet is required when selecting leptons coming from W or Higgs boson decays in leptonic exclusive categories. Finally jets with a  $\Delta R < 0.5$  within either photon are rejected to avoid the possibility of photons included in jet reconstruction.



Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and the scale is found in the simulation to be within 5% to 10% of the true momentum over the whole transverse momentum spectrum and detector acceptance. The jet energy measurement is calibrated, to correct for instrumental effects, using samples of dijet,  $\gamma$  + jets and Z+jets events [53]. Energy coming from pileup interactions is subtracted using an  $\eta$ -dependent transverse momentum density, evaluated event-by-event. The jet energy resolution typically amounts to 15% (8%) at 10 (100) GeV, to be compared to about 40% (12%) obtained when the calorimeters alone, instead of the particle flow algorithm, are used for jet clustering.

Jets, in the 8 TeV dataset, are also required to pass a cut based pile-up identification which removes contributions from particles originating from multiple interactions. This identification is based on the compatibility of the tracks with the primary vertex and jet shape and has been tested on Z+jets events [54].

Jets originating from bottom quarks (b-jets) can be identified since b-quarks hadronize in particles with a lifetime sufficient to travel some distance in the tracker before decaying. This means that tracks of particles forming the b-jet do not originate from the interaction vertex but from a displaced secondary vertex. The decay vertex is identified using the combined secondary vertex algorithm (CSV) [55]. This sophisticated tagging technique exploits all known variables, which can distinguish b from non-b jets, combining information about impact parameter significance, the secondary vertex and jet kinematics.

Throughout this thesis two different working points of this tagger are used:

- the *loose* working point provides an efficiency for identifying b jets of about 80% and a misidentification probability for jets from light quarks and gluons of about 10%;
- the *medium* working point provides an efficiency for identifying b jets of about 70% and a misidentification probability for jets from light quarks and gluons of about 1%;

## 3.3 Leptons

The presence of well isolated leptons and muons, coming from decays of vector or Higgs bosons, are required in different steps of this analysis to create low-background categories. Indeed, the presence of leptons usually allows to suppress QCD backgrounds by order of magnitudes.

Throughout this thesis the word *lepton* indicates only electrons and muons, since  $\tau$  leptons are unstable and are revealed experimentally throughout their decay products.

### 3.3.1 Muons

Muons are measured in the pseudorapidity range  $|\eta| < 2.4$ , with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with  $20 < p_T < 100$  GeV of

1.3–2.0% in the barrel and better than 6% in the endcaps. The  $p_T$  resolution in the barrel is better than 10% for muons with  $p_T$  up to 1 TeV [56].

Muon identification requirements are summarized in table 3.4. A tight selection on the track, based on its quality and on the compatibility with the primary vertex, is applied. Consistency between tracks observed in the tracker and those in the muon chambers is required, in order to reduce the contamination of muons originating from hadrons decay.

**Table 3.4.** Muon ID requirements

$\chi^2/\text{NDF}$ of the global-muon track fit	< 10
At least one muon chamber hit included in the global-muon track fit	
Muon segments in at least two muon stations	
$d_{xy} < 2$ mm (w.r.t. primary vertex)	
$d_z < 5$ mm (w.r.t. primary vertex)	
Number of pixel hits	> 0
Cut on number of tracker layers with hits	> 5

### 3.3.2 Electrons

Electrons are identified as clusters of energy in ECAL, associated to a track in the silicon tracker. Electron candidates are required to have an ECAL supercluster within the same fiducial region as for photons.

The electron selection criteria are summarized in table 3.5.

The electron track has to fulfill requirements on the transverse and longitudinal impact parameter with respect to the electron vertex and cannot have more than one missing hit in the innermost layers of the tracker for conversion rejection. Electron isolation is computed with the same particle-flow based algorithm used for photons.

After this preselection, multivariate techniques are used for electron identification. Input variables are related to topological properties of the track, shower shape and isolation in the calorimetric system and in tracker. The highest MVA score electron is considered as the electron tag candidate. In order to choose the cut on the MVA output variable, an optimization which uses the  $S/\sqrt{B}$  as figure of merit has been performed, choosing a cut at 0.9.

$d_0$ w.r.t. electron vertex	< 0.02 cm
$d_z$ w.r.t. electron vertex	< 0.2 cm
Combined relative PF isolation	0.15
Vertex fit probability (conv. rej)	$10^{-6}$
Missing hits (conv. rej.)	0 or 1
MVA Cut	0.9

**Table 3.5.** Electron identification requirements.

## Chapter 4

# The $H \rightarrow \gamma\gamma$ SM Analysis and Coupling Measurements

### 4.1 Introduction

The diphoton decay channel provides a clean final-state topology: two energetic and isolated photons reconstructed in the electromagnetic calorimeter. For this reason, it plays a major role for a low-mass Higgs boson, despite its low branching ratio. The CMS electromagnetic calorimeter has been specifically designed for achieving the best sensitivity possible in this channel.

The  $H \rightarrow \gamma\gamma$  process offers the possibility to fully reconstruct the decay products of the Higgs boson with high resolution. The number of signal and background events is estimated performing a fit to the diphoton invariant mass  $m_{\gamma\gamma}$ , defined as

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - \cos\theta)}$$

where  $E_1$  and  $E_2$  are the energy of the photons and  $\theta$  is the angle between the two. As can be seen from this formula a good resolution on energy reconstruction and vertex determination, to determine  $\theta$ , is needed. A dedicated photon identification and reconstruction has been presented in paragraph 3.1. The diphoton vertex determination will be described in details in paragraph 4.2.1.

Events in this analysis are classified in different exclusive categories targeting different production modes. This classification depends on the presence of additional objects in the final state and on the expected mass resolution. Since this separation allows a good sensitivity in coupling measurements, particular attention is given in this chapter to the description of exclusive categories. Given the importance of  $t\bar{t}H$  production mode in theoretical models, a combination of the diphoton channel with another analysis targeting the same production mode in other decay channels, is described separately at the end of the chapter.

### 4.2 Data Samples and simulated events

This analysis is performed on the full dataset collected by the CMS experiment, in proton-proton collisions, during 2011 and 2012. It corresponds to an integrated luminosity of  $5.1 \text{ fb}^{-1}$  at a center of mass energy of 7 TeV and  $19.7 \text{ fb}^{-1}$  at 8 TeV.

The production and decay of the Higgs in two photons and all background processes are obtained using simulations. The simulation of the detector response utilizes a detailed description of CMS produced with the GEANT4 software [57]. Simulated events include multiple interactions taking place in each bunch crossing (pileup). The default distribution was initially different from the one observed in data. The profile of the number of pileup events in simulations is therefore corrected, to match the actual distribution in data.

Signal is simulated separately for the four different production modes. The two processes with the highest cross section, ggH and VBF, are obtained using the next-to-leading order (NLO) matrix-element generator POWHEG [58–62] interfaced with PYTHIA [63]. Events at 7 TeV are weighted to reproduce the transverse momentum spectrum of the Higgs boson from ggH computed at higher order by the HQT program [64–66]. HQT performs the resummation of the large logarithmic contributions appearing at transverse momenta  $q_T$  much smaller than the mass of the Higgs boson. At 8 TeV, POWHEG has been tuned to directly reproduce the HQT spectrum. VH and  $t\bar{t}H$  processes are instead generated with PYTHIA at leading-order while higher order diagrams are taken into account by the parton showering. The SM Higgs boson cross sections and branching ratios used can be found in [67].

Simulated samples of  $Z \rightarrow e^+e^-$ ,  $Z \rightarrow \mu^+\mu^-$  and  $Z \rightarrow \mu^+\mu^-\gamma$  events used for data/MC comparison and photon energy calibration are generated with MADGRAPH [68], SHERPA [69] and POWHEG. Simulated background samples are used for multivariate discriminants and classification optimization. At  $\sqrt{s} = 7$  TeV the diphoton processes are obtained using a combination of MADGRAPH interfaced to PYTHIA, for all processes apart from the gluon-fusion box diagram, for which PYTHIA alone is used. At  $\sqrt{s} = 8$  TeV the diphoton processes are simulated using SHERPA. The remaining backgrounds, i.e.  $\gamma$ +jets events where one photon candidate arises from a misidentified jet, are simulated using PYTHIA and cross sections of the processes are scaled using  $k$ -factors derived from CMS measurements [70, 71].

#### 4.2.1 Diphoton vertex assignment

The mean number of vertexes per bunch crossing is 9 in the 7 TeV dataset and 21 in the 8 TeV one. The luminous region, in the longitudinal direction  $z$ , has a spread of about 5 cm. The resolution on the photons opening angle makes a negligible contribution to mass resolution, compared to energy resolution, only if the vertex is known within about 10 mm.

Since diphoton events used in this analysis usually have a low number of charged tracks, the vertex is determined exploiting kinematic properties of the diphoton system. Additional tracks coming from photon conversions or objects in the final state are also considered in vertex determination.

Three variables are used in vertex determination:

- $\sum \vec{p}_T^2$
- $-\sum(\vec{p}_T \cdot \frac{\vec{p}_T^{\gamma\gamma}}{|\vec{p}_T^{\gamma\gamma}|})$
- $(|\sum \vec{p}_T| - |\vec{p}_T^{\gamma\gamma}|)/(|\sum \vec{p}_T| + |\vec{p}_T^{\gamma\gamma}|)$

where  $\vec{p}_T$  is the transverse momentum of the charged tracks and  $\vec{p}_T^{\gamma\gamma}$  the transverse momentum of the diphoton system. An additional variable  $g_{conv}$  is used for converted photons, since additional electron tracks coming from conversions can be used for vertex determination. From these tracks an estimate of vertex position  $z_e$  can be obtained with a given uncertainty  $\sigma_e$ . The variable  $g_{conv}$  is defined as the pull between  $z_e$  and  $z_{vtx}$ , the longitudinal position of the reconstructed vertex:  $g_{conv} = |z_e - z_{vtx}|/\sigma_e$ .

These variables are used as inputs to a multivariate discriminator using BDT techniques, referred as “vertex assignment BDT”. The vertex finding efficiency, i.e. the efficiency that the chosen vertex is within 10 mm of the true location, is estimated on  $Z \rightarrow \mu^+\mu^-$  events, removing the muon tracks to mimic diphoton events, and comparing the chosen vertex with the true one from the muon pair. The use of converted photons in vertex reconstruction is instead validated on  $\gamma + \text{jets}$  events. The vertex finding efficiency, for a 125 GeV Higgs boson, is computed to be 85.4 (79.6)% in the 7 (8) TeV dataset.

A second vertex discriminant (BDT<sup>vtx</sup>) is used to estimate, event-by-event, the probability of correct vertex assignment within 10 mm of the diphoton interaction. It is used, together with the event-by-event estimate of the energy resolution, to estimate the diphoton mass resolution for each event.

The inputs of this vertex probability BDT are:

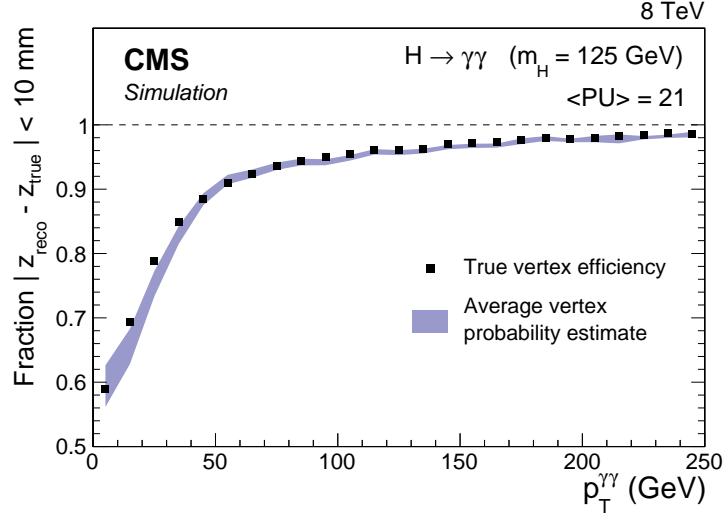
- the values of the vertex assignment BDT for the three vertexes in the event with highest score,
- the total number of reconstructed vertexes in the event,
- the transverse momentum of the diphoton system  $p_T^{\gamma\gamma}$ ,
- the distance between the chosen vertex and the second- and third-best vertexes,
- the number of converted photons.

Performances of this BDT are determined using  $Z \rightarrow \mu^+\mu^-$  events, assuming the vertex of the dimuon pair as the true vertex. Validation for events with converted photons is done on  $\gamma + \text{jets}$  events. In Figure 4.1 the fraction of diphoton vertexes correctly assigned by the vertex BDT and the probability of correctly locating the vertex on simulated signal events are shown as a function of  $p_T^{\gamma\gamma}$ .

### 4.2.2 Event Classification overview

The photon reconstruction and identification used in this analysis are those described in paragraph 3.1. The additional preselection described in section 3.1.5 is applied to further suppress diphoton and  $\gamma + \text{jets}$  events with one jet misreconstructed as photon. For photons in simulated Higgs boson events the efficiency of these preselection criteria ranges from 92% to 99% depending on  $\eta$  and  $R_9$ .

The analysis uses events with two photon candidates, requiring for both of them a score for MVA photon identification greater than -0.2. This requirement is almost fully efficient on signal events passing all the other selection requirements while removes around 25% of the background. Events with  $m_{\gamma\gamma}$  outside the window



**Figure 4.1.** Fraction of diphoton vertices (solid points) assigned by the vertex assignment BDT to a reconstructed vertex within 10 mm of their true location in simulated Higgs boson events as a function of  $p_T^{\gamma\gamma}$ . The probability of correctly locating the vertex is also shown as a band. The mean prediction, estimated from the vertex probability BDT, is calculated in  $p_T^{\gamma\gamma}$  bins and the width of the band represents the event-by-event uncertainty in the estimates.

100-180 GeV are discarded and a  $p_T$  greater than  $m_{\gamma\gamma}/3$  ( $m_{\gamma\gamma}/4$ ) is required for the leading (trailing) photon. The use of  $p_T$  thresholds scaled by  $m_{\gamma\gamma}$  is needed to avoid the distortion of the shape at the low edge of the  $m_{\gamma\gamma}$  spectrum. In few cases where two diphoton pairs are present the one with the highest  $p_T^{\gamma 1} + p_T^{\gamma 2}$  is chosen.

To achieve the best sensitivity, events are divided into classes based on their mass resolution and relative probability to be due to signal rather than background. In total 14 event classes are used for the 8 TeV analysis and 11 for the 7 TeV one. These classes are exclusive, which means that events are tested against class selection requirements in a fixed order. Classes tagging a specific associate production mode (VBF, VH and  $t\bar{t}H$ ) are checked first, looking for additional objects in the final state. Untagged classes collect the remaining events ( $\sim 99\%$  of the total). Once selected, events are no longer candidate for inclusion in other classes. Selection criteria used in each category are described in details in section 4.3. The ordering is that shown in Table 4.1, with a summary of the main requirements.

If an event does not satisfy the requirements of any of the exclusive categories it falls in the untagged ones. Untagged events are categorized according to a diphoton multivariate discriminator ( $BDT^{\gamma\gamma}$ ), described in the next paragraph. Also events in VBF categories are split according to a multivariate discriminant ( $BDT^{VBF}$ ). It is a combined BDT exploiting both diphoton and dijet kinematics (cfr. paragraph 4.3.1 for details).

Events in VH and  $t\bar{t}H$  classes are classified according to the objects present in the final state and different cuts on the diphoton multivariate discriminant. In addition to the cuts described later in detail for each tag the following cuts are applied on

**Table 4.1.** Event classes for the 7 and 8 TeV datasets and some of the selection requirements. Events are tested against the selection requirements of the classes in the order they are listed here.

Label	No. of classes		Main requirements
	7 TeV	8 TeV	
ttH lepton tag	★	1	$p_T^{\gamma 1} > m_{\gamma\gamma}/2$ 1 b-tagged jet + 1 electron or muon
VH tight $\ell$ tag	1	1	$p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$ [e or $\mu$ , $p_T > 20$ GeV, and $E_T^{\text{miss}} > 45$ GeV] or [2e or $2\mu$ , $p_T^\ell > 10$ GeV; $70 < m_{\ell\ell} < 110$ GeV]
VH loose $\ell$ tag	1	1	$p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$ e or $\mu$ , $p_T > 20$ GeV
VBF dijet tag 0-2	2	3	$p_T^{\gamma 1} > m_{\gamma\gamma}/2$ 2 jets; classified using combined diphoton-dijet BDT
VH $E_T^{\text{miss}}$ tag	1	1	$p_T^{\gamma 1} > 3m_{\gamma\gamma}/8$ $E_T^{\text{miss}} > 70$ GeV
t $\bar{t}$ H multijet tag	★	1	$p_T^{\gamma 1} > m_{\gamma\gamma}/2$ 1 b-tagged jet + 4 more jets
VH dijet tag	1	1	$p_T^{\gamma 1} > m_{\gamma\gamma}/2$ jet pair, $p_T^j > 40$ GeV and $60 < m_{jj} < 120$ GeV
Untagged 0-4	4	5	The remaining events, classified using diphoton BDT

★ For the 7 TeV dataset, events in the ttH lepton tag and multijet tag classes are selected first, and combined to form a single event class.

BDT $^{\gamma\gamma}$ : 0.17 for the two VH lepton classes and t $\bar{t}$ H lepton class, 0.62 for the VH  $E_T^{\text{miss}}$  tagged class, 0.76 for the VH dijet-tagged class, 0.48 for t $\bar{t}$ H multijet class, where the numerical scale is the classifier shown in figure 4.2. These boundaries have been obtained by minimizing the expected uncertainty in the measurement of signal strength of the process, using data in control regions to estimate the background, and MC signal samples to estimate the signal efficiency.

### 4.2.3 Multivariate event classifier

A multivariate event classifier, called diphoton BDT (BDT $^{\gamma\gamma}$ ), is built to fulfill the following criteria:

1. It should assign a high score to events that have
  - (a) good diphoton mass resolution
  - (b) high probability of being signal rather than background
2. It should be mass independent

This classifier is constructed using a Boosted Decision Tree implemented in TMVA. It incorporates as input variables those related to diphoton kinematics, mass resolution and photon identification.

The complete list of variables used is:

- the transverse momentum, divided by  $m_{\gamma\gamma}$ , and pseudorapidity of photons,
- the cosine of the angle between the two photons in the transverse plane ( $\cos(\Phi_{\gamma 1} - \Phi_{\gamma 2})$ ),
- the expected relative diphoton mass resolutions under the two hypotheses of selecting the correct or the wrong interaction vertex,
- the probability of selecting the correct vertex,
- the photon identification BDT classifier for both photons.

The choice of these variables has been made following the principle that the Higgs signal over background ratio,  $S/B$ , depends both on kinematic properties of the diphoton system and on photon identification. Indeed, if both photons are in the barrel, have high  $p_T^{\gamma\gamma}$  and high values of identification BDT, the  $S/B$  ratio is expected to be higher. The diphoton BDT is trained using as signal simulated Higgs boson events, having a mass  $m_H = 123$  GeV which is a value near the centre of the mass range of this analysis, and other MC simulations of other processes as background.

The diphoton BDT output distribution for the 8 TeV dataset is shown in figure 4.2. The classifier output has been transformed such that the sum of signal events from all processes has a flat distribution. This allows an easier estimation of BDT performances directly from the plot. The vertical dashed lines indicates the boundaries of untagged event classes. These boundaries have been set minimizing the expected uncertainty in the overall signal strength measurement.

As can be seen, the signal-to-background ratio increases considerably with the classifier score. Given the harder  $p_T$  of the diphoton system compared to inclusive production, events produced through VBF, W/ZH and  $t\bar{t}H$  peak at higher values of the classifier output.

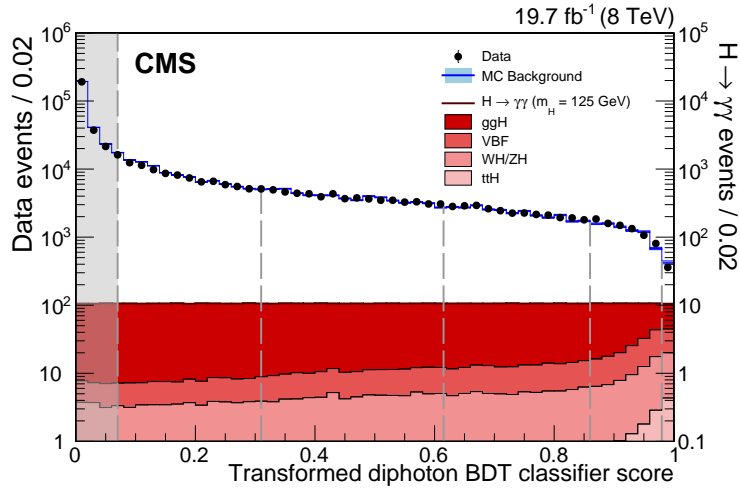
The diphoton BDT has also been checked on  $Z \rightarrow e^+e^-$  events where electrons are reconstructed as photons. In figure 4.3 the data/MC comparison for the BDT output is shown. The good data/MC agreement for  $Z \rightarrow e^+e^-$  represents an important cross check which ensures that the modeling of the BDT input variables and their correlations in the simulations are accurate. The band in the plot indicates the systematic uncertainty coming from propagation of photon BDT score and uncertainty on per photon estimator of energy uncertainty.

### 4.3 Exclusive Categories Tagging for Associate Production Modes

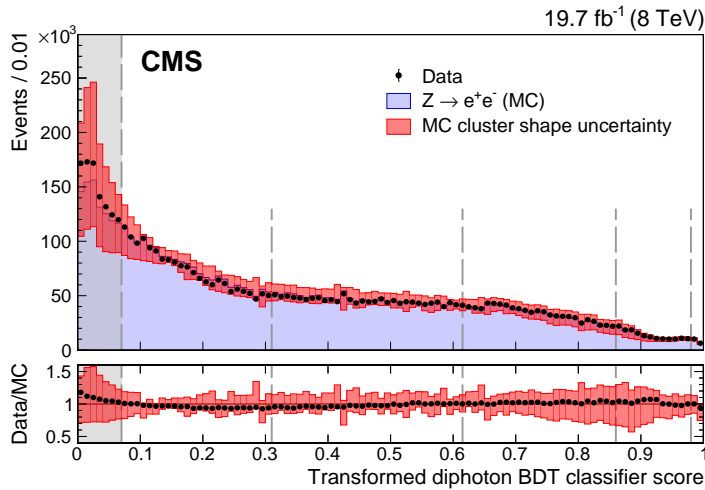
The presence of additional objects in the final state allows to construct categories enriched in a single production mechanism. This categorization is crucial to achieve the best possible sensitivity in couplings measurement. For each production mode the different signatures are exploited to create an optimized selection.

Higgs bosons produced via VBF are accompanied with a pair of energetic jets in the forward region. According to the different decays of the W/Z boson, Higgs





**Figure 4.2.** Transformed diphoton BDT classifier score in 8 TeV datasets. Data are represented as points with error bars (the scale is set by left axis), while simulated background events are reported with outlined histogram. Signal events, splitted in different production processes are reported as solid filled histograms (the scale is set by right axis). The vertical dashed lines represent the boundaries of untagged event categories. Events in the shaded region at low diphoton BDT output are discarded and not used in the final analysis.



**Figure 4.3.** Transformed diphoton BDT output score for  $Z \rightarrow e^+e^-$  events in 8 TeV data and in MC simulation in which electrons are reconstructed as photons. The histogram represents the simulated events while the points with error bars the data. In the lower plot, for each bin the ratio of number of events in data and to simulated events is shown. The bands represent the systematic uncertainties. Vertical dashed lines represent the boundaries of untagged categories

bosons produced with VH mechanism can be tagged requiring in the final state one or more charged leptons, large  $E_T^{\text{miss}}$ , or jets. Finally events resulting from

$t\bar{t}H$  production have a characteristic signature with b quarks in the final state and additional jets and leptons from W decays.

The  $p_T$  spectrum of the photons coming from the Higgs boson produced by VBF, VH and  $t\bar{t}H$  is significantly harder than for those produced by ggH or background photons. This is due to the fact that the scale of the process is set by the mass of the heavy particles present in the final state. Tighter  $p_T$  requirements are therefore used to reject background processes and to achieve purity in one single production mode.

In this paragraph the optimized selections for each production mode are reviewed. Particular attention will be given to the categories tagging VH production with hadronic decays of the vector boson and  $t\bar{t}H$ . With the creation of these new categories, not present in the  $H \rightarrow \gamma\gamma$  analysis at the time of the discovery, all possible production modes have been exploited for achieving the best sensitivity possible in couplings measurements.

### 4.3.1 VBF Tag

VBF-like events are characterized by the presence of two jets, originating from the two scattered quarks. Jets are first reconstructed as described in section 3.2, then a dijet preselection is applied on events passing the diphoton selection described in the previous paragraph.

The requirements of dijet preselection are:

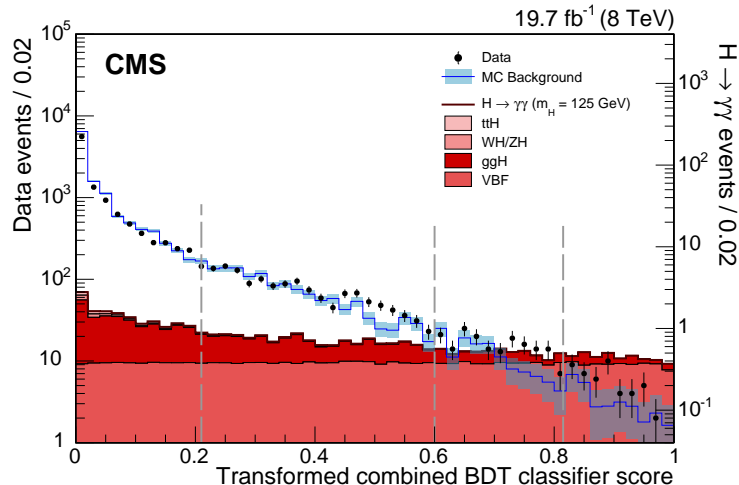
- the leading and subleading jet in the event have to be in the region  $|\eta| < 4.7$  and respectively a  $p_T$  greater than 30 and 20 GeV is required,
- an invariant mass of the dijet system,  $m_{jj}$ , greater than 250 GeV,
- an additional tighter cut on the leading photon:  $p_T^{\gamma^1} > m_{\gamma\gamma}/2$ .

After this preselection a dijet BDT ( $\text{BDT}^{dijet}$ ) is trained to separate VBF signal from diphoton background (and ggH signal) using events satisfying this dijet preselection. The input variables to this BDT are related to diphoton and dijet kinematics to exploit the peculiar topology of VBF events. The input variables used are: the scaled transverse momenta of the photons,  $p_T^{\gamma^1}/m_{\gamma\gamma}$  and  $p_T^{\gamma^2}/m_{\gamma\gamma}$ , the transverse momenta of the leading and subleading jets,  $p_T^{j1}$  and  $p_T^{j2}$ , the dijet invariant mass,  $m_{jj}$ , the difference between the pseudorapidity of the two jets,  $|\Delta\eta_{jj}|$ , the difference between the average pseudorapidity of the two jets and the one of the diphoton system,  $|\eta_{\gamma\gamma} - (\eta_{j1} + \eta_{j2})/2|$  [72], and the absolute difference in the azimuthal angle between the diphoton system and the dijet system,  $\Delta\phi_{\gamma\gamma jj}$ .

The score of the  $\text{BDT}^{dijet}$  is used as input to a combined BDT ( $\text{BDT}^{VBF}$ ), together with the score of the diphoton BDT,  $\text{BDT}^{\gamma\gamma}$ , and the scaled transverse momentum of the diphoton system. This output is used to define the dijet-tagged classes. To define boundaries between categories, the expected uncertainty in the signal strength associated with VBF production mechanism is minimized. Given the limited MC statistics, for the 7 TeV dataset the number of categories is limited to two and the lower boundary of the lowest dijet-tagged class is fixed so that the same efficiency is obtained for VBF events as the 8 TeV dataset. As a result of this

optimization procedure, two VBF-tagged classes are defined for the 7 TeV dataset, while three classes are used for the 8 TeV dataset.

In figure 4.4 the combined BDT for the 8 TeV dataset is shown. Simulated events, data and the VBF-tagged classes boundaries are shown. Events with score below the lower boundary fail the dijet selection, but can be selected by other classes.



**Figure 4.4.** Score of the combined BDT ( $\text{BDT}^{VBF}$ ) for events satisfying the dijet preselection in 8 TeV events. Points with errors represent the data, simulated signal events are separated in four histograms according to the production mode. The outlined histogram is for simulated background events, with the statistical uncertainty reported as a shaded error band. The vertical dashed lines represent the boundaries of the dijet classes, with the leftmost one representing the score below which events are not included in the VBF-tagged classes, but remain candidates for inclusion in other classes. The classifier is transformed such that signal events produced by VBF mechanism have a uniform flat distribution.

### 4.3.2 Tags for VH selection

Associate production of the Higgs boson with a W or Z boson is investigated using optimized categories according to possible decays of the vector boson.

Three kinds of categories are therefore present:

- VH  $\ell$  tag targeting leptonic decays,
- VH  $E_T^{\text{miss}}$  tag for invisible decays and leptonic decays with unreconstructed leptons,
- VH dijet tag for hadronic decays.

Selection requirements for all these classes have been obtained by minimizing the expected uncertainty in the measurement of signal strength of the process targeted, using data in control regions to estimate the background and MC signal samples to estimate signal efficiency.

### 4.3.2.1 VH Leptonic Tag

The categories of VH  $\ell$  tag require the presence of at least a high- $p_T$  electron or muon in the final state, coming from leptonic decays of the vector bosons in VH mechanism.

The tightly selected lepton class (“VH tight  $\ell$ ”) is characterized by the full signature of  $W \rightarrow \ell\nu$  or  $Z \rightarrow \ell^+\ell^-$  decays, requiring the presence of  $E_T^{\text{miss}}$  or another lepton of the same flavor of the first with an opposite sign. In case of one lepton and  $E_T^{\text{miss}}$  the  $p_T$  of the lepton is required to be greater than 20 GeV, while in the dilepton case the  $p_T$  requirement is lowered to 10 GeV. In the second case a cut is applied on the invariant mass of the lepton pairs, requiring it to be between 70 and 110 GeV.

The loose selected lepton class (“VH loose  $\ell$ ”), targeting the same kind of decays, has looser requirements. Indeed, it requires the presence of a single electron or muon only. Additional requirements are used to reduce the background coming from Z prompt production with initial- or final-state radiation. Muons and electrons are required to be separated from the closest photon by  $\Delta R > 1.0$  and the invariant mass of electron and photon pairs is required to be at least 10 GeV away from the Z boson mass. A tighter conversion veto is also applied to electrons to reject those originating from photon conversions.

For both categories the cut on the leading photon has been tightened with respect to preselection requiring  $p_T^{\gamma^1} > 3m_{\gamma\gamma}/8$ .

### 4.3.2.2 VH $E_T^{\text{miss}}$ Tag

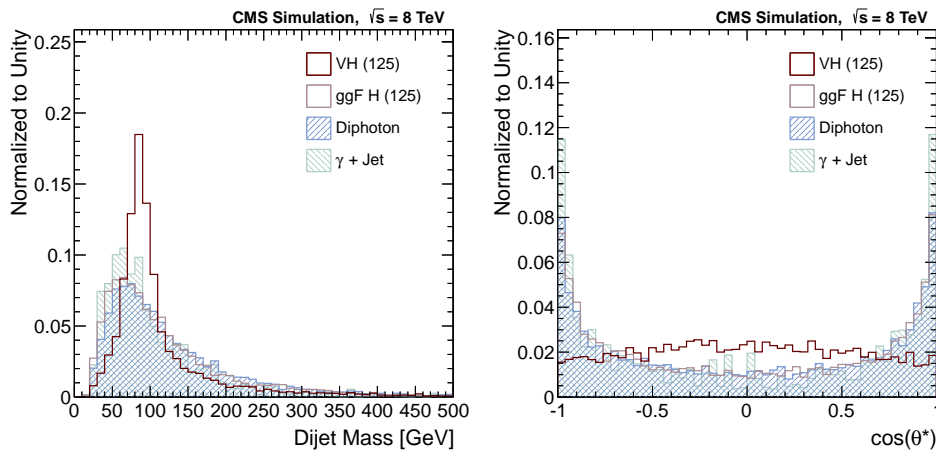
The VH  $E_T^{\text{miss}}$  tag targets W leptonic decays where the lepton from W is not reconstructed, due to experimental or acceptance effects, and invisible decays of the Z boson ( $Z \rightarrow \nu\nu$ ). In both cases, large  $E_T^{\text{miss}}$  is expected in the final state.

Since the direction of the Higgs candidate and  $\vec{E}_T^{\text{miss}}$ , in absence of additional radiated jets, are balanced in the transverse plane, an additional cut is applied on the azimuthal angular separation between the diphoton system and  $E_T^{\text{miss}}$ :  $|\Delta\Phi_{\gamma\gamma E_T^{\text{miss}}}| > 2.1$ . Discrepancies between data and simulated events in direction and magnitude of  $E_T^{\text{miss}}$  vector have been carefully studied and a set of corrections, to simulations or data, are applied to achieve good agreement. In addition to this, a cut on the direction of the diphoton system and the leading jet in the event is applied, requiring  $|\Delta\Phi_{\gamma\gamma j1}| < 2.7$  to get rid of tails in  $E_T^{\text{miss}}$  distribution where the agreement of data and simulation is not sufficient. The corrected  $E_T^{\text{miss}}$  is required to satisfy  $E_T^{\text{miss}} > 70$  GeV.

### 4.3.2.3 VH Hadronic Tag

The selection used in this category aims at a full reconstruction of VH events with hadronic decays of the vector boson. This category attempts to reconstruct the full final state of the decay, therefore explicitly requires the presence of two jets in addition to the two photons.

Background in this channel is mainly constituted by diphoton events produced in association with two hard jets arising from pile up, radiation, or underlying event activity.



**Figure 4.5.** Dijet mass (left) and  $\cos(\theta^*)$  (right) distributions. The expected shape of signal events (dark brown line) is compared to the shape of gluon-gluon fusion Higgs boson production (light brown line) and the major backgrounds: diphoton continuous production (blue) and photon+jet (green). All distributions are normalized to unity.

At least a pair of jets with  $p_T > 40 \text{ GeV}$  is required within the region  $|\eta| < 2.4$ . Since these two jets in the signal come from vector boson decay, their invariant mass is resonant in the boson mass. This is not true for backgrounds where selected jets are almost uncorrelated. In the left plot of Figure 4.5 the distribution of dijet invariant mass in simulated events is shown for the VH and ggH production modes and for the main backgrounds of the analysis (continuum diphoton prompt production and  $\gamma + \text{jets}$ ). It can be seen that requiring the dijet invariant mass to be within the range  $60 < m_{jj} < 120 \text{ GeV}$  is almost fully efficient on signal while it suppresses background and ggH contributions.

Strong handles to fight the background to this channel come from the total reconstruction of the  $V^* \rightarrow VH$  decay chain. The angle  $\theta^*$  that the diphoton system makes, in the diphoton-dijet centre-of-mass frame, with respect to the direction of motion of the diphoton-dijet system in the lab frame, is computed for this purpose. The distribution of  $\cos \theta^*$  is expected to be flat for VH events, since a two-body decay is isotropic. In the case of background events and ggH production, where the dijet and the diphoton system do not have the same correlations, the distribution of  $\cos \theta^*$  is expected to peak at  $|\cos \theta^*| = 1$ , since emission of jets in the boost direction is favoured for radiation. As can be seen from the right plot of figure 4.5, this variable proves to be strongly discriminating between signal and background, and is expected to be faintly correlated with other kinematic variables of the event. Consequently  $|\cos \theta^*| < 0.5$  is required.

As in other VH tags harder cuts on photons are used:  $p_T > m_{\gamma\gamma}/2$  and  $p_T^{\gamma\gamma} > 13m_{\gamma\gamma}/12$  is required.

### 4.3.3 Tags for $t\bar{t}H$ selection

The production of the Higgs boson in association with a pair of top quarks represents a unique opportunity to probe directly the top-Higgs coupling. Given the low cross

section, only a handful of events is expected from this production mode.

The final state is characterized by the presence of at least two b-quarks and additional jets or leptons coming from W decays. Thanks to this very characteristic topology, a selection based on jets and b-jets multiplicity allows to separate  $t\bar{t}H$  from background and other Higgs production mechanisms. To maximize signal efficiency, selections collecting both leptonic (one or more leptons in the final state) and hadronic (no leptons in the final state) decays of the top quarks are devised, defining a lepton-tagged and a multijet-tagged  $t\bar{t}H$  class.

The main backgrounds in  $t\bar{t}H$  searches are events with top quarks with either genuine or misidentified photons and the production of high- $p_T$  photons in association with many jets. Background contribution will be estimated using data which contains a mixture of these processes. It is useful, for optimization purposes, to test the background model in a data-driven control sample. The control sample is constructed using events where one photon fails the photon identification requirements for the Higgs boson signal. To take into account the fact that efficiency and photon isolation are not constant as a function of  $p_T$  and  $\eta$ , a two-dimensional reweighting procedure is applied for such events. The reweighting is performed so as to match the photon  $p_T$  and  $\eta$  spectra to the ones of photons populating the signal region. A control sample with similar kinematic properties as the data, yet statistically independent, is thus obtained.

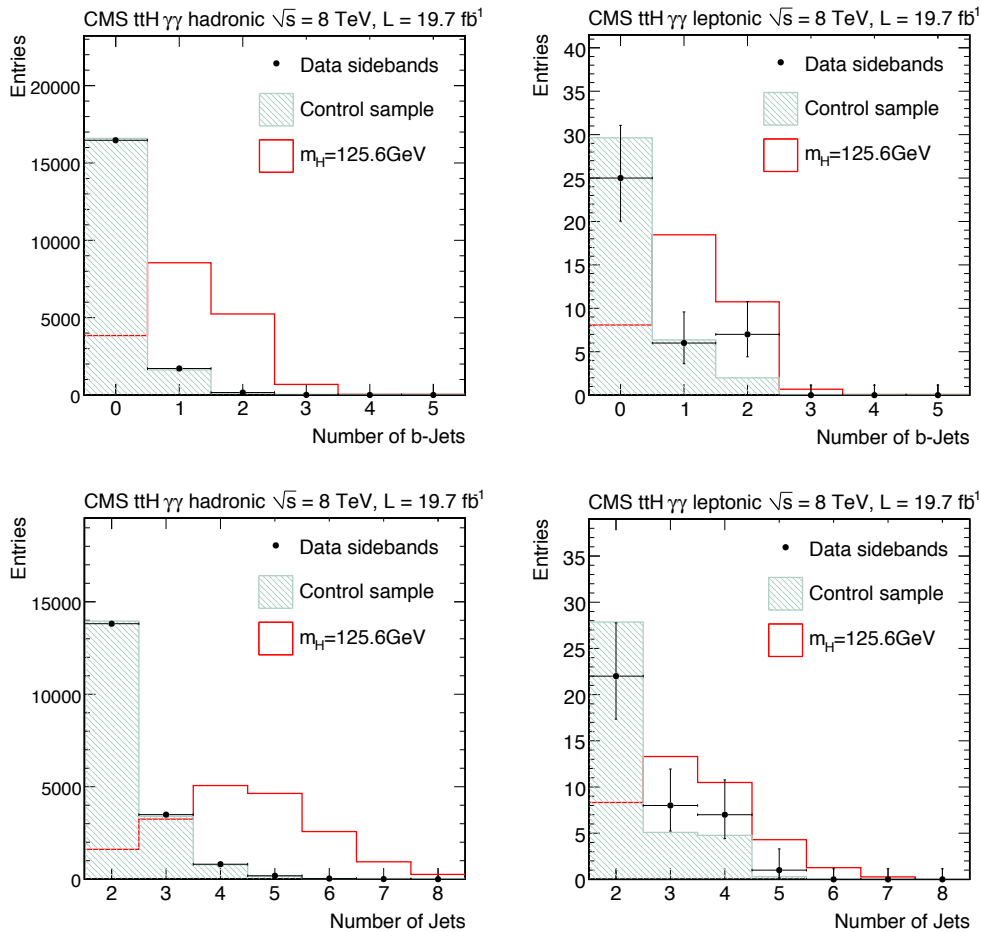
This control sample can be used to model the selection according to the difference in the number of jets and b-jets between signal and background events. In figure 4.6 these two key kinematic distributions are shown, requiring at least two high  $p_T$  jets in the final state. Events are split requiring a lepton or not in the final state. The black markers show the data sample, the green histogram is the control sample, and the red line represents simulated  $t\bar{t}H$  events. All distributions are normalized to the number of events observed in data for shape comparison purposes. In these plots, events where the diphoton invariant mass is consistent with the Higgs boson mass within a 10 GeV window are removed in data and control sample.

From these distributions, it can be seen that, given the large jet activity in  $t\bar{t}H$ , hard cuts on these variables will provide good discrimination between signal and background. According to these considerations separate optimizations are provided for leptonic and hadronic tag. For both categories the presence of two photons with the leading one passing the cut  $p_T > m_{\gamma\gamma}/2$  is required. Jets are required to have  $p_T > 25$  GeV and both classes require the presence of at least one additional b-tagged jet.

The lepton tag is then defined requiring at least one more jet in the event and at least one electron or muon with  $p_T > 20$  GeV. The multijet tag is defined instead by the requirement of at least four more jets in the event and no leptons. For the 7 TeV dataset, due to limited number of events, the events in the two classes are combined after selection to form a single  $t\bar{t}H$  event class.

#### 4.3.4 Signal Model

For each event class a parametric signal model must be constructed for each production mechanism from a fit to simulated invariant mass distribution of  $H \rightarrow \gamma\gamma$  events. All corrections, described in the previous chapters, derived in  $Z \rightarrow e^+e^-$

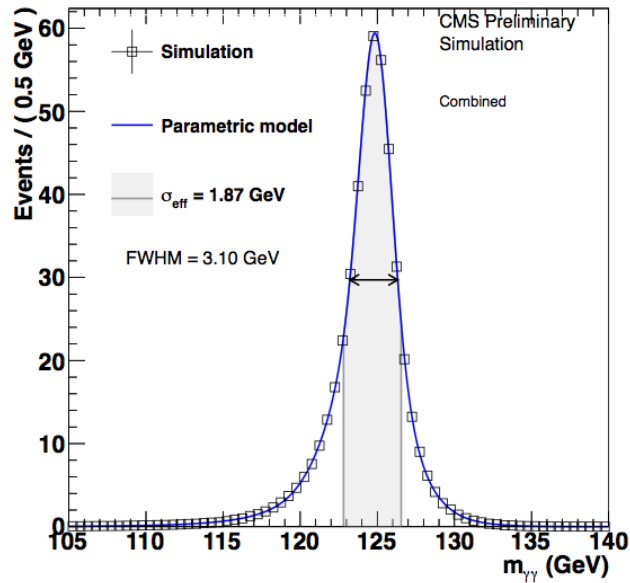


**Figure 4.6.** Distributions of the b-tagged jet multiplicity (top row) and jet multiplicity (bottom row) for events passing a selection which requires two photons and at least two jet in the final state, in the hadronic (left) and leptonic (right) channels. The plots compare the data events (black markers) and the data from the control sample (green filled histogram) to simulated  $t\bar{t}H$  events (red open histogram). For events in the control sample only one of the two photons passes the stringent identification requirements of the standard  $H \rightarrow \gamma\gamma$  selection. Both signal and background histograms are normalized to the total number of data events observed in this region to allow for a shape comparison.

and  $Z \rightarrow \mu^+\mu^-\gamma$  are applied. Using a sum of gaussian functions, a good description of the distributions, including the tails, can be achieved. Up to five gaussians are used depending on the category and production mechanism, even if in most cases two or three are enough for a good fit. In figure 4.7 the result of the fit to simulated data is shown combining together all the categories of the analysis.

The overall acceptance times efficiency for this analysis is close to 50% for both 7 and 8 TeV datasets.

Thanks to the strategy presented in this thesis, thanks the new corrections and regressions used in photon reconstruction, the mass resolution achieved improved significantly with respect to analyses of this decay mode previously published by



**Figure 4.7.** Full parametrized signal shape integrated over all classes in simulated  $H \rightarrow \gamma\gamma$  events. The black points are weighted Monte Carlo events while the blue lines correspond to the analytic model. The effective  $\sigma_{\text{eff}}$  coming from the fit and the corresponding interval are also shown in grey.

CMS [13]. For events in which both photons are in the barrel, the width of the mass distributions has been reduced by around 5% in 7 TeV data and by 20% in 8 TeV data. When at least one photon is in the endcap the improvement is around 20% in 7 TeV data, and more than 30% in 8 TeV data.

#### 4.4 Fit Methodology

A simultaneous binned maximum-likelihood fit to the diphoton invariant mass distributions is performed in all event classes to extract results and measurements of this analysis. The bin size chosen is small enough, 250 MeV, compared to mass resolution, that no information is lost in the fit procedure. Signal model is derived through fits to simulated events as explained in the previous paragraph. Background is evaluated fitting  $m_{\gamma\gamma}$  in data in the range  $100 < m_{\gamma\gamma} < 180$  GeV, without any use of MC simulations.

The likelihood in the fit, evaluated in the signal-plus-background hypothesis, will be therefore:

$$\mathcal{L} = \mathcal{L}(\text{data} | s(p, m_{\gamma\gamma}) + f(m_{\gamma\gamma})) \quad (4.1)$$

where  $p$  represents parameters of the signal, such as  $m_H$  or the signal strength, that are free to vary in the fit procedure,  $s(p, m_{\gamma\gamma})$  is the parametric signal model and  $f(m_{\gamma\gamma})$  the background fit function.

Profile likelihood ratio is used as test statistic and systematic uncertainties are incorporated as nuisance parameters and treated according to the frequentist



paradigm. The full description of this methodology widely used in particle physics can be found in [73,74] and the results are obtained using the ROOTS package [75].

The background fitting function is not a single function, but the so called “envelope method” is used in this analysis. This method consists in a discrete profiling method, where the uncertainty from the choice of the function used to fit the background is handled in a similar way to systematic uncertainties. The choice of the background fit function, in each class, is included as a discrete nuisance parameter in the likelihood used to extract the result. The families of functions considered in this process cover a wide spectra of possible behavior of the data points: exponentials, power-law functions, polynomials in the Bernstein basis and Laurent series. In the likelihood minimization process, all functions in these families are tried and a penalty term is added to account for the number of free parameters in the fitting functions.

The penalized likelihood function,  $\tilde{\mathcal{L}}_f$ , is therefore defined, for each fitting function  $f$ , as:

$$-2 \ln \tilde{\mathcal{L}}_f = -2 \ln \mathcal{L}_f + k N_f, \quad (4.2)$$

where  $\mathcal{L}_f$  is the unpenalized likelihood function defined in equation 4.1,  $N_f$  is the number of free parameters in  $f$ , and  $k$  is a constant. The likelihood ratio  $q(p)$ , is used when measuring a quantity  $p$ . In this formalism, it is given by:

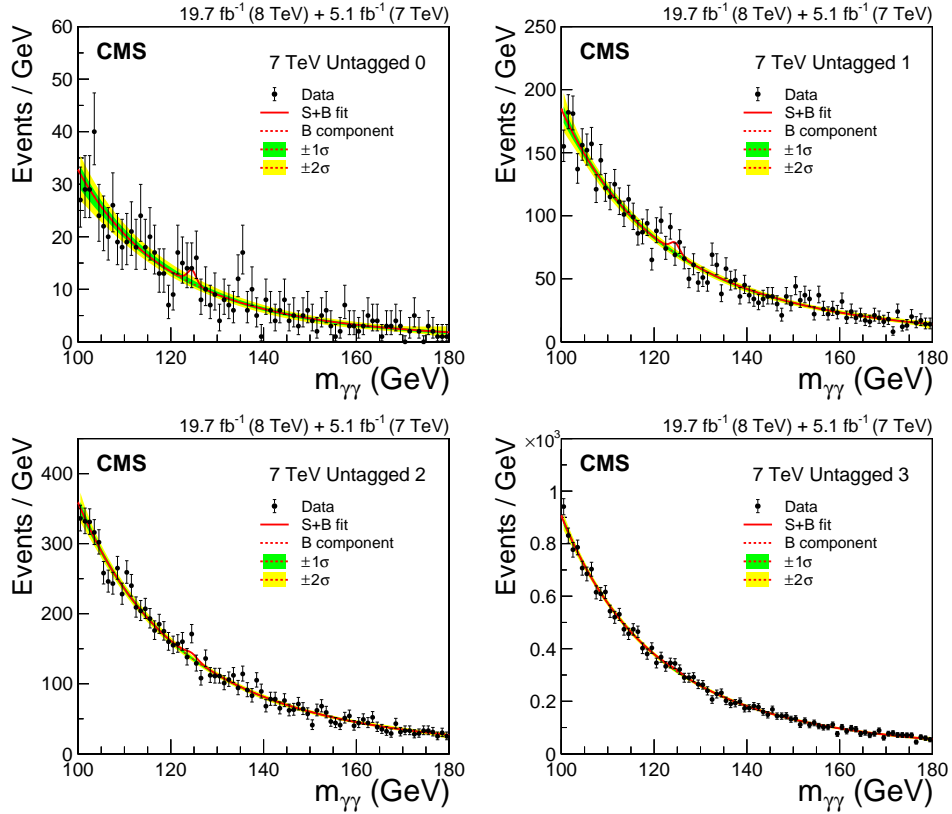
$$q(p) = -2 \ln \frac{\tilde{\mathcal{L}}(\text{data}|p, \hat{\theta}_p, \hat{f}_p)}{\tilde{\mathcal{L}}(\text{data}|\hat{p}, \hat{\theta}, \hat{f})}, \quad (4.3)$$

where the numerator is the likelihood function computed in the maximum for a given  $p$ , achieved for the best-fit values of nuisance parameters,  $\theta = \hat{\theta}_p$ , and a particular background function,  $f = \hat{f}_p$ . The denominator is instead the likelihood function computed in the global maximum, i.e. for  $p = \hat{p}$ ,  $\theta = \hat{\theta}$ , and  $f = \hat{f}$ .

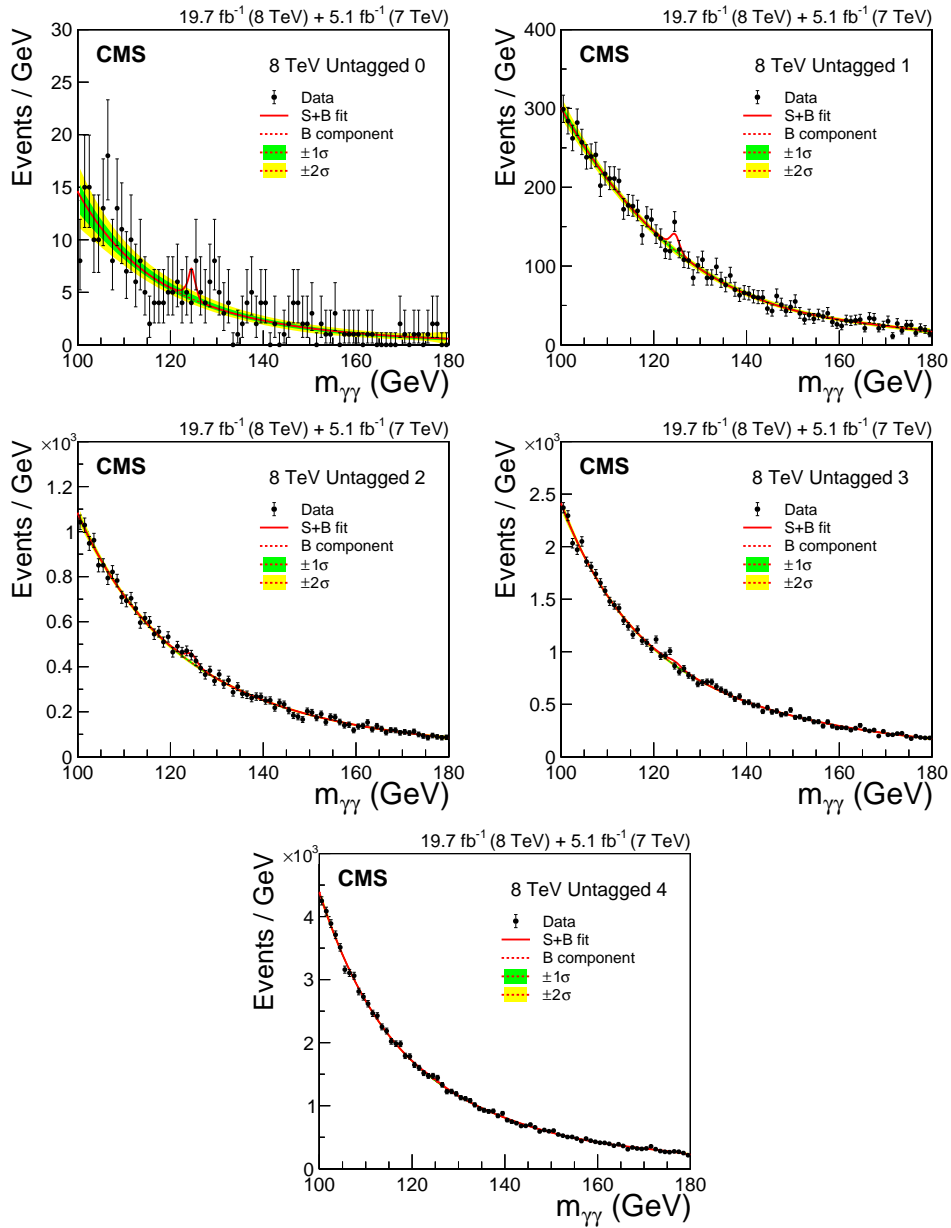
The  $k$  in equation 4.2 sets the relative importance between the unpenalized term and the penalty term for increasing number of degrees of freedom. The values  $k = 1$  and  $k = 2$  are both justified by theory [76]. Studies made using pseudo-experiments confirmed that using  $k = 1$  gives consistent results and negligible bias.

The  $m_{\gamma\gamma}$  distributions of categories of 7 and 8 TeV analysis, together with the results of the fits in the signal-plus-background hypothesis are shown in figures 4.8-4.15. The signal-plus-background fit is done simultaneously to all classes with a single overall value of the signal strength  $\mu$  free to vary. The  $1\sigma$  and  $2\sigma$  uncertainty bands are computed from the variation in pseudo-experiments on the fitted background yields. These bands include uncertainty coming from the choice of the background fit function and the uncertainty in the fitted parameters.

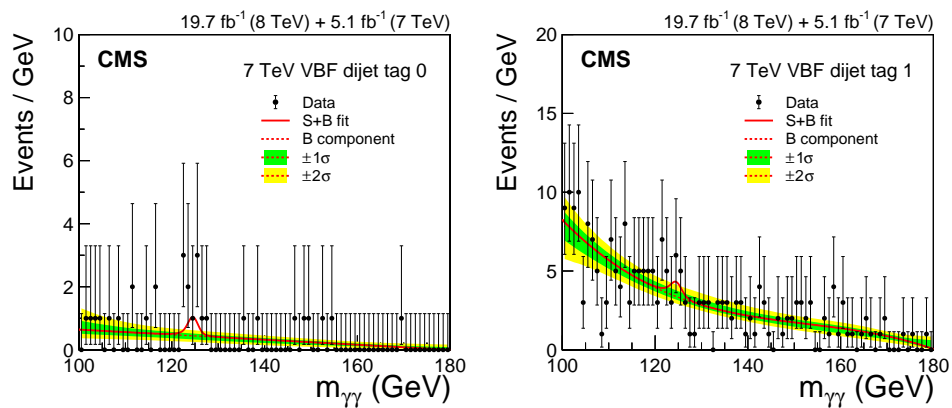
In Figure 4.16 event display for events in VH hadronic and  $t\bar{t}H$  categories are shown in a 3D view of the CMS detector. The peculiar signature of this kind of events, jet activity in the tracker and in the calorimetric system, is clearly visible. The chosen events have a value of  $m_{\gamma\gamma}$  close to 125 GeV and are therefore compatible with Standard Model Higgs production.



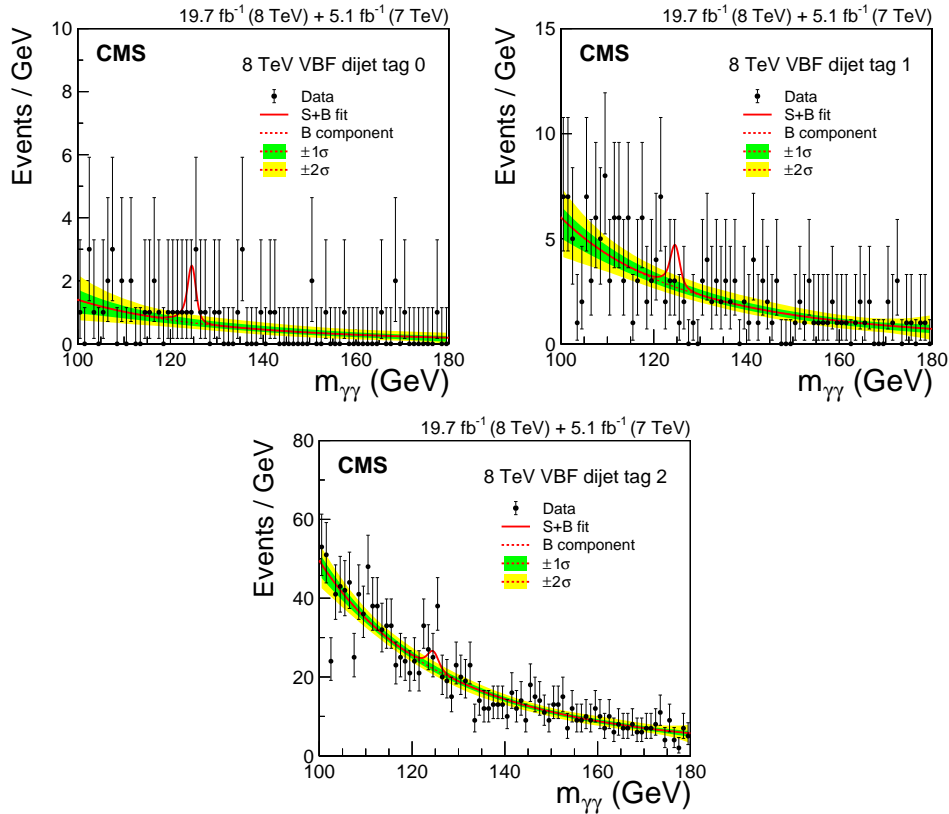
**Figure 4.8.** Diphoton invariant mass distributions for the events in the four untagged classes of the 7 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



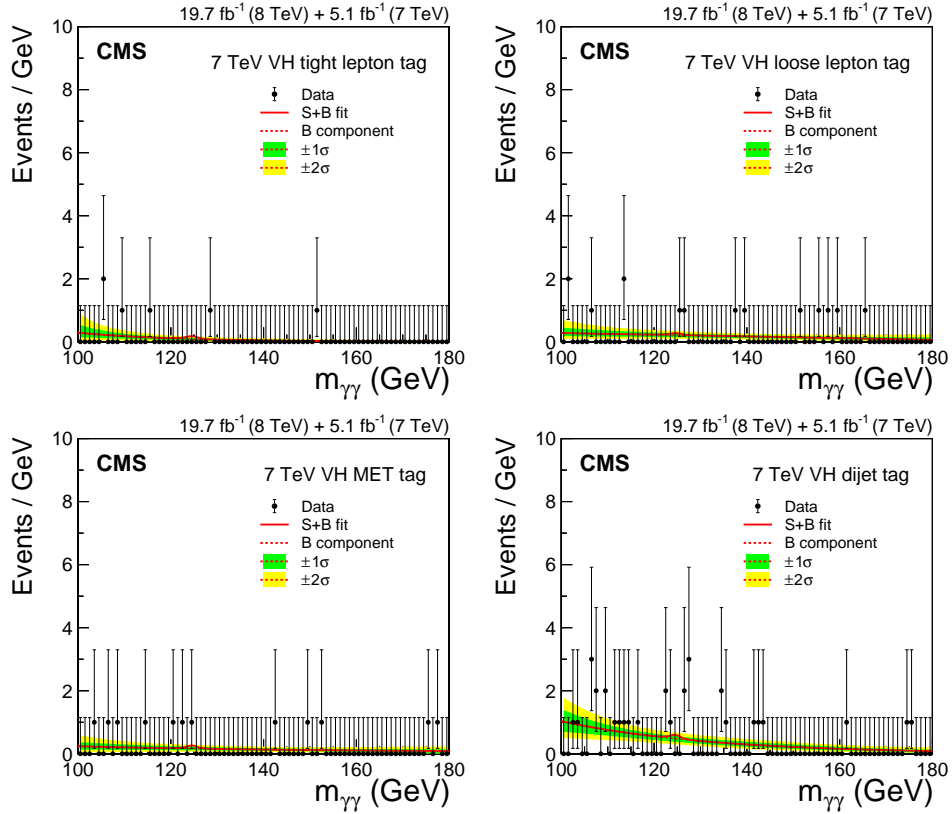
**Figure 4.9.** Diphoton invariant mass distributions for the events in the five untagged classes of the 8 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



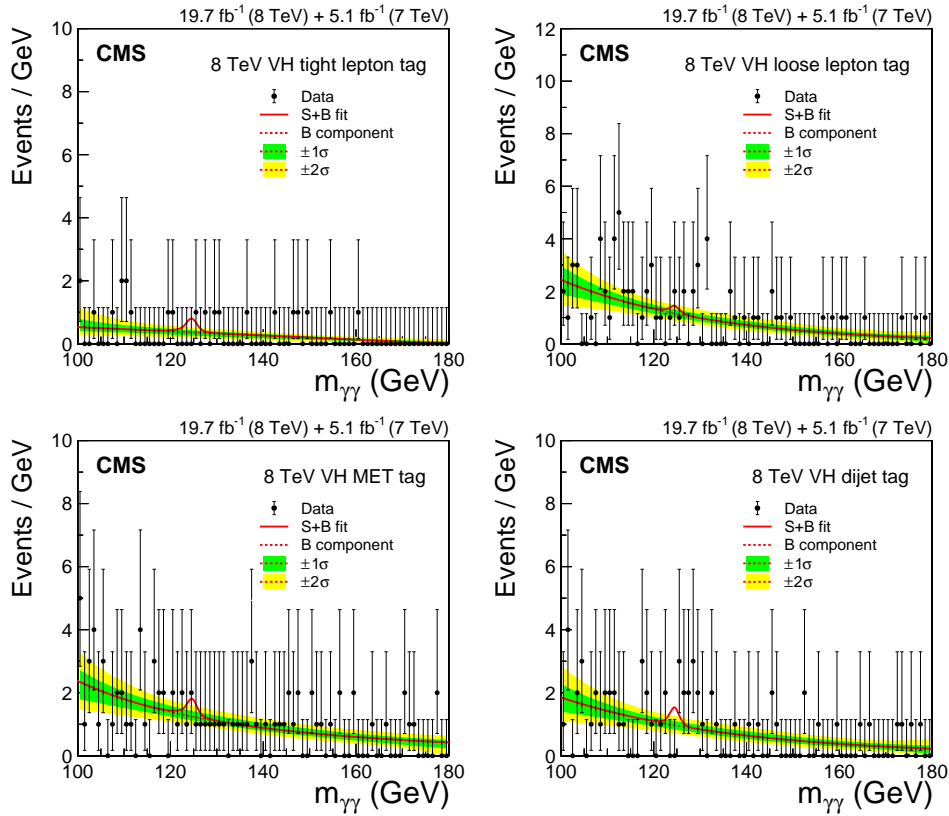
**Figure 4.10.** Diphoton invariant mass distributions for the events in the two VBF dijet-tagged classes of the 7 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



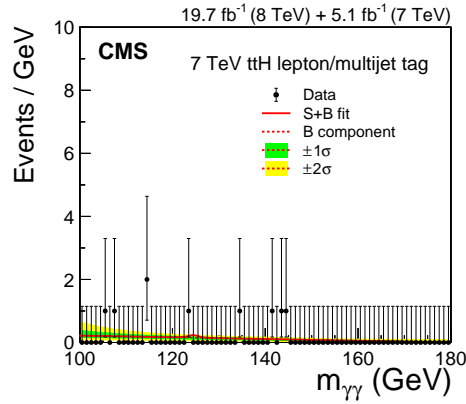
**Figure 4.11.** Diphoton invariant mass distributions for the events in the three VBF dijet-tagged classes of the 8 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



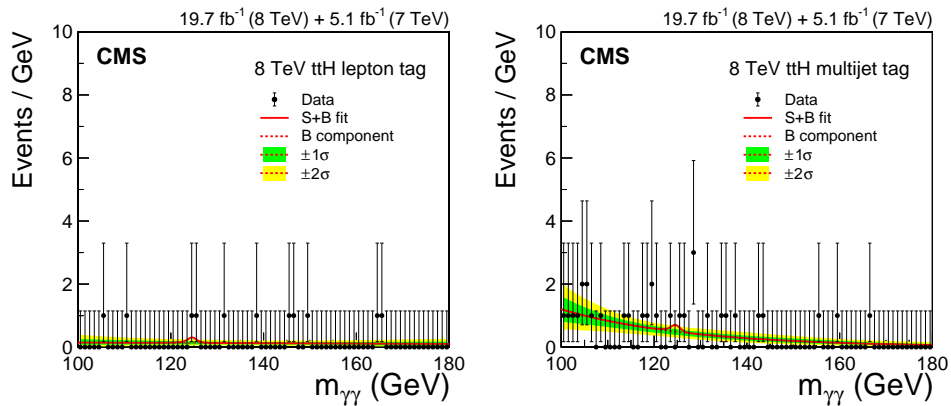
**Figure 4.12.** Diphoton invariant mass distributions for the events in the VH-tagged classes of the 7 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



**Figure 4.13.** Diphoton invariant mass distributions for the events in the VH-tagged classes of the 8 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit are computed from the fit uncertainty in the background yield in bins corresponding to those used to display the data. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



**Figure 4.14.** Diphoton invariant mass distributions for the events in the  $t\bar{t}H$ -tagged class of the 7 TeV dataset, together with the result of a fit of the signal-plus-background model for  $m_H = 124.7$  GeV. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



**Figure 4.15.** Diphoton invariant mass distributions for the events in the two  $t\bar{t}H$ -tagged classes of the 8 TeV dataset, together with the result of a fit of the signal-plus-background model. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit are computed from the fit uncertainty in the background yield in bins corresponding to those used to display the data. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated.



## 4.5 Systematic Uncertainties

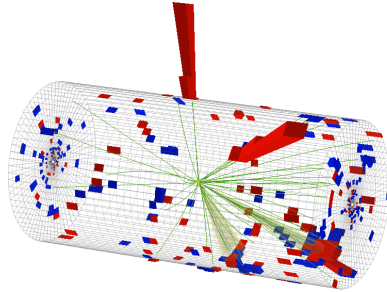
Different sources of systematic uncertainties are considered in this analysis. An overview of the main systematic uncertainties is given here grouping the effects in three groups: systematic uncertainties assigned to all events, uncertainties related to individual photons and uncertainties related to additional objects in classes targeting exclusively production modes. The overall effect of these systematic uncertainties on the measurement of signal strength modifier  $\mu$  is reported as an example in Table 4.2.

Systematic uncertainties common to all events are:

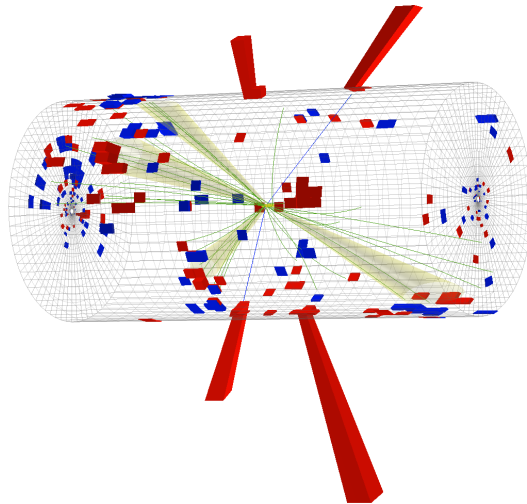
- *Theory uncertainties:* this kind of uncertainties is related to parton distribution functions of protons (PDF), the QCD scale chosen, production cross sections and diphoton branching ratio. They are evaluated following the recommendations of the LHC Higgs Cross Section Working Group [30, 32]. This kind of uncertainties are dominated by uncertainty on the ggH process cross section, coming from missing higher orders calculations and uncertainties on parton distribution functions. They range from 1 to 15%, depending on the production process.
- *Integrated luminosity:* This uncertainty affects the signal yields in both 7 and 8 TeV dataset. It has been evaluated as described in [77, 78] and amounts to 2.2% for 7 TeV and 2.6% for 8 TeV.
- *Vertex finding efficiency:* Since data/MC scale factors are measured using  $Z \rightarrow \mu^+ \mu^-$  events, the uncertainty on this derivation is taken into account as systematic uncertainty. An additional 1%, related to the amount of activity resulting in charged tracks in signal events, estimated varying PYTHIA tunes, is added. A further uncertainty of 0.2% is finally added since vertex-finding efficiency depends on  $p_T^{\gamma\gamma}$  and  $p_T$  measurement has its proper uncertainty.
- *Trigger efficiency:* Uncertainty in trigger efficiency is evaluated using a tag-and-probe technique in  $Z \rightarrow e^+ e^-$  events and amounts roughly to 1%.

Systematic uncertainties related to individual photons are:

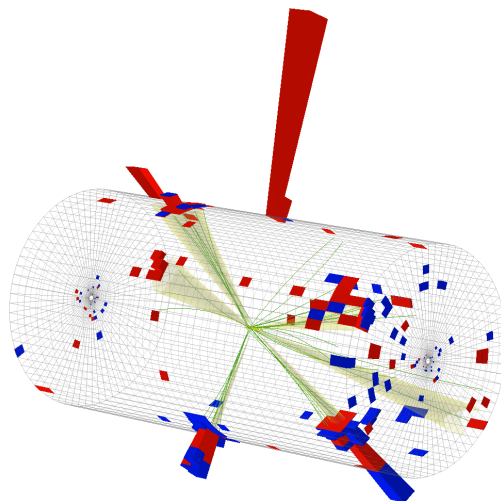
- *Photon energy scale uncertainty resulting from electron/photon differences:* Energy scale corrections are evaluated using  $Z \rightarrow e^+ e^-$  studies and then propagated to the Higgs signal. Differences between electrons and photons in MC simulations can arise from different causes. The dominating one is an imperfect description of the material between the interaction point and ECAL. In particular, CMS studies suggest a deficit of material in the simulation, probably localized in specific regions. The data/MC discrepancies correspond to what would be caused by a 10% uniform deficit of material in the region  $|\eta| < 1.0$  and a 20% uniform deficit in the region  $|\eta| > 1.0$ . The resulting uncertainty on energy scale has been therefore evaluated using samples in which the tracker material is artificially increased respectively by 10 and 20% in different  $\eta$  bins. They range from 0.03% in central barrel up to 0.3% in the outer endcap.



CMS Experiment at LHC, CERN  
Data recorded: Thu Jul 26 06:45:25 2012 CEST  
Run/Event: 199699 / 380253151  
Lumi section: 313



CMS Experiment at LHC, CERN  
Data recorded: Thu Nov 1 02:13:01 2012 CEST  
Run/Event: 206446 / 1072391444  
Lumi section: 784



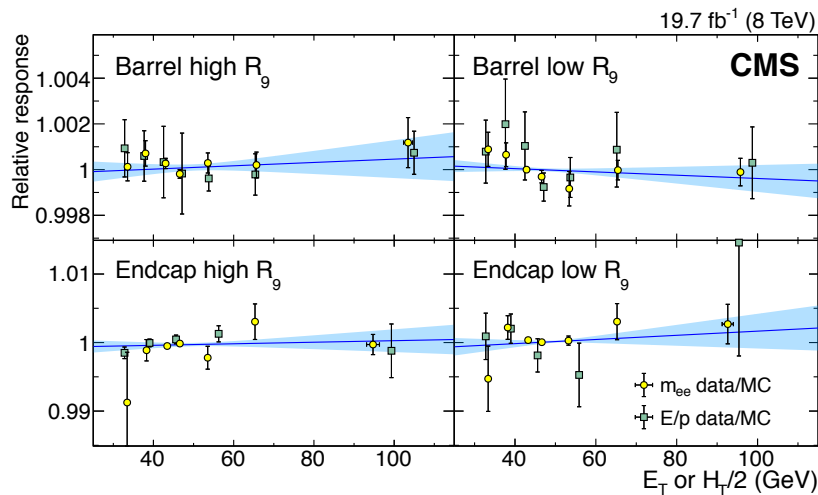
CMS Experiment at LHC, CERN  
Data recorded: Sat Nov 24 19:16:36 2012 CEST  
Run/Event: 207889 / 771018991  
Lumi section: 783

**Figure 4.16.** Event displays for events in VH hadronic (top),  $t\bar{t}H$  leptonic (center) and  $t\bar{t}H$  hadronic (bottom) categories.

Another source of uncertainties related to electron/photon differences is the modelling of the varying fraction of scintillating light reaching the photodetector as a function of the longitudinal depth in the crystal. Crystals were produced with a good degree of uniformity, however the level of uniformity changed slightly during crystals production. Furthermore uniformity is affected by transparency loss caused by radiation [79,80]. The simulation of this effect highlighted a difference in the energy scale between electrons and unconverted photons which leads to an additional uncertainty of about 0.05%.

A further small uncertainty originates from different possible modelizations of electromagnetic shower, by the GEANT4 software, that change the difference between photon and electron energy scale, leading to an effect of 0.05%.

- *Energy scale nonlinearity*: Differences between simulation and data are estimated using electrons coming from  $Z \rightarrow e^+e^-$  decays. Therefore, possible non linear effects have to be considered in the extrapolation from energies typical of electrons in  $Z \rightarrow e^+e^-$  decays to those of photons from  $H \rightarrow \gamma\gamma$ . These effects have been investigated checking the  $Z \rightarrow e^+e^-$  peak position by binning events according to  $H_T$  (the scalar sum of  $E_T$  of the two electron showers), and measuring, in bins of  $E_T$ , the ratio between the measured energy and the measurement of the electron momentum from the tracker on  $Z \rightarrow e^+e^-$  and  $W \rightarrow e\nu$  events. The data/MC ratio of these two quantities as a function of  $E_T$  or  $H_T/2$  is shown in figure 4.17 for the four categories in  $\eta$  and  $R_9$  of photon identification. Linearity assumption has been checked using a parabola as a fitting function, and uncertainties are found to be of the order of 0.1%.



**Figure 4.17.** Differential non-linearity in the four photon identification categories. Each graph shows the ratio of the estimated position of the E/p (green) or of the  $m_{ee}$  (yellow) invariant mass peaks in data to the corresponding position estimated in MC simulation. On the abscissa, the  $E_T$  ( $H_T/2$ ) of the electron is shown for the E/p (mee analysis).

- *Energy scale in data and MC measurement and corrections*: Additional systematic uncertainties arise from the derivation of energy scale and resolution in

data with electrons coming from  $Z \rightarrow e^+e^-$ , with the methodology described in section 3.1.2, related to imperfect data-MC agreement. These are derived in bins of  $\eta$  and  $R_9$  and range from 0.05% for unconverted photons in central barrel up to 0.1% for endcap converted photons. Energy dependence of the smearings and possible misdescription of the  $Z \rightarrow e^+e^-$  line-shape are also considered.

- *Photon identification BDT score and estimate of per photon energy resolution:* As described in section 3.1.5 the agreement between data and simulation of input variables to the diphoton BDT is checked on  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  events. Among the input variables to the BDT, the distribution of the photon identification BDT score and per-photon energy resolution show non-negligible differences between data and simulation. These differences are fully covered by a variation of  $\pm 0.01$  on the photon identification BDT together with an uncertainty in the per-photon energy resolution parametrized as a rescaling of the resolution estimate by  $\pm 10\%$  about its nominal value. These uncertainties are reported in Table 4.2 as “shower shape modeling”.

Systematic uncertainties related to additional objects in the final state for exclusive categories are:

- *Jet tagging efficiency:* Jets are present in different categories tagging VBF, VH and  $t\bar{t}H$  production modes. Theoretical uncertainties on jet production, concerning the probability of having additional jets in gluon-gluon production mechanism are the dominating uncertainties. Uncertainties on the yields are quantified using the recommendations of the LHC Higgs Cross Section Working Group.

An additional uncertainty is considered in the yield of ggH events in  $t\bar{t}H$  multijet-tagged class, arising from the uncertainty in the probability of gluon splitting in a pair of b-quarks. It is estimated from data-POWHEG simulations discrepancies in  $t\bar{t}$ +jets samples with two charged leptons in the final state [81]. For  $t\bar{t}H$  multijet category, since few events from simulated ggH events are selected, an additional contribution for the limited sample size is considered.

For VBF dijet-tagged class, VH dijet-tagged class and  $t\bar{t}H$  multijet-tagged class an uncertainty in the effect of the algorithm used to reject jets from pileup is considered (in the 8 TeV dataset only).

Further small experimental uncertainties are due to jet energy scale and resolution corrections. They are evaluated by varying the measured jet quantities within their uncertainties and taking the relative yield variation.

- *Lepton identification efficiency:* For both electron and muons this uncertainty is computed varying the data/MC scale factors by their uncertainties. The resulting differences range from 0.2% to 0.5% depending on the category.
- *$E_T^{\text{miss}}$  identification efficiencies:* For WH production mode where real  $E_T^{\text{miss}}$  is expected the difference in efficiency (2.6%) applying or not the  $E_T^{\text{miss}}$  corrections is considered as a systematic uncertainty. For ggH, VBF and  $t\bar{t}H$  events the

fraction of events in the tail of the  $E_T^{\text{miss}}$  distribution, causes an additional uncertainty. This is evaluated by comparing diphoton data and simulated events in control samples enriched by  $\gamma + \text{jets}$  events, which have a similar  $E_T^{\text{miss}}$  distribution to the Higgs signal. The overall effect amounts to 4%.

- *b-tagging efficiency*: The uncertainty on b-tagging efficiency has to be considered in ttH-tagged classes where at least one b-tagged jet is required. It is evaluated varying the b-tagging efficiency scale factors within their uncertainty. The resulting uncertainty is 1.3% in the lepton-tagged class and 1.1% in the multijet one.

**Table 4.2.** Magnitude of the uncertainty in the best fit signal strength,  $\mu$ , induced by the systematic uncertainties in the signal model. Statistical uncertainty is subtracted quadratically from the positive and negative uncertainties separately. The values quoted are the average magnitudes of the positive and negative uncertainties. The statistical uncertainty includes all uncertainties in the background modelling.

Source of uncertainty	Uncertainty in $\hat{\mu}$
PDF and theory	0.11
Shower shape modelling	0.06
Energy scale and resolution	0.02
Other	0.04
All syst. uncert. in the signal model	0.13
Statistical	0.21
Total	0.25

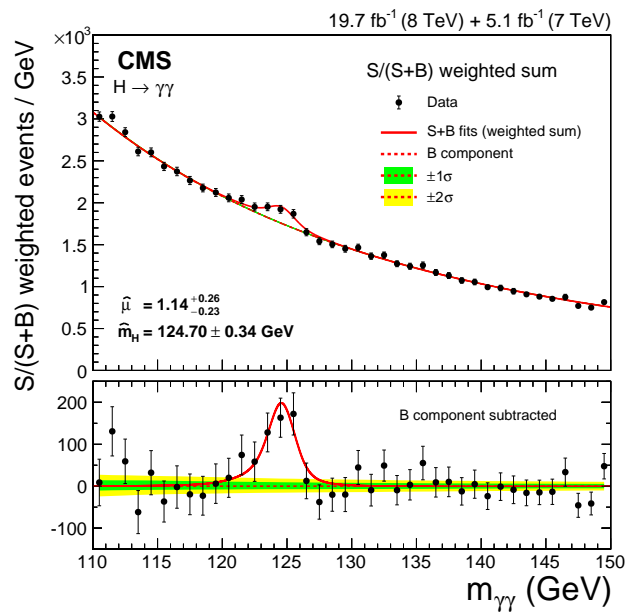
## 4.6 Results

### 4.6.1 Standard Model Analysis Results

Figure 4.18 shows the  $m_{\gamma\gamma}$  distribution of the combination of all classes of 7 and 8 TeV analysis. Events are weighted by the ratio  $S/(S+B)$  in each event class, to take into account that the analysis is performed through a simultaneous fit to different categories [82]. In the lower plot the background component is subtracted and a clear excess is visible in data around  $m_{\gamma\gamma} = 125$  GeV.

The best-fit mass is  $m_H = 124.7$  GeV. Table 4.3 shows the number of expected signal events for a 125 GeV SM Higgs boson and background estimation for each event class. The background estimation is obtained from the fit and it is expressed as a differential rate (events/GeV). Signal is splitted for different production processes (as predicted by MC simulations). For each category the mass resolution, measured both by half the width of the narrowest interval containing 68.3% of the  $m_{\gamma\gamma}$  distribution ( $\sigma_{\text{eff}}$ ) and half the full width at half maximum of the distribution divided by 2.35 ( $\sigma_{\text{HM}}$ ) is reported.

The local p-value, which quantifies the probability for the background to produce a fluctuation of such magnitude, is shown in figure 4.19. The p-value has been

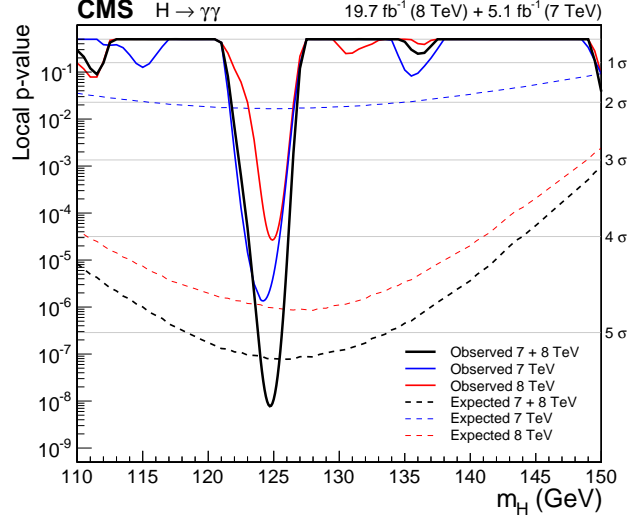


**Figure 4.18.** Combined diphoton mass spectrum weighted by the ratio  $S/(S+B)$  in each event class, together with the background subtracted weighted mass spectrum. The  $1\sigma$  and  $2\sigma$  uncertainty bands shown for the background component of the fit include the uncertainty due to the choice of function and the uncertainty in the fitted parameters. These bands do not contain the Poisson uncertainty that must be included when the full uncertainty in the number of background events in any given mass range is estimated. The lower plot shows the residual data subtracting the fitted background component.

**Table 4.3.** Expected number of SM Higgs boson events ( $m_H = 125$  GeV) and estimated background (“Bkg.”) at  $m_{\gamma\gamma} = 125$  GeV for all event classes of the 7 and 8 TeV datasets. The composition of the SM Higgs boson signal in terms of the production processes and its mass resolution is also given. The number corresponding to the production process with the largest contribution in each event class is highlighted in boldface. Numbers are omitted for processes representing less than 0.05% of the total signal. The variables used to characterize the mass resolution,  $\sigma_{\text{eff}}$  and  $\sigma_{\text{HM}}$ , are defined in the text.

Event classes		Expected SM Higgs boson signal yield ( $m_H=125$ GeV)							Bkg. (GeV <sup>-1</sup> )	
		Total	ggH	VBF	WH	ZH	ttH	$\sigma_{\text{eff}}$ (GeV)		$\sigma_{\text{HM}}$ (GeV)
7 TeV 5.1 fb <sup>-1</sup>	Untagged 0	5.8	<b>79.8%</b>	9.9%	6.0%	3.5%	0.8%	1.11	0.98	11.0
	Untagged 1	22.7	<b>91.9%</b>	4.2%	2.4%	1.3%	0.2%	1.27	1.09	69.5
	Untagged 2	27.1	<b>91.9%</b>	4.1%	2.4%	1.4%	0.2%	1.78	1.40	135.
	Untagged 3	34.1	<b>92.1%</b>	4.0%	2.4%	1.3%	0.2%	2.36	2.01	312.
	VBF dijet 0	1.6	19.3%	<b>80.1%</b>	0.3%	0.2%	0.1%	1.41	1.17	0.5
	VBF dijet 1	3.0	38.1%	<b>59.5%</b>	1.2%	0.7%	0.4%	1.65	1.32	3.5
	VH tight $\ell$	0.3	—	—	<b>77.2%</b>	20.6%	2.2%	1.61	1.31	0.1
	VH loose $\ell$	0.2	3.6%	1.1%	<b>79.1%</b>	15.2%	1.0%	1.63	1.32	0.2
	VH $E_T^{\text{miss}}$	0.3	4.5%	1.1%	41.5%	<b>44.6%</b>	8.2%	1.60	1.14	0.2
	VH dijet	0.4	27.1%	2.8%	<b>43.7%</b>	24.3%	2.1%	1.54	1.24	0.5
ttH tags	0.2	3.1%	1.1%	2.2%	1.3%	<b>92.3%</b>	1.40	1.13	0.2	
8 TeV 19.7 fb <sup>-1</sup>	Untagged 0	6.0	<b>75.7%</b>	11.9%	6.9%	3.6%	1.9%	1.05	0.79	4.7
	Untagged 1	50.8	<b>85.2%</b>	7.9%	4.0%	2.4%	0.6%	1.19	1.00	120.
	Untagged 2	117.	<b>91.1%</b>	4.7%	2.5%	1.4%	0.3%	1.46	1.15	418.
	Untagged 3	153.	<b>91.6%</b>	4.4%	2.4%	1.4%	0.3%	2.04	1.56	870.
	Untagged 4	121.	<b>93.1%</b>	3.6%	2.0%	1.1%	0.2%	2.62	2.14	1400.
	VBF dijet 0	4.5	17.8%	<b>81.8%</b>	0.2%	0.1%	0.1%	1.30	0.94	0.8
	VBF dijet 1	5.6	28.5%	<b>70.5%</b>	0.6%	0.2%	0.2%	1.43	1.07	2.7
	VBF dijet 2	13.7	43.8%	<b>53.2%</b>	1.4%	0.8%	0.8%	1.59	1.24	22.1
	VH tight $\ell$	1.4	0.2%	0.2%	<b>76.9%</b>	19.0%	3.7%	1.63	1.24	0.4
	VH loose $\ell$	0.9	2.6%	1.1%	<b>77.9%</b>	16.8%	1.5%	1.60	1.16	1.2
	VH $E_T^{\text{miss}}$	1.8	16.3%	2.7%	34.4%	<b>35.4%</b>	11.1%	1.68	1.17	1.3
	VH dijet	1.6	30.3%	3.1%	<b>40.6%</b>	23.4%	2.6%	1.31	1.06	1.0
ttH lepton	0.5	—	—	1.6%	1.6%	<b>96.8%</b>	1.34	1.03	0.2	
ttH multijet	0.6	4.1%	0.9%	0.8%	0.9%	<b>93.3%</b>	1.34	1.03	0.6	

computed separately for 7 and 8 TeV dataset and then combined. The values in the plot are not corrected by the “look-elsewhere effect” [83]. The expected significance has been computed using the background expectation obtained from the fit in the signal-plus-background hypothesis, the so-called *post – fit* expectation.



**Figure 4.19.** Local  $p$ -values as a function of  $m_H$  for the 7 TeV, 8 TeV, and the combined dataset. The values of the expected significance, calculated using the background expectation obtained from the signal-plus-background fit, are shown as dashed lines.

The significance of the minimum of the local  $p$ -value corresponds to  $5.7\sigma$  where  $5.2\sigma$  is expected for a SM Higgs boson.

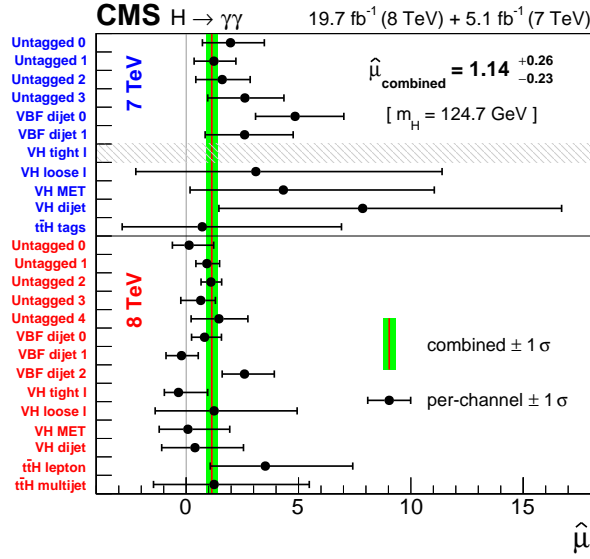
#### 4.6.1.1 Couplings measurement

The results presented in the previous paragraph for the search of the Higgs boson in the diphoton decay channel clearly state the presence of an excess around 125 GeV in  $m_{\gamma\gamma}$  distribution, compatible with the SM Higgs boson. A measurement of the couplings of this particle can be given in the parametrization reported in paragraph 1.6.1.

In Figure 4.20 the best-fit signal strength,  $\mu$ , is reported for each category of the analysis with its uncertainty. The result for the VH  $\ell$  tight tag at 7 TeV is not reported because the the signal-plus-background fit for that particular category does not converge when done alone since no events are observed in data in the region where the signal is observed. The measured value for  $\mu$ , combining all classes together is  $1.14^{+0.26}_{-0.23}$  and it is shown in the figure as a vertical line with a band representing its uncertainty. This measured value is compatible with SM expectations and a good agreement is observed between classes tagging different production modes.

In figure 4.21 production modes associated with fermion ( $ggH$  and  $t\bar{t}H$ ) or vector couplings (VBF and VH) are split. The  $1\sigma$  and  $2\sigma$  contours, computed as variations around the likelihood maximum for  $\mu_{ggH,t\bar{t}H}$  and  $\mu_{VBF,VH}$  are shown. In the fits, both signal strength modifiers are allowed to vary and  $m_H$  is treated as an unconstrained parameter. The best fit value and the expected uncertainty in case of a SM Higgs





**Figure 4.20.** Values of  $\mu$  measured individually for all event classes in the 7 and 8 TeV datasets, fixing  $m_H = 124.7$  GeV. The horizontal bars indicate  $\pm 1\sigma$  uncertainties in the values, and the vertical line and the green band indicate the best-fit signal strength in the fit to the data and its uncertainty.

are reported in Table 4.4.

**Table 4.4.** Expected and observed best-fit values of the signal strength modifiers  $\mu_{ggH, t\bar{t}H}$  and  $\mu_{VBF, VH}$  for a SM Higgs boson signal together with their uncertainties, indicating the expected uncertainty in the measurement and the best-fit values obtained from the data.

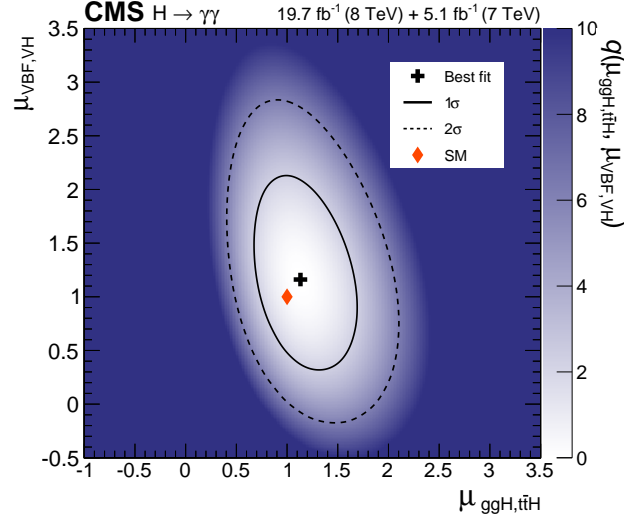
	Expected	Observed
$\hat{\mu}_{ggH, t\bar{t}H}$	$1.00^{+0.34}_{-0.30}$	$1.13^{+0.37}_{-0.31}$
$\hat{\mu}_{VBF, VH}$	$1.00^{+0.57}_{-0.51}$	$1.16^{+0.63}_{-0.58}$

The signal strength  $\mu$  can be varied independently for the four production mechanisms. As shown in figure 4.22, the different production mechanisms give compatible results with SM expectations, within the statistical uncertainty. Numerical values for each production mode are reported in Table 4.5.

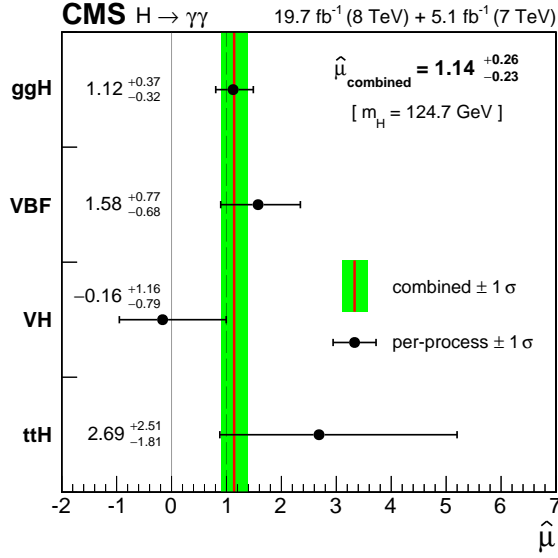
As explained in section 1.6.1 it is interesting to consider one single scale factor for fermions ( $\kappa_f$ ) and another one for bosons ( $\kappa_V$ ). The measurement of effective coupling modifiers to gluons ( $\kappa_g$ ) and to photons ( $\kappa_\gamma$ ) are also of paramount importance to constraint new physics. Figure 4.23 shows two-dimensional likelihood scans in these two different planes, fixing  $m_H = 124.7$  GeV. The best-fit points are  $(\kappa_f, \kappa_V) = (1.05, 1.06)$  and  $(\kappa_\gamma, \kappa_g) = (1.14, 0.90)$ .

#### 4.6.2 Combination with Other Channels

The most precise determination of the Higgs boson couplings has been obtained in CMS combining the  $H \rightarrow \gamma\gamma$  result, presented in this thesis, with the other analyses



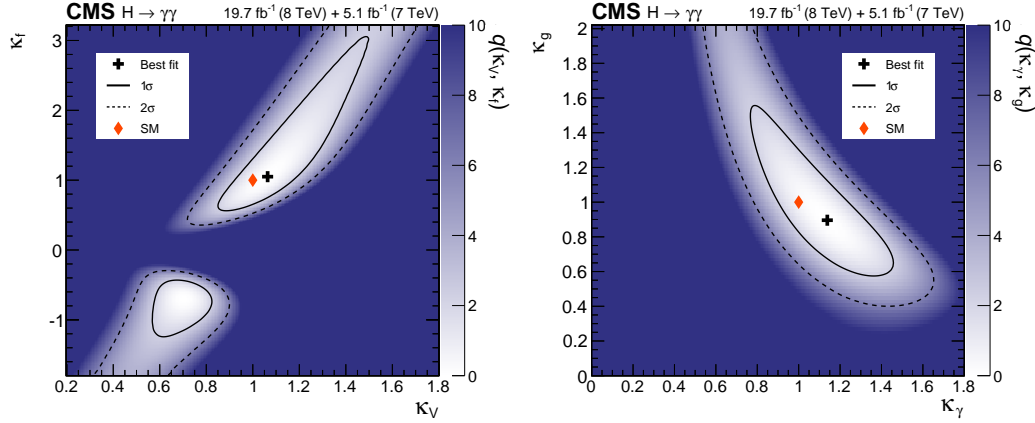
**Figure 4.21.** Map of the likelihood ratio  $q(\mu_{\text{ggH}, \text{ttH}}, \mu_{\text{VBF}, \text{VH}})$  with  $m_H$  treated as an unconstrained parameter. The  $1\sigma$  and  $2\sigma$  uncertainty contours are shown. The cross indicates the best-fit values,  $(\hat{\mu}_{\text{ggH}, \text{ttH}}, \hat{\mu}_{\text{VBF}, \text{VH}}) = (1.13, 1.16)$ , and the diamond represents the SM expectation.



**Figure 4.22.** Best-fit signal strength,  $\hat{\mu}$ , measured for each of the production processes in a combined fit where the signal strengths of all four processes are allowed to vary independently in the fit. The signal mass, common to all four processes, is treated as an unconstrained parameter in the fit. The horizontal bars indicate  $\pm 1\sigma$  uncertainties in the values for the individual processes. The band corresponds to  $\pm 1\sigma$  uncertainties in the value obtained from the combined fit with a single signal strength.

**Table 4.5.** Best-fit signal strength modifiers for the four production processes. The total uncertainty for each process is separated into statistical (stat) and systematic contributions. The systematic uncertainty has been separated, where feasible, into the contributions from theoretical (theo), and experimental (exp) uncertainties. Statistical uncertainty is subtracted quadratically from the positive and negative uncertainties separately. The quoted values are the mean values of the positive and negative uncertainties.

Process	$\hat{\mu}$	Uncertainty			
		total	stat	systematic theo	exp
ggH	$1.12^{+0.37}_{-0.32}$	0.34	0.30	0.13	0.09
VBF	$1.58^{+0.77}_{-0.68}$	0.73	0.69	0.20	0.15
VH	$-0.16^{+1.16}_{-0.79}$	0.97	0.97	0.08	
t $\bar{t}$ H	$2.69^{+2.51}_{-1.81}$	2.2	2.1	0.4	



**Figure 4.23.** Maps of the likelihood ratio  $q(\kappa_V, \kappa_f)$  (left), and  $q(\kappa_\gamma, \kappa_g)$  (right), showing the  $1\sigma$  and  $2\sigma$  uncertainty contours. The crosses indicate the best-fit values, and the diamonds indicate the SM expectation.

targeting different decay channels. Around 125 GeV almost all the possible decays have a sizeable branching ratio and can contribute to the final combination. CMS has exploited decays of the Higgs particles to:  $\gamma\gamma$ ,  $ZZ$ ,  $WW$ ,  $\tau\tau$  and  $bb$  and the full 7 and 8 TeV dataset. The full details of the combination techniques can be found in [84]. In the first part of this chapter results on couplings obtained combining all these channels will be presented. All available analyses targeting the four possible production mechanisms (ggH, VBF, VH and  $t\bar{t}H$ ) are used in the combination. Given the importance of the top-Higgs coupling, searches targeting  $t\bar{t}H$  production mechanism, even if already used in the results presented here, will be described in more details in the next paragraph.

The channels  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ \rightarrow 4\ell$  play a major role given the possibility of reconstructing completely the final state. In both channels the signal over background ratio is good and the mass of the particle can be measured with high precision. The  $H \rightarrow WW$  channel, with leptonic decays of the W pair, has high sensitivity but poor mass resolution because of the presence of two neutrinos. The  $\tau\tau$  and  $bb$  channels are affected by large background contributions and poor mass resolution. Therefore they have a lower sensitivity compared to other channels.

$H \rightarrow ZZ \rightarrow 4\ell$  analysis [85] is performed reconstructing the invariant mass of the four-lepton system. Indeed, since muons and electrons can be reconstructed with high precision, a narrow peak at the Higgs mass value over a continuum background is expected. Several categories are devised in this channel according to the possible combinations of Z boson decays:  $4e$ ,  $4\mu$ ,  $2e2\mu$ . Given the different mass resolution and background rates, these categories are analyzed separately.  $ZZ$  production, with Z bosons decaying leptonically is a dominant irreducible background and is estimated from simulation. Reducible backgrounds with misidentified leptons, from Z+jets,  $t\bar{t}$  and WZ+jets are estimated from data. Events are further categorized according to the number of jets present in the final state. Angular correlations and kinematic variables of the decay products are used to build discriminators to further suppress background contributions.

$H \rightarrow WW \rightarrow \ell\nu\ell\nu$  analysis [86] measures an excess of events with two oppositely charged leptons or three leptons with a total charge  $\pm 1$ , moderate  $E_{\text{T}}^{\text{miss}}$  and up to two jets. Main background contributions in this channel are represented by WW non-resonant production, top-quark production and Drell-Yan production ( $q\bar{q} \rightarrow \gamma^*/Z^* \rightarrow \ell\ell$ ).

$H \rightarrow \tau\tau$  analysis [87] searches for an excess of events using multiple final states, targeting leptonic and hadronic decays of  $\tau$  leptons. Categorization of events is done according to the number of reconstructed jets in the final state and, in leptonic categories, to the  $p_{\text{T}}$  of leptons. The main irreducible background comes from  $Z \rightarrow \tau\tau$  prompt production, while reducible contributions come from W+jets and multijet production.

$H \rightarrow bb$  [88] analyses are affected by a huge background of prompt jet QCD production. Searches in this channel are therefore performed exploiting only the VH and  $t\bar{t}H$  production mechanism. The Higgs boson candidate is reconstructed by requiring two b-tagged jets and a multivariate regression is used to achieve a final dijet mass resolution of about 10% for  $m_H = 125$  GeV.

The combination of the analyses is done, as in  $H \rightarrow \gamma\gamma$  standalone analysis, using a test-statistic based on the profile likelihood ratio to determine how signal-like

**Table 4.6.** The best-fit values for the signal strength at  $m_H = 125.0$  GeV, of the VBF and  $VH$ , and of the  $ggH$  and  $t\bar{t}H$  production mechanisms,  $\mu_{\text{VBF},VH}$  and  $\mu_{\text{ggH},t\bar{t}H}$ , respectively. The observed and median expected results for the ratio of  $\mu_{\text{VBF},VH}$  to  $\mu_{\text{ggH},t\bar{t}H}$  together with their uncertainties are also given for the full combination.

Channel grouping	Best fit ( $\mu_{\text{ggH},t\bar{t}H}, \mu_{\text{VBF},VH}$ )
$H \rightarrow ZZ$ tagged	(0.88, 1.75)
$H \rightarrow \gamma\gamma$ tagged	(1.07, 1.24)
$H \rightarrow WW$ tagged	(0.87, 0.66)
$H \rightarrow \tau\tau$ tagged	(0.52, 1.21)
$H \rightarrow b\bar{b}$ tagged	(0.57, 0.96)
Combined best fit $\mu_{\text{VBF},VH}/\mu_{\text{ggH},t\bar{t}H}$	
	Observed (expected)
	$1.25^{+0.63}_{-0.45}$ ( $1.00^{+0.49}_{-0.35}$ )

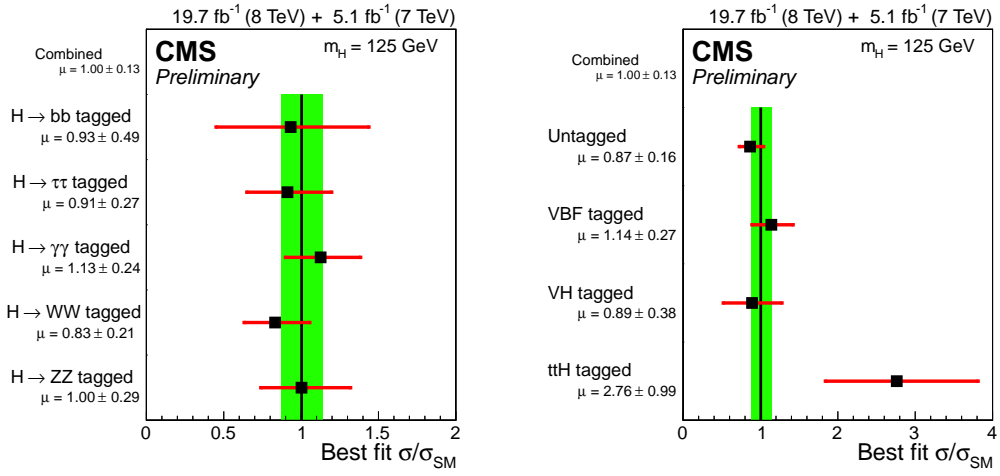
or background-like the data are. Systematic uncertainties are incorporated in the analysis as nuisance parameters and the correlations within different analyses are taken into account.

All channels observe an excess over background-only expectations compatible with the presence of a SM Higgs boson. Figure 4.24 shows the measured signal strength grouping analyses depending on decay channels or production modes. The largest deviation is observed in the  $t\bar{t}H$  production mode and will be further characterized in the next paragraph. Table 4.6 shows the measured signal strength for different channels grouping production modes associated with fermions ( $ggH$  and  $t\bar{t}H$ ) or vector couplings (VBF and  $VH$ ).

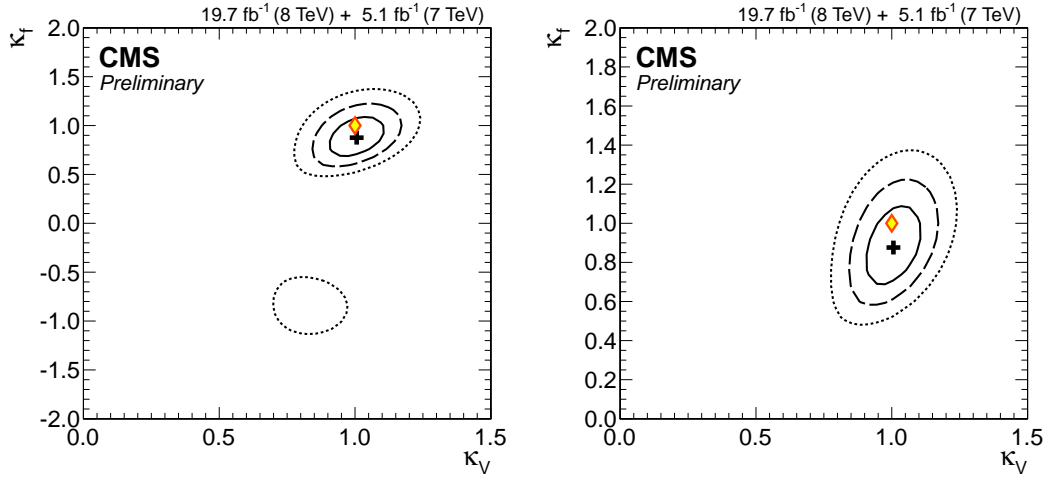
In analogy to what already done in  $H \rightarrow \gamma\gamma$  analysis, compatibility of couplings of the Higgs boson to vector bosons and fermions can be tested. Maps of the likelihood ratio in the plane of couplings to vector bosons and fermions can be made, as shown in figure 4.25. The 95% CL intervals for  $\kappa_V$  and  $\kappa_f$  are each obtained from a 1D scan where the other parameter is fixed to unity, and are equal to [0.88,1.15] and [0.64,1.16], respectively.

In the same way, the likelihood scan can be shown in the 2D plane of effective couplings to photons and gluons ( $\kappa_\gamma, \kappa_g$ ). Partial widths associated with tree-level production are assumed to be unaltered. The result of the 2D likelihood scan is shown in figure 4.26, the best fit point corresponds to (1.14,0.88) and the 95% CL intervals for these quantities are [0.89, 1.42] for  $\kappa_\gamma$  and [0.69,1.10] for  $\kappa_g$ .

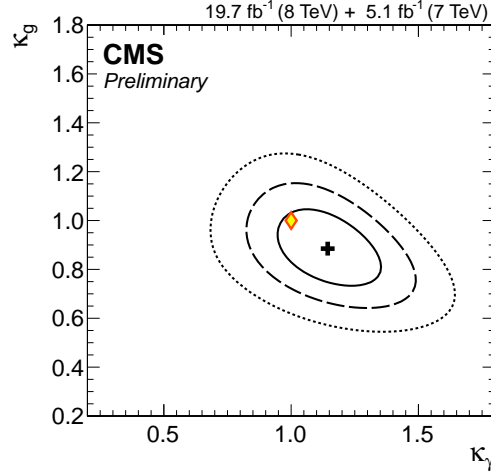
Since  $H \rightarrow \gamma\gamma$  decay and  $ggH$  production proceed through loop diagrams, the values of  $\kappa_\gamma$  and  $\kappa_g$  are expected to be significantly altered by the presence of BSM particles. Indeed any unknown particle, carrying electromagnetic charge, can contribute to the loop inducing a big change in these two couplings with respect to SM predictions. The value of the branching ratio of new physics,  $\text{BR}_{\text{BSM}} = \Gamma_{\text{BSM}}/\Gamma_{\text{tot}}$  can be constrained, profiling  $\kappa_\gamma$  and  $\kappa_g$  with other nuisance parameters. Figure 4.27 shows the likelihood scan versus  $\text{BR}_{\text{BSM}}$ . The likelihood profile allows to constrain  $\text{BR}_{\text{BSM}}$  in the interval [0.00,0.32] at 95% CL.



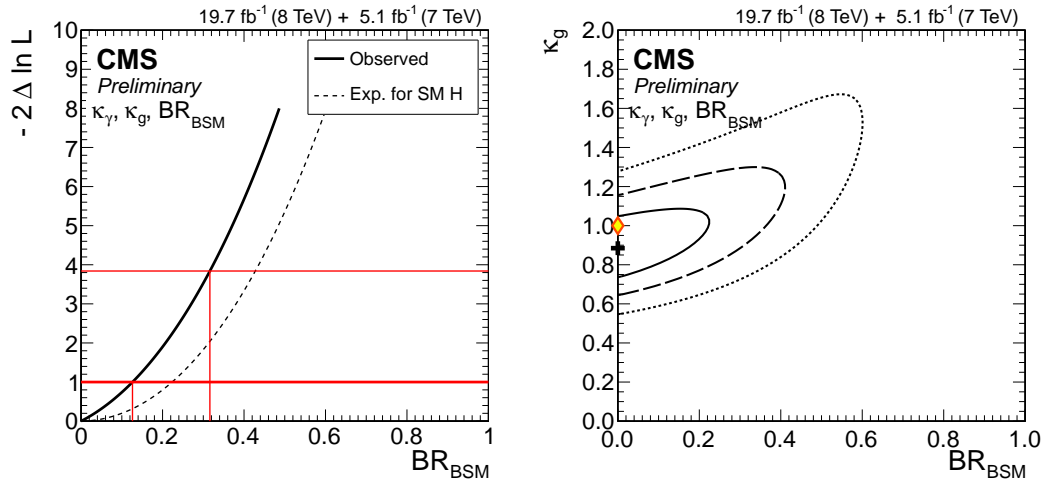
**Figure 4.24.** Values of the best-fit  $\mu = \sigma/\sigma_{\text{SM}}$  for the combination (solid vertical line), for individual channels, and for subcombinations by predominant decay mode or production mode tag. The vertical band shows the overall  $\mu$  uncertainty. The horizontal bars indicate the  $\pm 1$  standard deviation uncertainties in the best-fit  $\sigma/\sigma_{\text{SM}}$  values for the individual modes; they include both statistical and systematic uncertainties. On the left subcombinations by predominant decay mode are shown, while on the right subcombinations by analysis tags targeting individual production mechanisms are reported.



**Figure 4.25.** Results of 2D likelihood scans for the  $\kappa_V$  and  $\kappa_f$  parameters. The cross indicates the best-fit values. The solid, dashed, and dotted contours show the 68%, 95%, and 99.7% CL regions, respectively. The yellow diamond shows the SM point  $(\kappa_V, \kappa_f) = (1, 1)$ . The left plot shows the likelihood scan in two quadrants,  $(+, +)$  and  $(+, -)$ . The right plot shows the likelihood scan constrained to the  $(+, +)$  quadrant.



**Figure 4.26.** The 2D likelihood scan for the  $\kappa_g$  and  $\kappa_\gamma$  parameters, assuming that  $\Gamma_{\text{BSM}} = 0$ . The cross indicates the best-fit values. The solid, dashed, and dotted contours show the 68%, 95%, and 99.7% CL regions, respectively. The yellow diamond represents the SM expectation,  $(\kappa_\gamma, \kappa_g) = (1, 1)$ . The partial widths associated with the tree-level production processes and decay modes are assumed to be unaltered ( $\kappa = 1$ ).



**Figure 4.27.** (Left) The likelihood scan versus  $\text{BR}_{\text{BSM}} = \Gamma_{\text{BSM}}/\Gamma_{\text{tot}}$ . The solid curve represents the observation and the dashed curve indicates the expected median results in the presence of the SM Higgs boson. The partial widths associated with the tree-level production processes and decay modes are assumed to be unaltered ( $\kappa = 1$ ). (Right) Correlation between  $\kappa_g$  and  $\text{BR}_{\text{BSM}}$ . The solid, dashed, and dotted contours show the 68%, 95%, and 99.7% CL regions, respectively.

#### 4.6.2.1 $t\bar{t}H$ Combination with Other Channels

Given the theoretical interest to directly probe top-Higgs coupling through  $t\bar{t}H$  direct production, a big effort has been done in the CMS collaboration to combine all possible decay channels to achieve the best sensitivity in  $\mu_{t\bar{t}H}$  measurement. Given the low cross section of this process, exploiting all different decay modes and dedicated analysis is needed to have a sensitive result with the current amount of data.

CMS  $t\bar{t}H$  analyses can be divided in three families according to the Higgs decay mode:  $H \rightarrow \text{hadrons}$  [89],  $H \rightarrow \text{leptons}$  [90] and  $H \rightarrow \text{photons}$  (already presented in previous sections). These searches are different given the wide spectra of possible final states they investigate. Anyway, a strategy common to all of them is to categorize events according to the number of jets and b-jets in the final state, since, when looking for  $t\bar{t}H$  production mechanism, the signal-over-background ratio increases requiring jet activity.

$H \rightarrow \text{hadrons}$  analyses target decays of the Higgs boson into a pair of b-quarks or hadronically decaying  $\tau$  leptons. They are characterized by a high rate, given that the  $H \rightarrow b\bar{b}$  channel contributes almost 60% of the total Higgs decay width. The main limitation of this channel is the presence of the big irreducible background made of  $t\bar{t}+b\bar{b}$  prompt production. For this reason searches in these channels focus on the lepton + jets and dilepton decay of the  $t\bar{t}$  system, using single or double lepton triggers respectively. The background modeling is done on MC simulations while control samples are used to constrain the normalization of some processes. Multivariate techniques are used to separate signal process from irreducible and reducible backgrounds with input variables related to object kinematics and b-tag. A multivariate discriminator is also used to tag jets coming from a b quark or a  $\tau$ . The main systematic uncertainties of these analyses are on the modeling of  $t\bar{t}+b\bar{b}$  events since large theoretical uncertainties arise in the computation of top pair production with heavy flavor emission.

$H \rightarrow \text{leptons}$  analysis requires the presence of at least two leptons in the final state. The Higgs boson decay modes investigated here are those in a pair of vector bosons (W or Z) and in a pair of leptonically decaying  $\tau$ . The expected rate is lower compared to hadronic final states, but the presence of leptons allows to suppress the background from top events. Lepton charge correlations and kinematic variables are used to improve sensitivity. A dedicated lepton identification, using multivariate techniques, has been specifically developed to suppress events with jets misreconstructed as leptons, i.e.  $t\bar{t}+\text{jets}$ , which constitute the main background. Background modeling is done using simulations; specific correction factors for lepton- and jet- tagging are applied to rescale simulations to match data distributions. The data/MC agreement is cross-checked in control regions and systematic uncertainties related to these corrections are considered. Three categories are used according to the number of leptons in the final state: 2 (same sign), 3 or 4 leptons.

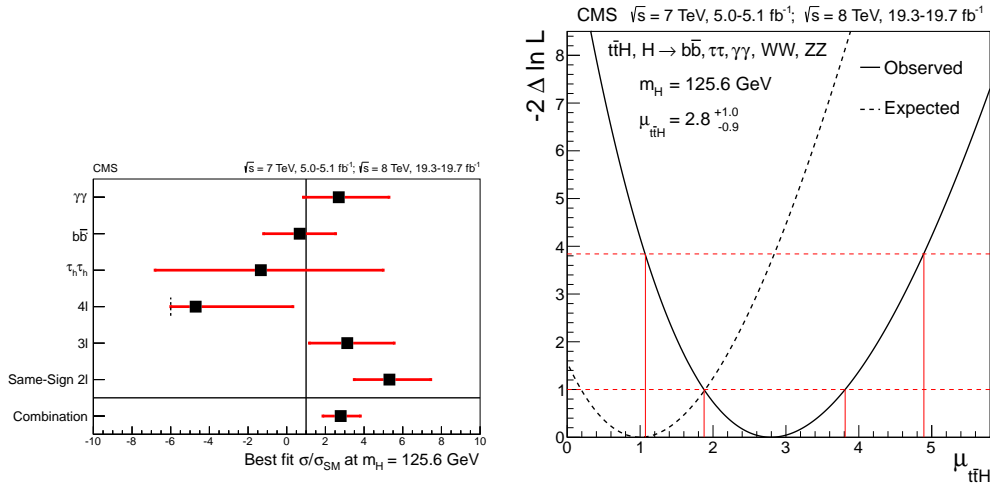
These analyses are combined together and results are extracted using the same statistical approach described in section 4.4. In brief, a binned likelihood using all channels is built, where  $\mu_{t\bar{t}H}$  is extracted assuming SM branching ratios [91]. The best-fit value for  $\mu_{t\bar{t}H}$  in individual channels and their combination are given in table 4.7 and figure 4.28. Combining all channels, the best-fit value is  $\mu = 2.8^{+1.0}_{-0.9}$ .



It represents roughly a  $3\sigma$  excess with respect to background only expectations and it is compatible with the SM expectations at  $2\sigma$  level. All the results are given for a Higgs mass of 125.6 GeV which corresponds to the best-fit value of the CMS collaboration at the time the result was obtained. This excess is mainly driven by the multilepton and diphoton channel. The biggest excess on background only expectations can be seen in the dimuon same-sign category.

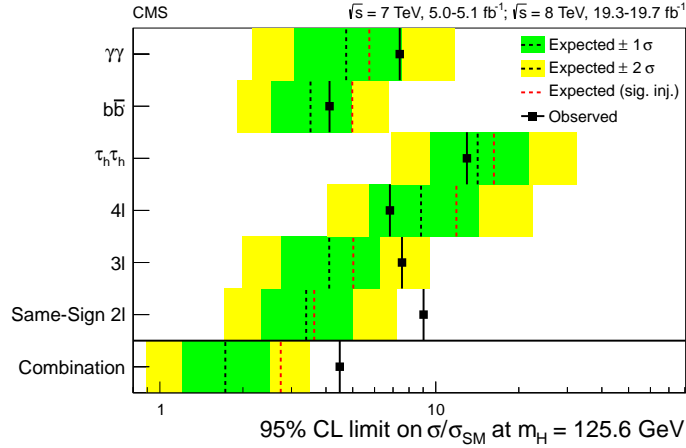
**Table 4.7.** The best-fit values of the signal strength parameter  $\mu = \sigma/\sigma_{\text{SM}}$  for each  $t\bar{t}H$  channel at  $m_H = 125.6$  GeV. The observed and expected 95% CL upper limits on the signal strength parameter  $\mu = \sigma/\sigma_{\text{SM}}$  for each  $t\bar{t}H$  channel at  $m_H = 125.6$  GeV are also shown.

t $\bar{t}H$ channel	Best-fit $\mu$	95% CL upper limits on $\mu = \sigma/\sigma_{\text{SM}}$ ( $m_H = 125.6$ GeV)				
		Observed	Observed	Median signal-injected	Median	68% CL range
$\gamma\gamma$	$+2.7^{+2.6}_{-1.8}$	7.4	5.7	4.7	[3.1, 7.6]	[2.2, 11.7]
$b\bar{b}$	$+0.7^{+1.9}_{-1.9}$	4.1	5.0	3.5	[2.5, 5.0]	[1.9, 6.7]
$\tau_h\tau_h$	$-1.3^{+6.3}_{-5.5}$	13.0	16.2	14.2	[9.5, 21.7]	[6.9, 32.5]
4l	$-4.7^{+5.0}_{-1.3}$	6.8	11.9	8.8	[5.7, 14.3]	[4.0, 22.5]
3l	$+3.1^{+2.4}_{-2.0}$	7.5	5.0	4.1	[2.8, 6.3]	[2.0, 9.5]
Same-sign 2l	$+5.3^{+2.1}_{-1.8}$	9.0	3.6	3.4	[2.3, 5.0]	[1.7, 7.2]
Combined	$+2.8^{+1.0}_{-0.9}$	4.5	2.7	1.7	[1.2, 2.5]	[0.9, 3.5]



**Figure 4.28.** Left: The best-fit values of the signal strength parameter  $\mu = \sigma/\sigma_{\text{SM}}$  for each  $t\bar{t}H$  channel at  $m_H = 125.6$  GeV. The signal strength in the four-lepton final state is not allowed to be below approximately  $-6$  by the requirement that the expected signal-plus-background event yield must not be negative in either of the two jet multiplicity bins. Right: The 1D test statistic  $q(\mu_{t\bar{t}H})$  scan vs. the signal strength parameter for  $t\bar{t}H$  processes  $\mu_{t\bar{t}H}$ , profiling all other nuisance parameters. The lower and upper horizontal lines correspond to the 68% and 95% CL, respectively. The  $\mu_{t\bar{t}H}$  values where these lines intersect with the  $q(\mu_{t\bar{t}H})$  curve are shown by the vertical lines.

It is also useful to quote the results in terms of a 95% CL upper limit on the expected  $t\bar{t}H$  cross section, to set constraints on BSM models predicting  $t\bar{t}H + X$  final states. The upper limit for individual channels and the combined value are reported in table 4.7 and in figure 4.29.



**Figure 4.29.** The 95% CL upper limits on the  $t\bar{t}H$  signal strength parameter  $\mu = \sigma_{t\bar{t}H} / \sigma_{t\bar{t}H}^{\text{SM}}$ . The black solid and dotted lines show the observed and background-only expected limits, respectively. The red dotted line shows the median expected limit for the SM Higgs boson with  $m_H = 125.6$  GeV. The green and yellow areas show the  $1\sigma$  and  $2\sigma$  bands, respectively.

## 4.7 Summary and conclusions

In this chapter searches for a SM Higgs in the diphoton decay channel have been presented using the full CMS dataset. The main improvement in couplings measurement, with respect to past analyses, comes from the inclusion of dedicated categories tagging associate production modes. Thanks to this strategy, the precision on this measurement has significantly increased.

The data clearly show the presence of an excess over background-only predictions in  $m_{\gamma\gamma}$  spectrum. This excess is compatible with a SM Higgs boson and a full characterization of its couplings to fermions and bosons has been given in different benchmark models. No striking deviations from theoretical predictions have been observed; the measured signal strengths for different production modes are compatible with SM expectations. A mild excess, roughly two sigmas over SM expectations in the combination of all decay channels, is seen in  $t\bar{t}H$  production mode.

The results presented are given in such a general way that can be directly translated into constraints on theoretical models involving the Higgs sector. A sizeable part of the phase-space of BSM theories can be excluded and projections for future LHC runs can be produced using the current sensitivity. This will be important to elaborate new theories and highlight the most promising channel for BSM searches with Higgs bosons in the final state.

## Chapter 5

# Direct searches beyond Standard Model in the Higgs sector

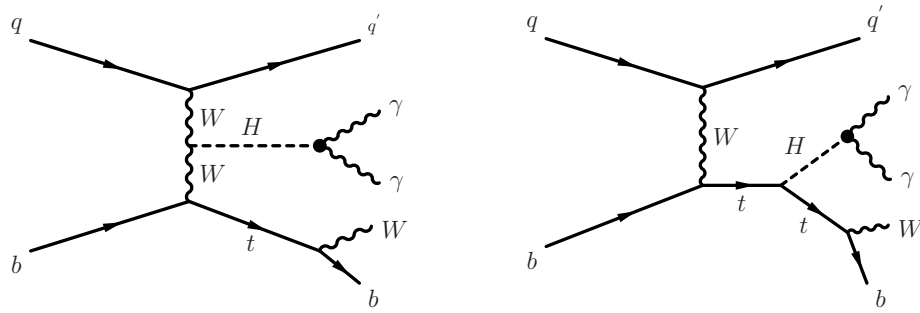
In the previous chapter, measurements of Higgs boson couplings to other particles in the diphoton decay channel have been presented. The results obtained can be used to constrain several BSM physics models, like those presented in chapter 1.

In this chapter a different and complementary approach is followed: instead of setting indirect constraints, direct searches for BSM particles involving the Higgs boson in the final state are presented. The chapter is divided in two parts: the first part presents a search for anomalous values of the top-Higgs coupling in the single top plus Higgs process, the second one describes instead a search for heavy vector-like partners of the top quark decaying in top and Higgs. Both searches are motivated by different BSM scenarios that predict the existence of a new strong sector, like composite or little Higgs models. Both analyses are performed in the diphoton decay channel and, unless where it is explicitly said, use the same object reconstruction and identification criteria of the SM Higgs analysis. In particular, both analyses exploit the strategy of fitting the diphoton invariant mass to extract the number of signal and background events. The only relevant difference is that a cut-based photon identification, described in paragraph 3.1.4, is used, instead of the multivariate one. This is preferable to the multivariate approach for BSM searches, since more model-independent.

The full dataset collected by CMS at a center-of-mass energy of 8 TeV is used in both analyses, and events are required to pass the same diphoton trigger requirements as the SM  $H \rightarrow \gamma\gamma$  analysis.

### 5.1 Associate Production of a Single Top Quark and a Higgs boson

The most straightforward way to study the magnitude of the coupling between the top quark and the Higgs boson ( $C_t$ ) is the associate pair production ( $t\bar{t}H$ ), presented in the previous chapter. Single top associate production ( $tHq$ ) is an interesting



**Figure 5.1.** Dominant Feynman diagrams for the production of tHq events: the Higgs boson is typically irradiated from the heavier legs of the diagram, i.e. the W boson (left) or the top quark (right).

way to prove, not only the magnitude, but also the sign of this coupling. This process mainly proceeds through the t-channel where the Higgs is emitted from the W boson propagator or directly from the top leg. Figure 5.1 shows the dominant Feynman diagrams for this Higgs production mode. The cross section of the process is proportional to the magnitude and the relative sign of the two couplings to top quark and W boson:

$$\sigma_{\text{tHq}} \propto |C_W \mathcal{M}_W + C_t \mathcal{M}_t|^2$$

where  $\mathcal{M}_W$  and  $\mathcal{M}_t$  are the matrix element corresponding to the two possible diagrams for tHq production.

The Standard Model predicts opposite sign for the coupling of the Higgs to W and the one to the top. This leads to a negative interference of the two diagrams and the predicted production cross section for this process, at Next-to-Leading Order (NLO), is only 18 fb at 8 TeV. As can be seen from the left plot of figure 4.23, the possibility of a Higgs boson having negative coupling to fermions ( $C_t = -1$ ) is not yet excluded with the current sensitivity of the SM  $H \rightarrow \gamma\gamma$  analysis. Given the interference of the two diagrams in single top production, a big enhancement in the production cross section is expected if  $C_t$  has an opposite sign with respect to SM predictions.  $C_t = -1$  would mean an increase of the cross section of a factor 15.

The single top quark plus Higgs boson production cross section could be enhanced by other modifications of the SM. For example, new physics appearing at high mass scales could appear as flavor-changing-neutral-currents (FCNC) involving the top quark and the Higgs boson. Vertex such as tHu, tHc would be possible in this scheme, giving rise to final states overlapping with the final state under study here. While there is a large body of work in the literature on the phenomenology of such FCNC at colliders [92], the ATLAS and CMS collaborations produced only some preliminary investigations [93, 94].

Composite Higgs scenarios would predict the existence of a heavy top quark partner with large branching fraction into a top quark and a Higgs boson. The single top partner production cross section exceeds the pair production cross section under a large variety of assumptions of electroweak top quark partners couplings, for top quark partner masses above the 700-800 GeV [95]. This exotic production mode would give rise to final states that would overlap with the ones needed for

studying single top+Higgs production in the Standard Model.

It must be noted that the condition  $C_t = -1$  not only enhances the cross section of tHq events, but the branching ratio of the Higgs to photons would increase by a factor 2.4, because of the modified interference between the decay loops mediated through W bosons and top quarks in the  $H \rightarrow \gamma\gamma$  decay. For this reason, also the yields of standard H processes increase. These additional contributions will be considered as signal for the purposes of the statistical interpretation of the results.

The interpretation of the results of this search will be given in the  $C_t = -1$  scenario, even if the full spectra of possible values could be interesting.

### 5.1.1 Analysis strategy

This analysis searches for a Higgs boson produced in association with a top quark. The Higgs boson decay in two photons and the leptonic decay of the top quark are required.

The process studied is therefore:

$$\text{tHq} \rightarrow (\text{t} \rightarrow \text{b}\ell\nu)(\text{H} \rightarrow \gamma\gamma)q \quad \text{with } \ell = \text{e}, \mu$$

The final state is composed by the diphoton system, a lepton (electron or muon) and a jet which is expected to have high rapidity, since the Higgs is radiated off one of the heavier legs of this t-channel process. As in SM Higgs analysis, the invariant mass of the diphoton system,  $m_{\gamma\gamma}$ , is used to search for a signal-like excess and to estimate non-resonant background contribution.

The backgrounds in this search are of two types: non-resonant and resonant in  $m_{\gamma\gamma}$ . The non-resonant one includes processes with photons and jets in the final state like continuous diphoton production ( $\gamma\gamma + \text{jets}$ ), single photon production with jets misreconstructed as photons ( $\gamma + \text{jets}$ ) and diphoton events with top quarks ( $\text{t}\gamma\gamma$  and  $\text{t}\bar{\text{t}}\gamma\gamma$ ). Resonant backgrounds have a Higgs boson decaying in two photons, from SM processes with higher cross sections. These processes will appear as a background under the Higgs peak, and are dominated by  $\text{t}\bar{\text{t}}\text{H}$  and  $\text{VH}$  production.

A cut-and-count analysis is performed, in which the contribution of non-resonant backgrounds is evaluated by counting the events in the  $m_{\gamma\gamma}$  sidebands ( $100 < m_{\gamma\gamma} < (m_{\text{H}} - 3 \text{ GeV})$  and  $(m_{\text{H}} + 3 \text{ GeV}) < m_{\gamma\gamma} < 180 \text{ GeV}$ ) whereas the contribution of resonant backgrounds is estimated using simulations. The signal region is defined as the  $\pm 3 \text{ GeV}$  region around the nominal mass position (125 GeV).

The production of tHq events has been generated with MADGRAPH 5 [68], interfaced to PYTHIA 6.4.2 [63] for parton shower and hadronization. For VH and  $\text{t}\bar{\text{t}}\text{H}$  processes, the same samples of the SM analysis, described in the previous chapter, are used.

As non-resonant backgrounds are evaluated directly in the data  $m_{\gamma\gamma}$  sidebands, Monte Carlo simulations are not used to obtain the final result. However, the use of simulated non-resonant backgrounds will play a role in the studies performed on the control sample described in Section 5.1.3. For that reason, diphoton plus jets events simulated in SHERPA [69] are used, as well as single-photon and jet-production events simulated in PYTHIA, top quark-antiquark production simulated in MADGRAPH+PYTHIA, and, finally, diphoton events produced in association with top quarks in WHIZARD [96, 97].

Photon identification and reconstruction are done using the cut-based approach described in paragraph 3.1.4.

### 5.1.2 Event Selection

The final state of Higgs associate production with single top is composed by three objects: the Higgs (decaying in a pair of photons), the top quark and the additional quark jet (called ‘qJet’ in the following). The two photons coming from the Higgs have high transverse momentum, like in  $t\bar{t}H$  or  $VH$  production, and, since the leptonic decay channel of the top is targeted, at least one electron or muon from  $W$  decays can be required. Regarding jets, the quark jet is expected to be forward, given the production mechanism, and a  $b$ -quark should be present from top decays.

The event selection therefore requires:

- two photons, with  $p_T > 50 \cdot m_{\gamma\gamma}/120$  and 25 GeV respectively;
- exactly one lepton (electron or muon) with  $p_T > 10$  GeV;
- a  $b$ -jet with  $p_T > 20$  GeV;
- a ‘qJet’ candidate: it is identified as the hardest additional jet in the event and must have  $p_T > 20$  GeV and  $|\eta| > 1$ ;

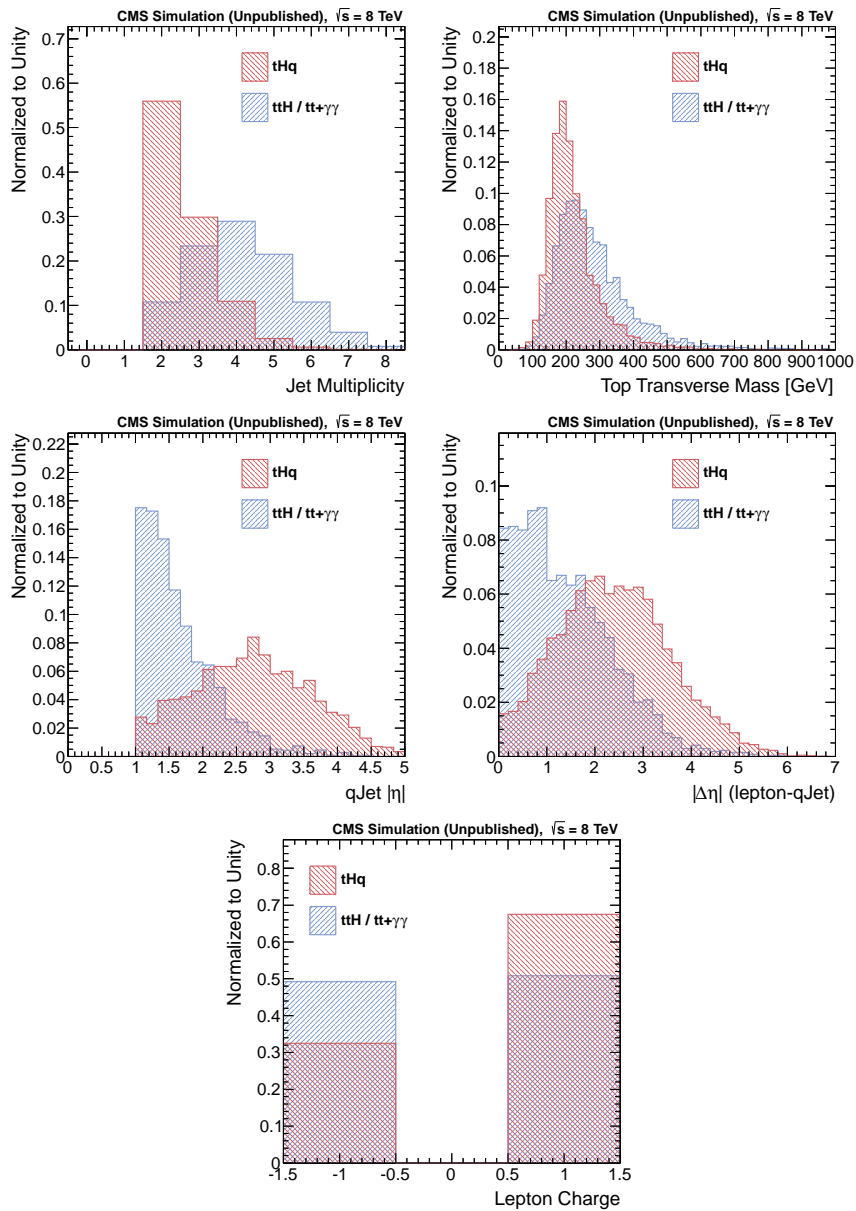
The requirement on the leading photon is almost fully efficient (98%) on signal and suppresses non-resonant background where photons are generally soft. No stringent requirements are applied to the lepton to be as efficient as possible on signal. If more than a jet is positively  $b$ -tagged, the hardest one is taken as the  $b$ -jet candidate. The definition of the ‘qJet’ candidate allows to correctly match the parton in 90 % of the cases.

After these requirements the resonant background contribution,  $t\bar{t}H$ , is still relevant. A likelihood product discriminator (LD) is therefore used to further suppress this background.

Five discriminating variables between  $tHq$  and  $t\bar{t}H$  are used as input variables:

- jet multiplicity,
- transverse mass of the top candidate,
- pseudorapidity of the qJet candidate,
- rapidity gap between the lepton and the qJet,
- charge of the lepton candidate.

The distributions of these variables for  $tHq$  and  $t\bar{t}H$  events are shown in figure 5.2. Events entering those plots are required to have at least two jets with  $p_T > 20$  GeV in the range  $|\eta| < 5$ , one of which  $b$ -tagged, and one lepton. The two photons are required to pass  $p_T$  cuts compatible with trigger thresholds: 33 GeV for the leading and 25 GeV for the trailing one. As can be seen, these variables have discriminating power given the different jet and lepton kinematics between  $tHq$  and  $t\bar{t}H$ . For instance, in  $t\bar{t}H$  process a higher number of jets is expected with respect to  $tHq$ ,

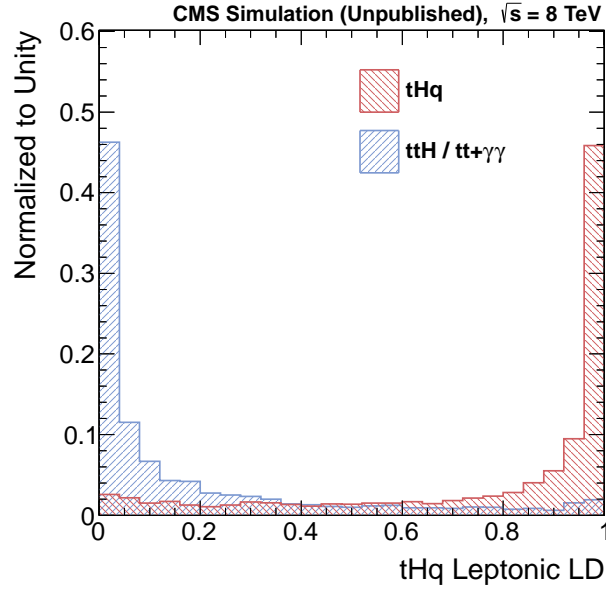


**Figure 5.2.** Expected distributions for input variables to the likelihood discriminant (LD): jet multiplicity (top left), top transverse mass (top right), pseudorapidity of the  $q_{\text{Jet}}$  candidate (center left), rapidity gap between the lepton and the  $q_{\text{Jet}}$  (center right), charge of the lepton candidate (bottom). Signal events are reported in red, while background contribution corresponds to the blue histogram. Both histograms are normalized to unity.

given the presence of two top quarks, and the  $q_{\text{Jet}}$  candidate is expected to be central since coming from  $W$  decays or additional radiation. The variables used are almost not correlated (their linear coefficient were found to be less than 10 % for both processes) so the use of a simple likelihood discriminant is expected to have close-to-optimal performances. The possibility of using boosted decision trees or

add more variables has been investigated without any significant gain in sensitivity.

The likelihood discriminant output is shown in figure 5.3. As can be seen, a good discrimination between signal, peaking at high values, and resonant  $t\bar{t}H$  background is achieved. To achieve a  $t\bar{t}H$  contamination less than 10 % under the Higgs peak, a cut is added to event selection requiring the likelihood discriminant to be greater than 0.25. The full analysis requirements are summarized in table 5.3.



**Figure 5.3.** Likelihood discriminant output for  $tHq$  (red) and  $t\bar{t}H$  (blue) simulated events. Both contributions are normalized to unity.

Table 5.1 shows the expected number of events passing the selection requirements, for signal and resonant background contributions. Additional contributions to  $t\bar{t}H$  and  $VH$  from negative top-Higgs coupling in SM processes are also reported. The analysis selection has an efficiency of 17% on  $tHq$  events in the leptonic decay channel.

Figure 5.4 shows the diphoton mass spectrum for events passing the full selection requirements. As can be seen from the plot, after the full selection, no events are found in data. As the number of background events and the related uncertainty needs to be estimated using sidebands, the measurement of zero events should be treated carefully and will be discussed in the next paragraph.

### 5.1.3 Background Estimation

In the diphoton decay channel, the main advantage is that a sideband analysis can be performed to estimate the contribution of non-resonant background. A sideband analysis relies on the definition of a signal-free control region (the sidebands) where to measure the background and extrapolate this measurement into an estimate of the background in the signal counting window.

In this analysis the variable used for the cut-and-count sideband analysis is  $m_{\gamma\gamma}$ . The knowledge of the background shape  $f_{bg}(m_{\gamma\gamma})$  is a key element to estimate the



**Table 5.1.** Expected MC yields in  $19.7\text{fb}^{-1}$  of data passing full requirements. Yields are counted in the 122-128 GeV range. The additional contributions from the ‘standard’ Higgs processes which would arise from the condition  $C_t = -1$  are marked with a dagger ( $\dagger$ ).

Process	Yield
$tHq$ ( $C_t = -1$ )	0.67
$t\bar{t}H$	$0.03 + 0.05^\dagger$
VH	$0.01 + 0.01^\dagger$
other $H$	0

non resonant background contribution under the Higgs peak. Indeed, after having counted the number of events in data sidebands ( $N_{side}$ ) the number of non-resonant background events in the signal region is given by:

$$N_{bg} = \alpha N_{side}$$

where  $\alpha$  is the ratio of integrals of  $f_{bg}$  in the signal and sidebands region:

$$\alpha = \frac{\int_{signal\ region} f_{bg}(m_{\gamma\gamma}) dm_{\gamma\gamma}}{\int_{sideband\ region} f_{bg}(m_{\gamma\gamma}) dm_{\gamma\gamma}} \quad (5.1)$$

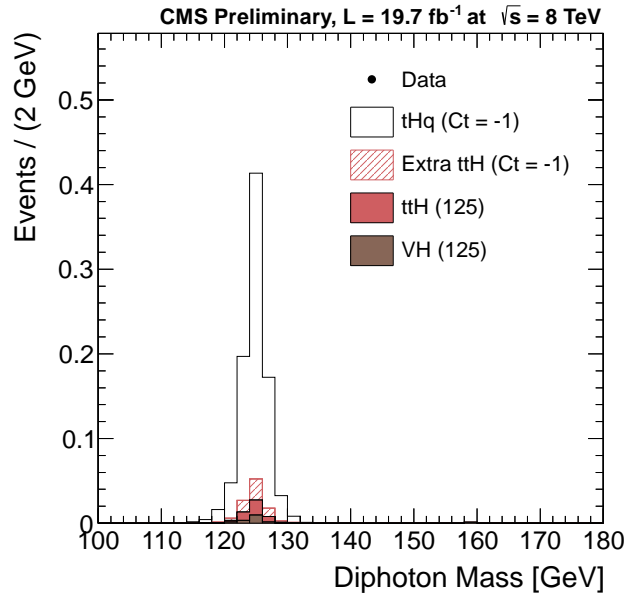
Given that zero events are observed in data after full selection,  $f_{bg}$  cannot be estimated fitting directly data, as in SM  $H \rightarrow \gamma\gamma$  analysis. A different approach is followed here, using control regions to derive a robust estimate of the shape of non-resonant background.

Four control samples are derived from data, which have looser selection requirements compared to the nominal selection:

- **CSVL**: similar to the nominal selection, but defining the b-jet candidate as the hardest jet in the event which passes the ‘loose’ CSV requirements, instead of the ‘medium’ one used in the nominal selection (cfr. par. 3.2);
- **CSV0**: same selection as the signal region, but defining the b-jet candidate as the hardest jet in the event with  $|\eta| < 2.4$ , with no b-tagging requirement;
- **InvID CSVL**: same selection as **CSVL**, but with inverted photon ID criteria on one of the two photons;
- **InvID CSV0**: same selection as **CSV0**, but with inverted photon ID criteria on one of the two photons.

It must be noted that the control samples are not statistically independent, as for instance **CSVL** is a subset of **CSV0**. The inversion of photon identification criteria follows what was done in  $t\bar{t}H$  category of the SM analysis (cfr. 4.3.3) and the same reweighting procedure is followed.

The event samples identified with these control regions have high enough yields that fits can be performed to the sideband regions. The chosen functional form of the



**Figure 5.4.** Invariant mass of the diphoton system for events passing full selection requirements. The data (black markers) are compared to Monte Carlo simulations (stacked histograms). As can be seen, zero events are observed in the data.

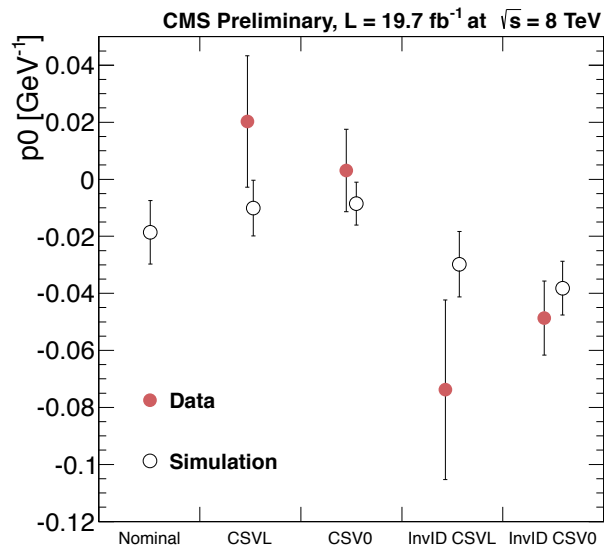
background is an exponential, and the fit is performed with an unbinned maximum likelihood technique. As a cross-check, the same fit is then performed on background simulations, and the results of these fits are compared in Figure 5.5, where the fitted slope of the exponential for events passing the nominal selection and in the four control regions is shown. As can be seen a good agreement, within statistical precision, is found between the fits performed in the data (solid red markers) and in the simulation (open black markers).

The fitted shape of the CSV0 control region, which is the one with the lowest statistical uncertainty for the exponential shape, is taken as the nominal background shape. This yields  $\alpha = 7.7\%$ . The full difference between using this shape and using the InvID CSV0 shape ( $\alpha = 10.2\%$ ) is taken as a systematic uncertainty on the contribution of non-resonant backgrounds under the signal mass peak. Therefore, the systematic uncertainty on the expected nominal background amounts to 33%.

#### 5.1.4 Systematic Uncertainties

Table 5.2 summarizes the systematic uncertainties of this analysis. Some of them are in common with the SM  $H \rightarrow \gamma\gamma$  analysis and the details on how they have been derived have been described in paragraph 4.5. The derivation of systematic uncertainty on the background shape has been described in the previous paragraph.

In addition to these uncertainties, an uncertainty is considered on tHq signal model. Indeed, the simulated signal sample has been generated with a five-flavor parton scheme, in which the only production diagram is  $qb \rightarrow tHq$ . A possible systematic uncertainty related to this choice is derived by computing the expected efficiency on a signal sample generated in a four-flavor parton scheme, in which



**Figure 5.5.** Fitted exponential slopes ( $p_0$ ) for events passing the nominal selection and in the four control regions in the data (red markers) and in the simulation (open black markers).

the signal is produced via  $qq' \rightarrow tHqb$ . The selection efficiency on  $qq' \rightarrow tHqb$  is found to be about 17% lower than on  $qb \rightarrow tHq$ . The former process is expected to contribute to 31% of the total cross section [98], so this translates into a difference of 5.5%, which is taken as a systematic uncertainty.

### 5.1.5 Results

As already said, after applying selection requirements zero events are observed in data. Resonant background contribution is estimated through simulations while the non-resonant one is obtained using a background shape estimation done on control samples.

The analysis, considering systematic uncertainties, has an expected significance of  $1.2 \pm 1.2$  standard deviations on tHq discovery, and an expected upper limit on tHq cross section for  $C_t = -1$  of 4.1 times the theoretical cross section. In this evaluation,  $C_t = -1$  is assumed and all extra H yields arising from this condition are considered as signal. The statistical approach followed in this evaluation is the same used in Section 4.4.

As no event is observed in the signal region, the observed UL coincides with the expected one: that is an observed 95% UL is set at 4.1 times the expected cross section (with  $C_t = -1$ ).

**Table 5.2.** Summary of the adopted systematic uncertainties.

	tHq	$t\bar{t}H$	VH	Continuous BG
Luminosity	$\pm 2.6\%$	$\pm 2.6\%$	$\pm 2.6\%$	-
PDF	+3.1/-2.5 %	$\pm 8\%$	$\pm 11\%$	-
QCD Scale	+4.8/-4.3 %	+11/-14 %	$\pm 2.3\%$	-
Signal Model	$\pm 5.5\%$	-	-	-
Photon Energy Resolution	+4/-2 %	+4/-2 %	+4/-2 %	-
Photon Energy Scale	+1/-4 %	+1/-4 %	+1/-4 %	-
Photon ID Efficiency	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	-
Vertex Efficiency	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.1\%$	-
HLT	$< 0.1\%$	$< 0.1\%$	$< 0.1\%$	-
JEC	$\pm 1.5\%$	+3/-5 %	$\pm 8\%$	-
JER	$\pm 0.5\%$	$\pm 3\%$	+8/-0 %	-
$b$ -tagging	$\pm 2\%$	$\pm 1.5\%$	$\pm 0.1\%$	-
PU ID	$\pm 2\%$	$\pm 0.5\%$	$\pm 2\%$	-
Lepton Reconstruction	$\pm 1\%$	$\pm 1\%$	$\pm 1\%$	-
BG shape	-	-	-	33%

**Table 5.3.** Summary of the event selection.

Leading photon with $p_T > 50 \cdot m_{\gamma\gamma}/120$ GeV
Subleading photon with $p_T > 25$ GeV
Exactly one lepton (e/ $\mu$ ) with $p_T > 10$ GeV
At least one b-jet with $p_T > 20$ GeV
Hardest jet (not the b-Jet) $p_T > 20$ GeV and $ \eta  > 1$
LD > 0.25

## 5.2 Search for Pair Production of Heavy Vector-Like Partners of the Top Quark

As seen in Paragraph 1.5.2 several theoretical models predict the existence of a new strong sector of heavy vector-like quarks. Searches presented here focus on partners of the top quark (T).

Different decays of this particle are possible, depending on different assumptions on its nature. If the T quark is a weak isospin doublet, it can decay through flavor changing neutral currents to a top quark and a Higgs or Z boson. If it is a weak isospin singlet, it can additionally decay to a W boson and a b quark similarly to a sequential fourth generation of quarks.

Older searches for heavy vector-like quarks focused separately either on the  $T \rightarrow bW$  or  $T \rightarrow tZ$  final states [99–103]. The precise knowledge of the Higgs boson mass now allows to target the  $T \rightarrow tH$  decay as well.

In this chapter, searches for T pair production where at least one of them decays to a top quark and a Higgs boson are presented. The search is focused on the  $H \rightarrow \gamma\gamma$  decay mode. The presence of the two photons in this search ensures that a Higgs boson is actually present in the final state. This is useful, in case of a discovery, to precisely measure the key parameters of the theory.

The final state, given the complex decay chain, is crowded: in addition to the photons coming from the Higgs, further jets and leptons arise from the decay of the other top and Higgs. This topology is similar to SM  $t\bar{t}H$  production and analogous techniques are used, even if an optimized selection is used according to different kinematics and to reject SM background.

Throughout the analysis, the SM branching ratio of Higgs to photons,  $\mathcal{B}(H \rightarrow \gamma\gamma) = 0.0023$ , is used. This corresponds to the assumption that the presence of top quark partners in the Higgs to photons loop would not change the branching ratio. General variations of the above branching ratio are anyway possible and the results can be interpreted by scaling up or down the values of  $\sigma(T\bar{T}) \times \mathcal{B}$  used here.

### 5.2.1 Analysis Strategy

The strategy used in this top partners search shares the same approach of other diphoton analysis: background estimation is performed using a fit to the diphoton mass distribution. Expected signal and resonant contributions, instead, are estimated using simulations.

Signal samples with varying top quark partner mass and decay modes have been obtained through MADGRAPH version 5 [68] for the hard scatter and PYTHIA version 6.4 [63] for parton shower/hadronization and underlying event modeling. The cross sections for  $T\bar{T}$  production have been computed at next-to-next-to-leading order (NNLO) + next-to-next-to-leading logarithm (NNLL) soft gluon resummation using the Top++2.0 computer program [104], and MSTW2008nnlo68cl parton density functions (PDFs) [105] and are tabulated in table 5.4. As can be seen from the table, the expected cross sections are sizeable in a wide range of  $m_T$  and therefore an evidence of these particles could be possible even with the current amount of data. Uncertainties in the cross section due to the uncertainty in the underlying parton density functions and from the choice of the renormalization and factorization

scale are approximately 2.1% and 4.3%, respectively, in the top quark mass range of interest for the analysis [106].

**Table 5.4.** Cross section  $\sigma$  in pb for  $T\bar{T}$  pair production computed for different  $pp$  center of mass energies  $\sqrt{s}$  at NNLO+NNLL [106].

$m_T(\text{GeV})$	$\sqrt{s} = 8 \text{ TeV}$	$\sqrt{s} = 13 \text{ TeV}$
400	2.386	11.67
450	1.153	6.091
500	0.590	3.359
550	0.315	1.938
600	0.174	1.161
650	0.0999	0.717
700	0.0585	0.455
750	0.0350	0.295
800	0.0213	0.196

As in single top associate production, SM Higgs production is a background resonant in  $m_{\gamma\gamma}$  in this BSM search. The only relevant production mechanism for this search is  $t\bar{t}H$  production and its cross-section is estimated through Monte Carlo simulation (cfr. Par. 4.2).

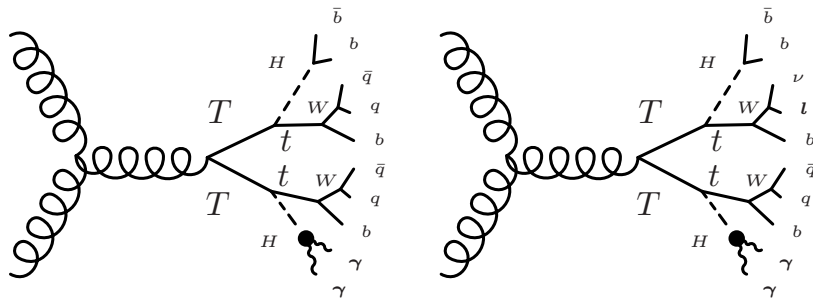
Non-resonant background contributions in this analysis are the same as in SM exclusive categories: direct QCD production as well as emission in top quark production ( $\gamma\gamma$ +jets,  $t\bar{t}+\gamma\gamma$ ,  $t+\gamma\gamma$ ). Jets in  $t\bar{t}$  events can also be misreconstructed as photons due to the high number of jets in the final state. No simulations are used to estimate these contributions, since data-driven control samples are used for optimization while data are used for limit extraction.

To achieve best sensitivity, two different classes of possible final states are targeted, according to the decay mode of the top quark partner.

Two search channels are therefore defined:

- the **hadronic channel** searches for events with a pair of photons and no isolated electrons or muons. This channel targets events in which a top quark partner decay chain leads only to quarks, and the top and Higgs boson, coming from the other top partner, decay both hadronically. An example of this kind of diagrams is given in the left side of Fig. 5.6;
- the **leptonic channel** searches for events with a pair of photons and at least one isolated high- $p_T$  electron or muon. This channel targets events where at least one charged lepton appears either in the top quark decay chain, or in the decay chain stemming from the second top quark partner. A Feynman diagram representing one such example is depicted in the right side of Fig. 5.6.

Different event selections will be optimized for the two categories of events and the final results will be extracted combining the two.



**Figure 5.6.** Example Feynman diagrams for top quark partners production and decay in the hadronic (left) and leptonic (right) channels

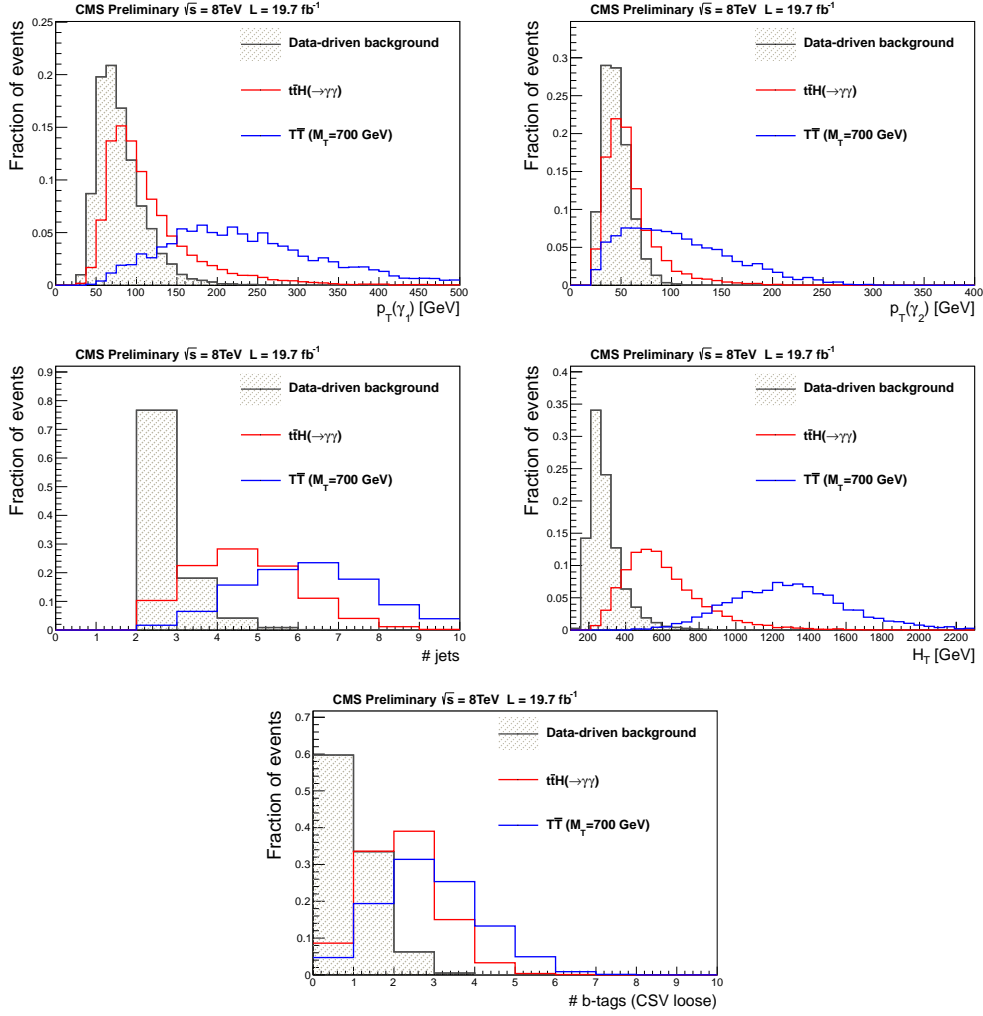
### 5.2.2 Event Selection

The aim of the event selection is to separate signal process from the reducible non-resonant background contribution. Given that simulations are not fully reliable in the phase space of interest of this analysis (events with high number of jets and leptons in the final state have big uncertainties in simulations and not yet completely understood discrepancies with data) the non-resonant background contribution for optimization is estimated using a data-driven control sample. This control sample is derived inverting photon identification criteria, following what was done in the  $t\bar{t}H$  category of the SM analysis (cfr. 4.3.3) and the same reweighting procedure is followed.

The optimization is performed, separately for the leptonic and the hadronic channel, using  $T\bar{T}$  Monte Carlo simulations with  $m_T=700$  GeV as signal and the sum of the data-driven background and  $t\bar{t}H$  simulation as background. Different kinematic variables, for both photons and jets, are investigated and the most discriminating and less correlated ones are used as input variables for a multivariate optimization using TMVA. Figure 5.7 shows the distributions of these variables for different background contributions and signal. The variable  $H_T$  is defined as the sum of the transverse momenta of all the objects in the final state: photons, jets, and leptons. This variable is particularly discriminating given the high number of high-energy particles in  $T\bar{T}$  final state with respect to background processes. As can be seen from these distributions, the signal is characterized by high- $p_T$  photons, the  $p_T$  spectrum is even harder than  $t\bar{t}H$  production, high jet and b-jet multiplicity.

The optimization process provides a series of cut values for each variable maximizing background rejection for a given signal efficiency. The optimal working point (WP) is chosen using as a figure of merit the 95 % confidence level (CL) expected upper limit (UL) on  $\sigma(T\bar{T})$ . The final selection for the two channels, corresponding to the chosen WP, is reported in Table 5.5.

After having defined the selection criteria, since the background is extracted from a fit to the diphoton invariant mass, an appropriate fitting function has to be chosen. Studies performed on pseudo-experiment proved that a simple exponential can be used without introducing any relevant bias in background estimation.



**Figure 5.7.** Distributions for  $p_T$  of the leading and subleading photon (top right and left), number of jets and  $H_T$  (center left and right) and number of b jets (bottom). All distributions are normalized to unity. Signal with  $m_T=700 \text{ GeV}$  (blue),  $t\bar{t}H$  (red) and the data-driven non resonant background (grey) contributions are shown. All contributions are normalized to unity.

### 5.2.3 Systematic Uncertainties

All simulation-derived experimental systematic uncertainties apply to signal and to  $t\bar{t}H$  resonant background. They are summarized in Table 5.6. All of them are in common with  $H \rightarrow \gamma\gamma$  and single top Higgs associate production analysis and are derived following the same procedures.

### 5.2.4 Results

An unbinned maximum-likelihood fit is performed to the diphoton mass distribution, in the range 100-180 GeV, under the hypothesis of no signal, in both categories. This fit is used to estimate non resonant background contribution. The result of the



**Table 5.5.** Final selection for leptonic and hadronic channel

Variable	Leptonic channel	Hadronic channel
$p_T(\gamma_1)$	$> \frac{1}{2}m_{\gamma\gamma}$ GeV	$> \frac{3}{4}m_{\gamma\gamma}$ GeV
$p_T(\gamma_2)$	25 GeV	35 GeV
$n_{\text{jets}}$	$\geq 2$	$\geq 2$
$H_T$	$\geq 770$ GeV	$\geq 1000$ GeV
leptons	$\geq 1$	0
b tags	-	$\geq 1$

**Table 5.6.** Summary of the adopted systematic uncertainties.

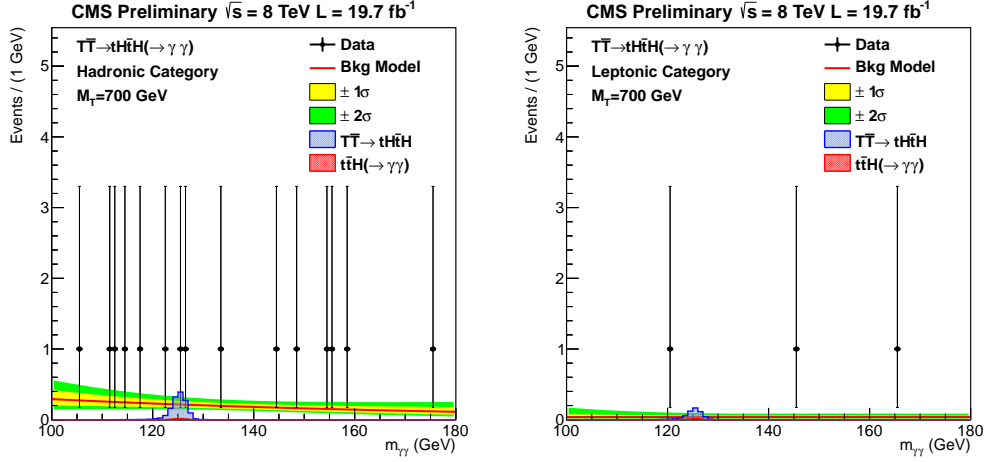
	T $\bar{\text{T}}$	t $\bar{\text{t}}\text{H}$
Luminosity	$\pm 2.6\%$	$\pm 2.6\%$
PDF	-	$\pm 8.1\%$
QCD scale	-	+4/-9 %
Photon Energy Resolution	+4/-2 %	+4/-2 %
Photon Energy Scale	+1/-4 %	+1/-4 %
Photon ID Efficiency	$\pm 2\%$	$\pm 2\%$
Trigger	$< 0.1\%$	$< 0.1\%$
JEC	$\pm 2\%$ (had) $\pm 1\%$ (lep)	$\pm 7\%$ (had) $\pm 5\%$ (lep)
JER	$\pm 1\%$	$< 0.5\%$
<i>b</i> -tagging	$< 0.5\%$ (had)	$< 0.5\%$ (had)
Pile-up identification	$\pm 2\%$	$\pm 2\%$
Lepton Reconstruction	$\pm 1\%$ (lep)	$\pm 1\%$

fits is shown in figure 5.8. The error bands show the uncertainty on the background shapes associated with the statistical uncertainties of the fits. Signal contribution, after full selection, is estimated using simulations. The expected yields for signal (with  $m_T = 700$  GeV) and backgrounds are summarized in Tab. 5.7. All the yields are computed in a window of 1.5 GeV around the signal peak, which corresponds to 1 full width at half maximum of the  $m_{\gamma\gamma}$  distribution for the Higgs boson.

**Table 5.7.** Expected yields for signal, t $\bar{\text{t}}\text{H}$  and non-resonant background (from the fit to data) in  $19.7 \text{ fb}^{-1}$  after full event selection. All the yields are computed in a window of 1 full width at half maximum (FWHM =  $\pm 1.5$  GeV) around 125 GeV.

Process	Leptonic	Hadronic
T $\bar{\text{T}}(m_T = 700 \text{ GeV})$	0.43	1.05
t $\bar{\text{t}}\text{H}$	0.039	0.042
Background	$0.11^{+0.07}_{-0.03}$	$0.65^{+0.16}_{-0.13}$
Observed Data	0	2

No significant excess is observed in data, with respect to background-only predic-



**Figure 5.8.** Diphoton invariant mass distribution for candidate  $T\bar{T}$  events for the hadronic (left) and leptonic (right) channel. The signal is normalized to the predicted theoretical cross section corresponding to  $m_T = 700$  GeV. Background predictions coming from the fit are shown as a red line, and bands corresponding to 68% (yellow) and 95% (green) are added.

tions, in the signal window. An upper limit on top partner production cross section can be therefore set, using the modified frequentist  $CL_s$  approach, taking the profile likelihood as a test statistic [107, 108]. To derive exclusion limits, an assumption should be made on the branching ratio of the top partners: all the exclusion limits results are firstly obtained under the assumption of  $\mathcal{B}(T \rightarrow tH) = 100\%$ . This hypothesis is dropped later to allow for alternative interpretations in the phase space of  $T$  branching ratios.

Figure 5.9 shows the 95% confidence exclusion limit on the production cross section times branching ratio, for both search channels, as a function of the top partner mass. The expected limit is shown as a dotted black line, and the bands corresponding to 68% (yellow) and 95% (green) probability are added. The observed limit is represented by a black line. As can be seen, the hadronic channel expected exclusion reaches top quark partner masses up to 538 GeV while in the leptonic channel the expected exclusion region arrives up to 522 GeV. The two channels can be combined together and the overall exclusion limit for top partner production in the diphoton decay channel is shown in figure 5.10. The expected and observed upper limit on the top quark partner cross section for the two categories and their combination are summarized in table 5.8. The final observed (expected) exclusion bound to the existence of top quark partners reaches a mass mass of 540 (607) GeV.

The assumption  $\mathcal{B}(T \rightarrow tH) = 100\%$  can be relaxed for a more general interpretation of the results. Indeed, exclusion limits can be derived not only as a function of  $m_T$  but also leaving the branching ratio to top and Higgs as a free parameter. In Fig. 5.11 the expected and observed exclusion region is shown in the 2-dimensional space of  $\mathcal{B}(T \rightarrow tH)$  and  $m_T$ . The variable represented by the color palette on the right of the plot corresponds to the excluded cross section ( $\sigma^{UL}$ ), normalized to the theoretical one,  $\sigma^{theo}$ , for a given top partner mass. The solid line corresponds to

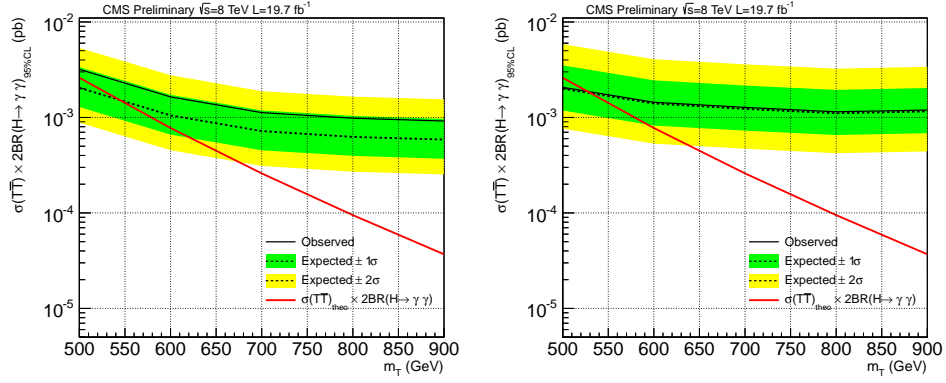


Figure 5.9. 95% C.L. upper limit on heavy vector-like partners of the top quark (T), with  $T \rightarrow tH(\rightarrow \gamma\gamma)$ , production in the hadronic (left) or leptonic (right) channel.

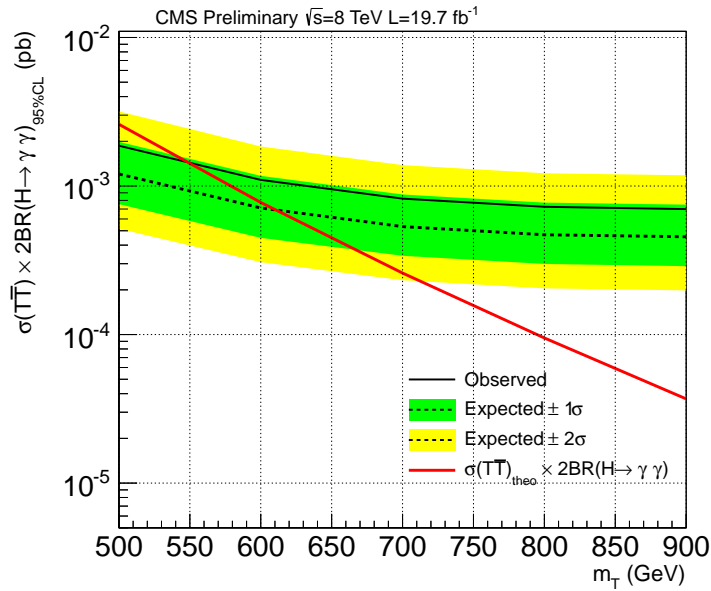


Figure 5.10. 95% C.L. upper limit on heavy vector-like partners of the top quark (T), with  $T \rightarrow tH(\rightarrow \gamma\gamma)$ , production is shown for the combination of the two channels

**Table 5.8.** Summary of observed and expected 95% upper limits (UL) to the production cross section of heavy vector-like partners of the top quark (T), with  $T \rightarrow tH(\rightarrow \gamma\gamma)$ : the observed and expected limit are given for the hadronic channel, the leptonic channel, and their combination.

	Observed UL	Expected UL for $m_T = 700$ GeV
Hadronic Channel	0.25 pb	0.16 pb
Leptonic Channel	0.29 pb	0.27 pb
Combined	0.19 pb	0.12 pb

the boundary of the excluded region, where the excluded cross section corresponds to the theoretical one, which means  $\frac{\sigma^{UL}}{\sigma^{theo}} = 1$ .

In the same phase space, expected exclusion with different assumptions on branching ratio values can be computed. In figure 5.12 this is shown for two different hypotheses:  $\mathcal{B}(T \rightarrow Wb) = 0$  and  $\mathcal{B}(T \rightarrow tZ) = 0$ . As can be seen from the plot, the analysis is almost only sensitive to  $\mathcal{B}(T \rightarrow tH)$  and different assumptions on other branching ratios do not change notably the expected exclusion region.

Nevertheless, a complete interpretation of the result can be given also in the phase space of the three possible branching ratios:  $\mathcal{B}(T \rightarrow tH)$ ,  $\mathcal{B}(T \rightarrow bW)$  and  $\mathcal{B}(T \rightarrow tZ)$ . The two triangles in Fig. 5.13 shows the expected and observed limits on T quark mass as a function of the branching ratios in the different decay channels. In the white region of the triangle, which corresponds to the part where  $\mathcal{B}(T \rightarrow tH)$  is too small, the presence of any T quark with a mass greater of 500 GeV cannot be excluded.

In conclusion, searches for  $T\bar{T}$  in the diphoton decay channel are very performing and the results can be interpreted in very different scenarios. This search is particularly sensitive to the scenario where  $\mathcal{B}(T \rightarrow tH)$  is close to unity, while small sensitivity has been found to other possible branching ratios of top partners decays. No significant excess has been found, since all observed results are compatible with background-only expectations. Under the hypothesis that  $\mathcal{B}(T \rightarrow tH) = 1$  values of the top quark mass up to 540 (607) GeV are observed (expected).

### 5.2.5 Combination with Other channels

Given the wide range of possible decays of fermionic top partners, very different analyses can be combined together to obtain overall exclusion limits on the production of these particles.

CMS has designed different analyses to target all possible decay modes:  $T \rightarrow bW$ ,  $T \rightarrow tH$  and  $T \rightarrow tZ$ . No assumption is made on the relative weight of these branching ratios and limits on production cross section are set for different possible combinations. Five different T quark searches, optimized for different topologies, have been performed. One of the analyses [109] is inclusive and sets limits for all possible branching ratios scenarios. The other four analyses have better sensitivity in their respective optimized regions, but they do not cover the full range of branching ratios. A common strategy to all searches is to impose hard cuts on the number

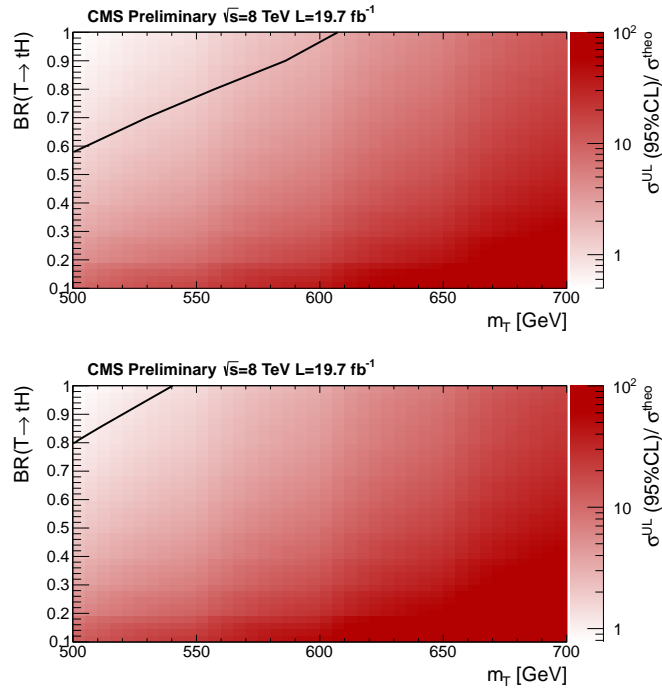


Figure 5.11. Expected (top) and observed region (bottom) of exclusion in the 2D phase space of  $\mathcal{B}(T \rightarrow tH)$  and  $m_T$ . The solid black line corresponds to  $\frac{\sigma^{UL}}{\sigma^{theo}} = 1$

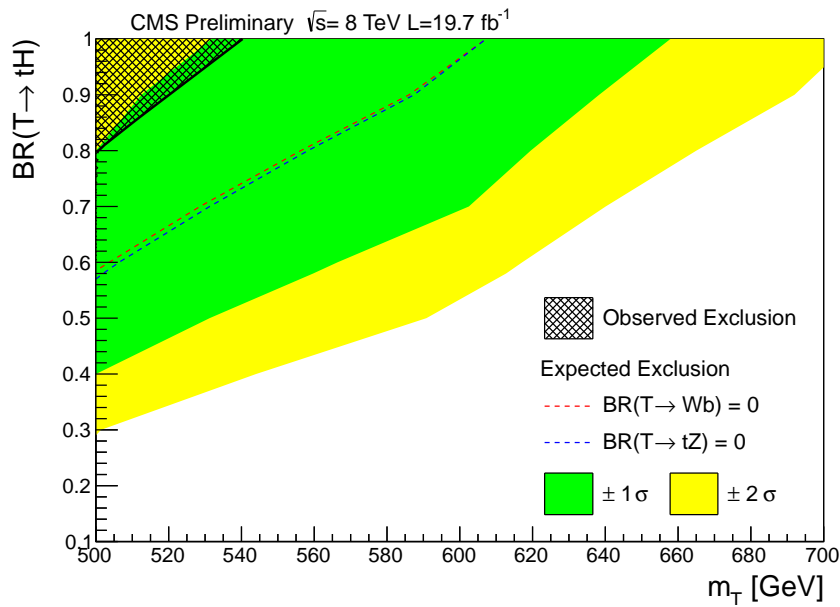
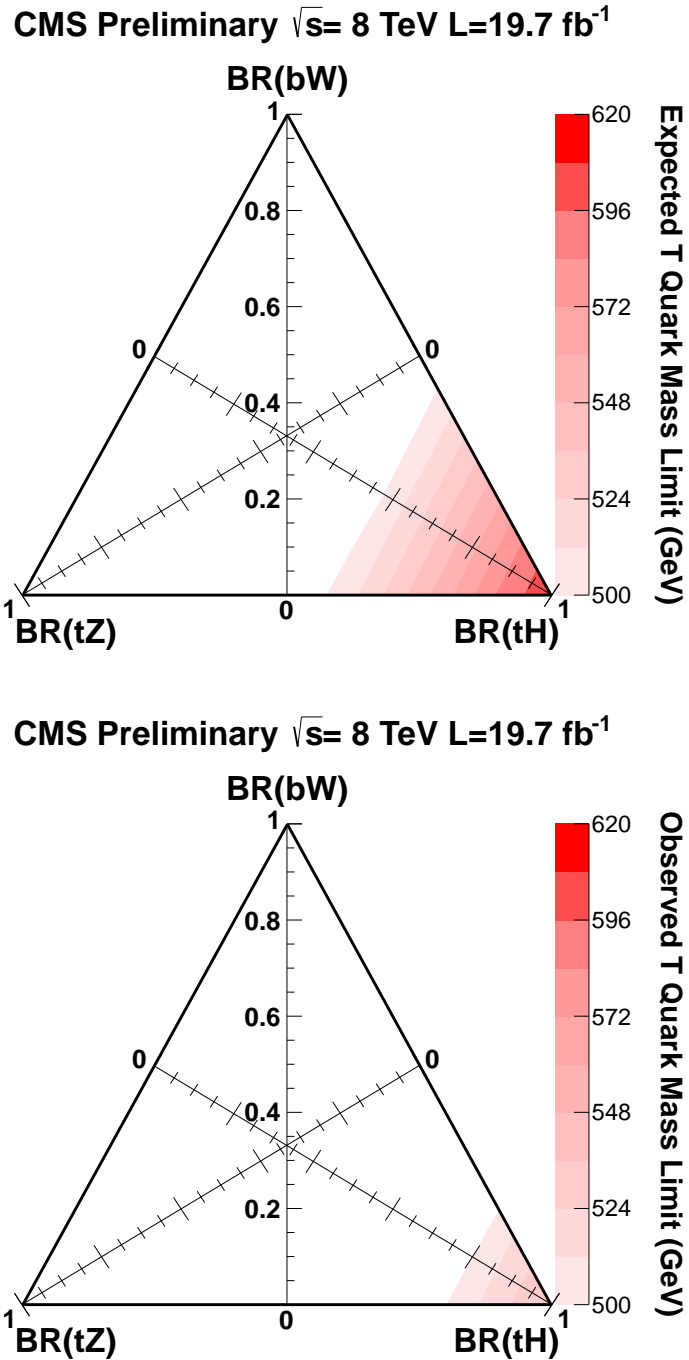


Figure 5.12. Expected and observed region of exclusion in the 2D phase space of  $\mathcal{B}(T \rightarrow tH)$  and  $m_T$ . Different dashed lines correspond to different assumptions on the branching ratios of the top partner:  $\mathcal{B}(T \rightarrow Wb) = 0$  (and consequently  $\mathcal{B}(T \rightarrow tZ) = 1 - \mathcal{B}(T \rightarrow tH)$ ) in red and  $\mathcal{B}(T \rightarrow tZ) = 0$  (and consequently  $\mathcal{B}(T \rightarrow Wb) = 1 - \mathcal{B}(T \rightarrow tH)$ ) in blue.



**Figure 5.13.** Branching fraction triangle with expected (top) and observed (bottom) limits for the T quark mass. We cannot exclude any T quark with a mass greater of 500 GeV in the white region of the triangle.

of jets and b-jets in the final state and on  $H_T$ , according to  $T\bar{T}$  topology, while different techniques are used depending on the decay channel targeted.

Inclusive searches are performed in two different channels, according to the presence of one or multiple leptons in the final state [109]. The main backgrounds in these searches are SM prompt production of a pair of top quarks,  $t\bar{t}$ , or of W- and Z-boson. Multivariate techniques, using jet and b-jet multiplicity and kinematics, are employed in single-lepton searches. In multilepton channels, charge correlations are also used to reduce background contribution.

Specific searches for  $T \rightarrow bW$  decays are performed in single-lepton [110] and all-hadronic [111] final states. Tools for inspecting substructure of jets [112] are used in these searches, since decay products of the T quark are expected to be highly boosted and secondary hadronic decays may get merged into single jets (fat jets) with intrinsic substructure. Top- and W- tagging techniques are exploited to identify jets from hadronic decays of these particles [113–116]. In the single-lepton searches, one lepton is required in the final state and a kinematic fit, constraining the mass of decay products to those of the W boson, is used in the hypothesis  $T\bar{T} \rightarrow bW^+\bar{b}W^-$ . All-hadronic searches in the  $T \rightarrow bW$  channel target, instead, events where the W decays hadronically. In addition to two W-tagged, additional jets and b-jets are required to fully reconstruct the T pair. QCD multijet events contribute the largest fraction to the standard model background in this search and the modeling of this kind of background is derived from data.

The  $T \rightarrow tH$  decay mode, in addition to the diphoton analysis already presented, is also investigated in the all-hadronic decay mode [117]. Substructure requirements on jets are strongly used to devise an optimized selection reconstructing the final state in its entirety.

All the analyses in the different channels are combined together. Results are extracted using the statistical approach used in previous chapters. Given that no striking excess over background-only predictions is observed, exclusion limits on the expected heavy top partners cross section can be set, in different hypotheses of branching ratios.

Figure 5.14 shows the expected limits for individual analysis and their combination in the case of exclusive branching ratio. Observed limits, along with the expected uncertainty bands, are displayed instead in Figure 5.15.

Table 5.9 summarizes the expected (observed) exclusion limits obtained determining the crossing point between expected (observed) limits with the theoretical predictions. These results can be visualized using triangular planes, as already shown for the diphoton analysis. Combining all searches in these plots allows to see the complete picture of CMS exclusion limits for this kind of particle. Expected and observed triangle plots are shown in figure 5.16.

### 5.2.6 Summary of T searches

Searches for  $T\bar{T}$  in the diphoton decay channel have been presented in the first part of this section. No significant excess has been found analyzing 8 TeV dataset, and the observed excluded region, under the hypothesis that  $\mathcal{B}(T \rightarrow tH) = 1$ , reaches 540 GeV.

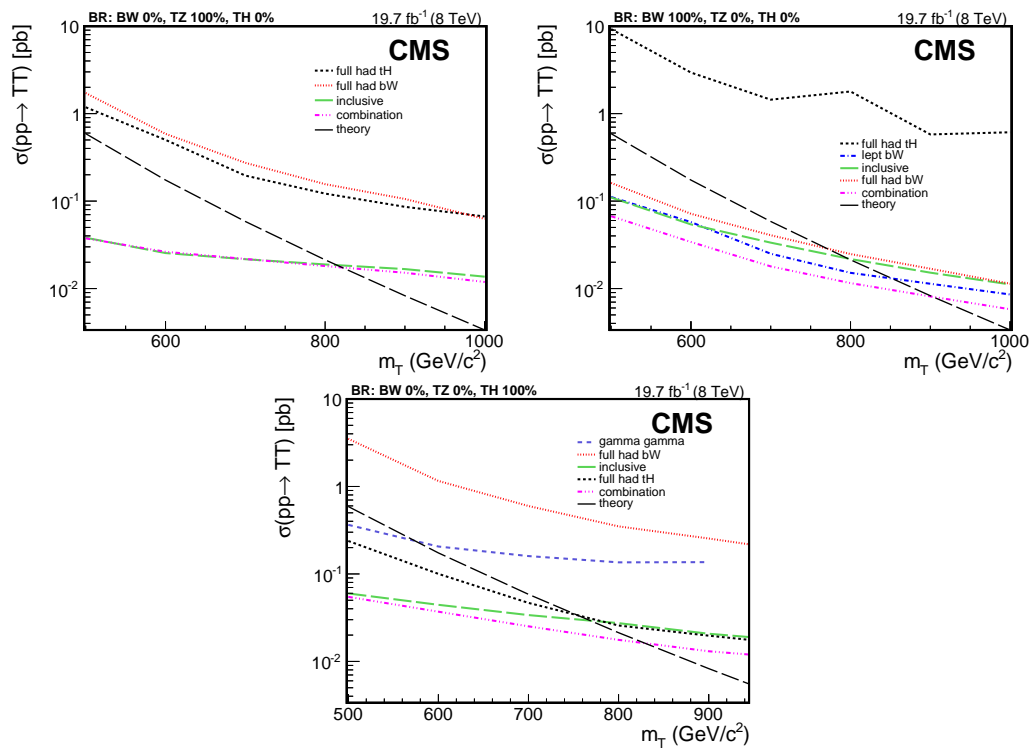
Since the hypothetical T quark has three decay modes ( $T \rightarrow bW$ ,  $T \rightarrow tH$

$\mathcal{B}(bW)$	$\mathcal{B}(tZ)$	$\mathcal{B}(tH)$	obs. limit (GeV)	exp. limit (GeV)	expected $1\sigma$ (GeV)
0.0	1.0	0.0	787	831	[783,878]
0.0	0.8	0.2	776	822	[776,875]
0.0	0.6	0.4	765	818	[775,872]
0.0	0.4	0.6	764	823	[773,877]
0.0	0.2	0.8	766	822	[775,881]
0.0	0.0	1.0	765	843	[782,891]
0.2	0.8	0.0	780	822	[777,876]
0.2	0.6	0.2	753	811	[767,871]
0.2	0.4	0.4	746	807	[760,870]
0.2	0.2	0.6	751	798	[761,874]
0.2	0.0	0.8	749	821	[769,880]
0.4	0.6	0.0	758	812	[766,875]
0.4	0.4	0.2	744	799	[753,874]
0.4	0.2	0.4	741	798	[748,872]
0.4	0.0	0.6	734	808	[758,879]
0.6	0.4	0.0	749	817	[765,882]
0.6	0.2	0.2	744	810	[755,880]
0.6	0.0	0.4	730	818	[760,884]
0.8	0.2	0.0	907	866	[797,918]
0.8	0.0	0.2	896	867	[792,920]
1.0	0.0	0.0	946	905	[855,981]

**Table 5.9.** Lower limits for the T quark mass, at 95% CL, for different combinations of T quark branching fractions.



and  $T \rightarrow tZ$ ) different exclusive searches can be considered and combined together in different hypotheses of branching ratios. No excess has been observed over background expectations in the whole mass range and limits on the production cross-section of vector-like  $T$  quarks are set. The expected limits vary between 800 GeV and 890 GeV depending on the branching ratio of the  $T$  quark. For a branching ratio of  $\mathcal{B}(bW) = 100\%$  an expected (observed) limit of 905 GeV (946 GeV) is found. For  $\mathcal{B}(tH) = 100\%$  the expected (observed) limit is 843 GeV (765 GeV) while for  $\mathcal{B}(tZ) = 100\%$  the expected (observed) limit is 831 GeV (787 GeV).



**Figure 5.14.** Expected limits of the individual analyses in comparison to the result of the combination for different branching ratios scenarios

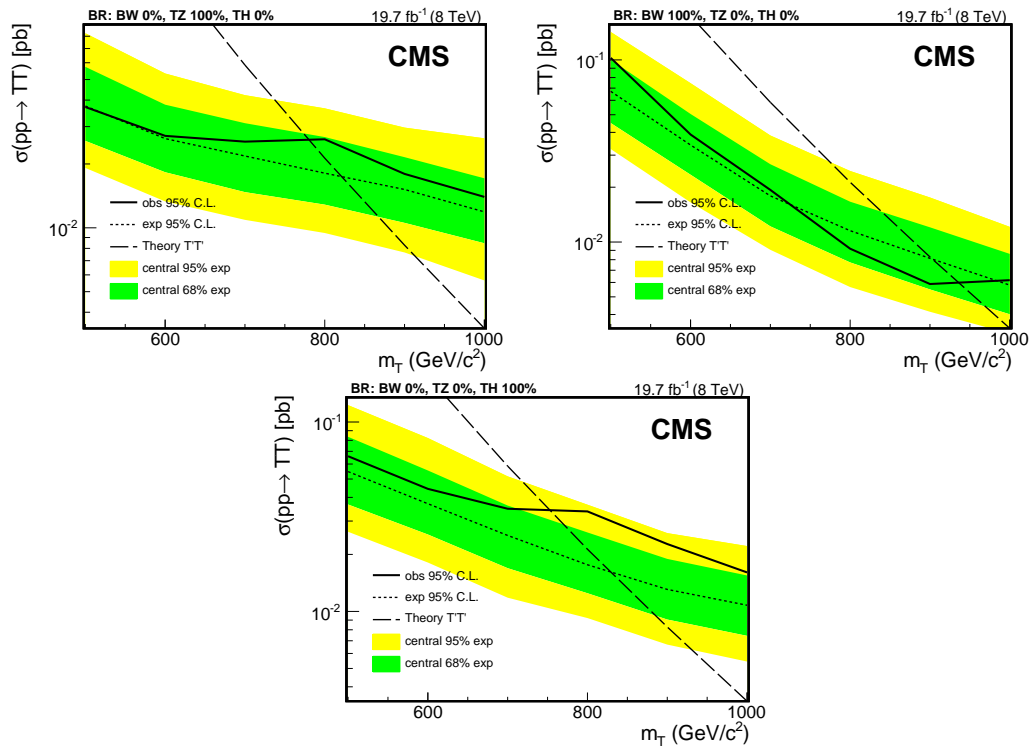


Figure 5.15. Observed limits (solid black line) along with the expected one-sigma and two-sigma uncertainties for different branching ratios scenarios.

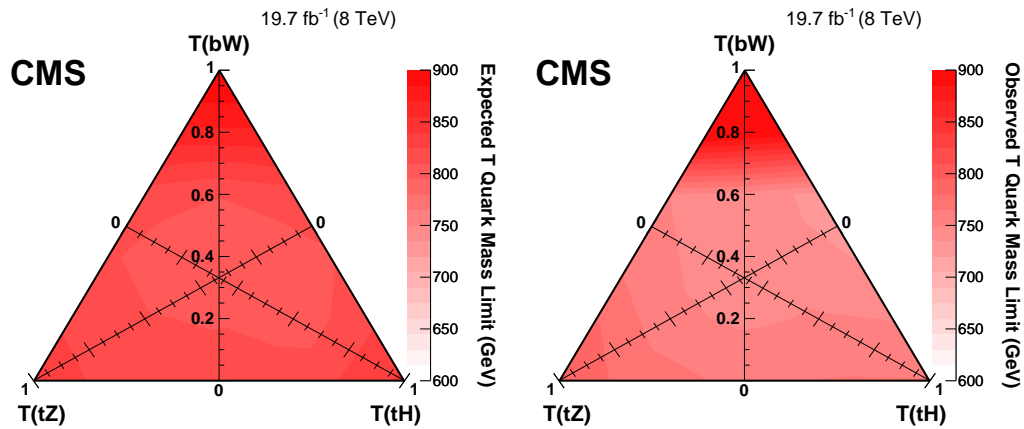


Figure 5.16. Expected (left) and observed (right) limits of the combined analysis, visualized in a triangle representing the decay branching ratios.



## Chapter 6

# Conclusions

This thesis analyzes in details the discovery potential of the diphoton decay of the Higgs boson for new physics searches. Both direct and indirect searches are investigated in the  $H \rightarrow \gamma\gamma$  decay channel.

Given the presence of two high energy photons, the strategy used in all searches is to identify, in the diphoton mass distribution, the narrow Higgs boson mass peak over a continuum background spectrum. The main advantages of this channel are: the possibility to estimate background contribution directly from data, the good signal over background ratio and the possibility to measure the mass. A good accuracy on photon energy and vertex reconstruction is required in order to achieve the mass resolution needed for this kind of searches. The main background contribution comes from prompt production of events with photons and jets in the final state. Consequently, dedicated photon identification and reconstruction criteria have been designed to reduce these backgrounds.

Measurements of the Higgs boson couplings to other particles have been presented in this thesis. These measurements have been done devising exclusive categories, tagging different production modes of the Higgs boson. The presence of additional objects has been used to further suppress the background and increase the sensitivity. Results are presented in different benchmark models, grouping production mechanisms and using different possible scale factors for Higgs boson production cross sections and decay branching ratios. The approach is as model independent as possible, and only reasonable general assumptions are made in the search of possible deviations from the SM predictions.

No striking deviations from theoretical predictions have been observed; the measured signal strengths for different production modes are compatible with SM expectations. A mild excess, roughly two sigmas over SM expectations in the combination of all decay channels, is seen in  $t\bar{t}H$  production mode. Thanks to the inclusion of dedicated categories for associate production modes in  $H \rightarrow \gamma\gamma$  analysis the precision on couplings measurements has significantly increased. These measurements can be used to constrain several BSM models such as 2HDM and VLQ.

Direct searches of new physics effects have been performed focussing on two different processes: associate production of a single top quark and a Higgs boson ( $tHq$ ), and pair production of heavy vector-like partners of the top quark (T) decaying

in top and Higgs. The first process can highlight anomalous top-Higgs coupling while the second one could be an evidence a new strong sector of heavy quarks, useful to solve the hierarchy problem. These processes have similar final states to associate SM Higgs production. Optimized searches have been designed taking advantage of experimental techniques developed for Higgs coupling measurements.

No excess over background-only expectations is found in both analyses. An upper limit of 4.1 times the theoretical cross section is therefore set for anomalous  $t\bar{t}Hq$  production. Instead, the existence of vector-like partners of the top quarks, decaying completely in top and Higgs, is excluded up to a mass of 550 GeV in the diphoton decay channel. The combination with other decay channels increases the lower bound on their mass up to 750 GeV.

The presence of one or more top quarks in the final state has a key role in creating optimized selections for the different searches presented in this thesis. This particle, given its unexpectedly high mass for a SM quark, has a great theoretical importance in Higgs physics and the precise measurement of its coupling to the Higgs boson is one of the physics goals of LHC. Top quarks have an important role in many BSM models, since often produced in association with Higgs bosons as decay products of heavy particles. The experimental signature of this kind of events with top quarks and Higgs bosons is particularly complex, given the presence of lots of different types of objects in the final state. The results presented here highlight the versatility of the CMS detector, which permits to reconstruct this kind of events with great accuracy in a difficult experimental ambient.

The original results contained in this thesis were published in several articles of the CMS Collaboration [91, 118–120] and are the most updated to quote for these searches in the  $H \rightarrow \gamma\gamma$  channel. Both direct and indirect searches presented in this thesis have pointed out the important role of the  $H \rightarrow \gamma\gamma$  process in new physics searches. The processes searched in this thesis will become of primary importance for the next LHC Run, which is foreseen to start at a center-of-mass energy of 13 TeV. The Standard Model  $t\bar{t}H$  production, for instance, is close to observation and limits on vector-like partners of the top quark could be extended up to the TeV scale, confirming or excluding most of BSM models in the first years of data-taking.

# Bibliography

- [1] S. L. Glashow, Partial-symmetries of weak interactions, Nuclear Physics 22 (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.
- [2] S. Weinberg, A model of leptons, Phys. Rev. Lett. 19 (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.
- [3] A. Salam, Weak and Electromagnetic Interactions, Conf.Proc. C680519 (1968) 367–377.
- [4] P. W. Higgs, Broken symmetries, massless particles and gauge fields, Phys.Lett. 12 (1964) 132–133. doi:10.1016/0031-9163(64)91136-9.
- [5] F. Englert, R. Brout, Broken symmetry and the mass of gauge vector mesons, Phys. Rev. Lett. 13 (1964) 321–323. doi:10.1103/PhysRevLett.13.321. URL <http://link.aps.org/doi/10.1103/PhysRevLett.13.321>
- [6] G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, Global conservation laws and massless particles, Phys. Rev. Lett. 13 (1964) 585–587. doi:10.1103/PhysRevLett.13.585. URL <http://link.aps.org/doi/10.1103/PhysRevLett.13.585>
- [7] N. Cabibbo, Unitary symmetry and leptonic decays, Phys. Rev. Lett. 10 (1963) 531–533. doi:10.1103/PhysRevLett.10.531. URL <http://link.aps.org/doi/10.1103/PhysRevLett.10.531>
- [8] M. Kobayashi, T. Maskawa, CP Violation in the Renormalizable Theory of Weak Interaction, Prog.Theor.Phys. 49 (1973) 652–657. doi:10.1143/PTP.49.652.
- [9] F. Mandl, G. Shaw, Quantum Field Theory, A Wiley-Interscience publication, John Wiley & Sons, 2010. URL <http://books.google.ch/books?id=Ef4zDW1V2LkC>
- [10] R. Barate, et al., Search for the standard model Higgs boson at LEP, Phys.Lett. B565 (2003) 61–75. arXiv:hep-ex/0306033, doi:10.1016/S0370-2693(03)00614-2.
- [11] Higgs boson studies at the tevatron, Phys. Rev. D 88 (2013) 052014. doi:10.1103/PhysRevD.88.052014. URL <http://link.aps.org/doi/10.1103/PhysRevD.88.052014>

- [12] Observation of an Excess of Events in the Search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Tech. Rep. ATLAS-CONF-2012-093, CERN, Geneva (Jul 2012).
- [13] Observation of a new boson with a mass near 125 GeV, Tech. Rep. CMS-PAS-HIG-12-020, CERN, Geneva (2012).
- [14] M. M. Kado, C. G. Tully, The searches for higgs bosons at lep, Annual Review of Nuclear and Particle Science 52 (2002) 65–113. doi:10.1146/annurev.nucl.52.050102.090656.
- [15] J. Alcaraz, Precision Electroweak Measurements and Constraints on the Standard Model arXiv:0911.2604.
- [16] T. Group, Design Report Tevatron 1 project.
- [17] Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in higgs boson searches at the tevatron, Phys. Rev. Lett. 109 (2012) 071804. doi:10.1103/PhysRevLett.109.071804. URL <http://link.aps.org/doi/10.1103/PhysRevLett.109.071804>
- [18] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, M. Spira, Effective Lagrangian for a light Higgs-like scalar, JHEP 1307 (2013) 035. arXiv:1303.3876, doi:10.1007/JHEP07(2013)035.
- [19] A. Falkowski, F. Riva, A. Urbano, Higgs at last, JHEP 1311 (2013) 111. arXiv:1303.1812, doi:10.1007/JHEP11(2013)111.
- [20] J. R. Ellis, Supersymmetry for Alp hikers (2002) 157–203 arXiv:hep-ph/0203114.
- [21] J. M. Maldacena, The Large N limit of superconformal field theories and supergravity, Adv.Theor.Math.Phys. 2 (1998) 231–252. arXiv:hep-th/9711200.
- [22] R. Rattazzi, A. Zaffaroni, Comments on the holographic picture of the Randall-Sundrum model, JHEP 0104 (2001) 021. arXiv:hep-th/0012248, doi:10.1088/1126-6708/2001/04/021.
- [23] R. Contino, L. Da Rold, A. Pomarol, Light custodians in natural composite Higgs models, Phys.Rev. D75 (2007) 055014. arXiv:hep-ph/0612048, doi:10.1103/PhysRevD.75.055014.
- [24] M. S. Carena, E. Ponton, J. Santiago, C. E. Wagner, Light Kaluza Klein States in Randall-Sundrum Models with Custodial SU(2), Nucl.Phys. B759 (2006) 202–227. arXiv:hep-ph/0607106, doi:10.1016/j.nuclphysb.2006.10.012.
- [25] N. Arkani-Hamed, A. G. Cohen, H. Georgi, Electroweak symmetry breaking from dimensional deconstruction, Phys.Lett. B513 (2001) 232–240. arXiv:hep-ph/0105239, doi:10.1016/S0370-2693(01)00741-9.
- [26] N. Arkani-Hamed, A. Cohen, E. Katz, A. Nelson, The Littlest Higgs, JHEP 0207 (2002) 034. arXiv:hep-ph/0206021, doi:10.1088/1126-6708/2002/07/034.



- [27] M. Perelstein, Little Higgs models and their phenomenology, *Prog.Part.Nucl.Phys.* 58 (2007) 247–291. [arXiv:hep-ph/0512128](#), [doi:10.1016/j.pnpnp.2006.04.001](#).
- [28] R. Contino, Y. Nomura, A. Pomarol, Higgs as a holographic pseudoGoldstone boson, *Nucl.Phys. B* 671 (2003) 148–174. [arXiv:hep-ph/0306259](#), [doi:10.1016/j.nuclphysb.2003.08.027](#).
- [29] A. David, et al., LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle [arXiv:1209.0040](#).
- [30] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables, CERN-2011-002 [arXiv:1101.0593](#).
- [31] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), Handbook of LHC Higgs Cross Sections: 2. Differential Distributions, CERN-2012-002 [arXiv:1201.3084](#).
- [32] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, R. Tanaka (Eds.), Handbook of LHC Higgs Cross Sections: 3. Higgs Properties, CERN-2013-004 [arXiv:1307.1347](#).
- [33] *The ALICE experiment at the CERN LHC*, 2008. [doi:10.1088/1748-0221/3/08/S08002](#).
- [34] G. Aad, et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* 3 (2008) S08003. [doi:10.1088/1748-0221/3/08/S08003](#).  
URL <http://www-library.desy.de/cgi-bin/spiface/find/hep/www?irn=7955545>
- [35] R. Adolphi, et al., The CMS experiment at the CERN LHC, *JINST* 3 (2008) S08004. [doi:10.1088/1748-0221/3/08/S08004](#).
- [36] T. L. Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* 3 (08) (2008) S08005.  
URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08005>
- [37] G. Bayatian, et al., CMS technical design report, volume II: Physics performance, *J.Phys. G* 34 (2007) 995–1579. [doi:10.1088/0954-3899/34/6/S01](#).
- [38] The CMS magnet project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, 1997.
- [39] The CMS tracker project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, 1997.
- [40] The CMS electromagnetic calorimeter project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, 1997.
- [41] P. Adzic, Energy resolution of the barrel of the cms electromagnetic calorimeter, *Journal of Instrumentation* 2 (04) (2007) P04004.  
URL <http://stacks.iop.org/1748-0221/2/i=04/a=P04004>

- [42] S. Chatrchyan, et al., Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at  $\sqrt{s} = 7$  TeV, JINST 8 (2013) P09009. [arXiv:1306.2016](#), [doi:10.1088/1748-0221/8/09/P09009](#).
- [43] The CMS hadron calorimeter project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, 1997.
- [44] The CMS muon project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, 1997.
- [45] *CMS TriDAS project: The trigger systems*, Technical Design Report CMS, CERN, 1997.
- [46] ECAL response variations due to LHC irradiation.
- [47] S. Chatrchyan, et al., Measurement of the inclusive W and Z production cross sections in pp collisions at  $\sqrt{s} = 7$  TeV with the CMS experiment, JHEP 10 (2011) 132. [arXiv:1107.4789](#), [doi:10.1007/JHEP10\(2011\)132](#).
- [48] S. Chatrchyan, et al., Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV, JHEP 06 (2013) 081. [arXiv:1303.4571](#), [doi:10.1007/JHEP06\(2013\)081](#).
- [49] CMS Collaboration, Particle-flow event reconstruction in CMS and performance for jets, taus, and  $E_T^{\text{miss}}$ , CMS Physics Analysis Summary CMS-PAS-PFT-09-001 (2009).  
URL <http://cdsweb.cern.ch/record/1194487>
- [50] CMS Collaboration, Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector, CMS Physics Analysis Summary CMS-PAS-PFT-10-001 (2010).  
URL <http://cdsweb.cern.ch/record/1247373>
- [51] H. Voss, A. Höcker, J. Stelzer, F. Tegenfeldt, TMVA: Toolkit for Multivariate Data Analysis with ROOT, in: XIth International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT), 2007, p. 40. [arXiv:physics/0703039](#).  
URL [http://pos.sissa.it/archive/conferences/050/040/ACAT\\_040.pdf](http://pos.sissa.it/archive/conferences/050/040/ACAT_040.pdf)
- [52] M. Cacciari, G. P. Salam, G. Soyez, The Anti-k(t) jet clustering algorithm, JHEP 0804 (2008) 063. [arXiv:0802.1189](#), [doi:10.1088/1126-6708/2008/04/063](#).
- [53] S. Chatrchyan, et al., Determination of jet energy calibration and transverse momentum resolution in CMS, JINST 6 (2011) P11002. [doi:10.1088/1748-0221/6/11/P11002](#).
- [54] CMS Collaboration, Pileup Jet Identification, CMS Physics Analysis Summary CMS-PAS-JME-13-005 (2013).  
URL <http://cdsweb.cern.ch/record/1581583>

- [55] S. Chatrchyan, et al., Identification of b-quark jets with the CMS experiment, JINST 8 (2012) P04013. [arXiv:1211.4462](#), [doi:10.1088/1748-0221/8/04/P04013](#).
- [56] S. Chatrchyan, et al., Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV, JINST 7 (2012) P10002. [arXiv:1206.4071](#), [doi:10.1088/1748-0221/7/10/P10002](#).
- [57] S. Agostinelli, et al., GEANT4 - a simulation toolkit, Nucl. Instrum. Meth. A 506 (2003) 250. [doi:10.1016/S0168-9002\(03\)01368-8](#).
- [58] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, JHEP 11 (2004) 040. [arXiv:hep-ph/0409146](#), [doi:10.1088/1126-6708/2004/11/040](#).
- [59] S. Frixione, P. Nason, C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, JHEP 11 (2007) 070. [arXiv:0709.2092](#), [doi:10.1088/1126-6708/2007/11/070](#).
- [60] S. Alioli, P. Nason, C. Oleari, E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP 06 (2010) 043. [arXiv:1002.2581](#), [doi:10.1007/JHEP06\(2010\)043](#).
- [61] S. Alioli, P. Nason, C. Oleari, E. Re, NLO Higgs boson production via gluon fusion matched with shower in POWHEG, JHEP 04 (2009) 002. [arXiv:0812.0578](#), [doi:10.1088/1126-6708/2009/04/002](#).
- [62] P. Nason, C. Oleari, NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG, JHEP 02 (2010) 037. [arXiv:0911.5299](#), [doi:10.1007/JHEP02\(2010\)037](#).
- [63] T. Sjöstrand, S. Mrenna, P. Z. Skands, PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026. [arXiv:hep-ph/0603175](#), [doi:10.1088/1126-6708/2006/05/026](#).
- [64] G. Bozzi, S. Catani, D. de Florian, M. Grazzini, The  $q_T$  spectrum of the Higgs boson at the LHC in QCD perturbation theory, Phys. Lett. B 564 (2003) 65. [arXiv:hep-ph/0302104](#), [doi:10.1016/S0370-2693\(03\)00656-7](#).
- [65] G. Bozzi, S. Catani, D. de Florian, M. Grazzini, Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC, Nucl. Phys. B 737 (2006) 73. [arXiv:hep-ph/0508068](#), [doi:10.1016/j.nuclphysb.2005.12.022](#).
- [66] D. de Florian, G. Ferrera, M. Grazzini, D. Tommasini, Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC, JHEP 11 (2011) 064. [arXiv:1109.2109](#), [doi:10.1007/JHEP11\(2011\)064](#).
- [67] LHC Higgs Cross Section Working Group, Handbook of LHC Higgs cross sections: 3. Higgs Properties, CERN Report CERN-2013-004 (2013). [arXiv:1307.1347](#), [doi:10.5170/CERN-2013-004](#).

- [68] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, MadGraph 5 : going beyond, JHEP 06 (2011) 128. [arXiv:1106.0522](#), [doi:10.1007/JHEP06\(2011\)128](#).
- [69] T. Gleisberg, S. Hoche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, J. Winter, Event generation with SHERPA 1.1, JHEP 02 (2009) 007. [arXiv:0811.4622](#), [doi:10.1088/1126-6708/2009/02/007](#).
- [70] S. Chatrchyan, et al., Measurement of the Production Cross Section for Pairs of Isolated Photons in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, JHEP 01 (2012) 133. [arXiv:1110.6461](#), [doi:10.1007/JHEP01\(2012\)133](#).
- [71] S. Chatrchyan, et al., Measurement of the differential dijet production cross section in proton-proton collisions at  $\sqrt{s} = 7$  TeV, Phys. Lett. B 700 (2011) 187. [arXiv:1104.1693](#), [doi:10.1016/j.physletb.2011.05.027](#).
- [72] D. L. Rainwater, R. Szalapski, D. Zeppenfeld, Probing color singlet exchange in  $Z + \text{two jet}$  events at the CERN LHC, Phys. Rev. D 54 (1996) 6680. [arXiv:hep-ph/9605444](#), [doi:10.1103/PhysRevD.54.6680](#).
- [73] ATLAS and CMS Collaborations, LHC Higgs Combination Group, Procedure for the LHC Higgs boson search combination in Summer 2011, Tech. Rep. ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005, CERN (2011).  
URL <http://cdsweb.cern.ch/record/1379837>
- [74] S. Chatrchyan, et al., Combined results of searches for the standard model Higgs boson in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, Phys. Lett. B 710 (2012) 26. [arXiv:1202.1488](#), [doi:10.1016/j.physletb.2012.02.064](#).
- [75] L. Moneta, K. Belasco, K. S. Cranmer, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, M. Wolf, The RooStats project, in: 13<sup>th</sup> International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010), SISSA, 2010. [arXiv:1009.1003](#).  
URL [http://pos.sissa.it/archive/conferences/093/057/ACAT2010\\_057.pdf](http://pos.sissa.it/archive/conferences/093/057/ACAT2010_057.pdf)
- [76] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (1974) 716. [doi:10.1109/TAC.1974.1100705](#).
- [77] CMS Collaboration, Absolute calibration of the luminosity measurement at CMS: Winter 2012 update, CMS Physics Analysis Summary CMS-PAS-SMP-12-008 (2012).  
URL <http://cdsweb.cern.ch/record/1434360>
- [78] CMS Collaboration, Cms luminosity based on pixel cluster counting - summer 2013 update, CMS Physics Analysis Summary CMS-PAS-LUM-13-001 (2013).  
URL <http://cdsweb.cern.ch/record/1598864>
- [79] R. Paramatti, CMS ECAL group, Crystal properties in the electromagnetic calorimeter of CMS, AIP Conf. Proc. 867 (2006) 245. [doi:10.1063/1.2396960](#).

- [80] E. Auffray, Overview of the 63000 PWO barrel crystals for CMS ECAL production, *IEEE Trans. Nucl. Sci.* 55 (2008) 1314. doi:10.1109/TNS.2007.913935.
- [81] Measurement of Jet Multiplicity Distributions in Top Quark Events With Two Leptons in the Final State at a centre-of-mass energy of 7 TeV, Tech. Rep. CMS-PAS-TOP-12-023, CERN, Geneva (2012).
- [82] R. Barlow, Event classification using weighting methods, *J. Comp. Phys.* 72 (1987) 202. doi:10.1016/0021-9991(87)90078-7.
- [83] E. Gross, O. Vitells, Trial factors for the look elsewhere effect in high energy physics, *Eur. Phys. J. C* 70 (2010) 525. arXiv:1005.1891, doi:10.1140/epjc/s10052-010-1470-8.
- [84] Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model, Tech. Rep. CMS-PAS-HIG-14-009, CERN, Geneva (2014).
- [85] S. Chatrchyan, et al., Measurement of the properties of a Higgs boson in the four-lepton final state, *Phys. Rev. D* 89 (2014) 092007. arXiv:1312.5353, doi:10.1103/PhysRevD.89.092007.
- [86] S. Chatrchyan, et al., Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states, *JHEP* 01 (2014) 096. arXiv:1312.1129, doi:10.1007/JHEP01(2014)096.
- [87] S. Chatrchyan, et al., Evidence for the 125 GeV Higgs boson decaying to a pair of  $\tau$  leptons, *JHEP* 05 (2014) 104. arXiv:1401.5041, doi:10.1007/JHEP05(2014)104.  
URL <http://dx.doi.org/10.1007/JHEP05%282014%29104>
- [88] S. Chatrchyan, et al., Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks, *Phys. Rev. D* 89 (2014) 012003. arXiv:1310.3687, doi:10.1103/PhysRevD.89.012003.
- [89] Search for Higgs Boson Production in Association with a Top-Quark Pair and Decaying to Bottom Quarks or Tau Leptons, Tech. Rep. CMS-PAS-HIG-13-019, CERN, Geneva (2013).
- [90] Search for the standard model Higgs boson produced in association with top quarks in multilepton final states, Tech. Rep. CMS-PAS-HIG-13-020, CERN, Geneva (2013).
- [91] V. Khachatryan, et al., Search for the associated production of the Higgs boson with a top-quark pair arXiv:1408.1682.
- [92] J. Aguilar-Saavedra, Top flavor-changing neutral interactions: Theoretical expectations and experimental detection, *Acta Phys. Polon. B35* (2004) 2695–2710. arXiv:hep-ph/0409342.

- [93] A. Collaboration, Search for flavour changing neutral currents in top quark decays top to charm and Higgs, with Higgs to photons, and limit on the  $t\bar{c}H$  coupling with the ATLAS detector at the LHC, ATLAS Conference Note ATLAS-CONF-2013-081.
- [94] C. Collaboration, A search for anomalous production of events with three or more leptons using 8 TeV data, CMS Conference Note CMS-SUS-13-002.
- [95] J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, M. Perez-Victoria, A handbook of vector-like quarks: mixing and single production, Phys.Rev. D88 (2013) 094010. [arXiv:1306.0572](https://arxiv.org/abs/1306.0572), [doi:10.1103/PhysRevD.88.094010](https://doi.org/10.1103/PhysRevD.88.094010).
- [96] W. Kilian, T. Ohl, J. Reuter, Whizard - simulating multi-particle processes at lhc and ilc, The European Physical Journal C 71 (9) (2011) 1–29. [doi:10.1140/epjc/s10052-011-1742-y](https://doi.org/10.1140/epjc/s10052-011-1742-y).  
URL <http://dx.doi.org/10.1140/epjc/s10052-011-1742-y>
- [97] M. Moretti, T. Ohl, J. Reuter, O'Mega: An Optimizing matrix element generator [arXiv:hep-ph/0102195](https://arxiv.org/abs/hep-ph/0102195).
- [98] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, A. Thamm, Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson, JHEP 1305 (2013) 022. [arXiv:1211.3736](https://arxiv.org/abs/1211.3736), [doi:10.1007/JHEP05\(2013\)022](https://doi.org/10.1007/JHEP05(2013)022).
- [99] S. Chatrchyan, et al., Search for pair produced fourth-generation up-type quarks in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with a lepton in the final state, Phys.Lett. B718 (2012) 307–328. [arXiv:1209.0471](https://arxiv.org/abs/1209.0471), [doi:10.1016/j.physletb.2012.10.038](https://doi.org/10.1016/j.physletb.2012.10.038).
- [100] G. Aad, et al., Search for pair-produced heavy quarks decaying to  $Wq$  in the two-lepton channel at  $\sqrt{s} = 7$  TeV with the ATLAS detector, Phys.Rev. D86 (2012) 012007. [arXiv:1202.3389](https://arxiv.org/abs/1202.3389), [doi:10.1103/PhysRevD.86.012007](https://doi.org/10.1103/PhysRevD.86.012007).
- [101] S. Chatrchyan, et al., Search for heavy, top-like quark pair production in the dilepton final state in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, Phys.Lett. B716 (2012) 103–121. [arXiv:1203.5410](https://arxiv.org/abs/1203.5410), [doi:10.1016/j.physletb.2012.07.059](https://doi.org/10.1016/j.physletb.2012.07.059).
- [102] G. Aad, et al., Search for pair production of heavy top-like quarks decaying to a high-pT  $W$  boson and a  $b$  quark in the lepton plus jets final state at  $\sqrt{s}=7$  TeV with the ATLAS detector, Phys.Lett. B718 (2013) 1284–1302. [arXiv:1210.5468](https://arxiv.org/abs/1210.5468), [doi:10.1016/j.physletb.2012.11.071](https://doi.org/10.1016/j.physletb.2012.11.071).
- [103] S. Chatrchyan, et al., Search for heavy quarks decaying into a top quark and a  $W$  or  $Z$  boson using lepton + jets events in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, JHEP 1301 (2013) 154. [arXiv:1210.7471](https://arxiv.org/abs/1210.7471), [doi:10.1007/JHEP01\(2013\)154](https://doi.org/10.1007/JHEP01(2013)154).
- [104] M. Czakon, A. Mitov, Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders [arXiv:1112.5675](https://arxiv.org/abs/1112.5675).
- [105] A. Martin, W. Stirling, R. Thorne, G. Watt, Parton distributions for the lhc, Eur. Phys. J. C63 (2009) 189–285. [arXiv:0901.0002](https://arxiv.org/abs/0901.0002).

- [106] M. Czakon, P. Fiedler, A. Mitov, The total top quark pair production cross-section at hadron colliders through  $\mathcal{O}(\alpha_s^4)$ , *Phys.Rev.Lett.* 110 (2013) 252004. [arXiv:1303.6254](https://arxiv.org/abs/1303.6254), [doi:10.1103/PhysRevLett.110.252004](https://doi.org/10.1103/PhysRevLett.110.252004).
- [107] A. Read, Modified frequentist analysis of search results (the  $CL_s$  method), Tech. Rep. CERN-OPEN-2000-005, CERN (2000).  
URL <http://cdsweb.cern.ch/record/451614>
- [108] T. Junk, Confidence level computation for combining searches with small statistics, *Nucl. Instrum. Meth. A* 434 (1999) 435. [arXiv:hep-ex/9902006](https://arxiv.org/abs/hep-ex/9902006), [doi:10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2).
- [109] S. Chatrchyan, et al., Inclusive search for a vector-like t quark with charge  $2/3$  in pp collisions at  $\sqrt{s} = 8$  tev, *Phys. Lett. B* 729 (0) (2014) 149 – 171. [doi:10.1016/j.physletb.2014.01.006](https://doi.org/10.1016/j.physletb.2014.01.006).  
URL <http://www.sciencedirect.com/science/article/pii/S037026931400015X>
- [110] Search for top quark partners with kinematic reconstruction in single-lepton final states, Tech. Rep. CMS-PAS-B2G-12-017, CERN, Geneva (2014).
- [111] Search for top-quark partners decaying to bW in the all hadronic final state, Tech. Rep. CMS-PAS-B2G-12-013, CERN, Geneva (2014).
- [112] Y. L. Dokshitzer, G. D. Leder, S. Moretti, B. R. Webber, Better Jet Clustering Algorithms, *JHEP* 08 (1997) 001. [arXiv:hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323), [doi:10.1088/1126-6708/1997/08/001](https://doi.org/10.1088/1126-6708/1997/08/001).
- [113] D. E. Kaplan, K. Rehermann, M. D. Schwartz, B. Tweedie, Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks, *Phys.Rev.Lett.* 101 (2008) 142001. [arXiv:0806.0848](https://arxiv.org/abs/0806.0848), [doi:10.1103/PhysRevLett.101.142001](https://doi.org/10.1103/PhysRevLett.101.142001).
- [114] T. Plehn, M. Spannowsky, Top Tagging, *J. Phys. G* 39 (2012) 083001. [arXiv:1112.4441](https://arxiv.org/abs/1112.4441), [doi:10.1088/0954-3899/39/8/083001](https://doi.org/10.1088/0954-3899/39/8/083001).
- [115] CMS Collaboration, Boosted top jet tagging at cms, CMS Physics Analysis Summary CMS-PAS-JME-13-007, <http://cds.cern.ch/record/1577417> (2013).
- [116] CMS Collaboration, Identifying Hadronically Decaying W Bosons Merged into a Single Jet, CMS Physics Analysis Summary CMS-PAS-JME-13-006.
- [117] Search for top-Higgs resonances in all-hadronic final states using jet substructure methods, Tech. Rep. CMS-PAS-B2G-14-002, CERN, Geneva (2014).
- [118] V. Khachatryan, et al., Observation of the diphoton decay of the Higgs boson and measurement of its properties, *The European Physical Journal C* 74 (10). [doi:10.1140/epjc/s10052-014-3076-z](https://doi.org/10.1140/epjc/s10052-014-3076-z).  
URL <http://dx.doi.org/10.1140/epjc/s10052-014-3076-z>

- [119] Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at  $\sqrt{s} = 8$  TeV, Tech. Rep. CMS-PAS-HIG-14-001, CERN, Geneva (2014).
- [120] Search for vector-like top quark partners produced in association with Higgs bosons in the diphoton final state, Tech. Rep. CMS-PAS-B2G-14-003, CERN, Geneva (2014).



## Acknowledgments

*A Peppe, Lolli, Maria e Cesare,  
per i loro occhi troppo belli e per la loro grafia da bambino di quinta elementare.*

*A Nando e Rosanna,  
per tutto, perché non giudicano ma provano a capire e perché dicono "copriti" quando esco di casa.*

*A Maria Rita,  
perché dovunque tu sia, dovunque io vada, saremo sempre unici. Sei per me quel cucchiaino di Totti contro l'Olanda nel 2000.*

*A Nicola, Marina e Giulia,  
sangue del mio sangue, nervi dei miei nervi. Per aver dormito nel mio stesso letto almeno due volte l'anno da quasi vent'anni.*

*A Bozzo, Glenda (e Castalia),  
perché ci sono sempre e perché una volta, da sbronzo, in un vicololetto di San Lorenzo ho detto loro di fare un figlio con un nome strano. E l'hanno fatto davvero.*

*A Ivanotalamo e Pandolf,  
perché non vedo l'ora di essere in punto di morte per fare la scena della morte del Perozzi di Amici Miei.*

*A chi era parte della mia vita e ne è uscito per colpa mia ("Ho mentito ogni volta che ho detto 'sto bene'. Tutte le volte. E questo lo sapevano tutti."), per colpa sua ("ancora tra cent'anni ci sarà chi curerà questa ferita che non guarisce mai") o per colpa di nessuno. Il mio numero è sempre quello e soffro ancora di insonnia se ve lo state chiedendo.*

*A chi nella mia vita c'è sempre stato e sempre ci sarà. E anche a chi entra nella mia vita all'improvviso, cambiando le carte in tavola. I vostri nomi qui non ci sono ma voi vi riconoscerete sicuro.*

*A Beppe Fenoglio, scrittore e partigiano.*

*Ai Cccp-Csi-Pgr e a Gipi, per l'ansia.*

*Alla Toscana, per essere il posto più bello del mondo. Alla Roma fuori dalle mura aureliane (esclusa Roma Nord).*

*Alla vita, che, quando non è schifosa, è splendida.*